FACE MASK DETECTION FOR COVID-19 STANDARD OPERATING PROCEDURE BY USING DEEP LEARNING

TAN ZHI WEI

A project report submitted in fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic Systems)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

FEBRUARY 2022

DEDICATION

This project report is dedicated to my family, who have supported and encourage me unconditionally throughout the process. They have taught me the value of hard work and unceasingly support me to strive towards success.

ACKNOWLEDGEMENT

I spoke with a lot of people while writing this thesis, including researchers, academics, and practitioners. They have helped me to better comprehend and think about things. I'd like to express my gratitude to Assoc. Prof. Ir. Ts. Dr. Nasrul Humaimi Mahmood, my major thesis supervisor, for his support, advice, criticism, and friendship. This thesis would not have been the same if it hadn't been for his continual attention and encouragement.

Next, I cherish the valuable feedback and thought from my panel of examiners Assoc. Prof. Ts. Dr. Eileen Su Lee Ming and Ts. Dr. Nor Aini Zakaria during the presentation phase.

My colleague postgraduate student's help should also be acknowledged. My sincere gratitude also goes out to all of my co-workers and those who have helped me on various occasions. Their suggestions and opinions are beneficial to this project report. Unfortunately, there isn't enough place to mention them all in this area. Finally, I want to express my gratitude to every one of my family members.

ABSTRACT

In 2019, a new and highly infectious disease emerged in Wuhan, China, and quickly spread around the world. SARS-Cov 2 (also known as COVID-19) is the illness. The coronavirus COVID-19 pandemic is wreaking havoc on the world's health system, infecting over 180 million people, and killing over 3.8 million people. The virus transmission method was spread through respiratory droplets when an infected person coughs, sneezes or even speaks Even though a vaccine is available, there is still no effective treatment for this condition, and even if one is vaccinated, one can still be diagnosed. Therefore, the best approach to deal with it is to avoid it, and many medical experts have recommended wearing a face mask as one of the most effective ways to stop the virus from spreading. Aside from that, numerous countries throughout the world have enacted new laws or guidelines requiring individuals to wear face masks on a regular basis. However, some people continue to refuse to use a face mask when visiting public areas, especially in crowded places. As a result, stationing a security guard at the entry to monitor visitors appears to be the alternative. This approach, however, not only puts the guards in risk, but it also has the potential to cause overcrowding at the gate due to its inefficiency. Machine learning is undeniably the key to averting this downfall by minimizing direct human participation. Over the year, in field of image processing and computer vision, the spotlight was more focus on only face detection rather than face mask detection therefore the vulnerabilities of face mask detection technologies have not been properly addressed. Hence, the first objective of this paper is to implement the deep learning in image recognition for face mask detection. Next, the objective will be developing a system that able to detect whether a person is wearing a face mask or not by utilizing Convolutional Neural Network (CNN). In this project, the CNN architecture, MobileNetV2 is being utilised due to its low computational cost. A total of 3486 images of face masked and without face masked datasets are created from various online open-sourced datasets and fed to the model. Several optimizers, including SGD, RMSProp, and Adam, are evaluated to obtain the optimal network model. Finally, the Adam optimizer is chosen, and optimization techniques such as epoch size, batch size, and initial learning rate are gradually tuned and applied to the model. The validation accuracy could reach 99% throughout the tuning process, and the validation loss was decreased from 4.13 % to 2.86 %. The result model was then compared to another state-of-the-art CNN model, VGG-16, and the results reveal that the MobileNetV2 model did indeed utilise fewer computing resources, as it consumed 11% less memory, had a 5 times smaller result model, and took 5 times less time to train than that in VGG-16. When the model was put to the test with 50 real-life example images, the model able to achieve accuracy of 86% which the model able to detect the face in the image and correctly labelled it. In the end, the system able to detect face and distinguish the face with or without face mask and thus help in face mask detection to prevent the spread of COVID-19.

