



Human vs. machine: evaluation of fluorescence micrographs

Tim W. Nattkemper^{a,*}, Thorsten Twellmann^a, Helge Ritter^a, Walter Schubert^{b,c}

^a*Neuroinformatics Group, Faculty of Technology, University of Bielefeld, P.O. Box 100131,
D-33501 Bielefeld, Germany*

^b*Neuroimmunology and Molecular Pattern Recognition Group, Institute of Medical Neurobiology,
University of Magdeburg, D-39120, Magdeburg, Germany*

^c*Mel Tec, ZENIT-Building, Leipziger Str. 44, D-39120, Magdeburg, Germany*

Received 26 April 2002; accepted 28 August 2002

Abstract

To enable high-throughput screening of molecular phenotypes, multi-parameter fluorescence microscopy is applied. Object of our study is lymphocytes which invade human tissue. One important basis for our collaborative project is the development of methods for automatic and accurate evaluation of fluorescence micrographs. As a part of this, we focus on the question of how to measure the accuracy of microscope image interpretation, by human experts or a computer system. Following standard practice we use methods motivated by receiver operator characteristics to discuss the accuracies of human experts and of neural network-based algorithms. For images of good quality the algorithms achieve the accuracy of the medium-skilled experts. In images with increased noise, the classifiers are outperformed by some of the experts. Furthermore, the neural network-based cell detection is much faster than the human experts.

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Functional proteomics; Fluorescence microscopy; High-throughput screening (HTS); Neural networks; Receiver operator characteristics (ROC)

1. Introduction

In the last several years, use of optical microscopy in cell biology has drawn the attention of biomedical research, drug–cell interaction and drug discovery, because of two developments in these

* Corresponding author. Tel.: +49-521-106-6058; fax: +49-521-106-6011.

E-mail address: tnattkem@techfak.uni-bielefeld.de (T.W. Nattkemper).

fields. First, the cell has been regarded as an important object of study in toxicology, new drug development, and patient–drug interaction [1]. Second, the strategies in microscopic analysis of biological cells have been expanded by ongoing automation of sample preparation, field selection, focusing and imaging [2]. Consequently, scientists are able to collect large sets of microscope images (*micrographs*), with constant quality.

In our collaborative research venture we use optical microscopy to identify surface proteins of biological cells and their topological distribution in tissue. To this end, we use multi-parameter immunofluorescence labeling. Since this is an optical approach, several inherent problems of data interpretation arise. These include (i) inconsistent staining of the sample, (ii) similar staining of unrelated like surrounding tissue, (iii) non-uniform illumination, (iv) out of focus blur and (v) shape variability of the objects of interest.

Both, computers and human experts, must have the ability to compensate for (i)–(v) to give reliable, significant, and reproducible data [3]. The identification and evaluation of image data is a difficult task in general. Decision support systems for evaluation by human experts are a substantial help for the improvement of biomedical studies on the cellular level [4].

In contrast to the advances in automation of microbiological image evaluation, the advances in computational evaluation of macrobiological image domains as reported by Shiffman and Shortliffe [5] and summarized by Duncan and Ayache [6], are enormous. These computational advances are mainly driven by innovative macrobiological imaging techniques, which have become standard diagnostic tools. Thus, there is an increased commercial interest in computer tools for image processing, registration and evaluation, that accelerates the advances in these fields. In microbiological imaging a large diversity of experimental setups can be observed. This diversity reflects the complexity and versatility of the cellular functions the biological researchers strive to understand. Thus, the development of techniques for the post-processing of microbiological images has been restricted to the academic world to a large extent.

This work addresses the question of how to assess the accuracy of a fluorescence micrograph evaluation. We believe that any advance in automatic micrograph evaluation depends strongly on significant measurements of the accuracy of any evaluation, be it done by a technical system or by human experts.

2. Methods

The micrographs in this study are evaluated by determining at which positions fluorescent cells are located, which are affected by factors (i)–(v). The identified cells are to be marked at referential (x, y) -image positions, that are to be saved in a database.

To discuss and analyze the evaluation performance of human experts and the computer system, records of cell detection in micrographs have been produced by five experts (called E_1 – E_5) and by two different computer algorithms. To compare and discuss the obtained results, a *perfect* evaluation result, the so-called *gold standard* is defined. The results from human experts and the computer system are analyzed as a contribution to discussions about the interpretation of fluorescence micrographs in the fields of cell biology and functional proteomics. In this collaborative work, fluorescence microscopy is used to identify cell surface proteins on lymphocytes in tissue section. The lymphocytes are stained iteratively with m individual monoclonal antibodies (mAb) that bind each to a

