A Spatial SEM-Based Shallow Neural Network for Electromagnetic Inverse Source Modeling

Abdelelah Alzahed¹, Said Mikki², *, and Yahia Antar¹

Abstract—We derive and verify a new type of low-complexity neural networks using the recently introduced spatial singularity expansion method (S-SEM). The neural network consists of a single layer (shallow learning approach to machine learning) but with its activation function replaced by specialized S-SEM radiation mode functions derived by electromagnetic theory. The proposed neural network can be trained by measured near- or far-field data, e.g., RCS, probe-measured fields, array manifold samples, in order to reproduce the unknown source current on the radiating structure. We apply the method to wire structures and show that the various spatial resonances of the radiating current can be very efficiently predicted by the S-SEM-based neural network. Convergence results are compared with Genetic Algorithms and are found to be considerably superior in speed and accuracy.

1. INTRODUCTION

Inverse scattering and source modeling problems, such as radar applications, remote sensing, earth modeling, electromagnetic compatibility (EMC), etc., commonly involve the estimation of the material and/or geometrical constitution of some unknown objects, e.g., dielectric profile, positions, and orientation, by using radio transceivers to collect measured returned far- or near-field data [1]. Mainstream approaches as in [2, 3] depend crucially on the backscattered fields of these objects such as their radar cross sectional area (RCS) when illuminated by a field source. However, other approaches proposed target detection mechanisms based on imaging techniques where the whole process depends solely on data extracted from images, e.g., field pattern, rather than working directly with the much more complex 3-dimensional field-theoretic electromagnetic problem [4–11].

In regard to the most recent work conducted in this area of research, we propose an alternative physics-based estimation of targets’ properties via the spatial electromagnetic data of the unknown object under consideration. This we propose to attain via the recently-introduced spatial-singularity expansion method (S-SEM) information [12] combined with an intelligent machine learning agent [13]. More specifically, we outline a new technique based on retrieving the S-SEM data information of an unknown source using a new type of feedforward backpropagation artificial neural network (FFBP-NN) [14]. The concept will be discussed by demonstrating within a machine learning framework for inverse scattering [15, 16] an alternative methodology for estimating the S-SEM data using specially designed neural network (NN), see Fig. 1. The feasibility of the proposed S-SEM approach to inverse source modeling problems was investigated earlier using a genetic algorithm (GA) search processor to reconstruct the S-SEM data of some wire systems [15]. However, for real-time scenarios, the GA is very slow and the previous method may experience difficulties in modeling objects consisting of large number of wire segments (e.g., 3D wire grid).
To mitigate these difficulties, we focus in this work on solving the inverse source problem (ISP) using neural networks, which are efficiently trained by local gradient methods instead of the global optimization GA-based search processor. The key idea here is importing into the NN a built-in physics-based prediction mechanism based on the ability of the S-SEM to provide new data interpolating/extrapolating capabilities. This we will attain through building the various necessary learning, training and testing/simulating phases of the proposed machine learning approach to ISPs as follows.

(i) In contrast to the now popular deep learning approaches, our new S-SEM-based NN is composed of only a single layer, and hence part of the often neglected though powerful strategy in machine learning called \textit{Shallow Learning}. The small size of the neural network in shallow learning leads to considerable reduction in training cost, and consequently the overall design cost, while the NN computational implementation cost is in general low for feedforward operational modes like in testing.

(ii) In contrast to most existing NNs, the activation function (the “brain” of each neuron), is fully electromagnetic. In other words, we do not use standard activation function popular in statistical learning theory like rectifier or sigmoid functions. Instead, our NN is designed based on exact theoretical derivation coming from the rigorous full-wave EM physics as seen through the eyes of the spatial SEM [12].

(iii) We use the parallel backpropagation gradient descent algorithm to train the NN. The derived S-SEM NN is applied to analysis of 1-dimensional (wire-like) targets and is shown to be very robust and efficient numerically in predicting the spatial form of the source current. Comparison with the GA-previous non-NN method shows that the proposed S-SEM-NN is faster by several orders of magnitude.

