Deep-Learning Architecture-Based Approach for 2-D-Simulation of Microwave Plasma Interaction

Mihir Desai, Pratik Ghosh, *Student Member, IEEE*, Ahlad Kumar, *Senior Member, IEEE*, and Bhaskar Chaudhury

***Abstract*— This article presents a convolutional neural network (CNN)-based deep-learning (DL) model, inspired from UNet with a series of encoder and decoder units with skip connections, for the simulation of microwave–plasma interaction. The microwave propagation characteristics in complex plasma medium pertaining to transmission, absorption, and reflection primarily depend on the ratio of electromagnetic (EM) wave frequency and electron plasma frequency, and the plasma density profile. The scattering of a plane EM wave with fixed frequency (1 GHz) and amplitude incident on a plasma medium with different Gaussian density profiles (in the range of $1 \times 10^{17} - 1 \times 10^{22}$ m^{-3}) have been considered. The training data associated with microwave–plasma interaction has been generated using 2-D finite-difference time-domain (FDTD)-based simulations. The trained DL model is then used to reproduce the scattered electric field values for the 1-GHz incident microwave on different plasma profiles with an error margin of less than 2%. We propose a complete DL-based pipeline to train, validate, and evaluate the model. We compare the results of the network, using various metrics like structural similarity index metric (SSIM) index, average percent error, and mean square error, with the physical data obtained from well-established FDTD-based EM solvers. To the best of our knowledge, this is the first effort toward exploring a DL-based approach for the simulation of complex microwave–plasma interaction. The DL technique proposed in this work is significantly fast when compared to the existing computational techniques and can be used as a new, prospective, and alternative computational approach for investigating microwave–plasma interaction in a real-time scenario.**

***Index Terms*— Computational electromagnetics (CEM), convolutional neural network (CNN), deep learning (DL), finite-difference time-domain (FDTD), microwave–plasma interaction, multiphysics, structural similarity index metric (SSIM).**

I. Introduction

MICROWAVE–PLASMA interaction has remained one of the widely researched domains for several decades due to its various applications in electromagnetic (EM) reflectors and absorbers [1], plasma antennas [2], plasma stealth technology [3], plasma metamaterials [4], plasma-based limiters [5], switching and protection [6], [7], [8], [9], plasma diagnostics [10], microwave breakdown applications [11], [12], [13], RF heating of fusion plasmas, and microwave rocket [14]. The EM wave propagation into a plasma (complex dispersive medium) primarily depends on the spatial distribution of plasma density and the frequency of the incident EM wave [15]. The scattering features of microwaves by an inhomogeneous plasma due to wave reflection and absorption is usually used to investigate microwave–plasma interaction. In the case of a collisional unmagnetized plasma, the EM wave can propagate in an underdense plasma while being attenuated due to electron-neutral collisions. However, when the plasma density is significantly high (greater than critical density) the EM wave gets reflected and can be completely blocked by an overdense plasma. This interplay of transmission, reflection, and absorption of an EM wave propagating in an inhomogeneous plasma leads to complex EM scattering patterns. Several studies have been performed to investigate the effects of different electron density profiles on scattering patterns of an incident EM wave of a given frequency. The study of microwave–plasma interaction becomes challenging when the plasma is inhomogeneous and the density profile is complicated due to the presence of gradients such as a Gaussian density profile.

Several theoretical and computational methods exist for the study of EM wave propagation in a plasma [15], [16], [17]. The finite-difference time-domain (FDTD)-based computational electromagnetics (CEM) technique has remained one of the most preferred methods to model and accurately simulate the microwave–plasma interactions [15], [18], [19], [20]. Most of the traditional CEM approaches whether iterative or direct are computationally challenging due to stringent numerical criteria that lead to high memory usage and longer simulation time as the problem size increases [21]. To overcome such challenges inherent to traditional EM solvers, different approaches mostly sought either the use of advanced parallelization techniques or selective meshing to reduce long simulation time without loss in accuracy of results [21], [22], [23]. The high computational cost associated with CEM techniques becomes prohibitive where the real-time analysis of the EM-plasma interaction is of utmost importance. Therefore, it is extremely valuable to explore alternative approaches that can address the problem of the high computational cost associated with traditional EM solvers.

As described earlier, primarily two sets of data (plasma density profile and scattered EM wave pattern) are associated with microwave–plasma interaction problem, and therefore neural networks can be potentially used to learn the nonlinear

Manuscript received 30 May 2022; revised 9 September 2022; accepted 1 October 2022. Date of publication 9 November 2022; date of current version 12 December 2022. This work was supported by the Department of Science and Technology-Science and Engineering Research Board (DST-SERB), Government of India under Project CRG/2018/003511. *(Corresponding author: Bhaskar Chaudhury.)*

The authors are with the Group in Computational Science and HPC, Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar 382007, India (e-mail: bhaskar_chaudhury@daiict.ac.in).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMTT.2022.3217138.

Digital Object Identifier 10.1109/TMTT.2022.3217138

mappings between these two sets of data and once the network is trained, it can give the outputs in roughly $O(1)$. In the last decade, there have been vast improvements toward the development of large and powerful deep neural networks (DNNs), which have been applied to solve complex problems in the areas of computer vision and image processing. Physics-informed neural networks, a DNN framework, can also be used as a black box to approximate a physical system [24] and recent results have shown that DNNs with many layers perform a surprisingly good job in modeling partial differential equation-based complicated physics problems in terms of both speed and accuracy [25]. Off late, machine-learning (ML)/deep-learning (DL) have also been used to successfully address different complex problems in the areas of plasma physics and CEM. Deep reinforcement learning has been applied for tokamak magnetic controller design to produce new plasma configurations [26], the potential of AI/ML in predicting disruptive instabilities in controlled fusion plasmas has been established in several studies [27], [28], feasibility of applying ML models for modeling, diagnostics, and control of nonequilibrium plasmas has been discussed in [29] and DL has been also used for extracting electron scattering cross sections from plasma swarm data [30].

A convolutional neural network (CNN)-based architecture is learned to solve full-wave inverse scattering problems [31]. The visualizations generated from the problems can be used to train and get the results from the neural network which can potentially help solve and accelerate the traditional equation-based solvers [32]. DL as applied to EMs, antenna, and EM wave propagation has been well reviewed in [33]. EM-Net [34] is a modified end-to-end CNN architecture with residual blocks and skip connections inspired from the UNet [35], a robust network with encoder–decoder-like structure that generates an image as an output, is widely used in image segmentation problems. Qi et al. [34] of EM-Net predict the EM field scattered by the complex geometries. An unsupervised DL model is used for solving time-domain EM simulations, encoding the initial and boundary conditions as well as the Maxwell's equations when training the network [36]. CNNs have been also explored for plasma tomography and disruption prediction from bolometer data [37]. Cheng et al. [38] compare various CNN-based architectures like UNet, MSNet to solve the 2-D Poisson equation for electric field computation in plasma simulations.

Our problem requires us to predict the scattered EM wave as an image, so we need a network that generates an image as the output. The UNet architecture provides good acceptable results and its capabilities are discussed in the existing literature [35], [39]. UNet has several advantages over similar networks, like FCN [40] and SegNet [41], such as it is relatively simple, fast, and works with smaller training data compared to other networks, hence UNet is suitable to analyze the given problem. This work is aimed at exploring the feasibility of using UNet to accelerate the accurate simulation of microwave–plasma interaction.

