Computer Methods and Programs in Biomedicine xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Simulator-generated training datasets as an alternative to using patient data for machine learning: An example in myocardial segmentation with MRI

Christos G. Xanthis[a,b,*], Dimitrios Filos[a], Kostas Haris[a], Anthony H. Aletras[a,b]

[a] Laboratory of Computing, Medical Informatics and Biomedical-Imaging Technologies, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki, Greece
[b] Department of Clinical Physiology, Clinical Sciences, Lund University and Lund University Hospital, Lund, Sweden

ARTICLE INFO

Article history:
Received 3 June 2020
Accepted 21 October 2020
Available online xxx

Keywords:
Magnetic resonance imaging
simulation
machine learning
supervised techniques
segmentation

ABSTRACT

Background and Objective: Supervised Machine Learning techniques have shown significant potential in medical image analysis. However, the training data that need to be collected for these techniques in the field of MRI 1) may not be available, 2) may be available but the size is small, 3) may be available but not representative and 4) may be available but with weak labels. The aim of this study was to overcome these limitations through advanced MR simulations on a realistic computer model of human anatomy without using a real MRI scanner, without scanning patients and without having personnel and the associated expenses.

Methods: The 4D-XCAT model was used with the coreMRI simulation platform for generating artificial short-axis MR-images for training a neural-network to automatic delineate the LV endocardium and epicardium. Its performance was assessed on real MR-images acquired from eight healthy volunteers. The neural-network was also trained on real MR-images from a publicly available dataset and its performance was assessed on the same volunteers' data.

Results: The proposed solution demonstrated a performance of 94% (endocardium) and 90% DICE (epicardium) in real mid-ventricular slices, whereas a 10% addition of real MR-images in the artificial training dataset increased the performance to 97% DICE. The use of artificial MR-images that cover the entire LV yielded 85% (endocardium) and 88% DICE (epicardium) when combined with real MR data with an 80%-20% mix respectively.

Conclusions: This study suggests a low-cost solution for constructing artificial training datasets for supervised learning techniques in the field of MR by using advanced MR simulations without the use of a real MRI scanner, without scanning patients and without having to use specialized personnel, such as technologists and radiologists.

1. Introduction

Machine Learning (ML) has matured and can be used for solving several practical problems in medicine. In the field of radiology, ML has proven to be an effective tool to automate the analysis and diagnosis for medical images [1]. Several ML approaches have been proposed today, which not only assist in image-based diagnosis but also in disease prognosis and risk assessment [2]. Supervised ML techniques have shown significant potential in medical image analysis and can potentially reduce the burden on radiologists in daily practice [3–5].

However, as for any other system that learns from training datasets, the training data that need to be collected for an adequate generalization is a key issue. This issue becomes even more important in the field of medical image analysis and particularly in radiology since the training data sets 1) may not be available, 2) may be available but of small size (small data sets), 3) may be available but not representative and 4) may be available but with weak labels (non-well-annotated data sets).

In the past, the availability of large datasets with diagnostic labels has altered the course of research not only in the field of disease diagnosis but also in the field of disease prognosis [2].

* Corresponding author.
E-mail addresses: cxanthis@gmail.com (C.G. Xanthis), dimfilos@auth.gr (D. Filos), konharis@gmail.com (K. Haris), aletras@hotmail.com (A.H. Aletras).

https://doi.org/10.1016/j.cmpb.2020.105817
0169-2607/

However, the availability of only small sized labelled training datasets does not often allow the development and improvement of supervised ML techniques [6]. If large enough training datasets are not available, methods such as virtual sample generation [7] and data augmentation have previously been proposed [8] to compensate.

Other remaining challenges that further hinder the application of ML techniques in medical imaging are the existence of a) non-representative as well as of b) non-annotated training data sets. In the first case, supervised ML techniques may achieve high performance when both training and test data sets are subgroups of the same image distribution, but they will exhibit low performance when the test images have been acquired from a different imaging protocol (such as different scanner, different imaging center, different technologist, different patient population characteristics, etc.). In the second case of non-annotated training data, most of the supervised ML techniques require the availability of training images where experts have previously designed the necessary contours and labelled the images. However, performing this task is not always easy since the quality of the available images may not be suitable due to various reasons (such as underlying disease, image artifacts, limited spatial and temporal resolution of images, etc.) or because there is a lack of sufficiently specialized personnel. Moreover, even when appropriately trained personnel is available, visual assessment and manual annotation remain subjective and may deviate from reality.

In the field of Magnetic Resonance Imaging (MRI), the process of creating training datasets is considered costly, both in time and money. Creating training datasets requires not only acquisition of medical images that cover a broad spectrum of medical parameters but also manual annotation of the medical images by experts. In addition, generating the training dataset requires the use of an MRI scanner and the availability of personnel (such as technologists and radiologists). Moreover, the complexity of the underlying MRI physics hinders this process even more. Furthermore, image acquisition is performed under a specific MRI protocol and MRI system configuration. While supervised learning techniques have demonstrated good performance on relatively controlled experiments with standardized imaging protocols, their performance may deteriorate in cases where images are acquired with a different imaging protocol or on a different MRI system [2]. In addition to this, acquisition of data that cover the entire range of parameters that describe the anatomy and physiology of the population is impractical. Last, despite that several other approaches have recently been proposed in the field of MR image synthesis, they may be based on a kspace-corruption approach for the generation of synthetic MR data [9,10] that are “not entirely representative” for a wide range of real cases, and/or require the availability of thousands of real MR images [11]. The latter is a quite common approach in the field of medical image synthesis [12].

The specific aim of this study was to propose a solution to these limitations encountered today in generating training datasets for supervised ML techniques in the field of Magnetic Resonance Imaging. We hypothesized that this can be achieved without the use of a real MRI scanner, without scanning patients and without having personnel (such as technologists and radiologists), but through the application of advanced MR simulations on a realistic computer model of human anatomy. This study is a proof of concept that artificial MR-images can be used as training datasets in supervised learning approaches in the MR field. For this purpose, an example in myocardial segmentation of end-diastolic images with MRI and existing neural-networks was utilized. Since this work is not aimed per se at presenting LV segmentation methods and due to space considerations, end-systolic examples are not shown. This work is only the initial step to a wide range of potential applications (e.g. tissue characterization, image reconstruction, etc.) with or without various pathologies that will have to be developed and validated.

2. Materials and methods

2.1. Computer model

The 4D-XCAT anatomical model [13] was utilized as the main, highly detailed, computer model that represents the whole-body human anatomy. The 4D-XCAT model was modified to become compatible with MR simulations by assigning magnetic properties (T1, T2 and PD) to all the 78 available tissue types at 1.5T.

The long axis of the left-ventricle was identified, and eleven perpendicular-to-long-axis slices were extracted (from the apex to the base of the left ventricle with slice gap = 10 mm). For every short-axis slice, a number of rotated slices was considered, where their angle with the vertical plane ranged from 10 to -10 degrees. The tissues located at positions outside the selected slice along its vertical direction were then removed and the computer model of the slice of interest remained with only the tissues at the selected plane.

Two datasets of 120 slices each were extracted at the end-diastole of the cardiac cycle. The first dataset represented a normal anatomy of myocardial tissue whereas the second dataset represented a pathological reduced myocardial thickness at the apex of the left ventricle. Regional LV wall thinning may be associated with transmural chronic myocardial infarction and scar tissue. Both datasets included the papillary muscles within the blood pool of the left ventricle whereas the corresponding maps of the endocardium and epicardium of the left-ventricle within the slices of interest were extracted automatically from the 4D XCAT model. From the total 120 slices of each dataset, 24 slices were labelled as apical, 72 slices were labelled as mid-ventricular and 24 slices were labelled as basal.

