 https://doi.org/10.24191/jeesr.v21i1.005

Classification of Left Ventricle and Non- Left Ventricle Segment for Cardiac Assessment Using Deep Convolutional Neural Network

Dayang Suhaida Awang Damit*, Siti Noraini Sulaiman, Muhammad Khusairi Osman, Noor Khairiah A. Karim and Samsul Setumin

***Abstract*—In large-scale medical imaging, selecting the best image to extract relevant imaging biomarkers for image assessment is crucial. Segmentation of the left ventricle (LV) and myocardium are performed in computer-aided analysis usually at short-axis slices of cardiac magnetic resonance (MR) image to quantify cardiovascular disease assessment, such as myocardial scarring, LV ejection fraction and LV mass. The need to correctly identify a short-axis slice range for efficient quantification is preferred for automatic classification of the slice range of interest. The goal of this research is to establish an image processing method for the segmentation of Left ventricle scar from late gadolinium-enhanced (LGE) MR images. In order to achieve the main purpose, the work is divided into two parts, the first is to identify the cardiac Left ventricle segment (LVS) in the stack of short-axis LGE MR images and the second part; detecting the scar in between the LV myocardium area. This paper will present the outcomes of the first part by utilizing a deep convolutional neural network (DCNN) to construct an automatic system for classifying LVS and Non- Left Ventricle Segment (Non-LVS) in MR images. The same image dataset will be used for a comparative analysis with six DCNN models designed from scratch and three famed pre-trained networks, Alexnet, GoogleNet and SqueezeNet. Each model is trained up to 35 epochs using the Cardiac Atlas dataset and cross-validation method. The outcome from this work demonstrated that the DCNN3Y performs well over small training data with an average accuracy of 94.49%. Whereas SqueezeNet outperformed the three pre-trained networks with an average accuracy of 96.96%. It has also been discovered that increasing the number of filters and their subsequent configuration slightly influences the network's performance. This produces very promising results showing that it is ready to be used in the second part of this research. The outcome of this research can compensate for the deficiencies of manual detection in the original image detection system, increase detection efficiency, reduce detection misjudgments, and advance the development of automated and intelligent detection in the medical field.**

***Index Terms*— Cardiac, CNN, Feature Map, Left Ventricle, LGE MRI**

This manuscript is submitted on 12th January 2022 and accepted on 29th June 2022.

Dayang Suhaida Awang Damit[1], Siti Noraini Sulaiman[1,2], Muhammad Khusairi Osman[1] and Samsul Setumin[1] are with the [1]Centre for Electrical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, Permatang Pauh, 13500 Pulau Pinang, MALAYSIA and [2]Integrative Pharmacogenomics Institute (iPROMISE), UiTM Puncak Alam Campus, Bandar Puncak Alam, Puncak Alam, Selangor, 42300, Malaysia (e-mail: dayang671@uitm.edu.my, sitinoraini@uitm.edu.my, khusairi@uitm.edu.my and samsuls@uitm.edu.my).

Noor Khairiah A. Karim is with the Advanced Medical and Dental Institut, Universiti Sains Malaysia, Malaysia (e-mail: drkhairiah @ usm.my)

*Corresponding author
Email address: dayang671@uitm.edu.my

I. INTRODUCTION

ACCORDING to the World Health Organization (WHO) [1], cardiovascular disease (CVD) caused the deaths of 17.9 million people in 2019, accounting for 32% of all fatalities worldwide. Heart attacks were responsible for one-third of these deaths and can be categorized as a silent killer. Furthermore, recent investigations by researchers in [2-5] addressed various forms of cardiovascular involvement including myocardium scarring in patients with post covid 19. It is a sign that the number of people undergoing cardiac analysis is growing every year. As a result, rapid and precise assessment should be carried out in order to improve the growing interest in medical image analysis. In the field of cardiac magnetic resonance imaging (MRI), a stack of short-axis cardiac MR images frequently includes outmost slices of the left ventricle (LV) coverage [6]. However, in the process of examining the patient's myocardium scar, not all slices will be considered (usually only the middle slices will be considered as it contains useful information, and the rest will be neglected).

Fig. 1 shows an example of a stack of short-axis slices with class labelling encompassing the full LV. The slices labelled as LVS normally will be taken as slices of interest while LAP and LBS will be neglected. The segmentation step is usually preceded by selecting a short-axis slice range; i.e. the LVS slices, which this selecting short-axis slice range stage may require a manual procedure and time-consuming. This demonstrates the importance of using proper and accurate slice selection steps to ensure the success of the next process. The need to correctly identify a short-axis slice range for efficient quantification also has been emphasized in standard cardiac imaging postprocessing guidelines by [7].

