A Framework for Image Analysis and Object Recognition in Industrial Applications with the Ensemble of Classifiers

Bogusław Cyganek
AGH University of Science and Technology
Department of Electronics
Krakow, Poland
cyganek@agh.edu.pl

Michał Woźniak
Wroclaw University of Technology,
Wybrzeże Wyspiańskiego 27, 50-370
Wrocław, Poland
Michal.Wozniak@pwr.wroc.pl

Abstract

The paper presents a work-in-progress on a classification system for object detection in vision based industrial applications. The main idea of the presented system is application of the ensemble of one-class classifiers trained with specific features of objects of the whole scene. During recognition stage the ensemble tries to recognize the trained patterns. The system is well suited for parallel processing and therefore it can operate in real-time conditions. Preliminary experiments in automotive applications show promising results.

1. Introduction

Visual signals in different spectra are frequently used for object detection and recognition in industrial applications, such as line production surveillance, automotive, underwater, etc. However, despite the acquisition system, what really counts is an accurate and fast response of such systems [9][13][5][14].

In this paper a system of cooperating classifiers is proposed in various industrial applications aimed at object detection and scene analysis based on visual signals in different spectra. Such system of cooperating base classifiers is called an ensemble of classifiers. However, despite its relatively weak member classifiers, thanks to their cooperation and diversity, it was shown that such ensembles achieve superior results when compared to even complex but single classifiers [12][10]. In our proposition for base classifiers we use the one-class support vector machines (OC-SVM) [15], which are versions of the well known SVMs originally proposed by Cortes and Vapnik [3].

The specificity of the one-class classifiers is that they can deal with data of only one class. Such situations arise frequently if exemplars of other classes are too numerous or unknown. As an example can serve fault detection systems in which frequently only data of the normal conditions of operation are available since acquiring data of faulty operation would demand destroying of the system. Such situation is depicted in Figure 1. A task here is to detect objects which in some sense are different than the other for which the system was trained.

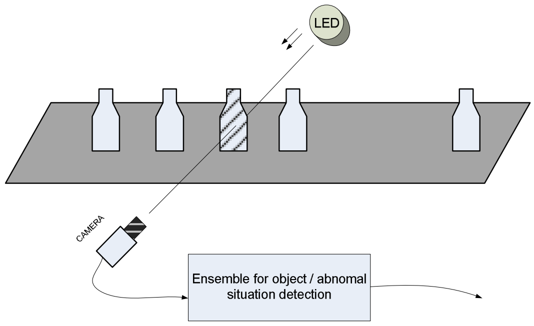

Figure 1. An example of a line production under surveillance of the computer vision system.

Another example is object detection in digital images in which only a model of an object is available, whereas other are unknown. Such situations are frequently encountered in real applications. An example here is the seat occupation detection system for proper passenger airbag deployment, discussed in Section 4.

The presented here methods is based on our previous works on the one-class pattern classification, scene analysis, as well as on image segmentation [6][7].

Last but not least, in this paper serial and parallel implementations of the proposed systems are discussed. This addresses the problem of real-time processing, which frequently is a necessary assumption in real industrial applications.

2. System Description

As already mentioned, the ensemble can be trained with data specific to an object of interest. However, each member classifier of the ensemble receives different chunk of data. These can be obtained with the feature bagging, clustering, or spatial partitioning of an object. Such a situation is depicted in Figure 2. In this case it is assumed that an object is represented by a set of sparse but salient points and features gathered around these.

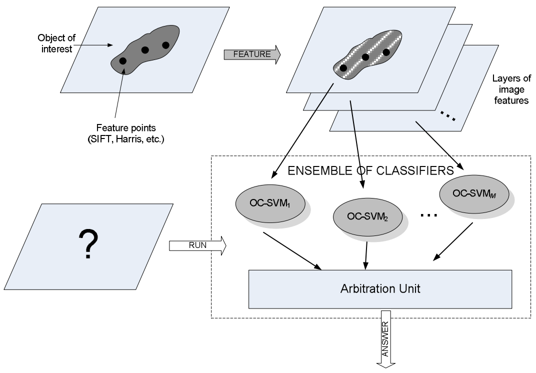

Figure 2. Specific object detection in a scene based on the ensemble of classifiers trained with different object features. Black dots denote salient points.

