Chapter 8 Summary and future work

Cine and 4D flow cardiac MRI are two important non-invasive MR imaging techniques to assess cardiac function and diagnose cardiovascular diseases. Cine MRI offers great soft tissue detail which allows clinical experts to evaluate structure and function of the heart. 4D flow MRI further has the ability of three-dimensional time-resolved acquisition of blood flow velocity, which can be used to derive intra-cardiac hemodynamic parameters. In this thesis, we developed deep learning-based approaches to analyze cine and 4D flow cardiac MRI. In this chapter, we summarize the previous chapters and discuss potential directions of future work.

8.1 Summary

In **Chapter 1**, we provided a general introduction about cine and 4D flow cardiac MRI and deep learning applications in the field of cardiac MRI. In Chapter 2, we proposed a sampling inspection network combining specially designed data augmentation methods to assess CMR image quality. The proposed method showed a competitive performance against the other methods in the CMRxMotion challenge. In Chapter 3, we proposed temporal and spatial stacks to incorporate temporal or spatial information using stack attention mechanism for left ventricle segmentation in short-axis cine MRI. In Chapter 4, we further studied the concept of domain generalization in the setting of right ventricle segmentation in unseen datasets, such as data with differences in acquisition protocol, across different centers, scanner vendors and diseases. In Chapter 5, we investigated the feasibility of using deep learning-based approaches to segment the LV directly from 4D flow MRI and explored the performance of integrating features extracted from magnitude and velocity images. A transformer based feature fusion model was developed to improve the performance of LV segmentation from 4D flow MRI in Chapter 6. Chapter 7 aimed to train a CNN model to predict blood flow velocity from long-axis cine MRI using the corresponding 4D flow data as ground truth.

Chapter 2 CMR may suffer from motion-related artifacts resulting in non-diagnostic quality images. Visual inspection of image quality is time-consuming and also relies on experienced radiologists. In this chapter, we proposed an automatic method for CMR image quality assessment. Given limited data and an unbalanced class ratio, we proposed three specially designed data augmentation methods to enlarge the dataset including generating transition phases between ED and ES phases, generating images using different levels of respiratory motion and generating images using histogram matching and linear interpolation. To mimic the sampling inspection, we randomly take two subsamples from one 3D volume to estimate the quality of a 3D volume. In the developed model, which was adapted from ResNet, channel attention is used to explore the intra-channel relationship for the features extracted from each subsample. Subsequently, a feature fusion module is introduced to fuse features from two subsamples to predict the image quality. The proposed method is validated in the 2022 CMRxMotion competition, achieving a mean accuracy of 75% and 72.5% in training and validation dataset, respectively. Additionally, our method ranked at the 4th place in the testing dataset which was hidden by the organizer.

Chapter 3 In this chapter, we leveraged the spatiotemporal information from neighboring slices to improve the segmentation accuracy. The target image is stacked with its spatial or temporal neighboring images as the input. Then a stack attention is developed to extract and weigh the relevant features using the target image as a guide. The stack attention is inserted into U-Net to automatically segment the LV and myocardium from multi-slice short-axis cardiac MRI. An internal data set from one center and one public data set of the 2017 Automated Cardiac Segmentation Challenge (ACDC) were involved in evaluating and validating the proposed method. The model is trained on the internal data set first and then fine-tuned on the public data set. The method achieved a Dice of 0.91 and Hausdorff Distance of 3.37 mm on the in-house data set. The performance on the ACDC data set achieved a Dice of 0.92, 0.89 and Hausdorff Distance of 9.7 mm and 7.1 mm on ED and ED phases, respectively, which confirms a good generalization. Additionally, the results in both data sets show high correlation of LVEF and myocardium mass derived from the model and manual segmentation, demonstrating a potential valuable application in clinical practice.

