Training Based Cell Detection from Bright-Field Microscope Images

Tuomas Tikkanen*, Pekka Ruusuvuori†‡, Leena Latonen‡, and Heikki Huttunen*
*Department of Signal Processing, Tampere University of Technology, Finland
†Tampere University of Technology, Pori, Finland
‡BioMediTech, University of Tampere, Finland

Abstract—This paper proposes a framework for cell detection from bright-field microscope images. The method is trained using manually annotated images, and it uses Support Vector Machine classifiers with Histogram of Oriented Gradient features. The performance of the method is evaluated using 16 training and 12 test images with altogether 10736 human prostate cancer cells. Both the implementation and the annotated image database are released for download. The experiments consider various parameters and their effect on performance, and reaches accurate detection results with cross-validated AUC over 0.98, and mean relative deviation of 9% from manually counted annotations in the growth curve over six days.

Keywords—Cell Detection, Histogram of Oriented Gradients, Support Vector Machine, Growth Curve, Supervised Learning

I. INTRODUCTION

The rate of cell division can be observed with a growth curve, which is a useful statistical tool in cancer treatment research. Growth curve is an empirical model, providing a way to study evolution of cell culture over time in terms of number of cells [1]. The nature of the experiment, requiring days of culturing and multiple imaging rounds, prevents the use of fluorophores, making bright-field microscopy the common choice. For routine use, growth curve analysis requires automated cell counting from microscope images. This task is an important and challenging image processing problem in bioimage informatics [2], which is referred to as cell segmentation or cell detection. Cell segmentation aims at identifying cell boundaries from multi-cell image, whereas in cell detection the object is to detect only the location instead of the area of each cell. The use of bright-field imaging means the cells appear almost transparent and have low contrast, making the analysis challenging, but also offers several benefits such as avoiding toxicity or stress to the cells through fluorophores, and no special optics or equipment needed compared to phase contrast or differential interference contrast microscopy [3].

Several methods exist for counting the cells from microscope images. The difficulty of cell segmentation or detection task, and accordingly, choosing the correct algorithm depends much on the characteristics of cells being targeted. If the cells are well separated from each other and have uniform intensity, simple thresholding or watershed algorithms are popular choices of approach [4], [5]. If the cells are packed together, algorithms which account for cell shape and size are preferred [6]. Another approach uses training-based methods, where examples of cells and background are shown to the detector, which then learns their most important characteristics [7], [8], [9]. Learning-based approaches have been implemented also in some of the publicly available cell image analysis software tools/platforms, such as ICY [10] and ilastik [11]. While such tools enable the use of learning based methods for segmentation, studies on the use of various machine learning methods and feature descriptors will be needed in order to fully capitalize the potential of training based methods in image segmentation.

Our method is a training based approach using Histogram of Oriented Gradient (HOG) features [12] for detecting the cells. The framework consists of calculating the HOG descriptors and applying the support vector machine (SVM) classifier for detection. The HOG feature descriptor was introduced by Dalal and Triggs in 2005 [12] for pedestrian detection. HOG describes features based on local histograms of gradient orientations weighted with gradient magnitudes. Since then, it has been successfully used for various image recognition tasks, such as in face recognition and vehicle orientation detection [13], [14]. The HOG has also been used as the basis of more advanced detection approaches such as the state-of-the-art Latent SVM object detection algorithm [15].

However, the applications of HOG in microscopy have centered around fluorescence imaging [16], [17], [18], and there are not many documented experiments with the use for cell detection from bright-field microscope images. All existing literature on HOG based cell detection from brightfield [19] uses only a small amount of test data. This paper provides a practical evaluation, where the method is trained and applied to a research-scale large collection of human prostate cancer cell images spanning a six day cell culturing experiment.

To summarize, the main contributions of this paper are the following: 1) We propose a HOG-based live cell detector for brightfield microscope images; 2) Our experiments show that our method is superior to a recent state-of-the-art brightfield cell counting method, and the reported accuracies are comparable to those typically obtained using more expensive and destructive imaging modalities; 3) We provide the annotated dataset with 10736 human prostate cancer cells together with the implementation of our HOG detector for public use.

The rest of this paper is organized as follows: In Section II we describe the HOG features and the SVM training procedure. Section III defines the experiment image database. Section IV concentrates on experimental results, and in Section V we discuss the results and the practical applicability of the method.
II. CELL DETECTION FRAMEWORK

The proposed cell detection framework is illustrated in Figure 1. The complete pipeline consists of preprocessing, feature extraction and classification steps. The preprocessing step standardizes the input images by gamma and color normalization. After that the actual HOG features are extracted as in [12]. The gradient orientations are computed and collected into a histogram, which are subsequently grouped together with neighboring histograms to blocks. The block-histograms are finally normalized and lastly classified with the SVM. Our implementation of the HOG pipeline uses evenly spaced undirected orientation bins in the range of $[0^\circ, 180^\circ]$. R-HOG blocks with L2-Hys normalization, linear SVM kernel, and the parameters specified in Table I.

