Initial Experimental Tests of an ANN-Based Microwave Imaging Technique for Neck Diagnostics

Chiara Dachena, Alessandro Fedeli, *Member, IEEE*, Alessandro Fanti, *Member, IEEE*, Matteo B. Lodi, *Member, IEEE*, Giorgio Fumera, *Member, IEEE*, Matteo Pastorino, *Fellow, IEEE*, and Andrea Randazzo, *Senior Member, IEEE*

***Abstract*— In this letter, a microwave imaging strategy based on an artificial neural network (ANN) is applied, for the first time, to experimental data gathered from simplified neck phantoms. The ANN is used for solving the underlying inverse scattering problem, with the aim of retrieving the dielectric properties of the neck for monitoring and diagnostic purposes. The ANN is trained using simulated phantoms, to overcome the limited availability of experimental data. First, a simple configuration with a liquid-filled glass beaker is tested. Then, simplified 3-D-printed models of the human neck are considered. The preliminary findings indicate the possibility of training the network with numerical simulations and testing it against experimental measurements.**

***Index Terms*— Artificial neural networks (ANNs), biomedical imaging, inverse scattering, machine learning (ML), microwave imaging (MWI).**

I. Introduction

MICROWAVE imaging (MWI) emerged several decades ago as a promising diagnostic technique for biomedical applications [1]. Recently, besides the widespread applications related to breast tumor [2]–[5] and brain stroke detection [6]–[8], the use of MWI for other parts of the body, such as torso and arms [9], [10], has gained attention, too. Specifically, this work deals with MWI for diagnosing and monitoring neck diseases, such as cervical myelopathy [11] and tumors [12], [13]. Indeed, although magnetic resonance imaging (MRI) and computerized tomography (CT) represent the gold standard, they are either expensive or relying on ionizing radiations. MWI represents a complementary technique that, due to its nonionizing nature and the possibility of using cheap components, may allow frequent monitoring. Moreover, since it aims at reconstructing the dielectric properties from scattered-field measurements, it also potentially provides additional diagnostic information.

MWI is based on the solution of a strongly ill-posed and nonlinear inverse problem [1]. Several solving approaches, based on quantitative [3], [14] or qualitative methods [15], [16], have been developed. In this context, we developed in [11] a quantitative technique based on a regularization procedure in Lebesgue spaces for the diagnosis of cervical myelopathy. Inversion methods based on machine learning (ML) paradigms have also been recently devised [17], considering both convolutional [18] and fully connected artificial neural networks (ANNs) [19]. In particular, the applicability of a fully connected ANN for neck tumor detection has been studied, through numerical simulations, by us in [12]. Such an ANN can be in principle applied also to other parts of the body, provided that a proper training dataset is adopted. It is worth noting that other ANNs have been tried on other body parts, e.g., for breast cancer [19] and brain stroke detection [20].

Although ANNs have been found to be quite effective, one of the main obstacles in their practical application is the amount of data required in the training phase [21]. In fact, in MWI, the requirement of massive quantities of data for the training phase does not allow training ANNs with experimental data only. For this reason, most works are limited to numerical analyses or use data from well-known datasets (e.g., the Fresnel one [22]–[28]), and few contributions make use of custom experimental data in the test phase [20]. Therefore, the application of ANNs to experimental data still represents a great challenge, which needs to be faced to uncover the full potential of ML for MWI.

The aim of this letter is to report some initial tests against experimental data of an ANN that has been trained with numerical models only. To the best of the authors' knowledge, this is the first time that a deep learning approach is used with experimental data gathered from neck-mimicking phantoms.

This letter is organized as follows. An overview of the MWI method and the ANN architecture is provided in Section II. Preliminary experimental results are reported and discussed in Section III. Section IV reports some concluding remarks.

II. Overview of the Imaging Method

Let us consider the tomographic imaging system summarized in Fig. 1. S equally spaced antennas are located around the neck on a circumference of radius r_n and acquire

Manuscript received 20 June 2022; accepted 22 July 2022. *(Corresponding author: Alessandro Fedeli.)*

Chiara Dachena, Alessandro Fedeli, Matteo Pastorino, and Andrea Randazzo are with the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN), University of Genoa, 16145 Genoa, Italy (e-mail: chiara.dachena@edu.unige.it; alessandro.fedeli@unige.it; matteo.pastorino@unige.it; andrea.randazzo@unige.it).

