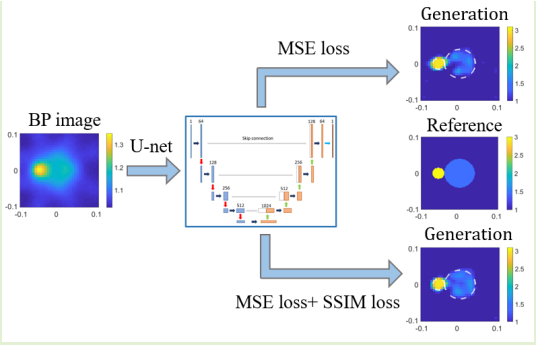
Deep Learning-Based Inverse Scattering With Structural Similarity Loss Functions

Youyou Huang, Rencheng Song, *Member, IEEE*, Kuiwen Xu, *Member, IEEE*, Xiuzhu Ye, *Senior Member, IEEE*, Chang Li, *Member, IEEE*, and Xun Chen, *Senior Member, IEEE*

***Abstract*—Deep learning based inverse scattering (DL-IS) methods attract much attention in recent years due to advantages of fast speed and high-quality reconstruction. The loss functions of neural networks in DL-IS methods are commonly based on a pixel-wise mean squared error (MSE) between the reconstructed image and its reference one. In this article, we introduce a structural similarity (SSIM) loss function to combine with the MSE loss for reconstructing dielectric targets under a DL-IS framework. The SSIM loss imposes a further regularization on the target at the perceptual level. Numerical tests for both synthetic and experimental data verify that this new perceptually-inspired loss function can effectively improve the imaging quality and the generalization capability of the trained model.**



***Index Terms*—Inverse scattering, convolutional neural network, structural similarity loss.**

I. Introduction

ELECTROMAGNETIC inverse scattering problem (ISP) is a very challenging yet fundamental problem commonly required by biomedical imaging [1], geophysics [2], through-wall imaging [3], remote sensing [4] and non-destructive testing [5] etc. In general, the ISP is highly nonlinear and ill-posed. Over years, researchers have proposed various reconstruction methods to determine the position, shape and constitutive parameters of unknown scatterers from measured scattered field. Conventional ISP methods generally retrieve the unknown parameters through a nonlinear optimization with regularization terms [6]. Typical nonlinear ISP methods include the distorted Born iterative method (DBIM) [7], the contrast source inversion (CSI) method [8], and the subspace optimization method (SOM) [9] etc. Those nonlinear inverse scattering methods achieve good performance in many applications. However, the computational costs are usually expensive and the imaging quality may degrade greatly for complex cases with high nonlinearities.

In the past few years, deep learning techniques have been widely used in image processing and computer vision. Inspired by the great success in these areas, researchers have also applied deep learning methods to solve the electromagnetic ISP in recent years [10], [11]. For example, Li *et al.* [12] proposed a 'DeepNIS' algorithm based on a complex convolutional neural network technology (CNN) to build up a mapping between the rough image got by the backpropagation (BP) [13] method and the target profile. Wei and Chen [14] proposed a dominant current scheme (DCS) to improve the quality of input image, and therefore simplified the nonlinear mapping of input-and-output pairs.

The existing results of deep learning based inverse scattering (DL-IS) methods have demonstrated significant improvements on imaging quality and speed compared to traditional nonlinear ISP methods. However, there is still a big roof to improve. It is well known that the loss function has a significant impact on the reconstruction quality. However, there is little attention

Manuscript received August 14, 2020; revised October 1, 2020; accepted October 2, 2020. Date of publication October 12, 2020; date of current version January 15, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61922075 and Grant 41901350, in part by the Provincial Natural Science Foundation of Anhui under Grant 2008085QF285, and in part by the Fundamental Research Funds for the Central Universities under Grant JZ2019HGBZ0151. The associate editor coordinating the review of this article and approving it for publication was Dr. Marko Vauhkonen. *(Corresponding author: Rencheng Song.)*

Youyou Huang, Rencheng Song, and Chang Li are with the Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China (e-mail: huangyy@mail.hfut.edu.cn; rcsong@hfut.edu.cn; changli@hfut.edu.cn).

Kuiwen Xu is with the Engineering Research Center of Smart Microsensors and Microsystems, Ministry of Education, Hangzhou Dianzi University, Hangzhou 310018, China (e-mail: kuiwenxu@hdu.edu.cn).

Xiuzhu Ye is with the School of Information and Electronics Engineering, Beijing Institute of Technology, Beijing 100081, China (e-mail: xiuzhuye@outlook.com).

Xun Chen is with the Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230026, China (e-mail: xunchen@ustc.edu.cn).

Digital Object Identifier 10.1109/JSEN.2020.3030321

paid on the loss functions in existing DL-IS methods, and the mean squared error (MSE) loss or the mean absolute error (MAE) loss is commonly taken to optimize the network parameters [14]–[17]. As known, the MAE or MSE loss enforces a pixel-wise match between the reconstructed image and its reference. On the other hand, the target images composed by dielectric parameters in ISP usually have apparent structural feature. Hence, the reconstructed image not only needs to match with the reference on a pixel-wise level but also needs to match it on the perceptual level. This is considered to reduce artifacts in the reconstruction with the help of intrinsic structural information of targets.

