International Journal of Innovative
Computing, Information and Control
Volume **15**, Number **3**, June **2019**

ICIC International ©2019 ISSN 1349-4198
pp. 923–938

PEDESTRIAN DETECTION BASED ON IMAGE ENHANCEMENT

Ling Xiao, Yongjun Zhang*, Yuewei Li and Qian Wang

Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province
College of Computer Science and Technology
Guizhou University
No. 2708, Huaxi Road, Huaxi District, Guiyang 550025, P. R. China
{ xl201304; hesjed; wangqyzm }@163.com; *Corresponding author: zyj6667@126.com

Received August 2018; revised December 2018

Abstract. *Pedestrian detection is one of hot topics in the field of computer vision and pattern recognition, which is of great value to video surveillance, intelligent traffic and human-computer interaction, etc. and how to improve detection rate and speed is the key. In this paper, a pedestrian detection method based on image enhancement is proposed to improve the detection rate. Before extracting Local Binary Pattern (LBP) feature, we perform interpolation for regions of interest to enhance images. While LBP feature based on grayscale images loses some information on color channels, therefore, this paper further proposes Image Enhancement-Colorized Local Binary Pattern (IE-CLBP) based on color spaces after image enhancement, which describes texture information of images more comprehensively. The experimental results show that the detection rate of IE-LBP (Image Enhancement-Local Binary Pattern) increases by 3.2% compared to that of Uniform Local Binary Pattern (ULBP) feature. Besides, the detection rate of IE-CLBP feature improves by 2.9% than CLBP (Colorized Local Binary Pattern) feature obtained by cascading ULBP feature on color spaces, and the combination of LBP and HOG, CSS features on SVM and HIKSVM classifiers after image enhancement has great performance improvement.*

Keywords: Pedestrian detection, Image enhancement, Local binary pattern, Colorized local binary pattern

1. **Introduction.** At present, pedestrian detection [1-3] has been applied in unmanned systems, but there are still some problems to be solved in daily life because of the surrounding environment and complexity of accidents [4-6]. Therefore, it is very meaningful for us to improve the accuracy and robustness of pedestrian detection. Pedestrian detection algorithms mainly include statistical learning method [7] and background modeling method [8,9]. The former has higher detection rate and robustness, so the statistical learning method receives more attention. The statistical learning method mainly consists of two steps: extracting of features and training of pedestrian detection models and classifiers.

The features commonly used in pedestrian detection have HOG, LBP, CSS, CENTRIST, ICF, etc. and frequently-used classifiers are SVM (Support Vector Machine) and Adaboost. Dalal and Triggs proposed the Histogram of Oriented Gradient (HOG) feature [10] that was widely used in pedestrian detection because it could describe contour of pedestrians well, and it was not easy to be influenced by change of lights. In order to overcome slow calculation of HOG feature, Zhu et al. [11] put forward an accelerated method using the integral histogram to improve the speed. Wojek and Schiele [12] also improved the speed by extracting HOG feature by virtue of GPU. LBP feature [13] proposed by

DOI: 10.24507/ijicic.15.03.923

Ojala and Harwood had been widely used in object detection, for the computation was simple and it had characteristics of grayscale and rotation invariants. The combination of multiple features can describe information of pedestrians more comprehensively, so Walk et al. [14] had achieved good results in pedestrian detection with the combination of CSS (Color Self-Similarities) and HOG features. Wang et al. [15] proposed the cascade of HOG and LBP features and exploited SVM classifier to deal with partial occlusion problems. Wu and Nevatia [16] exploited the combination of HOG feature and covariance feature that achieved significant results in pedestrian detection. Later, Wu et al. [17,18] proposed the Census Transform Histogram (CENTRIST) feature describing global scene information, utilized the integral histogram to speed up the calculation and combined it with Adaboost classifier, which achieved better results. The Integral Channel Feature (ICF) proposed by Dollár et al. [19] combined multiple information of images from different angles and reduced computational complexity by a small number of parameters, but the ICF feature needs multiscale detection, resulting in slow speed of detection. Later, Dollár et al. [20] put forward the Crosstalk method based on the ICF feature to accelerate the sliding window scanning, so that the speed of window scanning increased by 4-30 times. Benenson et al. [21] sped up the detection by training multiscale templates to directly detect images, but only for limited scenes. Zhang et al. [22] on the basis of the ICF feature divided the erected human body into three rectangle detection templates, which effectively improved the detection rate of pedestrian detection. Costea and Nedevschi [23] extracted the HOG, LBP and color features of original images by the dictionary training method, which further improved the detection rate and speed.

The above are traditional methods based on statistical learning that obtain pedestrian detection classifier according to a large number of training samples by learning the transformations of positive and negative samples. Therefore, the method based on statistical learning has a wide range of applicability. Among them, HOG feature is the most effective single feature descriptor of pedestrian, and SVM+HOG method is the pedestrian detection algorithm with better speed and effect balance performance. LBP feature is a simple but very effective texture descriptor which has strong robustness to illumination and is a very powerful feature in texture classification because it can capture the details of the image very well. For its strong discrimination and simple calculations, it has been applied in different scenarios. CSS feature is easily affected by light, so it is often combined with features such as LBP and HOG. However, CENTRIST and ICF features contain lots of redundant information and have less application in pedestrian detection.

In recent years, the development of deep learning has become more and more rapid. Sermanet et al. [24] achieved good results by the Convolution Neural Network (CNN) to detect pedestrians and Chan et al. [25] used the automatic learning filter group to convolve images in order to improve the accuracy of pedestrian classification. Girshick et al. [26] proposed R-CNN (Region-CNN) algorithm for pedestrian detection that exploited the superpixel method to find Regions of Interest (ROI) of each image and then used the CNN algorithm to extract features of the ROI region. Redmon et al. [27] proposed YOLO (You Only Look Once) method based on deep learning which achieved fast pedestrian detection. However, models of deep learning are poor in generalization ability and prone to overfitting and gradient disappearance.

In order to improve the detection rate of pedestrian detection, this paper proposes the method based on image enhancement. The experimental results show that the image enhancement method can enhance texture information of images, and the detection rate of LBP feature extracted from enhanced images is significantly higher than that of original images. While LBP feature extracted from grayscale images will lose some information of color channels, this paper further extracts IE-CLBP feature based on color spaces

in RGB, HSL, HSV, CIELAB, YCbCr, Opponent and L'C'C' space respectively after image enhancement. The experimental results show that the IE-CLBP feature increases the detection rate compared with the CLBP feature, and the detection rates of IE-LBP, IE-CLBP features combined with HOG, CSS features are also improved.

The paper is organized as follows. Section 2 introduces the method proposed in this paper. Firstly, the basic steps of this method are briefly introduced. Then the image enhancement method and the extraction of LBP feature based on the enhanced image are described in detail. Section 3 gives analyses of the method proposed in this paper. Section 4 shows minutely the experimental results. Finally, this paper is summarized in Section 5.

2. **The Proposed Method.**

In this paper, a pedestrian detection method based on image enhancement is proposed. In this section, we first introduce the basic steps of the proposed method, and then detailedly describe the image enhancement method and the extraction method of LBP feature based on the enhanced image.

2.1. **Main steps of the proposed method.**

The main innovation of this paper is to exploit the ideology of image enhancement to enhance the details of the image so as to better process the image and then the LBP feature is extracted from the enhanced image and combined with the HOG, CSS features. The main steps are as follows.

