Using uncertainty information to combine soft classifications

Luisa M S Gonçalves[1,2], Cidália C Fonte[2,3], Mario Caetano[4,5]

[1] Polytechnic Institute of Leiria, School of Technology and Managment, Department of Civil Engeneering, Portugal
[2] Institute for Systems and Computers Engineering at Coimbra, Portugal
[3] Department of Mathematics, University of Coimbra, Portugal
[4] Portuguese Geographic Institute (IGP), Remote Sensing Unit (RSU), Lisboa, Portugal
[5] CEGI, Instituto Superior de Estatística e Gestão de Informação, ISEGI, Universidade Nova de Lisboa, 1070-312 Lisboa, Portugal
luisa.goncalves@ipleiria.pt
cfonte@mat.uc.pt
mario.caetano@igeo.pt

Abstract. The classification of remote sensing images performed with different classifiers usually produces different results. The aim of this paper is to investigate whether the outputs of different soft classifications may be combined to increase the classification accuracy, using the uncertainty information to choose the best class to assign to each pixel. If there is disagreement between the outputs obtained with the several classifiers, the proposed method selects the class to assign to the pixel choosing the one that presents less uncertainty. The proposed approach was applied to an IKONOS image, which was classified using two supervised soft classifiers, the Multi-layer Perceptron neural network classifier and a fuzzy classifier based on the underlying logic of the Minimum-Distance-to-Means. The overall accuracy of the classification obtained with the combination of both classifications with the proposed methodology was higher that the overall accuracy of the original classifications, which shows that the methodology is promising and may be used to increase classification accuracy.

Keywords: Soft classifiers, uncertainty information, combining soft classifications

1 Introduction

A variety of different classification outputs can be obtained applying different classifiers to the same image with the same training sets. The classifiers have different capabilities and their performance depends of the application fields and image characteristics [1]. Through the combination of the outputs of a set of classifiers it is possible to obtain a classification that is often more accurate than the individual classifications ([1], [2], [3], [4]). To this aim several approaches have already been proposed. For example, [5] used an approach in which the class membership values for each class, derived from different methods, were summed and the class with the highest combined value is the one assigned to the pixel. In [6], Lu integrated classification results derived from individual classifiers using Dempster-shafer's theory of evidence. In [1], the authors applied two methods to improve accuracy of hard classifications, one, that he called a consensus builder system, to adjust classification output in the case of disagreement in classification between the maximum likelihood classifier, an expert system classifier and a neural network classifier. The second method integrated a rule-based expert system and a neural network classifier. The output of the expert system classifier was used as an additional new input layer in the neural network classifier. Doan and Foody [4] applied four methods for combining soft classifications. These methods were based on: 1) the selection of the most accurate prediction on a class-specific basis; 2) the average of the outputs of the individual classifications for each case; 3) the direct combination of classifications using reasoning and 4) the adaptation of the outputs to enable the use of a conventional (hard classification) ensemble approach.

Although several approaches have been proposed for combining hard classifications, the development of methods to combine soft classifications is still a field of investigation and the application of uncertainty information in this process is at an early stage.

The use of soft classifiers to perform image classification enables the generation of possibility or probability distributions for each pixel, depending of the classifiers used, where each probability or possibility is associated with a class of the nomenclature. The spatial units are assigned to the class presenting the larger degree of possibility or probability. The additional information provided by these classifiers may be used as indicators of the classifier difficulty to assign only one class to the spatial unit, and, together with the application of uncertainty measures, may provide valuable information that can be used in combined classification methods.

This study tests whether the proposed combining approach, that uses the uncertainty information obtained with two soft classifiers, improves the classification accuracy. The approach developed includes the following steps: 1) pixel-based soft classification; 2) application of an uncertainty measure to the outputs of the previous step to obtain the ambiguity information; 3) evaluation of the accuracy of the classification obtained in the first step; 4) development of rules to combine the soft classifications, that incorporate the information provided by the previous pixel-based classification and the results given by the uncertainty measure; 5) evaluation of the combined classification accuracy.

2 Data

The study was conducted in a rural area with a smooth topographic relief, occupied mainly by agriculture, pastures, forest and agro-forestry areas. The dominant forest species in the region are eucalyptus, coniferous and cork trees. An image obtained by the IKONOS sensor was used, with a spatial resolution of 4m in the multi-spectral mode (XS). The product acquired was the Geo Ortho Kit and the study was performed using the four multi-spectral bands. The geometric correction of the multi-spectral image consisted of its orthorectification. The average quadratic error obtained for the geometric correction was of 1.39m, less than half the pixel size, which guarantees an accurate geo-referencing.

