A Study of New Features for Motorcycle Detection in Nighttime

Anonymous ICCV submission

Paper ID : 2429

Abstract

Traffic accidents mostly occur at night. It is understandable since at night, we have low visibility. Effort to reduce accidents at night have been reported by developing tools to detect nearby vehicles to avoid crashes. However, most of them worked only on detecting cars. They did not focus on motorcycle due to its lack of a pair of lamps lighting features usually utilized in detecting cars. Meanwhile, motorcycle consists of a rider, a taillight, and area surrounding license plate in the nighttime traffic images. To get these properties, we propose four features (red pixel, edge pixel, edge ROI, and active contour) that are extracted from red and edge maps. The red map is used for recognizing the spreading out of taillight on the image. The edge map is used to recognize the rider, back part of the motorcycle, and the whole curve of object. To see the effectiveness of our features, we selected 3 commonly used classifiers (Artificial Neural Network, Decision Tree Algorithm, and Support Vector Machine) in the experiment. The result shows that some classifiers have achieved more than 80% accuracy rate.

1. Introduction

The autonomous driving systems (ADS) are sometimes found in our daily traffic. The driver assistance systems (DAS) are being used by many vehicle industries to provide more convenience to driver. These systems require high accuracy and speed on detection of surrounding objects during day- and nighttime.

Nighttime object recognition and detection are difficult challenges in computer vision due to their low contrast, low brightness, and low signal to noise [7]. Instead of daytime environment which can provide many attributes, it has several obstacles to find key properties from image to serve for object recognition, classification, detection, etc.

There are some methods worked on this kind of problem including vehicle and motorcycle. The blind spot warning system [5], tries to detect car and motorcycle during nighttime by using head light detection. They use binary thresholding to find all possible lights and detect its contour. In grouping process, the paired light are labeled as car if they have the same vertical position and match with 3D size and distance. The rests are checked to remove reflected and close related lights. The remained lights are computed by 3D size and distance, and labeled as motorcycles.

Hulin Kuang et al [8], propose method to provide a set of proposals from nighttime images. They introduce the Bayes Saliency-based to generate the object proposal of vehicle. There are three features extracted from image such as local contrast, taillight, and luminance to serve for saliency detection. The Edge detection through prior estimation is used to separate the background and salient objects. The prior estimation also determines threshold for feature maps and Bayes rule.

The Harr-like based classifier is used to detect the vehicle candidates from both dark and bright environment by using light pair identification to check whether they are belonged to vehicle or not [6]. So, it can help to produce the region of interest(ROI) for alert signal analysis. The HSV (hue, saturation, and value) color system thresholds to discover vehicle lights for ROI. The luminance analysis looks for additional area of light scattering which is became the frequency by Discrete Cosine Transform (DCT). The magnitude is trained by Adaboost to detect the turn signal.

The taillights are detected by adapting the threshold directly from HSV color system and paired by cross-correlation symmetry analysis [10]. It is also tracked by Kalman filtering to be applicable with the distortions that are affected by other lights. On the other hand, their HSV threshold value will be efficient only whenever the environment is darker, the taillight is exactly clear to human-eyes.

The headlights are detected by AdaBoost classifier after training to reduce the false reflection and grouped by using maximal independent set framework [14]. They are considered as the hint to identify the object of vehicle. However, The video frames are coming from the stable camera such as CCTV and located at the opposite traffic direction. This means, the ROI is sensitively limited at the specific area.

All these existing researches for nighttime vehicle

ISSN : 2355-9365 e-Proceeding of Engineering : Vol.6, No.2 Agustus 2019 | Page 10014

detections are trying to pair either taillight or headlight as the vital attribute to identify and recognize whether it is the desired object or not. They depend on it to be the vital feature and sufficient to differentiate car from motorcycle, road lamp, traffic light, traffic sign, and other noises. However, most of object detection works in traffic images ignore the motorcycle and concentrate on car detection.

Some researchers are not interested in motorcycle because of several reasons. They predicted that this kind of transportation will vanish from our daily traffic in the future. Another reason, like [8], they focuses on car because it is the most popular private transportation in Taiwan.

