



THE CLASSIFICATION OF CORNEAL ARCUS IMAGES BY USING IMPROVED CONVOLUTIONAL NEURAL NETWORK



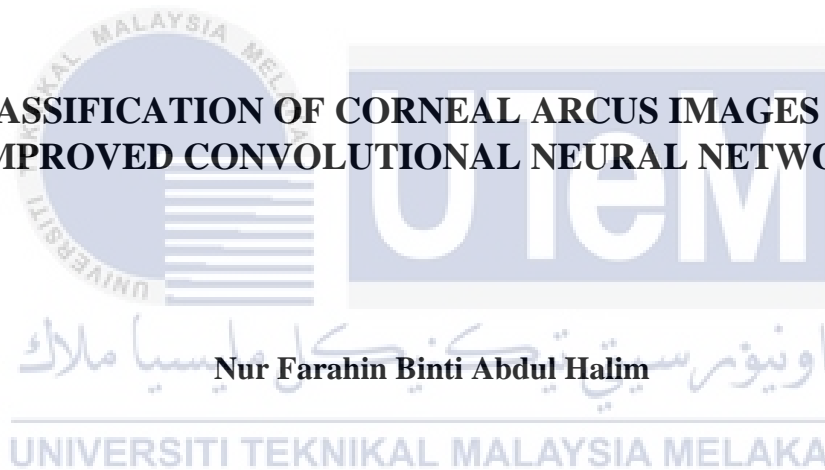
MASTER OF SCIENCE IN ELECTRONIC ENGINEERING

2024



**Faculty of Electronics and Computer Technology and
Engineering**

**THE CLASSIFICATION OF CORNEAL ARCUS IMAGES BY USING
IMPROVED CONVOLUTIONAL NEURAL NETWORK**



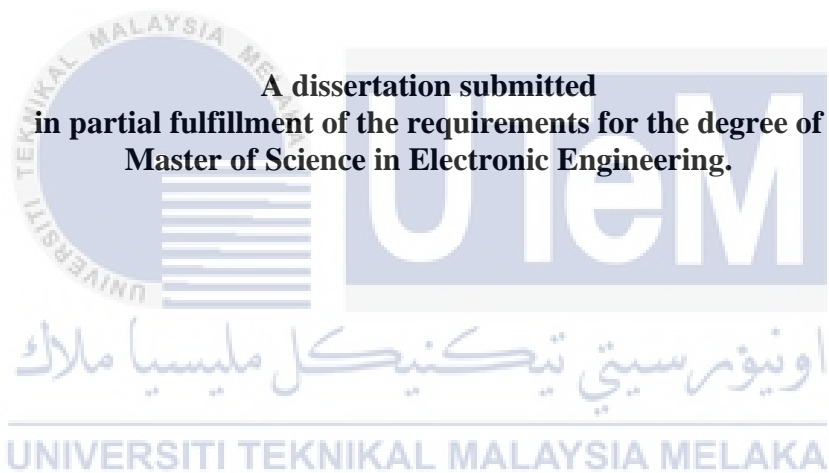
Nur Farahin Binti Abdul Halim

Master of Science in Electronic Engineering

2024

**THE CLASSIFICATION OF CORNEAL ARCUS IMAGES BY USING
IMPROVED CONVOLUTIONAL NEURAL NETWORK**

NUR FARAHIN BINTI ABDUL HALIM



Faculty of Electronics and Computer Technology and Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2024

DEDICATION

All praises to Allah, the Most Gracious and the Most Merciful. To my beloved husband, family, friends and my supervisor Ir. Dr. Ridza Azri Bin Ramlee. This accomplishment is not just my own, but a testament to the collective efforts of all those who have supported me. I dedicate this achievement to them, with a heart full of gratitude. Thank you very much.



