HAL
open science

Automatic Quality Assessment of Cardiac MR Images with Motion Artefacts using Multi-task Learning and K-Space Motion Artefact Augmentation

Tewodros Weldebirhan Arega, Stéphanie Bricq, Fabrice Meriaudeau

To cite this version:

Tewodros Weldebirhan Arega, Stéphanie Bricq, Fabrice Meriaudeau. Automatic Quality Assessment of Cardiac MR Images with Motion Artefacts using Multi-task Learning and K-Space Motion Artefact Augmentation. Medical Image Computing and Computer Assisted Intervention (MICCAI 2022) STACOM workshop, Sep 2022, Singapore, Singapore. hal-03880574

HAL Id: hal-03880574

https://hal.science/hal-03880574v1

Submitted on 1 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automatic Quality Assessment of Cardiac MR Images with Motion Artefacts using Multi-task Learning and K-Space Motion Artefact Augmentation

Tewodros Weldebirhan Arega(✉), Stéphanie Bricq, and Fabrice Meriaudeau

ImViA Laboratory, Université Bourgogne Franche-Comté, Dijon, France
✉tewdrosw[at]gmail.com

Abstract. The movement of patients and respiratory motion during MRI acquisition produce image artefacts that reduce the image quality and its diagnostic value. Quality assessment of the images is essential to minimize segmentation errors and avoid wrong clinical decisions in the downstream tasks. In this paper, we propose automatic multi-task learning (MTL) based classification model to detect cardiac MR images with different levels of motion artefact. We also develop an automatic segmentation model that leverages k-space based motion artefact augmentation (MAA) and a novel compound loss that utilizes Dice loss with a polynomial version of cross-entropy loss (PolyLoss) to robustly segment cardiac structures from cardiac MRIs with respiratory motion artefacts. We evaluate the proposed method on Extreme Cardiac MRI Analysis Challenge under Respiratory Motion (CMRxMotion 2022) challenge dataset. For the detection task, the multi-task learning based model that simultaneously learns both image artefact prediction and breath-hold type prediction achieved significantly better results compared to the single-task model, showing the benefits of MTL. In addition, we utilized test-time augmentation (TTA) to enhance the classification accuracy and study aleatoric uncertainty of the images. Using TTA further improved the classification result as it achieved an accuracy of 0.65 and Cohen's kappa of 0.413. From the estimated aleatoric uncertainty, we observe that images with higher aleatoric uncertainty are more difficult to classify than the ones with lower uncertainty. For the segmentation task, the k-space based MAA enhanced the segmentation accuracy of the baseline model. From the results, we also observe that using a hybrid loss of Dice and PolyLoss can be advantageous to robustly segment cardiac MRIs with motion artefact, leading to a mean Dice of 0.9204, 0.8315, and 0.8906 and mean HD95 of 8.09 mm, 3.60 mm and 6.07 mm for LV, MYO and RV respectively on the official validation set. On the test set, the proposed segmentation method was ranked in second place in the segmentation task of CMRxMotion 2022 challenge.

Keywords: Cardiac MRI · Multi-task Learning · Quality Control · Aleatoric Uncertainty · Segmentation · Deep Learning · Motion Artefact

1 Introduction

Cardiovascular diseases (CVDs) are the number one cause of death globally. More people die annually from CVDs than from any other causes [20]. Advanced medical imaging techniques are employed in clinical practice for the diagnosis and prognosis of cardiac diseases. Cardiac magnetic resonance (CMR) is a set of magnetic resonance imaging used to provide anatomical and functional information about the heart.

Deep learning-based methods have recently achieved promising results in cardiac structures (left ventricular blood pool, myocardium, and right ventricular blood pool) segmentation from CMR images [4, 5]. As part of the ACDC challenge [4], Baumgartner et al. [3] used a 2D U-Net with a cross-entropy loss to segment the cardiac structures from CMR images. Isensee et al. [10] implemented an ensemble of 2D and 3D U-Net segmentation models with Dice loss, and Khened et al. [11] proposed 2D Dense U-Net with inception module to segment the cardiac structures. Li et al. [13] utilized a two-stage FCNs method that first localizes the heart region as a region of interest (ROI) and then segments the left ventricular blood pool, myocardium and right ventricular blood pool from the localized region of the CMR image. In addition, [8, 16] utilized data augmentation-based solutions, including histogram matching, contrast modification, and image synthesis to tackle domain shift or distribution shift in CMR images segmentation as part of M&Ms challenge [5]. However, in clinical settings, the CMR images are prone to motion artefacts due to respiratory motion or patient movement during CMR acquisition. This degrades the CMR image quality, which can challenge the diagnostic value of the image and can lead to incorrect analysis.

