

Article

Reliable Off-Resonance Correction in High-Field Cardiac MRI Using Autonomous Cardiac B_0 Segmentation with Dual-Modality Deep Neural Networks

Xinqi Li ^{1,2}, Yuheng Huang ^{3,4}, Archana Malagi ¹, Chia-Chi Yang ¹, Ghazal Yoosefian ³, Li-Ting Huang ¹, Eric Tang ¹, Chang Gao ⁵, Fei Han ⁵, Xiaoming Bi ⁵, Min-Chi Ku ², Hsin-Jung Yang ^{1,*} and Hui Han ^{6,*}

- ¹ Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA; xinqi.li@mdc-berlin.de (X.L.); archana.malagi@cshs.org (A.M.); chia-chi.yang@cshs.org (C.-C.Y.); li-ting.huang@cshs.org (L.-T.H.); erictang0220@gmail.com (E.T.)
- ² Berlin Ultrahigh Field Facility (B.U.F.F.), Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany; min-chi.ku@mdc-berlin.de
- ³ Krannert Cardiovascular Research Center, Indiana University School of Medicine, Indianapolis, IN 46202, USA; yuhenghuang@g.ucla.edu (Y.H.); gyoosefi@iu.edu (G.Y.)
- ⁴ Bioengineering, University of California, Los Angeles, Los Angeles, CA 90095, USA
- ⁵ MR R&D Collaborations, Siemens Medical Solutions Inc., Los Angeles, CA 90048, USA; chang.gao@siemens-healthineers.com (C.G.); fei.han@siemens-healthineers.com (F.H.); xiaoming.bi@siemens-healthineers.com (X.B.)
- ⁶ Department of Radiology, Weill Medical College of Cornell University, New York, NY 10065, USA
- * Correspondence: hsin-jung.yang@cshs.org (H.-J.Y.); huh4006@med.cornell.edu (H.H.)
- † These authors contributed equally to this work.



Citation: Li, X.; Huang, Y.; Malagi, A.; Yang, C.-C.; Yoosefian, G.; Huang, L.-T.; Tang, E.; Gao, C.; Han, F.; Bi, X.; et al. Reliable Off-Resonance Correction in High-Field Cardiac MRI Using Autonomous Cardiac B_0 Segmentation with Dual-Modality Deep Neural Networks. *Bioengineering* **2024**, *11*, 210. <https://doi.org/10.3390/bioengineering11030210>

Academic Editors: Crescenzo Gallo and Gianluca Zaza

Received: 18 January 2024

Revised: 9 February 2024

Accepted: 18 February 2024

Published: 23 February 2024

Abstract: B_0 field inhomogeneity is a long-lasting issue for Cardiac MRI (CMR) in high-field (3T and above) scanners. The inhomogeneous B_0 fields can lead to corrupted image quality, prolonged scan time, and false diagnosis. B_0 shimming is the most straightforward way to improve the B_0 homogeneity. However, today's standard cardiac shimming protocol requires manual selection of a shim volume, which often falsely includes regions with large B_0 deviation (e.g., liver, fat, and chest wall). The flawed shim field compromises the reliability of high-field CMR protocols, which significantly reduces the scan efficiency and hinders its wider clinical adoption. This study aims to develop a dual-channel deep learning model that can reliably contour the cardiac region for B_0 shim without human interaction and under variable imaging protocols. By utilizing both the magnitude and phase information, the model achieved a high segmentation accuracy in the B_0 field maps compared to the conventional single-channel methods (Dice score: 2D-mag = 0.866, 3D-mag = 0.907, and 3D-mag-phase = 0.938, all $p < 0.05$). Furthermore, it shows better generalizability against the common variations in MRI imaging parameters and enables significantly improved B_0 shim compared to the standard method (SD(B_0 Shim): Proposed = $15 \pm 11\%$ vs. Standard = $6 \pm 12\%$, $p < 0.05$). The proposed autonomous model can boost the reliability of cardiac shimming at 3T and serve as the foundation for more reliable and efficient high-field CMR imaging in clinical routines.

Keywords: cardiac MRI; B_0 shim; B_0 field map; dual modality image segmentation



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cardiac Magnetic Resonance Imaging (CMR) represents a pivotal advancement in cardiac care, offering a comprehensive and non-invasive approach to assessing heart structure, function, and myocardial tissue characterization. CMR utilizes the magnetic resonance signal from water protons in the heart to provide pathologically sensitive signals without exposing patients to ionizing radiation. This allows the application of CMR to extend beyond mere anatomical visualization and play a crucial role in the evaluation of myocardial perfusion, ventricular contractility, myocardial viability, and myocardial tissue

composition [1–5], making it the preferred modality in the diagnosis and management of a variety of cardiac conditions, such as cardiomyopathies, heart failures, and congenital heart diseases [6–10].

1.1. B_0 Inhomogeneity Effect on 3.0T CMR

Since the FDA approved the use of 3.0T scanners for whole-body clinical applications in 2002, the adoption of high-field scanners has been rising rapidly, particularly for neuroimaging. In general, 3.0T provides higher SNR, spatial resolution, and reduced scan time than 1.5T. In some facilities, 3.0T may be the only available field strength. However, 3.0T adoption for CMR has been relatively slow. For CMR, the superior SNR at 3.0T provides high potential to facilitate accelerated imaging with enhanced spatial and temporal resolution, which can be further optimized through techniques such as compressed sensing and deep learning [11–13]. The increased T1 relaxation times at 3.0T augment T1-weighted imaging, improving the diagnostic quality of late gadolinium enhancement (LGE) and first-pass perfusion methods [14]. This leads to improved myocardial tissue characterization. Moreover, the amplified T2* contrast inherent to 3.0T MRI allows for a more effective assessment of iron deposition [11,12], hemorrhage [13], and oxygen consumption [14,15], which provides critical information for comprehensive evaluation of cardiac pathology. Furthermore, the improved spectral separation at 3.0T enhances metabolic imaging, magnetization transfer, and magnetic resonance spectroscopy imaging, allowing for refined chemical exchange saturation transfer (CEST) imaging [15,16] and more effective fat suppression in coronary imaging. Collectively, these advantages have the potential for 3.0T systems to elevate the diagnostic capabilities of CMR.

However, despite its numerous benefits, high-field cardiac MRI poses unique challenges that hinder its wider clinical adoption. A major challenge for 3.0T CMR remains the increased B_0 inhomogeneity from the amplified main field, particularly at the tissue–air interface due to susceptibility variations [17–19].

B_0 field homogeneity is critical for optimal CMR imaging, particularly when leveraging the high signal-to-noise (SNR) benefits of 3.0T systems. Steady-state free precession (SSFP), a cornerstone CMR sequence at 1.5T due to its rapid acquisition and high SNR, encounters significant challenges at 3.0T. B_0 inhomogeneity at this higher field strength results in substantial signal variability and banding artifacts [17,18], as shown in Figure 1. Similarly, echo-planar imaging (EPI)—despite its efficiency—is vulnerable to B_0 inhomogeneity, leading to image distortion and signal loss at 3.0T [20]. This inhomogeneity further undermines the consistency of T2*-based sequences [2] and the effectiveness of fat suppression techniques, both imperative for detailed myocardial and coronary artery visualization [21]. To capitalize on the SNR advantages of 3.0T CMR, reliable B_0 shimming techniques and sequence adaptations are necessary to acquire high-quality diagnostic images.

1.2. Cardiac B_0 Shimming to Improve Image Quality of High-Field CMR

Active B_0 shimming is the most direct way to correct for B_0 field inhomogeneities [22–24]. By creating a correction B_0 field from shim coils, B_0 shimming adjusts the static magnetic field across the imaging volume to ensure field uniformity. Recent advancements in shimming technology have led to the development of advanced shim hardware, like RF coils with integrated B_0 shimming [24,25], capable of generating high-order shim fields to correct for the unique off-resonance patterns in the heart. However, the heart's shape, location, and tissue composition make accurately measuring the B_0 field challenging. In today's standard clinical practice, manual selection of a shim box is used to identify shim volumes for B_0 off-resonance estimation [26]. Yet, there are common confounders, such as the false inclusion of non-cardiac off-resonance sources (e.g., liver and chest wall) and fat-induced chemical shifts in the shim box. These confounders can significantly degrade the accuracy of shim field estimation, leading to failed B_0 shimming and further degradation of the image quality [19,27,28], as shown in Figure 2. To facilitate accurate B_0 shimming, it's critical to precisely delineate the heart region and ensure the shim coils can generate the optimized cancellation shim field in the heart and

improve cardiac image quality [27]. While prior works have shown manual contouring of the cardiac region can improve the shim robustness and shimming accuracy, manually contouring the region of interest in CMR images is time-consuming and expensive, making it impractical for clinical settings. An autonomous and reliable segmentation method for CMR B_0 maps is desired to improve B_0 shimming robustness and facilitate reliable high-field CMR scans.

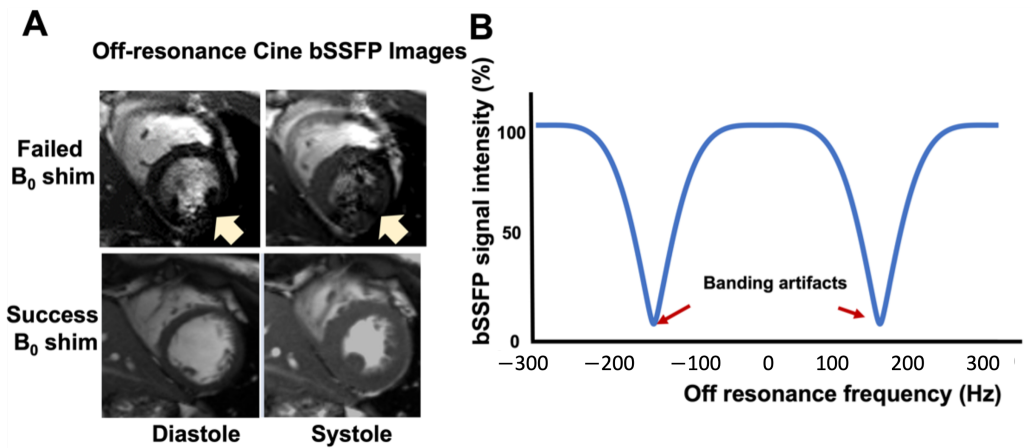


Figure 1. Banding artifacts in the clinical SSFP cine images under failed B_0 shim can corrupt the image quality. Representative short-axis cine images with failed and successful B_0 shim are presented in (A). The banding artifact in the blood pool leads to severe imaging artifacts (arrows) and makes the images unreadable. (B) The signal intensity of the SSFP sequence in the presence of B_0 off-resonance (TR = 3.3 ms, phase cycle = 180 d).

