



Título artículo / Títol article: Spectral–Spatial Pixel Characterization Using Gabor Filters for Hyperspectral Image Classification

Autores / Autors Olga Rajadell, Pedro García-Sevilla, and Filiberto Pla

Revista: Geoscience and Remote Sensing Letters, IEEE

Versión / Versió: Versió pre-print

Cita bibliográfica / Cita bibliogràfica (ISO 690): RAJADELL, Olga; GARCÍA-SEVILLA, Pedro; PLA, Filiberto. Spectral–Spatial Pixel Characterization Using Gabor Filters for Hyperspectral Image Classification. *Geoscience and Remote Sensing Letters, IEEE*, 2013, vol. 10, no 4, p. 860 - 864.

url Repositori UJI: <http://hdl.handle.net/10234/88009>

Spectral-spatial pixel characterization using Gabor filters for hyperspectral image classification

Olga Rajadell, Pedro García-Sevilla, and Filiberto Pla

Abstract—This paper presents a spectral-spatial pixel characterization method for hyperspectral images. The characterization is based on textural features obtained using Gabor filters over a selected set of spectral bands. This scheme aims at improving land-use classification results decreasing significantly the number of spectral bands needed in order to reduce the dimensionality of the task thanks to an adequate description of the spatial characteristics of the image. This allows to require less data and avoid the curse of dimensionality. Very promising results are obtained which are similar or better than previous classification results provided by other spectral-spatial methods, but here also reducing the complexity using a reduced number of spectral bands.

I. INTRODUCTION

Current hyperspectral sensors can have high spectral and spatial resolution. Some sensors can cover spectral resolutions higher than 10nm, reaching 1m per pixel for spatial resolution (e.g. some images provided by ROSIS sensor). As a result, hyperspectral images are composed of a high number of correlated bands that may cause a dimensionality problem. When the spatial resolution was not so high, main efforts were focussed at the classification stage. In particular, Support Vector Machines (SVM) proved to obtain good performances in this task [1]. With the increase of the spatial resolution, a joint spectral-spatial analysis was identified as a desired goal [2]. Spectral-spatial characterization aims at obtaining one feature vector for each pixel in the image based on the spectral measurements (spectral information) and a series of values extracted from spatial operations involving neighboring pixels (spatial information). However, as in all classification problems, it should not be forgotten that increasing the number of features used does not provide an endless improvement because of the well-known curse of dimensionality problem [3].

Nowadays, a wide range of techniques is used to include spatial information into the image characterization, such as morphological profiles [4] or Markov fields [5]. However, these methods introduce a scale selection problem. Recently, several proposals have risen to face the over-segmentation problem and the scalability with very good results. Tarabalka et al. [6] presented a spectral-spatial classification scheme that consists of a pixel wise classification and a partitional clustering by a majority vote with adaptive neighborhoods. The result is a segmentation map that needs a spatial post regularization to reduce the noise. This provides more homogeneous regions than a simple pixel wise classification process but it is not yet suitable for images containing small classes since they may

be missed. The same problem is observed in [7] where an extension of the watershed segmentation algorithm for hyperspectral images was presented in order to define the spatial structures. To deal with the segmentation of small regions, the same authors suggested in [8] to select the most reliable pixels from a pixel wise classification as markers to be used in a Minimum Spanning Forest Grown obtaining a spectral-spatial classification map refined afterwards by majority voting within the spatially connected regions.

The characterization of spatial structures in an image has been studied in detail when dealing with the analysis of visual textures [9]. However, most of these methods were developed mainly for grey level images and their extension for multi-channel images has been generally faced as a multi-dimensional extension of the mono-channel techniques. Healey et al. [10] made one of the first proposals on how to use spatial information across spectral bands using Gabor filters. Opponent features were first described for color images [10] and extended to be used over multi-channel images [11]. They combine spatial information across spectral bands at different scales by combining the responses of the filters applied separately to each channel. Lately they also used three-dimensional Gabor filter banks [12]. However, all these methods have been always applied to patches of stationary textures and no analysis has been done about the characterization of individual pixels which allows segmentation of images using this spatial information.

