



Heriot-Watt University  
Research Gateway

# Overcoming Measurement Inconsistency In Deep Learning For Linear Inverse Problems: Applications In Medical Imaging

## Citation for published version:

Vella, M & Mota, JFC 2021, Overcoming Measurement Inconsistency In Deep Learning For Linear Inverse Problems: Applications In Medical Imaging. in *2021 International Conference on Acoustics, Speech, and Signal Processing*. IEEE, pp. 8113-8117, 46th International Conference on Acoustics, Speech, and Signal Processing 2021, 6/06/21. <https://doi.org/10.1109/ICASSP39728.2021.9414173>

## Digital Object Identifier (DOI):

[10.1109/ICASSP39728.2021.9414173](https://doi.org/10.1109/ICASSP39728.2021.9414173)

## Link:

[Link to publication record in Heriot-Watt Research Portal](#)

## Document Version:

Peer reviewed version

## Published In:

2021 International Conference on Acoustics, Speech, and Signal Processing

## Publisher Rights Statement:

© 2021 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 collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

## General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

## Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [open.access@hw.ac.uk](mailto:open.access@hw.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

# OVERCOMING MEASUREMENT INCONSISTENCY IN DEEP LEARNING FOR LINEAR INVERSE PROBLEMS: APPLICATIONS IN MEDICAL IMAGING

Marija Vella

João F. C. Mota

Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh, UK

## ABSTRACT

The remarkable performance of deep neural networks (DNNs) currently makes them the method of choice for solving linear inverse problems. They have been applied to super-resolve and restore images, as well as to reconstruct MR and CT images. In these applications, DNNs invert a forward operator by finding, via training data, a map between the measurements and the input images. It is then expected that the map is still valid for the test data. This framework, however, introduces measurement inconsistency during testing. We show that such inconsistency, which can be critical in domains like medical imaging or defense, is intimately related to the generalization error. We then propose a framework that post-processes the output of DNNs with an optimization algorithm that enforces measurement consistency. Experiments on MR images show that enforcing measurement consistency via our method can lead to large gains in reconstruction performance.

**Index Terms**— Neural networks, linear inverse problems, medical imaging, optimization, total variation.

## 1. INTRODUCTION

Many applications in science and engineering require solving inverse problems in which the number of available measurements is much smaller than the number of parameters to be estimated. Examples include various medical imaging modalities, remote sensing, image restoration, seismography, and LiDAR depth estimation. In such problems, we have access to  $m$  linear measurements of a vector  $x^* \in \mathbb{C}^n$  that we wish to estimate. Formally, we have the linear system

$$b = Ax^* + e, \quad (1)$$

where  $b \in \mathbb{C}^m$  is the vector of known measurements,  $A \in \mathbb{C}^{m \times n}$  is a known measurement matrix with more columns than rows, i.e.,  $m < n$ , and  $e \in \mathbb{C}^m$  is a noise vector. As the number of measurements in (1) is smaller than the number of variables, there is often an infinite number of vectors  $x \in \mathbb{C}^n$  satisfying (1).

The classical approach of inferring  $x^*$  from (1) formulates an optimization problem for finding the simplest solution of (1) according to the known structure of  $x^*$ . Such structure is captured by regularizers such as the  $\ell_1$ -norm [1], which enforces sparsity in a given domain, or different versions like the total variation (TV) norm [2] and structural or hierarchical priors [3]. In recent years, however, these optimization-based methods for linear inverse problems have been surpassed by data-driven approaches, specifically, deep (convolutional) neural networks (DNNs). By leveraging large datasets of input-output pairs  $(x^*, b)$  during training, DNNs are able to automatically learn the structure of typical signals. During deployment, this enables them not only to reconstruct  $x^*$  with quality better than optimization approaches, but also to do it faster. This

phenomenon has been observed in several linear inverse problems, including single-image super-resolution [4], denoising [5], biomedical imaging [6], and LiDAR depth estimation [7].

Despite these successes, DNNs still suffer from important drawbacks that have slowed down their application in critical domains, such as autonomous driving or fully automated medical diagnosis. These include overfitting, lack of sharp theoretical guarantees, and instability with respect to small data perturbations [8]. A related drawback, which is particularly important in linear inverse problems and which we explore in this paper, is measurement inconsistency.