2.2. Simulation platform

The publicly available, cloud-based, coreMRI simulation platform (www.coremri.org) [14] was utilized for the generation of the artificial MR-images. A single-shot bSSFP pulse sequence [15,16] and the computer models of the slices of interest were uploaded to the cloud through the frontend of the coreMRI platform. A more detailed description of the simulation platform can be found in the literature [14,17,18] and in the Appendix of the current study (Section “coreMRI simulation platform”).

The pulse sequence parameters were the following: the TR/TE was 2.72/1.36 ms, the bSSFP readout used a three-lobe 490 μs sinc-shaped RF pulse with 8 mm slice thickness and 63° excitation flip angle, a half-alpha preparation was used prior to the bSSFP readout, the field of view was 384(FE)x288(PE) mm, the scan matrix size was 192(FE)x144(PE), the receiver BW was 200 kHz and the acquisition was performed on a transversal slice at the isocenter of the virtual MR scanner. The simulated MR-images were generated with no noise, with Gaussian distributed noise in I and Q channels with zero mean and standard deviation equal to 25, and with Gaussian distributed noise in I and Q channels with zero mean and standard deviation equal to 40. All images were reconstructed on the Gadgetron reconstruction framework using a simple 2D Fourier-Transform (FT) MRI reconstruction as it was described by the Gadgetron configuration file default.xml [19]. The simulated MR-images were further filtered with a 2-D Gaussian smoothing kernel with standard deviation of 0.9 whereas the contrast of the output images remained unchanged or was further increased through the saturation of the bottom 1% and the top 1% of all pixel values. The simulations were performed on a p2.xlarge instance type provided by Amazon AWS equipped with 1 GPU

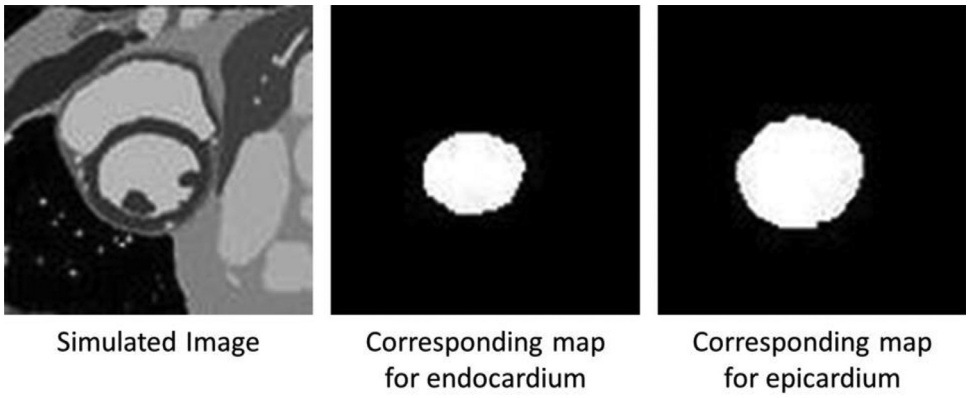

Fig. 1. Simulated short-axis image of a mid-ventricular slice of the anatomical model that represents a normal anatomy of myocardial tissue and the corresponding maps for the endocardium and epicardium as being used in the neural network.

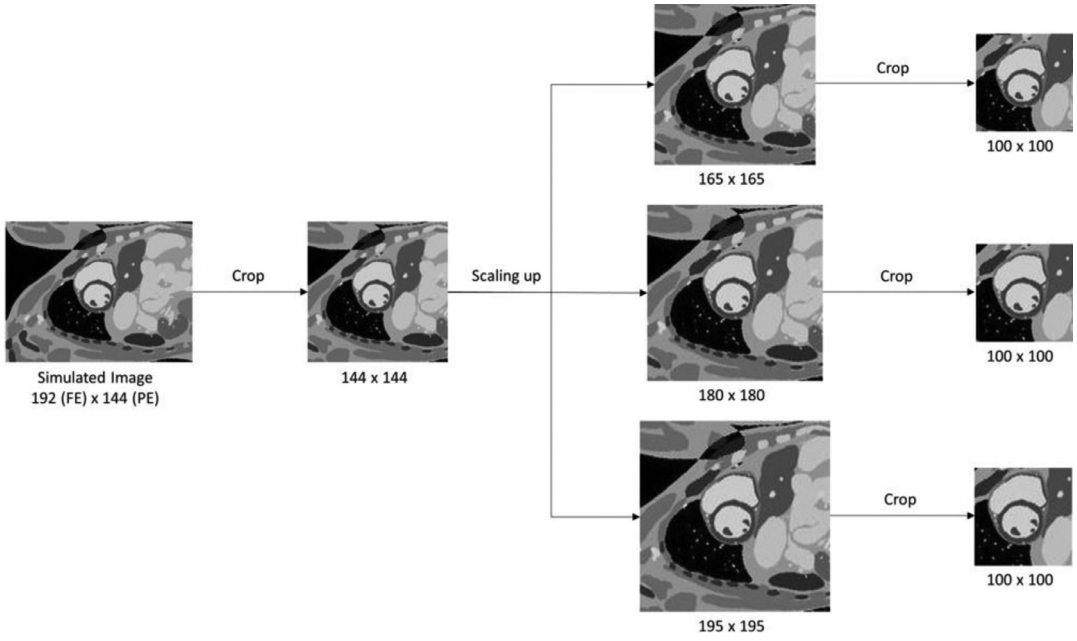

Fig. 2. The pre-processing performed on the artificial MR-images. The initial artificial MR image (192 × 144 pixels) was cropped at a size of 144 × 144 pixels and then scaled to three different sizes: 165 × 165, 180 × 180 and 195 × 195 pixels. Last, a square (100 × 100 pixels) center crop of each scaled image was performed to define the image that would be utilized as an input to the NN.

(NVIDIA-K80). The total simulation time was around 40 sec for every slice.

2.3. Neural-network

A previously developed fully convolutional neural-network (FCN) [20] was utilized in this study for LV segmentation. The FCN consisted of 15 stacked convolution layers and three layers of overlapping max pooling of stride 2. Each convolution layer was followed by a Rectified Linear Unit and a Mean-Variance normalization operation.

The artificial training dataset consisted of the simulated images of the slices of interest and the corresponding maps of the endocardium and epicardium of the left-ventricle within these slices (Fig. 1). The training dataset was used to train the FCN for LV segmentation of the endocardium and epicardium. In cases where real MR data were used for the training of the neural-network (NN), real MR-images along with the corresponding ground truth contours of endocardium and epicardium were incorporated in the training dataset.

Preprocessing was performed on all artificial MR-images that were utilized as inputs to the NN (training, validation and test datasets). In particular, the initial artificial MR-images (192 × 144 pixels) were cropped at 144 × 144 pixels and then scaled to three different sizes: 165 × 165, 180 × 180 and 195 × 195 pixels. Last, a square (100 × 100 pixels) center crop of each scaled image was performed to define the image that would be utilized as an input to the NN. Similar preprocessing was performed on the corresponding maps of the endocardium and epicardium of the left-ventricle. Fig. 2 presents the pre-processing that was performed on the artificial MR-images.

The training datasets were further augmented by geometric transformations, such as rotation, translation, scaling, vertical and horizontal flipping. Augmentation was performed in all cases (in-vivo and artificial MR images) so as to avoid overfitting, to similarly improve the performance of the NN but also to exclude augmentation as a confounding factor when comparing different methods. The number of training epochs used was 40 and the evaluation metrics used were accuracy, Dice index and Jaccard index. The SimpleITK toolkit [21] was utilized for the calculations of these

Table 1
Acquisition parameters of the CINE pulse sequences utilized for the healthy volunteers. TE: Echo Time; TR: Repetition Time; FOV: Field Of View; FE: Frequency Encoding; PE: Phase Encoding.