Chapter 4 This chapter focuses on model generalization, in which the aim is to develop a model that performs well on unseen data sets from different centers, vendors or different diseases. The M&Ms-2 Challenge is motivated to segment the right ventricle based on a multi-disease, multi-view and multi-center samples of 360 cardiac MRI datasets. The most straightforward approach to tackle this problem is to collect more data to train a model. Given limited labeled data, we first introduce an intensity-based registration method to propagate the available labels from the end-diastolic and end-systolic phases to the other unlabeled phases. We subsequently investigate the performance of different input modalities including single 2D image, multi-channel 2D image and 3D volume. The multi-channel 2D image is constructed using the spatial and temporal stack proposed in Chapter 3. On the validation data set, our method achieved a Dice of 0.92 and 0.92, Hausdorff Distance of 9.5 mm and 5.3 mm in short-axis and long-axis view, respectively. Our method also generates a good performance on the hidden testing dataset, yielding a Dice of 0.93, 0.92 and Hausdorff Distance of 10.6 mm, 6.0 mm in short-axis and long-axis view, respectively. The experimental results demonstrate that the multi-channel 2D image provides more information for the segmentation. Combining volume input and label propagation can further improve the generalization ability.

Previously reported 4D flow segmentation approaches rely on the registration between cine MRI and 4D flow data, which requires high computational cost. **Chapter 5** and **Chapter 6** focus on LV segmentation directly from 4D flow MRI without being dependent on additional cine MRI. In **Chapter 5**, we explored using the combination of magnitude and velocity images together in 4D flow data as input. The poor contrast between the heart chambers and myocardium will result in inherent uncertainty in the segmentation results. Therefore, Monte Carlo dropout method is introduced to assess the segmentation uncertainty. Additionally, five deep learning based models are compared to investigate the effect of using different network architectures, data pre-processing, inputs and feature fusion methods on the segmentation performance. Based on the results, the proposed method was shown to be highly accurate. Additionally, the clinical parameters derived from the best model show a high correlation with results derived from manual annotations, confirming the feasibility of LV segmentation from 4D flow MRI directly.

Chapter 6 presents a transformer based efficient feature fusion method to fuse the information from magnitude and velocity images and to improve the segmentation performance in 4D flow MRI. The network is an encoder-decoder structure based on U-Net. In the encoder, the magnitude and velocity images are considered as the inputs of two branches separately. The features from the same level are integrated using the feature fusion module. The cross- and self-fusion layer in the feature fusion module aim to explore the inter- and intra-relationship between those features. The fused features are added into the original features. The paired multi-level features are concatenated along the channel dimension followed by a convolutional layer as the input of the decoder. The decoder is kept the same as that in U-Net. The proposed methods achieve the best performance compared to the other models and get significant improvement in clinical parameters, yielding a Pearson correlation coefficient of 83.3%, 97.4%, 96.97% and 98.92% for LVEF, EDV, ESV and KE, respectively. The proposed feature fusion method therefore facilitates to aggregate the features from different modalities in an efficient manner.

Chapter 7 In this chapter, we designed and evaluated a deep learning based method to predict the intra-cardiac blood flow pattern from long-axis cine MRI using the velocities derived from 4D flow data as the ground truth. The network, a variant of U-Net and ResNet, takes a subsequence of cine MR images as the input to extract the displacement of blood over the cardiac frames. Although the averaged predicted velocity was shown to be under-estimated by 26.69%, the global time-varying blood flow pattern shows a high correlation with the 4D flow derived velocities. A potential application of the proposed method is to estimate the E/A ratio. The results indicated that the E/A ratio can be estimated without significant bias and can further classify the diastolic function with a high accuracy. Our study is the first to employ deep learning for blood flow prediction from cine MRI. After further improvement of the model this work could potentially be valuable in clinical applications to visualize the intra-cardiac blood flow without additional 4D flow data.

8.2 Discussion and Future work

The work presented in this thesis aims to develop deep learning based methods for automated analysis of cardiac cine and 4D flow MRI.

