Untrained Neural Network Priors for Inverse Imaging Problems: A Survey

Adnan Qayyum¹, Inaam Ilahi², Fahad Shamshad², Farid Boussaid², Mohammed Bennamoun², and Junaid Qadir²

¹Information Technology University Punjab
²Affiliation not available

October 30, 2023
Untrained Neural Network Priors for Inverse Imaging Problems: A Survey
Adnan Qayyum, Inaam Ilahi, Fahad Shamshad, Farid Boussaid, Mohammed Bennamoun, and Junaid Qadir,

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 handcrafted approaches. Traditionally, analytical methods have been used to solve 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 covering more than 140 recent papers of such studies and various applications of UNNP for different tasks and elaborate upon 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 there exist multiple solutions for the same IIP. Reconstructing a unique solution that fits these observations is 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 the IIPs including denoising, super-resolution, inpainting, image deblurring, MRI reconstruction, and many more [2], [3], [4]. These priors are nothing more than regularization techniques.

Conventional algorithms to solve the IIPs are based on the simple mathematical models that are hand-crafted from the domain knowledge. These domain knowledge-based IIP recovery algorithms carry out the inference based on knowledge of the underlying forward model associated to 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) [5].

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 [6], [7]. 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. The DL-based approaches have shown to outperform their hand-crafted counterparts on numerous imaging reconstruction tasks and also with a much faster inference, making these approaches suitable for real-time imaging applications. In particular, convolutional neural networks (CNNs) have been shown to be very effective for image processing and computer vision applications [8]. 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 domains like medical imaging. 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 voluminous reliable labeled data.

To bridge the gap between conventional hand-crafted priors (which suffer from a poor discriminative capability) and DL-based approaches (which require large scale labeled datasets), untrained neural network priors (UNNPs) have recently emerged as a promising line of research. These priors have shown to outperform the conventional hand-crafted priors and offer a comparable performance compared to their DL-based counterparts, while integrating the knowledge of the forward model without requiring massive labeled datasets [9]. Specifically, these priors assume that

- A. Qayyum, I. Ilahi, F. Shamshad, and J. Qadir are with the Information Technology University (ITU) of Punjab, Lahore, Pakistan
- F. Boussaid and M. Bennamoun are with the University of Western Australia, Perth, Australia
plenty of 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 inverse image problems especially in medical imaging where labeled datasets are hard to obtain. 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 in the IIPs. We note here that Ulyanov et al. [9] named the term UNNPs as deep image prior (DIP) and since then this alternative term has been largely used in the literature.

Contributions of this paper: 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 UNNP to inverse imaging problems. We present a comprehensive review of more than 140 papers that appeared since the emergence of UNNP, in the leading conferences and journals, thus covering the most recent and advanced progress in the field.

2) We provide a detailed coverage of different applications of UNNP for different IIPs. We specifically categorized them into two groups, i.e., general and medical IIPs. Our focus on medical applications stems from the ethical constraints and privacy concerns, which make the field struggle with 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 UNNP and provide insights into the possible future directions for the new researchers in this field.

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

2 BACKGROUND

In this section, we briefly discuss the background related to inverse imaging problems and the commonly used models to model inverse imaging problems, i.e., handcrafted priors, sparse coding and dictionary learning based methods, DL-based methods, and untrained neural networks. Finally, we will provide a brief overview and cover the necessary background related to untrained neural networks.

2.1 Inverse Imaging Problems

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) $x_0 \in \mathbb{R}^n$ is observed via $m$ noisy measurements $y \in \mathbb{R}^m$. More specifically, we can write

$$y = A(x_0) + \eta,$$

where $A$ is the forward measurement operator, i.e., the physical model underlying the measurement process, and $\eta$ is noise perturbation modelled by $\mathcal{N}(\mu, \sigma^2)$. Determining the output $y$ for a given input $x_0$ for the known forward operator $A$ represents the forward model. Similarly, finding the input $x_0$ for a given output $y$ and the operator $A$ represents the solution to the inverse problem. Inverse problems that typically arise in imaging include compressive sensing, image deblurring, super-resolution, denoising, etc. 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 $x_0$ 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

$$\arg\min_x \frac{1}{2}||y - A(x)||^2 + \lambda R(x),$$

where $R(x)$ is the regularization term (also known as image prior), $||.||^2$ is the $L_2$ norm, and $\lambda$ is the regularization weight (hyperparameter). In Eq. (2), the first term is the data fidelity 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. 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 forward operator in a typical image processing pipeline is shown in Figure 1. A summary of forward operators used in different inverse problems is presented in Table 1. As shown in that table, the inverse imaging problems can be of different types, e.g., inpainting, denoising, super-resolution, and compressed sensing as depicted in Figure 2.

2.2 Methods for Solving Inverse Imaging Problems

In this section, we briefly describe different methods that are usually used for modeling natural images namely, handcrafted prior-based methods, sparse coding or dictionary learning, DL-based methods, and UNNP-based methods, which are shown in Figure 3.

2.2.1 Handcrafted Priors

One natural approach to model natural images is to view them as sparse to a particular handcrafted basis, e.g., wavelets like basis, as shown in Figure 3(a). This type of method does not require training data and it assumes that
**Fig. 1:** Typical image processing pipeline. In a forward model, the transformation $A(.)$ is applied to an input image. The inverse problem aims to obtain an estimate of the input image from the observation via a reconstruction algorithm by leveraging the prior knowledge about the true (target) image.

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

**Fig. 2:** Examples of different inverse imaging problems (IIPs).

The image being modeled is in the linear span of a few elements in a space and is not necessarily the best representation, i.e., it is a linear model. Such image reconstruction methods exploit some prior knowledge about the true image $x$ such as sparsity, smoothness, total variation, non-local self similarity, Markov-tree models on wavelet coefficients, geometric properties etc. The reconstruction amounts to finding a solution $\hat{x}$ that is a good fit to the observations $y$ 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 described above most of the handcrafted methods are linear). As a consequence, a large majority of unnatural signals also satisfy the constraints specified by these hand-designed priors.

### 2.2.2 Sparse Coding or Dictionary Learning

In sparse coding or dictionary learning approach, we use the training data to learn a few images that can be used to recover a particular desired image as shown in Figure 3 (b). Like the handcrafted basis method, this approach is also a linear model. More information about using sparsity and dictionary learning to solve inverse imaging problems can be found in [10], [11], [12].