Volunteer	TE (ms)	TR (ms)	FOV (mm) (FE x PE)	Flip Angle (o)	Matrix Size (FE x PE)
1	1.36	2.72	384 × 264	67	192 × 132
2	1.36	2.72	384 × 228	70	192 × 114
3	1.36	2.72	384 × 336	69	192 × 168
4	1.36	2.72	384 × 240	71	192 × 120
5	1.36	2.72	384 × 228	70	192 × 114
6	1.36	2.72	384 × 288	63	192 × 144
7	1.36	2.72	384 × 312	69	192 × 156
8	1.36	2.72	384 × 324	72	192 × 162

metrics. The performance of NN was reported using the Dice index and Hausdorff Distance (HD – computed in pixels).

2.4. Real MR data

The performance of the NN was evaluated on real cardiac MR data (end-diastolic, short-axis, cardiac MR-images) that were acquired for eight [8] healthy volunteers with no medical history (4 men, 4 women, age 26±4 years). The study was approved by the regional ethics committee and all subjects provided written informed consent (The regional ethics committee, Lund, Sweden. Ethics applications numbers: 541/2004 and 815/2016). All subjects underwent cardiac MR on a MAGNETOM Aera 1.5T scanner (Siemens Healthcare, Erlangen, Germany) using an 18-channel coil. A segmented CINE (TR/TE equal to 2.72/1.36ms, slice thickness equal to 8mm, without parallel imaging) was used to acquire short axis images covering the entire volume of the left-ventricle. Table 1 summarizes the CINE parameters per volunteer.

All images were analysed using the software Segment, version 2.2R7052 (http://segment.heiberg.se) [22]. The delineations of endocardium and epicardium from end-diastolic images were performed by an experienced reviewer. The real MR dataset consisted of 99 images in total (30 apical images, 57 mid-ventricular images and 12 basal images). End-systolic, short-axis, cardiac MR-images were also acquired but were not used in this study because the aim was to demonstrate using simulated datasets instead of real datasets for training NNs rather than a new segmentation method per se.

Preprocessing was also performed on all real MR-images that were used as inputs to the NN (training, validation and test datasets). In cases where the size of the acquired MR-images was different than 192 × 144 pixels, zero-padding or cropping was performed. These images were cropped at a size of 144 × 144 pixels whereas special care was taken to ensure that the LV center was always at the center of the image. The images were then scaled at a size of 195 × 195 pixels and, last, a square 100 × 100 pixels was cropped at the center of each scaled image.

3. Experiments

3.1. Mid-ventricular slices only

In order to investigate the effect that the quality of the artificial training dataset had on NN performance, 6 different configurations were applied, as described below, for generating the artificial training dataset from the mid-ventricular slices of the first dataset only (normal anatomy of myocardial tissue) whereas the NN performance was evaluated on the healthy volunteers' mid-ventricular images only. The size of the training dataset was 216 images.

In particular, the entire dataset of the mid-ventricular images that underwent through all the different scales (165 × 165, 180 × 180 and 195 × 195 pixels, 216 total images) was initially prepared with a 2-D Gaussian smoothing kernel with a standard deviation of 0.9 and then further examined for all the combinations of noise on the simulated MR signal (no noise, Gaussian distributed noise with standard deviation equal to 25 and Gaussian distributed noise with standard deviation equal to 40) and contrast improvement (unchanged contrast vs. contrast increased through the saturation of the bottom 1% and the top 1% of all pixel values).

The six different configurations were the following:

1. Noise00 without noise and without contrast improvement
2. Noise00_ci without noise and with contrast improvement (_ci)
3. Noise25 with gaussian distributed noise with standard deviation equal to 25 and without contrast improvement
4. Noise25_ci with gaussian distributed noise with standard deviation equal to 25 and with contrast improvement (_ci)
5. Noise40 with gaussian distributed noise with standard deviation equal to 40 and without contrast improvement
6. Noise40_ci with gaussian distributed noise with standard deviation equal to 40 with contrast improvement (_ci)

Moreover, to further investigate the effect that the augmentation of the artificial training dataset had on NN performance, the aforementioned 6 different configurations were also applied for generating the artificial MR-images from the mid-ventricular slices only of the first (normal anatomy of myocardial tissue) and the second (pathological reduced myocardial thickness at the apex of the left ventricle) dataset whereas the NN performance was evaluated on the healthy volunteers' mid-ventricular images only. The size of the training dataset was 432 images

Last, real MR-images (extracted from the pool of the healthy volunteers' mid-ventricular acquisitions) were added into the artificial training dataset and the performance of the NN was reassessed on the healthy volunteers' remaining mid-ventricular images. The incorporation was performed under a 90%-10% and 80%-20% configuration, where the 90% and 80% represented the proportion of the simulated MR-images in the training dataset whereas the 10% and 20% represented the proportion of the real MR-images in the training dataset. In these experiments, the simulated MR-images were generated using all the mid-ventricular slices of the first dataset (normal anatomy of myocardial tissue) that underwent through all the different scales, Gaussian distributed noise with standard deviation equal to 25 and 40, gauss filtering with standard deviations of 0.9 and contrast improvement (unchanged contrast and contrast increased through the saturation of the bottom 1% and the top 1% of all pixel values).

The 90%-10% configuration consisted of 159 artificial MR-images and the real mid-ventricular MR-images from 3 healthy volunteers (19 MR-images) whereas the 80%-20% configuration consisted of 159 artificial MR-images and the real mid-ventricular MR-images from 6 healthy volunteers (41 MR-images). For comparison purposes, training of the NN was also performed on the 159 artificial MR-images only. The assessment of the NN performance was performed on the remaining 16 real mid-ventricular MR-images acquired from the remaining 2 healthy volunteers and remained the same across the different experiments.

Table 2
Details of the six different experiments that involved trainings on the publicly available dataset and tests on the healthy volunteer's real MR-images. LV: Left-ventricle.

	Training (publicly available dataset)	Size of training dataset	Testing (eight healthy volunteers)	Size of test dataset
Mid-ventricular slices only	Mid-ventricular only, entire population, end-diastolic images	187	Mid-ventricular MR-images	57
	Mid-ventricular only, normal subjects only, end-diastolic images	34	Mid-ventricular MR-images	57
	Mid-ventricular only, 159 randomly selected images from the entire population, end-diastolic images	159	On the remaining 18 real mid-ventricular MR-images (see 90%-10% and 80%-10% experiments for mid-ventricular images only)	18
Entire LV volume	Entire LV volume, entire population, end-diastolic images	420	Entire LV volume images	99
	Entire LV volume, normal subjects only, end-diastolic images	78	Entire LV volume images	99
	Entire LV volume, 279 randomly selected images from the entire population, end-diastolic images	279	On the remaining 28 real MR-images of the entire LV volume (see 90%-10% and 80%-10% experiments for entire LV volume)	28

3.2. All slices covering the entire LV volume

In order to assess the performance of the NN on short-axis MR-images that covered the entire LV volume, training of the NN was performed on the simulated MR-images that were generated using all slices of the first dataset (normal anatomy of myocardial tissue) that underwent 1) through all the different scales, Gaussian distributed noise with standard deviation equal to 25, gauss filtering with standard deviations of 0.9 and no contrast improvement, and 2) through all the different scales, Gaussian distributed noise with standard deviation equal to 40, gauss filtering with standard deviations of 0.9 and contrast improvement through the saturation of the bottom 1% and the top 1% of all pixel values. The size of each training dataset was 279 artificial MR-images.

Moreover, real MR-images (extracted from the pool of the healthy volunteers' acquisitions) were added into the previously-created artificial training datasets under a 90%-10% and 80%-20% configuration. The 90%-10% configuration consisted of 279 artificial MR-images and 31 real MR-images (12 apical, 12 mid-ventricular and 7 basal MR-images) whereas the 80%-20% configuration consisted of 279 artificial MR-images and 71 real MR-images (16 apical, 48 mid-ventricular and 7 basal MR-images). For comparison purposes, training of the NN was also performed on the 279 artificial MR-images only. Last, training of the NN with only the 71 real MR-images (16 apical, 48 mid-ventricular and 7 basal MR-images) from the 80%-20% configuration was performed for comparison purposes as well.

