Automated Cardiac Resting Phase Detection Targeted on the Right Coronary Artery

Seung Su Yoon   seung.su.yoon@fau.de
Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen 91058, Germany
Magnetic Resonance, Siemens Healthcare GmbH, Erlangen 91052, Germany

Elisabeth Preuhs   elisabeth.preuhs@gmail.com
Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen 91058, Germany

Michaela Schmidt   michaela.schmidt@siemens-healthineers.com
Magnetic Resonance, Siemens Healthcare GmbH, Erlangen 91052, Germany

Christoph Forman   christoph.forman@siemens-healthineers.com
Magnetic Resonance, Siemens Healthcare GmbH, Erlangen 91052, Germany

Teodora Chitiboi   teodora.chitiboi@siemens-healthineers.com
Siemens Medical Solutions USA, Inc., Princeton, NJ, USA

Puneet Sharma   sharma.puneet@siemens-healthineers.com
Siemens Medical Solutions USA, Inc., Princeton, NJ, USA

Juliano L. Fernandes   jlaraf@terra.com.br
Jose Michel Kalaf Research Institute, Sao Paulo, Brazil

Christoph Tillmanns   tillmanns@diagnostikum-berlin.de
Diagnostikum Berlin, Berlin, Germany

Jens Wetzl   jens.wetzl@siemens-healthineers.com
Magnetic Resonance, Siemens Healthcare GmbH, Erlangen 91052, Germany

Andreas Maier   andreas.maier@fau.de
Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen 91058, Germany

Abstract

Static cardiac imaging such as late gadolinium enhancement, mapping, or 3-D coronary angiography require prior information, e.g., the phase during a cardiac cycle with least motion, called resting phase (RP). The purpose of this work is to propose a fully automated framework that allows the detection of the right coronary artery (RCA) RP within CINE series. The proposed prototype system consists of three main steps. First, the localization of the regions of interest (ROI) is performed. Second, the cropped ROI series are taken for tracking motions over all time points. Third, the output motion values are used to classify RPs. In this work, we focused on the detection of the area with the outer edge of the cross-section of the RCA as our target. The proposed framework was evaluated on 102 clinically acquired dataset at 1.5T and 3T. The automatically classified RPs were compared with the reference RPs annotated manually by a expert for testing the robustness and feasibility of the framework. The predicted RCA RPs showed high agreement with the experts annotated RPs with 92.7% accuracy, 90.5% sensitivity and 95.0% specificity for the unseen study dataset. The mean absolute difference of the start and end RP was 13.6 $\pm$ 18.6 ms for the validation study dataset (n=102). In this work, automated RP detection has been introduced by the proposed framework and demonstrated feasibility, robustness, and applicability for static imaging acquisitions.

Keywords: Resting Phase Detection, AI Workflow Automation, Static Cardiac Imaging

1. Introduction

In cardiovascular magnetic resonance (CMR) imaging, static cardiac imaging techniques, such as late gadolinium enhancement (LGE) (Kellman and Arai, 2012; Kellman et al., 2002; Akçakaya et al., 2012; Basha et al., 2017), mapping (Kellman and Hansen, 2014; Messroghli et al., 2017; Aherne et al., 2020), or three-dimensional (3-D) whole heart coronary angiography (Munoz et al., 2020; Greil et al., 2017; Cruz et al., 2017; Forman et al., 2014) are increasingly being performed to qualitatively and quantitatively assess the cardiac anatomy and function. It is important to acquire the data during the phase of the cardiac cycle with least motion, called a resting phase (RP), especially mid- or end-diastolic (ED) phases (Kramer et al., 2020; Isma'eel et al., 2009; Kramer et al., 2013), or in patients with a fast heart rate during the end-systolic (ES) phase.

In standardized CMR protocols (Kramer et al., 2020, 2013), the guidelines recommend using the diastolic RP with a duration of less than 200 ms as the acquisition window for static cardiac imaging. In certain situations, e.g., high heart rate or patients with arrhythmias, especially in terms of mapping acquisition, the systolic RP is preferably chosen. As outlined in (Kim et al., 2001; Seifarth et al., 2007; Hofman et al., 1998), electrocardiogram-based heuristics enable the ED phase selection based on a trigger time at 75% of the RR interval for most patients, however it can be suboptimal due to the magnetohydrodynamic effect (Abi-Abdallah et al., 2007), and not generalizable, especially for patients with high or irregular heart rates. For advanced applications such as high-resolution angiography, accurate determination of RP is necessary. As different structures in the heart rest at different times of the cardiac cycle, ideally a targeted RP for the anatomy of interest should be determined. For coronary angiography, for example, it is suggested to accurately determine the RP of the right coronary artery (RCA) (Kramer et al., 2020, 2013; Shechter et al., 2005; Johnson et al., 2004; Wang et al., 2001).

The selection of the RPs is typically performed based on visual inspection on a CINE series acquired prior to the static imaging. In current clinical practice, an expert is required to select either the end-systolic, mid- or end-diastolic phase for acquisition. To tackle the complex and time-consuming manual task of RP selection for the static cardiac imaging, some previous studies have been introduced to perform the RP determination automatically. In previously conducted studies (Wang et al., 2001; Stuber et al., 1999), a calibration scan-based approach using navigator echoes has been presented, however this approach requires significant user experience and interaction in order to accurately plan the navigator position. A threshold-based clustering algorithm was proposed to track low-motion periods (Suever et al., 2011). However to track the RCA area, manual tracking is required. Another approach (Jahnke et al., 2005) was proposed using image based cross-correlation of CINE series for the automatic selection of RPs, proving to be advantageous in terms of image quality. An extension of the previous method (Ustun et al., 2007) calculates the myocardial displacement from the cross-correlation calculation. These methods, however, also require user interaction to position the region of interest (ROI) enclosing the heart. An alternative technique with automated RCA positioning (Sato et al., 2009) using template matching algorithm was proposed to automatically select RP based on image intensity differences. The intensity difference calculation as well as the template matching algorithm can be sensitive to artifacts. Further, a method attempts to determine the cardiac motion

resolution-independently (Huang et al., 2014) using intensity standard deviation calculation, and additionally proposed the motion extraction based on deformable model-based registration. Another idea was proposed in which the feasibility of motion correction algorithm for quantifying the rest periods of the coronary arteries is shown (Shechter et al., 2005). However, the computed RP is based on the entire field-of-view and do not provide localized RPs for specific anatomies, and the detection is limited to two RPs due to the two local minima search. In a recent study (Asou et al., 2018), an automated RP selection algorithm was introduced based on a motion area map generated from the high-speed component of the motion within a CINE series, however the high-speed component is not necessarily related to the anatomical structures.

Deep Learning approaches have the potential to automate clinical workflows, and the state-of-the-art methods using convolutional neural networks (CNN) are currently used for image localization and segmentation (Krizhevsky et al., 2012; Szegedy et al., 2015; Simonyan and Zisserman, 2014; Ronneberger et al., 2015). The CNN-based models are particularly used for learning the optimal spatial features from input data, especially images, thus performing a specific task such as localization or segmentation. Beside the power of CNN models for learning the spatial information, 3-D convolutional operators are used to learn spatial and temporal information for capturing features in a 3-D input data (Ji et al., 2012; Tran et al., 2015). Registration approaches are commonly used when it comes to motion analysis, object tracking, etc. (Dyke et al., 2019; Chefd'Hotel et al., 2002; Szeliski and Shum, 1996; Spinei et al., 1998; Rueckert et al., 1999), by estimating a smooth correspondence function mapping between the coordinates from a reference image and those in a target image. These techniques can be used for calculating the motion of a target with the deformation fields within CINE series quantitatively.

In this work, we propose a fully automated prototype system combining the advantages of the 3-D CNN and registration algorithms for detecting localized RPs of the RCA from 4-chamber view (4CH) CINE series. As the CINE series are time-resolved images, a 3-D CNN based model is trained to perform landmark detection over the cardiac phases. The proposed system combines the deep neural network for landmark detection and a registration algorithm. The motion within a localized anatomy is quantified in order to automatically classifying the systolic and diastolic RPs of the RCA. The duration of the RPs is quantified in both cases. To test the robustness and feasibility, the proposed framework was integrated into the scanner software and validated on patient data acquired on 1.5T and 3T scanners at multiple centers and different CINE sequences.

2. Methods

2.1 System Overview

The proposed prototype system consists of three main steps that are executed consecutively (Figure 1). The first step is to localize the ROI (see details in section 2.2) from the input which is a 4CH CINE series. ROI can be chosen for any anatomical structures, that are of interest displayed in 4CH images, such as the RCA. The localization can be performed by neural networks trained for landmark detection tasks. In case of landmark detection, the output can be pixel coordinates that are used for cropping the image to the localized anatomy (see details in section 2.3). The cropped series containing the ROI are used for
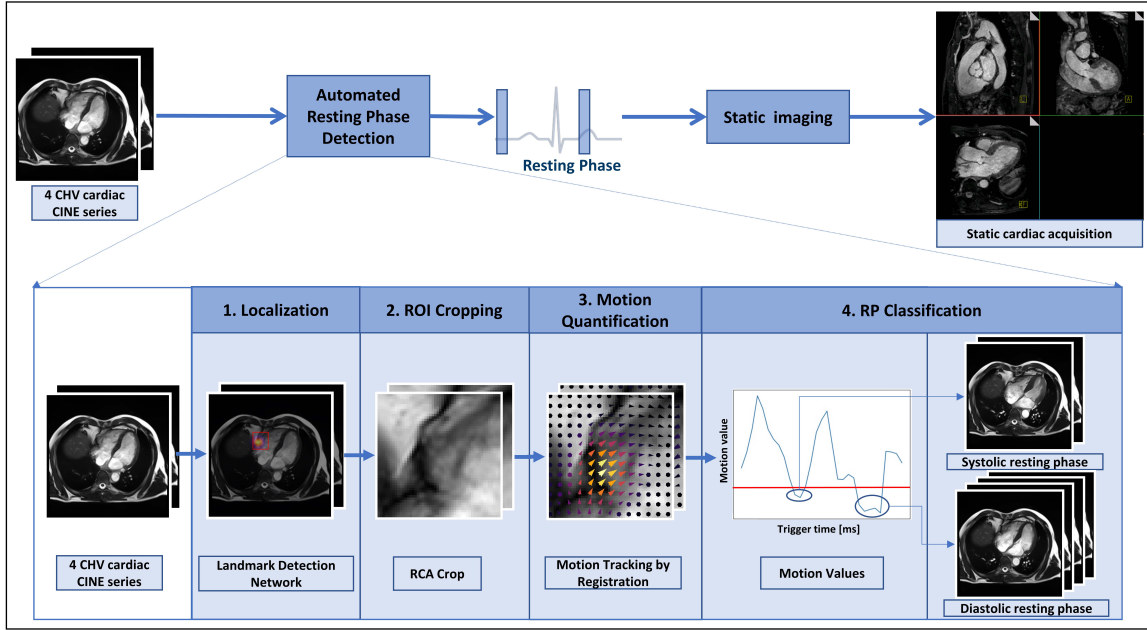
Figure 1: Overview of the proposed system. Prior to a static imaging, a RP must be defined to prevent motion artifacts. The automated RP detection system consists of localization, ROI cropping, motion quantification and RP classification steps and provides the RPs of the target of interest within a 4CH CINE series.