### 2.2.3 Deep Learning for Solving Inverse Problems

For a given image modeling task, we typically train DL models using a training dataset in an end-to-end fashion, e.g., we can use generative adversarial networks (GANs) for image segmentation and reconstruction, as shown in Figure 3 (d). Recently, DL-based approaches have led to the state of the art results in solving various inverse imaging problems at the expense of being trained on massive labelled datasets, e.g., see Figure 4. 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 $y$ to $x$ without any knowledge of the forward operator $A$ during the training process. More specifically, these approaches aim to minimize the objective
function $\|x - F_\theta(y)\|$, where $F_\theta(\cdot)$ denotes a deep neural network having parameters (weights) denoted as $\theta$. The general principle is that, given enough training data, one should be able to learn everything one needs to know about $A$ to successfully estimate $y$. This straightforward approach has shown impressive results on numerous inverse imaging problems including superresolution, blind image deblurring, magnetic resonance imaging, and numerous other inverse imaging problems. We refer the interested reader to the comprehensive review of DL-based approaches for inverse imaging problems [7]. DL-based approaches can effectively circumvent the limited discriminative capability of the handcrafted 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 inverse imaging problems and different parameter settings of the forward model. This means even a slight change in the forward acquisition model such as noise level (e.g., for denoising) or sampling rate (e.g., in compressed sensing) requires costly retraining of these DL models. Note that apart from the supervised learning, DL-based approaches have also been used in unsupervised setups to solve numerous inverse imaging problems. Examples include generative models, and unrolled optimization-based approach [14].

2.2.4 Untrained Neural Networks Priors (UNNPs)

Instead of training DL models on a large-scale datasets, untrained DL models are used to model natural images. More specifically, the DL model (initialized with random parameters) is given a random noise as input and is tasked to model a particular natural image using the random input and its parameters (Figure 3 (c)). The parameters of the network are 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 UNN 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 [9] 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 network itself 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
Fig. 5: Applications of UNNP in numerous inverse imaging problems along with baseline approaches. It can be seen that UNNP based approaches give superior qualitative performance compared to handcrafted based methods such as NLM and competitive performance compared to DL-based approaches. Individual figures sources: (a) and (b) from [9], (c) from [15], (d) from [16], (e) from [17], and (f) from [18].

 datasets. Instead, UNNP can operate on a single (degraded) image. Furthermore, the authors demonstrated that UNN has high impedance to noise and low impedance to the true image. 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 random parameters were optimized to maximize their likelihood given a task-specific observation model and the specific degraded image. A performance comparison of UNNP with baseline approaches is presented in Figure 5. It can be seen that UNNP based approaches give superior qualitative performance compared to handcrafted based methods such as NLM and competitive performance compared to DL-based approaches. The mathematical expression for solving inverse problems is given by energy minimization of the type

\[
\hat{x} = \arg\min_x E(x; x_0) + R(x),
\]

where, \( E \) is task-specific function, \( x_0 \) and \( x \) denote the input and the generated images, respectively. \( R(x) \) denotes the regularizer term. UNNP removes the explicit regularization term \( R(x) \) by assuming that the unknown image \( x \)
should be an image generated from the generator network such that

\[ \hat{\theta} = \arg\min_{\theta} E(f_{\theta}(z); x_0), \quad \hat{x} = f_{\theta}(z), \quad (4) \]

where, \( E(f_{\theta}(z); x_0) \) is a task specific term (e.g., denoising, deblurring, etc.), \( x_0 \) and \( x \) represent the degraded image and the restored image, respectively. \( f_{\theta}(z) \) is the generator function that maps the random code vector \( z \) to an input image \( x \), i.e., \( x = f_{\theta}(z) \). A suitable optimizer such as gradient descent is used for obtaining the minimizer \( \theta \) from randomly initialized neural network parameters and the output of the image reconstruction (i.e., restoration) process for a given \( \theta^* \) is given by \( \hat{x} = f_{\theta}(z) \). The iterative process used by UNNP to recover the input image is depicted in Figure 6. UNNP starts the reconstruction process from iteration 0 using random weights \( \theta_0 \) and iteratively update them to minimize the task-specific objective (i.e., Eq. 4). In each iteration, the parameters are mapped to an image \( x = f(z) \), where \( z \) is a fixed code vector. DNNs perform the mapping using its parameters. The task-dependent loss \( E(x; x_0) \) is computed using the image \( x \). 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 UNNP at different iterations (i.e., 0, 50, 250, and 100) for different inverse imaging tasks is shown in Figure 7.

![Fig. 6: Illustration of different iterations of untrained neural network priors (UNNP) for denoising input image \( x_0 \) over different iterations (Figure adapted from [9]).](image)

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 UNNP. A number of works have been recently proposed to circumvent the issue of early stopping. These works include the deep decoder [19] and a Bayesian alternative of UNNP [20]. 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 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. [21] demonstrated that the projectional approach to UNNP alleviates the need for early stopping and provides a valid and plausible reconstruction, which is achieved by \( f_{\theta}(z) \) after the minimization for reconstruction. Similarly, the use of hybrid deep priors for alleviating the problem of overfitting in UNNP is explored in [22]. The authors proposed two algorithms to incorporate an implicit prior (such as denoising algorithm) or explicit prior (such as total variation (TV)) with UNNP to avoid overfitting.

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 [23], 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

\[ \hat{\theta} = \arg\min_{\theta} E(f_{\theta}(z); x_0) + \lambda TV(f_{\theta}(z)), \]

\[ \text{s.t.} \quad x^* = f_{\theta^*}(z), \quad (5) \]

where, \( TV \) is the total variation regularizer, which is controlled by the multiplying factor \( \lambda \).

