

On-Road Vehicles Detection using Appearance and Texture Information

M. Hassaballah¹, Mourad A. Kenk^{*2}, Ibrahim M. Elhenawy³

¹Computer Science Department, Faculty of Computers and Informatics, South Valley University, Luxor, Egypt

²Mathematics Department, Faculty of Science, South Valley University, Qena, 83523, Egypt.

³Computer Science Department, Faculty of Computers and Informatics, Zagazig University, Zagazig, Egypt
mourad.ahmed@sci.svu.edu.eg

Abstract

Autonomous vehicles and traffic monitoring systems have gained more attention during the last decades. Accurate features are the corner stones in these systems, where the performance of automatically detecting on-road vehicles systems mainly depends on these features. In this paper, we propose a vehicle detection method based on Local Binary Pattern (LBP) derivatives for feature extraction and then fuse them to align histograms for random sampled local regions. The aligned histograms are in turn used to measure the dissimilarity between regions of all training images. Evaluating the fit between histograms is built in clustering forests which are used to search between different LBP derivatives filters for latent features of vehicles and cluster them to build a discriminative fusion codebook of latent features. Finally, the task of vehicles detection is performed using saliency map built on the learnt latent fusion features to estimate the vehicles locations in test image. The experimental results show that the proposed method achieves significant and consistent improvements over the state-of-the-art methods.

Keywords: *Vehicles Detection; Features Extraction; Local Binary Pattern; Histogram; Clustering Forest.*

1. Introduction

In the last few years, detecting and tracking on-road vehicles are considered as an important researches area in computer vision [1-3] with real traffic surveillance applications counting traffic management, traffic flow control system at an intersection. For example, video surveillance analysis systems for urban traffic roads provide rapid effective information producing in increased safety and traffic flow such as: stranded vehicles, lane crossing, vehicles parked cross the roads, traffic congestion, count passing vehicles as well as determine their direction and the lane which they are taking, furthermore, the plate number, type, and speed of vehicles [4,5]. Moreover, vehicle detection is an origin for vehicle validation, vehicle classification, vehicle verification and vehicle tracking. The purpose of vehicle detection is to identify the possible vehicle locations in the image and mark them as regions of interest (ROI). Unfortunately, the detection task for vehicles using video surveillance cameras is a computer vision challenging problem because vehicles have a large intra-class variation and their appearance changes considerably. Furthermore, varying illumination, occlusion and a clutter background make the problem very difficult than other object detections [6].

Several vehicle detection methods have been proposed in the literature. For systems with stereo cameras such as in [7,8], the disparity map computed from the stereo images is used to find the possible vehicle locations. For a monocular-based system, the determination of vehicle locations has to be done by analyzing the vehicle's motion or appearance.

Additionally, appearance-based technique [9] detects vehicles based on some specific appearances of vehicles such as: rear and front view, side view. Where, a cascade search tree classifier is trained based on clustering groups of car samples in an unsupervised way. Also, surveillance camera position is affecting on what scale and what range that can be useable. Thus, variation of the appearance of vehicles need a robust pattern descriptor that can describe the vehicles features in different scenarios. For detecting the vehicles in a tunnel surveillance application, [10] use the Haar-like features and AdaBoost cascade classifier. The edge information is employed to highlight the sides of the vehicles in [11]; where, the edge and corner features are used with dynamic Bayesian network classifier. In [12], a vehicle detection method in high resolution aerial images is proposed based on using sparse representation and superpixel segmentation. In more recent works, [13] combines deep convolutional neural network (DNN) based feature learning to learn discriminative features with exemplar-SVMs classifier to improve the robustness of the detection.

Efficient features extraction and representation of vehicle features is an important factor and has a direct influence on the performance of the detection. Therefore, feature descriptors should extract highly discriminative features to be robust against image transformations such as rotation, lighting changes, and noise. Extraction of image features is studied extensively in the literature [14]. In this regard, local binary pattern (LBP) [15], as feature extraction technique, has become extensively used in various applications such as ear recognition [16], image analysis, and image classification [17] due to its high discriminative strength, tolerance in contrast to illumination variation and ordinary of computations alongside these applications. Also, saliency detection has become an extremely attractive technique in computer vision, and it has been applied in many domains including object detection [18] and object segmentation [19]. The principal of saliency detection is as follows. First, distinct features are adopted from the input image, such as: intensity, color, and texture, then saliency maps are built using these features. Saliency map by definition is a gray scale image where higher values are characterizing regions that are more distinct for objects or places, while lower values are corresponding to background. Hence, this information can be used as a prior for a classification problem to classify and detect objects.

