PATTERN RECOGNITION OF HEART VALVE IN ECHOCARDIOGRAM USING CONVOLUTIONAL NEURAL NETWORK

MUHAMMAD HANIF BIN AHMAD NIZAR

PROJECT SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE MASTER DEGREE

INSTITUTE OF GRADUATE STUDIES
UNIVERSITY OF MALAYA
KUALA LUMPUR

2017

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Muhammad Hanif bin Ahmad Nizar (I.C No:)

Registration/Matric No: KQB160004

Name of Degree: Masters of Biomedical Engineering

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Pattern recognition of heart valve in echocardiogram using convolutional neural network

Field of Study: Biomedical Engineering

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 21/8/2017

MUHAMMAD HANIF BIN AHMAD NIZAR

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:
Designation:

ACKNOWLEDGEMENT

I praise to the almighty God for giving me the strength and patience to complete this project. This accomplishment could not have been possible without the support I receive from my family especially my mother. I would also like to express my gratitude to my supervisor, Ir. Dr. Lai Khin Wee for his guidance, advice, and enormous support in this endeavour. This project's accomplishment will not be possible without the help of my friends Yong Ching Wai and Dilkumar Marimuthu. Last but not least, I would like to express my gratitude to University Malaya and the National Heart Institute for their support in this project.

ABSTRACT

One of the causes of heart failure is valvular heart disease which can be diagnosed using echocardiogram. However, using this machine requires a degree of skill in order to locate the heart valves. We propose the use of a trained neural network to locate the position of the aortic valve from an echocardiogram image as an assistive technology during echocardiogram examination. The neural network AlexNet was used in this study which was trained using a deep learning platform, NVIDIA DIGITS. 58 of patients' echocardiogram were used to train the AlexNet which were obtained from the National Heart Institute. After training the AlexNet, it was tested with 25 images and the resulted images were validated with a sonographer. Testing the AlexNet within the deep learning platform showed it was able to achieve an accuracy of 99.87% for aortic valve and 99.69% for background images. The qualitative comparison from the sonographer with the resulted image was that the trained neural network was able to localize the image accurately but may not be able to segment the valve precisely. This study was able to demonstrate the possibility of utilizing neural network to develop assistive technology for medical devices such as echocardiogram. Recommendations to increase the performance of the neural network such as using a neural network with more layers and providing a larger dataset for training were also explained.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF GRAPHS

PAGE

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

In 2014, there were 52,650 people that had died due to heart disease in Malaysia (World Health Organization, 2014). Valvular heart disease is one of the factors that may lead to heart failure. A diseased valve may be altered causing narrowing of the valve or inability of the valve to prevent blood backflow during closure of the valve (Ladich, Yahagi, Romero, & Virmani, 2016). Most of the cases of are due to aging factor leading to degeneration of the heart valves (d'Arcy et al., 2016; Nkomo et al., 2006). Other factors such as congenital abnormalities and rheumatic heart disease may also lead to the failure of the heart valves (Boudoulas, Vavuranakis, & Wooley, 1994). Out of the 4 heart valves, the aortic valve is the most prevalent of being affected by diseases as the valve is subjected to a higher pressure than other valves (Maganti, Rigolin, Sarano, & Bonow, 2010).

One of the method to diagnose valvular heart disease is by using an echocardiogram which uses ultrasound waves to visualize the heart and its valves (Rodés-Cabau, Taramasso, & O'Gara, 2016). Magnetic resonance imaging (MRI) and computed tomography (CT) scan may also be used but echocardiogram is preferred as it is faster to obtain data and less harmful effect to the patients. However, the diagnosis of heart disease depends largely on locating the defect which may be difficult due to the limitation of the echocardiogram and the skills of the sonographer (Matsumoto et al., 1976; Stevenson, Kawabori, Dooley, & Guntheroth, 1978).

Convolutional neural network (CNN) is a part of artificial intelligence that uses deep learning and is specialize in visual imagery analysis (Karpathy et al., 2014). The mathematical model of the neural network is based on how the neurons of the brain functions (Matsugu, Mori, Mitari, & Kaneda, 2003). Due to its ability to recognize patterns in visual imagery, the neural network had been used in various applications such as facial recognition, sentence classifications and object recognition (Briske-Anderson, Finley, & Newman, 1997; Goodfellow, Bulatov, Ibarz, Arnoud, & Shet, 2013; Kim, 2014; Lawrence, Giles, Tsoi, & Back, 1997). Thus, it is possible for CNN to assist in identifying structures of the heart especially heart valves.

What methods that can be used to enhance or assist the sonographer in identifying the heart anatomy during echocardiogram examination? The solution may lie on utilising CNN as it can be trained to identify heart structures form an echocardiogram images. A reliable assistive technology can potentially decrease the time required for an echocardiogram examination. The assistive technology may also decrease the learning time required for beginner sonographers or doctors to develop their skill in echocardiography. Eventually, this will decrease the workload burden in medical field. Thus, we aim to train a CNN to be able to identify and localize the heart valve.

1.2 OBJECTIVES

The aim of this study is to train a CNN to localize the aortic valve. The objectives are:

1. To obtain patient's echocardiogram at parasternal short axis view for training the neural network.
2. To train CNN using deep learning platform such as NVIDIA DIGITS.
3. To validate the performance of the neural network to localize heart valve.

1.3 PROJECT OUTLINE

Including the current introductory, this paper consists of six chapters. The contents are described in the following:

Chapter 2 describes a literature review of topics that is related to the current research. This review includes the history of machine learning, the architecture of CNN, AlexNet and the current trend of CNN in the medical field.

Chapter 3 explains the methods used to obtain images to be used to train the neural network. In this chapter, it also explains how the CNN architecture, AlexNet was trained and utilized to analyse test images of echocardiogram.

Chapter 4 contains the result of the trained AlexNet performance within the deep learning platform. The chapter also includes the qualitative comparison of the sonographer against the resulted image of the trained AlexNet.

Chapter 5 includes the discussion regarding the data obtained and how it can be further improved for future studies.

The final chapter 6 describe the conclusion that is made regarding this study.

CHAPTER 2

LITERATURE REVIEW

2.1 MACHINE LEARNING

Researchers had often pursued the field of artificial intelligence to allow the creation of intelligent machines that reacts and function like a human being (Russell & Norvig, 2003). A machine with artificial intelligence should be able to have a thought process or behaviour that is considered to be rational. Several testing methods were developed to prove the capabilities of artificial intelligence. One of the famous test was the Turing Test, which was developed by Alan Turing in 1950 (Turing, 1948). It was a test for the machine to give a response that is indistinguishable from another intelligent being such as a human being. This test involves a human interrogator posing a written question to both human and the machine. The interrogator will be isolated from both the machine and human subject. Turing test requires the computer to have natural language processing, knowledge representation ability, automatic reasoning and machine learning. The machine passes the test if the human interrogator could not differentiate the written response coming from a human or a machine. The Turing test was further developed by Steven Harnad to include computer vision and robotic limbs which became the Total Turing Test (Harnad, 1991). This new test allows the machine to be able to perceive physical objects and interact with the objects using its robotic limbs.

Machine learning was developed in order to allow machines the ability to learn from any given data (Munoz, 2014). Machine learning involves using algorithms and models to be able to recognize patterns and make prediction on given data (Kohavi & Provost, 1998). There are several categories in machine learning which depends on the nature of the data and the type of response or feedback that is desired from the machine. These categories are supervised learning, reinforcement learning and unsupervised learning (Russell & Norvig, 2003). Supervised learning is when the machine is required to learn the general rule of the inputs and outputs from the human expert. This is in contrast to unsupervised learning where there will be no human expert and the machine is required to achieve its goal by developing its own rule. From these categories, there are many machine learning models that had been developed over the years. Examples of machine learning models are decision tree learning, association rule learning, artificial neural network (ANN) and inductive logic programming.

ANN is a mathematical model that mimics the biological nervous system that consists of interconnecting neurons. The neurons of ANN are connected by weighted links and is able to propagate signals from one neuron to another. Figure 2.1 shows a simple representation of an ANN model that consists of 3 layers. In 1943, Warren McCulloch and Walter Pitts were the first to propose ANN and demonstrated its capability to learn from the data(McCulloch & Pitts, 1943). This was further improved by Frank Rosenblatt which proved the perceptron convergence theorem. Frank proposed that the connection strength of the perceptron can be adjusted using a learning algorithm (Rosenblatt, 1958). This lead to a more efficient learning method for the neural network. Further researches had led to maturation of the neural network by introducing backpropagation learning algorithm and reinforcement learning (Barto, Sutton, & Anderson, 1983; Rumelhart, Hinton, & McClelland, 1986).

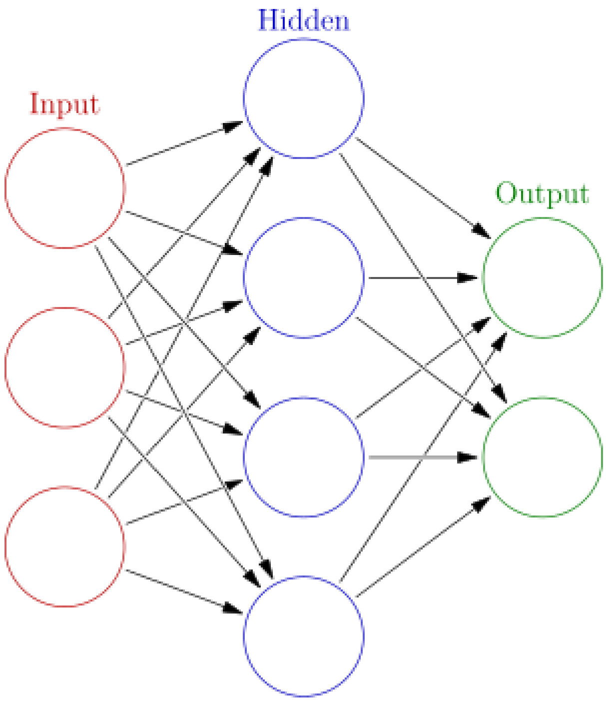


Figure 2.1: Simple ANN model with neurons arranged in layer. There are 3 layers: input, hidden and output. Source from https://en.wikipedia.org/wiki/Artificial_neural_network

Another branch of machine learning is Deep Learning which is essentially any neural network that consists of more than 1 hidden layer. The additional hidden layers allow the neural network to identify progressively more complex feature of a given data (Bengio, 2009). However, different scenarios pose different complexity and problems involving data pattern and recognition. Thus, researchers develop various Deep Neural Network (DNN) architectures to solve the issue. Examples of these architectures are Autoencoders, Deep Belief Networks, Boltzmann Machines and Recurrent Neural Networks (Ackley, Hinton, & Sejnowski, 1985; Hinton, 2009; Li, Luong, & Jurafsky, 2015; Mikolov, Karafiát, Burget, Cernocký, & Khudanpur, 2010).

