

**AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR
DETECTING INFECTION SEVERITY FROM THICK BLOOD
SMEAR IMAGES**

**IBITOYE, OLAJUMOKE GRACE
(22PBF02394)**

**B.Tech Biochemistry, Ladoke Akintola University of Technology,
Ogbomoso**

AUGUST, 2024

**AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR
DETECTING INFECTION SEVERITY FROM THICK BLOOD
SMEAR IMAGES**

BY

**IBITOYE, OLAJUMOKE GRACE
(22PBF02394)**

**B.Tech Biochemistry, Ladoke Akintola University of Technology,
Ogbomoso**

**A DISSERTATION SUBMITTED TO THE SCHOOL OF
POSTGRADUATE STUDIES IN PARTIAL FUFILMENT OF THE
REQUIREMENT OF THE (M.SC) DEGREE IN BIOINFORMATICS,
DEPARTMENT OF COMPUTER AND INFORMATION SCIENCES,
COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT
UNIVERSITY, OTA, OGUN STATE, NIGERIA**

AUGUST, 2024

ACCEPTANCE

This is to attest that this dissertation is accepted in partial fulfilment of the requirements for the award of the degree of Master of Sciences in Bioinformatics in the Department of Computer and Information Sciences, College of Sciences 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 hereby declare that **IBITOYE, OLAJUMOKE GRACE (22PBF02394)**, carried out this research entitled “An Automated Malaria Diagnosis System for Detecting Infection Severity from Thick Blood Smear Images”. It was carried out under the supervision of Dr. Itunuoluwa Isewon. Concepts of this research project are the results of the research carried out by Ibitoye, Olajumoke Grace and ideas of other researchers have been fully recognised.

IBITOYE, OLAJUMOKE GRACE

Signature and Date

CERTIFICATION

This is to certify that this dissertation titled “**AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR DETECTING INFECTION SEVERITY FROM THICK BLOOD SMEAR IMAGES**” is original research carried out by **IBITOYE, OLAJUMOKE GRACE (22PBF02394)** in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria under the supervision of Dr. Itunuoluwa Isewon. We have examined and found this work acceptable as part of the requirements for the award of Master of Science (M.Sc.) in Bioinformatics.

Dr. Itunuoluwa M. Isewon
(Supervisor)

Signature and Date

Prof. Olufunke O. Oladipupo
(Head of Department)

Signature and Date

Prof. Olusegun Folorunso
(External Examiner)

Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies)

Signature and Date

DEDICATION

I dedicate this project to God Almighty my creator, my strong pillar, my source of inspiration, wisdom, knowledge and understanding

ACKNOWLEDGEMENTS

I am extremely grateful to God for giving me the strength and resources to complete this project and the program successfully. His unwavering support and guidance were crucial during the study, especially during challenging times. I owe a debt of gratitude to my supervisor, Dr. Itunuoluwa Isewon, whose invaluable guidance, expertise, and patience were instrumental in shaping this thesis. I thank all the Department of Computer and Information Science faculty and staff for their input. Their insightful feedback and constructive criticism were invaluable in improving the quality of my work. I thank Mr Taiwo and Dr. Benson of the Department of Biological Science for their benevolence. I extend my appreciation to CApIC-ACE for providing the necessary resources and facilities for me to focus on my research and complete this thesis.

I want to acknowledge my friends and colleagues, Faith, Temi, Emmanuel, Zeendi, Promise, Ogo, Ife, Mercy, Julius, Ashuza, Ademolu, Naomi, Wisdom, Layo, and Favour for their meaningful discussions, collaboration, and encouragement throughout my academic journey. Your diverse perspectives enriched my understanding and inspired me to strive for excellence. I also want to thank Mr. Emmanuel for his help and encouragement during my Master's degree journey and express my gratitude to Mr Gbenga, Adeola, and Deborah for their timely encouragement. Special appreciation goes to Pastor Judah for his advice and prayer.