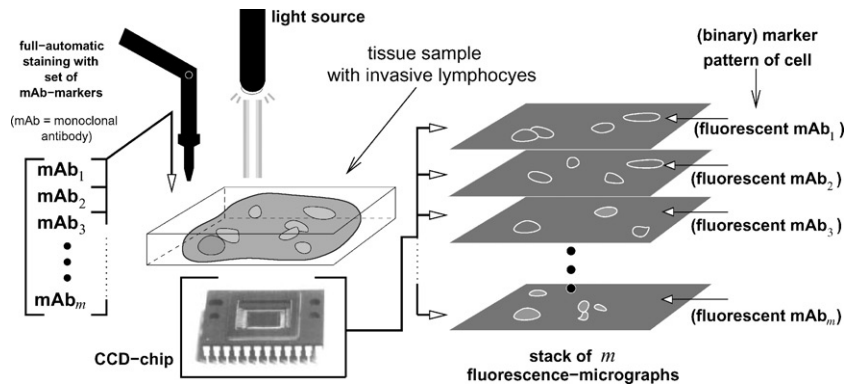


Fig. 1. Simplified sketch of the experiment and the imaging process: The lymphocytes are prepared with m markers in an iterative procedure. At each iteration step, the fluorescence excitation lets those lymphocytes appear in the image that are bound to the marker. A digital intensity image is taken by an integrated CCD-camera. After all m steps of preparation, fluorescence excitation and recording at a time are completed, a stack of m images with different subsets of lymphocytes can be evaluated.

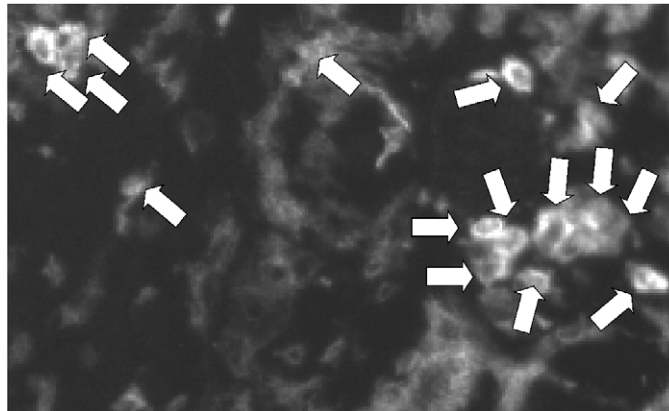


Fig. 2. An enlarged region from one micrograph is shown. The white arrows point at some fluorescent lymphocytes in focus on this plane and examples of many others present in the site.

specific cell surface protein. The cells that carry this specific protein on their surface are made visible by fluorescence excitation. This subset of all cells appears fluorescent in the micrograph, which is recorded with an integrated CCD camera. This is done for all m markers, so a stack of m fluorescence micrographs is recorded and transferred to a database (see Fig. 1). One important feature of our method is that the positions and shapes of the lymphocytes are not affected. Thus, a simultaneous measurement of proteins can be done by evaluation of locally corresponding signals through the stack. The technique is called MELK (multi-epitope-ligand-cartography) and its basic principles are described in more detail by Schubert [7,8]. In Fig. 2, an enlarged subregion of one micrograph is shown.

3. Micrograph evaluation by human experts

The four micrographs were given to five experts, E_1 – E_5 , who evaluated the micrographs. The micrographs are 658×517 pixels and each contains between 100 and 400 fluorescent cells. They are referred to by the respective applied mAb as CD3, CD4, CD7 and CD8. The experts inspected the images on a computer screen and marked the cell positions with a computer mouse cursor in the cell centers. Each of them evaluated the four micrographs blind from the others, so he had no idea about the other expert's results. As output of this experiment, for each micrograph five lists of cell positions were recorded and saved in a database. Caused by the subjectivity of human perception of the gray scale images, two positions for one cell marked by two different experts can vary by up to about two pixels, with an average cell diameter of 12 pixels.

Expert E_1 is one of the authors (W.S.), who has a long experience in the interpretation of visualized fluorescence signals for cellular protein identification. Furthermore, he is the designer of this applied multi-parameter microscopy technique. Experts E_2 – E_4 are members of the biomedical research group and have medium skills in visual inspection and interpretation of fluorescence micrographs. Expert E_5 is the first author of this paper (T.N.), a computer scientist with novice skills in interpreting micrographs.

In Fig. 3 the five detection results on the CD3-micrograph are separately visualized. Around all (x, y) -positions of a list recorded by one expert, white circles are drawn into the micrograph. The five visualizations show that the cell detection performances show considerable variation, even for the CD3-micrograph which contains a relatively low number of cells and is of good contrast.

4. Automatic cell detection with artificial neural networks

In general, the application of artificial neural networks (ANN) in the field of biomedical image segmentation and object detection is still exceptional. For instance a multi-layer-perceptron (MLP) can be applied to the detection of cell bodies in a micrograph by pointwise evaluation of the micrograph [9]. In the field of single cell image classification, successful applications of ANNs are reported regularly [9,10]. ANN have also been successfully applied for the exploration and classification of gene expression data [11,12].

In this work, the system acquires the expert knowledge from a set of hand-labeled image patches $\{\mathbf{X}_x\}$ of $N \times N$ pixels ($N = 15$). One-fourth of this set are patches that show fluorescent cells and is supplied by a single expert, who marks the cells in one example micrograph. This way, the primal experience of the expert is utilized, because he selects this training set the same way as marking cells in a *daily* image evaluation process. The rest of the training set consists of randomly chosen image regions that show no cells.

