

Extension of training set using mean shift procedure for aerospace images classification

Yuriy N. Sinyavskiy^{1,*}, Pavel V. Melnikov¹, and Igor A. Pestunov¹

¹ Institute of Computational Technologies of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Abstract. An effective method of training set extension for aerospace images classification is proposed. The method is based on mean shift procedure with respect to spatial information. It allows considering the unlabeled data structure. The results of experimental study using the Salinas hyperspectral image are presented, proving the effectiveness of the proposed method.

1. Introduction

In image classification tasks, a training set (labeled data) is necessary for constructing a decision rule or training neural networks [1, 2]. The process of obtaining a training set (TS) for aerospace image is often associated with significant material and time costs. Therefore, in practice, TS is available only for a small number of classes which is of interest to the user and is non-representative at the same time. Some classes can be represented by several labeled pixels.

It is known [3, 4] that in order to ensure an acceptable classification quality the minimum number of TS points for parametric classifiers should be of $10k$ per class order (where k is the dimensionality of the feature space), and for non-parametric classifiers should be of $50k$ per class. Therefore, the problem of obtaining a representative TS is especially relevant while processing hyperspectral images for which the number of spectral channels (features) is measured in hundreds.

At the same time, a characteristic feature of aerospace images classification tasks is that a large amount of unlabeled data is always available while solving them. In such circumstances the training set can be extended with the help of unlabeled data using methods based on clustering algorithms. Parametric methods based on the EM algorithm were most widely spread among such methods [5]. However, when processing aerospace images any *a priori* data on the probabilistic characteristics of classes are absent, as a rule. Therefore, the application of these methods can lead to unsatisfactory results.

A method of training set extension based on the nonparametric algorithm soft-PARZEN is proposed in [6, 7]. In this method “rigid” constraints on the form of conditional density function are not required. However, its application requires the presence of training sets for entirely all classes in the image. In practice, this condition is provided very rarely. In

* Corresponding author: yorikmail@gmail.com

addition, its application for processing aerospace images is associated with unacceptably high computational costs.

A common approach to the training set extension is based on the mean shift procedure. In [8] a semi-supervised support vector machine method is proposed. The support vector machine is applied to unlabeled data in such a way as to minimize an error when classifying both labeled and unlabeled data. The main disadvantage of the method is that the objective function for such a task is hard to optimize because it is not a convex one [9]. To solve this problem, a computationally effective gradient descent optimization method is suggested in [10].

In this work the method of training set extension based on mean shift segmentation with respect to the image spatial information is presented [11]. To demonstrate the efficiency of the proposed algorithm the constructed training sets were applied in order to classify the hyperspectral image of Salinas [12] with the Support Vector Machine (SVM) method using radial basis functions.

2. The proposed method of training set extension

The algorithm of the training set extension with the parameters $\{h_r, h_s, \alpha, p\}$ can be divided into three stages.

At the first stage, the image is segmented using the mean shift procedure with respect to spatial information. For this purpose mean shift procedure is applied to each pixel of the image. Only the pixels located at a distance of no more than h_r in the image domain and no more than h_s in the feature space [11] are taken into account when calculating the coordinates of a new center. The Euclidean distance between vectors is used as the distance between pixels. The distance is determined by the spectral brightness vectors in the feature space, and by the pixel coordinates vectors in the image domain. After that, segments with close centers (at a distance of no more than h_r in the image domain and no more than $h_s/2$ in the feature space) are combined and a representative is calculated for each segment (the average value of the feature vectors of all the pixels related to the segment).

At the second stage, a set of segments is formed for each class of the initial TS containing the points of the initial training set for this class. The segments from the obtained sets that simultaneously contain points from several classes of the initial TS are deleted. The appearance of such segments is possible in 2 cases. First, the parameters of the mean shift procedure can be unsuccessful. Second, errors can appear in the initial TS. Using the points from such segments in TS extension can lead to classification errors. Then, each set is extended by the segments located in the feature space at a distance of no more than α from it. The distance to a set is calculated as the smallest of the distances to its elements. The Euclidean distance between representatives is used as a measure of the distance between segments. After that, the segments that are simultaneously included in several sets are deleted again. These segments are located in the feature space on the boundary of the classes represented in the initial TS.

At the final stage of the algorithm, randomly selected points from the segments (p percent of the segment size) which are included in the set formed for this class are added to the initial TS for each class.

