A Physics-Assisted Deep Learning Microwave Imaging Framework for Real-Time Shape Reconstruction of Unknown Targets

Álvaro Yago Ruiz, Graduate Student Member, IEEE, Marta Cavagnaro, Member, IEEE, and Lorenzo Crocco, Senior Member, IEEE

Abstract—In this article, an innovative approach to microwave imaging, which combines a qualitative imaging technique and deep learning (DL), is presented. The goal is to develop a tool for reliable and user-independent retrieval of the shape of unknown targets from the knowledge of the scattered fields. Qualitative imaging methods are powerful inverse scattering tools, as they provide morphological information in real time. However, their outcome is a continuous map, which has to be hard-thresholded to clearly identify the targets. This thresholding unavoidably results in case-dependent, often user-biased, results. To deal with this issue, a DL approach, based on a physics-assisted deep neural network, is proposed to automatically classify image pixels, i.e., to generate binary masks, separating the targets (foreground) from the background. In particular, the proposed network binarizes the output of a qualitative imaging inversion technique known as the orthogonality sampling method. For the sake of comparison, a DL method is also exploited, which generates the binary masks directly from the scattered fields without any qualitative imaging aid. A quantitative assessment of the performances of both methods and a test on experimental data are provided.

Index Terms—Automatic classification, convolutional neural networks (CNNs), deep learning (DL), inverse scattering, microwave imaging (MWI), multifrequency, U-Net.

I. INTRODUCTION

MICROWAVE imaging (MWI) technology enables accurate inspection and exploration of unknown scenarios not directly accessible. As such, it is relevant to several applications, such as biomedical imaging [1], [2], subsurface sensing [3], food security monitoring [4], or through-wall imaging [5].

The mathematical problem underlying MWI is a nonlinear and ill-posed inverse scattering problem (ISP), and many techniques have been proposed in the literature to achieve reliable and accurate solutions [6]. Among them, qualitative imaging approaches, which aim at recovering just the presence, position, and shape of the unknown targets from the knowledge of the field they scatter, are worth mentioning [7]. As a matter of fact, in many practical cases, retrieving the morphological information of a target is sufficient information [1], and qualitative methods pursue this goal smartly, without introducing approximations and in real time [6], [7].

The typical outcome of qualitative imaging methods is a continuous map of an indicator function, which usually takes on large values where the target is supposed to be located and low values elsewhere [7]. While this can be sufficient to provide qualitative visual information about the target, it does not provide a clear indication of the actual morphological properties. This is particularly important when an objective characterization of the target is needed, or when the morphological information is exploited as a priori knowledge by a quantitative imaging technique aimed at achieving the estimation of the electromagnetic properties of the target. The issue of deriving the morphological information from the indicator is typically tackled by hard-thresholding the continuous map at some ad hoc defined values. However, this is a case-dependent approach, often biased by the user’s expertise [8].

In this article, an innovative approach to tackle the shape retrieval problem is presented. The proposed approach consists of enhancing qualitative imaging with automated classification, based on deep learning (DL), to enable a user-independent procedure.

In recent years, machine learning, and especially its sub-field of DL, proved to be extremely successful in solving complex problems and attracted the attention of many researchers who explored its potential in the solution of ISPs [9]. In this regard, families of approaches can be identified according to the role of DL in the devised framework [9]. For instance, direct learning methods provide a reconstruction of a target directly from the scattered fields. In this family, works have been carried out to improve the accuracy of
well-known neural network architectures, e.g., incorporating a switch layer into an existing architecture [10]. Another option is learning assisted loss function methods, where the goal is to use DL to find the optimal parameters of an iterative optimization algorithm in the solution of the ISP. Notable examples of these methods employ the learning power to find the best searching directions for an iterative solver [11] or retrieve prior information on the scatterers [12], [13]. Finally, with the incorporation of domain knowledge into either the inputs or the internal structure of the DL architecture, the learning is assisted by the physics of the ISP. Most relevant contributions belonging to this family supply neural networks with prior information on the currents induced on the scatterers to retrieve a refined version of those currents [14], [15] but also to recover the dielectric properties of the targets [16], [17].

This article is focused on a physics-assisted approach. In particular, a convolutional neural network (CNN) [18] is exploited to build the DL model, to which domain knowledge achieved through a qualitative imaging algorithm is incorporated. CNNs have been extensively used for classification purposes, in different fields involving imaging [19]–[21]. As a matter of fact, the problem at hand can be seen as a pixel-wise classification (segmentation), where the end goal is the location of structures and their boundaries, instead of a regular image classification problem.

As far as the qualitative imaging approach is concerned, the proposed processing framework exploits the orthogonality of a collection of targets invariant along one direction (say the z-axis). The targets are embedded in a homogeneous and lossless medium of relative permittivity \( \varepsilon_b \), and each target is characterized by a relative dielectric permittivity \( \varepsilon(\vec{r}) \) and an electric conductivity \( \sigma(\vec{r}) \) with \( \vec{r} = (x, y) \). All materials are supposed to be nonmagnetic, i.e., the magnetic permeability is everywhere that of vacuum, \( \mu_0 \).

