SVM-based Human Cell Detection Technique using Histograms of Oriented Gradients

TUDOR BARBU
Institute of Computer Science of the Romanian Academy
T. Codrescu, 2, cod 700481, Iasi
ROMANIA
tudbar@iit.tuiasi.ro http://www.iit.tuiasi.ro/~tudbar

Abstract: - An automatic human cell detection system is proposed in this article. Histograms of Oriented Gradients are used for cell feature extraction. A robust search procedure, using variable sized sliding windows, is performed for cell localization in grayscale images. The proposed sliding-window based search algorithm is used in combination with a non-linear SVM-based feature vector classification technique. Some successful experiments are described in this paper.

Key-Words: - object-class detection, cell detection, HOG features, window-based search algorithm, SVM classifier

1 Introduction

Object-class detection represents an important computer vision domain that deals with identifying instances of semantic objects of a certain class in digital images and video sequences [1]. The most object-class detection techniques identify popular classes of objects, such as humans, faces, cars, buildings or animals.

These object detection methods use various image features, such as Wavelet-based features [2], Haar-like characteristics [3] or HOG-based features [4]. Support Vector Machines [5] or Boost classifiers [6], like AdaBoost and GentleBoost, are usually used by these approaches in the classification stage.

We approached the object detection domain in our previous works. Thus, several skin detection [7], face detection [8] and person detection techniques [9] have been proposed by us. In this paper we focus on another human-related class of image objects: human cells [10]. A cell detection method is thus provided.

Histograms of Oriented Gradients (HOG) represent a powerful tool for performing human detection [4]. We use them successfully for cell featuring in this work. Our HOG-based feature extraction is described in next section.

Then, a SVM-based classification approach is proposed in the third section. A sliding-window based search algorithm is used in combination with the SVM classifier for cell localization. The proposed search procedure is presented in the fourth section.

The performed cell detection experiments and method comparisons are discussed in the fifth section of this paper. The conclusions are elaborated in the last section, and the paper ends with a list of references.

2 HOG-based Feature Extraction

As mentioned in the introduction we consider a HOG-based cell feature extraction. Histogram of Oriented Gradients represents a robust feature descriptor used in computer vision area for object detection. They prove to be very useful for pedestrian detection [4].

We compute HOG characteristics for the subimages corresponding to cells' bounding boxes. A HOG-based feature vector is modeled for such an image.

First, one determines the image gradient values, representing directional changes in the intensity or color. The gradient vector is formed by combining the partial derivatives of the image I in the x and y directions:

$$\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right), \quad (1)$$

The gradients in the two directions can be computed by applying the $1D$ centered, point discrete derivative mask in the horizontal and vertical directions:

$$\begin{cases} \dfrac{\partial I}{\partial x} = I * [-1\ 0\ 1] \\ \dfrac{\partial I}{\partial y} = I * [-1\ 0\ 1]^T \end{cases} \quad (2)$$

The gradient orientations of the image are computed as $\theta = \arctan\left(\dfrac{\partial I}{\partial x}, \dfrac{\partial I}{\partial y}\right)$. The image I is then divided into cells, and for each cell, one computes a local $1D$ histogram of gradient directions (orientations) over the pixels from that cell [4,11].

We consider 9 bins for the local histogram. The histogram channels are evenly spread over 0 to 180 degrees, so each histogram bin corresponds to a 20 degree orientation interval. The obtained cell histograms are then combined into a descriptor vector of the image.

First, these cells have to be locally contrast-normalized, due to the variability of illumination and shadowing in the image. That requires grouping the cells together into larger, spatially-connected blocks. Once the normalization is performed, all the histograms can be concatenated in a single feature vector, representing the HOG descriptor.

We use [3x3] cell blocks of [6x6] pixel cells with 9 histogram channels. The feature vector of the image is computed as its HOG descriptor, with 81 coefficients. This could be expressed as $V(I) = HOG(I)$.

