



OPEN

# Deep learning for tumor margin identification in electromagnetic imaging

Amir Mirbeik<sup>1,2</sup> & Negar Ebadi<sup>2,3</sup>✉

In this work, a novel method for tumor margin identification in electromagnetic imaging is proposed to optimize the tumor removal surgery. This capability will enable the visualization of the border of the cancerous tissue for the surgeon prior or during the excision surgery. To this end, the border between the normal and tumor parts needs to be identified. Therefore, the images need to be segmented into tumor and normal areas. We propose a deep learning technique which divides the electromagnetic images into two regions: tumor and normal, with high accuracy. We formulate deep learning from a perspective relevant to electromagnetic image reconstruction. A recurrent auto-encoder network architecture (termed here DeepTMI) is presented. The effectiveness of the algorithm is demonstrated by segmenting the reconstructed images of an experimental tissue-mimicking phantom. The structure similarity measure (SSIM) and mean-square-error (MSE) average of normalized reconstructed results by the DeepTMI method are about 0.94 and 0.04 respectively, while that average obtained from the conventional backpropagation (BP) method can hardly overcome 0.35 and 0.41 respectively.

Surgical resection is the mainstream treatment for solid tumors (tissues that normally do not contain cysts or liquid areas). The resected specimens are sent to pathology for evaluation and subsequently an additional piece of tissue is removed if cancer is observed in the margins of the resected pieces. In this procedure, sampling error may be present as less than 2% of the total sample is normally examined microscopically due to tissue processing and sectioning<sup>1</sup>. Therefore, the presence of tumor at the margins is sometimes underestimated or even missed. On the other hand, often times some normal tissue is unnecessarily removed during surgery. Since lesions generally occur on sun-exposed body parts, for instance the face, minimizing tissue loss is important for improving the quality of life of cancer patients.

However, the current methods to distinguish malignant from healthy tissue are primarily limited to tactile and visual cues as well as the surgeon's experience. As a result, a positive surgical margin or the presence of residual tumor may be left behind after resection. It is well-documented that the presence of residual tumor can negatively impact treatment outcomes and survival, as well as pose an economic burden<sup>2</sup>.

Imaging techniques would be highly valuable in assisting surgeons to identify tumor margins pre- or intra-operatively to evaluate the remainder of the resected lesion on the body. Several non-invasive imaging tools such as neuro-navigation, magnetic resonance imaging (MRI), ultrasound (US), Raman spectroscopy (RS), and optical fluorescence imaging (FLI) have been developed for this purpose<sup>3-14</sup>. However, they each have limitations on their uses and provide optimal outcomes only under certain conditions. No imaging tool has been widely adopted in the clinics. There is a definite need for an imaging tool that could visualize tumor profiles with sufficient contrasts throughout the depth of the skin at an affordable cost.

In Ref.<sup>15</sup>, the authors developed a high-resolution millimeter-wave imaging (HR-MMWI) system for early-stage detection of skin cancer. In-vivo measurements of more than 136 cancerous and benign skin lesions were performed. In particular, the diagnosis of skin lesions with clinical interest, i.e., cancer/precancer versus benign lesions, was concerned. The findings established that real-time millimeter-wave imaging can distinguish between malignant tissues and benign skin lesions with high diagnostic accuracy (97% sensitivity and 98% specificity). As the next step, the capability of the system in identifying tumor margins pre- or intra-operatively needs to be verified. This capability will potentially simplify the tumor removal surgery to a single-layer excision procedure by visualizing the borders of the cancerous tissue for the surgeon. Nonlinear electromagnetic (EM) inverse scattering is the conventional imaging technique for accurately reconstructing dielectric properties of a target. However, there are important challenges arising from intrinsic strong nonlinearity, ill-posed-ness, and high computational costs.

<sup>1</sup>RadioSight LLC, Hoboken, NJ 07030, USA. <sup>2</sup>Department of Electrical and Computer Engineering, Stevens Institute of Technology, 1 Castle Point Ter, Hoboken, NJ 07030, USA. <sup>3</sup>Stanford University School of Medicine, Stanford, CA, USA. ✉email: negar.ebadi@stevens.edu

To tackle these difficulties, in this work, we propose a novel method for tumor margin identification in 3-D electromagnetic imaging. To identify tumor margins, the border between the normal and tumor parts needs to be identified. Therefore, the images need to be segmented into tumor and normal areas. The proposed method is a learned technique which divides the electromagnetic images into two regions, i.e. tumor and normal, with high accuracy. We establish a fundamental connection between a deep neural network architecture and iterative methods utilized for nonlinear EM inverse scattering problems. Inspired by this connection, we then develop a novel DNN architecture tailored for tumor margin identification (termed here DeepTMI), which consists of a cascade of multilayer neural network modules.

The performance of the proposed method is validated by an experimental demonstration. We examine the effectiveness of the algorithm by segmenting the reconstructed images of an experimental skin phantom<sup>16</sup>. Specifically, it is shown that the method is a promising tool for efficiently tackling nonlinear inverse scattering problems, which are impractical to be solved by using conventional methods.

### Problem statement

Over the past few years, deep learning has consolidated as one of the most powerful approaches in several areas of regression and classification, due to ease of availability of vast amounts of data and the ever-increasing computational power<sup>17,18</sup>. Deep neural networks (DNN) specifically have attracted attention in image processing and computer vision, such as semantic segmentation<sup>19</sup> depth estimation<sup>20</sup>, image deblurring<sup>21</sup>, and image super resolution<sup>22,23</sup>. DNN was also demonstrated to be advantageous over traditional machine learning approaches in the automated analysis of high-content microscopy data<sup>24</sup>. Deep learning was also shown to aid the design and realization of advanced functional materials<sup>25</sup> and high-accuracy reconstruction from compressed measurements<sup>26,27</sup>. Most recently, DNN algorithms have been applied in biomedical imaging, e.g., magnetic resonance imaging<sup>28</sup>, X-ray computed tomography<sup>29</sup>, and computational optical imaging<sup>30,31</sup>. It has been empirically found that NN-based<sup>32,33</sup> and DNN-based strategies can outperform conventional image reconstruction techniques in terms of improved image quality and reduced computational costs<sup>28–33</sup>.

