International Journal of Microwave and Wireless Technologies

cambridge.org/mrf

Hyperparameters optimization of neural network using improved particle swarm optimization for modeling of electromagnetic inverse problems

Debanjali Sarkar, Taimoor Khan and Fazal Ahmed Talukdar

Department of Electronics and Communication Engineering, National Institute of Technology Silchar, Silchar, India

Research Paper

Cite this article: Sarkar D, Khan T, Ahmed Talukdar F (2022). Hyperparameters optimization of neural network using improved particle swarm optimization for modeling of electromagnetic inverse problems. *International Journal of Microwave and Wireless Technologies* **14**, 1326–1337. https://doi.org/10.1017/S1759078721001690

Received: 9 August 2021
Revised: 22 November 2021
Accepted: 24 November 2021
First published online: 17 December 2021

Key words:
Artificial neural network (ANN); electromagnetic bandgap (EBG); genetic algorithm (GA); monopole antenna; particle swarm optimization (PSO); ultra-wideband (UWB)

Author for correspondence:
Taimoor Khan, E-mail: ktaimoor@ieee.org

Abstract

Optimization of hyperparameters of artificial neural network (ANN) usually involves a trial and error approach which is not only computationally expensive but also fails to predict a near-optimal solution most of the time. To design a better optimized ANN model, evolutionary algorithms are widely utilized to determine hyperparameters. This work proposes hyperparameters optimization of the ANN model using an improved particle swarm optimization (IPSO) algorithm. The different ANN hyperparameters considered are a number of hidden layers, neurons in each hidden layer, activation function, and training function. The proposed technique is validated using inverse modeling of two meander line electromagnetic bandgap unit cells and a slotted ultra-wideband antenna loaded with EBG structures. Three other evolutionary algorithms viz. hybrid PSO, conventional PSO, and genetic algorithm are also adopted for the hyperparameter optimization of the ANN models for comparative analysis. Performances of all the models are evaluated using quantitative assessment parameters viz. mean square error, mean absolute percentage deviation, and coefficient of determination (R^2). The comparative investigation establishes the accurate and efficient prediction capability of the ANN models tuned using IPSO compared to other evolutionary algorithms.

Introduction

Rapid evolution has occurred in the field of wireless communication over recent years. Ultra-wideband (UWB) wireless technologies, intended for short-distance communication have gained popularity since the Federal Communications Commission (FCC) released 7.5 GHz of the unlicensed band for commercial applications. According to the FCC rulings, the UWB device occupies fractional bandwidth (FBW) greater than 20% and can commercially operate in the frequency range of 3.1–10.6 GHz. As UWB systems utilize a huge amount of bandwidth, they must share the spectrum with other wireless services and applications. To address spectrum access coordination, UWB regulatory organizations from various countries developed their own UWB radio spectrum regulations as listed in Table 1 [1]. UWB antennas being the key component of UWB systems have wide impedance bandwidth, compact size, low cost, uniform omnidirectional radiation pattern, and provide high-speed data rate. Several narrow frequency bands operate in the UWB frequency band such as WiMAX (3.3–3.7 GHz), C-band (downlink 3.7–4.2 GHz, uplink 5.925–6.425 GHz), WLAN (5.15–5.35 GHz and 5.725–5.825 GHz), and X-band (downlink 7.25–7.75 GHz, uplink 7.9–8.4 GHz). The communicating devices operating in the narrow bands may interfere with the devices operating in UWB. To avoid this electromagnetic interference, UWB antennas are designed with band rejection characteristics by etching slots on the patch or ground plane, using parasitic components, or using tuning stubs. Electromagnetic band-gap (EBG) structures are being introduced to nullify the mutual coupling interference and independently regulate the notch band. EBG structures are a periodic arrangement of metal conductors and dielectric material to intercept the transmission of specific bandwidth at a certain frequency band. EBG structures exhibit stopband characteristics that can direct the radiation of the antenna and prevent the scattering of surface waves [2–8].

Several electromagnetic (EM) simulators such as IE3D, FEKO, HFSS, and CST Studio are available for the analysis and synthesis of electromagnetic devices. These simulators are based on the finite difference method (FDM) and finite element method (FEM) for solving differential equations and method of moment (MoM) for solving integral equations of electromagnetic problems. However, the huge computational time and resources required by these EM simulators are their main constraints. To overcome these limitations, computational intelligence (CI) techniques are proved to be an alternate solution. CI techniques are being extensively used in the field of microwave engineering for modeling and optimization of microwave structures. These techniques are efficient for solving intensive nonlinear problems

© The Author(s), 2021. Published by Cambridge University Press in association with the European Microwave Association

CAMBRIDGE UNIVERSITY PRESS

Table 1. UWB Regulation standards of different countries

Country	Frequency range	Effective isotropic radiated power (EIRP)
USA	3.1–10.6 GHz	−41.3 dBm/MHz
Europe	6.0–8.5 GHz	−41.3 dBm/MHz
Korea	1.0–10 GHz	−66.5 dBm/MHz
Japan	3.4–4.8 GHz, 7.25–10.2 GHz	−41.3 dBm/MHz
Singapore	2.2–10.6 GHz	−35 dBm/MHz

as they require limited computational resources. Artificial neural network (ANN) has been widely used in modeling complex EM structures because of their ability to learn from experience and accurate predictions compared to numerical techniques [9–14].

ANNs are characterized by hyperparameter values which are responsible for defining the network structure, regularization parameters, and learning rate. The performance of the NN model depends on the optimal selection of these user-specified hyperparameter values. Usually, the hyperparameters are selected using trial-and-error or by grid search. Grid search is a hyperparameter tuning method based on scanning each possible combination of parameters and evaluating the NN model accordingly [15]. However, these traditional approaches are time-consuming and not feasible for assessing a higher number of hyperparameters. Random search algorithms outperform the traditional techniques in terms of efficiency as it tries to find a global optimum. Each hyperparameter is statistically distributed such that the values may be randomly sampled. The drawback of random search is that it has a higher rate of convergence and yields high variance during computation [16]. Evolutionary algorithms have been widely adopted to find the optimum parameter value of the network model. In [17, 18], a genetic algorithm (GA) has been used to configure ANN topology and tune the network weights and biases. The optimal regularization hyperparameters and activation functions of a multi-layer NN model are determined using GA in [19, 20]. GA has also been used to optimize the architecture of a convolution neural network (CNN) in [21]. Particle swarm optimization (PSO) has also been opted to determine the optimal hyper-parameters for designing CNN-LSTM network, deep neural network (DNN), and back-propagation network [22–25].

In PSO, each particle's movement is dictated by its local best position in order to reach the global best position by computing its fitness. However, if a particle does not improve its fitness, the previous velocity vector has an effect on the next velocity vector. In order to overcome this, a mutation mechanism is subjected to the velocities of the particles that are unable to locate a better position [26]. The PSO algorithm based on velocity mutation has been utilized for optimizing the geometry of a reconfigurable antenna array and complementary split-ring resonator [27, 28]. A fractal antenna and a log-periodic dipole array have also been optimized using this algorithm for obtaining low S_{11} and high-gain radiation pattern respectively [29, 30]. The contribution of this work is the integration of the PSO algorithm based on velocity mutation with ANN which has been conceptualized for the first time to the best of the authors' knowledge. In the presented work, ANN-based modeling is proposed for inverse modeling two EBG unit cells and an EBG loaded antenna. ANN modeling for the two EBG unit cells is implemented to predict the geometrical parameters from resonant frequency and its corresponding reflection coefficient. These two unit cells are incorporated into a slotted monopole antenna to realize penta notched band characteristics. The geometrical parameters of the proposed antenna are predicted from the UWB impedance bandwidth and multiband notch frequencies of the antenna. Hyperparameters of the NN models, such as number of hidden layers, number of hidden neurons in each layer, activation function, and training function are tuned using PSO algorithm based on velocity mutation mechanism termed in this work as improved PSO (IPSO).