Using this training set, a classifier is trained to recognize if an image patch contains a fluorescent cell body or not. In this work, we implemented two different classifiers. One is a support vector machine (SVM), because in recent years SVMs were shown to be very efficient in learning and performing binary classification tasks and outperformed many other classifiers [13]. The other classifier is a variant of the self-organizing map (SOM), the local linear map (LLM). This classifier is also very efficient in learning object recognition tasks from even small training sets. Using the trained classifier, the cells in one micrograph are detected in two steps: First, for each micrograph pixel

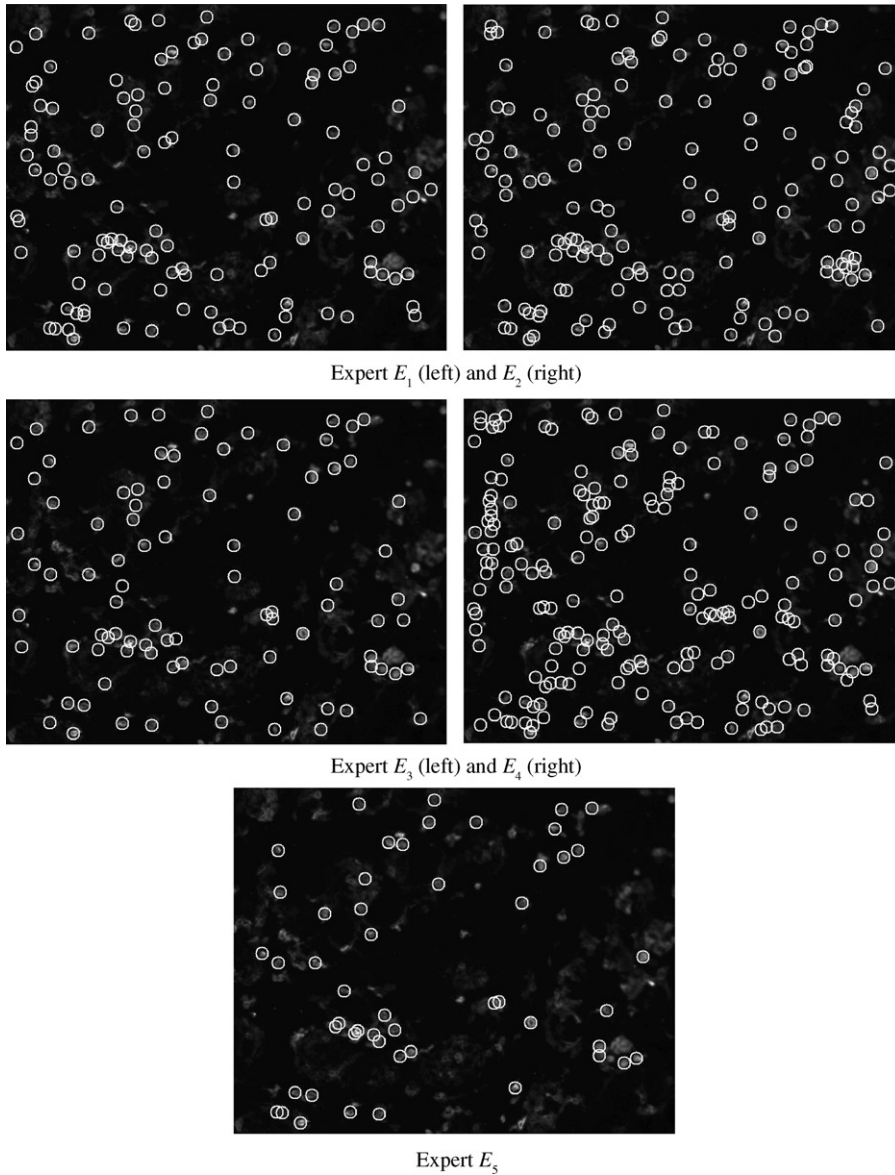


Fig. 3. In one example micrograph, the cells detected by the five experts E_1 – E_5 are visualized as white circles. Even if the cells cannot be identified in this plane, one can see that—even in the CD3-micrograph that contained a small number of fluorescent cells—the detection results vary to a considerable extent.

(x, y) a so-called *confidence value* is computed. This is done by mapping its $N \times N$ neighborhood $\mathbf{X}_{x,y}$ to the output value by the trained neural network. For each pixel of a micrograph, its confidence value is written to its corresponding position in a new image, which is called *confidence map*. In this map, the confidence values form mountains with large peaks that represent cell positions. A fast and easy to implement processing of the confidence value mountains gives a list of cell positions.

4.1. LLM-classifier

The LLM-approach [14] was originally motivated by Kohonen’s SOM [15,16] with the aim to obtain a good map resolution even with a small number of units. The LLM combines unsupervised and supervised learning in contrast to the widely used strictly supervised architectures, as the MLP trained with back-propagation [17]. From the training set $\Gamma = \{(\mathbf{x}_\alpha, \mathbf{y}_\alpha)\}_{\alpha=1, \dots, n_t}$ of n_t (input, output)-pairs, the LLM constructs a mapping $C : \mathbb{R}^{d_{\text{in}}} \mapsto \mathbb{R}^{d_{\text{out}}}$ by (i) a vector quantization of the input feature vector distribution (from 5 reference vectors in the present study) and (ii) an adaptation of local linear mappings into the output space, which are attached to each reference vector. Each input vector \mathbf{x}_α is computed from one image patch \mathbf{X}_α and is of much lower dimensionality, i.e., $d_{\text{in}} \ll N^2$, as described in the following section. The output values are set to $\mathbf{y}_\alpha = 1$ for positive training samples (cells) and $\mathbf{y}_\alpha = 0$ for negative ones (environmental tissue, background). After training, the LLM implements a mapping $C(\mathbf{x}_{(x,y)}) \mapsto [0, 1]$, with $\mathbf{x}_{(x,y)}$ as the feature vector computed from an $N \times N$ -sized image neighborhood of pixel (x, y) . For a more detailed description of the LLM learning, the reader may see Nattkemper et al. [18].

