

Classification of Natural Images Using Supervised and Unsupervised Classifier Combinations

Leena Lepistö¹, Iivari Kunttu¹, Jorma Autio², and Ari Visa¹

¹ Tampere University of Technology, Institute of Signal Processing,
P.O. Box 553, FI-33101 Tampere, Finland
{Leena.Lepisto, Iivari.Kunttu, Ari.Visa}@tut.fi
<http://www.tut.fi/>

² Saanio & Riekkola Consulting Engineers,
Laulukuja 4, FIN-00420 Helsinki, Finland
Jorma.Autio@sroy.fi
<http://www.sroy.fi>

Abstract. Combining classifiers has proved to be an effective solution to several classification problems in pattern recognition. In this paper we use classifier combination methods for the classification of natural images. In the image classification, it is often beneficial to consider each feature type separately, and combine the classification results in the final classifier. We present a classifier combination strategy that is based on classification result vector, CRV. It can be applied both in supervised and unsupervised manner. In this paper we apply our classifier combination method to the classification of rock images that are non-homogenous in terms of their color and texture properties.

1 Introduction

Considering real life classification problems it is usual that the features are spread in many difficult ways. Different classifiers may classify the same sample in different ways and hence there are differences in the decision surfaces. However, it has been found that a consensus decision of several classifiers can give better accuracy than any single classifier [1],[10].

The goal of combining classifiers is to form a consensus decision based on opinions provided by different base classifiers. The base classifiers can differ from each other in many ways [6]. In the base classifiers, there can be differences in initializations, parameter choices, architectures, classification principle, training sets, or feature sets [6]. Combined classifiers have been applied to several classification tasks, for example to the recognition of faces [14] or handwritten characters [4], person identification [3], and fingerprint verification [9].

In image classification, a number of visual descriptors are used to classify the images. In the images, there are different types of visual features, like color, texture, and shape. The feature space is typically high dimensional and the categories of images are often overlapping in the feature space. A common approach in classification is to combine all the selected descriptors into a single feature vector. The similarity between these vectors is defined using some distance metric and the most similar images

are then classified (labeled) to the same category. However, when different types of descriptors are combined into the same feature vector, some large-scaled features may dominate the distance, while the other features do not have the same impact on the classification. Furthermore, various descriptor types use their specific distance measures in classification, which makes it even more problematic to combine these descriptors into single vector. Especially, in the case of high dimensional descriptors combination into the same vector can be problematic and may yield to remarkable drawbacks in classification performance. This is known as “the curse of dimensionality” [5]. Therefore, it is often more reasonable to consider each descriptor separately. This can be done using combined classifiers.

In this paper, we present a method to the classification of non-homogenous rock images using combined classifiers. In this method, separate base classifiers use different feature sets. The final classification can be obtained based on the combination of separate base classification results. In the base classification, supervised or unsupervised approaches can be used. In the experiments, we test these approaches in the classification of real rock images. The results are also compared to commonly used classifier combinations.

2 Combining Classifiers in Image Classification

The use of classifier combinations has been a subject of an intensive research work during last ten years. Popular solution on this field has been bagging [2], which subsamples the training data set. Another common algorithm, boosting [7], also manipulates the training data, but it emphasizes the training of samples which are difficult to classify. These methods, however, sub-sample the same feature set and therefore they cannot be applied to combination tasks with separate feature sets.

Recently, the probability-based classifier combination strategies have been popularly used in pattern recognition. Kittler et al. [10] have presented several common strategies for combining base classifiers. These strategies are e.g. product rule, sum rule, max rule, min rule, and median rule. These rules make final decision based on a posteriori probabilities generated by the base classifiers.

If the probability distributions of the base classifiers are not available, a simple combination strategy is voting. In the voting-based classifier combinations, the majority of the base classifier outputs decide the final class of the sample. In voting, the base classifiers can be regarded as black boxes, because any internal information about the base classification is not needed. Voting-based classifier combinations have been used in pattern recognition in [13].