2.2 CONVOLUTIONAL NEURAL NETWORK

One of the most popular DNN architecture is the Convolutional Neural Network (CNN). As shown in Figure 2.2, a CNN model is made of specialized layers which consists of convolutional layer, pooling layer and fully connected layer. This neural network is suitable to be used for image classification and require less pre-processing compared to other algorithms (LeCun, 2015). CNN was popularised in 2012 due to the development AlexNet (Krizhevsky, Sutskever, & Hinton, 2012). AlexNet was able to reduce its classification error from 26% to 15%, which earns it the winner in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). This had sparked researchers to further pursue the development of CNN models which lead to a large increase of participants in ILSVRC 2013.

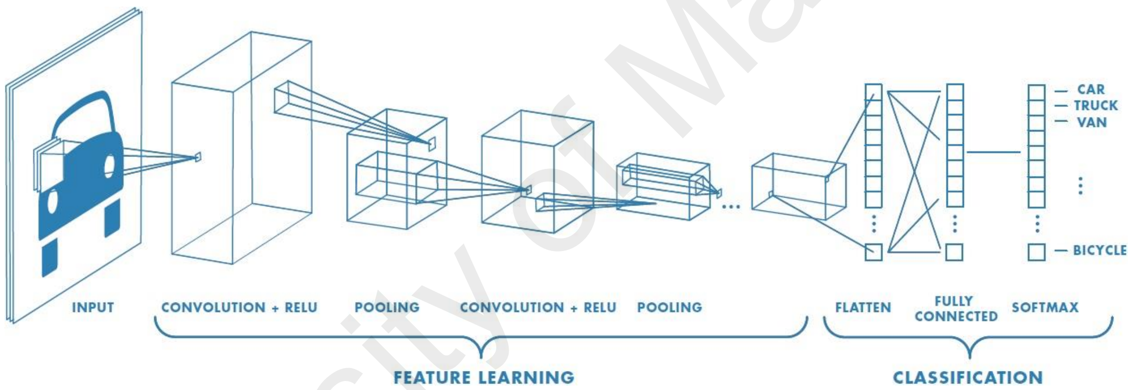


Figure 2.2: An example of CNN model containing multiple layers of neuron. Source from https://www.mathworks.com/discovery/convolutional-neural-network.html

Digital images are composed of array of pixels with specific values which are determined by the resolution and size of the image (Goodman, 2016). The data within the digital image will first be analysed by the convolutional layer, the first layer of CNN (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016). The convolutional layer consists of a filter and a feature map. The filters consist of array of numbers representing the weight and has receptive field to perceive data from previous layer. Feature map is the output from the filter that was applied to the previous layer. The method of how convolutional layer operate can be demonstrated using an example of an image with 32x32 pixel and receptive field of 5x5. Each filter can only obtain input from its own receptive field at a time. As the filter slides or convolving around the digital image with 1 stride length, it will result in a feature map of 28x28 output values. This procedure is shown in Figure 2.3. A convolutional layer represents a certain feature of interest in the image and sends output according to whether the feature is present or absent. Increasing the number of filters in a convolutional layer also increases the number of features that can be classified.

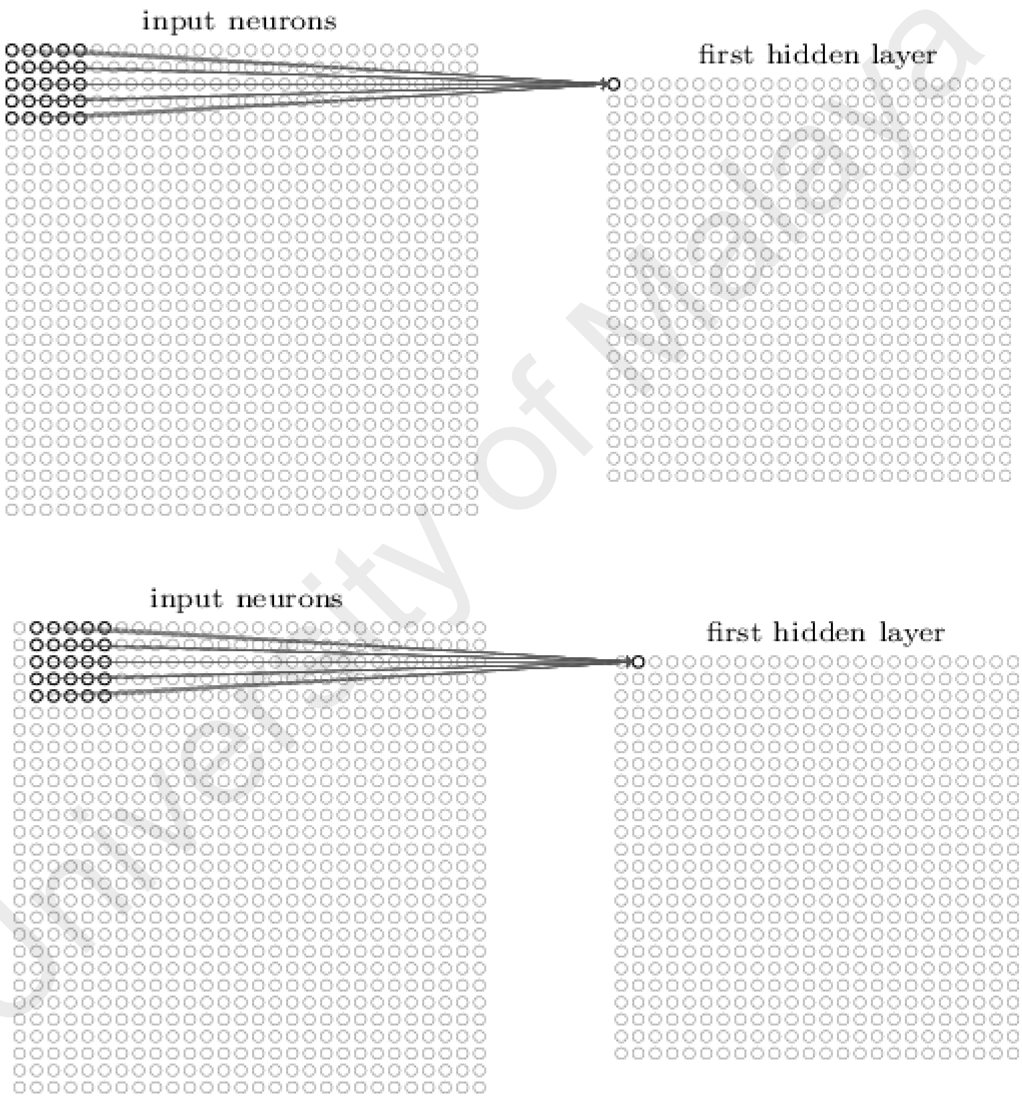


Figure 2.3: Example of how convolutional layer obtain data. The receptive field is sliding with 1 stride length at a time. Source from http://neuralnetworksanddeeplearning.com/chap6.html

The output from the convolutional layer will then be send to the pooling layer. Pooling layer will simplify the information received from the convolutional layer by compressing the data (Szegedy et al., 2016). The layer also has receptive field which are considerably smaller than convolutional layer. There are various methods of pooling such as max-pooling and L2 pooling.

The final layer in CNN is the fully connected layer. The layer connects all the neuron from the pooling layer to all its output neuron. In CNN, this layer is used to make the final combination of features that is present on the image data and for making predictions.

2.3 ALEXNET

The performance of CNN depends largely on the size of its network levels and the number of neurons at each level (Szegedy et al., 2015). However, there are drawbacks when increasing the size of the network. Firstly, larger neural network increases the number of parameters. This in turn may result in overfitting of data especially if the training set is limited. Another drawback is the increase in computational performance in order to use the neural network.

For this study, AlexNet is the preferred choice of CNN model as it performs well and will not require a high-end graphic processing unit from the computer. This network is made of 5 convolutional layers, max-pooling layers, 3 fully connected layers and dropout layers (Krizhevsky et al., 2012). The architecture of the network allows it to be trained with two graphic cards simultaneously if required. Training an AlexNet can be simplified by using a deep learning platform such as NVIDIA DIGITS. In order to use DIGITS more effectively, a Linux operating system is required.

Other notable CNN models are ZF Net, VGG Net and GoogLeNet (Simonyan & Zisserman, 2014; Szegedy et al., 2015; Zeiler & Fergus, 2014). ZF Net was developed in 2013 by Matthew Zeiler and Rob Fergus. The model is an improvement to AlexNet with better optimization within its neural network. VGG Net appeared a year later and contains 19 layers which is considerably more than previous models. In 2015, GoogLeNet had further improved on the technology by increasing the layers to 22. Furthermore, GoogLeNet introduces the Inception module which enables it to have significantly less parameters than AlexNet (Zeiler & Fergus, 2014). Thus, GoogLeNet is able to achieve a higher accuracy while avoiding overfitting of data. However, AlexNet is chosen in this study as it is simpler and requires less computing power than other models. Additionally, AlexNet will still be able to achieve a good accuracy.

2.4 NVIDIA DIGITS

In order to perform deep learning, various frameworks are available to develop the neural network. Examples of the most common frameworks are Caffe, Torch and Tensorflow. With the introduction of NVIDIA DIGITS, deep neural network can be trained more rapidly. It is a free webapp program that requires the computer to have graphic card with CUDA cores which is only available with NVIDIA graphic cards. The program supports all the common framework and also contains well established pre-trained models such as AlexNet, GoogleLeNet, and VGG. Currently, NVIDIA DIGITS works best in a Linux-based operating system (OS). It is still possible to install DIGITS in a Microsoft Window OS but certain features will not be available such as DetectNet. Thus, Linux-based machine will be used in this study to avoid any complications during training.

