J Infrared Milli Terahz Waves (2025) 46:14
https://doi.org/10.1007/s10762-025-01033-z

RESEARCH

Using Neural Networks to Design Leaky-Wave Antennas for Terahertz Wireless Links

Joshua Neronha[1] · Hichem Guerboukha[2] · Daniel M. Mittleman[1]

Received: 15 May 2024 / Accepted: 17 January 2025 / Published online: 24 January 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract

A primary challenge of terahertz communications is the high free-space path loss experienced by THz signals, rendering omnidirectional antennas inadequate for communications needs in this spectral range. One possible solution lies in using types of antennas that emit radiation in specific directions, such as leaky-wave antennas (LWAs), in order to direct power in the exact direction of the receiver. However, for an arbitrary arrangement of receivers, designing an antenna that emits power in all of their directions while avoiding wasting power elsewhere is a significant challenge. In this paper, we propose an irregular, aperiodic design schema for LWAs that can generate emission maxima simultaneously at a variety of angles. We combine this novel antenna architecture with the power of neural networks in order to build a model that can accurately and quickly select a design from among the extremely large set of options, that will produce a desired far-field signal. Combined with the ability to perform experimental tests using rapid-prototyping methods, we propose an end-to-end solution for antenna design that allows researchers to both generate and fabricate a slot design for some chosen far-field signal within minutes, reducing the need for lengthy finite-element calculations or tedious trial-and-error experiments.

Keywords Terahertz · Antenna · Leaky-Wave Antennas · Neural Network · Machine-Learning

✉ Hichem Guerboukha
hichem.guerboukha@umkc.edu

[1] School of Engineering, Brown University, Providence, RI, USA

[2] School of Science and Engineering, University of Missouri-Kansas City, Kansas City, MO, USA

1 Introduction

The predictive power of deep learning and machine learning techniques more generally have been applied to a vast array of problems ever since the neural network was proposed as a computational technique loosely modeling the brain's architecture in the mid-twentieth century [1]. This, of course, includes communications technology: deep learning has been used extensively in this field because of its ability to approximate accurate solutions to nonlinear problems. This includes the optimization of massive multiple-input multiple-output (MIMO) systems [2], end-to-end communication using autoencoders [3], and predicting a channel state [4]. Likewise, deep learning has been extensively used in antenna design because of its strength in predicting complex, nonlinear, and multivariate relationships that underlie many problems in electromagnetics [5].

Deep learning models, a specific type of neural network categorized by multiple layers of weights and biases [6], are particularly useful in situations where analytical solutions cannot be obtained and/or numerical simulations are expensive [7, 8]. Crucially, deep neural network models do not simply "fit a curve" like regression methods but are able to uncover complex nonlinear patterns with remarkable precision. In the context of optics, these types of problems are usually split into two cases: forward and inverse problems. The former refers to situations where models attempt to generate the electromagnetic response given some parametric input, while the latter attempts the reverse: predicting those parameters given some signal [9]. Examples of direct problems include approximating an antenna's output given a geometry [10] or modeling the response induced by a metasurface [11]. Inverse problems consider the opposite direction, for example, reconstructing an image from scattered light [12].

Of particular interest here, deep learning methods have also been used extensively in the design of antennas [13]. Machine learning methods like support vector machines and neural networks have been used in antenna parameter optimization, for example in rectangular microstrip antennas based on measured values [14] and finding the optimal design parameters in Wband slotted waveguide antenna arrays [15]. Deep learning has also been used in inverse-scattering problems for antenna design [8, 16, 17], for example, to predict the proper geometry in multimode resonant antennas and triband printed antennas [9].

In this work, we are particularly interested in the inverse problem associated with the design of terahertz leaky-wave antennas (LWAs) because of their applicability to the emerging field of THz communications. With the promise of breaking the terabit-per-second (Tbps) barrier [18], THz communications ($\geq$ 100 GHz) are often seen as an important component of future 6G systems [19]. Many challenges arise when working at such high frequencies. One of them is the large free-space path loss that can reach ~80 dB over a 1-m distance at 300 GHz. To counter these losses, highly directional beams generated from high-gain antennas are required. This can be achieved, for example, using a LWA, a single-mode waveguide with an aperture to permit some of the radiation to leak out into free space. By tailoring the aperture, one can sculpt the far-field radiation pattern, even generating a narrow pencil-like

beam ideally suited for signal distribution in a high-frequency directional network [20]. This recent work emphasized the importance of designing the LWA aperture for a particular desired far-field radiation pattern. However, there is an enormous parameter space available for such designs, resulting in a challenging optimization problem. Here, we consider the following question: given some arbitrary desired far-field pattern, how can we rapidly find a LWA design that meets the specific needs of a wireless LAN operating in the THz range? This problem is well-suited for an approach based on deep learning given its nonlinear nature and the large parameter space at hand.

In this study, we focus on a particular LWA architecture based on a parallel-plate metal waveguide, which has proven to be a very effective platform for exploring LWA capabilities in the terahertz range. In the past, this type of LWA, which emits radiation with a frequency-dependent angle of emission as shown in Fig. 1a, has been shown to be valuable for diverse THz applications including link discovery [21], multiplexing and demultiplexing [22, 23], and for object detection [24–26]. We note that, with a few exceptions [27], the majority of the previous works on the design of LWAs aimed at generating a single output beam. This approach is justified at lower frequencies, where a single beam might have enough coverage in the far-field. At THz frequencies, however, the highly directional nature of the required links (angular divergence of perhaps only a few degrees), strongly suggests that high-gain multi-beam antennas will be very valuable [28–30]. In that regard, LWAs with periodic slot design are especially interesting because of the characteristic multi-beam emissions at various angles, both in the forward and backward directions [31]. And while periodic LWAs have been studied before [32], few works have been dedicated to their multi-beam application [33, 34]. To the best of our knowledge, no studies have used machine learning for leaky-wave antenna multi-beam optimization.

