RESEARCH ARTICLE

Advancing Sensor-Data Based PAT Image Reconstruction Through Efficient and Intelligible Unrolled Networks

MARY JOHN AND IMAD BARHUMI[ID], (Senior Member, IEEE)
Department of Electrical Engineering, United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates

Corresponding author: Imad Barhumi (imad.barhumi@uaeu.ac.ae)

ABSTRACT Despite recent advancements, the challenge of deep-dense tissue imaging with high resolution and contrast persists in breast cancer diagnosis; however, photoacoustic tomography (PAT) imaging addresses this issue by capturing both anatomical and functional information, including small masses as tiny as 3mm. Compressive sensing coined with PAT allows for a faster reconstruction because it requires fewer sensors and samples. Although CS-PAT algorithms are efficient they rely on application-specific parameter tuning and system matrix modeling. This paper introduces a time-efficient approach of algorithm unrolling, based on CS algorithms, to directly reconstruct high-resolution PAT images from sensor data, eliminating the need for explicit parameter assignment and system matrix modeling. The study proposes two unrolled deep learning networks based on split-Bregman total-variation and relaxed-basis-pursuit with ADMM implementation, followed by a denoising network for further enhancement. The model-free unrolled deep-learning approach successfully reconstructs high-resolution PAT images, even in the presence of noise, with low validation root mean square error. An enhancer network based on U-Net improves image quality to 0.91 and significantly reduces mean square error by 95%. Overall, the proposed algorithm unrolling method demonstrates the promising potential for practical clinical applications, particularly in early disease detection, offering rapid image reconstruction without explicit system matrix modeling or parameter tuning. The inclusion of a U-Net denoising network enhances the approach's resilience and adaptability, suggesting possibilities for improved disease diagnosis and treatment outcomes, especially for early detection in dense tissues like breasts.

INDEX TERMS Compressive sensing, deep learning, photoacoustic tomography, U-Net denoising, unrolled networks.

I. INTRODUCTION

Over the past ten years, photoacoustic tomography (PAT) has attracted a lot of attention. It has a wide range of uses in biological (in vivo and ex vivo) imaging [1] such as the detection of hemoglobin [2], functional brain imaging [3], oxygen saturation monitoring [4], and pre-clinical cancer diagnoses [5]. Industrial imaging [6], including imaging of fractures, flaws [7] and lithium batteries [8] are just a few of the fascinating imaging applications made feasible by PAT imaging. The basis of PAT imaging is the photoacoustic

The associate editor coordinating the review of this manuscript and approving it for publication was Qingchun Chen[ID].

phenomenon, whereby sound waves are detected as a result of optical illumination. The advantages of traditional optical and ultrasound imaging are available with PA imaging without the severe disadvantages of either technique. PAT imaging demonstrates that it is non-invasive and is capable of obtaining high-resolution deep tissue images [9].

In 1800, Graham Bell discovered that a medium may absorb electromagnetic waves, which led to the development of the photoacoustic effect. Later, Gregor in [10] claimed that generalizable electromagnetic energy of small duration was regionally absorbed within a turbid and immensely scattering medium, and photoacoustic echoes could be captured with a high spatial resolution with shorter-duration

radiation dose and higher-frequency transducer designs. Then it was expanded to include the tissue tomography approach, which allows for the imaging of blood vessels in highly scattering substances and the determination of blood concentrations [11]. Recently, this has extended the research to the case of detection of breast cancer which is one of the common cancers in women [12], [13]. The death rate shifts towards more obvious higher grading and rises by an average of 1.3% for every millimeter of tumor size at detection. These data provide credence to the idea that prognosis is significantly impacted by early identification [14]. The literature review explores four noteworthy studies on breast cancer detection and control [15], [16], [17], [18].

A novel sensing technique called compressive sensing (CS) enables us to sample the signal much more slowly than the Nyquist sampling rate. Donoho, Candes, Romberg, and Tao introduced CS in 2004 [19]. It is mostly employed for the acquisition of naturally sparse signals. The usage of CS in a number of domains, including magnetic resonance imaging, high-speed video collection, and ultra-wideband communication, has sparked study interest as a result of this feature [20], [21]. When employing the Nyquist criterion, the sampling rate is determined by the signal's highest frequency component, but in CS, the signal sparsity determines the sampling rate. The orthonormal basis of a signal is assumed to be sparse and is the primary reason for the success of standard compression tools [19] and it eliminates the need for huge storage requirements and expedites transmission. Faster acquisition times, fewer sensors, and hence lower costs are further benefits of CS [22]. Because of the many well-liked benefits of CS in medical imaging, it has been employed alongside magnetic resonance imaging (MRI), X-ray, optical coherence tomography (OCT), and PAT. Reduced data requirements for image creation have been a major driving force behind the advent of CS in MRI [23]. However, MRI does not create deep tissue images and takes a long time, and it requires the subject to be stationary. Similarly, CS is advantageous for X-ray computed tomography (CT) imaging due to the reduced X-ray dose because it requires fewer views than the traditional approach, enabling quicker acquisition and removing the need for the patient to remain still for extended periods of time [24]. Nevertheless, ionizing X-rays are mostly used to view bones. OCTs are not employed for other imaging purposes and are primarily used to image eyes [25]. In light of all the drawbacks and benefits we seek, PAT answers the need for non-invasive, high-resolution deep-tissue imaging. In comparison to conventional optical modalities, PAT can provide substantially superior spatial resolution at depths below the optical ballistic regime (1 mm in the skin). This is the driving force behind adopting PAT, which when used in conjunction with CS enables a faster reconstruction.

Compressive sensing photoacoustic tomography (CS-PAT) algorithms are widely used for localizing sparse signal emitters and have found applications in various settings.

In general, they fall into the categories of greedy algorithms [26], threshold type [27], combinational type [28], convex [29] and non-convex optimization [30] formulations. However, these algorithms often suffer from high computational costs, and their efficiency can vary significantly depending on the employed techniques [33]. Moreover, when dealing with uncertain imaging systems, optimization parameters may need to be modified, adding further complexity. Additionally, the specialized design of CS-PAT algorithms primarily focuses on leveraging sparsity, often neglecting factors such as cell size and density. Iterative approaches used in CS-PAT require the modeling of the system matrix and may not account for important contextual information. To overcome these limitations, a promising alternative is to replace the iterative algorithms with trained neural networks that perform the same mathematical operations. This procedure is known as algorithm unrolling or unfolding. For sparse recovery, this work was initially put forth by Gregor and LeCun in 2010 [34]. The unrolled network is built upon the architecture of traditional CS-PAT iterative algorithms while offering several key advantages. In iterative approaches, parameter tuning poses challenges, requiring adjustments for different applications, and system matrix modeling is time-consuming. However, the unrolled network directly reconstructs images from pressure/sensor data, sidestepping the need for explicit system matrix modeling, and leading to more efficient reconstructions. By leveraging neural networks, it considers crucial contextual factors often overlooked by iterative methods, resulting in accurate reconstructions without manual parameter tuning or sparsity assumptions. Overall, the unrolled network combines the interpretability of iterative approaches with the adaptability and efficiency of learning methods, overcoming limitations, and advancing image reconstruction in various settings.

This research article presents the development of an unrolled network by concatenating the iteration block (IB) of the present iteration to the IB of the next iteration. These IBs are based solely on the algorithm's architecture and its augmented variables. Additionally, the classic iterative algorithms' stated parameters and coefficients are now learned from the training data during backpropagation, making it more generalizable. Furthermore, since iterative algorithms only use a limited number of parameters or coefficients, a smaller training set is required for rolling them out. In image processing, image denoising and enhancement techniques are widely applied to improve image quality by reducing noise. Proxy images, which are degraded versions of the expected output images, can be processed with these techniques to enhance the accuracy and quality of the final output images. Using denoising networks trained on proxy images has become increasingly popular, especially in deep learning-based image reconstruction methods, as they can remove noise and artifacts, resulting in more precise images suitable for further analysis. U-Net [35] is one of the deep learning models used for image denoising and enhancement.

Although other models such as convolutional networks (CNNs), autoencoders, and generative adversarial networks [36] also improve image quality, U-Net is distinguished by its unique architecture that includes skip connections. These connections allow U-Net to capture both low-level and high-level features of an image and preserve finer details during the reconstruction process.

This research proposes two novel unrolled networks, one is based on split Bregman total variation (SBTV) and the other is based on relaxed basis pursuit alternating direction method of multipliers (rBP-ADMM). These networks can effectively reconstruct an image from PAT sensor data, even with a limited training set, and show improved image quality compared to the learned iterative shrinkage thresholding algorithm (LISTA) [34].

The significance of utilizing PAT with laser pulses for tissue excitation lies in its ability to visualize deeper tissues, surpassing conventional modalities commonly used for breast cancer diagnosis, such as mammography, ultrasound, MRI, and CT scans. The proposed image reconstruction approach, combining the unrolling method with a U-net based denoiser, directly produces high-contrast and high-resolution images from sensor data. By eliminating the need for a predefined measurement model, this method offers an efficient and accurate way to visualize tissues. The findings have substantial implications for breast cancer diagnosis and early disease detection.

This paper is organized as follows. Initially, a brief review of existing compressive sensing algorithms and algorithm unrolling is provided in Section II, followed by the proposed unrolled algorithms and their implementation in Section III. Then, the comparison of the proposed unrolled networks with an existing unrolled algorithm is provided in Section IV. A comparative evaluation of CS algorithms and unrolled-denoiser networks is discussed in Section V. Finally, Section VI provides a summary and a description of the research's future direction.