Alessandro Fanti is with the Department of Electrical and Electronic Engineering, University of Cagliari, 09123 Cagliari, Italy, and also with the Istituto Nazionale di Fisica Nucleare-CA, Complesso Universitario di Monserrato, 09042 Cagliari, Italy (e-mail: alessandro.fanti@unica.it).

Matteo B. Lodi and Giorgio Fumera are with the Department of Electrical and Electronic Engineering, University of Cagliari, 09123 Cagliari, Italy (e-mail: matteob.lodi@unica.it; fumera@unica.it).

Color versions of one or more figures in this letter are available at https://doi.org/10.1109/LMWC.2022.3194805.

Digital Object Identifier 10.1109/LMWC.2022.3194805

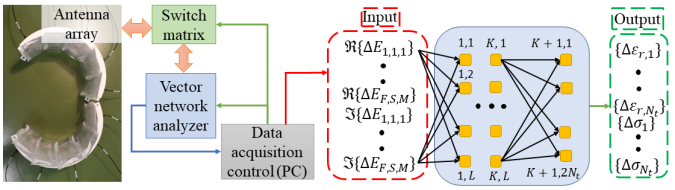

Fig. 1. Structure of the proposed MWI system and the inversion approach based on a fully connected neural network.

measurements of the electric field with a multi-illumination, multistatic, and multifrequency approach. In this configuration, S views are considered, with $M = S - 1$ receiving antennas for each view, and data are measured at F frequencies. Between the antennas and the neck, a layer of matching medium (with dielectric permittivity ε_{mm} and electric conductivity σ_{mm}) is present. The developed approach aims at reconstructing the differences in the dielectric properties inside the investigation domain D_t with respect to a given reference configuration. In the following, $\Delta E_{f,s,m}$ denotes the difference between the z-component of the total electric field due to the actual and reference configurations (for the fth frequency, the sth view, and the mth measurement point). Moreover, $\Delta\varepsilon_{r,n}$ and $\Delta\sigma_n$ are the differential dielectric properties in the nth subdomain ($n = 1, \ldots, N_t$) in which D_t is discretized. As it is well known, the retrieval of $\Delta\varepsilon_{r,n}$ and $\Delta\sigma_n$ starting from $\Delta E_{f,s,m}$ constitutes a nonlinear and ill-posed inverse problem [1]. An ANN-based inversion approach, originally proposed in [12], is adopted in this letter. As shown in Fig. 1, the input of the network is an array containing the real and imaginary parts of $\Delta E_{f,s,m}$, i.e., $\mathbf{I} = [\Re\{\Delta E_{1,1,1}\}, \ldots, \Re\{\Delta E_{F,S,M}\}, \Im\{\Delta E_{1,1,1}\}, \ldots, \Im\{\Delta E_{F,S,M}\}]^t$. The ANN is composed of K hidden layers, each one containing L neurons characterized by a weight vector $\mathbf{W}_{k,l}$ and a bias $b_{k,l}$, with $k = 1, \ldots, K$ and $l = 1, \ldots, L$. The output $\mathbf{O}_k$ of the kth layer is computed as in [12]. The output layer has $2N_t$ neurons (with weight vector $\mathbf{W}_{K+1,l}$ and a bias $b_{K+1,l}$) and a regression function is added to predict the dielectric properties in the investigation domain, which are contained in the output array $\mathbf{O} = [\Delta\varepsilon_{r,1}, \ldots, \Delta\varepsilon_{r,N_t}, \Delta\sigma_1, \ldots, \Delta\sigma_{N_t}]^t$. The output is computed as $\mathbf{O} = \mathbf{R}[\mathbf{W}^t_{K+1,1}\mathbf{O}_K + b_{K+1,1}, \ldots, \mathbf{W}^t_{K+1,2N_t}\mathbf{O}_K + b_{K+1,2N_t}]^t$, where $\mathbf{R}$ is the rectified linear unit (ReLu) activation function. The Adam method [29] is used in the training phase to minimize the loss function $\mathcal{L} = (1/D_{tr})\big[\sum_i(\|\Delta\varepsilon^i_{r,r} - \Delta\varepsilon^i_{r,a}\|/\|\Delta\varepsilon^i_{r,a}\|) + \sum_i(\|\Delta\sigma^i_r - \Delta\sigma^i_a\|/\|\Delta\sigma^i_a\|)\big]$, where D_{tr} is the number of samples, $\Delta\varepsilon^i_{r,r}$ and $\Delta\varepsilon^i_{r,a}$ are the reconstructed and actual real parts of the relative dielectric permittivity of the ith sample, and $\Delta\sigma^i_r$ and $\Delta\sigma^i_a$ are the corresponding electric conductivities.

