Untrained Neural Network Priors for Inverse Imaging Problems: A Survey

Adnan Qayyum, Inaam Ilahi, Fahad Shamshad, Farid Boussaid, Mohammed Bennamoun, and Junaid Qadir, *Senior Member, IEEE*

Abstract—In recent years, advancements in machine learning (ML) techniques, in particular, deep learning (DL) methods have gained a lot of momentum in solving inverse imaging problems, often surpassing the performance provided by hand-crafted approaches. Traditionally, analytical methods have been used to solve inverse imaging problems such as image restoration, inpainting, and superresolution. Unlike analytical methods for which the problem is explicitly defined and the domain knowledge is carefully engineered into the solution, DL models do not benefit from such prior knowledge and instead make use of large datasets to predict an unknown solution to the inverse problem. Recently, a new paradigm of training deep models using a single image, named untrained neural network prior (UNNP) has been proposed to solve a variety of inverse tasks, e.g., restoration and inpainting. Since then, many researchers have proposed various applications and variants of UNNP. In this paper, we present a comprehensive review of such studies and various UNNP applications for different tasks and highlight various open research problems which require further research.

Index Terms—Inverse imaging problems, untrained neural networks priors, deep learning

1 Introduction

Inverse imaging problems (IIPs) have been recently attracting increasing attention from researchers due to their applications in numerous domains including computer vision, medical imaging, remote sensing, and autonomous driving to name a few. IIPs aim to reconstruct an unknown image from the possibly noisy observations. These observations are obtained from the unknown real data by a forward process that is typically ill-posed and with multiple possible solutions for the same IIP. Reconstructing a unique solution that fits these observations ismay be very difficult or impossible without reliable prior information about the data [1]. Thus designing effective priors has been the subject of substantial research in the image processing community as a variety of image reconstruction tasks fit under the umbrella of IIPs including denoising, super-resolution, inpainting, image deblurring, MRI reconstruction, and many more [2]. These priors are essentially regularization techniques.

Conventional algorithms to solve IIPs are based on simple mathematical models that are *hand-crafted* from the domain knowledge. These domain knowledge-based recovery algorithms carry out the inference based on knowledge of the underlying forward model associated with the observed measurements. These algorithms usually do not rely on data to learn their mapping. However, these simple models often suffer from a poor discriminative capability and as a consequence, a large majority of unnatural images also satisfy the constraints specified by the forward model (since the solution to the ill-posed problem is not unique) [3].

Recently, deep learning (DL) has provided major breakthroughs in solving IIPs as compared to *hand-crafted* priors-based approaches that were so far unable to effectively tackle these problems [4]. For instance, recently, generative models have shown their ability to reconstruct a high-resolution image from a fraction of samples in the compressed sensing problem. Such a high compression ratio was not previously possible with *hand-crafted* priors like sparsity. However, these DL-based approaches owe much of their success to the availability of large labeled imaging datasets that are difficult to obtain in many application areas, such as medical imaging. Data labels can be annotations for a classification task and reference images for inverse problems, e.g., paired bad and good quality images for training end-to-end DL models for image enhancement. Therefore, the widespread deployment of such models is constrained for many practical applications due to the lack of expert annotators, time, ethical constraints, and the financial resources required to create sufficiently large reliable labeled data. Also, it is important to note that not all DL methods will require labeled images, e.g., [5]. However, such an approach (in [5]) would still require a large amount of clean images to train the generative models before pre-trained generative models can be used to solve IIPs. In contrast to the approach adopted in [5], untrained neural network priors (UNNPs) can recover a faithful estimate of the clean sample (without having any prior

- *Adnan Qayyum and Inaam Ilahi are with the Information Technology University (ITU) of Punjab, Lahore 54000, Pakistan. E-mail: {adnan.qayyum, inaam.ilahi}@itu.edu.pk.*
- *Fahad Shamshad is with the Information Technology University (ITU) of Punjab, Lahore 54000, Pakistan, and also with the MBZ University of Artificial Intelligence, Abu Dhabi, UAE. E-mail: fahad.shamshad3@gmail.com.*
- *Farid Boussaid and Mohammed Bennamoun are with the University of Western Australia, Perth, WA 6009, Australia. E-mail: {farid.boussaid, mohammed.bennamoun}@uwa.edu.au.*
- *Junaid Qadir is with the Qatar University, Doha 2713, Qatar. E-mail: jqadir@qu.edu.qa.*

Manuscript received 12 March 2021; revised 28 July 2022; accepted 15 August 2022. Date of publication 5 September 2022; date of current version 3 April 2023.
(Corresponding author: Junaid Qadir.)
Recommended for acceptance by Y. Chi.
Digital Object Identifier no. 10.1109/TPAMI.2022.3204527

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

knowledge of the ground truth itself) while using a single corrupted measurement.

To bridge the gap between conventional *hand-crafted* priors (which suffer from a poor discriminative capability) and DL-based approaches (which usually require large scale labeled datasets, i.e., paired datasets for end-to-end training of deep models), UNNPs have recently emerged as a promising line of research. These priors have been shown to outperform the conventional *hand-crafted* priors and offer comparable performance to their DL-based counterparts while integrating the knowledge of the forward model without requiring massive labeled datasets [6]. More specifically, these priors assume that rich image statistics are captured by the structure of the randomly initialized convolutional neural networks (CNNs), with the randomly initialized network weights serving as a parametrization of the restored image. Since their inception, UNNPs have been widely used in numerous IIPs. Despite their remarkable success, the literature related to UNNPs remains cluttered posing significant challenges to the researchers in this field. To fill in this gap, this paper provides a comprehensive survey on the applications of UNNPs to IIPs. We note here that Ulyanov et al. [6] introduced a different term for UNNPs and named it as deep image prior (DIP) and since then this alternative term has been largely used in the literature. However, the terminologies used for referring to UNNPs are expanding beyond just DIP. We provide a summary of different alternative terms used to describe untrained neural networks-based image priors in Table 1.

Specifically, the following are the major contributions of this survey paper.

1) To the best of our knowledge, this is the first survey paper that comprehensively covers applications of UNNPs to IIPs. We present a comprehensive review of different developments made upon UNNP thus covering the most recent and advanced progress in the field.
2) We provide detailed coverage of different applications of UNNPs for different IIPs. We specifically categorized them into two groups, i.e., general and medical IIPs. Our focus on medical applications is due to ethical constraints, privacy concerns, and the scarcity of labeled datasets. This limits the effectiveness of data-driven methods and makes UNNP ideally suited for medical applications.
3) We highlight different open research problems related to UNNPs and provide insights into possible future directions for new researchers in this field.

The rest of paper is organized as follows. Section 2 presents a detailed background of UNNPs and its different variants. Section 3 provides a broad introduction to UNNPs. Section 4 presents different applications of UNNPs for various IIPs and the applications of UNNPs in medical imagning tasks are presented in Section 5. Insights and pitfalls are discussed in Section 6. Section 7 presents various open research problems which require further development. Finally, we conclude the paper in Section 8.

TABLE 1
Summary of Different Alternative Terms Used to Describe Untrained Neural Networks-Based Priors Over the Past Few Years, They All Rely on the Same Underlying Principle

Year	Term (s)	Reference (s)
2018	- Deep Image Prior	[6], [7], [8], [9],
	- Deep Decoder	[10], [11]
	- Deep Prior	[12]
		[13], [14]
2019	- Deep Image Prior	[15], [16]
	- Untrained Neural Network Priors	[17], [18]
	- Untrained Network Priors	[19], [20]
	- Deep Prior	[21], [22]
	- Deep Network Prior	[23]
2020	- Deep Image Prior	[24], [25]
	- Untrained Network Priors	[26]
	- Untrained Neural Networks	[27]
	- Deep Decoder	[28]
	- Deep Prior	[29], [30], [31], [32]
	- Untrained Deep Neural Network	[33], [34]
	- Untrained Neural Network Priors	[35]
2021	- Deep Image Prior	[36], [37]
	- Untrained Neural Networks	[38], [39], [40], [41],
	- Untrained Networks	[42], [43]
	- Untrained Deep Decoder Network	[44]
	- Deep Prior	[45]
	- Untrained Neural Network Priors	[46]
	- Untrained Network Priors	[47], [48], [49], [50]
	- Untrained Deep Neural Networks	[51]
	- Deep Randomized Neural Networks Priors	[52], [53], [54] [55]

2 Inverse Imaging Problems (IIPs): Background and Methods

2.1 IIPs: An Introduction

The reconstruction of an unknown image or multidimensional signal/tensor from observed measurements is regarded as an inverse problem. To be more precise, we consider inverse problems in which an unknown n-pixel image (in vectorized form) $\boldsymbol{x}_0 \in \mathbb{R}^n$ is observed via m noisy measurements $\boldsymbol{y} \in \mathbb{R}^m$. More specifically, we can write

$$\boldsymbol{y} = \mathcal{A}(\boldsymbol{x}_0) + \boldsymbol{\eta}, \tag{1}$$

where $\mathcal{A}$ is the forward measurement operator, i.e., the physical model underlying the measurement process, and $\boldsymbol{\eta}$ is noise perturbation modelled by $\mathcal{N}(\mu, \sigma^2)$. Determining the output $\boldsymbol{y}$ for a given input $\boldsymbol{x}_0$ for the known forward operator $\mathcal{A}$ represents the forward model. Similarly, finding the input $\boldsymbol{x}_0$ for a given output $\boldsymbol{y}$ and the operator $\mathcal{A}$ represents the solution to the inverse problem. Furthermore, as the measurement process is often costly, the underdetermined regime (where $m \ll n$) has gained much attention during the last two decades. Solving these underdetermined systems is challenging as they are often ill-posed and do not possess a unique solution. Some prior knowledge about the true image $\boldsymbol{x}_0$ is usually required to solve these problems efficiently. More specifically, in the presence of additive white Gaussian noise and some prior information about the true image, the maximum a posteriori (MAP) leads to the following formulation for solving inverse problems

$$\arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x})\|_2^2 + \lambda R(\boldsymbol{x}), \tag{2}$$

where $R(\boldsymbol{x})$ is the regularization term (also known as image prior), $\|.\|_2^2$ is the ℓ_2 norm, and λ is the regularization weight (hyperparameter). In Eq. (2), the first term is the data fidelity

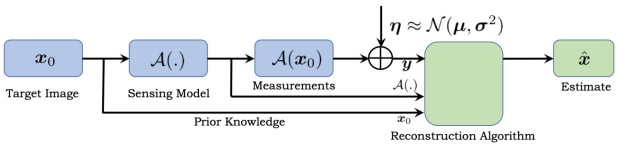

Fig. 1. Pipeline to solve IIPs. In the forward model, the transformation $\mathcal{A}(.)$ (i.e., sensing model) is applied to an input image x_0 to acquire measurements. The inverse problem aims to obtain an estimate of x_0 from the observation via a reconstruction algorithm that leverages the *prior knowledge* about the true (target) image x_0 and $\mathcal{A}(.)$. Normal noise η is added to the measurements which are fed to the reconstruction algorithm.

term and the second term denotes the latent image prior. It has been observed that the success of recent imaging inverse methods mainly stems from the development of effective image priors that stabilize the degradation inversion and direct the outcome towards a more plausible image [4]. Also, the forward model specifies the type of a typical inverse problem at hand, i.e., when the forward operator is linear then the inverse problem will be linear. Otherwise, it will be non-linear. The role of the forward operator in a typical image processing pipeline is shown in Fig. 1. A summary of different forward operators used in different inverse problems is presented in Table 2. IIPs can be of different types, e.g., inpainting, denoising, super-resolution, and compressed sensing as depicted in Fig. 2. It is worth noting the difference between the 'true' and discretized measurement process often used while solving IIPs. Neglecting this difference may lead to the 'inverse crime' [56]. This term was first coined in [57] and has been used in the IIPs literature to describe the use of approximate forward operators in inverse tasks, i.e., when approximately similar forward operators are used in the forward and inverse operations in an inverse modeling task. For example, in the case of magnetic resonance imaging (MRI) reconstruction, the continuous Fourier transform is used in the forward model while the discrete Fourier transform is used in the inverse task. This difference, which is often neglected, can result in reconstruction errors that may affect the interpretation of results. However, the authors argued the non-uniqueness of the solution to the inverse problem is somehow revealed by the notion of this inverse crime, and avoiding this crime can induce the unique solution.

TABLE 2
A Summary of Different Inverse Problems and Associated Forward Operators (Adapted From [4])

Inverse Problem Task	Forward Operator	Description
Denoising	$\boldsymbol{A} = \boldsymbol{I}$	$\boldsymbol{I}$ is the identity operator.
Super-resolution	$\boldsymbol{A} = \boldsymbol{SK}$	$\boldsymbol{S}$ is a sub-sampling operator (identity matrix with missing rows). $\boldsymbol{K}$ is a blurring operator which corresponds to a convolution with a blur kernel.
Inpainting	$\boldsymbol{A} = \boldsymbol{S}$	$\boldsymbol{S}$ is a diagonal matrix where the diagonal elements are set to 1 for sampled pixels.
Phase Retrieval	$\mathcal{A}(\boldsymbol{x}) = \lvert\boldsymbol{Ax}\rvert^2$	$\lvert.\rvert$ denotes the absolute value, the square is taken element-wise, and $\boldsymbol{A}$ is an application-dependent measurement matrix which is often a variant of a discrete Fourier transform (FT) matrix.
Fourier Ptychography	$\mathcal{A}(\boldsymbol{x}) = \lvert\boldsymbol{F}^{-1}\boldsymbol{P}_\ell\boldsymbol{Fx}\rvert^2$	$\boldsymbol{F}$ denotes the 2D FT and $\boldsymbol{P}_\ell$ denotes ℓ the pupil mask that acts as a bandpass filter in the Fourier domain.
Deconvolution	$\mathcal{A}(\boldsymbol{x}) = \boldsymbol{k} * \boldsymbol{x}$	$\boldsymbol{k}$ is a known blur kernel and * denotes the convolution operator. For a known $\boldsymbol{k}$, the problem is called non-blind deconvolution.
Compressed Sensing	$\boldsymbol{A} = \boldsymbol{SF}$ or a random Gaussian matrix	$\boldsymbol{S}$ is a sub-sampling operator (identity matrix with missing rows). $\boldsymbol{F}$ is a discrete FT matrix.
HDR Imaging	$\mathcal{A}(\boldsymbol{x}) = c.\boldsymbol{x}^\gamma$	$\gamma > 1$ and c is a positive scaling constant.
MRI Imaging	$\boldsymbol{A} = \boldsymbol{SFD}$	$\boldsymbol{S}$ is a sub-sampling operator (identity matrix with missing rows), $\boldsymbol{F}$ is the discrete FT matrix, and $\boldsymbol{D}$ is a diagonal matrix representing a spatial domain multiplication with the coil sensitivity map.
Computed Tomography	$\boldsymbol{A} = \boldsymbol{R}$	$\boldsymbol{R}$ is the discrete Radon transform.

2.2 Methods for Solving IIPs

2.2.1 Hand-Crafted Priors

One natural approach to model natural images is to view them as sparse to a particular hand-crafted basis, e.g., wavelets-like basis. This type of method does not require training data and it assumes that the image being modeled is in the linear span of a few elements in a space, which may not provide the best representation. Such image reconstruction methods exploit some prior knowledge about the true image $\boldsymbol{x}$ such as sparsity [58], smoothness [59], total variation [60], non-local self similarity [61], Markov-tree models on wavelet coefficients [62], geometric properties [63], etc. The reconstruction amounts to finding a solution $\hat{\boldsymbol{x}}$ that is a good fit to the observations $\boldsymbol{y}$ given the prior knowledge. These conventional priors often fail to capture the rich structure that many natural signals exhibit due to their limited discriminative capability. As a consequence, a large majority of unnatural signals also satisfy the constraints specified by these hand-crafted priors.

2.2.2 Sparse Coding or Dictionary Learning

In sparse coding or dictionary learning, we use the training data to learn a few images that can be used to recover a particular desired image. The majority of sparse coding and dictionary learning methods are based on l_0-norm or l_1-norm based optimization [66]. Like the hand-crafted basis method, this

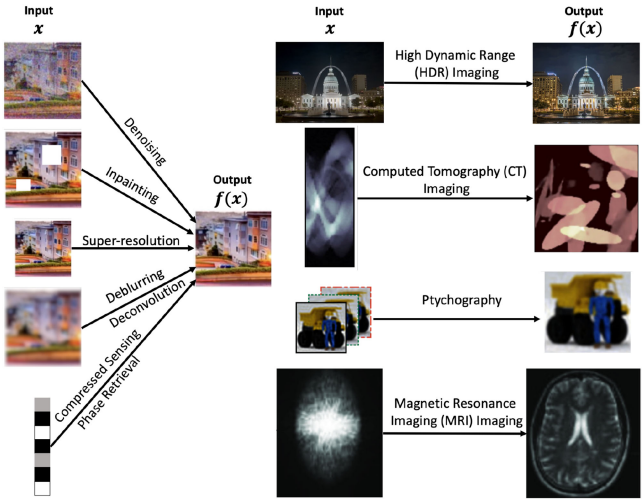

Fig. 2. Examples of inverse imaging problems (IIPs). Individual images adapted from: Left: [18]; Right (top to bottom): first[1]; second [24]; third [64]; fourth [65].

approach is also linear. Sparse coding algorithms can be categorized into four classes, i.e., constrained optimization, proximity algorithm-based optimization, greedy strategy approximation, and homotopy algorithm-based sparse representation, the empirical evaluation of various methods within each category of aforementioned sparse coding algorithms has been performed in [66].