The structural similarity (SSIM) index [18] is a well-known metric to measure the perceptual distance of two images based on luminance, contrast and texture. Since SSIM is differentiable, it can be used as a loss function to evaluate the perceptual distance of the output image and the corresponding reference. For example, Snell *et al.* [19] employed the SSIM loss in the image generation task and they achieved a superior performance of using a single SSIM loss compared to the common MAE or MSE loss. Zhao *et al.* [20] further evaluated the performance of combining SSIM loss with the MAE loss in the image restoration task. The results showed that the hybrid loss achieved better performance over that using a single pixel-wise loss or a single perceptual loss.

The above studies on SSIM losses all belong to the field of natural image processing and they have demonstrated outstanding performance in related tasks. Inspired by these works, we investigate the role of SSIM loss to solve ISP under the DL-IS framework. We need to point out that this study is meaningful compared to existing works considering the following reasons. It aims to bring attention to alternative choices of loss functions for DL-IS reconstructions. Although the SSIM index has been employed in existing DL-IS methods [12], [14], [21] to evaluate the results, the use of SSIM as a perceptual loss in the training process has not yet been investigated for solving ISP. Since the SSIM loss can measure the perceptual distance between the reconstructed image and the target image, it is expected to enhance the imaging quality of the DL-IS method. On the other hand, the obtained model is thought to be more robust against noise because the structural similarity metric can suppress the interference of noise. It is important to explore the influence of SSIM loss for the DL-IS method under different noise levels.

Based on above reasons, in this article, we investigate the role of SSIM loss to solve ISP under a general deep learning framework. In detail, the BP method is used to obtain a rough input image from measured scattered fields. The low-resolution input image is then mapped through a U-net [22] CNN to generate a high-resolution output image that matches with the reference image. The SSIM loss is combined with the commonly-used MSE loss in the objective function to train the U-net model. This can enforce the network to further learn high-level perceptual features of targets and hence improve the reconstruction quality. We compare the inversion results of the obtained model with and without use of the SSIM loss for both synthetic and experimental data. The results show that the use of SSIM loss can effectively improve the reconstruction quality as well as the generalization capability of the trained deep model under different levels of noise.

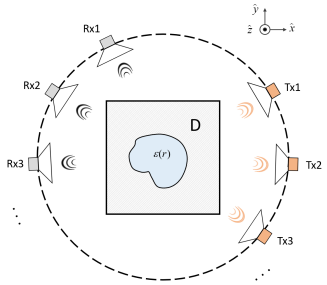


Fig. 1. Typical setups for inverse scattering problems.

II. Description of the Algorithm

A. Forward Problem

In this article, we consider a two-dimensional (2-D) transverse-magnetic (TM) [23] case of ISP. As illustrated in Fig. 1, the unknown lossless dielectric scatterers are located in the domain of interest (DOI) with a free-space background. The shape, position and the relative permittivity of the scatterers need to be determined with the measured scattered fields. The transmitters $T_{\mathrm{x}i}(i = 1, 2, \ldots, N_T)$ and the receivers $R_{\mathrm{x}i}(i = 1, 2, \ldots, N_R)$ are uniformly distributed along a circle S outside the DOI.

We use the method of moment (MOM) [24] to simulate the scattered field. Suppose the DOI is discretized into $N \times N$ grids. The governing Lippmann-Schwinger equation [6] in the DOI can be described as

$$\overline{E}^{\mathrm{tot}} = \overline{E}^{\mathrm{inc}} + \overline{\overline{G}}_D \cdot \overline{\overline{\xi}} \cdot \overline{E}^{\mathrm{tot}}, \tag{1}$$

where $\overline{E}^{\mathrm{inc}}$ and $\overline{E}^{\mathrm{tot}}$ represent the incident and total fields, respectively, and $\overline{\overline{G}}_D$ is the matrix of two-dimensional Green's function mapping the contrast source $\overline{J} = \overline{\overline{\xi}} \cdot \overline{E}^{\mathrm{tot}}$ to the scattered filed in the DOI D. The diagonal matrix $\overline{\overline{\xi}}$ indicates the permittivity contrast of the scatterer, and the nth diagonal element of $\overline{\overline{\xi}}$ is expressed as $\overline{\overline{\xi}}(n, n) = \overline{\overline{\varepsilon}}_r(r_n) - 1$, where $\overline{\overline{\varepsilon}}_r(r_n)$ is the relative permittivity at subunit r_n. The equation in (1) is also denoted as the state equation in ISP papers.

After getting the induced current in domain D, the scattered field $\overline{E}^{sca}$ at the receiver locations can be got as

$$\overline{E}^{sca} = \overline{\overline{G}}_S \cdot \overline{\overline{\xi}} \cdot \overline{E}^{\mathrm{tot}}, \tag{2}$$

where $\overline{\overline{G}}_S$ is the matrix of two-dimensional Green's function mapping the contrast source from the target area D to the scattered fields measured at the receiving antennas located at S.

The goal of ISP is to determine the relative permittivity of unknown scatterers from the observed scattered fields following (1) and (2).

B. DL-IS Method With SSIM Losses

As illustrated in Fig. 2, the full DL-IS framework reconstructs the images from scattered fields following two procedures, i.e., the non-iterative model-based inversion method

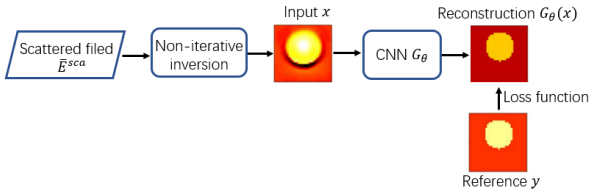


Fig. 2. The flowchart of the DL-IS method.

to map the scattered filed to a coarse input image, and the CNN to further enhance the resolution of input. It combines the advantages of the fast speed of non-iterative inversion and the powerful capability of CNN to enhance the resolution of scatterer images.

