TRAFFIC SIGN RECOGNITION

BY

Willy Chai Yi Xian

Wong Ci Shen

Toh Wei Xuan

Yin Zi Keen

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

BACHELOR OF COMPUTER SCIENCE (HONS)

Faculty of Information and Communication Technology
(Kampar Campus)

MAY 2022

Abstract

Done by: Wong Ci Shen, Willy Chai Yi Xian

Traffic signs are symbols or signs that covey roads information and also helps to instruct related movement of traffic on the roads. With the development of advanced technology, computer vision will soon become the essential factor in order to develop automation car driving. Traffic signs detection system will become the one of the most important contributions towards making intelligent vehicles. Image detection system, more commonly known as image processing, is a composite process as it consists of multiple fundamental stages. The major stages include image pre-processing, image detection, image segmentation and finally image classification. This paper will be focusing on the image segmentation stage, where the proposed method will be working on utilizing HSV color space to perform image segmentation alongside with k-means clustering algorithm. Image segmentation is the process of extracting the region of interest (ROI) from the image, which then allows the image to be easily processed in the classification stage. In the case of traffic sign recognition, the proposed system should be able to identify the region containing the traffic sign, perform masking to isolate the traffic sign, giving an output of a segmented image containing the detected traffic sign only. The reason of using HSV color space is that it provides more information regarding illumination and saturation of the images, compared to RGB color space where it only provides information regarding the color of the image.

Table Of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Problem Statement and Motivation

Done By: Yin Zi Keen

In this century, car is one of the greatest inventions in the world and there is a report by [1] which has stated that the number of registered vehicles in Malaysia which around 33.3 million has overtaken the population of the country which is 32.6 million . In here, there will be rise of an issue in this situation which is the vehicle accident cases is increased followed by increase in number of vehicles. Thus, traffic sign may play an important role in relieve this issue. Traffic sign may give awareness to driver such as informing them the condition of the rood or informing them animals like to appear in certain area or even informing them there are some facilities that need to take care about such as school which has many children nearby it.

However, the traffic sign may be in unclear condition such as paint stripping in traffic signpost, uneven in light condition, or due to weather condition like raining day. Hence, the driver unable to recognize the traffic sign and cannot receive the message that delivered. Furthermore, some drivers that are myopia may overlook the traffic sign so that car accident may occurs easily. Although nowadays there are some existing technologies has been used to recognize the traffic signs, but if the image captured is blur, the image captured is affected by light and weather condition, or there is damage or defect in the traffic signpost, the image still cannot be analysis completely and thus lead to the traffic sign unrecognized.

Therefore, the motivation of the project is to allow the driver to know well situation of the road so that they can drive safely and confidently. Hence, with the problem that stated above, a system for traffic sign segmentation is proposed.

1.2 Project Objectives

Done By: Yin Zi Keen

The goal of this proposed project is to allow the system to detect the traffic sign from captured image no matter the image condition, obtain sufficient data from the captured image, and classify them into correct category so that the traffic sign can be recognized correctly in the end of the process and the system can provide the information of the recognized traffic sign to driver correctly. Hence, these are some objectives respectively:-

1. Implement a segmentation technique so that the data being gained has high level of completeness which means there is no extra data or insufficient of data for recognition occurs .
2. Ensure that the segmentation technique applied is time-efficient in such a way that the operation should take less than approximately 2 seconds to complete for each image.

1.3 Project Scopes

Done By: Willy Chai Yi Xian

This project will develop a traffic sign segmentation system. The primary scope of this system is to segment the traffic sign from the captured image properly so that it can be used in image classification systems to provide information about the recognized traffic sign to the driver so that they can always know well the condition of road which make help them to make a correct decision when driving. In other words, this system may reduce vehicle accident cases directly. This project may make improvements in three phases which is detection, segmentation and classification. This is to improve the correctness of the final deliverable which is the recognized traffic sign. The recognition process is start with the detection of the traffic sign from captured image and then proceed to segmentation which ensure there is sufficient of data being segmented and lastly is classification which classify the processed traffic sign into correct category and further processed it. This proposed project is expected to be done within few weeks only. In the process of development, it is assumed that a laptop is enough for develop the system. The final deliverable of this project is a correct traffic sign. With the proposed system, although viewing condition is not well, but the driver still can get the correct message by the traffic sign that recognized by the system and can know well the condition of the road.

1.4 Impact, Significance and Contribution

Done By: Toh Wei Xuan

Innovative technologies are now being developed in the fields of image processing, notably in the area of image segmentation. The process of segmenting an image is both crucial and challenging. The most important stage of picture analysis is the segmentation procedure. Image segmentation is the process of dividing the image into homogeneous sections based on certain criteria and ideally, correlating to actual objects in the scene. Furthermore, the main objective of segmentation is to reduce data for an easier analysis procedure. The area of image processing still has unsolved and fascinating research questions around picture segmentation. For researchers and developers alike, creating a generalized method for image segmentation continues to be challenging. Meanwhile this project will encourage readers to explore more challenges on segmenting the traffic sign in color-based or shape-based features segmentation. There are many criteria needed for a good segmentation such as the area of pixels are all the part of the traffic sign, an area is considered linked if any two pixels inside a given area can be connected by a line that does not leave the area, each area is homogeneity with regard to a certain feature which might be semantically based on intensity, texture and color, not feasible to create a single homogenous area by combining adjacent areas, there is also no overlap occur between the areas [2].

Besides, this project aims to improve the accuracy and enhance the segmentation performance through preprocessing, uses several detections based on color rate and choose the faster method to propose the whole segmentation. Thus, we will use prior knowledge of traffic sign in the HSV system to achieve image segmentation based on color feature. Hue and saturation component of an HSI color space was act as the color classifier and start the segmentation. Those method is used to reduce the error color rate and the SVM algorithm also used to boost precision and reduce false positives. By having this model, users can segment the traffic sign with high accuracy without any error in color rate.

1.5 Report Organizations

Done by: Wong Ci Shen

The details of this project are shown in the following chapters. In Chapter 2, some similar existing segmentation and recognition methods are reviewed. The previous work done by other existing segmentation and recognition method can be known in this chapter. This chapter will also indicate the advantages of traffic sign method and the critical remarks of the previous work. After that, methodology, system block diagram and components specifications will be presented in Chapter 3. Methodology describes on the general work procedures of the project whereas system diagram indicates the overview of the system and describe the purpose of each method. In chapter 4, tools to use in the proposed project will be discussed and the challenges which faced during the implementation of the project will be defined. Next, chapter 5 describes the testing setup and result and the comparison with previous works. Lastly, Chapter 6 will summarize all the things that have been described in the previous chapter such as problem statement, motivation, proposed method to solve the issue and others.

Chapter 2: Literature Review

2.1 Traffic Sign Segmentation and Recognition using MACHINE Learning

2.1.1 Traffic sign segmentation and classification using statistical learning methods

Done by Willy Chai Yi Xian

The proposed system in [3] is a three-staged system which is capable of recognizing both chromatic and achromatic signs. The first stage is segmentation stage, where the image data are segmented into chromatic and achromatic groups. For chromatic segmentation, the images are segmented using L*a*b* color space, as chromatic colors can be better distinguished using the a*b* space. A k_m-means clustering algorithm was then applied to the a*b* color space to obtain a set of centroids for each representing class, namely red, green, blue, and black. Two machine learning models were applied to the dataset and compared, which are Support Vector Machine (SVM) and k-Nearest Neighbors (k-NN), with k-NN being the better machine learning model due to a low error rate and low computational burden. The dataset classified in the black class will then undergo achromatic segmentation, which uses HSI color space to utilize the saturation component S for the segmentation process. Simple thresholding is then used to perform achromatic segmentation by thresholding the saturation component S.

The second stage is post-processing stage, which mainly involves removing the non-interest regions, connecting fragmented signs and separating signs located at the same post. Different procedures were used for chromatic and achromatic groups. For chromatic post-processing, non-interest regions such as sky or vegetation classified in the blue and green group is filtered by removing the regions where the ROI is connected to the image border. Image region with very small or large ROI were also removed from the image. Morphological closing was applied to solve the sign fragmentation issue. Additional procedure has also been applied to further segment and separate co-located signs. For achromatic post-processing, to eliminate the false positive areas, contour of the images were emphasized by a variance filter and the edges were localized through thresholding. After that, the relative color difference, rcd were computed, and pixels with rcd exceeding a certain threshold were removed. Similar to chromatic post-processing, regions that were connected to outside of the image were removed, and regions with very small or large ROI were also removed.