(a) All images of the test sample are enhanced by interpolating columns with the mean values of adjacent two columns and then the interpolated images are resized to the original size by bilinear interpolation method, which is called the image enhancement. The detailed algorithm is shown in Section 2.2.
(b) LBP and CLBP features are extracted from enhanced images and the specific process can be seen in Section 2.3.
(c) Combine the features above with HOG, CSS feature.
(d) Exploit SVM and HIKSVM classifiers to train samples to get pedestrian detection model.
(e) Exploit the above pedestrian classifier to detect the input images.

2.2. **The image enhancement method.**

This section will introduce in detail the method of image enhancement proposed in this paper.

Image enhancement is to process input images so that enhanced images can be better identified. In this paper, before extracting features, we interpolate images by exploiting the mean values of two adjacent pixels between columns as a new pixel to increase details. For an image of $M \times N$ sizes, the size of the image after interpolation is $(2M-1) \times N$, which increases the number of ROI in the multiscale sliding window scanning stage. Consequently, the bilinear interpolation method resizes the interpolated images to the original size in order to reduce the time of sliding scanning.

Bilinear interpolation carries out one-dimensional linear interpolation in two directions. Supposing that the $M(x_1, y_1)$ and $N(x_2, y_2)$ on a straight line are shown on the left of Figure 1, if x is known, y of K point on the straight line can be solved according to (1).

$$y = k \cdot (y_2 - y_1) + y_1, \quad k = \frac{x - x_1}{x_2 - x_1} \tag{1}$$

When $x_1 = j$, $x_2 = j + 1$, $x = j + u$, the value of y is shown in (2).

$$y = u \cdot (y_2 - y_1) + y_1, \quad 0 < u < 1 \tag{2}$$

Bilinear interpolation is based on one-dimensional linear interpolation, and gray values of four neighboring points are weighed to get the gray value of the target point. As shown on the right of Figure 1, the x and y axes are the length and width of the image

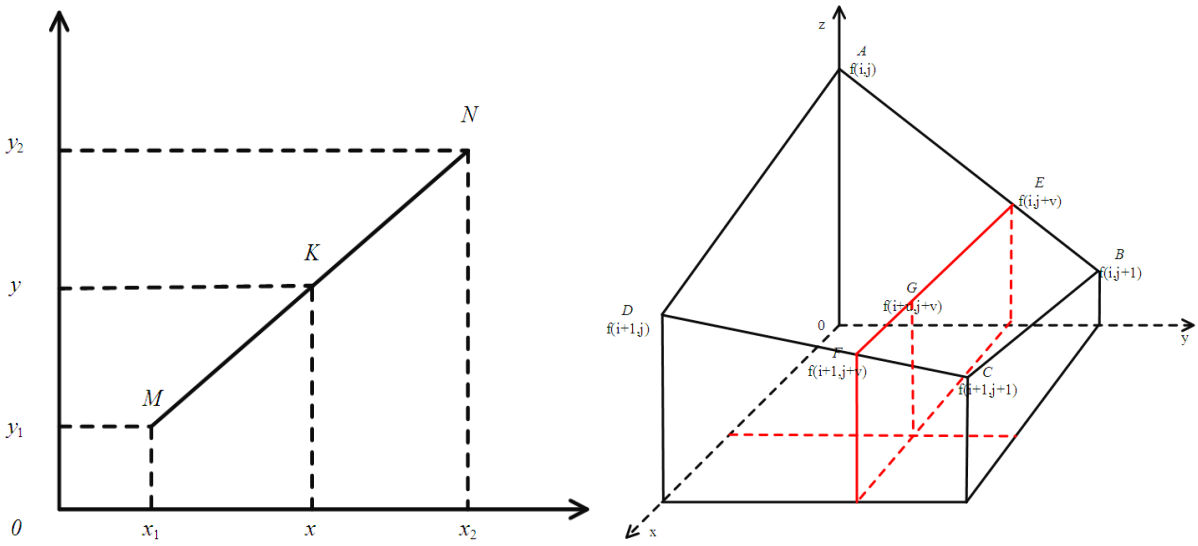


FIGURE 1. One-dimensional linear interpolation and bilinear interpolation

respectively, and the z axis is the gray value of a point in the image. According to (2), we can get gray values of E and F points, as shown in (3) and (4).

$$f(i, j+v) = v \cdot [f(i, j+1) - f(i, j)] + f(i, j),\ 0 < v < 1 \tag{3}$$

$$f(i+1, j+v) = v \cdot [f(i+1, j+1) - f(i+1, j)] + f(i+1, j),\ 0 < v < 1 \tag{4}$$

Similarly, we can obtain the gray value of the target point G to be interpolated, as shown in (5) where $0 < u, v < 1$. Bilinear interpolation considers the influence of four neighborhood points around the target point and for the LBP feature that calculates block sizes of 3×3, and it strengthens the texture characteristics of the image.

$$\begin{aligned} f(i+u, j+v) = {} & u \cdot v \cdot f(i+1, j+1) + u \cdot (1-v) \cdot f(i+1, j) \\ & + v \cdot (1-u) \cdot f(i, j+1) + (1-u) \cdot (1-v) \cdot f(i, j) \end{aligned} \tag{5}$$

2.3. **Extracting LBP and CLBP features from enhanced images.** After introducing the image enhancement method presented in this paper, this section will concretely describe the extraction method of LBP and CLBP features for enhanced test samples.

2.3.1. *LBP feature of image enhancement.* After the above image enhancement is carried out, relevant features are extracted from enhanced images, which proves by experiments that the enhanced images reinforce texture information of images, and the detection rate of LBP feature extracted from the enhanced images is significantly higher than that of original images.

(1) Basic pattern LBP feature of image enhancement. The basic local binary pattern (BasicLBP) feature is an effective texture feature because it is not easily affected by illumination change and describes local texture information of images well. The detailed algorithm to extract LBP feature from enhanced images is shown below.

According to the above image enhancement method, all the test samples are interpolated by mean values of two adjacent columns of pixels, and then the interpolated images are scaled to the original size by bilinear interpolation. Supposing that the gray value of the pixels in the original image is represented by $f(x, y)$, the gray values of the i row pixels are $f(1, i), \ldots, f(k, i)$. The gray value of the image is represented by $g(x, y)$ after interpolation using the mean values of adjacent columns, and the gray values of the i row pixels are $g(1, i), \ldots, g(2k-1, i)$. The relationship between the gray value of the pixels

after interpolation and the gray value of the original image is shown as (6), where N_+ is a positive integer.

$$g(x,i) = \left\{ \begin{array}{ll} f(k,i), & x = 2k-1 \ (k \in N_+) \\ \dfrac{f(k,i)+f(k+1,i)}{2}, & x = 2k \ (k \in N_+) \end{array} \right\} \tag{6}$$

After interpolation, the image is scaled to the original image size by bilinear interpolation. The gray value of the scaled image pixels is shown in (7), where w_1 is the width of the image interpolated by mean values of adjacent columns, and w_2 is the width of the original image size scaled by bilinear interpolation. Assuming the original image size is $M \times N$, then $w_1 = 2M - 1$, $w_2 = M$.

$$h(z,m) = a \cdot g\left(\left\lceil \frac{w_1}{w_2} \cdot z \right\rceil, m\right) + (1-a) \cdot g\left(\left\lfloor \frac{w_1}{w_2} \cdot z \right\rfloor, m\right), \ a = \frac{w_1}{w_2} \cdot z - \left\lfloor \frac{w_1}{w_2} \cdot z \right\rfloor \tag{7}$$