II. THEORY

When a region is heated there occurs a change in pressure and as a result, acoustic waves are generated. The linear approximation of the acoustic pressure obeys the wave equation, and it is solved using Green's function (GF) approach. GF of a PA equation represents the response (increase in pressure due to absorption of laser pulse) of a point absorber to heating [37]. By evaluating the pressure at different time instants for different sensor locations, the system matrix can be formed. The computation is performed over two grids; the inner grid and outer grid with equal grid spacing d. The inner grid is used to define initial pressure distribution with $N_{in} \times N_{in}$ grid points and the outer grid is used for k-space computation with $N_{out} \times N_{out}$ grid points. For all time instants (N_t), the measurement of raw data is found using the equation:

$$\mathbf{y} = \mathbf{M}\mathbf{x}, \qquad (1)$$

where $\mathbf{y} \in \mathbb{R}^{N_s N_t \times 1}$ is the measurement vector obtained with N_s is the number of sensors, $\mathbf{M} \in \mathbb{R}^{N_s N_t \times N_{in}^2}$ is the system matrix and $\mathbf{x} \in \mathbb{R}^{N_{in}^2 \times 1}$ is the image to be reconstructed (in vectorized form) [38]. Thus, using the resulting system matrix, CS algorithms may be created, and their effectiveness can be assessed. The sparsity required by CS to rebuild an image from fewer data is enforced by applying a regularization parameter and prior knowledge. Subsection II-A discusses some of the existing and commonly used CS algorithms.

A. COMPRESSIVE SENSING ALGORITHMS

- l_2-regularized least squares: The least squares (LS) is an approach to mathematical regression that identifies the line that fits a given collection of data the best. l_2-regularized LS also known as Tikhonov regularization is formulated as

$$\mathbf{x}^* = \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_2^2, \qquad (2)$$

where λ is the regularization parameter. This type of regularization causes the error to be distributed and is computationally expensive [33], [39].

- Basis pursuit (BP): it solves an optimization problem of the form

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \ \mathbf{M}\mathbf{x} = \mathbf{y}. \qquad (3)$$

Although it has a high computational cost, the BP approach [40] has the advantage of offering good resolution. It takes substantially longer time to compute with some of the current solvers, including CVX [41] and l_1-MAGIC [22], [42]. In this paper, an unrolled relaxed-BP using ADMM is proposed and is presented and found to produce superior outcomes than the use of an iterative algorithm.

- Total variation (TV) regularized LS: The ability of TV regularization to maintain image edges makes it a preferred choice for image restoration and reconstruction [43], [44]. It is formulated as

$$\mathbf{x}_{TV} = \min_{\mathbf{x}} \|\mathbf{D}\mathbf{x}\|_p \quad \text{s.t.} \ \mathbf{M}\mathbf{x} = \mathbf{y}, \qquad (4)$$

where $\mathbf{D}$ is a difference matrix defined as:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & & \\ & -1 & 1 & \\ & & \cdots & \cdots \\ & & & -1 \end{bmatrix}. \qquad (5)$$

When $p = 1$, (4) is referred to as an anisotropic TV problem and as isotropic-TV when $p = 2$ [45], [46], [47]. In this study, a TV regularized LS problem is formulated using an unrolled split-Bregman formulation, which yields results that are similar to those of an unrolled relaxed BP formulation.

The next subsection provides a brief theory about unrolled networks and their implementation.
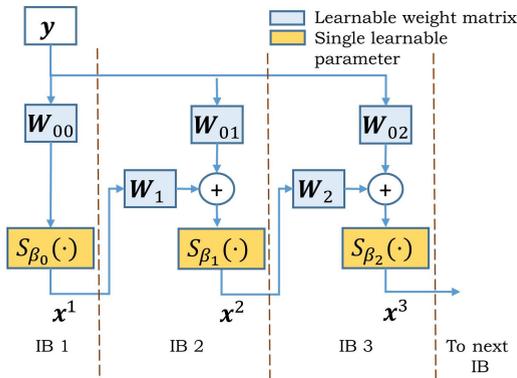
FIGURE 1. U-ISTA DL network block diagram showing the output of first three IBs.

B. ALGORITHM UNROLLING-BASICS

The first work of algorithm unrolling was done by Gregor in [34] and the main focus of that paper and our work is to increase the computational efficiency by using a trained network for reconstruction. The efficient implementation of iterative shrinkage thresholding algorithm (ISTA) using the method of algorithm unrolling to obtain LISTA is also demonstrated in [34]. The implementation of ISTA is shown in Algorithm 1.

Algorithm 1 ISTA

Require: $\mathbf{y}$, $\mathbf{M}$, λ, L, $kmax$, tol.
  Initialize: $k = 0$, NMSE $= \infty$
  while $k < kmax$ or NMSE $>$ tol **do**
    $\mathbf{x}^{k+1} \leftarrow S_{\frac{\lambda}{L}}(\mathbf{x^k} - 2L\mathbf{M}^T(\mathbf{Mx^k} - \mathbf{y}))$
    NMSE $\leftarrow \frac{\|\mathbf{x}^{k+1}-\mathbf{x}^k\|_2^2}{\|\mathbf{x}^{k+1}\|_2^2}$
    $k \leftarrow k + 1$
  end while
  $\hat{\mathbf{x}} = \mathbf{x}^k$

Based on Algorithm 1, unrolled ISTA (U-ISTA) deep learning (DL) network is constructed as follows. The update of $\mathbf{x}$ at iteration $k + 1$ is

$$\mathbf{x}^{k+1} = S_{\beta_k}(\mathbf{W}_k\mathbf{x}^k + \mathbf{W}_{0k}\mathbf{y}), \tag{6}$$

where k is the iteration index, $\mathbf{W}_k$ is the weight matrix (resembles a fully connected network) with $\mathbf{W}_0 = \mathbf{0}$, $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{W}_k$ resembles $(\mathbf{I} - 2L\mathbf{M}^T\mathbf{M})$, and $\mathbf{W}_{0k}$ resembles $2L\mathbf{M}^T$. In the unrolled network, each iteration is referred to as an IB. $S_{\beta_k}(\cdot)$ is the element-wise shrinkage operator and β_k as the learnable threshold value.

The shrinkage or soft thresholding operation is implemented using two ReLU operations as

$$S_\beta(\mathbf{x}) = ReLU(\mathbf{x} - \beta) - ReLU(-\mathbf{x} - \beta), \tag{7}$$

where β is the learnable threshold value of the shrinkage operator. The block diagram of U-ISTA DL network implementation is shown in Fig. 1. In [34] it has been stated that LISTA is twenty times faster than accelerated ISTA and uses much fewer iterations.

III. PROPOSED WORK

In this section, two unrolled deep learning algorithms are proposed. The first one is unrolled split Bregman (SB) TV (U-SBTV) and the second one is unrolled relaxed BP-ADMM (U-rBP-ADMM) DL network. The unrolled DL network implementation using a deep neural network is discussed in this section.

A. UNROLLED SBTV

A TV regularization is primarily chosen as it provides a solution sparser than l_1-regularization. It is essentially the gradient of the l_1 norm. Regularizations impose a limit on the solution set and should be chosen such that they maintain and enforce sparsity (as they allow for fast computation and less memory storage). Since a TV regularized optimization problem is non-convex, splitting techniques must be used to divide the original optimization problem into equivalent sub-problems rather than trying to solve it as a single optimization problem. Iterative re-weighted technique [48] and quadratic programming [49] are some of the two methods that can be used to solve a TV problem, but were primarily used for denoising and restoration applications and are computationally expensive. Split-Bregman TV (SBTV) algorithm is an iterative method employed in image reconstruction, specifically for tasks such as image denoising and deblurring. It extends the widely recognized Bregman iteration technique and leverages total variation (TV) regularization principles. An SBTV problem can be formulated as

$$\mathbf{x}_{TV} = \min_{\mathbf{x}} \|D\mathbf{x}\|_1 + \frac{\beta}{2}\|\mathbf{Mx} - \mathbf{y}\|_2^2. \tag{8}$$

An auxiliary variable $\mathbf{d} = \mathbf{Dx}$, is introduced along with the split Bregman parameter $\mathbf{b}$. So, the SBTV problem can be decoupled in terms of the $\mathbf{x}$, $\mathbf{d}$ and $\mathbf{b}$ update as follows,

$$\mathbf{x}^{k+1} = \min_{\mathbf{x}} \frac{\beta}{2}\|\mathbf{Mx} - \mathbf{y}\|_2^2 + \frac{\gamma}{2}\left\|\mathbf{d}^k - \mathbf{Dx} - \mathbf{b}^k\right\|_2^2, \tag{9}$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{d}^{k+1} - \mathbf{Dx}^{k+1}. \tag{10}$$

The SBTV formulation allows for efficient decoupling of the TV regularized optimization problem into smaller equivalent sub-problems, enabling iterative exact minimization. This feature makes it more efficient than conventional implementations. The SBTV algorithm, presented in Algorithm 2, involves solving these sub-problems iteratively until the convergence criteria are met.

