A Non-parametric Approach for Accurate Contextual Classification of LIDAR and Imagery Data Fusion

Jorge Garcia-Gutierrez, Daniel Mateos-Garcia, and Jose C. Riquelme-Santos

Department of Computer Languages and Systems, University of Seville,
Reina Mercedes s/n, 41012 Seville, Spain
{jorgarcia,mateosg,riquelme}@us.es

Abstract. Light Detection and Ranging (LIDAR) has become a very important tool to many environmental applications. This work proposes to use LIDAR and image data fusion to develop high-resolution thematic maps. A novel methodology is presented which starts building a matrix of statistics from spectral and spatial information by feature extraction on the available bands (RGB from images, and intensity and height from LIDAR). Then, a contextual classification is applied to generate the final map using a support vector machine (SVM) to classify every cell and the nearest neighbor (NN) rule to sequentially reclassify each cell. The results obtained by this novel method, called SVMNNS (SVM and NN Stacking), are compared with non-contextual and contextual SVMs. It is shown that SVMNNS obtains the best results when applied to real data from the Iberian peninsula.

Keywords: Remote sensing, supervised learning, contextual classifiers.

1 Introduction

Light Detection and Ranging technology (LIDAR) is a remote sensing laser-based technology that, as a main characteristic, can determine the distance from the source to an object or surface providing, not only the x-y position, but also the coordinate z for every impact. The distance to the object is determined by measuring the time between the pulse emission and the detection of the reflected signal, taking into account the position of the emitter. The main applications of LIDAR are related to digital elevation model extraction [2], forest inventories [1] and fuel models [5].

Thematic maps are remote sensing products used to study and manage geographical areas of interest according to a special theme. Although thematic maps have traditionally been generated from aerial and satellite images, the appearance of new sensors has led to increasing interest in data fusion to obtain high resolution products. In this way, LIDAR data-fused thematic maps can be seen in literature to map fire risk [14] or plant communities relations [17].

Classification techniques are traditionally applied to generate thematic maps.

If the information about each element's neighbour in addition to the element

itself is used, the classification process is said to be contextual [4]. One of the latest examples of a contextual classifier can be seen in [15]. The authors propose a technique called SVMMRF for hyperspectral classification consisting of two steps. In the first step, a probabilistic support vector machine (SVM) is applied to a pixelwise classification of hyperspectral images. In the second step, spatial contextual information is used for refining the classification results with a Markov random field (MRF) regularization.

In this study, we show a novel methodology to develop high resolution maps from LIDAR and imagery data fusion. This methodology is based on a contextual classifier technique for data fusion, we have called SVMNNS (SVM and Nearest Neighbour Stacking). SVMNNS tries to improve the contextual regularization with a non-parametric technique such as the nearest neighbour rule used in previous methodologies [8] but in an iterative manner which is the main difference with our previous approaches [7]. In addition, a fair comparison with another cutting-edge contextual technique (not addressed in our previous work) is shown when a high resolution classification is required.

This article is organized in the following manner. Section 2 provides a description of the methodology. Section 3 shows and analyzes the results achieved. Finally, Section 4 presents the main conclusions of the study and suggests future lines of investigation.

2 Method

This article proposes a novel method, called SVMNNS, which is contextual. The method has the objective of generating thematic maps from data fusion, supported by the use of several families of classifiers (SVM and Nearest Neighbour). Each SVMNNS step is described in detail in the following subsections.

2.1 Data Description

To carry out this study, data were collected from two zones of the Iberian Peninsula (see Figure 1). LIDAR and orthophotography data were obtained from each zone. In both cases, the altitudes of the LIDAR data were normalized through a digital elevation model generated by morphological filters on the LIDAR data [9]. Then, to make the LIDAR intensities usable, it was also necessary to perform normalization [11]. However, due to the scarce density of returns per square meter, the pulses with multiple impacts were not eliminated.