Yet another mode of operation is presented in Figure 3. In this case, the goal is to encode a scene of interest by an ensemble of ensemble of OC-SVM classifiers. More concretely, at first the input image is divided into a predefined number of tiles. Then, in each tile features are detected, which are then used to train the ensemble assigned to that tile. Once again, type of features, as well as the way of their distribution between member classifiers of the ensemble (such as bagging or clustering), are application dependent as will be discussed.

Also, the tiles can be obtained not in an arbitrary fashion described above, but for example from the prior image segmentation.

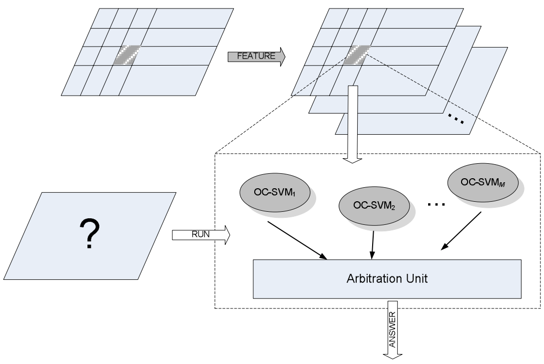

Figure 3. Scene representation with a set of ensembles of classifiers. Training images are split into tiles from which characteristic features are extracted. Each set of features of each tile is used to train its specific ensemble. The whole scene is represented by a set of ensembles, each responsible for a single tile.

Answer of the system depends on the aforementioned type of operation. In the first case, after detection of salient points and feature extraction, the ensemble provides its response to each possible configuration. This way an object can be detected. In the second mode, shown in Figure 3, an unknown image is fed to the ensembles which provide their answer on a percentage points which belong to specific one-class definitions [6].

3. Structure of the Ensemble of Classifiers for Object Detection and Scene Representation

In this section building blocks of the ensemble of classifiers are presented and discussed.

3.1. Image Feature Extraction

As already pointed out, feature extraction plays the major role in object/scene definition. Depending on an application, features can be prepared from object prototypes either as dense data or sparse representations around salient points. These, in turn, can be obtained from the Harris corner detector or more advanced methods, such as SIFT, etc. In either case, the final features are specific to an object of interest. These should be as discriminative as possible. The well known methods are orientation, color, texture histograms, log-polar representation, just to name the few [9][4][5][13].

3.2. One-Class Base Classifiers

As alluded to previously, for base classifier the OC-SVM are assumed [15][7]. These classifiers are characteristic of many useful properties from which the most important are listed below:

- Only one-class data are necessary for training.
- After training, the whole class is represented exclusively by the not so numerous support vectors (data can be discarded).
- Kernel mapping allows better separation of the class of interest.
- Very fast response time (real-time operation).

Figure 4 depicts the hypersphere enclosing the training class data. It can be characterized by its centre **a** and a radius r. The volume of the hypersphere, which is proportional to r^n, should be as minimal as possible to tightly encompass the data.

The above requirement boils down to the minimization with respect to r^2. In effect, the minimization functional Θ is as follows [15]

$$\Theta(\mathbf{a}, r) = r^2 \tag{1}$$

with the constraint

$$\underset{i}{\forall}: \quad \left\| \mathbf{x}_i - \mathbf{a} \right\| \le r^2 , \tag{2}$$

where $\mathbf{x}_i$ are data points. To be more realistic, the slack variables ξ_i are introduced to allow points distant further than r (the outliers).

This is introduced to **(1)** as follows

$$\Theta\left(\mathbf{a},r\right)=r^2+C\sum_i \xi_i \tag{3}$$

with the new constraints

$$\forall_i : \ \left\|\mathbf{x}_i-\mathbf{a}\right\|\le r^2+\xi_i\,, \qquad \xi_i\ge 0\ ,. \tag{4}$$

In the above, C denotes a parameter that controls the optimization process. The larger C, the less outliers are possible, at the larger volume of the hypersphere.

