Motion Artifact Recognition and Quantification in Coronary CT Angiography using Convolutional Neural Networks

T. Lossau (née Elss)[a,b], H. Nickisch[a], T. Wissel[a], R. Bippus[a], H. Schmitt[a], M. Morlock[b], M. Grass[a]

[a] *Philips Research, Hamburg, Germany*
[b] *Hamburg University of Technology, Germany*

Abstract

Excellent image quality is a primary prerequisite for diagnostic non-invasive coronary CT angiography. Artifacts due to cardiac motion may interfere with detection and diagnosis of coronary artery disease and render subsequent treatment decisions more difficult. We propose deep-learning-based measures for coronary motion artifact recognition and quantification in order to assess the diagnostic reliability and image quality of coronary CT angiography images. More specifically, the application, steering and evaluation of motion compensation algorithms can be triggered by these measures. A *Coronary Motion Forward Artifact model for CT data* (CoMoFACT) is developed and applied to clinical cases with excellent image quality to introduce motion artifacts using simulated motion vector fields. The data required for supervised learning is generated by the CoMoFACT from 17 prospectively ECG-triggered clinical cases with controlled motion levels on a scale of 0 to 10. Convolutional neural networks achieve an accuracy of $93.3\% \pm 1.8\%$ for the classification task of separating motion-free from motion-perturbed coronary cross-sectional image patches. The target motion level is predicted by a corresponding regression network with a mean absolute error of 1.12 ± 0.07. Transferability and generalization capabilities are demonstrated by motion artifact measurements on eight additional CCTA cases with real motion artifacts.

Keywords: Cardiac CT, Motion Artifact Measure, Coronary Angiography, Convolutional Neural Network

1. Introduction

Non-invasive coronary computed tomography angiography (CCTA) has become a preferred technique for the detection and diagnosis of coronary artery disease (CAD) (Budoff et al., 2017; Foy et al., 2017; Camargo et al., 2017; Liu et al., 2017), but high quality imaging for small and moving vessels is still challenging. ECG-controlled acquisition is used to enable the reconstruction of heart phases with small motion level and gating windows are limited to the temporal projection range required for back-projection. However, hardware constraints restrict the temporal resolution of the reconstructed CT image volumes. Despite ECG-triggering and -gating, cardiac motion frequently leads to artifacts in the reconstructed CT image volumes (Ghekiere et al., 2017). These artifacts manifest in typical patterns containing intensity undershoots and arc-shaped blurring due to the CT reconstruction geometry (see Figure 1) and potentially limit or even preclude the evaluation of parts of coronary arteries or cause misinterpretations.

Thus, motion correction algorithms to improve image quality of the coronary arteries have been an important research area for years. Several approaches have been developed which are based on motion estimation via 3-D/3-D registration of multiple heart phases and subsequent motion-compensated filtered back-projection (MC-FBP) (van Stevendaal et al., 2008; Isola et al., 2010; Bhagalia et al., 2012). An iterative motion compensation approach dealing with motion vector field (MVF) estimation by minimization of handcrafted motion artifact measures (MAMs) has been introduced by Rohkohl et al. (2013).

Due to possible failure modes and their substantial computational footprint, motion correction methods can benefit from a reliable measure of motion artifacts. First, the recognition and quantification of motion artifacts in the coronary artery tree during CCTA could decide whether and where mo-

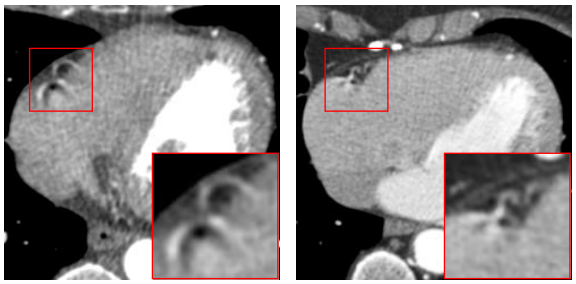
tion correction is required to enable the diagnosis of CAD and prevent misinterpretations. Second, measures for motion artifacts could be used to verify the success of a motion compensation method. Finally, an integration of motion artifact measures in the motion compensation process as shown by Rohkohl et al. (2013) is conceivable.

Furthermore, assessment of the scan quality might be useful in automatic analysis of the coronary arteries, e.g. by reporting on the reliability of the coronary artery calcium score. A deep-learning-based system for the identification of coronary artery calcifications which are strongly affected by cardiac motion artifacts has been introduced by Šprem et al. (2017).

Most handcrafted measures for motion artifacts (Rohkohl et al., 2013; McGee et al., 2000) such as entropy and positivity are best suited for relative assessment, i.e. for the comparison of the same image region at different motion states. An absolute measure for consistent artifact quantification across patients and vessel segments has to be robust to possible variations in noise level, background intensity, vessel structure and contrast agent enhancement. The ability of five handcrafted metrics to quantify absolute motion artifact levels at the coronary arteries has been investigated by Ma et al. (2018). Beside the MAMs entropy and positivity, the three metrics normalized circularity, Fold Overlap Ratio and Low-Intensity Region Score are considered. These rely on a prior segmentation of the blurring artifacts or the intensity undershoot areas.

We propose a deep-learning-based method for the quantification of the absolute motion artifact levels directly from coronary cross-sectional image patches. By a data-driven methodology, machine learning holds the promise to circumvent the challenging task of designing an appropriate handcrafted measure. Over the past few years, Convolutional Neural Networks (CNNs) have been driving advances in many image-related tasks such as pattern recognition, classification, segmentation, generation, synthesis, style transfer, and translation (Krizhevsky et al., 2012; Chen et al., 2016; Gatys et al., 2016). Also in the medical domain, CNN-based predictive models showed great results (Litjens et al., 2017; Zreik et al., 2018; Ronneberger et al., 2015).

In the preliminary work (Elss et al., 2018a,b), clinical data with synthetic motion artifacts is generated and used for a subsequent supervised learning process. In (Elss et al., 2018b) the feasibility of motion artifact recognition using CNNs has already been demonstrated. The most recent work (Elss et al., 2018a) furthermore deals with single-slice motion estimation based on the coronary artifact appearance and the angular reconstruction range.



(a) Axial plane of a *step-and-shoot* case reconstructed at diastolic rest phase. Severe motion artifacts at the right coronary artery (RCA) are highlighted in red.

(b) Axial plane of a *helical* case reconstructed at diastolic rest phase. Severe motion artifacts at the right coronary artery (RCA) are highlighted in red.

Figure 1: Cardiac motion leads to differently shaped artifacts in helical and step-and-shoot CT scans. Motion artifacts in step-and-shoot CT scans are less complex as merely one coherent angular segment is used for the reconstruction of each voxel. Due to table movement and multi-cycle reconstruction, motion artifacts have a different appearance in CT scans with helical acquisition mode.

Extending this work, we propose a *Coronary Motion Forward Artifact model for CT data* (CoMoFACT). Beyond motion artifact recognition, we quantify motion artifact levels with CNNs. Our deep-learning-based motion artifact measures are created by the following steps:

1. The CoMoFACT generates the required input and label data for supervised learning by introducing simulated and hence controlled motion to clinical cases with excellent image quality, see Section 3.1.
2. Following Elss et al. (2018b), CNNs are trained to classify motion-free and motion-perturbed coronary cross-sectional image patches, see Sections 3.2 & 3.2.3.
3. CNNs are trained to predict the artifact level of coronary cross-sectional image patches, see Sections 3.2 & 3.2.4.

Finally, several experiments are performed to investigate the generalization capabilities of the resultant deep-learning-based motion artifact measures

to clinical data with real artifacts (see Sections 4.2 & 4.3.2).

2. Material

2.1. Reference data

Contrast-enhanced cardiac CT data sets with excellent image quality are the basis of the proposed deep learning procedure. In addition to the reconstructed CT image volumes, we require the corresponding coronary artery trees and the raw projection data. Slice-by-slice visual inspection is performed to gather CCTA data sets which exhibit no coronary motion artifacts in the reconstructed cardiac CT image volume. In total, 17 prospectively ECG-triggered clinical data sets from different patients are selected as reference point determining the *no motion* state.

In all reference cases, acquisition was performed with a 256-slice CT scanner (Brilliance iCT, Philips Healthcare, Cleveland, OH, USA) using a step-and-shoot protocol and a gantry rotation speed of 0.272 sec per turn. The restriction to step-and-shoot cases offers the advantage to generate artifacts in a well-controlled situation without table movement or multi-cycle reconstruction (see Figure 1). The mean heart rates of the patients $\mathrm{HR}_{\mathrm{mean}}$ ranged from 45.2 bpm to 66.0 bpm during the acquisition. The cardiac CT image volumes are reconstructed at the mid-diastolic quiescent cardiac phase. The center of the cardiac gating window for the aperture-weighted cardiac reconstruction (AWCR) (Koken and Grass, 2006; van Stevendaal et al., 2007), hereafter called the reference cardiac phase r is chosen between 70% and 80% R-R interval, respectively.

The coronary artery tree of each case is segmented using the Comprehensive Cardiac Analysis Software (IntelliSpace Portal 9.0, Philips Healthcare, Cleveland, OH, USA). It includes a set of centerline points $\vec{c} \in C$ with associated information on the corresponding cross-section, such as lumen contour and normal vector $\vec{n}_{\vec{c}}$ (centerline direction). As illustrated in Figure 2, the required input and label data for supervised learning is generated by applying the forward model presented in Section 3.1 to these 17 reference cases.