**Measurement inconsistency in DNNs.** Typically, DNNs for linear inverse problems are trained by minimizing a real-valued loss function  $\ell : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{R}$ , usually the  $\ell_2$ -norm, over all the  $T$  samples  $\{x^{(t)}\}_{t=1}^T$  of the training set. That is, if  $f_\theta : \mathbb{C}^m \rightarrow \mathbb{C}^n$  represents a DNN with parameters  $\theta$ , one finds the optimal set of parameters  $\theta^*$  by solving

$$\underset{\theta}{\text{minimize}} \quad \frac{1}{T} \sum_{t=1}^T \ell(x^{(t)}, f_\theta(Ax^{(t)})), \quad (2)$$

where, to simplify, we assumed that the noise  $e$  in (1) is negligible. Hence, the argument of  $f_\theta$  in (2) represents the  $t$ th instance of the measurement vector  $b$  in (1) with  $e = 0$ . Once trained, given a vector of measurements  $b$  from a sample  $x^*$  absent from the training set, the DNN estimates  $x^*$  by simply applying a forward pass to  $b$ :

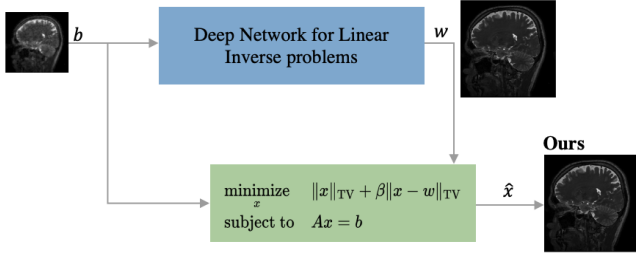
$$w := f_{\theta^*}(b). \quad (3)$$

We say that the trained DNN  $f_{\theta^*}$  is *measurement inconsistent* when  $Aw \neq b$ .<sup>1</sup> In fact, we show in Proposition 1 that measurement inconsistency is a by-product of generalization errors, even when the DNN is trained to minimize consistency, i.e., when the loss function in (2) is  $\ell(x, w) := \|Ax - Aw\|_2^2$ .

This calls for a new framework that not only enforces data consistency, but also can harness the excellent performance of DNNs. And while the data consistency problem has been identified as an important one, for example, in medical imaging, the existing solutions consist mostly of modifying the network architecture. For instance, [9, 10] introduce skip-connections between the input and output, and [11] trains a DNN using different models of  $A$ . These solutions still fall under the framework of (2) and, thus, suffer from the problem highlighted in Proposition 1. Our approach to solve this problem consists of post-processing the output of the DNN,  $w$  in (3), with an optimization algorithm that enforces consistency.

**Our approach and contributions.** Fig. 1 shows a diagram of the scheme we propose. The scheme addresses the measurement consistency problem by formulating an optimization problem,

<sup>1</sup>This definition can be extended to the case in which  $e$  is non-negligible but, for example, bounded:  $\|e\|_2 \leq \sigma$ . In this case, we say the DNN is measurement inconsistent when  $\|Aw - b\|_2 > \sigma$ .



**Fig. 1:** Our framework. Measurement consistency is enforced by post-processing the output of a DNN designed for a linear inverse problem with an optimization problem with  $Ax = b$  as a constraint.

named *TV-TV minimization*, that explicitly enforces consistency via constraints, while minimizing a cost function that balances a small total variation (TV) of the reconstructed image — as in classical optimization-based approaches — and proximity to the output of the DNN, as measured by the TV-norm. This last term is the key element of our approach: it allows combining the major benefit of optimization-based methods, i.e., the ability to enforce constraints explicitly, with the major benefit of data-driven methods, i.e., their outstanding performance. Our experiments show that our framework improves the reconstruction performance of state-of-the-art DNNs for MRI reconstruction, e.g., MoDL [12], by 5 dB in PSNR. We summarize our contributions as follows:

- We show that the conventional method for training DNNs according to empirical loss minimization (2) cannot solve the measurement consistency problem in linear inverse problems; this is done in Proposition 1 and illustrated with experiments.
- We propose a framework that addresses the measurement consistency problem by combining an optimization-based method, TV-TV minimization, with DNNs.
- We apply our framework to MRI and show that it can lead to significant gains with respect to state-of-the-art DNNs.

