VEHICLE LOGO CLASSIFICATION USING BAG OF WORD DESCRIPTOR AND SUPPORT VECTOR MACHINE CLASSIFIER

NG JIA PHUI

A project report submitted in partial fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic System)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

JANUARY 2019

DEDICATION

This project report is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

Throughout the process in making this report, I received countless help and support from many people, lecturers and friends. Without their contribution, this report will never be completed. The very first person I wish to express my sincere gratitude is my supervisor, Prof. Dr. Syed Abdul Rahman Bin Syed Abu Bakar. He gave me a lot of guidance, support and critics to ensure my project can be completed in timely manner. Without him, completing this report would have been difficult.

I am also indebted to Universiti Teknologi Malaysia (UTM) for all the teaching and knowledge throughout the 2 years courses of part time Master study in this university. I would also like to express my appreciation to Intel Microelectronic Sdn. Bhd. for sponsoring my part time Master study in UTM. Among my colleagues, I would like to thanks Mr. Teh Eng Keong and Mr. Ng Tze Han, as well as my former manager Mr. Yeo Kah Ping, for constantly cover for my work in the company so that I have sufficient time to perform my research and write this report.

Last but not least, I would like to express my sincere appreciation to my family members, especially my parents who gave me limitless support and encouragement. Thank you.

ABSTRACT

Intelligent Transportation Systems play an important role in traffic areas such as to record vehicular traffic data. In order to improve transportation safety and security, a system with the ability to automatically extract and recognize a vehicle is needed apart from the existing plate number recognition system. The detection and recognition of the vehicle type or model can be helpful in determining whether the vehicle is registered with the department of motor vehicle. Hence, this project aims at providing extra information with respect to the vehicle which is to determine the maker of the vehicles. In this project, the classification system is trained with 10 training images for each vehicle's manufacturer. The common features for each logo will be extracted using the Speeded-Up Robust Features algorithm and then feature points will be grouped and arranged using Bag of Word representations which will then be clustered using K means clustering method. The vehicle's classification will be determined by using Support Vector Machine classifier to classify and identify the logo of the vehicle. From the experimental results, the classification system achieved 87% and 77% for front view and side view images respectively with 1500, number of cluster.

ABSTRAK

Sistem Pengangkutan Pintar memainkan peranan penting dalam bidang trafik seperti mencatat data trafik kenderaan. Untuk meningkatkan keselamatan dan keselamatan pengangkutan, sistem yang mempunyai keupayaan untuk mengekstrak dan mengenali kenderaan secara automatik diperlukan selain daripada sistem pengenalan nombor plat sedia ada. Pengesanan dan pengiktirafan jenis kenderaan atau model boleh membantu menentukan sama ada kenderaan itu didaftarkan dengan jabatan kenderaan bermotor. Oleh itu, projek ini bertujuan menyediakan maklumat tambahan berkenaan dengan kenderaan yang menentukan pembuat kenderaan. Dalam projek ini, sistem klasifikasi dilatih dengan 10 gambar latihan untuk setiap pengeluar kenderaan. Ciri-ciri umum bagi setiap logo akan diekstrak dengan menggunakan algoritma Ciri Berkuasa yang Dipercepatkan dan kemudian titik ciri akan dikumpulkan dan disusun menggunakan perwakilan Bag of Word yang kemudiannya akan dikelompokkan menggunakan kaedah K berarti pengelompokan. Klasifikasi kenderaan akan ditentukan dengan menggunakan pengeluar Mesin Sokongan Vektor untuk mengklasifikasi dan mengenal pasti logo kenderaan. Dari hasil eksperimental, sistem klasifikasi mencapai 87% dan 77% untuk pandangan depan dan pandangan sisi masing-masing dengan 1500, bilangan kluster.