To detect and reduce motion artefact in Cardiac MR images, Lorch et al. (2017) [14] introduced a random forest-based method that used box, line, histogram, and texture features as inputs. However, only artificially created motion artifacts were used to evaluate their approach. Oksuz et al. (2018) [19] proposed a deep learning based method to detect cardiac MR motion artefacts. They utilized 3D spatio-temporal CNNs to identify motion artefacts in 2D+time short axis CMR sequences. To increase their training dataset size, they employed k-space based data augmentation. Lyu et al. (2021) [15] utilized a recurrent generative adversarial network for cardiac MR motion artefact reduction. Oksuz et al. (2020) [18] presented a deep learning framework that jointly detects, corrects, and segments CMR images with motion artefacts.

In this paper, we proposed a fully automatic classification network to detect cardiac MRIs with motion artefact. The cardiac images were acquired using different breath-hold instructions. Noticing that there might be a contextual similarity between predicting the level of motion artefact and breath-hold type (as some breath-hold types may cause more motion artefact than others), we propose a multi-task learning (MTL) based classification network to learn better features for the main task and reduce overfitting through shared representations. To robustly segment cardiac structures from cardiac MRIs with respiratory motion artefact, we propose a fully automatic deep learning segmentation model

that benefits from k-space based motion artefact data augmentation to increase training appearance variability and improve the generalization. We evaluated our method on Extreme Cardiac MRI Analysis Challenge under Respiratory Motion (CMRxMotion 2022) challenge dataset. For the classification task, the MTL-based method achieved better results compared to the single-task method, showing the advantage of sharing features between related tasks. For the segmentation task, the motion artefact augmentation and compound loss of Dice and polynomial loss enhanced the segmentation performance of the baseline method.

2 Dataset

The Extreme Cardiac MRI Analysis Challenge under Respiratory Motion (CMRxMotion 2022) [1] consists of 360 cine CMR images which have extreme cases with different levels of respiratory motions. The images were acquired using Siemens 3T MRI scanner (MAGNETOM Vida). To acquire the images, scan parameters with a spatial resolution of $2.0 \times 2.0\ mm^2$, slice thickness of $8.0mm$, and slice gap $4.0mm$ were used. In this challenge, the standard of procedure (SOP) mandates the 45 healthy volunteers to follow specific breath-holds guidelines. Each of the volunteers undergoes a 4-stage scan over the course of a single visit: 1) adhere to the breath-hold instructions; 2) halve the breath-hold period; 3) breathe freely; 4) breathe intensively. For each scan, short-axis CMR acquired at End-Systolic (ES) and End-Diastolic (ED) time frames are provided. From the total 360 (45 volunteers $\times$ 4 scans $\times$ 2 frames) short-axis CMR images, 160 of them were used for training, 40 for validation, and 160 for test.

The challenge has two tasks. The first one focuses on the quality assessment of the CMR images under respiratory motion artefacts. The quality of the images is labeled by radiologists using a standard 5-point Likert scale. Based on these scores, three levels of motion artefacts were defined: mild motion artefacts (label 1), intermediate motion artefacts (label 2) and severe motion artefacts (label 3). The second task focuses on the segmentation of cardiac structures from CMR images. All images with diagnostic quality are segmented by an experienced radiologist, including contours for left ventricular blood pool (LV), myocardium (MYO) and right ventricular (RV) blood pool.

3 Methods

3.1 Cardiac Image Quality Classification

Multi-task learning aims to solve multiple tasks simultaneously by taking advantage of the commonalities and differences across these tasks. Compared to training the models separately, it can lead to improved regularization, learning efficiency, and prediction accuracy for the task-specific models [6,7]. To leverage the advantages of multi-task learning, we proposed a multi-task deep learning