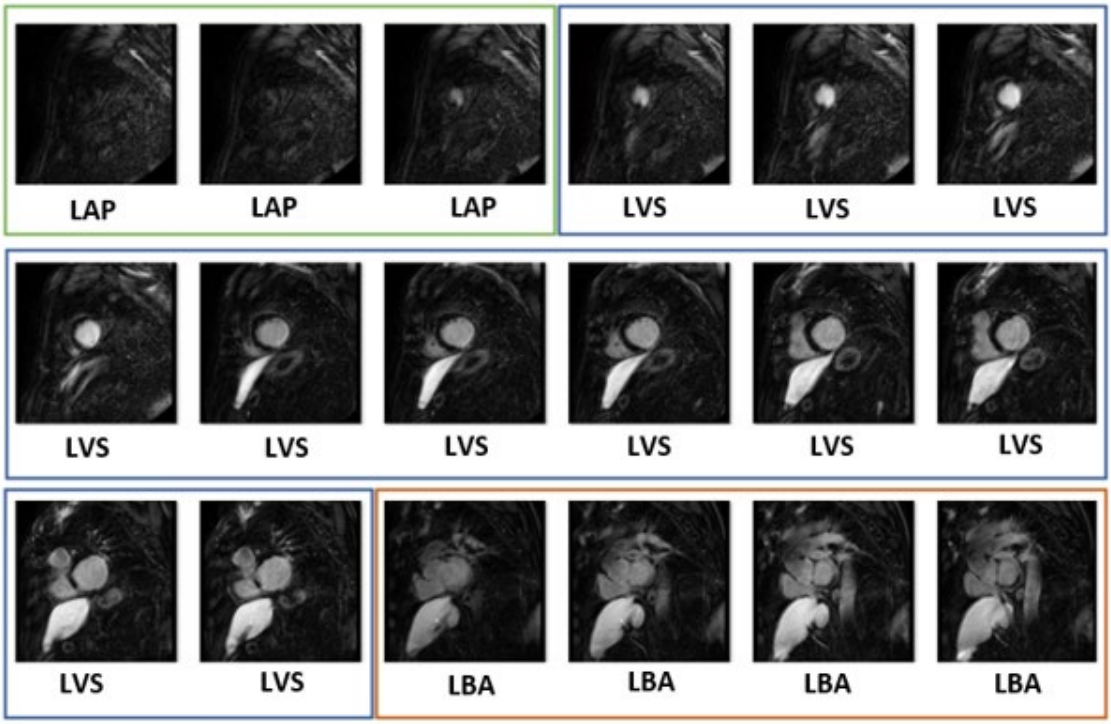

Fig. 1. Cardiac MRI short-axis slices with example of three class labelling

The potential of a convolutional deep learning network is no doubt a challenge to improve automated medical image classification. However, it is being actively explored. Training a deep convolutional neural network (DCNN) from scratch is likely to overfit in circumstances when the number of images for training is insufficiently large. Data augmentation is a well-known method of overcoming this problem [8]. However, hyperparameters such as the factors that dictate the network structure and the variables that govern how the network is trained, can influence the network optimization process. Hence, further research is needed on this topic. It is also worth noting that batch normalization and dropout improved network performance as batch normalization overcame the internal covariate shift and dropout tackling the overfitting issue [9]. Besides, using a pre-trained network instead of training from scratch is an alternative. This version of DCNN can also be deployed using an approach known as transfer learning [10-11] of the pre-trained model.

Currently, several DCNN configurations are available, and the fundamental distinctions between these designs are represented by the patterns of layer interconnection and the total number of layers. The exploitation of depth and different structural adaptations is significantly improved in the DCNN learning capability. Abdullah et al. [12] investigated the effectiveness of adding convolutional layers to the DCNN model in classifying thorax and non-thorax regions. Three DCNN models with different network depths have been proposed to classify CT scan images for lung cancer detection. The outcome shows that the DCNN with two convolution layer structures, which has a 99.42% accuracy, is the most efficient network.

Namgyu Ho et al. [13] investigated the potential of using transfer learning to categorise short-axis slices in cardiac cine MR images using 18 DCNN models and three models trained from scratch and the outcome from their study indicated that the transfer learning VGG16 model produced the highest values in all evaluation categories considered and appeared to be the most appropriate choice for the cardiac slice range classification. However, from those two research, no additional research endeavour into CNN hyperparameter was presented.