In order to segment and classify hyperspectral land cover images, we classify individual pixels to get a classification map following a lately trend [8] [6] [5]. This task has already been faced using a large amount of data. However, when devices improve, dealing with an increasing amount of data also increases the risk of reaching the accuracy ceiling. Thus, we also aim at using a very small number of features to obtain the same or even better results found in literature leaving then room for adding new features that may improve the classification. To pursue this objective, a band selection method will be first used over the whole set of bands provided by the spectrometer. Then the pixel characterization methods will be applied over the selected spectral bands. Three different pixel characterization methods based on Gabor filters will be used here.

The rest of the paper is organized as follows. First, filter bank characterization methods are introduced in Section II, Gabor filters have been used for texture characterization and we will propose their use for pixel classification in two different ways. Besides we will adapt a method from the literature to perform pixel classification and compare all of

them. In Section III, we present the database used in this paper followed by the classification set up that will be further used, a comparison between the three characterization methods and a study of the relation between the characterization and the scales of the filters within the filter bank. The supervised segmentation results are presented in Section IV as images, also providing the per class accuracies. Eventually, conclusions can be found in Section V.

II. FILTER BANK CHARACTERIZATION METHODS

Several features have been suggested in the literature for the description of spatial (or texture) information (see [9] for a survey). In this paper, features are obtained by filtering the input image with a set of filters (filter bank). The vector of features per each pixel corresponds to all the responses of the pixel to the filter bank.

For an image of B bands, let I^i be the i^{th} band. Let f_k be the k^{th} filter in the filter bank F . The response to the filter when applied over the i^{th} band is given by $h_k^i = I^i * f_k$, where $*$ stands for the convolution operator.

We chose to use a Gabor filter bank. This is a set of Gabor filters of M different scales (spatial frequencies) and N orientations designed to cover the frequency domain:

$$F = \{f_{m,n}\}_{m=1,n=1}^{M,N} \quad (1)$$

They consist of sine and cosine functions modulated by a Gaussian envelope that achieve optimal joint localization in space and frequency [13]. They can be defined by:

$$f_{mn}^{real}(x, y) = \frac{1}{2\pi\sigma_m^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma_m^2}\right\} \times \cos(2\pi(u_m x \cos \theta_n + u_m y \sin \theta_n)) \quad (2)$$

$$f_{mn}^{imag}(x, y) = \frac{1}{2\pi\sigma_m^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma_m^2}\right\} \times \sin(2\pi(u_m x \cos \theta_n + u_m y \sin \theta_n)) \quad (3)$$

where m is the index for the scale, n for the orientation and u_m is the central frequency of the scale [14].

For real signal values, the outputs for orientations θ_n and $\theta_n + \pi$ will be complex conjugates. These pair of filters are usually joined together, cancelling in this way the imaginary parts of the outputs and dealing only with real value outputs.

A. Opponent features

Opponent features [10] use Gabor filters and combine the filtered results (spatial information) across spectral bands at different scales. According to the authors, this is related to processes in human vision. They are computed from the responses to Gabor filters as the difference of responses between two different filters. In other words, the spectral bands are first individually filtered and their responses are combined afterwards to obtain the opponent features aiming at introducing inter-channel information into the characterization process. The combination among responses [11] is made for all pair of spectral bands i, j , with $i > j$, and two scales m and m' , such that $0 \leq (m - m') \leq 1$, as follows:

$$d_{mm'n}^{ij} = h_{mn}^i - h_{m'n}^j \quad (4)$$

In our case, instead of computing the energies for whole image patches, a feature vector for each individual pixel is obtained as the set of all opponent features computed for it. In this way we obtain opponent features for each individual pixel applying the filter bank only once over the whole image. If a texture patch was considered around each pixel in the image, the filter bank must be applied over each patch. As each pixel will belong to several patches, it will be repeatedly analyzed. In this way, we expect to obtain similar results but reducing the computational effort required.

B. Gabor filters over individual bands

We propose a simplified version where each spectral band is analyzed separately and each pixel is characterized with the responses to the filter bank used. This will result in a smaller number of features per pixel keeping the spatial information but missing the inter-channel information.