Orthophotographies of both areas were provided by the Regional Ministry of Andalusia in the case of Huelva, and the Laborate Group of the University of Santiago in the case of Trabada and they were taken in 2007 and 2004, respectively. In the case of Huelva, images have a resolution of 0.2 m whilst Trabada images have a resolution of 0.5 m. In both cases, they just contained RGB information.

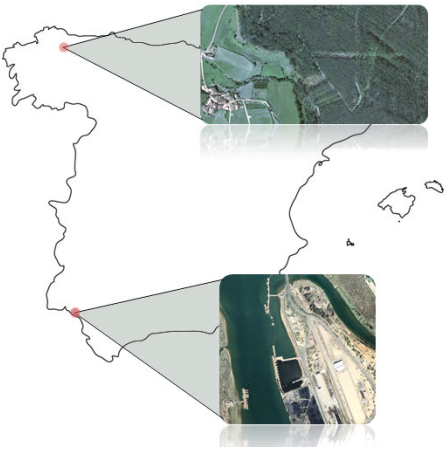


Fig. 1. Study areas of Trabada in the north and Huelva in the south

2.2 Feature Set Extraction

The first step in creating a thematic map is the generation of a matrix that covers the surface to be classified. In our case, the matrix has a vector of features associated to each of its cells which are calculated from images and LIDAR pulses. This fact involves selecting a resolution to set the size of cell in the matrix. The resolution is closely related to the density of the LIDAR pulses. In the case of Huelva, because the data had a low pulse density, the resolution value was set at 3 m^2, whereas in Trabada, where the data had a greater density, the resolution was set at 1 m^2.

In the next step, the features associated with each cell of the matrix are calculated. Seventy-one features were defined based on various other studies [6]. In Table 1, 54 features (9 measures for each of the 6 bands) are shown that are calculated from the available bands: RGB from the images, LIDAR intensity, LIDAR normalized altitudes and the SNDVI (Simulated Normalized Difference Vegetation Index). The SNDVI band simulates the NDVI (Normalized Difference Vegetation Index), substituting the value of the infrared for a value of the LIDAR return intensity. In addition, 17 other features were defined from the distributions of the LIDAR impacts (Table 2).

2.3 Training Set Selection

With a feature matrix calculated, training data are generated. A set of cells of the matrix is classified assigning them a label according to their land cover through a visual inspection of the aerial images or from other data sources. We selected 618 and 332 cells in the case of Huelva and Trabada, respectively. In both cases, the quantity of training instances was approximately 1% of the total. Every instance for the classification is then formed by the vector of stacked textures described in the previous subsection and a label that identifies its class.

Table 1. Band-based texture set

Symbol	Description	Symbol	Description
MAX	Maximum	MIN	Minimum
RANG	Range	STD	Standard Deviation
VAR	Variance	MEAN	Mean
KURT	Kurtosis	CV	Coefficient of Variation
SKEW	Skewness		

Table 2. LIDAR-distribution-based texture set

Symbol	Description
PCTR1	Percentage of first return
NOTFIRST	Number of non-first impacts
PCTR2	Percentage of second return
NEMP	Number of empty neighbours
PCTR3	Percentage of third or later return
TOTALR	Total number of impacts
PCTR21	PCTR2 over PCTR1
PEC	Penetration coefficient
PCTR31	PCTR3 over PCTR1
IID	Difference among first and last impacts in the pixel
PCTR32	PCTR3 over PCTR2
PCTN1	Percentage of single impact
EXTRASLP	Slope among every neighbor
PCTN2	Percentage of double impact
INTRASLP	Slope in the pixel
PCTN3	Percentage of triple or more impact
CRR	Canopy Relief Ratio

Table 3. Parameters C and γ for the radial basis function kernel applied to the SMO algorithm in each area

Area	C	γ
Huelva	8.12	6.17
Trabada	7.93	4.15

As is advised in various studies [18], a pre-process should be applied to data before generating a classification model. Thus, three types of filters are used. First, the missing values are replaced by the corresponding mean value. Second, the data are normalized. Finally, to avoid problems with dimensionality, a feature selection based on correlations is applied. The three filters are executed in Weka

[10] with default parameters. Once the filtered database has been obtained, the next phase is the execution of the SVMNNS algorithm.