After obtaining the enhanced images above, the image is divided into blocks of the same size, and the LBP value of each point in the block is calculated. The standpoint of calculating the LBP value of each point is that the gray value of the center pixel is regarded as the threshold value of the local area. Then calculating the difference between the threshold value and neighborhood values around the center pixel obtains a binary encoding, and the binary encoding is converted to a decimal number. Finally, the decimal value is used as the LBP value of the point. For $h(x,y)$ point of the image, the point is regarded as center point h_c of the local region and a 3×3 region is formed with the h_c as the center. Then 8 neighborhood points around h_c are defined as $h_1, \ldots, h_8$. Subtracting gray values of the 8 neighborhood points from the gray value of the center point obtains 8 differences, as shown in (8).

$$LBP \approx V\left(t\left(h_1 - h_c\right), \ldots, t\left(h_8 - h_c\right)\right), \quad t(x) = \left\{ \begin{array}{ll} 1, & x > 0 \\ 0, & x \le 0 \end{array} \right\} \tag{8}$$

We can get a binary encoding of 8 bits by (8) and the LBP value of this point is weighed for each binary number, as shown in (9).

$$LBP(x,y) = \sum_{i=1}^{8} t(h_i - h_c) \cdot 2^{(i-1)} \tag{9}$$

According to the above formula, we can obtain the LBP value of a pixel point after image enhancement. The calculation of LBP feature extracted from enhanced images is shown in Figure 2, and the specific calculation process is as follows.

(a) When calculating a block in 3×3 domain, the mean values of adjacent columns are interpolated between columns, as shown on the second graph in Figure 2.
(b) The bilinear interpolation is used to scale the interpolated image to 3×3 sizes, as shown on the third graph in Figure 2.
(c) The threshold value of the block is the gray value 49 of the center point. Gray values of 8 neighborhood points of the block severally subtract the threshold value. When the result is greater than 0, the value of this point is 1 and when the result is less than or equal to 0, the value of this point is 0, as shown on the fourth graph in Figure 2.
(d) Taking the point above the center of the block as the starting point, the values multiply $2^0, 2^1, \ldots, 2^7$ according to clockwise direction respectively.
(e) The binary encoding obtained by the above calculation is 10001010 and decimal value is 138, so the LBP value of this point is 138.

(2) ULBP feature of image enhancement. The BasicLBP feature is defined on the square region size of 3×3. The neighborhood is too local to obtain multi-scale texture

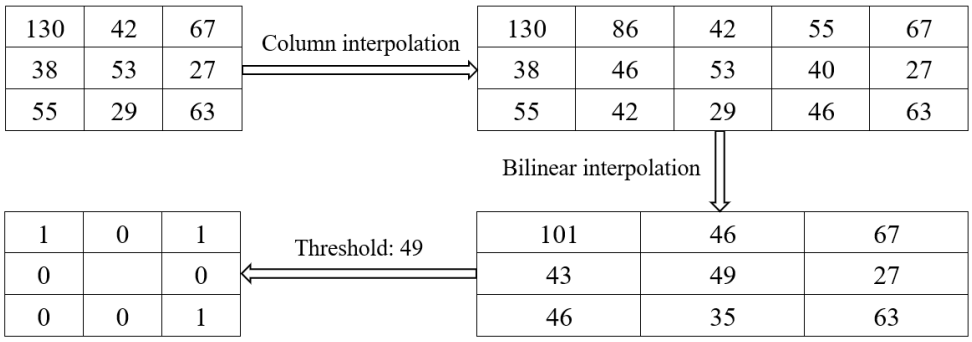


FIGURE 2. The calculating example of LBP feature extracted from enhanced images

information and there is no rotation invariant characteristic. In view of these defects, Ojala et al. proposed a circular neighborhood system that adaptively adjusted the size of the circular region and extended LBP feature to the multi-scale direction, which enhanced the robustness of the image rotation transformation. However, the dimension of BasicLBP feature is too high. For the local area with a radius of r, if there are p sampling points, there will be 2^p species generated, so Ojala et al. put forward ULBP feature to reduce dimension.

ULBP feature extracted from the enhanced image is similar to the BasicLBP feature in processing the image. The calculation of ULBP feature is as follows: a binary encoding is obtained on the basis of BasicLBP feature. Then the frequencies of transformation from 0 to 1 or 1 to 0 in the binary encoding are counted. When the frequencies are less than or equal to 2, each binary encoding corresponds to an equivalent pattern, and other binary encodings are summed up as a pattern. The frequency of transformation U is shown as (10).

$$U = \sum_{i=1}^{p-1} |t(h_{i+1} - h_c) - t(h_i - h_c)| \tag{10}$$

Among them, the definition of h_i references (8). When U is less than or equal to 2, all patterns belong to uniform mode. When the local area is 3×3 and there are 8 sampling points, the classes of binary patterns are 58. These 58 homogeneous patterns are classified into 1 class and the other is 1 category, so the dimension of ULBP feature is 59. Compared with BasicLBP feature, ULBP feature enormously reduces dimension.

The dimension of BasicLBP feature is too high, which leads to larger time complexity. Therefore, the following methods of image enhancement are based on ULBP feature.

2.3.2. *CLBP feature of image enhancement.* LBP feature based on grayscale images loses information on other color channels, so IE-CLBP feature is extracted from multiple color spaces after image enhancement. The IE-CLBP feature is to process images by the image enhancement method above and then the ULBP feature is extracted from each channel of the processed image. Finally, the ULBP feature vectors of the three channels are cascaded.

In this paper, ULBP feature is extracted from the enhanced images in RGB, HSL, HSV, CIELAB, YCbCr, L'C'C' and Opponent color spaces respectively. The RGB is composed of red, green and blue channels. HSL and HSV are relatively similar spaces, consisting of hue, saturation, lightness and hue, saturation, brightness. CIELAB is a color space based on physiological characteristics, consisting of luminance, red-green and yellow-blue intervals. YCbCr has lots of applications in image compression, consisting of brightness, blue and red chromaticity. L'C'C' and Opponent color spaces are obtained by linear

