nature portfolio

Peer Review File

Deep-learning Enabled Generalized Inverse Design of Multi-Port Radio-frequency and Sub-Terahertz Passives and Integrated Circuits

Corresponding Author: Professor Kaushik Sengupta

This manuscript has been previously reviewed at another journal. This document only contains reviewer comments, rebuttal and decision letters for versions considered at Nature Communications.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript presents a novel approach for the inverse design of complex multi-port electromagnetic structures using deep learning. This approach is proposed to significantly expedite the design process while opening up new design spaces previously inaccessible through conventional methods.

The paper addresses a significant problem in the design of millimeter-wave and terahertz frequency circuits—namely, the complexity and resource-intensiveness of current design methodologies. The proposed deep-learning enabled generalized inverse design method is innovative in reducing design times and exploring new design spaces. The application of a deep-learning model to this domain is original and also highly relevant to current technological needs.

The manuscript effectively highlights the challenges and innovations related to the design of millimeter-wave and terahertz frequency circuits using deep learning. While the focus on the high-frequency regime is timely due to its relevance in upcoming technologies, the manuscript would benefit from a discussion on how this approach compares or could be adapted to other frequency domains, such as photonics or lower-frequency electronic circuits. Clarifying why the high-frequency regime is uniquely suited for this deep learning approach, or discussing the method's limitations and potential in other domains, would significantly enhance the paper's breadth and impact.

The manuscript presents a methodology based on convolutional neural networks (CNNs) to predict scattering and radiative properties of complex electromagnetic structures. While the use of CNNs is justified, the rationale behind the exclusive choice of this model type over other potentially suitable machine learning techniques, such as generative models or reinforcement learning (RL), is not thoroughly explained.

Generative models could potentially offer advantages in designing novel circuit layouts, and RL could be explored for its ability to iteratively improve design solutions based on performance feedback. A more comprehensive analysis comparing these methods, discussing their advantages and disadvantages in the context of electromagnetic structure design, would not only strengthen the justification for using CNNs but also illuminate potential avenues for future research. This discussion is crucial for understanding the decision-making process in the methodology and evaluating the full potential of machine learning in this innovative application.

As additional suggestions to enhance the robustness of the manuscript, consider adding:

Include case studies or simulations comparing designs generated by CNNs against those created by RL or generative models. This could empirically highlight the strengths or weaknesses of each approach in practical scenarios.

Provide a deeper theoretical analysis or simulation data that supports the specific advantages of CNNs in capturing the spatial relationships and complexities of multi-port electromagnetic structures, which might be less effectively modeled by other techniques.

Reviewer #2

(Remarks to the Author)
In the past few years, there have been extensive studies utilizing machine learning techniques for the inverse design of optical/EM structures. Among other things, these techniques are usually based on the pre-training of forward deep neural network as a fast simulator, followed by its combination with an inverse-design model. In this paper, Karahan et al. similarly demonstrate a genetic algorithm (GA)-based inverse-design method for RF/sub-THz photonic structures using deep learning-based EM simulator, and verify the results through fabrication and experiment. The manuscript is overall well-written and presents substantial results that can be of interest to researchers in the specific field. In my opinion, however, it lacks sufficient novelty in terms of the methodology to attract general readership of Nature Communications, at least in its current form.
Rather, I find it more intriguing that the authors chose the GA as an inverse-design model and that they also co-designed the circuit elements with EM structures, which were not that adequately emphasized in the manuscript. Regarding these points, I believe the authors could justify the novelty by (1) providing a detailed comparison of GA and tandem networks [Ref. 19] in terms of computation cost, accuracy, and so on. Since it is already well known that deep tandem structures beat traditional design methods, a mere comparison against "traditional optimization" is insufficient. On top of that, (2) including details of the co-design and its importance would be beneficial. In Fig. 7, for example, the reasons for exploiting asymmetrical structure and the separated amplifiers are unclear, as being not familiar with electronics. Assuming these improvements are made, on the other hand, (3) I do not think all the previous figures and descriptions are necessary; indeed, they are quite redundant. This is because I think the use of the word "generalized" or "universal" seems quite exaggerated, which merely refers to the application of method to a few different platforms. The right panel of Fig. 1a can be misleading; the authors have not shown any simultaneously radiative (f1,3) and non-radiative (f2) structure, which I initially thought was meant by "generalized." After the authors successfully enhance the novelty of the work, I would happily recommend its publication in Nature Communications.
Minor comments:
1. (1st paragraph, p.2) "The passive structures include include …" -> "The passive structures include …"
2. (3rd paragraph, p.4) I guess it will not affect the result, but strictly speaking, a test set, which is literally for the final performance measure, should not be used for choosing the best model.
3. Figure 2d would be better in Supplementary information.
4. (Fig. 5) In the table, (out of a total combination of 2N x 2N) -> probably 2^(N2) ?

Reviewer #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #4

(Remarks to the Author)
This paper presents deep-learning-based inverse design methodology for multi-port electromagnetic structures co-designed with circuits which can offer the synthesis of high-frequency on-chip passives and circuits. The paper looks valuable for millimeter-wave circuit designers and deserves publication when the following concerns and questions are addressed in the revision.

1) Difference between this paper and Ref [36] by the same authors should be emphasized in the introduction, since quite similar figure s are used for both.

2) What is the optimal size of pixelated square elements? Wavelength/100?

3) Why experimental results for antennas (Fig. 6) deviate from the design? Are there any issues associated with the proposed methodology against radiative devices?

4) Since all the design examples are below 100 GHz, use of "sub-THz" is not appropriate.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Based on the rebuttal letter and the detailed responses to both my comments and those of the other reviewers, the authors

have satisfactorily addressed the concerns raised. They have clarified their choice of CNNs over other machine learning techniques, elaborated on the distinctions between their approach and related work, and made substantial revisions, particularly in relation to the complexity of the co-design process and the integration of active devices with passive structures.

In light of these thorough clarifications and improvements, I recommend accepting the paper for publication.

Reviewer #2

(Remarks to the Author)
I appreciate the authors' effort to improve the manuscript, especially by including the design comparison with tandem network (generative AI) results. I'm now in favor of the acceptance.

