*Journal of Engineering Technology (ISSN. 0747-9964)*
*Volume 8, Issue 1, Jan. 2019, PP.1-12*

RESEARCH ARTICLE -ENGINEERNG TECHNOLOGY ASEE

# A Kernel Functions Analysis For Support Vector Machines To Identify ARGAN Forest Using Sentinel-2 Images

**Omar EL KHARKI[1], Miriam Wahbi[1], Jamila Mechbouh[2], Danielle Ducrot[3], Ali Essahlaoui[4], Khalid Auhmani[5]**

[1]Faculty of Sciences and Techniques of Tangier, BP 416, Tangier, Morocco.
[2]RIFGeomatique, Tangier,Morocco.
[3]CESBIO, Toulouse, France.
[4]Faculty of Sciences, Meknes, Morocco.
[5]National School of Applied Sciences, Safi, Morocco.

**Abstract:** Support vector machines (SVM) is a non-parametric supervised classification derived from statistical learning theory that has been recently used in the remote sensing field. Accuracy of SVM method is influenced by the type of kernel function and its parameters. The purpose of this research is to investigate the sensitivity of SVM architecture including internal parameters and kernel types on land cover classification accuracy of Sentinel-2 for the study of area in South-Western Morocco, exhibiting to intense degradation of Argan forest. 256 models for each kernel (radial basis function (RBF), polynomial and sigmoid) were investigated by a grid search method using cross validation approach for the selection of the optimal parameters (penalty C, gamma $\gamma$ and degree d) for each kernel. The results suggest that the choice of model parameters and kernel type have an important impact on the accuracy of the SVM classification.

Three kernels with their estimated optimum parameters were applied for the classification of Sentinel-2 image and the results were analyzed. The Results showed that the best model of RBF kernel with (C=2^7, $\gamma$=2^7), outperforms all SVMs models in terms of overall and individual class accuracies. This study allowed us the verification of the effectiveness and the robustness of RBF kernel in monitoring and mapping Argan forest over large areas.

**Keywords:** SVM, Kernel, RBF, Polynomial, Sigmoid, Argan forests, Sentinel-2, Classification, Grid Search, Cross Validation.

## 1 Introduction

Argan tree (Argania spinosa) is a species of tree endemic to the calcareous semi-desert Souss-Massa valley of southwestern Morocco. The Argan forest supports many important ecological and socioeconomic functions. Currently, in spite of the biosphere reserve label attributed by UNESCO in 1998, the threat of the degradation of the sparse Argan forest is a main concern for both local population and scientists. Since several decades, a decrease of the extension area of the Argan forest and of the tree density has been observed.

The development of remote sensing technology has increasingly facilitated the acquisition of such information. How to extract accurate and timely knowledge about land use/ land cover from remote sensing imagery relies upon not only on the data quality and resolution, but also the classification techniques used. Therefore, the improvement of remote sensing classification accuracy is always a point in case worth being debated.

During the past three decades, many classification techniques have been developed to automate the identification of Argan forest, including supervised maximum likelihood classifiers [1]. It is a parametric approach that assumes the class signature has normal distribution. Although this assumption is generally valid, it is invalid for classes consisting of several subclasses or for classes having different spectral features [2],[3]. To overcome this problem, some non-parametric classification techniques such as artificial neural networks, decision trees and Support vector machines (SVMs) have been recently introduced. Support vector machines (SVMs) have been also proposed to improve classification performance due to their characteristics of generalization [4]. SVMs are the supervised learning method used for classification and regression problems based on statistical learning theory [5].

Recent studies have suggested that support vector machine (SVM) can provide good results for hyperspectral remote sensing classification and superior results have been reported compared to traditional remote sensing classification algorithms such as maximum likelihood (ML), k-nearest neighbor, and artificial neural networks (ANN) [6], [7], [8]. The most appealing property of SVM was the high capacity for generalization with relatively small numbers of training data [8]. However, the potential of SVM has not received much attention for SENTINEL-2 data classification. Specifically, the performance of SVM classification has not been thoroughly assessed by the type of kernel function and its parameters, and the application of this approach in Argan forest environments is comparatively understudied.

In this study, SVMs were used for land use/land cover classification of Agadir region (Morocco) using Sentinel-2 acquired on 2015. The study area is the region of Argan forest which is the subject of intense degradation. The objective of this study is to evaluate the performances of different kernel functions and the effects of its parameters. It should be pointed out that since the selection of kernel parameters have an effect on the generalization ability of SVMs, the determination of their optimum parameters is regarded as critical for the success of the classification.

This paper is organized as follows. In Section two, we introduce some related background including some basic concepts of SVM, kernel function selection, and parameters selection of SVM. In Section three, we present the study area and the adopted methodology. Next, results and discussions will be presented in section four. Finally, a conclusion and future works are presented in section five.

## 2. Support Vector Machines

In this section we introduce the reader to some basic concepts of SVM, kernel function, and the selection of parameters.

