BEN- GURION UNIVERSITY OF THE NEGEV
THE FACULTY OF ENGINEERING SCIENCES
DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

ACCELERATED MAGNETIC RESONANCE IMAGING BY GENERATIVE ADVERSARIAL NEURAL NETWORKS

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE MASTER OF SCIENCES DEGREE

OHAD SHITRIT

UNDER THE SUPERVISION OF: DR TAMMY RIKLIN-RAVIV

Feb 27, 2018

BEN- GURION UNIVERSITY OF THE NEGEV

THE FACULTY OF ENGINEERING SCIENCES

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

ACCELERATED MAGNETIC RESONANCE IMAGING BY GENERATIVE ADVERSARIAL NEURAL NETWORKS

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE MASTER OF SCIENCES DEGREE

OHAD SHITRIT

UNDER THE SUPERVISION OF: DR TAMMY RIKLIN-RAVIV

Signature of student: _____ Date: 19/2/2018

Signature of supervisor: תמי ריקלין רביב Date: 19/2/2018

Signature of chairperson

Of the committee for graduate studies: _____ Date: _____

Abstract

A main challenge in Magnetic Resonance Imaging (MRI) for clinical applications is speeding up scan time. Beyond the improvement of patient experience and the reduction of operational costs, faster scans are essential for time-sensitive imaging, where target movement is unavoidable, yet must be significantly lessened, e.g., fetal MRI, cardiac cine, and lungs imaging. Moreover, short scan time can enhance temporal resolution in dynamic scans, such as functional MRI or dynamic contrast enhanced MRI. Current imaging methods facilitate MRI acquisition at the price of lower spatial resolution and costly hardware solutions.

We introduce a practical, software-only framework, based on deep learning, for accelerating MRI acquisition, while maintaining anatomically meaningful imaging. This is accomplished by partial MRI sampling, while using an generative adversarial neural networks to directly estimate the missing k-space samples. The inter-play between the generator and the discriminator networks enables the introduction of an adversarial cost in addition to fidelity loss used for optimizing the peak signal-to-noise ratio (PSNR). Promising image reconstruction results are obtained for 3T and 1.5T brain MRI, from a large publicly available dataset, where only 40%, 25%, and 16.6% of the raw samples of each scan are used. To assess the clinical usability of the reconstructed images we also performed tissue segmentation and compared the results to those obtained by using the original fully-sampled images. Segmentation compatibility, measured in terms of Dice scores and Hausdorff distances, demonstrate the quality of the proposed MRI reconstruction with respect to other methods, including the widely-used Compressed Sensing.

Contents

1 Acknowledgments **5**

2 Introduction **6**
2.1 Magnetic Resonance Imaging (MRI) 6
2.2 MRI Acceleration 7
2.2.1 Classic methods 7
2.2.2 Machine Learning Methods 8

3 Method **10**
3.1 K-space 10
3.1.1 Notations and Problem Definition 10
3.1.2 K-space GAN 11
3.1.3 System Architecture 11

4 Experimental Results **14**
4.1 Random Gaussian Sampling 14
4.1.1 PSNR and SSIM 15
4.1.2 Brain Extraction - Skull Stripping 17
4.1.3 Brain Segmentation 18
4.2 Phase Encoding Sampling 22

5 Summary and Discussion **25**
5.1 Discussion 25
5.1.1 Acknowledgment 26

List of Tables

4.1 Error in PSNR, without masking 15
4.2 Error in PSNR, with masking 16
4.3 SSIM for different sampling factors 17
4.4 MHD - brain extraction 18
4.5 Segmentation Dice scores for different sampling ratios 19
4.6 Error in PSNR 23

List of Figures

3.1 Under-sampling artifacts: the arrows illustrate Cartesian sampling 10
3.2 Framework architecture: G and D are the generator and discriminator networks, respectively. See text for details. 12
3.3 Networks architecture. Both the discriminator and the generator inputs are two-channel image for the real and imaginary components. The number above each layer represents the output channels. All convolutional kernels are 3×3. 13

4.1 $2D$ Gaussian sampling masks 15
4.2 PSNR without masking error-bar 16
4.3 Example of brain extraction 18
4.4 Examples of reconstructed MR images from under-sampled k-space - Factor 6. 20
4.5 Examples of reconstructed MR images from under-sampled k-space - Factor 4. 21
4.6 Examples of reconstructed MR images from under-sampled k-space - Factor 2.5. 22
4.7 Examples of reconstructed MR images from under-sampled k-space. 24

.

Publications

1 Ohad Shitrit, Tzachi Hershkovich ,Tamar Shalmon, Ilan Shelef, and Tammy Riklin Raviv. Probabilistic model for 3D interactive segmentation, In MICCAI workshop on Interactive Medical Image Computing (IMIC), 2014.

2 Tzachi Hershkovich ,Ohad Shitrit, Tammy Riklin Raviv et al., T.R., 2016. Probabilistic model for 3D interactive segmentation. Computer Vision and Image Understanding, 151, pp.47-60. Vancouver

3 Ohad Shitrit and Tammy Riklin Raviv. Accelerated Magnetic Resonance Imaging by Adversarial Neural Network. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 30-38). Springer, Cham.

4 Submission to IEEE journal.

Chapter 1

Acknowledgments

This was a great journey, challenging and educational. Thanks to all the amazing people that support, encouraged and guided me through these years.

First of all, I would like to thank to my supervisor, Dr. Tammy Riklin-Raviv, for guiding me in the last four years. Starting from the my undergraduate final project, during my master and my thesis research. Thanks for letting me be a part of one of the most fascinating filed of research. I admire your professionalism, dedication and your curiosity, which always been an inspiration for me. Tammy, thank you for the long hours of conversations, tips and guidances at any time, explaining theory and ideas in a simple and easy way to understand. All these are made me a better researcher and engineer.

