

Graph-CNNs for RF Imaging: Learning the Electric Field Integral Equations

Kyriakos Stylianopoulos¹, Panagiotis Gavrilidis¹, Gabriele Gradoni², and George C. Alexandropoulos¹

¹Department of Informatics and Telecommunications, National and Kapodistrian University of Athens,

Panepistimiopolis Ilissia, 16122 Athens, Greece

²Institute for Communication Systems, University of Surrey, Guildford GU27XH, UK

e-mails: {kstylianop, pangavr, alexandg}@di.uoa.gr, g.gradoni@surrey.ac.uk

Abstract—Radio-Frequency (RF) imaging deals with the digital recreation of the surfaces of scene objects based on the scattered field at distributed receivers. To solve this difficult inverse scattering problem, data-driven methods are often employed that extract patterns from similar training examples, while offering minimal latency. In this paper, we first provide an approximate yet fast electromagnetic model, which is based on the electric field integral equations, for data generation, and subsequently propose a Deep Neural Network (DNN) architecture to learn the corresponding inverse model. A graph-attention backbone allows for the system geometry to be passed to the DNN, where residual convolutional layers extract features about the objects, while a UNet head performs the final image reconstruction. Our quantitative and qualitative evaluations on two synthetic data sets of different characteristics showcase the performance gains of the proposed advanced architecture and its relative resilience to signal noise levels and various reception configurations.

Index Terms—Radio-frequency imaging, inverse scattering problems, deep learning for electromagnetics, graph neural networks, electric field integral equations.

I. INTRODUCTION

Radio-Frequency (RF) imaging is widely used in many areas, such as in medical diagnostics [1], security applications [2], and diverse engineering fields, including structural tomography [3] and ground-penetrating radar [4]. Multi-static imaging setups, where multiple Transmitters (TXs) emit RF waves that interact with surrounding objects before being recorded by distributed Receivers (RXs), enable imaging through digital signal processing or techniques realized directly in the Electromagnetics (EM) domain. However, typical systems require specialized hardware, high computational resources, and operate in high power and bandwidth regimes.

Next-generation wireless networks are expected to offer imaging opportunities through communications infrastructure, leveraging base station densification, device-to-device networking, and millimeter-wave technology. Applications such as integrated sensing and communications [5], digital twinning [6], and autonomous guided vehicles benefit from high-resolution channel propagation characterization [7] and optimized beamforming [8]. Integrating, however, imaging functionalities into wireless networks imposes stringent require-

ments on latency, computation, and communication resources, challenging existing methodologies.

In RF imaging, Synthetic Aperture Radar (SAR) techniques, such as the range migration algorithm [2], [9], [10], reconstruct scenes by coherently combining multiple received signals across frequency or spatial domains. While enabling high-resolution imaging, SAR requires large bandwidth, struggles in near-field scenarios, and relies on simplified reflection models that overlook diffraction and multi-bounce effects. Moreover, SAR primarily captures outer surfaces, making it ineffective for occluded objects within a Domain of Interest (DoI). Alternatively, RF imaging can directly leverage EM theory, capitalizing on the fact that received signals satisfy Maxwell's frequency-domain equation [11], inherently encapsulating the DoI's EM properties. This approach circumvents SAR's limitations, but presents its own challenges: solving nonlinear Partial Differential Equations (PDEs) is computationally expensive, prone to degenerate solutions due to ill-posedness, and depends on numerical models lacking accuracy.

Machine Learning (ML) methodologies provide an attractive alternative in overcoming such shortcomings. The plethora of real life data sets of medical data has been instrumental to the successful applications of Deep Neural Networks (DNNs) in medical imaging (see [12] and references therein) through the exploitation of diverse computer vision schemes, such as image denoising, super resolution, and unsupervised learning. Additionally, ample synthetic data sets may be generated in simulations using EM models to train ML approaches for general-purpose inverse scattering problems [11], providing a stream of training examples with de facto labeling. As a prominent example, the work in [13] deployed the UNet architecture [14] to capitalize on the DNN's ability to output images comprised of distinct regions. More recently, there has been an ongoing interest in combining ML with EM models to the purpose of designing Physics-Informed Neural Networks (PINNs), which enhance the fidelity of image reconstruction. This may be achieved through pre- or post-processing of the input and output variables of DNN with physics-compliant models, by incorporating loss functions that conserve EM properties (e.g., [15]), or by approximating parts of EM models with DNN layers; refer to [11] for an in-depth analysis of PINNs for RF imaging. However, despite the promising results of PINNs in performance gains and data efficiency, there