Finally, the third stage is shape classification stage, where the segmented data are furthered grouped into six different classes, circle, triangle, square, rectangle, arrow and semicircle. Two phases were included in this stage, which are feature extraction and shape classification. For feature classification, contour-based methods were used, more specifically, the Fourier Descriptors method. This method was used due to its efficiency compared to other contour-based methods for computation. For shape classification, three different models were proposed, voting k-NN, linear SVM and nonlinear SVM. Nonlinear SVM was chosen as the final model as it has the lowest error rate out of all the proposed models. Figure 2-1-1 illustrates the entire process of segmentation, post-processing and classification of the proposed system by [3].

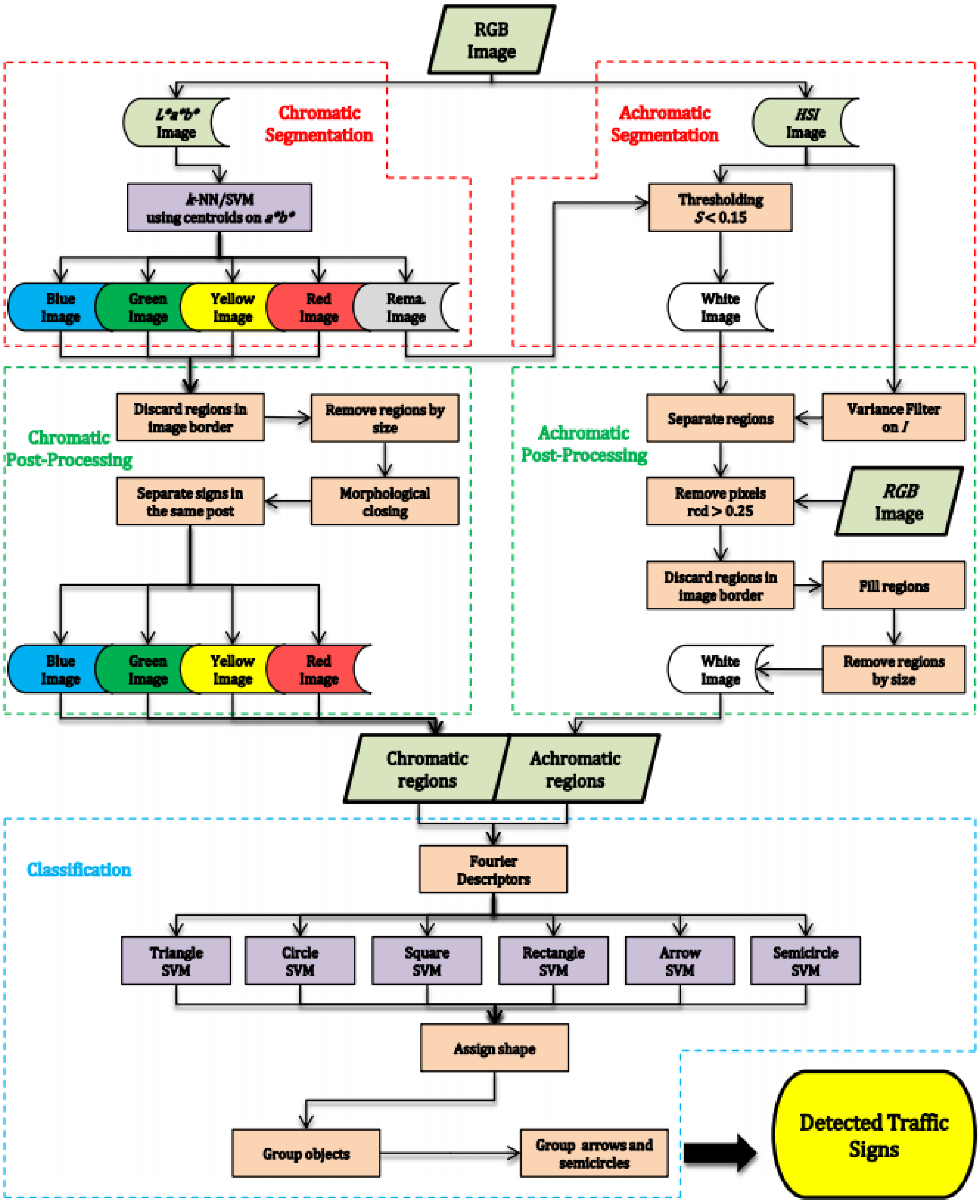


Figure 2-1-1: Block Diagram of the Proposed System [3]

2.1.2 Road Sign Detection and Recognition Based on SVM

Done by Toh Wei Xuan

In [4], algorithms for recognizing traffic signs typically include two stages which are detection and recognition. The segmentation step by thresholding with a specific color space to extract the sign color from the picture makes up the initial block of the detection system. The RGB color space used to detect strong colors in the sign. Although other space is used, the HIS system also used to provide modest variances for items of interest with a comparable color. The YUV system is explored to recognize blue rectangular signs. Even yet, there are some pieces where thresholding is not explicitly done using a particular color space. Simple vector filter (SVF) is used in thresholding by looking for chromatic and achromatic colors. Besides, SVF contains features that allow it to simultaneously remove all outlines and extract a single color.

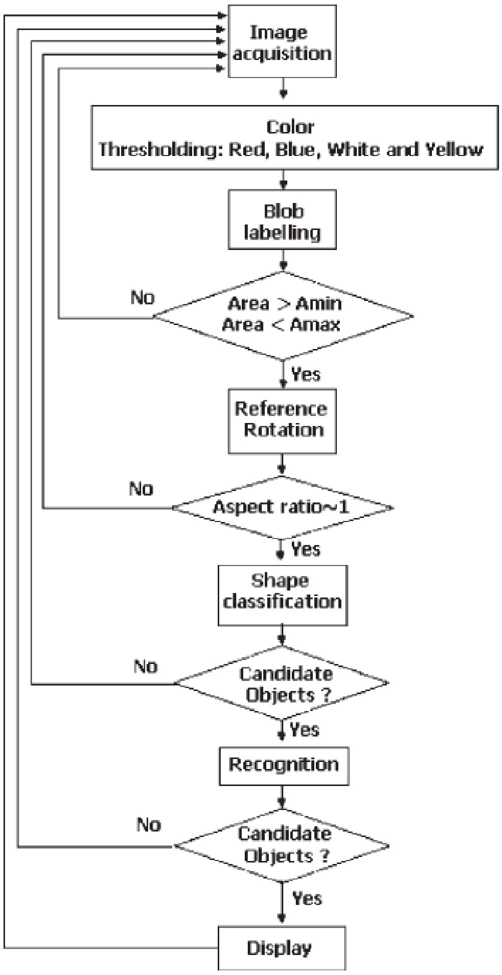


Figure 2-1-2: Overview of algorithms used based on SVM [4]

Based on edge detection, more detection techniques have been created. A shape-based algorithm based on radical symmetry applied and tailored to triangular, square, diamond, octagonal, and circular shapes and it needs operates on the gradient of grayscale image. its advantage is to operate in real time as it detects shapes based on edges, the technique is vigorous to changing the illumination images. In paper [4], the Bayes classifier selects a color based on the maximum conditional probability of each hue, assuming that the colors, red, yellow, green, blue, and white have identical a priori probabilities. Once the candidate regions have been separated from the image, the following classification will start to apply based on shape. The classification can divide based on color and shape of traffic sign such as red circular, red triangle, and blue circular. For further extraction, color extraction is supplemented with shape features using two neutral networks (NNs).

