The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods

Article *in* Geocarto International · May 2019

DOI: 10.1080/10106049.2019.1622597

CITATIONS
3

READS
88

2 authors:

Akhtar Jamil
Istanbul Sabahattin Zaim University
23 PUBLICATIONS **52** CITATIONS

SEE PROFILE

Bulent Bayram
Yildiz Technical University
87 PUBLICATIONS **264** CITATIONS

SEE PROFILE

Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: https://www.tandfonline.com/loi/tgei20

The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods

Akhtar Jamil & Bulent Bayram

To cite this article: Akhtar Jamil & Bulent Bayram (2019): The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods, Geocarto International, DOI: 10.1080/10106049.2019.1622597

To link to this article: https://doi.org/10.1080/10106049.2019.1622597

Accepted author version posted online: 21 May 2019.

The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods

Akhtar Jamil*[a] and Bulent Bayram[b]

[a] *Department of Computer Engineering, Istanbul Sabahattin Zaim University, Istanbul Turkey.*

[b] *Department of Geomatics Engineering, Yildiz Technical University, Istanbul, Turkey*

Email: *akhtar.jamil@izu.edu.tr (0000-0002-2592-1039), bayram@yildiz.edu.tr

The delineation of tea gardens from high resolution digital orthoimages using mean-shift and supervised machine learning methods

Rize district is an important tea production site in Turkey, which is known for high quality tea. Determining the temporal changes is very crucial from the viewpoint of agricultural management and protection of tea areas. In addition, delineation of tea gardens using photogrammetric evaluation techniques for a single orthoimage takes approximately 8 hours of labour work, which is both costly and time-consuming process. To overcome these issues, a method is proposed for demarcation of tea gardens from high-resolution orthoimages. In this paper, a hierarchical object-based segmentation using mean-shift (MS) and supervised machine learning methods are investigated for delineation of tea gardens. First, the mean-shift algorithm was applied to partition the images into homogeneous segments (objects) and then from each segment, various spectral, spatial and textural features were extracted. Finally, four most widely used supervised machine learning classifiers, support vector machine (SVM), artificial neural network (ANN), Random Forest (RF), and Decision Trees (DTs), were selected for classification of objects into tea gardens and other types of trees. Photogrammetrically evaluated tea garden borders were taken as reference data to evaluate the performance of the proposed methods. The experiments showed that all selected supervised classifiers were effective for delineation of the tea gardens from high-resolution images.

Keywords: tea garden extraction; mean-shift segmentation; support vector machine; artificial neural network; decision trees; random forest;

1. Introduction

The advent of high spatial resolution imaging technology has made it possible to perform detailed image analysis even at fine level with good accuracy. Although, high resolution multispectral images encompass an in-depth land surface reflectance information, yet, they also induce several problems such as, inter-class spectral

similarity and intra-class spatial variability etc. (Camps-Valls and Bruzzone 2005). Generally, object-based feature descriptors are more effective for remote sensing data analysis compared to the individual pixel-based counterpart (Costa, Foody, and Boyd 2018). Therefore, object-based approaches have widely been used in last few decades for analysis of high resolution images (Thomas Blaschke et al. 2014). The rationale of this approach is to group individual pixels into meaningful objects based on some similarity metric (Pham, Brabyn, and Ashraf 2016). This representation allows extraction of additional information which is not possible with pixel-based approaches, such as contextual, structural and textural features (T. Blaschke 2010). The raw pixels are transformed into an object-based representation by using segmentation techniques.

Image segmentation is the basis of object-based image analysis (OBIA). Various methods have been proposed in literature for segmentation of raw pixels into meaningful objects. For instance, (Pipaud and Lehmkuhl 2017) proposed an object-based method by using mean-shift and SVM for classification of alluvial fans. The final output was obtained by applying a fuzzy membership on SVM classification result. Similarly, (Su et al. 2015) employed mean-shift for extraction of crop lands from high resolution images. In (Qin 2014) mean-shift vector-based shape features (MSVSF) were derived for classification of high resolution images. MSVSF is considered as a shape descriptor for spectrally homogeneous local regions. Adaptive mean-shift algorithm and SVM classifier are proposed for delineation of objects from hyperspectral images in (Huang and Zhang 2008). Mean-shift segmentation and maximum likelihood classifier are used for classification of six different crops in (Ozdarici Ok and Akyurek 2012).

Combining OBIA with supervised machine learning (ML) approaches offer enormous potential for remote sensing applications, such as tea / hazelnut tree delineation (Dihkan et al. 2013; Akar and Güngör 2015; Jamil, Bayram, and Seker

2019), extraction of green space areas from Pleiades-1A images (Zylshal et al. 2016), vegetation classification (Street 2010), land cover classification (Qian et al. 2015), buildings and tree species extraction (Zarea and Mohammadzadeh 2015), etc. Similarly, in (Pal and Mather 2005) the performance of multi-class SVMs is compared with maximum likelihood and ANN classifiers for remote sensing data classification. (Sasaki et al. 2012) employed DTs for classification of land cover classes and dominant trees from LiDAR data. In the context of high resolution image classification, ML methods have produced significantly optimal results (Jamil and Bayram 2018). Further details about widely used supervised machine learning classifiers for classification of remote sensing data can be found in (Maxwell, Warner, and Fang 2018).

In this study, an object-based method for segmentation using mean-shift followed by supervised classifiers is investigated for delineation of tea-gardens from high resolution digital orthoimages. Mean-shift algorithm was used as basis for object-based segmentation while the performance of fours relatively mature supervised classifiers namely, SVM, ANN, RF and DTs, were investigated to classify the objects into tea gardens and other types of trees. As a secondary objective, further experiments were performed to find the minimum amount of training data size required to train the classifiers (SVM, ANN, RF and DTs). It is also shown that for SVM, RBF kernel performed better than polynomial kernel on all images. Similarly, for ANN, sigmoid activation function produced slightly more accurate results in terms of overall accuracy than tanh activation function.

2. Methodology

The schematic diagram of the proposed method is shown in Figure 1. The main steps are: 1) NDVI extraction 2) mean-shift segmentation 3) morphological processing 4)

feature extraction 5) classification. In the following sections, each step is described in details.

Figure 1

2.1 Study Area

Turkey is one of the major consumer and producer of tea, therefore, tea garden extraction is useful to estimate the extent and distribution of tea gardens. Rize district was selected as the study area which is located in northeast of Turkey near the eastern Black Sea coast. Most of the area is covered with forest and agricultural land. Figure 2 a) shows map of Turkey and location of Rize district on the map while b) shows a sample orthoimage of the area in RGB colour space and c) shows its reference data where white pixels indicate delineated tea gardens.