3. Experimental research

The proposed method of the TS extension is implemented in the C++ programming language (using Microsoft Visual Studio 2017 IDE) using the OpenMP standard.

A Salinas hyperspectral image (the Salinas Valley, California) measuring 512×217 pixels obtained from the AVIRIS sensor on October 8, 1998 was used in the experiment [12]. The features constructed by the principal component analysis were used for processing. Fig. 1 shows the image in false colors (the first three principal components) and the ground truth image for 16 classes.

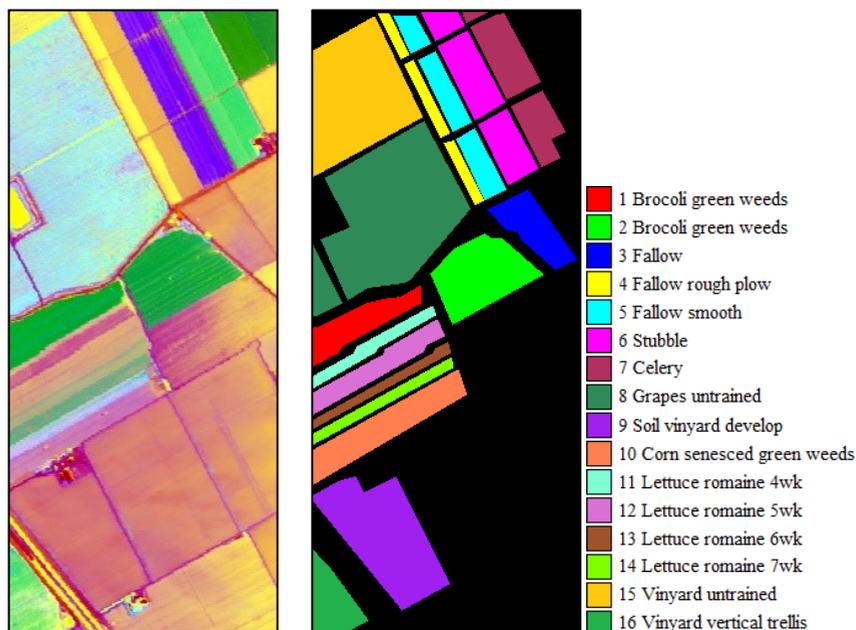


Fig. 1. Salinas image in false colors (first 3 principal components) and ground truth image for 16 classes.

The initial training sets containing 48 points (3 points per each class represented on the image), 80 points (5 points per class), and 160 points (10 points per class) were randomly formed according to the ground truth image. Each of these sets was extended by 5, 10, 15 and 20%. The parameter value $\alpha = 10$ was fixed. The extension of the sets was carried out using the first 4 principal components. The transition to the principal components is due to the high computational complexity of the mean shift procedure associated with the need for multiple calculation of the Euclidean distance in a multidimensional feature space. The resulting sets were used to classify the image using the SVM algorithm based on radial basis functions. The classification was carried out in the feature space obtained by the principal components analysis. The implementation of the SVM algorithm included in the Exelis ENVI software package was used. The default values were selected for all configurable parameters of the algorithm. The accuracy of the classification was defined by comparing the obtained classification images with ground truth one. Black pixels on ground truth image were ignored.

To average the results, the initial training sets were formed three times and for each of these sets the experiment with the same set of parameters was repeated three times. The obtained accuracy values of the classification (in percents) are given in the table. Fig. 2 shows the dependence of the average and maximum classification accuracy on the value of parameter p . The value $p = 0$ corresponds to the initial TS.

Accuracy values (in percents) for Salinas image classification with SVM algorithm using resulting TSs. The parameter value $p = 0$ corresponds to the initial TS.

Initial TS size	Value of parameter p												
	0	5			10			15			20		
48	12.0	66.2	64.1	65.0	67.5	68.6	68.1	70.6	69.0	69.6	69.9	69.5	69.5
	8.9	67.8	67.4	67.2	71.0	72.1	72.4	73.1	72.8	73.0	73.7	73.6	73.6
	9.5	60.7	61.2	60.6	61.5	61.7	61.7	62.8	62.5	62.6	64.3	63.7	62.9
80	33.2	73.9	74.9	73.9	77.0	76.6	76.7	77.2	77.7	76.4	77.4	78.1	77.6
	36.1	74.2	74.7	74.2	77.6	77.3	76.7	77.4	77.8	78.1	79.1	79.0	78.8
	39.8	64.0	64.0	65.7	67.4	67.3	66.2	67.8	68.1	67.1	67.9	67.5	68.3
160	51.2	62.9	62.3	61.7	63.6	63.7	63.1	63.8	63.6	63.4	64.0	63.3	63.6
	47.9	74.9	74.8	75.7	76.6	76.7	77.0	77.8	77.1	77.1	78.0	77.6	77.3
	53.8	66.4	67.3	67.5	70.5	70.9	71.1	72.2	72.2	75.5	73.2	73.5	73.5