Second, I want to thank to my lab colleagues, for a fruitful brainstormings, technical support and most important, being my friends. Great thanks to Tzachi Hershkovich, that was my partner of my undergraduate final project, I have learned a lot from you.

Last but not least, I want to thank to my life partner, my beloved wife Lee, for all your support during this journey. You encouraged me to be who I am today.

Chapter 2

Introduction

2.1 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI) is a non-ionizing imaging modality, and is therefore widely used in diagnostic medicine and biomedical research.

From MRI theory, we know that raw data collected from MR machines, the k-space values, which are RF signals at a predefined range of frequencies and phases.

Let's define G_{FE}, G_{PE} as the frequency and phase encoding gradients respectively. K-space is in a spatial frequency domain. The complex array of raw data points is treated as a two-dimensional grid of points (k_x, k_y) where k_x, k_y are defined as follows:

$$\begin{cases} k_x = \gamma G_{FE} t \\ k_y = \gamma G_{PE} T_{PE} \end{cases} \tag{2.1}$$

T_{PE} denotes the total time in which the respective gradient is active and $r\gamma$ is the gyromagnetic constant. The desired signal, a 2D MR image, is defined as the inverse Fourier transform of the encoded measured signal s_0, where $\Omega \subset \mathbb{R}^2$:

$$I(x, y) = \iint_{\Omega} s_0(k_x, k_y) e^{+j2\pi(k_x x + k_y y)} dk_x dk_y \tag{2.2}$$

As known from signal processing theory, lack of information in k-space will cause artifacts such as aliasing, blurring, and results in low resolution MR images. In some cases, decreasing the acquisition time is necessary for detecting temporal changes which can indicate certain illnesses.

2.2 MRI Acceleration

2.2.1 Classic methods

Magnetic Resonance Imaging (MRI) is a non-ionizing imaging modality, and is therefore widely used in diagnostic medicine and biomedical research. The physical principles of MRI are based on a strong magnetic field and pulses of radio frequency (RF) electromagnetic radiation. Images are produced when hydrogen atoms, which are prevalent in living organisms, emit the absorbed RF energy that is then received by antennas in close proximity to the anatomy being examined. Spatial localization of the detected MRI signals is obtained by varying the magnetic field gradients. The discretized RF output is presented in a Fourier space (called k-space), where the x-axis refers to the frequency and the y-axis to the phase. An inverse fast Fourier transform (IFFT) of the k-space is then used for generating anatomically meaningful MRI scans. Figure 3.1 presents k-space traversal patterns used in conventional imaging. Each row of the k-space is acquired after one RF excitation pulse. The number of rows multiplied by the number of slices (z-axis) determines the total scan time.

The duration of standard single structural MRI acquisition is approximately 5 minutes. Usually, several scans of different modalities or a sequence of scans are acquired such that the overall scan time is much longer. The lengthy imaging process reduces patient comfort and is more vulnerable to motion artifacts. In cases where motion is inevitable, e.g., fetal MRI, cardiac cine, and lungs imaging, scan time must be significantly shortened, otherwise the produced images might be useless. Moreover, in dynamic MRI sequences, acquisition must be brief such that the temporal resolution of the sequence would allow capturing significant temporal changes, e.g., instantaneous increment of the contrast-enhanced material concentration in DCE-MRI or differences in hemodynamic response expressed in fMRI [1].

A straight forward reduction of the scan time can be obtained by sampling fewer slices, thus reducing the spatial resolution in the z-axis. Spatial distances between adjacent slices of fetal MRI or fMRI, for example, are often as large as 0.5 centimeters. Therefore, a significant portion of the potential input is not covered through imaging. On the other hand, under-sampling in the x-y domain leads to aliasing, as predicted by the Nyquist sampling theorem.

Numerous research groups as well as leading MRI scanner manufacturers make significant efforts to accelerate the MRI acquisition process. Hardware solutions allow parallel imaging by using multiple coils [2] to sample k-space data. There exist two major approaches [3] that are currently implemented in commercial MRI machines. Both reconstruct an image from the under-sampled k-space data provided by each of the coils. The sensitivity encoder (SENSE) transforms the partial k-spaces into images, then merges the resulting aliased images into one coherent image [4]. The GeneRalized Autocalibrating Partial Parallel Acquisition (GRAPPA) techniques [5] operate on signal data within the complex frequency domain before the IFFT.

The compressed sensing (CS) technique [6] allows efficient acquisition and re-

construction of a signal with fewer samples than the Nyquist-Shannon sampling theorem requires, if the signal has sparse representation in a known transform domain. Using CS for MRI reconstruction by sampling a small subset of the k-space grid had been proposed in [7]. The underlying assumption is that the undersampling is random, such that the zero-filled Fourier reconstruction exhibits incoherent artifacts that behave similarly to additive random noise. The smooth reconstruction obtained by the CS may result in a loss of noise-like texture that can be anatomically meaningful.

2.2.2 Machine Learning Methods

Recently, different data driven machine learning techniques were used for MRI acceleration. Ravishankar and Bresler [8] proposed a dictionary-based learning method, to exploit the sparsity of overlapping image patches emphasizing local structure, to reconstruct MRI from highly under-sampled k-space data. In [9], this concept was extended using spatio-temporal patches for dynamic MRI reconstruction. Compressed manifold learning based on Laplacian eigenmaps has been applied in [10] for respiratory motion estimation and in [11] for cardiac MRI reconstruction.