### 2.1 Overview of SVM

Support vector machines (SVM) classifier is a non-parametric supervised classification derived from statistical learning theory. SVM was developed in the late 1970s, but its popularity in remote sensing only, started increasing about a decade ago [9]. The underlying theory and the detailed mathematical explanation of SVM have been demonstrated in many previous studies [10].

This technique consists of finding the optimal separation surface between classes due to the identification of the most representative training samples called the support vectors. If the training dataset is not linearly separable, a kernel method is used to simulate a non-linear projection of the data in a higher dimensional space, where the classes are linearly separable [11].

In the two-class case, the goal of the SVM is to find a hyperplane that maximizes the distance from the members of each class to the optimal hyperplane. Let the training data of two separable classes with k samples be represented by $\{(X_1, Y_1), \dots, (X_n, Y_n)\}$, where $X_i \in R^N$, is an n-dimensional vector and $Y_i \in \{-1, +1\}$ is the class label. Optimum hyperplane works in the manner that maximizes the margin between the classes. As it can be seen from this hyperplane is defined as $w.x_i + b = 0$, where x is a point lying on the hyperplane, parameter w determines the orientation of the hyperplane in space, b is the bias that the distance of hyperplane from the origin. For the linearly separable case, a separating hyperplane can be defined for two classes as:

$$w.x_i + b \geq +1 \;\; \forall\, y = +1 \qquad (1)$$

These inequalities can be combined into a single inequality:

$$y_i(w.x_i + b) \geq 1 \qquad (2)$$

The training data points on these two hyperplanes, which are parallel to the optimum hyperplane and defined by the functions $w.x_i + b = \pm 1$, are the support vectors [12]. If a hyperplane exists that satisfies Eq. (2), the classes are linearly separable. Therefore, the margin between these planes is equal to $2/_{\|w\|}$. As the distance to the closest point is $2/_{\|w\|}$, the optimum separating hyperplane can be found by minimizing $\|w\|^2$under the constraint Eq. (2). Thus, the determination of optimum hyperplane is required to solve optimization problem given by:

$$\begin{cases} \qquad\qquad min\dfrac{1}{2}\|w\|^2 \\ subjecttoconstraints, \qquad y_i(w.x_i + b) \geq 1 \end{cases} \qquad (3)$$

This concept can be extended to the case when the classes are not linearly separable. A slack variable, $\xi_i$ i =1,…, k can be introduced such that (2) can be written as :

$$y_i(w.x_i + b) - 1 + \xi_i \geq 0 \qquad (4)$$

While the objective function is supplemented to keep the constraint violation as small as possible:

$$\underset{w,\xi_i,b}{min} \left\{\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}\xi_i\right\} \qquad (5)$$

Where C is a regularization constant or penalty parameter

This result is a quadratic optimization problem which is solved by a standard numerical optimization package. The minimization procedure uses Lagrange multipliers and Quadratic Programming optimization methods. If $\lambda_i$, i = 1, . . ., k are the non-negative Lagrange multipliers associated with constraint (5), the optimization problem becomes one of maximizing (Osuna et al., 1997):

$$L(\lambda) = \sum_i \lambda_i - \frac{1}{2}\sum_{i,j}\lambda_i\lambda_j y_i y_j (x_i \,.\, x_j) \qquad (6)$$

under constraints $\lambda_i$ = 0, i = 1, . . ., k.

The training vectors x are only used in inner products which can be replaced by a kernel function $k(x_i, x_j)$ that satisfy Mercer’s condition. This is equivalent to mapping the feature vectors into a high-dimensional feature space before using a hyperplane classifier there.

### 2.2 Kernel Selection of SVMs

The kernel function allows for the training data to be projected in a larger space where it may be increasingly possible to discover a superior separating margin for the optimal separating hyperplane. There are many kernel functions in SVM, so how to select a good kernel function is also a research issue. However, for general purposes, there are some popular kernel functions [13]:

- Linear kernel: $k(x_i, x_j) = x_i^T x_j$
- Polynomial kernel: $k(x_i, x_j) = (\gamma x_i^T x_j + r)^d \quad , \ \gamma > 0$
- RBF kernel: $k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|\right), \ \gamma > 0$
- Sigmoid kernel: $k(x_i, x_j) = \tanh\left(\gamma(x_i x_j) - r\right), \gamma > 0$

Here, γ, r and d are kernel parameters. The choice of kernel used for a problem and the parameters selected can have an effect on the speed and accuracy of the classification.

### 2.3 Parameters Selection of SVM

Parameter selection is also an important issue in SVM. Recently, SVM have shown good performance in data classification. Its success depends on the tuning of several parameters which affect the generalization error. If you use the linear SVM, you only need to tune the cost parameter C. Unfortunately, linear SVM are often applied to linearly separable problems. The remotely sensed data are not linearly separable. Therefore, we often apply nonlinear kernel to solve classification problems, so we need to select the cost parameter (C) and kernel parameters (γ, d) [14].