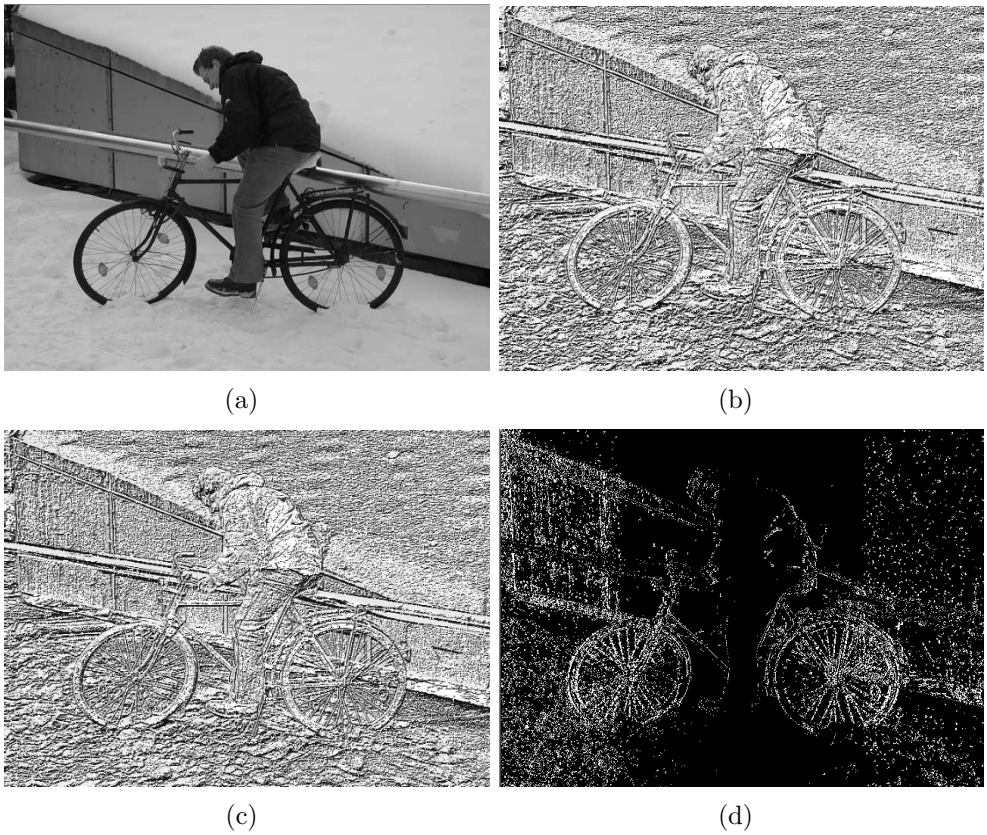
transformation of RGB space, and calculation methods are shown in (11) and (12).

$$\begin{aligned} L' &= (0.299R + 0.587G + 0.114B)/255.0 \\ C_1' &= (0.5R + 0.5G - 1.0B)/255.0 \\ C_2' &= (0.866R - 0.866G)/255.0 \end{aligned} \tag{11}$$

$$\begin{aligned} O_1 &= (R - G)/\sqrt{2.0} \\ O_2 &= (R + G - 2B)/\sqrt{6.0} \\ O_3 &= (R + G + B)/\sqrt{3.0} \end{aligned} \tag{12}$$

3. **Analyses of the Proposed Method.** In this section, we will analyze the LBP feature extraction method based on image enhancement, and compare the LBP feature extracted from the original image with the enhanced image.

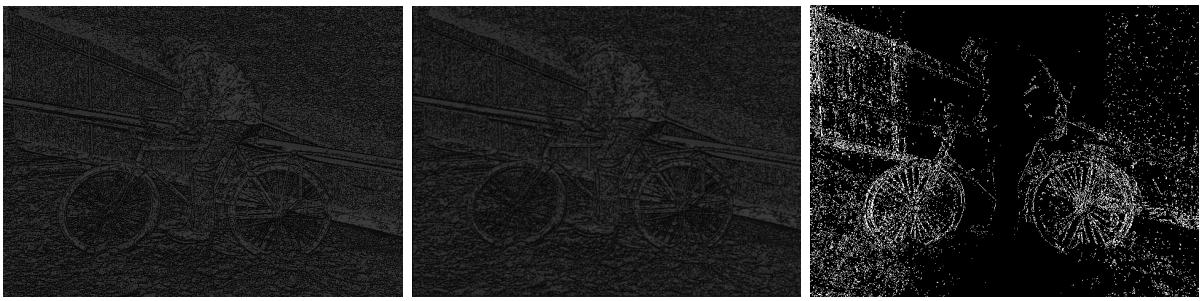
The images are expanded by using the mean values of adjacent columns to make interpolation, which makes up for the lost information of the original images. Then, the interpolated images are resized to the original size by the bilinear interpolation method, which reduces the number of ROI regions in the phase of sliding scanning windows and also enhances details of the original images. The experimental results show that the image enhancement method obviously enhances texture information of images, as shown in Figure 3 and Figure 4.



(a) (b)

(c) (d)

FIGURE 3. Comparison of BasicLBP feature extracted from the original image and the enhanced image: (a) grayscale image, (b) BasicLBP feature extracted from the original image, (c) BasicLBP feature extracted from the enhanced image and (d) the difference value of (b) subtracted (c)



(a) (b) (c)

FIGURE 4. Comparison of ULBP feature extracted from the original image and the enhanced image: (a) ULBP feature extracted from the original image, (b) ULBP feature extracted from the enhanced image and (c) the difference value of (a) subtracted (b)

Figure 3 shows the comparison of BasicLBP feature extracted from the enhanced image and the original image. The (a) is the grayscale image of the original image and (b), (c) are BasicLBP feature extracted from the original image and the enhanced image. (d) is the difference value of (b) subtracted (c), since the (b) and (c) are not distinctly different from the surface, but the result of the difference between the enhanced image and the original image is extraordinarily obvious. The contours and details of objects such as pedestrians and bicycles can be clearly seen from (d), indicating that the extracting BasicLBP feature after the image enhancement can obtain more abundant information than the original image.

Figure 4 shows the comparison of ULBP feature extracted from the original image and the enhanced image, where (a), (b) are ULBP feature extracted from the original image and the enhanced image respectively. (c) is the difference value of (a) subtracted (b). We can also get information about objects such as pedestrians and bicycles in (c), which illustrates that ULBP feature after the image enhancement can acquire more information than the original image.

4. **Experimental Results.** In this section, we will carry out a series of experiments to prove that the proposed image enhancement method can enhance the texture information of the image.

The experimental platform is Intel i7@3.6GHZ CPU and 8G RAM, and the experimental environment is Visual Studio 2013 + OpenCV2.4.10. The experimental sample is the INRIA data set that has 2416 positive training samples, 1218 negative training samples, 1126 positive test samples and 453 negative test samples. Each positive sample extracts with the size of 64×128 ROI and each negative sample is randomly cut out 10 ROI with the size of 64×128.

4.1. **Comparison of three image enhancement methods.** After interpolating, nearest neighbor interpolation, bilinear interpolation and cubic interpolation are used to resize the interpolated image to the original image size respectively. As shown in Figure 5, when FPPW $= 10^{-4}$, the performance of the nearest neighbor interpolation and the bilinear interpolation is almost identical, and the detection rate using the cubic interpolation is the lowest. The nearest neighbor interpolation exploits the gray value of the nearest point from target point as the value. For calculating the local region of LBP feature, the continuity of gray values of points will be more stable. Bilinear interpolation exploits gray values of the 2×2 neighborhood of the target point, which enhances the texture

