Learning-based mitotic cell detection in histopathological images

Christoph Sommer
Institute for Biochemistry, ETH Zürich
christoph.sommer@bc.biol.ethz.ch

Luca Fiaschi
HCI Heidelberg
luca.fiaschi@iwr.uni-heidelberg.de

Fred A. Hamprecht
HCI Heidelberg
fred.hamprecht@iwr.uni-heidelberg.de

Daniel W. Gerlich
Institute for Biochemistry, ETH Zürich
present address: daniel.gerlich@imba.oeaw.ac.at

Abstract

Breast cancer grading of histological tissue samples by visual inspection is the standard clinical practice for the diagnosis and prognosis of cancer development. An important parameter for tumor prognosis is the number of mitotic cells present in histologically stained breast cancer tissue sections.

We propose a hierarchical learning workflow for automated mitosis detection in breast cancer. From an initial training set a pixel-wise classifier is learned to segment candidate cells, which are then classified into mitotic and non-mitotic cells using object shape and texture features. Our workflow banks on two open source biomedical image analysis software: "ilastik" and "CellCognition" which provide a user user friendly interface to powerful learning algorithms, with the potential of making the pathologist work an easier task.

We evaluate our approach on a dataset of 35 high-resolution histopathological images from 5 different specimen (provided by International Conference for Pattern Recognition 2012 contest on Mitosis Detection in Breast Cancer Histological Images). Based on the candidate segmentation our approach achieves an area-under Precision-Recall-curve of 70% on an annotated dataset, with good localization accuracy, little parameter tuning and small user effort. Source code is provided.

1. Introduction

The standard approach for breast cancer diagnosis relies on visual inspection of histopathological samples stained with hematein and eosin (H&E). Among measurable characteristics, the mitotic cell index is currently the best single predictor for long-term progno-

sis [3] for breast carcinomas. Manual counting of mitotic cells is a time consuming and tedious task; it is limited by subjectivity and hence prone to low reproducibility [7]. Despite recent advancements in semi-automatic and automatic methods, the manual counting of mitotic cells is still today the standard procedure in clinical practice.
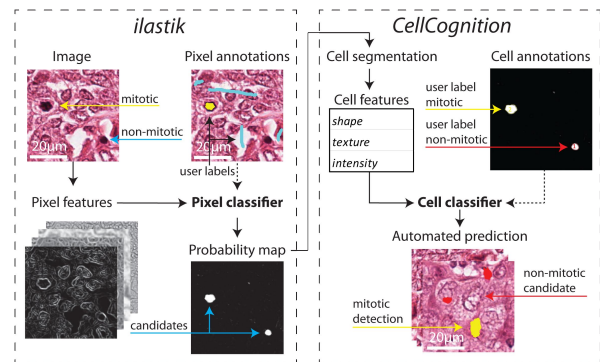
Figure 1. Overall workflow of the mitotic detection system.

Several automatic methods based on image analysis have been proposed. The majority of these methods employs cell segmentation as a first step followed by statistical analysis. Cell segmentation is achieved by traditional image processing techniques such as thresholding [10], watershed [11], morphological operations [12], and active contour models [6]. Then, segmented cells are classified based on previously annotated examples [4, 5]. In a clinical setting, however, many of these methods are of limited use because they require careful parameter tuning, complex user interactions and lack an open source, ready to use, implementation.

In order to overcome these shortcomings, our

learning-based framework addresses the challenging problem of mitotic cell counting demonstrating accurate results, and also facilitates a pathologist's analysis banking on the usage of two user friendly and open-source software packages: "ilastik" [14] and "CellCognition" [9]. In particular we combine these two flexible tools and cast the problem into a hierarchical workflow based on two steps. On the first level, using ilastik, the image pixels are classified on the base of their surrounding local intensities and their gradient information. Then, candidate cells are segmented with a robust threshold of the pixel probability map. On the second level, using CellCognition, object features expressing global properties of the cells, such as shape and texture, are used to classify each segment as mitotic or non-mitotic. The combined workflow is illustrated in Figure 1.

Our paper is organized as follows: in the next section we introduce the benchmark dataset. Section 2 describes pixel-wise classification and cell segmentation, while details on the final candidate classification are given in Section 3. The performance of our workflow is demonstrated in Section 4.

