Bayesian-optimized Neural Network for Evaluating the Operational Range of Freeform Microstrip Antennas

Michał Wyciszkiewicz [0009-0009-6473-0807] and Bartosz Czaplewski [0000-0001-7904-5567]

Department of Teleinformation Networks, Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland
bartosz.czaplewski@pg.edu.pl

Abstract. This paper introduces Bayesian optimization into the training process of a neural network classifier model that allows for determining the operating range of freeform microstrip antennas for randomly generated designs. The proposed approach addresses the challenge of manual hyperparameter tuning and leads to improved classification accuracy compared to existing implementations.. Moreover, the optimization process includes selected architectural parameters, resulting in a more generalizable solution that does not require expert knowledge in electromagnetic simulation procedures. As a result of the conducted experiments, a multilayer perceptron network architecture was developed, achieving high classification accuracy for frequency band intervals of the designed microstrip antennas. The application of Bayesian optimization simultaneously allowed for the effective minimization of the number of model architecture parameters.

Keywords: Bayesian Optimization, Neural Networks, Planar Microstrip Antennas.

1 Introduction

In recent years, neural networks have been increasingly implemented into existing telecommunications solutions to enhance their functionality and performance. Their application spans a wide range of tasks, including detection [1] and protection [2] of wireless networks against cyberattacks, classification [3] and traffic analysis [4] in telecommunications networks. Neural networks also play a significant role in the development of 5G systems [5], optical networks [6], and antenna design [7]. This progress was made possible by the rapid development of various neural network architectures and hyperparameter optimization techniques.

Conventional antenna design process is typically based on manual, iterative procedures involving substrate selection, geometry definition, and parametric adjustment, which are time-consuming, often suboptimal, and susceptible to designer bias. To overcome these limitations, automated and specification-driven design frameworks employing optimization algorithms have been proposed [8–11], although they introduce new challenges related to geometry representation, high-dimensional design spaces, and substantial computational requirements. Automatically generated antenna

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

structures are commonly represented either as combinations of simple geometric primitives, such as rectangles or triangles [12,13], or through coordinate-based descriptions, including splines and line segments [14–16]. This study focuses on the latter representation, which offers greater geometric flexibility but also leads to difficulties such as self-intersecting shapes, nontrivial initialization of antenna topology, and increased dimensionality. While metaheuristic optimization methods partially mitigate these issues [16,17], their practical applicability is still constrained by the high cost of electromagnetic (EM) simulations required for accurate performance evaluation. The lack of analytical models for particular antenna topology often leads to the use of heuristic or literature-based initialization strategies, which may introduce design bias. Although surrogate modeling approaches have been proposed to reduce computational burden [14,18], challenges associated with dimensionality and bias remain unresolved [19]. A comprehensive overview of these methods is provided in [20–24]. In this context, machine learning techniques offer a promising alternative by enabling rapid performance estimation and reducing dependence on full-wave EM simulations. The present work addresses the problem of identifying promising antenna geometries for a target operating frequency range, with the objective of supporting subsequent optimization stages while minimizing reliance on computationally expensive EM solvers and limiting traditional engineering bias.

The most popular neural network architectures include multilayer perceptrons (MLPs) [25-27], convolutional neural networks (CNNs) [28-30], and recurrent neural networks (RNNs) [31,32]. Each of these architectures exhibits distinct strengths and is suited to specific domains. MLPs are particularly effective for tabular data and can achieve satisfactory performance even with relatively limited training datasets. CNNs are well suited for image-related tasks due to their ability to exploit spatial correlations, whereas RNNs are designed for sequential data processing, as they can capture temporal dependencies by retaining information from previous states.

The effectiveness of these models strongly depends on the appropriate selection of hyperparameters. Many methods have been developed to optimize them, the most commonly used being random search, grid search and Bayesian Optimization (BO). Random search explores the hyperparameter space by randomly sampling configurations without relying on an explicit model of the objective function. This increases the risk that the best hyperparameter combination may not be found during optimization. Grid search involves defining a range of values for each hyperparameter. All possible combinations of the defined hyperparameters are then tested, regardless of whether a given configuration is likely to perform well. This makes the method computationally expensive. In contrast, BO constructs a probabilistic surrogate model to approximate the objective function and uses it to guide the search toward promising regions of the parameter space. Previous studies have shown that despite significantly shorter model tuning times, random search yielded the worst model tuning results. Grid search and BO methods had similar model tuning results, with a slight bias towards BO [33].

