Customized Descriptor for Various Obstacles Detection in Road Scene

Haythem Ameur
Laboratory of Micro-Optoelectronics and Nanostructures
(LMON), University of Monastir,
Monastir, Tunisia

Abdelhamid Helali
Laboratory of Micro-Optoelectronics and Nanostructures
(LMON), University of Monastir
Monastir, Tunisia

J. Ramírez
2Dept. of Signal Theory, Networking and Communications,
University of Granada,
Granada, Spain

J. M. Gorriz
2Dept. of Signal Theory, Networking and Communications,
University of Granada,
Granada, Spain

Ridha Mghaieth
Laboratory of Micro-Optoelectronics and Nanostructures
(LMON), University of Monastir
Monastir, Tunisia

Hassen Maaref
Laboratory of Micro-Optoelectronics and Nanostructures
(LMON), University of Monastir
Monastir, Tunisia

Abstract—Recently, real-time object detection systems have become a major challenge in the smart vehicle. In this work, we aim to increase both pedestrian and driver safety through improving their recognition rate in the vehicle's embedded vision systems. Based on the Histogram of Oriented Gradients (HOG) descriptor, an optimized object detection system is presented in order to achieve an efficient recognition system for several obstacles. The main idea is to customize the weight of each bin in the HOG-feature vector according to its contribution in the description process of the extracted relevant features. Performance studies using a linear SVM classifier prove the efficiency of our approach. Indeed, based on the INRIA datasets, we have improved the sensitivity rate of the pedestrian detection by 11% and the vehicle detection by 5%.

Keywords—ADAS; customized HOG; linear SVM; obstacle detection

I. INTRODUCTION

Nowadays, pattern recognition has become an interesting task in several applications such as Advanced Driver Assistance System (ADAS) specifically for pedestrian and vehicle detection. The need for such system is motivated, unfortunately, by the number of pedestrians killed in road accidents each year. With 1.25 million deaths, each year [1], the World Health Organization describes traffic accidents as one of the major causes of death and injuries around the world.

To enhance pedestrian's safety and prevent vehicles collision, several pairs of sensors were used in ADAS applications such as, camera and RADAR [2], camera and LIDAR [3], thermal camera [4], stereovision [5], [6], etc. Most of the ADAS are based on one vision sensor with generally another active sensor. Currently, the recent advances in image resolution and power computing platforms, computer vision systems are becoming increasingly available for ADAS. Some new high-end cars are already equipped with several on-vehicle sensors to prevent danger cases. In this context, our application is integrated in order to detect and recognize different obstacles in an urban environment and aimed at helping drivers to see the road environment and reduce traffic accidents with an automotive monocular camera.

Pedestrian and vehicle detection tasks have dominated the recent works in ADAS. They represent the most complex objects, since they have a significant inter-variability in the shape, size, color, and appearance found in typical driving scenarios. This type of obstacle has made the detection process a major challenge so far. Consequently, it is necessary to investigate more powerful feature extraction methods to address the obstacle recognition challenge. The main idea in this work is to build a dedicated descriptor for each type of obstacle without changing the process of recognition. Personalizing the parameters of single descriptor to extract features and recognizing several type of objects, makes it possible to gain in speed and area consumed by implementation process.

The structure of the detection task for typical computer vision systems using a monocular camera is illustrated in Fig. 1. In the obstacle detection chain, images are acquired through a camera: a sliding window function scans the entire image and generates several sub-windows named Regions of Interest (ROIs). First, the descriptor extracts the significant features, namely shape, local distribution of gradient intensity and edge directions presented in each sub-window. Second, it supplies the classifier to decide whether the desired obstacle is present or not for each ROI.
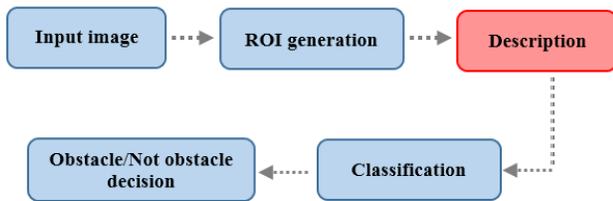
Fig. 1. Obstacle detection chain for typical computer vision systems.

Our obstacle detection process is included in the conventional passive supervised machine learning. Supervised learning method takes a known dataset (images in our case) and known responses to the data named labels (positive examples/negative examples), and tries to build a predictive system which can be used for mapping a new unknown image. In literature, various combinations of descriptor/classifier pairs are commonly used to recognize a special obstacle. Furthermore, we find some descriptors that are more suitable to characterize an object among others. We mention for examples, the HOG descriptor developed by N. Dalal and B. Triggs [7] for pedestrian detection, Haar wavelets by viola et al.[8]for face detection, the LBP descriptor T. Ojala et al. [9] characterized by their low computational cost, and finally the combination between several descriptors as in [10]-[12].

According to the state of the art, the two best known classifiers are Adaboost and Support Vector Machine (SVM). Several weak classifiers are combined into a stronger one to define an Adaboost classifier, while the SVM constructs one or a set of hyperplanes in a high dimensional space in order to achieve a good separation (the largest margin) between the positive and negative training dataset. In more general overviews, most of the proposed works that focus on an obstacle detection system (based on supervised learning machine) combine the HOG descriptor with the SVM classifier [13]-[17] and the Haar features with AdaBoost classifier [18]-[20]. These combinations achieve a better result owing to the logarithmic adaptation between the constituents of every pair. In this paper, we will be interested by the pedestrian and vehicle detection at once; something that is not enough developed in recent works. We will use a modified HOG [21] with a linear SVM as a descriptor / classifier pair to detect and identify the desired obstacles.

This paper, present an analyzing and customizing of the HOG model presented by N. Dalal et al. [7] in order to create a dedicated descriptor for each type of obstacle without changing the process of recognition. The remaining of this paper is organized as follows: in the next two sections, we summarize some related works and describe briefly the computation steps of the standard HOG descriptor. The proposed customized HOG for each obstacle will be presented and discussed in Section 4. Experimental results for pedestrian and vehicle recognition are given and discussed in section 5. Finally, Section 6 will conclude this work.