2.4 SVMNNS

SVMNNS is characterized by applying two levels of classification by means of a stacking of two well-known classifiers: SVM and k-nearest neighbor (k-NN). Thus, from the training data, an SVM is initially generated by the SMO optimization algorithm [13] implemented in Weka (which was optimized by evolutionary computation [12] with the parameters in Table 3). The SVM performs a first-level classification for every cell of the matrix, assigning a value to its labels.

The next step is the application of a k-NN to reclassify each cell. The value of k is a parameter that the user has to set up at the beginning of the execution. This second level of classification is carried out with a implementation of the k-NN algorithm (IBk in Weka). For each cell, the k-NN is trained with just the values of the features of the adjacent neighbours in the matrix. Thus, SVMNNS introduces the context of each instance in its own classification.

For the zones that were used in this study, it was shown through experimental analysis that the best results were obtained with an 8-adjacency and a value of $k = 7$ (after testing on $k = 3, k = 5, k = 7$), that is, each NN was based on the 7 nearest neighbors from the 8 adjacent elements of every pixel. It should be kept in mind that the k-NN algorithm is very fast for a small number of training instances T. In our case, T depends on the number of adjacent elements where $T <= 8$ is satisfied. Knowing that for all k-NN, $k <= T$ is satisfied, we can conclude that the method is computationally tractable.

The number of refinements, which can be seen as a number of iterations, is also a configurable parameter n. The value of n is also defined by the user at the beginning of the execution of the SVMNNS method. With regard to the areas of this study, the number of iterations was set at 6 to assure the convergence which is reached at the fifth iteration in classical Markovian contextual classifiers [4].

Finally, after going through the final iteration, the algorithm generates a thematic map in which each pixel is associated with the label of the corresponding cell in the same position in the matrix.

3 Experiments

3.1 Results

To confirm the quality of the SVMNNS method, a comparison was established with two other classifiers: SVMs and SVMMRFs. To make a fair comparison of the results, the contextual algorithms (SVMNNS, SVMMRF) used the same initial classification obtained by the non-contextual SVM (optimized by the SMO algorithm) and the same number of subsequent refinements. The comparison is founded on a well-known testing strategies in remote sensing: a hold-out process.

Table 4. Hold-out percentage results for Huelva

		SVM		SVMMRF		SVMNNS	
Class	Num. instan.	Comm. error	Omiss. error	Comm. error	Omiss. error	Comm. error	Omiss. error
Water	2151	3.00	4.10	0.00	3.50	0.00	0.30
Marsh	1266	14.60	27.50	3.00	25.50	3.50	18,40
Roads	1083	23.30	10.70	6.40	4.00	4.50	4,30
Low Veg.	686	24.10	18.10	20.80	9.30	14.40	11.10
Mid. Veg.	464	49.80	48.50	29.50	13.50	30.20	11.20
High Veg.	329	37.10	42.50	19.50	21.40	4.30	12.50
Buildings	1314	20.20	22.90	11.70	4.80	4.10	3.40
Landfills	209	66.50	27.10	81.80	0.00	63.60	0.00
KIA		0.77		0.87		0.91	
Accuracy		81.02		89.66		92.90	