In this paper, a method for vehicle detection is proposed considering both appearance and texture information to represent the vehicle images based on combination of pattern descriptors LBP; a non-linear dimension reduction and clustering forests (decision trees whose leaves define a spatial partitioning or grouping that groups vehicles regions samples in an unsupervised way). Where each region is aligned to a single feature histogram, these histograms are thresholded by Chi-Square dissimilarity measurement which define spatial clusters. Then, a robust tree-structured detector is built based on the saliency map. The proposed method is evaluated on comprehensive car data sets and experimental results show that it outperforms state-of-art methods.

The rest of the paper is organized as follows: Section 2 explains details and implementation of the proposed vehicle detection method. Section 3 presents the experimental results and Section 4 concludes the paper.

2. The Proposed Method

2.1 LBP

The original LBP is turned up by Ojala et al. [15]; where the pixels of an image is being thresholded by the 3x3 neighbors of each pixel with the center value and treating the result as a binary digit. Then, the histogram can be used as a pattern descriptor. Subsequently, the LBP operator is extended to use neighbors of different scope. Using circular neighbors and bi-

linearly interpolating the pixel values allow any radius and number of pixels in the neighbors. For neighbors (P,R) which mean P existence points on a circle of radius R. Assumed any central pixel in the image, a pattern number is computed by emulating its value with its neighbors' values:

$$LBP_{P,R} = \sum_{\ell=0}^{P-1} s(\mathcal{G}_\ell - \mathcal{G}_c)2^\ell, \quad s(\cdot) = \begin{cases} 1, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where $\mathcal{G}_c, \mathcal{G}_\ell$ are the gray value of the central pixel and the value of its neighbors respectively. Where, the coordinates of \mathcal{G}_ℓ are given using bi-linearly interpolation as:

$$x_\ell = \left(x + R \cos\left(\frac{2\pi\ell}{P}\right) \right), y_\ell = \left(y - R \sin\left(\frac{2\pi\ell}{P}\right) \right) \quad (2)$$

The other gray values of neighbors that are not in the center of grids can be estimated by interpolation. After representing the input image $N \times M$ by the LBP for each pixel (i,j) , the histogram builds by:

$$\mathcal{H}1(k) = \sum_{i=0}^N \sum_{j=0}^M \mathcal{G}(LBP_{P,R}(i,j), k)_{k \in [0,K]}, \mathcal{G}(x,y) = \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where K is maximum value of the LBP pattern.

2.2 Histograms of contrast and variance

Nevertheless, the LBP is brought in as an integral measure for local image contrast which is a property of texture usually overlooked as an important key for vision systems, but the original LBP operator overpass the significance of gray level differences. In general, texture is distinct not only the texture of patterns but also the amplitude of the patterns. So that, these two measures can support each other in a very useful way for resisting even intra-image illumination variation. Local variance $LV_{P,R}$ can be defined as invariant against shifts in the gray level

$$LV_{P,R} = \frac{1}{P} \sum_{\ell=0}^{P-1} (\mathcal{G}_\ell - \mathcal{V})^2, \quad \text{where } \mathcal{V} = \frac{1}{P} \sum_{\ell=0}^{P-1} \mathcal{G}_\ell \quad (4)$$

Then, the joint distribution $LBP_{P,R}/LV_{P,R}$ produce LBP variance (LPBV) that can describe the local image texture better than using LBP alone [20]. Hence, the $LV_{P,R}$ can be utilized as a modifying weight to adapt the contribution of the LBP in the LPBV histogram computing:

$$\mathcal{H}2(k) = \sum_{i=0}^N \sum_{j=0}^M \mathcal{U}(LBP_{P,R}(i,j), k), \quad k \in [0, K] \quad (5)$$

where,

$$\mathcal{U}(LBP_{P,R}(I,j), k) = \begin{cases} LV_{P,R}(I,j), & LBP_{P,R}(I,j) = k \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

In spatial histograms, there is an effectively description for the vehicle features on three different levels of representation:

- Histograms contain information about the patterns on a pixel-level,
- Histograms contain different distance over neighbor produce information on a regional-level,
- Histograms contain the effect of contrast for pattern description on a strength-level.