Specially, in this article, we take the BP as the non-iterative method and the U-net as the CNN. We denoted this full method as the 'BP-U-net' hereafter. It is well-known that the BP reconstruction will be distorted for strong scatterers, those with high contrast and (or) electrically large size. This shortcoming can be partially remedied by the powerful nonlinear fitting ability of U-net to build complicated one-to-one mappings between the input and output images. However, for challenging cases with strong multiple scattering effects, we may choose to construct the input image with more advanced methods such as the dominant current scheme (DCS) or the non-iterative methods based on singular value decomposition (SVD) [25].

In detail, a rough image x is firstly reconstructed from the measured scattered field $\overline{E}^{sca}$ by the BP method. The low-resolution x is then processed by the U-net G_θ to enhance the quality. The high-resolution output image $G_\theta(x)$ is matched with the ground truth y under some loss metric, where the parameters of G_θ are then updated through a gradient based optimization. Next, we will introduce the related details.

1) The Generation of Input by BP: In this article, the BP method from [13] is taken to generate the rough image x. The BP is a fast linear inversion method, which can be described as follows. First, we need to determine the induced current $\overline{J}_j$ corresponding to the jth incident wave under an assumption that $\overline{J}_j = \gamma_j \cdot \overline{\overline{G}}_S^{\mathrm{H}} \cdot \overline{E}_j^{sca}$, where the superscript H denotes the conjugate transpose operation, and γ_j is the unknown complex scaling coefficient. For simplicity, we omit the subscript j of the incident field in the following formula.

The complex γ can be explicitly obtained through minimizing the following objective function

$$F(\gamma) = ||\overline{E}^{sca} - \overline{\overline{G}}_S \cdot \gamma \cdot \overline{\overline{G}}_S^{\mathrm{H}} \cdot \overline{E}^{sca}||^2, \tag{3}$$

where γ can be solved as

$$\gamma = \frac{\left\langle \overline{E}^{sca}, \overline{\overline{G}}_S \cdot (\overline{\overline{G}}_S^{\mathrm{H}} \cdot \overline{E}^{sca}) \right\rangle_S}{\left\| \overline{\overline{G}}_S \cdot (\overline{\overline{G}}_S^{H} \cdot \overline{E}^{sca}) \right\|_S}. \tag{4}$$

Here $\langle \cdot, \cdot \rangle_S$ denotes the inner product of two vectors in the domain S.

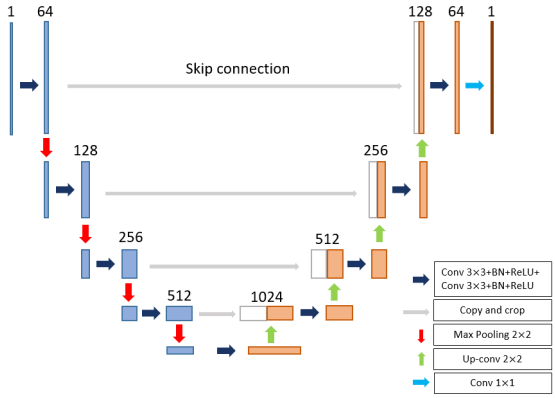


Fig. 3. The structure of the U-net CNN.

Next, the total electric field $\overline{E}^{\mathrm{tot}}$ in D can be calculated by (1). The above process is repeated for all N_T incident waves. Finally, following the relation of $J(r) = \xi(r) \cdot E^{tot}(r)$ at location r, we take into account of all the obtained total fields and contrast sources to get the contrast $\xi(r)$. This can be considered as a least-squares problem and we get the $\xi(r)$ as

$$\xi(r) = \frac{\sum_{p=1}^{N_T} J_p(r) \cdot [E_p^{tot}(r)]^{\mathrm{H}}}{\sum_{p=1}^{N_T} [E_p^{tot}(r)]^2}. \tag{5}$$

The rough input image x of relative permittivities can then be obtained using the contrasts $\xi(r)$ in (5).

2) The U-Net CNN: In this article, the well-known U-net is chosen as the CNN G_θ to build the nonlinear mapping between the rough input x and the reference image y as illustrated in Fig. 2. The structure of the U-net is shown in Fig. 3. It can be seen the U-net architecture consists of two branches. The encoding branch on the left is to extract different levels of feature maps, while the decoding branch on the right is to reconstruct the image through deconvolutions of corresponding features. The skip connections are taken to concatenate feature maps of convolutional and deconvolutional blocks from the same level. The settings of U-net are given in Fig. 3. Since it has been used in many studies, the readers can also refer to papers [14], [22] for more details.

3) The SSIM Loss Function: In order to determine the weight parameters of the neural network G_θ, we define a hybrid loss function as

$$L_{full}(\hat{y}, y) = L_{mse}(\hat{y}, y) + \alpha L_{ssim}(\hat{y}, y), \tag{6}$$

where $\hat{y} = G_\theta(x)$ is the reconstructed image by G_θ, y denotes the target image, L_{ssim} is the SSIM metric to measure the perceptual distance of two images, L_{mse} is a pixel-wise MSE loss, and α is the weight of L_{ssim} loss.

