AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR DETECTING LIFE CYCLE STAGES OF PLASMODIUM IN THIN BLOOD SMEAR IMAGES

OBA, EMMANUEL BABATUNDE (22PBF02396) B.Sc Medical Physiology, Bayero University, Kano

AUGUST, 2024

AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR DETECTING LIFE CYCLE STAGES OF PLASMODIUM IN THIN BLOOD SMEAR IMAGES

BY

OBA, EMMANUEL BABATUNDE (22PBF02396) B.Sc Medical Physiology, Bayero University, Kano

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 BOI-INFORMATICS IN THE DEPARTMENT OF COMPUTER AND INFORMATION SCIENCES, COLLEGE OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY, OTA, OGUN STATE

ACCEPTANCE

This is to attest that this dissertation is accepted in partial fulfillment of the requirements for the award of the Master of Sciences in Bio-informatics degree 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, OBA, EMMANUEL BABATUNDE, (22PBF02396), declare that the research entitled "AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR DETECTING LIFE CYCLE STAGES OF PLASMODIUM IN THIN BLOOD SMEAR IMAGES." Was carried out by me under the supervision of Dr. Itunuoluwa M. Isewon of the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria. I attest that the dissertation has not been submitted anywhere else, in whole or in part, for the purpose of receiving a degree. This dissertation includes complete acknowledgments for all data and scholarly information sources.

OBA, EMMANUEL BABATUNDE

Signature and Date

CERTIFICATION

This is to certify that this dissertation titled "AN AUTOMATED MALARIA DIAGNOSIS SYSTEM FOR DETECTING LIFE CYCLE STAGES OF PLASMODIUM IN THIN BLOOD SMEAR IMAGES" is an original research work carried out by OBA, EMMANUEL BABATUNDE, (22PBF02396) in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria under the supervision of Dr. Itunuoluwa M. 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)

Prof. Olufunke O. Oladipupo (Head of Department)

Prof Olusegun Folorunso (External Examiner)

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

Signature and Date

Signature and Date

Signature and Date

DEDICATION

I present this work as a tribute to the Almighty God, whose boundless wisdom, grace, and love have been the beacon illuminating my path through life. Furthermore, I devote this work to my dear parents, who have exhibited steadfast commitment and extraordinary efforts in preparing me for achievement.

ACKNOWLEDGEMENTS

To begin with, I would like to express my gratitude to God Almighty for granting me the strength and mercy necessary to conduct this research study effectively. I would also like to extend a special thanks to my parents for their unwavering love and support throughout this process, as well as a special tribute to my mom. Dear Mom, I'm so grateful for the way I was raised. You were never afraid to be my mother, to put your foot down, to teach me right from wrong, and to make sure I always did my best. You made it clear that I had a mom who cared about me, who believed in me, and who was always there for me, no matter what. Without you, I wouldn't be here today. I wouldn't be the person I am, and I wouldn't have the life I have. You are everything to me and more, and there's no way I could ever quantify the importance you hold in my life. Though you're no longer with us, you are deeply missed. **"Mom, you rock my world."**

I would also like to thank the Head of the Department, Prof. Olufunke Oladipupo, and I would like to acknowledge the invaluable contributions of my supervisor, Dr. Itunuoluwa M. Isewon. Her guidance, patience, and understanding were instrumental in ensuring the success of this study.

I am also grateful for the support from CApIC-ACE to help me pursue my dreams. I appreciate the work of Prof. Emeka E.J.Iweala, Prof. E. Adetiba, and Prof. Ezekiel Adebiyi; I am very grateful to Prof. Olubanke Ogulana for her unwavering kindness, mentorship, and motherly support, and big thanks to Ms. Helen Jevwegaga for always making sure that we are comfortable and we got all we need to succeed in the academic environment.

I am thankful to my colleagues in the program who welcomed and treated me like a family. Furthermore, I would like to thank the Management and Chancellor of Covenant University for providing an enabling environment that allowed me to pursue my academic aspirations in a peaceful and stimulating setting.

Lastly, I want to express my appreciation to my colleagues and friends in the postgraduate program, including Carrin Fiamitia, Ndifon Naomi Sije-Okim, Kuderha Rifin Ashuza, Ibitoye Olajumoke Grace, and others. Their unwavering support and encouragement during challenging times were greatly appreciated, and I would like to thank myself; for all the hard work and overnights I put into this work.