This work has been supported by the SNS JU project 6G-DISAC under the EU's Horizon Europe research and innovation program under Grant Agreement No 101139130 and the 2023 Greek-British Cooperation on Short-Term Scholarships Programme.

remain yet unresolved challenges. First, the physics-compliant loss functions, or operations, are typically computationally expensive to evaluate, especially if such computations involve evaluating or solving PDEs. Additionally, while one is free to choose from a wide variety of approximate EM models and their corresponding properties to infuse into PINNs, it is difficult to determine a priori which of those will be beneficial for training and what DNN layers are effective in such regimes.

In this paper, we argue that developing efficient DNN models with pertinent training methodologies, that better exploit all the available system information, is of paramount importance, either as standalone physics-agnostic imaging algorithms, or as components of PINNs. The specific contributions of this work are as follows: An approximate yet computationally fast EM model is first presented, which is based on the Electric Field Integral Equations (EFIEs) to produce the received signals scattered by objects within the DoI. The model is then used to generate two synthetic RF imaging data sets addressing different challenges. Next, a novel DNN architecture for RF imaging is proposed, which is based on graph neural networks that capture geometrical information about the system's transceivers and provide highly expressive feature extraction, followed by ResNet [16] and UNet structures for effective image generation. This architecture is evaluated on the generated data sets under varying Signal-to-Noise Ratio (SNR) conditions and numbers of RXs to showcase its performance gains over simpler conventional DNN approaches.

II. SYSTEM AND PROBLEM FORMULATION

A. EM Modeling

We consider a 2D disk DoI of radius $\rho = 5\lambda$, with λ being the wavelength of the propagating electric field, which is centered at the origin. N_r RX units are placed in equally spaced positions around the periphery of the DoI, which are assumed to be perfectly synchronized. Contrary to most RF imaging techniques [11], [13], [15], [17], we assume a single TX located at a distance of 20λ from the origin, at $\mathbf{p}_t = [20\lambda, 0]^\top$. Despite making the problem more challenging, using a single TX greatly simplifies the involved synchronization processes and reduces the measurement latency, offering a more attractive scenario for wireless communication systems.

Inside the DoI, there exist one or more disjoint objects of perfect electrical conductors, that are represented as a discrete set of infinitesimal elements of coordinates $\bar{\mathbf{s}} = [x_{\bar{\mathbf{s}}}, y_{\bar{\mathbf{s}}}]^\top$, sampled with spatial resolution ρ_0 inside the boundaries of each object. Let $\mathcal{S}$ be the set of all potential points in the DoI, and $\bar{\mathcal{S}}$ be the set of all coordinates comprising the objects, which will be henceforth referred to as a surface, despite containing possibly disjoint objects. We assume the transmitting field to travel in the (x, y) -plane, being polarized in the z -direction, as $E_t(\mathbf{p}) = E_0 \exp(-j\mathbf{k}^\top(\mathbf{p} - \mathbf{p}_t)) / (4\pi\|\mathbf{p} - \mathbf{p}_t\|)$ with $\mathbf{p} = [x_p, y_p]^\top$ being an arbitrary point in the DoI, $\mathbf{k} \in \mathbb{R}^{2 \times 1}$ denoting the free space wave vector, and E_0 being a constant amplitude term. Following first physics principles and considering time-harmonic electric field (i.e., $\mathcal{E}(t, \mathbf{p}) = E_0 \cos(\omega t - \mathbf{k}^\top(\mathbf{p} - \mathbf{p}_t))$), we formulate the relationship that

relates the electric field at any point in space with the current density $J(\mathbf{s})$ of the DoI via the EFIE, as follows:

$$E_r(\mathbf{p}) = \iint_{\mathcal{S}} G(\mathbf{p}, \mathbf{s}) J(\mathbf{s}) d\mathbf{s} + E_t(\mathbf{p}), \quad (1)$$

where $G(\mathbf{p}, \mathbf{s})$ is the Green's function. For the TM_z propagating wave considered here, this function is defined via the Hankel function of the 0-th order of the 2nd kind, $H_0^{(2)}$, as $G(\mathbf{p}, \mathbf{s}) = -\frac{j}{4} H_0^{(2)}(k_0 \|\mathbf{p} - \mathbf{s}\|)$ [18, eq. (11-11a)], where k_0 denotes the free-space wavenumber. Computing the electric field reduces to determining the current density, which can be done by adopting N (known) orthogonal basis functions $\{\Phi_n(\mathbf{s})\}_{n=1}^N$ and accompanying intensities $\{\alpha_n\}_{n=1}^N$. The current density is then expressed as $J(\mathbf{s}) = \sum_{n=1}^N \alpha_n \Phi_n(\mathbf{s})$.