In [24], 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 UNNP. The objective function of their proposed approach is given below.

\[ x^*, \theta^* = \arg\min_{x, \theta} \|P_n(y) - f_{\theta}(z)\|, \quad (6) \]

\[ \text{s.t.} \quad x = f_{\theta}(z), \]

Similarly, in [25], 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

\[ \theta^* = \arg\min_{\theta} \sum_{n=1}^{N} |P_n(y) - f_{\theta}(z)|, \quad (7) \]

\[ \text{s.t.} \quad x = f_{\theta}(z), \]

where operator \( P_n(\cdot) \) extracts the \( n^{\text{th}} \) patch from the input image and \( x_n \) is the denoised \( n^{\text{th}} \) patch. The final
estimate of the image is reconstructed by combining all the denoised patches, \( \hat{x} = \hat{P}(\hat{x}_1, \hat{x}_2, ..., \hat{x}_n) \), where \( \hat{P}(\cdot) \) is the function that reconstructs the image back from its patches. Similarly, the use of Tikhonov functionals rather than deep networks was proposed in [26]. 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. [27] 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 and these optimized network architectures have been shown to enhance the performance of classic UNNP methods. Instead of using handcrafted neural networks for UNNP, Chen et al. [28] 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 [29], [30], [31], which involve the search of optimal neural networks that give the top performance on large datasets. Their approach differs from NAS-FPN (proposed in [29]) as the target here is the recovery of feature maps with higher spatial resolution in the decoder, whereas the aim in [29] is to learn pyramidal feature representations for object detection. Their network design is based upon the standard U-Net architecture and 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 [32]. 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 network architecture search by proposing a Bayesian perspective on UNNP [20]. On the other hand, Uezato et al. [32] argued that the network architecture of UNNP [9] 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 [33]. Chakrabarty et al. [34] proposed a method to analyze trajectories generated by 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. A systematic study focused on highlighting neural architecture search challenges and their potential solutions can be found in [35].
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. [36] 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 [37]. 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. [38] evaluated the performance of the UNNP framework for the denoising task using their proposed Stein’s Unbiased Risk Estimator (SURE) loss instead of using ℓ2 loss.

3.5 Coupling Multiple UNNPs

The multiple UNNP concept is based on the fact that various computer vision tasks aim to decompose an image into its components. More specifically, the authors in [39] proposed an unsupervised Deep framework to 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 such 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.

3.5.1 Image Decomposition

In [39], 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, 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 infinite 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.

For fence removal from an image, Shang et al. [40] proposed a supervised learning-based method to detect the mask of the fences and then employed unsupervised learning for the robust restoration of the background image free from the fences. They used three UNNP networks to model the three different parts of an image, i.e., background image, fences layer, and fences mask. Furthermore, they proposed to adopt the Laplacian smoothness loss function to refine the detected fences mask. This makes the restoration of the background image more robust to detection errors of the fence mask. Their proposed algorithm can generate a more visually plausible background image than the previous state-of-the-art techniques.

Text-based CAPTCHA is the most popular security mechanism used to prevent malicious attacks on websites. Over the past few years, several mechanisms have been proposed to highlight the vulnerabilities of these CAPTCHA based mechanisms. To solve CAPTCHA, Tian et al. [41] proposed to integrate representation learning and supervised learning together with noisy background removal using double-UNNP.

3.5.2 Blind Image Deblurring

Blind deconvolution remains challenging in many real-world applications. The use of fixed and handcrafted priors in traditional maximum a posteriori (MAP) based methods leads to an 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 [42], the authors extended the idea of Double UNNP for blind image deblurring by employing two untrained UNNPs, one to produce the clean image and one for the blur kernel. They use an asymmetric autoencoder with skip connections and a FCN to respectively capture the deep priors of latent clean image and blur kernel, respectively, 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 suggested 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 and no ground-truth clean image.

3.5.3 Reflection Separation

Kim et al. [43] proposed a learning-free method for unsupervised reflection separation from the background using two cross-coupled deep CNNs to separate the image into background and reflection. Their proposed approach has a performance comparable to supervised learning techniques, and higher than other unsupervised learning methods. They argue that UNNP [9] may not always produce good results for reflection separation from natural images as it is only able to capture the low-level image statistics. As a remedy, they propose a new architecture of the network called 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 are jointly optimized for reflection separation tasks. Furthermore, they argued that the original double-UNNP (proposed in [39]) will only work if the two different images are not correlated which is not applicable in natural images. Similarly, Chandramouli et al. [44] also proposed a UNNP-based method for reflection separation for face images. Unlike [9], 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.

Lei et al. [45] 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.

3.5.4 N-Layer Decomposition

Lu et al. [46] argued that the Double-UNNP approach proposed by Gandelsman et al. [39] has limited control over their output layers and entirely relies on the CNN properties to produce a meaningful decomposition. Their proposed method gives control over the decomposition which is required for re-timing people in videos. They further argued that their method is designed 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. Their technique is not able to effectively handle dynamic backgrounds. As a remedy, 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

In the original UNNP paper [9], Ulyanove et al. demonstrated the effectiveness of their method for different inverse imaging problems, i.e., image denoising, super-resolution, and inpainting. Since then UNNP has been used for a wide range of inverse problems. In this section, we present an overview of such applications. A taxonomy is presented in Figure 8. A summary of different applications of UNNP for different inverse problems can also be seen in Table 2.

4.1 3D Shape Reconstruction

3D shape reconstruction is a challenging problem because the search space of 3D shape is very large. Gadelha et al. [47] argued that this search over natural shapes can be replaced by a search over the parameters of the neural network using a UNNP network. 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 applied their proposed technique to a variety of reconstruction problems such as tomographic reconstruction, visual hull reconstruction, and 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 [116]. A survey on 3D shape reconstruction with a prime focus on DL-based methods is presented in [117].

4.2 Compressed Sensing

In [15], the authors leverage a small amount of training data for regularization, and used the UNNP approach to solve the compressive sensing problem. In contrast to other pre-trained models, this approach does not require access to ground truth, i.e, labelled data. Similarly, Ren et al. [118] combined UNNP and compressed sensing for the recovery of 1D signals generated from soils’ data, which are used to measure the quality of soil for agricultural purposes. In [119], 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. The proposed method is evaluated on different datasets including facial images, medical and multi-band astronomical images.

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. [49] proposed the use of UNNP for high dynamic range (HDR) image reconstruction without training data. They proposed a combination of UNNP and TV regularization to reconstruct low-light images. Their results demonstrate a significant improvement over previous traditional dynamic range enhancement techniques. Similarly, Cascarano et al. [50] proposed the use of a total variation 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 as ADMM-UNNPTV) outperforms two of the art techniques (UNNP [9] and UNNPTV [23]) 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. [120] proposed a learning-free framework to investigate the relationship between network 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 that how the image restoration tasks are affected by the construction of untrained networks. Instead of using the randomly initialized deep network as a handcrafted prior for image restoration, they used their proposed learning-free approach based on handcrafted structures for image restoration. They argued that the non-use of training samples to learn the image prior in UNNP [9] 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. [51] proposed an illumination normalization method to generate photorealistic face textures while preserving the face iden-
inity. 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. The output of these components has a peeled-off appearance which is corrected using UNNP in order to generate realistic textures. They conduct experiments on public datasets and show the robustness of their proposed method to illumination. Furthermore, they show that their proposed technique is able to outperform previous state-of-the-art techniques.