ABSTRAK

Pada tahun 2019, penyakit baharu dan berjangkit muncul di Wuhan, China, dan merebak dengan cepat ke seluruh dunia. SARS-Cov 2 (juga dikenali sebagai COVID-19) adalah penyakit tersebut. Pandemik coronavirus COVID-19 menyebabkan malapetaka kepada sistem kesihatan dunia, ia menjangkiti lebih 180 juta orang, dan menyebabakan kematian lebih 3.8 juta orang. Kaedah penularan virus ini disebarkan melalui titisan pernafasan apabila seseorang yang dijangkiti batuk, bersin atau bercakap Walaupun vaksin tersedia, tiada rawatan berkesan untuk keadaan ini, dan walaupun seseorang telah divaksin, dia masih boleh didiagnosis. Oleh itu, tindakan terbaik untuk menanganinya adalah mengelakkannya, dan ramai pakar perubatan telah mengesyorkan memakai topeng muka sebagai salah satu cara paling berkesan untuk menghentikan virus daripada merebak. Selain itu, banyak negara di seluruh dunia telah melaksanakan undang-undang atau garis panduan baharu untuk mewajibkan individu memakai topeng muka. Namun, segelintir orang terus enggan menggunakan topeng muka ketika mengunjungi kawasan awam terutama di tempat yang sesak. Akibatnya, menempatkan pengawal di pintu masuk untuk memantau pengunjung nampaknya menjadi alternatif. Tindakan ini, bukan sahaja menyebabkan pengawal dalam risiko, tetapi ia juga berpotensi menyebabkan kesesakan di pintu pagar kerana ketidakcekapannya. Pembelajaran mesin tidak dapat dinafikan adalah kunci untuk mengelakkan masalah ini dengan meminimumkan penyertaan manusia secara langsung. Sepanjang tahun, dalam bidang pemprosesan imej dan penglihatan komputer, tumpuan lebih tertumpu pada pengesanan muka sahaja berbanding pengesanan topeng muka oleh itu kelemahan teknologi pengesanan topeng muka tidak ditangani dengan teliti. Oleh itu, objektif pertama kertas kerja ini adalah untuk melaksanakan pembelajaran mendalam dalam pengecaman imej untuk pengesanan topeng muka. Seterusnya, objektif kedua adalah membina sistem yang dapat mengesan sama ada seseorang itu memakai topeng muka atau tidak dengan menggunakan Convolutional Neural Network (CNN). Dalam projek ini, seni bina CNN, MobileNetV2 telah digunakan kerana kos pengiraannya yang rendah. Sebanyak 3486 imej set data bertopeng muka dan set data tanpa topeng muka telah dicipta daripada pelbagai set data sumber terbuka dalam internet dan disalurkan kepada model. Beberapa optimizers, termasuk SGD, RMSProp dan Adam, telah dinilai untuk mendapatkan model rangkaian yang optimum. Akhirnya, “optimizers” Adam dipilih, dan teknik optimization seperti “epoch size”, “batch size” dan “initial learning rate” tealh ditala secara beransur-ansur dan digunakan pada model. Ketepatan pengesahan model boleh mencapai 99% sepanjang proses penalaan, dan kehilangan pengesahan dikurangkan daripada 4.13% kepada 2.86%. Model hasil juga dibandingkan dengan model CNN tercanggih yang lain, seperti VGG-16, dan hasilnya mendedahkan bahawa model MobileNetV2 sememangnya menggunakan lebih kurang sumber pengkomputeran, kerana ia menggunakan 11% kurang memori, mempunyai 5 kali lebih kecil. model hasil, dan mengambil masa 5 kali lebih kurang berbanding dengan VGG-16. Apabila model diuji dengan 50 imej contoh kehidupan sebenar, model ini dapat mencapai ketepatan 86% yang mana model itu dapat mengesan wajah dalam imej dan melabelnya dengan betul. Akhirnya, sistem ini dapat mengesan muka dan membezakan muka dengan atau tanpa topeng muka dan dengan itu, ia dapat membantu dalam pengesanan topeng muka untuk mencegah penularan COVID-19.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

WHO	-	World Health Organization
AI	-	Artificial Intelligence
CNN	-	Convolutional Neural Network
ReLU	-	Rectified linear unit layer
ROI	-	Region of interest
NMS	-	Non-maximum suppression
SVM	-	Support vector machine
RPN	-	Region proposal network
YOLO	-	You Only Look Once
SSD	-	Single Shot Multi-Box Detector
PCA	-	Principal Component Analysis
FPN	-	Feature Pyramid Network
RMFD	-	Real-World Masked Face Dataset
SMFD	-	Simulated Masked Face Dataset
LFW	-	Labeled Faces in the Wild
MMD	-	Medical Masks Dataset
FMD	-	Face Mask Dataset
FMLD	-	Face-Mask Label Dataset
MAFA	-	MAsked FAces
IoU	-	mean Intersection over Union
DSFD	-	Dual Shot Face Detector
BAIR	-	Berkeley AI Research
SGD	-	Stochastic Gradient Descent
RMSProp	-	Root Mean Square Propagation
Adam	-	Adaptive Moment Estimation

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Problem Background

In 2019, a new and contagious disease had broken out in Wuhan, China and eventually spread out to the whole world. This disease is the SARS-Cov 2 also known as COVID-19. In 2020, the rapid spreading of COVID-19 has forced the World Health Organization (WHO) to declare COVID-19 as a global pandemic. In figure 1.1, as of 27 June 2021, COVID-19 has infected more than 179 million people worldwide and caused over 3.8 million deaths [1].