Deriving a complete EM model involves finding appropriate values of $\{\alpha_n\}_{n=1}^N$ for a given $\bar{\mathcal{S}}$, which can be obtained via the point-matching method [18, Section 12.2.4]. It is first noted that (1) can be applied to any point in 2D space, including the RXs, empty space, as well as any $\bar{\mathbf{s}} \in \bar{\mathcal{S}}$ in particular. We thus exploit the fact that the set $\bar{\mathcal{S}}$ contains perfect electric conductors, which dictates the following boundary condition onto the surface points:

$$E_r(\bar{\mathbf{s}}) = 0, \quad \forall \bar{\mathbf{s}} \in \bar{\mathcal{S}}. \quad (2)$$

By assuming known $\Phi_n(\mathbf{s})$ and by evaluating (1) on N different $\bar{\mathbf{s}}_i$, $\forall i = 1, \dots, N$ points, we obtain a linear system of N equations with N unknowns $\mathbf{a} \triangleq [\alpha_1, \dots, \alpha_N]^\top$ through (2). Specifically, let $\mathbf{e}_t \triangleq [-E_t(\bar{\mathbf{s}}_1), \dots, -E_t(\bar{\mathbf{s}}_N)]^\top$ and the $N \times N$ matrix $\mathbf{M}$ as $[\mathbf{M}]_{i,n} \triangleq \iint_{\mathcal{S}} G(\bar{\mathbf{s}}_i, \mathbf{s}) \Phi_n(\mathbf{s}) d\mathbf{s}$ $\forall i, n = 1, \dots, N$. Assuming $\mathbf{M}^{-1}$ is defined, the system of equations deduced from (2) can be solved as $\mathbf{a} = \mathbf{M}^{-1} \mathbf{e}_t$, thus, completing the model. For the data-generation process we have selected the following procedure for the creation of Φ_n : An $\sqrt{N} \times \sqrt{N}$ grid of equally spaced points $\hat{\mathbf{s}}_n$ is constructed. Each basis Φ_n is defined as a 2D square pulse of unit amplitude centered around $\hat{\mathbf{s}}_n$ with a width of $2\rho/\sqrt{N}$.

B. RF Imaging Problem Formulation

For each instance t of the problem, the DoI contains a distinct surface $\bar{\mathcal{S}}^{(t)}$. Let the (fixed) positions of the RXs be denoted as $\mathbf{r}_k = [x_{r_k}, y_{r_k}]^\top \quad \forall k = 1, \dots, N_r$, and the corresponding vector of received signals as $\mathbf{e}^{(t)} \triangleq [E_r(\mathbf{r}_1), \dots, E_r(\mathbf{r}_{N_r})]^\top$. The model described in Section II-A is referred to as the *forward* model, as it provides the mapping $F : \bar{\mathcal{S}}^{(t)} \rightarrow \mathbf{e}^{(t)}$. As a general formulation, the imaging objective can be described as the optimization problem $\mathcal{S}_{\text{opt}} \triangleq \arg \max_{\bar{\mathcal{S}}} \|\mathbf{e}^{(t)} - F(\bar{\mathcal{S}})\|$, which seeks the surface that, when processed through the forward model, best matches the received electric fields. However, the above is analytically intractable in large DoI's resolutions, as it involves searching over the space of all possible surfaces.