4.1.1. PCA or ICA features?

To compute lower-dimensional feature vectors, two approaches were implemented. First, the $N \times N$ -sized image patches were mapped to a six-dimensional feature vector \mathbf{x} using a principal component analysis (PCA). In this approach, the image patches were multiplied with the six eigenvectors of the six largest eigenvalues of the correlation matrix of the image patches from the training set $\{\mathbf{X}_\alpha\}$, that show fluorescent cells [18].

In a second approach, image patches are mapped to 10-dimensional feature vectors \mathbf{x} by scalar multiplication with ten filters, computed by an independent component analysis (ICA) on the set of cell patches of the training set (see Kämpfe et al. [19] for details). For computing the ICA-filters we choose the INFOMAX-approach, proposed by Bell and Sejnowski [20].

Comparing the PCA-based cell detection with the ICA-based showed that both approaches led to similar results for images of good contrast. Remarkably, in case of strong noise and lower contrast, the ICA-based feature computation outperformed the PCA approach. Thus, we discuss only the ICA-based classification in this work.

4.2. SVM classifier

For the last 5 years, SVM have undertaken a triumphal procession in the fields of (supervised) machine learning, pattern recognition, and classification. SVM implement a classification function by finding the data items $\{\mathbf{X}_i\}$ from the training set $\{(\mathbf{X}, \mathbf{Y})_\alpha\}$, that describe the best class separating hyperplane, which is found by solving an optimization problem. One could say, that those items are the “most informative”. In the context of this work, the input data are N^2 -dimensional gray value vectors of patches \mathbf{X}_α .

To find the hyperplane, or the descriptive set of support vectors $\{\mathbf{X}_i\}$, the data item is transformed into a higher-dimensional feature space, using a so-called *kernel function*. In this work, we used the prevalent radial basis function kernel. For training, we used the SMO algorithm, proposed by Platt [21] and enhanced as described in Keerthi et al. [22]. Best results were obtained with 30% of the training items as support vectors.

The training and application of the SVM to fluorescence micrographs is described in more detail in Twellmann et al. [23]. For a general introduction into SVM, the reader may see the articles by Smola and Schölkopf [24] or Burges [25] and for a comprehensive introduction the book by Christianini and Shawe-Taylor [26].

For training the SVM, the same set is used as for the LLM. Cells are detected in a micrograph by computing SVM output values for the neighborhood $\mathbf{X}_{x,y}$ of each image pixel (x,y) . The output values $C_{x,y}$ are written to the corresponding positions in a confidence map and processed to obtain a list of cell positions, as described below.

4.3. Computation of cell positions

The confidence map, computed either by LLM or SVM, is visualized as a gray value image by normalizing the values to $[0,255]$. The higher peaks of the confidence mountains correspond to centers of fluorescent cells. Thus, a point (x,y) in the confidence map is identified as the center of a fluorescent cell, if its confidence value $C_{x,y}$ obeys

$$C_{x,y} > t_c \wedge C_{x,y} = \arg \max_{d((x',y'),(x,y)) < N/2} \{C_{x',y'}\}, \tag{1}$$