3 Methodology

Two soft classification methods were used in this application: 1) the neural network Multi-Layer Perceptron (MLP); 2) a pixel-based supervised fuzzy soft classifier based on the underlying logic of Minimum-Distance-to-Means (FMDM). Both classifiers were trained using the same sampling protocol that included 100 pixels per-class. The classes used in this study are: Eucalyptus Trees (ET); Cork Trees (CKT), Coniferous Trees (CFT); Shadows (S); Shallow Water (SW), Deep Water (DW), Herbaceous Vegetation (HV), Sparse Herbaceous Vegetation (SHV) and Non-Vegetated Area (NVA). These classification methods assign, to each pixel, different degrees of assignment, in the case of MLP, and different degrees of possibility, in the case of FMDM, to the several classes under consideration. This extra data provide additional information at the pixel level which allows the assessment of the classification uncertainty.

3.1 Classifiers

The MLP is a non-parametric method and is the most commonly used neural network in remote sensing. Details of the MLP can be found in [7] and in [8]. The MLP provides an activation level for every output class of each pixel, and for hard classifications each pixel is allocated to the class with the largest activation level. A soft classification may be derived from this classifier by considering the activation levels of the network output units for each pixel. These activation levels range from 0 to 1, and may be used as the measures of class membership that reflect the class composition of the pixel [9] or indicators of the uncertainty associated with the pixel allocation to the classes. The second interpretation is used in this paper and the output values assigned to the pixels are used to compute classification uncertainty measures.

The second classification method used in this study is a pixel-based supervised fuzzy classifier based on the underlying logic of the Minimum-Distance-to-Means classifier. The underlying logic of this method is that the mean of a given signature represents the ideal point for the class, where fuzzy set membership is one. The fuzzy set membership is calculated based on a standardized Euclidean distance from each pixel reflectance, on each band, to the mean reflectance for each class signature, using a sigmoid membership function ([10]; [11]). When distance increases, fuzzy set membership decreases, until it reaches the user-defined Z-score distance where fuzzy set membership decreases to zero. To determine the value to use for the standard deviation unit, the information of the training data set was used to study the spectral separability of the classes and to determine the average separability of the classes.

With this classification methodology, the sum of the degrees of membership of each pixel to each class may sum up to any value between zero and the number of classes. Since fuzzy sets induce possibility distributions [12], a possibility distribution associated to each pixel is obtained.

Unlike traditional hard classifiers, the output obtained with these classifiers is not a single classified map, but rather a set of images (one per class) that expresses the probability, for the first classifier, and the possibility, for the second one, that each pixel belongs to the class in question.

To evaluate the classification accuracy of the two individual soft classifications a stratified random sampling with about 100 pixels per class was selected considering the entire image scene, which also included mixed pixels. The number of pixels was chosen to obtain a standard error of 0.05 for the estimation of the accuracy indexes of each class [13]. Each land cover class was sampled independently and the accuracy assessment was made with an error matrix.

3.2 Combination of classifiers

The outputs of the two individual soft classifiers were combined through the use of an uncertainty information measure. If the output classes for each individual pixel differed, the uncertainty information was compared and the class assigned with the lower value of uncertainty is chosen to be the one assigned to the pixel. In this approach the uncertainty measure E, developed by [14], was used to quantify the uncertainty at each spatial unit. This measure is given by

$$\mathrm{E} = 1 - p(x_1) \quad \mathbf{(1)}$$

where $p(x_1)$ is the largest degree of possibility or probability of the possibility distributions or probability distributions assigned with the pixel. This measure is also called ambiguity measure [15].

The first phase of the algorithm developed to combine classifications checks whether the same class is assigned to each pixel by both classifiers. If this condition is satisfied the class is accepted. If the two classifiers have different results for a certain pixel, the ambiguity information is used to make a judgement. The class with the lower ambiguity value is taken as the output for the pixel.

To evaluate if the combined classification improves the results, the accuracy assessment was made with the same protocol used with the single classifiers and the results were compared.

4 Results

4.1 Individual soft classifications

The accuracy assessment for both classifications was made with an error matrix and was undertaken with the same testing datasets. The error matrixes are generated assigning each pixel to the class with highest degree of possibility or activation level (in the case of the MLP classifier), corresponding to hard versions of the classifiers. The Global Accuracy was computed as well as the Users' Accuracy (UA) and the Producers' Accuracy (PA) for all classes. In terms of the overall accuracy, the classifications were similar. With the FMDM classifier method the overall accuracy was 66% and with the MLP classifier 65%. However, on a per–class basis, differences in accuracy are more evident. The results are shown in Fig.1 and Fig.2. The pixels that are not classified (NC) are also considered.