To this end, a data-driven objective is employed instead. For notation purposes, let us first define the operator $I(\bar{\mathcal{S}})$ that transforms the surface $\bar{\mathcal{S}}$ to a discretized binary image of the DoI of arbitrary spatial resolution, where each pixel has

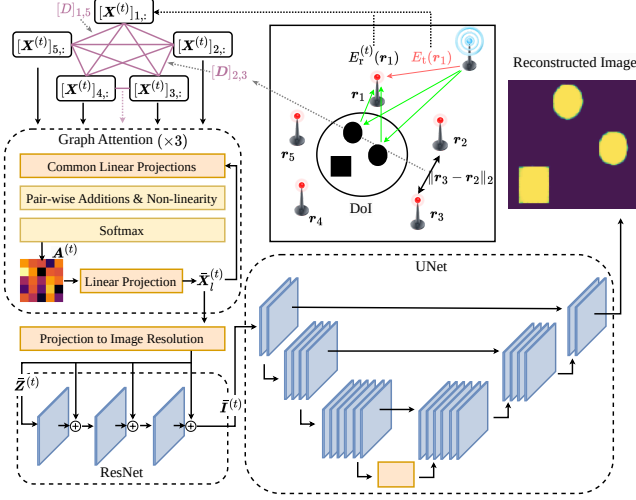


Fig. 1. GAT-Res-UNet architecture for RF imaging via inverse scattering.

the value of 1 if and only if there is at least one point in $\bar{S}$ that is spatially covered by the pixel. Assume that $\bar{S}^{(t)}$'s elements are conditionally dependent and are thus sampled by a non-uniform distribution, which motivates the need for ML applications. The learning objective can be defined as approximating the *inverse* model $F^{-1} : e^{(t)} \rightarrow \bar{S}^{(t)}$ as:

$$F_{\text{opt}}^{-1} \triangleq \arg \min_{F^{-1}} \frac{1}{|\mathcal{D}|} \sum_{t=1}^{|\mathcal{D}|} \text{CE}(I(\bar{S}^{(t)}), I(F^{-1}(e^{(t)}))), \quad (3)$$

where the Cross-Entropy (CE) function $\text{CE}(x, y) = -\sum_j [x]_j \log[y]_j$ is the objective function ($I(\cdot)$ is treated as a vector) and $\mathcal{D}$ is a collected set of input-target data tuples $\{e^{(t)}, I(\bar{S}^{(t)})\}$. For ML approaches, $F^{-1}(\cdot)$ is a parameterized neural network, hence, problem (3) can be solved through Stochastic Gradient Descent (SGD), by optimizing the DNN's trainable weight matrices over batches of set $\mathcal{D}$.

III. PROPOSED DEEP NEURAL NETWORK ARCHITECTURE

The proposed DNN architecture consists of three parts. First, a Graph-Attention Network (GAT) [19], [20] is used for feature extraction. The imposed graph structure allows for the incorporation of systemic and geometric information (that remains fixed throughout the data set) as node and edge features, while the attention mechanism is tasked to learn correlations between the electric field of each data instance inputs and those system variables. In particular, let the input data matrix for the k -th data instance $\mathbf{X}^{(t)} \in \mathbb{R}^{N_r \times 4}$ contain the received field at the RXs ($e^{(t)}$) and the (fixed) electric field contribution solely by the TX as $[\mathbf{X}^{(t)}]_{i,:} \triangleq [\Re(E_r^{(t)}(\mathbf{r}_i)) \ \Im(E_r^{(t)}(\mathbf{r}_i)) \ \Re(E_t(\mathbf{r}_i)) \ \Im(E_t(\mathbf{r}_i))]^\top$. Define the edge weights between the i -th and j -th RXs via their Euclidean distance as $[D]_{i,j} \triangleq \|\mathbf{r}_i - \mathbf{r}_j\|_2$. The attention coefficients of the first GAT layer $[\mathbf{A}^{(t)}]_{i,j}$ are computed as:

$$[\mathbf{A}^{(t)}]_{i,j} \triangleq \text{softmax} \left(\text{LeakyReLU}(\mathbf{w}_s^\top \Theta_s [\mathbf{X}^{(t)}]_{i,:}^\top + \mathbf{w}_t^\top \Theta_t [\mathbf{X}^{(t)}]_{j,:}^\top + \mathbf{w}_e^\top \Theta_e [D]_{i,j}) \right), \quad (4)$$

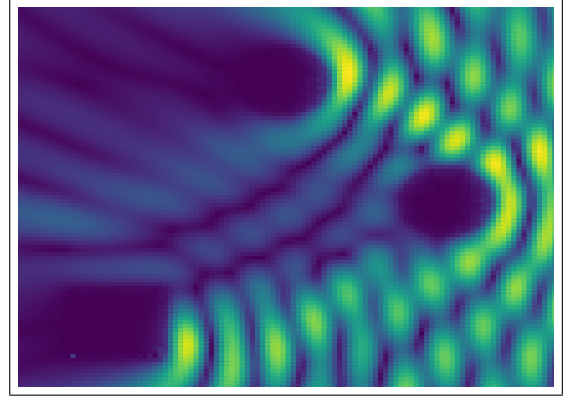


Fig. 2. Visualization of the magnitude of the electric field $|E_t(\mathbf{s})|$ in arbitrary units over the DoI in the presence of three perfect electrical conductors of side and diameter 1.2λ , demonstrating the complex interactions captured by the employed EFIEs. Non-overlapping pulse functions with $\lambda/20$ spacing in between them were used as bases.

