Pertanika J. Sci. & Technol. 33 (2): 875 - 891 (2025)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Better Network Optimization Through Batch Normalization in Left Ventricle Chamber Classification

Dayang Suhaida Awang Damit[1], Siti Noraini Sulaiman[1,2,3]*, Muhammad Khusairi Osman[1], Noor Khairiah A Karim[4,5] and Samsul Setumin[1]

[1]Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA Cawangan Pulau Pinang, Permatang Pauh Campus,13500 Pulau Pinang, Malaysia
[2]Integrative Pharmacogenomics Institute (iPROMISE), UiTM Puncak Alam Campus, Bandar Puncak Alam, 42300 Puncak Alam, Selangor, Malaysia
[3]Advanced Rehabilitation Engineering in Diagnostic and Monitoring Research Group (AREDiM), Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA Cawangan Pulau Pinang, Permatang Pauh Campus,13500 Pulau Pinang, Malaysia
[4]Department of Biomedical Imaging, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, 13200 Kepala Batas, Pulau Pinang, Malaysia
[5]Heart Failure Research Initiative, Advanced Medical & Dental Institute, Universiti Sains Malaysia, Bertam, 13200 Kepala Batas, Pulau Pinang, Malaysia

ABSTRACT

Convolutional neural networks (CNNs) have emerged as a prominent deep learning technique for medical image classification. This study investigated the impact of batch normalization layer placement on the performance of the CNNs model in classifying the left ventricle segment in Delayed-enhancement cardiac magnetic resonance (De-CMR) image slices. Three batch normalization arrangements, including one without a batch normalization layer, were examined to assess their impact. Additionally, the influence of three learning rates (0.0001, 0.001, 0.01) from two different types of optimizers, namely Adam and Sgdm, was explored to identify the optimal configuration for our proposed CNN model. A model without batch normalization was used as a baseline for comparison. The results show that placing batch normalization after the convolutional layers, combined with the Adam optimizer and a learning rate of 0.0001, yielded the best performance, improving classification accuracy from 83.1% to 88.4%. These results highlight the significance of batch normalization layers with optimal configuration in enhancing

ARTICLE INFO

Article history:
Received: 07 June 2024
Accepted: 18 November 2024
Published: 07 March 2025

DOI: https://doi.org/10.47836/pjst.33.2.15

E-mail addresses:
dayang671@uitm.edu.my (Dayang Suhaida Awang Damit)
sitinoraini@uitm.edu.my (Siti Noraini Sulaiman)
khusairi@uitm.edu.my (Muhammad Khusairi Osman)
drkhairiah@usm.my (Noor Khairiah A Karim)
samsuls@uitm.edu.my (Samsul Setumin)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

© Universiti Putra Malaysia Press

the performance in the classification of the left ventricle and non-LV chambers in De-CMR images, thereby facilitating improvements in the streamlined workflow for automated myocardial infarction diagnosis.

Keywords: Batch normalization, classification, convolution neural network, delayed-enhancement cardiac magnetic resonance, learning rate

INTRODUCTION

Ischemic heart disease stands as the top cause of death globally (World Health Organization, 2021). It is caused by clogged cardiac arteries, predominantly affecting the left ventricular (LV) region, leading to heart tissue death. This condition leads to heart attacks, or is also known as myocardial infarction. The exponential growth in the number of MRI scans performed can be attributed to recent research addressing a variety of cardiovascular involvements, including infarcted muscle in patients with post-COVID-19 (Huang et al., 2020; Li et al., 2023; Toro et al., 2021). This surge in demand has placed an overwhelming workload on radiologists, making it challenging to efficiently analyze each image and provide timely and accurate diagnoses. Delayed-enhancement Cardiac Magnetic Resonance (De-CMR) is a well-known method for characterizing cardiac tissue, especially for assessing regional scar formation and myocardial fibrosis by enhancing abnormal cardiac tissue. De-CMR acquires a stack of short-axis CMR image slices that cover the entire cardiac region, including the outermost slices (Im et al., 2019), as presented in Figure 1. However, not every image contains clinically relevant segments, as only slices that include left ventricular structures are essential for the early diagnosis of myocardial infarction. The uppermost basal and lowermost apical images do not contain the LV chamber. These images are highlighted in the yellow box in Figure 1 and are classified as non-LV chamber regions, while the intermediate slices contain the LV chamber. The standard diagnostic process used by radiologists involves manually reviewing multiple short-axis slices to identify infarcted tissue, with a primary focus on the LV chamber, where infarctions are most likely to occur (Karim et al., 2016). This process is time-consuming because each patient generates a stack of 15-45 image slices, depending on the image thickness and size of the heart. Each slice must be carefully examined to assess the condition of the myocardium, which can be labor-intensive, especially in high-demand clinical environments. This manual process can be further complicated by factors such as image quality variability due to patient movement or noise artifacts, anatomical variations, and overlapping structures. For instance, the basal and apical regions of the LV chamber are often difficult to distinguish from surrounding tissues, adding to the diagnostic challenge.

With the advancement of computer vision and image processing techniques, many studies focus on the segmentation of LV chambers but rely on manual image selection as a starting point (Chen et al., 2020; Leong et al., 2019; Shaaf et al., 2022). Surprisingly, the

initial step of identifying relevant image slices has been overlooked. Figure 2 illustrates our proposed automated LV chamber classification as a key component of the overall workflow for fully automated myocardial infarction diagnosis. This step eliminates the need for manual selection, ensuring that only the necessary images proceed to the segmentation phase. By developing an automated classification method that identifies the LV chamber before segmentation, this research enables a fully automated and streamlined process for diagnosing myocardial infarction, addressing a key limitation in existing methods.