Reviewer #3

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #4

(Remarks to the Author)
The authors replied to reviewer's concerns and questions and revised their manuscript almost satisfactorily. The manuscript is now ready for publication.

We would like to thank the editor and the reviewers for their comments and suggestions in improving the paper. We really appreciate the time and effort they have put in. We have addressed the reviewers' questions, and updated the manuscript accordingly.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript presents a novel approach for the inverse design of complex multi-port electromagnetic structures using deep learning. This approach is proposed to significantly expedite the design process while opening up new design spaces previously inaccessible through conventional methods.

The paper addresses a significant problem in the design of millimeter-wave and terahertz frequency circuits—namely, the complexity and resource-intensiveness of current design methodologies. The proposed deep-learning enabled generalized inverse design method is innovative in reducing design times and exploring new design spaces. The application of a deep-learning model to this domain is original and also highly relevant to current technological needs.

The manuscript effectively highlights the challenges and innovations related to the design of millimeter-wave and terahertz frequency circuits using deep learning. While the focus on the high-frequency regime is timely due to its relevance in upcoming technologies, the manuscript would benefit from a discussion on how this approach compares or could be **adapted to other frequency domains, such as photonics or lower-frequency electronic circuits.** Clarifying why the high-frequency regime is uniquely suited for this deep learning approach, or discussing the method's limitations and potential in other domains, would significantly enhance the paper's breadth and impact.

Thank you for your comments. We appreciate it.

As the reviewer points out, the relevance of the frequency regime stems from its 1) unique application regime for 6G wireless communication and sensing, 2) full on-chip integration capabilities and 3) active EM-circuit co-design that makes this distinctly different from the design methodologies pursued in photonics.

In the photonics domain, the inverse designed devices are mostly passive structures with standard waveguide interfaces that achieve a certain functionality with a tailored frequency response. In contrast, in the presented works, multiple active devices that provide gain interact closely with these passive structures, and also with each other to create novel functionalities. The complex impedances of these devices, their ability to amplify and their mutual interactions, need to be taken into account to realize an optimized end-to-end design process. Therefore, the passives are not designed separately as functional blocks, but collectively with the integrated devices—the latter may number from tens to tens of thousands depending on the complexity of circuits that are being realized, all of which can still be integrated on-chip. Second, the RF/sub-THz passive structures mostly support quasi-TEM modes with metal surfaces (realized in the back-end-of-the-line metal layers inside the chip) sandwiched within the chip dielectric. The losses per wavelength are much higher compared to photonics, and minimization is a critically important consideration. Building a robust model that captures the loss factors accurately, at the resolution of each pixel (~λ/100) is also very challenging. When compared to low-frequency RF, on the other hand, it can be pointed out that at sub-THz and THz frequencies, the passives are small enough that they can be integrated on-chip directly with the active devices, allowing these complex systems to be fully integrated.

We have added this in the manuscript in the Introduction section.

The manuscript presents a methodology based on convolutional neural networks (CNNs) to predict scattering and radiative properties of complex electromagnetic structures. While the use of CNNs is justified, the rationale behind the exclusive choice of this model type over other potentially suitable machine learning techniques, **such as generative models or reinforcement learning (RL), is not thoroughly explained.** Generative models could potentially offer advantages in designing novel circuit layouts, and RL could be explored for its ability to iteratively improve design solutions based on performance feedback. A more comprehensive analysis comparing these methods, discussing their advantages and disadvantages in the context of electromagnetic structure design, would not only strengthen the justification for using CNNs but also illuminate potential avenues for future research. This discussion is crucial for understanding the decision-making process in the methodology and evaluating the full potential of machine learning in this innovative application.

Thank you for your comments. The paper demonstrates a CNN based forward model for prediction of scattering parameters of arbitrary shaped passive EM structures, and an approach to inverse synthesis utilizing this forward model. One can use this forward model with a generative AI framework or with RL approach for synthesis. Both of these methods relate to the design, and not to the modeling aspect of it, and therefore, can be used with the CNN-based forward model. Utilizing the developed forward model for such synthesis will reduce the training time.

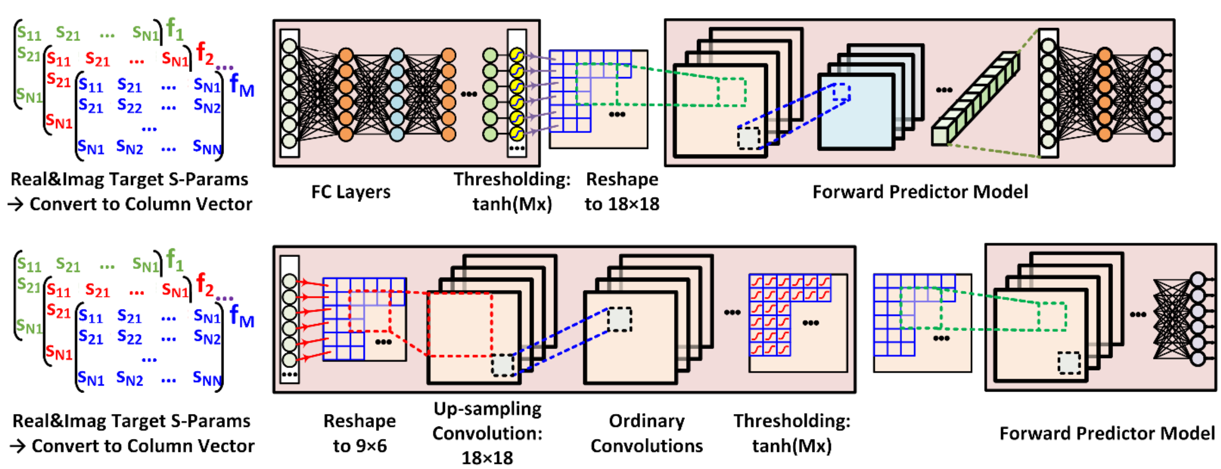

Fig. 1. Example architecture for inverse synthesis of EM structures with a generative AI framework.