This paper aims to extend the idea in [13], which the first part will focus on the designation of six DCNN models with different network depths, including the influence of the number of kernels on the DCNN model using late gadolinium enhanced (LGE) MR images. This idea is to assist beginners with a better understanding of basic deep learning architecture and how all of the element's work. Thus, this fundamental network will lead to a novel outcome while enabling more simple models to succeed. To the best of our knowledge, this is the first paper tackling the classification of short-axis slices in the LGE MR image modality. The second part is to observe the performance of three popular pre-trained model namely AlexNet, GoogleNet and SqueezeNet with the same image dataset. Then, the performance of the proposed DCNN model and the three pre-trained models will next be compared. The rest of the paper is organized as follows. Section 1 gives some introduction and reviews some related literature works. In Section 2, the detail of our methodology is presented. Section 3 is the results and discussion whereby Section 4 is the conclusion of this work and some recommendation for future work.

II. METHODOLOGY

This section presents a methodology to classify left and non-left ventricle segments for the cardiac assessment using the deep

convolutional neural network approach. It comprises four sub-sections; dataset, proposed DCNN from scratch, training using a pre-trained DCNN and performance criteria.

A. Dataset

All the DL models were trained on the image database consists of 15 LGE CMR images of humans/patients provided with an expert observers' consensus ground truth which is publicly available via the Cardiac Atlas project challenge website: https://www.cardiacatlas.org/challenges/ventricular-infarct-segmentation.

In this paper, LGE CMR images are classified into three categories: (1) the section which contains LV region (LVS), (2) Loss Apical (LAP) which is the Non-LV cardiac segment refers to the section which does not contain LV region located below the apical, (3) Loss Basal (LBA) which is also the Non-LV cardiac segment refers to the section which does not contain LV region located at the basal. The slices labelled as LVS will be taken as slices of interest while LAP and LBA will be neglected in clinical practice. The criterion used to determine a correct apical and basal slice position is observable using ground truth given with the dataset [14].

These images are used as the dataset for the developed DL models and have undergone the data augmentation process. Table I displays the total number of images in each class before and after data augmentation. The image was oversampled corresponding to the LAP and LBA classes using a simple augmentation approach that applied random rotations between 90, -90 and 15 degrees for each image to overcome the class imbalance and overfitting issue. Each class was selected equally so that the DCNN perceives a balanced dataset with 255 images during the five-fold cross-validation process. The total image containing image slices from 15 patients was randomly divided into five distinct subsets for training and validation set. One subset was used for validation for each fold, and the remaining four were used for training. Therefore, to assess the performance of a single parameter choice, a total of five models were trained and validated.

TABLE I
THE TOTAL NUMBER OF IMAGES IN EACH CLASS BEFORE AND AFTER DATA AUGMENTATION

	Original Dataset				After Data Augmentation.				Training and Validation Data			
	LVS	LAP	LBA	TOTAL	LVS	LAP	LBA	TOTAL	LVS	LAP	LBA	TOTAL
Image Count	255	62	81	398	255	310	324	889	255	255	255	765
Percentage	64%	16%	20%	100%	29%	35%	36%	100%	33.30%	33.30%	33.30%	100%

B. Proposed DCNN From Scratch

As previously stated, the interests will be more on detecting LVS in the CMR LGE image dataset. The top and bottom of Non-LVS Cardiac MRI slices are also classified using the proposed neural network inherited from the standard framework of CNN. Besides, current approaches such as the parametric rectified linear unit, dropout, and batch normalization adopted in [16] were rearranged in the layer. Each convolutional layer is stacked with a batch normalization layer and a ReLU activation function. The ReLU activation function induces some sort of non-linearity for better operation of subsequent layers. Six DCNN models were trained for comparison. It consists of different types of layers, including an input layer, convolutional layers, pooling layers, FC layers, and an output layer. Analytically, the number of convolution layer was varied from 1 to 4, and the network with the best performance was selected. Fig. 2 illustrates the general framework of the proposed DCNN model. The original MR images are scaled to 227×227×3 pixels and then used as input.

The hyper-parameters of a convolutional layer are the variables that determine the network structure (Example: kernel size, number of kernels, stride, and padding) and also the factor that governs how the network is trained (e.g. learning rate, batch size, number of epoch). These parameters must be chosen carefully in order to provide a highly accurate and fast output. The parameters were set up to specify the training option in the DCNN in MATLAB is shown in Table II. Kandel et al. [17] in their study highlighted that higher learning rates necessitate larger batch sizes. While the use of a large batch size value is not recommended in our study because the dataset is insufficiently large, batch size of 16 and a Stochastic Gradient Descent (SGD) with momentum optimizer [18] was selected to train all DCNN models at a fixed learning rate of 0.0001.

