



Detection of Bikers without Helmet Using Image Texture and Shape Analysis

R. Badaghi^a, H. Hassanpour^{*a}, T. Askari^b

^a Faculty of Computer Engineering and IT, Shahrood University of Technology, Shahrood, Iran

^b Computer Science Department, Faculty of Mathematics and Computer, Higher Education Complex of Bam, Bam, Iran

PAPER INFO

Paper history:

Received 11 November 2020

Received in revised form 04 January 2021

Accepted 14 January 2021

Keywords:

Helmet

Biker

Detection

Local Binary Pattern

Local Variance

Histogram of Oriented Gradient

ABSTRACT

Helmet are essential for preventing head injuries in bikers. Traffic laws are applied in most countries to bikers who don't wear a helmet. Manually checking bikers for the usage of a helmet is a very costly and tedious task. In this regard, several helmet detection methods were developed in literature for detecting bikers violating the law in recent years. This paper proposes an image processing method based on the Local Binary Pattern (LBP), Local Variance (LV), and Histogram of Oriented Gradient (HOG) descriptors for detection of bikers without a helmet. The innovation of the proposed method is mainly on the feature extraction step, which leads the classification towards appropriately discriminating between the two classes of helmet and non-helmet. The experimental results show our method is superior to the existing methods for helmet detection. The accuracy of the proposed helmet detection method is 98.03% using the Support Vector Machine classifier.

doi: 10.5829/ije.2021.34.03c.09

1. INTRODUCTION

Biker safety helmets are very effective at preventing head injuries in road accidents [1]. Due to its effectiveness, laws in many countries enforce bikers to wear a safety helmet. Unfortunately, some bikers refuse to wear it for various reasons.

Image processing techniques can be used to automatically detect violators [2]. Several studies have been conducted on the recognition of bikers without helmets using image processing techniques. The low quality of traffic images is one of the dilemmas for the detection task. To reduce the processing time for helmet detection, only the area containing the biker's head is considered. In most methods, this issue is considered by using the 1/5 (one-fifth) top of the image as the Region of Interest (ROI) (the area containing the biker's head) [3-6]. The process of helmet detection contains three steps: pre-processing, feature extraction, and classification. Different lighting and climate conditions cause unwanted data, such as noise, on traffic images. Therefore, a pre-process is required before the feature extraction step. Feature extraction is the most important step in the bikers' helmet detection. Most of the existing

approaches use several descriptors, based on color, texture, and the geometric shape of the helmet, to extract the features from ROI. Local Binary Pattern (LBP), Histogram of Oriented Gradient (HOG), Circle Hough Transform (CHT), Scale-Invariant Feature Transform (SIFT), and Haar Wavelet are common descriptors in this context [2-6].

This paper proposes a method to detect the biker without a helmet. After a simple pre-processing of images, LBP, Local Variance (LV), and HOG descriptors are used to extract features from the ROI. In order to evaluate the extracted features, a Support Vector Machine (SVM) classifier is used. The rest of the paper is organized as follows. In Section 2, we review previous works on helmet detection. The proposed method is described in Section 3. The results of applying the proposed method on a dataset are provided in Section 4. Finally, the paper is concluded in Section 5.

2. RELATED WORKS

Dinesh et al. provided a method for automatic detection of bikers' non-use of safety helmet [3]. After

*Corresponding Author Institutional Email:
h.hassanpour@shahroodut.ac.ir (H. Hassanpour)

differentiating motorcycles from the other vehicles, the candidate area (head area) from the biker's image is considered for helmet detection. SIFT, LBP, and HOG descriptors were used to extract the features from the head area. The achieved feature vector was employed for classification using SVM. Silva *et al.* suggested the use of CHT technique to determine the position of the biker's head from 1/5 (one-fifth) of the top image [7]. The LBP, HOG, and Wavelet Transform (WT) descriptors were used for feature extraction. They used several classifiers to evaluate the output of the descriptors. Although the computational complexity of the method is significant, the accuracy of the detection is acceptable. Shine and Jiji provided a method for the detection of bikers without helmets [8]. After determining the ROI, the extracted features using the LBP, HOG, and Haarlick descriptors were used for classification. The accuracy of this method is higher than the method proposed by Silva et al. [7].

In another study, Talaulikar et al. used image processing and machine learning techniques to detect the biker's helmet, [9]. They applied median filtering, flood fill, erode and dilate on the ROI. The region area was divided into four quadrants, and the average of intensities and hues were measured in each quadrant. The Principal Component Analysis (PCA) method was applied on the derived features. The accuracy of the method was suitable.

