A Deep Learning Architecture for *Augmented* Shape Reconstruction via Microwave Imaging

Álvaro Yago Ruiz*§, Marija Nikolic Stevanovic‡, Marta Cavagnaro§*, Lorenzo Crocco*,
*CNR-IREA National Research Council of Italy, Institute for Electromagnetic Sensing of Environment,
Napoli, Italy, crocco.l@irea.cnr.it, yago.a@irea.cnr.it
§DIET, Sapienza University of Rome, Rome, Italy, marta.cavagnaro@uniroma1.it
‡University of Belgrade, Belgrade, Serbia, mnikolic@etf.bg.ac.rs

***Abstract*—In this paper, an innovative microwave imaging approach that combines deep learning techniques and qualitative inversion methods is presented. In particular, the proposed approach is meant for imaging piece-wise homogeneous targets and aims at providing an *augmented* morphological reconstruction, which not only retrieves the shape of the targets, but also the spatial variations of the permittivity values. Such an information is not displayed by qualitative inversion methods; however it is efficiently encoded in the gradient of the unknown contrast. In particular in this paper, a physics-assisted deep learning technique, where domain knowledge is given in the inputs of a U-Net architecture, is developed. The domain knowledge is provided by the qualitative image of the unknown targets obtained using the orthogonality sampling method, thus allowing the architecture to provide, once trained, a fully automated and real-time prediction. An initial assessment for the approach with synthetic data is provided.**

I. Introduction

Microwave imaging (MWI) technology enables accurate inspections and explorations of unknown scenarios not directly accessible. As such, it is relevant to several applications, e.g., biomedical imaging, subsurface sensing, food security monitoring or through-wall imaging. Many techniques have been proposed in the literature to cope with the underlying non-linear and ill-posed inverse scattering problem (ISP) [1]. Among them, qualitative imaging approaches, which aim at recovering just morphological properties of unknown targets from the knowledge of the field they scatter, are worth mentioning [2]. More recently, deep learning (DL) has attracted the attention of many researchers who have explored its potential in the solution of inverse scattering problems [3].

In a previous study [4], a combined framework using qualitative imaging and deep learning (DL) was developed to set an automatic framework for the reconstruction of the shape of unknown targets from their scattered fields encoded as binary images. The chosen qualitative imaging technique, the orthogonality sampling method (OSM), was selected for the advantage over other typical qualitative imaging techniques for its computational efficiency and the fact that it provides a user-independent regularized solution. The DL architecture of choice was a fully convolutional network called U-Net. Among the positive features of the architecture, its non-iterative nature allowed it to work jointly with OSM. Also, the fact it performs image-to-image processing makes it a remarkable option in image enhancement frameworks.

In this paper, an improved framework combining OSM and DL is developed to provide an *augmented* reconstruction of the targets. The term *augmented* refers to the fact that the resulting image not only predicts the morphology of the targets, but also the relative changes in the permittivity values. The rationale behind the proposed approach lies on two circumstances. First, assuming piece-wise homogeneous targets, the gradient of the contrast function efficiently encodes both the information on the target's shape and the different permittivity values. Secondly, the OSM indicator function is related to the radiating part of the contrast sources [5], thus embeds both information on the shape and the (relative) spatial variations of the permittivity. Accordingly, the proposed architecture exploits the OSM as physics-assisted learning [3] and assumes as ground truth the gradient of the unknown contrast function.

In the following, the proposed automated quantitative inversion framework is detailed and assessed in the canonical 2D scalar problem (TM polarized fields) in free space.

II. Formulation of the problem

Let Ω denote the imaging domain which hosts the cross-section Σ of a collection of lossless targets invariant along one direction (say the z-axis). The targets are embedded in a homogeneous and lossless medium of relative permittivity ε_b and each target is characterized by a relative dielectric permittivity $\varepsilon(\underline{r})$, with $\underline{r} = (x, y)$. All materials are supposed to be non-magnetic, i.e., the magnetic permeability is everywhere that of vacuum, μ_0.