When the whole filter bank is applied, the feature vectors for the pixels in the image will be obtained by simply taking the responses to all filters for all bands:

$$\psi = \{h_k^i\}_{k=1,i=1}^{K,B} \quad (5)$$

In this way, hyperspectral images will be simply decomposed into separated bands and the same feature extraction process will be performed over each band. By filtering with such a filter bank, the response of one pixel to each filter is a decomposition of the spectral measurement in the amounts corresponding to a each spatial frequency range and orientation used to define the filter bank.

C. Gabor filters over complex bands

Filtering each band individually misses the inter-channel information proposed by Healey et al. [10]. In order to test its significance, we propose a variation of the characterization method described above that introduces inter-channel data. To pursue this, two real bands are merged into one complex band, one as the real part and the other one as the imaginary part. Now, each Gabor filter will be applied over a complex band as follows:

$$h_{mn}^{ij} = (I^i + I^j \mathbf{i}) * f_{m,n} \quad (6)$$

with $\mathbf{i} = \sqrt{-1}$ where I^i and I^j are the i^{th} and j^{th} spectral bands respectively.

All pairs of spectral bands will be considered and filtered. Inter-channel information is included here because two spectral bands are filtered at once so the response to the filters combines information from these two bands. Since the bands to be analyzed are no longer real, now filters for orientations θ_n and $\theta_n + \pi$ will provide different outputs and, therefore, they are not joined together. As a consequence, the number of orientations will be the double of the number used for individual bands. Note also that now each output will be a complex number that will be represented using two real values, while in previous cases each output was represented by just one real number. This doubles the number of features required in this case.

III. CLASSIFICATION EXPERIMENTS

The scheme here proposed combines a band selection method with the spectral-spatial pixel characterization methods previously proposed. Among the different band selection methods, WaLuMI [15] has been chosen for preserving the original bands, providing as output a subset of them. It is based on a clustering of bands that pursues, as a whole, to maximize the mutual information among bands in each cluster and to minimize the inter-cluster correlation. However, any other band selection method that fulfills similar criteria can be used instead. In this section all classification experiments are tested over the Indian Pine hyperspectral dataset (AVIRIS). Two classifiers, SVM with a third order polynomial kernel and a 3-nearest neighbour classifier, are used.

A. Dataset

Hyperspectral image 92AV3C was provided by the spectrometer AVIRIS and acquired over the Indian Pine Test Site in Northwestern Indiana in 1992. The image has a spatial dimension of 145×145 pixels. Spatial resolution is 20m per pixel. Spectral coverage ranges from 0.38 to $2.50\mu m$ with 220 spectral bands. Classes range from 20 to 2468 pixels. Due to the small size of some classes, this database is suitable for testing if the proposed methods can also success at the classification of small areas which are often missed in highly unbalanced datasets.

B. Experiment set up

For the characterization of the data, a Gabor filter bank is designed with four orientations and six scales. The four orientations (0° , 45° , 90° , 135°) are the minimum number of orientations recommended to get textural information [9]. Gabor filter scales are chosen to be dyadic, being the first scale of width 1. Hence, given the size of the image used, the maximum number of scales is $M = 6$. Besides, Gabor filters were designed to overlap each other when achieving a value of 0.5 following the recommendation in [16]. The filter bank is applied according to one of the methods defined in the previous section and each pixel is characterized with the responses to it. This leads to a different number of features per pixel regarding the method used (see Fig. 1). This is an important issue because of the so-called Hughes phenomenon [3], which also leads to the fact that increasing the number of dimensions does not necessary leads to an improvement.

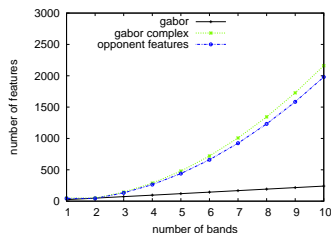


Fig. 1. Number of features per method vs. the number of spectral bands.

For the classification experiments, the labeled pixels in the image database were divided into twenty non-overlapping

sets keeping the a priori probability of each class. Therefore, no redundancies were introduced. Ten classification attempts were carried out and the mean of the error rates of these attempts was taken as the performance. For each attempt one set was used as training and another for testing and sets were never used twice. This methodology was already used in [15] and [17] in order to increase the statistical independence among the classification attempts.