We point out that we proposed a particular instance of this framework in [13, 14] for single-image super-resolution. There, the observed gains in PSNR were much smaller ( $< 1$  dB) than the ones observe here for MRI. In this paper, we address the consistency problem theoretically. And to apply TV-TV minimization to MRI, we had to generalize the algorithm to handle complex vectors and matrices.

After overviewing prior work in Section 2, we present our theoretical result on consistency in Section 3. Section 4 describes our framework, and Section 5 our experiments on MRI reconstruction.

## 2. RELATED WORK

**Optimization-based approaches.** Inverse linear problems have traditionally been solved via optimization algorithms. The associated optimization problem is formulated such that its solution simultaneously conforms to the observed measurements and is sufficiently simple according to some prior knowledge. The latter is usually encoded by sparsity in some domain, such as wavelet or DCT representations [1], or in gradient space [2]. Sparsity can be encoded directly via an  $\ell_0$ -pseudo-norm, which leads to nonconvex problems that can be solved via greedy approaches [15, 16], or via a convex relaxation such as the  $\ell_1$ -norm, which leads to convex problems that can be provably solved [17]. Convex formulations also usually have strong theoretical reconstruction guarantees [18]. Indeed, [18] generalizes

the concept of simplicity via sparsity to *atomic norms*, which apply to a wider range of problems. In the field of MRI reconstruction, a common choice for simplicity is sparsity in the gradient domain, which is captured by a total variation (TV) norm [2, 19, 20] or generalizations of the TV-norm [21].

**DNN-based methods.** Deep neural networks (DNNs) have been successfully applied in various linear inverse problems [4, 5, 6, 7, 10, 22]. These networks learn to invert a forward model by leveraging several input-output pairs obtained from a single measurement operator. Once trained, they can be used to reconstruct the input of the operator from its output by a simple forward pass. This operator, however, often fails to guarantee measurement consistency. Attempts to solve this problem include adding skip-connections to share information from the input to the final layer [9, 10, 23], and embedding a data consistency layer in the network [22, 24]. Another approach is to unroll iterative optimization algorithms that alternate between data consistency and a nonlinear operation related to prior knowledge. This idea was first proposed [25], and it spawned different lines of research, e.g., [26, 12, 27]. For example, [12] trains a DNN denoiser whose output is then fed to an optimization block that acts as a data consistency layer [12]. As our experiments show, this is still not enough to guarantee consistency.

**Instability in DNNs.** In recent years, several issues have been identified in the deployment of DNNs, including overfitting and data memorization [28], and adversarial examples [29]. For linear inverse problems, [8] studied how small perturbations to the measurement or to the sampling method can lead to artefacts in the reconstruction. Moreover, the reconstructed outputs may miss fine details that are present in measurements, a feature that can have critical consequences in medical imaging.

To overcome this, we present a framework that leverages the good performance of DNNs, while enforcing measurement consistency. Consequently, the risk of missing such details is reduced. Before presenting our framework, we first see how inconsistency in DNNs is related to the generalization error.

## 3. THE PROBABILITY OF INCONSISTENCY IN DNNs

Here we adopt a probabilistic setting to analyze the measurement inconsistency problem of DNNs. Henceforth, we assume the model in (1) with negligible noise, i.e.,  $e = 0$ . Let  $X \in \mathbb{C}^n$  represent a vector of  $n$  random complex variables whose  $i$ th component is  $X_i = \text{Re}\{X_i\} + j\text{Im}\{X_i\}$ . To avoid technicalities, we assume all the functions we deal with are measurable, including the DNN  $f_\theta : \mathbb{C}^m \rightarrow \mathbb{C}^n$ , for any  $\theta$ . Without loss of generality, we assume the following squared  $\ell_2$ -norm loss:  $\ell(x, w) := \|Ax - Aw\|_2^2$ . Hence, the *expected loss* associated to  $f_\theta$  is defined as

$$\ell_{\text{exp}}(f_\theta) := \mathbb{E} \left[ \|AX - Af_\theta(AX)\|_2^2 \right], \quad (4)$$

where expectation is with respect to  $X$ . The *empirical loss* of  $f_\theta$  on the training set  $\mathcal{T} := \{x^{(t)}\}_{t=1}^T$ , where  $x^{(t)}$  is a realization of  $X$ , is

$$\ell_{\text{emp}}(f_\theta; \mathcal{T}) := \frac{1}{T} \sum_{t \in \mathcal{T}} \|Ax^{(t)} - Af_\theta(Ax^{(t)})\|_2^2. \quad (5)$$