2.2.3 Deep Learning for Solving Inverse Problems

Recently, supervised DL-based approaches have achieved state-of-the-art performance in various IIPs, but at the cost of massive labelled training datasets. These DL-based approaches typically learn an inverse mapping from signal measurements by minimizing the reconstruction loss on a set of training examples. More specifically, these approaches learn a mapping from $\boldsymbol{y}$ to $\boldsymbol{x}$ without any knowledge of the forward operator $\mathcal{A}$ during the training process. Particularly, these approaches aim to minimize the objective function $\|\boldsymbol{x} - \mathcal{F}_\theta(\boldsymbol{y})\|$, where $\mathcal{F}_\theta(.)$ denotes a deep neural network having parameters (weights) denoted as θ. The general principle is that given enough training (labelled) data, one should be able to learn everything one needs to know about $\mathcal{A}$ to successfully estimate $\boldsymbol{y}$. This straightforward approach has shown impressive results on numerous IIPs including super-resolution, blind image deblurring, magnetic resonance imaging, and numerous other IIPs. DL-based approaches can effectively circumvent the limited discriminative capability of the hand-crafted priors by leveraging the power of large datasets. However, DL-based training, popularly known as discriminative learning, makes the network task-specific. As a consequence, we need to retrain the network for various IIPs and different parameter settings of the forward model. This means that even a slight change in the forward acquisition model such as noise level (e.g., for denoising task) or sampling rate (e.g., in compressed sensing) requires costly retraining of these DL models. It is worth noting that apart from the supervised learning based DL approaches, there are unsupervised DL methods that can work with unlabelled data.

1. https://docs.opencv.org/3.4/d2/df0/tutorial_py_hdr.html

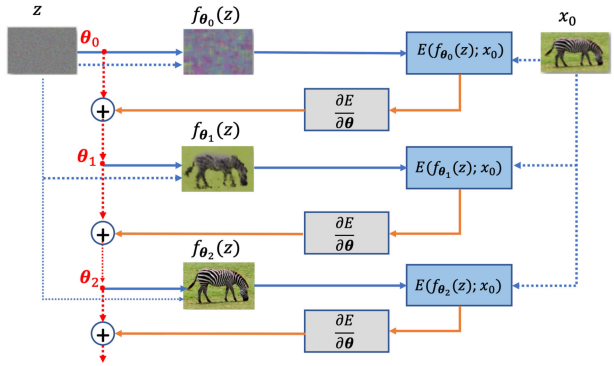

Fig. 3. Denoising input image $\boldsymbol{x}_0$ over different iterations of UNNP, $\boldsymbol{z}$ is the fixed random noise which is given as an input to UNNP (Figure adapted from [6]).

2.2.4 Untrained Neural Networks Priors (UNNPs)

Instead of training DL models on large-scale datasets, untrained DL models are used to model natural images. More specifically, the DL model is given a random noise (i.e., $\boldsymbol{z}$, as shown in Fig. 3) as input and is tasked to model a particular natural image using the random input (i.e., $\boldsymbol{z}$) and its parameters. The network is initialized with random parameters, which are subsequently optimized iteratively to recover the desired image. This approach is very similar to the wavelets basis-based method in which a few wavelets are combined to get a particular image. In both of these methods, there is no learning, i.e., in UNNP we have one random input and random parameters that the neural network maps into a particular image.

3 Untrained Neural Network Priors: An Introduction

The recent work by [6] is an original contribution in the intersection of inverse problems and DL. The authors proposed a new strategy for handling the regularization task in inverse problems. More specifically, the deep model's architecture is used as the regularizer to the inverse problem. The authors showed that the architecture of the DNN is biased to natural images and is capable of capturing low-level image statistics without being explicitly trained using large-scale datasets. Instead, UNNPs can operate on a single (degraded) image. Furthermore, the authors demonstrated that UNNPs have a high impedance to noise and low impedance to the true image. Since the inception of UNNPs (i.e., DIP) by Ulyanov et al. [6], a number of UNNP frameworks have been proposed in the literature. The most prominent variants are deep decoder [12] and double DIP [7]. DIP and double DIP rely on an over-parameterization (i.e., the number of neural network parameters is greater than the number of pixels in the input image) approach while the Deep Decoder adopts an under-parameterization (i.e., using fewer parameters than the number of pixels in the input image) approach. More recently, new versions of UNNPs have been proposed in the literature. For instance, the Deep Matching Prior [67] uses an UNNP to learn priors for semantically similar pairs of input images. Under-

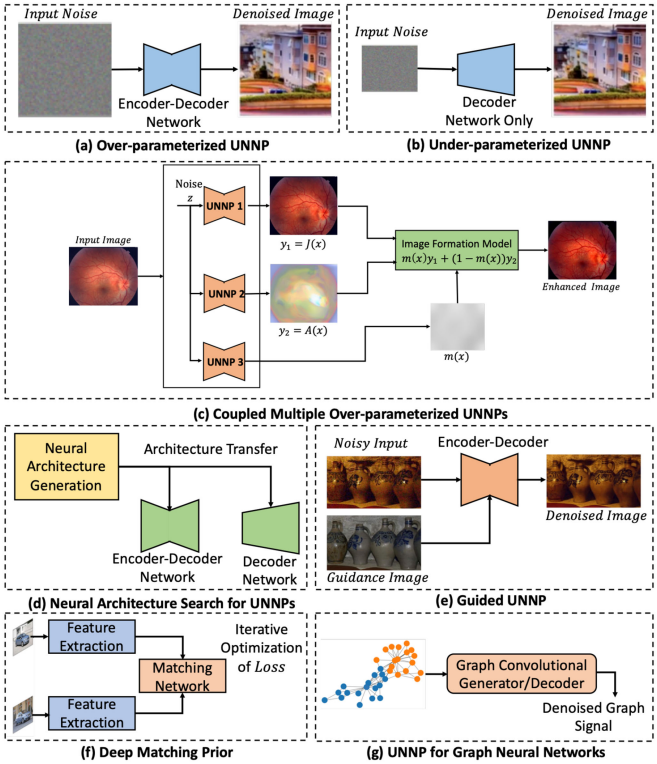

Fig. 4. Different UNNP architectures proposed in the literature. Relevant papers: (a) [6] ; (b) [12] ; (c) [7], [25], [49] ; (d) [37], [68], [69] ; (e) [28], [70] ; (f) [67] ; (g) [39].

parameterized UNNP-inspired untrained graph neural networks were introduced in [39]. The illustration of different UNNP architectures proposed in the literature can be found in Fig. 4. In addition, to these frameworks, numerous efforts have also been made to circumvent the limitations of UNNPs and to improve their performance, e.g., eliminating the need for early stopping in UNNPs (to avoid overfitting), improving performance by introducing regularization in UNNPs, and finding optimal neural network architectures for a specific problem. In the subsequent sections, we will provide an overview of such efforts.

In the original UNNP paper, the authors solved various linear inverse problems including denoising, inpainting, super-resolution, and flash-no-flash reconstruction using an untrained (having randomly initialized parameters) CNN-based generative model. To get the restored version of the degraded image, the parameters were optimized to maximize their likelihood given a task-specific observation model and the specific degraded image. The mathematical expression for solving inverse problems is given by energy minimization of the type

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}}{\operatorname{argmin}} \; E(\boldsymbol{x}; \boldsymbol{x}_0) + R(\boldsymbol{x}), \tag{3}$$

where E is a task-specific function, $\boldsymbol{x}_0$ and $\boldsymbol{x}$ denote the input and the generated images, respectively. $R(\boldsymbol{x})$ denotes the regularizer term. UNNP removes the explicit regularization term $R(\boldsymbol{x})$ by assuming that the unknown image $\boldsymbol{x}$ should be an image generated from the generator network such that

$$\hat{\boldsymbol{\theta}} = \underset{\theta}{\operatorname{argmin}} \; E(f_{\boldsymbol{\theta}}(\boldsymbol{z}); \boldsymbol{x}_0), \quad \hat{\boldsymbol{x}} = f_{\hat{\boldsymbol{\theta}}}(\boldsymbol{z}), \tag{4}$$

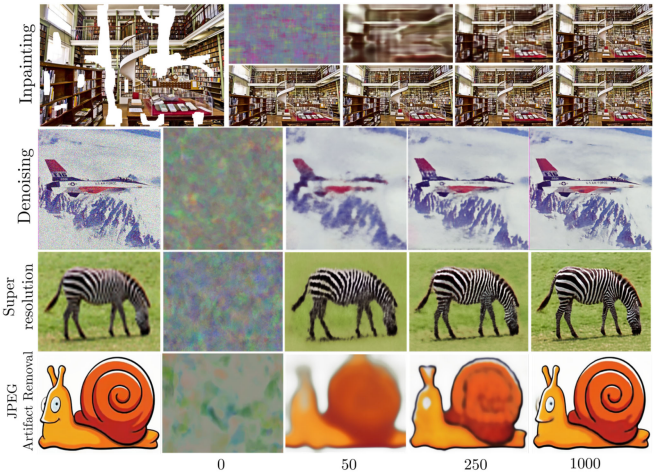

Fig. 5. As the optimization process of UNNP progresses, the recovery of the image is progressively improved while the degradation (e.g., holes, noise) reduces gradually. *Figures reproduced using the publicly available source code.[2]

where, $E(f_{\boldsymbol{\theta}}(\boldsymbol{z}); \boldsymbol{x}_0)$ is a task specific term (e.g., denoising, deblurring, etc.), $\boldsymbol{x}_0$ and $\boldsymbol{x}$ represent the degraded image and the restored image, respectively. $f_{\boldsymbol{\theta}}(\boldsymbol{z})$ is the generator function that maps the random code vector $\boldsymbol{z}$ to an input image $\boldsymbol{x}$, i.e., $\boldsymbol{x} = f_{\boldsymbol{\theta}}(\boldsymbol{z})$. A suitable optimizer such as gradient descent is used for obtaining the minimizer $\boldsymbol{\theta}$ from randomly initialized neural network parameters and the output of the image reconstruction (i.e., restoration) process for a given $\boldsymbol{\theta}^*$ is given by $\hat{\boldsymbol{x}} = f_{\hat{\boldsymbol{\theta}}}(\boldsymbol{z})$. The iterative process used by UNNPs to recover the input image $\boldsymbol{x}_0$ is depicted in Fig. 3. UNNPs start the reconstruction process from iteration 0 using random weights $\boldsymbol{\theta}_0$. They then iteratively update the weights to minimize the task-specific objective (i.e., Eq. (4)). In each iteration, the parameters are mapped to an image $\boldsymbol{x} = f(\boldsymbol{z})$, where $\boldsymbol{z}$ is a fixed code vector. DNNs perform the mapping using its parameters. The task-dependent loss $E(\boldsymbol{x}; \boldsymbol{x}_0)$ is computed using the image $\boldsymbol{x}$. The gradient of the loss with respect to the parameters is then computed and used to update them. Moreover, the quality of images being reconstructed using a UNNP at different iterations (i.e., 0, 50, 250, and 100) for different inverse imaging tasks is shown in Fig. 5.

3.1 Early Stopping in UNNP

In the original UNNP work, the authors demonstrated that fitting the weights (parameters) of an over-parameterized deep convolutional network to a single image, together with strong regularization by the early stopping of the optimization, performs competitively on a variety of image restoration problems. More specifically, the authors relied on early stopping to avoid the inherited overfitting issue of UNNPs. A number of works have been recently proposed to circumvent the issue of early stopping. These works include the deep decoder [12] and a Bayesian alternative of UNNP [16]. Unlike UNNP, the deep decoder is an under-parameterized non-convolutional network that not only represents images well but at the same time cannot fit noise no matter how long it is being optimized (i.e., it circumvents

2. https://github.com/DmitryUlyanov/deep-image-prior

the need for early stopping). Bayesian UNNP presents a novel Bayesian view of the UNNP, which parameterizes a natural image as the output of a CNN with random parameters and random input. More specifically, the authors provide an approach to avoid overfitting by adding suitable priors over the network parameters and then using posterior distributions to quantify the uncertainty. Dittmer et al. [71] demonstrated that the projectional approach to UNNP alleviates the need for early stopping and provides a valid and plausible reconstruction, which is achieved by $f_{\boldsymbol{\theta}}(\boldsymbol{z})$ after the minimization for reconstruction. Similarly, the use of hybrid deep priors for alleviating the problem of overfitting in UNNP is explored in [72]. The authors proposed two algorithms to incorporate an implicit prior (such as a denoising algorithm) or explicit prior (such as total variation (TV)) with UNNP to avoid overfitting. In a recent study [73], authors analyzed UNNP optimization using effective degrees of freedom and then proposed an effective early stopping strategy. Also, they proposed to incorporate stochastic temporal ensemble (STE) to further enhance the efficacy of UNNP for image denoising.

3.2 Adding Further Regularization to UNNP

While UNNP has shown to be quite an effective unsupervised approach, its performance still falls short when compared to state-of-the-art alternatives. A few recent works aim to enhance the performance of the UNNP framework by adding an explicit prior. This enriches the overall regularization effect and provides better-reconstructed images. In [74], the authors proposed the use of TV regularization to improve the basic UNNP approach. The results of image reconstruction for the task of image denoising and deblurring demonstrate that TV regularization provides high-quality results, compared to the basic UNNP approach. More specifically, they aim to solve the following optimization problem

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ E(f_{\boldsymbol{\theta}}(\boldsymbol{z}); \boldsymbol{x}_0) + \lambda \rho_{TV}(f_{\boldsymbol{\theta}}(\boldsymbol{z})), \\ \text{s.t.}\quad \boldsymbol{x}^* &= f_{\boldsymbol{\theta}^*}(\boldsymbol{z}), \end{aligned} \tag{5}$$

where, ρ_{TV} is the total variation regularizer, which is controlled by the multiplying factor λ.

In [75], the authors boost the performance of the UNNP by augmenting it with the power of the trained denoisers. They evaluated their proposed approach on image denoising, super-resolution, and deblurring, showing the clear benefit that regularization by denoisers provides over classical the UNNP. The objective function of their proposed approach is given below.

$$\begin{aligned} \boldsymbol{x}^*, \boldsymbol{\theta}^* &= \underset{\boldsymbol{x}, \boldsymbol{\theta}}{\operatorname{argmin}}\ E(f_\theta(\boldsymbol{z}); \boldsymbol{x}_0) + \frac{\lambda}{2} \boldsymbol{x}^T (\boldsymbol{x} - f(\boldsymbol{x})), \\ \text{s.t.}\quad \boldsymbol{x} &= f_{\boldsymbol{\theta}}(\boldsymbol{z}), \end{aligned} \tag{6}$$

Similarly, in [76], the authors showed that strong prior enforced by the UNNP can be augmented with the information that recurs (i.e., repetitive information) in different patches of a natural image to boost the reconstruction performance. They minimized the following loss function

$$\begin{aligned} \boldsymbol{\theta}^* &= \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ \sum_{n=1}^{N} \|\mathcal{P}_n(\boldsymbol{y}) - f_{\boldsymbol{\theta}}(\boldsymbol{z}))\|, \\ \text{s.t.}\quad \boldsymbol{x} &= f_{\boldsymbol{\theta}}(\boldsymbol{z}), \end{aligned} \tag{7}$$

where operator $\mathcal{P}_n(\cdot)$ extracts the n^{th} patch from the input image and $\hat{\boldsymbol{x}}_n$ is the denoised n^{th} patch. The final estimate of the image is reconstructed by combining all the denoised patches, $\hat{\boldsymbol{x}} = \widetilde{\mathcal{P}}(\hat{\boldsymbol{x}}_1, \hat{\boldsymbol{x}}_2, \ldots, \hat{\boldsymbol{x}}_n)$, where $\widetilde{\mathcal{P}}(\cdot)$ is the function that reconstructs the image back from its patches.

Similarly, the use of Tikhonov functionals rather than deep networks was proposed in [29]. The authors provided empirical evidence and showed that their method is equivalent to regularization techniques.