ABSTRACT

Corneal arcus (CA), also known as arcus senilis or simply arcus, refers to a condition characterized by the accumulation of lipids, particularly cholesterol, in the cornea of the eye. The presence of CA may be an indicator of underlying systemic conditions such as high cholesterol in human body. When CA is detected during an eye examination, it may prompt further investigation into the patient's lipid profile and overall cardiovascular health. Individuals with CA may be advised to undergo blood tests to measure cholesterol and triglyceride levels. The most common test is known as a lipid panel or a cholesterol blood test. Therefore, leveraging image processing offers a non-invasive and painless alternative to traditional blood tests for detecting corneal arcus. This research is about the implementation of convolutional neural network (CNN) in detecting cholesterol's presence by classifying normal and CA images. A dataset of 459 images comprising 237 for normal and 222 for CA images were formed. There are three different CNN models were proposed for feature extraction and classifying the normal and CA images which are CNN, Resnet-50 and VGG-19. From the parameter evaluation, it can be concluded that batch size of 20 and learning rate of 0.0001 suit with Resnet-50 and CNN model, while VGG-19 suit with batch size of 10 and learning rate of 0.00001 to classify with the normal and CA dataset. The best result was exhibited by Resnet-50 with 10-fold cross-validation producing high average detection in terms of sensitivity, specificity, and accuracy of 100%. Thus, deeper networks implementation is recommended in the future to further improve CA localisation in cholesterol detection.

اوتنور سیتی تکنیکل ملیسیا ملاک

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

KLASIFIKASI IMEJ ARKUS KORNEA DENGAN MENGGUNAKAN RANGKAIAN NEURAL PELINGKARAN YANG DIPERBAIKI

ABSTRAK

Kornea Arkus (KA), juga dikenali sebagai arkus senilis atau ringkasnya arkus, merujuk kepada keadaan yang dicirikan oleh pengumpulan lipid, terutamanya kolesterol, dalam kornea mata. Kehadiran KA boleh menjadi penunjuk keadaan di mana kolesterol tinggi dalam badan manusia. Apabila KA dikesan semasa pemeriksaan mata, pakar perubatan akan menasihati pesakit untuk membuat pemeriksaan kesihatan yang lanjut ke atas profil lipid pesakit dan kesihatan kardiovaskular secara keseluruhan. Individu yang mempunyai KA akan dinasihatkan untuk menjalani ujian darah untuk mengukur paras kolesterol dan trigliserida. Ujian darah yang dikenali sebagai lipid panel atau ujian darah kolesterol akan mengukur pelbagai jenis kolesterol dan trigliserida dalam darah. Penyelidikan ini adalah mengenai pelaksanaan Rangkaian Neural Pelingkaran (CNN) dalam mengesan kehadiran kolesterol dengan mengklasifikasikan imej normal dan KA. Oleh itu, set data 459 imej yang terdiri daripada 237 untuk imej biasa dan 222 untuk imej KA telah dibentuk. Terdapat tiga model CNN berbeza telah digunakan untuk mengelaskan imej biasa dan KA iaitu CNN, Resnet-50 dan VGG-19. Daripada penilaian parameter, dapat disimpulkan bahawa kumpulan saiz 20 dan kadar pembelajaran 0.0001 sesuai dengan model Resnet-50 dan CNN, manakala VGG-19 sesuai dengan kumpulan saiz 10 dan kadar pembelajaran 0.00001 untuk mengklasifikasikan dengan normal dan CA dataset. Ringkasnya, hasil terbaik dipamerkan oleh Resnet-50 dengan pengesahan silang 10 kali ganda menghasilkan pengesanan purata tinggi dari segi kepekaan, kekhususan dan ketepatan 100%. Oleh itu, pelaksanaan rangkaian yang lebih mendalam disyorkan pada masa hadapan untuk menambah baik pengelasan KA dalam pengesanan kolesterol.

