applied sciences

MDPI

Article

Cycle Consistent Generative Motion Artifact Correction in Coronary Computed Tomography Angiography

Amal Muhammad Saleem [1], Sunghee Jung [1,*], Hyuk-Jae Chang [2] and Soochahn Lee [1]

[1] Department of Electronics Engineering, Kookmin University, Seoul 02727, Republic of Korea; amalsaleem@kookmin.ac.kr (A.M.S.) ; sclee@kookmin.ac.kr (S.L.)
[2] Division of Cardiology, Severance Cardiovascular Hospital, Yonsei University College of Medicine, Yonsei University Health System, Seoul 03722, Republic of Korea; hjchang@yuhs.ac
* Correspondence: sh.jung@kookmin.ac.kr

Abstract: In coronary computed tomography angiography (CCTA), motion artifacts due to heartbeats can obscure coronary artery diagnoses. In this study, we introduce a cycle-consistent adversarial-network-based method for motion artifact correction in CCTA. Our methodology involves extracting image patches and using style transfer for synthetic ground truth creation, followed by CycleGAN network training for motion compensation. We employ Dynamic Time Warping (DTW) to align extracted image patches along the artery centerline with their corresponding motion-free phase patches, ensuring matched pixel correspondences and similar anatomical features for accuracy in subsequent processing steps. Our quantitative analysis, using metrics like the Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD), demonstrates CycleGAN's superior performance in reducing motion artifacts, with improvements in image quality and clarity. An observer study using a 5-point Likert scale further validates the reduction of motion artifacts and improved visibility of coronary arteries. Additionally, we present a quantitative analysis on clinical data, affirming the correction of motion artifacts through metric-based evaluations.

Keywords: coronary computed tomography angiography; motion artifact correction; deep learning; generative adversarial networks; Dynamic Time Warping

Citation: Saleem, A.M.; Jung, S.; Chang, H.-J.; Lee, S. Cycle Consistent Generative Motion Artifact Correction in Coronary Computed Tomography Angiography. *Appl. Sci.* **2024**, *14*, 1859. https://doi.org/10.3390/app14051859

Academic Editor: Vladislav Toronov

Received: 1 January 2024
Revised: 14 February 2024
Accepted: 22 February 2024
Published: 23 February 2024

1. Introduction

Coronary computed tomography angiography (CCTA) is a widely utilized non-invasive technique for diagnosing coronary artery disease (CAD), offering diagnostic accuracy on par with conventional invasive coronary angiography but with a markedly reduced risk of complications [1]. Despite these advantages, the acquisition of CCTA images is susceptible to motion artifacts, potentially leading to errors in coronary artery tracking and segmentation.

Several approaches have been proposed for motion correction using image-based motion estimation methods [2–7]. These methods leverage motion vector fields derived through techniques like registration and artifact metric minimization. However, motion estimation itself is an extremely difficult problem, perhaps even more so in the presence of motion artifacts.

As deep learning methods [8,9] have demonstrated revolutionary performance in various domains including image super-resolution [10], image denoising [11], and image deblurring [12], deep-learning-based motion compensation methods have also been proposed [13–17]. One approach is to apply deep learning to improve the motion estimation used for motion correction [15,16].

However, the straightforward approach to applying learning-based methods for motion correction in CCTA is to learn to generate artifact-less images based on a large set of training images. Thus, in the work of Jung et al. [13,14], a convolutional neural

network (CNN) is trained to generate images without motion artifacts. In the work of Zhang et al. [16], the pix2pix [18] generative adversarial network (GAN) method is used.

In this paper, we present a new method for coronary motion correction based on GAN [19] with cycle consistency [20]. Compared to previous work, we introduce a more robust methodology for mitigating motion artifacts in CCTA images. Initially, we create a dataset that includes spatially aligned 2D image patches. These 2D cross-sectional image patches are extracted along the RCA (Right Coronary Artery) centerline from 3D CT volumes at different phases within a 4D CT. Figure 1 showcases representative images from the dataset, illustrating the coronary artery's various phases throughout a complete cardiac cycle as captured by the 4D CT scans. A distinctive aspect of our approach is the application of Dynamic Time Warping (DTW) for the precise alignment of image patches. This method ensures that each patch extracted from the motion-affected phases is accurately matched with its corresponding patch from a phase with minimal motion artifacts. This alignment is crucial for maintaining the anatomical accuracy and consistency across the patches, forming a solid foundation for the subsequent motion correction process.

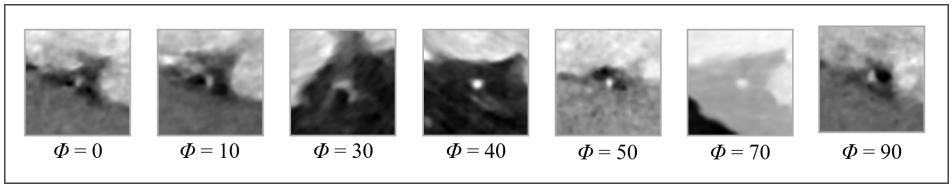


Figure 1. Various phases of the coronary artery throughout a complete cardiac cycle, as captured by a 4D CT scan.

Following this, style transfer is implemented using a patch from a phase with the least motion artifact as the target to generate synthetic ground truths for motion-affected patches. The goal is to generate a synthetic motion-corrected ground truth (GT) image patch as the fusion of the content of the source image, such as local anatomical elements, with the motion artifact characteristics of the target image.

Leveraging the motion-free and motion-affected pairs, we train a CycleGAN model [20] to produce realistic artifact-free patches. At the last stage, these 2D patches are reinserted into our 3D CT scan with volumetric interpolation to obtain the final motion-corrected 3D CT volume. This dual-step image generation process ensures the preservation of the original internal structures, while simultaneously maintaining a realistic shape for the RCA. At the testing phase, we assume the 3D CT scan is motion-perturbed, and the RCA centerline is annotated. Based on this, we proceed to extract 2D patches, correct them, and subsequently reintegrate them into the scan.

The main contributions of our work are as follows. First, we propose a coronary artery motion correction technique that employs the cycle consistency GAN model on 2D image patches extracted from a CT image set comprising different cardiac phases, each presenting various forms of motion artifacts. Second, we utilized the Dynamic Time Warping technique to reduce the error involved in extracting corresponding image patches of the coronary artery from CT images across various cardiac phases. Third, we conducted thorough quantitative and qualitative assessments, comparing the effectiveness of our proposed method with previous works.

We include a detailed analysis of the performance improvements of our proposed method, highlighting the advancements over prior methodologies. We quantitatively evaluated the proposed method by measuring motion artifact metrics [21]. For comparison with related research, we conducted quantitative evaluations using metrics such as the Dice score to provide a comprehensive analysis. Additionally, we conducted an observer study to qualitatively score the degree of motion artifacts.

2. Previous Works

In addressing our problem, two prevalent deep learning methodologies are commonly employed. The first involves using neural networks to determine the underlying motion vectors of the artifacts. Lossau et al. [15] introduced constant linear motion in the axial plane on artifact-free CT scans to yield sets of motion-affected image patches along with their associated 2D motion vectors. Then, they trained CNNs to estimate the motion vectors. In Maier et al. [16], a deep neural network is trained to estimate the motion vector field (MVF), which is then used for motion-compensated image reconstruction.

The second methdology involves using neural networks to directly generate motion-corrected images based only on the input image appearance, without any explicit motion estimation. In Jung et al. [13,14], a CNN structure originally proposed for super-resolution [10] was used to generate artifact-removed images. In Ren et al. [22], the pix2pix [18] method, proposed for image-to-image translation, was used to improve CCTA images from raw CCTA imaging data using reconstructed CCTA images as targets. The reconstructed images were obtained from Snapshot Freeze (GE Healthcare), a commercial motion correction algorithm that leverages information from adjacent cardiac phases within a single cardiac cycle to mitigate motion artifacts. Zhang et al. [17] also utilized the pix2pix method to convert motion-perturbed images into motion-free images. They used an automated phase selection algorithm, Smart Phase (GE Healthcare), to determine the optimal target phase within the R-R interval from four sets of image reconstructions. Then, similar to Ren et al. [22], they employed Snapshot Freeze (GE Healthcare) to the optimal target phase to generate an image reconstruction that is assigned as the ground truth to train the pix2pix network. Both GAN-based approaches demonstrated significantly improved performance on quantitative metrics such as circularity, Dice Similarity Coefficient (DSC), and Hausdorff Distance (HD) for GAN-generated images compared to raw images when evaluated against the reference phase.

