ISSN: 2148-9173 Vol: 6 Issue:1 April 2019

International Journal of Environment and Geoinformatics (IJEGEO) is an international, multidisciplinary, peer reviewed, open access journal.

Optimization of Remote Sensing Image Attributes to Improve Classification Accuracy

Dilek KÜÇÜK MATÇI & Uğur AVDAN

Abstracting and Indexing: DOAJ, Index Copernicus, OAJI, Scientific Indexing Services, JF, Google Scholar

IJEGEO International Journal of Environment and Geoinformatics 6(1): 50-56 (2019)

Reaserch Article

Optimization of Remote Sensing Image Attributes to Improve Classification Accuracy

Dilek Küçük Matçı* **Uğur Avdan**

Research Institute of Space and Earth Sciences, Anadolu University, Eskişehir, TR

Corresponding author: D. Küçük Matçı
E-mail:dkmatci@anadolu.edu.tr

Received 03 Oct 2018
Accepted 01 March 2019

How to cite: Küçük Matçı & Avdan (2019). Optimization of Remote Sensing Image Attributes to Improve Classification Accuracy, *International Journal of Environment and Geoinformatics (IJEGEO)*, 6(1): 50-56. DOI: 10.30897/ijegeo.466985

Abstract

Remote sensing technologies provide very important big data to various science areas such as risk identification, damage detection and prevention studies. However, the classification processes used to create thematic maps to interpret this data can be ineffective due to the wide range of properties that these images provide. At this point, there arises a requirement to optimize the data. The first objective of this study is to evaluate the performance of the Bat Search Algorithm which has not previously been used for improving the classification accuracy of remotely sensed images by optimizing attributes. The second objective is to compare the performance of the Genetic Algorithm, Bat Search Algorithm, Cuckoo Search Algorithm and Particle Swarm Optimization Algorithm, which are used in many areas of the literature for the optimization of the attributes of remotely sensed images. For these purposes, an image from the Landsat 8 satellite is used. The performance of the algorithms is compared by classifying the image using the K-Means method. The analysis shows a 10-22% increase in overall accuracy with the addition of attribute optimization.

Keywords: Remote Sensing, Classification, Optimization, Unsupervised Classification

Introduction

Rapid development of technology allows remote sensing systems to provide increasingly detailed data, with hundreds of properties, to areas such as agriculture, the military, meteorology, etc. The use of this detailed data requires processing time and expertise, and in some cases there is an adverse effect of higher accuracy rates. It is especially difficult for scientists without image processing expertise to use this data effectively. In these cases, it is extremely important to optimize features by determining which features are important for classification. Land use mapping is one of the major applications in remote sensing. How to accurately acquire the land use information by classifying remotely sensed data have been extensively studied in the past decades of time (Bo, 2008; Gazioğlu, 2018). The concept of optimization includes the whole set of methods used to achieve specific goals with the resources at hand, in the best possible way (Lindley 1991). Optimization methods are used in many different areas including agriculture (Blumenstein, 2018; Siegmeier et al. 2018, Huang & Song 2018; Li 2018), production (Saraç & Özdemir 2003), health (Simav, et al, 2015; Acar & Butt 2016; İncekara et al., 2017), scheduling (Sagir & Ozturk 2010) and image processing (Zhang & Izquierdo 2006). Image classification is one of the foundation applications of remotely sensed images in relation to the analysis of landcover (McCloy & Bocher, 2007; Esetlili, et al., 2018). One method that provides quick solutions to complex problems is the Bat Search Algorithm (BSA). This algorithm is inspired by the sonar location systems that bats use to determine their location (Yang 2010). The BSA has been used in areas such as data mining (Khan, Nikov et al. 2011), image processing (Akhtar, Ahmad et al. 2012) and fuzzy logic (Khan, Nikov et al. 2011, Lemma & Hashim 2011). The rules of the algorithm are based on the way bats find their prey by echolocation (Yang 2010).

The Cuckoo Search Algorithm (CSA), also used in this study, was developed by Yang and Deb (2009), and has been used in many areas including production optimization (Walton, Hassan et al. 2011), software testing and test data generation (Perumal, Ungati et al. 2011) and nurse assignment problems (Tein & Ramli 2010). CSA is based on the reproductive strategies of cuckoos (Gandomi, 2013; Yang et al. 2009-2010).