motion quantification (see details in section 2.4) by performing the elastic image registration (Chefd'Hotel et al., 2002). By taking the median of the magnitude of the deformation fields from consecutive frames, the motion values over time points are calculated. Finally, the RPs are classified by an absolute threshold, defined based on the correlations of the predicted RPs by varying the thresholds with the expert annotations (see details in section 2.5). The frames which have lower motion values than the threshold value, are selected as RPs. The detected RPs are then used to plan the subsequent static image acquisition.

2.2 Localization

In this work, the landmark detection neural network is built to automatically detect the location of RCA in a 4CH time-resolved series. The densely connected neural network (3-D DenseNet) architecture is trained to regress the x- and y- coordinates of the RCA over time and it is described in detail here. As a preprocessing step, the 4CH CINE series are interpolated to a fixed spatial and temporal size of $224 \times 224 \times 32$ to be independent from resolution. Further, the min-max pixel intensity normalization was applied to rescale the different intensity range in [0,1]. The 3-D DenseNet proposed in (Huang et al., 2017) was trained under supervised learning. The weights of the network were updated by using the Adam optimizer (Kingma and Ba, 2014) with $\lambda = 10^{-3}$ and the mean-squared-error loss

Table 1: The extended 3-D DenseNet built for the RCA detection. Each Conv consists of the successively executed layers of 3-D batch normalization, rectified linear unit activation function and 3-D convolutions. For the RCA detection, the 3-D DenseNet with the total number of 122 convolutional layers are built as follows: $d = 32$, $h, w = 224$ and $c_{1,2,3,4} = 6, 12, 24, 16$.

Layers	Output Size	3-D DenseNet
Convolution	$d \times \frac{h}{2} \times \frac{w}{2}$	$3 \times 3 \times 3$ Conv, $1 \times 2 \times 2$ stride
Pooling	$\frac{d}{2} \times \frac{h}{4} \times \frac{w}{4}$	$3 \times 3 \times 3$ max-pool, $2 \times 2 \times 2$ stride
Dense block	$\frac{d}{2} \times \frac{h}{4} \times \frac{w}{4}$	$\begin{bmatrix} 1 \times 1 \times 1 \text{ Conv} \\ 3 \times 3 \times 3 \text{ Conv} \end{bmatrix} \times c_1$
Transition block	$\frac{d}{4} \times \frac{h}{8} \times \frac{w}{8}$	$1 \times 1 \times 1$ Conv, $2 \times 2 \times 2$ avg-pool, $2 \times 2 \times 2$ stride
Dense block	$\frac{d}{4} \times \frac{h}{8} \times \frac{w}{8}$	$\begin{bmatrix} 1 \times 1 \times 1 \text{ Conv} \\ 3 \times 3 \times 3 \text{ Conv} \end{bmatrix} \times c_2$
Transition block	$\frac{d}{8} \times \frac{h}{16} \times \frac{w}{16}$	$1 \times 1 \times 1$ Conv, $2 \times 2 \times 2$ avg-pool, $2 \times 2 \times 2$ stride
Dense block	$\frac{d}{8} \times \frac{h}{16} \times \frac{w}{16}$	$\begin{bmatrix} 1 \times 1 \times 1 \text{ Conv} \\ 3 \times 3 \times 3 \text{ Conv} \end{bmatrix} \times c_3$
Transition block	$\frac{d}{16} \times \frac{h}{32} \times \frac{w}{32}$	$1 \times 1 \times 1$ Conv, $2 \times 2 \times 2$ avg-pool, $2 \times 2 \times 2$ stride
Dense block	$\frac{d}{16} \times \frac{h}{32} \times \frac{w}{32}$	$\begin{bmatrix} 1 \times 1 \times 1 \text{ Conv} \\ 3 \times 3 \times 3 \text{ Conv} \end{bmatrix} \times c_4$
Classification block	$d \times 4$	3-D adaptive avg-pool, $1 \times 1 \times 1$ Conv

function (MSE) as follows:

$$MSE = \frac{1}{N} \sum_{n=1}^{N} \left(\frac{1}{T} \, \|(\hat{y}_{t,n} - y_{t,n}\|^2 \right) \qquad (1)$$

where $\hat{y}_{t,n}$ is the predicted pixel coordinates at the time t from n dataset, the $\hat{y}_{t,n}$ ground truth, T the number of frames, and N the number of the datasets. The ground truth is generated in a semi-supervised manner, where the RCA pixel coordinates in the first frame are manually annotated and propagated to the next frames using the deformation fields describing the displacement between $\hat{y}_{t,n}$ and $\hat{y}_{t+1,n}$. The deformation field are generated by using the elastic image registration (Chefd'Hotel et al., 2002). Each frame was then corrected manually if the propagated coordinates were not accurate. The total number of convolutional layers (Conv) is 122, and before each Conv a 3-D batch normalization (BN) (Ioffe and Szegedy, 2015) and rectified linear unit (ReLU) activation functions (Nair and

5

Hinton, 2010) are applied. After the initial 3-D Conv and max pooling operator, the feature maps were forwarded through 4 concatenated dense blocks (DB) and transition blocks (TB). The number of 4 concatenated DBs are set to 6, 12, 24, 16, in which after each DB, a TB is applied. In each DB, 2 Convs, each followed by 3-D BN and ReLU operators with the increase of the feature maps with 12 are applied. Each layer obtains additional inputs from all preceding layers and forwards the feature maps to all subsequent layers. In each TB, a Conv with BN and ReLU followed by 3-D average pooling operator is applied for spatial and temporal down-sampling. As the last step, a global average pooling and $1 \times 1 \times 1$ convolutional operator are used to regress the coordinates from the extracted features maps. The detailed architectural details can be found in the Table 1.

2.3 Cropping

From the output of the localization task, a ROI can be simply selected from the pixel coordinates. The detected pixel coordinates are transformed back to the original coordinate system, and then the cropping is performed. Given the predicted pixel coordinates of RCA by 3-D DenseNet, the bounding box is defined by taking the minimum and maximum x- and y-coordinates of the points in the coordinate plane from a time-resolved series and calculating the average of these x and y-coordinates. The size of the bounding box is selected based on prior knowledge about the size of the anatomy, chosen as $50 \times 50 \ mm^2$.

2.4 Motion Quantification

The motion values are quantitatively determined using elastic image registration (Chefd'Hotel et al., 2002). Consecutive frames of the CINE series, $s_t(x)$ and $s_{t+1}(x)$ for all timepoints t, are registered to obtain deformation fields $d_t(x)$ such that $s_{t+1}(d_t(x))$ minimizes the dissimilarity measure related to $s_t(x)$. The motion curve $m(t)$ describing the amount of RCA motion is then computed as the median of the weighted magnitudes of the deformation fields $\|d_t(x)\|$ as follows:

$$m(t) = median\{G_t(x) \cdot \|d_t(x)\|^2\} \tag{2}$$

where $G_t(x)$ is a Gaussian weighting function centered at the midpoint of the detected location of the RCA between $\hat{y}_t$ and $\hat{y}_{t+1}$ at the time point t:

$$G_t(x) = \exp\left(-\frac{\|x - p_t\|^2}{\sigma^2}\right) \tag{3}$$

while p_t denotes the midpoint of the detected RCA position. This Gaussian weighting ensures that the motion curve represents mainly the motion of the RCA, while still being robust to slight imprecisions of the localization results. The standard deviation was empirically chosen as $\sigma = 12$. Figure 2 shows the Gaussian weighting functions overlaid on the anatomical images at each time point, as well as the weighted deformation fields corresponding to subsequent image pairs. The quantification of motion can be considered in different ways, and the detailed analysis of obtaining the RCA motion values can be found in section 3.2.
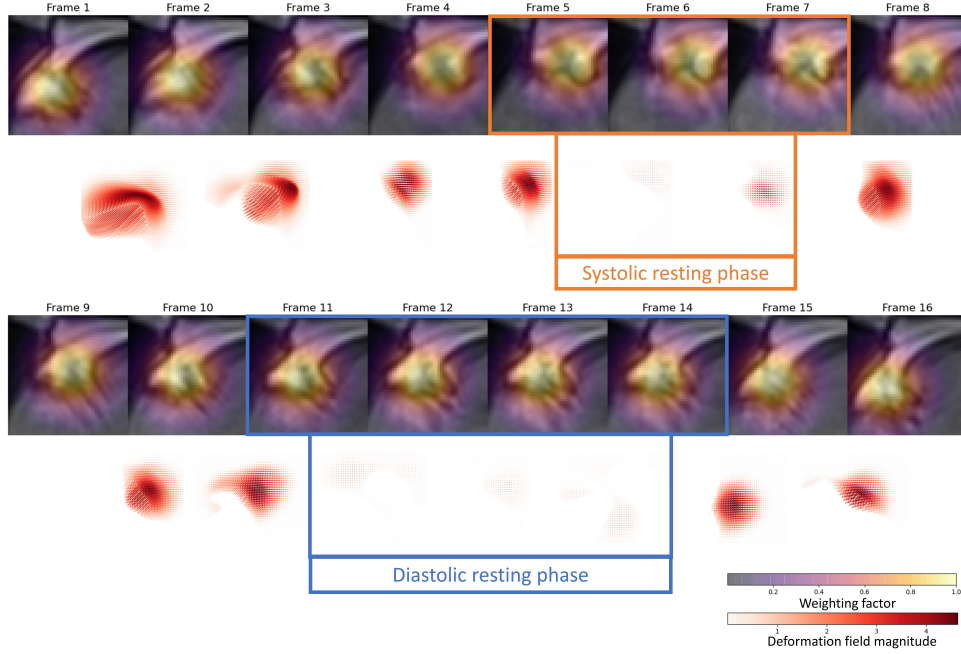
Figure 2: An example of RCA ROI overlayed with weighted heatmaps, and below each two frames the magnitude of the weighted consecutive deformation fields are illustrated. The upper color bar corresponds to the weighted heatmaps, and the below one to the magnitude of the deformation vectors. The predicted systolic resting phase is marked in orange, and the predicted diastolic resting phase in blue. For the sake of simplicity, the last frame is not visualized.