In the remaining parts of this paper, the above methodology is first proved by deriving the new NN in Section 2, followed by implementation, testing, and verification in Section 3. The results are briefly discussed together with a report on performance toward the end.

2. DERIVATION OF THE S-SEM-BASED NEURAL NETWORK (S-SEM-NN)

Imagine a real radar detection scenario where a radar station, say an omnidirectional sensor or a mechanically/electronically scanning array system, is scanning for targets in free-space. Once the radar station detects the target, it will capture its far-field pattern. The S-SEM data are conventionally formed by extracting key spectral information from the spatial data of each target [12] and then use them to reconstruct the entire spatial form current distribution [15]. As a result, they are as complex data, namely the S-SEM poles and their strengths. Since S-SEM data are considered good candidates for

\footnote{For simplicity, we assume in this paper that the desired target is placed too far at a far-field point.}
classifying problems, from a data processing perspective the ultimate future objective of the proposed NN can be seen as to classify targets using the S-SEM radiation function presented in [12, 17].

To provide a proof of concept, let us assume that only 1-dimensional (wire-like) sources are considered. Our goal is to reconstruct the entire current on the 1D current segment using only partial far-field data. We start from the accurate electromagnetic S-SEM formula [12]

\[
E_{\text{rad}}(\hat{r}) = \sum_{m=1}^{M} \sum_{n=1}^{N} \alpha_{mn} f_{mn}(\theta, \varphi; L_m, s_{mn}),
\]

(1)

where

\[
f_{mn}(\theta, \varphi; L_m, r_m; s_{mn}) := \hat{L}_m \cdot \left[\hat{I} - \hat{r} \hat{r}\right] e^{ik\hat{r} \cdot r_m} \frac{e^{(ik\hat{L}_m + s_{mn})L_m}/2} {e^{(-ik\hat{L}_m - s_{mn})L_m/2}},
\]

(2)

Here, the vector given by

\[
\hat{r}(\theta, \varphi) := \hat{x} \cos \varphi \sin \theta + \hat{y} \sin \varphi \sin \theta + \hat{z} \cos \theta
\]

(3)

is the radial unit vector \(r/\|r\|\), \(\hat{I}\) is the unit dyad, and \(k = \omega/c\), where \(\omega\) is the angular frequency, and \(c\) is the speed of light. \(M\) is the number of wire segments, each with the parameters \(L_m, \hat{L}_m, r_m, s_{mn}, \alpha_{mn}\), as the \(m\)th wire length, orientation, position, S-SEM poles, and strengths, respectively. The order of all S-SEM models is fixed as \(N\) for simplicity though it may vary from one segment to another. The current distribution on the \(m\)th segment is given by the S-SEM expansion [12]

\[
I_m(r; \omega) = \sum_{n=1}^{N} \alpha_n(\omega) e^{s_n(\omega)l},
\]

(4)

where \(l\) is a local position parameter on the segment under consideration. The expansion in Eq. (4) was developed in [17–20].

In this paper, we fix the geometry for simplicity and attempt to find the source current by training a specialized NN to find \(I(l)\) given in terms of the S-SEM information \(\alpha_{mn}\) and \(s_{mn}\) as per Eq. (4). The proposed NN architecture is described in Fig. 1. It differs from mainstream NNs in the fact that we replace the activation function by the S-SEM radiation modes \(f_{mn}\) defined by Eq. (2). In the estimation process of the proposed S-SEM-NN, \(f_{mn}\) are nested in each neuron within the NN layers to reproduce the proper S-SEM information by replicating the formal structural analogy between Eq. (2) and the NN architecture of Fig. 2. Since we are working with complex data representation, i.e., the S-SEM data \(\alpha_{mn}\) and \(s_{mn}\) in Eq. (2), while traditional NNs do not operate with complex data, the S-SEM-NN inner architecture will be modified to perform a parallel data processing on the real and imaginary data independently.