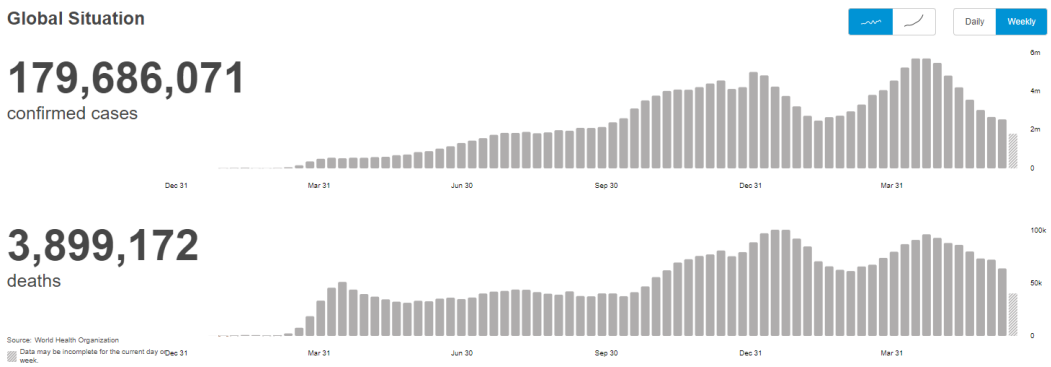

Figure 1.1 Confirmed cases and death counts for COVID-19 [1]

COVID-19 is very infectious, and one of the ways it spreads is by droplets of saliva or nasal secretions when an infected person coughs, sneezes, or even talks. It may spread quickly, especially in crowded environments. COVID-19 infections must be restricted by preventive approaches because there is no therapy for COVID-19 [2]. The World Health Organization (WHO), medical professionals, and governments throughout the world are now recommending that people wear face masks in public areas to help prevent the spread of this new coronavirus. Wearing masks has been demonstrated to minimize the likelihood of coronavirus transmission [3], making it one of the most effective preventive measures now available. The correct method to

wear a mask, according to the World Health Organization (WHO), is to adjust it to cover the mouth, nose, and chin. If masks are not worn properly, protection is substantially diminished.

As a result of the coronavirus pandemic, new levels of worldwide scientific cooperation have emerged. Artificial Intelligence (AI) based on Machine Learning and Deep Learning can contribute to the combat against COVID-19 in a variety of ways. Researchers and medical experts are able to analyze large amounts of data to predict the hotspot of the spreading of COVID-19 through machine learning and the outcome can act as an early alert system for potential pandemics and identify risk groups. As previously discussed, wearing a face mask is one of the most efficient ways to help restrict the spread of the virus; in many countries, people are required by law to wear face masks in public. The process of monitoring large amounts of people, on the other hand, is difficult to implement and challenging. In response to these advances, the computer vision community has recently focused on face mask detection research, to build automatic detection models that can aid society in limiting the COVID-19 pandemic.

1.2 Problem Statement

Face mask detection has been popular because of the COVID-19 pandemic, which demands people to wear face masks, keep social distance, and wash their hands with hand sanitizers. While other precautionary measures such as social distancing and sanitization issues have been handled, face mask detection has yet to be properly addressed. COVID-19 infections can only be prevented at this time because there is no specific therapy for it. Various vaccines, such as Pfizer, Sinovac, AstraZeneca, and others, have been developed to avoid infection and severe COVID-19 effects. However, since the COVID-19 virus continues to evolve, it still can infect a completely vaccinated individual [4]. As a result, wearing a mask during this pandemic is both a necessary precaution and crucial especially when preserving social distance is difficult. A mask is required for anybody who in danger of serious sickness from COVID-19 illnesses, especially those who are at higher risk. COVID-19 is discovered

to transmit mostly among persons who are near one another around 2 meters, but those who have no symptoms and are clueless that they are infected might spread the virus too.

