# Interpretable deep learning for multimodal super-resolution of medical images*


Evaggelia Tsiligianni[1][0000−0003−2876−5788], Matina Zerva[1][0000−0002−6420−5377], Iman Marivani[2,3][0000−0002−2928−4014], Nikos Deligiannis[2,3][0000−0001−9300−5860], and Lisimachos Kondi[1][0000−0002−0678−4526]

[1] Department of Computer Science and Engineering, University of Ioannina, Ioannina, Greece
{etsiligia, szerva, lkon}@cse.uoi.gr

[2] Department of Electronics and Informatics, Vrije Universiteit Brussel, Brussels, Belgium
{imarivan,ndeligia}@etrovub.be

[3] imec, Kapeldreef 75, B-3001, Leuven, Belgium



**Abstract.** In medical image acquisition, hardware limitations and scanning time constraints result in degraded images. Super-resolution (SR) is a post-processing approach aiming to reconstruct a high-resolution image from its low-resolution counterpart. Recent advances in medical image SR include the application of deep neural networks, which can improve image quality at a low computational cost. When dealing with medical data, accuracy is important for discovery and diagnosis, therefore, interpretable neural network models are of significant interest as they enable a theoretical study and increase trustworthiness needed in clinical practice. While several interpretable deep learning designs have been proposed to treat unimodal images, to the best of our knowledge, there is no multimodal SR approach applied for medical images. In this paper, we present an interpretable neural network model that exploits information from multiple modalities to super-resolve an image of a target modality. Experiments with simulated and real MRI data show the performance of the proposed approach in terms of numerical and visual results.



**Keywords:** medical image super-resolution · interpretable neural networks · deep unfolding · coupled sparse representations.


## 1 Introduction

Image super-resolution (SR) is a well-known inverse problem in imaging applications. Depending on the number of the employed imaging modalities, SR


* This work has been co-funded by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATE (project code: T1EDK03895)

techniques can be divided into single modal and multimodal. Single modal SR aims to reconstruct a high-resolution (HR) counterpart of a given low-resolution (LR) image of the same modality. Multimodal SR uses complementary information from multiple modalities to recover a target modality. In medical imaging, multiple modalities are coming from different scanning devices or different hardware configurations. Acquisition time constraints, hardware limitations, human body motion etc., result in low-resolution images; therefore, applying a post-processing SR technique to improve the quality of the modality of interest is a considered approach in medical applications [6].

Existing image reconstruction approaches include conventional methods and data-driven techniques. Conventional methods model the physical processes underlying the problem and incorporate domain knowledge; however, the associated iterative optimization algorithms, typically, have a high computational complexity. Among data-driven techniques, deep learning (DL) is popular [12, 17, 21, 29, 24, 30] as it can dramatically reduce the computational cost at the inference step [11]. Nevertheless, neural networks have generic architectures and it is unclear how to incorporate domain knowledge. As a result, one can hardly say what a model has learned. When dealing with medical data, the accuracy and trustworthiness of reconstruction is critical for discovery and diagnosis. Therefore, finding a balance between accuracy and latency raises a significant challenge [27].

Bridging the gap between conventional methods and DL has motivated the design of interpretable neural networks [15, 18]. Recently, a principle referred to as deep unfolding has received a lot of attention [3, 7, 22]. The idea is to unfold the iterations of an inference algorithm into a deep neural network, offering interpretability of the learning process. The model parameters across layers are learned from data and the inference is performed at a fixed computational cost. The approach has been applied to medical imaging in [1, 19, 26]; however, existing works deal with unimodal data. To the best of our knowledge, no interpretable DL design has been reported for multimodal medical image reconstruction.

In this paper, we assume that the similarity between different imaging modalities can be captured by coupled sparse representations that are similar by means of the $\ell_1$-norm. We formulate a coupled sparse representation problem which can be solved with an iterative thresholding algorithm. The algorithm is unfolded into a neural network form, resulting in a learned multimodal convolutional sparse coding model (LMCSC). We incorporate LMCSC into a network that can reconstruct an HR image of a target modality from an LR input with the aid of another guidance modality.