In detail, the L_{mse} loss in (6) is defined as

$$L_{mse}(\hat{y}, y) = \frac{1}{W * H} \sum_{i=1}^{W} \sum_{j=1}^{H} (\hat{y}_{i,j} - y_{i,j})^2, \tag{7}$$

where W and H indicate the width and hight of images, respectively.

The L_{ssim} in (6) is defined as

$$L_{ssim} = 1 - \mathrm{SSIM}(\hat{y}, y) = 1 - \frac{(2\mu_{\hat{y}}\mu_y + C_1)(2\sigma_{\hat{y}y} + C_2)}{(\mu_{\hat{y}}^2 + \mu_y^2 + C_1)(\sigma_{\hat{y}}^2 + \sigma_y^2 + C_2)}, \quad (8)$$

where $\mu_{\hat{y}}$ denotes the mean of $\hat{y}$, $\sigma_{\hat{y}}^2$ is the variance of $\hat{y}$, and $\sigma_{\hat{y}y}$ indicates the covariance of $\hat{y}$ and y. The C_1 and C_2 in (8) are small constants to avoid zero in the denominator, where $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$ with $K_1 = 0.01$ and $K_2 = 0.03$ as two hyperparameters, and L as the dynamic range of pixel values of the target image y.

The SSIM index measures the perceptual distance of two images considering the similarity of luminance, contrast and structure information. This is because the $\mu_{\hat{y}}$ and $\sigma_{\hat{y}}$ can be taken as estimates of the luminance and contrast of $\hat{y}$, respectively, and $\sigma_{\hat{y}y}$ represents the structural similarity of $\hat{y}$ and y.

It should be noted that the SSIM loss is actually implemented on patches of image. Suppose the full image is divided into N_p patches with each patch occupying $M \times M$ pixels. Then the loss function defined for the full image can be written as

$$L_{full}(\hat{y}, y) = \sum_{j=1}^{N_p} L_{mse}(\hat{y}_j, y_j) + \alpha L_{ssim}(\hat{y}_j, y_j), \quad (9)$$

where $\hat{y}_j$ and y_j are the jth patch of the images $\hat{y}$ and y, respectively.

Therefore, in order to improve the quality of reconstruction, the U-net will be optimized with the full loss L_{full} defined in (9), which combines the pixel-wise L_{mse} and the perceptual L_{ssim} losses.

III. Experiments

In this section, we choose three examples including both synthetic and experimental data to verify the effectiveness of the SSIM loss in the reconstruction. We will compare the performance of BP-U-net using the combined L_{full} loss (denoted as 'BP-U-net+L_{full}') with the one using only the pixel-wise L_{mse} loss (denoted as 'BP-U-net+L_{mse}'). Meanwhile, we further verify the pros and cons of the presented method compared to the existing deterministic algorithms. The subspace optimization method (SOM) is chosen for comparison and more details of SOM can be found in [26].

A. Training Details

In the proposed method, the handwritten digits in Mixed National Institute of Standards and Technology database (MNIST) are used as the training data and validation data for all examples. The MNIST database has been widely used in deep learning based inverse scattering (DL-IS) research [12], [15], [21]. Since the main purpose of this article is to investigate the role of SSIM loss to DL-IS methods, we choose similar training data sets as the exiting works. In general, the number of training samples depends on the complexity of the network we use. The more complex the neural network is, the larger number of training samples should be used to train

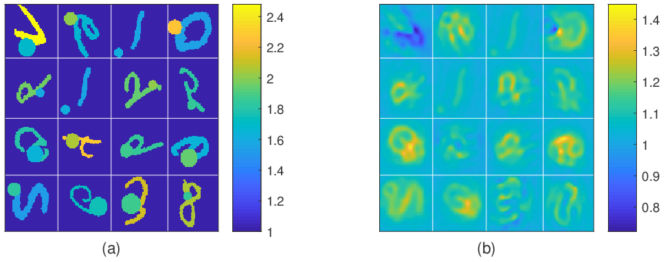


Fig. 4. Training samples of digit-like objects. (a) 16 ground truths (b) 16 input images by the BP method.

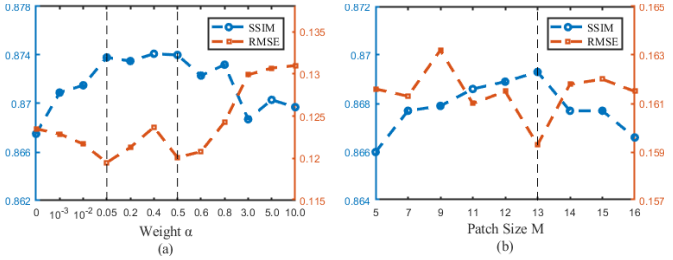


Fig. 5. Validation results under different patch sizes and different weight parameters α.

the parameters in the network. We finally randomly selected 5000 images from MNIST as the training data set and another 2500 images as the validation data set considering both the capacity and complexity of the network. In order to enhance the generality capability of the model, we randomly rotated the digit and also add a random circle to the DOI D to represent possible multiple scatters as referred to [21]. In detail, a random circle with radius from 0.1m to 0.5m is incorporated into each digit to improve the model generalization capability. Meanwhile, each digit is randomly rotated with an angle between $-170°$ and $170°$ to account for the spatial diversity of the scatterers. All scatterers are assumed to be lossless dielectrics with relative permittivity ε_r randomly distributed between 1.5 and 2.5. In Fig.4, we demonstrate 16 randomly selected training samples, where the ground truth images are presented in Fig.4 (a) and the related rough BP input images are illustrated in Fig.4(b).