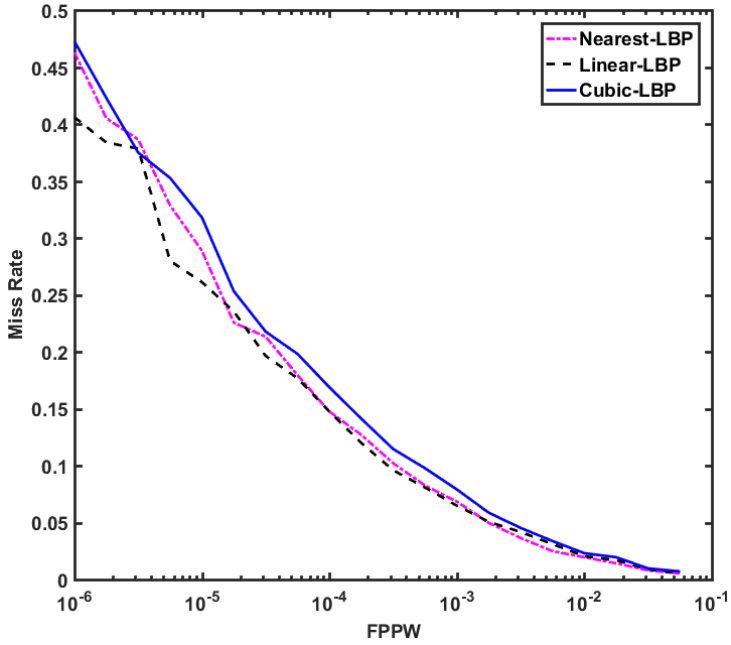


FIGURE 5. Comparison of three image enhancement methods for IE-LBP

characteristics of the 3×3 regions for LBP feature. Cubic interpolation is weighed by gray values of the 4×4 neighborhood to get the value of the target point. When calculating the 3×3 sizes of blocks of the LBP feature, it produces inferior results because small local region has a certain continuity in gray values and if the area is bigger, the difference of gray values will be greater, which will make gray values of the 3×3 sizes of blocks discontinuous. Besides, the cubic interpolation has higher computational complexity than the bilinear interpolation and the nearest neighbor interpolation, and the detection rate is lower. While stability and the detection rate for the bilinear interpolation are higher than the nearest neighbor interpolation, consequently, the method in this paper is based on the bilinear interpolation.

4.2. **Comparison of IE-LBP and other LBP features.** ULBP feature extracted from enhanced images is compared with ULBP and BasicLBP features extracted from original images. In this paper, square (neighborhood points in the square area) and circular (neighborhood points in the circular region) LBP features are extracted respectively. As shown in Figure 6, when FPPW $= 10^{-4}$, the detection rate of square IE-LBP increases by 3.2% and 1.8% respectively than those of square ULBP and BasicLBP, and the detection rate of circular IE-LBP is 1.3% higher than that of circular ULBP. Their detection rates are shown in Table 1. Because the continuity of the local gray value for the calculated LBP feature block will be more stable after above image enhancement, and the texture characteristics of the local area will be further enhanced. For LBP feature extracted from the circular or square local area, the detection rate of the enhanced image is higher than that of the original image. Furthermore, the detection rate of the circular IE-LBP is 4.4% higher than that of the square because the square neighborhood system is too local to obtain multi-scale texture information, while the circular region has the characteristic of rotation invariance. However, the circular LBP feature has high computational complexity, and the subsequent extraction of LBP feature is based on the square area.

4.3. **Comparison between IE-LBP+HOG and ULBP+HOG.** The combination of multiple features can further improve the detection rate of pedestrian detection, so

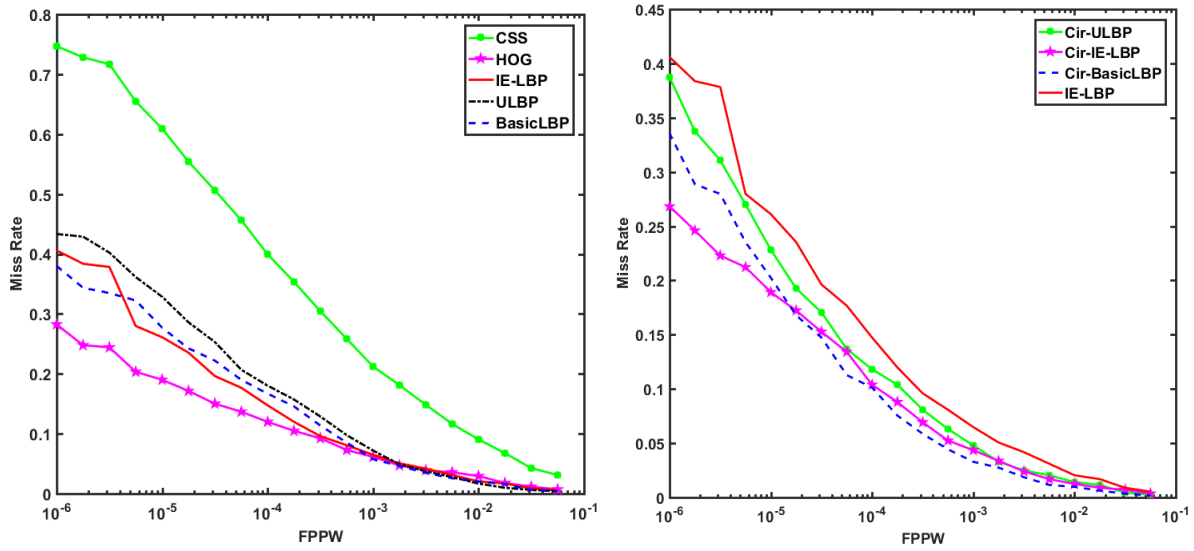


FIGURE 6. IE-LBP compared with other LBP features

TABLE 1. The detection rates of LBP features

Feature	Dimension	Detection rate
Circular IE-LBP	6195	89.5%
Circular ULBP	6195	88.2%
Circular Basic LBP	26880	89.7%
Square IE-LBP	6195	85.1%
Square ULBP	6195	81.9%
Square Basic LBP	26880	83.3%

this paper combines ULBP feature with HOG feature to detect pedestrians after image enhancement. As shown in Figure 7, when FPPW = 10^{-4}, the detection rate of IE-LBP+HOG increases by 9.25% than that of IE-LBP and the detection rate of IE-LBP+HOG is 1.15% higher than that of ULBP+HOG, which verifies the robustness of the image enhancement method. Because LBP feature describes the texture characteristics of pedestrians, while HOG feature describes edge contour information of pedestrians, the combination of the two features can obtain more information and greatly improve the detection rate. Furthermore, the above mentioned image enhancement method evidently enhances the local texture information of the image, and does not reduce the texture information of the image after combining with HOG feature, so the detection rate of combining LBP feature with HOG feature extracted from the enhanced image is still higher than the original image.

4.4. **Comparison of IE-LBP+CSS and ULBP+CSS.** As mentioned above, the combination of multiple features can further improve the detection rate, and CSS feature is easily affected by light, so it is frequently combined with other features in pedestrian detection. As shown in Figure 8, when FPPW = 10^{-4}, the detection rate of CSS feature is only 60% and the detection rates of ULBP, IE-LBP features combined with CSS feature are 89%, 90% respectively. It can be seen that the detection rate of the combination of CSS and LBP features has been significantly improved, as LBP feature has good stability in light change and can make up for the defects of CSS feature affected by illumination change. Moreover, it is similar to the combination of LBP and HOG features, and the

