Classification of Cell Types in Feulgen Stained Cytologic Specimens using Morphologic Features

Martin Zarzycki[1], Timna Schneider[1], Dietrich Meyer-Ebrecht[1]
and Alfred Böcking[2]

[1]Institute for Image Processing, RWTH Aachen, 52074 Aachen
[2]Institute for Cytopathology, Heinrich-Heine-Universität, 40225 Düsseldorf
Email: mz@lfm.rwth-aachen.de

Abstract. We examine the effectiveness of morphologic features in their ability to discriminate between different cell types existent in samples from serous effusion. Using two major feature selection methods (*ranking* and *wrapper*) the original set of features is being reduced to smallest possible subsets with the best classification performance for each individual cell type. The evaluation of features and their combinations is based on scatter matrices and the k nearest neighbors classifier. The wrapper approach is driven by the floating search algorithm.

1 Introduction

Our analysis of methods for cell classification is motivated by the development of a new cytologic cancer detection system[1]. The idea of the overall system is a multimodal cell analysis where identical cell images can be examined in different stains. The multimodal approach allows new examination methods with substantially reduced demand for cell material. Usually the specimens contain different types of cells. In order to run an examination, relevant cells of different types need to be classified by an expert. This is a time consuming process and the automation of this task would make the application of the system more feasible.

We examine the classification of cells in Feulgen stained samples using morphologic features computed for nuclei. The selection of appropriate features for each individual classification task is essential in achieving good results. We use two common feature selection techniques to examine the performance of individual features and feature sets. On the one hand a simple method called *ranking* is applied to evaluate the performance of each feature. On the other hand a *wrapper* approach combined with the floating search algorithm is used to find strong feature combinations. Both selection algorithms use either scatter matrices or a k nearest neighbors classifier as a measure of the features' discriminative power.

2 Previous Work

So far there are numerous publications dealing with the classification of cells in microscopic images. In most of the papers - and in many other more general

classification tasks - some sort of feature selection is performed before applying a classifier. The number of available features often exceeds the maximum dictated by the *peaking phenomenon* so selecting a small and efficient set is necessary to ensure optimal results. An extensive survey over feature selection algorithms in general can be found in [2]. T. Wittenberg et al.[3] report on a practical application of feature selection on textural features in the context of cancer detection based on images with multiple cells. B. Fischer et al.[4] compare floating search with a genetic algorithm in their ability to find the best feature sets for classification of cancer in dermal tissue images. Finally M.Beller et al.[5] describe their approach to classify lymphocytes in blood. However they do not apply any feature selection.

3 Algorithms and Methods

Many advanced methods have been developed to analyze how features cope with a certain classification task. We focus on two common feature selection techniques.

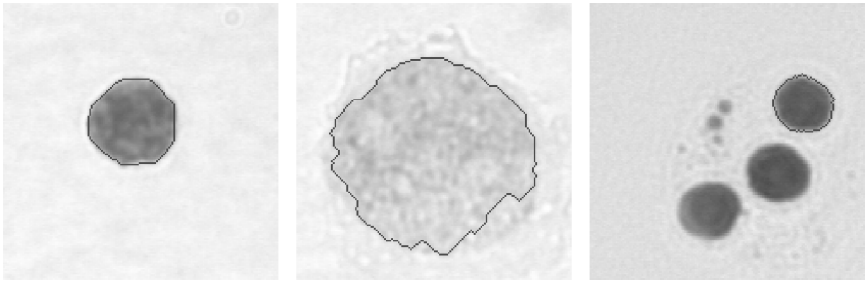
3.1 Ranking

The filter approach or ranking is a simple way to get an insight into the relevance of individual features. Each feature in the original set is being evaluated in respect to some criterion function. It is assumed that the features with the highest score are the best candidates for an optimal subset. The main advantage of this method is its linear scalability. The computation can be easily performed on collections containing thousands of features. Because each feature is evaluated separately there is no problem with the peaking phenomenon. However due to correlation between features and specific properties of the given criterion function a combination of the best individual features does not necessarily result in a good performance of the subset[2].

3.2 Wrapper Approach

An effective way of testing subsets of features for their relevance is incorporated in the wrapper approach. In most problems the number of features in the original set is so big that an evaluation of all possible subset combinations becomes computationally very expensive. In order to overcome this burden the wrapper approach makes use of a search algorithm that explores only a fraction of all possible subsets. Which subsets are tested depends on the evaluation results of previous subsets. The scores are the basis for the search algorithm in making its decision for testing new subsets. This feedback system that wraps around the results of one search step in order to guide the consecutive search, allows a very effective exploration of promising feature combinations. In this case a strategy called floating search[6] was implemented. Despite a greedy search and depending on the chosen criterion function the evaluation can take a considerable amount of time.

Fig. 1. Examples of the 3 different cell types: lymphocyte (a), mesothelial cell (b), granulocytes (c).



(a) (b) (c)

3.3 Criterion Functions

In order to make a decision about the relevance of a feature, some sort of criterion needs to be defined. One possibility of evaluation is to build a classifier and test the performance of a feature subset in terms of error rate using cross-validation. We have chosen the k nearest neighbors (knn) classifier for this task. As an alternative to the error rate of the knn classifier we implemented a function based on scatter matrices[7] also known as Fisher's criterion which measures class separability directly from the sample distributions.