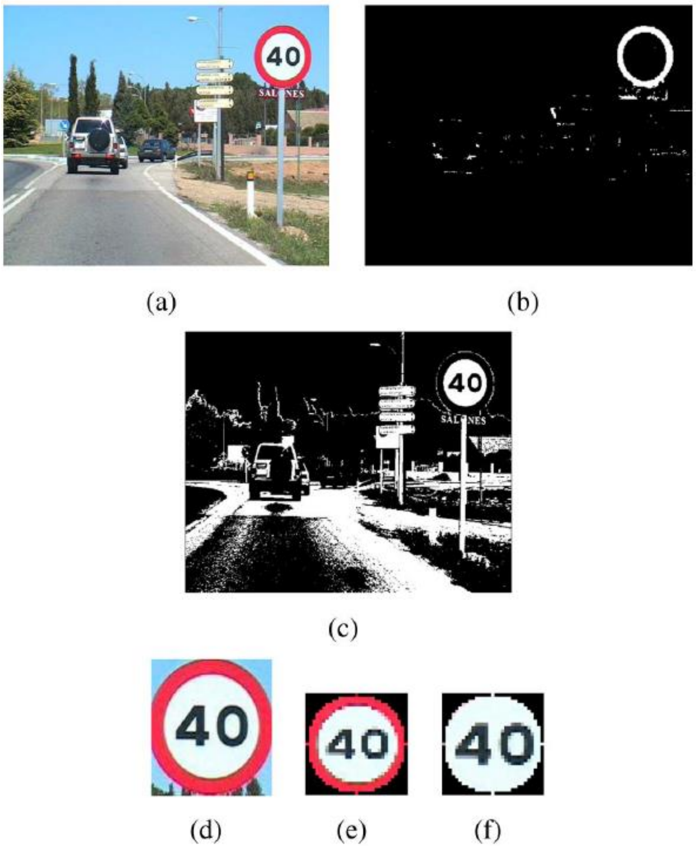


Figure 2-1-3: a) original image, b) and c) segmentation masks by red and achromatic colors, d) extracted sign, e) and f) outer and inner regions of normalized sizes [4]

2.2 Traffic Sign Segmentation and Recognition using DEEP Learning

2.2.1 Traffic Sign Recognition with Multi-Scale Convolutional Networks

Done by Willy Chai Yi Xian

Proposed method in [5] uses Convolutional Network architecture to perform the traffic sign classification. Data validation and data pre-processing were performed in the data preparation phase. The dataset images were standardized into the size of 32x32 and converted to YUV space. Next, features were extracted from a set of randomly initialized architectures with different capacities. The top classifier is then trained using these features as input, repeated with a range of different capacities. After evaluation, the optimal feature size determined are 108-108 feature size and followed by 108-200 feature size. Supervised data training is then carried out using these optimal architectures on the training set. Table (_) shows that the highest accuracy obtained from the model was 99.17% for feature size of 108-108, followed by 98.85% for feature size of 108-200. The architecture also implements colorless recognition, which has shown to be more accurate compared to color recognition methods.

#	Team	Method	Accuracy
	sermanet	**EBLearn 2LConvNet ms 108-108 + 100-feats CF classifier + No color**	**99.17%**
197	IDSIA	cnn_hog3	98.98%
196	IDSIA	cnn_cnn_hog3	98.98%
178	**sermanet**	**EBLearn 2LConvNet ms 108-108**	**98.97%**
	sermanet	**EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color**	**98.85%**
	sermanet	**EBLearn 2LConvNet ms 108-200 + 100-feats CF classifier + No color + Ramdom features + No jitter**	**97.33%**

Table 2-2-1: Evaluated Results for Different Feature Size on ConvNet [5]

2.2.2 Fast traffic sign recognition using color segmentation and deep convolution network Done by Wong Ci Shen

In [6], the process for traffic sign recognition can be separated into three parts: segmentation detection, recognition and filtering. The first stage can be known as color segmentation for the purpose of traffic sign detection. Shape and color of the sign are the two variables that helps to detect the signs. The reason of using shape as the factor to detect traffic sign is to handle the challenges presented in a typical street scene. With the two factors, it able to reduce the initial search space and fasten up the computational time to detect traffic signs. Despite that, Improved Hue Saturation and Luminance(IHSL) color space is used for the purpose of extraction ROI. By implementing IHSL, it is very useful for taking into lighting changes since chromatic and achromatic components are independent to IHSL. Moreover, a experimental analysis shown that color information distribution in daylight condition have a peak values at the right side whereas night scene have a higher peaks on the left side and foggy conditions are having great number of pixels at the middle. Besides that, Normalised Hue-Saturation (NHS) method and post processing steps are used to find the potential pixels to road signs. Erosion and dilation morphological operators are applied on the binary image for noise removal. After removing noise value, pixels will be grouped to contours and bounding box will be computed.

The next stage is recognition where module is based on CNN concepts. Images are randomly rotated, scaled and translated or enlarged it. Referring to the literature, deep learning categorized to two different architectures which are single-scale CNN and multi-scale CNN. Single scale has two convolutional layer, one pooling layer and two sequential fully connected layer with rectification and last layer is softmax linear classifier. Single scale CNN is implemented and there are two ramifications occur in the project which are feed-forward fashion and input for full connect convolution layer.

The last step is spatial filtering. The purpose of spatial filtering is to further recognize the traffic sign. A probability value is assigned to each of the patches that are not recognized as specific traffic sign and it will be taken consideration through their neighbour patches.

2.2.3 Extended Single Shoot Multibox Detector for Traffic Signs Detection and Recognition in Real-time.

Done by Toh Wei Xuan

In paper [7], traffic sign recognition approached using either deep learning methods or sliding windows with picture pyramids such as R-CNN, YOLO and SSD. To classify traffic sign, sliding window-based methods frequently combine HOG and Linear Discriminant Analysis (LDA) with SVM or Adaboost. Given the characteristics of the colour of traffic signs and the plate rim, needed converted colours photos into grayscale images before applying a multi-scale shape filter. SVM used as an additional step to reduce the number of false positives and boost precision. These conventional methods frequently fall short in a variety of difficult circumstances. Besides, the underlying convolutional networks, trained on ImageNet for feature extraction, is integrated into deep learning techniques. Regions of the picture that may include traffic signs are created using interest region proposal algorithms, and then the ROI output is identified. There are several frameworks that make use of the R-CNN family and the classifier to attain high accuracy. Popular detection frameworks YOLO and SDD have lately received a lot of attention as candidates for general detection jobs. Their accuracy is not as good as the family of R-CNN, but their speed is ideal for real-time applications. The goal is to use the benefits of both techniques to produce a pedestrian detector that is effective in both fps and accuracy.

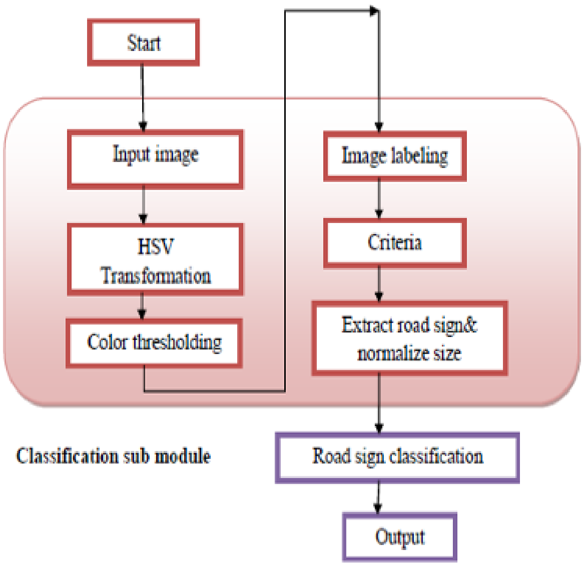


Figure 2-2-1: Overview block diagram based on SVM [7]

2.2.4 Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles Done by Yin Zi Keen

[8] proposed a traffic sign recognition that performed in three phases which is image segmentation, traffic sign detection and classification. Each phase have been improved using different algorithm or method in order to achieve the goal which is improve the traffic sign recognition.

