MONASH University

School Of Information Technology

FIT - 3161

Computer Science Project 1

Semester 1

by

Nabil Ahmed (ID: 25364170)

Kevin Setiawan (ID:25895710)

Project Proposal

Classification of Diabetic Retinopathy using Retinal Images (Convolutional Neural Network with MATLAB + Fusion Techniques)

1.0 Introduction

One the biggest causes of blindness in the world is Diabetic Retinopathy (DR). If a person is diagnosed with DR at an earlier stage, there are some effective treatments of DR that can be applied to prevent that blindness. For this reason, classification DR in real time is a vital procedure which can save a lot of people from going blind. The accuracy of the classification is very important in this context which if classified wrongly defeats the purpose of this project. It needs a trained clinician to detect DR. DR classification contains weighting of numerous features and also their location in the retina. This will take a lot of time for clinicians. But, the problem can be solved by computers. Computers can achieve quicker, more accurate, and more consistent results if they are trained correctly. Related work has been done to DR classification by a method that uses a k-NN classifier and vector machines. This method usually only classifies DR into two classes which is DR and no DR. Our project will classify DR into three classes which are no DR, mild DR and severe DR. So, it will be easier for people who has responsibility in medical can treat the patient correctly.

Convolutional Neural Networks(CNN) is division of deep learning that has extraordinary result for image classification. The previous work, CNN used for image analysis and interpretation that related into retinal images to analyze the some features that cause DR. By computing in high graphical processor unit(GPU), CNN possible for complicated image detection problem. Nowadays, CNN are successfully complete a hard image detection with a lot of classes. But CNN also has disadvantages. The first drawbacks are on achieving a required offset in sensitivity and specificity. Where sensitivity is for patient properly has DR

and specificity is for patient properly has not DR. This is quite hard because the project will identify the patient from retinal images into three classes (no DR, mild DR, severe DR). Next, the other disadvantages are overfitting in neural network. A big data sometimes extremely tilted. Because it might be a big difference amount of data for each class.

In this project, we propose a classification DR with CNN architecture and fusion technique in MATLAB programming language. Because CNN become famous in image classification and many successful researches show impressive result. Fusion technique itself can increase the accuracy of image classification. So, by combining DR classification with CNN and fusion technique, it should be have a better result in classify DR into three classes. Retinal fundus images as an input will be provided by the supervisor. Total of 654 input images will be preprocessed and will be augmented before go to the networks. All the steps will be clearly stated below at methodology section.

2.0 Background

Diabetes may be a well-known illness that arises once the body cannot process the hormone secreted by the pancreas or once the pancreas miscarries to secrete enough of it (Verma, Deep et al. 2011). If the patient cannot control the blood sugar and suffer diabetes in long term, chances of having DR is getting higher as well. It will affect the retina of the patient which can cause mild vision problem until blindness (Doshi, Shenoy et al. 2016).

World Health Organization(WHO) made the survey relate to DR, there more than 347 million people in this world suffering diabetic retinopathy. They also predict that it still can be increase up to more than 500 million people by 2030. DR classification method that used by clinicians is by detecting lesion with vascular abnormalities, but it only can be done by highly trained clinicians and expert in this field. Not only the clinicians but also some certain equipment is needed to classify DR. DR become a major challenge in underdeveloped and overpopulated countries such as China and India. In underdeveloped country, number of trained clinicians who can expert in DR classification is so small and the equipment may not be provided. Because of that, machine learning can gain a better method to detect DR in real time (Pratt, Coenen et al. 2016). Therefore, automatic detection for DR classification is well-known to hold the increasing number of patient who suffering DR

Diabetic retinopathy can be divided into two big classes which is early DR and advanced DR. Early DR also acknowledged as Non-Proliferative Diabetic Retinopathy (NPDR). In this class, blood vessel walls of retina are getting weaker. Micro aneurysm made the blood vessel wall become small because if blood leaking and fluid directly into the eye. This class still can be

separate to different classes from mild DR until severe DR depends on how much blood vessel blocked. In getting worse in advanced DR, if the patient does not do any treatment or cure to the retina, it will lead the patient to blindness.

DR classification is very challenging because it need a highly-trained clinician and many specific equipment to accurately and effectively inspect DR (Doshi, Shenoy et al. 2016). The method that used by clinicians might be already effective and accurate. But if the demands are high, the clinician cannot fulfil all the demands. Because detecting DR into different classes require a lot of time to do. Since it highly time consuming for them, computer be able classify DR very fast and accurately when it is trained correctly.
As stated earlier, DR can be classified into different classes depending on how much the blood vessels is blocked and also some features that stay in the retina.

Blood Vessel and Hemorrhage Extraction

By using RGB fundus images, it will be easier to extract some image feature extraction. This picture below use adaptive histogram to improve the contrast.

Fig.1. The blood vessels shown in white pixels against the black background.

Fig.2. After applying some filter which transformed to binary corresponding with global threshold

Fig.3. Existence of broken lines were observed.

Fig.4. The density of blood vessels increase by many times. The density can be calculated by looking at total area of blood vessels

Fig.5. The image was thresholder using a universal thresholding value

Fig.6. The hemorrhages were detected and their density was calculated by finding the number of white pixels in the image

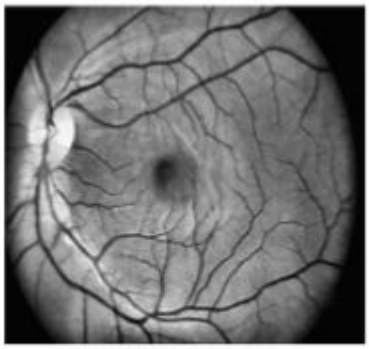


Figure 1. Retinal image after removing the noise.

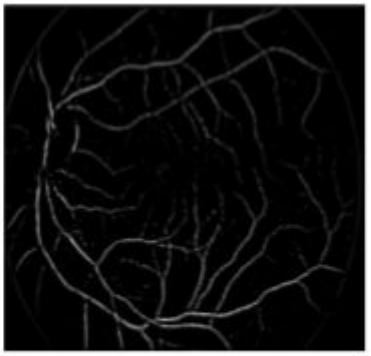


Figure 2. Image obtained after passing through the matched filter.

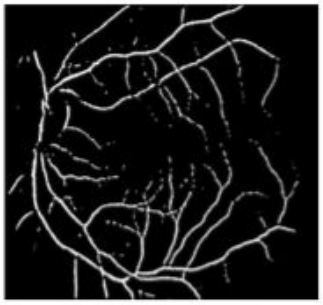


Figure 3. Figure 3: Binary image after thresholding.

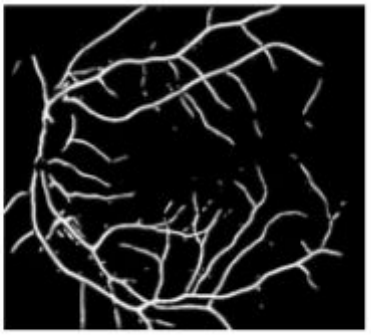


Figure 4. Image after perception based binarisation.