2.5 CNN IN MEDICAL FIELD

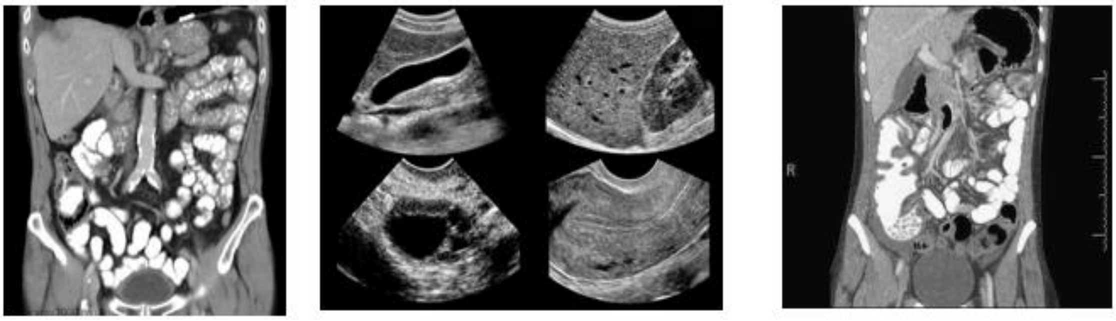
Various researcher had developed neural network to perform image segmentation or classification on medical images. This includes magnetic resonance image (MRI) of the brain, MRI of breast lesion and segmentation of cancerous cell from microscopic specimens (Chen, Zhou, & Wong, 2006; Meinel, Stolpen, Berbaum, Fajardo, & Reinhardt, 2007; Prasoon et al., 2013; Shen, Sandham, Granat, & Sterr, 2005). These studies showed that with neural network, we may be able to further improve diagnostic capabilities of diseases and help doctors in patient's management.

A study conducted at Stanford University compared the performance of AlexNet, GoogleLeNet and a shallow CNN model in classifying breast masses as benign or malignant (Lévy & Jain, 2016). The study used 1820 mammogram images containing breast mass from 997 patients in mediolateral oblique and craniocaudal views. GoogleLeNet (92.9% accuracy) was able to best both AlexNet (89.0% accuracy) and the shallow CNN (60.4% accuracy) by a significant margin. Furthermore, the GoogleLeNet was able to achieve a better recall (0.934) when compared to radiologist which achieve recalls between 0.745 and 0.923. Achieving a high recall is important as false negative patients (remain undiagnosed) is potentially more problematic compared to false positive patients (undergo breast biopsy). This further showed the promising future in medical diagnosis when using neural network.

In this study, we are aiming to assist the sonographer or doctors in localizing the aortic valve. Aortic valve was chosen for this study as it is more clinically significant compared to the other heart valves. The aortic valve is located at the left side of the heart and is subjected to most of blood pressure. As patients get older, this valve has a high tendency to develop calcification causing restriction in the blood flow (B. F. Stewart et al., 1997).

Using echocardiogram or any other ultrasound machine require skill and practice to be able to correctly identify any abnormalities during the scanning. Studies had shown that sonographer have to undergo extensive training and encounter many cases in order to become competent (Hertzberg et al., 2000; W. J. Stewart et al., 1999). This is also more crucial during emergency cases where quick diagnosis based on the ultrasound finding is needed. Thus, it may be beneficial to have a subsystem such as a neural network to assist in the scanning process.

Compared to other imaging modalities such as MRI and CT scan, echocardiogram image is considered more challenging to train for the neural network. One of the reason is because the quality of the image for echocardiogram is less detailed and contains more noise. Figure 2.4 shows a side-by-side comparison of image quality between ultrasound, CT scan and MRI scan of the abdomen. Despite the difference in image quality, ultrasound is still essential in the medical field as it does not expose patient to radiations and can provide diagnosis faster compared to the other image modalities. Studies had also proved that there may not be significant difference in reaching a diagnosis for the patient when using ultrasound against CT scan or MRI scan (Dueholm & Lundorf, 2007; Martínez, Ripollés, Paredes, Blanc, & Martí-Bonmatí, 2009). However, CT scan and MRI will still be used to accurately visualize the internal organs and to get better understanding of the patient’s disease. Another challenge in this study is echocardiogram data source are in video format compared to others which are in still images. The heart constantly moves and the valves are in either closed or open state from the video. Thus, there is considerably more variations within the data which may cause difficulty in CNN training.



MRI image | Ultrasound image | CT scan image

Figure 2.4: A side-by-side comparison of image quality of ultrasound, MRI and CT scan at the abdomen. Source from http://www.ultrasound-direct.com/uploads/images/0x0/Women-Upper-Abdominal-Pelvic-Scan.jpg, http://www.paediatricgastroenterologist.co.uk/application/files/4614/3453/9587/mri-abdomen11.jpg https://openi.nlm.nih.gov/imgs/512/345/3151403/PMC3151403_13181_2011_144_Fig2_HTML.png

CHAPTER 3

METHODOLOGY

3.1 IMAGE ACQUISITION

The echocardiograph images of the aortic valve were obtained from the National Heart Institute database with consent. 58 patients were chosen randomly from the database for training the neural network. Another set of echocardiograms from 5 patients were reserved for the trained CNN to localize the aortic valve and evaluated by a sonographer. To obtain a good view of the aortic valve, an echocardiogram image of the short axis view at the level of the aortic valve was obtained. Figure 3.1 shows an example of echocardiogram image at the short axis view. The machine used for the echocardiography is the Philips iE33. The mode has been set to 2D mode. The frequency of the probe was set to 17 Hz and depth was set to 14 cm.

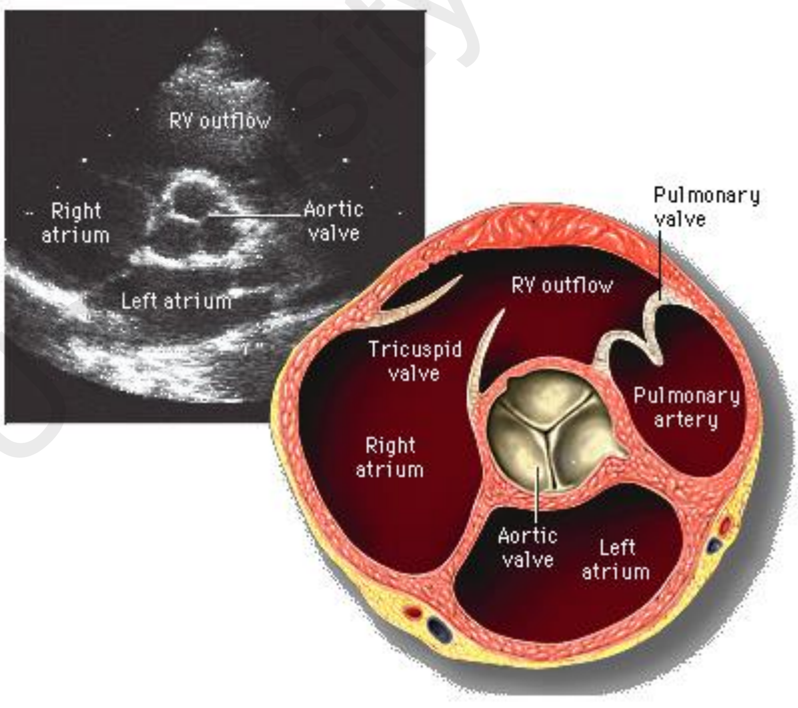


Figure 3.1: Comparison between an illustrated figure and a real echocardiogram at the short axis parasternal view at aortic level. Source from http://www.med.yale.edu/intmed/cardio/echo_atlas/contents/index.html

3.2 CNN TRAINING

The echocardiogram footages obtained were converted to Joint Photographic Experts Group (JPEG) files which resulted images ranging from 100 to 200 frames depending on the patients. Each frame had been resized to 351x351 pixels to include only the echocardiogram image and not the setting details of the machine. 50 to 60 frames were taken from each patient where the aortic valve was cropped out of the image and replaced with black background. The aortic valve images were categorized as “AorticValve”. From the 50 images, 10 were chosen and was sub cropped to 25 images for each image. These are put under the category of “Background”. To avoid overfitting and generalization of data of CNN, the aortic valve images were transformed to various configurations. The aortic images were subjected to horizontal flip, vertical flip, rotation to left and to the right. The workflow is shown in Figure 3.2 and an example of the work process is shown in Figure 3.3. This resulted in 14,980 aortic valve images and 16,686 background images. 25 images from the reserved 5 patients will be used for testing with the trained CNN model.

