Approach Based on SPEA2-Band Selection and Random Forest Classifier to Generate Thematic Maps from Hyperspectral Images

Completed Research

Diego Saqui
Federal University of São Carlos (UFSCar), Federal Institute of Mato Grosso do Sul (IFMS)
diego.saqui@ufscar.br

José H. Saito
Federal University of São Carlos (UFSCar), University Center of Campo Limpo Paulista (UNIFACCAMP)
saito@dc.ufscar.br

Daniel C. Lima
Federal University of São Carlos (UFSCar)
daniel.lima@dc.ufscar.br

Lucio A. de C. Jorge
Embrapa Instrumentation
lucio.jorge@embrapa.br

Steve T. M. Ataky
Université du Québec à Montréal (UQAM)
steve.ataky@nca.ufma.br

Abstract

Hyperspectral images (HIs) and segmentation have become a promising solution for different applications such as the production of thematic maps (TMs) of agricultural areas. However, problems such as the Hughes phenomenon and high demand for computational resources are related to the high number of bands of HIs. This study proposes a hybrid approach of Strength Pareto Evolutionary Algorithm 2 (SPEA2) and Random Forest classifier for producing TMs, aiming at band selection and improvement of average recall of segmentation. In experiments, the proposed approach reduced the number of bands on average from 220 to 30 in the Indian Pines image and from 224 to 42 in the Salinas image. The proposed approach was statistically whether identical or better than other approaches regarding the average recall of segmentation. Therefore, the proposed approach is promising as regards band selection and competitive in segmentation being a potential tool for generating TMs.

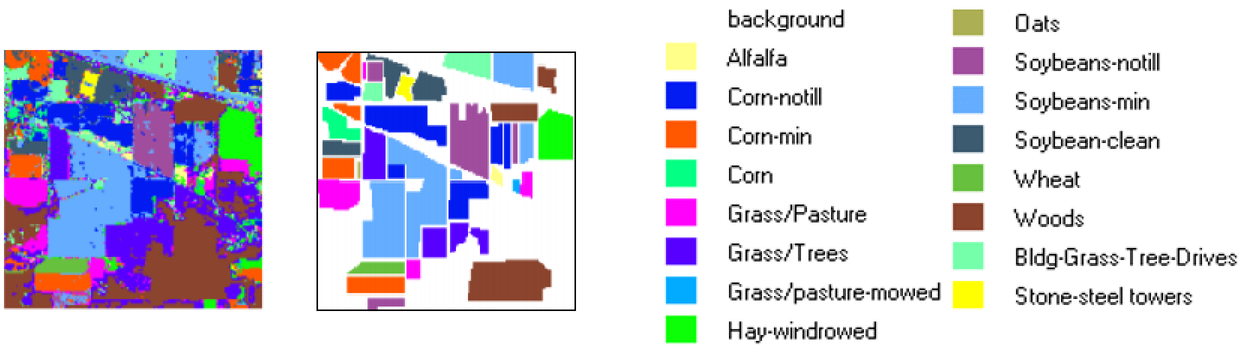
Keywords

Thematic Maps, Band Selection, Hyperspectral images, Remote Sensing.

Introduction

Hyperspectral images (HIs) contain spectral information of materials beyond the visible range of the spectrum. Each pixel can be represented by a spectral signature composed of a large number of bands that vary according to the kind of the object, allowing tasks of identification and analysis (Chang 2003). Among other types of images, HIs have the advantage of having unique bands that enable discrimination of different types of materials more easily (Attas et al. 2003; Khan 2018). Because of this capability, they are useful for medical imaging, remote sensing (RS), agricultural monitoring, among other applications (Gong, Zhang and Yuan 2015).

Applications with images and RS can occur by thematic maps (TMs). TMs show the spatial distribution of identifiable earth surface features; it provides an informational description over a given area, rather than a data description. The themes can range, for example, from categories such as soil, vegetation, and surface water in a general description of a rural area, to different types of soil, vegetation, and water depth or clarity for a more detailed description (Schowengerdt 2007). In particular, in agricultural monitoring, TMs can help identify crop problems, such as the presence of pests, and prevent pesticides from being applied in excess, impacting the economy and preserving the environment. TMs facilitate the understanding of the observed area and the communication between experts from different specialties. Figure 1 shows an example of TM for Indian Pines HI, the ground truth (GT) of this image that is provided by a specialist and represents the actual region of the captured scene and their respective classes.



(a) (b) (c)

Figure 1. (a) Thematic Map. (b) Ground Truth. (c) Classes of Indian Pines HI. Adapted from Tinoco et al. (2013).