However, the reviewer brings out an interesting point compare generative AI frameworks against the optimization based methods we present here. In our prior work [26], we have shown a tandem neural network approach, inspired from a variational autoencoder approach, to allow synthesis of single-port antennas. Extending to much more complex multi-port structures that interact with active devices, that can be radiative and/or non-radiative, however, requires far more extensive generalization capabilities. This might require considering physics-informed architecture. Here, we show some examples just to demonstrate the challenges of the classical generative framework. In Fig. 1, we show two exemplar inverse generative networks for 16x16 pixelated structures across 30-100 GHz.

While every EM structure has unique S-parameters, there can be multiple EM structures that have the same S-parameters. To avoid the non-uniqueness property and failure to converge, we apply the tandem approach where, we freeze the forward model. As shown in the Fig. 1, here the input vectors are the desired real and imaginary parts of S-parameters. In the intermediate layer, we have the output of inverse network, which is an 18x18 matrix denoting the pixelated structure. S-Parameters of this 18x18 matrix are then predicted by the forward model. Error is calculated as the difference between desired and achieved S-Parameter terms. The weights of the inverse network are updated via back propagation using this error.

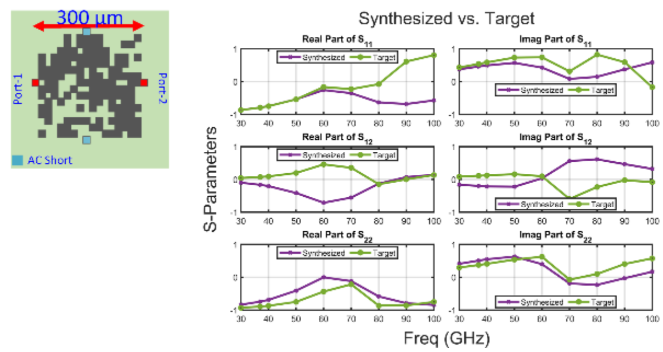

Fig.2. Successful synthesis of the EM structure with the generative framework.

In the first architecture, this vector is processed through fully connected (FC) layers. Eventual output from FC layers is a vector of length 324, which is passed through the thresholding function. This vector is then reshaped to 18x18 matrix, that is the expected input size for the forward predictor model. Lastly, port and AC short locations were assigned at fixed locations to avoid multiple locations being assigned as ports.

The second architecture performs reshaping of input vector at the first step. This was then converted to 18x18 using up sampling convolution operation. Subsequent convolution layers are processes the 18x18 structure. Lastly, element-wise thresholding operation is conducted using the thresholding function.

An important challenge is to avoid predictive neural network getting 'tricked' by the generative inverse network. For example, generative network can output an 18x18 matrix consisting of non-discrete values, which is not representative of a pixelated grid type of structure. While a hard thresholding function can ensure that output of the inverse network is solely

Inverse Network Outputs to Replicate GA Results

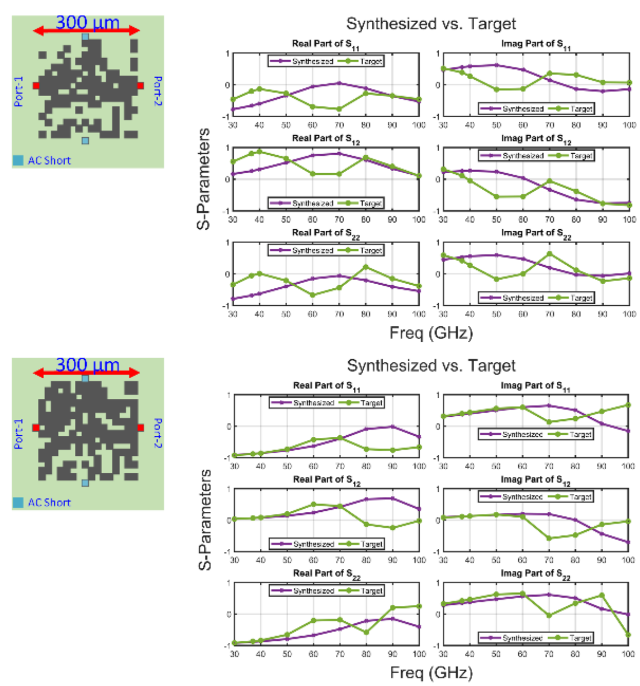

Fig.3. Inability to synthesize the S-parameters by the generative framework. The structures presented were created successfully by the optimization approaches using the CNN-based forward model.

1s and 0s, performing back-propagation derivative of hard threshold function can be challenging. Moreover, in the initial training steps, it is plausible to relax the discrete output requirement to let inverse network get trained. We approached this by changing the value of ‘M’ in the thresholding function at every 10 epochs. On the other hand, during the deployment of inverse network, we performed hard thresholding. Of the two architect, the first one gave more plausible results. Below we plot some test entries as input, resulting inverse network output and simulation result of this output.

We show some examples of synthesis. Fig. 2 shows an example of a band stop filter that was synthesized with the generative framework. The synthesized structure exhibits the desired response, showing successful synthesis. We show more examples, synthesized with the presented optimization methods, where unfortunately, generative farmwork fails. In Fig. 3 and Fig. 4, while inverse network tries to generate a structure that attempts to approximate the desired S-Parameter input, the achieved results often are unsatisfactory. It can be noted that these structures were all synthesized with optimization approach, and it shows some challenges in using the inverse neural network architecture. Finally, below is the overall error statistics that was performed with ~1200 EM simulation result compared with targeted S-Parameters:

