A Deep Learning Approach to Classify Fabry Cardiomyopathy from Hypertrophic Cardiomyopathy Using Cine Imaging on Cardiac Magnetic Resonance

Wei-Wen Chen, Ling Kuo, Yi-Xun Lin, Wen-Chung Yu, Chien-Chao Tseng, Yenn-Jiang Lin, Ching-Chun Huang, Shih-Lin Chang, Jacky Chung-Hao Wu, Chun-Ku Chen, Ching-Yao Weng, Siwa Chan, Wei-Wen Lin, Yu-Cheng Hsieh, Ming-Chih Lin, Yun-Ching Fu, Tsung Chen, Shih-Ann Chen, and Henry Horng-Shing Lu

1Institute of Computer Science and Engineering, National Yang-Ming University, Hsinchu, Taiwan
2Faculty of Medicine and Institute of Clinical Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan
3Division of Cardiology, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan
4Department of Biomedical Imaging and Radiological Sciences, National Yang-Ming University, Taipei, Taiwan
5Institute of Statistics, National Yang Ming Chiao Tung University, Hsinchu, Taiwan
6Department of Radiology, Taipei Veterans General Hospital, Taipei, Taiwan
7Department of Radiology, Taichung Veterans General Hospital, Taichung, Taiwan
8Department of Post-Baccalaureate Medicine, National Chung Hsing University, Taichung, Taiwan
9Cardiovascular Center, Taichung Veterans General Hospital, Taichung, Taiwan
10Department of Pediatric Cardiology, Taichung Veterans General Hospital, Taichung, Taiwan
11Children’s Medical Center, Taichung Veterans General Hospital, Taichung, Taiwan
12Department of Pediatrics, School of Medicine, National Chung-Hsing University, Taichung, Taiwan
13College of Medicine, National Chung Hsing University, Taichung, Taiwan
14Department of Statistics and Data Science, Cornell University, Ithaca, New York, USA

Correspondence should be addressed to Shih-Ann Chen; epsachen@ms41.hinet.net and Henry Horng-Shing Lu; henryhslu@nycu.edu.tw

Received 1 August 2023; Revised 20 March 2024; Accepted 23 March 2024; Published 26 April 2024

A challenge in accurately identifying and classifying left ventricular hypertrophy (LVH) is distinguishing it from hypertrophic cardiomyopathy (HCM) and Fabry disease. The reliance on imaging techniques often requires the expertise of multiple specialists, including cardiologists, radiologists, and geneticists. This variability in the interpretation and classification of LVH leads to inconsistent diagnoses. LVH, HCM, and Fabry cardiomyopathy can be differentiated using T1 mapping on cardiac magnetic resonance imaging (MRI). However, differentiation between HCM and Fabry cardiomyopathy using echocardiography or MRI cine images is challenging for cardiologists. Our proposed system named the MRI short-axis view left ventricular hypertrophy classifier (MSLVC) is a high-accuracy standardized imaging classification model developed using AI and trained on MRI short-axis (SAX) view cine images to distinguish between HCM and Fabry disease. The model achieved impressive performance, with an F1-score of 0.846, an accuracy of 0.909, and an AUC of 0.914 when tested on the Taipei Veterans General Hospital (TVGH) dataset. Additionally, a single-blinding study and external testing using data from the Taichung Veterans General Hospital (TCVGH) demonstrated the reliability and effectiveness of the model, achieving an F1-score of 0.727, an accuracy of 0.806, and an AUC of 0.918, demonstrating the model’s reliability and usefulness. This AI model holds promise as a valuable tool for assisting specialists in diagnosing LVH diseases.
1. Introduction

