Vehicle Color Recognition in The Surveillance with Deep Convolutional Neural Networks

Boyang Su, Jie Shao, Jianying Zhou, Xiaoteng Zhang, Lin Mei*
Internet of things technology department
The Third Research Institute of the Ministry of Public Security
Shanghai, P.R. China.
uestcsby@163.com*

Keywords: Vehicle Color Recognition, Deep Convolutional Nerual Networks, network in network.

Abstract.Vehicle information extraction is the key means in Intelligent Transportation System (ITS). Color plays an important role in vehicle recognition. The main challenge of vehicle color recognition is to find the dominant color. In this paper, we propose a color recognition method using convolutional neural network. We train the classifier with the network structure 'NIN' to increase the classification accuracy. The experiments are validated on our dataset and extra data, which are collected from city surveillance equipment. The proposed method outperforms other competing color recognition methods.

Introduction

Since the explosive growth of information, people receive lots of messages in every hour and moment. However, the news become so big for human to deal with now. So people analysis the mass data with the help of intelligent system. A famous application of intelligent system is Intelligent Transportation System (ITS).

In ITS, people always focus on the vehicle information first, for the personal information of vehicles are very useful for video surveillance and many aspects of public security. The user may need to know all kinds of vehicle information such the color, the type, the plate number, and the annual inspection marks etc. Among these information color is the most dominant cues for vehicle information. Vehicle color classification/recognition in natural scenes can provide useful information in vehicle detection, vehicle tracking and vehicle retrieval.

Of all image features, color represents one of the most widely used visual features, while it is one fundamental characteristic of the contents of all images. As its invariance on size, orientation and complexity, color is also an intuitive feature. To process color images, an appropriate color space is needed. Since the models of human perception of color differences are described in the form of color spaces [1], so the research on color image must be studied in a given color space, RGB, YIQ, YUV, HSV etc. Next, color histogram is employed to represent the distribution of colors in an image. The histogram shows the numbers of pixels of the colors fails in a certain range. If the set of color values is sufficiently small, each of those colors may be placed on a range by itself. The comparison between query image and image in database is achieved by the use of some measure which identifies the distance or similarity between the two histograms. The color histogram is available on any kind of color space.

However, traditional color feature still meets many difficulties:

1. Color feature is easily influenced by the change of naturel environment. Strong light makes the colors in close color spaces hard to distinguish. While the dim environment may

cause all colors blur. The weather change can also cause a visible color variation.

2. The quality of color recognition is limited by the quality of images and videos, such as low resolution, noise, overexposure, jitters, etc. These factors may cause color shift or even color change.
3. The vehicle body may have more than one color, the color recognition may mislead by some conspicuous parts. Parts with light colors such as red, yellow and green are more easily to be considered as main objects, so that mislead to a wrong recognition result.

Deep convolutional neural network have achieved a great progress in the field of computer vision, and other fields. Various work on convolutional neural networks (CNN) (Simonyan et al., 2013 [2], Girshich et al., 2014 [3], Chatfield et al., 2014a [4], Zeiler & Fergus, 2014 [5] have successfully demonstrated its powerful representation on visual recognition as well as a good feature descriptor for all kinds of features. So we believe that deep learning can perform better than color space in color recognition.

In this paper, we focus on the color classification task, which is a very important part of intelligent transportation system (ITS) or Smart City. Our goal is to perform a precise recognition that is qualified for the standard eight colors, as well as the extension colors from the standard eight colors. We train several models and test their color recognition capacities, in order to find the most suitable classification structure for color. Our model performance is measured by the rate of different color types, and the average rate of these colors. We also compare our model with other color descriptors. As the experimental result shown below, our model is robust and precise enough on illuminance and environment variation.

Related work and proposed method

In order to find the best presentation for color features, we build several DL models to compare their color classification abilities.

AlexNet. AlexNet is the fundamental and classical network for classification. While it can be considered as a mixed descriptor of image global and local features. An AlexNet has seven layers which contains five convolutional layers and two full-connected layers. Now many models are built referring to Alexnet structure, so we set AlexNet as a benchmark in our experiment

GoogleNet. GoogleNet was announced by google in ILSVRC-2014 [6]. The network's biggest characteristic is to promote the utilization of computing resources. Under the premise of the changeless of network's calculated amount, GoogleNet use a deeper and bigger network structure, which has 22 layers. However, the amount of parameters is only one twelfth of AlexNet, but the accuracy is higher than it of AlexNet. GoogleNet uses NIN (introduce later) to increase the power of the neural network. Since simply increase the network structure may lead to overfitting, as well as the consumption of computing performance, GoogleNet uses many 'Inception' to solve this problem.

Network In Network. Network In Network (NIN) [7] is actually a network structure which can enhance the model discriminability. The traditional convolutional layers use linear filters followed by a nonlinear activiation function to scan the input. NIN use more complex structures (micro neural network) instead to abstract the data. Deep NIN can be implemented by stacking mutiple of the above described structure. With enhanced local modeling via the micro network, we are able to utilize global average pooling over feature maps in the classification layer, which is easier to interpret and less prone to overfitting than traditional fully connected layers.

Experiment Result

Dataset. Since there is no ready-made public vehicle datasets for color recognition, we built our own dataset. The images are collected from the HD bayonets, cropped from the surveillance videos and resized to 256*256. The dataset is consist of 15016 vehicle images with different types, such as car, bus, SUV and truck. We set eight main colors for these images, for this eight types of color are the most common appearance of cars on the road.

