Radiology

ORIGINAL RESEARCH • CARDIAC IMAGING

Deep Learning–based Method for Fully Automatic Quantification of Left Ventricle Function from Cine MR Images: A Multivendor, Multicenter Study

Qian Tao, PhD • Wenjun Yan, BSc • Yuanyuan Wang, PhD • Elisabeth H. M. Paiman, MD • Denis P. Shamonin, MSc • Pankaj Garg, MD, PhD • Sven Plein, MD, PhD • Lu Huang, MD, PhD • Liming Xia, MD, PhD • Marek Sramko, MD, PhD • Jarsolav Tintera, PhD • Albert de Roos, MD, PhD • Hildo J. Lamb, MD, PhD • Rob J. van der Geest, PhD

From the Department of Radiology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, the Netherlands (Q.T., E.H.M.P., D.P.S., A.d.R., H.J.L., R.J.v.d.G.); Department of Electrical Engineering, Fudan University, Shanghai, China (W.Y., Y.W.); Multidisciplinary Cardiovascular Research Centre & Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, Leeds, England (P.G., S.P.); Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China (L.H., L.X.); and Departments of Cardiology (M.S.) and Radiology (J.T.), Institute for Clinical and Experimental Medicine, Prague, Czech Republic. Received March 23, 2018; revision requested April 30; final revision received August 23; accepted August 27. **Address correspondence to** Q.T. (e-mail: *q.tao@lumc.nl*).

Supported by the Dutch Technology Foundation (STW project no. 12899).

Conflicts of interest are listed at the end of this article.

See also the editorial by Colletti in this issue.

Radiology 2019; 290:81–88 • https://doi.org/10.1148/radiol.2018180513 • Content codes: CA IN MR

Purpose: To develop a deep learning–based method for fully automated quantification of left ventricular (LV) function from short-axis cine MR images and to evaluate its performance in a multivendor and multicenter setting.

Materials and Methods: This retrospective study included cine MRI data sets obtained from three major MRI vendors in four medical centers from 2008 to 2016. Three convolutional neural networks (CNNs) with the U-NET architecture were trained on data sets of increasing variability: *(a)* a single-vendor, single-center, homogeneous cohort of 100 patients (CNN1); *(b)* a single-vendor, multicenter, heterogeneous cohort of 200 patients (CNN2); and *(c)* a multivendor, multicenter, heterogeneous cohort of 400 patients (CNN3). All CNNs were tested on an independent multivendor, multicenter data set of 196 patients. CNN performance was evaluated with respect to the manual annotations from three experienced observers in terms of *(a)* LV detection accuracy, *(b)* LV segmentation accuracy, and *(c)* LV functional parameter accuracy. Automatic and manual results were compared with the paired Wilcoxon test, Pearson correlation, and Bland-Altman analysis.

Results: CNN3 achieved the highest performance on the independent testing data set. The average perpendicular distance compared with manual analysis was 1.1 mm ± 0.3 for CNN3, compared with 1.5 mm ± 1.0 for CNN1 ($P < .05$) and 1.3 mm ± 0.6 for CNN2 ($P < .05$). The LV function parameters derived from CNN3 showed a high correlation ($r^2 \geq 0.98$) and agreement with those obtained by experts for data sets from different vendors and centers.

Conclusion: A deep learning–based method trained on a data set with high variability can achieve fully automated and accurate cine MRI analysis on multivendor, multicenter cine MRI data.

Accurate quantification of left ventricular (LV) function, including end-diastolic volume, end-systolic volume, LV mass, and LV ejection fraction (LVEF), is highly clinically relevant for the diagnosis and prognostication of cardiovascular disease (1–3). Cine MRI is the current reference standard for LV function assessment and is routinely performed as part of cardiovascular MRI examinations (4,5). Multisection short-axis cine imaging covering the entire LV is typically used to derive quantitative parameters of LV function. Quantification of LV function requires careful review of and manual segmentation on many individual images, a time-consuming task hampered by variations in image quality and observer education and experience (6).

Automated or semiautomated computer methods for analyzing cine MR images have been in development for the past decades. Most methods relied on individual images or a limited training set and tried to identify the LV border on the basis of local edges, sometimes combined with shape models (7–11). In practice, the analysis involves two steps: LV detection, to detect the presence of LV, and LV segmentation, to identify the border of LV myocardium. The substantial variations across cine images (different patients, vendors, and centers) make it very challenging for traditional methods to reach a clinically acceptable balance of accuracy and generalizability for both tasks (ie, method developed on one data set can work with an "unseen" data set). As such, in current practice, the analysis of cine images still involves substantial manual input on computers or workstations, including contour tracing, and initialization and correction to aid semiautomated computer methods.

With the rapid development of deep learning techniques, convolutional neural networks (CNNs) are

Abbreviations

CNN = convolutional neural network, LV = left ventricle, LVEF = LV ejection fraction

Summary

A deep learning–based algorithm can achieve fully automated, accurate left ventricle segmentation from short-axis cine MR images when tested against a multivendor, multicenter data set.