III. Preliminary Experimental Results

Some preliminary experimental results are presented in this section to test the reconstruction capability of the proposed ANN. The adopted MWI system is described in [11]. It consists of $S = 10$ slotted cavity-backed bow-tie antennas and a vector network analyzer connected to the antennas through a switch matrix. A mixture of water and 70% glycerin (volumetric), enclosed in polyethylene bags interposed between the antennas and the phantom, is used as matching medium.

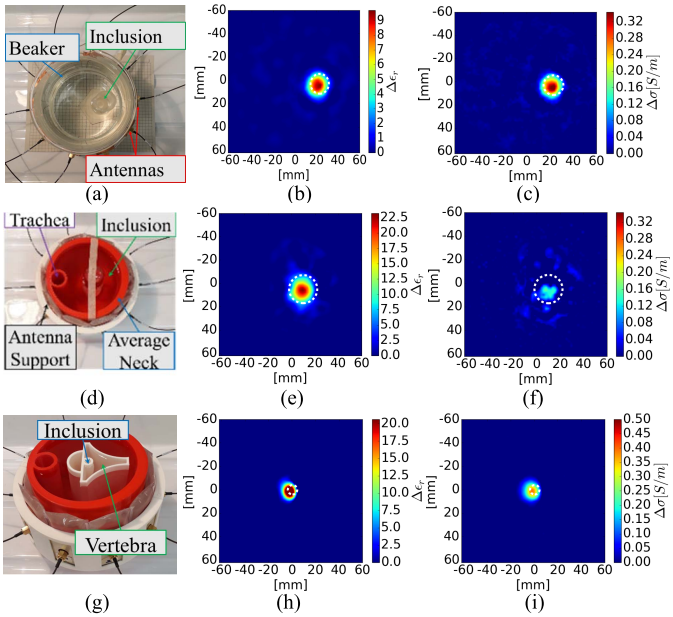

Fig. 2. Configuration of the experimental target and reconstructed maps. (a) Beaker. (b) $\Delta\varepsilon_r$. (c) $\Delta\sigma$. (d) 3-D printed neck "phantom 1" [11]. (e) $\Delta\varepsilon_r$. (f) $\Delta\sigma$. (g) 3-D printed neck "phantom 2" [11]. (h) $\Delta\varepsilon_r$. (i) $\Delta\sigma$.

$F = 4$ frequencies between 600 and 750 MHz with 50-MHz frequency step have been considered for the first two cases, whereas for the third case, $F = 7$ frequencies between 600 and 900 MHz have been used. Such parameters have been chosen based on the analysis performed in [11], where it has been found that frequencies between 500 and 1 GHz, with a matching medium having relative dielectric permittivity between 5 and 60, provide a good tradeoff between low reflection from skin and high penetration into the neck. The developed network has $K = 5$ hidden layers and $L = 448$ neurons for each layer. Such parameters have been chosen according to the analysis performed in [12]. The initial weights of the ANN are drawn from a Gaussian distribution with zero mean value and standard deviation equal to 10^{-2}, whereas the initial biases are set to zero. A constant L^2 norm regularization parameter equal to 10^{-4} is adopted to avoid overfitting, with a mini-batch size of 256. An initial learning rate equal to 10^{-3} is used in the Adam method, with a maximum of 500 epochs.