Deep neural networks (DNN) allow implicitly learning prior information from large datasets. This has been exploited in recent works for MRI reconstruction from sub-sampled data. A convolutional neural network (CNN) was introduced in [12] to learn the mapping between zero-filled and fully-sampled MRIs. In [13], MRI acceleration by super-resolution (SR) concepts was introduced. One of the key ideas is a siamese CNN structure that allows receiving multiple inputs acquired from different viewing planes. Improved low-resolution (LR) high-resolution (HR) mapping was obtained by residual architecture. A multi-scale residual network, also known as U-net [14], was used in [15] to address the CS optimization, formulated as a residual regression problem. Hayun et el. [16], applied U-net to perform anti-aliasing of an MRI generated from sub-sampled k-space. The U-net output image is then enhanced by integrating its k-space domain with the originally acquired raw samples. U-net CNN have been also used in [17] for dynamic MRI reconstruction. The spatio-temporal characteristics of sequential MRI data was recently exploited using recurrent neural network (RNN) by [18]. All these works address the reconstruction problem in the image domain rather than the k-space domain.

The proposed framework utilizes recent advances in deep learning, while addressing MRI reconstruction directly from the k-space. Specifically, we use generative adversarial networks (GAN) [19, 20, 21]. GANs are based on the inter-play between two networks: a generator and a discriminator. The generator is capable of learning the distribution over a data-base, and sample realizations of this distribution. The discriminator is trained to distinguish between 'generated' samples and real ones. This powerful combination has been used for MRI acceleration in [22, 23], and also for generating Computed Tomography (CT)-like images from MRIs [24]. In contrast, here, the generator is used for reconstruction of the entire k-space grid from under-sampled data. Its loss is

a combination of an adversarial loss, calculated on the image domain based on the discriminator output, and a fidelity loss with respect to the fully sampled k-space. Unlike [16], the entire process, including the IFFT, is carried out and optimized in an end-to-end manner by the proposed GAN architecture. As the MRI data is complex, both real and imaginary components are incorporated in both the generator and the discriminator.

We compare the proposed method reconstructions to those obtained by using a conventional compressed sensing method CS-MRI [7] and zero-filling. In addition, for comparison, we tested two other networks: 1. CNN-L2 - a k-space generator (with same architecture as the proposed method) trained with only L2 loss; 2. IM-CNN-L2 - a network that performs on the image-space and is optimized to remove the artifacts caused by under-sampling and zero padding. The method is applied to the Information eXtraction from Images (IXI) dataset, which is a large publicly available dataset, including 3T and 1.5T scans. Raw data is down-sampled such that only 40%, 25%, and 16.6% of the raw samples of each scan are used. To assess the clinical usability of the reconstructed images we perform an extended validation, composed of anatomical quantitative measurements. In addition to the commonly used PSNR and Structure Similarity Index (SSIM), we perform tissue segmentation and compare the results to those obtained by using the original fully-sampled images. Segmentation compatibility, measured in terms of Dice scores and Hausdorff distances, demonstrate the quality of the proposed MRI reconstruction with respect to other methods. The proposed method outperforms all other tested methods in all evaluation metrics. Moreover, compared to the widely used CS-MRI, the proposed method is faster by two orders of magnitude. In fact, it can work in real time, and is therefore practical for routine uses.

Chapter 3

Method

3.1 K-space

3.1.1 Notations and Problem Definition

We aim to address MRI reconstruction from under-sampled k-space data using DNN. We hereby present the proposed methodology as well as the network architecture. Let $\boldsymbol{u}$ denote the desired signal, a 2D MR image, obtained by IFFT of the complex k-space signal s_0, i.e.,

$$\boldsymbol{u} = F^H s_0 \tag{3.1}$$

where H is the Hermitian transpose operation and F^H is an orthonormal 2D IFFT operator, such that $F^H F = I$. Partial k-space sampling, results in MR images that are subject to artifacts and aliasing. Figure **3.1**.d exemplifies artifacts caused by under-sampling and zero padding in the phase axis using a Gaussian mask (40%).

(a) raw k-space (b) original (c) sampling mask (d) reconstructed

Figure 3.1: Under-sampling artifacts: the arrows illustrate Cartesian sampling

Let $s_p = M_p \odot s_0$ denote the under-sampled k-space, where $\odot$ denotes element-wise multiplication and M_p is the sampling mask. Given a model f, defined by the set of parameters Θ, our goal is to estimate the missing k-space

samples such that:

$$\Theta = \arg\min_{\Theta} L(F^H f\left(s_p;\Theta\right), \boldsymbol{u}) \tag{3.2}$$

where $L\left(\cdot\right)$ is a loss function. L2 loss is widely used for natural image reconstruction. However, as mentioned in [21], L2 minimization provides a blurry solution, resulting in a loss of fine details. In our case, where spatial frequencies in the form of k-space are considered, using this loss alone may be inappropriate, resulting in poor reconstruction. To address this problem, we propose using in addition an adversarial loss, based on a GAN.

3.1.2 K-space GAN

We train our model using an adversarial strategy, as described in [19, 20]. This method is based on a generator G, which gets noise z with uniform distribution $p_u\left(z\right)$ as input and generates samples from the observed data distribution. A discriminator D is trained to distinguish between "true" examples from the data and generated ("fake") examples from G. During the training process, we optimize G to maximize the discriminator's probability of error. Simultaneously, D is enhanced and provides more accurate predictions.

Let s_0 denote a true k-space sample from the distribution $p_r\left(s_0\right)$. The following optimization process can be described by two-players min-max value function $V(G,D)$:

$$\min_{G}\max_{D} V(G,D) = \mathbb{E}_{\mathrm{s_0\sim p_r(s_0)}}\log\left[\mathrm{D}\left(\mathrm{s_0}\right)\right] + \mathbb{E}_{\mathrm{z\sim p_u(z)}}\log\left[1-\mathrm{D}\left(\mathrm{G(z)}\right)\right] \tag{3.3}$$

where z is a noise drawn from a uniform distribution $p_u\left(z\right)$.