4 Experiments

The three types of cells in this scenario and their quantity were lymphocytes (478), mesothelial cells (478) and granulocytes (107), all from serous effusion specimens. The microscopic images were made with a x63 lens and a resolution of 768x576 pixels. An example of each cell type is shown in figure 1. The 26 extracted features are part of an extensive feature collection for morphologic cell analysis proposed by K.Rodenacker and E.Bengtsson[8]. All of the features are at least position independent, some are additionally invariant to rotation. Finally in order to obtain more numerical stability the feature values were scaled to a uniform range throughout all three classes.

In the first step we used Fisher's criterion to create feature subsets. Using the filter approach, scatter values were calculated for each feature. Assuming that a high score represents a relevant feature, the results were sorted in decreasing order. The construction of promising subsets was implemented by successively grouping the ranked features into sets of increasing size. An alternative subset search was performed using the wrapper approach. Finally a knn classifier tested the subsets by calculating the error rate for each class using the *leaving-one-out* cross-validation technique.

Table 1. Classification results of best subsets found using Fisher's criterion.

granulocytes

	scatter value	knn result	subset size
ranking	4.98	**88.79**% (k=3)	10
floating search	**5.61**	87.85% (k=3)	15

lymphocytes

	scatter value	knn result	subset size
ranking	5.22	**96.53**% (k=3)	14
floating search	**5.56**	96.02% (k=1)	13

mesothelial cells

	scatter value	knn result	subset size
ranking	3.67	96.44% (k=3)	6
floating search	**5.29**	**97.28**% (k=3)	10

In a second step the same feature subset analysis was repeated but instead of Fisher's criterion we used directly the error rate of a knn classifier as the criterion function.

5 Results

Table 1 shows the results of the best subsets in terms of classification rate when using a search strategy guided by Fisher's criterion. Although the floating search algorithm succeeds in finding subsets with higher scatter values than the ranking approach, those scores obviously do not relate to the classification rate of a knn classifier. The reason for this is the irregularity in the sample distribution of some features. Especially features based on high order moments show a strong tendency to outliers, which degrades the predictive quality of scatter values. In those cases most samples of each class are scattered nearby clearly separated median values. The presence of outliers however, cripples the calculation of class means and variances, which are the base for a final scatter score. As a result, features that might discriminate well between classes, are being discarded by the search algorithm because of low scatter scores.

Much better results were obtained when instead of Fisher's criterion the classification rate of a knn classifier was used as an objective function. In this configuration the best subsets consisted only of one or two features. Additional features merely downgraded the classification performance. Table 2 gives an overview of the best features for each classification task. As mentioned before, particularly features based on moments show exceptional high classification rates. For more detailed information on features indicated in the table please refer to [8].

Table 2. Best features found using knn as criterion function for feature selection. B2.RAD is a geometric feature describing the radius of the largest fitting circle. The label B4 stands for features based on invariant moments.

	classification rate	feature type
granulocytes	96.26% (k=1)	B2.RAD
lymphocytes	100% (k=3)	B4.MM5
mesothelial cells	100% (k=3)	B4.IM4, B4.NM4

6 Conclusion

We have analyzed the performance of morphologic features in a cytologic classification task. The results in our scenario were very promising and show that morphologic features can perform very well. In the presented case of Feulgen stained cells only the cell nuclei are being classified, so the capability of morphologic features can not be transfered to other stains. Certainly more research on this aspect has to be done in future. Another conclusion of this work is the importance of a thorough feature selection in order to achieve optimal classification results. This includes the analysis of the distribution functions of feature values and the choice of appropriate criterion functions for the selection algorithm.

This project is supported by the *Viktor-and-Mirka-Pollak-Fund.*

References

1. Böcking A, Stockhausen J, Meyer-Ebrecht D. Towards a single cell cancer diagnosis. Multimodal and monocellular measurements of markers and morphology. Cellular Oncology 2004.
2. Guyon I, Elisseeff A. An introduction to variable and feature selection. In: Journal of Machine Learning Research 3; 2003. p. 1157–1182.
3. Wittenberg T, Neubauer K, Küblbeck C, Permanyer I, Schmidt R. Automatische Tumorerkennung bei unterschiedlichen Organen mittels Berechnung und Klassifikation von Texturmerkmalen. In: Bildverarbeitung für die Medizin; 2001. p. 377–381.
4. Fischer B, Palm C, Lehmann T, Spitzer K. Selektion von Farbtexturmerkmalen zur Tumorklassifikation dermatoskopischer Fotografien. In: Bildverarbeitung für die Medizin; 2002.
5. Beller M, Stotzka R, Gemmeke H, Weibezahn KF, Knedlitschek G. Bildverarbeitung für ein motorisiertes Lichtmikroskop zur automatischen Lymphozytenidentifikation. In: Bildverarbeitung für die Medizin; 2003.
6. Pudil P, Novovičová J, Kittler J. Floating search methods in feature selection. Pattern Recognition Letters 1994;15(11):1119–1125.
7. Fukunaga K. Statistical Pattern Recognition. Academic Press; 1990.
8. Rodenacker K, Bengtsen E. A feature set for cytometry on digitized microscopic images. Analytical Cellular Pathology 2003.