## 3. Study Area and Methodology

### 3.1 Study area

The study area selected for this research is located in the southwestern Morocco. It is situated between 9° 11’ 40’’ and 9° 1’ 30’’ West and between 30° 10’ 15’’ and 30° 18’ 30’’ North. Land is principally covered by agriculture, greenhouses and grassland. Also, a large part of the study area is covered by Argan forests. Built areas are often discontinuous and considerably less in proportion.

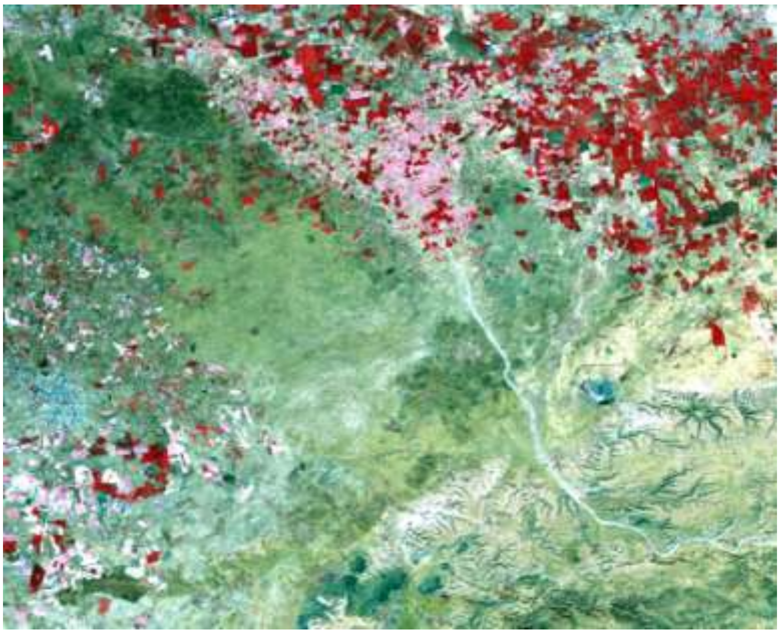


**Figure. 1.**Sentinel-2 Image of the study area

### 3.2 Sentinel-2 Data

Sentinel-2 is a state of the art sensor launched on 23 June 2015 by the European Space Agency (ESA). Sentinel-2 mission is a land monitoring constellation of two identical satellites (Sentinel-2A and Sentinel-2B) that deliver high resolution optical imagery. The system provides a global coverage of the Earth's land surface and is characterized by its high revisit time (ten days with one satellite and 5 days when Sentinel-2B becomes operational; its launch is scheduled for 2017). The system is designed to collect data at 10 m (blue, green, red and near-infrared-1) and respectively, 20 m (red edge 1 to 3, near-infrared-2, short wave infrared 1 and 2). Three additional bands for atmospheric correction are collected at 60 m (for retrieval of aerosol, water vapor, cirrus), resulting in a total of 13 spectral bands. For our study we designed to exploit the key feature of Sentinel-2 imagery, which is the availability of four 10m spatial resolution bands: B2 (490 nm), B3 (560 nm), B4 (665 nm) and B8 (842 nm).

### 3.3 Methodology

The SVMs classifier was applied to the Sentinel-2 image for mapping the land use/ land cover of the study area using the training data. SVM was implemented using three kernels which are: polynomial, sigmoid and radial basis function (RBF).