In equilibrium, the generator G imitates the true data, which is in our case the missing k-space samples. The generator loss is a linear combination of L2 fidelity loss and an adversarial loss as follows:

$$\begin{aligned} L_G = \alpha\cdot E_{\mathrm{z\sim p_u(z)}}\log\left[1-\mathrm{D}\left(\mathrm{F}^{-1}\left(\hat{\mathrm{s}}_0\right)\right)\right] + \\ \beta\cdot||\left(1-M_p\right)\odot\left(\hat{s}_0-s_0\right)||_2^2 \end{aligned} \tag{3.4}$$

where $\hat{s}_0$ is the estimated k-space and $\alpha=1$, $\beta=1$ are hyperparameters tuned by a cross-validation process. We choose to use the Wasserstein distance as the GAN loss function (WGAN) [25], which improves the training stability and provides better reconstruction compared to the classic GAN loss we used in our preliminary work [26].

3.1.3 System Architecture

The proposed framework is illustrated in Figure 3.2.

Figure 3.2: Framework architecture: G and D are the generator and discriminator networks, respectively. See text for details.

The generator's input is a two-channel signal representing the real and the imaginary parts of the partially sampled k-space image, s_p. Each missing sample (i, j) is initialized by uniform i.i.d. noise as follows:

$$G_{in}(i,j) = s_{p_{i,j}} + (1 - M_p)_{i,j} z_{i,j} \tag{3.5}$$

G produces reasonable k-space samples from a given samples and noise distribution $p_u(z)$. In order to utilize the remaining sampled data, s_p, and estimate only the missing samples we use a residual network [27] as in [13], such that:

$$\hat{s}_0 = s_p + (1 - M_p) \odot G_{out} \tag{3.6}$$

where G_{out} is the generator output. The discriminator's two-channel (real and imaginary) input is either the reconstructed MR image obtained by IFFT of the generator's estimated or the true MR image obtained by IFFT of the original fully-sampled k-space data. By that, we are incorporating the IFFT into the network performing end-to-end optimization. The discriminator's output defines the WGAN loss.

We use a standard discriminator architecture, composed of convolutional layers, batch normalization, and leaky-ReLU as suggested in [20]. For the generator, we compose a dedicated architecture, detailed in Section 4.1. Both the discriminator and the generator inputs are two-channel image for the real and imaginary components. Figure 3.3 illustrates their architectures. During the training phase, k_d discriminator update steps are carried out for every step of the generator. The training ends when the discriminator cannot accurately distinguish between the reconstructed ("fake") and the true inputs.

Figure 3.3: Networks architecture. Both the discriminator and the generator inputs are two-channel image for the real and imaginary components. The number above each layer represents the output channels. All convolutional kernels are 3×3.

Chapter 4

Experimental Results

4.1 Random Gaussian Sampling

The training data consists of 500 3D brain MRI (T1-weighted) scans of different patients from the IXI dataset.[1] The data was acquired by three MR machines, Philips 1.5T, 3T and GE 1.5T using different imaging protocols. From each 3D volume we extracted 93 2D saggital slices. All slices were interpolated to 256×256 pixels. We used $37.2k$ (80%) 2D slices for training and $9.3k$ (20%) for testing (100 3D volumes). In order to create k-space images for training, inverse orthonormal 2D FFT was applied to the fully-sampled MR images. Unlike our preliminary work [26], which employed a 1D Cartesian sampling mask along the phase axis; here, we perform k-space sampling using a 2D random Gaussian mask. By this, we exploit the sparsity in both phase and frequency dimensions and are therefore able to accommodate higher sampling factors (Figure 4.1). MRI reconstruction is performed using 40%, 25%, and 16.6% of the original k-space data.

The generator is composed of 5 blocks of CONV-BatchNorm-ReLU, with output channels $16, 32, 64, 32, 8$, respectively (see Figure 3.3). The last CONV layer is composed of two output channels for real and imaginary components. The discriminator is composed of 4 blocks of CONV-Pool-BatchNorm-LReLU with output channels $8, 16, 32, 16$ and one fully-connected layer. All CONV layers' kernel size is 3×3. All weights were initialized using Xavier [28]. We used the RMSprop [29] solver with a fixed learning rate of 5e−6 and set k_d to 1.

We compared the proposed method reconstructions to those obtained by using a conventional compressed sensing method-CS-MRI [7] and zero-filling. In addition, we tested two other networks. To assess the use of an adversarial loss, we trained a k-space generator, with the same architecture as the proposed method, with only an L2 loss (CNN-L2). We also implemented the method proposed in [16] and constructed a network IM-CNN-L2, which is applied directly

[1] `http://brain-development.org/ixi-dataset/`

to the image space rather than the k-space. Specifically, as in [16], we used the U-net architecture [14] and applied an FFT on each of its output images to get an estimated k-space. The complete k-space is then recovered by integrating the estimated samples with the sampled raw data according to the sampling mask. The same sampling masks were used for all methods.

To demonstrate the clinical usability of the reconstructed images we performed an extended validation, composed of anatomical quantitative measurements. In addition to the commonly used PSNR and Structure Similarity Index (SSIM) [30] metrics, we performed tissue segmentation and compared the results to those obtained by using the original fully-sampled images. The segmentation compatibility is measured in terms of Dice scores and Hausdorff distances.

(a) Factor 2.5 (40%) (b) Factor 4 (25%) (c) Factor 6 (16.66%)

Figure 4.1: $2D$ Gaussian sampling masks

4.1.1 PSNR and SSIM

PSNR and SSIM are the most commonly used measures for evaluation of image reconstruction quality. PSNR measures the Mean Squared Error (MSE) between the fully-sampled MR image and the reconstructed image. Quantitative evaluation is presented in Table 4.1 and a graphical visualization in Figure 4.2. The PSNR is calculated for the entire image without masking. Note that the proposed method outperforms the others. To support this, we also calculated the two-sample t-Test for the table's scores. The differences are statistically significant (p-values < 0.005) for all pairwise comparisons between the results of our method and the others.