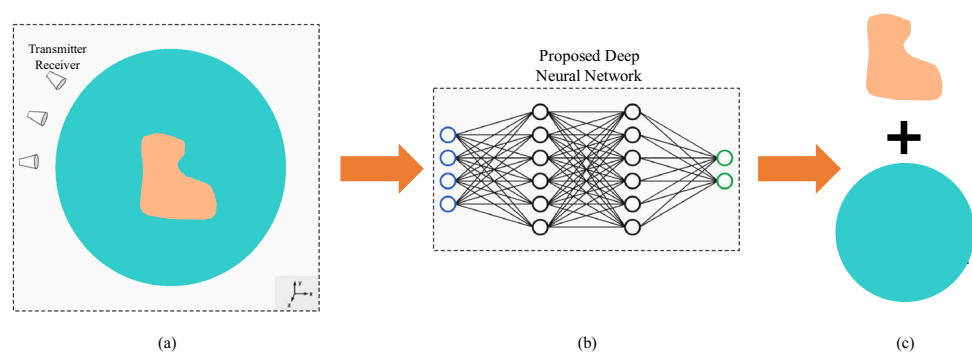
Motivated by the capability of neural networks, this paper focuses on adapting a new architecture, namely DeepTMI to iterative optimizations utilized for the nonlinear EM inverse scattering problems. The proposed method is a learned technique which divides the electromagnetic images into two regions, i.e. tumor and normal, with high accuracy and approximately characterizes the multi-scattering physical mechanism. The proposed neural network module is a straightforward extension of the conventional neural networks<sup>22</sup>, which is an end-to-end map from an input rough image to the refined solution of a nonlinear inverse scattering problem. The input data of the first module of the network comes from the backpropagation (BP) image. For the remaining modules of the network, the input is the output of the previous module.

We initiate the discussion by unveiling the connection between the DNN architecture of interest and iterative methods for nonlinear EM inverse scattering. Since the iterative solution of a nonlinear EM inverse scattering requires convolutions and should account for nonlinearities, DNN may offer an efficient alternative solution.

### Connection between DNN and nonlinear EM inverse scattering

With reference to the measurement configuration in Fig. 1a, we illustrate our strategy in the context of a multiple-input multiple-output measurement configuration. For computational imaging, the investigation domain is uniformly divided into pixels such that the total electric fields, the contrast currents, and the contrast functions are assumed uniform in each pixel. In the area of image processing, it has become a consensus that most of the natural images have some structure. This underlying structure allows for a sparse representation in some transformed domain, which also assists in regularization.

In ultra-wideband imaging also, the presence and location of significant backscatters such as a malignant tumor is identified through solving the inverse problem of (1), i.e. a reconstruction technique. The main challenge in image reconstruction is devising an algorithm which provides high resolutions and suppressions of



**Figure 1.** Basic configuration of the developed deep learning-based solver. (a) The complex-valued input comes from conventional reconstruction algorithms, and the output is the image of inverse scattering. The proposed method (b) divides the electromagnetic image into two regions, tumor and normal, with high accuracy as shown in (c).

strong artifacts and noise. Various image reconstruction algorithms have been proposed for ultra-wide band imaging<sup>34,35</sup>. These techniques essentially employ the delay-and-sum (DAS) concept (with various modifications) in which the backscattered signals are time-shifted and summed to create a synthetic focal point of the target.

We can model the received signal ( $S$ ) as follows<sup>16</sup>:

$$S = [S(r)] = \mathcal{F}[s] = \iint s(x, f) \exp\{-2jkR^x(r)\} df dx = Fs \quad (1)$$

where  $x$  is the three-dimensional position of the scatterer,  $f$  is the frequency,  $s(x, f)$  is the reflectivity function of scatterer,  $R^x$  is the electrical distance between the scatterer and the scanning point  $r$ , and  $k = \omega/c$  is the wave-number ( $\omega$  is the temporal angular frequency and  $c$  is the speed of light).

Iterative sparsity promoting algorithms have been recently proposed for image reconstruction. As demonstrated in (1), the scattering data of a target at high frequencies can be represented as a sum of responses from individual scattering centers. These scatterers are sparsely distributed in the sensing scene i.e., most of the coefficients in the target's signature vector are zeros or nearly zeros. We use an iterative approach to reconstruct the level set and corresponding reflectivity profile from a set of measured and computed data. A common non-linear image formation method is to regularize a quadratic data likelihood function and solve the following optimization problem.

The use of sparsity in signal processing frequently calls for the solution to the minimization problem<sup>36</sup>:

$$\hat{s} = \text{Argmin}_C(s) = \frac{1}{2} \|Fs - S\|_2^2 + \lambda \|s\|_1 \quad (2)$$

where  $C$  is a regularized least-squares cost function used to assess the similarity of the measured received signal ( $S$ ) to the total computed reflected signal for the reconstructed reflectivity profile. The problem in (2) can be solved iteratively using a forward-backward splitting algorithm. First a gradient descent update is performed for the smooth component of the objective function (the  $\lVert \cdot \rVert_2$ -norm term) and a projection is done for the non-smooth term. This results in the following iteration:

$$s^{k+1} = \mathcal{P}_\lambda \left( s^k - \alpha F^H (Fs^k - S) \right) \quad (3)$$

where  $\mathcal{P}_\lambda$  is the shrinkage operator and performs a soft thresholding of the input (i.e., the proximity operator for the  $\lVert \cdot \rVert_1$ -norm).

### DNN for nonlinear EM inverse scattering

The iterative imaging algorithm of (3) emulates a feed-forward neural network which is the most general architecture for deep learning. In a feed-forward neural network, each layer of the network represents an iteration and the inner product with a weight matrix plus bias is fed as the input. The neural network is formed by unfolding the iterative optimization method<sup>37</sup>. Motivated by this recent application of deep learning to inverse problems and the power of deep neural networks to approximate non-linear mappings, we investigate the use of deep learning in reconstructing the level-set function. A level set is a real-valued function defined in the whole imaging domain. An example of a single level set segmenting a domain into distinct regions is shown in Fig. 1a, where The goal is to accurately identify the sections with dielectric properties corresponding to cancer tissue, and regions with dielectric properties corresponding to normal tissue as shown Fig. 1b,c.