The unknown targets are probed with TM-polarized incident fields \( E_{inc} \), transmitted by a set of antennas located in \( \mathcal{L} \), with \( \Gamma \) being a closed curve located in the far-zone of \( \Omega \). For each transmitter, the interaction between the incident field and the targets gives raise to the scattered field \( E_s \), and 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 \( \Gamma \) as well, with the receiver position being \( \vec{r}_l \).

The overall phenomenon is cast through a Fredholm type integral equation as

\[
E_s(\vec{r}_l, \vec{r}) = \int_{\Omega} G(\vec{r}_l, \vec{r}_s) \tau(\vec{r}_s) E(\vec{r}_s, \vec{r}) d\vec{r}_s
\]

where \( G(\vec{r}_l, \vec{r}_s) \) is Green’s function of the assumed homogeneous background medium and \( \tau(\vec{r}) = \varepsilon_{eq}(\vec{r})/\varepsilon_b - 1 \) is the contrast function encoding the properties of the targets. \( \varepsilon_{eq}(\vec{r}) = \varepsilon(\vec{r}) + j\sigma(\vec{r})/\omega\varepsilon_0 \) denotes the relative complex permittivity of the targets, with \( j \) being the imaginary unit, \( \omega = 2\pi f \) being the angular frequency for the corresponding working frequency \( f \), and \( \varepsilon_0 \) being the dielectric permittivity of vacuum.

The problem that is faced is the retrieval of the shapes \( \Sigma \) of the targets from measurements of the fields that they scatter. While simpler than the problem of achieving a complete characterization of the targets (i.e., not only their shape but also their dielectric properties), this problem is still nonlinear and ill-posed if (1) is directly solved.

III. QUALITATIVE IMAGING: ORTHOGONALITY SAMPLING METHOD

The above issue has been addressed in the literature by resorting to the so-called qualitative imaging methods [7]. These methods pursue the goal of solving (1) by resorting to an auxiliary problem, which allows to reformulate it in terms of a linear integral equation. However, such an auxiliary problem, while being simpler than the original one, is still ill-posed, thus requiring the introduction of a suitable regularization strategy. This is a nontrivial issue since the estimation of the proper regularizer is an optimization problem often more cumbersome than the solution of the auxiliary problem itself [27].

A remarkable exception is represented by OSM [22], in which regularization is enforced by means of the proper definition of an indicator function [23]. In particular, the...
OSM indicator function is built exploiting the reduced field, $E_{\text{red}}$, which, for each received scattered field, is computed as [22], [23]

$$
E_{\text{red}}(\mathbf{r}_p, \mathbf{L}_s) = \langle E_s(\mathbf{r}_p, \mathbf{L}_s), G(\mathbf{r}_p, \mathbf{L}_s) \rangle \Gamma
$$

(2)

where $\langle , \rangle$ denotes the scalar product on $\Gamma$ and $\mathbf{L}_p$ denotes a point of an arbitrary grid sampling the imaging domain $\Omega$. The OSM indicator function is then obtained as

$$
\mathcal{I}(\mathbf{L}_p) = \| E_{\text{red}}(\mathbf{r}_p, \mathbf{L}_s) \|_{\Gamma}^2
$$

(3)

with $\| \cdot \|$ denoting the $L^2$ norm computed on $\Gamma$. As shown in [23], the reduced field can be seen as an implicit way to enforce the regularization of the underlying inverse problem, as $E_{\text{red}}$ is linked to the radiating component of the currents induced in the targets by the incident fields, i.e., the one that is not affected by ill-conditioning. Having implicit regularization bestows the method with numerical stability, thus making the OSM robust against noise [28].

Taking into account the aforementioned properties of OSM, two advantages over other techniques seem obvious. First, the introduction of an auxiliary linear problem allows a wider range of validity as opposed to Born-like approximations, where assumptions on the targets’ “weak-scattering” nature are adopted. Then, no matter what linear approximation, the problem will be still ill-posed, requiring the enforcement of an explicit regularization (such as the Tikhonov penalty or spectral truncation), whereas OSM does not require such an effort.

Although linear approximations can be partially avoided using backpropagated sources, explicit regularization must be dealt with anyways. Furthermore, the number of backpropagated sources is dictated by the number of transmitters, reducing the applicability of the technique in the proposed DL framework.

The indicator $\mathcal{I}$ provides an estimation of the geometrical properties of the unknown target as it usually assumes large values when $\mathbf{L}_p \in \Sigma$ and low values elsewhere. As anticipated, the evaluation of $\mathcal{I}$ is straightforward and avoids any explicit inversion process, as it only requires the computation of a scalar product in each sampling point to test the orthogonality between the scattered field pattern and Green’s function. Moreover, calculations are performed considerably fast; therefore, it can be assumed that image formation is performed in real time.