The objective of this study is to integrate BO into a classifier for assigning the appropriate operating frequency band of an antenna. Specifically, the proposed approach aims to: (1) identify an effective classifier architecture, (2) reduce the complexity of the neural network by minimizing the number of architectural parameters, and (3)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

improve the classification performance reported in [34]. The main contributions of this work are as follows: (1) the identification of training process parameters that enable high classification performance, (2) the implementation and detailed description of the BO algorithm, (3) the determination of effective ranges for the optimized parameters to facilitate practical deployment of the proposed solution, and, (4) an extensive experimental evaluation validating the proposed approach.

Although the proposed approach builds upon the framework introduced in [34], the key difference lies in the replacement of grid search with Bayesian optimization for hyperparameter and architecture tuning. This modification enables a more efficient exploration of the high-dimensional search space and leads to measurable improvements in model performance. In particular, the optimized model achieves higher classification accuracy, which enhances its usefulness as a reliable evaluation tool for antenna design. At the same time, the resulting architecture contains fewer parameters, reducing memory requirements and computational complexity during inference, while also mitigating the risk of overfitting. It is worth noting that the primary objective of introducing Bayesian optimization was to maximize predictive performance, whereas model compactness emerged as a secondary but beneficial outcome.

It should be noted that the use of Bayesian optimization introduces additional computational cost during the training stage due to the need for multiple optimization trials. However, this cost is incurred offline and is of secondary importance compared to the primary objective, which is the identification of an effective model architecture and hyperparameter configuration. In contrast to exhaustive grid search, Bayesian optimization enables more efficient exploration of the search space, while typically achieving better results than random search due to its guided sampling mechanism.

The structure of the paper is organized as follows. Section 2 reviews the related work. Section 3 presents the proposed methodology. Section 4 describes the Bayesian optimization approach employed in this study. Section 5 reports and discusses the obtained results. Finally, Section 6 concludes the paper.

2 Related work

In [34], a method for classifying the operating range of planar microstrip antennas was developed. The solution enabled effective classification of antenna operating bands without the need for electromagnetic (EM) simulations. This solution accelerates the antenna classification process and reduces the need for specialized knowledge. An MLP model was used, whose hyperparameters were optimized using grid search. As a result, two solution architectures were obtained. The first, more advanced, achieved slightly better accuracy results at the expense of a larger number of parameters. The second solution, on the other hand, struck a balance between the number of parameters and classification accuracy.

In [35], the authors introduced a convolutional neural network designed to estimate the operating frequency ranges of antennas using 256×256 binary representations of pseudo-random geometries composed of 50 vertices. The problem was formulated as a multi-label classification task, where each design could be assigned to multiple fre-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

quency bands. The accompanying dataset consisted of 136,351 samples, annotated across ten frequency intervals of 0.5 GHz each, spanning from 3 GHz to 8 GHz. To handle the high-dimensional input, the proposed architecture included six convolutional layers followed by four fully connected layers. The developed model achieved a macro-averaged test accuracy of 0.9387.

Most existing works focus on regression of antenna parameters (e.g., S_{11}) or classification between conventional well-established antenna shapes, which differs fundamentally from the multi-label classification formulation adopted in this study.

The paper [36] was devoted to optimizing the antenna design process. The resulting solution accelerates design time and simplifies the process, ensuring that antenna performance requirements are met. A trained artificial neural network (ANN) and simulated annealing (SA) were combined for this purpose. The ANN was used to predict electromagnetic performance by constructing a cost function. The SA, in turn, searched for a wider antenna bandwidth based on the ANN cost function. The resulting model enabled the design of a broadband patch antenna with a 19% relative bandwidth, confirming the feasibility of the method.

In [37] a method for automated antenna sizing and radiation pattern prediction is presented. This solution significantly reduces the design, modeling, and optimization time of antenna designs compared to conventional methods. A hybrid structure consisting of a CNN-RNN and a generative adversarial network (GAN) was used to build the network. The CNN-RNN structure enables antenna sizing and returns specifications for gain and reflectance (S_{11}). The GAN was used to predict the antenna radiation pattern. The method successfully estimated the parameters, which was verified by designing and optimizing an array antenna and comparing it with simulation results.