TABLE OF CONTENTS

TITLE		PAGE
DECLARATION		**ii**
DEDICATION		**iii**
ACKNOWLEDGEMENT		**iv**
ABSTRACT		**v**
ABSTRAK		**vi**
TABLE OF CONTENTS		**vii**
LIST OF TABLES		**x**
LIST OF FIGURES		**xi**
LIST OF ABBREVIATIONS		**xiii**
LIST OF SYMBOLS		**xiv**
CHAPTER 1	**INTRODUCTION**	**1**
1.1	Project Background	1
1.2	Problem Statement	2
1.3	Research Objectives	3
1.4	Scope of Work	4
1.5	Project Report Outline	4
CHAPTER 2	**LITERATURE REVIEW**	**6**
2.1	Introduction	6
2.2	Existing Features Extraction Techniques	6
	2.2.1 Scale-Invariant Feature Transform	8

2.2.2 Principal Component Analysis (PCA) Algorithm 9

2.2.3 Speeded Up Robust Features 11

2.2.4 Zernike Moments Phase-Based Logo Feature Vector 14

2.3 Logo Recognition 16

2.3.1 Template Matching 16

2.3.2 K Nearest Neighbor 18

2.3.3 Support Vector Machine 19

2.4 Related Work 21

CHAPTER 3 RESEARCH METHODOLOGY 25

3.1 Introduction 25

3.2 Design Model 25

3.3 Input Acquisition and Pre-Processing 26

3.4 Training Process 30

3.4.1 Feature Extraction 30

3.4.2 Bag of Word Descriptor 31

3.4.3 Support Vector Machine 35

3.5 Testing Process 38

CHAPTER 4 EXPERIMENT RESULT 39

4.1 Introduction 39

4.2 Histogram of Training Data 39

4.3 Vehicle's Manufacturer Classification Analysis 42

CHAPTER 5 CONCLUSION AND FUTURE WORK 46

5.1 Conclusion 46

5.2 Future Works 46

REFERENCES **48**

Appendices A - B 53 - 57

LIST OF TABLES

TABLE NO.	TITLE	PAGE
Table 2.1	Four common kernel function	21
Table 2.2	Overall summary of the literature review	21
Table 4.1	The testing result of the number of clusters	43
Table 4.2	The accuracy of classification for each vehicle's model in frontal view with number of cluster, 1500	44
Table 4.3	The accuracy of classification for each vehicle's model in side view with number of cluster, 1500	45

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 2.1	Illustrative image local features (a) input image, (b) corners, (c) edges and (d) regions	7
Figure 2.2	(a) Approximation of Gaussian derivatives using box filters; (b) scale space of different sub_pixels in parallel	13
Figure 2.3	Classify object for K= 4	19
Figure 2.4	Illustration of SVM principle	20
Figure 3.1	The design of vehicle's logo classification system	26
Figure 3.2	Cropping of logo from vehicle image	27
Figure 3.3	Training dataset arrangement by labelling	28
Figure 3.4	Testing dataset for front view logo arrangement by labelling	29
Figure 3.5	Testing dataset for side view logo arrangement by labelling	29
Figure 3.6	Create grid of the training image	31
Figure 3.7	Block diagram of the process of learning a vocabulary of visual words	32
Figure 3.8	Block diagram of create histogram for each training data	34
Figure 3.9	Illustration of create SVM model for Honda and Jaguar	35
Figure 3.10	Block Diagram on classifying the testing image	38
Figure 4.1	10 categories of training dataset	40

Figure 4.2 Histogram of each logo 41

Figure 4.3 The number of clusters versus the accuracy of classification with frontal view logo and side view logo 43

LIST OF ABBREVIATIONS

BoW	-	Bag of Word
ITS	-	Intelligent Transport System
LPR	-	License Plat Recognition
KM	-	Krawtchouk Moment
KMI	-	Krawtchouk Moment Invariant
KNN	-	K Nearest Neighbor
OST	-	Orthogonal Structure Tree
PCA	-	Principal Component Analysis
RBF	-	Gaussian Radial Basis Function
SIFT	-	Scale Invarient Feature Transform
SURF	-	Speeded Up Robust Feature
SVM	-	Support Vector Machine
VLR	-	Vehicle Logo Recognition
VMR	-	Vehicle Model Recognition
ZMs	-	Zernike Moment