where $\mathbf{w}_s, \mathbf{w}_t, \mathbf{w}_e \in \mathbb{R}^{d \times 1}$ and $\Theta_s, \Theta_t, \Theta_e \in \mathbb{R}^{d \times 4}$ are trainable weights with an arbitrary embedding dimension d . The output of the GAT layer is given as $[\bar{\mathbf{X}}^{(t)}]_{i,:} = \sum_{j=1}^{N_r} [\mathbf{A}]_{i,j} \Theta_t [\mathbf{X}^{(t)}]_{j,:}^\top$. For the considered data sets, $L_1 = 3$ GAT layers with $d = 32$ are applied in sequence so that their respective inputs are $\mathbf{X}_l^{(t)} = \bar{\mathbf{X}}_{l-1}^{(t)} \ \forall l = 2, \dots, L_1$.

Next, $\mathbf{X}_{L_1}^{(t)}$ is flattened to $\mathbf{z}^{(t)} \in \mathbb{R}^{dN_r \times 1}$ and linearly projected to the $\mathbb{R}^{HW \times 1}$ space as $\bar{\mathbf{z}}^{(t)} \triangleq \sigma(\mathbf{W}_z \mathbf{z}^{(t)} + \mathbf{b}_z)$, with $H \times W$ being the number of pixels of the desired image, $\mathbf{W}_z$ and $\mathbf{b}_z$ being trainable parameters of conformable dimensions, and $\sigma(\cdot)$ denotes the sigmoid activation. Then, $\bar{\mathbf{z}}^{(t)}$ is seen as an image $\bar{\mathbf{Z}}^{(t)} \in \mathbb{R}^{1 \times H \times W}$ and the two computer vision modules are applied.

Proceeding, $L_2 = 5$ typical ResNet blocks of a single channel are applied on $\bar{\mathbf{Z}}$ with the intention of capturing information about the location, size, and shape of the objects, resulting in an intermediate image $\bar{\mathbf{I}}^{(t)} \in \mathbb{R}^{1 \times H \times W}$. This information is assumed to exhibit high spatial locality due to the spatially aware graph introduced by the GAT layers, therefore, the convolutional layers of ResNet provide an intuitive choice. Finally, $\bar{\mathbf{I}}^{(t)}$ is processed by a UNet of $L_3 = 3$ downsampling-upsampling layers with 16, 32, and 64 channels, respectively. The purpose of the UNet is to efficiently encode information about the shapes of the objects in the image during downsampling to a bottleneck vector, by increasing the channel and decreasing the image dimensions. The converse procedure of the upsampling layers is designed to decode this information toward assigning each pixel of the image to its corresponding object (or background). Recall that this work deals with perfect conducting objects, indicating that the output image is binary. This implies that it suffices to apply a sigmoid activation to obtain the final image estimation $\hat{\mathbf{I}}^{(t)} \in \mathbb{R}^{H \times W} \equiv I(\bar{S}^{(t)})$, before performing SGD on (3). The overall DNN architecture is illustrated in Fig. 1.

IV. EVALUATION

Using the forward model of Section II-A and $\lambda = 0.125$ m, we have produced two data sets of 10^4 data points, each

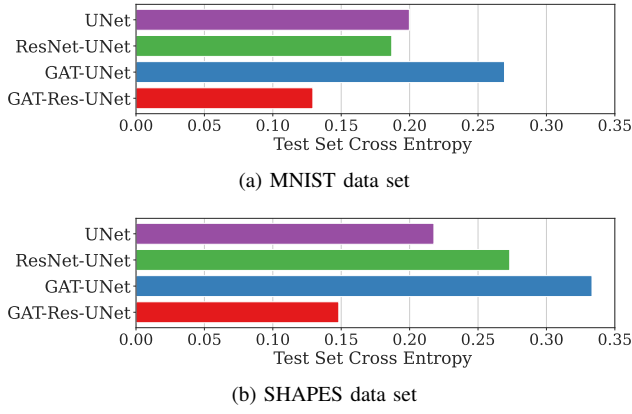


Fig. 3. Comparison of validation scores for the proposed GAT-Res-UNet and baseline methods on MNIST and SHAPES data sets (lower is better).