Furthermore, despite the fact that laws and rules requiring individuals to wear face masks while entering public areas have been enacted, some people continue to refuse to do so. There is a cause for this, such as clashing messages that have confused individuals. Those who may be conflicted about whether to wear a mask may be influenced by opposing messages in the media and from authoritative figures such as government authorities. Furthermore, some mask refusers view the matter solely in terms of their own comfort; out of thoughtlessness or selfishness, they are unaware and ignore the effect of face mask on preventing the virus from spreading not only to them but also to their family members and community. They fail to recognize the value of doing something for the larger good, even if it is uncomfortable and inconvenient [5]. At the moment, one method of resolving this problem is to station security guards in public locations to urge people to wear masks. This approach, however, not only exposes the guards to virus-infected air but also causes overcrowding at the entrance due to its inefficiency. Therefore, a quick and effective solution such as automated face mask detection is needed to deal with this situation.

In the field of image processing and computer vision, face mask detection has shown to be a difficult challenge. Face detection has a variety of applications, ranging from face recognition to collecting facial movements, the latter of which necessitates revealing the face with incredible accuracy. For example, some existing face detector model like Dlib frontal face detector, Haar cascades, MTCNN, etc, has shown great performance in detecting a face. Face recognition without a mask is simpler, but face recognition with a mask is more difficult because masked face feature extraction is more complex than regular face feature extraction. Many facial characteristics, such as the nose, lips, and chin, are missing from the covered face. Despite significant advancements in the field of machine learning algorithms, the vulnerabilities of face mask detection technologies have not been properly addressed.

1.3 Research Objectives

The objectives of this project are:

(a) To implement the deep learning in image recognition for face mask detection.

(b) To develop a system that able to detect whether a person is wearing a face mask or not by utilizing Convolutional Neural Network (CNN).

(c) To analyses the accuracy of the system in face mask detection.

1.4 Project Scope

The scopes of this project are:

(a) The CNN model MobileNetV2 is chosen to do the image classification and the algorithm is implemented in Python.

(b) The face detector model is using a pretrained model provided in OpenCV.

(c) To compile a face masked and no face masked dataset for input of network from various online open-source datasets.

1.5 Thesis Outline

The thesis is organised into five chapters. The 1st chapter begins with an introduction that provides a brief overview of the COVID-19 outbreak. In addition, this chapter contains the thesis's problem statements, research objectives, and research scope. Next, Chapter 2 presents a literature review of the fundamentals of deep learning on image classification and object detection, as well as some research on related work on deep learning face mask detection. Following that, Chapter 3 explains the methodology of this study on data collecting, data pre-processing, and data

analysis. There are two main phases of this methodology: the first phase is the training of the face detector, and the second phase is the deployment of the trained face detector. In Chapter 4, it shows the results, analysis, and discussion for the experimental controlled factors. The attributes of different optimizers are introduced with preliminary data. The different optimization approaches are then applied to the basic model. Tables and graphs are used to examine each tuning iteration. Furthermore, the comparison against the other CNN model is conducted with judgement. Lastly, Chapter 5 concludes the findings of the project report and suggests a few recommendations for possible further research.

REFERENCES

[1] W. H. Organization, "WHO Coronavirus (COVID-19) Dashboard," 27 June 2021. [Online]. Available: https://covid19.who.int/. [Accessed 27 June 2021].

[2] Q. Cai, M. Yang, D. Liu, J. Chen, D. Shu, J. Xia and X. Liao, "Experimental treatment with favipiravir for COVID-19: an open-label control study.," *Engineering*, vol. 6, no. 10, pp. 1192-1198, 2020.

[3] S. E.Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich and A. B.Gumel, "To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic," *Infectious Disease Modelling*, vol. 5, pp. 293-308, 2020.

[4] Bernama, "Covid: Fully vaccinated Malaysian man among 18 new cases reported in Singapore on June 12," 13 June 2021. [Online]. Available: https://www.theedgemarkets.com/article/covid-fully-vaccinated-malaysian-man-among-18-new-cases-reported-singapore-june-12. [Accessed 13 June 2021].

[5] C. Gillespie, "Why Do Some People Refuse to Wear a Face Mask in Public?," 5 January 2021. [Online]. Available: https://www.health.com/condition/infectious-diseases/coronavirus/face-mask-refuse-to-wear-one-but-why. [Accessed 10 June 2021].