Lastly, I want to thank Pastor Ibitoye, A/P (Mrs.) Ibitoye, Olaposi, Gbolahan, Dolapo, and Joy 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 CONTENTS

CONTENTS	PAGES
TITLE PAGE	i
COVER PAGE	ii
ACCEPTANCE	iii
DECLARATION	iv
CERTIFICATION	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	xi
LIST OF TABLES	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	4
1.4 Research Methodology	5
1.5 Significance of the Study	6
1.6 Organization of the Dissertation	6
CHAPTER TWO: LITERATURE REVIEW	7
2.1 Preamble	7
2.2 Malaria	7
2.3 Malaria Diagnosis and Parasitaemia	7
2.3.1 Clinical Diagnosis	8
2.3.2 Microscopic Diagnosis	8
2.3.3 Molecular Diagnosis	10
2.3.4 Antigen Detection	10
2.4 Classification	11
2.5 Machine Learning	12
2.5.1 Traditional Machine Learning	12
2.5.2 Deep Learning	14
2.6 Transfer Learning	23
2.7 Malaria Image Processing Techniques	23
2.7.1 Image Acquisition	23
2.7.2 Preprocessing	25

2.7.3 Erythrocyte Segmentation	25
2.7.4 Feature Extraction	26
2.7.5 Classification	27
2.8 Related Works	28
2.9 Summary of Literature Review	33
CHAPTER THREE: METHODOLOGY	35
3.1 Preamble	35
3.2 Dataset	35
3.2.1 Data collection	35
3.2.2 Data Preprocessing	36
3.2.3 Data Split	36
3.3 Transfer Learning and Hyperparameter Tuning	36
3.4 Model Selection and Customized Layers	37
3.5 Model Evaluation	42
3.6 Deploying the Model on a Web-based Server	43
3.7 Environment for Model Building and Implementation	44
CHAPTER FOUR: RESULTS AND DISCUSSION	46
4.1 Introduction	46
4.2 Data Collection and Pre-processing	46
4.3 Transfer Learning Models with Customized Classification Layer	47
4.4 Evaluation of the model with or without augmentation	50
4.5 Evaluation of the model across the classes	57
4.6 Summary of the Experiments	59
4.7 Web Application for deploying the best performing model	60
CHAPTER FIVE: CONCLUSION AND RECOMMENDATION	63
5.1 Summary	63
5.2 Contribution to Knowledge	63
5.3 Recommendation	64
5.4 Limitation	64
5.5 Conclusion	65

LIST OF FIGURES

FIGURES LIST OF FIGURES	PAGES
Figure 1.1: Computer-Aided Diagnosis	1
Figure 2.1: Microscopic view of thin and thick blood smear	9
Figure 2.2:Microscopic view of thin and thick blood smear	11
Figure 2.3:CNN Architecture	15
Figure 2.4: ResNet-50	18
Figure 2.5:VGG16	20
Figure 2.6:Architecture of VGG19	20
Figure 2.7:Architecture of AlexNet	21
Figure 2.8:Architecture of Xception Block Network	22
Figure 2.9: Architecture of CapsuleNet	22
Figure 2.10: Appearance of malaria stages in Giemsa stained thin and thick blood films	24
Figure 3.1: Architecture of VGG19	38
Figure 3.2: Architecture of ResNet50	39
Figure 3.3: Architecture of InceptionV3	39
Figure 3.4: Architecture of VGG16	40
Figure 3.5: Architecture of DesNet201	41
Figure 3.6: UML Diagram of Web-based Application	44
Figure 3.7: Methodology Workflow	45
Figure 4.1: Pascal Format Annotation File for the Malaria Image	47
Figure 4.2: Sample of the severe, mild and uninfected malaria slide images respectively	47
Figure 4.3: VGG16 with customized layer model summary	48
Figure 4.4:VGG19 with customized layer model summary	49
Figure 4.5:ResNet50 with customized layer model summary	49
Figure 4.6:InceptionV3 with customized layer model summary	50
Figure 4.7:DenseNet201 with customized layer model summary	50
Figure 4.8: VGG16 accuracy and loss without augmentation	51
Figure 4.9: VGG16 classification report without augmentation	51
Figure 4.10:VGG16 layer accuracy and loss with augmentation	52
Figure 4.11:VGG16 classification report with augmentation	52