The Indian Pines HI in Figure 1 was captured by RS and was obtained from an agricultural area. This HI is used for large-scale monitoring and this monitoring is more efficient than manual methods, especially with the expansion of RS devices as drones. Therefore, applications that are capable of generating TMs are promising for agriculture and other areas that need monitoring.

For creating TMs, algorithms with peculiar statistical properties are required for classification and segmentation. The classification categorizes a pixel or a set of the pixel of an image in its corresponding area. Some examples of classifiers are Support Vector Machines (SVM), Gaussian Maximum Likelihood Classifier (GMLC) (Lillesand, Kiefer and Chipman 2014), and Random Forest (RF) (Dinç and Aygün 2013; Amini et al. 2018; Xia et al. 2018; Sun et al. 2018). The pixels of the same category when connected form regions called segments, that is, it characterizes the segmentation.

Although TM generated from HIs have many advantages, some challenges are related to the data volume: (a) information redundancy due to the high correlation of neighboring bands; (b) the curse of dimensionality; (c) high demand for computational resources, such as storage, processing and data transmission; and (d) the high cost of sensors to capture HIs. Because of these problems, the selection of relevant bands for further classification and thematic mapping process is required.

Methods that reduce features of the data are described as dimensionality reduction and can be categorized into feature extraction and feature selection (Kumar 2004). Feature selection is more feasible for dimensionality reduction in HIs because it preserves the physical information of the original data (Gong, Zhang and Yuan 2015) and for this purpose, it is called band selection. Another categorization divides feature selection into Filter, Wrapper or Embedded approaches.

Filter approaches score variables and then eliminate some before constructing a model and the wrapper approaches use a predictive measure of the model, such as accuracy, and directly evaluate the value of the feature set (Tuv et al. 2009). Owing to the fact that wrappers approaches are directly related to the model, they can produce better results than filters, on the other hand, they demand for more computing resources.

Some examples of these approaches are based on Genetic Algorithms (GAs) (Zhang, Sun and Li 2009; Saqui et al. 2016; Vaiphasa et al. 2017; Zhuo et al. 2008) and Particle Swarm (Zhang et al. 2012; Zang, Ma and Gong 2017), that can return satisfactory results during a given number of iterations. These algorithms operate with a single objective function which is usually a metric that leads to a good segmentation such as pixel classification (for wrappers) or region separation (for filters). They can also combine one of these metrics with reducing the number of bands as shown in Zhuo et al. (2008), but the problem is that they are subject to unbalanced solutions because of the existence of conflicting objectives such as reducing the number of bands and improving the classification.

Multiobjective optimization band selection (MOBS) search for balanced solutions that improve segmentation and reduce the number of bands simultaneously (Gong, Zhang and Yuan 2015). Different filters MOBS can be found in the literature such as Non-dominated Sorting Genetic Algorithm 2 (NSGA2) (Kumar 2004) and Multiobjective Evolutionary Algorithm Based on Decomposition (MOEA/D) with Tchebycheff Decomposition (Gong, Zhang and Yuan 2015; Xu, Shi and Pan 2017). Another algorithm that can be applied as filter-based feature selection is the Strength Pareto Evolutionary Algorithm 2 (SPEA2) (Xue et al., 2013; Zitzler, Laumanns and Thiele, 2001). Studies have shown that SPEA2 can operate better than the traditional NSGA2 for a small number of iterations, small population, and noisy data as obtained by sensors, that is, HIs (Bui et al. 2005; Cholodowicz and Orlowski 2017).

Embedded approaches use all the variables to generate and analyze a model to infer the importance of the variables (Tuv et al. 2009). Examples of embedded approaches are decision trees (DTs) and RF. RF is a supervised algorithm which can produce promising results in HI segmentation (Dinç and Aygün 2013; Amini et al. 2018). A hybrid approach of DTs and feature selection with GA showed superior results about the use of the DT approach without prior band selection technique. The authors suggest that this is because the hybrid approach can focus on relevant features (Stein et al. 2005). This performance improvement may have occurred due to feature selection before the application of RF (Tuv et al. 2009; Sun et al. 2018).

In this paper, we propose a hybrid approach of SPEA2-based band selection (wrapper) and RF classifier to generate TMs. The proposed approach is modeled as a supervised MOBS dealing with the high dimensionality of HIs by band selection and maintaining the average recall performance of segmentation. Therefore, the main purposes of this study are:

- Select relevant bands of HIs with the proposed approach;
- Produce TMs with the selected bands;
- Structure a supervised MOBS to search tradeoff solutions;
- Evaluate the proposed approach by comparing it with GA-based supervised approach that combines two metrics (Zhuo et al. 2008) and the direct application of RF in HIs with all bands.