3.3 Neural Architecture Search for UNNP

The efficacy of UNNP depends upon the architecture of the neural networks being used. Ho et al. [68] proposed a neural architecture search (NAS) technique to boost the performance of the unsupervised learning capabilities of the UNNP framework. They evaluated their proposed technique on different tasks namely image denoising, inpainting, and super-resolution. They demonstrated that the configuration and the meta-parameters of the generator network are automatically optimized by using evolutionary search. These optimized network architectures have been shown to enhance the performance of classic UNNP methods. Instead of using hand-crafted neural networks for UNNP, Chen et al. [69] proposed the use of deep reinforcement learning (DRL) to search for the best possible neural network architecture for a specific problem. Their work is inspired by the NAS algorithms [77], [78], [79], which involve the search of optimal neural networks that give the top performance on large datasets. Their approach differs from NAS-FPN (proposed in [77]) as their target is the recovery of feature maps with higher spatial resolution in the decoder, whereas the aim in [77] is to learn pyramidal feature representations for object detection. Their network design is based upon the standard U-Net architecture. They used DRL to search for an optimal upsampling cell and a pattern of cross-level feature connections by treating the obtained PSNR as the reward of the DRL algorithm. They performed extensive experimentation and demonstrated that the performance of their proposed technique is better than the state-of-the-art learning-free techniques and comparable to the state-of-the-art learning-based methods.

An alternative architecture similar to the UNNP network named guided deep decoder (GDD) for unsupervised image fusion was proposed in [28]. The authors demonstrated that GDD provides state-of-the-art performance in solving a variety of image fusion tasks. Similarly, Cheng et al. proposed to explore the network architecture search by proposing a Bayesian perspective of UNNPs [16]. On the other hand, Uezato et al. [28] argued that the UNNP network architecture [6] does not fully exploit the semantic features of the guidance image and therefore, it remains unclear how the network architecture is conditioned on the features of the guidance image. These uncertainties limit UNNP's ability to achieve state-of-the-art performance in various image fusion problems. To interpret the UNNP framework, manifold modeling in embedded space using a

novel denoising autoencoder combined with multi-way delay embedding transform was proposed in [80]. Chakrabarty et al. [81] proposed a method to analyze trajectories generated by a UNNP and demonstrated two key observations, i.e., the convolution layers in the encoder-decoder network decouple the frequency components of an image while learning at varying rates, and the model starts by fitting the lower frequencies. As a result, enforcing the model to stop early is similar to a low pass filter. In a recent study [48], authors investigated UNNP from the perspective of neural architecture, task-specific model (e.g., compressed sensing, inpainting, and denoising), and the input data type. They also proposed two methods for the selection of optimal hyperparameters for UNNP. Furthermore, they demonstrated that the choice of optimal hyperparameters significantly varies with the problem at hand, leading to degraded performance when used for the wrong task.

3.4 Improving UNNP's Activation Function

In the literature, a number of techniques have been proposed to improve the overall performance of the UNNP network. For instance, Segawa et al. [82] presented a new activation function named RSwish and evaluated its performance versus *LeakyReLU* in UNNP for the task of super-resolution. A comparative analysis of different activations functions, i.e., rectified linear unit (ReLU), leaky rectified linear unit (LeakyReLU), and the randomized leaky rectified linear unit (RReLU) is presented in [83]. The authors considered the tasks of super-resolution, denoising, and inpainting and found that RReLU performs best for the task of denoising and inpainting, whereas ReLU performs better for the task of super-resolution. Metzler et al. [84] evaluated the performance of the UNNP framework for the denoising task using their proposed Stein's Unbiased Risk Estimator (SURE) loss instead of using ℓ_2 loss.

3.5 Theoretical Development

Since UNNP inception, substantial effort has been devoted to building an understanding of why convolutional generators are biased towards natural images and why the untrained network is capable of denoising noisy images without being explicitly trained. In this regard, Heckel et al. [85] have investigated how a specific network architecture enforces an over-parameterized UNNP to fit a single degraded input. In particular, they empirically proved that a convolutional generative network (with fixed convolutional filters) fits natural images faster than pure noise when optimized via gradient descent. In [86], the same authors provided theoretical evidence and empirical proof for a self-regularizing intriguing property of UNNP in recovering a clean image from minimal measurements of the corrupted image without any regularization. Furthermore, they identified that an over-parameterized UNNP is governed by the spectral properties of their Jacobian mapping that contains singular vectors, which can be approximated by orthonormal trigonometric basis functions.

3.6 Coupling Multiple UNNPs

The concept of coupling multiple UNNPs is based on the fact that various computer vision tasks aim to decompose an image into its components. More specifically, the authors in [7] proposed an unsupervised Deep framework that decompose a single image into its layers, such that the distribution of "image elements" within each layer is simple. More specifically, they demonstrated that coupling multiple UNNPs provides a powerful tool to decompose images into their basic components, for a wide variety of applications including image-dehazing, Fg/Bg segmentation, watermark-removal, transparency separation in images and video. These capabilities are achieved in an unsupervised way, with no training examples other than the input image/video itself. Different applications of coupling multiple UNNP networks are described as follows.

3.6.1 Image Decomposition

In [7], the authors proposed a double-UNNP (a general-purpose unsupervised DL framework to decompose a complex input image into simpler layers) and showed that coupling multiple UNNP networks can be used to decompose images into basic layers by exploiting the fact that the internal statistics of a mixture of layers is more complex than the statistics of its components. They evaluated the capability of their proposed method for different image decomposition tasks, including image-dehazing, foreground-background segmentation, watermark-removal, and transparency separation in images and videos. The distribution of small patches within each layer is simpler than the joint distribution of the mixed image along with having weak similarity among the patches of the two layers. The self-similarity of patches within each layer of the image has been exploited by the fact that a single UNNP network shares its filter weights across the entire image due to being fully convolutional. Furthermore, they defined the following criteria for the decomposition to be reasonable as there could be an infinite number of possible decompositions of an image: *(i)* the recombination of the recovered layers must lead to the original image; *(ii)* each layer must show strong self-similarity of patches; *(iii)* each layer must be mutually independent (as possible). These criteria are enforced through a reconstruction loss employing separate UNNPs for each layer, and an exclusion loss between the outputs of the different UNNPs for minimizing their correlation.

In [87], authors employed three UNNPs to model three different image components, i.e., background, fence, and fence mask, in the fence removal problem. The initial mask was estimated using a recurrent network, which was then further refined using an UNNP network that uses a Laplacian smoothness loss function. The proposed approach can generate a visually plausible background image while removing the fence. However, no comparison was performed with state-of-the-art defencing methods. The proposed UNNP-based approach was only compared with existing image inpainting methods, which showed superior performance in fence removal problems. However, this comparison was not fair as inpainting and defencing are distinct problems. Moreover, the proposed approach requires data-driven training of the mask generation network. It also significantly increases the execution time of the proposed method as the training and inference time of the mask generation network is additive to the optimization

time of the three UNNP networks. Tian et al. [88] proposed to use multiple UNNPs for the decomposition of the text layer and noisy background. The enhanced images were then used for CAPTCHA recognition using a representation learning and supervised learning strategy. The UNNPs-based approach provided improved performance in recovering the basic layers of CAPTCHA images as compared to existing baselines. However, UNNPs still exhibited limitations when decomposing images in which the background and character layer possess strong similarities.

3.6.2 Blind Image Deblurring

Blind deconvolution remains challenging in many real-world applications. The use of fixed and hand-crafted priors in traditional maximum a posteriori (MAP) based methods lead to insufficient characterization of clean images and blur kernels. On the other hand, DL-based motion deblurring networks learn from massive datasets but are limited to handling only simpler types of kernels. Recently, in [21], the authors extended the idea of multiple UNNP for blind image deblurring by employing two UNNPs, one to produce the clean image and one for the blur kernel. They used an asymmetric autoencoder with skip connections and an FCN to respectively capture the deep priors of latent clean image and blur kernel, with the SoftMax nonlinearity applied to the output of the FCN to meet the nonnegative and equality constraints of the blur kernel. Moreover, a joint optimization algorithm is proposed to solve the unconstrained neural blind deconvolution model. The optimization process is a kind of zero-shot self-supervised learning, where the generative networks are trained using only a blurry test image without ground-truth clean image.

3.6.3 Reflection Separation

Kim et al. [89] argued that the standard UNNP framework may not always produce good results for the task of reflection separation and removal from natural images. This is because it is only able to capture the low-level image statistics. As a remedy, they proposed the Perceptual UNNP, which can contain high-level semantic information by embedding feature maps extracted from a pre-trained image classification network. Two perceptual UNNPs with cross-feedback were jointly optimized for reflection separation tasks. Furthermore, they argued that the original multiple-UNNP framework ([7]) will only work if the two different images are not correlated which is not the case in natural images. Chandramouli et al. [90] also proposed a UNNP-based method for reflection separation for face images. Unlike [6], they used the output of the previous iteration as the input to the UNNP, and this cross-feedback enhanced the ability of exclusion for both perpetual UNNPs. Through experiments, they showed that the use of the cross-feedback loss not only increases the convergence speed but also increases the robustness of the model. Similarly, Lei et al. [91] argued that the double-UNNP technique fails when used for reflection separation from a given image. The rationale behind this assumption is that the input image is composed of two images with spatial-invariant coefficients. This would not be true in the case of natural images. To overcome this issue, they proposed a perceptual normalized cross-correlation (PNCC) loss to minimize the correlation between the estimated reflection and the transmission at different feature levels.

3.6.4 N-Layer Decomposition

Lu et al. [92] argued that the Double-UNNP approach (proposed in [7]) has limited control over their output layers and entirely relies on the CNN properties to produce a meaningful decomposition. Furthermore, they proposed a method that gives control over the decomposition which is required for re-timing people in videos. However, their method works for only two-layer decomposition and requires N-times the learnable features for an N-layer decomposition. On the other hand, their proposed approach can produce an arbitrary number of features with the same number of parameters. As their technique is not able to effectively handle dynamic backgrounds. They proposed to add an extra layer for dynamic backgrounds in addition to the layers for the foreground objects.

4 Untrained Neural Network Priors: Applications

A summary of different applications of UNNP to different IIPs is presented in Table 3 and are described below.

4.1 3D Shape Reconstruction

3D shape reconstruction is a challenging problem because the search space of 3D shapes is very large, as it aims at reconstructing 3D representations from multiple 2D scenes. Gadelha et al. [22] argued that this search over natural shapes can be replaced by a search over the parameters of the neural network. They proposed to use UNNP for neural parameters search for 3D shape reconstruction. Furthermore, they proposed differentiable projection operators for the reconstruction of shapes from noisy and incomplete projections. These operators when combined with UNNP, generate deep shape priors allowing efficient inference through gradient descent without requiring task-specific training. They evaluated their proposed method on a variety of reconstruction problems such as tomographic reconstruction, visual hull reconstruction, and 3D shape reconstruction. They employed 3D convolutions instead of 2D in UNNP's architecture for 3D shape reconstruction. They showed that their proposed differentiable technique is quite faster than the Bayesian inference method using Markov Chain Monte Carlo (MCMC) techniques [120]. Similarly, the UNNP has been used for 3D mesoscopic imaging using a non-fixed phone camera, where the UNNP was used for the unsupervised restoration of height maps [121]. The authors in [122] demonstrated that UNNPs can be used to solve non-convex optimization problems arising from terahertz (THz) imaging. They incorporated a 3D model-based autoencoder into the UNNP framework, which was used for parameter estimation of the THz model for real and synthetic data under low signal-to-noise ratio (SNR) and shot noise conditions.

TABLE 3
Summary of Various Applications of Untrained Neural Network Priors (UNNPs) for IIPs

Application	Ref	Methodology	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
3D Shape Reconstruction	[22]	Integrated differential operators with UNNP	- ModelNet40 - Intersection over Union (IoU)	UNNP outperformed several handcrafted and procedural priors for image and volumetric reconstruction.	Volumetric representations for shapes incur high memory and storage requirements and significantly increase overall execution time.
Image Enhancement	[93]	Combined UNNP with TV regularization	- San Francisco, Red Rock, & MNIST - Normalized MSE & SSIM	UNNP outperformed traditional hand-crafted methods.	Due to the underparameterized nature of UNNP, sparsity priors were more beneficial at higher sampling rates.
	[94]	Proposed an UNNP variant by combining under-parameterized UNNP with TV and weighted TV regularization.	- Self-Prepared Dataset - PSNR & SSIM	Modified UNNP outperformed standard UNNP for denoising real, synthetic, natural, & medical images.	The proposed approach requires the selection of good hyperparameters. Also, it becomes vulnerable to overfitting at higher iterations.
Illumination Normalization	[10]	Used illumination regression filter with accelerated proximal gradient algorithm along with UNNP framework.	- Extended-YaleB, CAS-PEAL, & Multi-PIE - PSNR, RMSE & SSIM	UNNP outperformed conventional hand-craft priors and provided comparable performance with learning-based methods.	Only considered lighting variations and not variations in pose and facial expressions.
Image Denoising	[95]	Integrated MobileNet-based blind image quality assessment network with UNNP to prevent overfitting.	- PolyU and Nam - PSNR	Outperformed hand-crafted approaches with a good noise suppression capability. Also, achieved 36% reduction in iterations compared to standard UNNP.	Requires separate training of blind image quality assessment network, which require labelled data and can be time-consuming.
	[96]	Proposed Variational UNNP	- Kodak Image Dataset, and McMaster Dataset - CPSNR, SSIM & FSIM	Superior quantitative performance of variational UNNP as compared to standard UNNP and learning-based DL methods.	High execution time as compared to standard UNNP as it uses a training step for each incoming color filter array image.
Image Deblurring	[97]	Used L_0 regularization with UNNP	- Self-Defined Dataset - PSNR & VIF	Outperformed hand-crafted methods quantitatively and qualitatively on simulated and real data.	Proposed UNNP framework only works with Poisson noise and cannot work for non-uniform noise.
	[98]	Incorporated back-propagation loss into UNNP to improve the deblurring performance.	- Set14 Dataset - PSNR	Provided superior quantitative results than standard UNNP with fast convergence.	Performance of conventional back projection method is superior when blur kernel size is small.
Image Decomposition	[7]	Used multiple coupled UNNP networks	- Self-Defined Dataset - PSNR	Outperformed hand-crafted priors and provided comparable performance to learning-based methods.	Highly dependent on the choice of hyperparameters, especially for the task of segmentation.

TABLE 3
(*Continued*)

Application	Ref	Methodology	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
HSI Image Super-resolution	[99]	Used UNNP as the prior of the latent HR HSI	- CAVE and Washington DC datasets - RMSE, PSNR, SAM & SSIM	Provided superior performance as compared to hand-crafted approaches on two benchmark hyperspectral imaging datasets.	High computational cost due to the volumetric nature of data. Also, as no early stopping was used, therefore, performance drops at higher iterations.
	[100]	Integrated UNNP with super resolution convolutional neural network (SRCNN)	- Self-chosen hyperspectral images - Overall Accuracy & Kappa index	UNNP-based super resolution mapping outperformed three state-of-the-art methods.	SRM and SR on a larger scale could have a sophisticated impact on transfer learning performance and applicability.
	[101]	Trained UNNP with both external dataset and internal information of the spatial-spectral restricted input coded image.	- CAVE, the Harvard, and ICVL datasets - PSNR, SSIM & SAM	Better generalization ability, quantitative, and qualitative results on synthetic and real data compared to hand-crafted and learning-based approaches.	External learning module requires large dataset for training.
	[102]	Integrated image registration into HSI super-resolution for joint unsupervised learning	- CAVE, and Harvard datasets - RMSE, PSNR, SAM & SSIM	Effectively exploits the spatial-spectral structures and outperforms conventional approaches in terms of four quantitative metrics.	Despite improved performance, it incurs extra computational cost due to addition of spatial transformer network.
Depth Map Super Resolution	[103]	Integrated a visual appearance based loss function with UNNP for 3D image reconstruction.	- SimGeo, ICLNUIM, Middlebury, SUN RGBD, and ToF-Mark - DSSIM & LPIPS	UNNP with proposed loss function outperformed two learning-free and four learning-based state-of-the-art image super resolution (SR) methods.	Early stopping was not employed, hence UNNP optimization becomes prone to overfitting at higher iterations.
	[9]	Relaxed the constraints of the original UNNP, enforcing it to capture compact prior knowledge for a given task.	- Set5 dataset - PSNR	Provided slightly less performance as compared to a state-of-the-art DL-based method in noiseless SR and outperformed standard UNNP in noisy SR.	UNNP struggles to reconstruct fine details such as eyelashes.
Image Inpainting	[104]	Integrated human guidance with UNNP via a feedback mechanism.	- Mogao Grottoes - LMSE & DSSIM	Provided superior performance as compared to four learning-based methods and hand-crafted method.	Irrelevant inputs from human can adverserly affect the performance. Also, human involvement increases, which incurs more time.
	[105]	Integrated UNNP with a depth reconstruction loss and a view-constrained photo-consistency loss	- Tanks and Temples (TnT), KITTI stereo, & NYU depth V2 - Precision, Recall & F-score	UNNP-based method outperformed quantitatively and qualitatively different baseline approaches, including hand-crafted and learning-based methods.	Limitations in recovering small image details. Also, 3D point cloud fusion incurs increased computational complexity and longer execution time, thus limiting real-time applicability.