The DOI D is chosen as a square domain with size of 2.0 m×2.0 m in a free space background. Moreover, 16 linearly polarized transmitters, which are located uniformly over a circle S with radius $R = 3.0$ m, successively illuminate the investigation domain. Meanwhile, 32 co-polarized receivers are used to collect the scattered field. The operating frequency is set as 400 MHz. For each incidence, we calculate the scattered field using the MoM, where D is discretized into 100×100 grids. To avoid the inverse crime, the grids used for reconstruction are changed to 64×64. The scattered field used in the training stage is always noiseless for all examples.

The Adam optimization method is taken to optimize the U-net with the exponential decay rates set as $\beta_1 = 0.9$ and $\beta_2 = 0.999$, respectively. Some other hyperparameters used for training are listed as follows. The batch size is taken as 1 and totally 20 epochs are used to train the model. The learning rate is initially set to 0.0002 during the

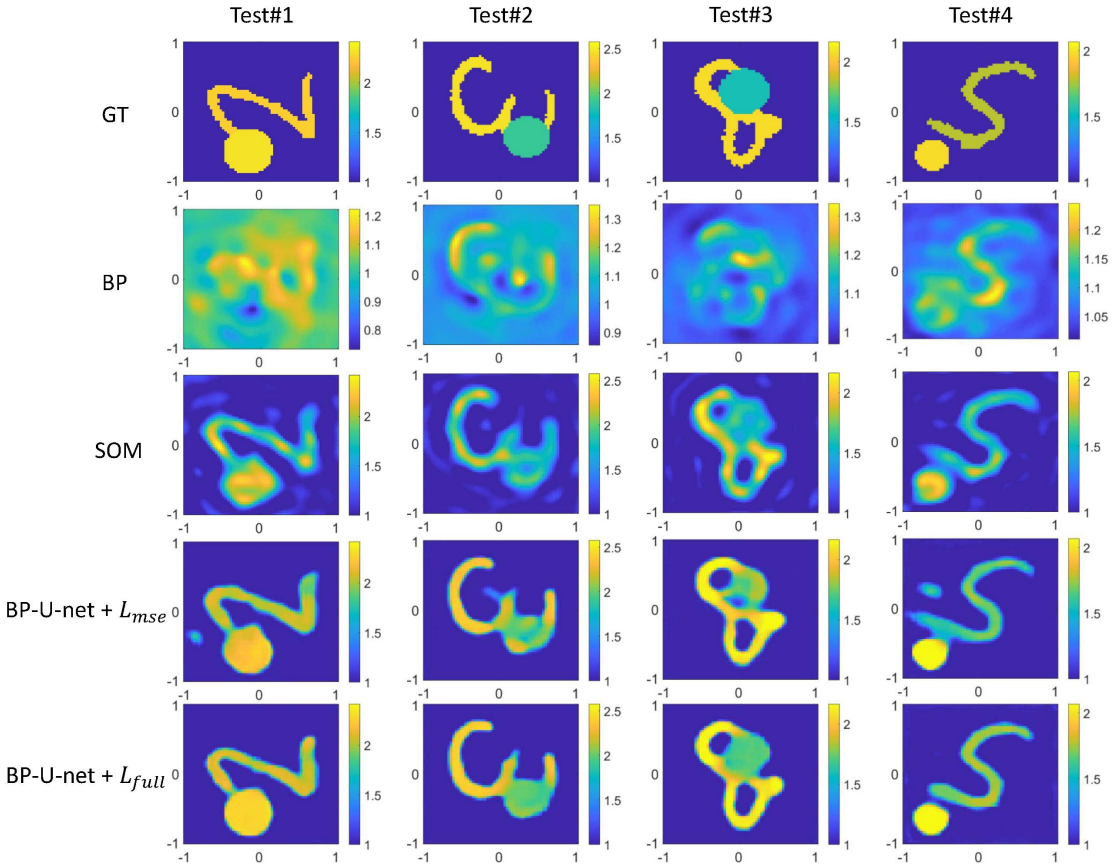


Fig. 6. Reconstruction results of Test#1 to Test#4 from the MNIST data set with 10% white Gaussian noise.

first 10 epochs and it sequentially decreases to zero from the 11th to the last epoch. All network training and testing are done on a workstation (Intel Core i7-6800K and 64 GB RAM) with a single Nvidia GeForce GTX 1080 Ti using PyTorch framework, and the related SSIM loss is implemented with an open source code 'pytorch-msssim'.[1]

We set the dynamic range L in SSIM loss (equation (8)) as 2.5, since the relative permittivity ε_r is randomly distributed between 1.5 and 2.5 in the training set. Another parameter that affects the performance of SSIM loss is the patch size M in (9). During the training process, we determine an optimal M with the validation set using a parameter sweeping method. Suppose the weight parameter α of the hybrid loss in (9) is fixed as 1.0, which is relatively large so that it can highlight the role of SSIM loss. We train the U-net model with the patch size M gradually increased from 5 to 13 and evaluate the performance of models on the validation set. The corresponding results are presented in Fig. 5 (b). The Root-Mean-Square Error (RMSE) and SSIM are two metrics used to compare the reconstruction results. We observe that the model achieves overall the best performance on the validation set when $M = 13$. Hence, we set the patch size as 13×13 for all examples.

