

Universiti Teknologi MARA

Text To Image Generation Using Stable Diffusion

Muhammad Aizaq Bin Azman

**This is submitted in fulfilment of the requirements for Bachelor of
Computer Science (Hons.) College of Computing, Informatics and
Mathematics**

January 2024

ACKNOWLEDGEMENT

I give praise and appreciation to Allah for His sovereignty and the various talents He has bestowed upon me, since I was able to complete this study in the allocated time. I would want to thank my adviser first and foremost for this work. would not have been possible without the assistance and guidance of Sir Roger Canda, who not only gave his time and effort to help me complete this study, but also aided me along the way. In addition, I would like to thank my lecturer for CSP 600/CSP650, Madam Ummu Fatihah binti Mohd Bahrin, for sharing her expertise with me, supporting me, and inspiring me during the whole semester.

My family members have been the most significant people in my pursuit of this undertaking. A particular thank you should go out to my cherished parents for their unwavering support—both psychologically and physically—as well as for their encouragement and, when I most needed it, their financial aid. I am appreciative to everyone I have had the pleasure of working with on this endeavour. In order to successfully complete the final year project, I would like to thank my favourite classmates for their helpful assistance and emotional support.

ABSTRACT

The results of a study on the text to image generation using stable diffusion are presented in this publication. The goal of the project was to create a system that could create a real human face based on the user description. The proposed system was developed in 3 phases consists of preliminary phase, design and implementation phase, and evaluation phase. The study utilized a dataset that has in LAION-5. The pre-trained model, v1-5-pruned-emaonly in hugging face is used as base model because this project applies transfer learning. The app is designed to be webpage by using the visual studio code. The model is evaluated by using 3 ratio train-test split performance and the highest performance and accuracy is chosen to be the final model.

TABLE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	iii
STUDENT DECLARATION	iv
ACKNOWLEDGEMENT	v
ABSTRACT	vi
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER ONE	1
INTRODUCTION	1
1.1 Background of the study	1
1.2 Problem Statement	2
1.3 Objective	3
1.4 Scope	5
1.5 Project significance	6
1.6 Research Framework	7
1.7 Conclusion	9
CHAPTER TWO	10
LITERATURE REVIEW	10
2.1 Artificial Intelligence	10
2.2 Image Processing	11
2.2.1 Image Generation	11
2.2.2 Computer Vision	13
2.2.3 Application of Image Processing in Law Enforcement	13
2.3 Algorithm	14
2.3.1 What is Stable Diffusion	15
2.3.2 Advantages and Disadvantages of Stable Diffusion	17
2.3.5 What is Generative Adversarial Networks	19

2.3.6 Advantages and Disadvantages Generative Adversarial Networks	20
2.3.7 How Stable Diffusion Work	21
2.4 Implementation of Stable Diffusion Algorithm in Various Problem	28
2.5 Similar Works	37
2.6 Implication of Literature Review	43
2.7 Conclusion	44
CHAPTER THREE	45
METHODOLOGY	45
3.1 Overview of Research Framework Methodology	45
3.1.1 Detailed Research Methodology	45
3.2 Preliminary Study	50
3.2.1 Problem Statement Formulation	50
3.2.2 Literature Review	50
3.2.3 Knowledge of Acquisition	51
3.2.4 Objective and Goals Formulation	51
3.2.5 Research Exploration	52
3.3 Data Collection and Preparation	52
3.3.1 User Requirement Analysis	54
3.3.2 Data Collection	54
3.3.3 Data Description	54
3.3.4 Data Preprocessing	55
3.4 Design and Implementation Phase	57
3.4.1 Software Development Life Cycle (SDLC)	57
3.4.2 Prototype Architecture	58
3.4.3 Flowchart	59
3.4.4 Interface Design	61
3.4.5 Pseudocode of Selected Algorithm	62
3.4.6 Prototype Implementation	62
3.5 Performance Evaluation	63
3.5.1 Dataset Training	64
3.5.2 Find Accuracy	66
3.5 Gantt Chart	67
3.6 Conclusion	68
CHAPTER 4	69
RESULT AND FINDINGS	69
4.1 Conceptual Framework	69