4.4 Image Denoising

In the original paper of UNNP [9], 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 is either known or unknown, i.e., blind image denoising. Zou et al. [55] designed a MobileNet based blind image quality assessment network for denoising of endoscopic images. Their method reduces by 36% the number of UNNP iterations required for image reconstruction. Park et al. [58] 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 [9] but proposed a new loss function that incorporates both constant and varying noise derived from a Gaussian distribution. The use of UNNP for the restoration of corrupted remote sensing images, especially where multi-temporal snapshots are not available was proposed in [121].

4.5 Image Deblurring

Although noise filters can effectively remove noise from a blurred image, they can also damage the blurred information by introducing more serious blur. Feng et al. [59] proposed the use of UNNP to restore the noisy and blurred images using a single degraded image. 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. [60] proposed BP-UNNP by combining UNNP with a backpropagation (BP) fidelity term to be used in place of the MSE loss (usually used in UNNP). They experimentally demonstrated that their proposed method is not only able to achieve a higher performance (in terms of PSNR value) compared to previous works but has also a better inference run-time. Their results show that BP-UNNP yields a higher PSNR value and reaches its peak PSNR in relatively fewer iterations than Least Square (LS)-UNNP. However, they argued that early stopping is still required to avoid the overfitting problem.

4.6 Image Super-resolution

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 by using existing techniques are based on the assumption that both of the input images are clean which is too idealistic for real cases. To address the problem of noisy HSI and MSI input, Nie et al. [61] proposed a UNNP-based HSI super-resolution method. UNNP is employed as the prior of the latent HR HSI which converts it into an end-to-end DL problem. Then it can be solved using back-propagation and can capture better statistics of the latent HR HSI. Furthermore, they demonstrate the effectiveness of their proposed technique by evaluating it on two benchmark datasets: the CAVE dataset and the Washington DC dataset. Ma et al. [62] proposed a combination of UNNP with a super-resolution convolutional neural network (SRCNN) to estimate fine resolution fraction images for each land cover type. Similarly, a UNNP-based approach for coded HSI reconstruction was proposed by Zhang et al. [63]. The approach 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. [67] 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 real-world, few efforts have taken this issue into account. As a solution to this, Nie et al. [67] 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 implement a UNNP-based spatial transformer network (STN) that avoids overfitting by constraining the STN. 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. [68] 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 using their proposed loss. Furthermore, they showed UNNP can be used to simultaneously solve the super-resolution and inpainting problems.

4.6.3 UNNP Variant for SR
Sagel et al. [73] argued that the optimization in the original UNNP limits the output image being reconstructed to be represented by a convolutional neural network, and this might neglect prior knowledge and may make certain regularizers ineffective in specific cases. As a solution, they suggest 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 and require extensive training and a large-scale annotated training data. To circumvent this issue, Weber et al. [75] 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. To reconstruct a depth map from a noisy and incomplete depth map, Ghosh et al. [18] used UNNP coupled with a depth reconstruction loss and a view-constrained photo-consistency loss (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 they demonstrated that dense 3D models of higher quality are produced by maps refined by their technique. Traditionally, handcrafted approaches have been used for depth map estimation, which suffers from their technique. Traditionally, handcrafted 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 [122].

4.8 Phase Retrieval
Untrained deep neural network priors have been used for linear as well as non-linear inverse problems. For instance, Jagatap et al. [99] 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 $n \ll d$. The authors solved this problem using two approaches, i.e., using gradient descent and projected gradient descent. Similarly, in [100], the authors proposed to use UNNP for unwrapping the phase in 2D quantitative phase imaging. The proposed method was evaluated on organoids images that are acquired using digital holographic microscopy. In [101], the authors used UNNP for Fourier ptychography, which is a special case of phase retrieval. Similarly, Boominathan et al. [102] proposed a UNNP based method for phase retrieval in Fourier ptychography.

4.9 Image Deconvolution
Wang et al. [104] combined the idea of learning-free deconvolution methods with neural networks and proposed a variant of UNNP named deep image kernel prior (DIKP). They showed that DIKP improves the performance of image deconvolution and outperforms traditional learning-free regularization-based priors for image deconvolution. Ren et al. [105] proposed a hybrid method for blind image deconvolution. They used UNNP (an autoencoder 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 [123], UNNP was used for deconvolution in electrochemical impedance spectroscopy (EIS).

4.10 Seismic Data Reconstruction
In [124], a UNNP based algorithm named DSPRecon was proposed for the reconstruction of seismic data. 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 integrates projection onto convex sets (POCS) based regularization with UNNP for the reconstruction of seismic data [125]. In a similar study, UNNP with multiple U-nets was used for the interpolation of seismic data to reconstruct 3D shot gathers [124].

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 [107]. 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) [126], basic iterative method (BI), and least-likely class iterative method (LLCI) [127]. A similar type of behavior in early iterations was observed in [108], 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. [109] 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 is also investigated in [110]. The authors considered two adversarial ML attacks FGSM and Carlini and Wagner (C&W) [128] in both black and white-box settings.

4.11.2 Crafting Adversarial Perturbations

In contrast with the aforementioned studies, Gittings et al. [111] proposed a method to generate robust adversarial image examples using a UNNP network. 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. [112] investigated the adversarial vulnerability of super-resolution deep priors against three attacks, i.e., style transfer attack, classification attack, and caption attack.

4.11.3 Generation of Counterfactual Explanations

Narayanaswamy et al. [113] proposed a regularization strategy based on an auxiliary loss estimator, which efficiently guides the prior 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.

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. [80] proposed a UNNP-based video inpainting algorithm that is 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 network to learn a “flow prior”. A consistency loss is then defined on the base 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 in order to improve the visual sharpness of the generated images.

