

DEEP LEARNING ANALYSIS IN DEVELOPMENT OF HANDWRITTEN AND
PLAIN TEXT CLASSIFICATION API

By

Danny Gani

11502016

BACHELOR'S DEGREE

in

Information Technology

Faculty of Engineering and Information Technology



SWISS GERMAN UNIVERSITY

The Prominence Tower

Jalan Jalur Sutera Barat No. 15, Alam Sutera

Tangerang, Banten 15143 - Indonesia

August 2019

Revision after the Thesis Defense on July 19th, 2019

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.

Danny Gani

Student

Date

Approved by:

James Purnama, M.Sc.

Thesis Advisor

Date

Dipl.-Inf. Kho I Eng

Thesis Co-Advisor

Date

Dr. Dipl.-Ing. Maulahikmah Galinium, S. Kom, M.Sc

Dean

Date

Danny Gani

ABSTRACT

DEEP LEARNING ANALYSIS IN DEVELOPMENT OF HANDWRITTEN AND PLAIN TEXT CLASSIFICATION API

By

Danny Gani

James Purnama, M.Sc., Advisor
Dipl.-Inf. Kho I Eng, Co-Advisor

SWISS GERMAN UNIVERSITY

Optical Character Recognition (OCR) and Handwritten Text Recognition (HTR) are technologies that enable text recognition. The difference between OCR and HTR is one designed specifically for digital text and one designed for handwritten text. There are already various implementations of OCR and HTR online. However, such systems do not guarantee the systems are in premises. To solve this problem, the OCR and HTR system must be built from the scratch. The purpose of this research is to improve the recognition by separating the text whether it is a handwritten or a printed text, which will later be forwarded into the appropriate recognition system. An application program interface (API) was also created in order to finalize the classification system into real world usage. In this research, the classification system being developed using convolutional neural network (CNN) method. To be able to reach the highest accuracy of the classification system, the experimentation and improvement about hyperparameters, dataset format, data augmentation and analysis on 3 CNN architectures were conducted. In the end of this research, there are 2 architectures in a tight competition, one is VGG-16 with 90.63% accuracy and one is AlexNet with 90.17% accuracy on ideal data testing. However, AlexNet is chosen as the winner after the testing with real data.

Keywords: CNN, text classification, text categorization, in premises, OCR, HTR, API.



© Copyright 2019
by Danny Gani
All rights reserved

SWISS GERMAN UNIVERSITY

DEDICATION

I dedicate this work for my parents, family, friends and lecturer who have supported me with their wisdom, patience and perseverance during this research.



ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor: James Purnama M.Sc., co-advisor: Dipl.-Inf. Kho I Eng and my lecturers: Dr. Dipl.-Ing. Maulahikmah Galinium, S.Kom, M.Sc and Dr. Charles Lim, Msc., CTIA, CHFI, EDRP, ECSA, ECSP, ECIH, CEH, CEI for guiding, supporting, the patience and the knowledge given along this project. Their involvement in this project is more than it can be seen in this paper.

I would also like to thank my family and friends for the support and patience during the time this research conducted and a special thanks to Wilbert to help me in crucial time when this paper might be lost forever in the end of the project and Gino to cross check this paper.

The logo for Swiss German University (SGU) features the letters 'S', 'G', and 'U' in a large, stylized font. The 'S' is light blue, the 'G' is a darker blue, and the 'U' is light blue. A registered trademark symbol (®) is located to the right of the 'U'. Below the logo, the text 'SWISS GERMAN UNIVERSITY' is written in a bold, blue, sans-serif font.

SWISS GERMAN UNIVERSITY

TABLE OF CONTENTS

STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
DEDICATION.....	5
ACKNOWLEDGEMENTS.....	6
TABLE OF CONTENTS.....	7
LIST OF FIGURES.....	12
LIST OF TABLES.....	16
CHAPTER 1 – INTRODUCTION.....	18
1.1. Background.....	18
1.2. Research Problems.....	21
1.3. Research Objectives.....	21
1.4. Significance of Study or Contribution.....	21
1.5. Research Questions.....	21
1.6. Hypothesis.....	21
1.7. Scope and Limitation.....	22
CHAPTER 2 - LITERATURE REVIEW.....	23
2.1. Theoretical Definition.....	23
2.1.1. Machine Learning.....	23
2.1.2. Deep Learning.....	23
2.1.3. Optical Character Recognition.....	24
2.1.4. Handwritten Text Recognition.....	24
2.1.5. Text Classification.....	24
2.1.6. Image Pre-processing.....	25
2.1.7. Supervised Learning.....	25
2.1.8. Binary Classification.....	25
2.1.9. Support Vector Machine.....	25
2.1.10. On Premises Software.....	25
2.1.11. Low Quality Image and OCR DPI.....	25