Particle Swarm Optimization (PSO) is another method used for optimization in this study. It was developed by Kennedy and Eberhart (Eberhart and Shi 2000, Poli, Kennedy et al. 2007, Kennedy 2011). PSO is used in areas such as feature selection (Huang & Dun 2008), scheduling (Allahverdi & Al-Anzi 2006), etc. This algorithm starts with a population of random solutions called particles. Each particle is spread in an n-dimensional space. The velocity is updated by evaluating the best position value of each particle and the best particle value so far. Then the new positions of the particles are created. This iteration continues until the stop condition is satisfied (Eberhart & Shi 1998; Shi & Eberhart 1998). The Genetic Algorithm (GA), used for feature optimization in this study, was developed by John Holland (Holland 1975) inspired by evolution in the form of chromosomes and gene transfer simulations. This algorithm usually starts with the production of a chromosome population at random. These algorithms use crossover and mutation operators to arrive at solutions. In GA, a process of obtaining the best solutions and transferring them to the next generation is performed (Pal and Wang 2017).

This study investigates the performance of these optimization methods in the featured optimization problem, including their accuracy and speed of problem solving. To test the performance of these algorithms in improving classification accuracy, a Landsat 8 satellite image is classified by the K-Means method. The contribution of this study is optimizing the properties of a remotely sensed image using the optimization methods frequently used in various areas of the literature. We propose that multiple satellite image features can be optimized by evaluating the performance of these methods.

Study Area and Materials

In this study, feature optimization is applied to a multi-band satellite image. The image covers an area affected by a forest fire that took place in Adrasan and Kumluca regions in Antalya province on June 2016 (Figure 1). Dense forest fires are encountered in the Mediterranean region, which is at the 1st degree risk of forest fire, especially in the summer months. In the Kumluca fire area, 17 houses were burned, about 200 decares of greenhouse and 300 decares of garden were damaged (pomegranate, olive, citrus) and 60 large and small ruminants were destroyed. Fire, irrigation, water pumps, water pumps and irrigation systems have been destroyed. Adrasan has been destroyed by about 100 large and small ruminants in the fire area(Derneği 2016). Detection of burned forest areas is important both in determining the damage and planning the renovation work to be done.

The use of satellite images for this purpose provides economical and easily updatable data. For this reason, free images obtained from the Landsat 8 satellite are used. The bands used in the study are multispectral bands with 30m resolution, blue, green, red, near infrared, shortwave infrared 1 and shortwave infrared 2, and a panchromatic band with 15m resolution. The panchromatic and multispectral bands are fused using the PHANSHARP2 algorithm to improve the spatial resolution of the multispectral bands from 30 to 15m.

