Vehicle Color Recognition Using Deep Learning for Hazy Images

Aarathi K S, Anish Abraham
Computer Science and Engineering
Govt. Engineering College Thrissur
Kerala, India

Abstract— **Recent years the number of vehicles increases tremendously. Because of that to identify the vehicle is significant task. Vehicle color and number plate recognition are various ways to identify the vehicle. So Vehicle color recognition essential part of an intelligent transportation system. There are several methods for recognizing the color of the vehicle like feature extract, template matching, convolutional neural network (CNN), etc. CNN is emerging technique within the field of Deep learning. The survey concludes that compared to other techniques CNN gives more accurate results with less training time even for large dataset. The images taken from roads or hill areas aren't visible because of haze. Consequently, removing haze may improve the color recognition. The proposed system combines both techniques and it adopts the dark channel prior technique to remove the haze, followed by feature learning using CNN. After feature learning, classification can be performed by effective classification technique like SVM.**

Keywords— Convolutional Neural Network, Deep Learning, Dark channel prior, Intelligent transportation system

I. INTRODUCTION

Vehicle data recognition is a key part of Intelligent Transportation Systems [1]. Vehicle data recognition include vehicle color recognition, number plate recognition etc. Here we tend to discuss mainly regarding vehicle color recognition. Color plays a very important role in vehicle identification. Vehicle color recognition in natural scenes will offer helpful data in vehicle detection, vehicle tracking and automatic driving system. However, identifying vehicle color in uncontrolled environments could be difficult task. Vehicle color recognition can be treated as supervised learning problem consists of preprocessing, feature extraction of images and classification of resulting features along with labels. The input images of the system are the images taken from roadside cameras. Sometimes the images taken in roadside cameras don't seem to be visible as a result of haze, snow etc. Removal of the haze from the image using haze removal techniques is major step for achieving good results. There are many haze removal algorithms dark channel prior, edge detection, and color attenuation prior etc. Here we are using dark channel prior algorithm for haze removal.

Deep learning [2] is a part of neural network that accepts an input and passes on a changed version of input to consecutive layer. There are number of layers. In between input layer and output layer, these layers are called hidden layers. Every layer finds out more complex features than previous layer. There are many deep learning architectures convolutional neural network (CNN), deep belief networks, continual neural networks etc.. Convolution Neural Network is a kind of feed forward artificial neural network, they're like standard neural network. They're created from neurons that have learn-able weights and biases. Every neuron receives some inputs and performs some operation. There are three kinds of layers in CNN, convolution layer, pooling layer and fully connected layer.

- Convolution Layer: It will calculate the output of neurons that are connected to local regions within the input, each compute a dot product between the weights and biases.
- Pooling Layer: max pooling, avg pooling, sum pooling etc. are several pooling layer operations. Commonly use max pooling.
- Fully connected layer: Each neuron in this layer connected to previous layer neurons.

CNN is a new rising technique within the field of deep learning. It contains an additional number of layers and every layer identifies more complex features and it provides a more accurate result.

Proposed system consists of identification of color of the vehicle using the convolutional neural network from haze and haze free images.. Then haze free images pass to convolutional neural network for feature learning followed by classification by Support Vector Machine.

II. RELATED WORKS

Vehicle info Recognition is a major part of Intelligent Transportation System (ITS). Vehicle information recognition contains Automatic Number Plate Recognition [3], [4] and Vehicle Color Recognition. Here we discuss about Vehicle Color Recognition. Different methods for vehicle color recognition are described below.

TABLE I
ANALYSIS OF EXISTING APPROACHES

Author	Algorithm Used	Accuracy (%)
Chuanping Hu	CNN	94.6
Pan Chen	Feature Context	90.68
Wang et al	Template matching	86.4
Dule et al.	ROI	84.8
Sermanet and LeCun	CNN	92.6

Chuanping Hu, Xiang Bai[5], proposed an approach for vehicle color recognition using deep learning. During this system CNN is adopted as the feature extractor, that outputs a vector for each input vehicle image and SVM is used as the classifier, which predicts the color of the vehicle. First the image is rescaled into an appropriate size. In training section, the weights and biases of the network are updated by using backpropagation algorithm. These trained parameters are stored in the network. Finally trained result used for classification. SVM classifies the output using the training result and input image. The input of the system is haze-free image and output of the system is that the color of the vehicle (black, white, blue...).

Pan Chen, Xiang Bai[6], proposed an approach to identify the color of the vehicle using feature context. In this method the input is taken from roadside cameras, so the input is not clear due to haze, snow etc. in this project first select the portion of the vehicle from the input image and then apply haze removal technique to remove the haze. This paper uses Dark Channel Prior algorithm for haze removal. After that features are selected from the image. These features are passing to support vector machine and SVM classifies the color of the vehicle.

