Sens Imaging (2017)18:20
DOI 10.1007/s11220-017-0173-8

ORIGINAL PAPER

Vehicle Color Recognition with Vehicle-Color Saliency Detection and Dual-Orientational Dimensionality Reduction of CNN Deep Features

Qiang Zhang[1] · Jiafeng Li[1] · Li Zhuo[1,2] · Hui Zhang[1] · Xiaoguang Li[1]

Received: 22 February 2017 / Revised: 24 April 2017

Abstract Color is one of the most stable attributes of vehicles and often used as a valuable cue in some important applications. Various complex environmental factors, such as illumination, weather, noise and etc., result in the visual characteristics of the vehicle color being obvious diversity. Vehicle color recognition in complex environments has been a challenging task. The state-of-the-arts methods roughly take the whole image for color recognition, but many parts of the images such as car windows; wheels and background contain no color information, which will have negative impact on the recognition accuracy. In this paper, a novel vehicle color recognition method using local vehicle-color saliency detection and dual-orientational dimensionality reduction of convolutional neural network (CNN) deep features has been proposed. The novelty of the proposed method includes two parts: (1) a local vehicle-color saliency detection method has been proposed to determine the vehicle color region of the vehicle image and exclude the influence of non-color regions on the recognition accuracy; (2) dual-orientational dimensionality reduction strategy has been designed to greatly reduce the dimensionality of deep features that are learnt from CNN, which will greatly mitigate the storage and computational

✉ Li Zhuo
zhuoli@bjut.edu.cn

Qiang Zhang
sammyrow@163.com

Jiafeng Li
lijiafenga@163.com

Hui Zhang
huizhang@bjut.edu.cn

Xiaoguang Li
lxg@bjut.edu.cn

1 Signal and Information Processing Lab, Beijing University of Technology, Beijing, China

2 Collaborative Innovation Center of Electric Vehicles in Beijing, Beijing, China

burden of the subsequent processing, while improving the recognition accuracy. Furthermore, linear support vector machine is adopted as the classifier to train the dimensionality reduced features to obtain the recognition model. The experimental results on public dataset demonstrate that the proposed method can achieve superior recognition performance over the state-of-the-arts methods.

Keywords Vehicle color recognition · Vehicle-color saliency detection · Dual-orientational dimensionality reduction · CNN deep features

1 Introduction

Vehicle recognition plays an important role in Intelligent Traffic System (ITS), which can automatically recognize vehicle attributes by using computer vision technology, e.g. model [1, 2], license plate [3], color [4], logo [5], and type [6] as well. Color is one of the most stable attributes of vehicles and can be widely applied in some applications, such as traffic accident investigation, market survey, and so on. In complex environments, the visual characteristics of vehicle colors are easily affected by variations of illumination, weather, noise and other environmental factors, which results in vehicle color recognition being a challenging task.

Various vehicle color recognition methods have been proposed in recent years. Generally, the research work on vehicle color recognition can be divided in two stages: In the first stage, researchers aim to find robust handcrafted features to represent the color characteristics of the vehicle images and select suitable classifiers to train the features to achieve higher performance [4, 7–11], such as support vector machine (SVM). Color histogram is the most commonly used feature. For example, Baek et al. [7] extracted the color histogram in HSV color space and a two-dimensional feature vector using two components, H and S have been obtained. Kim et al. [8] extracted color histogram features from each component of HSI (Hue, Saturation, Intensity) color space. They selected an appropriate number for the color histogram bins of each component (e.g. 8 bins for H, 4 bins for S, and 4 bins for I). Compared with the conventional histogram methods where the same number of bins is extracted from each component, this method is able to reduce processing time and promote recognition accuracy.

Color features are easily affected by illumination, noise and other factors. Thus these methods usually cannot achieve satisfactory vehicle color recognition accuracy in actual complex environments. In order to enhance the robustness of vehicle color recognition, Chen et al. [9] proposed a method to implicitly select Regions of Interest (ROIs) from the vehicle image for vehicle color recognition. They combined different color histogram features in several color spaces and adopted BoW(Bag of Words) model to compactly represent the local features. Then, SVM is exploited to train the vehicle color recognition model. Hsieh et al. [10] took the effect of sunlight and car windows into account, and proposed a method to reduce the influence of illumination on recognition accuracy through mapping the inter-relation of multiple frames to solve the problem of color distortion.