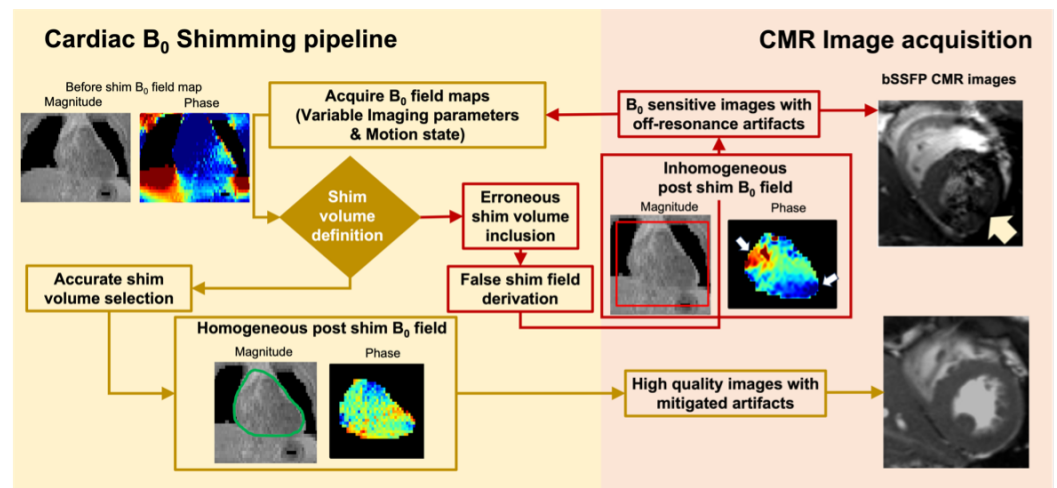


Figure 2. Cardiac B_0 shimming pipeline. A schematic flow chart of the CMR B_0 shimming workflow is presented. In today’s clinical practice, the manual selection of a shim box is used to derive the shim currents. The use of a rigid shim box can include undesired off-resonance fields outside of the heart and lead to failed B_0 shimming for B_0 -sensitive CMR images.

1.3. State-of-the-Art CMR Segmentation Models Are Not Optimized for B_0 Field Maps

Image segmentation has been a crucial task in CMR applications. Manual segmentation is time-consuming and labor-expensive so conventionally there are some semi-automatic techniques such as threshold [29], region grow [30] and contour-based methods [31].

With the rapid evolution of deep learning techniques, multiple automatic segmentation models have been proposed in recent years [32–37]. The U-Net structure, characterized by a symmetric encoder and decoder with skip connections, has succeeded greatly across various medical imaging domains [35]. In the encoder, multiple convolutional and down-

sampling layers are used continuously to extract the feature information. The decoder up-samples the extracted feature within a large receptive field to the input resolution to enable pixel-level semantic prediction. The skip-connection between each layer of encoder and decoder helps mitigate the information degradation during down-sampling and up-sampling [38]. Following this elegant design, model structures such as 3D U-Net [39], U-Net++ [36], nnU-Net [33] have been developed. The nnU-Net framework, in particular, is an automated configuring framework that provides a standardized and efficient pipeline with highly accurate segmentation results [33]. In the evolving landscape of deep learning-based automatic segmentation methods, its application in CMR images has also emerged as a pivotal tool for CMR image analysis [40,41]. Although convolutional neural networks have achieved huge success in segmentation tasks, transformer based methods and generative models have been popular recently due to their great success in other computer vision tasks. For example, BerDiff [42] utilized the conditional Bernoulli Diffusion model, ViT-FRD [43] that combines a visual transformer and a CNN through knowledge refinement, and Swin-UNETR [32] that incorporated the vision transformer into a U-structure.

Segmentation models have traditionally been designed for single-channel signal magnitude variations. However, regarding B_0 field maps in MRI, the single-channel approach faces limitations due to their unique contrast characteristics. Standard MRI segmentation models do not work well with B_0 field maps as these images are often proton-density weighted and have low soft tissue contrasts. The challenge is particularly evident in cardiac B_0 field maps, where the heart and liver are in proximity and have unclear boundaries. As a result, these models struggle to provide accurate cardiac boundary delineation in cardiac B_0 field maps for shim field derivation. Since MRI is based on the magnetic resonance of water protons, MRI images consist of both magnitude and phase signal data. Phase images in MRI provide crucial spatial information by reflecting the chemical composition and local field inhomogeneity within the tissue. The phase images of the heart and liver can exhibit distinctive features due to local frequency changes influenced by their respective location, shape, and orientation in relation to both the air-filled space and the B_0 direction. However, this aspect of MRI is often overlooked despite its potential for improving the reliability and robustness of segmentation algorithms.

In this study, we hypothesize that by incorporating phase images, which are inherently sensitive to organ boundaries, segmentation accuracy can be significantly enhanced. The main contributions are:

1. We developed a dual-channel CNN model to improve cardiac segmentation for B_0 shimming in high-field CMR by combining magnitude and phase images.
2. We thoroughly evaluated the performance of the proposed model under different imaging parameters and compared it with state-of-the-art medical image segmentation techniques. Besides, we demonstrated the generalizability of dual-channel module on different existing models to improve the performance.
3. We further demonstrated the application of this dual-channel segmentation model in providing the foundation of high-quality B_0 shimming in the heart.

2. Methodology

2.1. Image Preprocessing

Given that various imaging parameter sets were employed in experiments, a standardized image pre-processing routine was applied to calibrate the data. The data pre-processing included the following steps:

- (1) *Background removal* : The raw data contained some redundant air introduced during the image acquisition and reconstruction. As a first step, Otsu's method [44] was derived from the magnitude maps with number of threshold values equal to 2. The rough mask was generated based on the threshold level and followed by a post-processed operation using morphological closing. The structuring element was a disk-shape one defined by the resolution of the image and applied to both the

magnitude and phase map. It helped effectively segregate the region of interest from extraneous air.

- (2) *Resolution and FOV Alignment:* The voxel spacing within our acquired data was heterogeneous. The large spacing might cause the loss of detailed information, while the small spacing requires a larger computational budget. To reconcile this, we established a target voxel spacing based on the median spacing observed across all subjects for each axis. Given the anisotropic nature of our dataset, we resampled all images to the uniform target voxel spacing using third-order spline interpolation. Subsequently, images were either cropped or padded to match the dimension at the center region, if necessary.
- (3) *Noise Standardization:* All images were normalized based on mean and standard deviation values per case. The normalization step ensured all data conformed to a consistent scale and distribution.
- (4) *Dataset split:* The T1-w data set, including 54 subjects, was randomly partitioned into a training set (comprising 40 volumes) and a test set (comprising 14 volumes). Additionally, 10 PD-w volumes from subjects not included in the training set were reserved for an independent test set to validate the model's generalizability across varied imaging protocols.

2.2. Model Architecture

In this study, we proposed a two-channel segmentation model built based on the nnU-Net, integrating both the magnitude and phase information for heart segmentation in CMR images. As shown in Figure 3, the magnitude map cannot provide a clear contour for the heart region, while the phase map can provide additional information to delineate the heart. The general pipeline is shown in Figure 4. The magnitude and phase map will be pre-processed and concatenated into a 4D matrix with the additional channel dimension. Then we trained three different models, namely 2D-mag-net, 3D-mag-net, 3D-mag-phase-net, where 2D-mag-net was a 2D U-Net model, 3D-mag-net was a single channel nnU-Net model only using magnitude information and 3D-mag-phase-net was a dual-channel nnU-Net based model using magnitude and phase information. As we used the cross-fold validation in training, during the inference, we ensemble the softmax probabilities of 5 folds to predict the segmentation. Besides, we applied the connected component-based post-processing [45] to eliminate the obvious false positives and generate the final prediction.

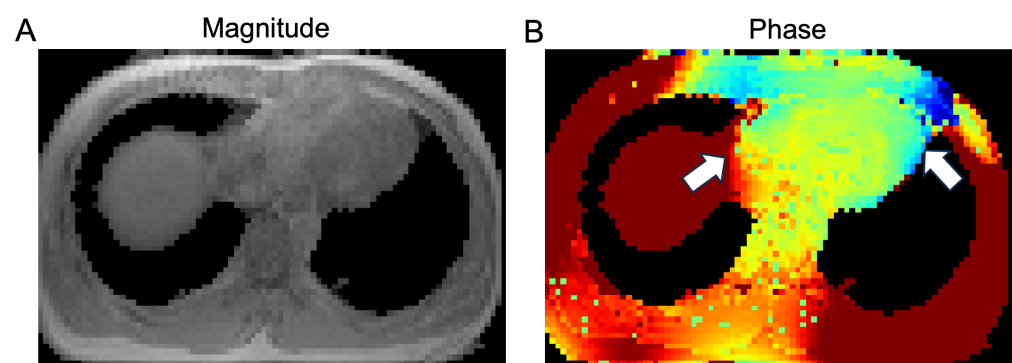


Figure 3. Magnitude and phase map from axial view. The magnitude map (A) image does not show a differentiable boundary as the white arrows show (the heart-liver boundary and the heart-lung interface), while the phase map (B) can provide the distinctive differences due to local frequency changes influenced by their respective location, shape and orientation in relation to both the air-filled space and the B_0 direction.