Notice that these definitions differ from conventional definitions of expected and empirical loss in two ways (see, e.g., [30, 27] for examples in linear inverse problems). First, the probability distribution is defined over the output space  $X$  only. Indeed, under the assumption that the measurements in (1) are noiseless, the input random

variable  $B \in \mathbb{C}^m$  is completely specified by  $X$ :  $B = AX$ . Second, the  $\ell_2$ -norm is applied to the output rather than to the input space: that is, in (4), we consider  $\|AX - Af_\theta(AX)\|_2^2$  rather than  $\|X - f_\theta(AX)\|_2^2$ . This reflects a training strategy to minimize inconsistency, and captures the type of regularization terms used in deep prior or network unrolling models, e.g., MoDL [12]. This is without loss of generality, as the loss can include additional terms. The following proposition bounds the probability of the DNN outputting an inconsistent result as a function of the generalization error  $c := \ell_{\text{exp}}(f_\theta) - \ell_{\text{emp}}(f_\theta; \mathcal{T})$ .

**Proposition 1.** Consider  $f_{\theta^*}$  with a parameter  $\theta^*$  that achieves an empirical loss  $\epsilon := \ell_{\text{emp}}(f_{\theta^*}; \mathcal{T}) > 0$ . Assume the random variable  $Y := \|AX - Af_{\theta^*}(AX)\|_2^2$  is upper bounded by  $C$  almost surely. Assume a positive generalization error  $c := \ell_{\text{exp}}(f_\theta) - \epsilon > 0$ . Then, for any  $\delta$  such that  $0 < \delta < c + \epsilon$ ,

$$\mathbb{P}\left(\|AX - Af_{\theta^*}(AX)\|_2^2 \geq \delta\right) \geq 1 - \exp\left(-2\frac{(c + \epsilon - \delta)^2}{C^2}\right). \quad (6)$$

*Proof.* By assumption,  $0 \leq Y \leq C$  almost surely, which implies that  $Y$  is sub-Gaussian with parameter  $\sigma := C/2$  [31]. Therefore, for any  $t > 0$ ,

$$\mathbb{P}(Y < \delta) = \mathbb{P}(Y - \mathbb{E}[Y] < \delta - \mathbb{E}[Y]) \quad (7)$$

$$= \mathbb{P}(Y - \mathbb{E}[Y] < \delta - c - \epsilon) \quad (8)$$

$$\leq \mathbb{P}(|Y - \mathbb{E}[Y]| > c + \epsilon - \delta) \quad (9)$$

$$\leq \exp\left(-t(c + \epsilon - \delta) + \frac{t^2\sigma^2}{2}\right). \quad (10)$$

From (7) to (8), we used the definition of  $c = \mathbb{E}[Y] - \epsilon$ . From (8) to (9), we used the fact that  $\delta < c + \epsilon$ . And from (9) to (10), we applied a Chernoff bound [32, §4.2] taking into account that  $Y$  is sub-Gaussian with parameter  $\sigma$  [31, §2.2.1]. Setting  $t = (c + \epsilon - \delta)/\sigma^2$ , replacing  $\sigma = C/2$ , and taking the complementary event in (10) yields (6).  $\square$

The left-hand side of (6) expresses the probability that the output of the DNN is inconsistent, and the right-hand side increases as a function of the generalization gap  $c$  as well as the empirical loss  $\epsilon$ . The assumption that the random variable  $Y = \|AX - Af_{\theta^*}(AX)\|_2^2$  is bounded above can be easily relaxed: the proof can be adapted if we assume  $Y$  is sub-Gaussian or even sub-exponential (fat tails).

