VEHICLE'S COLOR AND LOGO IDENTIFICATION

DAMITHA SENA BANDARA TILAKARATNA

A THESIS REPORT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING PROGRAM IN AUTOMOTIVE ENGINEERING
INTERNATIONAL COLLEGE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2017
KMITL-2017-IC-M-004-003

THESIS TITLE	Vehicle's Color and Logo Identification
STUDENT NAME	Damitha Sena Banadara Tilakaratna
STUDENT ID	58610012
DEGREE	Master of Engineering
PROGRAME	Automotive Engineering
	(International Program)
ADVISOR	Dr. Ukrit Watchareeruetai
CO-ADVISORS	Dr. Supakorn Siddhichai
	Assoc. Prof. Masaki Yamakita

Abstract

Vehicle identification is an essential criterion used by police force to detect traffic and criminal activity. Vehicle license plate recognition (LPR) is one of the common methods of vehicle identification. This method has its own drawbacks when it comes to the visibility of number plate and usage of embellish number plates. Moreover, the LPR method is not applicable for vehicles that use illegal number plates. Though it is a simple task to swap number plates, it is difficult to change the color and make of the vehicle. Therefore vehicle model, make and color information should be considered to increase the accuracy of vehicle identification.

This work discusses the implementation of vehicle color and logo identification methods. In color identification part, several segmentation methods and classification methods were discussed. Mainly, histogram back projection method was used with k-nearest neighbor (KNN) and support vector machine (SVM) classifiers. Furthermore, logo identification part has been divided in to two sub-parts as logo localization and logo classification. In logo identification part SVM, multilayer perceptron (MLP) and convolution neural network (CNN) were used.

Proposed vehicle color recognition methods were able to successfully recognize ve-

hicle colors with the highest accuracy of 87.41% in SVM based method. However, there were some difficulties in histogram and back projection based method while classifying the colors due to shadows and multi-colored appearance of the vehicles.

Proposed vehicle make recognition methods were able to successfully localize the vehicle logo with an accuracy of 88.72% and recognize vehicle make with an accuracy of 97.65% in CNN based method. When both systems were combined together, automatic vehicle logo cropping with logo classification achieved an accuracy of 82.45%.

This work has introduced several methods to improve the accuracy of vehicle identification systems. In color recognition, a segmentation method has been suggested on precisely extracting the vehicle metallic area which encloses the color of the vehicle. In vehicle logo identification, sliding window based method has been suggested on logo localization. In future work, this work can be further developed to identify vehicle make along with its model as well to recognize multicolored vehicles such as taxis and trucks.

Acknowledgments

I would like to express my gratitude to supervisors, Dr. Ukrit Watchareeruetai of International College, King Mongkut's Institute of Technology Ladkrabang (KMITL) and Dr. Supakorn Siddhichai of Image Technology Laboratory, NECTEC at National Science and Technology Development Agency who was abundantly supportive and offered invaluable assistance and guidance to me. Furthermore, my deepest appreciation are also due to the members of the supervisory committee Dr. Nattachai Watcharapinchai of Image Technology Laboratory, NECTEC at National Science and Technology Development Agency. Without their knowledge and assistance this study would not have been successful.

Special thanks also go to my Japanese advisor Associate Professor Masaki Yamakita who I was fortunate enough to work under in the exchange program. My sincere gratitude also goes to the staff of KMITL and NECTEC who assisted me in numerous ways throughout the two years of my master life.

I would also like to convey my appreciations to the KMITL for the financial support given to follow my master's program. Last but not least, I owe more than thanks to my parents, for their financial support and encouragement throughout my life. Without their support, it would have been impossible for me to pursue my education seamlessly

Damitha S.B. Tilakaratna

Contents

Appendices

A Publication

List of figures

List of tables

Chapter 1

Introduction

In this chapter, the motivation and problem description are described. The chapter also explains the structure of this thesis document.

1.1 Motivation and problem description

Vehicle identification with the use of vehicle license plate recognition system is currently implemented on the traffic system in Thailand. In Thailand, a vehicle license plate consists of letters and numbers. The current system has a good accuracy but it has some drawbacks. Visibility of vehicles number plates have a huge influence on the vehicle identification accuracy. Bad visibility of vehicle number plate can lead to errors in number plate localization and character recognition. Additionally, vehicle identification using license plate recognition (LPR) system is not applicable in identifying vehicles that use illegal license plates. In order to increase the accuracy of vehicle identification, information such as vehicle make, model and color can be used. Recently, vehicle make and color recognition (VMCR) systems are developed as a solution for this issue and a way of recognizing vehicles more accurately.

Chen et al. were able to develop a good algorithm for locating the region of interest

(ROI) in order to do the classification of colors in [13]. Though they have a high accuracy in locating the region, their color classification robustness is low as their method fails in different lighting conditions. Different light conditions affect the color classification accuracy as the pixel values of the same color changes in respect to the lighting condition. Kryjak et al. have implemented a system to work in different lighting conditions in [12], so that the color classification will be more robust in real life usage. Similar to most of the implemented methods, the method in [12] has considered only rich colors where shades of colors were not taken into consideration. In color classification another important aspect is the feature vector. For example, color can be represented in R, G, B values or H, S, V values. Dule et al. have studied about different feature vectors that can be used in color classification in [8].

Vehicle make classification has been done in different ways. Some work have been done by the vehicle logo classification and some were done by vehicle model and logo classification together. In vehicle logo classification, localization of the logo hugely affects the accuracy of the classification system. Liu and Li have proposed a vehicle logo localization method based on edge detection in [14]. Liorca et al. have used support vector machines (SVM) for the logo classification in [15].

There are so many methods that have been implemented in vehicle color and logo identification. Most of the color identification methods have focused on classifying pure colors. Mainly, non-colors such as black, gray, white and silver were not considered to be separate classes, where non-colors were mostly defined as black and white color classes. This work focuses to identify vehicle color including those non-colors as separate classes. Moreover, most of the work in vehicle logo identification have considered logos that have regular quadrilateral area. In this work, logos with regular and irregular quadrilateral areas were taken in to consideration in vehicle logo identification.

1.2 Objectives and scope

There are so many ways that have been implemented in order to identify vehicles. The main objective of this work is to identify vehicle logo and color information. In identifying vehicle color, vehicle metallic area segmentation methods will be proposed. Additionally black, white, silver and gray colors will be considered as separate classes to propose a good vehicle color identification method. Several classification methods such as k-nearest neighbor (KNN) and SVM will be evaluated using Thailand traffic cam dataset [24]. Furthermore, license plate coordinate information will be acquired from the LPR system.

Furthermore, in vehicle logo identification part, a logo localization and logo classification methods will be proposed. Moreover, classification methods such as SVM, multilayer perceptron (MLP) and convolution neural network (CNN) will be evaluated with the proposed methods to find out the best classifier. Moreover, vehicle logos with regular and irregular quadrilateral areas will be considered in logo localization and classification. Two datasets, i.e. Thailand traffic cam dataset [24] and COMPCAR dataset [21], will be used to evaluate methods to check the robustness of the identification system.

1.3 Thesis structure

The remaining of this thesis is organized as follows: Chapter 2 provides background knowledge and describes some related works. Chapter 3 explains the proposed methods. Chapter 4 presents and discusses experimental results. Finally, Chapter 5 concludes the thesis.

Chapter 2

Background knowledge

This chapter provides background knowledge of vehicle recognition systems that are being currently used as well as methods that were used on those systems. This would assist readers to get an overall knowledge of a vehicle recognition system.

2.1 Feature extraction methods

In image processing, features represent some information about a ROI of an image, that can be used to solve some computational task related to a process. Features can be specific structures such as edges or objects. Moreover, features can also be a result of general neighborhood operation or feature detection applied on an image. Features play a very important role in the area of image processing. Prior to feature extraction, various image processing techniques such as binarization, thresholding, re-sizing and normalization are applied on the sampled image. After that, feature extraction techniques are applied to get features that will be useful in classification and recognition of images.

Feature extraction is an important step in the construction of any pattern classification and aims at the extraction of the relevant information that characterizes each class. In this process, relevant features are extracted from objects to form feature vectors.

These feature vectors are then used by classifiers to recognize the input unit with target output unit. It becomes easier for the classifier to classify between different classes by looking at these features as it is fairly easy to distinguish. Feature extraction is the process to retrieve the most important data from the raw data. Feature extraction is finding the set of parameters that define the shape of an object precisely and uniquely. In feature extraction phase, each object is represented by a feature vector, which becomes its identity.

As explained in [13], a good feature set contains discriminating information, which can distinguish one object from other objects. It must be as robust as possible in order to prevent generating different feature codes for the objects in the same class. The selected set of features should be a small set of values which efficiently discriminate among patterns of different classes, but are similar for patterns within the same class. Features can be classified in to two parts:

1. Local features, which are usually geometric (e.g. concave/convex parts, number of endpoints, branches, joints etc.).
2. Global features, which are usually topological (connectivity, projection profiles, number of holes, etc.) or statistical (invariant moments etc.).

When it comes to feature extraction methods, there are pre-built methods available in digital image processing. However, these feature extraction methods can be divided into three major groups in respect to feature types.

1. Statistical features

   Distribution of points in an image will be considered in deriving these features. Statistical features provide high speed and low complexity and take care of noise to some extent. These features may also be used in reducing the feature set dimension.

2. Global transformation and series expansion features

   The main advantage of these type of features are that it is invariant to global deformations like translation and rotations. For classification purpose generally more information is represented in a continuous signal. Linear combination of a series of simpler well-defined functions is a way to represent a signal. Compact encoding provided by the coefficients of the linear combination is known as series expansion. Fourier transforms, Hough transform, moments and rapid transform are some examples of common transform and series expansion features.

3. Geometrical and topological features

   These features have high tolerances to distortions and variations and may contain global and local properties. These features will help to understand the contour of the object or what sort of components make up that object.

   Some of the common feature extraction methods are as follows:

   - Scale invariant feature transform (SIFT) [16]

     These features describe local features in an image. This is commonly used in object recognition, image stitching, gesture recognition and video tracking. This feature descriptor is invariant to scaling, illumination changes, orientation and partially invariant to affine distortion.

   - Histogram of oriented gradients (HOG) [5]

     This method is used in object detection. This method counts the number of occurrences of gradient orientation in small regions of an image. This feature extraction method differs from edge orientation histogram, SIFT and shape context because it is computed on a dense grid of uniformly spaced cells and it uses overlapping local contrast normalization to improve accuracy.

- Speeded up robust features (SURF) [2]

  This is a patented local feature detector and descriptor mainly used in object detection and image registration. SURF is faster than SIFT and also robust against different image transformation than SIFT.

2.2 Classification methods

In machine vision, classification is the most vital and challenging task. Mostly, the classification is based on image texture, similarity or description of items or things. Classification can be divided into two main approaches.

1. Supervised classification

   In this approach, user inputs sample images that represent specific classes and then input those images to the processing software in order to be used as a reference for classification of all other new images. Training sets are selected and grouped by the user and labeled according to the group.

2. Unsupervised classification

   In this approach, the outcomes depend on the analysis of an image by the software without user providing any sample classes. Software itself uses techniques to analyze the image and determine which pixels are similar and group them into classes. In this algorithm, the number of classes can be specified by the user. This is commonly used when no labels are available for the dataset.

There have been several image classification methods that have been developed in image processing. Some of the common classification methods can be explained as follows:

2.2.1 Support vector machine (SVM)

SVM [10] is a well known for multi-class classification by constructing a set of hyper-planes in a high dimensional space. This classification method comes under supervised learning model. A SVM training algorithm builds a model that assigns new examples to one category or the other while making it a non-probabilistic binary linear classifier. In a simpler way, SVM model is a way of representing examples as some points in feature space, so that example from other categories are divided by a clear gap that is wide as possible. Therefore the new examples will be mapped into the same feature space and from that it will predict the relevant category based on which side of the gap they fall into. SVM was developed from the theory of structural risk minimization. It is to transform a low-dimensional feature space into a high-dimensional feature space and find out the maximum division margin between classes. Figure 2.1 is an example of a SVM, trained with samples from two classes.

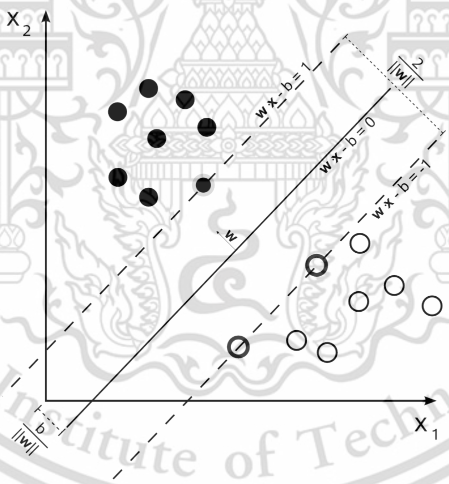

Figure 2.1: SVM trained with samples from two classes

(Source: https://en.wikipedia.org/wiki/Supportvectormachine)

2.2.2 Multilayer perceptron (MLP)

MLP [17] is a class of feedforward artificial neural network, consists of at least three layers of nodes. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. These neurons in all layers are fully-connected by weighted connections as shown in Fig.2.2. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. MLP can distinguish data that is not linearly separable.