Besides the above color recognition approaches, which take all pixels into consideration, Hu et al. [11] creatively obtained a foreground image by estimating RGB values of the car body. They tried to filter the non-color parts (e.g. wheels, windows etc.) automatically and obtained the prominent body color region that will be directly used for color recognition. The above methods have the advantages of simple principle, fast implementation, but recognition accuracy is usually not high. The features used are handcrafted, the designers should have specialized and professional prior knowledge about the specific task. Therefore, it is difficult to manually design appropriate features while facing new data or new tasks.

The second phase of vehicle color recognition is to obtain classification model or deep features from big data using deep learning. The existing work shows that compared to the traditional handcrafted features, the high-level deep features learnt from big data is capable of improving performance of vehicle color recognition significantly [12, 13]. Rachmadi et al. [12] converted the input images into two different color spaces, HSV and CIE Lab, and used them to train the vehicle color recognition model with a parallel cross convolutional neural network (CNN) architecture. In this method, Alex-Net [14] is applied and a 4096-dimensional feature vector is generated. The experimental results show that the recognition accuracy of the method can be 2% higher than that of literature [9]. Hu et al. [13] proposed a method in which the fifth-layer convolution features of Alex-Net have been exacted on 4 sub-regions of Spatial Pyramid Matching (SPM) [15] partitioning and the original image respectively, which will be then used to train SVM classifier. The experimental results on public vehicle color dataset have demonstrated that the proposed method can achieve a recognition accuracy of 94.69%, superior to the recognition performance of the handcrafted features based methods.

The advantage of deep learning based vehicle color recognition methods is that they are able to avoid the process of handcrafted features design, and automatically learn the deep features from big data. The high-level deep features can express the important information of the category while suppressing the irrelevant background information, and thus, can improve the robustness of recognition model. But it should be noted that the high recognition performance is usually achieved at the cost of repeated iterations and extremely high computational complexity.

In summary, the existing methods usually use the whole image to recognize the vehicle color and the dimensionality of the deep features is very high, which results in heavy storage and computational burden on the following processing.

In this paper, a novel vehicle color recognition method using CNN deep features has been proposed. Compared with the existing approaches, the novelty of the proposed method includes two parts. Firstly, vehicle-color saliency detection method has been proposed, which can exclude the negative influence of non-color regions on the recognition accuracy. Secondly, a dual-orientational dimensionality reduction strategy is proposed to reduce the dimensionality of the deep features while retaining their discriminative ability. Finally, linear SVM is adopted as the classifier for its efficiency and high precision. The experimental results on public vehicle color dataset demonstrate that the proposed method can achieve superior performance over the state-of-the-arts methods.

The rest of this paper is organized as follows: Sect. 2 describes the main idea and implementation details of the proposed vehicle color recognition method. Section 3 introduces the experimental results and analysis. Finally, the conclusions are drawn in Sect. 4.

2 The Proposed Vehicle Color Recognition Method

The framework of the proposed method is depicted in Fig. 1, which can be divided into four parts: vehicle-color saliency detection, deep feature extraction, dual-orientational dimensionality reduction, and classifier training.

Firstly, the specular-free image is extracted from the original image, which can effectively highlight the visual characteristics of the vehicle color, and then the vehicle-color saliency regions are detected according to the color information distribution of the specular-free image. The sub-images corresponding to the vehicle-color saliency regions are input to CNN to obtain the deep features through using Pre-training + Fine-tuning process. A high-dimensional 2D deep feature map can be generated after repeated iterations. To reduce the dimensionality of the deep features, 2D-PCA [16] is applied twice, to obtain a low-dimensional feature representation of the vehicle-color saliency regions. Finally, SVM is used as a classifier to train the recognition model on the dimensionality reduced features. Specifically, the vehicle-color saliency detection based on a specular-free image is described in Sect. 2.1, and the dual-orientational dimensionality reduction strategy is described in Sect. 2.2.

2.1 Vehicle-Color Saliency Detection

Actually, in many parts of the vehicle image, such as vehicle windows, wheels and background, there is no color information, but occupy a large part of the image, which often brings interference to the vehicle color recognition. Therefore, a vehicle-color saliency detection method is proposed in this paper to exclude the non-color regions in the vehicle image. Figure 2 shows the flowchart of vehicle-color saliency detection.

The dichromatic reflection model, which is generally used to describe the reflection of the majority of uneven materials, is firstly adopted to extract specular-free image, which can effectively remove the non-color parts, and highlight the vehicle color areas to a certain degree [17]. Then, the distribution characteristic of the color regions is computed by statistical analysis, and adaption threshold segmentation is applied to obtain the local vehicle color regions.