Image classification, a fundamental task in computer vision, has gained significant traction in the field of medical imaging (Zorkafli et al., 2019; Abdullah et al., 2020). In recent years, Convolutional Neural Networks (CNNs) have emerged as a powerful tool for analyzing medical images due to their ability to automatically learn discriminative

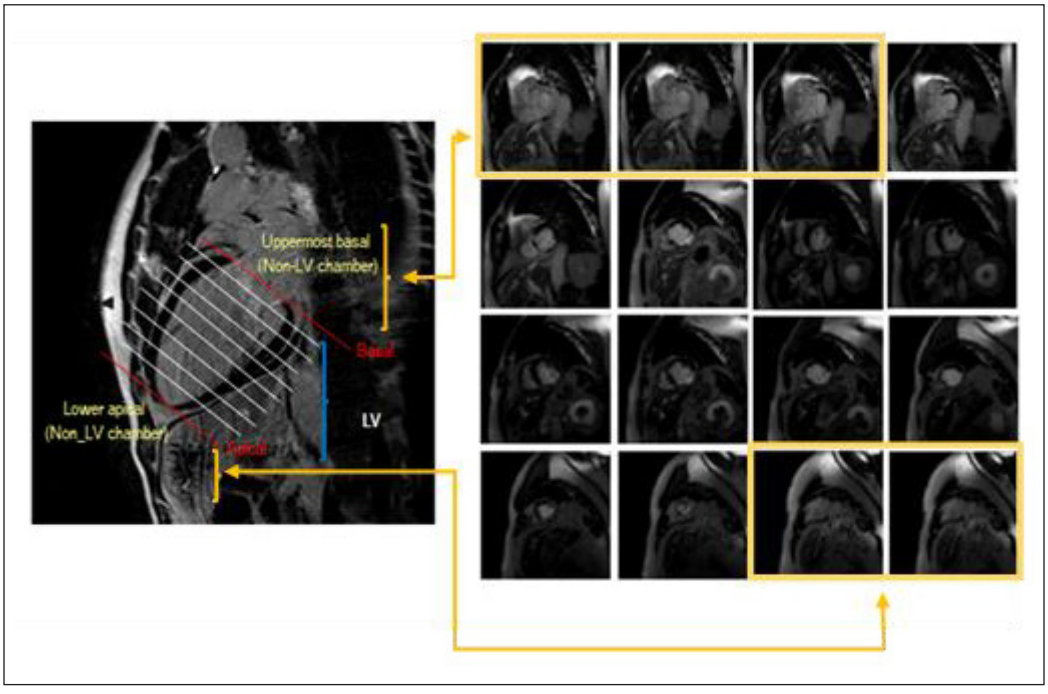


Figure 1. Example of De-CMR image sequences per patient. Uppermost basal and lower apical images comprise the non-LV chamber class

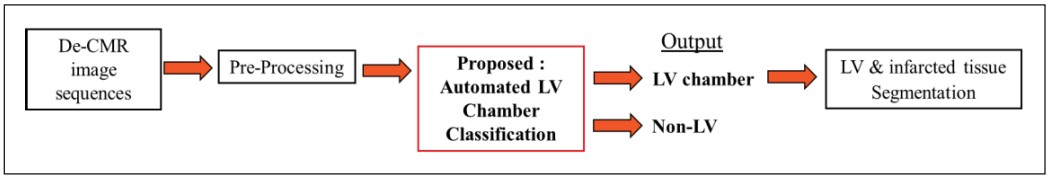


Figure 2. The Proposed automated LV chamber classification in overall workflow for fully automated myocardial infarction diagnosis

features from the input data (Abdullah et al., 2023; Khairandish et al., 2022; Harron et al., 2022). In the cardiac study, Zhang et al. (2016) examined unnecessary CMR images using two distinct 5-layer CNN models by identifying a missing apical or basal slice. In another study, Ho and Kim (2021) explored the effectiveness of transfer learning for CINE CMR images. They evaluated 18 pre-trained Deep CNN models and three models trained from scratch. Among these, the VGG16 model achieved the highest accuracy with 86% across all evaluation categories, suggesting its potential as a suitable choice for cardiac classification. However, no additional study into CNN hyperparameters was presented as a result of those two studies. In contrast to the previous two investigations, which utilized the CINE Cardiac MRI modality, our study employs the De-CMR modality. Although pre-trained models are known for their efficiency and rapid deployment, there is a unique aspect to developing a custom CNN from scratch. By learning from a myriad of different data, pre-trained models can become quite complicated and hard to understand. Conversely, the process of designing a CNN model from scratch offers a straightforward and clear, making it easier to grasp and adaptable to necessary modifications. Moreover, the performance of CNNs can be further enhanced by incorporating additional techniques and optimizations.

Batch normalization is one such technique that has garnered attention in recent years. The batch normalization layer helps reduce dependence on specific weight initializations and acts as a regularizer, preventing overfitting and promoting generalization to unseen data. However, the impact of the batch normalization layer on CNNs varies depending on their arrangement and configurations. Hasani and Khotanlou (2019) investigated the placement of batch normalization in the network using three well-known pre-trained networks on three image datasets (CIFAR10, Cifar100 and Tiny Imaget). Since the results vary for various networks, their study proposed that tests should be conducted on various networks and image datasets to determine the ideal configuration. There is ongoing debate in the field of machine learning regarding the optimal placement of an additional layer in a network to achieve the greatest acceleration in the training process (Garbin et al., 2020; Laurent et al., 2016). Further investigation in this study is motivated by the advantages of the batch normalization layer.

Accordingly, the primary objective of this study is to enhance the performance of our new classification model for segregating images with LV chambers from De-CMR images using the CNN classification model from scratch. To the best of our knowledge, the influence of incorporating a batch normalization layer and its arrangement on the performance of our image classification tasks has not been thoroughly studied and analyzed in the training CNNs. Furthermore, these enhancements form the foundation for the key contributions of our paper:

- Novel Classification Model: We present a new shallow convolutional neural network model specifically designed to automate the classification of LV and non-

LV chambers in De-CMR image slices, aiming to improve myocardial infarction assessment.

- Optimal Network Configuration: Through extensive experimentation, we identified the best configuration of batch normalization layers and learning rates across different optimizers, ensuring optimal performance in LV chamber classification.
- Enhanced Workflow for Radiologists: Our proposed model significantly improves radiologists' workflow by providing an efficient and automated solution, reducing the need for manual selection of LV chamber slices in De-CMR for myocardial infarction diagnosis.