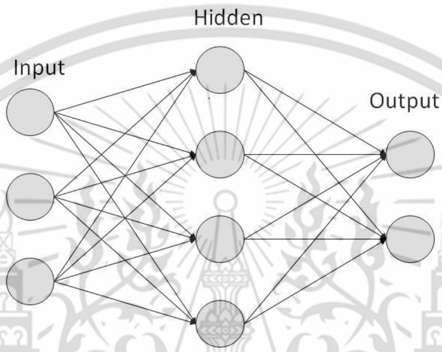

Figure 2.2: A simple multilayer perceptron network
(Source: https://www.tutorialspoint.com/)

2.2.3 Convolution neural network (CNN)

CNN [11] is a class of deep learning method in machine learning. Feed forward artificial neural networks have been successfully applied to analyzing image datasets. This classification method has invariant characteristics and shared weight architecture. CNN uses less pre-processing compared to other methods. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. Figure 2.3 is an example for a CNN.

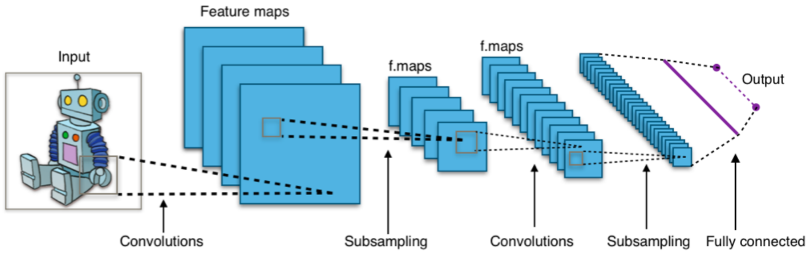

Figure 2.3: A simple convolution neural network
(Source: https://en.wikipedia.org/wiki/Convolutionalneuralnetwork)

2.3 Color classification

Chen et al. have proposed a novel scheme for the color identification of vehicles using the locating algorithm of ROI as well as the color histogram features from still images [13]. A coarse-to-fine strategy was adopted to efficiently locate the ROIs for various vehicle types. The red patch labeling shown in Fig.2.4, geometrical-rule filtering and a texture-based classifier were cascaded to locate the valid ROIs. A color space fusion together with a dimension reduction scheme were designed for color classification in their method. Color histograms in ROIs were extracted and classified by a trained classifier. Seven different color classes were used in classification. The average rates of ROI location and color classification were 98.45% and 88.18% respectively. Moreover, the classification efficiency of the the method was up to 18 frames per second.

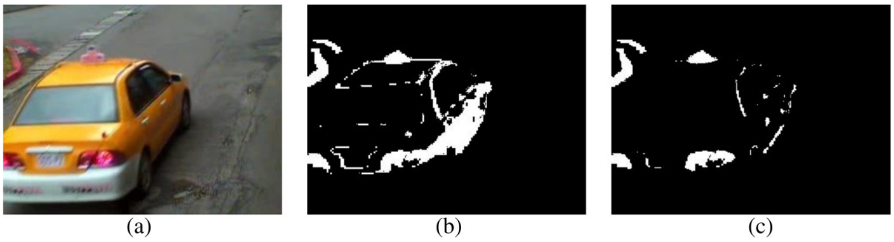

Figure 2.4: The red patch labeling for a taxi: (a) a taxi image, (b) the mis detected red pixels using simple rules, and (c) the labeled results verified by an SVM classifier

(Source: [13])

Dule et al. have studied the performance of different feature sets obtained by various color spaces and different classification methods in order to improve the outdoor vehicle color recognition [8]. Moreover, different ROIs and feature vector construction methods have been developed to gain better performance. They have examined two ROIs (smooth hood peace and semi front vehicle), three classification methods (*k*-nearest neighbors, ANN, and SVM), and all possible combinations of 16 color space components as different feature sets. This work has obtained 83.50% success in experiments.

Kryjak et al. have proposed a novel color correction technique for classifying vehicles under different lighting conditions [12]. To reduce the lighting effects, a reference image is first selected in order to building the mapping function between the current frame and the reference image. With this mapping function, the color distortions between frames can be reduced to minimum. In addition to lighting changes, the effect of sun light will make the vehicle window become white and lead to the errors of vehicle classification. To reduce this effect, they have proposed a window removing method for making vehicle pixels with the same color more concentrated on the foreground region. Then, vehicles can be more accurately classified to their categories even though strong sun light casts on them. To tackle the confusion problem with some vehicle colors that are too similar, e.g., deep-blue and deep-green, a novel tree-based classifier is proposed for classifying vehicles to more detailed labels.

2.4 Make classification

Sittampalam and Ramanan have illustrated a framework to classify 25 distinct vehicle logos using a classifier-free codebook-based classification technique [19]. Their method not only speeds up the recognition process but also reduces the storage requirement when comparing with the traditional bag-of-features (BoF) approach [9]. Their approach is free from representing the features as fixed length histograms and then classifying the histograms to predict the class label of a given test image. Therefore the classification rate is faster compared to the traditional BoF approach. According to their results comparison, their method achieved an accuracy of 93.6% which was very similar to or slightly better than the traditional BoF approach. In addition to the computational and classification performance of their method, the approach is also fundamentally different from the traditional BoF approach which makes use of standard classifiers in classifying vehicle logos. Figure 2.5 demonstrates a clear comparison between BoF method and the approach in [19].

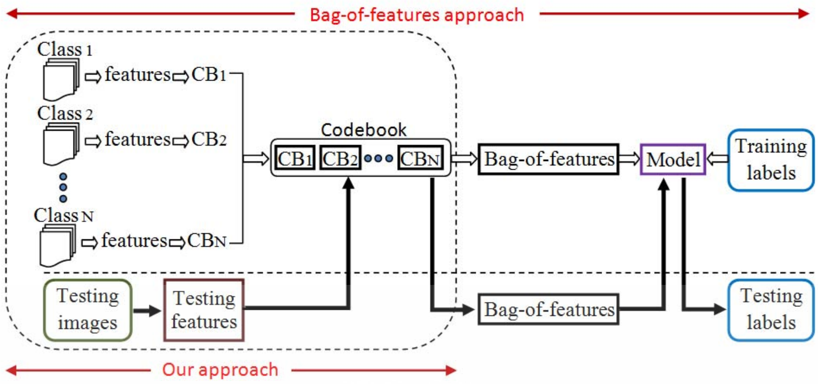

Figure 2.5: Bag of features versus approach in [19]
(Source: [19])

In Sittampalam and Ramanan approach, the training stage is applied to the training set to obtain a locally merged global codebook, while, in BoF method, the training stage is applied to the training set to obtain a codebook and a classifier. In addition, the testing stage is applied to the test set using the learned global codebook to obtain a label for every test image by means of a majority voting strategy applied on the visual descriptors, while, in BoF method, testing stage is applied to the test set using the previously computed codebook and classifier to obtain a label for every test image.

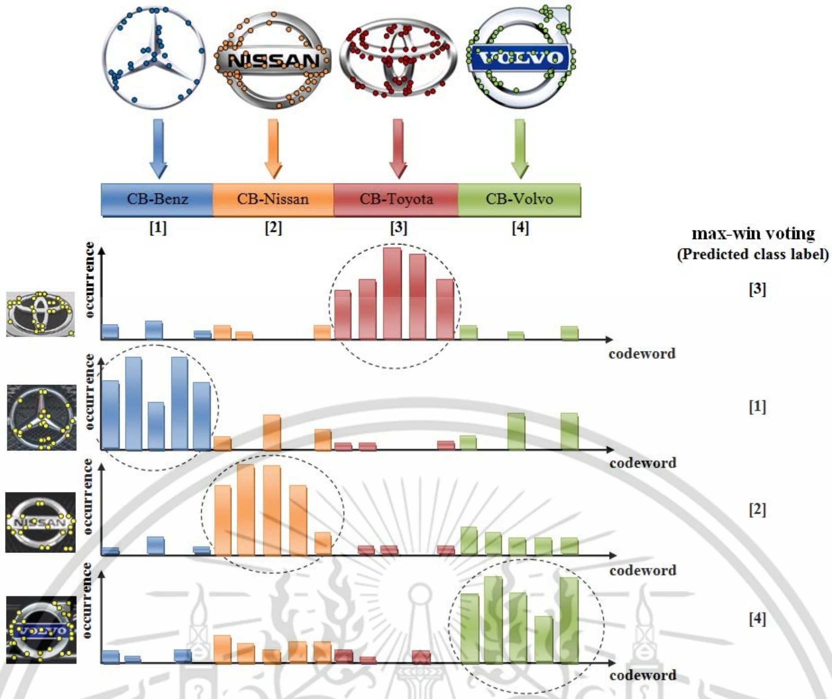

Figure 2.6: Codebook based classification

(Source: [19])

Figure 2.6 demonstrates the majority voting strategy. The method in [19] measures the L2-norm for each visual descriptor of a test image. The relative position of the codeword in the global codebook that results in minimum distance is recorded in a list. Then, to predict the class label of the test image, a majority voting strategy is applied.

Liorca et al. have proposed a HOG/SVM framework for vehicle logo recognition using images captured by traffic cameras [15]. Logo detection is assisted by a previously developed LPR stage. A sliding window approach is applied in a ROI, which is defined above the detected license plate as shown in Fig.2.7. Local binary patterns (LBP), SIFT and HOG have been studied as features to represent the vehicle logo. A multi-class SVM were used to classify all the regions provided by the sliding window stage. Finally, a majority vote approach is implemented to estimate the logo using the binary outputs given by the SVM. The method in [15] was assessed on a set of 3,579 vehicle images, captured by two different traffic cameras, that belong to 27 distinctive vehicle manufacturers. They have managed to obtain 92.59% performance from the method.

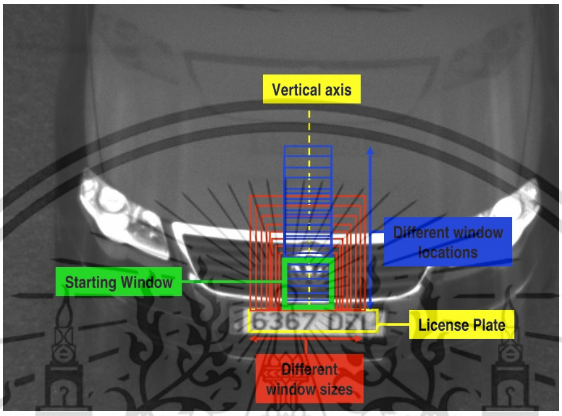

Figure 2.7: Overall view of the sliding window approach in [15]
(Source: [15])

Zheng et al. have proposed a method of vehicle logo recognition based on bag-of-words [22]. They have used dense-SIFT to extract stable features, quantized features by soft assignment and computed histogram with spatial information to improve the performance. In [22], vehicle logo dataset consists of 840 small size images (about 30×30 pixels) from 14 classes. Their frame work is shown in Fig.2.8. Their method has 97.3% recognition accuracy and also the recognition speed is rather fast, about 0.0226 seconds per image.

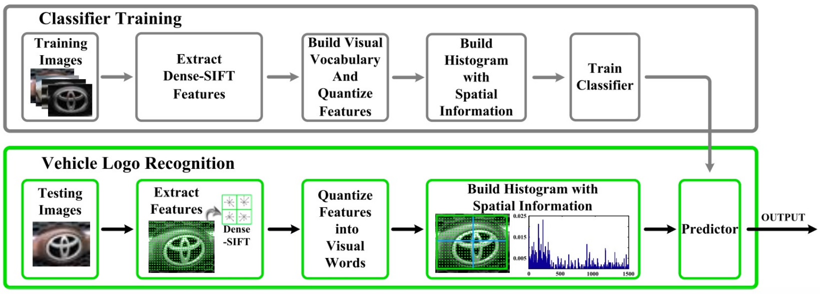

Figure 2.8: Overall view of the sliding window approach in [22]

(Source: [22])

Anakavej et al. have proposed a new framework for vehicle searching system and a new approach for vehicle model and logo recognition from frontal view of vehicles [1]. In this framework, users have to input name or detail of vehicle model that they want to search in videos or image databases. The system will automatically collect sample images of the desired vehicle model from internet. As a result, the training samples for any vehicle model are not necessary to be prepared in advance. For vehicle model and logo recognition, they have proposed a novel scheme based on eigenfaces [4] and pyramid of HOG. Their method can provide the recognition rates of 98.32% and 94.00% for vehicle logo and model recognition, respectively.

Liu and Li have proposed a vehicle logo location recondition system [14]. They have successfully located the position of the number plate from the LPR system. Moreover, they have taken the coordinates of right bottom point of the number plate as a reference from LPR system. With the use of license plate location and dimensions, they have proposed a method to define an area of the image that includes the vehicle logo. Furthermore, they have used Sobel operator to detect the texture of the grille and suppress the noise of grill texture. Then they have dilated the image and successfully located the area of the logo. They have used 200 vehicle images with 180 vehicle images either horizontal or vertical grille texture and 20 images with messy or complicated grille texture. Their method is effective when they deal with vehicle logos with horizontal and vertical grille texture. However, messy grille texture has caused some errors in accurately detecting the logo location.