Inverse modeling

Multilayer perceptron

ANN is a computational model for surrogate modeling aimed at reducing computational time and resources. It bypasses the requirement of lengthy and tedious analysis and mathematical calculations. NN consists of artificial neurons in multiple layers and maps the input data with the target output. Neurons of each layer are connected with each other and carry weights that are responsible to excite or inhibit the input signals. One of the widely used ANN architecture is the multilayer perceptron (MLP) consisting of an arbitrary number of hidden layers between an input and output layer. MLP is a feed-forward network and is mostly implemented for supervised learning problems. It is used to process the correlation between input and target output using backpropagation. In the forward pass, the input data flows through the intermediate layers to the output layer. Error difference between the target and predicted output is calculated. The partial derivative of the error function with respect to weights and biases are backpropagated through the network. Weights and biases are adjusted in each iteration until the network reaches a state of convergence.

The equation representing the input-output relationship of a generalized MLP is shown in Fig. 1 can be expressed as,

$$y_o = f^{l+1}\left(\sum_{p=1}^{q} w_{o,p}^{(l+1)}\left(f^{l}\left(\sum_{k=1}^{s} w_{p,k}^{(l)}\left(\ldots\ldots\left(f^{1}\left(\sum_{i=1}^{n} w_{j,i}^{(1)}x_i + b_j^{(1)}\right)\right)\ldots..\right) + b_p^{(l)}\right)\right) + b_o^{(l+1)}\right) \quad (1)$$

where x_i is the input to i^{th} neuron and y_o is the output of o^{th} neuron of MLP; f^1, f^l, and f^{l+1}represents the activation function of the first hidden layer, l^{th} hidden layer, and output layer respectively; $w_{j,i}^{(1)}$ denotes the connection weights between i^{th} input neuron and j^{th} hidden neuron; $w_{p,k}^{(l)}$ denotes the connection weights between k^{th} and p^{th} hidden neuron; $w_{o,p}^{(l+1)}$ denotes the connection

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

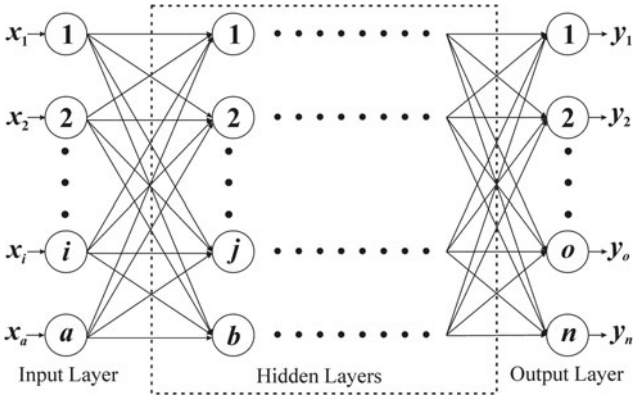


Fig. 1. Generalized architecture of MLP.

weights between p^{th} hidden and o^{th} output neuron; $b_j^{(1)}$, $b_p^{(l)}$, and $b_o^{(l+1)}$ represents the biases of hidden and output layer, respectively.

An MLP can be created with multiple hidden layers with a large number of neurons in each layer. However, an increase in the number of hidden layers and neurons may lead to generalization capability loss [31]. Also, few numbers of hidden layers and neurons fail to map complex input-output relationships and leads to underfitting. To achieve faster convergence, proper selection of activation function is necessary. There are several training functions used to train the MLP, which have a significant influence on its performance. Over the years, the trial and error approach is widely used to find the hyperparameters of MLP. In the proposed work, we have explored the capabilities of IPSO to determine appropriate MLP architecture (MLP-IPSO) with better hyperparameter configuration.

Design of MLP-IPSO

PSO is a search-based optimization algorithm inspired by the social behavior of swarms. An initial set of the population with random position and velocity are selected and moved around in a multidimensional search space. The movement of each particle is governed by its local best position to reach the global best position. The proximity of the particle to the global best is measured using a fitness function. The position and velocity of each particle are updated at each iteration until a global optimum solution is achieved. However, if a particle fails to achieve better fitness, then the previous velocity vector affects the next velocity vector. In order to avoid that, velocity mutation is introduced in this IPSO algorithm [26].

Let the MLP configuration be represented as $\lambda = (\lambda_N, \lambda_H)$ where λ_N denotes the network architecture and λ_H denotes the hyperparameter configuration respectively. The goal of the optimization is to tune $\lambda_H = (\lambda_D, \lambda_C) \in K$ where K is the decision set of all hyperparameters. The hyperparameters to be optimized are the number of hidden layers, the number of neurons in each hidden layer, activation function, and training function. The discrete hyperparameters λ_D include the number of hidden layers and the number of neurons in each hidden layer. Activation function and training function are categorical or non-ordinal hyperparameters λ_C. Every λ_N is trained using corresponding λ_H in order to minimize $err_{MLP}[(\lambda_N, \lambda_H)]$ which is calculated as,

$$err_{MLP}[(\lambda_N, \lambda_H)] = \frac{\sum_{i=1}^{N} [f_p(x_i) - f_t(x_i)]^2}{N} \tag{2}$$

where $f_p(x_i)$ is the predicted model output and $f_t(x_i)$ is the target output.

For MLP-IPSO, an initial population of particles with random position and velocity is selected. Each p_i particle has its position $x_i \in S$ and velocity $v_i \in S$ where S is the search space of all the hyperparameters. After population initialization, each particle p_i determines an MLP configuration, $p_i = (net_{IPSO}, H_{IPSO})$ where, net_{IPSO} and H_{IPSO} denotes the MLP architecture and hyperparameter configuration, respectively. The corresponding MLP model is then trained using the cross-validation technique [32]. k-fold cross-validation technique is widely applied to divide the dataset into k random independent subsets. k-1 subsets are used for training purposes and one of the k subsets is used for testing. During all the k-fold runs, each subset is selected as hold out at a time. The final performance of the model is then assessed by averaging the k recorded errors. The modeling performance of the MLP is assessed using mean square error (MSE) given by eq. 2.

The current fitness of p_i is compared with the previous best solution of the particle's position *pbest* and is updated if the current fitness is larger. Similarly, the global best solution of the swarm's position *gbest* is updated if the current fitness value is larger than the previous *gbest*. After updating *pbest* and *gbest*, the position and velocity of each particle is then modified using,

$$\begin{aligned} v_i^{iter+1} = k\{&v_i^{iter} + \varphi_1 rand^{iter}(pbest_i^{iter} - x_i^{iter}) \\ &+ \varphi_2 rand^{iter}(gbest_i^{iter} - x_i^{iter})\} \end{aligned} \tag{3}$$

$$x_i^{iter+1} = x_i^{iter} + v_i^{iter+1} \tag{4}$$

where, x_i and v_i represents n-dimensional position and velocity of i^{th} particle, respectively; *iter* denotes the iteration index; *rand* represents random numbers uniformly distributed in [0,1]; φ_1, φ_2 denotes the cognitive and social coefficients respectively, and k is constriction coefficient calculated as,

$$k = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|} \tag{5}$$

where, $\varphi = \varphi_1 + \varphi_2$.

At the end of iteration *iter*, if an i^{th} particle is not able to improve its fitness, then its velocity components v_i^{it} are mutated by a factor F_g, given as [26],

$$F_g = (0.6 + 0.1g)(2rand - 1) \tag{6}$$

where g denotes the number of iterations for a particle with no fitness improvement. In this case, the velocity of the particle is updated using [26],

$$\begin{aligned} v_i^{iter+1} = k\{&F_g v_i^{iter} + \varphi_1 rand^{iter}(pbest_i^{iter} - x_i^{iter}) \\ &+ \varphi_2 rand^{iter}(gbest_i^{iter} - x_i^{iter})\} \end{aligned} \tag{7}$$

The process is continued until the maximum numbers of iterations are over. The search space of all the hyperparameters to be optimized by IPSO is listed in Table 2. The optimal solutions of the algorithm obtained are considered as the tuned hyperparameters

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

Table 2. Search space for hyperparameter optimization

Hyperparameters	Search space
No. of hidden layers	[1, 2]
No. of hidden neurons	[1, 60]
Activation function	{'tansig", "logsig", "elliotsig'}
Training function	{'trainlm", "trainscg", "trainbfg", "trainrp", "trainbr", traincgb", "traincgf", "traincgp'}

of the MLP. Figure 2 represents the flowchart of the proposed MLP-IPSO approach.

Application examples

In this section, three inverse NN models based on MLP-IPSO approach have been proposed for two EBG unit cells and an EBG loaded antenna.