2.5 Resting Phase Classification

After the motion curve is quantified, a classification is required to obtain the RP. This classification is performed within a time interval $T_{valid} = [\alpha, \omega]$, i.e. the first α and the last ω ms of a cardiac cycle are excluded. α accounts for the time needed for preparation pulses before the data acquisition window can start. ω is a safety margin at the end of the cardiac cycle to account for heart rate variability, i.e. the detected resting phase should not be too late in the cardiac cycle in case subsequent cardiac cycles during the 3-D acquisition are shorter, which can lead to unstable measurements (Kim et al., 2001; Seifarth et al., 2007). Based on the ground truth annotation of RPs, the α, and ω are empirically chosen to be 80 ms. The RPs can be determined with an absolute threshold from the motion values. The frames with motion values lower than the absolute threshold value are assigned as RPs which can be described as follows:

$$RP(t) = \begin{cases} 1, & m(t) < \tau, t \in T_{valid} \\ 0, & m(t) \geq \tau, t \in T_{valid} \end{cases}$$

where τ is the absolute threshold value, and $m(t)$ is the obtained motion value at trigger time t. The threshold value is obtained based on sensitivity (true positive rate) and specificity

7

(true negative rate) analysis from manual annotations. The optimal absolute threshold is chosen as the one achieved with best balanced accuracy.

2.6 Data

2.6.1 TRAINING AND VALIDATION DATASET FOR THE RCA DETECTION NETWORK

Data used for training and evaluating the RCA detection network was acquired on 1.5T and 3T clinical MRI scanners (MAGNETOM Aera, Avanto, Prisma, Skyra, Trio TIM; Siemens Healthcare, Erlangen, Germany) at multiple centers (n=1000). The dataset was split into 70% training, 15% validation and 15% testing set for the RCA detection. Details about the datasets used for training, validating, testing the RCA detection network, and evaluating the classified resting phases are shown in Table 2. A medical expert with more than 10 years of cardiac MRI experience manually annotated the RPs on 76 datasets from the testing set for the analysis of the system.

2.6.2 STUDY DATASET

Data used for evaluating the proposed system was acquired on 1.5T and 3T clinical MRI scanners (MAGNETOM Aera, Avanto fit, Skyra, Skyra fit, Sola, Vida; Siemens Healthcare, Erlangen, Germany) at multiple centers (n=102). The proposed system was integrated into the scanner software and tested online. The dataset from the study was not used for training, testing the RCA network or the threshold analysis, and was not mixed with the RCA Detection Network dataset. The test datasets consist of non-pre-selected patient data with a minimum heart rate of 35 and a maximum heart rate of 97 who are clinically referred for coronary assessment. None of the patients have not received beta-blockers or a target heart rate was determined. Details about the study datasets are listed in Table 2.

2.7 Experiments

2.7.1 LOCALIZATION

The RCA detection is validated by calculating the mean and standard deviation of the Euclidean distance between the predicted pixel coordinates and the ground truth pixel coordinates.

$$\text{Distance Error} = \frac{1}{N} \sum_{n=1}^{N} \left(\frac{1}{T} \sum_{t=1}^{T} \|\hat{p}_{t,n} - p_{t,n}\|^2 \right) \tag{4}$$

where N denotes the number of annotated test dataset, T the number of time frames in each CINE series and $\hat{p}$ the predicted and p the ground truth RCA position. The robustness and performance of the network was qualitatively validated on 12 oblique and diverse oriented 4CH CINE and 9 clinically acquired unseen study dataset with different field strength scanners. Further, the network was evaluated by a box plot showing the performance of the prediction at each frame.

Table 2: Statistics about the data population and acquisition used for training and testing the RCA detection network, and of the additional unseen study dataset.

	RCA Detection Network Dataset (Resting Phase annotation)	Study Dataset
Number of patients	1000 (76)	102
Age	55.0 ± 19.0 (59.0 ± 17.2)	39.3 ± 10.9
Gender	64 % male (68% male)	69% male
Heart Rate [bpm]	64.7 ± 11.6 (68.0 ± 13.2)	70.1 ± 12.3
Field Strength	$25\,\%\,1.5\,T$, $75\,\%\,3\,T$ $(46\,\%\,1.5\,T,$ $54\,\%\,3\,T)$	$22\,\%\,1.5\,T$, $78\,\%\,3\,T$
Spatial Resolution $[mm^2]$	1.4 ± 0.1 (1.5 ± 0.2)	1.7 ± 0.1
Temporal Resolution $[ms]$	33.8 ± 9.9 (35.8 ± 6.8)	37.3 ± 7.8
FOV [mm x mm]	311.3 ± 30.6 344.4 ± 26.9 $(312.2 \pm 31.7$ $357.4 \pm 23.6)$	283.5 ± 9.2 345.9 ± 8.9
Number of Frames	25.8 ± 2.3 (25.6 ± 1.6)	26.0 ± 2.0

2.7.2 MOTION QUANTIFICATION

In order to find the best approach to quantify motion values, several approaches were evaluated on the annotated datasets for quantifying RCA motion from a cropped CINE series. The first approach is to quantify motion based on the distance between detected pixel coordinate over each adjacent time point as follows:

$$m_{\mathrm{dist}(t)} = \|p_t - p_{t+1}\| \tag{5}$$

The second is to aggregate the magnitudes of the deformation fields within the ROI without the Gaussian weighting using percentile or mean:

$$m_{\mathrm{pct}}(t) = \eta_n\{\|d_t(x)\| \,|\, x \in \mathrm{ROI}, \text{ or}\} \tag{6}$$

$$m_{\mathrm{mean}}(t) = mean\{\|d_t(x)\| \,|\, x \in \mathrm{ROI}\} \tag{7}$$

where η_n is the nth percentile.
Our last proposed approach is to aggregate the weighted deformation field magnitudes to calculate the motion values as described in the following:

$$m_{\mathrm{wpct}}(t) = \eta_n\{G_t(\vec{x}) \cdot \|d_t(x)\| \,|\, x \in \mathrm{ROI}, \text{ or}\} \tag{8}$$

9

$$m_{\text{wmean}}(t) = mean\{G_t(x) \cdot \|d_t(x)\| \,|\, x \in \text{ROI}\},\tag{9}$$

where $\|d_t(x)\|$ is the magnitudes of the deformation fields. The percentile analysis is performed to quantify motion from the deformation fields, which is based on calculating the balanced accuracy, sensitivity and specificity by varying n from 10th to 100th by 10th percentile steps. Further, the mean value is calculated as well, and compared with the percentile analysis. In addition, the motion values of the above mentioned 9 clinically acquired dataset are extracted and qualitative validated.

2.7.3 Threahold Analysis

The analysis for finding the optimal threshold value to determine the RPs from the quantified motion value is performed by the binary classification task with varying the threshold τ from 0.01 to 1 by 0.01 steps by calculating the sensitivity and specificity. The analysis is performed separately for 1.5T and 3T and using the annotated datasets. The performance with the selected threshold τ is analyzed based on area under curve (AUC) from the receiver operating characteristic (ROC) curve (Swets, 1979), and confusion matrices are evaluated on testing and study datasets. For all different approaches of motion quantification, the threshold analysis is performed such that each approach can be fairly compared. Based on thresholding, the accuracy of classified RPs is evaluated as described in section 2.5.