II. RELATED WORK

Numerous studies have been conducted in order to address the detection of pedestrians, vehicles, road signs or other objects that can be presented in a road scene. But, only a few of them considered the detection of various obstacles at once by the same technique, specifically for a pedestrian and vehicle detection tasks. In this work, we will focus on the pedestrian and vehicle detection problem simultaneously, something that has not been explored enough in recent works on computer vision systems.

Over the past few years, several feature extraction processes have been done. We will mention a few of them based on the HOG descriptor for pedestrian and vehicle detection. The HOG descriptor has been initiated by N. Dalal et al. [10], it is a powerful feature extraction method dedicated for the human shape. A modified approach proposed by G. Ballesteros et al. [22] yielded a reduced set of HOG features. In this way, the dimensionality of the feature vector was decreased significantly. The mechanism proposed by Zhang et al. [23] adapts the cell size in the descriptor entry by a limiting ratio (length / width = 2), then each image is divided in 8×16 cells per average. Jia et al. [24] integrated the HOG descriptor in the Viola's face detection Framework (viola et al. [8] at the end to achieve the descriptor effectiveness and the Framework speed. X. Wang et al. [25] combined a Local Binary Patterns descriptor (LBP) with the HOG algorithm to define a new descriptor called HOG-LBP. The performance of their algorithm for the pedestrian detection exceeds that of standard HOG. Q. Zhu et al. [26] developed a real-time system by integrating a cascade of rejectors with HOG features to achieve a fast and accurate human detection system. In [27] a new descriptor called Scale Space Histogram of Oriented Gradients (SS- HOG) was considered. The authors have used the multiple scale property to describe an object.

Influenced by the high performance reported by the HOG descriptor, some other research have considered the advantages and extracted features for other objects like face, head, bicycle, car, etc. Some works related to vehicle detection are mentioned below. A typical preceding system for vehicle detection using a standard HOG descriptor and SVM classifier has been presented by M. Ling et al. [28] and X. Li et al. [29]. While, Arróspide et al. [30] have proposed an HOG-like gradient-based descriptor for vehicle verification with an exploitation of the known rectangular shape of vehicle rears. To detect vehicles in videos, a combination of Haar-features and HOG-features has been presented by H. Youpan et al. [31]. The authors have expressed that their method can classify and detect the vehicles in multi-orientations with good classification results. The same procedure was proposed by P. Negri et al [32], but with a comparative study between the Haar-like features, the HOG features and their fusion. The results show that the fusion combines the advantages of the first two detectors. Known that the standard parameters of HOG are optimized for human recognition, a re-optimization of the HOG parameters for vehicle detection has been presented by G. Ballesteros et al. [22]. They have tested various combinations in their experiments, and the results show that $[-\pi, \pi]$ as orientation range, (n=4) as the number of cells, (p=16) as the number of orientation bins and a nonlinear kernel on SVMs are the most suitable choice for vehicle detection.

In this paper, an innovative technique is proposed to customize the standard HOG for each obstacle and then a comparison between our approach and other works will be presented.

III. OVERVIEW OF HOG FEATURES DESCRIPTOR

The HOG-features extraction approach could be used to describe a specific gradient orientation in local parts of the image. Such algorithm calculates the gradient direction in small areas of an image, then it assembles the information obtained from all regions into a single vector. N. Dalal et al. [7] have subdivided the image into regions of 8 × 8 pixels that are named cells. Indeed, the HOG feature extraction method consists of calculating the cell-histogram vectors (each vector contains 9 bins and represent the histogram of orientated gradients in one cell), then concatenating them in a single vector. To increase immunity against light variations and lighting conditions, the authors in [7] have normalized all 2 ×2 neighboring cells (which were called a block) to an L2-norm using the following equation:

$$V \rightarrow \frac{v}{\sqrt{\|v\|^2 + \varepsilon^2}} \qquad (1)$$

where *V* is the normalized vector, *v* is the non-normalized vector and ε is a very small constant.

The final HOG feature vector is the collection of the normalized vectors for all the blocks, with an overlapping of 50% per cell. Considering a sliding window of 64 × 128 pixels presented in Fig. 2, it contains 7 ×15 blocks. The assembly of normalized vectors for all blocks into a single 1-D vector then gives 3780 components (36 × 7 × 15 = 3780). The first observation reveals that this feature extraction is a dense representation that maps local image regions to high-dimensional feature spaces. They will be used to train a linear SVM classifier.
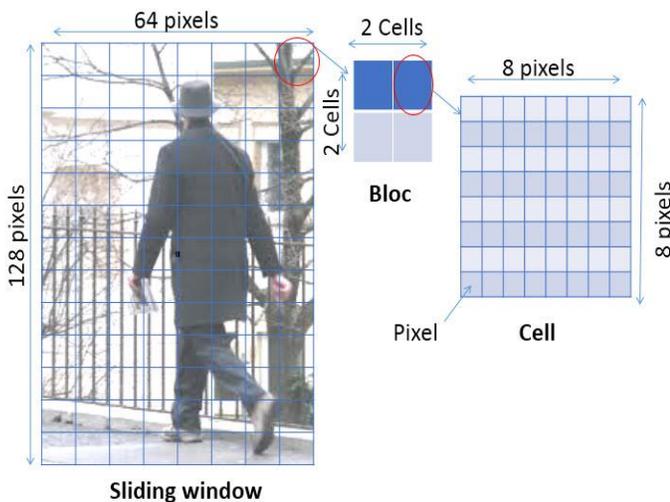

Fig. 2. Overview of HOG feature configuration: window, blocks and cells configuration.