Figure 4.12: VGG19 accuracy and loss without augmentation	53
Figure 4.13: VGG19 classification report without augmentation	53
Figure 4.14: VGG19 accuracy and loss with augmentation	53
Figure 4.15: VGG19 classification report with augmentation	54
Figure 4.16: InceptionV3 accuracy and loss without augmentation	54
Figure 4.17: InceptionV3 classification report without augmentation	54
Figure 4.18: InceptionV3 accuracy and loss with augmentation	55
Figure 4.19: InceptionV3 classification report with augmentation	55
Figure 4.20: DenseNet201 accuracy and loss without augmentation	55
Figure 4.21: DenseNet201 classification report without augmentation	56
Figure 4.22: DenseNet201 accuracy and loss with augmentation	56
Figure 4.23: DenseNet201 classification report with augmentation	56
Figure 4.24: ResNet50 Accuracy and Loss Performance	57
Figure 4.25: VGG16 confusion matrix with augmentation and without augmentation	57
Figure 4.26: VGG19 confusion matrix Without augmentation with augmentation	58
Figure 4.27: InceptionV3 confusion matrix without augmentation and with augmentation	58
Figure 4.28: DenseNet201 confusion matrix without augmentation and with augmentation	59
Figure 4.29: Initial landing page	60
Figure 4.30: Selection of images for classification	61
Figure 4.31: Image upload	61
Figure 4.32: Severe thick malaria image result	62
Figure 4.33: Image upload prompt	62

LIST OF TABLES

TABLES	LIST OF TABLES	PAGES
Table 3.1:	Hyper-parameters and their types	37
Table 4.1:	Dataset Distribution	46
Table 4.2:	Model Performance Metrics Without Augmentation	59
Table 4.3:	Model Performance Metrics With Augmentation	60

ABSTRACT

Malaria is a significant and life-threatening disease that poses a threat to global health and the economy. Accurate diagnosis is crucial in combating this disease. Microscopy is considered the gold standard for diagnosing malaria parasites and requires a skilled slide reader to examine blood slides for the presence of parasites and careful quantification. Therefore, automating the process would be beneficial. The accessibility and availability of large curated datasets for quantifying malaria parasite images in thick blood smears have posed a significant challenge. The absence of an accessible, fully automated, and fully functional web-based system for multiclass classification is also a notable gap. This study aims to develop an automated system for quantifying malaria parasites in routine blood films. Blood samples collected from Covenant University Medical Centre and AceMedicare clinic were used to generate a dataset of 1518 labeled thick blood smear images. The annotations were recorded from the health facilities and used to correctly label the images. The dataset includes 458 uninfected, 427 mild, and 633 severe thick blood smear images. The images were digitized using the Olympus CX33 trinocular microscope. Data processing techniques such as resizing, rescaling, and data augmentation were used to prepare the dataset for training. Five transfer models - VGG16, VGG19, ResNet50, InceptionV3, and DenseNet201 - were trained and evaluated on the newly generated images. The models' accuracies were VGG16- 81.2%, VGG19- 82.5%, InceptionV3- 81.0%, DenseNet201- 81.8%, and ResNet- 68.3%. The best predictive model across all performance metrics was VGG19 with an accuracy of 82.5%, precision of 82.0%, recall of 82.0%, F1-score of 81.0%, and 0.38 loss. The best-performing predictive model was deployed as a web-based application. The study proffers deep learning model capable of classifying malaria thick blood smear images into multi class classification based on their level of severity.

Keywords: Malaria, Thick Blood Smear Image, Deep Learning, Transfer Learning, Multiclass Classification, Malaria Severity