The key contributions of this article are as follows.

1) To the best of our knowledge, for the first time, we propose a DL-based approach for investigating the interaction of microwaves with an inhomogeneous unmagnetized collisional plasma.
2) An end-to-end DL model consisting of UNet architecture is used to significantly accelerate the solution of Maxwell's equations coupled with plasma current density term.
3) A wide range of 2-D Gaussian plasma density profiles with different peak electron density values associated with transmission, absorption, and reflection of microwaves by the plasma have been considered for generating the training data.
4) Extensive computational experiments have been carried out to demonstrate the effectiveness of the proposed approach through various ablation studies.

The remainder of the article is organized as follows. Section II provides the detailed physics of EM–plasma interaction, the physical model, and its numerical implementation. In Section III, the proposed DL methodology has been discussed. Section IV discusses the experimental work including the dataset generation, the criteria for parameter selection, the loss function, and, subsequently, the simulation results are used to provide the effectiveness of the proposed technique by comparing it with the existing results. Finally, we discuss the future work followed by a conclusion in Section V.

II. Simulation of Microwave–Plasma Interaction

A substantial number of theoretical, numerical, and experimental studies have been carried out to investigate the EM wave propagation characteristics in an unmagnetized collisional plasma. When an EM wave such as a microwave is incident on a weakly ionized unmagnetized plasma it is subjected to scattering as well as absorption. The complex relative dielectric permittivity of a collisional plasma can be expressed as

$$\epsilon(\omega) = \left(1 - \frac{\omega_p^2}{\omega^2 + \nu_m^2}\right) - i\left(\frac{\omega_p^2}{\omega^2 + \nu_m^2}\right)\left(\frac{\nu_m}{\omega}\right). \tag{1}$$

The real part in the above equation decides the permittivity, and the conductivity is determined by the imaginary part. The conductivity is given by the formula, $\sigma = \epsilon_0 \nu_m(\omega_p^2/(\omega^2+\nu_m^2))$, where $\omega_p = (n_e e^2/m_e\epsilon_0)^{1/2}$ is the plasma frequency, ω is the wave angular frequency, ν_m is the electron-neutral collision frequency, n_e is the local electron density that varies with position, and e and m_e represents the electron charge and mass, respectively. Loss of EM wave energy due to energy transfer to charged particles and subsequently to neutral particles by elastic/inelastic collisions leads to absorption. Wave scattering is determined by the density variations within the plasma. In the case of a microwave whose energy is lower than the ionization potential of the background gas, the wave lacks sufficient energy to further ionize the gas. The plasma acts as a Debye dispersive media which responds to the incident EM wave with varying dielectric properties based on the wave frequency and local density. For a fixed wave frequency, the conductivity increases as plasma density increases, which results the plasma to behave as a conductor. The coupling of the EM energy to the plasma is decided by the plasma density. At critical

density (n_{critical}), when the $\omega_p \approx \omega$, the EM wave starts getting reflected. Furthermore, for a collisional plasma, if the plasma density approaches the cutoff density ($n_{\text{cutoff}} = n_{\text{critical}}(1 + (\nu_m/\omega)^2)$), the plasma shields the incoming microwave resulting in minimum skin depth of EM wave into plasma. The plasma density profile controls the different regimes of operation of the plasma from dielectric to a conductor, depending on relation $\omega > \omega_p$, $\omega \approx \omega_p$ or $\omega < \omega_p$, corresponding to transmission, absorption, reflection, and minimum penetration into the plasma also referred as skin depth (the distance over which the E-field of the wave decays $1/e$ of its initial strength).

To generate the training data required for ML-based approach, we have used the well-established EM–plasma fluid model to simulate the physics of EM–plasma interaction [42]. The plasma under consideration is a steady-state, nonuniform, cold, weakly ionized, unmagnetized, and collisional. The plasma is assumed to be quasi-neutral, and ion contribution to the current density is negligible due to the heavier mass of ions. The model primarily comprises Maxwell's equations with the electron current density ($\boldsymbol{J}$) term

$$\frac{\partial \boldsymbol{E}}{\partial t} = \frac{1}{\epsilon_0}(\nabla \times \boldsymbol{H}) - \frac{1}{\epsilon_0}(\boldsymbol{J}) \tag{2}$$

$$\frac{\partial \boldsymbol{H}}{\partial t} = -\frac{1}{\mu_0}(\nabla \times \boldsymbol{E}) \tag{3}$$

$$\frac{\partial \boldsymbol{v_e}}{\partial t} = -\frac{e\,\boldsymbol{E}}{m_e} - \nu_m\,\boldsymbol{v_e} \tag{4}$$

where μ_0 and ϵ_0 represent the permeability and electrical permittivity of vacuum, respectively, $\boldsymbol{J}$ is the plasma current density ($\boldsymbol{J} = -e\,n_e\,\boldsymbol{v_e}$) in (A m^{-2}), e is the electron charge ($e = 1.602 \times 10^{-19}$ C), n_e is the electron density in (m^{-3}), $\boldsymbol{v_e}$ is the electron velocity in (m/s), m_e is the mass of the electron ($m_e = 9.1 \times 10^{-31}$ kg), ν_m is the electron-neutral collision frequency in (s^{-1}) [for air plasma considered here, $\nu_m = 5.3 \times 10^9$ p, where p is the ambient pressure in (torr)] [42].

The solution to the EM–plasma fluid model can be numerically achieved by using an FDTD-based computational solver that solves Maxwell's equations (2) and (3), and the plasma momentum transfer equation (4), both are coupled using the current density term J. FDTD is an explicit second-order accurate time-domain method using centered finite differences on a uniform Cartesian grid, yielding the spatio-temporal variation of the E and H fields and has been applied to a wide variety of EM scattering problems [43]. The velocity equation is discretized by the direct integration scheme.

Fig. 1 represents the 2-D Cartesian grid on which the discretized form of the system of Maxwell's equations, E_x, E_y, and H_z, as well as the charge particle velocity (v_e) equation gets updated. The discretized equations are as follows:

$$\frac{E^{n+1}_{x,i,j} - E^{n}_{x,i,j}}{\Delta t} = \frac{1}{\epsilon_0}\left\{\left(\frac{H^{n-1/2}_{z,i,j} - H^{n-1/2}_{z,i,j-1}}{\Delta_y}\right) - \left(\frac{J^{n+1}_{x,i,j} + J^{n}_{x,i,j}}{2}\right)\right\} \tag{5a}$$

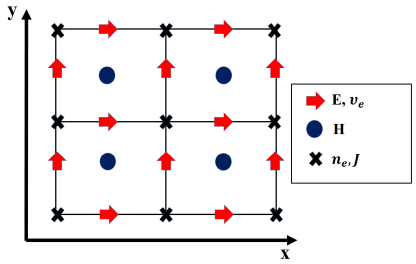

Fig. 1. Two-dimensional Cartesian grid representation of FDTD (E- and H-fields) and plasma fluid model: plasma velocity (v_e) and current density (J) solver. The square 4 × 4 grid, each grid size (Δ). The locations of E-field, H-field, plasma density (n_e), J and v_e are indicated. Proper spatial averaging is required to evaluate n_e, J, and v_e at the same locations of the E-field.