The remainder of this paper is organized as follows: The second section presents a detailed description of this approach. The experiments are conducted and results are reported in the third section. At last, the main contribution of this paper is summarized and future work direction is pointed.

Approach based on SPEA2-band selection and Random Forest

Multiobjective Optimization Band Selection (MOBS)

GA is an adaptive optimization search methodology based on Darwin's theory of natural selection and genetics in biological systems. It is an alternative to conventional heuristic methods.

GAs can represent solutions (individuals) to a problem by a vector, and an example of representation for HIs and bands is to use the number 1 for presence and 0 for the absence of the band equivalent to the position of these values in the vector. With this representation, the GA performs a search process trying different combinations of the bands in the solutions in order to improve the value obtained from the so-called fitness function.

GA can be started with a random set of solutions and try to optimize them by some iterations (called generations) or until meet a requirement. The optimization happens through the selection of the solutions with better values obtained from the fitness function and the application of crossover and mutation operators. The crossover is the action of combining parts of solutions, generating new solutions. Mutation consists of changing the values of solutions (such as 0 to 1) and can be applied so that the search process does not get stuck in local optima (Zhuo et al. 2008).

Usually, the fitness function used in GA can improve a single metric, such as overall accuracy, entropy or other metrics related to segmentation. This function can also be used to combine two metrics as shown in Zhuo et al. (2008). One problem with this single objective functions is that they may not be able to operate with conflicting metrics. For example, a reduction in the number of bands cannot occur if only a fitness function for segmentation improvement is applied. This happens because, during the search process, bands that improve and not improve the segmentation be present together in the solutions and the latter are never eliminated since the fitness function is not designed for this. MOBS were proposed to try to solve these problems since they are suitable to operate with conflicting objectives searching for tradeoff solutions.

The term MOBS was the first time introduced by Gong, Zhang and Yuan (2015) to refer to multiobjective algorithms of feature selection applied in the band selection of HIs but a similar study had already been developed by Kumar (2005). The main algorithms found in this category were the MOEA/D and the NSGA2. All these approaches operated with metrics that characterize them as filters, that is, they did not incorporate a supervised classification algorithm during their execution. Another common feature among these studies is the use of the Pareto frontier (PF) theory.

The best solution set of the search space composes the PF and solutions of this set cannot be said better among each other. For the establishment of the PF, it is necessary to compare all solutions in the search set and for this, algorithms like the NSGA2 and SPEA2 use the non-dominance relation given in (1) that means:

- The solution x dominates the solution y because x is better than y in at least one objective and equal or better in all other objectives

$$x \succ y \tag{1}$$

Solutions not dominated by any other are from the PF.

SPEA2, which is of interest in this study, is also a multiobjective optimization algorithm that can be modeled as MOBS. Reports in the literature, describe that this algorithm works well for noisy data, as in the case of images acquired by sensors and therefore is explored in this study.

For the establishment of a ranking of the solutions of the PF and other solutions the SPEA2 uses a specific function, where, first for each solution x, the raw fitness $R(x)$ in (2) is calculated.

$$R(x) = \sum_{y \succ x} S(y) \tag{2}$$

$S(y)$ is strength value, given by the total number of solutions that y dominates, established using (1), in the population and the archive. Then, SPEA2 uses a density estimation technique represented in (3).

$$D(x) = 1/(\sigma_x^k + 2) \tag{3}$$

where σ_x^k is the Euclidean distance of x concerning to the k-th solution of a list of distances. This list is in ascending order and contains the distances of all solutions in archive and population relative to x . k is given by the square root of the sum of the population size and archive size.

Finally, the value of each solution used in the selection operation is calculated by the function shown in (4) (Zitzler, Laumanns and Thiele, 2001).

$$F(x) = R(x) + D(x) \tag{4}$$

An important observation is that (1) is established with a set of fitness functions that must be previously established, such as the function to evaluate the segmentation and another to evaluate the number of bands.

One of the functions used in this study is based on the average recall that to be calculated it is necessary to apply an algorithm for classification where the RF is used.

Random Forest (RF)

RF is an algorithm that creates an ensemble of DTs to obtain a more accurate and more stable prediction (Breiman 2001). Supervised algorithms such as C4.5 can model DTs, but it is required a set of training samples with defined classes provided by a specialist. The modeling usually happens by evaluating how important a feature (or band) is for classification in those samples, and metrics used for this evaluation are the Gini impurity and Information Gain Entropy. In addition to the classification, the RF also performs a feature selection while structuring each leaf node of the tree and therefore it is explored in this study.