2.2 Dataset

Flights were made on March 27-28, 2013 in the Rize district to obtain images using an UltraCamX digital aerial camera. The multispectral images obtained had four spectral bands namely: red, green, blue and near infrared. The acquisition of images and creation of reference orthoimages was done by EMI Group Inc. Turkey. For this study, ten digital orthoimages were used, which were obtained from different geographic locations from Rize district. Table 1 summarizes the technical specification of the aerial camera system used to capture the images.

The photogrammetrically evaluated tea garden border were transformed to binary images, where white pixels represented the delineated tea gardens and black pixels represent the background (Figure 1). These images were then used as reference data for accuracy assessment of the proposed method. Since supervised machine learning

approaches require data from all potential classes, therefore, to train the classifiers, two classes were defined: tea garden and other trees. Training objects for each class were selected from randomly selected five images while for testing rest of the objects from all ten images were considered.

Figure 2

2.3 Normalize Difference Vegetation Index (NDVI) Extraction

The tea garden extraction step was carried in two stages: first the vegetation areas were obtained by eliminating the non-vegetation areas. In the second step, the vegetation areas were further classified into tea gardens and other types of trees by applying ANN and SVM classifiers.

For the first step, NDVI index (Tucker 1979) was calculated from the original digital orthoimages. It is obtained by taking ratio of difference and sum of red (R) and near infrared (NIR) band of the images. Mathematically,

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad (1)$$

The output values of NDVI are in the range [-1, +1]. Higher values indicate stronger vegetation while lower values indicate lack or absence of vegetation. Since, the tea gardens and other types of vegetation assumed to have NDVI values greater than zero, so NDVI values were filtered and only values greater than 0.2 were retained while other values were discarded. As it can be seen in Figure 3 that there is overlap in NDVI profile between tea gardens and other trees, however, the profiles for tea garden have low variations compared to other trees. In other words, tea gardens provided consistent spectral information which produced more consistent homogeneous segments when mean-shift algorithm was applied on it.

Figure 3

2.4 Mean Shift Segmentation

Image segmentation is the prerequisite of object-based image analysis (OBIA). The idea is to accumulate individual pixels into groups of homogeneous regions called objects (segments) based on their spectral similarity with their neighbouring pixels. This provides more meaningful representation of the image objects and analysis can be relatively easy to perform (Pedram Ghamisi et al. 2017). In this paper, mean shift algorithm was employed as basis for object-based image segmentation.

Mean shift is a nonparametric gradient estimation algorithm that is extremely versatile and can be used for applications like mode seeking, clustering, image segmentation, visual object tracking, and space analysis (Cheng 1995). Although it was first introduced by (Fukunaga and Hostetler 1975), however, it received more attention in the computer vision community when (Comaniciu and Meer 2002) proved the convergence of the mean shift procedure towards a stationary point, thus enabled to find the mode of the density function. The idea is to consider the feature space as a probability density function (PDF), where the dense regions corresponds to the local maxima (mode) of the underlying PDF (Comaniciu and Meer 2002). According to(Comaniciu and Meer 2002) fundamental mathematical formulation of mean shift method are given below:

Given $n \in R^d$ data points represented by $x_i \;\; , where\ i = 1,2, \dots n,$ then the kernel density estimation function can be written as

$$f(x) = \frac{c}{nd^d} \sum_{i=1}^{n} K\left(\frac{x - \overline{x_i}}{h}\right) \tag{2}$$

Where $K(x)$ is the kernel function, which provides a measure of correlation between the data point, $\overline{x_i}$ is the data centre, h is the bandwidth parameter, and c is the

normalization constant. Suppose, x_s be spatial domain and x_r be the range domain of the image, then the multivariate kernel can be modelled as:

$$f(x) = \frac{c}{n h_s^{2h_r^p} d^d} \sum_{i=1}^{n} K\left(\left\| \frac{x_s - \overline{x_{si}}}{h_s} \right\|^2 + \left\| \frac{x_r - \overline{x_{ri}}}{h_r} \right\|^2 \right) \quad (3)$$

Where h_s and h_r represent bandwidths of spectral and spatial domains respectively while p is the number of bands in the multispectral image. Further details about mathematical formulation of mean shift can be found in (Comaniciu and Meer 2002)

Mean shift algorithm was applied on the NDVI image as the objective was to identify one of the vegetation types, i.e. tea gardens. The bandwidth parameters of the algorithm were empirically calculated by running it on the input images and visualizing the segmentation results. It was observed that smaller values for h_r were sensitive to noise and resulted in substantial number of small segments. On the other side, larger values tend to ignore small but important gradient information which resulted in merging small segments (Qin 2014). Similarly, large values for h_s mere increased the computational complexity as it covered larger area, and vice versa. It is interesting to note that h_s was less sensitive than h_r for segmentation of images.

From implementation point of view, apart from bandwidth, two more parameters were required: number of iterations and minimum segment size. The former indicates the number of times mean shift algorithm is repeated on the input image, while the later parameter indicates a threshold on the size of the segments. Segments smaller than this threshold were merged with neighbouring segments. For this study, the number of iterations were set to one while the minimum segment size was set to fifty for all images. For implementation, Emgu CV library and C# were used. Figure 4 shows the impact of the bandwidth selection on the mean shift segmentation for varying values of

(h_s, h_r). For current study, the bandwidth parameter with values (5,5) produced optimal results on the used dataset.

2.5 Morphological Processing

The mean-shift resulted in an object-based representation of image. Since, the candidate tea garden objects were relatively larger compared to other vegetation object such as individual trees, small grassy areas, etc., therefore, morphological processing and geometrical constraints were applied to remove such objects from the processed images.

Figure 4

Morphological processing is a nonlinear operation that can be used to analyse the spatial relationship between neighbouring pixels (Gu et al. 2016). Structural information can be extracted by selecting a special element called structuring element (SE) with known shape and size. The size of SE was empirically calculated (5x5). The two fundaments morphological operations (dilation and erosion) were used in this study to remove small objects smaller than the selected SE and fill holes within objects.

Geometrical constraints were applied to eliminate small objects which were not removed by morphological operation such as small trees, and isolated green areas. To achieve this, threshold values were empirically obtained for width, height, and area of the bounding box for each candidate object by analysing the size of individual target objects in the segmented image.

2.6 Feature Extraction and Selection

Spectral, spatial and textural features were extracted from each object for classification. The spatial/spectral features were directly obtained from the spectral reflectance information while texture features were derived from grey-level co-occurrence matrix (GLCM). The GLCM was constructed for each object in the segmented image using

object's bounding rectangle and then features were calculated.

Initially, 25 features were considered for each object. 16 were spectral/spatial features, which include mean, standard deviation, minimum and maximum value for each object in all four bands. Similarly, from GLCM, 8 features were extracted, which include homogeneity, contrast, energy, correlation, local homogeneity, sum of entropy, cluster prominence, and cluster shade. Additionally, the NDVI value was also included in the feature vector.

