Advanced Remote Sensing

http://publish.mersin.edu.tr/index.php/arsej

e-ISSN 2979-9104

Crop classification from multi-temporal PolSAR data with regularized greedy forest

Mustafa Ustuner *1, Fusun Balik Sanli 2

[1]*Artvin Coruh University, Department of Geomatics Engineering, Artvin, Türkiye, mustuner@artvin.edu.tr*
[2]*Yildiz Technical University, Department of Geomatics Engineering, Istanbul, Türkiye, fbalik@yildiz.edu.tr*

Cite this study: Ustuner, M., & Sanli, F. B. (2021). Crop classification from multi-temporal PolSAR data with regularized greedy forest. Advanced Remote Sensing, 1(1), 10-15

Keywords
Remote Sensing
Polarimetric SAR
Ensemble Learning
Random Forest
Agriculture

Research Article
Received: 12.11.2021
Revised: 10.12.2021
Accepted: 17.12.2021
Published:30.12.2021

Abstract

Polarimetric Synthetic Aperture Radar (PolSAR) images are considered as an important data source for the crop mapping and monitoring especially for the time-critical agricultural applications. The objective of this paper is to evaluate the potential of a novel ensemble learning algorithm, Regularized Greedy Forest (RGF), for crop classification from multi-temporal quad-pol PolSAR data. For the classification of crops (maize, potato, wheat, sunflower, and alfalfa) in the study site, the polarimetric features of Cloude–Pottier decomposition (a.k.a H/A/α decomposition) were used as the input data. The performance of RGF was compared to Random Forests (RF) and Support Vector Machines (SVM) in terms of overall accuracy and Kappa values. Our experimental results demonstrated that RGF can yield higher accuracy (with an overall accuracy of 0.78) than RF and SVM for crop classification using PolSAR images. Moreover, it can be concluded that polarimetric features of Cloude–Pottier decomposition are of efficient for the discrimination of crops using multi-temporal PolSAR data.

1. Introduction

Crop classification is one of the vital and important applications in remote sensing since this information could be used as an input data for crop yield estimation, agricultural planning as well as spatio-temporal monitoring of crops. Spaceborne SAR images are usually preferred for the time-critical agricultural applications because SAR signals are sensitive to the crop structure and dielectric properties. In particular, PolSAR images provide more detailed information for agronomic characteristics as they record the complete characteristics of the scattering in each polarization for the natural targets [1-4].

Polarimetric target decompositions (or target decompositions) are used for easier understanding and simpler interpretation of the complex scattering characteristics of natural and man-made targets [3, 5-7]. In this experimental research, we implemented the Cloude–Pottier decomposition (a.k.a. H/A/α decomposition) that is a type of eigenvector-based decomposition.

Over the last two decades, a wealth of ensemble learning algorithms has been utilized in remote sensing such as random forests [8], extremely randomized trees [9-10] (a.k.a. extra trees), canonical correlation forest [11-12], extreme gradient boosting (XgBoost) [13], Light Gradient Boosting Machine (LightGBM) [14] and deep forest [15]. We chose Regularized Greedy Forest (RGF) in this experimental research since the regularized greedy forest algorithm has not been fully explored yet for the crop classification using multi-temporal PolSAR data. Furthermore, we compared the classification performance of RGF with the two popular and well-established machine learning algorithms in remote sensing, namely RF and SVM.

In this paper, we consider the following questions. (1) Can RGF yield higher accuracy than RF and SVM for PolSAR image classification in our experimental research? (2) Are polarimetric features of Cloude–Pottier decomposition sufficient for crop discrimination from multi-temporal PolSAR data? The major contributions of our experimental study can be shortly summarized as follows.

(1) We investigated the regularized greedy forest algorithm for the crop classification using the polarimetric features from Cloude–Pottier decomposition.

(2) The performance of RGF in comparison to RF and SVM was evaluated for crop classification from multi-temporal PolSAR images.

The rest of the paper is organized as follows. Section 2 introduces the study area and data. PolSAR data processing and classification models are summarized in Section 3. The details of the experimental results and their discussion are presented in Section 4. And following in Section 5, the final conclusions and some important remarks are provided.

2. Study area and data

In this section, the study area site and the details of the multi-temporal PolSAR dataset will be presented.