## 4. OUR FRAMEWORK

To address the measurement inconsistency problem studied in the previous section, we propose the framework represented in Fig. 1. In order to understand how it differs from previous approaches, recall that conventional DNN-based methods for linear inverse problems use empirical loss minimization (2) to find a good-enough parameter  $\theta^*$  of a DNN  $f_\theta : \mathbb{C}^m \rightarrow \mathbb{C}^n$ . Then, during testing, they simply apply  $w := f_{\theta^*}(b)$  to measurements  $b = Ax^*$  of unseen data  $x^*$ .<sup>2</sup> This framework applies to DNNs designed for a specific inverse problem, e.g., single-image super-resolution [4], as well as unrolled networks, e.g., [25, 27]. However, as shown in Proposition 1, a generalization error typically implies inconsistency of the DNN output to the measurements, i.e.,  $Af_{\theta^*}(b) \neq b$ .

**TV-TV minimization.** To overcome this problem, and as shown in Fig. 1, we feed the measurements  $b$  together with the output  $w =$

$f_{\theta^*}(b)$  of a given DNN, which is generally very close to the desired  $x^*$ , but measurement inconsistent, into an optimization problem that we call TV-TV minimization:

$$\begin{aligned} & \underset{x}{\text{minimize}} && \|x\|_{\text{TV}} + \beta\|x - w\|_{\text{TV}} \\ & \text{subject to} && Ax = b. \end{aligned} \quad (11)$$

In the objective function,  $\|\cdot\|_{\text{TV}}$  stands for the 2D TV semi-norm of an image  $\bar{x} \in \mathbb{C}^{M \times N}$  [whose vectorization is  $x \in \mathbb{C}^n$  with  $n = M \cdot N$ ]. It is defined as  $\|x\|_{\text{TV}} := \sum_{i=1}^M \sum_{j=1}^N |v_{ij}^\top x| + |h_{ij}^\top x| = \|Dx\|_1$ , where  $v_{ij}, h_{ij} \in \mathbb{R}^n$  are real vectors that extract the vertical and horizontal differences at pixel  $(i, j)$  of  $\bar{x}$ , and  $D \in \mathbb{R}^{2n \times n}$  is the vertical concatenation of  $v_{ij}^\top$  and  $h_{ij}^\top$  for all  $i = 1, \dots, M$  and  $j = 1, \dots, N$ . The first term,  $\|x\|_{\text{TV}}$ , encodes the assumption that the image to reconstruct has a small number of edges compared to its dimension. This is a standard approach in optimization-based methods for image restoration [2, 33] and MRI reconstruction [20]. The second term,  $\beta\|x - w\|_{\text{TV}}$ , specifies that the solution of (11) should be close to the output  $w$  of the DNN, in a TV-norm sense. Here,  $\beta$  balances between the two terms of the objective. According to the theory in [34] and the experiments in [13, 14], a value close to  $\beta = 1$  yields the best results. Finally, consistency is achieved by constraining the solution of (11) to satisfy  $Ax = b$ . Notice that (11) can be modified to accommodate noisy measurements as in (1). However, as will be shown section 5, we found that even when considering a noiseless (and thus inaccurate) model, post-processing the output  $w$  of a state-of-the-art DNN for MRI reconstruction via (11) leads to significant performance gains.

**Algorithm for solving (11).** We apply ADMM [35] to a reformulation of (11). The main idea and algorithm are described in [14]. But to apply the resulting algorithm to MRI reconstruction, we had to slightly change the reformulations in order to handle complex vectors and matrices. Details will be described in a forthcoming paper.