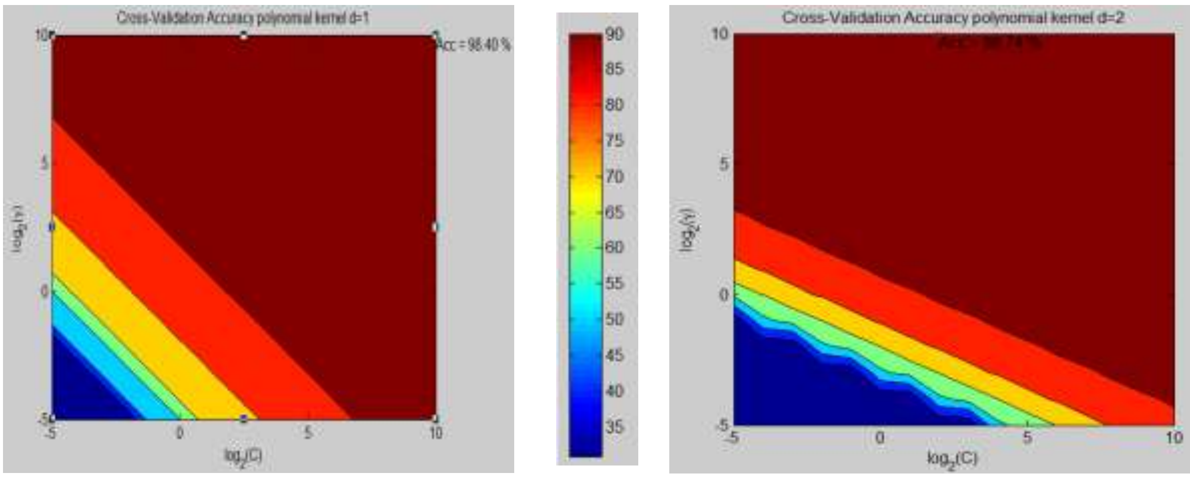
For these three kernels, the penalty parameter $C$ and gamma parameter $\gamma$, have an important role on decision boundary. The penalty parameter $C$ , also referred to as the soft margin constant, creates a soft margin that permits some misclassifications [10]. A small value of C permits ignorance of points closes to the decision boundary and thus creates a large margin between classes. The parameter $\gamma$ controls the "smoothness" of the decision boundary. It also increases in $\gamma$ lead to greater curvature of the decision boundary [10]. Large values of $C$ and $\gamma$ tend to classify training dataset accurately, but with a risk of overfitting, yielding a SVM model that could not generalize well to the rest of the data. Therefore, $C$ and $\gamma$ must be determined carefully to achieve the optimal generalization performance classification result for a given data set. Thus, optimum parameter search must be performed. In this study, the parameters of RBF, sigmoid and polynomial kernels were determined by a grid search method using cross validation approach. The main idea behind the grid search method is that different pairs of parameters are tested and the one with the highest cross validation accuracy is selected. The penalty C and γ parameters being tested in the log 2 domain within the range of [-5 to 10] stepped by 1, and therefore a total of 256 trials are made for each kernel. The parameters C and γ that achieve the highest overall accuracy are selected as the optimal parameters. In the polynomial kernel, there is an additional kernel parameter which is d (the degree of the polynomial function). We vary degree d from 1 to 6.

We firstly applied 5-fold cross validation method to the training dataset to select the best parameters C and γ . The training dataset and the selected parameters were then used to develop SVM models, which were applied to the validation dataset to examine their predictive abilities. We use LibSVM software [15] to perform SVM classification.

## 4 Results and Discussions

In this study, we focused our effort to search the optimal model for C and γ from $2^{-5}$ to $2^{10}$, and the order of the polynomial (d) ranged from 1 to 6. These are identified from the grid-search using 5-fold cross-validation.

Figure.2 shows the impact of d on the polynomial kernel performance. The grid points are chosen on a logarithmic scale and the accuracy of the classifier is estimated for each point on the grid. The optimal pair of (C, γ) is found at the brick red zone. The polynomial kernel of d=6 and ($C=2^{-2}, \gamma = 2^5$) showed the best accuracy. Note: No obvious trend in precision was observed when the polynomial order d increased from 2 to higher values. Note also that the value of the optimal pair of (C, γ) varies with the degree of the polynomial.