1.1. Benchmark dataset

The MITOS-dataset [13] is published for the *International Conference for Pattern Recogniton* 2012 contest on *Mitosis Detection in Breast Cancer Histological Images* and provides a set of 5 breast cancer biopsy slides. The slides are stained with H&E. In each slide the pathologists selected 10 images (termed high power fields in the contest description) at $40\times$ magnification. A image has a size of $512\mu m \times 512\mu m$. For each image, all mitotic cells were manually labeled by two expert pathologists. In total the entire dataset contains more than 300 mitoses in 50 images (35 images released as training data). The dataset covers all phases of mitosis, yielding high variability of shapes and texture. The dataset has been recorded by three different scanners termed A, H, and a 10 bands multi-spectral microscope M in the contest description. According to the contest regulation, we opt to perform our experiments using only the raw data (RGB images) generated from scanner A. Figure 2 (A) shows examples of mitotic cells, including prometa, meta and anaphase for scanner A. The images in (B) depict two regions of the raw data where no mitotic cell was labeled. Cell nuclei in the input RGB-images appear as spotted structures dark purple in color. Other surrounding structures such as stroma appear red and fat tissue white.
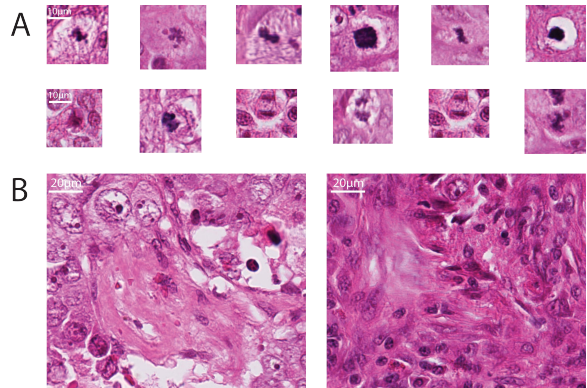
Figure 2. The MITOS-dataset: (A) selected mitotic examples from scanner A including examples from prometa, meta adatas-dfnd anaphase. (B) Raw data from two different specimen containing no mitotic cell.

2. Segmentation of candidate cells

On the first workflow level we seek to detect and segment cell candidates from the heterogeneous background. To avoid complex user interaction for object detection we choose a simple bottom-up approach. A pixel classifier is learned to produce pixel probability maps for candidate cells. These are then thresholded to segment cell candidates. We exploit ilastik [14] for this step since it offers several advantages: an efficient and fully multi-threaded implementation, high accuracy with little parameter tuning, and a graphical user interface with interactive feedback during labeling.

The pixel-wise mitotic ground-truth annotations (provided with the dataset) are imported as foreground (i.e. candidate) labels. Background pixel labels (non-candidate) are obtained by manual annotation. We used the graphical user interface of ilastik to interactively add labels by giving brush strokes to regions containing no candidate cells (see Figure 1).

A number of standard image analysis filters, over varying spatial scales, are used as features. These filters comprise Gaussian smoothing (at $\sigma = 1.6, 3.5$), Gaussian gradient magnitude ($\sigma = 1.6$), difference of Gaussians ($\sigma_1 = 1, \sigma_2 = 1.6$), Laplacian of Gaussian ($\sigma = 1.6$), eigenvalues of the Hessian matrix ($\sigma = 1.6, 3.5$), and the eigenvalues of the structure tensor ($\sigma = 1.6, 3.5$). In total, the feature vector for each pixel of a RGB image has a length of 39 (i.e. 13 per color channel).

ilastik uses a random forest [1] classifier for learning. Random forests consist of many decision trees.

During prediction, each pixel is classified by collecting the votes of each individual tree. The ratio of the tree votes is converted into a probability map and provides the input for mitosis detection in the second processing stepusing CellCognition. Local adaptive threshold is applied on the probability maps followed by a split algorithm [15] in order to separate touching cells.

3. Mitosis classification

On the second workflow level, segmented candidate cells are classified into mitotic and non-mitotic. For each segment a set of object descriptors is computed comprising intensity, shape (e.g. circularity) and texture features (e.g. haralick [8], statistical geometric features [2]) (a full list of features can be found in [9]). All features are computed per color channel separately and are concatenated to a feature vector length of 717 (239 per color channel). For the classification, we use a Support Vector Machine (SVM) with an Gaussian kernel. The parameters (i.e. cost and gamma) are optimized by grid-search. We use the graphical user interface of CellCognition to provide the object labels for training. Mitotic labels are imported from the ground-truth, whereas non-mitotic cells are labeled manually.