We introduce here a deep-learning-designed LWA that can generate a multi-beam profile using an aperiodic slot design. Specifically, we suggest an antenna design where a rectangular slot in an LWA is separated into a series of adjacent sub-slots, each of which can either be metallic, transparent, or somewhere in between (i.e., partially transparent), as shown in Figs. 1b and c. Given the large number of possible

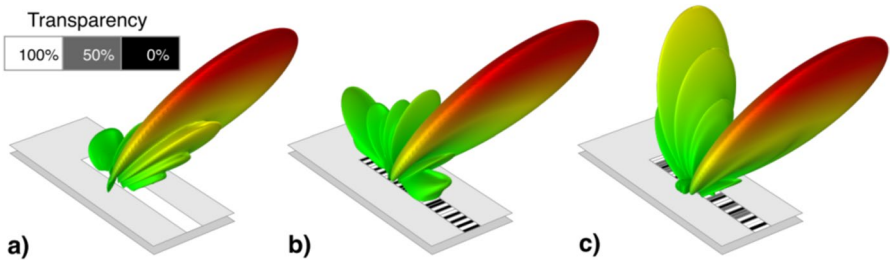

Fig. 1 Leaky-wave antenna geometries under consideration in this investigation. The antenna is a simple parallel-plate waveguide with non-uniformly periodic sub-slots. As we transition from **a** a simple open-slot LWA to **b** a periodically modulated binary or **c** grayscale waveguide, multiple emission peaks can be generated

slot configurations (and associated radiation patterns) that can be generated with such a geometry, a complete characterization of all of them is very far from practical. Instead, we develop a deep-learning-based model to predict the optimal LWA geometry for producing a given (desired) far-field pattern. This antenna geometry could realistically be fabricated with active elements, for example using electrically gated graphene patches that modulate the effective slot transparency [35, 36]. As a proof of concept, we use a hot stamping technique [37] to experimentally prototype antennas predicted by our model and compare the (static) measured radiation patterns to the desired target patterns.

This paper is organized as follows: we start by numerically and experimentally reviewing the behavior of a standard LWA with a periodically varying slot aperture before introducing two separate aperiodic schemes that can modulate the LWA to obtain desired far-field patterns. We build a deep-learning model for each scheme and examine its performance by collecting experimental measurements for designs predicted by the model. Finally, we consider shortcomings of our approach and consider how further research in this field might proceed.

2 Results and Discussion

2.1 Uniformly Periodic Leaky-Wave Antennas

As a prototype platform, we consider a LWA based on a parallel plate metal waveguide with plate separation of 1 mm, operating in the fundamental transverse electric (TE1) mode [38], with a rectangular slot in one plate such that the slot aperture varies periodically, alternating between transparent and opaque sections arranged with some periodicity Λ. According to Floquet theory [31], a periodic array of subslots excites an infinite number of spatial harmonics each with its distinct dispersion constant β_p such that:

$$\beta_p = \sqrt{k_0^2 - \left(\frac{\pi}{b}\right)^2} + \frac{2\pi p}{\Lambda} \tag{1}$$

where k_0 is the free space wave number, $p = 0, \pm 1, +2, \ldots$ is the Floquet mode number, and b is the plate separation. This result assumes that the slot width perpendicular to the wave vector of the guided wave is small, such that we can consider the electric field to be uniform across its width [20]. Each different Floquet mode produces a radiation pattern which peaks in the far field at some angle θ defined by $\theta = \cos^{-1}(\beta_p/k_0)$, conditional on the fact that the excited modes can be considered fast-wave—in other words, the mode can only leak if $\left|\beta_p\right| < k_0$ [31]. Given a periodicity Λ, and a mode number p, there may be more than one leaky mode.

We begin by performing numerical simulations with a finite element solver for a variety of uniform periodicities and plot their far-field power response in Fig. 2. In general, any mode with a mode number p that allows $\left|\beta_p\right| < k_0$ will outcouple and produce a distinct peak. For example, the $\Lambda = 1$ mm period allows the mode β_0 to

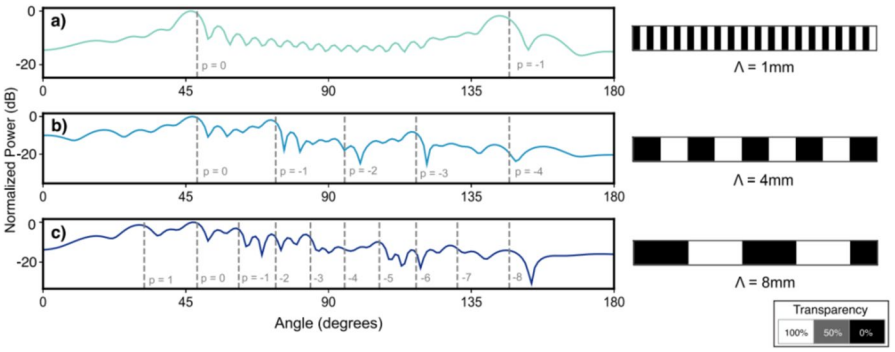

Fig. 2 Floquet theory predictions for different geometries of uniform periodicities from numerical simulations: **a** $\Lambda = 1$ mm, **b** $\Lambda = 4$ mm, and **c** $\Lambda = 8$mm. Regardless of periodicity, each geometry exhibits a main non-periodic peak around 48 degrees; as the periodicity increases from 1 to 8 mm, there are more possible Floquet modes (2 in **a** and 10 in **c**), albeit at smaller magnitudes than the primary peak at $p = 0$. We exploit these variable Floquet modes by superimposing them in an aperiodic aperture

outcouple at an angle of 48 degrees (main peak of the fundamental mode), but also the β_{-1} mode to outcouple at 147 degrees. It is also possible to output more than two peaks, as shown for example in the case of a period of $\Lambda = 4$ mm. It is worth noting that the angles can also be smaller than that of the main peak ($p = 0$) given than a positive value of p is allowed, as shown in the case of $\Lambda = 8$ mm.

These results confirm that modulating the periodicity of the slot geometry enables the generation of peaks in the far-field radiation pattern at a wide variety of angles. In the next section, we introduce an aperiodic slot design to allow more flexibility in the produced multi-beam pattern.

2.2 Binary Sub-Slot Scheme

Our goal is to design an antenna whose far-field signal has peaks at any desired set of angles. Figure 2 shows that antennas with different aperture periodicities yield peaks at different angles. As a result, we can produce an antenna with peaks at multiple angles if we superpose multiple periodicities in a single slot, which is achieved by a non-uniformly periodic slot aperture. The key to our strategy is the discretization of the slot into 36 sub-slots, each of 0.5 mm length, that can be modulated in their transparency, as illustrated in Fig. 1b. This design allows us to excite multiple Floquet modes, producing peaks at multiple angles. By Fourier decomposition, we note that any non-periodic slot design can be viewed as a linear combination of periodic designs with wave number β_p, each producing a far-field peak at an angle $\cos\theta = \beta_p / k_0$. Our discretized geometry imposes limits on the possible excited Floquet modes by constraining the values of Λ—only a finite discretized number of periodicities, and thus a finite number of emission peaks, can exist.

Incidentally, due to the fast-wave requirement (i.e., $\left|\beta_p\right| < k_0$), only a few select modes will be able to leak out of the waveguide, meaning that for a given geometry only some of the peaks will be present. To calculate the possible peak angles, we consider the dispersion relation in Eq. 1. We deduce the possible angles that satisfy the fast-wave condition given the set of possible Λ in the discretized geometry and the set of mode number p. For the example geometry considered here, there are almost 60 possible modes which correspond to about 35 distinct peaks (we ignore duplicate angles).