Chapter 3

Proposed methods

This chapter presents methods which were used for vehicle identification using logo and color information on this work. The structure of this Chapter consists of four main sections, including system overview, color recognition, logo localization and logo classification.

Under those main sections several methods were implemented. Color recognition has been done using techniques such as automated ROI based method and histogram back projection based method. Furthermore, logo localization was carried out in two methods, namely HOG feature based method and CNN based method. Furthermore, CNN based approach was implemented in logo identification.

3.1 System overview

In vehicle identification, the common way is to use the LPR system. The current LPR system in Thailand is capable in localizing the number plate and reading the number plate characters. Though LPR is good in vehicle identification, it has drawbacks when it comes to bad visibility and usage of embellish number plates.

Vehicles can be identified by its color, shape, make and model. As a solution for drawbacks in the LPR system, this work proposes a vehicle logo and color identification system as a way of acquiring more information about a vehicle.

This work was carried out in three main parts as follows which is also graphically represented in Fig.3.1.

1. Color recognition
2. Logo localization
3. Make classification

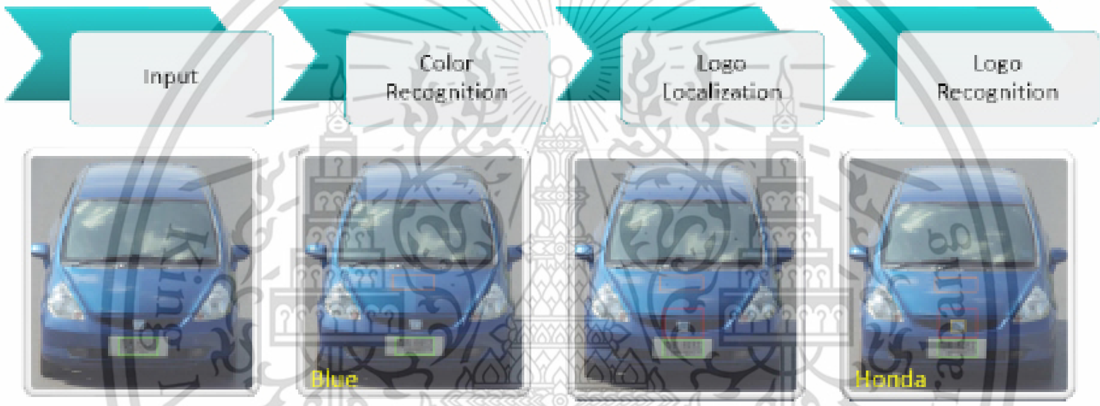

Figure 3.1: Overall view of the vehicle identification system

The current LPR system accurately localizes the vehicle license plate from the input video and stores the vehicle image alongside with the license plate information. This LPR system can give out the top left corner point, height and the width of the vehicle license plate. These parameters were used throughout this work as inputs from LPR system to vehicle make and color recognition (VMCR) system.

3.2 Color recognition

As LPR system captures the frame of the vehicle, it stores that image in the database. Color recognition was done with the use of those stored images as it has a variety in colors and a large number of images.

These images were manually grouped in to color classes. The list of color classes was prepared according to the vehicle color categories which is being used in Thailand land transportation department. After studying the database, it was found that vehicles with silver, gray and white colors can be found frequently. Because of this reason, this work paid an indispensable focus on classification between silver, gray and white colored vehicles as those colors are even similar to the human eye. Vehicle color recognition was mainly done under two methods in order to achieve a better accuracy in color recognition. Methods will be discussed as follows:

1. Automated ROI based method
2. Histogram back projection based method

3.2.1 Automated ROI based method

The input image contains noises such as the road, lines and shadows which have to be filtered out in order to achieve a good accuracy. This part of the work was done to create a mask that will only extract the vehicle metallic (painted) area from the input image.

In order to reduce the noise of the input image, ROI is defined right above the license plate with the use of license plate coordinates as shown in Fig. 3.2.

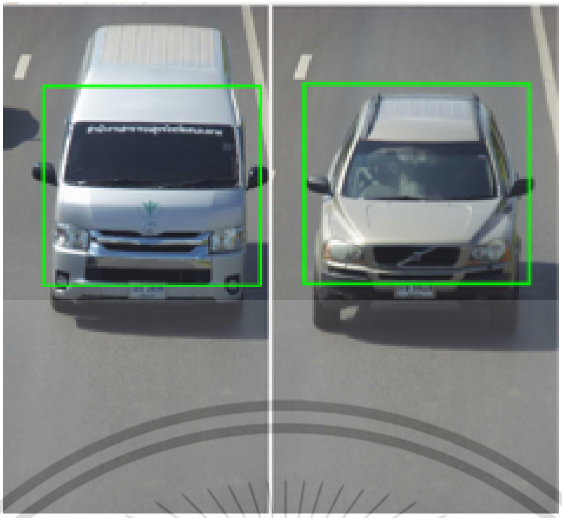

Figure 3.2: First ROI in automated ROI based method

After extracting the region of interest, as shown in the Fig.3.2, the system then creates two matrices. One matrix contains the maximum value from each RGB channels and the other matrix contains the minimum value from each RGB channels. For instance, a white vehicle will contain high RGB values. Therefore the maximum matrix will contain 255 as the values for the colored area and low values for other areas of the image. Moreover, in the minimum matrix the road area will contain the lowest values.

After calculating these two matrices, matrices will be subtracted and the difference will be stored in a matrix. The higher values depicted in the matrix which were created after the subtraction of two minimum and maximum matrices, are from the vehicle and the lower values depicted are from the background. Then by thresholding, a mask was created as shown in Fig.3.3.

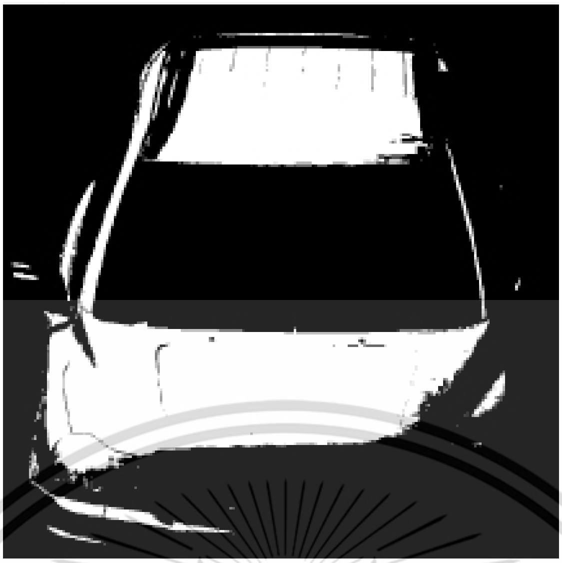

Figure 3.3: Example of vehicle mask created from automated ROI based method

The noise has been removed entirely and the image mask will only represent the vehicle metallic area as shown in Fig.3.3. This mask was used to extract the pixel values only from vehicle metallic area to do the color classification.

3.2.2 Histogram back projection based method

This method is another way of creating an accurate mask in order to reduce the noise of the input image. In this method, pre-defined ROI is used to calculate the histogram which will help in creating a mask that can extract pixel values from vehicle metallic area. Figure 3.4 presents the major steps that were used in this method.

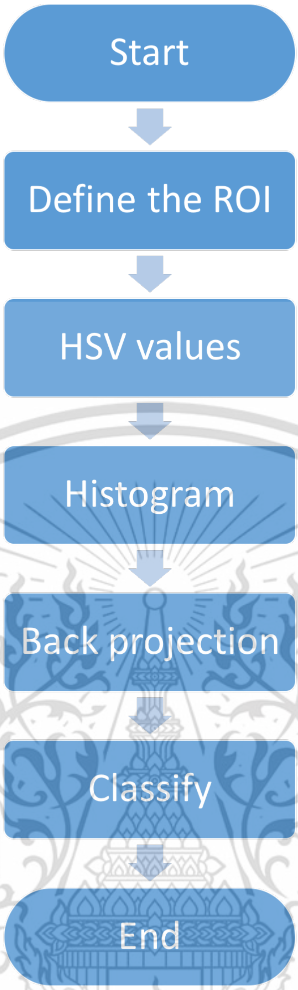

Figure 3.4: Overview of the histogram back projection based method

A rectangular ROI is defined on the hood area with a size of 50×200 pixels. The position of the ROI is determined by the license plate position and size. The program will extract the pixels from the defined 50×200 sized ROI on the hood area as shown in Fig. 3.5. This ROI is then converted into HSV color space using the Eqs. 3.1–3.6 [23].

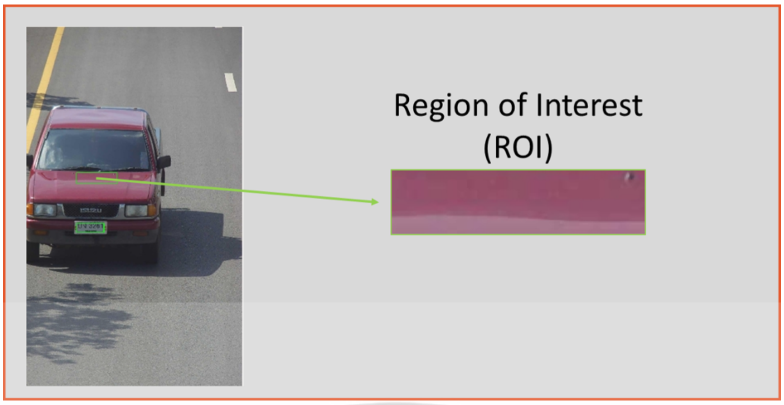

Figure 3.5: ROI in histogram back projection based method

$$C_{max} = max(R, G, B) \tag{3.1}$$

$$C_{min} = max(R, G, B) \tag{3.2}$$

$$\Delta = C_{max} - C_{min} \tag{3.3}$$

$$H \begin{cases} 0 & if C_{max} = 0 \\ (60 \times \frac{G-B}{\Delta} + 360)\, mod 360 & if R = C_{max} \\ 60 \times \frac{B-R}{\Delta} + 120 & if G = C_{max} \\ 60 \times \frac{R-G}{\Delta} + 240 & if B = C_{max} \end{cases} \tag{3.4}$$

$$S = \frac{\Delta}{C_{max}} \tag{3.5}$$

$$V = \frac{C_{max}}{255} \tag{3.6}$$

The converted ROI will then be used to calculate the histogram model for the color on that region. The histogram is normalized to be in the range of $[0, 255]$, in order to get an accurate histogram model as shown in Fig. 3.6.

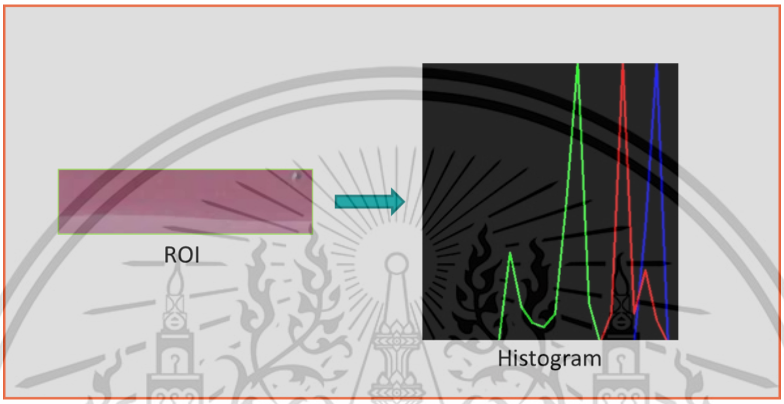

Figure 3.6: Histogram model

After calculating the histogram model for the ROI, the model was used for back projecting as shown in Fig.3.7. Histogram back projection[18] is a way of recording how well the pixel values of the given image fit the distribution of pixels in the histogram model. In other words, this system calculates histogram model for ROI and then use the model to find these features in the entire input image. Therefore the same colored areas will pop out as a mask as shown in Fig.3.7.

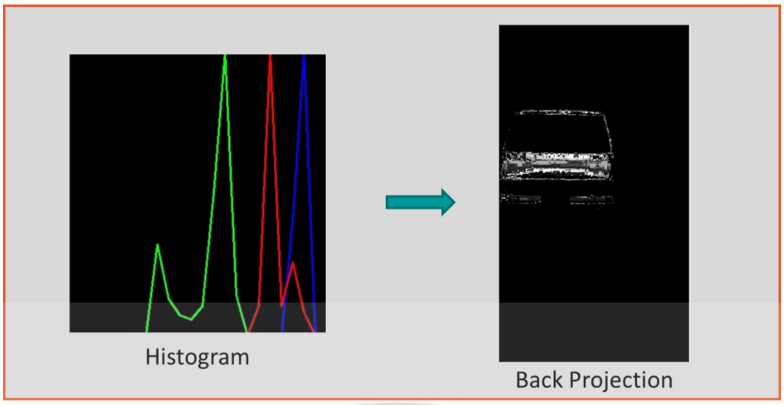

Figure 3.7: Histogram back projection

The back projection method collects information of each pixel of the image and then it will find the corresponding bin location of the histogram model. Subsequently, the bin value will be saved in a new image. Finally, a mask is created by thresholding. Therefore this mask will contain a majority of vehicle metallic area which represents the overall color of the vehicle. Using this mask, the system can reduce the noise of the input image by eliminating areas such as background and windshield which are not useful for the color recognition system. As shown in Fig.3.8, color pixels will be extracted from the image.