The assessment of the NN performance was performed on the remaining 28 real MR-images acquired from the healthy volunteers (14 apical, 9 mid-ventricular and 5 basal MR-images) and remained the same across the different experiments.

3.3. Publicly available dataset

One publicly available annotated dataset (Sunnybrook dataset) of short-axis MR-images [23] was utilized in this study in order to assess the performance of the NN on images that were acquired through the regular and most resource-demanding approach. The publicly available dataset consisted of 420 end-diastolic, short-axis images acquired from a mixed group (n=45, 32 men, 13 women, age 61±15 years) of normal subjects (n=9, 6 men, 3 women, age 60±18 years) and patients with hypertrophy (n=12, 7 men, 5 women, age 58±14 years), patients with heart failure with infarction (n=12, 11 men, 1 woman, age 61±12 years) and patients with heart failure without infarction (n=12, 8 men, 4 women, age 65±17 years). The images were acquired from a 1.5T GE Signa MRI system with a cine, SSFP based pulse sequence (slice thickness=8mm, slice gap=8mm, FOV=320 × 320 mm and matrix-size=256 × 256) whereas delineations of the endocardium and epicardium were drawn and confirmed by two experienced cardiologists. End-systolic, short-axis, cardiac MR-images were also available in this publicly available dataset but were not used in this study because the aim was to demonstrate using simulated datasets instead of real datasets for training NNs rather than a new segmentation method per se.

The publicly available dataset was utilized as the training dataset so as to assess the performance of the NN that was trained on a dataset acquired under a different configuration than the one used for the acquisition of the test dataset. Therefore, six different trainings on the publicly available dataset and tests on the healthy volunteer's data were performed in this study and compared against the relevant experiments that involved artificial MR-images. Table 2 presents the details of these six different experiments.

4. Results

Fig. 3 presents simulated short-axis MR-images covering the entire LV that were generated from the coreMRI simulation platform using the XCAT anatomical model and the single-shot bSSFP pulse sequence. Fig. 4 shows the same mid-ventricular, short-axis, artificial MR image for different combinations of noise on the simulated MR signal, gauss filtering and contrast improvement.

4.1. Mid-ventricular slices only

Table 3 shows the performance of the NN in terms of delineating the LV endocardium and epicardium in mid-ventricular slices acquired from 8 healthy volunteers. The NN was trained on 100% simulated images.

First, out of the 6 different trainings on normal anatomy alone, the configurations Noise25_ci and Noise40_ci (bold typeface in lines 4 and 6 of Table 3) yielded the best combined performance. The configuration Noise40_ci also yielded the best performance for delineating the LV endocardium (DICE index 94%) whereas the configuration Noise25 (line 3 in Table 3) yielded the best performance (DICE index 90%) in delineating the LV epicardium.

Second, out of the 6 different trainings on normal anatomy and pathologically reduced myocardial thickness at the apex of the left ventricle, the configuration Noise25 (line 3 at Table 3) yielded the best performance in delineating both the LV endocardium and epicardium (93% DICE index for endocardium and 90% DICE index for epicardium).

The corresponding performance (DICE index) of the NN when trained on the mid-ventricular MR-images of the entire population (healthy and patients, 187 images) of the publicly available dataset

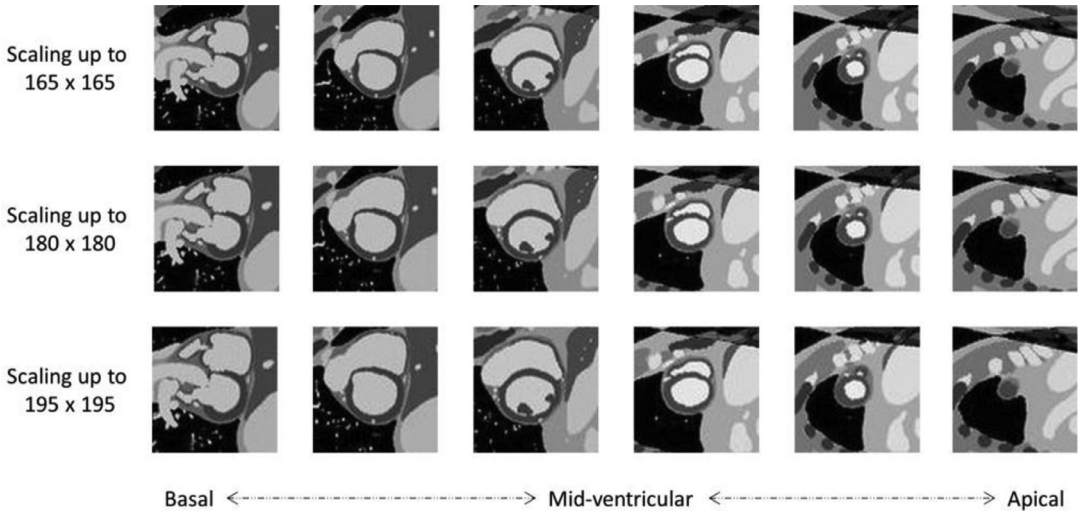

Fig. 3. Simulated short-axis MR-images covering the entire LV volume for three different scale factors. The first row presents the simulated short-axis MR-images that were scaled up to a size of 165 × 165, the second row presents the same simulated short-axis MR-images that were scaled up to a size of 180 × 180 and the third row presents the same simulated short-axis MR-images that were scaled up to a size of 195 × 195.

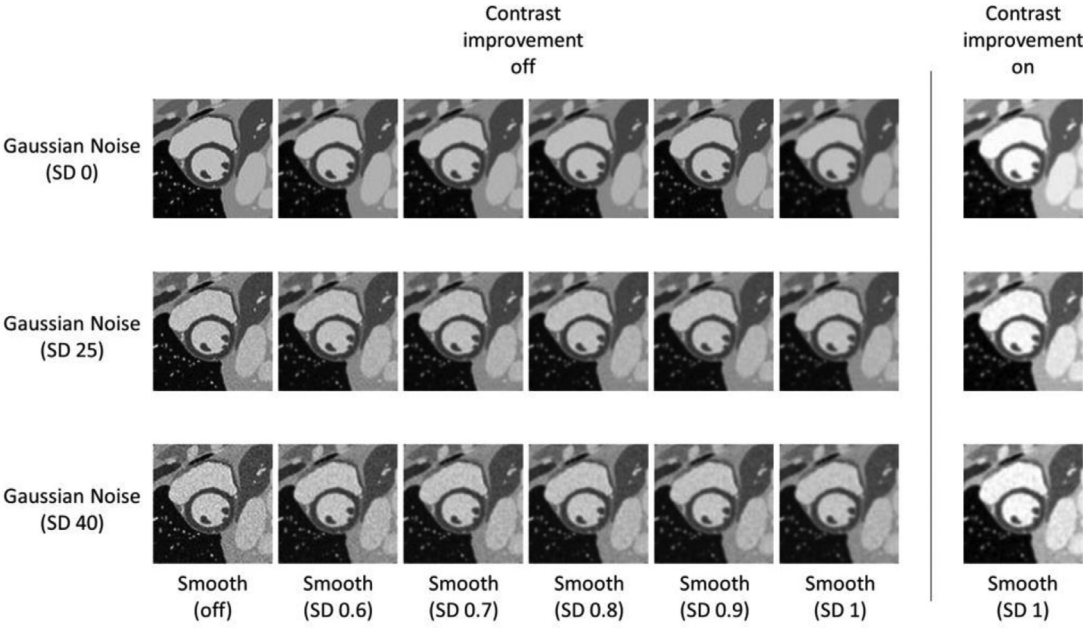

Fig. 4. A simulated, mid-ventricular, short-axis, MR image for different combinations of noise on the simulated MR signal (no noise, Gaussian distributed noise with zero mean and standard deviation equal to 25, and with Gaussian distributed noise with zero mean and standard deviation equal to 40), gauss filtering (2-D Gaussian smoothing kernel with standard deviations of 0.6, 0.7, 0.8, 0.9 and 1) and contrast improvement.