2.1.1 Specular-Free Image Generation

Specular-free is an effective way to precisely distinguish between chromatic area and non-chromatic area in the image by preserving the saturation values of all pixels

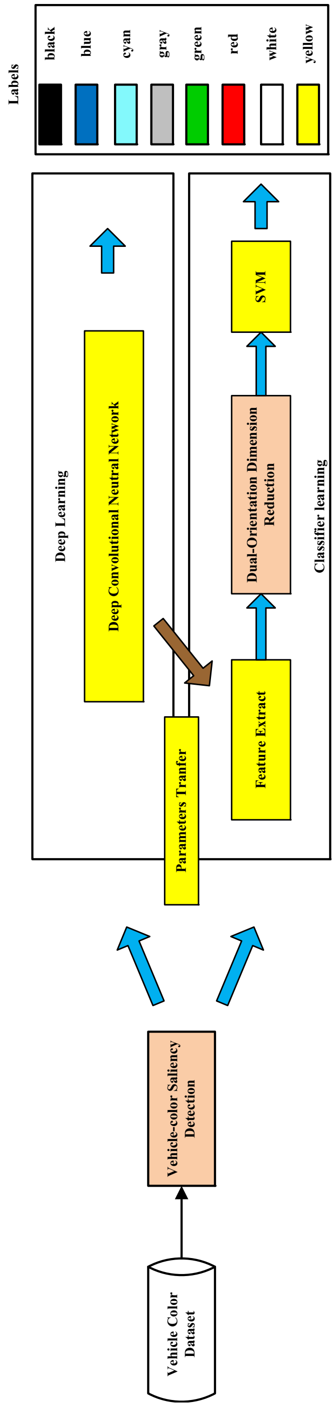

Fig. 1 Framework of the proposed method for vehicle color recognition

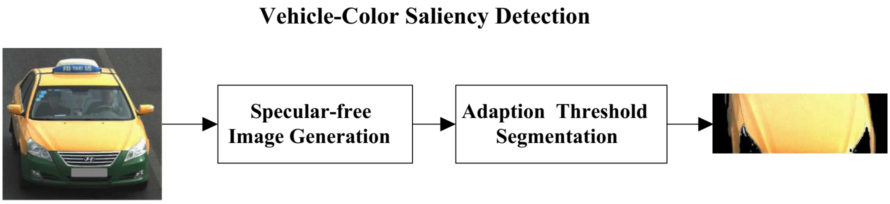

Fig. 2 The flowchart of vehicle-color saliency detection

constant while retaining their hues [18]. The specular-free image I^{spec} can be generated according to Eq. (1):

$$I^{spec}_{r(g,b)} = I_{r(g,b)} - \frac{\sum_{r,g,b} I_i - \frac{\tilde{I}(3\tilde{C}-1)}{\tilde{C}(3\Lambda-1)}}{3} \tag{1}$$

where $\tilde{I} = \max(I_r, I_g, I_b)$, $\tilde{C} = \frac{\tilde{I}}{I_r+I_g+I_b}$ and $[I_r, I_g, I_b]$ denotes RGB values of a pixel. The parameter Λ is set to 0.6. Each pixel value of I^{spec} can be obtained, and thus, Specular-free image can be generated. T_h is a pre-set threshold to extract the chromatic areas. When the pixel value of I^{spec} in any component of the original image is larger than T_h, this pixel will be treated as a chromatic pixel.

Figure 3 shows the comparison between a specular-free image and the original image. It can be easily seen that, in the specular-free image, non-color parts are all removed and the vehicle color has been highlighted, which will facilitate to improve the color recognition performance.

2.1.2 Adaption Threshold Segmentation

Based on the generated Specular-free image, an adaption threshold segmentation method is then used to determine the vehicle color regions using horizontal projection. The segmentation can be performed using Eq. (2):

Fig. 3 Comparison of specular-free image and the vehicle-color saliency detection results

$$r_i^s = \begin{Bmatrix} remove, & h_i < T \\ r_i^o, & h_i \geq T \end{Bmatrix},$$
$$here,\ h_i = \sum_{j=1}^{N} c_{ij}, \quad T = \lambda * \max(h_i), \tag{2}$$
$$i = 1, 2, \ldots, M, \quad j = 1, 2, \ldots, N$$

where r_i^o represents the pixel value of ith row of the original image, r_i^s represents the pixel value of segmented color regions, and λ represents the chromaticity information amount control factor, which is set as 0.8 in this paper. When the pixel at (i, j) point contains color information, $c_{ij} = 1$, otherwise $c_{ij} = 0$. According to this method, the vehicle-color saliency region can be detected. As shown in Fig. 3, it can be seen that the saliency region fully contains vehicle color information.