The weight parameters α balances the MSE and the SSIM losses. We also explore the impact of α using the validation set with α gradually increasing from 0.005 to 5.0. The results from validation set are shown in Fig. 5 (a). We observe that the metrics of SSIM and RMSE have both been clearly improved when α is within the range of $[0.05, 0.5]$. Particularly, the best results are achieved at the two ends when α is chosen as 0.05 or 0.5. It indicates that there is little impact of SSIM loss on the reconstruction results if α is smaller than 0.05. Otherwise, it may overweight the SSIM term and lead to a worse result if α exceeds 0.5. The weight α can be selected from the interval of $[0.05, 0.5]$ according to the specific problem.

[1] https://github.com/VainF/pytorch-msssim

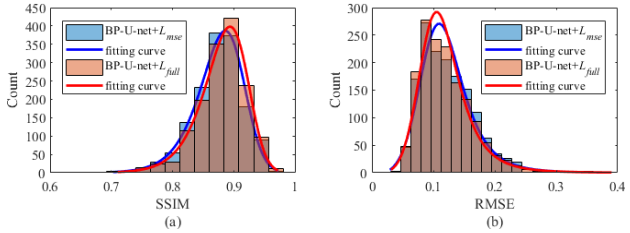


Fig. 7. Statistical histograms of the reconstructed results. (a) SSIM (b) RMSE.

B. Test With MNIST Data

In the first example, we test all the candidate methods with another 1500 images randomly selected from MNIST data set. The scattered fields in all testing samples are contaminated with 10% white Gaussian noise.

TABLE I
COMPARISON OF RECONSTRUCTION RESULTS FOR SOM, BP-U-NET MODEL WITH A L_{mse} LOSS AND BP-U-NET MODEL WITH A L_{full} LOSS IN THE MNIST TESTING DATA SET

Test	SSIM			RMSE		
	SOM	BP-U-net+L_{mse}	BP-U-net+L_{full}	SOM	BP-U-net+L_{mse}	BP-U-net+L_{full}
#1	0.694	0.792	0.846	0.192	0.209	0.190
#2	0.720	0.831	0.865	0.181	0.187	0.177
#3	0.783	0.876	0.896	0.132	0.118	0.109
#4	0.811	0.818	0.873	0.116	0.132	0.108
MNIST (1500)	0.778	0.873	0.880	0.142	0.123	0.119

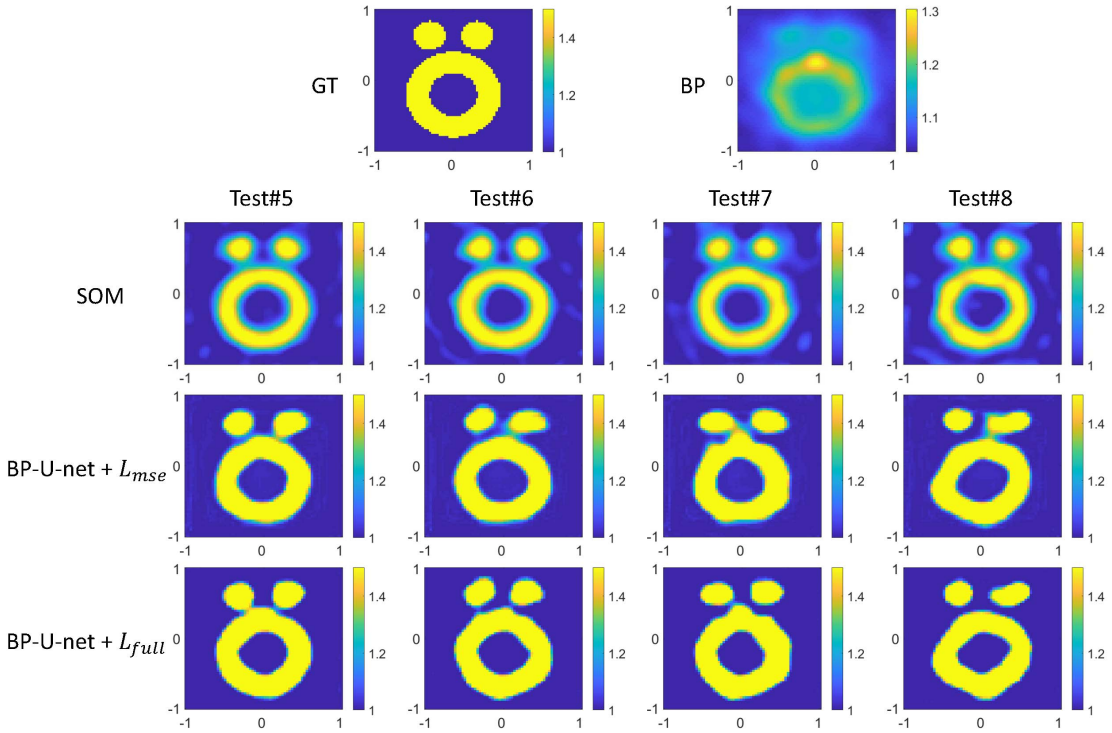


Fig. 8. Reconstruction results of "*Austria*" profiles under different noise levels. Test#5 to Test#8 are obtained with additional 10%, 15%, 20% and 25% white Gaussian noises, respectively. The first row shows the ground truth image (left) and the reconstructed result by BP with 10% noise (right).