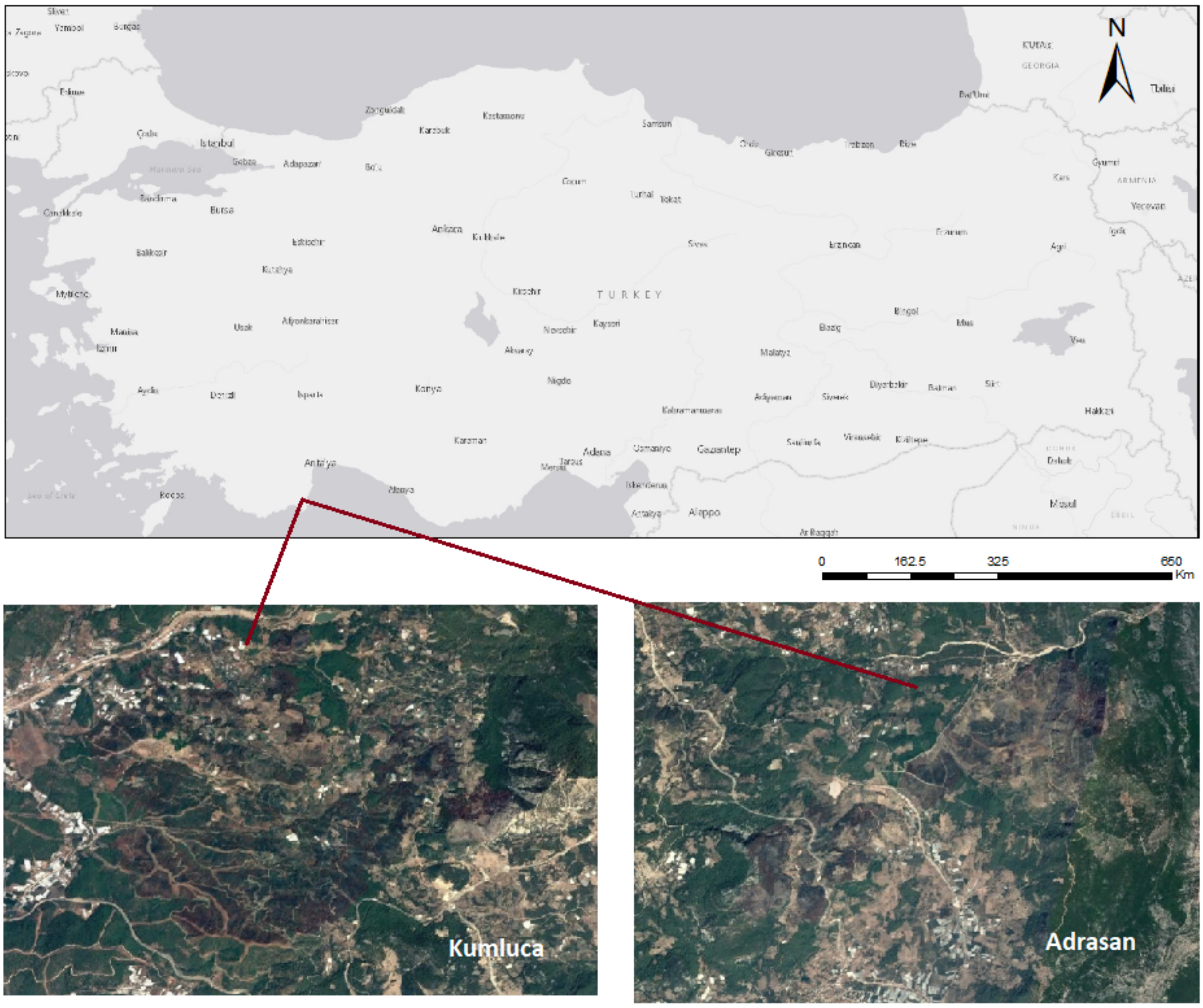


Figure 1. The study area

Methods

The process steps applied in the study are shown in Figure 2. Pansharpening is performed first to obtain a higher resolution. Then the image is segmented by the multi-resolution segmentation method (Baatz 2000).

Spectral plant indexes and burned area indexes were calculated, because the forest fires directly affects the reflectance values of plants. These calculated indexes were added to the image. The calculated spectral indices are Normalized Difference Vegetation Index (NDVI), Normalized Burn Ratio (NBR), BAI, Normalized Burn Ratio 2 (NBR2), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Mid Infrared Burn Index (MIRBI) and Normalized Difference Moisture Index (NDMI).