2.1. Study area

The study site corresponds to the agricultural fields close the province of Konya, Turkey, illustrated in Figure 1. The region has a flat topography and favorable climate conditions for precision farming. The main crop types covering the study site are alfalfa, maize, potato, summer wheat and sunflower. In-situ data was collected simultaneously at the acquisition dates of the PolSAR images.

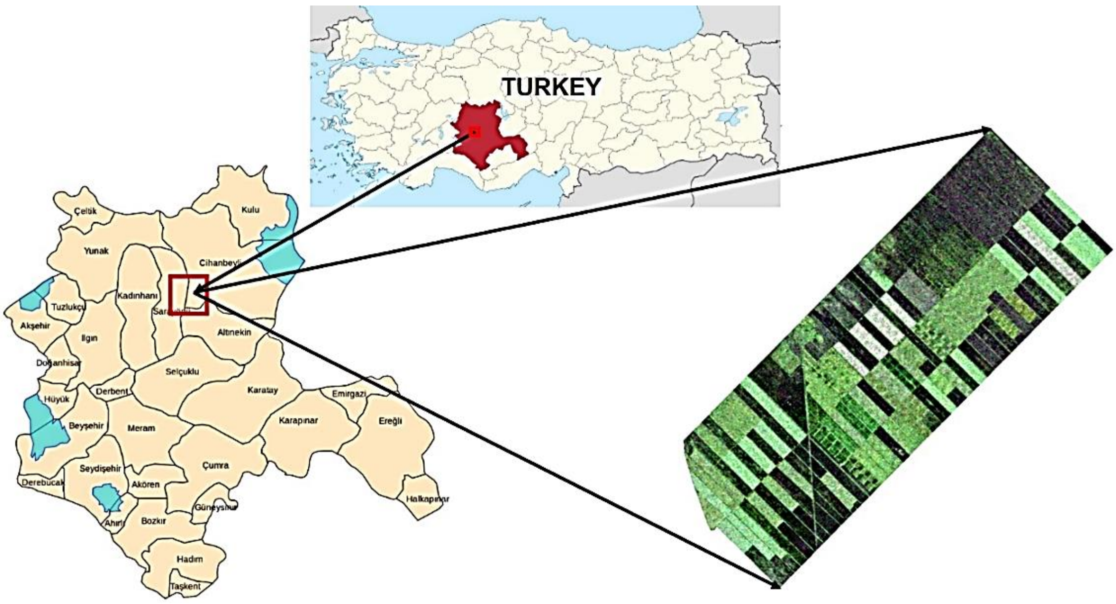

Figure 1. Study area [14]

2.2. Multi-temporal PolSAR dataset

The Multi-temporal quad-polarimetric RADARSAT-2 data (single look complex with fine quad-polarization acquisition mode) was used in our experimental study. The data was acquired for the key dates of the crops as follows: June 13, July 7, July 31 and August 24 of 2016. The data specifications are presented in Table 1.

Table 1. PolSAR Data Specifications

Specifications	Description
Wavelength	C band - 5.6 cm
Resolution (in m)	4.7 x 5.1 (rg x az)
Incidence angle	400
Pass direction	Descending
Acquisition type	Fine quad pol
Polarization	Quad polarimetric
Product type	Single look complex

3. Methods

3.1. PolSAR data processing

Some pre-processing steps for PolSAR data processing are required in order to extract the relevant and proper polarimetric features from target decompositions. In our experimental study, the data pre-processing includes the following steps: (1) data calibration; (2) matrix generation (from coherency matrix); (3) extraction of polarimetric features from Cloude-Pottier decomposition; and (4) orthorectification. All pre-processing steps were implemented using open-source SNAP (The Sentinel Application Platform) v6.0 toolbox, provided by European Space Agency. The data processing steps were illustrated in Figure 2.