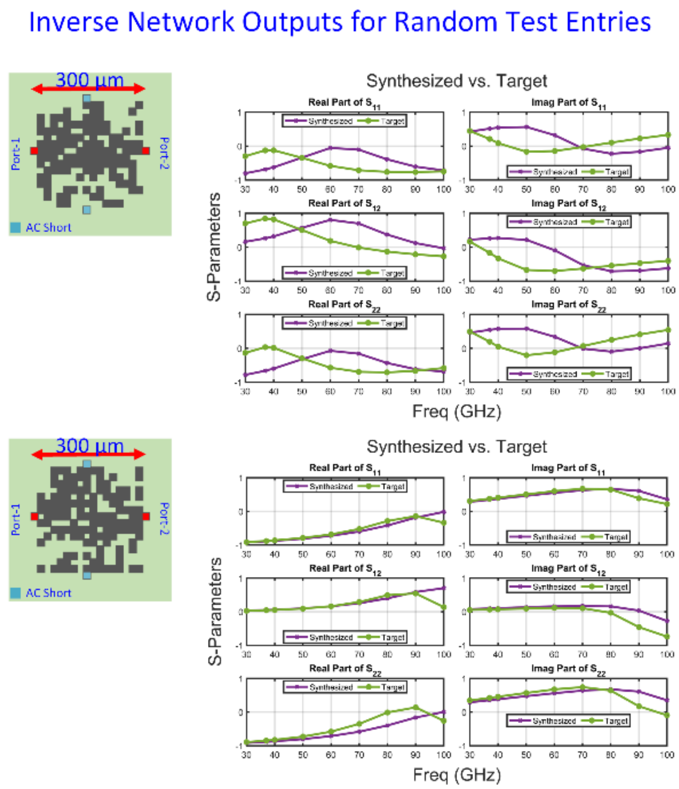

Fig. 4. Failure to generalize results in the generative framework failing to synthesized the structures.

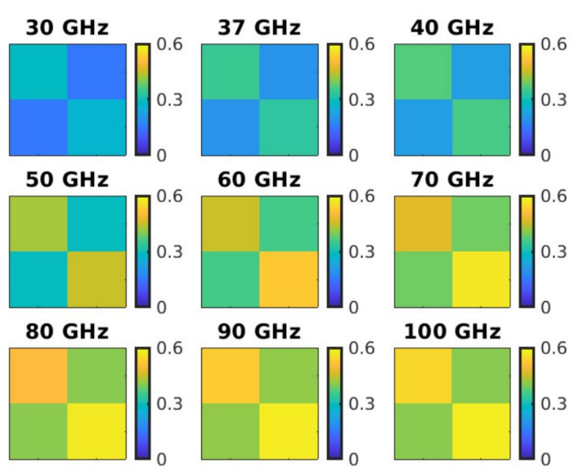

Fig. 5. High error in synthesis of randomized set of multi-port EM structures with desired S-parameters through the tandem approach.

We can observe that inverse network is not able to generalize well. While inverse network attempts to approximate the desired result, due to poor generalization ability, results fall short of target functionality.

In general, generative networks are a lot more data intensive to train and prone to mode collapses and training stagnation. We should note that generative AI models, such as transformers and diffusion models are notoriously more data and training intensive, but can generalize better. This is an area we would like to explore.

We have added this in the main manuscript under Generalized Inverse Synthesis and also in the Supplementary material.

As additional suggestions to enhance the robustness of the manuscript, consider adding:

Include case studies or simulations comparing designs generated by CNNs against those created by RL or generative models. This could empirically highlight the strengths or weaknesses of each approach in practical scenarios.

Thank you for the comment. We have addressed this in the previous comment.

Provide a deeper theoretical analysis or simulation data that supports the specific advantages of CNNs in capturing the spatial relationships and complexities of multi-port electromagnetic structures, which might be less effectively modeled by other techniques.

Thank you for the comment. An extensive comparison of all neural network architectures is beyond the scope of this paper. We can briefly summarize these and state their relevancy to our work. CNNs can automatically learn and extract features. This is in contrast to classical machine learning tools (such as regression methods and support vector machines) where a set of pre-defined and pre-extracted features are needed for training and inference. We should note that since our approach aims to move away from an intuition based parameterized designs, providing a set of pre-determined features would be in fact counteracting this goal.

Unlike FC networks, connections in CNNs are local, which speeds up convergence and reduces number of learnable parameter. This also puts an emphasis on the localized features, which can be spread over different locations of the input image. Since convolution operations are shift invariant scanning over the entire image with convolutional filters captures these features. The relative location of different features within an electromagnetic structure is the relevant analogy to an image processing context. As an example, with a shift invariant feature detector, relative distance between an excitation port and AC short location can be detected regardless of their absolute location in the image. Therefore, such geometric properties of convolutional networks supplements their applicability to physics problems where structures with equivalent geometries are frequently encountered.

We have added this in section Results.

Reviewer #2 (Remarks to the Author):

In the past few years, there have been extensive studies utilizing machine learning techniques for the inverse design of optical/EM structures. Among other things, these techniques are usually based on the pre-training of forward deep neural network as a fast simulator, followed by its combination with an inverse-design model. In this paper, Karahan et al. similarly demonstrate a genetic algorithm (GA)-based inverse-design method for RF/sub-THz photonic structures using deep learning-based EM simulator, and verify the results through fabrication and experiment. The manuscript is overall well-written and presents substantial results that can be of interest to researchers in the specific field. In my opinion, however, it lacks sufficient novelty in terms of the methodology to attract general readership of Nature Communications, at least in its current form.

Rather, I find it more intriguing that the authors chose the GA as an inverse-design model and that they also co-designed the circuit elements with EM structures, which were not that adequately emphasized in the manuscript. Regarding these points, I believe the authors could justify the novelty by (1) providing a detailed comparison of GA and tandem networks [Ref. 19] in terms of computation cost, accuracy, and so on. Since it is already well known that deep tandem structures beat traditional design methods, a mere comparison against "traditional optimization" is insufficient.

Thanks a lot for your comments. Please allow us to address the novelty of the work and distinguish from previous works that have focused on inverse designs of EM structures particualry in the photonics domain such as in [21,22].