2.2. Test data

We collect eight clinical cases from different patients which exhibit real motion artifacts for testing purposes to complement the artifact-free reference cases. Step-and-shoot data (five vessels) as well as helical data (three vessels) from the Brilliance iCT are considered and corresponding centerlines are extracted by the Comprehensive Cardiac Analysis Software. The helical CCTA scans have an extended temporal scan range which enables retrospectively ECG-gated reconstruction of multiple heart phases. The centerlines are extracted for these cases from the reconstructed CT image volume with 75% R-R as reference gating phase.

3. Methods

3.1. CoMoFACT

The *Coronary Motion Forward Artifact model for CT data* (CoMoFACT) takes a reconstructed CT image volume with corresponding raw projection data and segmented coronary artery tree as input and delivers locally motion-perturbed CT image volumes as output. The introduced motion level is determined by the control parameter $s \in \mathbb{R}^+$, hereafter called the target motion strength.

The CoMoFACT introduces simulated motion by applying the motion compensated filtered back-projection (MC-FBP) algorithm (van Stevendaal et al., 2008; Schäfer et al., 2006) which is briefly explained in Section 3.1.1. For each centerline point $\vec{c} \in C$ in the coronary artery tree, an continuous MVF $\vec{d}_{\vec{c}}$ with pre-selected target motion strength s is created. The Subsections 3.1.2, 3.1.3, 3.1.4 detail the design of the synthetic MVF and underlying motion models. Subsequent MC-FBP delivers a CT image volume which is locally motion-perturbed around the corresponding centerline point. The reversing motion trajectory $\vec{d}_{\vec{c}}^{-1}$ corresponds to the simulated heart motion during acquisition.

After application of the CoMoFACT, one cross-sectional image patch is sampled perpendicular to the centerline (see Section 3.2.1) and finally added to the input data of the supervised learning process. As described in Section 3.2.3 and 3.2.4, ground truth labels are defined by means of the utilized target motion strength s. In this way, the CoMoFACT enables the generation of the required input and label data for supervised learning.

3.1.1. Motion-compensated filtered back-projection

The MC-FBP algorithm is an extension of the AWCR method. Both concepts are compared in Figure 3. In the AWCR, the attenuation coefficient

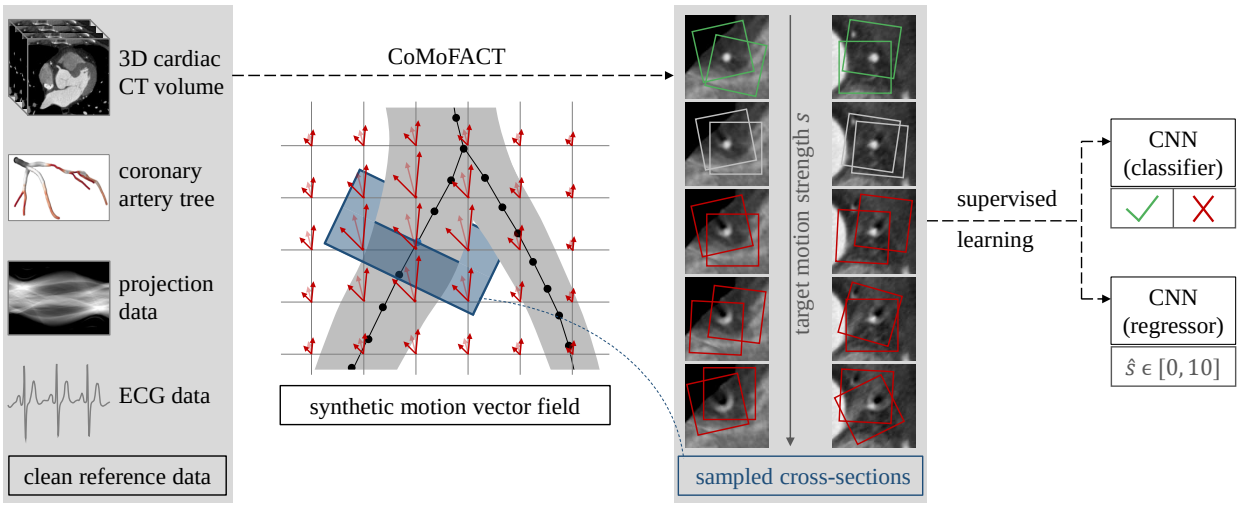


Figure 2: A *Coronary Motion Forward Artifact model for CT data* (CoMoFACT) is developed which enables to transform cardiac CT data sets with excellent image quality to locally motion-perturbed CT image volumes. Motion is introduced around each centerline point of the coronary artery tree by respectively creating a synthetic continuous MVF and applying the MC-FBP algorithm. The target motion strength s is a control parameter which scales the length of the displacement vectors in the CoMoFACT. A coronary cross-sectional patch (highlighted in blue) is sampled perpendicular to the centerline from the locally motion-perturbed image sub-volume and used as input data for supervised learning. Corresponding ground truth labels are defined by means of the target motion strength s. Finally, CNNs are trained for motion artifact classification and artifact level regression on randomly rotated, mirrored and cropped cross-sectional patches (highlighted in green, gray and red).

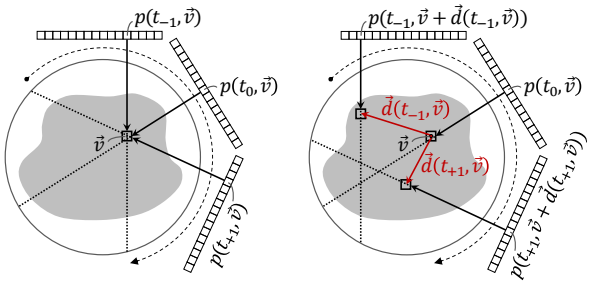
$\mu(\vec{\nu})$ of each voxel $\vec{\nu} \in \Omega$ in the field of view $\Omega \subset \mathbb{R}^3$ is calculated by:

$$\mu(\vec{\nu}) = \int_{t_{\text{start}}}^{t_{\text{end}}} w_{\text{AWCR}}(t, \vec{\nu})\; p_{\text{filt}}(t, \vec{\nu})\; \mathrm{d}t \qquad (1)$$

The weighting function w_{AWCR} includes aperture weighting for avoidance of cone-beam artifacts, angular weighting for gated reconstruction and pi-partner normalization. The projection integral $p_{\text{filt}}(t, \vec{\nu})$, which passes through the voxel $\vec{\nu}$ at time point $t \in [t_{\text{start}}, t_{\text{end}}]$, is re-binned to wedge geometry and high-pass filtered with a ramp filter. The variables t_{start} and t_{end} denote the start time and the end time of the CT scan. The MC-FBP furthermore takes into account the estimated displacements $\vec{d}(t, \vec{\nu})$ of each voxel during acquisition:

$$\mu(\vec{\nu}) = \int_{t_{\text{start}}}^{t_{\text{end}}} w_{\text{AWCR}}(t, \vec{\nu}+\vec{d}(t, \vec{\nu}))\; p_{\text{filt}}(t, \vec{\nu}+\vec{d}(t, \vec{\nu}))\; \mathrm{d}t \qquad (2)$$

For this purpose, a reference motion state has to be chosen. MVFs are usually calculated by registration and interpolation to approximate the motion each image voxel has undergone between the reference time t_0 and the time each specific projection was acquired. Each voxel is moved accordingly before back-projection is actually done. So, MC-FBP leads to a compensation of correctly estimated motion, whereas the application of the MC-FBP with an artificial MVF on high quality cases induces motion artifacts.



(a) AWCR (b) MC-FBP

Figure 3: Schematic drawing of voxel-driven back-projection without and with motion compensation. (a) The simple back-projection procedure of AWCR relies on the assumption, that the object is static during acquisition. Inconsistent projection data caused by motion leads to artifacts in the reconstructed CT image volume. (b) In case of MC-FBP, moving voxel positions are considered and line integrals are spatially corrected in the back-projection procedure.

3.1.2. Synthetic motion vector field

The continuous MVF $\vec{d}_{\vec{c}}\colon [0\%, 100\%] \times \Omega \to \mathbb{R}^3$ describes the displacement of each voxel coordinate $\vec{\nu} \in \Omega \subset \mathbb{R}^3$ in the CT volume at each time point $t_{cc} \in [0\%, 100\%]$ in millimeters. Due to the simulation of periodic motion, time is measured in percent cardiac cycle here. The artificial MVF is defined by three separable components:

$$\vec{d}_{\vec{c}}(t_{cc}, \vec{\nu}) = s \cdot m_{\vec{c}}(\vec{\nu}) \cdot \vec{\delta}_{\vec{c}}(t_{cc}) \tag{3}$$

The first component is the pre-selected target motion strength s which scales the length of each displacement vector in the continuous MVF. The role of s as motion level regulator is illustrated in Figure 2. Whenever $s = 0$, no motion is introduced and the CoMoFACT delivers the input CT image volume without motion artifacts as output.