PSNR Method	Factor 2.5 (40%)	Factor 4 (25%)	Factor 6 (16.66%)
Zero-filled	32.555 ± 2.294	26.791 ± 1.856	16.781 ± 1.737
CS-MRI	39.053 ± 1.750	33.088 ± 1.929	26.491 ± 2.587
IM-CNN-L2	39.663 ± 1.934	35.042 ± 2.042	28.134 ± 2.181
CNN-L2	39.394 ± 1.985	33.829 ± 2.034	31.403 ± 2.040
Proposed	**40.211 ± 1.902**	**35.133 ± 1.870**	**32.040 ± 2.110**

Table 4.1: Error in PSNR, without masking

Figure 4.2: PSNR without masking error-bar

An additional validation of our model is performed by calculating the PSNR measure on different brain tissues. We applied the commonly used FAST [31] image segmentation algorithm on the original fully-sampled MR image and then used it for calculating the masked-PSNR on gray matter (GM), white matter (WM), and the Cerebrospinal Fluid (CSF). Results are presented in Table 4.2. The differences are statistically significant with p-values < 0.0005 for CS-MRI and zero-filled for sampling factors 2.5 and 4, and also for IM-CNN-L2 with p-values < 0.0001 for sampling factor 6.

PSNR	Factor 2.5 (40%)			Factor 4 (25%)			Factor 6 (16.66%)		
Method	White	Gray	CSF	White	Gray	CSF	White	Gray	CSF
Zero-filled	41.49±3.48	36.73±3.32	38.54±2.45	37.27±2.18	34.37±2.52	36.36±1.87	24.88±1.62	24.98±1.91	29.33±2.08
CS-MRI	44.37±4.20	40.04±3.84	40.74±3.16	41.65±3.12	37.24±2.99	39.11±2.43	35.59±3.15	33.11±2.91	36.64±2.26
IM-CNN-L2	45.66±4.60	41.06±4.27	41.51±3.58	43.47±3.22	39.36±3.19	40.16±2.51	37.34±3.19	34.18±2.85	36.61±2.10
CNN-L2	45.72±4.70	41.59±4.16	41.68±3.69	43.26±2.95	40.15±3.15	40.24±2.47	41.11±2.48	39.27±2.65	39.10±1.93
Proposed	**46.20±4.77**	**41.94±4.34**	**41.61±2.45**	**43.85±3.16**	**40.11±3.25**	**40.49±2.57**	**41.91±2.77**	**39.01±2.88**	**39.36±2.17**

Table 4.2: Error in PSNR, with masking

We next used SSIM, which is a full reference metric, and is considered a perception-based model, to measure the similarity between the original fully-sampled images and the reconstructed ones. It quantifies the degradation caused by sampling artifacts based on pixels' inter-dependencies in the image. SSIM scores for different sampling factors, for all methods compared, are presented in Table 4.3. The advantage of our method is clearly shown. All pairwise comparisons between the results obtained by our method and the others are statistically significant (p-values < 0.0001) for all sampling factors.

SSIM Method	Factor 2.5 (40%)	Factor 4 (25%)	Factor 6 (16.66%)
Zero-filled	0.698 ± 0.0452	0.590 ± 0.040	0.255 ± 0.021
CS-MRI	0.865 ± 0.029	0.757 ± 0.037	0.617 ± 0.030
IM-CNN-L2	0.884 ± 0.030	0.796 ± 0.036	0.606 ± 0.023
CNN-L2	0.885 ± 0.030	0.749 ± 0.043	0.682 ± 0.042
Proposed	**0.917 ± 0.019**	**0.818 ± 0.034**	**0.726 ± 0.038**

Table 4.3: SSIM for different sampling factors

While PSNR and SSIM are widely used, these measures are not sensitive enough to differences in fine details and complex anatomical structure topologies. For example, a model can provide a good reconstruction in the sense of PSNR, yet with blurry edges and therefore loss of high frequency information. This may significantly affect diagnostic usability. We therefore further compare segmentation quality between the fully-sampled and the reconstruction images. Results of the analysis obtained for our method and all others are presented next.

4.1.2 Brain Extraction - Skull Stripping

Brain extraction (skull stripping) is an algorithm that delineates the brain boundary (Figure 4.3). It is a necessary preliminary step for brain analysis algorithms. We examine the different reconstruction methods by applying the Brain Extraction Tool (BET) [32] to both MR reconstructed and original (fully-sampled) images. Compatibility of the skull stripping results between the original and the generated datasets is measured via the Modified Hausdorff Distance (MHD) [33], defined as follows. Let $C_1, C_2 \in R^2$ denote the brain boundaries extracted by the BET algorithm, using the original and the reconstructed scans, respectively. The MHD between C_1 and C_2 is:

$$\begin{aligned} MHD(C_1, C_2) = 0.5 \left[\frac{1}{|C_1|} \sum_{\sigma_1 \in C_1} d(\sigma_1, C_2) + \frac{1}{|C_2|} \sum_{\sigma_2 \in C_2} d(\sigma_2, C_1) \right] \\ d(\sigma_1, C_2) = \min_{\sigma_2 \in C_2} ||\sigma_1 - \sigma_2|| \\ d(\sigma_2, C_1) = \min_{\sigma_1 \in C_1} ||\sigma_2 - \sigma_1|| \end{aligned} \tag{4.1}$$

where σ_1, σ_2 are points on the contours C_1, C_2, respectively (Figure 4.3.c) and $|C_i|,\ i = 1, 2$, are the respective cardinalities. MHD values for all tested methods and sampling ratios are presented in Table 4.4. The MHD results obtained by our methods are better than all others and are significant (p-values < 0.0001) for CS-MRI and zero-filled.