4. Experiments and Discussion

For the pixel-wise training we select 6 images out of the full training dataset containing 35 images. We label the pixels in two classes: candidate foreground mitotic pixels and candidate non-mitotic background pixels. The graphical user interface of ilastik allows to interactively observe and correct the output of a random forest classifier. Therefore, we adopt an active learning strategy during labeling: first, we import the cells labeled by pathologist as pixels labels for the mitotic class, second, we start subsequently adding new background labels in regions where the response of the classifier is wrong. We qualitatively bias the classifier to have a maximum accuracy and recall for the foreground pixels. In total 32453 pixel are labeled as foreground and 77920 pixels as background.

We train a random forest classifier consisting of 100 decision trees. After training all images are predicted by the random forest and the resulting probability maps are segmented. With this procedure we achieve a 93% pixel recall for the mitotic class.

In the first training round on the cell level classification we annotate 208 mitotic cells according to the ground-truth and select 700 non-mitotic manually from the candidate segmentation by using the cell annotation


Original image

Ground truth

Automated prediction
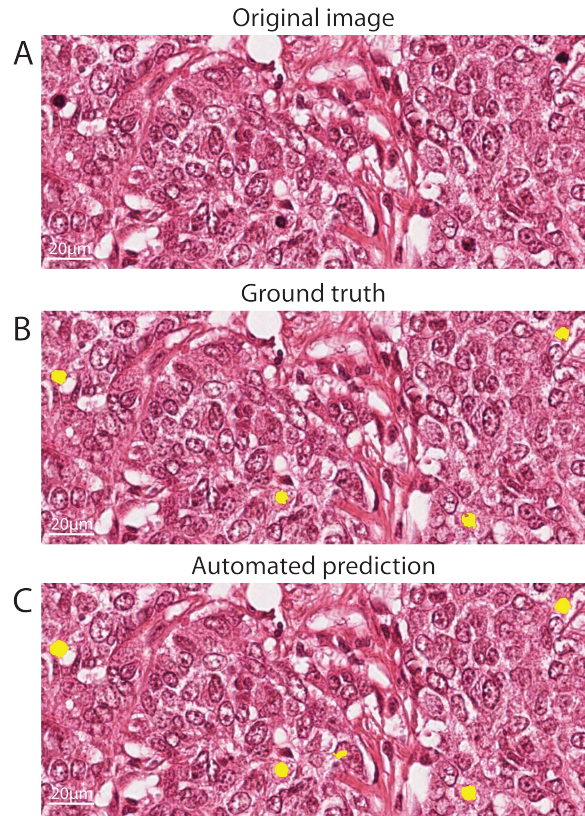
Figure 3. Mitotic cell detection overlayed on test image: (A) raw data, (B) ground-truth annotation by pathologist, and (C) mitotic cell nuclei as found by the proposed method.

tool from CellCognition. In the second training round we include 796 additional and previously false positive detections to the final training set of 1704 examples. We train a SVM with an Gaussian kernel and optimized its parameters (i.e. cost and gamma) by grid-search. Figure 3 (A) shows a selected region of the raw data with its ground-truth annotation in Figure 3 (B). The resulting detection of mitotic cells with three true-positive and one false-positive detections is shown in Figure 3 (C).

For better comparison and future reference we evaluate our performance according to the contest evaluation metric. The overall performance is measured by the area-under Precision-Recall (AUC-PR = 70%) curve. Figure 4 shows the PR-curve. The area under the receiver operating characteristic curve (AUC-ROC) is 0.906 (not shown). We choose a classifier at a threshold of 0.189 which correspond to a recall of 0.798 precision of 0.519 (F-measure of 0.629) to predict the entire set. The pixel accuracy is measured by the mean localiza-

