DEVELOPMENT OF A LIGHTWEIGHT MODEL FOR COVID-19 FACE MASK WEARING-POSITION DETECTION

BY

ORUMA, SAMSON OGHENEOVO
(19PCK01991)

B.Eng Electrical / Electronics Engineering, University of Benin, Benin-City.

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF ENGINEERING (M.Eng) DEGREE IN INFORMATION AND COMMUNICATION ENGINEERING IN THE DEPARTMENT OF ELECTRICAL AND INFORMATION ENGINEERING, COLLEGE OF ENGINEERING, COVENANT UNIVERSITY.

SEPTEMBER, 2021

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 Engineering in Information and Communication Engineering in the Department of Electrical and Information Engineering, College of Engineering Covenant University, Ota, Nigeria

Mr John A. Philip
(Secretary, School of Postgraduate Studies) Signature and Date

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

DECLARATION

I, **ORUMA, SAMSON OGHENEOVO (19PCK01991)** declare that this research was carried out by me under the supervision of Prof. Francis E. Idachaba of the Department of Electrical and Information Engineering, College of Engineering, Covenant University, Ota, 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.

ORUMA, SAMSON OGHENEOVO **Signature and Date**

CERTIFICATION

We certify that this dissertation titled **"DEVELOPMENT OF A LIGHTWEIGHT MODEL FOR COVID-19 FACE MASK-WEARING POSITION DETECTION"** is an original research work, carried out by **ORUMA, SAMSON OGHENEOVO (19PCK01991)** in the Department of Electrical and Information Engineering, College of Engineering, Covenant University, Ota, Ogun State, Nigeria under the supervision of Prof. Francis E. Idachaba. We have examined and found this work acceptable as part of the requirements for the award of Master of Engineering.

Prof. Francis E. Idachaba
(Supervisor) **Signature and Date**

Prof. Emmanuel Adetiba
(Head of Department) **Signature and Date**

Prof. Adeseko B. Ayeni
(External Examiner) **Signature and Date**

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

DEDICATION

I dedicate this dissertation to my wife Blessing and my son Emmanuel, whose love and support kept me focused during this work.

ACKNOWLEDGEMENTS

My profound acknowledgement goes to the National Information Technology Development Agency (NITDA) for sponsoring this program through the National Information Technology Development Education Fund (NITDEF) scholarship scheme. Thanks for the transparency in the selection process and for the overwhelming support we received throughout this program.

Special thanks to my supervisor Prof. Francis Idachaba, for his guidance and direction throughout this work. My sincere thanks go to my other faculty and mentors of the Department of Electrical and Information Engineering for their selfless services. To my fellow EIE coursemates, I acknowledge the wonderful time we shared while on campus. You will always be in my heart!

Thanks to Blessing and Emmanuel Oruma for their understanding and patience for all the days I was absent from home in the course of this dissertation.

TABLE OF CONTENTS

CONTENT		Page
COVER PAGE		
TITLE PAGE		**i**
ACCEPTANCE		**ii**
DECLARATION		**iii**
CERTIFICATION		**iv**
DEDICATION		**v**
ACKNOWLEDGEMENTS		**vi**
LIST OF FIGURES		**ix**
LIST OF TABLES		**xi**
LIST OF ABBREVIATIONS		**xii**
ABSTRACT		**xiii**
CHAPTER ONE:	**INTRODUCTION**	**1**
1.1	Background of Study	1
1.2	Statement of the Research Problem	3
1.3	Aim of Study	4
1.4	Scope and Limitation of Study	5
1.5	Justification for the study	5
1.6	Dissertation Organization	6
CHAPTER TWO:	**LITERATURE REVIEW**	**7**
2.1	Computer Vision	7
2.2	Artificial Intelligence, Machine Learning and Deep Learning	10
2.3	Object Detection	14
2.4	The YOLO (You Only Look Once) Series	16
2.5	Related Works	36

CHAPTER THREE: METHODOLOGY **40**
3.1 Approval from Covenant Health Research Ethics Committee (CHREC) 40
3.2 Dataset Collection 42
3.3 Data Preprocessing 44
3.4 Model Selection 48
3.5 Model Training and Validation 49
3.6 Model Testing 54
3.7 Model Deployment 54
3.8 Model Evaluation Metrics and Visualization Tools 55

CHAPTER FOUR: RESULTS AND DISCUSSION **59**
4.1 Generated Dataset: Black Face Mask Dataset 59
4.2 Training and Validation Results 64
4.3 Testing and Benchmarking Results 68
4.4 Model Deployment 69

CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS **72**
5.1 Summary 72
5.2 Conclusion 73
5.3 Contribution to Knowledge 73
5.4 Recommendations 74

REFERENCES **74**

Appendix A: Nigerian National Code for Health Research Certificate **83**
Appendix B: Informed Consent Form **84**
Appendix C: CHREC Approval Letter **85**
Appendix D: Python Script for train.py **86**
Appendix E: Python Script for detect.py **98**
Appendix F: Python Script for val.py **104**
Appendix G: Python Script for export.py **111**
Appendix H: Python Script for hubconf.py **115**