2.1 Classification Result Vector (CRV)

The method for combining supervised classifiers is to combine the outputs of the base classifiers into a feature vector that is called classification result vector (CRV) [12]. The CRV contains the opinions of each single base classifier. Like in the voting-

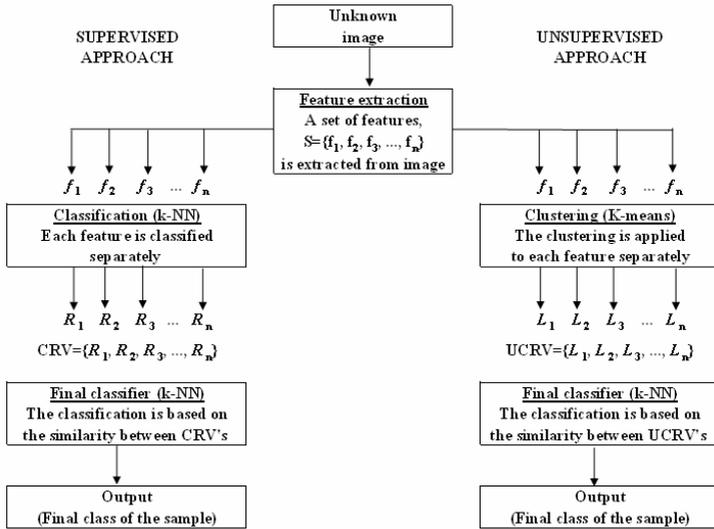


Fig. 1. The outline of the supervised and unsupervised classifier combination approaches

based approaches, also in the case of CRV method, the base classifiers can be treated as black boxes. This makes it possible to combine the classifiers without knowledge of their probability distribution. The CRV that contains the class labels provided by separate base classifiers can be used as the basis of the final classification. Hence, instead of using the majority decision of base classifier opinions, all the outputs of the base classifiers are used as the basis of final classification. This is the main difference between CRV method and voting.

The outline of the method is presented in figure 1. In this method, a set of n descriptors $S=\{f_1, f_2, f_3, \dots, f_n\}$ is extracted from an unknown sample. In the supervised approach, each feature is classified separately using supervised base classifier (e.g. k -NN classifier). The output of each base classifier is the number of class, to which the classifier assigns the sample. The class numbers are marked with $R_1, R_2, R_3, \dots, R_n$. The outputs of the separate base classifiers are then combined by collecting them into a vector. This vector is called classification result vector (CRV), and it is used in the final classification. The final classification of the unknown sample is made based on CRV's defined for each sample. Hence, the class numbers collected into CRV's form a new feature space, in which the final classification is made. In the final classification, the similarity between the CRV's is measured using Hamming distance.

In this approach, the visual descriptors extracted from the sample image are used in classification in the base classification, which is performed for each descriptor separately. The separate base classifiers classify similar samples in similar way, which is utilized in the final classification. Hence, the CRV characterizes the behavior of the base classifiers in the case of each input pattern, which makes it possible to use CRV's to describe the content of the unknown sample. Therefore, in the final classification the samples with similar base classification results are assigned to belong to the same category. This is made by comparing the CRV's of the samples. In contrary to

voting, in CRV method the individual features do not directly affect the final classification result. Therefore, classification result is not sensitive to variations and non-homogeneities of any single features.

2.2 Unsupervised Classification Result Vector (UCRV)

Accurate classification results can also be achieved by using unsupervised base classifiers. Using unsupervised base classifiers, the base classification represents the categories, to which the unknown sample is classified without supervision. Thus the base classification obeys the natural division of each input feature. In the base classification, some clustering algorithm [8] can be used. When the clustering is performed for each feature separately, the opinions of the unsupervised base classifiers represent the natural division of each descriptor of the unknown sample. These opinions are used as features in the final classification.

