Improving the Generalizability of Cardiac MRI Segmentation Models Across Multi-Center and Multi-Vendor Datasets

부산대학교 정보컴퓨터공학부

지도교수: 감진규

팀명: Cyber

202055633 이슬람 살리흐

202155631 케네스 예라슬

202155629 누가예바 알트나이

Content

Table of Contents

Project Overview

Introduction and Problem Background

The advent of deep learning has significantly improved the state-of-the-art in cardiac magnetic resonance (CMR) segmentation. Numerous techniques have been introduced in recent years, pushing the accuracy of automated segmentation to near-human levels. However, these models are frequently trained and validated on cardiac imaging samples from individual clinical centers or uniform imaging protocols. Thus, Campello et al. [1] states that this practice has hindered the creation and validation of models that can generalize effectively across various clinical centers, imaging conditions, and scanner vendors.

Another literature indicates that deep learning techniques have achieved or exceeded previous benchmarks in numerous cardiac segmentation tasks. This progress has been fueled by the availability of larger public datasets, advancements in network architectures, and the availability of high-performance computing hardware. Despite these advancements, questions arise regarding the direct applicability of deep learning methods in real-world clinical settings to alleviate clinician workload. DL-based methods commonly face the challenge of limited generalization when confronted with new or unseen samples, such as data from unfamiliar scanners or cases not represented in the training set. In essence, these models are often influenced by the characteristics of their training data, restricting their applicability in real-world scenarios and reducing their potential to enhance clinical workflows. Existing research [2] suggests that there is still considerable room for improvement in this regard.

Objectives

This assignment project aims to achieve the following objectives:

1. Assess the Generalization Capability of Deep Learning Models: Evaluate the extent to which existing deep learning models for cardiac MRI segmentation can generalize across diverse datasets, including those from different scanners, imaging protocols, and clinical centers.
2. Identify Factors Affecting Model Generalization: Investigate the key factors that contribute to or hinder the generalization capability of deep learning models in cardiac MRI segmentation, such as dataset diversity, model architecture, and training strategies.
3. Develop Strategies to Enhance Model Generalization: Propose and explore novel techniques and approaches to improve the generalization performance of deep learning models for cardiac MRI segmentation, aiming to mitigate the effects of dataset bias and domain shift.
4. Evaluate the Performance on Unseen Data: Conduct rigorous evaluations to assess the performance of the developed models on previously unseen datasets, particularly focusing on their ability to accurately segment cardiac structures in varied clinical scenarios.
5. Contribute to the Advancement of Clinical Applications: Contribute insights and methodologies that can facilitate the deployment of deep learning models in real-world clinical settings, with the ultimate goal of enhancing diagnostic accuracy and efficiency in cardiac imaging workflows.

Requirements Analysis and Constraints

Requirements Analysis

Data Requirements: There should be an access to a wide range of cardiac MRI datasets from multiple clinical centers, featuring various imaging protocols and scanner types. For the purpose of proper training and model fitting, the dataset from M&M's Challenge [1] is considered to be taken, as it promises to be 375 heterogeneous CMR datasets acquired by using four different scanner vendors in six hospitals and three different countries (Spain, Canada and Germany). Such diverse properties are crucial since images provided by wide range of clinical centers and vendors may differ. For instance, in the Fig. 1 below, the visual apperance of a CMR short axis middle slice for anatomically similar subjects in the four different vendors is considered. Besides their anatomic similarity a clear differences can be observed due to diverse scanning vendors.

