Published in final edited form as:

NMR Biomed. 2021 December ; 34(12): e4606. doi:10.1002/nbm.4606.

Automated segmentation of bi-ventricular contours in tissue phase mapping using deep learning

Daming Shen[1,2,*], **Ashitha Pathrose**[1,*], **Roberto Sarnari**[1], **Allison Blake**[1], **Haben Berhane**[1,2], **Justin J Baraboo**[1,2], **James C Carr**[1], **Michael Markl**[1,2,+], **Daniel Kim**[1,2,+]

[1]Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, USA

[2]Biomedical Engineering, Northwestern University McCormick School of Engineering and Applied Science, Evanston, USA

Abstract

Tissue phase mapping (TPM) is an MRI technique for quantification of regional biventricular myocardial velocities. Despite the potential, its clinical use is limited due to the requisite labor-intensive manual segmentation of cardiac contours for all time frames. The purpose of this study was to develop a deep learning (DL) network for automated segmentation of TPM images, without significant loss in segmentation and myocardial velocity quantification accuracy compared with manual segmentation. We implemented a multi-channel 3D (2D + time) dense U-Net that trained on magnitude and phase images and combined cross-entropy, dice, and Hausdorff distance loss terms to improve the segmentation accuracy and suppress unnatural boundaries. The dense U-Net was trained and tested with 150 multi-slice, multi-phase TPM scans(114 scans for training, 36 for testing) from 99 heart transplant (HTx) patients(44 females, 1–4 scans/patient), where the magnitude and velocity-encoded (V_x, V_y, V_z) images were used as input and the corresponding manual segmentation masks were used as reference. The accuracy of DL segmentation was evaluated using quantitative metrics (dice scores, Hausdorff distance) and linear regression and Bland-Altman analyses on the resulting peak radial and longitudinal velocities (V_r and V_z). The mean segmentation time was ~2 hours per patient for manual and 1.9 ± 0.3 seconds for DL. Our network produced good accuracy (median dice = 0.85 for left ventricle [LV], 0.64 for right ventricle [RV], Hausdorff distance = 3.17 pixels) compared with manual segmentation. Peak V_r and V_z measured from manual and DL segmentations were strongly correlated ($R \geq 0.88$) and in good agreement with manual analysis (mean difference and limits of agreement for V_z and V_r were −0.05±0.98cm/s and −0.06±1.18cm/s for LV, and −0.21±2.33cm/s and 0.46±4.00cm/s for RV, respectively). The proposed multi-channel 3D dense U-Net was capable of reducing the segmentation time by 3600-times, without significant loss in accuracy in tissue velocity measurements.

Please send correspondence to: Daming Shen, PhD, Department of Radiology, Northwestern University, 737 N. Michigan Avenue, Suite 1600, Chicago, IL 60611, DamingShen2017@u.northwestern.edu.

*Ashitha Pathrose and Daming Shen contributed equally to this work

+Michael Markl and Daniel Kim contributed equally to this work

Keywords

deep learning (DL); tissue phase mapping (TPM); image segmentation; multi-channel 3D dense U-Net

INTRODUCTION

Cine cardiovascular magnetic resonance (CMR) using balanced steady-state free precession (1) is the gold-standard imaging modality for the evaluation of cardiac function (2,3). Global cardiac functional parameters such as end diastolic volume, end systolic volume, stroke volume, ejection fraction, and mass of the left ventricle (LV) and right ventricle (RV) are routinely evaluated from cine CMR images. However, for most diseases, intramyocardial abnormalities may precede global functional abnormalities (e.g. ejection fraction). CMR has the ability to non-invasively characterize intramyocardial function using techniques such as myocardial tagging (4,5), strain-encoded (SENC) (6), cine displacement-encoded imaging with stimulated echoes (DENSE) (7,8), and tissue phase mapping (TPM) (9,10). These techniques have proven useful in diagnosing and monitoring abnormal myocardial function in a wide range of cardiovascular conditions including congenital heart diseases (11,12), cardiotoxicities (13,14), cardiac transplantation (15–17), left ventricular dyssynchrony (18,19), and various cardiomyopathies (20–22).

In this study, we will focus on TPM, which is a black-blood 2D cine phase contrast MRI pulse sequence with three-directional tissue velocity encoding for quantification of regional biventricular myocardial velocities (23). Recently, several studies have demonstrated significant differences in biventricular global and regional velocities in heart transplant (HTx) patients (adult and pediatric) with and without transplant rejection (15,16). Despite these studies demonstrating feasibility, the clinical translation of TPM is hampered by the requisite labor-intensive manual segmentation of the myocardial contours for all cardiac frames. In general, manual segmentation of TPM images is more challenging for the RV than the LV, because the former has thinner wall, highly trabeculated myocardium, and epicardial fat partially superimposing on the myocardial wall. For clinical translation of TPM, there is a need to automate segmentation of bi-ventricular contours.

The task of developing a post-processing tool to automate the segmentation of myocardial contours in TPM data from HTx patients is challenging due to several factors. First, the black-blood preparation module is not perfect and often produces residual blood signal, which makes it challenging to define the endocardial border. Second, metal artifacts from open chest surgery makes it challenging to define the epicardial contours, especially for the RV. Third, TPM with gradient echo readouts is inherently a low signal-to-noise-ratio (SNR) pulse sequence, thereby making segmentation sensitive to noise. Previous studies have reported automatic or semi-automatic segmentation tools using traditional machine learning or atlas based methods for TPM images with various degrees of success (24–28). A major disadvantage of such methods is that they often require prior knowledge to achieve satisfactory accuracy. To date, none of them have enabled TPM to be translated into clinical practice.

Deep learning (DL) has found many applications in medical imaging, including image reconstruction (29–31), segmentation (32–34), and disease classification (35–37). The major advantage of DL over traditional segmentation methods is that neural networks are good at automatically discovering intricate features from data for object detection and segmentation. Another advantage of DL over traditional segmentation methods is that the inference processing time is several orders of magnitude faster. In this study, we sought to develop a fully automated segmentation method for TPM images using DL and evaluate its accuracy compared with manual LV and RV contour delineation.

METHODS

Patient Demographics

This study was conducted in accordance with protocols approved by our institutional review board and was Health Insurance Portability and Accountability Act (HIPAA) compliant. All subjects provided informed consent in writing and agreed to future analysis of their data. We retrospectively identified 99 patients with heart transplantation (mean age = 50 ± 15 years; 55 males; 44 females) who participated in a longitudinal study, where each patient underwent 1–4 CMR scans for post HTx cardiac monitoring (median duration post HTx: 4.4 years; range: 6 days to 30 years). In total, 150 CMR scans were included in this study. For patients with more than one scan, there was a gap of at least three months between consecutive scans.

MRI Hardware

TPM scans were conducted on a 1.5T whole-body MRI scanner (MAGNETOM Aera or Avanto, Siemens Healthcare, Erlangen, Germany). The scanners were equipped with a gradient system capable of achieving a maximum gradient strength of 45 mT/m and maximum slew rate of 200 T/m/s. Body coil was used for radio-frequency excitation. Both body matrix and spine coil arrays (30–34 elements in total) were used for signal reception.

Pulse Sequence

TPM data were acquired in three short-axis slices at basal, mid-ventricular, and apical locations using a prospectively ECG-gated, black-blood prepared 2D phase-contrast sequence with three-directional velocity encoding (11,38–40) (VENC = 25 cm/s). Spatiotemporal imaging acceleration using Parallel MRI with Extended and Averaged GRAPPA Kernels (PEAK-GRAPPA) (41) with an undersampling factor of 5 permitted breath-hold data acquisitions with scan time = 24–28 heart beats per slice. Other relevant imaging parameters included: temporal resolution = 19–24 ms, in-plane spatial resolution = 2.0–2.3 mm^2, slice thickness = 8 mm, TE = 3.2–3.8 ms, TR = 4.8–6.1 ms, receiver bandwidth = 460–840 Hz/pixel, flip angle 10° or 15°.

