POLITECNICO DI TORINO
Repository ISTITUZIONALE

A Deep Learning Method for Optimal Undersampling Patterns and Image Recovery for MRI Exploiting Losses and Projections

Original
A Deep Learning Method for Optimal Undersampling Patterns and Image Recovery for MRI Exploiting Losses and Projections / Martinini, F.; Mangia, M.; Marchioni, A.; Rovatti, R.; Setti, G.. - In: IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING. - ISSN 1932-4553. - STAMPA. - 16:4(2022), pp. 713-724. [10.1109/JSTSP.2022.3171082]

Availability:
This version is available at: 11583/2972702 since: 2022-10-30T21:35:47Z

Publisher:
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC

Published
DOI:10.1109/JSTSP.2022.3171082

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2022 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

12 October 2024

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

A Deep Learning Method for Optimal Undersampling Patterns and Image Recovery for MRI Exploiting Losses and Projections

Filippo Martinini, Mauro Mangia, *Member, IEEE,* Alex Marchioni, *Student Member, IEEE,* Riccardo Rovatti, *Fellow, IEEE,* and Gianluca Setti, *Fellow, IEEE*

***Abstract*—Compressed Sensing was recently proposed to reduce the long acquisition time of Magnetic Resonance Imaging by undersampling the signal frequency content and then algorithmically reconstructing the original image. We propose a way to significantly improve the above method by exploiting a deep neural network to tackle both problems of frequency subsampling and image reconstruction simultaneously, thanks to the introduction of a new loss function to drive the training and the addition of a post-processing non-neural stage. Furthermore, we highlight how some of the quantities along the processing chain can be used as a proxy of the quality of the recovered image, thus allowing a self-assessment of the whole technique.**

All improvements hinge on the possibility of identifying constraints to which the final image must obey and suitably enforce them. The effectiveness of our approach is tested on real-world MRI acquisitions from the fastMRI public database and achieves an appreciable improvement in Peak Signal-to-Noise Ratio with respect to the original CS-based proposal with speed-up factors 4 and 8.

***Index Terms*—fast MRI, Compressed Sensing, Deep Neural Network, U-NET, measurement constraint.**

I. Introduction

MAGNETIC Resonance Imaging (MRI) [1] is a widely-used non-invasive technique to obtain precise pictures of human internal organs and tissues. Roughly speaking, MRI relies on the possibility of measuring the frequency response to very intense magnetic fields of hydrogen nuclei contained in the subject under test. Hence, the acquired information is a set of frequency components from which the final clinical image is obtained via inverse Fourier transform.

Unfortunately, the acquisition of an MRI scan is a complex operation, which leaves the patient in a situation of appreciable discomfort for quite a long time. Furthermore, the space where the subject is located during the test is noisy and claustrophobic, and it is not rare to have individuals refusing to perform an MRI and even more often being unable to remain still during the scan, thus creating undesired motion artefacts in the final image, which may result in need of new acquisitions.

F. Martinini, A. Marchioni, M. Mangia and R. Rovatti are with the Department of Electrical, Electronic, and Information Engineering, University of Bologna, 40136 Bologna, Italy, and also with the Advanced Research Center on Electronic Systems, University of Bologna, 40125 Bologna, Italy (e-mail: filippo.martinini@unibo.it, alex.marchioni@unibo.it, mauro.mangia2@unibo.it, riccardo.rovatti@unibo.it).

G. Setti is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: gianluca.setti@polito.it).

Several techniques have been proposed to solve this critical issue, including the implementation of an *open* MRI [2], which certainly eliminates the claustrophobic environment, but at the expense of an undesirable reduction in the resolution of the acquired image. As such, the most effective solutions rely on techniques to reduce the MRI acquisition time. In this scenario, Compressed Sensing (CS) appeared [3], shortly after its introduction [4], [5] as a methodology worth applying. CS allows reconstructing a whole image starting from its undersampled version with respect to classical Shannon-Nyquist theory, thus enabling a faster acquisition. Interesting enough, the use of CS for MRI acquisition was recently allowed by FDA [6].

In its general formulation [4], [5], CS relies on the use of a linear encoder to compress the signal or image to acquire, resulting in a low dimensional measurement vector. A decoder is necessary to invert the encoding process and recover the original information. In the MRI context [3], CS encoding is pursued by applying a binary mask that selects only the most relevant frequencies to be physically acquired, while decoding starts from the undersampled frequency content and applies denoising algorithms based on prior knowledge. The output is an image close to the fully sampled/ground truth one.

Between encoding and decoding, in MRI applications, the latter stage has been studied more extensively. Classic CS decoder approaches involve the solution of an optimization problem [3]. A wide range of approaches solve the optimization problem via regularization terms, such as total-variation penalty, sparsity-inducing norms on the coefficients of the transformed representations (i.e. wavelet [7], sparsifying dictionaries [8]), or non-local patch-based operators [9]. Using these methods results in good performances but requires high computational cost and hyper-parameter tuning, severely limiting their practical impact.

Concerning the encoder, commonly used undersampling patterns follow simple and non-data-driven schemes, e.g., Spectrum based [10], Variable density [11], Cartesian sampling [12] or uniform random undersampling [13]. For a desired increase in the acquisition rate, there are countless possible patterns one could use for image undersampling, which unfortunately lead to very different reconstruction performance. Therefore, it is of paramount importance to determine the undersampling patterns that, coupled with the decoder, result in the optimal quality of the recovered images. This direction requires implementing a "data-driven" encoder and

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

leads to an overall improvement in reconstruction performance with respect to classic CS solutions where the encoder design is agnostic about the class of signal to be acquired.

Approaches based on encoder/decoder co-design exploit neural networks (NN) to outperform classic methods and effectively combine high-dimensional sensing and signal processing. An example of this family of approaches is reported in [14], while a better performing one, referred to as Learning-based Optimization of the undersampling PattErn (LOUPE) is proposed in [15], [16], and solves the encoder and the decoder design problem at the same time by adopting a signal-aware strategy.

This work develops on LOUPE by enhancing subsampling selection and improving the use of measurements, leading to a performance improvement in terms of quality reconstruction or speed-up. In detail, LUOPE corrects the undersampling artefacts with a deep neural network trained to minimize the difference between the full-resolution image and the produced output. This mechanism ignores whether the output matches the true information content passing through the undersampling mask. LOUPE does not guarantee that the k-space of the reconstructed image is coherent with the observed measurements. To improve this match, we propose here some possible techniques:

- regularization terms that either promote similarity between the two representations in the k-space or promote similarity between the k-space frequencies that are not acquired (available at training time) and the components that the decoder creates to fill the k-space;
- a slight modification of the U-NET [17] architecture in the decoder that allows checking if the decoder output matches the ground-truth undersampling k-space;
- a new decoder structure evolving the one proposed in [15] that forces the equality between the two k-space representations;
- a post-processing stage forcing the final output to be compliant with the ground-truth k-space representation. This block operates by applying Dykstra's projection algorithm [18], [19] at inference time.

A further contribution is identifying a feature correlated with the output quality that can be used to estimate the goodness of reconstruction without the need to know the ground truth.

- Self-assessment provides each reconstructed image with a score to measure the similarity with the ground truth. This ability makes the whole framework more robust, and to the best of our knowledge, this is the first tentative of presenting such a tool in the fast-MRI.

As a final remark, the proposed framework preserves the LOUPE's characteristics: *i*) it preserves the undersampling patterns generality, e.g., it can work with both Cartesian and non-Cartesian undersampling patterns *ii*) it is designed for single-coil acquisitions (even if the multi-coil acquisition is not precluded).

The rest of the paper is organized as follows. Section II and III introduce the basics of CS and detail some related works. Section IV describes the reference acquisition architecture while the proposed frameworks are discussed in Section V. Before conclusion numerical evidences and a comparison with other methods are presented in Section VI.

II. Compressed Sensing for MRI

Compressed Sensing (CS) is a general signal acquisition scheme that allows the reconstruction of the original signal $\boldsymbol{x}$ from a number of observations that is smaller than those needed according to classic Shannon-Nyquist theory [4], [5].

It can be thought as the coupling of an *encoding* stage with a *decoding* stage. The encoder $\mathrm{Enc}(\cdot)$ that models the acquisition process is a linear operator producing the measurement vector $\boldsymbol{y} = \mathrm{Enc}(\boldsymbol{x})$. The decoder $\mathrm{Dec}(\cdot)$ expands the measurements into an estimation of the acquired signal $\hat{\boldsymbol{x}} = \mathrm{Dec}(\boldsymbol{y})$.

The *compression* property comes from the fact that the dimension of $\boldsymbol{x}$ is larger than the dimension of $\boldsymbol{y}$ and, though, in principle, this would make the exact retrieval of $\boldsymbol{x}$ impossible as $\mathrm{Enc}(\cdot)$ becomes non-injective. The recovery of the original signal can be guaranteed if prior knowledge of its structure is available.

The most common setting is the one in which $\boldsymbol{x}$ is n-dimensional while $\boldsymbol{y}$ is m-dimensional, with $m < n$, and it is known that $\boldsymbol{x}$ has a *sparse* representation, i.e., that a set of $W \geq n$ vectors $\boldsymbol{w}_0, \ldots, \boldsymbol{w}_{W-1}$ exists such that in the linear combination $\boldsymbol{x} = \xi_0 \boldsymbol{w}_0 + \xi_1 \boldsymbol{w}_1 + \cdots + \xi_{W-1} \boldsymbol{w}_{W-1}$ only a small number $k < m < n$ of coefficients ξ_j are non-zero.

If m is propoerly chosen accordingly to n and k and if $\mathrm{Enc}(\cdot)$ satisfies quite general assumptions, then the original signal can be recovered by [4], [5]

$$\hat{x} = \mathrm{Dec}(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \sum_{j=0}^{W-1} |\xi_j| \ \text{ s.t. } \ \mathrm{Enc}(\boldsymbol{x}) = \boldsymbol{y} \qquad (1)$$

where the objective function can be proved to be, in this case, a good proxy of the number of non-zero coefficients, and the constraint ensures that the recovered signal is compatible with the observed measurements.

The applicability of CS schemes to MRI comes from two key observations. First, intuitively, for CS to work, $\mathrm{Enc}(\cdot)$ must be such that each component of $\boldsymbol{y}$ depends on many components of $\boldsymbol{x}$ so that $\boldsymbol{y}$ contains information on the whole $\boldsymbol{x}$ despite being a shorter vector. This assumption is intrinsic in MRI systems as the physical quantities acquired by the sensor are the Fourier transform of the image we want to reconstruct. Second, the reduction of dimensionality implicit in $m < n$ can be seen as an undersampling, i.e., the possibility of acquiring only a subset of all the possible Fourier coefficients.

Common undersampling patterns follow simple and non-adaptive schemes such as uniform random sample skipping [13], or variable density [11] and evenly spaced Cartesian with skipped lines [12]. They have the advantage of simplicity but are not very effective.