In the first phase which is image segmentation, color play an important role in image segmentation because it can locate the traffic sign quickly. HSV color space is used for spatial threshold segmentation. This is because HSV color space has the faster detection speed and less affected by illumination among HSV, HSI and RGB color space. There are three components inside HSV color space which is H is the hue that represent the color change of the image in which different color has different angles, S represent the proportion of current color purity to the maximum purity in the range of 0-1, and v represent the brightness change in the image in the range of 0(black)-1(white). [8]

Color	*H*	*S*	*V*
Red	$H > 0.90 \| H < 0.10$	$S > 0.40$	$V > 0.35$
Yellow	$0.50 < H < 0.70$	$S > 0.40$	$V > 0.40$
Blue	$0.09 < H < 0.18$	$S > 0.35$	$V > 0.40$

Figure 2-2-2: Threshold range for three main colors [8]

This paper has also showed the corresponding threshold range for three main colors which is red, yellow, and blue colors that mainly exist in traffic sign. This paper has stated that RGB image need to convert into HSV image. After that, H and S value can be obtained, and these values would be used to calculate the means and standard deviation of H and S. After that, it will undergo threshold segmentation process [8]. In the process of threshold segmentation, the image is binarized and the pixels within the set threshold range are set to white, otherwise they are set to black.

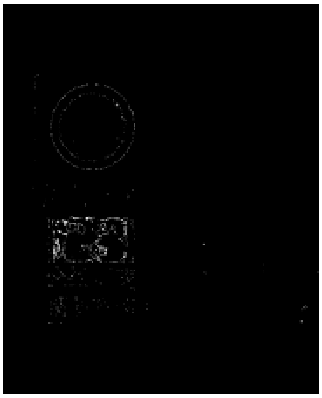


Figure 2-2-3: Threshold rough segmentation image [8]

In the second phase of traffic sign recognition, traffic sign detection. It will start the process with filtering the useless data or interference in image. First and foremost, the binary image undergoes image corrosion and expansion. Image corrosion will eliminate the useless pixels that exist on the edge of image while expansion is used to enlarge the area of ROI. In a result, a prominent shape of traffic sign will be produced. Secondly, filling process which help in complete the place in image of traffic sign which has be damaged or blocked by some obstacle in actual road has been conducted. Lastly, contour filtering is conducted by contour analysis of connected area and aim to filter the large irregular interference areas. In contour filtering, the image is set with all same pixel points and the circumference and area of contour of connected area are calculated and then later is used to compared with the standard circular mark. The contours that meet the requirements will remained, otherwise will be discarded[8]. Thus, the remaining part are the detected traffic sign and it is the final deliverable of this phase.

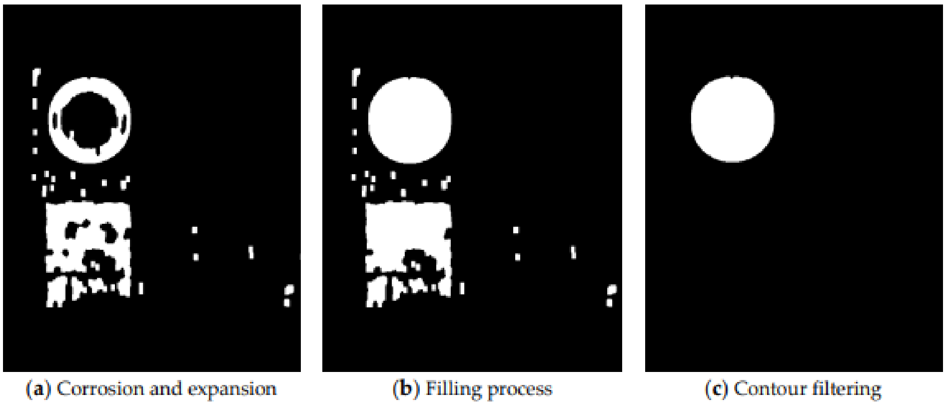


Figure 2-2-4: Processes in traffic sign detection [8]

In the last phase of traffic sign recognition, classification and recognition. In here, a LeNet-5 CNN model which improved by Gabor kernel, adding the BN processing and Adam method as optimizer algorithm. The traffic sign classification and recognition are based on

German Traffic Sign Recognition Benchmark (GTSRB) dataset. Inside GTSRB, there are 51839 images divided as 75% and 25% into training sets and testing sets respectively. Each image contains one traffic sign which is not necessary located at center of image. Image size also unequal which is from the range of 15 x 15 pixels to 250 x 250 pixels.[8] Initially, the classic LeNet-5 CNN model has a good classification and recognition effects on single target but not traffic sign recognition training. Thus, an improved version with embedded Gabor kernel is being proposed. This is because Gabor kernel can solve the problem such as the defect or damage on traffic sign and the exist of obstacles easily and it is insensitive to changes in light so that it means it has good adaptability to light. After that, BN is also using for data normalization because it can solve the gradient dispersion that occurs. In last, the Adam method is used as optimizer algorithm which require less memory space. This is suitable for solving the optimization problems with large-scale data and parameters. In the end, it has shown that the accuracy recognition rate of traffic sign reaches 99.75% which is very high accuracy and the average processing time per frame is 5.4 ms.[8]

2.3 Critical Remarks of Previous Works

Done by Willy Chai Yi Xian, Wong Ci Shen, Toh Wei Xuan, Yin Zi Keen

In [3], for chromatic segmentation, two different machine learning models were applied to compare their performance, namely Support Vector Machine (SVM) and k-Nearest Neighbors (k-NN). After comparison, it is shown that k-NN is more efficient due to its lower error rate along with computational burden. The clear advantage of doing so is to ensure the that more efficient machine learning model is applied to the segmentation process. As a result, high F1-score rate were obtained from the results, signifying that the model is able to segment and classify the datasets accurately. While [3] was able to perform segmentation on both chromatic and achromatic dataset, the achromatic signs are still harder to identify compared to chromatic signs, meaning there is a higher rate of error when it comes to achromatic sign segmentation and classification.

[4] mentioned that uses thresholding of RGB color space can be used for image segmentation as RGB color space can be used to detect strong colors of the sign. Another thresholding method which is Simple Vector Filter can also be used for chromatic and achromatic colors. While RGB color space appears to work well with sign segmentation, it is also prone to error when noises exist in the form of strong colors in the image.

As for Convolutional Network architecture in [5], the result from the proposed architecture is a high accuracy model, which uses colorless recognition, meaning there is no need for color segmentation stage. While the model itself gives high accuracy, colorless recognition may be less effective and reliable when compared to color recognition when applied in real-life situations. [6] mainly emphasizes on two factors for segmentation stage, shape and color. Using shape and color as variables to detect the sign, it will be able to reduce the initial search space and fasten the computational time to detect traffic signs. IHSL color space is also used for ROI extraction, which takes lighting changes into consideration.

For image segmentation in [8], it uses HSV color space which is then segmented using spatial threshold segmentation. The advantage of using HSV color space is that it has faster detection speed and less sensitive to light change. This means that HSV space will be able to adapt to surrounding illumination change better compared to other color spaces, resulting in a better segmentation result.

In [7], several deep learning methods have been implemented and compared, which are R-CNN, YOLO and SSD. R-CNN is shown to give the highest accuracy out of the three models, with the drawback of slower computational speed, which might not be ideal for real-time applications.

Chapter 3: System Design

For the proposed system, the color space is first converted into HSV color space. *K-means* model is then used to segment the input image into different segments, and it has a shortcoming where the optimal k value must be determined, which is the number of clusters for segmented image. To resolve this, Elbow method is introduced to determine the optimal k value. Contour is then used to extract the region of interest (ROI). Several conditions has been applied to filter out the noise contours so that the sign can be properly segmented from the background. In short, the proposed system uses HSV color space with k-means segmentation for image segmenting and contour for ROI extraction.