Figure 3 shows experimental measurements of the emitted radiation patterns from two different LWAs—one that is uniformly periodic (Fig. 3a) and one that is not (Fig. 3b)—when the waveguides are excited with broadband input signals that span the entire single-mode waveguide band for a parallel plate waveguide with a $b = 1$ mm plate spacing (150–300 GHz). The two slot geometries are illustrated schematically below the respective plots (see the "Methods" section for more details on the measurement technique). Over the spectrum, the results show that the periodic slot design produces two distinct peaks, which are highlighted by the red ($p = 0$) and black ($p = -1$) curves. This contrasts with the aperiodic case where multiple additional peaks can be observed in addition to the fundamental mode (red curve). This rich and complex radiation pattern illustrates the strategy that we will exploit to design an antenna with a multi-beam profile of our choosing. The question, then, becomes how to find the particular aperiodic antenna geometry that produces a desired far-field pattern with specified amplitudes at each possible peak location

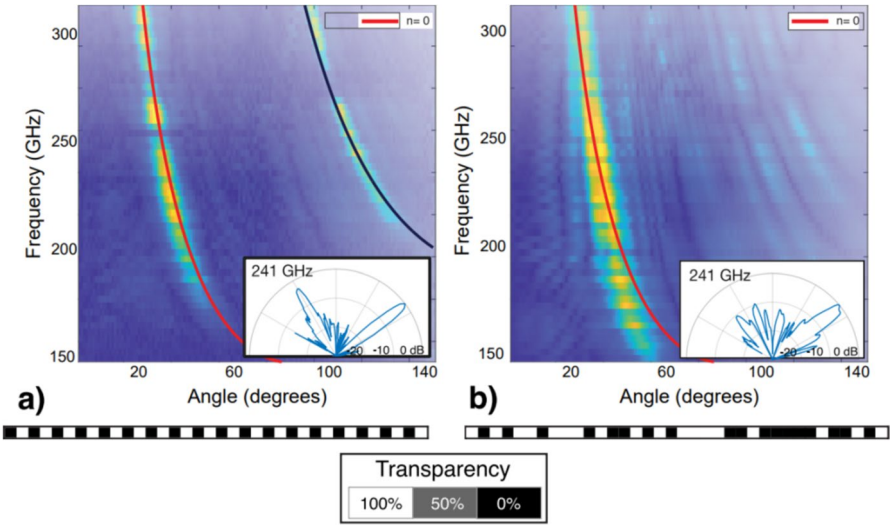

Fig. 3 Experimental time-domain spectroscopy (TDS) measurements for two different slot designs; in **a**, there is only one periodicity but **b** contains a linear combination of periodicities, evident from the non-uniform nature of the sub-slots. As a result, in **a**, we see two discrete modes, whereas in **b**, we see a wide array of lower-power modes, as we would expect from a non-uniformly periodic slot. This illustrates how we exploit periodicity in order to achieve various peaks at different angles

(angle). Solving this inverse problem, however, is difficult; deriving an analytical expression is difficult if not impossible, and solving numerically would require iteration through multiple antenna designs until arriving at an optimal solution, essentially searching for a needle in a haystack of $2^{32} \approx 4 \times 10^9$ possible antenna geometries. This is where the neural network model enters. Deep learning methods are effective at describing nonlinear relationships of complex systems simply by providing training data from which the model can "learn" [5]. Thus, we build a model that predicts the slot geometry for some input of a desired far-field signal using a neural network architecture. In order to build a model that can predict a slot for a given far-field pattern, we first must generate training data. In this work, for simplicity, we neglect the effect of frequency. Our optimization strategy is designed for a particular frequency of 200 GHz, which lies in the single-mode operation band. Using a finite element solver, we build a numerical simulation of a leaky-wave antenna with a slot of length 18 mm divided into 36 sub-slots, each of which is 0.5 mm as in Fig. 1a. The sub-slots are randomly divided between two categories (i.e., opaque or transparent) so that the number of each type of slot is equal. We generate 20,000 simulations to assemble the training data set. Each individual simulation consists of a far-field signal and the slot that generated it, serving as the inputs and labels of the neural network model. The model is trained by extracting the peaks that are predicted by Floquet theory and optimizing the weights and biases of the neural network architecture in the schematic shown in Fig. 4. We utilize a native layered neural network architecture designed after extensive testing given the unique needs of this model rather than a pre-built architecture like CvxNet. Full details of our native architecture can be found in the "Methods" section.

Once the model has been trained, we can test its performance. First, we select an arbitrary antenna configuration and compute its far-field radiation pattern using the finite element solver. Then we provide this far-field pattern to the neural network as a desired (objective) far-field response. The model extracts the Floquet peaks of the objective and uses them to predict each individual sub-slot's probability of being transparent using a sigmoid activation function. We then compare the objective pattern to the one generated by the antenna that is designed by the model. For

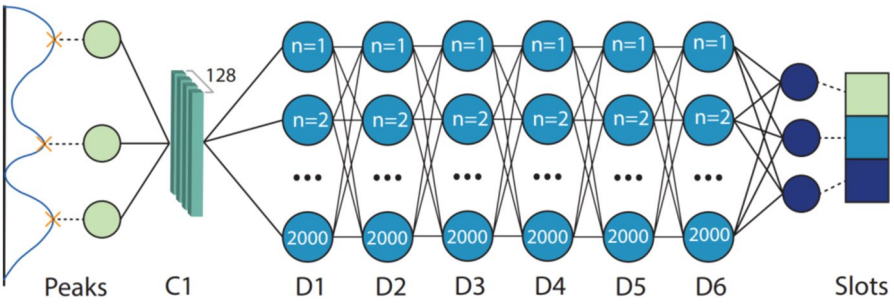

Fig. 4 Neural network architecture schematic. The amplitudes at the angle of each possible Floquet mode are used as the input data and passed through one convolutional layer and six hidden layers within a neural network architecture to predict the slot design that will yield the closest possible signal. Details on network parameters and design can be found in the "Methods" section.