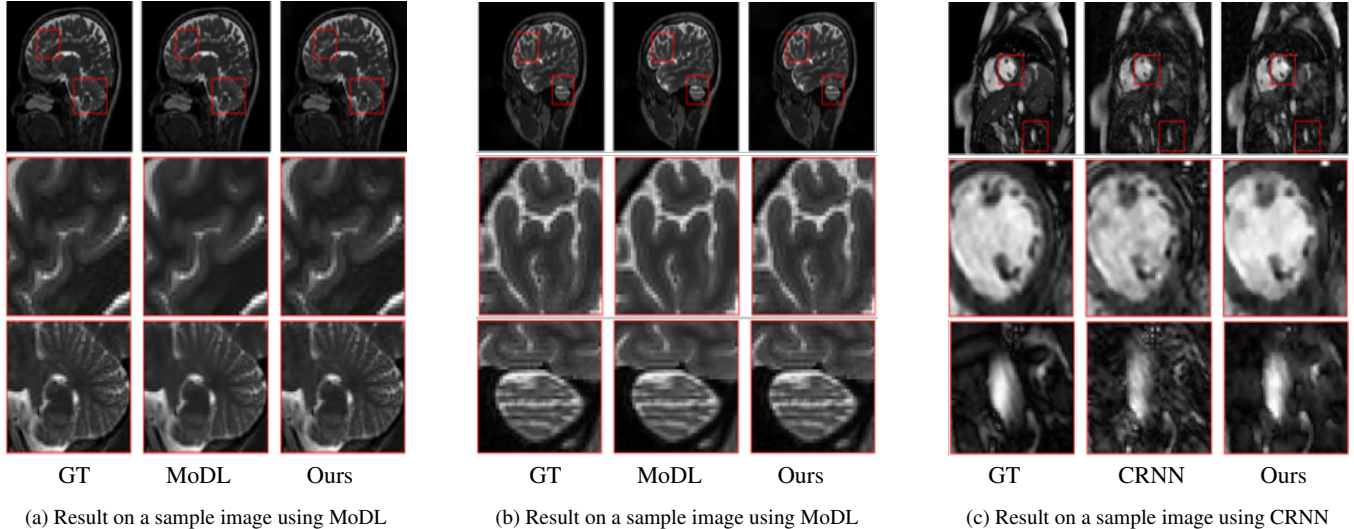
## 5. APPLICATION TO MRI RECONSTRUCTION

We now describe our experiments for MRI reconstruction. After explaining the setup, we illustrate how the proposed framework solves the measurement consistency problem and how this leads to significant gains in reconstruction performance.

**Experimental setup.** We applied our framework to two state-of-the-art MRI reconstruction DNNs: MoDL [12] and CRNN [24]. MoDL reconstructs multichannel MRI images in which data is acquired by a 12-channel head coil, and was trained on the multichannel brain dataset acquired by the authors of [12]. For testing, we used 164 slices from a single subject, resulting in a test dataset of dimensions  $256 \times 232 \times 164 \times 12$  (rows  $\times$  columns  $\times$  slices  $\times$  number of coils). CRNN, in turn, applies to data acquired by a single-channel coil, and was trained on cardiac images from various subjects [24]. To avoid retraining CRNN, we used a pretrained version of the network, which was trained for a single subject. As we will see, this led to low quality outputs. Both methods process the real and imaginary parts of the complex MRI data separately. We set an acceleration factor of 6 for MoDL and of 4 for CRNN. During testing, we added no artificial noise to the data.

For MoDL (resp. CRNN), we set  $\beta = 1$  (resp.  $\beta = 0.8$ ) in (11) and ran our algorithm a maximum number of 100 (resp. 50) ADMM iterations. The measurement matrix  $A$  in (1) for MoDL was the product of a sampling mask  $S$ , a coil sensitivity map  $C$ , and a 2D discrete Fourier transform  $F$ , i.e.,  $A = SFC$ , and for CRNN it was just  $A = SF$ , as this network operates on a single coil only.

<sup>2</sup>As mentioned, here, we consider the acquisition noise  $e$  in (1) negligible.



**Fig. 2:** Reconstruction of two test images (a)-(b) for MoDL [12], and (c) one test image for CRNN [24]. Each figure shows the full image in the top row, and two zoomed regions in the two bottom rows; the left column shows the groundtruth (GT), and the right one our method.

**Table 1:** Measurement consistency of MoDL, CRNN (2nd column), and corresponding consistency of our method (3rd column).