Implications for Patient Care

- Deep learning–based image segmentation enables fast and accurate left ventricle function assessment from short-axis cine MR images without requiring any user interaction.
- Integration of the developed deep learning method as a fully automated in-line processing module would allow faster reporting of cardiac MRI studies, without the delay caused by time-consuming manual image analysis.

pushing the limit of automated image analysis to an unprecedented level (12–14). In this work, we sought to develop and evaluate a deep learning–based method that can achieve fully automatic detection and segmentation of the LV myocardium from cine MR images. We hypothesized that, with a large set of data in training covering sufficient variability, a CNN may address simultaneously the accuracy and generalizability issue. Therefore, we tested the hypothesis on cine MRI data acquired with multiple MR machines from multiple vendors, in multiple centers, and from a heterogeneous cohort of patients with cardiovascular disease.

Materials and Methods

Cine MRI Data

For training and testing our algorithm, cine MRI data from three major MR vendors were collected from four medical centers (Leiden University Medical Center, Leiden, the Netherlands; University of Leeds, Leeds, England; Tongji Hospital, Wuhan, China; Institute for Clinical and Experimental Medicine, Prague, Czech Republic). For our retrospective study, in three of the four centers (University of Leeds, Institute for Clinical and Experimental Medicine, and Tongji Hospital), the use of MRI data was approved by the local institutional review board and written informed consent was obtained from all patients. In one of the four centers (Leiden University Medical Center), the local institutional review board approved the use of MRI data but waived the requirement to obtain written informed consent.

In total, 596 cine MRI data sets were included. Data sets were acquired from 2008 to 2016. We randomly divided each data set into training and testing sets in a ratio of 2:1. The details about the data sets are shown in Table 1. All data used for our study were strictly anonymized.

Manual Annotation

Cine images in the data sets were analyzed by three observers (E.H.M.P., M.S., and R.J.v.d.G., with 4, 7, and 20 years of experience in cardiac MRI, respectively). All annotations followed the same protocol: Papillary muscles and trabeculations were included in the LV blood pool, and the most basal section was defined as the section that showed myocardial tissue for at least 50% of the circumference. All annotations were finally reviewed by an experienced observer (R.J.v.d.G.).

Manual annotation was performed by using software (MASS, research version 2017; Leiden University Medical Center, Leiden, the Netherlands). In partial annotation, the observers traced the endocardial and epicardial contours of the LV at the end-diastolic and end-systolic phase; in full annotation, contours in the end-diastolic and end-systolic phases were first manually annotated and then automatically propagated to all other cardiac phases. The observers then reviewed all phases and made corrections whenever necessary.

Training the CNNs

An established CNN architecture, called U-NET (15), was used to train the cine MRI analysis network. Figure 1 illustrates the entire system, including preprocessing, U-NET, and postprocessing. Preprocessing involved *(a)* normalization of the in-plane resolution to 2 mm; *(b)* cropping of the image to a uniform field of view of 256 × 256 mm, centered to the original image; and *(c)* normalization of the signal intensity to a uniform range of 0–256. Postprocessing involved converting the resulting label map to contour in the original image resolution and applying Fourier smoothing to reduce the pixel effect (16).

To investigate how data impact CNN performance, we trained the network on three training data sets with increasing variability. CNN1 was trained on a single-vendor, single-center, homogeneous cohort, namely, data from 100 patients with chronic myocardial infarction (training data set 1, total number of images with label: 25 389). CNN2 was trained on a single-vendor, multicenter, heterogeneous cohort including 150 patients with chronic myocardial infarction and 50 with acute myocardial infarction (training data sets 1 and 2, total number of images: 27 488). CNN3 was trained on a multivendor, multicenter, heterogeneous cohort of 400 patients (training data sets 1, 2, 3, and 4; total number of images: 41 593).

The label image for training consists of three labels based on expert annotation: the blood pool, myocardium, and background. In images beyond the LV base and apex, a background label was assigned. These images were also included in the training, such that the resulting CNN can perform detection and segmentation simultaneously.

Testing the CNNs

After the three CNNs were trained, they were tested on the same independent testing data set: the 196 cine MR acquisitions from all vendors and centers (Table 1). Manual annotation was used as the reference standard to evaluate LV segmentation accuracy. To evaluate the LV detection performance, in a subset of 100 cine data, fully automated annotation was performed; that is, the contours were traced in all images where LV myocardium was present.

Table 1: Specifications of the Multicenter, Multivendor Cine MRI Data Sets

Data Set No.	Training Data Set	Testing Data Set	Cohort	MRI Unit	Typical Parameters
1 (*n* = 150)	100 patients, 21 624 annotated images, 3765 background images	50 patients, 10 694 annotated images	Chronic myocardial infarction	1.5-T Intera (Philips Medical Systems, Best, the Netherlands)	Field of view, 400 × 320 mm^2; imaging matrix, 256 × 206; reconstructed resolution, 1.68 × 1.68 mm; section thickness, 10 mm; no section gap; flip angle, 35°; echo time, 1.71 msec; repetition time, 3.41 msec; number of phases, 35
2 (*n* = 150)	100 patients, 1815 annotated images, 284 background images	50 patients, 812 annotated images	50% acute myocardial infarction, 50% chronic myocardial infarction	1.5-T Ingenia (Philips Medical Systems)	Field of view, 400 × 400 mm^2; imaging matrix, 228 × 198; reconstructed resolution, 1.19 × 1.19 mm; section thickness, 8 mm; section gap, 2 mm; flip angle, 60°; echo time, 1.46 msec; repetition time, 2.92 msec; number of phases, 30
3 (*n* = 146)	100 patients, 9529 annotated images,1328 background images	46 patients, 4453 annotated images	Dilated cardiomyopathy	1.5-T Avanto (Siemens Medical Solutions, Erlangen, Germany)	Field of view, 300 × 300 mm^2; imaging matrix, 192 × 144; reconstructed resolution, 1.56 × 1.56 mm; section thickness, 8 mm; no section gap; flip angle, 70°; echo time, 1.16 msec; repetition time, 65.52 msec; number of phases, 24
4 (*n* = 150)	100 patients, 2610 annotated images, 638 background images	50 patients, 1639 annotated images	33% hypotrophy, 15% dilated cardiomyopathy, 15% no cardiovascular disease, 11% ischemic cardiomyopathy, 7% pulmonary hypertension, 9% other	1.5-T HDxt (*n* = 135), 3.0-T Discovery (*n* = 15) (GE Medical Systems, Waukesha, Wis)	Field of view, 380 × 380 mm^2; imaging matrix, 224 × 224; reconstructed resolution, 1.40 × 1.40 mm; section thickness, 8 mm; no section gap; flip angle, 45°; echo time, 1.55 msec; repetition time, 3.47 msec; number of phases, 20