However, motorcycle is a common mode of transport in Southeast Asian countries. It is considered as the most dangerous method of transportation due to its characteristics such as not stable, small, fast moving, and easily broken [5]. Traffic accidents involving motorcycles contribute to the highest death category with 36\% in Indonesia [11]. Most accidents happened at night and impacted on road user by individuals, communities, and countries.

Therefore, nighttime motorcycle detection is very necessary in computer vision for intelligent transport systems (ITS), since nighttime traffic accidents mostly involve motorcycles.

Since now, there is no specific research to develop method for detecting motorcycle during nighttime yet. The major problem is feature. It comes from blur image of nighttime environment and small size of motorcycle.

To solve this problem, we propose a set of new features to be used for motorcycle detection in nighttime images. As seen by back rider or camera, motorcycle is small compared with other vehicles, has one taillight, and one headlight. It is easily confused with many kinds of noises such road lamp, banner, and traffic light (see Figure 1). It is different with car which has two taillights and large size. The taillight is not only important for human as a driver but also systems to recognize the back-view of vehicles. When the motorcycle is ridden on the road, its taillight will be a key hint for another drivers. Furthermore, the rider's back and the lower part of the motorcycle can also be additional clues to help identify the motorcycle.

To obtain those attributes from traffic image, red area is noticeable to infer its intensity to taillight. Either below or above taillight, there are some reflections which become edges. All these edges are exploited to figure out pertinent area of actual object including rider, taillight, and other parts of motorcycle by active contour process.

Our contributions in this paper are described as follows. The red area and edges are exploited to become vital features from nighttime image. Generally, red and edge pixels are counted to be real numbers. In addition, we take the advantages of potential light as known as taillight for motorcycle. The light is used to help determine ROI and active contour. We propose twenty active points to initiate the active contour of motorcycle. All edge pixels within ROI are specially counted. The final contour area is calculated. All numbers are averaged by image size. Eventually, all features (red pixel, edge pixel, edge ROI, active contour) are in range of [0, 1].

The structure of the paper is organized as follows. The detail of our proposed features are introduced in section 2. The experiment is described in section 3. Finally, section 4 illustrates the conclusion and future work.

Figure 1: A sample image of motorcycle during nighttime.

2. Feature extraction

Dark or blur images gives more problem to differentiate between background and foreground, between object and noise. It also poses low intensity gradient. Thus, the possible information to be used as features from nighttime image are less than daytime. Suppose, there is a bounding box known as ground truth which is cropped from RGB image with the position of potential light as in Figure 3a and 4a. It is considered as raw input data to be used for training and testing in any machine learning algorithm to classify whether it is motorcycle or not. It has to be transformed to a way that machine can learn and predict. We propose the features shown in Figure 2 to perceive some attributes from box image to be used by learning algorithm.

The features that are described in section 2.1, 2.2, 2.3 demand to normalize the amount of pixels in the box image meeting the criteria postulated by ***f(x,y)***.

$$Feature(Image) = (\sum_{x=0}^{W-1}\sum_{y=0}^{H-1} f(x,y))/(W \times H) \quad (1)$$

where ***W*** and ***H*** are denoted as width and height of image; ***f(x,y)*** represents different criteria.