ACKNOWLEDGEMENTS

ALLAH SWT is the best planner. I am decided to continue my study on the year of COVID-19. The beginning was so hard for me to keep on doing the research from home because I do not have a suitable source to run my project. It takes time almost 1 year after the lockdown for me to access the facility provided by the faculty for me to run the project. As time goes by, I managed to run this project using the facility provided in Research Lab 3 (MLSP). Despite, many difficulties and constraints I had faced, there are many people I would like to thank you for giving me support and encouragement to complete this research and without whom I would not have made it through my master degree! First of all, I would like to express my sincere gratitude to my supervisor, Ir. Dr. Ridza Azri Bin Ramlee whose insight and knowledge into the subject matter steered me through this research. My colleagues for giving me consistent support and had to put up with my stresses and moans for the past three years of study! and my biggest appreciation to my husband, Muhammad Irshad Bin M. Abu Zaki, my parent and family for all the support you have shown me and always pray for my continued success in no matter what it takes or how long it takes to complete this research even though my mom always ask me to complete the master as quickly as possible. Next, thank you to all that directly or indirectly include in the completion of this master research.

TABLE OF CONTENTS

	PAGES
DECLARATION	
APPROVAL	
DEDICATION	
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
LIST OF APPENDICES	xii
LIST OF ABBREVIATIONS	xiv
LIST OF PUBLICATIONS	xvii
CHAPTER	
1. INTRODUCTION	1
1.1 Project Overview	1
1.2 Problem Statement	7
1.3 Research Objective	10
1.4 Research Scopes	11
1.5 Research Significance	12
1.6 Thesis Outline	13
2. LITERATURE REVIEW	16
2.1 Prevalence and Types of Eye Abnormalities	16
2.2 Invasive Technique Detecting Cholesterol	18
2.3 Datasets	23
2.3.1 UBIRIS Dataset	25
2.3.2 Abnormal Images	26
2.4 Image Processing Technique	27
2.4.1 Application of Image Enhancement	27
2.4.1.1 Contrast Limited Adaptive Histogram	28
2.4.1.2 Unsharp Masking	30
2.4.2 Image Quality Analysis (IQA)	31
2.4.2.1 Mean Squared Error (MSE)	32
2.4.2.2 Peak Signal-to-Noise Ratio (PSNR)	32

2.4.3	Application of Image Augmentation	33
2.5	Deep Learning Technique	34
2.5.1	Convolutional Neural Network (CNN)	36
2.5.1.1	Convolutional Layer	37
2.5.1.2	Pooling/Subsampling Layer	38
2.5.1.3	Fully Connected Layer	40
2.5.1.4	Dropout Layer	41
2.5.1.5	Softmax Layer	42
2.5.2	Visual Geometry Group (VGG)	43
2.5.3	Residual Network (ResNet)	44
2.6	Justification of Deep Learning Model Selection	47
2.7	Performance Measure	49
2.7.1	Confusion Matrix	49
2.7.2	Standard Deviation	51
2.8	Conclusion	51
3.	METHODOLOGY	53
3.1	Hardware and Software Requirement	53
3.2	Project Flowchart	54
3.3	Sample Preparation	57
3.4	Image Pre-processing	58
3.4.1	Image Augmentation	58
3.4.2	Image Enhancement	60
3.4.2.1	CLAHE	60
3.4.2.2	Unsharp Masking	62
3.5	Proposed Deep Learning – Convolutional Neural Network	63
3.5.1	Fine-tuning ResNet-50	64
3.5.2	Fine-tuning VGG-19	68
3.5.3	CNN	72
3.6	Parameter Optimization	74
3.7	Performance Evaluation	76
3.7.1	Training Options	76
3.7.2	Confusion Matrix	77
3.7.3	Mean Average	78
3.7.4	Standard Deviation	78
3.8	Conclusion	80
4.	RESULT AND DISCUSSION	82
4.1	Image Pre-Processing Evaluation	82
4.2	Training-validation of Model Analysis	85
4.2.1	Resnet-50 Performance	85
4.2.2	VGG-19 Performance	90
4.2.3	CNN Performance	96
4.3	Comparison of Models Performances	102