We apply our model to multi-constrast Medical Resonance Imaging (MRI). MRI images with different contrast mechanisms (T1-weighted, T2-weighted, FLAIR) provide different structural information about body tissues. However, the long acquisition process can result in motion-related artifacts. To reduce the acquisition time, a compromise is to generate an LR T2W image and a corresponding HR T1W (or FLAIR) image with a short acquisition time and then obtain an HR T2W image by using multimodal SR methods. Our experiments are conducted on two benchmark datasets, showing that the proposed model achieves state-of-the-art performance.

The paper is organized as follows. Section 2 provides the necessary background on sparse modelling, and Section 3 reports related work on deep unfolding. The proposed model is presented in Section 4, while experiments are included in Section 5. Finally, conclusions are drawn in Section 6.

## 2 Sparse Modelling for Image Reconstruction

Linear inverse problems in imaging are typically formulated as follows [20]:

$$\boldsymbol{y} = \boldsymbol{L}\boldsymbol{x} + \boldsymbol{\eta}, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^k$ is a vectorized form of the unknown source image, $\boldsymbol{y} \in \mathbb{R}^n$ denotes the degraded observations and $\boldsymbol{\eta} \in \mathbb{R}^n$ is the noise[4]. The linear operator $\boldsymbol{L} \in \mathbb{R}^{n\times k}$, $n < k$, describes the observation mechanism. In image SR, $\boldsymbol{L}$ can be expressed as the product of a downsampling operator $\boldsymbol{E}$ and a blurring filter $\boldsymbol{H}$ [25].

Even when the linear observation operator $\boldsymbol{L}$ is given, problem (1) is ill-posed and needs regularization. Following a sparse modelling approach, we assume that $\boldsymbol{x} = \boldsymbol{D}_x\boldsymbol{u}$, with $\boldsymbol{D}_x \in \mathbb{R}^{k\times m}$, $k \leq m$, denoting a representation dictionary, and $\boldsymbol{u} \in \mathbb{R}^m$ being a sparse vector. Then, (1) can be written as $\boldsymbol{y} = \boldsymbol{A_x}\boldsymbol{u} + \boldsymbol{\eta}$, with $\boldsymbol{A}_x = \boldsymbol{L}\boldsymbol{D}_x$, $\boldsymbol{A}_x \in \mathbb{R}^{n\times m}$, and finding $\boldsymbol{x}$ reduces to the sparse approximation problem

$$\min_{\boldsymbol{u}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}_x\boldsymbol{u}\|_2^2 + \lambda\|\boldsymbol{u}\|_1, \tag{2}$$

where $\lambda$ is a regularization parameter, and $\|\boldsymbol{u}\|_1 = \sum_{i=1}^m |u_i|$ is the $\ell_1$-norm, which promotes sparsity. Sparse approximation was first used for single image SR in [25].

According to recent studies [16], the accuracy of sparse approximation problems can be improved if a signal $\boldsymbol{\omega}$ correlated with the target signal $\boldsymbol{x}$ is available; we refer to $\boldsymbol{\omega}$ as side information (SI). Let $\boldsymbol{\omega} \in \mathbb{R}^d$ have a sparse representation $\boldsymbol{z} \in \mathbb{R}^m$ under a dictionary $\boldsymbol{D}_\omega \in \mathbb{R}^{d\times m}$, $d \leq m$; assume that $\boldsymbol{z}$ is similar to $\boldsymbol{u}$ by means of the $\ell_1$-norm. Then, given the observations $\boldsymbol{y}$, we can obtain $\boldsymbol{u}$ as the solution of the $\ell_1$-$\ell_1$ minimization problem

$$\min_{\boldsymbol{u}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}_x\boldsymbol{u}\|_2^2 + \lambda(\|\boldsymbol{u}\|_1 + \|\boldsymbol{u} - \boldsymbol{z}\|_1). \tag{3}$$

Similarity in terms of the $\ell_1$-norm holds for representations with partially common support and a number of similar nonzero coefficients; we refer to them as *coupled sparse representations.*

[4] Notation: Lower case letters are used for scalars, boldface lower case letters for vectors, boldface upper case letters for matrices and boldface upper case letters in math calligraphy for tensors.

