MMV-Net: A Multiple Measurement Vector Network for Multi-frequency Electrical Impedance Tomography

Citation for published version:
Chen, Z, Xiang, J, Bagnaninchi, PO & Yang, Y 2022, 'MMV-Net: A Multiple Measurement Vector Network for Multi-frequency Electrical Impedance Tomography', *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1-12. https://doi.org/10.1109/TNNLS.2022.3154108

Digital Object Identifier (DOI):
10.1109/TNNLS.2022.3154108

Published In:
IEEE Transactions on Neural Networks and Learning Systems

MMV-Net: A Multiple Measurement Vector Network for Multi-frequency Electrical Impedance Tomography

Zhou Chen, *Student Member, IEEE*, Jinxi Xiang, *Student Member, IEEE,* Pierre Bagnaninchi, and Yunjie Yang, *Member, IEEE*

***Abstract*—Multi-frequency Electrical Impedance Tomography (mfEIT) is an emerging biomedical imaging modality to reveal frequency-dependent conductivity distributions in biomedical applications. Conventional model-based image reconstruction methods suffer from low spatial resolution, unconstrained frequency correlation and high computational cost. Deep learning has been extensively applied in solving the EIT inverse problem in biomedical and industrial process imaging. However, most existing learning-based approaches deal with the single-frequency setup, which is inefficient and ineffective when extended to the multi-frequency setup. This paper presents a Multiple Measurement Vector (MMV) model based learning algorithm named MMV-Net to solve the mfEIT image reconstruction problem. MMV-Net considers the correlations between mfEIT images and unfolds the update steps of the Alternating Direction Method of Multipliers for the MMV problem (MMV-ADMM). The non-linear shrinkage operator associated with the weighted $l_{2,1}$ regularization term of MMV-ADMM is generalized in MMV-Net with a cascade of a Spatial Self-Attention module and a Convolutional Long Short-Term Memory (ConvLSTM) module to better capture intra- and inter-frequency dependencies. The proposed MMV-Net was validated on our *Edinburgh mfEIT Dataset* and a series of comprehensive experiments. The results show superior image quality, convergence performance, noise robustness and computational efficiency against the conventional MMV-ADMM and the state-of-the-art deep learning methods.**

***Index Terms*—Deep learning, Electrical Impedance Tomography (EIT), multi-frequency, Multiple Measurement Vector (MMV), image reconstruction.**

I. Introduction

BIO-IMPEDANCE as an indicator of the physiological status of biological tissues varies with frequency. Electrical Impedance Tomography (EIT) is a non-intrusive and non-destructive imaging modality for revealing the cross-sectional conductivity distribution from a sequence of boundary current injections and induced differential voltage measurements [1]. The EIT inverse problem reconstructs the conductivity inside the Region Of Interest (ROI) based on the voltage measurements. EIT is fast, low-cost, portable, non-intrusive, label-free and radiation-free, making it an up-and-coming candidate in biomedical imaging. Emerging applications include functional lung imaging [2], stroke diagnosis [3], [4], and biological tissue imaging [5], [6]. As the impedance spectra of biological tissues are frequency-dependent, differences in electrical properties between various tissues can be exploited to benefit physiological and pathological diagnostics for tissue differentiation, early cancer detection, and tumor or stroke imaging [6], [7]. This motivates the development of multi-frequency EIT (mfEIT) [8], which measures the bio-impedances under different frequencies of interest and employ them to reconstruct a set of multi-frequency conductivity images related to tissue properties.

The mfEIT-image-reconstruction problem concerns the simultaneous reconstruction of multiple conductivity images of selected frequencies. The problem is fundamentally challenging because of its non-linearity, severe ill-posedness, sensitivity to modelling and measurement errors, and high computational cost. However, existing literature mainly focuses on advancing mono-frequency EIT image reconstruction algorithms based on the Single Measurement Vector (SMV) model [9]–[13]. The effort in developing effective and efficient mfEIT image reconstruction algorithms has been relatively limited. A straightforward way is to recover the image under each frequency individually using the SMV-based methods while ignoring the correlations among conductivity images over the selected frequencies. In contrast, the Multiple Measurement Vector (MMV) model [14] as an extension of SMV considers the mfEIT measurements as a whole and can simultaneously reconstruct multiple conductivity images of given frequencies with higher quality by exploiting the inherent correlation between multi-frequency images. Specifically, it enforces the joint-sparsity constraint to each pixel within the ROI to exploit the spatial structure. The hypothesis is that solutions for all frequencies have common profiles but unnecessarily similar magnitudes. To optimize the MMV model, the MMV-ADMM algorithm adopts the extensively utilized Alternating Direction Method of Multipliers (ADMM) [15], [16], which has been introduced in hyperspectral imaging [17] and multi-frequency Electrical Capacitance Tomography (ECT) [18]. Alternatively, Liu *et al.* [19] and Xiang *et al.* [20] extended the SMV-based Sparse Bayesian Learning (SBL) framework based on the MMV model for multi-frequency tomographic imaging. However, the MMV-SBL approaches rely heavily on the sparsity assumption, leading to considerably degraded performance in non-sparse scenarios. These existing MMV algorithms

Z. Chen and Y. Yang are with the Intelligent Sensing, Analysis and Control Group, Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh, UK, EH9 3JL (E-mail: y.yang@ed.ac.uk).

J. Xiang is with the the Department of Precision Instrument, Tsinghua University, Beijing 100084, China.

P. Bagnaninchi is with the Centre for Regenerative Medicine, Institute for Regeneration and Repair, The University of Edinburgh, Edinburgh EH16 4UU. (E-mail: Pierre.Bagnaninchi@ed.ac.uk).

Manuscript received xx, 2021.

generally require manual determination of regularization terms and fine-tuning of hyper-parameters, such as penalty parameters and step sizes. In addition, the computational cost is considerable, preventing their wide adoption in biomedical applications that desire real-time imaging capability.

Recently, deep learning has proven its effectiveness for medical tomographic image reconstruction with noteworthy improvement of image quality and speed [21]. Deep neural networks composed of a stack of layers automatically learn complicated functions from large-scale training datasets without requiring manually designed priors. Learning-based methods are also desirable for real-time imaging due to the faster execution time against conventional model-based algorithms. Existing learning-based EIT image reconstruction algorithms can be classified into three categories [22]: (a) fully learning approaches that directly map measurement data to a conductivity image [23]; (b) image post-processing approaches that employ a trained network to eliminate artifacts of a preliminary conductivity image obtained from model based algorithms [24], [25]; (c) model-based deep learning approaches that unroll a finite number of iterations of the model based methods into a network, e.g. FISTA-Net [22], MoDL [26], ADMM-CSNet [27] and ISTA-Net [28].

The main drawback of generic deep networks in (a) and (b) via end-to-end training is the lack of interpretability and the underlying structures neglect the physical processes of the inverse problem. In contrast, the unrolling approaches in (c) are capable of incorporating the advantages of physical model-based methods and deep neural networks. However, the state-of-the-art model-based deep learning approaches in EIT only address the mono-frequency image reconstruction problem. Model-based learning for mfEIT image reconstruction remains an open problem. Therefore, this paper focuses on developing an effective model-based deep learning approach to solve the mfEIT inverse problem. We aim to exploit intra- and inter-frequency dependencies to improve multi-frequency image quality. The proposed approach, named as MMV-Net, unrolls the MMV-ADMM algorithm into a single pipeline (see Fig. 2). MMV-Net is composed of multiple blocks, each of which corresponds to one iteration. The non-linear shrinkage operation in each block is approximated and generalized by a deep network. MMV-Net is fundamentally different from ISTA-Net [29] and FISTA-Net [30]. Although MoDL, ADMM-CSNet and MMV-Net have a similar structure based on ADMM, MoDL and ADMM-CSNet are exclusively designed for SMV-based imaging. In contrast, MMV-Net goes beyond to provide simultaneously multiple conductivity images constrained by frequency correlations to promote image quality. The main contributions of this paper are as follows:

1) A novel model-based deep learning approach is proposed for simultaneous mfEIT image reconstruction. The proposed MMV-Net tackles the inherent limitations of the conventional MMV-ADMM approach. Parameters across all iteration blocks are shared and learned through end-to-end training.
2) A dedicated network is developed to substitute the non-linear shrinkage operator, which learns a more general regularizer to incorporate spatial and frequency correlations between mfEIT images. The unique design could boost the reconstruction performance of mfEIT.
3) The proposed MMV-Net has much fewer parameters than the state-of-the-art model-based learning approaches, e.g. MoDL [26] and FISTA-Net [22] (by 12.8 and 8.5 times, respectively), making it easier to train even with a smaller dataset.
4) The *Edinburgh mfEIT Dataset* is constructed for mfEIT image reconstruction. The dataset mimics tissue engineering applications and comprises $4\times 12,414$ randomly generated multi-object, multi-conductivity phantoms at four distinct frequencies.
5) MMV-Net is thoroughly evaluated on the *Edinburgh mfEIT Dataset* and various real-world experiments, and achieves superior reconstruction quality, convergence performance, and noise robustness, compared to the state of the art, such as MMV-ADMM [18], MoDL [26], and FISTA-Net [22].

The remainder of this paper is organized as follows. In Section II, we present the problem formulation of mfEIT and elaborate the proposed MMV-Net. Section III describes the *Edinburgh mfEIT Dataset*, experiment implementation, and simulation and experimental results. Section IV draws conclusions and discusses future work.