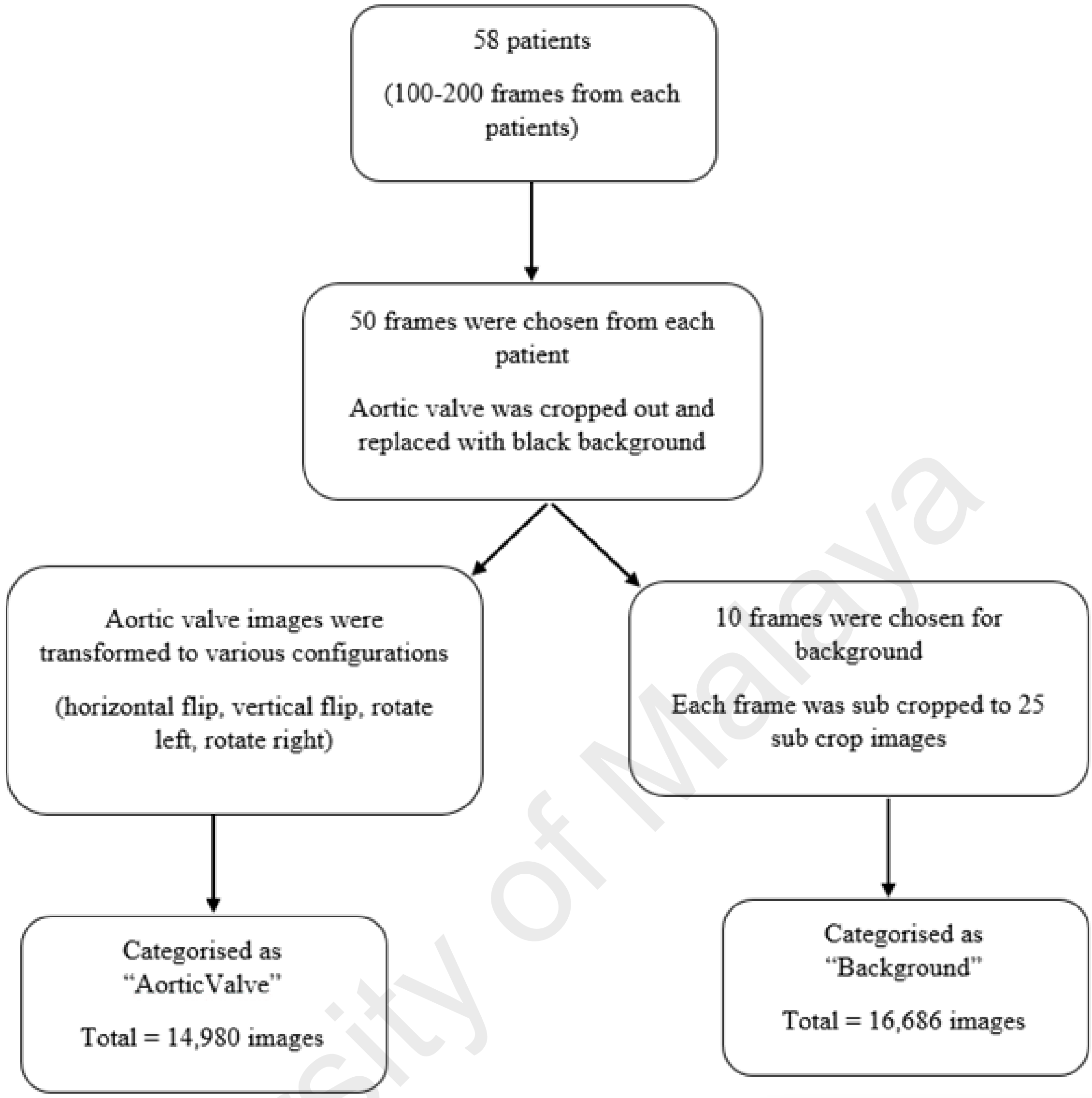


Figure 3.2: Workflow for data acquisition and categorisation

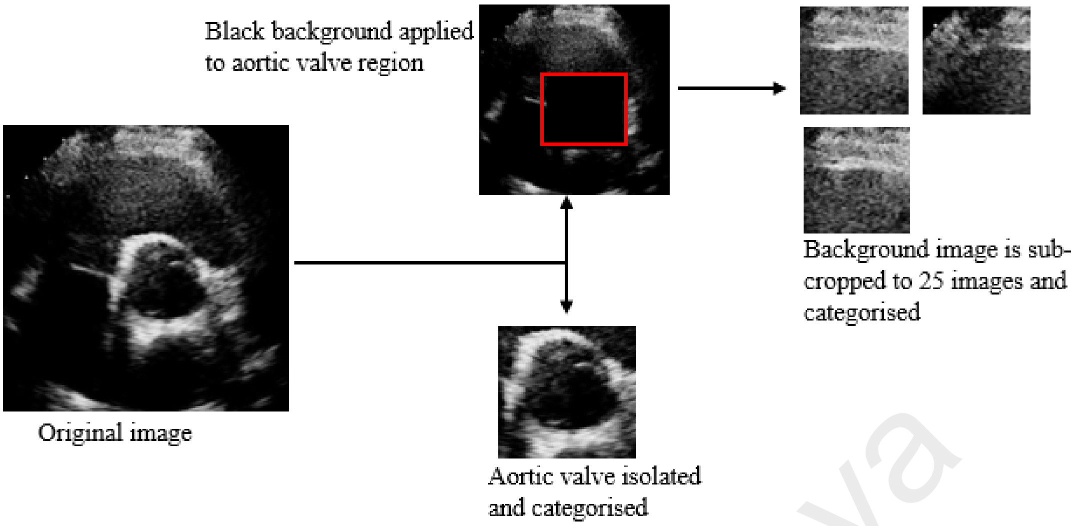


Figure 3.3: Example of the work process. The red box is to highlight the area where the black background was patched to the image. The red box is not present during the actual work process.

The computer used for CNN training had a processor of Intel Core i7-3720QM and Nvidia GTX660m which had 384 CUDA cores, clocked at 835MHz and 2GB of memory. The graphic card specifications are important factor in determining the duration of the CNN training and how many data can be processed. The NVIDIA DIGITS deep learning platform was used in this study to train the AlexNet model. For the parameters, the dataset input was set to 100x100 pixels with grayscale. Grayscale was chosen as the ultrasound images are in black and white. Choosing grayscale will also decrease the time of training as there is less data to process compared to a coloured image. 100x100 pixels were chosen as bigger sized pixels will require more time to process and may not achieve a significant improvement on the results. The deep learning platform will automatically resize the images without losing the overall data within the image. The learning rate for the model was 0.01 and the epoch was set to 30. From the dataset, 25% will be used for validation purposes and another 25% for testing within the deep learning platform.

3.3 ANALYSIS OF DATA USING ALEXNET

In order for the trained AlexNet to localize the aortic valve from echocardiogram images, various technique including pixelated probability map (PPM), multi-level Otsu Threshold and corrected region was used. This can be done using coding and Microsoft Visual Studio Community 2017 installed with OpenCV 3.2.0. The flowchart of the code is shown in Figure 3.4.

A similar sliding kernel window technique from convolutional neural network was implemented in this study. The PPM requires frequency of probability and cumulative probability; represented in 2D as Counter Map (CM) and Cumulative Probability Map (CPM) respectively. An element-wise division on CPM with CM will result in PPM as seen in Equation 3.1.

$$Prob_{(j,i)} = CumProb_{(j,i)} \odot \frac{1}{Count_{(j,i)}}$$

Equation 3.1: *j* and *i* represent the rows and columns of the map while $\odot$ represents the element wise division operation.

In order to isolate part of the images with the highest probability of containing aortic valve, multi-level Otsu Threshold was performed (Otsu, 1979). It is a method to reduce the grey level and convert it to a binary image. Afterwards, a technique to remove small connected region was applied to reduce noise in the resulted image. After all the 25 images from the reserved patients had been analysed by the trained AlexNet, the resulted images were shown to a qualified sonographer from IJN for comparison.