LVH is a common finding in transthoracic echocardiography (TTE) in clinical practice and is often associated with poor cardiovascular outcomes and ventricular arrhythmias [1]. LVH can be caused by arterial hypertension, aortic stenosis, HCM, Fabry disease, or cardiac amyloidosis. However, distinguishing between these different etiologies can be challenging based on morphological features alone. Cardiac magnetic resonance imaging (MRI) can help differentiate LVH caused by different cardiomyopathies through tissue characterization of the myocardium. For instance, native T1 mapping and extracellular volume (ECV) mapping can effectively differentiate Fabry cardiomyopathy from cardiac amyloidosis or HCM, with a decrease in native T1 value in Fabry cardiomyopathy and a significant increase in T1 value and ECV in cardiac amyloidosis [2–4]. Among various cardiomyopathies with LVH, HCM is the most prevalent, with a reported incidence of 1:500 (0.2%) [5, 6]. In Taiwan, Fabry disease is particularly interesting due to its incidence of 1/1600 in male newborns, with 82% having a specific gene mutation of IVS4+919G>A [7]. Fabry disease with this mutation is associated with a late-onset cardiac phenotype, with LVH developing in the midfifties [8]. Enzyme replacement therapy is available to treat the deficiency of galactosidase-A enzyme activity in Fabry disease [9]. Nowadays, cardiac MRI is recommended for early identification of Fabry cardiomyopathy from HCM, which is clinically beneficial for timely managements [5, 10]. However, there is limited availability of parametric mapping on cardiac MRI worldwide, and the standard ranges of native T1 values can vary between different MRI scanners [11]. In this study, we aim to use deep learning techniques to differentiate HCM from Fabry cardiomyopathy by analyzing cine images using a universal cardiovascular magnetic resonance sequence in the absence of myocardial tissue characterization. There are challenges associated with training the model, such as variable depth size and frame rate in the data and the need for large GPU memory [12, 13]. Nevertheless, there have been significant advancements in deep learning techniques for medical tasks and domains [13–18], and 3D residual neural networks (3D ResNet) are effective in analyzing 3D medical images, including cardiac MRI [19, 20]. A previous study has used CNN after segmenting SAX cine images to distinguish HCM from other hypertrophic mimics, including cardiac amyloidosis, Fabry disease, and hypertensive hypertrophy [21]. This study will develop the machine learning model to classify HCM from Fabry cardiomyopathy without segmentation of SAX cine images.

1.2. Cardiac MRI Acquisition Protocol. Cardiac MRI was performed on 1.5T scanner (GE Optima MR450w, GE Healthcare, Waukesha, Wisconsin, USA) and on 3T scanner (Discovery MR750, GE Healthcare, Waukesha, Wisconsin, USA). Each study applied a cardiac phased array receiver surface coil and ECG gating. Cine images were obtained by using a steady-state free precession sequence (echo time (TE): 1.2–1.6 ms, repetition time (TR): 3.2–3.6 ms) in a stack of 8 mm thick short-axis slices encompassing the whole ventricles after gadolinium injection and in long-axis slice. On the other hand, patients with device utilized gradient echo sequence to minimize artifact. The late gadolinium enhancement image acquisition occurred 10–15 minutes after intravenous administration of 0.15 mmol/kg and 0.10 mmol/kg gadobutrol (Gadovist, Bayer, Germany) for 1.5T and 3.0T scanner, respectively, using an inversion-recovery gradient-echo pulse sequence with individually adjusted inversion time to optimize nulling of normal myocardium (typical TI: 310–380 ms) [26]. Field of view was typically set to 300 mm × 300 mm (may varied depending on patient size), typical voxel size of the images was 1.6 × 2.0 × 8 mm, the TE was 3.1–3.5 ms, and the TR was 6.2–7.6 ms for 1.5T scanner and TE 2.5–3.1 ms and TR 5.46.6 ms for 3T scanner. All patients had obtained informed consent for cardiac MRI, and patient with cardiac implantable electronic device was monitored according to standard procedure [27].

1.3. Statistical Analysis. We conducted a statistical analysis to demonstrate the baseline characteristics and LV parameters of study cohort. Continuous variables are expressed as the mean ± SD, and categorical variables are expressed as counts with percentages. Independent sample t-test and chi-square test were used to compare normally distributed...
continuous and categorical variables. All statistical analyses were performed using SPSS version 24 (SPSS Inc., Chicago, IL, USA). Statistical significance was set at two-tailed \( P < 0.05 \). All statistical analyses were performed using SPSS version 24 (SPSS Inc., Chicago, IL, USA). Statistical significance was set at two-tailed \( P < 0.05 \).