An unrolled SBTV DL network is developed using the architecture of Algorithm 2 as

$$\mathbf{x}^{k+1} = \mathbf{W}_{0k}\mathbf{y} + \gamma_k\mathbf{W}_{1k}\mathbf{D}^T(\mathbf{d}^k - \mathbf{b}^k), \tag{11}$$

$$\mathbf{d}^{k+1} = S_{\beta_k}(\mathbf{Dx}^{k+1} + \mathbf{b}^k), \tag{12}$$

and

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \mathbf{d}^{k+1} - \mathbf{Dx}^{k+1}, \tag{13}$$

where k is the number of iterations. Here $\mathbf{W}_{0k}$ is a learnable weighting matrix that resembles $(\mathbf{M}^T\mathbf{M} + \frac{\gamma}{\beta}\mathbf{D}^T\mathbf{D})^{-1}\mathbf{M}^T$,

Algorithm 2 SBTV

Require: : $\beta, \gamma > 0$, tol, $\mathbf{y}$, $\mathbf{M}$, $kmax$.
　Initialize: $k = 0$, NMSE $= \infty$, $\mathbf{d}^0 = \mathbf{b}^0 = \mathbf{0}$
　while $k < kmax$ or NMSE $>$ tol **do**
　　$\mathbf{x}^{k+1} \leftarrow (\beta \mathbf{M}^T \mathbf{M} + \gamma \mathbf{D}^T \mathbf{D})^{-1} (\beta \mathbf{M}^T \mathbf{y} + \gamma \mathbf{D}^T (\mathbf{d}^k - \mathbf{b}^k))$
　　$\mathbf{d}^{k+1} \leftarrow S_{\frac{2}{\gamma}} (\mathbf{D} \mathbf{x}^{k+1} + \mathbf{b}^k)$
　　$\mathbf{b}^{k+1} \leftarrow \mathbf{b}^k + \mathbf{d}^{k+1} - \mathbf{D} \mathbf{x}^{k+1}$
　　NMSE $\leftarrow \frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2}{\|\mathbf{x}^{k+1}\|_2^2}$
　　$k \leftarrow k + 1$
　end while
　$\hat{\mathbf{x}} = \mathbf{x}^k$
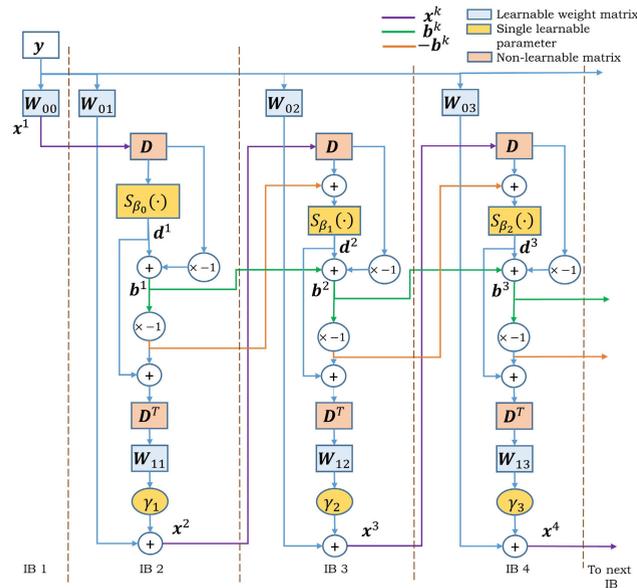
FIGURE 2. U-SBTV DL network block diagram showing the output of first four iteration blocks.

for each iteration k. More flexibility is provided by learning it separately for each IB rather than defining it the same for all iterations. $\mathbf{W}_{1k}$ is another learnable weighting matrix that resembles $(\beta \mathbf{M}^T \mathbf{M} + \gamma \mathbf{D}^T \mathbf{D})^{-1}$. The parameter γ is designed as a learnable parameter γ_k. Matrix $\mathbf{D}$ is defined as a non-learnable function matrix as in equation (5) and its transpose is $\mathbf{D}^T$. The shrinkage operation in the $\mathbf{d}$-update with a threshold value of $\frac{2}{\gamma}$ is implemented using a shrinkage operation, as in equation (7), with a learnable threshold β_k for IB k. The threshold value is learned separately for each IB rather than explicitly defining it the same for all iterations.

The U-SBTV DL network is implemented in Fig. 2. An IB is numbered based on the output of $\mathbf{x}^k$-updates for kth iteration. Combining the IBs leads to the formation of U-SBTV DL network. The number of IBs required is determined after training using the required dataset.

B. UNROLLED rBP-ADMM

In BP-ADMM [50], the initial BP problem is broken down into equivalent sub-problems that could be solved iteratively to reconstruct a PAT image. However, this approach fails

to reconstruct a PAT image due to the high sensitivity to system modeling errors. In this paper, to overcome this problem, a relaxation factor is incorporated in the update of the data vector $\mathbf{x}$, such that it updates more gently than other parameters. The modified BP-ADMM, called relaxed BP-ADMM (rBP-ADMM), is able to recover the PAT image from pressure measurements. The algorithm can also be employed with an accelerated variation (similar work done in [31] and [32]), in which the parameter γ is multiplied by an acceleration factor, chosen to be less than 1. This adjustment results in fewer iterations, thus reducing computation time, all while preserving the algorithm's performance. The value of the parameter ρ is chosen as 1 but is still maintained for the unrolled network to provide flexibility and interpretability. Cholesky factorization can also be used in the algorithm to execute the inverse faster. In [33], it is demonstrated that the implementation of rBP-ADMM, which contains a relaxation parameter and Cholesky factorization, is significantly more effective than the implementation of the traditional BP-ADMM. The implementation of rBP-ADMM is presented in Algorithm 3. Based on Algorithm 3, the U-rBP-ADMM

Algorithm 3 rBP-ADMM

　Require: $\rho, \alpha, \gamma, r, \mathbf{y}, \mathbf{M}, kmax$, tol.
　Initialize: $k = 0$, $\mathbf{z}^0 = \mathbf{u}^0 = 0$, NMSE $= \infty$
　while $k < kmax$ or NMSE $>$ tol **do**
　　$\mathbf{x}^{k+1} \leftarrow \left(\rho \mathbf{I} + \alpha \mathbf{M}^T \mathbf{M}\right)^{-1} \left(\rho \mathbf{z}^k + r \mathbf{u}^k + \alpha \mathbf{M}^T \mathbf{y}\right)$
　　$\mathbf{z}^{k+1} \leftarrow S_{\frac{\gamma}{\rho}} \left(\frac{\mathbf{u}^k}{\rho} + \mathbf{x}^{k+1}\right)$
　　$\mathbf{u}^{k+1} \leftarrow \mathbf{u}^k + \rho \left(\mathbf{x}^{k+1} - \mathbf{z}^{k+1}\right)$
　　NMSE $\leftarrow \frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2}{\|\mathbf{x}^{k+1}\|_2^2}$
　　$k \leftarrow k + 1$
　end while
　$\hat{\mathbf{x}} = \mathbf{x}^k$

DL network can be implemented using deep neural network layers that perform similar updates as in the algorithm.

The updates at $(k + 1)$th IB of the U-rBP-ADMM DL network are designed as follows.

$$\mathbf{x}^{k+1} = \mathbf{C}_{1k} \mathbf{z}^k + \mathbf{C}_{2k} \mathbf{u}^k + \mathbf{W}_{1k} \mathbf{y}, \qquad (14)$$

$$\mathbf{z}^{k+1} = S_{\beta_k}(\zeta_k \mathbf{u}^k + \mathbf{x}^{k+1}), \qquad (15)$$

and

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \tau_k(\mathbf{x}^{k+1} - \mathbf{z}^{k+1}). \qquad (16)$$

Here $\mathbf{C}_{1k}$ resembles $(\mathbf{I} + \frac{\alpha}{\rho} \mathbf{M}^T \mathbf{M})^{-1}$, $\mathbf{C}_{2k}$ resembles $(\rho \mathbf{I} + \alpha \mathbf{M}^T \mathbf{M})^{-1} r$, and $\mathbf{W}_{1k}$ resembles $(\frac{\rho}{\alpha} \mathbf{I} + \mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T$. The blocks $\mathbf{C}_{1k}$ and $\mathbf{C}_{2k}$ are designed as a learnable weighting matrices. The constant $\frac{1}{\rho}$ is designed using ζ_k, which is a learnable parameter. The shrinkage operator with threshold value $\frac{\gamma}{\rho}$ is implemented using the layer $S_{\beta_k}(\cdot)$, with β_k being the learnable threshold value, learned separately for each IB. The network implementation is shown in Fig. 3.
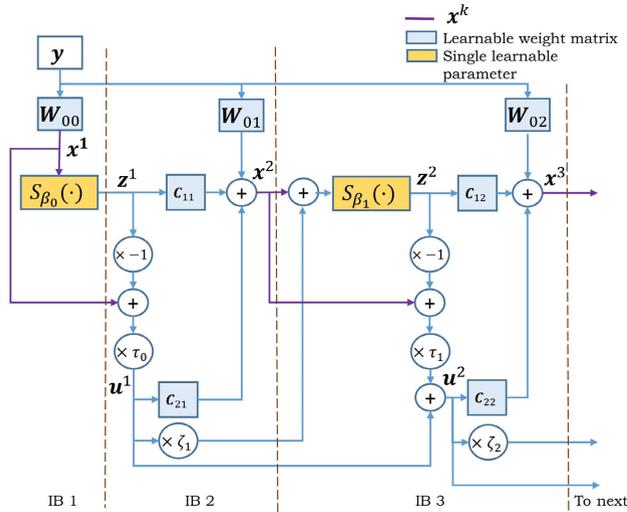
FIGURE 3. U-rBP-ADMM block diagram showing the output of first three blocks.
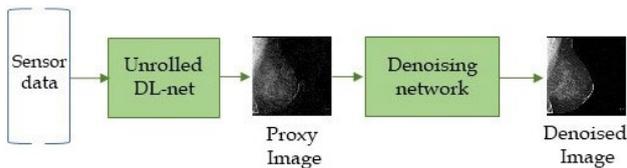
FIGURE 4. System implementation for handling significant differences between test and training sensor data.

C. IMPROVING IMAGE RECONSTRUCTION IN UNROLLED DEEP NETWORKS WITH DENOISING NETWORKS

Unrolled networks are a type of physics-free neural networks that rely solely on data without any prior knowledge or assumptions about the underlying physical principles that govern it. However, the challenge with unrolled networks is that they may struggle to accurately reconstruct images when there are significant differences between the test and the training sensor data. Such discrepancies can arise due to changes in the environment, equipment, anomalies/errors, or inherent noise, thereby limiting the network's ability to generalize well beyond the training data. An additional significant challenge arose due to the scarcity of the training set, significantly impacting the performance of the networks. To address these issues, a denoising network can be applied to the output image of the unrolled network. The denoising network is trained on a set of proxy images (which are degraded versions of the expected output images produced by the unrolled network) as input and clean high-resolution images as desired outputs. The utilization of denoising networks in image reconstruction can provide several benefits. Firstly, incomplete or noisy data utilized for image reconstruction can lead to artifacts and noise that negatively impact image quality. By applying a denoising network, such artifacts, and noise can be reduced or removed, resulting in a clearer and more precise image. Secondly, a denoising network can enhance the visual appearance of images, particularly in applications that rely on visual interpretation such as medical imaging or remote sensing.