C. Comparison of the characterization methods proposed

In this section, the different characterization methods described in section II are compared. The settings for these experiments are the ones described in Sect. III-B. The value of B (number of spectral bands) varies from one to ten in each experiment. The set of bands is provided by the WaLuMI algorithm.

The classification results using a SVM can be found in Fig. 2. The results using only spectral information were also included as a baseline reference. In all cases the mean rate of 10 experiments is shown. The variance between experiments was really small (less than 3%).

All spectral-spatial features clearly outperformed the spectral features. Also, we can see that there is almost no difference between the three spectral-spatial methods considered. Experiments using Gabor filters over texture patches around each pixel were also carried out providing similar classification rates. This means that the spatial information is much more important than the inter-channel information for the appropriate characterization of the pixels in the image. It is important to note that the initial information used in all experiments is exactly the same because the spatial features are directly computed from the spectral data.

Results with KNN are slightly lower than the ones obtained with SVM (an average of 2% lower) with a small difference of 1.5% in favour of Gabor filters over complex bands.

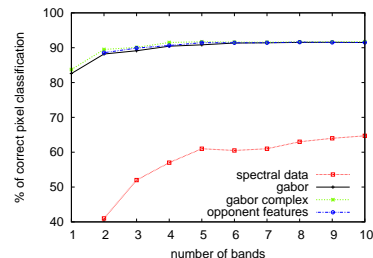


Fig. 2. Pixel classification rates for the proposed characterization methods over the AVIRIS database using a SVM classifier.

Thus, we can conclude that spatial information improves the classification but the addition of inter-channel information is not relevant enough and does not justify the increase in the dimension of the classification space. Considering this conclusion, for the next series of experiments we suggest the use of Gabor filters over individual bands.

For the number of bands, observe that after $B = 3$ no significant improvement is achieved when increasing the number of bands. An experiment using all bands available ($B = 220$) was performed with a result of 87% using a SVM

and 86.6% for the 3-NN, which are below the maximum result presented in Fig. 2. Although B is a parameter to set for the process, the performance usually reaches a maximum and adding more bands does not improve the classification results. Hence, the selection of B is not critical as long as we choose a value greater than the one needed to reach the flat zone of the learning curve.

D. Scale analysis

In a Gabor filter bank, those filters with different orientation and the same scale provide information corresponding to the same range of spatial frequencies. It is known that different frequencies provides a different analysis of the scenario, for example high frequencies contain most of the noise present in the image. The following experiment is a classification using solely the features obtained from each set of filters with the same scale but with different orientations. These results are presented in Fig. 3(left). The settings for the classification experiments are the same as in Sect. III-C, except for the scales used.

Note that the lower the scale is, the better the result. This was expected because most of the areas to classify are quite homogenous.

As a further analysis, we also run an experiment performing a progressive combination of features from different scales. First, only features using the filters that are defined with the first range of spatial frequencies are taken. In each step the features from the following scale are combined with the previous by adding one scale at each step until covering the whole set of scales. These classification results are shown in Fig. 3(right). Observe that when we join the features of the first two scales, the classification rate improves. When adding the first, second, and third scales, the results are quite similar. However, when adding more than three scales, the results progressively worsen. Recall that the higher scales may mainly contain noise and they do not help in the characterization of the pixels. This highlights the fact that the discriminant piece of information for this sort of images is in the first scales because they contain well-defined areas of low spatial frequencies.

IV. SEGMENTATION EXPERIMENTS

To get a supervised segmentation from the pixel wise classification we split our data in a set of samples with known labels and a test set to be classified. The resulting labels create a classification map. Unlike the previous experiments, the set of labeled pixels is here directly split in two. A 5% of samples from the whole data set, keeping the a priori probabilities, forms the training set and the rest the test set. Again results using a SVM classifier are shown.

The classification results in Fig. 2 showed that the improvement has a maximum. Because any value of B over 3 will provide a similar result, raising the number of features (dimensions) in this problem is not convenient. Hence, the set of bands for $B = 3$ were selected using the WaLuMi algorithm. Besides, as seen in Section III-D, adding features from more scales is neither improving the characterization. For that reason we chose to perform this experiment with the

features coming from $B = 3$ and combining the features from the two first scales. We chose to reduce the number of features to show that 24 features can provide a result as good or even better than a much higher number of dimensions.