In the first case, a circular glass beaker with external radius $r_b = 53.5$ mm filled with a 70% glycerin:water mixture and containing a circular inclusion with radius $r_{i1} = 8.5$ mm placed at (23, 3) mm is considered [see Fig. 2(a)]. The inclusion is filled with an 80% glycerin:water mixture. The dielectric properties of these mixtures have been estimated from reflection measurements performed on a liquid-filled section of a short-circuited coaxial line [11]. The network has been trained using a dataset of $D = 10\,000$ numerically simulated configurations. Since the aim is to detect just the inclusion inside the phantom, each target is modeled as a circular cylinder with radius $r_n = 6.2$ cm (to account for beaker and matching medium) and dielectric properties $\varepsilon_{\text{mm}} = 43\varepsilon_0$ S/m and $\sigma_{\text{mm}} = 0.8$ S/m [11], in which a cylindrical inclusion with random radius $r_{t1} \in (5, 10)$ mm and position (inside the beaker) is present. The dielectric permittivity of the inclusion randomly ranges from $30\varepsilon_0$ to $40\varepsilon_0$ and its electric conductivity is in the range $[0.4, 0.7]$ S/m. The investigation domain is discretized into $N_t = 3024$ square cells of side

2 mm. The scattered fields are computed with a custom solver based on the method of moments [30]. A subset of 90% of samples is used to train the network, whereas the remaining ones are used for validation. The obtained values of the loss function are 1.03 and 1.1 for the training and validation sets, respectively. The reference configuration is the phantom without the inclusion, i.e., a homogenous circular structure with dielectric properties equal to those of the matching medium.

In order to use the experimental data for testing the ANN, a scattered-field calibration has been performed [31], where the fields measured in the presence of a known target are used to derive a set of scaling coefficients, one for each transmitter–receiver pair and frequency. The calibrated measurements are then obtained as $\Delta E_{f,s,m} = \Delta E^{M}_{f,s,m} \Delta E^{SI,R}_{f,s,m} \Big/ \Delta E^{M,R}_{f,s,m}$, where $\Delta E^{M}_{f,s,m}$ are the uncalibrated data and $\Delta E^{M,R}_{f,s,m}$ and $\Delta E^{SI,R}_{f,s,m}$ are the experimental and simulated fields for the calibration target, respectively. In particular, in this case, the calibration target is a circular inclusion with radius $r_{i2} = 6.2$ mm that has been placed inside the beaker at $(23, 5)$ mm. The reconstructions of $\Delta\varepsilon_r$ and $\Delta\sigma$ are reported in Fig. 2(b) and (c). The internal inclusion is correctly localized, and a quite good estimation of the dielectric properties is obtained although $\Delta\varepsilon_r$ is underestimated.

In the second test case [see Fig. 2(d)], a 3-D-printed model of the human neck is considered ("phantom 1" of [11]). The considered printed structure contains some of the main anatomical details, such as the air-filled trachea in the anterior part of the neck and the external fat layer [11]. The 3-D-printing material is polylactic acid (PLA), whose dielectric properties are $\varepsilon_{\text{PLA}} = 3\varepsilon_0$ and $\sigma_{\text{PLA}} = 0.001$ S/m. The trachea is characterized by $\varepsilon_{\text{tra}} = \varepsilon_0$, $\sigma_{\text{tra}} = 0$ S/m, and the neck filling has $\varepsilon_{\text{neck}} = 43\varepsilon_0$ and $\sigma_{\text{neck}} = 0.8$ S/m. An inclusion with radius $r_{s1} = 12$ mm has been placed inside the neck phantom at $(10, 5)$ mm. The network has been trained by using a dataset built from a numerical model of the phantom (containing the external layer and the trachea), in which inclusions with different random positions (inside the inner part) and radius $r_{t2} \in (3, 15)$ mm have been located. The dielectric properties vary in the same ranges of the previous case. The reference scenario is the phantom without inclusions. To calibrate the data, an inclusion with radius $r_{s2} = 5$ mm has been placed at $(10, 9)$ mm. The reconstructed values of $\Delta\varepsilon_r$ and $\Delta\sigma$ are shown in Fig. 2(e)–(f). In this case, too, the inclusion is well detected, although $\Delta\varepsilon_r$ is slightly overestimated and $\Delta\sigma$ is underestimated.