TABLE II
DETAILS OF THE TRAINING PARAMETER SETUP

CNN Parameter	Specifications
Optimizer	SGDM
Learning Rate	0.0001
MaxEpochs	35
Mini Batch Size	16
Convolution layer	1/2/3/4

To investigate the influence of the number of kernels in the network, the performance over the input images with a fixed kernel window size of 3x3 and stride value that used in all experiments is 1 have been examined using different numbers of kernel sequences [19]. The number of kernels in the layer for each DCNN model was set up as in Table III. As DCNN is dealing with the multi-classification problem, the fully connected layer was set to the dense layer with the value of 3 (classes) and softmax function was deployed at the output layer with 0.7 dropout [20].

 https://doi.org/10.24191/jeesr.v20i1.005

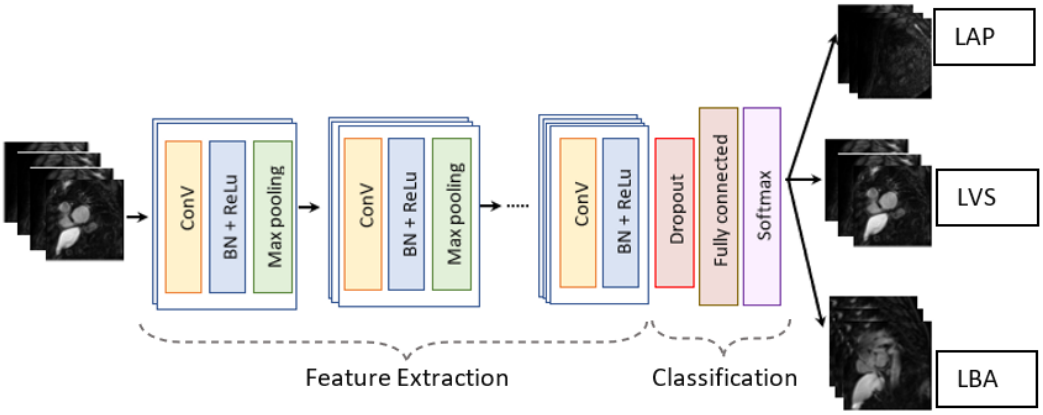

Fig. 2. Overview of the proposed DCNN model for cardiac MRI LV assessment

Table III, IV, V, VI, VII and VIII show the network structure architecture for DCNN1, DCNN2, DCNN3X, DCNN3Y, DCNN4X and DCNN4Y respectively. These six different networks were developed to find the best accuracy for the system. Other performance parameters such as precision, recall, and F1 score are considered to ensure reliable outcomes.

TABLE III
NETWORK STRUCTURE ARCHITECTURE FOR DCNN1 MODEL

Layer	DCNN1			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
fc	-	-	-	-

TABLE IV
NETWORK STRUCTURE ARCHITECTURE FOR DCNN2 MODEL

Layer	DCNN2			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
Conv2	3x3	16	1	'same'
Relu	-	-	-	-
Pool2	2	16	2	0
fc	-	-	-	-

TABLE V
NETWORK STRUCTURE ARCHITECTURE FOR DCNN3X MODEL

Layer	DCNN3X			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
Conv2	3x3	16	1	'same'
Relu	-	-	-	-
Pool2	2	16	2	0
Conv3	3x3	16	1	'same'
Relu	-	-	-	-
Pool3	2	16	2	0
fc	-	-	-	-

TABLE VI
NETWORK STRUCTURE ARCHITECTURE FOR DCNN3Y MODEL

Layer	DCNN3Y			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
Conv2	3x3	16	1	'same'
Relu	-	-	-	-
Pool2	2	16	2	0
Conv3	3x3	32	1	'same'
Relu	-	-	-	-
Pool3	2	32	2	0
fc	-	-	-	-

TABLE VII
NETWORK STRUCTURE ARCHITECTURE FOR DCNN4X MODEL

Layer	DCNN4X			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
Conv2	3x3	16	1	'same'
Relu	-	-	-	-
Pool2	2	16	2	0
Conv3	3x3	16	1	'same'
Relu	-	-	-	-
Pool3	2	16	2	0
Conv4	3x3	32	1	'same'
Relu	-	-	-	-
Pool4	2	32	2	0
fc	-	-	-	-

TABLE VIII
NETWORK STRUCTURE ARCHITECTURE FOR DCNN4Y MODEL

Layer	DCNN4Y			
	Kernel size	No. of kernel	Stride	Pad
Conv1	3x3	8	1	'same'
ReLu	-	-	-	-
Pool1	2	8	2	0
Conv2	3x3	16	1	'same'
Relu	-	-	-	-
Pool2	2	16	2	0
Conv3	3x3	32	1	'same'
Relu	-	-	-	-
Pool3	2	32	2	0
Conv4	3x3	32	1	'same'
Relu	-	-	-	-
Pool4	2	32	2	0
fc	-	-	-	-

C. Training Using a Pre-Trained CNN

Pretrained networks are a common and very effective alternative for small image data samples [21]. Therefore, this study compared the proposed network to the three well-known CNN designs for transfer learning methods, which are AlexNet [22], GoogLeNet [23], and SqueezeNet [24]. To compare our proposed model to transfer learning models, all pre-trained networks are fixed with the same learning rate, batch size, and optimizer as same as the proposed model. The last fully connected layer was fine-tuned in order to accommodate the three image categories. The number of epochs was fixed at 35 to ensure consistency in all network results.

