PERFORMANCE EVALUATION OF TRANSFER LEARNING ALGORITHMS WITH OPTIMISERS FOR CLASSIFYING MALARIA CELL IMAGES

FOLORUNSO, FAVOUR BOLADE
(21PCG02287)
B.Sc. Computer Science, Redeemers University, Ede, Osun state

AUGUST, 2023

PERFORMANCE EVALUATION OF TRANSFER LEARNING ALGORITHMS WITH OPTIMISERS FOR CLASSIFYING MALARIA CELL IMAGES

BY

FOLORUNSO, FAVOUR BOLADE
(21PCG02287)
B.Sc. Computer Science, Redeemers University, Ede, Osun state

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE (M.SC) DEGREE IN COMPUTER SCIENCE IN THE DEPARTMENT OF COMPUTER AND INFORMATION SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY, OTA, OGUN STATE, NIGERIA

AUGUST, 2023

ACCEPTANCE

This is to attest that this dissertation has been accepted in partial fulfilment of the requirements for the award of the degree of Master of Science in Computer Science in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Miss Adefunke F. Oyinloye
(Secretary, School of Postgraduate Studies) **Signature and Date**

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DECLARATION

I**, FOLORUNSO, FAVOUR BOLADE (21PCG02287),** declare that this research was carried out by me under the supervision of Prof. Olufunke O. Oladipupo of the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria. I attest that the dissertation has not been presented either wholly or partially for the award of any degree elsewhere. All sources of data and scholarly information used in this dissertation are duly acknowledged.

FOLORUNSO, FAVOUR BOLADE

Signature and Date

CERTIFICATION

We certify that this dissertation titled "**PERFORMANCE EVALUATION OF TRANSFER LEARNING ALGORITHMS WITH OPTIMISERS FOR CLASSIFYING MALARIA CELL IMAGES**" is an original research work carried out by **FOLORUNSO, FAVOUR BOLADE (21PCG02287)** in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Ogin State, Nigeria under the supervision of Prof. Olufunke O. Oladipupo. We have examined and found this work acceptable as part of the requirements for the award of Master of Science (M.Sc.) in Computer Science.

Prof. Olufunke O. Oladipupo
(Supervisor) **Signature and Date**

Prof. Olufunke O. Oladipupo
(Head of Department) **Signature and Date**

Prof. Adebukola S. Onashoga
(External Examiner) **Signature and Date**

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies) **Signature and Date**

DEDICATION

I humbly dedicate this work to the Almighty God, whose boundless wisdom, grace, and love have been a guiding light throughout my life's journey.

ACKNOWLEDGEMENTS

First and foremost, I want to thank God for granting me the strength, perseverance, and guidance throughout this challenging journey. Without divine blessings, this endeavour would not have been possible. I am indebted to my supervisor Professor Oladipupo Olufunke, whose invaluable guidance, expertise, and patience were crucial in shaping the direction of this thesis.

Your insightful feedback and constructive criticism have been invaluable in improving the quality of my work. I extend my gratitude to the CApIC-ACE for providing the necessary resources, and facilities that allowed me to focus on my research and complete this thesis. I would like to acknowledge my friends and Colleague Abiodun, Peace Durojaiye, Damilola Adeniji, Opeyemi Odetola, Paul Owolabi, Uche Nnaji, and Faith Adegoke, for their meaningful discussions, collaboration, and camaraderie throughout this academic journey. Your diverse perspectives enriched my understanding and inspired me to strive for excellence. I would also like to appreciate Dr. Adam Yagoub, Prof Ezekiel Adebiyi, and Dr Marion Adebiyi for their help and guidance so far during my whole thesis journey. Lastly, I want to thank my family Pastor and Pastor (Mrs) Folorunso, Flourish, Sharon, and Shalom for their love, encouragement, and unwavering belief in my abilities. Your constant support has been the driving force behind my academic pursuits. To all those mentioned above and many others who have supported me, directly or indirectly, in this endeavour, I extend my heartfelt appreciation. Your contributions have played a vital role in shaping the outcome of this dissertation.

