Decision Tree Classification System for Brain Cancer Detection using Spectrographic Samples

Inmaculada Dópido, Cassandra Déniz, Himar Fabelo, Gustavo Callicó, Sebastian López, Roberto Sarmiento Institute for Applied Microelectronics

University of Las Palmas de Gran Canaria, Spain {idopido,cdeniz,hfabelo,gustavo,seblopez,roberto}@ulpgc.iuma.es

Abstract— Hyperspectral imaging is an active research field for remote sensing applications. These images provide a lot of information about the characteristics of the materials due to the high spectral resolution. This work is focused in the use of this kind of information to detect tumour tissue, particularly brain cancer tissue. In recent years, the study of this kind of tumour has been a challenging task due to the nature of these tissues. The neurosurgeon usually finds several problems to detect tumour tissues by the naked eye. In order to address this problem, this work makes use of high spectral resolution samples in the range from 400 nm to 6000 nm, provided by an Agilent Resolutions Pro V.5 spectrometer that has been diagnosed by histopathology. This instrument can sample a single pixel with a very high spectral resolution. The high spectral resolution allows a reliable separation between the different tissues in brain tumour. The proposed approach is based on a hierarchical decision tree. This approach is composed by several systems of Support Vector Machine classifiers. The 225 used samples come from 25 adults (males and females) and have been taken at different surgical procedures at the University Hospital of Southampton. The main goal is to discriminate between tumour tissue and normal tissue. Specifically, it assigns priority to the group of classes known a priori to the classification showed accordingly to the level of detail. The experimental results indicate that the use of the proposed new decision tree approach could be a solution to effectively discriminate between tumour and normal tissue and additionally provide information about the specific tissue for these classes. For our data set, a sensitivity of 100% and a specificity of 99.27% have been obtained when healthy and tumour samples are discriminated. These results clearly indicate that the use of high dimensionality spectral data is a promising and effective technique to indicate if a brain sample is or not affected by cancer with a high reliability.

Keywords— Support Vector Machine; Hyperspectral Imaging, Brain Cancer Detection

I. Introduction

Gliomas are the commonest primary tumours of the central nervous system. Gliomas are graded based on their histological appearance, into grades I-IV. Grades I and II are considered 'low-grade', and are managed differently to grades III and IV ('high-grade'), both operatively and post-operatively. The pre- or intra-operative distinction between high and low-grade tumours is therefore important. Gliomas are further termed based on their predominant cell type often astrocytomas or oligodendrogliomas. Normal brain cortex is formed of grey and white matter. White matter is

Diederik Bulters, Elizabeth Casselden, Harry Bulstrode

University Hospital Southampton United Kingdom dbulters@nhs.net, lizziecasselden@gmail.com, hjcjb@mac.com

so-called due to the presence of myelin - it will therefore have significant molecular differences based on lipid content and as such the distinction is reassuring to ongoing investigation. High grade lesions are incurable and in general rapidly fatal despite best medical and surgical management (13 months median survival [1]). There is growing evidence that life expectancy increases with a more extensive resection; a gross total resection of high-grade gliomas is associated with a three month longer survival than subtotal resection [2]. With low-grade gliomas that have a longer survival time, the time to tumour progression and malignant transformation are key, and are thought to be delayed with a more extensive resection [3]–[5].

Gross total resection (GTR) is usually therefore the goal of surgery, but achieving it can be challenging due to the infiltrative nature of gliomas. Confidence that the resection is not straying into adjacent normal brain depends on accurate assessment of tumour borders despite this diffuse infiltration. This is a judgement made throughout the operation on the basis of anatomical considerations and assessment of resected tissue; a histologically-complete resection is thought to be achieved in fewer than 20 % [6]. If cure is therefore unfeasible, then quality of life becomes paramount. It is therefore important not to damage areas of eloquent cortex through over-zealous resection. With this in mind, many patients undergo a subtotal resection. There are therefore several important distinctions to be made:

- 1) Normal vs tumour tissue
- 2) High-grade vs low-grade glioma
- 3) White vs grey matter

These are reflected in our decision tree approach as explained below.

