Article

Performance Evaluation of Region-Based Convolutional Neural Networks Toward Improved Vehicle Taillight Detection

Zhenzhou Wang ¹, Wei Huo ¹, Pingping Yu ¹, Lin Qi ¹, Shanshan Geng ² and Ning Cao ^{3,4,*}¹ School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050000, China² School of Information Technology, Hebei University of Economics and Business, Shijiazhuang 050061, China³ School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China⁴ School of Internet of Things and Software Technology, Wuxi Vocational College of Science and Technology, Wuxi 214028, China

* Correspondence: ning.cao2008@hotmail.com; Tel.: +86-1860-326-3678

Received: 28 May 2019; Accepted: 4 September 2019; Published: 8 September 2019

Abstract: Increasingly serious traffic jams and traffic accidents pose threats to the social economy and human life. The lamp semantics of driving is a major way to transmit the driving behavior information between vehicles. The detection and recognition of the vehicle taillights can acquire and understand the taillight semantics, which is of great significance for realizing multi-vehicle behavior interaction and assists driving. It is a challenge to detect taillights and identify the taillight semantics on real traffic road during the day. The main research content of this paper is mainly to establish a neural network to detect vehicles and to complete recognition of the taillights of the preceding vehicle based on image processing. First, the outlines of the preceding vehicles are detected and extracted by using convolutional neural networks. Then, the taillight area in the Hue-Saturation-Value (HSV) color space are extracted and the taillight pairs are detected by correlations of histograms, color and positions. Then the taillight states are identified based on the histogram feature parameters of the taillight image. The detected taillight state of the preceding vehicle is prompted to the driver to reduce traffic accidents caused by the untimely judgement of the driving intention of the preceding vehicle. The experimental results show that this method can accurately identify taillight status during the daytime and can effectively reduce the occurrence of confused judgement caused by light interference.

Keywords: driving intention analysis; vehicle detection; taillight detection; taillight semantic recognition

1. Introduction

With the explosive growth of the number of vehicles, many social problems have become increasingly prominent, such as traffic accidents, traffic congestion and deterioration of the traffic environment. According to traffic accident data released by relevant departments, the proportion of rear-end collisions in traffic accidents reached 30–40% [1]. Frequent traffic accidents have caused huge losses to people's lives and property. Recently, many experts and scholars in the field of active vehicle safety have studied the related technologies of detecting and identifying abnormal driving behavior of the preceding vehicles. An Intelligent Transportation System (ITS) integrates control system, information system, communication system and computer network technology. It has the characteristics of real-time, accurate and wide-scale. This system can provide a method to ease urban traffic congestion and improve driver safety. Meanwhile, it can reduce emission pollution and it is

important to promote urban economic development. ITS technological development includes two aspects: vehicle and vision [2,3].

As an important part of the ITS, an auxiliary driving system uses machine vision technology to collect an image of the traffic environment and information pertaining to the preceding vehicle. It can also be used to deal with noise pollution and identify the region of interest (ROI). Therefore, in an emergency, it can activate an alarm quickly and automatically [4]. In the realm of complex and diverse traffic accidents, traffic accidents caused by rear-end vehicle collisions account for a large proportion. If the driver has access to road information before the rear-end accident occurs, they can interpret the driving intention of the preceding vehicle and respond rapidly to prevent a rear-end accident [5,6].

The traditional automobile intelligent collision avoidance system ignores the comprehensive influence of the driver, the vehicle condition and the road traffic environment on the driving safety state. Therefore, it cannot adapt to the complicated and varied road traffic environment. In recent years, information fusion and situational awareness theory has been introduced into the field of traffic safety to monitor the safety of traffic in complex road traffic environments. The application of Internet of Things (IoT) technology in the automotive and transportation sectors forms an independent network system. It mainly includes sensor technology and information integration, open intelligent vehicle terminal and voice recognition technology. Service terminal computing technology and communication technology are also included in it.

The goals of the study are to analyze and predict the driving intentions of the preceding vehicle, with the focus on taillight semantic recognition. This system uses image processing-related methods to achieve image de-noising, filtering and grayscale conversion. It converts images in Red-Green-Blue (RGB) to other color spaces for further research and analysis. Convolutional neural network (CNN) is used to realize the feature extraction of vehicle. After the vehicle detection is completed, the system implements taillight detection and light recognition based on dynamic thresholds and image histogram features in the color space.

The system uses the Faster Region CNN (RCNN) to identify vehicles on the road. The Faster RCNN can use the training set images to achieve faster training and higher detection and identification rates. CNN is constructed and trained by manually marking the ROI on the image dataset. The test set is then used to test the performance of CNN, and the results show that the algorithm can achieve high precision. The system converts images from RGB color space to Hue-Saturation-Value (HSV) color space. This achieves a more precise segmentation of taillights and solves the difficulty of detecting red taillights in RGB space. It extracts taillight pairs by using features of a histogram of oriented gradients, position and color correlation, and realizes taillight state detection using the histogram features of the taillight area image. Thus, it can further improve the accuracy of taillight detection. Compared with the conventional method of extracting taillights directly in the image without using the vehicle detection, the method further improves the accuracy of the taillight signals recognition. The recognition rate of the brake light can be increased to more than 90%, and the recognition rate of the turn signal can be increased from about 80% to 87%.

The remainder of the paper is organized as follows. Studies related to the current research are described in the Section 2. Vehicle detection-related work is described in the Section 3. Taillight extraction and detection, as well as light signal identification are described in the Section 4. Finally, the discussions and conclusions are given.