METHODOLOGY

Data Preparation

This research utilized datasets from the De-CMR modality chosen at random from the Radiology Unit of the Imaging Department, Advanced Medical and Dental Institute. Approval from the Human Research Ethics Committee of USM was necessary to access this private dataset. All selected images, originally in Digital Imaging and Communications in Medicine (DICOM) format, were anonymized and then converted to bitmap (bmp) format. To ensure a balanced dataset for each class, the non-LV chamber images, which were fewer in number compared to the LV chamber images, were augmented using geometric transformation techniques. These techniques included rotations of +90, +180, and +270 degrees. A total of 1420 images were used in this study. All experimental network models were trained and validated on 80% (1136 images) (Bichri et al., 2024) of the data using five-fold cross-validation (Hou et al., 2024). This method involves splitting the training data into five folds, using four folds for training and one for validation in each iteration. This process is repeated five times, ensuring all data is used for both training and validation, leading to a more robust estimate of model performance. We reserved the remaining 20% (284 images) of the total dataset to avoid evaluation bias as a separate test set. These unseen images were not included in the training process and were used solely for model test evaluation.

Network Architecture

CNN's most essential layer is the convolution layer, which extracts features. Ghosh et al. (2020) summarize the convolution layer's operating principle. CNN's weights are learned from back-propagation, except for the initial condition when given randomly. The three convolution layers depicted in Figure 3 of our proposed CNN architecture from scratch were generated experimentally, as described in our prior research by Awang Damit et al. (2022). However, no additional experiments or discussions regarding the use of batch

normalization were presented in our previous study. Therefore, the current work aims to fill this gap by focusing on the inclusion of batch normalization in the model and thoroughly investigating its effects through experimental and in-depth discussions.

Certain values, such as the number of layers and filters, were selected as the acceptable accuracy for each dataset due to the limited computational resources. The network was kept simple and reliable to fulfill the next research objective: to hybrid with another network model to automatically segment the infarcted tissue in the LV chamber. MATLAB software was used to specify the network parameters. The proposed model input layer expects input images with a resolution of 227 × 227 pixels. Kandel and Castelli (2020) state that higher learning rates require bigger batch sizes. Thus, a fixed learning rate of 0.0001, 16 batch size, and the Adam optimizer were used to train the first experiment regarding the presence of batch normalization layer because a greater learning rate is not advised in this study due to an insufficiently large dataset.

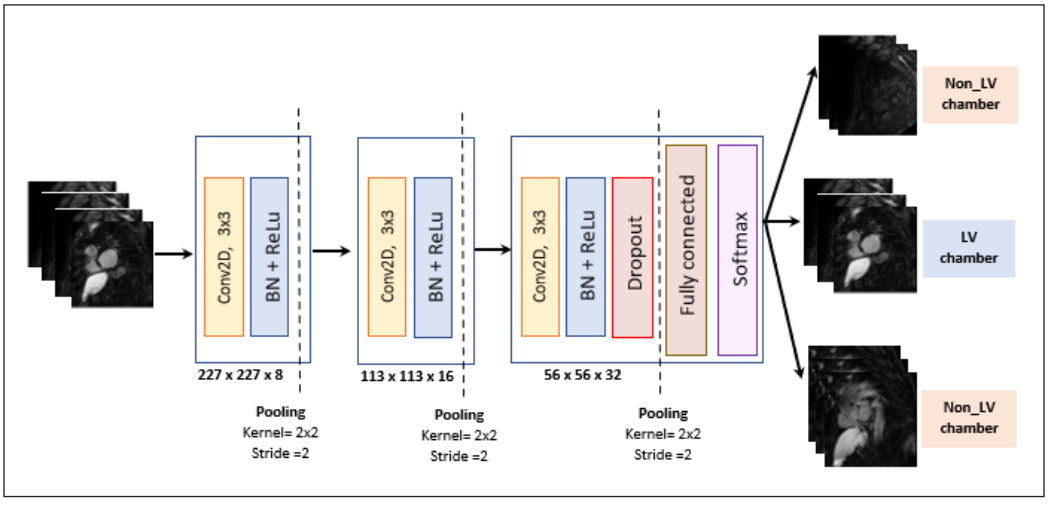


Figure 3. Network architecture for the proposed CNN model

Training Scenario with Batch Normalization

Three different arrangements of stacking the layers in our proposed CNN with the appearance of a batch normalization layer were systematically tested in this work. Figure 4 illustrates these three CNN models, one without batch normalization layer (A1) and the other with different batch normalization layer arrangements. The first arrangement of batch normalization layers was to place them after the convolution layer (A2). In contrast, the second arrangement placed them before the convolution layer, where it will normalize the input (A3). The test dataset was used to generate the final metrics after all models had been trained using the training dataset with the hyperparameters tuned by the validation

dataset. Adam optimizer was adopted as our chosen optimizer with an initial learning rate of 0.0001. The optimum performance model configuration shown in the previous experiment was then examined utilizing three learning rates from two different types of optimizers, Adam and Sgdm (Zohrevand & Imani, 2022). Furthermore, our best-proposed model with the optimal parameter configuration from this experiment was further evaluated against well-known models, including GoogleNet (Szegedy et al., 2015), SqueezeNet (Avanzato & Beritelli, 2023), and AlexNet (Krizhevsky et al., 2017). The same parameter settings and testing datasets were used across all models to ensure a fair comparison and avoid potential biases. The images were resized to match the required input size for each network to adapt these pre-trained models to our task. GoogleNet requires input images of size 224 × 224 × 3, whereas SqueezeNet and AlexNet require 227 × 227 × 3. The final layer of each network was fine-tuned by replacing it with a new layer for the two-class classification task of LV and non-LV chambers.

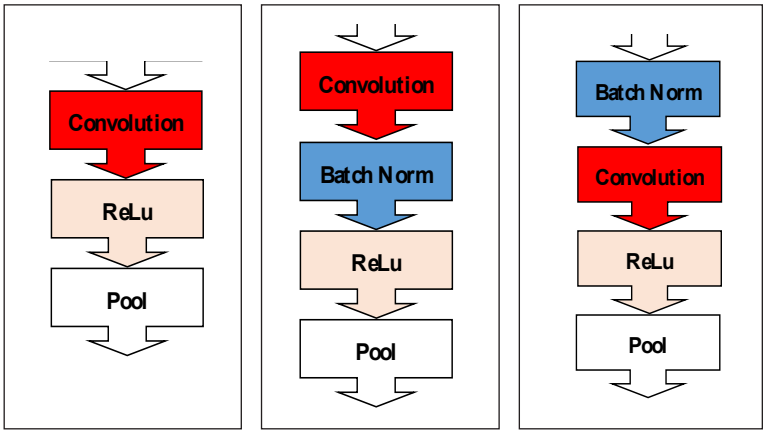


Figure 4. Illustration of the three CNN models: (a) Model A1 without a batch normalization layer, (b) Model A2 with batch normalization after the convolution layer, and (c) Model A3 with batch normalization before the convolution layer

