On Spatial Regularization for Semisupervised Hyperspectral Image Segmentation Using Hybrid Extreme Rotation Forest

Borja Ayerdi and Manuel Graña

Computational Intelligence Group
UPV-EHU, Spain

Outline

Outline

Introduction

- The generation of thematic maps from hyperspectral images by
  - classification of the pixel spectra.
- Scarcity of labeled information
  - semi-supervised training
- Combining both spatial and spectral processing.

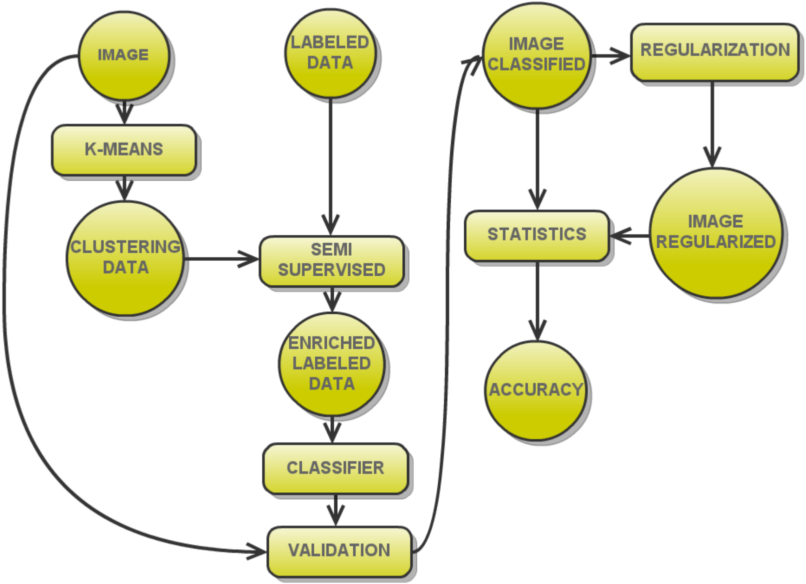
Contributions

We propose:

1. Spectra classification.
   1. Hybrid Extreme Rotation Forest (HERF)
2. A semisupervised training,
   1. k-means clustering and image spatial neighborhood.
3. Spatial regularization
   1. most frequent class in the neighborhood.

Outline

General Pipeline



- Heterogenous ensemble of classifiers
  - Extreme Learning Machines (ELM)
  - Decision Trees
- Partial adaptation to the problem domain

Hybrid Extreme Rotation Forest

For $i = 1 \dots L$

Computation of rotation matrix R_i^{α}:

- Partition F into K random subsets: $F_{i,j}; j = 1 \dots K$
  - For $j = 1 \dots K$
    - Let $X_{i,j}$ be the data set X for features in $F_{i,j}$.
    - $C_{i,j}$ obtained applying PCA on $X_{i,j}$
  - Compose $R_{i,j}^{\alpha}$ using matrices $C_{i,j}$.
- Decide if D_i is a DT or an ELM
- Train classifier D_i on training set (XR_i^{α}, Y).

Classification Phase

Decision by majority voting

For a given $\mathbf{x}^{test}$,

$$d_i = D_i(x^{test} R_i^{\alpha})$$

$$c^{test} = \max_i \{d_i, i = 1, \ldots, L\}$$

Outline

Semisupervised classification

input $X_L = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_L, y_L)\}$

1. initial HERF classifier $C_L : \mathbb{R}^d \rightarrow \Omega$, $C(\mathbf{x}_i) = \hat{y}_i$.
2. K-means: k_i the cluster assigned to sample $\mathbf{x}_i$.
3. $\mathcal{N}_j(r)$ spatial neighborhood of $\mathbf{x}_j \in X_L$ of radius r

   extended training set $X_{L+U} = X_L \cup X_U$

$$X_U = \{(\mathbf{x}_i, y_j) \,|\mathbf{x}_i \in \mathcal{N}_j(r) \wedge k_i = k_j \text{for some}\, \mathbf{x}_j \in X_L\}.$$

4. semisupervised classifier $C_{L+U} : \mathbb{R}^d \rightarrow \Omega$
5. classify whole image: $\hat{Y} = \{\hat{y}_i = C_{L+U}(\mathbf{x}_i)\}_{i=1}^{N}$.

Spatial regularization

1. most frequent class inside the spatial neighborhood of each pixel:

$$\tilde{y}_i = \arg\max_{y} \left| \left\{ \hat{y}_j \in \mathcal{N}_i \left(r \right) \right\} \right| .$$

Outline

Experimental Design

Real hyperspectral image data sets collected by AVIRIS sensor.