2.3 Selective regions models

When the image has been divided into regions, it can be expected that some of the regions contain more useful information than others in terms of vehicle features. For example, headlight, logo, side mirror, wheels, license number seem to be an important cue in vehicle tracking for traffic management. To take advantage of this, we use the uniform distribution to divide image into regions \mathcal{R}_j covering whole vehicle features. Let \mathcal{A} is a set of regions $\mathcal{A} = \{\mathcal{R}_j^{\text{train}}(\Gamma_j, \mathcal{H}_j, l_j, d_j)\}$, where each region is represented by three space channels $\Gamma = \{\gamma^1, \gamma^2, \gamma^3\}$ extracted by different derivative of LBP operators. Simultaneously, \mathcal{H} is a set of histograms $\mathcal{H} = \{\mathcal{H}1, \mathcal{H}2, \mathcal{H}3\}$ which have been generated concurrence with LBP spatial. Several possible dissimilarity measures have been proposed for histograms: Log-likelihood statistic, Histogram intersection and Chi square statistic (χ^2). The log-likelihood measure has a good performance with many situations but may be sensitive with small regions. Because of the histograms maybe contain many zeros. So that the Chi-square distance usually is stable with small regions. Finally, displacement vector d_j for regions have been associated and the class label $l_i = \{1,0\}$ for vehicle and background respectively. Figure 1 shows the basic processes in the training stage of the proposed method.

2.4 Clustering forests

We developed clustering forests (CF) for building visual codebooks. CF consists of randomized decision trees that are built recursively top down. At each node i corresponding to set of regions $\mathcal{A}_i = \{\mathcal{R}_j^{\text{train}}\}$, two children nodes L, R are created by choosing a binary test defined as $\mathcal{T}(\Gamma) \rightarrow \{0,1\}$ that divides \mathcal{A}_i into two disjoint set of regions by

$$\Omega_i = \arg \max \Delta (\mathcal{A}_i, \mathcal{A}_i^R, \mathcal{A}_i^L, \mathcal{T}_i) \quad (7)$$

where $\mathcal{A}_i = \mathcal{A}_i^L \cup \mathcal{A}_i^R$ with $\mathcal{A}_i^L \cap \mathcal{A}_i^R = \emptyset$, and $\Delta = \Delta 1$ or $\Delta 2$. In this context, $\Delta 1$ measures the impurity of displacement vector d_j for regions, and

$$\Delta 1 = \sum_{l_i=1} \|d_j - d_m\|^2 \quad (8)$$

While, $\Delta 2$ measures the impurity of class label l_i uncertainty between vehicle regions and the background, where

$$\Delta 2(\mathcal{A}_i) = |\mathcal{A}_i| \cdot \Phi(l_i), \quad \Phi(l_i, \mathcal{A}_i) = - \frac{\sum_h P(l_i | \mathcal{A}_i) \log P(l_i | \mathcal{A}_i)}{\sum_h P(l_i | \mathcal{A}_i)} \quad (9)$$

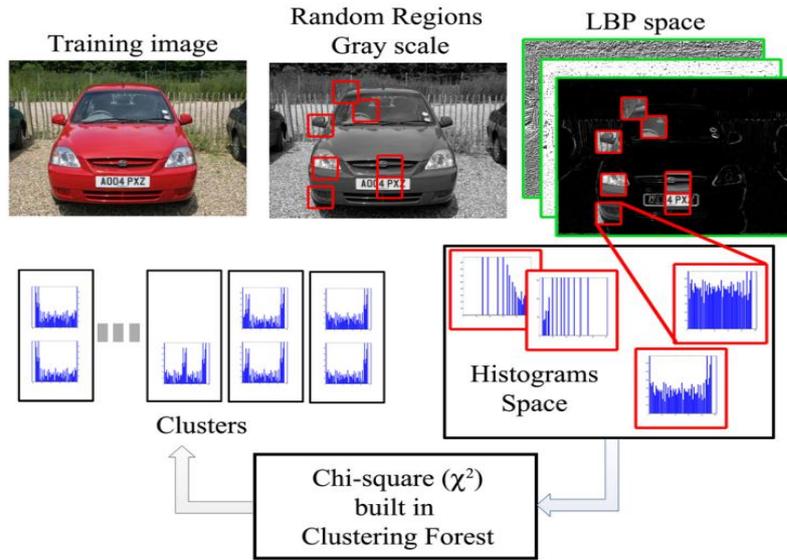


Figure 1. The training process of the proposed method.

where Φ is the Aczél and Daróczy Entropy (ADE) [21] used to maximizing the joint probability of the class label for classification information gain. Hence, we define a probability density function (PDF) as the saliency map, where the high probability vote of regions is representing the high saliency.