Manual TPM Data Analysis

TPM data post-processing and myocardial velocity estimations were made using a custom-made software package programmed in MATLAB (The Mathworks Inc, Natick, Mass). First, the velocity data were pre-processed by correcting eddy currents (42) and bulk-motion (38,39,43). For manual segmentation, the magnitude images in short-axis views were used

to place approximately 10–20 coordinate points, each for epicardial LV, endocardial LV, epicardial RV, and endocardial RV borders for the base, mid, and apex for all time frames, after which the software uses spline fitting to close the contour. The anterior and inferior LV-RV intersections were automatically detected for all time frames and used to remove the septum from the RV masks. The Cartesian velocity (V_x and V_y) within the segmented LV and RV masks were converted into velocities along the three principal directions of the heart - radial shortening (V_r), tangential/circumferential shortening (V_ϕ), and longitudinal shortening (V_z). For simplicity, only the V_r and V_z are considered for further statistical analyses. The expanded 16 LV +10 RV American Heart Association (AHA) model (44) was used to report segmental end-systolic and end-diastolic peak velocities. Global LV and RV peak velocities were obtained by averaging the segmental values for each ventricle. End-systole was detected automatically as the time frame with the smallest endocardial LV volume and end-diastole as the time frame with the largest LV volume (summed over all three slices). Analyzing each study manually from pre-processing to deriving velocities would take ~2 hours per patient with the most effort (95%) spent on manual placement of coordinate points.

Deep Learning Architecture

Consistent with prior studies (45,46), we have split our training and testing data to have approximately 3:1 ratio. For training, we randomly selected 114 scans (342 slices, 20–42 time frames per slice, 10,096 2D images) and for testing, the remaining 36 scans (108 slices, 23–36 time frames per slice, 3,288 2D images) were used. The pre-processed images after correction for eddy currents and bulk motion were used. While a fraction of image series (n=55 [12.6%] slices) had suboptimal image quality due to poor breath holding, we elected to include all image series in this study (n=39 [11.4%] for training; n=16 [14.8%] for testing.) As shown in Figure 1, the manual contours from three observers (AP, RS, AB; medical fellows with 2 to 8 years of experience) were transformed into multi-layer masks and used as the reference (i.e. labelled 0–3 for each pixel, 0: background, 1: blood pool, 2: RV myocardium, and 3: LV myocardium). A 3D (2D + time) dense U-Net (Figure 2) was used to learn the segmentation process, while 2D max pooling (2×2×1) was used to allow arbitrary number of time frames.

Three different ways of utilizing the TPM images as inputs are compared: 1) magnitude image alone; 2) magnitude image and a combined velocity image ($V_{sum} = \sqrt{V_x^2 + V_y^2 + V_z^2}$); 3) magnitude and three dimensional velocity-encoded (V_x, V_y, V_z) images as independent input channels (i.e. stacked in the channel dimension). Input 3) was used for other comparisons.

For the loss term, in addition to cross-entropy loss, we investigated into multi-class dice loss (47), and Hausdorff distance (48) loss to achieve better results. For the dice loss, each class was calculated separately by Equation 1.

$$L_{DSC} = 1 - \frac{2\sum_i s_i r_i}{\sum_i s_i + \sum_i r_i} \quad \text{Eq. 1,}$$

where s_i is the DL segmentation result, and r_i is the ground truth at each voxel i. The dice losses of all three classes (i.e. blood pool, LV myocardium and RV myocardium) are weighted equally. The Hausdorff distance is calculated for the boundary of the entire segmentation (i.e. three classes combined). The total loss is described by Equation 2, where K is the total number of classes (i.e. 3) and L_{DSCk} is the dice loss for class k.

$$L_{total} = L_{CE} + \frac{1}{K}\sum_{k=1}^{K} L_{DSCk} + L_{HD} \quad \text{Eq. 2}$$

Four different combinations of loss terms are compared: 1) L_{CE} alone; 2) $L_{CE} + L_{HD}$; 3) $L_{CE} + L_{DSC}$; 4) $L_{CE} + L_{DSC} + L_{HD}$. The loss terms were weighted equally in this study. Loss term 4) was used for other comparisons.

The training took 19.7 hours. As part of our efforts to ensure transparency and reproducibility, we have made available our dense U-Net architecture programmed in Pytorch (see the Data Availability Statement section). To check for overfitting, we performed a 5-fold cross validation experiment by repeating the training and testing as described above.

The DL-generated masks were converted into coordinate points and loaded onto the software for velocity estimations. Further analysis used the same semi-automatic procedure as for the manual analysis, interpolated splines were generated and LV-RV intersection points were identified.

Computer Hardware

For training and testing of the DL network, we used a GPU workstation (Tesla V100 32GB memory, NVIDIA, Santa Carla, California, USA; 32 Xeon E5–2620 v4 128 GB memory, Intel, Santa Clara, California, USA) equipped with Pytorch (Version 1.4, Berkeley Software Distribution), and MATLAB (R2020b, The Mathworks Inc, Natick, MA, USA) running on a Linux operating system (Ubuntu16.04).

Quantitative Analysis

To assess the accuracy of DL-based segmentation, we calculated the dice scores for LV and RV with manually contoured masks as reference. The Hausdorff distance is calculated for the entire heart segmentation to access the offsets of the boundaries. Peak-systolic and peak-diastolic LV and RV segmental velocities were compared between the manual segmentation and DL segmentation. A second independent observer (IO) manually analyzed 12 scans randomly selected from the 36 testing scans to evaluate inter-observer variability.

Statistical Analysis

The statistical analyses were conducted by two investigators (DS, AP) using MATLAB. We tested for normality of variables using the Shapiro-Wilk test. Normally distributed data were reported as mean ± standard deviation; non-normally distributed data were reported as median and interquartile range (IQR); 25^{th} percentile, 75^{th} percentile. Analysis of variance

or Kruskal-Wallis test with Bonferroni correction were used to compare the quantitative metrics among different input groups and different loss term groups. Paired t-tests or Wilcoxon signed rank tests were used to compare the quantitative metrics between the manual and DL segmentations. Pearson correlation (r) and Bland Altman analyses were used to compare velocities derived from the manual and DL segmentations. A $p < 0.05$ was considered significant for each statistical test.

RESULTS

The mean segmentation time was approximately 2 hours per patient (3 slices per patient) for manual and 1.9 ± 0.3 seconds for DL segmentation. According to Shapiro-Wilk test, all three image quality metrics were not normally distributed (statistic = 0.874, $p<0.001$ for LV dice; statistic = 0.931, $p <0.001$ for RV dice; statistic = 0.895, $p <0.001$ for Hausdorff distance). Therefore, the Kruskal-Wallis and Wilcoxon signed rank tests were used to compare quantitative metrics.

Table S1 in Supplementary Materials summarizes the quantitative metrics measured on the 36 testing cases comparing the deep learning outcomes using different inputs. All three quantitative metrics were not significantly ($p>0.37$) different between the three groups. As shown in Figure S1, the magnitude + phase (V_x, V_y, V_z) as input produced better segmentation than other input cases, which contained noticeable discontinuity on LV or RV myocardium masks. Given that magnitude + phase (V_x, V_y, V_z) group produced better median quantitative metrics, we elected to use it throughout.

Table S2 in Supplementary Materials summarizes the quantitative comparison with different loss terms. While the LV and RV dice scores were not significantly ($p>0.38$) different among between four groups, the Hausdorff distance was significantly ($p<0.03$) different between the four groups. As shown in Figure S2 in Supplementary Materials, the CE + HD + Dice loss term produced better results than other loss terms. Thus, we elected to use it throughout.

As summarized in Table 1, the median dice scores for the LV and RV DL segmentations were 0.85 and 0.64, respectively. The median Hausdorff distance of the testing set is 3.17 pixels. As shown in Table 2, for the 12 scans including analyses by two observers, the dice score was significantly better for DL than the second independent observer for LV (DL: 0.86; manual IO: 0.80; $p<0.001$), but they were not significantly different for RV (DL: 0.62; second IO: 0.60; $p=0.32$). The median Hausdorff distance was not significantly different between DL and second IO (DL: 3.33 pixels; second IO: 3.58 pixels; $p=0.23$).