## 3 Deep Unfolding

Deep unfolding was first proposed in [7] where a sparse coding algorithm was unfolded into a neural network form. The resulting model, coined LISTA, is a learned version of the iterative soft thresholding algorithm (ISTA) [4]. Each layer of LISTA computes:

$$\boldsymbol{u}^t = \phi_\gamma\big(\boldsymbol{S}^t\boldsymbol{u}^{t-1} + \boldsymbol{W}\boldsymbol{y}\big), \tag{4}$$

where $\phi_\gamma$ denotes the soft-thresholding operator $\phi_\gamma(\alpha_i) = \text{sign}(\alpha_i)(|\alpha_i| - \gamma)$, $i = 1, \ldots, k$; the parameters $\boldsymbol{S}^t$, $\boldsymbol{W}$, $\gamma$ are learned from data. The authors of [10] integrated LISTA into a neural network design for image SR, obtaining an end-to-end reconstruction architecture that incorporates a sparse prior.

Multimodal image SR via deep unfolding was first addressed in [13], where the authors introduced the assumption that correlated images of multiple modalities can have coupled sparse representations. According to this assumption, given an HR image $\boldsymbol{\omega}$ of a guidance modality, we can compute a sparse representation $\boldsymbol{u}$ from the observations of the target modality $\boldsymbol{y}$ by solving a problem of the form (3). The multimodal network presented in [13] incorporates LeSITA [23], a deep unfolding design that learns coupled sparse representations. Implementing iterations of a side-information-driven thresholding algorithm that solves (3), each layer of LeSITA computes:

$$\boldsymbol{u}^t = \xi_\mu\big(\boldsymbol{S}^t\boldsymbol{u}^{t-1} + \boldsymbol{W}\boldsymbol{y}; \boldsymbol{z}\big), \tag{5}$$

where $\xi_\mu$ is a proximal operator [23] that integrates the information coming from another modality (in the form of $\boldsymbol{z}$) into the reconstruction process.

## 4 A Multimodal Convolutional Deep Unfolding Design for Medical Image Super-Resolution

Due to the large size of images, sparse modelling techniques are typically applied to image patches. Alternatively, Convolutional Sparse Coding (CSC) [28] can be directly applied to the entire image. Let $\boldsymbol{X} \in \mathbb{R}^{n_1 \times n_2}$ be the image of interest. A sparse modelling approach with respect to a convolutional dictionary $\boldsymbol{\mathcal{D}}^X \in \mathbb{R}^{p_1 \times p_2 \times k}$ has the form $\boldsymbol{X} = \sum_{i=1}^{k} \boldsymbol{D}_i^X * \boldsymbol{U}_i$, where $\boldsymbol{D}_i^X \in \mathbb{R}^{p_1 \times p_2}$, $i = 1, ..., k$, are the atoms of $\boldsymbol{\mathcal{D}}^X$, and $\boldsymbol{U}_i \in \mathbb{R}^{n_1 \times n_2}$, $i = 1, ..., k$, are the corresponding sparse feature maps; the symbol $*$ denotes a convolution operation. Then, an observation of $\boldsymbol{X}$ can be written as $\boldsymbol{Y} = \sum_{i=1}^{k} \boldsymbol{A}_i^X * \boldsymbol{U}_i$, with $\boldsymbol{A}_i^X = \boldsymbol{L}\boldsymbol{D}_i^X$.

When, besides the observation of the target image modality, another image modality $\boldsymbol{\Omega}$, correlated with $\boldsymbol{X}$ is available, we can reconstruct $\boldsymbol{X}$ by solving a convolutional form of (3), that is,

$$\min_{\boldsymbol{U}_i} \frac{1}{2}\|\boldsymbol{Y} - \sum_{i=1}^{k} \boldsymbol{A}_i^X * \boldsymbol{U}_i\|_F^2 + \lambda\big(\sum_{i=1}^{k} \|\boldsymbol{U}_i\|_1 + \sum_{i=1}^{k} \|\boldsymbol{U}_i - \boldsymbol{Z}_i\|_1\big), \tag{6}$$

where $\boldsymbol{Z}_i \in \mathbb{R}^{n_1 \times n_2}$, $i = 1, ..., k$, are the sparse feature maps of the modality $\boldsymbol{\Omega}$ with respect to a convolutional dictionary $\boldsymbol{\mathcal{D}}^{\Omega} \in \mathbb{R}^{p_1 \times p_2 \times k}$.