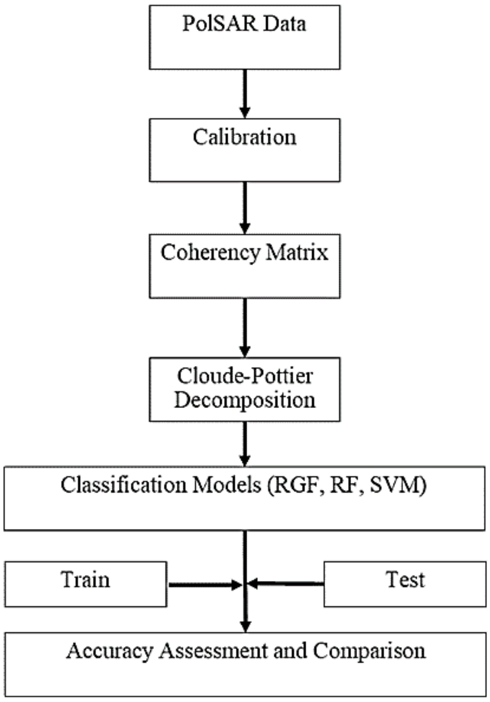

Figure 2. PolSAR Data Processing and Classification

The Cloude-Pottier decomposition (also known as H/A/α decomposition) is an eigenvector-based decomposition of coherency matrix and separates the total scattering mechanism into three polarimetric features (averaged parameters) which are entropy (H), anisotropy (A) and alpha angle (α). Each feature type provides different information regarding the scattering mechanism such as entropy measures the randomness level of scattering and the alpha angle defines the scattering type of targets (i.e., surface, double-bounce and volume scattering). And the latter parameter, anisotropy, is helpful to demonstrate the differences between the scattering mechanisms [4-5,14,16].

3.2. Image classification

In this experimental study, three different machine learning algorithms were implemented for the crop classification from multi-temporal C-band PolSAR images. The details for training and testing samples were provided in Table 2. The brief summary of the classification models is provided in the following paragraphs.

Table 2. Ground Truth Information

Class	Training	Testing	Total
Alfalfa	1918	3542	5460
Maize	5581	14217	19798
Potato	2275	10604	12879
Sunflower	3729	8915	12644
Wheat	3524	6338	9862

Regularized Greedy Forest (RGF) is a type of tree-based ensemble learning algorithm, developed by [17]. RGF builds decision forests via fully-corrective regularized greedy search by using the underlying forest structure. Fully-corrective regularized greedy search algorithm recursively re-optimizes the coefficients of all decision rules. The novelty of this method is that it combines two ideas: (1) tree-structured regularization into the learning formulation and (2) fully-corrective regularized greedy algorithm. The classification was performed by using the python wrapper of RGF [18].

The Random Forest is one the most frequently used ensemble learning algorithms in remote sensing image classification. RF creates a set of decision trees to make a prediction and the final output of the classifier is determined by the majority voting of the trees [19]. Support Vector Machines are one of the popular kernel-based learning algorithms and based on statistical learning theory. SVM use the kernels to map the data into higher dimensional space for the linear separation of classes. Radial Basis Function kernel was used in our experiment [20-21]. SVM and RF classifications were performed using the open-source Scikit-learn (v 0.19) module in Python v3.6.4 [22].

4. Experimental results and discussion

The classification performance of Regularized Greedy Forest in comparison to RF and SVM was analyzed in our experimental study for the classification of crops from multi-temporal PolSAR data. The overall accuracy of the classified images was derived from the error matrix and the comparison of the methods was assessed in terms of overall accuracy and kappa coefficients. Table 3 presents the overall accuracies and kappa coefficients for the classification algorithms. The highest classification accuracy (overall accuracy of 78.65% and kappa coefficient of 0.72) was produced with RGF while lowest classification accuracy (overall accuracy of 75.08% and kappa coefficient of 0.67) was obtained by SVM. The classified images for each method were presented in Figure 3.