The proposed dual-channel nnU-Net-based model followed the original U-Net [35] structure with 5 encoder and decoder layers. Each layer included two convolutional blocks. Within each convolutional block in the encoder and decoder, we incorporated a $3 \times 3 \times 3$ convolution with stride 2, an instance normalization and a leaky ReLU nonlinearity with a negative slope of 0.01. Notably, we utilized the leaky ReLU activation function, which differs from standard ReLU because of its smaller slope for negative values. In the encoder, strided convolutions with a stride size of 2 were employed for down-sampling, while the transposed convolution was used in the decoder to up-sample the feature map. We employed the dice loss function, defined as follows:

$$L_{Dice} = - \frac{2 \sum_i o_i y_i}{\sum_i o_i + \sum_i y_i}, \quad (1)$$

where o_i represents the voxel's value from the labeled volume and y_i represents the voxel's value from the predicted volume.

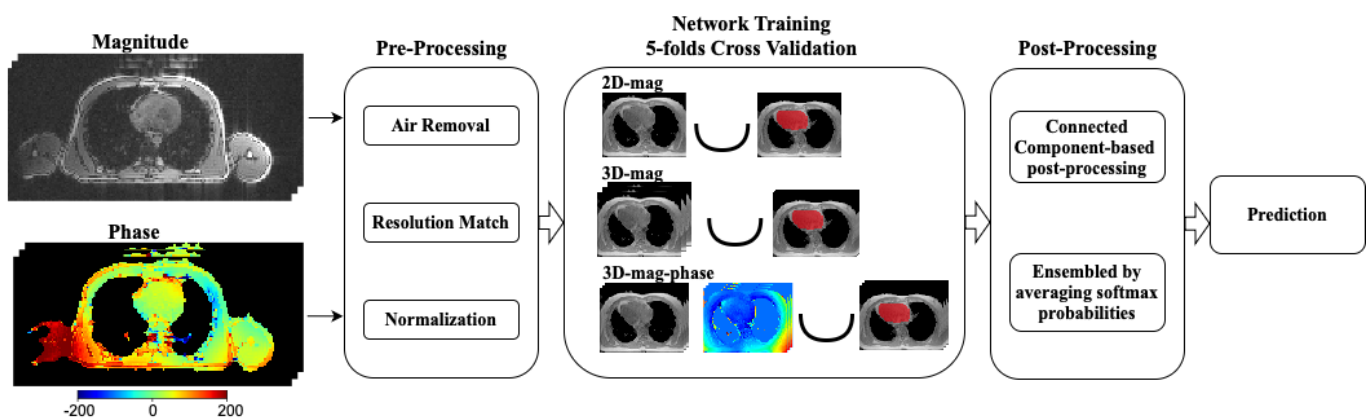


Figure 4. The general workflow of our experiments. The magnitude and phase maps are preprocessed and fed into several neural networks. We train 2D-mag-net, 3D-mag-net and 3D-mag-phase-net using 5-fold cross-validation. The segmentation outcomes are generated through the post-processing based on the output of the softmax layer.

2.3. Training Strategy

Previous studies demonstrated that a large patch size is important for model training, as a small batch size leads to noisier gradients during the training, and a larger patch size allows the aggregation of more information [46]. To accommodate large patch sizes, we maintained a modest batch size of 2, with the patch size tailored for different configurations. The stochastic gradient descent with Nesterov momentum and an initial learning rate of 0.01 was used to learn weights. Each network was trained for 1000 epochs, with each epoch consisting of 250 mini-batches. To prevent the drastic reduction of the number of samples that can be used for learning, we implemented cross-validation during training, using different portions of the training set as train and validation data, allowing better evaluation of the performance. The training set was divided into k ($k = 5$) smaller sets and one subset as validation set was used each time during training. During the inference, k models were averagely ensembled to predict the segmentation.

2.4. Data Augmentation

In medical images, previous studies have proved that data augmentation is critical given the limited data samples and the complexity of medical images. Several data augmentation strategies were applied on the fly by probability during the training phase: rotation, scaling, Gaussian noise, simulation of low resolution, and mirroring. We found that scaling was significant in our case due to the nature of different patient sizes. Comprehensive details are provided in the Experiments section. The augmentation was implemented using TorchIO [47]. In addition to the standard augmentation parameters, the apparent size of the

subjects is a common variation in medical images (Figure 5). Although the problem of object size variation has been extensively studied in machine learning algorithms, especially for object detection tasks [48], its application in medical imaging models is more complicated with the limited field of view and partially covered body parts. To mitigate this issue, we further simulated the diverse patient sizes. This was accomplished by either cropping or padding the image from its center and rescaling it to induce the FOV coverage effect.

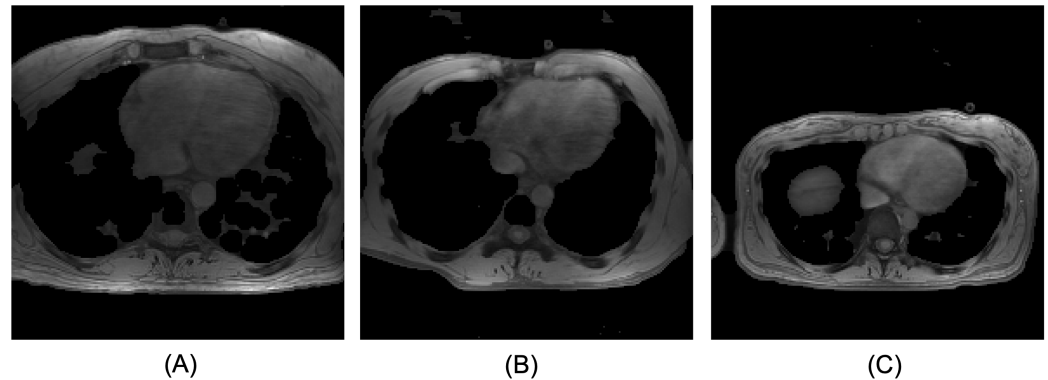


Figure 5. Axial views of normalized magnitude image. Subject (A) from back to chest distance is 229.69 mm, subject (B) is 203.13 mm and subject (C) is 154.12 mm. Various patient sizes exist and might affect the segmentation performance if it has not been addressed properly.

2.5. Evaluation Metrics

We measured the performance based on two key metrics: (1) Dice Score. The Dice score computed the overlap between two volumes, varying from 0 (mismatch) to 1 (perfect match). It is defined as:

$$\mathcal{D}(A, B) = 2 \frac{A \cap B}{A + B}, \quad (2)$$

where A and B correspondingly denote the sets of heart voxels in ground-truth and predicted volumes. (2) 95% Hausdorff Distance. The Hausdorff distance (HD) calculated the maximum distance between two volumes. It is defined as:

$$\mathcal{H}(A, B) = \max_{A, B} \{d_{AB}, d_{BA}\} = \max_{A, B} \left\{ \max_{x \in A} \min_{y \in B} d(x, y), \max_{y \in B} \min_{x \in A} d(x, y) \right\}, \quad (3)$$

where d represents Euclidean distance. To calculate the 95% HD, the calculation is based on the 95th percentile of the distances between boundary points in A and B . Lower values of the 95% HD indicate superior segmentation performance. (3) Jaccard Index. The Jaccard Index or named as Jaccard similarity coefficient was defined as the Intersection over Union (IoU) between the ground-truth and predicted results [49].

2.6. Statistical Analysis

All the data are represented as mean \pm standard deviation (SD). We performed paired t-test for paired comparisons and repeated measures of Analysis of Variance (ANOVA) for 3 way comparisons (2D-mag, 3D-mag, 3D-mag-phase), using the Scipy [50] and Statsmodels [51] in Python 3.10. The post-hoc analysis based on the pairwise t-tests with Bonferroni correlation.

3. Results

Three neural networks (2D-mag-net, 3D-mag-net, 3D-mag-phase-net) were implemented using PyTorch based on nnU-Net. These networks were initially trained on a dataset including 40 T1-w CMR images. Subsequently, models were applied to test data without fine-tuning or further training. The developed model is available at <https://github.com/lixinqi98/DynamicShim>, accessed on 18 January 2024.

3.1. Datasets

All images were acquired in a 3T MR systems (Biograph mMR, Siemens Medical Solutions, Erlangen, Germany). Healthy volunteers ($n = 64$) were recruited per the protocol reviewed and approved by the institutional review board (IRB), No.23469. Every participant was competent, provided written informed consent, and had no history of coronary artery disease, lung disease, abnormal cardiac rhythm and rate, kidney or liver disease, and was not contraindicated for cardiac MR examinations (i.e., they completed a detailed cardiac MRI questionnaire). To explore the generalizability of the proposed method, B_0 field maps were acquired with two sets of imaging parameters. One set (T1 weighted B_0 maps, T1-w) included 54 different healthy subjects' cardiac volume with TE1/TE2 = 1.31/3.53 ms; Flip angle = 16° ; FOV = $400 \times 300 \times 250 \text{ mm}^3$; Spatial resolution = $3.57 \times 3.57 \times 5.2 \text{ mm}^3$. Echo spacing = 2.1 ms. Another set (Proton density weighted B_0 maps, PD-w) includes 10 healthy subjects acquired with TE1/TE spacing = 1.42/2.01 ms; number of echoes = 6; flip angle = 8° ; FOV = $300 \times 300 \times 120 \text{ mm}^3$; spatial resolution = $1.56 \times 1.56 \times 5 \text{ mm}^3$. The ground-truth segmentations were manually drawn by experts (MRI scientist and Radiologist). All the data follows the image preprocessing in Section 2.1, matched to median resolution $3.57 \times 3.57 \times 5.2 \text{ mm}^3$ and size of $96 \times 96 \times 40$. According to the data augmentation in Section 2.4, the summary of data before and after augmentation can be found in Table 1.