The reconstruction results of Test#1 to Test#4 are shown in Fig. 6. We can observe that the BP-U-net model with the hybrid L_{full} loss has achieved the best reconstruction results compared to that with only the MSE loss and the SOM. The quantitative evaluation results for all samples are summarized in Table I, and we can clearly see that the proposed method outperforms the other two methods on both the average SSIM and RMSE metrics. In order to compare the performance of BP-U-net models obtained with the single L_{mse} and the hybrid L_{full} losses more intuitively, the metric distributions of all testing samples are illustrated in Fig. 7. We can see that the BP-U-net method with the SSIM loss has achieved better imaging quality for the MNIST data compared to the BP-U-net results with only the pixel-wise MSE loss.

C. Test With "Austria" Profile Under Different Noise Levels

In the second example, we verify the generalization capability of the model using the well-known "*Austria*" profile [27] under different noise levels. The same BP-U-net model as the first example is taken to reconstruct the scatterer. The "*Austria*" profile has a relative permittivity of ε_r as 1.5. We separately add 10%, 15%, 20% and 25% white Gaussian noises to the scattered field to test the robustness of the model.

In Fig.8, we present the reconstruction results of all the three methods. It can be seen that the BP-U-net with hybrid loss L_{full} has achieved better imaging quality for all the cases under different noise levels. The SSIM and RMSE metrics for Test#5 to Test#8 are also listed in Table II, where the results of BP-U-net model with the hybrid loss L_{full} consistently outperform those ones with only the MSE loss. It is indicated that the perceptual SSIM loss effectively improves the reconstruction quality for the challenging "*Austria*" profile. This example also proves that the SSIM loss maintains a good generalization capability of the trained model under different noise levels.

D. Test With Experimental Data

Finally, we also validate the effectiveness of perceptual SSIM loss using experimental data measured by Institue Fresnel [28]. A "*FoamDielExt*" profile under a TM case is

TABLE II
COMPARISON OF RECONSTRUCTION RESULTS FOR "*Austria*" PROFILE IN FIG.8

Test	Noise Level	SSIM			RMSE		
		SOM	BP-U-net+L_{mse}	BP-U-net+L_{full}	SOM	BP-U-net+L_{mse}	BP-U-net+L_{full}
#5	10%	0.763	0.864	0.900	0.073	0.070	0.066
#6	15%	0.756	0.858	0.877	0.075	0.071	0.069
#7	20%	0.727	0.823	0.854	0.081	0.085	0.080
#8	25%	0.715	0.799	0.810	0.083	0.090	0.087

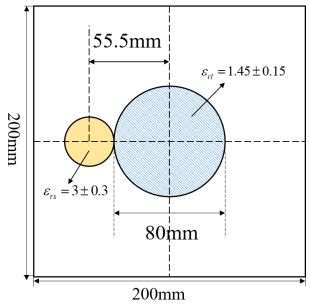


Fig. 9. The "*FoamDielExt*" profile from Fresnel experimental data.

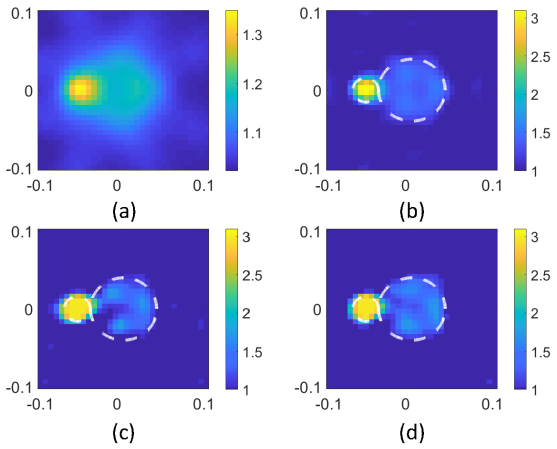


Fig. 10. Reconstruction results of "*FoamDielExt*" profile at 3 GHz with (a) BP (b) SOM (c) BP-U-net+L_{mse} and (d) BP-U-net+L_{full}.

taken to test as shown in Fig. 9. The "*FoamDielExt*" profile consists of two cylinders, where the blue foam cylinder has a diameter of 80.0mm with ε_r as 1.45 ± 0.15, and the yellow copper tube has a diameter of 28.5mm with ε_r as 3 ± 0.3. There are 8 linearly polarized transmitters and 241 co-polarized receivers located at a circle of radius as $d_e = 1.67$ m.

In this example, we use the same MNIST training profiles as the above two examples. But we need to change the frequency from 400MHz to 3GHz to be consistent with the experimental data. Accordingly, the size of DOI is also changed from 2.0 m $\times$ 2.0 m to 0.2 m $\times$ 0.2 m. In this example, we set the weight parameters α in the hybrid loss function as 0.05.

The reconstructed results of "*FoamDielExt*" profile are shown in Fig. 10. It can be seen that the BP-U-net method with hybrid losses achieves much better image quality compared to the one with only a pixel-wise MSE loss in the experimental data. The corresponding SSIM index is increased from 0.855 to 0.886, while the RMSE is reduced from 0.176 to 0.158. It indicates that the perceptual SSIM loss is also effective when dealing with real ISP data. But we also observe that the SOM method gets a litter better results compared to the other two BP-U-net methods, where the SSIM is 0.914 and the RMSE is 0.100. It indicates that the SOM method has good generalizations to different profiles since it is a model-based inversion method. Whereas, considering the training set used by BP-U-net is totally different from the testing profile herein, the reconstruction performance of BP-U-net with hybrid losses is still satisfied. On the other hand, the SOM runs for 50 iterations and it takes 39.370 seconds to get the results in Fig. 10. In contrast, the BP-U-net method with hybrid losses only takes 0.716 seconds to achieve similar reconstructions. This example further indicates the pros and cons of BP-U-net methods under the DL-IS framework.