In the 5-fold cross validation experiment, all three quantitative metrics were not significantly ($p>0.1$) different among the five groups (see Table S3 in Supplementary Materials). Therefore, we used the first experiment results throughout.

Figure 3 shows four representative cases of TPM segmentation using manual contouring and deep learning. These example results show good agreement between manual and DL segmentation. For dynamic display of Figure 3, see Video S1 in Supplemental Materials. Figure 4 shows example V_r and V_z time curves of LV and RV for manual and DL contours on the right and the corresponding velocity maps at end-systolic and end diastolic

time frames on the left. This example shows good agreement in time-resolved velocity measurements derived from manual and DL segmentation. For dynamic display of Figure 4, see Video S2 in Supplemental Materials.

Figure 5 shows scatter plots resulting from linear regression analysis illustrating strong correlation between manual and DL segmentation methods ($R \geq 0.88$) and between two independent observers ($R \geq 0.89$) for peak V_r and V_z (LV and RV, systole and diastole). All 26 segments (LV and RV 16+10 segment AHA model) are plotted with the basal, mid-ventricular, and apical segments color-coded as red, blue and green, respectively. Figure 6 shows Bland-Altman plots illustrating good agreement between manual and DL segmentation for LV V_z (mean = 3.85 cm/s; mean difference = −0.05 cm/s (1.3 % of mean); the upper and lower limits of agreement [LOA] = −0.05 ± 0.98 cm/s); LV V_r (mean = 3.38 cm/s; mean difference = −0.06 cm/s (1.8% of mean); and the upper and lower LOA = −0.06 ± 1.18 cm/s); RV V_z (mean = 2.96 cm/s; mean difference = −0.21 cm/s (7.1% of mean); and the upper and lower LOA = −0.21 ± 2.33 cm/s); RV V_r (mean = 3.95 cm/s; mean difference = 0.46 cm/s (11.7% of mean); and the upper and lower LOA = 0.46 ± 4.00 cm/s).

Figure 6 also shows good agreement between independent observers for LV V_z (mean = 3.54 cm/s; mean difference = −0.07 cm/s (2.0% of mean); and the upper and lower LOA = −0.07 ± 1.10 cm/s); LV V_r (mean = 3.33 cm/s, mean difference = −0.13 cm/s (3.9% of mean); and the upper and lower LOA = −0.13 ± 1.21 cm/s); RV V_z (mean = 2.82; mean difference = −0.32 cm/s (11.4% of mean); and the upper and lower LOA = −0.32 ± 2.92 cm/s); RV V_r (mean = 4.11 cm/s; mean difference = 0.47 cm/s (11.4% of mean); and LOA = 0.47 ± 4.06 cm/s).

DISCUSSION

This study describes the implementation and evaluation of a DL-based automated image segmentation method for TPM images. As expected, the processing time is 3,600-times shorter for DL (~2 s) than manual segmentation (~7,200 s). The resulting accuracy (LV dice = 0.85, RV dice = 0.64, Hausdorff distance = 3.17 pixels) in segmentation with DL was slightly better than inter-observer agreement (LV dice = 0.80, RV dice = 0.60, Hausdorff distance = 3.58 pixels).

This study has several points that warrant further explanations. First, we used a 3D dense U-Net architecture in this study. Compared to a traditional convolution layer with increasing channel sizes in the deeper layers, the dense layer uses a series of convolutions of relatively small sized channels (i.e. 16) and concatenates the feature maps from previous convolutions (49,50). This largely reduces the overall number of parameters (3D U-Net: 18.8M, ours: 2.7M) by efficiently utilizing every feature extracted throughout the CNN, thereby reducing the significant computational demand (e.g. GPU memory) for 3D CNNs (51–53). Second, by utilizing the magnitude and phase (V_x, V_y, V_z) images together, the dense U-Net produced better results than using magnitude alone (see Figure S1 and Video S3 in Supplementary Materials). Previously, our team found similar results in a cohort of 26 pulmonary hypertension patients and 8 healthy controls (54). A 4-channel network using the magnitude and 3 phase images together showed better results than a 1-channel

network using only the magnitude images and a 2-channel network using the magnitude and a combined velocity ($V_{sum} = \sqrt{V_x^2 + V_y^2 + V_z^2}$). While the phase images are hard to be used by human, they can be utilized to enrich the image features learned by DL networks, thereby improve the performance of DL networks. Third, incorporating cross-entropy loss, dice loss and Hausdorff distance into the loss function produced better results (see Figure S2 and Video S4 in Supplementary Materials). Forth, the dice scores were significantly better for DL than second IO for LV (DL: 0.86; second IO: 0.80; $p<0.001$) but not for RV (DL: 0.62; second IO: 0.60; $p=0.49$). For more accurate segmentation, DL results can be used as an initial guess and then further improved with minor manual adjustments, at the expense of increased processing time. This semi-automated method may still be fast enough for clinical translation of TPM.

DL-based RV myocardial segmentation is challenging due to its complex crescent shape that varies across slices and phases. The inhomogeneity in shape and myocardial signal intensity in our cohort of post-cardiac transplant subjects adds to this problem. Previous works on RV myocardial segmentation are limited. The finalists of the MICCAI 2012 Right Ventricle Segmentation Challenge used either automated or semi-automated approaches, three of which are atlas-based methods, two are prior-based methods, and two are prior-free, image-driven methods that make use of the temporal dimension of the data (55). An end-to-end DL-based CNN architecture was implemented by Tran et al. to segment the LV and RV myocardium from short-axis cine slices (56). All of these prior works were conducted using cine CMR images, which have a high blood-to-myocardial contrast. A future work is warranted to improve the RV segmentation performance in TPM images.

This study has several limitations that warrant further discussion. First, we did not document the manual segmentation time from prior analysis. For this study, one observer (AS) with prior experience with TPM analysis repeated the analysis for three training datasets to derive an approximate segmentation time of 2 hours per patient. Second, the DL segmentation results may be influenced by poor image quality (i.e. image artifacts caused by breathing motion, metals), whereas manual contours is less sensitive to artifacts because trained observers may use prior knowledge to over read poor image quality. One approach to improve DL based segmentation is by incorporating shape models. Third, as shown in Figure 6, the agreement in Vz and Vr was significantly worse for RV than LV. This may be due to: 1) partial volume averaging, as the RV myocardium is thinner than the LV; 2) susceptibility artifact from open chest surgery (i.e. signal void caused by sternal wires), which is closer to the RV; 3) epicardial fat signal, which makes it harder to determine epicardial RV boundary. The agreement in Vr for RV was worse for diastole with negative velocities than systole with positive velocities. This may be due to: 1) partial volume averaging, as the RV is thinner at diastole than systole; 2) proximity to signal void caused by sternal wires, which is closer at diastole than systole. Fourth, we did not compare our DL-based method to previously published semi-automated or automated methods for TPM (24–27) due to lack of access to such methods. A future study is warranted to conduct a head-to-head comparison for segmentation accuracy and computational efficiency. Fifth, this study used data obtained from a single site, scanner vendor, and field strength, which may limit generalizability to other sites, scanner vendors and field strengths. Sixth, we used equal

weight for the different loss terms in this study. A future study is warranted to determine the optimal weight for each loss term to achieve best results.

In summary, this study describes an automated image segmentation method for biventricular TPM images with deep learning that is significantly faster than manual contouring, without significant loss in segmentation accuracy and TPM parameters, thereby verifying clinical translatability.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank funding support from the National Institutes of Health (R01HL116895, R01HL138578, R21EB024315, R21AG055954, R01HL151079, R01HL117888, T32EB025766, R21EB030806) and American Heart Association (19IPLOI34760317).