Surgical adjuncts used to guide resection include image guidance systems, but these suffer limitations related to calibration, intra-operative brain shift, and the lack of 'real-time' information. Spectroscopic techniques promise to assist distinction of tumour from normal brain tissue, at high resolution and in real-time. This promise can be realized only through development.

For this reason, hyperspectral imaging seems to be a good alternative for the detection of cancer [7]–[9]. The main characteristic of hyperspectral images is the large amount of spectral information (hundreds of narrow bands) and so they

can be considered as an extension of the concept of digital image [10]. It allows many different applications such as classification, spectral unmixing, target detection, etc. In order to detect cancer tissue in the spectrographic samples classification it is necessary to separate materials into spectrally similar groups. Hyperspectral image classification is a very active area of research in recent years in which the main objective is to assign a unique label to each pixel [8], [11]. Several techniques have been successfully used to perform hyperspectral data classification, particularly supervised techniques such as kernel methods, which can deal effectively with the Hughes phenomenon [12], [13]. It is important to emphasize that the analysis of hyperspectral images is not an easy task; this is due to the great variability of hyperspectral signatures and the high dimensionality of the data. In this work, the main goal is focused on brain cancer detection using supervised classification techniques. Given the problems to detect brain tumour tissues our motivation is focused on exploiting the information that offers the spectrographic samples (very high spectral resolution), it can be a great opportunity to develop a new and efficient technique able to detect tumour tissue using limited training samples [14].

The remainder of this paper is organized as follows. Section II presents an introduction about the existing technology and the process of obtaining medical samples. Section III provides our experimental setting, with emphasis on describing the different hierarchies and the classification systems in order to validate the presented technique. Section IV concludes with some remarks and hints at plausible future research lines.

II. METHODS

This section is organized as follows. Subsection II-A introduces the instrument and the medical samples used for validating the proposed approach. Subsection II-B describes the data pre-processing applied to the medical samples to acquire their full spectral data using the spectrometer. Finally, subsection II-C describes the new proposed approach developed to detect brain tumour tissues.

A. Spectrographic System and Medical Samples

The set of samples used in these experiments was collected by a spectrometer. For this purpose, we used an Agilent Resolutions Pro V.5 spectrometer which has been pioneered in areas such as biotechnology, food, agriculture and security. The 600-IR spectrometers are especially suitable for research applications or method development. The samples (single pixels) obtained by this spectrometer are absorptions in the range from 400 nm to 6000 nm, comprising 2906 bands, with a spectral resolution of 1.92 nm. They come from 25 different adults (males and females) and have been taken at different surgical procedures. In Table I, the samples used to perform these experiments are shown. A total of 213 samples have been used, including 198 tumours and 15 normal samples. The shortage of normal brain samples has been one of the main difficulties encountered during the development of the classification algorithm, due to the difficulty in obtaining normal brain samples. Fig. 1 shows several spectrographic samples, which comprise tumour samples and healthy samples.

B. Data Pre-processing

Samples were collected from appropriately consented patients undergoing craniotomy for suspected glioma. Tissue was identified as tumour by one of two surgeons, under image-guidance technology. En bloc resection specimens were collected, along with tissue derived from the ultrasonic aspirator system, which is used extensively in developing a plane around the tumour, and debulking the tumour mass. Samples were washed in sterile medical-grade water 0.9 % saline to remove visible blood traces where appropriate. They were then weighed and air-dried at approximately 41.7 until they reached a consistency compatible with grinding, and once dry were ground into a homogenous powder with KBr powder, using a pestle and mortar. The resulting mixture was then fractionated into 0.5 g portions and pressed at ~ 10 tonnes in a pellet press with a vacuum facility, to create a solid pellet suitable for mounting in the spectrometer. Samples were stored individually at -20 °C. with a sachet of silica gel in order to reduce water absorption by KBr.