During training, splitting recursion continues until further subdivision is impossible: clustering max level reach or either all surviving training regions examples belong to one similarity measurement threshold, where all regions have identical values for all histograms. To this end the Chi-square χ^2 statistic dissimilarity measurement can be considered as the crux of the binary test:

$$\mathcal{T}_{(\mathcal{S}, \mathcal{M}, \omega)}(\Gamma) = \begin{cases} 1, & \chi^2(\mathcal{S}, \mathcal{M}) < \omega \\ 0, & \text{otherwise} \end{cases}, \quad \chi^2(\mathcal{S}, \mathcal{M}) = \sum_i \frac{(\mathcal{S}_i - \mathcal{M}_i)^2}{\mathcal{S}_i + \mathcal{M}_i} \quad (10)$$

where \mathcal{S}, \mathcal{M} are two histograms indexed in the same LBP operator. The tests are selected randomly as follows: a histogram index $\mathcal{J}(\mathcal{H}_i) = \{\mathcal{H}1, \mathcal{H}2, \mathcal{H}3\}$ for appearance of regions extract by LBP operator $\Gamma = \{\gamma^1, \gamma^2, \gamma^3\}$ is chosen randomly. Thus, a random elementary threshold ω is generated at the beginning of the training stage for tests as dissimilarity measurement, thereafter the threshold value ω for each test is chosen uniformly from the range of differences between histograms observed on data randomly. Clustering forests are utilized in recent studies for some tasks such as prediction of breast cancer [22].

2.5 Saliency detection

During the detection stage, each region histogram from the test image is matched to the clusters of histograms, similar to the case for appearance-based object detection [19], when an entry in the clustered forest is activated, it casts votes for all possible vehicles centroids based on the associated vectors. Finally, candidate vehicle centroids are detected as high saliency in the saliency map using Mean-Shift Mode Estimation. For each member histogram $(\mathcal{H}_j)_i$ in cluster i , we store their linked Region position relative to the vehicle center. Thus, a cluster entry records the following information:

- Similarity measure threshold value
- LBP derivative descriptor type
- Number of similarities histograms Regions
- Relative Regions to vehicle center ($y_m, m = 1, \dots, M_i$)

At last, probability of positive training regions at each cluster is being computed for saliency map.

The proposed detection process is illustrated in Figure 2. Firstly, we measure the similarity of each region histogram in the test image located at x to the clustered histograms in forest using Chi-square distance. The cluster entry is declared to be similar if $\chi^2(\cdot) < \omega$. For each matched cluster entry j , we cast votes for possible locations of the vehicle centers ($y_m, m = 1, \dots, M_i$). The described saliency map that represents a PDF can then be used as prior knowledge to detect vehicles. Then, vehicle detection is accomplished by searching the saliency map using mean-shift mode estimation. Formally, let $p(x)$ is the probability that vehicle appears at candidate position x_n in the test image, since the voting is casted from each individual member in activated clusters entries \mathcal{J} , votes are casted for all possible locations of the vehicle centers y_m . Thus, $p(x_n|\mathcal{J})$ can be marginalized as:

$$p(x_n|\mathcal{J}) = \sum_k p(x_n|y_m) p(y_m|\mathcal{J}) \tag{11}$$

without prior knowledge of y_m , they are treated equally and assume $p(y_m|\mathcal{J})$ is a uniform distribution, i.e., $p(y_m|\mathcal{J}) = \frac{1}{M_i}$. Since we only vote at the location of possible object centers, we have $p(x_n|y_m) = \delta(x_n - y_m)$ where $\delta(z)$ is the Dirac delta function. The vehicle detection hypotheses $E(V)$ coming from all trees in the forest $\{F_t\}_{t=1}^T$ by:

$$E(V) = \sum_{x_n} p(x_n|\mathcal{J}; \{F_t\}_{t=1}^T) \tag{12}$$

To handle scale variations, the size of the detected vehicle bounding boxes is assumed to be fixed to $W \times H$ during both training and testing. Then, the hypothesized $\overline{E(V)}$ bounding box in the original image is centered at the point $\frac{\bar{x}_n}{s}$ has the size $\frac{W}{s} \times \frac{H}{s}$, where s is the scale value for detection.