A. Gradients and oriented gradients computation

Gradient computation is the first step to extract the HOG features. To calculate the pixel gradients, several techniques have been previously presented. Among these techniques, the use of a centered derivative mask [-1, 0, 1] turns out to be the best result [7]. The application of the selected gradient operator provides the edge intensity and orientation value for each pixel. The horizontal gradient *dx(x,y)* and vertical gradient *dy(x,y)* of the pixel *I(x,y)* are calculated through equations (2) and (3), while the magnitude M(x,y) is calculated through equation (4).

$$dx(x, y) = \begin{bmatrix} -1, & 0, & 1 \end{bmatrix} * I(x, y) \qquad (2)$$

$$dy(x, y) = \begin{bmatrix} -1, & 0, & 1 \end{bmatrix} * I(x, y) \qquad (3)$$

$$M(x, y) = \sqrt{dx(x, y)^2 + dy(x, y)^2} \qquad (4)$$

Furthermore, the gradient orientation θ(x,y)is given by equation (5):

$$\theta(x, y) = \arctan(\frac{dx(x, y)}{dy(x, y)}) \qquad (5)$$

B. Spatial / Orientation Binning

The histograms show the partition of the orientated gradient elements over the cells. In [7], authors have divided the gradients orientation [0°-180°] "unsigned gradient" into 9 intervals with the same range (20° for everyone) as showing in Fig. 3.
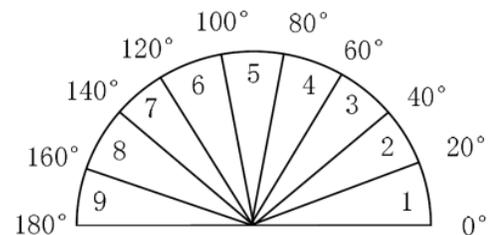

Fig. 3. Orientation range of each bin.

Each interval will be represented by a bin that codes the occurrence's frequency of the gradients orientation in a cell. In practice, each pixel in the cell contributes with a vote to the two closest histogram channels, weighted according to the gradient magnitude at the location (x, y). To summarize, the histogram of oriented gradients is a histogram of neighborhood pixels according to their gradient orientation and is weighted by their gradient magnitude.

C. SVM classifier

By using Supervised Learning Machine in obstacle detection systems, the common characteristics of the samples belong to the same class (training phase) can be determined, which allows the system to subsequently recognize the class of a new unknown sample (decision phase). The SVM

classifier belongs to the class of Supervised Learning Machines. . Such algorithm tries to build an optimal hyper-plane in order to separate the examples of two different classes during the learning phase (*B.E. Boseret al [33]* and *V. Vapnik [34]*. Thus, the decision is made using the previously constructed hyper-plane.

Considering the following set of learning examples and associated class labels $\{X_k, Y_k\}$

Wher,e X_k denote the HOG vectors and $Y_k \in \{-1, 1\}$ the class labels.

Initially, the method ensures the transformation of X_k in a larger space using a kernel function $\varphi(x)$. Then it tries to find a decision function which is given by equation (6):

$$f(x) = w * \varphi(x) + b \qquad (6)$$

Where, the decision function $f(x)$ is optimal in the sense that it maximizes the distance between the nearest point $\varphi(x_i)$ and the hyper-plane. The class label of the HOG vector is then obtained by considering the sign of $f(x)$. Solving the optimization problem is obtained by using the following equations:

$$\min_{w\xi} \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{m}\xi_i \qquad (7)$$

subject to the constraints:

$$y_i(w^T x_i + b) \geq 1 - \xi_i \qquad \text{for } i = (1 \dots m) \qquad (8)$$

Where, the variables are ξ_i known as slack variables. The regularization parameter C is a positive constant that controls the relative influence of the two competing terms. In our experiments, we will use the linear SVM as our binary classifier due first to the large number of HOG features (one may not need to map data to a higher dimensional space) and second to its faster computation.

IV. IMPROVEDHOG APPROACH

In the following subsections, we describe in depth the complete framework of the proposed detection system. The main goal of our approach is to increase the accuracy of the road-obstacle detection system. Our study has presented an improvement for the two most common obstacles in the road (pedestrian and car), but not only limited to these two types. In fact, the method can be applied to other road obstacles such as; buses, bikes, animals (dogs, cats, antelope…) or to recognize the traffic signs. Pedestrians and cars are the most complex obstacles for the detection and identification task, due to their change in appearance and position previously mentioned. The steps involved in the proposed approach are the following.

Firstly, we modified the histogram building method of the standard HOG algorithm to get an average histogram of oriented gradients for each selected eigenvector. Secondly, we apply a new procedure to extract the bins that better

characterize the desired object features. Finally, we amplify the selected bins in the new customized HOG algorithm that will be included in the main chain of the vision system. A general overview of the complete framework can be seen in Fig. 3. More explanations for each phase are presented in the following subsection (Fig. 4).
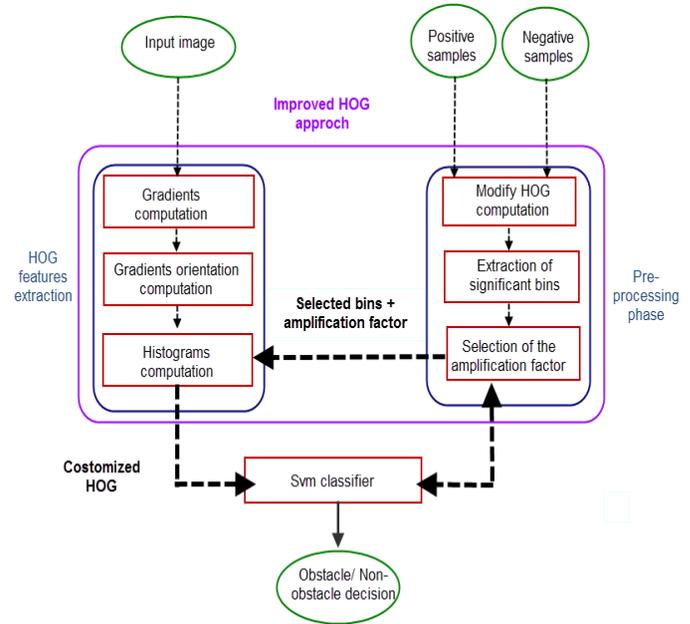

Fig. 4. Whole system of the customized HOG approach.

A. Modified computational method

The histograms that better characterize the desired object are selected by adding all the obtained cell-histogram vectors instead of its concatenation as in the original HOG algorithm. This process gives an average vector that containing 9 bins in the whole image instead of 3780 components. Based on the adopted Dalal's approach [10], the local normalization block was maintained in order to guarantee the immunity against lighting conditions. Nevertheless, the overlap of the cells was removed, having negligible effect in this stage.