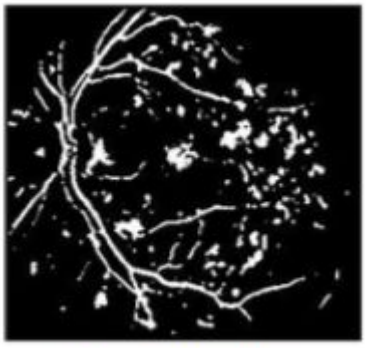


Figure 5. Thresholded image.

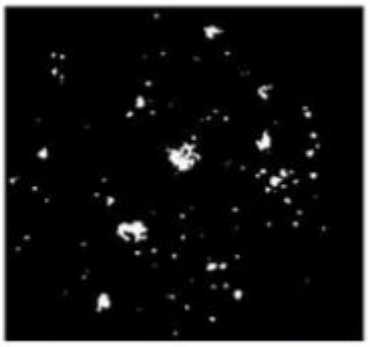


Figure 6. Hemorrhages in the retinal image.

From six pictures above, its shows the blood vessels and hemorrhages were extracted. The white pixel over the black background is blood vessels and hemorrhages. Both extracted features can be an input for the classifier to detect DR stages.

In order to classify DR, there are some features need to be extracted as stated above. Meaning that we can train the machine to learn how to detect it. There are several method to classify DR. For example,by using Neural Networks, Gardener et al having 84.4% and 83.5% for sensitivity and specificity result which only detecting DR and no DR. They also only use a limited dataset that only consist 200 images which also be divided into pieces before goes to SVM implementation.Neural Networks must boot been used in three-class classification of DR. Nayak et al used choices like area the globe the realm of exudates and the territory of blood vessels at the feature of texture parameters. choices area unit entered the neural network to classify photos into ancient, non-proliferative retinopathy and proliferative retinopathy. The neural network used these choices as input for classification. The detection results were valid by examination with grading from knowledgeable ophthalmologists. They incontestable a classification accuracy of ninety-three, sensitivity of ninetieth and specificity of 100%. This was assigned on a dataset of 140 photos and have extraction was required of all photos in every employment and testing which can be time overwhelming.

The remarkable majority of study on the five-class classification that has been assigned has used support vector machines (SVMs). Acharya et al have formed automatic procedure for characteristic the five classes. Features, that unit extracted from the knowledge using a higher-order varieties methodology, unit fed in to the SVM classifier and take the variation inside the shapes and contours inside the images.This SVM procedure according a mean accuracy of eighty-two, sensitivity eighty-two face up to specificity of half a mile. Acharya et al to boot created a five- class classification approach by perceptive the areas of the many choices like hemorrhages, microaneurysms, exudate and vessel. the choices determined to be the leading crucial; blood vessels, micro-aneurysms, exudates, and hemorrhages, were extracted from the raw photos practice image method techniques. These were then fed to

the SVM for classification. A sensitivity of eighty-two specificity of 86%and accuracy of 85.9% was achieved practice this method. These ways in which were per- designed on relatively very little datasets and therefore the telephone call sensitivity and specificity was probably attributable to the progressive nature of the five classes downside.

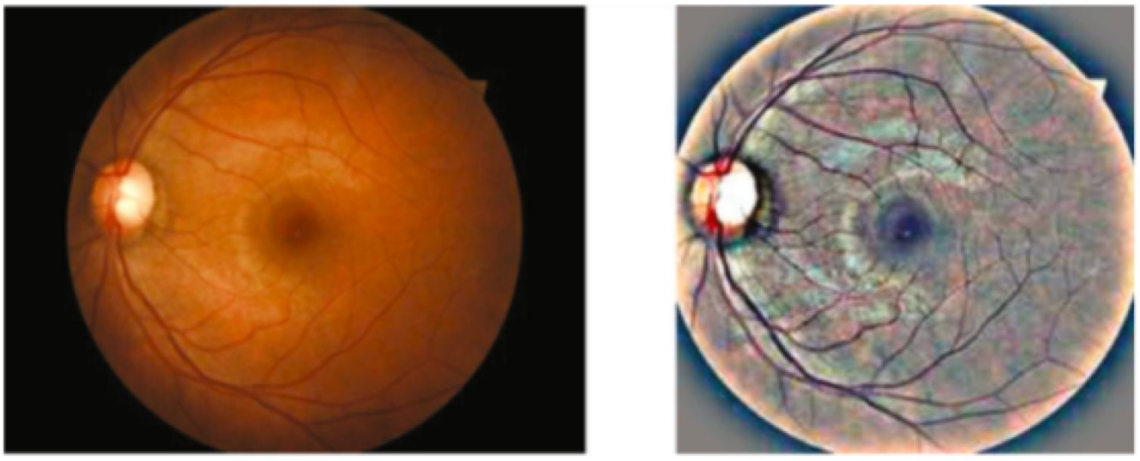
This project will more focus on DR classification using CNN approach. CNN has different network compare to other method that already stated above. In brive discussion, CNN genuine measure base on the convolutional layers in the constructions. The convolutional layers will calculate trial for examination native choices together with location of the input images. To get indigenous features, all node in convolutional layer is connected to a low set of spatially neurons within the input picture. To regulate the are seeking for an equivalent native feature throughout the enter channels, the association substances square measure common between the nodes inside the convolutional layers. each set of shared weights is named a kernel, or a convolution kernel. Thus, a convolutional layer with kernels learns to examine native choices whose power across the input pictures is visible within the ensuing function maps. To limit computer first-rate and convey home the bacon a gradable set of image options, every structure of convolution layers is observed via a pooling layer, a development corresponding to straightforward and elaborate cells within the main cortical region. The significant physical damage pooling layer reduces the scale of function maps through selecting the utmost feature response in overlapping or non-overlapping native neighborhoods, discarding the location of such most reactions. As a result, significant bodily damage pooling will any improve translation in unpredictability. CNN commonly incorporates many pairs of convolutional and pooling layers, accompanied with the aid of range of consecutive fully connected layers, and ultimately a softmax layer, or regression layer, to get the special outputs. in additional stylish CNN architectures, machine efficiency is carried out by using trade the pooling layer with a convolution layer with a progress larger than one. Like multilayer perceptions, CNN square measure skilled with the back-propagation rule by using reducing the following value.

CNN architecture is well-known as one of the method on images classification. Some pre-trained CNN such as AlexNet, VGG-16, GoogleNet, etc. has successfully classify images with a promising result. All pre-trained CNN have some differences on their network.

Meaning that they are using a different amount of layer on the architecture. For example, some architecture use a lot of convolutional layer but less of pooling layer and fully connected layer. The other use less convolutional layer but put a lot of layer of fully connected layer.