Method	$\ Aw - b\ _2$	$\ A\hat{x} - b\ _2$
MoDL [12]	$3.10 \times 10^{-1}$	<b><math>9.88 \times 10^{-5}</math></b>
CRNN [24]	$2.06 \times 10^{-6}$	<b><math>7.71 \times 10^{-15}</math></b>

**Table 2:** PSNR and SSIM in the format average  $\pm$  std, min/max. The best values are highlighted in bold.

Method	PSNR	SSIM
MoDL	39.06 $\pm$ 1.58, 33.86/40.91	0.97 $\pm$ 0.02, 0.84/0.99
Ours	<b>45.96 <math>\pm</math> 3.94, 35.48/53.45</b>	<b>0.98 <math>\pm</math> 0.02, 0.85/1.00</b>
CRNN	24.08 $\pm$ 0.59, 22.91/25.29	0.71 $\pm$ 0.03, 0.64/0.78
Ours	<b>25.45 <math>\pm</math> 0.71, 24.17/26.70</b>	<b>0.76 <math>\pm</math> 0.02, 0.71/0.80</b>

Because background noise in the images introduced some variation in the results, we evaluated the performance metrics, PSNR and SSIM, on images cropped to the relevant anatomy content.

**Measurement consistency.** Table 1 displays the consistency metric  $\|Ax - b\|_2$  for the outputs of MoDL and CRNN (2nd column), and the respective metric after post-processing with our algorithm (3rd column). The input and output images are displayed in Fig. 2a for MoDL and in Fig. 2c for CRNN. It can be seen that our algorithm reduces this metric by 4 orders of magnitude for MoDL, and by 9 orders of magnitude for CRNN. Although  $\|Aw - b\|_2$  for CRNN is already very small, this does not necessarily translate into good reconstruction performance, as we will see next.

**Reconstruction performance.** Table 2 shows the results we obtained on the brain dataset for MoDL, and on the 30 cardiac test set for CRNN. The 2nd (resp. 3rd) column displays the average PSNR (resp. SSIM) and respective standard deviation over the test images. The first (resp. last) two rows refer to the performance of MoDL

(resp. CRNN) and of the subsequent processing with our method.

In the case of MoDL, it can be seen that our post-processing increased the PSNR performance by more than 5dB. We also observe a marginal degradation in SSIM. The reason may be because the SSIM values for MoDL were already large and, being constrained to the interval  $[0, 1]$ , were difficult to increase. However, Figs. 2a-2b, which show two examples of test images, demonstrate visually that our method preserves edges, whereas MoDL over-smooths them. The results in Table 1 indicate that this is a by-product of enforcing consistency.

For CRNN, the last two rows of Table 2 show that our method improved both the PSNR and SSIM values. The gains, however, were much smaller, likely because CRNN enforces consistency better than MoDL. The table also shows that the reconstruction performance using CRNN is much worse than using MoDL, in part because this network was trained on a single subject. Fig. 2c demonstrates visually that our method enables preserving details better.

## 6. CONCLUSIONS

We studied the phenomenon of measurement inconsistency in DNNs for linear inverse problems. We achieved this by relating the probability of obtaining an inconsistent output to the generalization error. To overcome this problem, we then proposed a post-processing algorithm that improves the output of DNNs by enforcing consistency. Experimental results on MRI reconstruction showed that applying our algorithm not only leads to better consistency, but also to significant reconstruction gains. And the better the improvement in consistency, the larger the gains.

## 7. REFERENCES

- [1] S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” *SIAM J. Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1998.
- [2] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation

