

Universiti Teknologi MARA

**Color-Based Butterfly Species Identification Using
Support Vector Machine**

Norina Arliana Binti Sopian

**Thesis submitted in fulfilment of the requirements for
Bachelor of Computer Science (Hons.)
Faculty of Computer and Mathematical Sciences**

July 2017

STUDENT DECLARATION

I certify that this thesis and the project to which it refers is the product of my own work and that any idea or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.

.....

NORINA ARLIANA BINTI SOPIANA
2014856302

ABSTRACT

Previously, few of expert found that the butterfly identification was complicated, more time consuming and difficult. Butterfly species are difficult to identify because they have different in size, color, and shape. Butterfly Species Identification is an identification of butterfly species that involves a butterfly image. The purpose of this project is to identify butterfly species. This project proposed an appropriate method for the butterfly species identification which is to identify the species of the butterfly images. The scope of this project covers of two species of butterfly which are Vanessa atalanta and Aglais io. The data has been collected from Datasets for Computer Vision Research website. For project methodologies, there are three phases such as data collection, processing and post-processing. For data collection, two species are involved Vanessa atalanta and Aglais io that have 100 data image for each species. In processing phase, there have feature extraction and butterfly species identification. For feature extraction, color histogram technique was used to extract feature from butterfly image. The features include mean, variance, standard deviation and skewness. These color features are calculated from the color component of red, green and blue. This feature classify using Support Vector Machine (SVM) was applied as a technique for identify two group of butterfly species. Testing and evaluation are includes in post-processing phase. The results proved that it significantly works on two butterfly species of Vanessa atalanta and Aglais io to classify that butterfly images. Therefore, this prototype will be significantly benefits to the users.

TABLE OF CONTENT

CONTENT	PAGE
SUPERVISOR APPROVAL	ii
STUDENT DECLARATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENT	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	x
CHAPTER ONE: INTRODUCTION	
1.1 Background of Study	1
1.2 Problem Statements	2
1.3 Objectives	2
1.4 Scope	3
1.5 Significance	4
1.6 Organization of the Thesis	4
1.7 Conclusion	7
CHAPTER TWO: LITERATURE REVIEW	
2.1 Introduction	8
2.2 Butterfly Species	8
2.3 Image Processing	10
2.4 Feature Extraction	10
2.5 Identification	17
2.6 Conclusion	22
CHAPTER THREE: METHODOLOGY	
3.1 Introduction	23
3.2 Framework	23
3.3 Data Collection	24
3.4 Feature Extraction	25
3.4.1 Color-Based Feature Extraction	26

3.5	Butterfly Species Identification	28
3.5.1	Support Vector Machine (SVM) Classifier	29
3.6	Testing and Evaluation	30
3.7	Hardware and Software Requirements	31
3.8	Conclusion	32
CHAPTER FOUR: PROTOTYPE DESIGN AND IMPLEMENTATION		
4.1	Introduction	33
4.2	Prototype	33
4.2.1	Prototype Initialization	33
4.2.2	Open New Image	34
4.2.3	Generate the Step of Feature Extraction Process	36
4.2.4	Identification Result	37
4.3	Conclusion	38
CHAPTER FIVE: RESULTS AND DISCUSSIONS		
5.1	Introduction	39
5.2	Feature Extraction	39
5.3	Butterfly Species Identification	40
CHAPTER SIX: CONCLUSIONS AND FUTURE WORK		
6.1	Introduction	43
6.2	Summary of the Project	43
6.3	Strengths and Limitations of the Prototype	44
6.4	Recommendations and Future Works	45
6.5	Conclusion	45
REFERENCES		46
APPENDICES		
APPENDIX A: RESULTS OF BUTTERFLY SPECIES IDENTIFICATION FOR VANESSA ATALANTA		51
APPENDIX B: RESULTS OF BUTTERFLY SPECIES IDENTIFICATION FOR AGLAIS IO		73