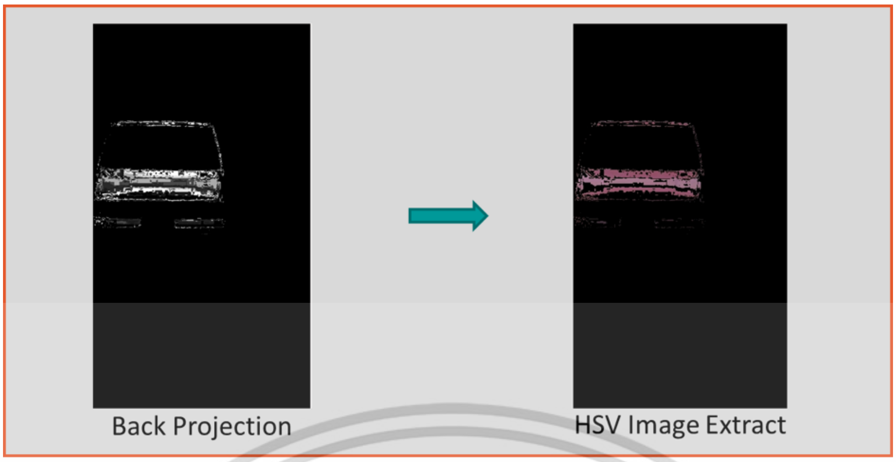

Figure 3.8: Metallic area extraction from histogram back projection based method

With the use of pixel values from the metallic area of the image, color classification is done using KNN and SVM classifiers. Majority voting basis was then applied to classify the color of the vehicle as shown in Fig.3.9.

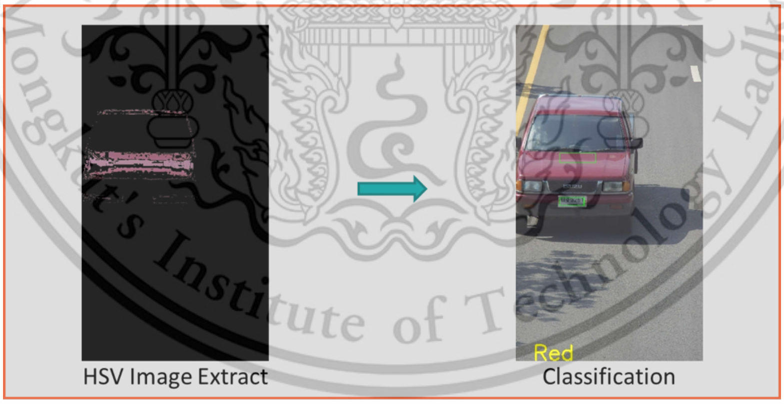

Figure 3.9: Color classification

This method has its own pros and cons. As shown in Fig. 3.10, this method successfully classifies major colors. However, as shown in Fig. 3.11, white color classification gives an error as the method fails to segment the vehicle metallic part.

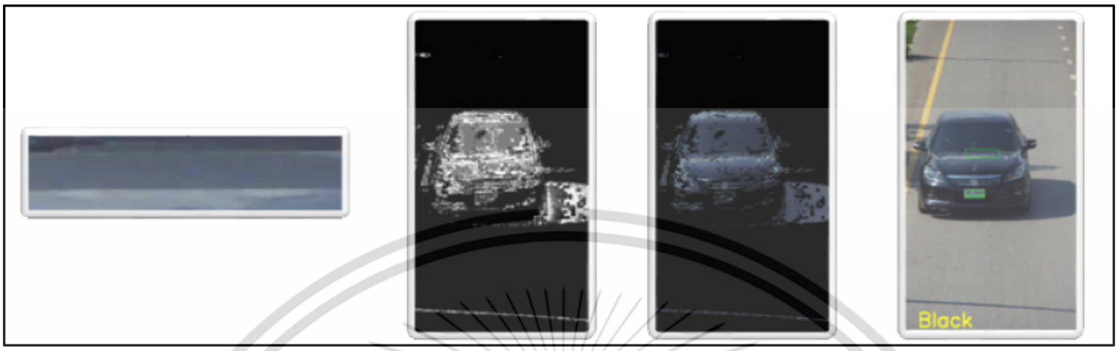

Figure 3.10: Correct color classification

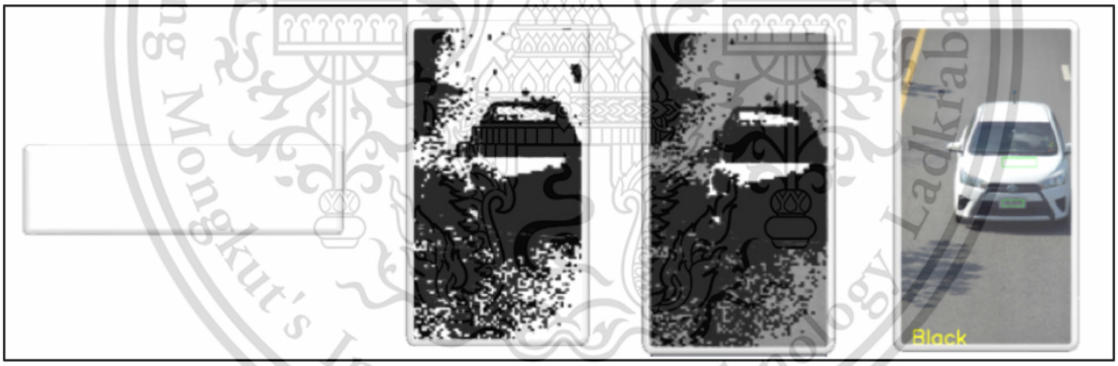

Figure 3.11: White color classification error

As shown in Fig.3.13, this method fails to create an accurate mask for white color vehicles. As a solution for this, the system will first extract the ROI from the vehicle hood area and then it will determine if the color is white using pre-defined values. If it is classified as white the system will output the vehicle as white colored vehicle. If not, the system will proceed with the histogram back projection method to create an accurate mask and then do the vehicle color classification using the pixels that are being extracted using the mask.

After completing vehicle metallic area extraction, classifier is used to classify the vehicle color. Feature vectors were created in order to be used for the classification. To create feature vectors, an area of 10×10 pixels were considered per time. Within this area, each pixel were used to calculate the feature vector. Figure 3.12 is an example for feature extraction.

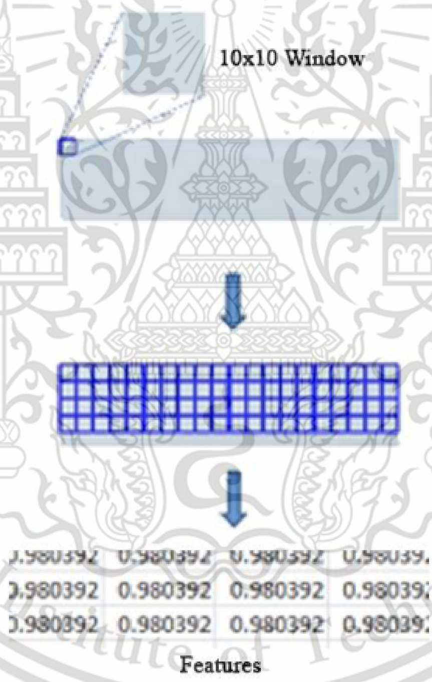

Figure 3.12: Extracting features

In order to create an accurate feature vector, several color spaces were considered in extracting features. Each color space has its own advantages and disadvantages. For example, HSV color space is good with intensity variation but poor in classification on non-colors such as black, white and sliver. However, when RGB color space is considered, it is good with colors but poor with intensity variations. As this work aimed on non-colors as well as intensity variations, RGB and HSV color spaces were chosen to extract features. Two types of feature combinations were considered in this work.

- Four features (R, G, B, V)
- Six features (R, G, B, H, S, V)

When four features are extracted from a pixel, there will be 400 features in the feature vector for the respective area of 10×10 pixels. Similarly, when six features are extracted from a pixel, there will be 600 features in the feature vector for the respective area of 10×10 pixels.

3.3 Logo localization

Logo localization is an essential part in make classification. This section will explain the methods that were carried out in order to accurately determine the logo location which will be used for make classification later on. Mainly three classification methods were used as follows:

1. SVM
2. MLP
3. CNN

In each machine learning method, different features were used in order to achieve a high accuracy as the make classification relies on the result of logo localization. The dataset was created manually by extracting logo areas and non-logo areas as shown in Fig.3.13. For this manual extraction, vehicle images were randomly chosen in order to have a variety in logo representations.

Figure 3.13: Manually extracted logo and non-logo areas

The logo area (marked in green) is manually selected and extracted from the image to be saved in the logo dataset. In parallel to that, 4 adjacent square areas are extracted to create the non-logo dataset. As the logo size varies from vehicle to vehicle, the non-logo areas are changed automatically in order to keep the dataset valid through out any vehicle type. As for the non-logo image dataset, it contains with plain color, lines, grill area and, sometimes, a bit of vehicle number plate area, as shown in the Fig.3.13. As to be a standard throughout both datasets, all the images were re-sized in to 80×80 pixels image. Moreover, all the images that were used in logo localization and logo classification were taken in day time condition.

Figures 3.14 and 3.15 show some samples of positive and negative datasets. These datasets were used in training the models for logo localization methods and logo classification methods.

Figure 3.14: Example of positive samples

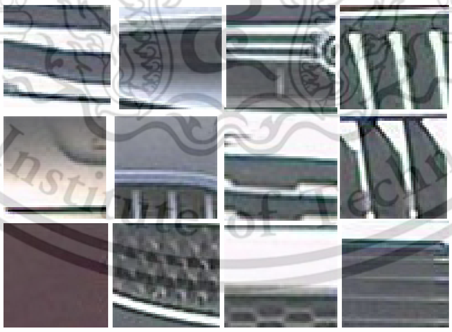

Figure 3.15: Example of negative samples

In order to locate the vehicle logo, a method of sliding window was implemented. This sliding window approach was applied within a ROI, which was defined using the license plate coordinates and width as shown in Fig.3.16. The defined ROI was then used for the sliding window approach as shown in Fig. 3.17.