Overall, the need for a denoising network in an unrolled deep network arises when the unrolled network is unable to reconstruct a proper image due to the variability in the input data. By applying a trained denoising network to the proxy image, the denoising network can produce a cleaner, more accurate image that is suitable for further analysis or processing. Fig. 4 illustrates the implementation process for improving image reconstruction in unrolled deep networks using a denoising network. A U-Net based denoising network, as described in Fig. 5, that accepts an input image of size 64 by 64 with an encoder depth of 3 has a total of 7 stages, which include 3 encoder stages (ES), a bridge, and 3 decoder stages (DS). Each encoder stage consists of a series of convolutional layers with increasing feature maps, followed by a max-pooling layer. The first encoder stage has the highest resolution, with 64 feature maps, while subsequent encoder stages have 128 and 256 feature maps, respectively. The bridge connects the final encoder stage to the first decoder stage and typically consists of a single convolutional layer. Each decoder stage consists of a series of upsampling layers, followed by a series of convolutional layers with decreasing feature maps. The first decoder stage has 256 feature maps, while the subsequent decoder stages have 128 and 64 feature maps, respectively. Skip connections are used to connect the corresponding encoder and decoder stages, allowing the network to reconstruct fine details and textures. The output of the final decoder stage is the denoised image with a size of 64 by 64. During training, the U-Net based denoising network is optimized to minimize the difference between the noisy input and the clean output, typically using a mean squared error loss function. Once the denoising network is trained, it can be used to denoise new images by passing them through that network. Thus starting with the sensor data as input, the unrolled network produces a proxy image which is then passed through a denoising network to generate a denoised/clean image suitable for various purposes such as visual interpretation, feature extraction, or any other downstream analysis.

IV. RESULTS

In this section, the performance of the proposed U-rBP-ADMM DL network and U-SBTV DL network with unrolled ISTA is compared to their iterative implementations. The performance and tolerance of the unrolled networks are evaluated when images are reconstructed from noisy sensor data, which is an important consideration for real-world applications. Additionally, the effectiveness of a U-net-based denoising network in enhancing the quality of reconstructed PAT images is evaluated when the test sensor data exhibits higher variability than the training data.

A. PERFORMANCE METRICS AND SYSTEM REQUIREMENTS

The following measures are used to evaluate the performance.
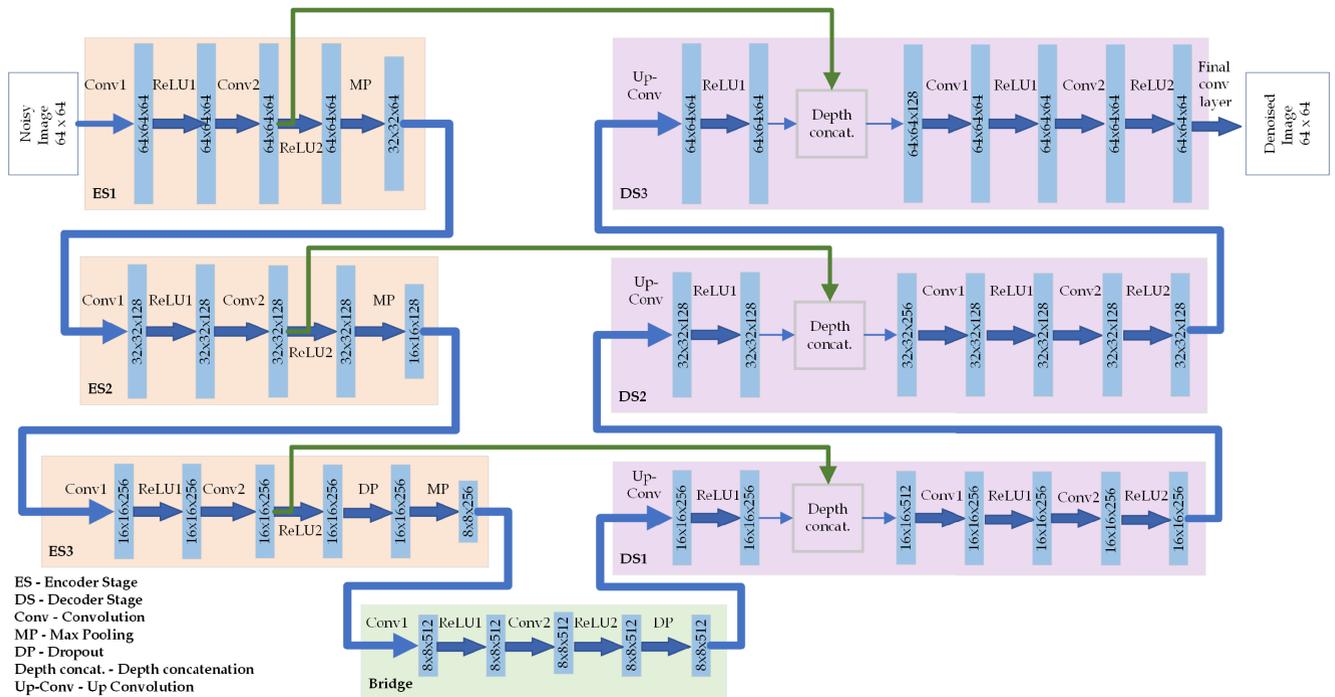
FIGURE 5. U-Net based image denoising network.

1) Structural similarity index (SSIM) [51]: This metric is used to assess the reconstructed image quality by comparing it to a reference image. Its value ranges from -1 to 1. When the reconstructed image closely resembles the original image, its value is near 1, else it will be close to 0 or less than 0. In this simulation, SSIM is computed using Matlab® built-in functions.

2) Peak signal-to-noise ratio (PSNR): It is calculated as

$$PSNR = 10 \cdot \log_{10}(\frac{m_i^2}{MSE}), \qquad (17)$$

where m_i is the maximum possible pixel value of the image and MSE is the mean-squared error between the vectorized reconstructed image (or the reconstructed image from measurements) $\hat{x}$ and the original image x. The MSE is defined as

$$MSE = \frac{||x - \hat{x}||_2^2}{N_{in}^2}. \qquad (18)$$

A higher PSNR value denotes better image quality. It is measured in decibels (dB).

3) Number of IBs: Neural network layers are added to form an IB, which reflects the depth of the DL network formed. The complexity of the network grows with the network depth.

4) Validation root mean square error (RMSE): The network's validation RMSE is also considered to measure the accuracy of the training process and hence the trained network's reconstruction capability. It is acceptable when its value is close to zero [52], [53], [54].

5) Computation time: The computation time is the amount of time needed for an algorithm or network to reconstruct an image from sensor data. The most desirable value is less than 1 second, and it is typically expressed in seconds.

B. SIMULATION SETUP

The motivation of this research is to create a network that can reconstruct an image directly from the pressure data such that the system is compatible with practical applications (such as integration with existing ultrasound imaging devices). In this sense, the simulation setup is divided into three phases. In the first phase, the photoacoustic modeling toolbox (k-Wave [55]) is used to obtain the pressure sensor data. In the second phase, the proposed unrolled DL networks are trained using the sensor data obtained from the first phase and their corresponding output images (vectorized). In the third phase, the output generated by the unrolled deep learning network is enhanced using a U-Net, which serves as a denoising network.

1) INITIAL SETUP

A computational grid, which performs the computation in cartesian mesh, with an inner grid of size 64×64 and an outer grid of size 160×160 with a grid spacing $d = 0.1mm$ is created using k-Wave Matlab® toolbox. A special boundary layer called a perfectly matched layer (PML) is used to absorb the acoustic waves that reach close to the edge of the computational domain. Sound speed is chosen to be $1500m/s$. The highest frequency that the simulation grid can propagate

TABLE 1. Unrolled ISTA DL network.

Number of IBs	Validation RMSE	Number of layers	Data inside training set		Data Outside training set	
			SSIM	PSNR (dB)	SSIM	PSNR (dB)
1	2.4864	4	0.8905	28.76	0.8070	28.27
2	1.5699	8	0.9263	34.20	0.8615	32.91
3	1.5667	12	0.9547	38.16	0.8585	33.18
4	1.5716	16	0.9522	35.25	0.8325	31.38
5	3.5118	20	0.8385	29.32	0.5335	23.41
6	64.464	24	0.0093	6.53	0.0081	4.53

TABLE 2. Unrolled SBTV DL network.

