

Land Cover Classification using Compressive Covariance Sampling: A Case Study at Valle de San José, Colombia

Jonathan Monsalve, Ariolfo Camacho, Tatiana Gelvez-Barrera, Henry Arguello
Universidad Industrial de Santander
Bucaramanga, Colombia

Abstract

Land cover classification supports environmental protection, urban planning, and natural resource management applications. Remote sensing platforms acquire spectral images (SIs) with spectral data processed by specialized algorithms. However, extensive data can pose computational challenges when acquiring, transmitting, and processing SIs. Compressive spectral imaging technique (CSI) reduces the acquisition and transmission load by capturing compressed versions of the SIs. Further, compressive covariance sensing (CCS) can reduce the processing load by operating over the covariance matrix estimated directly from the compressed SI. This paper presents a land cover classification scheme based on CCS, composed of a deep-learning feature extractor and a support vector machine classifier. The proposal is evaluated as a case study in Valle de San José, Santander, Colombia, considering the interest of the governmental institutions in monitoring the wide crop variety in the region's land cover. Experimental results show the classification scheme effectiveness over a real scenario, achieving a comparable performance of literature methods using complete data but a reduced computational load.

1. Introduction

Land cover classification inventories record comprehensive information on the materials that cover the surface of the earth's landmass [1]. A remote sensing platform builds a land cover map using satellites to acquire spectral images (SIs), providing a rich understanding of the surface features useful for precisely classifying land cover types, such as vegetation, soil, and urban structures [2].

Remote sensing platforms are computationally limited when acquiring, transmitting, and processing vast SI data. Hence, compressive spectral imaging (CSI) has emerged as a mathematical tool that solves acquisition and transmission data overload by acquiring a compressed version of the SI while maintaining the essential spectral features [3]. After the acquisition, CSI employs deep-learning or optimization-based computational algorithms, taking advantage of SI

prior information such as sparsity, low-rank, or smoothness to perform reconstruction [4], target detection [5], or classification tasks [6]. Beyond that, the compressive covariance sensing technique (CCS) increases the processing speed, estimating the covariance matrix (CM), directly from the compressed measurements, and classifying using a low-rank PCA-based approximation [7].

This paper proposes a land cover classification scheme involving four components: (i) the compressed SI modeling, (ii) the SI fast low-rank approximation based on the CCS technique (iii) the feature extractor learning, training a deep-learning model with several SI publicly available datasets, and (iv) the classification step, taking advantage of the support vector machine to classify with a few in-situ acquired samples. The proposed classification scheme is evaluated as a case study at Valle de San José in Santander, Colombia, considering the crop variety of the region and ease of access to the rural areas. Specifically, a set of SIs acquired with the Sentinel 2B satellite were employed for the compressive acquisition simulation. Further, the computational simulations were performed using a fast reconstruction from the estimated CV. The experiments section analyzes the obtained accuracy when varying the compression ratio, showing that the CCS-based classification performs comparably to the classification using full data, while using a lower computational load.

2. Area of Studio and Classes Definition

The studio was conducted in Valle de San Jose, Santander, Colombia, presenting a mountainous geography characterized by the cultivation of coffee, sugar cane, and cattle pastures, coffee being the most significant crop with the highest presence. The selected region is centered on the WGS84 geographic coordinates of $6^{\circ} 59' 1300''$ N, $73^{\circ} 40' 6200''$ W, whose referenced image was acquired on September 10, 2022, using the Sentinel-2 VNIR sensor [8]. For the study, we selected a subregion of 440×680 spatial pixels and 10 spectral channels covering the 490 nm to 2190 nm spectral range. Originally, the spatial resolution is of

10m for the 2,3,4 and 8 spectral channels; and 20m for the 5,6,7,11, and 12 spectral channels. The spatial resolution was adjusted at 10m for all channels by sub-sampling the latter channels.

2.1. Definition of the Classes

The identification of the predominant classes was carried out by a visit in-situ in the area of studio, analyzing the ground truth of the vegetation at some specific locations to create a land-cover inventory. The region contains a large agricultural area, mostly corresponding to coffee crops, followed by sugar cane, cocoa, and pastures. Hence, we define the following five classes: (i) Coffee crops and trees class. (ii) Agricultural vegetation class, excluding coffee. (iii) Bare soil class. (iv) Urban areas class. (v) Water class.

Figure 1 shows the selected classes, geo-referenced with control points taken with a sub-meter GPS and photographs to identify the pixels in the Sentinel-2b SI. Through this visit, around 500 pixels per class were identified in the satellite image using the captured GPS points on-site.