After demonstrating the natural connection between the DNN architecture and nonlinear EM inverse scattering, we now develop a complex valued DNN to solve the nonlinear EM inverse scattering problem. For the sake of DNN computational complexity, the proposed DeepTMI can be designed as a cascade of CNN modules where the input data of the network comes from the BP image. For the remaining modules of the network, the output of the previous CNN module is the input of the next module. Each module consists of several up-sampling convolution layers, each layer consisting of three steps: in the first step, the input is convolved with a set of learned filters, resulting in a set of feature (or kernel) maps; in the second step, the maps undergo a point-wise nonlinear function, resulting in a sparse outcome; an optional third down-sampling step (termed as pooling) is applied on the result to reduce its dimensions, thus forming the multilayer structure.

### Algorithm implementation

Deep learning has recently become immensely popular in a wide variety of traditional signal processing tasks such as image segmentation<sup>38</sup>, denoising<sup>39</sup>, and point source localization in the presence of noise<sup>40</sup>, displaying extremely impressive results. In the area of remote sensing and RF imaging, deep learning is mostly applied to terrain surface classification<sup>41,42</sup>, target segmentation<sup>43,44</sup>, object detection<sup>45</sup>, antenna selection in cognitive radar<sup>46</sup>, interference mitigation<sup>47</sup>, vehicle detection<sup>48</sup> in automotive applications, and activity recognition in indoor monitoring<sup>49-51</sup>. In all these works, the neural networks operate on images that have already been formed using other image formation approaches. Despite the success in optical imaging, the huge potential of deep learning in UWB image reconstruction remains locked.

Here, we review the most relevant deep learning models, point out possible pitfalls by analyzing special characteristics of electromagnetic scattering data, evaluate the performance of the state-of-the-art of deep learning when applied to remote sensing and inverse problems, and conclude with an optimized framework to be used in our imaging experiments.

### CNNs

Convolutional neural networks (CNNs) have attracted worldwide attention and are currently used for many image understanding tasks, such as image classification, object detection, and semantic segmentation. Residual neural network (ResNet)<sup>52</sup>, U-Net<sup>53</sup>, and DenseNet<sup>54</sup> were the subsequent major CNN architectures. Their main feature concerned the idea of connecting not only neighboring layers but any two layers in the network by using skip connections. This helped reduce information loss across networks, mitigated the problem of vanishing gradients, and facilitated the design of deeper networks. U-Net is among the most common image segmentation networks.

### RNNs

Besides CNNs, recurrent neural networks (RNNs)<sup>55</sup> are a major class of deep networks. Their main building blocks are recurrent units, which take the current input and the output of the previous state as their input. They provide state-of-the-art results for processing data of variable lengths, including text and time-series information. Their weights can be replaced with convolutional kernels for visual processing tasks, such as image captioning and predicting future frames/points in visual time-series data.

### GANs

Proposed by Ian Goodfellow et al.<sup>56</sup>, generative adversarial networks (GANs) are among the most popular inventions in the field of deep learning. Based on game theoretic principles, they consist of two networks, i.e. a generator and a discriminator. The generator's objective is to learn a latent space through which it can create samples from the same distribution as the training data, while the discriminator tries to learn to distinguish whether a sample is from the generator or the training data.

The learning procedure in all the above-mentioned networks takes place by optimizing a cost functional (typically an  $L^2$ -norm of the mismatch between the network output and the ground truth) with respect to network weights. This optimization is mostly carried out using the stochastic gradient descent (SGD) algorithm. The optimization problem is typically high-dimensional and highly non-convex, often consisting of many saddle points and non-optimal local minima. Therefore, a critical aspect would be the initialization of the network weights. Prior knowledge of the input-to-output mapping initializing the back propagation increases the chance of reaching an optimal solution. In this research, we will use the physical forward model of UWB imaging to obtain suitable initialization.

The weights of the network provide a parametrization of the operation that the network performs, whereas the non-linear units introduce the capacity to approximate complex, non-linear mappings between input and output spaces. These sophisticated transformations are generally explained in terms of the universal approximation theorem or probabilistic inference<sup>37,57</sup>. Although the universal approximation theorem states that a single hidden layer neural network is a universal approximator, the depth of the architecture introduces new strengths and capabilities<sup>58</sup>. This results in deep networks gaining more expressive power and the ability to represent more complex mappings. This is a direct consequence of the fact that ANNs are formed by the composition of multiple non-linear mappings.

The output at each layer of the neural network can be characterized as

$$h^{(k+1)} = \sigma(\tilde{W}^k \cdot \tilde{h}^k + w_b^k) \quad (4)$$

where  $\tilde{h}^k$  is the input vector,  $\tilde{W}^k$  is the weight matrix, and  $w_b^k$  is the bias weight vector of the  $k$ th layer,  $\sigma$  is an element-wise, non-linear activation function, and  $h^{k+1}$  is the output of the  $k$ th layer, and subsequently the input of the  $(k+1)$ th layer of the network.

In our case, we use the auto-encoder framework combined with a recurrent neural network (RNN) to form a recurrent auto-encoder<sup>59</sup>. An RNN fundamentally differs from other architectures by fixing the weight matrix across all layers. Essentially, an RNN can be interpreted as an implementation of an iterative algorithm for a fixed number of iterations. This makes it ideal as an architecture for applying deep learning to the level-set image reconstruction problem as it allows for the utilization of existing image reconstruction methods such as the DAS algorithm. Furthermore, RNNs can be trained in a similar fashion to other network architectures. Specifically, a stochastic gradient descent with the DAS algorithm can be used to analytically calculate the derivatives.

We set up the RNN by unfolding the algorithm up to a fixed number of iterations. To demonstrate how (3) relates to (4), we first define  $\mathbf{K} = \mathbf{I} - \alpha \mathbf{F}^H \mathbf{F}$ . Using this notation, we can express (3) as:

$$s^{k+1} = \mathcal{P}_\lambda \left( \mathbf{K} s^k + \alpha \mathbf{F}^H \mathbf{S} \right) \quad (5)$$

It is clear that the weight matrix in (4) is  $\mathbf{K}$ , the bias is  $\mathbf{F}^H \mathbf{F}$ , and the non-linear activation function is the shrinkage operator. The iterates in (5) produce a level-set estimate for a sufficiently deep network. The depth is necessary since the algorithm requires many iterations to converge. In practical applications, due to phase errors, the forward propagation produces defocused and smeared images. To suppress artifacts due to phase errors, we refine the ideal forward model  $\mathbf{F}$  by learning  $\mathbf{F}^H$  and  $\mathbf{K}$  through the DAS algorithm.