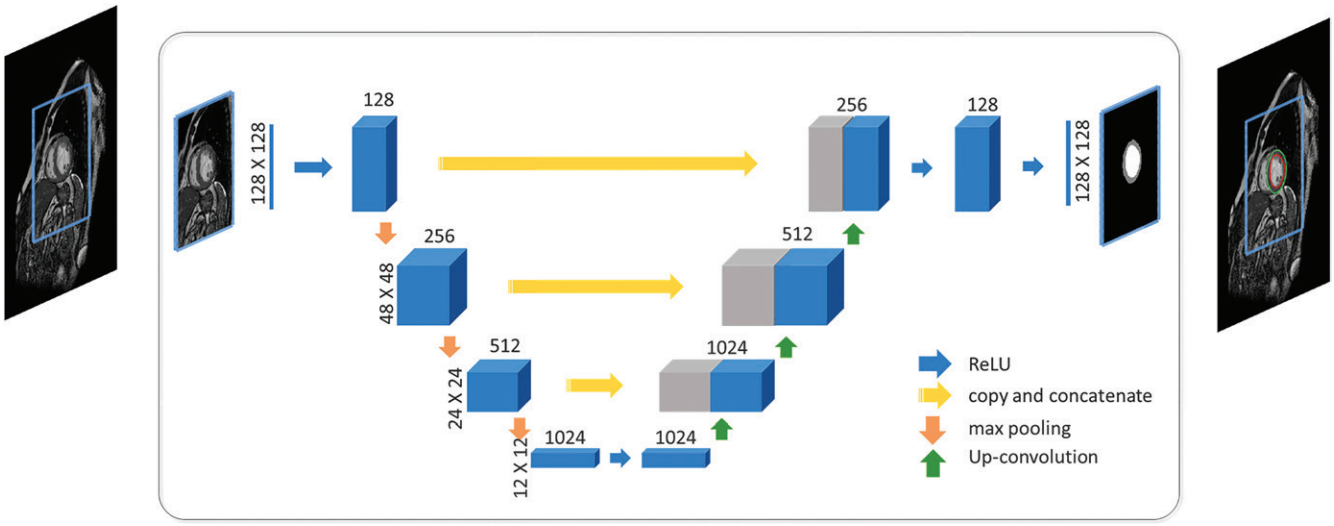

Figure 1: Diagram shows complete workflow of left ventricular (LV) segmentation network. Each blue box corresponds to a convolutional layer, with matrix size (from 128 to 12) and depth (from 1 to 1024) indicated at each layer. Colored arrows indicate different operations: blue = rectified linear unit (*ReLU*), yellow = copy and concatenate from downsampling path to upsampling path, orange = maximum pooling, and green = up-convolution. The input is any single image from a cine MRI acquisition, whereas the output is the endocardial and epicardial contours if LV is detected on the image.