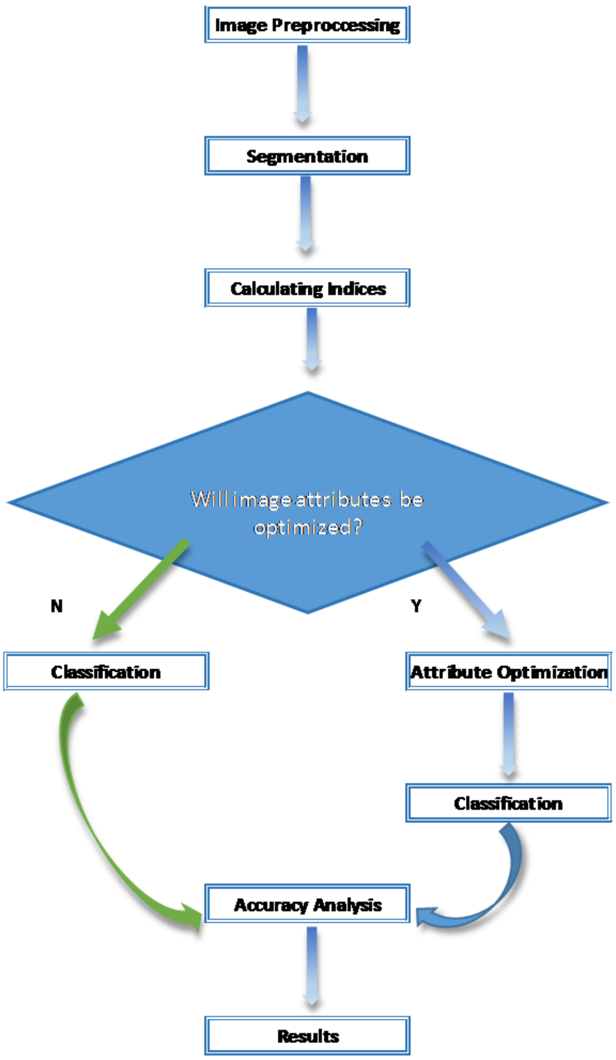


Figure 2. The workflow followed in the study.

After all image preprocessing steps is completed, the image was classified without selecting the attributes by optimization methods. The classification process is carried out with the K-Means algorithm. This method classifies a given pixel set in the image over a certain number of sets (k sets). The process starts with identifying the k number of centers randomly. Each pixel in the image is assigned to the nearest center. When all the pixels in the image are classified, the new center of the generated groups is calculated. This process is repeated until the stage that none of the centers move 27. The accuracy analysis is based on the overall accuracy. In order to determine the accuracy of the classification, 1491 control points are randomly assigned to the area. The real values of the generated points and the result values obtained from the classification processes are compared, and an error matrix is created.

Objective function $f(\mathbf{x})$, $\quad \mathbf{x} = (x_1, ..., x_d)^T$
Initialize the bat population $\mathbf{x}_i$ $(i = 1, 2, ..., n)$ *and* $\mathbf{v}_i$
Define pulse frequency f_i *at* $\mathbf{x}_i$
Initialize pulse rates r_i *and the loudness* A_i
while *(t <Max number of iterations)*
Generate new solutions by adjusting frequency,
and updating velocities and locations/solutions
  if $(rand > r_i)$
  Select a solution among the best solutions
  Generate a local solution around the selected best solution
  end if
  Generate a new solution by flying randomly
  if $(rand < A_i \ \& \ f(\mathbf{x}_i) < f(\mathbf{x}_*))$
  Accept the new solutions
  Increase r_i *and reduce* A_i
  end if
Rank the bats and find the current best $\mathbf{x}_*$
end while
Postprocess results and visualization

Figure 3. The pseudo code of Bat Search Algorithm (Yang 2010)

Table 1. Selected attributes with CSA, PSO, GA and BSA

Classes	Forest		Water		Urban Area		Soil		Burned Area		OVERALL
Evaluation	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	
K-means	0.69	0.86	1	0.35	0.99	0.99	0.9	0.85	0	0	0.63

In the second phase of the study, the attributes of the objects obtained after the segmentation process are optimized using BSA. The BSA algorithm has been proposed, inspired by the direction and distance detection behavior of an object, utilizing the echoing of sound called echolocation. The pseudo code of BSA algorithm is given in Figure 3. After determining the features to be classified, classification and accuracy analysis was performed.