TABLE 3
(*Continued*)

Application	Ref	Methodology	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
Video Inpainting	[106]	Used flow prior and a consistency loss with UNNP	- Vidoes from DAVIS dataset - FID, PSNR & SSIM	Proposed UNNP-variant outperformed standard UNNP and a baseline learning-free method.	Cannot fill large holes and cannot handle motion artifacts. Also, each video is processed for several hours.
Video Stabilization	[107]	Shift in frames was modelled as the dense optical flow field of consecutive frames	- Videos from different sources - Accumulated optical flow, cropping ratio, global distortion, frequency domain stability & the smoothness of frame motion.	UNNP-based approach provided an improved quantitative performance as compared to baseline methods. However, in some examples, there was a negligible difference in qualitative performance with slightly improved quantitative performance.	UNNP does not converged for video segments. Also, processing of temporal information, incurs much larger time as compared to processing 2D images.
Temporal Consistency	[108]	Used UNNP for achieving blind video temporal consistency	- DAVIS dataset - Temporal consistency & data fidelity	Provided a comparable quantitative performance as compared to one learning-based and learning-free method.	UNNP struggles with recovering flickering artifacts in videos and require higher iterations as that of normal videos. Also, video processing, significantly alleviates execution time and memory utilization.
Video Restoration	[109]	Jointly solved speckle noise reduction and image inpainting using an UNNP-based approach.	- Self defined dataset made from noisy old and new movies - PSNR	Proposed approach outperformed the hand-crafted, learning-based method, and as well as standard UNNP.	UNNP struggles in recovering minute spatio-temporal inconsistencies in videos. Also, video processing using UNNP, significantly alleviates execution time and memory utilization
Super-Pixel Segmentation	[110]	Designed specialized objective function for UNNP that was used for super-pixel segmentation.	- BSDS500 dataset - Achievable segmentation accuracy (ASA) & boundary recall (BR)	UNNP with proposed objective function outperformed three hand-crafted approaches for the task of super-pixel segmentation.	Generation of superpixels are highly dependent on initial parameters. Also, UNNP's reconstruction loss may induce independent pixels to same superpixel.
Pansharpening	[32]	Used UNNP and dual-attention residual network (DARN) for HSI pansharpening. UNNP was employed for super-resolution task and DARN was trained in a data-driven strategy.	- CAVE, Pavia Center, Botswana, and Los Angeles datasets - SAM, RMSE, ERGAS & PSNR	Proposed UNNP-based approach provided superior performance as compared with existing hand-crafted and learning-based methods.	Training time of DARN model increases with the increasing residual blocks.
Phase Retrieval	[17]	Used UNNP for compressive phase retrieval and optimized it using gradient descent and projected gradient descent.	- MNIST & CelebA - nMSE	UNNP-based approach outperformed baseline hand-crafted ompressive phase retrieval methods.	Initialization of parameters is problem specific and impacts performance.

TABLE 3
(*Continued*)

Application	Ref	Methodology	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
	[111]	Used UNNP for unwrapping the phase in 2D quantitative phase imaging.	- Self-defined Dataset - Regressed SNR (RSNR)	Outperformed learning-based DL models in specific analyses and underperformed in some cases.	Evaluation performed on simulated data.
	[26]	Used UNNP for Fourier ptychography	- Self-defined Dataset - PSNR & SSIM	Significantly outperforms sparsity-based approach and trained generative models in terms of quantitative metrics and visual quality at low sampling rate.	Performance evaluation and results are only reported for simulated datasets.
	[112]	Used autoencoder in UNNP framewrok for the phase retrieval in Fourier ptychography.	- INRIA Holidays data-set - PSNR & SSIM	Proposed UNNP variant outperformed baseline DL-based methods and as well as standard UNNP.	UNNP cannot handle low overlap cases for which a supervised GAN was used. Also, it fails in efficiently handling images with bad illumination.
Image Deconvolution	[113]	Inspired by learning free deconvolution methods, modified classical UNNP's objective to deconvolution energy function.	- Self-defined Dataset - MSE, PSNR & KL divergence	Results only compared on six standard benchmark images and no comparison was performed with existing baselines.	Proposed UNNP variant was not able to effectively handle Gaussian degraded deconvolution.
	[31]	Used an autoencoder in UNNP framework for denoising and a FCN network to model blur kernel with TV regularization.	-Public data - PSNR & SSIM	Proposed hybrid UNNP approach outperformed nine baseline methods and standard UNNP.	As proposed hybrid method involves optimization of two generator networks, it incurs larger execution time.
Adversarial Defense	[114]	Used UNNP to clean adversarial noise from adversarial perturbations to withstand adversarial attacks on ML/DL classifiers.	- CIFAR-10 - Classification Accuracy	UNNP provided improved performance in cleaning adversarial noise generated by three well-known methods, i.e., FGSM, BIM, & LLCI as compared to a randomization-based baseline.	All methods aiming at using UNNP for denoising of adversarial noise suffer from one major limitation, i.e., UNNP is only able to recover the clean image in early iterations, and after that it gets overfitted to adversarial noise. Therefore, an optimal early stopping criteria is required to stop optimization. Also, when UNNP optimization is stopped in early iterations, the quality of the reconstructed clean image will be poor (as optimization is usually required to be stopped in the first few hundred iterations, e.g., 300-500).
	[115]	Used UNNP for generation of adversarially robust features and proposed adaptive early stopping strategy.	- CIFAR-10 & ImageNet (subset) - Classification Accuracy	UNNP outperformed five state-of-the-art methods in denoising adversarial perturbations generated using FGSM, PGD JSMA, Momentum, & STA (with varying noise level).	
	[116]	Used UNNP for cleaning adversarial perturbations	- Self-defined Dataset - Classification Accuracy	Only FGSM-based adversarial perturbations were considered for denoising.	
	[117]	Used UNNP for cleaning adversarial perturbations	- CelebA Dataset - Accuracy, precision & Recall	Successfully denoised adversarial perturbations generated using FGSM and C&W.	

TABLE 3
(*Continued*)

Application	Ref	Methodology	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
Crafting Adversarial Perturbations	[118]	Used UNNP for crafting adversarial perturbations	-ImageNet - Misclassification Rate	UNNP can be used to reconstruct such adversarial perturbations that are robust to affine deformations.	Generates different perturbations for each run. However, it does not affect performance significantly and takes longer time.
	[13]	Explored the vulnerability of super-resolution UNNP to adversarial attacks.	- ImageNet & MSCOCO - Acc, BLEU, ROUGE & CIDEr	UNNP-based reconstructed images were successfully able to evade three DL models.	UNNP cannot effectively exploit minute image details in low-resolution image.
Counterfactual Explanations	[119]	Integrated an auxillary loss estimator trained with predictor that guides UNNP in performing image reconstruction.	- ISIC 2018 lesion data-set - No metric used	Performs better in synthesizing meaningful counterfactuals as compared to standard UNNP.	The quantitative evaluation was not performed. Also, performance degrades at higher iterations of UNNP due to overfitting issue.

4.2 Compressed Sensing

Compressed sensing aims to reconstruct an unknown signal/image of dimension n from a small set of its m linear and noisy measurements, where $m \ll n$. UNNPs have been successfully applied to compressed sensing tasks. For instance, Van et al. [123] proposed to use UNNPs for the compressed sensing-based reconstruction of images from three databases, i.e., chest X-ray images, MNIST, and START. In their proposed method, they used the DCGAN architecture in the UNNP framework and integrated learned regularization to allow processing of non-linear measurements, which decreases reconstruction error as well. The proposed UNNP framework outperformed different baseline methods (i.e., TVAL3 and Lasso-DCT) except BM3D-AMP for higher m. However, despite using the unsupervised UNNP framework, they still relied on a small amount of training data for learned regularization. Similarly, Ren et al. [124] combined a UNNP (with a DCGAN's generator architecture) and compressed sensing for the recovery of 1D signals generated from soils' data, whose function is to measure the quality of soil. In [125], the authors proposed to train a deep unsupervised model by partitioning the images into several sub-bands. A UNNP-based alternative algorithm was used to perform reconstruction by imposing constraints on the parameters of the DNN. Furthermore, to boost the performance of UNNPs, they injected image-based priors, which were extracted from the training data. The proposed method was evaluated on different datasets including facial images, medical and multi-band astronomical images. In a recent study [126], the robustness aspects of three different MRI reconstruction methods namely, end-to-end trained DL models, UNNPs (DIP and Deep Decoder), and traditional compressed sensing methods were investigated against three vulnerabilities, i.e., adversarial perturbations, distribution shifts, and recovering details. The authors found that all three MRI reconstruction methods are vulnerable to adversarial perturbations and performance against distribution shift was linearly correlated with the in-distribution performance.

Meng et al. [50] proposed a plug-and-play framework that leverages UNNP for spectral snapshot compressed sensing along with other traditional priors. During the optimization process, conventional and UNNP-based priors complement each other. Furthermore, the authors proposed an alternative optimization strategy to jointly solve the reconstruction and optimization of network parameters optimization. Similarly, Sun et al. [36] presented a plug-and-play prior and augmented Lagrangian formulation of the problem (at hand) within the UNNP framework. In [127], a learning free generative modelling framework (named APGen) combining UNNP and a sparse regularizer, was proposed. The authors demonstrated the effectiveness of APGen for the compressed sensing of neural action potentials. The results show that APGen outperforms existing data-driven and model-based in terms of time efficiency, reconstruction performance, and robustness to action potential misalignment overlap. Similarly, in [45], the authors showed that under-parameterized UNNPs integrated with simple regularizers (such as energy minimization or l_2 regularization) can be used for the reconstruction of Gabor holograms. In their proposed UNNP framework, the input random noise was kept constant. In a similar study [54], an under-parameterized UNNP was used for the reconstruction of dual-wavelength in-line holographic images. A complete task-specific physical model (representing dual-wavelength in-line holography) was incorporated in the UNNP framework to help the network effectively eliminate the effect of amplified noise. Monakhova et al. investigated both under-parameterized and over-parameterized UNNPs for compressed sensing based lensless 2D imaging [44]. They showed that over-parameterized UNNPs provide superior performance compared to under-parameterized networks. As the average reconstruction time was 1.5 hours, early stopping was used to limit computational time.

4.3 Image Restoration

4.3.1 Image Enhancement

The enhancement of images taken in poor lighting conditions is a well-known problem. This enhancement can be performed by extending (post-acquisition) the dynamic range of the captured images. Jagatap et al. [93] proposed the use of UNNP for high dynamic range (HDR) image reconstruction without training data. To enhance the performance of UNNP, they used TV regularization for the reconstruction of low-light images and they used Deep Decoder architecture, which is a variant of UNNP. Their results demonstrate a significant improvement over previous traditional dynamic range enhancement techniques. Similarly, to improve the performance of UNNP, Cascarano et al. [94] proposed the use of a total variation and weighted total regularizer in the UNNP framework to promote the gradient-sparsity of the solution. The minimization problem is then solved using the alternating direction method of multipliers (ADMM) optimization framework. Furthermore, the authors demonstrated that the proposed technique (named ADMM-DIPPTV) outperforms two state-of-the-art techniques (i.e., DIP [6] and DIP-TV [74]) on several image restoration tasks in terms of PSNR and SSIM values. Furthermore, they claimed that the use of this ADMM splitter guarantees the algorithm's stability. This was demonstrated, without a loss of focus on small details, that their proposed technique increases sharpness over the edges of the objects in an image.

Mastan et al. [128] proposed a learning-free framework to investigate the relationship between UNNP (architecture) construction and image restoration. Their proposed framework is based on multi-level extensions of encoder-decoder networks (MED) and allows various network structures, e.g., by modifying the skip connections and the network depth, composing encoder-decoder sub-networks, and cascading the network input into intermediate layers. Furthermore, they illustrated how the image restoration tasks are affected by the construction of neural network architecture in UNNP. Instead of using the randomly initialized deep network as a hand-crafted prior for image restoration, they used their proposed learning-free approach based on hand-crafted structures for image restoration. They argued that the non-usage of training samples to learn the image prior in UNNP [6] causes it to miss local level features in the output image. Nonetheless, it has been shown to generate better images than bicubic sampling.

4.3.2 Illumination Normalization for Face Images

Illumination normalization is a major factor that impacts face recognition. To solve this problem, Han et al. [10] proposed an illumination normalization method to generate photorealistic face textures while preserving the face identity. In their proposed technique, they combine an illumination regression filter with an accelerated proximal gradient algorithm. Both of these components remove several illumination components from the image, while the latter also reduces noise. However, the output of these components has a peeled-off appearance. To restore these components they employed a UNNP that was able to generate realistic textures. They conducted experiments on public datasets and demonstrated the robustness of their proposed method to illumination. However, the proposed approach was not able to handle pose and expression variations.

When an image is affected by multiple distortions, the classical UNNP may not provide the expected performance in image restoration. To address this limitation, a novel strategy named dual prior learning (DPL) was proposed in a recent study [129]. DPL employs two networks, a UNNP to model image specific prior (using random noise) as well as another network to capture distortion prior, which is used to learn information about the different distortions. Reported results show that DPL significantly outperforms the conventional UNNP in the restoration of images having multiple distortions.

4.4 Image Denoising

In the original UNNP paper [6], the authors showed that the parametrization approach of UNNP demonstrates high impedance to image noise. Thus, UNNP can be naturally used for image restoration, whose goal is to recover clean image x from noisy observation x_0. The degradation model can either be known or unknown, i.e., blind image denoising. Zou et al. [95] used UNNP for the denoising of endoscopic images. They integrated a trained MobileNet model for blind image quality assessment (which was used for stopping UNNP iterations). Also, they employed transfer learning strategy, which was able to reduce UNNP iterations by 36%. Park et al. [96] proposed a variational UNNP for jointly solving the problem of demosaicing and denoising. They used the same U-net architecture as in the original UNNP [6] but proposed a new loss function that incorporates both constant (z_c) and varying noise (z_v) derived from a Gaussian distribution. z_c remains constant until p^{th} iterations before z_v is added to z_c. Introducing variational noise in UNNP optimization provides better convergence, compared to the original UNNP in terms of denoising performance. However, their proposed framework has higher execution cost than the standard UNNP approach. The use of UNNPs for the restoration of corrupted remote sensing images, especially where multi-temporal snapshots are not available was proposed in [130]. The key purpose of using UNNP was to fill gaps in corrupted remote sensing images (similar to a typical image inpainting problem). In [131], UNNP was used for the denoising of a synthesized multiplane image (MPI). The authors found that a UNNP integrated with multilayer perceptrons results in a better regularized MPI while providing superior performance compared to standard optimization without UNNP. In [132], authors integrated physics-inspired sensor modeling with UNNP for the denoising and super-resolution of thermal images. Moreover, given that their approach is modular, they suggested that any UNNP framework can be used in the proposed approach, e.g., Deep Decoder.

4.5 Image Deblurring

Although noise filters can effectively remove noise from a blurred image, they can also damage the blurred information by introducing a more serious blur. Feng et al. [97] proposed the use of UNNP to restore the noisy and blurred images using a single degraded image. They incorporated a RED regularizer for optimizing UNNP for the task of

Poissonian image deblurring, as the classical UNNP approach is not effective in handling Poisson noise [97]. Their learned denoiser (a neural network) can then be used as a regularizer to constrain the latent clear image. They combine this prior with the L_0 regularization before establishing a restoration model for the Poisson image. They performed an experimental evaluation on several real and simulated images and showed that their proposed method achieves competitive results. Furthermore, their experiments show that their proposed method is not only able to suppress the staircase effects in the image but is also able to preserve the details. Zukerman et al. [98] proposed BP-DIP by combining UNNP with a back projection (BP) fidelity term used in place of the standard MSE loss (usually used in UNNP). TV regularization was also incorporated in the proposed framework to circumvent the noise sensitivity of BP. They experimentally demonstrated that their proposed method is not only able to achieve higher performance (in terms of PSNR value) compared to previous works but also a better inference run-time. Their results show that BP-DIP yields higher PSNR and reaches its peak PSNR in relatively fewer iterations than standard MSE based UNNP. However, they argued that early stopping is still required to avoid the overfitting problem of UNNP.