2. Related Work

Because of the non-uniformity of vehicle taillight shapes and positions as well as their relatively small outline sizes, it is a challenge to detect and recognize the taillight semantic under real urban road environment condition in daytime. Therefore, there are relatively few studies on taillight detection and light semantic recognition in daytime urban traffic. Although, more interference and complex situations exist on the road during day, there is an advantage that the vehicle outline can be located clearly. Based on that outline, the taillight outline can be determined.

In general, algorithms of taillights detection can be divided into the following categories: feature-based methods, machine-learning-based methods, and multi-sensor fusion methods.

(1) Feature-based methods

Rezaei et al. used chain coding to maintain the original accuracy in the process of taillight outline detection and analyze the geometric rules of the taillight outline. They used virtual symmetry detection technology to locate the taillight position [7]. Based on the cognitive theory, Weis et al. decomposed a video input stream into color, shape and other features which are closely combined with an image model and created the pixel values of interest in taillight areas according to the characteristics of taillights and atmospheric effects caused by external lighting and weather conditions [8]. Jen et al. analyzed the taillight area color and brightness to find the invariant characteristics of the light scattering areas. Then, they used a classifier to train its dynamic frequency response and the size of the light scattering region to detect the taillight's brake light signal. However, this method did not consider the influence of illumination variation on light scattering in the taillight area [9]. Some scholars have extracted the correct taillight status by calculating the correlation of regional brightness values of the possible taillight areas of the vehicles on the road. They used the codebook theory to realize the all-weather tracking and detection of taillights, and the detection and identification of brake lights, left-turn lights, and right-turn lights [10–12]. The algorithm based on a priori information uses a relatively specific feature to detect the taillight. The advantage of this method is that it is faster to detect and can be applied to detection in a dynamic background. However, the disadvantage is that the relatively low level of the feature leads to a decrease in the ability to distinguish the target, which leads to an increase in false detection.

(2) Machine learning-based methods

Ming et al. selected 8 directions and 5 different scales of Gabor filters to extract features of the vehicle taillight images and trained a BP neural network to extract the taillight distribution characteristics [13]. This method needs to further improve the matching effect of the taillights and solve the problem of poor detection of the red vehicles. Wang et al. used a vehicle's rear appearance image to learn "brake light mode" through a multi-layer perceptron neural network in a large database and trained a depth classifier to judge whether the taillight was in the normal or braking state [14]. Wang et al. proposed a method to optimize the faster RCNN for vehicle detection by improving multi shapes receptive field generation, anchor generation optimization and ROI assignment to improve detection speed and accuracy [15]. Zhang et al. proposed a single deep neural network for vehicle detection in urban traffic surveillance. Different feature extractors for target classification were used in this article, which improves detection speed and accuracy. The feature pyramid was used to accurately generate the vehicle bounding box and accurately classify the vehicle category [16]. Machine learning-based algorithms generally use feature operators and classifiers, where the feature operators are manually extracted. This method is generally based on the characteristics of a single scene design. So, the detection results will be influenced and changed, when the scene changes.

(3) Multi-sensor fusion methods

Jin et al. combined millimeter wave radar and machine vision. This method can effectively identify the preceding vehicle at night. By extracting the vehicle characterization features and integrating information using D-S evidence theory, the vehicle taillights can be detected. However, this method is affected when the target is occluded or overlaps with another object [17]. Manuel et al. proposed a feature-based method for on-road vehicle detection in urban traffic and determined a rough upper bound of the intensity for shadow which reduces false positives rate [18].

In order to solve the problem that taillight detection is greatly affected by complex traffic environment, this paper proposes a method of combining vehicle detection and taillight detection. The vehicle detection of the preceding vehicle is completed using the Faster R-CNN network. On this

basis, the taillight area is extracted in the HSV color space, and the judgment of the taillight lighting state is completed.

3. Vehicle Detection

3.1. Faster RCNN Model

The traditional method of detecting taillights is to detect taillights directly in the image using feature matching or image processing techniques. This method will cause more interference in the image, which is not conducive to the detection of the tail light. Therefore, this paper introduces the technology of vehicle detection. After the vehicle detection is completed, the taillight detection can effectively reduce the interference in the image.

In the process of vehicle detection, there may be image quality and angle, as well as illumination and other factors, resulting in image quality degradation or target scale deformation. The vehicle detection method based on Faster RCNN can effectively eliminate interference and has relatively strong generalization ability. Neural networks have always been one of the most popular research topics in the field of artificial intelligence. It models the process of information processing by abstractly simulating the human brain and establishes probabilities or classification models based on different methods of neuron connections.

In recent years, the research on the structure of the CNN has been deepening. The CNN is widely used in various fields, such as behavior recognition, pedestrian detection, human posture recognition, and so on [19,20]. Recently, the CNN has been developed toward artificial intelligence, such as speech recognition and natural language processing, and so on.

According to current research and investigation, CNN can learn high-level image group features from a massive amount of image information. AlexNet has achieved excellent results in the field of large-scale image classification. It was designed in 2012 by the winners of ImageNet competitions, Hinton and Krizhevsky [21]. RCNN based on region feature extraction transforms the target detection problem through the region proposal method with the help of the good feature extraction and classification performance of CNN. This is a milestone in applying the CNN method to the object detection problem [22]. A Fully Convolutional Network (FCN) achieves the pixel-level classification of images, thus resolving the problem of semantics-level image segmentation [23].