LIST OF SYMBOLS

C	-	Overfitting Parameter
k	-	Number of Cluster Centre
γ	-	Kernel Gamma

CHAPTER 1

INTRODUCTION

1.1 Project Background

As the current traffic density increases, traffic congestions, accidents and other traffic problems can be seen everywhere. According to the statistics on total road accidents in Malaysia from Ministry of Transport Malaysia [1], a total number of 521 466 accidents were recorded in 2016, it is increased from 489 606 accidents in 2015. Based on the accident statistics between year of 2015 and 2016, the number of road accident has increased from 26, 301 to 27, 613. The existing traffic surveillance video is applied in the traffic road to capture the conditions of the vehicle and traffic only. However, the system couldn't recognize the vehicle automatically [2].

Therefore, Intelligent Transportation Systems (ITS) is very important in traffic areas such as highway traffic, toll station, parking lot management and other fields. The application of ITS included License Plate Recognition (LPR), Vehicle Logo Recognition (VLR), and Vehicle Model Recognition (VMR) to monitor the traffic condition. VLR is important in both the former and latter tasks as supplementary information. Vehicle's account of possible unscrupulous activity on the part of the plate user is identified by

using vehicle logo identification [3] [4]. Hence, the combination of VLR and LPR can improve the accuracy of recognition and information of the vehicle. Besides, the safety performance in traffic road can be improved by differentiating car manufacturers and models. The information of VLR can be useful for businesses and transportation authorities. In addition, the VLR can be efficiently narrowed by considering manufacturer of the vehicle and also the characteristics of the vehicles.

In conclusion, VLR is important in ITS because it can assist in vehicle identification through the vehicle logo. Hence, vehicle logo recognition is very useful for commercial investigations and document retrieval.

1.2 Problem Statement

Today, it is very useful to know the number and type of vehicles on road. The fundamental for ITS applications is used to record and identify vehicle traffic data. Therefore, plate number recognition system is one of the image processing technology which uses number plate to identify the vehicle. Through ITS, officer in-charge is able to check whether the plate number is a valid registered number or not [1].

However, using LPR to identify the vehicle information is not fully reliable because the appearance of worn license plate and fake

license plate of the vehicle. In order to improve transportation safety and security, the system need to be enhanced with the ability to automatically extract and recognition a vehicle's model using the logo and analyze whether the given plate number matches with the vehicle type or model registered with the department of motor vehicle [4].

Hence, vehicle logo can be used as important feature because the logo essentially provides the information on the maker of the vehicle. If the vehicle logo system could be accurately recognized, then vehicle classification in ITS will effectively improve in recognition accuracy and vehicle classification by combining vehicle logo with other vehicle characteristics [5].

1.3 Research Objectives

The aim of the project is to classify the vehicle's logo using Bag of Words (BoW) Descriptor and Support Vector Machine (SVM) Classifier. The project's aim can been accomplished through the following efforts:

(1) To determine the feature point of vehicle logo by using Speeded-Up Robust Features (SURF) algorithm.

(2) To classify the vehicle's logo accurately by using Bag of Words Descriptor and Support Vector Machine Classifier.

(3) To determine the performance of the classifier based on different view of the logo.

1.4 Scope of Work

The scope of the project is mainly focused on the classification part which applies BoW Descriptor and SVM Classifier. The feature points of the logo will be determined by using SURF algorithm which will be used in the vehicle's logo classification. Besides that, dataset will be prepared in either png or jpg format. The dataset will be collected from website (Google Image). The 10 type of dataset is collected from local or international vehicle's logo which is Honda, Jaguar, Kia, Mazda, Mercedes, Nissan, Perodua, Peugeot, Proton, Toyota and Volvo. The collected vehicle's logo will be at front view for training dataset. MATLAB will be the main tool in this project to design vehicle's logo classification system.