[1] `http://cmr.miccai.cloud/`

based classification network to classify the cardiac image quality based on their respiratory motion artefact level. The proposed method is designed to simultaneously conduct two classification tasks: image quality prediction (main task) and patient's breath-hold type prediction (auxiliary task). The main task has three classes: mild motion artefacts (class 1), intermediate motion artefacts (class 2) and severe motion artefacts (class 3). The auxiliary task focuses on predicting the type of breath-hold the patient followed during the CMR acquisition. It has four classes: adhere to the breath-hold instructions (class 1), halve the breath-hold period (class 2), breathe freely (class 3), and breathe intensively (class 4). The loss function of the multi-task model is computed as follows:

$$L_{Total} = L_{MainTask} + \beta L_{AuxTask} \tag{1}$$

where $L_{MainTask}$ is a weighted categorical cross-entropy loss for the main classification task, $L_{AuxTask}$ is a categorical cross-entropy loss for the auxiliary task and β is a hyper-parameter value that controls the contribution of $L_{AuxTask}$ to the total loss (Eq. 1).

For the network architecture, we modified the 3D ResNet-18 [2] architecture by adding a second classification branch, as shown in Fig. 1. More specifically, after the global average pooling of ResNet-18, we added a fully connected layer (512) with ReLU activation then, it is followed by two classification branches for the two tasks. During pre-processing, all the volumes were resampled to 0.664mm $\times$ 0.664mm $\times$ 9.60mm and the intensity of every volume was normalized to have zero-mean and unit-variance. In order to improve the robustness of the model, we utilized data augmentation such as random cropping (with size of $(300, 300, 8)$) , random rotation ($degrees = (-15, 15)$ with probability of 0.5), random flipping (horizontal and vertical flipping with probability of 0.5) and random scaling (range $(0.9, 1.2)$ with probability of 0.5). To enhance the results further and study the data-dependent (aleatoric) uncertainty [2, 22], we employed test-time augmentation (TTA). For TTA, we used M different types of data augmentations during testing. The augmentations utilized for TTA include horizontal flipping, vertical flipping, random rotation ($degrees = (-10, 10)$) and random scaling ($range = (0.75, 1.25)$). Then the mean prediction of the M augmented images is used as the final prediction, and the variance or entropy of these predictions is considered as the uncertainty measure.

3.2 Cardiac Segmentation

For the Cardiac MRI segmentation, we employed a 3D segmentation network which is based on 3D nnU-Net framework [9]. The U-Net's encoder and decoder consist of 12 convolutional layers where each convolution is followed by instance normalization and Leaky ReLU (negative slope of 0.01) activation function.

Dice loss (L_{Dice}) is a region-based loss that directly optimizes the Dice coefficient metric. PolyLoss [12] redesigns loss functions as a linear combination

[2] `https://github.com/Project-MONAI`

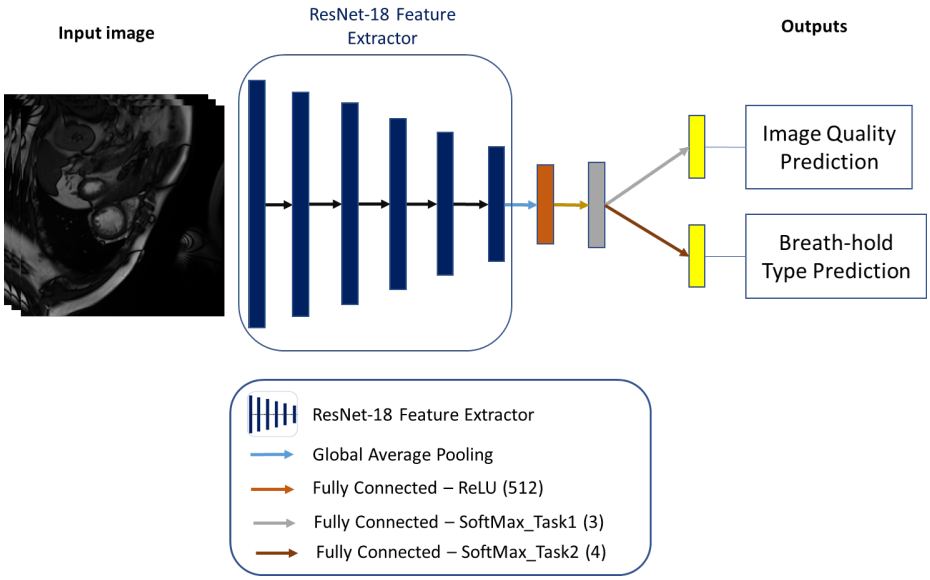


Fig. 1. Overview of the multi-task classification network

of polynomial functions. By dropping the higher order polynomials and adding terms that perturb the polynomial coefficients, Leng *et al.* (2022) [12] came up with a simplified version of the polynomial loss, which is called Poly-1. As can be observed in Eq. 2, this loss function modifies the cross-entropy (L_{CE}) by just adding one hyper-parameter (ϵ) [12], where p is the prediction probability of the target class. In this work, we utilized a hybrid loss of Dice loss and a Polynomial version of cross-entropy loss ($L_{DicePolyCE}$) to segment the heart structures from cardiac MRIs, as shown Eq. 3.