CNN is a powerful feature extraction method. Szegedy et al. tried to regard object detection as a regression problem, but the detection effect on VOC2007 is average [24]. Sande et al. used the candidate region method to solve the detection problem and obtained a good effect on VOC2007 [25]. Fast RCNN was proposed in 2015, and it mapped candidate regions to the last layer feature map of CNN, which solves the problem of many repeated computations in RCNN and improves the speed of object detection [26]. In the same year, the Faster RCNN was proposed, whose basic structure includes the convolutional layer, the Region Proposal Network (RPN) layer, the ROI pooling layer and the classification layer [27], as shown in Figure 1. It further improves the speed and accuracy of object detection by using the shared features of RPN and Fast RCNN. Faster RCNN, one of the mainstream frameworks for object detection, was built in and has been used since late 2015, although Mask RCNN and other improved frameworks have been developed since then. The basic structure of Mask RCNN has not changed much and Faster RCNN still has the best accuracy in practical applications.

Figure 1. Basic structure of Faster RCNN.

3.2. The Construction of Faster RCNN

The CNN layer is mainly composed of the convolution layer, the pooling layer, and the ReLU layer. The convolutional layer uses specific models to extract features from the target image. Then, the target image is sorted, identified, and predicted by these extracted features. It assumes that the size of the input image is $M \times N$, the size of the filter kernel is $K \times K$, the stride is S , and the number of padding pixels is P . Then, the method for calculating the size of the feature map of the convolutional layer output is as follows:

$$O = (I - K + 2P) / S + 1 \quad (1)$$

In this paper, the parameters of the convolutional layer are defined as: kernel_size = 3, pad = 1. Then the original image becomes $(M + 2) \times (N + 2)$. The convolution operation is applied with the filter kernel size of 3×3 and the output size will be $M \times N$. This method ensures that the convolutional layer in the CNN structure does not change the sizes of the input and output matrices. Pooling layer is usually connected after the convolutional layer to aggregate the features and reduce computational complexity by reducing dimensions. First, the proposal feature maps must go through the full connect layer and softmax layer to calculate the classification proposals and output the detected probability vector cls_prob . Next, the position offset $bbox_pred$ of each proposal is obtained by bounding box regression, which is used to regress a more accurate object detection boundary. It assumes that the input image is size $M \times N$, the filter kernel is size $K \times K$, the stride is S , and the padding is P . Then, the method for calculating the size of the feature map of the pooling layer output is as follows:

$$O = (I - K) / S + 1 \quad (2)$$

In this paper, the parameters of the pooling layer are defined as: kernel size = 2 and stride = 2. A matrix of size $M \times N$ becomes $(M/2) \times (N/2)$ after the pooling layer operation.

The RPN network is composed of two structures. One part classifies anchors using the softmax classifier. After the classification, the foreground (the object of detection) and the background can be obtained. Another part is used to calculate the bounding box regression offset of the anchors, which can achieve more accurate proposals. The proposal layer, integrated with foreground anchors and bounding box regression offsets, is used to obtain the boundary candidate boxes for vehicle target. Certain unsuitable proposals are excluded. Then the RPN achieves the target localization by using the proposal layer.

The anchors are rectangles generated by a group of RPN. The four values ($x1, y1, x2, y2$) of each row in the matrix represent the coordinates of the top left corner and the bottom right corner of the rectangular frame.

There are nine rectangular frames in the matrix with different ratios of length to width. Therefore, anchors introduce a multi scales method for target detection. These nine anchors traverse CNN and obtain feature maps after calculation. Each point uses these nine anchors as the initial detection frame, but the results obtained are inaccurate.

The bounding box regression image example is shown in Figure 2. The correct Ground Truth (GT) data is shown by the external thick boundary, and the extracted foreground anchors are shown by the internal thin boundary. In this process, the thin boundary detection range should be adjusted slightly so that the identified foreground anchors are closer to GT. Generally, the window is represented by four-dimensional vectors (x, y, w, h), which represent the coordinate position, width and height of the central point of the window, respectively. Assume boundary A represents the original foreground anchors. This process is to find a mapping so that boundary A can be mapped to a regression window GT' that is closer to the actual window GT.

Figure 2. Boundary regression of the target.

3.3. Experimental Methods and Results: Faster RCNN construction

3.3.1. Comparison of the Detection Performance of Different Models

This paper mainly tests the vehicles in the road and produces 2500 vehicles image data sets collected in the real road environment. The camera resolution of the acquisition equipment is 1280×720 . Due to the limitation of data workload, the data set mainly includes sedans and sport utility vehicles. Then we use an image labelling tool to mark the image dataset with the ROI with each image containing 1 to 2 marked target vehicles, as shown in Figure 3. Figure 3 shows the marked vehicles, and the target position of the vehicle mark. The small data set can establish the RCNN training process faster. We randomly select 1500 sheets as the training set and 1000 sheets as the test set.

Figure 3. Production of vehicle data sets.

The article's vehicle inspection data set has fewer data problems, so the data is supplemented and improved. The article combines a homemade dataset with a subset of the standard dataset. In many public driving data sets, the BDD100K dataset from Berkeley contains a variety of weather data, which contains approximately 100,000 images of 1280×720 containing vehicle targets. The article selects 6000 images as the test set and 4000 images as the test set. Therefore, there are 7500 images in the training set and 5000 images in the test set.

To verify the detection capability of the Faster RCNN model, three representative vehicle detection methods based on CNN are selected in this paper. The methods are the CNN, the RCNN and the Fast RCNN. The different method approaches are presented in Table 1.

Table 1. Features of the various vehicle detection methods.