TABLE OF CONTENT

CONTENT **PAGES**

COVER PAGE	**i**
TITLE PAGE	**ii**
ACCEPTANCE	**iii**
DECLARATION	**iv**
CERTIFICATION	**v**
DEDICATION	**vi**
ACKNOWLEDGEMENTS	**vii**
LIST OF FIGURES	**x**
LIST OF TABLES	**xii**
LIST OF ABBREVIATIONS	**xiii**
ABSTRACT	**xiv**
CHAPTER ONE: INTRODUCTION	**1**
1.1 Background to the Study	1
1.2 Statement of the Problem	3
1.3 Aim and Objectives of the Study	3
1.4 Significance of the Study	5
1.5 Organization of the Study	5
CHAPTER TWO: LITERATURE REVIEW	**6**
2.1 Preamble	6
2.2 Conceptual Review	6
2.2.1 Malaria	6
2.2.2 Influence of machine learning on malaria detection	8
2.2.3 Transfer Learning	9
2.2.4 Ensemble Learning	10
2.3 Methodological Review	12
2.3.1 Model Evaluation/ Performance metrics	12
2.3.2 Malaria Detection Approaches	14
2.3.4 Transfer Learning Approaches	18
2.3.5 Ensemble Learning Approaches	20
2.3.6 Optimizers	29
2.4 Related Works	31
CHAPTER THREE: RESEARCH METHODOLOGY	**35**
3.1 Preamble	35
3.2 Data collection and Pre-processing	35

3.2.1 Data Collection 35

3.2.2 Data Pre-Processing 37

3.3 Building of Classifiers 39

3.4 Evaluation of the classifiers 42

3.5 Programming tools for developing the Web Application 43

3.6 Environment for Model building and Implementation 44

CHAPTER FOUR: RESULTS AND DISCUSSION 45

4.1 Preamble 45

4.2 Data Collection and Pre-processing 45

4.3 Building of the Classifiers 48

4.3.1 Transfer learning models with customized classification layer 49

4.3.2 Transfer learning models with and without Data Augmentation 57

4.3.3 Retrained Transfer Learning model summary 65

4.3.4 Retrained/ Fine-tuned Transfer Learning algorithm with different optimizers 72

4.3.5 Ensemble of the retrained models 94

4.3.6 Comparative Summary of the Experiments 95

4.4 Pipeline for Classification of Malaria Red Blood Cell images 98

4.5 Web Application for the Recommended Pipeline 99

CHAPTER FIVE: CONCLUSION AND RECOMMENDATION 102

5.1 Summary 102

5.2 Contribution to Knowledge 102

5.3 Recommendations 103

5.4 Limitations 103

5.5 Conclusion 103

REFERENCES 104

LIST OF FIGURES

FIGURES	TITLE OF FIGURES	PAGES
2.1	Life Cycle of a Mosquito (https://www.cdc.gov/mosquitoes/about/life-cycles/anopheles.html)	7
2.2	Schematic of the Ross-Macdonald models	8
2.3	General framework of Ensemble learning	11
2.4	General framework for Homogeneous and Heterogeneous Ensemble Learning	12
2.5	ResNet architecture	18
2.6	VGG architecture	19
2.7	Illustration of DenseNet model	20
2.8	General framework of sequential Ensemble	22
2.9	General framework of Parallel Ensemble	22
2.10	Different strategies for ensemble learning	29
3.1	Research Process Flow	36
3.2	Transfer Learning models (Top layer)	42
3.3	Detailed Specifications of the Local Machine	44
4.1	Parasitized Malaria microscopic red blood cell Images	45
4.2	Uninfected Malaria microscopic red blood cell Images	46
4.3	Pre-processed Resized Images	47
4.4	Pre-processed Augmented Images	48
4.5	Summary of DenseNet201 and ResNet50V2 models Without Customized Top layer (Full Model)	51
4.6	Summary of (a) DenseNet201, (b) ResNet50V2, (c) VGG19, (d) VGG16, and (e) Xception models With Customized classification layer	56
4.7	Accuracy and Loss curves for (a) DenseNet201, (b) ResNet50V2, (c) VGG19, (d) VGG16, and (e) Xception models without Data Augmentation	59
4.8	Accuracy and Loss curves for (a) DenseNet201, (b) ResNet50V2, (c) VGG19, (d) VGG16, and (e) Xception models with Data Augmentation	61
4.9	Confusion matrix and Performance Evaluation of (a) DenseNet201, (b) ResNet50V2, (c) VGG19, (d) VGG16, and (e) Xception models with and without Data Augmetation	65