Unfortunately, corresponding CCTA images with and without motion artifacts cannot be acquired, since we cannot stop the heart to acquire images. Thus, special techniques must be employed to generate the training data necessary for supervised learning. In Jung et al. [13,14], neural style-transfer [23] is applied to convert images with artifacts into images without artifacts, which are considered as synthetic ground truth. Our proposed method applies a similar approach to that of Jung et al. [14] for generating ground truth. It is important to note that this method differs significantly from the data procurement process utilized by Ren et al. [22] and Zhang et al. [17]. This divergence in methodologies makes it challenging to directly compare our method with theirs. However, since they have both employed the popular generative adversarial network architecture, pix2pix, we have trained our own model to specifically evaluate and compare its performance against the proposed model, CycleGAN.

3. Methodology

Our proposed pipeline aims to enhance a motion-affected 3D CT volume and generate a motion-corrected version using neural networks. Therefore, we adhere to the following methodology:

1. We extract image patches along the artery centerline. These patches are aligned with their corresponding target motion-free phase patch using Dynamic Time Warping (DTW). DTW is an algorithm used to match correspondences between pixels across the sequence of patches. This ensures that the corresponding patch pairs exhibit similar anatomical features. This alignment is essential to ensure the accuracy and reliability of the subsequent steps in the process;
2. Utilizing the patch corresponding to slower motion as the target, we employ a style transfer network to produce a synthetic ground truth image (SynGT) for the motion-perturbed image. Style transfer is performed between pairs of corresponding patches in different phases in a 4D CT and alters only the style (local texture), not the content

(structure). Our assumption is that motion artifacts are closer to style and thus can be reduced using this process;
3. A CycleGAN network is trained to perform motion compensation, with the input consisting of pairs of motion-affected and SynGT patches;
4. Motion-corrected patches are reinserted and interpolated into the original 3D CT volume to compensate for the motion artifacts of the mid RCA.

Figure 2 illustrates a schematic representation of the training processes mentioned in the second and third points. The subsequent sections explain the technical details of each subprocess.

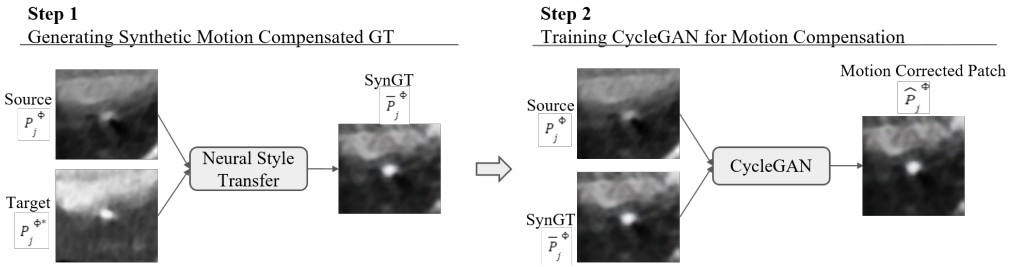


Figure 2. Workflow of the proposed method. Step 1, generate synthetic motion-corrected patch (SynGT) using style transfer method. Step 2, train the cycle consistency adversarial network(CycleGAN) using SynGT. Detailed descriptions of steps 1 and 2 are given in Sections 3.2 and 3.3, respectively.

3.1. Extracting Corresponding Coronary Patches from 4D CT

The 4D CT data is captured using retrospective gating with a dual-source CT scanner. The data is systematically reconstructed at 10% intervals across the heartbeat cycle, highlighting phases at 40% and 70% for their minimal motion disturbance in the coronary artery, while other phases are often plagued by significant motion artifacts. We base this process on that of Jung et al. [14].

Focusing on the highly motion-sensitive midsection of the right coronary artery (RCA), an experienced radiologist manually annotates this region on each temporally sampled 3D CT volume. This annotation, done using the commercial coronary analysis software (QAngioCT, Medis Medical Imaging Systems, Leiden, the Netherlands), spans from the first right ventricle branch to the acute marginal branch.

Adhering to a process parallel to our prior studies, we discretize the mid-RCA centerline at each phase ϕ into a set of ordered 3D coordinates, forming a linear approximation between these points. The total length of the mid-RCA centerline is calculated by summing the distances between adjacent point pairs. For extracting patches along the centerline, we first establish corresponding points aligned with the same anatomical landmarks across different phases. A fixed number of equidistant points are sampled along the centerline, with interpolation to precisely determine their 3D coordinates. Planar patches, centered at these points, have their normals defined by the tangent to the centerline at each respective point. Figure 3 illustrates this process, highlighting the alignment of corresponding points across various temporal phases of the 3D CT volumes.

The corresponding patches are then extracted by sampling the voxel intensities within the 3D CT volume on a grid, centered and aligned at each equidistant point along the mid-RCA. This approach ensures consistency in the extracted patches across different phases, which is crucial for effective motion artifact correction in subsequent steps.

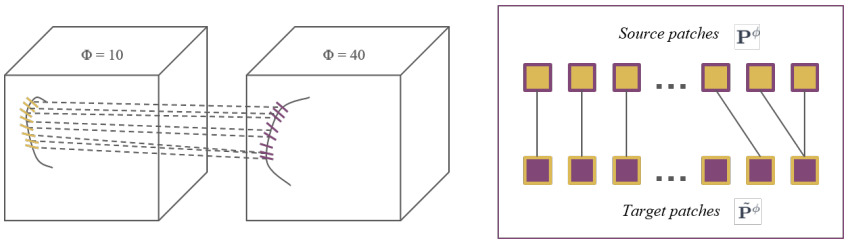


Figure 3. Using dynamic programming to find corresponding patches based on structural similarities between target and source patches.

3.2. Dynamic Time Warping for Synthetic Patch Alignment and Cross-Phase Style Transfer

The patches extracted from the coronary artery from different phases of a 4D CT scan are not only distinct in terms of severity of motion artifact but also in terms of the anatomical structures. The motion of the heart causes variability in the surrounding local features as well. Our goal is to obtain corresponding patches with similar neighboring distinctive features without motion disturbance, to train our CycleGAN model. Because this is clinically unattainable, we synthesize a same-phase-no-artifact patch, $\tilde{\mathbf{P}}_j^{\phi}$, using style transfer to the source patch $\mathbf{P}_j^{\phi}$ with a different-phase-no-artifact patch as the target patch $\mathbf{P}_j^{\phi\star}$. We refer to this process as *cross-phase style transfer*, where ϕ denotes the cardiac phase and $\phi\star$ denotes the phase within the heartbeat when the motion is the slowest, resulting in the least severe motion artifacts.

To optimally align the patches, we use the Dynamic Time Warping (DTW) algorithm. Let $\tilde{\mathbf{P}}^{\phi} = \{\tilde{\mathbf{P}}_1^{\phi}, \tilde{\mathbf{P}}_2^{\phi}, \ldots, \tilde{\mathbf{P}}_N^{\phi}\}$ be the target patch sequence, and $\mathbf{P}^{\phi} = \{\mathbf{P}_1^{\phi}, \mathbf{P}_2^{\phi}, \ldots, \mathbf{P}_M^{\phi}\}$ be the source patch sequence, where N and M are the lengths of the target and source sequences, respectively. We calculate the distance or similarity measure between each pair of patches $\tilde{\mathbf{P}}_i^{\phi}$ and $\mathbf{P}_j^{\phi}$:

$$\text{cost}(i,j) = \text{distance}(\tilde{\mathbf{P}}_i^{\phi}, \mathbf{P}_j^{\phi})$$

In the context of source and target patch matching with similar anatomical structures (background), we employ the Structural Similarity Index (SSIM) as the similarity measure. We chose this metric as it takes into account luminance, contrast, and structural information, which is crucial in our case. Mathematically, it is given by

$$\text{SSIM}(\tilde{\mathbf{P}}^{\phi}, \mathbf{P}^{\phi}) = \frac{(2\mu_{\tilde{\mathbf{P}}}\mu_{\mathbf{P}} + c_1)(2\sigma_{\tilde{\mathbf{P}}\mathbf{P}} + c_2)}{(\mu_{\tilde{\mathbf{P}}}^2 + \mu_{\mathbf{P}}^2 + c_1)(\sigma_{\tilde{\mathbf{P}}}^2 + \sigma_{\mathbf{P}}^2 + c_2)}$$

The mean values, $\mu_{\tilde{\mathbf{P}}}$ and $\mu_{\mathbf{P}}$, capture the average intensity of pixel values in $\tilde{\mathbf{P}}^{\phi}$ and $\mathbf{P}^{\phi}$. The covariance, $\sigma_{\tilde{\mathbf{P}}\mathbf{P}}$, measures the joint variability between the patches, while $\sigma_{\tilde{\mathbf{P}}}^2$ and $\sigma_{\mathbf{P}}^2$ represent their individual variances. The constants c_1 and c_2 prevent instability in the SSIM computation.