3.1 System Block Diagram

Done by Willy Chai Yi Xian, Wong Ci Shen, Toh Wei Xuan, Yin Zi Keen

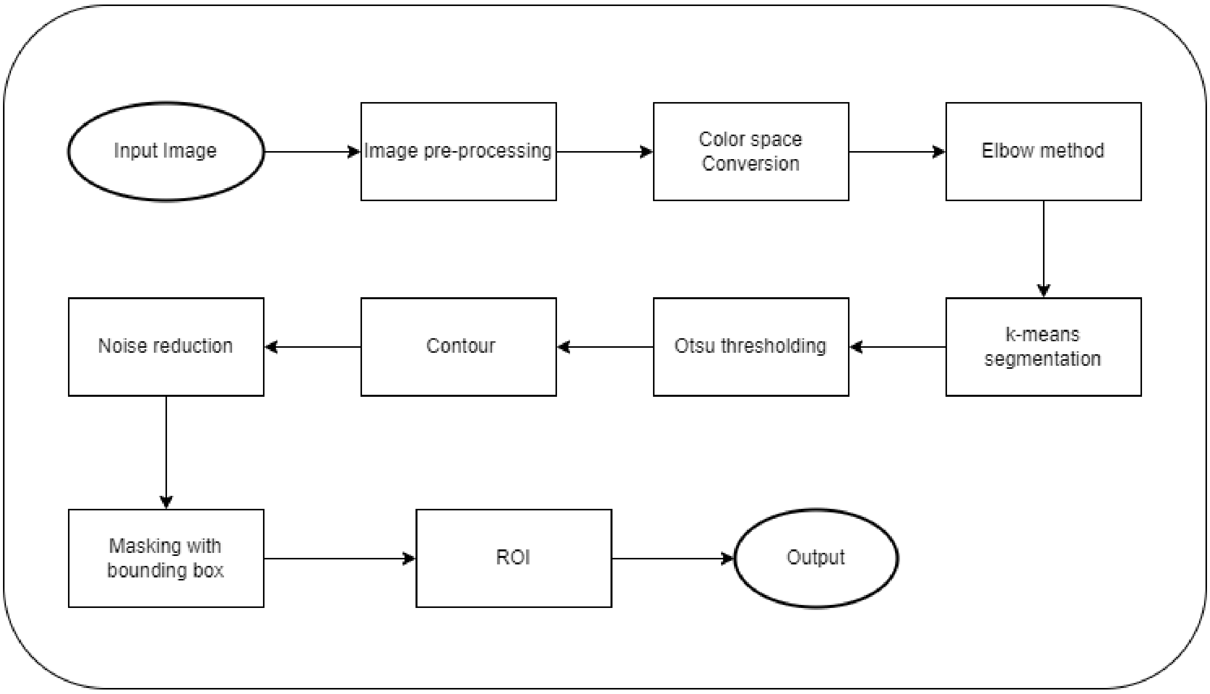


Figure 3-1-1: Block Diagram of Segmentation

3.2 System Components Specifications

Done by Toh Wei Xuan, Willy Chai Yi Xian

3.2.1 Image Pre-processing

Pre-processing is a crucial step in the first stage of proposed work which improves the quality of the input image through edge sharpening and noise filtering. The BGR input image is resized (200 x 200), followed by bilateral filtering to remove noise while the edges remain sharp. Furthermore, the image is then converted into HSV color space and CLAHE is applied on V space to enhance image contrast.

3.2.2 Elbow method

Finding the ideal number of clusters to divide the data into is a critical stage in any unsupervised technique. One of the most prominent techniques for figuring out this ideal value of k is the elbow method. To create clusters in such a way that the total intra-cluster variation, or in other words, the total within-cluster sum of square (WCSS), is minimized, is the fundamental notion behind partitioning methods like k-means clustering. The total WCSS should be as minimal as possible because it represents how compact the clustering is.

The steps of method as follow:

1. Initializing k to 2.
2. Create a for loop which repeats step 3 until it reaches maximum k value, 10.
3. Determine total of WCSS for each k value.
4. Plot the curve of WCSS against the k value.
5. Plot a straight line from starting k to ending k.
6. Find the distance between each point on the elbow graph and the straight line.
7. Plot the results in another graph, the largest distance indicates the optimal k value.

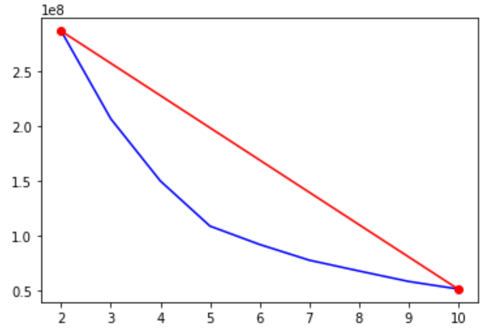


Figure 3-2-1: Elbow graph

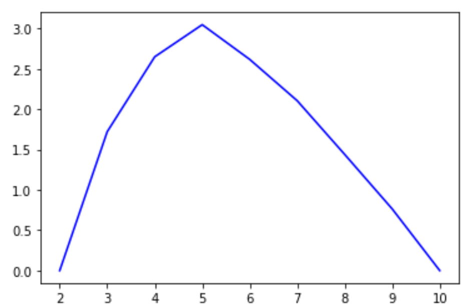


Figure 3-2-2: Optimum k value

3.2.3 *k-means* segmentation

Image segmentation is the division of an image into numerous meaningful, non-overlapping segments that may or may not share the same properties. An important algorithm in digital image processing is image segmentation. As a result, any application that needs this kind of visual information will function better or worse depending on how well the segmentation is done. Traditional image segmentation algorithms base their segmentation decisions on the following three image properties:

- boundary
- region
- threshold

The algorithms used in traditional picture segmentation methods are based on the cluster analysis hypothesis, which is used by people as they learn to recognize objects by continuously changing their subliminal clustering patterns. The proposed method uses the K-means segmentation method because of the excellent performance of traditional K-means cluster analysis in large-scale data. The K-means method fundamental idea is to divide the supplied data into groups according to the clustering number K, with the points in each cluster being the nearest to one another.

k-means is a centroid-based unsupervised classification system. In accordance with minimizing the error function, it classifies the data into a pre-set group. The popular feature space clustering techniques have a partitioning algorithm called the k-means clustering algorithm. Finding groupings based on similarities between the data and the number of groups *k* represents is the objective.

The steps of algorithms as follow:

1. Choose a random group of points from the data set to serve as the initial cluster centers.
2. Determine the distance between each sample and the cluster centers in turn. The sample is then positioned beneath the sample from the closest class.
3. Recalculate the cluster center in accordance with the results of the clustering. The new clustering center will be determined by taking the arithmetic mean of all the elements.
4. Re-cluster each component of the data set in accordance with the new center.
5. Repetition of Step 4 until no change in clustering is observed.
6. Display the outcome.

Stop repeat criteria: -

- There are no more iterations possible.
- The newly generated cluster's centroid remains unchanged.
- The same cluster is true for the data points.

It is necessary to group the data points into clusters, assuming that there are N total data points. J is to be minimized by the *k-means [9]*:

number of clusters　number of cases　centroid for cluster *j*

case *i*

$$\text{objective function} \leftarrow J = \sum_{j=1}^{k}\sum_{i=1}^{n} \underbrace{\left\| x_i^{(j)} - c_j \right\|^2}_{\text{Distance function}}$$

Figure 3-2-3: k-means function [9]

3.2.4 ROI extraction with contour

The steps of contour-based image segmentation are as follows. Choose an image for segmentation first, then apply thresholding and contour onto the thresholded image, filter out noise contours, drawing bounding box according to filtered contour, and perform masking techniques to segment out the sign from background.

The steps of this algorithm are as follows:

A. Thresholding

Convert the k-means segmented image into grayscale image. Then, perform Otsu thresholding onto the grayscale image so that it can be used for contouring later on.

B. Apply contour

Then, using the thresholded image, apply the standard contour model to the traffic sign image. This will result in a set of contours, where one of the contours will draw out the edge of the sign.

C. Filter contour

Filter out the contours accordingly so that only the contour of the edge of the sign remains. Contours that are touching the display border and contours will area smaller than threshold are removed. By filtering the contours, most of the noises can be cleared.

D. Bounding box

After filtering the noise, search for the contour with largest area and draw a rectangle bounding box which is used to compare with the ground truth for result evaluation later on.

E. Masking

The results from contour filtering will result in a final contour which gives a basic outline along the sign perimeter. Foreground and background are separated from one another in the photograph through masking. The final contour is used to draw out the mask. Dilation method is then used to enlarge the size of the mask due to some of the traffic signs did not include the border of segmented sign.