The sliding window procedure was run on single-scale using step size of $2 \times 2$. The detector window size was selected to have aspect ratio and size close to average of those in annotated cells, as presented in Section III. Resizing the sliding window size to various scales only lowered performance in our case. Multi-scale search enables detection of large cells which, however, exist in small numbers. The suggested values were selected based on thorough testing [20] and they serve as a reasonable starting point. The detection accuracy can be improved by increasing the number of orientation bins or by minimizing the cell size to $2 \times 2$.

The proposed cell detection framework consists of separate training and testing phases as illustrated in Figure 1. In the first training iteration, all annotated cells and an equal amount of randomly sampled background samples are extracted from the input images using ground truth annotations. Then, HOG features are generated from each of the cropped images and an initial linear SVM classifier is trained. The resulting classifier from the first iteration tends to have a poor performance due to many false positives. This is natural because the negative samples were randomly sampled, and thus representing mostly "easy cases".

Therefore, the training is continued by subsequent iterations, which search for hard background samples. More specifically, all false positives incorrectly detected by the classifier of the first iteration are appended to the training set as representatives of hard false positives. After this, the classifier is re-trained. Typically only a few retraining iterations are needed to decrease the false positive rate to an acceptable level. In Section IV we iterate until less than 5% of the initial amount of false positives are found. This is a compromise between accuracy and training time, as a strict requirement of zero false positives in the training set would require dozens of iterations with very little effect in the accuracy on the test set.

In the testing phase, unseen images are scanned with sliding window and HOG features are generated from the position of each window in the image. These features are

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptor window size</td>
<td>$32 \times 32$</td>
</tr>
<tr>
<td>R-HOG block size</td>
<td>$8 \times 8$</td>
</tr>
<tr>
<td>Block step size</td>
<td>$4 \times 4$</td>
</tr>
<tr>
<td>Cell size</td>
<td>$4 \times 4$</td>
</tr>
<tr>
<td>Number of orientation bins</td>
<td>9</td>
</tr>
<tr>
<td>Max number of detection window increases</td>
<td>1</td>
</tr>
</tbody>
</table>
classified as cell or non-cell with the linear SVM classifier that was constructed in the training phase. It is likely that multiple positive detections will be clustered, because windows can overlap with each other. Thus, clustered detections are merged together. Finally, the number of cells in the images of each day (growth curve) is estimated.

We implemented the proposed framework in Python on top of the OpenCV platform\(^1\) and the scikit-learn package\(^2\). The HOG features are readily calculated with OpenCV, but to the best of our knowledge, there are no publicly available implementations of the complete SVM training procedure with proper mining for hard examples as described in the original HOG paper [12]. The annotated database and the implementation of the cell detection framework presented in this paper can be downloaded from the supplementary website\(^3\). The implementation enables straightforward extension of the framework with other features available in OpenCV.

III. MICROSCOPE IMAGES

[Images of microscope images]

Fig. 3. Illustration of experimental data: 49 randomly selected cell (a) and non-cell (c) example images, median image of all cell (b) and non-cell (d) examples and visualization of their HOG features. The pixel intensities are scaled for better visualization.

The cell detection framework was evaluated on a data set consisting of grayscale bright-field images of PC3 human prostate cancer cell lines. Images were taken with QImaging Retiga-2000R camera using Olympus IX71 microscope and Objective Imaging Surveyor automated scanning and imaging software. The data set consists of images taken with 25 different focus levels on days 1-6 after passaging. Pixel resolution of the images is 1596 \(\times\) 1196. Fig. 2 presents cropped images of the same area in a sample throughout days 1-6 demonstrating the growth of cancer cell culture. It is worth noting that the cells could move around freely in the samples to some extent.

The proposed method was trained with 16 autofocus images from days 1-6 and tested with separate 12 autofocus images (2 per day). Although the number of images is low, the number cells, which are the actual samples, is relatively high. The total number of 4858 cell and 7198 non-cell training examples were collected for the first SVM training iteration. Training was iterated for 15 iterations until less than 5 % of the initial amount of false positives are found. During the iterations, total of 2564 hard examples were collected for the final classifier. Fig. 3 presents training examples from each class and median images of all training examples of each class with a visualization of HOG features showing the distribution of gradient orientations within HOG cells. The median cell image consists of almost perfectly round object in the middle of the image, whereas the median non-cell image does not show any pattern and consists of low amplitude noise only. All images used in this study, as well as the annotations used by the learning based method, are available on the supplementary site.