Table 1. Data summary before and after preprocessing augmentation. The orientation, resolution and scaling augmentation performed in a preprocessing manner. Other augmentation methods not listed such as Gaussian noise, rotation, mirroring were performed on-the-fly with probability during training.

Aspects	Before Augmentation	After Augmentation
Orientation	RAS	RAS, LAS
Resolution	$3.57 \times 3.57 \times 5.2 \text{ mm}^3$	$3.57 \times 3.57 \times 5.2 \text{ mm}^3$, $4.46 \times 4.46 \times 5.2 \text{ mm}^3$
Scaling	$\times 1$	$\times 1, \times 2, \times 3$

3.2. Model Performance

We tested the models' performance first on images with identical imaging parameters (T1-w images) to the training data set for assessing the segmentation ability against anatomical variations between subjects. For the dice score, the 2D-mag-net, 3D-mag-net, and 3D-mag-phase-net was 0.87 ± 0.04 , 0.91 ± 0.02 , and 0.94 ± 0.04 , respectively. For 95% HD, 2Dmag-net, 3D-mag-net, and 3D-mag-phase-net was 11.20 ± 5.90 , 7.78 ± 4.62 , and $6.20 \pm 3.61 \text{ mm}$, respectively. For Jaccard index, 2Dmag-net, 3D-mag-net, and 3D-mag-phase-net was 0.76 ± 0.06 , 0.83 ± 0.04 , and 0.87 ± 0.07 , respectively. As illustrated in Figure 6, under the null hypothesis that the predictions made by these models share the same distribution, a significant enhancement in the dice score is presented in the dual channel model. This indicated that our model (3D-mag-phase-net) outperformed other models with the help of additional phase information.

Furthermore, as we applied the cross-validation during the training, we evaluated the consistency across different folds in our cross-validation approach. We generated the segmentation employing the model from each fold. As illustrated in Figure 7, the performance of various models across each fold was examined. The bar plot demonstrated that, in most of the folds, our 3D-mag-phase model consistently exhibited superior performance. This finding underscores the robustness and reliability of the proposed method.

In Figure 8, we presented the training progress of different models. Notably, the 2D-mag-net converged fastest due to a larger batch size. The larger batch size also leads to a more stable training process. However, the performance of the 2D-mag-net, as previously demonstrated, suggested a potential overfitting of the training data. The batch size of the

two 3D models is the same. Our proposed model converged faster than the 3D-mag-net and maintained a more stable training process.

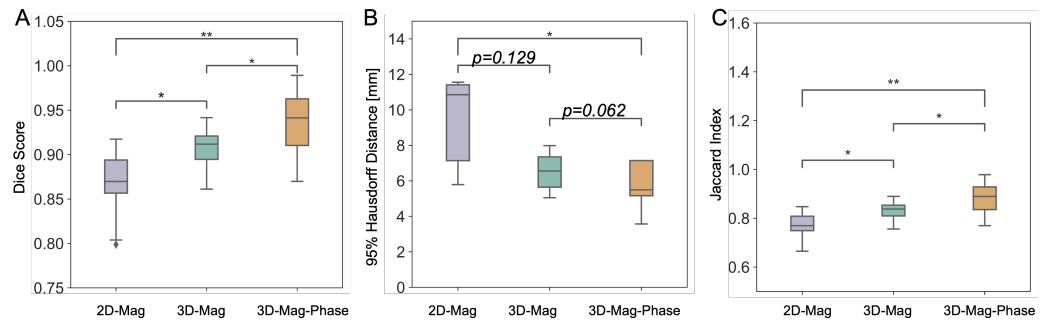


Figure 6. Comparative analysis of three different models on the T1-w dataset. The 3D-Mag-Phase net showed a significantly higher dice score than the other parameters (A) and a significantly higher Jaccard index than others (C). The 3D-Mag-Phase net is the only model that showed significant improvement in 95% HD compared to the conventional 2D model (B). (* indicates p -value < 0.05 and ** indicates p -value < 0.01.)

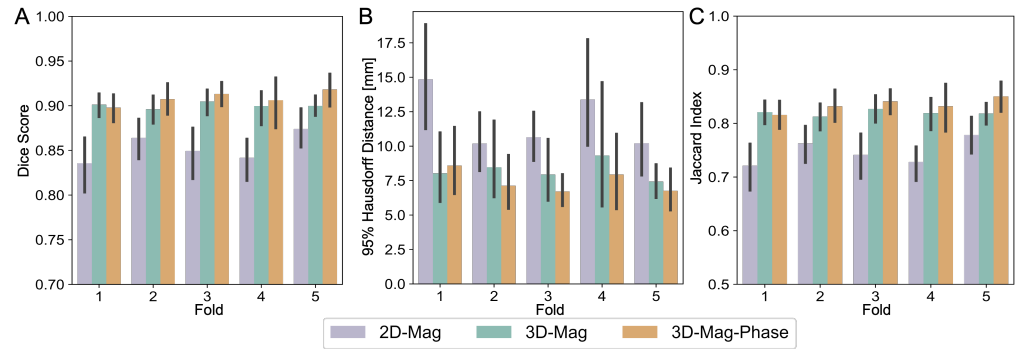


Figure 7. Comparison of different models on T1-w data among 5 folds in cross-validation. (A) shows the average dice score and standard deviation on T1-w test data, (B) shows the 95%HD, and (C) shows the Jaccard index.

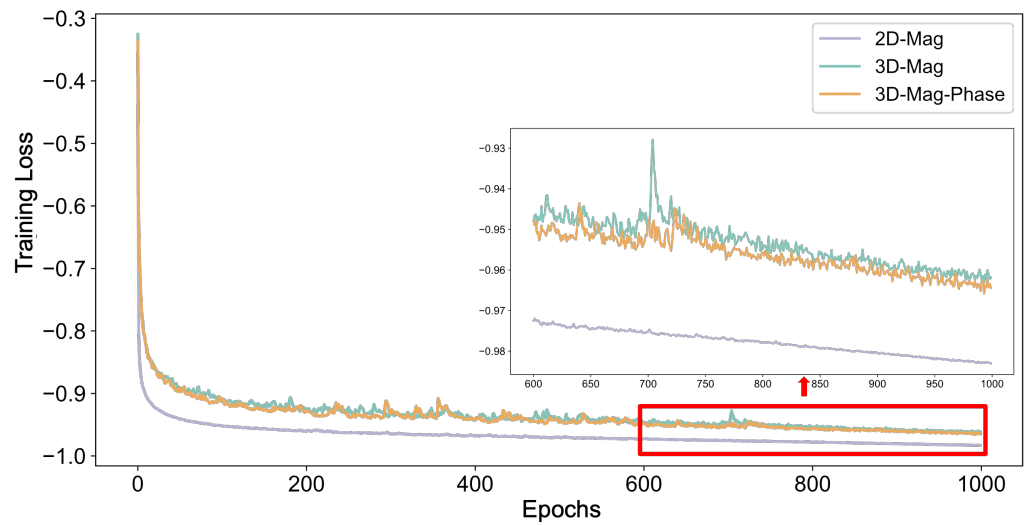


Figure 8. The training procedure of different models. This figure showed the training process of the 5th fold as an example.

3.3. Generalizability Analysis

To assess the generalizability of our models, we conducted the following ablation studies.

3.3.1. SNR Variations

We investigated the robustness of different models to the changes in the Signal-to-Noise Ratio (SNR). We simulated the potential noise during the data acquisition by adding Gaussian noise. The noise was sampled from a normal distribution with a mean of 0 and standard deviations ranging from 0.01 to 0.05. As shown in Figure 9, the 3D models demonstrated a significantly higher degree of robustness in the presence of noise, whereas the performance of the 2D-mag-net exhibited a significant decline. Notably, our 3D-mag-phase-net consistently outperformed other models, and the standard deviation of the performance was consistently smaller than the 3D-mag-net throughout all SNR levels.

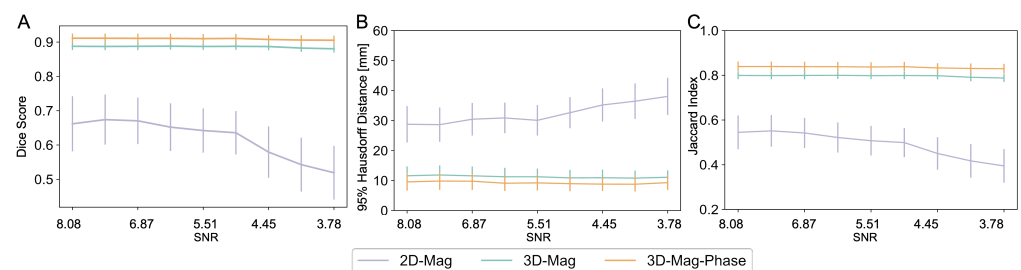


Figure 9. Model performance against field map SNR changes. In the 2D model, the segmentation performance was significantly reduced, corresponding to the increased noise level. On the contrary, the 3D models consistently performed with the SNR variation. In addition, the proposed dual-channel model (3D-mag-phase) demonstrates consistently improved segmentation results compared to the single-channel model (3D-mag). The (A–C) showed the corresponding segmentation performance in dice score, 95%HD and jaccard index.

3.3.2. Imaging Protocol Variations

In addition to the training dataset, we conducted experiments on field maps acquired with a different imaging protocol to test the model's performance against the variation of image contrast and image resolution. Proton density-weighted (PD-w) high-resolution images were acquired from 10 healthy subjects. Furthermore, since field maps can be acquired under different breath-holding states in different clinical practices, we further collected images under end-inspiration and end-expiration to investigate the influences of the respiratory position on the model performance. We predicted the contours using the trained models without further fine-tuning or training on the new dataset. The models' performance against the imaging protocol variation is compared in Figure 10. The proposed 3D-mag-phase-net showed significantly improved dice scores compared to the magnitude-only models under the new dataset with different image contrast and resolution. The summary results was shown in Table 2. The 95% HD shows a similar trend. (In end-expiratory, 2D-mag-net's average 95%HD and SD was 6.610 ± 0.921 , 3D-mag-net was 6.019 ± 0.792 , and 3D-mag-Phase-net was 5.775 ± 1.486 . In end-inspiratory, 2D-mag-net's average 95%HD and SD was 7.824 ± 2.421 , 3D-mag-net was 6.396 ± 1.080 and 3D-mag-phase-net was 5.922 ± 1.837 .) Furthermore, consistent results are shown in the end-expiratory or end-inspiratory cycle, demonstrating the models' robustness against respiratory motion.