2. Materials and Methods

2.1. Data Sources. The research study was conducted under the approval of the Institutional Review Board (IRB) at Taipei Veterans General Hospital (TVGH) and Taichung Veterans General Hospital (TCVGH). The TVGH and TCVGH datasets comprised digital imaging and communications in medicine (DICOM)-formatted images acquired using GE Healthcare Systems’ magnetic resonance imaging (MRI) technology. Specifically, the study focused on the short-axis (SAX) view cine images obtained from MRI scans. We collected 215 MRI images, including 156 images of patients with Fabry disease and 59 images of patients with hypertrophic cardiomyopathies. To evaluate our model’s accuracy and generalization ability, we randomly split the data into a training set of 80% and a testing set of 20%. Additionally, we collected 31 MRI images of patients manufactured by Siemens Healthcare Systems from TCVGH to validate our model’s generalization ability. All images used in the study had either \( 512 \times 512 \) or \( 256 \times 256 \) pixels in size. The TVGH SAX view cine images consisted of 8 mm or 10 mm thick slices with frame rates of 20, 30, or 40. Please refer to Figure 1 for a visual representation of the external verification process.

2.2. Implementation Details and Data Preprocessing. In the SAX view, cardiac images contain four dimensions—height, width, depth, and time. Since each patient’s heart size and heartbeat vary during the MRI scan, the resulting images have varying slice numbers and frame rates. We adopt several preprocessing steps to standardize the data configurations, including pixel sizes, frame rates, and slice numbers [13]. Firstly, we resize the images to \( 256 \times 256 \) using bilinear interpolation [28]. Subsequently, we crop the images between 50 and 200 to remove noncardiac organs. To preserve the cardiac cycle, we downsample the sampling rate to a frame rate of 20 using the Fourier method [29].

Next, we select five slices, including the middle and two on either side. While we could consider taking more slices, such as three on each side, this approach could capture unwanted areas, such as the atrium, or miss the heart, leading to a higher computational burden and lower information gain. Finally, we stack the time and depth dimensions together, resulting in a video with a shape of \( 100 \times 150 \times 150 \). We apply min-max normalization to the resulting video to rescale the values between 0 and 1 [30].

2.3. Designed Models. For this study, we divided the dataset into 80% for training and 20% for testing. To select the best hyperparameters for our model, we employed stratified 5-fold cross-validation [31]. We evaluated the \( F_1 \)-score metric to identify the optimal mean validation \( F_1 \)-score and epoch. Subsequently, we trained a final model using 80% of the training dataset and the identified optimal hyperparameters. Additionally, we calculated various metrics based on the mean and standard deviation of the 5-fold cross-validation results [21].

2.4. Data Processing Flow. This model is aimed at identifying the direction of a video shot in MRI. Depending on the scanning process, video files may be obtained in either a ventricle-to-atrium or an atrium-to-ventricle direction, which can affect the development of the model. To address this issue, we developed a model that automatically identifies the order in which the images were collected, thus enabling us to standardize the order from ventricle to atrium. We utilized a 3D convolution neural network (3D CNN) structure, specifically the 3D ResNet18 model, as shown in Figure 2 [20]. Since we needed to perform two-class classification and the output of 3D ResNet18 is in 2048 dimensions, we connected four dense layers and dropout layers after each dense layer to avoid bottlenecks caused by extreme compression [32]. To prevent overfitting, we also incorporated dropout layers, L2 regularization, label smoothing, and augmentation techniques, such as random rotation from \( -\pi/6 \) to \( \pi/6 \), which effectively doubled our dataset by creating a reverse order. It is important to note that we rotated the video (i.e., all images must rotate at the same angle), as illustrated in Figure 3. We used class weight [33] to address the imbalance problem. Specifically, the TVGH image dataset had 153 instances of atrium-to-ventricle and 63 instances of ventricle-to-atrium. In comparison, the TCVGH image dataset had 13 instances of atrium-to-ventricle and 18 instances of ventricle-to-atrium. We used 5-fold cross-validation to identify the optimal hyperparameters and fine-tune the model. Ultimately, this model can differentiate the order of shot images and standardize the order of videos. It is worth noting that all the algorithms for LVH disease classification were implemented using the TensorFlow 2.10.0 deep learning framework on a desktop computer running the Linux Ubuntu operating system. The TensorFlow platform was deployed on a system equipped with an Intel i7-8700K processor and 62 GB of memory and an NVIDIA GTX-1080 Ti GPU boasting 12 GB of memory.

2.5. LVH Disease Classification Model. The model is aimed at distinguishing between Fabry disease and HCM disease. The model structure in Figure 2 is the same as the identity order model, except that the fully connected (FC) layer has some differences and does not perform reverse order augmentation. After standardizing the imaging direction for all images (from the ventricle to the atrium), we train the model. There are two challenges with this dataset.