Therefore, the first phase provides 9-components mean vector X1, characterizing the image of the object to be detected. Regarding the large inter-variety between pedestrians, we must now generalize this vector through averaging it in the whole dataset that contains n pedestrian images. This step can be obtained through equation (9).

$$X = \frac{\sum_{k=1}^{n} X_k}{n} \qquad (9)$$

where *X* represents the mean vector of the whole positive database, X_k represents the vector for the image number *k* and *n* is the total number of the positive examples in the dataset.

MIT CBCL and INRIA pedestrian datasets are the two most commonly used databases in the field of computer vision machine for the pedestrian recognition task. All of them can

be publicly accessed. In our experiments, the mean vectors were obtained by averaging all the pedestrian images of the entire INRIA and MIT datasets. The same procedure was performed for the calculation of the average vector for the negative examples (not pedestrian images) in the INRIA dataset (all training and test negative examples). Fig. 5 illustrates the mean vectors calculated for the two datasets (pedestrian and non-pedestrian images).

B. Extraction of the significant bins

At present, we have two main vectors that define a pedestrian image and a random image through 9 bins for each one. Thereafter, we calculated the difference between the two histograms in order to extract the most frequent orientations presented in pedestrian images. The result vector is represented in Fig. 6.
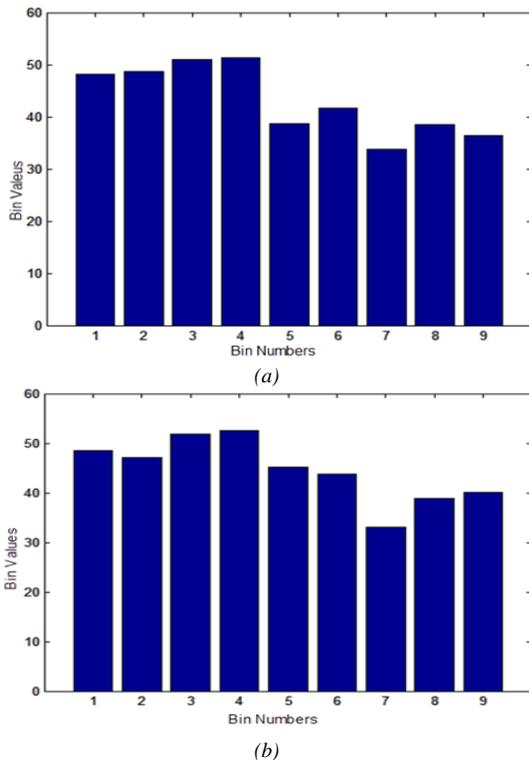

(a)

(b)

Fig. 5.   Mean vectors of the used dataset: (a) Mean vector of the pedestrian images, (b) Mean vector of the random images.
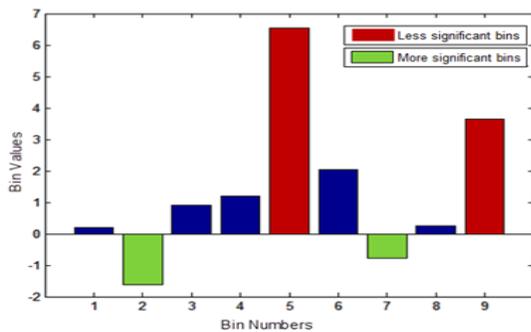

Fig. 6.   Difference between the two mean vectors of the used datasets.

As shown in Fig. 6, the subtraction between the mean vectors gives two special bins (2 and 7) whose values are reversed when compared to other bins. Indeed, these bins have larger gradient density in a pedestrian image than a random image in traffic environments. In other words, these bins encode the edge orientations that describe the shape of the human bodies. Thus, we called them the most significant bins. On the other hand, the bin numbers 5 and 9 represent the highest values in this histogram. These bins encode the least frequent oriented gradients for pedestrian images. Then, we called them the less significant bins. In the last phase of our proposed algorithm, a modification will take place in the vote partition of the oriented gradient elements that may be very promising.

C. Amplifying the extracted bins

The main idea is to amplify the most significant bins using an alpha parameter ($\alpha > 1$) in the cell-histograms building step. The physical significance of this amplification is to highlight the contrast of the contour for some specific orientations that describes the shape of the human body. Actually, the different bins of the HOG-feature vector will not share the same weight, and an amplification factor will be distributed for each bin in order to increase the weight bins that describe the relevant obstacle features.

V.   EXPERIMENTALRESULTS AND DISCUSSION

In order to well assess the measures performance, the experimental results are evaluated based on the three statistical measures test of a binary classification: Accuracy, sensitivity and specificity. Accuracy measures the proportion of actual positives and negatives samples which are correctly identified. Sensitivity measures the proportion of actual positive samples which are correctly identified (e.g. the percentage of pedestrian images which are identified as a true pedestrian image). Specificity measures the proportion of negative samples which are correctly identified (e.g. the percentage of non-pedestrian images that are identified as a true non-pedestrian image). Their expressions are:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (10)$$

$$Sensitivity = \frac{TP + TN}{TP + FN} \qquad (11)$$

$$Specificity = \frac{TP + TN}{TP + FP} \qquad (12)$$

where *TP* is the number of true positives; number of pedestrian images correctly classified; *TN* is the number of true negatives; number of non-pedestrian images correctly classified; FP is the number of false positives; number of pedestrian images classified as non-pedestrian; *FN* is the number of false negatives: number of non-pedestrian images classified as pedestrian.

A. Pedestrian detection

In this section, the impact of the proposed algorithm is analyzed. The used datasets to evaluate our approach are INRIA [35] and MIT [36]. The first dataset contains 2416 positive examples (1208 pictures with their reflections of

horizontal axis) and 1218 negative examples. It contains pedestrians in various postures, clothing as well as wide variety of backgrounds and lighting condition. This makes it one of the most complex databases for pedestrian detection. The MIT dataset contains only positive examples (709 images). The bodies of pedestrians are centered and they have almost the same size in the image. Additionally, a pedestrian is shown alone in a front or rear position. Therefore, these characteristics make the MIT dataset less complicated than the INRIA dataset. Fig. 7 shows some images from the datasets.

The fusion of the two databases in our learning system provides a greater efficiency in the general detection system, which is assessed in terms of 2% increase in the recognition rate of the system.