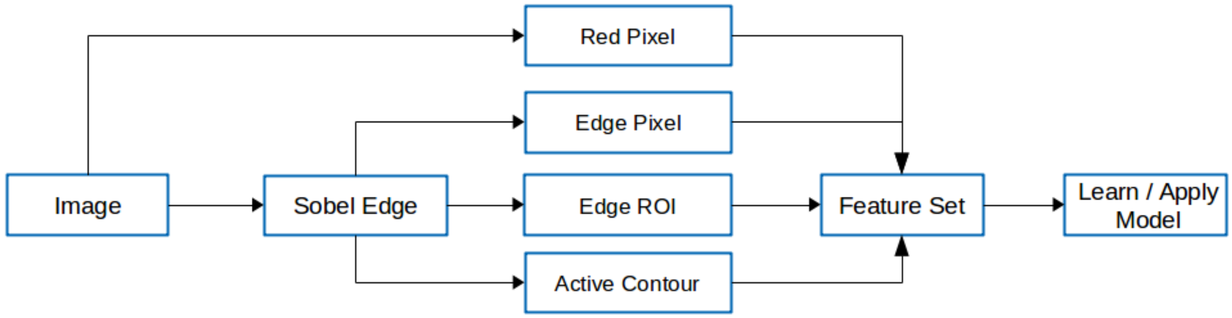


Figure 2: A proposed diagram of feature extraction

2.1. Red pixel

From observation, the images look like more red color if the motorcycles exist. Some parts of motorcycles get red reflection from its taillight or another vehicles on the road. That's why, the ground truth images that contain the motorcycle have red pixels more than noise images. Digital camera commonly produces three image channels with red, green, and blue (RGB) color system. The pixel belongs to red color whenever the maximum value is coming from red channel compare with green and blue channels. Therefore, the average ratio is to count every red pixels in image by using Equation 1. The procedure of identifying red pixel is decided by Equation 2.

$$f(x,y) = \begin{cases} 1, & \text{if } Red(x,y) > Blue(x,y) \wedge \\ & \quad Red(x,y) > Green(x,y) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where ***f(x,y)*** is the function to check the pixel at the ***x*** and ***y*** axes in the input RGB image whether it belongs to red pixel or not. The ***Red(x,y), Blue(x,y),*** and ***Green(x,y)*** are the color mapping functions to get its color value at the position ***x*** and ***y*** coordinates. Under summation of ***f(x,y)*** and averaging by Equation 1, the output value of this feature is in the range of [0, 1] from the input image to become one of feature set.

2.2. Edge pixel

Nighttime image provokes edge to become unclear because of low contrast and brightness. However, edge is still marked as notable pixels in nighttime object [8]. There is only one taillight on motorcycle. It can not be detected by using paired lamps. The taillight can be confused with road lamp but the difference is number of edges. Its bottom side consists of license plate, tires, turn lamps, and other parts of motorcycle. Its top side consists of rider's body and helmet. All these sides provide a lot of edges meanwhile the road lamp or noise does not have. So, the actual object contains more edges than non-object candidate.

First, we use Sobel Operator [13] to compute edge map of input image. This map is thresholded to be binary map E_{binary}. After that, we calculate the average ratio by using equation (1) respectively. The criteria of counting edge is formulated as follow.

$$f(x,y) = \begin{cases} 1, & \text{if } E_{binary}(x,y) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where ***f(x,y)*** is denoted as function to check edge at position ***x*** and ***y*** axes. $\mathbf{E_{binary}(x,y)}$ is a binary pixel value at ***x*** and ***y*** axes from edge map. Zero value is noted as background, otherwise edge pixel. The binary output of function ***f(x,y)*** is executed by general equation (1) to produce feature in range of [0, 1]. This value becomes the member of feature set in Figure 2.

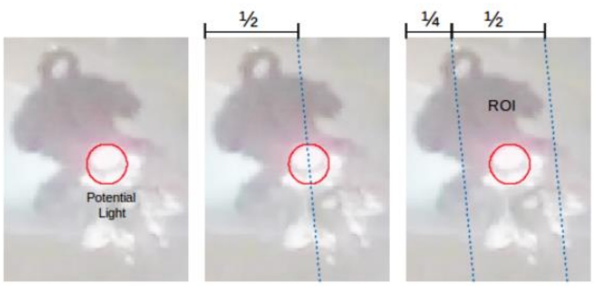
2.3. Edge in Region of interest (ROI)

To detect the back-view of motorcycle, we need some rotations to adjust the position of ROI. Because the motorcycle is not always ridden in the middle road. Sometimes, it is located at a bit left or right hand of camera. To focus on the object including rider and bottom part, the vertical line segment is proposed to rotate ROI. This line segment is generated by center of taillight and mid width of ground truth in Figure 3b and 4b. It is flexible to follow the potential light known as taillight for motorcycle. The adaptive orientation innovates ROI to correct position of object. In Figure 3c and 4c, the vertical line is used to enlarge 1/4 of width to left and right to get the actual ROI. This means that width of ROI takes almost half area of ground truth. It is coming from most images in our dataset.

ROI covers all pixels within distance 25% of width. It is highly suspected to contain actual object include rider and backside of motorcycle. So, all edges in ROI are marked to be counted by equation (1). The conditional pixel is shown by equation below.