LIST OF FIGURES

Figures	Title of Figures	Page
2.1	Human and computer vision compared.	8
2.2	Relationship between AI, ML and DL	10
2.3	Structure of a Neural Network	13
2.4	Neural network operation	13
2.5	Object detection approaches	15
2.6	The original YOLO model illustration with 7×7 grid cell input image (Redmon, Divvala, Girshick, & Farhadi, 2016)	18
2.7	YOLOv1 model architecture (Redmon, Divvala, Girshick, & Farhadi, 2016)	19
2.8	Prediction of bounding boxes in YOLOv2 using sigmoid activation function (Redmon & Farhadi, 2017)	22
2.9	Skip connection illustration in ResNet (He, Zhang, Ren, & Sun, 2016)	25
2.10	YOLOv3 multi-scale detection (Ayoosh, 2018)	27
2.11	Architecture of one-stage and two-stage object detectors (Bochkovskiy, Wang, & Liao, 2020)	27
2.12	Block diagram of state-of-the-art computer vision methods investigated by YOLOv4 authors (Jocher, 2020)	28
2.13	Structure of DenseNet (Z. Huang et al., 2020)	28
2.14	CSP and DenseNet illustrated	29
2.15	The DC-SPP-YOLO Model Architecture adopted in YOLOv4 (Z. Huang et al., 2020)	29
2.16	(a) Spatial Pyramid Pooling (SPP), (b) Improved Spatial Pyramid Pooling	30
2.17	PAN framework. (a) Backbone (b) Bottom-up path augmentation (c) Adaptive feature pooling (d) Box branch (e) Fully connected fusion (Shu Liu, Qi, Qin, Shi, & Jia, 2018b)	31
2.18	Modified PAN in YOLOv4 (Jocher, 2020)	31
2.19	Comparison of YOLOv3 and YOLOX head section. (Z. Ge, Liu, Wang, Li, & Sun, 2021)	36
3.1	Block diagram of research methodology	40
3.2	FujiFilm digital camera used for data capturing	42
3.3	Study locations for dataset collection. Source: Google Earth.	43
3.4	Sample dataset for both male and female	44
3.5	Variety of face mask type, colours and participants in the created dataset	45

3.6 Cropping operation with NCH PhotoPad Image Editor 46
3.7 Bulk image resizing operation with Pixillion Image Converter 46
3.8 Image annotations (a) with LabelImg software, (b) with roboflow.ai label assist tool 47
3.9 A sample of the dataset after data augmentation 48
3.10 YOLOv5 Model Architecture used for this study 49
3.11 Snapshots from Google Colab Notebook during model training 51
3.12 Model's parameters and Summary 53

4.1 Class balance analysis of Black Face Mask Dataset 62
4.2 Dataset annotation heatmap 63
4.3 Histogram of Object Count for the dataset 64
4.4 Train and validation metrics on Roboflow's platform 65
4.5 Confusion matrix for all classes 67
4.6 BBox Loss from Google Colab Training 67
4.7 Objectness Loss from Google Colab Training 68
4.8 Classification Loss from Google Colab Training 68
4.9 Precision, Recall and Mean Average Precision (mAP) from Google Colab Training 69
4.10 (a) Precision vs confidence curve (b) Recall vs confidence curve 69
4.11 Performance Metrics: (a) Precision vs Recall Curve (b) F1 vs Confidence Curve 69
4.12 Label correlogram 70
4.13 Sample of test images 70
4.14 Sample test images of benchmarking datasets 71
4.15 Deployment Outcome on Roboflow's Platform 71

A.1 Certificate obtained by principal investigator before the commencement of data collection 83

B.1 Informed Consent Form administered to participants before data collection 84

C.1 CHREC Approval Letter for Research Protocol 85

LIST OF TABLES

Tables	Title of Tables	Page
1.1	Variants of Covid-19 (CDC, 2020)	2
2.1	Machine learning and Deep learning compared	12
2.2	Summary of major Improvements in the YOLO series	17
2.3	Darknet-19 model structure (Redmon & Farhadi, 2017)	24
2.4	Architecture of Darknet-53 employed in YOLOv3 (Redmon & Farhadi, 2018)	26
2.5	Summary of commonly used face mask datasets	37
2.6	Summary of related works	39
3.1	List of classes for generated dataset	42
3.2	Digital Camera specification ("FujiFilm Finepixs4300 Full Specifications", 2021)	42
3.3	Geographical Co-ordinates of study location	43
3.4	Summary of generated dataset	45
3.5	Dataset size before and after data augmentation	47
3.6	Annotation formats available in Roboflow	50
4.1	Dataset class description	60
4.2	Annotation formats available in Black Face Mask Dataset	61
4.3	Training Validation and test results from Roboflow platform	65
4.4	Model's performance after training and validation	66
4.5	Summary of benchmarking dataset performance	70