II. METHODOLOGY

A. Multi-frequency EIT

Consider difference imaging of mfEIT [31], voltage changes $\Delta\mathbf{V} \in \mathbb{R}^{m\times l}$ at a set of excitation frequencies $\{f_1, f_2, \ldots, f_l\}$ are measured to reconstruct the conductivity changes $\Delta\boldsymbol{\sigma} \in \mathbb{R}^{n\times l}$. The MMV model [14] of mfEIT linearly approximates the relationship between $\Delta\mathbf{V}$ and $\Delta\boldsymbol{\sigma}$ by

$$\Delta\mathbf{V} = \mathbf{A}\Delta\boldsymbol{\sigma} \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ $(m \ll n)$ denotes the sensitivity matrix.

We define $\Delta\mathbf{V} = [\Delta\mathbf{v}_{f_1}, \Delta\mathbf{v}_{f_2}, \ldots, \Delta\mathbf{v}_{f_l}]$, where $\Delta\mathbf{v}_{f_i} \in \mathbb{R}^{m\times 1}(i = 1, \ldots, l)$ denotes the i^{th} column of $\Delta\mathbf{V}$. The first method to obtain voltage changes leverages Time-Difference (TD) measurements [32], i.e. $\Delta\mathbf{v}_{f_i} = \mathbf{v}_{f_i}(t_1) - \mathbf{v}_{f_i}(t_0)$, which requires mfEIT measurements at two time instants t_0 and t_1, i.e. $\mathbf{v}_{f_i}(t_1)$, and $\mathbf{v}_{f_i}(t_0) \in \mathbb{R}^{m\times 1}$. Another prevailing method is to utilize Frequency-Difference (FD) measurements [33]. The FD approach employs voltage changes at different frequencies, i.e. $\Delta\mathbf{v}_{f_i} = \mathbf{v}_{f_i} - \mathbf{v}_{f_0}$, where f_0 is the reference frequency.

For convenience, we use $\mathbf{B}$ and $\mathbf{X}$ as substitutes for $\Delta\mathbf{V}$ and $\Delta\boldsymbol{\sigma}$ respectively in the rest of the paper. Typically, the MMV model-based mfEIT image reconstruction problem can be solved by addressing the constrained optimization problem:

$$\begin{cases} \min\limits_{\mathbf{X}} \mathcal{R}(\mathbf{X}) \\ s.t.\, \mathbf{A}\mathbf{X} = \mathbf{B} \end{cases} \tag{2}$$

where $\mathcal{R}(\cdot)$ denotes the regularization function, which encodes the *a priori* knowledge of the conductivity distribution $\mathbf{X}$. The

MMV model uses the weighted $l_{2,1}$ regularization to promote the joint sparsity [18], [34]:

$$\begin{cases} \min_{\mathbf{X}} \|\mathbf{X}\|_{w,2,1} := \sum_{i=1}^{n} w_i \|\mathbf{X}_i\|_2 \\ s.t. \, \mathbf{AX} = \mathbf{B} \end{cases} \tag{3}$$

where $w_i (i = 1, \ldots, n)$ is a positive scalar and $\mathbf{X}_i$ is the i^{th} row of $\mathbf{X}$.

B. MMV-ADMM for mfEIT image reconstruction

The MMV-based mfEIT-image-reconstruction problem in (3) can be efficiently solved using the classic Alternating Direction Method of Multipliers (ADMM) [15], [16]. By introducing an auxiliary vector $\mathbf{Z} \in \mathbb{R}^{n \times l}$, the problem in (3) is equivalent to

$$\begin{cases} \min_{\mathbf{X},\mathbf{Z}} \|\mathbf{Z}\|_{w,2,1} \\ s.t. \, \mathbf{Z} = \mathbf{X}, \mathbf{AX} = \mathbf{B}. \end{cases} \tag{4}$$

The augmented Lagrangian problem of (4) is

$$\min_{\mathbf{X},\mathbf{Z}} \|\mathbf{Z}\|_{w,2,1} - \mathbf{\Lambda}_1^{\mathrm{T}}(\mathbf{Z} - \mathbf{X}) + \frac{\beta_1}{2}\|\mathbf{Z} - \mathbf{X}\|_2^2 - \mathbf{\Lambda}_2^{\mathrm{T}}(\mathbf{AX} - \mathbf{B}) + \frac{\beta_2}{2}\|\mathbf{AX} - \mathbf{B}\|_2^2 \tag{5}$$

where $\mathbf{\Lambda} = \{\mathbf{\Lambda}_1 \in \mathbb{R}^{n \times l}, \mathbf{\Lambda}_2 \in \mathbb{R}^{m \times l}\}$ are Lagrangian multipliers and $\beta = \{\beta_1, \beta_2 > 0\}$ are penalty parameters. The ADMM is then applied to solve (5) through the following steps:

$$\begin{cases} \arg\min_{\mathbf{X}} \mathbf{\Lambda}_1^{\mathrm{T}}\mathbf{X} + \frac{\beta_1}{2}\|\mathbf{Z} - \mathbf{X}\|_2^2 - \mathbf{\Lambda}_2^{\mathrm{T}}\mathbf{AX} \\ \qquad + \frac{\beta_2}{2}\|\mathbf{AX} - \mathbf{B}\|_2^2, \\ \arg\min_{\mathbf{Z}} \|\mathbf{Z}\|_{w,2,1} - \mathbf{\Lambda}_1^{\mathrm{T}}\mathbf{Z} + \frac{\beta_1}{2}\|\mathbf{Z} - \mathbf{X}\|_2^2, \\ \mathbf{\Lambda}_1 \leftarrow \mathbf{\Lambda}_1 - \gamma_1\beta_1(\mathbf{Z} - \mathbf{X}), \\ \mathbf{\Lambda}_2 \leftarrow \mathbf{\Lambda}_2 - \gamma_2\beta_2(\mathbf{AX} - \mathbf{B}), \end{cases} \tag{6}$$

where $\gamma = \{\gamma_1, \gamma_2 > 0\}$ are step lengths for the Lagrangian multipliers.

The first step in (6) is a convex quadratic problem, which has a closed-form solution:

$$\mathbf{X} = (\beta_1\mathbf{I} + \beta_2\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}(\beta_1\mathbf{Z} - \mathbf{\Lambda}_1 + \beta_2\mathbf{A}^{\mathrm{T}}\mathbf{B} + \mathbf{A}^{\mathrm{T}}\mathbf{\Lambda}_2) \tag{7}$$

where $\mathbf{I} \in \mathbb{R}^{n \times n}$ is an identity matrix.

To avoid large matrix inversion and reduce the computation cost, we in this work adopt the gradient descent method as a substitute for (7):

$$\mathbf{X} \leftarrow \mathbf{X} - \eta\nabla_{\mathbf{X};\mathbf{Z},\mathbf{\Lambda}_1,\mathbf{\Lambda}_2} \tag{8}$$

where η is the step size, and $\nabla_{\mathbf{X};\mathbf{Z},\mathbf{\Lambda}_1,\mathbf{\Lambda}_2}$ is the gradient of the first step in (6) with respect to $\mathbf{X}$ given $\{\mathbf{Z}, \mathbf{\Lambda}\}$, which is defined by:

$$\nabla_{\mathbf{X};\mathbf{Z},\mathbf{\Lambda}_1,\mathbf{\Lambda}_2} = (\beta_1\mathbf{I} + \beta_2\mathbf{A}^{\mathrm{T}}\mathbf{A})\mathbf{X} - (\beta_1\mathbf{Z} - \mathbf{\Lambda}_1 + \beta_2\mathbf{A}^{\mathrm{T}}\mathbf{B} + \mathbf{A}^{\mathrm{T}}\mathbf{\Lambda}_2) \tag{9}$$

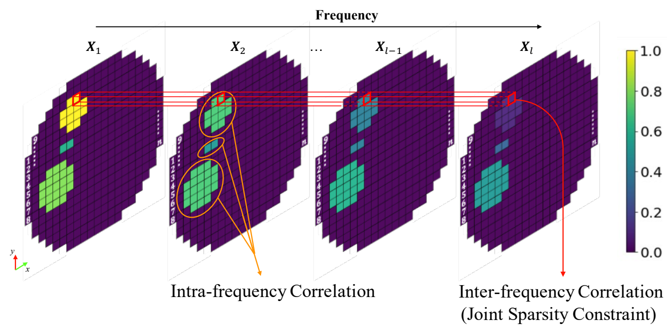

Fig. 1. Illustration of intra- and inter-frequency correlation between mfEIT images.