It is worth noting that the above formulation of the OSM is valid for single frequency data. When data at multiple frequencies are available, a collection of indicators can be generated, say $\mathcal{I}_k$ for each frequency $k = 1, \ldots, N_F$. Then, a multifrequency indicator can be calculated combining the single frequency indicators as [22]

$$
\mathcal{I}_{MF}(\mathbf{L}_p) = \sum_{k=1}^{N_F} \mathcal{I}_k(\mathbf{L}_p).
$$

(4)

While not requiring the determination of the optimal regularization parameter, the OSM still does not provide an objective estimation of the shape of the target, similar to other qualitative methods. As a matter of fact, the transition of $\mathcal{I}$ from low to high values is smooth, and there is no general criterion to exactly determine if the sampling points in the vicinity of the target’s boundary actually belong to $\Sigma$ or not. In practice, this is solved by hard-thresholding the indicator, which is obviously a suboptimal solution and leads to case-dependent and user-biased results.

IV. PHYSICS-ASSISTED DEEP LEARNING FOR OBJECTIVE SHAPE ESTIMATION

DL techniques are increasingly used to solve ISPs [9] so that it is worth considering them as alternative approaches to tackle the problem of estimating the shapes $\Sigma$ of the targets from measurements of the fields they scatter with the goal of providing an objective and effective estimation.

From the perspective of DL, the resolution of an inverse problem is driven by data [9]. In particular, assuming a supervised learning approach and denoting with $F_\theta$ the adopted DL architecture, the key step is the training process. In it, for a given set of $N$ pairs $(x_n, y_n)$, where $x_n$ denotes the unknown characteristics the network should learn to retrieve and $y_n$ denotes the corresponding available knowledge that leads to $x_n$, the set of parameters $\theta$ that characterize the network are iteratively optimized through (nonlinear) regression

$$
\arg \min_{\theta} \sum_{n=1}^{N} \mathcal{M}(F_\theta(y_n), x_n) + \mathcal{R}(\theta).
$$

(5)

In (5), $\mathcal{M}$ stands for the measure of the mismatch between the outputs of the architecture $F_\theta(y_n)$ and their corresponding ground-truth values $x_n$, and $\mathcal{R}(\theta)$ is the regularization applied to overcome overfitting. To exploit the above general scheme for the specific problem at hand, the training pairs $(x_n, y_n)$, the loss function $\mathcal{M}$, the regularization $\mathcal{R}$, and the architecture $F_\theta$ have to be determined.

As far as the architecture $F_\theta$ is concerned, since the problem can be devised as a classification task (i.e., assigning each pixel to either a target or the background), a special type of CNN, called fully convolutional network (FCN), can be exploited [29]. Due to its inner structure, FCN is capable of retrieving modified versions of an input image. In particular, FCN allows retrieving a direct segmentation map by means of an image-to-image mapping.

A peculiar implementation of FCN, called U-Net [16], [30], is explored here. The original implementation of U-Net is meant to generate a binary mask to allow the user objectively discriminating the foreground (the targets) from the background in the output image. As such, a convenient choice of the loss function is the binary cross-entropy [31], which enforces a maximization of the probability of correctly labeling the pixels of an input image. Moreover, to avoid introducing regularization terms in the loss function, overfitting is addressed via dropout [32], which is by randomly deactivating connections in the inner structure of the network during the training.

As for the training pairs, the first choice can be that of considering the pairs given by the scattered field collected for all illuminations and receivers, and the corresponding binary mask encoding the targets. While being straightforward, this
Fig. 1. Physics-assisted DL method for the automatic generation of binary masks with the MUMU approach. From the scattered fields, a stack of OSM indicator maps is built and fed into U-Net. The network automatically retrieves the predicted binary mask for those inputs. When U-Net is in the training stage, the feedback loop is used to optimize the inner parameters of the network by minimizing the binary cross-entropy between the ground truth and the prediction.

Among qualitative methods, OSM is suitable to implement an automatic and efficient physics-assisted classification architecture. In fact, as explained in Section III, OSM can supply an estimate of the unknown target shape $\Sigma$ in real time (thus not affecting the computational time of the training, neither the performance of the network once put in action). In addition, it does not need an explicit regularization, which would require fixing an input parameter for each sample and may either imply suboptimal or user-biased inputs, depending on the way in which the regularization is enforced.

V. NETWORK IMPLEMENTATION AND OPTIMIZATION

A. Training and Test Data Generation

To build the training set needed to optimize U-Net parameters, an approach similar to the one in [16] has been used since a similar problem (the 2-D canonical ISP in free space) is therein considered. Accordingly, the training is concerned with the simulation of a number of scattering experiments involving combinations of lossless homogeneous circular cylinders enclosed in the imaging domain. For each computed scattered field, the OSM indicator function is built according to (3) or (4) depending on the adopted approach.