Table 5. Per-class partial accuracies for Huelva

	SVM		SVMMRF		SVMNNS	
Class	Partial Accur.	F-Mea.	Partial Accur.	F-Mea.	Partial Accur.	F-Mea.
Water	0.97	0.964	1	0.982	1	0.999
Marsh	0.854	0.784	0.97	0.843	0.965	0.884
Roads	0.767	0.825	0.936	0.948	0.955	0.956
Low. Veg.	0.759	0.788	0.792	0.845	0.856	0.872
Middle Veg.	0.502	0.509	0.705	0.777	0.698	0.782
High Veg.	0.629	0.601	0.805	0.796	0.957	0.914
Buildings	0.798	0.784	0.883	0.916	0.959	0.962
Landfills	0.335	0.459	0.182	0.308	0.364	0.533
Minimun	0.335	0.459	0.182	0.308	0.364	0.533
Mean	0.661	0.686	0.717	0.747	0.791	0.826

To perform the hold-out test, 7501 test instances (cells) were chosen in the Huelva study area, and 2320 were chosen in the Trabada study area which means that approximately 10% of the total number of instances of each dataset was used as test data. Table 4 and Table 6 show the results obtained by the SVM, SVMMRF, and SVMNNS methods, respectively. The tables show the number of test instances per class, the values for the errors of commission and omission, the global accuracy, and the Kappa index of agreement (KIA). In Table 5 and Table 7, the true positives and the F-measure per class are shown.

Based on the results obtained for every test instance counted as "hit" or "fail", we carried out a statistical analysis in order to determine if the classifiers behave statistically different. With a Cochran's Q test, we found that there exists

Table 6. Hold-out percentage results for Trabada

		SVM		SVMMRF		SVMNNS	
Class	Num. instan.	Comm. error	Omiss. error	Comm. error	Omiss. error	Comm. error	Omiss. error
Roads	640	38.00	23.40	35.50	17.90	18.00	12.20
Low. Veg.	549	12.60	32.40	4.20	31.00	4.60	18.00
High Veg.	765	14.10	21.10	5.50	16.70	7.30	10.70
Buildings	366	40.20	15.40	49.50	1.10	21.90	1.00
KIA		0.67		0.72		0.84	
Accuracy		75.56		79.61		88.10	

Table 7. Per-class partial accuracies for Trabada

	SVM		SVMMRF		SVMNNS	
Class	Partial Accur.	F-Mea.	Partial Accur.	F-Mea.	Partial Accur.	F-Mea.
Roads	0.62	0.686	0.645	0.723	0.82	0.848
Low. Veg.	0.874	0.763	0.958	0.802	0.954	0.882
High Veg.	0.859	0.822	0.945	0.885	0.927	0.91
Buildings	0.598	0.701	0.505	0.669	0.781	0.873
Minimun	0.598	0.686	0.505	0.669	0.781	0.848
Mean	0.738	0.743	0.763	0.770	0.870	0.878

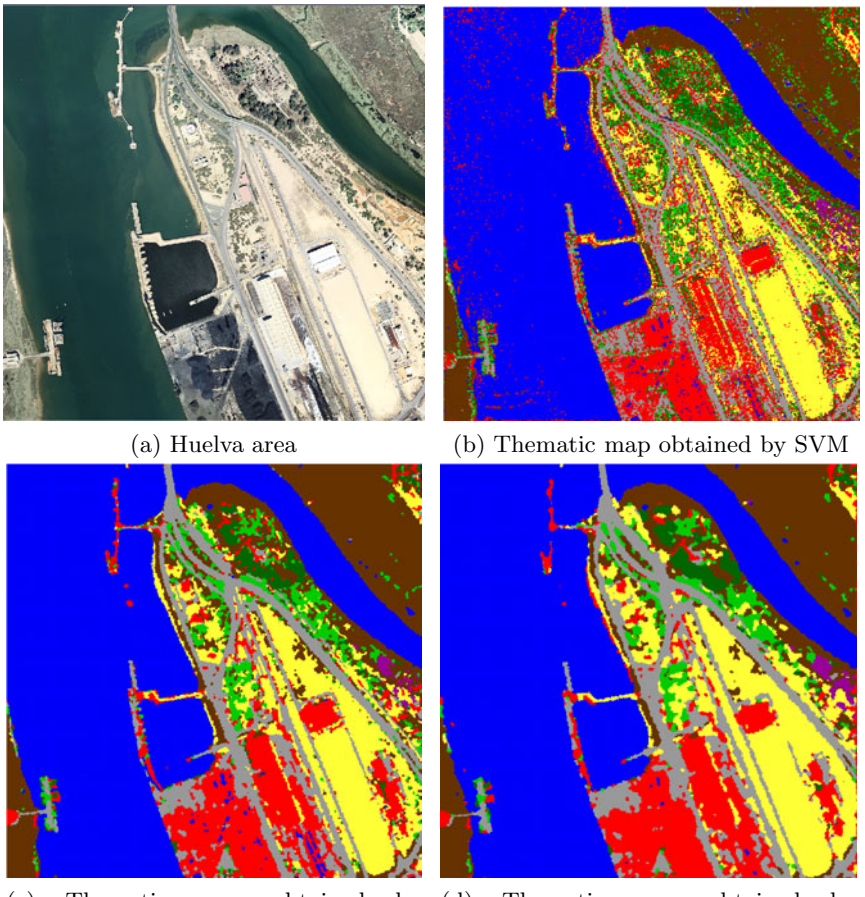
Table 8. Holm's adjusted p-values for the significance McNemar's tests