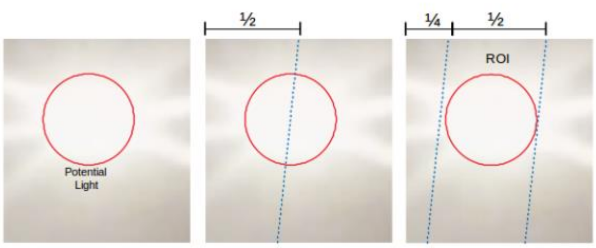
$$f(x,y) = \begin{cases} 1, & \text{if } E_{binary}(x,y) > 0, (x,y) \in ROI \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where ***f(x,y)*** is the function to check edge at position ***x*** and ***y*** axes respecting to its criteria. ***E_{binary}(x,y)*** is a pixel value at position ***x*** and ***y*** in edge map. It is considered as edge pixel whenever its value is greater than zero. The function ***f(x,y)*** will return number one if the edge pixel is within ROI. The outputs of this function are computed by equation (1) to get the feature value in range of [0, 1]. Therefore, the input image is transformed to real number for feature set in Figure 2.



(a) Ground Truth (b) Vertical line (c) ROI

Figure 3: The sample motorcycle of ROI



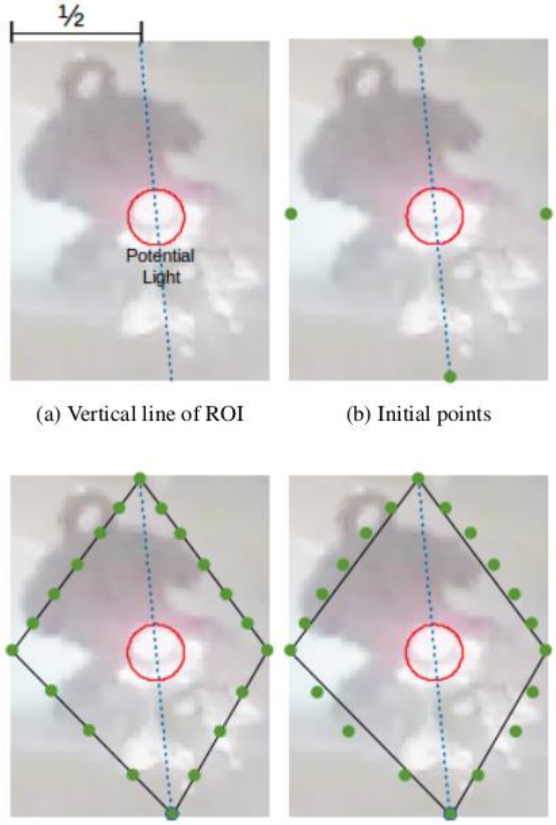
(a) Ground Truth (b) Vertical line (c) ROI

Figure 4: The sample noise of ROI

2.4. Active contour area

Nighttime environment seems like impossible to get the actual shape of object to fit any specific model. However, we find the whole shape of object including taillight, rider, and backside of motorcycle by using active contour.

In our case, active contour is the collection of active points moving to vertical line pixel by pixel when it does not meet the obstacle as known as edge. The more number of points, the more computation cost. The contour is closed and moved by each point to discover the actual curve of object. There are more edges within image which contain the motorcycle. So, the active contour stays around the large area. Its points also take short step to move. However, the final contour will be small or remained only the potential light if the ground truth is not the motorcycle. It might take more cost to travel each points from initial stage until breaking iteration. The processes are summarized as follows.