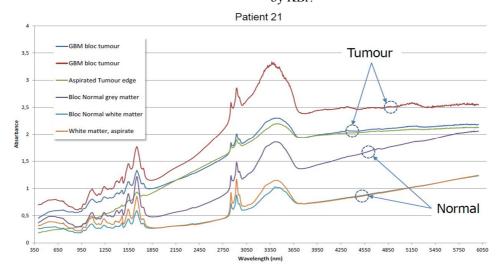


Fig. 1. Spectrographic samples associated to the Patient 21

TABLE I. SPECTROGRAPHIC MEDICAL SAMPLES USED IN THIS SET OF EXPERIMENTS RELATED WITH THE DIFFERENT CHARACTERISTICS OF THE AVAILABLE

Medical Samples								
			Sample Diagnosis	Number of Samples				
Tumour	Low Grade	Grade II	II Astro II Oligo	22				
	High Grade	Grade III	III Astro III Oligo GBM mixed astro/oligo	30				
		Grade IV	GBM astro IV gliosarcoma GBM	146				
Normal	Normal Grey		Normal grey matter	6				
	Norn	nal White	Normal white matter	9				

C. SVM Hierarchical Tree Approach

The rich spectral information available in remotely sensed hyperspectral images allows for the possibility to distinguish between spectrally similar materials [15]. For this reason, supervised classification techniques in hyperspectral images is a very challenging task due to the generally unfavourable ratio between the (large) number of spectral bands and the (limited) number of training samples available a priori, which results in the Hughes phenomenon [16]. In this study, the use of hyperspectral data sets is focused in order to detect brain tumour samples. For this reason, the supervised classification model is adopted in which kernel methods have been widely used due to their insensitivity to the curse of dimensionality [13]. However, the good generalization capability of machine learning techniques such as SVM [17] can still be enhanced by an adequate extraction of relevant features to be used for classification purposes [18], especially if limited training sets are available a priori.

This paper proposes a SVM Hierarchical Tree approach based on the well-known classifier support vector machine classifiers [19]. This classifier combines three ideas: 1) the optimal search technique hyperplanes as a solution, 2) the idea of convolution of scalar product, linear extension to non-linear functions, and 3) the notion of soft margins, which allows errors in the training patterns. This technique works with the Structural Risk Minimization or SRM, which

is better than ERM (Empirical Risk Minimization) and many other techniques. SVM can be reduced to a convex quadratic programming problem (QP). It offers some advantages in comparison to the classical methods and they seem to get a better performance (more robust) with high amounts of data. In the literature, SVM has been used in many fields, such as text categorization, image classification, bioinformatics, and of course in medical imaging [8], [11], which performs better accuracies compared to the classical techniques used.

In the proposed work, it has been developed a new technique in which the different levels of information are taken into account. The main advantage of the spectrographic samples is the spectral resolution of each pixel. This special characteristic allows exploiting a huge amount of information. The new technique is composed by a hierarchical tree with several classification systems as Fig. 2 shows. These classification systems use support vector machine algorithm in order to discriminate between the selected classes. This strategy can be divided in two steps depending on the hierarchical tree. In the first step, this approach is designed to provide information about the tumour tissues and normal tissues available in the input data sets. The principal goal of the new approach is focused on this hierarchical tree. It is important to obtain acceptable results in this part of the algorithm because in case of getting too many errors (low classification results) the next classifiers will not perform well, as they will make a classification based on erroneous input data. However, the proposed algorithm provides additional information about the behaviour of these tissues. In order to address this issue, the second step composed by the second decision tree is introduced. It is divided in two classification systems: second classification system and third classification system. In order to obtain information about the level of the tumour tissue, the second classification system is developed. The possibility to explore in detail the nature of the tumour tissues can be observed in this classification system. In the third classification system, normal tissue can be separated into normal gray and normal white matter. With this approach, more detailed information about the nature of the samples is obtained. It can be developed in more detail but it depends on the samples known a priori.

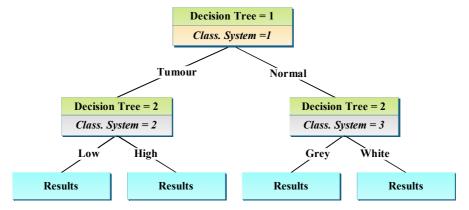


Fig. 2. Block diagram illustrating the new decision tree approach using several classification systems.

The new technique is developed in supervised fashion and a very small set of training samples has been selected. The process of the new proposed supervised strategy can be summarized by the flowchart given in Fig. 3. It graphically shows how the decision tree approach is divided in several classification systems over the medical samples known. This system is also composed by a set of training samples for the classification process and tested with the remaining samples.