To obtain high classification accuracy, however, an optimal set of features needed to be selected by removing redundant or less informative ones (Chehata et al. 2014). For this purpose, a search strategy and seperability criteria among features is required (Melgani and Bruzzone 2004). The former aims to reducing the original feature space to a lower dimensional space while retaining as much original information as possible. The later focuses on measuring the seperability among features using a statistical distance measure such as Bhattacharyya, Jeffries-Matusita (JM) distance etc. JM distance gradually saturates as it reaches optimal number of features (Dalponte et al. 2013).

This study employed a combination of sequential forward section algorithm and JM for feature selection which have been widely used for remote sensing applications. The feature selection approach resulted in selection of 13 optimal features from 25 original features set. Table 2 summarizes the optimal features used in this study. These features are not highly correlated as the linear correlations are less than $|r| \leq 0.68$ (see table 3 for correlation matrix).

2.5 Image Classification

ML methods have widely been used for remote sensing data classification (Jamil and Bayram 2018). In this study, four mostly widely used supervised classifiers were

selected: SVM,ANN, RF and DT. SVM is a nonparametric supervised classifier which was primarily designed to classify linearly separable data. However, to cope with non-linearity of data, kernel methods can be applied. The objective is to transforms the data into a higher dimensional space where it is assumed to be linearly separable, then it constructs a hyperplane to classify data into respective classes. SVMs have shown their effectiveness and produced promising results even in presence of small training data for large scale data classification (Camps-Valls and Bruzzone 2005). For a detailed description about mathematical background of SVM, refer to (Melgani and Bruzzone 2004).

Similarly, ANN is another nonparametric supervised classifier. It can learn to classify complex nonlinearly separable data but, the computational complexity of training is relatively higher (Pedram Ghamisi et al. 2017). Several approaches based on ANNs have been proposed in literature, but the most commonly used approach is multilayer feed-forward perceptron (MLP) that contains many interconnected layers of neurons (Raczko and Zagajewski 2017). Further details about ANN can be found in (Civco 1993; Tucker 1979).

DT is a non-parametric classifier. It consists of a root-nodes-branches-leaf flowchart that is created by splitting the data recursively using the highly informative attributes at each tree node. Since they are robust against noisy data, therefore, DTs have also been used widely for tree species extraction from remote sensing data, e.g. (Zhang and Hu 2012).

RF is based on ensemble machine learning techniques which has proven to be very effective for classification of remote sensing data (Shukla et al. 2017). The RF classifier is similar to DT however, it is robust against deviation from normal distribution, and can be used for both classifications and regressions (Berhane et al.

2018). Since, RF uses different subsets of the same training dataset therefore, it avoids errors of bias and variance. Refer to (Belgiu and Drăguţ 2016) for further details about its application in remote sensing domain.

The most commonly used one-against-all model of SVM was selected while both Gaussian and polynomial kernels were evaluated on our dataset. The values for regularization parameter (C), Gaussian kernel parameter γ and polynomial kernel d were determined by applying a grid search algorithm on a predefined range using ten-fold cross validation. Similarly, for ANN sigmoid and tanh activation functions were selected and the optimal values for learning rate and momentum were also obtained by applying grid search algorithm. Similarly, the required parameters for each classifier were fine-tuned using validation dataset. MATLAB © environment and LIBSVM (Chang and Lin 2011) were used to perform all experiments.

3. Experimental Results

Experiments were conducted using a dataset containing ten high-resolution digital orthoimages to evaluate the effectiveness of the proposed method. Two main classes were defined for classification of vegetation areas: tea gardens and other trees. For all experiments, a balanced dataset was considered for both types of classes. 5 images were selected from geographically different regions to account for varying reflectance information for selection of training objects while rest of the objects from all ten images were used for testing. We employed ten-fold cross-validation to train the classifiers.

The SVM classifier required tuning of three parameters: cost (C) for SVM, RBF kernel parameter gamma γ and polynomial kernel parameter degree (d). To obtain these parameters, an exhaustive grid search was performed in a predefined range: $C = [2^{-5}, 2^{-3}, \ldots . 2^{9}]$, $\gamma = [2^{-9}, 2^{-7}, \ldots . . 2^{3}]$ and $d = [0,1, \ldots .5]$. It is also worth

mentioning here that the scaling of the features did not affect the classification accuracy of the classifier. Therefore, the features extracted were used without scaling for both training and testing phases.

Similarly, ANN also required tuning of two parameters: momentum and learning rate. The learning rate was searched between 0.001 and 0.5 while momentum between 0 and 0.9. The optimal values obtained were 0.1 and 0.8 for learning rate and momentum respectively. Other parameters for ANN, also known as topological parameters, were fixed for all experiments as they are not usually data dependent. The number of iterations was set to 1000. A higher value made sure that the model will converged before reaching the limit for number of iterations during training. In addition, only one hidden layer with 200 neurons was used which was empirically calculated. Similarly, the number of neurons in the input layer were set according to the number of input features (13) while the number of output neurons were set to 2 (according to the number of output classes). Further, we experimented with two most widely used activation functions (sigmoid and tanh).

RF required tuning of two parameters: the number of predictors (NP) and number of random trees (NT). To find the optimal RF model for classification, a range of values for both parameters were tested and evaluated: NP = 50:250 with a step size of 50 and NT= 1:13 with a step size of 1. The optimal values selected for NP and NT were 3 and 50 respectively.

Similarly, for DT two parameters were fine-tuned: tree depth (TD) and Maximum Features (MF). TD, which represents the depth of the tree, was empirically calculated as 11 while MF, which represents the number of features to consider when performing a best split at each node, was set to 13 (all features).

The outputs obtained from the proposed method were compared with reference data using an area-based method. It was based on: true positives (TP), false positive (FP), false negatives (FN) and true negatives (TN) metrics. Quantitative information were obtained from these metrics using precision (P), recall (R), F-measure (FM), overall accuracy (OA) and Kappa coefficient (KC). For SVM, RBF kernel resulted in 85.19% and 70.95% precision and recall respectively, while polynomial kernel produced 82.13% and 72.93% precision and recall respectively. For ANN, precision and recall for sigmoid function were 87.86% and 84.52% respectively, while for tanh functions precision and recall were 83.50% and 84.52 % respectively. Similarly, RF classifier produced precision 87.90% and recall 86.66% while DT resulted in 83.25% precision and 79.50% recall.

Sometimes, these two measures are combined into a single measure called f-measure, which is the harmonic mean of precision and recall. The f-measure produced for each method were, SVM (RBF: 77.05%, polynomical:76.99%) , ANN (sigmoid: 87.36% and tanh:84.00%), RF : 87.21% and DT: 81.93%. The average OA and KC obtained for tea gardens using RBF were 84.53% and 82.71 % respectively. Similarly, the same metrics for polynomial kernel were 79.15% and 79.22% respectively. For ANN, sigmoid activation function produced slightly better results (KC: 0.86, OA: 87.02%) compare to tanh function (KC:0.80, OA: 83.19%). The results of RF were similar to ANN (KC: 0.85, OA: 82.21%). DT produced KC and OA as 82.21% and 0.80 respectively. Table 4 summarizes the overall results obtained for the proposed method to extract tea gardens from high-resolution digital orthoimages.