3 A SVM Classification Approach

The Support Vector Machines (SVM) represent supervised machine learning models, widely used in object detection tasks [5]. They can perform efficiently both linear and non-linear feature vector classification. Using a set of training examples, each marked as belonging to one of two classes, a SVM training algorithm predicts, for each given input, which of two possible classes has to represent the correct output [12].

Obviously, the object-class detection task can be formulated as a binary linear classifier problem. It can be solved by considering two classes, objects and non-objects, and applying a SVM model on them.

We develop a non-linear SVM training model for human cell detection, because non-linear SVMs are consistently found to be better suited for the image object detection task. Thus, a large enough training set is constructed first. Our training data consists of two subsets: the set of positives, containing bounding boxes of biological cells, and the set of negatives, containing non-cell images of various sizes.

A training set example is depicted in the next two figures. Its positive samples are displayed in Fig.1, while the negative samples are displayed in Fig.2.

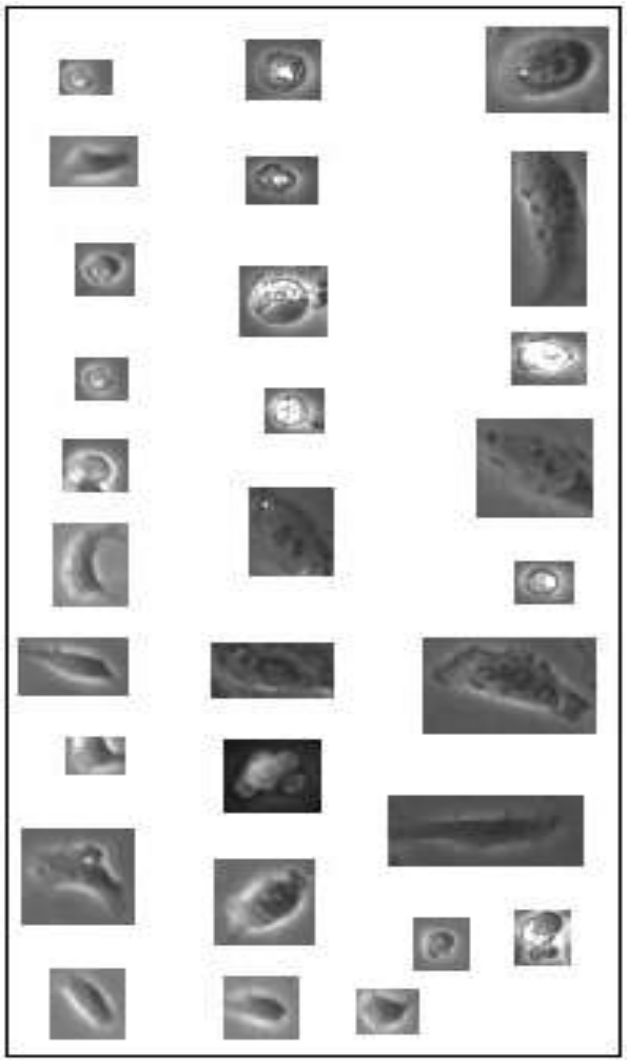

Fig. 1. The positive training sub-set

The training set could be expressed in the following form:

$$S = \{(T_1, x_1), \ldots, (T_n, x_n)\},\ x_i \in \{-1,1\} \quad (3)$$

where T_i represent the training samples and their labels are:

$$x_i = x(T_i) = \begin{cases} 1, \text{if } T_i - positive \\ -1, \text{if } T_i - negative \end{cases} \quad (4)$$

The HOG-based feature extraction described in previous section is performed on these training subsets, the feature training set of the system thus

being obtained. It is composed of the computed feature vectors $V(T_i), i = 1,...,n$.