Thus, to solve (4) with ADMM, we have the following updates at the k^{th} iteration:

$$\begin{cases} \mathbf{X}^{(k)} = \mathbf{X}^{(k-1)} - \eta\nabla_{\mathbf{X}^{(k-1)};\mathbf{Z}^{(k-1)},\mathbf{\Lambda}_1^{(k-1)},\mathbf{\Lambda}_2^{(k-1)}}, & \text{(10a)} \\ \mathbf{Z}^{(k)} = \mathcal{S}(\mathbf{X}^{(k)} + \frac{1}{\beta_1}\mathbf{\Lambda}_1^{(k-1)}, \frac{w}{\beta_1}), & \text{(10b)} \\ \mathbf{\Lambda}_1^{(k)} = \mathbf{\Lambda}_1^{(k-1)} + \gamma_1\beta_1(\mathbf{X}^{(k)} - \mathbf{Z}^{(k)}), & \text{(10c)} \\ \mathbf{\Lambda}_2^{(k)} = \mathbf{\Lambda}_2^{(k-1)} + \gamma_2\beta_2(\mathbf{B} - \mathbf{AX}^{(k)}), & \text{(10d)} \end{cases}$$

where $k \in \{1 \ldots K_s\}$ is the iteration index, and $\mathcal{S}(\cdot)$ represents a row-wise shrinkage operator associated with the weighted $l_{2,1}$ regularization:

$$\mathbf{Z}_i^{(k)} = \max\left\{\left\|\mathbf{X}_i^{(k)} + \frac{1}{\beta_1}(\mathbf{\Lambda}_1)_i^{(k-1)}\right\|_2 - \frac{w_i}{\beta_1}, 0\right\} \cdot \frac{\mathbf{X}_i^{(k)} + \frac{1}{\beta_1}(\mathbf{\Lambda}_1)_i^{(k-1)}}{\left\|\mathbf{X}_i^{(k)} + \frac{1}{\beta_1}(\mathbf{\Lambda}_1)_i^{(k-1)}\right\|_2}, for \ i = 1, \ldots, n. \tag{11}$$

In this work, we propose to approximate $\mathcal{S}(\cdot)$ through the data-driven method. Instead of directly applying (11) to only promote the joint sparsity, MMV-Net aims to learn a more general $\mathcal{S}(\cdot)$ to exploit both the spatial correlation and the inherent correlation across different frequencies (see illustration in Fig. 1), to simultaneously reconstruct the multi-frequency conductivity distributions effectively and efficiently.

C. MMV-Net for mfEIT image reconstruction

For the mfEIT-image-reconstruction problem, the conventional ADMM (MMV-ADMM) has limitations in three respects: a) typically it requires hundreds of iterations to achieve the optimum, which degrades the computational efficiency to a significant extent [17], [18]; b) the non-linear shrinkage operator $\mathcal{S}(\cdot)$ is only valid for specific image patterns (e.g. sparsity [35], group sparsity [18]); c) it is non-trivial to fine-tune the algorithm parameters $\{\beta, \gamma, \eta\}$.

To address the above issues, we propose a deep architecture named MMV-Net for mfEIT image reconstruction by unrolling the iterative MMV-ADMM algorithm. MMV-Net combines the model-based method and the deep neural network for mfEIT image reconstruction to exploit advantages from both sides. We map the four update procedures in (10a)-(10d) into an unfolded data flow with K_s iterations as illustrated in Fig. 2.

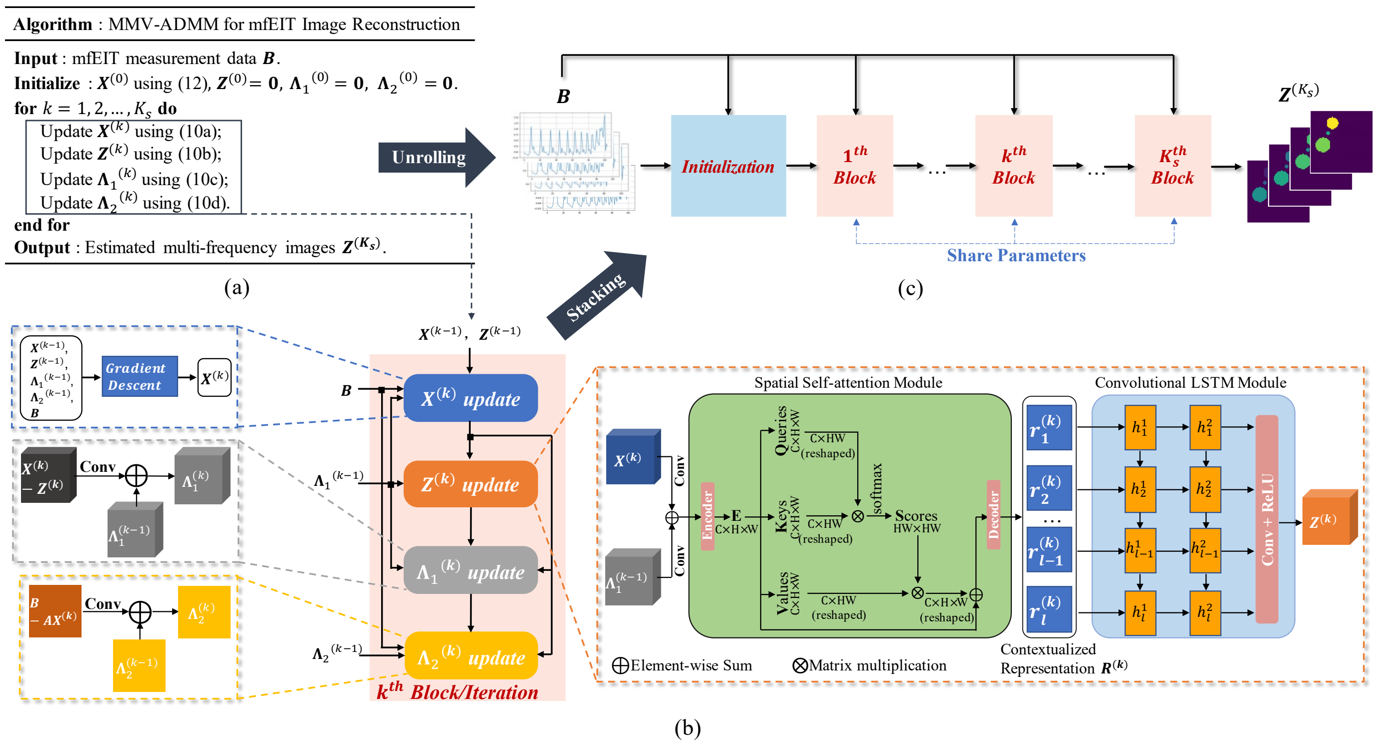

Fig. 2. (a) The iterative MMV-ADMM algorithm. (b) Illustration of the four updating steps at the k^{th} iteration corresponding to (10a)-(10d). (c) Overall architecture of the proposed MMV-Net. MMV-Net is an unrolled architecture for K_s iterations. It alternates among the four update steps. These update steps share parameters across all iterations.

The input of the MMV-Net is the mfEIT measurement data $\mathbf{B}$. $\mathbf{Z}^{(0)}$, $\mathbf{\Lambda}_1^{(0)}$, and $\mathbf{\Lambda}_2^{(0)}$ are initialized as zeros. $\mathbf{X}^{(0)}$ is obtained by employing the one-step Gaussian Newton solver with the Laplacian filter [36]:

$$\mathbf{X}^{(0)} = (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \lambda\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B} \tag{12}$$

where $\mathbf{L}$ is the Laplacian matrix, and λ is the regularization factor. The initialization results are then utilized to generate the final multi-frequency conductivity images $\mathbf{Z}^{(K_s)}$ after K_s iterations. The subsequent part of MMV-Net comprises of K_s blocks, where the k^{th} block corresponds to the k^{th} iteration of the MMV-ADMM algorithm. Each block consists of four update steps corresponding to one iteration of MMV-ADMM in (10a)-(10d), including gradient descent ($\mathbf{X}$) update, auxiliary variable ($\mathbf{Z}$) update, the first multiplier ($\mathbf{\Lambda}_1$) update, and the second multiplier ($\mathbf{\Lambda}_2$) update. The following parts discuss the four update steps at the k^{th} iteration in detail.

*1) $\mathbf{X}^{(k)}$ **update**:* This update step implements the gradient descent method. It generates the immediate result $\mathbf{X}^{(k)}$. Given $\mathbf{X}^{(k-1)}$, $\mathbf{Z}^{(k-1)}$, $\mathbf{\Lambda}_1^{(k-1)}$, and $\mathbf{\Lambda}_2^{(k-1)}$, which are obtained from the previous $(k-1)^{th}$ iteration, the output $\mathbf{X}^{(k)}$ is computed according to (10a).

*2) $\mathbf{Z}^{(k)}$ **update**:* This step updates the auxiliary variable by unrolling the generalized non-linear operator $\mathcal{S}(\cdot)$ in (10b), where the prior knowledge is integrated. For mfEIT, it is crucial to learn intra- and inter-image correlations simultaneously with respect to frequency channels (see Fig. 1). With this purpose, we propose to design $\mathcal{S}(\cdot)$ as a deep neural network, in particular, a cascade of a Spatial Self-Attention (SSA) module and a Convolutional Long Short-Term Memory (ConvLSTM) module. Fig. 2(b) illustrates how (10b) is mapped into a network. First, a more general combination of $\mathbf{X}^{(k)}$ and $\mathbf{\Lambda}_1^{(k-1)}$ is learned by two $1\times1\times1$ convolutional layers respectively and an element-wise sum. Afterwards, we apply the SSA and ConvLSTM to improve reconstruction performance of $\mathbf{Z}^{(k)}$. Let $Conv_1(\cdot)$ and $Conv_2(\cdot)$ be the two convolutional layers, $\mathcal{F}_{SSA}(\cdot)$ and $\mathcal{F}_{LSTM}(\cdot)$ be the function of SSA and ConvLSTM, respectively, then (10b) can be reformulated as:

$$\mathbf{Z}^{(k)} = \mathcal{F}_{LSTM}(\mathcal{F}_{SSA}(Conv_1(\mathbf{X}^{(k)}) + Conv_2(\mathbf{\Lambda}_1^{(k-1)}))) \tag{13}$$

The detailed design of the SSA and ConvLSTM is elaborated as follows.