Related to Sun, Wan et al. 2017, the standard classification from retinal pictures will constructs an honest distinction in machine-controlled DR detection. Additionally, with increasing of application of convenient complex body part camera, larger range of retinal pictures may be taken. With poor quality of training dataset may end up low accuracy of detection of DR. The researchers propose four CNN architectures to classify retinal complex body part pictures. Then, they choose and be a part of the highest two networks. Four architectures that employed in this methodology is AlexNet, GoogleNet, VGG-16, and ResNet-50.

First, the all input images are going to be resized to 256x256 pixels to avoid the disadvantages from lighting conditions on the preprocessing stage. Also, the images square measure clipped to ninetieth of original size to decrease the black house. The result from preprocessing stages may be seen below.



(a) Original image (b) Preprocessed image

Next methodology is to try to to data augmentation that principally employed in training of CNN. The retinal images square measure rotates haphazardly or verticals flip with random horizontal or vertical shifts with random horizontal. With this methodology, the training set

redoubled concerning eight times. On the specification, deep CNN is pre-train established on large dataset and weight of train deep CNN are going to be set earlier.

AlexNet is planned to induce smart performance in ImageNet. This CNN represent five convolutional layer, max pooling layers, dropout layers and three fully connected layers.

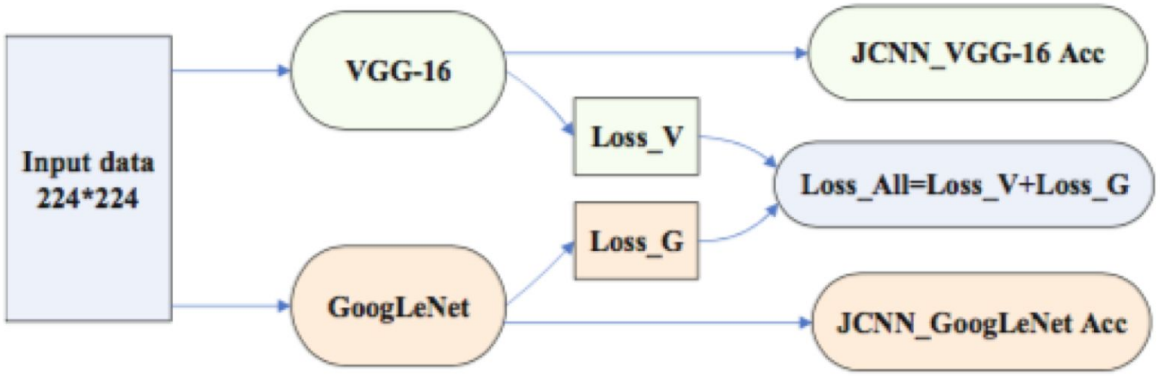
Google Net is that the design that uses totally different origin modules to create twenty-two layers.

VGG-16 solely uses three x three filters in convolutional layers and later mix as a sequence. This produce from thirteen convolutional layers, five max pooling layers and three fully connected layers.

ResNet-50 is that the methodology with solely three.6% error rate that contains residual block and skip association bypassing. ResNet has deeper substance which may consists a 101 or perhaps 152 layers. This design has vi modules. Conv1 as convolutional layer, conv2 x, conv3 x, conv4 x, conv5 x as residual block and fc as fully connected layer.

Join Fine-tuned CNN: they choose GoogleNet and VGG-16 as two high network. the whole loss of the network may be outlined from this

Loss_All = Loss_V + Loss_G



On the training part, by removing last absolutely connected layer and replace with two outputs owing to fully fine tune, solely the last one from layer is touched. the speed of learning last fully connected layer is redoubled ten times. For 4CNN design, the coaching information is place within the network at the side of pre-train weight parameters. Researches fine tune GoogleNet and VGG-16 at identical time with same input for every

channel. it'll backward circulation coaching technique to point out the classifier slope to any or all networks

Around 80,000 pictures square measure provided by Kaggle web site is effectively employed in the dataset. They choose 2894 pictures for sample and a 2170 pictures for coaching and check set. The experiment result shown below.

Algorithm	Acc	AUC
AlexNet	96.53%	0.993
GoogLeNet	97.04%	0.994
VGG-16	96.87%	0.995
ResNet-50	96.20%	0.992
JCNN_GoogLeNet	97.00%	0.995
JCNN_VGG-16	**97.12%**	**0.995**

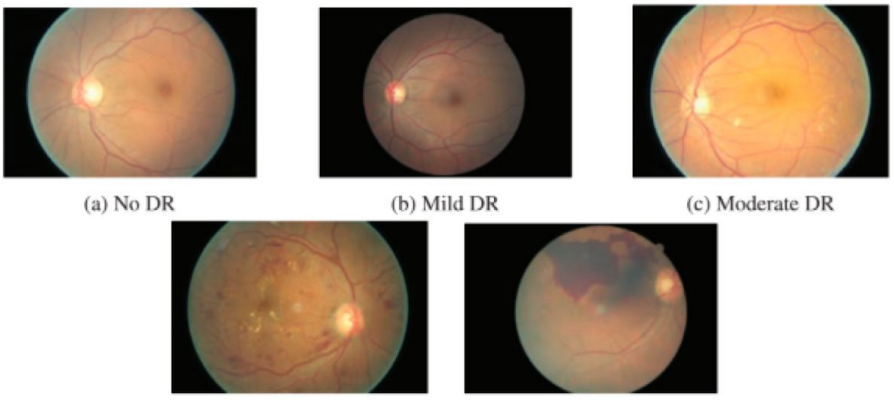
Algorithm	GoogLeNet	GoogLeNet-NP	GoogLeNet-NA
Accuracy	**97.04%**	96.12%	96.49%

As may be seen from table, all methodology performs an honest accuracy on the result. This shown that the information from taken from natural image may be transfer to create image classification effectively. Specially for JCNN between VGG-16 and GoogLeNet, each network already is that the best option network at the side of they build supported two field assumptions that create them cannot miss the classification behavior.

Pratt, Coenen et al. 2016 uses a CNN design for classifying the severity of Diabetic Retinopathy. They use digital complex body part pictures as input(testing). The options they target for the classification purpose square measure the microaneurysms, hemorrhages and exudate on the membrane because of these options square measure ascertained to be the foremost crucial once police work Diabetic Retinopathy (Acharya et al). once the input image is given the CNN design can mechanically classify the severity of DR supported the image. The design has already been trained by employing a high-end GPU (graphics

processor unit). The dataset used for coaching the design was taken from the publically accessible Kaggle dataset that consisted of 80,000 images. They show outstanding results of ninety fifth sensitivity and seventy fifth accuracy on 5000 validation pictures.