In particular, cylinders placed in groups of two or three with variable sizes, locations, and permittivities were considered for the generation of the training set. In doing so, no profile was
allowed to be partially outside of the imaging domain, while some targets could overlap. Details of the setup conditions are listed in Table I. From the table, it could be noted that the training set and initial validation were performed on lossless targets.

Different from [16], the training set does not consider the case of a single target, and given the fact the proposed network aims at retrieving the shape of the targets, a broader range of admissible permittivities was considered. Moreover, the measurement configuration adopted for the training is inspired by the one used in the Fresnel database [25], [26], which is used in Section VI to benchmark the proposed approach. For the training and the performance assessment, a total set of 2000 scattering experiments was simulated. Among them, 90%(1800) were used as the training set and 10%(200) as the test set. The selection of the samples belonging to either set was performed using a tenfold cross-validation scheme, leading to ten experiments with varying train and test sets to be used in the performance evaluation, as detailed in Section VI.

For each experiment, the input data–ground-truth pair for the network training was generated using a proprietary forward solver based on the method of moments. For the direct learning approach, a stack of $N_F$ matrices encoding the scattered field data at the different frequencies made the input data. Each matrix has size $N_R \times N_T$. For the physics-assisted learning, the input is the 64 × 64 pixels image of the indicator $I_{MF}$ for the SISI approach, while the input for MUMU is a stack of $N_F$ matrices each encoding the 64 × 64 pixels of the OSM indicator $I$ at each frequency.

B. Implementation Details

One important feature for faster convergence of the training is image normalization to [0, 1] [33]. Such normalization was performed differently in the two versions of the physics-assisted approach. For the SISI, the four OSM indicator functions were added up, and the final image (the multi-frequency OSM indicator) was normalized afterward, while, in the MUMU, the same four OSM indicators were individually normalized. Finally, the normalization of scattering matrices performed for the direct learning approach can be seen as an equalization of the information across the different frequencies.

Ground-truth 64 × 64 images were obtained from the binarization of the simulated targets. In order to make the output of U-Net suitable for the computation of the binary cross-entropy, a Softmax activation function [18] is employed in the last layer. When Softmax is used, the output becomes an N-stack of images, with N being the number of classes (2 for binary classification) instead of a single image. As a consequence, the ground truth is transformed as well to match the dimensions of the prediction. The transformation of an image where each pixel belongs to a class into an N-stack of images is known as one-hot encoding [34].

The optimization of the binary cross-entropy loss function was carried out using the Adam optimizer [35], with a learning rate of $10^{-5}$ and a batch size of 16. An optimal solution is found after several passes through all samples in the training set. A complete pass of the whole training set is known as an epoch. A 200 epoch-long training with a dropout rate of 0.5 was performed.

C. Performance Evaluation Metrics

To quantitatively assess the performances of the optimized models, several metrics were used.

The first considered metric is accuracy (ACC) [36], which is defined as

\[
ACC = \frac{TP + TN}{TP + TN + FP + FN}
\]

where $TP$, $TN$, $FP$, and $FN$ are true positives, true negatives, false positives, and false negatives, respectively. In the problem at hand, $TP$ was considered to be the count of all pixels correctly labeled as foreground (targets), whereas $TN$ was considered the count of all pixels correctly labeled as background.

However, ACC scores may be influenced by data unbalance, that is, when there are considerably more pixels in one class than in the others. Such unbalance results in $TN$ and $FN$ being much larger than $TP$ and $FP$, hence making the formers dominate (6). Accordingly, a second metric, the Dice similarity coefficient (DCS) [36], was employed. DCS is defined as

\[
DSC = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}.
\]

DSC does not rely on $TN$ to give an estimation of the performance; therefore, it is more robust against data unbalance. Furthermore, it doubles the contribution of $TP$, rewarding algorithms with better target detection capabilities.

Nevertheless, DSC can also suffer from overestimating the performance in situations where there is data unbalance with a higher proportion of pixels in the foreground than in the background, as in those situations in which $TP$ and $FP$ outnumber $FN$. The double contribution of $TP$ to $DSC$
TABLE II

| Method       | ACC  | DSC  | MCC  |
|--------------|------|------|------|
| Direct learning | 0.958 | 0.747 | 0.733 |
| SISI         | 0.978 | 0.884 | 0.874 |
| MUMU         | 0.989 | 0.943 | 0.937 |

worsens the estimation of the performance in such situations, which could lead to an even more optimistic DSC than ACC [37]. For these reasons, a further metric called Matthews correlation coefficient (MCC) [38] was calculated as well

\[ MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{DM}} \] (8)

where \( DM = (TP+FP)(TP+FN)(TN+FP)(TN+FN) \).

Finally, to further appraise the performance of the various approaches, the behavior of the training and validation losses has been observed. The training losses are the values reached by the binary cross-entropy loss function during the training and provide a measure of the quality of the optimized parameters \( \theta \). Simultaneously, validation losses over the test set can be provided.