Spatial self-attention module $\mathcal{F}_{SSA}(\cdot)$: our previous work [37] observed that by explicitly learning the structural information of the conductivity image, significant image quality improvement can be achieved in terms of spatial resolution and accuracy. This idea is inherited by using a spatial self-attention module to determine the structural information under each frequency channel to extract inter-frequency correlations.

The family of attention modules is capable of modeling long-range dependencies in natural language processing and computer vision. As a variation of attention, self-attention mechanism was firstly proposed to extract global dependencies of inputs for machine translation [38]. [39] and [40] extended the self-attention mechanism to video classification and image segmentation, respectively. Self-attention is usually inserted in a network and generates importance maps to refine the high-level feature maps. As a result, important regions can be focused on and feature representations are enriched with contextual relationships for intra-image compactness.

Inspired by the self-attention mechanism, the proposed SSA adopts the structure of a symmetric encoder–decoder network, embedded with a self-attention mechanism to the encoder output. An encoder is typically composed of a stack of convolutional layers with activation layers to extract latent feature representations. For instance, Wang *et al.* [39] and Fu *et al.* [40] employed the pretrained ResNet [41] as the backbone. He *et al.* [42] used a lightweight backbone based on ShuffleNetV2 [43] followed by a deformable context feature pyramid network to improve the adaptive capability of multiscale features. As the reconstructed mfEIT images under all frequencies share the same structure, the encoder part first introduces a 3×3 convolutional layer producing one output channel. Then, we apply two 2×2 convolutional layers with a stride of 2. Each layer is followed by a Batch Normalization layer and an ELU layer. The encoder outputs a feature representation $\mathbf{E}$ with size of $C \times H \times W$, which is fed into three 3×3 convolutional layers to generate feature maps queries $\mathbf{Q}$, keys $\mathbf{K}$ and values $\mathbf{V}$, respectively. $\mathbf{Q}$, $\mathbf{K}$ and $\mathbf{V}$ now have the size as $\mathbf{E}$. They are all reshaped to $C \times P$, where $P = H \times W$. $\mathbf{Q}$ and $\mathbf{K}$ are multiplied and fed into a softmax layer to generate a score/attention map $\mathbf{S}$ with size of $P \times P$. Afterwards $\mathbf{S}$ and $\mathbf{V}$ are multiplied and reshaped back to $C \times H \times W$. We then perform residual learning through a skip connection to $\mathbf{E}$. Finally, the decoder is applied and it comprises two 2×2 deconvolutional layers with a stride of 2, each of which is followed by a BatchNorm layer and an ELU layer. The final output is the contextualized representation of structure information $\mathbf{R}^{(k)} \in \mathbb{R}^{m \times l}$.

Convolutional LSTM module $\mathcal{F}_{LSTM}(\cdot)$: based on the structure information $\mathbf{R}^{(k)}$, we then attempt to learn the inter-frequency correlations by reconstructing the trend of the varying conductivity contrast along the frequency domain, meanwhile preserving the general structures learned from the SSA.

We view the contextualized representation $\mathbf{R}^{(k)}$ as a set of sequential images, i.e. $\{\mathbf{r}_i^{(k)}\}_{i=1}^{l}$, where $\mathbf{r}_i^{(k)} \in \mathbb{R}^m$ represents the i^{th} column of $\mathbf{R}^{(k)}$. To tackle this sequence-to-sequence (seq2seq) problem, Recurrent Neural Network (RNN) and LSTM models [44], [45] are in a dominant position in the field of deep learning. One drawback of RNNs/LSTMs is that they require considerable memory to store intermediate cell gate parameters, especially for long sequences and high dimensional inputs, on account of the usage of full connections. Though powerful enough to capture temporal correlations, the fully-connected layers raise redundancy and distortion for spatial data. In contrast, Convolutional LSTM (ConvLSTM) [46] is more computationally efficient as it replaces the fully-connected layers with convolutional layers. This operation further preserves spatial correlations with much fewer parameters and better generalization, meaning that we could employ more parameters to construct the SSA.

To learn the changes of conductivity contrast along with the frequency, we take advantage of ConvLSTMs. The proposed ConvLSTM module has a stack of multiple ConvLSTM layers with a kernel size of 3×3. We set the layer number as two by default. We finally apply an additional 1×1 convolutional layer and a ReLU layer to generate $\mathbf{Z}^{(k)}$.

3) $\mathbf{\Lambda}_1^{(k)}$ ***update:*** The multiplier update step corresponds to (10c). As shown in Fig. 2(b), the residual $(\mathbf{X}^{(k)} - \mathbf{Z}^{(k)})$ first goes through a $1 \times 1 \times 1$ convolutional layer, which is expected to learn $\gamma_1\beta_1$. Then we perform an element-wise sum operation between the residual and $\mathbf{Z}^{(k-1)}$ to obtain the output $\mathbf{\Lambda}_1^{(k)}$.

4) $\mathbf{\Lambda}_2^{(k)}$ ***update:*** Fig. 2(b) also illustrates the multiplier update according to (10d) with inputs of $\mathbf{\Lambda}_2^{(k-1)}$ and $(\mathbf{B}-\mathbf{A}\mathbf{X}^{(k)})$. Similar to the update of $\mathbf{\Lambda}_1$, we decompose this operation to a 1×1 convolutional layer to learn the product $\gamma_2\beta_2$ and an element-wise sum to generate $\mathbf{\Lambda}_2^{(k)}$.

Given a training dataset $\mathcal{D}$ with $N_{\mathcal{D}}$ pairs of samples, we define the objective function as the Mean Square Error (MSE) between the predicted images $\mathbf{Z}^{(K_s)}$ and the ground truth $\mathbf{X}^{(gt)}$:

$$\mathcal{L} = \frac{1}{N_{\mathcal{D}}} \sum_{(\mathbf{B},\mathbf{X}^{(gt)}) \in \mathcal{D}} \left\| \mathbf{Z}^{(K_s)} - \mathbf{X}^{(gt)} \right\|^2. \tag{14}$$

D. Network Training

We train the MMV-Net using PyTorch and employ Adam [47] for optimization with the batch size of 6. Similar to [22], [26]–[28], the non-linear operator and parameters $\mathbf{\Theta} = \{Conv_1(\cdot), Conv_2(\cdot), \mathcal{F}_{SSA}(\cdot), \mathcal{F}_{LSTM}(\cdot), \beta, \gamma, \eta\}$ are all learned from training data, rather than hand tuning. We employ the parameter-sharing strategy to penalize the recursive network size for effective learning, where $\mathbf{\Theta}$ of the MMV-Net are shared across all iterations. Inspired by the training approach in [26], we adopt a three-step approach for training. We first train the auxiliary variable update ($\mathbf{Z}$) to learn $\mathbf{\Theta_Z} = \{Conv_1(\cdot), Conv_2(\cdot), \mathcal{F}_{SSA}(\cdot), \mathcal{F}_{LSTM}(\cdot)\}$. Then we train the entire parameters $\mathbf{\Theta}$ for only one iteration, initialized with the previously learned parameters $\mathbf{\Theta_Z}$. The trained parameters $\mathbf{\Theta}$ with single iteration serve as a starting point of training the MMV-Net with multiple iterations.

III. Experiments and Results

A. The Edinburgh mfEIT Dataset

We constructed the *Edinburgh mfEIT Dataset* (the dataset and code will be available in www.research.ed.ac.uk/en/datasets/) to train the proposed MMV-Net. It contains multiple imaging objects with continuously varying conductivity values along four frequencies ($l = 4$) within a circular 16-electrode EIT sensor. The forward problem was solved by using COMSOL Multiphysics and Matlab. We adopt the adjacent measurement strategy [48] and a completed non-redundant measurement cycle contains $m = 104$ voltage measurements. In solving the inverse problem, we divide the circular sensing region by a 64×64 quadrate mesh, which contains $n = 3228$ pixels.

The background substance is saline with a constant conductivity of 2 S/m, which does not change with frequency. One to three circular objects are simulated with their diameters randomly determined by the uniform distribution [0.05d, 0.3d] (d is the sensor diameter). Extra constrains are imposed to

TABLE I
GROUPS OF SIMULATED CONDUCTIVITY VALUES AT DIFFERENT FREQUENCIES.

Group Index	f_1 (S/m)	f_2 (S/m)	f_3 (S/m)	f_4 (S/m)
1	0.01	0.6	1.2	1.8
2	0.4	0.6	0.8	1.0
3	0.8	1.0	1.2	1.4

TABLE II
PERFORMANCE COMPARISONS (PSNR, SSIM AND RMSE) ON EDINBURGH MFEIT DATASET.

Metrics	*Frequency Channel*	MMV-ADMM [18]	MoDL [26]	FISTA-Net [22]	MMV-Net
PSNR	1	19.3950	21.4738	21.4444	**23.7423**
	2	22.0398	24.1978	24.4338	**26.1284**
	3	24.3093	26.1114	26.6960	**28.5507**
	4	26.6077	27.0675	27.6000	**28.6125**
	Average	23.0880	24.7126	25.0435	**26.7585**
SSIM	1	0.4347	0.8467	0.8784	**0.9354**
	2	0.5266	0.8676	0.9092	**0.9312**
	3	0.6175	0.8712	0.9182	**0.9469**
	4	0.6383	0.8569	0.9092	**0.9265**
	Average	0.5543	0.8606	0.9038	**0.9350**
RMSE	1	0.1175	0.0910	0.0933	**0.0700**
	2	0.0836	0.0651	0.0644	**0.0527**
	3	0.0650	0.0526	0.0499	**0.0404**
	4	0.0549	0.0512	0.0496	**0.0407**
	Average	0.0803	0.0650	0.0643	**0.0510**
No. of learning parameters		NA	112,517	75,045	8,780

Best results are highlighted in bold.

avoid overlap within the sensing region. We then design three possible groups of increasing conductivity values associated with the four frequencies as shown in Table I, from which the changing conductivity values of target objects along frequency are assigned randomly. A distinct conductivity group is further ensured for each circular object within a phantom. This setup was adopted to simulate potential target application scenarios in tissue engineering (e.g. cell culture imaging [37]).

