Lab session 2: Land feature extraction based on satellite images

Félix Jiménez Monzón and Javier Carnerero Cano
Univesidad Carlos III de Madrid
Avenida de la Universidad 30, 28911, Leganés (Madrid), Spain
{100290742, 100292885}@alumnos.uc3m.es

Abstract

In this work, we evaluate the performance of different feature extraction strategies to determine wich ones provides better performance to classify the crop type (16 classes) of a set of pixels in a multiband spectral image. Such spectral image was taken by sensor type Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) on board a satellite and it covers a region of Indiana, United States. We use different machine learning strategies, including linear and non-linear feature extraction techniques. From the obtained results, we could affirm that the Principal Component Analysis (PCA) and the Kernel Canonical Correlation Analysis (KCCA) methods achieve a better performance than the other linear and non-linear methods, respectively.

1 Introduction

Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) [1] is an optic sensor, unique in the world, that allows the obtaining of images in the emitted spectral radiation. It is able to measure the molecular absorption and particle scattering signatures in 224 adjacent spectral bands.

The main objective of the AVIRIS project is to identify, measure and monitor the Earth's surface and atmosphere in order to understand, monitor and predict the processes related to the global environment and climate change. In that sense, a spectral image from some crop fields from a region of Indiana (USA) is going to be used. This image was taken with the AVIRIS sensor on board a satellite, and 220 spectral bands have been used.

In the previous part of this work we evaluated different classifiers in order to see the performances of the same. Once we have selected the classifier with the best results from the previous work, different types of feature extraction techniques will be used to determine which ones provide better performance in this setting.

2 Objective

The objective of this project is to carry out a comparative study of the performance of several linear, non-linear, supervised and non-supervised feature extraction techniques. With this features, we will apply a Gaussian-kernel SVM, since that classifier provided us the best performance in our previous work. That classifier is able to classify and associate the pixels from a spectral picture, taken with a satellite, of various types of crop fields. The input data are spectral images of 145×145 pixels, which represent the molecular absorption and particle scattering signatures in 220 contiguous spectral bands. Each pixel in the image represents an area of 20 m^2. The classifier has been trained so that it can associate the 220 spectral bands of each pixel with a type of crop field (16 classes). Now in this work, we are going to try to reduce the dimensionality of the input data (220 dimensions) minimizing the loss of relevant information. We will take the 20% of the labelled image pixels for training.

First of all, we normalize the pre-processed data[1] to zero mean and unit standard deviation in each of the input variables. Then, in order to carry out this project, several linear, non-linear, supervised and non-supervised feature extraction techniques will be used. First, in terms of linear MVA methods, the performance of the Principal Component Analysis (PCA), Partial Least Squares (PLS), Canonical Correlation Analysis (CCA), and Linear Discriminant Analysis (LDA) algorithms will be studied. After that, non-linear feature extraction techniques based on kernel methods such as Kernel Principal Component Analysis (KPCA), Kernel Partial Least Squares (KPLS), and Kernel Canonical Correlation Analysis (KCCA) will be analyzed. Finally, a comparative analysis of the results will be made.

3 Linear MVA methods

Four kinds of linear MVA methods have been studied. These are: PCA, PLS, CCA, and LDA algorithms.

First of all, we would like to note that we have used the Gaussian-kernel SVM classifier in all the cases after applying the corresponding feature extraction technique. We have used the optimum parameters we obtained in our previous work for that classifier. They are: $C = 215.44$ and $\gamma = 8.84 \cdot 10^{-3}$. And the accuracy obtained is **85.01%**. We are aware that we should have validated those parameters for each new case, since the number of variables is different, but the computational cost for that is very high and it requires a huge amount of time. For this reason, we have used those parameters in all the cases.

3.1 PCA

Firstly, we carry out the classification applying in first place the PCA algorithm, which is an unsupervised algorithm that finds projections maximizing the variance of the projected data. We have applied the algorithm with a maximum number of features equals to 100. In order to find the optimum number of features, a validation process has been done, validating the performance of the algorithm (from 1 feature to 100). This evolution can be seen in Figure 1(a). We can see that the optimum number of features is **75**, and the accuracy obtained using that number is **84.34%**.

3.2 PLS

The next algorithm we have evaluated is the PLS algorithm, whose objective is to find the projections of the input and output data with maximum covariance. We have set the maximum number of features equals to the number of classes (16) and we have done a validation process in order to determine the optimum number of features. These results can be seen in Figure 1(b). According to those results, it can be seen that the optimum number of features is **13**, and the accuracy obtained is **83.58%**.