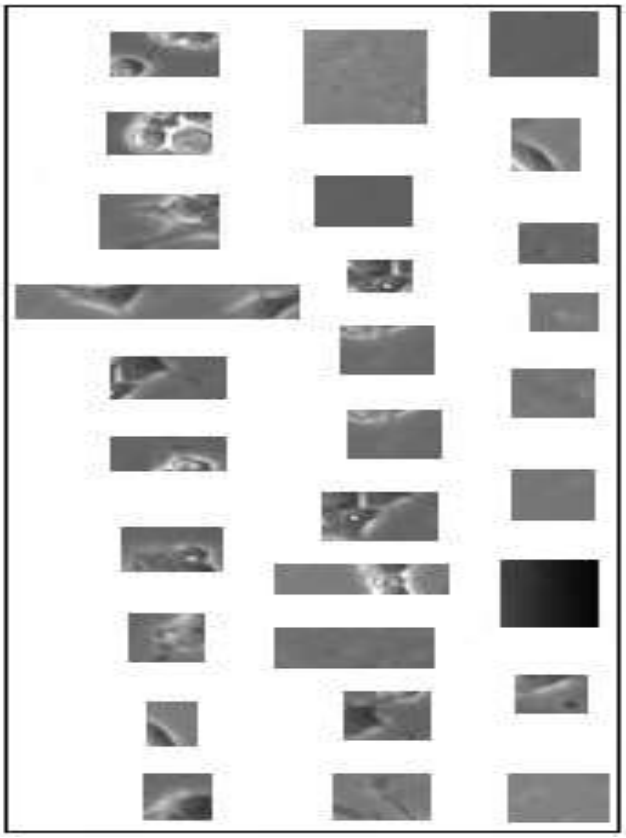

Fig. 2. The negative training sub-set

One has to identify the maximum-margin hyperplane that divides the objects characterized by $x_i = 1$ from those having $x_i = -1$. We consider a quadratic programming technique for separate hyperplane identification.

Also, our non-linear classification algorithm uses a nonlinear kernel function instead of the *dot product* used in the linear case. Thus a quadratic kernel is considered to map the training data *S* into the kernel space.

If this SVM classifier is applied to an input object represented by its bounding image *I*, the SVM-based classification of its feature vector will produce the object labeling. That can be expressed formally as:

$$x(I) = SVM\ (V(I)) \qquad (5)$$

4 A Novel Sliding-window based Search Algorithm

One has to solve the following computer vision task: locate all objects representing human cells from a given image *Im*. We detect the bounding boxes of those cells by performing a sliding window scanning over the grayscale version of the respective image [13].

Thus, we propose a cell search algorithm that scans *Im* using a variable sized sliding window. At each step, one computes the HOG-based feature vector of the sub-image corresponding to the current window. The SVM classification algorithm is applied to the feature vector and, as a result, the analyzed sub-image is labeled as either positive or negative.

If a positive is identified, the next search should be performed far enough from it. For this reason the positions of pixels corresponding to the detected cell are marked as already visited. That will reduce the complexity of the scanning process.

We consider that the widths and heights of the cells vary between some minimum and maximum values. Therefore, the sliding-window size has to vary accordingly. Let us note these values as following: $w_{\min}, w_{\max}, h_{\min}, h_{\max}$.

Let us introduce the following notation: *Subim*(*Im, x, y, w*$_1$, *w*$_2$, *h*$_1$, *h*$_2$) representing the set of all subimages of *Im* having the upper-left pixel *Im*(*x,y*), the width in interval [*w*$_1$, *w*$_2$] and the height in the interval [*h*$_1$, *h*$_2$]. This set of subimages can be determined recursively.