The first is an imbalance problem, where Fabry disease has 156 patients and HCM has 59 patients. To address this issue, we use class weights calculated based on the samples in each class in the training set [33].

The second challenge is overfitting, a common issue in deep learning with small datasets. To avoid overfitting, we add dropout layers after dense layers. Additionally, we add norm two penalties to the ResNet18 and the fully connected
The dropout rate is set to 0.4, the fraction of input units to drop. The CNN’s L2 regularization factor (λ) is 0.00003, while the L2 regularization factor (λ) of the fully connected layer is 0.005. We set the label smoothing to 0.1, and our loss function is focal loss (where γ is 3) [34]. These methods help to focus on difficult-to-distinguish samples.

Furthermore, we use a learning rate of 0.00005, a batch size of 8, and the Adam optimizer. We train the model for 83 epochs, and these parameters are tuned using 5-fold cross-validation.

3. Experimental Results

3.1. Baseline Characteristics and LV Parameters. The baseline characteristics and LV parameters measured by echocardiography and MRI are shown in Table 1. In brief, in patients with LVH, there were no significant differences in
age, body weight, body surface area, and LV mass index either measured by echocardiography or cardiac MRI in patients with Fabry disease and HCM. However, patients with HCM had thicker interventricular septal wall thickness (IVST) measured by echocardiography compared with patients with Fabry disease. In patients without LVH but only focal wall thickness ≥ 13 mm, patients with Fabry disease were younger, male predominant, thinner IVST, and LV mass index either measured by echocardiography or cardiac MRI.

### 3.2. Results for Identifying Order Models.

We evaluated the model's performance using 5-fold cross-validation, testing data, and TCVGH externally validated data during the experiments. The classification results in Table 2 demonstrate high accuracy, providing confidence that the model effectively combines SAX view cine images from the ventricle to the atrium when preprocessed.

### 3.3. Analyze Unified Direction versus Nonunified Direction Training Models.

After analyzing Table 3, it was observed that unifying the video direction led to better performance than when it was nonunified. This performance improvement was observed in the 5-fold cross-validation and the TVGH testing set. Consequently, based on this result, the decision was made to unify the training video direction.

### 3.4. Differentiate between Fabry Cardiomyopathy and HCM.

In Table 4, we present the performance analysis of our experiment for the classification of Fabry cardiomyopathy and HCM. The evaluation was performed using 5-fold cross-validation, internal testing at TVGH, and external validation at TCVGH. We employed a 5-fold cross-validation technique to evaluate the performance of our model and identify the optimal hyperparameters based on the highest F1-score obtained. The best F1-score achieved was 0.861, and we utilized this model for training our final classifier for predicting the testing set. Our final model achieved an F1-score of 0.861, an accuracy of 0.909, and an AUC of 0.914. However, when it comes to external verification, we face significant challenges due to variations in the equipment, personnel, and practices employed by manufacturers, doctors, and technicians. In addition, our test was a single-blinded study, and we only had one opportunity to demonstrate that our model's performance was acceptable. For the external verification, our model achieved an F1-score of 0.727, an accuracy of 0.806, and an AUC of 0.918 when tested on data from TCVGH. Despite the challenges of external validation, our model's performance is promising and merits further investigation.