D. Performance Criteria

The performance of the proposed classification model in this study is evaluated via the accuracy of the clinical test, which refers to how correctly a diagnostic test identifies and excludes a specific condition, precision, recall and F1-score. The measures are mathematically expressed as follows [25] using the confusion matrix generated by each model:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where the TP, TN, FP, and FN represent the number of true positive, true negative, false positive, and false negative in the confusion matrix generated from each model, respectively.

III. RESULTS AND DISCUSSION

This section presents the results obtained and its discussion. Table IX compares the training performance of five-fold cross-validation for the six DCNN models and the pre-train models in terms of precision, recall, F1-score, validation accuracy for each class, as well as the average for each performance metric. The first analysis focuses on how the convolution layer increment affects the performance of all evaluations. The network attained the highest score by three convolution layers architecture named DCNN3Y with 96.09% accuracy in LVS, 94.4%, and 93% accuracy in LAP and LBA, respectively. The result also shows a good agreement between accuracy and other performance scores. In the precision score, which relates to the low false positive rate, shows a high result obtained in the LVS class with 96.09%, followed by 94.35% in LAP and 93.04% in LBA. Recall score corresponds to the ratio of correctly predicted positive observations to all observations in the actual class. The recall performance of all models follows the same trend, with the LVS class having the lowest recall performance. However, the DCNN3Y model still has the highest recall score with 92.47% in the LVS class, 94.35% and 97.72% in the LAP and LBA class, respectively. The DCNN with two convolution layers, namely DCNN2, had the second-best network performance, which differed slightly from DCNN3Y in average precision, recall, F1-score, and validation accuracy by 0.46% 0.54%, 0.47%, and 0.46%, respectively.

Another important finding in this paper was the effect of the configuration number of filters on the model's performance. The increased number of kernels used in DCNN3Y architectures with the configuration of {8, 16, 32} improved its performance as compared to DCNN3X {8,16,16}. Although the number of kernels in the network has increased, the addition of four convolution layers has slightly degraded overall performance. This result is agreed upon by Weimer et al. [26] by saying that because the complexity of a defect is restricted, adding more neurons will result in a saturation effect, and the CNN will be unable to generate more complex features. From the overall result in Table IX, the DCNN3Y shows promising efficacy in classifying LVS and Non-LVS in a limited MRI dataset.

Next, in comparing the previous finding with the three pre-trained networks, the DCNN3Y model shows competitive results against AlexNet. The overall result can be seen in Table X. SqueezeNet provides the highest values in all average evaluation scores with a 0.15% slight gap difference with GoogleNet models for average precision, recall, F1-score and validation accuracy.

The performance of training/validation accuracy and loss using all pre-trained models and DCNN3Y were converged in line graphs shown in Fig, 3 and Fig. 4. It indicates that there is no sign of underfitting and overfitting. Among the three pre-trained models, GoogleNet and SqueezeNet show a small degree of loss on both the training dan validation set. The AlexNet had started with high validation loss as compared to GoogleNet and SqueezeNet. Even though the train from a scratch model, DCNN3Y shows a high difference in training and validation loss, this model did not exhibit patterns of a high degree of overfitting in the validation loss and obtain a good validation accuracy. Hence, in the future, this study will further refine the current structure of the model to reduce training time by tuning some training parameters, such as the deep learning optimizer, and verify this network with other LGE MR image datasets.

