Universiti Teknologi MARA

Image Analysing to Differentiate Human Users or Bots Using Convolutional Neural Network (CNN)

MUHAMAD ANIF IKMAL BIN RUSDI

Thesis submitted in fulfilment of the requirements for Bachelor of Computer Science (Hons.) College Of Computing, Informatics & Media

JULY 2023

ACKNOWLEDGEMENT

Alhamdulillah, praises and thank you to Allah because of His Almighty and His utmost blessings, I was able to complete this research within the designated timeline. I would like to extend my sincerest thanks to my supervisor, Sir Muhammad Atif Bin Ramlan for her guidance and support throughout the project journey. Without her, I would not have been able to complete this project with such success.

Without the participation and cooperation of so many people, some of whose names may not be numerated, Final Report CSP600 could not have been completed. I truly appreciate and gladly accept their contributions. But I also want to express my sincere gratitude to my dear lecturer, Madam Ummu Fatihah binti Mohd Bahrim, lecturer of Project Formulation (CSP600), for allowing me the chance to do research and for her crucial advice throughout my final report. I have been greatly inspired by her dynamism, vision, genuineness, and drive.

My parents and my friends' love, prayers, care, and sacrifices in helping to educate and prepare me for the future have made me incredibly grateful. I also like to thank my siblings for their help and helpful prayers. Finally, I would want to express my gratitude to everyone who helped me, directly or indirectly, to finish the final report before I was able to finish the project.

ABSTRACT

In today's digital landscape, telling the difference between human users and bots has become tricky. To tackle this issue, research focuses on creating a system that uses image analysis to identify and classify entities as either human users or bots. The approach involves collecting a dataset of images, processing the data, and training a model—like a Convolutional Neural Network (CNN)—to accurately distinguish between the two. The study demonstrates the effectiveness of using image analysis, particularly CNNs, in achieving high accuracy rates across various scenarios. The main tasks include gathering data, implementing image analysis techniques, training the model, and evaluating performance. The results emphasize the potential of image analysis-based systems for reliable differentiation, contributing to improved online security measures and prevention of malicious activities. This research aims to provide a straightforward solution to the challenge of distinguishing between human users and bots, with the ultimate goal of enhancing online security, particularly in the context of cybersecurity in Malaysia.

TABLE OF CONTENTS

CONTENT **PAGE**

TABLE OF CONTENTS i
LIST OF FIGURES iii
LIST OF TABLES v
CHAPTER 1 6
INTRODUCTION 6
1.1 Background of Study 6
1.2 Problem Statement 7
1.3 Research Objectives 9
1.4 Project Scope 9
1.5 Project Significance 10
1.6 Overview of Research Framework 11
1.7 Conclusion 11
CHAPTER 2 11
LITERATURE REVIEW 12
2.1 Deep Learning 12
2.1.1 Applications for Deep Learning in Image Recognition 13
2.2 Analyzing to differentiate human users or bots using CNN Based on Image 14
2.3 Convolutional Neural Network Algorithm 15
2.3.1 CNN and How Does It Works? 16
2.3.2 Types of Convolutional Network 17
2.3.3 Advantages and Disadvantages of Convolutional Neural Network 19
2.4 Implementation of Convolutional Neural Network Algorithm in Various Problem 21
2.5 Similar Works 31
2.6 The Implication of Literature Review 42

2.7 Conclusion 43
CHAPTER 3 44
METHODOLODY 44
3.1 Overview of Research Methodology 44
3.1.1 Detailed of Research Framework 44
3.2 Preliminary Phase 51
3.2.1 Literature Study 51
3.2.2 Data Pre-Processing 52
3.2.2.1 Data Collection 52
3.2.2.2 Data Cleaning 53
3.3 Design Phase 54
3.3.1 Prototype Architecture 54
3.3.2 Flowchart 55
3.3.3 Interface Design 56
3.3.4 Pseudocode of Selected Algorithm 57
3.4 Development 57
3.5 Performance Evaluation 58
Accuracy of Algorithm 58
3.5 Documentation 59
3.6 Conclusion 61
CHAPTER 4 62
RESULTS AND FINDINGS 62
4.1 Conceptual Framework 62
4.2 Experiment Settings 63
4.3 Training with Different Epoch 64
4.3.1 Experiment 1: Training 80% testing 20% 64
4.3.1.1 Training with 20 epochs 64
4.3.1.2 Training with 30 epochs 66
4.3.1.3 Training with 40 epochs 67
4.3.2 Summary of Experiment 1 68
4.3.3 Experiment 2: Training 70% Testing 30% 70
4.3.3.1 Training with 20 epochs 70
4.3.3.2 Training with 30 epochs 71