Number of IBs	Validation RMSE	Number of layers	Data inside training set		Data Outside training set	
			SSIM	PSNR (dB)	SSIM	PSNR (dB)
1	1.8108	3	0.9413	38.23	0.8336	29.64
2	0.4670	14	0.9557	41.24	0.9432	36.02
3	0.2884	26	0.9997	63.11	0.9517	36.66
4	0.1302	38	0.9997	64.21	0.9515	36.74
5	0.1773	50	0.9998	66.04	0.9511	36.71
6	0.6121	62	0.9994	61.25	0.9469	36.31
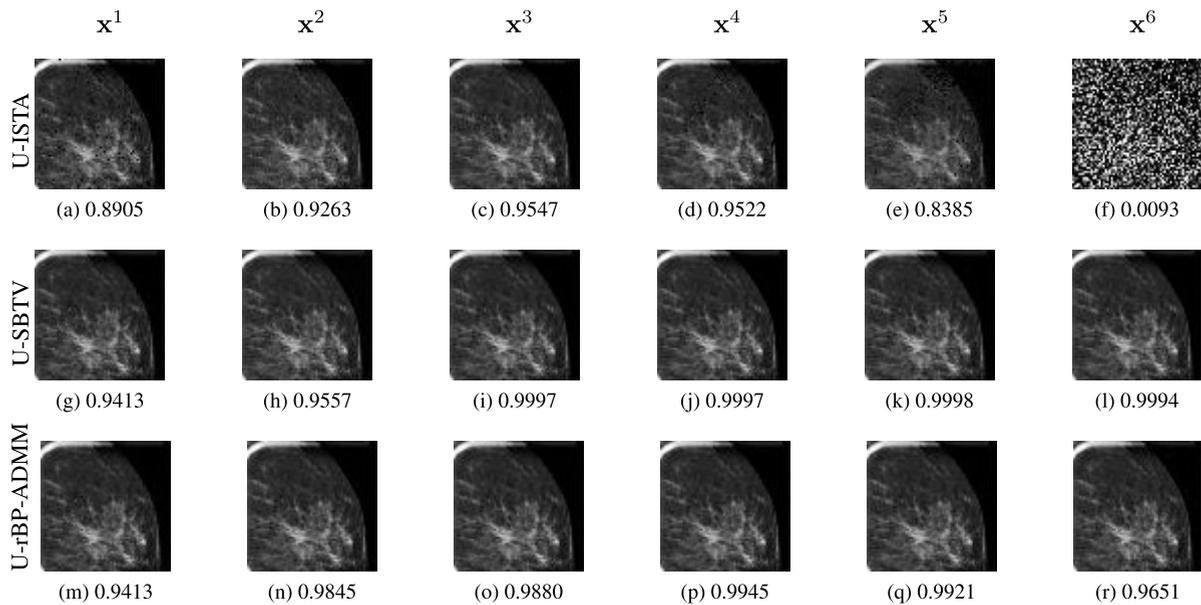
FIGURE 6. Reconstructed images, from sensor data chosen within the training set, and their corresponding SSIM, when Unrolled ISTA, SBTV, and rBP-ADMM networks are used, for iteration blocks 1 to 6 (starting from left).

depends on the grid spacing and sound speed, which is calculated as $f_m = \frac{c}{2d}$. Within the computational domain, a binary sensor mask is defined for recording the pressure field at each time step. Pressure data vector is acquired for 5μs at a rate of 15MHz which results in a total number of 75-time samples. The simulations use a total of 71 sensors, which are arranged in a square grid around the object. The processor used in the Matlab® simulation is Intel®Core™ i7-4790 CPU@3.60GHz.

2) DL NETWORK SETUP
Initially, the pressure data vector is obtained for 206 input images using the simulation setup in the first phase.

The training data set is generated by pairing pressure data vectors and image vectors. The unrolled DL network created using the architecture of iterative algorithms is then trained using Adam optimizer [56] with a learning rate of 0.001. The validation frequency is chosen as 50 and the maximum number of epochs is set to 500 with a mini-batch size of 32.

C. PERFORMANCE COMPARISON OF UNROLLED NETWORKS
In this subsection, the numerical results of the performance of the trained unrolled DL networks are discussed and compared.

Tab. 1 shows the numerical results of the U-ISTA DL network. First, we analyze the validation RMSE versus the number of IBs. It is shown that the validation RMSE value is minimum with 3 IBs. IBs less than 3, correspond to model underfitting. Whereas, IBs of more than 3 correspond to

TABLE 3. Unrolled rBP-ADMM DL net.

Number of IBs	Validation RMSE	Number of layers	Data inside training set		Data Outside training set	
			SSIM	PSNR (dB)	SSIM	PSNR (dB)
1	1.8108	3	0.9413	38.23	0.8336	29.64
2	0.3046	11	0.9845	45.84	0.9433	36.26
3	0.2113	22	0.9980	54.34	0.9522	36.71
4	0.2550	33	0.9945	48.79	0.9499	36.58
5	0.3143	44	0.9921	46.08	0.9493	36.50
6	0.5222	55	0.9651	41.85	0.9275	35.35
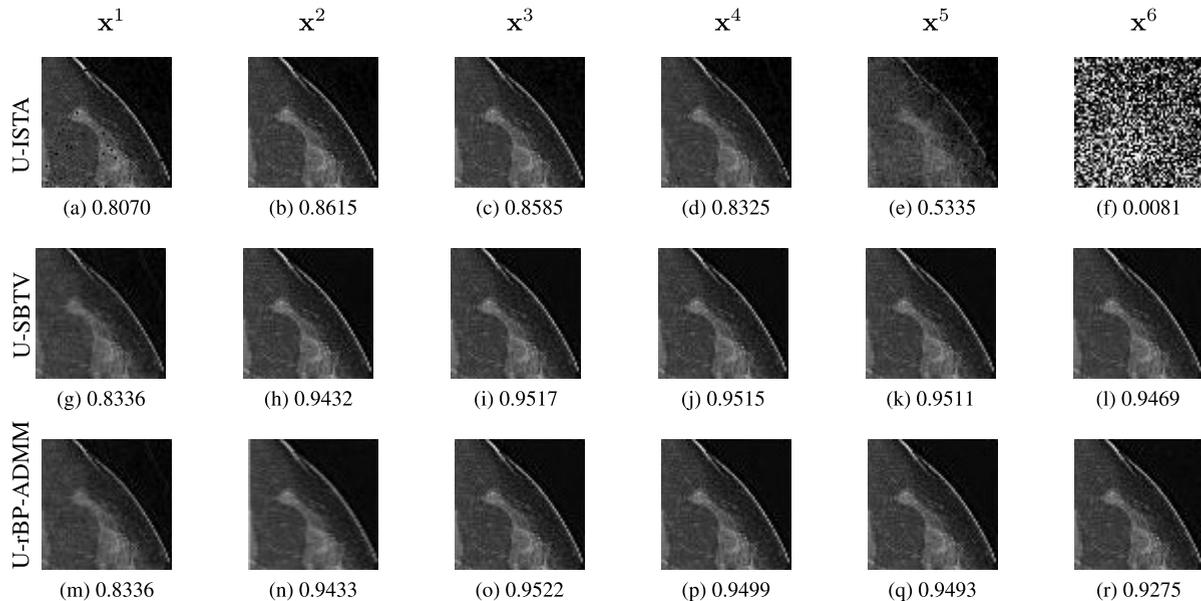
FIGURE 7. Reconstructed images, from sensor data chosen outside the training set, and their corresponding SSIM, when Unrolled ISTA, SBTV, and rBP-ADMM networks are used, for iteration blocks 1 to 6 (starting from left).

TABLE 4. Noisy image reconstruction results - using test data from training set.

Network used	SNR=10 dB		SNR=15 dB		SNR=20 dB	
	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)
U-ISTA	0.7760	28.71	0.9006	33.36	0.9358	35.73
U-SBTV	0.7976	31.07	0.9089	36.17	0.9627	40.88
U-rBP-ADMM	0.8044	31.17	0.9080	36.14	0.9654	40.89

model overfitting. Likewise, the SSIM and PSNR metrics are optimal or maximum; 0.9547 and 38.16 dB respectively, with 3 IBs given the test data chosen from the training set. However, the performance is slightly different, when the test data is chosen from outside the training set, where the optimal SSIM of 0.8615 is observed when the number of IBs is 2 (slightly better SSIM compared to when the number of IBs is 3) though the optimal PSNR of 33.18 dB is observed when the number of IBs is 3. Fig. 6(a-f) and Fig. 7(a-f) show the reconstructed images that were obtained using a different number of IBs.

The simulation results of the U-SBTV DL network are shown in Tab. 2. The validation RMSE is seen to change as the number of IBs goes from 1 to 6. The RMSE drops from 1.8108 to 0.1302 when 4 IBs are used and rises to 0.6121 when 6 IBs are used. The SSIM is almost the same for the number of IBs 3, 4, and 5, for both cases of test

data (0.999 and 0.951 respectively). In terms of PSNR, when the test data is chosen from inside the training set, the best performance is obtained at 5 IBs. Whereas for the case of test data chosen from outside the training set, the performance is the same for 3, 4, and 5 IBs (approximately 36 dB). The reconstructed images produced using the U-SBTV DL network when the number of IBs is increased from 1 to 6 are shown in Fig. 6(m-r) and Fig. 7(m-r).

The results of the U-rBP-ADMM DL network analysis are presented in Tab. 3. The validation RMSE is seen to drop when the IBs are increased from 1 to 3, then increases as the IBs are increased from 4 to 6. The use of three IBs yields the lowest validation RMSE, with an optimal result of SSIM 0.9980 and PSNR of 54.34 dB for an image reconstructed using data from inside the training set. Correspondingly, the use of three IBs yields the best results for image reconstruction from data outside the training set,

with SSIM of 0.9522 and PSNR of 36.71 dB. Fig. 6(g-l) and Fig. 7(g-l) show images that were obtained when U-rBP-ADMM DL network is used for reconstruction for different number of IBs.

When comparing the three developed unrolled networks, it is observed that unrolled SBTV DL network exhibits the lowest validation RMSE of 0.1305 when 4 IBs (38 layers) are used. Moreover, SBTV exhibits the highest SSIM of 0.9998 and the highest PSNR value of 66.04 dB, for reconstruction from test data inside the training set. The superiority of the unrolled SBTV-DL network over the unrolled rBP-ADMM network can be attributed to two key factors: TV regularization and the split Bregman formulation. Total variation regularization enables the network to leverage sparsity, preserving essential structural information while suppressing noise and artifacts. The split Bregman formulation enhances the optimization process by decoupling the problem into smaller subproblems, leading to faster convergence and improved denoising performance. By combining sparsity exploitation and efficient optimization, the unrolled SBTV-DL network achieves superior denoising results compared to the unrolled rBP-ADMM network, providing higher overall performance in preserving image structures and suppressing noise. However, the maximum comparable performance is attained by U-rBP-ADMM DL network for reconstruction when test data are chosen from outside the training set, with an SSIM of 0.9522 and a PSNR of 36.7 dB. Thus, U-SBTV and U-rBP-ADMM DL networks can be used for PAT image reconstruction, where SBTV unrolled network is considered at the cost of slightly higher complexity because of the additional number of layers.