Sizes of annotated cancer cells vary in our images within a fairly narrow range: width 39 ± 8 px, height 38 ± 9 px, aspect ratio 0.9 ± 0.4 px. Knowing that each image represents an area of 1190.8 \(\mu\)m \(\times\) 891.4 \(\mu\)m in real life, we get roughly the actual sizes of cells: width 29.10 ± 5.97 \(\mu\)m and height 28.32 ± 6.71 \(\mu\)m. However, it is important to notice though, that the annotations were slightly bigger than the cells, as the median cell image in Fig. 3 demonstrates.

IV. EXPERIMENTAL RESULTS

Iterative training was noticed to lower the number of false positive detections considerably. ROC curves in Fig. 4 a) show how AUC of SVM decreases from 0.997 to 0.982 during 15 iterations. At the same time the number of hard examples decreases from 1064 to 50. The level of generalization of the model gets higher and higher on each iteration as the model is trained with more and more hard examples.

Each day images prefer different detection sensitivities in terms of averaged \(F_1\)-score, shown in graph b) of Fig. 4. As the days progress, more liberal classifier is required. The most probable reason for requiring more sensitive detector is the increased number of partially occluded cells, since the cells overlap with each other increasingly as a function of time. Based on the peak \(F_1\)-scores that are indicated by dashed lines in Fig. 4 b), detection sensitivity of 0.6 was selected to be used when the growth curve in Fig. 4 c) was estimated with HOG. Since the optimal threshold is different for different days, the results could be improved further by choosing the threshold adaptively for each day using cross-validation with the training data. However, this is left for future study.

The estimated growth curves in Fig. 4 c) show that HOG produces more valid cell count estimate than an alternative, recently developed, approach for cell segmentation by Buggenthin et al. [21]. The mean relative deviation over each day with HOG is only 9 %, whereas the method of Buggenthin et al. produces mean relative deviation of 25 %. Additionally, the HOG approach produces less false positives (FP), less false

---

1. http://opencv.org
negatives (FN) and more true positives (TP) each day than that of Buggenthin et al.

We discovered that the accuracy of the algorithm of [21] is heavily dependent on the focus of the input image. In fact, our database contains 25 focus levels for each image, and the most focused image is automatically selected as the input for the HOG method. However, the results with the method of Buggenthin et al. were significantly poorer than those of the proposed method; particularly in cases where a part of the image was out of focus (due to cells stacking on top of each other). Therefore, we tuned the reference algorithm manually by selecting a more suitable focus plane in order to reach a performance comparable to the proposed method. After manual tuning, the focus levels were selected to be the 9th image (of 25) from the top for days 1-5 and the 11th for day 6.

Cell counts in Fig. 4 c) have been normalized according to the day 1 manual counts. The unnormalized manual cell counts in Fig. 4 c) grow almost linearly from 70 to 200 on during days 1-3. After that, the number of cells roughly doubles when measured on days 4 and 5. In the end, the cell quantity grows from ~800 to ~1400 when moved on from day 5 to day 6.

V. CONCLUSIONS

The results indicate that HOG features can be successfully applied to cell detection from bright-field microscope images. Growth curve can be estimated automatically, which agrees favorably with manual counts. The automated algorithm counted the cells in a more objective, consistent and faster manner than manual counting. The iterative training process was noticed to be a crucial step for eliminating false positive detections, producing cross-validated ROC AUC of 0.98.

We also compared the proposed method against a recently published alternative [21]. The results of the proposed method are superior, despite tuning the comparison method by manually selecting the focus level as to maximize the performance.

Perhaps the most important question lies in selecting the correct detection threshold when detecting cells with HOG. The ideal threshold varies from images to images, depending on the cell density and the number of cells in the image.

Therefore, the accuracy could be still improved with cross-validated selection of threshold for each day separately.

In summary, it is possible to implement a complete and robust cell detection framework with HOG features, which yields accurate results with relatively low error rate. Even though bright-field microscope images of prostate cancer cells were specifically used in this study, we believe the detection framework can be successfully applied also to other kinds of microscopy techniques and cell types.

Our future work will involve developing the training based cell detector further, and we will also explore the use of measurements from multiple focus levels. It has been shown (see, e.g., [3], [22]) that the use of images taken with out-of-focus settings in addition to the conventionally used in-focus layer may provide useful information for cell detection and segmentation purposes.

REFERENCES