In the photonics domain, the inverse designed devices are mostly passive structures with standard waveguide interfaces that achieve a certain functionality with a tailored frequency response. In contrast, in the presented works, multiple active devices that provide gain interact closely with these passive structures, and also with each other to create novel functionalities. The complex impedances of these devices, their ability to amplify and their mutual interactions, need to be taken into account to realize an optimized end-to-end design process. Therefore, the passives are not designed separately as functional blocks, but collectively with the integrated devices---the latter may number from tens to tens of thousands depending on the complexity of circuits that are being realized, all of which can still be integrated on-chip. Second, the RF/sub-THz passive structures mostly support quasi-TEM modes with metal surfaces (realized in the back-end-of-the-line metal layers inside the chip) sandwiched within the chip dielectric. The losses per wavelength are much higher compared to photonics, and minimization is a critically important consideration. Building a robust model that captures the loss factors accurately, at the resolution of each pixel (~λ/100) is also very challenging. When compared to low-frequency RF, on the other hand, it can be pointed out that at sub-THz and THz frequencies, the passives are small enough that they can be integrated on-chip directly with the active devices, allowing these complex systems to be fully integrated.

Let us comment on the synthesis aspect. The paper demonstrates a CNN based forward model for prediction of scattering parameters of arbitrary shaped passive EM structures, and an approach to inverse synthesis utilizing this forward model. One can use this forward model with a generative AI framework or with RL approach for synthesis. Both of these methods relate to the design, and not to the modeling aspect of it, and therefore, can be used with the CNN-based forward model. Utilizing the developed forward model for such synthesis will reduce the training time.

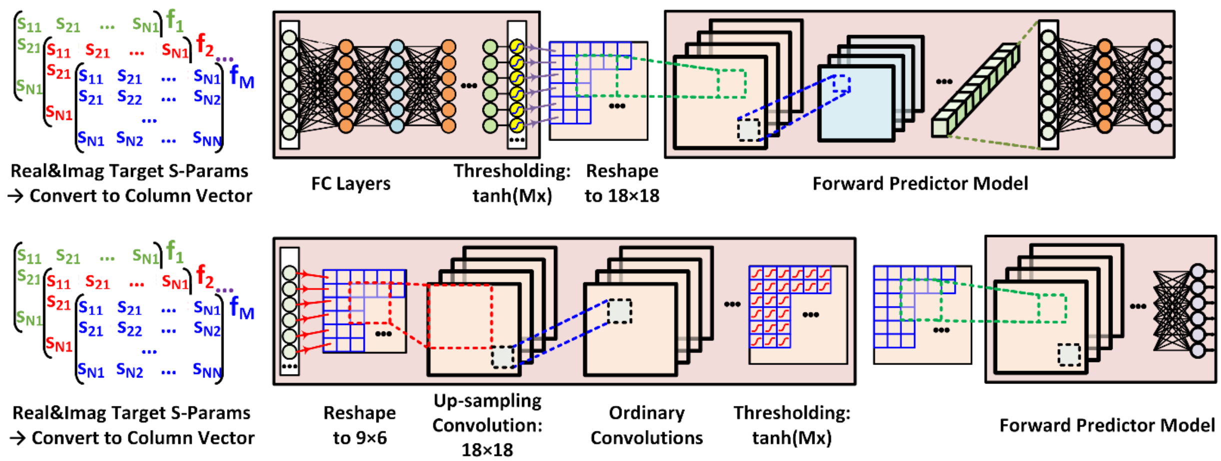

Fig. 6. Example architecture for inverse synthesis of EM structures with a generative AI framework.

However, the reviewer brings out an interesting point compare generative AI frameworks against the optimization based methods we present here. In our prior work [26], we have shown a tandem neural network approach, inspired from a variational autoencoder approach, to allow synthesis of single-port antennas. Extending to much more complex multi-port structures that interact with active devices, that can be radiative and/or non-radiative, however, requires far more extensive generalization capabilities. This might require considering physics-informed architecture. Here, we show some examples just to demonstrate the challenges of the classical generative framework. In Fig. 6, we show two exemplar inverse generative networks for 16x16 pixelated structures across 30-100 GHz.

While every EM structure has unique S-parameters, there can be multiple EM structures that have the same S-parameters. To avoid the non-uniqueness property and failure to converge, we apply the tandem approach where, we freeze the forward model. As shown in the Fig. 6, here the input vectors are

the desired real and imaginary parts of S-parameters. In the intermediate layer, we have the output of inverse network, which is an 18x18 matrix denoting the pixelated structure. S-Parameters of this 18x18 matrix are then predicted by the forward model. Error is calculated as the difference between desired and achieved S-Parameter terms. The weights of the inverse network are updated via back propagation using this error.

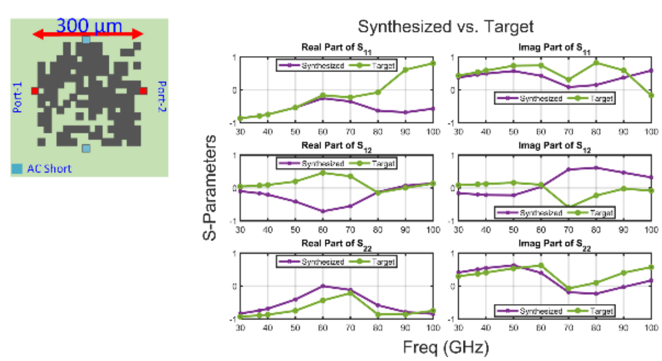

Fig. 7. Successful synthesis of the EM structure with the generative framework.

In the first architecture, this vector is processed through fully connected (FC) layers. Eventual output from FC layers is a vector of length 324, which is passed through the thresholding function. This vector is then reshaped to 18x18 matrix, that is the expected input size for the forward predictor model. Lastly, port and AC short locations were assigned at fixed locations to avoid multiple locations being assigned as ports.

The second architecture performs reshaping of input vector at the first step. This was then converted to 18x18 using up sampling convolution operation. Subsequent convolution layers are processes the 18x18 structure. Lastly, element-wise thresholding operation is conducted using the thresholding function.