	algorithm	χ^2	p-value	Holm's
2	SVMMRF	300.62	1.32^{-10}	0.025
1	SVM	908.40	3.17^{-10}	0.05

a significant difference in usage among the three methods we surveyed ($p_{value} < 10^{-9}), \alpha = 0.05$). A posterior post-hoc analyis based on pairwise comparison using McNemar's tests (traditional non-parametric test in remote sensing [15]) with Holm's correction revealed that SVMNNS significantly obtained different accuracies for both areas as can be seen in Table 8.

Having found that the differences among the methods were significantly different for $\alpha = 0.05$, the analysis of the results concluded that the accuracy of the SVMNNS method was significantly better than that of its competitors for the data of the study areas.

Finally, the aerial images and the obtained thematic maps of Huelva and Trabada are shown in Figures 2 and 3 for visual comparison.



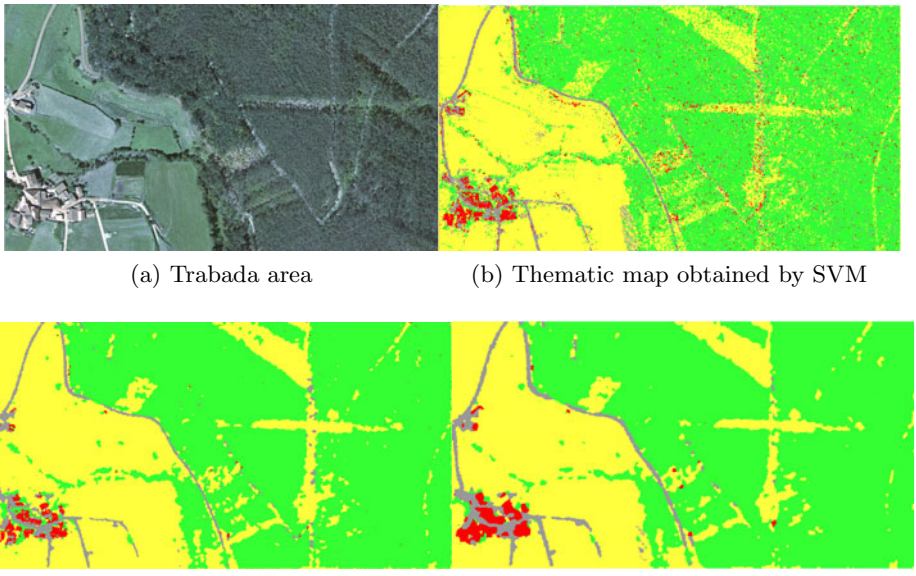
(a) Huelva area

(b) Thematic map obtained by SVM

(c) Thematic map obtained by SVMMRF

(d) Thematic map obtained by SVMNNS

Fig. 2. Final classification obtained for Huelva by each model. Water in blue, marshland in brown, roads and railways in grey, low vegetation and bare earth in yellow, middle vegetation in light green, eucalyptus in green, buildings in red and landfills in purple.