The unknown targets are probed with TM-polarized incident fields E_{inc}, transmitted by a set of antennas located in $\underline{r}_t \in \Gamma$, with Γ being a closed curve located in the far-zone of Ω.

The antennas radiate in the frequency band $F = [f_{min}, f_{max}]$. For each transmitter, the interaction between the incident field and the targets gives raise to the scattered field E_s. The superposition of these two fields is the total field $E = E_{inc} + E_s$, which is measured by a set of receivers that, without any loss of generality, is assumed to be located on Γ as well, with the receiver position being $\underline{r}_s$.

For each frequency $f \in F$, the overall phenomenon is cast through a Fredholm type integral equation as:

$$E_s(\underline{r}_s, \underline{r}_t) = \int_{\Omega} G(\underline{r}_s, \underline{r}')\tau(\underline{r}')\, E(\underline{r}', \underline{r}_t)\, d\underline{r}', \qquad (1)$$

where $G(\underline{r}_s, \underline{r}')$ is the Green's function of the assumed homogeneous background medium and $\tau(\underline{r}) = \varepsilon(\underline{r})/\varepsilon_b - 1$ is the contrast function encoding the properties of the targets.

The problem which is faced is the retrieval of the changes in contrast $\nabla\tau$ of targets, from measurements of the scattered fields E_s. However, directly tackling equation (1) implies solving a non-linear and ill-posed problem.

III. Outline of the Proposed Approach

A. Qualitative Inversion: Orthogonality Sampling Method

OSM is a qualitative method which exploits the measured scattered field E_s and the knowledge of the Green's function to form an image of the unknown target shape [5], [6].

The OSM indicator function is built exploiting the *reduced field* E_{red}, which, for each scattered field, is computed as [5], [6]:

$$E_{red}(\underline{r}_p, \underline{r}_t) = < E_s(\underline{r}_s, \underline{r}_t), G_{(\underline{r}_s, \underline{r}_p)} >_{\Gamma} \qquad (2)$$

where $<,>$ denotes the scalar product on Γ and $\underline{r}_p$ a point of an arbitrary grid sampling the imaging domain Ω. The OSM indicator function is then obtained as:

$$\mathcal{I}(\underline{r}_p) = ||E_{red}(\underline{r}_p, \underline{r}_t)||^2_{\Gamma}, \qquad (3)$$

with $||\ ||$ denoting the L^2 norm. The indicator $\mathcal{I}$ provides an estimation of the morphological properties of the unknown target as it assumes large values where a target is located, $\underline{r}_p \in \Sigma$, and low values elsewhere. Furthermore, the intensity values of the indicator carry information on the electromagnetic properties as well. As can be deduced from Equations (2) and (3), the computation of $\mathcal{I}$ is quick, hence a real-time framework using this technique is viable.

As shown in [5], the reduced field can be seen as an implicit way to enforce the regularization of the underlying inverse problem, as E_{red} is linked to the radiating component of the currents induced in the targets by the incident fields. In addition, since these induced currents depend on the spatial distribution of the electromagnetic properties of the targets, the reduced field also embeds information on the electromagnetic properties of the targets besides the morphological ones.

B. Deep Learning

DL techniques are increasingly used to solve inverse scattering problems [3], thus it is worth considering them as an alternative approach to tackle the problem of estimating $\nabla\tau$ of the targets, by automatically and quickly (real-time) turning the smooth transition of the OSM indicator into sharp boundaries, and visualizing the jumps in contrast present in the imaging domain. In Fig. 1, a diagram of the proposed approach is shown. One advantage of this implementation scheme, mapping the scattered field E_s to a final prediction of the gradients magnitude $\| \nabla\tau \|$ is that reconstructing the gradient instead of the actual contrast τ benefits from the predominantly sparse nature of the gradient which facilitates the task of estimating its values since the majority of the pixels will be put to 0.