Table 2. The mean and standard deviation of segmentation performance on PDw data. The dice score, 95%HD and Jaccard index were reported under different respiratory cycle and for all the comparison, all $p < 0.05$.

Motion States	Model	Dice Score ↑	95%HD [mm] ↓	Jaccard Index ↓
End-Expiration	2D-Mag	0.85 ± 0.04	6.61 ± 0.92	0.80 ± 0.06
	3D-Mag	0.89 ± 0.02	6.02 ± 0.80	0.86 ± 0.03
	3D-Mag-Phase	0.93 ± 0.02	5.78 ± 1.49	0.94 ± 0.03
End-Inspiration	2D-Mag	0.83 ± 0.11	7.82 ± 2.42	0.78 ± 0.13
	3D-Mag	0.89 ± 0.02	6.40 ± 1.08	0.86 ± 0.04
	3D-Mag-Phase	0.93 ± 0.03	5.92 ± 1.84	0.93 ± 0.05

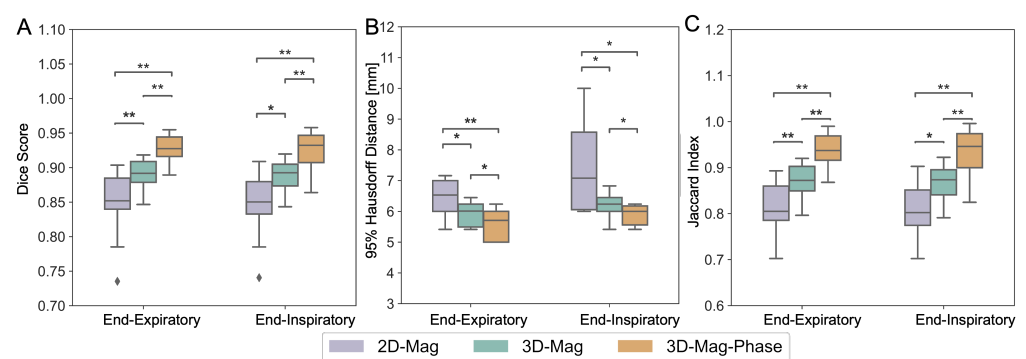


Figure 10. Comparison of different models on the proton-density weighted dataset. The dice score using different models in end-expiratory and end-inspiratory was shown in (A), 95%HD in (B) and jaccard index in (C). Our proposed 3D-mag-phase-net consistently and significantly outperformed others (* indicates p -value < 0.05 and ** indicates p -value < 0.01).

3.3.3. Comparisons between Model Architectures

To evaluate the ability of different model structures to utilize the phase information, we extended our work to incorporate the dual-channel into other existing deep learning-based segmentation models. Specifically, we implemented the naïve UNet and the Swin UNETR models based on the MONAI [52] framework. The training strategy remained consistent with our prior description, and we evaluated the segmentation results on all the same testing datasets. In addition, we finetuned a state-of-the-art transformer-based single-channel model (SAM-Med3D [53]) based on our training dataset to test its performance on the task of field map segmentation. The models' performance is shown in Table 3. In all U-net-based models (U-Net, Swin UNETR, and ours (3D-mag-phase-net)), the magnitude-phase model showed improved dice scores compared to the magnitude-only models. This indicates that the additional information from phase maps can be extracted regardless of the model structures. In addition, the proposed 3D-mag-phase-net outperformed all other models (all $p < 0.05$) and showed an accurate segmentation result across the different imaging protocols. It is worth noting that, in recent studies, transformer-based models have been recognized to outperform convolution-based models when trained with large heterogeneous datasets [54]. However, due to the limited training data size, the SAM-Med3D model and Swin UNETR did not show the desired performance on this task in our study. In contrast, despite the limited training data, our model reliably performs and shows significant advantages in model convergence.

Table 3. The mean and standard deviation of dice scores using different types of models. The additional phase information can improve the model performance regardless of the model architecture, which highlights the importance of phase information. Additionally, our proposed model reported the highest dice score and reported a smaller SD (* indicates p -value < 0.05).

Models	Magnitude Only	Magnitude-Phase
SAM-Med3D	0.5814 (0.051)	-
U-Net	0.7988 (0.064)	0.8252 (0.049)
Swin UNETR	0.8571 (0.045)	0.8623 (0.044)
Ours (3D-mag-phase-net) *	0.9065 (0.023)	0.9379 (0.038)

3.3.4. Cardiac B_0 Shimming Experiments and Performance Comparison

The B_0 shimming ability of the proposed method is evaluated and compared to the standard manual-selected box shim volume in Figure 11. B_0 shimming was derived using the 2nd-order spherical harmonic shim coils that are equipped with state-of-the-art 3T clinical scanners. Representative images of the cardiac region from axial and coronal views demonstrated the shimming performance was improved with the proposed autonomous shimming pipeline and is shown in Figure 11A. To validate the visualized improvement, the shimming performance was evaluated quantitatively using standard deviation (SD) and interquartile range (IQR) of the B_0 field in the heart and compared to the standard manual box shim in Figure 11B. Significantly more homogeneous B_0 from the proposed model is presented in SD and IQR and reflect a more reliable performance of B_0 shimming without human interaction. Cardiac B_0 SD ($SD(B_0\text{Shim})/SD(B_0)$) decreased $15 \pm 11\%$ using our proposed autonomous shimming, while $6 \pm 12\%$ using standard manual shimming. For IQR, $IQR(B_0\text{Shim})/IQR(B_0)$ decreased $21 \pm 12\%$ using our autonomous ROI, while $14 \pm 18\%$ using standard manual box volume (all p -value < 0.05).

3.3.5. Ablation Study

To better understand the contribution of each part, except comparing the 2D-mag, 3D-mag and 3D-mag-phase net, we further investigated the following components for ablation: instance normalization and augmentation. For each setting we re-trained the model and evaluate the performance using cross-fold validation. The performance on validation set was quantified on each component. In our implementation, we utilized the instance normalization instead of batch normalization after each convolution operation. In our experiments, the change from batch normalization to instance normalization will decrease the average validation dice score from 0.93 to 0.91 but drastically accelerate the process. The small batch size (batch size of 2 in our implementation) limited the batch normalization's ability to speed up and stabilize training. While the instance normalization can deal with the noiser mean and variance when the batch size is very small. In our proposed model, we finally chose the instance normalization, as the segmentation accuracy was more important in shimming application. Additionally, we experimented the importance of data augmentation, we found that without scale augmentation, the dice score decreased averagely 2.15% on the validation set, which is consistent with our observation in Section 2.4 that various patient size in test set will affect the model performance.

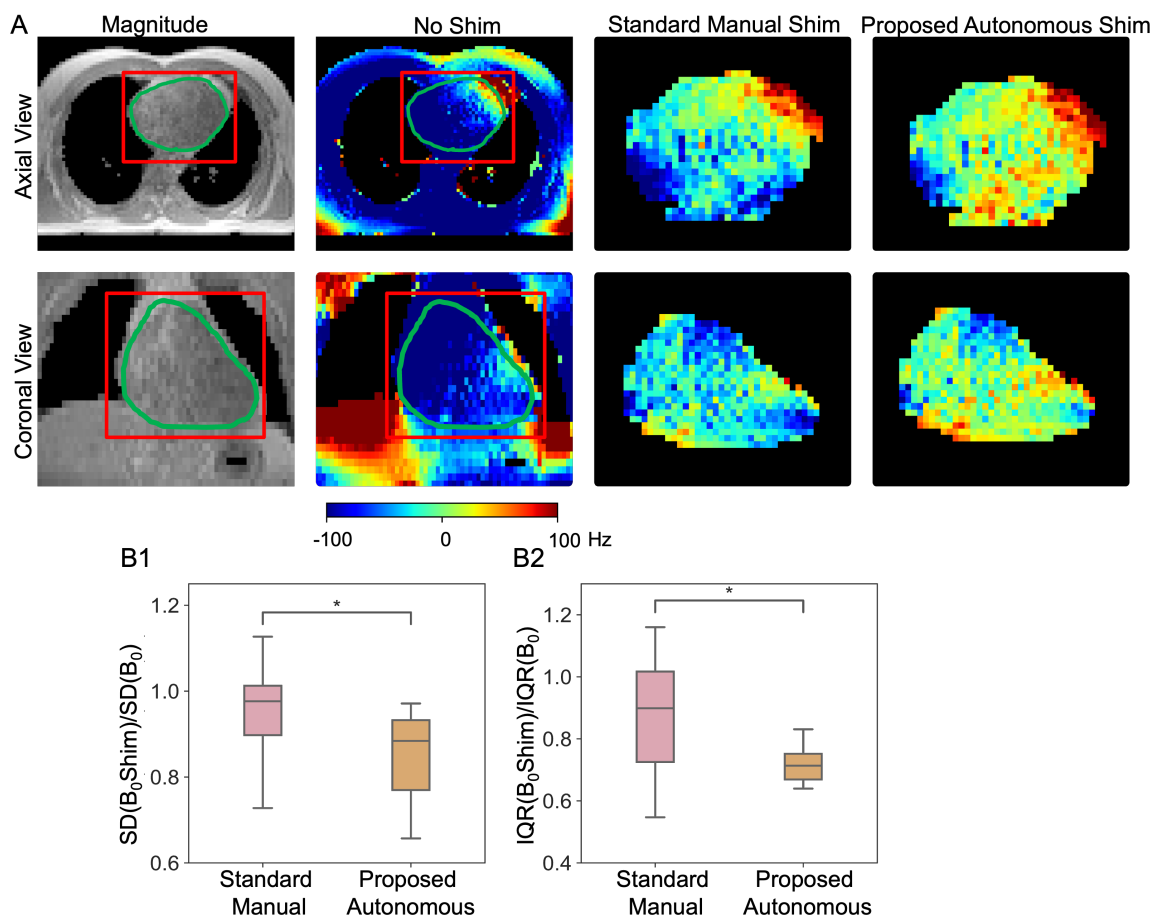


Figure 11. B_0 shimming comparisons between the manual box shim and autonomous contour shim. The representative figure (A) demonstrated the magnitude map, original phase map (No shim), and shimmed heart ROI using standard manual box shim and shimmed heart ROI using the proposed autonomous shim in axial and coronal view. The red square indicates the manual shimming box during standard manual shim, and the green contour indicates the auto-generated contour for our proposed method. The statistical results demonstrating the SD (B1) and IQR (B2) ratio after and before shimming, showed that our proposed method improved the shimming process significantly (* indicates p -value < 0.05).