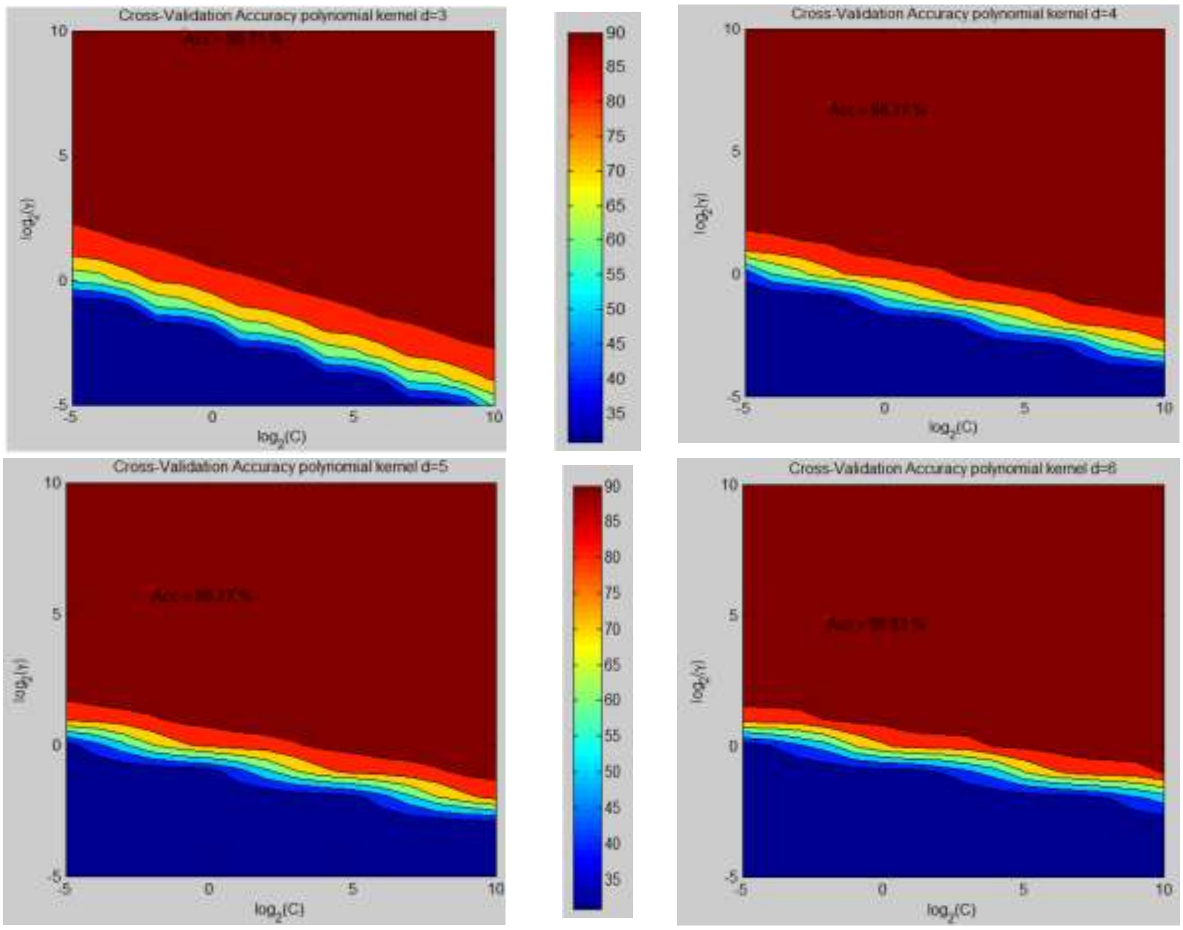
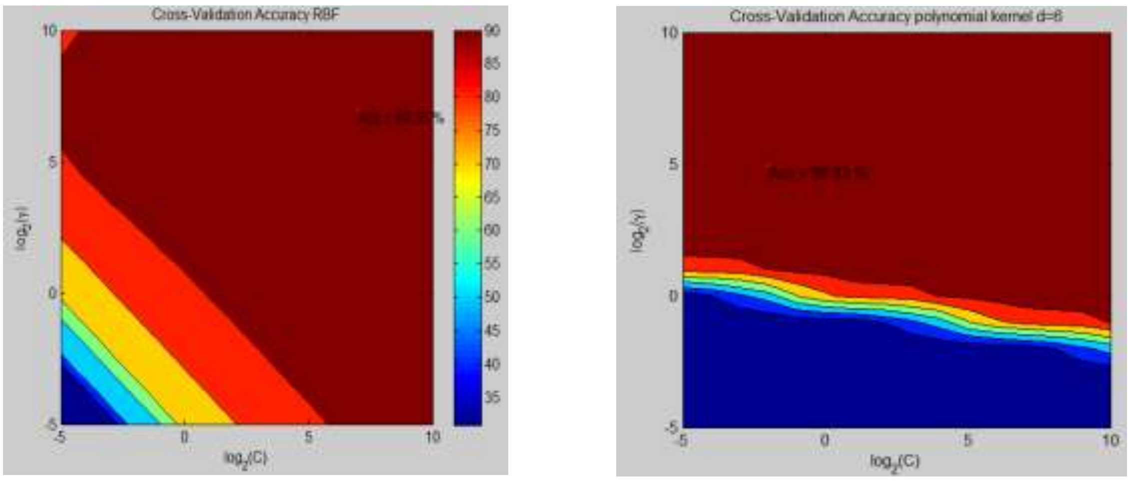



**Figure. 2.**Results obtained for polynomials kernels of various degrees



*Journal of Engineering Technology (ISSN. 0747-9964)*
*Volume 8, Issue 1, Jan. 2019, PP.1-12*

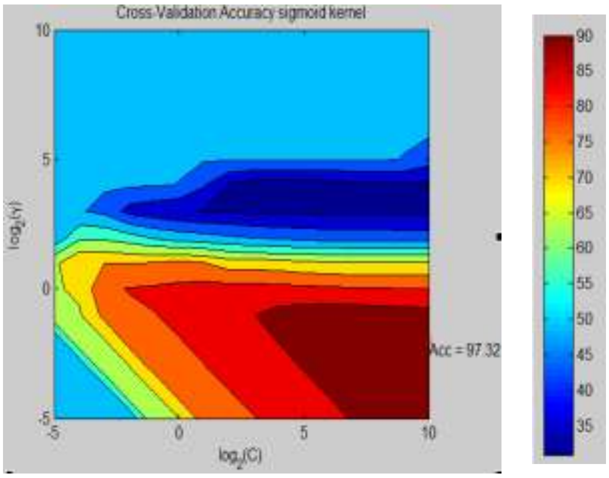




**Figure. 3.**Results obtained of comparing the three kernels

The figure. 3 presents the result of comparing the three kernels. In this figure, we present the optimal parameters C and γ, and the corresponding accuracy rates. A comparison between the three kernels reveals that the performance of the RBF and polynomial kernel is less affected by C then the sigmoid kernel. For the three kernels, the impact of γ is more significant especially for the RBF kernel. But the impact of the combinations (C, γ) is more significant. The results show that the best accuracy of the different kernels varies between 97.32% and 99.35%. The lowest value corresponds to the sigmoid kernel. The highest value has been achieved using a SVM with the radial basis kernel, with (C=$2^7$, $\gamma = 2^7$).