From the perspective of DL, the solution of an inverse problem is driven by data [3]. In particular, denoting with F_θ the adopted DL architecture, the key step is the training process in which, for a given set of N data pairs $(\underline{x}_n, \underline{y}_n)$, in which $\underline{x}_n$ denotes the information the network has to be able to retrieve after the training, and $\underline{y}_n$ the corresponding available knowledge that should lead to $\underline{x}_n$, the set of parameters θ which characterizes the network are iteratively optimized through (non-linear) regression.

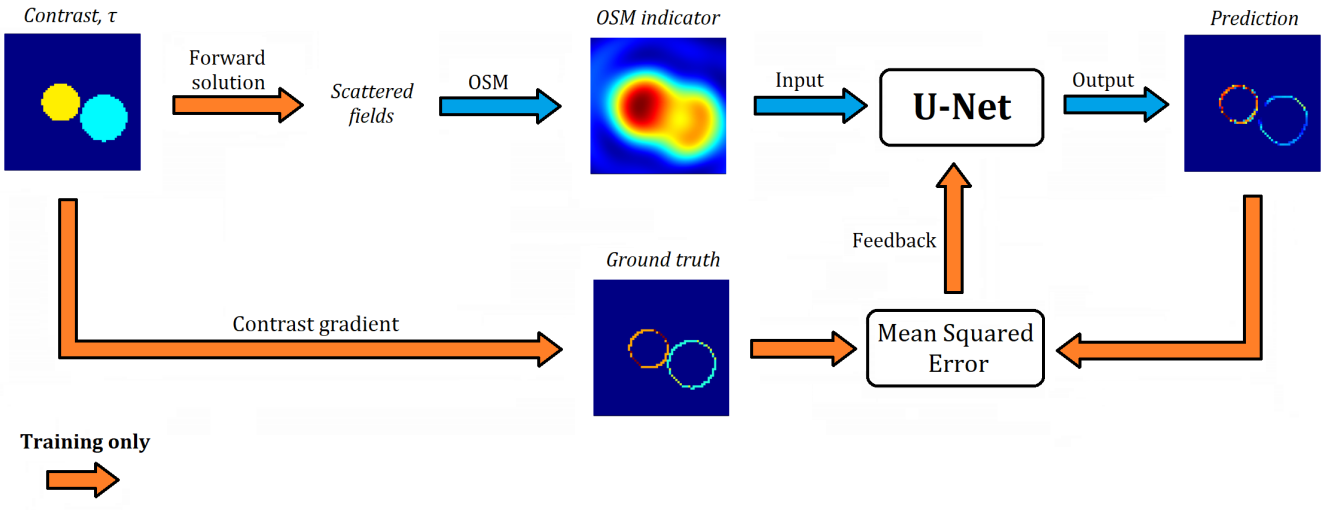

Fig. 1. Physics-assisted deep learning method for fast automatic retrieval of contrast gradient. From the scattered field, OSM indicator maps are built and fed into U-Net. The network automatically retrieves the predicted gradient for those inputs. When U-Net is on the training stage, the feedback loop is used to optimize the inner parameters of the network by minimizing mean squared error between the ground truth and the prediction of the samples.

As far as the architecture F_θ is concerned, since the problem can be seen as a pixel-wise estimation of $\| \nabla\tau \|$ (i.e., assigning a gradient value to each pixel), a Fully Convolutional Network (FCN) can be exploited [7]. A peculiar implementation of FCNs, called U-Net [8], [9] is explored for our problem. U-Net can be easily configured to generate a gradient map, allowing the user to quantitatively asses the jumps in contrast occurring in the imaging domain, as a consequence of changes in the electromagnetic properties, hence shaping the targets and identifying differences in contrast among them.

In regression problems, a convenient choice of loss function to be minimized is the mean squared error [10]:

$$MSE = \arg\min_{\theta} \frac{1}{N_P} \sum_{i=1}^{N_P} (x_i - \hat{x}_i)^2 \tag{4}$$

where N_P is the total number of pixels per image. Moreover, to avoid introducing regularization terms in the cost function, overfitting is dealt with via dropout [11], that is by randomly deactivating connections in the inner structure of the network during the training. Therefore the estimation of the output becomes $\underline{\hat{x}}_n = F_{\theta,drop}(\underline{y}_n)$.