Once one fixes the target reduction of the acquisition rate (i.e. undersampling level), the number of compatible undersampling patterns is vast and picking the right scheme may

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

significantly improve the performance of the whole acquisition chain.

Whatever the choice, the optimal encoding must be paired with a proper decoder. Hence, a further issue regards the exploitation of a priori knowledge on $\boldsymbol{x}$ to cope with the non-injective nature of undersampling. Sparsity can be still called into play when, for example, the grey-level histogram of the original image is bi-modal, revealing that most of the pixels are either black or white, thinking of one of the two colours as the *zero* filling most of the components. It is also worth stressing that more refined forms of sparsity have also been identified through wavelet or dictionary-based representations [20].

III. Compressed Sensing and Deep Neural Networks for MRI

A possible path to overcome the issue related to the straightforward application of CS to MRI relies on using neural networks specifically adapted to suit encoder-decoder architectures for CS. Proposals that separately address the encoder-decoder CS problems exist: those are algorithms that either specialize in finding the best undersampling pattern or improving the decoding stage only.

As far as decoding is concerned, [21] reports the use of a neural network to directly map the undersampled scans in the frequency domain to the reconstructed image in the pixel domain. An example of NN applied to solve, in a machine-learning-fashion, a conventional CS MRI problem was proposed in [22]. NN have also been used along with older works such as GRAPPA [23] to find new hybrid adaptations [24]. Several of the proposed solutions in the Literature exploits the so-called U-NET [17], which is a variation of a classic contracting-expanding convolutional network that differs from it by introducing a set of skip connections and concatenation layers.

Examples are [25], [26], which have been recently used as the baseline accuracy in the famous "fastMRI" competition. In [27], authors compare a net built using complex convolution layers with a net that uses classic convolution operation, demonstrating appreciable improvements. Many recent works are inspired by the successes of the adversarial neural network (GAN) [28] and present their neural architecture adaptation to the MRI reconstruction problem [29]. Examples are [30] that uses a modified cyclic loss, [31] that uses un-pared scans during training to tackle the problem of a small-size dataset, [32] that uses a Wasserstein loss, [33] that introduces the attention mechanism inside the classic GAN architecture, and [34] that carefully studies how to exploit at best all the acquired frequency regions while reducing a novel structural loss. [35], [36] also contribute to the GAN-based model research direction by respectively proposing combined innovative losses and architectural modifications. An important class of publications focuses on exploring how a multi-domain framework improves denoising [37], [38]. In addition, [39] leverages both the magnitude and the phase information content by creating two different neural networks at once, and [40] reconstructs the undersampled image exploiting the frequency domain and the pixel representations simultaneously by creating an adapted recurrent neural network. Notable works adopting recurrent-based solutions are [41], [42], where novel models are deployed to leverage the dependencies between samples at best. Interesting advancements can also be found in [43], which lays the foundation for a simultaneous combined metric optimization, and from [44], which proposes a new geometrically expressivity enhanced design technique.

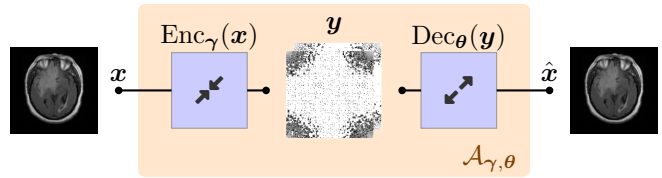


Fig. 1. The cascade of encoding and decoding as a unique autoencoding structure $\mathcal{A}_{\boldsymbol{\gamma},\boldsymbol{\theta}}(\cdot)$.

A great deal of effort has also been paid to search for the best undersampling pattern (encoder), and many data-driven inspired approaches have been presented. As an example, [45] studies the solution of a Bayesian inference problem formulation, [10] uses the power spectrum of a set of examples images as a reference to optimize the acquisition.

More recent contributions tackle the possibility to adapt encoding and decoding simultaneously as this has been demonstrated to be a profitable approach in other contexts [46]. For instance, [47] and [48] deploy a Model-based deep learning (MoDL) [49] framework (an interesting novel approach alternative to CS), reporting outstanding results. The MoDL algorithm restores the undersampled image, adapting to the input data, while the undersampling pattern is jointly optimized within the conventional NN training.

Among the novel contributions relying on data-driven encoder/decoder simultaneous optimization, LOUPE [15], [16] can be considered as the state-of-the-art CS-based approach that uses a Deep Neural Network (DNN) to speed up MRI. More specifically, in [16], authors present a regularized-training version, while in [15], the regularization is removed, and a smart rescaling function is used instead, enabling better performances. The latter is the starting point of our work, which offers a significant enhancement of the decoding stage.

IV. Basic Architecture

We assume that the encoder (decoder) depends on some parameters $\boldsymbol{\gamma}$ ($\boldsymbol{\theta}$) and that the two stages are embedded in a single generalized autoencoding structure [50] $\mathcal{A}_{\boldsymbol{\gamma},\boldsymbol{\theta}}(\cdot) = \mathrm{Dec}_{\boldsymbol{\theta}}\big(\mathrm{Enc}_{\boldsymbol{\gamma}}(\cdot)\big)$ that we interpret as a single DNN.

We assume that the original target image is normalized to be $\boldsymbol{x} \in [0,1]^{N\times N}$. The reconstructed acquisition is $\hat{\boldsymbol{x}} = \mathcal{A}_{\boldsymbol{\gamma},\boldsymbol{\theta}}(\boldsymbol{x})$ and the autoencoder is trained to be as close as possible to the identity to set $\hat{\boldsymbol{x}} = \boldsymbol{x}$.

The encoder is characterized by parameters $\boldsymbol{\gamma}$, which control the generation of a random undersampling mask $\boldsymbol{M}_{\boldsymbol{\gamma}} \in [0,1]^{N\times N}$ so that

$$\boldsymbol{y} = \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{x}) = \boldsymbol{M}_{\boldsymbol{\gamma}} \circ \mathcal{F}(\boldsymbol{x})$$

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

where $\circ$ indicates the Hadamard entrywise product between matrices, and $\mathcal{F}(\cdot)$ indicates the two-dimensional Fourier transform.

Actual undersampling is only achieved by using a binary mask $\boldsymbol{M_\gamma} \in \{0,1\}^{N\times N}$. Hence, to have a speed-up of R, one needs a mask $\boldsymbol{M_\gamma}$ whose amount of non-zero elements is exactly rN^2, where $r = 1/R$ and N^2 are the total entries.

$\mathrm{Enc}_{\boldsymbol{\gamma}}(\cdot)$ generates $\boldsymbol{M_\gamma}$ differently depending on whether $\mathcal{A}_{\boldsymbol{\gamma},\boldsymbol{\theta}}(\cdot)$ is under training, or is used for real acquisitions. In both cases $\boldsymbol{M_\gamma}$ depends on a matrix $\boldsymbol{T}(\gamma) \in [0,1]^{N\times N}$ that is random at training time and fixed at inference time.

To compute $\boldsymbol{T}(\gamma)$, define the sigmoid function

$$\sigma_t(\cdot) = \frac{1}{1+\exp(-t\cdot)} \tag{2}$$

parametrized by t and set $\boldsymbol{S} = \sigma_t(\gamma) \in [0,1]^{N\times N}$ where $\boldsymbol{\gamma} \in \mathbb{R}^{N\times N}$ and $\sigma_t(\cdot)$ is applied to each component separately. The value of t is not critical as this mapping is only used to guarantee $\boldsymbol{\gamma}$ to contain generic real numbers mapped into the $[0,1]$ range.

Once $\boldsymbol{S}(\boldsymbol{\gamma})$ is given, one computes its average $\langle \boldsymbol{S}(\boldsymbol{\gamma})\rangle = \frac{1}{N^2}\sum_{i,j}\boldsymbol{S}_{i,j}(\boldsymbol{\gamma})$ and sets

$$\boldsymbol{T}(\boldsymbol{\gamma}) = \begin{cases} \frac{r}{\langle S(\boldsymbol{\gamma})\rangle}\boldsymbol{S}(\boldsymbol{\gamma}) & \text{if } \langle \boldsymbol{S}(\boldsymbol{\gamma})\rangle \geq r \\ \boldsymbol{1} - \frac{1-r}{1-\langle \boldsymbol{S}(\boldsymbol{\gamma})\rangle}\left(\boldsymbol{1}-\boldsymbol{S}(\boldsymbol{\gamma})\right) & \text{if } \langle \boldsymbol{S}(\boldsymbol{\gamma})\rangle < r \end{cases} \tag{3}$$

where $\boldsymbol{1}$ is the $N\times N$ matrix with all unit entries and the formula ensures that $\langle \boldsymbol{T}(\gamma)\rangle = r$.

As long as $\mathcal{A}_{\boldsymbol{\gamma},\boldsymbol{\theta}}(\cdot)$ is under training, masks are generated by drawing a random matrix $\boldsymbol{U} \in [0,1]^{N\times N}$ with independent and uniformly distributed entries and setting

$$\boldsymbol{M_\gamma} = \sigma_s\left(\boldsymbol{T}(\boldsymbol{\gamma}) - \boldsymbol{U}\right) \tag{4}$$

for a suitable parameter s. The magnitude of s is important since there is a trade-off between smaller values easing back-propagation and larger values that tend to guarantee substantially binary masks.

In its most generic implementation, the above scheme generates unconstrained masks that are often indicated as non-Cartesian subsampling trajectories. Yet, [15] shows that a scheme like LOUPE is also able to cope with Cartesian trajectories that subsample the k-space with subsequent parallel lines. Since we improve LOUPE relying on the same basic principles, our framework straightforwardly applies to non-Cartesian trajectories if $\boldsymbol{\gamma}$ is thought as an N dimensional array in which the i-th element encodes the probability of sensing the entire i-th line of the k-space.

The mask generation mechanism is the one in [15] and can be implemented using four layers. The first layer produces $\boldsymbol{S}(\boldsymbol{\gamma}) = \sigma_t(\boldsymbol{\gamma})$. The second layer computes $\boldsymbol{T}(\boldsymbol{\gamma})$ from $\boldsymbol{S}(\boldsymbol{\gamma})$ by (3). The last two layers generate the random matrix $\boldsymbol{U}$ and apply (4). We represent the whole mechanism as a unique block named mask generator.

Contrary to the original proposal, once $\boldsymbol{T}(\gamma)$ is fixed by training and the model is used for acquisitions, we use the deterministic binary mask defined componentwise as

$$(\boldsymbol{M_\gamma})_{j,k} = \begin{cases} 0 & \text{if } \boldsymbol{T}_{j,k}(\boldsymbol{\gamma}) < 0.5 \\ 1 & \text{if } \boldsymbol{T}_{j,k}(\boldsymbol{\gamma}) \geq 0.5 \end{cases} \quad j,k = 0,\ldots,N-1 \tag{5}$$

Note that in (5), $\boldsymbol{M_\gamma}$ is the instance with the highest probability among all the possible training-time masks for $s \to \infty$. This aspect is also reported in Fig. 2, showing proposed encoders and decoders, where the encoder $\mathrm{Enc}^{\mathrm{t}}_{\boldsymbol{\gamma}}(\cdot)$ with the mask generator is for training, while $\mathrm{Enc}^{\mathrm{i}}_{\boldsymbol{\gamma}}(\cdot)$ is for testing/physical testing.