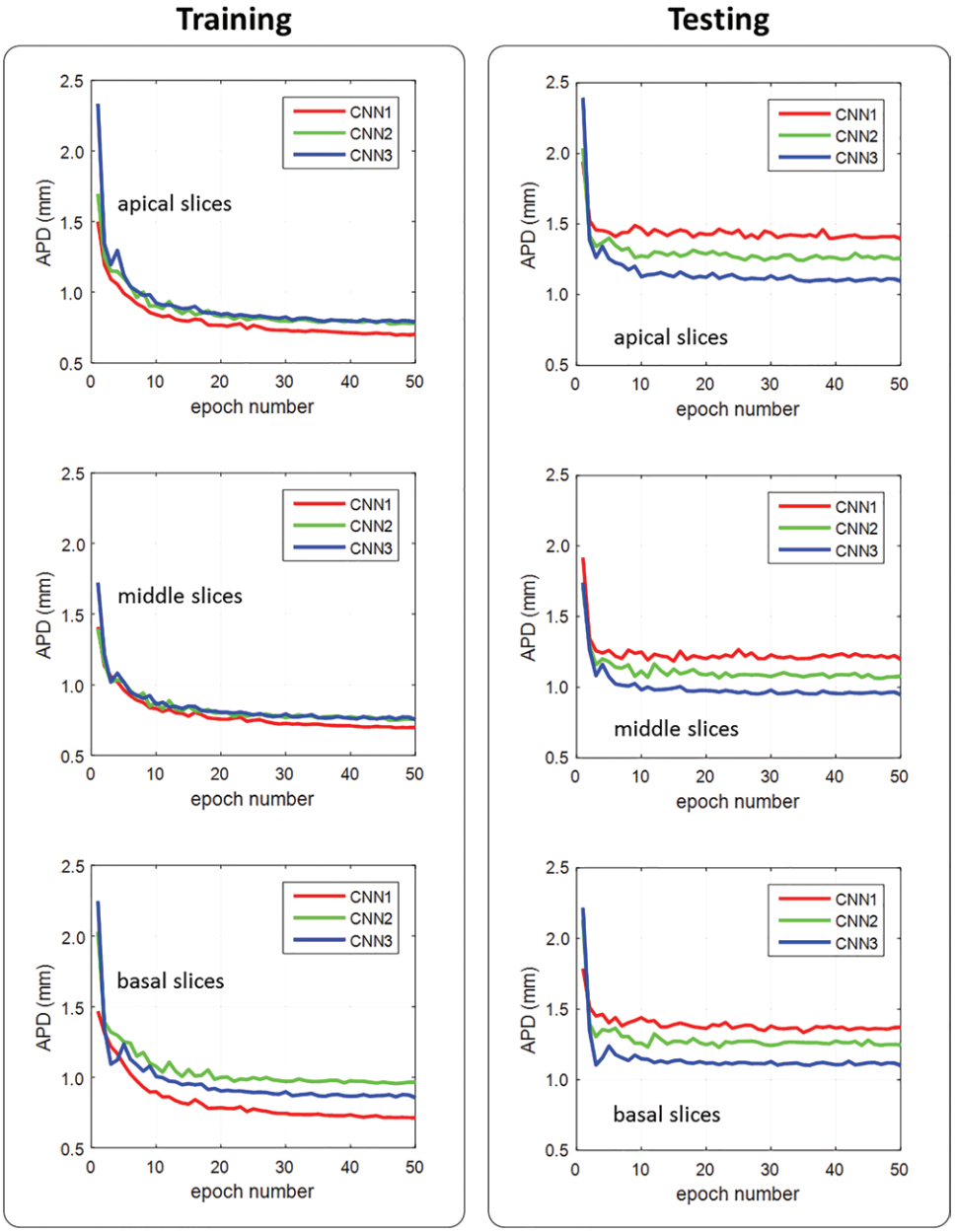

Figure 2: Performance of automated left ventricular (LV) endocardium segmentation in terms of average perpendicular distance (*APD*) relative to contours drawn by experts at apical, middle, and basal parts of LV. Left column shows performance on training data set; right column shows performance on independent testing data set. *CNN1* = convolutional neural network (CNN) trained on single-vendor, single-center, homogeneous data set; *CNN2* = CNN trained on single-vendor, multicenter, heterogeneous data set; *CNN3* = CNN trained on multivendor, multicenter, heterogeneous data set.

The CNNs were trained and tested on the Google TensorFlow platform (17) with use of a specialized graphic processing unit (GeForce GTX 1080; Nvidia, Santa Clara, Calif). Fifty epochs of training were performed.

Evaluation Criteria

The performance of the trained CNNs was evaluated in three aspects: detection, segmentation, and accuracy of clinical parameters.

Detection was defined as the accuracy of detecting the LV in a fully automated fashion, without any user intervention. This was evaluated by determining the ratio of falsely detected and falsely rejected LV in all cine images of a patient, in the subset with full annotations (n = 100).

If the LV was detected, the accuracy of LV segmentation was quantified by the average perpendicular distance between contours determined by the CNN and contours determined manually (in millimeters) and the Dice index (the ratio of two overlapping areas relative to their mean) of the segmented endocardial and epicardial areas. Both metrics were evaluated on a per-patient basis (n = 196) by using expert contours as the reference standard. We differentiated the basal, middle, and apical parts of the LV (each as approximately one-third of the imaging planes).

The accuracies of clinical parameters derived from the automatic detection and segmentation per patient (n = 196), namely, end-diastolic volume, end-systolic volume, LV mass, and LVEF, were calculated.

Statistical Analysis

Continuous variables are expressed as means ± standard deviations. Paired variables were compared with the Wilcoxon signed-rank test, without assuming the underlying distribution. $P < .05$ was considered indicative of a statistically significant difference. The Pearson correlation coefficient r^2 was used to evaluate the correlation between variables. Bland-Altman analysis was performed to compare the four LV function parameters. Both evaluation and statistical analysis were performed in the Matlab environment (R2015b; MathWorks, Natick, Mass).