In summary, the proposed method is tested on three examples including the MNIST digits, the "*Austria*" profile and the Fresnel experimental data. All the results verify the effectiveness of the perceptual structural similarity loss to enhance the quality of reconstructions. It should be noted that we have also got similar conclusions for the SSIM loss to combine with the MAE loss. To save space, we omitted related results here.

IV. CONCLUSION

In this article, we have systematically evaluated the effectiveness of using the SSIM loss to improve the reconstruction quality of ISP under the common DL-IS framework. The hybrid loss with a perceptual SSIM term has been proven to consistently outperform the original MSE one. The results for both synthetic and experimental data validate that the use of SSIM loss can effectively reduce artifacts of the reconstructions and thereby enhance the imaging quality. Besides, a good generalization capability has also been validated even for the challenging ISP cases.

Although this article uses the SSIM loss as an example, other differentiable perceptual metrics can also be considered to build the loss functions. Overall, the current study provides a new idea to improve the reconstruction quality of DL-IS methods. The use of perceptual loss is expected to play an important role in real ISP applications, especially for imaging targets with complex texture information.

REFERENCES

[1] T. Deisboeck and J. Y. Kresh, *Complex Systems Science in Biomedicine*. New York, NY, USA: Springer, 2007.
[2] R. Persico, *Introduction to Ground Penetrating Radar: Inverse Scatterin and Data Processing*. Hoboken, NJ, USA: Wiley, 2014.

[3] L.-P. Song, C. Yu, and Q. Huo Liu, "Through-wall imaging (TWI) by radar: 2-D tomographic results and analyses," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2793–2798, Dec. 2005.
[4] E. Anterrieu, S. Gratton, and B. Picard, "Inverse problem in remote sensing by aperture synthesis imaging," *Trait. Signal*, vol. 21, no. 1, pp. 1–16, 2004.
[5] R. Zoughi, *Microwave Non-Destructive Testing and Evaluation Principles*, vol. 4. Dordrecht, The Netherlands: Springer, 2000.
[6] X. Chen, *Computational Methods for Electromagnetic Inverse Scattering*. Hoboken, NJ, USA: Wiley, 2018.
[7] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted born iterative method," *IEEE Trans. Med. Imag.*, vol. 9, no. 2, pp. 218–225, Jun. 1990.
[8] P. M. V. D. Berg and R. E. Kleinman, "A contrast source inversion method," *Inverse Problems*, vol. 13, no. 6, pp. 1607–1620, Dec. 1997.
[9] Y. Zhong and X. Chen, "Twofold subspace-based optimization method for solving inverse scattering problems," *Inverse Problems*, vol. 25, no. 8, pp. 85003–85013, 2009.
[10] A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "DNNs as applied to electromagnetics, antennas, and propagation—A review," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2225–2229, Nov. 2019.
[11] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Prog. Electromagn. Res.*, vol. 167, pp. 67–81, 2020.
[12] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1819–1825, Mar. 2019.
[13] K. Belkebir, P. C. Chaumet, and A. Sentenac, "Superresolution in total internal reflection tomography," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 22, no. 9, pp. 1889–1897, 2005.
[14] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, Apr. 2019.
[15] H. M. Yao, W. E. I. Sha, and L. Jiang, "Two-step enhanced deep learning approach for electromagnetic inverse scattering problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2254–2258, Nov. 2019.
[16] Y. Sanghvi, Y. Kalepu, and U. K. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 46–56, 2020.
[17] L. Zhang, K. Xu, R. Song, X. Z. Ye, G. Wang, and X. Chen, "Learning-based quantitative microwave imaging with a hybrid input scheme," *IEEE Sensors J.*, early access, Jul. 27, 2020, doi: 10.1109/JSEN.2020.3012177.
[18] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.
[19] J. Snell, K. Ridgeway, R. Liao, B. D. Roads, M. C. Mozer, and R. S. Zemel, "Learning to generate images with perceptual similarity metrics," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 4277–4281.
[20] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss functions for image restoration with neural networks," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 47–57, Mar. 2017.
[21] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, Sep. 2019.
[22] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
[23] L. Li, L. G. Wang, J. Ding, P. K. Liu, M. Y. Xia, and T. J. Cui, "A probabilistic model for the nonlinear electromagnetic inverse scattering: TM case," *IEEE Trans. Antennas Propag.*, vol. 65, no. 11, pp. 5984–5991, Nov. 2017.
[24] A. Lakhtakia, "Strong and weak forms of the method of moments and the coupled dipole method for scattering of time-harmonic electromagnetic fields," *Int. J. Modern Phys. C*, vol. 3, no. 3, pp. 583–603, Jun. 1992.
[25] T. Yin, Z. Wei, and X. Chen, "Non-iterative methods based on singular value decomposition for inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4764–4773, Jun. 2020.
[26] X. Chen, "Subspace-based optimization method for solving inverse-scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 42–49, Jan. 2010.
[27] K. Belkebir and A. Tijhuis, "Using multiple frequency information in the iterative solution of a two-dimensional nonlinear inverse problem," in *Proc. Prog. Electromagn. Res. Symp. (PIERS)*, Innsbruck, Germany, 1996, p. 353.
[28] J.-M. Geffrin, P. Sabouroux, and C. Eyraud, "Free space experimental scattering database continuation: Experimental set-up and measurement precision," *Inverse Problems*, vol. 21, no. 6, pp. S117–S130, Dec. 2005.