$$\frac{E^{n+1}_{y,i,j} - E^{n}_{y,i,j}}{\Delta t} = \frac{1}{\epsilon_0}\left\{\left(\frac{H^{n-1/2}_{z,i-1,j} - H^{n-1/2}_{z,i,j}}{\Delta_x}\right) - \left(\frac{J^{n+1}_{y,i,j} + J^{n}_{y,i,j}}{2}\right)\right\} \tag{5b}$$

$$\frac{H^{n+1/2}_{z,i,j} - H^{n-1/2}_{z,i,j}}{\Delta t} = \frac{1}{\mu_0}\left\{\left(\frac{E^{n}_{x,i,j} - E^{n}_{x,i,j+1}}{\Delta_y}\right) - \left(\frac{E^{n}_{y,i+1,j} - E^{n}_{y,i,j}}{\Delta_x}\right)\right\} \tag{5c}$$

$$\frac{v^{n+1}_{e_{x,i,j}} - v^{n}_{e_{x,i,j}}}{\Delta t} = \left\{\frac{e}{m_e}\left(\frac{E^{n+1}_{total_{x,i,j}} + E^{n}_{total_{x,i,j}}}{2}\right) - \nu_m\left(\frac{v^{n+1}_{e_{x,i,j}} + v^{n}_{e_{x,i,j}}}{2}\right)\right\} \tag{5d}$$

$$\frac{v^{n+1}_{e_{y,i,j}} - v^{n}_{e_{y,i,j}}}{\Delta t} = \left\{\frac{e}{m_e}\left(\frac{E^{n+1}_{total_{y,i,j}} + E^{n}_{total_{y,i,j}}}{2}\right) - \nu_m\left(\frac{v^{n+1}_{e_{y,i,j}} + v^{n}_{e_{y,i,j}}}{2}\right)\right\}. \tag{5e}$$

The grid size ($\Delta_x = \Delta_y = \Delta$) is decided based on the minimum of EM wave frequency (wavelength) and gradient of plasma density to be resolved [15]. The time step for each iteration (Δt) satisfies CFL criteria for a stable FDTD solution. The root mean square (rms) of the E-field obtained from microwave–plasma scattering has been used for data generation purpose. E_{rms} is the time-averaged E-field over one EM wave period, $E_{\text{rms}} = ((1/N)\sum_{i=1}^{N} E^2_{\text{total}})^{1/2}$, where i is the number of iterations going up to N, corresponds to 1 EM wave period and E_{total}, total E-field, $E_{\text{total}} = E_{\text{scattered}} + E_{\text{incident}}$ [43]. The solution to the above-discretized coupled microwave plasma fluid model, represented by (5) (a)–(e), can be alternatively solved using the DL-based technique which will be discussed in Section III.

III. Proposed Deep-Learning Architecture

In this section, a DL-based architecture is proposed for solving the microwave–plasma scattering problem given the plasma density and an EM wave of fixed frequency. A general flowchart followed for carrying out the experiments in the next section is shown in Fig. 2 and is explained as follows:

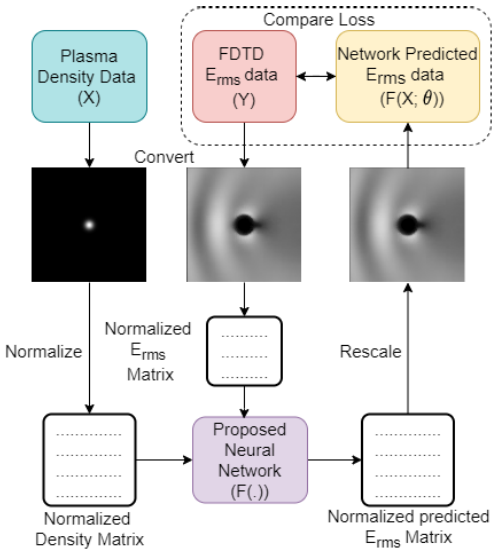

Fig. 2. Flowchart for training the proposed architecture.

DL models are trained on image datasets, and therefore the data obtained from the FDTD solver needs to be converted into images. To generate the data and to train the DL model, the plasma density and E_{rms} data from the 2-D FDTD-based simulations is normalized between 0 and 1 using the maximum value of the plasma density and E_{rms} obtained from the complete dataset, respectively. Both of these maximum values for plasma density (in m^{-3}) and E_{rms} (in V/m) are saved and used for scaling up the normalized output which is generated by the trained neural network. The normalized dataset is scaled in the range (0–255) and gray-scale images are generated for training the proposed DL network.

The generated pair of plasma density (X) and E_{rms} (Y) images are then used to train the proposed DL model. The model is then evaluated on the testing data. Plasma density image X is given as input to the trained network which outputs the predicted E_{rms} image (denoted by $F(X;\theta)$), where F represents the DL model and θ is the trained model weight matrix. The predicted E_{rms} image is converted to the physical E_{rms} values (in V/m) by scaling the normalized output by the global maximum of the dataset as discussed earlier. The predicted E_{rms} values from the DL model are then compared with the actual E_{rms} values from the 2-D FDTD-based computational solver.

The proposed architecture is a CNN-based UNet [35] where the input to the network is the single-channeled, gray-scaled, normalized image of plasma density X and the output to the network is the corresponding single-channeled, gray-scaled, normalized image of the E_{rms} data. The architecture of the network is shown in Fig. 3. It can be seen that the grayscale plasma density image is given as an input to the network. The model consists of an encoder- and decoder-like structures. The encoder consists of a series of convolutional and max pooling layers that learns the features from the image and reduces the dimensions in each layer. It helps the network to learn training weights and the reduction in image dimension decreases the complexity of the model. There are six encoder units each having a convolutional layer with n filters, where n is twice the number of filters than the previous unit having 3×3 kernel size. The output of each layer is followed by the ReLU activation function.

Correspondingly, there are six decoder units with each unit having the transposed convolution operation layer with a kernel size of 2×2 followed by the ReLU activation function. The decoder layer will upsample the features to construct the output image of the network. The input to the decoder layer is connected directly to the output of the encoder. Each layer of the decoder is connected to the corresponding encoder unit's output using a skip connection as shown in Fig. 3. The skip connections are implemented by concatenating the output of one layer to the other layer to which it is connected. The output of the final decoder unit is the predicted E_{rms} image from the proposed architecture.