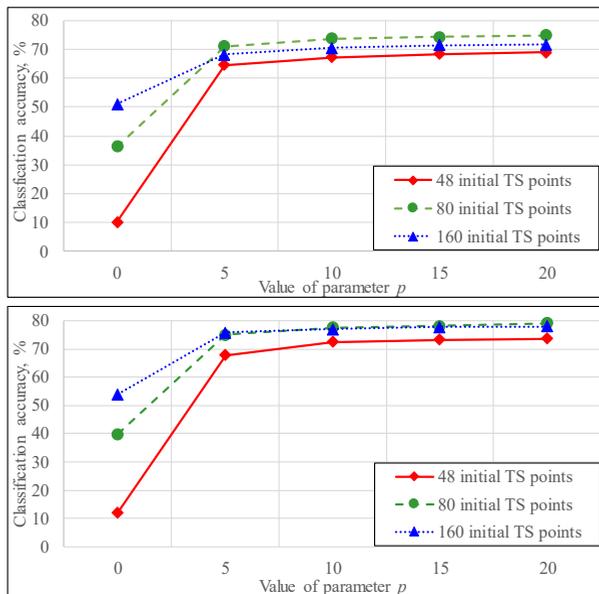


Fig. 2. The average (top) and maximum (bottom) accuracy of the Salinas image classification by SVM algorithm depending on the value of parameter p for different initial TS size. The value $p = 0$ corresponds to initial TS.

4. Conclusion

The work suggests an effective method of the training set extension for the tasks of aerospace images classification that allows to take into account the structure of unlabeled data. An experimental research using Salinas hyperspectral image showed that the proposed method is effective even for a small size of the initial set (3 points per class).

The work was carried out with partial financial support of the Integrated Program of Fundamental Research of the SB RAS “Interdisciplinary Integration Studies” for 2018-2020 (project No. 37) and the Russian Foundation for Basic Research (grant No. 18-37-00492).

References

1. A. Oliver, A. Odena, C. Raffel, E.D. Cubuk, I.J. Goodfellow, *Realistic evaluation of deep semi-supervised learning algorithms*, <https://arxiv.org/pdf/1804.09170.pdf> (2018)
2. Z. Dai, Z. Yang, F. Yang, W.W. Cohen, R.R. Salakhutdinov, *Adv. Neural Inform. Processing Systems*, 6510-6520 (2017)
3. Sh. Yu. Raudis, *Statistical Problems of Control*, **66**, 9-42 (Mathematics and Cybernetics Institute of LTS AS, Vilnius, 1984)
4. J.A. Richards, *Remote Sensing Digital Image Analysis* (Springer-Verlag, Berlin, 1999)
5. M.M. Dundar, D.A. Landgrebe, *IEEE Trans. on Geosci. and Remote Sensing* **42**, 264-270 (2004)
6. P. Juszczak, R.P.W. Duin, *Proc. 4th Intern. Conf. on Comp. Recogn. Syst.*, 203-210 (Springer-Verlag, 2005)
7. E.A. Kulikova, I.A. Pestunov, *KazNU Bulletin. Mathematics, Mechanics, Computer Science Series*, **13** (3), 284-290 (2008)
8. K.P. Bennett, A. Demiriz, *Advances in Neural Information Processing Systems*. 368-374 (1999)
9. X. Zhu, *Comp. Sci.* **2**, 4 (University of Wisconsin-Madison, 2006)
10. F. Gieseke, A. Airola, T. Pahikkala, O. Kramer, *Neurocomputing* **123**, 23-32 (2014)
11. D. Comaniciu, P. Meer, *IEEE Trans. Patt. Anal. Mach. Intell.* **24**, 603-619 (2002)
12. *Hyperspectral Remote Sensing Scenes*.
http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Salinas_scene (2014)