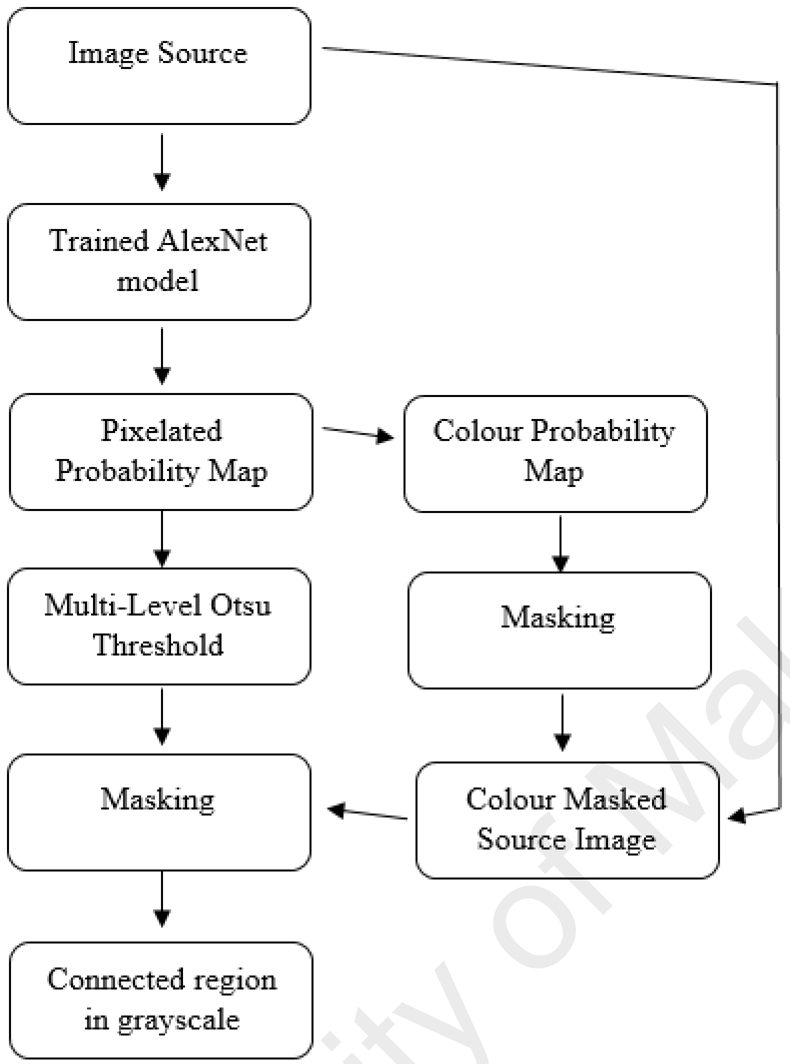


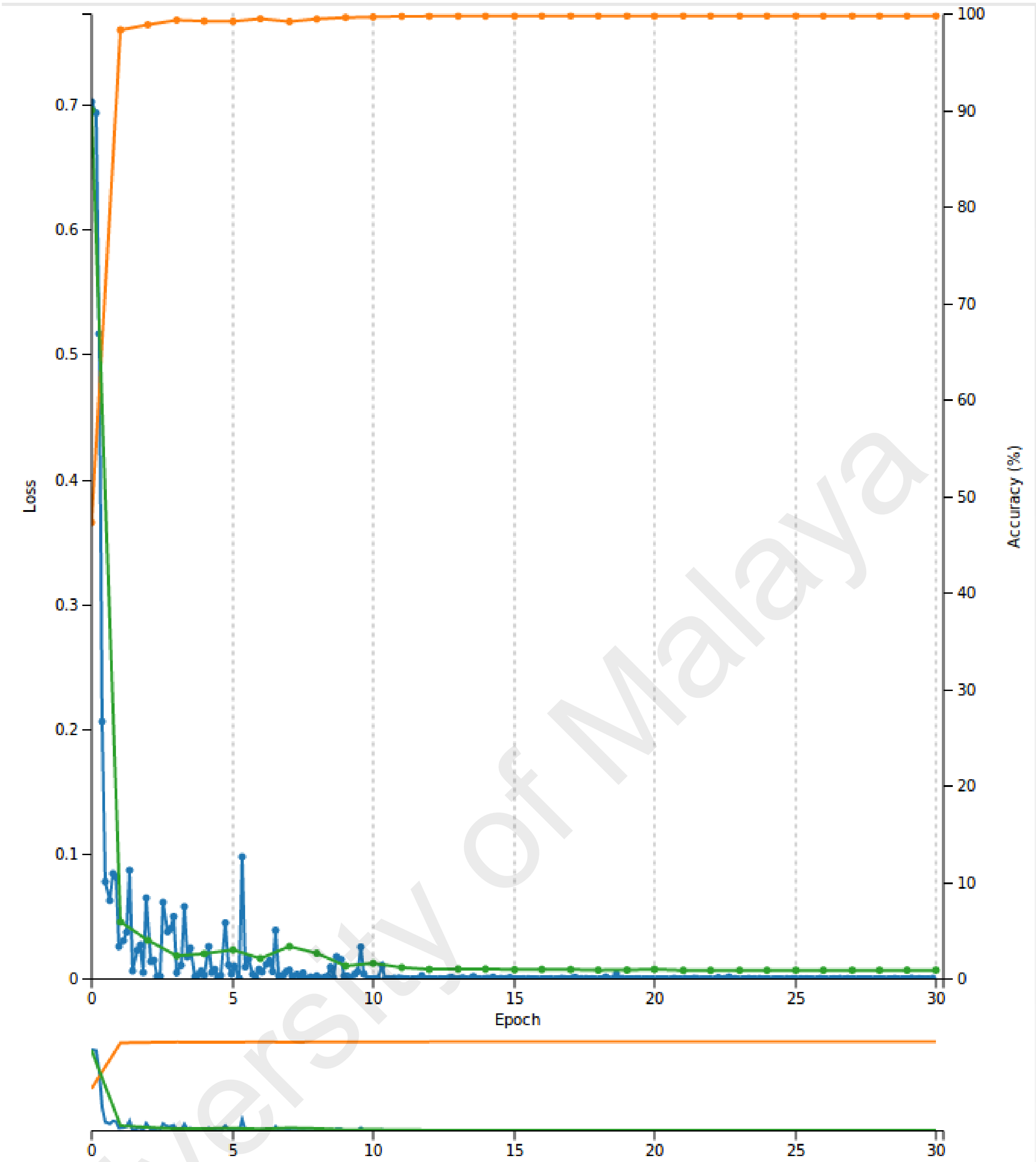
Figure 3.4: Coding flowchart

CHAPTER 4

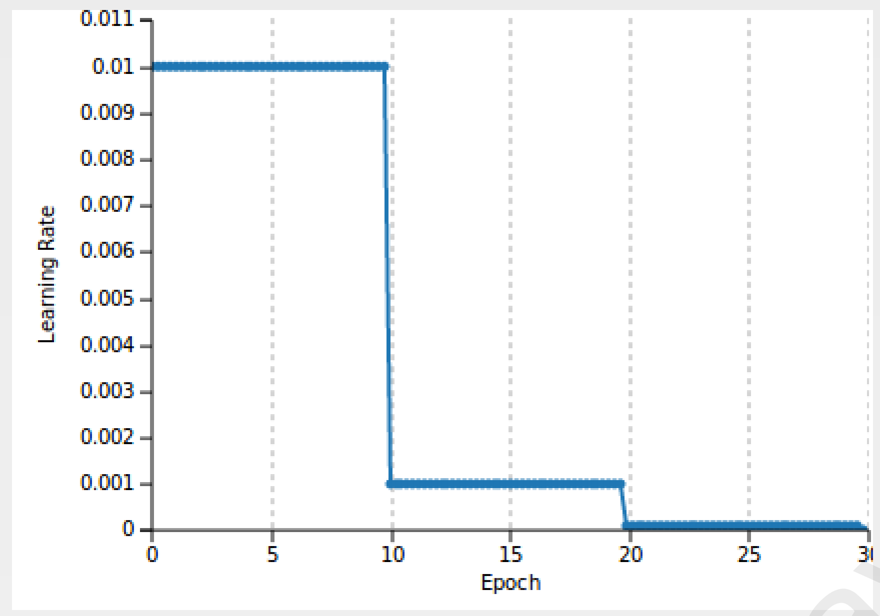
RESULTS

4.1 TRAINED ALEXNET MODEL

The training of the CNN model takes 1 hour to complete. It was able to achieve an accuracy of 99.81%. However, it should be noted that the accuracy started to reach its baseline of 99% starting at epoch 10 as seen in Graph 4.1. The learning rate had also decreased dramatically at epoch 10 and remain constant at 0.0001 after epoch 20. The learning rate is shown in Graph 4.2.



Graph 4.1: Graph showing a real-time performance of the CNN model during training. Orange line represents accuracy; green line represents loss(val); blue line represents loss(train).



Graph 4.2: Graph the learning rate of the CNN model at each epoch.

Performance of the trained AlexNet model can be measured directly from NVIDIA DIGITS itself. 25% of the dataset was previously set for testing which consists of 4,171 background images and 3,745 aortic valve images. The result of confusion matrix is shown in Table 4.1. The CNN model was able to achieve an accuracy of 99.87% for aortic valve and 99.69% for background images.

	Aortic Valve (Predicted)	Background (Predicted)	Per-Class Accuracy
Aortic valve (Actual)	3740	5	99.87%
Background (Actual)	13	4158	99.69%

Table 4.1: Confusion Matrix of the trained CNN model

4.2 IMAGE LOCALIZATION

Localization of the aortic valve using the trained neural network was done using OpenCV. Figure 4.1 shows the different stages of how the image is being analysed. As seen in the figure, the PPM helps in categorising the location that has a high probability of containing aortic valve. The Otsu Threshold helps in filtering out areas with low probability. Small connected region was then successfully removed. The analysis was applied to all 25 images and the final result of the image was examined by the sonographer.

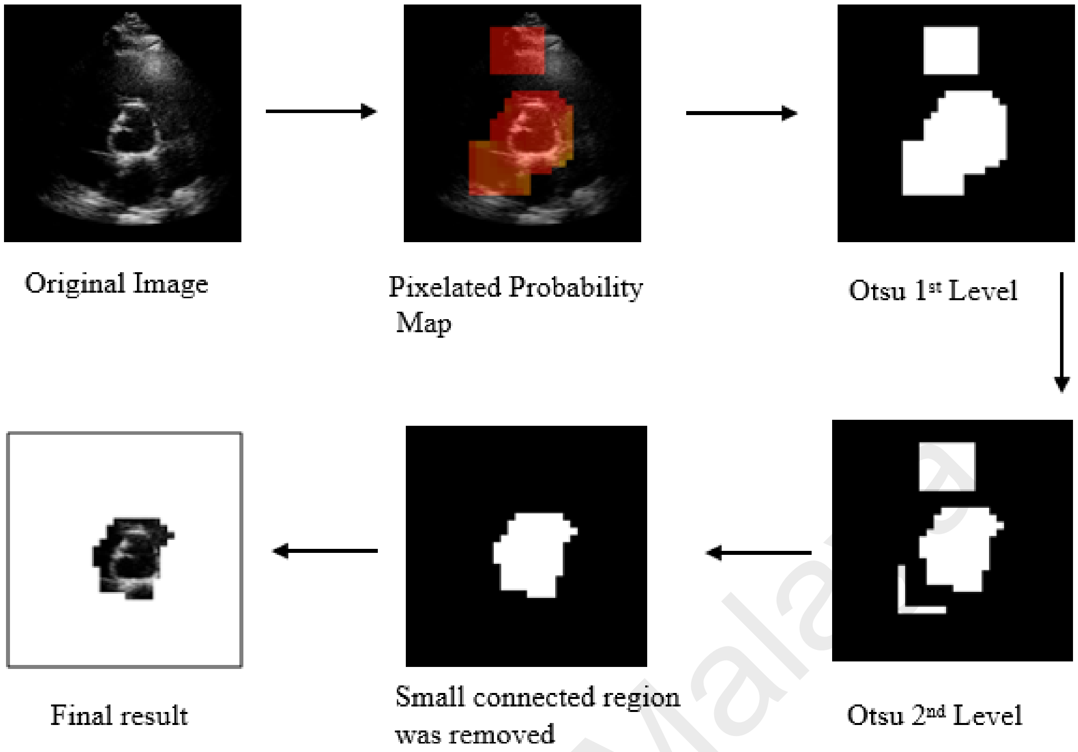


Figure 4.1: This shows the process of image analysis and localisation of the aortic valve

In the evaluation process, the sonographer classified the image whether it is precise and accurate. An accurate result is considered if the final image contains the presence of aortic valve. If the final image was able to properly segment the heart valve and not contain any other structure it is considered precise. The result of the examination is shown in Table 4.2. Localization images of the aortic valve that was performed by the neural network is shown starting from Figure 4.2 to Figure 4.6. Figure 4.7 and 4.8 shows the calculations in the first convolutional layer and the fully connected layer when analysing an image. Based on the sonographer, the neural network was able to localize the aortic valve for all except 1 image (96%). The image affected was image 17. 32% of the 25 images contains only aortic valve. The rest of the images have additional structure other than aortic valve or may not contain anything at all. The results showed that the trained CNN was able to accurately locate the heart valve but may have difficulty in segmenting the actual valve itself.

Image number	Accuracy	Precision
1	Yes	Yes
2	Yes	No
3	Yes	No
4	Yes	Yes
5	Yes	No
6	Yes	Yes
7	Yes	No
8	Yes	No
9	Yes	Yes
10	Yes	No
11	Yes	Yes
12	Yes	No
13	Yes	No
14	Yes	Yes
15	Yes	Yes
16	Yes	No
17	No	No
18	Yes	No
19	Yes	No
20	Yes	No
21	Yes	Yes
22	Yes	No
23	Yes	No
24	Yes	No
25	Yes	No
Total (Percentage)	96%	32%

Table 4.2: Results of the examination done by the sonographer on the final images.