- based noise removal algorithms,” *Physica D*, vol. 60, pp. 259–268, 1992.
- [3] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [4] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016.
- [5] K. Zhang, W. Zuo, and L. Zhang, “FFDNet: Toward a fast and flexible solution for CNN-based image denoising,” *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, 2018.
- [6] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, 2017.
- [7] J. Uhrig, N. Schneider, U. Franke, T. Brox, and A. Geiger, “Sparsity invariant CNNs,” in *Intern. Conf. 3D Vision (3DV)*, 2017, pp. 11–20.
- [8] V. Antun, F. Renna, C. Poon, B. Adcock, and A. Hansen, “On instabilities of deep learning in image reconstruction and the potential costs of AI,” *Proc. National Academy of Sciences*, 2020.
- [9] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, 2017.
- [10] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “Enhanced deep residual networks for single image super-resolution,” in *Proc. CVPR Workshop*, July 2017.
- [11] K. Zhang, W. Zuo, and L. Zhang, “Learning a single convolutional super-resolution network for multiple degradations,” in *Proc. CVPR*, 2018, pp. 3262–3271.
- [12] H. K. Aggarwal, M. P. Mani, and M. Jacob, “Modl: Model-based deep learning architecture for inverse problems,” *IEEE Trans. Med. Imaging*, vol. 38, no. 2, pp. 394–405, 2019.
- [13] M. Vella and J. F. C. Mota, “Single image super-resolution via cnn architectures and TV-TV minimization,” in *British Machine Vision Conference*, 2019.
- [14] M. Vella and J. F. C. Mota. (2020) Robust single-image super-resolution via CNNs and TV-TV minimization. ArXiv:2004.00843.
- [15] D. L. Donoho, Y. Tsaig, I. Drori, and J. Starck, “Sparse solution of underdetermined systems of linear equations by stage-wise orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.
- [16] D. Needell and J. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [18] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky, “The convex geometry of linear inverse problems,” *Found. Comput. Analysis*, vol. 12, no. 6, pp. 805–849, 2012.
- [19] K. T. Block, M. Uecker, and J. Frahm, “Undersampled radial MRI with multiple coils. iterative image reconstruction using a total variation constraint,” *Magn. Res. in Medicine*, vol. 57, no. 6, pp. 1086–1098, 2007.
- [20] L. Weizman, Y. C. Eldar, and D. Ben-Bashat, “Reference-based MRI,” *Med. Phys.*, vol. 43, no. 10, pp. 5357–5369, 2016.
- [21] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, “Second order total generalized variation (TGV) for MRI,” *Magn. Res. in Medicine*, vol. 65, no. 2, pp. 480–491, 2011.
- [22] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, “A deep cascade of convolutional neural networks for dynamic MR image reconstruction,” *IEEE Trans. Med. Imaging*, vol. 37, no. 2, pp. 491–503, 2018.
- [23] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, “CNN-based projected gradient descent for consistent CT image reconstruction,” *IEEE Trans. Med. Imaging*, vol. 37, no. 6, pp. 1440–1453, 2018.
- [24] C. Qin, J. Schlemper, J. Caballero, A. N. Price, J. V. Hajnal, and D. Rueckert, “Convolutional recurrent neural networks for dynamic MR image reconstruction,” *IEEE Trans. Med. Imaging*, vol. 38, no. 1, pp. 280–290, 2019.
- [25] K. Gregor, Karol, and Y. LeCun, “Learning fast approximations of sparse coding,” in *Proc. ICML*, 2010, p. 399–406.
- [26] J. H. R. Chang, C. Li, B. Poczos, B. V. K. V. Kumar, and A. C. Sankaranarayanan, “One network to solve them all? solving linear inverse problems using deep projection models,” in *Proc. ICCV*, 2017, pp. 5889–5898.
- [27] H. Van Luong, B. Joukovsky, and N. Deligiannis. (2020) Interpretable deep recurrent neural networks via unfolding reweighted  $\ell_1$ - $\ell_1$  minimization: Architecture design and generalization analysis. ArXiv:2003.08334.
- [28] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, “Understanding deep learning requires rethinking generalization,” in *Int. Conf. Learning Repr.*, 2017.
- [29] S.-M. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, “Deepfool: a simple and accurate method to fool deep neural networks,” in *Proc. CVPR*, 2017.
- [30] J. Amjad, Sokolić, and M. R. D. Rodrigues, “On deep learning for inverse problems,” in *Proc. EUSIPCO*, 2018, pp. 1895–1899.
- [31] M. J. Wainwright, *High-Dimensional Statistic: A Non-Asymptotic Viewpoint*. Cambridge University Press, 2019.
- [32] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.
- [33] C. Li, W. Yin, H. Jiang, and Y. Zhang, “An efficient augmented Lagrangian method with applications to total variation minimization,” *Comput. Optim. and Applications*, vol. 56, no. 3, pp. 507–530, 2013.
- [34] J. F. C. Mota, N. Deligiannis, and M. R. D. Rodrigue, “Compressed sensing with prior information: Strategies, geometry, and bounds,” *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4472–4496, 2017.
- [35] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. E. J., “Distributed optimization and statistical learning via the alternating method of multipliers,” *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.