TABLE OF CONTENTS

| CONT | TENTS | PAGES |
|--|---|---|
| TITLE COVE ACCE DECL CERT DEDIO ACKN TABLI LIST (LIST (ABSTI | E PAGE ER PAGE CPTANCE ARATION IFICATION CATION NOWLEDGEMENTS E OF CONTENTS OF FIGURES OF FIGURES OF TABLES RACT | i ii iii iv v v vi vii xi xii xii |
| СНАР | TER ONE: INTRODUCTION | 1 |
| 1.1 | Background to the Study | 1 |
| 1.2 | Statement of Research Problem | 3 |
| 1.3 | Aim and Objectives of the Study | 3 |
| 1.4 | Significance of the study | 5 |
| 1.5 | Scope of the study | 6 |
| 1.6 | Dissertation Organisation | 6 |
| СНАР | PTER TWO: LITERATURE REVIEW | 7 |
| 2.1 | Introduction | 7 |
| 2.2 | Life Cycle Stages of Plasmodium | 7 |
| | 2.2.1 Different Species of Plasmodium | 8 |
| 2.3 | Evaluating Diagnostic Techniques for Malaria | 9 |
| 2.4 | Advancement in Artificial Intelligence and Machine Learning for | Malaria |
| Diag | gnosis | 12 |
| | 2.4.1 Computational methods for image and object detection | 12 |
| 2.5 | Computer-aided diagnosis (CAD) | 13 |
| | 2.5.1 Image acquisition | 13 |
| | 2.5.2 Preprocessing | 13 |

| | 2.5.3 | Image segmentation | 14 |
|--|---|--|--|
| | 2.5.4 | Feature extraction | 14 |
| | 2.5.5 | Image Classification | 15 |
| 2.6 | Machine Learning (ML) | | |
| | 2.6.1 | Supervised learning | 16 |
| | 2.6.2 | Unsupervised learning | 16 |
| | 2.6.3 | Semi-supervised learning | 17 |
| | 2.6.4 | Machine Learning algorithms | 17 |
| 2.7. | Deep Learning | | 17 |
| | 2.7.1 | Convolutional Neural Network (CNN) | 18 |
| | 2.7.2 | Transfer learning (TL) | 19 |
| | 2.7.3 | Pre-Trained Convolutional Neural Network Models | 19 |
| 2.8 | Related | Works | 23 |
| 2.9 | Summa | ry of Findings | 30 |
| | 2.9.1 | Gaps | 30 |
| | | | |
| CHAP | TER TH | REE: RESEARCH METHODOLOGY | 31 |
| CHAP 3.1 | TER TH Introdu | REE:_RESEARCH METHODOLOGY ction | 31 31 |
| CHAP: 3.1 3.2 | TER TH Introdu Data Co | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition | 31 31 31 |
| CHAP: 3.1 3.2 | TER TH Introdu Data Co 3.2.1 | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition | 31 31 31 31 |
| CHAP: 3.1 3.2 3.3 | TER TH Introdu Data Co 3.2.1 Image I | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing | 31 31 31 31 31 |
| CHAP 3.1 3.2 3.3 3.4 | TER TH Introdu Data Co 3.2.1 Image I Model s | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection | 31 31 31 31 31 31 32 |
| CHAP 3.1 3.2 3.3 3.4 3.5 | TER TH Introdu Data Co 3.2.1 Image I Model s Model 7 | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning | 31 31 31 31 31 32 33 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 1 | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation | 31 31 31 31 31 32 33 33 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 | TER TH Introdu Data Co 3.2.1 Image I Model S Model I Develop | REE:_RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application | 31 31 31 31 31 32 33 33 33 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 1 Develop Ethical | REE: RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application Approval | 31 31 31 31 31 32 33 33 33 34 35 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 CHAP | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 2 Develop Ethical | REE: RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application Approval | 31 31 31 31 31 32 33 33 34 35 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 CHAP | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 2 Develop Ethical | REE: RESEARCH METHODOLOGY ction blection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application Approval UR: RESULTS AND DISCUSSION | 31 31 31 31 31 31 32 33 33 34 35 36 36 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 CHAP 4.1 4.2 | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 2 Develop Ethical TER FO Preamb | REE: RESEARCH METHODOLOGY ction ollection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application Approval UR: RESULTS AND DISCUSSION le | 31 31 31 31 31 32 33 33 34 35 36 36 36 |
| CHAP 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 CHAP 4.1 4.2 | TER TH Introdu Data Co 3.2.1 Image I Model 2 Model 2 Develop Ethical TER FO Preamb Data Co | REE: RESEARCH METHODOLOGY ction blection and Image Acquisition Image Acquisition Preprocessing selection Training and Fine-tuning Evaluation pment of Web Application Approval UR: RESULTS AND DISCUSSION le blection and Pre-processing | 31 31 31 31 31 32 33 33 34 35 36 36 36 36 36 |

| | 4.2.2 | Visual Representation of Different Stages | 37 |
|------|---------------------------|---|---------|
| 4.3 | Buildir | ng Classifiers | 39 |
| | 4.3.1 | Customize Transfer learning models with fine-tuned classification | ı layer |
| | | | 39 |
| | 4.3.2 | Comparative Summary of the Experiment | 47 |
| 4.4 | Interac | tive Image Classification Web Application | 48 |
| | | | |
| CHAP | FER FI | VE: CONCLUSION AND RECOMMENDATION | 51 |
| 5.1 | Summa | ary | 51 |
| 5.2 | Contribution to Knowledge | | 51 |
| 5.3 | Recommendations | | 51 |
| 5.4 | Limita | tions | 52 |
| 5.5 | Conclu | ision | 52 |
| | | | |