(a) Vertical line of ROI (b) Initial points

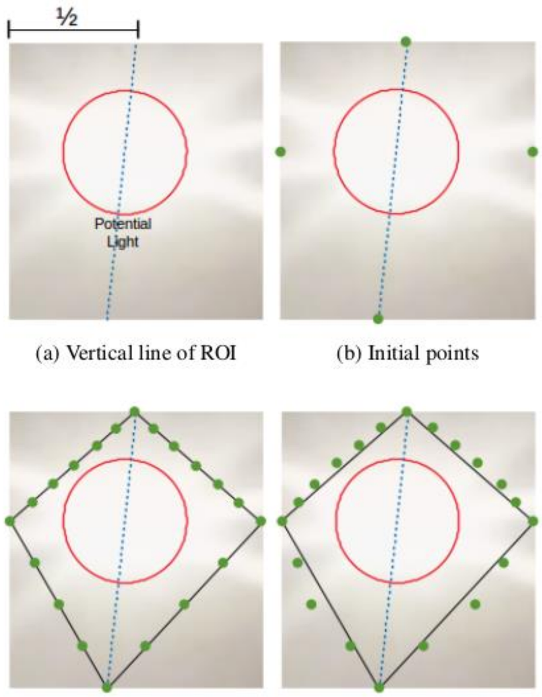
(c) Active points (d) Weighted points

Figure 5: The sample motorcycle of active contour

The line segment is created from the central point of potential light with the mid width of ground truth image. It is almost vertical within ROI as shown in Figure 5a. Thus, there are four points of polygon generated. Two are came from vertical line segment. Another two is from horizontal light's position in Figure 5b.

Between top to left and right points, we generate five points for each equally. From bottom to left and right points, we also create three points for each individual. These optimal numbers are determined by our observation after performing experiment on our dataset with different criterion accordingly.

First, we need to create another ten points for top side basing on X and Y variances. The Y variance is calculated by distance between top point to either left or right point with Y direction and then divide by six. Because we need six spaces for five points equally. For X variance, the image width is divided by twelve, because the top point is the result of dividing the image width by two and required to six spaces more for another individual points. So, we need to divide the image width by twelve (2 x 6). Those variances *X* and *Y* are used to create active points on the top side of potential light which is known as taillight in case of motorcycle.



(a) Vertical line of ROI (b) Initial points

(c) Active points (d) Weighted points

Figure 6: The sample noise of active contour

Second, there are another six points required for bottom side. We calculate Y variance vertically from either left or right to bottom point and divide by four. For X variances are different, because the bottom point is not stable or balance between left and right. So, we need to define X variance for left and right individually. Afterward, we use those variances to determine active points linearly for bottom side. (See Figure 5c)

Third, we need to adjust weight through X direction on some important points to fit with ellipse shape. For five points at top right, we add an X variance in positive direction to third point, half for second and fourth, and 1/3 of X variance to first and fifth points. The points at top left are undergone the same to top right in the negative direction. For three points at bottom right side, we add further by its X variance for second point, and half for first and second points. This procedure is also applied for bottom left side but opposite direction. (See Figure 6d)

Fourth, those active points are iteratively move to vertical line of ROI horizontally except top and bottom points that are travel indent following to line orientation. The point will stop moving if it meed the any edge or vertical line, otherwise continue. After movement finished, we apply Douglas Peucker algorithm [4] to approximate the contour and compute its area using Green's theorem[1]. Thus, the feature is produced in range [0, 1] by the result of dividing the area of contour and size ground truth.

3. Experiment

In classification's job, machine learns from feature set to build its own model. To see the result how our proposed features perform in motorcycle classification during nighttime image, the experiments are conducted by various environment and large enough size of sample. Those features are extracted to become vectors for training and testing with some classification tools such as Artificial Neural Network (ANN) [12], Decision Tree (DTree) [1], and Support Vector Machine (SVM) [2]. The features are performed with Principle Component Analysis (PCA)(Jolliffe, 1986)[2] which is a dimensional reduction technique. This linear algebra based, PCA occurs to find a good pattern and significant characteristic of data in high dimension [9]. Thus, the learning tool can be improved it efficiency.

3.1. Experimental setup

By focusing on features for classification, the dataset has 1111 images(1280 x 720) extracted from frame of 11 videos during nighttime on the road in Indonesia. Among those, there are 3337 samples consist of its position with potential lamp and two class labels including motorcycle and non-motorcycle. The dataset also is separated by 80-20 into two various subsets are training and testing sets.

PCA is used to help the classifier to improve its performance. It clarifies and manipulates the most important feature among all vectors. We suppose our training set as 2D matrix. Each column represents the feature vector and becomes dimension in PCA space. The eigenvector and eigenvalue are the outputs of PCA. The eigenvalue tells us the variance level of eigenvector; the higher the better. The eigenvector gives the platform to transform original feature to another pattern by multiplying the matrix transpose of eigenvector with training or testing set. Thus, we have a new feature set.