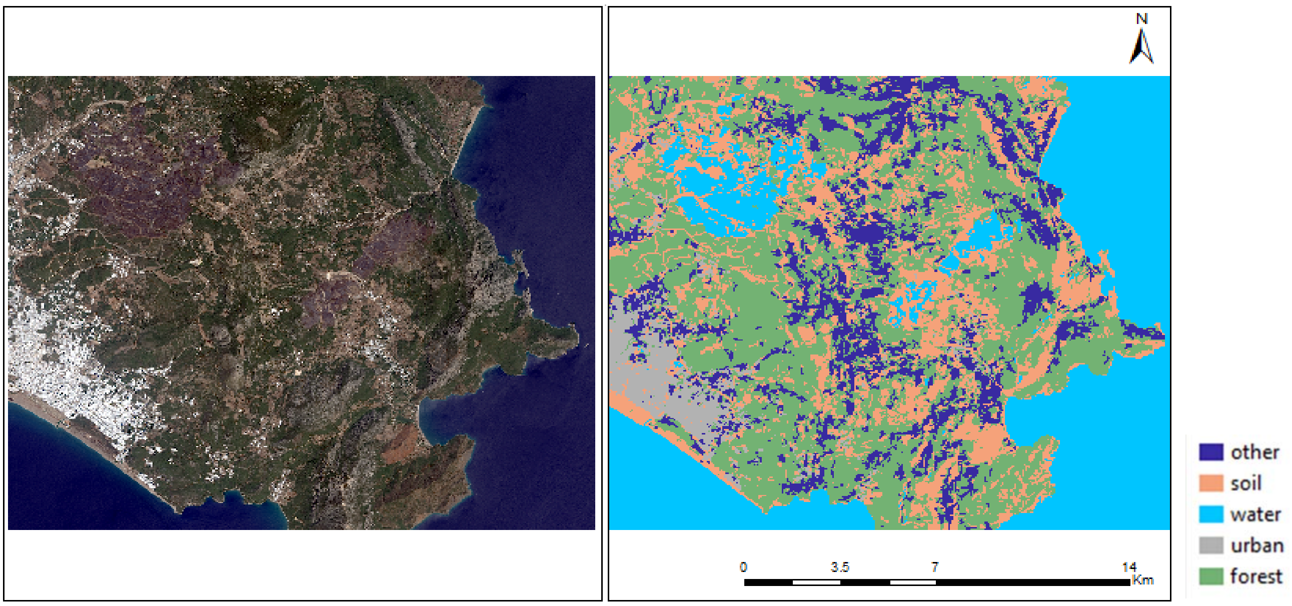
Results and Discussion

The study was carried out in an area including various land coverings such as barren lands, bare soil, open water areas, forests, greenhouses, urban areas and burnt areas. The classification process by using K-Means aims to obtain five classes, soil, water, forest, urban area and burnt area. The resulting map of the classification by the K-Means algorithm using all the properties obtained is presented in Figure 4. The results obtained as a result of the accuracy analysis are given in Table 1. According to these results, overall accuracy was obtained as 63%. The accuracy of the classification is evaluated by means of producer's accuracy (PA), user's accuracy(UA) and the overall accuracy parameters.

As a result of the qualitative analysis carried out in the result map, it was determined that the urban areas are classified with high accuracy. However, the fact that burned forest areas and wetlands are not separated is seen as the biggest reason for the low accuracy rate.

The attributes of the satellite image are optimized by using CSA, BSA, GA and PSO to see the effect of optimizing image attributes to the classification process. The selected attributes given as a result of the property optimization are shown in Table 2.

The satellite image is again classified by the K-Means method using the attributes obtained by the attributes optimization process. The result maps are shown in Figure 5.



a b

Figure 4. (a) Classification map of the K-means algorithm using all attributes; (b) Satellite image