Bayesian optimization was successfully applied in [38]. It was used to tune the hyperparameters of a network intended for radio location. The authors argue that in designing a network for radio localization purposes, the selection of hyperparameters is crucial, particularly when the deployment environment changes and the model must be retrained. The solution they introduced meets these requirements. A neural network composed of fully connected layers was used. Bayesian optimization was used to optimize the hyperparameters, whose objective function is modeled using a Gaussian process (GP). The implemented acquisition function is expected improvement (EI). The model was be able to reduce the location prediction error by approximately 10% for WiFi networks and by approximately 20% for ultra-wideband (UWB) compared to an analogous model using RS to optimize the hyperparameters.

In [39], Bayesian optimization was used in a network designed to precisely determine the maximum power point of a photovoltaic (PV) system. The presented solution outperforms not only unoptimized models but also the traditional perturbation and observation (P&O) method used for tracking the maximum power point (MPTT) in photovoltaic systems. A regression neural network with Bayesian optimization was used to optimize its hyperparameters. Training was performed using K-fold cross-validation. The results obtained by the optimized network showed that it was able to track the maximum power point more effectively than a similar network without optimization and P&O. The network performed better regardless of the temperature at which the test was conducted or changes in solar irradiance.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

3 Methodology

To determine the antenna's operating bandwidth, an EM simulation should be conducted. An antenna is considered a promising candidate [34] for further optimization if its relative bandwidth exceeds $\theta_W = 0.1$, and its reflection coefficient remains below $\theta_R = -3$ dB within this frequency range. The proposed method offers an efficient alternative to computationally expensive EM simulations.

The 1Ddesign-to-Label-1.8.3 dataset [40] of 106,351 pseudo-random antenna designs was used to train the network. Each design has 103 parameters. The first is a scaling factor (<24, 36>), the next two are normalized feed point coordinates (<-1, 1>), and the remaining 50 are normalized coordinates (<-1, 1>) of the antenna shape vertices. After scaling, these values represent the physical dimensions in millimeters. The design method [34] was quasi-random and ensured the absence of intersections between connections and correct power supply connection. Examples are in Figure 1.

Each antenna design is also paired with a 10-element vector describing the antenna's bandwidth. The vector elements take values from the set {0;1}, where each element corresponds to a 0.5 GHz frequency band. The total frequency range is from 3 GHz to 8 GHz. Any number of labels may be assigned to a single design.