EBG unit cells

Two symmetrical defected spiral lines, and two L-shaped defected lines with two defected spiral lines, are employed to design EBG_1. EBG_2 is designed using four symmetrical spiral lines connected from the center. Both the EBG unit cells, shown in Fig. 3, are designed on FR4 dielectric substrate with a thickness of 1.6 mm using Ansys HFSS. The geometrical parameters of EBG_1 (a_1, b_1, c_1, d_1, e_1, a_2, b_2, c_2, d_2, e_2, l) and height of the dielectric substrate (h) are varied to obtain the resonance frequency and its corresponding $|S_{11}|$. Extensive parametric analysis has been performed on each parameter to obtain the training and testing datasets. Similarly for EBG_2, different geometrical parameters of the spiral line (a, b_1, b_2, c, d, e, f, g), and height of the dielectric substrate (h) are varied using parametric analysis to obtain the datasets. Table 3 lists out the sampling procedure for generating the datasets. A total of 1220 and 1098 datasets are obtained for EBG_1 and EBG_2, respectively. The simulated performance of both EBG_1 and EBG_2 is shown in Fig. 4. EBG_1 gives dual resonance at 5.7 and 11.1 GHz, respectively whereas EBG_2 is resonating at 7.5 GHz. The Brillouin-zone based dispersion diagrams of EBG_1

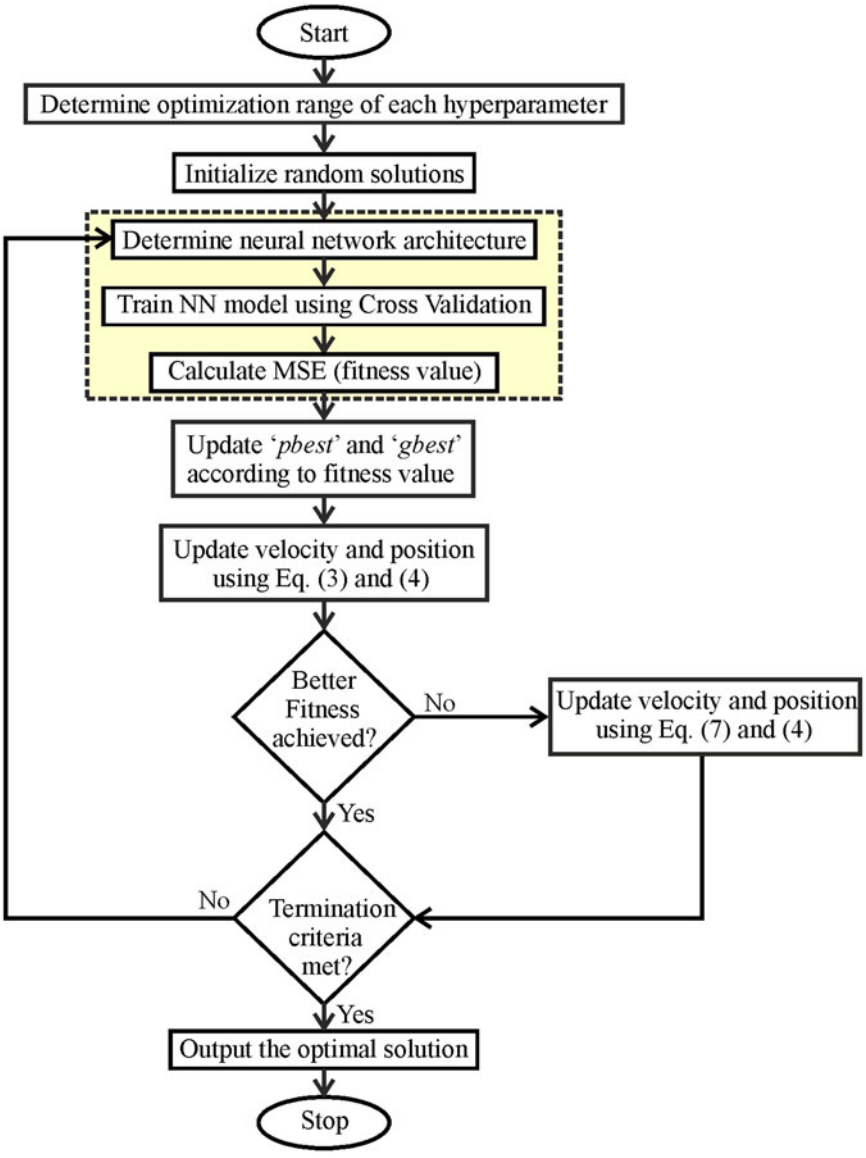


Fig. 2. Flowchart of MLP-IPSO.

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

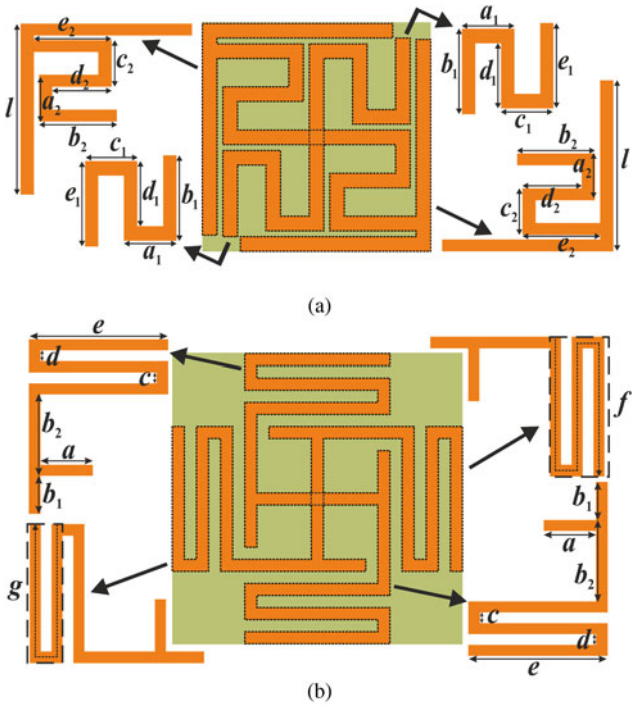


Fig. 3. EBG Unit Cells (a) EBG_1, (b) EBG_2.

and EBG_2 are shown in Fig. 5. For EBG_1, two bandgaps are exhibited between 5.66–5.85 GHz and 10.8–11.5 GHz, and the bandgap for EBG_2 falls between 7.21 and 7.79 GHz. A comparative study between the EBG structures and the existing literature in terms of types of EBG design, size, and bandgaps is conducted as listed in Table 4.

For modeling EBG_1, and EBG_2, two NN models, NN_1, and NN_2 are proposed. NN_1 has been implemented for obtaining 12-dimensional response $[G_1]$ of EBG_1 for the 4-dimensional input $[I_1]$ where $[G_1] = [a_1, b_1, c_1, d_1, e_1, a_2, b_2, c_2, d_2, e_2, l, h]$ and $[I_1] = [f_{r1}, |S_{11}|_1, f_{r2}, |S_{11}|_2]$. NN_2 is proposed for EBG_2, where 2-dimensional excitation $[I_2]$ is processed for getting 9-dimensional response $[G_2]$. Here, $[I_2] = [f_r, |S_{11}|]$ and $[G_2] = [a, b_1, b_2, c, d, e, f, g, h]$.

EBG loaded UWB antenna

A conventional UWB antenna with a rectangular radiating surface and a partial ground plane is taken as the reference antenna. The simulated performance of the antenna is shown in Fig. 6 and it can be observed that the antenna achieved an impedance bandwidth from 2.9 to 10.5 GHz. For obtaining triple-band notch characteristics, three modified *U*-shaped slots as shown in Fig. 7 are incorporated onto the radiating surface. The effect of notch frequencies is validated by individually simulating each modified *U*-shaped slot. The slot positions are optimized in such a way that it rejects three interfering bands intended for ISM, radar surveillance, and WiMAX applications.