Proposed approach

In this approach, the SPEA2 conducts the band selection by some iterations (or generations). In each generation, it establishes a population and an external archive that contains solutions that are analyzed in the search process. A binary vector characterizes each solution/HI.

The following fitness functions are applied to each solution:

- **Fitness 1:** Increasing the average recall of the pixel classification performed by RF in images generated by each solution as shown in (5),

$$Average\ Recall = \frac{1}{n}\sum_{i=1}^{n} tp_i/(tp_i + fn_i) \tag{5}$$

  where n is the number of classes, i is the class, tp_i is the total of pixels correctly classified, and fn_i is the total of pixels erroneously classified.
  The recall is the rate of correctly classified pixels in each class and if this value is bad for some class then the average recall of all classes can be penalized. Therefore, improving the average recall of the classes implies increasing the segmentation performance, and so this metric was chosen as fitness.
- **Fitness 2:** Reducing the number of bands of HIs where it is considered the smallest possible value.

The process of the proposed approach is adapted from the original SPEA2 and shown in Algorithm 1.

Algorithm 1 SPEA2-based approach for band selection

Input: N (population size), $\bar{N}$ (archive size), T (number of generations), GT, $OrigHI$ (HI with all bands), M (mutation rate), and R (recombination rate).

Output: Solution with best average recall of the last generation.

Step 1 - Initialization: Creates an initial population P_0 and empty external archive $\bar{P}_0 = \emptyset$. Set $t = 0$. For P_0 are generated solutions with random combinations of bands from $OrigHI$.

while t <= T **do:**

  Step 2* - Classification: Train and classify the pixels of HIs generated from new solutions using RF.

  Step 3 - Fitness assignment: Establish fitness values of solutions in P_t and $\bar{P}_t$, calculating the average recalls in (5), the number of bands, and the fitness values in (4).

  Step 4 - Environmental selection: Copy all nondominated solutions in P_t and $\bar{P}_t$ to $\bar{P}_{t+1}$. If size of $\bar{P}_{t+1}$ exceeds $\bar{N}$ then reduce $\bar{P}_{t+1}$ by means of the truncation operator, otherwise if size of $\bar{P}_{t+1}$ is less than $\bar{N}$ then fill $\bar{P}_{t+1}$ with dominated solutions in P_t and $\bar{P}_t$.

  Step 5 - Mating selection: Performs binary tournament selection with replacement on $\bar{P}_{t+1}$ to fill the mating pool.

  Step 6- Variation: Apply recombination and mutation to the mating pool and set P_{t+1} to the resulting population. Increment the counter ($t = t + 1$).

Step 7- Select one solution: Selects the solution with best average recall from $\bar{P}_t$.

In algorithm 1, initially, population and archive sizes, HI with all bands ($OrigHI$), GT, and mutation and recombination rates are defined. In step 1, the population of solutions is created with a random combination of bands generated from HI with all bands.

Then, in step 2, for each solution, 3-fold cross-validation is applied for training and validation, and the average value of the average recall obtained is used as fitness. The samples (pixels) are obtained by systematic sampling from the GT through an iterative process by selecting pixels alternately for each fold. The training samples are used to establish a model based on the RF, which is later validated with the validation samples. *The use of the RF in step 2 is the main modification about the other MOBS proposed in the literature characterizing the proposed approach as the wrapper.

In step 3, fitness values and non-dominance relationships between all solutions are established.

In step 4 the environmental selection is performed, where all non-dominated solutions of the population and archive are copied to the new archive of the next generation. If this archive is smaller than the predetermined size, the dominated solutions are sorting according to (4) and then copied to fill up to the archive size. If the next generation archive has the number of non-dominated solutions that exceeds the predefined size, then the solutions to be removed are those that have less distance σ_x^k than others. If there are solutions with the same distance, SPEA2 considers the second minimum distance, and so on.

In step 5 the binary tournament selection for mating is performed and in step 6 the recombination and mutation operations are applied to compose the population of the next generation. A predetermined rate determines the occurrence of recombination between solutions, and in this paper, the one-point crossover is used. The mutation operation occurs in new solutions, where we apply the flip bit that allows changing the value of each vector position (from 1 to 0 or 0 to 1) based on a predetermined rate. In step 7 the solution that generated the image with the best average recall of the PF of the current population is selected.

After applying these steps, the selected bands can be used to construct an RF model that can later be applied to segmentation and generation of TM.