IV. Computational Experiments and Results

A. Dataset Generation

As discussed in Section III that DL architectures require training data. Therefore, in this section, a discussion about how the dataset is generated is carried out. Fig. 4(a) provides a phenomenological picture of the 2-D problem we are trying to simulate. Let us consider a linearly polarized plane EM wave propagating in air plasma in the X-direction. The simulation plane XY contains an electric field E and the wave vector k parallel to the X-direction. The magnetic field H is in the YZ plane perpendicular to the X-direction. This is equivalent to a Y-polarized, X-directed wave. A 2-D Gaussian plasma density has been considered given by $n_e(x, y) = n_0 \exp(-(\{x - x_0\}^2/\sigma_x^2 + \{y - y_0\}^2/\sigma_y^2))$, where x_0 and y_0 are the center of the plasma peak density (n_0) and the spread of the plasma is controlled by the plasma width σ_x and σ_y. Here, $\sigma_x = \sigma_y = \text{S}$. Different Gaussian plasma profiles can be defined by tuning the two important parameters, the width of the Gaussian S, and the peak plasma density n_0. Some representative plasma profiles for generating the dataset (pair of plasma density and E_{rms} field) to study the EM–plasma interaction are shown in Fig. 4(b) and (c). The dataset to train the network is prepared by keeping the frequency of the incident wave fixed and the shape of the 2-D Gaussian plasma profile is varied. The size of the computational domain is $1.5\lambda \times 1.5\lambda$ as per the setup shown in Fig. 4. For each instance of the profile, the data file for the plasma density and the corresponding 2-D scattered EM wave data (E_{rms}) is generated via our in-house developed FDTD computational solver [21]. The generated data is in the form of a 2-D grid and is visualized in Fig. 5 for varying peak density. It can be observed that the EM wave transmits through in the case of lower peak densities while it gets reflected in the case of higher peak densities.

B. Training Details

The proposed DL network is trained on the pair of the generated gray-scale images of the dataset. The plasma density image X is given as input to the network, and the network learns its parameters by minimizing the loss between the actual

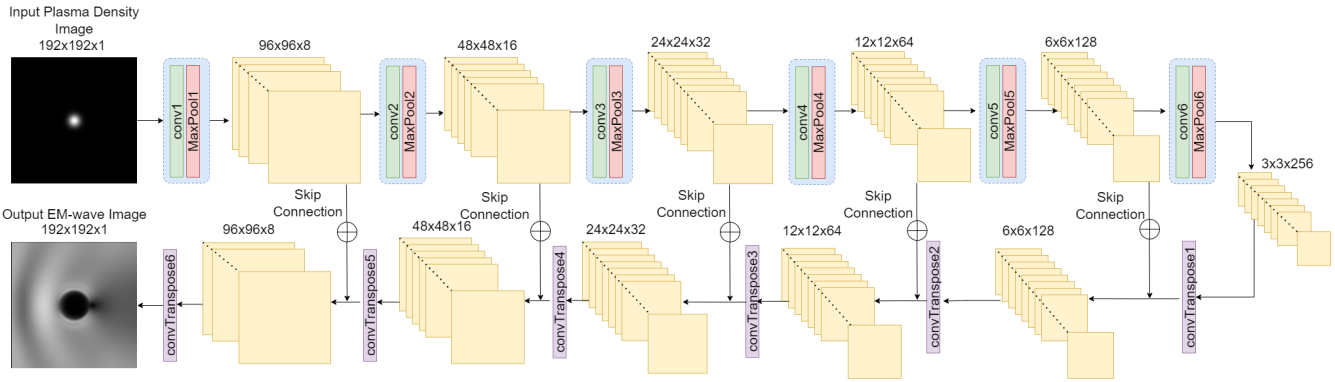

Fig. 3. Proposed architecture having encoder, decoder, and skip connections for EM–wave scattered by plasma density prediction.

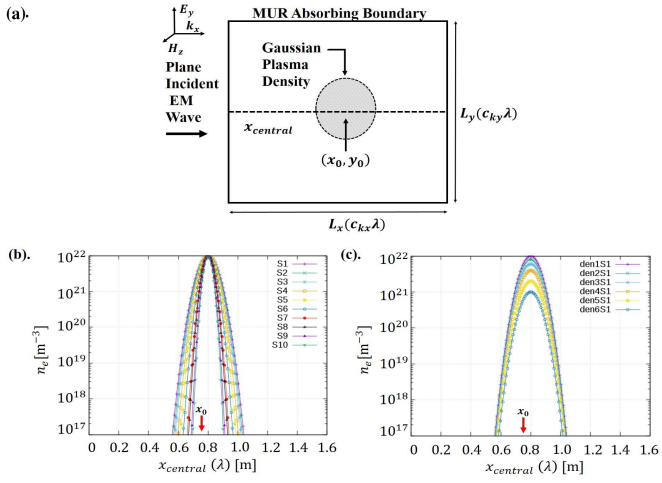

Fig. 4. (a) Schematic of the square computational domain $\{(c_{kx}, c_{ky}) \in \mathbb{Q}^+\}$; length of the domain L_x and L_y are taken in terms of the wavelength of the incident EM wave. The location x_0, y_0 is $0.5L_x$ and $0.5L_y$, respectively, and $\{c_{kx}, c_{ky}\} = \{1.5, 1.5\}$, where λ corresponds to freq = 1 GHz. The parameter space can be varied by changing two parameters of the 2-D Gaussian profile—width and peak density. (b) Plasma profile along the central x-axis (x_{central}) for different widths of Gaussian [S1:0.05λ (highest) to S10:0.02λ (lowest)] for a fixed peak plasma density, $n_0 = 10^{22}$ m^{-3}. (c) Different peak plasma densities for a fixed width of Gaussian. The den1S1: $n_0 = 10^{22}$ m^{-3} to den6S1: $n_0 = 10^{21}$ m^{-3}.

E_{rms} image denoted as Y and the output of the network which is the predicted E_{rms} image denoted as $F(X;\theta)$. The loss function for training the architecture is given as follows:

$$L(\theta) = \frac{1}{M}\sum_{i=1}^{M}\|F(X;\theta) - Y\|_2^2 + \lambda\sum_{j=1}^{l}\|W_j\|_1 \tag{6}$$

where M is the total number of training images, θ is the network weight parameter matrix, l is the total number of kernels used, and W_j is the weight of the jth kernel. The loss is minimized using the Adam optimizer [44] with learning rate $\eta = 1e-3$, $\beta_1 = 0.9$ (the exponential decay rate for first-order moment estimates), $\beta_1 = 0.999$ (the exponential decay rate for second-order moment estimates), and $\epsilon = 1e{-}7$. The kernel weights matrix for the convolution and transposed convolution layers are initialized with Glorot-uniform which draws samples from a uniform distribution. Here, L1 regularization is used to overcome the problem of overfitting with $\lambda = 1e-7$. There are 611 833 trainable parameters in the proposed DL model with six convolutional encoders, six convolutional transposed decoders, and five skip connections in between. The network is trained on NVIDIA Tesla K40c GPU using Keras API with Tensorflow running in the backend. The loss while training the model for 300 epochs is shown in Fig. 6. It can be observed that training and testing losses decrease with each epoch indicating that our network is learning.

C. Performance Comparison Metrics

To evaluate the performance of the proposed architecture in terms of predicting the image as well as the actual physical values of the E_{rms} data (in V/m), we have used existing metrics that are discussed in this section. In order to compare the quality of image reconstruction by the DL model with respect to the actual E_{rms} image, we use structural similarity index metric (SSIM) [45], where the image degradation is perceived change in structural information. For two images x and y of size $N \times N$, the similarity measure is given by

$$\text{SSIM} = \frac{(2\mu_x\mu_y + (k_1L)^2)(2\sigma_{xy} + (k_2L)^2)}{(\mu_x^2 + \mu_y^2 + (k_1L)^2)(\sigma_x^2 + \sigma_y^2 + (k_2L)^2)} \tag{7}$$

where μ denotes the mean, σ^2 denotes the variance, L is the dynamic range of pixel values, and $k_1 = 0.01$ and $k_2 = 0.03$. Values of the SSIM index are from 0 to 1. The closer the value of the SSIM metric to 1, the better the quality of image reconstruction.