A method was provided to identify construction workers that didn't use safety helmets [10]. Although the purpose of this system is different from our research, the feature extraction step gives useful insights. In this method, the CHT technique is used to detect circular objects inside the image. Since safety helmets used by construction workers are in certain unique colors, i.e. yellow, blue, or red, the circular areas detected by the CHT are searched for helmets in these colors. Color is not a reliable feature for helmet detection in bikers, because the biker's hair may also be the same color as the helmet. The use of CHT technique lonely increases the number of false-positive (FP) predictions in helmet detection, because both the helmet and head are circular.

A study was conducted on workers without helmets in construction sites [11]. Statistical features were calculated from the Gray Level Co-occurrence Matrix (GLCM) after achieving the LBP image. The statistical features include contrast, correlation, entropy, energy, and homogeneity. To evaluate the method, an Artificial Neural Network (ANN) was used. Although the image texture is well analyzed using statistical features, but high time complexity is the disadvantage of the method.

In another study, a method was provided for detecting power substation perambulatory workers without a safety helmet [12]. From the head area, color space transformation and color feature discrimination were extracted. For the color segmentation, the HSV color space (with three channels of Hue, Saturation, and Value)

is more robust than the other color spaces, because this model can easily distinguish the desired color from the range of other colors. Hence, the images were transformed from the RGB to HSV color space. In this method, to segment various colors, two different thresholds were used for the Hue and Saturation channels. The threshold value for the Hue channel (which represents the color) was set manually, whereas the threshold value for the Saturation channel was automatically determined by the OTSU thresholding algorithm. By employing the color segmentation on the two channels, it can be recognized whether the workers are wearing helmets or not.

A detection method for bikers without helmets was proposed using image processing and Convolutional Neural Network (CNN) [13]. Initially, the method identifies bikers using the HOG and the SVM classifier. Then, by applying a CNN to the area of interest, bikers without helmets are identified. While in another method, CNN is used in two steps: discriminating bikers from other vehicles; and the detection of the bikers without helmets. The algorithm was less accurate for motorcycles with a large number of riders or motorcycles with an uncommon passenger position [14].

3. PROPOSED METHOD

In this paper, we propose a new method for detecting bikers not wearing a helmet. After a simple pre-processing, we focus on the feature extraction step. Three descriptors are used to describe samples of the two classes, i.e. helmet and non-helmet.

The SVM is used for classification. Therefore, the proposed method consists of three steps: preprocessing, feature extraction and classification. Each of the three steps is described below.

The LBP extracts homogeneous patterns from the image and pays less attention to heterogeneous patterns. In contrast, heterogeneous patterns in the image are well extracted using the variance [15]. So, simultaneous use of these descriptors can be effective for image texture analyzing and for extracting useful information. On the other hand, the HOG specifies the local shape and the direction of the edges in the image. Hence, this descriptor can describe samples of the two classes. The final descriptor is the feature vector achieved from the descriptors detailed below.

3.1. Database

The database used in this work, thankfully provided by Silva et al. [7], contains 255 images of bikers' head. There are 152 images with helmet, and 103 images without helmet, in the database. Some of the images contain bikers wearing casual hats, which are classified into the non-helmet category. A number of these images are shown in Figure 1. The

images have been captured under different lighting and climate conditions. Also, surveillance cameras were installed far from the road, and they have a low image quality, hence, as seen, the captured images are in low resolution.

3. 2. Pre-processing Initially, images are resized to 40×40 pixels. Usually, shape, texture, and color are used for object feature extraction. Safety helmets are in different colors, but the biker’s hair may be in the same color range as the helmet. So the color is not a trustable feature for this purpose. The descriptors used in the next step are independent of the color space, therefore we obtained gray level images from the RGB color space.

3. 3. LBP Operator The LBP operator is a powerful descriptor to analyze texture information from low-resolution images. The operator is locally applied, reflecting the appearance and the structure of the various regions in the image. The original method is applied on windows of size 3×3 pixels. The LBP uses a thresholding method in which the value of the central pixel is considered as the threshold. The neighboring pixels within the window are labeled considering the threshold value. The pixel is labeled as 0 if its value is smaller than the threshold, otherwise it is labeled as 1. Then, the resulting LBP can be expressed in decimal form as follows:

$$LBP_{P,R} = \sum_{i=0}^{P-1} (S(g_i - g_c)) 2^i, \tag{1}$$

where g_i and g_c are the gray-level values of neighboring pixels and the central pixel, respectively. P is the number of neighborhoods and R is neighborhood radius, where $P = 8$ and $R = 1$ in this research. Function $S(x)$ is a sign function defined as:

$$S(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{2}$$

The decimal value obtained from the encoded bit string (in clockwise direction) replaces the central pixel in the window. Since the LBP is applied to the gray level image, the resulting value is between 0 and 255 (see Figure 2). This process is repeated for all pixels in the image and