1.5 Project Report Outline

The project report is organized as the following description. There are total of five chapters in this report. Chapter 1 is the introduction of the background and the encouragement to develop logo classification system. Project background, problem statement, objectives, scope of work, and the project report outline are discussed. Chapter 2 is the literature review of this research project. The studies

and research findings on existing features extraction techniques, for example SIFT and SURF algorithms and logo recognition techniques using SVM and KNN algorithm. The comparison of the existing features extraction and logo recognition techniques will be included in this chapter as well and related work on the existing research are presented in this chapter. Chapter 3 discusses the research methodology of this project, as well as the methods are used for achieving objectives. The tools used in this project also are discussed. Chapter 4 is the result and discussion of this project. The accuracy of classification system design is reported and discussed more detailed. Chapter 5 discussed on the conclusion for this project and recommendation on future work to improve current design

REFERENCES

[1] General Road Accident Data in Malaysia (1997 – 2016). Available online: https://www.miros.gov.my/1/page.php?id=17

[2] N. Farajzadeh, "Vehicle Logo Recognition using Image Matching and Textural Features," Journal of Artificial Intelligence in Electrical Engineering, Vol. 2, No. 5, pp. 82–86, 2014.

[3] K. Zhou, K. Zhou, K. M. Varadarajan, M. Vincze, and F. Liu, "Hybridization of Appearance and Symmetry for Vehicle Logo Localization," 15th International Conference on Intelligent Transportation Systems, pp. 1396-1401, 2012

[4] R. Chen, M. Hawes, L. Mihaylova, J. Xiao, and W. Liu, "Vehicle Logo Recognition by Spatial-SIFT Combined with Logistic Regression," 19th International Conference on Information Fusion, pp. 1-8, 2016

[5] P. Billa, "An Implementation of Effective Logo Matching and Detection using Multiple Descriptors to Enhance the Resolution," vol. 161, no. 5, pp. 24–27, 2017.

[6] M. Hassaballah, Aly Amin Abdelmgeid and Hammam A. Alshazly, "Image Features Detection, Description and Matching," Springer, Image Feature Detectors and Descriptors, pp. 11-45, 2016

[7] T. Tuytelaars, K. Mikolajczyk, "Local invariant feature detectors: a survey," Foundations and trends in computer graphics and vision, vol. 3, no. 3, pp. 177–280, 2008

[8] E. Salahat and M. Qasaimeh, “Recent Advances in Features Extraction and Description Algorithms : A Comprehensive Survey.” 8th International Conference on Industrial Technology, pp. 1-5, 2017

[9] M. A. Mohamad, “A Review on Feature Extraction and Feature Selection for Handwritten Character Recognition,” vol. 6, no. 2, pp. 204–212, 2015.

[10] S. Romberg, “From Local Features to Local Regions”, pp. 841-845, 2011.

[11] V. V Bhosle, “Automatic Logo Extraction and Detection for Document Verification using SIFT and SURF,” vol. 6, no. 5, pp. 555–560, 2017.

[12] V. Pali, “An Extensive Survey on Feature Extraction Techniques for Facial Image Processing,” Sixth International Conference on Computational Intelligence and Communication Networks, pp. 142-148, 2014.

[13] S. Singh, “Accurate Face Recognition Using PCA and LDA”, International Conference on Emerging Trends in Computer and Image Processing, pp. 62-68, 2011.

[14] J.-W. Wu, X.-G. Zhang, “A PCA Classifier and Its Application in Vehicle Detection,” IEEE, pp. 600-605, 2001.

[15] K. Filter, “A Comprehensive Motion Estimation Technique for Algorithm and Kalman Filter,” Sensor Book, vol. 16, no.486, pp.1-16, 2016.

[16] X. Zh, “A TV Logo Detection and Recognition Method Based on SURF Feature and Bag-of¬ Words Model,” 2nd IEEE International Conference on Computer and Communications, pp. 370–374, 2016.

[17] C. Science and C. Science, “Improve speed of Logo Detection and Recognition from the images Using SURF,” International Journal of Advancements in Research & Technology, vol. 4, no. 12, pp. 1–6, 2015.