4.12.2 Video Stabilization

Video destabilization occurs due to various problems such as lens distortion, dynamic objects, motion blur, and low illumination, etc. In this regard, a UNNP-based video stabilization system is proposed in [85]. 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, and they achieved visually more plausible and quantitatively better results than the previous state of the art.

4.12.3 Temporal Consistency

Lei et al. [90] 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. [92] proposed a UNNP-based hybrid method to recover old movies. Their proposed framework has two main components, i.e., speckle noise detection using spatiotemporal filtering techniques and UNNP-based image inpainting.

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 CNN. 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 (as shown in Figure 9). 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 3.

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
Similarly, in [132], the authors focused on the problem used for the reconstruction of time-dependent dynamic MRI with nonuniform fast Fourier transform (NuFFT) has been the proposed framework. In a similar study, UNNP along with learned regularization to improve the performance of network priors [15]. The authors also used early stopping based MRI image reconstruction using untrained neural networks [15]. Van et al. proposed a method for compressed sensing based MRI image reconstruction using untrained neural network priors [15]. The authors also used early stopping and learned regularization to improve the performance of the proposed framework. In a similar study, UNNP along with nonuniform fast Fourier transform (NuFFT) has been used for the reconstruction of time-dependent dynamic MRI [17]. Similarly, in [132], the authors focused on the problem of accelerated MRI reconstruction and proposed a variant of UNNP and deep decoder networks. Zhao et al. [134] proposed a UNNP based framework for MRI reconstruction from undersampled k-space data. Their proposed method uses a high-resolution MR image as an input reference to learn a structural prior during the training process.

5.1.1 MRI Reconstruction
Van et al. proposed a method for compressed sensing based MRI image reconstruction using untrained neural network priors [15]. The authors also used early stopping and learned regularization to improve the performance of the proposed framework. In a similar study, UNNP along with nonuniform fast Fourier transform (NuFFT) has been used for the reconstruction of time-dependent dynamic MRI [17]. Similarly, in [132], the authors focused on the problem of accelerated MRI reconstruction and proposed a variant of UNNP and deep decoder networks. Zhao et al. [134] proposed a UNNP based framework for MRI reconstruction from undersampled k-space data. Their proposed method uses a high-resolution MR image as an input reference to learn a structural prior during the training process.

5.1.2 Tomographic Image Reconstruction
5.1.2.1 PET Image Reconstruction: Gong et al. [135] proposed to leverage UNNP for PET image reconstruction that does not require any paired training data, except the patient-specific measured data. Specifically, they formulated the problem of PET image reconstruction as a constrained optimization problem and solved it using the alternating direction method of multipliers algorithm. Furthermore, the authors demonstrated that their proposed UNNP-based method provides a comparatively improved lesion contrast. Later, the authors extended this approach for low dose dynamic patlak PET image reconstruction [137]. Similarly, Yokota et al. [138] 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 clinically generated data.

5.1.2.2 Diffraction Tomography Reconstruction: Zhou et al. [139] proposed a UNNP based method for the reconstruction of high resolution 3D refractive index of thick biological samples using diffraction tomography. 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. [141] proposed to use the UNNP framework for the reconstruct-
tion of low-dose dual-energy CT images from noisy observations by jointly reconstructing low and high energy images. Similarly, Baguer et al. [142] proposed to use UNNP along with a classical TV regularizer 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, they do not use skip connections and 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 [147].

5.2 Medical Image Enhancement

5.2.1 Haze Removal in Fundus Imaging

In [39], 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. [16] 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. Their method does not require any training data neither paired nor unpaired and recovers an enhanced retinal image using a single degraded input.

5.2.2 Tomographic Image Denoising

Hashimoto et al. [148] proposed a UNNP based dynamic PET image denoising technique, which can handle unknown cases (unlike conventional methods). 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 algorithms. Similarly, Cui et al. [149] 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 of [154] and trained the network using a Limited Memory BFGS optimizer.

5.2.3 MRI Image Correction

Han et al. [150] 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 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 [151], a non-local UNNP is 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., 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 and performed a comparative analysis in terms of different standard metrics (i.e., contrast-to-noise ratio (CNR) and SSIM) [153].

6 Insights and Pitfalls

Based on our literature review on untrained neural networks, we summarize the salient insights related to the use of UNNPs for inverse imaging problems.

- The under-parameterized untrained neural networks (e.g., the deep decoder [19]) can only fit signals with low complexity, while over-parameterized untrained neural networks (e.g., UNNP [9]) 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. The smoothness and locality information, when modeling natural images using untrained neural networks, is enforced from CNN operations, i.e., upsampling and convolutions.

- Untrained neural network priors can be used to aid in the learning of priors. For instance, the authors in [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.

- It has been shown in [19] that the prior learned by untrained neural networks is asymptotically equivalent to handcrafted priors (e.g., Gaussian processes (GP)). However, GP are computationally more expensive than the stochastic gradient descent (SGD) optimization when we aim at modeling high resolution images.

- Although, untrained neural networks 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., total variation [23].

Limitations of UNNPs include:

- 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 for a theoretical understanding of the learning behavior of untrained neural networks.

- The prior estimated by untrained neural networks using SGD and stochastic gradient Langevin dynamics (SGLD) matches with the prior estimated by GP for smaller size models. However, it is still unknown whether this holds for larger networks [19].
- The literature suggests that the optimal stopping point for untrained neural networks depends on the structure of neural networks being used [27]. However, it is difficult to identify the optimal stopping point for any neural network architecture. Therefore, this remains an open ended problem.
- Yokota et al. [138] 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 for further developments. The authors proposed an optimal stopping strategy before the convergence of the parameters.

7 OPEN RESEARCH PROBLEMS

7.1 Exploring Network Architectures

The characteristics of untrained neural network priors are very similar to handcrafted priors based modeling, as in handcrafted-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 untrained neural network for a typical inverse problems relies on the preference of a specific neural network structure, i.e., a different architecture of CNN 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 problems at hand is still an open research problem.

7.2 Execution Time

Untrained neural network priors 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 untrained neural network priors 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 untrained neural network priors based unsupervised approaches.

7.3 Early Stopping

Despite their ability to effectively model natural images, untrained neural networks suffers from the issue of overfitting. In the original UNNP paper, 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 [19], Bayesian perspective of UNNP [20], and the projectional method [21]. These studies provided alternative approaches to alleviate the need of early stopping in untrained neural networks. 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 [116].