4.6 Image Super-Resolution

In an image super-resolution problem, given a low resolution (LR) RGB input image $x_0 \in \mathbb{R}^{3\times H\times W}$ and an upsampling factor t, the objective is to generate a high resolution (HR) RGB image $x \in \mathbb{R}^{3\times tH\times tW}$ such that when x is downsampled by a factor of t, the output is the same image as x_0. Hence, super-resolution is an ill-posed problem because there can be an infinite number of HR images that can correspond to the same LR when downsampled.

4.6.1 Hyperspectral Image Super Resolution

Fusing a low spatial resolution (LR) hyperspectral image (HSI) with a high spatial resolution (HR) multi-spectral image (MSI) is an effective way to achieve HSI super-resolution. The results generated using existing techniques are based on the assumption that both input images are clean. This is too idealistic for real cases. To address the problem of noisy HSI and MSI input, Nie et al. [99] proposed a UNNP-based HSI super-resolution method in which UNNP is employed as the prior of the latent HR HSI which converts it into an end-to-end DL problem and solved it using back-propagation to capture better statistics of the latent HR HSI. Furthermore, they demonstrated the effectiveness of their proposed technique by evaluating it on two benchmark datasets: the CAVE dataset and the Washington DC dataset. Ma et al. [100] proposed a combination of UNNP with a super-resolution convolutional neural network (SRCNN) to estimate fine resolution fraction images for each land cover type. The proposed UNNP-based approach was shown to be quite robust on small objects. It also results in reduced soft classification uncertainty. Similarly, a UNNP-based approach for coded HSI reconstruction was proposed by Zhang et al. [101]. UNNP learns the deep prior from the external dataset, as well as the internal information of the spatial-spectral restricted input coded image. They showed that their proposed method can sufficiently represent HSIs by effectively exploiting the spatial-spectral correlation. Using both quantitative metrics and perceptive consistency, their results proved that their proposed methodology outperforms the state-of-the-art.

Nie et al. [102] argued that the success of the existing HSI super-resolution methods based on fusion depends on the premise that the images used for fusion (i.e., the HSI low-spatial-resolution input and the multispectral image with low-spectral resolution) are exactly registered. While such a premise is too idealistic for the real world, few efforts have taken this issue into account. As a solution to this, Nie et al. [102] proposed the integration of image registration into HSI super-resolution for joint unsupervised learning. To learn the parameters of the affine transformation between the two input images, they used a UNNP-based spatial transformer network (STN) that avoids overfitting by constraining the STN. UNNP was used for the estimation of latent HSI image Z from two observed input images. Through experimental evaluation, they showed that their methodology can successfully deal with unregistered input images.

4.6.2 Depth Map Super Resolution

RGBD images typically offer high-resolution color and lower-resolution depth. Voynov et al. [103] argued that the low-resolution of the depth maps can be improved by using the color information from RGBD images, which can be effectively used for the 3D reconstruction of images. To leverage this, they proposed a novel visual appearance-based loss function and integrated it with different depth processing methods including UNNP. They showed that UNNP yields dramatically improved 3D shapes when optimized using their proposed visual difference-based loss function, as it does not suffer from false geometry artifacts, unlike the classical UNNP (that does not incorporate the aforementioned loss). Furthermore, they showed that UNNP can be used to simultaneously solve super-resolution and inpainting problems.

4.6.3 UNNP Variant for SR

Sagel et al.[9] argued that the optimization in the original UNNP limits the output image being reconstructed to be represented by a convolutional neural network. This might neglect prior knowledge and may make certain regularizers ineffective in specific cases. As a solution, they suggested an alternative approach that relaxes this constraint and takes full advantage of all prior knowledge. They demonstrated the effectiveness of their approach to the task of image super-resolution and showed that their algorithm provides a substantial improvement over the original UNNP algorithm.

4.7 Image Inpainting

Manual image inpainting requires much domain knowledge for supervised learning-based automatic image inpainting methods, which require extensive training and large-scale annotated training data. To circumvent this issue, Weber et al. [104] proposed a UNNP-based technique named "Interactive UNNP", which is a combination of manual and automated processes, i.e., it keeps a human in the loop during the inpainting process. The human acts as a

guide for the automated inpainting process by iteratively embedding the domain knowledge into it. They evaluated their proposed method with five other state-of-the-art techniques and empirically showed that their proposed method can generate better results compared to other state-of-the-art methods with even very little guidance. However, their proposed approach requires a human observer to inject subjective feedback during the (iterative) optimization of the UNNP framework, which increases the optimization time and limits the real-time practicability of the proposed method. To reconstruct a depth map from a noisy and incomplete depth map, Ghosh et al.[105] used UNNP for estimation of depth maps and incorporated it with a depth reconstruction loss and a view-constrained photo-consistency loss (which is measured using a geometrically calibrated camera taking images from surrounding viewpoints). They applied their technique to inpainting for both binocular and multi-view stereo pipelines and demonstrated that dense 3D models of higher quality are produced by maps refined by their proposed UNNP-based technique. Traditionally, hand-crafted approaches have been used for depth map estimation, which suffers from a number of imperfections, i.e., large uniform regions and textured areas and occlusions [133]. Similarly, in [46], UNNP was used for 3D image completion and reconstruction. Also, the authors interpreted the effectiveness of UNNP using a neural tangent kernel (NTK) perspective and found that UNNP was effectively able to complete missing image details (analyzed in NTK feature space). A sparse 3D CNN architecture was used in the proposed UNNP framework.

4.8 Phase Retrieval

Phase retrieval aims to reconstruct the phase of the signal/image using only its magnitude measurements. It is a non-linear and highly ill-posed inverse problem. To construct the target image in coherent diffraction imaging, the intensity of the diffraction pattern scattered from a target is first measured and then a phase retrieval algorithm is used to find its phase. Algorithmic phase retrieval provides an alternative way to recover the phase of signals without measuring the signal using a sophisticated method. For instance, Jagatap et al. [17] proposed to use UNNP for compressive phase retrieval, which is a non-linear inverse problem, whose goal is to reconstruct a d-dimensional signal from n magnitude-only measurements such that $n \ll d$. UNNP for the given task was optimized using two approaches, i.e., using gradient descent and projected gradient descent. Similarly, in [111], the authors proposed to use UNNP for unwrapping the phase in 2D quantitative phase imaging. The proposed method was evaluated on organoid images that are acquired using digital holographic microscopy. In [26], the authors used UNNP for Fourier ptychography, which is a special case of phase retrieval. Similarly, Boominathan et al. [112] proposed to use autoencoder architecture in UNNP framework for phase retrieval in Fourier ptychography. In [27], authors incorporated physical model for image formation into UNNP network. The modified UNNP was able to reconstruct single-beam phase imaging while only getting the single diffraction pattern of a phase object as an input. Standard mean square error (MSE) loss was used to update the neural network parameters, while the loss was measured between the output and input image.

4.9 Image Deconvolution

Wang et al. [113] combined the idea of learning-free deconvolution methods with neural networks and proposed a variant of UNNP named deep image kernel prior (DIKP), in which they modified the objective function of classical UNNP to deconvolution energy function. They showed that DIKP improves the performance of image deconvolution and outperforms traditional learning-free regularization-based priors for image deconvolution. Ren et al. [31] proposed a hybrid method for blind image deconvolution. They used UNNP (an autoencoder architecture with skip connections) to generate a clean image and a fully-connected network (FCN) for modeling blur kernel. Furthermore, they incorporated a TV regularizer and proposed two algorithms to solve the inverse problem, i.e., alternating optimization and joint optimization. In [30], UNNP was used for learning deep prior from seismic data interpolation, in the proposed UNNP framework Multi-Res U-Net architecture was used for learning the deep prior.

4.10 Seismic Data Reconstruction

This problem is concerned with the reconstruction of missing traces in seismic data, which is used for the extraction of geophysical information in structural analysis. This problem is very similar to compressed sensing as seismic data reconstruction is performed using sub-sampled data. In [134], a UNNP-based algorithm named DSPRecon was proposed for the reconstruction of seismic data. A U-Net with skip connections was used in the proposed UNNP framework that learns the deep prior for seismic image reconstruction. The authors demonstrated that the proposed method performs comparatively better than the spectrum analysis (SSA) and Cadzow-based reconstruction methods. Similarly, Park et al. proposed an approach that integrates projection onto convex sets (POCS) based regularization with UNNP for the reconstruction of seismic data [135]. Integration of POCS regularization within UNNP optimization eliminates the limitation of the classical UNNP approach in filling large gaps in the data.

4.11 Adversarial ML

4.11.1 Defending Adversarial ML Attacks

In the literature, it has been shown that reconstructing an adversarial image using a UNNP network cleans the adversarial perturbations from the final reconstructed image [114]. Kattamis et al. investigated the properties of the early outputs of the UNNP and demonstrated that these early iterations demonstrate invariance to adversarial perturbations by classifying progressive UNNP outputs and using a novel saliency map approach. Furthermore, they argued that using UNNP as a defense against adversarial attacks has great potential against three well known adversarial ML attacks, i.e., fast gradient-sign method (FGSM) [?], basic iterative method (BI), and least-likely class iterative method (LLCI) [136]. A similar type of behavior in early iterations of UNNP was observed in [115], where the UNNP was shown to produce adversarially robust features in earlier iterations compared to the features

learned by later iterations (which were found sensitive to adversarial perturbations). Furthermore, to leverage the UNNP network as a defense against adversarial attacks, the authors proposed an adaptive stopping strategy using a second-order exponential smoothing strategy for stopping UNNP-based reconstruction. Sutanto et al. [116] performed a similar type of analysis and proposed to use the UNNP network to remove FGSM-based adversarial perturbations. Similarly, the use of UNNP to clean adversarial perturbations was also investigated in [117]. The authors considered two adversarial ML attacks FGSM and Carlini and Wagner (C&W) [137] in both black and white-box settings. In [138], authors used UNNP to develop an attack and model agnostic defense for defending against adversarial attacks. They considered both black and white box adversarial attacks for evaluation.

4.11.2 Crafting Adversarial Perturbations

The problem of finding imperceptible adversarial samples by introducing some noise into the input samples is known as crafting adversarial perturbations. In contrast with the aforementioned studies, Gittings et al. [118] proposed a method to generate robust adversarial image examples using a UNNP framework. They empirically showed that using UNNP to reconstruct an image under adversarial constraints induces perturbations that are more robust to affine deformations. Yin et al. [13] investigated the adversarial vulnerability of UNNP for the super-resolution task against three attacks, i.e., style transfer attack, classification attack, and caption attack.

4.11.3 Generation of Counterfactual Explanations

Counterfactual explanations is one of the most emerging methods to interpret neural network's predictions. These explanations are used to describe input changes affecting the model's prediction to a predefined output. Narayanaswamy et al. [119] proposed a regularization strategy based on an auxiliary loss estimator, which efficiently guides the UNNP to recover natural pre-images. They performed experiments with a real-world International Skin Imaging Collaboration (ISIC) skin lesion detection problem and showed the effectiveness of their proposed technique in synthesizing meaningful counterfactuals. They argued that the standard UNNP inversion often proposes visually imperceptible perturbations to irrelevant parts of the image, while their proposed approach systematically introduces perturbations in the lesion-specific regions. This is strongly in line with the widely adopted signatures for lesion type detection. However, quantitative evaluation of the generated images was not performed.

4.12 UNNP for Video

4.12.1 Video Inpainting

Inpainting is an ill-posed problem due to the non-existence of a unique solution. Extending this problem to video brings more challenges as the inpainted content needs to be consistent across the frames of the video. Zhang et al. [106] leveraged UNNP for inpainting that was able to generate missing appearance and motion information while enforcing visually plausible textures. Furthermore, they showed that their proposed framework is able to ensure mutual consistency of both appearance and optical flow of the video. L_2 reconstruction loss is used for image reconstruction based on the known portions of the image. The network is augmented to predict the color and flow values at each pixel for 6 consecutive frames of the video. The flow generation loss is then defined based on the known regions to encourage the UNNP network to learn a "flow prior". A consistency loss is then defined on the basis of these two losses (i.e., L_2 and flow generation loss) to ensure the temporal consistency between the generated frame and the generated flow. Furthermore, a perceptual loss is defined over the extracted features to improve the visual sharpness of images generated by UNNPs.

4.12.2 Video Stabilization

Video destabilization occurs due to various problems such as lens distortion, dynamic objects, motion blur, and low illumination. A UNNP-based video stabilization system is proposed in [107]. The authors formulated the problem as a large-scale non-convex optimization problem and suggested an optimization routine to solve it, moving it to the realm of neural networks. Furthermore, they argued that it is better to model the shift in frames as a dense optical flow field of consecutive frames instead of using complex motion models. More specifically, they incorporated optical flow-based objective function in the UNNP framework to achieve visually more plausible and quantitatively better results than the previous state of the art.

4.12.3 Temporal Consistency

Lei et al. [108] argued that the application of image processing methods independently to each frame of a video might lead to inconsistencies in the resulting video. To overcome this issue, they proposed the use of UNNP to achieve blind video temporal consistency. They trained the model using a pair of original and processed videos and used an iterative reweighting training strategy to address the multi-modal inconsistency problem. Furthermore, they demonstrated the effectiveness of their approach for different tasks such as colorization, dehazing, image enhancement, style transfer, image-to-image translation, intrinsic decomposition, and spatial white balancing.

4.12.4 Video Restoration

Liu et al. [109] proposed a UNNP-based hybrid method to recover old movies. Their proposed framework has two main components, i.e., speckle noise detection using spatio-temporal filtering techniques and UNNP-based image inpainting. UNNP was used to reduce speckle noise and to fill missing image portions in videos of old movies.

5 Applications of UNNP in Medical Imaging

In realistic medical settings, the availability of large-scale representative training data is often very challenging due to annotations cost, time, and ethical constraints. UNNP provides a means to circumvent this issue by providing the ability of fitting medical images directly using the structure of CNNs. In this section, we provide a detailed discussion

on the applications of UNNP to different medical imaging tasks, i.e., medical image reconstruction and enhancement. Note that although these two tasks (i.e., medical image reconstruction and enhancement) come under the umbrella of inverse image reconstruction problem, we deliberately described the papers on reconstruction and enhancement separately for ease of understanding. A summary of the various applications of UNNP for different medical imaging tasks is also presented in Table 4.

5.1 Medical Image Reconstruction

In recent years, various DL learning-based approaches (e.g., generative models) have been proposed for the reconstruction of different medical modalities, e.g., magnetic resonance imaging (MRI) and positron emission tomography (PET). However, these methods mainly rely on extensive training using large-scale annotated datasets, e.g., training end-to-end GAN for the task of medical image enhancement using paired images (pair of corrupted and clean images). On the other hand, UNNP-based methods have achieved competitive results without the requirement of such large training datasets. Here, we elaborate upon the application of UNNPs for medical image reconstruction tasks.

5.1.1 MRI Reconstruction

In [139], UNNP along with nonuniform fast Fourier transform has been used for the reconstruction of time-dependent dynamic MRI from sparsely acquired measurements. Specifically, 1D manifold learning and a mapping network were used to improve the reconstruction performance. Similarly, in [140], the authors focused on the problem of accelerated MRI reconstruction and proposed a UNNP variant that leverages both over and under-parameterized UNNPs. The proposed method outperformed classical sparsity-based methods and provided a comparable performance as that of training-based methods. Zhao et al. [141] proposed a UNNP-based framework for MRI reconstruction from undersampled k-space data. Their proposed method uses a high-resolution MRI image as an input reference to learn a structural prior during the training process of UNNP.

5.1.2 Tomographic Image Reconstruction

5.1.2.1 PET Image Reconstruction. Gong et al. [11] proposed to leverage UNNP for PET image reconstruction that does not require any paired training data, except for the patient-specific measured data. In the proposed approach, UNNP was optimized using patient-specific measured data and prior information. More specifically, they formulated the problem of PET image reconstruction as a constrained optimization problem and solved it using the alternating direction method of multipliers. Furthermore, the authors demonstrated that their proposed UNNP-based method provides a comparatively improved lesion contrast as compared to Gaussian filter. Later, the authors extended this approach for low dose dynamic patlak PET image reconstruction [142]. UNNP was then optimized using measured sinogram data and patient-specific anatomical prior image. Similarly, Yokota et al. [15] incorporated non-negative matrix factorization with UNNP for the reconstruction of dynamic PET images using sinograms. They quantitatively evaluated their proposed method on both simulated data and clinical data. U-Net architecture was used in the UNNP framework that was trained to extract the spatial factor decomposed from the data matrix. Multiple U-Nets in the UNNP framework were also used for dynamic PET image reconstruction using sinograms data with the employed U-Nets ranging from 1-5.

5.1.2.2 Diffraction Tomography Reconstruction. Zhou et al. [143] proposed a UNNP based method for the reconstruction of high resolution 3D refractive index of thick biological samples using diffraction tomography. In the proposed framework, a phase retrieval algorithm is used to process the multi-angle data while UNNP (with 3D U-Net) was used for reparameterization for 3D image reconstruction. Furthermore, the authors demonstrated that their proposed UNNP based method efficiently solves the missing cone problem.