Experiments and results

Experiments

There is a limitation of free datasets of HIs with GT available, so for this study, two traditional HIs of agricultural areas were used, Indian Pines and Salinas. These HIs were obtained through the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) and correspond mainly to land covers with crops. Table I shows information about these datasets.

Name	# of pixels	# of classes	# of bands
Indian Pines	21025	16	220
Salinas	7138	6	224

Table 1. Information about datasets.

The proposed approach was compared with the GA-based approach that searches the best combination of bands with the average recall fitness as shown in equation (6).

$$f = w_a * average\ recall + (w_n / number\ of bands) \quad (6)$$

where w_a and w_n are the weights for the average recall and number of bands and values of 0.8 and 0.2 were used, respectively. This was the only wrapper approach of band selection found combining two metrics and therefore was chosen for comparison.

Both approaches were configured to run for 500 generations and with a population of 100 solutions. The recombination/crossover and mutation rates were 50% and 5% respectively. The recombination/crossover and mutation operation of the two approaches are the same, and the binary tournament selection was used.

For RF classifier, ten DTs generated from the Gini indices (measures the degree of heterogeneity of each node) and bootstrap samples were used. In the RF application, the rules previously presented of training and validation were used, as well as the criteria for selecting samples.

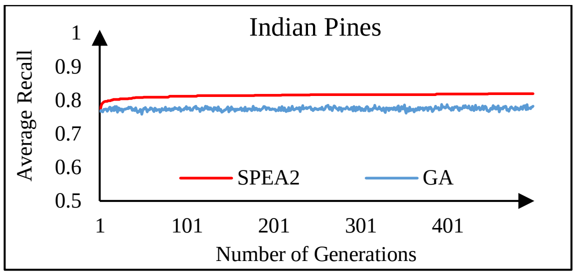
We also compared the proposed approach with the results of segmentation of GA-based approach and the original HIs with all bands. In this comparison, we used the last generation solution of each approach and evaluated the average recall and the accuracy of the segmentation.

The process of band selection (training and validation) and tests was organized in a 10-fold cross-validation for each image. We use 80% of the pixels for the band selection and 20% for final tests. The 20% pixels of each test fold were selected to be different from each other, and so was the 80% pixel of each fold used in band selection. Based on this structure, the different band selection algorithms were applied and after the band selection, statistics were obtained from the test data.

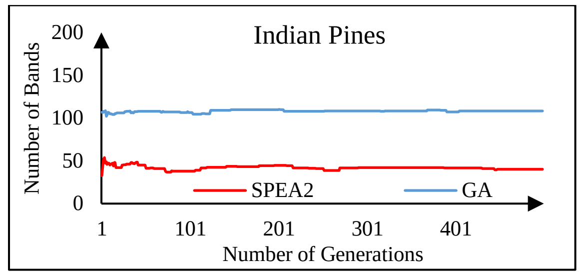
The algorithms used were developed using the Python 3.5 language and OpenCV, Scikit-Learn, Spectral Python and Distributed Evolutionary Algorithms in Python (DEAP) packages.

Results

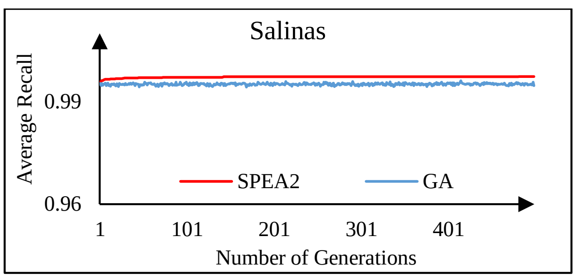
The following graphs show the average evolution of the solution with the best average recall in each generation obtained from the ten executions for the proposed approach (with SPEA2 and RF) and the GA-based approach. These graphs help analyze the behavior of each approach in the band selection process.



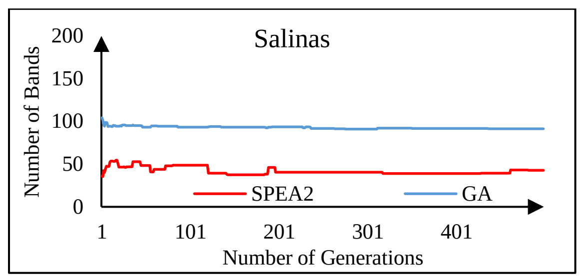
(a) Indian Pines - Average Recall



(b) Indian Pines - Number of bands



(c) Salinas - Average Recall



(d) Salinas - Number of bands

Figure 2. Graphs of the evolution of the average recall and number of bands for Indian Pines and Salinas HIs with band selection performed by proposed approach and GA.