TABLE IX
PERFORMANCE OF THE SIX DCNN MODELS WITH DIFFERENT LAYER AND NUMBER OF KERNELS

Deep Learning Architectures	Precision			Recall			F1 score			Accuracy			Average Precision	Average Recall	Average F1 score	Average Accuracy
	LVS	LAP	LBA	LVS	LAP	LBA	LVS	LAP	LBA	LVS	LAP	LBA				
DCNN1	0.9435	0.9087	0.9348	0.9118	0.9372	0.9389	0.9274	0.9227	0.9368	0.9435	0.9087	0.9348	0.9290	0.9293	0.9290	0.9290
DCNN2	0.9478	0.9174	0.9565	0.9316	0.9420	0.9565	0.9397	0.9295	0.9565	0.9478	0.9174	0.9565	0.9406	0.9434	0.9419	0.9406
DCNN3X {8,16,16}	0.9478	0.9217	0.9304	0.9198	0.9339	0.9469	0.9336	0.9278	0.9386	0.9478	0.9217	0.9304	0.9333	0.9336	0.9333	0.9333
DCNN3Y {8,16,32}	**0.9609**	**0.9435**	**0.9304**	**0.9247**	**0.9435**	**0.9772**	**0.9424**	**0.9435**	**0.9532**	**0.9609**	**0.9435**	**0.9304**	**0.9449**	**0.9484**	**0.9464**	**0.9449**
DCNN4X {8,16,16,32}	0.9435	0.9348	0.9304	0.9118	0.9513	0.9469	0.9274	0.9430	0.9386	0.9435	0.9348	0.9304	0.9362	0.9367	0.9363	0.9362
DCNN4Y {8,16,32,32}	0.9522	0.9348	0.9000	0.8939	0.9307	0.9673	0.9221	0.9328	0.9324	0.9522	0.9348	0.9000	0.9290	0.9306	0.9291	0.9290

TABLE X
PERFORMANCE OF THE PRE-TRAIN MODELS: ALEXNET, GOOGLENET & SQUEEZENET

Deep Learning Architectures	Precision			Recall			F1 score			Accuracy			Average Precision	Average Recall	Average F1 score	Average Accuracy
	LVS	LAP	LBA	LVS	LAP	LBA	LVS	LAP	LBA	LVS	LAP	LBA				
DCNN3Y {8,16,32}	0.9609	0.9435	0.9304	0.9247	0.9435	0.9772	0.9424	0.9435	0.9532	0.9609	0.9435	0.9304	0.9449	0.9484	0.9464	0.9449
AlexNet	0.9478	0.9478	0.9391	0.9046	0.9561	0.9774	0.9257	0.9520	0.9579	0.9478	0.9478	0.9391	0.9449	0.9460	0.9452	0.9449
GoogLeNet	0.9565	0.9870	0.9609	0.9524	0.9742	0.9779	0.9544	0.9806	0.9693	0.9565	0.9870	0.9609	0.9681	0.9682	0.9681	0.9681
SqueezeNet	**0.9565**	**0.9739**	**0.9783**	**0.9524**	**0.9782**	**0.9783**	**0.9544**	**0.9760**	**0.9783**	**0.9565**	**0.9739**	**0.9783**	**0.9696**	**0.9696**	**0.9696**	**0.9696**

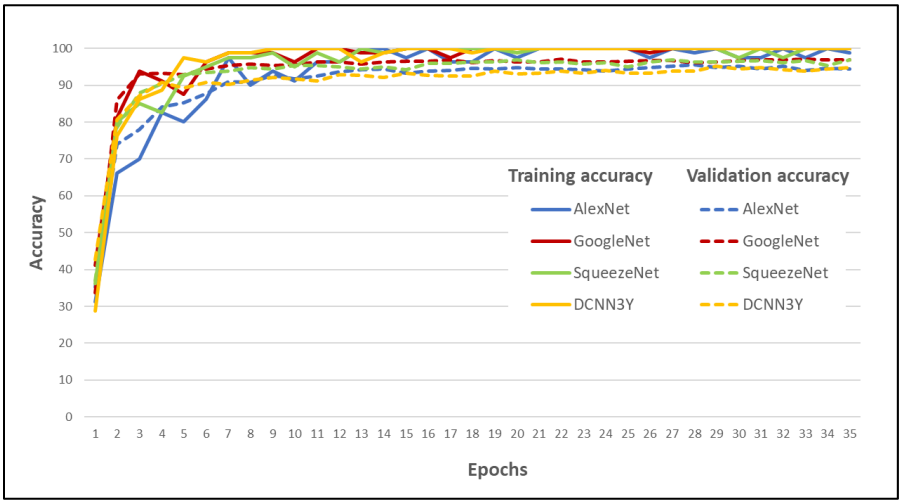

Fig. 3. Plots of training and validation accuracy in the average of five-fold cross-validation for the DCNN3Y proposed model and three pre-train model