The planned CNN classifies the severity of the DR into the 5 stages:



(a) No DR (b) Mild DR (c) Moderate DR

(d) Severe DR (e) Proliferative DR

Fig 1: Stages of diabetic retinopathy (DR) with increasing severity

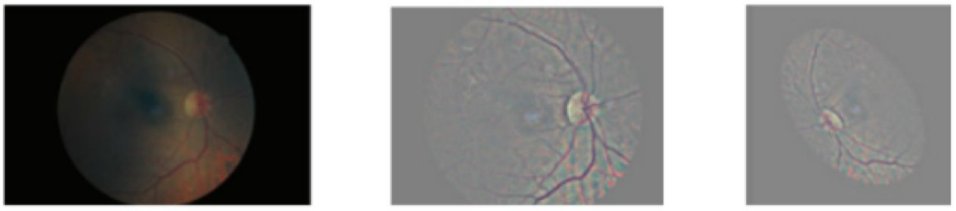
They additionally claim that they're the primary to classify DR into five stages employing a CNN approach.

The article is explored in additional details compared to different articles because it is totally and directly associated with our project literature review.

The Building of the CNN design for classifying the various stages of DR

Preprocessing is completed on all the photographs from the Kaggle dataset. The dataset consists of pictures from patients of variable age teams, individuals of various races and culture and intensely wide-ranged level of lighting within the complex body part photograph. This creates variations not associated with classification levels since the constituent intensity is affected. This downside was resolved by victimization color social

control. They additionally size the photographs to one explicit size to retain solely the complex options needed for distinguishing whereas doing classification. of these original preprocessed pictures, square measure then used for coaching the network once. Then, period information augmentation is employed throughout coaching to boost the localization ability.



(a) Original image (b) Preprocessed image (c) Augmented image

Fig 3: Illustration of the preprocessing and augmentation processes

The article proves that by victimization CNN it's attainable to classify DR into five totally different stages. It accurately classifies most of proliferative cases and cases with no DR. They showed that by employing a way more general dataset, their methodology made comparable leads to comparison to different (feature specific detection) strategies instead of CNN.

The network achieved this result with only 1 image yet. Although, the network does not face any issues whereas police work the image of a healthy eye, the article does not mention what happens once it uses image of associate degree unhealthy eye. The network additionally struggled to be told deep enough options of several the aspects of DR, that is ascertained from the sensitivity results of delicate and moderate categories. This was primarily thanks to the photographs within the dataset that wasn't cleaner and needed quite a substantial preprocessing step. However, overall the network did an honest job in classifying the photographs into the various classes.

In order to get a better accuracy on the network, there are some technique that can be applied. For example, Fusion Technique. There are several type of fusion technique that can improve the accuracy of the project.

The method that will be used is to combine the data from various sources. For example, to increase the video analysis, it requires visual and sound. For multimedia purpose, it will need textual and visual content to improve the quality. In addition, it requires textual and visual retrieval to increase the quality of some medical purpose. Or it might be combination of two or more same retrieval to get the better result. To do the fusion technique, there are three types of fusion methods

Early Fusion: Each feature from several approaches are directly combine into one feature that will be used as an input for classifier to get the result.

Late Fusion: Features from every channel are used in the classifier and obtain the score of classification. Then, all scores from the channel are concatenated to find final classification result. It might be different criteria such as mean, min, max, or weighted sum.

Intermediate Fusion: This fusion is not early fusion nor late fusion. The algorithm of this fusion should consist to concatenated kernel from every feature and fall into certain category

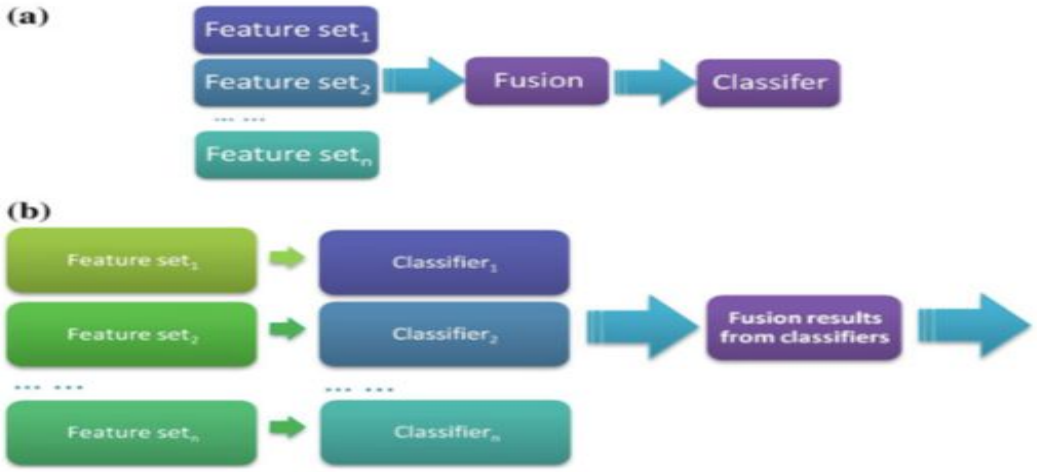


Fig. 1.1 A comparison between early (**a**) and late (**b**) fusion strategies

Every type of fusion technique has advantages and disadvantages. The advantage of early fusion is a simple and straightforward condition where just combining all the extracted features from several data to one vector representation. But it has large dimensionality also early fusion will hard to face when combining features from different type of data. Specially for intermediate fusion, it can present a good execution on several visual classification problems. But, intermediate fusion has high computation cost because of many parameters need to be learn. In late fusion strategies, it will deliver a compensation between computational efficiency with preservation of information. It also it will run faster than early fusion.

The other method that might increase the accuracy is doing fine tune. Fine tuning basically is replacing and retraining the classifier layer to the new dataset also will fine tune the weight of pre-train network by following the backpropagation. Principally all the layer that used is CNN can be fine-tuned. The purpose of fine tune is to get more generic feature that might be used in other layers. There are four scenarios to do fine tuning:

1. New dataset is small and like original dataset

It is not a good approach to fine tune the small data because of the overfitting issue. But because the data is look alike with the original data, its good idea use fine tune to get the exact features that needed.

2. New dataset is large and like original dataset

   Because of the data is big, it is good to have more assurance which will not over-fit when applying the fine tune to the whole network

3. New dataset is small and different with original dataset

   It is not a good approach to fine tune the small data because of the overfitting issue also might not be best to train the classifier that already has a lot of detailed features form the upper of the network

4. New dataset is large and different with original dataset

   Because of the amount of data, It will be beneficial to train the whole network from the scratch because the weights from pre-train model can be initialize also it possible to fine tune the whole network

In every project, it require the training and testing dataset to basically test the software is accurate or not. Cross validation is one of the approach to validate the whole software. Cross validation is better evaluation model compare to residual. The residual evaluation does not have the sign to indicate how good the testing phase did the job when it must make a prediction from new data that never be seen earlier. The solution of the problem of residual evaluation is to not use all data in training phase. By separating some data before the training phase, later the separated data can be used to test the functioning of the model to the new data. This method also known as cross validation. There are three types of cross validation.