$$L_{PolyCE} = L_{CE} + \epsilon(1 - p) \tag{2}$$

$$L_{DicePolyCE} = L_{Dice} + L_{PolyCE} \tag{3}$$

To enhance the robustness of the segmentation model in the cardiac MRI dataset with motion artefact, we utilized k-space motion artefact data augmentation in addition to intensity-based data augmentation and spatial data augmentation. We applied brightness, contrast, gamma, Gaussian noise, and blur as an intensity-based augmentation. Elastic deformation, random rotation, random scaling, random flipping, and low resolution simulation are employed as spatial augmentations. During MR image acquisition, signals from the MRI scanner are temporarily stored in the k-space. The inverse Fourier transform of the k-space is then computed to provide the MR image. To generate k-space motion artefact from the cardiac MRIs, the k-space is filled with random rigidly modified versions of the original image. Then the inverse transform of the compound k-space is computed [1, 21]. The k-space motion artefact augmentations were generated

using TorchIO library [3]. We applied the k-space motion artefact on randomly selected good-quality cardiac MR training images. The generated images were then added to the training dataset to increase the variety of the training images.

3.3 Training

The segmentation models were trained for 1000 epochs in a five-fold cross-validation scheme. The weights of the network were optimized using Stochastic gradient descent (SGD) with Nesterov momentum ($\mu = 0.99$) with an initial learning rate of 0.01 [9]. The learning rate was decayed using the "poly" learning rate policy. The mini-batch size was 2 for the segmentation model. For the classification task, the dataset was shuffled and randomly split into 65% training, and 35% validation. The models were trained for 300 epochs with an ADAM optimizer and a mini-batch size of 32. We used a value of 1 for epsilon (ϵ) in the PolyLoss (Eq. 2). For the multi-task loss, the weighting factor (β) (in Eq. 1) is empirically chosen to be 0.4. To compute the data-dependent uncertainty, we used a value of 5 for M (number of data augmentations). All the training was done on NVIDIA Tesla V100 GPUs using Pytorch deep learning framework.

4 Results and Discussion

For the evaluation of segmentation results, Dice coefficient and Hausdorff distance 95% percentile (HD95) metrics are employed. To assess the classification results, accuracy and Cohen's kappa metrics are utilized.

As can be seen in Table 1, the proposed multi-task network yielded a significantly better result compared to the single-task network. The performance increase was almost by 10% and 15% in terms of accuracy and Cohen's kappa, respectively. This shows that simultaneously learning both tasks can help the network learn better representations for the main task. Applying TTA to the multi-task model further improved the results, as shown in Table 1. When we ensemble the results of the models which use different loss functions with the results of single-task and multi-task models, it yielded an accuracy of 0.7 and Cohen's kappa of 0.507.

TTA can also be used to estimate heteroscedastic aleatoric uncertainty of images [2]. Comparing the aleatoric uncertainties on our own validation split ($n = 56$) in Fig. 2, one can observe that the input-dependent uncertainty (variance) of the incorrect predictions (0.073) is higher than the correct predictions (0.060). This indicates that images with higher aleatoric uncertainty are harder for the model to classify than images with lower aleatoric uncertainty.

[3] `https://github.com/fepegar/torchio`

Table 1. Image quality classification results for different methods on the official validation set ($n = 40$) of the challenge. The bold values are the best.