Finally, a more complex 3-D-printed model with vertebral column filled with glycerin ($\varepsilon_{\text{ver}} = 9.77\varepsilon_0$ and $\sigma_{\text{ver}} = 0.36$ S/m) is considered ["phantom 2" of [11], Fig. 2(g)]. A circular printed inclusion with radius $r_{v1} = 5$ mm has been located inside the vertebra at $(0.5, 0)$ mm. The numerical models used in the training procedure reproduce the printed structure and contain random inclusions inside the vertebra with radius in the range $r_{t3} \in (3, 12)$ mm. The ranges of the dielectric properties are the same as before, and the structure without inclusion is used as a reference scenario. To calibrate the data, an inclusion with radius $r_{v2} = 9$ mm has been considered inside the vertebra at $(5, 0)$ mm. The reconstructed maps of the dielectric properties are shown in Fig. 2(h)–(i). Even in this more involving case, the inclusion is well detected in both $\Delta\varepsilon_r$ and $\Delta\sigma$, and the reconstructed values are quite accurate.

TABLE I

RELATIVE RECONSTRUCTION ERRORS (DIMENSIONLESS)

Case	e_{tot,ε_r}	$e_{tot,\sigma}$	e_{b,ε_r}	$e_{b,\sigma}$	e_{in,ε_r}	$e_{in,\sigma}$	e_r	$e_{cd,x}$	$e_{cd,y}$
#1	0.007	0.013	0.004	0.008	0.169	0.259	0.001	0.026	0.100
#2	0.011	0.033	0.006	0.008	0.141	0.703	0.146	0.013	0.048
#3	0.008	0.006	0.006	0.005	0.262	0.328	0.250	0.020	0.001

For completeness, the performances of the approach have been evaluated using the mean relative errors $e_{\{tot,b,in\},\gamma} = (1/N_{\{t,b,in\}}) \sum_{\mathbf{r}_n \in D_{\{t,b,in\}}} (|\gamma_{\text{rec}}(\mathbf{r}_n) - \gamma_{\text{act}}(\mathbf{r}_n)|/|\gamma_{\text{act}}(\mathbf{r}_n)|)$, where γ_{rec} and γ_{act} are the reconstructed and actual values of $\Delta\varepsilon_r$ or $\Delta\sigma$, respectively, $\mathbf{r}_n$ is the center of the nth cell of the investigation domain, $D_{\{b,in\}}$ is the background (b) or the inclusion (in) region, and $N_{\{b,in\}}$ is the corresponding number of cells. Moreover, the relative errors on the estimated radius and center of the inclusion have been evaluated, too. In particular, the radius error, e_r, is computed as $e_r = |r_r - r_a|/|r_a|$, where r_a and r_r are the actual and estimated radiuses, respectively, the latter obtained by setting a threshold equal to 40% on the maximum reconstructed value of $\Delta\varepsilon_r$. Similarly, the center errors, $e_{cd,x}$ and $e_{cd,y}$, are the relative errors on the x and y coordinates of the reconstructed inclusion. The errors are reported in Table I for the three considered cases. In the second configuration, e_{tot,ε_r} and $e_{tot,\sigma}$ and e_{b,ε_r} and $e_{b,\sigma}$ are slightly higher than in the first case. Indeed, in Fig. 2(e)–(f), some artifacts in the background are present. Moreover, $e_{in,\sigma}$ is about twice the value of the first case, and this is motivated by the underestimation of the reconstructed values of $\Delta\sigma$. In all cases, low radius and position errors are obtained, thus allowing to identify and locate the inclusion suitably. In the last case, errors are generally comparable with the simplest configuration.

Finally, the robustness of the method with respect to uncertainties in the values of the dielectric properties used in the training set was tested (considering test cases #1 and #2). In particular, the network has been trained with a model having a variation of $\pm 5\%$ of the dielectric properties (outside the inclusion). The average increase in the relative errors is 2.45% for e_{tot,ε_r} and 9.15% for $e_{tot,\sigma}$.