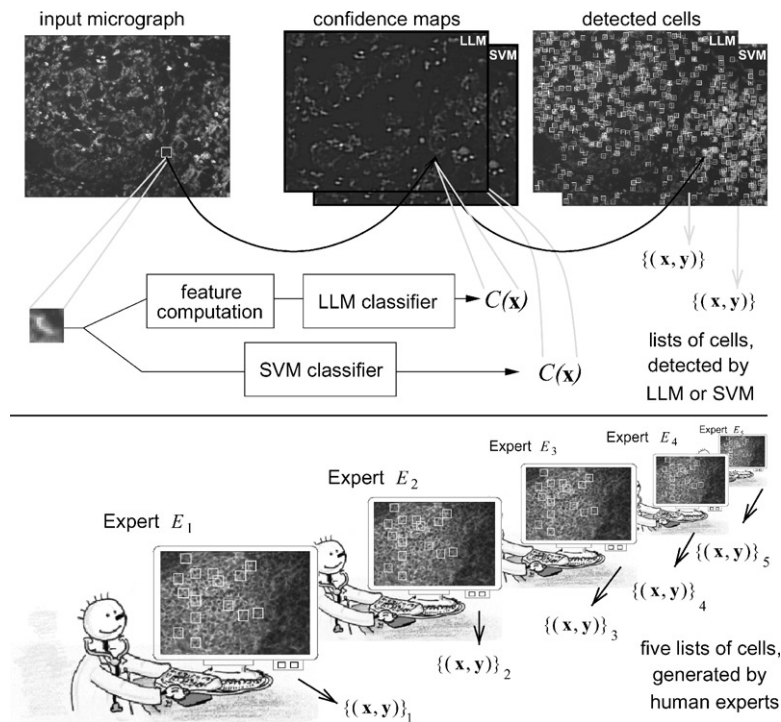


Fig. 4. One micrograph is evaluated seven times, leading to seven different lists of cell positions. Two lists are computed, using the artificial neural network classifiers. Each image point is mapped to a confidence value, representing the degree of belief that the pixel neighborhood shows a fluorescent cell. The confidence values are processed to obtain the final positions of fluorescent cells. For comparison, the micrograph is evaluated by five experts independently from each other.

where t_c is a given threshold value. In other words, all points which resemble a local maximum above a given threshold t_c in a circular region with diameter N are assumed to be cell centers. By pointwise evaluation of condition (1), the fluorescent cells are detected in the evidence map as the set of points that hold the condition. Fig. 4 illustrates how two lists of cell positions are computed by LLM and SVM and five lists of human expert counts are collected for comparison.

5. Accuracy evaluation

The *accuracy* of an evaluation has to be *defined* with respect to specific needs for analyzing the evaluation performance of human experts, and for inventing and optimizing an image processing system in biomedicine. So the first question is how to define the index for the accuracy of a system. In the literature many index terms like *sensitivity*, *specificity*, *efficiency*, *accuracy*, *utility*, *value*, *worth*, *effectiveness*, *usefulness*, *positive/negative predictive value* and *likelihood ratio*, to mention the most frequent, are listed. To assess the accuracy of an evaluation system, complementary pairs of terms (for example *sensitivity* against *specificity*) are plotted as so-called receiver operator characteristics (ROC).

In demonstrating the limits of a tool's or classifier's ability to discriminate between two alternative states, ROC occupy a unifying position. ROC are based on statistical decision theory and were introduced in the 1950s as reported by Metz [27] in the context of electronic signal detection and interpretation problems with radar. Later, the ROC were applied to medical imaging [28,29] and pap smear analysis [30]. The calculation of ROC for accuracy measurement is so widely used within the field of biomedical engineering, that its application has become a worldwide standard. To perform ROC analysis a test set of example input data has to be selected and classified into a set of positive and negative cases by an expert, concerning a considered property. This set of examples and their class labels (*positive*, *negative*) are called the *gold standard* or the *ground truth* and are considered as the correct classification result on the test set. To measure the accuracy of a human expert or a technical system each item of the test sample is given to that classifier, classifying the item to be positive or negative. This is done for all items of the test set and the results are compared with the gold standard. The classification results of the test sample can now be divided into four subsets:

TP: Positive cases, that are classified as positive (true positive).

FP: Negative cases, that are classified as positive (false positive).

TN: Negative cases, that are classified as negative (true negative).

FN: Positive cases, that are classified as negative (false negative).

The ROC values are calculated from the counts of these four values as explained below. For review on ROC the reader may refer to Zweig and Campbell [31].

In this work, human experts and neural cell classifiers are analyzed, as to how well they can discriminate between fluorescent cells and other structures in the micrographs that are not cells.

5.1. Generating the OR-expert as a gold standard

As indicated in the previous section, one has to possess a gold standard that represents the correctly classified test sample for measuring ROC. An observation of the human detection results

(an example case shown in Fig. 3), reveals that the experts show considerable inter-observer variation. This is often observed in the field of human visual inspection of biomedical image data [9,32–35]. Discussions with the experts led to the insight, that even those with advanced skills miss a large number of fluorescent cells in their evaluation. Thus, to get a more objective result, the five lists for each of the micrographs CD3, CD4, CD7 and CD8 are combined to one list per micrograph, which is called the OR-expert. The OR-expert list of a micrograph contains all cell positions in this micrograph that were detected by at least two experts. The lesser expertise of expert E_5 is not a source of error here, because he is by far the most precautionous evaluator.

5.2. ROC plots

The LLM and SVM classifiers are applied to each of the four micrographs. Thereby, two result lists per micrograph are generated in addition to the five lists per micrograph generated by the five human experts. The TP, FN, and FP are determined for each of the seven evaluations for each micrograph. These values have to be determined concerning the gold standard in the following way: A cell position of one result list is labeled to be a TP, if the gold standard list of positions contains a (x, y) -position with euclidian distance less than $N/3$ to this position. The parameter N is the image neighborhood size. The cell position in the gold standard list is labeled, to avoid it being counted twice. By evaluating the cell positions' labels, the TP, FN and FP are counted. Obviously, the value TN cannot be determined easily in this test scenario, so its evaluation is omitted. Based on evaluation of all result lists and both gold standards, the following two ROC values are computed:

SE: The sensitivity (SE) is the ratio of positive training samples that were correctly classified as positive. It is calculated as

$$SE = \frac{TP}{TP + FN}.$$

PPV: The positive predictive value (PPV) is the ratio of detected cells that are also items in the gold standard. It is calculated as

$$PPV = \frac{TP}{TP + FP}.$$

Based on the gold standard G_O , the SE and PPV are computed for all seven evaluations. In contrast to standard ROC plots, the threshold t_c is fixed, because we focus on the comparison between human and machine evaluation. The t_c is set at the intersection of SE and PPV for the training image, assuming the same costs for FP and FN. The resulting values of SE and PPV are plotted into one box per micrograph. In the box, the abscissa displays the SE and the ordinate displays the PPV. Thus, the ROC of one cell detection result is represented by a symbol at the point (SE, PPV) in that graph. The four result plots are shown in Fig. 5. The accuracy of a cell detection is defined by the euclidian distance of its symbol to the point (1, 1) in the graph.