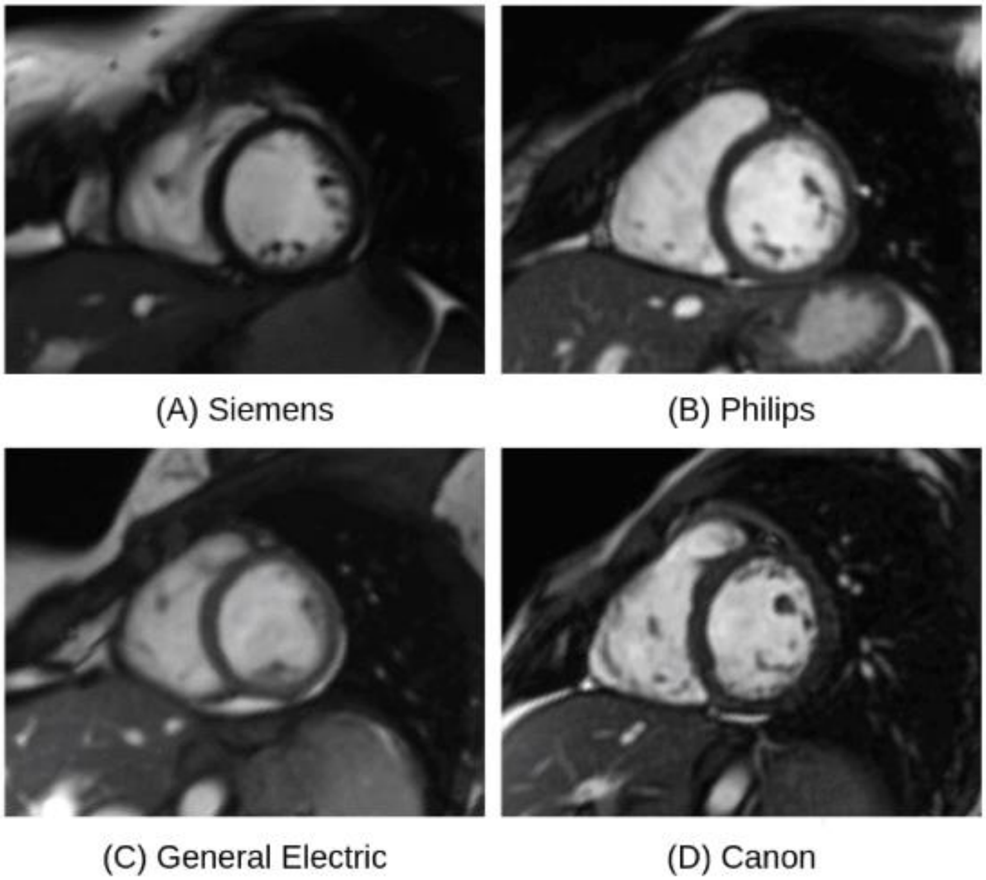

Figure 1. Images of similar cardiac slices from 4 different vendors

Another publicly available and extensively used dataset is from “Automatic Cardiac Diagnosis Challenge” (ACDC) [3]. However, the ACDC datasets were created using scans from 150 subjects at a single clinical center with a uniform imaging protocol. This restricts researchers' ability to develop and evaluate models that can generalize effectively across different centers and scanner vendors.

Training Requirements: Along with the heterogeneos dataset, techniques such as data augmentation and domain adaptation should be incorporated for enhancing model generalization.

- **Data Augmentation.** Data augmentation seeks to expand the size and diversity of training datasets by artificially creating new samples from existing labeled data. This is typically accomplished by applying various geometric or photometric transformations to the original image-label pairs [4]. Such transformations may include affine transformations, the addition of random noise, or adjustments to image contrast. Clear example could be observed in Fig. 2, where the top row displays the original image along with spatial augmentations, while the bottom row shows augmentations based on intensity adjustments.