5.1.2.3 CT Image Reconstruction. Gong et al. [144] proposed to use the UNNP framework for the reconstruction of low-dose dual-energy CT images from noisy observations. UNNP was used for the joint reconstruction of low and high energy images. Furthermore, to enhance the performance of UNNP based reconstruction, they introduced isotropic TV regularization in UNNP optimization. Similarly, Baguer et al. [24] proposed to incorporate classical TV regularizer the UNNP optimization framework for the reconstruction of low-dose and sparse angle CT images.

5.1.2.4 Magnetic Particle Imaging (MPI). Dittmer et al. proposed to use UNNP for 3D MPI reconstruction and performed a quantitative comparison of different regularization techniques such as [145]. In their proposed 3D UNNP model, unlike the original UNNP, they do not use skip connections but they also used 3D convolutions with ReLU as an activation function in the U-Net model.

5.1.2.5 Microscopic Image Reconstruction. To overcome the trade-off between spatial sampling and imaging speed encountered in photoacoustic microscopy, Vu et al. proposed to use UNNP for photoacoustic microscopic image reconstruction from sparsely sampled images [152]. Unlike classical UNNP optimization in which fixed noise is used, the authors considered perturbing noise. The proposed framework was evaluated using vascular and non-vascular data and a significant performance improvement was achieved.

5.2 Medical Image Enhancement

5.2.1 Haze Removal in Fundus Imaging

In [7], the authors demonstrated that coupled UNNP networks are capable of decomposing an input image into its individual elements, e.g., decomposing an image into foreground and background for a segmentation task. Qayyum et al. [25] leveraged this idea of image decomposition using coupled UNNP networks. They proposed an unsupervised framework for retinal fundus image enhancement using coupled UNNP networks by integrating dark channel prior loss into the overall loss of coupled UNNP networks. Their method does not require any training data neither paired (i.e., paired clean and corrupted reference image) nor unpaired (i.e., clean and corrupted image) and recovers an enhanced retinal image using a single degraded input. Building upon their previous work [25], the authors

TABLE 4
Summary of Medical Imaging Applications of Untrained Neural Network Priors (UNNPs)

Application	Ref	Method	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
Compressed Sensing	[123]	An UNNP framework compressed sensing application incorporating a learned regularizer to facilitate non-linear measurements.	- MNIST, Chest X-Ray, & STARE - Mean Square Error (MSE)	UNNP with learned regularization outperformed two baseline unlearned methods in terms of MSE, except BM3D-AMP for higher measurements.	Although learned regularization improved UNNP performance in terms of MSE as compared to standard UNNP, however, it requires a small labelled training set, which could be difficult to obtain in medical settings.
MRI Reconstruction	[139]	UNNP was used with nonuniform fast Fourier transform. Specifically, 1D manifold was used to learn the temporal dependencies and a mapping network was used to improve reconstruction performance.	- Retrospective Dataset & Real - Fetal Cardiac Data - Regressed SNR	Proposed UNNP-based framework provided a performance gain of approximately 20%, 4%, and 3% when compared with three baseline methods (non-learning).	In addition to UNNP optimization time, the slow forward model further increases the overall execution time. Moreover, the manifold selection and mapping network design has a direct impact on the performance for a given dataset.
	[140]	UNNP variant is used for accelerated MRI reconstruction, which is a combination of both under-parameterized and over-parameterized untrained neural networks.	- FastMRI - PSNR, SSIM, VIF, & Multi-Scan SSIM	Outperformed classical untrained methods and provided similar reconstruction performance as that of training-based methods.	Highly dependent on the chosen number of channels for each layer in the UNNP architecture and does not perform well if the number of channels is either too small or too large.
	[141]	Used high-resolution MR image as input reference for learning structural prior during optimization of UNNP. The reference image was used for guiding UNNP.	- Vivo MR - Relative error, PSNR & SSIM	Outperformed classical zero filtering-based method and provided a performance gain of $\approx$ 2% as compared to standard UNNP.	It involves iterative computation of undersampling at each iteration, which increases its overall execution time as compared to classical UNNP.
PET Image Reconstruction	[11]	Optimized UNNP for PET image reconstruction using the alternating direction method of multipliers method.	- 3D BrainWeb - CRC vs. STD	Outperformed a hand-crafted baseline method while generating better lesion contrast.	Higher computational overhead due to the use of 3D U-Net. Also, it cannot handle images acquired using different scans.
	[142]	Extension of [11] for low dose dynamic patlak PET image reconstruction.	- Self-prepared	No quantitative evaluation is performed.	High memory requirement and large execution time due to the use 3D U-Net.
	[15]	Multiple UNNPs were used for the extraction of spatial factor from decomposed data matrix (via non-negative matrix factorization) for dynamic PET reconstruction.	- Self-prepared & Simulated Data - Signal to Noise Ratio (SNR)	Outperformed three existing classical baseline methods and provided lower performance as compared to EM in the noise-free case.	Reconstruction performance degrades at higher iterations, as no early stopping was employed and the number of iterations was selected empirically.
Diffraction Tomography Reconstruction	[143]	UNNP with 3D U-Net was used for the reconstruction of high resolution 3D refractive index of thick biological samples using diffraction tomography.	- 3D isotropic EM images of hippocampal cells - RMSE & SSIM	Proposed UNNP framework outperformed standard UNNP in terms of reconstruction performance.	UNNP optimization diverges rapidly in some cases, thus indicating early stopping is required. Also, larger execution time due to the use of gigavoxel data.

TABLE 4
(*Continued*)

Application	Ref	Method	Dataset (s) & Metric (s)	Performance Comparison	Limitation (s)
CT Image Reconstruction	[144]	Used isotropic TV regularized UNNP for CT image reconstruction that works by joint reconstruction of low and high energy images.	- Self-prepared DECT - Contrast-to-Noise Ratio (CNR)	Additional regularization enhanced performance as compared to standard UNNP and outperformed one classical denoising method.	High execution time due to the use volumetric data, additional regularization, and underlying 3D neural network in UNNP framework.
	[24]	Used UNNP along with classical TV regularizer for CT image reconstruction.	- LoDoPaB-CT & Ellipses - PSNR & SSIM	Outperformed two classical learning free methods, i.e., TV and filtered back-projection.	High number of iterations were used (i.e., 8000, 5000, etc.), which incurs large execution time.
Magnetic Particle Imaging	[145]	U-net is used in UNNP without skip connections and 3D convolutions with ReLU as an activation function.	- 3D open MPI - PSNR & SSIM	Proposed 3D UNNP outperformed learning free and data-driven training-based methods.	High execution time due to the use of volumetric data and 3D neural networks in the UNNP framework.
Microscopic Image Reconstruction	[146]	Used UNNP to address the trade-off between spatial sampling and imaging speed in photoacoustic microscopy.	- Self-prepared - PSNR & SSIM	Outperformed hand-crafted methods and was competitive with data-driven DL methods.	UNNP was not able to recover structures that were missed during undersampling.
Haze Removal	[25]	Multiple UNNPs were used to decompose retinal images for haze removal. To improve the performance, dark channel prior loss was incorporated with UNNP.	- Five public datasets used - PSNR, SSIM	UNNP-based approach outperformed existing classical learning-free methods as well as training-based methods.	Integrating DCP loss increases the computational overhead, as it is calculated iteratively during each iteration of UNNP optimization.
Tomographic Image Denoising	[147]	Used UNNP for dynamic PET image dnoising using gray and white matter data.	- 3D BrainWeb - PSNR, SSIM & Regional TACs	Provides better statistical noise reduction and efficiently preserves the cortex as compared to hand-crafted baselines.	The reconstruction performance decreases in higher iterations, as no early stopping was employed and the number of iterations was selected empirically.
	[148]	Used 3D U-Net in UNNP and trained it using the Limited Memory BFGS optimizer for denoising PET images.	- Self-prepared dataset - CNR	Outperformed Gaussian and non-local mean filtering-based hand-crafted methods.	Does not provide good performance in handling large noise levels and large execution time due to the use of 3D data.
MRI Image Correction	[149]	Formulated MRI reconstruction as a Bayesian inference problem and used two UNNPs models to obtain an inhomogeneity-free image.	- 3D BrainWeb & OASIS3 - Norm. Cross Correlation, Coefficient of Variation	UNNP provided comparable quantitative performance to a well known baseline method (i.e., N4) while visual results were even better.	UNNP optimization process becomes unstable sometimes resulting in the generation of many artifacts in output images.
OCT Image Denoising	[150]	Used non-local UNNP for denoising OCT retinal images by incorporating an autocorrelation loss	- SDOCT image datasets - CNR, Equivalent No. Looks	Outperformed standard UNNP and as well as other learning and learning free baselines.	Modified loss function increases computation complexity of the iterative process and increase overall execution time.
	[151]	Used under-parametereized UNNP for denoising of OCT retinal B-scans with different losses.	- Self-prepared dataset - CNR & SSIM	UNNP performance was slightly lower than the data-driven deep learning model.	Poor performance with increasing noise level and optimization was prone to overfitting (as no early stopping applied).

investigated the use of dark and bright channel priors for atmospheric light estimation for retinal fundus image enhancement. Such an approach eliminates the need for separate UNNPs for atmospheric light estimation. The authors also investigated pretrained UNNP and found that pretraining gives the additional benefit of learning fine details in early iterations over the cost of slight drop in overall performance.

5.2.2 Tomographic Image Denoising

Hashimoto et al. [147] proposed a UNNP based dynamic PET image denoising technique, which can handle unknown cases. They evaluated their proposed method in terms of peak signal-to-noise ratio (PSNR), structural similarity (SSIM) index, and the regional TACs of the gray and white matter using both synthetically generated data and real data acquired from a living monkey brain with 18F-fluoro-2-deoxy-D-glucose (18F-FDG). Their results quantitatively show that the proposed UNNP method achieves a better reduction in statistical noise and better preserves the cortex compared to the existing hand-crafted baseline algorithms. Similarly, Cui et al. [148] proposed to use UNNP for denoising PET images by feeding the UNNP network with CT images as an input. Furthermore, they adopted the 3D U-Net architecture as proposed in [153] instead of U-Net with 2D convolutions (as used in classical UNNP framework) and trained the network using a Limited Memory BFGS optimizer. In [154], authors used 4D convolutional architecture in UNNP for denoising dynamic PET images.

5.2.3 MRI Image Correction

Han et al. [149] formulated the problem of MRI reconstruction as a Bayesian inference problem to obtain an inhomogeneity-free image. Their proposed framework has two UNNP networks that were used to learn the inhomogeneity field and priors of the image, while their likelihood was derived from the observed image itself. To demonstrate the effectiveness of the proposed method, the authors performed an experimental evaluation on both simulated and real data and compared the performance with a well-known inhomogeneity correction method named N4 [155].

5.2.4 OCT Image Denoising

In [150], a non-local UNNP was proposed to denoise OCT retinal images by incorporating an autocorrelation loss (that exploits sorted non-local statistics) into the overall reconstruction loss of UNNP. The proposed method was evaluated using both subjective (visual quality) and objective metrics i.e., a contrast to noise ratio (CNR) and an equivalent number of looks (ENL). Similarly, Hagan et al. investigated different network architectures and loss functions with UNNP to denoise OCT retinal B-scans [151]. They performed a comparative analysis in terms of different standard metrics (i.e., contrast-to-noise ratio (CNR) and SSIM).

6 Insights and Pitfalls

Using UNNPs for IIPs. The under-parameterized untrained neural networks (e.g., the deep decoder [12]) can only fit signals with low complexity, while over-parameterized untrained neural networks (e.g., UNNP [6]) can certainly model anything. However, the literature suggests that they perform better for low complexity signals. Untrained neural networks exploit the smoothness and locality information of natural images to recover them without using any training data such information generally does not exist in any arbitrary image, therefore, they are capable of effectively modeling natural images. When modeling natural images using untrained neural networks, the operations in CNN like upsampling and convolutions enforces the smoothness and locality information.

UNNPs can be used to aid in the learning of priors. For instance, Hussein et al. [156] used a trained GAN to provide a quick start to an untrained neural network. Similarly, in [157], the authors treated the UNNP as an image specific prior to learn high dimensional latent representations from trained neural networks. Moreover, it has been shown in [12] that the prior learned by UNNPs is asymptotically equivalent to hand-crafted priors (e.g., Gaussian processes (GP)). However, GP is computationally more expensive than stochastic gradient descent (SGD) optimization when we aim at modeling high-resolution images. Although, UNNPs effectively regularize the input being reconstructed (by optimizing its parameters), it has been shown that their performance can be improved by adding further regularization (, e.g., using total variation [74]).

Limitations of UNNPs. It is difficult to articulate what prior knowledge is being captured by the untrained neural networks for a particular input image. As described above, over-parameterized untrained neural networks perform better for low complexity signals. However, we cannot articulate what aspects of low complexity are desired by these networks. This demands a theoretical understanding of the learning behavior of untrained neural networks. Moreover, the prior estimated by untrained neural networks using SGD and stochastic gradient Langevin dynamics (SGLD) matches the prior estimated by GP for smaller size models. However, it is still unknown whether this holds for larger networks [12]. Yokota et al. [15] found that a randomly initialized neural network (i.e., U-Net) often generates edge-enhanced images before the convergence of the parameters update process. The authors articulated that these edge enhancements are not feasible for medical image reconstruction. To circumvent this issue, they proposed to combine outputs of multiple parallel models to reconstruct dynamic PET images. However, this increases the computational complexity of the overall framework and demands further developments. The authors proposed an optimal stopping strategy before the convergence of the parameters.

The literature suggests that the optimal stopping point for untrained neural networks depends on the structure of neural networks being used [68]. However, it is difficult to identify the optimal stopping point for any neural network architecture. Therefore, this remains an open-ended problem. Also, UNNPs are considerably slower than discriminative DL methods (and sometimes even slower than generative DL methods). For instance, the literature suggests that the optimization of UNNPs usually requires several minutes even on a single GPU [6]. Similar observations have been noted in different other papers, e.g., [7], [11], [74], [75], [117]. This hinders the real-time applicability of

UNNPs in such applications, e.g., real-time applications and applications constrained by limited computational resources. A few recent studies have attempted to address this issue by using pretraining, e.g., [49], [158].

Despite UNNPs' ability to achieve comparable performance to DL methods, they have certain limitations when used in some applications. In image super-resolution, UNNPs cannot recover fine image details from low-resolution images. In contrast, DL-based approaches can still leverage large datasets to recover such details. UNNPs are also less effective than end-to-end DL methods for inpainting large holes in images. In denoising applications, the performance of UNNPs degrades in the presence of large noise. However, early stopping can prevent overfitting while providing satisfactory results even in the presence of large noise. In image deblurring and deconvolution, UNNPs struggle to handle non-uniform deblurring and high noise levels. Similarly, it is difficult to balance image quality and perturbation removal when using UNNPs for adversarial noise-cleaning. UNNPs have prohibitively high execution times and memory requirements when processing video and 3D volumetric data. On the other hand, UNNPs often outperform DL-based methods in certain applications, e.g., blind video temporal consistency [108] and finding semantic correspondences between images [67].

7 Open Research Problems

7.1 Exploring Network Architectures

The characteristics of UNNPs are very similar to hand-crafted priors-based modeling. In hand-crafted-based modeling, we handcraft individual basis elements while in untrained neural networks we handcraft neural network architectures. For example, one can design a specific neural network architecture (of their own choice) for the modeling of a particular image. Thus, any application of an untrained neural network for typical inverse problems relies on the preference of a specific neural network structure, i.e., a different CNN architecture will provide different results. In the literature, a number of studies have proposed methods to construct the most relevant neural network architecture for a given specific task. However, exploring the most relevant and application-specific (i.e., optimal) neural architecture(s) for UNNP to model a particular inverse problem at hand is still an open research problem.

7.2 Execution Time

UNNPs provide near state-of-the-art performance without the requirement of massive training datasets. However, their iterative nature makes them unsuitable for use in real-time applications. This means that a particular application of UNNPs requires considerable time and computational resources due to their inherent iterative optimization process. Therefore, this has particular implications for their use in critical applications such as healthcare and surveillance applications where immediate outcomes may be required. In this regard, it is very important to explore new optimization approaches to reduce the execution time of the UNNPs-based unsupervised approaches.

7.3 Early Stopping

Despite their ability to effectively model natural images, UNNPs suffer from the issue of overfitting. In the original UNNP paper [6], the authors relied on early stopping to avoid overfitting. To overcome this overfitting issue, different methods have been proposed in the literature, e.g., the deep decoder [12], Bayesian perspective of UNNP [16], and the projectional method [71]. These studies provided alternative approaches to alleviate the need for early stopping in UNNPs. However, these methods are not generalized. The development of efficient and customized approaches to alleviate the issue of overfitting, therefore, requires further investigation. We refer the interested readers to a comprehensive tutorial on developing Bayesian methods for DL models [120].