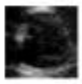
Image number	Original image	Resulted image
1		
2		
3		
4		
5		

Figure 4.2: Shows the localization of the aortic valve by the trained CNN for image 1 to 5.

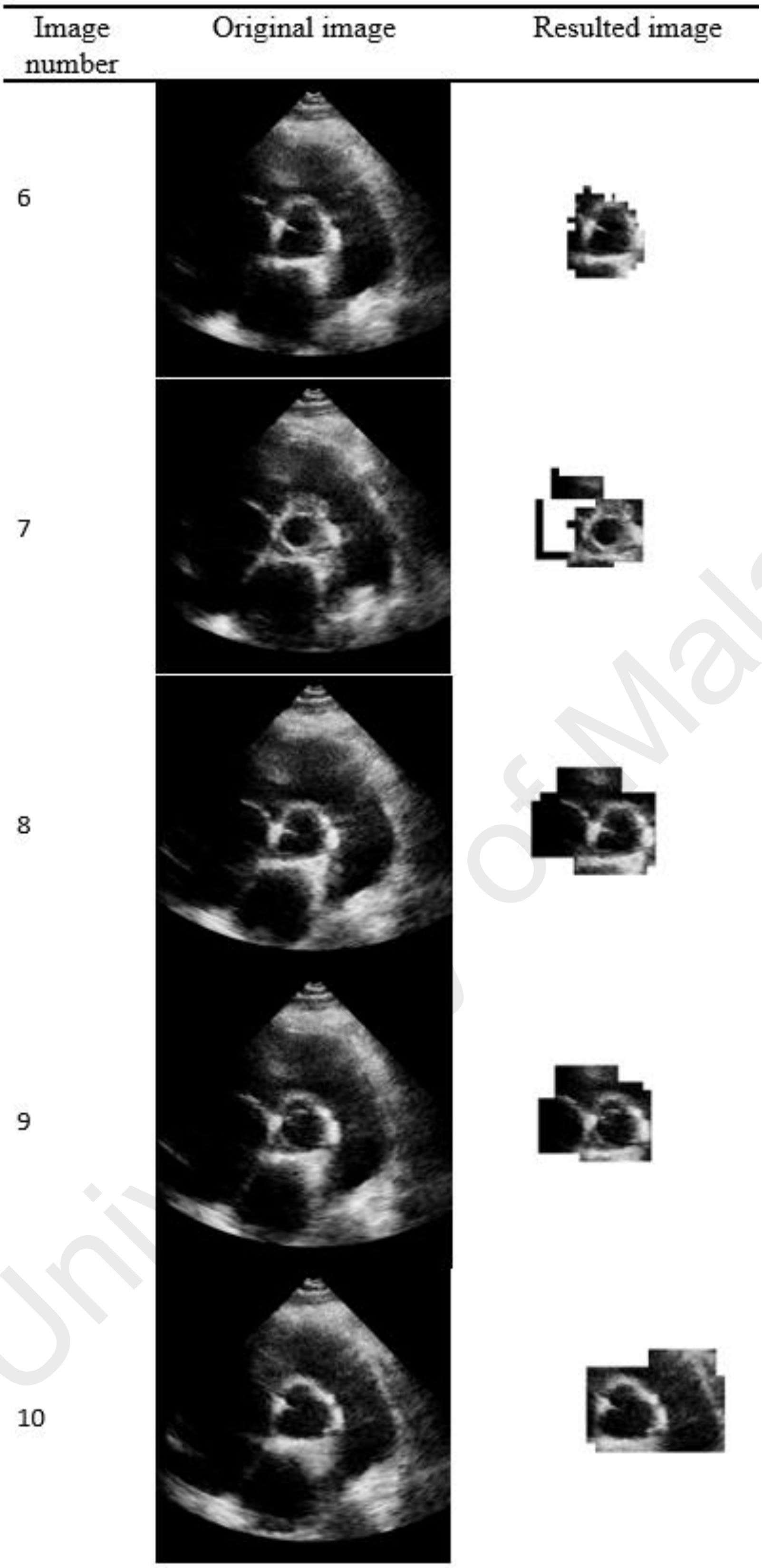


Figure 4.3: Shows the localization of the aortic valve by the trained CNN for image 6 to 10.

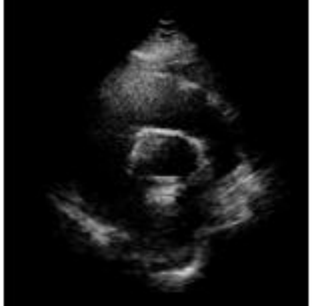
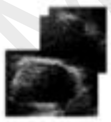
Image number	Original image	Resulted image
11		
12		
13		
14		
15		

Figure 4.4: Shows the localization of the aortic valve by the trained CNN for image 11 to 15.

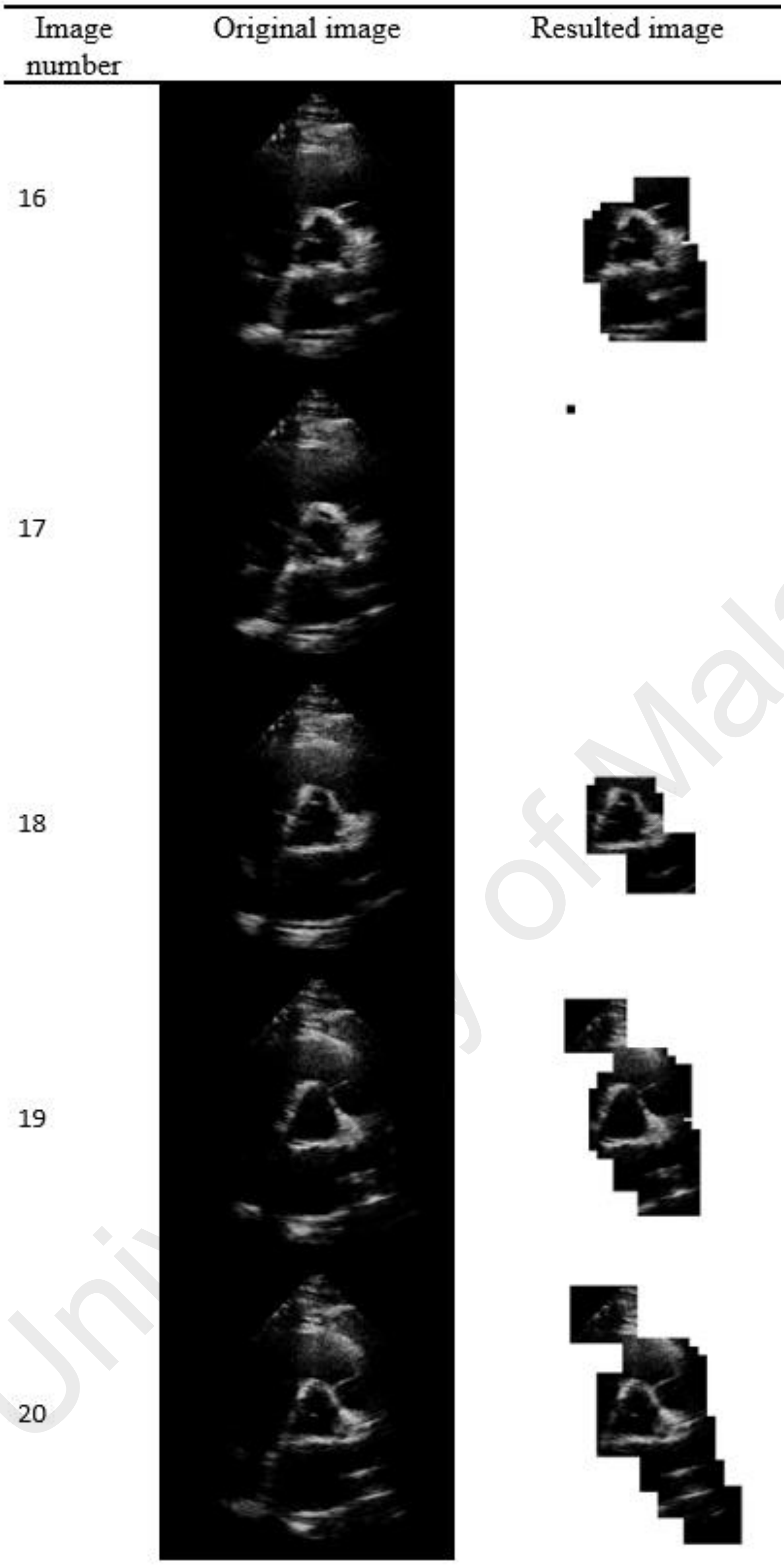


Figure 4.5: Shows the localization of the aortic valve by the trained CNN for image 16 to 20.