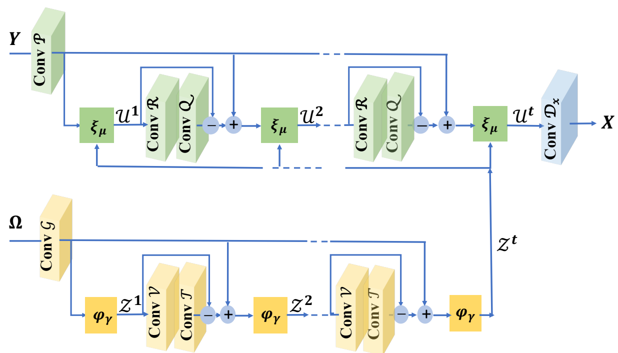



Fig. 1: The proposed multimodal SR model. The lower branch computes the sparse codes $\boldsymbol{\mathcal{Z}}$ of the guidance modality, while the upper (main) branch computes the sparse codes $\boldsymbol{\mathcal{U}}$ of the target modality with the aid of $\boldsymbol{\mathcal{Z}}$. The target HR image $\boldsymbol{X}$ is the result of a convolution operation between $\boldsymbol{\mathcal{U}}$ and a learned dictionary $\boldsymbol{\mathcal{D}}^X$.

The linear properties of convolution allow to write (6) in the form of (3). Then, LeSITA (5) can be used for the computation of the convolutional sparse codes. However, it is computationally more efficient to write (5) in a convolutional form [14], obtaining a learned multimodal convolutional sparse coding (LMCSC) model. LMCSC includes the following stages:

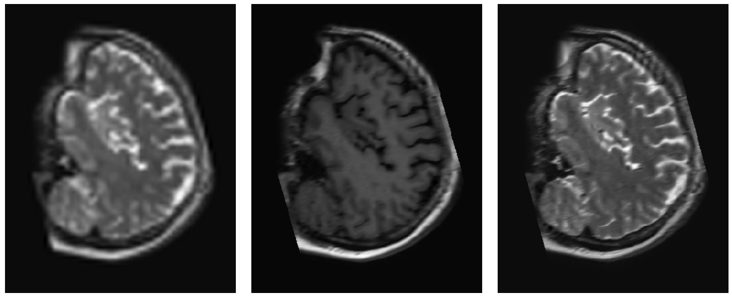
$$\boldsymbol{\mathcal{U}}^t = \xi_\mu(\boldsymbol{\mathcal{U}}^{t-1} - \boldsymbol{\mathcal{Q}} * \boldsymbol{\mathcal{R}} * \boldsymbol{\mathcal{U}}^{t-1} + \boldsymbol{\mathcal{P}} * \boldsymbol{Y}; \boldsymbol{\mathcal{Z}}), \tag{7}$$

with $\xi_\mu$ the proximal operator defined in [23]. The parameters $\boldsymbol{\mathcal{Q}} \in \mathbb{R}^{p_1 \times p_2 \times c \times k}$, $\boldsymbol{\mathcal{R}} \in \mathbb{R}^{p_1 \times p_2 \times k \times c}$, $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{p_1 \times p_2 \times c \times k}$ correspond to learnable convolutional layers; $c$ is the number of channels of the employed images; $\mu > 0$ is also learnable.