2.7.4 RP Classification

To evaluate the performance of RP classification, the mean absolute error (MAE) and the standard deviation between the system predicted $\widehat{RP}$ and annotated start and end time points of systolic and diastolic RP are calculated as follows:

$$MAE_{\lambda,\mu_{\text{type}}} = \frac{1}{N}\sum_{n=1}^{N}\left|\widehat{RP}_{\lambda,\mu_{N,\text{type}}} - RP_{\lambda,\mu_{N,\text{type}}}\right|,$$

$$\text{where } \lambda \in \text{start, end of RP}, \mu \in \text{window, frame}$$

The number of RP annotated dataset is denoted as N, and the type can be the classified systolic (sys) or diastolic (dia) RP. The performance was validated by two different measures, firstly the difference of the time window and secondly, the number of images between the predicted and ground truth annotation. The time window specifies the accuracy in milliseconds, whereas the frame in number of images. The validation was performed on the testing datasets in which the RPs are manually annotated by a medical expert used for the threshold analysis. The performance of the system was analyzed based on Bland-Altmann analysis. The RPs was not counted when it was very short (<30 ms, n=10), i.e., not resolvable by the temporal resolution of the acquisition. Additionally, the results of the system validation were presented by sensitivity, specificity and accuracy. To overcome the imbalanced classes (RP, no RP), the balanced accuracy is calculated based on true positive rate, and false true negative rate as follows: Accuracy $= \frac{(TPR+TNR)}{2}$, where TPR is true positive rate, and TNR true negative rate. Further, the ranges of each annotated RP type and predicted RP type were compared.
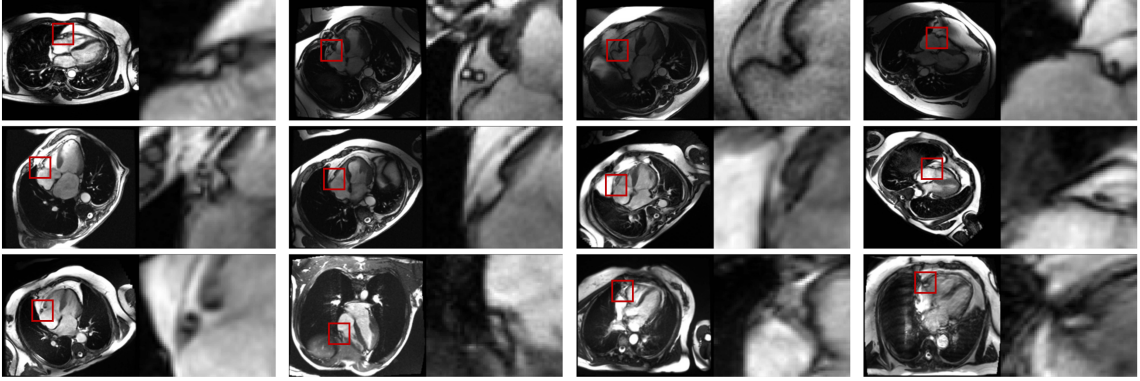
Figure 3: The robustness of the RCA detection network is shown in a subset of the unseen dataset. For each case, the 4CH input is shown with a bounding box in red and the cropped image next to it for a total of 12 cases.

To evaluate the robustness of the proposed system, different CINE sequences (Cartesian segmented, Cartesian segmented with small field-of-view, Cartesian segmented Compressed Sensing Prototype (CS), Cartesian Real-time CS, Radial real-time) were acquired and the predicted RPs of each sequence were compared with each of the expert annotation. Further, the computation time of the proposed system was measured at the beginning and the end of the proposed system.

3. Results

3.1 Localization

The mean and standard deviation error between the prediction and ground truth of the fully convolutional 3-D DenseNet with 122 layers was 4.6 ± 1.8 mm. The box plot in Figure 5, shows the Distance Error in mm between the $\hat{p}$ and the p in each frame. Robustness results for unseen datasets are presented in two ways, first in Figure 3 which shows the performance on oblique and diverse oriented cases (n=12) and second in Figure 4, which visualized the worst cases. The quantitative localization results of a part of the study dataset (n=9) acquired with breath-hold and free-breathing CINE sequences are shown in Figure 6 above. The first frame of each CINE series and the corresponding RCA cropped series are shown. Each case was visualized with the first frame of the corresponding CINE series marked with a red box showing the position of ROI defined based on the network prediction and beside it with the cropped series enclosing the area with the outer edge of the cross-section of the RCA overlayed with the generated heatmap.

3.2 Motion Quantification

The motion values obtained by calculating the distance between the predicted pixel co-ordinates in each adjacent time points achieved 61.1% accuracy for 1.5T, and 52.8% for 3T. As the motion quantification analysis in the Table 3 shows, the 50th percentile/median of Gaussian weighting achieved 90.1% accuracy, whereas the accuracy was 87.2% without
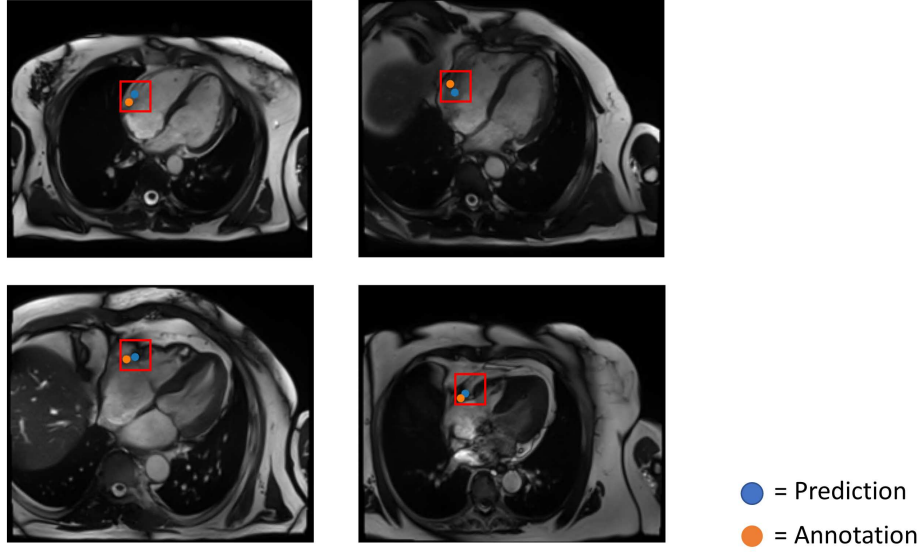
Figure 4: The worst cases on the test dataset of the RCA detection network are visualized (n=4). The blue dot represents the predicted landmark position, the orange dot, the expert's annotation and the red box, the ROI defined based on the predicted landmark positions, respectively. Despite incorrect detection, the annotation landmark is still invariably included in all ROI windows.
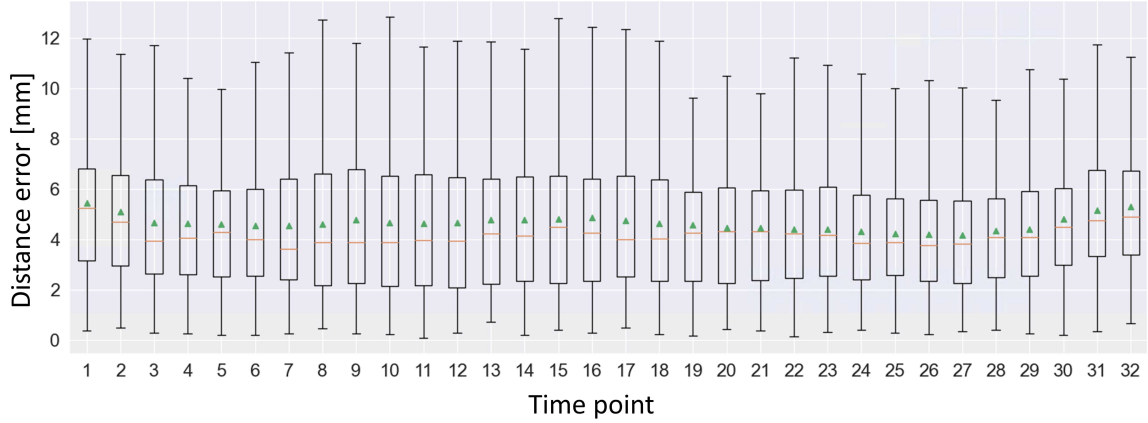

Figure 5: A box plot showing the performance of the 122-layer 3-D DenseNet in each time point is illustrated. Orange line represents the median value and the green triangle, the mean value.

weighting the deformation field. The median performed 91.0% accuracy for 1.5T and 88.9% for 3T. The best accuracy achieved by taking the mean metric was 89.3%, without Gaussian weighting. The motion values quantified based on the median approach with Gaussian weighting in 9 clinically acquired dataset are shown in Figure 6 bottom.
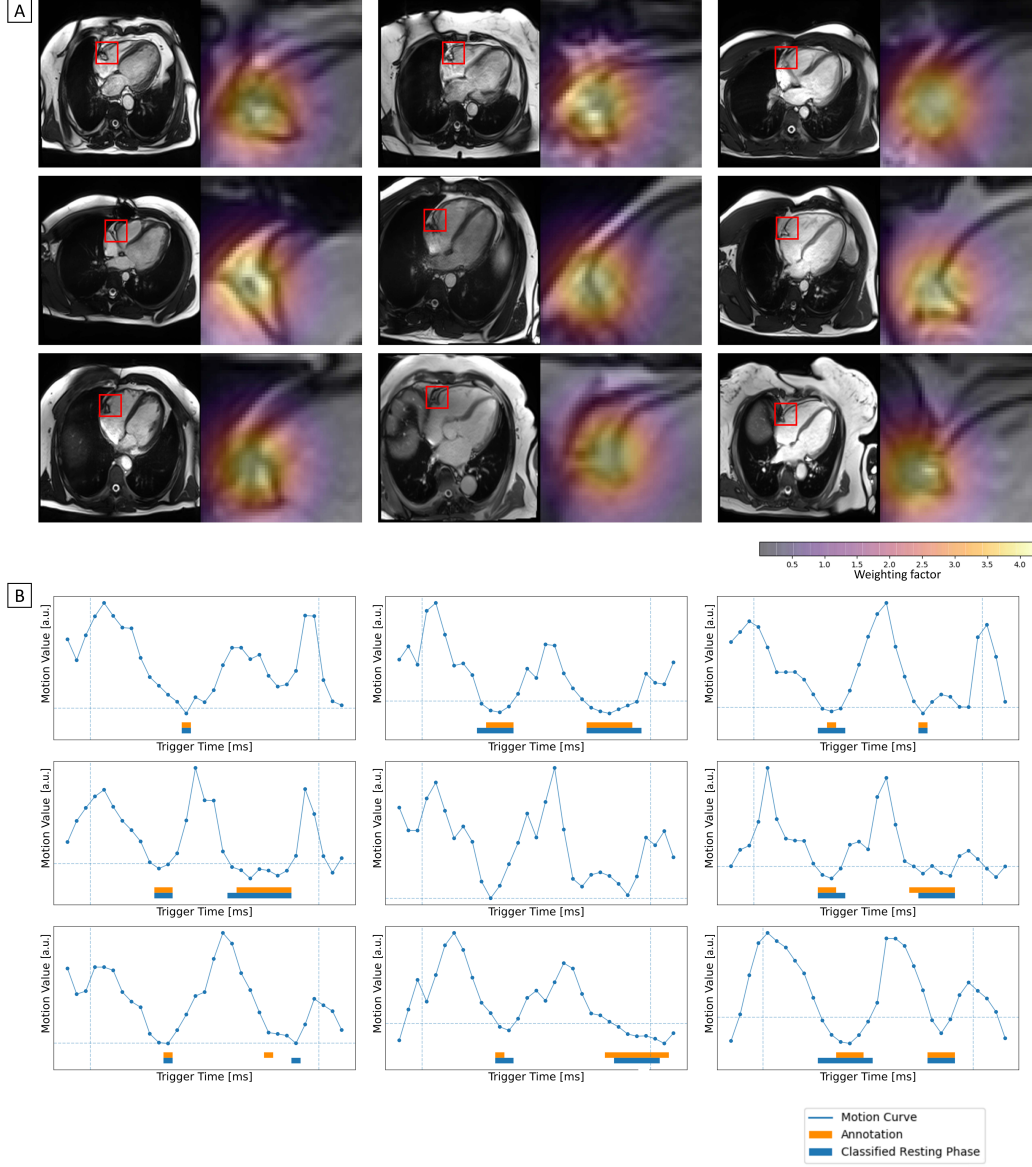
Figure 6: Overview of the first frame of 9 different RCA series highlighted with a generated heatmap, which center point is taken from the predicted coordinate by the 3-D DenseNet trained for detecting the RCA pixel coordinate (A). The quantified motion value and classified RPs corresponding to each dataset are shown below (B). The vertical dashed line represents the window of interest as described in 2.5. The horizontal dashed line represents the selected threshold value. The lower the motion value, the less motion exists in the frame.