REFERENCES

53

LIST OF FIGURES

| FIGURES | TITLE OF FIGURES | PAGES |
|-------------|---|-------|
| Figure 2.1 | AlexNet Convolutional Neural Network Architecture | 20 |
| Figure 2.2 | Inception Module Architecture | 21 |
| Figure 2.3 | Oxford University's Visual Geometric Group Architecture | 22 |
| Figure 2.4 | ResNet Architecture | 23 |
| Figure 2.5 | ResNet-50 Architecture | 23 |
| Figure 3.1 | Experimental Flowchart | 35 |
| Figure 4.1 | Ring Stage Image | 38 |
| Figure 4.2 | Trophozoite Stage Image | 38 |
| Figure 4.3 | Schizont Stage Image | 39 |
| Figure 4.4 | Gametocyte Stage Image | 39 |
| Figure 4.5 | Training and validation accuracy and Training and validation loss | 40 |
| Figure 4.6 | Confusion matrix for VGG16 | 41 |
| Figure 4.7 | VGG19 model architecture summary | 42 |
| Figure 4.8 | Training and validation accuracy and Training and validation loss | 42 |
| Figure 4.9 | Confusion Matrix | 43 |
| Figure 4.10 | Training and validation accuracy and Training and validation loss | 44 |
| Figure 4.11 | Confusion matrix | 44 |
| Figure 4.12 | MobileNet model architecture summary | 46 |
| Figure 4.13 | Training and validation accuracy and Training and validation loss | 46 |
| Figure 4.14 | Confusion matrix | 47 |
| Figure 4.15 | Web application First page | 49 |
| Figure 4.16 | Image upload of an infected cell | 49 |
| Figure 4.17 | Classification of Infected Cell | 55 |

LIST OF TABLES

| TABLE | TITLE OF TABLES | PAGES |
|-----------|--|-------|
| Table 1.1 | Table of Objectives VS Methodology | 4 |
| Table 4.1 | Malaria Image Dataset Composition | 37 |
| Table 4.2 | Dataset Splitting | 37 |
| Table 4.3 | Classification Report for VGG16 | 42 |
| Table 4.4 | Classification Report for VGG19 | 44 |
| Table 4.5 | Classification Report for ResNet50 | 45 |
| Table 4.6 | Classification Report for MobileNet | 47 |
| Table 4.7 | Comparison between each of the retrained model | 48 |

ABSTRACT

The worldwide influence of malaria has hastened the creation and execution of innovative diagnostic approaches aimed at combating the disease. Several attempts have been made to develop an automated malaria diagnosis system (AMDS), but most of these systems are trained to perform binary classification to distinguish between infected and uninfected Plasmodium falciparum parasites. This study investigates the application of deep learning techniques in developing an Automated Malaria Diagnosis System (AMDS) to enhance the accuracy and efficiency of malaria life cycle stage detection from thin blood smear images. The main objectives are to curate a novel dataset of thin blood smear images with the lifecycle stages of Plasmodium and to leverage pre-trained convolutional neural networks (CNNs) to identify the life cycle stages of Plasmodium, which are responsible for malaria. The study specifically focuses on evaluating the performance of different CNN architectures, including VGG16, VGG19, ResNet50, and MobileNet. The methodology involved curating a dataset of annotated thin blood smear images representing various life stages of the Plasmodium parasite. This dataset was then used to fine-tune the selected CNN models. The models were evaluated based on metrics such as accuracy, precision, recall, F1-score, and support. The results showed that VGG16 achieved the highest accuracy of 0.72, but its precision (0.47) and F1-score (0.49) indicated room for improvement in classification performance. VGG19, while slightly lower in accuracy at 0.71, demonstrated better precision (0.55) and recall (0.64), resulting in a higher F1-score (0.56). ResNet50, commonly recognized for its robustness in other domains, underperformed with an accuracy of 0.61 and a notably lower recall of 0.26. MobileNet displayed moderate results, with an accuracy of 0.68 and balanced precision and recall values. The findings suggest that VGG19 offers a more balanced performance for malaria stage classification, making it a promising candidate for deployment in a clinical setting. However, further optimization and refinement of these models are necessary to improve diagnostic precision and reliability. The study concludes that deep learning models, particularly VGG19, hold significant potential in supporting malaria diagnosis and contributing to more effective disease management and control strategies.

Keywords: Automated Malaria Diagnosis System, Convolutional Neural Networks, Deep Learning, Plasmodium, Thin Blood Smear Images, Web Application