Also $\mathrm{Dec}_{\boldsymbol{\theta}}(\cdot)$ works in stages [15]. The first takes the encoded image $\boldsymbol{y}$ and yields a baseline ground-truth approximation $|\mathcal{F}^{-1}(\boldsymbol{y})|$. Here the modulus is needed to bring $\boldsymbol{y}$ to the real domain, in fact it may contain imaginary components that are a consequence of the combined undersampling and Inverse Fourier Transform.

The artefacts of this first approximation are corrected by the second stage, which applies pixel-wise adjustments computed by a sub-DNN $\mathcal{D}^{\mathbb{R}}_{\boldsymbol{\theta}}$ that observes $\mathcal{F}^{-1}(\boldsymbol{y})$ to yield the decoder output (see Fig. 2)

$$\hat{\boldsymbol{x}} = \mathrm{Dec}^0_{\boldsymbol{\theta}}(\boldsymbol{y}) = |\mathcal{F}^{-1}(\boldsymbol{y})| + \mathcal{D}^{\mathbb{R}}_{\boldsymbol{\theta}}(\mathcal{F}^{-1}(\boldsymbol{y})) \tag{6}$$

where $\mathcal{D}^{\mathbb{R}}_{\boldsymbol{\theta}}$ depends on the trainable parameters $\boldsymbol{\theta}$, and is the slight modification of the U-NET DNN [17], as originally proposed in [15].

In its original conception [15], [16], the parameters of this architecture are trained by considering the result of the decoding $\hat{\boldsymbol{x}}$, defining the error

$$\delta\boldsymbol{x} = \|\hat{\boldsymbol{x}} - \boldsymbol{x}\|_1 \tag{7}$$

where $\|\cdot\|_1$ represents the ℓ_l norm, and using the loss function

$$L^0(\boldsymbol{\gamma},\boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{x}}[\delta\boldsymbol{x}] \tag{8}$$

where $\mathbf{E}_{\boldsymbol{x}}[\cdot]$ stands for expectation over all possible $\boldsymbol{x}$.

V. Improved Architectures

The original LOUPE architecture and training strategy can be improved along multiple directions. Our work relies on two major intuitions: the mismatch between the measurements and the encoded version of the reconstructed scan can be *i*) reduced, if not eliminated, to improve the final reconstruction quality; *ii*) exploited to infer properties of the final output.

To reach these goals, we first go through mathematical operations that, given a general scan $\hat{\boldsymbol{x}}$, force its undersampled frequency components to a desired set. We will then show how to embed such operations inside the model, proposing new architectures and training strategies.

Let us define the following convex subsets of the set of possibly complex images $\mathbb{I}_{\mathbb{C}} = \mathbb{C}^{N\times N}$.

$$\begin{aligned} \mathbb{I}_{\boldsymbol{y}} &= \{\boldsymbol{x} | \boldsymbol{y} = \boldsymbol{M_\gamma} \circ \mathcal{F}(\boldsymbol{x})\} \\ \mathbb{I}_{\mathbb{R}} &= \mathbb{R}^{N\times N} \\ \mathbb{I}_{[0,1]} &= [0,1]^{N\times N} \end{aligned}$$

Note that $\mathrm{Dec}^0_{\boldsymbol{\theta}}(\cdot)$ guarantees that $\hat{\boldsymbol{x}} \in \mathbb{I}_{\mathbb{R}}$ but not that $\hat{\boldsymbol{x}} \in \mathbb{I}_{[0,1]}$, and thus may generate invalid images.

For a generic subset of $\mathbb{I} \subset \mathbb{I}_{\mathbb{C}}$, we indicate with $\Pi_{\mathbb{I}}(\cdot)$ the projection of its argument onto $\mathbb{I}$. We clearly have

$$\begin{aligned} \Pi_{\mathbb{I}_{\mathbb{R}}}(\boldsymbol{x}) &= \mathrm{Re}\{\boldsymbol{x}\} \\ \Pi_{\mathbb{I}_{[0,1]}}(\boldsymbol{x}) &= \max\{0, \min\{1, \mathrm{Re}\{\boldsymbol{x}\}\}\} \end{aligned}$$

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

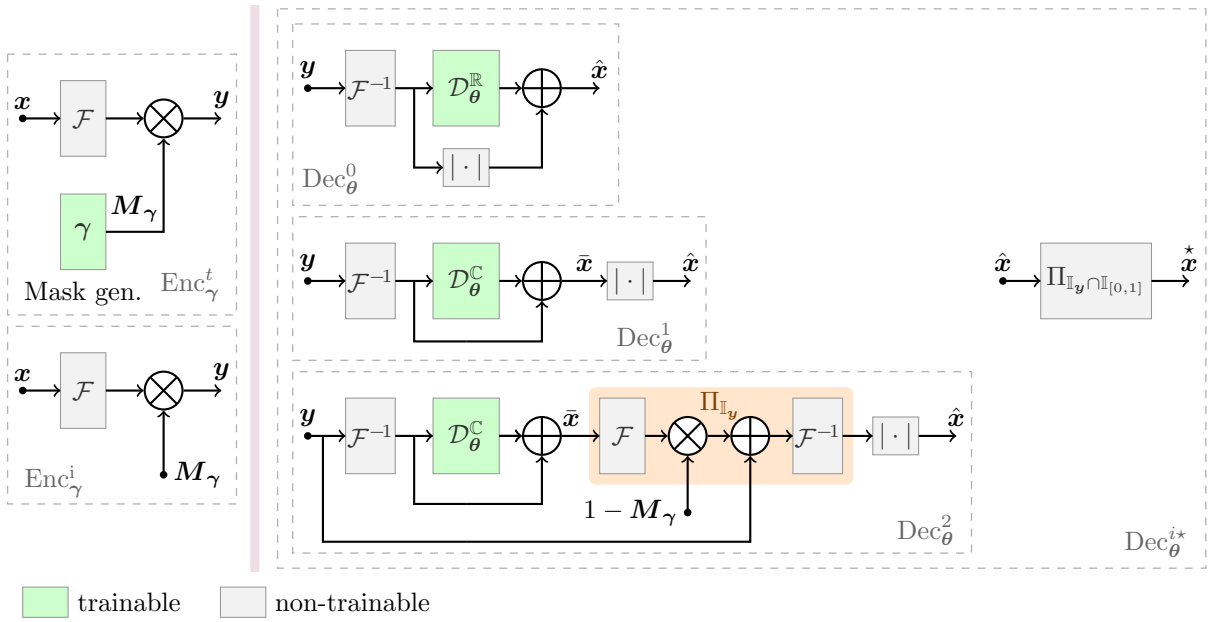


Fig. 2. Illustration of all the proposed architectures. On the left side we grouped the encoders: $\mathrm{Enc}^t_{\boldsymbol{\gamma}}$ is used at training time to learn the mask, while $\mathrm{Enc}^i_{\boldsymbol{\gamma}}$ is used at inference time and employs a fixed binary mask. On the right are the decoders: $\mathrm{Dec}^0_{\boldsymbol{\theta}}$ is one proposed in LOUPE [15], $\mathrm{Dec}^1_{\boldsymbol{\theta}}$ is a slight modification of $\mathrm{Dec}^0_{\boldsymbol{\theta}}$, $\mathrm{Dec}^2_{\boldsymbol{\theta}}$ embeds the operator $\Pi_{\mathbb{I}_{\boldsymbol{y}}}$. All the three decoders may benefit from an additional block that performs $\Pi_{\mathbb{I}_{\boldsymbol{y}}\cap\mathbb{I}_{[0,1]}}$ at inference time. When such a block is used we append a $\star$ to the decoder label: $\mathrm{Dec}^{i\star}_{\boldsymbol{\theta}}$, with $i = 0, 1, 2$.

where elementary operations are applied entry-wise to the whole matrix.

To give an expression for $\Pi_{\mathbb{I}_{\boldsymbol{y}}}$ note that $\mathrm{Enc}_{\boldsymbol{\gamma}}(\cdot) = \boldsymbol{M}_{\boldsymbol{\gamma}} \circ \mathcal{F}(\cdot)$ is a non-injective linear operator whose codomain is a $\left(N^2/R\right)$-dimensional coordinate subspace (i.e., a subspace defined by the fact that certain components of its vectors are zero). Given any $\boldsymbol{x} \in \mathbb{I}_{\mathbb{C}}$, the projection $\Pi_{\mathbb{I}_{\boldsymbol{y}}}(\boldsymbol{x})$ is the solution of

$$\Pi_{\mathbb{I}_{\boldsymbol{y}}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{\xi}} \|\boldsymbol{\xi} - \boldsymbol{x}\|_2 \quad \text{s.t.} \quad \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{\xi}) = \boldsymbol{y} \tag{9}$$

i.e., the vector

$$\Pi_{\mathbb{I}_{\boldsymbol{y}}}(\boldsymbol{x}) = \boldsymbol{x} + \mathrm{Enc}^{\dagger}_{\boldsymbol{\gamma}}\big(\boldsymbol{y} - \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{x})\big) \tag{10}$$

where $\mathrm{Enc}^{\dagger}_{\boldsymbol{\gamma}}(\cdot)$ is the Moore-Penrose pseudo-inverse of $\mathrm{Enc}_{\boldsymbol{\gamma}}(\cdot)$ whose feature deserves a brief discussion. Note in fact that, by defining the complementary encoding $\mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\cdot) = (\mathbf{1} - \boldsymbol{M}_{\boldsymbol{\gamma}}) \circ \mathcal{F}(\cdot)$, we have $\mathcal{F}(\cdot) = \mathrm{Enc}_{\boldsymbol{\gamma}}(\cdot) + \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\cdot)$. With these definitions, the Fourier transform of the projection in (10) is

$$\begin{aligned}\mathcal{F}\big(\Pi_{\mathbb{I}_{\boldsymbol{y}}}(\boldsymbol{x})\big) &= \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{x}) + \boldsymbol{y} - \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{x}) + \\ &\quad \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\boldsymbol{x}) + \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}\Big(\mathrm{Enc}^{\dagger}_{\boldsymbol{\gamma}}\big(\boldsymbol{y} - \mathrm{Enc}_{\boldsymbol{\gamma}}(\boldsymbol{x})\big)\Big) \\ &= \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\boldsymbol{x}) + \boldsymbol{y}\end{aligned}$$

where we have exploited the fact that $\mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}\Big(\mathrm{Enc}^{\dagger}_{\boldsymbol{\gamma}}(\cdot)\Big) = 0$ since, by definition, $\mathrm{Enc}^{\dagger}_{\boldsymbol{\gamma}}(\cdot)$ is an image whose Fourier transform has zeros in the positions masked by $\boldsymbol{M}_{\boldsymbol{\gamma}}$, while $\mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\cdot)$ puts to zero the Fourier components not masked by $\boldsymbol{M}_{\boldsymbol{\gamma}}$. Hence,

$$\Pi_{\mathbb{I}_{\boldsymbol{y}}}(\boldsymbol{x}) = \mathcal{F}^{-1}\left((\mathbf{1} - \boldsymbol{M}_{\boldsymbol{\gamma}}) \circ \mathcal{F}(\boldsymbol{x}) + \boldsymbol{y}\right) \tag{11}$$

meaning that the projection of $\boldsymbol{x}$ can be obtained by transforming $\boldsymbol{x}$ by Fourier, substituting the non-zero components of $\boldsymbol{y}$ in the results, and transforming it back to the image domain.