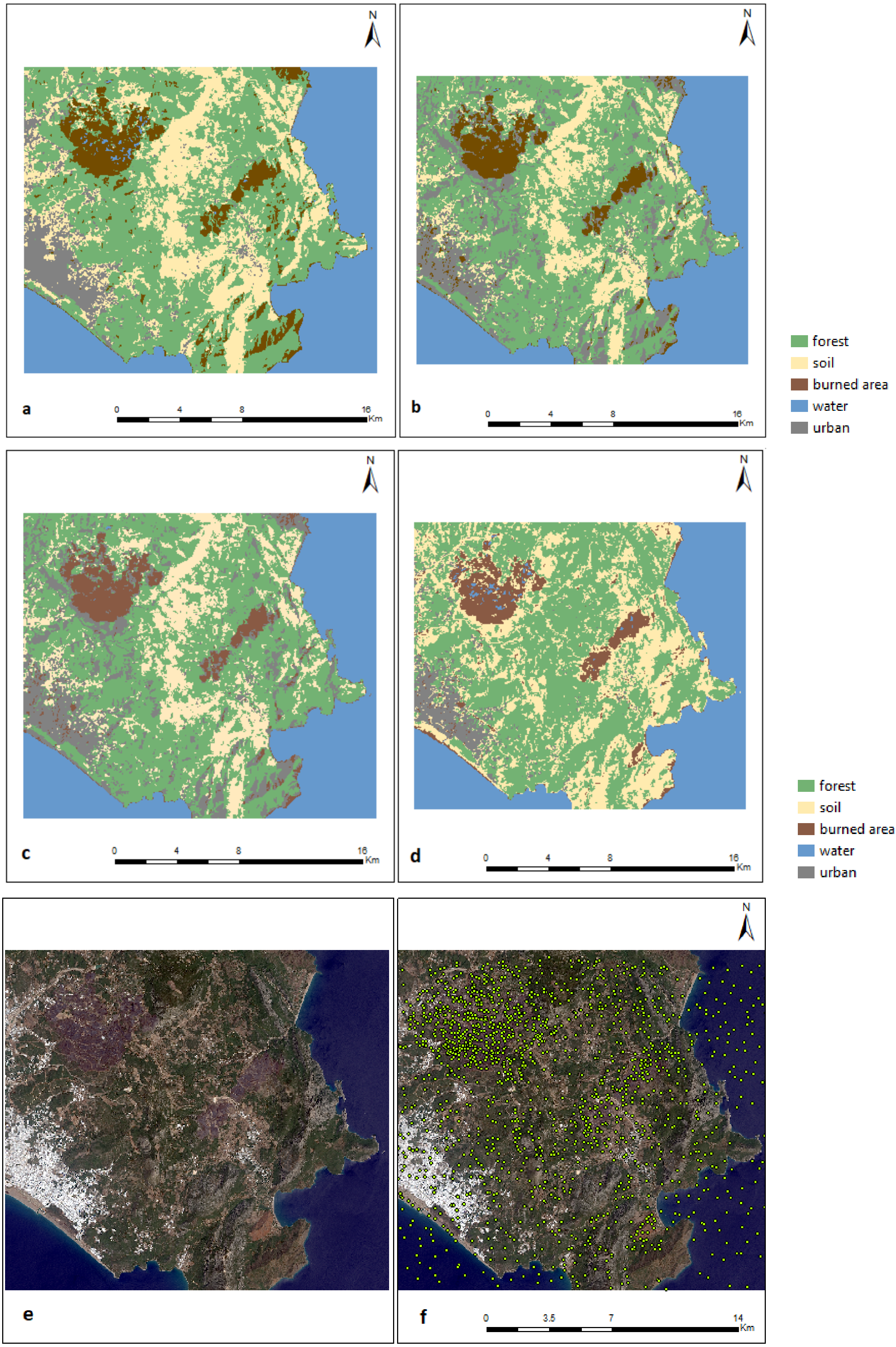


Fig 5. Classification results maps (a) CSA + K-Means; (b) BSA + K-Means; (c) GA + K-Means; (d) PSO + K-Means; (e) Satellite image with points; (f) Satellite image

Table 2. Selected attributes with CSA, PSO, GA and BSA

Features			
CSA	BSA	PSO	GA
BAI	EVI	Red Band Value	EVI
EVI	BAI	NBR2	MIRBI
	NDVI	SAVI	NDVI

Table 3 Accuracy analysis results.

Classes	Forest		Water		Urban Area		Soil		Burned Area		OVERALL
Algorithms	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	
BSA	0.82	0.62	1	0.98	0.99	0.78	0.49	0.75	0.93	0.96	0.78
CSA	0.86	0.63	1	0.85	0.99	0.79	0.48	0.81	0.88	0.94	0.77
PSO	0.78	0.69	1	0.79	0.92	0.97	0.59	0.68	0.82	0.91	0.77
GA	0.81	0.55	1	0.99	0.96	0.36	0.23	0.68	0.89	0.99	0.69
ALL	0.69	0.86	1	0.35	0.99	0.99	0.90	0.85	0	0	0.63

The accuracy analysis is undertaken by assigning 1491 random points to the area (Figure 3e). The results of the accuracy analysis are given in Table 2. The accuracy of the classification is evaluated by means of producer's accuracy (PA), user's accuracy (UA) and the overall accuracy parameters.