[18] N. V. Kumar, V. V. Kantha, K. N. Govindaraju, and D. S. Guru, “Features Fusion for Classification of Logos,” Procedia - Procedia Comput. Sci., vol. 85, no. Cms, pp. 370–379, 2016.

[19] Z. Chen and S. Sun, “A Zernike Moment Phase-Based Descriptor for Local Image Representation and Matching,” IEEE Transaction On Image Processing, vol. 19, no. 1, pp. 205–219, 2010.

[20] Z. Zhang, X. Wang, W. Anwar, and Z. L. Jiang, “A Comparison of Moments-Based Logo Recognition Methods,” Abstract and Applied Analysis, pp 1-6, 2014.

[21] C.-H. Teh and R. T. Chin, “On image analysis by the methods of moments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 4, pp. 496–513, Apr. 1988.

[22] A. Alaei, M. Delalandre, "A Complete Logo Detection/Recognition System for Document Images", 11th IAPR International Workshop on Document Analysis Systems, pp. 324-328, 2014

[23] I.-A. Aljarrah, A.-S. Ghorab and I.-M. Khater, “Object Recognition System using Template Matching Based on Signature and Principal Component Analysis”, International Journal of Digital Information and Wireless Communications, pp. 156-163, 2012.

[24] J. Laaksonen and E. Oja, “Classification with Learning k-Nearest Neighbors”. IEEE Proceedings of International Conference on Neural Networks, pp. 1480-1483, 2002.

[25] M. Akhil jabbar, B.L. Deekshatulu, P. Chandra. "Classification of Heart Disease”, 3rd International Conference on Computing for Sustainable Global Development, pp. 3107-3111, 2016

[26] H. Wang and K. Yang, “Normalization Methods of SIFT Vector for Object Recognition,” 10th International Symposium on Distributed Computing and Applications to Business, Engineering and Science, pp. 2–5, 2011.

[27] K. Zhou, K. Zhou, K. M. Varadarajan, M. Vincze, and F. Liu, “Hybridization of Appearance and Symmetry for Vehicle Logo Localization,” 15th International Conference on Intelligent Transportation Systems Anchorage, pp.1386-1401, 2012.

[28] Llorca, D. F., Arroyo, R., & Sotelo, M. A. “Vehicle logo recognition in traffic images using HOG features and SVM,” Conference on Intelligent Transportation Systems, Proceedings, pp. 2229–2234, 2013.

[29] Lipikorn, R., Cooharojananone, N., Kijsupapaisan, S., & Inchayanunth, T. “Vehicle logo recognition based on interior structure using SIFT descriptor and neural network,” International Conference on Information Science, Electronics and Electrical Engineering, pp. 1595–1599, 2014.

[30] Liu, Y., Xiang, W., & Xiao, J. “Vehicle logo recognition by weighted multi-class support vector machine ensembles based

on sharpness histogram features," International Image Processing, pp. 527–534, 2015.

[31] Farooq, J. "Object detection and identification using SURF and BoW model," International Conference on Computing, Electronic and Electrical Engineering, pp. 318–323., 2016.

[33] Zhang, J., Zhao, Y., Zhou, F., & Chi, M. "Visual saliency-based vehicle manufacturer recognition using auto encoder pre-training deep neural networks," International Conference on Imaging Systems and Techniques, pp. 1-6, 2017.

[34] Sotheeswaran, S., & Ramanan, A. "A classifier-free codebook-based image classification of vehicle logos," 9th International Conference on Industrial and Information Systems, pp. 1–6, 2015.

[34] D.Virmani, S. Taneja, G. Malhotra. "Normalization based K means Clustering Algorithm." International Journal of Computer Science and Information Technologies, Vol-2, pp. 36-40, 2015.

[35] J. C. Duan,W. Karl Härdle, J. E. Gentle. "Introduction to Support Vector Machines and Their Applications in Bankruptcy Prognosis," Springer. pp. 731-761, 2011