Furthermore, the statistical significance of differences between classifier results were computed using McNemar's test with a 95% confidence interval. It is a nonparametric test based on the standardized normal test statistics, as defined below:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}}$$

where f_{12} is the number of samples correctly classified by classifier 1 but misclassified by classifier 2, and f_{21} is the number of samples misclassified by classifier 1 but correctly classified by classifier 2 (Foody 2004). The difference in accuracy between the two classifiers is said to be statistically significant $|Z| > 1.96$. The sign of Z indicates whether classifier 1 is more accurate ($Z>0$) or vice versa ($Z< 0$) (Fauvel et al. 2013). Table 5 summarizes the paired statistical significance t-test results for different classifiers.

Additional experiments were performed to evaluate the minimum amount of training objects required to obtain good classification accuracy for each classifier. The amount of training data selected for each class (tea garden and other trees) was [5, 50,100, 300,500, 800, 1000]. The OA obtained for each classifier using varying number of training data samples are shown in Figure 5. As expected, relatively smaller number of sample objects (<300) resulted in less OA, while increasing the number of training sample objects for each class tend to produce higher accuracy. However, increasing the number of sample objects beyond 300 did not show any significant change in the classification accuracy as shown in the Figure 5. For the current dataset, this value (300) served as a threshold for training the classifiers.

Figure 6 (a) and (b) show a sample image and its reference data, (c) and (d) show the results obtained from SVM using RBF and polynomial kernel and (e) and (f) show the results obtained for ANN classifier with sigmoid and tanh activation functions respectively.

Figure 5

4 Discussion and Conclusions

In this study, we investigated an object-based method (mean shift) integrated with most widely used supervised machine learning approaches (SVM and ANN) for delineation of tea gardens from high resolution digital orthoimages. First, the image pixels were transformed into an object-based representation using mean-shift algorithm. Then, for each object, a set of textural and spatial/spectral features were extracted. Since, the efficiency of a supervised classifier is affected by the quality of features, therefore, a feature selection using forward sequential search was employed to select highly informative features. Finally, the optimal features were fed into SVM,ANN, RF and DT classifiers separately to generate classification maps for tea gardens and other types of trees. The results obtained on the data set confirmed the effectiveness of the proposed approach and achieved high classification accuracy, which is a valuable contribution for agricultural related task. The resulting classified images also contained some misclassifications. This is not unusual for remote sensing data classification due to similarity in textural and spectral information between two classes of interest.

Figure 6

As for many real applications, it is important to find the minimum amount of data required to train the classifiers. In this context, we performed several experiments to find the ideal amount of training objects required to training the classifiers, which could still produce optimal results. Although the overall response for both classifiers was relatively similar, yet after 300 training samples, the results were almost consistent and did not result in abrupt accuracy change. It indicates that, to get good classification results for current data, selection of 275 – 350 training samples is ideal size.

It is interesting to note that the RBF kernel produced higher overall accuracies compare to the polynomial kernel (~5.50%). In fact, higher polynomial degrees (3, 4, 5) not only produced unexpected low classification results but they were also costly in terms of processing time. Similarly, for ANN, sigmoid function resulted in higher classification accuracy compared to tanh activation function (~3.36%).

It is also worth mentioning here that presence of recently cultivated tea gardens reduced the classification accuracy of both classifiers. Visually they look similar to bare land (Figure 7). However, the photogrammetric experts marked them as tea gardens during reference data creation. Since, this study focused to classify vegetation areas only, therefore, recently cultivated areas were misclassified. In Figure 7, an example of recently cultivated tea garden is shown with red polygon in (a), (b) shows the reference data which shows that the area has been marked as white (representing tea garden) and (c) shows that proposed method classified these areas as non-tea gardens.

In future, we would like to extend the methodology for integration of higher spectral resolution image with LiDAR data for further improving the obtained results. LiDAR data can provide additional information, such as height, density etc., which is useful to distinguish between tea gardens and rest of trees. Moreover, automatic feature selection methods can be introduced to reduce the burden of manual feature extraction and selection. In addition, the current work did not take shadows into account. Their presence can increase the miss classification rate. Therefore, we will also develop a pre-processing method that could reduce the impact of shadows from multispectral imagery.

Figure 7

Acknowledgment

We are thankful to the EMI Group Turkey for providing the important data for this study which is a part of TEYDEP Project entitled “Development of Object Based Neural Network Image Processing System Determination of Vegetation and Forestry

Boundaries" (Project Nr. 7140512). It was supervised by EMI Group-Turkey, and consulted by Prof. Dr. Bulent Bayram.

Update the reference to include our MESEAP

Reference

Akar, and O. Güngör. 2015. "Integrating Multiple Texture Methods and NDVI to the Random Forest Classification Algorithm to Detect Tea and Hazelnut Plantation Areas in Northeast Turkey." *International Journal of Remote Sensing* 36 (2): 442–464. doi:10.1080/01431161.2014.995276.

Belgiu, Mariana, and Lucian Drăguţ. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114 (April): 24–31. doi:10.1016/j.isprsjprs.2016.01.011.

Berhane, Tedros, Charles Lane, Qiusheng Wu, Bradley Autrey, Oleg Anenkhonov, Victor Chepinoga, and Hongxing Liu. 2018. "Decision-Tree, Rule-Based, and Random Forest Classification of High-Resolution Multispectral Imagery for Wetland Mapping and Inventory." *Remote Sensing* 10 (4): 580. doi:10.3390/rs10040580.

Blaschke, T. 2010. "Object Based Image Analysis for Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (1). Elsevier B.V.: 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Blaschke, Thomas, Geoffrey J Hay, Maggi Kelly, Stefan Lang, Peter Hofmann, Elisabeth Addink, Raul Queiroz Feitosa, et al. 2014. "Geographic Object-Based Image Analysis – Towards a New Paradigm." *ISPRS Journal of Photogrammetry and Remote Sensing* 87 (January). International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS): 180–191. doi:10.1016/j.isprsjprs.2013.09.014.

Camps-Valls, Gustavo, and Lorenzo Bruzzone. 2005. "Kernel-Based Methods for Hyperspectral Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 43 (6): 1351–1362. doi:10.1109/TGRS.2005.846154.

Chang, Chih-Chung, and Chih-Jen Lin. 2011. "LIBSVM : A Library for Support Vector Machines." *ACM Transactions on Intelligentnt Systems and Technology* 2 (3). ACM: 1–27. doi:10.1145/1961189.1961199.