Table 3. Classification Accuracies

Methods	Overall Accuracy (%)	Kappa
SVM	75.08	0.67
RF	76.52	0.69
RGF	78.65	0.72

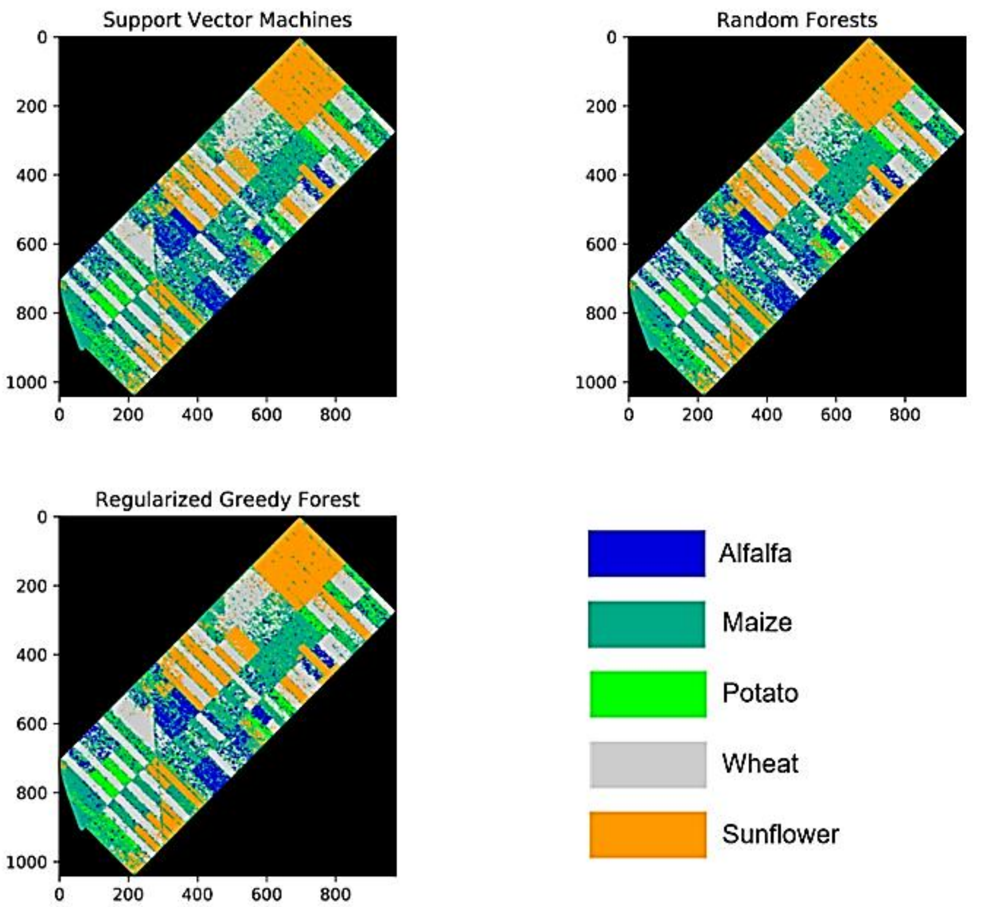

Figure 3. Classified Images

Individual class accuracies were assessed based on F1-score which is the harmonic mean of the user accuracy and producer accuracy values. Table 4 presents the individual class accuracies (based on F1-score) for each classification method.

Table 4. Individual class accuracies (F1-score)

Methods	RGF	RF	SVM
Alfalfa	0.26	0.26	0.28
Maize	0.76	0.74	0.72
Potato	0.70	0.63	0.66
Sunflower	0.99	0.99	0.97
Wheat	0.95	0.95	0.93

Sunflower is the most accurate classified class (0.99) while alfalfa is the least accurate classified class (0.26) in which F1-score could not reach up to the 0.30. SVM predicted the alfalfa class more accurate than other methods though it yielded the lowest classification accuracy. Moreover, SVM obtained higher F1-score than RF for potato and alfalfa classes. Wheat is the second most accurate predicted class in our experimental study where it was predicted above the F-1 score values of 0.90.

5. Conclusion

This research investigated the performance of a novel ensemble learning algorithm, RGF, in comparison to RF and SVM for the crop classification from multi-temporal PolSAR images. The Cloude-Pottier decomposition was implemented for the extraction of the polarimetric features. Our experimental results demonstrated the following conclusions: 1) RGF can yield higher classification accuracy than RF and SVM for the classification of multi-temporal PolSAR images 2) the polarimetric parameters derived from Cloude-Pottier decomposition are suitable for the crop classification. Our future research will focus on the extensive analysis of the polarimetric features derived from incoherent polarimetric decompositions for the crop classification.

Acknowledgement

The authors gratefully acknowledge the data support from TAGEM (Project No: TAGEM/TSKAD/14/A13/P05/03). Mustafa Ustuner received the research fellowship in 2018 by TUBITAK 2214/A International Doctoral Research Fellowship Programme (Grant No #1059B141700579) and would like to thank TUBITAK as well as the Department for Earth Observation of Friedrich-Schiller University of Jena, Germany for hosting him in their department as a guest researcher in 2018.