List of financial Support:

National Institutes of Health (R01HL116895, R01HL138578, R21EB024315, R21AG055954, R01HL151079, R01HL117888, T32EB025766, R21EB030806) and American Heart Association (19IPLOI34760317)

None of the authors have relationships with industry related to this study

Data Availability Statement

The Pytorch codes used for implementing our dense U-Net can be found in GitHub (https://github.com/dsc936/DenseUnet_for_TPM_segmentation).

List of Abbreviations

AHA	American Heart Association
CMR	Cardiovascular Magnetic Resonance
DENSE	Displacement-Encoded imaging with Stimulated Echoes
DL	Deep Learning
DICOM	Digital Imaging and Communications in Medicine
ECV	Extracellular Volume
FOV	Field of View
GPU	Graphic Processing Unit
HIPAA	Health Insurance Portability and Accountability Act
HTx	Heart Transplant
LOA	Limits of Agreement
LV	Left Ventricular

IO	Independent Observer
RV	Right Ventricular
SENC	Strain-Encoded
SNR	Signal-to-Noise-Ratio
TPM	Tissue Phase Mapping

Bibliography & References Cited

1. Carr JC, Simonetti O, Bundy J, Li D, Pereles S, Finn JP. Cine MR angiography of the heart with segmented true fast imaging with steady-state precession. Radiology 2001;219(3):828–834. [PubMed: 11376278]
2. American College of Cardiology Foundation Task Force on Expert Consensus D, Hundley WG, Bluemke DA, Finn JP, Flamm SD, Fogel MA, Friedrich MG, Ho VB, Jerosch-Herold M, Kramer CM, Manning WJ, Patel M, Pohost GM, Stillman AE, White RD, Woodard PK. ACCF/ACR/AHA/NASCI/SCMR 2010 expert consensus document on cardiovascular magnetic resonance: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents. J Am Coll Cardiol 2010;55(23):2614–2662. [PubMed: 20513610]
3. Grothues F, Smith GC, Moon JC, Bellenger NG, Collins P, Klein HU, Pennell DJ. Comparison of interstudy reproducibility of cardiovascular magnetic resonance with two-dimensional echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. The American journal of cardiology 2002;90(1):29–34. [PubMed: 12088775]
4. Zerhouni EA, Parish DM, Rogers WJ, Yang A, Shapiro EP. Human heart: tagging with MR imaging--a method for noninvasive assessment of myocardial motion. Radiology 1988;169(1):59–63. [PubMed: 3420283]
5. Axel L, Dougherty L. MR imaging of motion with spatial modulation of magnetization. Radiology 1989;171(3):841–845. [PubMed: 2717762]
6. Osman NF, Sampath S, Atalar E, Prince JL. Imaging longitudinal cardiac strain on short-axis images using strain-encoded MRI. Magnetic Resonance in Medicine 2001;46(2):324–334. [PubMed: 11477637]
7. Aletras AH, Ding S, Balaban RS, Wen H. DENSE: displacement encoding with stimulated echoes in cardiac functional MRI. J Magn Reson 1999;137(1):247–252. [PubMed: 10053155]
8. Kim D, Gilson WD, Kramer CM, Epstein FH. Myocardial tissue tracking with two-dimensional cine displacement-encoded MR imaging: development and initial evaluation. Radiology 2004;230(3):862–871. [PubMed: 14739307]
9. Hennig J, Schneider B, Peschl S, Markl M, Laubenberger TKJ. Analysis of myocardial motion based on velocity measurements with a black blood prepared segmented gradient-echo sequence: Methodology and applications to normal volunteers and patients. Journal of Magnetic Resonance Imaging 1998;8(4):868–877. [PubMed: 9702889]
10. Jung B, Föll D, Böttler P, Petersen S, Hennig J, Markl M. Detailed analysis of myocardial motion in volunteers and patients using high-temporal-resolution MR tissue phase mapping. Journal of Magnetic Resonance Imaging 2006;24(5):1033–1039. [PubMed: 16947325]
11. Ruh A, Sarnari R, Berhane H, Sidoryk K, Lin K, Dolan R, Li A, Rose MJ, Robinson JD, Carr JC, Rigsby CK, Markl M. Impact of age and cardiac disease on regional left and right ventricular myocardial motion in healthy controls and patients with repaired tetralogy of fallot. Int J Cardiovasc Imaging 2019;35(6):1119–1132. [PubMed: 30715669]
12. Kempny A, Fernández-Jiménez R, Orwat S, Schuler P, Bunck AC, Maintz D, Baumgartner H, Diller G-P. Quantification of biventricular myocardial function using cardiac magnetic resonance feature tracking, endocardial border delineation and echocardiographic speckle tracking in patients with repaired tetralogy of Fallot and healthy controls. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance 2012;14(1):32–32. [PubMed: 22650308]

13. Houbois CP, Nolan M, Somerset E, Shalmon T, Esmaeilzadeh M, Lamacie MM, Amir E, Brezden-Masley C, Koch CA, Thevakumaran Y, Yan AT, Marwick TH, Wintersperger BJ, Thavendiranathan P. Serial Cardiovascular Magnetic Resonance Strain Measurements to Identify Cardiotoxicity in Breast Cancer: Comparison With Echocardiography. JACC Cardiovasc Imaging 2020.

14. Fallah-Rad N, Walker JR, Wassef A, Lytwyn M, Bohonis S, Fang T, Tian G, Kirkpatrick ID, Singal PK, Krahn M, Grenier D, Jassal DS. The utility of cardiac biomarkers, tissue velocity and strain imaging, and cardiac magnetic resonance imaging in predicting early left ventricular dysfunction in patients with human epidermal growth factor receptor II-positive breast cancer treated with adjuvant trastuzumab therapy. J Am Coll Cardiol 2011;57(22):2263–2270. [PubMed: 21616287]

15. Sarnari R, Blake AM, Ruh A, Abbasi MA, Pathrose A, Blaisdell J, Dolan RS, Ghafourian K, Wilcox JE, Khan SS, Vorovich EE, Rich JD, Anderson AS, Yancy CW, Carr JC, Markl M. Evaluating Biventricular Myocardial Velocity and Interventricular Dyssynchrony in Adult Patients During the First Year After Heart Transplantation. J Magn Reson Imaging 2020.

16. Berhane H, Ruh A, Husain N, Robinson JD, Rigsby CK, Markl M. Myocardial velocity, intra-, and interventricular dyssynchrony evaluated by tissue phase mapping in pediatric heart transplant recipients. Journal of Magnetic Resonance Imaging 2020;51(4):1212–1222. [PubMed: 31515865]

17. Shenoy C, Romano S, Hughes A, Okasha O, Nijjar Prabhjot S, Velangi P, Martin Cindy M, Akçakaya M, Farzaneh-Far A. Cardiac Magnetic Resonance Feature Tracking Global Longitudinal Strain and Prognosis After Heart Transplantation. JACC: Cardiovascular Imaging 2020;13(9):1934–1942. [PubMed: 32563650]

18. Delfino JG, Bhasin M, Cole R, Eisner RL, Merlino J, Leon AR, Oshinski JN. Comparison of myocardial velocities obtained with magnetic resonance phase velocity mapping and tissue Doppler imaging in normal subjects and patients with left ventricular dyssynchrony. J Magn Reson Imaging 2006;24(2):304–311. [PubMed: 16786564]

19. Westenberg JJ, Lamb HJ, van der Geest RJ, Bleeker GB, Holman ER, Schalij MJ, de Roos A, van der Wall EE, Reiber JH, Bax JJ. Assessment of left ventricular dyssynchrony in patients with conduction delay and idiopathic dilated cardiomyopathy: head-to-head comparison between tissue doppler imaging and velocity-encoded magnetic resonance imaging. J Am Coll Cardiol 2006;47(10):2042–2048. [PubMed: 16697323]