LIST OF ABBREVIATIONS

Abbreviations	Meaning
2D	Two-dimensional
3D	Three-dimensional
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BCE	Binary Cross-Entropy
BFMD	Black Face Mask Dataset
CDC	Centres for Disease Control and Prevention
CHREC	Covenant Health Research Ethics Committee
CIoU	Complete Intersection over Union
CITI	Collaborative Institutional Training Initiative
CLIP	Contrastive-Language Image Pretraining
CmBN	Cross-mini-Batch Normalization
CNN	Convolutional Neural Networks
Covid-19	Coronovirus disease of 2019
CPU	Central Processing Units
CSP	Cross Stage Partial
CSV	Comma Seperated Values
CVPR	Computer Vision and Pattern Recognition
DenseNet	Dense Networks
DIoU	Distance Intersection over Union
DIoU-NMS	Distance Intersection over Union with Non-Maximum Suppression
DL	Deep Learning
DNN	Deep Neural Network
DPM	Deformable Part-based Model
EMA	Exponential Weighted Average
FCOS	Fully Convolutional One Stage Object Detection
FMDD	Face Mask Detection Dataset
FMLD	Face Mask Labelled Dataset
FN	False Negatives
FP	False Positives
FPN	Feature Pyramid Networks
fps	frame per seconds
GAN	Generative Adversarial Networks
GFLOPs	Giga Floating Points Operations Per Seconds
GIoU	Generalized Intersection over Union
GPU	Graphics Processsing Units
HIoU	Harmonic Intersection over Union
HOG	Histogram of Oriented Gradients
IDE	Integrated Development Environment
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
IOU	Intersection over Union
JSON	JavaScript Object Notation
LFW	Labelled Face in the Wild
LIDAR	Light Detection and Ranging

Abbreaviations	Meaning
MAFA	Masked Faces
mAP	mean Average Precision
MFDD	Masked Face Detection Dataset
MFN	MaskedFaceNet
MiWRC	Multi-input Weighted Residual Connection
ML	Machine Learning
MLP	Multi-Layer Perceptrons
MMD	Medical Mask Dataset
MS COCO	Microsoft Common Object in Context
NLP	Natural Language Processing
PAN	Path Aggregation Networks
PANet	Path Aggregation Networks
PIoU	Pixel Intersection over Union
Pr	Probability
Q-learning	Quality (of estimated utility function) - learning
R-CNN	Region-based Convolutional Neural Networks
ReLU	Rectified Linear Units
ResNet	Residual Networks
RFB	Receptive Field Block Network
R-FCN	Region-based Fully Connected Networks
RMFRD	Real-world Mask Face Recognition Dataset
RPN	Region Proposal Networks
SAM	Spatial Attention Module
SIFT	Shift Invariant Feature Transforms
SMFRD	Simulated Masked Face Recognition Dataset
SPP	Spatial Pyramid Pooling
SSD	Single Shot Multibox Detectors
SSL	Semi-Supervised Machine Learning
SVM	Support Vector Machines
TP	True Positives
TXT	Text
VOC	Visual Object Classes
WF	WilderFace
WHO	World Health Organization
XML	Extensible Markup Language
YOLO	You Only Look Once
YOLOv1	You Only Look Once version 1
YOLOv2	You Only Look Once version 2
YOLOv3	You Only Look Once version 3
YOLOv4	You Only Look Once version 4
YOLOv5	You Only Look Once version 5
YOLOv5l	You Only Look Once version 5 large size
YOLOv5m	You Only Look Once version 5 medium size
YOLOv5s	You Only Look Once version 5 small size
YOLOv5x	You Only Look Once version 5 extra-large size
YOLOX	You Only Look Once version X

ABSTRACT

The Corona Virus Disease (Covid-19) spread has led to many infection cases with several resulting deaths. The increasing number of new Covid-19 variants has reinforced the need to develop a proactive critical preparedness, readiness response action plan. This study aims to develop a lightweight model for detecting face mask-wearing positions using a locally generated dataset of black people. A six-fold methodology of dataset generation, data preprocessing, model selection, model training and validation, and model deployment was adopted for this study. A dataset of black people from three universities in Nigeria was generated for males and females, indicating four distinct face mask-wearing positions and eight classes, using a digital camera, smartphones and medical face mask. The images were subject to preprocessing such as cropping, resizing, labelling and data augmentation. The generated dataset was used to train a modified YOLOv5s model and deployed using Roboflow's webcam platform and local PC with Pycharm IDE. The developed model achieved 94.2%mAP, 94% recall, and 79% precision on Roboflow's platform after training for 250 epochs. Training on Google Colab platform for 100 epochs resulted in 91.5% mAP, 91.4% recall and a precision of 86.8%. A dataset called "Black Face Mask Dataset" was generated from this study, with 13 different annotation formats. This study's outcome will be beneficial to researchers in computer vision and the government of developing countries. The generated dataset can be merged with the existing face mask detection dataset to achieve a better model with good black people representation. the eight classes can be merged into smaller classes based on the application requirement to produce higher object numbers per class. The developed model can be cloned from GitHub for easy integration without the need for retraining.

Keywords: Object detection, Image classification, Covid-19 prevention, Face mask detection, Black-coloured people, Deeplearning.