The resulting SSIM scores are stored in a cost matrix of size $N \times M$, which represents the accumulated cost of the optimal alignment up to position (i,j). To find the minimum cost path from the starting point to each point (i,j), an accumulated cost matrix of size $N \times M$, where $\text{AccCost}(i,j)$, is calculated. This accumulates costs using the recurrence relation:

$$\text{AccCost}(i,j) = \text{cost}(i,j) + \min(\text{AccCost}(i-1,j), \text{AccCost}(i,j-1), \text{AccCost}(i-1,j-1))$$

Then, this is traced back from (N,M) to $(1,1)$ to find the optimal alignment path. This path represents the matching between the target and source patches.

The technique of transforming a source image's style—but not its content—into a target image's style is known as style transfer. Texture, color, and contrast—both local and global—are components of the style. Conversely, content usually refers to the shapes, hues, and textures needed to identify the scene, including any particular people or objects. We

consider local anatomical features to be content in our framework, but motion artifacts as part of the style.

We apply the *neural style transfer* method [23], similar to the process in Jung et al. [14]. The pipeline of this method is as follows. A VGGNet [24]—pre-trained on the ImageNet database [25]—is used to compute local image features, which are the described as the numerical representation of the content. If we denote the tensor of the CNN features at layer l as $\mathbf{F}_x^1$ and $\mathbf{F}_c^1$ for the synthesized image $\vec{I}_x$ and content reference image $\vec{I}_c$, respectively, the loss function for the content is defined as

$$\mathcal{L}_{content}\left(\vec{I}_x, \vec{I}_c\right) = \frac{1}{2}||\vec{I}_x - \vec{I}_c||_2^2. \tag{1}$$

In the next step, the inner product between different CNN features at layer l is defined by the Gram matrix $\mathbf{G}^l$, as

$$\mathbf{G}_{ij}^l = \mathbf{F}_i^l \cdot \mathbf{F}_j^l, \tag{2}$$

where $\mathbf{G}_{ij}^l$ denotes the element at row i, column j of $\mathbf{G}^l$, and $\mathbf{F}_i^l$ and $\mathbf{F}_j^l$ denote the i_{th} and j_{th} features, respectively, corresponding to the i_{th} and j_{th} convolutional kernels, respectively, at layer l. The loss function for style is, hence, defined as

$$\mathcal{L}_{style}\left(\vec{I}_x, \vec{I}_s\right) = \frac{1}{2N_x^{l\,2} \times 2N_s^{l\,2}}||\mathbf{G}_x^l - \mathbf{G}_s^l||_2^2, \tag{3}$$

where $\mathbf{G}_x^l$ and $\mathbf{G}_s^l$ are the Gram matrices, and N_x^l and N_s^l are the number of features at layer l, for $\vec{I}_x$ and style-reference image $\vec{I}_s$, respectively.

Finally, $\vec{I}_x$ is determined by using a gradient descent to minimize the balanced loss, defined as

$$\mathcal{L}_{total}\left(\vec{I}_x, \vec{I}_c, \vec{I}_s\right) = \alpha\mathcal{L}_{content}\left(\vec{I}_x, \vec{I}_c\right) + \beta\mathcal{L}_{style}\left(\vec{I}_x, \vec{I}_s\right) \tag{4}$$

where α and β are coefficients to balance the effect between the content and style loss terms. The loss in Equation (4) is not optimized while training, it is fixed. Instead, a modified version $\vec{I}_x$ of the input images is generated [23].

To summarize, $\tilde{\mathbf{P}}_j^\phi$, $\mathbf{P}_j^\phi$, and $\mathbf{P}_j^{\phi\star}$ correspond to $\vec{I}_x$, $\vec{I}_c$, and $\vec{I}_s$, respectively. Whereas the phase $\phi\star$ with the minimum amount of motion is determined manually, the patches from all other phases ϕ can be assigned as the source, i.e., the reference patch for content $\mathbf{P}_j^\phi$.

3.3. Training and Applying the Cycle Consistent Generative Motion Artifact Correction Network

The proposed CycleGAN model is designed for image-to-image translation. It learns the pixel mapping and color distributions of the source and target patches for subsequent data synthesis. Typically, the GAN model comprises of two deep learning models—a generator model and a discriminator model. The learning process alternates between updating the generator to trick the discriminator more effectively and updating the discriminator to better identify generated images as real or fake, such that at the end the generator synthesizes realistic images.

As depicted in Figure 4, the training process of the CycleGAN aims to map features from source images X, and target images Y to accurately synthesize realistic RCA. In our study, source domain X images are the motion-perturbed patches, while the SynGT, from the previous step, are the target domain Y images. This method involves two cyclic image mapping functions (generators), denoted as $G : X \rightarrow Y$ and $F : Y \rightarrow X$, and two discriminators, denoted as D_x and D_y. In the CycleGAN architecture, the mapping functions G and F exhibit cycle consistency, signifying that the image translation cycle can restore an image x from domain X back to the original image ($F(G(x)) \rightarrow x$), and similarly, it can restore an image y from domain Y back to the original image ($G(F(y)) \rightarrow y$).

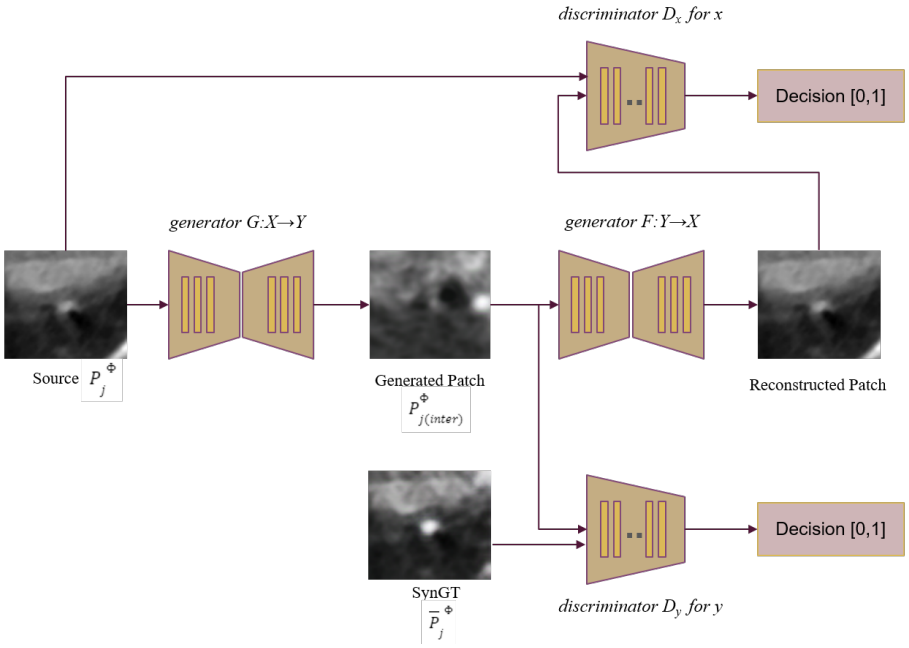


Figure 4. A schematic of our proposed method based on CycleGAN [20].