The global classification accuracy obtained was 92.99% using a SVM (note that the result using a 3-NN classifier was the same). This result is slightly higher than the ones in [7] [8], where the same problem for the AVIRIS dataset was addressed, obtaining a 91.80% of correct classification. Besides, in these cited works, a fixed number of samples per class were picked as training set, thus the a priori probabilities were not kept and small classes were over-represented in the training and all spectral bands were used there. Therefore, the results presented here have been obtained in more realistic conditions, taking into account that real unbalanced data is a harder classification problem.

The producer's accuracy per class for the AVIRIS dataset is shown in Table I. Notice that 7 of the 16 classes are usually ignored in this sort of experiments because they contain a very small number of pixels [1]. However, we include them in our experiments and the results are fairly good considering the difficulties when treating with such unbalanced classification problem. For example, the class representing oats has only 20 pixels, and only one pixel was used for training. Therefore, an important amount of classification errors is expected. Nonetheless, it is remarkable that other small areas corresponding to Alfalfa, Bldg-Grass-Trees-Drives, Grass/Pasture-mowed, Corn, and Wheat were fairly well classified. We can also see the same results on the image in Fig. 4 where the errors are represented in white.

class	training/total	Per class accuracy (%)
Stone-steel towers	4/95	76.92
Hay-windrowed	25/489	99.14
Corn-min till	42/834	96.59
Soybeans-no till	48/968	89.23
Alfalfa	2/54	100.0
Soybeans-clean till	30/614	87.32
Grass/pasture	25/497	90.04
Woods	65/1294	95.77
Bldg-Grass-Tree-Drives	19/380	97.22
Grass/pasture-mowed	2/26	91.66
Corn	11/234	92.82
Oats	1/20	52.63
Corn-no till	72/1434	92.51
Soybeans-min till	124/2468	91.93
Grass/trees	37/747	94.92
Wheat	10/212	99.50
Overall accuracy		92.99
kappa		0.92

TABLE I
PRODUCER'S ACCURACY PER CLASS FOR THE AVIRIS DATASET USING 24 FEATURES (4 ORIENTATIONS, 2 SCALES, AND 3 SPECTRAL BANDS).

V. CONCLUSIONS

A hyperspectral pixel classification scheme that combines a band selection procedure with a spatial feature extraction process has been presented. The increase of the spatial resolution in hyperspectral sensors encouraged this idea. It has been experimentally proven that the proposed scheme provides

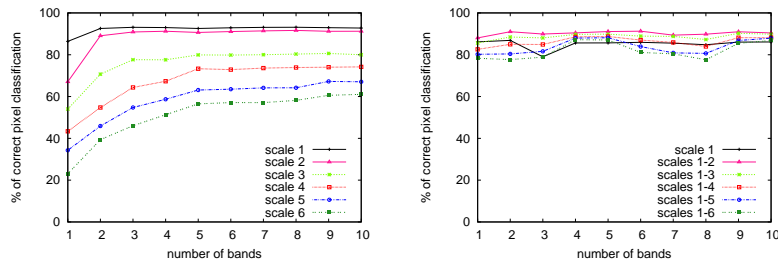


Fig. 3. Pixel classification rates for the AVIRIS dataset using spatial features derived from Gabor filters and a SVM classifier. Left: the analysis of individual scales. Right: joining of features from consecutive ascending scales.

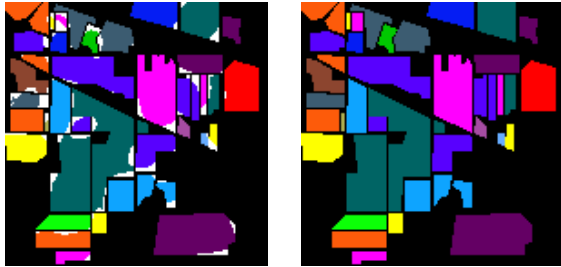


Fig. 4. Left: Classification results for the AVIRIS dataset using 24 features (4 orientations, 2 scales, and 3 spectral bands). Overall Accuracy: 84.836 Kappa coefficient: 0.791. Right: Groundtruth.

better classification rates than other state of the art spectral-spatial methods. Furthermore, the approach presented here uses a reduced set of selected spectral bands, simplifying the representation while keeping the classification rates with respect to other approaches. This is important in order to avoid the problems caused by the curse of dimensionality and also because it leaves room for other features to be used to improve the characterization.