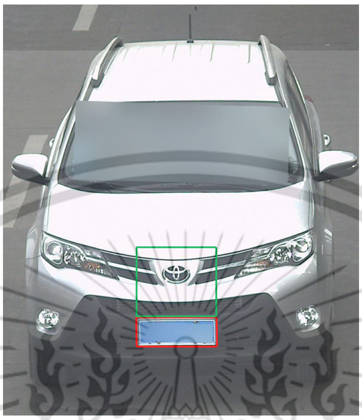

Figure 3.16: ROI in sliding window approach

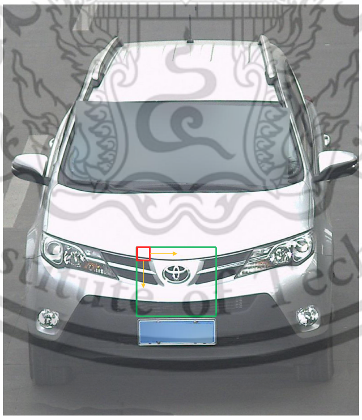

Figure 3.17: Sliding the window within the ROI

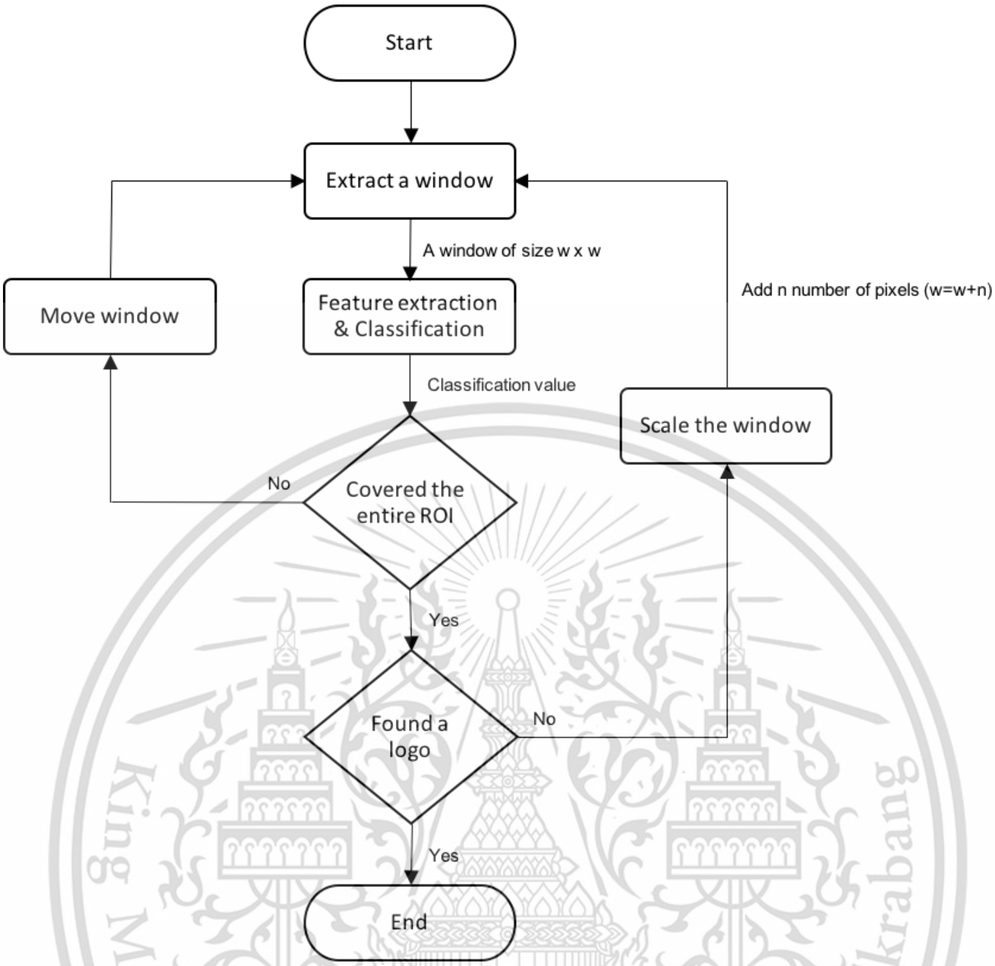

Figure 3.18: Sliding window approach

The sliding window approach will start by extracting a window with initial size of $w \times w$ pixels as shown in Fig.3.18. This widow will start from the top left corner of defined ROI as shown in Fig.3.17. The window will extract a region from ROI and features will be used in classifier to get the classification value to that particular window. This process will be repeated all throughout the defined ROI. If this $w \times w$ pixels sized window fails to locate a logo, the sliding window size is scaled by adding n number of pixels to width and height of previous window size. After scanning throughout the ROI, the window with maximum classification value will be extracted as the positive logo area.

Several classification methods were used in order to classify the sliding window which was explained. Mainly classification can be divided as follows;

1. HOG features based
   (a) SVM
   (b) MLP
2. CNN based method

CNN is a multilayer feed forward neural network which uses its learning to extract invariant multi stage features from the image data. By cause of its own ability to extract features, CNN requires less pre-processing than other classifiers. For SVM and MLP classifiers HOG features were used as these classifiers need pre-processing to extract features from the image.

3.3.1 HOG feature based method

HOG features describe the direction of the edge distribution of an image. These features are implemented by dividing image into small regions and compiling a histogram of edge orientations for pixels within each and every region that is being divided. These features are then combined in order to capture local shape properties. HOG has a main advantage which is its robustness against noise. Figure 3.19 is an example process of HOG feature extraction.

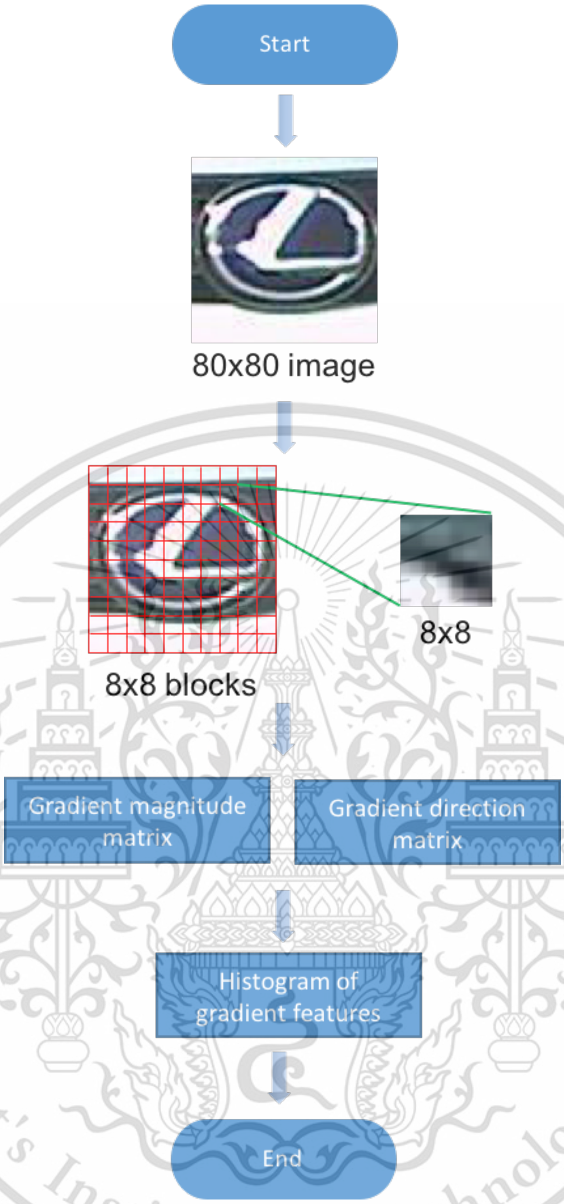

Figure 3.19: HOG feature extraction

Input image from sliding window approach were re-sized in to 80×80 pixels as to be a standard throughout the logo localization method. 8×8 pixels block size was used to calculate HOG features. To calculate the final feature vector for the entire input image, all the histogram of oriented gradients were concatenated into one vector. This feature vector was used with SVM and MLP classifiers to determine if the extracted window (from the sliding window approach) is a positive logo.

3.3.2 CNN based method

As explained in Section 2.2.3, CNN is a multilayer feed forward architected which use its learning to extract invariant multi stage features from image data. For vehicle logo localization, CNN has an architecture of two convolution layers each followed by a max pooling layer. Pooling layers were used to minimize the features. Drop out layers were used to make the CNN model robust against geometrical distortion. A fully connected layer was used to learn nonlinear combinations for high level features from convolution layers. Table3.1 shows the information of the CNN structure.

Table 3.1: Logo localization CNN structure

Parameter	Value
Input image size	80×80 pixels
No. of convolution layers	2
max pooling	every layer
Drop out	every layer
Batch size	32
Epoch	10

3.4 Logo classification

After localization of the logo, this work focused on classifying the vehicle make by vehicle logo. Logos were manually extracted as in Section 3.3. This work focused on classifying 12 logo classes. Each class has around 600 images for training purpose. Manual extraction of the logo area and some samples of training logo images are illustrated in Figs.3.20 and 3.21 respectively.

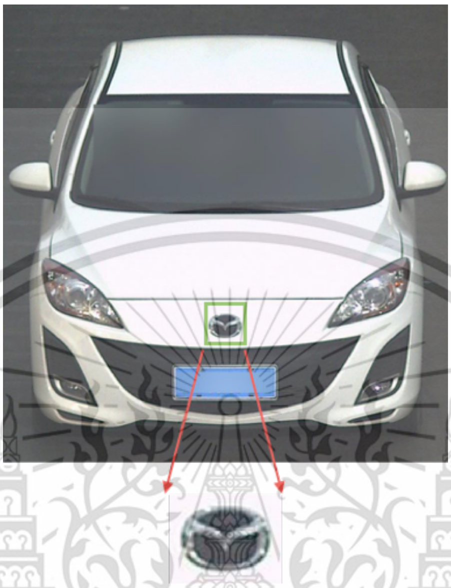

Figure 3.20: Manual extracting of logo area

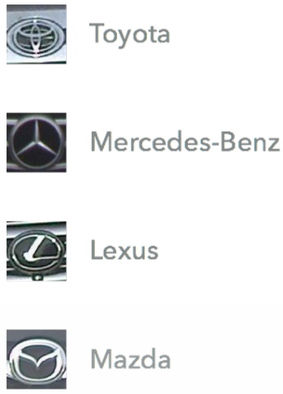

Figure 3.21: Sample of logos

The modeled CNN has an architecture of 5 hidden layers each followed by a max pooling layer to reduce the resolution. Each layer consists of convolution layers. Dropout layers are added in order to improve the robustness of the CNN model. Dropout layers are used to deactivate neurons in the training process to improve the generalization, by forcing the layer to learn with different neurons. Finally, after several convolutional and max pooling layers, the high-level reasoning in the neural network is done via fully connected layer. Table 3.2 shows the information of modeled CNN structure.

Table 3.2: Logo classification CNN structure

Parameter	**Value**
Input image size	80×80 pixels
No. of convolution layers	5
max pooling	every layer
Drop out	every layer
Batch size	32
Epoch	10

Chapter 4

Experimental results and discussion

Proposed methods were tested in an experimental manner in order to gain a good accuracy with each method. Moreover, test results were compared to find out the best method out of proposed methods. This section explains the experiment setup and results that were achieved from testing.

4.1 Image datasets

For vehicle color classification, Thailand traffic cam dataset [24] was used and for vehicle color classification COMPCAR dataset [21] was used. Table 4.1 shows details of these two datasets. Moreover, Figs. 4.1 and 4.2 show examples of each datasets.

Table 4.1: Information of datasets

Parameter	Thailand traffic cam dataset	COMPCAR dataset
Image size	1920×1080	Cropped images from 1920×1080
No. of images	5500	8500
Lighting condition	Day time	
Captured view	Front view	

Figure 4.1: Example image Thailand traffic cam dataset

Figure 4.2: Example image COMP CAR dataset

4.2 Experiment 1: Evaluation of vehicle color recognition methods

4.2.1 Objective

To find out the best classification method from each proposed method to accurately classify vehicles with respect to its body color. This experiment mainly aims on color classification of the entire vehicle by assuming the vehicle is single colored.

4.2.2 Experiment set-up

As explained in Section 4.1, Thailand traffic cam dataset was used in this experiment set-up. Figure 4.3 shows the color distribution of the dataset.

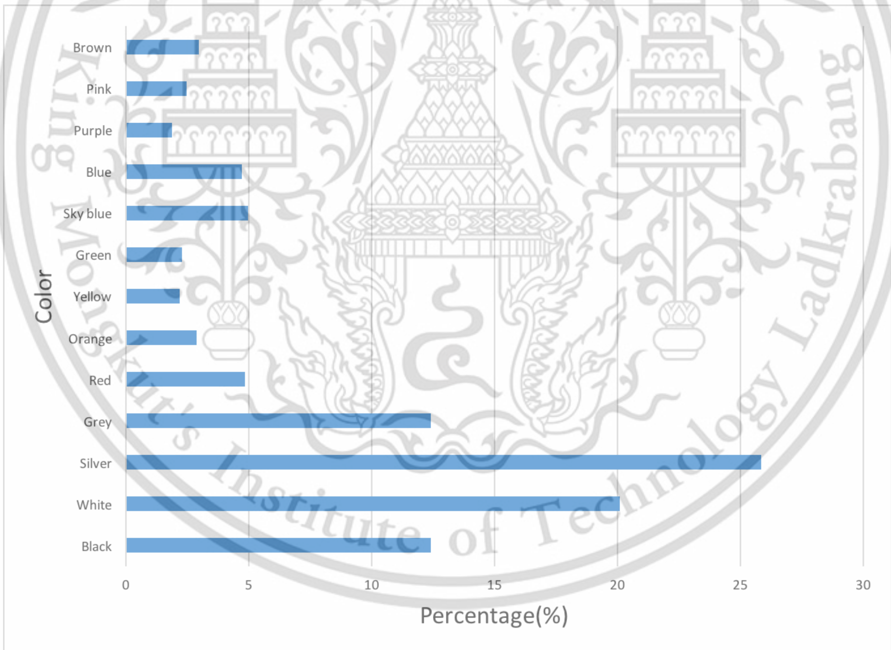

Figure 4.3: Vehicle color distribution in dataset

For SVM, RBF kernel was used and in KNN, K value of 13 was used. KNN and SVM classifiers were tested with 2500 images from Thailand traffic cam dataset in order to find out the best method. The accuracy of each method was calculated from Eqn. 4.4 where, $n_c = no.\ of\ correctly\ recognized\ images$ and $n = total\ no.\ of\ images$

$$Accuracy = \frac{n_c}{n} \tag{4.1}$$

4.2.3 Results and discussion

After testing, it was found out that automated ROI based method performs well in segmenting vehicle metallic area, that has a brighter color. Figures 4.4 and 4.5 show examples of correct segmentation and wrong segmentation respectively.

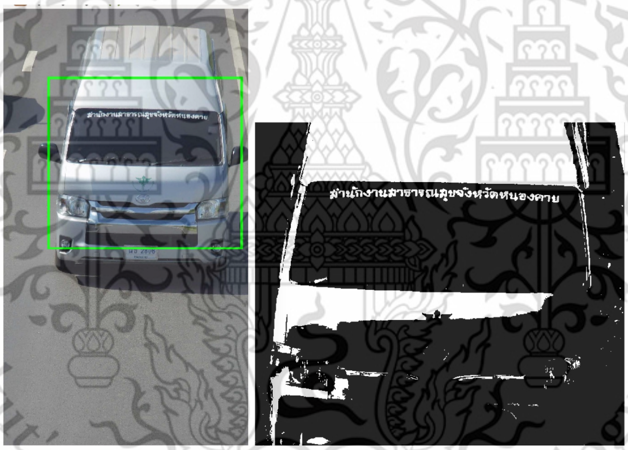

Figure 4.4: Example of correct output from automated ROI based method

Figure 4.5: Example of wrong output from automated ROI based method

Test was done on histogram back projection based method. This method was able to segment vehicle metallic area successfully for most of the colors. However, this method was unsound in segmentation of white colored vehicle metallic area. Figures 4.6 and 4.7 show examples of correct segmentation and wrong segmentation in histogram back projection based method.