(a) Trabada area (b) Thematic map obtained by SVM

(c) Thematic map obtained by SVMMRF (d) Thematic map obtained by SVMNNS

Fig. 3. Final classification obtained for Trabada by each model. Roads in grey, low vegetation and bare earth in yellow, eucalyptus in green and buildings in red.

3.2 Discussion

The results for Huelva show a high level of accuracy (Table 4), with global accuracies above 81% and the contextual treatment improved the results in more than 8 points. Comparing both contextual methods, we can see SVMNNS obtained the best results for accuracy, both globally (almost three percentage points of difference from SVMMRF) and partially (see Table 5), except for the classes for medium vegetation and marshlands, categories for which SVMMRF performed better but with SVMNNS very close. If F-measure is selected as comparing measure in the place of partial accuracy, we will see that SVMNNS outperformed SVMMRF in every single category.

From a visual point of view (Figure 2, it can be observed that the contextual treatment improved the general results of SVM overcoming "salt and pepper" problems. An important issue related to this dataset is its clear problem of imbalance [3] (e.g., 2.151 instances of water and 209 of landfills). In any case, SVMNNS presents better behavior for the minority classes, in contrast to the SVMMRF and SVM methods (see F-measure for landfills in Table 5).

The Trabada area only had four classes, and although one could presume that the results would be better, the accuracies did not quite reach 80% except for SVMNNS (Table 6). This fact can be attributed to a lower level of separability from the available bands among classes and the influence of atypical instances that produce mistakes. The difference between SVMNNS and SVMMRF is much

higher in this case. As SVMNNS do not depend on the initial classification except for the first iteration, its results outperformed SVMMRF. However, it is important to stress that the use of the context notably improves the results of SVM (almost 4 points of difference in the worst case).

Visually (Figure 3), the final classification shows that SVMNNS is better than SVMMRF, which tends to introduce areas of eucalyptus in zones of buildings. Finally, it is worth noting the problems found with classification for every classifier when working on the road label. This may be due to the relatively similar infrared response that this class has compared with other classes.

By jointly observing the results, other conclusions can be made. First, it is worth underlining the synergy between LIDAR and orthophotography to separate the classes of both study areas. It is important to keep in mind that high levels of accuracy have been reached ($> 85\%$) for SVMMRF and SVMNNS with both datasets. In addition, SVMNNS improved the SVM and SVMMRF results not only globally but also in general for each label according to the partial accuracies. This improvement with respect to SVMMRF was greater for the Trabada data than for the Huelva data. That is, better performance is obtained with the greater resolution of evaluated data. The strategy of the classifier combination turned out to be more appropriate in this context than the parametric MRF approach to develop a contextual regularization, which, as has been observed, may generate a very high level of errors in certain cases. However, when the training set is easier to handle, the differences between the two contextual approaches are reduced, although SVMNNS maintains a higher level of accuracy for the results.

4 Conclusion

This study presented a novel method, called SVMNNS, to generate thematic maps based on data fusion with the objective of improving the return of investment for LIDAR and orthophotography joint flights. The results obtained by the SVMNNS method were compared with other classifiers (SVM and SVMMRF), and this comparison showed that SVMNNS obtained the best results in two distinct areas of the Iberian Peninsula.

Despite the good results that were obtained, there are still problems to be resolved. The most important issues are related to the method of extracting training data and their possible imbalances. The introduction of errors, the lack of accuracy, and the lack of completeness of the data provided by experts are problems that severely affect classification. To solve these problems, SVMNNS should evolve from supervised learning to active learning [16] introducing a detection phase for singular points, considering whether to eliminate them if they are possible outliers or to suggest their introduction within the training database if they are possible key instances. In addition, the application of new optimization techniques, such as evolutionary computation, may improve the final results by assigning weights to features, thus correcting problems associated with certain data sources e.g., those caused by intensities from multiple LIDAR returns.