The second component is the location-dependent weighting mask $m_{\vec{c}}\colon \Omega \to [0, 1]$ which restricts the motion to a limited area around the currently processed centerline point $\vec{c}$ and additionally forces the MVF to be spatially smooth. It is defined as a 3D trapezoidal function generated by binary dilation of the centerline point and subsequent uniform filtering. A kernel radius of 15 mm for dilation and a uniform filter size of 12.4 mm × 12.4 mm are chosen in the following experiments. The limitation of the motion area is required to prevent undesired motion artifacts from peripheral structures like bones. The smoothing is necessary to avoid reconstruction artifacts as elastic tissue structure forbids abrupt changes of motion in a local neighborhood.

The third component $\vec{\delta}_{\vec{c}}\colon [0\%, 100\%] \to \mathbb{R}^3$ defines the motion direction for each point in time. It is obtained by piecewise linear interpolation between five sample vectors $\vec{\delta}_i \in \mathbb{R}^3$, $i \in \{-2, -1, 0, +1, +2\}$. The corresponding phase points $t_i \in \{r - 10\%, r - 5\%, r, r + 5\%, r + 10\%\}$ are assigned around the reference heart phase r of the input CT volume with a temporal distance of 5% cardiac cycle. The temporal projection range required for reconstruction depends on the heart rate and the gantry rotation speed. For the given data sets, the angular weighting window is narrower than 20% cardiac cycle respectively, so no extrapolation has to be performed.

A schematic drawing of the artificial MVF is given in Figure 2. The displacement vectors (light red arrows) are linearly interpolated in time domain from the sample vectors (dark red arrows) to obtain the motion state at some $t_{cc} \in [r - 10\%, r + 10\%]$. For a phase point t_{cc}, the motion directions are spatially constant, while the displacement length decreases with increasing distance to the currently processed centerline point $\vec{c}$ (highlighted in blue).

Two model variants are presented in the following Subsections 3.1.3 and 3.1.4 which differ in terms of the sample vector definition. Both concepts are compared in Figure 4.

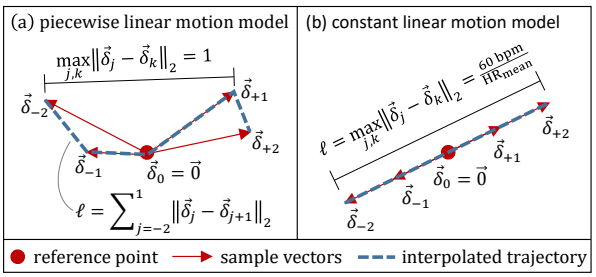


Figure 4: Schematic drawing of the time-dependent motion trajectories (dashed blue lines) determined by the sample vectors $\vec{\delta}_i$ (red arrows) for both sub-models. (a) The piecewise linear motion model comprises random displacement directions and varying velocities. (b) The constant linear motion model is restricted to a predefined motion direction and equidistant sample vectors.

3.1.3. Piecewise linear motion model

The first model variant of piecewise linear motion has been introduced in Elss et al. (2018b) and was developed for the classification task of separating *no-artifact* and *artifact* coronary cross-sectional patches. The sample vectors are calculated by:

$$\vec{\delta}_i = \frac{\vec{\rho}_i}{\max_{j,k} \|\vec{\rho}_j - \vec{\rho}_k\|_2} \tag{4}$$

The motion directions are given by random uniform vectors $\vec{\rho}_i \sim \mathcal{U}[-1, 1]^3$ for $i \in \{-2, -1, +1, +2\}$ and the center heart phase defines the reference state of no motion, i.e. $\vec{\rho}_0 = \vec{0}$. Normalization of the random uniform vectors $\vec{\rho}_i$ is performed so that the target motion strength s finally corresponds to the maximal displacement during 20% R-R interval in millimeters (see Figure 4a).

The choice of random uniform vectors for $\vec{\rho}_i$ enables complex motion trajectories and application of the CoMoFACT leads to realistic-looking motion artifact pattern with differently shaped blurring and intensity undershoots (see Figure 5a). The target motion strength s can be observed as weak surrogate of the visual artifact level. For deep-learning-based motion artifact quantification, the

motion model is adapted with the objective of increasing the correlation between selected target motion strength s and the resulting visual artifact level. Compared to (Elss et al., 2018b), the CoMoFACT is extended for cross-section-wise motion corruption under consideration of the mean heart rate and the angular reconstruction range.

3.1.4. Constant linear motion model

The second model variant is developed for the regression task of predicting the artifact level in coronary cross-sectional patches. Multiple factors beside the motion level during acquisition have an impact on the artifact level. The phantom study in Figure 6 shows that the relation between motion direction and the angular reconstruction range is essential. Most severe artifacts occur in case of motion which is orthogonal to the mean reconstruction direction (highlighted in gray). In addition, the visual artifact level depends on surrounding background intensities, the temporal resolution required for reconstruction and the relation between motion direction and vessel orientation.

In contrast to the classification model variant presented in Section 3.1.3, severe restrictions are made to consider each of the aforementioned influencing factors except for surrounding background intensities. The sample vectors are calculated now by the following formula:

$$\vec{\delta}_i = \frac{60 \text{ bpm}}{\text{HR}_{\text{mean}}} \frac{i}{4} \cdot \frac{\vec{\rho}_{\text{orth}}}{\|\vec{\rho}_{\text{orth}}\|_2} \tag{5}$$

As illustrated in Figure 4b, the regression model is limited to constant linear motion. It takes the mean heart rate HR_{mean} of each data set during acquisition into account to force homogeneous velocities among the clinical input cases. In contrast to the previous classification model, the motion direction now depends on the currently processed centerline point $\vec{c}$.

The motion direction determined by $\vec{\rho}_{\text{orth}}$ is defined as the cross product of the normal vector $\vec{n}_{\vec{c}}$ of the corresponding centerline segment and the mean reconstruction direction in axial plane (see Figure 6, right). The mean reconstruction direction is computed by means of the gantry rotation angle at the center of the cardiac gating window and is constant for each voxel reconstructed by the same circular scanning shoot. In case of helical acquisition trajectories, voxels are not necessarily reconstructed by one coherent angular segment. So, in contrast to the classification model, the regression model with its orthogonal displacement directions is not directly transferable to helical cases.

Figure 5b shows coronary cross-section images with varying target motion strength s generated by the constant linear motion model. In comparison to the corresponding outputs of the piecewise linear model in Figure 5a, the data looks more consistent and the target motion strength s can be interpreted as an approximate measure for the artifact level. However, the severely restricted constant linear motion model merely allows for a specific artifact appearance (banana-shaped blurring) whereas the more complex motion trajectories of the piecewise linear motion also generates bird-shaped artifacts. Therefore, the risk of overfitting should be considered in the evaluation of the regression networks.

3.1.5. Sub-volume reconstruction

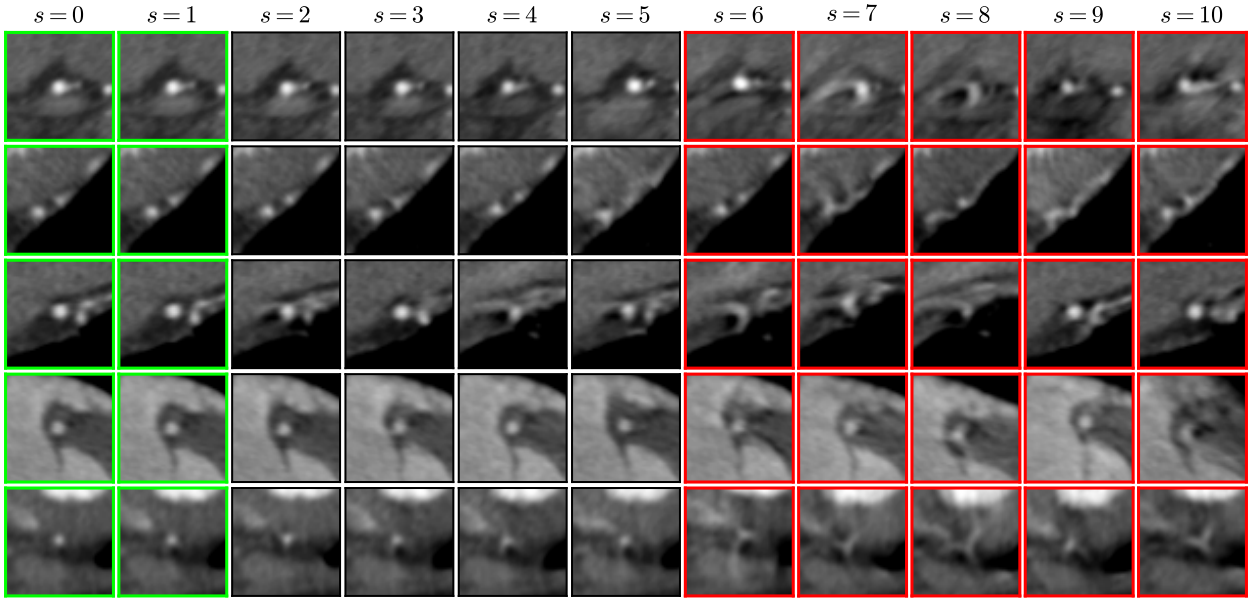
For the final extraction of one cross-sectional image patch per motion-perturbed image volume, merely a limited area around the currently processed centerline point $\vec{c}$ is of interest. Therefore, the FOV $\Omega = \Omega(\vec{c})$ is restricted in the CoMoFACT, in order to speed-up the data generation process. The selected centerline point $\vec{c}$ defines the center of the restricted FOV and the FOV size is determined by the patch size during the subsequent sampling process (see Section 3.2.1). In addition, the reconstruction of a sub-volume instead of the full input CT image geometry is reasonable in terms of memory requirements.