Table 3
NN trained 100% on simulated MR-images: performance on delineating the LV endocardium and epicardium in the mid-ventricular images acquired from 8 healthy volunteers. Training was performed a) on normal anatomy a.k.a. first dataset and b) on normal anatomy as well as on pathological reduced myocardial thickness at the apex of the LV a.k.a. second dataset. All experiments in this table refer to configurations of images with all scales (165 × 165, 180 × 180 and 195 × 195 pixels) and a 2-D Gaussian smoothing kernel with a standard deviation of 0.9. The experiments in bold represent the best performing configurations (3,4 and 6). The corresponding DICE indexes of the NN were 93% (endocardium) and 91% (epicardium) when trained on the mid-ventricular MR-images of the entire population of the publicly available dataset and 10% (endocardium) and 15% (epicardium) when trained on the mid-ventricular MR-images of the healthy subjects alone. HD: Hausdorff Distance; endo: endocardium; epi: epicardium; SD: standard-deviation.

				Normal anatomy training (216 images)				Normal and abnormal anatomy training (432 images)			
Nr.	Experiment	Noise SD	Contrast improvement	DICE endo (%)	DICE epi (%)	HD endo	HD epi	DICE endo (%)	DICE epi (%)	HD endo	HD epi
1	Noise00	0	No	90	80	4.46	12.4	92	86	3.77	10.78
2	Noise00_ci	0	Yes	91	84	4.2	12.08	89	87	4.83	11.4
3	**Noise25**	25	No	86	90	5.65	8.51	**93**	**90**	2.98	7.59
4	**Noise25_ci**	**25**	**Yes**	**93**	**87**	3.43	9.19	87	85	5.47	11.53
5	Noise40	40	No	87	88	5.44	9.27	92	85	4.17	11.33
6	**Noise40_ci**	**40**	**Yes**	**94**	**86**	2.88	11.4	92	87	3.57	10.44

Table 4
NN trained with a mix of simulated MR-images and real images: performance on delineating the LV endocardium and epicardium in the mid-ventricular images acquired from 2 healthy volunteers. All experiments in this table refer to configurations of images with all scales (165 × 165, 180 × 180 and 195 × 195 pixels) and a 2-D Gaussian smoothing kernel with a standard deviation of 0.9. The experiment in bold represents the best performing solution. The corresponding DICE indexes of the NN were 96% (endocardium) and 94% (epicardium) when the NN was trained on 159 randomly selected mid-ventricular images from the entire population (healthy and patients) of the publicly available dataset. HD: Hausdorff Distance; endo: endocardium; epi: epicardium; SD: standard-deviation; mid: mid-ventricular.

Nr.	Experiment	Noise SD	Contrast improvement	Simulated vs Real	Short-axis slices	Size training dataset	Size artificial dataset	Size real dataset	DICE endo (%)	DICE epi (%)	HD endo	HD epi
1	Noise25_8020	25	No	80%-20%	Mid	200	159	Mid: 39	97	97	1.55	2.11
2	**Noise40_ci_8020**	40	Yes	80%-20%	Mid	200	159	Mid: 39	97	97	1.59	1.86
3	Noise25_9010	25	No	90%-10%	Mid	178	159	Mid: 18	97	95	1.74	2.91
4	Noise40_ci_9010	40	Yes	90%-10%	Mid	178	159	Mid: 18	96	95	1.6	1.99
5	Noise25_sim	25	No	100%-0%	Mid	159	159	0	87	90	6.81	9.98
6	Noise40_ci_sim	40	Yes	100%-0%	Mid	159	159	0	87	90	6.61	7.39

Table 5
Performance of the NN for delineating the LV endocardium and LV epicardium in real short-axis images that covered the entire LV volume (test dataset: 28 images) when trained with different mixes of simulated and real MR-images. The experiment in bold represents the best performing solution (line 1). The corresponding DICE indexes of the NN were 78% (for both endocardium and epicardium) when the NN was trained with 71 in-vivo MR-images from the dataset of the 8 healthy volunteers and 83% (endocardium) and 81% (epicardium) when trained on 279 randomly selected real images (from the entire LV volume) from the entire population (healthy and patients) of the publicly available dataset. LV: Left-ventricle; Mid: mid-ventricular; Ap: Apical; B: Basal; HD: Hausdorff Distance; endo: endocardium; epi: epicardium; SD: standard-deviation.

Nr.	Experiment	Noise SD	Contrast improvement	Simulated vs Real	Short-axis slices	Size training dataset	Size artificial dataset	Size real dataset	DICE endo (%)	DICE epi (%)	HD endo	HD epi
1	**Noise25_8020**	25	No	80%-20%	Entire LV	350	279	Mid: 48, Ap: 16, B: 7	85	86	4.23	5.07
2	Noise40_ci_8020	40	Yes	80%-20%	Entire LV	350	279	Mid: 48, Ap: 16, B: 7	83	88	4.94	4.66
3	Noise25_9010	25	No	90%-10%	Entire LV	310	279	Mid: 12, Ap: 12, B: 7	82	84	5.57	6.68
4	Noise40_ci_9010	40	Yes	90%-10%	Entire LV	310	279	Mid: 12, Ap: 12, B: 7	83	84	6.72	7.22
5	Noise25_sim	25	No	100%-0%	Entire LV	279	279	0	59	47	18.68	23.44
6	Noise40_ci_sim	40	Yes	100%-0%	Entire LV	279	279	0	67	49	11.59	21.45
7	In-vivo images (72% of all images from 8 healthy volunteers)	-	-	In-vivo	Entire LV	71	0	Mid: 48, Ap: 16, B: 7	78	78	7.01	10.69

was 93% and 91% for endocardium and epicardium respectively. When trained on the mid-ventricular MR-images of the healthy subjects alone (34 images) the performance dropped down to 10% and 15% for endocardium and epicardium, respectively.

Table 4 shows the performance of the NN in delineating the LV endocardium and LV epicardium in mid-ventricular images (16 images) acquired from two volunteers. Training of the NN was performed with a different mix of simulated and real MR-images. The addition of 10% real MR-images to the 159 simulated images used for training helped raise the performance to 97% (DICE) for endocardium and 95% (DICE) for epicardium (lines 3 and 4 in Table 4). Similar to the aforementioned performance was obtained when the NN was trained on 159 randomly selected mid-ventricular images from the entire population (healthy and patients) of the publicly available dataset and tested on the same 16 mid-ventricular images of the two healthy volunteers (96% and 94% for endocardium and epicardium respectively). Section "Tables 4 and 5 plots" in Appendix presents plots of DICE and Hausdorff scores for the experiments presented in Table 4.

4.2. All slices covering the entire LV volume

Table 5 shows the performance of the NN in delineating the LV endocardium and LV epicardium in the real short-axis images that cover the entire LV from base to apex (test dataset: 28 images) for different mixes of simulated and real MR-images in the training dataset. An addition of 10% real MR-images in the artificial training dataset (lines 3 and 4 in Table 5) yielded a performance of 83% and 84% (DICE index) for endocardium and epicardium respectively. The addition of 20% real MR-images in the artificial training dataset (lines 1 and 2 at Table 5) yielded a performance of 85% and 88% (DICE index) for endocardium and epicardium, respectively. With 0% real MR-images added to the artificial training dataset the performance of the NN was only 67% and 49% for endocardium and epicardium respectively (line 6 at Table 5). Typical results of the performance of the NN (for experiment 1 in Table 5) on real MR-images acquired from 8 healthy volunteers are shown in Fig. 5. Last, when the NN was trained with 71 in-vivo MR-images from the dataset of the 8 healthy volunteers (corresponding to 72% of the total images acquired from this population) then the performance was 78% (DICE index) for both endocardium and epicardium (line 7 at Table 5).