7.4 Theoretical Understanding

Despite the impressive success of untrained neural networks in numerous inverse imaging problems, there is still a lack of theoretical understanding of why untrained neural networks 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 untrained neural networks, e.g., Heckle et al. [19] had provided the theoretical reasoning behind their proposed variant of UNNP (i.e., deep decoder). Similarly, Yokota et al. proposed a manifold modeling in an embedding space to interpret UNNP [33]. However, more questions still need to be investigated about the working process of UNNP. 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 identify what 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.

8 CONCLUSIONS

Inverse problems arise in many distinct settings for 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 by clustering them into two groups, 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 UNNs which require further investigation.

References

[1] Mario Bertero and Patrizia Boccacci. Introduction to inverse problems in imaging. CRC press, 2020.
[2] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. Regularization of inverse problems, volume 375. Springer Science & Business Media, 1996.
[3] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. arXiv preprint arXiv:1801.03992, 2018.
[4] Christian Clason. Regularization of inverse problems. arXiv preprint arXiv:2001.00617, 2020.
[5] Chinmay Hegde. Algorithmic aspects of inverse problems using generative models. In 2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pages 166–172. IEEE, 2018.
[6] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. Introduction to inverse problems.
[7] Mohammad Bennamoun. A guide to convolutional neural networks for computer vision. Synthesis Lectures on Computer Vision, 8(1):1–207, 2018.
[8] Dmytro Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 9446–9454, 2018.
[9] Zheng Zhang, Yong Xu, Jian Yang, Xuelong Li, and David Zhang. A survey of sparse representation: algorithms and applications. IEEE access, 3:490–530, 2015.
[10] M Vivek, Joonki Paik, et al. Imaging inverse problem using sparse representation with adaptive dictionary learning. In 2015 IEEE International Advanced Computing Conference (IACC), pages 1247–1251. IEEE, 2015.
[11] Lucas Rencker, Francis Bach, Wenwu Wang, and Mark D Plumbley. Sparse recovery and dictionary learning from nonuniform measurements. IEEE Transactions on Signal Processing, 67(21):5659–5670, 2019.
[12] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. Globally and locally consistent image completion. ACM Transactions on Graphics (ToG), 36(4):1–14, 2017.
[13] Steve Diamond, Vincent Sitzmann, Felix Heide, and Gordon Wetzstein. Unrolled optimization with deep priors. arXiv preprint arXiv:1705.08041, 2017.
[14] Dave Van Veen, Ajil Jalal, Mahdi Soltankotabi, Eric Price, Srimat Vishwanath, and Alexandros G Dimakis. Compressed sensing with deep image prior and learned regularization. arXiv preprint arXiv:1806.06438, 2018.
[15] Aditya Qureshi, Waqs Sultani, Fahad Shamshad, Junaid Qadir, and Rashid Tufail. Single-shot retinal image enhancement using deep image priors. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23, pages 636–646. Springer, 2020.
[16] Kyong Hwan Jin, Harsh Gauba, Jerome Yerly, Matthias Stuber, and Michael Unser. Time-dependent deep image prior for dynamic MRI. arXiv preprint arXiv:1910.01684, 2019.
[17] Pallabi Ghosh, Vibhan Vineet, Larry S Davis, Abhinav Shrivastava, Sudipta Sinha, and Noel Joshi. Deep depth prior for multi-view stereo. arXiv preprint arXiv:2001.07791, 2020.
[18] Reinald Heckel and Paul Hand. Deep decoder: Concise image representations from untrained non-convolutional networks. arXiv preprint arXiv:2011.01748, 2020.
[19] Zehzhou Cheng, Mathews Gadelha, Subhransu Maji, and Daniel Sheldon. A bayesian perspective on the deep image prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5442–5451, 2019.
[20] Sören Dittmer and Peter Maass. A projectional ansatz to reconstruction. arXiv preprint arXiv:1907.04675, 2019.
[21] Zhaodong Sun. Solving inverse problems with hybrid deep image priors: the challenge of preventing overfitting. arXiv preprint arXiv:2001.04776, 2020.
[22] Christian Clason. Regularization of inverse problems by a deep prior approach for inverse problems. Journal of Mathematical Imaging and Vision, 62(3):456–470, 2020.
[23] Muhammad Asim, Fahad Shamshad, and Ali Ahmed. Patchdup exploiting patch redundancy in deep image prior for denoising. NeurIPS 2019 Workshop Deep Inverse Program Chairs, 2019.
[24] Mohammad Dittmer, Tobias Oeter and Daniel Oeter. Regularization by architecture: A deep prior approach for inverse problems. Journal of Mathematical Imaging and Vision, 62(3):456–470, 2020.
[25] Yinfu Chen, Chen Gao, Esther Robb, and Jia-Bin Huang. Nas-dip: Learning deep image prior with neural architecture search. arXiv preprint arXiv:2008.11713, 2020.
[26] Cheng Zhang, Ghiass, Tsung-yi Lin, and Quoc V Le. Nas-fpn: Learning scalable feature pyramid architecture for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7036–7045, 2019.
[27] Piyush Tang, Shiyu Chang, Yifan Jiang, and Zhangyang Wang. Autogen: Neural architecture search for generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, pages 3224–3234, 2019.
[28] Barret Zoph and Quoc V Le. Neural architecture search with reinforcement learning. arXiv preprint arXiv:1610.01576, 2016.
[29] Tatsuki Uezato, Danfeng Hong, Naoto Yokoya, and Wei He. Guided deep decoder: Unsupervised image pair fusion. In European Conference on Computer Vision, pages 87–102. Springer, 2020.
[30] Satoshi Iizuka, Hidenori Tanaka, Qibin Zhao, and Andrzei Cichocki. Manifold modeling in embedded space: A perspective for interpreting “deep image prior”. arXiv preprint arXiv:1908.02995, 2019.
[31] Prithvijit Chakrabarty and Subhransu Maji. The spectral bias of the deep image prior. arXiv preprint arXiv:1912.00805, 2019.
[32] Pengzhen Ren, Yun Fu, Cuong Nguyen, Yizhou Huang, Zhuhui Li, Xiaojia Qiang, and Xin Wang. A comprehensive survey of neural architecture search: Challenges and solutions. arXiv preprint arXiv:2006.02903, 2020.
[33] Ryo Segawa, Hitoshi Hayashi, and Shohei Fujiy. Proposal of new activation function in deep image prior. IEEE Transactions on Electrical and Electronic Engineering, 15(S):1248–1249, 2020.
[34] Shohei Fujii and Hitoshi Hayashi. Comparison of performance by activation functions on deep image prior. In 2019 International Conference on Information and Communication (ICALIC), pages 255–258. IEEE, 2019.
[35] Christopher A Metzler, Ali Mousavi, Reinald Heckel, and Richard G Baraniuk. Unsupervised learning with stein’s unbiased risk estimator. arXiv preprint arXiv:1805.10531, 2018.
[36] Yossi Gandelsman, Assaf Schochter, and Michal Irani. “double-dip”: Unsupervised image decomposition via coupled deep-image priors. arXiv preprint arXiv:1812.00467, 2018.
[37] Wei Wang, Pengfei Zhu, and Dongwei Ren. Semi-supervised learning to remove fences from a single image. In Chinese Conference on Pattern Recognition and Computer Vision (PRCV), pages 79–90. Springer, 2020.
[38] Cheng Tian and Tao Xiong. A generic solver combining unsupervised learning and representation learning for breaking text-based captchas. In Proceedings of The Web Conference 2020, pages 860–871, 2020.
[39] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu, and Wangmeng Zuo. Neural blind deconvolution using deep priors. arXiv preprint arXiv:1908.02187, 2019.
[40] Suhong Kim, Hamed RahmaniKhezri, Seyed Mohammad Nourbaksh, and Mohsen Heifeeda. Unsupervised single-image reflection separation using perceptual deep image priors. arXiv preprint arXiv:2009.07072, 2020.
[41] Paramand Chandramouli and Kanchana Vaishnavi Gandikota. Blind single image reflection suppression for forensic imaging using deep generative priors. In Proceedings of the IEEE International Conference on Computer Vision Workshops, pages 0–0, 2019.
[42] Chenyang Lei, Xuhua Huang, Mengdi Zhang, Qiong Yan, Wenxiu Sun, and Qifeng Chen. Polarized reflection removal with perfect alignment in the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1750–1758, 2020.
[43] Erika Lu, Forrester Cole, Tali Dekel, Weidi Xie, Andrew Zisserman, David Salesin, William T Freeman, and Michael Rubinstein. Layered neural rendering for retiming people in video. arXiv preprint arXiv:2009.07833, 2020.
[44] Mathews Gadelha, Rui Wang, and Subhransu Maji. Shape reconstruction using differentiable projections and deep priors. In Proceedings of the IEEE International Conference on Computer Vision, pages 7715–7724, 2019.
[45] Zhihong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Lingluang Zhang, Xiaou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3797–3805, 2019.
conference on computer vision and pattern recognition, pages 1912–1920, 2015.