Chehata, N., C. Orny, S. Boukir, D. Guyon, and J.P. Wigneron. 2014. "Object-Based Change Detection in Wind Storm-Damaged Forest Using High-Resolution Multispectral Images." *International Journal of Remote Sensing* 35 (13): 4758–4777. doi:10.1080/01431161.2014.930199.

Cheng, Yizong. 1995. "Mean Shift, Mode Seeking, and Clustering." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (8): 790–799. doi:10.1109/34.400568.

Civco, Daniel L. 1993. "Artificial Neural Networks for Land-Cover Classification and Mapping." *International Journal of Geographical Information Systems* 7 (2): 173–186. doi:10.1080/02693799308901949.

Comaniciu, Dorin, and Peter Meer. 2002. "Mean Shift: A Robust Approach toward Feature Space Analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (5): 603–619. doi:10.1109/34.1000236.

Costa, Hugo, Giles M. Foody, and Doreen S. Boyd. 2018. "Supervised Methods of Image Segmentation Accuracy Assessment in Land Cover Mapping." *Remote Sensing of Environment* 205 (December 2016). Elsevier: 338–351. doi:10.1016/j.rse.2017.11.024.

Dalponte, Michele, Hans Ole Orka, Terje Gobakken, Damiano Gianelle, and Erik Naesset. 2013. “Tree Species Classification in Boreal Forests with Hyperspectral Data.” *IEEE Transactions on Geoscience and Remote Sensing* 51 (5): 2632–2645. doi:10.1109/TGRS.2012.2216272.

Dihkan, Mustafa, Nilgun Guneroglu, Fevzi Karsli, and Abdulaziz Guneroglu. 2013. “Remote Sensing of Tea Plantations Using an SVM Classifier and Pattern-Based Accuracy Assessment Technique.” *International Journal of Remote Sensing* 34 (23). Taylor & Francis: 8549–8565. doi:10.1080/01431161.2013.845317.

Fauvel, Mathieu, Yuliya Tarabalka, Jón Atli Benediktsson, Jocelyn Chanussot, and James C. Tilton. 2013. “Advances in Spectral-Spatial Classification of Hyperspectral Images.” *Proceedings of the IEEE* 101 (3): 652–675. doi:10.1109/JPROC.2012.2197589.

Foody, Giles M. 2004. “Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy.” *Photogrammetric Engineering & Remote Sensing* 70 (5): 627–633. doi:10.14358/PERS.70.5.627.

Fukunaga, K., and L. Hostetler. 1975. “The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition.” *IEEE Transactions on Information Theory* 21 (1): 32–40. doi:10.1109/TIT.1975.1055330.

Gu, Yanfeng, Tianzhu Liu, Xiuping Jia, Jon Atli Benediktsson, and Jocelyn Chanussot. 2016. “Nonlinear Multiple Kernel Learning with Multiple-Structure-Element Extended Morphological Profiles for Hyperspectral Image Classification.” *IEEE Transactions on Geoscience and Remote Sensing* 54 (6): 3235–3247. doi:10.1109/TGRS.2015.2514161.

Huang, Xin Huang Xin, and Liangpei Zhang Liangpei Zhang. 2008. “An Adaptive Mean-Shift Analysis Approach for Object Extraction and Classification From Urban Hyperspectral Imagery.” *IEEE Transactions on Geoscience and Remote Sensing* 46 (12): 4173–4185. doi:10.1109/TGRS.2008.2002577.

Jamil, Akhtar, and Bulent Bayram. 2018. “Tree Species Extraction and Land Use/Cover Classification From High-Resolution Digital Orthophoto Maps.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11 (1): 89–94. doi:10.1109/JSTARS.2017.2756864.

Jamil, Akhtar, Bulent Bayram, and Dursun Zafer Seker. 2019. “Mapping Hazelnut Trees from High Resolution Digital Orthophoto Maps : A Comparative Analysis of an Object and Pixel-Based Approach.” *Fresenius Environmental Bulletin* 28 (2): 561–567.

Maxwell, Aaron E., Timothy A. Warner, and Fang Fang. 2018. “Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review.” *International Journal of Remote Sensing* 39 (9). Taylor & Francis: 2784–2817. doi:10.1080/01431161.2018.1433343.

Melgani, Farid, and Lorenzo Bruzzone. 2004. “Classification of Hyperspectral Remote Sensing Images with Support Vector Machines.” *IEEE Transactions on Geoscience and Remote Sensing* 42 (8): 1778–1790. doi:10.1109/TGRS.2004.831865.

Ozdarici Ok, Asli, and Zuhal Akyurek. 2012. “A Segment-Based Approach to Classify Agricultural Lands by Using Multi-Temporal Optical and Microwave Data.” *International Journal of Remote Sensing* 33 (22): 7184–7204. doi:10.1080/01431161.2012.700423.

Pal, M, and P M Mather. 2005. “Support Vector Machines for Classification in Remote Sensing.” *International Journal of Remote Sensing* 26 (5): 1007–1011. doi:10.1080/01431160512331314083.

Pedram Ghamisi, Javier Plaza, Yushi Chen, Jun Li, and Antonio Plaza. 2017. “Advanced Supervised Spectral Classifiers for Hyperspectral Images: A Review.” *IEEE Geoscience and Remote Sensing Magazine* 5 (1): 8–32. http://homepage.hit.edu.cn/chenyushi.

Pham, Lien T.H., Lars Brabyn, and Salman Ashraf. 2016. “Combining QuickBird, LiDAR, and GIS Topography Indices to Identify a Single Native Tree Species in a Complex Landscape Using an Object-Based Classification Approach.” *International Journal of Applied Earth Observation and Geoinformation* 50 (August). Elsevier B.V.: 187–197. doi:10.1016/j.jag.2016.03.015.

Pipaud, Isabel, and Frank Lehmkuhl. 2017. “Object-Based Delineation and Classification of Alluvial Fans by Application of Mean-Shift Segmentation and Support Vector Machines.” *Geomorphology* 293 (May). Elsevier: 178–200. doi:10.1016/j.geomorph.2017.05.013.

Qian, Yuguo, Weiqi Zhou, Jingli Yan, Weifeng Li, and Lijian Han. 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.

Qin, Rongjun. 2014. “A Mean Shift Vector-Based Shape Feature for Classification of High Spatial Resolution Remotely Sensed Imagery.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (5): 1974–1985. doi:10.1109/JSTARS.2014.2357832.

Raczko, Edwin, and Bogdan Zagajewski. 2017. “Comparison of Support Vector Machine, Random Forest and Neural Network Classifiers for Tree Species Classification on Airborne Hyperspectral APEX Images.” *European Journal of Remote Sensing ISSNOnline) Journal European Journal of Remote Sensing* 50 (1). Taylor & Francis: 144–154. doi:10.1080/22797254.2017.1299557.