1) Alpha parameter study

The system has been trained with INRIA and MIT training datasets and tested with INRIA Test Dataset. The global recognition system, presented in Fig. 8, was increased by amplifying the most significant bins extracted in the first phase (bin No 2 and 7). This comes as expected due to the amplification of the characteristics concerning the pedestrian shape in the image. By varying the value of the amplification factor α, the sensitivity rate changes significantly, while the specificity still globally unaffected.

On the other hand, the amplification and the attenuation of the least significant bins (bin No 5 and 9) reduce respectively the recognition rate by 1.5% and 0.5%. Therefore, we have maintained their values without modifications.
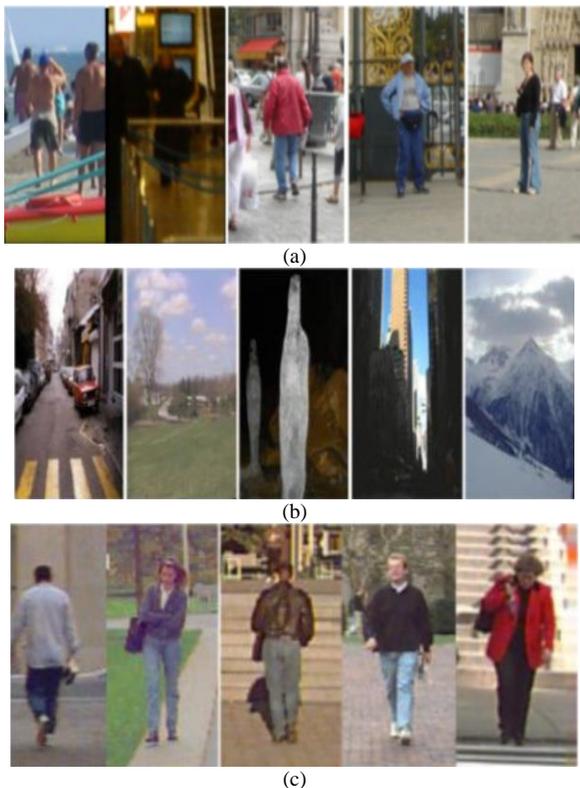

(a)

(b)

(c)

Fig. 7. Image examples from datasets (a) Positive examples of INRIA Dataset, (b) Negative examples of INRIA Dataset and (c) Positive examples of MIT Dataset.
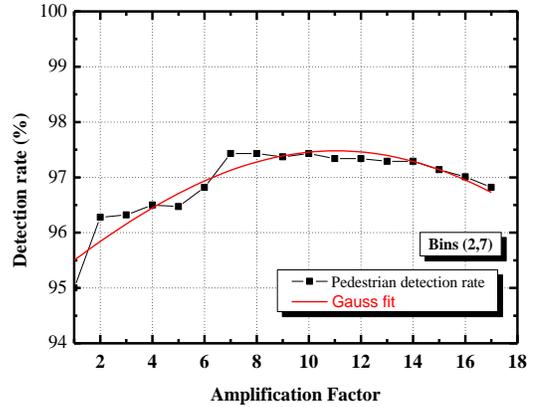

Fig. 8. The system performance according to α value.

2) Experimental results

The details of the used database for learning and testing the performance of the pedestrian detection system are presented in Table 1.

A comparison between the results of our approach and further works based on HOG descriptor is shown in Table 2.

TABLE I. DETAILS OF THE USED DATABASE

Data set	Learning		Test	
	Positive Examples	Negative Examples	Positive Examples	Negative Examples
INRIA	2416 (128×64)	1218 (320×240)	1126 (128×64)	453(different)
MIT	709 (128×64)	--	214 (128×64)	--
Used	3125 (128×64)	6090 (128×64)	MIT214 (128×64)	INRIA 2265 (128×64)
			INRIA 1126 (128×64)	

TABLE II. COMPARISON OF THE EXPERIMENTAL RESULTS

Descriptor	Datasets			
	Sensitivity MIT	*Sensitivity INRIA*	*Specificity INRIA*	*Accuracy MIT + INRIA*
HOG *N. Dalal et al. (2005)*	81.39%	85.9%	99.10%	88.79 %
Optimized HOG *G. Zhang et al. (2010)*	100%	99.07%	98.89%	99.32%
Integral HOG *Y. Said et al. (2012)*	---	94.48%	97.56%	96.02%
This work	**100%**	**97.42%**	**100%**	**99.14%**

It can be concluded from Table 2 that a perfect recognition rate (100%) for the Negative INRIA dataset and the MIT dataset is obtained, together with a respectable percentage for the Positive INRIA dataset. As a conclusion, the proposed

system yields a significant performance in the characterization of pedestrian features, when compared to the other works [7], [40].

B. Vehicle detection

1) Database

To build a vehicle recognition system in the conventional supervised learning, the positive training examples consist of vehicles, and the negative training examples consist of random non-vehicles. The datasets used in our system are MIT cars [37], INRIA cars [38] and Markus cars [39] as positive examples and non-pedestrians INRIA datasets as negative examples. Fig. 9 shows some positive and negative examples. We manually delete the images for non-pedestrian examples that contain cars in order to use them as negative examples for learning. We have obtained 988 car images with their reflections (1 976 samples in total) as positive examples and 4 236 samples extracted from 1 059 not-car images as negative examples. 1/3 of each database was intended for test and 2/3 were intended for learning the system.

2) Re-optimizing the HOG parameters for vehicle detection

After the validation of our approach in pedestrian detection system by increasing the system's accuracy, we will now generalize the proposed approach through the detection of other various obstacles in an urban environment. Based on the same principle (increase of significant bins of each specific obstacle), we will be interested in detecting and identifying the vehicle obstacles. In an image, pedestrian and car have various different characteristics. The HOG descriptor is primarily built for pedestrian detection. Therefore, we need to re-optimize several parameters of the standard HOG descriptor to get the best results for the car detection. Then we added the process of our approach.
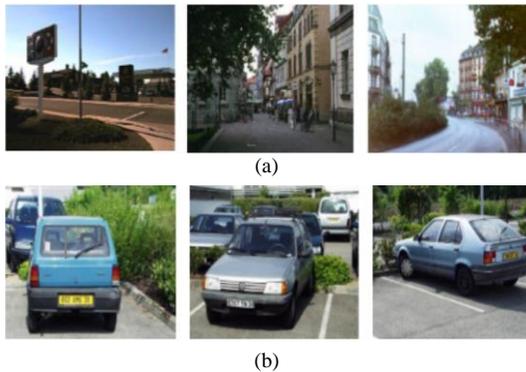
Fig. 9.   Image examples from datasets: (a) Negative examples, (b) Positive examples.