Therefore, it is now possible to process the S-SEM from each sub-network by combining the real and imaginary outputs outside the network. The proposed NN is configured to supply the user with the S-SEM data during the field reconstruction and classification. In other words, it acts as a reinforcement network that has two major tasks: the first is to estimate/classify targets given their far-field data, while the second is to obtain the S-SEM data (in this paper we focus on the second task). Note that in the proposed S-SEM-NN approach, the major improvement of solving inverse source modeling problems hinges on the strong connection between the far-field and S-SEM data as originally revealed by Eq. (1) originally obtained in [12].

3. IMPLEMENTATION AND VERIFICATION

3.1. General Remarks on the Neural Network Data Processing and Machine Learning Framework

Data processing in typical ANN environment is achieved through a systematic procedure. Here, the authors would like to shed light on some common practices in developing a properly tuned artificial neural network that will be applied into our proposed approach. ANNs consist of three major data
elements: Input, Output and Target data. Besides, it also consists of different processing phases known as learning, training and testing phases to simulate and examine the network performance.

The first phase, which is called the learning phase, dictates the initial formation/guess of the ANN by inserting a known input and target data. The idea is to guide the network to apply the weights and biases through the right direction such that a perfect mapping between input and target data can be achieved. This can also be referred to as the regression curve that shows how perfectly the input data are mapped with the target data. In a typical learning phase, the curves should be on top of each other with a 100% accuracy. In our system, we used two sets of data representing the phasor fields $E(\theta, \phi)$ of two wire systems.

The learning setup is executed by applying the field of each wire system on both ends of the network (input and target). Based on our prior knowledge of the S-SEM representation of each system and by virtue of the nested S-SEM basis function on each neuron, the ANN, besides the data classification, serves as a reinforcement network, where it supplies the values of the S-SEM coefficients (poles and strengths). To completely test the ANN performance, an additional process can be carried out to validate the learning phase with an input corrupted data using an additive white Gaussian noise of a defined variance. In this case, the noise will be added linearly to the phasor fields and the previous process is repeated. The previous step will be further discussed in a future publication. Later, a testing phase will be performed, which will be carried out in a future publication where both reinforcement and classification scenarios will be made.

### 3.2. Implementation and Main Results

For the preparation of the S-SEM-NN system, we first begin by initiating the proper weights (S-SEM data) assigned to the NN as depicted in Fig. 2 per wire segment. The flow of the data starts from the input layer that carries a number of $N$ neurons, where $N$ is also the total number of unknown S-SEM data to be recovered by training the NN (i.e., the order of the S-SEM model is half the order of the S-SEM-NN for each wire.) This assumption will facilitate the process of estimating unknown S-SEM data when classifying targets given their collected far-field data since each neuron is responsible for injecting one S-SEM datum $s_{mn}$ and $\alpha_{mn}$. The dataset size is expected to vary according to the particular system under consideration, but we maintain here a fixed number of neurons such that every

---

† Note that we can also add the geometrical data $r_m$, $L_m$, into the set of NN model parameters. However, for brevity, we focus in this paper on the current distribution reconstruction through the S-SEM poles and their strengths.
wire section carries \( N/2 \) S-SEM poles \( s_{mn} \) and strengths \( \alpha_{mn} \) due to the symmetry of the S-SEM data.\(^5\)

The pre-learning process provides the most important details in forming the NN, where a captured field is injected to the input layer, while at the output layer the expected field is produced (Fig. 2). To perform the training part of the learning process, a suitable minimum squared error (MSE) cost function is introduced with the purpose of choosing the NN internal model parameters (S-SEM data \( s_{mn} \) and \( \alpha_{mn} \) in our case) such that the MSE-measure reaches its minimal threshold. Hence, the purpose of this network is to test the ability of the newly derived NN to reproduce the correct measured field. This MSE cost is defined by

\[
C := \frac{1}{D} \int_{\Omega_{tr}} \| E_{rad}^{S-SEM}(\Omega) - E_{rad}^{tr}(\Omega) \|^2 \, d\Omega, \tag{5}
\]

where \( E_{rad}^{tr} \) is the far-field training dataset (e.g., measured field data collected through radars or sensors). On the other hand, \( E_{rad}^{S-SEM} \) is the S-SEM neural network produced by means of Eq. (2). Here, \( D \) is the size of the training field dataset used in the experiment, while \( \Omega_{tr} \) is the training far-field angles (training dataset features.) Three data sets were used in training the network for the symmetric, L-shape fields, and a random field set. Each data set has a size of 720 points on the elevation scanning direction of each system. In what follows, the classical stochastic back-propagation gradient algorithm [21] will be used to perform the training (optimization) problem.