D. PERFORMANCE COMPARISON OF UNROLLED NETWORKS WITH NOISY SENSOR DATA

When data is acquired using ultrasound sensors, thermal noise is likely to be present. A fundamental noise model that simulates the effects of several random processes that take place in sensing is additive white Gaussian noise (AWGN). In the following noisy pressure data is considered, where WGN is added with a signal-to-noise ratio (SNR) = 10, 15, and 20 dB. The noisy pressure data are used for image reconstruction (not for training), where the same model is used for testing with the original pressure data. The results are presented in Tab. 4 and Tab. 5.

The SSIM and PSNR values are shown to increase as the SNR increases. U-SBTV and U-rBP-ADMM DL networks perform similarly for both sets of test data when SNR is 15 and 20 dB. However, U-rBP-ADMM performs somewhat better in the scenario where the SNR=10 dB. Thus, both of the unrolled networks that are proposed can be employed, albeit U-rBP-ADMM may be preferred in low SNR scenarios.

E. PERFORMANCE COMPARISON OF UNROLLED NETWORKS AND ITERATIVE ALGORITHMS

In this subsection, the performance of the proposed unrolled DL networks is compared with the performance of the iterative algorithms (Algorithms 1, 2, and 3 are also presented in our earlier work [33]). In the implementation of ISTA, the regularization parameter $\lambda = 0.008$ and the Lipschitz constant $L = 2$. In the implementation of the SBTV algorithm, the value of $\beta = 1200$ and $\gamma = 40$. For rBP-ADMM algorithm implementation, $\rho = 1$, $\alpha = 1$, $\gamma = 0.2$, and the relaxation factor $r = 0.001$. The algorithm stopping criteria is when the normalized mean square error (NMSE) becomes less than the tolerance value (set as 10^{-6}) or when the maximum number of iterations is reached ($kmax = 150$). Various values of r were investigated for performance comparison, and it was observed that values in the order of 10^{-3} yielded superior results. Furthermore, an accelerated variant of the rBP-ADMM algorithm can be adopted, wherein the value of γ is updated in each iteration such that $\gamma^{k+1} = \mu\gamma^k$ with typical values of μ in the range of [0.98, 0.99], for accelerated performance. This adaptation accelerates the algorithm and concurrently decreases computation time. When applying an acceleration μ of 0.99 to γ, which is the soft thresholding limit value, it improves the performance by increasing the SSIM from 0.9935 to 0.9941 (0.0604% increase) and PSNR from 48.61 to 50.00 (2.90% increase). More importantly, the number of iterations reduced from 201 to 121 (39.80% decrease) which thereby reduced the computation time from 35.47 to 21.13 seconds (40.43% decrease). Thus, employing an acceleration to the parameter γ allowed to further exploit the advantage of the rBP-ADMM algorithm for PAT image reconstruction.

ISTA achieves an SSIM of 0.9326 with 120 iterations and a computation time of more than 20 seconds, while U-ISTA achieves SSIMs of 0.9547 and 0.8585 with 3 IBs, respectively, for both cases of selecting the test data, with reconstruction times of about 1 second (excluding training time). With 11 iterations and a computation time of 5 seconds, SBTV achieves an SSIM of 0.9900. U-SBTV has an SSIM of 0.9997 and 0.9517 when using three IBs, and selecting test data from both inside and outside the training set, respectively, with a reconstruction time of 1 second. In comparison, the U-rBP-ADMM achieves SSIMs of 0.9980 and 0.9522 at 3 IBs, and 1 second, for test data chosen from inside the training set and outside it, respectively, while the rBP-ADMM does so by using 50 iterations and a computation time of 6 seconds to reach an SSIM of 0.9898. Overall, it has been observed that ISTA performs much worse for both iterative algorithms and unrolled networks. Additionally, iterative ISTA requires more time and iterations than SBTV and rBP-ADMM. The results are shown in Tab. 6. Unrolled-I refers to testing images chosen from inside the training set, while unrolled-O pertains to testing with images from outside the training set. These distinctions allow us to evaluate the model's performance in familiar and unfamiliar data scenarios, providing insights into its generalization ability.

Although the performances of U-ISTA and U-rBP-ADMM are practically identical to those of their iterative method

TABLE 5. Noisy image reconstruction results - using test data outside training set.

Network used	SNR=10 dB		SNR=15 dB		SNR=20 dB	
	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)
U-ISTA	0.7168	28.61	0.8020	31.18	0.8404	32.45
U-SBTV	0.7922	30.45	0.8923	33.69	0.9313	35.64
U-rBP-ADMM	0.8063	30.95	0.8937	33.77	0.9324	35.55
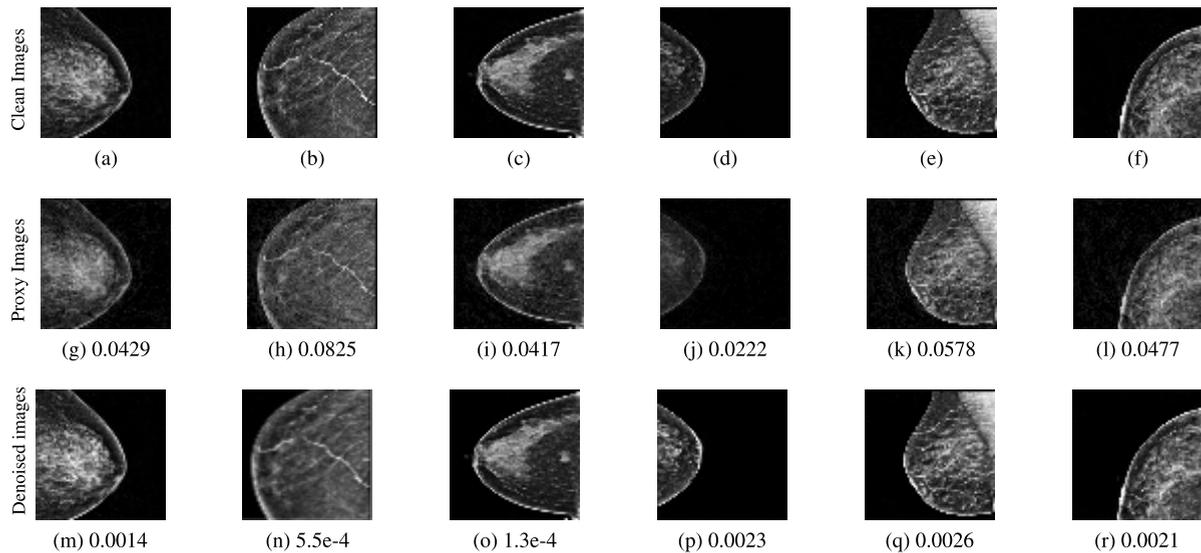
FIGURE 8. Clean images ((a-f)), proxy images ((g-l) produced by an unrolled deep learning network) and the corresponding denoised images ((m-r) produced by U-Net denoising network), with the corressponding MSE values.

TABLE 6. Comparison of SSIM, number of iteration (k) and computation (comp.) time (approximately 1 second for unrolled networks) between iterative algorithms and unrolled networks.

Method used	Iterative Algorithm SSIM; k; Comp. time	Unrolled-I SSIM; k;	Unrolled-O SSIM; k;
ISTA	0.9326; k=120; >20 sec	0.9547; k=3	0.8585; k=3
SBTV	0.9900; k=11; 5 sec	0.9997; k=3	0.9517; k=3
rBP-ADMM	0.9898; k=50; 6 sec	0.9980; k=3	0.9522; k=3

TABLE 7. Improvements in SSIM and MSE when a denoising network is used on a proxy image.

Dataset used	Proxy vs. Clean		Denoised vs. Clean		SSIM imp.%	MSE red.%
	SSIM	MSE	SSIM	MSE		
Training set	0.6931	0.0429	0.9685	0.0014	39.7	96.7
	0.6866	0.0825	0.9816	5.5e-4	43.0	99.4
	0.6809	0.0417	0.9816	1.3e-4	44.2	99.7
	0.5211	0.0500	0.9572	2.4e-4	83.7	99.5
Test set	0.4649	0.0559	0.8029	0.0025	72.7	95.5
	0.7061	0.0222	0.9190	0.0023	30.2	89.6
	0.6316	0.0578	0.8983	0.0026	42.2	95.5
	0.5427	0.0477	0.9093	0.0021	67.6	95.6

implementation, the real-time computation time savings is the advantage when comparing the unrolled algorithms for both scenarios of selecting the test data. This is due to the fact that iterative algorithms must do all computations in real-time, but unrolled networks are trained offline and require around 1 second for reconstruction. The substantial pre- and post-processing associated with iterative techniques are reduced when neural networks are used for reconstruction, which is a notable benefit of the unrolled networks, and is well reflected in the reconstruction time. The thresholds are also implicitly learned rather than explicitly established, allowing the network to be more adaptable to reconstruction when pressure data from outside the training set are used to produce an image.