Our proposed CycleGAN framework utilizes a ResNet6-based generator that features six residual blocks. Each block comprises convolutional layers with skip connections, allowing for efficient gradient flow during training. The generator utilizes a series of deconvolutional layers for upsampling, progressively refining the translated images. The number of filters in the first convolutional layer is set to 64. The cycle-consistency loss $\mathcal{L}_{\text{cyc}}$ for the generator is mathematically defined as:

$$\begin{aligned}\mathcal{L}_{\text{cyc}}(G, F, X, Y) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\|F(G(x)) - x\|_1] \\ + \mathbb{E}_{y\sim p_{\text{data}}(y)}[\|G(F(y)) - y\|_1]\end{aligned}$$

For the discriminator, we have adopted a 70×70 PatchGAN model. This divides the input image into non-overlapping patches of 70×70 pixels, evaluating each independently. Mathematically, the adversarial loss $\mathcal{L}_{\text{GAN}}$ for the discriminator is defined as:

$$\begin{aligned}\mathcal{L}_{\text{GAN}}(D, X, Y) = \mathbb{E}_{y\sim p_{\text{data}}(y)}[\log D_Y(y)] \\ + \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log(1 - D_Y(G(x)))]\end{aligned}$$

This approach provides fine-grained feedback, enabling the generator to focus on localized details and reducing computational complexity. The model undergoes training for a total of 100 epochs with an initial learning rate set at 0.0002. Following this initial training phase, the learning rate is linearly decayed to zero over the next 200 epochs. Figure 5 provides a detailed illustration of the layers within the described generator and discriminator architecture.

We chose CycleGAN because of (1) its cycle consistency, which means that after translating patches to correct motion artifacts, the reverse translation should bring them back to their original state, helping in maintaining relevant anatomical details while mitigating artifacts, and (2) PatchGAN's ability to handle patch-wise translations, which aligns well with the need to correct artifacts in only specific regions of the image.

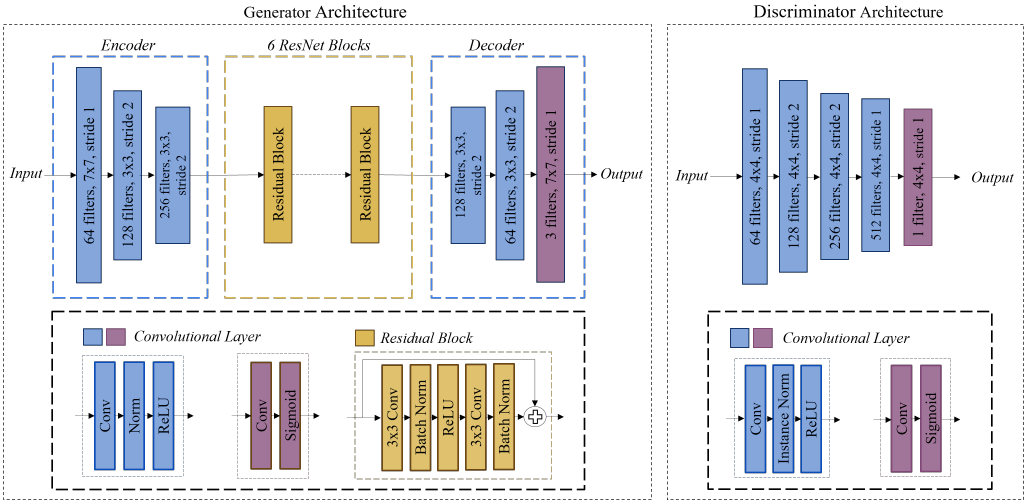


Figure 5. Detailed representation of the CycleGAN generator and PatchGAN discriminator architecture [20].

3.4. Reinsertion and Volumetric Interpolation of Motion-Corrected Patches into 3D CT Volume

Lastly, we reinsert the 2D patches generated by the CycleGAN model into the CT volume to obtain the motion-corrected scan. We employ the methodology outlined in our earlier work [14] and, therefore, provide only a concise overview here.

For a smooth and continuous appearance of the reinserted patches and the volume, we perform volumetric interpolation. Given our prior knowledge, of the center point $\vec{q}_j$ and the normal $\vec{n}_j$ of each patch through the patch extraction process described in Section 3.1, we reinsert the output patches $\mathbf{P}_j$ using the inverse of the known projection transforms.

For two adjacent patches $\mathbf{P}_j$ and $\mathbf{P}_{j+1}$, volumetric interpolation is on the bounding box enclosed by the reinserted patch coordinate grids $\mathbf{Q}_j$ and $\mathbf{Q}_{j+1}$. For each voxel with coordinate q within the bounding box, and in the volume between the two patches, we determine the vector $v^{k\star}_{j,j+1}$ among $\{v^k_{j,j+1}, 1 \leq k \leq R^2\}$ that is closest to the voxel coordinate, i.e., that has minimum point-to-line distance $d_{p2l}(q, v^k_{j,j+1}) = \frac{\left|(q-q^k_j)\times(q-q^k_{j+1})\right|}{\left|q^k_{j+1}-q^k_j\right|}$. The final value for voxel q is the weighted average defined as $I(q) = w^q_j \mathbf{P}_j(\tilde{q}^\dagger_j) + w^q_{j+1}\mathbf{P}_{j+1}(\tilde{q}^\dagger_{j+1})$, where $\mathbf{P}_j(q^\dagger_j)$ is the intensity obtained from bilinear interpolation on $\mathbf{P}_j$ at 2D non-integer coordinate $\tilde{q}^\dagger_j$, and w is the weight defined by the point-to-plane distances d_{p2p} of q to the planes containing $\mathbf{Q}_j$ and $\mathbf{Q}_{j+1}$.

4. Results

To facilitate a comprehensive comparison with other methods, we first conducted an evaluation using the CAVAREV platform [26] through a phantom study. This platform allowed us to benchmark our proposed method against both existing non-deep-learning (DL)-based and DL-based techniques previously evaluated on the same platform. In addition to the phantom study, we also assessed our method's performance on clinical data. For this evaluation, we specifically chose recent DL-based methods for comparison, including the work of Jung et al. [14] and methods based on the pix2pix framework. Although Ren et al. [22] and Zhang et al. [17] also employed pix2pix, direct comparisons were challenging due to differences in the training data distributions. To overcome this, we trained a pix2pix model on our dataset.

4.1. Phantom Study

Our evaluation was conducted on the CAVAREV platform [26], leveraging simulated dynamic projections from the 4D XCAT phantom, featuring contrast-enhanced coronary arteries modeled after patient data. The dataset, which focuses solely on cardiac motion, was selected for the assessment of our proposed methods. Geometry calibration was sourced from a genuine clinical angiographic C-arm system. This evaluation was chosen to facilitate a direct comparison with previous work, allowing us to benchmark the effectiveness of our method in motion correction against established techniques.

To evaluate the proposed method using the CAVAREV platform, we generated 10 C-arm CT volumes on a 256^3 grid, ensuring an isotropic voxel size of 0.5 mm, across ten distinct heart phases targeted for reconstruction. In line with prior works referenced on the CAVAREV website [27], the volume reconstructed at the heart phase of 0.9, characterized by minimal motion, served as the target phase for cross-phase style transfer, a process detailed in Section 3.2 of our methodology. The dataset comprised mid-RCA patches collected from 25 centerline points extracted from 10 CT volumes. Given data limitations, we allocated seven volumes for training, reserving two for test purposes.

Our method's efficacy in motion correction was quantified using the 3D metric (Q3D) as defined by CAVAREV [26], complemented by Dice similarity coefficients (DSCs) to gauge the overlap between binary images, with values ranging from zero (no overlap) to one (perfect match). The motion-corrected volumes were binarized and assessed against the ground truth, represented by the segmentation mask of the coronary artery from the volume reconstructed at a quiescent heart phase. The performance of the proposed method was evaluated using the average DSC across two test volumes. This served as the performance metric, and a comparative analysis was conducted against existing methodologies, including Jung et al. [14], pix2pix, and other methods featured on the CAVAREV website [27], as documented in Table 1. The comparison methods [28,29] are cost-minimization-based approaches, while those in [30–32] are 2D–2D- or 3D–3D-registration-based approaches.

According to Jung et al. [14], time efficiency is critical in emergency medical scenarios. In this context, our proposed DL-based method offers a notable benefit in terms of processing speed, while the pix2pix model shows performance on par with that of Jung et al. [14]. Our proposed method also exhibits an enhanced DSC score, indicating improved performance. Figure 6 displays a test example from a phase ($\phi = 50$) characterized by significant motion artifacts. Despite the initial poor image quality, the proposed method successfully diminishes these artifacts, showcasing its efficacy through both quantitative and qualitative improvements demonstrated on the CAVAREV dataset.