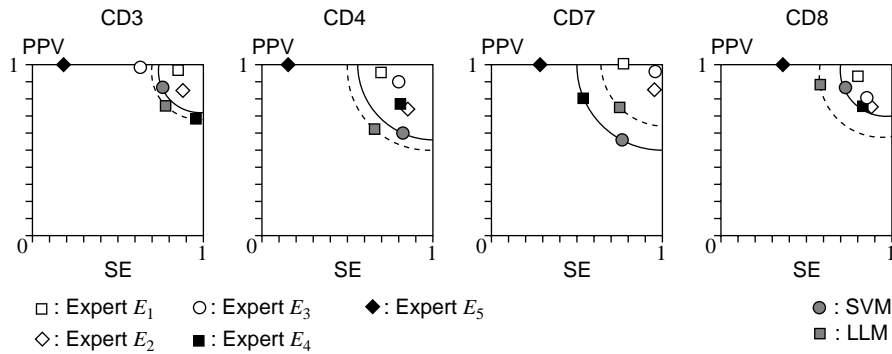


Fig. 5. The SE and PPV are computed for five human experts and two neural network-based systems are plotted for each micrograph (CD3, CD4, CD7, CD8) separately.

6. Results

In Fig. 5 the ROC plots, based on the generated gold standard G_O , are displayed. All human experts show a satisfactory accuracy, except the overly cautious expert E_5 , which was to be expected. The most experienced expert E_1 performs with highest accuracy for the CD3 and CD8 micrographs, which are of similar type concerning contrast and number of cells. In micrograph CD4 and CD7 he loses to expert E_3 , in micrograph CD7 also to expert E_2 . But across all four micrographs expert E_1 and E_2 show the most constant performance, because E_3 and E_4 have outliers in CD3 or CD7, respectively. Thus, E_1 's expertise is reflected by the stability of his evaluation performance. Among the two classifiers, the SVM-based algorithm outperforms the LLM classifier for images CD3, CD4 and CD8. For CD7, which is the micrograph of lowest contrast showing a large number of cells (> 300), it gets a much lower PPV value than the LLM classifier, caused by a large number of false positives. Comparing the SVM with the human experts, the data shows that—for the micrographs CD3 and CD8—it equals or outperforms up to three of the human experts.

7. Discussion

We draw two conclusions from the data: First, ANN-based classifiers can be trained and applied to microscopy data with satisfactory results. That means they can achieve the performance of a medium-trained human expert. If one considers a real world laboratory scenario with large sets of image data, the most experienced expert is much too expensive for an exhaustive visual inspection of all the data. In a best case, a medium-skilled expert would evaluate the data, needing at average 1 h per micrograph. Our algorithms need much less time, which is 30 min for the SVM and less than 2 min for the LLM.

Second, the performances of the classifiers show that, if the microscopy technique can be advanced to record micrographs of more constant quality, the SVM classifier should be applied in the future. If micrographs of low contrast like CD7 are exceptional, the SVM computes the more accurate data. For micrographs of varying quality, the LLM is the better choice, because it is more resistant to variations in the image domain, which is clearly shown in our experimental data.

8. Summary

For several years, growing attention has been paid to new methods of optical microscopy and immunofluorescence labeling in the field of functional proteomics. This has created a demand for efficient (semi-)automatic evaluation of large sets of fluorescence micrographs. Because of the inherent problems of noise, variation in size and shape of cells and artefacts, this task of micrograph evaluation is very demanding for a human expert or for a signal processing system. This leads to the necessity of a measurement of the accuracy of a micrograph evaluation.

The micrographs, in this study, show invasive lymphocytes in tissue, labeled following a highly standardized protocol with different antibody markers. Each micrograph shows between 100 and 400 cells and is of size 658×517 pixel.

Four exemplary micrographs were blindly evaluated by five experts. The experts selected positions of cells using an interactive computer program. For each micrograph, five lists of (x, y) -positions were recorded.

Because the lists show strong discordance among the experts, for each micrograph the five lists are merged to one simulated OR-expert. This expert reflects the correct cell positions best, as stated by the biomedical experts, that overlooked a large number of cells themselves.

The cell detection algorithm processes one micrograph in two steps. In the first step, a trained classifier computes for each pixel a confidence value that represents the degree of belief, that a fluorescent cell is centered at this position. The whole confidence values of all points give a new image, the confidence map. In this map, large values indicate positions of cells. In the second step, the confidence map is post-processed to compute the cell positions, using a local maximum criterion and a threshold.