Figure 2 shows the location of pixels whose spectral signatures were extracted for the training and testing process for each defined class. Note the high similarities between the spectral response of the defined classes, indicating that the classification of such image is a challenging task.



Figure 1. Some classes selected at Valle de San José.

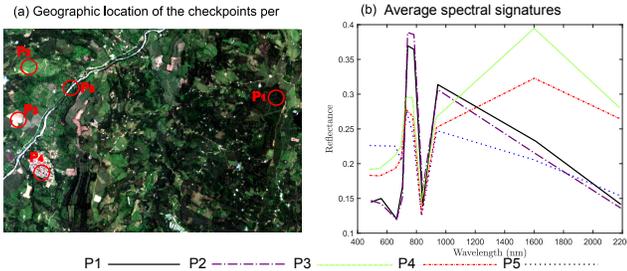


Figure 2. Average spectral response of the five defined classes. (a) Location of the area where the pixels can be found. (b) Coffee crop, P_1 . (c) Agricultural vegetation without coffee, P_2 . (d) Bare soil, P_3 . (e) Urban areas, P_4 , and (f) River, P_5 .

3. Proposed Classification Method

The proposed semi-supervised land cover classification scheme based on compressive covariance sensing (CCS) involves four components: (i) The SI compressive acquisition. (ii) The SI low-rank approximation based on CCS. (iii) The learning of a feature extractor with a convolutional neural network (CNN) model. And (iv) a support vector machine (SVM) classifier.

3.1. Compressive Spectral Imaging

Let $\mathbf{F} \in \mathbb{R}^{l \times n}$, denote the matrix form of a SI with l spectral bands and n spatial pixels. The acquisition of a compressed version $\mathbf{Y} \in \mathbb{R}^{m \times n}$, with $m \ll n$ can be modelled by the following CSI forward model

$$\mathbf{Y} = \mathbf{P}^T \mathbf{F} + \mathbf{N}, \quad (1)$$

where $\mathbf{P} \in \mathbb{R}^{l \times m}$ denotes a random projection matrix, and $\mathbf{N} \in \mathbb{R}^{m \times n}$ models acquisition noise [7,9].

3.2. Compressive Covariance Sensing Recovery

Given the compressed SI, CCS recovers the complete SI based on the CM, reducing the computational load of traditional CSI reconstruction algorithms. Nonetheless, in most cases the CM \mathbf{S} is unknown and must be estimated from \mathbf{Y} . Therefore, this work is based on the fast CM estimation approach presented in [7,9], which demonstrated that splitting the signal \mathbf{F} into p disjoint subsets $\mathbf{F}_i \in \mathbb{R}^{l \times n/p}$ and projecting them onto different subspaces $\mathbf{P}_i \in \mathbb{R}^{l \times m}$ allows to accurately estimate the CM \mathbf{S} from compressed measurements by solving the optimization problem

$$\mathbf{S}^* = \underset{\mathbf{S} \in \mathbf{D}}{\operatorname{argmin}} \sum_{i=1}^p \|\tilde{\mathbf{S}}_i - \mathbf{P}_i^T \mathbf{S} \mathbf{P}_i\|_F^2 + \tau \psi(\mathbf{S}), \quad (2)$$

where \mathbf{D} is the set of positive semi-definite matrices, ψ promotes a low-rank structure, and τ is a weighting parameter.

After estimating the CM \mathbf{S} , a low-rank approximation of the SI is computed as follows

$$\tilde{\mathbf{F}} = (\mathbf{P}^T \mathbf{W})^\dagger \mathbf{Y}, \quad (3)$$

where $\tilde{\mathbf{F}} \in \mathbb{R}^{k \times n}$ denotes the SI low-rank approximation with the first k PCA coefficients, and \mathbf{W} is a matrix containing the first k eigenvectors of \mathbf{S} [10].

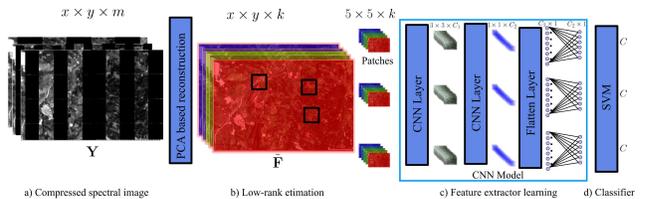


Figure 3. Schematic representation of the semi-supervised patch-based classification approach from compressed measurements.