Table 3: The analysis of motion quantification. Two approaches by using the Gaussian weighting of the magnitude of the deformation fields, or without any weighting have been compared for varying the percentile and taking the mean value.

Metric	Percentile	Accuracy	Sensitivity	Specificity	Threshold [a.u.]
m_dist	-	71.4	50.5	92.3	0.50
m_pct	10th	75.1	55.5	94.6	0.08
m_pct	20th	81.5	67.3	95.8	0.13
m_pct	30th	84.1	72.2	96.0	0.18
m_pct	40th	86.9	77.6	96.2	0.23
m_pct	50th	87.2	78.1	96.4	0.28
m_pct	60th	87.7	80.1	95.3	0.34
m_pct	70th	86.5	76.2	96.7	0.38
m_pct	80th	83.2	69.0	97.5	0.42
m_pct	90th	81.7	65.8	97.7	0.49
m_pct	100th	78.3	61.6	95.1	0.68
m_mean	-	89.3	84.7	93.9	0.34
m_wpct	10th	62.4	29.5	95.2	0.01
m_wpct	20th	72.8	51.9	93.6	0.03
m_wpct	30th	82.7	74.2	91.2	0.07
m_wpct	40th	87.4	80.8	94.0	0.12
$\mathbf{m_\mathbf{wpct}}$	$\mathbf{50^\mathbf{th}}$	**90.1**	**85.4**	**94.8**	**0.20**
m_wpct	60th	88.0	78.9	97.2	0.28
m_wpct	70th	85.8	74.4	97.2	0.42
m_wpct	80th	85.1	73.4	97.0	0.62
m_wpct	90th	83.6	71.2	96.0	0.98
m_wpct	100th	66.4	33.2	99.5	1.0
m_wmean	-	86.2	76.6	95.8	0.36

3.3 Threshold Analysis

For each percentile and mean analysis, the threshold τ selected based on binary classification task is listed in the right column in the Table 3. From the motion values obtained by taking the median, the selected τ was 0.2. The resulting ROC curve is plotted in Figure 7, and the accuracy over each threshold step is plotted in Figure 7 in the top row, in which the threshold τ is marked by an orange vertical line. On the below row in Figure 7 shows the confusion matrices of each annotated datasets. On the above one, the performance of the threshold on the testing dataset is shown, while on the below one the performance of the threshold on the study dataset is displayed.

3.4 RP Classification

The detailed results about the performance of the predicted systolic and diastolic RP on the study datasets are listed in Table 4, Table 5. The Bland-Altmann plots showing the performance of start and end detected time point for each systolic and diastolic RP is shown in Figure 8. $MAE_{start,end\ window/frame_{sys}}$ and $MAE_{start,end\ window/frame_{dia}}$ for 1.5T datasets was 11.8 ± 16.5 ms (0.38 ± 0.53 frame) and 14.2 ± 18.9 ms (0.48 ± 0.63 frame) for 3T datasets. By using the selected τ, the proposed system resulted in 93.4% accuracy, sensitivity at 90.1% and specificity 96.8% for 1.5T and 92.6%, 90.7% and 94.5% for 3T
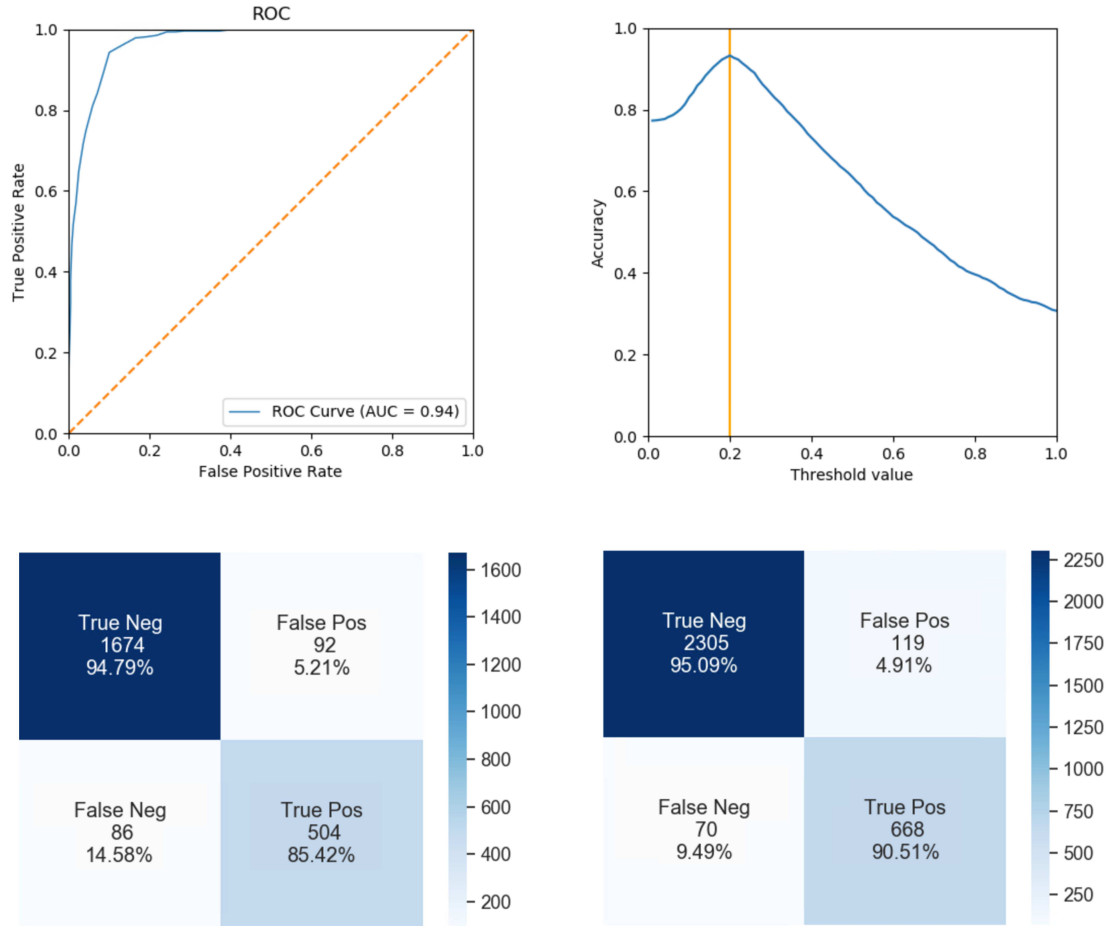
Figure 7: Upper Left: the ROC curve from the threshold analysis is shown. Upper Right: the accuracy plot over the threshold values is shown. The best performed threshold value, 0.2 is marked in orange vertical line. Below Left: the confusion matrix showing the performance of binary classification task by taking the best threshold value on testing datasets. Below Right: the confusion matrix showing the performance by taking the best threshold value on study datasets. In each measure, the counts and rate in percentage are listed.

Table 4: The results of the system on the unseen study datasets are listed. In the first row, the difference between the start and end RP and expert annotations are shown. In the second row, the difference of start and end systolic and diastolic RPs are shown in frame.

Magnetic Field	Number datasets	Threshold	Start Systolic RP $[ms]$	End Systolic RP $[ms]$	Start Diastolic RP $[ms]$	End Diastolic RP $[ms]$
1.5 T	N=22	0.20	14.4 ± 19.8	15.0 ± 16.9	5.7 ± 13.1	12.2 ± 16.1
3 T	N=80	0.20	19.7 ± 22.4	12.2 ± 16.6	15.8 ± 21.3	8.9 ± 15.2
1.5 T & 3 T	N=102	0.20	18.7 ± 22.1	12.7 ± 16.7	13.2 ± 20.0	9.7 ± 15.5

Magnetic Field	Number datasets	Threshold	Start Systolic RP $[Frame]$	End Systolic RP $[Frame]$	Start Diastolic RP $[Frame]$	End Diastolic RP $[Frame]$
1.5 T	N=22	0.20	0.46 ± 0.63	0.54 ± 0.63	0.16 ± 0.36	0.37 ± 0.48
3 T	N=80	0.20	0.70 ± 0.82	0.43 ± 0.59	0.49 ± 0.63	0.28 ± 0.49
1.5 T & 3 T	N=102	0.20	0.66 ± 0.80	0.45 ± 0.60	0.40 ± 0.59	0.31 ± 0.49

Table 5: Accuracy, sensitive and specificity are listed with the optimally defined threshold.