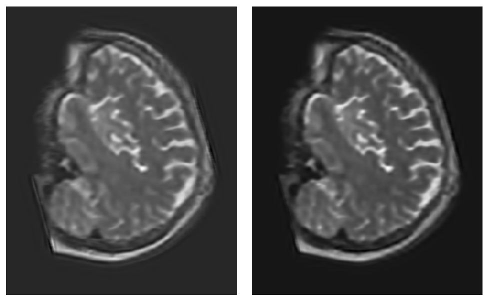
We will apply this model for the super-resolution of LR T2W images with the aid of HR T1W (or FLAIR) images. We assume that images of both modalities have coupled sparse representations under different convolutional dictionaries. We also assume that the LR and HR T2W images can have the same sparse representation under different convolutional dictionaries. Therefore, the reconstruction of the HR T2W image reduces to the computation of the convolutional coefficients of the corresponding LR image. The final HR T2W image can be obtained by a convolutional operation between the sparse coefficients and a convolutional dictionary. The model is depicted in Fig. 1. The training process results in learning the convolutional dictionary $\boldsymbol{\mathcal{D}}^X$ as well as the parameters of the unfolded algorithm (7). The sparse codes $\boldsymbol{\mathcal{Z}}$ of the guidance modality are obtained using a convolutional LISTA model [22], computing at the $t$-th layer:

$$\boldsymbol{\mathcal{Z}}^t = \phi_\gamma(\boldsymbol{\mathcal{Z}}^{t-1} - \boldsymbol{\mathcal{T}} * \boldsymbol{\mathcal{V}} * \boldsymbol{\mathcal{Z}}^{t-1} + \boldsymbol{\mathcal{G}} * \boldsymbol{\Omega}), \tag{8}$$

with $\boldsymbol{\mathcal{T}} \in \mathbb{R}^{p_1 \times p_2 \times c \times k}$, $\boldsymbol{\mathcal{G}} \in \mathbb{R}^{p_1 \times p_2 \times c \times k}$, $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{p_1 \times p_2 \times k \times c}$ and $\gamma$ learnable parameters.



(a) LR T2W (b) HR T1W (c) ground truth



(d) coISTA [5] (e) proposed LMCSC

Fig. 2: A $\times 6$ SR example from the MS-MRI dataset. Reconstruction of an HR T2W image from an LR T2W image (PSNR = 17.51 dB) with the aid of an HR T1W image. Reconstruction PSNR values are 37.53 dB for coISTA and 38.45 dB for LMCSC.

## 5 Experiments

We have used LMCSC with two multimodal MRI databases from the Laboratory of Imaging Technologies[5], namely, a brain MR database [2], which contains simulated data from 20 patients, and an MR Multiple Sclerosis (MS) database [9], which contains real data from 30 patients. Both databases include co-registered T1W, T2W and FLAIR 3D images. From each database, we reserve data from five patients for testing. We create the training dataset by selecting cropped image slices of size $44 \times 44$, and apply data augmentation by flipping and rotating images, obtaining 22K training samples for the MS-MRI dataset and 25K samples for the brain-MRI dataset. We use whole image slices for testing. Each T2W image is blurred with a $3 \times 3$ Gaussian filter and downsampled. We obtain an input LR image of the desired dimensions after bicubic interpolation.

We implement the proposed model with three unfolding stages for each network branch. The size of the learned parameters is set to $7 \times 7 \times 1 \times 85$ for $\boldsymbol{\mathcal{P}}$, $\boldsymbol{\mathcal{Q}}$, $\boldsymbol{\mathcal{T}}$, $\boldsymbol{\mathcal{G}}$, and $7 \times 7 \times 85 \times 1$ for $\boldsymbol{\mathcal{R}}$, $\boldsymbol{\mathcal{V}}$, $\boldsymbol{\mathcal{D}}^X$; a random gaussian distribution

[5] http://lit.fe.uni-lj.si/tools.php?lang=eng

Table 1: Super-resolution of T2W with the aid of T1W or FLAIR images (SI). Results are presented in terms of PSNR (in dB) for two multimodal datasets.

| SI | T1W | | | | FLAIR | | | |
|---|---|---|---|---|---|---|---|---|
| dataset | Brain-MRI | | MS-MRI | | Brain-MRI | | MS-MRI | |
| SR-scale | $\times 4$ | $\times 6$ | $\times 4$ | $\times 6$ | $\times 4$ | $\times 6$ | $\times 4$ | $\times 6$ |
| bicubic interpolation | 28.56 | 25.30 | 17.16 | 17.15 | 28.55 | 25.30 | 17.18 | 17.14 |
| coISTA [5] | 32.57 | 28.34 | 40.54 | 36.56 | **32.39** | 28.10 | 40.28 | 36.66 |
| LMCSC | **34.86** | **31.97** | **40.94** | **37.28** | 32.11 | **28.44** | **40.66** | **36.80** |