After testing the performances of the trained SVMs with different kernel functions and verifying their effectiveness, the best SVMs models were applied to the Sentinel-2 image covering the study area in order to produce thematic maps. Results of Sentinel-2 image classification using SVMs with radial basis function kernel, sigmoid kernel and polynomial kernel are presented in Figure. 4. Some important findings could be easily extracted from the thematic maps even after a visual interpretation. These classifiers performed well in terms of describing the spatial distribution of the different land cover classes of the study area. Firstly, it is observed that the agriculture areas were mainly located in north-east part of the study area. Secondly, it is delineated that dense Argan forest areas are mainly located in north-west part of the study area.

The quality of the thematic maps produced using SVMs classifications with different kernels are analyzed statistically by overall and individual accuracies. Tables 1, 2 and 3 summarize the classification accuracy assessment results. They showed that the SVMs classification with RBF kernel marginally outperformed the SVMs classification with the other kernels. For the SVMs classifier with RBF kernel, accuracy assessment on the validation dataset shows an overall accuracy of 96,05 % and the producer accuracy for all classes is greater than 68% (Table 2). All producers accuracies of classes are relatively low for soil-1 class.

The reached results indicate that irrespective of the used SVMs Kernel, the most accurately classified land covers types are vegetation, and soil2, due to their distinct spectral characteristics when compared with other classes. On the other hand, the classes of soil1 and water for the RBF kernel and the class-built area for the Polynomial and sigmoid kernel are assigned to the class greenhouses. This may be because of the transparent structure of the plastic leads to high spectral variability depending on the crop type grown in it and sometimes the roof is covered by Gypsum.

On the other hand, for Argan detection, using SVMs classification with RBF and sigmoid kernels provided quite accurate results. Comparing the producer's accuracies for this case, RBF kernel provided slightly higher value (91,28%) than sigmoid kernel (84,62%), while sigmoid kernel provided slightly higher value than polynomial kernel (77,44%).

In RBF kernel case, the producer's accuracy was computed as 91,28% for the class of Argan. Out of 195, 178 pixels were correctly classified as Argan, while 17 pixels were wrongly classified as soil2 (16 pixels) and greenhouse (1pixel).

The user's accuracy of Argan was computed as 99.44%, meaning that 178 pixels among 179 pixels that classified as Argan correspond to this class on the ground.

For the three kernels, the Argan class is confused with soil 2 due to the spectral similarity between them. The Argan tree in southwest Morocco is characterized by the fragmentation of its forest area. It is scattered, the effect of the underlying soil causes a confusion which does not make it possible to easily extract the Argan tree.