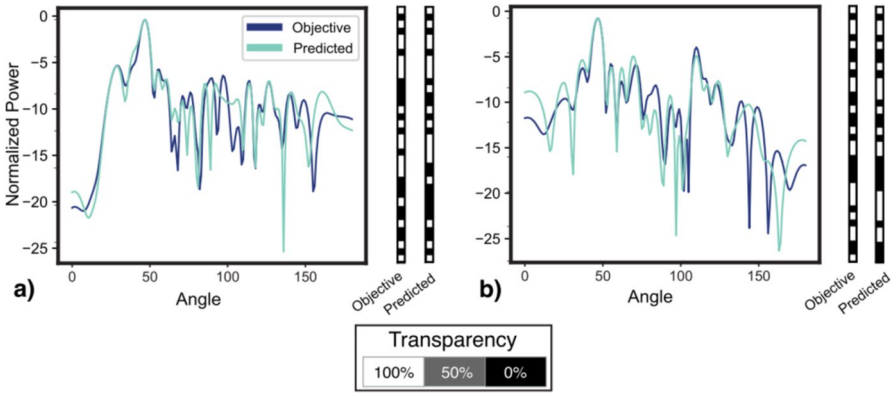

Fig. 5 Model predictions for two different peak profiles. In each panel, the blue line corresponds to a far-field signal that serves as the "objective" signal we want to replicate and is generated by the objective slot design. The green line is the signal corresponding to the predicted slot—in other words, the far-field signal created by the slot design produced by the model. The mean-square errors between the objective and predicted far-field signals of **a** and **b** are 0.0155 and 0.0250, respectively, corresponding to the first quadrant of the histogram in Fig. 6. Overall, we see relatively close agreement between the model and the ground truth

example, Fig. 5 shows the results for two particular objectives used as inputs for the model. The objective signals and the slots that created them are shown in both cases, together with the slot design predicted by the model and the far-field radiation patterns that those slots actually generate when used as an input to the FEM simulation.

We can also analyze the model's performance more generally. For 500 randomly chosen samples in the training set, we compare the objective far-field signals with the far-field signal generated by the model-predicted slot. We compute the mean square error (MSE) for each sample and plot a histogram of MSE in. This is compared to a control distribution where we compute the MSE between two randomly selected signals in the training set. The Gaussian profile of the distribution in Fig. 6a

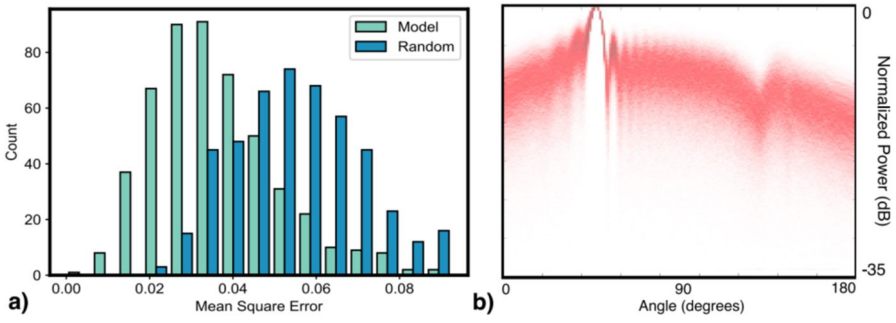

Fig. 6 **a** Mean square error (MSE) histogram. The horizontal displacement between the two histograms shows that the model predicts significantly more accurate antenna geometries than the control and is confirmed by a *Z*-test. **b** Density plot of 5000 samples across the angular range. A wide array of values are possible at each angle, indicating that the model may be effective over a wide range of far-field signals

matches that of similar work [11]. This distribution suggests that the model produces slot designs that consistently generate far-field patterns far closer to the objective than random slot selection would allow, indicating our model is useful. We also confirm this statistically by conducting a two-sided Z-test—the requirements of which are satisfied given that the histogram indicates a normal distribution and there are a sufficient number of samples—and obtain a p-value of 2.91×10^{-89}, indicating that there is a statistically significant difference between the model and random. This shows that the model does indeed hold predictive value. Figure 6b shows the power density across the angular range. We see a wide array of possible values across all angles (except for at the primary peak, where the normalized power is always at a maximum). At most angles, there is a dynamic range of approximately 20 dB, indicating that antennas can be produced with specific angular-varying amplitudes using our design schema and model. The two examples in Fig. 5 showcase the ability of the model to produce slots whose radiation patterns accurately approximate the desired far-field signal. We also wish to compare our numerical predictions to experimental observations for physical verification. We fabricate the predicted antenna geometry using hot-stamping and perform an experimental angular sweep of the far-field signal. We extract the experimental far-field signal at 200 GHz and compare to both the objective radiation pattern (i.e., what we were trying to achieve) and the expected far-field pattern (what numerical simulations predict for the slot geometry generated by the neural network model), as shown in Fig. 7. More details about the experimental setup can be found in the "Methods" section.

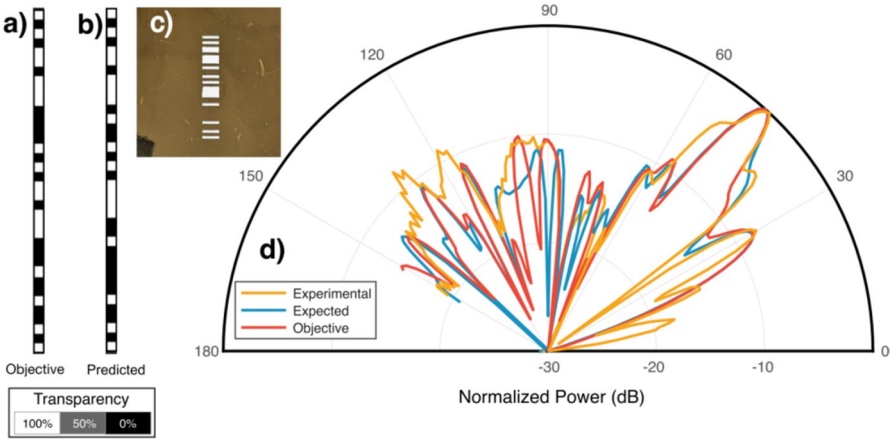

Fig. 7 Comparison of experimental and numerical results for non-uniform periodic slot. Panel **a** shows the objective slot design used to generate the objective signal, **b** shows the model-predicted slot design, and **c** shows the fabricated model-predicted slot design. In **d**, the red line corresponds to the objective far-field signal that we ask the model to replicate, the blue line represents the numerical far-field signal of the slot that the model predicts, and the orange line is the experimental measurements of the predicted slot using THz-TDS. We observe general agreement between the expected and objective results, indicating the model has suggested an appropriate slot; the experimental results also correspond well to the finite element prediction, indicating that the slot regime we have designed produces physically observable results

We draw two primary conclusions from Fig. 7. First, the slot predicted by the model generates a far-field signal which is quite similar to the objective function, indicating that the model works as expected. Second, the experimental data generally agrees with the simulation data, particularly near the largest peaks, indicating that our design schema is valid. One limiting factor at play is the relatively simple setup used to measure the experimental data, which has limited ability to measure back-scattered peaks (angles greater than 140 degrees) due to geometrical constraints. Additionally, hot-stamping is an imprecise technique that generates slot patterns with a reasonable degree of accuracy, but imperfections are clearly visible because of the nature of the method as a rapid prototyping technique. Nevertheless, these measurements serve as satisfactory proof of concept of our model and its usefulness in designing leaky-wave antennas.