MHD Method	Factor 2.5 (40%)	Factor 4 (25%)	Factor 6 (16.66%)
Zero-filled	1.111 ± 0.563	2.617 ± 1.214	3.121 ± 1.279
CS-MRI	0.701 ± 0.511	1.447 ± 1.027	3.114 ± 1.617
IM-CNN-L2	0.541 ± 0.388	0.724 ± 0.394	1.902 ± 1.437
CNN-L2	0.420 ± 0.270	0.715 ± 0.561	1.083 ± 1.052
Proposed	**0.391 ± 0.250**	**0.617 ± 0.306**	**1.050 ± 1.033**

Table 4.4: MHD - brain extraction

(a) MR image (b) Brain extraction (c) Hausdorff distance

Figure 4.3: Example of brain extraction

4.1.3 Brain Segmentation

We next tested and compared segmentation quality applying a brain tissue partitioning algorithm (FAST [31]) to the reconstructed images. Segmentation results were compared to those obtained for the original fully-sampled scans. We believe that these comparisons are indicative of the diagnostic quality of the reconstructed MRIs as tissue segmentation (WM, GM, and CSF) plays a key role in clinical brain analysis. We used the Dice score [34] to measure the compatibility between the segments obtained for the original (S_{REF}) and the reconstructed (S_{REC}) images, defined as follows:

$$DICE(S_{\text{REF}}, S_{\text{REC}}) = 2\frac{S_{\text{REF}} \cap S_{\text{REC}}}{S_{\text{REF}} + S_{\text{REC}}} \tag{4.2}$$

Dice scores are between zero and one, where one indicates complete overlap between the tested segments. Therefore, higher dice scores imply higher compatibility between reconstructed and original images. Dice scores for all tested methods and sampling factors are presented in Table 4.5. As for previous tests, our method achieves the best scores for all brain tissues and all sampling factors. Statistically significant values (p-values < 0.0005) were found for all methods using sampling factor of 2.5. In addition to the CS-MRI and zero-filled, the differences are statistically significant with p-values < 0.0001 for CNN-L2 using sampling factor 4 and for IM-CNN-L2 using sampling factor 6.

DICE	Factor 2.5 (40%)			Factor 4 (25%)			Factor 6 (16.66%)		
Method	White	Gray	CSF	White	Gray	CSF	White	Gray	CSF
Zero-filled	0.882±0.08	0.826±0.05	0.796±0.05	0.718±0.10	0.644±0.05	0.627±0.05	0.599±0.12	0.466±0.07	0.050±0.06
CS-MRI	0.942±0.05	0.910±0.03	0.868±0.03	0.871±0.10	0.805±0.08	0.770±0.05	0.708±0.14	0.639±0.08	0.621±0.07
IM-CNN-L2	0.947±0.02	0.917±0.02	0.887±0.02	0.903±0.03	0.853±0.02	0.828±0.02	0.741±0.14	0.681±0.08	0.674±0.07
CNN-L2	0.948±0.02	0.919±0.02	0.891±0.02	0.888±0.06	0.836±0.04	0.813±0.03	0.836±0.09	0.770±0.06	0.747±0.06
Proposed	**0.954±0.01**	**0.928±0.02**	**0.900±0.002**	**0.903±0.06**	**0.858±0.04**	**0.833±0.03**	**0.851±0.09**	**0.789±0.06**	**0.767±0.06**

Table 4.5: Segmentation Dice scores for different sampling ratios

MR reconstruction results for factor 6 (16.66%) as well as their respective brain segmentation results are visually presented in Figure 4.4 for a single representative scan. The left most column corresponds to the original image and columns 2-5 correspond to the reconstructed images using Zero-filled, CS-MRI, IM-CNN-L2, and the proposed method, respectively. The first row presents a $2D$ slice of the original and the reconstructed MRIs. The second row presents the zoom-in of brain regions in the red frames shown in the first row. The third row presents the corresponding brain extraction. The fourth row presents the automatic tissue segmentation where, dark blue is used for WM, blue for GM and light blue for CSF. Note that MRIs reconstructed using the suggested adversarial method have stronger contrast and no significant aliasing or artifacts. In addition, the brain extraction and tissue segmentation results are more compatible with those obtained from the original fully-sampled MRIs. Additional examples of different cases, using sampling factors of 4 (25%), and 2.5 (40%) are shown in Figures 4.5, 4.6, respectively.

Figure 4.4: Examples of reconstructed MR images from under-sampled k-space - Factor 6.

Figure 4.5: Examples of reconstructed MR images from under-sampled k-space - Factor 4.

Figure 4.6: Examples of reconstructed MR images from under-sampled k-space - Factor 2.5.

4.2 Phase Encoding Sampling

The training data consists of 55 $3D$ brain MRI (T1) scans of different patients, acquired by 1.5T MR machine with resolution of 256×256 pixels. We used $3k$ 2D slices for training and $1.2k$ for testing. In order to create k-space images for training, inverse orthonormal 2D FFT is applied to the fully-sampled MR images. We sample the k-space images along the phase axis using Cartesian binary mask, while keeping 5% from the centered phases, which is the area that contains 95% of the energy. Data augmentation is created by random offsets of the proposed mask and image flipping. This leads us to reconstruction of the MR image from only 52% of the original k-space data.

The generator is composed of 5 blocks of CONV-BatchNorm-ReLU, with output channels $16, 32, 64, 32, 8$, respectively. The last layer is CONV with two outputs channels (for real and imaginary parts). The discriminator is composed of 4 blocks of CONV-Pool-BatchNorm-LReLU with output channels $8, 16, 32, 16$ and one fully-connected layer. All CONV layers kernel size is 3×3. All weights

was initialized by Xavier [28]. We used Adam solver with fixed learning rate of 0.0005 and set k_g to 20.