4. Discussion

In this paper, we explored the integration of magnitude and phase information to enhance the accuracy of 3D segmentation models for CMR field maps and its ability to improve B_0 shimming compared to the standard manual shimming pipeline. We evaluated the model and demonstrated the robustness and generalizability of the proposed dual-channel model using CMR field maps acquired with different contrast weighting and imaging parameters. The proposed 3D-mag-phase-net, built based on the nnU-Net structure, successfully harnessed the complementary information from phase maps, especially benefiting the segmentation accuracy in regions with impaired tissue contrast in the magnitude images. It demonstrates reliable performance in real-world data through data variations commonly presented in the clinical setting.

Previous research has indicated that a conventional U-net model has the potential to predict cardiac contour and aid in cardiac B_0 shimming [55]. However, this approach has only been tested on 1.5T scanners, and its efficacy in higher field scanners remains untested. Moreover, the previous study was performed using fixed imaging protocols and did not explore the possibility of performance variation between field map acquisition parameters, which means that the model's generalizability is unclear.

In this study, we developed a dual-channel model combined with an advanced nnU-Net architecture to accommodate the potential variation of field map acquisition. We tested the model at 3.0T, where the B_0 off-resonance is stronger and affects its daily application in the clinical setting. The integration of magnitude and phase images allows for structural and functional insights, making it a powerful tool for the segmentation task in MR images. Notably, in segmenting the cardiac field map, a major challenge for conventional methods is to separate the connected heart and liver at the apical portion of the left ventricle. This is particularly important for B_0 shimming as strong off-resonance artifacts are commonly presented in this region due to the unfavorable heart-lung anatomy. Because the conventional magnitude images (both T1 or proton-density weighted) exhibit similar signal intensity between the heart and the liver. The magnitude-only segmentation methods often fail in this region and significantly affect the shim results. In contrast, phase maps showed strong phase differences between the organs, reflecting their local frequency changes influenced by their respective location, shape, and orientation in relation to both the air-filled space and the B_0 direction, as shown in Figure 12. The phase map provides clear delineation at the heart-liver interface, which can facilitate reliable segmentation for the heart.

In addition to the organ boundaries, tissue composition is critical for field map segmentation. Particularly, the epicardial fat can cause contrast variations in the T1w image and introduce contrast changes between imaging parameters, in Figure 12. This can compromise the generalizability of the magnitude-only models. The phase maps provide a consistent frequency profile of the fat signal, which can assist in identifying the fat tissue and keep the consistency of cardiac field map segmentation.

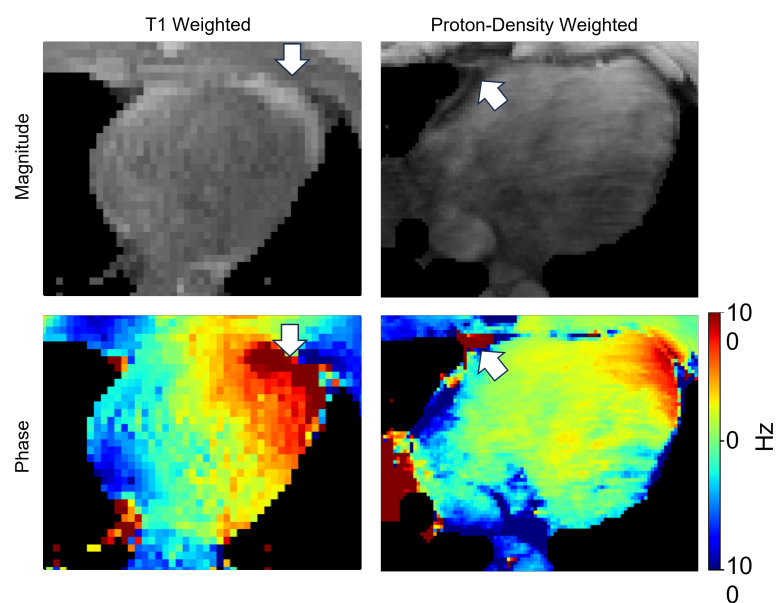


Figure 12. The contrast difference of fat region in different MRI parameter images. Magnitude images' contrast changes on fat region when using different MRI parameters, as the white arrows show. However, the fat region is differentiable according to the phase maps.

In our study, the generalizability of the models between imaging parameters is tested under different MRI acquisition protocols. We showed the ability of the proposed model to maintain accurate segmentation capability under different SNR, resolution, and MRI imaging contrasts. In the conventional single-channel techniques, the change in imaging contrast is usually a domain-transferring task in segmentation models. Fortunately, the phase maps of MR field maps with multi-echoes are quantitative and resilient to imaging parameter changes. This provides a consistent domain for segmenting the target organs in MR field maps.

One potential trade-off of this integration is the dependence on the quality of the phase maps. Artifacts caused by phase wrapping and motion-induced phase errors can degrade the performance of the segmentation. These challenges in the phase domain will propagate the errors and even mislead the segmentation results. Mitigating such artifacts is crucial, and future work may also include developing more advanced algorithms for phase unwrapping and motion correction to fully leverage the integrated information [28,56,57]. Another consideration is the additional computational demand of the dual-channel model. The processing of additional channels inherently requires more computation resources, such as larger RAM to feed the data, potentially limiting its clinical applicability. The current model is compatible with a single workstation equipped with an NVIDIA GeForce RTX 4090 GPU with 24 GB RAM. Using a relatively small batch size and median patch size as mentioned in the methods, we can successfully deploy the model in a state-of-the-art scanner's host computer. To enable broader applications, more investigation into computational efficiency without drastically compromising performance can be done in the future. This can help the development of more streamlined clinical applications in scanners with less computation powers.

The U-net structure based on convolutional blocks has shown its success in various medical image-related tasks [58], with the help of its U-structure, to capture the semantic information within the image. That being said, the transformed-based backbone utilizing the attention mechanism is catching popularity in the computer vision field [59]. The attention mechanism allows the transformer model capture the long-range and global context information while convolutional layers usually local semantic information [60]. Although transformer-based models showed promising results [32,61,62] in medical image analysis, due to the large model size, the networks are more difficult to train and require larger training data [63]. A common practice for the transformed-based network is to start from the pre-trained model on large-scale datasets [54], which allows for fine-tuning specific tasks with smaller data input. However, fine-tuning a large model is not trivial work [64]; it requires careful hyperparameter tuning, including learning rates, batch sizes, and normalization techniques, making the finetuning of the transformer models on the limited medical imaging datasets challenging. In this study, we found that our fine-tuning of the SAM-Med3D model [53] performed much worse than the proposed U-net-based model in the test dataset. This might be due to the unique imaging contrast in the B_0 field maps and the limited size of the training data, which does not provide enough diversity and complexity for the model.

The application of our proposed model in cardiac B_0 shimming is shown in a high-field in-vivo experiment by creating autonomous contours of the field maps. Our data demonstrated that the model's ability to produce accurate, motion-robust contours on cardiac field maps quickly could significantly enhance the B_0 homogeneity with the standard clinical shimming hardware. Furthermore, it is worth noting that the shim volume accuracy is particularly important for the fast-developing multi-coil shimming hardware and combined shim-RF coils [24,25,65–67]. Because of the increased flexibility of the shimming fields, an erroneous shimming ROI can lead to strong overfitting to the false B_0 field and corrupt the overall robustness of the procedure. The developed autonomous shimming pipeline can be a crucial step for implementing high-order shimming hardware and its clinical applications.

5. Conclusions

Accurate B_0 shimming is crucial for successful high-field cardiac MRI studies. The developed dual-channel model incorporating phase and magnitude information has achieved high segmentation accuracy in cardiac B_0 field maps that are insensitive to imaging acquisition protocol changes. The integration of the developed autonomous pipeline into clinical scanners could serve as the foundation for reliable high-field CMR imaging and widen its clinical adoption. Furthermore, future works to combine the developed method with

high-order shimming hardware can further boost the B_0 homogeneity and enable advanced imaging contrast for a wide range of cardiovascular diseases.

Author Contributions: Conceptualization, X.L. and H.-J.Y.; methodology, X.L., Y.H., X.B., F.H. and C.G.; software, X.L.; formal analysis, X.L., A.M., L.-T.H. and Y.H.; data curation, Y.H., A.M., C.-C.Y., G.Y., L.-T.H. and E.T.; Drafting, review and editing, Y.H., A.M., C.-C.Y., G.Y., L.-T.H., G.Y., M.-C.K., H.-J.Y. and H.H.; supervision, M.-C.K. and H.-J.Y.; project administration, H.-J.Y. and H.H.; funding acquisition, H.-J.Y. and H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Heart, Lung, and Blood Institute (NHLBI) and the National Institutes of Health (NIH) grants 1R01 HL165211, R01 HL156818 and R01 NS121544.