Detection Method	Candidate Box Extraction Method	Feature Extraction Method	Classifier
CNN	Sliding window	CNN	SVM
RCNN	Selective Search Algorithms	CNN	SVM
Fast RCNN	Selective Search Algorithms	CNN	Softmax
Faster RCNN	Region Proposal Network	CNN	Softmax

In this paper, training sets are used to train various models, and test sets are used to test the effect, as shown in Table 2. Among them, the accuracy rate calculation method is: correct detection/(correct detection + false detection). The recall rate calculation method is: correct detection/(correct detection + missed detection).

Table 2. Comparison of vehicle detection methods.

Detection Method	Correct Detection	False Detection	Missed Detection	Precision Rate/(%)	Recall Rate/(%)
CNN	3136	1058	806	74.8	79.6
RCNN	3581	742	677	82.8	84.1
Fast RCNN	3657	606	737	85.8	83.2
Faster RCNN	4017	394	589	91.1	87.2

The experimental results show that the principle and mechanism of a method will lead to different experimental results. The sliding window method of CNN obtains the sliding window features in different image positions, but this method is less efficient than CNN. RCNN uses a selective search algorithm to extract a candidate boundary and it combines a Support Vector Machine (SVM) to extract regions. RCNN has a better extraction effect. Compared with RCNN, Fast RCNN also uses a selective search algorithm to extract a candidate boundary, but it employs the softmax classifier, which leads to a better classification result because it introduces an interclass competition mechanism. Compared with Fast RCNN, Faster RCNN uses RPN to extract the candidate boundary and omits a selective search algorithm while achieving end-to-end training. Faster RCNN uses the convolution feature and RPN sharing method based on Fast RCNN, and its detection time averages about 94 ms.

3.3.2. Detection Performance with Different Scenes

The standard sample set contains images in different weathers, including 53,535 sunny days, 7125 rainy days, 7888 snowy days, and 181 fog days. This article selects images from some different weather conditions from the standard sample set. The partial data of the standard sample set and the self-made data are combined to form the test set of this paper. In the data set, there are 1272 images on sunny days, 537 images on foggy days, 279 images on snowy days, and 412 images on rainy days. The number of data in different weather conditions is shown in Table 3.

From the test set, 150 frames of images during the sun, cloudy, foggy and rainy days were selected for testing. The test data is shown in Table 4.

Table 3. The number of data in different weather conditions.

Scenes	Sunny Days	Rainy Days	Foggy Days	Snowy Days
Standard sample set	2578	1588	113	1721
Homemade sample set	1272	412	537	279
Total	3850	2000	650	2000

Table 4. Detection results for different weather conditions.

Scenes	Correct Detection	False Detection	Missed Detection	Precision Rate/(%)	Recall Rate/(%)
Sunny Days	124	12	14	91.2	89.9
Rainy Days	111	22	17	83.5	86.7
Foggy Days	98	25	27	79.7	78.4
Snowy Days	117	19	14	86.0	89.3

In different road environments, trained neural networks can better detect vehicles. These include factors like excessive light intensity, complex road background, and excessive distance from preceding vehicle. The vehicle detection results are shown in Figure 4. The correct test results are shown in Figure 4a,b. Figure 4c shows the result of false detection. Figure 4d shows the result of the missed test.

Figure 4. Vehicle detection results. (a) Correct vehicle test results; (b) Correct vehicle test results; (c) Correct vehicle test results; (d) Correct vehicle test results.

4. Taillight Detection and Light Signal Identification Based on Image Processing

This chapter is mainly to complete the taillight detection and light recognition work of the vehicle. By recognizing the taillight signals, the driving intention of the preceding car can be effectively understood. Under the normal driving conditions of urban roads, the most frequently used ones are brake lights and turn signals. This paper mainly designs approaches for taillight detection based on the correlation among the three parameter and the histograms of the two taillight states. The method of identifying taillight state by parameters. The procedure of this method is shown in Figure 5.

Figure 5. The process of taillight detection and light recognition.

4.1. Color Space Conversion

Due to the complex and varying lighting conditions during the day, it is difficult to process rapidly changing color information by simply using the R channel thresholds in the RGB color model to extract the taillight regions. Detecting the taillights with the R-channel threshold results in inaccurate segmentation of the taillight area for red vehicles, as shown in Figure 6.

Figure 6. Results of color space conversion. (a) Image of vehicle in RGB space; (b) The segmentation result of the R channel threshold; (c) The segmentation result of the HSV color space.

Figure 7a is an RGB color space image, and Figure 7b is an HSV color space image after conversion. From the images, the taillight area in the target image in Figure 7b is more obvious.

Figure 7. Image processing result. (a) Image of vehicle in RGB space; (b) Image of vehicle in HSV space.

The image threshold segmentation is performed on the HSV color space image. Image segmentation is performed using a specific threshold. A pixel in which the gradation value is greater than or equal to a certain threshold is judged to belong to the target object, and a pixel whose gradation value is smaller than the threshold is excluded from the target.

Then select Sobel for edge detection in the taillight area. The image is denoised as much as possible and the internal holes are narrowed to reconnect the adjacent areas. The image is subjected to a morphological closing operation to eliminate narrow discontinuities and small voids in the taillight region. A morphological opening operation is performed on the image to smooth the image boundary. The detection result is shown in Figure 8.

Figure 8. Grayscale adaptive threshold segmentation results in HSV space.

4.2. Taillight Detection

To detect the correct taillight pixel pair, we first need to determine the centroid of the pixel cluster in the image. Then we select the side of the image as the primary point, the other particles as the dependent points, and calculate the vertical distance between the primary point and the other dependent points. Histograms of directional gradient correlation, distance correlation, and color correlation are used to detect vehicle taillight pairs.