Table 1. Comparative analysis of Dice scores utilizing the CAVAREV dataset across various methods.

Method	Approach [a]	DSC
Dynamic Level Set [28]	CM	0.691
Streak-Reduced ECG-Gated FDK [30]	REG3D	0.744
Residual Motion Compensation [31]	REG2D	0.776
Motion Compensation [32]	REG2D	0.823
Spatio-temporal TV [29]	CM	0.876
Jung et al. (MAC-net) [14]	DL	0.829
pix2pix	DL	0.830
Proposed Method (CycleGAN)	DL	0.841

[a] CM: cost minimization, REG2D: 2D–2D registration, REG3D: 3D–3D registration, DL: deep learning.

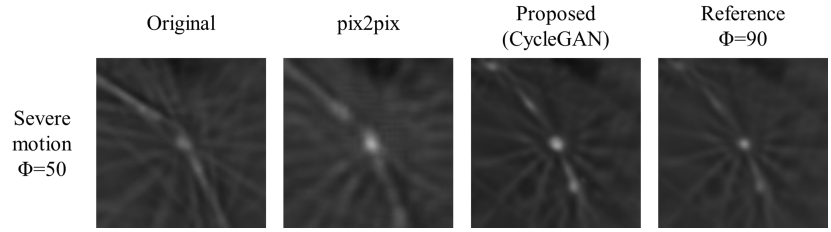


Figure 6. A qualitative example from the CAVAREV dataset test data showcasing severe motion artifacts. The set comprises the original mid-RCA test patch, the outcome from the pix2pix model, the result of the proposed CycleGAN method, and the corresponding reference patch. Phase information (ϕ from the 4D CT) is indicated above the patch. Radial artifacts are attributed to the limited number of projections ($n = 133$).

4.2. Clinical Data

4.2.1. Datasets

The 4D CT data utilized in this study were retrospectively collected using a dual-source CT scanner (SOMATOM Definition Flash, Siemens Healthineers, Forchheim, Germany) at Severance Hospital, South Korea, with the acquisition dates ranging from December 2015 to March 2016. The dataset comprises scans from 140 patients, systematically reconstructed at 10% intervals throughout the cardiac cycle. The training data comprised patches from 20 points in eight phases in 100 4D CT volumes. The number of phases $K = 8$ was determined by the number of temporal-phase quantizations, 10, minus the 40% and 70% phases designated as the targets for cross-phase style transfer. Several 4D CT volumes had K values of less than eight, in which phases with extreme motion artifacts were excluded because manual annotation of the coronary artery was impossible. The final dataset comprised a total of 5868 pairs of mid-RCA patches, which were then augmented to 35,208 using vertical and horizontal flips and rotations. Each patch sampled from the 3D volume was constructed to be of size 64×64 pixels. Validation and test sets, comprising 2152 patches from 30 4D CT volumes and 734 patches from 10 4D CT volumes, were similarly constructed.

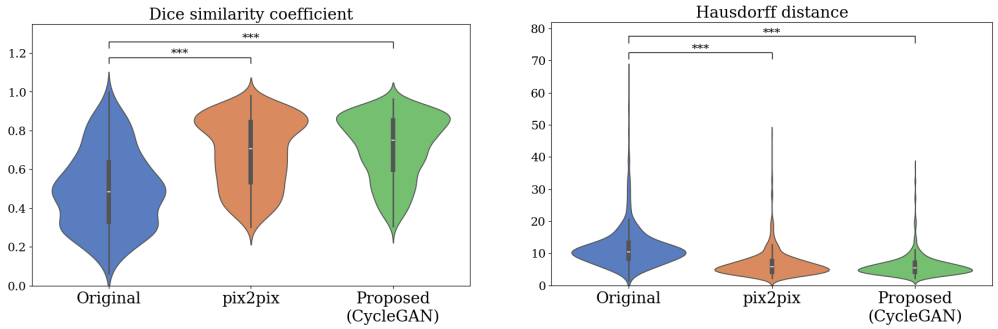
4.2.2. Quantitative Evaluation: DSC and HD

We present a quantitative evaluation of GAN-based motion artifact correction methods using the Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD) metrics. Our assessment encompasses image patches affected by motion artifacts, compared to reference patches, and evaluates the performance of the pix2pix model and CycleGAN model in correcting these artifacts. To accurately measure the DSC and HD metrics, we segmented the vessel region using a seed-based region-growing technique that allows for semi-automatic annotation of the vessel. The reference patches were extracted from the phases of the 4D CT scans that exhibited the least motion artifacts, typically the cardiac phases with minimal motion disturbance. By comparing the corrected patches against these motion-minimized references, we were able to quantitatively assess the effectiveness of the pix2pix and CycleGAN models in correcting motion artifacts.

For a visual representation of these metrics, refer to Figure 7, which includes two subfigures—(a) displaying a violin plot of the DSC, and (b) a violin plot for the HD—illustrating the performance comparisons between the methods. CycleGAN demonstrates the highest median DSC (0.751), surpassing pix2pix, indicating its superior effectiveness in mitigating motion artifacts. In terms of HD, CycleGAN achieves the lowest median value (5.385), signifying superior patch similarity to pix2pix. Detailed numerical values and statistical significance values of these findings are summarized in Table 2, which provides a comprehensive comparison of the DSC and HD values for pix2pix and CycleGAN.

We applied the Wilcoxon signed-rank test to assess performance differences between pix2pix and CycleGAN. This non-parametric test confirms significant improvements in both DSC and HD for both methods over the motion-affected patch baseline (p-values < 0.001). Notably, CycleGAN achieves significantly higher DSC and lower HD than pix2pix (p-value < 0.001), demonstrating its superior efficacy in mitigating motion artifacts.

In summary, CycleGAN outperforms pix2pix in motion artifact correction, with higher median DSC and lower median HD values. The reduced variability, indicated by narrower Interquartile Ranges (IQRs), underscores CycleGAN's robustness in mitigating motion artifacts across various image patches. These findings emphasize substantial and statistically significant performance improvements by CycleGAN over pix2pix in terms of DSC and HD metrics.



(**a**) Dice similarity coefficient (**b**) Hausdorff distance

Figure 7. Quantitative Evaluation Using DSC and HD Metrics: Two violin plots, showcasing the performance of the pix2pix model and the proposed (CycleGAN) model in motion artifact correction. (**a**) illustrates the distribution of Dice Similarity Coefficient (DSC) values, where CycleGAN demonstrates a higher median DSC. (**b**) displays the Hausdorff Distance (HD) metric, with CycleGAN achieving a lower median HD. Both plots include median and interquartile ranges. The "***" markers denote statistically significant differences as confirmed by the Wilcoxon signed-rank test.

Table 2. Performance Evaluation Metrics: DSC and HD.

	Method	**Median**	**IQR**	***p*-Value**
Dice similarity coefficient	Original	0.488	0.330–0.626	
	pix2pix	0.706	0.534–0.843	<0.001
	Proposed (CycleGAN)	0.751	0.598–0.853	<0.001
Hausdorff distance	Original	10.440	8.246–13.200	
	pix2pix	5.831	4.243–7.616	<0.001
	Proposed (CycleGAN)	5.385	4.123–7.000	<0.001

IQR: Interquartile Range, representing the range between the 25th and 75th percentiles.

4.2.3. Quantitative Evaluation: Motion Artifact Metrics

Next, we extended our investigation to include the application of GAN-based methods, specifically pix2pix and CycleGAN, to enhance the performance metrics for coronary artery imaging. Our quantitative analysis focused exclusively on normal coronary arteries, omitting diseased ones due to the challenging nature of these metrics in handling plaque.

We employed the motion artifact metrics as proposed in Jung et al. [14], measuring isotropy, fold overlap ratio (FOR), low-intensity region score (LIRS), and motion artifact score (MAS). Isotropy, quantifying the level of motion artifacts, is evaluated based on the ratio of two eigenvalues (λ_1, λ_2) of the vessel region's shape, aligning with methods used in Jeon et al. [33]. FOR, assessing the symmetry of vessel regions, and LIRS, measuring the shading effect of motion artifacts, are calculated following the methodologies outlined in Ma et al. [21].