3.2. Supervised learning

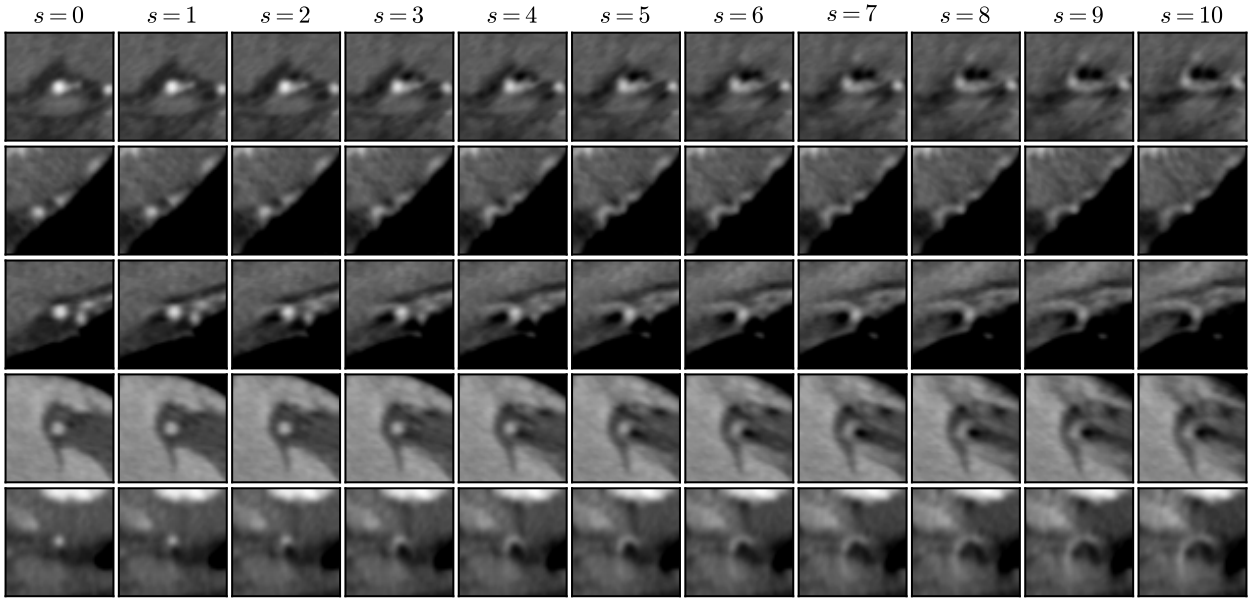
The proposed CoMoFACT enables the generation of multiple motion-perturbed CT image (sub)volumes $I_{\Omega}^{\vec{c},s}$ with controlled motion level s at specific coronary centerline points $\vec{c} \in C$. On the basis of velocity measurements at the coronary arteries by Vembar et al. (2003), the data generation process for the supervised learning task is limited in the following experiments to maximal displacements of 10 millimeters during 20% cardiac circle.

3.2.1. Patch sampling

The sampling process of the input data is illustrated in Figure 2. One cross-sectional image patch $I_{100}^{\vec{c},s}$ of size $100 \times 100 \times k$ voxels (blue box) is sampled by trilinear interpolation from each output CT image volume $I_{\Omega}^{\vec{c},s}$ of the CoMoFACT with a resolution of $0.4 \times 0.4 \times 0.4$ millimeters per voxel. The first



(a) Output patches of the piecewise linear motion model from Section 3.1.3.



(b) Output patches of the constant linear motion model from Section 3.1.4.

Figure 5: Coronary cross-sectional image patches are sampled from motion-perturbed CT image volumes, which are generated by means of the CoMoFACT. Each row shows the same cross-section of size 60×60 pixels in different motion states. Both sub-models of the CoMoFACT are compared and reveal the trade-off between motion model complexity and suitability of s as artifact level surrogate. (a) The piecewise linear motion model from Section 3.1.3 leads to a wide range of coronary artifact appearances, but the visual coronary artifact level is not monotonically increasing with the underlying target motion strength s. Patches highlighted in green and red are assigned to the classes *no artifact* or *artifact* respectively. Non-highlighted patches with a target motion strength s between two and five are excluded from the learning process. (b) The constant linear motion model from Section 3.1.4 leads to a limited range of coronary artifact appearances, but the underlying target motion strength s highly correlates with the visual artifact level at the coronary arteries. All patches are included as input data in the regression learning process with s as their corresponding label.

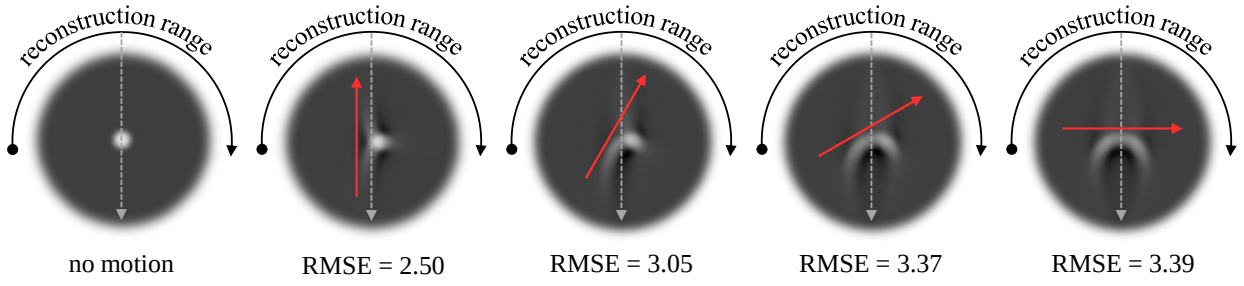


Figure 6: Constant linear motion is introduced in the axial plane of a vessel phantom using the CoMoFACT. Depending on the relation of motion direction (red arrow) and mean reconstruction direction (gray dashed arrow), motion artifacts with varying level occur. The artifact level is measured by the root mean squared error (RMSE) with respect to the motion-free phantom plane (left). Orthogonal motion (right) leads to most severe motion artifacts in the reconstructed image.

two patch dimensions are aligned perpendicular to the centerline segment while the third dimension is oriented along the normal vector $\vec{n}_{\vec{c}}$. The center slice of each image patch covers the processed centerline point $\vec{c}$. Single-slice ($k = 1$) and multi-slice ($k > 1$) patches are tested as input data in the following experiments. The grey values of each patch are clipped to the relevant intensity range with a window/level setting of 900/200 HU and additionally normalized (from $[-250, 650]$ to $[-1, 1]$).

3.2.2. Data augmentation

Due to the patch similarity of adjacent centerline points, the data for training, validation and testing are case-wise separated with a ratio of 9 : 4 : 4. The database during the training process is extended by online data augmentation. Motion artifacts are variable in shape, orientation and position. In order to build this invariance into the neural network, the following transformations are performed in cross-sectional plane, i.e. limited to the first two dimensions:

Cropping: The CoMoFACT may cause small vessel shifts compared to the original coronary centerline position. Image translation by cropping is necessary to avoid a bias from the in-plane coronary position. Therefore, sub-patches of the size $85 \times 85 \times k$ are randomly cropped from $I^{\vec{c},s}_{100}$.

Rotating: The sub-patches of the size $85 \times 85 \times k$ are randomly rotated by 0 to 360 degrees. The center patch of the size $60 \times 60 \times k$ is finally cropped to ensure full image contents.

Mirroring: Horizontal mirroring is performed with a probability of 0.5.

The final image patches $I^{\vec{c},s}_{60}$ of size $60 \times 60 \times k$ voxels are used as input data for supervised learning. During validation and testing merely center cropping is performed.

3.2.3. Classification

Following Elss et al. (2018b), the database for the classification task of separating *artifact* and *no artifact* cross-sectional patches is generated by applying the proposed piecewise linear motion model of Section 3.1.3 seven times (with different target motion strength s) per coronary centerline point. The target label y_{class} (0: *no artifact*, 1: *artifact*) of each image patch $I^{\vec{c},s}_{60}$ is defined by the corresponding utilized target motion strength:

$$y_{\text{class}} = \begin{cases} 0, & \text{if } s \in \{0, 1\} \\ 1, & \text{if } s \in \{6, 7, 8, 9, 10\} \end{cases} \tag{6}$$

The gap in s is chosen to assure better class separation. Merely a subset of two fifths of the samples from class *artifact* is randomly selected and included in the learning process, in order to force balanced classes. By this procedure, a total amount of 14 724 samples is collected as classification database.

Multiple hyperparameter settings and network architectures including the ResNet (He et al., 2016) and VGG inspired networks (Simonyan and Zisserman, 2015) were tested by extensive cross-validation. Highest validation results are achieved by a feed-forward 20-layer ResNet which is employed in all subsequent experiments. Figure 7 illustrates the network architecture. The learning process is driven by the cross-entropy loss. The Adam optimizer (Kingma and Ba, 2015) with an initial learning rate of 0.05, a minibatch size of 32 and a momentum of 0.8 is defined as the learning setup.