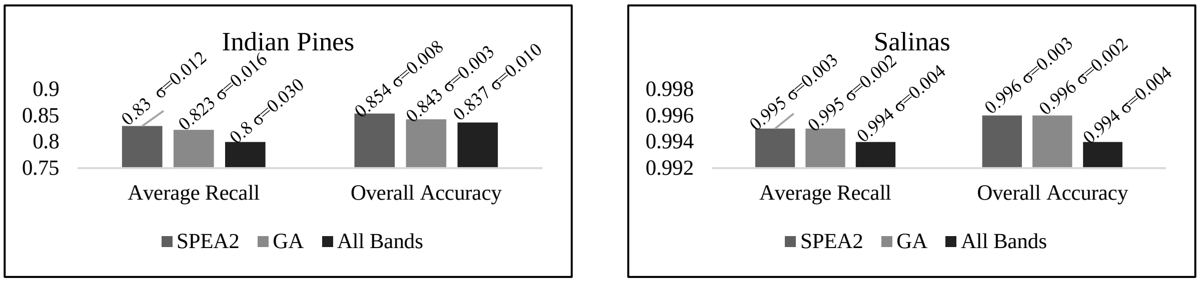
In the graph of Figure 2a, which refers to Indian Pines, it is shown that the proposed approach had an average evolution of the average recall better than the GA-based approach. The graph in Figure 2b shows the average evolution of the number of bands under the same conditions of the graph of Figure 2a, and when analyzed together, we can see that the proposed approach reduces the number of bands whilst practically maintaining the average recall, whereas the GA-based approach cannot reduce the number of bands. The improvement of the average recall and reduction of the number of bands performed by the proposed approach is noticeable mainly at the beginning of the generations (between the first and the one hundred fiftieth generation). This characterizes the behavior of MOBS in the proposed approach where tradeoff solutions were found.

Figures 2c and 2d show the average evolution of the average recall and the number of bands respectively for Salinas image. In these graphs the behavior of both algorithms for Salinas is very similar to the results of the Indian Pines, maintaining a high average recall value, and the proposed approach reduces the number of bands more than the GA-based approach.

An observation in Figures 2b and 2d where the number of bands is shown, for the proposed approach, the function is not monotonically decreasing because the best average recall solution is selected from the Pareto set of each generation.

The characteristics shown in the average evolution of these graphs (Figure 2) suggest an improvement in the treatment of the dimensionality and the removal of redundant bands without degrading the average recall. Under these conditions, the proposed approach was better than the GA-based approach.

In Figure 3 are shown the results of the average recall and overall accuracy for each HI with bands selected by the methods compared, obtained from the classification result by the algorithm GMLC. The values shown above the columns are the averages obtained from the 10-Fold cross-validation with their respective standard deviations.



(a) Indian Pines (b) Salinas

Figure 3. Average values of average recall and overall accuracy obtained from the classification result by the GMLC.

A paired t-test with the significance of 0.05 was applied in the average recall and overall accuracy, and, statistically, the proposed approach had the same mean as the other cases compared, thus equal performance for these metrics. Just for Indian Pines image, the proposed approach was statistically better than the use of all bands for overall accuracy.

Name	Indian Pines	Salinas
Proposed Approach	39	42
GA	108	91
HI with all bands	220	224

Table 2. Average number of bands selected at the end of the executions of each approach and total of original bands.