4.4	Testing the Model	103
4.4.1	Original Dataset	104
4.4.2	CLAHE + Unsharp Masking Dataset	105
4.4.3	CLAHE Dataset	106
4.5	Conclusion	107
5.	CONCLUSION AND FUTURE WORK	110
5.1	Summary and Contributions of the Research	110
5.2	Recommendations for Future Works	113
	REFERENCES	115
	APPENDICES	125



LIST OF TABLES

TABLE	TITLE	PAGE
1.1	The summary of the total dataset for normal and abnormal eye images.	12
2.1	The comparison of non-invasive method in detecting cholesterol.	21
2.2	The selection of deep learning model from previous CA research using deep learning approach.	49
3.1	Hardware specification used in this research.	53
3.2	The Parameters of Resnet-50 model.	67
3.3	The parameters of VGG-19 model.	70
3.4	The parameters of CNN model.	74
3.5	The proposed value of learning rate and batch size.	75
3.6	The confusion matrix for CA classification.	77
4.1	The overall performance of image quality assessment.	84
4.2	The accuracy performance evaluation of Resnet-50.	85
4.3	The specificity (Sp.), F1-score (F1-Sc.) and negative predictive value (NPV) of training-validation evaluation Resnet-50 model.	89
4.4	The overall performance evaluation of VGG-19.	90
4.5	The specificity (Sp.), F1-score (F1-Sc.) and negative predictive value (NPV) of training-validation evaluation VGG-19 model.	95
4.6	The overall performance evaluation of CNN.	96
4.7	The specificity (Sp.), F1-score (F1-Sc.) and negative predictive value (NPV) of training-validation evaluation CNN model.	101

4.8	The comparison between batch size 20 and 30 through validation loss and time (s).	101
4.9	The evaluation performance of the selected batch size and learning rate parameter.	103
4.10	The elapsed time of each model to train-validation.	103
4.11	The testing performance on each model using original dataset.	104
4.12	The loss of testing per iteration on each model using original dataset.	105
4.13	The testing performance on each model using CLAHE and unsharp masking dataset.	105
4.14	The loss of testing per iteration on each model using CLAHE and unsharp masking dataset.	106
4.15	The testing performance on each model using CLAHE dataset.	107
4.16	The loss of testing per iteration on each model using CLAHE dataset.	107

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	The illustration of type of cholesterols.	3
1.2	The comparison of normal and abnormal eye.	6
1.3	An illustration of clear cornea region.	6
2.1	Type of eye disability (a) Corneal Arcus (Panahi-Bazaz et al., 2014), (b) Cataract (Ang and Afshari, 2021), (c) Pterygium (Zulkifley, Abdani and Zulkifley, 2019), (d) Glaucoma (Mitchell C. Latter, M.D.: Board-Certified Ophthalmologist, 2024), (e) Conjunctivities (NHS, 2021).	18
2.2	The sample of UBIRIS v2 Database (Proença et al., 2010)	25
2.3	The sample of images acquired from several journals; (a) (Mahesh Kumar and Gunasundari, 2016), (b) (Fernández et al., 2007), (c) (Panahi-Bazaz et al., 2014),.	26
2.4	The general overview subset of machine learning and artificial intelligence.	35
2.5	The brief model of VGG neural network model.	44
2.5	A brief architecture of Resnet-50.	46
3.1	Block diagram of the proposed CA classification using deep learning model.	55
3.2	Research project flowchart.	56
3.3	The sample of faulty images from UBIRIS dataset.	58