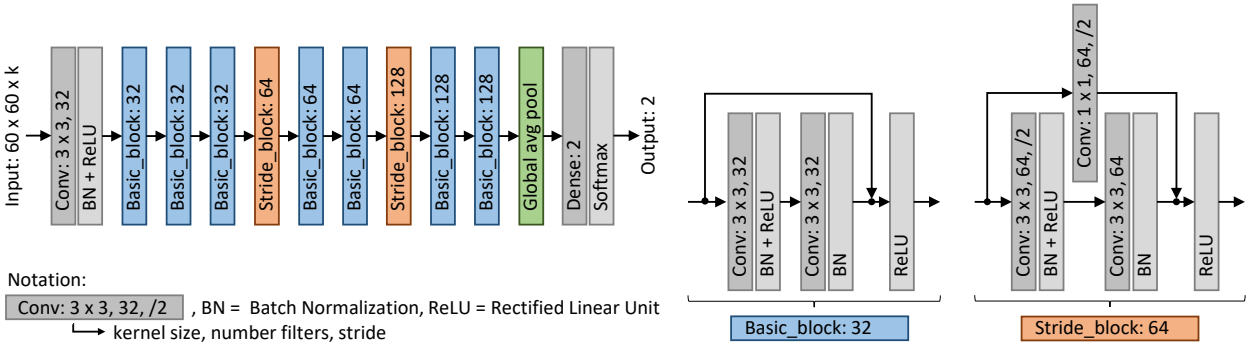


Figure 7: The 20-layer ResNet takes coronary cross-sectional image patches of the size $60 \times 60 \times k$ as input. In the last layer the soft-max function is used as activation function of the two (positive and negative) output nodes. The projection shortcuts are realized as 1×1 convolutions. In case $k > 1$ the convolutional kernel sizes are enlarged from $(3 \times 3 \times 1)$ to $(3 \times 3 \times 3)$.

Finally, the trained neural network $\mathrm{NN}_{\mathrm{class}}$ takes a cross-sectional patch $I_{60}^{\vec{c},s}$ as input and delivers a predicted artifact probability as output.

3.2.4. Regression

The database for the regression task of predicting the artifact level in cross-sectional patches is generated by applying the proposed constant linear motion model of Section 3.1.4 eleven times per coronary centerline point. The target label $y_{\mathrm{regr}} = s \in \{0, 1, \ldots, 10\}$ of each image patch $I_{60}^{\vec{c},s}$ is set equal to the corresponding utilized target motion strength. By this discrete equidistant labeling procedure, a total amount of 40 491 samples is collected as regression database.

Except for the reduction of output neurons in the last layer from two to one and the replacement of the soft-max function by simple linear activation, no adaption of the network architecture is done compared to Figure 7. The initial learning rate is changed to $5 \cdot 10^{-4}$ while the remaining hyperparameters remain unchanged.

The neural network $\mathrm{NN}_{\mathrm{regr}}$ takes a cross-sectional patch $I_{60}^{\vec{c},s}$ as input and delivers a prediction $\hat{y} \in \mathbb{R}$ as output. The learning process is driven by a piecewise L1 loss function:

$$l(y_{\mathrm{regr}}, \hat{y}) = \begin{cases} \max(0, \hat{y}), & \text{if } y_{\mathrm{regr}} = 0 \\ |y_{\mathrm{regr}} - \hat{y}|, & \text{if } 0 < y_{\mathrm{regr}} < 10 \\ \max(0, 10 - \hat{y}), & \text{if } y_{\mathrm{regr}} = 10 \end{cases} \tag{7}$$

This loss function considers adaptive penalization at the boundaries for predictions outside the target interval $[0, 10]$ to avoid for too conservative predictions. In comparison to network training with the simple L1 loss, the regressors more often dare output values near zero or ten. Clipping of the network output finally delivers the predicted artifact level $\hat{s} = min(max(0, \hat{y}), 10) \in [0, 10]$ which is used for the following evaluation.

4. Experiments and Results

4.1. Quantitative error analysis

For all experiments, the Microsoft Cognitive Toolkit (CNTK v2.5, Microsoft Research, Redmond, WA, USA) is used as deep learning framework. A bagging approach is applied for quantitative evaluation which comprises the following steps:

1. Four validation cases and four test cases are randomly sampled.
2. Network training is performed based on the remaining nine clinical cases.
3. After every epoch of the learning process the generalization capability is examined by means of the validation set.
4. The model with the highest performance on the validation set during 60 epochs of training is selected for calculation of the test accuracy (or test error).
5. Steps 1.-4. are performed five times in total.
6. The mean test accuracy (or the mean test error) over the five splits is calculated.
7. Steps 5.-6. are performed for $k \in \{1, 3, 5, 7\}$ (with the same separations in training, validation and testing for comparability).

The test results of the classification and the regression networks are summarized in Table 1. The networks performances increase with the number of

Table 1: Test results including mean and standard deviation of the classification accuracy and the absolute regression error for single-slice and multi-slice input data.

number slices	classification accuracy	regression error
$k = 1$	$91.64\% \pm 1.63\%$	1.38 ± 0.17
$k = 3$	$92.08\% \pm 2.12\%$	1.16 ± 0.06
$k = 5$	$92.70\% \pm 2.18\%$	1.14 ± 0.07
$k = 7$	$\mathbf{93.26\% \pm 1.82\%}$	$\mathbf{1.12 \pm 0.07}$

input slices. The additional information in multi-slice patches seem to provide a benefit, e.g. in the differentiation between bifurcations and blurring artifacts. But, higher memory requirements and execution time have to set against it. In case of $k = 7$, the classification result splits into a ratio of $46.90\% : 46.36\% : 3.64\% : 3.10\%$ for the rates TN : TP : FN : FP, where *positive* refers to the class *artifact*. Figure 8 shows the confusion matrix of the corresponding regression network. A clear diagonal structure with few scattering of the predicted labels is observable.

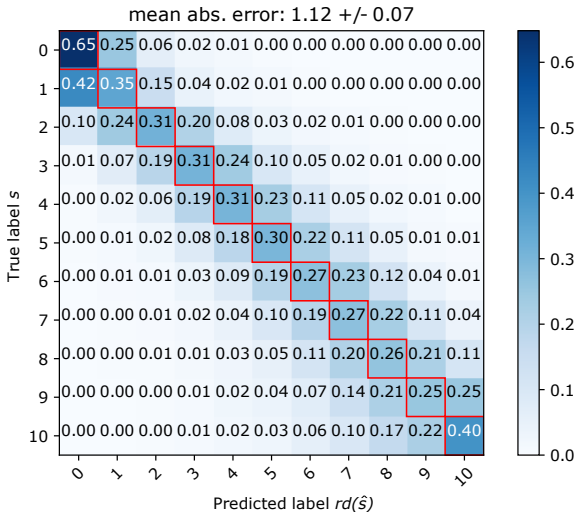


Figure 8: Confusion matrix of the regression network for multi-slice input data ($k = 7$) with rd denoting the rounding operator.

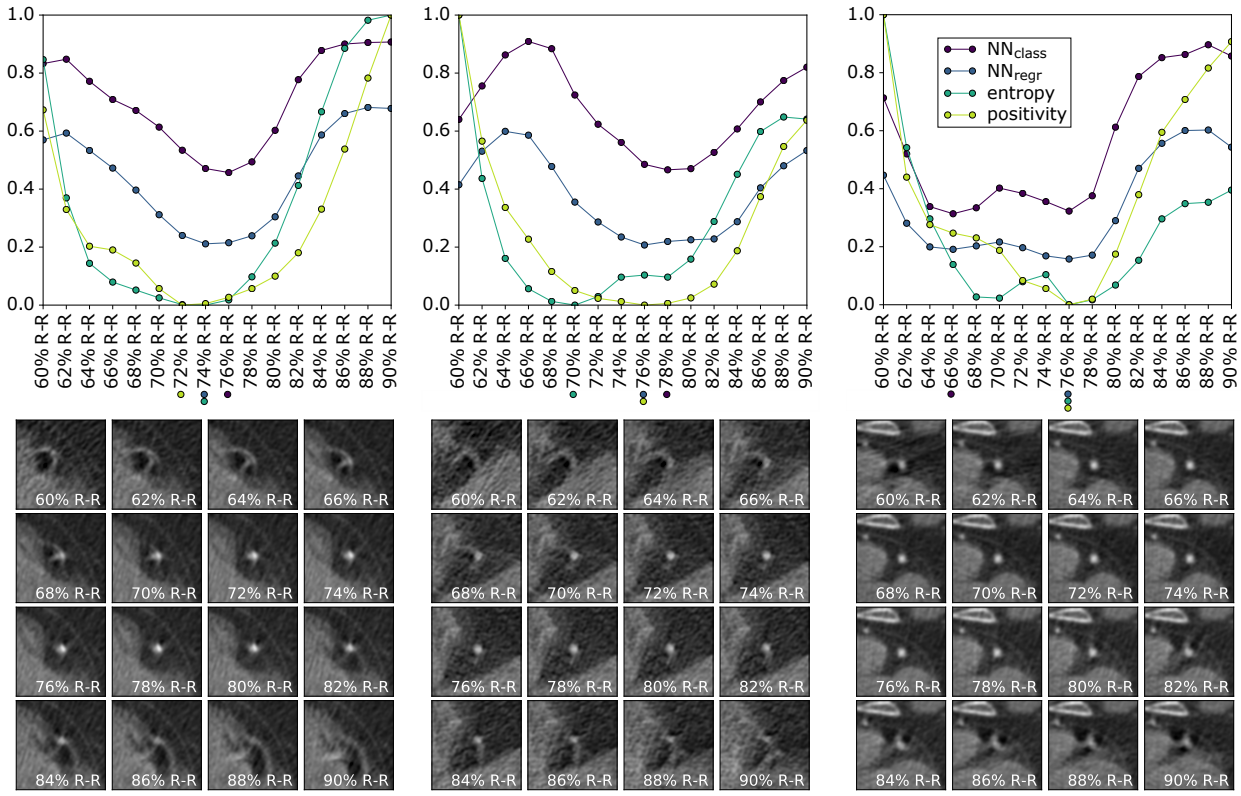
The presented quantitative results prove that CNNs are able to identify artifact pattern from synthetically introduced motion. To further evaluate generalization capabilities and the performance of the learned motion artifact measures in clinical practice additional qualitative experiments are performed. In comparison with the handcrafted MAMs entropy and positivity from Rohkohl et al. (2013), the abilities for measuring relative and absolute levels of motion artifacts are verified.