3.3. Feature Extractor Learning using a Convolutional Neural Network

A classification CNN usually compresses the input into a latent space before applying a dense layer with a softmax to assign a class. The learned latent space can act as a feature extractor useful to fine-tune the model for a specific task [11, 12].

The proposed scheme learns the latent space using a patch-based classification approach, which extracts and labels spatial patches of the image, promoting spatial priors and smooth classifications [13]. The used CNN architecture consists of two convolutional layers and two dense layers, a simple architecture for fast and accurate SI classification with few trainable parameters. The input of the CNN corresponds to the first PCA coefficients of each patch, given the low-rank approximation estimated in the previous component, reducing the computational complexity.

3.4. Support Vector Machine Classifier

The fine-tuning removes the last layer of the CNN model used to learn the feature extractor and trains the model with new classes and data, taking advantage of the SVM generalization capabilities when dealing with a few samples [14]. Additionally, the computational complexity is also reduced since only the SVM needs to be trained using the latent projection of the signal generated by the CNN model.

Let $g_\theta : \mathbb{R}^l \rightarrow \mathbb{R}^s$ be a CNN model whose last layer has been removed, and (\mathbf{f}_i, y_i) a labeled pair of data for training. The optimization problem for the SVM is given by

$$\min_{\mathbf{u}, b} \frac{1}{2} \|\mathbf{u}\|^2 + C \sum_{i=1}^N \max(1 - y_i(\mathbf{u}^T g_\theta(\mathbf{f}_i) + b), 0)^2, \quad (4)$$

where \mathbf{u} contains the normal vectors to the hyperplane separating the classes [14].

Figure 3 illustrates the proposed semi-supervised land-cover classification using CCS. In a), the input is the compressed SI \mathbf{Y} in (1). In b), the CM is estimated by solving (2), and a low-rank approximation of the SI is obtained using the matrix \mathbf{W} with the first k eigenvectors of \mathbf{S}^* , following (3). In c), a patch-based classification CNN is used to learn a latent space used as the feature extractor, where the CNN architecture consists of two convolutional layers and two dense layers, trained with literature SI datasets covering various classes. Finally, in d), an SVM replaces the last layer of the CNN to fine-tune the classification task over the few samples acquired at Valle de San José using the training process according to the solution of (4).

4. RESULTS AND DISCUSSION

The experiments use four SI datasets. The Indian Pines, acquired by the AVIRIS sensor from the Northwestern Indian Pines test site in June 1992 with 145×145 spatial

pixels and 200 spectral bands [15]. The Pavia University and Pavia Center, acquired by the ROSIS sensor over Pavia, Northern Italy with 610×340 spatial pixels and 103 spectral bands [16]. And Salinas, acquired by the 224-band AVIRIS sensor over Salinas Valley with 512×217 spatial pixels [16].

Metrics: average accuracy (AA) computing the average accuracy for each class useful to evaluate unbalanced datasets; the overall accuracy (OA) measuring the correct number of pixels across the classes; the F1 score computed as a ratio between precision and recall, (the F1 score is computed as the average of the F1 for each class); and, kappa statistics considering that assigning labels randomly has a certain degree of accuracy, making this metric a robust way to measure classification performance [17].

4.1. Feature Extractor Training

The training process to learn the feature extractor uses three of the four datasets, addressing a completely supervised training (University of Pavia, Salinas, Indian Pines). The model is trained progressively using each dataset, applying a PCA, and adjusting the model with the corresponding classes. Specifically, 1) Apply PCA to obtain a dimensionality-reduced version using six components. 2) Adjust the dense layer to match the number of classes in the dataset. 3) Run 1000 iterations to retrain the feature extractor with the current classes. The process is repeated until the model converges, no matter which image is used.

4.2. Classification with Pavia Center dataset

The remaining dataset (Pavia Center) is used for semi-supervised training using the SVM as the classification layer keeping the CNN model weights fixed, used as the feature extractor. For the training, 100 pixels of each class were used to fine-tune the CNN model and to train the SVM classifier in the experiments. This experiment tests two scenarios, using full data with no compression other than the PCA dimensionality reduction and applying a random compression on the data using the CCS approach as described in Section 3 acquiring only 10% of the data.

Table 1 shows the results for different scenarios labeled as follows: using full data and the CNN model with a dense layer and a Sigmoid activation as a classifier (CNN-full); using full data and the CNN model with the SVM as a classifier (SVM-full); using compressed data and the CNN model with a dense layer and a Sigmoid activation as a classifier (CNN-CCS); and using compressed data and the CNN model with the SVM as a classifier (SVM-CCS).