Figure 1. Samples of images from the database used in the paper

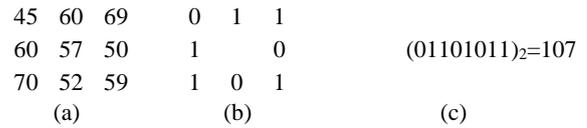


Figure 2. Example of the LBP descriptor calculation. (a) a sample window. (b) Thresholding considering the central pixel. (c) Pattern computed from the threshold result

the histogram obtained from the LBP image is used as a descriptor.

A Uniform Local Binary Pattern (ULBP) that can describe uniform patterns, is one of the basic LBP extensions. In the LBP, a pattern with a maximum of two bitwise transitions is called uniform. For example, the patterns 00000000 (0 transition), 11111110 (1 transition) and 11100111 (2 transitions) are uniform, whereas 00110010 (4 transitions), 10110110 (5 transitions) are not. ULBP operator is based on the original LBP in which neighborhood radius and the number of neighbors can be more than the basis (3×3 neighborhood). The descriptor considers a bin for each uniform pattern. Hence, the number of bins from 256 (different labels that can be obtained with the basic LBP) is reduced to a smaller number [8]. The number of patterns obtained using the descriptor depends on the number of neighbors. For P neighbors, $P \times (P - 1) + 3$ bins are obtained, where the last bin is used for all non-uniform patterns. For example, for $P = 8$, 59 patterns are obtained [16]. We used the LBP_{U2} where $U2$ means uniform, P is the number of neighbors and R is the neighborhood radius (in this paper the best value of P and R are 8 and 1 that were obtained empirically (see Figure 3)).

A biker head image can be represented as a combination of micro-patterns by an LBP histogram. Using one histogram for the whole image cannot encode the shape information and indicate locations of these micro-patterns in the image [17]. Hence, the image is divided into neighboring cells (in a 5 × 5 window) and the LBP histogram is calculated in each cell. In this research, 59 features were extracted. Eventually, these histograms are concatenated into a single histogram feature vector (as shown in Figure 4).

3. 4. Local Variance The Local Variance (LV) defines the gray level distribution between pixels within

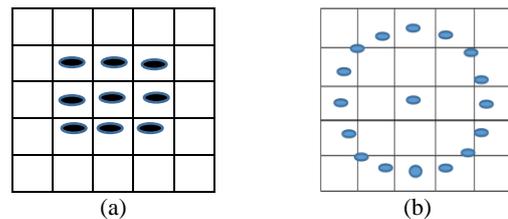


Figure 3. Examples of the Uniform LBP operator; (a) the circular (8, 1) and; (b) the circular (16, 2) neighborhoods

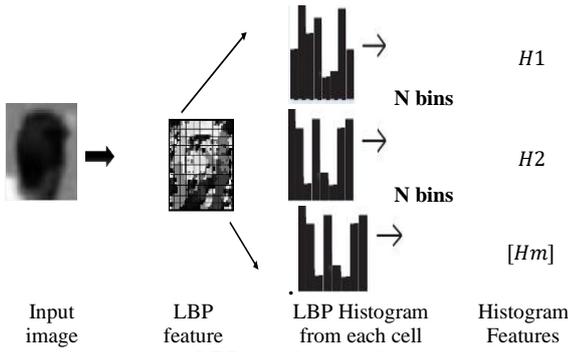


Figure 4. LBP-based image descriptor.

a neighborhood. The smaller the variance, the closer the gray level of pixels within the neighborhood, and vice versa [16]. As mentioned earlier, image variance can extract non-homogeneous patterns. On the other hand, ULBP analyzes homogeneous patterns. The descriptors can be good supplements for extracting the image’s textural patterns. In our approach, a local variance is used with a window size of 4. The window moves on each row of the image, horizontally across each column, and the variance of each image pixel is computed. The second row of Figure 5 shows the heterogeneous regions. Equation (3) shows the calculation of a pixel variance, where A is a vector made up of N scalar observations. In this paper, the value of N is considered as 4.

$$V = \frac{1}{N-1} \cdot \sum_{i=1}^N |A_i - \mu|^2 \tag{3}$$

$$\mu = \frac{1}{N} \cdot \sum_{i=1}^N A_i \tag{4}$$

Also, μ is the mean in the vector, shown in Equation (4). The dimensions of this vector is 1×36 . A number of variance images are shown in Figure 5.