1. Holdout method

The simplest method of cross validation. It will divide the dataset into two set, training and testing set. The function only goes to the training set for the first time. After that, the function go to testing set to get the output result.

Advantage: preferable to residual evaluation and not require any implementation

Disadvantage: The method can have a big variance. Also, the evaluation has more dependencies to training set compare the testing set.

2. K-fold cross validation

Improvement of holdout method. This evaluation will divide the data set into k set. Means that doing holdout method of k times. Every time when k set become testing set, k-1 set will be a training set. Later, the average of error can be computed across k tries.

Advantage: it does not matter how many times the data will be divided, all the data will be used exactly once because of the dividing part of testing and training set. So, it will decrease the variance as number of k increased.

Disadvantage: The algorithm will run k times. It will take more times to finish the evaluation.

3. Leave-one-out cross validation

Basically is K-fold cross validation with number of k is equal to N. Means that the function will train all the data except one point. The one point that is not trained will be used as prediction. This method looks like take more time than other cross validation method. But, computing leave one out cross validation will take less time even compare to residual evaluation.

3.0 Project Management Plan

3.1. Project overview

Diabetic Retinopathy is one of the biggest causes blindness. As longer patient has DR, chance of patient get blind also get higher. Classification Diabetic Retinopathy is required trained clinicians. The result from clinicians can be accurate but it highly time consuming and the clinicians cannot always be there as they are needed. The accuracy of classification DR also can be decreased by factor of human error.

This project propose to classify DR in real time also as accurate as possible. The DR class can be divided into three classes which no DR, mild DR, and sever DR. Using CNN architecture that one of the best method for image classification, this project should get a high accuracy and efficiency. A lot of successful project can lead the project to publishing in a respectable and prestigious research journal.

First step to start the whole project is pre-processing input retinal images. The input data is given from the supervisor. The retinal images already separated into three different classes which is no DR, mild DR and severe DR.There are 654 total input images that already divided into 3 different classes (153 images of no DR , 247 images of mild DR and 254 images of severe DR).All the data will resize to the same size and also will be re-colored to same color images. The purpose is to get the same details of the data which can lead the decrease the error of the whole project. After that, all data will be augmented by using rotating, flipping,etc. This data augmentation will increase the size of data to eight times from the original input images. After all the process, the total amount of data will 5232 images with in the same characteristic.

Next, the retinal input images will go through CNN architecture to extract some features that can indicate the DR stages. Feature that can be extracted is blood vessel,exudate,microaneurysms and hemorrhages. In addition, the fusion technique will combine two or more approaches to increase the accuracy of the project.
The whole code implementation will be written in MATLAB programming language. Some user interface also will be created as simple as possible,so it will be easier for user to use the software.

The project is started from week six of semester one and will be finished in the end of semester two. The whole project take around eighteen weeks for literature review, project proposal and also the code implementation. The code implementation in semester two will include preprocessing images, training data set, testing data set and the software will be finish after all process.

3.2. Project Scope

3.2.1. Project deliverables

The primary aim of this project is to deliver a simple software with a very simple architecture. Since, the project only consists of two people working on it and only one semester time (12 weeks) is given to implement it, the primary focus is to complete the main functionality as neatly as possible. The interface itself is also very simple which is explained in details in section 4.0 (External Design). With more time, resources and manpower implementing a more complex software would be possible, however, that is something that can be kept for future development. The topic on which this project is proposed has the potential to grow more in the future.

The project also requires a solid plan which has to be carried out over the next semester and results must be seen at the end of each of those activities. Completing, each of these small tasks are imminent for the success of the project. Following the timeline is essential. A gant in section 3.5 (Schedule) shows the plan of how the sub-activities will be carried out throughout the semester.

3.2.2. Product characteristics and requirements

The product is a software which needs to be run on a computer. An external GPU (hardware) is required for training the images. This is vital for the project because the

project deals with thousands of images and training a CNN architecture with them requires with high computational complexity which is very costly i.e requires high end processor and good GPU (expectedly GTX 1080 or higher).

Although, the project does not require any external costs, it does require information from published journal articles on the proposed subject. The project also requires software like Matlab (a good software for complex mathematical calculations mostly used by scientists) and some other softwares mostly for implementing the User Interface.

3.2.3. Product user acceptance criteria

The query image entered by the user has to be classified correctly. Our software deals with retinal images taken by fundus camera. So, if the query image is not an image of the retina then the program should not give an output and should state invalid data. If query image is a valid image of the retina of the eye then the program should classify it the image correctly into one of the three classes (No DR, Mild DR, Severe DR). However, there is a bit of pre-processing that has to be done on the query image before it can be classified because the architecture is basically taking in a new image and trying to detect the new things on it which it has learned when it was trained. So, the medium of the query image has to be converted to an image resolution or size through which the machine was trained.

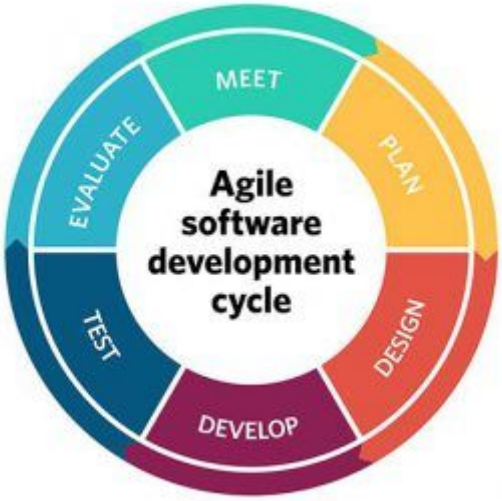
3.3. Project Organisation

3.3.1. Process Model

The whole project will use adaptive approach. There are several types of adaptive approach that can be used in the project. For this project, we will using agile methodology. It has some advantages that can taken from the methodology such as

- The client can be satisfied by rapid which is continuous delivery of useful software
- It is better to having people and interaction to emphasized than process and tools
- Meeting each other in face to face method is best method to communicate
- Always trying to get excellence and better design
- Changing circumstances is normal
- Late changes still can be apply

Means that this project try to find the best accuracy of DR classification. As time goes by, more knowledge we can get. By using this approach, we can easily change some circumstances to getting a better result.



3.3.2. Project Responsibilities

Task	Due	Responsibility
Pre-Processing Images	Week 3 Semester 2	Kevin and Nabil
Ensure the result	Week 3 Semester 2	Kevin
Training Dataset	Week 6 Semester 2	Kevin and Nabil
Ensure the result	Week 6 Semester 2	Nabil
Test Dataset	Week 9 Semester 2	Kevin and Nabil
Ensure the result	Week 9 Semester 2	Kevin and Nabil
Testing the whole software	Week 10 Semester 2	Kevin and Nabil
Documentation	Week 12 Semester 2	Kevin and Nabil

3.4. Management Process

3.4.1. Risk Management

All the project always having a risk that will be faced. The risk can have different probabilities and impact to the project. Because of that, there is some strategy to reduce the risk probability and plan if the risk happened. The first risk that might be face on this project is the project is not finish on time. This highly impact to the whole project and the probability of this risk is happened is not high. By having a good time management and work breakdown agreement, this risk can be solved. Next, the implementation is not working as well as predicted and not efficient as the objectives. The probability of these risk happened is quite high and highly impact the project. The strategies that can be used to reduce risk is do more testing for the implementation for more encouraging result also try to debug the code until it efficient and execute well. In addition, the result of classification of DR is not accurate. It should not be happened because most of the previous work result, CNN is one of the best method to classify images. If this risk come out in this project, the strategies that might be used is to try debug the implementation or try other CNN approaches.