IV. Numerical Assessment

A. Training and Test Data Generation

In a similar way to what was done in [9], the training of the network was carried out by simulating a number of scattering experiments involving combinations of lossless homogeneous circular cylinders. In particular, two cylinders with variable size, location and permittivity were considered for each simulation. More details of the setup conditions are listed in Table I.

TABLE I
SIMULATIONS FOR TRAINING DATA GENERATION

Size of imaging domain	25×25 cm^2
Image size in pixel	64×64
Pixel size	0.1526 cm^2
Background medium	Air
Number of illuminating antennas	8
Angular spacing between emitters	45^o
Number of receiving antennas	241
Angular spacing between receivers	$2/3^o$
Distance of the source from the center of Ω	167 cm
Distance of the receiver from the center of Ω	167 cm
Frequency band	$[2, 9]$ GHz
Frequency step	1 GHz
Number of frequencies	8
SNR	70 dB
Target radius range	$[1.2, 5]$ cm
Target relative permittivity range	$[1.3, 3.5]$

For the training and the performance assessment, a total set of 4000 scattering experiments was simulated. Among them, 90% were used as the training set and 10% as test set. 4 random samples of the test set are shown in Fig. 2. For each

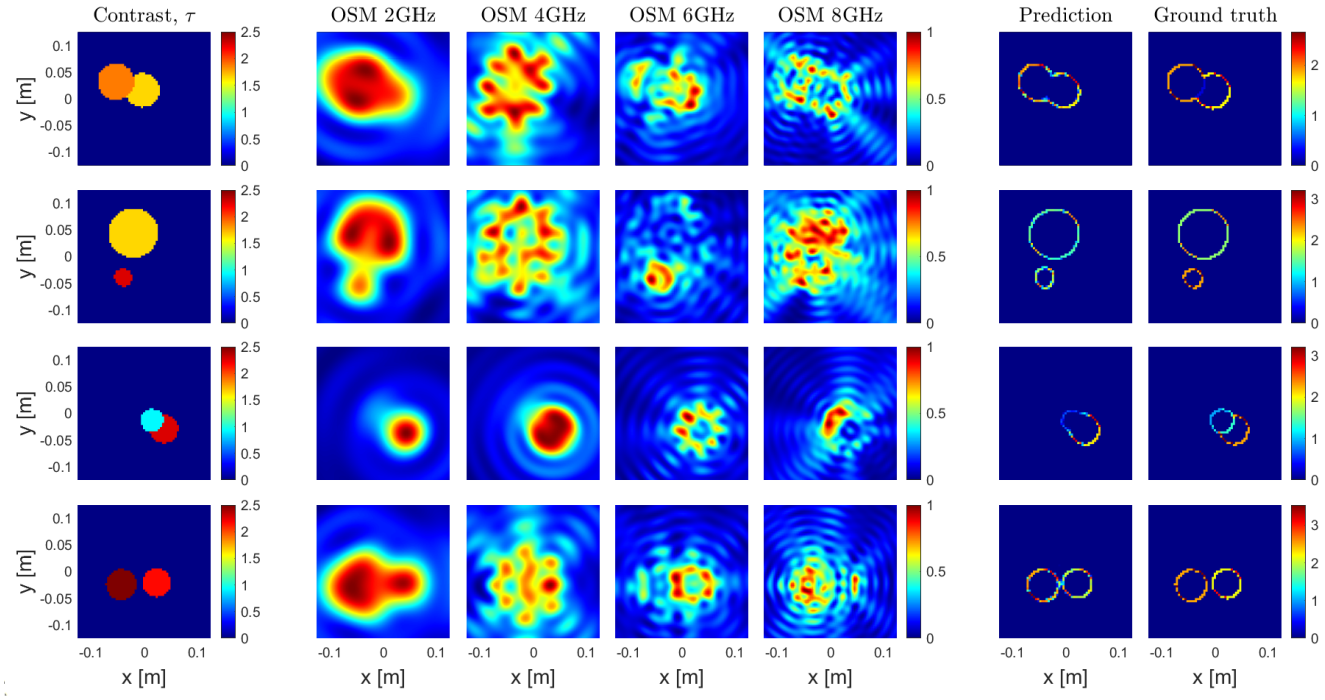

Fig. 2. Imaging results of 4 test samples. One sample per row, with the 1st column depicting the contrast used to compute the forward solution. The following 4 columns represent the OSM indicators from measured scattered fields at 2, 4, 6 and 8 GHz (4 other indicators not shown). The 6th column shows the prediction for the gradient of the contrast made by U-Net, while the last column shows the ground truth gradient.

experiment, an input data-ground truth pair for the network training was generated. The input is a stack of 8 images 64×64 with the obtained OSM indicator maps at the 8 simulated frequencies (Table I). The ground truth is made by 64×64 resulting from the calculation of the discrete gradient of the known contrast spatial distribution.

The optimization of the objective function was carried out for 200 epochs, using Adam optimizer with a learning rate of 10^{-5} and a batch size of 32 [12], whereas the dropout regularization parameter was set to 0.5 .

B. Results and Performance Evaluation

The trained network was assessed with test data. To assess the performance, the mean squared error in equation (4) was calculated and averaged across the 400 samples of the test set. A MSE of 2.2% is reported. However, the unbalance between zero-valued pixels versus the rest present in the images weights down the error and therefore its interpretation should be done cautiously. Such a low error could be interpreted as boundary detection accuracy metric, as clearly proven in Fig. 2.