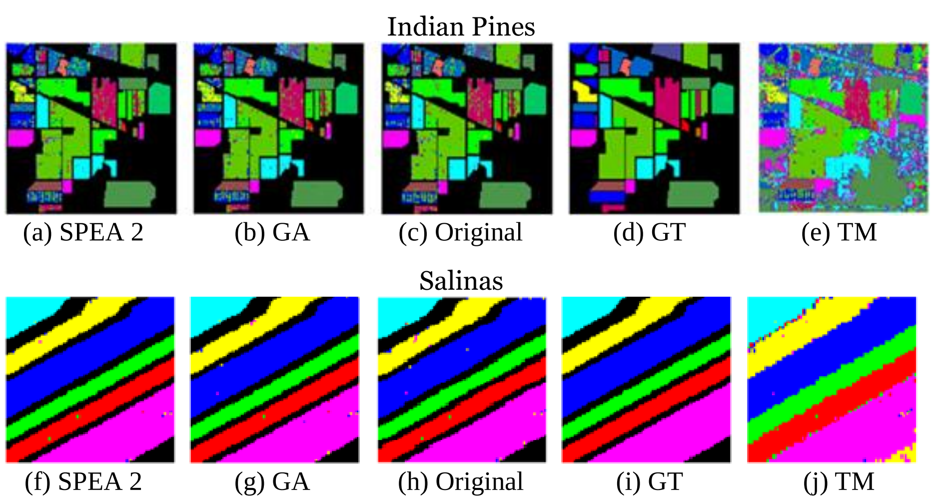


Figure 4. HIs segmented: (a) Indian Pines with bands selected by SPEA2. (b) Indian Pines with bands selected by GA (c) Indian Pines with all bands. (d) GT Indian Pines. (e) TM for Indian Pines. (f) Salinas with bands selected by SPEA2. (g) Salinas with bands selected by GA (h) Salinas with all bands. (i) GT Salinas. (j) TM for Salinas

The main advantage of the proposed approach was in the reduction of bands and the final average number of bands selected by each approach is shown in Table II. Together with the results obtained in the pixel classification, it shows that the proposed approach can be applied to generate thematic maps from HIs.

Figure 4 shows examples of TM obtained from the band selection of the first execution of each approach. In Figures 4a-4c and 4f-4h the segmented images of Indian Pines and Salinas are respectively shown. The areas not present in the GT were removed. Figures 4d and 4i, respectively show the GT of the Indian Pines and Salinas. These regions are important because they are areas of interest in thematic maps shown in Figures 4e and 4j. The result of this segmentation is obtained from the training of RFs with 30% of the pixels of each class, which were replaced in the HIs for later classification.

The results shown in this section corroborate that the proposed approach can improve the selection of bands performed by the RF, contributing to a good classification result and allowing the generation of TMs.

Conclusion

A novel supervised MOBS using SPEA2 and RF for the production of thematic maps from HIs has been described herein. Given the relationship between quality of segmentation and number of bands, the idea of the approach is to search for tradeoff solutions with the objectives of improving the average recall of the pixel classification and reduction of the number of bands. Therefore, the feature space used by RF to generate their DTs can be optimized reducing the number of bands for classification without loss of quality in the segmentation.

The proposed approach was compared with the direct application of RF as well as a GA-based approach. Experiments on the Indian Pines and Salinas HIs showed that the proposed approach was competitive in the average recall and outperformed in the reduction of the number of bands. Therefore, the approach can be applied to generate thematic maps from HIs to reduce the computational cost of processing these images without loss of quality.

As the approach has always selected the best average recall solution of the Pareto Frontier, we believe that a decision-making algorithm can improve the performance of reducing the number of bands and will be exploited in future works. Other subjects for future works are implementing a strategy to correct misclassified pixels in the segments, testing other fitness functions with metrics such as accuracy, precision, recall, F1-score and Area Under the Curve (AUC) for the Receiver Operating Characteristic (ROC) and making comparisons with other approaches in the literature.

REFERENCES

Amini, S., Homayouni, S., Safari, A., Darvishsefat, A. A. 2018. “Object-based classification of hyperspectral data using Random Forest algorithm”. In: Geo-spatial Information Science, vol. 21, pp. 127-138.

Attas, M., Cloutis, E., Collins, C., Goltz, D., Majzels, C., Mansfield, J. R., Mantsch. H. H. 2003. “Near-infrared spectroscopic imaging in art conservation: Investigation of drawing constituents”, Journal of Cultural Heritage, vol. 4, no. 2, pp. 127-136.

Breiman, L. 2001. “Random forests”. In: Machine Learning, vol. 45, no. 1, pp. 5–32.

Bui, L. T., Essam, D., Abbas, H. A., Green D. 2005. “Performance analysis of evolutionary multiobjective optimization methods in noisy environments”. In: Proceedings of The 8th Asia Pacific Symposium on Intelligent and Evolutionary Systems, pp. 29-39.

Chang, C. I. 2003. Hyperspectral Imaging: Techniques for spectral Detection and Classification. Springer Science & Business Media.

Cholodowicz, E., Orlowski, P. 2017. “Comparison of SPEA2 and NSGA-II Applied to Automatic Inventory Control System Using Hypervolume Indicator”. Studies in Informatics and Control, vol. 26, pp. 67-74.

Dinç, S., Aygün, R. S. 2013. “Evaluation of Hyperspectral Image Classification Using Random Forest and Fukunaga-Koontz Transform”. In: International Workshop on Machine Learning and Data Mining in Pattern Recognition. MLDM 2013: Machine Learning and Data Mining in Pattern Recognition. Lecture Notes in Computer Science, vol 7988. Springer, pp. 234-245.