A total of $4 \times 12,414$ (where 4 is the number of current frequencies) pairs of voltage-conductivity samples were generated through finite element modelling simulation. Considering phantom complexity, we generated $4 \times 3k$ one-object samples, $4 \times 4k$ two-object samples, $4 \times 5,414$ three-object samples. They are partitioned into $4 \times 8,700$ training set, $4 \times 1,900$ validation set, and $4 \times 1,814$ testing set for network training.

To eliminate the influence of systematic defect, we calibrate and normalize the voltage measurements and conductivity in the dataset, following:

$$\mathbf{B} = \frac{\mathbf{V}_{mea} - \mathbf{V}_{ref}}{\mathbf{V}_{ref}}, \tag{15}$$

$$\mathbf{X} = \frac{\boldsymbol{\sigma}_{mea} - \boldsymbol{\sigma}_{ref}}{\boldsymbol{\sigma}_{ref}}, \tag{16}$$

where $\boldsymbol{\sigma}_{ref}$ and $\mathbf{V}_{ref}$ denote the reference conductivity distributions and corresponding measurement data respectively with only background substance (discussed in Section II-A); $\boldsymbol{\sigma}_{mea}$ and $\mathbf{V}_{mea}$ denote respectively the conductivity distribution and measurement with perturbations.

B. Evaluation on Simulation Data

In this sub-section, we evaluate the performance of the proposed MMV-Net using simulated mfEIT data.

1) Performance Comparison: we compare the proposed MMV-Net with three state-of-the-art image-reconstruction methods for mfEIT, i.e. MMV-ADMM [18], MoDL [26], and FISTA-Net [22] on the *Edinburgh mfEIT Dataset*. MMV-ADMM is a conventional MMV-based method with the ADMM solver that can be adjusted and applied for mfEIT image reconstruction. MoDL and FISTA-Net are model based deep learning methods targeted at single measurement vector based tomographic imaging. MoDL unrolls the traditional ADMM algorithm, while FISTA-Net is based on the FISTA framework [30]. The number of trainable parameters of the latter two model-based deep learning methods are given in the last row of Table II.

Table II shows quantitative comparison based on average Peak Signal to Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Root Mean Square Error (RMSE) on all the testing data. MMV-Net outperforms all competing approaches at all frequencies. Note that there is an explicit improvement of PSNR, SSIM and RMSE from f_1 to f_4. This is due to the higher sensitivity of these metrics to larger conductivity contrasts.

Fig. 3 demonstrates reconstructions of two simulated phantoms for qualitative comparison. MMV-ADMM can hardly reconstruct lower conductivity contrasts, especially at f_3 and f_4, whereas MoDL and FISTA-Net perform better. In contrast, the proposed MMV-Net can restore the most consistent structures/shapes and the conductivity changes along the frequency domain more smoothly. In addition, MMV-Net can distinguish fairly close objects more effectively than the other methods, which demonstrates the advantages of SSA and ConvLSTM used in MMV-Net. However, all methods failed to yield accurate conductivity values of each object. Even the best performing MMV-Net tends to assign similar values to all objects, although the shapes estimated are close to the ground truth. The potential reason is that the approximated linearization in (1) is unable to handle such non-linear circumstances, i.e. the sensitivity matrix $\mathbf{A}$ suffers from errors when interpreting conductivity levels from the measurement data.

Fig. 4 shows two typical failure cases of MMV-Net. For case 1, MMV-Net fails to reconstruct the smallest object closely positioned to two larger objects. Case 2 contains a noticeable artifact at the center of the reconstructed images of each frequency. This artifact is possibly inherited from the one-step initialization, indicating that the quality of the initial guess may also affect the final reconstruction.

2) Ablation Studies: In MMV-Net, we employ a Spatial Self-Attention (SSA) Module and a Convolutional LSTM (ConvLSTM) Module to capture intra- and inter-frequency correlations for high-performance mfEIT reconstructions. To verify the performance of the two modules, we conduct ablation studies (see Table III). MMV-Net with only the

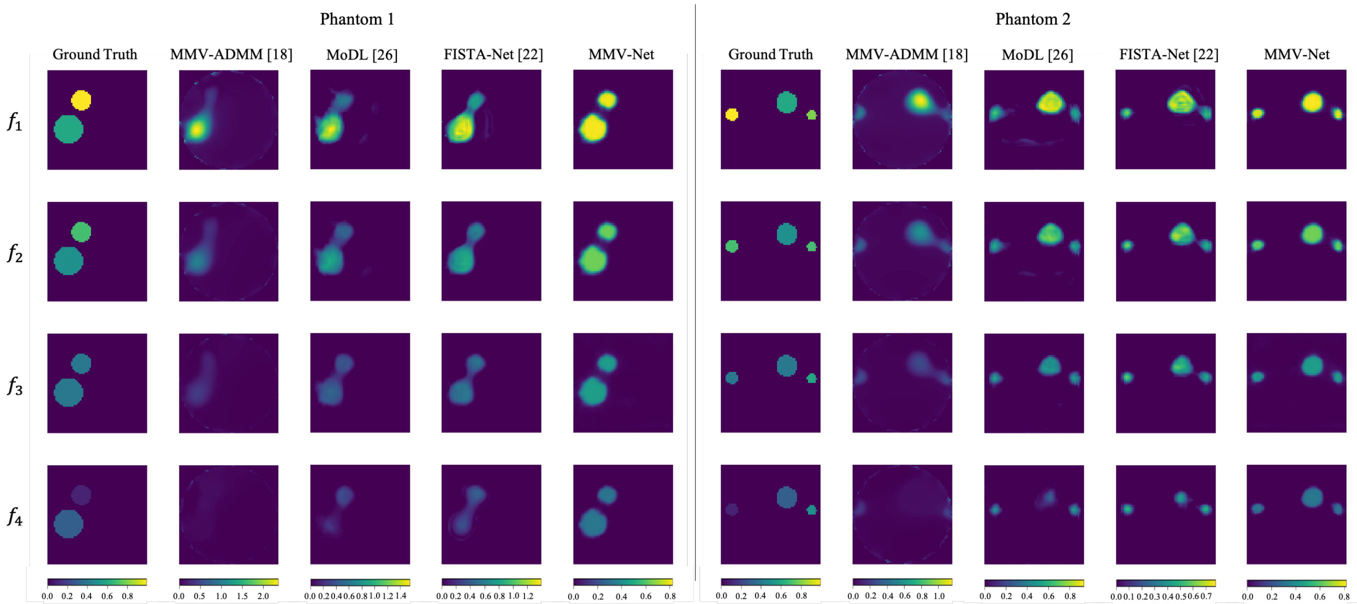

Fig. 3. Comparison of the proposed MMV-Net with the state-of-the-art imaging approaches on two simulated phantoms in the testing set.

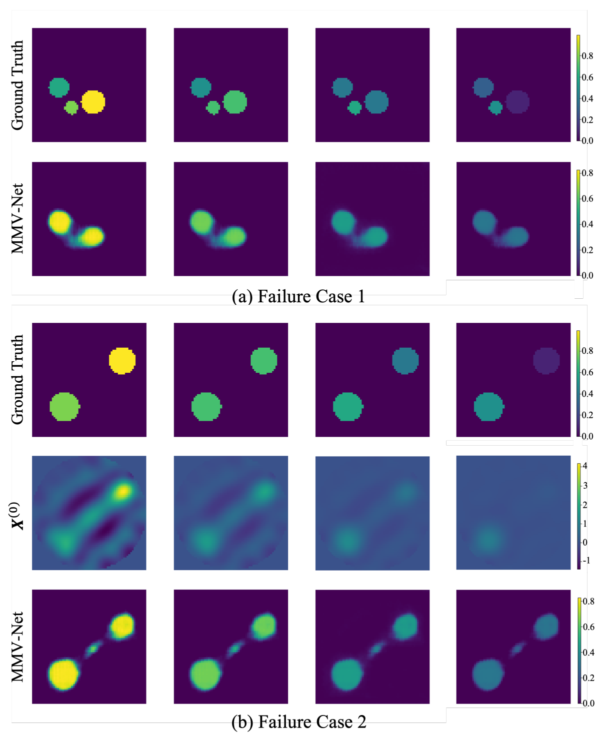

Fig. 4. Two failure cases of MMV-Net.

TABLE III
ABLATION STUDY ON THE VALIDATION SET. (SSA: SPATIAL SELF-ATTENTION MODULE; CONVLSTM: CONVOLUTIONAL LSTM MODULE.