We start our verification by an example of thin-wire system as in Fig. 3(a). Such wire can serve as a basic building block to create geometric approximation of arbitrary 3D bodies in the form of wire grids [12]. Following [12, 15], the wire dipole system is modeled as two segments of the type treated in Eq. (1) such that each wire has a free end at which theoretically the surface current vanishes.

![Figure 3](image)

**Figure 3.** (a) The symmetric half-wavelength wire system that is used in the learning phase and (b) its far-field polar representation in comparison to a MoM computation [22].

The S-SEM-NN network performance is assessed through two measures: (i) the MSE versus the number of iterations, and (ii) the regression curve [23]. The latter measure evaluates the quality of the attained NN’s mapping between the input and the desired data.

The authors have found that the MSE evolution data exhibit an error function convergence occurring after about 20 iterations when utilizing the S-SEM-NN with computation time taking less than 5 seconds when performed on a standard local machine while a minimum error occurred after about 60 iterations in more than 10 minutes using the Genetic Algorithm (GA)-based S-SEM processor [15]. The reported MSE calculation using the NN is about \( 3.4 \times 10^{-7} \), which is five orders of magnitude less than the

\(^5\) In asymmetric wire radiation problem, the current can still be predicted using \( N \) poles instead of \( N/2 \) without any change in the algorithm.
Figure 4. A comparison of the estimated S-SEM data and the actual ones obtained from a direct method (MoM or measurement) of the symmetric wire. Here, the MoM code WIPL-D [22] is used to generate test $f_{mn}$ data.

GA-based cost. Moreover, by carefully choosing the S-SEM-NN’s learning and training functions to best operate with the new activation functions $f_{mn}$, it is now possible to reasonably estimate the real S-SEM data of the target as shown in Fig. 4, where a good agreement between reconstructed data in comparison to the target or source data (e.g., through measurement or MoM in our case) is observed. In order to obtain these results, we used the Levenberg-Marquardt back-propagation (trainlm) training function [24] and the gradient descent as a learning processor [21].

It is important to note that since the random field used for testing the network is generated using a Gaussian distribution, the possibility that the NN is memorizing the labels/weights is excluded. The noise power level was varied, and the NN was found still capable of predicting the right field, though with increasing computational cost when the SNR deteriorates. We have used in the testing stage different input datasets (far-field angles and plane cuts), partitioned into separate training and testing datasets, and found that the NN trained on one plane cut is capable of predicting completely new plane cuts not used in the original training process. According to conventional ML, this constitutes a good indication that the NN did not memorize the labels since the NN has not been even shown those data.

Here, the noise is defined as a Gaussian random distribution with a specified variance that is added linearly to the phasor field data of the EM system. The idea is similar to adding noise to incoming fields in direction-of-arrival estimation techniques such that the noise is generated only once using a random Gaussian distribution (randn), then added to the incoming fields and mitigated on the receiving sensors. We adopted the same principle in testing our network against normal noise scenarios where the ANN regression analysis acts as a noise cancellation filter.

The second example assumes a relative tilting of one wire section with respect to the other. An L-shaped wire with an inclination angle of $90^\circ$ as depicted in Fig. 5(a) is used to test the NN’s performance in estimating the tilting angle besides the S-SEM information. To do so, we now include the geometrical data $L_m$ in in the internal model parameters set of the S-SEM-NN architecture. These will appear as internal (tuning or fixed) parameters in the proposed S-SEM-NN’s new activation function $f_{mn}$, see Eq. (2).