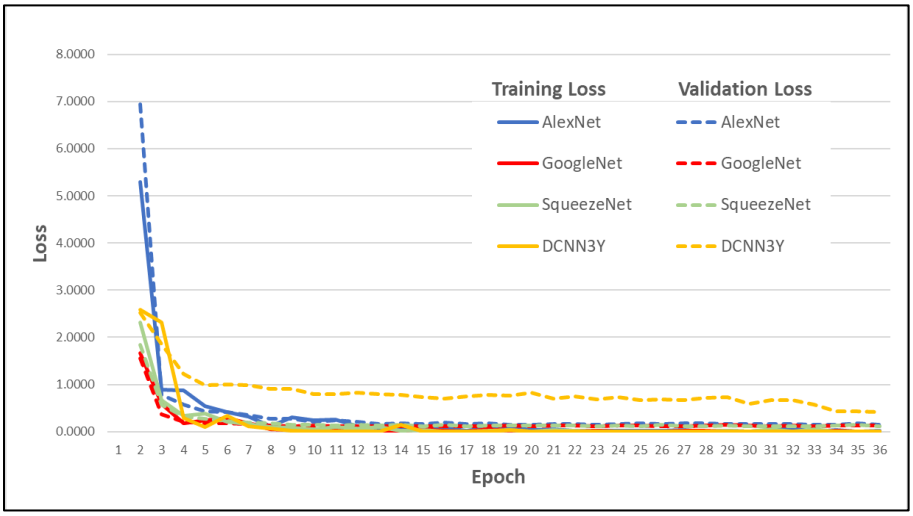

Fig. 4. Plots of training and validation loss in the average of five-fold cross-validation for the DCNN3Y proposed model and three pre-train model

IV. CONCLUSION

A comparison of six DCNN models with different numbers of convolution layers was carried out and the influence of the number of kernels in the network was observed. The results show that the three-convolution layer, DCNN3Y with the increasing number of kernels configuration {8,16,32} outperforming all DCNN models from scratch with 94.49% and 94.64% in average accuracy and F1-score, respectively. Nevertheless, the impact of adding the number of kernels is comparatively low. The higher the number of kernels, the higher the number of abstractions that the network can extract from image data. Still, the limitation must be observed to avoid degradation in the performance due to the saturation effect cause the CNN cannot generate more complex features. When the DCNN3Y model was compared to the three well-known pre-trained networks, the DCNN3Y model showed competitive results against AlexNet. However, the pre-trained SqueezeNet models provided the highest values in all evaluation categories. This result indicates a positive start to the proposed model because DCNN3Y shows a competitive outcome with a simpler design than the pre-train networks. It also shows that there is room for further improvement as well. Thus, SqueezeNet and DCNN3Y model appeared to be the most appropriate choice for the classification of LVS and non-LVS in the cardiac LGE MRI short-axis slice range for the next stage of this research.

As mentioned, the main goal of this research is to establish an image processing method for the segmentation of Left ventricle scar from LGE MR images. In order to achieve the main purpose, the work is divided into two parts, the first is to identify the cardiac Left ventricle segment (LVS) in the stack of short-axis LGE MR images and the second part, detecting the scar between in the LV myocardium area. Considerably more work will need to be done, this paper presents the primary results of the first part. Hence, the following aspect would be focusing on the future: validating the best-proposed network with different LGE MRI dataset and improving the training through various deep learning optimizers and kernel sizes in the convolutional layers of the network.

ACKNOWLEDGMENT

The authors would like to express their gratitude to members of the Advanced Control System and Computing Research Group (ACSCRG), Advanced Rehabilitation Engineering in Diagnostic and Monitoring Research Group (AREDiM), Integrative Pharmacogenomics Institute (iPROMISE), and Centre for Electrical Engineering Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang for their assistance and guidance during the fieldwork. The authors are grateful to Universiti Teknologi MARA, Cawangan Pulau Pinang for their immense administrative and financial support. Special thanks to the Imaging Department, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Kepala Batas, Pulau Pinang for the professional consultation and expert guide.

REFERENCES

[1] World Health Organization. "*Cardiovascular diseases (CVDs),"* 11 June 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). [Accessed: 29-Nov-2021].

[2] L. Huang et al., "Cardiac Involvement in Patients Recovered From COVID-2019 Identified Using Magnetic Resonance Imaging," *JACC Cardiovasc. Imaging, vol. 13, no. 1*1, pp. 2330–2339, 2020.

[3] C. Butter and J. Sperzel, "Gerätebasierte rhythmologische Diagnostik und Therapie in COVID-19-Zeiten," *Der Kardiol., vol. 15, no. 3*, pp. 272–281, Jun. 2021.

[4] M. Araz, Ç. Soydal, G. Sütçü, B. Demir, and E. Özkan, "Myocardial perfusion SPECT findings in postCOVID period," Eur. J. Nucl. Med. Mol. Imaging, no. 0123456789, pp. 1–6, 2021.

[5] U. A. Hasnie, P. Bhambhvani, A. E. Iskandrian, and F. G. Hage, "Prevalence of abnormal SPECT myocardial perfusion imaging during the COVID-19 pandemic," *Eur. J. Nucl. Med. Mol. Imaging, vol. 48, no. 8,* pp. 2447–2454, Jul. 2021.