Figure 4.6: Correct output from histogram back projection based method

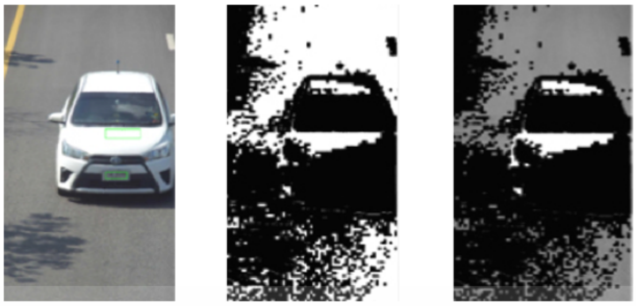

Figure 4.7: Wrong output from histogram back projection based method

Mainly 3 tests were done in order to evaluate classification methods as shown in Tables 4.2, 4.3 and 4.4. When the test results are compared, there is a big difference among grey, silver and white color classification in each method. Out of these three tests, SVM classifier gave the best results in classification among gray, silver and white colors as shown in Table 4.4.

Table 4.2: Confusion matrix of KNN classifier with 4 features

		Actual													
		Brown	Pink	Purple	Blue	Sky Blue	Green	Yellow	Orange	Red	Grey	Silver	White	Black	Total
Prediction	Brown	**72**	0	2	0	0	0	0	0	0	0	0	0	3	77
	Pink	0	**56**	0	0	0	0	0	0	0	0	0	0	0	56
	Purple	1	0	**43**	0	0	0	0	0	0	0	0	0	0	44
	Blue	0	0	0	**113**	1	0	0	0	0	2	0	0	0	116
	Sky Blue	0	0	0	3	**121**	1	0	0	0	0	8	1	2	136
	Green	0	0	0	0	0	**54**	1	0	0	0	0	0	0	55
	Yellow	0	0	0	0	0	1	**53**	3	0	0	0	0	0	57
	Orange	0	0	0	0	0	0	1	**68**	1	0	0	0	0	70
	Red	2	1	5	0	0	0	0	1	**121**	0	0	0	2	132
	Grey	0	0	0	0	0	0	0	0	0	**242**	84	1	3	330
	Silver	0	0	0	2	2	1	0	0	0	45	**398**	182	3	633
	White	0	0	0	0	0	0	0	0	0	21	152	**322**	0	495
	Black	0	0	0	0	0	0	0	0	0		2	0	**297**	299
	Total	75	57	50	118	124	57	55	72	122	310	644	506	310	**2500**

Table 4.3: Confusion matrix of SVM classifier with 4 features

		Actual													
		Brown	Pink	Purple	Blue	Sky Blue	Green	Yellow	Orange	Red	Grey	Silver	White	Black	Total
Prediction	Brown	**71**	0	2	0	0	0	2	3	2	8	0	0	5	93
	Pink	0	**57**	1	0	0	0	0	0	0	0	0	0	0	58
	Purple	1	0	**42**	0	0	0	0	0	3	0	0	0	0	46
	Blue	0	0	0	**105**	5	0	0	0	0	0	0	0	0	110
	Sky Blue	0	0	0	6	**108**	0	0	0	0	0	7	3	0	124
	Green	0	0	0	0	0	**53**	0	1	0	0	0	0	0	54
	Yellow	0	0	0	1	0	3	**50**	0	1	0	0	0	0	55
	Orange	1	0	1	0	0	0	3	**66**	0	0	0	0	0	71
	Red	2	0	4	0	0	0	0	2	**112**	0	0	0	0	120
	Grey	0	0	0	3	2	0	0	0	0	**235**	73	18	2	333
	Silver	0	0	0	2	8	1	0	0	0	24	**406**	87	1	529
	White	0	0	0	1	1	0	0	0	0	2	158	**398**	0	560
	Black	0	0	0	0	0	0	0	0	4	41	0	0	**302**	347
	Total	75	57	50	118	124	57	55	72	122	310	644	506	310	**2500**

Table 4.4: Confusion matrix of SVM classifier with 6 features

		Actual													
		Brown	Pink	Purple	Blue	Sky Blue	Green	Yellow	Orange	Red	Grey	Silver	White	Black	Total
Prediction	Brown	**72**	0	1	0	0	0	0	0	0	2	0	0	6	81
	Pink	0	**56**	0	0	0	0	0	0	0	0	0	0	0	56
	Purple	1	0	**44**	0	0	0	0	0	1	0	0	0	1	47
	Blue	0	0	0	**111**	4	0	0	0	0	0	0	0	0	115
	Sky Blue	0	0	0	3	**108**	0	0	0	0	0	3	3	0	117
	Green	0	0	0	0	0	**53**	2	0	0	0	0	0	0	55
	Yellow	0	0	0	0	0	3	**51**	1	0	0	0	0	0	55
	Orange	0	0	0	0	0	0	1	**65**	2	0	0	0	0	68
	Red	2	1	3	0	0	0	0	6	**115**	0	0	0	0	127
	Grey	0	0	0	0	0	0	1	0	0	**268**	34	13	7	323
	Silver	0	0	0	0	8	0	0	0	0	22	**524**	65	1	620
	White	0	0	0	0	4	1	0	0	0	0	83	**425**	0	513
	Black	0	0	2	4	0	0	0	0	4	18	0	0	**295**	323
	Total	75	57	50	118	124	57	55	72	122	310	644	506	310	**2500**

Color recognition methods are compared in the Table 4.5. SVM based color classification was able to classify vehicle color much accurately than the other methods. As SVM provides a more complicated and robust model, the effect of lighting variation did not affect in classifying the correct color. Moreover, using 6 features for color classification showed better results as feature description is good for on-color representation. Additionally, in histogram back projection method the ROI provides a more accurate and larger area which helps the classifiers to achieve a result with better accuracy.

Table 4.5: Vehicle color recognition method comparison

Method	Accuracy (%)
Automated ROI based method	64.86
Histogram and back projection based method	78.34
SVM based method (4 features)	80.14
SVM based method (6 features)	87.41

4.3 Experiment 2: Evaluation of vehicle logo localization

4.3.1 Objective

To find out the best classification method from each proposed method. Moreover, to find out how the final accuracy get changed with respect to the initial window size and window scaling size of the sliding window approach.

4.3.2 Experiment set-up

As explained in Section 4.1, COMPCAR dataset was used in this experiment set-up. For SVM and MLP classifiers HOG features were used as explained in Section 3.3.1.

Moreover, CNN classifier was trained with images directly as CNN can extract its own features by learning. All the classifiers were trained using positive and negative datasets. Table 4.6 shows information of the dataset.

Table 4.6: Dataset information

Parameter		Value
Dataset size	Positive	700 images
	Negative	1000 images
Image size	Original size	50×50 - 90×90
	Re-sized	80×80
Lighting condition		Day time

SVM model was trained using RBF kernel with the training dataset described in Table 4.6. Positive data corresponds to logo images which were manually extracted. Negative data corresponds to non-logo images which were extracted around the vehicle front grill area excluding logo area. Moreover, MLP classifier was modeled with 5 layers including the input layer, 3 hidden layers and the output layer. Back-propagation method was used to train the MLP model. As the activation function, sigmoid function was used. Moreover, CNN classifier was modeled using Keras library (Source: `https://keras.io`) with the use of TensorFlow backend (Source: `https://www.tensorflow.org`). Batch size of 32 and 10 epochs were used in training the CNN classifier. For the evaluation of each method Eqs. 4.2 and 4.3 were used. Distance was measured from the midpoint of original logo to the midpoint of localized logo to determine if the logo localization was correct.

$$d = \sqrt{(x-x_r)^2 + (y-y_r)^2} \tag{4.2}$$

$$Accuracy = \frac{n_c}{n} \tag{4.3}$$

Where,

- $d = distance\ from\ midpoint\ of\ the\ actual\ logo$
- $(x,y) - midpoint\ of\ the\ actual\ logo$
- $(x_r,y_r) - midpoint\ of\ the\ recognized\ logo$
- $n_c = no.\ of\ correctly\ recognized\ images$
- $n = total\ no.\ of\ images$

4.3.3 Results and discussion

Several tests were conducted using 1,000 images to evaluate logo localization methods. Table 4.7 shows the effect of the initial window size of sliding window approach on each logo localization accuracy.

Table 4.7: Effect on initial window size on localization accuracy

Method	**Accuracy according to initial window size (%)**			
	40 × 40	**50 × 50**	**60 × 60**	**70 × 70**
SVM based method	68.24	82.23	75.39	62.15
MLP based method	73.48	78.93	83.66	76.43
CNN based method	85.23	88.72	83.52	75.22

SVM based method and CNN based method show a good accuracy when the initial windows size is 50×50 pixels while, MLP based method performs well when the initial windows size is 60×60 pixels. The initial window size highly affects the logo localization accuracy because of variations in vehicle logo sizes.

In sliding windows approach, the window scales itself when a logo was not found after scanning throughout the ROI. This scaling was done by adding n number of pixels to the previous window width and height. Table 4.8 shows the results of the test that was conducted in order to determine the effect of this no. of scaling pixels (n).

Table 4.8: Effect of window scaling size on localization accuracy

Method	**Accuracy according to window scaling size (%)**			
	2	**4**	**6**	**8**
SVM based method	73.65%	77.24%	82.23%	77.74%
MLP based method	81.23%	83.66%	81.45%	74.86%
CNN based method	85.86%	88.72%	84.63%	82.26%

CNN based method and MLP based method perform well when the window was scaled with 4 pixels where SVM based method performs well when the window was scaled with 6 pixels. The accuracy mainly depends on the noise of the input image which mostly is the background of the logo. For an example, a logo of size 53×53 pixels and the initial window size is 50×50 pixels, the initial window will not capture the full area of the logo. Hence, if the window is scaled with 8 pixels, the new windows size will be 58×58 pixels. In the occasion of a window scaling size is too big the noise will be high if the logo is smaller than the window size. Therefore this window will contain more background which will affect on the classification.

Table 4.9 shows that, SVM based method had its best accuracy when the initial window size is 50×50 pixels and window scaling size of 6 pixels. Furthermore, MLP based method got its best accuracy when the initial window size is 60×60 pixels and window scaling size of 4 pixels. CNN based method shows the highest accuracy in logo localization. CNN performed well when the initial window size is 50×50 pixels and window scaling size of 4 pixels.

Table 4.9: Highest accuracy of each method

Method	Accuracy (%)
SVM based method	82.23
MLP based method	83.66
CNN based method	88.72

4.4 Experiment 3: Evaluation of vehicle logo classification

4.4.1 Objective

This experiment was done to evaluate the accuracy of logo classification method that was implemented in this work. Moreover, experiments were done in order to determine the logo classification accuracy with automatically cropped logos and manually cropped logos.

4.4.2 Experiment set-up and results

COMPCAR dataset was used in this experiment set-up. This work focused on classifying 12 logo classes. Figure 4.8 show the distribution of vehicle logos in the dataset.

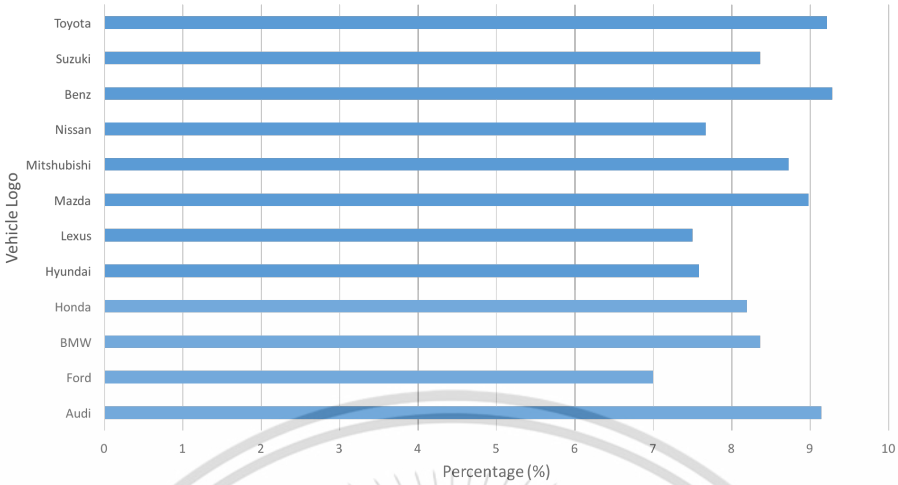

Figure 4.8: Vehicle logo distribution in dataset

CNN classifier was modeled using Keras library (Source: `https://keras.io`) with the use of TensorFlow backend (Source: `https://www.tensorflow.org`). For each vehicle logo class 600 images were used to train the CNN classifier with a batch size of 32 and 10 epochs. For the evaluation of the method Eqn. 4.4. was used, where $n_c = no.\ of\ correctly\ recognized\ images$ and $n = total\ no.\ of\ images$.

$$Accuracy = \frac{n_c}{n} \tag{4.4}$$

4.4.3 Results and discussion

Tests were done using COMPCAR dataset and Thailand traffic cam dataset in order to find out the robustness of the developed method. Tables 4.10 and 4.11 show the confusion matrix of testing. Testing results show that the proposed method is robust as tests showed an accuracy of 97.65% and an accuracy of 94.69% in COMPCAR dataset and Thailand traffic cam dataset respectively.