2.2 Dual-Orientational Dimensionality Reduction

Since 2012, deep learning has been investigated comprehensively and achieved outstanding performance in face recognition, image classification, and other fields, which has aroused great research enthusiasm. Learning features from large-scale data using deep learning has become a feasible approach for breaking through the limitations of manual design features. Deep Belief Net (DBN) [19, 20], CNN [14] and other methods have been proposed and applied in various fields. Deep learning is able to discover multilayer features from big data. High-level features can characterize the deeper nature of the data, which is hard for the conventional handcrafted features and shallow learning methods. However, deep learning will generate high-dimensional feature vector in the high layer. To avoid the problem of "curse of dimensionality", a dual-orientational dimensionality reduction strategy has been proposed to reduce the dimensionality of CNN deep features in this paper, thus improving the classification efficiency.

2.2.1 *Deep Feature Learning With CNN*

CNN is one of the most commonly used deep neural network architecture derived from LeNet-5 [21], which is composed of stacked convolution layers and its optimization layers, contrast normalization layer, pooling layer, and one or more fully-connected layers at the end of the architecture. The architecture avoids the extraction process of traditional manual features and can process the original images directly. The recognition results can be obtained via convolution feature extraction and mapping. It is also efficient and robust for vehicle color recognition under complex environments.

In recent years, several network variations of CNN, such as Alex-Net [14], GoogLeNet [22] and VGG-Net [23], etc. have been proposed and achieved excellent performance in image classification, image recognition and other fields. In this paper, Alex-Net is adopted to learn deep features for our task. Alex-Net takes RGB images with the size of 227 × 227 × 3 as input. The sizes of the output feature map of all the convolution layers are $55 \times 55 \times 96$, $27 \times 27 \times 256$, $13 \times 13 \times 384$, $13 \times 13 \times 384$, and $13 \times 13 \times 256$, respectively. Among all of these features, the smallest feature dimensionality is 43,264 while the largest is 290,400. The high-dimensional feature vector of Alex-Net brings heavy storage and computational burden on the following processing. Therefore, in this paper, a dual-orientational dimensionality reduction strategy has been proposed, which use 2D-PCA to reduce the dimensionality of the deep features.

2.2.2 Dual-Orientational Dimensionality Reduction of Deep Features

Dimensionality reduction is to preserve or strengthen some properties of the data by linear or non-linear mapping from high-dimensional space to low-dimensional space. There are several popular dimensionality reduction methods, such as Principal Component Analysis (PCA) [24], Locality Preserving Projections (LPP) [25], Linear Local Embedding (LLE) [26], SLE (Supervised Laplacian Eigen maps) [27] and SRE [28]. The convolution feature is composed of many sub-features (*i.e.* feature map) and the number of sub-features is equal to that of output neurons. The size of each sub-feature is decided by the convolution kernel size and parameters setting of current layer. 2D-PCA algorithm is adopted in this paper to reduce the dimensionality of deep features. Figure 4 shows the dual-orientational dimensionality reduction process, where each color represents a feature map, and the entire convolution feature consists of multiple feature maps. 2D-PCA algorithm is used to reduce the horizontal and vertical dimensionality of feature maps respectively. The redundancy of each feature map can be removed. The implementation process will be described as follows.

Firstlylet X denote an *n*-dimensional unitary column vector. The idea of 2D-PCA [16] algorithm is to project image A, an $m \times n$ random matrix, onto X by the following linear transformation.

$$Y = AX \tag{3}$$

In this paper, the feature map is denoted as $\boldsymbol{A}$, and an *m*-dimensional projected vector as $\boldsymbol{Y}$, which is called the projected feature vector of feature map **A**. A family of projected feature vectors, $[Y_{1,}\ldots, Y_{\mathrm{D1}}]$ can obtained using Eq. (3), which are called the principal component (vectors) of the feature map $\boldsymbol{A}$. The principal