4.2.1. Correlation Detection Based on Histogram of Oriented Gradients

A histogram of oriented gradients (HOG) is a kind of feature histogram. For an image $f(x, y)$, its gradient magnitude $|\nabla f(x, y)|$ and direction angle Φ at the point (x, y) are defined as:

$$\begin{aligned} |\nabla f(x, y)| &= \text{mag}(\nabla f(x, y)) = (G_x^2 + G_y^2)^{1/2} \\ \Phi(x, y) &= \arctan(G_y/G_x) \end{aligned} \quad (3)$$

G_x and G_y respectively represent the gradient in the x direction and the gradient in the y direction.

If we divide the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ in the gradient direction in Equation (6) into k uniform intervals (bin), use bin_k to represent the number k gradient direction and use K to represent the dimension values. Then the gradient magnitude weight projection function $Q_k(x, y)$ in the number k gradient direction of the pixel point $Q_k(x, y)$ can be expressed as:

$$Q_k(x, y) = \begin{cases} |\nabla f(x, y)| & \varphi(x, y) \in \text{bin}_k \\ 0 & \varphi(x, y) \notin \text{bin}_k \end{cases} \quad (4)$$

$$k = 0, 1, \dots, K - 1$$

The gradient magnitude weight projection function $Q_k(x, y)$ in Equation (4) is the gradient magnitude of the pixel, which reflects the edge information of the pixel to some extent. According to Equation (4), the gradient feature of each pixel point (x, y) is a K dimensional vector. The gradient direction histogram of the image is the histogram statistics of the K dimensional gradient features of all the pixels in the image. The candidate taillight areas obtained by the image processing method is as shown in Figure 9.

Figure 9. The taillight candidate areas extracted by the gradient feature.

The correlation coefficients of the gradient histograms of the two frames in the X and Y directions are calculated, as shown in Table 5. Analyzing the correlation coefficient of the taillight candidate area, the taillight pair candidate coefficients on both sides have the highest gradient histogram correlation coefficient values in the X direction and the Y direction, which are considered to be the correct taillight pair matching results.

Table 5. Correlation coefficient for the taillight candidate regions.

Taillight	Gradient_x	Gradient_y
Picture (a)_left\Picture (a)_right	0.9646	0.9884
Picture (a)_left\Picture (a)_other	0.8917	0.9380
Picture (a)_right\Picture (a)_other	0.9178	0.9638
Picture (b)_left\Picture (b)_right	0.9184	0.9642
Picture (b)_left\Picture (b)_other	0.8432	0.7691
Picture (b)_right\Picture (b)_other	0.8461	0.7943

4.2.2. Correlation Detection Based on Location Relationship

In the distance-pairing restrictions, the taillights are generally symmetrically distributed. On a flat road, although a camera angle offset may exist, the lights on the two sides are roughly in the same horizontal line. The centroids of the two taillights image randomly selected are shown in Figure 10.

Figure 10. Centroid of the taillight areas.

Considering the restrictions of the centroid of the two regions in the vertical and horizontal directions, Figure 11 shows the height difference between the centroid coordinates of the two points. In the vertical direction Y , the absolute value of the distance difference between the two centroids is less than the maximum value of the height of the cluster pairs between the two centroids. In the horizontal direction X , the distance between the two points is less than the width of the car body.

Figure 11. Symmetry condition in the vertical direction.

4.2.3. Correlation Detection Based on Color

To further reduce possible false detections in the color correlation restriction conditions, a correlation detection on the pixel cluster pairs satisfying the conditions is carried out. Pixel cluster pairs are extracted from the original image. The linear correlation coefficient formula is used, as follows:

$$r(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var[X]Var[Y]}} \quad (5)$$

The sum of the correlation coefficient vectors of the mean value of the three monochrome channels (red, green, and blue channels) in the candidate taillight pair region is calculated to determine whether taillight pairs belong to the same vehicle. $Var[X]$ is the variance of X . $Var[Y]$ is the variance of Y . $Cov(X, Y)$ is the covariance of X and Y . The calculation method is expressed as:

$$Cor(X, Y) = E\{[X - E(X)][Y - E(Y)]\} \quad (6)$$

$E(X)$ represents the expectations of component X . $E(Y)$ represents the expectations of component Y .

Using the original image based on the RGB color space to calculate the sum of the correlation of R, G, and B channels, and setting the dynamic threshold to a_{colour} , as follows:

$$r(R_1, R_2) + r(G_1, G_2) + r(B_1, B_2) < \alpha_{colour} \quad (7)$$

a_{colour} represents the dynamic threshold.

If the sum of the correlation is less than the threshold, the two regions are recognized as the correct taillight matching items.

The taillight detection method using the positional relationship constraint alone and the taillight detection method used in this article are compared as shown in Figure 12. The taillight detection method using the positional relationship constraint alone cannot effectively eliminate the interference points that are close in positional relationship. The method used in this paper can effectively eliminate the interference with similar position and accurately detect the taillight.

Figure 12. The detection results of the taillight pairs. (a) The result of the taillight pair detection using only the positional relationship constraint; (b) The results of the detection using the taillights of the method in this article.

4.3. Taillight Signal Identification

4.3.1. Histogram Characteristic Parameter Extraction

The gray level histogram feature parameters of the image are measured at certain specific pixels or their neighborhoods, which can well describe the gray level features in the image. The histogram shows a global description of the grayscale image, but the histogram is not directly used as a feature. Instead, its mean, variance, energy, and entropy are used as the characteristics of the differences between the categories.