For automated classification we used two different ANN classifiers. Those are one variant of the SOM and one SVM.

The neural network-based cell detection is performed on a representative set of images with both classifiers in comparison, both with fixed parameters for all images. For each micrograph, the two result lists of cell positions, together with the five human result lists are compared to the OR-expert gold standard. Two ROC-motivated measures are computed for each evaluation, the SE and the PPV. These measures characterize the evaluation accuracy most appropriate.

Analysis of the SE- and the PPV-plots reveals that, for a system with considerable variation in image quality, the SOM-based classifier leads to better results than the SVM. Compared to the human experts, the algorithmic approaches equal the accuracy of a medium-skilled expert.

In standard laboratory work, the large sets of micrograph data cannot be evaluated by the most experienced expert, because it would be too expensive. Thus, the proposed approach is definitely of high practical relevance.

References

- [1] D.L. Taylor, E.S. Woo, K.A. Giuliano, Real-time molecular and cellular analysis: the new frontier of drug discovery, *Curr. Opin. Biotech.* 12 (2001) 75–81.
- [2] D.L. Taylor, M. Nederlof, F. Lanni, A.S. Waggoner, The new vision of light microscopy, *Am. Sci.* 90 (4) (1992) 322–335.

- [3] S.H. Ong, X.C. Jin, R.S. Jayasooriah, Image analysis of tissue sections, *Comput. Biol. Med.* 26 (1996) 269–279.
- [4] D. Comaniciu, P. Meer, D.J. Foran, Image-guided decision support system for pathology, *Mach. Vision Appl.* 11 (1999) 213–224.
- [5] S. Shiffman, E.H. Shortliffe, Biomedical imaging and the evolution of medical informatics, *Comput. Med. Imag. Graph.* 20 (4) (1996) 189–192.
- [6] J.S. Duncan, N. Ayache, Medical image analysis: progress over two decades and the challenges ahead, *IEEE Trans. PAMI* 22 (2000) 85–105.
- [7] W. Schubert, Antigenic determinants of *t*-lymphocyte $\alpha\beta$ receptor and other leucocyte surface proteins as differential markers of skeletal muscle regeneration: detection of spatially and timely restricted patterns by MAM microscopy, *Eur. J. Cell Biol.* 58 (1992) 395–410.
- [8] W. Schubert, Molecular semiotic structures in the cellular immune system: key to dynamics and spatial patterning? in: W.Z.J. Parisi, S.C. Mueller (Eds.), *A Perspective Look at Nonlinear Media in Physics, Chemistry and Biology*, Lecture Notes in Physics, Springer, Berlin, 1997.
- [9] P. Sjöström, B. Frydel, L. Wahlberg, Artificial neural network-aided image analysis system for cell counting, *Cytometry* 36 (1999) 18–26.
- [10] M.V. Boland, R.F. Murphy, A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of hela cells, *Bioinformatics* 17 (2001) 1213–1223.
- [11] M.P.S. Brown, W.N. Grundy, D. Lin, N. Cristianini, C. Sugnet, T.S. Furey, M.J. Ares, D. Haussler, Knowledge-based analysis of microarray gene expression data using support vector machines, *Proc. Natl. Acad. Sci.* 1 (1997) 262–267.
- [12] J. Quackenbush, Computational analysis of microarray data, *Nat. Rev. Genetics* 2 (2001) 418–427.
- [13] B. Moghaddam, M.-H. Yang, Sex with support vector machines, in: T.K. Leen, T.G. Dietterich, V. Tresp (Eds.), *Advances in Neural Information Processing Systems*, Vol. 13, MIT Press, Cambridge, MA, 2001, pp. 960–966.
- [14] H. Ritter, Learning with the self-organizing map, in: T. Kohonen (Ed.), *Artificial Neural Networks*, Vol. 1, Elsevier, Amsterdam, 1991.
- [15] T. Kohonen, Self-organized formation of topologically correct feature maps, *Biol. Cybernetics* 43 (1982) 59–69.
- [16] T. Kohonen, *Self-Organization and Associative Memory*, Springer, Berlin, 1989.
- [17] D. Rumelhart, G. Hinton, R. Williams, Learning internal representations by error propagation, in: D.E. Rumelhart, J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition 1: Foundations*, 1986, pp. 318–362.
- [18] T.W. Nattkemper, H. Ritter, W. Schubert, A neural classifier enabling high-throughput topological analysis of lymphocytes in tissue sections, *IEEE Trans. ITB* 5 (2001) 138–149.
- [19] T. Kämpfe, T.W. Nattkemper, H. Ritter, Combining independent component analysis and self-organizing maps for cell image classification, *Proc. DAGM Lecture Notes in Computer Science* no. 2191 (2001) 262–268.
- [20] T. Bell, T. Sejnowski, An information-maximization approach to blind source separation and blind deconvolution, *Neural Comput.* 6 (1995) 1129–1159.
- [21] J. Platt, Fast training of support vector machines using sequential minimal optimization, in: B. Schölkopf, C. Burges, A. Smola (Eds.), *Advances in Kernel Methods—Support Vector Learning*, MIT press, Cambridge, MA, 1998.
- [22] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy, Improvements to Platt's SMO algorithm for SVM classifier design, Technical Report TR CD-99-14, National University of Singapore, 1999.
- [23] T. Twellmann, T.W. Nattkemper, W. Schubert, H. Ritter, Cell detection in micrographs of tissue sections using support vector machines, in: *Proceedings of the ICANN: workshop on Kernel & Subspace Methods for Computer Vision*, Vienna, Austria, 2001, pp. 79–88.
- [24] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, *Neurocolt Technical Report nc-tr-98-030*, Royal Holloway College, University of London, UK, 1998.
- [25] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, *Knowledge Discovery Data Min.* 2 (1998), pp. 121–167.
- [26] N. Cristianini, J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge University Press, Cambridge, 2000.
- [27] C. Metz, ROC methodology in radiologic imaging, *Invest. Radiol.* 21 (1986) 720–733.
- [28] L. Lusted, Signal detectability and medical decision making, *Science* 171 (1971) 1217–1219.