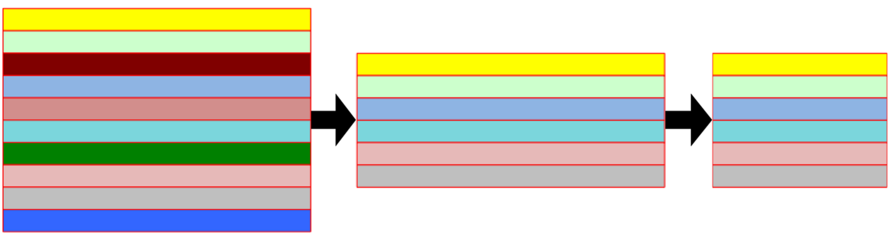

Fig. 4 Processing of dual-orientational dimensionality reduction

component vectors obtained are used to form an $m \times d$ matrix F_1, which is the low-dimensional feature of the feature map $\boldsymbol{A}$.

Secondly, 2D-PCA algorithm is applied to the second dimensionality reduction to obtain the final low-dimensional feature. This process can be expressed by Eq. (4):

$$Y' = (F_1^T X')^T \tag{4}$$

where $\boldsymbol{T}$ denotes the transpose operation of the matrix, $\boldsymbol{X'}$ represents the best projection vector when 2D-PCA transformation is performed on $\boldsymbol{F}_1$. The final low-dimensional feature $\boldsymbol{F}_2$ is obtained by selecting the first $\boldsymbol{D}_2$ component with the largest eigenvalue of $\boldsymbol{Y'}$, and the feature size of $\boldsymbol{F}_2$ *is* $\boldsymbol{D_2} \times \boldsymbol{D_1}$.

Specifically, take the feature map of the fifth convolution layers (C5) for instance. The number of output neurons in this layer is 256 and the feature size is 13×13, *i.e.* the size of sub-feature is 169. 2D-PCA is performed on the 2D feature map to generate a low-dimensional 2D feature vector, which the feature size has reduced to $169 \times \mathrm{D}_1$. Next, 2D-PCA is used again to further reduce the dimensionality to obtain the final $\mathrm{D}_2 \times \mathrm{D}_1$-dimensional feature. And linear SVM is exploited to train the vehicle color classification model. The features fed to SVM are the dimensionality reduced CNN deep features mentioned in Sect. 2.2. The best configuration of SVM parameters can be tuned by cross validation.

3 Experimental Results and Analysis

To verify the effectiveness of the proposed vehicle color recognition method, the experiments are conducted on the dataset constructed by Chen et al. [9]. The experimental results are compared with the state-of-the-arts vehicle color recognition methods. The comparative experiments are performed on the platform, which is set as following: 3.3-GHZ 4-core CPU, 16 GB-RAM, Tesla-K20C GPU, and Ubuntu 64-bit operating system.

3.1 Dataset

Vehicle Color dataset provided by Chen et al. [9] is one of the commonly used dataset for comparative experiments. The dataset contains 15,601 images of vehicles, including eight colors: black, blue, cyan, gray, green, red, white, and yellow. Some examples are shown in Fig. 5. There are 282 cyan vehicle images with the minimal proportion, and 4743 white vehicles with the maximal proportion. The images of the dataset are collected in the front view from the equipment installed on the urban roads (or less angle changed), and each image contains only a vehicle. The dataset is characterized by varying environments (such as illumination, weather, and so on) and contains a variety of vehicle types, such as trucks, cars, buses, etc., which brings greater challenges to rightly identify the color of the vehicles. In the experiments, for fair comparison, the settings are the same as those in [9, 12, 13], and the dataset is divided into training data and test data randomly at a ratio of 1: 1.

Fig. 5 Examples of the vehicle color dataset

3.2 Implementation Details

In this paper, ILSVRC-2012 [29] dataset is used for pre-training using Alex-Net. ILSVRC-2012 dataset contains a subset of the large hand-labeled ImageNet dataset (10,000,000 labeled images depicting 10,000 + object categories). Alex-Net is implemented on Caffe (Convolutional Architecture for Fast Feature Embedding) platform [30]. The output feature map of the fifth convolution layer is taken out to use as the image feature. Its size is 169×256. The linear SVM is used as the classifier. The experimental results are given and analyzed in the following sections.