The improvements we propose start from a slight rearrangement of the original decoder (see the intermediate block in Fig. 2) that yields

$$\hat{\boldsymbol{x}} = \mathrm{Dec}^1_{\boldsymbol{\theta}}(\boldsymbol{y}) = |\bar{\boldsymbol{x}}| = \left|\mathcal{F}^{-1}(\boldsymbol{y}) + \mathcal{D}^{\mathbb{C}}_{\boldsymbol{\theta}}(\mathcal{F}^{-1}(\boldsymbol{y}))\right| \tag{12}$$

where the quantity $\bar{\boldsymbol{x}}$ remains implicitly defined and $\mathcal{D}^{\mathbb{C}}_{\boldsymbol{\theta}}$ has the same structure as $\mathcal{D}^{\mathbb{R}}_{\boldsymbol{\theta}}$ but produces a two-channel output accounting for both a real and a imaginary part of the corrections.

Such a rearrangement allows to define a further error figure that accounts for the ability of the intermediate estimation $\bar{\boldsymbol{x}}$ to reproduce the same measurements that were observed starting from the original image $\boldsymbol{x}$, i.e.,

$$\delta\bar{\boldsymbol{y}} = \left\|\boldsymbol{y} - \mathrm{Enc}_{\boldsymbol{\gamma}}(\bar{\boldsymbol{x}})\right\|_2 \tag{13}$$

One may observe that good reconstructions $\bar{\boldsymbol{x}}$ should be such that $\delta\bar{\boldsymbol{y}}$ is small. It is possible to add this information to the training procedure of $\mathrm{Dec}^1_{\boldsymbol{\theta}}(\cdot)$ by defining a new loss function

$$L^1(\boldsymbol{\gamma}, \boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{x}}\left[(1-\phi)\delta\boldsymbol{x} + \phi\delta\bar{\boldsymbol{y}}\right] \tag{14}$$

which, by using the user-defined balancing coefficient $\phi \in [0, 1]$, favors solutions in which not only $\hat{\boldsymbol{x}}$ is close to $\boldsymbol{x}$ but also the measurements that would have been produced by $\bar{\boldsymbol{x}}$ are close to those produced by $\boldsymbol{x}$.

As an alternative option, one may introduce the measurement matching criterion into the signal chain and define (see third decoder structure in Fig. 2)

$$\hat{\boldsymbol{x}} = \mathrm{Dec}^2_{\boldsymbol{\theta}}(\boldsymbol{y}) = \left|\Pi_{\mathbb{I}_{\boldsymbol{y}}}\left(\mathcal{F}^{-1}(\boldsymbol{y}) + \mathcal{D}^{\mathbb{C}}_{\boldsymbol{\theta}}(\mathcal{F}^{-1}(\boldsymbol{y}))\right)\right| \tag{15}$$

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

that is normally trained by means of $L^0(\boldsymbol{\gamma}, \boldsymbol{\theta})$. Note that $\mathrm{Dec}^2_{\boldsymbol{\theta}}(\cdot)$ includes a projection onto $\mathbb{I}_{\boldsymbol{y}}$ of $\bar{\boldsymbol{x}}$ and thus improves the measurement matching criterion at processing time.

Finally, independently of the inner decoder, one may append a final stage guaranteeing that the recovered image satisfies all requirements, thus defining

$$\overset{\star}{\boldsymbol{x}} = \mathrm{Dec}^{i\star}_{\boldsymbol{\theta}}(\boldsymbol{y}) = \Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}} \left(\mathrm{Dec}^{i}_{\boldsymbol{\theta}}(\boldsymbol{y}) \right) \tag{16}$$

where $\Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}}$ can be computed by the Dykstra's alternating projections algorithm [18] [19] that uses the elementary expressions of $\Pi_{\mathbb{I}_{\boldsymbol{y}}}$ and $\Pi_{\mathbb{I}_{[0,1]}}$.

The adoption of a final projection stage means that the final image reconstruction is able to reproduce the observed measurements. This is the best one can do at decoding-time as the Fourier components of $\boldsymbol{x}$ that are not measurements are not known and cannot be used in computations.

Yet, at training time all the Fourier components are available. Hence, one may think of improving the performance of $\mathrm{Dec}^{i\star}_{\boldsymbol{\theta}}(\cdot)$ by training it with a loss function considering the measurements not taken by defining the error term

$$\delta \boldsymbol{y}^{\perp} = \left\| \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\bar{\boldsymbol{x}}) - \mathrm{Enc}^{\perp}_{\boldsymbol{\gamma}}(\boldsymbol{x}) \right\|_2 \tag{17}$$

that accounts for the ability of the intermediate estimation $\bar{\boldsymbol{x}}$ to produce the same Fourier components that $\boldsymbol{x}$ would produce if they were acquired. This can be used in an alternative loss function

$$L^2(\boldsymbol{\gamma}, \boldsymbol{\theta}) = \mathbf{E}_{\boldsymbol{x}} \left[(1-\psi)\delta \boldsymbol{x} + \psi \delta \boldsymbol{y}^{\perp} \right] \tag{18}$$

for a proper user-defined weight $\psi \in [0, 1]$. Worth noting this option gives substantial performance improvements only when applied along with $\mathrm{Dec}^{1\star}_{\boldsymbol{\theta}}(\cdot)$.

All our contributions, in particular the architectural modifications, introduce a negligible complexity overhead in term of number parameters and number of Multiply-and-Accumulate operations (MAC). In Dec^0, U-NET dominates the complexity count with $86.9\,\mathrm{GMAC}$. In Dec^1 the U-NET structure undergoes a small modification leading to an increase of $6.55\,\mathrm{MMAC}$. Dec^2 adds a projection block that increases the previous computational cost by a further $13.6\,\mathrm{MMAC}$. Similarly, Dykstra's algorithm (the $\star$ in $\mathrm{Dec}^{i\star}$) introduces $13.6\,\mathrm{MMAC}$ for every iteration; in our case we use 20 iterations. For what concerns the memory footprint the increase is minimal, in fact Dec^1 and Dec^2 only add 65 more parameters to the $31.5\mathrm{M}$ parameters of Dec^0.

A. Decoder self-assessment

All the improved decoders we propose exploit the fact that good reconstructions should correspond to small $\delta \bar{\boldsymbol{y}}$. Yet, in an ideal reconstruction chain, each estimation of the image should be as close as possible to the original one and thus virtually produce the same measurements.

Hence, it is sensible to assume that by defining

$$\delta \hat{\boldsymbol{y}} = \left\| \boldsymbol{y} - \mathrm{Enc}_{\boldsymbol{\gamma}}(\hat{\boldsymbol{x}}) \right\|_2 \tag{19}$$

when the reconstruction fails then $\delta \bar{\boldsymbol{y}}$ and/or $\delta \hat{\boldsymbol{y}}$ are substantially larger than 0.

Consequently, $\delta \bar{\boldsymbol{y}}$ and $\delta \hat{\boldsymbol{y}}$, which can be computed starting from the knowledge of the actual measurements $\boldsymbol{y}$ only, have a magnitude that can be seen as a proxy of the decoder performance.

This proxy can be exploited in a self-assessment stage that follows reconstruction and whose aim is to give the user additional information on the quality of the output image at inference time, i.e., when the true image $\boldsymbol{x}$ is unknown and the physical system only captures the frequency components in $\boldsymbol{y}$.

In particular, to assess the capability of $\delta \bar{\boldsymbol{y}}$ and/or $\delta \hat{\boldsymbol{y}}$ to predict decoder performances, we use the Pearson correlation coefficient $\rho(a, b) = \mathrm{cov}(a, b)/(\sigma_a \sigma_b)$, where $\mathrm{cov}(\cdot, \cdot)$ is the covariance between two vectors while $\sigma_{\cdot}$ stands for standard deviation. We adopt $\rho(\delta \hat{\boldsymbol{x}}, \delta \hat{\boldsymbol{y}})$ for $\mathrm{Dec}^0_{\boldsymbol{\theta}}(\cdot)$, $\rho(\delta \overset{\star}{\boldsymbol{x}}, \delta \hat{\boldsymbol{y}})$ for $\mathrm{Dec}^{0\star}_{\boldsymbol{\theta}}(\cdot)$ and $\rho(\delta \overset{\star}{\boldsymbol{x}}, \delta \bar{\boldsymbol{y}})$ for the remaining proposed approaches with $\delta \overset{\star}{\boldsymbol{x}} = ||\boldsymbol{x} - \overset{\star}{\boldsymbol{x}}||_1$.

VI. Numerical Evidence

A. Dataset

To ease comparison of our methods[1] with [15], we coherently test our models using a subset of the public NYU fastMRI dataset [51], a benchmark commonly used for international competitions [25], [26]. From this dataset, we select the emulated single-coil proton density weighted scans (derived from raw 15-channel multi-coil data) by the Biograph mMR scanner.

Scans are grouped in *volumes*, each volume being composed of slices that are grayscale images with $N = 320$ that we normalize into $\mathbb{I}_{[0,1]}$ dividing pixel values by the maximum magnitude within each volume.

The image set is split into three subsets: 50 volumes used for training, 5 volumes used for validation and 5 volumes used for testing. Each training volume has a number of slides ranging from 34 to 42, validation volumes from 34 to 40, while test volumes from 35 to 45 such that 1895 slices compose the training set, 188 the validation and the remaining 186 build the test set.

B. Training and Evaluation

We train our models on an Nvidia V100 using Adam optimizer, an initial learning rate of 0.01, and a batch size of 16. Evaluation is performed over the test set using an adapted, fixed undersampling mask.