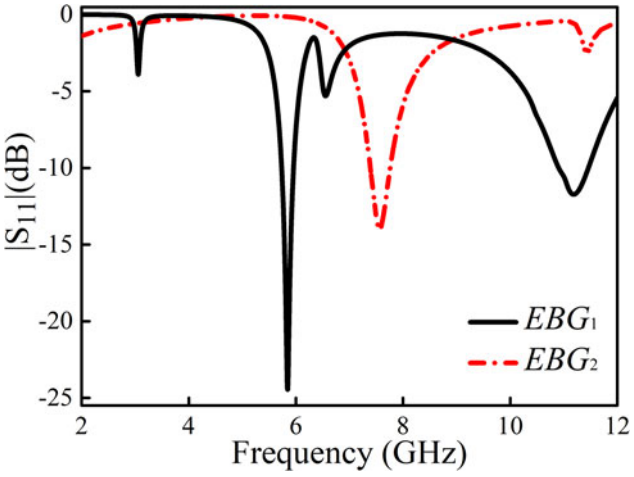


Fig. 4. Simulated performance of EBG structures.

To obtain the datasets of the reference UWB antenna, a parametric analysis is performed by varying four geometrical variables, L_P, W_P, W_f, and L_g. $Slot_1$ is introduced on the radiator and parametric analysis is performed by varying L_1 and W_1. $Slot_2$ is then incorporated on the radiating surface keeping $Slot_1$ at its optimized position. Parametric analysis is done on $Slot_2$ by varying L_2, and W_2. Finally, $Slot_3$ is placed keeping $Slot_1$ and $Slot_2$ at their optimized place and parametric analysis is performed on L_3, and W_3. For validating the effectiveness of EBG structures in rejecting frequency bands, the two EBG unit cells are incorporated onto the slotted antenna. As shown in Fig. 7, EBG_1 is etched from the radiator at a distance of d_1 and d_2 from $Slot_1$. The three modified *U*-shaped slots are kept at their optimized position and d_1 and d_2 are varied for obtaining the datasets. Similarly, EBG_2 is introduced at a distance of p_1 from the feed line as depicted in Fig. 7. Parametric variation of p_1 and p_2 is performed by keeping other geometrical variables at their optimized values.

Table 3. Sampling approach for dataset generation of EBG structures

	Parameters	Sampling range (mm)		Parameters	Sampling range (mm)
EBG_1	a_1, c_1, a_2, c_2	0.68–1.93	EBG_2	a	0.75–1.05
	b_1, e_1, b_2, e_2	0.60–2.05		b_1	0.10–0.85
	d_1, d_2	0.10–1.60		b_2	1.35–1.60
	l	1.50–4.90		c,d	0.01–0.30
	h	0.10–3.20		e	0.20–3.70
				f	5.60–8.90
				g	3.00–6.30
				H	0.10–3.20

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

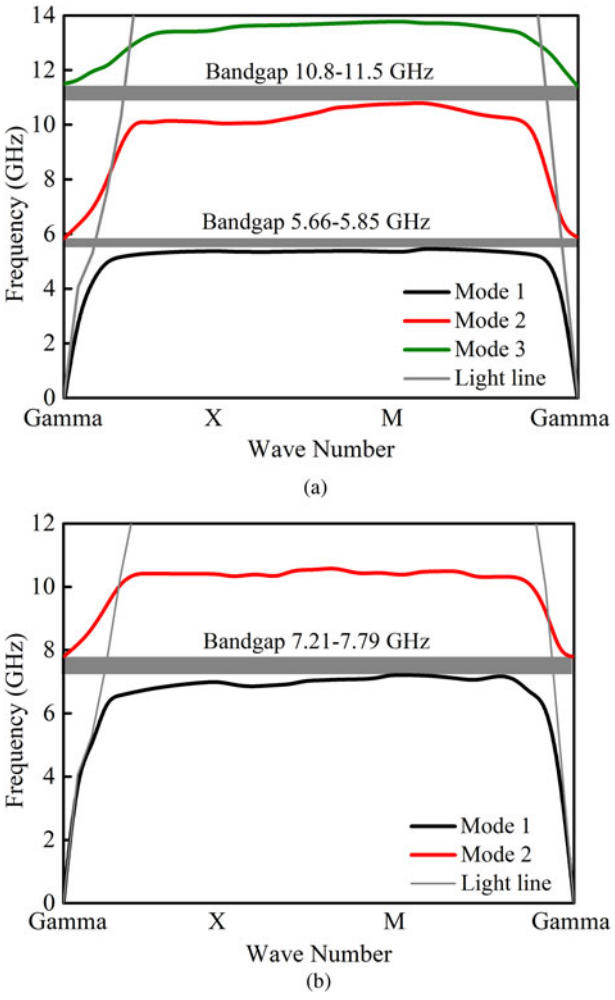


Fig. 5. Dispersion diagram of (a) EBG_1, (b) EBG_2.

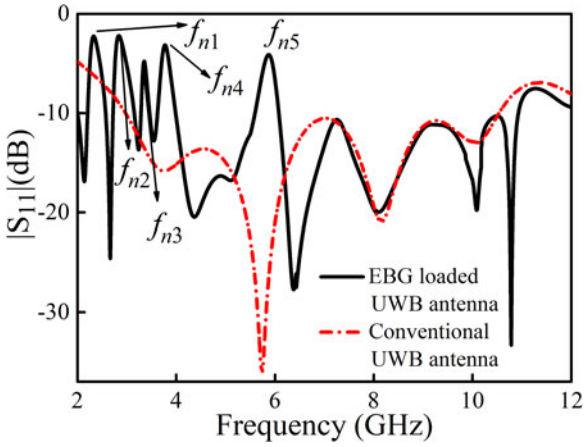


Fig. 6. Simulated performance.

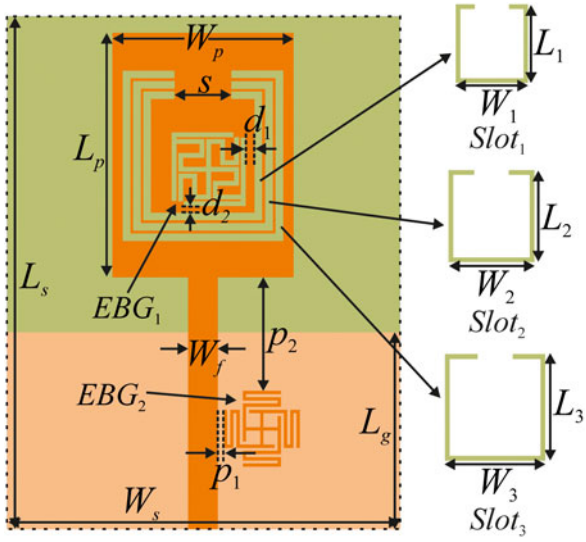


Fig. 7. Proposed EBG loaded antenna design.

Table 4. Comparison of EBG structures with existing literature

Ref.	Type of EBG	Size (mm^2)	Bandgaps (GHz)
[5]	SRS EBG	12 × 18	5.2–8.85, 10.54–15.28
[6]	TVS EBG	7 × 7	3.01–3.38
[7]	TVDS EBG	7 × 7	2.5–2.95, 5.13–5.36
[8]	TVS EBG	5.9 × 5.9	5.55–5.95
EBG_1 EBG_2	Meander line EBG	5.4 × 5.46 × 5	5.66–5.85, 10.8–11.5 7.21–7.79

Figure 7 is the final proposed antenna geometry, with two meander line EBG cells, EBG_1 and EBG_2, which produces penta notch-band characteristics. The sampling method to obtain the datasets is listed in Table 5. Figure 6 shows that the antenna satisfies the bandwidth requirement of UWB applications from f_{c1} = 2 GHz to f_{c2} = 10.74 GHz and the notches are achieved at 2.33, 2.83, 3.35, 3.87, and 5.87 GHz. The prototype of the optimized antenna geometry is fabricated using FR4 substrate sheet of 1.6 mm thickness for validating its performance. The fabricated prototype and the far-field measurement setup are shown in Fig. 8. The measured $|S_{11}|$ result of the fabricated prototype obtained using a vector network analyzer (VNA) is compared with the simulated results in Fig. 9. A good agreement between the simulated and measured performance characteristics is achieved. The slight deviations might be due to fabrication tolerances, soldering effect, or human error during the process of fabrication and/or measurement. The simulated and measured normalized radiation patterns for the antenna in *H*-plane and *E*-plane are plotted in Fig. 10. The antenna achieves an omnidirectional radiation pattern in the *H*-plane, and a dipole-like radiation pattern in the *E*-plane at the working frequencies of 4.2, 6.9, and 9.5 GHz. Figure 11 depicts that the antenna achieved stable group delay and relatively flat gain. The sudden transitions validate the radiation prohibition at the notch bands. For validating the creation of notched bands by etching slots and loading EBG structures, the surface current distribution at five notch frequencies of the antenna is plotted in Fig. 12. From the figure, it is clear that the current distribution is strongly confined around the respective slots at 2.33, 2.83, and 3.35 GHz. Besides, at 3.87, and 5.87 GHz, the surface current remains concentrated near the EBG_1 and EBG_2, respectively.