Our approach to combine unsupervised base classifiers is quite similar to the CRV method presented in section 2.1. Also in this approach, n features $\{f_1, f_2, f_3, \dots, f_n\}$ are extracted from an unknown sample (figure 1). The features are then fed into a clustering algorithm (e.g. K -means clustering) separately. In addition to the input features, the algorithm is also provided with the number of the clusters (K). The clustering algorithm performs the unsupervised classification of the input features. The output of the unsupervised classifier is a set of cluster labels defined for each feature separately, $\{L_1, L_2, L_3, \dots, L_n\}$. This set represents the unsupervised classification result vector, UCRV. In the final classification, the similarity between the UCRV's is defined using Hamming distance. The final classification is also in the case of UCRV supervised.

Also in this approach, the features extracted from the sample pattern are used in classification in the separate base classifiers, which in this case are unsupervised. Because the base classification is unsupervised, any information of real classes of the samples is not used. However, the clustering algorithm assigns similar samples in the same clusters in different feature spaces. Hence the samples which are similarly clustered can be assumed to be similar. Hence, in the final classification the samples with similar UCRV's are assigned to belong to the same category.

3 Classification of Natural Images

In this part we present experiments, in which we compare our method to other, commonly used classifier combination strategies. The application field in this case is classification of natural images. The division of natural images like rock, stone, clouds, ice, or vegetation into classes based on their visual similarity is a common task in many machine vision and image analysis solutions. The classification of natural images is demanding, because they are seldom homogenous.

Rock represents typical example of non-homogenous natural image type. This is because there are often strong differences in directionality, granularity, or color of the rock texture, even if the images represented the same rock type [11]. One typical application area of the rock imaging is bedrock investigation. In this kind of analysis,

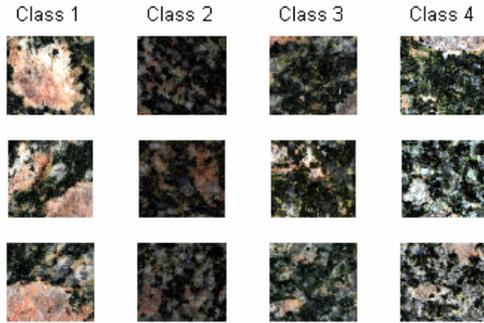


Fig. 2. Three examples from each class of the images in test set I

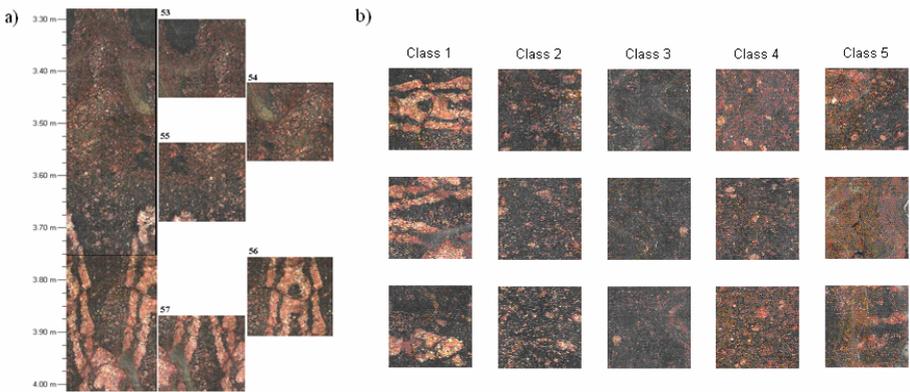


Fig. 3. a) The images of test set II are extracted from the borehole image. b) Three examples from each image class.

rock properties are analyzed by inspecting the images which are collected from the bedrock. In geological research, the rock properties are inspected based on the borehole images. Different rock layers can be recognized from the borehole images based on the color and texture properties of rock. Therefore, there is a need for an automatic classifier that is capable of classifying the rock images into visually similar classes. The use of classifier combinations has proved to be effective for this purpose [12].

As a testing database, we use two sets of rock images. Test set I consists of 336 images that are obtained by dividing large borehole images into parts. These images are manually divided into four classes by an expert. The division is based on their color and texture properties. Figure 2 presents three example images of each four class. In classes 1-4, there are 46, 76, 100, and 114 images in each class, respectively. In test set II, there are 66 non-homogenous rock images that are extracted from two boreholes. The division of a borehole image is presented in figure 3a. The example images of five types of rock in test set II are presented in figure 3b.