20. Orwat S, Kempny A, Diller GP, Bauerschmitz P, Bunck A, Maintz D, Radke RM, Baumgartner H. Cardiac magnetic resonance feature tracking: a novel method to assess myocardial strain. Comparison with echocardiographic speckle tracking in healthy volunteers and in patients with left ventricular hypertrophy. Kardiol Pol 2014;72(4):363–371. [PubMed: 24293146]

21. Foell D, Jung B, Germann E, Staehle F, Bode C, Markl M. Hypertensive heart disease: MR tissue phase mapping reveals altered left ventricular rotation and regional myocardial long-axis velocities. European Radiology 2013;23(2):339–347. [PubMed: 22886534]

22. Buss SJ, Breuninger K, Lehrke S, Voss A, Galuschky C, Lossnitzer D, Andre F, Ehlermann P, Franke J, Taeger T, Frankenstein L, Steen H, Meder B, Giannitsis E, Katus HA, Korosoglou G. Assessment of myocardial deformation with cardiac magnetic resonance strain imaging improves risk stratification in patients with dilated cardiomyopathy. European Heart Journal - Cardiovascular Imaging 2014;16(3):307–315. [PubMed: 25246506]

23. Menza M, Föll D, Hennig J, Jung B. Segmental biventricular analysis of myocardial function using high temporal and spatial resolution tissue phase mapping. Magnetic Resonance Materials in Physics, Biology and Medicine 2018;31(1):61–73.

24. Zhu Y, Pelc NJ. A spatiotemporal model of cyclic kinematics and its application to analyzing nonrigid motion with MR velocity images. IEEE Trans Med Imaging 1999;18(7):557–569. [PubMed: 10504090]

25. Bergvall E, Cain P, Arheden H, Sparr G. A fast and highly automated approach to myocardial motion analysis using phase contrast magnetic resonance imaging. J Magn Reson Imaging 2006;23(5):652–661. [PubMed: 16568435]

26. Cho J, Benkeser PJ. Cardiac segmentation by a velocity-aided active contour model. Computerized Medical Imaging and Graphics 2006;30(1):31–41. [PubMed: 16378714]

27. Meyer FG, Constable RT, Sinusas AJ, Duncan JS. Tracking myocardial deformation using phase contrast MR velocity fields: a stochastic approach. IEEE Transactions on Medical Imaging 1996;15(4):453–465. [PubMed: 18215927]
28. Chitiboi T, Schnell S, Collins J, Carr J, Chowdhary V, Honarmand AR, Hennemuth A, Linsen L, Hahn HK, Markl M. Analyzing myocardial torsion based on tissue phase mapping cardiovascular magnetic resonance. Journal of Cardiovascular Magnetic Resonance 2016;18(1):15. [PubMed: 27062364]
29. Hyun CM, Kim HP, Lee SM, Lee S, Seo JK. Deep learning for undersampled MRI reconstruction. Physics in medicine and biology 2018;63(13):135007. [PubMed: 29787383]
30. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. IEEE Trans Med Imaging 2018;37(2):491–503. [PubMed: 29035212]
31. Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, Liu F, Arridge S, Keegan J, Guo Y, Firmin D, Keegan J, Slabaugh G, Arridge S, Ye X, Guo Y, Yu S, Liu F, Firmin D, Dragotti PL, Yang G, Dong H. DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Trans Med Imaging 2018;37(6):1310–1321. [PubMed: 29870361]
32. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. 2016. Springer. p 424–432.
33. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. 2015. Springer. p 234–241.
34. LaLonde R, Xu Z, Irmakci I, Jain S, Bagci U. Capsules for biomedical image segmentation. Medical image analysis 2020;68:101889. [PubMed: 33246227]
35. Craik A, He Y, Contreras-Vidal JL. Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering 2019;16(3):031001. [PubMed: 30808014]
36. Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, Filippi M. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage Clinical 2019;21:101645. [PubMed: 30584016]
37. Al-Antari MA, Al-Masni MA, Kim TS. Deep Learning Computer-Aided Diagnosis for Breast Lesion in Digital Mammogram. Advances in experimental medicine and biology 2020;1213:59–72. [PubMed: 32030663]
38. Hennig J, Schneider B, Peschl S, Markl M, Krause T, Laubenberger J. Analysis of myocardial motion based on velocity measurements with a black blood prepared segmented gradient-echo sequence: methodology and applications to normal volunteers and patients. J Magn Reson Imaging 1998;8(4):868–877. [PubMed: 9702889]
39. Jung B, Föll D, Böttler P, Petersen S, Hennig J, Markl M. Detailed analysis of myocardial motion in volunteers and patients using high-temporal-resolution MR tissue phase mapping. Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine 2006;24(5):1033–1039.
40. Markl M, Rustogi R, Galizia M, Goyal A, Collins J, Usman A, Jung B, Foell D, Carr J. Myocardial T2-mapping and velocity mapping: changes in regional left ventricular structure and function after heart transplantation. Magn Reson Med 2013;70(2):517–526. [PubMed: 23008092]
41. Bauer S, Markl M, Föll D, Russe M, Stankovic Z, Jung B. K-t GRAPPA accelerated phase contrast MRI: Improved assessment of blood flow and 3-directional myocardial motion during breath-hold. J Magn Reson Imaging 2013;38(5):1054–1062. [PubMed: 23908094]
42. Walker PG, Cranney GB, Scheidegger MB, Waseleski G, Pohost GM, Yoganathan AP. Semiautomated method for noise reduction and background phase error correction in MR phase velocity data. J Magn Reson Imaging 1993;3(3):521–530. [PubMed: 8324312]
43. Markl M, Schneider B, Hennig J. Fast phase contrast cardiac magnetic resonance imaging: Improved assessment and analysis of left ventricular wall motion. Journal of Magnetic Resonance Imaging 2002;15(6):642–653. [PubMed: 12112514]
44. Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK, Pennell DJ, Rumberger JA, Ryan T, Verani MS. Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart. Circulation 2002;105(4):539–542. [PubMed: 11815441]

45. Kofler A, Dewey M, Schaeffter T, Wald C, Kolbitsch C. Spatio-Temporal Deep Learning-Based Undersampling Artefact Reduction for 2D Radial Cine MRI With Limited Training Data. IEEE Trans Med Imaging 2020;39(3):703–717. [PubMed: 31403407]
46. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Carapella V, Kim YJ, Suzuki H, Kainz B, Matthews PM, Petersen SE, Piechnik SK, Neubauer S, Glocker B, Rueckert D. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. Journal of cardiovascular magnetic resonance : official journal of the Society for Cardiovascular Magnetic Resonance 2018;20(1):65. [PubMed: 30217194]
47. Shen C, Roth HR, Oda H, Oda M, Hayashi Y, Misawa K, Mori K. On the influence of Dice loss function in multi-class organ segmentation of abdominal CT using 3D fully convolutional networks. arXiv preprint arXiv:180105912 2018.
48. Karimi D, Salcudean SE. Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks. IEEE Trans Med Imaging 2020;39(2):499–513. [PubMed: 31329113]
49. Huang G, et al. Densly Connected Convolutional Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016:4700–4708.
50. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. 2017. p 4700–4708.
51. Chmelik J, Jakubicek R, Walek P, Jan J, Ourednicek P, Lambert L, Amadori E, Gavelli G. Deep convolutional neural network-based segmentation and classification of difficult to define metastatic spinal lesions in 3D CT data. Med Image Anal 2018;49:76–88. [PubMed: 30114549]
52. Ambellan F, Tack A, Ehlke M, Zachow S. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the Osteoarthritis Initiative. Med Image Anal 2019;52:109–118. [PubMed: 30529224]
53. Mehrtash A, Ghafoorian M, Pernelle G, Ziaei A, Heslinga FG, Tuncali K, Fedorov A, Kikinis R, Tempany CM, Wells WM, Abolmaesumi P, Kapur T. Automatic Needle Segmentation and Localization in MRI With 3-D Convolutional Neural Networks: Application to MRI-Targeted Prostate Biopsy. IEEE Trans Med Imaging 2019;38(4):1026–1036. [PubMed: 30334789]
54. Shen D, Pathrose A, Baraboo JJ, Gordon DZ, Cuttica MJ, Carr JC, Markl M, Kim D. Automated Segmentation for Myocardial Tissue Phase Mapping Images using Deep Learning. In: Proceedings of the 28th Annual Meeting of the ISMRM. 2020. Abstract 3522.
55. Petitjean C, Zuluaga MA, Bai W, Dacher JN, Grosgeorge D, Caudron J, Ruan S, Ayed IB, Cardoso MJ, Chen HC, Jimenez-Carretero D, Ledesma-Carbayo MJ, Davatzikos C, Doshi J, Erus G, Maier OM, Nambakhsh CM, Ou Y, Ourselin S, Peng CW, Peters NS, Peters TM, Rajchl M, Rueckert D, Santos A, Shi W, Wang CW, Wang H, Yuan J. Right ventricle segmentation from cardiac MRI: a collation study. Med Image Anal 2015;19(1):187–202. [PubMed: 25461337]
56. Tran PV. A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI. 2016. p arXiv:1604.00494.