4.2. Relative artifact measurement

In the following, it is investigated whether the deep-learning-based artifact measures are able to identify the cardiac phase of a clinical data set with least motion artifacts. A quantitative study by Vembar et al. (2003) has shown that minimum velocities at the right coronary artery (RCA) can be observed in the mid-diastolic cardiac phase (between 70% and 80% R-R). Therefore, increasing temporal distance to this cardiac phase should go along with increasing artifact levels.

The three helical test cases (see Section 2.2) are reconstructed at multiple cardiac phases using the AWCR algorithm. The selected cardiac phases are arranged around 75% R-R interval (mid-diastole) with a temporal distance of 2% R-R interval. Cross-sectional patches of size $60 \times 60 \times 7$ are sampled along the RCA as input for the deep-learning-based and the handcrafted motion artifact measures. Finally, the mean motion artifact measures across the entire vessel are computed for each reconstructed phase image. In this experiment, the bagging ensembles of the five classification and the five regression networks with $k = 7$ are selected. The calculated motion artifact measures are scaled to the value range $[0, 1]$, to provide comparability. Therefore, the predicted artifact levels $\hat{s} \in [0, 10]$ of the regression network $\mathrm{NN}_{\mathrm{regr}}$ are down-scaled by a factor of ten. The handcrafted MAMs are normalized to the interval $[0, 1]$ by the minimal and maximal output over all motion states. As the classification network $\mathrm{NN}_{\mathrm{class}}$ already delivers predicted artifact probabilities, these remain unchanged.

Figure 9 shows the results of the proposed multi-phase experiment. In the first case (Figure 9a), all motion artifact measures (handcrafted and deep-learning-based) provide similar results. The predicted best cardiac phases around $72 - 76\%$ R-R comply with the visual inspection. In the second case (Figure 9b), only the positivity and the deep-learning-based measures deliver predictions of the best cardiac phase which concur with the visual impression. However, a weakness of the neural networks can be discerned. The modulation of the radiation dose leads to lower signal-noise-ratios (SNRs) at the marginal cardiac phases around 60% R-R. The trained neural networks seem to be more



(a) Case 1 (b) Case 2 (c) Case 3

Figure 9: The mean motion artifact level of the RCA is analyzed in three cardiac CT cases with helical acquisition mode by the deep-learning-based and the handcrafted motion artifact measures. For each reconstructed CT image volume within the phase interval $[60\%, 90\%]$, cross-sectional image patches are sampled based on the centerline segmented at the reference phase 75% R-R. The outputs of the motion artifact measures are averaged, scaled to the interval $[0, 1]$ and visualized in the upper row. The predictions of the best cardiac phases are indicated by colored dots, respectively. One axial slice per selected cardiac phase is given below for visual inspection.

vulnerable to such SNR fluctuations than the handcrafted measures. In the third case (Figure 9c), the regression network is in agreement with both handcrafted MAMs with respect to the predicted best cardiac phase at 76% R-R. The classification network selects an earlier stage around 66% R-R in which also hardly artifacts occur. A temporally extended rest phase is observable which is discovered by the trained neural networks.

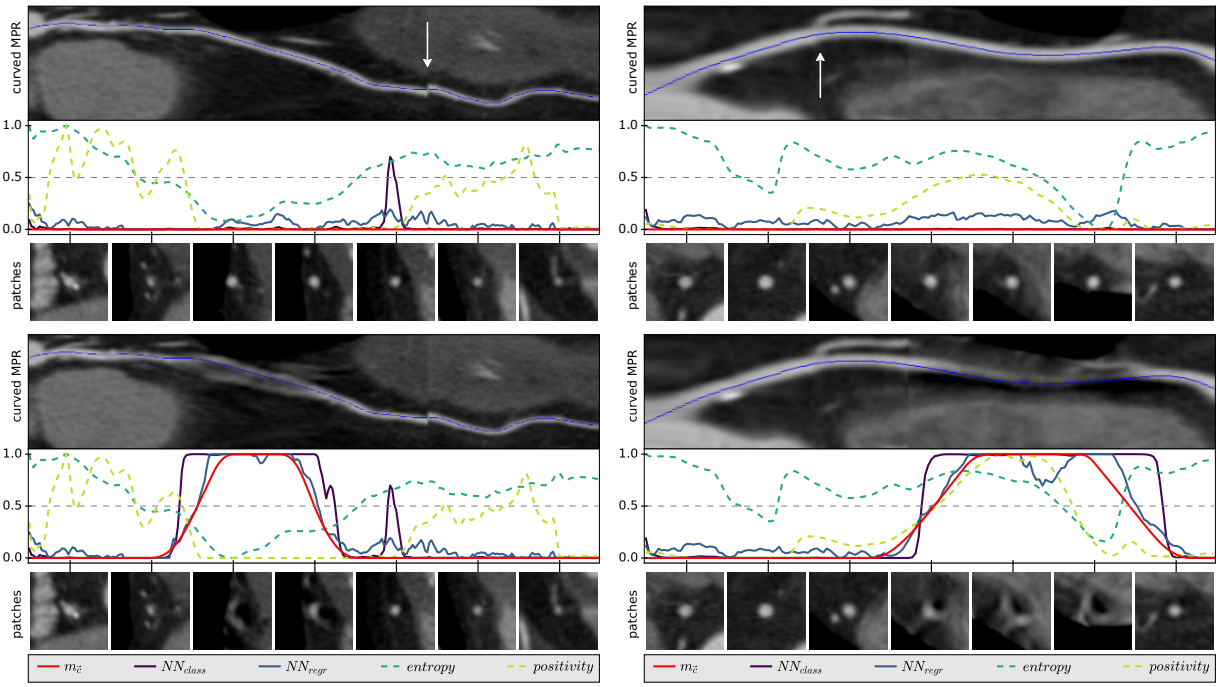
The results of this multi-phase experiment are promising given the fact, that the deep-learning-based measures are solely trained on step-and-shoot data which is perturbed by constant or piecewise linear motion. Generalization capabilities of the CNNs and transferability to helical data sets with real motion artifacts are demonstrated by this experiment.

4.3. Absolute artifact measurement

In the following, it is investigated whether the deep-learning-based artifact measures are able to detect a region of motion, given the approximate location of the coronary artery.

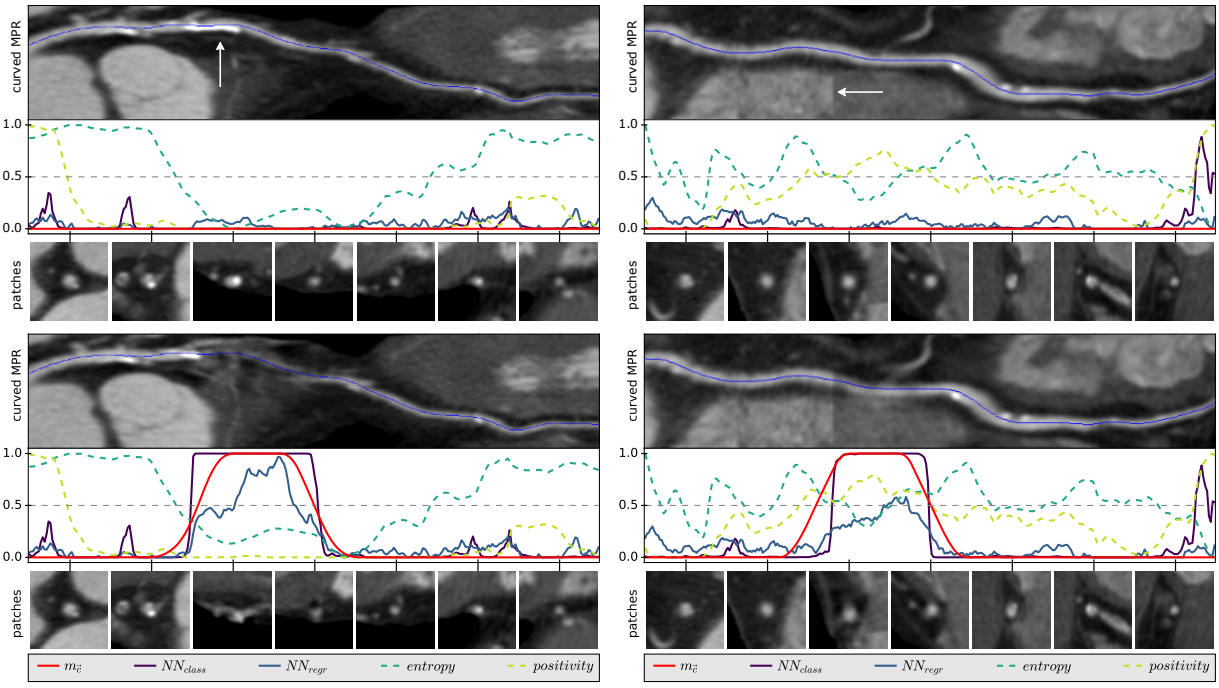
4.3.1. Evaluation on synthetic motion artifacts