F. IMPACT OF DENOISING NETWORKS ON IMAGE RECONSTRUCTION ACCURACY OF UNROLLED NETWORKS

A series of experiments were conducted to evaluate the effectiveness of denoising networks in improving image reconstruction accuracy on a test dataset of degraded images obtained from the unrolled deep network. The purpose of these experiments is to determine whether the application of denoising network could help mitigate the negative impact of sensor data with high variability on image reconstruction accuracy. The denoising effectiveness is specifically assessed using SSIM, mean squared error (MSE), and the percentage reductions of MSE and improvements in SSIM, before and after denoising. In particular, we evaluated the image reconstruction accuracy of unrolled deep networks with and without denoising networks for test sensor data that exhibits high variability and differs significantly from the training sensor data. This issue is addressed using the proposed denoising network based on a U-Net architecture with an

encoder depth of three. The denoising network was trained on a dataset of 190 images with a training:test:validation ratio split of 0.7:0.15:0.15. Results presented in a Tab. 7 show a significant improvement in the quality of the processed images. The corresponding proxy images and their denoised versions are displayed in Fig. 8, with the MSE values noted below each image. For the first image, the SSIM metric showed a 39.7% improvement, increasing from 0.6931 to 0.9685, while the MSE metric decreased from 0.0429 to 0.0014, indicating a 96.7% enhancement. Tab. 6 demonstrates that the use of different training images resulted in SSIM improvements ranging from 39 to 83 percent and a decrease in MSE between 96 to 99 percent. On the other hand, for test images, the SSIM improvement was between 30 to 72 percent, and the decrease in MSE ranged from 89 to 95 percent. Overall, the results suggest that the proposed denoising network can consistently enhance the performance of the unrolled deep learning network across different test images, and thereby improve image reconstruction accuracy. The use of denoising networks can be particularly useful in scenarios where sensor data can differ significantly from the training data due to changes in the environment, equipment, anomalies/errors, or inherent noise.

V. COMPARATIVE EVALUATION AND PRACTICAL APPLICATIONS

In this section, a detailed comparative evaluation of the proposed unrolled-denoiser network with the traditional iterative reconstruction approach is provided. The goal is to assess their respective performances in the context of more realistic images containing intricate features, such as nerves or tissue layers. Additionally, we discuss the practical applications and implications of our findings.

A. VISUAL COMPARISON

To begin the comparison, we selected a practical image with a high level of complexity, representing a challenging scenario for image reconstruction methods. Both the iterative reconstruction approach and the unrolled-denoiser network were applied to this image, and the reconstructed outputs were visually compared. The results are presented in Fig. 9, where the image reconstructions by both methods are shown, along with the corresponding MSE values noted below each image. Based on the MSE analysis, we can observe that the iterative method using SBTV achieves the lowest MSE, followed by rBP-ADMM and ISTA. Conversely, for the unrolled-denoiser network, the best performance is achieved with SBTV, followed by rBP-ADMM and ISTA.

To quantitatively measure the fidelity of the reconstructed images, we employed the HaarPSI [57] metric, which offers a perceptual assessment of the similarity between the reconstructions and the ground-truth images. The HaarPSI scores for the iterative reconstruction and unrolled-denoiser network were calculated and analyzed. Notably, the iterative method achieved an impressive HaarPSI score of 1, indicating a high-quality reconstruction. Meanwhile, the unrolled
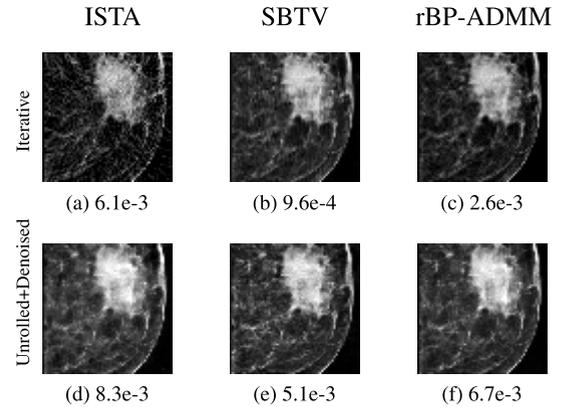
FIGURE 9. Reconstructed images of iterative and unrolled-denoiser implementations of ISTA, SBTV and rBP-ADMM with the corresponding MSE values (specified below each image).

network obtained a slightly lower but still excellent score of 0.999, demonstrating its ability to faithfully reconstruct the underlying features.

B. RECONSTRUCTION TIME AND PRACTICAL APPLICABILITY

Beyond reconstruction quality, a critical consideration for practical applications is the time and complexity involved in implementing the respective methods. The proposed unrolled-denoiser network has a clear advantage in this regard, as it requires only a training phase for deployment. On the other hand, the iterative reconstruction approach demands meticulous parameter tuning, system matrix definition, and modeling, resulting in a significantly longer and more resource-intensive implementation process.

C. IMPORTANCE AND APPLICATIONS OF THE UNROLLED-DENOISER NETWORK

The unrolled-denoiser network presents a compelling choice for real-world applications due to its visually pleasing reconstructions and minimal parameter adjustments. Its computational efficiency allows real-time image reconstruction, benefiting quick decision-making during medical procedures. Moreover, the data-driven approach enhances image quality, enabling precise diagnoses and reducing radiation exposure. The method's adaptability to diverse sensor data also allows integration with other imaging techniques, advancing medical research and patient care.

VI. CONCLUSION

In conclusion, this study proposed two efficient unrolled deep learning networks for high-resolution photoacoustic tomography image reconstruction from sensor data, which demonstrated their potential for early disease detection, particularly in breast cancer. The networks showed low validation RMSE and high SSIM and offered significant computational savings, flexibility, and adaptability compared to iterative methods. The U-Net-based enhancer network

further improved image quality for cases with high variability between test and training sensor data. Future work can extend this approach to detect and classify other types of tumors using different deep-learning networks, exploring larger datasets, and optimizing real-time applications for improved diagnostic accuracy and clinical usability. Collaborating with medical experts for validation studies and clinical trials would be essential for potential clinical implementation, benefiting breast cancer diagnosis and medical imaging practices. Furthermore, the proposed approach holds potential for real-time diagnosis and treatment monitoring in clinical settings. Overall, the significance of this work lies in the development of efficient and flexible deep-learning networks for high-resolution PAT image reconstruction from sensor data, which has the potential to improve early disease detection and clinical decision-making.

REFERENCES

[1] C. Fang, H. Hu, and J. Zou, "A focused optically transparent PVDF transducer for photoacoustic microscopy," *IEEE Sensors J.*, vol. 20, no. 5, pp. 2313–2319, Mar. 2020.

[2] J. Xia and L. V. Wang, "Small-animal whole-body photoacoustic tomography: A review," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1380–1389, May 2014, doi: 10.1109/TBME.2013.2283507.

[3] M. Li, Y. Tang, and J. Yao, "Photoacoustic tomography of blood oxygenation: A mini review," *Photoacoustics*, vol. 10, pp. 65–73, Jun. 2018, doi: 10.1016/j.pacs.2018.05.001.

[4] M. A. Naser, D. R. T. Sampaio, N. M. Muñoz, C. A. Wood, T. M. Mitcham, W. Stefan, K. V. Sokolov, T. Z. Pavan, R. Avritscher, and R. R. Bouchard, "Improved photoacoustic-based oxygen saturation estimation with SNR-regularized local fluence correction," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 561–571, Feb. 2019, doi: 10.1109/TMI.2018.2867602.

[5] Y. Zhang and L. Wang, "Video-rate ring-array ultrasound and photoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 39, no. 12, pp. 4369–4375, Dec. 2020, doi: 10.1109/TMI.2020.3017815.

[6] A. Bychkov, V. Simonova, V. Zarubin, E. Cherepetskaya, and A. Karabutov, "The progress in photoacoustic and laser ultrasonic tomographic imaging for biomedicine and industry: A review," *Appl. Sci.*, vol. 8, no. 10, p. 1931, Oct. 2018, doi: 10.3390/app8101931.

[7] W. Guo and Y. Liu, "Detection and function of elastic wave tomography of foundation piles of high-rise buildings under the background of Internet of Things," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–13, Jul. 2022, doi: 10.1155/2022/6585189.

[8] H. Liu, Y. Zhao, S.-H. Bo, and S.-L. Chen, "Application of photoacoustic imaging for lithium metal batteries," *Proc. SPIE*, vol. 11549, pp. 75–82, Oct. 2020, doi: 10.1117/12.2575184.

[9] J. Zhang, F. Duan, Y. Liu, and L. Nie, "High-resolution photoacoustic tomography for early-stage cancer detection and its clinical translation," *Radiol., Imag. Cancer*, vol. 2, no. 3, May 2020, Art. no. e190030, doi: 10.1148/rycan.2020190030.

[10] V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder, "Looking and listening to light: The evolution of whole-body photonic imaging," *Nature Biotechnol.*, vol. 23, no. 3, pp. 313–320, Mar. 2005, doi: 10.1038/nbt1074.

[11] C. G. A. Hoelen, F. F. M. de Mul, R. Pongers, and A. Dekker, "Three-dimensional photoacoustic imaging of blood vessels in tissue," *Opt. Lett.*, vol. 23, no. 8, pp. 648–650, Apr. 1998, doi: 10.1364/OL.23.000648.

[12] D. S. Kolah, A. Sajadi, A. R. Radmard, and H. Khademi, "Five common cancers in Iran," *Arch. Iranian Med.*, vol. 13, no. 2, pp. 143–146, Mar., 2010.

[13] E. J. Watkins, "Overview of breast cancer," *J. Amer. Acad. PAs*, vol. 32, no. 10, pp. 13–17, Oct. 2019, doi: 10.1097/01.JAA.0000580524.95733.3d.

[14] S. H. Heywang-Köbrunner, A. Hacker, and S. Sedlacek, "Advantages and disadvantages of mammography screening," *Breast Care*, vol. 6, no. 3, pp. 199–207, 2011, doi: 10.1159/000329005.

[15] F. Yasmin, M. M. Hassan, M. Hasan, S. Zaman, C. Kaushal, W. El-Shafai, and N. F. Soliman, "PoxNet22: A fine-tuned model for the classification of monkeypox disease using transfer learning," *IEEE Access*, vol. 11, pp. 24053–24076, 2023, doi: 10.1109/ACCESS.2023.3253868.