Method	**Accuracy**	**Cohen's Kappa**
Single-task (image quality)	0.525	0.209
Multi-task (image quality, breath-hold type)	0.625	0.367
Multi-task +TTA	0.65	0.413
Ensemble	**0.7**	**0.507**

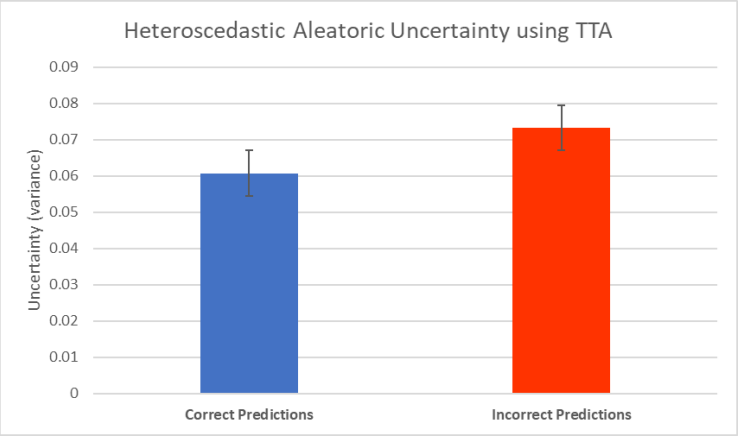
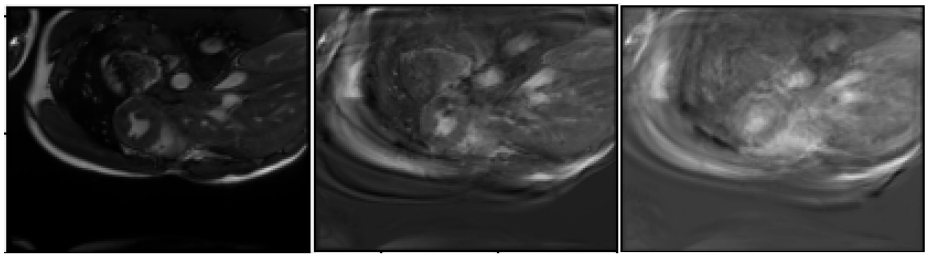


Fig. 2. Comparing the aleatoric uncertainty of correctly and incorrectly predicted images on our local validation split

For the segmentation task, the baseline method is the default nnUNet network. It uses light data augmentation and a compound loss that combines Dice loss with cross-entropy Loss (DiceCE). We compared the performance of the proposed DicePolyCE loss with two common segmentation losses: DiceCE loss (baseline) and DiceFocal loss [17] (a hybrid loss that combines Dice loss with Focal loss). Regarding data augmentations, we divided the experiments into light data augmentation (baseline), moderate motion artefact augmentation (moderate MAA) and heavy motion artefact augmentation (heavy MAA). The light data augmentation uses elastic deformation ($\alpha : (0., 200.), \sigma : (9., 13.)$, random rotation ($-30, +30$ degrees), scaling $(0.85, 1.25)$, mirroring, brightness ($\mu : 0, \sigma : 0.1$), and gamma $(0.7, 1.5)$. For moderate MAA, we added ten images with artificial motion artefact (number of transforms = 2). For heavy MAA, we increased the number of images to 30 and the number of simulated movements to 6 (number of transforms = 6), where larger values generate more distorted images. Some visual examples of motion artefact augmented images are shown in Fig. 3.

Comparing the performance of the data augmentation types, using a moderate MAA achieved the best result in terms of Dice and HD95 of the cardiac structures compared to the baseline and heavy MAA, as can be seen from Table 2. This shows the advantage of adding a moderate number of artificially generated motion artefacts into the training images to improve the model's generalizability on a dataset with respiratory motion artefacts. While using heavy MAA, which generates unrealistic motion artefacts that are more distorted from the original training images, leads to poor segmentation performance. Regarding the performance of the different hybrid losses, DiceFocal loss increased the segmentation accuracy of the baseline (DiceCE). Utilizing the proposed DicePolyCE loss outperformed the other two loss functions by achieving Dice scores of 0.9204, 0.8315, and 0.8906 and HD95 of 8.09 mm, 3.60 mm and 6.07 mm for LV, MYO and RV respectively (Table 2). This shows the robustness of the proposed loss that combines region-based Dice loss with a polynomial version of cross-entropy on cardiac MR segmentation with motion artefacts.



a) b) c)

Fig. 3. Visual appearance of a) original CMR image b) moderate motion artefact augmented CMR image c) heavy motion artefact augmented CMR image

Since the challenge allows the usage of an external public dataset, we added ACDC dataset [4] to CMRxMotion dataset to diversify the input images and improve the generalization of the segmentation model. For the ACDC dataset, as a preprocessing step, we resampled the images to the median voxel spacing of CMRxMotion and applied moderate MAA. As shown in Table 2, the addition of ACDC (CMRx+ACDC) improved the result of the baseline. When we changed the loss function from DiceCE to the proposed loss (DicePolyCE), it enhanced the segmentation performance further and achieved the best result with Dice of 0.9236, 0.8341, and 0.8967 and HD95 (in mm) of 8.31, 3.46 and 5.86 for LV, MYO and RV respectively.