Method	SSA	ConvLSTM	PSNR	SSIM	RMSE
	✓		24.9132	0.7640	0.0622
MMV-Net		✓	26.5144	0.9300	0.0530
	✓	✓	26.9817	0.9364	0.0498

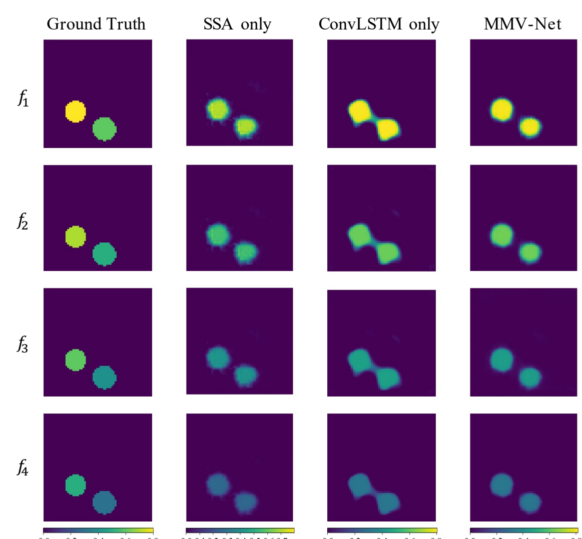

Fig. 5. An example of image reconstruction results from ablation study.

ConvLSTM module outperforms MMV-Net with individually the SSA module by 6% in PSNR, 22% in SSIM, and 15% in RMSE. Integration of the two modules into MMV-Net brings further improved performance of 26.9817 dB in PSNR, 0.9364 in SSIM, and 0.0498 in RMSE.

Fig. 5 illustrates the visual effects of the two modules. As expected, the SSA module itself manages to split the two objects and provide rough shapes but relatively vague boundaries, whereas the ConvLSTM module enhances the continuity but focuses less on shapes. By taking advantage of both modules, the proposed MMV-Net demonstrates superior performance among all the given approaches.

3) Iteration Analysis: Table IV shows the impact of the iteration number K_s. It can be observed that average PSNR,

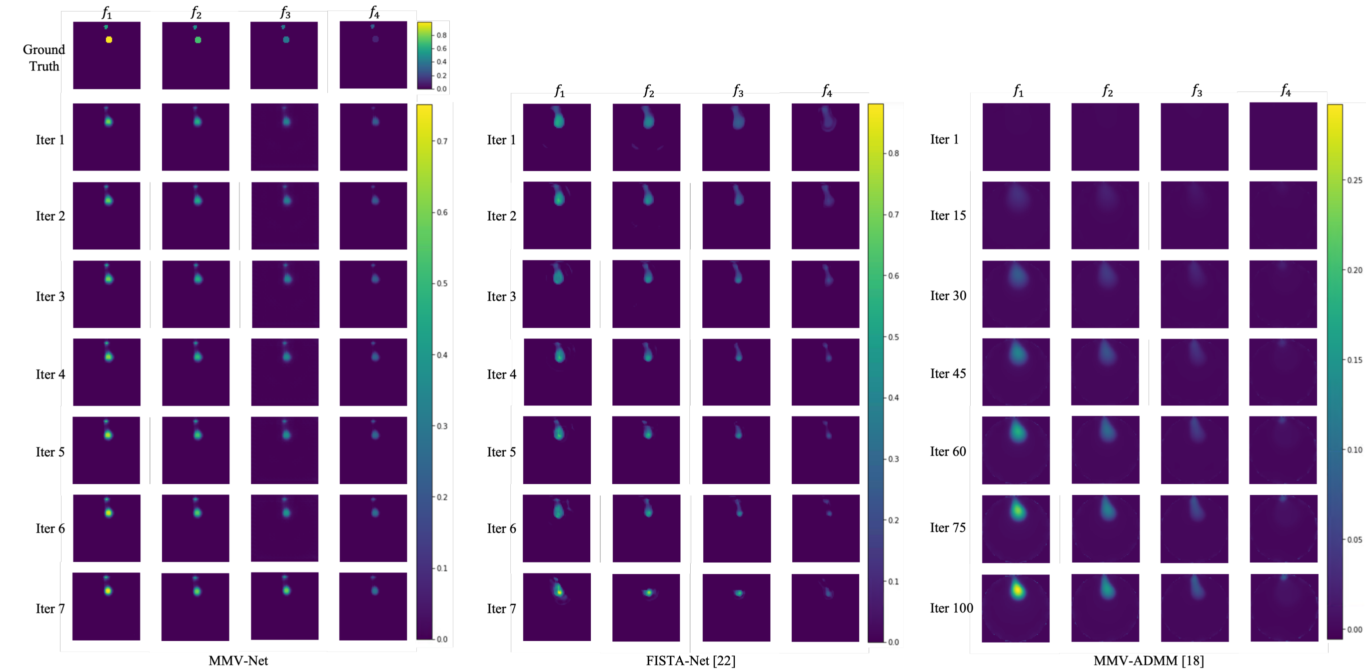

Fig. 6. Intermediate reconstructed images by MMV-Net, FISTA-Net and MMV-ADMM at different iterations.

TABLE IV
IMPROVEMENT IN RECONSTRUCTION QUALITY ON VALIDATION DATA WITH INCREASING NUMBER OF ITERATIONS OF THE NETWORK.

No. of Iterations	5	6	7	8	9
PSNR	26.8428	26.9422	26.9817	27.0174	27.0177
SSIM	0.9341	0.9342	0.9364	0.9391	0.9394
RMSE	0.0506	0.0501	0.0498	0.0496	0.0496

SSIM, and RMSE values on the validation set are improved with increasing iterations. These improvements slow down considerably when $K_s \geq 7$. Therefore, we use $K_s = 7$ for configuration as a compromise of performance and computational cost.

We show the intermediate reconstructed images of MMV-Net, FISTA-Net and MMV-ADMM at different iterations in Fig. 6. Each row corresponds to an iteration under different frequencies. The reconstruction quality of MMV-Net improves gradually with iteration. More specifically, details in the structural information are clearer and more accurate while tiny changes in conductivity values raise as it goes deeper. It might be because we put more parameters in the SSA modules to learn the structural information.

4) Generalization Ability: We demonstrate the generalization ability of MMV-Net by adding different levels of noise to the measurement data and evaluate the image quality based on PSNR. Fig. 7 shows the average PSNR values of different methods. Degraded performance can be observed for all methods, whilst the proposed MMV-Net is the most robust against noise and FISTA-Net suffers a rapid decay. MoDL and FISTA-Net are more robust than MMV-ADMM with lower noise levels (e.g. 45dB) but MMV-ADMM exceeds both MoDL and FISTA-Net when the SNR is smaller than around 37dB.

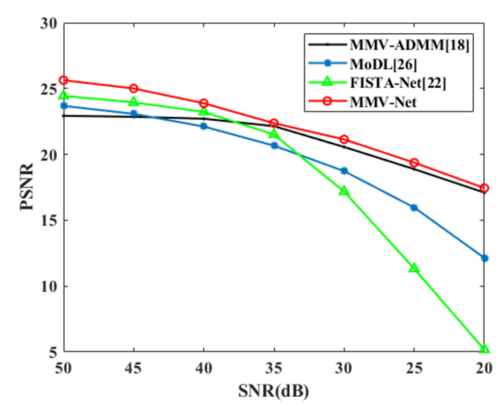

Fig. 7. Generalization ability study of different noise levels.

However, it is worth mentioning that all three learning-based methods are trained with only noise-free data. We believe our MMV-Net will retain more advantage upon MMV-ADMM and the other two model-based learning approaches should demonstrate more robustness to noise if sufficient noisy data are added in the training stage.

5) Convergence Performance: Fig. 8 illustrates the convergence performance of MMV-ADMM, MoDL, FISTA-Net, and MMV-Net. MMV-ADMM runs 100 iterations and ultimately converges to a certain level, while much faster convergence can be achieved by all network-based approaches. They adopt only 7 iterations and saturate at a much lower RMSE. We also observe that FISTA-Net shows smoother convergence than MoDL though they finally stop at similar values. MMV-Net

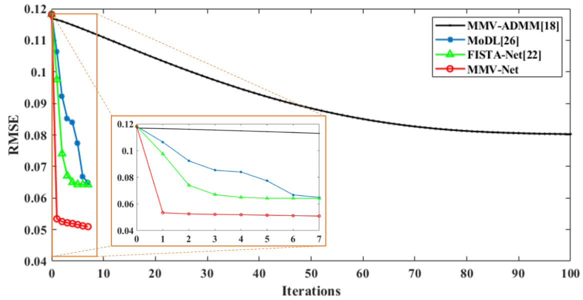

Fig. 8. Convergence performance analysis on the testing set.

TABLE V
COMPLEXITY EVALUATION. FPS MEANS FRAME PER SECOND. FLOPS MEANS FLOATING POINT OPERATIONS PER SECOND

Method	Model Size (MB)	FPS	FLOPS
MMV-ADMM [18]	-	1.39	-
MoDL [26]	0.46	0.48	12.93
FISTA-Net [22]	0.32	1.87	12.91
MMV-Net	**0.06**	**7.92**	**0.04**

Best results are highlighted in bold.

TABLE VI
NUMERICAL COMPARISONS OF EXPERIMENTAL RESULTS BASED ON SSIM.