		Reference Label of the Classification with FMDM									UA (%)
		DW	SW	NVA	ET	S	HV	CKT	CFT	SHV	
MAP LABEL	NC	4	1	7	1	5	2	1	1	2	
	DW	96	0	0	0	1	0	0	0	0	99.0
	SW	6	92	2	0	0	0	0	0	0	92.0
	NVA	0	0	81	0	0	19	0	1	51	53.3
	ET	0	0	0	39	1	6	11	18	7	47.6
	S	3	0	0	0	87	0	10	1	1	85.3
	HV	0	0	1	0	0	107	0	6	10	86.3
	CKT	0	0	11	14	8	0	54	11	25	43.9
	CFT	0	0	0	8	0	24	6	31	4	42.5
	SHV	0	1	13	2	0	9	11	2	36	48.6
PA (%)		88.1	97.9	70.4	60.9	85.3	64.1	58.1	43.7	26.5	**65.5%**

Fig. 1. Error matrixes of the classifications obtained with the FMDM.

		Reference Label of the Classification with MLP									UA (%)
		DW	SW	NVA	ET	S	HV	CKT	CFT	SHV	
MAP LABEL	NC	0	0	0	0	0	0	0	0	0	
	DW	104	1	2	0	1	0	0	0	0	96.3
	SW	4	92	1	0	0	0	0	0	0	94.8
	NVA	0	0	56	0	0	0	2	0	1	94.9
	ET	0	0	0	8	0	0	0	2	0	80.0
	S	1	0	4	1	83	0	18	0	1	76.9
	HV	0	0	0	2	0	78	0	0	2	95.1
	CKT	0	1	5	18	18	0	40	9	16	37.4
	CFT	0	0	1	32	0	59	22	56	20	29.5
	SHV	0	0	46	3	0	30	11	4	96	50.5
PA (%)		95.4	97.9	48.7	12.5	81.4	46.7	43.0	78.9	70.6	**64.5%**

Fig. 2. Error matrixes of the classifications obtained with the MLP classifiers.

Fig. 3 a) and Fig. 3 b) show the classification results when each pixel is assigned to the class with higher degree of possibility with the FMDM classifier, and with the largest activation level with the MLP classifier.

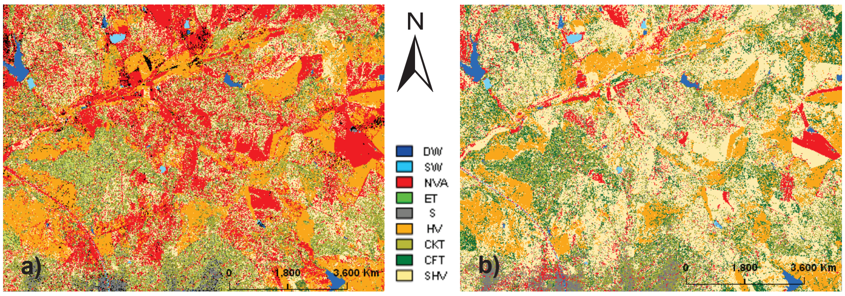

Fig. 3. – Hard version of the classification results with the a) FMDM and b) MLP.

The error matrix shows that water classes (DW and SW) were well identified by both classifiers. Forestry species were often confused between each other and with other classes, such as Sparse Herbaceous Vegetation (SHV) and Herbaceous Vegetation (HV).

With the MLP classifier the class with the smaller value of PA is Eucalyptus Trees (ET) (12.5%), which means it is the class with more omission error. With the FMDM classifier the class with the smaller value of PA is Sparse Herbaceous Vegetation (SHV) (26.5%). The class with smaller UA for both classifiers is Coniferous Trees (CFT). The MLP classification results for the UA was 29.5% and with the FMDM was 42.5%, and therefore it is the class with more commission errors.

The results obtained also shows that different classifications outputs were derived from the application of these two classifiers (Fig. 3). For example, with the MLP classifier the class NVA presents more omission errors then commission error and with the FMDM classifier it's the opposite. With the FMDM classifier a great amount of sites that should have been assigned to other classes, such as SHV and HV, were assigned to NVA, and were therefore absent from those classes, increasing their omission errors. With the MLP classifier a great amount of sites that should have been assigned to the NVA class were assigned to SHV class.