Magnetic Field	Number datasets	Threshold	Accuracy	Sensitivity	Specificity
1.5 T	N=22	0.20	93.4	90.1	96.8
3 T	N=80	0.20	92.6	90.7	94.5
1.5 T & 3 T	N=102	0.20	92.7	90.5	95.0

datasets. $MAE_{start,end\ window/frame_{sys}}$ and $MAE_{start,end\ window/frame_{dia}}$ was 13.6 ± 18.6 ms (0.46 ± 0.62 frame) when using the datasets independent from field strength for analysis. The accuracy, sensitivity and specificity achieved by the defined absolute τ, <0.2, was 92.7%, 90.5% and 95.0%. The automatically classified RPs resulted in a mean max error of 30 ms, meaning that it deviates by roughly one frame. The datasets with RPs with less than 30 ms were discarded (n=9). Further, there was no systolic RP annotated by the expert in 14 cases, and no diastolic RP in 12 cases. These phases were excluded from the analysis. The $\widehat{RP}_{start\ window_{sys}}$ matched with the annotation, or was off one frame in 93.6%, the $\widehat{RP}_{end\ window_{sys}}$ in 97.5%, the $\widehat{RP}_{start\ window_{dia}}$ in 93.8% and $\widehat{RP}_{end\ window_{dia}}$ in 96.3%, respectively. $\widehat{RP}_{start\ window_{sys}}$ was detected earlier/later than the expert's annotation in 27.8%/5% and $\widehat{RP}_{start\ window_{dia}}$ in 20.9%/16.0%. $\widehat{RP}_{end\ window_{sys}}$ was detected earlier/later than the expert's annotation in 11.4%/29.1% and $\widehat{RP}_{end\ window_{end}}$ in 15.2%/14.8%. In average, the $\widehat{RP}_{start\ window}$ was selected earlier than the ground truth in 24.3% of the cases and later in 10.6%. In 10 cases, outliers were present, off by 2 or more frames. In 10 cases, there were outliers present for the $\widehat{RP}_{start\ window}$, off by 2 or more frames. For the $\widehat{RP}_{end\ frame}$, it was off by 2 or more frames in 5 cases.

The range of annotated systolic RP was 61.1 ± 24.1 ms and the predicted range of systolic RP was 75.5 ± 32.9 ms. Further the range of annotated diastolic RP was 156.0 ± 102.1 ms
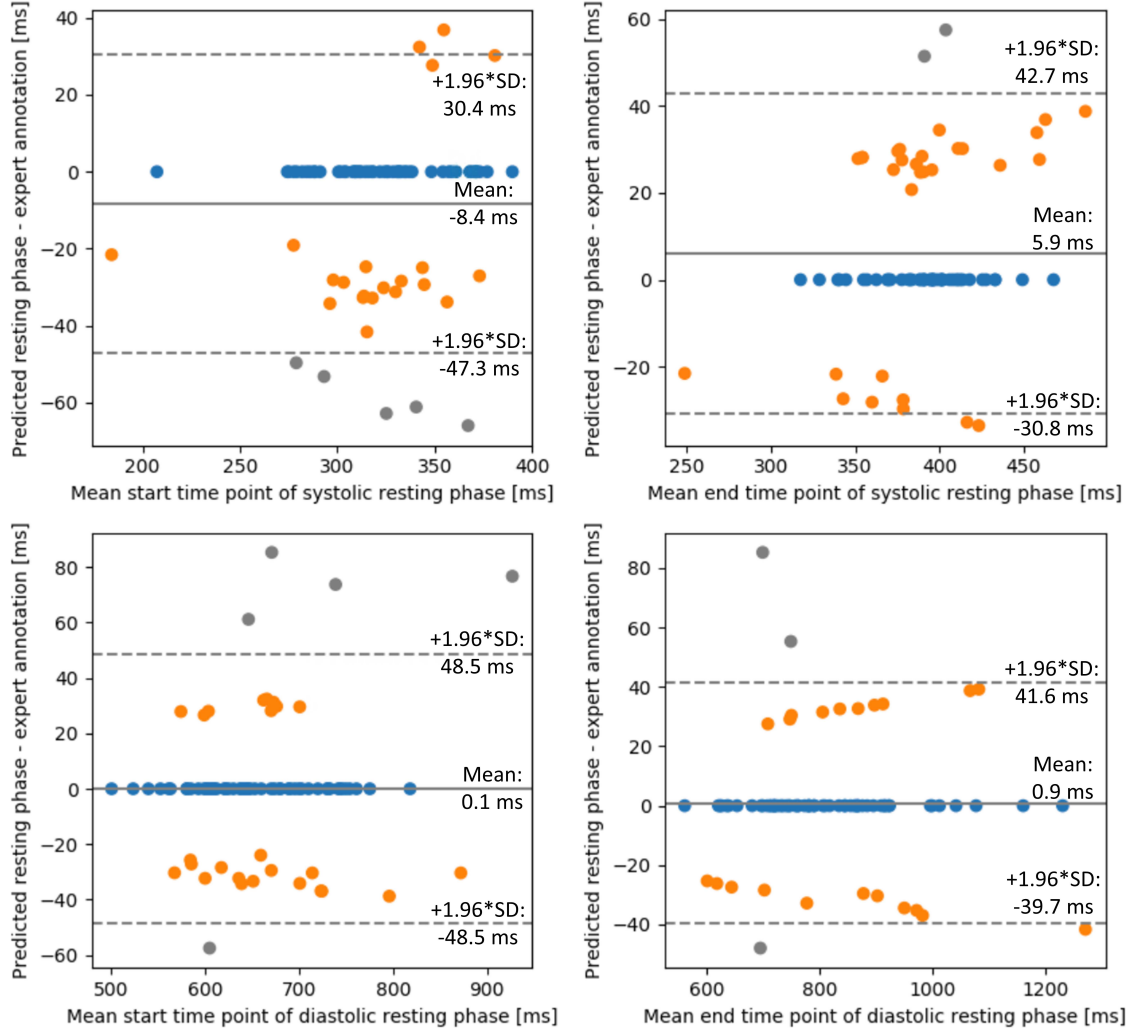
Figure 8: Top: the difference between the predicted and expert annotations of start and end systolic RP are shown in Bland-Altmann plot. Bottom: the difference between the predicted and expert annotations of start and end diastolic RP are shown. The blue dots represent the exact match between the predicted and annotation, and the orange dots show when there is one frame difference. The gray dots represent when the difference is more than 2 frames.
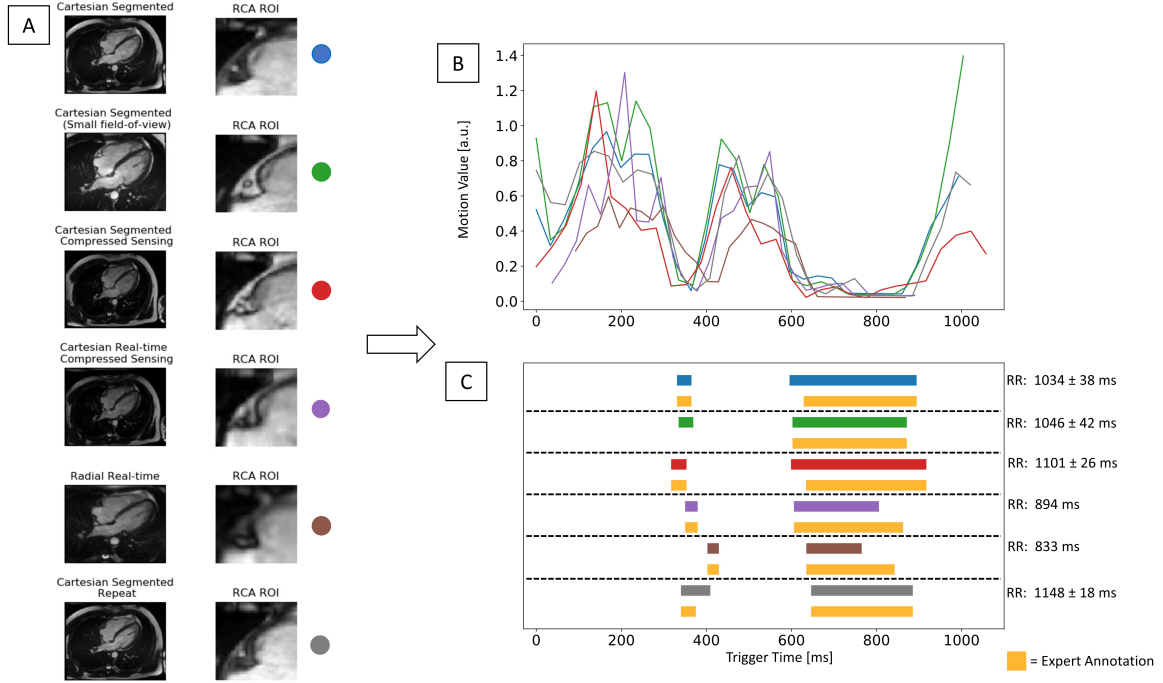
Figure 9: An example of a study case with 6 different CINE sequences. On the left (A), the first frame of each CINE series and the output of the cropped series are shown. Each color represents one sequence acquisition. On the top right (B), the quantified motion values of each RCA cropped series are plotted over the time points. On the bottom right (C), the classified RPs of each are shown with the expert's annotation, which is marked in orange color. The RR interval of the real-time CINE acquisitions varied from the segmented CINE, which caused early detection of the RPs.

and the predicted range was 158.2 ± 104.3 ms. In Figure 9, the robustness of the proposed system is shown in which the system was tested in different sequences including a rescan from a single volunteer. The visualized different CINE outputs are acquired in the order from top to bottom and with 5min between the first and last acquisition. The predicted start and end systolic phases were matched with annotation in most cases, except in one CINE sequence, where the systolic RP was detected by the system but not by the expert, and in a repeat scan, the end time point was off one frame. The start diastolic RP was detected one frame earlier in 2 cases, and end diastolic RP was detected two frames off in real-time sequences. In an example case, the automatically detected RPs were used for the later 3-D static cardiac acquisition targeted to the RCA. The 3-D RCA visualization with the automatically classified RPs showed no residual motion artifacts (Figure 10). The computation time of the proposed system was averaged 1.5 seconds.