D. Evaluation

In order to quantitatively evaluate the performance of the proposed approach, it is necessary to define some figures of merit. For this work, it has been adopted several standard metrics described below:

- Sensitivity: relates to the tests ability to identify a condition correctly. It is obtained as the number of true positives divided by the total number of true positives and false negatives in population.
- Specificity: relates to the tests ability to exclude a condition correctly. It is obtained as the number of false negatives divided by the total number of true negatives and false positives in population.
- Average Accuracy: the average accuracy is calculated as the sum of the sensitive and specificity figures divided by the number of classes in the test set.
- Overall Accuracy: is the total number of correctly classified samples divided by the total number of test samples. So, it is the probability that a sample will be correctly classified by a test.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The experimental settings used to validate the described new technique using real spectrographic samples are described in this section. The proposed approach is mainly based on the support vector machine classifier. For this process, a very limited training set was selected for each classifier, in order to study the worst case for the proposed approach, considering that these samples are difficult and expensive to acquire. The reported measurements correspond to the average of the results after running 10 independent Monte Carlo runs with respect to the training set in order to demonstrate the stability of the proposed approach. Kernel parameters were optimized by a grid search procedure, and the optimal parameters were selected using 10-fold crossvalidation. Particularly in this set of experiments, it has been chosen the Linear, Polynomial and Gaussian Kernel. The LIBSVM library was used for the experiments (available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/).

The classification system is currently in the implementation stage in order to reach real time performance. Two platforms have been considered: the NVIDIA GPU Tesla K40 [20] and the Kalray many-core platform [21]. This paper shows different tables in which can be appreciated the evaluation metrics (percentage and standard deviation), the training set (first class/second class), its corresponding decision tree and its corresponding classification system. This section is organized as follows. Subsection III-A introduces the first hierarchy decision tree results obtained in these experiments. Subsection III-B describes the experiments conducted with a second hierarchy decision tree. And finally the general study of the impact of this approach is showed in the subsection III-C.

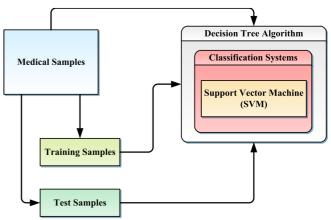


Fig. 3. General process of the proposed approach for brain cancer detection.

A. First Decision Tree Hierarchy

The first decision tree hierarchy is the most important part of this technique in which the main goal in cancer detection is to optimally separate normal and tumour tissue.

For the validation process, it has used the standard metrics previously described and several sets of training samples are used for the classification process as can be observed in Table II.

Table II reveals competitive accuracy results for spectrographic medical samples in cancer detection (tumour and normal tissue) using Linear, Polynomial and Gaussian kernel. In all the cases, the first decision tree is seen to be a robust framework in our classification system, even when we used a very small set of training samples. For instance, with 6 and 4 labelled samples in Gaussian kernel the classification results are around 99%. It can observe that the proposed technique can efficiently separate tumour and normal tissue.

TABLE II. FIGURES OF MERIT OBTAINED USING THE FIRST DECISION TREE OF THE PROPOSED APPROACH FOR BRAIN CANCER DETECTION WITH SEVERAL SET OF TRAINING SAMPLES AND DIFFERENT KERNELS

Decision Tree 1							
	Classification System 1						
	Linear Kernel		Polynomial Kernel		Gaussian Kernel		
Training Set	6/4	9/6	6/4	9/6	6/4	9/6	
Average Accuracy (%)	100	100	100	100	99.64 ± 1.80	100	
Overall Accuracy (%)	100	100	100	100	99.95 ± 0.22	100	
Sensitivity (%)	100	100	100	100	100	100	
Specificity (%)	100	100	100	100	99.27 ± 3.60	100	

B. Second Decision Tree Hierarchy

The proposed approach in the second decision tree hierarchy is composed by two classification systems. In this subsection, it has been described the experimental results obtained for this hierarchy tree. The second classification system is used to provide information about the level of the tumour tissue while the third classification system offers information about the nature of the normal tissue. Table III summarizes experimental accuracy results obtained after applying the second decision tree algorithm for medical samples. In the top of this table, it can be appreciated the set of training samples used for the classification process (number of samples for first class/number of samples for second class) and its correspondent kernel.