During this system, CNN is adopted as Template matching is another vital sort of ways to determine the color of a vehicle. Wang et al. [7] extracted the tail region by detecting automobile lamps. The vehicle color is classified based on the distances of the values of multiple color areas between the region of interest (ROI) and those of training samples. Yang et al. [8] given a system that detects the hood and roof region supported an edge map first and matches the templates using the RGB color histogram. Since template matching strategies are comparatively sensitive to variances of car images, learning-based strategies are more acceptable for color recognition. Dule et al. [9] chosen the hood area as the ROI and then used K-NN, ANN, and SVM to train the classifier for color recognition.

In the field of intelligent transportation systems, a deep CNN also obtains the progressive performance on the task of traffic sign recognition [10], [11]. Sermanet and LeCun [12] made a network with only 2 convolutional layers and one fully-connected layer. Not like the standard CNN architectures, the outputs of the 2 convolutional layers are put together for the consecutive fully-connected layer. Ciresan et al. [13] adopted a deeper CNN [15], which leads to higher recognition accuracy.

From the above survey, the accuracy is more for CNN compared to other methods. The other methods like Feature Context, Template matching, and ROI techniques have accuracy around ninety percentage. In the case of Convolutional Neural Network accuracy nearly 95 %. Dataset of CNN is large compared to other methods. So it requires Graphical Processing Unit (GPU) for processing and it takes less time for training. Because of these reasons this paper uses CNN for feature learning and extraction.

III. METHODOLOGY

This system is to spot the color of the vehicle using the convolutional neural network for each haze and haze free images. This project is more helpful for criminal detection, roadside assistance, traffic signal monitoring etc. The input images are collected from roadside cameras, so typically it's not visible because of haze, snow etc. therefore 1st take away the haze from input images. In preprocessing stage first, remove the haze from the images. Then apply CNN for haze-free pictures. Finally, classification is done by using Support Vector Machine. The overall working of the system can be displayed in figure 1.

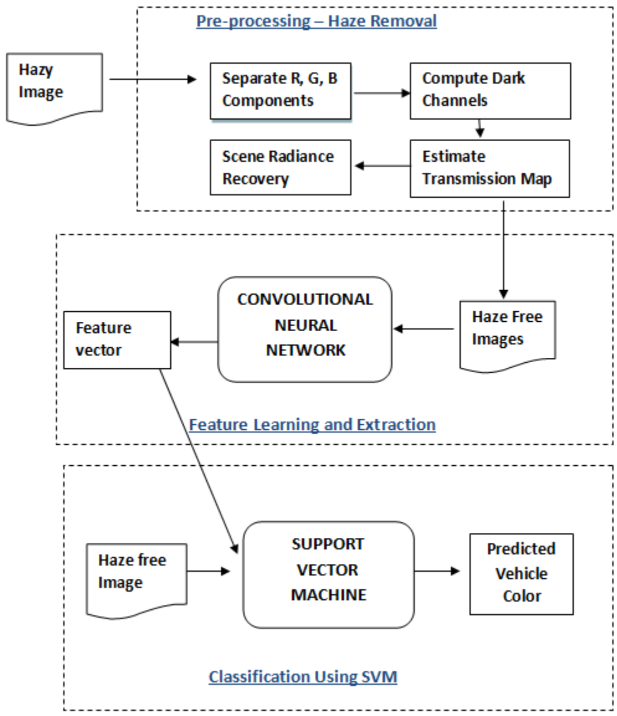

Fig. 1. Overall Architecture

A. Preprocessing - Haze Removal

Dark Channel Prior [16] is an algorithm which removes haze from images. This technique is based on the observation on haze free images, in most of the non-sky patches; a minimum of one color channel has terribly low intensity at some pixels. These points are known as dark channels.

Algorithm for Haze Removal described in Algorithm 1 The working of dark channel prior is described below. First find out the dark channels from the image. In that image, most of the non-sky patches, at least one color channel has very low intensity at some pixels. In other words, the minimum intensity in such a patch should have a very low value. Formally, for an image J, we define

$$J^{dark}(x) = \min_{c\in\{r,g,b\}} \left(\min_{y\in\Omega(x)} (J^c(y))\right) \qquad (1)$$

Algorithm 1 Haze Removal

Input: Hazy Image

Output:Haze free Image

1. Read Input vehicle Color Image
2. Divide Color Image into Red, Green and Blue components.
3. Find out Dark Channels from images.
4. Let f be a given image represented as pixel matrix. L is the various possible intensity values, often 256.
5. Assuming that the transmission in a local patch n is constant the transmission map, t is transmission map, w is local patch with minimum value of all elements.
6. Estimate Transmission map
7. Refine transmission map with a maximum value of omega
8. The airlight and the transmission map are estimated appropriately, the scene radiance can be obtained Estimate airlight and the transmission map and find out scene radiance