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

Table 5. Sampling approach for dataset generation of EBG loaded UWB antenna

Parameters	Optimized values (mm)	Parameter range (mm)
L_P	16.5	12.00–17.00
W_P	12.2	7.00–17.00
W_f	15.1	10.00–20
L_g	2.00	1.00–5.00
L_1	8.70	8.00–16.00
W_1	8.00	7.00–12.00
L_2	10.1	9.90–16.00
W_2	9.40	9.10–12.00
L_3	11.5	11.20–16.00
W_3	10.8	10.50–12.00
d_1	0.60	0.10–1.70
d_2	0.40	0.10–2.40
p_1	0.19	0.10–5.50
p_2	6.25	1.50–11.50

A comparative study between the presented geometry and existing literature is conducted in terms of the performance characteristics viz. impedance bandwidth, the number of notches obtained, rejected frequency, notch bands, and antenna size, as listed in Table 6. Although, [33] and [3] have proposed a more compact structure, it has achieved only single and dual notches, respectively. The suggested geometry is found to be more miniaturized than [2, 4, 34–36], and has also obtained notch characteristics at five frequency bands.

For modeling the proposed antenna structure, NN_3 has been presented to predict 14-dimensional output $[G_3]$ from 7-dimensional input $[I_3]$ where $[G_3] = [L_P, W_P, W_f, L_g, L_1, W_1, L_2, W_2, L_3, W_3, d_1, d_2, p_1$ and $p_2]$ and $[I_3] = [f_{c1}, f_{c2}, f_{n1}, f_{n2}, f_{n3}, f_{n4}, f_{n5}]$.

Computed performance

To verify the effectiveness of the MLP-IPSO approach mentioned in Section II, the performance of the three proposed models is evaluated in this section. The optimal hyperparameters obtained after the models are tuned using IPSO are listed in Table 7. It is observed that two hidden layers are proved to be optimum for all the models. The optimal activation function is tansig for NN_1, and NN_3, and logsig for NN_2. Trainbr is proved to be the optimum training function for NN_1, and NN_2, whereas for NN_3, trainlm is the optimum hyperparameter.

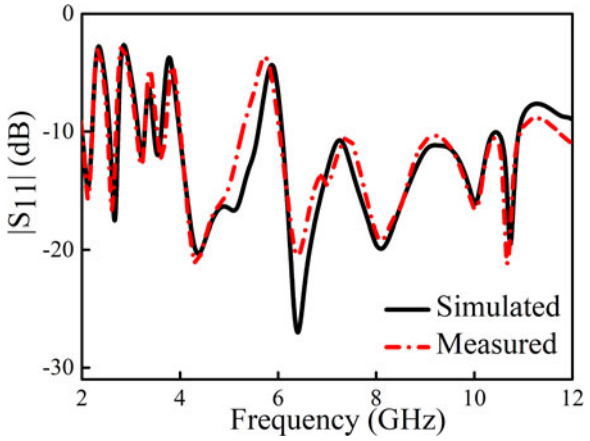


Fig. 9. Simulated and measured $|S_{11}|$ comparison.

The convergence of MLP-IPSO is shown in Fig. 13. The MSE calculated using eq. 2 is plotted against the number of iterations measured on the entire training dataset, to obtain the convergence characteristics. The effectiveness of the algorithm for NN_1, NN_2, and NN_3 is clearly depicted in Fig. 13, as the curve consistently converges toward the minimum. Tables 8 and 9 gives a comparison between the simulated and computed performance of all the models. It is observed from Table 8, that an error percentage of less than 1.5% is achieved for all the output parameters of EBG_1 and EBG_2, respectively. For the EBG loaded antenna, the error is below 1.1% for all the output parameters as shown in Table 9.

For a comparative analysis, the three MLP models are also tuned using other evolutionary algorithms, viz. hybrid PSO (HPSO) [37], conventional PSO [38], and GA [39]. The parameters of the algorithms mentioned in Table 10 are selected based on an initial parametric analysis performed with each algorithm individually. The performance of all the models has been analyzed in terms of MSE, mean absolute percentage deviation (MAPD), and coefficient of determination (R^2). The training

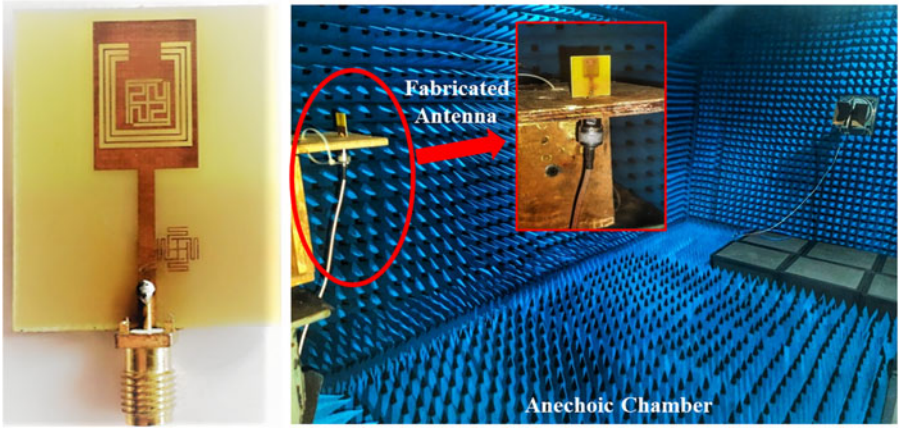


Fig. 8. Fabricated prototype and the measurement setup in anechoic chamber.