IV. CONCLUSION

In this letter, initial experimental tests of an ANN approach for quantitative MWI of the human neck have been discussed. The considered ANN, which is composed of fully connected layers, has been trained by means of numerical simulations and then tested, for the first time, against experimental data. Three different phantoms with inclusions have been considered: a liquid-filled glass beaker, and a simplified and a complex 3-D-printed model of the human neck. The preliminary results, although obtained with simplified phantoms in a laboratory environment, are promising. In all cases, it has been possible to identify with quite good accuracy position, dimension, and dielectric properties of the inclusion, as summarized in Table I. Future works will be aimed at a more extensive analysis of the performance, also including more advanced phantoms and possibly real necks. More complex neural network architectures will be considered, too. The potential extensions of the approach to other frequency bands (with the necessary tradeoffs in terms of attainable accuracy) and to different parts of the body will be addressed, too.

REFERENCES

[1] M. Pastorino and A. Randazzo, *Microwave Imaging Methods and Applications*. Boston, MA, USA: Artech House, 2018.

[2] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE Microw. Mag.*, vol. 12, no. 7, pp. 78–94, Dec. 2011, doi: 10.1109/MMM.2011.942702.

[3] N. Abdollahi, I. Jeffrey, and J. LoVetri, "Improved tumor detection via quantitative microwave breast imaging using eigenfunction-based prior," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1194–1202, 2020, doi: 10.1109/TCI.2020.3012940.

[4] M. A. Aldhaeebi, K. Alzoubi, T. S. Almoneef, S. M. Bamatraf, H. Attia, and O. M. Ramahi, "Review of microwaves techniques for breast cancer detection," *Sensors*, vol. 20, no. 8, p. 2390, Apr. 2020, doi: 10.3390/s20082390.

[5] S. Hosseinzadegan, A. Fhager, M. Persson, and P. M. Meaney, "Application of two-dimensional discrete dipole approximation in simulating electric field of a microwave breast imaging system," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 3, no. 2, pp. 80–87, Jun. 2019, doi: 10.1109/JERM.2018.2882689.

[6] A. T. Mobashsher and A. Abbosh, "Development of compact directional antenna utilising plane of symmetry for wideband brain stroke detection systems," *Electron. Lett.*, vol. 50, no. 12, pp. 850–851, Jun. 2014, doi: 10.1049/el.2014.0616.

[7] I. Sarwar *et al.*, "Low-cost low-power acceleration of a microwave imaging algorithm for brain stroke monitoring," *J. Low Power Electron. Appl.*, vol. 8, no. 4, p. 43, Nov. 2018, doi: 10.3390/jlpea8040043.

[8] A. Fedeli, V. Schenone, A. Randazzo, M. Pastorino, T. Henriksson, and S. Semenov, "Nonlinear S-parameters inversion for stroke imaging," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 3, pp. 1760–1771, Mar. 2021, doi: 10.1109/TMTT.2020.3040483.

[9] A. Zakaria, A. Baran, and J. LoVetri, "Estimation and use of prior information in FEM-CSI for biomedical microwave tomography," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1606–1609, 2012, doi: 10.1109/LAWP.2012.2237537.

[10] D. Gibbins, D. Byrne, T. Henriksson, B. Monsalve, and I. J. Craddock, "Less becomes more for microwave imaging: Design and validation of an ultrawide-band measurement array," *IEEE Antennas Propag. Mag.*, vol. 59, no. 5, pp. 72–85, Oct. 2017, doi: 10.1109/MAP.2017.2731198.

[11] C. Dachena, A. Fedeli, A. Fanti, M. B. Lodi, M. Pastorino, and A. Randazzo, "Microwave imaging for the diagnosis of cervical diseases: A feasibility analysis," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 5, no. 3, pp. 277–285, Sep. 2021, doi: 10.1109/JERM.2020.3042711.

[12] C. Dachena *et al.*, "Microwave imaging of the neck by means of artificial neural networks for tumor detection," *IEEE Open J. Antennas Propag.*, vol. 2, pp. 1044–1056, 2021, doi: 10.1109/OJAP.2021.3121177.

[13] M. T. Bevacqua, G. G. Bellizzi, T. Isernia, and L. Crocco, "A method for effective permittivity and conductivity mapping of biological scenarios via segmented contrast source inversion," *Prog. Electromagn. Res.*, vol. 164, pp. 1–15, 2019, doi: 10.2528/PIER18071704.