The results show that pix2pix and CycleGAN demonstrated substantial improvements across all evaluated metrics, surpassing the performance of Jung et al. [14], as detailed in Table 3. A notable increase using GAN-based methods was observed for all metrics including Isotropy, FOR, LIRS, and MAS. Furthermore, the proposed approach using CycleGAN resulted in slightly improved performance compared to pix2pix, demonstrating the effectiveness of our method.

We also confirmed the statistical significance of the measurements, ascertained through a Wilcoxon signed-rank test, with p-values smaller than 0.001. This robust statistical evidence firmly supports the superiority of pix2pix and CycleGAN over the previous method, illustrating their marked effectiveness in reducing motion artifacts.

For the effective calculation of these metrics, it is crucial to accurately identify specific areas, including the vessel region affected by motion, areas exhibiting low-intensity shading artifacts, and the myocardium region. To streamline this process, we implemented a seed-based region-growing technique for semi-automatic annotation of the vessel and low-intensity artifact regions. This method enhances efficiency and assists in clearly defining the segment boundaries while ensuring consistent intensity across the region. Conversely, the myocardium regions are annotated manually due to their variable appearances and shapes, which makes them less suitable for automated region growing processes.

Table 3. Comparison of Methods Across Different Metrics.

Motion Artifact Metric	Method	Median	IQR	p-Value
Isotropy	Original	0.40	0.30–0.58	
	Jung et al. [14]	0.63	0.46–0.75	<0.001
	pix2pix	0.75	0.65–0.83	<0.001
	Proposed (CycleGAN)	0.74	0.63–0.84	<0.001
Fold Overlap Ratio (FOR)	Original	0.59	0.54–0.62	
	Jung et al. [14]	0.63	0.59–0.67	<0.001
	pix2pix	0.79	0.78–0.81	<0.001
	hlProposed (CycleGAN)	0.8	0.79–0.81	<0.001
Low-Intensity Region Score (LIRS)	Original	0.92	0.89–0.95	
	Jung et al. [14]	0.97	0.93–1.0	<0.001
	pix2pix	1.0	1.0–1.0	<0.001
	Proposed (CycleGAN)	1.0	1.0–1.0	<0.001
Motion Artifact Score (MAS)	Original	0.54	0.48– 0.59	
	Jung et al. [14]	0.61	0.55–0.67	<0.001
	pix2pix	0.79	0.78–0.81	<0.001
	Proposed (CycleGAN)	0.8	0.79–0.81	<0.001

4.2.4. Qualitative Evaluation

One experienced reader evaluated the degree of motion artifacts in coronary artery images using a 5-point Likert scale. The assessment was blinded; the reader was unaware of whether the images were pre- or post-application of the proposed motion correction methods. Images were randomly presented without any identification, including both original and motion-corrected patches.

Table 4 shows the distribution of the Likert scale scores for the test patches before and after motion correction using different methods, namely, Original, Jung et al., pix2pix, and CycleGAN. Initially, a significant majority of images (98.5%) were rated in the lower categories of the Likert scale (1, 2, and 3), indicating a high degree of motion artifacts. However, after applying the motion correction methods, this percentage substantially decreased to 35% for Original and further reduced with the application of advanced methods like pix2pix and CycleGAN, as demonstrated by the lower frequency of categories 1, 2, and 3 (0.5%, 3.5%, and 32.1% for pix2pix; 0.5%, 3.1%, and 30.6% for CycleGAN, respectively).

Moreover, the mean score based on the Likert scale improved markedly from 1.43 (±0.66) in the original images to 3.80 (±0.87) for Jung et al. [14], 3.82 (±0.84) for pix2pix, and 3.88 (±0.85) for CycleGAN, indicating a significant reduction in motion artifacts. The statistical significance of these improvements is confirmed with *p*-values smaller than 0.001.

In our comparative analysis of pix2pix and CycleGAN, we observed a notable superiority in the performance of CycleGAN, particularly in the context of motion artifact correction in coronary artery imaging. While both methods showed significant improvements in reducing motion artifacts, CycleGAN demonstrated a slight edge over pix2pix in several key metrics. For instance, CycleGAN's ability to better preserve the structural integrity of the coronary arteries was evident, as indicated by its higher scores in the Likert scale evaluation. Specifically, CycleGAN achieved a mean score of 3.88 (±0.85) compared to pix2pix's 3.82 (±0.84), suggesting a more effective reduction in noticeable motion artifacts.

Table 4. Motion artifact before and after applying proposed method

	Original	**Jung et al. [14]**	**pix2pix**	**Proposed (CycleGAN)**	***p*-Value**
Likert scale					<0.001
1 = Completely unreadable	66.5%	2.5%	0.5%	0.5%	
2 = Significant motion	26%	2.5%	3.5%	3.1%	
3 = Apparent motion	6%	30%	32.1%	30.6%	
4 = Minor motion	1.5%	41.5%	41.1%	39.8%	
5 = No motion	0%	23.5%	22.7%	25.9%	
Score					<0.001
Mean	1.43	3.80	3.82	3.88	
Standard deviation	±0.66	±0.87	±0.84	±0.85	

The sample results showcasing the comparative effectiveness of pix2pix and CycleGAN in motion artifact correction are illustrated in Figures 8 and 9. Accompanying each image set are the Likert score changes and source information, providing a comprehensive overview of the improvements.

Figure 8 highlights the varying degrees of enhancements achieved. These range from significant, where image quality substantially improves from completely unreadable to no motion, to more moderate improvements, such as reductions in apparent motion. Notably, pix2pix and CycleGAN both significantly enhance the clarity and visibility of the coronary arteries, yet CycleGAN often outperforms pix2pix, particularly in maintaining the integrity of the arterial structure.

In Figure 9, instances where pix2pix and CycleGAN are directly contrasted are presented. Section (a) of the figure illustrates cases where pix2pix inadvertently enhances areas outside the targeted coronary region, as indicated by yellow arrows. Conversely, section (b) shows the samples where pix2pix fails to rectify motion artifacts, whereas CycleGAN successfully mitigates these issues. These comparative examples underline the superior capability of CycleGAN in addressing complex motion artifacts.

Figure 10 illustrates the efficacy of our proposed motion correction algorithm, especially highlighting the seamless integration of 2D-patch-based corrections back into the 3D CT volume. The upper row of the figure displays original slices from the CT volume that were previously affected by motion artifacts. Notably, these slices include the right coronary artery (RCA), a region typically susceptible to motion-induced distortions.

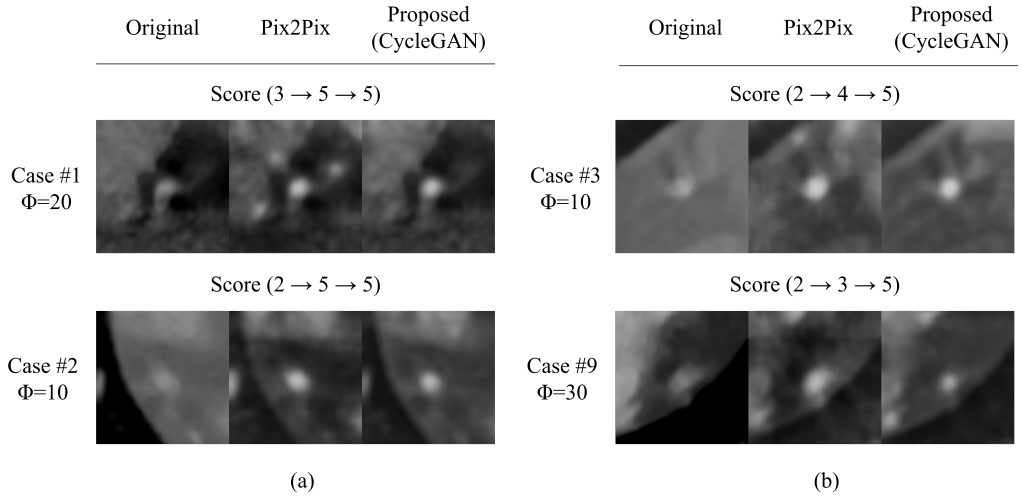


Figure 8. Qualitative examples of the original patch containing the mid-RCA, alongside the results obtained from applying pix2pix and the proposed method, CycleGAN. These samples showcase a range of improvements, with expert evaluations based on a 5-point Likert scale (1 = completely unreadable, 2 = significant motion, 3 = apparent motion, 4 = minor motion, 5 = no motion). Additionally, source information such as case number and phase ϕ in 4D CT is displayed to the left of each sample. (**a**) represents instances where the scores remained consistent between pix2pix and the proposed method, CycleGAN, while (**b**) highlights cases where the application of the proposed method, CycleGAN, resulted in improved scores compared to pix2pix.