The proposed searching algorithm is expressed by the following pseudocode:

```
Detect_cells (Im)
 mark = 0 matrix with the same size [M × N] as Im
 S = φ;
 for i = 1 to M - h_min
  for j = 1 to N - w_min
   if mark (i,j) = 0
    for each I ∈ Subim(Im, i, j, w_min, w_max, h_min, h_max)
     Compute V (I) = HOG(I);
     Apply SVM classifier to V(I) as in (5):
     x(I) = SVM(V(I));
     if x(I) = 1 then
      S = S ∪ {I} ;
      for each pixel location (a,b) ∈ I
          mark(a,b) = 1; (mark location as visited)
      end
     end
    end
   end
  end
 end
 Return S.
End
```

The result of a cell searching process is displayed in Fig. 3. The stem cells from the depicted grayscale image are identified and marked with red rectangles. In that figure, one can see also a blue rectangle marking a subimage representing a non-cell example.

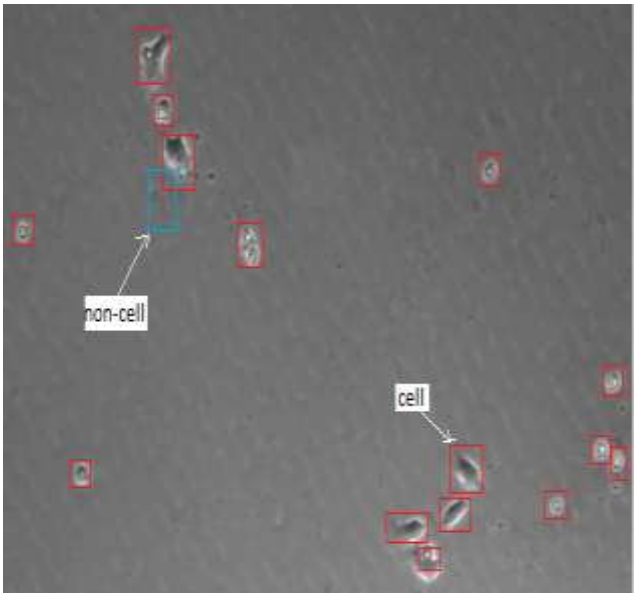

Fig. 3. Example of a cell detection result

5 Experiments

We have performed numerous cell detection experiments using the detection technique provided here. The numerical tests performed on various image datasets, containing various types of cells, have produced satisfactory detection results that prove the effectiveness of our proposed approach. It has achieved a high detection rate, of approximately 90%.

Also, high values are obtained for the performance parameters. Thus, values around 0.9 have been achieved for both the *Precision* and *Recall* parameters. This means that almost all the detected objects returned by our approach are relevant, very few false positives being returned by it. Also, the high *Recall* value indicates that almost all true positives are returned.

We have developed and tested many appropriate SVM training sets, containing cells and non-cell objects. One of those training data sets is described in Fig. 1 and Fig. 2, being composed of 29 positives and 29 negatives.

While the proposed detection technique provides very good cell identification results, it does not execute fast enough. That is because of the high complexity and computational cost of the sliding-window based search algorithm. In our experiments we apply this algorithm with the following parameters: $w_{\min} = 20, w_{\max} = 50, h_{\min} = 20, h_{\max} = 50$ M = 512 and N = 672. Lower values could reduce its complexity, but also its effectiveness. The values of parameters related to HOG computing, described in section 2, also may influence the computational complexity of the detection system.

Method comparisons have also been performed. Our cell searching procedure is considerably faster than methods based on Exhaustive Search (ES), but slower than approaches that do not perform image scanning.

Thus, we have also compared the obtained detection results with those produced by other cell identification techniques that are based on image segmentation, using edge detection or threshold values and some morphological operations [10, 14]. From our tests, we have found that the approach described here is more performant, obtaining a higher *Recall* value, but also it is somewhat slower. While the segmentation-based methods often fail to return all the positives, producing a higher number of missed biological cells, they runs faster than the technique proposed in this paper.

6 Conclusion

An automatic cell detection system has been proposed in this paper. It performs a proper HOG-based feature extraction, then it applies a non-linear SVM-based algorithm, using a quadratic kernel function, in the classification stage.

Obviously, we have approached a particular object-class detection task: cell identification in images. The main purpose of this paper was to prove that the HOG characteristics and the SVMs, used mainly in human detection domain, work satisfactory for other classes of objects too.