The classification accuracy results in the second classification system provide several problems in the case of the low level. However the classifier offers better accuracy results in the case of the high level. This behaviour is due to the unbalance between the numbers of samples known a priori. The use of support vector machine as a classifier

algorithm provides a high number of advantages but in this case it generates some problems because the number of samples between classes is unbalanced.

On the other hand, Table III also reveals competitive experimental results in the third classification system. This system is quite robust as the achieved classification results are very similar to those found in the optimal case.

C. General Analysis

Finally, Fig. 4 shows a classification comparison for algorithm using a reduced set of labelled samples when the classifiers are applied in supervised fashion. Fig. 4 displays the average of the classification results (out of 10 runs) obtained after applying the SVM to each classification system considered for the spectrographic samples. Improvements in classification accuracy can be appreciated (particularly, in the first decision tree part). Overall, the results reveal that, the new technique can provide adequate classification results in a challenging classification problem for brain cancer detection using spectrographic samples.

TABLE III. FIGURES OF MERIT OBTAINED USING THE SECOND DECISION TREE OF THE PROPOSED APPROACH FOR BRAIN CANCER DETECTION WITH SEVERAL SET

			Decision Tree 2	2				
	Classification System 2							
	Linear Kernel		Polyno	mial Kernel	Gaussian Kernel			
Training Set	6/8	8/12	6/8	8/12	6/8	8/12		
Average Accuracy (%)	77.47 ± 7.94	81.30 ± 8.51	75.90 ± 7.90	80.24 ± 9.97	69.93 ± 8.12	70.62 ± 9.45		
Overall Accuracy (%)	75.75 ± 10.33	80.21 ± 10.17	76.90 ± 9.14	81.08 ± 9.87	75.87 ± 9.57	81.08 ± 10.74		
Sensitivity (%)	75.34 ± 11.53	79.99 ± 11.21	77.13 ± 10.48	81.25 ± 10.82	77.27 ± 12.77	83.24 ± 13.46		
Specificity (%)	79.60 ± 12.66	82.62 ± 14.31	74.67 ± 15.53	79.23 ± 17.80	62.59 ± 25.63	58.00 ± 26.52		
	Classification System 3							
	Linear Kernel		Polynomial Kernel		Gaussian Kernel			
Training Set	3/5	4/6	3/5	4/6	3/5	4/6		
Average Accuracy (%)	98.50 ± 6.00	99.33 ± 3.30	99.50 ± 2.47	99.67 ± 2.36	96.00 ± 8.62	100		
Overall Accuracy (%)	98.29 ± 6.85	99.20 ± 3.96	99.43 ± 2.83	99.60 ± 2.83	96.57 ± 7.39	100		
Sensitivity (%)	100	100	100	100	92.00 ± 17.25	100		
Specificity (%)	97.00 ± 11.99	98.67 ± 6.60	99.00 ± 4.95	99.33 ± 4.71	100	100		

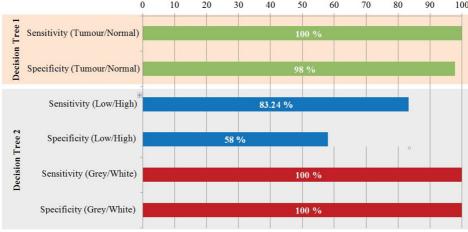


Fig. 4. The comparative accuracy results (specificity and sensitivity) for medical samples using a very small set of training samples using the proposed techniques

IV. CONCLUSIONS

In this work, a new approach for brain tumour tissue detection using spectrographic data in which a hierarchical decision tree based on support vector machine classifiers has been developed. In this context, the proposed approach offers information about the nature of the tissues. The main goal is to discriminate between tumour tissue and normal tissue which is achieved in the first decision tree. The behaviour of this technique has proven to be efficient in a set of spectrographic samples. One of the main advantages is the possibility of providing more detailed information about tumour and nor- mal tissue which could be useful in surgical situations. The accuracy of results is high when the number of samples of each class is well balanced. Further researches are focused on introducing spatial information in the acquisition data process in order to complete the applications of this proposed approach and improve the accuracy of results.

ACKNOWLEDGMENT

This work has been supported in part by the European Commission through the FP7 FET Open programme ICT-2011.9.2, European Project HELICoiD "HypErspectral Imaging Cancer Detection" under Grant Agreement 618080.