[6] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

[7] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way," Towards Data Science, 17 December 2018. [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53. [Accessed 25 May 2021].

[8] Alex Krizhevsky, S. Ilya and E. H. Geoffrey, "ImageNet classification with deep convolutional neural networks.," *Commun. ACM,* vol. 60, no. 6, p. 84–90, 6 June 2017.

[9] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand and M. &. A. H. Andreetto, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[11] M. Hollemans, "MobileNet version 2," [Online]. Available: https://machinethink.net/blog/mobilenet-v2/. [Accessed 26 May 2021].

[12] L. Hulstaert, "A Beginner's Guide to Object Detection," DataCamp Community, 19 April 2018. [Online]. Available: https://www.datacamp.com/community/tutorials/object-detection-guide. [Accessed 31 May 2021].

[13] Z. Zou, Z. Shi, Y. Guo and J. Ye, "Object Detection in 20 Years: A Survey," *CoRR*, vol. abs/1905.05055, 16 May 2019.

[14] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and," in *Proceedings of the IEEE conference on computer vision and pattern recognition,*, 2014.

[15] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015.

[16] S. Ren, K. He, R. Girshick and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal," *Advances in neural information processing systems,* p. 91–99, 2015.

[17] J. Redmon, R. G. S. Divvala and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.

[18] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "SSD: Single shot multibox," in *European conference on computer vision*, Springer, 2016.

[19] A. Nieto-Rodríguez and M. Mucientes, "System for Medical Mask Detection in the Operating Room Through Facial Attributes," *Pattern Recognition and Image Analysis,* vol. 9117, pp. 138-145, 2015.

[20] M. S. Ejaz, M. R. Islam, M. Sifatullah and A. Sarker, "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition," in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019.

[21] M. Jiang, X. Fan and H. Yan, "RetinaMask: A Face Mask detector," *CoRR,* vol. abs/2005.03950, 8 Jun 2020.

[22] M. Loey, G. Manogaran, M. H. N.Taha and N. E. M.Khalif, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement,* vol. 167, 2021.

[23] M. Loey, G. Manogaran, M. H. N.Taha and N. E. M.Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustainable Cities and Society,* vol. 65, 2021.

[24] B. QIN and D. LI, "Identifying Facemask-Wearing Condition Using Image Super-Resolution with Classification Network to Prevent COVID-19," *Sensors,* vol. 20, no. 18, p. 5236, 2020.

[25] B. Batagelj, P. Peer, V. Š. and S. Dobrišek, "How to Correctly Detect Face-Masks for COVID-19 from Visual Information?," *Applied Sciences* , vol. 11, no. 5, p. 2070, 2021.

[26] borutb-fri, "Face Mask Label Dataset (FMLD)," 25 February 2021. [Online]. Available: https://github.com/borutb-fri/FMLD. [Accessed 1 June 2021].

[27] P. Bhandary, "Mask Classifier," Github, 15 Apr 2020. [Online]. Available: https://github.com/prajnasb/observations. [Accessed 25 May 2021].

[28] X-zhangyang, "Real-World Masked Face Dataset, RMFD," GitHub, 12 Jan 2021. [Online]. Available: https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset. [Accessed 26 May 2021].

[29] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," *ArXiv,* vol. abs/1712.04621, 2017.

[30] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. B. Girshick, S. Guadarrama and T. J. Darrell, "Caffe: Convolutional Architecture for Fast Feature Embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014.

[31] V. Agarwal, "Face Detection Models: Which to Use and Why?," Towards data science, 3 July 2020. [Online]. Available: https://towardsdatascience.com/face-detection-models-which-to-use-and-why-d263e82c302c. [Accessed 29 May 2021].

[32] G. E. Hinton, "A Practical Guide to Training Restricted Boltzmann Machines," *Neural Networks: Tricks of the Trade,* vol. 7700, pp. 599-619, 2012.

[33] Y. Bengio, "Practical Recommendations for Gradient-Based Training of Deep Architectures," *Neural Networks: Tricks of the Trade,* vol. 7700, pp. 437-478, 2012.

[34] A. Rosebrock, Deep Learning for Computer Vision with Python, PYIMAGESEARCH, 2017.

[35] J. Niesen, J. Martinot-Lagarde, G. Peña-Castellanos, G. Vaillant, S. Silvester, C. v. d. Meeren and divenex, "Spyder memory profiler plugin," 18 Apr 2020. [Online]. Available: https://github.com/spyder-ide/spyder-memory-profiler.