Chapter 4: System Implementation

4.1 Hardware Setup

Done by Yin Zi Keen

The hardware that used in this project:-

i. Laptop

Description	Specifications
Model	Lenovo ideapad 330
Processor	AMD Ryzen 3 2200U with Radeon Vega Mobile Gfx
Operating System	Windows 10
Graphic	AMD Radeon Graphics Processor (0x15DD)
RAM	12GB DDR4 RAM
Storage	1TB SATA HDD 128GB SATA SSD

Table 4-1-1: Specifications of laptop

4.2 Software Setup

Done by Willy Chai Yi Xian

The software that used in this project:-

i. OpenCV

Version: 4.5.5

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing. By using it, image and videos can be processed to identify objects, faces, or even handwriting of a human.

ii. Jupyter Notebook

Version: 6.3.0

Jupyter Notebook is used for running all coding that related to process of traffic sign recognition.

4.3 Implementation Issues and Challenges

Done by Yin Zi Keen, Wong Ci Shen

The primary factor that affects traffic sign recognition is the lighting and weather conditions. Daylight weather will affect excessive light from background and bring the consequences of reflection and shades whereas absence of light occurs in the dark night. This issue causes a serious problem to binary thresholding value where background color will not be entirely same. Therefore, finding a precise global thresholding value for every image is needed. Despite that, different thresholding techniques need to be used while conducting image processing as global thresholding is not suitable to be used for dark night scenario. Thus, the ultimate challenges will be finding suitable thresholding techniques and value.

Apart from that, images that have more than one traffic sign which is not equal in size cannot be segmented. Sometimes, the image is captured at the angle that can capture two traffic signs on a same frame accidently. Usually, one of the traffic signs in the image is big or far away from another traffic sign. The challenge is unable to segment two traffic signs completely. This issue may lead to inaccurate result of segmentation.

Last but not least, the most problematic challenges are the differences of traffic signs over every country. The traffic signs are giving the similar meaning to the driver, but color layout might have slightly different in every country. For example, some of the countries are having "No Entry" signs with white background color whereas most of the countries are having red background color. Therefore, shapes of the traffic sign might need to be consider improving the accuracy for classification process.

4.4 Timeline

Done by Yin Zi Keen

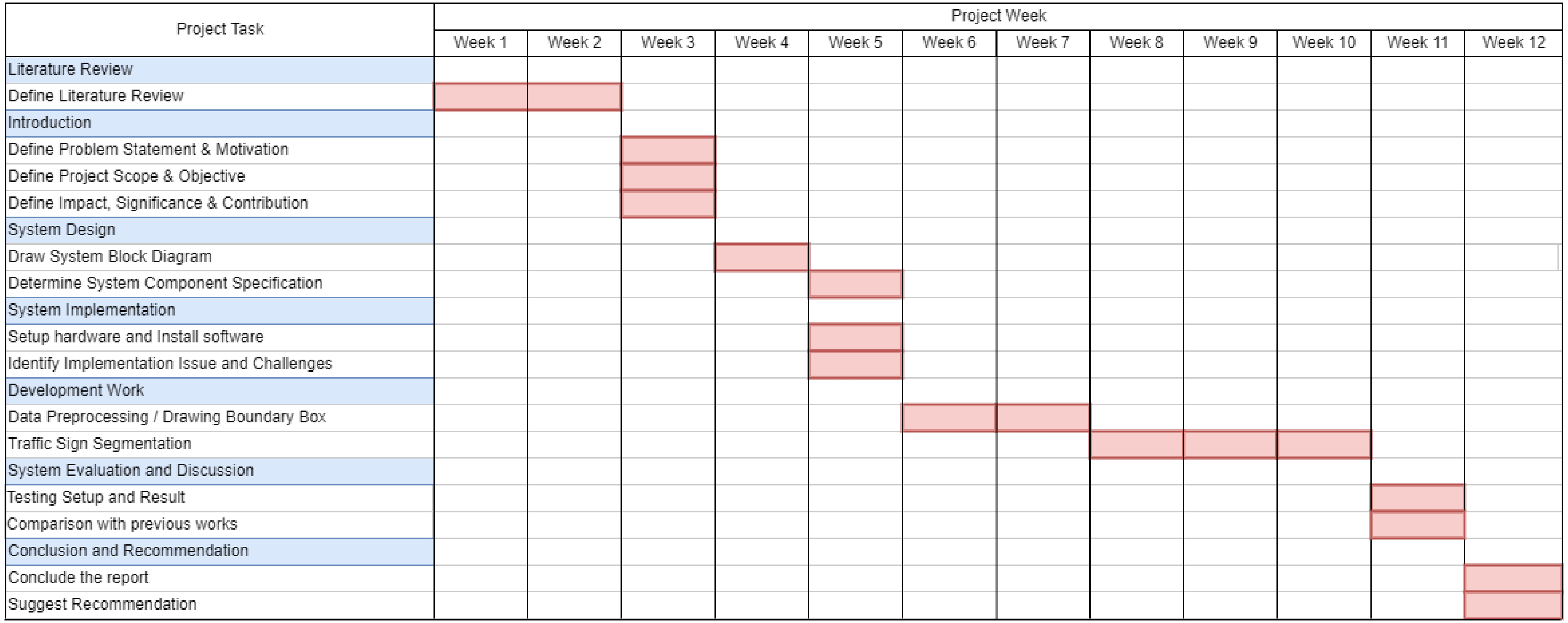


Figure 4-4-1: Project Timeline

Chapter 5: System Evaluation and Discussion

5.1 Testing Setup and Results

Done by Wong Ci Shen

Before proceeding to the image preprocessing stage, reading images is the initial stage that needs to be completed. After reading the image, all images will be resized to a total length of 200 width and height pixels. Cubic interpolation is being considered to implement rather than linear interpolation due to the better accuracy when resizing images. After resizing the images, bilateral filtering is used to perform smooth filtering images. Bilateral filtering is an advanced version of Gaussian blurring where it not only reduces the noises but also keeps the edges sharply.

(a)

(b)

Figure 5-1-1 (a) shows the original image whereas (b) shows the image after implementing bilateral filtering

After performing bilateral filtering, filtered images from RGB format turn to HSV format. The reason that we changing to HSV format is to separate image luminance. By separating the luminance, it makes the process easier when detecting the colour space. Despite that, histogram equalization is implemented for the purpose of improving the contrast of the image. CLAHE is known as Contrast Limited Adaptive Histogram Equalization where it only increases the region which has lower contrast. For instance, the clipLimit parameter is used to increase the level of contrast whereas titleGridSize is used to determine how many regions need to be split for comparing the contrast.

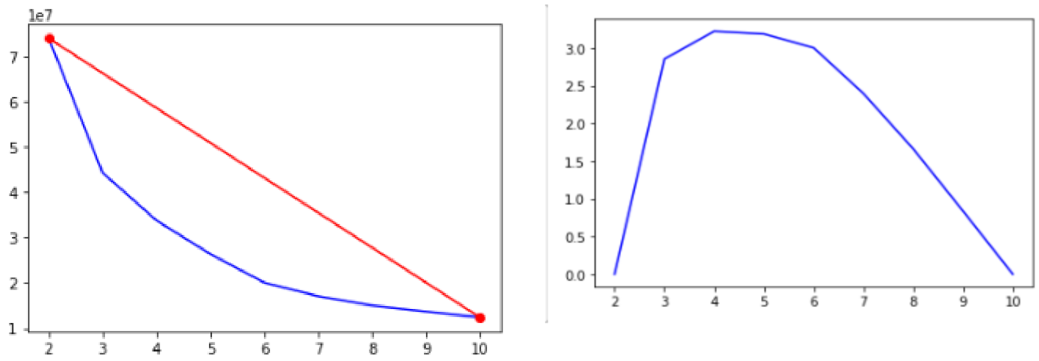
(a) **(b)** **(c)**

Figure 5-1-2 (a) is the original image, (b) the image after implementing bilateral filtering and (c) shows filter image implements CLAHE result

Due to the differentiation of every image, finding an optimal K value is significant in order to segment the traffic sign perfectly. Thus, the method used to determine the optimal K value is utilizing the elbow method.