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

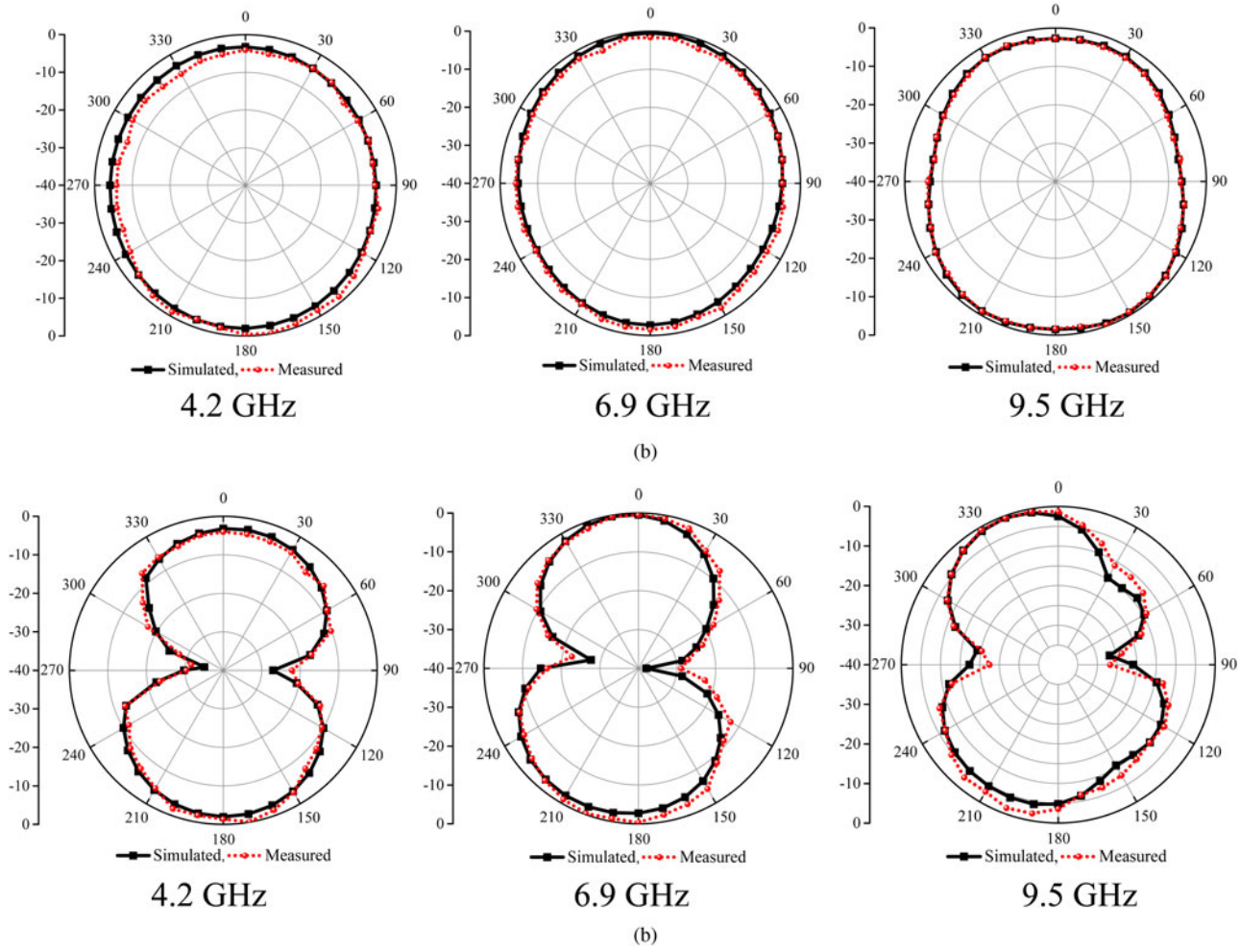


Fig. 10. Radiation pattern comparisons (a) *H*-Plane, (b) *E*-Plane.

MSE of three MLP models tuned using IPSO, HPSO, PSO, and GA are depicted in Fig. 14. Training MSE of NN_1 and NN_3 is the largest when tuned using PSO followed by GA, HPSO, and IPSO. NN_2 when tuned using HPSO and PSO has achieved almost equal training MSE. It is observed from the figure, for all the MLP models, IPSO has achieved the least MSE compared to other algorithms.

The accuracy of the models is also evaluated using MAPD and R^2; mathematical formulae for which are given as,

$$MAPD = \frac{100\%}{N}\sum_{i=1}^{N}\frac{f_p(x_i) - f_t(x_i)}{f_t(x_i)} \tag{8}$$

MAPD measures prediction accuracy and gives relative deviation from true values whereas R^2 assesses the ability of the model to adequately fit the data. As shown in Table 11, training and testing MAPD of MLP-IPSO models are less than that of MLP-HPSO, MLP-PSO, and MLP-GA models. MAPD of NN_1, and NN_2 is the worst when the hyperparameters are tuned using PSO, whereas, in the case of NN_1, training MAPD is the least when IPSO is used for tuning the hyperparameters, followed by HPSO, PSO, and GA. R^2 of all the models, listed in Table 11, depicts a stronger correlation between the target and predicted output. Although R^2 of all the models is greater than 0.97, training and testing R^2 are better for MLP-IPSO models as compared to MLP-HPSO, MLP-PSO, and MLP-GA models.

$$R^2 = \left(\frac{N\sum_{i=1}^{N} f_p(x_i) f_t(x_i) - \sum_{i=1}^{N} f_p(x_i)\sum_{i=1}^{N} f_t(x_i)}{\sqrt{\left[N\sum_{i=1}^{N}(f_p(x_i))^2 - \left(\sum_{i=1}^{N} f_p(x_i)\right)^2\right]\left[N\sum_{i=1}^{N}(f_t(x_i))^2 - \left(\sum_{i=1}^{N} f_t(x_i)\right)^2\right]}}\right)^2 \tag{9}$$

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

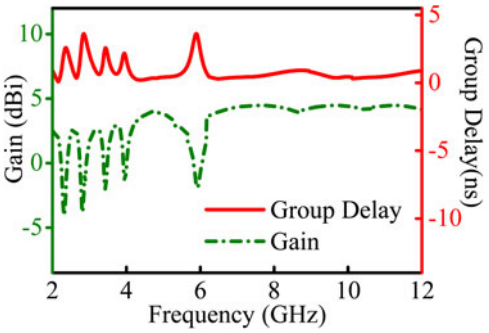


Fig. 11. Simulated group delay and gain.

Conclusion

This paper presented an IPSO algorithm to optimize the hyperparameters of an ANN-based model for solving inverse electromagnetics problems. The proposed algorithm is validated with two EBG unit cells and an EBG loaded slotted monopole antenna. The geometrical parameters of the two EBG unit cells are predicted using resonant frequency and corresponding $|S_{11}|$ as the input to the NN models. However, to predict geometrical parameters of EBG loaded slotted antenna, cut off frequencies of the UWB frequency band and notch frequencies are considered as the input for the NN model. Hyperparameters of different NN model configurations investigated in this work

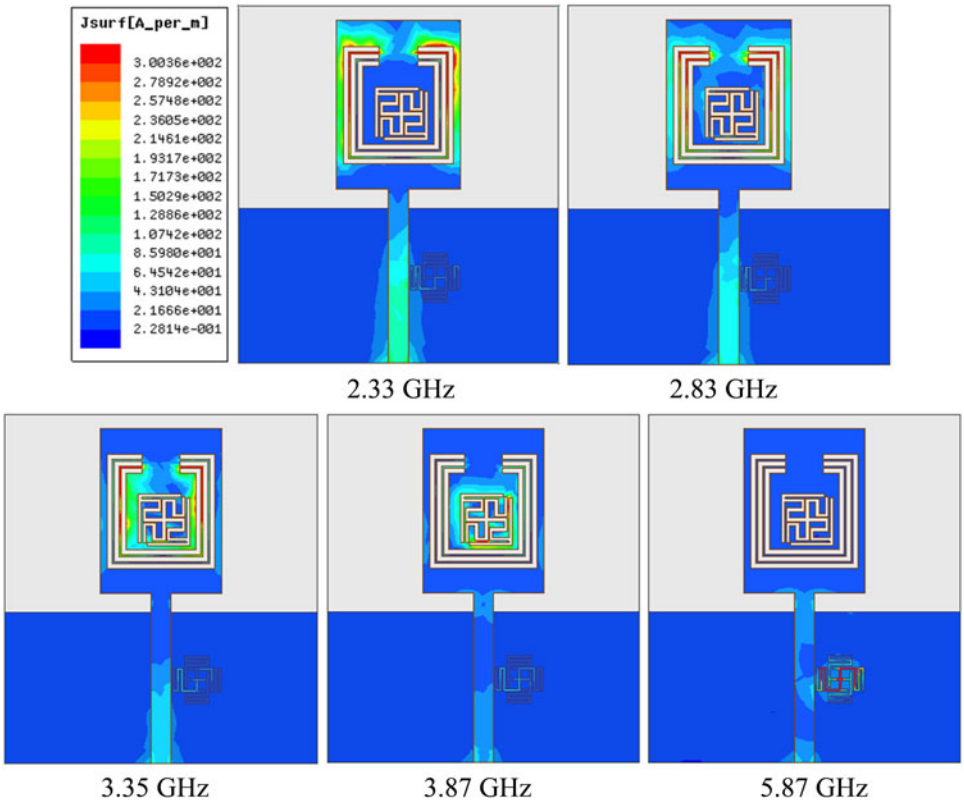


Fig. 12. Surface current distribution.

Table 6. Comparison of optimized design with existing literature

Ref.	Impedance BW (GHz)	No. of notches	Notch frequency (GHz)	Notch bands	Volume (mm^3)
[33]	3.1–10.6	1	5–6	WLAN	31 × 30 × 1.5 (= 1395)
[2]	–	2	3.5/5.5	WiMAX/WLAN	30 × 40 × 1.6 (= 1920)
[3]	3.1–9.2	2	5.2/8.2	WLAN/ITU	32 × 24 × 0.76 (= 583.68)
[4]	2.4–11.6	2	3.5/5.5	WiMAX/WLAN	50 × 50 × 0.8 (= 2000)
[34]	2.0–12.5	2	5.3/7.4	WLAN/X-band	38.5 × 46.4 × 1 (= 1786.40)
[35]	3.2–11.9	3	3.4/4.23/6.35	WiMAX/WLAN	55 × 54 × 1.59 (= 4722.3)
[36]	1.4–11.3	3	2/3.5/5.8	AWS1-AWS2/WiMAX/IEEE 802.11/ HIPERLAN	34 × 33 × 1.6 (= 1795.2)
Proposed	2–10.74	5	2.33/2.83/3.35/3.87/ 5.87	ISM/Radar Surveillance/WiMAX/ C-Band/WLAN	34.9 × 31.3 × 1.6 (= 1747.79)

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

Table 7. Optimized hyperparameter values.