There, it can be seen that the classification performance when using compressed versions of the dataset achieves comparable results in terms of classification accuracy with those obtained using full data. Furthermore, using the SVM achieves comparable or even improves the Kappa metric and F1 score, indicating that the SVM approach provides

a better classification agreement and balance between the recall and precision metrics.

Figure 4 shows a visual comparison between the four evaluated scenarios. It can be seen that the results of the compressed datasets are comparable to those obtained using full data, especially when the SVM classifier is used as visualized in the boxed region where the CNN scenario presents artifacts.

4.3. Experiments with Valle de San José data

The Valle de San José data was compressed following equation (1) using 60% of the information. The classification results are compared against the results obtained using full data. Figure 5 illustrates the obtained classifications over the real data, where using the compressed measurements achieves comparable results with those obtained with full data. Since no ground truth is available, we present a qualitative analysis of the predicted classes. Based on the knowledge of the zone and the visit in situ, we observed a high correlation with the expected distribution of

Model	AA(%)	OA(%)	Kappa	F1 score
CNN-Full	98.6 ± 0.2	93.8 ± 0.8	91.5 ± 1	87.0 ± 1.3
SVM-Full	98.6 ± 0.2	93.8 ± 0.9	91.5 ± 1.3	86.7 ± 1.8
CNN-CCS	98.5 ± 0.4	93.3 ± 1.8	90.6 ± 2.5	85.2 ± 2.4
SVM-CCS	98.6 ± 0.1	93.8 ± 0.8	91.3 ± 1.1	85.8 ± 1.1

Table 1. Classification quantitative results Pavia Center Image.

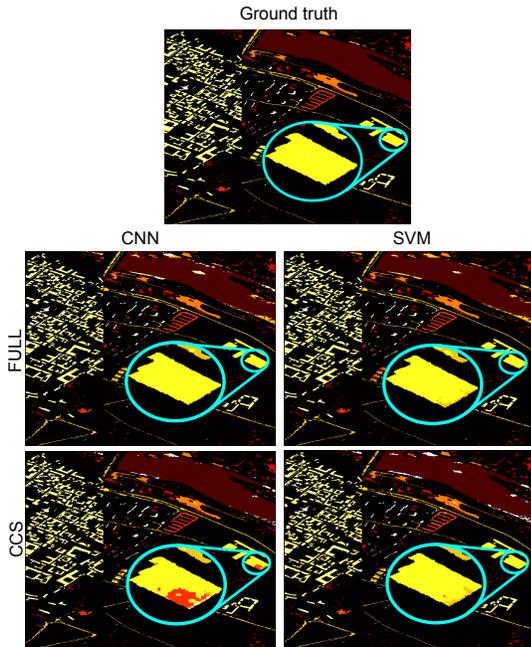


Figure 4. Visual comparison of the classification results. Note that the black regions are unlabeled zones hence, it was masked out in the results for interpretability

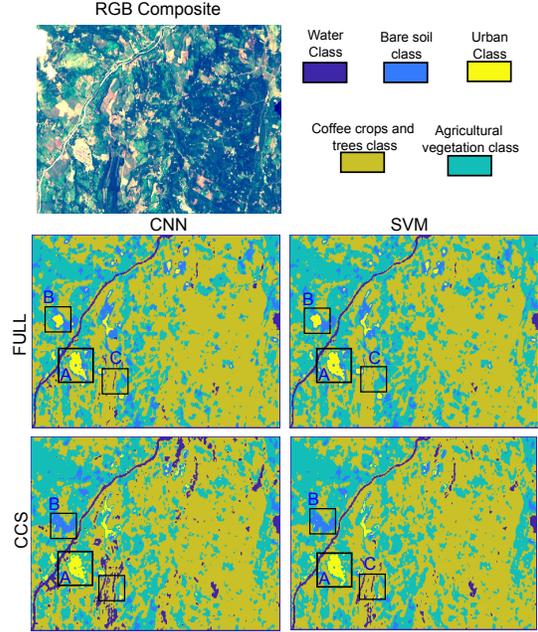


Figure 5. Visual comparison of the results obtained using the data from valle de san josé

the classes. The (A) region in figure 5 corresponds to the location of the town which is correctly predicted. Region B corresponds to bare soil, it is interesting that using the compressed measurements the algorithm was able to predict it correctly meanwhile, using full data misclassified it. Finally, region C was predicted as water in both compressed measurements, but there is no water in that zone. Overall, the use of an SVM with compressed data exhibits a better performance than using only a CNN model.