Procedure for calculating optimal K value

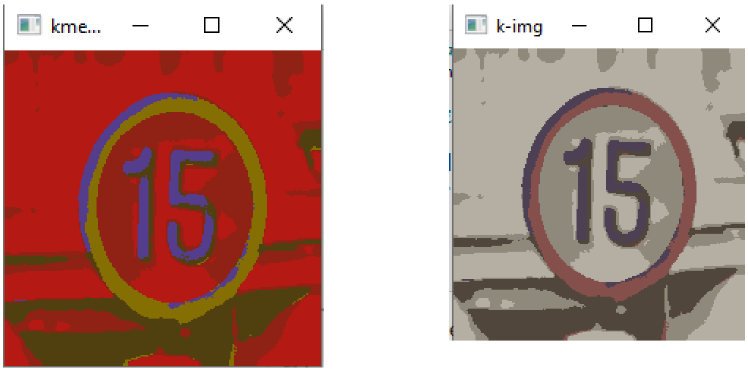
- Plot a straight line from starting point x(2) to ending point x(10)
- Find the longest distance between the points in the elbow graph and the line plotted.
- The longest distance of x value will be determined as the K optimum value.



(a) **(b)**

Figure 5-1-3 (a): elbow function (b): graph to determine the optimal K value

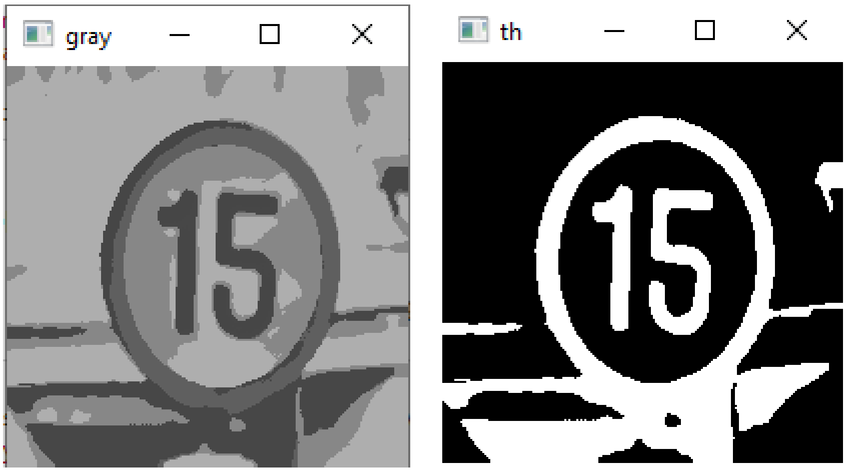
Hence, the finding of the optimal value of K assist the process of k-means segmentation. Instead of using K, K+1 is being replaced as the value of K. This is because the increase of K helps to overcome some images which cannot be segmented properly. Conversion of HSV to RGB format will occur again in order to view the images.



(a) **(b)**

Figure 5-1-4 (a): k-means segmented image in HSV (b): k-means segmented image in BGR

After completing k-means segmentation, the image still shows a lot of noises in the background. Hence, applying contour is necessary in order to extract the ROI region and remove the background noises. Thus, the k-mean segmented image is required to convert to a grayscale image for the purpose of performing thresholding. Otsu Thresholding is used to perform thresholding as it provides better accuracy than global thresholding. The usage of thresholding images is to find all possible contours that can be drawn.



(a) **(b)**

Figure 5-1-5 (a): grayscale k-mean image (b): Otsu thresholded image

Figure 5-1-6 (a) shows the drawn contours that have been detected in the image. In order to extract the contour which outlines the traffic sign, it is important that some rules need to be declared to reduce the drawn contours. By declaring the rules, most of the background noise can be cleared successfully.

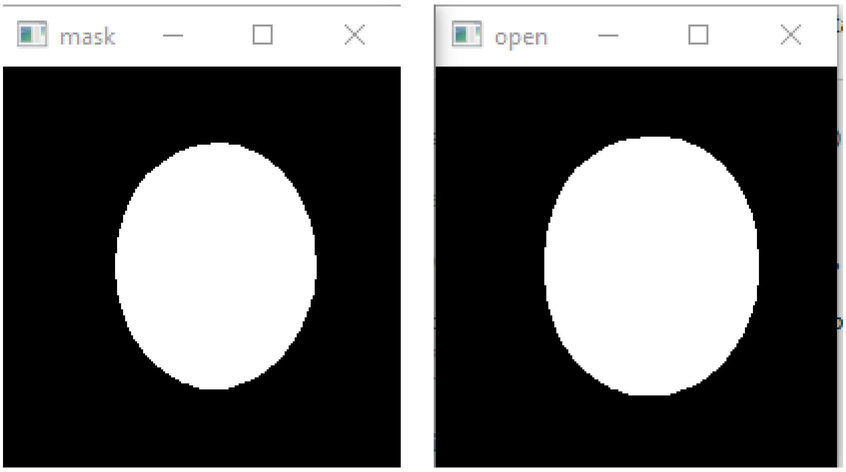
Filtered out the drawn contours

- Remove the contours in which touching to the display border
- Remove contours that have a smaller area than the threshold (Area <=2000)

(a) **(b)**

Figure 5-1-6 (a): contour drawn (b): result contour after filtering

By drawing the contour of the images, masking process need to be used for the purpose of segmentation. The main reason for using masking is to show the region that is interested in. Dilation method is also being implemented for the purpose of enlarging the masking image as some of the segmentation does not include the sign border.



(a) **(b)**

Figure 5-1-7 (a): masking image (b): dilation of masking image

Lastly, bitwise and operation is used to combine the masking of the image and the image that drawn by the contours.

(a) **(b)**

Figure 5-1-8 (a): Input image (b): Segmented image

5.2 Comparison with previous works

After conducting the image segmentation, Intersection over Union (IOU) method is used for the purpose of evaluating the segmented image. A threshold value of 0.8 is fixed in order to indicate the correct segmented image. If the IOU of the particular image is lower than 0.8, the image is considered not to be segmented correctly.

According to the display result, the accuracy of having correctly segmented images is 76%. As a comparison with [5], the correctly segmented image for that Multi-Scale Convolutional Network (Deep Learning Model) is 99.17%. A total difference of 23.17% accuracy between the HSV k-means contour model and deep learning model.

```
Number of total images: 100
Number of segmented images with over IoU of 0.8: 76
Accuracy of the system (segmented images/total images): 0.76
```

Figure 5-2-1: Accuracy of the system

Chapter 6: Conclusion and Recommendation

6.1 Conclusion

Done by Yin Zi Keen

This proposed project is to develop a system that can detect the traffic sign from captured image. This system should be able to segment and recognize the traffic sign no matter the image condition is blur or clear. Therefore, the system can provide correct information of the recognized traffic sign to driver correctly. The result that gained after segmentation should be consist of high level of completeness which means there is no extra data or insufficient of data for recognition occurs. Moreover, the segmentation method should be time-efficient which should take less than 2 seconds for whole segmentation of each captured image.

The segmentation process is executed using k-means algorithm. After reading the image, bilateral filtering is executing to perform smooth filtering images and followed by changing RGB of image into HSV of image. Contrast Limited Adaptive Histogram Equalization (CLAHE) also being used to increase the region which has lower contrast. After that, the optimal k value will be automatic calculated using elbow method and the k value will be used in k-means algorithm later. After k-means segmentation completed, contour will be used to extract ROI region and remove background noises. Otsu thresholding being used later for finding all possible contour that can be drawn and bounding box will be drew. In the end, masking will be executed, and segmentation is being done. The total traffic sign image that used is 100 images. For the result, the accuracy for the proposed model is 0.76, where out of 100 input images, there were 76 images that have been segmented with over 80% of intersection over union (IoU).

6.2 Recommendation

Done by Yin Zi Keen

To improve the overall correctness of the segmented result and to identify the problem that possible occur during segmentation, more test set should be provided. Increasing the number of test set can provide a more accurate result because it means it may identify more bugs that existed due to different test set. The traffic sign image is the test set in this project. Therefore, the traffic sign image that captured better consist of different traffic sign in each image or different environment with same traffic sign in each image. Other than that, avoid using duplicate image during image segmentation.