3.4	An example of abnormal image rotation (a) The original CA infected eye image, (b) Samples of horizontal rotation.	59
3.5	The shift translation on training-validation dataset, (a) The original training-validation CA image, (b) The sample of 180°rotation on CA image, (c) The sample of 250°rotation on CA image.	60
3.6	The contrast enhancement using CLAHE, a) Raw images, b) Enhancement using CLAHE.	62
3.7	A result of CLAHE eye image applied with unsharp masking, a) CLAHE eye image, b) CLAHE with unsharp masking.	63
3.8	The comparison of Resnet-50 and fine-tuning Resnet-50, (a) The brief model of ResNet-50 (b) The fine-tuning of ResNet-50.	66
3.9	The comparison of VGG-19 and fine-tuned VGG-19 (a) The brief model of VGG-19 (b) The fine-tuning of VGG-19.	69
3.10	The CNN model (proposed method).	73
3.11	The training options for evaluating the training progress.	76
4.1	The accuracy result of Resnet-50, (a) The training accuracy of Resnet-50, (b) The validation accuracy of Resnet-50.	87
4.2	The standard deviation accuracy on each batch size of different learning rate on Resnet-50 model (BS10,30,40 and 50 refered to primary axis. BS30 refered to secondary axis).	88
4.3	The overall result of Resnet-50.	90
4.4	The accuracy results of VGG-19, (a) The training accuracy of VGG-19, (b) The validation accuracy of VGG-19.	93

4.5	The standard deviation accuracy on each batch size of different learning rate VGG-19, (a) The batch size of 10, (b) The batch size of 20, (c) The batch size of 30, (d) The batch size of 40, (e) The batch size of 50.	94
4.6	The overall result of VGG-19	96
4.7	The accuracy results of CNN, (a) The training accuracy of CNN, (b) The validation accuracy of CNN.	99
4.8	The standard deviation accuracy on each batch size of different learning rate CNN.	100
4.9	The overall result of CNN.	102



LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Verification of CA images.	125
B	The overall performance evaluation of Resnet-50 with batch-size 10.	127
C	The overall performance evaluation of Resnet-50 with batch-size 20.	130
D	The overall performance evaluation of Resnet-50 with batch-size 30.	133
E	The overall performance evaluation of Resnet-50 with batch-size 40.	135
F	The overall performance evaluation of Resnet-50 with batch-size 50.	137
G	The overall performance evaluation of VGG-19 with batch-size 10.	139
H	The overall performance evaluation of VGG-19 with batch-size 20.	141
I	The overall performance evaluation of VGG-19 with batch-size 30.	143
J	The overall performance evaluation of VGG-19 with batch-size 40.	146
K	The overall performance evaluation of VGG-19 with batch-size 50.	148

L	The overall performance evaluation of CNN with batch-size 10.	150
M	The overall performance evaluation of CNN with batch-size 20.	152
N	The overall performance evaluation of CNN with batch-size 30.	155
O	The overall performance evaluation of CNN with batch-size 40.	158
P	The overall performance evaluation of CNN with batch-size 50.	160

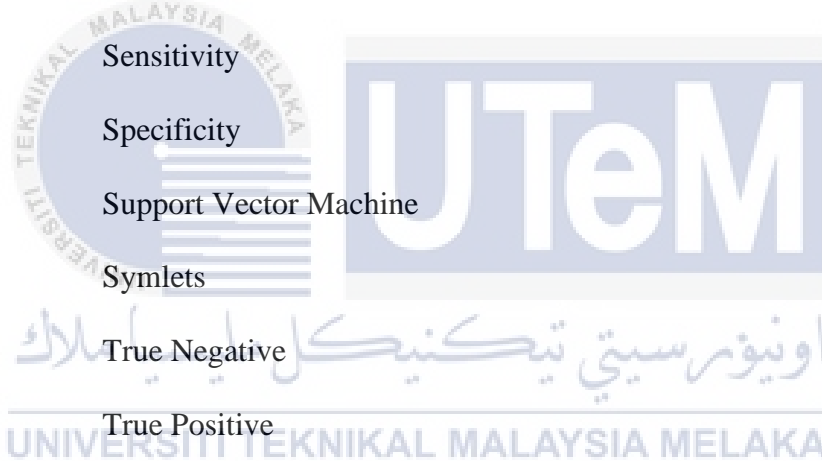


LIST OF ABBREVIATIONS

Acc	Accuracy
AG	Angiography
AHE	Adaptive Histogram Equalization
AI	Artificial Intelligence
BG	Biorthogonal
BPN	Backpropagation Neural Network
BR	Bayesian Regulation
CA	Corneal Arcus
CASIA	Chinese Academy of Science
CDF	Cumulative Distribution Function
CDR	Cholesterol Detection Regression
CHD	Coronary Heart Disease
CL	Connected Layer
CLAHE	Contrast Limited Adaptive Histogram Equalization
CLHE	Contrast Limited Histogram Equalization
CNN/ConvNet	Convolutional Neural Network
COCO	Common Objects in Context
COVID	Corona Virus
CPU	Central Processing Unit
DA	Data Augmentation
DB	Daubechies
dB	Decibels