Jointly appraising training and validation losses can be useful to identify overfit models, whose training losses tend to continuously drop over the epochs without a corresponding drop in the validation losses. In fact, while training losses for overfitting models consistently drop over the epochs, their corresponding validation losses start to increase after a certain number of epochs.

VI. ASSESSMENT WITH SIMULATED DATA

When training is complete, U-Net can be used to make predictions on new data. Accordingly, the 200 samples excluded from the training in each of the tenfold experiments were fed into the network, and the corresponding predictions were retrieved.

For each experiment, the three considered metrics were calculated and averaged over the ten folds. The average performance scores for the direct learning approach, SISI, and MUMU are shown in Table II. Among the experiments, an average proportion of foreground pixels around 10.7% of the total, with a minimum of 4.3% and a maximum of 29.2%, was found, thus, revealing some degree of data unbalance; therefore, unreliable ACC scores should be expected. Consequently, U-Net showed a very optimistic ACC, while the DSC score was lower than ACC, as expected. On the other hand, the MCC score was close to DSC, thus showing the reliability of the two considered metrics in the proposed scenarios.

The binary cross-entropy losses evaluated during the training for the cross-validation fold with the MCC score closest to average (fourth fold) are plotted in Fig. 2. The training losses show a regular decay with an advantage of SISI and MUMU over the direct learning approach. When validation losses are compared, it is clearly evident that direct learning reaches the loss plateau earlier, followed by SISI and MUMU, which is consistent with the performance metrics in Table II. For the sake of comparison, the same number of epochs was used in the three cases though SISI and MUMU were not fully in the plateau.

It is worth mentioning that the validation loss is usually calculated over a subset of the training set different from the test set. This allows for an unbiased evaluation of the models, and it is a good practice when fine-tuning the non-trainable parameters of the network (hyperparameters) [18]. Nevertheless, since SISI and MUMU reached a level of performance sufficiently accurate in the first implementation attempts, no further hyperparameter tweaking was carried out, and the test set was directly employed for the calculation of the validation loss.
Fig. 3. Imaging results of four test samples. One sample per row, with the first column depicting the contrast values, followed by four columns representing the OSM indicator function from simulated scattered fields at 2, 4, 6, and 8 GHz, and the multifrequency OSM indicator in the sixth column. The seventh column depicts the prediction made by U-Net using the direct learning approach. The eighth and ninth columns depict the prediction made by U-Net when using the physics-assisted approach, SISI, and MUMU, respectively. The last column shows the ground-truth binary mask used to evaluate the scattering data.

| TABLE III | PERFORMANCE METRICS OF THE OSM-DL INTEGRATED APPROACH FOR TEST SIMULATED DATA |
|-----------|---------------------------------------------------------------------------------|
|           | Direct                        | SISI                        | MUMU                        |
| ACC       | 0.913                         | 0.950                      | 0.971                        |
| DSC       | 0.632                         | 0.695                      | 0.856                        |
| MCC       | 0.600                         | 0.668                      | 0.844                        |
| ACC       | 0.995                         | 0.992                      | 0.974                        |
| DSC       | 0.786                         | 0.955                      | 0.881                        |
| MCC       | 0.782                         | 0.951                      | 0.867                        |
| ACC       | 0.982                         | 0.992                      | 0.983                        |
| DSC       | 0.935                         | 0.950                      | 0.914                        |
| MCC       | 0.926                         | 0.946                      | 0.906                        |

Four U-Net prediction samples of the fourth fold test set are shown in Fig. 3. The samples were thoroughly chosen to show the processing performance with different contrasts. Their corresponding individual metrics are reported in Table III. It can be seen that these results are consistent with both the individual metrics and average performance metrics of Table II, with particular reference to DSC and MCC. Most notably, the direct learning approach is outperformed by the physics-assisted approach, whose outputs appear closer to the ground truths.

The individual metrics on Table III demonstrate that SISI and especially MUMU achieve comparable high reconstruction accuracies with different combinations of high and low contrast targets, as well as overlapping shapes.

In commenting on such a result, the better performance of MUMU comes with a lack of flexibility since the number of images provided to the network must forcibly be the number of selected frequencies. In contrast, SISI exhibits more flexibility as it only requires a single image as input, independently of the number of selected frequencies.

An input could be constructed by adding up more OSM indicator functions than the four considered in the experiments, thus possibly allowing the network to improve the predicted binary mask. Having a MUMU-like framework with a wider selection of frequencies, however, will certainly be limited by the available memory. Such differences suggest that the two techniques can be conveniently exploited in a cross-validated framework.

In order to evaluate the numerical stability of the OSM method, further testing was performed corrupting the scattered fields with different levels of additive white Gaussian noise prior to the calculation of the OSM. Direct learning was not included in this experiment due to its previously reported worse performance. From the reported metrics in Table IV, it can be seen that the performances remain almost unchanged, with slight degradation at 0 dB.

VII. ASSESSMENT AGAINST EXPERIMENTAL DATA

As a final validation, U-Net was tested with the experimental data collected from [25], [26]. The aim of this testing was to provide an assessment of samples collected in a completely different way than those used for training.