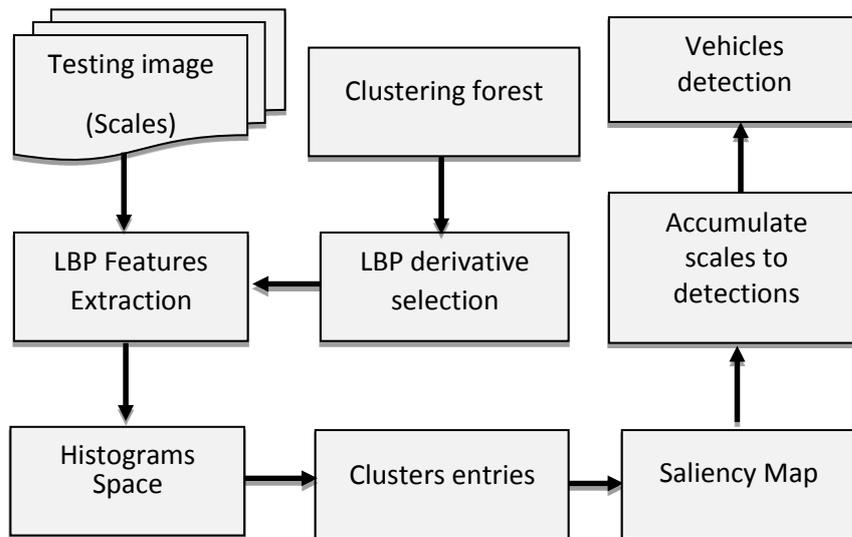


Figure 2. The proposed vehicle detection method pipeline.

3. Experiments and Results

3.1 Evaluation metrics

We use the Detection Accuracy (DA), Multiple Object Detection Accuracy (MODA), Multiple Object Detection Precision (MODP), and different Overlap Ratio (O_Ratio) with Non-binary thresholding between the predicted bounding box $D_i^{(z)}$ and ground truth bounding box $G_i^{(z)}$

$$DA(z) = \frac{O_Ratio}{\left\lfloor \frac{N_G^{(z)} + N_D^{(z)}}{2} \right\rfloor}, \quad O_Ratio = \sum_{i=1}^{N_{mapped}^{(z)}} \frac{|G_i^{(z)} \cap D_i^{(z)}|}{|G_i^{(z)} \cup D_i^{(z)}|} \quad (13)$$

Where, $N_G^{(z)}$ and $N_D^{(z)}$ denote the number of ground-truth vehicles and the number of detected vehicles in image z , respectively. DA computes the spatial overlap between the ground truth and detected objects as a ratio of the spatial intersection between the two objects and the spatial union of them. While, the O_Ratio is the sum of all the overlaps that is normalized over the average of the number of ground truth and detected objects for a single image z where $N_{mapped}^{(z)}$ is the number of mapped ground truth and detected vehicle pairs for image z .

3.2 Performance of LBP derivatives and its fusion

Evaluation the performance of the proposed method is carried out using the same setting of clustering forests training for the stopping criteria of 15 levels with maximum clustering and 20 regions of the same features and three different size of forests trees {1,3,5}. Positive samples are collected with 587 images in the training set of the CompCars dataset [23] with resolution 100×45, example images are given in Figure. 3. Negative training set are randomly cropped from background images with 5329 images. We individually investigate the LBP derivatives, then their randomly fused together by train the clustering forests with 29,350 positive regions and 266,450 negative regions samples.



Figure 3. Training samples: Negative training dataset (right) CompCars training dataset (left)

Table 1 shows the important information about clustering process, the original LBP (OLBP) channel scores the highest clusters number in all forests sizes. On the other hand, the circular LBP (CLBP), and the LBPV score the lowest clusters number in different forests sizes. The data are in Table 1 with bold font. The fusion features of LBP derivatives in LBP-F is considered randomly without supervision, the fusion ratio between them in all clustering forests tree is shown in Figure. 4 and the thresholding variations across training stage is shown in Figure 5; where the LBP-F characterized with high thresholding values between features and fewer clusters.

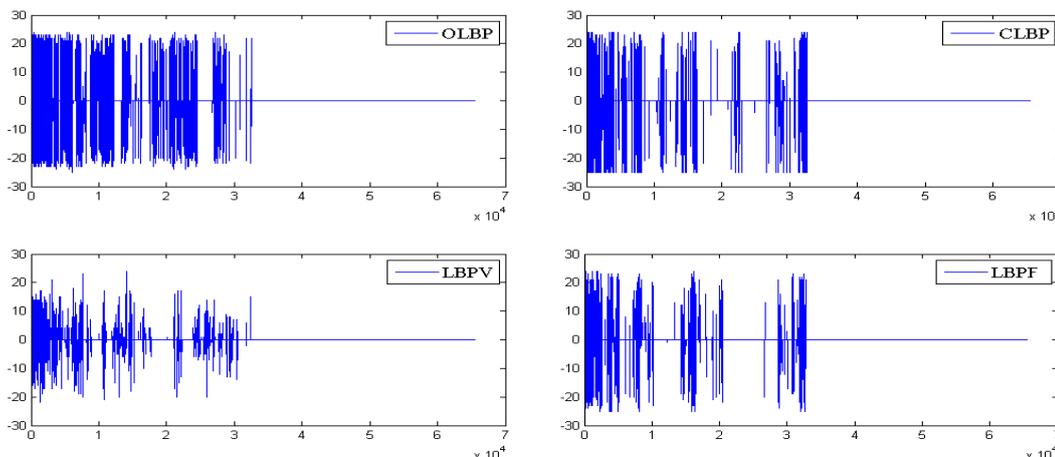


Figure 4. Thresholding variations across training stage where LBP-F characterized with high thresholding values between features and fewer clusters (right bottom).