- [29] L. Lusted, Decision making studies in patient management, *N. Engl. J. Med.* 284 (1984) 416–424.
- [30] L. Lusted, ROC recollected, *Med. Decision Making* 4 (1984) 131–135.
- [31] H.M. Zweig, G. Campbell, Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, *Clin. Chem.* 39 (1993) 561–577.
- [32] Y.Y.Y. Shimosato, K. Yamagishi, Experience and present status of telepathology in the national cancer center hospital, *Zentralbl. Pathol.* 138 (1992) 413–417.
- [33] A. Ackerman, Discordance among expert pathologists in diagnosis of melanocytic neoplasms, *Hum. Pathol.* 27 (1996) 1115–1116.
- [34] F.J.W.-M. Leong, A.K. Graham, T. Gahm, J.O. McGee, Telepathology: clinical utility and methodology, in: D.G. Lowe, J.C.E. Underwood (Eds.), *Recent Advances in Histopathology*, Vol. 18, Churchill Livingstone, London, 1999.
- [35] J. Viedma Contreras, Leucocyte activation markers in clinical practice, *Clin. Chem. Lab. Med.* 37 (6) (1999) 607–622.

Tim Wilhelm Nattkemper studied mathematics, physical chemistry and computer science at the University of Bielefeld, Germany. After post-graduate work for the Neuroinformatics Group and the Department of Macro-Economic Dynamics at the University of Bielefeld, he received a Ph.D. in Computer Science at the university of Bielefeld in 2001. He is a member of the Neuroinformatics Group, the graduate program “Structure Formation Phenomena” and the “International Graduate School in Bioinformatics and Genome Research”. Since 2002 he is head of the newly founded group “Applied Neuroinformatics” at the Bielefeld University, Faculty of Technology.

In his research T.W.N. is developing methods for image analysis, multi-parameter visualization and datamining in bio-medicine using neural networks. This includes the analysis of fast learning networks for the detection and segmentation of cells in tissue samples, text categorization, and 3D segmentation/visualization. Additionally, he is a co-worker in a project for the analysis of multi-parameter protein vectors of cells by sonification methods.

Thorsten Twellmann received the diploma in Computer Science from the University of Bielefeld, Germany in 2001. He is currently pursuing a Ph.D. program in Computer Science at the University of Bielefeld, working within the Neuroinformatics Group. His main research activities are in kernel-based machine learning and multi-parameter image processing.

Helge J. Ritter studied physics and mathematics at the Universities of Bayreuth, Heidelberg and Munich and received a Ph.D. in Physics from the Technical University of Munich in 1988. Since 1985, he has been engaged in research in the field of neural networks. In 1989 he moved as a guest scientist to the Laboratory of Computer and Information Science at Helsinki University of Technology. Subsequently, he was assistant research professor at the then newly established Beckman Institute for Advanced Science and Technology and the Department of Physics at the University of Illinois at Urbana-Champaign. Since 1990, he is professor at the Faculty of Technology, Bielefeld University, where he is head of the neuroinformatics research group.

Helge Ritter’s main interests are principles of neural computation, in particular self-organizing and learning systems, and their application to machine vision, robot control, intelligent man–machine interfaces and data analysis.

Walter Schubert has studied medicine at the universities of Aachen and Bonn, Germany, where he was graduated in 1976 and received a doctoral degree in 1977 on the basis of an experimental doctoral thesis in the field of neurobiology of the epiphysis testis axis. After post-graduate work in surgery and internal medicine, and military service as medical officer he specialized in neurology and psychiatry from 1979 to 1985 and was leader of the laboratory for neuromuscular disorders at the university clinic of neurology, Bonn, Germany from 1983 to 1988. In 1988 he joined as a research associate the Center of Molecular Biology (ZMBH), biological faculty, at the university of Heidelberg, Germany.

Since 1994 he leads as associate professor for medical neurobiology the Molecular Pattern Recognition Research Group at the Otto-von-Guericke-University of Magdeburg, Germany. The scientific work concerns the analysis of function of protein systems. Results of scientific activities comprise the detection of molecular markers of muscle regeneration, a role for the Alzheimer amyloid precursor proteins in synapses, and the development of a new molecular and robotic technology to analyze the functional information within the cells proteome. This technology is the first to enable continuous “reading” of large protein networks in single cells.