To assess the quality of each reconstruction we adopt three different metrics:

- The Peak Signal-to-Noise Ratio (PSNR):

$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{\max\{\boldsymbol{x}\}}{\mathrm{MSE}} \right) \tag{20}$$

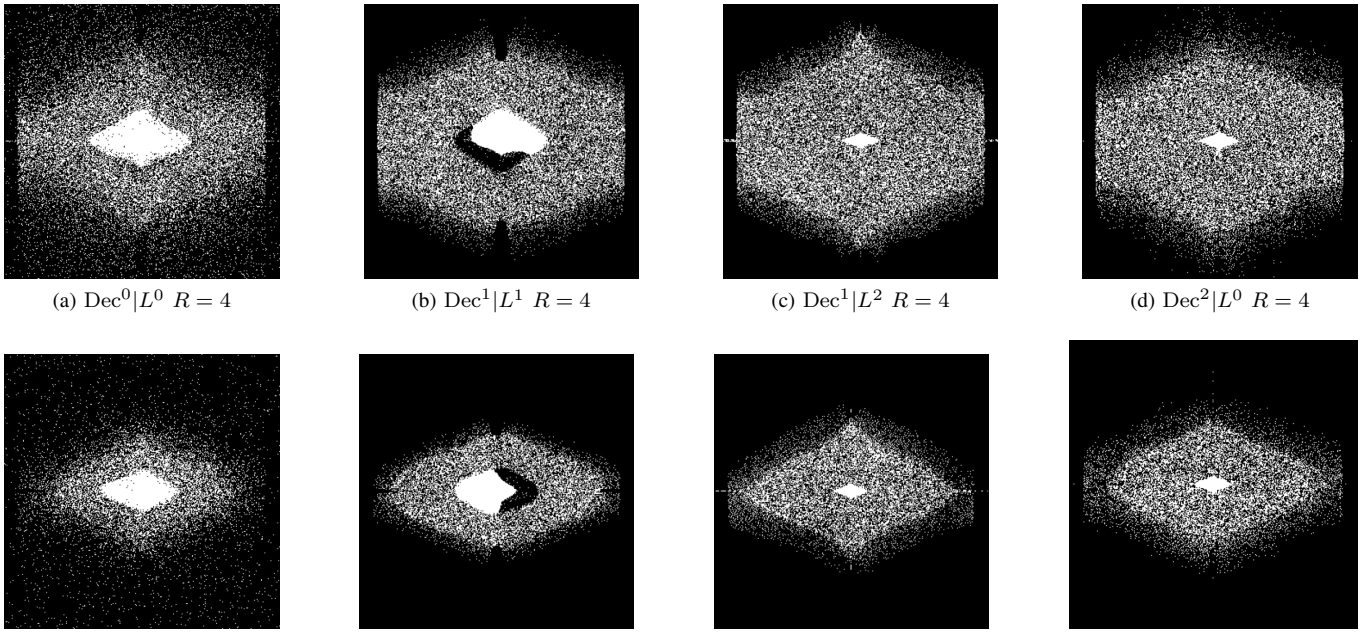
  where MSE is the mean square error between $\boldsymbol{x}$ and the decoder output.
- The Structural Similarity Index (SSIM):

$$\mathrm{SSIM} = \frac{(2\mu_{\boldsymbol{x}}\mu_{\hat{\boldsymbol{x}}} + c_1)(2\sigma_{\boldsymbol{x}\hat{\boldsymbol{x}}} + c_2)}{(\mu^2_{\boldsymbol{x}} + \mu^2_{\hat{\boldsymbol{x}}} + c_1)(\sigma^2_{\boldsymbol{x}} + \sigma^2_{\hat{\boldsymbol{x}}} + c_2)} \tag{21}$$

[1]Code implementing proposed methods available at https://github.com/SSIGPRO/LOUPE_evolutions

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082



(a) $\mathrm{Dec}^0|L^0$ $R=4$ (b) $\mathrm{Dec}^1|L^1$ $R=4$ (c) $\mathrm{Dec}^1|L^2$ $R=4$ (d) $\mathrm{Dec}^2|L^0$ $R=4$

(e) $\mathrm{Dec}^0|L^0$ $R=8$ (f) $\mathrm{Dec}^1|L^1$ $R=8$ (g) $\mathrm{Dec}^1|L^2$ $R=8$ (h) $\mathrm{Dec}^2|L^0$ $R=8$

Fig. 3. undersampling masks $\boldsymbol{M}_{\boldsymbol{\gamma}} \in [0,1]^{N\times N}$ after training for all considered decoders (starred and non-starred versions share the same undersampling mask).

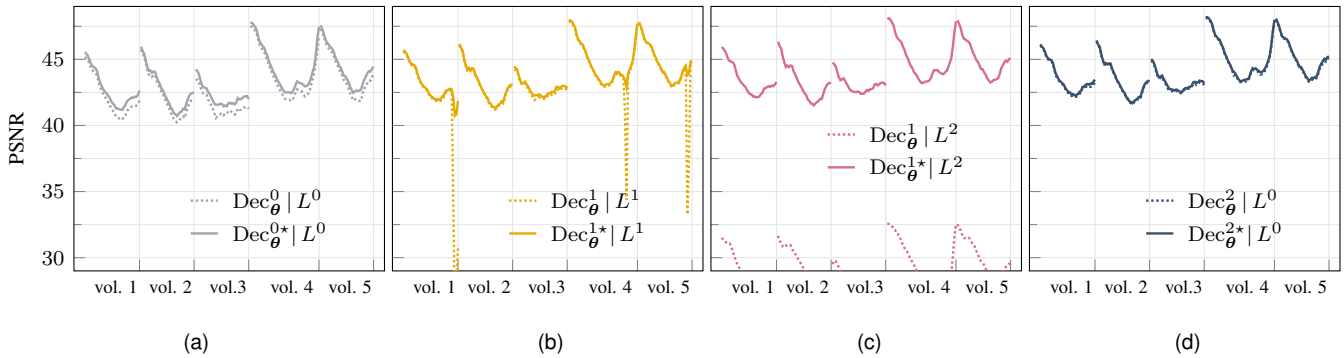


Fig. 4. Effect of final projection on various decoders for $R=4$.

where $\mu_{\boldsymbol{x}}$ and $\mu_{\hat{\boldsymbol{x}}}$ are defined as *local* averages for the original and reconstructed images, $\sigma_{\boldsymbol{x}}$ and $\sigma_{\hat{\boldsymbol{x}}}$ are the *local* variances, while $\sigma_{\boldsymbol{x}\hat{\boldsymbol{x}}}$ is the *local* covariance between the decoder output and the ground-truth. *local* means computed on sliding windows with size 11×11 where each pixel is weighted with a gaussian weight with standard deviation 1.5 pixels. In the above formula, $c_1=(k_1L)^2$ and $c_2=(k_2L)^2$, where L is the dynamic range of the pixel values, $k_1=0.01$ and $k_2=0.03$ [52].

- The high-frequency error norm (HFEN), which is a metric quantifying how well original edges are preserved by reconstruction. Following [53], a Laplacian of Gaussian (LoG) filter is first applied with a standard deviation of 1.5 pixels and a window size of 13×13. HFEN is computed as the norm ℓ_2 of the difference between the LoG filtered ground-truth and the LoG filtered reconstructed image.

We focus on speed-up factors $R=4$ and $R=8$. For a given R, we consider four possible combinations of decoder architectures and training loss functions, namely: $\mathrm{Dec}^0|L^0$, which is the original LOUPE [15]; $\mathrm{Dec}^1|L^1$ in which the loss function emphasizes the ability to reproduce the physical measurements; $\mathrm{Dec}^1|L^2$ in which the loss function emphasizes the ability to reproduce the Fourier coefficients that are neglected in the acquisition; and $\mathrm{Dec}^2|L^0$ in which an approximate measurement constraint is embodied in the network architecture. For all these options the performance of the *starred version* $\mathrm{Dec}^{i\star}$ is also considered.

In particular, the best values of ϕ in $\mathrm{Dec}^1|L^1$ is experimentally determined to be $\phi=10^{-5}$ for $R=4$ and $\phi=10^{-4}$ for $R=8$, respectively. An even more unbalanced configuration results from the numerical optimization of ψ in $\mathrm{Dec}^1|L^2$, which yields the best results with $\psi=1$ for both R=4 and R=8. This value implies that the loss reduces to $L^2(\boldsymbol{\gamma},\boldsymbol{\theta})=$

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

TABLE I
AVERAGE (μ), MINIMUM (min) AND MAXIMUM (max) VALUES OF PSNR [dB] SSIM AND HFEN FOR THE THE TEST SET AND $R = 4$.

		$\mathrm{Dec}^0\|L^0$ (LOUPE)	$\mathrm{Dec}^{0\star}\|L^0$	$\mathrm{Dec}^{1\star}\|L^1$	$\mathrm{Dec}^{1\star}\|L^2$	$\mathrm{Dec}^{2\star}\|L^0$
HFEN	μ	0.162	0.099	0.031	**0.022**	0.025
	min	0.112	0.056	0.015	**0.011**	0.012
	max	0.316	0.241	0.083	**0.065**	0.073
SSIM	μ	0.961	0.965	0.969	0.971	**0.971**
	min	0.944	0.949	0.955	0.958	**0.958**
	max	0.979	0.982	0.983	**0.983**	0.983
PSNR	μ	42.8	43.4	43.8	44.0	**44.1**
	min	40.3	40.7	41.2	41.5	**41.6**
	max	47.5	47.8	48.0	**48.1**	48.1

TABLE II
AVERAGE (μ), MINIMUM (min) AND MAXIMUM (max) VALUES OF PSNR [dB] SSIM AND HFEN FOR THE THE TEST SET AND $R = 8$.

		$\mathrm{Dec}^0\|L^0$ (LOUPE)	$\mathrm{Dec}^{0\star}\|L^0$	$\mathrm{Dec}^{1\star}\|L^1$	$\mathrm{Dec}^{1\star}\|L^2$	$\mathrm{Dec}^{2\star}\|L^0$
HFEN	μ	0.264	0.224	0.173	0.144	**0.144**
	min	0.154	0.118	0.083	**0.071**	0.071
	max	0.614	0.504	0.430	0.388	**0.369**
SSIM	μ	0.936	0.940	0.944	0.948	**0.948**
	min	0.902	0.908	0.915	0.920	**0.921**
	max	0.968	0.973	0.974	0.974	**0.974**
PSNR	μ	40.5	40.8	41.1	41.4	**41.4**
	min	37.7	38.0	38.3	38.6	**38.7**
	max	45.6	46.0	46.1	46.2	**46.2**

$\mathbf{E}_{\boldsymbol{x}}\left[\delta \boldsymbol{y}^{\perp}\right]$, i.e., it only focuses on predicting the frequencies discarded at sensing time. Such an approach is expected to reach consisting performances only when coupled with a final projection stage as the actual value of measurements would be otherwise completely neglected.

C. Undersampling Masks

To show how encoder adaptation works, Fig. 3 reports the fixed binary masks $\boldsymbol{M}_{\boldsymbol{\gamma}}$ in (5) resulting from the training of the four architectures we consider. White dots correspond to Fourier coefficients that are acquired by the systems. Clearly, their number decreases when R increases. Yet, it is evident that, since the encoder and the decoder are trained together, different adaptation policies and decoder architectures yield different under-sampling masks.