DL	Deep Learning
FCL	Fully Connected Layer
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
GHz	Gigahertz
GLCM	Gray Level Co-Occurrence Metrics
GLM	Gray Level Mean
HDL	High Density Lipoprotein
HE	Histogram Equalization
HEF	High-Frequency Emphasis Filtering
ICE	Iris Challenge Evaluation
ILSVRC	Imagenet Large Scale Visual Recognition Challenge
IT	Information Technology
LDL	Low Density Lipoprotein
min	Minutes
ML	Machine Learning
MLP	Multilayer Perceptron
MLP-BP	Multilayer Perceptron
MMU	Multimedia University
MNF	Minimum Noise Function Transformation
MPL	Max-Pooling Layer
MSE	Mean Squared Error
NHANES	National Health and Nutrition Examination
NIR	Near Infrared

NM	Not Mention
NPV	Negative Predictive Value
P	Precision
PSNR	Peak Signal to Noise Ratio
RAM	Random Access Memory
RBF	Radial Basis Function
ReLU	Rectified Linear Regularization Unit
ResNet	Residual Network
RGB	Red-Green-Blue
s	Seconds
Sn	Sensitivity
Sp	Specificity
SVM	Support Vector Machine
SYM	Symlets
TN	True Negative
TP	True Positive
UBIRIS	University Of Beira Interior
UM	Unsharp Masking
US	United States
VGG	Visual Geometry Group
WV	Wavelets
WVU	West Virginia University



LIST OF PUBLICATIONS

The followings are the list of publications related to the work on this thesis:

Farahin, N., Halim, B. A., Azri, R., Ramlee, B., Zaki, M., Mas'ud, B., and Jamaludin, A., 2022. Enhancement of automatic classification of arcus senilis-nonarcus senilis using convolutional neural network. *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, 28(1), 210–220. <https://doi.org/10.11591/ijeecs.v28.i1.pp210-220> (SCOPUS indexed).

AH, N. F., Ramlee, R. A., Mas'ud, M. Z., and Alias, M. A. 2022. A Comparative Study of Deep Learning Parameters for Arcus Senilis Classification. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 14(4), 21-24.



CHAPTER 1

INTRODUCTION

This chapter mainly will provide an overview of corneal arcus condition as well as the background study of classification technique through deep learning method. The objectives of the project and challenges or problems that led to the development of this project are highlighted and discussed. Overall, this chapter presents the organization of this research study.

1.1 Project Overview

A healthy lifestyle means physical and mental health in a person. In many cases, the physical and mental are closely related to each other. Cholesterol, a fundamental lipid molecule, plays a pivotal role in the physiology of living organisms, serving as an indispensable component of cell membranes and a precursor for vital bioactive molecules. While essential for various physiological processes, an imbalance in cholesterol levels can lead to profound health implications, notably cardiovascular diseases, stroke, diabetes, high blood pressure and many more. High and low cholesterol in the blood vessels is one of the physical health issues that we must consider.

There are two types of cholesterol in lipoprotein transporters which are low density lipoproteins (LDL) and high-density lipoproteins (HDL). LDL is considered bad cholesterol that carries cholesterol particles throughout the body and cholesterol will accumulate in the

wall of the arteries. HDL considered as good cholesterol, it will pick up the excess cholesterol and take it to liver. Figure 1.1 below shows the illustration of LDL and HDL in human body. Usually, people that will be associated from high risk of developing high cholesterol are those who are living with obesity, consuming a lot of saturated and trans-fat like fast food, have a limited physical activity, smoking, having a family history of high cholesterol or having diabetes, kidney failure or hypothyroidism.