Where J c is a color channel of J and Ω(x) is a local patch centered at x. Our observation says that except for the sky region, the intensity of J dark is low and tends to be zero, if J is a haze-free outdoor image. We call J dark the dark channel of J, and we call the above statistical observation or knowledge the dark channel prior. Next we want to estimate the transmission map. Equation for transmission map is

$$t(x) = \min_{c}\left(\min_{y\in\Omega(x)} (I^c(y)/A^c)\right) \qquad (2)$$

Last step is Scene radiance recovery. But the direct attenuation term J(x)t(x) can be very close to zero when the transmission t(x) is close to zero. The directly recovered scene radiance J is prone to noise. Therefore, we restrict the transmission certain amount of haze are preserved in very dense haze regions. The final scene radiance J(x) is recovered by:

$$J(x) = \frac{I(x) - A)}{max(t(x), t_0)} + A \qquad (3)$$

The value of t 0 is taken as 0.1.

B. Feature Learning And Extraction

Here we are using CNN for feature extraction and learning. The architecture of convolutional neural network displayed in figure 2.

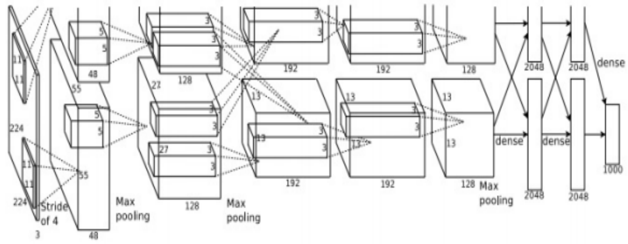

Fig. 2. CNN Alexnet Architecture

The CNN alexnet architecture consists two base networks and eight layers for every base network with total sixteen layers. The primary 2 layers of CNN architecture consists of a convolutional layer followed by normalization and pooling. The equation for convolution is

$$Z_j = \sum_{i=1toM_f} w_i X_{j+i-1} \qquad (4)$$

The convolutional layer is a layer that does convolution process that same as convolution process in image processing algorithm. The choice of activation function in the convolutional layer has a huge impact for the networks. There a several choices of activation function including tanh and ReLU (Rectified Linear Unit). In our CNN networks, we use ReLU activation function for all layers including the fully-connected layers. The last process in two first layers is the pooling process. There are two types of pooling, max pooling and mean pooling.

Equation for Max pooling is

$$P_{i,m} = \max_{1toG} q_{i,(m-1)*s} \qquad (5)$$

Equation for Mean Pooling is

$$P_{i,m} = \sum_{1toG} q_{i,(m-1)*s} \qquad (6)$$

P pooling matrix, q input matrix, G: Pooling size, s: Shift size. Each type has a different approach, max pooling will take the maximum response from the convolutional process which is the shape with sharp edges and mean pooling will take the average of the convolutional process response which summarizes the shape in the neighborhood.

Algorithm 2 Feature Learning And Extraction

Input: Haze Free Images
Output: Feature Vector

1: Read Input vehicle Color Image
2: Change the input to Fixed size 227x227x3
3: Then the image pass through 5 convolution layers, and 1st 2nd and 5th convolutions followed by pooling layer and normalization layer.
4: Equation for Convolution is

$$Z_j = \sum_{i=1toM_f} w_i X_{j+i-1} \tag{8}$$

5: Max pooling is performed in pooling layer. Equation for Max pooling is

$$P_{i,m} = \max_{1toG} q_{i,(m-1)*s} \tag{9}$$

6: Last 3 layer are fully connected Layers, all neuron in one layer connected to all neuron in another layer.
7: Last Fully connected layer gives the output feature vector.

In this paper, we use max pooling with size 3x3 and stride 2 for overlapping pooling. The second, fourth and fifth layer are grouping into two group which each group is independent each other. The third and fourth layer is also a convolutional layer but without pooling and normalization process. The output of third and fourth layer is same as input because we use the 3x3 kernel with pad 1 added for each border. The fifth layer is the convolutional layer with only pooling process without normalization. Before going into fully-connected layers, the pooling output of the fifth layer from two base networks is concatenated and flattened into one long vector. The sixth and seventh layer is a fully-connected layer employed dropout regularization method to reduce overfitting. Overall, our CNN architecture consists 2 base networks, 8 layers each with total 16 layers. The first layer use an 11x11@3 kernel with total 48 kernels, the second layer uses a 3x3@48 kernel with total 128 kernels, third use 3x3@128 kernel with total 192 kernels, fourth layer use the 3x3@192 kernel with total 192 kernels and fifth layer use 3x3@192 with total 128 kernels. The pooling process is employed in first, second, and the fifth layer with the same parameter, the pooling size of 3x3 with 2-pixel stride. Sixth, seventh, and eighth layer are the fully-connected layer with each 4096-4096- 8 neuron with dropout regularization method employed in the sixth and seventh layer. The network input is a 3 channel image with 150,228 dimensional or 227x227@3 resolution. Total neuron involved in the networks is 658,280 neurons.