[14] C. Estatico, A. Fedeli, M. Pastorino, A. Randazzo, and E. Tavanti, "A phaseless microwave imaging approach based on a Lebesgue-space inversion algorithm," *IEEE Trans. Antennas Propag.*, vol. 68, no. 12, pp. 8091–8103, Dec. 2020, doi: 10.1109/TAP.2020.2999789.

[15] X. Li and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microw. Wireless Compon. Lett.*, vol. 11, no. 3, pp. 130–132, Mar. 2001, doi: 10.1109/7260.915627.

[16] D. Byrne, M. Sarafianou, and I. J. Craddock, "Compound radar approach for breast imaging," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 40–51, Jan. 2017, doi: 10.1109/TBME.2016.2536703.

[17] M. Li *et al.*, "Machine learning in electromagnetics with applications to biomedical imaging: A review," *IEEE Antennas Propag. Mag.*, vol. 63, no. 3, pp. 39–51, Jun. 2021, doi: 10.1109/MAP.2020.3043469.

[18] Y. Huang, R. Song, K. Xu, X. Ye, C. Li, and X. Chen, "Deep learning-based inverse scattering with structural similarity loss functions," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4900–4907, Feb. 2021, doi: 10.1109/JSEN.2020.3030321.

[19] M. Ambrosanio, S. Franceschini, V. Pascazio, and F. Baselice, "Microwave breast imaging via neural networks for almost real-time applications," 2021, *arXiv:2103.12522*.

[20] A. Al-Saffar, A. Bialkowski, M. Baktashmotlagh, A. Trakic, L. Guo, and A. Abbosh, "Closing the gap of simulation to reality in electromagnetic imaging of brain strokes via deep neural networks," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 13–21, 2021, doi: 10.1109/TCI.2020.3041092.

[21] K. Edwards *et al.*, "Stored grain inventory management using neural-network-based parametric electromagnetic inversion," *IEEE Access*, vol. 8, pp. 207182–207192, 2020, doi: 10.1109/ACCESS.2020.3038312.

[22] L. Zhang, K. Xu, R. Song, X. Ye, G. Wang, and X. Chen, "Learning-based quantitative microwave imaging with a hybrid input scheme," *IEEE Sensors J.*, vol. 20, no. 24, pp. 15007–15013, Dec. 2020, doi: 10.1109/JSEN.2020.3012177.

[23] K. Xu, C. Zhang, X. Ye, and R. Song, "Fast full-wave electromagnetic inverse scattering based on scalable cascaded convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–11, 2022, doi: 10.1109/TGRS.2021.3093100.

[24] Z. Lin *et al.*, "Low-frequency data prediction with iterative learning for highly nonlinear inverse scattering problems," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 10, pp. 4366–4376, Oct. 2021, doi: 10.1109/TMTT.2021.3098769.

[25] R. Song, Y. Huang, K. Xu, X. Ye, C. Li, and X. Chen, "Electromagnetic inverse scattering with perceptual generative adversarial networks," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 689–699, 2021, doi: 10.1109/TCI.2021.3093793.

[26] L. Xiao, J. Li, F. Han, W. Shao, and Q. H. Liu, "Dual-module NMM-IEM machine learning for fast electromagnetic inversion of inhomogeneous scatterers with high contrasts and large electrical dimensions," *IEEE Trans. Antennas Propag.*, vol. 68, no. 8, pp. 6245–6255, Aug. 2020, doi: 10.1109/TAP.2020.2990222.

[27] 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, doi: 10.1109/TAP.2018.2885437.

[28] 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, doi: 10.1109/TGRS.2018.2869221.

[29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[30] J. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Trans. Antennas Propag.*, vol. AP-13, no. 3, pp. 334–341, May 1965, doi: 10.1109/TAP.1965.1138427.

[31] M. Ostadrahimi *et al.*, "Analysis of incident field modeling and incident/scattered field calibration techniques in microwave tomography," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 900–903, 2011, doi: 10.1109/LAWP.2011.2166849.