3.3 Vehicle-Color Saliency Detection

To validate the impact of vehicle-color saliency detection on the vehicle color recognition performance, in this paper, the experiments are conducted using the methods with or without vehicle-color saliency detection respectively. Without vehicle-color saliency detection, the features of the original images are directly extracted from feature map of CNN. Table 1 shows the comparison results of recognition performance on vehicle color dataset. It can be seen from Table 1 that, vehicle-color saliency detection can facilitate to improve the recognition accuracy

Table 1 Performance of vehicle-color saliency detection

Image	Black	Blue	Cyan	Gray	Green	Red	White	Yellow	AP
Origin image	0.9570	0.9650	0.9928	0.8391	0.8467	0.9886	0.9443	0.9689	0.9378
Local image	0.9762	0.9719	0.9929	0.8827	0.8008	0.9848	0.9522	0.9897	0.9439

by about 0.61% on average. And the recognition accuracy is up to 94.39%. The experimental results demonstrate that the performance of the proposed vehicle-color saliency detection has positive effect on the improvement of vehicle color recognition performance.

3.4 Dimensionality Reduction

In order to verify the impact of the proposed dual-orientational dimensionality reduction strategy on the performance of vehicle color recognition, the experiments are conducted and the experimental results are shown in Table 2, where six sets of experimental results are listed. Each experiment is conducted with different dimensionalities. Without dual-orientational dimensionality reduction, the recognition accuracy is up to 94.39%. But if using dual-orientational dimensionality reduction, when the dimensionalities are reduced to be 12 and 128 respectively, the comparable recognition accuracy can be achieved. Even more, the recognition accuracy can be slightly improved when the dimensionality is 16 $\times$ 110. And the recognition accuracy is up to 94.86%. The accuracy can be improved about 0.47%. Figure 6 shows the example results of dimensionality reduced feature map of C5 layer.

Although the proposed algorithm works well under various challenging conditions, it is far from perfect. It might make mistakes or give wrong predictions in certain cases. As shown in Fig. 7, a majority of the incorrect predictions are caused by various illuminations, indistinguishable colors and occlusions. This means that there is still room for further improvement in vehicle color recognition.

3.5 Performance Comparison with State-of-the-Arts Methods

In order to verify the recognition performance of the proposed method, we compare it with the state-of-the-arts methods on the same dataset. Three methods are used for comparison. In Chen et al. [9] combined the different histogram features extracted in several color spaces, then used BoW model and SVM to achieve the goal of promoting vehicle color recognition. Rachmadi et al. [12] applied a parallel cross

Table 2 Performance of dual-orientational dimensionality reduction

Method		Black	Blue	Cyan	Gray	Green	Red	White	Yellow	AP
D_2	D_1									
169	256	0.9762	0.9719	0.9929	0.8827	0.8008	0.9848	0.9522	0.9897	0.9439
169	128	0.9657	0.9521	0.9714	0.8785	0.8672	0.9897	0.9258	0.9724	0.9404
48	128	0.9727	0.9669	0.9571	0.8359	0.8548	0.9897	0.9490	0.9759	0.9377
24	128	0.9843	0.9724	0.9714	0.8549	0.8589	0.9907	0.9540	0.9862	0.9466
12	128	0.9750	0.9711	0.9857	0.8603	0.8401	0.9928	0.9571	0.9716	0.9443
16	**110**	**0.9942**	**0.9797**	**0.9786**	**0.8332**	**0.8631**	**0.9742**	**0.9861**	**0.9793**	**0.9486**

The significance of bold is an emphasis on the test results which dimension (16 $\times$ 110) is the best for the average prediction accuracy

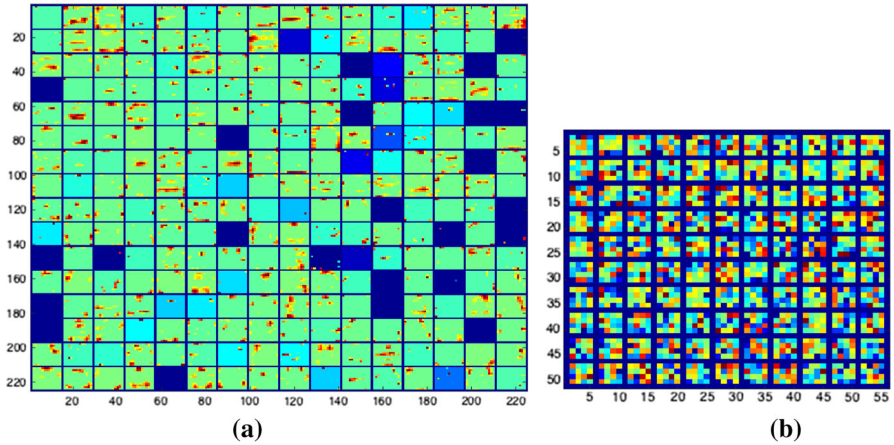

Fig. 6 Example results of dimensionality reduced feature map of C5 layer. **a** Feature maps visualization of C5 where size is 169 $\times$ 256, **b** Dimensionality Reduced feature maps with the size of 16 $\times$ 110. Here are a total of 110 neurons, each neuron (*small rectangle*) size is 16