2.3 Gray Sub-Slot Scheme

We now explore the possibility that the sub-slots in our LWA aperture could be partially transmitting, rather than merely opaque or transparent. This additional variable opens up many new options for the targeted far-field radiation pattern. To extend our model to encompass this possibility, we generate training data and build a model just as we did with the binary case, except that now each sub-slot can have any of three different opacity values: 0%, 50%, and 100% opacity. In other words, there is now an "in-between" option for each sub-slot, halfway between opaque and transparent, which we will call the "gray" case. The idea is that by increasing the number of possible states for each sub-slot, we increase the number of degrees of freedom by six orders of magnitude. By increasing the number of degrees of freedom, we would similarly expect a wider range of possible peak profiles the model can obtain, increasing the versatility of the model.

We again generate approximately 20,000 samples using numerical simulation where each sub-slot can take on these values using a transition boundary layer that emulates the transparency through a change in the sub-slot conductivity (see the

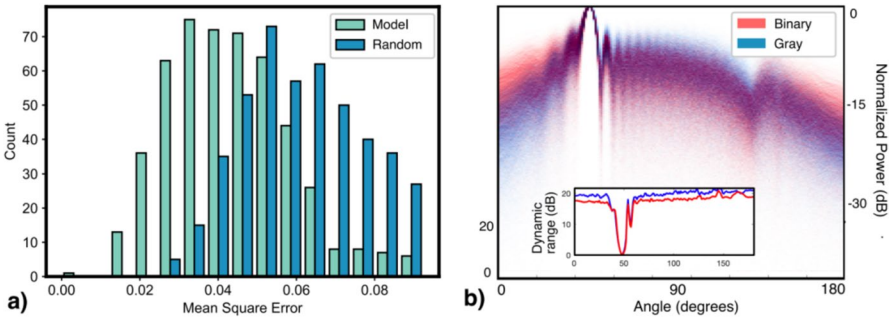

Fig. 8 **a** The mean square error histogram for both the gray model performance and a control sample. The horizontal offset validates the model's predictive value and is confirmed by a *Z*-test. **b** A density plot for both the binary and gray models. The inset shows the higher dynamic range of the gray model, measured from the 1st percentile to the 100th percentile, as a function of angle

"Methods" section for more details). We compile the model using a similar architecture as before. For 500 samples in the training set, we again compute the MSE between the objective signal and the far-field signal generated by the slot predicted by the model. We plot the histogram of this distribution in Fig. 8a.

As with the binary case, we again note the horizontal offset in the histogram between the model and control, intuitively suggesting that our model can produce useful antenna designs to produce a desired signal compared to a random control. We again verify statistically by conducting a two-sided *Z*-test between two samples and again establish a statistically significant difference between the model and random with *p*-value of $1.86 \times 10 - 72$, indicating statistical significance and thus the true predictive value of the model. The average MSE in the gray case is lower than in the binary case, which is not unexpected: due to computational constraints, we generate the same number of training samples as the binary case despite the increased degrees of freedom (3^{36} instead of 2^{36}). That means in the gray scheme, we have a smaller amount of data in relation to the number of possible outcomes, which may explain the decreased accuracy. While the accuracy is lower, the gray scheme is able to design slots for a larger range of far-field signals as shown by plot of dynamic range in the inset of Fig. 2b. A more flexible model that can account for more sorts of inputs is more useful for obtaining a specific far-field profile, and the accuracy of the model would likely improve if more data was added to the training set (although this would require more computational effort).

As before, we are interested in comparing the results of our model and numerical simulations to experimental data. We then produce our LWAs using the exact same methodology as in the binary case, except that we produce the semi-transparent slots by printing in a grayscale value calibrated to correspond to approximately 50%

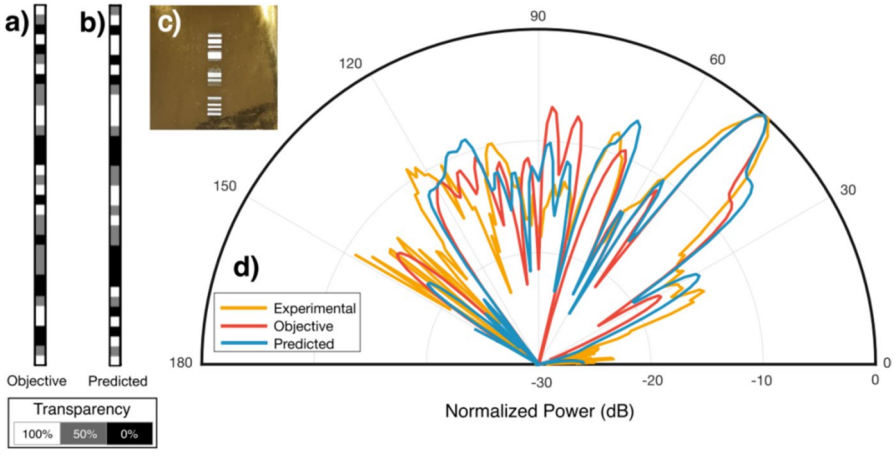

Fig. 9 An example of the gray model's performance and comparison to experimental results. The left slot design in **a** shows the slot used to produce the objective far-field signal and the right slot design in **b** shows the model's prediction and is actually fabricated in **c**. The polar plot in **d** shows the far-field signals computed numerically for each of these slots, overlaid with the experimental results collected using TDS. The numerical predicted and objective far-field signals closely agree with an MSE of 0.0311, despite slots that while sharing similarities are somewhat different, and match the experimental data well

transmission. (Additional details about the fabrication are included in the "Methods" section.)

An example of the model's predictive performance is shown in Fig. 9, along with a comparison to the experimental results for the predicted slot. As in Fig. 7, the red line represents the objective far-field signal for which we instruct the model to design a slot, the blue line is the numerical far-field signal of the slot that the model predicts, and the orange line is the experimental measurements of the predicted slot using THz-TDS. We draw several conclusions from Fig. 9. First, the predicted and objective far-field signals generally agree quite well and have very similar shapes, despite having slot patterns that are not exactly the same. It is not surprising that the slot patterns need not be identical in order to yield fairly similar far-field patterns; the combination of periodicities is what determines the signal, not their specific location. Secondly, the experimental results agree fairly closely with the numerical predicted values. While there are some differences, particularly in the backwards direction, it is important to remember that these results were collected using hot-stamping, an inherently inexact rapid-prototyping technique. Its imprecision is amplified by the fact that we can only generate an approximately 50% transparent slot subject to the whims of the laser printer Overall, the experiment confirms the validity of producing LWAs in this fashion, and we expect the experimental accuracy to be higher using more advanced methods of producing patterned metallic surfaces, including actively-modulated metasurfaces.