As a result of the accuracy analysis, the overall accuracy of the classification without property optimization is 63%. The overall accuracies of the classification by PSO, CSA, BSA and GA are 77%, 77%, 78% and 69%, respectively. Although the overall accuracy rate is low, the most successful result in the extraction of burnt areas was obtained with GA. According to these results, optimizing the features increased the overall accuracy rate by 10% to 24%. Burnt areas, which could not be separated as a result of classification performed without feature optimization, could be extracted with a success rate of up to 99%.

Conclusions

The most commonly used method of obtaining useful information about the earth via satellite images is classifying images. In satellite images, different objects have different spectral reflection values. This difference provides the advantage of distinguishing ground objects (Lillesand, et al. 2014).

This study explores how optimization methods are used to make the most of the available resources to classify satellite images with the best results. For this purpose, a Landsat 8 image containing the regions affected by forest fires occurring in 2016 is used. Then the image properties were optimized with different methods and the classification results were compared. Accuracy analysis shows that the optimization of the attributes increases the overall accuracy rate by 10-24%. One of the most interesting outcomes of this study is that the burnt areas cannot be distinguished by the classification process performed by using all the features in the image, whereas the burned areas can be extracted with a success rate of up to 99% as a result of classification by attribute optimization.

References

Acar, I. & Butt, S.E. (2016). Modeling nurse-patient assignments considering patient acuity and travel distance metrics. *Journal of Biomedical Informatics, 64,* 192-206.

Akhtar, S., Ahmad, A.R. & Abdel-Rahman, E.M. (2012). A metaheuristic bat-inspired algorithm for full body human pose estimation. *Computer and Robot Vision (CRV), 2012 Ninth Conference on.* IEEE.

Allahverdi, A. & Al-Anzi F.S. (2006). A PSO and a Tabu search heuristics for the assembly scheduling problem of the two-stage distributed database application. *Computers & Operations Research 33(4),* 1056-1080.

Baatz, M. (2000). Multi resolution Segmentation: an optimum approach for high quality multi scale image segmentation. In. *Beutrage zum AGIT-Symposium.* Salzburg: Heidelberg.

Blumenstein, B., Siegmeier, T., Selsam, F. & Moller, D. (2018). A case of sustainable intensification: Stochastic farm budget optimization considering internal economic benefits of biogas production in organic agriculture. *Agricultural Systems, 159,* 78-92.

Bo, Y. (2008). A Method to Improve the Accuracy of Remote Sensing Data Classification by Exploiting the Multi-Scale Properties in the Scene, *Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences Shanghai, P. R. China, June 25-27, 2008, 183-18.*

Eberhart, R. C. & Shi, Y. (1998). Comparison between genetic algorithms and particle swarm optimization. *International conference on evolutionary programming.* Springer.

Eberhart, R. C. & Y. Shi (2000). Comparing inertia weights and constriction factors in particle swarm optimization. *Evolutionary Computation Proceedings of the 2000 Congress on*. IEEE.

Esetlili, MT., Bektaş Balçık, F., Balık Şanlı, f., Üstüner, M., Kalkan, K., Göksel, Ç., Gazioğlu, C. & Kurucu, Y. (2018). Comparison of Object and Pixel-Based Classifications For Mapping Crops Using Rapideye Imagery: A Case Study of Menemen Plain, Turkey, *International Journal of Environment and Geoinformatics (IJEGEO),* 5(2), 231-243.

Gandomi, A.H., Yang, X.S. & Alavi, A.H (2013). Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. *Engineering with computers, 29(1),* 17-35.

Gazioğlu, C (2018). Biodiversity, coastal protection, promotion and applicability investigation of the ocean health index for Turkish seas. *International Journal of Environment and Geoinformatics (IJEGEO).* 5(3), 353-367.