3. 5. Histogram Descriptor The HOG feature descriptor is an efficient and popular method for object detection. The main idea behind the algorithm is to identify the local object shape and appearance using the distribution of local intensity gradients or edge directions [18]. This descriptor works even without knowing the edge’s precise position. The steps for implementation of the HOG descriptor are as follows:

- (1) Dividing the image into small regions (cells). In the paper, cell size is 8×8 pixels and the cell histogram is computed using nine bins.

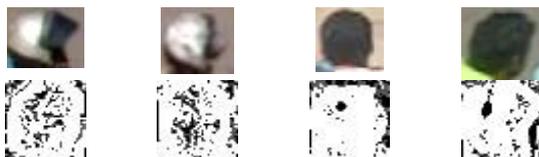


Figure 5. A number of images (top row) with their variance (bottom row)

- (2) Calculating the gradient in x and y directions for each pixel within the cell, and putting them into a x -bin histogram. Value of x depends on the gradient orientation. For the unsigned gradients, this value is 9, else it is 18. Because larger values of x cannot get the important information, we used the unsigned gradients [19]. The range of orientation is from 0 to 180 degrees. The mask array used for computing the gradient is $[-1, 0, +1]$.
- (3) For invariance to illumination and shadows, a group of adjacent cells is considered as a block, then all the cells in the block are normalized (as shown in Equation (5) and Equation (6)). We used blocks with 2×2 cells, where each of them contains 16×16 pixels; for more robust results, blocks have a 50% overlap. In total, each block contains a histogram with 36 bins (4×9). See Figure. 6.

3. 6. Image Classification

Our dataset consists of bikers wearing and not-wearing a helmet, so we use binary classifiers to classify the images. The classifier’s input is the extracted feature vector from the hybrid descriptor, which consists of LBP, HOG and LV. The linear SVM classifier is used for the classification. SVM is a supervised learning model which is used in classification and regression. Data is divided into two categories: train and test. The technique produces a model based on training data and their labels, and uses this model to predict test data labels. If the data are linearly separated, SVM uses a linear hyper plane, otherwise, using a kernel function, SVM can separate the data using a non-linear hyper plane. The data is transferred to a higher-dimensional space; in which data can be linearly separable in the new space [20].

$$H_{B1} = \left[\frac{HC1}{NB1}, \frac{HC2}{NB1}, \frac{HC5}{NB1}, \frac{HC6}{NB1} \right] \tag{5}$$

$$FV = [H_{B1,2}, H_{B3}, \dots] \tag{6}$$

4. EXPERIMENTAL RESULTS

By applying the proposed method on the introduced data, we evaluated the effectiveness of this method. The cross

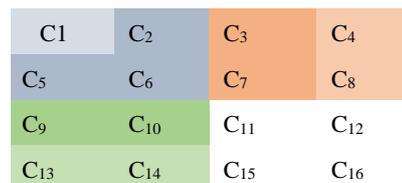


Figure 6. Example of cells and blocks, C_1 is a cell and $[C_1, C_2, C_5, C_6]$ are a block

validation technique has been used to separate train and test samples. In this technique, the original samples are divided into K equal parts, the model learns from the training samples ($K-1$ parts), and is evaluated using the test samples (one part). The process of model learning and testing is repeated K -times, so that each time the $K-1$ parts are used for training and one part is used for the test [21]. Each sample is used once for testing which makes the results reliable. Eventually, the average of this K -iterations is reported as the final result. We used the basic state of this technique with the value of K equal to 10. This method is evaluated using SVM classifier. We evaluated SVM with different kernels (linear kernel, tangent sigmoid, polynomial and radial basis function) and the linear kernel gave the best result.