Three spatial features have been suggested for the characterization of individual pixels, all of them based on features derived from Gabor filters. We have shown that the spatial information provides an appropriate characterization of the pixels for classification tasks. These features lead to good classification rates. We have also shown that the spatial information influences the characterization process much more than the inter-channel information. No big differences have been found between the three sort of spatial features analyzed although they have big differences in the number of features used to describe each pixel, being the method proposed by applying Gabor filters over individual bands the most appropriate because of its simplicity and smaller dimensionality.

We have also studied the influence of the different scales in the feature extraction process and found that, when only areas of low spatial frequencies compose the image, the first scales provide the best characterization and the addition of the last scales tends to worsen the classification results. However, if we have to deal with non-homogeneous regions, the use of the medium scales may improve the characterization.

In the segmentation experiments, we found that most of the miss-classified pixels fall in the borders of the labeled regions where the spatial features can be confused due to the background information or due to the transitions between different

classes in the image plane. However, the segmentation of the inner part of the regions was always remarkably homogeneous, despite the fact that no further spatial regularization is applied to the pixel-based classification proposed.

ACKNOWLEDGMENT

This work has been supported by FPI PREDOC/2007/20 from Fundació Bancaixa and projects CSD2007-00018 (Consolider Ingenio 2010) and AYA2008-05965-C04-04 from the Spanish Ministry of Science and Innovation.

REFERENCES

- [1] A. Plaza et al., "Recent advances in techniques for hyperspectral image proc." *Remote sensing of environment*, vol. 113, pp. 110–122, 2009.
- [2] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ: Wiley, 2003.
- [3] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. on IT*, vol. 14, pp. 55–63, 1968.
- [4] M. Fauvel et al., "Spectral and spatial classification of hyperspectral data using svms and morphological profiles," *IEEE TGRS*, vol. 46, no. 10, pp. 3804–3814, 2008.
- [5] J. Li, J. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE TGRS*, vol. 48, no. 11, pp. 4085 – 4098, 2010.
- [6] Y. Tarabalka, J. Chanussot, and J. Benediktsson, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE TGRS*, vol. 47, no. 8, pp. 2973–2987, 2009.
- [7] —, "Segmentation and classification of hyperspectral images using watershed transf." *Pattern Rec.*, vol. 43, no. 7, pp. 2367–2379, 2010.
- [8] —, "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers," *IEEE Trans. on SMC*, vol. 40, no. 5, pp. 1267–1279, 2010.
- [9] M. Petrou and P. García-Sevilla, *Image Processing: Dealing with Texture*. John-Wiley and Sons, 2006.
- [10] A. Jaim and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE TIP*, vol. 7, pp. 124–128, 1998.
- [11] M. Shi and G. Healey, "Hyperspectral texture recognition using a multiscale opponent representation," *IEEE TGRS*, vol. 41, pp. 1090–1095, 2003.
- [12] T. Bau, S. Sarkar, and G. Healey, "Hyperspectral region classification using three-dimensional gabor filterbank," *IEEE TGRS*, vol. 48, no. 9, pp. 3457–441, 2010.
- [13] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological Cybernetics*, vol. 61, pp. 103–113, 1989.
- [14] A. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern Recognition*, vol. 24, pp. 1167–1186, 1991.
- [15] A. Martínez-Usó, F. Pla, J. M. Sotoca, and P. García-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE TGRS*, vol. 45, pp. 4158–4171, 2007.
- [16] F. Bianconi and A. Fernández, "Evaluation of the effects of gabor filter parametres on texture classification," *Pattern Recognition*, vol. 40, pp. 3325–3335, 2007.
- [17] J. S. Sánchez, R. A. Mollineda, and J. M. Sotoca, "An analysis of how training data complexity affects the nearest neighbor classifiers," *Pattern Analysis and Applications*, vol. 10, pp. 189 – 201, 2007.