TABLE I
PERFORMANCE COMPARISON IN TERMS OF CE UNDER DIFFERENT SNR LEVELS AND N_r VALUES.

Data set	MNIST				SHAPES			
	0	10	20	30	0	10	20	30
UNet	<u>0.25</u>	0.20	0.18	0.19	<u>0.34</u>	<u>0.29</u>	0.26	0.25
ResNet-UNet	<u>0.25</u>	0.20	0.20	0.20	0.35	0.36	0.37	0.32
GAT-Res-UNet	<u>0.25</u>	<u>0.18</u>	<u>0.17</u>	<u>0.16</u>	0.35	0.30	<u>0.21</u>	<u>0.19</u>
N_r	8	16	32	48	8	16	32	48
UNet	0.27	0.24	0.21	0.19	0.36	<u>0.34</u>	0.30	0.31
ResNet-UNet	0.26	0.25	0.26	0.20	0.37	0.36	0.32	0.28
GAT-Res-UNet	<u>0.25</u>	<u>0.21</u>	<u>0.17</u>	<u>0.16</u>	<u>0.35</u>	0.36	<u>0.28</u>	<u>0.22</u>

with 0.2 test split, that were designed to test different system characteristics. First, we sampled MNIST digits [21] to investigate the ability of the designed DNNs to identify complex shapes. Each 28×28 image was conceptually centered on the DoI origin so that its total height and width are 3λ each, and each pixel occupies an area of $0.1\lambda \times 0.1\lambda$. Using the spatial resolution $\rho_0 = \lambda/20$, we sampled points within the area occupied by the pixels of the digits, to obtain each $\bar{\mathcal{S}}^{(t)}$ surface, which contains 282 discrete points on average. $\bar{\mathcal{S}}^{(t)}$ was then used in (1) to obtain $e^{(t)}$ while setting $N = |\bar{\mathcal{S}}^{(t)}|$. Next, we created a synthetic data set, referred to as “SHAPES,” which was designed to assess the detection of multiple disjoint objects with larger image resolution. To that end, a square and two discs of width and diameters uniformly sampled in $[\lambda, 1.2\lambda]$ were randomly positioned within the DoI disk to create each $\bar{\mathcal{S}}^{(t)}$. The discretization results in 651 points on average and the target images had dimensions of 128×128 pixels. Figure 2 provides a visual inspection of the our EFIE method in the generation of the data set.

We have compared the proposed “GAT-Res-UNet” architecture against different architectures of the same three components: UNet, ResNet-UNet, and GAT-UNet. The UNet baseline closely follows the benchmark approach of [13], while the rest of the approaches constitute modifications of it. Other combinations of the same layers did not exhibit notable performance. For the approaches that do not contain GAT as the first layer, only $e^{(t)}$ was used as input, without the system

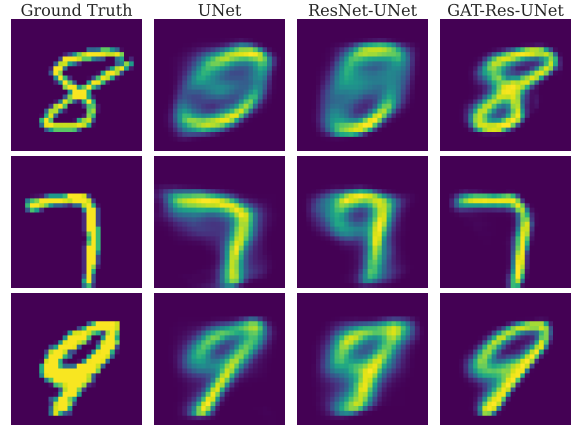


Fig. 4. Visual comparison of the reconstruction of three MNIST images.