While the current network architecture imitates image formation methods, it has limitations when it comes to supervised learning. In inverse imaging problems, supervised learning may not be applicable since the exact object generating the data, including precise tumor margins, is not fully known in real-world scenarios. Only the measured data and knowledge of the mapping are available. Additionally, the availability of training data may be limited. To address these challenges, we adopt an unsupervised approach using an autoencoder structure to learn from the data.

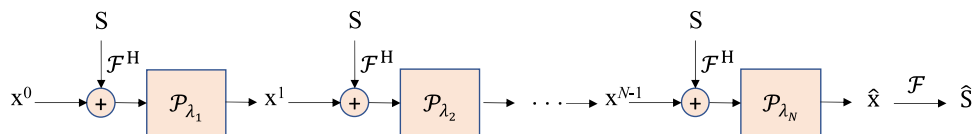
To transform the above RNN architecture into an autoencoder, we add a new linear layer after the last RNN layer. This additional layer maps the estimated image back to the data using function  $F$ , enabling unsupervised training. The resulting architecture is a recurrent auto-encoder with  $N$  layers. The first  $(N - 1)$  layers are based on Eq. (5), while the final  $N$ th layer performs forward propagation. Figure 2 illustrates the final architecture, where the regularization parameter can vary at each layer.

Although our training framework provides an effective approach, there are two challenges to consider. Firstly, storing the network parameters can be challenging due to the size of the images. Secondly, RNN training encounters the vanishing or exploding gradients problem, which arises when the weight matrix is poorly conditioned<sup>60</sup>. This issue is characterized by either oscillatory behavior or extremely slow convergence. To overcome this problem, the introduction of long short-term memory (LSTM) RNNs has proven beneficial<sup>61</sup>. LSTM consists of a memory cell along with input, output, and forget gates, which help address the exploding and vanishing gradient problem by modifying the image reconstruction process to be additive instead of multiplicative. LSTM is a popular RNN architecture known for its ability to store values from past instances and mitigate the issue of gradient diminishing. In the context of our imaging approach using wideband pulses, RNNs, including LSTM, are natural choices for processing time-series information.

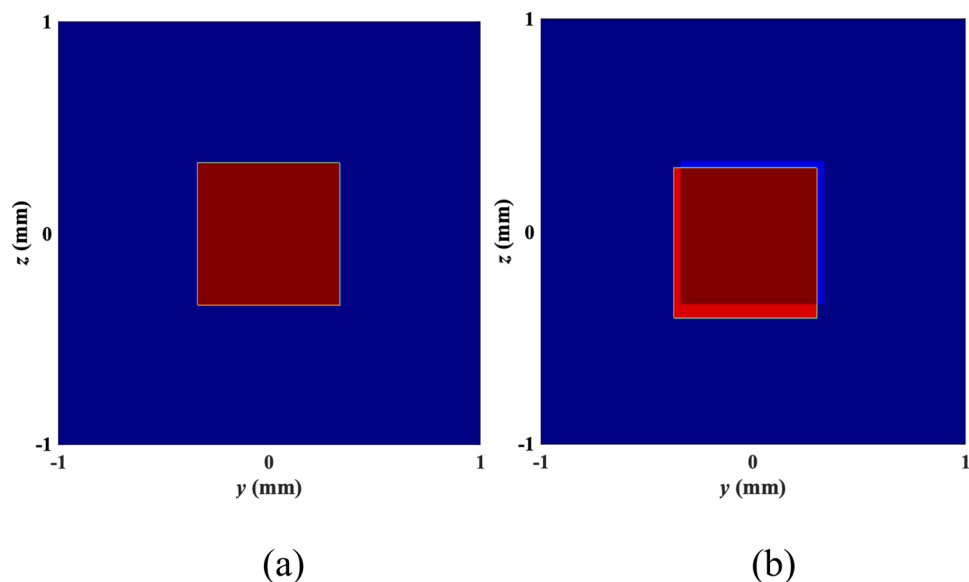
### Experimental results

The effectiveness of our algorithm is first demonstrated by segmenting the reconstructed images of numerical phantoms containing tumors. The data are derived from our prior work on the development of skin-equivalent phantoms for mimicking interactions of millimeter waves with the human skin and skin tumors<sup>62</sup>. These phantoms closely mimic the dielectric properties of normal skin and cancer tissues at millimeter-wave frequencies. We test our algorithm to reconstruct the normal skin and the tumor tissue, while keeping the corresponding electrical properties similar to Ref.<sup>62</sup>.

We use a  $2 \times 2 \text{ m}^2$  scene with a single stationary point target discretized into  $30 \times 30$  pixels with the origin of the coordinate system located at the center of the scene as shown in Fig. 3a. The forward model explained in “Connection between DNN and nonlinear EM inverse scattering” was used to create a bistatic received signal. The circular trajectory was navigated by two antennas. The transmitted waveform had a bandwidth of 40 GHz and a center frequency of 90 GHz. To build a training set of 100 samples, we uniformly sampled the signal in slow time. Each measurement in the training set included additive phase error with a uniform distribution and a variance of 3 m, representing the antenna trajectory error. Additionally, a test set of 30 measurements was



**Figure 2.** The recurrent auto-encoder architecture capable of training with measurements. The linear stages are represented as arrows. The non-linear activation functions are represented as boxes. At each layer, the linear gradient descent step is followed by the shrinkage operation.



**Figure 3.** (a) Actual geometry. (b) Joint contrast and shape recovery.

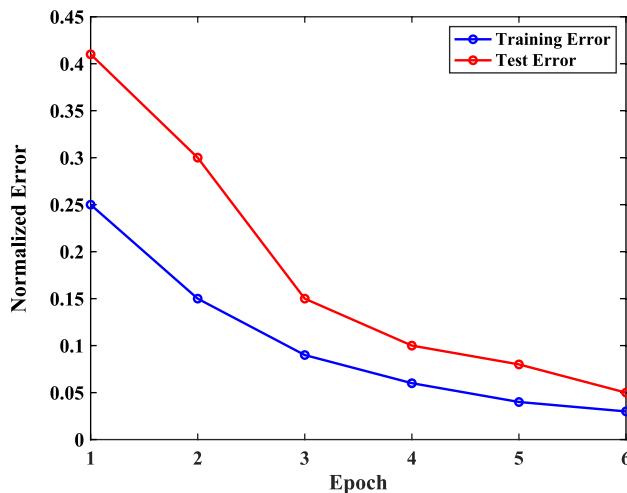
generated in the same manner. For our analysis, we constructed an 8-layer network following the architecture detailed in “DNN for nonlinear EM inverse scattering”.