Primary, most of the vehicles have rectangular shapes and they have a larger size then a pedestrian, justifying the choice of (128×128) pixels per window in the learning system. Second, changing the number of pixels per cell, the number of cells per block and the overlapping ratio does not affect the system's performance.

Then we will keep the same parameter's values (8 × 8 Pixels per cell, 2×2 Cells per block) proposed in the standard HOG, that turns out to be effective to express the features of

cars in images. Finally, vertical orientations for a car are characterized by an acute and accurate angle, which does not change within its movement at variance with the pedestrians. That leads to minimizing the scale of bins by increasing her number in [0 Π] plan. The simulation results for different values of bin's number are shown in Table 3.

The simulation provides the best result for 60 bins through an accuracy rate equal to 97.3% with an enhancement of 1.47%. However, it represents the most complex and greedy simulation: resource intensive, memory consuming, execution time... Since such application target an automotive embedded system, working with higher feature-dimension will slow the learning step and may be risking the over-fitting of the SVM classifier in the hardware implementation. Therefore, in the following, we will apply our approach for a number of bins equal to 18; first, in order to save simulation time, second to target an efficient hardware implementation of real time vehicle detection and finally to demonstrate the efficiency of our approach since this case represents the lower sensitivity.

The sample size used in our experiment is a window of 128 × 128 pixels that define the car and non-car images. The HOG-feature extraction process based on 18 bins gives a features vector of 16200 dimensions as shown Fig. 10.

TABLE III.   RECOGNITION RATE ACCORDING TO THE NUMBER OF BINS IN THE HISTOGRAMS

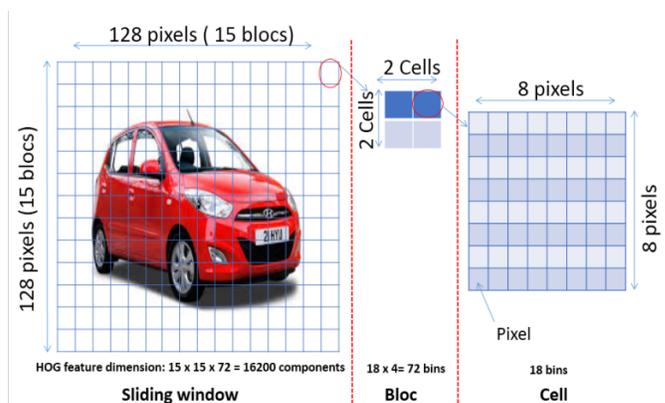| Number of bins	6	9	18	24	36	60	72
HOG Feature length	5400	8100	**16200**	21600	32400	**54000**	64800
Sensitivity rate (%)	93.44	93.88	**91.26**	93.8865	96.94	**97.4**	96.3
Specificity rate (%)	97.50	97.78	**99.12**	98.32	97.58	**97.2**	96.7
Accuracy rate (%)	95.47	95.83	**95.19**	96.1	97.26	**97.3**	96.5

Fig. 10. Overview of the HOG features configuration for the vehicle detection: window, blocks and cells.

3) Experimental results

To better extract the car features, we have applied our approach on the whole datasets through these three steps:

- Select the significant bins that better describe a car feature from other obstacles

- Optimize the amplification factor for each selected bin

- Inject the proposed change in the standard HOG algorithm

The results of each step are presented below.

4) Extraction of significant bins

As shown in Fig. 11, the subtraction between the two mean vectors of the negative and positive examples for the car datasets used in our experiments (INRIA, MIT, MARKUS) give four bins (2,7,15 and 17) whose values are reversed compared to the others bins. By the same logic we explain that these bins have larger gradients orientation density in a car image than a random image in traffic environments. On the other hand, the bin numbers (6, 10 and 14) represent the highest values in this histogram. These bins encode the least frequent oriented gradients for a car image.

Afterwards, we will amplify several combinations between the most significant bins, the least significant bins and their mixture, in the purpose of extracting the best possible system's accuracy.
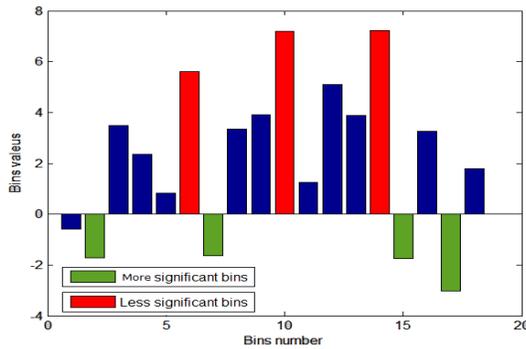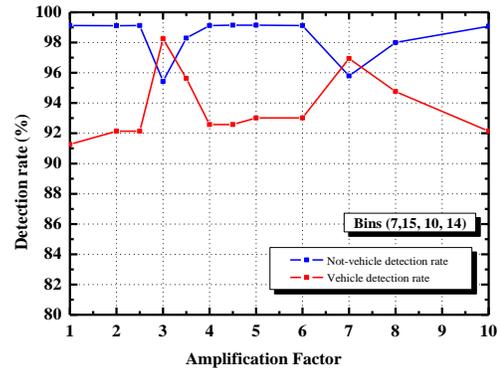

Fig. 11. Most significant bins for a car features extraction based on 18 bins.
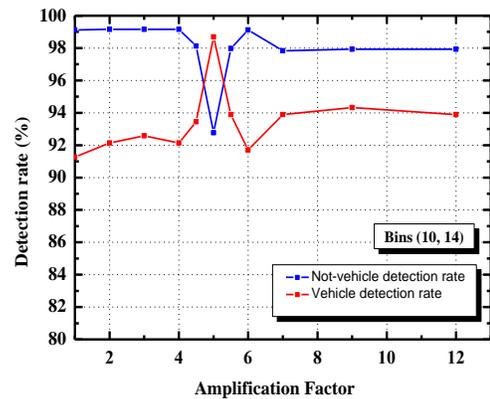
5) Select the amplification factor