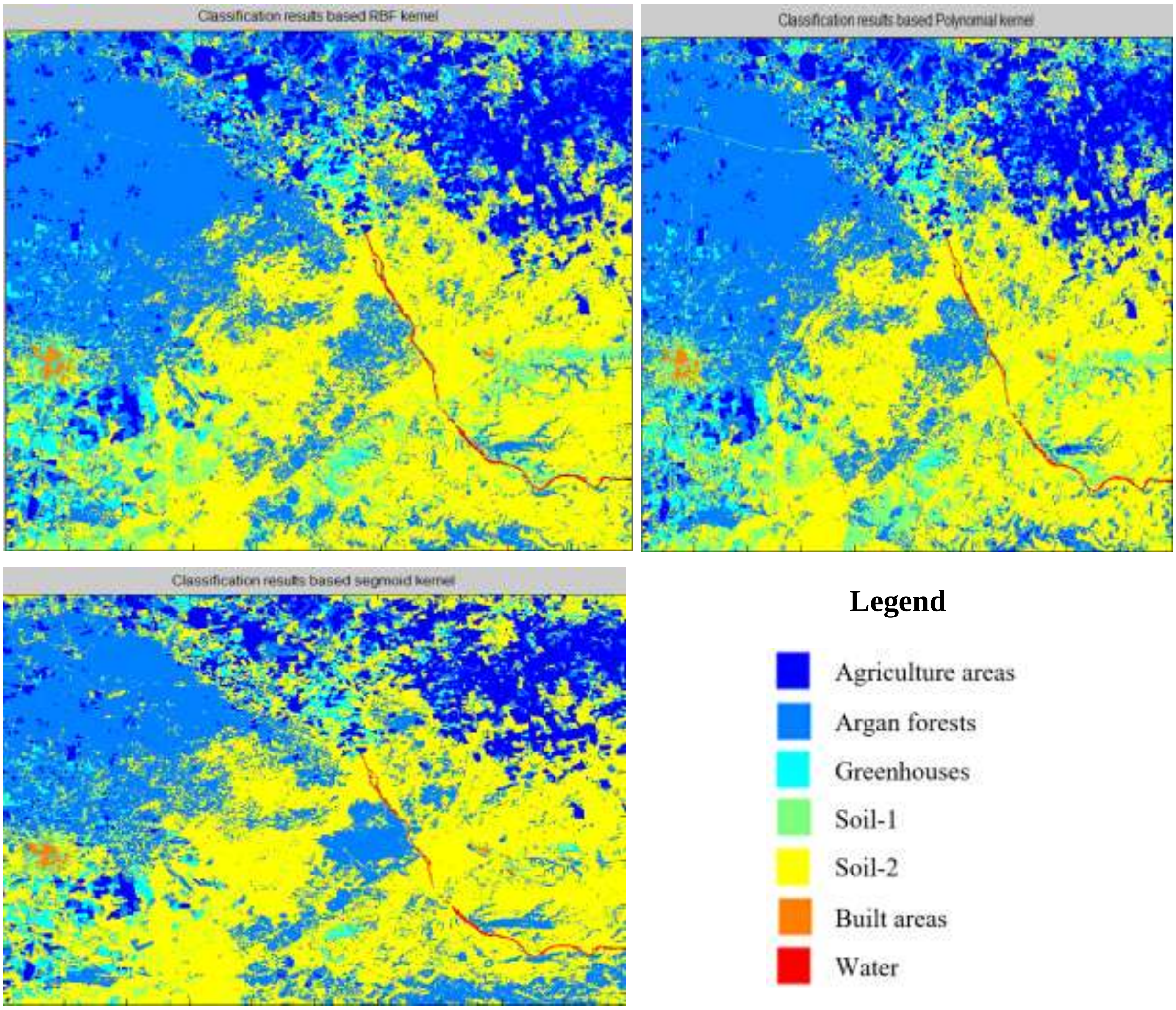



**Figure. 4.** Thematic maps derived from the classification of Sentinel-2 image using SVMs with RBF, Polynomial and sigmoid kernel.

*Journal of Engineering Technology* (ISSN. 0747-9964)
Volume 8, Issue 1, Jan. 2019, PP.1-12
RESEARCH ARTICLE -ENGINEERNG TECHNOLOGY ASEE

**Table 1.**Confusion matrix for SVM classifier with RBF kernel

| | | Ground truth (%) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Agriculture | Argan forest | greenhouses | Soil-1 | Soil-2 | Built area | Water |
| Classes | Agriculture | 100,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| | Argan forest | 0,00 | 91,28 | 0,00 | 0,85 | 0,00 | 0,00 | 0,00 |
| | greenhouses | 0,00 | 0,51 | 94,64 | 18,64 | 0,00 | 6,52 | 11,64 |
| | Soil-1 | 0,00 | 0,00 | 0,00 | 68,64 | 0,00 | 0,00 | 0,00 |
| | Soil-2 | 0,00 | 8,21 | 0,00 | 10,17 | 100,00 | 0,00 | 0,86 |
| | Built area | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 89,13 | 0,00 |
| | Water | 0,00 | 0,00 | 5,36 | 1,69 | 0,00 | 4,35 | 87,50 |
| | Total | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 |
| | Overall accuracy (%) | 96,05 | | | | | | |

**Table 2.** Confusion matrix for the SVM classifier with Polynomial kernel

| | | Ground truth (%) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Agriculture | Argan forest | greenhouses | Soil-1 | Soil-2 | Built area | Water |
| Classes | Agriculture | 100,00 | 11,79 | 0,13 | 0,85 | 0,00 | 0,00 | 0,00 |
| | Argan forest | 0,00 | 77,44 | 0,00 | 4,24 | 0,00 | 0,00 | 0,00 |
| | greenhouses | 0,00 | 0,00 | 91,69 | 0,00 | 0,00 | 47,83 | 9,91 |
| | Soil1 | 0,00 | 0,00 | 0,00 | 83,05 | 0,00 | 0,00 | 0,00 |
| | Soil2 | 0,00 | 10,77 | 0,00 | 8,47 | 100,00 | 0,00 | 1,29 |
| | Built area | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 43,48 | 0,00 |
| | Water | 0,00 | 0,00 | 8,18 | 3,39 | 0,00 | 8,70 | 88,79 |
| | Total | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 |
| | Overall accuracy (%) | 94,51 | | | | | | |