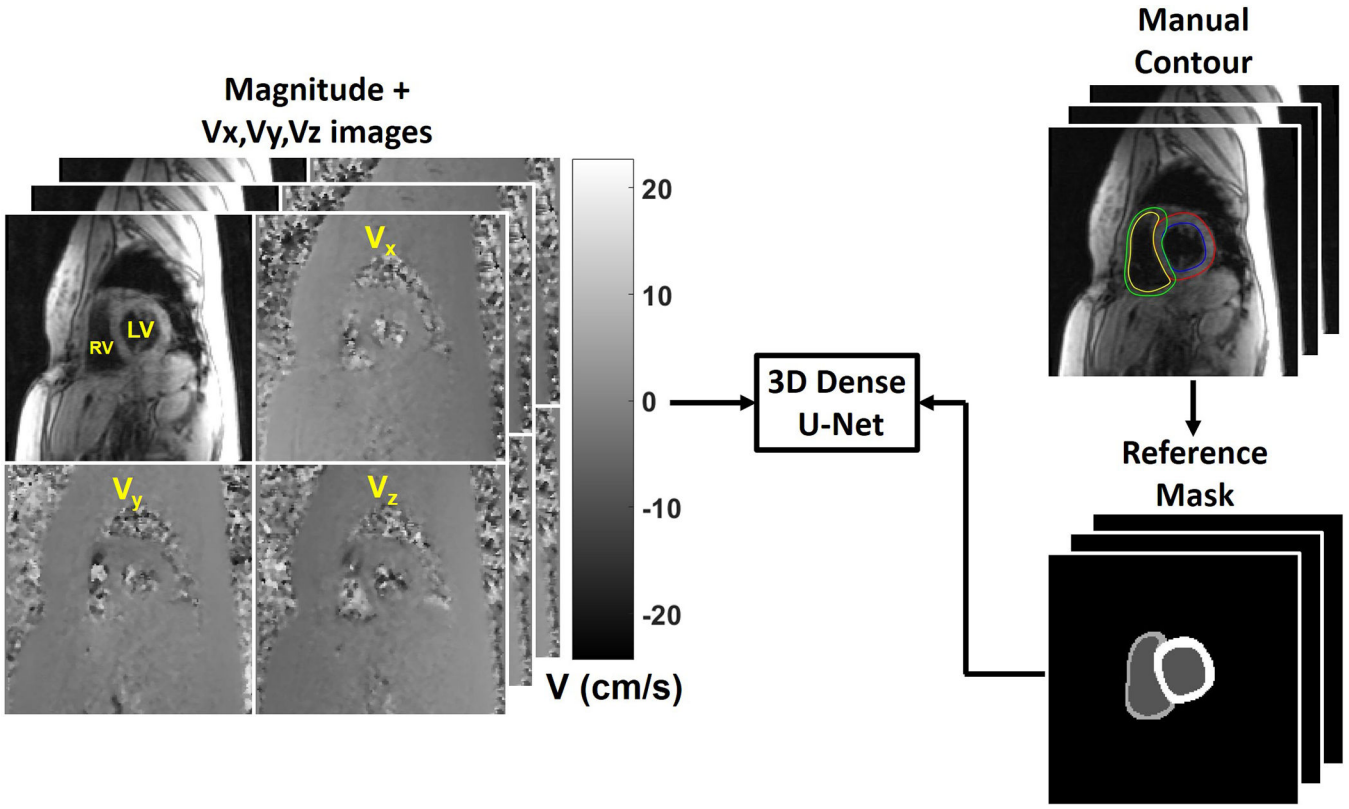

Figure 1.

The manual contours were transformed into multi-layer masks (i.e. 0 for background, 1 for blood pool, 2 for RV myocardium, and 3 for LV myocardium). We used the magnitude image and three dimensional velocity-encoded (V_x, V_y, V_z) images as independent input channels and the multi-class masks as the reference to train a 3D dense U-Net.

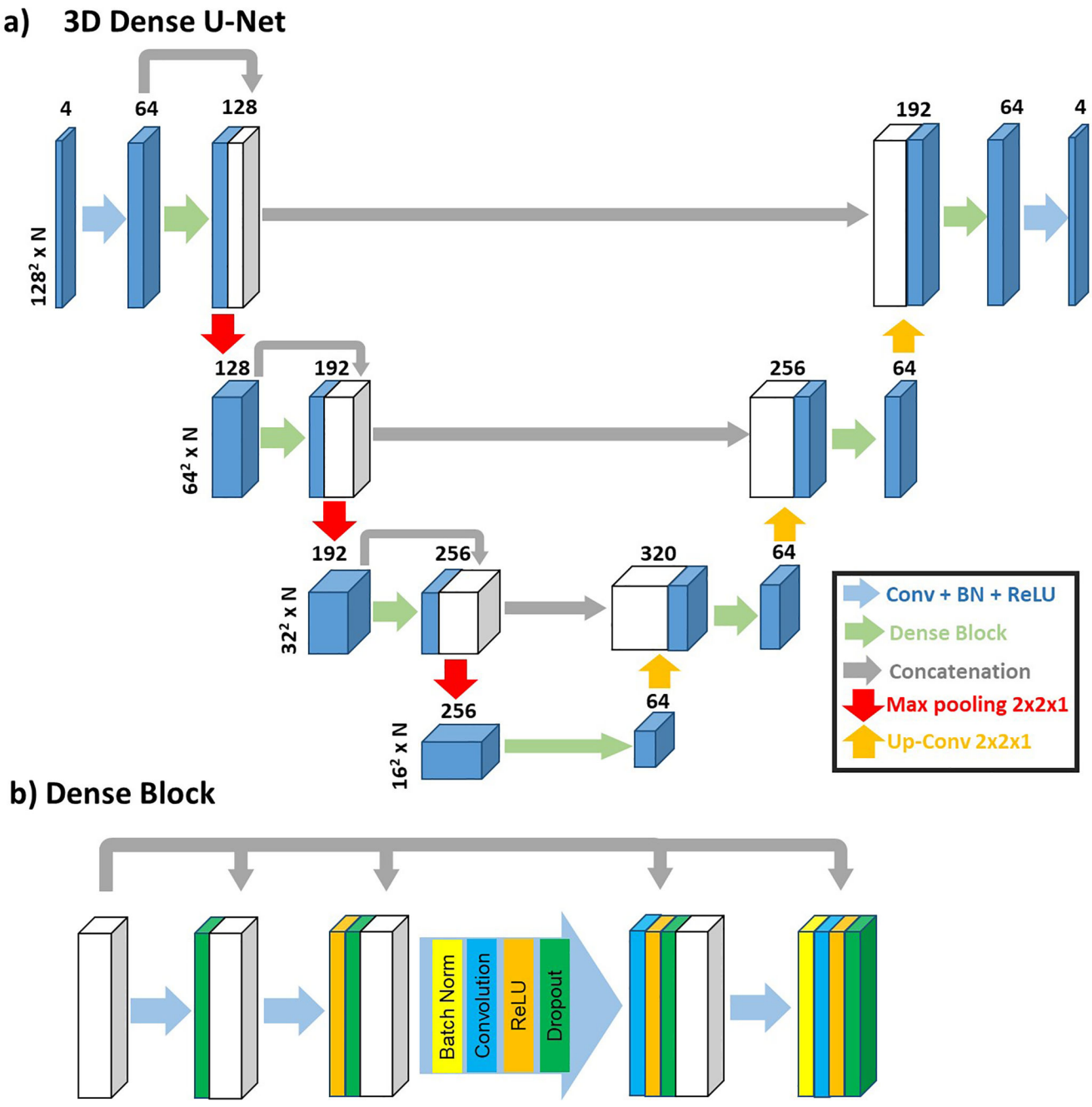

Figure 2.

a) A 3D (2D + time) Dense U-Net was used to learn the segmentation process, while max pooling (2×2×1) was used to allow arbitrary number (N) of time frames. b) The structure of dense block used in the 3D Dense U-Net.