Institutional Review Board Statement: This study was authorized by the Institutional Review Board (IRB), No.23469, at Cedars Sinai Medical Center, Los Angeles. All participants in the study provided their consent after receiving complete information about the research. The anonymity and confidentiality of all participants were strictly maintained throughout the research process. The IRB reviewed and approved the study protocol and procedures to protect the participants' rights and welfare.

Informed Consent Statement: All participants gave informed consent before participating in the study. They were informed of the purpose, procedures, potential hazards and benefits, confidentiality measures, and their rights as research subjects. They were free to disengage from the study at any time. Each participant provided written consent, and anonymity and confidentiality were maintained throughout the study. The study protocol complied with ethical standards and was approved by the relevant Institutional Review Board.

Data Availability Statement: Data available on request due to restrictions of privacy. The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the Research Imaging Core (RIC) at Cedars-Sinai for their valuable support.

Conflicts of Interest: The authors, H.J.Y. and H.H., own equity in Lucidity Medical LLC, and X.B. F.H., G.C. are employees of Siemens Healthineers. These financial and professional affiliations are disclosed to ensure transparency. The authors assert that these interests do not influence the scientific content of their work.

Abbreviations

The following abbreviations are used in this manuscript:

MRI	Magnetic Resonance Imaging
CMR	Cardiac Magnetic Resonance Imaging
LGE	Late gadolinium enhancement
CEST	Chemical Exchange Saturation Transfer
SNR	Signal-to-Noise Ratio
SSFP	Steady-state Free Precession
EPI	Echo-Planar Imaging
CNN	Convolutional Neural Network
FOV	Field of View
ReLU	Rectified Linear Unit
HD	Hausdorff Distance
PDw	Proton density-weighted

References

1. Rajiah, P.S.; François, C.J.; Leiner, T. Cardiac MRI: State of the Art. *Radiology* **2023**, *307*, e223008. [[CrossRef](#)]
2. Kali, A.; Tang, R.L.Q.; Kumar, A.; Min, J.K.; Dharmakumar, R. Detection of Acute Reperfusion Myocardial Hemorrhage with Cardiac MR Imaging: T2 versus T2*. *Radiology* **2013**, *269*, 387–395. [[CrossRef](#)]
3. Jenista, E.R.; Wendell, D.C.; Azevedo, C.F.; Klem, I.; Judd, R.M.; Kim, R.J.; Kim, H.W. Revisiting how we perform late gadolinium enhancement CMR: Insights gleaned over 25 years of clinical practice. *J. Cardiovasc. Magn. Reson.* **2023**, *25*, 18. [[CrossRef](#)]

4. Yang, H.J.; Dey, D.; Sykes, J.; Butler, J.; Biernaski, H.; Kovacs, M.; Bi, X.; Sharif, B.; Cokic, I.; Tang, R.; et al. Heart Rate-Independent 3D Myocardial Blood Oxygen Level-Dependent MRI at 3.0 T with Simultaneous ¹³N-Ammonia PET Validation. *Radiology* **2020**, *295*, 82–93. [[CrossRef](#)]
5. Kali, A.; Choi, E.Y.; Sharif, B.; Kim, Y.J.; Bi, X.; Spottiswoode, B.; Cokic, I.; Yang, H.J.; Tighiouart, M.; Conte, A.H.; et al. Native T1 Mapping by 3-T CMR Imaging for Characterization of Chronic Myocardial Infarctions. *JACC Cardiovasc. Imaging* **2015**, *8*, 1019–1030. [[CrossRef](#)]
6. Carr, J.C.; Simonetti, O.; Bundy, J.; Li, D.; Pereles, S.; Finn, J.P. Cine MR angiography of the heart with segmented true fast imaging with steady-state precession. *Radiology* **2001**, *219*, 828–834. [[CrossRef](#)]
7. Haaf, P.; Garg, P.; Messroghli, D.R.; Broadbent, D.A.; Greenwood, J.P.; Plein, S. Cardiac T1 Mapping and Extracellular Volume (ECV) in clinical practice: A comprehensive review. *J. Cardiovasc. Magn. Reson.* **2016**, *18*, 89. [[CrossRef](#)]
8. Nojiri, A.; Hongo, K.; Kawai, M.; Komukai, K.; Sakuma, T.; Taniguchi, I.; Yoshimura, M. Scoring of late gadolinium enhancement in cardiac magnetic resonance imaging can predict cardiac events in patients with hypertrophic cardiomyopathy. *J. Cardiol.* **2011**, *58*, 253–260. [[CrossRef](#)]
9. Eitel, I.; Friedrich, M.G. T2-weighted cardiovascular magnetic resonance in acute cardiac disease. *J. Cardiovasc. Magn. Reson.* **2011**, *13*, 13. [[CrossRef](#)]
10. Triadyaksa, P.; Oudkerk, M.; Sijens, P.E. Cardiac T2* mapping: Techniques and clinical applications. *J. Magn. Reson. Imaging* **2020**, *52*, 1340–1351. [[CrossRef](#)]
11. Oshinski, J.N.; Delfino, J.G.; Sharma, P.; Gharib, A.M.; Pettigrew, R.I. Cardiovascular magnetic resonance at 3.0 T: Current state of the art. *J. Cardiovasc. Magn. Reson.* **2010**, *12*, 1–13. [[CrossRef](#)]
12. Bottomley, P.A.; Foster, T.H.; Argersinger, R.E.; Pfeifer, L.M. A review of normal tissue hydrogen NMR relaxation times and relaxation mechanisms from 1–100 MHz: Dependence on tissue type, NMR frequency, temperature, species, excision, and age. *Med. Phys.* **1984**, *11*, 425–448. [[CrossRef](#)]
13. Edelstein, W.; Glover, G.; Hardy, C.; Redington, R. The intrinsic signal-to-noise ratio in NMR imaging. *Magn. Reson. Med.* **1986**, *3*, 604–618. [[CrossRef](#)]
14. Sharma, P.; Socolow, J.; Patel, S.; Pettigrew, R.I.; Oshinski, J.N. Effect of Gd-DTPA-BMA on blood and myocardial T1 at 1.5 T and 3T in humans. *J. Magn. Reson. Imaging Off. J. Int. Soc. Magn. Reson. Med.* **2006**, *23*, 323–330. [[CrossRef](#)]
15. Zhou, Z.; Nguyen, C.; Chen, Y.; Shaw, J.L.; Deng, Z.; Xie, Y.; Dawkins, J.; Marbán, E.; Li, D. Optimized CEST cardiovascular magnetic resonance for assessment of metabolic activity in the heart. *J. Cardiovasc. Magn. Reson.* **2017**, *19*, 1–7. [[CrossRef](#)]
16. Wen, Y.; Nguyen, T.D.; Liu, Z.; Spincemaille, P.; Zhou, D.; Dimov, A.; Kee, Y.; Deh, K.; Kim, J.; Weinsaft, J.W.; et al. Cardiac quantitative susceptibility mapping (QSM) for heart chamber oxygenation. *Magn. Reson. Med.* **2018**, *79*, 1545–1552. [[CrossRef](#)]
17. Noeske, R.; Seifert, F.; Rhein, K.H.; Rinneberg, H. Human cardiac imaging at 3 T using phased array coils. *Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med.* **2000**, *44*, 978–982.
18. Reeder, S.B.; Faranesh, A.Z.; Boxerman, J.L.; McVeigh, E.R. In vivo measurement of T* 2 and field inhomogeneity maps in the human heart at 1.5 T. *Magn. Reson. Med.* **1998**, *39*, 988–998. [[CrossRef](#)]
19. Hock, M.; Terekhov, M.; Stefanescu, M.R.; Lohr, D.; Herz, S.; Reiter, T.; Ankenbrand, M.; Kosmala, A.; Gassenmaier, T.; Juchem, C.; et al. B0 shimming of the human heart at 7T. *Magn. Reson. Med.* **2021**, *85*, 182–196. [[CrossRef](#)]
20. Deux, J.F.; Maatouk, M.; Vignaud, A.; Luciani, A.; Lenczner, G.; Mayer, J.; Lim, P.; Dubois-Randé, J.L.; Kobeiter, H.; Rahmouni, A. Diffusion-weighted echo planar imaging in patients with recent myocardial infarction. *Eur. Radiol.* **2011**, *21*, 46–53. [[CrossRef](#)]
21. Nezafat, M.; Henningsson, M.; Ripley, D.P.; Dedieu, N.; Greil, G.; Greenwood, J.P.; Börnert, P.; Plein, S.; Botnar, R.M. Coronary MR angiography at 3T: Fat suppression versus water-fat separation. *Magn. Reson. Mater. Phys. Biol. Med.* **2016**, *29*, 733–738. [[CrossRef](#)]
22. Sengupta, S.; Welch, E.B.; Zhao, Y.; Foxall, D.; Starewicz, P.; Anderson, A.W.; Gore, J.C.; Avison, M.J. Dynamic B0 shimming at 7 T. *Magn. Reson. Imaging* **2011**, *29*, 483–496. [[CrossRef](#)]
23. Schwerter, M.; Hetherington, H.; Moon, C.H.; Pan, J.; Felder, J.; Tellmann, L.; Shah, N.J. Interslice current change constrained B0 shim optimization for accurate high-order dynamic shim updating with strongly reduced eddy currents. *Magn. Reson. Med.* **2019**, *82*, 263–275. [[CrossRef](#)]
24. Han, H.; Song, A.W.; Truong, T. Integrated parallel reception, excitation, and shimming (iPRES). *Magn. Reson. Med.* **2013**, *70*, 241–247. [[CrossRef](#)]
25. Hsin-Jung, Y.; John, S.; Linda, A.; Waishing, L.; Meng, L.; Yuheng, H.; Yoosefian, G.; Skyler, S.; Richard, H.; Yujie, S.; et al. Whole Heart High-Order B0 Shimming at 3T Using a UNified Coil (UNIC) for RF receive and shimming. In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), Sydney, Australia, 8–14 August 2020; Volume 2183.
26. Lee, J.; Lustig, M.; Kim, D.; Pauly, J.M. Improved shim method based on the minimization of the maximum off-resonance frequency for balanced steady-state free precession (bSSFP). *Magn. Reson. Med.* **2009**, *61*, 1500–1506. [[CrossRef](#)]
27. Li, X.; Huang, Y.; Guan, X.; Zhang, X.; Yoosefian, G.; Bi, X.; Han, F.; Lee, H.; Christodoulou, A.; Li, D., et al. Correcting motion induced B0 shim failure at 3T CMR using a deep learning-enabled 3D motion-resolved B0 shimming. In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), Toronto, ON, Canada, 3–8 June 2023; Volume 4977.
28. Huang, Y.; Guan, X.; Zhang, X.; Tang, L.; Ghazal, Y.; Bi, X.; Han, F.; Lee, H.; Han, H.; Christodoulou, A.; et al. The Effect of Respiratory and Cardiac Motion States on B0 Shimming at 3T. In Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), Toronto, ON, Canada, 3–8 June 2023; Volume 1150.