Results

LV Detection Accuracy

We evaluated the detection performance of three networks per patient in 100 patients with fully automatic annotation. The detection error included *(a)* false acceptance, which usu-

Table 2: Accuracy of Left Ventricle Segmentation with Manual and Automatic Contours

	CNN1		CNN2		CNN3	
Segment	Endocardium	Epicardium	Endocardium	Epicardium	Endocardium	Epicardium
Apex	82 ± 16	83 ± 13	85 ± 13	88 ± 11	88 ± 9	91 ± 8
Middle	90 ± 11	91 ± 7	92 ± 7	93 ± 5	95 ± 2	96 ± 2
Base	88 ± 8	89 ± 7	90 ± 8	92 ± 9	93 ± 6	94 ± 5

Note.—Data are mean Dice indexes (ratio of two overlapping areas relative to their mean) ± standard deviations (in percentages). CNN1 = convolutional neural network (CNN) trained with a single-vendor, single-center, homogeneous cohort; CNN2 = CNN trained with a single-vendor, multicenter, heterogeneous cohort; CNN3 = CNN trained by multivendor, multicenter, heterogeneous cohort.

ally occurred in basal sections where the ventricular myocardium was out of plane while the CNN still located a tiny rim, and *(b)* false rejection, which usually occurred in apical sections where the CNN failed to detect the myocardium. The mean false acceptance rates were 6.1% ± 4.9, 4.9% ± 3.8, and 0.4% ± 1.2 for CNN1, CNN2, and CNN3, respectively. The mean false rejection rates were 2.0% ± 3.2, 2.0% ± 3.1, and 0.8% ± 0.8 for CNN1, CNN2, and CNN3, respectively.

LV Segmentation Accuracy

The LV segmentation accuracy was evaluated in images where both manual and automatic segmentation were present. Figure 2 shows the average perpendicular distance after each epoch of CNN training for both the training and testing data sets. The performance of all three CNNs is plotted for comparison, and performance in different LV parts is shown. After 50 epochs of training, the average perpendicular distance compared with manual analysis was 1.1 mm ± 0.3 for CNN3, compared with 1.5 mm ± 1.0 for CNN1 ($P < .05$) and 1.3 mm ± 0.6 for CNN2 ($P < .05$). Table 2 reports the Dice overlap index after 50 epochs of training. Figures 3 and 4 show a heterogeneous set of examples of fully automated LV segmentation in the testing set by using the deep learning–based method.

LV Function Parameters

For CNN3, we evaluated four automatically derived parameters with reference to manually derived ones, namely, end-diastolic volume, end-systolic volume, LV mass, and LVEF ($n = 196$). Figure 5 shows that all four parameters were highly correlated to manually derived results ($r^2 = 0.99$, 0.99, 0.98, and 0.98, respectively). The limit of agreement between the CNN and manual results was −18 to 10 mL for end-diastolic volume, −17 to 12 mL for end-systolic volume, −17 to 13 mL for LV mass, and −5.3% to 4.8% for LVEF. The Wilcoxon signed-rank test showed that the CNN underestimated end-diastolic volume, end-systolic volume, and LV mass ($P < .001$), whereas LVEF was not significantly different from the manual results ($P = .74$).

Computation Time

The time of training was linearly dependent on the number of training images and the number of epochs. Each epoch took an average of 30 minutes. Figure 2 shows that the testing performance stabilized after 10 epochs of training. Once trained, the CNN can be applied to an unseen cine MRI data set, with fully automatic processing of a complete cine MRI data set (with approximately 300 images) in about 1 second.

Discussion

We developed a fully automatic cine MRI analysis system based on CNN and evaluated it in an extensive multivendor, multicenter, heterogeneous cohort. Once trained with data of sufficient variability, the system can produce fast, accurate, and fully automated LV detection and segmentation for a wide range of cine MRI inputs.

Our results demonstrated that, when the CNN used training data sets of increasing variability, the performance also improved in an unseen data set. CNN3 had the largest number of training samples and highest variability, and it achieved the best performance on an unseen heterogeneous testing data set. The final average perpendicular distance of approximately 1 mm was comparable to the intra- and interobserver variability (7,18,19) and largely surpassed previous machine learning methods (approximately 2 mm) (8–10). False rejection and false acceptance occurred mostly in apical and basal sections, where it was also most difficult for human observers to define LV. With the relatively small size of the LV there, however, the error did not have a significant impact on functional parameter estimations, especially LVEF.

We used an incremental CNN training strategy and evaluated its benefits. The increasing performance of the CNN with increasing training samples suggests that such a network may be constantly refreshed and strengthened with data. This can be done in practice by using the previous CNN as an initialization to train an updated network when new data with annotations are available, resembling a constant human learning process. Yet, further study is warranted to establish an efficient way to feed the data, regulate the training process, and find the limit.

A limitation of the work is that all training and testing was performed with retrospective data, whereas its prospective clinical use in clinic still must be evaluated, especially on data sets with a broader spectrum of cardiovascular abnormalities and imaging artifacts. Furthermore, as three