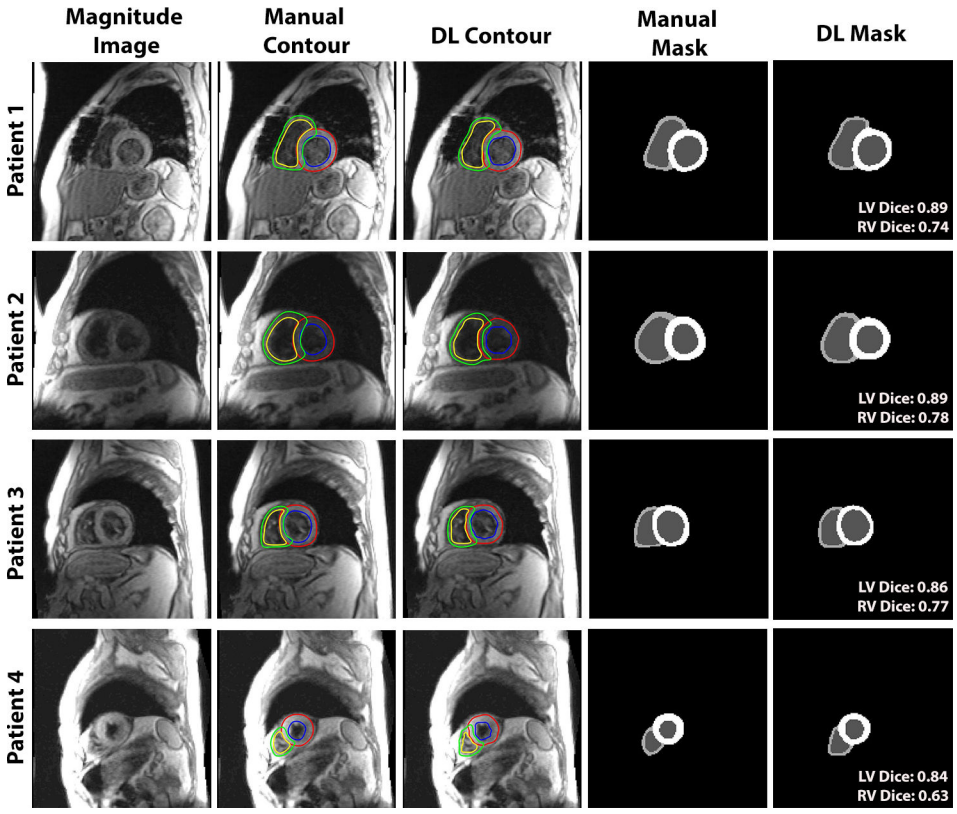

Figure 3.

Four representative patients comparing manual and DL segmentations. (Left column) The magnitude image; color-coded contours produced by manual (second column) and DL (third column): LV epicardium (red), LV endocardium (blue), RV epicardium (green) and RV endocardium (yellow). (4^{th} and 5^{th} column from the left) The multi-layer masks produced by manual (fourth column) and DL (fifth column). The dice scores of LV and RV myocardium for each case are labeled on the lower right corner of the DL masks.

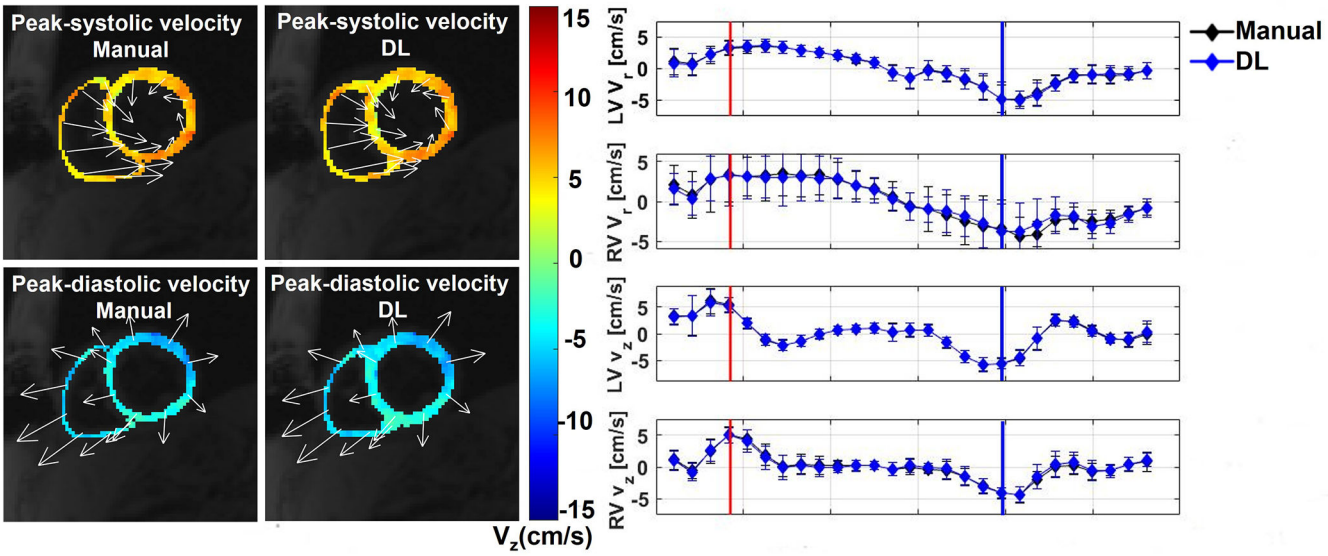

Figure 4.

Linear regression plots illustrating strong correlation between segmentation methods (top row, manual vs. DL, 36 testing cases, $R \geq 0.88$) and between independent observers (bottom row, 12 manual IO cases, $R \geq 0.88$) for peak V_r and V_z (LV and RV, systole and diastole). All 26 segments are plotted with the basal, mid-ventricular, and apical segments color-coded as red, blue and green, respectively.