In particular, there are differences regarding measurement configurations with respect to simulations described in Section V. Also, the Fresnel database includes cases, depicted in Fig. 4, which are not considered in the training, such as single targets, metallic targets, and targets with concave and rectangular shapes. Moreover, while the targets considered in the experiments were lossless, the dielectric materials employed to build the Fresnel targets are not exactly lossless.

Taking into account the better performance of the OSM-based physics-assisted approach, only SISI and MUMU
were employed in this analysis. The results of the analysis are shown in Fig. 5, while the metrics are reported in Table V. Such metrics were calculated using generated ground truths, as depicted in the last column of Fig. 5, based on the description of the Fresnel database. Therefore, depending on the interpolation method used for the ground truth, performance scores may vary.

In case (b), two cylinders, there is a known discrepancy between the nominal and the actual position of the targets (see special issue [26]). Such a discrepancy complicates a proper computation of the ground truth. Accordingly, the quality of the reconstruction for this target was assessed by computing the estimated radii of the targets, which are 13.7 and 15.6 mm, whereas the ground-truth radius for both targets is 15.0 mm.

As can be seen, the developed framework, while not optimized for the considered experimental data, successfully tackles most of the targets, estimating their shape or providing useful morphological information. Particularly, the following holds.

1) The shape of the targets (a) and (b) are properly retrieved, despite the fact that the case of a single cylinder like (a) was not included in the training.

2) For targets in (c) and (d), a sort of convex envelope is retrieved, in agreement with the fact that these targets have shapes (rectangular, concave) that are different from what the network was trained with, and therefore, U-Net tends to round their corners and fill the gaps. This explains the low \(DSC\) and \(MCC\) values. Moreover, these targets are metallic ones. Hence, this is an interesting example that, while being unavoidably driven by the data used in the training, still, the physics-assisted approach shows a certain degree of robustness.

3) Target (e) is made up of two almost concentric cylinders whose external radius is underestimated. The inner one has a permittivity of \(3.0 \pm 0.3\), while the permittivity of the outer one is \(1.45 \pm 0.15\). Such a large jump in permittivity influences the behavior of the OSM indicator [23], making the lower permittivity being confused with the background and partially lost.

4) As for target (f), it is interesting to observe how the network predictions improve the estimate of the shape as compared to the input indicator maps.

5) Targets (g) and (h) are hard to retrieve. The former because nested targets change the OSM indicator.
Fig. 5. Physics-assisted imaging results of eight experimental samples. The first four columns represent the OSM indicator functions at 2, 4, 6, and 8 GHz, followed by the multifrequency OSM indicator. In the sixth and seventh columns, the prediction made by U-Net when SISI (single image) and MUMU (four input images) are used is shown, while the last column depicts the ground-truth binary mask.

function, but its ground-truth binary mask remains the same as for a nonnested profile with similar dimensions. Hence, it is difficult for the network to process these types of targets. The latter is hardly retrieved because the network was not trained to deal with metallic targets. It is interesting to notice that, nevertheless, SISI retrieves a slightly accurate result for the (h) profile.

The assessment against the Fresnel data shows the capabilities of the proposed approach with cases not derivable from the training set and nonsimulated data. Moreover, it also provides an example of the flexibility of the framework with respect to the measurement configuration used to collect the data. As a matter of fact, while the network has been trained with a full-aperture configuration (wherein both the transmitters and the receivers are located on a circumference), the Fresnel measurement configuration is limited in aspect since for each transmitter position the receivers only cover an angular sector [25], [26].

VIII. CONCLUSION

In this work, an innovative microwave imaging framework for effective, reliable, and user-independent retrieval of the shapes of unknown targets from scattered fields was presented. To this end, a processing approach, which combines qualitative imaging and DL, was provided.
As well known, qualitative imaging methods are powerful tools for solving ISPs. However, their outcome is a continuous map that requires postprocessing in order to provide an objective output to clearly identify the targets. Thresholding the continuous map is typically employed to deal with this issue but unavoidably results in case-dependent and user-biased postprocessing.

The above shortcomings were overcome by adopting U-Net as a DL tool and the orthogonality sampling method as the qualitative imaging approach. In particular, the capability of OSM to provide images at multiple frequencies was exploited to provide U-Net with information at different frequencies. Information was supplied to the network in two ways, jointly in a single image (SISI) or as a stack of images (MUMU). In this article, only four frequencies were used, but more frequencies could potentially be employed.

Performance analysis of the processing framework was carried out on simulated data and quantitatively assessed by means of several metrics, showing, on average, excellent capabilities and robustness against noise.

The two versions of the physics-assisted approach were tested on experimental samples, some of which were considerably different from the samples used in the optimization of the network, showing compelling capabilities. Even so, performance should be expected to worsen the more the conditions differ from the training ones, as shown with the experimental samples (c) and (d).