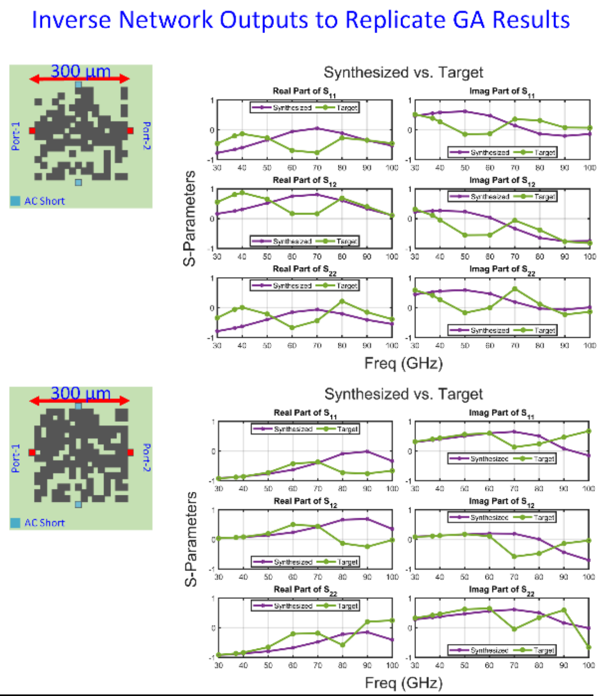

Fig. 8. Inability to synthesize the S-parameters by the generative framework. The structures presented were created successfully by the optimization approaches using the CNN-based

An important challenge is to avoid predictive neural network getting ‘tricked’ by the generative inverse network. For example, generative network can output an 18x18 matrix consisting of non-discrete values, which is not representative of a pixelated grid type of structure. While a hard thresholding function can ensure that output of the inverse network is solely 1s and 0s, performing back-propagation derivative of hard threshold function can be challenging. Moreover, in the initial training steps, it is plausible to relax the discrete output requirement to let inverse network get trained. We approached this by changing the value of ‘M’ in the thresholding function at every 10 epochs. On the other hand, during the deployment of inverse network, we performed hard thresholding. Of the two architect, the first one gave more plausible results. Below we plot some test entries as input, resulting inverse network output and simulation result of this output.

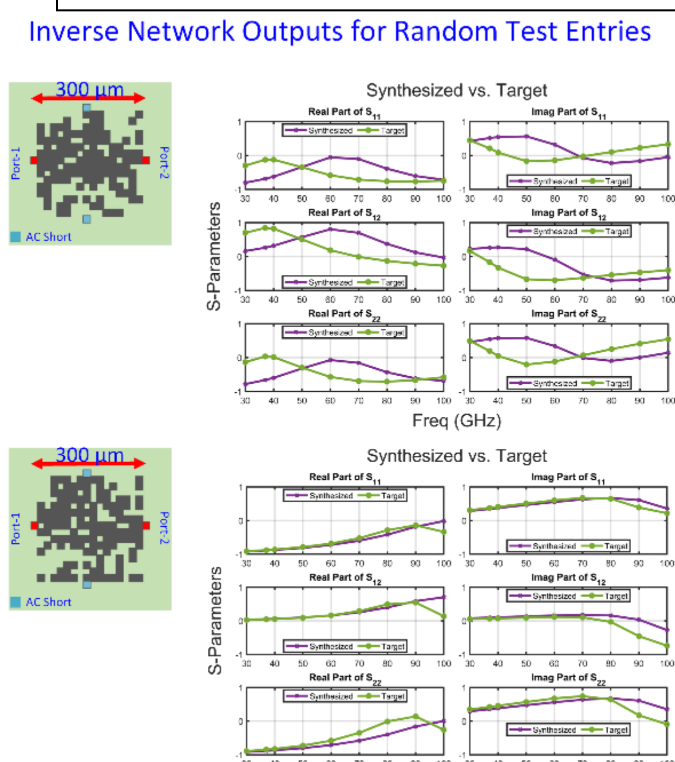

Fig. 9. Failure to generalize results in the generative framework failing to synthesized the structures.

We show some examples of synthesis. Fig. 7 shows an example of a band stop filter that was synthesized with the generative framework. The synthesized structure exhibits the desired response, showing successful synthesis. We show more examples, synthesized with the presented optimization methods, where unfortunately, generative farmwork fails. In Fig. 8 and Fig. 9, while inverse network tries to generate a structure that attempts to approximate the desired S-Parameter input, the achieved results often are unsatisfactory. It can be noted that these structures were all synthesized with optimization approach, and it shows some challenges in using the inverse neural network architecture. Finally, below is the overall error statistics that was performed with ~1200 EM simulation result compared with targeted S-Parameters:

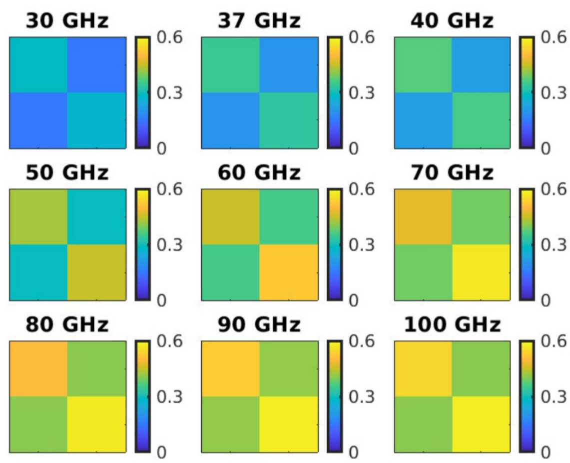

Fig. 10. High error in synthesis of randomized set of multi-port EM structures with desired S-parameters through the tandem approach.

We can observe that inverse network is not able to generalize well. While inverse network attempts to approximate the desired result, due to poor generalization ability, results fall short of target functionality.

In general, generative networks are a lot more data intensive to train and prone to mode collapses and training stagnation. We should note that generative AI models, such as transformers and diffusion models are notoriously more data and training intensive, but can generalize better. This is an area we would like to explore.

We have added this in the main manuscript under Generalized Inverse Synthesis and also in the Supplementary material.

On top of that, (2) **including details of the co-design and its importance would be beneficial**. In Fig. 7, for example, **the reasons for exploiting asymmetrical structure and the separated amplifiers are unclear, as being not familiar with electronics**.

Thank you for the comment. Detailed analysis for asymmetric structures and its bandwidth extension properties were elaborated in our previous work [16]. To facilitate understanding, we take the reviewer's suggestions and we summarize this in Supplementary section for the implemented PA.