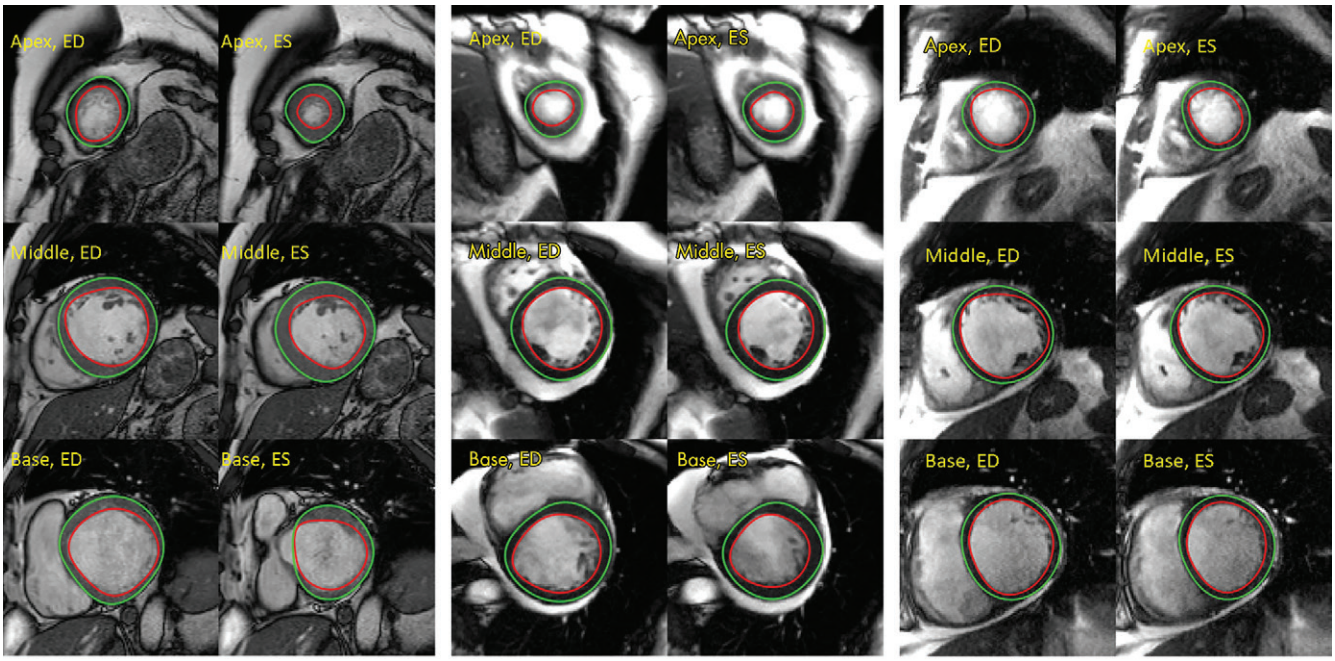

Figure 3: Examples of automated left ventricular segmentation from convolutional neural network. Apical, middle, and basal sections are shown at end-diastolic (*ED*) and end-systolic (*ES*) phases. **(a)** Data set 1. Images obtained at 1.5 T (Intera; Philips Medical Systems, Best, the Netherlands) in patient with ischemic cardiomyopathy. **(b)** Data set 2. Images obtained at 1.5 T (Ingenia, Philips) in patient with ischemic cardiomyopathy. **(c)** Data set 3. Images obtained at 1.5 T (Avanto; Siemens Medical Solutions, Erlangen, Germany) in patient with dilated cardiomyopathy.