The performance of all learnt detectors is evaluated using UIUC car dataset [24] on single-scale part. The single-scale part has 170 images containing 200 car instances with fixed resolution as 100×40 . We set {40%, 50%, 60%} as three different overlap ratios to investigate the detection accuracy. The performance of LBP-F is shown in Table 2. It is clear that the performance in case of fusion LBP derivatives is higher than using individual LBP feature. Examples of detection results on the UIUC dataset using the proposed method are shown in Figure. 6.

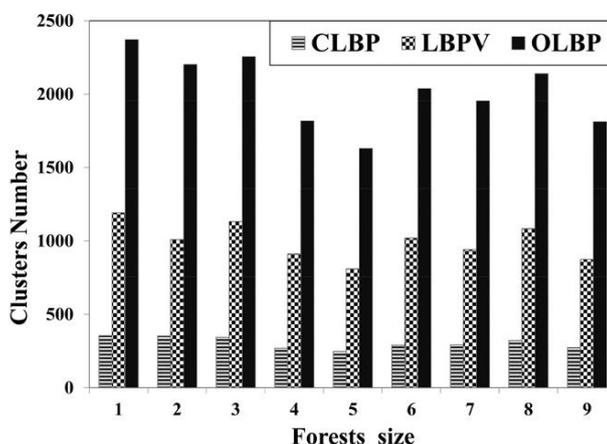


Figure 5. Fusion ratio of LBP derivatives in LBP-F training features.

Table 1. Number of splitting nodes and clusters for three forest size F=1, 2, 3.

	Splitting Nodes Number and Clusters Number					
	F=1		F=3		F=5	
	Nodes	Cluster	Nodes	Cluster	Nodes	Cluster
OLBP	32,773	4348	32,776	4601	32,637	4714
CLBP	32,000	2914	31,511	2700	31,913	2354
LBPV	29,707	2885	32,451	2760	30,379	3026
LBP-F	31,839	3209	31,496	2858	32,330	2192

Table 2. Performance evaluation of LBP and its derivatives on the UIUC car dataset.

Detection Accuracy		Overlap Ratio Threshold		
		40%	50%	60%
OLBP	F=1	43.8%	41%	39%
	F=3	45.8%	43.7%	42%
	F=5	50%	48.2%	45.3%
CLBP	F=1	52.5%	49%	48%
	F=3	55.9%	54%	52.7%
	F=5	57.2%	56%	55%
LBPV	F=1	51%	49.3%	49%
	F=3	58.6%	58%	57%
	F=5	65.6%	62.7%	62%
LBP-F	F=1	86%	83%	81%
	F=3	93.1%	91%	88.4%
	F=5	99.9%	99.6%	95.7%