Another limitation, specifically relevant for MUMU-like frameworks, is the need for knowing in advance the number of frequencies to be used and forcibly stick to those ones after the training is carried out. This should not be a burden in specific applications where an optimal frequency selection can be done.

It is worth noting here that the proposed approach can even be used to retrieve targets that are not in free space or embedded in a nonhomogeneous background, provided that the training would be carried out accordingly.

In order to address the listed limitations, further research with different targets, to verify the performance with less controlled experiments, possibly from medical scenarios involving lossy media, will be carried out. Furthermore, the scope of the framework will be extended to 3-D. Also, regarding MUMU implementation, a study investigating the implications of using a larger number of frequencies is needed. Specifically, an analysis of possible performance improvements, memory limitations, or processing speed burden will be performed.

References

[1] L. Crocco, I. Karanasiou, M. L. James, and R. C. Conceição, Emerging Electromagnetic Technologies for Brain Diseases Diagnostics, Monitoring and Therapy, Cham, Switzerland: Springer, 2018.

[2] An Introduction To Microwave Imaging For Breast Cancer Detection, R. C. Conceição ao., M. J. Johan, and M. O’Halloran Eds. Basel, Switzerland: Springer, 2016.

[3] A. S. Turk, K. A. Hocaoglu, and A. A. Vertiy, Subsurface Sensing, vol. 197. Hoboken, NJ, USA: Wiley, 2011.

[4] F. Vipiana, L. Crocco, and J. LoVerdi, “Electromagnetic imaging and sensing for food quality and safety assessment [guest editorial],” IEEE Antennas Propag. Mag., vol. 62, no. 5, pp. 16–17, Oct. 2020.

[5] M. G. Amin, Through-the-Wall Radar Imaging. Boca Raton, FL, USA: CRC Press, 2017.

[6] M. Pastorino, Microwave Imaging, vol. 208. Hoboken, NJ, USA: Wiley, 2010.

[7] F. Cakoni and D. L. Colton, A Qualitative Approach to Inverse Scattering Theory, vol. 767. Cham, Switzerland: Springer, 2014.

[8] N. Senthilkumaran and S. Vaiithi, “Image segmentation by using thresholding techniques for medical images,” Comput. Sci. Eng., Int. J., vol. 6, no. 1, pp. 1–13, Feb. 2016.

[9] X. Wei, Z. Wei, M. Li, and P. Rocca, “A review of deep learning approaches for inverse scattering problems (invited review),” Prog. Electromagn. Res., vol. 167, pp. 67–81, 2020.

[10] Y. Khoo and L. Ying, “SwitchNet: A neural network model for forward and inverse scattering problems,” SIAM J. Scientific Comput., vol. 41, no. 5, pp. A3182–A3201, Jan. 2019.

[11] R. Guo, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, “Supervised descent learning technique for 2-D microwave imaging,” IEEE Trans. Antennas Propag., vol. 67, no. 5, pp. 3550–3554, May 2019.

[12] G. Chen, P. Shah, J. Stang, and M. Moghaddam, “Learning-assisted multimodality dielectric imaging,” IEEE Trans. Antennas Propag., vol. 68, no. 3, pp. 2356–2369, Mar. 2020.

[13] K. Edwards, V. Khoshdel, M. Asefi, J. LoVerdi, C. Gilmore, and J. Jeffrey, “A machine learning workflow for tumour detection in breasts using 3D microwave imaging,” Electronics, vol. 10, no. 6, p. 674, Mar. 2021.

[14] Z. Wei and X. Chen, “Physics-inspired convolutional neural network for solving full-wave inverse scattering problems,” IEEE Trans. Antennas Propag., vol. 67, no. 9, pp. 6138–6148, Sep. 2019.

[15] X. Sanghvi, Y. Kalepu, and U. Khankhoje, “Embedding deep learning in inverse scattering problems,” IEEE Trans. Comput. Imag., vol. 6, pp. 46–56, 2020.

[16] 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. 2018.

[17] Z. Wei and X. Chen, “Uncertainty quantification in inverse scattering problems with Bayesian convolutional neural networks,” IEEE Trans. Antennas Propag., vol. 69, no. 6, pp. 3409–3418, Jun. 2021.

[18] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org.

[19] S. Xu, H. Wu, and R. Bie, “CXNet-m1: Anomaly detection on chest X-rays with image-based deep learning,” IEEE Access, vol. 7, pp. 4466–4477, 2018.

[20] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, and J. Wang, “Applications of deep learning to MRI images: A survey,” Big Data Mining Anal., vol. 1, no. 1, pp. 1–18, Mar. 2018.

[21] B. Gerazov and R. C. Conceicao, “Deep learning for tumour classification in homogeneous breast tissue in medical microwave imaging,” in Proc. IEEE EUROCON 17th Int. Conf. Smart Technol., Jul. 2017, pp. 564–569.

[22] R. Potthast, “A study on orthogonality sampling,” Inverse Problems, vol. 26, no. 7, Jul. 2010, Art. no. 074015.