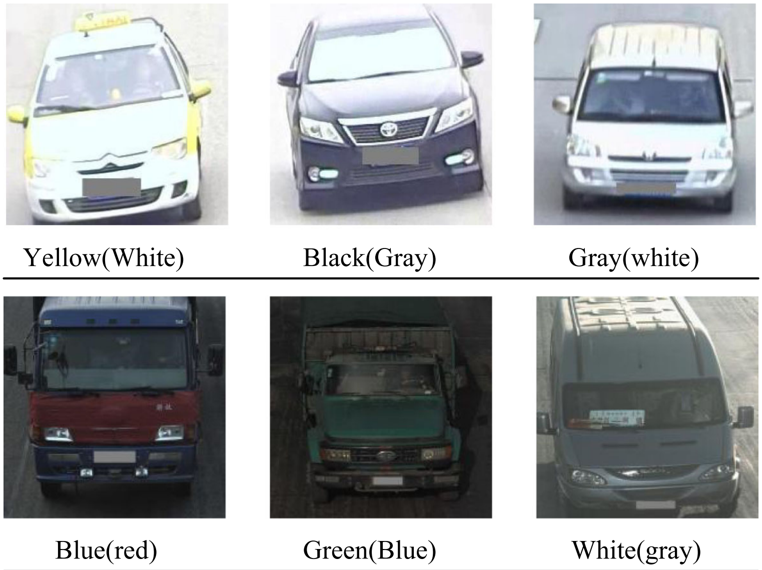

Fig. 7 Failure cases. *Word before bracket* ground truth. *Word in bracket* predicted color type

CNN model for vehicle color recognition. Hu et al. [13] utilized the deep features learn from Alex-Net, which are extracted from the original image and four sub-images of SPM partitioning, and kernel SVM to perform vehicle color recognition.

The experimental results using four methods are shown in Table 3. Compared with other three methods, the proposed method can significantly improve the recognition performance and achieve state-of-the-art recognition accuracy. The

Table 3 Comparison results of recognition performance using different methods

Method	Black	Blue	Cyan	Gray	Green	Red	White	Yellow	AP
Baseline [9]	0.9713	0.9451	0.9787	0.8461	0.7834	0.9876	0.9414	0.9457	0.9249
Parallel CNN [12]	0.9738	0.9410	0.9645	0.8608	0.8257	0.9897	0.9666	0.9794	0.9447
Deep Feature + SPM + kernel SVM [13]	0.9796	0.9642	0.9886	**0.8686**	0.8406	**0.9926**	0.9619	0.9787	0.9469
ur Proposed Method (linear SVM)	**0.9942**	**0.9797**	0.9786	0.8332	0.8631	0.9742	**0.9861**	**0.9793**	**0.9486**

The significance of bold is an emphasis on the test results of each sub-class which is the best for the average prediction accuracy recognition accuracy reaches 94.86%, and increasing by at least 0.17% over the other three methods.

In addition, compared with the method proposed by Hu et al. [13], except for slightly unsatisfying recognition on the vehicle colors of blue and gray, our method has obvious advantages on other six vehicle colors. However, it should be noted that the recognition accuracy of our proposed method is achieved when the dimensionality of the feature is only 1760, whereas the dimensionality of the features used in Ref. [13] is 216,320. The dimensionality of the features we adopt is only 0.8136% of that in Ref. [13]. It demonstrates that the proposed method can achieve better recognition accuracy with the low-dimensional features and higher efficiency.

4 Conclusion

In this paper, a high accuracy vehicle color recognition method with vehicle-color saliency detection and dual-orientational dimensionality reduction of deep features is proposed. Compared with the conventional methods, the proposed method shows outstanding performance on recognition accuracy and speed. Meanwhile, in a relatively complex environment, the proposed method can achieve better recognition accuracy and robustness on vehicle color recognition. In the following research, we will further study the influence of different dimensionality reduction methods on the recognition performance. The method will be applied to practical Intelligent Traffic System to further demonstrate its performance.