NN models	No. of hidden layers	No. of hidden neurons	Activation function	Training function
NN_1	2	[40 18]	Tansig	trainbr
NN_2	2	[34 27]	Logsig	trainbr
NN_3	2	[53 36]	Tansig	trainlm

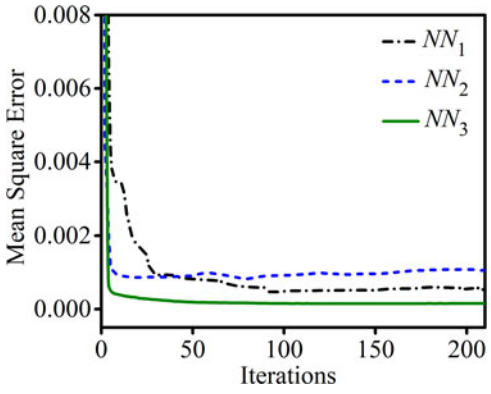


Fig. 13. Convergence curve.

have also been optimized using HPSO, PSO, and GA. Performances of all the models are compared on the basis of MSE, MAPD, and R^2. MLP-IPSO models have outperformed all other evolutionary algorithms in terms of statistical approach. The computed results of MLP-IPSO models are observed to be in close agreement with the simulated results. Though ANN has been widely used to predict the performance of microwave components for a long time, hyperparameter optimization of ANN

Table 9. Performance comparison of EBG loaded UWB antenna

Parameters	Simulated values	Computed values	% Error
L_P	16.50	16.580	0.485
W_P	12.20	12.108	0.754
W_f	15.10	15.132	0.212
L_g	2.00	2.013	0.650
L_1	8.70	8.684	0.184
W_1	8.00	7.949	0.638
L_2	10.10	10.118	0.178
W_2	9.40	9.378	0.234
L_3	11.50	11.518	0.157
W_3	10.80	10.773	0.250
d_1	0.60	0.598	0.333
d_2	0.40	0.401	0.250
p_1	0.19	0.193	1.053
p_2	6.25	6.269	0.304

Table 10. Settings of the evolutionary algorithms

Algorithms	Parameters
IPSO	Particle size = 20, $\phi_1 = \phi_2 = 2.05$
HPSO	Particle size = 20, cognitive and social coefficients = 2.0, inertia weight = 0.9–0.4 (real part)/ 1(binary part)
PSO	Particle size = 20, inertia weight = 0.6, constriction factor = 2
GA	Population size = 25, crossover rate = 0.7, mutation rate = 0.1

Table 8. Performance comparison of EBG_1 and EBG_2

EBG_1				EBG_2			
parameters	Simulated values	Computed values	% Error	parameters	Simulated values	Computed values	% Error
a_1	1.32	1.316	0.303	a	0.90	0.892	0.888
c_1	1.32	1.314	0.454	b_1	0.70	0.705	0.714
a_2	1.32	1.329	0.454	b_2	1.50	1.488	0.800
c_2	1.32	1.317	0.227	c	0.20	0.202	1.000
b_1	1.97	1.959	0.558	d	0.20	0.198	1.000
e_1	1.97	1.981	0.558	e	2.40	2.413	0.542
b_2	1.97	1.986	0.812	f	7.60	7.648	0.632
e_2	1.97	1.987	0.863	g	5.00	5.060	1.200
d_1	1.50	1.486	0.933	h	1.60	1.619	1.188
d_2	1.50	1.516	1.066	–	–	–	–
l	4.83	4.845	0.311	–	–	–	–
h	1.60	1.589	0.688	–	–	–	–

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

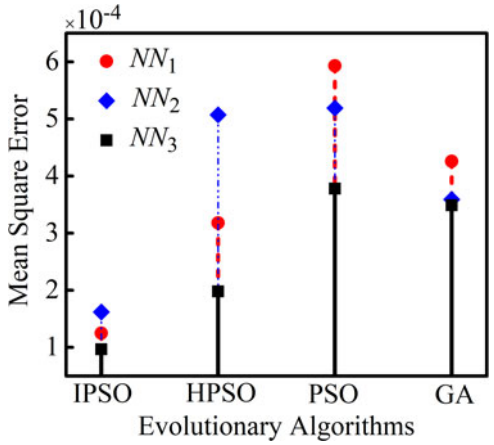


Fig. 14. MSE performance.

Table 11. Performance metrics comparison

		MAPD (%)		R^2	
		Training	Testing	Training	Testing
NN_1	IPSO	2.37	3.45	0.9862	0.9856
	HPSO	4.01	5.26	0.9850	0.9827
	PSO	5.49	6.52	0.9736	0.9702
	GA	4.79	5.54	0.9826	0.9798
NN_2	IPSO	2.35	3.98	0.9826	0.9820
	HPSO	2.92	4.69	0.9822	0.9798
	PSO	4.42	6.28	0.9728	0.9686
	GA	2.89	4.18	0.9799	0.9732
NN_3	IPSO	1.07	2.56	0.9992	0.9968
	HPSO	1.10	2.49	0.9989	0.9970
	PSO	2.05	2.93	0.9816	0.9790
	GA	2.67	3.73	0.9813	0.9835

for such a problem has not been addressed to the best of the authors' knowledge.

Acknowledgement. This work was supported by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Govt. of India (GoI) under a research grant No. SB/S3/EECE/093/2016.