7.4 Theoretical Understanding

Despite the impressive success of UNNPs in numerous IIPs, there is still a lack of theoretical understanding of why UNNPs are able to effectively model natural images, why the CNN architecture enforces a strong prior on natural images, and why UNNP cannot model any arbitrary image. Notwithstanding a few studies have provided some theoretical guarantees about the capabilities of UNNPs, e.g., Heckel et al. [12] provided the theoretical reasoning behind their proposed variant of UNNP (i.e., deep decoder). Similarly, Yokota et al. proposed manifold modeling in an embedding space to interpret UNNP [80]. However, more questions still need to be investigated about the working process of UNNPs. Therefore, the development of a proper theoretical framework to understand when they work, how to properly regularize them and interpret them, and how to measure their complexity remains an open research problem. For example, it would be interesting to investigate what type of information a typical untrained neural network prefers most and what it does not, while capturing the underlying image prior. This will eventually help to develop customized neural network architectures for particular tasks at hand.

7.5 Investigating UNNPs for High-Level Vision Tasks

Since their inception [6], UNNPs have been mainly used for solving low-level vision problems such as denoising, super-resolution, deblurring, etc. Their applications to high-level vision problems such as detection have received little attention. For instance, a recent study has shown that UNNPs can effectively be used to detect human faces [52]. Their work makes it a promising future direction to explore. Specifically, via extensive experiments, the authors showed that units sensitive to human faces emerge in the high and mid-level layers of randomly initialized neural networks, thereby enabling them to accomplish face detection tasks. This is one of the foremost works that highlight the effectiveness of UNNPs for high-level vision tasks thereby providing an avenue for many areas of exploration.

8 Conclusion

Inverse problems arise in many distinct applications such as restoration, denoising, inpainting, deconvolution, and medical

imaging. The use of untrained network priors to solve inverse problems is attractive since it can provide competitive results without requiring any training data. This is particularly appealing in settings where the data is scarce and difficult to obtain for logistical reasons or privacy concerns (as in healthcare). In this paper, we have presented the first comprehensive review on the use of untrained neural network priors (UNNPs) based methods for inverse imaging applications. Moreover, we have developed a taxonomy of different applications of UNNPs and presented their review across two dimensions, i.e., general and medical applications. Our literature review has found that UNNPs based approaches are particularly attractive for medical applications due to the scarcity of annotated data. Finally, this paper outlines various open research issues related to UNNPs that require further investigation.

ACKNOWLEDGMENTS

Open Access funding provided by the Qatar National Library.

REFERENCES

[1] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Boca Raton, FL, USA: CRC, 2020.
[2] C. Clason, "Regularization of inverse problems," 2020, *arXiv: 2001.00617*.
[3] C. Hegde, "Algorithmic aspects of inverse problems using generative models," in *Proc. IEEE 56th Annu. Allerton Conf. Commun., Control Comput.*, 2018, pp. 166–172.
[4] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 1, pp. 39–56, May 2020.
[5] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 537–546.
[6] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 9446–9454.
[7] Y. Gandelsman, A. Shocher, and M. Irani, "Double-DIP": Unsupervised image decomposition via coupled deep-image-priors," 2018, *arXiv:1812.00467*.
[8] G. Jagatap and C. Hegde, "High dynamic range imaging using deep image priors," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2020, pp. 9289–9293.
[9] A. Sagel, A. Roumy, and C. Guillemot, "Sub-dip: Optimization on a subspace with deep image prior regularization and application to superresolution," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2020, pp. 2513–2517.
[10] X. Han, Y. Liu, H. Yang, G. Xing, and Y. Zhang, "Normalization of face illumination with photorealistic texture via deep image prior synthesis," *Neurocomputing*, vol. 386, pp. 305–316, 2020.
[11] K. Gong, C. Catana, J. Qi, and Q. Li, "PET image reconstruction using deep image prior," *IEEE Trans. Med. Imag.*, vol. 38, no. 7, pp. 1655–1665, Jul. 2019.
[12] R. Heckel and P. Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," 2018, *arXiv:1810.03982*.
[13] M. Yin, Y. Zhang, X. Li, and S. Wang, "When deep fool meets deep prior: Adversarial attack on super-resolution network," in *Proc. 26th ACM Int. Conf. Multimedia*, 2018, pp. 1930–1938.
[14] R. Heckel, W. Huang, P. Hand, and V. Voroninski, "Deep denoising: Rate-optimal recovery of structured signals with a deep prior," 2018, *arXiv:1805.08855*.
[15] T. Yokota, K. Kawai, M. Sakata, Y. Kimura, and H. Hontani, "Dynamic PET image reconstruction using nonnegative matrix factorization incorporated with deep image prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3126–3135.
[16] Z. Cheng, M. Gadelha, S. Maji, and D. Sheldon, "A Bayesian perspective on the deep image prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5438–5446.
[17] G. Jagatap and C. Hegde, "Phase retrieval using untrained neural network priors," in *Proc. NeurIPS Workshop Deep Inverse Prog. Chairs*, 2019.
[18] O. Leong and W. Sakla, "Low shot learning with untrained neural networks for imaging inverse problems," 2019, *arXiv:1910.10797*.
[19] G. Jagatap and C. Hegde, "Algorithmic guarantees for inverse imaging with untrained network priors," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 14 832–14 842.
[20] F. Shamshad, A. Hanif, and A. Ahmed, "Subsampled fourier ptychography via pretrained invertible and untrained network priors," in *Proc. Deep Inverse Workshop Neural Inform. Process. Syst.*, 2019.
[21] D. Ren, K. Zhang, Q. Wang, Q. Hu, and W. Zuo, "Neural blind deconvolution using deep priors," 2019, *arXiv:1908.02197*.
[22] M. Gadelha, R. Wang, and S. Maji, "Shape reconstruction using differentiable projections and deep priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 22–30.
[23] M. Michelashvili and L. Wolf, "Audio denoising with deep network priors," 2019, *arXiv:1904.07612*.
[24] D. O. Baguer, J. Leuschner, and M. Schmidt, "Computed tomography reconstruction using deep image prior and learned reconstruction methods," *Inverse Problems*, vol. 36, no. 9, 2020, Art. no. 094004.
[25] A. Qayyum, W. Sultani, F. Shamshad, J. Qadir, and R. Tufail, "Single-shot retinal image enhancement using deep image priors," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Interv.*, 2020, pp. 636–646.
[26] F. Shamshad, A. Hanif, and A. Ahmed, "Subsampled fourier ptychography using pretrained invertible and untrained network priors," 2020, *arXiv:2005.07026*.
[27] F. Wang et al., "Phase imaging with an untrained neural network," *Light: Sci. Appl.*, vol. 9, no. 1, pp. 1–7, 2020.
[28] T. Uezato, D. Hong, N. Yokoya, and W. He, "Guided deep decoder: Unsupervised image pair fusion," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 87–102.
[29] S. Dittmer, T. Kluth, P. Maass, and D. O. Baguer, "Regularization by architecture: A deep prior approach for inverse problems," *J. Math. Imag. Vis.*, vol. 62, no. 3, pp. 456–470, 2020.
[30] F. Kong, F. Picetti, V. Lipari, P. Bestagini, and S. Tubaro, "Deep prior-based seismic data interpolation via multi-res U-Net," in *Proc. SEG Int. Expo. Annu. Meeting*, 2020, pp. 3159–3163.
[31] D. Ren, K. Zhang, Q. Wang, Q. Hu, and W. Zuo, "Neural blind deconvolution using deep priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 3338–3347.
[32] Y. Zheng, J. Li, Y. Li, J. Guo, X. Wu, and J. Chanussot, "Hyperspectral pansharpening using deep prior and dual attention residual network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8059–8076, Nov. 2020.
[33] E. Balevi, A. Doshi, and J. G. Andrews, "Massive MIMO channel estimation with an untrained deep neural network," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2079–2090, Mar. 2020.
[34] E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, "Deep phase decoder: Self-calibrating phase microscopy with an untrained deep neural network," *Optica*, vol. 7, no. 6, pp. 559–562, 2020.
[35] M. Z. Darestani and R. Heckel, "Accelerated MRI with untrained neural networks," 2020, *arXiv:2007.02471*.
[36] Z. Sun, F. Latorre, T. Sanchez, and V. Cevher, "A plug-and-play deep image prior," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2021, pp. 8103–8107.
[37] M. E. Arican, O. Kara, G. Bredell, and E. Konukoglu, "ISNAS-DIP: Image-specific neural architecture search for deep image prior," 2021, *arXiv:2111.15362*.
[38] B. V. Boas, W. Zirwas, and M. Haardt, "Transfer learning capabilities of untrained neural networks for MIMO CSI recreation," 2021, *arXiv:2111.07858*.
[39] S. Rey, S. Segarra, R. Heckel, and A. G. Marques, "Untrained graph neural networks for denoising," 2021, *arXiv:2109.11700*.
[40] B. V. Boas, W. Zirwas, and M. Haardt, "Machine learning for CSI recreation based on prior knowledge," 2021, *arXiv: 2111.07854*.
[41] J. Tang, J. Wu, J. Zhang, Z. Ren, J. Di, and J. Zhao, "Single-shot diffraction autofocusing: Distance prediction via an untrained physics-enhanced network," *IEEE Photon. J.*, vol. 14, no. 1, Feb. 2022, Art. no. 5207106.

[42] S. Pan, M. Liao, W. He, Y. Zhang, and X. Peng, "Untrained neural network for cryptanalysis of a phase-truncated-fourier-transform-based optical cryptosystem," *Opt. Exp.*, vol. 29, no. 26, pp. 42 642–42 649, 2021.

[43] Z. Li, X. Xiao, Y. Ren, N. Zhang, H. Shan, and W. Zhou, "Image dehazing with an untrained neural network," in *Proc. 5th Int. Conf. Electron. Inf. Technol. Comput. Eng.*, 2021, pp. 439–444.

[44] K. Monakhova, V. Tran, G. Kuo, and L. Waller, "Untrained networks for compressive lensless photography," *Opt. Exp.*, vol. 29, no. 13, pp. 20 913–20 929, 2021.

[45] F. Niknam, H. Qazvini, and H. Latifi, "Holographic optical field recovery using a regularized untrained deep decoder network," *Sci. Rep.*, vol. 11, no. 1, pp. 1–13, 2021.

[46] L. Chu, H. Pan, and W. Wang, "Unsupervised shape completion via deep prior in the neural tangent kernel perspective," *ACM Trans. Graph.*, vol. 40, no. 3, pp. 1–17, 2021.

[47] H. Lan, J. Zhang, C. Yang, and F. Gao, "Compressed sensing for photoacoustic computed tomography based on an untrained neural network with a shape prior," *Biomed. Opt. Exp.*, vol. 12, no. 12, pp. 7835–7848, 2021.

[48] Y. Sun, H. Zhao, and J. Scarlett, "On architecture selection for linear inverse problems with untrained neural networks," *Entropy*, vol. 23, no. 11, 2021, Art. no. 1481.

[49] A. Qayyum, W. Sultani, F. Shamshad, R. Tufail, and J. Qadir, "Single-shot retinal image enhancement using untrained and pretrained neural networks priors integrated with analytical image priors," *Comput. Biol. Med.*, vol. 148, 2022, Art. no. 105879.

[50] Z. Meng, Z. Yu, K. Xu, and X. Yuan, "Self-supervised neural networks for spectral snapshot compressive imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 2602–2611.

[51] S. Kafle, G. Joseph, and P. K. Varshney, "One-bit compressed sensing using untrained network prior," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2021, pp. 2875–2879.

[52] S. Baek, M. Song, J. Jang, G. Kim, and S.-B. Paik, "Face detection in untrained deep neural networks," *Nature Commun.*, vol. 12, no. 1, pp. 1–15, 2021.

[53] G. Kim, J. Jang, S. Baek, M. Song, and S.-B. Paik, "Visual number sense in untrained deep neural networks," *Sci. Adv.*, vol. 7, no. 1, 2021, Art. no. eabd6127.

[54] C. Bai, T. Peng, J. Min, R. Li, Y. Zhou, and B. Yao, "Dual-wavelength in-line digital holography with untrained deep neural networks," *Photon. Res.*, vol. 9, no. 12, pp. 2501–2510, 2021.

[55] C. Gallicchio and S. Scardapane, "Deep randomized neural networks," in *Recent Trends in Learning From Data*. Berlin, Germany: Springer, 2020, pp. 43–68.

[56] A. Wirgin, "The inverse crime," *math-ph/0401050*, 2004.

[57] D. L. Colton, R. Kress, and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, vol. 93. Berlin, Germany: Springer, 1998.

[58] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1647–1654, Sep. 2007.

[59] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, and A. K. Katsaggelos, "SoftCuts: A soft edge smoothness prior for color image super-resolution," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 969–981, May 2009.

[60] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Variational Bayesian blind deconvolution using a total variation prior," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 12–26, Jan. 2009.

[61] W. Hu, Z. Fu, and Z. Guo, "Local frequency interpretation and non-local self-similarity on graph for point cloud inpainting," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 4087–4100, Aug. 2019.

[62] H. Choi, J. Romberg, R. Baraniuk, and N. Kingsbury, "Hidden Markov tree modeling of complex wavelet transforms," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2000, pp. 133–136.

[63] M.-A. Charmi, S. Derrode, and F. Ghorbel, "Fourier-based geometric shape prior for snakes," *Pattern Recognit. Lett.*, vol. 29, no. 7, pp. 897–904, 2008.

[64] J. Holloway et al., "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Trans. Comput. Imag.*, vol. 2, no. 3, pp. 251–265, Sep. 2016.

[65] C. B. Paschal and H. D. Morris, "K-space in the clinic," *J. Magn. Reson. Imag.: Official J. Int. Soc. Magn. Reson. Med.*, vol. 19, no. 2, pp. 145–159, 2004.

[66] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.

[67] S. Hong and S. Kim, "Deep matching prior: Test-time optimization for dense correspondence," 2021, *arXiv:2106.03090*.

[68] K. Ho, A. Gilbert, H. Jin, and J. Collomosse, "Neural architecture search for deep image prior," 2020, *arXiv:2001.04776*.

[69] Y.-C. Chen, C. Gao, E. Robb, and J.-B. Huang, "NAS-DIP: Learning deep image prior with neural architecture search," 2020, *arXiv:2008.11713*.

[70] S. Liu, S. Miao, J. Su, B. Li, W. Hu, and Y.-D. Zhang, "UMAG-Net: A new unsupervised multiattention-guided network for hyperspectral and multispectral image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7373–7385, 2021.

[71] S. Dittmer and P. Maass, "A projectional ansatz to reconstruction," 2019, *arXiv:1907.04675*.

[72] Z. Sun, "Solving inverse problems with hybrid deep image priors: The challenge of preventing overfitting," 2020, *arXiv:2011.01748*.

[73] Y. Jo, S. Y. Chun, and J. Choi, "Rethinking deep image prior for denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 5087–5096.

[74] J. Liu, Y. Sun, X. Xu, and U. S. Kamilov, "Image restoration using total variation regularized deep image prior," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2019, pp. 7715–7719.

[75] G. Mataev, P. Milanfar, and M. Elad, "DeepRED: Deep image prior powered by red," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019.

[76] M. Asim, F. Shamshad, and A. Ahmed, "Patchdip exploiting patch redundancy in deep image prior for denoising," in *Proc. NeurIPS Workshop Deep Inverse Prog. Chairs*, 2019.

[77] G. Ghiasi, T.-Y. Lin, and Q. V. Le, "NAS-FPN: Learning scalable feature pyramid architecture for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7029–7038.

[78] X. Gong, S. Chang, Y. Jiang, and Z. Wang, "AutoGAN: Neural architecture search for generative adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3223–3233.

[79] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," 2016, *arXiv:1611.01578*.

[80] T. Yokota, H. Hontani, Q. Zhao, and A. Cichocki, "Manifold modeling in embedded space: A perspective for interpreting "deep image prior"," 2019, *arXiv:1908.02995*.

[81] P. Chakrabarty and S. Maji, "The spectral bias of the deep image prior," 2019, *arXiv:1912.08905*.

[82] R. Segawa, H. Hayashi, and S. Fujii, "Proposal of new activation function in deep image prior," *IEEJ Trans. Elect. Electron. Eng.*, vol. 15, no. 8, pp. 1248–1249, 2020.

[83] S. Fujii and H. Hayashi, "Comparison of performance by activation functions on deep image prior," in *Proc. Int. Conf. Artif. Intell. Inf. Commun.*, 2019, pp. 255–258.

[84] C. A. Metzler, A. Mousavi, R. Heckel, and R. G. Baraniuk, "Unsupervised learning with Stein's unbiased risk estimator," 2018, *arXiv:1805.10531*.