Acknowledgements The work in this paper is supported by the National Natural Science Foundation of China (Nos. 61531006, 61372149, 61370189, 61471013, and 61602018), the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (Nos. CIT&TCD20150311, CIT&TCD201404043), the Beijing Natural Science Foundation (Nos. 4142009, 4163071), the Science and Technology Development Program of Beijing Education Committee (Nos. KM201410005002, KM201510005004), Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality.

References

1. Gao, Y., & Lee, H. J. (2016). Local tiled deep networks for recognition of vehicle make and model. *Sensors, 16*(2), 226.
2. Chen, L. C., Hsieh, J. W., Yan, Y., & Chen, D. Y. (2015). Vehicle make and model recognition using sparse representation and symmetrical SURFs. *Pattern Recognition, 48*(6), 1979–1998.
3. More, N., & Tidke, B. (2015). License plate identification using artificial neural network and wavelet transformed feature selection. In *2015 IEEE international conference on pervasive computing (ICPC)* (pp. 1–5).
4. Wang, Y. C., Han, C. C., Hsieh, C. T., & Fan, K. C. (2014). Vehicle color classification using manifold learning methods from urban surveillance videos. *EURASIP Journal on Image and Video Processing, 2014*(1), 1.
5. Chen, R., Hawes, M., Mihaylova, L., Xiao, J., & Liu, W. (2016). Vehicle logo recognition by spatial-sift combined with logistic regression. In *Proceedings of fusion 2016*.
6. Huttunen, H., Yancheshmeh, F. S., & Chen, K. (2016). Car type recognition with deep neural networks. arXiv preprint arXiv:1602.07125.

7. Baek, N., Park, S. M., Kim, K. J., & Park, S. B. (2007). Vehicle color classification based on the support vector machine method. In *International conference on intelligent computing* (pp. 1133–1139). Berlin: Springer.
8. Kim, K. J., Park, S. M., & Choi. Y. J. (2008). Deciding the number of color histogram bins for vehicle color recognition. In *Asia–Pacific services computing conference, 2008. APSCC'08, IEEE* (pp. 134–138).
9. Chen, P., Bai, X., & Liu, W. (2014). Vehicle color recognition on urban road by feature context. *IEEE Transactions on Intelligent Transportation Systems, 15*(5), 2340–2346.
10. Hsieh, J. W., Chen, L. C., Chen, S. Y., & Chen, D. Y. (2015). Vehicle color classification under different lighting conditions through color correction. *IEEE Sensors Journal, 15*(2), 971–983.
11. Hu, W., Yang, J., Bai, L., & Yao, L. (2013). A new approach for vehicle color recognition based on specular-free image. In *Sixth international conference on machine vision (ICMV 13), international society for optics and photonics* (pp. 90671Q–90671Q-5).
12. Rachmadi, R. F., & Purnama, I. (2015). Vehicle color recognition using convolutional neural network. arXiv preprintarXiv:1510.07391.
13. Hu, C., Bai, X., Qi, L., Chen, P., Xue, G., & Mei, L. (2015). Vehicle color recognition with spatial pyramid deep learning. *IEEE Transactions on Intelligent Transportation Systems, 16*(5), 2925–2934.
14. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
15. Lazebnik, S., Schmid, C., & Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *2006 IEEE computer society Conference on computer vision and pattern recognition* (Vol. 2, pp. 2169–2178).
16. Yang, J., Zhang, D., Frangi, A. F., & Yang, J. Y. (2004). Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 26*(1), 131–137.
17. Shafer, S. A. (1985). Using color to separate reflection components. *Color Research and Application, 10*(4), 210–218.
18. Tan, R. T., & Katsushi, I. (2005). Separating reflection components of textured surfaces using a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(2), 178–193.
19. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507.
20. Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation, 18*(7), 1527–1554.
21. LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Back propagation applied to handwritten zip code recognition. *Neural Computation, 1*(4), 541–551.
22. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).
23. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
24. Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics, 15*(2), 265–286.
25. Niyogi, X. (2004). Locality preserving projections neural information processing systems. *MIT, 16,* 153.
26. Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science, 290*(5500), 2323–2326.
27. Raducanu, B., & Dornaika, F. (2012). A supervised non-linear dimensionality reduction approach for manifold learning. *Pattern Recognition, 45*(6), 2432–2444.
28. Timofte, R., & Van Gool, L. (2011). Sparse representation based projections. In *Proceedings of the 22nd British machine vision conference-BMVC 2011* (pp. 61.1–61.12). BMVC Press.
29. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision, 115*(3), 211–252.
30. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., & Darrell, T. (2014). Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on multimedia* (pp. 675–678).