The corresponding performance (DICE index) of the NN that was trained on 279 randomly selected real images (from the entire LV volume) from the entire population (healthy and patients) of the publicly available dataset and tested on the same 28 images as above (entire LV volume) of the eight healthy volunteers dataset was 83% and 81% for endocardium and epicardium respectively.

Section "Tables 4 and 5 plots" in Appendix presents plots of DICE and Hausdorff scores for the experiments presented in Table 5.

Moreover, the training of the NN on all short-axis, end-diastolic, MR-images (entire LV volume) from the entire population (healthy and patients, 420 images) of the publicly available dataset resulted in a performance of 86% (DICE index) for the delineation of both the endocardium and epicardium in the dataset of 28 images of the eight healthy volunteers. The corresponding performance of the NN that was trained on all short-axis, end-diastolic, MR-images (entire LV volume) from the healthy subjects only (78 images) was 82% and 71% for the delineation of endocardium and epicardium, respectively.

4.3. Discussion

This study demonstrated the first application of a simulator for generating datasets for use in training with supervised ML techniques. In particular, a physics simulator was used to generate

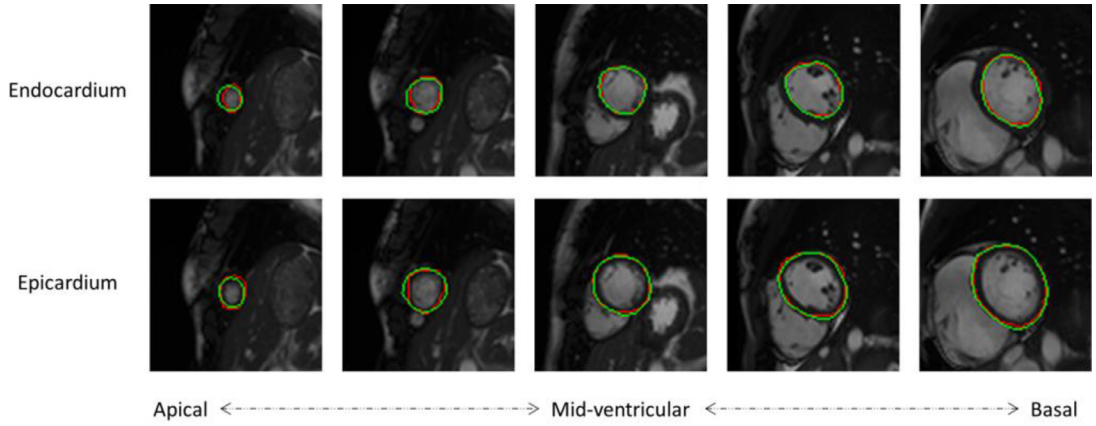

Fig. 5. Typical delineations on real MR-images acquired from the population of eight healthy volunteers. The top row presents the delineations of the endocardium whereas the bottom row presents the corresponding delineations of the epicardium. Red contours indicate ground truth whereas green contours indicate the predictions given by the NN. The NN was trained with a 20% addition of real MR-images in the artificial training dataset (experiment Noise25_8020, line 1 at Table 5).

cardiac MR images for training neural networks in myocardial segmentation.

In this study, the 4D-XCAT anatomical model [13] was used with the coreMRI simulation platform (www.coremri.org) [14] for generating artificial MR-images. Artificial datasets of short-axis MR-images were generated and utilized for training a fully convolutional neural-network in order to automatic delineate the endocardium and the epicardium of the LV. A set of several different parameters for generating artificial images was utilized and the performance of the neural-network was assessed on real MR-images acquired from eight healthy volunteers. For comparison purposes, the neural-network was also trained on real MR-images from a publicly available dataset and its performance was assessed on the same healthy volunteers' data.

The proposed solution showed that a dataset of simulated MR-images has the potential to substitute datasets of real MR-images for training a NN for automatic LV segmentation. The research community has long recognized more technical difficulties in the segmentation of basal and apical slices compared to mid-ventricular slices [20,24,25]. For this reason, two series of experiments were performed in this study. Experiments with mid-ventricular only slices were initially performed to evaluate the performance of the proposed solution against conventional approaches on a simpler framework. The proposed solution demonstrated a performance of 94% DICE (endocardium) and 90% DICE (epicardium) in real mid-ventricular slices, whereas a 10% addition of real MR-images in the artificial training dataset demonstrated a performance of 97% DICE (endocardium) and 95% DICE (epicardium). The corresponding performance of the NN that was trained on the mid-ventricular MR-images of a publicly available dataset was 96% and 94% for endocardium and epicardium, respectively. In the set of more advanced experiments that cover the entire LV, the highest performance was achieved under the 80%-20% configuration and was equal to 85% and 88% (DICE index) for the delineations of endocardium and epicardium, respectively, in real short-axis images covering the entire LV volume. The corresponding performance of the NN that was trained on the entire publicly available dataset was 86% for the delineations of both endocardium and epicardium. Last, the performance of the NN that was trained on 71 in-vivo MR-images from the dataset of the 8 healthy volunteers was 78% DICE for both endocardium and epicardium whereas the corresponding performance of the NN with the addition of 279 simulated images went up to 85% and 88% DICE, respectively.

Despite the high performance achieved in this study with the proposed approach, the lower performance of the NN for delineating the entire LV volume could be associated to the fuzzy boundaries in basal and apical slices that could have a negative influence on the segmentation accuracy in these locations. However, a similar performance drop was observed using the publicly available dataset for training the NN suggesting that this was not a limitation of the proposed approach. Although these technical challenges in LV segmentation have long been recognized by the research community [20,24,25], the lower performance could also be associated to the level of realism that the selected simulation model exhibited in these anatomical locations. This limitation may be due to the low number of available tissue types within the computer model but also to the selected magnetic properties of the available tissue types due to lack of measurements in the literature. Moreover, the proposed simulation framework is highly dependent on a prior knowledge of the characteristics of the real MR experiment (from magnetic properties of the tissues to the applied noise level on the simulated signal). Any deviation of the simulation model from real conditions may affect the performance of the proposed solution. Last, the different orientation of the simulated MR-images compared to the real MR-images could be considered as a limitation of the current study that could adversely affect the performance of the NN [26].

This study showed that the generation of an artificial dataset for training a NN is a cost-effective solution that has the potential to outperform the traditional approach of acquiring real MR data from healthy volunteers and patients. In the field of LV segmentation, the proposed solution outperformed the utilization of a publicly available dataset that was acquired from 45 subjects (healthy subjects and patients) and required the availability of MR scanners, technologists and radiologists. Compared to this conventional approach, the proposed solution could generate the artificial dataset and the corresponding delineations within a couple of hours on a single-GPU computer on the cloud. Moreover, the high performance achieved in this study could be attributed to the generation of the artificial dataset using a similar pulse sequence design as the one used for the acquisition of the real test dataset from healthy volunteers. The use of representative training datasets has become a necessity for achieving high performance in NN, however, this option remains today costly, resource intensive and may not always be available. The suggested methodology provides a flexible solution that allows the generation of training datasets for multiple variations of both the imaging protocol and the computer model, which is not easy today with a real MRI experiment configuration (MRI system, patient recruitment, etc.). In addition, the digitized and highly-customizable nature of the anatomical model

allows for the concurrent production of well-annotated data in the form of tissue masks. The annotation of the artificial data is always objective and depicts the real tissue characteristics without being affected by various factors that may deteriorate the quality of the medical image, such as noise, limited spatial and temporal resolution due to the pulse sequence design, etc. The automatic generation of well-annotated tissue masks may also aid in training recurrent NNs that utilize temporal information from images acquired at different phases along the cardiac cycle [27]. Moreover, variants where noise and the presence of other artifacts that can make the data look close to real can easily be accommodated. Last, compared to publicly available datasets, the current work proposes an alternative approach on the effort needed today to setup public databases of reliable data. Advanced MR simulations can help in limiting the number of datasets needed to be made available publicly in such databases. Also, they could potentially help in cases where data are hard to come by or are not easily sharable due to regulatory issues. Moreover, when sharing such real data, it is most of the times difficult to also keep track of the particular protocols that were used to acquire the images, which is important in determining image characteristics in MR. The datasets from MRI simulations could potentially reduce the effort needed for setting up publicly available datasets or provide an additional value to them. With the proposed solution, the training sets can be produced faster and easier, without the use of a real MRI scanner, without scanning patients and without requiring expensive specialized personnel (such as technologists for running the MR scanner, radiologists for evaluating the MR-images or data scientists to properly label publicly available data). Even in cases of more advanced MR applications where the incorporation of real data may be beneficial (such as the cases of basal and apical LV segmentation presented in this study), this study suggests that the use of less subjects and personnel-effort within the proposed framework may achieve the same or even better performance compared to the sole use of real data (Table 5).