Images shown in Fig. 4 correspond to the spatial distribution of the ambiguity E committed when the pixel is assigned to the class corresponding to the largest degree of assignment. The regions with larger ambiguity (dark zones) are the ones where the assignment degrees were lower.

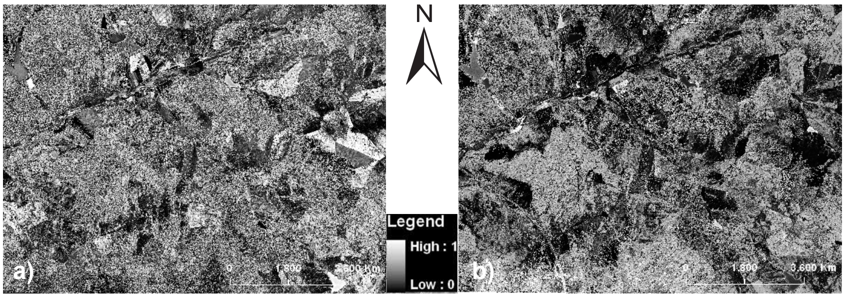

Fig. 4. Spatial distribution of ambiguity for the classifications obtained with: a) FMDM classifier b) MLP classifier.

The comparison of the mean ambiguity per class shows that forest species, such as CKT and CFT were assigned to the pixels with similar ambiguity by both classifiers (see Fig. 5). The class DW was assigned to the pixels with lower ambiguity with FMDM classifier, but all the other class presenter higher values of ambiguity with this classifier.

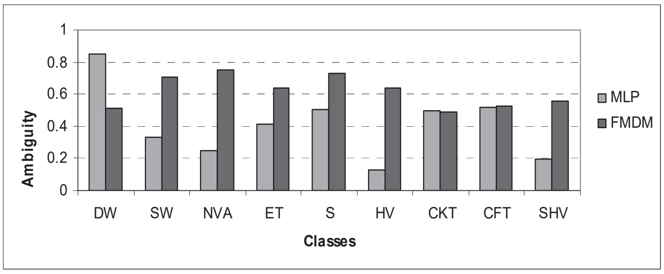

Fig. 5. Mean uncertainty per class.

4.2 Combined classifications

The accuracy assessment of the combined classification was made with the same testing datasets used to evaluate the individual classifications. The overall accuracy of the combined output was 4.5% higher than that of the most accurate individual classification. An improvement in some individual class accuracy was also observed (see Fig. 7 and Fig. 8). For example, the UA of classes SW, NVA, HV, CKT, SHV increased when compared with those of the most accurate individual classification (Fig. 1, Fig. 2, Fig. 6, Fig. 7). However, for some classes, the UA and PA of the combined classification didn't improve when compared to one of the initial classifications, such as the UA of the class ET when compared to the UA obtained with the MLP, or the UA of the class S when compared to the UA obtained with the FMDM. Although, the mean value of the UA and PA of all classes is higher than the mean values obtained for either of the initial classifications.

		Reference Label of Combining Classification									UA (%)
		DW	SW	NVA	ET	S	HV	CKT	CFT	SHV	
M	NC	0	0	0	0	0	0	0	0	0	
A	DW	104	1	1	0	1	0	0	0	0	97.2
P	SW	2	92	2	0	0	0	0	0	0	95.8
	NVA	0	0	56	0	0	0	2	0	0	96.6
L	ET	0	0	0	29	0	2	6	6	2	64.4
A	S	3	0	2	1	91	0	14	0	1	81.3
B	HV	0	0	0	2	0	85	0	0	2	95.5
E	CKT	0	1	7	20	10	0	54	10	15	46.2
L	CFT	0	0	1	10	0	50	10	53	18	37.3
	SHV	0	0	46	2	0	30	7	2	98	53.0
PA (%)		95.4	97.9	48.7	45.3	89.2	50.9	58.1	74.6	72.1	70%

Fig. 6. Error matrix of the combined classification.

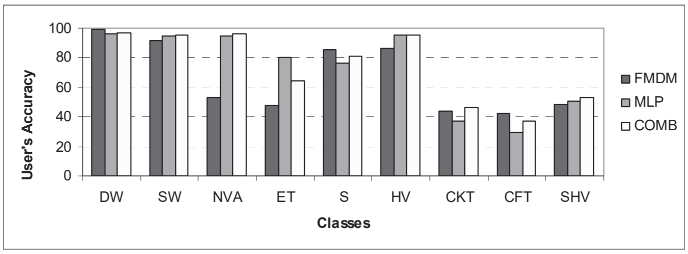

Fig. 7. User's Accuracy of the classes obtained with the FMDM, MLP and combined (COMB) classifications.