The networks developed in chapter 3 and 4 focused primarily on automated segmentation in cine MRI. In chapter 3, we demonstrated that extracting temporal or spatial information from neighboring slices can benefit the segmentation performance in short-axis view cine MRI. The performance derived from introducing spatial features is better than using temporal information. Because the spatial stack can provide more information about the position, size and shape of the heart. While the images in the temporal stack are similar to each other and contain comparable features. In these studies deep learning has shown its promising applications in cardiac MRI segmentation. However, the developed approaches are validated in cases where the testing data is from the same domain as the training data. In a realistic scenario an significant performance drop can be observed when a trained model is applied on data from another domain. For example, when our model trained on the Leeds University dataset (LUD) is applied to the ACDC dataset directly, the segmentation accuracy drops from 90% to 70%. This can be explained by the population bias from different sites, ages, genders, races and diseases, and image appearance differences from various vendors, protocols, and magnetic field strengths resulting in data distribution heterogeneity. The heterogeneity cannot be eliminated completely using data pre-processing. The model needs to be fine-tuned or re-trained on the new data set to achieve a good performance. Therefore, domain generalization is a technical bottleneck for deploying deep learning in real-world clinical environments. Collecting and labeling vast amount of data from various centers and vendors it the most straight-forward solution. However, it is prohibitively expensive to obtain high-quality manual annotations for every domain, as it requires expert knowledge and it is also impossible to cover full spectrum of data. In chapter 4 we introduced registration to propagate the available segmentation labels to unlabeled images in order to enlarge the training data. Additionally, data augmentation techniques are used to increase the variety of training data to improve the model's robustness. The results show that the model has a good generalization on data with unseen pathologies. As a promising direction of future research, it's also worth investigating self-learning and semi-supervised learning to extract more prior knowledge to improve model's generalization ability when the training data is limited.

In chapter 5, we compared several models for LV segmentation directly from 4D flow MRI without relying on the registration between cine and 4D flow MRI. In chapter 6, we further improve the performance using a novel feature fusion method. However, there are also other considerations that need to be taken into account. Firstly, although most studies focus on introducing novel algorithms, data pre-processing including correction, enhancement, resample and normalization is also significantly important. For example, as shown in chapter 5 the performance derived from resliced data volumes is better than using the original data without data pre-processing. Similarly, the model of nnUnet (no new Unet), a self-configuring method, surpasses most existing deep learning-based segmentation approaches on 23 public datasets. The strong performance is not achieved by introducing a new network architecture for each type of data, but is the result of the carefully designed process of automatic self-configuration. Secondly, it is important to be aware that the final evaluation measurements should be valuable and reliable for the quantitative assessment of model's performance. The proposed models in Chapter 5 achieved similar results in terms of Dice and ASD as reported in Table 5.2, which makes it difficult to select the best model. In general, clinical relevant metrics derived from the segmentation results provide meaningful and actionable information for diagnosis and treatment. As compared to the other state-of-arts, the proposed method in Chapter 6 improves the Dice by only 2%, but the Pearson correlation coefficient in EDV, ESV and KE got improved by 9%, 7% and 16%. Therefore, the clinical parameters should be involved when comparing the performance of different models to ensure that the algorithms are reliable for use in medical applications. It would be possible to develop deep learning based multi-task networks to jointly perform the task of cardiac segmentation task and the regression of volume or ejection fraction prediction. Thirdly, in chapter 6 a transformer based feature fusion module was presented and achieved the best performance. The module can integrate information extracted from two different modalities or views efficiently. It can be adapted to the other applications such as integrating short- and long-axis view cine MRI for disease diagnosis, or combining apical four chamber (A4C) and two chamber (A2C) acquisitions in echocardiography data.

8.3 General conclusions

In conclusion, this thesis proposes deep learning based methods for quantifying cardiac MRI. The described methods can be applied for cine MR image quality classification and ventricle segmentation without any human interactions. Investigating combining and fusing magnitude and velocity images can be helpful for left ventricle segmentation in 4D flow MRI, which is not fully explored yet.

Moreover, we proposed a network to predict the blood flow pattern from the cine MRI. By combining visualization of the blood flow and myocardial motion in the routinely acquired standard CMR exams, the method can be potentially used in clinical studies. All the deep learning methods described in this thesis were evaluated on MRI data, but can potentially also be applied in other imaging modalities such as computed tomography and echocardiography.