Table 2. Comparison of cardiac MR segmentation performances using various data augmentations and compound loss functions on validation set ($n = 40$) of the challenge. Dice: Dice score, HD95: Hausdorff distance 95% percentile in mm, moderate MAA: moderate motion artefact augmentation, heavy MAA: heavy motion artefact augmentation. The bold values are the best.Asterisk $*$ indicates that the dataset used is a mixture of CMRxMotion and ACDC datasets.

Method	**Dice LV**	**Dice MYO**	**Dice RV**	**HD95 LV**	**HD95 MYO**	**HD95 RV**
Baseline	0.9147	0.8298	0.8886	8.64	3.74	6.11
Moderate MAA	0.9168	0.8306	0.8918	8.49	3.70	6.04
Heavy MAA	0.9141	0.8298	0.8927	8.45	3.69	5.83
DiceFocal Loss	0.9181	0.8303	0.8927	8.24	3.67	5.62
DicePolyCE Loss	0.9204	0.8315	0.8906	**8.09**	3.60	6.07
Baseline + ACDC∗ (CMRx+ACDC)	0.9196	0.8325	0.8961	8.39	3.64	**5.57**
CMRx+ACDC∗ with DicePolyCE Loss	**0.9236**	**0.8341**	**0.8967**	8.31	**3.46**	5.86

5 Conclusion

In this paper, we proposed fully automatic deep learning methods to detect motion artefact levels and segment cardiac structures from cardiac MRIs with respiratory motion artefact. For the classification task, the proposed network simultaneously predicts both the motion artefact level (main task) and breath-hold type to leverage the benefits of multi-task learning. From the results, we observed that training a multi-task network encourages the network to extract useful features for both tasks and results in better prediction performance. The proposed method greatly increased the classification result in terms of accuracy and Cohen's kappa compared to the single-task model. Applying test-time augmentation also helped the model to improve the classification result further and to estimate the aleatoric uncertainty of the data. In the segmentation task, we showed that moderate k-space based motion artefact augmentation and hybrid loss of Dice and polynomial loss can enhance the segmentation performance and improve the model's generalization capability on cardiac MRIs with motion artefacts.

6 Acknowledgements

This work was supported by the French National Research Agency (ANR), with reference ANR-19-CE45-0001-01-ACCECIT. Calculations were performed using HPC resources from DNUM CCUB (Centre de Calcul de l'Université de Bourgogne) and from GENCI–IDRIS (Grant 2022-AD011013506). We also thank the Mesocentre of Franche-Comté for the computing facilities.