In our simulations, the network was trained for 6 epochs at which point we observed that the algorithm converged. We used the normalized  $\ell^2$ -norm error of the data residual as our metric. Figure 4 demonstrates the average test and training errors using 100 training samples and 30 test samples. As expected, the average errors for both groups decrease with respect to the epoch number. The test error decreases at a lower rate compared to the training error which is also typical.

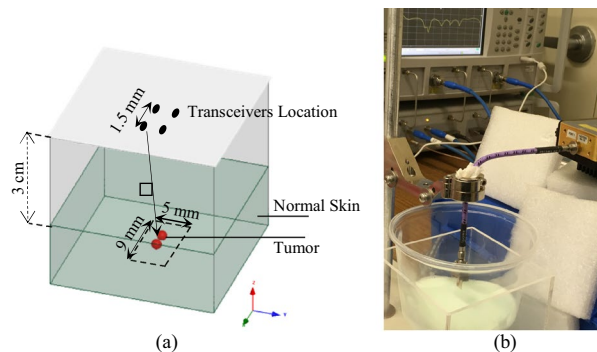
For visual comparison between the results obtained, we display a reconstructed test image formed after the application of the proposed DeepTMI method. The reconstructed image is shown in Fig. 3b. To quantify our observations, we use the Intersection over Union (IoU) performance metric for shape estimation and the relative error metric to characterize the reflectivity function estimation. IoU is defined as the ratio of the volume of overlap between recovery and original to the volume of the union of the two. In the perfect shape recovery, IoU is 1. We use the average root-mean-square (rms) error as the second metric. The quantitative shape error for the cube is 5% and the relative error is 0.01%.

The effectiveness of our algorithm is further demonstrated by segmenting the reconstructed images of numerical phantoms containing tumors. Experimental setups similar to Fig. 5a are considered for the evaluation of the performance of the proposed method. Data collection is performed in a monostatic manner in which the antennas are scanned over an aperture plane placed in front of the target. The operating frequency range of 75–110 GHz is considered. The antennas are SIW-based Vivaldi antennas previously developed by our group<sup>65</sup>. The antennas scan over the aperture plane across  $2 \times 2$  positions in a rectangular pattern as shown in Fig. 5a. The spacing between consecutive scan positions is 1.5 mm, which is the maximum spacing that satisfies the Nyquist criterion. The distance between the aperture plane and the skin surface is 3 cm. The Nyquist criterion requires the maximum spacing of ( $\sim \lambda/2$ ) to be met between two consecutive scans. Considering the smallest wavelength of  $\lambda = 3$  mm in our imaging system, the Nyquist spacing corresponds to 1.5 mm.

The phantoms were fabricated from our prior work on developing skin phantoms to mimic the interactions of electromagnetic waves with human skin and skin tumors<sup>62</sup>. These phantoms closely emulate the dielectric



**Figure 4.** Averaged training and test loss vs epochs.



**Figure 5.** (a) Schematic layout and (b) testbed of experimental setup for evaluation of the performance of the proposed method.

properties of normal skin and cancer tissues at millimeter-wave frequencies. Figure 5b shows the artificial phantoms fabricated and installed on the developed imaging testbed. Realistic skin phantoms serve as an invaluable tool for exploring the feasibility of new technologies and improving design concepts related to millimeter-wave skin cancer detection methods. Normal and malignant skin tissues are separately mimicked by using appropriate mixtures of deionized water, oil, gelatin powder, formaldehyde, TX-150 (a gelling agent, widely referred to as ‘super stuff’), and detergent.

Dispersive skin-mimicking phantoms available in the literature<sup>62</sup> with a finite thickness of 2 mm were fabricated using a mixture of materials (water, gelatin, oil, salt, and surfactant) following the procedure presented in Ref.<sup>62</sup>. To emulate the resolution experiments, two canonical-spherical tumors with diameters of 400  $\mu\text{m}$  were also fabricated and inserted within the phantoms. The tumor-mimicking mixture had 30% (in weight) higher oil content than the skin-mimicking mixture. This increase in oil percentage results in the closest possible dielectric properties to those of malignant skin tissues (specifically malignant SCC tissues) as reported in Ref.<sup>64</sup>.

We used single-pole Cole–Cole fitting to efficiently represent the measured data of biological tissues as:

$$\varepsilon_r(\omega) = \varepsilon_r'(\omega) - j\varepsilon_r''(\omega) = \varepsilon_{r0}(\omega) + \frac{\Delta\varepsilon_r}{1 + (j\omega\tau)^{1-\alpha}} + \frac{\sigma_s}{j\omega\varepsilon_0} \quad (6)$$