A report from the American Heart Association 2021 stated that according to National Health and Nutrition Examination Survey (NHANES) in 2015 to 2018, among US adults ≥ 20 years of age, the age-adjusted prevalence of obesity was 39.9% in males and 41.1% in females; the prevalence of extreme obesity was 6.2% in males and 10.5% in females; the overall prevalence of obesity among youth 2 to 19 years of age was 19.0% (Virani et al., 2021). Apart from this risk factor, it is wise to keep in mind that people all of ages, genders and ethnicities can have high cholesterol. This disease happened when the plaque (extra cholesterol that build up at the arteries) became hardened and narrowed at the wall of arteries which will block the flow of blood into heart. This condition is known as atherosclerosis which is a serious condition of that can limit the blood flow through the arteries and raise a risk of developing dangerous blood clot (Rafieian-Kopaei et al., 2014). Without any treatment or medication, hence, it will lead to many life-threatening complications like stroke, heart attack, chest pain or angina, high blood pressure peripheral vascular disease and chronic kidney disease.

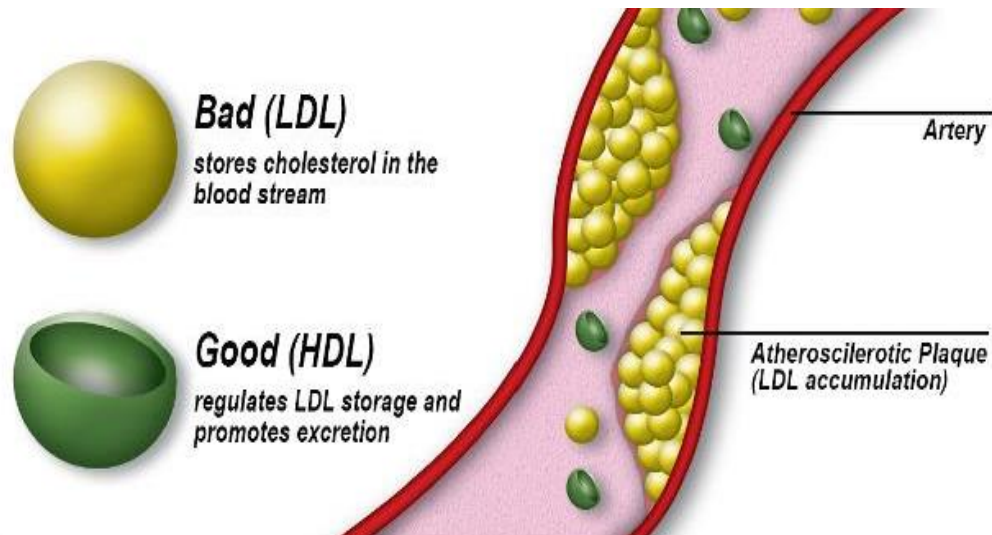


Figure 1.1: The illustration of type of cholesterol (Hosur, 2022).

Essentially, the focus of this research is to suggest non-intrusive way in detecting the cholesterol. Currently, there are a few ways in detecting level of cholesterol in blood vessel for instance lipid panel test. The test required patient to fast at least 10 hours fasting and if the patient suffered from heart attack, surgery, infection, or pregnancy they need to wait 2 months before having the test. A blood test will be performed using a syringe to extract the blood from the vein at the arm or finger prick in order to gain a blood sample to be test out the level of cholesterol. Sometime, there is a people that have a fear to needle also called needle phobia or trypanophobia. People who suffer from this phobia at times debilitating condition usually will experience symptoms like hypertension, rapid heart rate, fainting or loss consciousness, besides, they even fear associated with doctors and nurses. This will prevent them from seeking treatment when facing from any serious ailments for instances high cholesterol in body.

In conclusion, high cholesterol can be one of the root causes in triggering other disease to attract so, it is a must to have a preliminary detection on detecting cholesterol.