References

1. Arega, T.W., Legrand, F., Bricq, S., Meriaudeau, F.: Using mri-specific data augmentation to enhance the segmentation of right ventricle in multi-disease, multi-center and multi-view cardiac mri. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 250–258. Springer (2021)
2. Ayhan, M.S., Berens, P.: Test-time data augmentation for estimation of heteroscedastic aleatoric uncertainty in deep neural networks (2018)
3. Baumgartner, C.F., Koch, L.M., Pollefeys, M., Konukoglu, E.: An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 111–119. Springer (2017)
4. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., Cetin, I., Lekadir, K., Camara, O., Ballester, M.A.G., Sanromá, G., Napel, S., Petersen, S.E., Tziritas, G., Grinias, E., Khened, M., Kollerathu, V.A., Krishnamurthi, G., Rohé, M.M., Pennec, X., Sermesant, M., Isensee, F., Jäger, P.F., Maier-Hein, K., Full, P.M., Wolf, I., Engelhardt, S., Baumgartner, C.F., Koch, L.M., Wolterink, J.M., Igum, I., Jang, Y., Hong, Y., Patravali, J., Jain, S., Humbert, O., Jodoin, P.M.: Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved? IEEE Transactions on Medical Imaging **37**, 2514–2525 (2018)
5. Campello, V.M., Gkontra, P., Izquierdo, C., Martín-Isla, C., Sojoudi, A., Full, P.M., Maier-Hein, K., Zhang, Y., He, Z., Ma, J., Parreño, M., Albiol, A., Kong, F., Shadden, S.C., Acero, J.C., Sundaresan, V., Saber, M., Elattar, M., Li, H., Menze, B.H., Khader, F., Haarburger, C., Scannell, C.M., Veta, M., Carscadden, A., Punithakumar, K., Liu, X., Tsaftaris, S.A., Huang, X., Yang, X., Li, L., Zhuang, X., Viladés, D., Descalzo, M., Guala, A., Mura, L.L., Friedrich, M.G.W., Garg, R., Lebel, J., Henriques, F., Karakas, M., Cavus, E., Petersen, S.E., Escalera, S., Seguí, S., Palomares, J.F.R., Lekadir, K.: Multi-centre, multi-vendor and multi-disease cardiac segmentation: The m&ms challenge. IEEE Transactions on Medical Imaging **40**, 3543–3554 (2021)
6. Caruana, R.: Multitask learning. Machine Learning **28**, 41–75 (2004)
7. Chen, C., Bai, W., Rueckert, D.: Multi-task learning for left atrial segmentation on ge-mri. In: STACOM@MICCAI (2018)
8. Full, P.M., Isensee, F., Jäger, P.F., Maier-Hein, K.: Studying robustness of semantic segmentation under domain shift in cardiac mri. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 238–249. Springer (2020)
9. Isensee, F., Jaeger, P., Kohl, S., Petersen, J., Maier-Hein, K.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nature Methods **18**, 1–9 (02 2021). https://doi.org/10.1038/s41592-020-01008-z
10. Isensee, F., Jaeger, P.F., Full, P.M., Wolf, I., Engelhardt, S., Maier-Hein, K.H.: Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. In: International workshop on statistical atlases and computational models of the heart. pp. 120–129. Springer (2017)
11. Khened, M., Varghese, A., Krishnamurthi, G.: Densely connected fully convolutional network for short-axis cardiac cine mr image segmentation and heart diagnosis using random forest. In: STACOM@MICCAI (2017)
12. Leng, Z., Tan, M., Liu, C., Cubuk, E.D., Shi, X., Cheng, S., Anguelov, D.: Polyloss: A polynomial expansion perspective of classification loss functions. ArXiv **abs/2204.12511** (2022)

13. Li, C., Tong, Q., Liao, X., Si, W., Chen, S., Wang, Q., Yuan, Z.: Apcp-net: Aggregated parallel cross-scale pyramid network for cmr segmentation. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) pp. 784–788 (2019)
14. Lorch, B., Vaillant, G., Baumgartner, C.F., Bai, W., Rueckert, D., Maier, A.K.: Automated detection of motion artefacts in mr imaging using decision forests. Journal of Medical Engineering **2017** (2017)
15. Lyu, Q., Shan, H., Xie, Y., Kwan, A.C., Otaki, Y., Kuronuma, K., Li, D., Wang, G.: Cine cardiac mri motion artifact reduction using a recurrent neural network. IEEE Transactions on Medical Imaging **40**, 2170–2181 (2021)
16. Ma, J.: Histogram matching augmentation for domain adaptation with application to multi-centre, multi-vendor and multi-disease cardiac image segmentation. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 177–186. Springer (2020)
17. Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., Martel, A.L.: Loss odyssey in medical image segmentation. Medical image analysis **71**, 102035 (2021)
18. Oksuz, I., Clough, J.R., Ruijsink, B., Anton, E.P., Bustin, A., Cruz, G., Prieto, C., King, A.P., Schnabel, J.A.: Deep learning-based detection and correction of cardiac mr motion artefacts during reconstruction for high-quality segmentation. IEEE Transactions on Medical Imaging **39**, 4001–4010 (2020)
19. Öksüz, I., Ruijsink, B., Puyol-Antón, E., Bustin, A., Cruz, G., Prieto, C., Rueckert, D., Schnabel, J.A., King, A.P.: Deep learning using k-space based data augmentation for automated cardiac mr motion artefact detection. In: MICCAI (2018)
20. Organization, W.H.: Cardiovascular diseases (cvds) (2017), `https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)`
21. Pérez-García, F., Sparks, R., Ourselin, S.: Torchio: A python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. Computer Methods and Programs in Biomedicine **208** (2021)
22. Wang, G., Li, W., Aertsen, M., Deprest, J.A., Ourselin, S., Vercauteren, T.K.M.: Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing **335**, 34 – 45 (2019)