### Table 1: The baseline characteristics and LV parameters evaluated by echocardiography and MRI.

<table>
<thead>
<tr>
<th>Variables</th>
<th>LVH (n = 113)</th>
<th>HCM (n = 40)</th>
<th>P value</th>
<th>No LVH but focal wall thickness ≥ 13 mm (n = 102)</th>
<th>Fabry (n = 83)</th>
<th>HCM (n = 19)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yrs)</td>
<td>58.6 ± 10.0</td>
<td>61.1 ± 13.2</td>
<td>0.27</td>
<td></td>
<td>49.1 ± 13.3</td>
<td>56.4 ± 14.4</td>
<td>0.04</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>41 (56.2)</td>
<td>21 (52.5)</td>
<td>0.84</td>
<td></td>
<td>63 (75.6)</td>
<td>9 (47.4)</td>
<td>0.02</td>
</tr>
<tr>
<td>Body weight (kg)</td>
<td>64.6 ± 12.7</td>
<td>67.6 ± 13.5</td>
<td>0.24</td>
<td></td>
<td>67.9 ± 13.1</td>
<td>69.2 ± 11.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Body surface area (kg/m²)</td>
<td>1.70 ± 0.20</td>
<td>1.74 ± 0.22</td>
<td>0.27</td>
<td></td>
<td>1.76 ± 0.20</td>
<td>1.780 ± 0.19</td>
<td>0.64</td>
</tr>
<tr>
<td>IVST (mm)</td>
<td>14.6 ± 3.81</td>
<td>17.5 ± 4.12</td>
<td>&lt;0.001</td>
<td></td>
<td>10.7 ± 2.81</td>
<td>12.9 ± 3.64</td>
<td>0.008</td>
</tr>
<tr>
<td>LV mass index by TTE (g/m²)</td>
<td>163.4 ± 68.8</td>
<td>161.1 ± 38.1</td>
<td>0.84</td>
<td></td>
<td>84.3 ± 14.8</td>
<td>96.5 ± 11.7</td>
<td>0.003</td>
</tr>
<tr>
<td>LV mass index by MRI (g/m²)</td>
<td>80.2 ± 28.7</td>
<td>78.6 ± 19.9</td>
<td>0.57</td>
<td></td>
<td>53.2 ± 13.6</td>
<td>66.8 ± 10.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LVEF by TTE (%)</td>
<td>63.3 ± 9.10</td>
<td>64.3 ± 8.09</td>
<td>0.76</td>
<td></td>
<td>66.1 ± 6.59</td>
<td>66.4 ± 10.6</td>
<td>0.91</td>
</tr>
<tr>
<td>LVEF by MRI (%)</td>
<td>54.2 ± 8.28</td>
<td>57.0 ± 7.91</td>
<td>0.08</td>
<td></td>
<td>56.8 ± 5.02</td>
<td>60.4 ± 5.33</td>
<td>0.01</td>
</tr>
</tbody>
</table>

IVST: interventricular septal wall thickness; LV: left ventricular; LVEF: left ventricular ejection fraction; MRI: magnetic resonance imaging; TTE: transthoracic echocardiography.

### Table 2: The F1-score and accuracy results for 5-fold cross-validation, TVGH testing, and TCVGH external validation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>F1-score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-fold</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>TVGH</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>TCVGH</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### Table 3: Accuracy and F1-score were compared between unified and nonunified directions using five-fold cross-validation and TVGH testing.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Direction</th>
<th>F1-score</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-fold</td>
<td>Unified</td>
<td>0.861</td>
<td>0.924</td>
</tr>
<tr>
<td>5-fold</td>
<td>Nonunified</td>
<td>0.857</td>
<td>0.924</td>
</tr>
<tr>
<td>TVGH</td>
<td>Unified</td>
<td>0.846</td>
<td>0.909</td>
</tr>
<tr>
<td>TVGH</td>
<td>Nonunified</td>
<td>0.769</td>
<td>0.864</td>
</tr>
</tbody>
</table>

### Table 4: The F1-score, accuracy, sensitivity, specificity, and AUC were compared on 5-fold cross-validation, TVGH testing data, and TCVGH external validation data.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>F1-score</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-fold</td>
<td>0.861</td>
<td>0.924</td>
<td>0.895</td>
<td>0.992</td>
<td>0.967</td>
</tr>
<tr>
<td>TVGH</td>
<td>0.846</td>
<td>0.909</td>
<td>0.917</td>
<td>0.906</td>
<td>0.914</td>
</tr>
<tr>
<td>TCVGH</td>
<td>0.727</td>
<td>0.806</td>
<td>0.727</td>
<td>0.850</td>
<td>0.918</td>
</tr>
</tbody>
</table>
Furthermore, a blinded reader from TCVGH (SW. Chan) interpreted 44 cardiac MRI SAX cine images, consisting of 32 patients with Fabry cardiomyopathy and 12 patients with HCM. The overall diagnostic accuracy was 54.5%. For patients with Fabry cardiomyopathy and HCM, the diagnostic accuracy was 37.5% and 66.7%, respectively. SAX cine images on cardiac MRI can only provide wall motion and morphologic features without tissue characteristics of myocardium, which makes the differentiation between Fabry cardiomyopathy from HCM challenging for human reader.