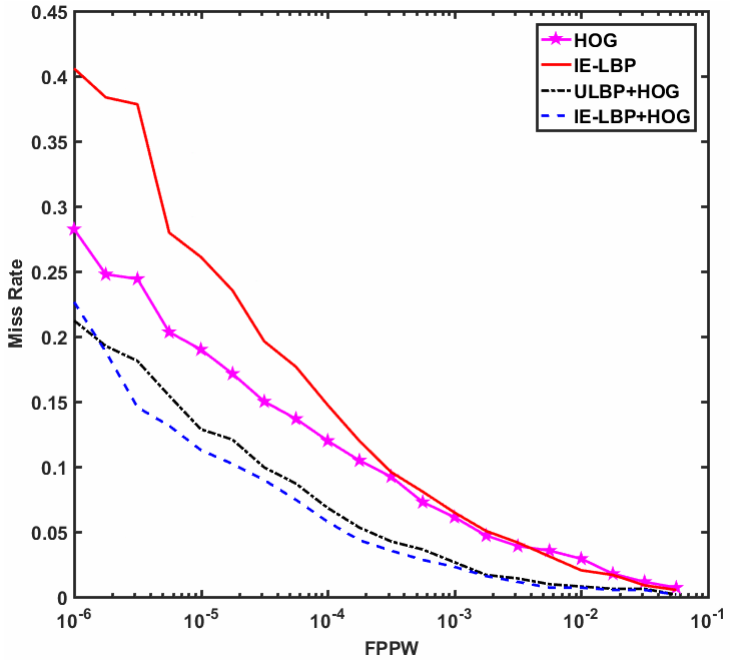


FIGURE 7. IE-LBP+HOG compared with other features

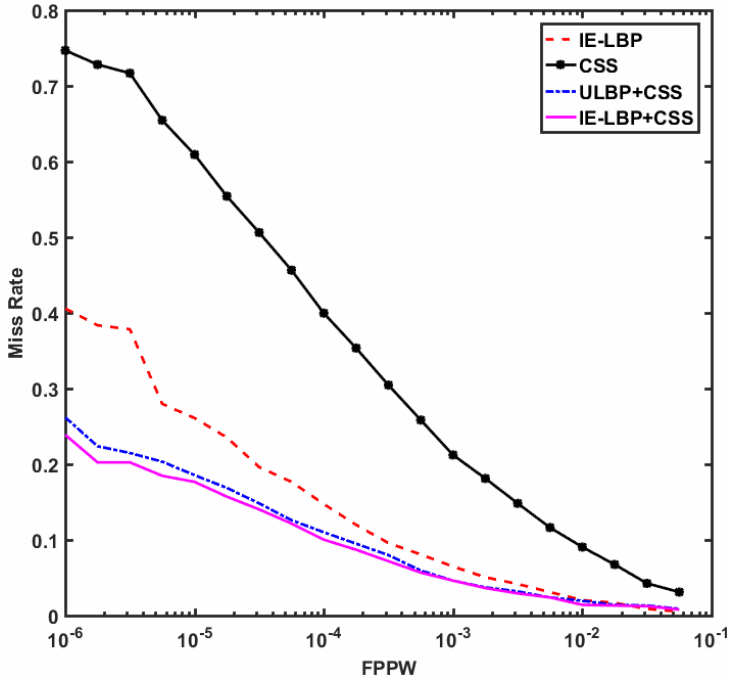


FIGURE 8. IE-LBP+CSS compared with other features

texture information of the image also can be enhanced by extracting LBP feature from the enhanced image and then combining it with CSS feature, so the combination of LBP and CSS features extracted from enhanced images is superior to original images, thus further proving that the image enhancement method can enhance texture information of images.

4.5. **Comparison of IE-LBP, ULBP, IE-LBP+HOG and ULBP+HOG on SVM classifier and HIKSVM classifier.** SVM is a common classifier in pedestrian detection,

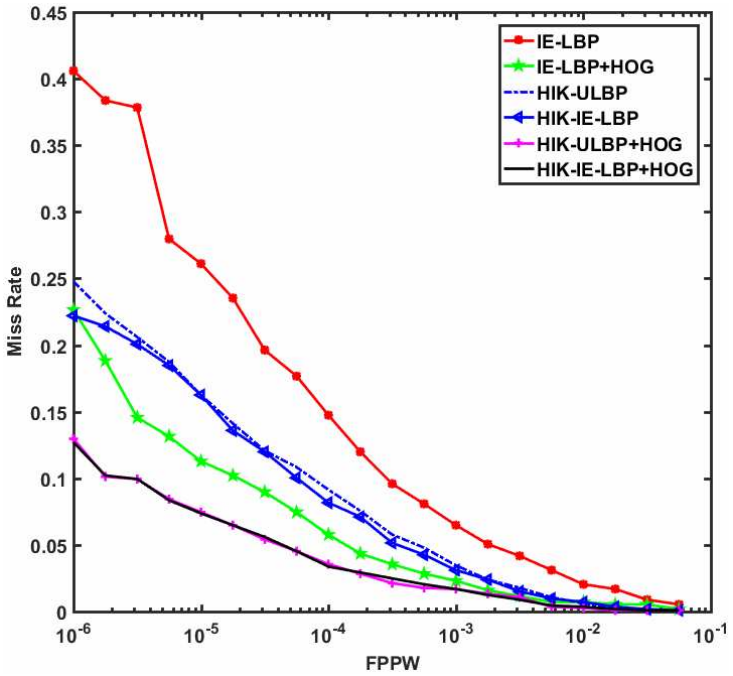


FIGURE 9. Comparison between IE-LBP and IE-LBP+HOG on SVM and HIKSVM classifiers

TABLE 2. The detection rates of different features on SVM and HIKSVM classifiers

	Feature	Dimension	Detection rate
SVM	ULBP+HOG	9975	93.2%
	IE-LBP+HOG	9975	94.35%
	ULBP+CSS	14323	89%
	IE-LBP+CSS	14323	90%
HIKSVM	ULBP	6195	90.7%
	IE-LBP	6195	91.8%
	ULBP+HOG	9975	96.4%
	IE-LBP+HOG	9975	96.6%

which has good classification effect, while HIKSVM classifier has better classification performance. Because HIKSVM iterates over negative samples in the training process and looks for some images which are easily mistaken for pedestrians as negative samples, the accuracy of weight coefficients is higher. However, the computational complexity is greater. As shown in Figure 9, the detection rates of IE-LBP, IE-LBP+HOG combined with the HIKSVM classifier are 91.8%, 96.6% and increase by 6.7%, 2.25% than the SVM classifier respectively. Besides, the detection rates of LBP, LBP+HOG features combined with the HIKSVM classifier after image enhancement are 1.1%, 0.2% higher than those of original images. Their detection rates are shown in Table 2. The classifier does not affect the effect of image enhancement, so the result of extracting LBP and LBP+HOG features combined with HIKSVM classifier after image enhancement is also better than the original image. However, when HIKSVM classifier is combined with LBP+HOG, the detection rate has reached a high state, so the improvement is not very obvious.