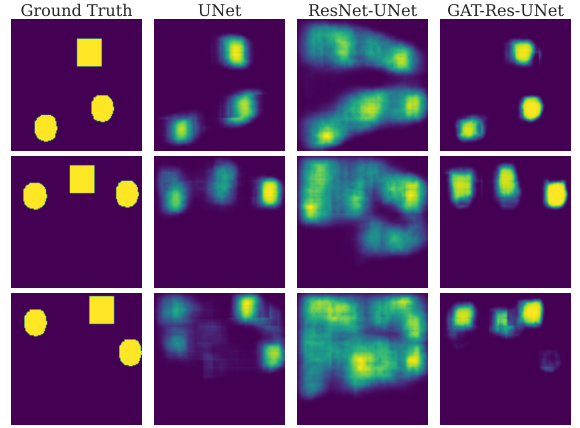
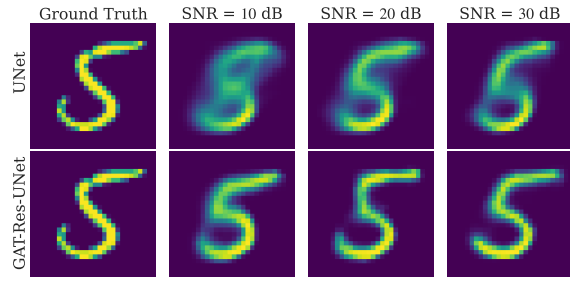


Fig. 5. Visual comparison of the reconstruction of three SHAPES images.

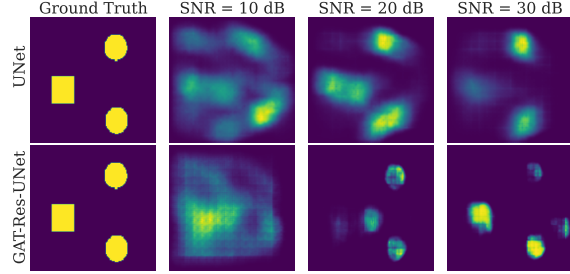
information of $E_t(\mathbf{r}_i)$ and $\mathbf{D}$. A linear layer was introduced in these cases instead of GAT to convert the input vectors to the image dimensions similar to the $\mathbf{z}^{(t)} \rightarrow \bar{\mathbf{z}}^{(t)}$ projection of Section III, so that the convolutional modules can be applied.

For $N_r = 64$ RXs and noise-free received signals, the CE error scores for the test sets of all compared methods are given in Figs. 3a and 3b. Clearly, the proposed GAT-Res-UNet approach shows lower reconstruction errors, with the performance difference with respect to the baseline approaches being increased for SHAPES images of higher resolution. To investigate scenarios where the received signals exhibiting Additive White Gaussian Noise (AWGN), as well as in setups with varying number of RXs, the models were re-trained for each particular SNR and N_r case, with the results being shown in Table I. As observed, the proposed method outperforms all benchmarks in almost all cases. It is shown that their differences become less pronounced when their resources deteriorate (fewer RXs and lower SNR), where no method produces effective results. The SHAPES data set presents the more challenging case, due to the higher resolution.

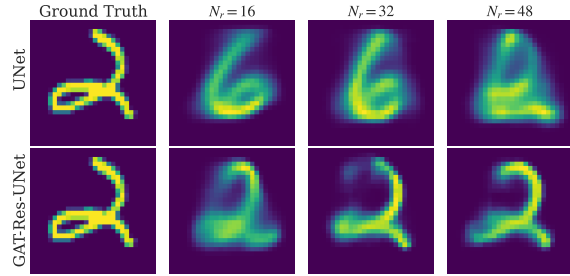
From a qualitative perspective, reconstructed images from the test sets of the two data sets are depicted in Figs. 4 and 5, where it can be seen that GAT-Res-UNet produces sharper images, especially for the case of MNIST. For the challenging case of SHAPES images, GAT-Res-UNet is mostly able to



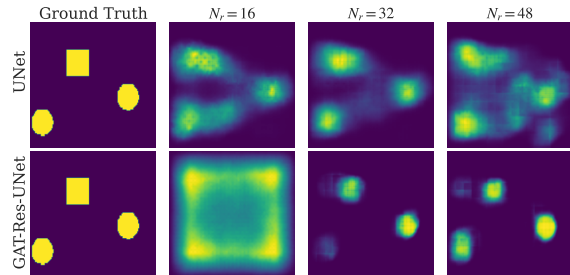
(a) MNIST image under different SNR conditions.



(b) SHAPES image under different SNR conditions.



(c) MNIST image under different numbers of RXs.



(d) SHAPES image under different numbers of RXs.