4. Discussion

The detection of the RCA ROI is successfully and robustly performed by the 3-D DenseNet on the testing dataset and on the study cases. The 3-D based Conv networks leverage the
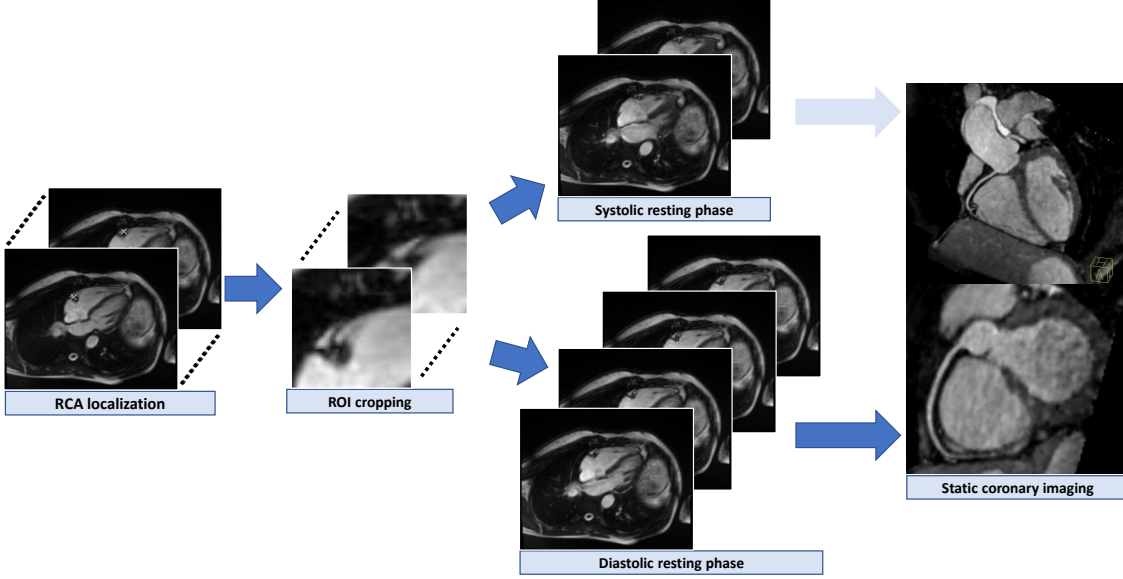
Figure 10: An example of a volunteer study illustrated with the main steps of the framework pipeline. The outputs are generated directly from the scanner after the proposed system was integrated online. The RCA localization series is the original CINE series with an RCA position marked by a cross. The ROI cropping generated the cropped series based on the RCA localization. From these cropped series, the motion is quantified, from which the RPs are determined. The series that represent the systolic and diastolic resting phase are generated as well. Here, the diastolic resting phase window (dark blue arrow) is applied for the static coronary imaging.

spatial and temporal information from the time-resolved input, rather than learning only spatial information per time point. The size of the fixed ROI (50×50mm^2) was sufficient in all cases for depicting the RCA at each time point. Further, the network was robustly performed on diverse oriented cases (Figure 3) and different CINE sequences (Figure 9) which allows the proposed system to be integrated into different clinical protocols.

In terms of quantifying motion values, the approach of taking the Euclidean distance from the predicted RCA pixel coordinate over cardiac phases highly relies on the performance of the network, furthermore, taking the pixel distance measurement for the displacement metric of the anatomy-of-interest between the consecutive frames was not accurate as shown in Table 3. The approaches deriving from the motion values by taking the deformation fields defined by the elastic image registration show a clear advantage, from which the highest accuracy was the one using the weighted deformation fields (see in Table 3). The approach with deformation fields is clearly more robust to slight inaccuracies of localization results. The Gaussian weighting further improves the performance of the system, as it allows to focus on the target-of-interest, and eliminates the area which is not of interest, such as the blood flow in the atrium contained in the RCA ROI.

The metric for assessing the motion values from the weighted deformation fields was rea-

sonably chosen as 50th percentile by the accuracy analysis. The absolute threshold value is selected based on AUC-ROC analysis, evaluated on the testing datasets from the different 1.5T, 3T scanners and CINE sequences and no specific data selection, thus shows versatility in the results (Figure 7). The classification of the RP by the absolute threshold value is possible due to the quantitative outputs of the deformation field-based approach which enables the detection of the phases with minimal motion, which can be either end-systolic or mid-diastolic or end-systolic and mid-diastolic RPs. In Figure 7, the quantified motion values of the clinically acquired dataset with different CINE sequences and the corresponding predicted RPs were well matched with the expert's annotation. As shown in Table 4, Table 5, the proposed framework performed robust in different field strengths as well. Interestingly, in the dataset visualized in the center, there was no RP found by a medical expert, and the proposed system was also able to classify the non-RP case demonstrating the advantage of the system. In such cases, the system gives the user the possibility to take the quantified curve as a reference and select the phase with minimum motion based on the motion curve.

Based on the Bland-Altmann plots (Figure 8), the systolic RP did not perform as well as the diastolic RP, especially the classification of the start systolic RP was challenging. The predicted systolic RP window was usually slightly longer than the experts' annotations. However, the mean range difference was 15ms, which is negligible since the temporal resolution of CINE series is 30ms. In a recent study, the proposed system was evaluated in a clinical validation, in which the interobserver variability was extensively validated and revealed that the automatically detected RPs were consistent with those of the experts (Ogawa et al., 2022). As shown in Figure 10, the RCA was sharply visualized without severe residual motion artifacts that was acquired during the automatically detected RPs.

The study presents some limitations. First, the proposed method depends on two factors: the performance of the RCA localization network and the registration algorithm that may lead to inaccurate results in the presence of severe artifacts in reconstructed CINE series, even though the network performed well on difficult cases as shown in Figure 3 and Figure 4. Second, the absolute thresholding used for classification, can be further investigated, whether other algorithms can be used for binary classification. Third, in this study, the focus was highly on method evaluation on patient dataset, however the clinical feasibility and interobserver variability study using this method was performed in (Ogawa et al., 2022). In the following research, a further clinical study can be of interest to compare the image quality between once acquired with the automatic approach and the other with manually determined resting phases.

Previous methods have shown the feasibility to detect the imaging acquisition window automatically using the shim box volume positioning and cross-correlation calculation (Jahnke et al., 2005; Ustun et al., 2007). However, these semi-automated methods still require the careful positioning of the shim volume coverage. Further, the RP detection is targeted to the whole heart instead of a specific anatomy of interest. These methods were validated on healthy in vivo subjects on a single field strength. Moreover, approaches based on the standard deviation of pixel intensity (Huang et al., 2014) or difference of gradient magnitudes (Piccini et al., 2017) were introduced, allowing the RP detection in real time. This method however performs the RP detection globally based on the entire field-of-view, and the detected RPs are always two RPs as the search is done by two local minima. Several

approaches have been proposed for the automated determination of targeted RP on regions such as RCA (Sato et al., 2009; Asou et al., 2018). A template matching was performed for finding the area with the outer edge of the cross-section of the coronary artery, however the template was defined based on randomly selected five datasets (Sato et al., 2009). In a more recent study, the regions were extracted from the high-speed component within CINE series by the frequency domain analysis, however by doing so, the cardiac anatomy structures can be disregarded. The authors stated that this method was validated on healthy volunteers, and uncertain the performance on large dataset especially with high heart rates. In Adam et al. (2020), the deep learning-based RP detection network is built by combining the CNN and Long short-term memory models taking the CINE series as input, and outputs the binary output which is either a RP or not a RP. Therefore, this method is not quantitative and unclear which cardiac structure is weighted, or whether the network tries to detect global RP.

5. Conclusions

To our knowledge, this proposed study is the first to present the fully automated localized RP detection framework from a CINE series that was validated with a large dataset with multiple 1.5T and 3T scanners acquired with different CINE protocols, such as with free-breathing or breath-held techniques. We investigated the robustness and feasibility of the proposed system for fully automated systolic and diastolic RP detection. The proposed system can improve the workflow efficiency, automation, and standardization of the static cardiac imaging that broaden the applicability towards any static cardiac imaging. The RP detection system can be applied in various applications, such as 2-D and 3-D LGE, mapping, 3-D coronary imaging, or any other applications in which the information of RP of heart can be useful. Future work will focus on clinical validations, improving the accuracy of RP classification and integration of automatic detection of other regions, such as the atria and ventricles.

Ethical Standards

The local ethics committee approved the study, and a written informed consent was provided by all subjects. The declaration of consent to participate in the study was signed by all subjects.

Conflicts of Interest

The authors declare no conflict of interest and no financial disclosure in relation to this study. A. Maier is Associate Editor of the Journal of Machine Learning for Biomedical

References

Dima Abi-Abdallah, Vincent Robin, Agnès Drochon, and Odette Fokapu. Alterations in human ecg due to the magnetohydrodynamic effect: a method for accurate r peak detection in the presence of high mhd artifacts. In 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 1842–1845. IEEE, 2007.

Naledi Adam, James Clough, Ronald Mooiweer, Phuoc Duong, Li Huang, Reza Razavi, Kuberan Pushparajah, Amedeo Chiribiri, Andrew King, and Sébastien Roujol. Fully automated detection of the quiescent phases of the cardiac cycle from cine images using deep learning. In Proceedings of the Joint Annual Meeting ISMRM-ESMRMB (28th Annual Meeting & Exhibition), IS for Magnetic Resonance in Medicine, Ed, 2020.

Emily Aherne, Kelvin Chow, and James Carr. Cardiac t1 mapping: techniques and applications. Journal of Magnetic Resonance Imaging, 51(5):1336–1356, 2020.

Mehmet Akçakaya, Hussein Rayatzadeh, Tamer A Basha, Susie N Hong, Raymond H Chan, Kraig V Kissinger, Thomas H Hauser, Mark E Josephson, Warren J Manning, and Reza Nezafat. Accelerated late gadolinium enhancement cardiac mr imaging with isotropic spatial resolution using compressed sensing: initial experience. Radiology, 264(3):691–699, 2012.