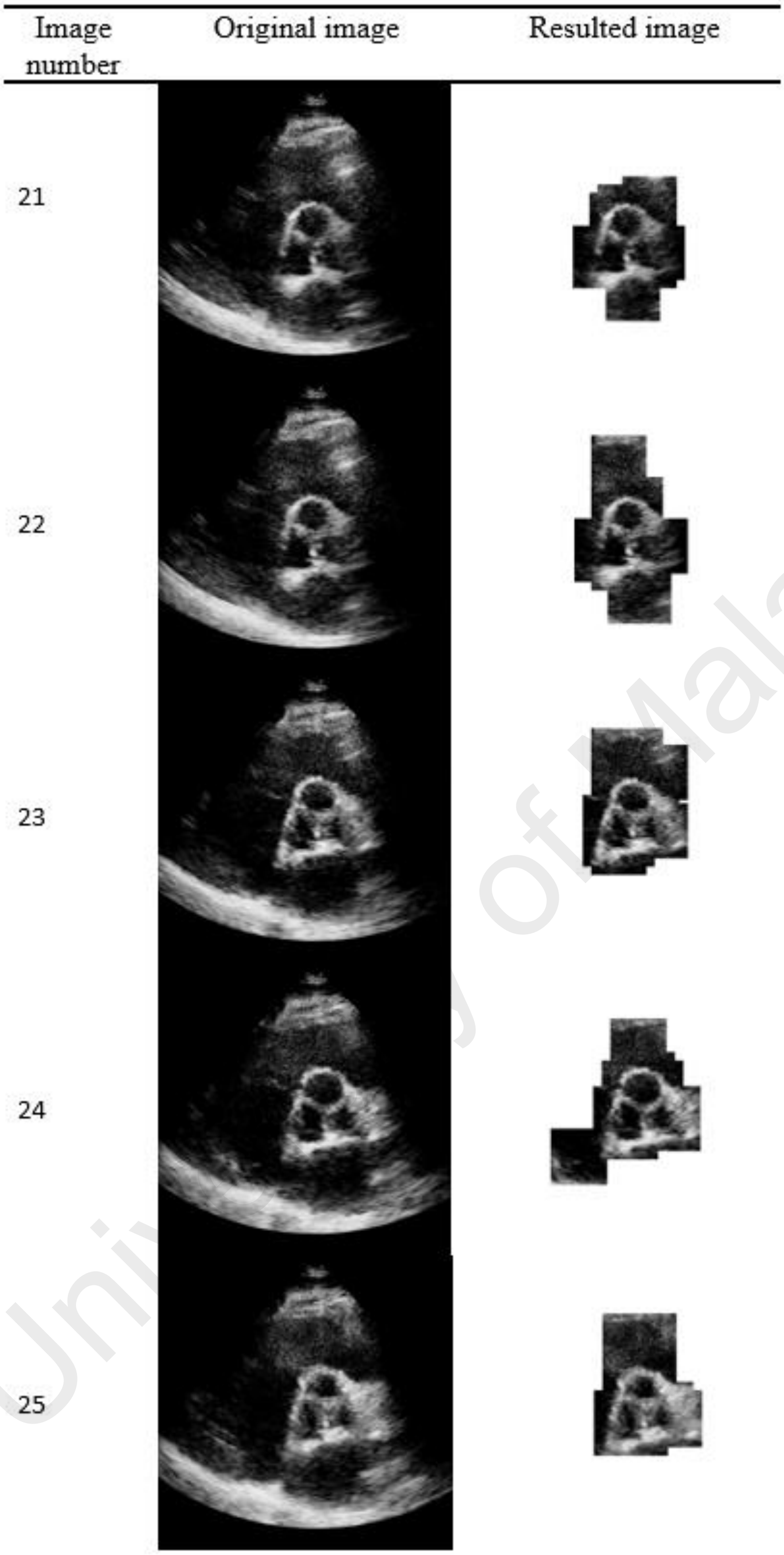


Figure 4.6: Shows the localization of the aortic valve by the trained CNN for image 21 to 25.

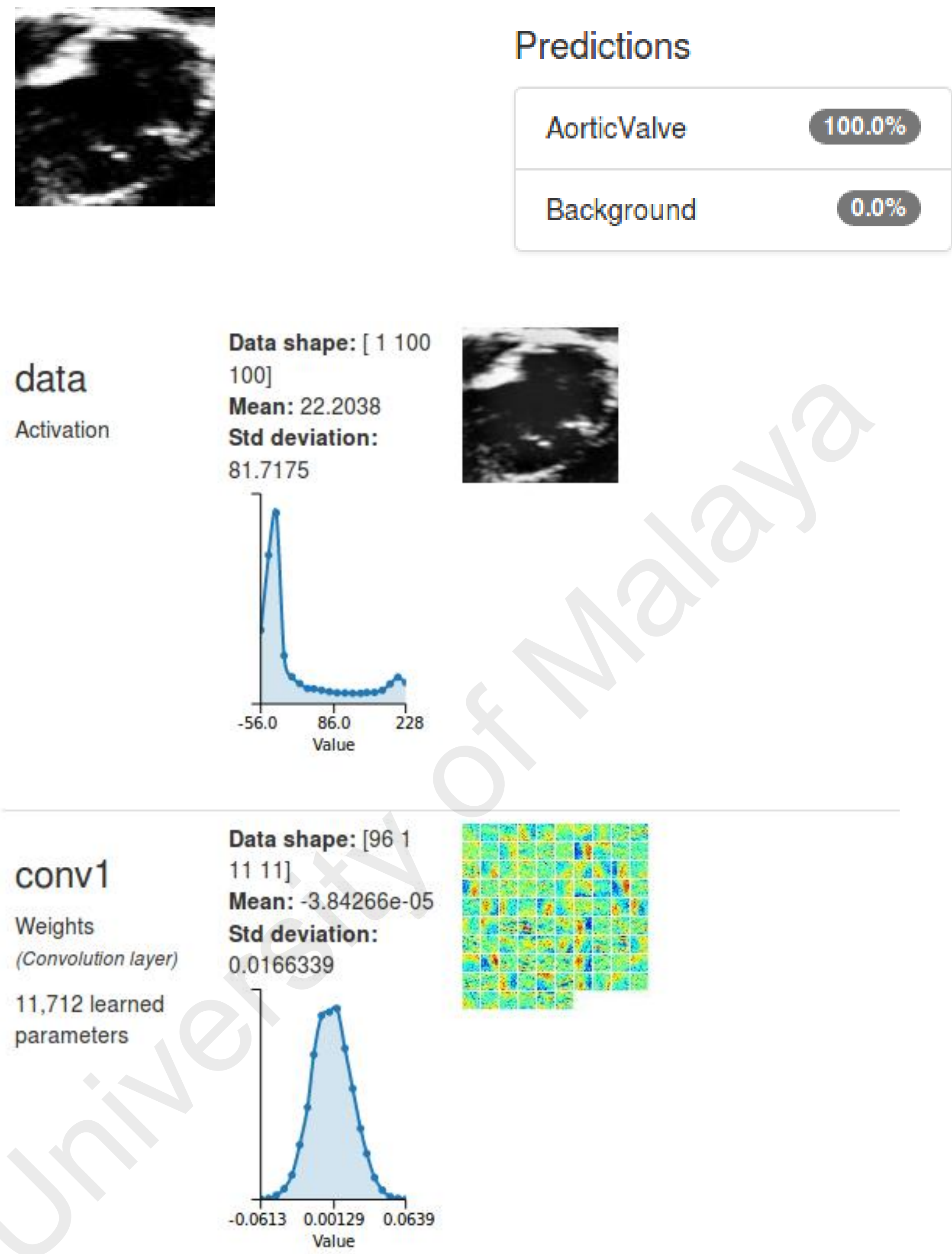


Figure 4.7: Example of how CNN model predicts the test image. The first convolutional layer is shown here.

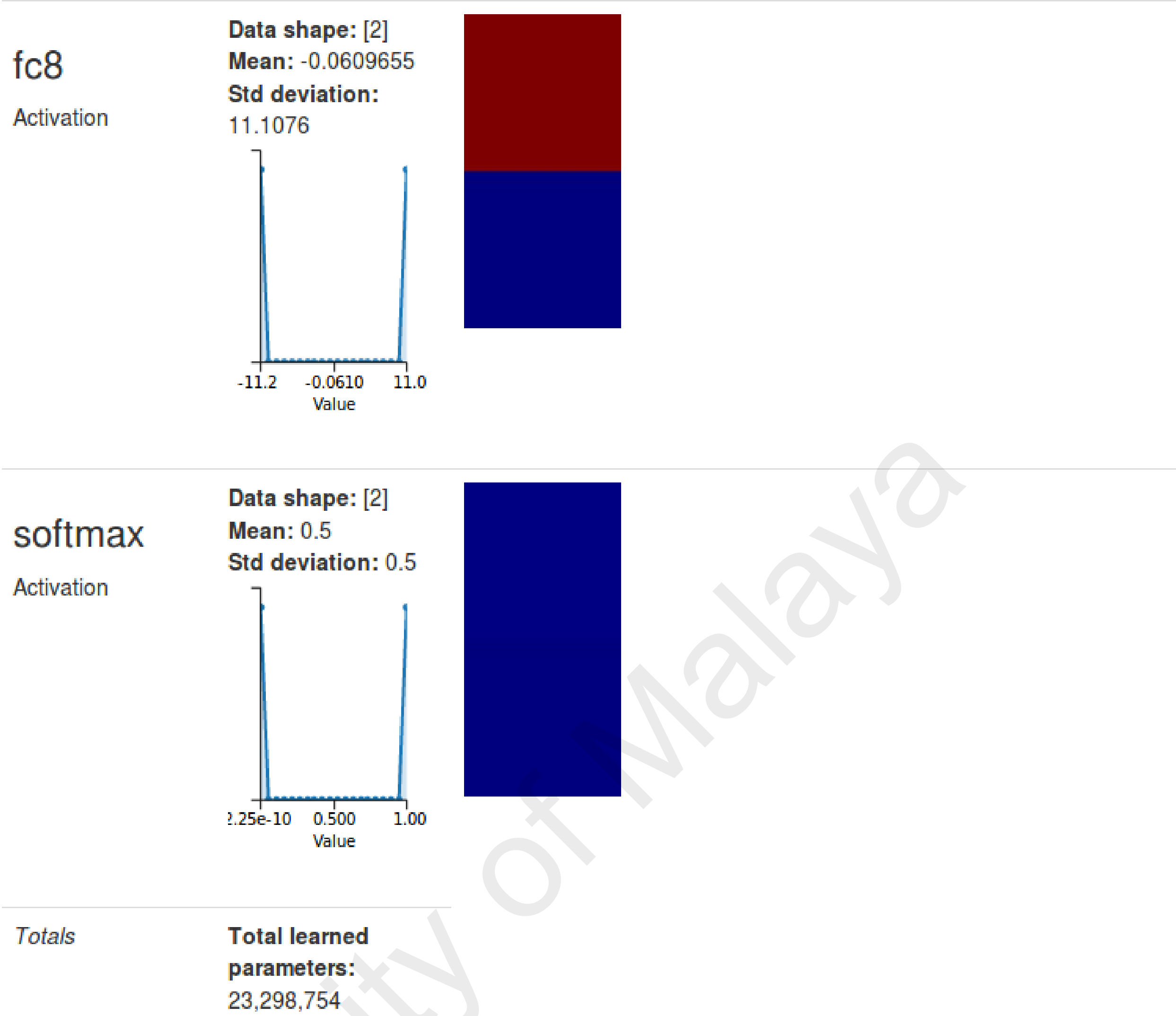


Figure 4.8: Example of how CNN model predicts the test image. Fully connected layer is shown here.