We first conduct the NN’s training operation by supplying the known far field pattern (here computed using MoM) of the L-shape wire system, where angles are fed at the input terminals while the complex $f_{mn}$ values are required at the output layers. Using the gradient descent algorithm to

$$n = \sqrt{\sigma^2} [\text{rand}(N) + j\text{rand}(N)]/\sqrt{2},$$

where $\sigma^2$ is the noise variance that varies between 0.1 to 0.001 and $N$ is the size of the dataset.

\footnote{The Gaussian noise is defined as follow}
search for optimum $\hat{L}, s_{mn}$, and $\alpha_{mn}$, good reconstruction of the S-SEM data (see Fig. 6) similar to Fig. 4 was obtained, with the MSE training cost converging after approximately 200 iterations in less than 10 seconds. The tilting angle found by the S-SEM-NN is about $90.627^\circ$ in comparison to $82^\circ$ using the S-SEM-GA. The S-SEM-GA running time is approximately 2–3 min with a total number of 1000 iterations.

3.3. Performance Evaluation

In order to evaluate the network performance, two important factors have to be addressed. The first is the ability of the network to restore the S-SEM data in which can be seen from the pole plot. Again, this part is associated with the network in reinforcing the user with these data during the classification process. The second is the regression plot that certifies a perfect mapping between the input data to an output data from a library dataset. These basic regression performance data are given in Fig. 7 for both symmetric and L-shaped wires, and they both indicate excellent performance.

The selected wire systems are operating at 1 GHz with each of an electric length of half-wavelength. Note that previous published research papers on S-SEM and T-SEM have extensively investigated the extraction of frequency data from the S-SEM pole map and relating it to T-SEM and vice-versa. However, this is not the focus and the intention of the proposed work but rather performing the inverse source problem from known datasets. The algorithm was tested on a local machine processor of 3.2 Gb/s.
Figure 7. Regression analysis plot (a) symmetric wire (b) L-shape wire.

and a RAM size of 32 GB. All GPU and parallel pools were deactivated to acquire the least network performance under usual conditions.

Our network is considered as a shallow network and most importantly with special neurons that carry an electromagnetic signature, hence the network accomplishes the training phase is few seconds between 1–5 secs. Note that the data are structured as a serial stream with one system running at a time. Consequently, the training time should fairly increase when performing data parallelization, i.e., one matrix for all systems with each column representing an EM system.

4. CONCLUSION

We proposed a new genre of neural networks (NN) derived from electromagnetic theory and suitable for applications to inverse source modeling where captured far field data are available for processing. The NN is based on the S-SEM technique and provides a convenient efficient and powerful computational machine learning tool capable of generalizing from partial far-field datasets to complete spatial current distribution on the unknown source, making them suitable for sensor and radar applications. The very robust and highly efficient stochastic back-propagation gradient descent algorithm was utilized to train the derived S-SEM-NN for two basic applications involving thin-wire radiators. It was found that the S-SEM-NN can predict the S-SEM data (spatial resonances and their strengths), which determine the radiating current distribution. The S-SEM-NN was found to be several orders of magnitude faster than the GA in performing the inverse modeling machine learning task.

REFERENCES

1. Ishimaru, A., Electromagnetic Wave Propagation, Radiation, and Scattering: From Fundamentals to Applications, Wiley-IEEE Press, 2017.
2. Wang, Y. and X. Chen, “3-D interferometric inverse synthetic aperture radar imaging of ship target with complex motion,” IEEE Transactions on Geoscience and Remote Sensing, Vol. 56, No. 7, 3693–3708, Jul. 2018.
3. Mooney, J. E. and L. S. Riggs, “Robust target identification in white gaussian noise for ultra wide-band radar systems,” IEEE Transactions on Antennas and Propagation, Vol. 46, No. 12, 1817–1823, Dec. 1998.
4. Bialkowski, K. S., J. Marimuthu, and A. M. Abbosh, “Low-cost microwave biomedical imaging,” 2016 International Conference on Electromagnetics in Advanced Applications (ICEAA), 697–699, Sept. 2016.
5. Christodoulou, C. and M. Georgiopoulos, *Applications of Neural Networks in Electromagnetics*, Artech House, 2001.