To improve the quality of the segmented result, comparison among different model or method on segmentation can be applied. For example, machine learning models such as Support Vector Machine (SVM) and k-Nearest Neighbors (k-NN), deep learning model such as R-CNN, thresholding method which include HSI thresholding, RGB color space thresholding and Simple Vector Filter can be applied. Then, same test set will be used for each model or method and algorithm to calculate the accuracy will be remain the same. Due to only way to segment traffic sign from captured image is not different, the result from the comparison among the method should be good to reference. Different model or method of segmentation may produce different accuracy. The model with higher accuracy should be used for the segmentation part in the system. Moreover, comparison make the limitation of each model or method of segmentation may be easier to find out such as slow computational speed, higher error rate, more sensitive to light change, etc.

Moreover, method that involved the combination of histogram equalization and color constancy method may be used to preprocess the HSV color space in future works to filter out noises and sharpen the image for better results. Ebner's color constancy algorithm can be used but only for parallel processing. Figure below shows an example on how the said method is done on RGB colour space. The process is start with the captured image is being separated into 3 different image and everyone in each of the RGB channels separately. After that, each RGB channel will undergo histogram equalization and followed by color constancy algorithm. This is to extract the true colour of the road signs [10].

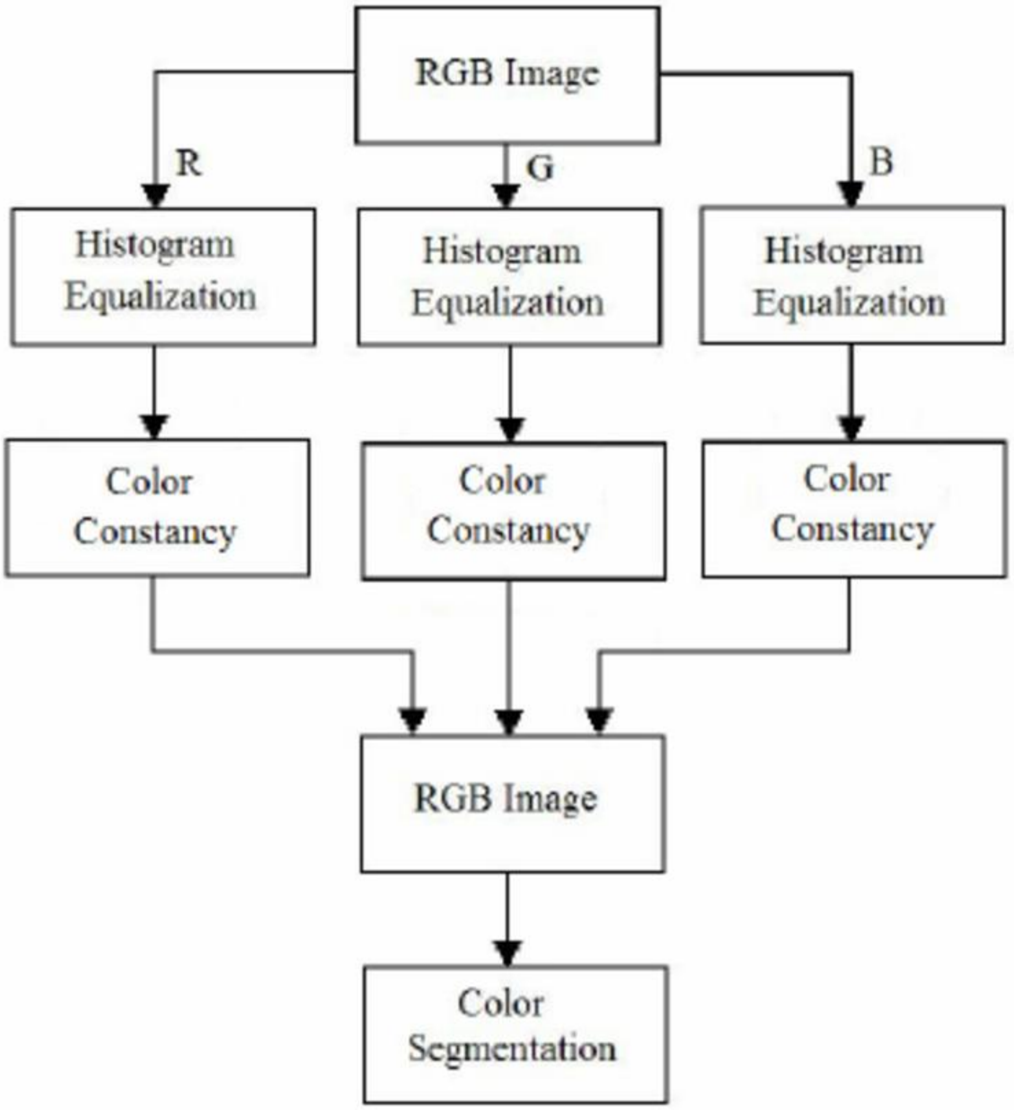


Figure 6-2-1: Block Diagram of Color Segmentation Algorithm

Bibliography

[1] A. Lim, "More cars than people in Malaysia at present? Not likely, as 33 mil figure is total registrations since 1946 - paultan.org", *Paul Tan's Automotive News*, 2022. [Online]. Available: https://paultan.org/2022/06/13/more-cars-than-people-in-malaysia-at-present-not-likely-as-33-mil-figure-is-total-registrations-since-1946/#:~:text=Last%20week%2C%20it%20was%20reported,which%20stands%20at%2032.6%20million. [Accessed: 27- Jul- 2022].

[2] S. Abdulateef and M. Salman, "A Comprehensive Review of Image Segmentation Techniques", *Iraqi Journal for Electrical and Electronic Engineering*, vol. 17, no. 2, pp. 166-175, 2021. Available: 10.37917/ijeee.17.2.18 [Accessed 27 July 2022].

[3] J. Lillo-Castellano, I. Mora-Jiménez, C. Figuera-Pozuelo and J. Rojo-Álvarez, "Traffic sign segmentation and classification using statistical learning methods", *Neurocomputing*, vol. 153, pp. 286-299, 2015. Available: 10.1016/j.neucom.2014.11.026 [Accessed 27 July 2022].

[4] A. Hechri and A. Mtibaa, "Two-stage traffic sign detection and recognition based on SVM and convolutional neural networks", *IET Image Processing*, vol. 14, no. 5, pp. 939-946, 2020. Available: 10.1049/iet-ipr.2019.0634 [Accessed 27 July 2022].

[5] B. Mamatkulovich, "LIGHTWEIGHT RESIDUAL LAYERS BASED CONVOLUTIONAL NEURAL NETWORKS FOR TRAFFIC SIGN RECOGNITION", *European International Journal of Multidisciplinary Research and Management Studies*, vol. 02, no. 05, pp. 88-94, 2022. Available: 10.55640/eijmrms-02-05-19 [Accessed 27 July 2022].

[6] I. Journal, "Traffic Sign Recognition and Classification Using Deep Convolution Neural Network", *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 06, no. 06, 2022. Available: 10.55041/ijsrem14156 [Accessed 27 July 2022].

[7] A. Arora, A. Grover, R. Chugh and S. Reka, "Real Time Multi Object Detection for Blind Using Single Shot Multibox Detector", *Wireless Personal Communications*, vol. 107, no. 1, pp. 651-661, 2019. Available: 10.1007/s11277-019-06294-1 [Accessed 27 July 2022].

[8] J. Cao, C. Song, S. Peng, F. Xiao and S. Song, "Improved Traffic Sign Detection and Recognition Algorithm for Intelligent Vehicles", *Sensors*, vol. 19, no. 18, p. 4021, 2019. Available: 10.3390/s19184021 [Accessed 27 July 2022].

[9] S. C. Nagesh , "Introduction to image segmentation with K-means clustering," *KDnuggets*, 09-Aug-2019. [Online]. Available: https://www.kdnuggets.com/2019/08/introduction-image-segmentation-k-means-clustering.html. [Accessed: 11-Sep-2022].

[10] F. Hasan, "Traffic Signs Color Detection and Segmentation in Poor Light Conditions," *ResearchGate*, Jan-2005. [Online]. Available: https://www.researchgate.net/publication/29751219_Traffic_Signs_Color_Detection_and_Segmentation_in_Poor_Light_Conditions. [Accessed: 11-Sep-2022].