To evaluate the performance of the model in predicting the E_{rms} values (in V/m) in comparison, the values obtained from the 2-D FDTD-based solver, we use two metrics. The first metric is the average of the percentage error over all the E_{rms} values on the 2-D grid. Let A_{ij} and B_{ij} denote the E_{rms} values obtained from the 2-D FDTD solver and DL-based approach at (i, j)th point on a $N \times N$ 2-D grid respectively. The average percentage error is given by

$$\text{Avg. percent error} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{|B_{ij} - Aij|}{A_{ij}}. \tag{8}$$

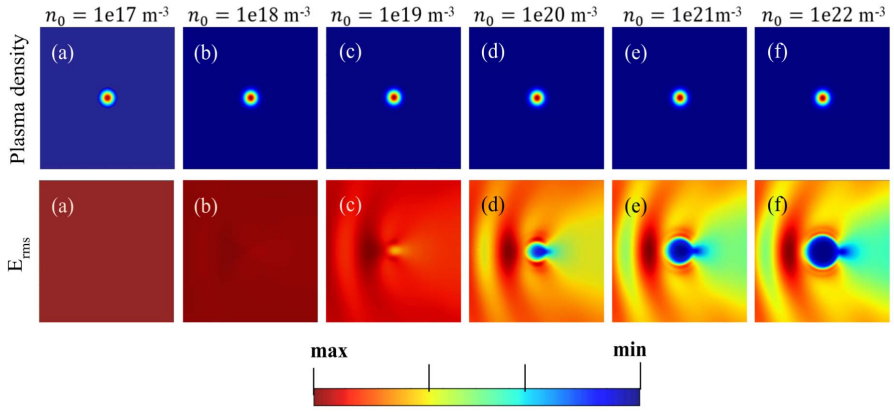

Fig. 5. (a)–(f) Generated dataset of plasma density and corresponding scattered E_{rms} for varying peak density. The colorbar represents the maxima, and minima correspond to both plasma density and E_{rms}. The maxima for plasma density is indicated by n_0 and the minima is 0. For E_{rms}, maxima are 7.07, 7.15, 7.74, 9.37, 10.36, and 10.47 V/m, respectively, and minima is 0. The skin depth of the microwave into the plasma profile reduces as n_0 increases indicated by the visibility of the exact scatterer dimension (2-D Gaussian profile) from E_{rms} (a) to (f).

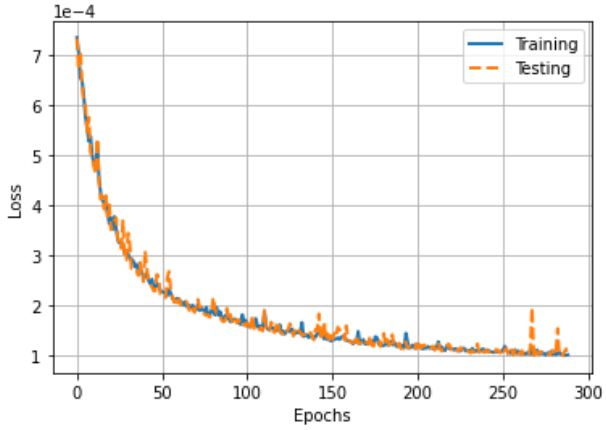

Fig. 6. MSE loss in training the model.

The second metric is the mean squared error (MSE) which is defined as

$$\text{MSE} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(B_{ij} - Aij\right)^2. \tag{9}$$

D. Experimental Results

In the microwave–plasma interaction study, the 2-D FDTD-based method discretizes the 2-D computational domain of size $1.5\lambda \times 1.5\lambda$, where λ corresponds to free space wavelength, using Yee approximation [46]. The data provided to the proposed DL-based model is such that the number of grid points per wavelength (λ) of the EM wave is 128 resulting in 192×192 grid points in the XY plane to accurately resolve the gradients in the E-field and the plasma density. In the simulation, the plane EM wave having the amplitude of 10 V/m is incident from the left-hand side of the domain as shown in Fig. 4. The frequency of the EM wave is 1 GHz.

The data used in training the network is varied by changing the peak plasma density from $n_0 = 1e21$ m^{-3} $\rightarrow$ $1e22$ m^{-3} to $n_0 = 1e17$ m^{-3} $\rightarrow$ $1e22$ m^{-3} gradually. The results of the test cases with varying dataset sizes in training are shown in Table I. It can be observed that the average SSIM index increases with the increase of the dataset range. The average percent error is observed to be less than 2%. The results obtained from the FDTD-based solver and DL-based approach for different plasma densities have been compared in Fig. 7. The first row in Fig. 7 represents density profiles changing from underdense (leftmost) plasma to overdense plasma (rightmost). The second row represents the scattered EM wave pattern obtained using the conventional FDTD-based solution. The third row shows the generated images from our proposed architecture and it can be observed that the scattering patterns closely match with that of the FDTD-based solver data shown in the second row. The fourth row shows the quantitative comparison of the scattered EM wave values shown in the second and third rows. It can be observed that the intensity values obtained by taking a 1-D profile across the central x-axis closely match with each other (FDTD versus our proposed approach).

TABLE I
EM-WAVE SCATTERING PREDICTED DATA COMPARISON WITH THE ACTUAL DATA

Dataset Range(in density m^{-3})	Avg. SSIM (image)	Avg. percent error (physical)	Avg. MSE (physical)
1e21 - 1e22	0.9894	1.8751%	0.00857
1e20 - 1e22	0..9932	1.1017%	0.00467
1e19 - 1e22	0.9935	1.3389%	0.00722
1e18 - 1e22	0.9946	1.0172%	0.00613
1e17 - 1e22	0.9955	0.9036%	0.00564

The aforementioned microwave–plasma interaction study in a $1.5\lambda \times 1.5\lambda$ computational domain using the FDTD based technique with 128 cells per λ takes approximately 18 s on Intel Xeon CPU E5-2640 V3 @ 2.60 GHz with x86_64 architecture for a physical time duration of 15 wave periods. The stable scattering pattern is obtained after the EM wave has attained a steady-state condition. Whereas it takes 0.0576 s on average per test case for the DL-based approach executed on the Intel(R) Xeon(R) CPU @ 2.20 GHz with x86_64 architecture.