Table 4.10: Confusion matrix of logo classification with manually cropped logos (Thailand traffic cam dataset)

		Actual												
		Audi	Ford	BMW	Honda	Hyundai	Lexus	Mazda	Mitsubishi	Nissan	Benz	Suzuki	Toyota	Total
Prediction	Audi	**52**	0	2	0	0	1	0	0	0	0	0	1	56
	Ford	0	**134**	4	0	0	0	0	1	0	1	0	0	140
	BMW	0	0	**80**	0	0	0	1	0	0	0	0	0	81
	Honda	1	2	0	**78**	2	1	0	1	1	0	0	0	86
	Hyundai	6	1	0	3	**104**	0	1	0	1	1	1	0	118
	Lexus	0	0	0	0	0	**36**	0	0	0	0	0	0	36
	Mazda	4	0	0	0	0	1	**82**	1	0	1	0	0	89
	Mitsubishi	0	1	2	0	1	0	0	**120**	0	0	0	0	124
	Nissan	2	0	0	1	0	0	0	0	**117**	0	0	0	120
	Benz	0	2	0	1	0	0	1	1	0	**91**	0	0	96
	Suzuki	0	2	0	0	1	0	0	0	1	0	**112**	0	116
	Toyota	1	1	0	0	0	3	0	0	0	0	0	**143**	148
	Total	66	143	88	83	108	42	85	124	120	94	113	144	**1210**

Table 4.11: Confusion matrix of logo classification with manually cropped logos (COMP CAR dataset)

		Actual												
		Audi	Ford	BMW	Honda	Hyundai	Lexus	Mazda	Mitsubishi	Nissan	Benz	Suzuki	Toyota	Total
Prediction	Audi	**112**	0	0	0	0	1	0	0	0	0	0	1	114
	Ford	0	**92**	2	0	0	0	0	0	0	1	0	0	95
	BMW	0	0	**106**	0	0	0	0	0	0	0	0	0	106
	Honda	1	0	0	**105**	2	1	0	0	1	0	0	0	110
	Hyundai	1	1	0	0	**96**	0	0	0	1	0	1	0	100
	Lexus	0	0	0	0	0	**93**	0	0	0	0	0	0	93
	Mazda	2	0	0	0	0	1	**96**	1	0	1	0	0	101
	Mitsubishi	0	1	0	0	0	0	0	**109**	0	0	0	0	110
	Nissan	2	0	0	1	0	0	0	0	**96**	0	0	0	99
	Benz	0	0	0	0	0	0	1	1	0	**118**	0	0	120
	Suzuki	0	0	0	0	0	0	0	0	1	0	**107**	0	108
	Toyota	2	1	0	0	0	1	0	0	0	0	0	**118**	122
	Total	120	95	108	106	98	97	97	111	99	120	108	119	**1278**

Table 4.12 shows that the logo classification achieved a good accuracy with proposed CNN based method. If the automatically cropped logo contains a partial area of the logo, it will affect on the logo classification accuracy. However, if an image containing 80% of the logo area is fed into the system, method will successfully classify the vehicle logo.

Table 4.12: Logo classification accuracy

Method	**Accuracy (%)**
Automatically cropped	82.45
Manually cropped	97.65

Chapter 5

Conclusion and future works

This chapter concludes the thesis and gives some ideas for future research. This chapter explains about the final conclusion of the work done and how this work can be developed in future projects so that the vehicle identification will be more accurate and robust.

5.1 Conclusion

As a solution for vehicle identification, this work proposed methods for vehicle color and logo identification. This work proposed two methods in vehicle color classification. Moreover, a sliding window approach was proposed in logo localization. Thus, two logo localization methods were proposed to use with this sliding window approach to localize the correct position of logo. Vehicle logo identification was done using CNN based method.

The color recognizing techniques discussed in the proposed methods were able to successfully recognize vehicle color. SVM based approach was the most effective method with an accuracy of 87.41% from all proposed methods. Moreover, from proposed logo localization methods, CNN based method was the most productive one with an accuracy of 88.72%. Finally, CNN based approach was able to successfully classify vehicle logos with an accuracy of 97.65%.

5.2 Future works

In future, vehicle color and logo information can be used with the license plate recognition system and can be tested to achieve better accuracy in vehicle identification. Also, this work can be modified or developed to recognize the vehicle make as well as the model which will be beneficial in vehicle identification.

Bibliography

[1] T. Anakavej, A. Kawewong and K. Patanukhom, "Internet-vision based vehicle model query system using eigenfaces and pyramid of histogram of oriented gradients," Proceedings of 2013 International Conference on Signal-Image Technology and Internet-Based Systems, pp.179–186, 2013.

[2] Bay, Herbert, Ess, Andreas, Tuytelaars, Tinne, V. Gool and Luc, "Speeded-Up Robust Features (SURF)," Journal on Computer vision and image understanding, vol. 110, no. 3, pp. 346-359, 2008.

[3] W. Y. Chen, H. C. Chuan, H. C. Ta and F. K.Chin "Vehicle color classification using manifold learning methods from urban surveillance videos," EURASIP Journal on Image and Video Processing, vol. 2014, no.1, pp.48–50, 2014

[4] H. Chul, Kim, Daijin, Bang and S. Yang, " Face recognition using the mixture-of-eigenfaces method," Journal of pattern recognition letters, vol. 23, no. 13, pp. 1549–1558, 2002.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 886-893, 2005.

[6] C. Dhaware and K. H. Wanjale "Survey on image classification methods in image processing," International Journal of Computer Science Trends and Technology (IJCS T), vol. 4, no. 3, pp.246–248, 2016.

[7] S. Draghici, "A neural network based artificial vision system for license plate recognition," International Journal of Neural Systems. vol. 8, no. 1 pp.113–126, 1997.

[8] E. Dule, M. Gokmen and M. S. Berayoglu "A convenient feature vector construction for vehicle color recognition," Recent advances in neural networks, fuzzy systems and evolutionary computing, pp.250–255, 2008.

[9] A. Faheema and S. Rakshit, "Feature selection using bag-of-visual-words representation," 2010 IEEE 2nd International Advance Computing Conference, pp. 151-156., 2010.

[10] Hearst and A. Marti, "Support Vector Machines," Journal of IEEE Intelligent Systems, vol. 13, no. 4, pp. 18-28 1998.

[11] D. H. Hubel, T. N. Wiesel,"Receptive fields and functional architecture of monkey striate cortex", The Journal of Physiology, vol. 195, no. 1, pp. 215243. 1968.

[12] T. Kryjak, M. Komorkiewicz and M. Gorgon, "Real-time hardwaresoftware embedded vision system for ITS smart camera implemented in Zynq SoC," Journal of Real-Time Image Processing, vol. 11, no. 4, pp.1–37, 2016.

[13] G. Kumar, P. K. Bhatia, "A detailed review of feature extraction in image processing systems," Proceedings of 4th International Conference on Advanced Computing and Communication Technologies, pp.5–12, 2014.

[14] Y. Liu and S. Li, "A Vehicle-logo location approach based on edge detection and projection," Proceedings of 2011 IEEE International Conference on. IEEE, pp.165–168, 2011.

[15] D. F. Llorca, R. Arroyo and M. A. Sotelo, "Vehicle logo recognition in traffic images using HOG features and SVM," Proceedings of IEEE Conference on Intelligent Transportation Systems (ITSC), pp.2229–2234, 2013.

[16] D. G Lowe, "Object recognition from local scale-invariant features", Proceedings of the 7th IEEE International Conference on Computer Vision, pp.1150-1157, 1999.

[17] R. A. Rosario and N. Tepedelenlioglu, "A rapid multi-layer perceptron training algorithm," Proceedings of IJCNN International Joint Conference on Neural Networks, pp. 824-829, 1992.

[18] M.J. Swain, D.H. Ballard (1992) Indexing via Color Histograms. In: A.K. Sood , H. Wechsler (eds) Active Perception and Robot Vision. NATO ASI Series (Series F: Computer and Systems Sciences), vol 83. Springer, Berlin, Heidelberg

[19] S. Sittampalam and A. Ramanan, "A classifier-free codebook-based image classification of vehicle logos," Proceedings of 9th International Conference on Industrial and Information Systems (ICIIS 2014), pp.1–6, 2015.

[20] R. Verma, "A leaf recognition algorithm for plant classification using probabilistic neural network," Proceedings of 2007 IEEE International Symposium on Signal Processing and Information Technology, pp.532–234, 2007.

[21] L. Yang, P. Luo, C. C. Loy, X. Tang. “A large-scale car dataset for fine grained categorization and verification,” in Computer vision and pattern recognition (CVPR), 2015.

[22] S. Yu, S. Zheng, H. Yang and L. Liang, “Vehicle logo recognition based on Bag-of-Words,” 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2013), pp.353–358, 2013.

[23] https://en.wikipedia.org/wiki/HSL_and_HSV#cite_note-FR841335-3

[24] “Thailand traffic cam dataset,” NECTEC, National Science and Technology Development Agency Thailnd.

Appendices

Appendix A

Publication

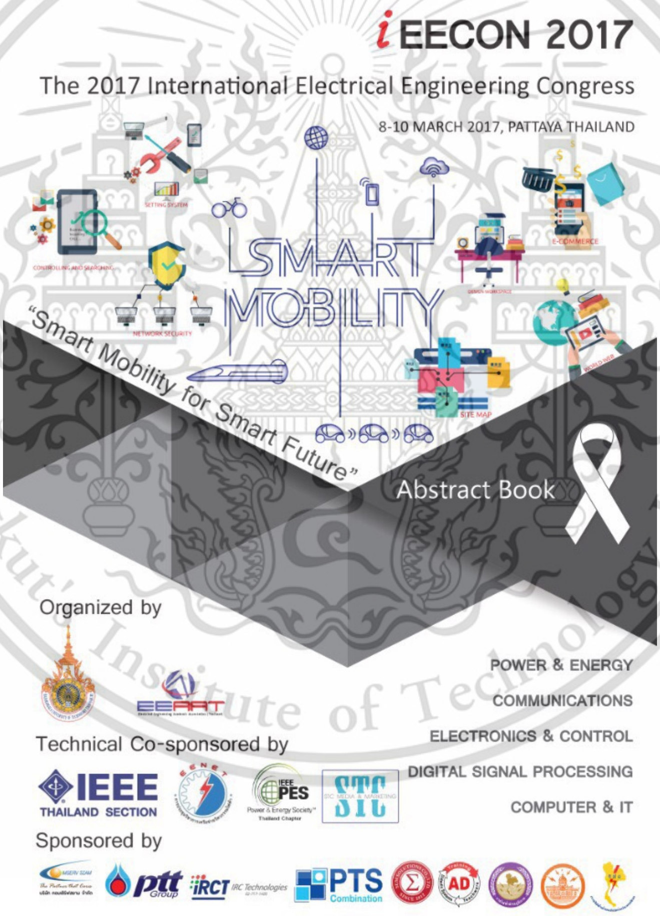

Image Analysis Algorithms for Vehicle Color Recognition

Damitha S.B. Tilakaratna and Ukrit Watchareeruetai
International College
King Mongkut's Institute of Technology Ladkrabang
Chalongkrung Road, Ladkrabang,
Bangkok 10520 Thailand
damithasbt@gmail.com, ukrit.wa@kmitl.ac.th

Supakorn Siddhichai and Nattachai Natcharapinchai
Image Technology Laboratory, NECTEC
National Electronics and Computer Technology Center
Phahonyothin Road, Khlong Nueng, Khlong Luang,
Pathumthani, 12120, Thailand
supakorn.siddichai, nattachai.watcharapinchai@nectec.or.th

***Abstract*— This work discusses about the implementation of a vehicle color recognition system to be used with the vehicle license plate recognition system. Because of the complexity of Thai alphabet, current license plate recognition system fails in correctly recognizing vehicles that have the same number but different Thai characters. Also the current system cannot identify illegal number plates that are being used. This color recognition system will help to resolve these problems and increase the accuracy of the vehicle recognition system. In addition, this system will provide a wide range of color classification which includes 13 colors including white, silver and gray colors. Color recognition is done by two methods: one with machine learning and one without machine learning. A best accuracy of 87.52% is given when using SVM to classify colors.**

Keywords*— *vehicle color recognition; color histogram; back projection; support vector machines

I. Introduction

With the development in automobile technology, digital image processing has been implemented in the industry. As a result of this autopilot vehicles have been introduced to the market. Some major automobile companies have implemented an adaptive suspension system with the use of digital image processing.