4.10	Summary of (a) DenseNet201, (b) ResNet50V2, (c) VGG19, (d) VGG16, and (e) Xception model with trainable layers	71
4.11	Retrained DenseNet201 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	74
4.12	Comparison of the various optimizer loss and accuracy curves	74
4.13	Confusion Matrix and Performance Evaluation of DenseNet201 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	76
4.14	Retrained ResNet50V2 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers.	78
4.15	Comparison of the different optimizer loss and accuracy curves	79
4.16	Confusion Matrix and Performance Evaluation of ResNet50V2 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	81
4.17	Retrained VGG19 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	83
4.18	Comparison of the different optimizer loss and accuracy curves	84
4.19	Confusion Matrix and Performance Evaluation of VGG19 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	85
4.20	Retrained VGG16 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	87
4.21	Confusion Matrix and Performance Evaluation of VGG16 with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	89
4.22	Retrained Xception with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	91
4.23	Comparison of the different optimizer loss and accuracy curves	92
4.24	Confusion Matrix and Performance Evaluation of Xception with (a) Adam, (b) RMSprop, (c) SGD, and (d) Adamax optimizers	93
4.25	Recommended pipeline for Classification of Malaria Red Blood Cell Images	98
4.26	Landing page of the Web application	99
4.27	Selection of multiple Images for classification	100
4.28	Diagnosed blood smears	100
4.29	Record or History of all Diagnosis made	101
4.30	Selecting Multiple images for Labelling	101
4.31	Labelling unlabelled Malaria Red Blood Cell Dataset	101

LIST OF TABLES

TABLES	TITLE OF TABLES	PAGES
1.1	Objectives and the Methodology	4
2.1	Ensemble methods	28
3.1	Augmentation techniques and their values	39
3.2	Hyper-parameters and their values	42
3.3	Confusion Matrix	43
4.1	The folders of the Malaria red blood cells dataset	45
4.2	Dataset Splitting	46
4.3	DensetNet201 with the various optimizers	76
4.4	ResNet50V2 with the various optimizers	81
4.5	VGG19 with the various optimizers	86
4.6	VGG16 with different optimizers	89
4.7	Xception with the various optimizers	94
4.8	Different Ensemble Combination	94
4.9	Performance metrics of the different ensembles	94
4.10	Comparative Performance of models with or without Data Augmentation	96
4.11	Comparison between the best optimizer of each of the retrained models	96
4.12	Comparison of proposed model with existing models	97

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
SGD	Stochastic Gradient Descent
RMSprop	Root mean square propagation
Adagrad	Adaptive Gradient Algorithm
Adam	Adaptive Moment Estimation
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
PCR	Polymerase Chain Reaction
RDT	Rapid Diagnostic Test
CNN	Convolutional Neural Network
RNN	Recurrent Neural Networks
MLP	Multilayer Perceptron Neural Network
Resnet	Residual Network
VGG	Visual Geometry Group
DenseNet	Densely Connected Network
NIH	National Institute of Health
CSS	Cascading Style Sheet
HTML	Hypertext Markup Language

ABSTRACT

Malaria, a devastating disease transmitted by mosquitoes, continues to impose a significant global health and economic burden. Despite this, traditional diagnostic methods such as microscopy encounter limitations due to human error and high workload. Accurate identification of malaria-infected red blood cells is pivotal for effective management, and machine learning—particularly transfer learning algorithms—has shown promise in enhancing diagnosis. However, the optimal pairing of transfer learning architectures with suitable optimizers for precise classification remains unclear. To tackle these challenges, this study presents a performance evaluation of transfer learning algorithms alongside optimizers for classifying malaria red blood cell images. The study utilizes five transfer learning algorithms, including DenseNet201, ResNet50, VGG19, VGG16, and Xception, to investigate their performance in malaria red blood cell image classification. The dataset comprises 27,558 cell images, divided into Parasitized and Uninfected categories with 13,779 images each, sourced from the National Institutes of Health (NIH) dataset. Data pre-processing involves resizing, data splitting, label encoding, and data augmentation techniques to enhance the dataset before model training. Five distinct classifiers are developed using the transfer learning models, employing fine-tuning methods and hyperparameters to achieve exceptional accuracy. The top three classifiers are combined using the max voting ensemble method. The proposed ensemble model demonstrates significant potential in enhancing classification with an accuracy, precision, recall, f1-score of 97.50%, 96.33%, 98.67%, and 97.49% respectively.

Keywords: Malaria red blood cell detection, Deep learning, Transfer Learning, Data Augmentation, Ensemble Learning.