Hiroya Asou, Naoyuki Imada, Yuichi Nishiyama, Tomoyasu Sato, and Katsuhiro Ichikawa. Automated determination of cardiac rest period on whole-heart coronary magnetic resonance angiography by extracting high-speed motion of coronary arteries. Clinical imaging, 52:183–188, 2018.

Tamer A Basha, Mehmet Akcakaya, Charlene Liew, Connie W Tsao, Francesca N Delling, Gifty Addae, Long Ngo, Warren J Manning, and Reza Nezafat. Clinical performance of high-resolution late gadolinium enhancement imaging with compressed sensing. Journal of Magnetic Resonance Imaging, 46(6):1829–1838, 2017.

Christophe Chefd'Hotel, Gerardo Hermosillo, and Olivier Faugeras. Flows of diffeomorphisms for multimodal image registration. In Proceedings IEEE International Symposium on Biomedical Imaging, pages 753–756. IEEE, 2002.

Gastao Cruz, David Atkinson, Markus Henningsson, Rene M Botnar, and Claudia Prieto. Highly efficient nonrigid motion-corrected 3d whole-heart coronary vessel wall imaging. Magnetic resonance in medicine, 77(5):1894–1908, 2017.

Roberto M Dyke, Yu-Kun Lai, Paul L Rosin, and Gary KL Tam. Non-rigid registration under anisotropic deformations. Computer Aided Geometric Design, 71:142–156, 2019.

Christoph Forman, Davide Piccini, Robert Grimm, Jana Hutter, Joachim Hornegger, and Michael O. Zenge. Reduction of Respiratory Motion Artifacts for Free-Breathing Whole-Heart Coronary MRA by Weighted Iterative Reconstruction. Magnetic Resonance in Medicine, pages 1–11, 2014. doi: 10.1002/mrm.25321. URL https://www5.informatik.uni-erlangen.de/Forschung/Publikationen/2014/Forman14-ROR.pdf.

Gerald Greil, Animesh Aashoo Tandon, Miguel Silva Vieira, and Tarique Hussain. 3d whole heart imaging for congenital heart disease. Frontiers in pediatrics, 5:36, 2017.

Mark BM Hofman, Samuel A Wickline, and Christine H Lorenz. Quantification of in-plane motion of the coronary arteries during the cardiac cycle: implications for acquisition window duration for mr flow quantification. Journal of Magnetic Resonance Imaging, 8 (3):568–576, 1998.

Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4700–4708, 2017.

Teng-Yi Huang, Yu-Sheng Tseng, and Tzu-Chao Chuang. Automatic calibration of trigger delay time for cardiac mri. NMR in Biomedicine, 27(4):417–424, 2014.

Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning, pages 448–456. PMLR, 2015.

Hussain Isma'eel, Yasmin S Hamirani, Ramona Mehrinfar, Songshuo Mao, Naser Ahmadi, Vahid Larijani, Subu Nair, and Matthew J Budoff. Optimal phase for coronary interpretations and correlation of ejection fraction using late-diastole and end-diastole imaging in cardiac computed tomography angiography: implications for prospective triggering. The international journal of cardiovascular imaging, 25(7):739–749, 2009.

Cosima Jahnke, Ingo Paetsch, Kay Nehrke, Bernhard Schnackenburg, Axel Bornstedt, Rolf Gebker, Eckart Fleck, and Eike Nagel. A new approach for rapid assessment of the cardiac rest period for coronary mra. Journal of Cardiovascular Magnetic Resonance, 7 (2):395–399, 2005.

Shuiwang Ji, Wei Xu, Ming Yang, and Kai Yu. 3d convolutional neural networks for human action recognition. IEEE transactions on pattern analysis and machine intelligence, 35 (1):221–231, 2012.

Kevin R Johnson, Salil J Patel, Amy Whigham, Alex Hakim, Roderic I Pettigrew, and John N Oshinski. Three-dimensional, time-resolved motion of the coronary arteries. Journal of Cardiovascular Magnetic Resonance, 6(3):663–673, 2004.

Peter Kellman and Andrew E Arai. Cardiac imaging techniques for physicians: late enhancement. Journal of magnetic resonance imaging, 36(3):529–542, 2012.

Peter Kellman and Michael S Hansen. T1-mapping in the heart: accuracy and precision. Journal of cardiovascular magnetic resonance, 16(1):1–20, 2014.

23

Peter Kellman, Andrew E Arai, Elliot R McVeigh, and Anthony H Aletras. Phase-sensitive inversion recovery for detecting myocardial infarction using gadolinium-delayed hyperenhancement. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 47(2):372–383, 2002.

W Yong Kim, Matthias Stuber, Kraig V Kissinger, Niels Trolle Andersen, Warren J Manning, and René M Botnar. Impact of bulk cardiac motion on right coronary mr angiography and vessel wall imaging. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine, 14(4):383–390, 2001.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

Christopher M Kramer, Jörg Barkhausen, Chiara Bucciarelli-Ducci, Scott D Flamm, Raymond J Kim, and Eike Nagel. Standardized cardiovascular magnetic resonance imaging (cmr) protocols: 2020 update. Journal of Cardiovascular Magnetic Resonance, 22(1): 1–18, 2020.

CM Kramer, J Barkhausen, SD Flamm, RJ Kim, and E Nagel. Society for cardiovascular magnetic resonance board of trustees task force on standardized p. standardized cardiovascular magnetic resonance (cmr) protocols 2013 update. J Cardiovasc Magn Reson, 15 (1):91, 2013.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25: 1097–1105, 2012.

Daniel R Messroghli, James C Moon, Vanessa M Ferreira, Lars Grosse-Wortmann, Taigang He, Peter Kellman, Julia Mascherbauer, Reza Nezafat, Michael Salerno, Erik B Schelbert, et al. Clinical recommendations for cardiovascular magnetic resonance mapping of t1, t2, t2* and extracellular volume: a consensus statement by the society for cardiovascular magnetic resonance (scmr) endorsed by the european association for cardiovascular imaging (eacvi). Journal of Cardiovascular Magnetic Resonance, 19(1):1–24, 2017.

Camila Munoz, Aurelien Bustin, Radhouene Neji, Karl P Kunze, Christoph Forman, Michaela Schmidt, Reza Hajhosseiny, Pier-Giorgio Masci, Martin Zeilinger, Wolfgang Wuest, et al. Motion-corrected 3d whole-heart water-fat high-resolution late gadolinium enhancement cardiovascular magnetic resonance imaging. Journal of Cardiovascular Magnetic Resonance, 22(1):1–13, 2020.

Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In Icml, 2010.

Ryo Ogawa, Tomoyuki Kido, Yasuhiro Shiraishi, Yuri Yagi, Seung Su Yoon, Jens Wetzl, Michaela Schmidt, and Teruhito Kido. Neural network–based fully automated cardiac resting phase detection algorithm compared with manual detection in patients. Acta Radiologica Open, 11(10):20584601221137772, 2022.

Davide Piccini, Robin Demesmaeker, Gabriella Vincenti, Tobias Kober, and Matthias Stuber. Automated cardiac resting phase detection in 2d cine mr images for acquisition window selection in high-resolution coronary mri. Prod. Intl. Soc. Mag. Reson. Med, 25: 2862, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.

Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. IEEE transactions on medical imaging, 18(8):712–721, 1999.

Tetsuo Sato, Tomohisa Okada, Shigehide Kuhara, Kaori Togashi, and Kotaro Minato. An approach for automatic selecting of optimal data acquisition window for magnetic resonance coronary angiography. In Medical Imaging 2009: Image Processing, volume 7259, page 72592A. International Society for Optics and Photonics, 2009.

Harald Seifarth, Susanne Wienbeck, Michael Pusken, Kai-Uwe Juergens, David Maintz, Christian Vahlhaus, Walter Heindel, and Roman Fischbach. Optimal systolic and diastolic reconstruction windows for coronary ct angiography using dual-source ct. American Journal of Roentgenology, 189(6):1317–1323, 2007.

Guy Shechter, Jon R Resar, and Elliot R McVeigh. Rest period duration of the coronary arteries: implications for magnetic resonance coronary angiography. Medical physics, 32 (1):255–262, 2005.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Adrian Spinei, Denis Pellerin, and Jeanny Hérault. Spatiotemporal energy-based method for velocity estimation. Signal processing, 65(3):347–362, 1998.

Matthias Stuber, René M Botnar, Peter G Danias, Kraig V Kissinger, and Warren J Manning. Submillimeter three-dimensional coronary mr angiography with real-time navigator correction: comparison of navigator locations. Radiology, 212(2):579–587, 1999.

Jonathan D Suever, Pierre J Watson, Robert L Eisner, Stamatios Lerakis, Robert E O'Donnell, and John N Oshinski. Time-resolved analysis of coronary vein motion and cross-sectional area. Journal of Magnetic Resonance Imaging, 34(4):811–815, 2011.

John A Swets. Roc analysis applied to the evaluation of medical imaging techniques. Investigative radiology, 14(2):109–121, 1979.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015.

Richard Szeliski and Heung-Yeung Shum. Motion estimation with quadtree splines. IEEE Transactions on pattern analysis and machine intelligence, 18(12):1199–1210, 1996.

Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, and Manohar Paluri. Learning spatiotemporal features with 3d convolutional networks. In Proceedings of the IEEE international conference on computer vision, pages 4489–4497, 2015.

Ali Ustun, Milind Desai, Khaled Z Abd-Elmoniem, Michael Schar, and Matthias Stuber. Automated identification of minimal myocardial motion for improved image quality on mr angiography at 3 t. American Journal of Roentgenology, 188(3):W283–W290, 2007.

YI Wang, Richard Watts, Ian R Mitchell, Thanh D Nguyen, Jeffrey W Bezanson, Geoffrey W Bergman, and Martin R Prince. Coronary mr angiography: selection of acquisition window of minimal cardiac motion with electrocardiography-triggered navigator cardiac motion prescanning—initial results. Radiology, 218(2):580–585, 2001.