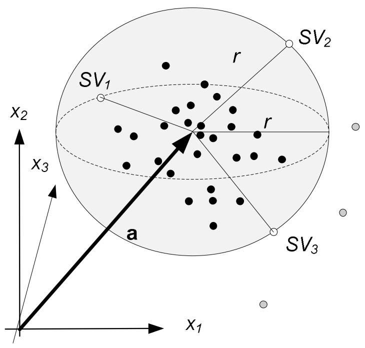

Figure 4. Hypersphere enclosing data points from a single class (black dots). Support vectors are the points on the surface (white). Points outside are outliers (gray).

Given a set of training points $\{\mathbf{x}_i\}$, solution to the equations (3) and (4) can be computed with the Lagrange multipliers [3]. From this a distance d from the centre $\mathbf{a}$ of the hypersphere to a test point $\mathbf{x}_x$ can be found [15].

During classification it can be assumed that an unknown point $\mathbf{x}_x$ is classified as belonging to the class enclosed by this hypersphere if the following is fulfilled

$$d^2\left(\mathbf{x}_x,\mathbf{a}\right)\le r^2, \tag{5}$$

After solving (3) and (4), it can be shown that the above classification rule can be expressed as follows

$$\sum_{i\in Idx(SV)}\alpha_i K\left(\mathbf{x}_x,\mathbf{x}_i\right)\ge\sum_{i\in Idx(SV)}\alpha_i K\left(\mathbf{x}_s,\mathbf{x}_i\right)=\delta\,. \tag{6}$$

where $Idx(SV)$ denotes a set of computed support vectors (see Figure 4) and α_i are scalar coefficients returned by the training procedure. K denotes a kernel function, which in our system are the Gaussian kernels, as follows

$$K_{RBF}\left(\mathbf{x}_i,\mathbf{x}_j\right)=e^{-\gamma\left\|\mathbf{x}_i-\mathbf{x}_j\right\|^2}, \tag{7}$$

where γ denotes a spread parameter. The right side of (6) is constant in the recognition stage. Thus it can be precomputed to a value δ which denotes a cumulative kernel-distance of a SV to all other SVs. Equation (6) is used to test a pattern $\mathbf{x}_x$ if it belongs to a class represented by a set of SVs. Thus, if real-time operation is required, equation **(7)** needs fast implementation, as will be discussed.

3.3. Ensemble Arbitration

The role of the arbitration unit is to orchestrate partial responses of the member classifier and provide a single answer out of the system. The simplest versions of this process that were used in our experimental systems relied on the majority and the weighted majority voting. Nevertheless, more advanced options are also possible. For details see [10][11].

4. Case Study

In this section we discuss applications of the presented system in the automotive domain.

These are as follows:

- Skin segmentation for driver behavior monitoring
- The seat occupation detector

Figure 5 depicts some results of application of the ensemble of OC-SVMs to the first task.

Input color image　　　Face segmentation map

Figure 5. Human skin detection with an ensemble of classifiers for automotive applications.

The classifiers were trained with skin color samples. The system can operate in real time and in difficult lighting conditions encountered in the cars.

Figure 6 presents a system of seat occupation detection. Its main task is to assess parameters for proper airbag deployment in a car depending on measurements of seat occupation. In this application a person is observed by a NIR camera, although some simpler detectors can be also used [1].

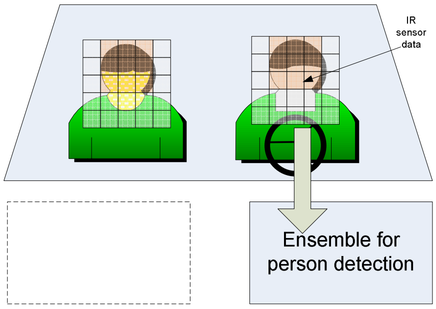

Figure 6. An example of seat occupation detection for proper passenger airbag deployment.

