Performance Analysis and Dynamic Evolution of Deep Convolutional Neural Network for Electromagnetic Inverse Scattering

Lianlin Li, Senior Member, IEEE, Long Gang Wang, and Fernando L. Teixeira, Fellow, IEEE

Abstract—The solution of electromagnetic (EM) inverse scattering problems is hindered by challenges, such as ill-posedness, nonlinearity, and high computational costs. Recently, deep learning was shown to be a promising tool in addressing these challenges. In particular, it is possible to establish a connection between a deep convolutional neural network (CNN) and iterative solution methods of EM inverse scattering. This led to the development of an efficient CNN-based solution to EM inverse problems, termed DeepNIS. It has been shown that DeepNIS can outperform conventional inverse scattering solution methods in terms of both the image quality and computational time. In addition, we probe the dynamic evolution behavior of DeepNIS by examining its near-isometry property and show that, after a proper training stage, the proposed CNN has near-optimal stability properties.

Index Terms—Convolutional neural network (CNN), electromagnetic (EM) inverse scattering, machine learning.

I. INTRODUCTION

The solution of electromagnetic (EM) inverse scattering problems is of interest in a number of applications [1]–[11]. These solutions are able to take into account multiple scattering effects inside the probed scene [1], [3]. Many inverse scattering algorithms have been developed over the years [4], [5]; however, their application to large, realistic scenes is hampered by high computational costs. Recently, with the emergence of compressive sensing (CS) theory, some sparseness-aware inverse scattering algorithms were proposed to mitigate the ill-posedness of underlying inverse problem [9], [12]–[14]. The CS-based approach relies on the so-called sparse representation in some known specified or learned transform domain. Although these methods can produce acceptable results for scenes with moderate size and contrast, it remains an outstanding challenge to deploy them in large and realistic scenes due to high computational costs. To address this difficulty, several learning-based approaches [15], [16] have been investigated such as artificial neural network [16]. In recent years, deep convolutional neural networks (CNN) have shown to be a promising tool for solving inverse problems due to the increasing availability of very large datasets and a concomitant increase in the available computational power [17], [18]. For example, CNN-based strategies have been successfully applied to magnetic resonance imaging, X-ray computed tomography [19], and computational optical imaging [20]–[22]. It has been found that they can typically outperform conventional image reconstruction techniques in terms of improved image quality and computational speed [19]–[22]. More recently, Li et al. [23] investigated the connection between deep CNNs and iterative methods for EM inverse scattering. Based on this connection, they proposed a complex-valued CNN architecture for tackling EM inverse scattering, termed “DeepNIS.” A complex-valued CNN is a straightforward extension of conventional real-valued CNN. DeepNIS is a noniterative solver, which greatly reduces the computational costs compared with iterative techniques. In parallel to this, Wei and Cheng [24] have recently applied a U-net-based deep neural network to EM inverse scattering problems, where three different input methods have been comprehensively studied. In contrast to CS-based approaches, the supervised CNN-based approach relies on the availability of a large amount of labeled training samples.

In this letter, we quantitatively evaluate the performance of DeepNIS as a function of the number of layers using the structure similarity index measure (SSIM) and mean square error (MSE) metrics. The performance is evaluated using the MNIST dataset [25]. In addition, we probe the dynamic evolution behavior of DeepNIS by examining its near-isometry property. After a proper training stage, this probing study shows that the proposed CNN is close to optimal in terms of the stability and generalization ability. The results show the potential of DeepNIS in tackling EM inverse scattering problems.

II. THEORETICAL FOUNDATIONS

This section summarizes the connection between the CNN architecture and Born iterative method for solving EM inverse scattering problems. Throughout this letter, the time dependence factor \( \exp(-i\omega t) \) is implied and omitted.
We illustrate our strategy in a two-dimensional multiple-input–multiple-output data acquisition setup in Fig. 1. The investigation domain $\Omega$ is successively illuminated by TM-polarized incident waves $E_{\text{inc},\theta}(r;\omega)$ (with $\theta$ being the index of the $\theta$th illumination). The transmitters and receivers are both located in the observation domain $\Gamma$. For the $n$th illumination ($n=1,2,\ldots,N$) and the $m$th ($m=1,2,\ldots,M$) receiver, the scattered electrical field $E_{s,\theta}(r;\omega)$ at $r$ is obtained by the coupled equations [5]–[8]