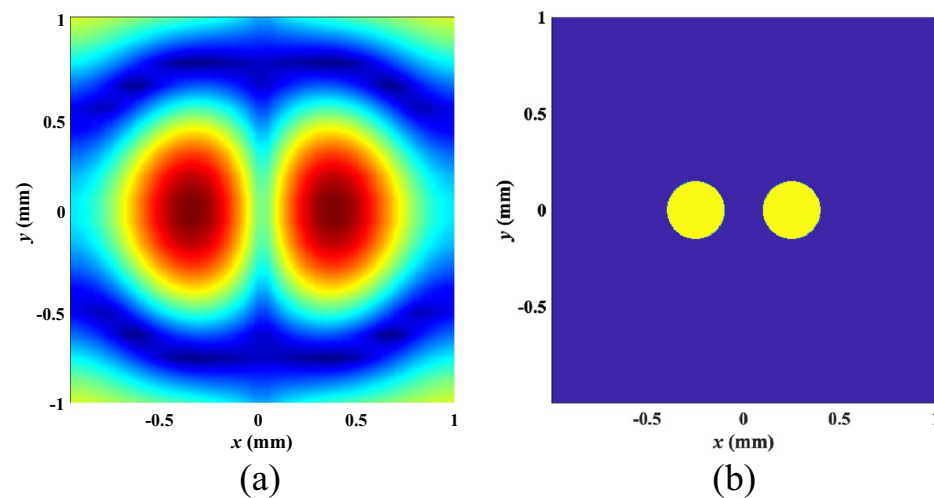
where  $\varepsilon_r'(\omega)$  and  $\varepsilon_r''(\omega)$  are the real and imaginary parts of the frequency-dependent relative dielectric permittivity ( $\varepsilon_r(\omega)$ ),  $\omega = 2\pi f$  ( $f$  is the frequency of operation),  $j = (-1)^{1/2}$ ,  $\tau$  is the relaxation time,  $\Delta\varepsilon_r$  is the magnitude of the dielectric dispersion of the skin,  $\varepsilon_{r0}$  is the permittivity of skin at optical frequencies,  $\varepsilon_0$  is the permittivity of free space,  $\alpha$  is a measure of the broadening of the dispersion, and  $\sigma_s$  is the ionic conductivity of the skin. The Cole–Cole parameters of the dielectric properties for each phantom are presented in Table 1.

For the dataset, 50 arrangements were used for training. Different spherical-shaped tumor models with different sizes (diameters of 0.1–0.4 mm) and different distances (center distances 0.5–1 mm) are considered. The proposed network was trained using the Adam Optimization<sup>65</sup>. The regularization parameter was set to  $\lambda = 10^{-7}$  and the learning rate was set to  $10^{-4}$ . The number of hidden layers used was 8. In our experiments, the network was trained for 6 epochs at which point we observed that the algorithm converged.

For visual evaluation of the results obtained, in Fig. 6, we demonstrate the reconstructed image formed after the application of a reconstruction algorithm we developed in Ref.<sup>16</sup> and the corresponding level-set function after application of the proposed learning method. The target consists of two spherical tumors as similar to Fig. 5. The defocusing effect when applying the conventional reconstruction algorithm is clearly displayed, which is also quantified by a poor T/C measure (the ratio of the peak tumor signal to the peak clutter). For this test sample, the T/C ratios are 5.5 dB and 22 dB for the images reconstructed using our conventional and the level-set approaches, respectively.

Parameter	$\varepsilon_0$	$\Delta\varepsilon_r$	$\tau$ (ps)	$\sigma_s$	$\alpha$
Normal skin	4.25	45.02	5.96	0.0005	0.15
Malignant skin	10	34.99	3.04	0.0004	0.05

**Table 1.** Cole–Cole parameters of the dielectric properties of the fabricated phantoms.



**Figure 6.** Image reconstructed by (a) the algorithm in Ref.<sup>16</sup> and (b) the proposed deep learning-based level-set method.

To quantify our observations, we use the structure similarity measure (SSIM) and mean-square-error (MSE) of the reconstructed images. SSIM and MSE are approximately 0.94 and 0.04 for the proposed method, while the values obtained from conventional frequency-domain reconstruction methods can hardly overcome 0.35 and 0.41, respectively.

## Discussion

In our previous imaging studies, we used an ultra-wideband image reconstruction technique developed in Ref.<sup>16</sup> to create three-dimensional reconstructions of objects. Although the developed reconstruction technique takes the dispersive behavior of the target over the ultra-wide imaging bandwidth of nearly 100 GHz into account, it involves large amounts of data due to the high sampling rate needed to produce high-resolution images. This leads to an increased number of slow-time pulses and hence, increased computational complexity in the imaging algorithm. In addition to requiring high sampling rates, the algorithm obtains high-resolution images by transmitting pulses with ultra-wide bandwidths; the wider the pulse bandwidth in frequency, the narrower the pulse in time, improving depth resolution. Thus, it is desired to reduce the number of slow-time pulses and increase the bandwidth to maintain high-resolution imaging. However, the data acquisition process is frequently plagued by phase errors. Due to inaccuracies in the scanning measurement as well as possible existence of moving targets (such as arm and hand movements of patients during our clinical examinations) in the observed scene, the acquired data will contain phase errors. These phase errors in turn result in a defocusing effect in the formed images.

Many different algorithms have been developed over time to form images from backscattered data. These algorithms use several different approaches such as back projection<sup>66</sup>, compressed sensing<sup>67</sup>, or signal processing<sup>68</sup>. For sparse target scenes, adding the target detection process in the raw scattering process is an effective method to improve the imaging process. It is well known that the scattering data of a target at high frequencies can be represented as a sum of responses from individual scattering centers. These scatterers are sparsely distributed in the sensing scene i.e., most of the coefficients in the target's signature vector are zeros or nearly zeros.

The theory of compressed sensing (CS) states that a high-dimensional signal can be accurately and robustly recovered from its lower-dimensional projections if the signal is sparse or can be sparsely represented. CS theory can be conveniently applied to many ultra-wideband (UWB) imaging applications due to the natural presence of sparsity in these problems<sup>69</sup>. Sparsity-based signal processing methods have achieved great success in both suppressing interference<sup>70,71</sup> and performing spectral gap extrapolation<sup>72,73</sup> to combat frequency notches. However, the available sparsity-based techniques in UWB imaging and sensing cannot accurately represent the diffraction scattering behavior for wide relative bandwidths since the model order cannot be estimated accurately in those situations.

To overcome the above-mentioned challenges, the method proposed in this paper offers an alternative approach to the traditional image formation algorithms. We solve the image formation problem in the presence of phase uncertainties. Such uncertainties arise due to target movement and errors in chip scanning trajectories among other sources. Using an iterative imaging algorithm, we will form a new neural network that emulates this process and further refines the parameters via training for improved performance.

In our proposed learning-based approach, as shown in Fig. 2, all steps of the image reconstruction processing are integrated into a deep neural network. Input to the network is raw scattering data, while the network output is the focused image. With our approach, all computational operations within the scheme are performed through the forward propagation of a single neural network. There is no need to pass processed data from one processing component to another, greatly simplifying the processing effort.

The main limitation of the presented work was the small size of the training dataset. In our future work, we will use more diverse experimental setups to train and evaluate the performance of the proposed method.

## Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 5 April 2023; Accepted: 12 September 2023

Published online: 23 September 2023

## References

- Rigel, D. S. Non-invasive gene expression testing to rule out melanoma (2018).
- American Cancer Society. *Cancer Facts and Figures 2019*. (2022). <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2022/2022-cancer-facts-and-figures.pdf>. Accessed 2022.
- Nimsky, C. Intraoperative MRI in glioma surgery: proof of benefit? *Lancet Oncol.* **12**, 982 (2011).
- Sastry, R. *et al.* Applications of ultrasound in the resection of brain tumors. *J. Neuroimaging.* **27**, 5 (2017).
- Unsgaard, G. *et al.* Intra-operative 3D ultrasound in neurosurgery. *Acta Neurochir.* **148**, 235 (2006).
- Laing, S., Jamieson, L. E., Faulds, K. & Graham, D. Intra-operative 3D ultrasound in neurosurgery. *Nat. Rev. Chem.* **1**, 0060 (2017).
- Mirbeik-Sabzevari, A. *et al.* Synthetic ultra-high-resolution millimeter-wave imaging for skin cancer detection. *IEEE Trans. Biomed. Eng.* **66**(1), 61–71 (2019).
- van Dam, G. M. *et al.* Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor- $\alpha$  targeting: First in-human results. *Nat. Med.* **17**, 1315 (2011).
- Nguyen, Q. T. & Tsien, R. Y. Fluorescence-guided surgery with live molecular navigation. *Nat. Rev. Cancer.* **13**, 653 (2013).
- Vahrmeijer, A. L., Hutteman, M., van der Vorst, J. R., van de Velde, C. J. & Frangioni, J. V. Image-guided cancer surgery using near-infrared fluorescence. *Nat. Rev. Clin. Oncol.* **10**, 507 (2013).
- van der Vorst, J. R. *et al.* Near-infrared fluorescence-guided resection of colorectal liver metastases. *Cancer* **119**, 3411 (2013).
- Owens, E. A., Lee, S., Choi, J., Henary, M. & Choi, H. S. NIR fluorescent small molecules for intraoperative imaging. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **7**, 828 (2015).



13. Kim, T. H., O'Brien, C., Choi, H. S. & Jeong, M. Y. Targeted near-infrared fluorescence imaging for regenerative medicine. *Appl. Spectrosc. Rev.* **53**, 349 (2018).
14. Park, G. K., Kim, G. S., Hwang, N. S. & Choi, H. S. Optical spectroscopic imaging for cell therapy and tissue engineering. *Appl. Spectrosc. Rev.* **53**, 360 (2018).
15. Mirbeik, A. *et al.* Real-time high-resolution millimeter-wave imaging for in-vivo skin cancer diagnosis. *Sci. Rep.* **12**, 4971 (2022).
16. Mirbeik-Sabzevari, A., Ashinoff, R. & Tavassolian, N. High-contrast, low-cost, 3D visualization of skin cancer using ultra-high-resolution millimeter-wave imaging. *IEEE Trans. Med. Imaging* **38**(9), 2188–2197 (2019).
17. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**(28), 436–444 (2015).
18. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* 10–13 (MIT Press, 2016).
19. Long, J., Shelhamer, E., Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440 (2015).
20. Eigen, D., Puhrsch, C., Fergus, R. Depth map prediction from a single image using a multi-scale deep network. In *Proceedings of the Neural Information Processing Systems* 2366–2374 (2014).
21. Sun, J., Cao, W., Xu, Z., Ponce, J. Learning a convolutional neural network for non-uniform motion blur removal. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 769–777 (2015).
22. Dong, C., Loy, C. C., He, K. & Tang, X. Image superresolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(2), 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281> (2016).
23. Liu, D. *et al.* Robust single image super-resolution via deep networks with sparse prior. *IEEE Trans. Image Process.* **25**(7), 3194–3207 (2016).
24. Kraus, O. Z. *et al.* Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.* **13**(4), 924 (2017).
25. Kalinin, S. V., Sumpster, B. G. & Archibald, R. K. Big-deep-smart data in imaging for guiding materials design. *Nat. Mater.* **14**(4395), 973–980 (2015).
26. Mousavi, A., Baraniuk, R. Learning to invert: Signal recovery via deep convolutional networks. In *Proceedings of ICASSP 2272–2276* (2017).
27. Kulkarni, K., Lohit, S., Turaga, P., Kerviche, R., Ashok, A. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA* 449–458 (2016).
28. Han, Y. S., Yoo, J., Ye, J. C. Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis. <https://arxiv.org/abs/1611.06391> (2016).
29. Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**(9), 4509–4522 (2017).
30. Sinha, A., Lee, J., Li, S. & Barbastathis, G. Lensless computational imaging through deep learning. *Optica* **4**(9), 1117–1125 (2017).
31. Kamilov, U. S. *et al.* Learning approach to optical tomography. *Optica* **2**(6), 517–522 (2015).
32. Marashdeh, Q., Warsito, W., Fan, L.-S. & Teixeira, F. L. Nonlinear forward problem solution for electrical capacitance tomography using feed-forward neural network. *IEEE Sens. J.* **6**(2), 441–449 (2006).
33. Marashdeh, Q., Warsito, W., Fan, L.-S. & Teixeira, F. L. A nonlinear image reconstruction technique for ECT using a combined neural network approach. *Meas. Sci. Technol.* **17**(8), 2097–2103 (2006).
34. Fear, E. C., Li, X., Hagness, S. C. & Stuchly, M. A. Confocal microwave imaging for breast cancer detection: Localization of tumors in three dimensions. *IEEE Trans. Biomed. Eng.* **49**(8), 812–822 (2002).
35. Bond, E. J., Li, X., Hagness, S. C. & van Veen, B. D. Microwave imaging via space-time beamforming for early detection of breast cancer. *IEEE Trans. Antennas Propag.* **51**(8), 1690–1705 (2003).
36. Bach, F., Jenatton, R., Mairal, J. & Obozinski, G. Optimization with sparsity-inducing penalties. *Found. Trends Mach. Learn.* **4**, 1–106 (2012).
37. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **4**, 251–257 (1991).
38. Han, Y. & Ye, J. C. Framing U-net via deep convolutional framelets: Application to sparse-view CT. *IEEE Trans. Med. Imaging* **37**(6), 1418–1429 (2018).
39. Long short-term memory. Wikimedia. [https://upload.wikimedia.org/wikipedia/commons/thumb/3/3b/The\\_LSTM\\_cell.png/1280px-The\\_LSTM\\_cell.png](https://upload.wikimedia.org/wikipedia/commons/thumb/3/3b/The_LSTM_cell.png/1280px-The_LSTM_cell.png). Accessed 2022.
40. Yang, Y., Zheng, K., Wu, C. & Yang, Y. Improving the classification effectiveness of intrusion detection by using improved conditional variational autoencoder and deep neural network. *Sensors* **19**(11), 2528 (2019).
41. Parikh, H., Patel, S. & Patel, V. Classification of SAR and PolSAR images using deep learning: A review. *Int. J. Image Data Fusion* **11**(1), 1–32 (2020).
42. Chen, S., Wang, H., Xu, F. & Jin, Y.-Q. Target classification using the deep convolutional networks for SAR images. *IEEE Trans. Geosci. Remote Sens.* **54**(8), 4806–4817 (2016).
43. Furukawa, H. Deep learning for end-to-end automatic target recognition from synthetic aperture radar imagery. [arXiv:1801.08558 \[cs\]](https://arxiv.org/abs/1801.08558) (2018).
44. Furukawa, H. SAVERS: SAR ATR with verification support based on convolutional neural network. [arXiv:1805.06298 \[cs\]](https://arxiv.org/abs/1805.06298) (2018).
45. Chen, S., Wang, H. SAR target recognition based on deep learning. In *Proc. Int. Conf. Data Sci. Adv. Anal. (DSAA)*, 541–547 (2014).
46. Creswell, A. *et al.* Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **35**(1), 53–65 (2018).
47. Zitnik, M., Agrawal, M. & Leskovec, J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics* **34**(13), 457–466 (2018).
48. Huang, B., Carley, K. M. Residual or gate? Towards deeper graph neural networks for inductive graph representation learning. [arXiv:1904.08035](https://arxiv.org/abs/1904.08035) (2019).
49. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998).
50. Alioscha-Perez, M., Berenguer, A. D., Pei, E., Oveneke, M. C., Sahli, H. Neural architecture search under black-box objectives with deep reinforcement learning and increasingly-sparse rewards. In *2020 International Conference Artificial Intelligence in Information and Communication (ICAIC), Fukuoka, Japan* 276–281 (2020).
51. LeCun, Y., Cortes, C., Burges, C. Mnist handwritten digit database. 2010. <http://yann.lecun.com/exdb/mnist/>. Accessed 2022.
52. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770–778 (2016).
53. Ronneberger, O., Fischer, P., Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention* 234–241 (2015).
54. Huang, G., Liu, Z., Weinberger, K., Maaten, L. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern Recognition (CVPR)* 2261–2269 (2017).
55. Pearlmutter, B. A. Learning state space trajectories in recurrent neural networks. *Neural Comput.* **1**(2), 263–269 (1989).
56. Goodfellow, I. *et al.* Generative adversarial nets. In *Proceedings of the International Conference on Neural Information Processing Systems* 2672–2680 (2014).
57. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **2**, 303–314 (1989).
58. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