Holland, J. (1975). Adaptation in natural and artificial systems: an introductory analysis with application to biology. *Control and artificial intelligence.*

Huang, C.-L. & Dun, J.F. (2008). A distributed PSO–SVM hybrid system with feature selection and parameter optimization. *Applied soft computing, 8(4),* 1381-1391.

Huang, J. S. & Song, J. (2018). Optimal inventory control with sequential online auction in agriculture supply chain: an agent-based simulation optimisation approach. *International Journal of Production Research, 56(6),* 2322-2338.

Incekara, A., Seker, D. Z., Tezcan, C. S., Bozkutoglu, E., & Gazioglu, C. (2017). Interpreting temperature-based discontinuity and roughness of rock surfaces by using photogrammetric technique. *International Journal of Environment and Geoinformatics (IJEGEO),4*(3), 206–213.

Kennedy, J. (2011). Particle swarm optimization.In. *Encyclopedia of machine learning,* (p. 760-766). Springer.

Khan, K., Nikov, A. & Sahai, A. (2011). A fuzzy bat clustering method for ergonomic screening of office workplaces. In. *Third International Conference on Software, Services and Semantic Technologies S3T.* Springer.

Lemma, T.A. & F.B. M. Hashim (2011). Use of fuzzy systems and bat algorithm for exergy modeling in a gas turbine generator. In. *Humanities, Science and Engineering (CHUSER).* IEEE.

Li, J. H. (2018). Optimization and Operation Mechanism of Agriculture Products Electricity Supplier Logistics Distribution Based on Supply Chain Strategic Coordination. *Journal of Advanced Oxidation Technologies 21(2).*

Lillesand, T., Kiefer, R.W. & Chipman, J. (2014). *Remote sensing and image interpretation.* John Wiley & Sons.

Lindley, D. V. (1991). *Making decisions.* London: Wiley.

McCloy, KR. & Bøcher, PK (2007). Optimizing Image Resolution to Maximize the Accuracy of Hard Classification *Photogrammetric Engineering & Remote Sensing* Vol. 73, No. 8, 893–903.

Pal, S. K. & Wang, P. P. (2017). *Genetic algorithms for pattern recognition.* CRC press.

Perumal, K., Ungati, J.M., Kumar, G., Jain, N., Gaurav, R. & Srivastava, P.R. (2011). Test data generation: a hybrid approach using cuckoo and tabu search. In. *International Conference on Swarm, Evolutionary, and Memetic Computing.* Springer.

Poli, R., Kennedy, J. & Blackwell, T. (2007). Particle swarm optimization. *Swarm intelligence 1(1),* 33-57.

Sagir, M. & Ozturk Z. K. (2010). Exam scheduling: Mathematical modeling and parameter estimation with the Analytic Network Process approach. *Mathematical and Computer Modelling 52(5-6),* 930-941.

Saraç, T. & Özdemir, M. S. (2003). A genetic algorithm for 1, 5 dimensional assortment problems with multiple objectives. In. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems.* Springer.

Shi, Y. & Eberhart, R. C. (1998). Parameter selection in particle swarm optimization. In. *International conference on evolutionary programming.* Springer.

Simav, Ö., Şeker, D.Z., Tanık, A. & Gazioğlu, C. (2015). Determining the endangered fields of Turkish coasts with coastal vulnerability index. *Journal of Map,* 153: 1- 8.

Walton, S., Hassan, O., Morgan, K. & Brown, M.R. (2011). Modified cuckoo search: a new gradient free optimisation algorithm. *Chaos, Solitons & Fractals, 44(9),* 710-718.

Yang, X.S. & Deb, S. (2009). Cuckoo search via Lévy flights. In. *Nature & Biologically Inspired Computing* (*NaBIC 2009).* World Congress on, IEEE.

Yang, X.S. (2010). A new metaheuristic bat-inspired algorithm. In. *Nature inspired cooperative strategies for optimization (NICSO 2010),* (p. 65-74). Springer.

Yang, X.S. (2010). *Nature-inspired metaheuristic algorithms.* Luniver press.

Zhang, Q. & Izquierdo, E. (2006). A multi-feature optimization approach to object-based-image classification. *Image and Video Retrieval, Proceedings, 4071,* 310-319.