29. Osuna-Enciso, V.; Cuevas, E.; Sossa, H. A comparison of nature inspired algorithms for multi-threshold image segmentation. *Expert Syst. Appl.* **2013**, *40*, 1213–1219. [[CrossRef](#)]
30. Pohle, R.; Toennies, K.D. Segmentation of medical images using adaptive region growing. In Proceedings of the Medical Imaging 2001: Image Processing, San Diego, CA, USA, 17–22 February 2001; Volume 4322, pp. 1337–1346.
31. Freedman, D.; Zhang, T. Interactive graph cut based segmentation with shape priors. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005; Volume 1, pp. 755–762.
32. Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H.R.; Xu, D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Springer International Publishing: Cham, Switzerland, 2022; pp. 272–284. [[CrossRef](#)]
33. Isensee, F.; Jaeger, P.F.; Kohl, S.A.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [[CrossRef](#)]
34. Tang, Y.; Yang, D.; Li, W.; Roth, H.R.; Landman, B.; Xu, D.; Nath, V.; Hatamizadeh, A. Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022. [[CrossRef](#)]
35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015, Proceedings, Part III 18*; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. [[CrossRef](#)]
36. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer International Publishing: Cham, Switzerland, 2018; pp. 3–11. [[CrossRef](#)]
37. Li, X.; Huang, Y.; Yoosefian, G.; Hui, H.; Yang, H.-J. Autonomous cardiac field map segmentation for B0 shimming pipeline using a dual-modality Deep-learning model. In Proceedings of the 26th Annual Scientific Sessions of the Society for Cardiovascular Magnetic Resonance (SCMR), San Diego, CA, USA, 25–28 January 2023.
38. Wang, H.; Cao, P.; Wang, J.; Zaiane, O.R. UCTransNet: Rethinking the Skip Connections in U-Net from a Channel-Wise Perspective with Transformer. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 22 February–1 March 2022; Volume 36, pp. 2441–2449. [[CrossRef](#)]
39. Çiçek; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016: Proceedings of the 19th International Conference, Athens, Greece, 17–21 October 2016, Proceedings, Part II 19*; Springer International Publishing: Cham, Switzerland, 2016; pp. 424–432. [[CrossRef](#)]
40. Yu, L.; Cheng, J.Z.; Dou, Q.; Yang, X.; Chen, H.; Qin, J.; Heng, P.A. Automatic 3D Cardiovascular MR Segmentation with Densely-Connected Volumetric ConvNets. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2017: Proceedings of the 20th International Conference, Quebec City, QC, Canada, 11–13 September 2017, Proceedings, Part II 20*; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2017; pp. 287–295. [[CrossRef](#)]
41. Li, P.; Wu, W.; Liu, L.; Michael Serry, F.; Wang, J.; Han, H. Automatic brain tumor segmentation from Multiparametric MRI based on cascaded 3D U-Net and 3D U-Net++. *Biomed. Signal Process. Control* **2022**, *78*, 103979. [[CrossRef](#)]
42. Chen, T.; Wang, C.; Shan, H. BerDiff: Conditional Bernoulli Diffusion Model for Medical Image Segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI, Vancouver, BC, Canada, 8–12 October 2023; Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R., Eds.; Springer: Cham, Switzerland, 2023; pp. 491–501.
43. Fan, C.; Su, Q.; Xiao, Z.; Su, H.; Hou, A.; Luan, B. ViT-FRD: A Vision Transformer Model for Cardiac MRI Image Segmentation Based on Feature Recombination Distillation. *IEEE Access* **2023**, *11*, 129763–129772. [[CrossRef](#)]
44. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man, Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
45. Schulz-Menger, J.; Bluemke, D.A.; Bremerich, J.; Flamm, S.D.; Fogel, M.A.; Friedrich, M.G.; Kim, R.J.; von Knobelsdorff-Brenkenhoff, F.; Kramer, C.M.; Pennell, D.J.; et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance—2020 update. *J. Cardiovasc. Magn. Reson.* **2020**, *22*, 19. [[CrossRef](#)]
46. Kandel, I.; Castelli, M. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express* **2020**, *6*, 312–315. [[CrossRef](#)]
47. Pérez-García, F.; Sparks, R.; Ourselin, S. TorchIO: A Python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Programs Biomed.* **2021**, *208*, 106236. [[CrossRef](#)]
48. Touvron, H.; Vedaldi, A.; Douze, M.; Jegou, H. Fixing the Train-Test Resolution Discrepancy. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–14 December 2019; Volume 32.
49. Mohan, G.; Subashini, M.M. Chapter 4—Medical Imaging With Intelligent Systems: A Review. In *Deep Learning and Parallel Computing Environment for Bioengineering Systems*; Sangaiah, A.K., Ed.; Academic Press: Cambridge, MA, USA, 2019; pp. 53–73. [[CrossRef](#)]
50. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]

51. Seabold, S.; Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 57, pp. 10–25080.
52. Cardoso, M.J.; Li, W.; Brown, R.; Ma, N.; Kerfoot, E.; Wang, Y.; Murray, B.; Myronenko, A.; Zhao, C.; Yang, D.; et al. MONAI: An open-source framework for deep learning in healthcare. *arXiv* **2022**, arXiv:2211.02701.
53. Wang, H.; Guo, S.; Ye, J.; Deng, Z.; Cheng, J.; Li, T.; Chen, J.; Su, Y.; Huang, Z.; Shen, Y.; et al. SAM-Med3D. *arXiv* **2023**, arXiv:2310.15161v2.
54. Wang, Z.; Bai, Y.; Zhou, Y.; Xie, C. Can CNNs Be More Robust Than Transformers? In Proceedings of the Eleventh International Conference on Learning Representations, Kigali, Rwanda, 1–5 May 2023.
55. Edalati, M.; Zheng, Y.; Watkins, M.P.; Chen, J.; Liu, L.; Zhang, S.; Song, Y.; Soleymani, S.; Lenihan, D.J.; Lanza, G.M. Implementation and prospective clinical validation of AI-based planning and shimming techniques in cardiac MRI. *Med. Phys.* **2022**, *49*, 129–143. [[CrossRef](#)]
56. Chavez, S.; Qing-San, X.; An, L. Understanding phase maps in MRI: A new outline phase unwrapping method. *IEEE Trans. Med. Imaging* **2002**, *21*, 966–977. [[CrossRef](#)]
57. Zhou, H.; Cheng, C.; Peng, H.; Liang, D.; Liu, X.; Zheng, H.; Zou, C. The PHU-NET: A robust phase unwrapping method for MRI based on deep learning. *Magn. Reson. Med.* **2021**, *86*, 3321–3333. [[CrossRef](#)]
58. Siddique, N.; Paheding, S.; Elkin, C.P.; Devabhaktuni, V. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access* **2021**, *9*, 82031–82057. [[CrossRef](#)]
59. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929.
60. Xiao, H.; Li, L.; Liu, Q.; Zhu, X.; Zhang, Q. Transformers in medical image segmentation: A review. *Biomed. Signal Process. Control* **2023**, *84*, 104791. [[CrossRef](#)]
61. Lu, Y.; Fu, J.; Li, X.; Zhou, W.; Liu, S.; Zhang, X.; Wu, W.; Jia, C.; Liu, Y.; Chen, Z. Rtn: Reinforced transformer network for coronary ct angiography vessel-level image quality assessment. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Singapore, 18–22 September 2022; Springer: Berlin, Germany, 2022; pp. 644–653.
62. Tabatabaei, S.; Rezaee, K.; Zhu, M. Attention transformer mechanism and fusion-based deep learning architecture for MRI brain tumor classification system. *Biomed. Signal Process. Control* **2023**, *86*, 105119. [[CrossRef](#)]
63. Touvron, H.; Cord, M.; Douze, M.; Massa, F.; Sablayrolles, A.; Jegou, H. Training data-efficient image transformers & distillation through attention. In Proceedings of the International Conference on Machine Learning, Virtual, 18–24 July 2021; pp. 10347–10357.
64. Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 11966–11976. [[CrossRef](#)]
65. Chen, Q.; Luo, C.; Tie, C.; Cheng, C.; Zou, C.; Zhang, X.; Liu, X.; Zheng, H.; Li, Y. A 5-channel local B₀ shimming coil combined with a 3-channel RF receiver coil for rat brain imaging at 3 T. *Magn. Reson. Med.* **2023**, *89*, 477–486. [[CrossRef](#)]
66. Gao, Y.; Mareyam, A.; Sun, Y.; Witzel, T.; Arango, N.; Kuang, I.; White, J.; Roe, A.W.; Wald, L.; Stockmann, J.; et al. A 16-channel AC/DC array coil for anesthetized monkey whole-brain imaging at 7T. *NeuroImage* **2020**, *207*, 116396. [[CrossRef](#)]
67. Juchem, C.; Nixon, T.W.; McIntyre, S.; Boer, V.O.; Rothman, D.L.; De Graaf, R.A. Dynamic multi-coil shimming of the human brain at 7T. *J. Magn. Reson.* **2011**, *212*, 280–288. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.