Table 2: Performance of LMCSC [in terms of PSNR (in dB)] for varying number of unfolding stages. Experiments are conducted for $\times 4$ SR of T2W with the aid of T1W images on the MS-MRI dataset.

| model configuration | #stages=2 | #stages=3 | #stages=4 |
|---|---|---|---|
| LMCSC | 40.85 | **40.94** | 40.92 |

with standard deviation 0.01 is used for initialization. The initial value of the parameters $\gamma$, $\mu$ is set to 0.1. We set the learning rate equal to 0.0001 and use the Adam optimizer [8] with the mean square error loss function to train the network end-to-end for 100 epochs. As a baseline method, we use a convolutional form of the coISTA model proposed in [5]. We follow the same initialization and training procedure for coISTA. All experiments have been performed on a desktop with AMD Ryzen 5 1600 Six-Core 3.7 GHz CPU, 16GiB RAM, and an NVIDIA GeForce GTX 1070 GPU.

Numerical results, in terms of Peak Signal-to-Noise Ratio (PSNR), presented in Table 1 include $\times 4$ and $\times 6$ SR. Besides LMCSC and coISTA, we also report results for bicubic interpolation. The results show the superior performance of the proposed approach. A visual example presented in Fig. 2, shows that reconstruction with the proposed LMCSC results in a high-contrast and more clear image compared to coISTA [5].

We also report results for different realizations of the proposed model with varying number $t$ of unfolding stages as described by (7). We only vary the number of stages of the main network branch computing the representation of the target modality. The number of ACSC unfoldings is kept fixed, i.e., equal to three. Experiments for this study have been conducted on the MS-MRI dataset for $\times 4$ SR of T2W with the aid of T1W. As can be seen in Table 2, the best performance is achieved with three unfolding stages.

## 6 Conclusion

Interpretable deep learning is a promising approach for the recovery of medical images as it combines trustworthiness and fast inference. Following the principle of deep unfolding, we have presented LMCSC, an interpretable multimodal deep learning architecture that computes coupled convolutional sparse codes.

LMCSC was applied to super-resolve multi-contrast MRI images. The model is designed to address linear inverse problems with side information, therefore, it can be applied for other multimodal recovery tasks such as denoising or compressive sensing reconstruction, while it can also include other medical imaging modalities. We will investigate these applications in our future work.

## References


1. Adler, J., Öktem, O.: Learned primal-dual reconstruction. IEEE Transactions on Medical Imaging **37**(6), 1322–1332 (2018)
2. Aubert-Broche, B., Griffin, M., Pike, G.B., Evans, A.C., Collins, D.L.: Twenty new digital brain phantoms for creation of validation image data bases. IEEE Transactions on Medical Imaging **25**(11), 1410–1416 (2006)
3. Borgerding, M., Schniter, P., Rangan, S.: AMP-inspired deep networks for sparse linear inverse problems. IEEE Transactions on Signal Processing **65**(16), 4293–4308 (2017)
4. Daubechies, I., Defrise, M., Mol, C.D.: An Iterative Thresholding Algorithm for Linear Inverse Problems with a Sparsity Constrain. Communications on Pure and Applied Mathematics **57** (11 2004)
5. Deng, X., Dragotti, P.L.: Deep coupled ISTA network for multi-modal image super-resolution. IEEE Transactions on Image Processing **29**, 1683–1698 (2020)
6. Greenspan, H.: Super-resolution in medical imaging. The computer journal **52**(1), 43–63 (2009)
7. Gregor, K., LeCun, Y.: Learning Fast Approximations of Sparse Coding. In: Proceedings of the 27th International Conference on Machine Learning. pp. 399–406. ICML'10, Omnipress, USA (2010)
8. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
9. Lesjak, Ž., Galimzianova, A., Koren, A., Lukin, M., Pernuš, F., Likar, B., Špiclin, Ž.: A novel public MR image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus. Neuroinformatics **16**(1), 51–63 (2018)
10. Liu, D., Wang, Z., Wen, B., Yang, J., Han, W., Huang, T.S.: Robust Single Image Super-Resolution via Deep Networks With Sparse Prior. IEEE Transactions on Image Processing **25**(7), 3194–3207 (2016)
11. Lucas, A., Iliadis, M., Molina, R., Katsaggelos, A.K.: Using deep neural networks for inverse problems in imaging: beyond analytical methods. IEEE Signal Processing Magazine **35**(1), 20–36 (2018)
12. Mansoor, A., Vongkovit, T., Linguraru, M.G.: Adversarial approach to diagnostic quality volumetric image enhancement. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). pp. 353–356. IEEE (2018)
13. Marivani, I., Tsiligianni, E., Cornelis, B., Deligiannis, N.: Multimodal Image Super-resolution via Deep Unfolding with Side Information. In: European Signal Processing Conference (EUSIPCO) (2019)
14. Marivani, I., Tsiligianni, E., Cornelis, B., Deligiannis, N.: Multimodal deep unfolding for guided image super-resolution. IEEE Transactions on Image Processing **29**, 8443–8456 (2020)
15. Monga, V., Li, Y., Eldar, Y.C.: Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. arXiv preprint arXiv:1912.10557 (2019)