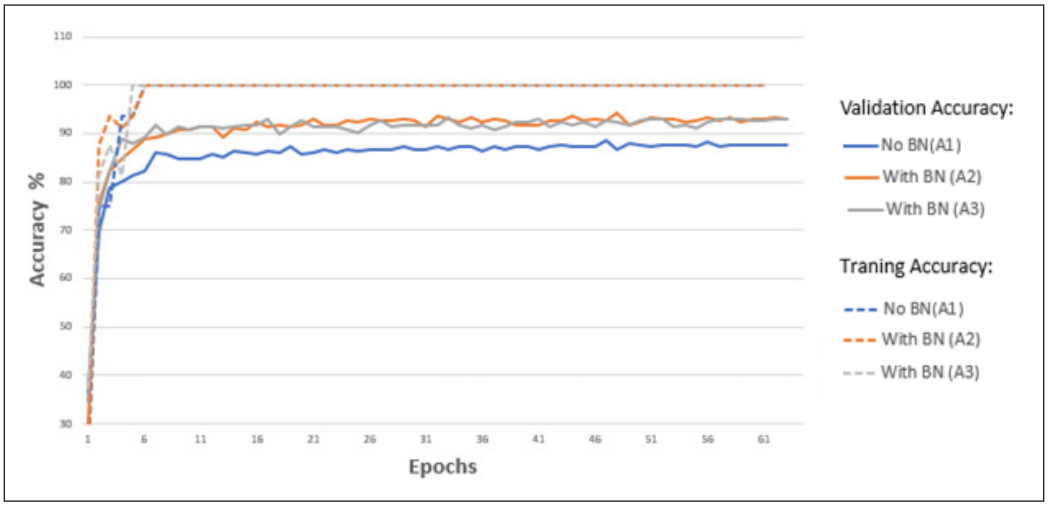
RESULTS

These experiments aimed to select the best layer architecture for the CNN model by utilizing a batch normalization layer for the De-CMR dataset.

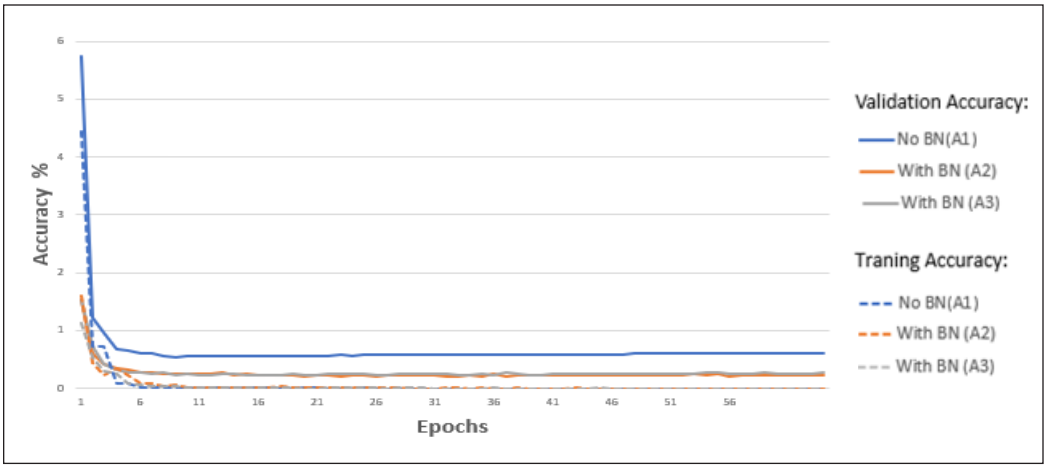
Results of Batch Normalization Implementation

The implications of the three distinct configurations mentioned earlier were compared to average accuracy performance. The model's validation and loss accuracy performance without a batch normalization layer and with two distinct batch normalization

layer arrangements converged in the line graphs of Figures 5(a) and 5(b). The training was stopped when no substantial gain was observed in accuracy after 64 epochs. The steady decrease in loss and consistent increase in accuracy, as depicted in the plots, provide strong evidence that the models were able to learn from the dataset effectively. While the validation accuracy follows a similar trend, none of the models show signs of overfitting. Figures 6(a) and 6(b) provide a detailed view of the convergence curves for training and validation accuracy, emphasizing the initial stages of the training process. The blue line represents accuracy over arrangement A1, while the orange and grey lines represent accuracy over arrangement A2 and A3, respectively. Both models with a batch normalization layer, A2 and A3, demonstrate an improved learning speed performance with a faster training



(a)

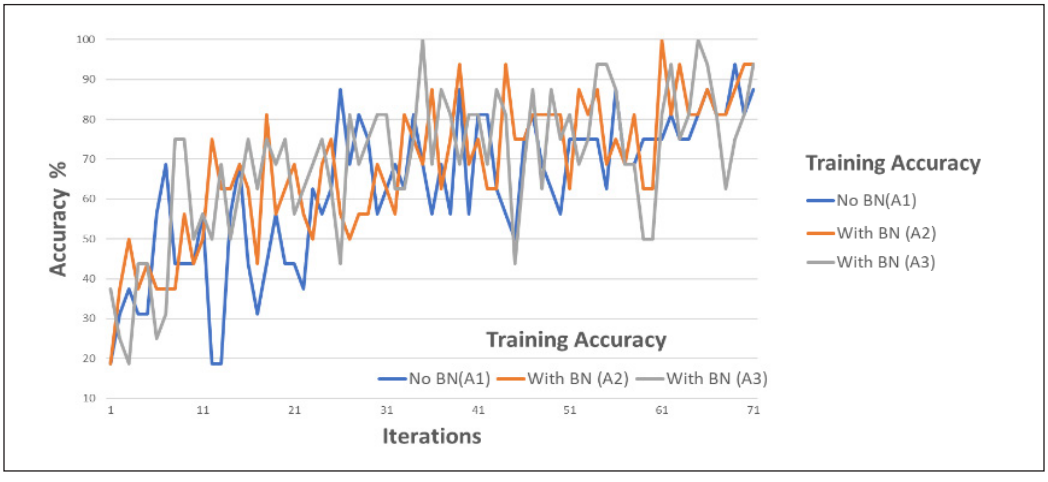


(b)