The contribution of this work in the field of MR image synthesis can be summarized as follows: 1) Realistic simulated MR data: The suggested method allows for the generation of training datasets through a realistic simulation framework. Such a simulation approach is as close as possible to a real MR experiment since it integrates realistic aspects of the MRI experiment from signal generation to image formation and solves the entire complex problem for densely spaced isochromats and for a densely spaced time axis. Other available simulation frameworks [28] are based on approximations via signal equations whereas basic and more complex artifacts cannot be taken into account. 2) High simulation performance: The publicly available, cloud-based and GPU-based simulation framework allows for the generation of artificial datasets and their corresponding delineations within a few hours. Others available solutions [17,18,29,30] may require purchase of advanced computer systems, advanced setup of a computer cluster, advanced technical knowledge, and the ability to modify the source code. 3) Well-annotated data: The digitized and highly-customizable nature of the anatomical model allows for the concurrent, automatic production of well-annotated data in the form of tissue masks. While other solutions that involve neural style transfer [31,32] may generate realistic MR images, the corresponding tissues masks may not follow the generated anatomy due to geometric distortions introduced in the target generated image by the style transfer mechanism [33].

In the field of MR image synthesis, the utilization of statistical atlases [11] could be considered as an alternative to the approach proposed in the current study. Statistical atlases are more accurate in the representation of a population than models built on a single subject or based on analytical expressions of general knowledge. Moreover, the synthesis of artificial images is usually accompanied by the generation of well-annotated tissue masks. However, in the field of MRI, representative statistical atlases usually require the availability of thousands of real MR images and an increased amount of manual intervention [11] whereas their synthetic images are usually confined to the same type of MR images as the ones used for the structure of the statistical atlas and cannot be generalized. On the other hand, the utilization of a representative statistical atlas as the main anatomical model in the proposed framework could yield more representative MR images. This could be a topic of further research in the future. In a similar manner, simulated MR images could be combined with real MR images in a style-transfer process to generate more realistic synthetic MR images. Preliminary results of the utilization of a GAN framework are presented in Section "GAN framework" in Appendix.

The work presented here is a proof-of-concept study that adds on the research that other groups have recently performed on medical image simulations and synthesis [34]. This proof-of-concept study suggests that a low-cost solution for the generation of artificial training datasets for supervised learning techniques is feasible in the field of MR by means of cloud-based and GPU-based, advanced MR simulations on a realistic computer model of the human anatomy. The proposed framework could be applied in other segmentation-related MR applications (such as MRI segmentation for tumor volume measurements, MRI segmentation of the human brain, etc.) [35,36] but also in more advanced medical applications that require the synthesis of medical images from one imaging modality to another [37,38].

This study does not intent to become another automatic LV segmentation method but mainly aims on showing how existing NNs on several other MR image processing methods may be utilized without acquiring real MR data or without using publicly available small benchmark datasets that may not be representative to the corresponding test dataset in terms of image quality and characteristics. Future work within the proposed framework should involve the development and incorporation of more advanced computer models that simulate pathological cases and specific diseases. Moreover, future work should also aim in the incorporation of different types of image artifacts that have been covered extensively in previous studies [14,17,18,39] so as more realistic artificial MR images to be generated and cover a wider spectrum of MR applications. We hope that this work will change the way training datasets are generated and inspire follow-up studies with applications in patients and specific diseases.

5. Conclusions

This study presented an alternative solution to the limitations encountered today in generating training datasets that are used in supervised learning techniques in MRI. Nowadays, the lack of large training sets does not allow for the development of advanced supervised ML techniques, whereas the availability of non-representative training sets may delay the application of these techniques in clinical practice. The proposed solution creates artificial training datasets through the application of advanced MR simulations on a realistic computer model of human anatomy without the use of a real MRI scanner, without scanning patients and without using expensive personnel (such as technologists, radiologists and data scientists).

Ethics approval and consent to participate

The study was approved by the Regional ethics committee and all subjects provided written informed consent (The regional ethics committee, Lund, Sweden. Ethics applications numbers: 541/2004 and 815/2016).

Consent for publication

All subjects provided written consent for publication of this study and accompanying images.

Availability of data and materials

Access to the source code of the MR simulator is not available, however, the coremri simulation platform (www.coremri.org) is publicly available to the research community. The datasets used and/or analyzed during the current study are available from the authors on reasonable request.

Funding

This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Program «Human Resources Development, Education and Lifelong Learning 2014-2020» in the context of the project "Simulation-based development of supervised machine learning algorithms" (MIS 5004853). One part of the Amazon-cloud resources utilized in this study was awarded through the "AWS Cloud Credits for Research" program.

Authors' contributions

CGX conceived and designed the study. CGX made the anatomical model compatible with MR simulations, designed the pulse sequence and run the MR simulations. DF extracted the short-axis slices from the anatomical model and the corresponding tissue maps. KH worked on the initial implementation of the NN and the initial design of the NN experiments. CGX run the NN experiments and collected the data presented in this study. CGX drafted the manuscript. CGX and AHA participated in funding acquisition. AHA had the management and coordination responsibility for the research activity planning and execution. All authors have contributed to final manuscript and approved it.

Competing interests

CGX and AHA are co-founders of Corsmed AB.

Acknowledgements

The authors would like to acknowledge the contribution of Aikaterini Karampasi for participating in the development of the neural-network and for running the experiments presented in this study, Dr. Felicia Seeman for the acquisition of the real MR-images from the group of eight healthy volunteers, Dr. Robert Jablonowski for the delineations of the real MR-images acquired from the group of eight healthy volunteers and Dr. Håkan Arheden for providing the resources needed for running this project within the Cardiac MR group, Lund, Sweden. The acknowledged contributors of this work would have been considered as co-authors for this study but this was not possible due to restrictions imposed by the organization providing the research funding.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2020.105817.