- Indian Pines -> 145×145 pixels, 224 spectral bands and 16 classes.
- Salinas C -> 217×512 pixels, 224 spectral bands and 16 classes.
- Salinas A -> 83×86 pixels, 224 spectral bands and 6 classes.

Comparative results

- Multinomial Logistics Regression (MLR) with active learning[1]
- We use the same size of the seed training set and validation by 100 Markov runs

[1] Jun Li, J.M. Bioucas-Dias, and A. Plaza, "Semisuper- vised hyperspectral image segmentation using multino- mial logistic regression with active learning," IEEE Transactions on Geoscience and Remote Sensing, vol. 48, no. 11, pp. 4085 –4098, Nov. 2010.

Numerical results - Salinas

Table : Results on the Salinas A data set at each step of the algorithm and corresponding results in the comparing publication.

SALINAS A	Our Method	MLR
Classification (L=18)	48.90 (18.00)	-
Classification [(L=18) + U]	95.1 (2.41)	90.86
Segmentation [(L=18) + U]	**99.13** (1.26)	**96.74**

Table : Results on the Salinas C data set at each step of the algorithm and corresponding results in the comparing publication.

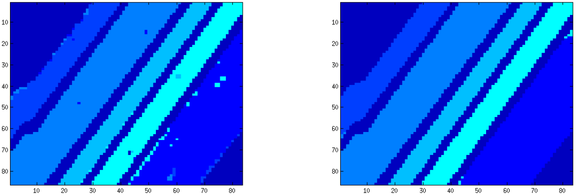
SALINAS C	Our Method	MLR
Classification (L=128)	81.18 (2.35)	81.97
Classification [(L=128) + U]	86.64 (1.30)	82.40
Segmentation [(L=128) + U]	**93.34** (1.58)	**89.61**

Numerical results - Indian Pines

Table : Results on the Indian Pines data set at each step of the algorithm and corresponding results in the comparing publication.

INDIAN PINES	Our Method	MLR
Classification (L=160)	51.96 (4.90)	63.19
Classification [(L=160) + U]	66.78 (3.03)	63.44
Segmentation [(L=160) + U]	**79.38** (4.04)	**75.60**

Visual results - Salinas A



(a) (b)

Figure : Visualization of classification results on Salinas A using 18 labeled samples. (a) After supervised classification with OA=97.58%. (b) After spatial regularization with OA=99.78%.

Visual results - Salinas C

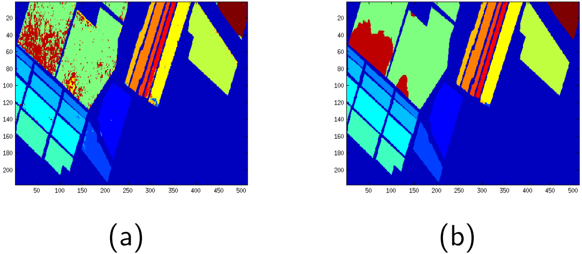


Figure : Visualization of classification results on Salinas C using 128 labeled samples. (a) After supervised classification with OA=88.22%. (b) After spatial regularization with OA=91.80%.

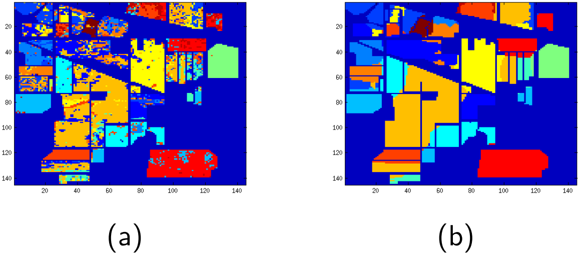


Figure : Visualization of classification results on Indian Pines. using 160 labeled samples. (a) After supervised classification with OA=66.03%. (b) After spatial regularization with OA=78.46%.

Outline

Conclusion

- New semisupervised approach involving
  - semisupervised training based on spectral clustering and spatial neighborhood
  - **two forms of spatial regularization**,
    - Selection of unlabled samples
    - Regularization over the final image segmentation
  - an innovative hybrid ensemble classifier **HERF**.

Computationally inexpensive

- classifiers used have **quick** learning algorithms and
- the regularization processes are computationally cheap.

Questions?

Thank you for your attention.

www.ehu.es/ccwintco