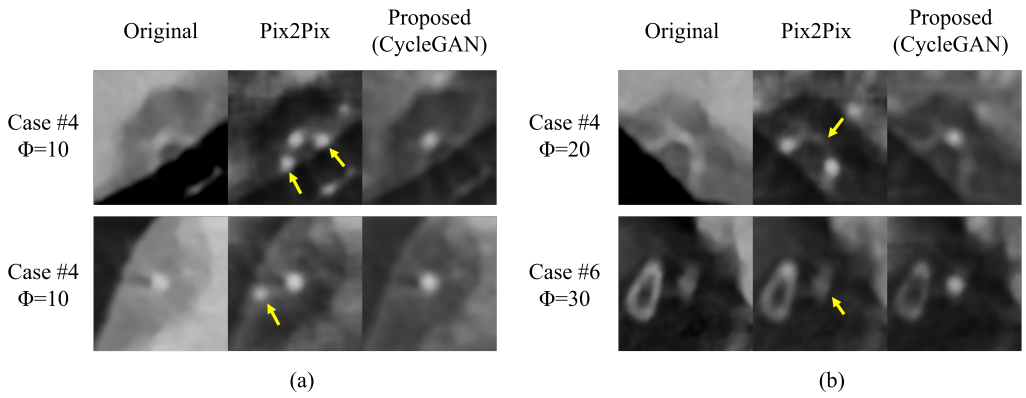


Figure 9. Comparative Analysis of pix2pix and CycleGAN in coronary motion correction. (**a**) Illustrates a case where pix2pix erroneously enhances regions outside the area of interest in the coronary artery, marked with yellow arrows to highlight the misenhanced areas. (**b**) Demonstrates cases where pix2pix fails to correct motion in the coronary artery, while CycleGAN successfully achieves motion correction. The regions where pix2pix fails are indicated with yellow arrows, contrasting with the successful correction by CycleGAN.

Following the application of our motion correction algorithm on these 2D patches, we reinserted them into the original 3D volume. The lower row of the figure showcases these post-correction slices. It is particularly remarkable to observe how the reinserted patches blend seamlessly with the surrounding areas, with no discernible artifacts or discontinuities. This seamless integration is further enhanced by the volumetric interpolation technique we employed, ensuring that the corrected patches maintain spatial and anatomical consistency with the adjacent uncorrected regions.

The RCA region, highlighted with green boxes in both rows, serves as a clear point of comparison. In the corrected slices, one can observe a marked reduction in motion artifacts, resulting in clearer and more defined images of the coronary artery. This comparison not only validates the effectiveness of our motion correction approach but also demonstrates

our algorithm's capability to maintain the integrity and continuity of the 3D structure in the CT volume.

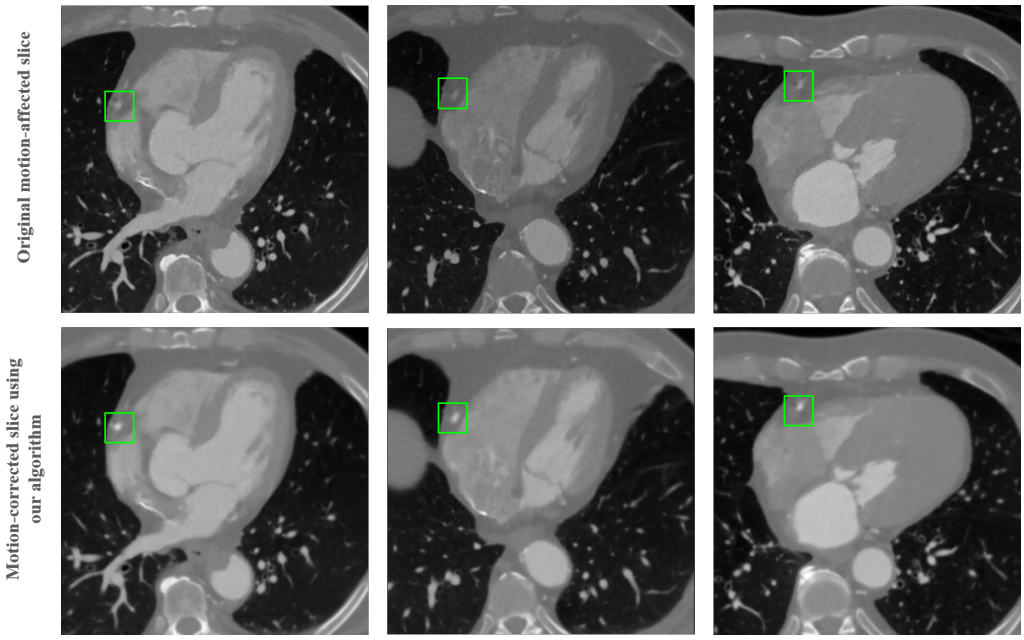


Figure 10. Comparison of axial views of a CT volume before and after applying the proposed coronary motion correction algorithm including re-insertion and volumetric interpolation. The upper row shows the original motion-affected slice and the lower shows the slice after motion correction using our algorithm. The RCA region is highlighted with green boxes.

5. Conclusions

In this study, we have advanced the field of motion correction in coronary computed tomography angiography (CCTA) by proposing a CycleGAN-based framework. Building upon the foundations laid by Jung et al. [14], this framework incorporates several key enhancements: (1) it focuses on 2D patches from the coronary artery for learning, (2) corresponding patches are extracted from different temporal phases in 3D volumes, capturing various degrees of motion, (3) CycleGAN is utilized for generating synthetic motion-corrected images, which serve as a superior alternative to the style transfer method used previously, (4) this approach enables the deep learning model to accurately learn motion correction from these synthetic images, and (5) during testing, motion-corrected patches are efficiently reinserted and interpolated back into the original 3D volumes.

Our comprehensive evaluation, both quantitative and qualitative, underscores the effectiveness of this updated method using phantom and clinical datasets. Through quantitative metrics such as the Dice Similarity Coefficient (DSC) and Hausdorff Distance (HD), CycleGAN demonstrated a remarkable improvement in motion artifact correction over the pix2pix model. Additionally, our qualitative analysis, conducted by an experienced reader using a 5-point Likert scale, reflected significant enhancements in image quality and clarity, especially in the mid-right coronary artery (RCA) region. The observed improvements in motion artifact metrics—such as Isotropy, Fold Overlap Ratio (FOR), and Low-Intensity Region Score (LIRS)—were statistically significant, with p-values less than 0.001, further affirming the superiority of our CycleGAN-based approach.

One of limitations of our study is the evaluation based on two specific datasets, which may not adequately represent the full spectrum of imaging environments and patient populations. This could restrict the broader applicability of our CycleGAN-based approach in diverse clinical settings. Additionally, the inherent complexity of deep learning models like CycleGAN raises challenges in interpretability, which could impact clinical trust and understanding.

In future research, we intend to expand the evaluation of our method by including a wider variety of datasets, enhancing our understanding of its robustness and adaptability in different clinical settings. Building on this, our subsequent efforts will concentrate on enhancing the transparency and interpretability of deep learning models, with a particular focus on our CycleGAN-based framework. In conclusion, the promising results achieved with our CycleGAN-based approach open new avenues for enhancing the clinical utility and diagnostic accuracy of CCTA. These advancements have the potential to significantly impact patient care, allowing for more accurate diagnosis and assessment of coronary artery diseases.

Author Contributions: Conceptualization, methodology, and writing, A.M.S., S.J. and S.L.; software and vaildation A.M.S. and S.J.; resources, S.J., H.-J.C. and S.L.; data curation, S.J. and H.-J.C.; supervision, project administration, and funding acquisition, S.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) through the Government of the Republic of Korea [Ministry of Science and ICT (MIST)] under Grant NRF-2021R1A2C2095452.

Institutional Review Board Statement: Ethical review and approval were waived for this study as it used existing data without collecting or recording personal identification information, referring to information already in existence at the time of the study and utilized retrospectively for research purposes.