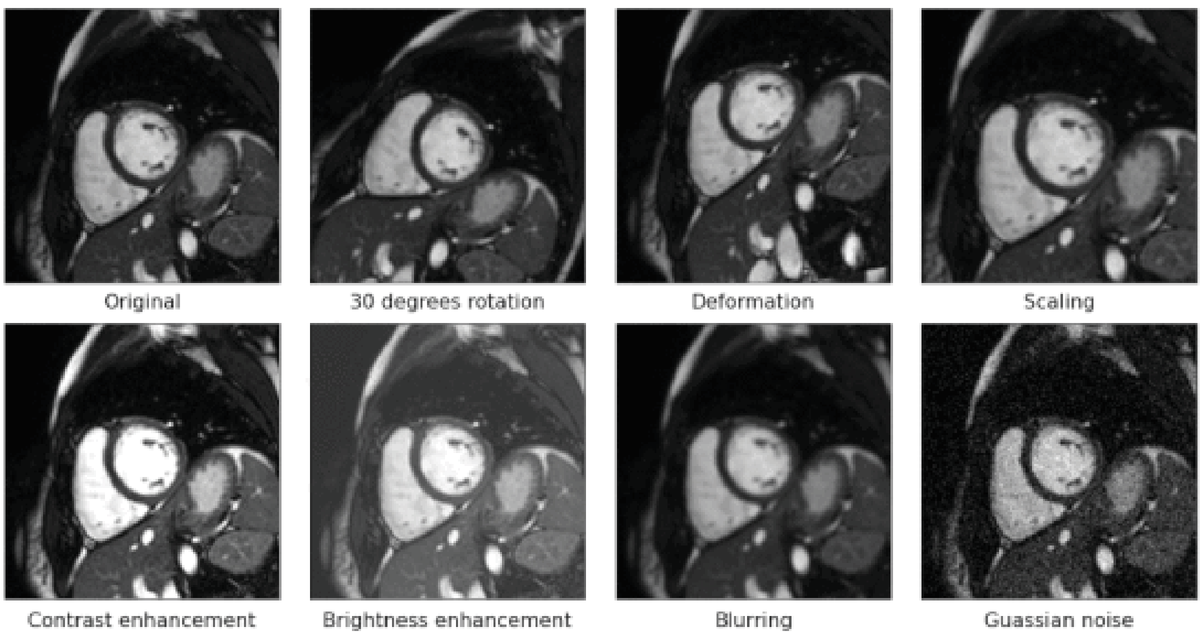

Figure 2. The impact of data augmentation on a single CMR slice

- **Domain Adaptation.** For the purpose of domain adaptation the technique introduced by Ganin et al. [5] is planned to be utilized. The **Domain Adversarial Training of Neural Networks** (DAAN) model includes a feature extractor network that integrates both a domain predictor and a label predictor. A gradient-reversal layer is situated between the feature extractor and the domain predictor; this layer reverses the gradient direction during backpropagation, which maximizes the domain prediction loss. This process minimizes the shift between domains while ensuring the model remains effective for the primary task of label prediction [6].

Constraints

- Data Constraints
  - Data Privacy and Security: Ensuring compliance with data protection regulations (e.g., GDPR, HIPAA) to maintain the confidentiality and security of patient data.
  - Limited Availability of Diverse Data: Potential scarcity of sufficiently diverse and annotated datasets, which may limit the training and validation process.
- Technical Constraints
  - Computational Limitations: High computational demands of training deep learning models may require significant investment in hardware and may be time-consuming.
  - Software Compatibility: Ensuring compatibility and integration of various software tools and libraries used in the project.

Research Directions and Implementation Plan

Research Directions

This study aims to explore several key research directions to improve the generalization of deep learning models for cardiac MRI segmentation:

- Domain Adaptation Techniques
  - Objective: Develop and evaluate domain adaptation methods to reduce the domain shift between training and unseen datasets.

    - Approach: Implement techniques such as domain adversarial neural networks (DANN), and domain-specific augmentation to enhance model robustness. Read related articles and research source code related to implementation.

- Data Augmentation Strategies
    - Objective: Increase the variability and size of training datasets through advanced data augmentation techniques.
    - Approach: Inspect publicly available source code and studies. Apply geometric and intensity-based augmentations, such as random rotations, affine transformations, noise addition, and contrast adjustments, to create more diverse training samples.
- Cross-Domain and Multi-Center Training
    - Objective: Train models on data from multiple clinical centers and imaging protocols to enhance generalization.
    - Approach: Aggregate datasets from M&Ms and M&Ms 2 data sources and implement cross-domain training strategies to create more robust models.