In this paper, the measures used for evaluation are: True Positive Rate or Sensitivity (S), Specificity (SP) or true negative rate, Positive Predictive Value (PPV) or Precision, Negative predictive value (NPV) and Accuracy (A). These measures are described below.

$$\text{Sensitivity(S)} = \text{TP} / (\text{TP} + \text{FN}) \quad (7)$$

$$\text{Specificity (SP)} = \text{TN} / (\text{TN} + \text{FP}) \quad (8)$$

$$\text{PPV} = \text{TP} / (\text{TP} + \text{FP}) \quad (9)$$

$$\text{NPV} = \text{TN} / (\text{TN} + \text{FN}) \quad (10)$$

$$\text{Accuracy(A)} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FN} + \text{TN}) \quad (11)$$

$$\text{F-measure (F)} = 2 \times (\text{S} \times \text{PPV} / (\text{S} + \text{PPV})) \quad (12)$$

By using the linear SVM classifier, 99 out of 103 images were correctly classified as bikers not wearing a helmet (True Positive -TP) and 151 out of 152 images were correctly classified as bikers wearing a helmet (True Negative - TN).

TABLE 1. Confusion matrix for SVM classifier

Method	S (%)	SP (%)	PPV (%)	NPV (%)	F-measure	A
The method introduced by Silva et al. [7]	94.00	-	91.61	91.00	92.81	91.37
The proposed method	96.11	99.34	99.00	97.41	97.53	98.03

TABLE 2. Images classification results

	Predicted	
	Positive	Negative
Actual	Positive	99
	Negative	4
		151

Hence, the values of TN, TP, FP and FN were obtained as 151, 99, 1 and 4, respectively. Confusion matrix for the classifier is shown in Table 1. The best result was obtained using the SVM classifier with the accuracy rate of 98.03.

We have also compared the performance of our method with the method proposed by Silvi *et al.* [7], and the results are shown in Table 2. The accuracy of the proposed method was improved more than 6% on the dataset by using the SVM classifier. To evaluate the computational complexity of the two approaches, both methods were implemented in MATLAB on a computer with an Intel(R) core(TM) i3 processor operating at 2.40 GHz clock frequency, and 6.00 GB of RAM. The execution time is 2.32 seconds and 8.37 seconds respectively for the proposed method and the method introduced by Silvi *et al.* [7].

5. CONCLUSION

In this paper, an approach was proposed for detecting bikers not wearing a helmet. The proposed method consists of three steps: preprocessing, feature extraction and classification. Novelty of the proposed method is mainly on the feature extraction step, which leads to the classification method appropriately discriminating the two classes of helmet and non-helmet. We used a hybrid descriptor which contained HOG, LBP and LV feature extractors. Experimental results have shown that the proposed method improved helmet detection in bikers both in terms of accuracy and computational complexity.

6. REFERENCES

1. Waranusas. R, Bundon. N, Tintong. V, Tangnoi. C, and Pattanathaburt. P, "Machine vision techniques for motorcycle safety helmet detection", 28th International Conference on Image and Vision Computing New Zealand (IVCNZ), (2013), 35-40. doi: 10.1109/IVCNZ.2013.6726989
2. Doungmala. P, and Klubsuwan. K, " Helmet wearing detection in Thailand using Haar like feature and circle hough transform on image processing", International Conference on Computer and Information Technology (CIT), (2011), 611-614. doi:10.1109/CIT.2016.87
3. Dahiya. K, Singh. D, and Mohan. C. K, "Automatic detection of bike-riders without helmet using surveillance videos in real-time", International Joint Conference on Neural Networks (IJCNN), 3046-3051, (2016), doi: 0.1109/IJCNN.2016.7727586
4. Silva. R, Aires. K, Santos.T, Abdala. K, Veras. R, and Soares. A," A. Automatic detection of motorcyclists without helmet", Latin American Computing Conference (CLEI), 1-7, (2013), doi: 10.1109/CLEI.2013.6670613
5. Silva. R, Aires. K, Veras. R, Santos. T, Lima. K, and Soares. A, "Automatic motorcycle detection on public roads", *CLEI Electronic Journal*, Vol. 16, No. 3, (2013), doi: 10.1109/clei.2012.6427165
6. Marayatr. T, and Kumhom.P, "Motorcyclist's Helmet Wearing Detection Using Image Processing", In *Advanced Materials*