3 Conclusion

In conclusion, we have demonstrated a novel method of designing a periodic leaky-wave antenna that allows us to exploit Floquet theory to achieve a wide variety of far-field peak profiles. We have shown that a deep neural network can be used to predict an appropriate antenna geometry that will achieve an approximate objective far-field signal and considered two different geometry schemes that can be modulated. Finally, we have validated that the aperiodic LWA performs as we expect experimentally, and the far-field patterns produced by the model-produced geometries often agree quite well with the objective signal. While rapid prototyping methods like hot stamping are useful for quickly testing antenna design, determining how to build an antenna that will produce some given signal is still a work in progress. It is easy to conceive of some far-field signal for some specific purpose (i.e., delivering a signal to receivers at some number of different locations but avoiding eavesdroppers [39] and jammers [40] at other locations) and also to build a prototype of the antenna for rapid experimentation, but the ideas presented here begin to bridge the gap between conception and experimentation. A user with an idea of some highly specific signal might use such a model to obtain a starting point for an antenna geometry, after which techniques like hot-stamping can be used to tweak the antenna until optimal prototype performance is obtained.

There surely remain many areas of further exploration. We present a baseline model that serves as a proof-of-concept, but it could surely be further improved and optimized by generating more training data and extensively tuning the model

hyper-parameters. Additionally, we explore only one additional method of modulation: making some sub-slots partially transparent. It is easy to conceive of other ideas, such as modulating the width of each sub-slot or further increasing the number of intermediate gray steps. Frequency dependence of the radiation patterns is an important element that we neglected in this study; a more thorough model could integrate the frequency effects in its prediction results. Perhaps most interestingly, this model has been built using numerical data due to the time involved with manually fabricating each antenna and scanning the angular range. However, using actively modulated antenna slots and an automated rotation stage, one could easily imagine using experimental data as the training data for a model similar to the one considered in this paper. An experimentally-trained model would allow us to close the inherent gap that always exists between numerical and experimental data by building a model that is able to perform simulations not by performing calculations but by inferences drawn from previous experiments, which would be quite interesting indeed. Such a system would have the benefits of rapid prediction similar to finite element methods, except based on experimental results and not meshed-base physics solvers, allowing far greater flexibility for researchers modeling novel antennas. The model and procedure considered here would be quite similar in such an investigation; the only change would be a difference in the origin of the training data set.

4 Methods

4.1 Computational Methods

The numerical simulation of the leaky-wave antenna is built in the finite-element software COMSOL Multiphysics. The model consists of an air-filled metal parallel-plate waveguide with a plate separation of 1 mm. We excite the input end of the waveguide with an electric field polarized parallel to the plates to excite the fundamental TE1 mode. A rectangular slot of 18 mm wide and 3 mm long made of elementary sub-slots with a length of 0.5 mm is introduced in the top plate. To simulate metal, semi-transparent, or transparent sub-slots, perfect-electric conductor boundary, transition boundary, and scattering boundary are respectively used. The transition boundary condition in particular allows us to model a discontinuity of the tangential electric field given a finite conductivity of a geometrically thin interface. We first realized a set of simulations in transmission geometry to obtain the value of the conductivity that allows a 50% power transmission; we found 130 S/m conductivity when assuming a 40-μm thin layer. The far-field pattern is then computed using the Stratton-Chu diffraction integral that takes contribution from the semi-transparent and transparent sub-slots [20].

In order to generate training data, we automate random geometry generation and simulation using MPh [41], an open-source Python package that enables controlling numerical simulation software using its API. Basic combinatorics theory tells us that there are $> 10^9$ and $> 10^{15}$ possible slot designs for the binary and gray case respectively. These large numbers clearly indicate the usefulness of deep neural networks

in this prediction task given the sheer number of possible outputs. Using only a few tens of thousands of sets of training data, we predict an optimal slot design for an arbitrary peak profile. Each simulation takes approximately one minute to complete, meaning that generating training data is quite slow, but once the model is built it can produce an ideal slot for some far-field pattern in milliseconds, reinforcing its usefulness as a complement to rapid prototyping techniques.

For the binary model, 20,000 instances of COMSOL simulation data were generated as the data source for the deep neural network, with a test-train split of 80–20 [42]. The network, which is built with TensorFlow [43], a popular deep-learning library, has a simple convolutional deep neural network architecture. There is one convolutional layer with 128 layers and a kernel size of 5 for feature extraction followed by a pooling layer [44], which is connected to six sequential fully-connected layers, each of which has 2000 neurons and uses a LeakyReLU activation function with $\alpha = 0.1$ [45] followed by a 10% dropout layer to avoid overfitting [46]. The model uses the Adam optimizer [47], has a learning rate of 0.00005, a batch size of 64, and trains the model over 20 epochs [48]. The final layer is a dense layer of size 36 (one for each sub-slot), which predicts the output of each individual sub-slot with a sigmoid activation function indicating the probability that a given sub-slot should be transparent. A binary cross-entropy loss function is used to compute the loss in each iteration and train the model. The gray model is constructed very similarly, with the only differences being that each dense layer has 3000 neurons to account for the increased degrees of freedom, the learning rate is 0.0001, the batch size is 128, and the number of training epochs is 35.

The tricky part of the inverse problem, as we consider here, is finding an appropriate loss function—the primary objective is to generate a slot that minimizes the difference in signal compared to what is desired, but in order to use this basis as our loss function, we would need to compute the far-field signal for each new slot we generate as we train the model, which would be extraordinarily (and unfeasibly) computationally expensive. Thus, instead, we turn to the naive idea of making the generated slot look as similar to the given training slot for a given peak profile as possible, although not ideal (one can easily imagine a situation where a slot design was simply translated by 1 unit of sub-slot in either direction, from which we would expect a very similar response yet would likely score poorly in accuracy). Modified loss functions that introduced a convolutional element in order to emphasize spatial relationships and not exact locations of a sub-slot were tested but performed poorly compared to the naive loss function. For the binary case, we face a binary classification problem for each sub-slot, we use binary cross-entropy (BCE) loss [49]. For the gray case, we use mean square error (MSE) loss because we no longer have a binary option between metallic and transparent and thus use a different loss function to compute the difference between the correct and predicted slot.