D. Performance

The plots showing the performance of the overall signal chain have all the same format. They report the result for every slice of each of the 5 volumes in the test set. Each slice corresponds to a position on the horizontal axis, and the first slice of the $(i+1)$-th volume follows the last slice of the i-th volume. Each reconstruction is described using three different figures of merit: SSIM, HFEN and PSNR. Different volumes and different slices in the same volume may have different quality scores, which reflects in the plots' non-constant profile.

Fig. 4 shows the effect of adding $\Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}}$ to change Dec^i into $\mathrm{Dec}^{i\star}$ for $R = 4$ in term of PSNR. The plots share the vertical axis, and dotted tracks correspond to the non-starred configurations while the solid tracks to the starred ones. The effect of the final projection is negligible only for $\mathrm{Dec}^2|L^0$ since the architecture of Dec^2 includes $\Pi_{\mathbb{I}_{\boldsymbol{y}}}$ that is the main component of $\Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}}$. In all other cases, it is clear that $\Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}}$ is most effective on the slices with lower PSNR and it brings a dramatic benefit when applied to models trained with L^2. Note that $\psi = 0$, hence the network is trained to exclusively learn how to reconstruct the Fourier coefficients that are not acquired (an effect that is complementary to the one of $\Pi_{\mathbb{I}_{\boldsymbol{y}} \cap \mathbb{I}_{[0,1]}}$).

Overall, best performances on this dataset are achieved by $\mathrm{Dec}^{1\star}|L^2$ and $\mathrm{Dec}^{2\star}|L^0$. To assess the improvement quantitatively over the original LOUPE ($\mathrm{Dec}^0|L^0$), we choose to consider HFEN, PSNR, SSIM, and we lay down Table I and Table II in which we report the average, minimum and maximum reconstruction qualities for the four starred configurations, along with the original LOUPE proposal in case of $R = 4$ and $R = 8$. With respect to LOUPE, our proposed architectures achieve non-negligible improvements in terms of both SSIM and PSNR, while HFEN values show a remarkable enhancement.

Finally, Fig. 5 compares proposed approaches (solid lines) with LOUPE, $\mathrm{Dec}^0|L^0$, (dashed line) for both $R = 4$ and $R = 8$. These results also show that passing from $R = 4$ to $R = 8$ causes a reduction of the quantified quality for all considered metrics, confirming results in Table I and Table II. Note also that the greatest improvements correspond to the slides with the most complex structure, i.e., the slides for which the reconstruction task is the hardest. With more details, for $R = 4/8$, the improvements are up to $2.1/1.7\,\mathrm{dB}$ in terms of PSNR, $0.016/0.020$ in terms of SSIM, and $0.101/0.079$ in terms of HFEN.

Moreover, it results that: *i*) $\mathrm{Dec}^{2\star}|L^0$ and $\mathrm{Dec}^{1\star}|L^2$ always obtains the best scores; *ii*) $\mathrm{Dec}^{0\star}|L^0$ always improves original LOUPE by applying the Dykstra's alternating projections algorithm to LOUPE reconstructed scans at inference time; *iii*) all proposed approaches improve with respect to original LOUPE with the only exception of $\mathrm{Dec}^{1\star}|L^1$ performing worse than $\mathrm{Dec}^0|L^0$ in the 1.6% of cases (3 out of 186 slices) with $R = 4$.

For a visual representation, reconstructed images with $\mathrm{Dec}^{2\star}|L^0$ and LOUPE, along with the ground truth, are in Fig. 6 where we chose $\mathrm{Dec}^{2\star}|L^0$ to represent our approaches.

E. Comparison with other approaches

In [15], authors compare LOUPE with state-of-the-art decoders and use many different undersampling patterns. Because our work evolves and improves LOUPE, we can infer to be superior to all those methods, but, as a further assessment, we test our model against some other state-of-the-art techniques, namely those in [34], [38], [41], [48]. For each of the above, we adapt our method to work on the same dataset and with the same speed-up factors. We then measure PSNR and SSIM and compare them with those reported in the cited references.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

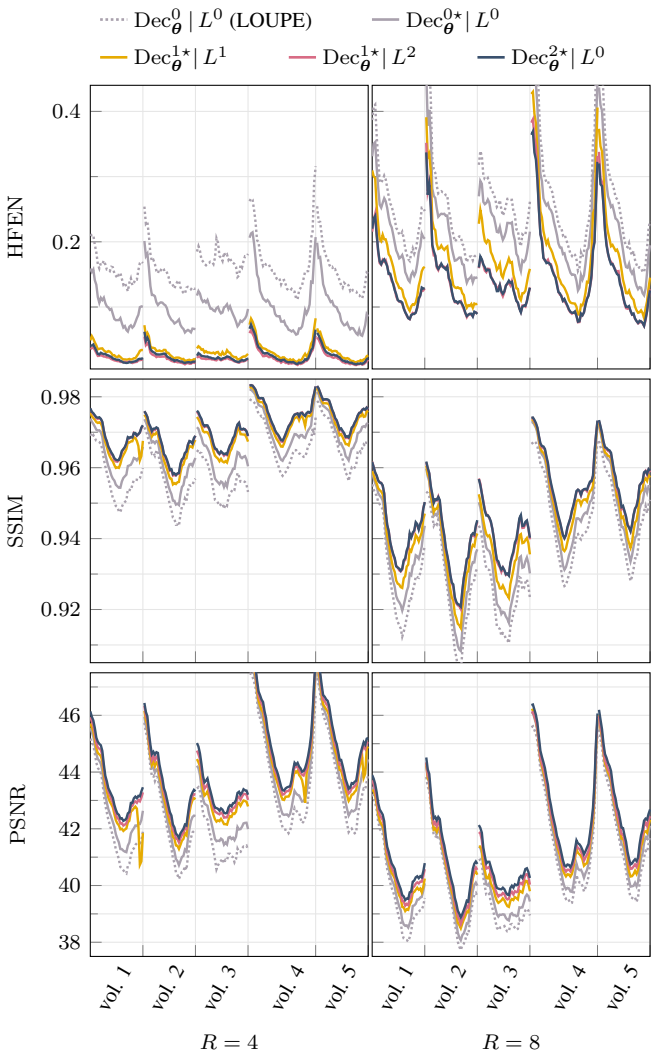


Fig. 5. Results for all considered decoders and for the 5 volumes composing the test set in terms of HFEN, SSIM, and PSNR. Volumes number 1,2,3 and 5 include 35 slices while 45 slides are in volume 4. Left plots are for $R = 4$ while $R = 8$ is used for plots on the right.

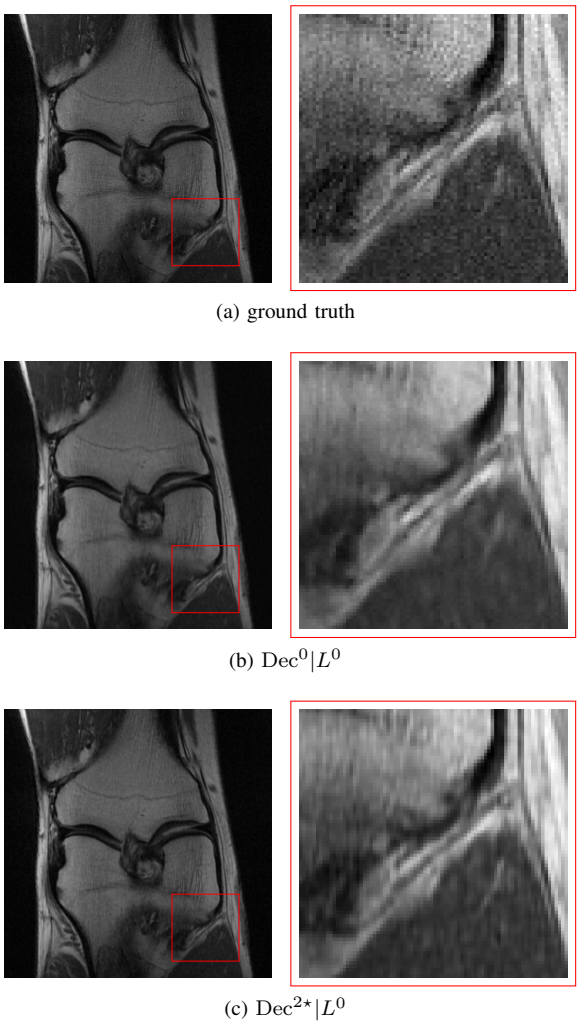


Fig. 6. Left side of (a) represents a single Knee image from the dataset "fastMRI", volume 2 and slice number 24 in the test set. Left sides of (b) and (c) correspond to reconstructed images by $\mathrm{Dec}^0|L^0$ with PSNR equal to 37.7 dB, and $\mathrm{Dec}^{2\star}|L^0$ with 38.9 dB. Reconstructions refer to $R = 8$. Right side images highlight the region of the images in the left sides in the red box.

The considered datasets are: *i*) the fast MRI proton density dataset (PD, 60 subjects, 2269 scans) [25], *ii*) the fast MRI proton density fat suppression dataset (PDFS, 70 subjects, 2440 scans) [25] and *iii*) IXI dataset [54] (170 subjects, 27750 scans), all of them considering the single-coil acquisition method. PD and PDFS contain fully sampled knees, while IXI is a collection of fully sampled brains. Following [38], we split IXI in test/validation set by randomly selecting 8/11 volumes. For PD and PDFS, we reproduce the partitioning proposed by [15], where test and validation sets are created by randomly extracting 5 volumes for each. In both cases, the remaining volumes are used to train the models.

We summarise in Table III all achieved results. Our methods, represented by $\mathrm{Dec}^{2\star}|L^0$, outperform every competitor in terms of both PSNR and SSIM, confirming the effectiveness of the proposed approaches.