Similarly to the skin detection, this system is trained to detect NIR signals which allow detection of a seating person.

5. Implementation Issues

The described system was implemented in C++ with help of the HIL library [5] and LIBSVM [2]. The experiments were carried out on the system with the Intel® quad-core processor i7 Q820 with a clock 1.73GHz and 8GB of the system RAM, and Windows 64-bit OS. The conducted experiments show that application of more than one classifier in an ensemble leads to the higher accuracy.

To allow real-time operation the part of the software implementing equations (6) and (7) were ported to the graphics card with help of the CUDA environment [16]. Thanks to this a speed-up of over two orders of magnitude was obtained, as reported in [8]. Also a hardware FPGA implementation is relatively straightforward, although in this case the fixed-point arithmetic should be employed [17].

6. Conclusions

The paper presents a work-in-progress on classification system for object detection and scene analysis in visual industrial applications. The main building block of the system is the ensemble of one-class classifiers trained with the specific features of an object of the whole scene. During recognition stage the ensemble tries to recognize the trained patterns. As shown, the proposed system is well suited for parallel processing thanks to which it can operate in real-time conditions. Preliminary experiments in automotive applications show promising results and real-time operation. Further research will concentrate on application of the system to other vision based tasks.

Acknowledgment

This work was financially supported in part by the funds of the Polish National Science Centre NCN, contract no. DEC-2011/01/B/ST6/01994. The work was also supported by the statutory funds of the Department of Systems and Computer Networks, Wroclaw University of Technology.

References

[1] L. Buydens, "Enhancing automotive safety through advanced IR sensor technology", *EETimes Europe*, pp. 30-32, September, 2012.

[2] C-C. Chang, C-J. Lin, "LIBSVM, a library for support vector machines", available at the Internet address http://www.csie.ntu.edu.tw/~cjlin/libsvm, 2001.

[3] C. Cortes and V. Vapnik, Support vector machines. *Machine Learning*, Vol. 20, 1995, pp. 273-297

[4] B. Cyganek, "Object Detection and Recognition in Digital Images. Theory and Practice", Wiley, 2013.

[5] B. Cyganek, J.P. Siebert, "An Introduction to 3D Computer Vision Techniques and Algorithms", Wiley, 2009.

[6] B. Cyganek and K. Wiatr, "Image Contents Annotations with the Ensemble of One-Class Support Vector Machines", *International Conference on Neural Computation Theory and Applications*, France, 2011

[7] B. Cyganek, "One-Class Support Vector Ensembles for Image Segmentation and Classification". *J. of Math. Imaging & Vision*, 42(2-3), Springer, pp. 103–117, 2012.

[8] B. Cyganek, K. Wiatr, "Design of parallel architectures of classifiers suitable for intensive data processing". The ninth international conference on Networking and Services, Lisbon, Portugal, pp. 14-19, 2013

[9] E. R. Davies, Machine Vision: Theory, Algorithms, Practicalities. Morgan-Kaufmann, 2005.

[10] L.I. Kuncheva, Combining Pattern Classifiers. Methods and Algorithms. Wiley Interscience, 2005.

[11] M. Kurzyński and M. Woźniak, Combining classifiers under probabilistic models: experimental comparative analysis of methods. *Expert Systems*, Vol. 29, No. 4, pp. 374–393, 2012.

[12] Polikar R. "Ensemble Based Systems in Decision Making", *IEEE Circuits and Systems Magazine*, Vol. 6, No. 3, pp. 21–45, 2006.

[13] W. E. Snyder, H. Qi, "Machine Vision", Cambridge University Press, 2010.

[14] R. Tadeusiewicz "Place and Role of Intelligent Systems in Computer Science. *Computer Methods in Materials Science*, Vol. 10, No. 4, pp. 193-206, 2010.

[15] D.M.J. Tax, R.P.W. Duin, "Support Vector Data Description", *Machine Learning* 54, pp. 45-66, 2004.

[16] www.nvidia.com

[17] www.xilinx.com