3.5. Model Visualization. Grad-CAM (gradient-weighted class activation mapping) [35] is a CNN visualization technique that enables the identification of model features and the avoidance of shortcuts to determine the model’s accuracy. By presenting slices and frames from both cases, Grad-CAM highlights the regions where the model pays the most attention, which helps infer the model’s reliability. As shown in Figure 4, the model is focused only on the middle and the heart, with the red portions representing the areas of most significant attention. As a result, the model does not take shortcuts when identifying relevant features.

4. Discussions

MRI is frequently used for diagnosing various diseases. However, it is often difficult for clinicians to determine the specific type of ventricular hypertrophy solely based on MRI images. A comprehensive assessment involving medical history, electrocardiogram (ECG), echocardiography, cardiac MRI, and cardiac stress testing is usually required to classify the ventricular hypertrophy type accurately. This process consumes healthcare resources and time. It would be exciting news if we could achieve the same level of accuracy using only one source of imaging data through the robust feature engineering and learning abilities of artificial intelligence (AI). A prior study [36–38] compared our medical imaging research findings for cardiac diseases. Madani et al. [36] primarily focused on the diagnosis of cardiac diseases using echocardiography images, while Zhou et al. [37] and Germain et al. [38] concentrated on the diagnosis of HCM and cardiac amyloidosis using cardiac cine images.

In contrast, our novel approach (MSLVHC) introduces a distinct method for classifying LVH specific to Fabry cardiomyopathy from HCM, utilizing MRI short-axis (SAX) view cine images. In recent years, medical imaging has made significant progress, with magnetic resonance imaging (MRI) being one of the most widely used techniques for diagnosing and monitoring various diseases. As the complexity and quantity of MRI data continue to grow, there is an increasing need for efficient and user-friendly tools to assist healthcare professionals in analyzing and interpreting these images.

4.1. Novel Deep Learning Approaches for Cardiac Disease Diagnosis. In this section, we compare the findings of our study with the referenced literature to provide the results and contributions of our research. The studies referenced cover a range of topics in medical imaging, including cardiac disease diagnosis, HCM mutation prediction, and cardiac amyloidosis diagnosis. The four kinds of research all present novel approaches to diagnosing cardiac diseases using deep learning. Madani et al. [36] focus on the diagnosis of cardiac disease using echocardiography images, while Zhou et al. [37] and Germain et al. [38] focus on the diagnosis of HCM and cardiac amyloidosis, respectively, using cardiac cine images. Our approach (MSLVHC) presents a new method to classify LVH for Fabry cardiomyopathy from HCM using MRI short-axis (SAX) view cine images. Table 5 compares the four approaches based on data sources, dataset, technology (AI method), model design, results, advantages, and disadvantages.

The effective handling of 3D cardiac MRI data is a crucial aspect in advancing the field of medical imaging, and each of the discussed papers provides unique insights into the application of deep learning for cardiac analysis. In comparing these approaches, it is essential to highlight the strengths and weaknesses of each method and the possible implications for future research and clinical application. Firstly, Madani et al. [36] address the limited labeled medical imaging data challenge through data-efficient deep learning models, specifically semisupervised GANs. The strengths lie in their data efficiency, segmentation for improved classification, and a comprehensive comparative analysis of model architectures. However, the limitations of small sample size and challenges with transfer learning underscore the need for further exploration and validation on larger, more diverse datasets.

Zhou et al. [37] introduce an innovative approach for predicting HCM genetic mutations using deep learning algorithms on cardiovascular magnetic resonance data. Their strengths include a reasonably sized and diverse dataset, comparison with established scores, and enhanced predictive performance when combined with clinical scores. Nevertheless, limitations such as a limited dataset and single-center data raise concerns about the model’s generalizability and broader utility in cardiac analysis.
There is a limitation in our study. In our study cohort, we included half of patients without LVH defined by LV mass index measured at papillary muscle level of parasternal short-axis view on transthoracic echocardiography according to the American Society of Echocardiography guideline [25] but only exhibited with focal wall thickness ≥ 13 mm, which is also reported as LVH. This is the clinical scenario that patients with Fabry disease or HCM can exhibit similar hypertrophic subtypes as concentric or asymmetric ventricular hypertrophy. In this subgroup, patients with HCM had thicker IVST and LV mass index compared with those with Fabry disease, which might constitute a bias for classification and assist our AI model to yield the high accuracy of the proposed mitigation plan, involving collecting additional data from diverse regions and testing on datasets from multiple sources, is crucial for addressing this limitation and enhancing the generalizability of the proposed method.