$$Z_j = \sum_{i=1toM_f} w_i X_{j+i-1} \tag{7}$$

C. *Classification Using SVM*

SVM is used for classification. The response maps of the convolutional layers are reshaped to feature vectors and fed to the sequential classifier for training or testing. Selecting SVM instead of fully-connected layers is as results of two reasons: 1) the performance of SVM is best as a result of the regularization constraint can facilitate combat overfitting. The matter of overfitting is believed to be the most issue of full-connected layers. 2) The number of parameters of SVM is a smaller amount than that of fully-connected layers that create the fine-tuning procedure in training abundant easier. Except for the parameters of the fully-connected layers, that are able to be learned within the training step, the number of neurons, weight decay, and dropout ratio needs to be set manually. Once the configuration of the parameters is reset, the entire deep CNN must be retrained with a waste of time.

IV. CONCLUSION

Convolution neural network is an emerging area within the field of machine learning. Using this methodology we are able to recognize the color of the vehicle more accurately. Similarly, the input image is sometimes not fully visible as a result of haze, snow etc. Thus in this situation we got wrong results. In this project we first apply dark channel prior method for haze removal. After removing haze apply convolutional neural network approach for feature learning. Then classify the color of the vehicle using Support Vector Machine. This system gives more accurate result for hazy images. And the system is more useful for traffic controlling, car parking, criminal vehicle detection etc.

References

[1] https://en.wikipedia.org/wiki/Intelligent transportation system

[2] Sergios Theodoridis, Nueral Networks and Deep Learning 2015.

[3] Ragini Bhat, Bijender Mehandia. ” RECOGNITION OF VEHICLE NUMBER PLATE USING MATLAB.” IJIREEICE, 2014 M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

[4] Chuin-Mu Wang, Jian - Hong Liui. ”Licence Plate Recognition System” 12th International conference on fuzzy systems and knowledge Discovery 2015.

[5] Chuanping Hu, Xiang Bai, Li Qi, Pan Chen, Gengjian Xue, and Lin Mei . ”Vehicle Color Recognition With Spatial Pyramid Deep Learning.” , IEEE Transactions on Intelligent Transportation Systems,2015

[6] Pan Chen, Xiang Bai, and Wenyu Liu. ”Vehicle Color Recognition on Urban Road by Feature Context.”, IEEE Transactions on Intelligent Transportation Systems, 2014.

[7] Y.-C.Wang, C.-T. Hsieh, C.-C. Han, and K.-C. Fan, The color identification of automobiles for video surveillance, in Proc. IEEE ICCST, 2011, pp. 15.

[8] M. Yang, G. Han, X. Li, X. Zhu, and L. Li, Vehicle color recognition using monocular camera, in Proc. IEEE Int. Conf. WCSP, 2011, pp. 15.

[9] E. Dule, M. Gkmen, and M. S. Beratoglu, A convenient feature vector construction for vehicle color recognition, in Proc. 11th WSEAS Int. Conf. NN, Iasi, Romania, 2010, pp. 250255.

[10] J. Ma, J. Zhao, J. Tian, X. Bai, and Z. Tu, Regularized vector field learning with sparse approximation for mismatch removal, Pattern Recog., vol. 46, no. 12, pp. 35193532, Dec. 2013.

[11] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, Robust point matching via vector field consensus, IEEE Trans. Image Process., vol. 23, no. 4, pp. 17061721, Apr. 2014

[12] P. Sermanet and Y. Lecun, Traffic sign recognition with multiscale convolutional networks, in Proc. IEEE IJCNN, 2011,pp. 28092813.

[13] D. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, A committee of neural networks for traffic sign classification, in Proc. IEEE IJCNN, 2011, pp. 19181921.

[14] J.-W. Son, S.-B. Park, and K.-J. Kim. ”A convolutional kernel method for color recognition” International Conference on Advanced Language A. Krizhevsky, I. Sutskever, and G. E. Hinton. ”ImageNet Classification with Deep Convolutional Neural Networks” NIPS 2012, pp: 10971105.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton. ”ImageNet Classification with Deep Convolutional Neural Networks” NIPS 2012, pp: 10971105.

[16] K. He, J. Sun, and X. Tang, ”Single image haze removal using dark channel prior” in Proc. IEEE Conf. CVPR, 2009, pp. 19561963.K. Elissa, “Title of paper if known,” unpublished.