Implementation Plan

Segmentation Architecture: U-Net, V-Net, nnU-Net

- Required Tools: TensorFlow/Keras

Classification: ResNet, DenseNet, Inception

- Required Tools: TensorFlow/Keras, scikit-learn

Data Augmentation: Albumentations, imgaug, Torchvision

- Required Tools: OpenCV, SimpleITK

Data Visualization: Matplotlib, Seaborn, Plotly

Data Preprocessing: pandas, NumPy, scikit-learn

Model Deployment: Flask

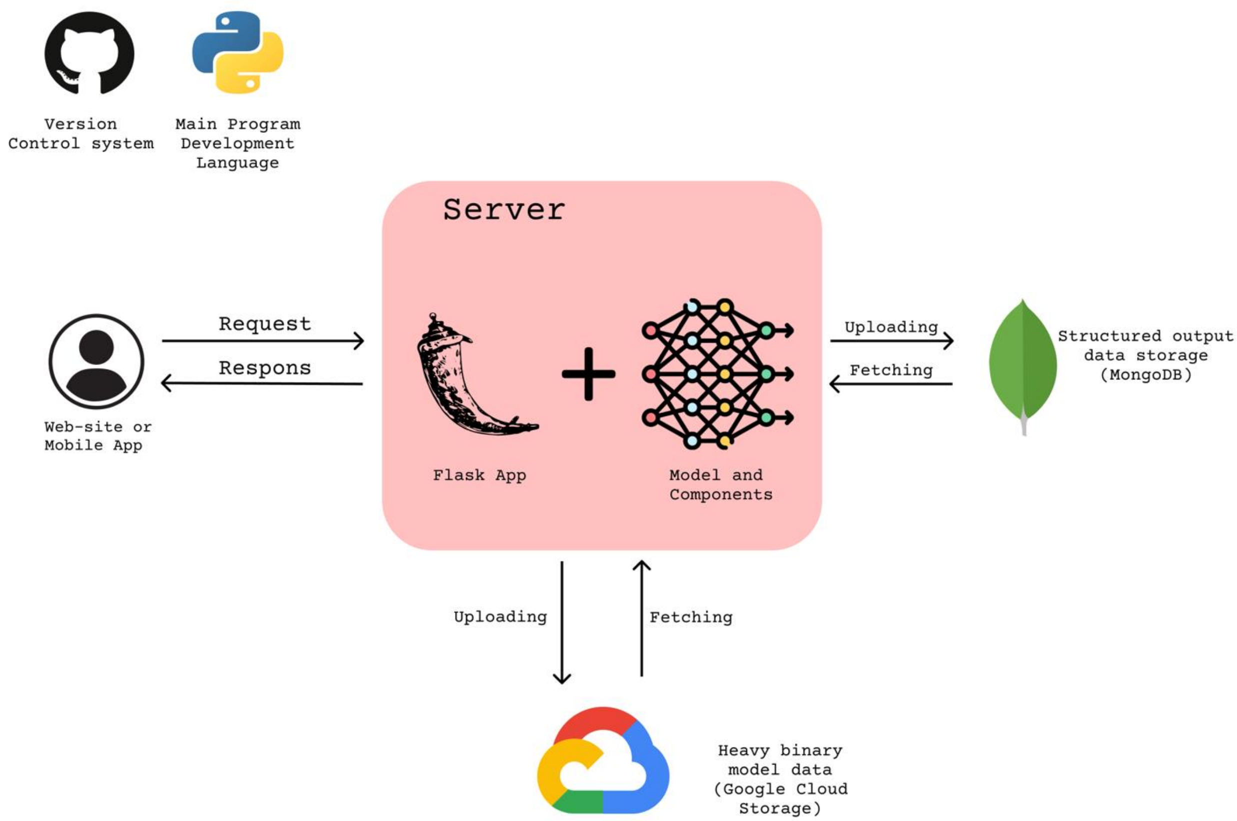

Figure 3. Workflow diagram

Development Schedule and Role Division

Development Schedule

구분	작업 일정																	
	5 월		6 월				7 월				8 월				9 월			
	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. **Literature Review and Research:** • Review research papers on cardiac MRI segmentation. • Summarize the information and identify the limitations of current approaches.																		
2. **Dataset Exploration and Analysis:** • Examine the dataset structure and its content.																		
3. **Data Preprocessing and Visualization:** • Apply data cleaning, normalization																		

and standardization. • Explore the cardiac MRI dimension and annotated segmentation labels using visualization tools.																		
4. **Model Development**: • Develop a new proposed model. • Train the developed model on the given dataset. • Evaluate the model performance. • Present the initial results and analyze the observations.																		
5. **Documentation and Presentation**: • Write and submit an intermediate report.																		
6. **Domain Adaptation for Generalizability**: • Research and implement domain adaptation techniques.																		

• Apply learning methods for increasing the robustness of the model.																		
7. **Model Evaluation and Analysis**: • Perform the quantitative model evaluation using various metrics. • Compute the qualitative evaluation of the model applying the visual inspection of earlier mentioned segmentation results.																		
8. **Application Integration**: • Creating the user interface for the application. • Front-end development. • Back-end development. • Integrate the trained model with the application. • Ensure the real-time performance.																		

9. **Application Testing and Debugging**: • Conduct the testing of the application. • Implement the necessary adjustments. • Execute performance testing of the application.																		
10. **Documentation and Presentation**: • Write and submit a final report. • Prepare a presentation and demonstration.																		