References

[1] S Wang, RM Summers, Machine learning and radiology, Medical image analysis 16 (5) (2012 Jul) 933–951 PubMed PMID: 22465077. Pubmed Central PMCID: 3372692.
[2] M. de Bruijne, Machine learning approaches in medical image analysis: From detection to diagnosis, Medical image analysis 33 (2016 Oct) 94–97 PubMed PMID: 27481324.
[3] Haskins G, Kruger U, Yan P. Deep Learning in Medical Image Registration: A Survey. arXiv preprint arXiv:190302026. 2019.
[4] A Işın, C Direkoğlu, M Şah, Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods, Procedia Computer Science 102 (2016) 317–324 2016/01/01/.
[5] H Li, ML Giger, BQ Huynh, NO Antropova, Deep learning in breast cancer risk assessment: evaluation of convolutional neural networks on a clinical dataset of full-field digital mammograms, J Med Imaging (Bellingham) 4 (4) (2017) 041304 PubMed PMID: 28924576. Epub 09/13. eng.
[6] , Small sample size effects in statistical pattern recognition: recommendations for practitioners and open problems, in: SJ Raudys, AK Jain (Eds.), [1990]Proceedings 10th International Conference on Pattern Recognition, 1990 16-21 June 1990.
[7] D-C Li, IH Wen, A genetic algorithm-based virtual sample generation technique to improve small data set learning, Neurocomputing 143 (2014) 222–230 2014/11/02/.
[8] J Salamon, JP. Bello, Deep convolutional neural networks and data augmentation for environmental sound classification, IEEE Signal Processing Letters 24 (3) (2017) 279–283.
[9] Oksuz I, Clough J, Ruijsink B, Puyol-Antón E, Bustin A, Cruz G, et al., editors. Detection and Correction of Cardiac MRI Motion Artefacts During Reconstruction from k-space2019; Cham: Springer International Publishing.
[10] R Shaw, C Sudre, S Ourselin, MJ Cardoso, et al., MRI k-Space Motion Artefact Augmentation: Model Robustness and Task-Specific Uncertainty, in: MJ Cardoso, F Aasa, G Ben, K Ender, O Ipek, U Gozde, et al. (Eds.), Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning; Proceedings of Machine Learning Research: PMLR, 2019, pp. 427–436.
[11] W Bai, W Shi, A de Marvao, TJ Dawes, DP O'Regan, SA Cook, et al., A bi-ventricular cardiac atlas built from 1000+ high resolution MR images of healthy subjects and an analysis of shape and motion, Medical image analysis 26 (1) (2015 Dec) 133–145 PubMed PMID: 26387054.
[12] Dermgan: Synthetic generation of clinical skin images with pathology, in: A Ghorbani, V Natarajan, D Coz, Y Liu (Eds.)Machine Learning for Health Workshop, 2020.
[13] WP Segars, G Sturgeon, S Mendonca, J Grimes, BM Tsui, 4D XCAT phantom for multimodality imaging research, Med Phys 37 (9) (2010 Sep) 4902–4915 PubMed PMID: 20964209. Pubmed Central PMCID: 2941518.
[14] CG Xanthis, AH Aletras, coreMRI: A high-performance, publicly available MR simulation platform on the cloud, PloS one 14 (5) (2019) e0216594 PubMed PMID: 31100074.
[15] O Bieri, K. Scheffler, Fundamentals of balanced steady state free precession MRI, Journal of Magnetic Resonance Imaging 38 (1) (2013) 2–11.
[16] H. Carr, Steady-state free precession in nuclear magnetic resonance, Physical Review 112 (5) (1958) 1693.
[17] CG Xanthis, IE Venetis, AH Aletras, High performance MRI simulations of motion on multi-GPU systems, Journal of cardiovascular magnetic resonance: official journal of the Society for Cardiovascular Magnetic Resonance 16 (2014) 48 PubMed PMID: 24996972. Pubmed Central PMCID: 4107941.
[18] Xanthis CG, Venetis IE, Chalkias AV, Aletras AH. MRISIMUL: A GPU-based Parallel Approach to MRI Simulations IEEE Transactions on Medical Imaging. 2014;3:607-17.
[19] MS Hansen, TS Sorensen, Gadgetron: an open source framework for medical image reconstruction, Magn Reson Med 69 (6) (2013 Jun) 1768–1776 PubMed PMID: 22791598.
[20] Tran PV. A fully convolutional neural network for cardiac segmentation in short-axis MRI. arXiv preprint arXiv:160400494. 2016.
[21] B Lowekamp, D Chen, L Ibanez, D Blezek, The Design of SimpleITK, Frontiers in Neuroinformatics 7 (45) (2013) English.
[22] E Heiberg, J Sjögren, M Ugander, M Carlsson, H Engblom, H Arheden, Design and validation of Segment - freely available software for cardiovascular image analysis, BMC medical imaging 10 (1) (2010) 1.
[23] Radau P, Lu Y, Connelly K, Paul G, Dick AJ, Wright GA, Wright G. Evaluation Framework for Algorithms Segmenting Short Axis Cardiac MRI. 2009.
[24] C Petitjean, J-N. Dacher, A review of segmentation methods in short axis cardiac MR images, Medical image analysis 15 (2) (2011) 169–184 2011/04/01/.
[25] V Tavakoli, AA. Amini, A survey of shaped-based registration and segmentation techniques for cardiac images, Comput Vis Image Underst 117 (9) (2013) 966–989.
[26] K Men, P Boimel, J Janopaul-Naylor, C Cheng, H Zhong, M Huang, et al., A study of positioning orientation effect on segmentation accuracy using convolutional neural networks for rectal cancer, Journal of Applied Clinical Medical Physics 20 (1) (2019) 110–117.
[27] Automated segmentation on the entire cardiac cycle using a deep learning work-flow, in: N Savioli, MS Vieira, P Lamata, G Montana (Eds.)2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), IEEE, 2018.
[28] L Wissmann, C Santelli, WP Segars, S Kozerke, MRXCAT: Realistic numerical phantoms for cardiovascular magnetic resonance, Journal of cardiovascular magnetic resonance: official journal of the Society for Cardiovascular Magnetic Resonance 16 (2014 Aug 20) 63 PubMed PMID: 25204441. Pubmed Central PMCID: 4422262.

[29] F Liu, JV Velikina, WF Block, R Kijowski, AA Samsonov, Fast Realistic MRI Simulations Based on Generalized Multi-Pool Exchange Tissue Model, IEEE Trans Med Imaging 36 (2) (2017 Feb) 527–537 PubMed PMID: 28113746. Pubmed Central PMCID: 5322984.
[30] T Stocker, K Vahedipour, D Pflugfelder, NJ Shah, High-performance computing MRI simulations, Magn Reson Med 64 (1) (2010 Jul) 186–193 PubMed PMID: 20577987.
[31] SU Dar, M Yurt, L Karacan, A Erdem, E Erdem, T Çukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks, IEEE transactions on medical imaging 38 (10) (2019) 2375–2388.
[32] A Sharma, G Hamarneh, Missing MRI pulse sequence synthesis using multi--modal generative adversarial network, IEEE transactions on medical imaging 39 (4) (2019) 1170–1183.
[33] X Yan, Y Xing, Z He, T Dai, Y Jiang, S-T Xia, Neural Style Transfer with Content Discrimination, in: IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2019, 2019, pp. 78–83.
[34] Medical image synthesis for data augmentation and anonymization using generative adversarial networks, in: H-C Shin, NA Tenenholtz, JK Rogers, CG Schwarz, ML Senjem, JL Gunter, et al. (Eds.)International Workshop on Simulation and Synthesis in Medical Imaging, Springer, 2018.
[35] L Clarke, R Velthuizen, M Camacho, J Heine, M Vaidyanathan, L Hall, et al., MRI segmentation: methods and applications, Magnetic resonance imaging 13 (3) (1995) 343–368.
[36] I Despotović, B Goossens, W Philips, MRI Segmentation of the Human Brain: Challenges, Methods, and Applications, Computational and mathematical methods in medicine 2015 (2015) 450341 2015/03/01.
[37] W Li, Y Li, W Qin, X Liang, J Xu, J Xiong, et al., Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy, Quantitative Imaging in Medicine and Surgery 10 (6) (2020) 1223–1236.
[38] C Siversson, F Nordström, T Nilsson, T Nyholm, J Jonsson, A Gunnlaugsson, et al., Technical Note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm, Medical Physics 42 (10) (2015) 6090–6097.
[39] CG Xanthis, S Bidhult, A Greiser, K Chow, RB Thompson, H Arheden, et al., Simulation-based quantification of native T1 and T2 of the myocardium using a modified MOLLI scheme and the importance of Magnetization Transfer, Magn Reson Imaging 48 (2018 May) 96–106 PubMed PMID: 29288037.