For this purpose we have considered some proper parameters for HOG computing and SVM, but more important we have constructed efficient training data sets for cell detection. Modeling an appropriate SVM training set is key to achieve satisfactory object-class detection results.

An important contribution of this article is also the proposed sliding-window based cell searching algorithm. The performed detection experiments prove its effectiveness, but we intend to further improve it, by reducing its complexity and consequently the running time.

The object-class detection approach provided here has also many important application areas, such as medical imaging, cell tracking, human cell database organizing, indexing and retrieval. Besides improving the cell searching procedure, our future research will focus on cell tracking in video

ISBN: 978-1-61804-148-7

sequences [15], which represents the next step beyond cell detection.

Acknowledgments

The research work described here has been supported by the project PN-II-PCE-20113-0027, financed by the Romanian Minister of Education and Technology.

We acknowledge also the research support of the Institute of Computer Science of Romanian Academy, Iasi branch, Romania.

References:

[1] D. Hall. A system for object class detection. *Cognitive Vision Systems*, 2004.

[2] R. N. Strickland, H. IlHahn. Wavelet transform methods for object detection and recovery. *IEEE Trans. Image Proc.*, 6, pp. 724-735, 1997.

[3] R. Lienhart, J. Maydt. An Extended Set of Haarlike Features for Rapid Object Detection. *Proceedings of IEEE ICIP 2002*, Vol. 1, pp. 900-903, Sep. 2002.

[4] N. Dalal, B. Triggs. Histograms of oriented gradients for human detection, *Computer Vision and Pattern Recognition*, San Diego, CA, June 20–25, 2005.

[5] S. Shah, S. H. Srinivasan, S. Sanyal. Fast object detection using local feature-based SVMs, In *MDM*, pp. 1-5, 2007.

[6] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features, In *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

[7] T. Barbu. Automatic Skin Detection Technique for Color Images, *Proceedings of the International Multidisciplinary Scientific Geo-Conference SGEM 2010*, Volume 1, pp. 1047-1052, June 20-25, 2010.

[8] T. Barbu, An Automatic Face Detection System for RGB Images, *International Journal of Computers, Communications & Control*, Vol. 6, No.1, pp. 21-32, 2011.

[9] T. Barbu. Novel Approach for Moving Human Detection and Tracking in Static Camera Video Sequences, *Proceedings of the Romanian Academy*, Series A, Volume 13, Number 3, pp. 269-277, July-September 2012.

[10] R. Koprowski, Z. Wrblewski. Automatic segmentation of biological cell structures based on conditional opening and closing, *Machine Graphics and Vision*, 14, pp. 285-307, 2005.

[11] T. Barbu. Template Matching based Video Tracking System using a Novel N-Step Search Algorithm and HOG Features, *Lecture Notes in Computer Science (ICONIP 2012, part V, Doha, Qatar, Nov. 12-15, 2012)*, Vol. 7667, pp. 328-336, Springer, Heidelberg, T. Huang et al. (Eds.), 2012.

[12] N. Cristianini, J. Shawe-Taylor. *An Introduction to Support Vector Machines and other kernel-based learning methods*, Cambridge University Press, 2000.

[13] D. Hoiem. *Object Category Detection: Sliding Windows, Computer Vision*, CS 543 / ECE 549 University of Illinois, 2011.

[14] S. Beucher, F. Meyer. The morphological approach to segmentation: The watershed transformation, *Mathematical Morphology in Image Processing*, Dougherty, E., ed., pp. 433-481, Marcel Dekker, 1993.

[15] H. Shen, G. Nelson, S. Kennedy, D. Nelson, J. Johnson, D. Spiller, M. White, D. Kell. Automatic tracking of biological cells and compartments using particle filters and active contours, *Chemometrics Intell Lab Syst.*, 82 (1–2), pp. 276-282, 2006.