REFERENCES

- [1] R. Stupp, W. P. Mason, M. J. Van Den Bent, M. Weller, B. Fisher, M. J. Taphoorn, K. Belanger, A. A. Brandes, C. Marosi, U. Bogdahn et al., "Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma," New England Journal of Medicine, vol. 352, pp. 987–996, 2005.
- [2] M. J. McGirt, K. L. Chaichana, M. Gathinji, F. J. Attenello, K. Than, A. Olivi, J. D. Weingart, H. Brem, and A. r. Quin´ones-Hinojosa, "Inde- pendent association of extent of resection with survival in patients with malignant brain astrocytoma: clinical article," Journal of neurosurgery, vol. 110, no. 1, pp. 156–162, 2009.
- [3] H. L. S. R. D.-I. D. T. F. e. a. Claus EB, Horlacher A, "Survival rates in patients with low-grade glioma after intraoperative magnetic resonance image guidance," Cancer.
- [4] C. S. B. D. B. D. B. J. e. a. Shaw EG, Berkey B, "Recurrence following neurosurgeon-determined gross-total resection of adult supratentorial low-grade glioma: results of a prospective clinical trial," J Neurosurg.
- [5] L. K. C. S. P. M. C. S. e. a. Smith JS, Chang EF, "Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas," J Clin Oncol.
- [6] M. T. W. O. Z. F. R. H. Stummer W, Pichlmeier U, "Fluorescence-guided surgery with 5-aminolevulinic acid for resection of malignant glioma: a randomised controlled multicentre phase iii trial," Lancet Oncol
- [7] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," Journal of biomedical optics, vol. 19, no. 1, pp. 010 901–010 901, 2014.
- [8] R. Pike, S. K. Patton, G. Lu, L. V. Halig, D. Wang, Z. G. Chen, and B. Fei, "A minimum spanning forest based hyperspectral image classification method for cancerous tissue detection," in SPIE Medical Imaging. International Society for Optics and Photonics, 2014, pp. 90 341W-90 341W.
- [9] B. Fei, H. Akbari, and L. V. Halig, "Hyperspectral imaging and spectral- spatial classification for cancer detection," in Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on. IEEE, 2012, pp. 62–64.
- [10] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for Earth remote sensing," Science, vol. 228, pp. 1147– 1153, 1985.

- [11] H. Akbari, L. V. Halig, D. M. Schuster, A. Osunkoya, V. Master, P. T. Nieh, G. Z. Chen, and B. Fei, "Hyperspectral imaging and quantitative analysis for prostate cancer detection," Journal of biomedical optics, vol. 17, no. 7, pp. 0 760 051–07 600 510, 2012.
- [12] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," Geoscience and Remote Sensing, IEEE Transactions on, vol. 42, no. 8, pp. 1778– 1790, 2004.
- [13] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyper-spectral image classification," Geoscience and Remote Sensing, IEEE Transactions on, vol. 43, pp. 1351–1362, 2005.
- [14] C. V. Dinh, M. Loog, R. Leitner, O. Rajadell, and R. P. Duin, "Training data selection for cancer detection in multispectral endoscopy images," in Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012, pp. 161–164.
- [15] A. Plaza, J. A. Benediktsson, J. W. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," Remote Sensing of Environment, vol. 113, pp. 110–122, 2009.
- [16] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers," IEEE Transactions on Information Theory, vol. 14, no. 1, pp. 55–63, 1968.
- [17] G. Camps-Valls, L. Gomez-Chova, J. Mun~oz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," IEEE Geoscience and Remote Sensing Letters, vol. 3, no. 1, pp. 93–97, 2006.
- [18] A. Plaza, P. Martine, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 466–479, 2005.
- [19] V. Vapnik, Statistical Learning Theory. New York: Wiley, 1998.
- [20] "Nvidia launches worlds fastest accelerator for supercomputing and big data analytics," http://www.nvidia.es/object/tesla-k40-gpu-forsupercomputers-nov18-2013-es.html, 2013.
- [21] "Supercomputing on a chip a new generation of high-performance, low- power, real-time processors," http://www.kalrayinc.com/, 2015.