There is a limitation in our study. In our study cohort, we included half of patients without LVH defined by LV mass index measured at papillary muscle level of parasternal short-axis view on transthoracic echocardiography according to the American Society of Echocardiography guideline [25] but only exhibited with focal wall thickness ≥ 13 mm, which is also reported as LVH. This is the clinical scenario that patients with Fabry disease or HCM can exhibit similar hypertrophic subtypes as concentric or asymmetric ventricular hypertrophy. In this subgroup, patients with HCM had thicker IVST and LV mass index compared with those with Fabry disease, which might constitute a bias for classification and assist our AI model to yield the high accuracy of the proposed mitigation plan, involving collecting additional data from diverse regions and testing on datasets from multiple sources, is crucial for addressing this limitation and enhancing the generalizability of the proposed method.

There is a limitation in our study. In our study cohort, we included half of patients without LVH defined by LV mass index measured at papillary muscle level of parasternal short-axis view on transthoracic echocardiography according to the American Society of Echocardiography guideline [25] but only exhibited with focal wall thickness ≥ 13 mm, which is also reported as LVH. This is the clinical scenario that patients with Fabry disease or HCM can exhibit similar hypertrophic subtypes as concentric or asymmetric ventricular hypertrophy. In this subgroup, patients with HCM had thicker IVST and LV mass index compared with those with Fabry disease, which might constitute a bias for classification and assist our AI model to yield the high accuracy of the proposed mitigation plan, involving collecting additional data from diverse regions and testing on datasets from multiple sources, is crucial for addressing this limitation and enhancing the generalizability of the proposed method.

5. Limitations

There are limitations in this study. First, this is a single-center, retrospective cohort study with limited patient datasets and external validation data. The small study cohort may impact the generalizability of our model. Second, there are
various hypertrophic phenotypes in patients with HCM or Fabry cardiomyopathy. Patients who exhibited unexplained focal LV wall thickening, apical hypertrophy, or asymmetrical hypertrophy might not meet the criteria to define LVH from the American Society of Echocardiography guidelines. However, our study cohort presented various real-world types of LV hypertrophy, which makes our model applicable in daily practice. Furthermore, we recognize the potential confounding effect of the distribution of patients among different magnetic resonance scanners. As discussed in the article by Kushol et al. [39], acknowledging the significance of scanner bias is crucial. Even though we have utilized cross-validation to enhance the robustness of our model in hopes of reducing the bias introduced by different scanners, the potential bias cannot be completely eliminated.

6. Conclusion

The study introduces an AI-driven MRI analysis approach, MSLVHC, distinguishing between HCM and Fabry disease with impressive internal validation results (F1-score 0.846, accuracy 0.909, AUC 0.914). External validation at Taichung Veterans General Hospital showed promising performance (F1-score 0.727, accuracy 0.806, AUC 0.918). Challenges include diverse data acquisition, which will be addressed by promoting automated techniques for DICOM files and collaboration with data providers. Ongoing research is aimed at refining models for early detection of LVH disease, potentially improving patient outcomes and healthcare decision-making.

Data Availability

The MRI SAX view cine image dataset that supporting this study is provided by Taipei Veterans General Hospital (TVGH) and Taichung Veterans General Hospital (TCVGH), but there are restrictions on the availability of these data, which were used under license for the current study, and so are not publicly available. This dataset is a restricted-access dataset. To access the files, you must be a credentialed user and sign the data use agreement for the project at Taipei Veterans General Hospital.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

WW. Chen and L. Kuo contributed equally to this work.

Acknowledgments

This work was supported in part by the Taipei Veterans General Hospital, Taiwan, under Grant no. V113B-002; in part by the Ministry of Science and Technology, Taiwan, under Grant nos. MOST 110-2811-M-A49-550-MY2, MOST 110-2118-M-A49-002-MY3, and MOST 110-2634-F-A49-005-SP5-2; in part by the National Science and Technology Council, Taiwan, under Grant nos. NSTC 110-2634-F-A49-005 and NSTC 111-2634-F-A49-014; in part by the Higher Education Sprout Project of the National Yang Ming Chiao Tung University and Ministry of Education, Taiwan; and in part by the Ministry of Education Yushan Scholar Program, Taiwan. The authors would like to thank Wan-Yi Tai for her great help and support. We are grateful to the National Center for High-Performance Computing for providing computing resources.

References