6. W. S. S., G. M. Araujo, E. A. B. da Silva, and S. K. Goldenstein, “Facial fiducial points detection using discriminative filtering on principal components,” *2010 IEEE International Conference on Image Processing*, 2681–2684, Sept. 2010.

7. Boerner, W., “Electromagnetic inverse methods and its applications to medical imaging-a current-state-of-the-art review,” *IEEE International Symposium on Circuits and Systems*, Vol. 2, 999–1006, May 1989.

8. Afsari, A. and A. Abbosh, “Fast onsite electromagnetic imaging method for medical applications,” *2018 Australian Microwave Symposium (AMS)*, 83–84, Feb. 2018.

9. Ambrosanio, M., P. Kosmasy, and V. Pascazio, “An adaptive multi-threshold iterative shrinkage algorithm for microwave imaging applications,” *2016 10th European Conference on Antennas and Propagation (EuCAP)*, 1–3, Apr. 2016.

10. Alqadah, H. F. and N. Valdivia, “Distributed radar imaging using a spatially enhanced linear sampling method,” *2013 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 425–428, Sep. 2013.

11. Mittra, R., W. L. Ko, and P. Harms, “Detection of high conductivity objects buried in sea oor sediments,” *Proceedings of IEEE Antennas and Propagation Society International Symposium and URSI National Radio Science Meeting*, Vol. 3, 1426–1429, Jun. 1994.

12. Mikki, S. M., A. M. Alzahed, and Y. M. M. Antar, “The spatial singularity expansion method for electromagnetics,” *IEEE Access*, Vol. 7, 124 576–124 595, Feb. 2019.

13. Mikki, S., A. Hanono, J. Persano, A. Alzahed, Y. Antar, and J. Aulin, “Theory of electromagnetic intelligent agents with applications to MIMO and DoA systems,” *2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, 525–526, Jul. 2017.

14. Schmidhuber, J., “Deep learning in neural networks: An overview,” *Neural Networks*, Vol. 61, 85–117, 2015.

15. Alzahed, A. M., Y. M. M. Antar, and S. M. Mikki, “Electromagnetic deep learning technology for radar target identification,” *2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 579–580, Jul. 2019.

16. Alzahed, A., S. Mikki, and Y. Antar, “Electromagnetic machine learning for inverse modeling using the spatial singularity expansion method,” *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, 1–1, 2020.

17. Mikki, S. M. and Y. M. M. Antar, *New foundations for Applied Electromagnetics: The Spatial Structure of Fields*, Artech House, 2016.

18. Mikki, S. M. and Y. M. M. Antar, “On the fundamental relationship between the transmitting and receiving modes of general antenna systems: A new approach,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 11, 232–235, 2012.

19. Mikki, S. M. and Y. M. M. Antar, “The antenna current Green’s function formalism: Part I,” *IEEE Transactions on Antennas and Propagation*, Vol. 61, No. 9, 4493–4504, Sept. 2013.

20. Mikki, S. M. and Y. M. M. Antar, “The antenna current Green’s function formalism: Part II,” *IEEE Transactions on Antennas and Propagation*, Vol. 61, No. 9, 4505–4519, Sept. 2013.

21. Goodfellow, I., *Deep Learning*, The MIT Press, Cambridge, Massachusetts, 2016.

22. Kohundzija, B. and M. Pavlovic, “Emulating magnetic ferrite tiles properties by wipl-d software suite,” *2017 11th European Conference on Antennas and Propagation (EUCAP)*, 3611–3613, Mar. 2017.

23. Anderson, J. A., *An introduction to Neural Networks*, MIT Press, 1995.

24. Sapna, S., “Backpropagation learning algorithm based on Levenberg Marquardt algorithm,” *Computer Science & Information Technology (CS & IT)*, 393–398, 2012.