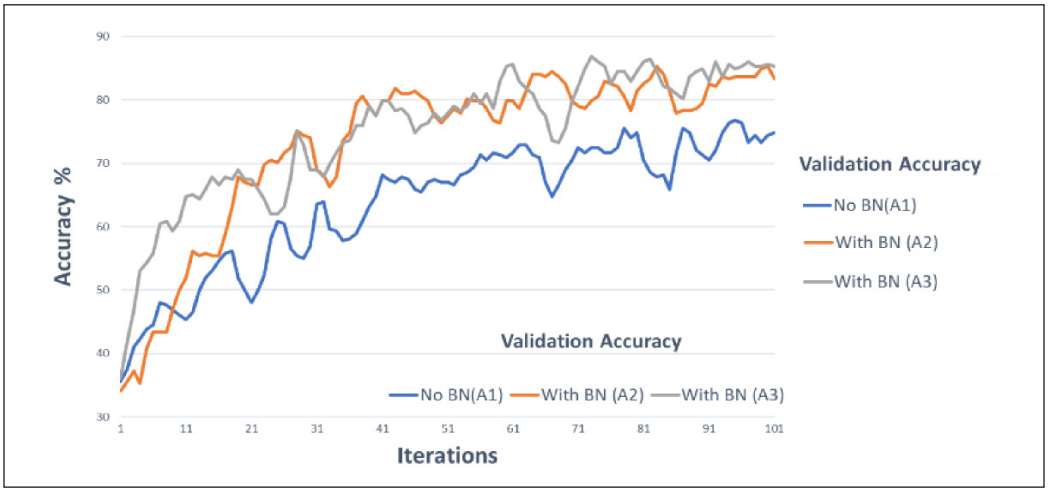
Figure 5. Plots of training and validation (a) accuracy and (b) loss for three CNN models in 64 epochs

and validation convergence curve than model A1. Additionally, model configuration A3 exhibits slightly faster convergence than A2. These results confirm that incorporating batch normalization slightly improves accuracy performance and leads to faster model learning compared to a model without batch normalization.

Table 1 presents the validation and testing accuracy performance of models A1, A2, and A3. Validation accuracy reflects the average testing accuracy achieved during the training process using five-fold cross-validation. In contrast, testing accuracy refers to the model's performance on a completely new set of images not included in the training process. This metric is crucial for understanding how well the model generalizes to unseen data. All models were trained using a learning rate 0.0001 and the Adam optimizer. Notably, A2



(a)



(b)

Figure 6. Plots of (a) training and (b) validation accuracy for three CNN models on early training convergence

stands out as the highest performer in terms of validation and test accuracy, achieving an exceptional 93.41% and 88.4%, respectively. A3 follows closely in testing accuracy with a slight difference of 1.4%. Conversely, the model without batch normalization, A1, exhibits the lowest performance, with a difference of 5.3% compared to A2.

Further investigating the performance of each model, Table 2 showcases the precision, recall, and F1-score metrics for the above three CNNs model configurations. Overall, both models with batch normalization, A2 and A3, significantly outperform the model without batch normalization, A1, across all three metrics, highlighting the positive impact of batch normalization on the model's ability to accurately classify the left ventricle segment. A2 achieves the highest F1-score (0.8773), indicating a well-balanced performance between precision (0.9291) and recall (0.8310). While A3 exhibits a slightly lower F1-score (0.8614), its precision remains above average (0.9200). These findings suggest that both batch normalization arrangements enhance the model's ability to correctly identify true positives while minimizing false positives, leading to a more accurate classification of the left ventricle segment. Further validation of these findings is provided by the confusion matrices presented in Figure 7, which compares the predicted classifications for all three CNN models (7(a) A1, 7(b) A2, and 7(c) A3) to the actual classifications for the two categories, namely LV and non-LV. It is important to note that the testing set contains an equal number of De-CMR images for both LV and non-LV categories (142 images each, resulting in a total of 284). This ensures a balanced representation of both classes. The first predicted row of each confusion matrix shows the number of correctly identified images: Model A1: 110 LV images and 126 non-LV images, Model A2: 118 LV images and 133 non-LV images, and Model A3: 115 LV images and 132 non-LV images. These values demonstrate that all models correctly identified a significant portion of the images. Notably, Figure 7 also

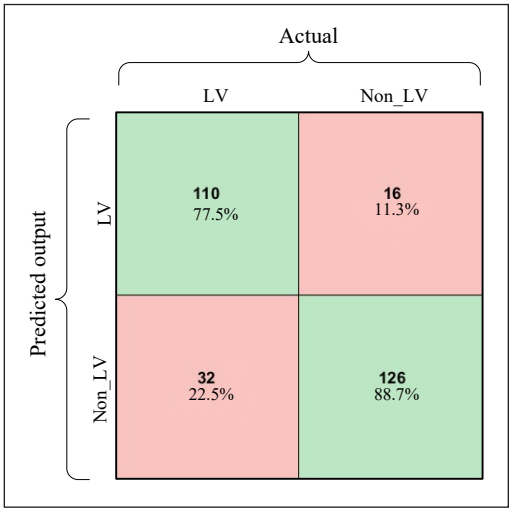
Table 1
Validation and test accuracy result and comparison of three different batch normalization arrangements with LR=0.0001

CNN Model	Validation Accuracy (%)	Test Accuracy (%)
A1 (No BN)	87.6	83.1
A2 (With BN)	**93.4**1	**88.4**
A3 (With BN)	92.64	87

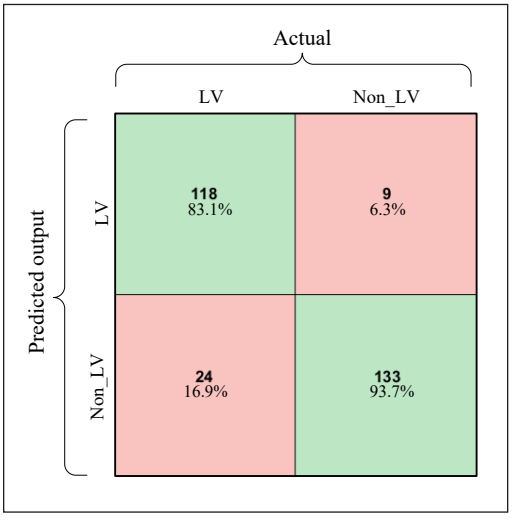
Table 2
Precision, recall and F1-Score performance of three different batch normalization arrangements with LR=0.0001