3.1 Classification

The classification is based on the color and texture properties of the test set images. The features used in this study consist of some MPEG-7 color and texture descriptors [16], and some other commonly used image features [12]. The color distribution of the images is characterized using color layout descriptor of MPEG-7 as well as hue and gray level histograms. In the description of texture properties of the images, homogenous texture descriptor and edge histogram descriptor of MPEG-7 are used. In addition, Gabor filtering method of Manjunath and Ma [15] is used.

In the classification experiments both supervised (CRV) and unsupervised (UCRV) approaches are tested. These approaches are compared to sum rule, max rule, median rule, and plurality voting [13], which have given good results in studies concerning classifier combinations. Product rule is not included into comparison, because the probability estimates of k -NN classifiers are sometimes zero, which may corrupt the result. The classification is based on the outline presented in figure 1. In the case of supervised classification (CRV), k -NN classification and leave one out validation methods are applied both in base classification and final classification. Unsupervised approach (UCRV) uses K -means clustering algorithm in the base classification, in which the value of K was selected to be 8, because it gave the most accurate classification. The final classification is similar to the supervised approach.

3.2 Results

The classification results are presented in tables 1a and 1b. The tables present mean classification result using separate base classifiers. In the lower part of the tables, the mean classification results using combined classifiers are presented. In this case, the classification results obtained from CVR and UCRV classification are compared to the results of the sum, max, and median rule and voting. In the classification, three values for nearest neighbors (k) are used. The results presented show that UCRV outperforms the other classifier combination methods in the classification of natural rock images. Both in databases I and II, the results were clearly best in the case of UCRV, whereas the CRV was the second in the comparison (database II). In the case of database I, CRV method was not able to give as accurate results, but it was still better than voting method.

4 Discussion

In this paper, we presented a method for combining classifiers in the classification of natural images. In the image classification, it is often beneficial to combine different visual descriptors to obtain the best possible classification result. Therefore, classifiers that use separate feature sets can be combined. We used natural rock images as testing material. Due to their non-homogenous nature, their classification is a difficult task.

Table 1. The classification results using separate features and classifier combinations for a) test set I and b) test set II

a)

Separate Features	Classification rate		
	k=1	k=3	k=5
ColorLayoutDescriptor	65.17 %	68.45 %	70.24 %
HomogenousTextureDescriptor	55.95 %	59.23 %	61.31 %
EdgeHistogramDescriptor	34.22 %	32.44 %	33.33 %
Gray level histogram	66.36 %	65.17 %	66.07 %
Hue histogram	69.35 %	67.56 %	65.18 %
Gabor features	61.31 %	62.80 %	61.90 %
Classifier combinations			
UCRV*	75.00 %	80.65 %	80.36 %
CRV	65.18 %	73.51 %	76.78 %
Sum rule	72.32 %	75.89 %	78.27 %
Max rule	65.18 %	74.70 %	75.00 %
Median rule	72.32 %	75.29 %	77.67 %
Voting	72.32 %	72.92 %	75.00 %

*K=8

b)

Separate Features	Classification rate		
	k=1	k=3	k=5
ColorLayoutDescriptor	59.09 %	54.55 %	57.58 %
HomogenousTextureDescriptor	46.97 %	46.97 %	46.97 %
EdgeHistogramDescriptor	31.82 %	36.36 %	36.36 %
GrayLevelHistogram	54.55 %	54.55 %	51.52 %
HueHistogram	63.64 %	62.12 %	62.12 %
Gabor features	48.48 %	53.03 %	56.06 %
Classifier combinations			
UCRV*	80.30 %	72.73 %	71.21 %
CRV	69.70 %	63.64 %	68.18 %
Sum rule	59.09 %	51.52 %	57.58 %
Max rule	59.09 %	56.06 %	56.06 %
Median rule	54.55 %	56.06 %	59.09 %
Voting	62.12 %	57.58 %	56.06 %

*K=8

The experimental results show that combined classifiers give accurate results also in the case of these kinds of images.