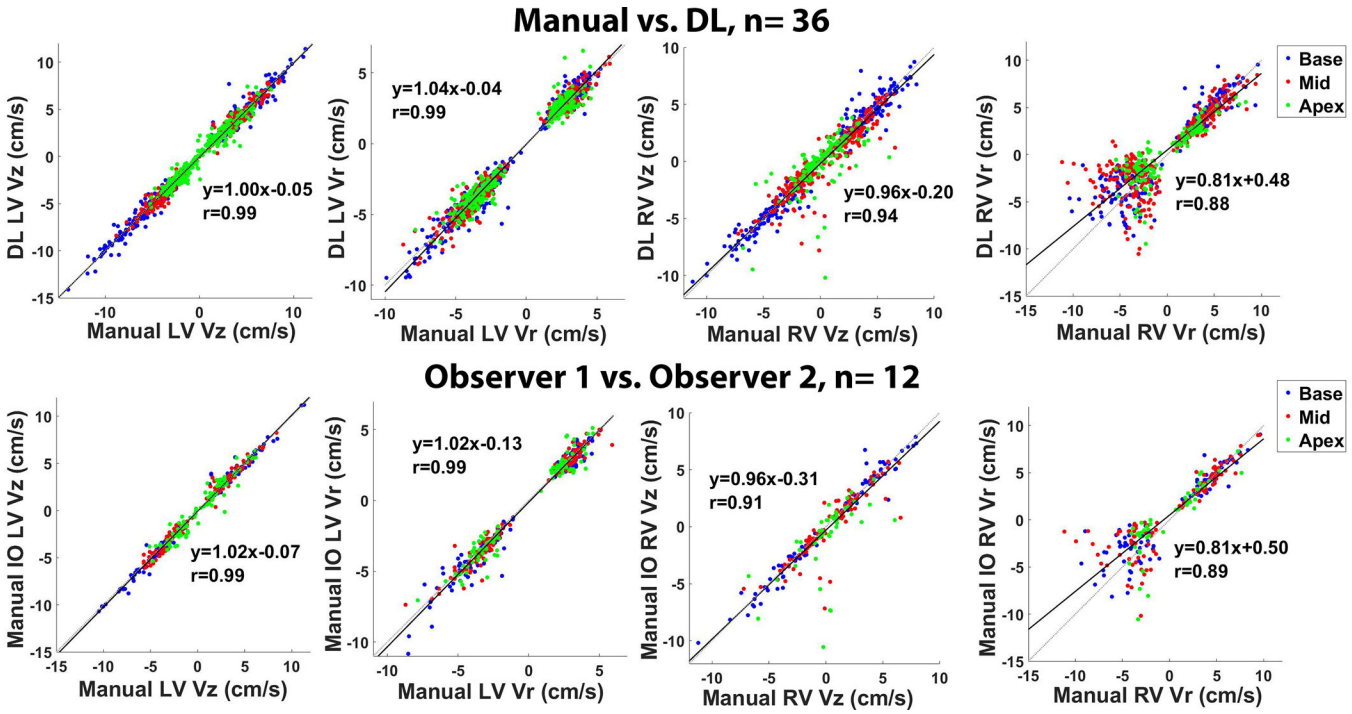

Figure 5.

Bland-Altman plots illustrating good agreement between segmentation methods (top row, manual vs. DL, 36 testing cases) and between independent observers (bottom row, 12 manual IO cases) for peak V_r and V_z (LV and RV, systole and diastole). All 26 segments are plotted with the basal, mid-ventricular, and apical segments color-coded as red, blue and green, respectively.

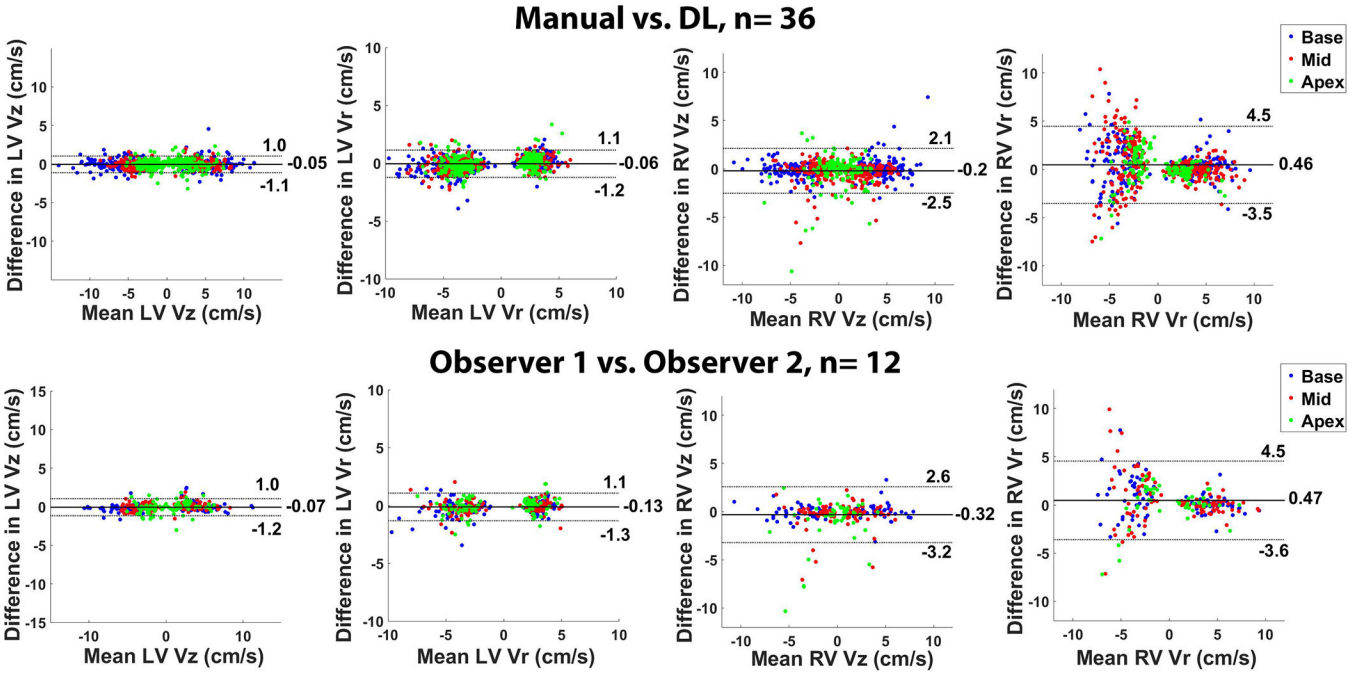

Figure 6.

(Left) Biventricular velocity maps derived from manual and DL segmentations at the peak-systolic and peak-diastolic time frames. Myocardial longitudinal velocities are color-coded and in-plane velocities are depicted by regionally averaged velocity vectors. (Right) Time resolved V_r and V_z curves of LV and RV. The red and blue vertical lines represent the time frames shown on the left (peak-systole and peak-diastole), while the black and blue curves represent manual and DL contours, respectively, with each time-frame represented by a rhombus.

Table 1.

Summary of quantitative metrics of 36 testing cases comparing deep learning segmentation versus manual segmentation. Reported values represent median and 25th to 75th percentiles (parenthesis).

	LV Dice Score	RV Dice Score	Hausdorff Distance
Basal	0.84 (0.81–0.89)	0.69 (0.62–0.76)	3.19 (2.48–4.06)
Mid	0.85 (0.82–0.89)	0.67 (0.60–0.73)	2.93 (2.49–3.53)
Apex	0.84 (0.78–0.85)	0.46 (0.37–0.62)	3.59 (2.66–3.94)
Combined	0.85 (0.80–0.88)	0.64 (0.47–0.73)	3.17 (2.52–3.93)

Table 2.

Summary of quantitative metrics (dice and Hausdorff distance) of image quality from all 12 testing cases compared with second independent observer (IO) as reference. Reported values represent median and 25th to 75th percentiles (parenthesis).

	LV Dice Score		RV Dice Score		Hausdorff Distance	
	DL	**Second IO**	**DL**	**Second IO**	**DL**	**Second IO**
Basal	0.88 (0.84–0.89)	0.80 (0.76–0.85)	0.73 (0.64–0.78)	0.70 (0.62–0.73)	2.81 (2.34–3.75)	3.14 (2.79–3.80)
Mid	0.86 (0.79–0.89)	0.80 (0.74–0.82)	0.67 (0.54–0.73)	0.64 (0.52–0.70)	3.04 (2.58–3.98)	3.11 (2.93–3.94)
Apex	0.83 (0.78–0.88)	0.74 (0.71–0.83)	0.39 (0.33–0.53)	0.38 (0.26–0.51)	3.82 (3.43–5.58)	4.36 (3.38–5.33)
Combined	0.86 (0.80–0.89)	0.80 (0.73–0.83)	0.62 (0.49–0.73)	0.60 (0.47–0.71)	3.33 (2.56–4.70)	3.58 (2.89–4.38)
P-value	<0.001		0.32		0.23	