All configurations for learning tools are basic and simple. ANN has 1 input layer with 4 neurons, 2 hidden layers with 16 and 8 neurons, and 1 output layer with two neurons (1 = motorcycle, 0 = non-motorcycle). DTree has no limitation for branch number. Its maximum depth is

[1] S. Steward, James. Calculus (6th ed). Thompson, Brooks/Cole.

[2] Jolliffe, I. T. (1986) Principal Component Analysis. New York: Springer.

3000, and maximum category is 10. SVM is trained with Linear kernel.

The experiments are performed on Ubuntu 16 LTS, OpenCV compiler, and C++ programing language.

3.2. Experimental scenario

There are four features proposed to do experiment in different cases to see the efficiency. The test stories are designed to follow particular and combination. The features require individually four cases of particular scenarios such as red pixel, edge pixel, edge ROI, and active contour. Those cases are randomly grouping together to turn another cases. There are two, three, and four features of combination. Among those, PCA is planed to implement on three and four combinations. The features are not repeated in one scenario. The possible number of combinational cases can be counted as follows.

- Particular: 4 cases
- Two combination: 6 cases
- Three combination: 4 cases
- Four combination: 1 case
- PCA with three & four combinations : 5 cases

The total cases is twenty. In each case, the experiment is conducted on three classifications. Therefore, the number of experimental scenarios is times three of total cases. It becomes sixty scenarios.

3.3. Experimental result

There are sixty experimental scenarios shown in table 1 on our proposed features for particular testing and combination. Among those, tens are above 80% accuracy classified by learning tools. Decision Tree contributes two results, and eights are from ANN. The highest accuracy is coming from the combination of three features including red pixel, edge pixel, and edge ROI. However, most of the rests are above 50% rate in this table.

Furthermore, the combinations of three and four features are performed with PCA. The table 2 is the result of applying PCA on those combinations of our proposed features that are transformed to another feature set.There are 4 results above 80% after doing 15 experimental tests. Those are classified by ANN at all. According to both of these table, PCA improves a little bit percentage for only some results of below 80% and remains the highest result of table 1 as well.

Therefore, the results of table 1 & 2 on accuracy of classification can claim that our proposed features are useful to work for machine learning in motorcycle detection during nighttime. Although, the experimental result provides different rate of accuracy following to individual learning algorithms, we do not need much concentration to compare the performance of machine learnings. We focus on feature set. From observation on all results, Edge ROI is the best feature in the experiment. It is always illustrated in every high accuracy. The less efficient feature is active contour which usually decreases the result.

Majority of above 80% accuracy are under combination of three and four features. This is acceptable results respecting to the amount image of dataset and nighttime environment.

Referring to all classifications' methods in both table 1 & 2, most of high result is provided by ANN, decision tree has two, and SVM does not show any result above 80%. The lowest accuracy rate of ANN is 79.11% in the table 2 while Decision Tree is having maximum 79.57% and 61.73% for SVM.

The experimental result shows that our features donate good results for ANN and Decision Tree, not SVM. In nonlinear revolution, Neural Network and Decision Tree were efficient learning for non-linear decision surface in 1980's [3]. The kernel based method like SVM does not present well for our features even if we have tried with another kernels such as polynomial, Radial basis, Sigmoid, exponential Chi2, and Histogram intersection kernels.

According to all high results of accuracy from those tables, the combination of three features (red pixel,edge pixel, edge ROI) has the best accuracy under classification of artificial neural network. All these things are expected to integrate with other methods in order to work for motorcycle detection.

4. Conclusion and future work

This paper indicates four features for classification on motorcycle image during nighttime. The red color and edges are noticeable from images to provide efficient features. The experiments show that machine learning can reaches 83.23% which is high accuracy in classification. We noted, Edge ROI is the best feature in the particular scenario during experiment. It also contributes the most efficient attribute in combination scenarios to distinct motorcycle and noise. Based on the proposed method, we can claim that it is possible to detect object of motorcycle during nighttime.

In the future, we plan to develop our method further to enhance the nighttime motorcycle detection. There are a lot of noises found in the traffic images. Road detection will be used to discard all those noises. It determines the road boundary and limits area of image for identifying motorcycle.