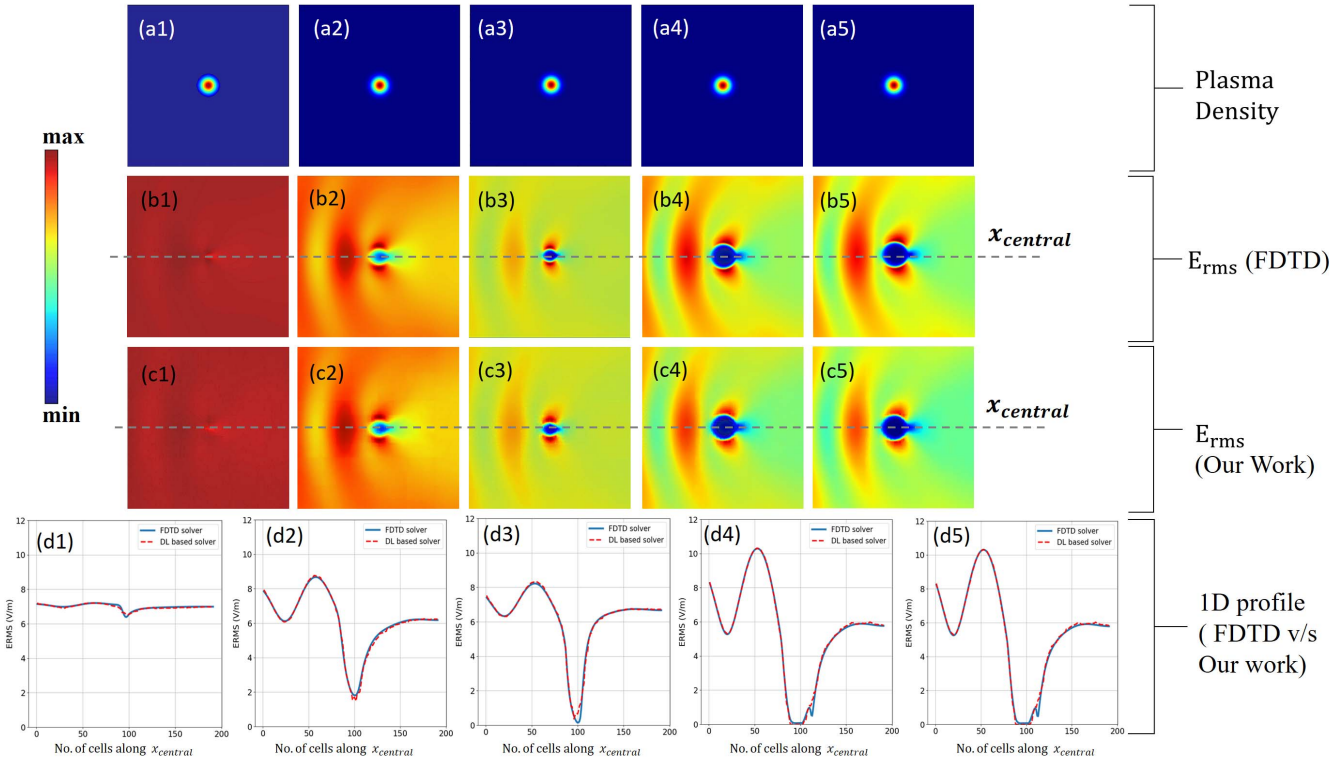

Fig. 7. Comparative study and results of example cases; row 1 (a1)–(a5): 2-D profile of input plasma density (for increasing n_0, $6e18$, $6e19$, $4e20$, $6e21$, and $8e21$ m^{-3} from left to right); row 2 (b1)–(b5): 2-D scattering (E_{rms}) pattern obtained from FDTD solver; row 3 (c1)–(c5): 2-D output E_{rms} image from the proposed DL-based architecture; row 4 (d1)–(d5): comparison of 1-D E_{rms} along the central x-axis (x_{central}) of the computational domain predicted from our work and FDTD-based solver.

For the $1.5\lambda \times 1.5\lambda$ problem size, we observe a speedup of around 350 times by using the DL-based approach compared to the FDTD-based technique. However, the computational time complexity of 2-D FDTD-based solver is $O(n^3)$, and if the problem size changes from $1.5\lambda \times 1.5\lambda$ to $3\lambda \times 3\lambda$ and finally to $6\lambda \times 6\lambda$, the execution time is around 93 and 744 s, respectively. However, in the case of DL-based approach, it will be much smaller.

E. Ablation Study

In order to explore the effect of changing the model parameters on the predicted E_{rms} values given the plasma density, the ablation studies on the effects of skip connections, number of encoder–decoder units, and method of upsampling are discussed below. In these experiments, all the models are trained (on the dataset having range $1e17$ m^{-3} – $1e22$ m^{-3}) on the same training and testing images.

1) Effect of Encoder–Decoder Units: In the proposed architecture, there are six encoder and decoder units. In this study, we perform additional experiments with four and five encoder–decoder units in the proposed architecture. The results are shown in Table II where it can be observed that adding more units will improve the results. For six unit pairs, we obtain an average SSIM value of more than 0.99 which is a good indicator for image reconstruction by the network. A similar observation is seen for average percentage and MSE metrics. Furthermore, adding more encoder–decoder units will just increase the computational cost.

TABLE II
IMPACT OF VARYING THE NUMBER OF ENCODER–DECODER UNITS

No. of encoder-decoder units	Avg. SSIM	Avg. Percent error	Avg. MSE
4 units	0.95881	4.919%	0.29512
5 units	0.97629	3.164%	0.09468
6 units	**0.99574**	**0.897%**	**0.00628**

TABLE III
IMPACT OF VARYING THE SKIP CONNECTIONS

No. of skip connections(SC)	Avg. SSIM	Avg. Percent error	Avg. MSE
All five SC	0.99574	0.897%	0.00628
Without SC	0.99533	0.955%	0.00805
With two alternate SC	**0.99601**	**0.857%**	**0.00600**

2) Effect of Skip Connections: In the proposed model, there are five skip connections that connect the output of the first five encoder and decoder units. For a second case, all the skip connections are removed. In the third case, there are two skip connections between the output of the second and fourth encoder and decoder units. We found the best results in the third case when there are two alternate skip connections as shown in Table III although the difference is much less.

3) Effect of Up-Sampling Method: The proposed architecture uses the transposed convolution operation in the decoder units. The transposed convolution operation can be replaced by a structure consisting of an up-sampling unit followed by a convolution unit. Unlike the transpose convolution which is

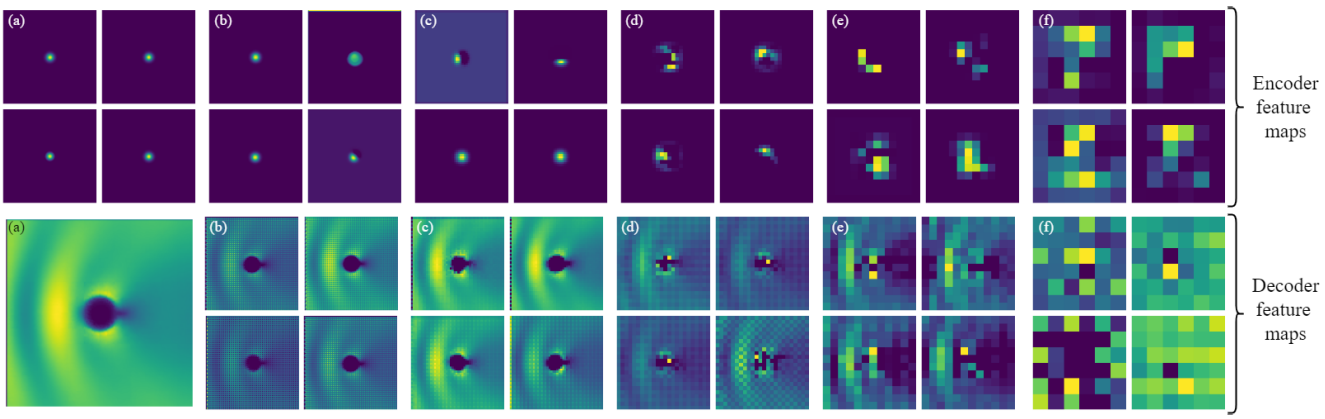

Fig. 8. Row 1: feature maps of the output of the six encoder units; row 2: feature maps of the output of the six decoder units. (a) Final output E_{rms} image of the network; the corresponding (a)–(f) pairs of the encoder–decoder pairs give the feature output in the same spatial dimensions. (For representational purpose only, four feature maps for a convolution unit are shown. The total number of feature maps for a unit is equal to the number of filters used in the convolution or transposed convolution layer.)