Phantom		MMV-ADMM [18]	MoDL [26]	FISTA-Net [22]	MMV-Net
1	f_1	0.39	0.71	0.71	**0.77**
	f_2	0.41	0.69	0.73	**0.77**
	f_3	0.45	0.69	0.72	**0.76**
	f_4	0.45	0.69	0.72	**0.75**
2	f_1	0.38	0.74	0.74	**0.79**
	f_2	0.41	0.75	0.76	**0.79**
	f_3	0.44	0.75	0.75	**0.79**
	f_4	0.43	0.74	0.75	**0.79**
3	f_1	0.53	0.61	0.79	**0.80**
	f_2	0.59	0.60	0.79	**0.80**
	f_3	0.61	0.59	0.80	**0.80**
	f_4	0.55	0.59	0.81	**0.81**
4	f_1	0.27	0.70	0.77	**0.87**
	f_2	0.40	0.71	0.86	**0.87**
	f_3	0.34	0.70	0.85	**0.87**
	f_4	0.36	0.74	0.85	**0.87**

Best results are highlighted in bold.

converges even faster than MoDL and FISTA-Net.

C. Computational Complexity Evaluation

Table V compares the model size and running speed of MMV-ADMM, MoDL, FISTA-Net, and MMV-Net as indicators for computational complexity. MMV-Net has the smallest model size while obtaining the best performance. In terms of FPS, MMV-ADMM is faster than MoDL and competitive to FISTA-Net, showing the advancement of the MMV model compared to the SMV model. The proposed MMV-Net achieves the highest model execution speed of 7.92 ft/s, making it suitable for real-time imaging. MMV-Net also achieves the lowest floating-point operations, which benefits from the lightweight backbone.

D. Evaluation on Experimental Data

In addition to simulation study, we carried out real-world experiments on two different EIT sensors [37], [49] to examine the generalization ability of the proposed method. The inner diameter of the first 16-electrode EIT sensor is 94 mm. We use a potato cylinder, a sweet potato cylinder, a metal cylinder and a plastic cylinder with different conductivity values as imaging targets. Fig. 9(a)-(f) show pictures and corresponding geometric distributions of the three phantoms based on the first EIT sensor, which contain combinations of different targets. The background substance is saline with a conductivity of 0.07 $S \cdot m^{-1}$. The excitation frequencies are $\{f_1, f_2, f_3, f_4, f_5\} = \{100, 80, 50, 40, 10\}kHz$, and $10kHz$ is selected as the reference frequency. The conductivity of metal and plastic hardly changes with frequency, whilst the conductivity of potato and sweet potato increases progressively with the increase of current frequency [49]. The second miniature EIT sensor has 16 planar electrodes and an inner diameter of 15 mm (see Fig. 9(g)). The background substance is cell culture media with a conductivity of 2 $S \cdot m^{-1}$. The imaging object is a triangular MCF-7 human breast cancer cell pellet, which is less conductive than the background substance and demonstrates an increasing conductivity with the increase of current frequency.

Fig. 10 illustrates the mfEIT image reconstruction results based on experimental data. We also compare the results quantitatively based on SSIM, which is listed in Table VI. Overall, only MMV-ADMM and the proposed MMV-Net manage to consistently provide a clear trend of conductivity values with respect to frequency. MMV-Net further produces more accurate shapes and less artifacts. Note that the potato cylinder and the sweet potato cylinder in experiment phantom 1 have the same location as that in experiment phantom 2, which is most successfully recovered by the MMV-Net. MMV, MoDL and FISTA-Net fail to identify the conductive metal cylinder in experiment phantom 1, whereas MMV-Net can roughly observe the metal cylinder but the shape is underestimated. For experiment phantom 3, MMV-ADMM, FISTA-Net and MMV-Net are more noise-resistant than MoDL. However, MoDL and MMV-Net can reconstruct the non-conductive plastic cylinder. Similarly, for experiment phantom 4, MMV-Net is the most effective in inhibiting artifacts and shows more shape consistency at all frequencies. The results suggest that MMV-Net generalizes well to real-world experiments and outperforms the conventional model-based method and state-of-the-art learning approaches due to the competitive capability of capturing both intra- and inter-frequency correlations.

IV. CONCLUSION

We proposed a model-based learning approach named MMV-Net to address the simultaneous image reconstruction problem of mfEIT. MMV-Net combines the advantages of the traditional MMV-ADMM algorithm and deep networks. All parameters are learned during training, rather than manually

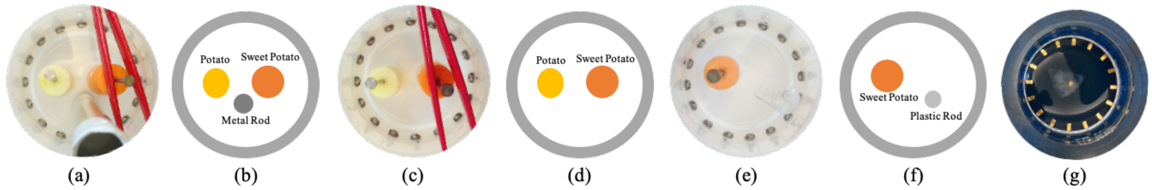

Fig. 9. Experiment phantoms using two different 16-electrode EIT sensors. (a) Phantom 1: potato rod, sweet potato rod and metal rod. (b) Geometric distribution of phantom 1. (c) Phantom 2: potato rod and sweet potato rod. (d) Geometric distribution of phantom 2. (e) Phantom 3: sweet potato rod and plastic rod. (f) Geometric distribution of phantom 3. (g) Phantom 4: MCF-7 cell pellet

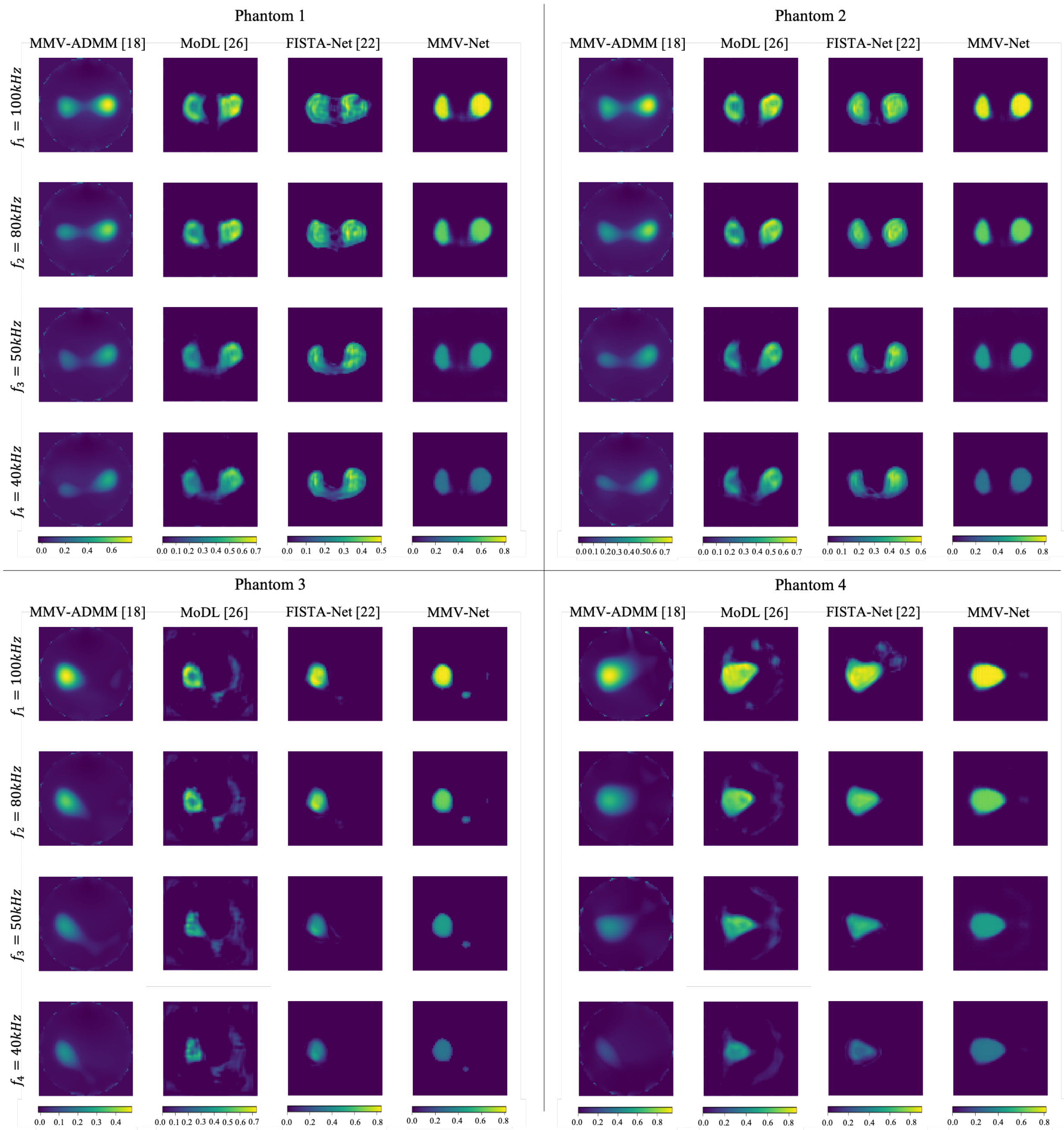

Fig. 10. mfEIT image reconstruction results of four experimental phantoms.

tuned. We generalized the regularizer of MMV-ADMM by introducing the spatial self-attention module and convolutional LSTM module to learn both spatial and frequency correlations between mfEIT images. Ablation experiments showed that cascading both modules strengthened the structural information effectively and provided superior results. Simulation and real-world experiments demonstrated that the proposed MMV-Net outperformed the state-of-the-art methods in terms of image quality, generalization ability, noise robustness and convergence performance. This work can be readily extended to solve other tomographic image reconstruction problems. The 3D version of MMV-Net will also be investigated in the near future for 3D cell culture imaging.