[49] Gauti Jagapati and Chinmay Hegde. High dynamic range imaging using deep image prior. In CVPR, 2019.

[50] Pasquale Casciaro, Andrea Sebastiani, and Maria Colomba Comes. Admm-diptv: combining total variation and deep image prior for image restoration, arXiv preprint arXiv:2009.11380, 2020.

[51] Xianjun Han, Yanli Liu, Hongyu Yang, Guanyu Xing, and Yanci Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. Sun rgb-d: a rgb-d scene understanding benchmark suite. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 567–576, 2015.

[52] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Gavrila. Contour detection and hierarchical image segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 746–760. Springer, 2012.

[53] Haotian Zhang, Long Mai, Ning Xu, Zhaowen Wang, John Col- lomosse, and Hailin Jin. An internal learning approach to video inpainting. In Proceedings of the IEEE International Conference on Computer Vision, pages 2720–2729, 2019.

[54] Miguel Granados, Kwan In Kim, James Tompkin, Jan Kautz, and Christian Theobalt. Backward inpainting for videos with dynamic objects and a free-moving camera. In Computer Vision and Pattern Recognition Conference (CVPR), volume 862–695. Springer, 2012.

[55] Miguel Granados, James Tompkin, Jan Kautz, and Christian Theobalt. How not to be seen—object removal from videos of crowded scenes. In Computer Graphics Forum, volume 31, pages 219–228. Wiley Online Library, 2012.

[56] Arno Knapitsch, Jaesik Park, Qian-Yi Zhou, and Vladlen Koltun. Unsupervised deep hyperspectral super-resolution with unregulated diversity. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3800–3808, 2018.

[57] Shuaicheng Liu, Lu Yuan, Ping Tan, and Jian Sun. Steadyflow: Spatially smooth optical flow for video stabilization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4209–4216, 2014.

[58] Shuaicheng Liu, Lu Yuan, Ping Tan, and Jian Sun. Bundled camera paths for video stabilization. ACM Transactions on Graphics (TOG), 32(4):1–10, 2013.

[59] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In Proceedings of European Conference on Computer Vision, pages 19–34. Springer, 2016.

[60] Jianguo Nie, Lei Zhang, Cong Wang, Wei Wei, and Yanning Zhang. Robust deep hyperspectral imagery super-resolution. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Sympo- sium, pages 847–850. IEEE, 2019.

[61] Chen Liu, Xiaoguang Li, Li Zhuo, Jiafeng Li, and Qingfeng Zhou. Deep blind denoising method for endoscopic images. In 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS), pages 1–4. IEEE, 2019.

[62] Yunxin Xu, Hui Li, Zhetong Liang, and Zhihua Wang. Normalization of face illumination with photorealistic texture via deep image prior synthesis. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 683–691, 2019.

[63] Jenny Zukerman, Tom Tirer, and Raja Giryes. Bp-dip: A backpro- jection based deep image prior. arXiv preprint arXiv:2003.05417, 2020.

[64] Shuaicheng Liu, Yazhou Xing, and Qifeng Chen. Blind video tempo- ral consistency via deep video prior. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages 7451–7459, 2018.

[65] Jianguo Nie, Lei Zhang, Wei Wei, and Yanning Zhang. Robust deep hyperspectral imagery super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 5859–5869, 2019.

[66] Fumihiyo Yasuma, Tomoo Mitsunaga, Daiyuse Ito, and Shree K Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. IEEE transactions on image processing, 19(9):2241–2253, 2010.