For an image f , if we assume the total number of pixels is N , the maximum gray level of the image (the grayscale image is 255) is L , the number of pixels with grey level k is N_k , then the gray-level histogram of f can be represented as:

$$h_k = \frac{N_k}{N} \quad k = 0, 1, \dots, L-1 \quad (8)$$

The mean value is the mean value of the grey level probability distribution, as follows:

$$\bar{f} = \sum_{k=0}^{L-1} k h_k \quad (9)$$

The variance is the measurement of the discrete type of the image grey value distribution, as follows:

$$\alpha_f^2 = \sum_{k=0}^{L-1} (k - \bar{f})^2 h_k \quad (10)$$

The energy represents the uniformity of the grey level distribution, as follows:

$$f_N = \sum_{k=0}^{L-1} (h_k)^2 \quad (11)$$

The entropy is the measurement of the amount of information in an image, as follows:

$$f_E = -\sum_{k=0}^{L-1} h_k \log_2 h_k \quad (12)$$

The average brightness of the image can be expressed by the mean value. The dispersion of the image grey level distribution can be expressed by the variance. Due to the influence of image sampling on the mean and the variance, the target image is usually normalized for classification. The energy is the secondary moment of the grey distribution to the original point. If the grey value of the image is in an equiprobability distribution, the energy is the smallest. Otherwise, the energy is larger. In information theory, the entropy reflects how much information is contained in an image.

4.3.2. Brake Light Identification Method

Traditionally, the methods of color threshold segmentation or shape matching are usually used to determine the taillight state. Although these methods can roughly detect the outline and area of the taillight, they cannot judge the current taillight lighting state. In this paper, by analyzing and comparing the on and off states of the brake light, and by observing, we find that when the brake light is on, the inner layer is yellowish because of its stronger brightness, while the outer layer is red because of its weaker brightness. When the brake light is off, its color is more uniformly red. To summarize, the taillight status can be monitored according to brightness and color information. Therefore, the histogram features of the original taillight area image are analyzed. Figure 13a shows the image when the taillight is off. Figure 13b shows the image when the taillight is turned on.

Figure 13. Vehicle taillight area image. (a) Unlit left taillight image; (b) Illuminated left taillight image.

The histogram of the lighting state statistics of the taillights is shown in Figure 14a. The histogram of the closing status of the tail light is shown in Figure 14b. There are obvious differences between the taillight distributions in the two states.

Figure 14. Gray-level histograms. (a) Taillight histogram in the lit state; (b) Taillight histogram in unlit state.

The on status of the brake light is a continuous lighting process. The 5-frame taillight image with an interval of 1 s is randomly selected, as shown in Figure 15a–e. We can see clearly the on and off process of the brake lights. In a period of 5 s, we calculate the histogram features of the left and right taillights in the image and those in the background image. If the taillights on the two sides satisfy the conditions of continuous lighting of the histogram characteristic parameters at the same time in the same time interval, the brake lights can be recognized as lit.

Figure 15. The change of the vehicle's brake lights.

If we segment the left taillight image in the five frames in Figure 15, as shown in Figure 16a–e, and calculate the statistics of the histogram characteristic parameters, as shown in Table 6. Within 5 s after the background frame is determined, if the histogram characteristic parameters of the vehicle taillight area images are larger than that of the background frames in two continuous frames or more, the taillights of the preceding vehicle are recognized as lit.

Figure 16. Brake light brightness change chart during on and off processes.

Table 6. Correlation coefficient for taillight candidate region.

-	Mean Value	Energy	Entropy	Variance
1	119.9437	0.0096	7.1629	$3.4068 \times 10^3 + 03$
2	112.1328	0.0104	7.0349	$3.0960 \times 10^3 + 03$
3	128.5328	0.0083	7.4066	$4.2965 \times 10^3 + 03$
4	123.1580	0.0064	7.5766	$4.3271 \times 10^3 + 03$
5	126.6490	0.0071	7.3157	$4.8005 \times 10^3 + 03$

According to the five images extracted at equal intervals within 5 s, the histogram feature parameter list mentioned above is obtained, and obvious changes can be noted. The first two points are the parameters corresponding to the two frames when the light is off, and the last three points are the parameters corresponding to the three frames when the light is on. By analyzing the histogram feature parameters, we can know that the average value, entropy, and variance of the taillights when they are not lit are less than the values when the taillights are lit. The energy when the taillights are lit is greater than the energy when the taillights are not lit. Whether the tail light is on can be determined by setting a dynamic threshold.

4.3.3. Identification Method of the Direction Light

The direction light of the vehicle flashes intermittently when it is on. The taillight images is shown in Figure 17. In this paper, the statistical results of the histogram distribution of the directional lights in the taillights on state are shown in Figure 18a, and the histogram distribution statistics of the directional lights in the taillights off state are shown in Figure 18b. The histogram distribution characteristics are obviously different when the direction light is on compared to when it is off. Therefore, the method of statistical histogram characteristic parameters can be used to analyze the lighting color and characteristics of the direction lights. The flickering of the direction lights causes a periodic change in the histogram characteristic parameters.

Figure 17. Vehicle taillight area image. (a) Unlit left taillight image; (b) Illuminated left taillight image.

Figure 18. Gray-level histogram. (a) Taillight histogram in the lit state; (b) Taillight histogram in unlit state.

Another important feature of the turn signal light is that there is an alternating process of turn-on and turn-off with a period of 1 s, as shown in Figure 19a–e. From the figure, it can be seen clearly that the direction light has a periodic on and off characteristic.

Figure 19. Brightness change chart of the direction light during on and off process.