[16] M. M. Hassan, M. M. Hassan, F. Yasmin, M. A. R. Khan, S. Zaman, K. K. Islam, and A. K. Bairagi, "A comparative assessment of machine learning algorithms with the least absolute shrinkage and selection operator for breast cancer detection and prediction," *Decis. Anal. J.*, vol. 7, Jun. 2023, Art. no. 100245, doi: 10.1016/j.dajour.2023.100245.

[17] J. M. Valencia-Moreno, J. Á. González-Fraga, J. P. Febles-Rodríguez, and E. Gutierrez-Lopez, "Review of intelligent algorithms for breast cancer control: A Latin America perspective," *IEEE Latin Amer. Trans.*, vol. 21, no. 2, pp. 226–241, Feb. 2023, doi: 10.1109/TLA.2023.10015215.

[18] S. Kabiraj, L. Akter, M. Raihan, N. J. Diba, E. Podder, and M. M. Hassan, "Prediction of recurrence and non-recurrence events of breast cancer using bagging algorithm," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, 2020, pp. 1–5, doi: 10.1109/ICCCNT49239.2020.9225440.

[19] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006, doi: 10.1109/TIT.2006.871582.

[20] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006, doi: 10.1109/TIT.2006.885507.

[21] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006, doi: 10.1109/TIT.2005.862083.

[22] E. Candes and J. Romberg. (Oct. 2005). *L1-Magic: Recovery of Sparse Signals Via Convex Programming*. [Online]. Available: http://www.acm.caltech.edu/l1magic/downloads/l1magic.pdf

[23] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, Mar. 2008, doi: 10.1109/MSP.2007.914728.

[24] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, May 2008, Art. no. 035020, doi: 10.1088/0266-5611/24/3/035020.

[25] X. Liu and J. U. Kang, "Sparse OCT: Optimizing compressed sensing in spectral domain optical coherence tomography," *Proc. SPIE*, vol. 7904, pp. 231–237, Feb. 2011, doi: 10.1117/12.874058.

[26] E. Liu and V. N. Temlyakov, "The orthogonal super greedy algorithm and applications in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2040–2047, Apr. 2012.

[27] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová, "Probabilistic reconstruction in compressed sensing: Algorithms, phase diagrams, and threshold achieving matrices," *J. Stat. Mech., Theory Exp.*, vol. 2012, no. 8, Aug. 2012, Art. no. P08009, doi: 10.1088/1742-5468/2012/08/P08009.

[28] S. Budhiraja, "A survey of compressive sensing based greedy pursuit reconstruction algorithms," *Int. J. Image, Graph. Signal Process.*, vol. 7, no. 10, pp. 1–10, Sep. 2015, doi: 10.5815/ijigsp.2015.10.01.

[29] S. S. Kumar and P. Ramachandran, "Review on compressive sensing algorithms for ECG signal for IoT based deep learning framework," *Appl. Sci.*, vol. 12, no. 16, p. 8368, Aug. 2022, doi: 10.3390/app12168368.

[30] S. D. Howard, A. R. Calderbank, and S. J. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order Reed–Muller codes," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst.*, Mar. 2008, pp. 11–15, doi: 10.1109/CISS.2008.4558486.

[31] J. Ying, H. Lu, Q. Wei, J. F. Cai, D. Guo, J. Wu, Z. Chen, and X. Qu, "Hankel matrix nuclear norm regularized tensor completion for N-dimensional exponential signals," *IEEE Trans. Signal Process.*, vol. 65, pp. 3702–3717, Apr. 2017.

[32] M. Tao, "Convergence study of indefinite proximal ADMM with a relaxation factor," *Comput. Optim. Appl.*, vol. 77, no. 1, pp. 91–123, Sep. 2020.

[33] M. J. John and I. Barhumi, "Fast and efficient PAT image reconstruction algorithms: A comparative performance analysis," *Signal Process.*, vol. 201, Dec. 2022, Art. no. 108691, doi: 10.1016/j.sigpro.2022.108691.

[34] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, Jun. 2010, pp. 399–406.

[35] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[36] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[37] B. T. Cox and P. C. Beard, "Fast calculation of pulsed photoacoustic fields in fluids using k-space methods," *J. Acoust. Soc. Amer.*, vol. 117, no. 6, pp. 3616–3627, Jun. 2005, doi: 10.1121/1.1920227.

[38] J. Provost and F. Lesage, "The application of compressed sensing for photo-acoustic tomography," *IEEE Trans. Med. Imag.*, vol. 28, no. 4, pp. 585–594, Apr. 2009, doi: 10.1109/TMI.2008.2007825.

[39] S. Djennadi, N. Shawagfeh, and O. A. Arqub, "A fractional Tikhonov regularization method for an inverse backward and source problems in the time-space fractional diffusion equations," *Chaos, Solitons Fractals*, vol. 150, Sep. 2021, Art. no. 111127, doi: 10.1016/j.chaos.2021.111127.

[40] S. Chen and D. Donoho, "Basis pursuit," in *Proc. 28th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, 1994, pp. 41–44, doi: 10.1109/ACSSC.1994.471413.

[41] M. Grant and S. Boyd. (Mar. 2014). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1.* [Online]. Available: http://cvxr.com/cvx
This is the citation format for CVX software as obtained from the software website. Please update the reference as mentioned above.

[42] J. Han, B. Russell, and J. Downton, "Seismic spectral decomposition via basis pursuit," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2017, pp. 2174–2179, doi: 10.1190/segam2017-17632255.1.

[43] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[44] M. J. John and I. Barhumi, "Total variation algorithms for PAT image reconstruction," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Chiang Mai, Thailand, Nov. 2022, pp. 1164–1168, doi: 10.23919/APSIPAASC55919.2022.9980048.

[45] W. Ring, "Structural properties of solutions to total variation regularization problems," *ESAIM, Math. Model. Numer. Anal.*, vol. 34, no. 4, pp. 799–810, Jul. 2000.

[46] X.-L. Zhao, W. Wang, T.-Y. Zeng, T.-Z. Huang, and M. K. Ng, "Total variation structured total least squares method for image restoration," *SIAM J. Sci. Comput.*, vol. 35, no. 6, pp. B1304–B1320, Jan. 2013, doi: 10.1137/130915406.

[47] F. Ortelli and S. van de Geer, "Prediction bounds for higher order total variation regularized least squares," *Ann. Statist.*, vol. 49, no. 5, pp. 2755–2773, Oct. 2021, doi: 10.1214/21-AOS2054.

[48] A. E. Beaton and J. W. Tukey, "The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data," *Technometrics*, vol. 16, no. 2, pp. 147–185, 1974, doi: 10.2307/1267936.

[49] F. Sha, L. Saul, and D. Lee, "Multiplicative updates for nonnegative quadratic programming in support vector machines," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 15, Dec. 2002, pp. 1–11.

[50] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends® Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[51] V. Bruni and D. Vitulano, "An entropy based approach for SSIM speed up," *Signal Process.*, vol. 135, pp. 198–209, Jun. 2017, doi: 10.1016/j.sigpro.2017.01.007.

[52] Y.-Q. Wang and N. Limare, "A fast C++ implementation of neural network backpropagation training algorithm: Application to Bayesian optimal image demosaicing," *Image Process. Line*, vol. 5, pp. 257–266, Sep. 2015, doi: 10.5201/ipol.2015.137.

[53] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, "Evolutionary extreme learning machine," *Pattern Recognit.*, vol. 38, no. 10, pp. 1759–1763, Oct. 2005, doi: 10.1016/j.patcog.2005.03.028.

[54] G. Lemaire, C. Francois, K. Soudani, D. Berveiller, J. Pontailler, N. Breda, H. Genet, H. Davi, and E. Dufrene, "Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3846–3864, Oct. 2008, doi: 10.1016/j.rse.2008.06.005.

[55] B. E. Treeby and B. T. Cox, "K-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.*, vol. 15, no. 2, 2010, Art. no. 021314, doi: 10.1117/1.3360308.

[56] Z. Zhang, "Improved Adam optimizer for deep neural networks," in *Proc. IEEE/ACM 26th Int. Symp. Quality Service (IWQoS)*, Banff, AB, Canada, Jun. 2018, pp. 1–2, doi: 10.1109/IWQoS.2018.8624183.

[57] R. Reisenhofer, S. Bosse, G. Kutyniok, and T. Wiegand, "A Haar wavelet-based perceptual similarity index for image quality assessment," *Signal Process., Image Commun.*, vol. 61, pp. 33–43, Feb. 2018, doi: 10.1016/j.image.2017.11.001.

MARY JOHN received the B.S. degree in electronics and communication engineering and the M.S. degree in optoelectronics and communication engineering from the Cochin University of Science and Technology, India, in 2013 and 2015, respectively. She is currently pursuing the Ph.D. degree in electrical and communication engineering with United Arab Emirates University, Al Ain, United Arab Emirates. Her research interests include signal processing, inverse imaging, image reconstruction, compressive sensing, and machine learning.

IMAD BARHUMI (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from Birzeit University, Birzeit, Palestine, in 1996, the M.Sc. degree in telecommunications from The University of Jordan, Amman, Jordan, in 1999, and the Ph.D. degree in electrical engineering from Katholieke Universiteit Leuven (KUL), Leuven, Belgium, in 2005. From 1999 to 2000, he was with the Department of Electrical Engineering, Birzeit University, as a Lecturer. After the Ph.D. graduation, he was a Postdoctoral Research Fellow for one year with the Department of Electrical Engineering, KUL. He is currently an Associate Professor with the Department of Electrical Engineering, United Arab Emirates University, Al Ain, United Arab Emirates. His research interests include signal processing for mobile and wireless communications, cooperative communications, resource allocation and management in wireless communications and networking, game theory, and compressive sensing.

• • •