References

1. **Nikookar H and Prasad R** (2009) *Introduction to Ultra Wideband for Wireless Communications*. Dordrecht, The Netherlands: Springer.
2. **Ghahremani M, Ghobadi C, Nourinia J, Ellis MS, Alizadeh F and Mohammadi B** (2019) Miniaturized UWB antenna with dual-band rejection of WLAN/WiMAX using slitted EBG structure. *IET Microwaves, Antennas & Propagation* **13**, 360–366.
3. **Yadav D, Abegaonkar MP, Koul SK, Tiwari V and Bhatnagar D** (2018) A compact dual band-notched UWB circular monopole antenna with parasitic resonators. *International Journal of Electronics Communications* **84**, 313–320.
4. **Lee DH, Yang HY and Cho YK** (2014) Ultra-wideband tapered slot antenna with dual band-notched characteristics. *IET Microwaves, Antennas & Propagation* **8**, 29–38.
5. **Alam MS, Islam MT and Misran N** (2012) A novel compact split ring slotted electromagnetic bandgap structure for microstrip patch antenna performance enhancement. *Progress in Electromagnetics Research* **130**, 389–409.
6. **Bhavarthe PP, Rathod SS and Reddy KTV** (2017) A compact two via slot-type electromagnetic bandgap structure. *IEEE Microwave and Wireless Components Letters* **27**, 446–448.
7. **Bhavarthe PP, Rathod SS and Reddy KTV** (2019) A compact dual-band gap electromagnetic band gap structure. *IEEE Transactions on Antennas and Propagation* **67**, 596–600.
8. **Dalal P and Dhull SK** (2020) Upper WLAN band-notched UWB monopole antenna using compact two via slot electromagnetic band gap structure. *Progress In Electromagnetics Research C* **100**, 161–171.
9. **Wang Z and Fang S** (2014) ANN synthesis model of single-feed corner-truncated circularly polarized microstrip antenna with an air gap for wideband applications. *International Journal of Antennas and Propagation* **2014**, 1–7.
10. **Gosal G, Almajali E, McNamara D and Yagoub M** (2016) Transmitarray antenna design using forward and inverse neural network modeling. *IEEE Antennas and Wireless Propagation Letters* **15**, 1483–1486.
11. **Xiao LY, Shao W, Jin FL and Wang BZ** (2018) Multiparameter modeling with ANN for antenna design. *IEEE Transactions on Antennas and Propagation* **66**, 3718–3723.
12. **Kapetanakis TN, Vardiambasis IO, Ioannidou MP and Maras A** (2018) Neural network modeling for the solution of the inverse loop antenna radiation problem. *IEEE Transactions on Antennas and Propagation* **66**, 6283–6290.
13. **Ustun D, Toktas A and Akdagli A** (2019) Deep neural network-based soft computing the resonant frequency of E–shaped patch antennas. *International Journal of Electronics Communications* **102**, 54–61.
14. **Sarkar D, Khan T and Laskar RH** (2020) Multi-parametric ANN modelling for interference rejection in UWB antennas. *International Journal of Electronics* **107**, 2068–2083.
15. **Bergstra J, Bardenet R, Bengio Y and Kegl B** (2011) Algorithms for hyperparameter optimization. *Advances in Neural Information Processing Systems* **24**, 2546–2554.
16. **Bergstra J and Bengio Y** (2012) Random search for hyper-parameter optimization. *Journal of Machine Learning Research* **13**, 281–305.
17. **Assuncao F, Lourenco N, Machado P and Ribeiro B**. Automatic generation of neural networks with structured grammatical evolution. 2017 IEEE Congress on Evolutionary Computation, San Sebastian, 1557–1564.
18. **Zhang S, Wang H, Liu L, Du C and Lu J**. Optimization of neural network based on genetic algorithm and BP. 2014 International Conference on Cloud Computing and Internet of Things, Changchun, 203–207.
19. **Itano F, Sousa MADAD and Hernandez EDM**. Extending MLP ANN hyper-parameters optimization by using genetic algorithm. 2018 International Joint Conf on Neural Networks, 1–8.
20. **Wang S, Roger M, Sarrazin J and Perrault CL** (2019) Hyperparameter optimization of two-hidden-layer neural networks for power amplifiers behavioral modeling using genetic algorithms. *IEEE Microwave and Wireless Components Letters* **29**, 802–805.
21. **Johnson F, Valderrama A, Valle C, Crawford B, Soto R and Nanculef R** (2020) Automating configuration of convolutional neural network hyper-parameters using genetic algorithm. *IEEE Access* **8**, 156139–156152.
22. **Kim TY and Cho SB**. Particle swarm optimization-based CNN-LSTM networks for forecasting energy consumption. 2019 IEEE Congress on Evolutionary Computation, 1510–1516.
23. **Ye F** (2017) Particle swarm optimization-based automatic parameter selection for deep neural networks and its applications in large-scale and high-dimensional data. *PLoS One* **12**, 1–36.
24. **Lorenzo PR, Nalepa J, Kawulok M, Ramos LS and Pastor JR** (2017) Particle swarm optimization for hyper-parameter selection in deep neural networks. Genetic and Evolutionary Computation Conference, Berlin, Germany 2017.

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press

25. **Sun C, Li C, Liu Y, Liu Z, Wang X and Tan J** (2019) Prediction method of concentricity and perpendicularity of aero-engine multistage rotors based on PSO-BP neural network. *IEEE Access* **7**, 132271–132278.
26. **Zaharis ZD, Gravas IP, Yioultsis TV, Laziridis PI, Glover IA, Skeberis C and Xenos TD** (2017) Exponential log-periodic antenna design using improved particle swarm optimization with velocity mutation. *IEEE Transactions on Magnetics* **53**, 1–4.
27. **Gravas IP, Zaharis ZD, Lazaridis PI, Yioultsis TV, Kantartzis NV, Antonopoulos CS, Chochliouros IP and Xenos TD** (2020) Optimal design of aperiodic reconfigurable antenna array suitable for broadcasting applications. *Electronics* **9**, 818.
28. **Reddaf A, Djerfaf F, Ferroudji K, Boudjerda M, Hamdi-Chérif K and Bouchachi I** (2019) Design of dual-band antenna using an optimized complementary split-ring resonator. *Applied Physics A: Solids and Surfaces* **125**, 1–9.
29. **Gravas IP, Sifakis NF, Zaharis ZD, Lazaridis PI and Xenos TD** (2020) Optimal fractal antenna for in-vehicle entertainment application. *Wireless Telecommunications Symposium*, 1–5.
30. **Zaharis ZD, Gravas IP, Lazaridis PI, Glover IA, Antonopoulos CS and Xenos TD** (2018) Optimal LTE-protected LPDA design for DVB-T reception using particle swarm optimization with velocity mutation. *IEEE Transactions on Antennas and Propagation* **66**, 3926–3935.
31. **Karystinos GN and Pados DA** (2000) On overfitting, generalization, and randomly expanded training sets. *IEEE Transactions on Neural Networks and Learning Systems* **11**, 1050–1057.
32. **Kwok TY and Yeung DY** (1995) Efficient cross-validation for feedforward neural networks. International Conference on Neural Networks, Australia 1995.
33. **Rahman M, Jahromi MN, Mirjavadi SS and Hamouda AM** (2019) Compact UWB band-notched antenna with integrated bluetooth for personal wireless communication and UWB applications. *Electronics* **8**, 1–13.
34. **Li WT, Hei YQ, Feng W and Shi XW** (2012) Planar antenna for 3G/Bluetooth/WiMAX and UWB applications with dual band-notched characteristics. *IEEE Antennas and Wireless Propagation Letters* **11**, 61–64.
35. **Srivastava K, Ashwani K, Kanaujia BK, Dwari S, Verma AK, Esselle KP and Mittra R** (2018) Integrated GSM-UWB Fibonacci-type antennas with single, dual, and triple notched bands. *IET Microwaves, Antennas & Propagation* **12**, 1004–1012.
36. **Iqbal A, Smida A, Mallat NK, Islam MT and Kim S** (2019) A compact UWB antenna with independently controllable notch bands. *Sensors* **19**, 1–12.
37. **Jin N and Samii YR** (2010) Hybrid real-binary particle swarm optimization (HPSO) in engineering electromagnetics. *IEEE Transactions on Antennas and Propagation* **58**, 3786–3794.
38. **Sengupta S, Basak S and Peters RA** (2018) Particle swarm optimization: a survey of historical and recent developments with hybridization perspectives. *Machine Learning and Knowledge Extraction* **1**, 157–191.
39. **Mirjalili S, Song Dong J, Sadiq AS and Faris H** (2020) Genetic algorithm: theory, literature review, and application in image reconstruction. In Mirjalili S, Song Dong J and Lewis A (eds), *Nature-Inspired Optimizers, Studies in Computational Intelligence*, vol. **811**. Cham: Springer, pp. 69–85.



Debanjali Sarkar is a research scholar in the Department of Electronics and Communication Engineering, at NIT Silchar. She received her B.E degree in electronics and communication engineering and M.Tech degree with a specialization in digital electronics and communication systems from Visvesvaraya Technological University. Her current research interest includes the design and modeling of printed ultra-wideband components using computational intelligence-based algorithms.



Taimoor Khan is an assistant professor in the Department of Electronics and Communication Engineering, NIT Silchar since 2014. Prior to that, he worked for more than 15 years in several organizations. His active research interests include printed microwave components, ultra-wideband technology, microwave energy harvesting and computational intelligence paradigms in electromagnetics. He has published over eighty-six research articles in international SCI/SCIE journals and conference proceedings of repute. He has successfully executed one major and two minor funded projects by SERB, AICTE and MHRD and currently executing two international collaborative SPARC and VAJRA research projects with QU Canada and CSUN USA. He has also been awarded a prestigious IETE-Prof SVC Aiya Memorial Award for his outstanding contribution for the year 2020.



Fazal Ahmed Talukdar received the B.E. degree in electrical engineering from the Regional Engineering College, Silchar, India, in 1987, the M.Tech. degree in energy studies from the IIT Delhi, New Delhi, India, in 1993, and the Ph.D. degree in power electronics from Jadavpur University, Kolkata, India, in 2003. He has been with the NIT Silchar, since 1991, where he is currently a professor with the Department of Electronics and Communication Engineering. His current research interests include power electronics, signal processing, and analog circuits. Dr. Talukdar is a member of the Project Review Steering Group, Ministry of Electronics and Information Technology. He is a fellow of the Institution of Engineers (India) and a Life Member of the Indian Society for Technical Education.

https://doi.org/10.1017/S1759078721001690 Published online by Cambridge University Press