Power combining is often needed in power amplifiers to generate high power. A classical approach to designing such power combining architecture is shown in Fig. S7 (a). The input signal gets split into two identical paths, gets amplified, and the output power gets combined symmetrically at the output port. The input splitter and output combiner are co-deigned with the active devices to ensure optimal matching for maximum power transfer. One key observation is that the two paths being identical, all the voltage nodes at each of the two identical paths is the same. As a result, in spite of having a second order network in each path of the combiner (i.e. a fourth order network in total), the combiner acts as a second-order network and therefore, limits the bandwidth.

In [16] we demonstrated an approach where we break the symmetry of the two paths as shown in Fig S7(b). The input splitter and output combiner in each of the two paths are not identical. Therefore, the voltages, currents at each nodes in the two paths are not identical, and the two paths cannot collapse into an equivalent one path. This results in a quasi-fourth order behavior at the input and output, which results in a broader bandwidth compared to the symmetrical case. The two different paths, however, need to be designed in such a way that the input signals, while splitting into two identical paths, experiences the same exact gain and collective phase shifts as they finally combine at the output. When such a phase condition is maintained (Fig.S7(c)), the bandwidth can be extended significantly without additional losses. However, there is no systematic approach to realizing these multi-port networks with asymmetric frequency-dependent S-parameters.

In this work, the AI-model and the synthesis approach determines this by itself. As shown in Fig. S8, the synthesis generates the asymmetric input and output networks while matching the total collective phase shifts between the paths. We pursue an active circuit-EM co-design to enable this. Fig.S8 provides detailed schematic of the multi-port mm-wave amplifier that was implemented in 90 nm BiCMOS process. For the optimization of output combiner, frequency dependent RC values corresponding to optimum termination impedances were calculated and stored in a lookup table. During the optimization, these parasitic capacitances (arising from the device and output pad) were lumped to the CNN predicted S-Parameters of the pixelated structure. Total transmissivity across the target frequencies were then calculated by assuming optimum phase and equal amplitude excitation. As a secondary objective, presence of a DC path between the input ports and AC shorting location was ensured.

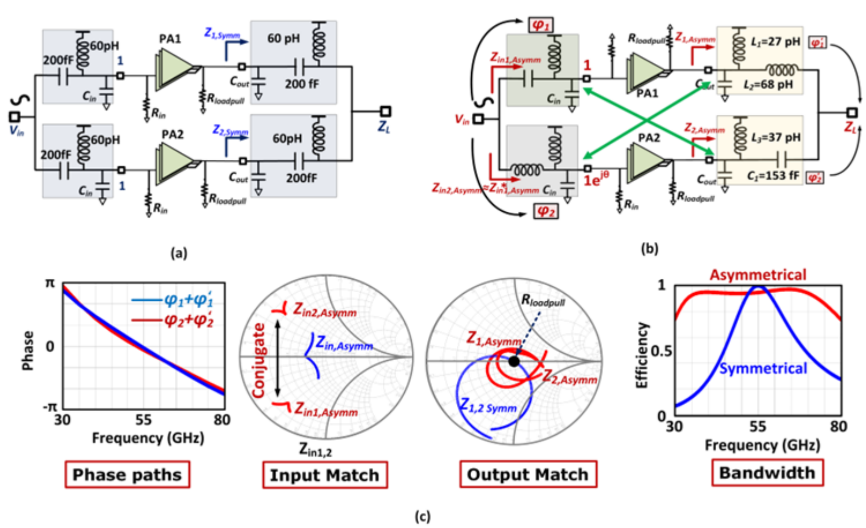

Fig. 11. Asymmetrical power amplifiers and its bandwidth extension compared to symmetric structures.

Assuming these improvements are made, on the other hand, (3) I do not think all the previous figures and descriptions are necessary; indeed, they are quite redundant. This is because **I think the use of the word "generalized" or "universal" seems quite exaggerated, which merely refers to the application of method to a few different platforms**. The right panel of Fig. 1a **can be misleading; the authors have not shown any simultaneously radiative (f1,3) and non-radiative (f2) structure, which I initially thought was meant by "generalized."** After the authors successfully enhance the novelty of the work, I would happily recommend its publication in Nature Communications.

Thank you for the comment. We use the word 'generalized' to emphasize that the method indeed is capable of synthesizing multi-port radiative and non-radiative structures with designer spectral and radiative properties. This is the first time such a methodology has been demonstrate to encompass the broad swathe of EM properties. We demonstrated this capability, as elaborated in this paper, both in simulation with several design examples, as well as in measurements with several fabricated examples and an end-to-end mmWave IC. In principle, achieving this simultaneously is possible as well, if it is so desired. The method makes no distinction between radiative and non-radiative structure that way.

Minor comments:

1. (1st paragraph, p.2) "The passive structures include include …" -> "The passive structures include …"

Thank you. We have corrected this.

2. (3rd paragraph, p.4) I guess it will not affect the result, but strictly speaking, a test set, which is literally for the final performance measure, should not be used for choosing the best model.

Thank you. The training is course done without the test data. We choose the best model after training based on it .

3. Figure 2d would be better in Supplementary information.

Thank you. We have kept it in Fig. 2D since it captures statistical error of random test structures, while the rest are individual examples.

4. (Fig. 5) In the table, (out of a total combination of 2N x 2N) -> probably 2^(N2) ?

Thank you. We have corrected this.

Reviewer #3 (Remarks to the Author): I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Thank you.

Reviewer #4 (Remarks to the Author): This paper presents deep-learning-based inverse design methodology for multi-port electromagnetic structures co-designed with circuits which can offer the synthesis of high-frequency on-chip passives and circuits. The paper looks valuable for millimeter-wave circuit designers and deserves publication when the following concerns and questions are addressed in the revision.

Thank you for your comments.

1) Difference between this paper and Ref [36] by the same authors should be emphasized in the introduction, since quite similar figure s are used for both.