Fig. 6. Reconstruction of sample images under various system parameters.

identify the positions and sizes of the shapes with moderate accuracy, but it does not succeed in producing precise shapes. Comparing with UNet for different resource budgets in Fig 6, a clear advantage of the proposed method can be inferred in all tests apart from low resources cases, where all methods are unsuccessful. While there is notable performance degradation as the system capabilities become limited, the proposed method is demonstrated to exhibit higher robustness.

V. CONCLUSION

In this paper, we designed a DNN architecture based on graph attention, residual convolutions, and UNet structures for effective RF imaging with reduced system requirements. An

EM modeling approach based on the EFIEs has been used for generation of two data sets of distinct features, and our numerical evaluation showcased notable performance gains of the proposed method compared to modified approaches from the literature, as well as relative resilience in decreased SNR levels and number of RXs. Such results further motivate us to investigate the incorporation of the presented EM model to ML methods in making DNN designs physics-informed, toward enhancing their performance and generalization capabilities.

REFERENCES

- [1] D. O'Loughlin, M. O'Halloran, B. M. Moloney, M. Glavin, E. Jones, and M. A. Elahi, "Microwave breast imaging: Clinical advances and remaining challenges," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2580–2590, 2018.
- [2] D. Sheen, D. McMakin, and T. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 9, pp. 1581–1592, 2001.
- [3] C. Maierhofer et al., *Non-Destructive Evaluation of Reinforced Concrete Structures*. Woodhead Publishing, 2010.
- [4] H. M. Jol, *Ground Penetrating Radar Theory and Applications*. Amsterdam, Netherlands: Elsevier, 2009.
- [5] E. Calvanese Strinati et al., "Towards distributed and intelligent integrated sensing and communications for 6G networks," *IEEE Wireless Commun.*, vol. 32, no. 1, pp. 60–67, 2025.
- [6] A. Masaracchia et al., "Towards the metaverse realization in 6G: Orchestration of RIS-enabled smart wireless environments via digital twins," *IEEE Internet Things Mag.*, vol. 7, no. 2, pp. 22–28, 2024.
- [7] L. Nisiotis et al., "Exploring gaming technologies, digital twins, and VR to visualise wireless propagation simulations," in *Proc. IEEE Annu. Comput., Software, Appl. Conf.*, Osaka, Japan, Jul. 2024.
- [8] K. Haneda et al., "5G 3GPP-Like channel models for outdoor urban microcellular and macrocellular environments," in *Proc. IEEE Veh. Technol. Conf.*, Nanjing, China, May 2016.
- [9] G. Wang et al., "Comparison between back projection algorithm and range migration algorithm in terahertz imaging," *IEEE Access*, vol. 8, pp. 18 772–18 777, 2020.
- [10] Y. Yang et al., "Automotive SAR imaging by range migration algorithm," in *Proc. Int. Conf. Microw. Millimeter Wave Technol.*, Harbin, China, Aug. 2022.
- [11] R. Guo et al., "Physics-embedded machine learning for electromagnetic data imaging: Examining three types of data-driven imaging methods," *IEEE Signal Process. Mag.*, vol. 40, no. 2, pp. 18–31, 2023.
- [12] G. Wang et al., "Deep tomographic image reconstruction: Yesterday, today, and tomorrow—editorial for the 2nd special issue "Machine Learning for Image Reconstruction"," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 2956–2964, 2021.
- [13] 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, 2019.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput. Assist. Interv.*, Munich, Germany, 2015.
- [15] H. M. Yao et al., "Two-step enhanced deep learning approach for electromagnetic inverse scattering problems," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2254–2258, 2019.
- [16] K. He et al., "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016.
- [17] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Prog. In Electromagn. Res.*, vol. 167, pp. 67–81, 2020.
- [18] C. A. Balanis, *Advanced Engineering Electromagnetics 2nd Edition*. Hoboken, NJ, USA: John Wiley & Sons, Ltd, 2011.
- [19] P. Veličković et al., "Graph attention networks," in *Proc. Int. Conf. Learn. Representations*, Vancouver, Canada, May 2018.
- [20] M. Fey and J. E. Lenssen, "GATConv — PyTorch Geometric 2.4.0 documentation," Online, 2024, available: <https://pytorch-geometric.readthedocs.io/en/latest/modules/gat.html>.
- [21] L. Deng, "The MNIST database of handwritten digit images for machine learning research," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 141–142, 2012.