[67] Ralf Grzeskowiak, Jan Tschirren, and Michael J. Black. Stereo datasets with subpixel-accurate ground truth. In German conference on pattern recognition, pages 31–42. Springer, 2014.

[68] Chen Liu, Zhihong Dang, and Paul Lichtenberg. Shape from illumination with two spec- tral rgb-d: A rgb-d scene understanding benchmark suite. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 55–59, 2011.

[69] Kuan Wu, Ankur Handa, and Christian Theobalt. Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence, 33(5):898–916, 2010.

[70] David Ferstl, Christian Reinbacher, Rene Ranftl, Matthias Ruther, and Horst Bischof. Image guided depth upsampling using anisotropic total generalized variation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 3531–3541, 2013.

[71] Alexander Sagel, Aline Roumy, and Christine Guillemot. Super-dip: Optimization on a subspace with deep image prior regularization and application to superresolution. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 8653–8657, 2020.

[72] Yunsong Wang, Wei Han, Christian Theobalt, and Jitendra Malik. Contour detection and hierarchical image segmentation. IEEE transactions on pattern analysis and machine intelligence, 33(5):898–916, 2010.
Kevin C Zhou and Roarke Horstmeyer. Diffraction tomography with a deep image prior. *Optics Express*, 28(9):12872–12896, 2020.

Aurélien Lucchi, Kevin Smith, Radhakrishna Achanta, Graham Knott, and Pascal Fua. Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE transactions on medical imaging*, 31(2):474–486, 2011.

Kuang Gong, Kyungsang Kim, Dufan Wu, Manudeep K Kalra, and Quanzheng Li. Low-dose dual energy ct image reconstruction using non-local deep image prior. *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–2. IEEE, 2019.

Daniel Otero Baguer, Johannes Leuschner, and Maximilian Schmidt. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *arXiv preprint arXiv:2003.04889*, 2020.

Johannes Leuschner, Maximilian Schmidt, Daniel Otero Baguer, and Peter Maas. The lodopab-c dataset: A benchmark dataset for low-dose ct reconstruction methods. *arXiv preprint arXiv:1910.01113*, 2019.

Johannes Leuschner, Maximilian Schmidt, and David Erzmann. Deep inversion validation library. 2019.

Tobias Knopp, Patryk Szewrgulis, Florian Greise, and Matthias Gräser. Omnipet: An initiative for freely accessible magnetic particle imaging data. *Data in brief*, 28:104971, 2020.

Tri Vu, Anthony DiSpirito III, Daiwei Li, Xizuan Zhang, Xiaoyi Zhu, Maomao Chen, Dong Zhang, Jianwen Luo, Yu Shrike Zhang, Roarke Horstmeyer, et al. Deep image prior for sparse-sampling photoacoustic microscopy. *arXiv preprint arXiv:2010.12041*, 2020.

Fumio Hashimoto, Kibo Ote, and Hideo Tsukada. Dynamic pet image denoising using deep convolutional neural network without prior training data. *Transactions on medical imaging*, pages 404–413. Springer, 2020.

Wenshi Fan, Hancheng Yu, Tianming Chen, and Sheng Ji. OCT image restoration using non-local deep image prior. *Electronics*, 9(5):784, 2020.

Leyuan Fang, Shuota Li, Ryan P McNabb, Qing Nie, Anthony N Kuo, Cynthia A Toth, Joseph A Izatt, and Sina Farsiu. Fast acquisition and reconstruction of optical coherence tomography images via sparse representation. *IEEE transactions on medical imaging*, 32(11):2034–2049, 2013.

Kristen Qgsen, David Li, and Jessica Loo. Noise reduction in optical coherence tomography using deep image prior. *Optical Engineering*, 52(10):100901, 2013.

Oğuz Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.

Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Tobias Knopp, Patryk Szwargulski, Florian Griese, and Matthias Gräser. Openmpidata: An initiative for freely accessible magnetic particle imaging data. *Data in brief*, 28:104971, 2020.

Adnan Qayyum is a Ph.D. candidate at the Information Technology University (ITU) of Punjab, Lahore, Pakistan. His research interests include inverse medical imaging problems, healthcare, and secure and robust ML. He received the Bachelor's degree in Electrical (Computer) Engineering from COMSATS Institute of Information Technology, Wah, Pakistan, in 2014 and M.S. degree in Computer Engineering (Signal and Image Processing) from the University of Engineering and Technology, Taxila, Pakistan, in 2016.

Inaam Ilahi is a research assistant at Information Technology University (ITU) of Punjab, Lahore, Pakistan. He has completed MS in Computer Science from ITU in 2020. He received bachelor degree in Electrical Engineering from National University of Sciences and Technology (NUST), Islamabad, Pakistan in 2017. His research interests include reinforcement learning, adversarial machine learning, and applications of deep reinforcement learning in networks.

Fahad Shamshad is Senior Machine Learning Engineer at Pakistani-US AI startup OMNO. He worked as a research associate at the Center of Artificial Intelligence and Computational Analytics, Information Technology University, Lahore. He received his MS from the National University of Sciences and Technology (NUST) and BS from the Institute of Space Technology (IST), Islamabad in Electrical Engineering. He is the recipient of the NUST merit-based scholarship and Pakistan's National Space agency scholarship during MS and BS, respectively. His research interests include inverse imaging problems, generative models, and optimization theory.

Farid Boussaid received the M.S. and Ph.D. 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 in the Department of Computer Science and Software Engineering at the University of Western Australia (UWA) and is a researcher in computer vision, machine/deep learning, robotics, and signal/speech processing. He has published 4 books (available on Amazon), 1 edited book, 1 Encyclopedia article, 14 book chapters, 150+ journal papers, 260+ conference publications, 15 invited and keynote publications. His h-index is 77 and his number of citations is 13,500+ (Google Scholar). He was awarded 150+ 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 is the director of the IHSAN Research Lab and a Professor at the Electrical Engineering Department at the Information Technology University (ITU) of Punjab in 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 research journals including IEEE Communication Magazine, IEEE Journal on Selected Areas in Communication (JSAC), IEEE Communications (CST), and IEEE Transactions on Mobile Computing (TMC). He was awarded the highest national teaching award in Pakistan—the higher education commission’s (HEC) best university teacher award—for the year 2012-2013. He has been appointed as ACM Distinguished Speaker for a three-year term starting from 2020. He is a senior member of both IEEE and ACM.