Vehicle Color Recognition with Spatial Pyramid Deep Learning

CHUANPING HU, XIANG BAI, SENIOR MEMBER, IEEE, LI QI, PAN CHEN, GENGJIAN XUE, LIN MEI

INTRODUCTION

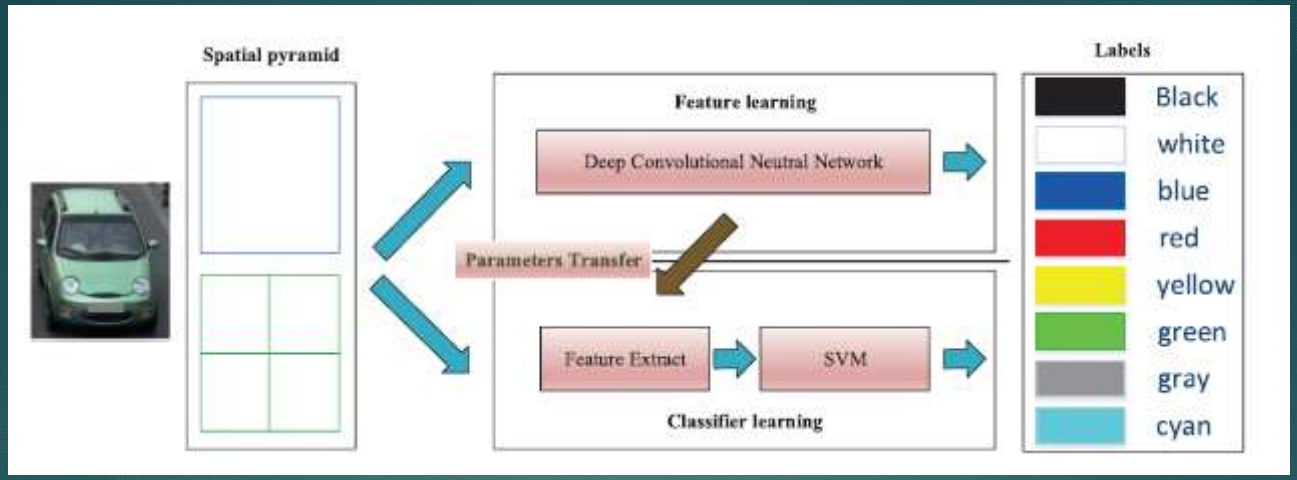
- License plates are usually used to monitor vehicle activities.
- But they need to be visible completely from every angle of camera, always; which is not the case.
- However, the vehicle color satisfies this.
- In this paper, a system that identifies the color of a vehicle using deep learning techniques has been proposed.
- This method provides better efficiency than the existing ones.

METHODOLOGY

- The CNN architecture is trained using the training examples and is later acting as a feature extractor that computes a feature vector for each input image.
- SVM is used as a classifier that predicts the color class of a given vehicle.
- Spatial information is utilized to capture the structural properties of vehicles and leads to higher recognition accuracy.

A. *Feature Learning With CNN*

- *The CNN network is composed of five convolution layers {C1, C2, C3, C4, C5} and three fully-connected layers {fc6, fc7, fc8}.*
- *Following each convolution layer, the contrast normalization, pooling, and nonlinear function are connected to it successively.*



- The output of the convolution layer is reshaped as a feature vector and fed to the fully-connected layers.
- After training the deep CNN, we choose the outputs of the last three layers, since these outputs of the first four layers are not discriminative enough, and the dimensions of the features are too high.

COLOR RECOGNITION WITH SP AND SVM

- The SP strategy is a classical method that embeds spatial information into a feature vector
- The SP is aimed at extracting features from different sub regions and aggregating the features of all the regions together to describe an image.
- This method computes the feature of each region in the convolutional layers in the learned CNN architecture.
- Unlike the CNN method, the output from convolution layers are fed to SVM classifier.
- 2 reasons for doings this :

1. The performance of SVM is better because it can help combat over fitting.
2. The number of parameters of SVM is less than that of fully-connected layers, which makes the fine-tuning procedure in training much easier.

Pros :

- The proposed feature is more informative than conventional features, even at an extremely small scale.
- The features from the training data are learnt automatically, and hence doesn't adopt manually designed features.
- The algorithm is run on raw pixels, and hence pre-processing of images to remove noise, hue, etc., is not required.
- Using spatial method for recognition increases the accuracy further.

Cons :

- The system is not 100% accurate.
- Severe illumination and several indistinguishable colours results in wrong predictions.

Implementation

- As a part of the mini project, a system using 5 convolution layers of CNN network will be built.
- The Spatial Pyramid strategy is used instead of fully connected layers, as the authors claims it to be more accurate.
- SVM classifier will then be used to classify the image vectors while training and recognise the colour while testing.
- The whole system will be built in python.