Local motion is introduced to test reference cases at arbitrary points in the coronary tree by means of the piecewise linear motion model proposed in Section 3.1.3. A target motion strength of $s = 8$ is selected in the following experiments. For each centerline point of the corresponding vessel, a cross-sectional patch of size $60 \times 60 \times 7$ is sampled from the locally motion-perturbed CT image volume $I_\Omega^{\vec{c},s}$ and the corresponding motion artifact measures are calculated. The classification and regression networks are selected so that the currently processed



(a) Case 1, vessel = LAD

(b) Case 1, vessel = RCA



(c) Case 2, vessel = LAD

(d) Case 2, vessel = RCA

Figure 10: Section-wise outputs of the deep-learning-based and the handcrafted motion artifact measures are calculated for four vessels without and with synthetic motion artifacts. In each subplot, the no motion state is given as reference above the local motion state. Stack transition artifacts (a)&(d), bifurcations (b) and severe calcifications (c) are indicated by white arrows in the MPRs. The weighting mask value $m_{\vec{c}}$ marked in red, corresponds to the true relative displacement width of each centerline point and determines the area of introduced motion. Corresponding cross-sectional image patches are given below for visual inspection. In contrast to the handcrafted MAMs, high predictions made by the neural networks are mostly correctly located at the regions of motion influence.

test case has neither been used for training nor for validation.

Figure 10 shows the results of this local motion experiment for two test cases. The left anterior descending artery (LAD) and the right coronary artery (RCA) are processed, respectively. In each subplot, the original no motion state is visualized as reference above the local motion state. The x-axis corresponds to the spatial coordinate along the centerline. The value range $[0, 1]$ of the y-axis is determined by the weighting mask values $m_{\vec{c}} \in [0, 1]$ which correspond to the level of introduced motion. In each subplot, calculated motion artifact measures are scaled accordingly, to provide comparability again.

As expected, the handcrafted MAMs are not suitable for section-wise motion artifact quantification, due to limited robustness regarding variations in background intensities. The deep-learning-based measures, by contrast, accurately detect regions of motion with few exceptions. Both networks are robust towards vessel shifts at stack transitions (see Figure 10a). Bifurcations (see Figures 10b), calcifications (see Figures 10c) and varying contrast levels between scanning sequences (see Figure 10d) do not affect the deep-learning-based measures either. The aforementioned image areas are highlighted in the multi-planar reformations (MPRs) by white arrows. The lowest artifact level in the motion area is predicted by the regression network in Figure 10d, which also confirms with visual inspection of the four motion-perturbed MPRs. This experiment already demonstrates generalization capability of the regression network which is trained on perturbed data with constant linear motion and tested on data with more complex piecewise linear motion trajectories.

4.3.2. Evaluation on real motion artifacts

The ability for absolute motion artifact measurement is additionally tested on eight clinical cases with real motion artifacts (see Section 2.2). Ensemble averaging ($k = 7$) is performed for the evaluation. Figure 11 shows the resulting artifact measurements and corresponding cross-sectional patches. The vessels are sorted by the maximum artifact level predicted by the regression network. Artifact areas identified by the classification networks (with running average for outlier removal) are highlighted in red.

Four separate observer studies were performed to rate the 120 cross-sectional image patches visualized in Figure 11 with respect to diagnostic reliability in a five point Likert scale (— : excellent, — : good, — : mixed, — : strong artifact, — : non-diagnostic). The eight vessels were presented in random order without indication of the motion artifact measures to the readers. It has to be noted that the readers were no radiologists, but research scientists with high level of expertise in reading cardiac CT images. The resulting annotations are visualized as color bars in the Figure 11.

In contrast to the handcrafted MAMs, the deep-learning-based measures deliver sensible results. The classification networks enable the rough detection of artifact areas and the regression network additionally allows one to assess the artifact severity. Noise (see Figure 11e) and vessel segments with tiny lumen (see Figure 11g) can be identified as potential sources of uncertainty. The human observer ratings show disagreements in these areas as well. Furthermore, confusions between bifurcations and blurring artifacts are observable in the reader scores (see first patch of Figure 11b and Figure 11h). Image patches with consistent artifact labels (red or orange) are correctly identified as motion-perturbed by the neural networks. This also holds for image patches with consistent artifact-free labels (light green and dark green). Hence, transferability from synthetic to real motion artifacts is demonstrated. The proposed experiments reveal potential and current limitations of deep-learning-based artifact measurement in clinical practice.

5. Discussion

The paper demonstrates the feasibility of accurate motion artifact quantification in the coronary arteries using deep learning. The results of the quantitative error analysis in Section 4.1 merely provide indications, as the target labels do not always constitute exact ground truth. Figure 5a shows multiple image patches with false positive target label, for instance originating from motion along the vessel orientation. The thresholds which define the margin in s are crucial parameters which might affect the classification performance. The reduced complexity of the constant linear motion model presented in Section 3.1.4 enables more consistent target labels, but these might also exhibit slight inaccuracies. Some label noise originates from approximations and simplifications in the CoMoFACT. The centerline and its normal vectors are

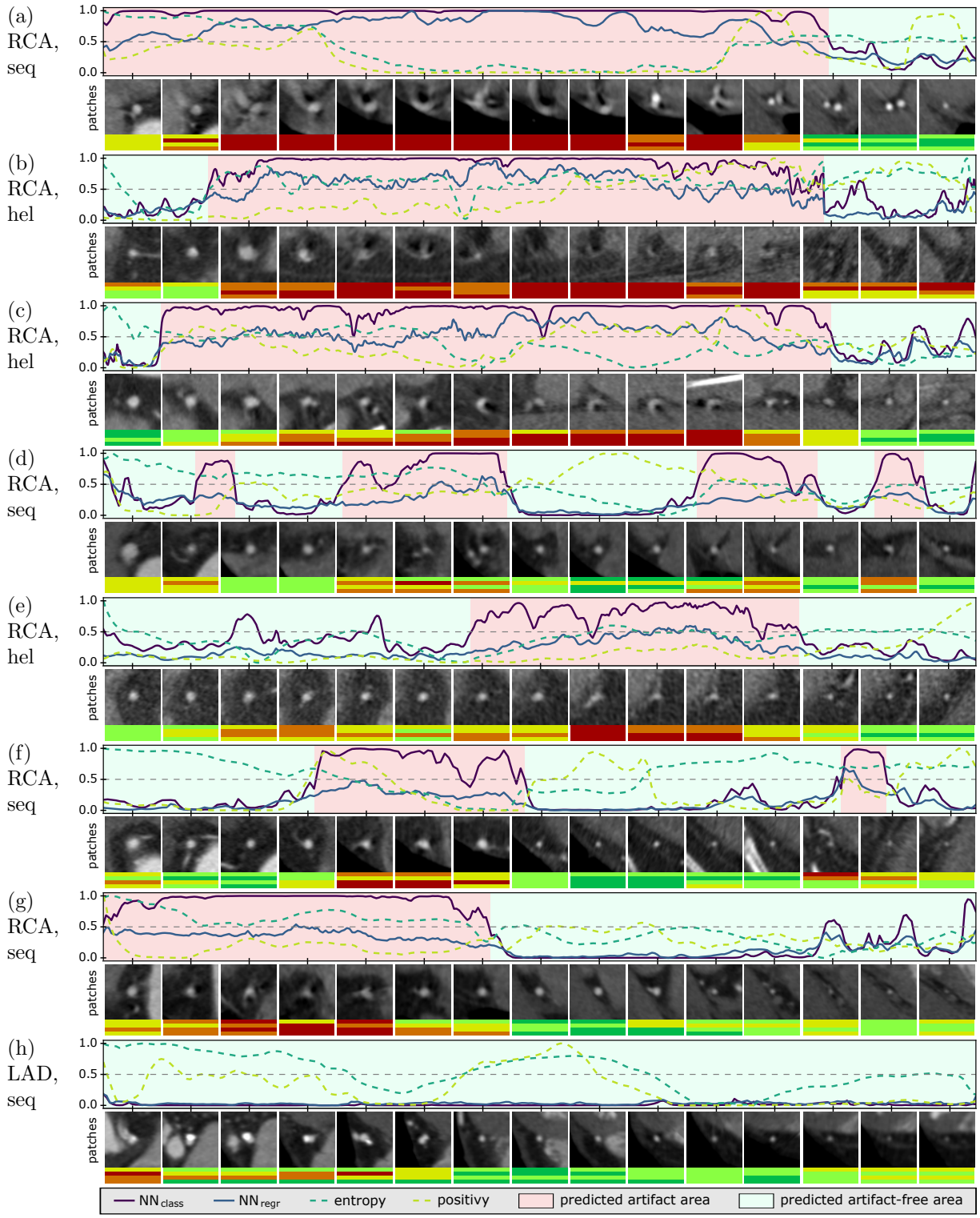


Figure 11: Section-wise outputs of the deep-learning-based and the handcrafted motion artifact measures are calculated for eight vessels with real motion artifacts. The predicted artifact level $\hat{s}$ of $\mathrm{NN}_{\mathrm{regr}}$, the entropy and the positivity are downscaled to $[0, 1]$. Corresponding cross-sectional image patches with four human observer ratings each (from — : excellent to — : non-diagnostic) are given below for visual inspection. The vessels are sorted by the maximum artifact level predicted by the regression network. Vessel type (RCA or LAD) and acquisition mode (seq: step-and-shoot or hel: helical) are specified.