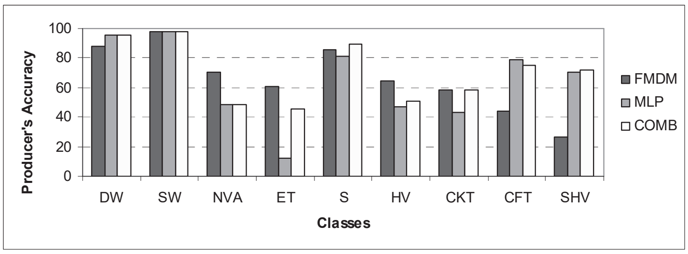

Fig. 8. Producer's Accuracy of the classes obtained with the FMDM, MLP and combined (COMB) classifications.

5 Conclusions

The classifiers tested in this study performed differently when applied to the same image, considering the same nomenclature and testing sets, and produced different results. Although the overall accuracy was similar for both individual classifications, on a per–class basis, differences were more evident. The proposed new classification methodology, integrating the results of both individual classifications, improved the overall accuracy of the classification. These results show that the information provided by the uncertainty measure was useful to determine the best class to assign to the pixels. Additional experiments will have to be made with other classifiers and uncertainty measures as well as the integration, in the combining decision process, of the uncertainty with the individual class accuracy information obtained with each classifier. However, the proposed approach seems to be promising, providing valuable information to the user, and deserves therefore further attention.

References

1. Liu, X.H., Skidmore, A.K., Van Oosten H.: Integration of classification methods for improvement of land-cover map accuracy. ISPRS Journal of Photogrammetry and Remote Sensing. 56, 257--268 (2002)
2. Liu, W.G., Gopal, S., Woodcock C.E.: Uncertainty and confidence in land cover classification using a hybrid classifier approach. Photogrammetric Engineering and Remote Sensing. 70, 963--971 (2004).
3. Huang, Z., Lees, G.: Combining non-parametric models for multisource predictive forest mapping. Photogrammetric Engineering and Remote Sensing. 70, 415--425 (2004)
4. Doan, H.T.X., Foody, G.M.: Increasing soft classification accuracy through the use of an ensemble of classifiers. International Journal of Remote Sensing. 28, 4609--4623, (2007)
5. Brown, D.G., Lusch, D.P., Duda, K.A.: Supervised classification of types of glaciated landscapes using digital elevation data. Geomorphology. 21, 233—250 (1998)
6. Lu, Y.: Knowledge integration in a multiple classifier system. Applied Intelligence. 6, 75—86 (1996)
7. Atkinson, P.M., Tatnall, A.R.L.: Neural networks in remote sensing. International Journal of Remote Sensing. 18, 699--709 (1997)
8. Brown, K.M., Foody, G.M., Atkinson P.M.: Estimating per-pixel thematic uncertainty in remote sensing classifications. International Journal of Remote Sensing. 30, 209--229 (2009)
9. Zhang, J., Foody, G.M.: Fully-fuzzy supervised classification of sub-urban land cover from remotely sensed imagery: statistical and artificial neural network approaches. International Journal of Remote Sensing. 22, 615--628 (2001)
10. Burrough, P. A., McDonnell, R. A.: Principles of geographical information systems. Oxford University Press (1998)
11. Kuncheva, I.L.: Fuzzy Classifier Design. Physica-Verlag, Springer-Verlag (2000)
12. Klir, G.: Generalized information theory: aims, results and open problems. Reliability Engineering and Systems Safety. 85, 21--38 (2004)
13. Stehman, S.V.: Statistical rigor and practical utility in thematic map accuracy assessment. Photogrammetric Engineering & Remote Sensing. 67, 727--734 (2001)
14. Chow, C. K.: On optimum error and reject tradeoff. IEEE Transactions on Information Theory. 16, 41–46 (1970)
15. Le Capitaine, H., Frélicot, C.: Classification with reject options in a logical framework: a fuzzy residual implication approach. In International Fuzzy Systems Association World Congress 2009 (IFSA 2009) and European Society for Fuzzy Logic and Technology Conference (EUSFLAT 2009), pp. 855--860 (2009)