TABLE III
COMPARISON IN TERMS OF PSNR AND SSIM OF $\mathrm{Dec}^{2\star}_{\boldsymbol{\theta}}|L^0$ WITH THE RECENTLY PROPOSED WORKS IN [34], [38], [41], [48] ON PD [25], PDFS [25], AND IXI [54] DATASETS.

dataset	R	PSNR [dB]			SSIM		
		$\mathrm{Dec}^{2\star}_{\boldsymbol{\theta}}\|L^0$	[41]	[48]	$\mathrm{Dec}^{2\star}_{\boldsymbol{\theta}}\|L^0$	[41]	[48]
PD	4	**44.1**	35.2	33.8	**0.971**	0.835	0.87
	6	**42.2**	34.2	-	**0.955**	0.807	-
PDFS	4	**35.6**	30.3	-	**0.868**	0.651	-
	6	**34.0**	29.7	-	**0.799**	0.604	-
		$\mathrm{Dec}^{2\star}_{\boldsymbol{\theta}}\|L^0$	[38]	[34]	$\mathrm{Dec}^{2\star}_{\boldsymbol{\theta}}\|L^0$	[38]	[34]
IXI	5	**48.3**	44.0	34.6	**0.995**	0.991	0.979
	10	**42.1**	37.6	31.2	**0.977**	0.964	0.961

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

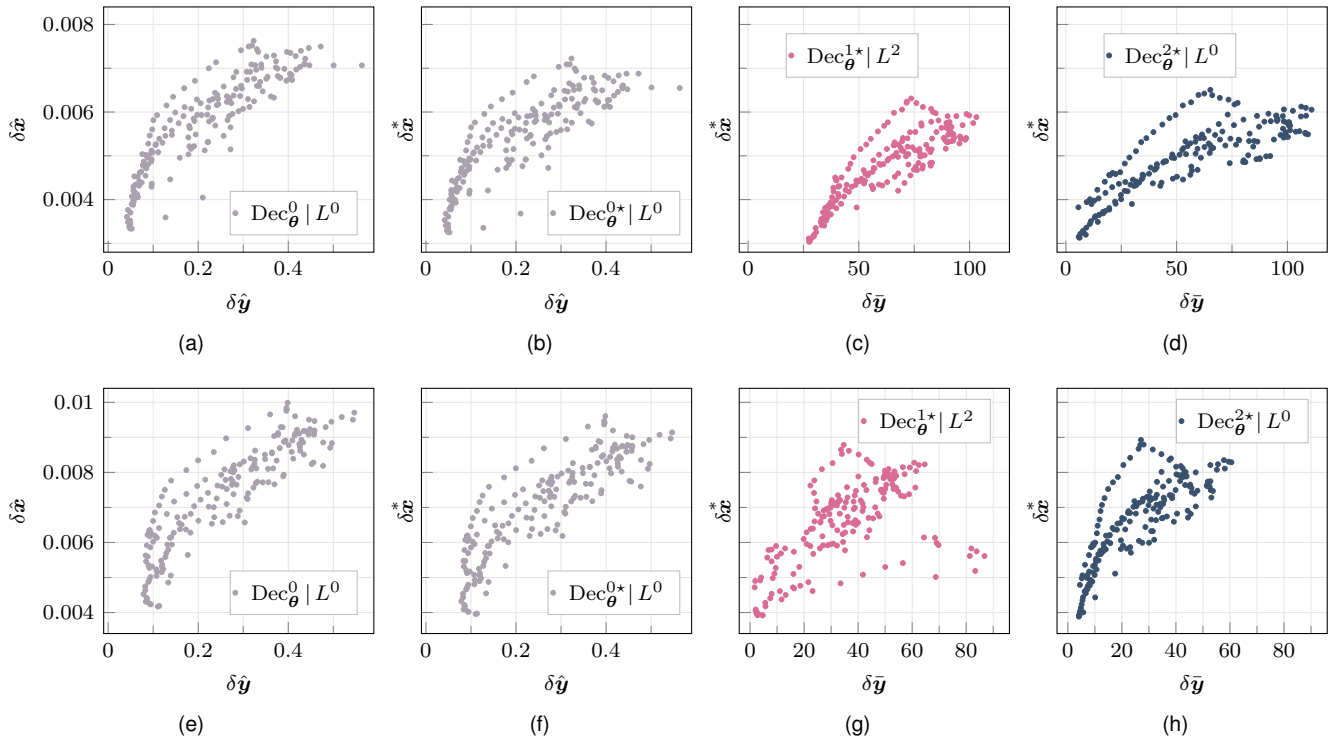


Fig. 7. Error in the measurement constraint fitting versus reconstruction error for the images composing the test set. Results are for all considered approaches except for $\mathrm{Dec}^{1*}|L^1$. Top plots are for $R = 4$ while bottom plots are for $R = 8$.

TABLE IV
PEARSON CORRELATION COEFFICIENT BETWEEN RECONSTRUCTION ERROR AND THE MEASUREMENT MISMATCH.

	$\mathrm{Dec}^0\|L^0$ (LOUPE)	$\mathrm{Dec}^{0\star}\|L^0$	$\mathrm{Dec}^{1\star}\|L^1$	$\mathrm{Dec}^{1\star}\|L^2$	$\mathrm{Dec}^{2\star}\|L^0$
	$\rho(\delta\hat{\boldsymbol{y}}, \delta\hat{\boldsymbol{x}})$	$\rho(\delta\hat{\boldsymbol{y}}, \delta\hat{\boldsymbol{x}}^\star)$	$\rho(\delta\bar{\boldsymbol{y}}, \delta\hat{\boldsymbol{x}}^\star)$	$\rho(\delta\bar{\boldsymbol{y}}, \delta\hat{\boldsymbol{x}}^\star)$	$\rho(\delta\bar{\boldsymbol{y}}, \delta\hat{\boldsymbol{x}}^\star)$
$R = 4$	0.877	0.837	0.138	0.853	0.821
$R = 8$	0.888	0.866	0.198	0.535	0.863

F. Self-assessment

In Table IV we present the Pearson correlation coefficient as defined in V-A.

Avoiding $\mathrm{Dec}^{1*}|L^1$, the results are overall promising and show a good correlation, i.e., it is possible to guess decoder performance starting from its capability to match the measurements constraint. We hope that this peculiarity could help us in the design of a detector capable of identifying flawed image reconstructions, a topic that will be explored in future communications. To further evidence the meaning of the computed correlation coefficients, Fig. 7 shows scatter plots connecting reconstruction errors with the fitting of the measurements constraint [2]. It is interesting to note that while all the graphs in Fig. 7 share the same y-axis, the x-axis ranges differ conspicuously, and this is because at training time $\mathrm{Dec}^0|L^0$ reduces $\delta\hat{\boldsymbol{y}}$, while the training of $\mathrm{Dec}^1|L^2$ and $\mathrm{Dec}^2|L^0$ does not affect $\delta\bar{\boldsymbol{y}}$.

[2] $\mathrm{Dec}^{1*}|L^1$ is not reported since it does not lead to interesting results.

VII. CONCLUSION

We evolve our method from a deep neural state-of-the-art technique called LOUPE [16], which simultaneously adapts both undersampling pattern and image recovery. We demonstrate that it is possible to enhance the reconstruction performance with some architectural changes. In particular *i*) we show how the previously neglected measurements constraint can be exploited in three different ways, always bringing advantages, and *ii*) we solve the optimization problem to find the image that, belonging to the image domain and the set of the counterimages, is optimally reconstructed. In terms of PSNR, we gain approximately 1 dB with most significant enhancements (up to 2.1 dB) in case of detail-rich images. Such improvements are also confirmed with perceptual metrics such as SSIM and HFEN.

We confirm that, in fast MRI, undersampling pattern and decoder benefit from joint optimization. In our view, this holds for any choice of encoder and decoder. In this sense, the framework here considered is only mildly dependent on the specific decoder selection (the U-NET) and may extend to other decoders.

Finally, we demonstrate a connection between the error committed on the whole reconstruction and the error committed on the measurement-only reconstruction of every scan. It should be possible to accurately estimate the first (otherwise unknown) simply by considering the second (always available). We will further explore this direction and plan to present achieved results in future communications.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

ACKNOWLEDGMENT

This work was partially supported by the SmartData@PoliTO center, and by the Italian Ministry for University and Research (MUR) under the program "Dipartimenti di Eccellenza (2018-2022)".

REFERENCES

[1] J. L. Prince and J. Links, "Medical Imaging Signals and Systems, 2nd edition", Pearson, 2015.

[2] D. Hailey "Open magnetic resonance imaging (MRI) scanners". Issues Emerging Health Technologies, 2006.

[3] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 4 2006.

[5] E. Candes and T. Tao, "Decoding by Linear Programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 12 2005.

[6] FDA. (2017, 2) Fda clears compressed sensing mri acceleration technology from siemens healthineers.

[7] Shiqian Ma, Wotao Yin, Yin Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *2008 IEEE Conference on Computer Vision and Pattern Recognition* IEEE, 6 2008, pp. 1–8.

[8] S. Ravishankar and Y. Bresler, "MR Image Reconstruction From Highly Undersampled k-Space Data by Dictionary Learning," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1028–1041, 5 2011.

[9] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Medical Image Analysis*, vol. 18, no. 6, pp. 843–856, 8 2014.

[10] F. Knoll, C. Clason, C. Diwoky, and R. Stollberger, "Adapted random sampling patterns for accelerated MRI," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 24, no. 1, pp. 43–50, 2 2011.

[11] Zhongmin Wang and G. Arce, "Variable Density Compressed Image Sampling," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 264–270, 1 2010.

[12] J. P. Haldar, D. Hernando, and Zhi-Pei Liang, "Compressed-Sensing MRI With Random Encoding," *IEEE Transactions on Medical Imaging*, vol. 30, no. 4, pp. 893–903, 4 2011.

[13] U. Gamper, P. Boesiger, and S. Kozerke, "Compressed sensing in dynamic MRI," *Magnetic Resonance in Medicine*, vol. 59, no. 2, pp. 365–373, 2 2008.

[14] R. Yang, F. Mentzer, L. Van Gool, and R. Timofte, "Learning for Video Compression with Recurrent Auto-Encoder and Recurrent Probability Model," *IEEE Journal on Selected Topics in Signal Processing*, vol. 15, no. 2, pp. 388–401, 2021.

[15] C. D. Bahadir, A. Q. Wang, A. V. Dalca, and M. R. Sabuncu, "Deep-Learning-Based Optimization of the Under-Sampling Pattern in MRI," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1139–1152, 2020.

[16] C. D. Bahadir, A. V. Dalca, and M. R. Sabuncu, "Learning-Based Optimization of the Under-Sampling Pattern in MRI," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, pp. 780–792.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234–241, 2015.

[18] R. L. Dykstra, "An algorithm for restricted least squares regression," *Journal of the American Statistical Association*, vol. 78, no. 384, pp. 837–842, 1983.

[19] S. Boyd and J. Dattorro, "Alternating projections," EE392o, Stanford University, 2003.

[20] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI: A look at how CS can improve on current imaging techniques," pp. 72–82, 2008.

[21] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, no. 7697, pp. 487–492, 3 2018.

[22] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Advances in Neural Information Processing Systems*, 2016, pp. 10–18.

[23] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 6 2002.

[24] A. Sriram, J. Zbontar, T. Murrell, C. L. Zitnick, A. Defazio, and D. K. Sodickson, "GrappaNet: Combining Parallel Imaging With Deep Learning for Multi-Coil MRI Reconstruction," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 6 2020, pp. 14 303–14 310.

[25] M. J. Muckley, B. Riemenschneider, A. Radmanesh, S. Kim, G. Jeong, J. Ko, Y. Jun, H. Shin, D. Hwang, M. Mostapha, S. Arberet, D. Nickel, Z. Ramzi, P. Ciuciu, J.-L. Starck, J. Teuwen, D. Karkalousos, C. Zhang, A. Sriram, Z. Huang, N. Yakubova, Y. W. Lui, and F. Knoll, "Results of the 2020 fastMRI Challenge for Machine Learning MR Image Reconstruction," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2021.

[26] F. Knoll, T. Murrell, A. Sriram, N. Yakubova, J. Zbontar, M. Rabbat, A. Defazio, M. J. Muckley, D. K. Sodickson, C. L. Zitnick, and M. P. Recht, "Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge," *Magnetic Resonance in Medicine*, vol. 84, no. 6, pp. 3054–3070, 12 2020.