Informed Consent Statement: Patient consent was waived due to the use of existing data without collecting or recording personal identification information, referring to information already in existence at the time of the study and utilized retrospectively for research purposes.

Data Availability Statement: Publicly available datasets were analyzed in this study. These data can be found here: https://www5.cs.fau.de/research/software/cavarev/, accessed: 13 February 2024. The clinical datasets presented in this article are not readily available because the data are part of an ongoing study. Requests to access the datasets should be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CCTA	Coronary Computed Tomography Angiography
GAN	Generative Adversarial Network
DTW	Dynamic Time Warping
RCA	Right Coronary Artery
DSC	Dice Similarity Coefficient
HD	Hausdorff Distance
SSIM	Structural Similarity Index
CNN	Convolutional Neural Network
SynGT	Synthetic Ground Truth
FOR	Fold Overlap Ratio
LIRS	Low-Intensity Region Score
MAS	Motion Artifact Score

References

1. Miller, J.M.; Rochitte, C.E.; Dewey, M.; Arbab-Zadeh, A.; Niinuma, H.; Gottlieb, I.; Paul, N.; Clouse, M.E.; Shapiro, E.P.; Hoe, J.; et al. Diagnostic performance of coronary angiography by 64-row CT. *N. Engl. J. Med.* **2008**, *359*, 2324–2336. [CrossRef] [PubMed]
2. Isola, A.A.; Grass, M.; Niessen, W.J. Fully automatic nonrigid registration-based local motion estimation for motion-corrected iterative cardiac CT reconstruction. *Med. Phys.* **2010**, *37*, 1093–1109. [CrossRef] [PubMed]
3. Tang, Q.; Cammin, J.; Srivastava, S.; Taguchi, K. A fully four-dimensional, iterative motion estimation and compensation method for cardiac CT. *Med. Phys.* **2012**, *39*, 4291–4305. [CrossRef] [PubMed]

4. Bhagalia, R.; Pack, J.D.; Miller, J.V.; Iatrou, M. Nonrigid registration-based coronary artery motion correction for cardiac computed tomography. *Med. Phys.* **2012**, *39*, 4245–4254. [CrossRef]
5. Rohkohl, C.; Bruder, H.; Stierstorfer, K.; Flohr, T. Improving best-phase image quality in cardiac CT by motion correction with MAM optimization. *Med. Phys.* **2013**, *40*, 031901. [CrossRef] [PubMed]
6. Kim, S.; Chang, Y.; Ra, J.B. Cardiac motion correction based on partial angle reconstructed images in X-ray CT. *Med. Phys.* **2015**, *42*, 2560–2571. [CrossRef] [PubMed]
7. Hahn, J.; Bruder, H.; Rohkohl, C.; Allmendinger, T.; Stierstorfer, K.; Flohr, T.; Kachelrieß, M. Motion compensation in the region of the coronary arteries based on partial angle reconstructions from short-scan CT data. *Med. Phys.* **2017**, *44*, 5795–5813. [CrossRef] [PubMed]
8. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
9. Hinton, G.; Deng, L.; Yu, D.; Dahl, G.E.; Mohamed, A.r.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.N.; et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.* **2012**, *29*, 82–97. [CrossRef]
10. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
11. Gondara, L. Medical image denoising using convolutional denoising autoencoders. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW), Barcelona, Spain, 12–15 December 2016; pp. 241–246.
12. Xu, L.; Ren, J.S.; Liu, C.; Jia, J. Deep convolutional neural network for image deconvolution. In Proceedings of the Advances in Neural Information Processing Systems (NIPS), Montreal, QC, Canada, 8–13 December 2014; pp. 1790–1798.
13. Jung, S.; Lee, S.; Jeon, B.; Jang, Y.; Chang, H.J. Deep learning based coronary artery motion artifact compensation using style-transfer synthesis in CT images. In Proceedings of the Simulation and Synthesis in Medical Imaging: Third International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 16 September 2018; Proceedings 3; Springer: Berlin/Heidelberg, Germany, 2018; pp. 100–110.
14. Jung, S.; Lee, S.; Jeon, B.; Jang, Y.; Chang, H.J. Deep learning cross-phase style transfer for motion artifact correction in coronary computed tomography angiography. *IEEE Access* **2020**, *8*, 81849–81863. [CrossRef]
15. Lossau, T.; Nickisch, H.; Wissel, T.; Bippus, R.; Schmitt, H.; Morlock, M.; Grass, M. Motion estimation and correction in cardiac CT angiography images using convolutional neural networks. *Comput. Med. Imaging Graph.* **2019**, *76*, 101640. [CrossRef] [PubMed]
16. Maier, J.; Lebedev, S.; Erath, J.; Eulig, E.; Sawall, S.; Fournié, E.; Stierstorfer, K.; Lell, M.; Kachelrieß, M. Deep learning-based coronary artery motion estimation and compensation for short-scan cardiac CT. *Med. Phys.* **2021**, *48*, 3559–3571. [CrossRef] [PubMed]
17. Zhang, L.; Jiang, B.; Chen, Q.; Wang, L.; Zhao, K.; Zhang, Y.; Vliegenthart, R.; Xie, X. Motion artifact removal in coronary CT angiography based on generative adversarial networks. *Eur. Radiol.* **2023**, *33*, 43–53. [CrossRef] [PubMed]
18. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
19. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 1–9.
20. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
21. Ma, H.; Gros, E.; Szabo, A.; Baginski, S.G.; Laste, Z.R.; Kulkarni, N.M.; Okerlund, D.; Schmidt, T.G. Evaluation of motion artifact metrics for coronary CT angiography. *Med. Phys.* **2018**, *45*, 687–702. [CrossRef] [PubMed]
22. Ren, P.; He, Y.; Zhu, Y.; Zhang, T.; Cao, J.; Wang, Z.; Yang, Z. Motion artefact reduction in coronary CT angiography images with a deep learning method. *BMC Med. Imaging* **2022**, *22*, 184. [CrossRef] [PubMed]
23. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image Style Transfer Using Convolutional Neural Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2414–2423. [CrossRef]
24. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
25. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE, Miami, FL,USA, 20–25 June 2009; pp. 248–255.
26. Rohkohl, C.; Lauritsch, G.; Keil, A.; Hornegger, J. CAVAREV—An open platform for evaluating 3D and 4D cardiac vasculature reconstruction. *Phys. Med. Biol.* **2010**, *55*, 2905. [CrossRef] [PubMed]
27. CAVAREV. Available online: https://www5.cs.fau.de/research/software/cavarev/ (accessed on 13 February 2024).
28. Keil, A.; Vogel, J.; Lauritsch, G.; Navab, N. Dynamic cone-beam reconstruction using a variational level set formulation. In Proceedings of the International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully3D), New York, NY, USA, 16–21 July 2009; pp. 319–322.
29. Taubmann, O.; Unberath, M.; Lauritsch, G.; Achenbach, S.; Maier, A. Spatio-temporally regularized 4D cardiovascular C-arm CT reconstruction using a proximal algorithm. In Proceedings of the IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 52–55.

30. Rohkohl, C.; Lauritsch, G.; Nottling, A.; Prummer, M.; Hornegger, J. C-arm ct: Reconstruction of dynamic high contrast objects applied to the coronary sinus. In Proceedings of the IEEE Nuclear Science Symposium Conference Record, Dresden, Germany, 19–25 October 2008; pp. 5113–5120.
31. Schwemmer, C.; Rohkohl, C.; Lauritsch, G.; Müller, K.; Hornegger, J. Residual motion compensation in ECG-gated interventional cardiac vasculature reconstruction. *Phys. Med. Biol.* **2013**, *58*, 3717. [CrossRef] [PubMed]
32. Schwemmer, C.; Rohkohl, C.; Lauritsch, G.; Müller, K.; Hornegger, J.; Qi, J. Opening windows-increasing window size in motion-compensated ECG-gated cardiac vasculature Reconstruction. In Proceedings of the International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology Nuclear Medicine, Lake Tahoe, CA, USA, 16–21 June 2013; pp. 50–53.
33. Jeon, B.; Hong, Y.; Han, D.; Jang, Y.; Jung, S.; Hong, Y.; Ha, S.; Shim, H.; Chang, H.J. Maximum a posteriori estimation method for aorta localization and coronary seed identification. *Pattern Recognit.* **2017**, *68*, 222–232. [CrossRef]