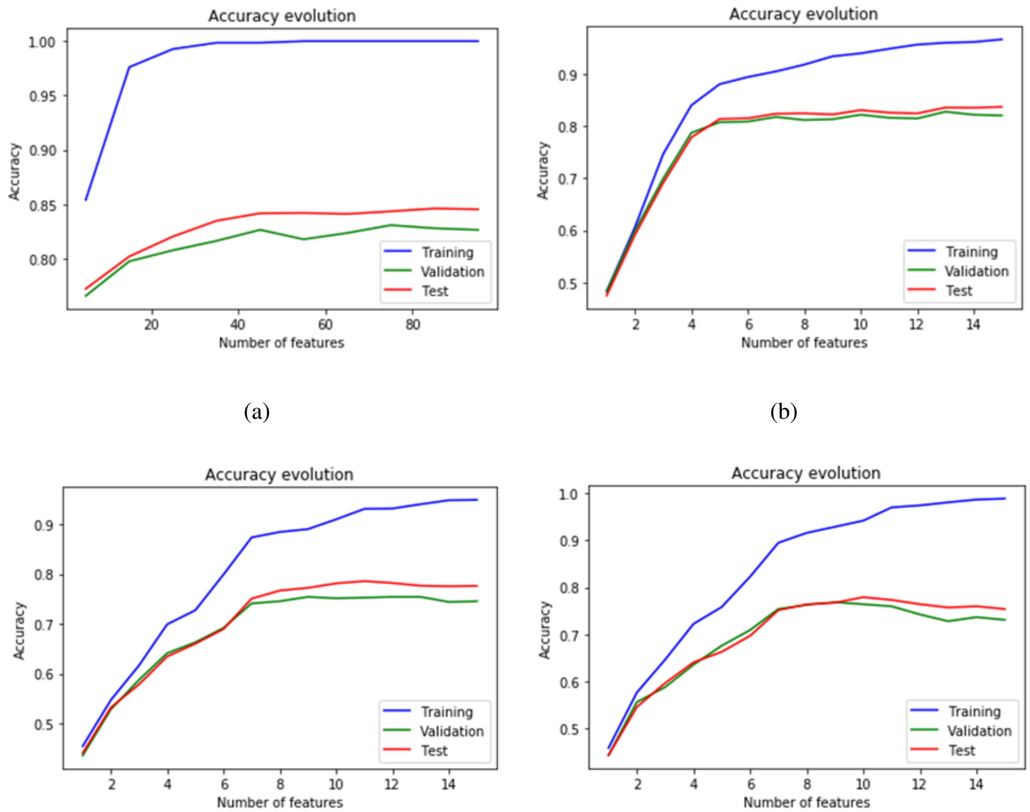
3.3 CCA

Thirdly, we have studied the performance of the CCA algorithm, which objective is to find the directions of maximum correlation between input and output data. In this case, we have set the maximum number of features equal to the number of classes minus one (15) and the *reg* parameter equal to 10^{-2}. Like in the previous cases, we have done a validation process to determine the optimum number of features. It can be seen in Figure 1(c). We can observe that the optimum number of features is **11** and the accuracy is **78.56%**.

3.4 LDA

Finally, the performance of the LDA algorithm has been evaluated. This algorithm considers that the data follow a Gaussian distribution with the same covariance matrix. Again, we have set the maximum number of features equals to the number of classes minus one (15) and we have determined the optimum number of features carrying out a validation process, which can be seen in Figure 1(d). We can see that the optimum number of features obtained is **10**, and the accuracy is **77.91%**.

[1] After eliminating the pixels corresponding to the image background, and the 20 highly noisy bands that cover water absorption region.



(a) (b)

(c) (d)

Figure 1: Accuracy evolution for the different linear feature extraction techniques. (a) PCA. (b) PLS. (c) CCA. (d) LDA.

3.5 Comparison

Table 1: Number of features and test accuracy for each algorithm.

Algorithm	Number of features	Test accuracy (%)
Baseline	220	85.01
PCA	75	84.34
PLS	13	83.58
CCA	11	78.45
LDA	10	77.91

We can observe that as the number of features decreases, the test accuracy also decreases. This is because of the information loss due to the reduction of the number of features. However, the fewer features we have, the less computational cost and complexity we obtain, so we have to make a trade-off between the number of features and the test accuracy.

4 Non-linear classifiers

Three types of non-linear (or kernel) techniques have been studied. They are based on the projection of the data into a high dimensional space, so that a linear algorithm runs in the "featured space" is non-linear in the input space. These methods are: KPCA, KPLS, and KCCA.

4.1 KPCA

First, we carry out the classification applying in first place the KPCA algorithm, which is an unsupervised algorithm that finds projections maximizing the variance of the data in the feature space. We have applied the algorithm with a maximum number of features equals to 100. The accuracy obtained using that number is **77.20%**.

4.2 KPLS

The next algorithm we have evaluated is the KPLS algorithm. We have set the maximum number of features equals to the number of classes (16). It can be seen that the the accuracy obtained is **69.31%**.

4.3 KCCA

In this case, we have set the maximum number of features equal to the number of classes minus one (15) and the *reg* parameter equal to 10^{-2}. The accuracy is **83.72%**.

4.4 Comparison

Table 2: Number of features and test accuracy for each algorithm.

Algorithm	Test accuracy (%)
Baseline	85.01
KPCA	77.20
KPLS	69.31
KCCA	83.72

We can see the best performance is achieved by the KCCA method.

5 Conclusions

In this project we have studied the performance of several linear and non-linear feature extraction techniques in order to see which one achieves the better result. The problem consisted in cataloguing the different types of crops (16 classes) from a spectral image taken with the AVIRIS sensor on board a satellite.

The PCA and KCCA methods achieve a better performance than the other linear and non-linear methods, respectively.

However, the accuracies obtained from all the feature extraction techniques are very similar, and these results correspond to a fix division of the train and test samples. If we want to see technique has the best performance, we have to repeat the process several times in order to calculate the average and the standard deviation of the accuracy of the classifier. We have not done this because of the extremely long computational time needed. So, only with these results we cannot assure that a technique is better than the others.

References

[1] NASA, "AVIRIS (Airborne Visible / Infrared Imaging Spectrometer)." Available: "https://aviris.jpl.nasa.gov/", 2017. [Online]. [Accessed February 28, 2017].

[2] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2007.

[3] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. Wiley & Sons, 2nd Edition, 2000.