Project Risk	Possibility	Effect	Strategy	Plan
Project is not finish	Average	Extreme	Have a good time management	-
Code implementation is not working well and not efficient	Average	High	Have a lot of testing and training data for get a better result	Debug the code until efficient and operate well
The result is not accurate	Low	Low	Use a reliable CNN approach	Try to use another CNN approaches

3.4.2. Monitoring and Controlling Mechanisms

3.4.2.1.Communication and reporting plan

This project is group project that consists two people which is Kevin Setiawan and Nabil Ahmed. We realize that this project cannot be done without a good teamwork. Based on the experienced from semester one, we met each other at least once a week to discuss about

the project. But if we have assignments, we usually met more than twice a week to discuss and do all the assignments. And after that we always double check the work of the others. So, we can know where the mistake and can improve our drawbacks.

Later, there is some report file that consist the plan about what should be done for each week from week one until the end of semester two. So, we can easily see that what we must do in this week and have a good time management in purpose to finish the project on time. Not only that, but also, we can control what we have done and can finish the project step by step that will lead us to having a better result in the project also.

The report will consist the objectives of every week and in the end that week, both of us will put some comments on the report. So, later if want to continue the project in the following week, we can see what we have done last week. So, it can be easily reviewed and continue proceed last week work. If there is some mistake, we straightforwardly identify it and fix it.

Type	Objective	Method	Frequency	Audience
Team Meeting	Review the project and planning	-Face to Face -Call	Weekly	Team Member
Technical Meeting	Implementation	Face to Face	Weekly	Team Member
Weekly Process Meeting	Discussion and report the status of the project	Face to Face	Weekly	Team Member
Monthly Process Meeting	Discussion and report the status of the project	Face to Face	Monthly	Team Member
Project Status	Discussion and	Face to Face	Every 2 weeks	Team Member

Meeting	report the status of the project			Supervisor

3.4.2.2.Review and audit mechanisms

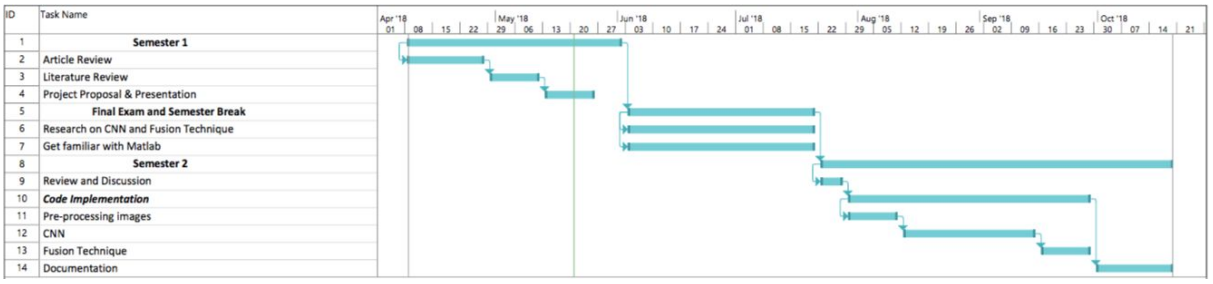
From the report that already created before, it easier to control the whole project. If there is something to update, it can be put on the report. Every work that has done during the project, will be put in the report.So, it will be clearly some works that already done and has not done. This will give a instruction to what to do next. The report also can be used as our simple documentation. All documentation also can be used as notes. The notes that create will be used to improve the work. So, every week the project will get better and better.

Quality of the project can be high from revise from the previous work. So, if there is something gone wrong, it can be identified easily by view from the report that already made and update it with the correct one. This might increase the quality of the project itself. Not only the project, but also, it can train us as group to increase quality of ourselves. We can learn from the mistake that made earlier to do better. So, we should not create a same mistake. It also will lead us to know where is our drawbacks. When each of us can give some advice to the other. After the report is made, it easily for us to review the project that we are done.

3.5. Schedule and Resource Requirements

3.5.1. Schedule

Task Name	Duration	Start	Finish
Semester 1	35 days	Mon 09/4/18	Fri 25/5/18
Article Review	15 days	Mon 09/4/18	Fri 27/4/18
Literature Review	10 days	Mon 30/4/18	Fri 11/5/18
Project Proposal and Presentation	10 days	Mon 14/5/18	Fri 25/5/18
Final Exam and Semester Break	40 days	Mon 28/5/18	Fri 20/7/18
Research on CNN and Fusion Technique	40 days	Mon 28/5/18	Fri 20/7/18
Get Familiar with Matlab	40 days	Mon 28/5/18	Fri 20/7/18
Semester 2	60 days	Mon 23/7/18	Fri 12/10/18
Review and Discussion	5 days	Mon 23/7/18	Fri 27/7/18
Code Implementation	40 days	Mon 30/7/18	Fri 21/9/18
Pre Processing Images	10 days	Mon 30/7/18	Fri 10/8/18
CNN	20 days	Mon 13/8/18	Fri 07/9/18
Fusion Technique	10 days	Mon 10/9/18	Fri 21/9/18
Documentation	15 days	Mon 24/9/18	Fri 12/10/18



As can be seen from Gantt Chart and Work Breakdown Structure, the project started from week six semester one of computer science project. It start with review article and create literature review to get more understanding of the topic. After literature review, we create a project proposal in the end of semester one. Next, all monash student having a final exam also semester break around seven to eight weeks. When we are having a semester break, we are planning to do more research on CNN and fusion technique. Not only the architecture, we also need to be familiar with matlab programming which all the code implementation will written in that language. This is very useful because when the semester start, we do not need spend more time to understand the language. So, we can straightforward implement the architecture. The code implementation will start from week two because at week one, we try to discuss and review what we have done in the semester break.Code implementation will include the pre-processing images, CNN implementation, fine tune also the fusion technique. When the software is done, we create the documentation for all the project that we already finish. The Work Breakdown Structure duration is slightly different with normal calculation. They already count five days for a week because Saturday and Sunday is not calculated. But in real live, we are still doing the project in that two off days.

3.5.2. Resource Requirements

Every project need several requirements to complete the whole thing. Not only the code implementation but also, some hardware is needed in this project. First, the project need big data of retinal images as an input. The images will be trained and tested to get to result of the classification. Because of the amount of data is big, it require high end GPU computer. This is for the efficiency when training and testing the data. Next, MATLAB also is one of the requirement. The whole code will implemented in MATLAB.