In our method, the feature vector that describes the image content is formed using the outputs of separate base classifiers. The class labels provided by the base classifiers form a new feature space, in which the final classification is made. In this space, the similarity between the feature vectors is defined using Hamming distance. Hence the final classification depends on the outputs of the separate classifiers, not the image features directly. This way the non-homogeneities of individual features do not have direct impact on the final result. The CRV can be formed using either supervised or unsupervised approach (UCRV). The unsupervised approach is a novel method that has proved to produce accurate classification results.

We compared the classification ability of our methods to other, commonly used classifier combination strategies in image classification. The results presented in table 1 show that our methods outperform the other classifier combination strategies in the classification of natural rock images. The results were clearly best in the case of UCRV, whereas the CRV proved also be accurate. It is also important that both the CRV and UCRV clearly outperform the voting method that is another black box approach in the classifier combination strategies.

In conclusion, the experimental results obtained from a practical image classification task show that methods presented in this paper are capable of accurate classification of real natural image databases. The methods are also computationally light and they are totally based on the outputs of separate base classifiers. Hence the base classifiers can be treated as black boxes, which make the proposed methods suitable for all kinds of classifier combinations.

Acknowledgment

The authors wish to thank Mr. Rami Rautakorpi from Helsinki University of Technology for evaluation of MPEG-7 descriptors for the test set images.

References

1. Alkoot, F.M., Kittler, J.: Experimental evaluation of expert fusion strategies, *Pattern Recognition Letters*, Vol. 20 (1999) 1361-1369
2. Breiman, L.: Bagging predictors. *Machine Learning*, Vol. 26, (1996) 123-140
3. Brunelli, R., Falavigna, D.: Person Identification Using Multiple Cues, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17 (1995) 955-966
4. Cao, J., Ahmadi, M., Shridhar, M.: Recognition of Handwritten Numerals with Multiple Feature and Multistage Classifier, *Pattern Recognition*, Vol. 28 (1995) 153-160
5. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd ed., John Wiley & Sons, New York (2001)
6. Duin, R.P.W.: The Combining Classifier: to Train or Not to Train, In: *Proceedings of 16th International Conference on Pattern Recognition*, Vol. 2 (2002) 765-770
7. Freund, Y., Shaphire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, Vol. 55 (1995) 119-139
8. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering, a review. *ACM Computing Surveys* Vol. 31 (1999) 265-323
9. Jain, A.K., Prabhakar, S., Chen, S.: Combining Multiple Matchers for a High Security Fingerprint Verification System, *Pattern Recognition Letters*, Vol. 20 (1999) 1371-1379
10. Kittler, J., Hatef, M., Duin, R.P.W., Matas J.: On Combining Classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20 (1998) 226-239
11. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Classification Method for Colored Natural Textures Using Gabor Filtering, In: *Proceedings of 12th International Conference on Image Analysis and Processing* (2003) 397-401
12. Lepistö, L., Kunttu, I., Autio, J., Visa, A.: Classification of Non-homogenous Textures by Combining Classifiers, *Proceedings of IEEE International Conference on Image Processing*, Vol. 1 (2003) 981-984
13. Lin, X., Yacoub, S., Burns, J., Simske, S.: Performance analysis of pattern classifier combination by plurality voting, *Pattern Recognition Letters*, Vol. 24 (2003) 1959-1969
14. Lu, X., Wang, Y., Jain, A.K.: Combining Classifiers for Face Recognition, In: *Proceedings of International Conference on Multimedia and Expo*, Vol. 3 (2003) 13-16
15. Manjunath, B.S., Ma, W.Y.: Texture Features for Browsing and Retrieval of image Data, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, (1996) 837-842
16. Manjunath, B.S., Ohm, J.-R., Vasuvedan, V.V., Yamada, A.: Color and Texture Descriptors, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11 (2001) 703-715