$$E_{s,\theta}(r;\omega) = \int_{D_{\text{inv}}} G(r,r';\omega) E_{\theta}(r';\omega) \chi(r') \, dr', \quad r \in \Gamma$$

and

$$E_{\theta}(r;\omega) = E_{\text{inc},\theta}(r;\omega) + \int_{D_{\text{inv}}} G(r,r';\omega) E_{\theta}(r';\omega) \chi(r') \, dr', \quad r \in D_{\text{inv}}$$

where $r'=(x',y')$ and $r=(x,y)$ denote source and observation points, respectively, and $G(r,r')$ is the free-space Green’s function. Additionally, $\chi = k^2 / \omega^2 - 1$ denotes the contrast, and $k$ and $k_0$ are the wavenumbers of the probed object and background medium, respectively.

For numerical implementation, the investigation domain $\Omega$ is discretized into a series of pixels and the discrete field values represented by column vectors. After that, (1) and (2) can be rewritten in compact form as follows:

$$E_{\text{scn}}^{(n)} = G_d E^{(n)} \chi$$

and

$$E^{(n)} - E_{\text{inc}}^{(n)} = G_s E^{(n)} \chi.$$

The l1-norm-based sparse Born iterative method, sBIM for short, can be applied to solve the EM inverse problem described by (3) and (4) as summarized in Algorithm 1, where $\rho$ is an empirical balancing factor, and (5) is solved by employing the so-called iterative soft threshold algorithm [3]. The critical stage is to solve the time-consuming and ill-posed optimization problem represented by (5). We attempt to address this difficulty by exploring the CNN strategy. To that end, we modify (5) as follows:

$$n = 0; E_{\theta}^{(n)} = E_{\theta}^{(\text{inc})}$$

WHILE NOT convergence

- $\chi^{(n+1)} = \arg\min_{\chi} \left\{ \sum_{\theta} \| E_{s,\theta} - G_d \text{diag} \left( E_{\theta}^{(n)} \right) \chi \|_2^2 + \rho \| \chi \|_1 \right\}$

- Update $E_{\theta}^{(n+1)}$ from the state equation (4).

$\n \leftarrow n + 1$

END WHILE

Algorithm 1: 11-Norm-Based Sparse Born Iterative Solution.

Algorithm 2: Modified Born Iterative Solution.

For $n = 1,2,\ldots,K$

- $\bar{p}^{(n)}(\theta) = \mathcal{S}\left\{ (I + \eta_\theta (A_{\theta}^{(n)}))^{*} A_{\theta}^{(n)} \chi^{(n)} - \eta_\theta (A_{\theta}^{(n)})^{*} E_{s,\theta} \right\}$

- $\chi^{(n)} = \sum_{\theta} \xi_{\theta}^{(n)}(\theta)$

- Update $E_{\theta}^{(n+1)}$

END FOR

$\chi^{(n+1)} = \arg\min_{\chi} \left\{ \sum_{\theta} \| E_{s,\theta} - G_d \text{diag} \left( E_{\theta}^{(n)} \right) \chi \|_2^2 + R(\chi) \right\}$