The 5 frames of the left taillight in Figure 19 are segmented as images which are shown in Figure 20, and the characteristic parameters are shown in Table 7. By analyzing the data in the histogram characteristic parameter list, the mean value, energy, and variance are more sensitive to the direction lights. Therefore, these three parameters are selected as the conditions for determining the direction light's on or off state. Within 5 s after determining the background frame, the histogram feature parameters are calculated once per second. If the parameters of the taillight area on both sides of the vehicle are greater than the parameters of the background frame, the direction light of the preceding vehicle is recognized as lit.

Figure 20. On and off process change chart of the direction light.

Table 7. Histogram characteristic parameter list.

-	Mean Value	Energy	Entropy	Variance
1	65.5098	0.0097	7.1047	$3.3507 \times 10^3 + 03$
2	93.5233	0.0193	7.0696	$5.9075 \times 10^3 + 03$
3	74.0808	0.0033	6.1351	$3.6327 \times 10^3 + 03$
4	89.4964	0.0258	6.8357	$6.9519 \times 10^3 + 03$
5	93.1196	0.0234	6.9524	$6.5517 \times 10^3 + 03$

According to the 5 images extracted at equal intervals within 5 s, the histogram feature parameter list mentioned above can be obtained. The first two points are the parameters corresponding to the two frames when the light is off, and the last three points are the parameters corresponding to the three frames when the light is on. By analyzing the histogram characteristic parameters, we find that the mean value, energy, and variance when the light is off are less than their values when the light is on, and the entropy value is not sensitive to whether the lamp is on.

4.4. Experiment Results and Analysis

By using the methods described in this paper, break light detection results and direction light result are shown in Figure 21. The time spent on vehicle detection and taillight detection takes about 200 ms.

Figure 21. Taillight detection results. (a) Brake light detection results; (b) Direction light detection results.

If the taillight is turned on but not detected, it is called a false alarm. If the taillight is off but is identified to be on, it is termed a missing alarm. In some cloudy and rainy weather conditions, the taillights can be detected, but in severe weather conditions, such as fog, heavy rain, and bright light, missed detections will occur. Some test results are shown in Figure 22. The experimental results are shown in Table 8.

Figure 22. Taillight detection result in different weather conditions. (a) The result of correct detection under the normal weather; (b) The result of missed detection under the intense illumination; (c) The result of the missed detection under the rainy weather.

Table 8. Experimental result list.

-	Experiment Frames	Accuracy Rate/(%)	False Alarm/(%)	Missing Alarm/(%)	Average Time
Brake light	200	92.5	6.0	1.5	107
Direction light	200	87.5	8.5	4.0	131

The longer the distance that can be identified, the more the system can give the driver an earlier warning and improve the safety of driving. In the experiment, under the normal driving condition of the urban road, the taillight of the preceding vehicle can be accurately recognized when the distance from the preceding vehicle is approximately within the range of 30 m. When the distance increases, the area of the taillight in the image is too small and consists of only a few pixels. In this case, the front taillight cannot be detected, as shown in Figure 23.

Figure 23. The distance of the preceding vehicle on the road. (a) The distance from the preceding vehicle of the road is about 3 m, and the taillight image is relatively clear; (b) The distance from the preceding vehicle of the road is about 30 m, and the taillight image can be distinguished; (c) The distance from the preceding vehicle of the road is about 60 m, and the taillight image is difficult to distinguish.

5. Discussion

There are some models where the taillights are complex polygons or irregular shapes. Because video frame images need to be processed by image pre-processing, image filtering and morphological processing, the binarization and grayscale transformation may cause incompleteness in the taillight area, which may result in an inability to identify the lamp area.

In some rainy weather conditions, the taillight area can be detected. However, in severe weather conditions, such as heavy fog, heavy rain, bright light, and so on, the image target area is partially or completely blocked, and objects caused by strong light distortion of the surface color will result in the inability to accurately capture and identify the image target area, resulting in a missed detection. In addition, the image data collected under different weather conditions needs to be further expanded, which can further verify the robustness of the system.

In the future, it will be necessary to design a more robust taillight detection method for different weather conditions and different taillight shapes. A distance detection sensor system will be considered to integrate into this system.

6. Conclusions

Effectively identifying taillight signals can help us understand the driving intentions of preceding vehicle. In this paper, CNN is used to detect preceding vehicles on the current road. The position, color and gradient histogram features of the taillight area are combined to detect the taillight areas on both sides. The change rules of the histogram characteristic parameters under different taillight lighting states are analyzed, and the taillight signal of the vehicle is identified.

This study collected images of vehicles taken under different road conditions and selected images from different scenes and marked areas of interest. The faster RCNN is then built and trained and tested through the image data set to achieve high-precision vehicle detection.

In this study, the taillight area segmentation threshold is obtained by weighting three channels in the HSV color model. The taillight pair matching is completed by using three correlation detections of gradient histogram feature, color feature and position space feature to improve the accuracy of taillight detection. The brake lights and turn lights are recognized according to the histogram characteristic parameters of the taillight region to improve the accuracy of the taillight detection.

The results show that the algorithm can detect vehicles and taillights better in daytime traffic environments, but there are still some shortcomings in the study of this issue. In future research work, it will be necessary to collect and produce more marked road vehicle images to train a more robust Faster RCNN neural network.

Author Contributions: Conceptualization, Z.W. and W.H.; Methodology, Z.W. and W.H. and P.Y.; Software, W.H.; Validation, W.H. and P.Y.; Formal analysis, Z.W.; Investigation, L.Q.; Resources, Z.W. and L.Q.; Data curation, Z.W.; Writing—original draft preparation, Z.W. and W.H.; Writing—review and editing, W.H., S.G., and N.C.; Supervision, Z.W. and N.C.; Project administration, Z.W. and N.C.; Funding acquisition, Z.W.