Sasaki, Takeshi, Junichi Imanishi, Keiko Ioki, Yukihiro Morimoto, and Katsunori Kitada. 2012. “Object-Based Classification of Land Cover and Tree Species by Integrating Airborne LiDAR and High Spatial Resolution Imagery Data.” *Landscape and Ecological Engineering* 8 (2): 157–171. doi:10.1007/s11355-011-0158-z.

Shukla, Gaurav, Rahul Dev Garg, Hari Shanker Srivastava, and Pradeep Kumar Garg. 2017. “Implementation of Random Forest Algorithm for Crop Mapping across an Aridic to Ustic Area of Indian States.” *Journal of Applied Remote Sensing* 11 (2): 026005. doi:10.1117/1.JRS.11.026005.

Street, George. 2010. “Evaluation of Spectral and Texture Features for Object-Based Vegetation Species Classification Using Support Vector Machines.” *Symposium A Quarterly Journal In Modern Foreign Literatures* XXXVIII (7A): 122–127. http://www.isprs.org/proceedings/XXXVIII/part7/a/pdf/122_XXXVIII-part7A.pdf.

Su, Tengfei, Hongyu Li, Shengwei Zhang, and Yongxiang Li. 2015. “Image Segmentation Using Mean Shift for Extracting Croplands from High-Resolution Remote Sensing Imagery.” *Remote Sensing Letters* 6 (12): 952–961. doi:10.1080/2150704X.2015.1093188.

Tucker, Compton J. 1979. “Red and Photographic Infrared Linear Combinations for Monitoring Vegetation.” *Remote Sensing of Environment* 8 (2): 127–150. doi:10.1016/0034-4257(79)90013-0.

Zarea, Asghar, and Ali Mohammadzadeh. 2015. “A Novel Building and Tree Detection Method From LiDAR Data and Aerial Images.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 1–12. doi:10.1109/JSTARS.2015.2470547.

Zhang, Kongwen, and Baoxin Hu. 2012. “Individual Urban Tree Species Classification Using Very High Spatial Resolution Airborne Multi-Spectral Imagery Using Longitudinal Profiles.” *Remote Sensing* 4 (6): 1741–1757. doi:10.3390/rs4061741.

Zylshal, Sayidah Sulma, Fajar Yulianto, Jalu Tejo Nugroho, and Parwati Sofan. 2016. “A Support Vector Machine Object Based Image Analysis Approach on Urban Green Space Extraction Using Pleiades-1A Imagery.” *Modeling Earth Systems and Environment* 2 (54). Springer International Publishing: 54. doi:10.1007/s40808-016-0108-8.

Table 1: Technical Specification of UltraCamX

Parameter Name	Value
Ground Sample Distance	30 cm
Capture Dates	13-03-2013
Side Laps	30%
End Laps	70%
Flight Altitude	4200 m
Sensor Type	Aerial
Radiometric resolution	8 bits

Table 2. Object-based spatial/spectral and textural features

Feature Name	Feature Description	Number of features
Mean Standard Deviation NDVI	Mean value for each band SD for each band NDVI value	4 4 1
Homogeneity Contrast Energy Correlation	GLCM texture features obtained for each object	4
	Total Features selected	13

Table 3 Correlation matrix of selected features.

Features	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
F2	0.24												
F3	0.47	0.61											
F4	0.56	0.63	0.53										
F5	0.24	0.17	0.20	-0.16									
F6	0.27	0.20	0.23	-0.19	0.58								
F7	0.34	0.26	0.28	-0.09	0.31	0.63							
F8	-0.09	-0.11	-0.10	-0.16	0.67	0.63	0.59						
F9	0.35	0.30	0.34	-0.13	0.43	0.47	0.51	-0.42					
F10	0.35	0.28	0.33	-0.15	0.43	0.38	0.68	0.60	-0.67				
F11	-0.05	0.03	-0.01	0.35	-0.69	-0.71	-0.69	-0.54	0.51	0.26			
F12	0.39	0.36	0.39	-0.01	0.61	0.47	0.28	0.36	-0.44	-0.52	-0.09		
F13	0.05	0.11	0.09	0.28	-0.46	-0.45	-0.44	-0.44	0.38	-0.12	0.45	-0.32	

Note: F1 = Mean Red; F2=Mean Green; F3 =Mean Blue; F4 = Mean NIR; F5 =Standard Deviation Red; F6 = Standard Deviation Green; F7 = Standard Deviation Blue; F8 = Standard Deviation NIR; F9=Homogeneity; F10=Energy; F11=Contrast; F13=Correlation;

Table 4. Results obtained for extraction of tea gardens (%)

Class	Classifier	Kernel Type	P	R	FM	OA	KC
Tea Fields	SVM	RBF	85.19	70.95	77.05	84.53	0.82
		Polynomial	82.13	72.93	76.99	79.15	0.79
	ANN	Sigmoid	87.86	86.96	87.36	87.02	0.86
		Tanh	83.50	84.52	84.00	83.19	0.80
	RF	-	87.90	86.66	87.21	86.66	0.85
	DT	-	83.25	79.50	81.93	82.21	0.80

Table 5 McNemar's statistical test to compare the classifiers with $|z| > 1.96$ at $\alpha = 0.05$