In addition, the project need convolutional neural network to classify retinal images into several class of DR. Example of CNN that can be used in this project such as AlexNet,GoogleNet or other network. After using CNN to classify input images, some fusion technique also will be needed in this project. Fusion technique is method to get a higher and better accuracy to classify DR from retinal fundus images.

4.0 External Design

4.1 User Interface(UI)

The User Interface is to be made as simple as possible. Our primary aim is to classify the level of DR in a given image. There will be a box for dragging, dropping or uploading images. The software will take that image and then let the client know that the image has been successfully uploaded. There is a button called “Classify” which will be clickable only if the user input image is valid. The software will also have a box for outputting the results or giving error messages.

If the image is not a valid image it will send an error message stating “Not a valid image” in the output.

If the user input image is a valid image it will make the Classify Button appear as clickable. The software will then classify the input image into any one of the three stages of DR (no DR, mild DR, severe DR) and will give it as the output.

Now, if the user wishes to classify another image, there is another button called “Upload New Image” which is also clickable. This button will only become clickable only if the output console is not clear and there is already an image that has already been classified. Clicking this button will clear everything out and the software would appear the way it appears when we first start it.

The user can repeat these steps to classify any number of images he wishes. If there are a lot of images that he wishes to classify at any given time all together then that is something that can be kept as a future work or motivation for improving the software.

Finally, there are two more buttons on the top right-hand-side. One is the minimize button “-” and another is the quit button “X”.

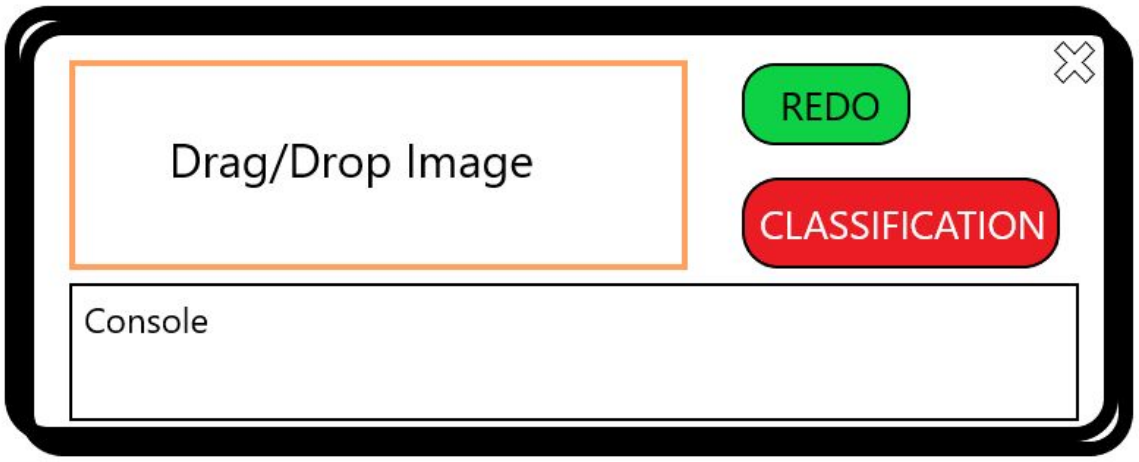


Figure: A sample user interface. It does not match completely with the description given above as it was designed by (Nabil Ahmed) roughly to give an idea of what to be expected from the software.

4.2 Dataset

The dataset has a total of 654 images. The data set contains three classes of images.

1. No DR
2. Mild DR
3. Severe DR

The dataset is provided by our supervisor (Monash University Malaysia Lecturer -DR. Mohammad Reza Zare). Not all the images from the dataset will be used for the training phase. Some of it will be kept for the testing phase. One of the reason for this is acquiring retinal images is quite challenging because it requires specialized fundus cameras that consist of an intricate microscope attached to a flash enabled camera and it also requires experienced clinicians to classify manually the level of DR of the patient by observing the images. Therefore, just to be on the safe side few images would be kept for testing. Since, we already know the result of the test image we can correctly test out the software for all the test cases. This is vital because there is no point in making a software for which we cannot compare its correctness. For example, our dataset is 4.87 Gigabytes(GB). So, Maybe we can use 4.5 GB of images for training and the rest we can keep for testing. However, this parameter is experimental and may change during the implementation phase in semester two.

4.3 Performance

CNN network architecture consists of many layers. For example, AlexNet has five convolutional layer, max pooling layers, dropout layers and three fully connected layers. Google Nets design uses completely different origin modules to create twenty-two layers. This means the algorithm will implement all the layers before the result of the DR classification can be computed. Compared to other techniques computing CNN requires more time. But, CNN architecture performed better compared to other architecture in image classification. This means time complexity of the CNN architecture is higher compared to others but it is also more efficient for classification of DR. So, we can expect to have a high time complexity (high run time) while testing the algorithm.

In addition, preprocessing the images will also take time because it has to train most of the input images (approximately 4.5 GB) in size.

4.4 Software Architecture

Matlab is platform that is specially designed for engineers and scientist to do programming. It is a matrix based language that allow the maximum natural expression of mathematics. There are some advantages and disadvantages by using this platform. Some advantages from matlab is complex math operation already built in inside the program because it is matrix based language. Next, can easily adding two list in one step, instead using a looping. Also, the functionally is significantly expanded by using toolboxes. On the other hand, matlab also has disadvantages such as it will use large amount of memory that can slowing down the performance of the computer. Also, matlab sits on top of Windows means that it highly depends on the CPU processors.

5.0 Methodology

5.1 Training

5.1.1 Pre-processing

The dataset consists of 654 images in total(153 images of no DR, 247 images of mild DR and 254 images of severe DR). The resolution of the images varies a lot. The lighting condition and the orientation differs as well. Since, there is no fixed resolution the images has to be normalized so that CNN architecture can train the data efficiently for giving better accuracy and results. An architecture can only perform better if there are a lot of high quality and relevant images being used to train it in the first place. The better the training the better the learning of the architecture.

The training images are scaled down to a fixed resolution of 256 x 256 pixels. 512 x 512 pixels is not desirable because the computational complexity would be too high and training a network with high-end images requires a lot of time. This is something that can be kept as a future work when the GPU prices would drop and more robust CNN architectures are introduced.

The contrast of the images are then enhanced evenly across pixels. This is done by using a technique call Histogram Equalization. The images are then normalized by using min-max normalization. The min-max normalization prevents the CNN from learning essential background noise in the image. [Doshi, Shenoy et al. 2016]

5.1.2 Data augmentation

The training set is augmented with random 0-360° rotation. Random horizontal and vertical flips and random horizontal and vertical shifts. The training set is increased by about 8 times [Sun, Wan et al. 2017].