Thank you for the comment. In [36](now [18]), we demonstrate synthesis of passive non-radiative two-port structures. Here, we present a universal approach for synthesis of arbitrary-shaped planar multi-port RF/sub-THz EM with arbitrary scattering and radiative properties. While there are pre-fixed templates for two-ports, there are no universal and generalized approach to synthesis for such complex multi-port high-frequency circuits. This opens up a new dimension of active RF circuits that could not be explored before.

We also compare our work with the inverse designed work in photonics. In the photonics domain, the inverse designed devices are mostly passive structures with standard waveguide interfaces that achieve a certain functionality with a tailored frequency response. In contrast, in the presented works, multiple active devices that provide gain interact closely with these passive structures, and also with each other to create novel functionalities. The complex impedances of these devices, their ability to amplify and their mutual interactions, need to be taken into account to realize an optimized end-to-end design process. Therefore, the passives are not designed separately as functional blocks, but collectively with the integrated devices---the latter may number from tens to tens of thousands depending on the complexity of circuits that are being realized, all of which can still be integrated on-chip. Second, the RF/sub-THz passive structures mostly support quasi-TEM modes with metal surfaces (realized in the back-end-of-the-line metal layers inside the chip) sandwiched within the chip dielectric. The losses per wavelength are much higher compared to photonics, and minimization is a critically important consideration. Building a robust model that captures the loss factors accurately, at the resolution of each pixel (~λ/100) is also very challenging. When compared to low-frequency RF, on the other hand, it can be pointed out that at sub-THz and THz frequencies, the passives are small enough that they can be integrated on-chip directly with the active devices, allowing these complex systems to be fully integrated.

We have added this in the introduction section.

2) What is the optimal size of pixelated square elements? Wavelength/100?

Thank you for the comment. In our previous work we investigated the effect of pixel size more in depth. While it is beneficial to enhance pixel count (hence smaller pixels) it comes as a cost for the training. Ultimately this choice depends on resources of the designer. In this sense, it is hard to find an optimal size since the amount of pixels is a trade-off between flexibility of synthesis and simulation requirements. Below is the relevant plot from our previous work where we compare matching network efficiency for different pixels (a) and number of simulations for different pixels (b):

3) Why experimental results for antennas (Fig. 6) deviate from the design? Are there any issues associated with the proposed methodology against radiative devices?

Thank you for the comment. For the antenna manufacturing, secondary effects, such as copper roughness and dielectric loss tangent become inflectional, especially in mmWave frequencies. Furthermore, dielectric properties of the prepreg are not as tightly controlled. These factors can cause slight deviations in the resonance frequency and incur extra losses. In comparison, we can show an edge feed mmWave antenna that was manufactured without prepreg. As shown below, measured and simulated S_{11} matches closely.

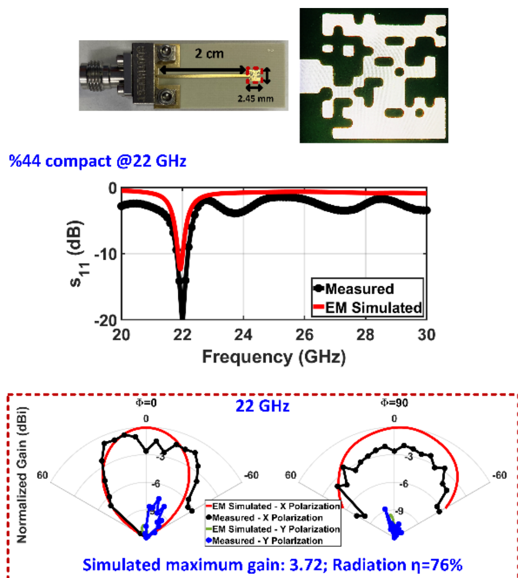

Fig. 12 Measured results of inverse designed antennas.

Second, we use a bulky end-launch connector to feed the antennas. Even though we design the microstrip feed lines going to antenna 2cm long, in order to isolate bulky connector from the near field, radiation pattern measurements are still influenced by this metal reflector. Also, due to the low frequency RF fields, complete shielding from all scattering structures was challenging.

4) Since all the design examples are below 100 GHz, use of "sub-THz" is not appropriate.

Thank you for the comment. The methods are easily translatable across the entire RF to THz frequency range. To demonstrate this we show filter synthesis examples for D-Band (140 GHz). We must emphasize that for the training of D-Band structures, we utilized transfer learning from an already existing forward model. As a result, relatively low number of simulations (25K in this case) was sufficient for training. In these examples, target is to pass the single frequency while rejecting other points as much as possible. We have added this in the Supplementary Section.

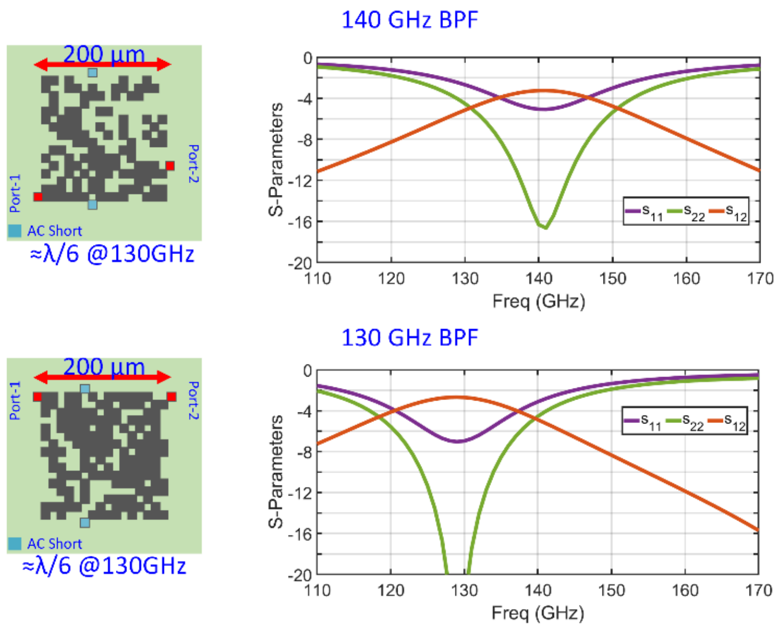

Fig. 13 Inverse designed sub-THz structures.