where $\{\xi_{\theta}^{(n)}\}$ denote illumination-dependent weighting factors used to adjust the contributions from different measurements. The regularization term $R(\chi)$ incorporates any prior on the contrast function $\chi$ and is used to mitigate the ill-posedness of the inverse problem. By applying so-called one-step first-order gradient-based approach [3] to solve (6), we obtain Algorithm 2, where $A_{\theta}^{(n)} = G_d \text{diag} (E_{\theta}^{(n)})$, and $\eta_\theta$ denotes the regularization factor. We assume that $E_{\theta}^{(n)}$ can be statistically approximated by its nearest stationary field $\bar{E}_{\theta}^{(n)}$. This implies that $(\bar{E}_{\theta}^{(n)})^{*} E_{\theta}^{(n)}$ is shift-invariant, and thus $w_{\theta}^{(n)} = I + \eta_\theta (A_{\theta}^{(n)})^{*} A_{\theta}^{(n)}$ behaves like a typical convolutional kernel. The iterative index $n$ can be understood as the layer index of the deep neural networks. $K$ corresponds to the total number of CNN layers, whereas the soft-threshold function $\mathcal{S}\{\cdot\}$ corresponds to the nonlinear activation function in deep learning [23]. In this sense, we establish the connection between the Born-iterative method and deep CNN, where $\bar{p}^{(n)}(\theta)$ extracts the illumination-dependent features. By comparing this strategy with conventional iterative inverse scattering methods, the expectation is that the learning method would be more efficient as it optimizes the weighting matrices and biases, and targets the reconstruction error with respect to the ground-truth images. The abovementioned observations suggest that deep CNN networks are naturally well suited for nonlinear EM inverse scattering problems.
III. EXPERIMENTAL RESULTS

In the following, we evaluate DeepNIS performance for solving EM inverse scattering problems. To evaluate the reconstructed image quality, we adopt the SSIM and MSE metrics. In addition, we examine its dynamic evolution.

A. Simulation Setup

We train and test the DeepNIS using the MNIST dataset [25]. With reference to Fig. 1, the region of interest $D_{\text{inv}}$ is a square domain of size $5.6 \lambda_0 \times 5.6 \lambda_0$ ($\lambda_0 = 7.5$ cm is the working wavelength in vacuum), which is uniformly divided into $110 \times 110$ pixels for the simulations. The MNIST dataset elements shown in the first column of Fig. 2(a) constitute randomly placed targets within each sample. A total of 36 linearly polarized transmitters uniformly spaced over the circumference $\Gamma$ with radius $R = 10 \lambda_0$ are used to successively illuminate the investigation domain. At the same time, 36 copolarized receivers are used to collect the scattered electric field. The MNIST dataset elements are assumed to be lossless dielectrics with relative permittivity $\varepsilon_r = 3$. In addition, white noise is added to the results of the forward simulations equivalent to a 30 dB signal-to-noise ratio [26]. A total of $10^4$ image pairs constituted by back-propagated images and original (ground truth) images randomly chosen from the MNIST dataset are divided into three sets: 7000 image pairs for training, 1000 for validation, and 2000 for blind testing. The training stage is done using the ADAM optimization method [27], with mini-batches size of 32, and epoch setting as 50. The learning rates are set to $10^{-4}$ and $10^{-5}$ for the first two layers and the last layer in each network, respectively, and halved once the error plateaus. The complex-valued weights and biases are initialized by random weights with zero mean Gaussian distribution and standard deviation of $10^{-3}$. The DeepNIS network training takes about 3.6 h. The computations are performed with AMD Ryzen Threadripper 1950X 16-Core processor, NVIDIA GeForce GTX 1080Ti, and 128 GB access memory. The networks are designed using the Tensor Flow library [28].

B. Numerical Results

Fig. 2(a) shows the images obtained by using the proposed CNN with the numbers of layers varying from one to nine. For reference, the corresponding ground truth data and the sBIM results are provided in the first column and second column, respectively. Throughout this letter, the total sBIM iterative number is set to be 25 because after that no further visible improvement on the reconstruction quality can be observed. In order to show the stability of the trained CNN, we also considered the rotated digitlike objects of MNIST dataset with 45° and −45°, as shown in the eighth and ninth rows of Fig. 2(a), respectively. Moreover, we also consider the 90°-rotated digitlike object after being mirrored, as shown in the bottom three rows of Fig. 2(a). These results clearly illustrate that well-trained CNNs with four layers or more can produce excellent reconstructions. In contrast, the sBIM provides relatively poorer reconstructions in this case. Fig. 2(b) and (d) examines quantitatively the quality of the
images by using the proposed CNNs with different number layers in terms of the SSIM and MSE metrics. Fig. 2(b) plots the dependence of the averaged SSIMs and MSEs as a function of the number of CNN layers. The averaged SSIMs and MSEs are computed over 2000 training samples and 2000 test samples. In Fig. 2(b), a sample of image reconstructions of the digit “5” is provided for reference. Fig. 2(d) reports the statistical histograms of the image quality in terms of SSIM corresponding to CNN with one, three, five, seven, and nine layers, respectively, over 2000 test images. The y-axis is normalized with respect to the total 2000 test images. Furthermore, in order to examine the performance of the proposed CNN network against the high-contrast objects, we have conducted another set of numerical experiments. Fig. 2(c) plots the dependence of the averaged SSIMs as a function of the object contrast. Some illustrated reconstructed results are inserted in Fig. 2(c) as well. Based on these results, several conclusions can be made. First, when the CNN has more than five layers, both MSEs and SSIMs converge to a stable level. It can be clearly seen that the DeepNIS with more than five layers can match the ground truth results very well. Second, if the CNN has more layers, the resultant discrepancy between training and testing performance increases. This is likely a consequence of the fact that deeper CNNs require more network parameters to be optimized over a given training sample and, hence, are more prone to overfitting.