Main methodology:

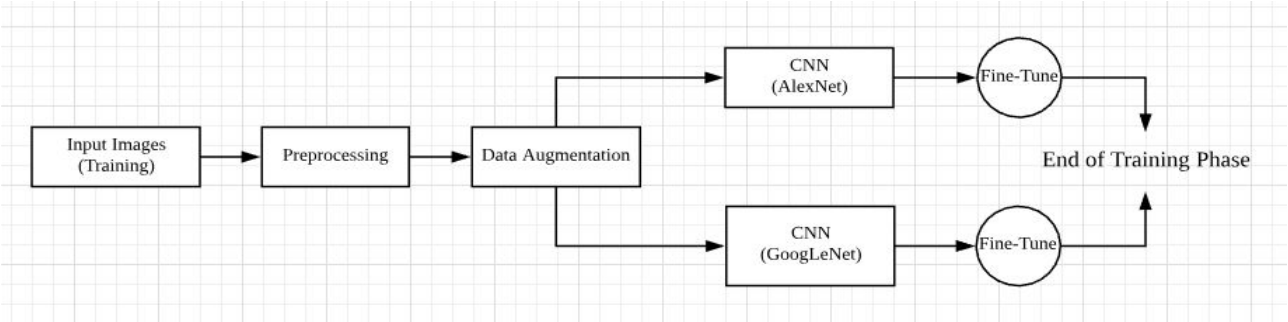


Figure: The above figure shows the training phase of the proposed model

First, the training images are preprocessed accordingly and then the preprocessed images are then augmented for better learning of the CNN architecture. The two CNN architecture chosen are the GoogleNet & AlexNet. This is because codes for these two architectures are available for their prominence in the field of machine learning and they have a good history when it comes to classification.

The training images are then fine tuned. Fine tuning is basically done by tweaking the layers (fully connected layers) at the end of the CNN architecture. Fine tuning helps to focus on the specific features that would like to extract based on our requirements. For our proposed model we are interested in detecting haemorrhages, exudate, blood-vessel and micro-aneurysms. This is the end of the training phase.

5.2 Testing

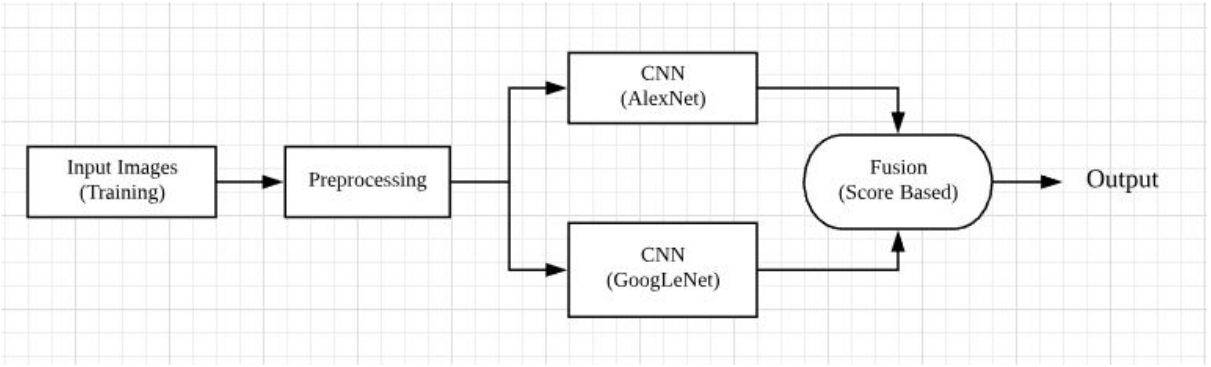


Figure: The above figure shows the testing phase of the proposed model

The testing phase is almost identical to the training phase. It only skips two steps(data augmentation, fine tuning). It has one additional step as well which is the score based fusion. This step is crucial because it finally classifies the query image based on the prediction from the pre trained CNN models.

6.0 Test Planning

6.1 Test Coverage

The whole algorithm must be tested to check if it extract the correct features or not. Because of DR classification required some extracted features such as microaneurysms, blood vessels or hemorrhages. All the amount of features will be calculated to classify DR into three different classes(no DR,mild DR and severe DR). Later, the output also need to be check either it the input images is reliable to be classified. If the quality of the input is really bad, it will affect the accuracy of DR classification. Next, the classification result also need to tested to ensure the DR classes. Because every DR class has different treatment to cure or prevent the blindness.

6.2 Test Methods

The first test method case is by checking that the algorithm extract some features that needed in DR classification. Also, how accurate the DR classification to separate it into three different classes. The accuracy, sensitivity and specificity also must be tested for the trusting rate. The higher accuracy result percentage, the software also become more and more reliable.

6.3 Sample Test Cases

The data that will be used is retinal images that already divided into three different classes. There more than 600 input images thats used in this project. Some images will be used for training stages and the other will be used for testing images. After the whole algorithm is implemented, the testing images can be an input for the software.

The output from this software is three DR classification which is no DR,mild DR and severe DR. Also the software can identify if the input images is invalid(is not a retinal images or very poor quality picture)

7.0 References

1. *Acharya, U., Lim, C., Ng, E., Chee, C., Tamura, T.. Computer-based detection of diabetes retinopathy stages using digital fundus images.*
   P I Mech Eng H 2009;223(5):545–553.
2. Andonová, M., et al. (2017). Diabetic retinopathy screening based on CNN. 2017 International Symposium ELMAR.
3. Tennakoon, R.B., Mahapatra, D., Roy, P., Sedai, S., & Garnavi, R. (2017). Image Quality Classification for DR Screening Using Convolutional Neural Networks.
4. García, G., et al. (2017). Detection of Diabetic Retinopathy Based on a Convolutional Neural Network Using Retinal Fundus Images. Artificial Neural Networks and Machine Learning – ICANN 2017, Cham, Springer International Publishing.
5. Doshi, D., et al. (2016). Diabetic retinopathy detection using deep convolutional neural networks. 2016 International Conference on Computing, Analytics and Security Trends (CAST).
6. Ghosh, R., et al. (2017). Automatic detection and classification of diabetic retinopathy stages using CNN. 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN).
7. Li, X., et al. (2017). Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI).
8. Pratt, H., et al. (2016). "Convolutional Neural Networks for Diabetic Retinopathy." Procedia Computer Science **90**: 200-205.
9. Sun, J., et al. (2017). Retinal Image Quality Classification Using Fine-Tuned CNN. Fetal, Infant and Ophthalmic Medical Image Analysis, Cham, Springer International Publishing.
10. Tajbakhsh, N., et al. (2016). "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?" IEEE Transactions on Medical Imaging **35**(5): 1299-1312.

11. Verma, K., et al. (2011). Detection and classification of diabetic retinopathy using retinal images. 2011 Annual IEEE India Conference.

12. Zhou, L., et al. (2018). "Deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images." IET Image Processing **12**(4): 563-571.