TABLE IV
IMPACT OF USING VARIOUS UP-SAMPLING METHODS

Up-sampling Methods	Avg. SSIM	Avg. Percent error	Avg. MSE
Transposed Convolution	0.99574	0.897%	0.00628
Bilinear interpolation	0.99568	0.960%	0.00749
Nearest neighbour interpolation	**0.99637**	**0.857%**	**0.00579**

trainable, the up-sampling layers follow a interpolation scheme that increases the dimension of the input. The combination of up-sampling with convolution is considered equivalent to the transposed convolution [47]. We use the bilinear and nearest neighbor in this study and observe that using the up-sampling with nearest neighborhood interpolation followed by convolution layers gives the best results as shown in Table IV.

F. Network Visualization Using Feature Maps

Fig. 8 shows the feature maps (outputs of the convolution units) for an example case which helps us visualize the way the network actually learns the features. It can be observed that the encoder section of the network (shown in row 1 in Fig. 8) learns the high-level features initially followed by learning low-level features (edges in the image) later. While the decoder of the network (shown in row 2 in Fig. 8) learns the low-level features first followed by learning the high-level features. The encoder down-samples the image from (a) to (f), while the decoder up-samples the image from (f) to (a). It can be observed that the decoder learns the scattering pattern unlike the encoder which just learns the structural features and down-samples the image resolution to give the structural information of the plasma density profile.

G. Future Work

We observe when the DL model is trained using a dataset containing a wide range of plasma density ($1e17 - 1e22$ m^{-3}) and the testing samples are taken from higher density regions ($1e21 - 1e22$ m^{-3}) where scattering is significant, there is a minor mismatch between FDTD- and DL-based 1-D results of E_{rms} along the central x-axis (Fig. 7). We believe that the DL-based model can be improved further to address such issues by either modifying the loss function of our current DL model or by using a PINN [48]-based additional loss function to consider the physics associated with microwave–plasma interaction.

V. CONCLUSION

This work presents a CNN-based DL model, inspired by UNet with a series of encoder and decoder units with skip connections, for the simulation of microwave plasma interaction. The scattering of a plane EM wave, with fixed frequency and amplitude, incident on a plasma medium with different Gaussian density profiles have been considered. The training data associated with microwave–plasma interaction have been generated using 2-D-FDTD-based simulations. The trained DL model is then used to reproduce the scattered EM wave from the plasma with an average percent error margin of less than 2%. The results obtained from the network have been evaluated using various metrics such as SSIM and MSE. Ablation studies along with network visualization using feature maps have also been discussed. This work can be further expanded by training the network on various shapes of plasma profiles. The DL technique proposed in this work is significantly fast compared to the existing FDTD-based computational techniques.

REFERENCES

[1] R. J. Vidmar, "On the use of atmospheric pressure plasmas as electromagnetic reflectors and absorbers," *IEEE Trans. Plasma Sci.*, vol. 18, no. 4, pp. 733–741, Sep. 1990.
[2] I. Alexeff, T. Anderson, S. Parameswaran, E. P. Pradeep, J. Hulloli, and P. Hulloli, "Experimental and theoretical results with plasma antennas," *IEEE Trans. Plasma Sci.*, vol. 34, no. 2, p. 166–172, Apr. 2006.
[3] B. Chaudhury and S. Chaturvedi, "Study and optimization of plasma-based radar cross section reduction using three-dimensional computations," *IEEE Trans. Plasma Sci.*, vol. 37, no. 11, pp. 2116–2127, Nov. 2009.
[4] O. Sakai and K. Tachibana, "Plasmas as metamaterials: A review," *Plasma Sources Sci. Technol*, vol. 21, no. 2, Jan. 2012, Art. no. 013001.