Finally, it is worth mentioning that it only takes a well-trained DeepNIS less than 1 s to construct an image in this case, whereas it takes Born iterative algorithm about nearly 1 h. Based on the above-mentioned results, it can be concluded that a properly trained DeepNIS solution clearly outperforms sBIM in terms of both image quality and computation time.

Fig. 3 provides further insights into the CNN-based EM inverse scattering solution. In this case, the number of CNN layers is fixed at five, and the ground truth corresponds to a digit “2”-like object. In this set of figures, nine randomly selected features extracted at different CNN depths are illustrated. It can be seen that with an increase in CNN depth, the extracted features approach gradually the ground truth. It can be conjectured that the extracted features more or less reflect the contrast function \( \hat{p}^{(n)}(\theta) \) under the different illumination conditions. In other words, these results suggest that CNN-based approach is not merely matching patterns, but actually has a learning capability to represent the underlying EM inverse scattering problem.

C. Dynamic Evolution Behavior of DeepNIS

Stability is a crucial issue for deep neural networks. Here, we examine the evolution behavior of feature propagation through the proposed CNN. More formally, the feature at the output layer of the CNN can be described as \( \chi^{(\text{out})} = F(\chi^{(\text{in})}, \Theta) \), where \( \Theta \) encapsulates the parameters of the network block. Inspired by the restricted isometry property widely examined in CS [3], we focus on the parameter \( \eta = \| \delta \chi^{(\text{out})} \|^2_2 / \| \delta \chi^{(\text{in})} \|^2_2 \), where \( \delta \chi^{(\text{in})} \) and \( \delta \chi^{(\text{out})} \) denote the input perturbation to the CNN network and the associated output perturbation, respectively. Evidently, the network with \( \eta \gg 1 \) is unstable since the output feature would be highly sensitive to data perturbations. Conversely, if \( \eta \ll 1 \), the corresponding network block lacks generalization ability since input features would be exponentially suppressed and not properly identified at the output. Consequently, an “optimal” neural network should have \( \eta \approx 1 \). In other words, a deep neural network with \( \eta \approx 1 \) is near optimal in terms of stability. Fig. 4 plots the dependence of averaged \( \eta \) as the function of iteration number for the CNN with different number layers (and with the other parameters set as before). It can be seen that the proposed CNNs with more than three layers, which are trained after ten iterations, indeed have \( \eta \approx 1 \).

IV. Conclusion

We have evaluated the performance of DeepNIS as a function of the number of layers based on different quantitative metrics quantitatively. We have shown that DeepNIS shows advantages over conventional inverse scattering methods in terms of image quality and computational time. We have also investigated the dynamic evolution of DeepNIS. The analysis shows that, following a proper training stage, the proposed CNN is near optimal in terms of the stability. Together, these results indicate the clear potential of DeepNIS in tackling EM inverse scattering problems. It is plausible that more advanced or tailored CNN architectures could yield better results, which will be explored in a future study.
REFERENCES

[1] G. Maire et al., “Experimental demonstration of quantitative imaging beyond Abbe’s limit with optical diffraction tomography,” Phys. Rev. Lett., vol. 102, 2009, Art. no. 213905.