16. Mota, J.F.C., Deligiannis, N., Rodrigues, M.R.D.: Compressed Sensing With Prior Information: Strategies, Geometry, and Bounds. IEEE Trans. Inf. Th. **63**(7), 4472–4496 (2017)
17. Nehme, E., Weiss, L.E., Michaeli, T., Shechtman, Y.: Deep-STORM: super-resolution single-molecule microscopy by deep learning. Optica **5**(4), 458–464 (2018)
18. Papyan, V., Romano, Y., Elad, M.: Convolutional neural networks analyzed via convolutional sparse coding. The Journal of Machine Learning Research **18**(1), 2887–2938 (2017)
19. Qin, C., Schlemper, J., Caballero, J., Price, A.N., Hajnal, J.V., Rueckert, D.: Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Transactions on Medical Imaging **38**(1), 280–290 (2018)
20. Ribes, A., Schmitt, F.: Linear inverse problems in imaging. IEEE Signal Processing Magazine **25**(4), 84–99 (2008)
21. Schlemper, J., Caballero, J., Hajnal, J.V., Price, A.N., Rueckert, D.: A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Transactions on Medical Imaging **37**(2), 491–503 (2017)
22. Sreter, H., Giryes, R.: Learned convolutional sparse coding. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 2191–2195. IEEE (2018)
23. Tsiligianni, E., Deligiannis, N.: Deep Coupled-Representation Learning for Sparse Linear Inverse Problems with Side Information. IEEE Signal Processing Letters (2019)
24. Xiang, L., Chen, Y., Chang, W., Zhan, Y., Lin, W., Wang, Q., Shen, D.: Deep-learning-based multi-modal fusion for fast MR reconstruction. IEEE Transactions on Biomedical Engineering **66**(7), 2105–2114 (2018)
25. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing **19**, 2861–2873 (2010)
26. Yang, Y., Sun, J., Li, H., Xu, Z.: Deep ADMM-Net for compressive sensing MRI. In: Advances in Neural Information Processing Systems (NIPS), pp. 10–18 (2016)
27. Yedder, H.B., Cardoen, B., Hamarneh, G.: Deep learning for biomedical image reconstruction: A survey. Artificial Intelligence Review pp. 1–37 (2020)
28. Zeiler, M.D., Krishnan, D., Taylor, G.W., Fergus, R.: Deconvolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2010)
29. Zeng, K., Zheng, H., Cai, C., Yang, Y., Zhang, K., Chen, Z.: Simultaneous single- and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. Computers in biology and medicine **99**, 133–141 (2018)
30. Zhou, B., Zhou, S.K.: DuDoRNet: learning a dual-domain recurrent network for fast MRI reconstruction with deep T1 prior. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4273–4282 (2020)