merely estimates and the mean reconstruction direction is limited to the axial plane, i.e. the z-axis is not considered. Furthermore, motion in the axial plane has different effects on the reconstructed image data than motion in z-direction. Also other factors like background intensities and image noise influence the visual artifact level. The majority of the generated image and corresponding label data constitute good approximations of the reality. For network training purposes, the label quality is sufficient, but exact test performance can not be determined. As a next step, quantitative comparison studies to hand-labeled data from radiologists should be performed.

For the quality assessment of CCTA images using the proposed motion artifact measures, the approximate locations of the coronary arteries have to be known. Since motion artifacts frequently inhibit fully automatic centerline segmentation, alternative approaches are required. Many tools for coronary centerline extraction are semi-automatic, i.e. they allow the user to guide the segmentation process. Furthermore, the creation of a coronary artery atlas which involves the probability density for the position of each vessel segment with respect to the heart segmentation and the deployment of deep-learning based centerline extraction are options which should be investigated. The proposed CoMoFACT enables one to evaluate the robustness of centerline extraction methods with regard to motion artifacts. In general, motion introduction by the CoMoFACT might be useful as a data augmentation strategy in several other CT data-driven learning tasks.

So far, the proposed measures are merely based on 17 clinical data sets. CT images for non-invasive coronary angiography are acquired with a wide variety of scanner types, imaging protocols and reconstruction algorithms. In order to increase robustness of the CNNs, collection of more data and network fine-tuning should be performed. Especially in order to increase the networks insensitivity to noise, an extension of the training database by including clinical cases with lower SNR or synthetic noise introduction would be required. Robustness might additionally be improved by measurement smoothing of adjacent centerline points in order to reduce appearing scatter and to avoid outliers.

6. Conclusions

Application of the MC-FBP algorithm using high quality cardiac CT cases and artificial MVFs as inputs, enabled the generation of CT image data with controlled motion levels at the coronary arteries. Subsequent supervised learning of CNNs delivered measures for motion artifact recognition and quantification. High predictive accuracy was achieved on data with synthetically introduced motion, thus demonstrating the advantage of deep-learning-based measures for motion artifact localization over existing handcrafted MAMs. Furthermore, the generalization capabilities of our measures have been shown by clinical data with real artifacts and helical acquisition mode. Quantitative validation studies are required to assess the transferability of these promising initial results to motion artifact prediction in clinical practice.

References

Bhagalia, R., Pack, J.D., Miller, J.V., Iatrou, M., 2012. Non-rigid registration-based coronary artery motion correction for cardiac computed tomography. Medical Physics 39, 4245–4254.

Budoff, M.J., Li, D., Kazerooni, E.A., Thomas, G.S., Mieres, J.H., Shaw, L.J., 2017. Diagnostic accuracy of noninvasive 64-row computed tomographic coronary angiography (CCTA) compared with myocardial perfusion imaging (MPI): the PICTURE study, a prospective multicenter trial. Academic Radiology 24, 22–29.

Camargo, G.C., Peclat, T., Souza, A.C., Lima, R.d.S.L., Gottlieb, I., 2017. Prognostic performance of coronary computed tomography angiography in asymptomatic individuals as compared to symptomatic patients with an appropriate indication. Journal of Cardiovascular Computed Tomography 11, 148–152.

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2016. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence doi:10.1109/TPAMI.2017.2699184.

Elss, T., Nickisch, H., Wissel, T., Bippus, R., Morlock, M., Grass, M., 2018a. Motion estimation in coronary CT angiography images using convolutional neural networks. Medical Imaging with Deep Learning (MIDL) .

Elss, T., Nickisch, H., Wissel, T., Schmitt, H., Vembar, M., Morlock, M., Grass, M., 2018b. Deep-learning-based CT motion artifact recognition in coronary arteries, in: Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 1057416. doi:10.1117/12.2292882.

Foy, A.J., Dhruva, S.S., Peterson, B., Mandrola, J.M., Morgan, D.J., Redberg, R.F., 2017. Coronary computed tomography angiography vs functional stress testing for patients with suspected coronary artery disease: A systematic review and meta-analysis. JAMA Internal Medicine 177, 1623–1631. doi:10.1001/jamainternmed.2017.4772.

Gatys, L.A., Ecker, A.S., Bethge, M., 2016. Image style transfer using convolutional neural networks, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414–2423.

Ghekiere, O., Salgado, R., Buls, N., Leiner, T., Mancini, I., Vanhoenacker, P., Dendale, P., Nchimi, A., 2017. Image quality in coronary CT angiography: challenges and technical solutions. The British Journal of Radiology 90, 20160567.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.

Isola, A.A., Grass, M., Niessen, W.J., 2010. Fully automatic nonrigid registration-based local motion estimation for motion-corrected iterative cardiac CT reconstruction. Medical Physics 37, 1093–1109.

Kingma, D., Ba, J., 2015. Adam: A method for stochastic optimization, in: International Conference on Learning Representations (ICLR).

Koken, P., Grass, M., 2006. Aperture weighted cardiac reconstruction for cone-beam CT. Physics in Medicine and Biology 51, 3433.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems (NIPS), pp. 1097–1105.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical Image Analysis 42, 60–88.

Liu, T., Maurovich-Horvat, P., Mayrhofer, T., Puchner, S.B., Lu, M.T., Ghemigian, K., Kitslaar, P.H., Broersen, A., Pursnani, A., Hoffmann, U., et al., 2017. Quantitative coronary plaque analysis predicts high-risk plaque morphology on coronary computed tomography angiography: results from the ROMICAT II trial. The International Journal of Cardiovascular Imaging , 1–9.

Ma, H., Gros, E., Szabo, A., Baginski, S.G., Laste, Z.R., Kulkarni, N.M., Okerlund, D., Schmidt, T.G., 2018. Evaluation of motion artifact metrics for coronary CT angiography. Medical Physics 45, 687–702.

McGee, K.P., Manduca, A., Felmlee, J.P., Riederer, S.J., Ehman, R.L., 2000. Image metric-based correction (autocorrection) of motion effects: analysis of image metrics. Journal of Magnetic Resonance Imaging 11, 174–181.

Rohkohl, C., Bruder, H., Stierstorfer, K., Flohr, T., 2013. Improving best-phase image quality in cardiac CT by motion correction with MAM optimization. Medical Physics 40.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 234–231.

Schäfer, D., Borgert, J., Rasche, V., Grass, M., 2006. Motion-compensated and gated cone beam filtered backprojection for 3-D rotational X-ray angiography. IEEE Transactions on Medical Imaging 25, 898–906.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations (ICLR).

van Stevendaal, U., von Berg, J., Lorenz, C., Grass, M., 2008. A motion-compensated scheme for helical cone-beam reconstruction in cardiac CT angiography. Medical Physics 35, 3239–3251.

van Stevendaal, U., Koken, P., Begemann, P.G., Koester, R., Adam, G., Grass, M., 2007. ECG gated circular cone-beam multi-cycle short-scan reconstruction algorithm, in: Medical Imaging 2007: Physics of Medical Imaging, International Society for Optics and Photonics. p. 65105P.

Vembar, M., Garcia, M., Heuscher, D., Haberl, R., Matthews, D., Böhme, G., Greenberg, N., 2003. A dynamic approach to identifying desired physiological phases for cardiac imaging using multislice spiral CT. Medical Physics 30, 1683–1693.

Šprem, J., de Vos, B.D., de Jong, P.A., Viergever, M.A., Išgum, I., 2017. Classification of coronary artery calcifications according to motion artifacts in chest CT using a convolutional neural network, in: Styner, M.A., Angelini, E.D. (Eds.), Medical Imaging 2018: Image Processing, International Society for Optics and Photonics. p. 101330R. doi:10.1117/12.2253669.

Zreik, M., Lessmann, N., van Hamersvelt, R.W., Wolterink, J.M., Voskuil, M., Viergever, M.A., Leiner, T., Išgum, I., 2018. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. Medical Image Analysis 44, 72–85.