REFERENCES

[1] A. Adler and A. Boyle, "Electrical impedance tomography: Tissue properties to image measures," *IEEE Trans. Biomed. Eng.,* vol. 64, no. 11, pp. 2494–2504, Nov. 2017.

[2] B. Schullcke, B. Gong, S. Krueger-Ziolek, M. Soleimani, U. Mueller-Lisse, and K. Moeller, "Structural-functional lung imaging using a combined CT-EIT and a Discrete Cosine Transformation reconstruction method," *Sci. Rep.*, vol. 6, pp. 25951, May. 2016.

[3] N. Goren, J. Avery, T. Dowrick, E. Mackle, A. Witkowska-Wrobel, D. Werring, and D. Holder, "Multi-frequency electrical impedance tomography and neuroimaging data in stroke patients," *Sci. Data,* vol. 5, Jul. 2018, Art. no. 180112.

[4] B. McDermott, A. Elahi, A. Santorelli, M. O'Halloran, J. Avery, and E. Porter, "Multi-frequency symmetry difference electrical impedance tomography with machine learning for human stroke diagnosis," *Physiol. Meas.,* vol. 41, no. 7, pp. 075010,2020.

[5] A. Sujin et al., "Validation of weighted frequency-difference EIT using a three-dimensional hemisphere model and phantom," *Physiological Meas.,* vol. 32, no. 10, pp. 1663–1680, 2011.

[6] A. Adler and A. Boyle, "Electrical impedance tomography: Tissue properties to image measures," *IEEE Trans. Biomed. Eng.,* vol. 64, no. 11, pp. 2494–2504, Nov. 2017.

[7] L. Yang et al., "The frequency spectral properties of electrode-skin contact impedance on human head and its frequency-dependent effects on frequency-difference EIT in stroke detection from 10 Hz to 1 MHz," *PloS One,* vol. 12, no. 1, Jan. 2017, Art. no. e0170563.

[8] Y. Yang and J. Jia, "A multi-frequency electrical impedance tomography system for real-time 2D and 3D imaging," Rev. Sci. Instrum., vol. 88, p. 085110, Aug. 2017.

[9] B. Jin, T. Khan, and P. Maass, "A reconstruction algorithm for electrical impedance tomography based on sparsity regularization," *Int. J. Numer. Methods Eng.*, vol. 89, no. 3, pp. 337–353, 2012.

[10] Y. Yang and J. Jia, "An image reconstruction algorithm for electrical impedance tomography using adaptive group sparsity constraint," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2295–2305, Sep. 2017.

[11] S. Liu, J. Jia, Y. D. Zhang, and Y. Yang, "Image reconstruction in electrical impedance tomography based on structure-aware sparse Bayesian learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 9, pp. 2090–2102, Sep. 2018.

[12] S. Ren, K. Sun, D. Liu, and F. Dong, "A statistical shape-constrained reconstruction framework for electrical impedance tomography," *IEEE Trans. Med. Imag.,* vol. 38, no. 10, pp. 2400–2410, Oct. 2019.

[13] D. Liu, D. Gu, D. Smyl, J. Deng, and J. Du, "Shape reconstruction using Boolean operations in electrical impedance tomography," *IEEE Trans. Med. Imag.,* vol. 39, no. 9, pp. 2954–2964, Sep. 2020.

[14] J. Ziniel and P. Schniter, "Efficient high-dimensional inference in the multiple measurement vector problem," *IEEE Trans. Signal Process.,* vol. 61, no. 2, pp. 340–354, 2012.

[15] M. Fukushima, "Application of the alternating direction method of multipliers to separable convex programming problems," *Comput. Optim. Appl.,* vol. 1, no. 1, pp. 93–111, Oct. 1992.

[16] J. Wang, J. Huang, F. Zhang, and W. Wang, "Group sparse recovery in impulsive noise via alternating direction method of multipliers," *Appl. Comput. Harmon. Anal.,* vol. 49, no. 3, pp. 831-862, 2020.

[17] Q. Qu, N. M. Nasrabadi, and T. D. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation," *IEEE Trans. Geosci. Remote Sens.,* vol. 52, no. 7, pp. 4404–4423, Jul. 2014.

[18] M. Zhang, L. Zhu, H. Wang, M. Soleimani, and Y. Yang, "Multiple Measurement Vector Based Complex-Valued Multi-Frequency ECT," *IEEE Trans. Instrum. Meas.,* vol. 70, 2021.

[19] S. Liu, Y. Huang, H. Wu, C. Tan, and J. Jia, "Efficient multitask structure-aware sparse Bayesian learning for frequency-difference electrical impedance tomography," *IEEE Trans. Ind. Informat.,* vol. 17, no. 1, 2021.

[20] J. Xiang, Y. Dong, and Y. Yang, "Multi-Frequency Electromagnetic Tomography for Acute Stroke Detection Using Frequency-Constrained Sparse Bayesian Learning," *IEEE Trans. Med. Imag.,* vol. 39, no. 21, 2020.

[21] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018.

[22] J. Xiang, Y. Dong, and Y. Yang, "FISTA-Net: Learning A fast iterative shrinkage thresholding network for inverse problems in imaging," *IEEE Trans. Med. Imag.,* 2020, arXiv:2008.02683. [Online].

[23] C. Tan, S. H. Lv, F. Dong, and M. Takei, "Image reconstruction based on convolutional neural network for electrical resistance tomography," *IEEE Sensors J.*, vol. 19, no. 1, pp. 196–204, Jan. 2019.

[24] S. J. Hamilton and A. Hauptmann, "Deep D-bar: Real-time electrical impedance tomography imaging with deep neural networks," *IEEE Trans. Med. Imag.,* vol. 37, no. 10, pp. 2367–2377, Oct. 2018.

[25] Z. Wei, D. Liu, and X. Chen, "Dominant-current deep learning scheme for electrical impedance tomography," *IEEE Trans. Biomed. Eng.,* vol. 66, no. 9, pp. 2546–2555, Sep. 2019.

[26] H. K. Aggarwal, M. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.,* vol. 38, pp. 394–405, 2019.

[27] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 42, pp. 521–538, 2020.

[28] J. Zhang and B. Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," in *CVPR,* pp. 1828–1837, 2018.

[29] J. M. Bioucas-Dias and M. A. Figueiredo, "A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Signal Process.,* vol. 16, no. 12, pp. 2992–3004, 2007.

[30] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sciences,* vol. 2, pp. 183–202, 2009.

[31] D. Liu et al., "A nonlinear approach to difference imaging in EIT; Assessment of the robustness in the presence of modelling errors," *Inverse Problems,* vol. 31, no. 3, p. 035012, 2015.

[32] B. H. Brown, "Electrical impedance tomography (EIT): A review," *J. Med. Eng. Technol.,* vol. 27, no. 3, pp. 97–108, 2003.

[33] B. Harrach and J. K. Seo, "Detecting inclusions in electrical impedance tomography without reference measurements," *SIAM J. Appl. Math.,* vol. 69, no. 6, pp. 1662–1681, 2009.

[34] W. Deng, W. Yin, and Y. Zhang, "Group sparse optimization by alternating direction method," in *Proc. SPIE*, Wavelets Sparsity XV, vol. 8858. Sep. 2013, pp. 88580R, DOI: 10.1117/12.2024410.

[35] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.,* vol. 53, no. 7, pp. 2477–2488, 2005.

[36] A. Adler et al., "GREIT: A unified approach to 2D linear EIT reconstruction of lung images," *Physiol. Meas.,* vol. 30, pp. S35–S55, 2009.

[37] Z. Chen, Y. Yang, and P. Bagnaninchi, "Hybrid Learning based Cell Aggregate Imaging with Miniature Electrical Impedance Tomography," *IEEE Trans. Instrum. Meas.,* vol. 70, 2021.

[38] A. Vaswani, N. Shazeer, N. Parmar, and J. Uszkoreit, "Attention is all you need," In *NIPS,* pages 6000–6010, 2017.

[39] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *CVPR,* Jun. 2018, pp. 7794–7803.

[40] J. Fu et al., "Dual attention network for scene segmentation," in *CVPR,* Jun. 2019, pp. 3146–3154.

[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR,* 2016, pp. 770-778.

[42] Q. He, X. Sun, Z. Yan, and K. Fu, "DABNet: Deformable contextual and boundary-weighted network for cloud detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.,* early access, Jan. 5, 2021, doi: 10.1109/TGRS.2020.3045474.

[43] N. Ma, X. Zhang, H. Zheng, and J. Sun, "Shufflenet V2: Practical guidelines for efficient CNN architecture design," in *Proc. Eur. Conf. Comput. Vis.,* Sep. 2018, pp. 122–138.

[44] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *EMNLP,* 2014, pp. 1724–1734.

[45] J. Donahue, L. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell, "Long-term recurrent convolutional networks for visual recognition and description," in *CVPR,* 2015, pp. 2625–2634.

[46] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. K. Wong, and W. C. Woo, "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting," in *NIPS,* 2015.

[47] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–41.

[48] B. H. Brown and A. D. Seagar, "The Sheffield data collection system," *Clin. Phys. Physiol. Meas.,* vol. 8, no. 4A, p. 91, Nov. 1987.

[49] Y. Yang and J. Jia, "A multi-frequency electrical impedance tomography system for real-time 2D and 3D imaging," *Rev. Sci. Instrum.,* vol. 88, no. 8, 2017, Art. no. 085110.