[85] R. Heckel and M. Soltanolkotabi, "Denoising and regularization via exploiting the structural bias of convolutional generators," 2019, *arXiv:1910.14634*.

[86] R. Heckel and M. Soltanolkotabi, "Compressive sensing with untrained neural networks: Gradient descent finds the smoothest approximation," 2020, *arXiv:2005.03991*.

[87] W. Shang, P. Zhu, and D. Ren, "Semi-supervised learning to remove fences from a single image," in *Proc. Chin. Conf. Pattern Recognit. Comput. Vis.*, 2020, pp. 79–90.

[88] S. Tian and T. Xiong, "A generic solver combining unsupervised learning and representation learning for breaking text-based CAPTCHAs," in *Proc. Web Conf.*, 2020, pp. 860–871.

[89] S. Kim, H. RahmaniKhezri, S. M. Nourbakhsh, and M. Hefeeda, "Unsupervised single-image reflection separation using perceptual deep image priors," 2020, *arXiv:2009.00702*.

[90] P. Chandramouli and K. V. Gandikota, "Blind single image reflection suppression for face images using deep generative priors," in *Proc. IEEE Conf. Comput. Vis. Workshops*, 2019, pp. 3315–3323.

[91] C. Lei, X. Huang, M. Zhang, Q. Yan, W. Sun, and Q. Chen, "Polarized reflection removal with perfect alignment in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 1747–1755.

[92] E. Lu et al., "Layered neural rendering for retiming people in video," 2020, *arXiv:2009.07833*.

[93] G. Jagatap and C. Hegde, "High dynamic range imaging using deep image priors," *IEEE Int. Conf. Acoustics, Speech Signal Process.*, pp. 9289–9293, 2020.

[94] P. Cascarano, A. Sebastiani, and M. C. Comes, "ADMM-DIPTV: Combining total variation and deep image prior for image restoration," 2020, *arXiv:2009.11380*.

[95] S. Zou, M. Long, X. Wang, X. Xie, G. Li, and Z. Wang, "A CNN-based blind denoising method for endoscopic images," in *Proc. IEEE Biomed. Circuits Syst. Conf.*, 2019, pp. 1–4.

[96] Y. Park, S. Lee, B. Jeong, and J. Yoon, "Joint demosaicing and denoising based on a variational deep image prior neural network," *Sensors*, vol. 20, no. 10, 2020, Art. no. 2970.

[97] Y. Feng, Y. Shi, and D. Sun, "Blind poissonian image deblurring regularized by a denoiser constraint and deep image prior," *Math. Problems Eng.*, vol. 2020, 2020, Art. no. 9483521.

[98] J. Zukerman, T. Tirer, and R. Giryes, "BP-DIP: A backprojection based deep image prior," 2020, *arXiv:2003.05417*.

[99] J. Nie, L. Zhang, C. Wang, W. Wei, and Y. Zhang, "Robust deep hyperspectral imagery super-resolution," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 847–850.

[100] X. Ma, Y. Hong, and Y. Song, "Super resolution land cover mapping of hyperspectral images using the deep image prior-based approach," *Int. J. Remote Sens.*, vol. 41, no. 7, pp. 2818–2834, 2020.

[101] T. Zhang, Y. Fu, L. Wang, and H. Huang, "Hyperspectral image reconstruction using deep external and internal learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8558–8567.

[102] J. Nie, L. Zhang, W. Wei, C. Ding, and Y. Zhang, "Unsupervised deep hyperspectral super-resolution with unregistered images," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2020, pp. 1–6.

[103] O. Voynov et al., "Perceptual deep depth super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5652–5662.

[104] T. Weber, H. Hußmann, Z. Han, S. Matthes, and Y. Liu, "Draw with me: Human-in-the-loop for image restoration," in *Proc. 25th Int. Conf. Intell. User Interfaces*, 2020, pp. 243–253.

[105] P. Ghosh, V. Vineet, L. S. Davis, A. Shrivastava, S. Sinha, and N. Joshi, "Deep depth prior for multi-view stereo," 2020, *arXiv:2001.07791*.

[106] H. Zhang, L. Mai, N. Xu, Z. Wang, J. Collomosse, and H. Jin, "An internal learning approach to video inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2720–2729.

[107] J. Yu and R. Ramamoorthi, "Robust video stabilization by optimization in CNN weight space," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3795–3803.

[108] C. Lei, Y. Xing, and Q. Chen, "Blind video temporal consistency via deep video prior," *Proc. Int. Conf. Neural Inf. Process. Syst.*, vol. 33, pp. 1083–1093, 2020.

[109] C. Liu, X. Li, L. Zhuo, J. Li, and Q. Zhou, "A novel speckle noise reduction algorithm for old movies recovery," in *Proc. IEEE 11th Int. Congr. Image Signal Process. Biomed. Eng. Informat.*, 2018, pp. 1–6.

[110] T. Suzuki, "Superpixel segmentation via convolutional neural networks with regularized information maximization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2020, pp. 2573–2577.

[111] F. Yang, T.-A. Pham, N. Brandenberg, M. P. Lutolf, J. Ma, and M. Unser, "Robust phase unwrapping via deep image prior for quantitative phase imaging," 2020, *arXiv:2009.11554*.

[112] L. Boominathan, M. Maniparambil, H. Gupta, R. Baburajan, and K. Mitra, "Phase retrieval for fourier ptychography under varying amount of measurements," 2018, *arXiv:1805.03593*.

[113] Z. Wang, Z. Wang, Q. Li, and H. Bilen, "Image deconvolution with deep image and kernel priors," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2019, pp. 980–989.

[114] A. Kattamis, T. Adel, and A. Weller, "Exploring properties of the deep image prior," in *Proc. Deep Inverse Workshop Neural Inform. Process. Syst.*, 2019.

[115] T. Dai et al., "DIPDefend: Deep image prior driven defense against adversarial examples," in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 1404–1412.

[116] R. E. Sutanto and S. Lee, "Adversarial attack defense based on the deep image prior network," in *Information Science and Applications*. Berlin, Germany: Springer, 2020, pp. 519–526.

[117] A. Gandhi and S. Jain, "Adversarial perturbations fool deepfake detectors," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2020, pp. 1–8.

[118] T. Gittings, S. Schneider, and J. Collomosse, "Robust synthesis of adversarial visual examples using a deep image prior," 2019, *arXiv:1907.01996*.

[119] V. Narayanaswamy, J. J. Thiagarajan, and A. Spanias, "Using deep image priors to generate counterfactual explanations," 2020, *arXiv:2010.12046*.

[120] L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun, "Hands-on Bayesian neural networks—A tutorial for deep learning users," *ACM Comput. Surv.*, vol. 17, no. 2, pp. 29–48, 2022.

[121] K. C. Zhou et al., "Mesoscopic photogrammetry with an unstabilized phone camera," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 7531–7541.

[122] T. M. Wong, H. Bauermeister, M. Kahl, P. H. Bolívar, M. Möller, and A. Kolb, "Deep optimization prior for THz model parameter estimation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2022, pp. 4049–4058.

[123] D. V. Veen, A. Jalal, M. Soltanolkotabi, E. Price, S. Vishwanath, and A. G. Dimakis, "Compressed sensing with deep image prior and learned regularization," 2018, *arXiv:1806.06438*.

[124] J. Ren, J. Liang, and Y. Zhao, "Soil PH measurement based on compressive sensing and deep image prior," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 4, no. 1, pp. 74–82, Feb. 2020.

[125] M. Ferrari, O. Taran, T. Holotyak, K. Egiazarian, and S. Voloshynovskiy, "Injecting image priors into learnable compressive subsampling," in *Proc. IEEE 26th Eur. Signal Process. Conf.*, 2018, pp. 1735–1739.

[126] M. Z. Darestani, A. Chaudhari, and R. Heckel, "Measuring robustness in deep learning based compressive sensing," 2021, *arXiv:2102.06103*.

[127] B. Sun, C. Mu, Z. Wu, and X. Zhu, "Training-free deep generative networks for compressed sensing of neural action potentials," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Apr. 08, 2021, doi: 10.1109/TNNLS.2021.3069436.

[128] I. D. Mastan and S. Raman, "Multi-level encoder-decoder architectures for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 1728–1737.

[129] X. Jin, L. Zhang, C. Shan, X. Li, and Z. Chen, "Dual prior learning for blind and blended image restoration," *IEEE Trans. Image Process.*, vol. 31, pp. 1042–1056, 2022.

[130] A. Petrovskaia, R. B. Jana, and I. V. Oseledets, "A single image deep learning approach to restoration of corrupted remote sensing products," *Proc. Int. Conf. Learn. Representations*, 2020.

[131] S. Wizadwongsa, P. Phongthawee, J. Yenphraphai, and S. Suwajanakorn, "NeX: Real-time view synthesis with neural basis expansion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 8530–8539.

[132] V. Saragadam, A. Dave, A. Veeraraghavan, and R. G. Baraniuk, "Thermal image processing via physics-inspired deep networks," in *Proc. IEEE Int. Conf. Comput. Vis. Workshop*, 2021, pp. 4040–4048.

[133] H. Laga, L. V. Jospin, F. Boussaid, and M. Bennamoun, "A survey on deep learning techniques for stereo-based depth estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 1738–1764, Apr. 2022.

[134] L. Liu, L. Fu, and M. Zhang, "Deep-seismic-prior-based reconstruction of seismic data using convolutional neural networks," *Geophysics*, vol. 86, no. 2, pp. 1–93, 2020.

[135] M. J. Park, J. Jennings, B. Clapp, and B. Biondi, "Seismic data interpolation using a POCS-guided deep image prior," in *Proc. SEG Tech. Prog. Expanded Abstr.*, 2020, pp. 3154–3158.

[136] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial examples in the physical world," 2016, *arXiv:1607.02533*.

[137] N. Carlini and D. Wagner, "Towards evaluating the robustness of neural networks," in *Proc. IEEE Symp. Secur. Privacy*, 2017, pp. 39–57.

[138] T. Dai, Y. Feng, B. Chen, J. Lu, and S.-T. Xia, "Deep image prior based defense against adversarial examples," *Pattern Recognit.*, vol. 122, 2022, Art. no. 108249.

[139] K. H. Jin, H. Gupta, J. Yerly, M. Stuber, and M. Unser, "Time-dependent deep image prior for dynamic MRI," 2019, *arXiv:1910.01684*.

[140] M. Z. Darestani and R. Heckel, "Can un-trained networks compete with trained ones for accelerated MRI?," in *Proc. 29th Annu. Meeting ISMRM*, 2021.

[141] D. Zhao, F. Zhao, and Y. Gan, "Reference-driven compressed sensing MR image reconstruction using deep convolutional neural networks without pre-training," *Sensors*, vol. 20, no. 1, 2020, Art. no. 308.

[142] K. Gong, C. Catana, J. Qi, and Q. Li, "Direct Patlak reconstruction from dynamic PET using unsupervised deep learning," in *Proc. 15th Int. Meeting Fully Three-Dimensional Image Reconstruction Radiol. Nucl. Med.*, 2019, Art. no. 110720R.

[143] K. C. Zhou and R. Horstmeyer, "Diffraction tomography with a deep image prior," *Opt. Exp.*, vol. 28, no. 9, pp. 12 872–12 896, 2020.

[144] K. Gong, K. Kim, D. Wu, M. K. Kalra, and Q. Li, "Low-dose dual energy CT image reconstruction using non-local deep image prior," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, 2019, pp. 1–2.

[145] S. Dittmer, T. Kluth, M. T. R. Henriksen, and P. Maass, "Deep image prior for 3D magnetic particle imaging: A quantitative comparison of regularization techniques on open MPI dataset," 2020, *arXiv:2007.01593*.
[146] T. Vu et al., "Deep image prior for sparse-sampling photoacoustic microscopy," 2020, *arXiv:2010.12041*.
[147] F. Hashimoto, K. Ote, and H. Tsukada, "Dynamic PET image denoising using deep convolutional neural network without training datasets," *J. Nucl. Med.*, vol. 60, pp. 242–242, 2019.
[148] J. Cui et al., "CT-guided PET image denoising using deep neural network without prior training data," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Proc.*, 2018, pp. 1–3.
[149] S. Han, J. L. Prince, and A. Carass, "Inhomogeneity correction in magnetic resonance images using deep image priors," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2020, pp. 404–413.
[150] W. Fan, H. Yu, T. Chen, and S. Ji, "OCT image restoration using non-local deep image prior," *Electronics*, vol. 9, no. 5, 2020, Art. no. 784.
[151] K. Hagan, D. Li, and J. Loo, "Noise reduction in optical coherence tomography using deep image prior,".
[152] T. Vu et al., "Deep image prior for undersampling high-speed photoacoustic microscopy," *Photoacoustics*, vol. 22, 2021, Art. no. 100266.
[153] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2016, pp. 424–432.
[154] F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, and Y. Ouchi, "4D deep image prior: Dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network," *Phys. Med. Biol.*, vol. 66, no. 1, 2021, Art. no. 015006.
[155] N. J. Tustison et al., "N4ITK: Improved N3 bias correction," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, Jun. 2010.
[156] S. A. Hussein, T. Tirer, and R. Giryes, "Image-adaptive GAN based reconstruction," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 3121–3129.
[157] S. Athar, E. Burnaev, and V. Lempitsky, "Latent convolutional models," 2018, *arXiv:1806.06284*.
[158] R. Barbano et al., "Is deep image prior in need of a good education?," 2021, *arXiv:2111.11926*.

Adnan Qayyum received the bachelor's degree in electrical (computer) engineering from the COMSATS Institute of Information Technology, Wah, Pakistan, in 2014, and the MS degree in computer engineering (signal and image processing) from the University of Engineering and Technology, Taxila, Pakistan, in 2016. He is currently working toward the PhD degree with the Information Technology University (ITU), Lahore, Pakistan. His research interests include inverse medical imaging problems, healthcare, and secure and robust ML.

Inaam Ilahi received the bachelor's degree in electrical engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, and the MS degree in computer science from Information Technology University (ITU), Lahore, Pakistan, in 2020. He is a research assistant with Information Technology University, Lahore, Pakistan. His research interests include reinforcement learning, adversarial ML, and applications of deep reinforcement learning in networks.

Fahad Shamshad received the BS degree in electrical engineering from the Institute of Space Technology (IST), Islamabad, and the MS degree from the National University of Sciences and Technology (NUST). He is senior machine learning engineer with OMNO AI. He worked as a research associate with Information Technology University, Lahore. His research interests include inverse imaging problems, generative models, and optimization theory.

Farid Boussaid received the MS and PhD degrees in microelectronics from the National Institute of Applied Science (INSA), Toulouse, France, in 1996 and 1999, respectively. He joined Edith Cowan University, Perth, Australia, as a postdoctoral research fellow, and a member of the Visual Information Processing Research Group in 2000. He joined the University of Western Australia, Crawley, Australia, in 2005, where he is currently a professor. His current research interests include neuromorphic engineering, smart sensors, and machine learning.

Mohammed Bennamoun is Winthrop professor with the Department of Computer Science and Software Engineering, University of Western Australia (UWA) and is a researcher in computer vision, machine/deep learning, robotics, and signal/speech processing. He has published four books (available on Amazon), one edited book, one Encyclopedia article, 14 book chapters, 150+ journal papers, 260+ conference publications, 16 invited and keynote publications. His h-index is 77 and his number of citations is 13,500+ (Google Scholar). He was Awarded 70+ competitive research grants, from the Australian Research Council, and numerous other Government, UWA and industry Research Grants. He successfully supervised +26 PhD students to completion. He won the Best Supervisor of the Year Award at Queensland University of Technology (1998), and received Award for research supervision at UWA (2008 and 2016) and Vice-Chancellor Award for mentorship (2016). He delivered conference tutorials at major conferences, including: IEEE CVPR 2016, Interspeech 2014, IEEE ICASSP, and ECCV. He was also invited to give a Tutorial at an International Summer School on Deep Learning (DeepLearn 2017).

Junaid Qadir (Senior Member, IEEE) is a professor of computer engineering with Qatar University, Doha, Qatar. Previously, he was the director of the IHSAN Research Lab and a professor with the Electrical Engineering Department, Information Technology University (ITU), Lahore, Pakistan. He has published more than 150 peer-reviewed articles at various high-quality research venues including more than 50 impact-factor journal publications at top international journals including the *IEEE Communication Magazine*, *IEEE Journal on Selected Areas in Communication* (JSAC), *IEEE Communications Surveys and Tutorials* (CST), and *IEEE Transactions on Mobile Computing* (TMC). He was Awarded Best University Teacher Award—for the year 2012-2013 by higher education commission of Pakistan. He has been appointed as ACM distinguished speaker for a three-year term starting from 2020. He is a senior member of the ACM.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/csdl.