[27] S. Wang, H. Cheng, L. Ying, T. Xiao, Z. Ke, H. Zheng, and D. Liang, "DeepcomplexMRI: Exploiting deep residual network for fast parallel MR imaging with complex convolution," *Magnetic Resonance Imaging*, vol. 68, no. September 2019, pp. 136–147, 5 2020.

[28] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, vol. 3, no. January, 2014, pp. 2672–2680.

[29] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310–1321, 6 2018.

[30] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network With a Cyclic Loss," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1488–1497, 6 2018.

[31] G. Oh, B. Sim, H. Chung, L. Sunwoo, and J. C. Ye, "Unpaired Deep Learning for Accelerated MRI Using Optimal Transport Driven CycleGAN," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1285–1296, 2020.

[32] M. Jiang, Z. Yuan, X. Yang, J. Zhang, Y. Gong, L. Xia, and T. Li, "Accelerating CS-MRI Reconstruction With Fine-Tuning Wasserstein Generative Adversarial Network," *IEEE Access*, vol. 7, pp. 152 347–152 357, 2019.

[33] J. Liu and M. Yaghoobi, "Fine-Grained Mri Reconstruction Using Attentive Selection Generative Adversarial Networks," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 6 2021, pp. 1155–1159.

[34] S. U. Dar, M. Yurt, M. Shahdloo, M. E. Ildiz, B. Tinaz, and T. Cukur, "Prior-guided image reconstruction for accelerated multi-contrast mri via generative adversarial networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1072–1087, 10 2020.

[35] J. Lv, C. Wang, and G. Yang, "PIC-GAN: A Parallel Imaging Coupled Generative Adversarial Network for Accelerated Multi-Channel MRI Reconstruction," *Diagnostics 2021, Vol. 11, Page 61*, vol. 11, no. 1, p. 61, 1 2021.

[36] Z. Yuan, M. Jiang, Y. Wang, B. Wei, Y. Li, P. Wang, W. Menpes-Smith, Z. Niu, and G. Yang, "SARA-GAN: Self-Attention and Relative Average Discriminator Based Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," *Frontiers in Neuroinformatics*, vol. 14, p. 58, 11 2020.

[37] R. Souza and R. Frayne, "A hybrid frequency-domain/image-domain deep network for magnetic resonance image reconstruction," *Proceedings - 32nd Conference on Graphics, Patterns and Images, SIBGRAPI 2019*, pp. 257–264, 2019.

[38] Z. Wang, H. Jiang, H. Du, J. Xu, and B. Qiu, "IKWI-net: A cross-domain convolutional neural network for undersampled magnetic resonance image reconstruction," *Magnetic resonance imaging*, vol. 73, pp. 1–10, 11 2020.

[39] D. Lee, J. Yoo, S. Tak, and J. C. Ye, "Deep Residual Learning for Accelerated MRI Using Magnitude and Phase Networks," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1985–1995, 9 2018.

[40] B. Zhou and S. Kevin Zhou, "Dudornet: Learning a dual-domain recurrent network for fast MRI reconstruction with deep T1 prior,"

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 4272–4281, 2020.

[41] E. Z. Chen, P. Wang, X. Chen, T. Chen, and S. Sun, "Pyramid Convolutional RNN for MRI Image Reconstruction," *IEEE transactions on medical imaging*, vol. PP, pp. 1–1, 2 2022.

[42] S. A. H. Hosseini, B. Yaman, S. Moeller, M. Hong, and M. Akcakaya, "Dense recurrent neural networks for accelerated mri: History-cognizant unrolling of optimization algorithms," *IEEE Journal on Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1280–1291, 10 2020.

[43] M. Seitzer, G. Yang, J. Schlemper, O. Oktay, T. Würfl, V. Christlein, T. Wong, R. Mohiaddin, D. Firmin, J. Keegan, D. Rueckert, and A. Maier, "Adversarial and perceptual refinement for compressed sensing MRI reconstruction," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11070 LNCS, pp. 232–240, 2018.

[44] E. Cha, G. Oh, and J. C. Ye, "Geometric Approaches to Increase the Expressivity of Deep Neural Networks for MR Reconstruction," *IEEE Journal on Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1292–1305, 10 2020.

[45] M. Seeger, H. Nickisch, R. Pohmann, and B. Schölkopf, "Optimization of k-space trajectories for compressed sensing by Bayesian experimental design," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 116–126, 2010.

[46] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Adapted Compressed Sensing: A Game Worth Playing," *IEEE Circuits and Systems Magazine*, vol. 20, no. 1, pp. 40–60, 2020.

[47] J. P. Haldar and D. Kim, "OEDIPUS: An Experiment Design Framework for Sparsity-Constrained MRI," *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1545–1558, 7 2019.

[48] H. K. Aggarwal and M. Jacob, "J-Modl: Joint model-based deep learning for optimized sampling and reconstruction," *IEEE Journal on Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1151–1162, 10 2020.

[49] I. A. Huijben, B. S. Veeling, and R. J. van Sloun, "Learning Sampling and Model-Based Signal Recovery for Compressed Sensing MRI," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 5 2020, pp. 8906–8910.

[50] Ian Goodfellow and Yoshua Bengio and Aaron Courville, *Deep Learning*. MIT Press, 2016.

[51] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana, Z. Zhang, M. Drozdzal, A. Romero, M. Rabbat, P. Vincent, N. Yakubova, J. Pinkerton, D. Wang, E. Owens, C. L. Zitnick, M. P. Recht, D. K. Sodickson, and Y. W. Lui, "fastMRI: An Open Dataset and Benchmarks for Accelerated MRI."

[52] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 4 2004.

[53] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1028–1041, 5 2011.

[54] (2022, 3) brain-development.org. [Online]. Available: https://brain-development.org/

Filippo Martnini received the M.S. degrees (Hons.) in electronic engineering from the University of Bologna, Bologna, Italy, in 2021; the same year he enrolled as a PhD student at the same university. His research interests include deep neural networks and biomedical applications, in particular fast MRI acquisition. It is the first time Filippo writes of himself using the 3-rd person: he feels like an important and mature person.

Alex Marchioni received the B.S. and M.S. degree (with honors) in electronic engineering and the Ph.D. degree in electronics, telecommunications and information technologies engineering from the University of Bologna, in 2011, 2015 and 2022, respectively. He is currently a postdoc with the Statistical Signal Processing Group of ARCES - University of Bologna. His research interests include compressed sensing, Internet of Things, signal processing, and machine learning with focus on dimensionality reduction and anomaly detection.

Mauro Mangia (S'09-M'13) received the B.Sc. and M.Sc. degrees in electronic engineering and the Ph.D. degree in information technology from the University of Bologna, Bologna, Italy, in 2005, 2009, and 2013, respectively. He was a Visiting Ph.D. Student with the Ecole Polytechnique Federale de Lausanne in 2009 and 2012. He is currently a Assistant Professor in electronics with the Department of Electrical, Electronic, and Information Engineering of the University of Bologna. He is also a member of the advance research center ARCES for the same university within the Statistical Signal Processing Group. His research interests are in nonlinear systems, machine learning, compressed sensing, anomaly detection, Internet of Things, Big Data analytics and optimization.

He was a recipient of the 2013 IEEE CAS Society Guillemin-Cauer Award and of the 2019 IEEE BioCAS Transactions Best Paper Award. He received the Best Student Paper Award at ISCAS2011. He was the Web and Social Media Chair for ISCAS2018.

Riccardo Rovatti (M'99-SM'02-F'12) received the M.S. degree in electronic engineering and the Ph.D. degree in electronics, computer science, and telecommunications from the University of Bologna, Italy, in 1992 and 1996, respectively. He is currently a Full Professor of electronics with the University of Bologna. He has authored approximately 300 technical contributions to international conferences and journals and two volumes. His research focuses on mathematical and applicative aspects of statistical signal processing and on the application of statistics to nonlinear dynamical systems.

He was Distinguished Lecturer of the IEEE CAS Society for the years 2017-2018. He was a recipient of the 2004 IEEE CAS Society Darlington Award, the 2013 IEEE CAS Society Guillemin-Cauer Award and the 2019 IEEE BioCAS Transactions Best Paper Award. He received the Best Paper Award at ECCTD 2005 and the Best Student Paper Award at the EMC Zurich 2005 and ISCAS 2011. He contributed to nonlinear and statistical signal processing applied to electronic systems.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Journal of Selected Topics in Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSTSP.2022.3171082

Gianluca Setti (S89,M91,SM02,F06) received a Ph.D. degree in Electronic Engineering and Computer Science from the University of Bologna in 1997. From 1997 to 2017 he has been with the School of Engineering at the University of Ferrara, Italy as an Assitant, Associate and, since 2009 as a Full Professor of Circuit Theory and Analog Electronics. Since December 2017 he is a Professor of Electronics for Signal and Data Processing at the Department of Electronics and Telecommunications (DET) of Politecnico di Torino, Italy. Since 2002 is also a permanent (in kind) faculty member of ARCES, University of Bologna. His research interests include nonlinear circuits, recurrent neural networks, statistical signal processing, electromagnetic compatibility, compressive sensing, biomedical circuist and systems, power electronics, design and implementation of IoT nodes.

Dr. Setti received the 1998 Caianiello prize for the best Italian Ph.D. thesis on Neural Networks. He is also receipient of the 2013 IEEE CAS Society Meritorious Service Award and co-recipient of the 2004 IEEE CAS Society Darlington Award, of the 2013 IEEE CAS Society Guillemin-Cauer Award, the 2019 IEEE Transactions on Biomedical Circuits and Systems best paper award, as well as of the best paper award at ECCTD2005, and the best student paper award at EMCZurich2005, ISCAS2011, PRIME2019 and EMCCOMPO 2019.

He held several editorial positions and served, in particular, as the Editor-in-Chief for the IEEE Transactions on Circuits and Systems - Part II (2006-2007) and of the IEEE Transactions on Circuits and Systems - Part I (2008-2009). He also served in the editorial Board of IEEE Access (2013-2015) and, since of the Proceedings of the IEEE (2015-2018). Since 2019 he served as the first non US Editor-in-Chief of the Proceedings of the IEEE, the flagship journal of the Institute.

Dr. Setti was the Technical Program Co-Chair of NDES2000 (Catania), ISCAS2007 (New Orleans), ISCAS2008 (Seattle), ICECS2012 (Seville), BioCAS2013 (Rotterdam) as well as the General Co-Chair of NOLTA2006 (Bologna) and ISCAS2018 (Florence).

He was a Distinguished Lecturer (2004-2005 and 2014-2015) of the IEEE CAS Society, as well as a member of its Board of Governors (2005-2008), and served as the 2010 CASS President. He held several other volunteer positions for the IEEE and in 2013-2014 he was the first non NorthAmerican Vice President of the IEEE for Publication Services and Products.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.