5. Conclusions

A land cover classification scheme based on CCS was presented. The scheme was developed to be used in a case study over a region in Vallé de San José, Santander, Colombia. For this, we carried out a visit in-situ to the region, analyzing the vegetation present at different locations and defining the predominant classes to build a land cover inventory. The developed classification scheme involves a deep-learning feature extractor and a support vector machine classifier, considering the few observed samples of the different classes of interest in the real study case. We conducted experiments over different configurations, showing the effectiveness of the proposed classification scheme to discriminate the predominant classes, achieving a comparable performance of literature methods using complete data at a reduced computational load.

References

- [1] T. Hermosilla, M. A. Wulder, J. C. White, and N. C. Coops, "Land cover classification in an era of big and open data: Optimizing localized implementation and training data selection to improve mapping outcomes," *Remote Sensing of Environment*, vol. 268, p. 112780, 2022. 1
- [2] D. Phiri, M. Simwanda, S. Salekin, V. R. Nyirenda, Y. Murayama, and M. Ranagalage, "Sentinel-2 data for land cover/use mapping: A review," *Remote Sensing*, vol. 12, no. 14, p. 2291, 2020. 1
- [3] X. Yuan, D. J. Brady, and A. K. Katsaggelos, "Snapshot compressive imaging: Theory, algorithms, and applications," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 65–88, 2021. 1
- [4] T. Gelvez, H. Rueda, and H. Arguello, "Joint sparse and low rank recovery algorithm for compressive hyperspectral imaging," *Applied optics*, vol. 56, no. 24, pp. 6785–6795, 2017. 1
- [5] Y. Oiknine, D. Gedalin, I. August, D. G. Blumberg, S. R. Rotman, and A. Stern, "Target detection with compressive sensing hyperspectral images," in *Image and Signal Processing for Remote Sensing XXIII*, vol. 10427. SPIE, 2017, pp. 192–197. 1
- [6] C. J. Della Porta, A. A. Bekit, B. H. Lampe, and C.-I. Chang, "Hyperspectral image classification via compressive sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 10, pp. 8290–8303, 2019. 1
- [7] J. Monsalve, J. Ramirez, I. Esnaola, and H. Arguello, "Covariance estimation from compressive data partitions using a projected gradient-based algorithm," *IEEE Transactions on Image Processing*, vol. 31, pp. 4817–4827, 2022. 1, 2
- [8] C. Dechoz, V. Poulain, S. Massera, F. Languille, D. Greslou, F. De Lussy, A. Gaudel, C. L'Helguen, C. Picard, and T. Trémas, "Sentinel 2 global reference image," in *Image and Signal Processing for Remote Sensing XXI*, vol. 9643. International Society for Optics and Photonics, 2015, p. 96430A. 1
- [9] J. Monsalve, J. Ramirez, I. Esnaola, and H. Arguello, "Covariance Estimation from Compressive Data Partitions using a Projected Gradient-based Algorithm," *arXiv:2101.04027*, pp. 1–10, jan 2022. [Online]. Available: <http://arxiv.org/abs/2101.04027> 2
- [10] J. E. Fowler, "Compressive-projection principal component analysis," *IEEE Trans. on Image Processing*, vol. 18, no. 10, pp. 2230–2242, Oct. 2009. 2
- [11] X. Zhai, A. Oliver, A. Kolesnikov, and L. Beyer, "S4l: Self-supervised semi-supervised learning," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1476–1485. 3
- [12] D. Hendrycks, M. Mazeika, S. Kadavath, and D. Song, "Using self-supervised learning can improve model robustness and uncertainty," *Advances in neural information processing systems*, vol. 32, 2019. 3
- [13] K. Makantasis, K. Karantzas, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," vol. 2015-November. IEEE, 7 2015, pp. 4959–4962. [Online]. Available: <https://ieeexplore.ieee.org/document/7326945/> 3
- [14] Y. Tang, "Deep learning using linear support vector machines," 6 2013. [Online]. Available: <http://arxiv.org/abs/1306.0239> 3
- [15] M. F. Baumgardner, L. L. Biehl, and D. A. Landgrebe, "220 band aviris hyperspectral image data set: June 12, 1992 indian pine test site 3," Sep 2015. [Online]. Available: <https://purr.purdue.edu/publications/1947/1> 3
- [16] C. I. G. (GIC), "Hyperspectral remote sensing scenes." [Online]. Available: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes 3
- [17] C. Hinojosa, J. Bacca, and H. Arguello, "Coded aperture design for compressive spectral subspace clustering," *IEEE Journal on Selected Topics in Signal Processing*, vol. 12, pp. 1589–1600, 2018. 3