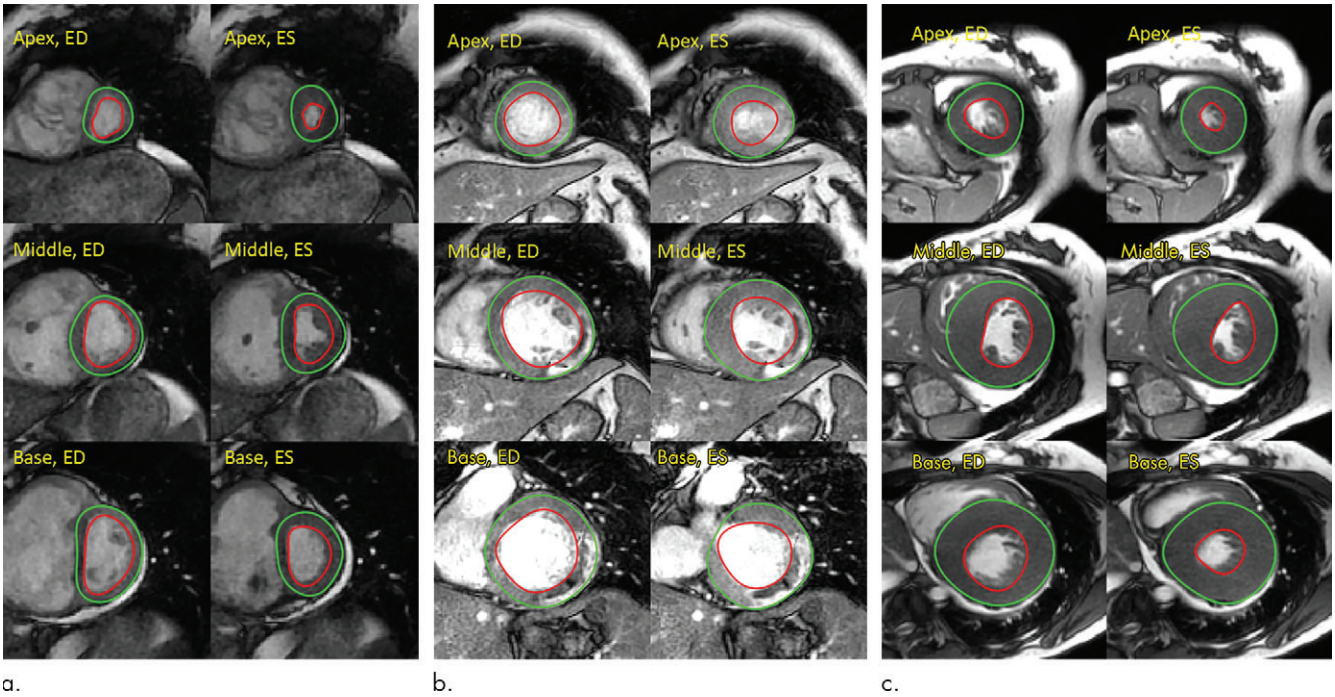

Figure 4: Examples of automated left ventricular segmentation from convolutional neural network. Six images are shown for each example. Apical, middle, and basal sections are shown at end-diastolic (*ED*) and end-systolic (*ES*) phases. **(a)** Data set 4. Image obtained at 1.5 T (HDxt; GE Medical Systems, Waukesha, Wis) in patient with pulmonary hypertension. **(b)** Data set 4. Images obtained at 1.5 T (HDxt, GE Medical Systems) in patient with ischemic cardiomyopathy after intravenous administration of gadolinium chelate. **(c)** Data set 4. Images obtained at 3.0 T (Discovery, GE Medical Systems) in patient with hypertrophic cardiomyopathy.

observers annotated both the training and testing set, the performance of the CNN with respect to interobserver variability was not thoroughly evaluated.

In conclusion, we developed a fully automated cine MRI analysis system based on a CNN and evaluated it in a real-world scenario by using multivendor, multicenter, heterogeneous

data from patients with cardiovascular disease. After being trained on a data set of sufficient variability, the deep-learning CNN was able to accurately detect and segment the LV in real time and to generate highly correlated volumetric measurements and accurate LVEFs compared with those from experienced observers, free of any user intervention.

Author contributions: Guarantors of integrity of entire study, Q.T., W.Y., Y.W.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, Q.T., W.Y., P.G., H.J.L., R.J.v.d.G.; clinical studies, W.Y., D.P.S., P.G., L.H., L.X., M.S., J.T., H.J.L.; statistical analysis, Q.T., W.Y., Y.W.; and manuscript editing, Q.T., W.Y., Y.W., D.P.S., P.G., S.P., L.H., L.X., M.S., A.d.R., H.J.L., R.J.v.d.G.

Disclosures of Conflicts of Interest: **Q.T.** disclosed no relevant relationships. **W.Y.** disclosed no relevant relationships. **Y.W.** disclosed no relevant relationships. **E.H.M.P.** disclosed no relevant relationships. **D.P.S.** disclosed no relevant relationships. **P.G.** disclosed no relevant relationships. **S.P.** disclosed no relevant relationships. **L.H.** disclosed no relevant relationships. **L.X.** disclosed no relevant relationships. **M.S.** disclosed no relevant relationships. **J.T.** disclosed no relevant relationships. **A.d.R.** disclosed no relevant relationships. **H.J.L.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: has grants/grants pending with Novo Nordisk; received payment for lectures including service on speakers bureaus from Philips Healthtech and Toshiba Medical Systems; received payment for development of educational presentations from Guerbet. Other relationships: disclosed no relevant relationships. **R.J.v.d.G.** disclosed no relevant relationships.