CNN Model	Precision	Recall	F1-Score
A1 (No BN)	0.8730	0.7746	0.8209
A2 (With BN)	**0.9291**	**0.8310**	**0.8773**
A3 (With BN)	0.9200	0.8099	0.8614

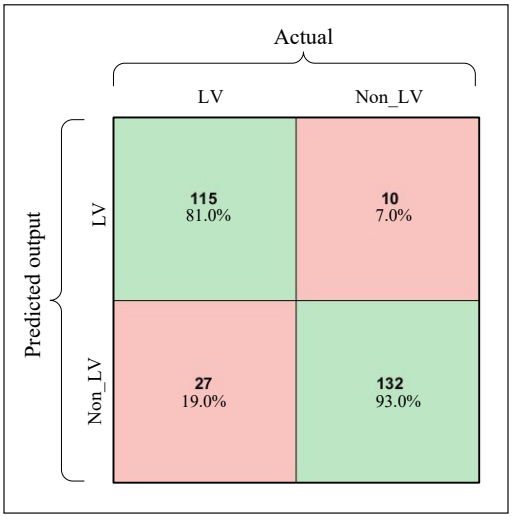
reveals a consistent pattern across all models: the number of false negatives and incorrectly classified positive images is higher than the number of false positives and incorrectly classified negative images. This is also proved by the data shown in the confusion matrix in Figure 7. The True Positives in the confusion matrix represent the number of correctly classified LV chamber images, while the True Negatives represent the correctly classified non-LV chamber images. The LV and non-LV De-CMR images are given equally in the testing set as 142 and 142 give the total of 284 of the testing set. In the first predicted row for all models, 110(77.5%), 118(83.1%) and 115(81%) out of 142 De-CMR images with LV chamber were correctly identified in models A1, A2 and A3, respectively and 126(88.7%),133(93.7%) and 132(93%) out of 142 De-CMR images with non- LV chamber were correctly identified in model A1, A2 and A3, respectively. These values demonstrate that all models correctly identified a significant percentage of the images. Notably, Figure 7 also reveals a consistent pattern across all models, where the number of false negatives (which are incorrectly classified LV images) is higher than the number of false positives (the incorrectly classified non-LV images). This



(a)



(b)



(c)

Figure 7. Testing confusion matrix for CNN Model (a) A1 (b) A2 and (c) A3

phenomenon can be partially attributed to the anatomical similarity between the utmost apical and basal regions of the LV chamber with the surrounding non-LV image, particularly in cases where the boundaries are unclear or the image quality is poor.

Results of Batch Normalization Implementation in Different Learning Rates

Figure 8 illustrates the impact of incorporating batch normalization and different learning rates on the performance of the proposed CNN models. It presents the accuracy achieved for each model configuration with the Adam and Sgdm optimizers across three different learning rates: 0.0001, 0.001, and 0.01. Across all learning rates and optimizers, models with batch normalization consistently achieve higher accuracy compared to models without batch normalization. This confirms the effectiveness of batch normalization in stabilizing the training process and improving the overall performance of the CNN models. The optimal learning rate appears to be 0.0001 for both Adam and Sgdm optimizers. At this learning rate, models with batch normalization achieve their highest accuracy at 88.4% with Adam and 86.1% with Sgdm. Using higher learning rates as 0.001 and 0.01 generally leads to a decline in accuracy for both models with and without batch normalization. This suggests that the chosen network architecture can be more sensitive to larger learning rates, potentially leading to instability and suboptimal performance. While both Adam and Sgdm optimizers benefit from batch normalization, Adam seems to achieve slightly higher accuracy with batch normalization across all learning rates. This indicates that the Adam optimizer is better suited for this specific network architecture and task.

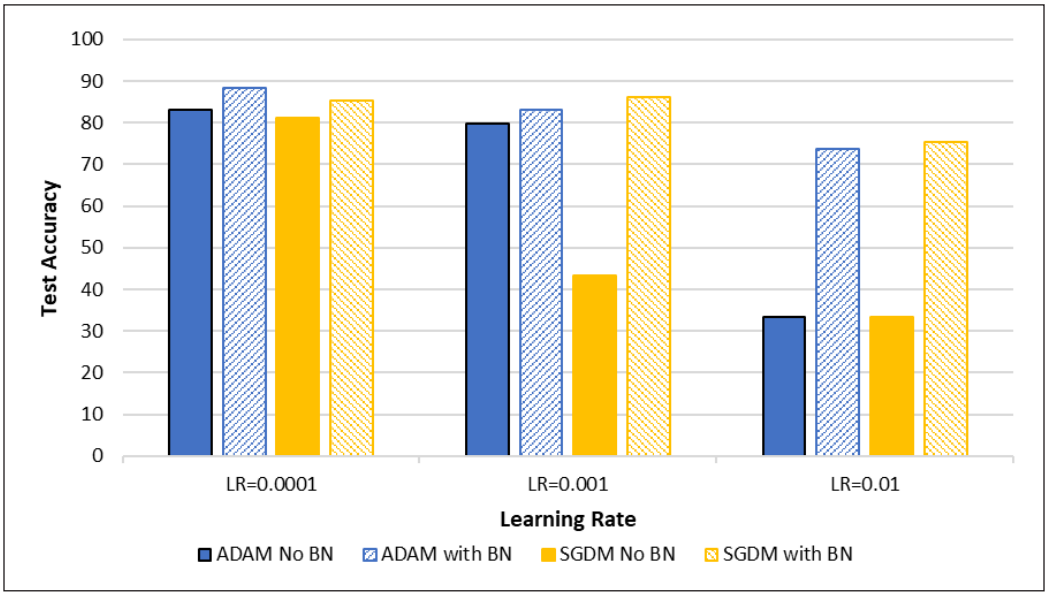


Figure 8. Accuracy performance for test data on Adam and Sgdm optimizers with different learning rates