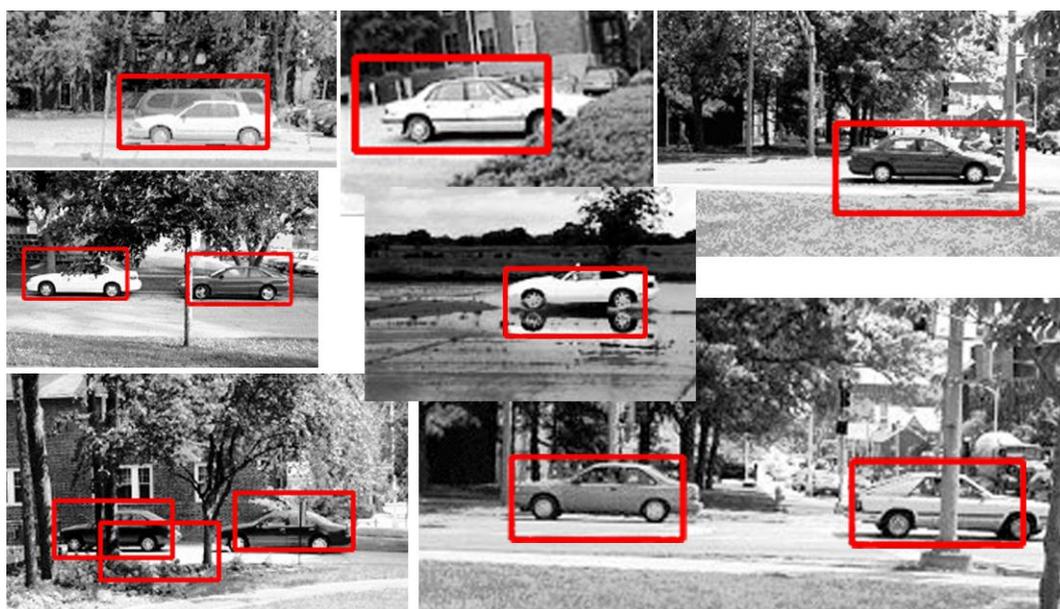


Figure 6. Example of detection results on the UIUC dataset using the proposed method.

3.3 Comparison with other methods

We use clustering forests of 5 trees with two stop conditions, 15 maximum level of clustering, and 20 regions in the cluster. After that, the discriminative saliency map is learnt. The performance of learnt clustering forests is evaluated on both single-scale and multi-scale part of UIUC car test set. Multi-scale part consists of 108 images containing 139 cars with resolution ranging from 89×36 to 212×85 . The comparison of the proposed method with previous ones is reported in Table 3. The Equal Error Rate (EER) of other methods are given in the Table 3. It can be seen that our method has competitive performance to the state-of-the-art methods. From the experiments, we can report that the proposed method can achieve

higher performance on single-scale part and multi-scale part of 99.6% and 96.9% on the UIUC car dataset, respectively.

Table 3. Comparison of the proposed method with state-of-the-arts on the UIUC dataset

Method	Single-scale	Multi-scale
Agarwal et al. [24]	76.5%	39.6%
Fergus et al. [25]	88.5%	-
Fritz et al. [26]	88.6%	87.8%
Kapoor&Winn [27]	94.0%	-
Mutch&Lowe [28]	99.9%	90.6%
Wu&Nevatia [29]	97.5%	93.5%
Leibe et al. [30]	97.5%	95.0%
Kuo& Ramakant [9]	98.5%	95.0%
The proposed method	99.6%	96.9%

4. Conclusion

In this work, a method for localizing and detecting on-road vehicle is proposed using the saliency map based on conducting local binary pattern histograms and clustering forests. The aligned histograms are used in measuring the dissimilarity by Chi-square statistic between regions of all training images. The performance of the proposed method is tested on the UIUC car dataset and the experimental results show that fusion several local binary pattern features together is very useful for the detection task; where, the detection rate reaches to 99% with threshold overlapping of 40 % and five forests compared to 50 % for using only a single binary pattern feature on the same setting. On the other hand, it is noted that the performance of the proposed method is competitive with several published methods and in some cases outperforms these methods. In future work, the study will focus on investigating other techniques for extracting highly discriminative image features.

References

- [1] Wang, X., Tang, J., Niu, J., Zhao, X., “Vision based two-step brake detection method for vehicle collision avoidance”. *Neurocomputing*, vol. 173, pp. 450–461, 2016.
- [2] Su, A., Sun, X., Zhang, Y., Yu, Q., “Efficient rotation-invariant histogram of oriented gradient descriptors for car detection in satellite images”. *IET Computer Vision*, vol. 33, pp. 1250–1265, 2016.
- [3] Asvadi, A., Garrote, L., Premevida, C., Peixoto, P., Nunes, U.J., “Multimodal vehicle detection: fusing 3D-LIDAR and color camera data”. *Pattern Recognition Letters*, vol. 115, no. 1, pp. 20-29, 2018.
- [4] Teoh, S. S., Bräunl, T., “Symmetry-based monocular vehicle detection system.” *Machine Vision and Applications*, vol. 23, no. 5, pp. 831-842, 2012.