**Table3.**Confusion matrix for SVM classifier with sigmoid kernel

| | | Ground truth (%) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | Agriculture | Argan forest | greenhouses | Soil-1 | Soil-2 | Built area | Water |
| Classes | Agriculture | 100,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 | 0,00 |
| | Argan forest | 0,00 | 84,62 | 0,00 | 0,85 | 0,00 | 0,00 | 0,00 |
| | greenhouses | 0,00 | 0,00 | 84,32 | 0,85 | 0,00 | 69,57 | 1,29 |
| | Soil-1 | 0,00 | 0,00 | 0,00 | 83,05 | 0,00 | 0,00 | 0,00 |
| | Soil-2 | 0,00 | 15,38 | 0,00 | 11,02 | 100,00 | 0,00 | 0,86 |
| | Built area | 0,00 | 0,00 | 0,00 | 1,69 | 0,00 | 15,22 | 0,00 |
| | Water | 0,00 | 0,00 | 15,68 | 2,54 | 0,00 | 15,22 | 97,84 |
| | Total | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 | 100,00 |
| | Overall accuracy (%) | 93,50 | | | | | | |

## 5. Conclusion

The performance of the following three kernels were compared in this study: RBF, polynomial and sigmoid from the SVM classification for land cover, with emphasis on Argan forest detection. The study was implemented in an area selected from Souss-Massa in Morocco, which included Argan forest using Sentinel-2 image. Results from this study revealed that kernel type and kernel parameter influence the performance of the SVM.

Using any of these kernels, most of the land classes can be detected adequately from Sentinel-2 image. However, the superiority of SVMs with RBF kernel (C=2^7, γ=2^7) compared to polynomial and sigmoid kernel was confirmed in this study in terms of overall accuracy (96,05%). Additionally, the RBF kernel has less hyperparameters than the other kernel.

Considering the detection of Argan forests, it can be stated that the RBF kernel usually outperforms the polynomial kernel and sigmoid kernel in term of accuracy (91,28%).

It can be concluded that error penalty and kernel type have critical importance on the sensitivity of SVM architecture. Optimum parameters vary from method to another method. Therefore, it is recommended that the optimum parameters for SVM models should be analyzed in detail before the selection of the final models for the best classification results is made.

In future work, we will apply the best models to other images, such as Sentinel-3 and other study areas.

## References

[1]. El Kharki, O., Mechbouh, J., Ducrot, D., Rouchdi, M., El Aboudi, A., Ngno, J. M. (2014). The effects of the scale level and of the spatial resolution on the object-oriented classification: Application to mapping Arganeraie (region Agadir, Morocco). International Journal of Science and Research, Vol 3 Issue 12.

[2]. Kavzoglu, T., Reis, S. (2008). Performance analysis of maximum likelihood and artificial neural network classifiers for training sets with mixed pixels. GIScience and Remote Sensing 45, 330–342.

[3]. El Kharki, O., Mechbouh, J., Ducrot, D., Rouchdi, M., NGNO, J. M. (2015). Panorama sur les méthodes de classification des images satellites et techniques d'amélioration de la précision de la classification. Revue Française de Photogrammétrie et Télédétection, N0 210.

[4]. Bishop, C.M. (2006). Pattern recognition and machine learning, first edition Springer.

[5]. Vapnik, V.N. (1995). The Nature of Statistical Learning Theory. Springer-Verlag, New York.

[6]. Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification'. International Journal of Remote Sensing, 23(4), 725-749.

[7]. Melgani, F., &Bruzzone, L. (2004). "Classification of hyperspectral remote sensing images with support vector machines". IEEE Transactions on Geoscience &Remote Sensing, 42(8), 1778-1790.

[8]. Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. International Journal of Remote Sensing, 26(5), 1007-1011.

[9]. Mountrakis, G., Im, J., Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66 (3), 247-259.

[10]. Ben-Hur, A., Weston, J. (2010). A user's guide to support vector machines. In: Carugo, O., Eisenhaber, F. (Eds.), Data Mining Techniques for the Life Sciences. Humana Press, 223–239.

[11].Ying Tian, QingE Wu &Pengfei Yue (2018). A comparison study of classification algorithms on dataset using WEKA tool. Journal of Engineering Technology, Vol 6, Issue 2, PP. 329-341.

[12]. Mathur, A., Foody, G.M. (2008). Multiclass and binary SVM classification: Implications for training and classification users. IEEE Geoscience and Remote Sensing Letters 5, 241–245.

[13]. Hsu, C., Chang, C., and Lin, C. (2007). "A Practical Guide to Support Vector Classification". Dept of Computer Science. National Taiwan Uni, Taipei, 106, Taiwan http://www.csie.ntu.edu.tw/~cjlin

[14]. Hsu, C. and Lin, C. (2002). A comparison of methods for multi-class support vector machines. IEEE Transactions on Neural Networks, 13(2), 415-425.

[15]. Chang, C. C., & Lin, C. J. (2016). LIBSVM: a library for support vector machines. Software available at https://www.csie.ntu.edu.tw/~cjlin/libsvm/ .