[2] Q. Marashdeh, W. Warsito, L.-S. Fan, and F. L. Teixeira, “A nonlinear image reconstruction technique for ECT using a combined neural network approach,” Meas. Sci. Technol., vol. 17, pp. 2097–2103, 2006.

[3] L. Li et al., “A survey on the low-dimensional-model-based electromagnetic imaging,” Found. Trends Signal Process., vol. 12, pp. 107–199, 2018.

[4] M. Pastorino, Microwave Imaging. New York, NY, USA: Wiley, 2010.

[5] L. Li et al., “Machine-learning reprogrammable metasurface imager,” Nature Commun., vol. 10, 2018, Art. no. 1082.

[6] D. Colton and R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory. Berlin, Germany: Springer, 2012.

[7] L. Di Donato, M. T. Bevacqua, L. Crocco, and T. Isernia, “Inverse scattering via virtual experiments and contrast source regularization,” IEEE Trans. Antennas Propag., vol. 63, no. 4, pp. 1669–1677, Apr. 2015.

[8] 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.

[9] A. E. Fouda and F. L. Teixeira, “Bayesian compressive sensing for ultra-wideband inverse scattering in random media,” Inverse Problems, vol. 30, no. 11, 2014, Art. no. 114017.

[10] P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa, “Evolutionary optimization as applied to inverse scattering problems,” Inverse Problems, vol. 25, 2009, Art. no. 125003.

[11] M. Pastorino, “Stochastic optimization methods applied to microwave imaging: A review,” IEEE Trans. Antennas Propag., vol. 55, no. 3, pp. 538–548, Mar. 2007.

[12] G. Oliveri, M. Salucci, N. Anselmi, and A. Massa, “Compressive sensing as applied to inverse problems for imaging: Theory, applications, current trends, and open challenges,” IEEE Antennas Propag. Mag., vol. 59, no. 5, pp. 34–46, Oct. 2017.

[13] N. Anselmi, M. Salucci, G. Oliveri, and A. Massa, “Wavelet-based compressive imaging of sparse targets,” IEEE Trans. Antennas Propag., vol. 63, no. 11, pp. 4889–4899, Nov. 2015.

[14] G. Oliveri, L. Poli, N. Anselmi, M. Salucci, and A. Massa, “Compressive sensing-based Born iterative method for tomographic imaging,” IEEE Trans. Microw. Theory Techn., vol. 67, no. 5, pp. 1753–1765, May 2019.

[15] A. Massa, G. Oliveri, M. Salucci, N. Anselmi, and P. Rocca, “Learning by-examples techniques as applied to electromagnetics,” J. Electromagn. Waves Appl., vol. 32, no. 4, pp. 516–541, 2018.

[16] S. Caorsi and P. Gamba, “Electromagnetic detection of dielectric cylinders by a neural network approach,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 2, pp. 820–827, Mar. 1999.

[17] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 28, pp. 436–444, 2015.

[18] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[19] Y. Han, J. Yoo, and J. C. Ye, “Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis,” 2016, arXiv:1611.06391.

[20] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” IEEE Trans. Image Process., vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[21] A. Sinha, J. Lee, S. Li, and G. Barbastathis, “Lensless computational imaging through deep learning,” Optica, vol. 4, no. 9, pp. 1117–1125, 2017.

[22] U. S. Kamilov et al., “Learning approach to optical tomography,” Optica, vol. 2, no. 6, pp. 517–522, 2015.

[23] 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.

[24] Z. Wei and X. Chen, “Deep-learning schemes for full-wave nonlinear inverse scattering,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 4, pp. 1849–1860, Apr. 2019.

[25] Y. LeCun, C. Cortes, and C. J. C. Burges, “MNIST database of handwritten digits,” 2013. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[26] M. F. Catedra, R. P. Torres, J. Basterrechea, and E. Gago, The CG-FFT Method: Application of Signal Processing Techniques to Electromagnetics. Boston, MA, USA: Artech House, 1995.

[27] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017, arXiv:1412.6980.

[28] M. Abadi et al., TensorFlow: Large-scale machine learning on heterogeneous systems white paper,” White Paper, 2015. [Online]. Available: https://www.tensorflow.org