59. Feng, W., Guan, N., Li, Y., Zhang, X., Luo, Z. Audio visual speech recognition with multimodal recurrent neural networks. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)* 681–688 (2017).
60. Pascanu, R., Mikolov, T. & Bengio, Y. On the difficulty of training recurrent neural networks. *ICML* **28**(3), 1310–1318 (2013).
61. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
62. Mirbeik-Sabzevari, A. & Tavassolian, N. Ultrawideband, stable normal and cancer skin tissue phantoms for millimeter-wave skin cancer imaging. *IEEE Trans. Biomed. Eng.* **66**, 176–186 (2018).
63. Mirbeik-Sabzevari, A. *et al.* W-Band micromachined antipodal vivaldi antenna using SIW and CPW structures. *IEEE Trans. Antennas Propag.* **66**, 6352–6357 (2018).
64. Mirbeik-Sabzevari, A., Ashinoff, R. & Tavassolian, N. Ultra-wideband millimeter-wave dielectric properties of freshly-excised normal and malignant human skin tissues. *IEEE Trans. Biomed. Eng.* **65**, 1320–1329 (2018).
65. Kingma, D. P., Ba, J. Adam: A method for stochastic optimization (2014) [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
66. Gorham, L. A., Moore, L. J. SAR image formation toolbox for MATLAB. In *Presented at the SPIE Defense, Security, and Sensing, Orlando, Florida 769906* (2010).
67. Wei, S.-J., Zhang, X.-L., Shi, J. & Xiang, G. Sparse reconstruction for sar imaging based on compressed sensing. *Progr. Electromagn. Res.* **109**, 63–81. <https://doi.org/10.2528/PIER10080805> (2010).
68. Range Doppler Algorithm. <https://earth.esa.int/handbooks/asar/CNTR2-6-1-2-3.html>. Accessed 2022.
69. Potter, L. C., Ertin, E., Parker, J. T. & Cetin, M. Sparsity and compressed sensing in radar imaging. *Proc. IEEE* **98**(6), 1006–1020 (2010).
70. Nguyen, L. H., Tran, T. D. Rfi-radar signal separation via simultaneous low-rank and sparse recovery. In *2016 IEEE Radar Conference (RadarConf) 1–5* (IEEE, 2016).
71. Song, Y., Jun, Hu., Dai, Y., Jin, T. & Zhou, Z. Estimation and mitigation of time-variant rfi in low-frequency ultrawideband radar. *IEEE Geosci. Remote Sens. Lett.* **15**(3), 409–413 (2018).
72. Cetin, M., Moses, R. L. Sar imaging from partial-aperture data with frequency-band omissions. In *Algorithms for Synthetic Aperture Radar Imagery XII. International Society for Optics and Photonics*, vol. 5808, 32–43 (2005).
73. Nguyen, L., Do, T. Recovery of missing spectral information in ultra-wideband synthetic aperture radar (sar) data. In *2012 IEEE Radar Conference 0253–0256* (IEEE, 2012).

### Author contributions

A.M. performed the technical analysis and simulations. He also wrote the paper. All authors reviewed the manuscript.

### Funding

This study was funded by National Science Foundation (Grant no. 2112267).

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to N.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023