We compare the proposed method to reconstruction results obtained by using a conventional compressed sensing method CS-MRI [7] and Zero-filling. In addition, we trained a generator (G) using only L2 loss (CNN-L2). The same sampling mask were used for all cases. Visual comparison is shown in Figure 4.7. MRIs reconstructed by using the suggested adversarial loss have stronger contrast and no significant aliasing or artifacts. Quantitative evaluation using PSNR values is presented in Table 4.6. Note that the proposed method outperforms the other methods.

PSNR	Mean	std
Zero-filled	30.48	0.13
CS-MRI	36.78	0.4
CNN-L2	37.12	0.82
Proposed	**37.95**	0.6

Table 4.6: Error in PSNR

Figure 4.7: Examples of reconstructed MR images from under-sampled k-space.

Chapter 5

Summary and Discussion

5.1 Discussion

The proposed deep learning framework for MRI reconstruction from highly under-sampled k-space (up to 16.66%), can be used as the foundation of a real-time, software-only solution for MRI acceleration. The inter-play between the generator and the discriminator plays a key role here. The generator is trained directly on the k-space, thus avoiding aliasing correction (due to under-sampling) as is done in other methods. Instead, the k-space distribution, and not its IFFT, is learned and generated. Moreover, combining both adversarial and fidelity loss in the DNN training, the method suggested provides high PSNR reconstructions for large and diverse brain MRI datasets acquired by different MRI machines (1.5T and 3.T) and different imaging protocols.

Until just a few years ago, CS techniques were considered the state-of-the-art for MRI reconstruction. Recently, deep learning methods, capable of learning imaging samples distribution from training datasets, were shown to obtain reconstruction with comparable or even higher quality. The strength of the proposed DNN is three-fold, combining direct k-space reconstruction, GAN architecture, and end-to-end optimization. This allows anatomically meaningful reconstruction with high fidelity to the original scan. Extensive comparative analysis shows that our method outperforms all other tested MRI reconstruction methods (CS and other DNN-based) for various quality metrics, showing more robustness to high acceleration factors.

A key advantage of the proposed method is its short runtime at test phase (less than 10 msec per slice), which is almost two orders of magnitude faster than CS. This makes it applicable for routine uses. As training can be performed off-line, prior to the acquisition sessions, the fact that it lasts a few days has no practical relevance.

Assessing clinical usability of the reconstructed data is takes a considerable part in the quality analysis performed as part of this research. We chose to test it by using standard tissue segmentation measures (such as Dice and MHD), as

we argue that compatibility of the extracted brain regions with those obtained by using the original fully-sampled MRIs indicate fine detail reconstructions that have apparent diagnostic significance.

Nevertheless, pathological cases were not considered in this study and to the best of our knowledge are seldom addressed. We believe that high quality reconstruction in the presence of brain atrophies is the true success criterion. In fact, addressing this challenge and measuring performances in this sense, should be the gold standard in future benchmarks.

5.1.1 Acknowledgment

This study was partially supported by The Israel Science Foundation (1638/16); The Israel Defense Forces (IDF) Medical Corps and Directorate of Defense Research & Development, Israeli Ministry of Defense (IMOD DDR&D); Israel Ministry of Science, Technology and Space (63551).

Bibliography

[1] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, K. Uğurbil, Multiband multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri, Magnetic Resonance in Medicine 63 (5) (2010) 1144–1153.

[2] P. B. Roemer, W. A. Edelstein, C. E. Hayes, S. P. Souza, O. Mueller, The nmr phased array, Magnetic resonance in medicine 16 (2) (1990) 192–225.

[3] A. Deshmane, V. Gulani, M. A. Griswold, N. Seiberlich, Parallel mr imaging, Journal of Magnetic Resonance Imaging 36 (1) (2012) 55–72.

[4] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al., Sense: sensitivity encoding for fast mri, Magnetic resonance in medicine 42 (5) (1999) 952–962.

[5] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, A. Haase, Generalized autocalibrating partially parallel acquisitions (grappa), Magnetic resonance in medicine 47 (6) (2002) 1202–1210.

[6] D. L. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306.

[7] M. Lustig, D. Donoho, J. M. Pauly, Sparse mri: The application of compressed sensing for rapid mr imaging, Magnetic resonance in medicine 58 (6) (2007) 1182–1195.

[8] S. Ravishankar, Y. Bresler, Mr image reconstruction from highly undersampled k-space data by dictionary learning, Medical Imaging, IEEE Transactions on 30 (5) (2011) 1028–1041.

[9] J. Caballero, A. N. Price, D. Rueckert, J. V. Hajnal, Dictionary learning and time sparsity for dynamic mr data reconstruction, Medical Imaging, IEEE Transactions on 33 (4) (2014) 979–994.

[10] M. Usman, G. Vaillant, D. Atkinson, T. Schaeffter, C. Prieto, Compressive manifold learning: Estimating one-dimensional respiratory motion directly from undersampled k-space data, Magnetic Resonance in Medicine 72 (4) (2014) 1130–1140.

[11] K. K. Bhatia, J. Caballero, A. N. Price, Y. Sun, J. V. Hajnal, D. Rueckert, Fast reconstruction of accelerated dynamic mri using manifold kernel regression, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, Springer, 2015, pp. 510–518.

[12] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, D. Liang, Accelerating magnetic resonance imaging via deep learning, in: Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on, IEEE, 2016, pp. 514–517.

[13] O. Oktay, W. Bai, M. Lee, R. Guerrero, K. Kamnitsas, J. Caballero, A. de Marvao, S. Cook, D. OARegan, D.Regan, D. Rueckert, Multi-input cardiac image super-resolution using convolutional neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 246–254.

[14] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[15] D. Lee, J. Yoo, J. C. Ye, Deep residual learning for compressed sensing mri, in: Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, IEEE, 2017, pp. 15–18.