CHAPTER 5

DISCUSSION

For the most part, the trained CNN model managed to locate the location of the aortic valve. It should be noted that ultrasound images are less detailed and contains more noise than other scanning devices such as MRI or CT scan. Certain part of the echocardiogram images had features that is almost similar to a heart valve. The situation can be seen in image test 16 to 20 where the surrounding of the aortic valves had parts that look similar to the valve. This may had resulted the trained CNN to have difficulties in precisely segmenting the aortic valve from the rest of the images.

Further improvement can be made if more data are used for the neural network train. There are many factors that may affect the quality of echocardiogram images. Examples of these factors are obesity of the patient, inability to immobilize the patient, improper use of echocardiogram and the type of echocardiogram machine itself (Shmulewitz, Teefey, & Robinson, 1993). Thus, providing the neural network with more dataset for training will help to resolve the variations within the data and increases its accuracy. Alternatively, other CNN models such as GoogLeNet and VGG Net can be used instead of AlexNet. Using these models may provide a better result due to the fact they have more layers in their network. However as previously stated, using such CNN models requires more powerful computational resources especially the graphic processing unit of the computer.

Future studies should also include quantitative analysis of the results obtained. Examples of methods that can be used are Dice coefficient and Hausdorff distance (Rockafellar & Wets, 2009; Sørensen, 1948). These methods allow comparison between the ground truth and the resulted image in pixel value. Quantitative analysis may further increase the objective value of the study. However, these methods require experts in the field such as doctors and sonographers to manually localize the heart valves to establish the ground truths.

In this study, we mainly use still images for the localization of the heart valve as it is more manageable to develop the neural network and decrease the complexity in coding. However, future studies should consider designing the system to be able to accept live-visual feed of echocardiogram machine. Analysis in real-time will prove more beneficial in the medical field. With the ability to localize the heart valves, the system will be able to guide junior doctors or sonographer to quickly learn the use of echocardiogram. Furthermore, the image can be further analysed to determine whether the valve is normal or diseased after the heart valves are located. Providing a "second opinion" regarding the patient condition may help doctors to better manage their patients.

CHAPTER 6

CONCLUSION

This study presents a method to utilise CNN for medical image analysis. AlexNet, a CNN model had been able to successfully locate the aortic valve using images of echocardiogram. The training of the neural network was performed using a deep neural network platform NVIDIA DIGITS. The resulted image from the trained CNN was validated by a sonographer to be accurate in locating the aortic valve. However, the trained neural network may not be able to precisely segment the aortic valve. Suggestions for improvements were to include larger set of data and use of different CNN architecture.

REFERENCES

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive Science*, *9*(1), 147–169.

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, (5), 834–846.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends® in Machine Learning*, *2*(1), 1–127.

Boudoulas, H., Vavuranakis, M., & Wooley, C. F. (1994). Valvular Heart Disease: The influence of changing etiology on nosology. *The Journal of Heart Valve Disease*, *3*(5), 516–526.

Briske-Anderson, M. J., Finley, J. W., & Newman, S. M. (1997). The Influence of Culture Time and Passage Number on the Morphological and Physiological Development of Caco-2 Cells. *Proceedings of the Society for Experimental Biology and Medicine*, *214*(3), 248–257. https://doi.org/10.3181/00379727-214-44093

Chen, X., Zhou, X., & Wong, S. T. C. (2006). Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Transactions on Biomedical Engineering*, *53*(4), 762–766.

d'Arcy, J. L., Coffey, S., Loudon, M. A., Kennedy, A., Pearson-Stuttard, J., Birks, J., … Wilson, J. (2016). Large-scale community echocardiographic screening reveals a major burden of undiagnosed valvular heart disease in older people: the OxVALVE Population Cohort Study. *European Heart Journal*, *37*(47), 3515–3522.

Dueholm, M., & Lundorf, E. (2007). Transvaginal ultrasound or MRI for diagnosis of adenomyosis. *Current Opinion in Obstetrics and Gynecology*, *19*(6), 505–512.

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., & Shet, V. (2013). Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv Preprint arXiv:1312.6082*.

Goodman, J. (2016). Digital images. *Salem Press Encyclopedia of Science*. Salem Press.

Harnad, S. (1991). Other bodies, other minds: A machine incarnation of an old philosophical problem. *Minds and Machines*, *1*(1), 43–54.

Hertzberg, B. S., Kliewer, M. A., Bowie, J. D., Carroll, B. A., DeLong, D. H., Gray, L., & Nelson, R. C. (2000). Physician training requirements in sonography: how many cases are needed for competence? *American Journal of Roentgenology*, *174*(5), 1221–1227.

Hinton, G. E. (2009). Deep belief networks. *Scholarpedia*, *4*(5), 5947.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725–1732).

Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv Preprint arXiv:1408.5882*.

Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning*, *30*(2–3), 271–274.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).

Ladich, E., Yahagi, K., Romero, M. E., & Virmani, R. (2016). Vascular diseases: aortitis, aortic aneurysms, and vascular calcification. *Cardiovascular Pathology*, *25*(5), 432–441.

Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, *8*(1), 98–113.

LeCun, Y. (2015). LeNet-5, convolutional neural networks. *URL: Http://yann. Lecun. Com/exdb/lenet*.

Lévy, D., & Jain, A. (2016). Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks. Retrieved from http://www.diglib.um.edu.my/interaktif/default.asp?url=http://search.ebscohost.com/login.aspx?direct=true&db=edsarx&AN=1612.00542&site=eds-live

Li, J., Luong, M.-T., & Jurafsky, D. (2015). A hierarchical neural autoencoder for paragraphs and documents. *arXiv Preprint arXiv:1506.01057*.

Maganti, K., Rigolin, V. H., Sarano, M. E., & Bonow, R. O. (2010). Valvular Heart Disease: Diagnosis and Management. *Mayo Clinic Proceedings*, *85*(5), 483–500.

Martínez, M. J., Ripollés, T., Paredes, J. M., Blanc, E., & Martí-Bonmatí, L. (2009). Assessment of the extension and the inflammatory activity in Crohn's disease: comparison of ultrasound and MRI. *Abdominal Imaging*, *34*(2), 141–148.

Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, *16*(5), 555–559.

Matsumoto, M., Matsuo, H., Beppu, S., Yoshioka, Y., Kawashima, Y., Nimure, Y., & Abe, H. (1976). Echocardiography diagnosis of ruptured aneurysm of sinus of Valsalva. Report of two cases. *Circulation*, *53*(2), 382 LP-389. Retrieved from http://circ.ahajournals.org/content/53/2/382.abstract

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115–133.

Meinel, L. A., Stolpen, A. H., Berbaum, K. S., Fajardo, L. L., & Reinhardt, J. M. (2007). Breast MRI lesion classification: Improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system. *Journal of Magnetic Resonance Imaging*, *25*(1), 89–95.

Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. In *Interspeech* (Vol. 2, p. 3).

Munoz, A. (2014). Machine Learning and Optimization. *URL: Https://www. Cims. Nyu. Edu/~ Munoz/files/ml_optimization. Pdf [Accessed 2016-03-02][WebCite Cache ID 6fiLfZvnG]*.

Nkomo, V. T., Gardin, J. M., Skelton, T. N., Gottdiener, J. S., Scott, C. G., & Enriquez-Sarano, M. (2006). Burden of valvular heart diseases: a population-based study. *The Lancet*, *368*(9540), 1005–1011.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66.

Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., & Nielsen, M. (2013). Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In *International conference on medical image computing and computer-assisted intervention* (pp. 246–253). Springer.

Rockafellar, R. T., & Wets, R. J.-B. (2009). *Variational analysis* (Vol. 317). Springer Science &

Business Media.

Rodés-Cabau, J., Taramasso, M., & O'Gara, P. T. (2016). Diagnosis and treatment of tricuspid valve disease: current and future perspectives. *The Lancet*, *388*(10058), 2431–2442. https://doi.org/http://dx.doi.org/10.1016/S0140-6736(16)00740-6

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.

Rumelhart, D. E., Hinton, G. E., & McClelland, J. L. (1986). A general framework for parallel distributed processing. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, *1*, 45–76.

Russell, S. J., & Norvig, P. (2003). Artificial Intelligence: A Modern Approach: Pearson Education. Inc.

Shen, S., Sandham, W., Granat, M., & Sterr, A. (2005). MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE Transactions on Information Technology in Biomedicine*, *9*(3), 459–467.

Shmulewitz, A., Teefey, S. A., & Robinson, B. S. (1993). Factors affecting image quality and diagnostic efficacy in abdominal sonography: A prospective study of 140 patients. *Journal of Clinical Ultrasound*, *21*(9), 623–630. https://doi.org/10.1002/jcu.1870210909

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv Preprint arXiv:1409.1556*.

Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol. Skr.*, *5*, 1–34.

Stevenson, J. G., Kawabori, I., Dooley, T., & Guntheroth, W. G. (1978). Diagnosis of ventricular septal defect by pulsed Doppler echocardiography. Sensitivity, specificity and limitations. *Circulation*, *58*(2), 322 LP-326. Retrieved from http://circ.ahajournals.org/content/58/2/322.abstract

Stewart, B. F., Siscovick, D., Lind, B. K., Gardin, J. M., Gottdiener, J. S., Smith, V. E., … Otto, C. M. (1997). Clinical factors associated with calcific aortic valve disease. *Journal of the American College of Cardiology*, *29*(3), 630–634.

Stewart, W. J., Douglas, P. S., Sagar, K., Seward, J. B., Armstrong, W. F., Zoghbi, W., … Schnittger, I. (1999). Echocardiography in emergency medicine: a policy statement by the American Society of Echocardiography and the American College of Cardiology. *Journal of the American Society of Echocardiography*, *12*(1), 82–84.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., … Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2818–2826).

Turing, A. M. (1948). Intelligent machinery, a heretical theory. *The Turing Test: Verbal Behavior as the Hallmark of Intelligence*, *105*.

World Health Organization. (2014). WHO NCD Malaysia Country Profile. Retrieved from http://www.who.int/gho/countries/mys/country_profiles/en/

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In