[5] A. Semnani, S. O. Macheret, and D. Peroulis, "A high-power widely tunable limiter utilizing an evanescent-mode cavity resonator loaded with a gas discharge tube," *IEEE Trans. Plasma Sci.*, vol. 44, no. 12, pp. 3271–3280, Dec. 2016.
[6] A. Semnani, S. Macheret, and D. Peroulis, "A quasi-absorptive microwave resonant plasma switch for high-power applications," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 8, pp. 1–9, May 2018.
[7] V. Semenov, M. Lisak, and D. Anderson, "Electric field enhancement and power absorption in microwave TR-switches," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 2, pp. 286–292, Mar. 1995.
[8] P. Ghosh and B. Chaudhury, "Computational investigation of microwave breakdown in HPM switching and protection," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Dec. 2021, pp. 1–4.
[9] M. Bäckström, U. Jordan, D. Andersson, A. V. Kim, M. Lisak, and O. Lundén, "Can intentional electrical discharges be used for HPM protection?" in *Proc. IEEE Int. Symp. Electromagn. Compat.*, Aug. 2011, pp. 752–757.
[10] M. N. Shneider and R. B. Miles, "Microwave diagnostics of small plasma objects," *J. Appl. Phys.*, vol. 98, Sep. 2005, Art. no. 033301.
[11] J.-P. Boeuf, B. Chaudhury, and G. Q. Zhu, "Theory and modeling of self-organization and propagation of filamentary plasma arrays in microwave breakdown at atmospheric pressure," *Phys. Rev. Lett.*, vol. 104, Jan. 2010, Art. no. 015002.
[12] B. Chaudhury, J.-P. Boeuf, G.-Q. Zhu, and O. Pascal, "Physics and modelling of microwave streamers at atmospheric pressure," *J. Appl. Phys.*, vol. 110, no. 11, Dec. 2011, Art. no. 113306.
[13] K. Frigui et al., "Microwave breakdown in waveguide filters theoretical and experimental investigations," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 12, pp. 3072–3078, Jan. 2009.
[14] M. Takahashi and N. Ohnishi, "Gas propellant dependency of plasma structure and thrust performance of microwave rocket," *J. Appl. Phys.*, vol. 125, no. 16, Apr. 2019, Art. no. 163303.
[15] B. Chaudhury and S. Chaturvedi, "Comparison of wave propagation studies in plasmas using three-dimensional finite-difference time-domain and ray-tracing methods," *Phys. Plasmas*, vol. 13, no. 12, Dec. 2006, Art. no. 123302.
[16] V. Ginzburg, J. Sykes, and R. Tatler, *The Propagation of Electromagnetic Waves in Plasmas* New York, NY, USA: Pergamon, 1971.
[17] T. Stix, *Waves Plasmas*. New York, NY, USA: American Institute of Physics, 1992.
[18] B. Chaudhury and J.-P. Boeuf, "Computational studies of filamentary pattern formation in a high power microwave breakdown generated air plasma," *IEEE Trans. Plasma Sci.*, vol. 38, no. 9, pp. 2281–2288, Sep. 2010.
[19] K. Kourtzanidis, F. Rogier, and J.-P. Boeuf, "ADI-FDTD modeling of microwave plasma discharges in air towards fully three-dimensional simulations," *Comput. Phys. Commun.*, vol. 195, pp. 49–60, May 2015.
[20] G. Wang, L. Zhang, F. He, and J. Ouyang, "Numerical study on microwave scattering by various plasma objects," *Plasma Sci. Technol.*, vol. 18, no. 8, pp. 791–797, Aug. 2016.
[21] P. Ghosh and B. Chaudhury, "Mesh refinement based simulation of complex plasma dynamics during high power millimeter wave breakdown," in *Proc. IEEE MTT-S Int. Conf. Numer. Electromagn. Multiphys. Modeling Optim. (NEMO)*, Dec. 2020, pp. 1–4.
[22] B. Chaudhury, A. Gupta, H. Shah, and S. Bhadani, "Accelerated simulation of microwave breakdown in gases on xeon phi based cluster-application to self-organized plasma pattern formation," *Comput. Phys. Commun.*, vol. 229, pp. 20–35, Aug. 2018.
[23] P. Ghosh and B. Chaudhury, "Efficient dynamic mesh refinement technique for simulation of HPM breakdown induced plasma pattern formation," 2021, *arXiv:2105.13276*.
[24] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, Nov. 2018.
[25] J. Han, A. Jentzen, and W. E, "Solving high-dimensional partial differential equations using deep learning," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 34, pp. 8505–8510, Aug. 2018.
[26] J. Degrave et al., "Magnetic control of tokamak plasmas through deep reinforcement learning," *Nature*, vol. 602, pp. 414–419, Feb. 2022.
[27] J. Kates-Harbeck, A. Svyatkovskiy, and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," *Nature*, vol. 568, no. 7753, pp. 526–531, Apr. 2019.
[28] N. Dalsania, Z. Patel, S. Purohit, and B. Chaudhury, "An application of machine learning for plasma current quench studies via synthetic data generation," *Fusion Eng. Design*, vol. 171, Oct. 2021, Art. no. 112578.
[29] A. Mesbah and D. B. Graves, "Machine learning for modeling, diagnostics, and control of non-equilibrium plasmas," *J. Phys. D, Appl. Phys.*, vol. 52, no. 30, May 2019, Art. no. 30LT02.
[30] V. Jetly and B. Chaudhury, "Extracting electron scattering cross sections from swarm data using deep neural networks," *Mach. Learn., Sci. Technol.*, vol. 2, no. 3, Sep. 2021, Art. no. 035025.
[31] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, Sep. 2019.
[32] M. Raissi, A. Yazdani, and G. E. Karniadakis, "Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations," *Science*, vol. 367, no. 6481, pp. 1026–1030, Feb. 2020.
[33] A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "DNNs as applied to electromagnetics, antennas, and propagation—A review," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2225–2229, Nov. 2019.
[34] S. Qi, Y. Wang, Y. Li, X. Wu, Q. Ren, and Y. Ren, "Two-dimensional electromagnetic solver based on deep learning technique," *IEEE J. Multiscale Multiphys. Comput. Techn.*, vol. 5, pp. 83–88, 2020.
[35] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention* (Lecture Notes in Computer Science), N. Navab, J. Hornegger, W. M. Wells and A. F. Frangi, Ed. Cham, Switzerland: Springer, 2015, pp. 234–241.
[36] P. Zhang, Y. Hu, Y. Jin, S. Deng, X. Wu, and J. Chen, "A Maxwell's equations based deep learning method for time domain electromagnetic simulations," *IEEE J. Multiscale Multiphys. Comput. Techn.*, vol. 6, pp. 35–40, 2021.
[37] D. R. Ferreira, P. J. Carvalho, and H. Fernandes, "Deep learning for plasma tomography and disruption prediction from bolometer data," *IEEE Trans. Plasma Sci.*, vol. 48, no. 1, pp. 36–45, Jan. 2020.
[38] L. Cheng, E. A. Illarramendi, G. Bogopolsky, M. Bauerheim, and B. Cuenot, "Using neural networks to solve the 2D Poisson equation for electric field computation in plasma fluid simulations," 2021, *arXiv:2109.13076*.
[39] N. S. Punn and S. Agarwal, "Modality specific U-Net variants for biomedical image segmentation: A survey," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5845–5889, Oct. 2022, doi: 10.1007/s10462-022-10152-1.
[40] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
[41] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[42] B. Chaudhury, J.-P. Boeuf, and G. Q. Zhu, "Pattern formation and propagation during microwave breakdown," *Phys. Plasmas*, vol. 17, no. 2, Dec. 2010, Art. no. 123505.
[43] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL, USA: Taylor & Francis, 1993.
[44] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[45] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.
[46] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, vol. AP-14, no. 3, pp. 302–307, May 1966.
[47] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," 2016, *arXiv:1603.07285*.
[48] Y. Chen, L. Lu, G. E. Karniadakis, and L. Dal Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Opt. Exp.*, vol. 28, no. 8, pp. 11618–11633, Apr. 2020. [Online]. Available: http://opg.optica.org/oe/abstract.cfm?URI=oe-28-8-11618

Mihir Desai received the B.Tech. degree (Hons.) in information and communication technology with a minor in computational science from the Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India, in 2022.

His research interests include image processing, machine learning, high-performance computing, computational finance, and data science.

Pratik Ghosh (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from ICFAI University, Dehradun, India, in 2012, and the M.Tech. degree in ECE (VLSI engineering) from Lovely Professional University, Phagwara, India, in 2016. He is currently pursuing the Ph.D. degree at the Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India.

He has worked as a Research Fellow (JRF and SRF) with the Government of India-funded DST-SERB project from 2019 to 2022. His current research interests include modeling and simulation, high-performance computing, machine learning, computational electromagnetics, and computational plasma physics.

Ahlad Kumar (Senior Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Jamia Millia Islamia, Delhi, India, in 2005, the M.Tech. degree from Atal Bihari Vajpayee Indian Institute of Information Technology and Management, Gwalior, India, in 2005, and the Ph.D. degree from the University of Malaya, Kuala Lumpur, Malaysia, in 2016.

From 2017 to 2019, he was a Post-Doctoral Research Fellow with Concordia University, Montreal, QC, Canada. He is currently an Assistant Professor at the Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, India. His research interests include image processing and analysis, machine learning, and deep learning.

Dr. Kumar has served as a reviewer for several international journals and conferences.

Bhaskar Chaudhury received the Ph.D. degree in computational physics from the Institute for Plasma Research, Ahmedabad, India, in 2008.

He is currently a Professor of computational and data sciences at the Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, India. Prior to joining the faculty at DA-IICT, he worked as a Researcher for around six years at the Laboratoire Plasma et Conversion d'Energie (LAPLACE), Centre National de la Recherche Scientifique (CNRS), Toulouse, France. He has been the Principal Investigator for several research projects sponsored by the Department of Atomic Energy, India; Department of Science and Technology, India; Indian Space Research Organization, India; National Supercomputing Mission, DST, India; GUJCOST, Gujarat, India; and National Science Foundation, USA. He has authored more than 100 research papers in reputed peer-reviewed journals/conference proceedings/book chapters. He has been involved in active research in the areas of computational science, data science, computational physics, machine learning, and high-performance computing.