To provide a further assessment, the MSE was computed and averaged only considering the pixels whose ground truths had non-zero gradient values. Such a modified MSE resulted in a 28.3% error which is representative of the error done by the network when predicting the quantitative values of the gradient.

V. Conclusion

In this work, an innovative MWI framework for quick and automatic retrieval of the shapes and partial information on the EM properties of targets from measured scattered fields has been presented. To this end, the processing approach combines qualitative imaging and deep learning.

In particular, assuming piece-wise homogeneous targets, the developed architecture was trained to retrieve the gradient of the unknown contrast function, as this embeds information on both the shape and the permittivity changes of the target. Aiming to reconstruct the gradient simplifies the optimization process since the network learns to identify sparse solutions. In doing so, the network takes advantage of the physics-assisted learning supplied by the qualitative reconstruction of the targets.

The performance of the processing architecture was tested with simulated data and quantitatively assessed, showing compelling capabilities for a real-time computing framework. Further examples will be presented at the conference.

Acknowledgment

This work has been supported by the EMERALD project funded by the European Union Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 764479.

References

[1] M. Pastorino, *Microwave imaging*. John Wiley & Sons, 2010, vol. 208.

[2] F. Cakoni and D. L. Colton, *A qualitative approach to inverse scattering theory*. Springer, 2014, vol. 767.

[3] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Progress In Electromagnetics Research*, vol. 167, pp. 67–81, 2020.

[4] A. Yago, M. Cavagnaro, and L. Crocco, "Deep learning-enhanced qualitative microwave imaging: Rationale and initial assessment," in *Proceedings of the IEEE Conference on Antennas and Propagation*, 2021.

[5] M. T. Bevacqua, T. Isernia, R. Palmeri, M. N. Akinci, and L. Crocco, "Physical insight unveils new imaging capabilities of orthogonality sampling method," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 5, pp. 4014–4021, 2020.

[6] R. Potthast, *Orthogonality sampling for object visualization*, 2007.

[7] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.

[9] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 1849–1860, 2018.

[10] M. Rimer and T. Martinez, "Classification-based objective functions," *Machine Learning*, vol. 63, no. 2, pp. 183–205, 2006.

[11] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[12] S. Shalev-Shwartz and S. Ben-David, "Stochastic gradient descent," in *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014, 150–166. DOI: 10 . 1017 / CBO9781107298019.015.