Gong, M., Zhang, M., Yuan, Y. 2015. “Unsupervised Band Selection Based Evolutionary Multiobjective Optimization for Hyperspectral Images”. In: IEEE Transactions on Geoscience and Remote Sensing, vol. 54, n. 1, pp. 544- 577.

Khan, M. J., Khan, H. S., Yousaf, A., Khurshid, K., Abbas, A. 2018. “Modern Trends in Hyperspectral Image Analysis: A Review”. In: IEEE Access, vol. 6, pp. 14118-14129.

Kumar, M. 2004. “Feature Selection for Classification of Hyperspectral Remotely Sensed data using NSGA-II”. In: Water Resources Seminar, Citeseer.

Lillesand, T. M., Kiefer, R. W., Chipman, J. W. 2004. “Remote Sensing and Image Interpretation”, 5th ed., Ed. John Wiley & Sons.

Saqui, D., Saito, J.H., Jorge, L.A.C, Ferreira, E.J., Lima, D.C., Herrera, J.P. 2016. “Methodology for band selection of Hyperspectral Images using Genetic Algorithms and Gaussian Maximum Likelihood Classifier”. In: International Conference on Computational Science and Computational Intelligence, Las Vegas, EUA, pp. 733-738.

Schowengerdt, R. A. 2007. *Thematic Classification*, Remote Sensing (Third Edition), Academic Press.

Stein, G., Chen, B., Wu, A. S., Hua, K. A. 2005. “Decision tree classifier for network intrusion detection with GA-based feature selection”. In: Proceedings of the 43rd Annual Southeast Regional Conference on - ACM-SE 43.

Sun, M., Wang, C., Wang, S., Zhao, Z., Li, X. 2018. “A New Semisupervised-Entropy Framework of Hyperspectral Image Classification Based on Random Forest”. Advances in Multimedia, vol. 2018.

Tinoco, S., Santos, A., Santos, H., Menotti, D., Santos, J. A. 2013. “Ensemble of classifiers for remote sensed hyperspectral land cover analysis: An approach based on linear programming and weighted linear combination”, Proc. IGARSS, pp. 4082-4085.

Tuv, E., Borisov, A., Runger, G. C., Torkkola, K. 2009. “Feature selection with ensembles, artificial variables, and redundancy elimination”. Journal of Machine Learning Research, vol. 10, pp. 1341–1366.

Vaiphasa, C., Skidmore, A. K., Boer W. F., Vaiphasa, T. 2007. “A hyperspectral band selector for plant species discrimination”. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 62, n. 3, p. 225-235.

Xia, J., Ghamisi, P., Yokoya, N., Iwasaki, A. 2018. “Random Forest Ensembles and Extended Multi-Extinction Profiles for Hyperspectral Image Classification”. In: IEEE Transactions on Geoscience and Remote Sensing, vol. 56, n. 1.

Xu, X., Shi, Z., Pan, B. 2017. “A New Unsupervised Hyperspectral Band Selection Method Based on Multiobjective Optimization”. IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 11, pp. 2112-2116.

Xue, B., Cervante, L., Shang, L., Browne, W., N., Zhang, M. 2013. “Multi-objective evolutionary algorithms for filter based feature selection in classification”. International Journal on Artificial Intelligence Tools, vol. 22.

Zhang, M., Ma, J., Gong, M. 2017. “Unsupervised Hyperspectral Band Selection by Fuzzy Clustering with Particle Swarm Optimization”. IEEE Geoscience and Remote Sensing Letters, vol. 14, n. 5, pp. 773-777.

Zhang, X., Sun, Q., Li, J. 2009. “Optimal band selection for high dimensional remote sensing data using genetic algorithm”. In: Proc. SPIE 7471, Second International Conference on Earth Observation for Global Changes.

Zhang, X., Wang, W., Li, Y., Jiao, L. C. 2012. “Pso-based automatic relevance determination and feature selection system for hyperspectral image classification”. Electronics Letters, vol. 48, n. 20, pp. 1263–1265.

Zhuo, L., Zheng, J., Wang, F., Li, X., Ai, B., Qian, J. 2008. “A Genetic Algorithm Based Wrapper Feature Selection Method for Classification of Hyperspectral Images Using Support Vector Machine”. In: Proceedings of SPIE 7147, Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification.

Zitzler, E., Laumanns, M., Thiele, L. 2001. “SPEA2: Improving the strength Pareto evolutionary algorithm”. In: Evolutionary methods for design optimization and control with applications to industrial problems, Athens, Greece, pp. 95–100.