[6] D. J. Im et al., "Guidelines for cardiovascular magnetic resonance imaging from the Korean society of cardiovascular imaging—part 3: Perfusion, delayed enhancement, and t1-and t2 mapping," *Korean J. Radiol., vol. 20, no. 12*, pp. 1562–1582, 2019.

[7] J. Schulz-Menger et al., "Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update: Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing," *J. Cardiovasc. Magn. Reson., vol. 22, no. 1,* pp. 1–22, Mar. 2020.

[8] Huynh-The and D. Kim, "Data Augmentation For CNN-Based 3D Action Recognition on Small-Scale Datasets," *2019 IEEE 17th International Conference on Industrial Informatics (INDIN), 2019*, pp. 239-244, doi: 10.1109/INDIN41052.2019.8972313.

[9] S. H. Wang, K. Muhammad, J. Hong, A. K. Sangaiah, and Y. D. Zhang, "Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization," *Neural Comput. Appl., vol. 32, no. 3,* pp. 665–680, 2020.

[10] C. Oksuz and M. K. Gullu, "Deep Feature Extraction Based Fine-Tuning," 2020 28th Signal Processing and Communications Applications Conference (SIU), 2020, pp. 1-4, doi: 10.1109/SIU49456.2020.9302108.

[11] J. Margeta, A. Criminisi, R. Cabrera Lozoya, D. C. Lee, and N. Ayache, "Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis., vol. 5, no. 5,* pp. 339–349, Sep. 2017.

[12] M. F. Abdullah, S. N. Sulaiman, M. K. Osman, N. K. A. Karim, I. S. Isa, and I. L. Shuaib, "Designation of Thorax and Non-Thorax Regions for Lung Cancer Detection in CT Scan Images using Deep Learning," J. *Electr. Electron. Syst. Res., vol. 17, no*. DEC 2020, pp. 41–49, Dec. 2020.

[13] N. Ho and Y. C. Kim, "Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification," *Sci. Rep., vol. 11, no. 1*, pp. 1–11, 2021.

[14] R.Karim et al. "Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images,", *Medical Image Analysis. Vol.30*, pp. 95-107 (2016)

[15] D. N. Metaxas and Z. Yan, "*Deformable models, sparsity and learning-based segmentation for cardiac MRI based analytics*," in Handbook of Medical Image Computing and Computer Assisted Intervention, Elsevier, 2020, pp. 273–292.

[16] Trivizakis, E., Ioannidis, G. S., Melissianos, V. D., Papadakis, G. Z., Tsatsakis, A., Spandidos, D. A., Marias, K. "A novel deep learning architecture outperforming 'off-the-shelf' transfer learning and feature-based methods in the automated assessment of mammographic breast density". *Oncology Reports 42.5 (2019),* pp. 2009-2015.

[17] I. Kandel and M. Castelli, "The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset," *ICT Express, vol. 6, no. 4*, pp. 312–315, Dec. 2020.

[18] Brahimi, M.; Arsenovic, M.; Laraba, S.; Sladojevic, S.; Boukhalfa, K.; Moussaoui, *A. Deep learning for plant diseases: Detection and saliency map visualisation*. In Human and Machine Learning; Springer: Berlin, Germany, 2018; pp. 93–117.

[19] W. S. Ahmed and A. A. A. Karim, "The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN," Proc. 2020 Int. Conf. Comput. Sci. Softw. Eng. CSASE 2020, no. January 2021, pp. 88–93, 2020.

[20] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting." *J. Mach. Learn. Res., 15,* pp.1929-1958,2014

[21] P. Sahu, A. Chug, A. P. Singh, D. Singh, and R. P. Singh, "Implementation of CNNs for Crop Diseases Classification: A Comparison of Pre-trained Model and Training from Scratch," *Int. J. Comput. Sci. Netw. Secur., vol. 20, no. 10*, pp. 206–215, 2020.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM, vol. 60, no. 6,* pp. 84–90, May 2017.

[23] Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[24] Iandola, F.N., Moskewicz, M.W., Ashraf, K., Han, S., Dally, W.J., & Keutzer, K. (2016). "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size." ArXiv, abs/1602.07360.

[25] A. Kulkarni, D. Chong, and F. A. Batarseh, "Foundations of data imbalance and solutions for a data democracy," in Data Democracy, Elsevier, 2020, pp. 83–106.

[26] D. Weimer, B. Scholz-Reiter, and M. Shpitalni, "Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection," *CIRP Ann., vol. 65, no.* 1, pp. 417–420, 2016.