DISCUSSION

This study investigated the performance of a newly proposed CNN model architecture, specifically focusing on the influence of batch normalization layer placement and learning rate on its ability to automatically classify images containing the LV chamber and non-LV chamber regions in De-CMR modality images. Our results demonstrated that incorporating a batch normalization layer after each convolution led to the highest accuracy compared to models without batch normalization or with batch normalization before the convolution layer. Although arrangement A3 accelerates the learning ability more quickly than arrangement A2, A2 eventually generates higher accuracy than A3. Additionally, experimentation with three different learning rates (0.0001, 0.001, 0.01) and two optimizers (Adam and Sgdm) identified 0.0001 as the optimal learning rate when using the Adam optimizer. Models trained with this lower learning rate demonstrated better generalization compared to those trained with larger rates (0.001 and 0.01), where a notable decline in accuracy was observed due to instability and suboptimal convergence. The lower learning rate allowed for more gradual and stable updates during gradient descent, preventing large oscillations in the loss function, which can occur with higher learning rates. The significant improvement in accuracy observed with batch normalization further highlights its importance. Through its ability to mitigate internal covariate shifts, batch normalization contributes to preventing overfitting by ensuring more consistent gradient updates. This ultimately leads to faster convergence and enhanced generalization performance. While the study focused solely on Adam and Sgdm optimizers, it is important to note that both these optimizers benefit from batch normalization. Exploring other optimization algorithms might reveal further insights. These findings highlight the importance of batch normalization in optimizing CNN architectures to improve accuracy in medical image analysis, leading to a future fully automated diagnosis system.

Based on the investigation, we evaluate our best-proposed model, A2, against AlexNet, SqueezeNet, and GoogleNet, which are well-established architectures. The accuracy metric is used to evaluate the performance of each model. Our proposed model achieves an accuracy of 88.4%, which slightly outperforms AlexNet (87%) but is lower than GoogleNet (92.96%) and SqueezeNet (92.61%) (Table 3). While our model exhibits slightly lower accuracy compared to pre-trained models like SqueezeNet and GoogleNet, it presents significant advantages in terms of simplicity and computational efficiency. Its less complex architecture translates to faster training times and lower resource requirements, making it a more practical choice for deployment in resource-constrained environments. Additionally, the model leverages automated feature extraction, eliminating the need for manual screening, which can be time-consuming. Notably, unlike pre-trained models that require fixed input image sizes, the proposed model offers greater adaptability to different image size datasets and potential future applications without the need for image resizing,

Table 3
Comparison of the proposed CNN with established pre-trained CNNs architectures

Author	CNN Model	Accuracy (%)
This study	A2 (Proposed)	88.4
Krizhevsky et al., 2017	AlexNet	87
Avanzato et al.,2023	SqueezeNet	92.61
Szegedy et al., 2015	GoogleNet	92.96

which can introduce noise and potentially degrade performance. Therefore, its simplicity, efficiency, and flexibility strengths make our proposed model a compelling alternative to pre-trained models, particularly in situations where computational resources are critical.

CONCLUSION AND RECOMMENDATIONS

In this study, we have examined the result of adopting the batch normalization layer in different positions of CNN models for LV chamber and non-LV chamber classification in De-CMR images as one of the pre-processing procedures in myocardial infarction assessment. This study aims to identify the optimal network layer configuration and learning rate combination for De-CMR image classification that maximizes accuracy while maintaining computational efficiency. The results demonstrate that incorporating a batch normalization layer after the convolution layer in arrangement A2 significantly improves the network's accuracy (88.4%) compared to models without batch normalization (83.1%) or with different layer placements. Additionally, the study reveals that using a lower learning rate (0.0001) with the Adam optimizer yields the best performance. While the proposed model achieves competitive accuracy compared to established architectures like AlexNet, this paper suggests more investigation into this model's behavior by increasing the training data volume for the diversity and complexity of medical conditions. Additionally, exploring the effectiveness of other architectures like inception modules, skip connections, or attention mechanisms could lead to further performance improvements. Overall, this study presents a promising initial approach for fully automated myocardial infarction assessment through LV chamber classification using a simple and efficient CNN model.

ACKNOWLEDGMENT

This research was funded by the Ministry of Higher Education through the Fundamental Research Grant Scheme (FRGS) with code FRGS/1/2023/SKK06/UITM/02/12, titled "New convolutional neural networks - autoencoder model with fusion correlation layer for left ventricle classification and scar tissue segmentation in cardiac magnetic resonance images of myocardial infarction disease." Ethical approval was granted on November 22nd, 2022, by the Human Research Ethics Committee of USM (JEPeM), Universiti Sains Malaysia

(USM/JEPeM/21090623). 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. Finally, the authors are grateful to Universiti Teknologi MARA, Cawangan Pulau Pinang, for their immense administrative support.

REFERENCES

Abdullah, M. F., Sulaiman, S. N., Osman, M. K., Karim, N. K. A., Setumin, S., Isa, I. S., & Ani, A. I. C. (2023). Geometrical feature of lung lesion identification using computed tomography scan images. *Jurnal Teknologi*, *85*(2), 149–156. https://doi.org/10.11113/JURNALTEKNOLOGI.V85.18828

Abdullah, M. F., Sulaiman, S. N., Osman, M. K., Karim, N. K. A., Shuaib, I. L., & Alhamdu, M. D. I. (2020). Classification of lung cancer stages from CT scan images using image processing and k-Nearest neighbours. In *2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC)* (pp. 68-72). IEEE Publication. https://doi.org/10.1109/ICSGRC49013.2020.9232492

Avanzato, R., & Beritelli, F. (2023). Thorax disease classification based on the convolutional network squeezenet. In *2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (Vol. 1, pp. 933-937). IEEE Publishing. https://doi.org/10.1109/IDAACS58523.2023.10348691.

Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the impact of train / test split ratio on the performance of pre-trained models with custom datasets. *International Journal of Advanced Computer Science and Applications, 15*(2), 331-339. https://doi.org/10.14569/IJACSA.2024.0150235

Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., & Rueckert, D. (2020). Deep learning for cardiac image segmentation: A review. *Frontiers in Cardiovascular Medicine*, *7*, Article 25. https://doi.org/10.3389/fcvm.2020.00025

Damit, D. S. A., Sulaiman, S. N., Osman, M. K., A. Karim, N. K., & Setumin, S. (2022). Classification of left ventricle and non- left ventricle segment for cardiac assessment using deep convolutional neural network. *Journal of Electrical & Electronic Systems Research*, *21*, 31–38. https://doi.org/10.24191/jeesr.v21i1.005

Garbin, C., Zhu, X., & Marques, O. (2020). Dropout vs. batch normalization: An empirical study of their impact to deep learning. *Multimedia Tools and Applications*, *79*(19), 12777–12815. https://doi.org/10.1007/s11042-019-08453-9

Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., & De, D. (2020). Fundamental concepts of convolutional neural network. In V. E. Balas, R. Kumar, & R. Srivastava (Eds.), *Recent Trends and Advances in Artificial Intelligence and Internet of Things* (pp. 519–567). Springer International Publishing. https://doi.org/10.1007/978-3-030-32644-9_36

Harron, N. A., Sulaiman, S. N., Osman, M. K., Isa, I. S., A. Karim, N. K., & Maruzuki, M. I. F. (2022). Deep learning approach for blur detection of digital breast tomosynthesis images. *Journal of Electrical & Electronic Systems Research, 21*, 39–44. https://doi.org/10.24191/jeesr.v21i1.006

Hasani, M., & Khotanlou, H. (2019). An empirical study on position of the batch normalization layer in convolutional neural networks. In *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)* (pp. 1-4). IEEE Publishing. https://doi.org/10.1109/ICSPIS48872.2019.9066113

Ho, N., & Kim, Y. C. (2021). Evaluation of transfer learning in deep convolutional neural network models for cardiac short axis slice classification. *Scientific Reports*, *11*(1), Article 1839. https://doi.org/10.1038/s41598-021-81525-9

Hou, R., Lo, J. Y., Marks, J. R., Hwang, E. S., & Grimm, L. J. (2024). Classification performance bias between training and test sets in a limited mammography dataset. *PLOS ONE, 19*(2), Article e0282402. https://doi.org/10.1371/journal.pone.0282402

Huang, L., Zhao, P., Tang, D., Zhu, T., Han, R., Zhan, C., Liu, W., Zeng, H., Tao, Q., & Xia, L. (2020). Cardiac involvement in patients recovered from COVID-2019 identified using magnetic resonance imaging. *JACC: Cardiovascular Imaging*, *13*(11), 2330–2339. https://doi.org/10.1016/J.JCMG.2020.05.004

Im, D. J., Hong, S. J., Park, E. A., Kim, E. Y., Jo, Y., Kim, J., Park, C. H., Yong, H. S., Lee, J. W., Hur, J. H., Yang, D. H., & Lee, B. Y. (2019). Guidelines for cardiovascular magnetic resonance imaging from the Korean society of cardiovascular imaging - Part 3: Perfusion, delayed enhancement, and T1- and T2 mapping. *Korean Journal of Radiology*, *20*(12), 1562–1582. https://doi.org/10.3348/kjr.2019.0411

Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, *6*(4), 312–315. https://doi.org/10.1016/j.icte.2020.04.010

Karim, R., Bhagirath, P., Claus, P., Housden, R. J., Chen, Z., Karimaghaloo, Z., Sohn, H. M., Rodríguez, L. L., Vera, S., Albà, X., Hennemuth, A., Peitgen, H. O., Arbel, T., Gonzàlez Ballester, M. A., Frangi, A. F., Götte, M., Razavi, R., Schaeffter, T., & Rhode, K. (2016). Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images. *Medical Image Analysis*, *30*, 95–107. https://doi.org/10.1016/j.media.2016.01.004

Khairandish, M. O., Sharma, M., Jain, V., Chatterjee, J. M., & Jhanjhi, N. Z. (2022). A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. *IRBM*, *43*(4), 290–299. https://doi.org/10.1016/j.irbm.2021.06.003

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386

Laurent, C., Pereyra, G., Brakel, P., Zhang, Y., & Bengio, Y. (2016). Batch normalized recurrent neural networks. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2657-2661). IEEE Publishing. https://doi.org/10.1109/ICASSP.2016.7472159

Leong, C. O., Lim, E., Tan, L. K., Aziz, Y. F. A., Sridhar, G. S., Socrates, D., Chee, K. H., Lee, Z. V., & Liew, Y. M. (2019). Segmentation of left ventricle in late gadolinium enhanced MRI through 2D-4D registration for infarct localization in 3D patient-specific left ventricular model. *Magnetic Resonance in Medicine*, *81*(2), 1385–1398. https://doi.org/10.1002/mrm.27486

Li, G., Zhang, W., Jia, D., Rong, J., Yu, Z., & Wu, D. (2023). Epidemic of the SARS-CoV-2 Omicron variant in Shanghai, China in 2022: Transient and persistent effects on Out-of-hospital cardiac arrests. *Resuscitation*, *186*, Article 109722. https://doi.org/10.1016/j.resuscitation.2023.109722

Shaaf, Z. F., Jamil, M. M. A., Ambar, R., Alattab, A. A., Yahya, A. A., & Asiri, Y. (2022). Automatic left ventricle segmentation from short-axis cardiac mri images based on fully convolutional neural network. *Diagnostics*, *12*(2), Article 0414. https://doi.org/10.3390/diagnostics12020414

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9). IEEE Publishing. https://doi.org/10.1109/CVPR.2015.7298594

Toro, A., Bozzani, A., Tavazzi, G., Urtis, M., Giuliani, L., Pizzoccheri, R., Aliberti, F., Fergnani, V., & Arbustini, E. (2021). Long COVID: long-term effects? *European Heart Journal Supplements*, *23*, E1–E5. https://doi.org/10.1093/eurheartj/suab080

World Health Organization. (2021). *Cardiovascular Diseases (CVDs)*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Zhang, L., Gooya, A., Dong, B., Hua, R., Petersen, S. E., Medrano-Gracia, P., & Frangi, A. F. (2016). Automated quality assessment of cardiac MR images using convolutional neural networks. In S. A. Tsaftaris, A. Gooya, A. F. Frangi, & J. L. Prince (Eds.), *Simulation and Synthesis in Medical Imaging* (pp. 138–145). Springer International Publishing.

Zohrevand, A., & Imani, Z. (2022). An empirical study of the performance of different optimizers in the deep neural networks. In *2022 International Conference on Machine Vision and Image Processing (MVIP)* (pp. 1-5). IEEE Publishing. https://doi.org/10.1109/MVIP53647.2022.9738743

Zorkafli, M. F., Osman, M. K., Isa, I. S., Ahmad, F., & Sulaiman, S. N. (2019). Classification of cervical cancer using hybrid multi-layered perceptron network trained by genetic algorithm. *Procedia Computer Science, 163*, 494–501. https://doi.org/10.1016/j.procs.2019.12.132