4.6. **Comparison of IE-CLBP and CLBP in different color spaces.** LBP feature based on grayscale images will lose information on the color channels, so this paper

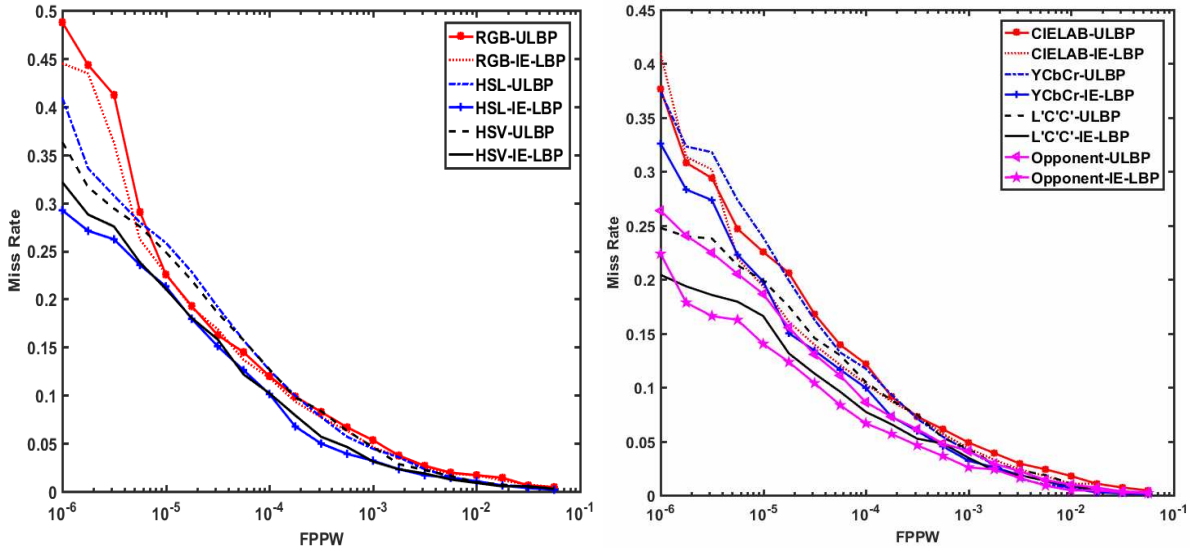


FIGURE 10. Comparison of IE-CLBP in different color spaces

TABLE 3. The detection rates of CLBP and IE-CLBP

	RGB	HSL	HSV	YCbCr	CIELAB	L'C'C'	Opponent
CLBP	87.8%	87%	87%	88%	87.8%	89.4%	91.2%
IE-CLBP	88%	89.9%	89.8%	90%	89.8%	92.1%	93.4%
Improvement	0.2%	2.9%	2.8%	2.0%	2.0%	2.7%	2.2%
Dimension	18585	18585	18585	18585	18585	18585	18585

extracts CLBP feature in RGB, HSL, HSV, CIELAB, YCbCr, L'C'C' and Opponent spaces of the original and enhanced images. As shown in Figure 10, the detection rate of IE-CLBP feature extracted from RGB space is only 0.2% higher than that of CLBP feature, indicating that the effect of the image enhancement method in RGB space is not very significant because the RGB channel cannot reflect the specific color information of the object very well. Nevertheless, the detection rates of IE-CLBP feature in HSL, HSV, CIELAB, YCbCr, L'C'C' and Opponent spaces increase by 2.9%, 2.8%, 2%, 2%, 2.7% and 2.2% respectively compared with those of CLBP feature, their detection rates are shown in Table 3, since these color spaces can process colors better digitally, and according to the local similarity principle of color, choosing the appropriate color space can handle images better. Among them, the detection rate of IE-CLBP feature based on Opponent color space is the highest and is 8.3% and 11.5% higher than those of IE-LBP feature and ULBP feature, respectively.

4.7. **Performance comparison of IE-CLBP combined with HOG, HOG+CSS.** As shown in Figure 10, IE-CLBP feature based on Opponent color space is the best, so this paper combines IE-CLBP feature based on the space with HOG, HOG+CSS features to further improve the detection rate, and their detection rates are shown in Table 4. As shown in Figure 11, the detection rate of IE-CLBP+HOG is 2.05% higher than that of IE-LBP+HOG. Furthermore, the detection rates of IE-CLBP+HOG+CSS, IE-LBP+HOG+CSS increase by 0.2%, 0.45% higher than those of IE-CLBP+HOG, IE-LBP+HOG, which also illustrates that the effect of combining CSS feature at this time is not obvious because IE-CLBP+HOG and IE-LBP+HOG have achieved high detection rates that almost reach the saturation state.

TABLE 4. Detection rates of IE-CLBP combined with HOG and HOG+CSS

Feature	Dimension	Detection rate
IE-CLBP+HOG	22365	96.4%
IE-LBP+HOG+CSS	18103	94.8%
IE-CLBP+HOG+CSS	30493	96.6%