Funding: This research was funded by Science and Technology Support Plan Project of Hebei Province, grant number 17210803D.

Acknowledgments: We are grateful for the assistances of the Fujian Province University Key Lab for Industry Big Data Analysis and Application.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Chen, T.; Fang, W.; An, L.F. Vehicle rear-end collision warning algorithm based on DSRC. *J. Automot. Saf. Energy* **2016**, *8*, 164–169.
- Susel, F.; Takayuki, I. Driver classification for intelligent transportation systems using fuzzy logic. In Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil, 1–4 November 2016; pp. 1212–1216.
- Liu, Y. Big Data Technology and Its Analysis of Application in Urban Intelligent Transportation System. In Proceedings of the 2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Xiamen, China, 25–26 January 2018; pp. 17–19.
- Xu, Y.; Yu, F. Research on the Image Acquisition and Camera Control of Machine Vision Camera Based on LabVIEW. In Proceedings of the 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, 26–27 August 2013; pp. 499–502.
- Jin, M. A New Vehicle Safety Space Model Based on Driving Intention. In Proceedings of the 2013 Third International Conference on Intelligent System Design and Engineering Applications, Hong Kong, China, 16–18 January 2013; pp. 131–134.
- Wu, H.; Li, Y.; Wu, C.; Ma, Z.; Zhou, H. A longitudinal minimum safety distance model based on driving intention and fuzzy reasoning. In Proceedings of the 2017 4th International Conference on Transportation Information and Safety (ICTIS), Banff, AB, Canada, 8–10 August 2017; pp. 158–162.
- Rezaei, M.; Terauchi, M.; Klette, R. Robust Vehicle Detection and Distance Estimation Under Challenging Lighting Conditions. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2723–2743. [[CrossRef](#)]
- Weis, T.; Mundt, M.; Harding, P.; Ramesh, V. Anomaly detection for automotive visual signal transition estimation. In Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–8.
- Jen, C.; Chen, Y.; Hsiao, H. Robust detection and tracking of vehicle taillight signals using frequency domain feature based Adaboost learning. In Proceedings of the IEEE International Conference on Consumer Electronics—Taiwan (ICCE-TW), Taipei, Taiwan, 12–14 June 2017; pp. 423–424.
- Casares, M.; Almagambetov, A.; Velipasalar, S. A Robust Algorithm for the Detection of Vehicle Turn Signals and Brake Lights. In Proceedings of the 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, Beijing, China, 18–21 September 2012; pp. 386–391.
- Almagambetov, A.; Casares, M.; Velipasalar, S. Autonomous tracking of vehicle rear lights and detection of brakes and turn signals. In Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defence Applications, Ottawa, ON, Canada, 11–13 July 2012; pp. 1–7.
- Cui, Z.; Yang, S.; Tsai, H. A Vision-Based Hierarchical Framework for Autonomous Front-Vehicle Taillights Detection and Signal Recognition. In Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems, Las Palmas, Spain, 15–18 September 2015; pp. 931–937.
- Qing, M.; Jo, K. Vehicle detection using tail light segmentation. In Proceedings of the 2011 6th International Forum on Strategic Technology, Harbin, China, 22–24 August 2011; pp. 729–732.
- Wang, J.; Zhou, L.; Pan, Y.; Lee, S. Appearance-based Brake-Lights recognition using deep learning and vehicle detection. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 19–22 June 2016; pp. 815–820.
- Wang, Y.; Liu, Z.; Deng, W. Anchor Generation Optimization and Region of Interest Assignment for Vehicle Detection. *Sensors* **2019**, *19*, 1089. [[CrossRef](#)] [[PubMed](#)]
- Zhang, F.; Li, C.; Yang, F. Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation. *Sensors* **2019**, *19*, 594. [[CrossRef](#)] [[PubMed](#)]
- Jin, L.; Fu, M.; Wang, M.; Yang, Y. Vehicle detection based on vision and millimeter wave radar. *J. Infrared Millim. Waves* **2014**, *33*, 465–471.
- Manuel, I.; Tardi, T.; Pérez-Oria, J.; Robla, S. Shadow-Based Vehicle Detection in Urban Traffic. *Sensors* **2017**, *17*, 975. [[CrossRef](#)]

19. Zhou, Z.; Duan, G.; Lei, H.; Zhou, G.; Wan, N.; Yang, W. Human behavior recognition method based on double-branch deep convolution neural network. In Proceedings of the 2018 Chinese Control and Decision Conference (CCDC), Shenyang, China, 9–11 June 2018; pp. 5520–5524.
20. Yang, Z.; Li, J.; Li, H. Real-time Pedestrian and Vehicle Detection for Autonomous Driving. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–29 June 2018; pp. 179–184.
21. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. In Proceedings of the 26th Annual Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 2–6 December 2012; pp. 1097–1105.
22. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
23. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the 2015 IEEE Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
24. Szegedy, C.; Toshev, A.; Erhan, D. Deep Neural Networks for Object Detection. In Proceedings of the Advances in Neural Information Processing Systems, Nevada, NV, USA, 5–10 December 2013; pp. 2553–2561.
25. Sande, K.; Uijlings, J.; Gevers, T.; Smeulders, A. Segmentation as selective search for object recognition. In Proceedings of the IEEE Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1879–1886.
26. Girshick, R. Fast R-CNN. In Proceedings of the IEEE Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 176–183.
27. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).