- Research**, Vol. 931, (2014), 588-592, doi:10.4028/www.scientific.net/AMR.931-932.588
7. Silva. R. V, Aires. K. R. T, and Veras. R. D.M, "Helmet detection on motorcyclists using image descriptors and classifiers", 27th Conference on Graphics, Patterns and Images, (2014), 141-148. doi:10.1109/sibgrapi.2014.28
 8. Shine. L, and Jiji. C.V," Automated detection of helmet on motorcyclists from traffic surveillance videos: a comparative analysis using hand-crafted features and CNN", *Journal of Multimedia Tools and Applications*, Vol. 1380, (2020), 1-21, doi: 10.1007/s11042-020-08627-w
 9. Alaulikar. A.S, Sanathanan. S, and Modi. C.N," An enhanced approach for detecting helmet on motorcyclists using image processing and machine learning techniques", In *Advanced Computing and Communication Technologies*, Vol. 702, (2019), 109-119, doi: 10.1007/978-981-13-0680-8_11
 10. Rubaiyat. A.H, Toma. T.T, Kalantari-Khandani. M, Rahman. S.A, Chen. L, Ye. Y, and Pan. C.S., "Automatic detection of helmet uses for construction safety", WIC/ACM International Conference on Web Intelligence Workshops (WIW), (2016), 135-142. doi: 10.1109/WIW.2016.045
 11. JIANG. X, XUE. H, ZHANG. L, and ZHOU. Y," A Study of Low-resolution Safety Helmet Image Recognition Combining Statistical Features with Artificial Neural Network", *International Journal of Simulation-Systems, Science and Technology*, Vol. 17, No. 37, (2016), 1473-8031, doi: 10.5013/IJSSST.a.17.37.11
 12. Li. K, Zhao. X, Bian. J, and Tan. M," Automatic Safety Helmet Wearing Detection", arXiv labs:experimental projects with community collaborators, 1802.00264, (2014).
 13. Raj. K. D, Chairat. A, Timtong. V, Dailey. M. N, and Ekpanyapong. M, "Helmet violation processing using deep learning", International Workshop on Advanced Image Technology (IWAIT), (2018), 1-4, doi: 10.1109/IWAIT.2018.8369734
 14. Siebert. F.W, and Lin. H," Detecting motorcycle helmet use with deep learning", *Journal of Accident Analysis and Prevention*, Vol. 134, (2020), doi:10.1016/j.aap.2019.105319
 15. Shakour.M. H and Tajeripour. F," Local entropy pattern to extract textural images features", Vol. 3, No. 2, (2016), 73-85.
 16. Lenc. L, and Král. P," Automatically detected feature positions for LBP based face recognition", In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, (2014), 246-255, doi: 10.1007/978-3-662-44654-6_24
 17. Huang.D, Shan. C, Ardabilian. M, Wang. Y, and Chen. L, "Local binary patterns and its application to facial image analysis: a survey", *Journal of Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 41, No. 6, (2011), 765-781, doi: 10.1109/TSMCC.2011.2118750.
 18. Dalal, N, and Triggs. B," Histograms of oriented gradients for human detection", Computer Society Conference on Computer Vision and Pattern Recognition, (2005), 886-893.
 19. Başa. B," Implementation of Hog Edge Detection Algorithm Onfpga's", *Procedia-Social and Behavioral Sciences*, Vol. 174, (2015), 1567-1575, doi: 10.1016/j.sbspro.2015.01.806
 20. Boser. B. E, Guyon. I. M, and Vapnik. V. N, "A training algorithm for optimal margin classifiers", In Proceedings of the fifth annual workshop on Computational learning theory, (1992), 144-152. doi: 10.1145/130385.130401
 21. Ellis. R. P, and Mookim. P. G, "Cross-validation methods for risk adjustment models", Mimeo, Boston University, (2008).

Persian Abstract

چکیده

کلاه ایمنی برای جلوگیری از آسیب های ناحیه سر موتورسواران ضروری است. در اکثر کشورها برای موتورسوارانی که از کلاه ایمنی استفاده نمی کنند، قوانینی وضع شده است. بررسی موتورسواران متخلف به صورت دستی کاری وقت گیر و پرهزینه است. در این راستا، در طی سال های اخیر چندین روش تشخیص موتورسواران بدون کلاه ایمنی ارائه شده است. در این مقاله یک روش بر حسب پردازش تصویر برای شناسایی موتورسواران بدون کلاه ایمنی پیشنهاد شده است. این روش از توصیفگرهای الگوی دودویی محلی، واریانس محلی و هیستوگرام مبتنی بر گرادینان استفاده می کند. نوآوری این روش اساسا در مرحله استخراج ویژگی است که باعث می شود طبقه بندی دو کلاس با و بدون کلاه ایمنی بهتر انجام شود. نتایج تجربی نشان می دهد که روش پیشنهادی ما نسبت به روش های موجود دیگر، دارای برتری در تشخیص و طبقه بندی است. دقت روش پیشنهادی برای تشخیص کلاه ایمنی با استفاده از دسته بند ماشینی بردار پشتیبان 98.03 درصد است.