[23] M. T. Bevacqua, T. Isemnia, R. Palmeri, M. N. Akinci, and L. Crocco, “Physical insight unveil new imaging capabilities of orthogonality sampling method,” IEEE Trans. Antennas Propag., vol. 68, no. 5, pp. 4014–4021, May 2020.

[24] A. Yago, M. Cavagnaro, and L. Crocco, “Deep learning-enhanced qualitative microwave imaging: Rationale and initial assessment,” in Proc. 15th Eur. Conf. Antennas Propag. (EuCAP), Mar. 2021, pp. 1–5.

[25] J.-M. Geffrin, P. Sabouroux, and C. Eyraud, “Free space experimental scattering database continuation: Experimental set-up and measurement precision,” Inverse Problems, vol. 21, no. 6, pp. S117–S130, Dec. 2005.

[26] K. Belkebir and M. Saillard, “Special section: Testing inversion algorithms against experimental data,” Inverse Problems, vol. 17, no. 6, pp. 1565–1571, 2001, doi: 10.1088/0266-5611/17/2F6/301.

[27] M. Benning and M. Burger, “Modern regularization methods for inverse problems,” Acta Numerica, vol. 27, pp. 1–111, May 2018.

[28] I. Harris and D.-L. Nguyen, “Orthogonality sampling method for the electromagnetic inverse scattering problem,” SIAM J. Sci. Comput., vol. 42, no. 3, pp. B722–B737, Jan. 2020.

[29] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2015, pp. 3431–3440.

[30] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015, pp. 234–241.
Álvaro Yago Ruiz (Graduate Student Member, IEEE) received the B.Sc. degree in biomedical engineering from the Universitat Politècnica de València, Valencia, Spain, in 2016, and the M.Sc. degree in electrical engineering from Aarhus University in 2019. He is currently pursuing the Ph.D. degree with Sapienza, Università di Rome, Rome, Italy.

He is currently involved in a European Project (EMERALD) and carries out his research at the Institute for the Electromagnetic Sensing of the Environment, National Research Council of Italy (IREA-CNR), Naples, Italy. For his master’s thesis, he studied the application of AI-based segmentation algorithms to CT images. Over the years, he has developed an interest in machine learning, optimization, and medical imaging.

Marta Cavagnaro (Member, IEEE) received the master’s degree (cum laude) in electronic engineering and the Ph.D. degree in electronic engineering from the Sapienza University of Rome, Rome, Italy, in 1993 and 1997, respectively.

She is currently an Associate Professor with the Department of Information Engineering, Electronics and Telecommunications. In 2017, she received the Full Professor qualification in electromagnetic fields from the Italian Ministry of Research and University. Her research interests are related to dosimetric aspects of the interaction between electromagnetic fields and biological systems, such as medical applications, the environmental impact of mobile communication systems, and measurements of dielectric properties of tissues.

Dr. Cavagnaro is an Associate Editor of the Bioelectromagnetics Journal and serves as Anonymous Reviewer for several scientific journals. She is an Italian Representative in the Management Committees of COST action devoted to the standardization of thermal therapies techniques MyWAVE (European network for advancing Electromagnetic hyperthermic medical technologies) and a member of the Thermal Medicine Standard Committee of ASME (The American Society of Mechanical Engineers). She is the author or a coauthor of more than 200 scientific articles.

Lorenzo Crocco (Senior Member, IEEE) is a Research Director with the Institute for the Electromagnetic Sensing of the Environment, National Research Council of Italy (IREA-CNR). His scientific activities mainly concern electromagnetic scattering with a focus on diagnostic and therapeutic uses of EM fields, through-the-wall radar, and GPR. On these topics, he has published more than 120 articles, given keynote talks and lectures, and led or participated in national and international research projects.

Dr. Crocco has been a member of the board for the European School of Antennas (ESOA), since 2015. Since 2017, he has been a member of the Board of Directors of the Italian Electromagnetic Society (SIEm). In 2018, he received the Full Professor Habilitation in electromagnetic fields from the Italian Ministry of Research and University. Since 2019, he has been a member of the Italian URSI Commission (International Union of Radio Science). In 2019, he has been elected to the Scientific Board of the Engineering Department (DIITET) of CNR. He was the recipient of the SIEm Barzilai Award for Young Scientists in 2004 and YSA at the URSI General Assembly held in New Delhi (India) in 2005. He is a fellow of the Electromagnetic Academy (TEA) and a URSI Senior Member. He is an Associate Editor for the IEEE JOURNAL OF ELECTROMAGNETICS, RF AND MICROWAVES IN MEDICINE AND BIOLOGY (IEEE J-ERM) and has edited a book on Electromagnetic Technologies for Brain Diseases Diagnostics, Monitoring and Therapy. Since 2013, he has been an Italian representative in the Management Committees of COST actions devoted to medical applications of EM fields (MiMed on microwave imaging and MyWAVE on therapeutic applications of electromagnetic waves).