- [5] Hsieh, J. W., Yu, S. H., Chen, Y. S., Hu, W. F. "Automatic traffic surveillance system for vehicle tracking and classification." *IEEE Trans. on Intelligent Transportation Systems*, vol. 7, no. 2, pp. 175-187, 2006.
- [6] Gupte, S., Masoud, O., Martin, R. K., Poulos, N. P. "Detection and classification of vehicles." *IEEE Trans. on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37-47, 2002.
- [7] Nedeveschi, S., Vatavu, A., Oniga, F., Meinecke, M. M., "Forward collision detection using a stereo vision system." In the 4th Conference on Intelligent Computer Communication and Processing (ICCP 2008), pp. 115–122, 2008.
- [8] Huh, K., Park, J., Hwang, J., Hong, D., "A stereo vision-based obstacle detection system in vehicles." *Optics Lasers Engineering*, vol. 46, no. 2, pp. 168–178, 2008.
- [9] Kuo, C-H., Ramakant, N., "Robust multi-view car detection using unsupervised sub-categorization." In the IEEE Workshop on Applications of Computer Vision (WACV), pp. 1-8, 2009.
- [10] Rios-Cabrera, R., Tuytelaars, T., Van Gool, L., "Efficient multi-camera vehicle detection, tracking, and identification in a tunnel surveillance application." *Computer Vision and Image Understanding*, vol. 116, pp. 742–753, 2012.
- [11] Cheng, H.Y., Weng, C.C., Chen, Y.Y., "Vehicle detection in aerial surveillance using dynamic Bayesian networks." *IEEE Trans. on Image Processing*, vol. 21, no. 4, pp. 2152–2159, 2012.
- [12] Chen, Z., Wang, C., Wen, C., Teng, X., Chen, Y., Guan, H., Luo, H., Cao, L., Li, J., "Vehicle detection in high-resolution aerial images via sparse representation and superpixels". *IEEE Trans. on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 103–116, 2016.
- [13] Cao, L., Jiang, Q., Cheng, M., Wang, C., "Robust vehicle detection by combining deep features with exemplar classification." *Neurocomputing*, vol. 215, pp. 225–231, 2016.
- [14] Hassaballah, M., Awad, A. I., "Detection and description of image features: An introduction." In: *Image feature detectors and descriptors*, pp. 1–8, 2016.
- [15] Ojala, T., Pietikainen, M., Maenpaa, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary pattern." *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.
- [16] Hassaballah, M., Hammam A. Alshazly, and Abdelmgeid A. Ali. "Ear recognition using local binary patterns: A comparative experimental study." *Expert Systems with Applications*, vol. 118, pp. 182-200, 2019.
- [17] Alsadeh, M., Joan, L., Qiang, X., Jinwen, M. "Optimizing Gabor Filter and Local Binary Patterns for multi-texture classification." *Egyptian Computer Science Journal*, vol. 42, no. 2, pp. 55-67, 2018.
- [18] Zhao-Guang, L., Yang, Y., Ji, X., "Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the YCbCr color space." *Signal, Image and Video Processing*, vol. 10, no. 2, pp. 277–284, 2016.
- [19] Chanh, J. Kim, C., "A unified spectral-domain approach for saliency detection and its application to automatic object segmentation." *IEEE Tran. on Image Processing*, vol. 21, no. 3 pp. 1272-1283, 2012.
- [20] Guo, Z., Zhang, L. Zhang, D., "Rotation invariant texture classification using LBP variance (LBPV) with global matching." *Pattern Recognition*, vol. 43, no.3, pp. 706-719, 2010.
- [21] Aczél, J., Daróczy, Z., "Charakterisierung der entropien positiver ordnung und der shannonschen entropie." *Acta Mathematica Hungarica*, vol. 14, no. 1, pp. 95-121., 1963.

- [22] Merhan, A.A., Ahmed, A.O., Mohamed, H.A., Mohamed, N.A. "Intelligent prediction of breast cancer: A comparative study." Egyptian Computer Science Journal, vol. 42, no. 3, pp. 29-43, 2018.
- [23] Yang, L., Ping, L., Chen, C. L., Xiaoou, T., "A Large-Scale Car Dataset for Fine-Grained Categorization and Verification." In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3973-3981, 2015.
- [24] Agarwal, S., Awan, A., Roth, D., "Learning to detect objects in images via a sparse, part-based representation." IEEE Tran. on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp. 1475–1490, 2004.
- [25] Fergus, R., Perona, P., Zisserman, A., "Object class recognition by unsupervised scale-invariant learning." In IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 264-271, 2003.
- [26] Fritz, M., Leibe, B., Caputo, B., Schiele, B., "Integrating representative and discriminant models for object category detection." In IEEE Conference on Computer Vision (ICCV), vol. 2, pp. 1363-1370, 2005.
- [27] Kapoor, A., Winn, J., "Located hidden random fields: Learning discriminative parts for object detection." In European Conference on Computer Vision (ECCV), pp. 302-315, 2006.
- [28] Mutch, J., Lowe, D., "Multiclass object recognition with sparse, localized features." In IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 11-18, 2006.
- [29] Wu, B., Nevatia, R., "Cluster boosted tree classifier for multi-view, multi-pose object detection." In IEEE Conference on Computer Vision (ICCV), pp. 1-8, 2007.
- [30] Leibe, B., Leonardis, A., Schiele, B., "Robust object detection with interleaved categorization and segmentation." International Journal of Computer Vision, vol. 77, no. 1, pp. 259-289, 2008.