2.1.12. Keras	26
2.1.13. TensorFlow	26
2.1.14. Theano.....	26
2.1.15. Microsoft CNTK.....	26
2.1.16. Tesseract	26
2.1.17. TensorBoard.....	27
2.1.18. Flask.....	27
2.1.19. ABBY FineReader	27
2.1.20. Esker	27
2.1.21. Training, Validation and Training Set	27
2.1.22. Hyperparameters	28
2.1.23. Loss Function.....	28
2.1.24. Epoch	28
2.1.25. Learning Rate.....	28
2.1.26. Convolution Neural Network.....	28
2.1.27. LeNet-5	28
2.1.28. AlexNet.....	29
2.1.29. VGG-16.....	29
2.1.30. IAM Datasets	29
2.2. Related Works	29
2.2.1. Image Classification with Support Vector Machines	29
2.2.2. Large-scale Learning with SVM and Convolutional Nets for Generic Object Categorization	30
2.2.3. Discrimination Between Printed and Handwritten Text in Documents.....	31
2.2.4. Comparison	31
2.3. Literature Framework.....	32
2.3.1. Research on Open Source Machine Learning Frameworks.....	32
2.3.2. Deep Learning Architecture Comparison	33
CHAPTER 3 - RESEARCH METHODS.....	34
3.1. Research Overview.....	34
3.2. System Overview	37
3.3. Preliminary Research	38

3.3.1. Research on Deep Learning Framework.....	38
3.3.2. Research on Deep Learning Architecture	38
3.4. Exploration of Hyperparameters and Dataset Behaviour	38
3.4.1. Design of Experiments.....	38
3.4.2. Experimental Procedure.....	39
3.4.2.1. Environment Setup.....	39
3.4.2.2. Download Sample Classification Project.....	39
3.4.2.3. Hyperparameters Exploration	39
3.4.2.4. Dataset Behaviour Exploration	40
3.5. Dataset Preparation.....	40
3.6. Design of Deep Learning and API Program	40
3.7. Develop Deep Learning Program.....	41
3.8. Configuration and Analysis of Deep Learning Program.....	41
3.8.1. Evaluation and Validation.....	41
3.8.1.1. Design of Experiments	42
3.8.1.2. Experimental Procedure	42
3.8.2. Compare and Select Deep Learning Architecture	43
3.8.2.1. Design of Experiments	43
3.8.2.2. Experimental Procedure	43
3.9. Develop API Program	43
3.10. Real Environment Testing	44
3.10.1. Design of Experiment	45
3.10.2. Experimental Procedure.....	45
CHAPTER 4 - RESULT AND DISCUSSION.....	46
4.1 Preliminary Research	46
4.1.1. Research on Deep Learning Framework.....	46
4.1.2. Research on Deep Learning Architecture	46
4.2. Exploration of Hyperparameters and Dataset Behaviour	47
4.2.1. Environment Setup.....	47
4.2.2. Download Sample Classification Project.....	47
4.2.3. Hyperparameters Exploration	48
4.2.4. Exploration of Dataset Behaviour.....	51

4.3. Dataset Preparation.....	53
4.4. Design of Deep Learning and API Program	59
4.5. Develop Deep Learning Program.....	60
4.6. Configuration and Analysis of Deep Learning Program.....	61
4.6.1. System Architecture.....	61
4.6.2. Evaluation and Validation.....	62
4.6.3. LeNet-5	62
4.6.3.1. Experiment 1	62
4.6.3.1.1. Evaluation and Validation	62
4.6.3.1.2. Result	63
4.6.3.2. Experiment 2	64
4.6.3.2.1. Evaluation and Validation	64
4.6.3.2.2. Result	65
4.6.3.3. Experiment 3	66
4.6.3.3.1. Evaluation and Validation	66
4.6.3.3.2. Result	69
4.6.3.4. Experiment 4	70
4.6.3.4.1. Evaluation and Validation	70
4.6.3.4.2. Result	70
4.6.4. AlexNet.....	72
4.6.4.1. Evaluation and Validation.....	72
4.6.4.2. Result.....	72
4.6.5. VGG-16.....	74
4.6.5.1. Evaluation and Validation.....	74
4.6.5.2. Result.....	74
4.6.6. Compare and Select Deep Learning Architecture	76
4.7. Develop API Program	78
4.8. Real Environment Testing.....	79
4.8.1. Image Taken from Mobile Device.....	79
4.8.2. API Reliability	81

CHAPTER 5 - CONCLUSIONS AND FUTURE WORKS	82
5.1. Conclusions	82
5.2. Future Works	82
GLOSSARY	83
REFERENCES	84
Curriculum Vitae	88