We separate this dataset by the proportion of ten to one as training data and test data. We also use additional test data on our best model.

Evaluation on Image Data set. We train three models which based on Alexnet, Googlenet and Colornet respectively. C1 uses Alexnet structure, C3 has the same structure as Googlenet. C9 is named as Colornet, which is an eight layers network and sets NIN as an addition means. The difference between C3 and C9 is the deep and size of the network. C9 has far less layers than C3. However, with almost the same performance, C9 can save lots of computing resource and time, and its model is smaller.

C9, add* is tested with the additional test data. There are wrong labeled images in the additional data set, so we fine-tuned our model with part of this data and re-tested with remained images, the performance significantly improved. We didn't test C1 on each color, for its low performance of average accuracy.

Table 1. Color recognition performance on image dataset. Each row lists the performance of a certain model, columns show the recognition rate of different color types.

Method	Black	Blue	Gray	Green	Red	Cyan	White	Yellow	Average
C1	-	-	-	-	-	-	-	-	0.7880
C3	0.9495	0.9810	0.9535	0.9455	0.9615	0.9394	0.9286	0.9583	0.9522
C9	0.9220	0.9857	0.9651	0.9636	0.9846	0.9091	0.9847	0.9444	0.9574
C9, add*	0.9094	0.9227	0.3874	0.8465	0.8864	0.7801	0.9618	0.9829	0.8347
C9,add*-ft	0.9355	0.9761	0.8490	0.8797	0.9938	0.9858	0.9456	0.9622	0.9410

We also set a comparison between different color recognition algorithms. Table 2 shows our outperformance on color recognition with the state-of-art strategies.

Table 2. Performance comparisons using different strategies.

	Method	Black	Blue	Gray	Green	Red	Cyan	White	Yellow	Average
Original Global Features	Layered Color Indexing [8]	0.8937	0.6998	0.6776	0.6058	0.9640	0.9291	0.9009	0.9038	0.8218
Original Local Features	Opponent Hist [9]	0.7735	0.4955	0.1501	0.5217	0.8910	0.7163	0.8352	0.7412	0.6406
BoW based Features	RGB Hist	0.7737	0.7209	0.5450	0.6149	0.9371	0.7623	0.7994	0.8160	0.7462
BoW based Features (SPM [10])	Transformed Color Hist[9]	0.9391	0.8323	0.7400	0.5013	0.9638	0.7971	0.9323	0.8609	0.8209
BoW based Features (FC [11])	Color Hist[12]	**0.9714**	0.9363	0.8218	0.7859	**0.9885**	**0.9660**	0.9415	0.9395	0.9189
CNN	Our Model	0.9220	**0.9857**	**0.9651**	**0.9636**	0.9846	0.9091	**0.9847**	**0.9444**	**0.9574**

Conclusion

This paper shows our study on color recognition. Powerful deep learning model can achieve better performance than traditional hand-craft feature descriptors. Through our study, we find that convolutional neural network is robust on the change of naturel environment, and can eliminate noise interference.

Acknowledgement

This work was supported in part by the National Science and Technology Major Project under Grant 2013ZX01033002-003, in part by the National High Technology Research and Development Program of China (863 Program) under Grant 2013AA014601, 2013AA014603, in part by National Key Technology Support Program under Grant 2014JSYJB006, and the project of Shanghai Municipal Commission of Economy and Information under Grant 12GA-19.

Reference

[1] Datta Ritendra, LiJia, Wang James Z. "Content-based image retrieval-Approaches and trends of the new age[A]," The 7th International Workshop on Multimedia Information Retrieval, in Conjunction with ACM International Conference on Multimedia[C], Singapore, 2005, 253-262.

[2] Simonyan, Karen, Vedaldi, Andrea, and Zisserman, Andrew. Deep inside convolutional networks: Visualising image classification models and saliency maps. 2013

[3] Girshick, Ross B., Iandola, Forrest N., Darrel, Trevor, and Malik, Jiendra. Deformable part models are convolutional neural networks. CoRR, abs/1409.5403, 2014

[4] Chatfield, Ken, Simonyan, Karen, Vedaldi, Andrea, and Zisserman, Andrew. Return of the devil in the datails: Delving deep into convolutional nets. CoRR, 2014a.

[5] Zeiler, Matthew D. and Fergus, Rob. Visualizing and understanding convolutional networks. In ECCV, pp. 818-833, 2014.

[6] Christian Szegedy, Wei Liu, Yangqing Jia and Andrew Rabinovich. Going deeper with convolutions. In *arXiv*:1409. 4842

[7] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *CoRR*, abs/1312.4400, 2013.

[8] G. Qiu and K. Lam, "Spectrally layered color indexing," in *Image and Video Retrieval*. Berlin, Germany: Springer-Verlag, 2002, pp. 100-107.

[9] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog*., 2006, vol. 2, pp. 2169-2178.

[10] X. Wang, X. Bai, W. Liu, and L. J. Latecki, "Feature context for image classification and object detection," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog., Jun. 2011, pp. 961-968.

[11] K. E. Van De Sande, T. Gevers, and C. G. Snoek, "Evaluating color descriptors for object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 32, no. 9, pp. 1582-1596, Sep. 2010.

[12] Pan Chen, Xiang Bai, and Wenyu Liu. Vehicle Color Recognition on Urban Road by Feature Context. In *Intell Trans Sys IEEE Trans*, 2014, 15(5): 2340-2346.