References

1. Anderson, J., Plourde, L., Martin, M., Braswell, B., Smith, M., Dubayah, R., Hofton, M., Blair, B.: Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. Remote Sensing of Environment 112(4), 1856–1870 (2008)
2. Brzank, A., Heipke, C., Goepfert, J., Soergel, U.: Aspects of generating precise digital terrain models in the Wadden Sea from lidar water classification and structure line extraction. ISPRS Journal of Photogrammetry and Remote Sensing 63, 510–528 (2008)
3. Chawla, N., Japkowicz, N., Kolcz, A.: Editorial: Special issue on learning from imbalanced data sets. ACM SIGKDD 6(1), 1–6 (2004)
4. Cortijo, F.J., de la Blanca, N.P.: Improving classical contextual classifications. International Journal of Remote Sensing 19(8) (1998)
5. Garcia, M., Riaño, D., Chuvieco, E., Danson, F.: Estimating biomass carbon stocks for a mediterranean forest in central Spain using LIDAR height and intensity data. Remote Sensing of Environment 114(4), 816–830 (2010)
6. Garcia-Gutierrez, J., Mateos-Garcia, D., Riquelme-Santos, J.C.: EVOR-STACK: a label-dependent evolutive stacking on remote sensing data fusion. Neurocomputing 75(1), 115–122 (2012)
7. Garcia-Gutierrez, J., Mateos-Garcia, D., Riquelme-Santos, J.C.: A SVM and k-NN Restricted Stacking to Improve Land Use and Land Cover Classification. In: Corchado, E., Graña Romay, M., Manhaes Savio, A. (eds.) HAIS 2010, Part II. LNCS, vol. 6077, pp. 493–500. Springer, Heidelberg (2010)
8. García-Gutiérrez, J., Mateos-García, D., Riquelme-Santos, J.C.: Evor-stack: A label-dependent evolutive stacking on remote sensing data fusion. Neurocomputing 75(1), 115–122 (2012)
9. Goncalves-Seco, L., Miranda, D., Crecente, R., Farto, J.: Digital terrain model generation using airborne LIDAR in forested area of Galicia, Spain. In: Accuracy 2006, pp. 169–180 (2006)
10. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: An update. SIGKDD Explorations 11(1) (2009)
11. Hofle, B., Pfeifer, N.: Correction of laser scanning intensity data: Data and model-driven approaches. ISPRS Journal of Photogrammetry and Remote Sensing 62(6), 415–433 (2007)
12. Huang, C.L., Wang, C.J.: A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications 31(2), 231–240 (2006)
13. Keerthi, S., Shevade, S., Bhattacharyya, C., Murthy, K.: Improvements to Platt's SMO algorithm for SVM classifier design. Neural Computation 13(3), 637–649 (2001)
14. Koetz, B., Morsdorf, F., van der Linden, S., Curt, T., Allgower, B.: Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data. Forest Ecology and Management 256, 263–271 (2008)
15. Tarabalka, Y., Fauvel, M., Chanussot, J., Benediktsson, J.: SVM- and MRF-based method for accurate classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters 7(4), 736–740 (2010)

16. Tuia, D., Pasolli, E., Emery, W.: Using active learning to adapt remote sensing image classifiers. Remote Sensing of Environment 115(9), 2232–2242 (2011)
17. Verrelst, J., Geerling, G., Sykora, K., Clevers, J.: Mapping of aggregated floodplain plant communities using image fusion of CASI and LiDAR data. International Journal of Applied Earth Observation and Geoinformation (11), 83–94 (2009)
18. Zhang, S., Zhang, C., Yang, Q.: Data preparation for data mining. Applied Artificial Intelligence 17(5/6), 375–381 (2003)