However, what determines the usefulness of our model is not the similarity of the slot but the similarity of the signal it generates to the objective. In order to consider the accuracy and usefulness of our model, we use the following procedure:

1) Using the testing dataset, extract the Floquet peaks from the far-field signal and pass them through the model to obtain slot geometry predictions
2) Perform numerical simulation of the predicted slot geometries to obtain their far-field patterns
3) For each testing sample, compare the far-field response between the objective from the training set and the far-field response from the model-generated slot from Step 2 and compute MSE This procedure allows us to compare both the slot designs and the far-field signals for both the objective and model-predicted results in order to gain a fuller sense of the model's performance.

4.2 Experimental Methods

In order to test the model experimentally, we produce simple hot-stamped leaky-wave antennas [37] and use a terahertz time-domain spectroscopy (TDS) system [50] to measure the experimental far-field response, scanning both the forward and backwards scattering directions by moving the receiver along a rotation stage. The antenna designs are hot stamped—meaning that the designs are printed in black and white using a standard laser printer and then passed through a laminator with a sheet of gold foil placed on top, a process that takes only a few minutes and enables rapid prototyping of LWAs. The hot-stamped surfaces are then attached to a metal plate, followed by a 1mm spacer and another plate of metal in order to form a parallel plate waveguide. This rapidly fabricated antenna is then mounted in the center of the rotation stage and aligned with the transmitter, lens, and coupler; the receiver is attached to the rotation stage and is moved to sweep the angular response of the far-field signal. While fabricating the binary antennas is simple—black corresponds to metal and white corresponds to transparency—the gray antenna case is a slightly more complicated experimental setup because we need to actually experimentally generate semi-transparent slots. We do this by using the aforementioned hot-stamping technique, but instead of using black ink for metallic slots and the absence of ink for transparent slots, we calibrate our TDS system using "gray" slots, the idea being that changing the ink color within the grayscale spectrum will directly correlate with the amount of metal deposited on the surface of the antenna. We first calibrate shades of gray with transmission values to determine the grayscale value between 0 and 255 that yields approximately 50% transmission; we find it to be 109. We then produce our LWAs using the exact same methodology as in the binary case.

Author Contribution J. N. and H. G designed and performed the experiments. J.N. prepared the figures. J.N, H. G. and D.M.M. wrote the main manuscript text.

Funding This work was supported by the US National Science Foundation (NSF-1954780 and NSF-2211616) and the Air Force Office of Scientific Research (FA9550-221–0412). HG acknowledges support from Fonds de Recherche du Quebec – Nature et Technologies (FRQNT).

Data Availability All relevant data are available from the corresponding author upon reasonable request.

Declarations

Ethical Approval Not applicable.

Competing Interests The authors declare no competing interests.

References

1. W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," The bulletin of mathematical biophysics, vol. 5, no. 4, pp. 115–133, 1943.
2. C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive mimo csi feedback," IEEE Wireless Communications Letters, vol. 7, no. 5, pp. 748–751, 2018.
3. T. Erpek, T. J. O'Shea, Y. E. Sagduyu, Y. Shi, and T. C. Clancy, Deep Learning for Wireless Communications. Cham: Springer International Publishing, 2020, pp. 223–266.
4. C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel state information prediction for 5g wireless communications: A deep learning approach," IEEE Transactions on Network Science and Engineering, vol. 7, no. 1, pp. 227–236, 2020.
5. M. Raissi, "Deep hidden physics models: Deep learning of nonlinear partial differential equations," The Journal of Machine Learning Research, vol. 19, no. 1, pp. 932–955, 2018.
6. Y. LeCun, Y. Bengio, and G. Hinton, 2015 "Deep learning," nature. 521 7553 436–444
7. Y. Kim, "Application of machine learning to antenna design and radar signal processing: A review," in 2018 International Symposium on Antennas and Propagation (ISAP), 2018, pp. 1–2.
8. A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "Dnns as applied to electromagnetics, antennas, and propagation—a review," IEEE Antennas and Wireless Propagation Letters, vol. 18, no. 11, pp. 2225–2229, 2019.
9. L.-Y. Xiao, W. Shao, F.-L. Jin, B.-Z. Wang, and Q. H. Liu, "Inverse artificial neural network for multiobjective antenna design," IEEE Transactions on Antennas and Propagation, vol. 69, no. 10, pp. 6651–6659, 2021.
10. H. M. Yao and L. J. Jiang, "Machine learning based neural network solving methods for the fdtd method," in 2018 IEEE International Symposium on Antennas and Propagation USNC/URSI National Radio Science Meeting, 2018, pp. 2321–2322.
11. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, 2019 "Deep learning for accelerated all-dielectric metasurface design," Opt. Express. 27 20 27 523–27 535
12. Y. Sun, Z. Xia, and U. S. Kamilov, 2018 "Efficient and accurate inversion of multiple scattering with deep learning," Opt. Express. 26 11 14 678–14 688
13. H. M. El Misilmani and T. Naous, 2019 "Machine learning in antenna design: An overview on machine learning concept and algorithms," in 2019 International Conference on High Performance Computing & Simulation (HPCS). IEEE. 600–607
14. N. T. Tokan and F. Gunes, "Support vector characterisation of the microstrip antennas based on measurements," Progress In Electromagnetics Research B, vol. 5, pp. 49–61, 2008.
15. J. Tak, A. Kantemur, Y. Sharma, and H. Xin, "A 3-d-printed w-band slotted waveguide array antenna optimized using machine learning," IEEE Antennas and Wireless Propagation Letters, vol. 17, no. 11, pp. 2008–2012, 2018.
16. L. Yuan, X.-S. Yang, C. Wang, and B.-Z. Wang, "Multibranch artificial neural network modeling for inverse estimation of antenna array directivity," IEEE Transactions on Antennas and Propagation, vol. 68, no. 6, pp. 4417–4427, 2020.
17. C. Zhang, J. Jin, W. Na, Q.-J. Zhang, and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," IEEE Transactions on Microwave Theory and Techniques, vol. 66, no. 8, pp. 3781–3797, 2018.
18. T. Kurner, D. Mittleman, and T. Nagatsuma, ¨ THz Communications: Paving the Way Towards Wireless Tbps. Springer. 2022
19. M. Polese, J. M. Jornet, T. Melodia, and M. Zorzi, "Toward end-to-end, full-stack 6g terahertz networks," IEEE Communications Magazine, vol. 58, no. 11, pp. 48–54, 2020.