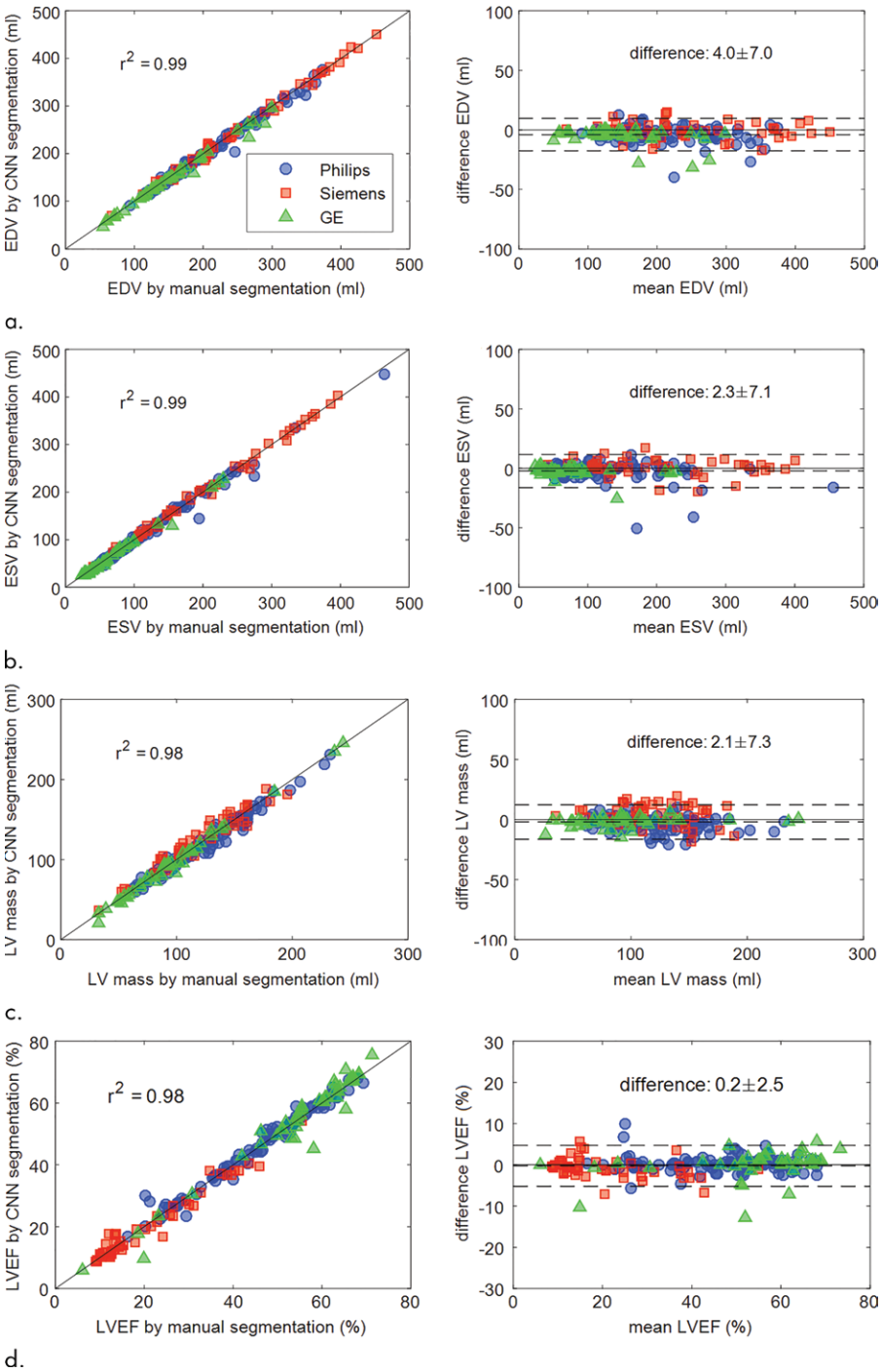

Figure 5: Correlation (left) and Bland-Altman (right) plots of functional parameters generated from manual and convolutional neural network (CNN) contours. **(a)** Left ventricular (LV) end-diastolic volume (*EDV*), **(b)** LV end-systolic volume (*ESV*), **(c)** LV mass, and **(d)** LV ejection fraction (*LVEF*). Different colors and shapes denote data from different vendors. GE = GE Medical Systems, Waukesha, Wis; Philips = Philips Medical Systems, Best, the Netherlands; Siemens = Siemens Medical Solutions, Erlangen, Germany.

References

1. Aljaroudi W, Alraies MC, Halley C, et al. Impact of progression of diastolic dysfunction on mortality in patients with normal ejection fraction. Circulation 2012;125(6):782–788.
2. Flachskampf FA, Biering-Sørensen T, Solomon SD, Duvernoy O, Bjerner T, Smiseth OA. Cardiac imaging to evaluate left ventricular diastolic function. JACC Cardiovasc Imaging 2015;8(9):1071–1093.
3. Yeboah J, Rodriguez CJ, Stacey B, et al. Prognosis of individuals with asymptomatic left ventricular systolic dysfunction in the multi-ethnic study of atherosclerosis (MESA). Circulation 2012;126(23):2713–2719.
4. Jerosch-Herold M, Kwong RY. Magnetic resonance imaging in the assessment of ventricular remodeling and viability. Curr Heart Fail Rep 2008;5(1):5–10.
5. de Roos A, Higgins CB. Cardiac radiology: centenary review. Radiology 2014;273(2 Suppl):S142–S159.
6. Sardanelli F, Quarenghi M, Di Leo G, Boccaccini L, Schiavi A. Segmentation of cardiac cine MR images of left and right ventricles: interactive semiautomated methods and manual contouring by two readers with different education and experience. J Magn Reson Imaging 2008;27(4):785–792.
7. Suinesiaputra A, Bluemke DA, Cowan BR, et al. Quantification of LV function and mass by cardiovascular magnetic resonance: multi-center variability and consensus contours. J Cardiovasc Magn Reson 2015;17(1):63.
8. Mitchell SC, Lelieveldt BP, van der Geest RJ, Bosch HG, Reiber JH, Sonka M. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. IEEE Trans Med Imaging 2001;20(5):415–423.

9. Nambakhsh CMS, Yuan J, Punithakumar K, et al. Left ventricle segmentation in MRI via convex relaxed distribution matching. Med Image Anal 2013;17(8):1010–1024.
10. Frangi AF, Niessen WJ, Viergever MA. Three-dimensional modeling for functional analysis of cardiac images: a review. IEEE Trans Med Imaging 2001;20(1):2–25.
11. Kaus MR, von Berg J, Weese J, Niessen W, Pekar V. Automated segmentation of the left ventricle in cardiac MRI. Med Image Anal 2004;8(3):245–254.
12. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–2324.
13. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.
14. Dreyer KJ, Geis JR. When machines think: radiology's next frontier. Radiology 2017;285(3):713–718.
15. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. Medical image computing and computer-assisted intervention – MICCAI 2015. MICCAI 2015. Lecture notes in computer science, vol 9351. Cham, Switzerland: Springer, 2015; 234–241.
16. Staib LH, Duncan JS. Boundary finding with parametrically deformable models. IEEE Trans Pattern Anal Mach Intell 1992;14(11):1061–1075.
17. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. https://arxiv.org/abs/1603.04467. Published online March 14, 2016.
18. Tao Q, Piers SRD, Lamb HJ, van der Geest RJ. Automated left ventricle segmentation in late gadolinium-enhanced MRI for objective myocardial scar assessment. J Magn Reson Imaging 2015;42(2):390–399.
19. Suinesiaputra A, Cowan BR, Al-Agamy AO, et al. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. Med Image Anal 2014;18(1):50–62.