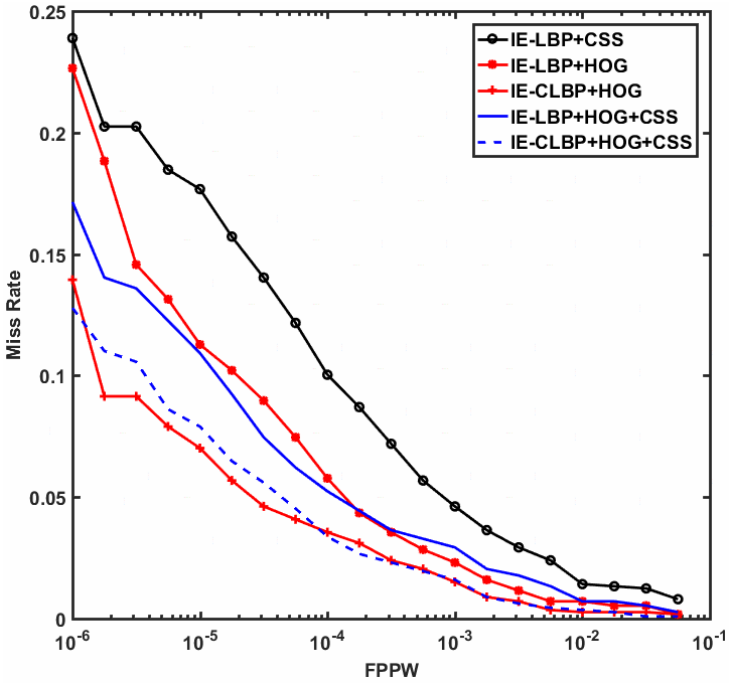


FIGURE 11. Combination of IE-CLBP and HOG, HOG+CSS

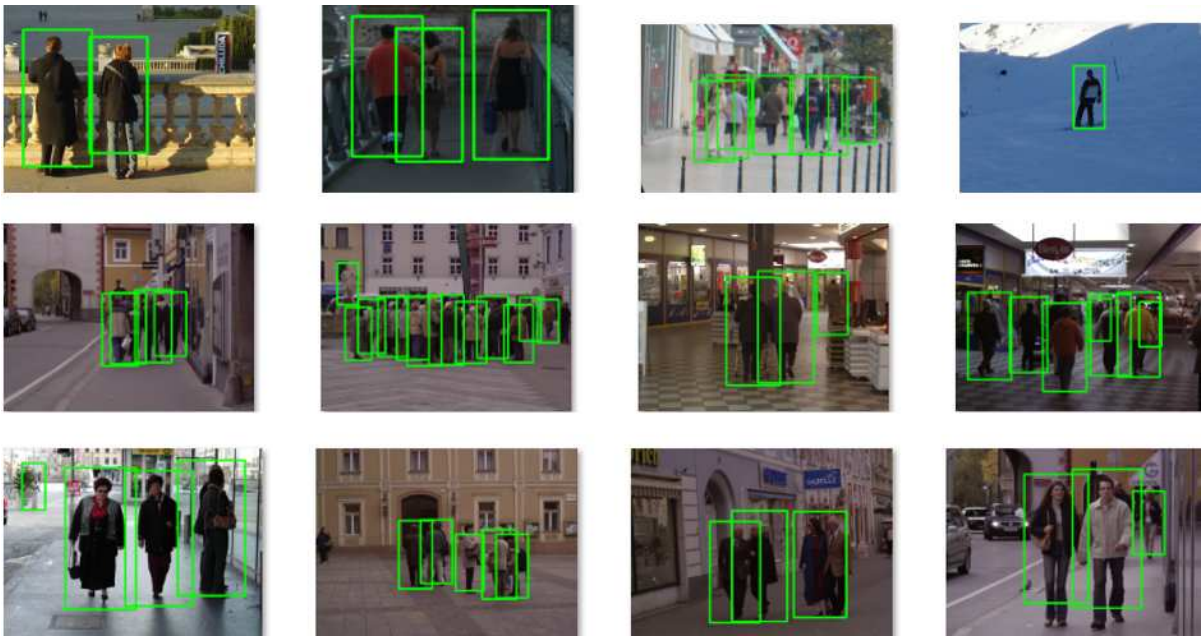


FIGURE 12. Part detection results of INRIA data set

4.8. **Detection results.** Extracting LBP and HOG features from enhanced images to detect test samples of INRIA is shown in Figure 12. As we can see, pedestrians can almost all be detected in images, but there are some errors due to some factors such as complexity of backgrounds. For example, the last image of the second line has two false pedestrians and the second picture of the third line has a missing pedestrian.

5. **Conclusions.** This paper proposes the pedestrian detection method based on image enhancement. After enhancing images, we combine LBP feature with HOG and CSS features respectively. Experimental results show that the image enhancement method can significantly improve texture information of images and LBP+HOG, LBP+CSS extracted from enhanced images with SVM and HIKSVM classifiers obtain higher detection rates than the same classifiers based on original images. LBP feature extracted from multi-color spaces can compensate for information loss of images. The experimental results also indicate that the detection rates of CLBP feature extracted from enhanced images in these color spaces are significantly higher than those original images. However, the combination of multiple features increases the dimension of features, and we will research how to improve the speed and reduce computational complexity with the premise of guaranteeing the detection rate.

Acknowledgment. This work was supported by the Joint Fund of Department of Science and Technology of Guizhou Province and Guizhou University under grant: LH [2014]7635, Research Foundation for Advanced Talents of Guizhou University under grant: (2016) No. 49, Key Supported Disciplines of Guizhou Province – Computer Application Technology (No. QianXueWeiHeZi ZDXX[2016]20), Specialized Fund for Science and Technology Platform and Talent Team Project of Guizhou Province (No. QianKeHePingTaiRenCai [2016]5609), and the work was also supported by National Natural Science Foundation of China (61462013, 61661010).

REFERENCES

[1] S. Zhang, R. Benenson, M. Omran, J. Hosang and B. Schiele, How far are we from solving pedestrian detection, *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp.1259-1267, 2016.

[2] J.-Q. Gui and Z.-M. Lu, Pedestrian detection based on modified dynamic background using Gaussian mixture models and HOG-SVM detection, *International Journal of Innovative Computing, Information and Control*, vol.14, no.1, pp.279-295, 2018.

[3] Y. Tian, P. Luo, X. G. Wang and X. Tang, Deep learning strong parts for pedestrian detection, *IEEE International Conference on Computer Vision*, Santiago, Chile, pp.1904-1912, 2016.

[4] R. Cheng, Y. Zhang and G. Wang, Haar-like multi-granularity texture features for pedestrian detection, *International Journal of Image & Graphics*, vol.17, no.4, 2017.

[5] Y. Zhang, Y. Zhao, G. Li, D. Wei and R. Cheng, Efficient and real-time pedestrian detection at night-time environments, *International Journal of Innovative Computing, Information and Control*, vol.11, no.2, pp.599-614, 2015.

[6] P. Mukilan and A. Wahi, An efficient human object detection and tracking with the aid of morphological operation and optimization algorithm, *International Journal of Innovative Computing, Information and Control*, vol.11, no.4, pp.1139-1153, 2015.

[7] T. Yang, J. Li, Q. Pan and Y. N. Zhang, Scene modeling and statistical learning based robust pedestrian detection algorithm, *Acta Automatica Sinica*, vol.36, no.4, pp.499-508, 2010.

[8] Z. Zhong, B. Zhang and G. Lu, An adaptive background modeling method for foreground segmentation, *IEEE Trans. Intelligent Transportation Systems*, vol.18, no.5, pp.1109-1121, 2017.

[9] Y. Xu, J. Dong, B. Zhang and D. Xu, Background modeling methods in video analysis: A review and comparative evaluation, *CAAI Trans. Intelligence Technology*, vol.1, no.1, pp.43-60, 2016.

[10] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, USA, pp.886-893, 2005.

[11] Q. Zhu, S. Avidan, M. C. Yeh and K. T. Cheng, Fast human detection using a cascade of histograms of oriented gradients, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, USA, pp.1491-1498, 2006.

[12] C. Wojek and B. Schiele, A performance evaluation of single and multi-feature people detection, *DAGM Symposium on Pattern Recognition*, Munich, Germany, pp.82-91, 2008.

[13] T. Ojala and I. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, vol.29, no.1, pp.51-59, 1996.

[14] S. Walk, N. Majer, K. Schindler and B. Schiele, New features and insights for pedestrian detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, USA, pp.1030-1037, 2010.

[15] X. Wang, X. Tony and S. Yan, An HOG-LBP human detector with partial occlusion handling, *International Conference on Computer Vision*, Kyoto, Japan, pp.32-39, 2009.

[16] B. Wu and R. Nevatia, Optimizing discrimination-efficiency tradeoff in integrating heterogeneous local features for object detection, *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, USA, pp.1-8, 2008.

[17] J. Wu and J. M. Rehg, CENTRIST: A visual descriptor for scene categorization, *IEEE Trans. Pattern Analysis & Machine Intelligence*, vol.33, no.8, pp.1489-1501, 2011.

[18] J. Wu, N. Liu, C. Geyer and J. M. Rehg, A real-time object detection framework, *IEEE Trans. Image Processing*, vol.22, no.10, pp.4096-4107, 2013.

[19] P. Dollár, Z. Tu, P. Perona and S. Belongie, Integral channel features, *British Machine Vision Conference*, London, UK, pp.1-11, 2009.

[20] P. Dollár, R. Appel and W. Kienzle, Crosstalk cascades for frame-rate pedestrian detection, *European Conference on Computer Vision*, Florence, Italy, pp.645-659, 2012.

[21] R. Benenson, M. Mathias, R. Timofte and G. L. Van, Pedestrian detection at 100 frames per second, *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, USA, pp.2903-2910, 2012.

[22] S. Zhang, C. Bauckhage and A. B. Cremers, Informed Haar-like features improve pedestrian detection, *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, pp.947-954, 2014.

[23] A. D. Costea and S. W. Nedevschi, Channel based multiscale pedestrian detection without image resizing and using only one classifier, *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, pp.2393-2400, 2014.

[24] P. Sermanet, K. Kavukcuoglu, S. Chintala and Y. Lecun, Pedestrian detection with unsupervised multi-stage feature learning, *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, USA, pp.3626-3633, 2013.

[25] T. H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng and Y. Ma, PCANet: A simple deep learning baseline for image classification?, *IEEE Trans. Image Processing*, vol.24, no.12, pp.5017-5032, 2014.

[26] R. Girshick, J. Donahue, T. Darrell and J. Malik, Region-based convolutional networks for accurate object detection and segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.38, no.1, pp.142-158, 2016.

[27] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You Only Look Once: Unified, real-time object detection, *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp.779-788, 2016.