Author contributions

Mustafa Ustuner: Conceptualization, Methodology, Data curation, Writing-Original draft preparation, Software and Validation. **Fusun Balik Sanli:** Visualization, Investigation, Writing-Reviewing, Editing and Validation

Conflicts of interest

The authors declare no conflicts of interest.

References

1. McNairn, H., & Brisco, B. (2004). The application of C-band polarimetric SAR for agriculture: A review. Canadian Journal of Remote Sensing, 30(3), 525-542.
2. Soria-Ruiz, J., Fernandez-Ordonez, Y., McNairm, H., & Bugden-Storie, J. (2007, July). Corn monitoring and crop yield using optical and RADARSAT-2 images. In 2007 IEEE International Geoscience and Remote Sensing Symposium (pp. 3655-3658). IEEE.
3. Jiao, X., Kovacs, J. M., Shang, J., McNairn, H., Walters, D., Ma, B., & Geng, X. (2014). Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data. ISPRS Journal of Photogrammetry and Remote Sensing, 96, 38-46.

4. Tamiminia, H., Homayouni, S., McNairn, H., & Safari, A. (2017). A particle swarm optimized kernel-based clustering method for crop mapping from multi-temporal polarimetric L-band SAR observations. International Journal of Applied Earth Observation and Geoinformation, 58, 201-212.
5. Cloude, S. R., & Pottier, E. (1997). An entropy based classification scheme for land applications of polarimetric SAR. IEEE transactions on geoscience and remote sensing, 35(1), 68-78.
6. Chen, S. W., Li, Y. Z., Wang, X. S., Xiao, S. P., & Sato, M. (2014). Modeling and interpretation of scattering mechanisms in polarimetric synthetic aperture radar: Advances and perspectives. IEEE Signal Processing Magazine, 31(4), 79-89.
7. Steele-Dunne, S. C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P. W., & Papathanassiou, K. (2017). Radar remote sensing of agricultural canopies: A review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(5), 2249-2273.
8. Pal, M. (2005). Random forest classifier for remote sensing classification. International Journal of Remote Sensing, 26(1), 217-222.
9. Barrett, B., Nitze, I., Green, S., & Cawkwell, F. (2014). Assessment of multi-temporal, multi-sensor radar and ancillary spatial data for grasslands monitoring in Ireland using machine learning approaches. Remote Sensing of Environment, 152, 109-124.
10. Samat, A., Persello, C., Liu, S., Li, E., Miao, Z., & Abuduwaili, J. (2018). Classification of VHR multispectral images using extratrees and maximally stable extremal region-guided morphological profile. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(9), 3179-3195.
11. Colkesen, I., & Kavzoglu, T. (2017). Ensemble-based canonical correlation forest (CCF) for land use and land cover classification using sentinel-2 and Landsat OLI imagery. Remote Sensing Letters, 8(11), 1082-1091.
12. Xia, J., Yokoya, N., & Iwasaki, A. (2016). Hyperspectral image classification with canonical correlation forests. IEEE Transactions on Geoscience and Remote Sensing, 55(1), 421-431.
13. Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., & Wolff, E. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. IEEE geoscience and remote sensing letters, 15(4), 607-611.
14. Ustuner, M., & Balik Sanli, F. (2019). Polarimetric Target Decompositions and Light Gradient Boosting Machine for Crop Classification: A Comparative Evaluation. ISPRS International Journal of Geo-Information, 8(2), 97. MDPI AG.
15. Yang, F., Xu, Q., Li, B., & Ji, Y. (2018). Ship detection from thermal remote sensing imagery through region-based deep forest. IEEE Geoscience and Remote Sensing Letters, 15(3), 449-453.
16. Cloude, S. R., & Pottier, E. (1996). A review of target decomposition theorems in radar polarimetry. IEEE Transactions on Geoscience and Remote Sensing, 34(2), 498-518.
17. Johnson, R., & Zhang, T. (2013). Learning nonlinear functions using regularized greedy forest. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(5), 942-954.
18. RGF Team, "Regularized Greedy Forest (RGF) library" [Online] (2018). Available: https://github.com/RGF-team/rgf, 2018. (Accessed on: December. 30, 2019)
19. Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. Pattern Recognition Letters, 27(4), 294-300.
20. Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 42(8), 1778-1790.
21. Zeybek, M. (2021). Inlier point preservation in outlier points removed from the ALS point cloud. Journal of the Indian Society of Remote Sensing, 49(10), 2347-2363.
22. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.