Table 1: Performance accuracy of machine learning on our proposed features.

Red pixel	Edge pixel	Edge ROI	Active contour	ANN (%)	DTree (%)	SVM (%)
✓				64.48	53.20	62.04
	✓			69.51	47.56	37.50
		✓		**82.62**	47.86	60.21
			✓	59.14	34.29	42.83
✓	✓			72.56	67.07	69.51
✓		✓		**83.07**	77.89	62.04
✓			✓	63.41	64.32	57.77
	✓	✓		**82.77**	75.15	66.31
	✓		✓	73.93	59.75	49.54
		✓	✓	**82.46**	73.17	42.83
✓	✓	✓		**83.23**	78.96	61.58
✓	✓		✓	78.50	72.56	50.60
✓		✓	✓	**83.07**	**80.33**	57.77
	✓	✓	✓	**83.07**	78.35	49.54
✓	✓	✓	✓	**82.92**	**82.46**	50.45

Table 2: Performance accuracy of machine learning on our proposed features with PCA.

Red pixel	Edge pixel	Edge ROI	Active contour	PCA	ANN (%)	DTree (%)	SVM (%)
✓	✓	✓		✓	**83.23**	78.50	61.73
✓	✓		✓	✓	79.11	69.81	50.60
✓		✓	✓	✓	**82.77**	78.20	57.62
	✓	✓	✓	✓	**83.07**	78.04	49.54
✓	✓	✓	✓	✓	**82.77**	79.57	50.45

References

[1] L. Breiman. Classification and regression trees. Routledge, 2017. 5

[2] C.-C. Chang. ” libsvm: a library for support vector machines,” acm transactions on intelligent systems and technology, 2: 27: 1–27: 27, 2011. http://www. csie. ntu. edu. tw/˜cjlin/libsvm, 2, 2011. 5

[3] N. Cristianini and B. Scholkopf. Support vector machines and kernel methods: the new generation of learning machines. Ai Magazine, 23(3):31–31, 2002. 7

[4] D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Cartographica: the international journal for geographic information and geovisualization, 10(2):112–122, 1973. 5

[5] C. Fernández, D. Llorca, M. Sotelo, I. Daza, A. Hellín, andS. Álvarez. Real-time vision-based blind spot warning system: Experiments with motorcycles in daytime/nighttime conditions. International Journal of Automotive Technology, 14(1):113–122, 2013. 1, 2

[6] C.-L. Jen, Y.-L. Chen, and H.-Y. Hsiao. Robust detection and tracking of vehicle taillight signals using frequency domain feature based adaboost learning. In Consumer Electronics-Taiwan (ICCE-TW), 2017 IEEE International Conference on, pages 423–424. IEEE, 2017. 1, 2

[7] H. Kuang, L. Chen, F. Gu, J. Chen, L. Chan, and H. Yan. Combining region-of-interest extraction and image enhancement for nighttime vehicle detection. IEEE Intelligent Systems, 31(3):57–65, 2016. 1

[8] H. Kuang, K.-F. Yang, L. Chen, Y.-J. Li, L. L. H. Chan, and H. Yan. Bayes saliency-based object proposal generator for nighttime traffic images. IEEE Transactions on Intelligent Transportation Systems, 2017. 1, 2, 3

[9] M. Marengoni and D. Stringhini. High level computer vision using opencv. In 2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials, pages 11–24. IEEE, 2011. 5

[10] R. O'Malley, E. Jones, and M. Glavin. Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. IEEE Transactions on Intelligent Transportation Systems, 11(2):453–462, 2010. 1, 2

[11] S.-E. A. REGION. Road safety in the south-east asia region 2015. 2016. 1

[12] M. Riedmiller and H. Braun. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In Proceedings of the IEEE international conference on neural networks, volume 1993, pages 586–591. San Francisco, 1993. 5

[13] I. Sobel and G. Feldman. A 3x3 isotropic gradient operator for image processing. a talk at the Stanford Artificial Project in, pages 271–272, 1968. 3

[14] Q. Zou, H. Ling, S. Luo, Y. Huang, and M. Tian. Robust nighttime vehicle detection by tracking and grouping headlights. IEEE Transactions on Intelligent Transportation Systems, 16(5):2838–2849, 2015. 1, 2