	SVM	ANN	RF	DT
SVM (RBF)	N/A	-	-	-
ANN (Sig)	5.53	N/A	-	-
RF	3.7926	1.66	N/A	-
DT	2.286	3.86	3.26	N/A

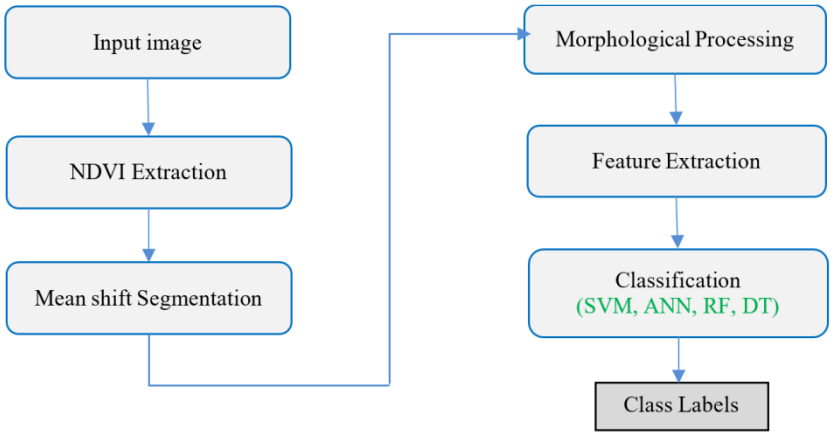

Figure 1. Scheme of the proposed method for tea garden extraction

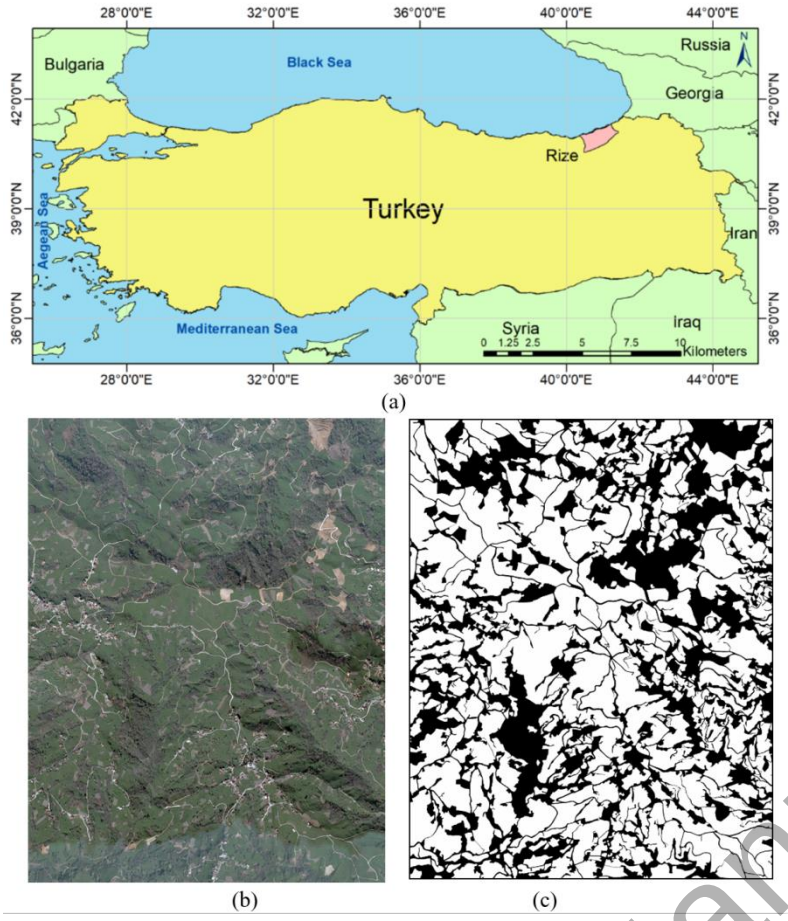

Figure 2. a) Map of Turkey showing location of the study area (Rize district) b) A sample digital orthoimage c) Reference image for sample image, where white pixels represent tea gardens

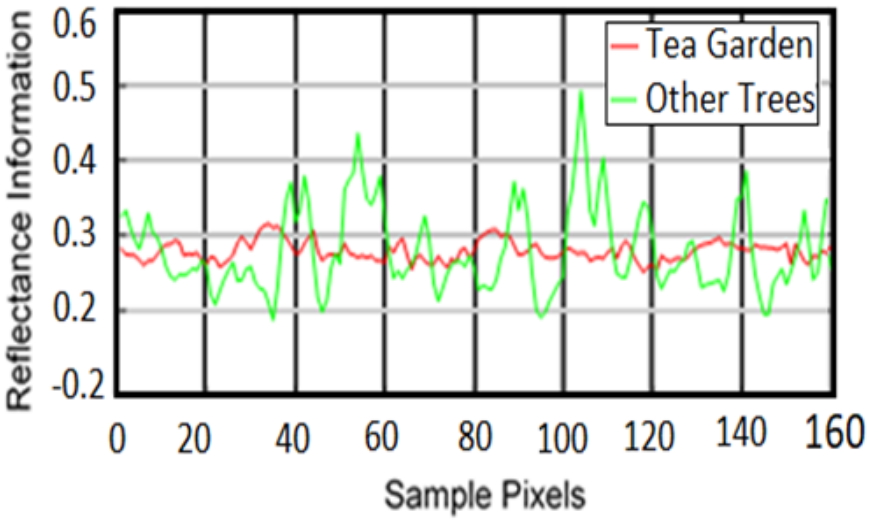

Figure 3. NDVI profiles for sample classes

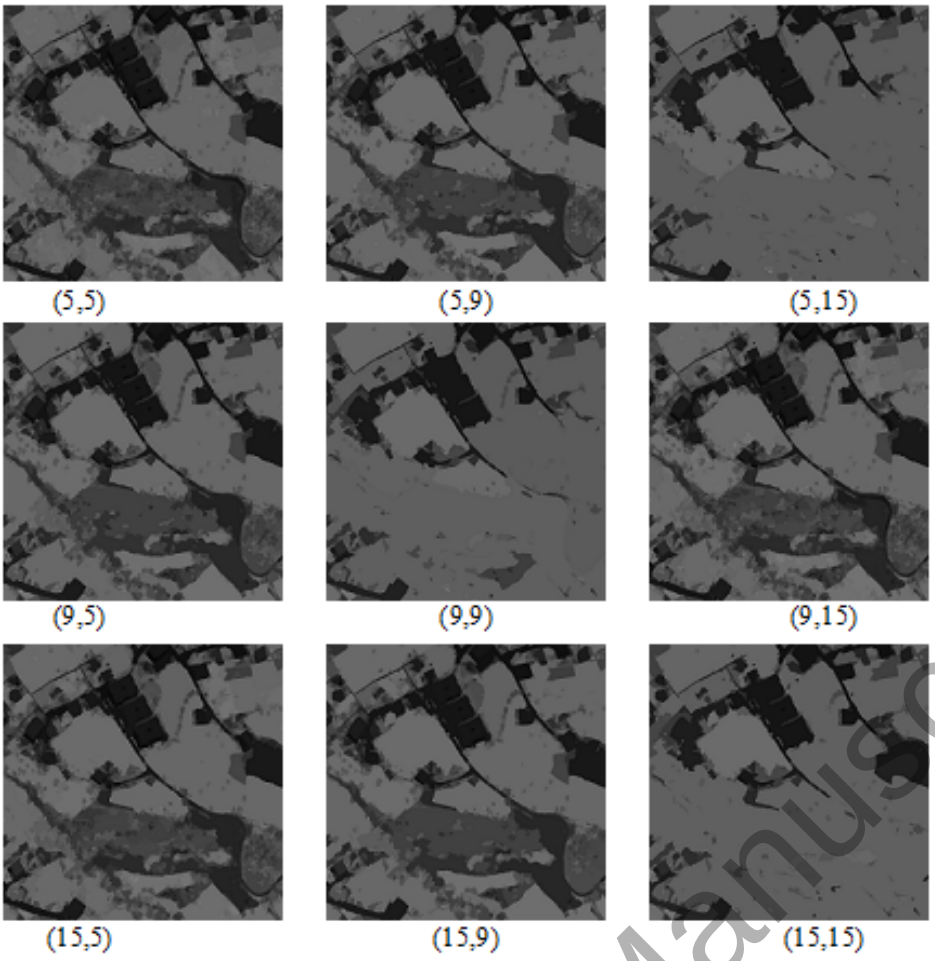

Figure 4. The impact of bandwidth (h_s, h_r) selection on segmentation of NDVI image using mean shift algorithm.