[16] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, J. K. Seo, Deep learning for undersampled mri reconstruction, arXiv preprint arXiv:1709.02576.

[17] C. M. Sandino, N. Dixit, J. Y. Cheng, S. S. Vasanawala, Deep convolutional neural networks for accelerated dynamic magnetic resonance imaging, preprint.

[18] C. Qin, J. Schlemper, J. Caballero, A. Price, J. V. Hajnal, D. Rueckert, Convolutional recurrent neural networks for dynamic mr image reconstruction, arXiv preprint arXiv:1712.01751.

[19] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in neural information processing systems, 2014, pp. 2672–2680.

[20] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434.

[21] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, A. A. Efros, Context encoders: Feature learning by inpainting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2536–2544.

[22] S. Yu, H. Dong, G. Yang, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, D. Firmin, et al., Deep de-aliasing for fast compressive sensing mri, arXiv preprint arXiv:1705.07137.

[23] M. Mardani, E. Gong, J. Y. Cheng, S. Vasanawala, G. Zaharchuk, M. Alley, N. Thakur, S. Han, W. Dally, J. M. Pauly, et al., Deep generative adversarial networks for compressed sensing automates mri, arXiv preprint arXiv:1706.00051.

[24] D. Nie, R. Trullo, C. Petitjean, S. Ruan, D. Shen, Medical image synthesis with context-aware generative adversarial networks, arXiv preprint arXiv:1612.05362.

[25] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein gan, arXiv preprint arXiv:1701.07875.

[26] O. Shitrit, T. R. Raviv, Accelerated magnetic resonance imaging by adversarial neural network, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2017, pp. 30–38.

[27] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[28] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks., in: Aistats, Vol. 9, 2010, pp. 249–256.

[29] T. Tieleman, G. Hinton, Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude, COURSERA: Neural networks for machine learning 4 (2) (2012) 26–31.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing 13 (4) (2004) 600–612.

[31] Y. Zhang, M. Brady, S. Smith, Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm, IEEE transactions on medical imaging 20 (1) (2001) 45–57.

[32] S. M. Smith, Fast robust automated brain extraction, Human brain mapping 17 (3) (2002) 143–155.

[33] M.-P. Dubuisson, A. K. Jain, A modified hausdorff distance for object matching, in: Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on, Vol. 1, IEEE, 1994, pp. 566–568.

[34] L. R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.

תקציר

אחד האתגרים הגדולים בתחום של דימות תהודה מגנטי (MRI) הוא האצת הרכישה לצרכים קליניים. מעבר לשיפור בנוחות המטופלים והוזלת עלויות, סריקות מהירות יותר הכרחיות לבדיקות אשר רגישות לשינויים בזמן, כמו למשל סריקות לב, ריאות ועוברים. סריקות אלו רגישות לתנועות של המטופל ובכך איכות הסריקה נפגעת.

יתר על כך, רכישה מהירה יתר תשפר רזולוציה זמנית בסריקות דינמיות, כמו fMRI ו-DC-MRI. השיטות הקיימות כיום אכן מקצרות את זמן הסריקה באיכות יחסית טובה, אך לעיתים אורכות זמן רב ואפילו מייקרות את החומרה הנדרשת לסריקה.

אנחנו מציגים פתרון פרקטי תוכנתי בלבד, המבוסס על למידה עמוקה להאצת רכישת ה- MRI של המוח, וזאת תוך כדי שימור המבנה האנטומי. פתרון זה מושג על ידי דגימה חלקית של המידע הגולמי ב-MRI הנקרא גם k-space, ושחזור בעזרת רשת נוירונים אדברסית המשערכת את הדגימות החסרות. הרשת בנויה על משחק סכום אפס בין שתי רשתות שונות. האחת משערכת דגימות חסרות והשנייה מייצרת מעין פידבק-משוב על תוצאת הראשונה. יחד עם שילוב אילוצי נורמת L2, השיטה שלנו מייצרת שחזורים באיכות טובה, על מאגר מידע גדול, עבור מספר מקדמי האצה, 2.5, 4 ו-6.

על מנת לוודא שאכן התמונות המשוחזרות שומרות על המבנה האנטומי, ועל השימוש הקליני שלהן, בצענו מספר ניסויים אשר יכולים להעיד על כך. השוונו תוצרי סגמנטציה בין התמונות המשוחזרות לבין התמונה המלאה בעזרת מדד ה Dice ומדד Hausdorff distance, ואלו המחישו את איכות השחזור באמצעות השיטה המוצעת, בהשוואה לשאר השיטות המוצגות בעבודה זו.

אוניברסיטת בן גוריון בנגב

הפקולטה למדעי ההנדסה

המחלקה להנדסת חשמל ומחשבים

האצת דימות תהודה מגנטית באמצעות רשתות נוירונים יוצרות-אדברסריות

חיבור זה מהווה חלק מהדרישות לקבלת התואר "מוסמך למדעי ההנדסה" (M.Sc.)

מאת: אוהד שטרית

בהנחיית: ד"ר תמי ריקלין-רביב

חתימת הסטודנט: _____ תאריך 19/2/2018

חתימת המנחה: תמי ריקלין רביב תאריך 19/2/2018

חתימת יו"ר הועדה המחלקתית: _____ תאריך _____

תשע"ח

אוניברסיטת בן גוריון בנגב
הפקולטה למדעי ההנדסה
המחלקה להנדסת חשמל ומחשבים

האצת דימות תהודה מגנטית באמצעות רשתות נוירונים יוצרות-יריבות

חיבור זה מהווה חלק מהדרישות לקבלת התואר "מוסמך למדעי ההנדסה" (M.Sc.)

מאת: אוהד שטרית
בהנחיית: ד"ר תמי ריקלין-רביב

תשע"ח
פבואר 27