In Fig. 12, we represent the best three combinations of the different significant bins. Equally to the case of the pedestrian detection system, we swept the amplification factor at the end to get the best recognition rate for the car detection. Through experimentations, we can reveal that the significant bins are sensitive to the amplification process. The best sensitivity rate has reached 98.69%, by amplifying the bins 10 and 14 with a factor equal to 5. However, we note a clear degradation for the specificity rate that attains 92.76%. The amplification of bins 2,7,15 and 17 has achieved the accuracy in all simulations; in fact, we attain a sensitivity non-vehicle recognition system. The amplification factor getting the highest rate is equal to 3.A comparison between our results and other ones (presented in Table 4), shows that the proposed approach outperforms recent works [28], [30], [31], [41]. However, we cannot rely on this comparison because we do not share the same database, seeing that a growing number of on-road equal to 98.25% and a specificity equal to 95.43% for the vehicle studies are reporting results from private video datasets.

Through the whole experimentation of the pedestrian and the vehicle detection, we have observed that the customized
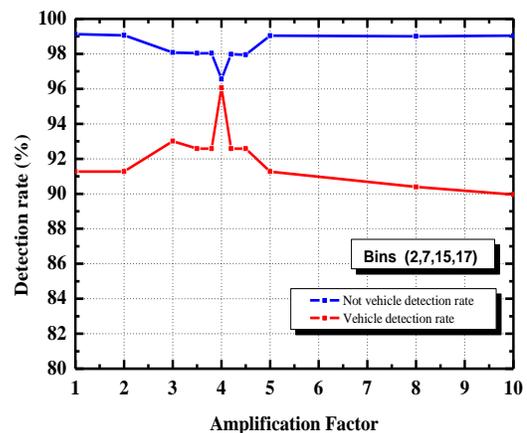
HOG-feature extraction method goes well with several types of obstacles. In addition, a tracking technique can be introduced to supply missing and false detection.

(a)

(b)

(c)

Fig. 12. Detection rate of the best three combinations for the different significant bins: (a) Amplification process for bins (7, 15, 10, 14), (b) Amplification process for bins (10, 14), (c) Amplification process for bins (2,7,15,17).

TABLE IV.   COMPARISON OF EXPERIMENTAL RESULTS

Methods	Sensitivity rate	Specificity rate	Accuracy rate
HOG-Like Gradient *J. Arrospide et al.(2012)*	92.48%	--	--
Haar-features and HOG-features *H. Youpan et al. (2013)*	97.2%	96.8%	97%
Standard HOG *Ling Mao et al. (2010)*	96.87%	97.33%	97.1%
HOG-HCT *S. Li et al. (2012)*	85.2%%	--	--
This work	**98.25%**	**96.32%**	**97.28%**

VI.   CONCLUSION

In this paper, we have proposed an improved version of the HOG feature extraction called Customized HOG. The main contribution is to extract and amplify the most significant bins that describe particularly the desired object. This technique presents a potential solution to the emerging problems related to the obstacle detection for ADAS as well as other applications. The performance evaluation shows that the proposed approach yields significant improvements for the characterization of pedestrians and vehicles features compared to other approaches. Future research works will focus on real-time object detection and its implementation on Field Programmable Gate Arrays (FPGAs) using the proposed customized HOG and some techniques to reduce the feature dimensionality.

REFERENCES

[1] Global status report on road safety 2015. World Health Organization, 2015

[2] Matthias Serfling, Otto Loehlein, Roland Schweiger and Klaus Dietmayert, 2009. Camera and Imaging Radar Feature Level Sensor fusion for Night Vision Pedestrian Recognition. IEEE Intelligent Vehicles Symposium, 597 – 603, Xi'an, 2009, .

[3] Daniel Clarke, Alois Knoll, Feihu Zhang: Vehicle Detection Based on LiDAR and Camera Fusio. 2014. Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on (October. 2014)

[4] Hao Sun, Cheng Wang and Boliang Wang.2011. Night vision pedestrian detection using a forward looking infrared camera. International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping, 1-4 , (Jan. 2011).

[5] Leu. A, Aiteanu. D and Graser. A. 2011. A novel stereo camera based collision warning system for automotive applications, Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on, 409 – 414 (May.2011).

[6] Mesmakhosroshahi. M, Joohee Kim, Yunsik Lee, and Jong-bok Kim. 2013. Stereo based region of interest generation for pedestrian detection in driver assistance systems. Image Processing (ICIP), 2013 20th IEEE International Conference on, 3386 – 3389, (Sept. 2013).

[7] N. Dalal and B. Triggs. 2005. Histograms of Oriented Gradients for Human Detection. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 886-893 (June. 2005)..

[8] Viola. P, Jones. M.2001. Rapid Object Detection using a Boosted Cascade of Simple Features. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on , vol 1, 511-518, 2001.

[9] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,"

in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, Jul 2002

[10] Y. Cao, S. Pranata, M. Yasugi, Z. Niu and H. Nishimura, "Stagged multi-scale LBP for pedestrian detection," 2012 19th IEEE International Conference on Image Processing, Orlando, FL, pp. 449-452, 2012.

[11] Jingjing Li, Yong Zhao, Dongbing Quan. 2013. The combination of C SLBP and LBP feature for pedestrian detection. International Conference on Computer Science and Network Technology (ICCSNT), 543 – 546, (Oct. 2013).

[12] Cosma. C, Brehar. R, Nedevschi. S. Pedestrians detection using a cascade of LBP and HOG classifiers. 2013. International Conference on Intelligent Computer Communication and Processing (ICCP), 69 – 75, (Sept. 2013).

[13] Laopracha. N, Thongkrau. T, Sunat. K, Songrum. P, Chamchong. 2014. Improving vehicle detection by adapting parameters of HOG and kernel functions of SVM. International Computer Science and Engineering Conference (ICSEC), 372 – 377, (July. 2014)..

[14] N. Dalal, B Triggs, and C Schmid. 2006. Human Detection Using Oriented Histograms of Flow and Appearance. 9th European Conference on Computer Vision, Graz, Austria. Computer Vision – ECCV, 428–441, 2006.