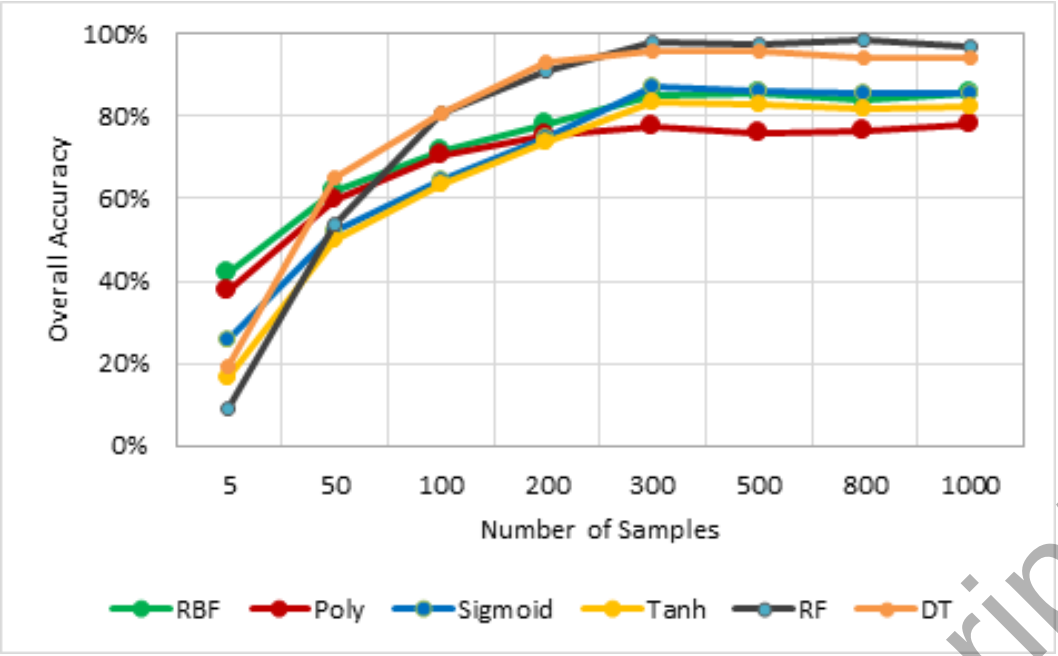

Figure 5. Analysing the effect of the training data size on the classification accuracy of the SVM and ANN classifier.

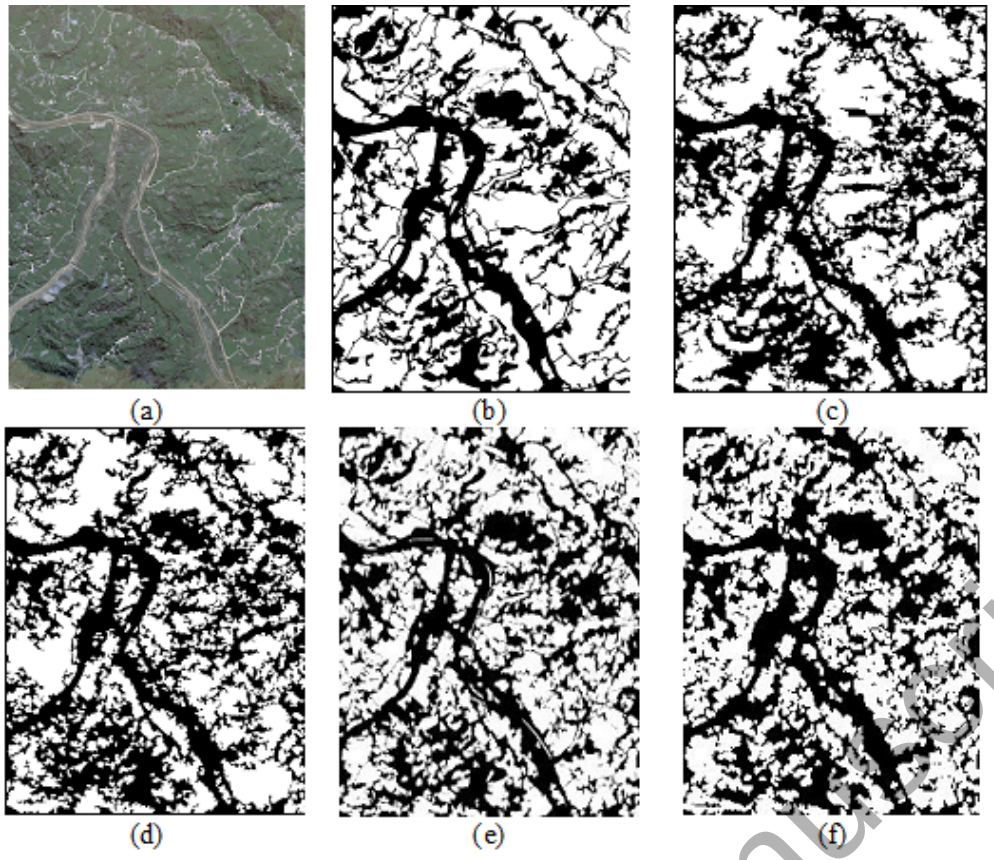

Figure 6. (a) A sample input image in RGB space, (b) Reference data for the sample image, The results obtained from each classifier (c) SVM (d) ANN (e) RF (f) DT

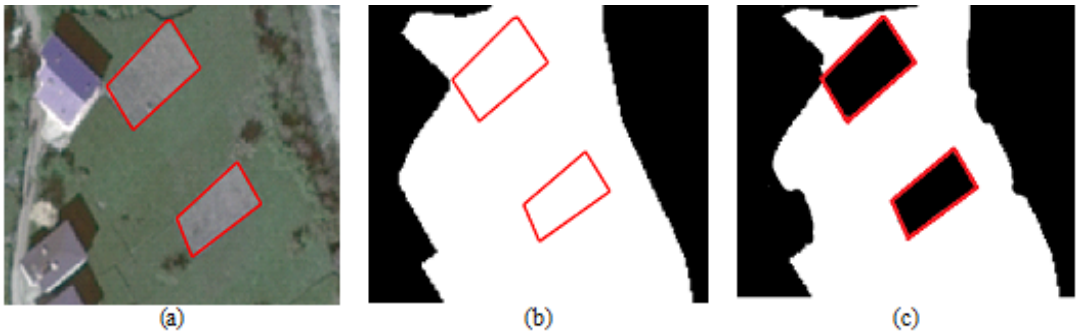

Figure 7. (a) A sample of recently cultivated tea garden (b) reference data (c) results obtained from proposed method