Role Division

이름	역할
이슬람 살리흐	• **Dataset Exploration and Analysis**: examine the multi-center dataset, its advantages and limitations.

	• **Data Preprocessing and Visualization:** data augmentation and sample generation using the visualization tools (applying affine transformations, the addition of random noise, adjustments to image contrast, etc.). • **Model Evaluation**: use different visual representation methods to assess the suggested method and its possible limitations.
케네스 예라슬	• **Application development**: integrating the trained model with the application, providing real-time results. • **Application Testing and Debugging:** validating the application's performance in terms of accuracy, robustness and processing speed.
누가예바 알트나이	• **Model Development**: develop and train a new suggested approach. • **Improving Generalizability**: applying domain adaptation techniques to improve model robustness. • **Model Evaluation**: Use different evaluation metrics to assess the performance and limitations of the suggested approach.
공통	• **Background study and Information gathering**: research the existing techniques and its challenges in cardiac MRI. • **Documentation**: writing reports, designing presentation and demonstration, defending the project.

References

1. V. M. Campello et al., "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge," in IEEE Transactions on Medical Imaging, vol. 40, no. 12, pp. 3543-3554, Dec. 2021, doi: 10.1109/TMI.2021.3090082.
2. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W and Rueckert D (2020) Deep Learning for Cardiac Image Segmentation: A Review. Front. Cardiovasc. Med. 7:25. doi: 10.3389/fcvm.2020.00025
3. O. Bernard et al., "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," in IEEE Transactions on Medical Imaging, vol. 37, no. 11, pp. 2514-2525, Nov. 2018, doi: 10.1109/TMI.2018.2837502.
4. Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*.
5. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of machine learning research*, *17*(59), 1-35.
6. Corral Acero, J., Sundaresan, V., Dinsdale, N., Grau, V., Jenkinson, M. (2021). A 2-Step Deep Learning Method with Domain Adaptation for Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Magnetic Resonance Segmentation. In: Puyol Anton, E., *et al.* Statistical Atlases and Computational Models of the Heart. M&Ms and EMIDEC

Challenges. STACOM 2020. Lecture Notes in Computer Science(), vol 12592. Springer,

Cham. https://doi.org/10.1007/978-3-030-68107-4_20