tion accuracy (i.e. displacement of the center-of-mass), the pixel recall and pixel overlap of each detection on a held-out image set. A mitotic detection is matched to the ground-truth if the center-of-mass lies within a radius of $5\mu m$ (≈ 21 pixel). With this we achieve a detection recall of 0.68 and precision of 0.38 with a mean localization accuracy of $0.57 \pm 0.38\mu m$, mean pixel recall of $89 \pm 8\%$ and a mean pixel overlap of $75 \pm 6\%$.
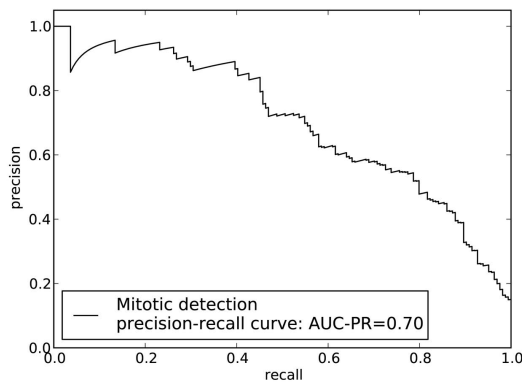
Figure 4. Overall performance for cell classification: PR-curve with 5-fold cross-validation

5. Conclusion

Automated analysis tools are a crucial step towards a more objective basis for tumor prognosis. In this paper we describe an automatic workflow for mitotic cell detection in histopathological images from breast cancer micro-sections. Our approach is based on supervised classification and combines two state-of-art machine-learning methods for pixel and object classification. By combining the open-source software ilastik and CellCognition, we enable pathologists to train an automatic mitosis detection system with small user effort facilitated by convenient graphical user interfaces. The evaluation on the MITOS-dataset shows convincing accuracy with good localization properties given the rather small number of annotated examples.

References

[1] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[2] Y. Chen, M. Nixon, and D. Thomas. Texture classification using statistical geometrical features. In *International Conference on Image Processing*, volume 3, pages 446–450. IEEE, 1994.

[3] F. Clayton. Pathologic correlates of survival in 378 lymph node-negative infiltrating ductal breast carcinomas. mitotic count is the best single predictor. *Cancer*, 68(6):1309–1317, 1991.

[4] E. Cosatto, M. Miller, H. Graf, and J. Meyer. Grading nuclear pleomorphism on histological micrographs. In *Proceedings of the 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008.

[5] J. Dalle, W. Leow, D. Racoceanu, A. Tutac, and T. Putti. Automatic breast cancer grading of histopathological images. In *30th Annual International Conference on Engineering in Medicine and Biology Society*, pages 3052–3055. IEEE, 2008.

[6] H. Fatakdawala, J. Xu, A. Basavanhally, G. Bhanot, S. Ganesan, M. Feldman, J. Tomaszewski, and A. Madabhushi. Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): Application to lymphocyte segmentation on breast cancer histopathology. *IEEE Transactions on Biomedical Engineering*, 57(7):1676–1689, 2010.

[7] T. Fuchs and J. Buhmann. Computational pathology: Challenges and promises for tissue analysis. *Computerized Medical Imaging and Graphics*, 2011.

[8] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(6):610–621, 1973.

[9] M. Held, M. Schmitz, B. Fischer, T. Walter, B. Neumann, M. Olma, M. Peter, J. Ellenberg, and D. Gerlich. Cellcognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature methods*, 7(9):747–754, 2010.

[10] H. Jeong, T. Kim, H. Hwang, H. Choi, H. Park, and H. Choi. Comparison of thresholding methods for breast tumor cell segmentation. In *Proceedings of the 7th International Workshop on Enterprise networking and Computing in Healthcare Industry*, pages 392–395. IEEE, 2005.

[11] A. Mouelhi, M. Sayadi, and F. Fnaiech. Automatic segmentation of clustered breast cancer cells using watershed and concave vertex graph. In *International Conference on Communications, Computing and Control Applications*, pages 1–6. IEEE, 2011.

[12] A. Nedzved, S. Ablameyko, and I. Pitas. Morphological segmentation of histology cell images. In *Proceedings of the 15th International Conference on Pattern Recognition*, volume 1, pages 500–503. IEEE, 2000.

[13] L. Roux, J. Klossa, F. Capron, and M. Gurcan. Mitosis detection in breast cancer. http://ipal.cnrs.fr/ICPR2012/, 2012.

[14] C. Sommer, C. Straehle, U. Koethe, and F. A. Hamprecht. ”ilastik: Interactive learning and segmentation toolkit”. In *8th IEEE International Symposium on Biomedical Imaging (ISBI)*, 2011.

[15] C. Wählby, I. Sintron, F. Erlandsson, G. Borgefors, and E. Bengtsson. Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections. *Journal of Microscopy*, 215(1):67–76, 2004.