Digital image processing has also been applied in safety and security systems. A system for vehicle license plate recognition for police vehicles has been developed to recognize vehicle license plate and then with the use of database it can retrieve driver's information. The system helps the police officer to identify the driver details easily.

Vehicle license plate recognition system has been implemented in Thailand. Vehicle color recognition is one of the effective ways of vehicle recognition. Some researchers have proposed a color recognition system using support vector machines (SVM) [1] and convolutional neural network (CNN) [2] which is good in rich colors but there method do not include white and silver color classification. Several well-known detection approaches [3-4] can provide an accurate bounding box for each vehicle.

One of the main reasons to exclude white, silver and gray colors in color classification is that pixel values of those three colors are so close to each other and cannot be classified using single color space. Although there is a high percentage of vehicles designed with white, silver and black colors, the researches conducted have not classified these three major colors for vehicle color recognition.

This work is to implement a vehicle color recognition system which can classify up to 13 colors including white, silver and gray colors to work with the current vehicle license plate recognition system and improve the system's accuracy.

The remaining of this paper is organized as follows: Section II describes the dataset used in this paper. Section III proposes two different methods for vehicle color recognition. Section IV presents and discusses experimental results. Section V concludes the paper.

II. Materials and Preparation

A. Dataset

Image datasets used in this paper were collected using traffic cameras which are mounted on Thailand motor ways. These cameras are used for the current License plate recognition (LPR) systems that have been deployed in Thailand. As a reference license plate information is extracted from the current LPR system to work with the color recognition system. Information about the data set is as follows:

- Dataset size – 7,000 images
- Test Dataset size – 2,500 images
- Image size – 1,920 × 1,080 pixels
- Condition – Day time
- Data source – Highway Traffic Camera

As shown in Table 1, this system has to classify 13 colors, including black, white, silver and gray and other popular vehicle colors in Thailand. This list was prepared with the information received from the vehicle information database at Thailand land transportation department.

TABLE I. VEHICLE COLOR DISTRIBUION IN DATASET

Color	Number	Percentage (%)
Silver	646	25.84
White	502	20.08
Black	310	12.40
Grey	310	12.40
Sky blue	124	4.96
Red	122	4.88
Blue	118	4.72
Brown	74	2.96
Orange	72	2.88
Pink	62	2.48
Green	58	2.32
Yellow	55	2.20
Purple	47	1.88
Total	**2,500**	**100.00**

B. *The Input*

The information regarding the vehicle plate and the image that is captured will be referenced from LPR system for the color and make recognition system implementation. This is because the current LPR system can accurately detect moving vehicles and localize the vehicle number plate of the image. As shown in Figure 1, the system then captures the frame and stores the license plate information (width and height) as well as the coordinate values of the license plate (top left corner x, y).

C. *Region of Interest*

With the use of the top left corner coordinates of the license plate, the system defines a region of interest (ROI) right above the number plate with the dimension of 50 × 200 pixels as shown in Figure 1.

Figure 1. Selection of ROI

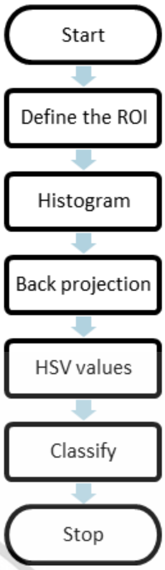

Figure 2. The flowchart of the histogram and back projection based method

III. COLOR RECOGNITION METHODS

As to successfully classify 13 colors, which was mentioned in Table 1, this work has implemented two main methods, as follows:

- Histogram and back projection based method
- Support vector machines (SVM) based method

These two methods use the same data set and same license plate data information. The first method is done without machine learning while the second method is based on machine learning. Each method has its own advantages as well as disadvantages. This work will compare these two methods in the aspect of accuracy.

A. *Method I: Histogram and Back Projection Based Method*

This method was implemented without any machine learning method. The model for each color was found after doing some research on color values of vehicle images. The method can be described by a flowchart, as shown in Figure 3.

As explained earlier, an ROI of size 50 × 200 pixels will be extracted from the hood area of the vehicle (Figure 3). Then the histogram is calculated for the extracted ROI in HSV color space (Figure 4). Then the histogram will be normalized in order to make it accurate and a histogram model is created.

With the use of the histogram model of ROI, histogram back projection is done as shown Figure 5. Back projection is a way of recording how well the pixel of the given image fit the distribution of pixels in the histogram model. In other words, this system calculates the histogram model for ROI and then uses it to find this feature in the entire image.

As shown in Figure 5, back projection will give a mask which contains the same features as the histogram model of

ROI. Therefore this mask will contain a majority of the vehicle color area as shown in Figure 6. This will eliminate the background and windshield of the vehicle which is not useful for the color recognition system.

After extracting the interested colored area of the image then it will be used to classify the color. Each color has a range of predefined values for the classification. So the HSV values for every pixel in the colored area will be voted according to the color. Then the vehicle color is classified as shown in Figure 7. Some examples for vehicle color classification using this method is shown in Figure 8.

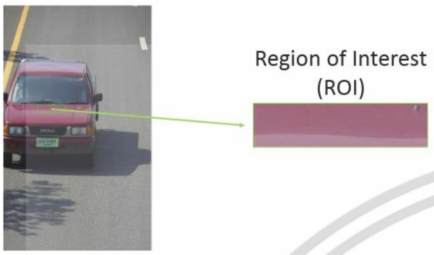

Figure 3. ROI extraction

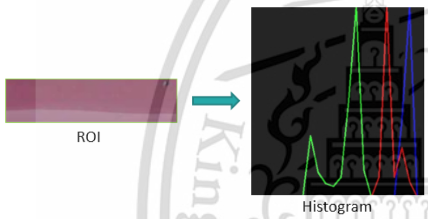

Figure 4. Model of color histogram in ROI

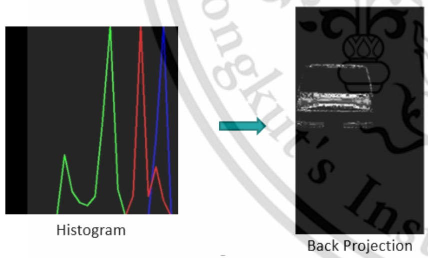

Figure 5. Example result of histogram back projection

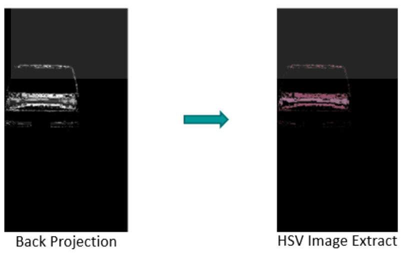

Figure 6. Extracted area

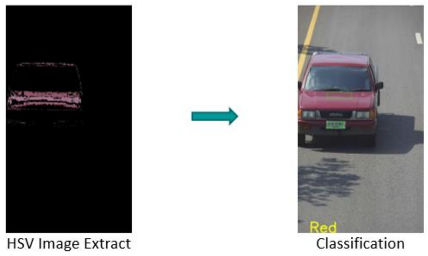

Figure 7. Color classification

Figure 8. Examples of color classificaiton result

B. Method II: SVM Based Method

SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training sample. Therefore the optimal separating hyperplane in a feature space maximizes the margin of the training data.

In this method, we used several feature combinations and carried out testing to get the results according to them. In this work, radial basis function (RBF) kernel was used in SVM classification. The same ROI was extracted in order to extract color features from the vehicle. This area will be processed with a sliding window of 10 × 10 pixels area. This sliding window will extract features from each pixel within its area (Figure 9).

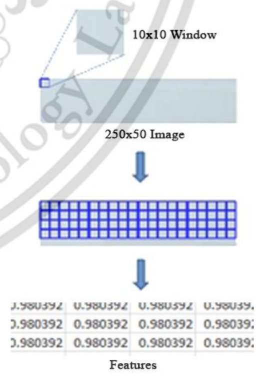

Figure 9. Feature extraction steps

In order to select accurate features for classifying colors some research was done using several features in this work. As the color classes contain with many combinations of primary colors more feature value should be selected. Each color space has its advantages and disadvantages. For an example HSV color space is good with intensity variation but poor in classification on non-colors such as black, white and silver. BGR is good with non-colors but bad with intensity. Consequently, in this work, the feature vector contains several color spaces in order to compensate errors of the each other color spaces.

In this method different feature combinations were tested in order to reach a good accuracy.

- Four features (R, G, B, V)
- Six features (R, G, B, H, S, V)

These feature vectors are classified using multi-class SVM. RBF has been used as it achieves the best accuracy when comparing with other kernels. Finally, the majority voting scheme is used to classify the final result as the class which has the largest number of votes.

TABLE II. ACCURACY OF METHODS

Color	**The number of vehicles correctly recognized (accuracy)**		
	Histogram Back Projection	***SVM Four features***	***SVM Six features***
Silver	398 (61.61)	406 (62.85)	524 (81.11)
White	322 (64.14)	398 (79.28)	425 (84.66)
Black	297 (95.81)	302 (97.42)	295 (95.16)
Grey	242 (78.06)	235 (75.81)	268 (86.45)
Sky blue	121 (97.58)	108 (87.10)	108 (87.10)
Red	121 (99.18)	112 (91.80)	115 (94.26)
Blue	113 (95.76)	105 (88.98)	111 (94.07)
Brown	72 (97.30)	71 (95.95)	72 (97.30)
Orange	68 (94.44)	66 (91.67)	65 (90.28)
Pink	57 (91.94)	57 (91.94)	57 (91.94)
Green	54 (93.10)	53 (91.38)	53 (91.38)
Yellow	53 (96.39)	50 (90.91)	51 (92.73)
Purple	43 (91.49)	42 (89.36)	44 (93.62)
Total	**1,961 (78.44)**	**2,005 (80.20)**	**2,188 (87.52)**

IV. RESULTS AND DISCUSSION

For each method various tests were done by varying its parameters and data sizes. Each and every test had its advantages and disadvantages.

In method I, histogram model with 32 bins in the range of 0-255 is used to create a mask that will then extract the colored metallic areas of the vehicle. This method is called back projection. In this method the image is compared with the ROI histogram model. When comparing this, Otsu's thresholding method [5] is used to automatically perform clustering-based image thresholding to maximize the extracted area acutely.

In this method I, it will extract the area which matches the histogram model from the ROI. So the extracted area will contain similar features as the give histogram model.

As shown in table II, the method I, is good in recognizing colors other than gray, silver and white. The reason for low accuracy is that gray, silver and white it has a small difference in pixel values in HSV color space.

In method II, SVM was used to classify each color. Experimental results are shown in Table II. In this method best accuracy was given when using six features for classification. For the case of using six features, it contains two color spaces and it successfully manages to classify colors with good accuracy for all the colors.

V. CONCLUSION

This paper proposes two methods in vehicle color recognition as well as a method to detect the vehicle metallic area using histogram back projection. SVM based classification method successfully solved the color classification on white, silver and gray color which had low accuracy in other methods that were proposed.

When using back projection method, it gives an accuracy around 78.44%. In this method, it successfully classifies colors when the lighting conditions are good. But when shadow is present on the vehicle this method fails to classify colors acutely.

In SVM maximum accuracy was given when using six features. It managed to give an accuracy of 87.52% and it managed to classify color even when it has shadows on vehicles. Overall SVM method gave the highest accuracy.

ACKNOWLEDGMENT

Firstly, I would like to express my sincere gratitude to Assoc. Prof. Masaki Yamakita of TIT for their guidance and support. Last but not least I wish to thank my parents who supported and motivated me immensely throughout my education.

REFERENCES

[1] C. Hu, X. Bai, L. Qi, P. Chen, G. Xue and L. Mei, "Vehicle Color Recognition With Spatial Pyramid Deep Learning," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 5, pp. 2925-2934, Oct. 2015.

[2] P. Chen, X. Bai, and W. Liu, "Vehicle color recognition on urban road by feature context," IEEE Transactions on Intelligent Transportation Systems, vol.15, no.5, pp.2340-2346, 2014.

[3] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part based models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.32, no.9, pp.1627-1645, 2010.

[4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), pp.511-518, 2001.

[5] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," Electronic Imaging, vol. 1, no. 13, p. 146–165, 2004.

[6] Y. Dong, M. Pei and X. Qin, "Vehicle Color Recognition Based on License Plate Color," 2014 Tenth International Conference on Computational Intelligence and Security, Kunming, 2014, pp. 264-26

AUTHOR BIOGRAPHY

Author: Mr. Damitha S.B. Tilakaratna

Degree: Master of Engineering

Date of Graduation: 16th March 2018

Date of Birth: 26th August 1992

Place of Birth: Kandy, Sri Lanka

Undergraduate and Graduate Education:

Master of Engineering in Automotive Engineering,
King Mongkut's Institute of Technology Ladkrabang, Bangkok

Bachelor degree in Mechatronics Engineering ,
Asian Institute of Technology, Bangkok

Major: Automotive Engineering

Presentations and Publications:

[1] D.S.B. Tilakaratna, U. Watchareeruetai, S. Siddhichai and N. Natcharapinchai "Image Analysis Algorithms for Vehicle Color Recognition", 2017, International Electrical Engineering Congress, Pattaya, Thailand.