20. H. Guerboukha, R. Shrestha, J. Neronha, O. Ryan, M. Hornbuckle, Z. Fang, and D. M. Mittleman, "Efficient leaky-wave antennas at terahertz frequencies generating highly directional beams," Applied Physics Letters, vol. 117, no. 26, p. 261103, 2020.
21. Y. Ghasempour, R. Shrestha, A. Charous, E. Knightly, and D. M. Mittleman, "Single-shot link discovery for terahertz wireless networks," Nature Communications, vol. 11, no. 1, p. 2017, 2020. [Online]. Available: https://doi.org/10.1038/s41467-020-15761-4
22. N. J. Karl, R. W. McKinney, Y. Monnai, R. Mendis, and D. M. Mittleman, "Frequency-division multiplexing in the terahertz range using a leaky-wave antenna," Nature Photonics, vol. 9, no. 11, pp. 717–720, 2015.
23. J. Ma, N. J. Karl, S. Bretin, G. Ducournau, and D. M. Mittleman, "Frequency-division multiplexer and demultiplexer for terahertz wireless links," Nature Communications, vol. 8, no. 1, p. 729, 2017. [Online]. Available: https://doi.org/10.1038/s41467-017-00877-x
24. Y. Amarasinghe, R. Mendis, and D. M. Mittleman, "Real-time object tracking using a leaky thz waveguide," Opt. Express. vol. 28 12 17 997–18 005, Jun 2020
25. Y. Amarasinghe, H. Guerboukha, Y. Shiri, and D. M. Mittleman, "Bar code reader for the thz region," Opt. Express. 29 13 20 240–20 249 Jun 2021
26. H. Matsumoto, I. Watanabe, A. Kasamatsu, and Y. Monnai, "Integrated terahertz radar based on leaky-wave coherence tomography," Nature Electronics, vol. 3, no. 2, pp. 122–129, 2020.
27. S. Jafar-Zanjani, S. Inampudi, and H. Mosallaei, "Adaptive genetic algorithm for optical metasurfaces design," Scientific Reports, vol. 8, no. 1, p. 11040, 2018. [Online]. Available: https://doi.org/10.1038/s41598-018-29275-z
28. Y. J. Guo, M. Ansari, R. W. Ziolkowski, and N. J. G. Fonseca, "Quasi-optical multi-beam antenna technologies for b5g and 6g mmwave and thz networks: A review," IEEE Open Journal of Antennas and Propagation, vol. 2, pp. 807–830, 2021.
29. D. Headland, W. Withayachumnankul, R. Yamada, M. Fujita, and T. Nagatsuma, "Terahertz multibeam antenna using photonic crystal waveguide and luneburg lens," APL Photonics, vol. 3, no. 12, p. 126105, 2018.
30. Y. Amarasinghe, R. Mendis, R. Shrestha, H. Guerboukha, J. Taiber, M. Koch, and D. M. Mittleman, "Broadband wide-angle terahertz antenna based on the application of transformation optics to a luneburg lens," Scientific Reports, vol. 11, no. 1, p. 5230, 2021.
31. F. Xu and K. Wu, "Understanding leaky-wave structures: A special form of guided-wave structure," IEEE Microwave Magazine, vol. 14, no. 5, pp. 87–96, 2013.
32. D. R. Jackson, C. Caloz, and T. Itoh, "Leaky-wave antennas," Proceedings of the IEEE, vol. 100, no. 7, pp. 2194–2206, 2012.
33. M. Ettorre, R. Sauleau, and L. Le Coq, "Multi-beam multi-layer leaky-wave siw pillbox antenna for millimeter-wave applications," IEEE Transactions on Antennas and Propagation, vol. 59, no. 4, pp. 1093– 1100, 2011.
34. W.-Y. Park and S. Lim, "Multi-beam leaky-wave antenna: Design, analysis, and experiments," Electromagnetics, vol. 31, no. 4, pp. 247–257, 2011.
35. X.-C. Wang, W.-S. Zhao, J. Hu, and W.-Y. Yin, "Reconfigurable terahertz leaky-wave antenna using graphene-based high-impedance surface," IEEE Transactions on Nanotechnology, vol. 14, no. 1, pp. 62–69, 2015.
36. M. Esquius-Morote, J. S. Gomez-Dıaz, and J. Perruisseau-Carrier, "Sinusoidally modulated graphene leaky-wave antenna for electronic beamscanning at thz," IEEE Transactions on Terahertz Science and Technology, vol. 4, no. 1, pp. 116–122, 2014.
37. H. Guerboukha, Y. Amarasinghe, R. Shrestha, A. Pizzuto, and D. M. Mittleman, 2020 "High-volume rapid prototyping technique for terahertz metallic metasurfaces," Opt. Express. 29 9 13 806–13 814
38. R. Mendis and D. M. Mittleman, "Comparison of the lowest-order transverse-electric (te1) and transverse-magnetic (tem) modes of the parallel-plate waveguide for terahertz pulse applications," Opt. Express, vol. 17, no. 17, pp. 14 839–14 850, Aug 2009.
39. J. Ma, R. Shrestha, J. Adelberg, C.-Y. Yeh, Z. Hossain, E. Knightly, J. M. Jornet, and D. M. Mittleman, "Security and eavesdropping in terahertz wireless links," Nature, vol. 563, no. 7729, pp. 89–93, 2018.
40. R. Shrestha, H. Guerboukha, Z. Fang, E. Knightly, and D. M. Mittleman, "Jamming a terahertz wireless link," Nature Communications, vol. 13, no. 1, p. 3045, 2022.
41. J. Hennig, M. Elfner, and J. Feder, 2022 "Mph-py/mph: Mph 1.1.5,"

42. A. Racz, D. Bajusz, and K. H ´ eberger, 2021 “Effect of dataset size and ´ train/test split ratios in qsar/qspr multiclass classification,” Molecules. 26 4
43. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, ´ C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, ´ P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org
44. S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 international conference on engineering and technology (ICET). Ieee, 2017, pp. 1–6
45. V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in Icml, 2010
46. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from over-fitting,” The journal of machine learning research, vol. 15, no. 1, pp. 1929–1958, 2014.
47. D. P. Kingma and J. Ba, 2014 “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980
48. L. N. Smith, 2018 “A disciplined approach to neural network hyperparameters: Part 1–learning rate, batch size, momentum, and weight decay,” arXiv preprint arXiv:1803.09820
49. S. Mannor, D. Peleg, and R. Rubinstein, 2005 “The cross entropy method for classification,” in Proceedings of the 22nd International Conference on Machine Learning, ser. ICML ’05. New York NY USA: Association for Computing Machinery. 561–568
50. H. Guerboukha, K. Nallappan, and M. Skorobogatiy, “Toward real-time terahertz imaging,” Adv. Opt. Photon., vol. 10, no. 4, pp. 843–938, Dec 2018.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.