[15] XianbinCao, Changxia Wu, Pingkun Yan, Xuelong Li.2011. Linear SVM classification using boosting HOG features for vehicle detection in low-altitude airborne videos. 18th IEEE International Conference on Image Processing (ICIP), 2421 – 2424, (Sept. 2011).

[16] Lorca. D.F, Arroyo. R, Sotelo. M.A.2013. Vehicle logo recognition in traffic images using HOG features and SVM. International Conference on Intelligent Transportation Systems - (ITSC), 2229 – 2234, (Oct. 2013)..

[17] ChengbinZeng, Huadong Ma, Anlong Ming. 2010. Fast human detection using mi-sVM and a cascade of HOG-LBP features. 17th IEEE International Conference on Image Processing (ICIP) 3845 – 3848, (Sept. 2010).

[18] M. P. Hoang, Le Dung, De Souza-Daw. T, Nguyen Tien Dzung, Thang Manh Hoang. 2012. Extraction of human facial features based on Haar feature with Adaboost and image recognition techniques. International Conference on Communications and Electronics (ICCE), 302 – 305, (Aug. 2012).

[19] Van-Dung Hoang, Vavilin. A, Kang-Hyun Jo.2012. Pedestrian detection approach based on modified Haar-like features and AdaBoost. 2012. International Conference on Control, Automation and Systems (ICCAS). 614 – 618, (Oct. 2012).

[20] Rakate, G.R, Borhade. S.R, Jadhav, P.S, Shah. M.S.2012. Advanced Pedestrian Detection system using combination of Haar-like features, Adaboost algorithm and Edgelet-Shapelet, International Conference on Computational Intelligence & Computing Research (ICCIC). 1-5, (Dec. 2012).

[21] H. Ameur, A. Helali, M.Nasri, H. Maaref. (2014), Improved feature extraction method based on Histogram of Oriented Gradients for pedestrian detection, Computer & Information Technology (GSCIT), pp 1 – 5, 2014

[22] Ballesteros. G, Salgado. L. 2014. Optimized hog for on-road video based vehicle verification. Signal Processing Conference (EUSIPCO), 805 – 809, (Sept. 2014).

[23] G. Zhang, F. Gao, C. Liu, W. Liu, H. Yuan. 2010. A Pedestrian Detection Method Based on SVM Classifier and Optimized Histograms of Oriented Gradients Feature, International Conference on Natural Computation (ICNC), 3257 – 3260, (Aug. 2010).

[24] Hui-Xing Jia, Yu-Jin Zhang. 2007. Fast Human Detection by Boosting Histograms of Oriented Gradients. International Conference on Image and Graphics, 683 – 688, (Aug. 2007).

[25] X. Wang, Tony X. Han, S. Yan. 2009. An HOG-LBP Human Detector with Partial Occlusion Handling. Computer Vision, 2009 IEEE 12th International Conference on, 32 – 39, (Sept. 2009)..

[26] Q. Zhu, S. Avidan, Mei-Chen Yeh, and K. Cheng. 2006. Fast Human Detection Using a Cascade of Histograms of Oriented Gradients.2006. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on (Volume 2), 1491 – 1498, 2006.

[27] N. He, J. Cao and L. Song. 2008. Scale Space Histogram of Oriented Gradients for Human Detection, International Symposium on Information Science and Engineering. 167 – 170 (Dec. 2008)..

[28] Ling Mao, Mei Xie, Yi Huang, Yuefei Zhang. 2010. Preceding Vehicle Detection Using Histograms of Oriented Gradients, International Conference on Communications, Circuits and Systems (ICCCAS), 354 – 358, (July. 2010).

[29] Xing Li, Xiaosong Guo. 2013. A HOG Feature and SVM Based Method for Forward Vehicle Detection With Single Camera, International Conference on Intelligent Human-Machine Systems and Cybernetics. 263 – 266. (Aug. 2013).

[30] Arrospide. J, Salgado. L, Marinas. J. 2012. HOG-like Gradient-based Descriptor for Visual Vehicle Detection, Intelligent Vehicles Symposium Alcalá de Henares, Spain. 223 – 228, (June. 2012).

[31] Youpan Hu, Qing He, Xiaobin Zhuang, Haibin Wang, Baopu Li, Zhenfu Wen, Bin Leng, Guan Guan, Dongjie Chen. 2013. Algorithm for vision-based vehicle detection and classification, International Conference on Robotics and Biomimetics (ROBIO), 568 – 572, (Dec.2013).

[32] Pablo Negri, Xavier Clady, Shehzad Muhammad Hanif, and Lionel Prevost. 2008. A Cascade of Boosted Generative and Discriminative Classifiers for Vehicle Detection, EURASIP Journal on Advances in Signal Processing, 12, 2008.

[33] B.E. Boser, I. M. Guyon, and V. N. Vapnik. 1992. A training algorithm for optimal margin classifiers, Proceeding COLT '92 Proceedings of the fifth annual workshop on Computational learning theory 144-152, 1992,.

[34] V. N. Vapnik. 1998. Statistical Learning Theory, September 1998.

[35] M. Oren and C.P. Papageorgiou and P. Sinha and E. Osuna and T. Poggio, Pedestrian Detection Using Wavelet Templates, cvpr, PP 193-99, 1997.

[36] N. Dalal, "INRIA Person Dataset". http://pascal.inrialpes.fr/data/human/

[37] CBCL Car database. http://cbcl.mit.edu/projects/cbcl/software-datasets/cars128x128.tar.gz

[38] http://www.cs.ubc.ca/~pcarbo/objrecls/data/cars.tar.gz

[39] Michael Fink and Pietro Perona, Technical report Caltech CSTR:2003.008, 2003.

[40] Y. Said, M. Atri, T. Saidani and R. Tourki. 2012. Détection d'Individus dans les Images Couleurs à base d'Histogramme Intégral de Gradient Orienté et SVM. Sciences of Electronics Technologies of Information and Telecommunications Sousse. (March. 2012).

[41] Sun Li, Bo Wang, ZhiHui Zheng, HaiLuo Wang. 2012. Multi-view vehicle detection in traffic surveillance combining HOG HCT and deformable part models. Wavelet Analysis and Pattern Recognition International Conference on. Xian, 202 – 207, (July. 2012).