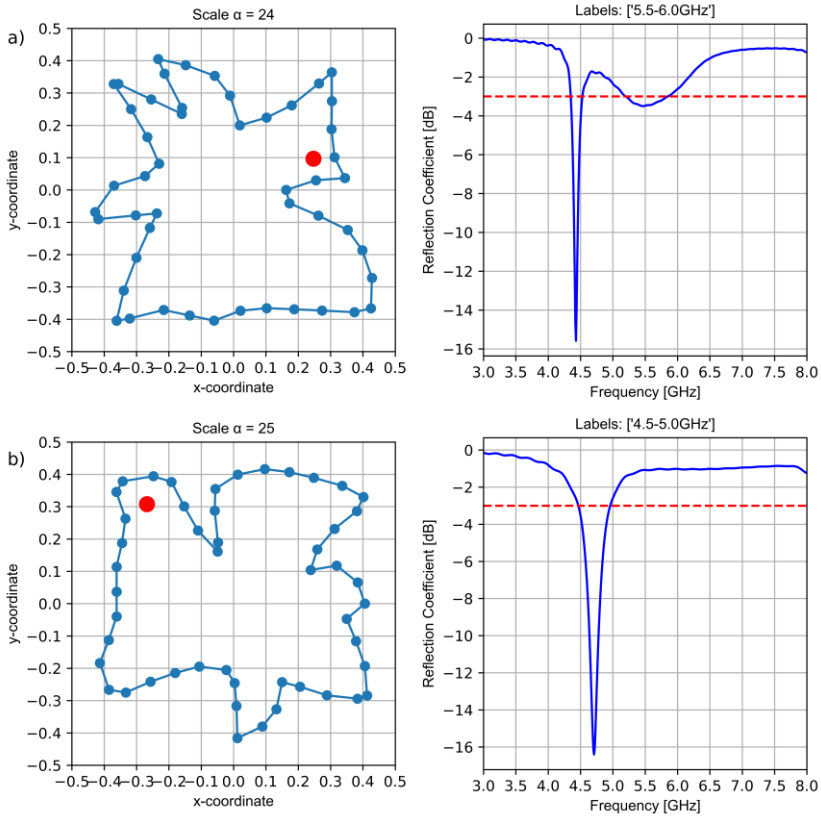

Fig. 1. Examples of antenna designs (blue – vertices, red – feed point) and antenna response.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

It is important to emphasize that the dataset used in this study is both large-scale and highly diverse, which supports the generalization capability of the proposed model within the considered design domain. The dataset contains over 100,000 pseudo-random antenna designs, generated using a procedure that ensures significant geometric variability while preserving physically valid structures. Each design is described by 50 vertices with continuous coordinate distributions, resulting in a high-dimensional and nontrivial design space representative of freeform antenna geometries. Moreover, the objective of this work is not to develop a universally applicable model for all antenna types, but rather a specialized classifier tailored to freeform microstrip antennas with complex geometries. This scope has been explicitly assumed throughout the study. Therefore, within this problem setting, the dataset provides sufficient coverage to ensure reliable model training and evaluation, and the obtained results can be considered representative for this class of antenna designs.

It is important to note that this is a multi-label classification problem (not multi-class). An antenna may operate over a wide range covering multiple subranges defined by the labels or have several disjoint ranges that meet the criteria. Therefore, each design in the dataset can be tagged with multiple labels: 77464 samples have only one active label, 20453 have two active labels, 6390 have three labels, 1776 have four labels, 253 have five labels, and 15 with six labels. To ensure uniform representation of all labels, each label appeared 15,000 times in the dataset, while taking into account the above characteristics of the multi-label classification problem.

The data was divided into three datasets: training, validation, and test. The training set included 80% of the designs, while the validation and test sets included 10% each. Training and validation data were used during the Bayesian optimization trials, while test data were used to evaluate the best trial. The training process involved selecting the MLP model's training parameters in each trial using Bayesian optimization. Subsequently, the training loop was executed on the training data for each trial. Simultaneously, the training process was verified for accuracy in each training epoch by checking the loss function and the model's prediction accuracy on the validation data. The goal was to monitor the training process to prevent overfitting. After the training process was completed, the loss value on the validation set was compared with the validation loss obtained in the best Bayesian optimization trial. If the result was better, meaning the function value was lower, then the model's parameters were saved to a separate file. Finally, after all trials were completed, the model parameters from the best trial were read to recreate the best model. This was evaluated using test data, to which the model had no access during any of the previous training stages. During training, the loss function values and the model's classification accuracy were calculated. In the testing phase, the final results were enhanced with calculations of precision, sensitivity, and F1 Score parameters.

Additionally to prevent overfitting, an early stopping mechanism was implemented. The mechanism terminated the training loop after detecting a series of several epochs in which further reduction of the training data loss function did not improve the results for the validation data. At that point, the model parameters were saved, ensuring the lowest loss function before early stopping was applied. A scheduler was also used for training, enabling dynamic changes to the learning rate to avoid skips. In

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

each training epoch, the data was divided into batches, which allowed for more effective training on a smaller number of designs. A complete list of training parameters and their corresponding values is presented in Table 1.

Table 1. Learning parameters.

Parameter	Value	Parameter	Value
Batch size	128	LR scheduler	ReduceLROnPlateau
Loss function	BCEWithLogitsLoss	LR scheduler factor	0.5
Solving algorithm	AdamW	LR scheduler patience	10
Initial LR	2.0528e−3	Max epochs	1000
Weight decay	4.0197e−6	Early stopping patience	17

Table 2. Structure of the proposed model

The Optimized NN (2,464,010 parameters)
Fully Connected (FC) layer, 103 neurons
FC, 2048 neurons, Batch Normalization, GELU, 0.263 dropout
FC, 256 neurons, Batch Normalization, GELU, 0.263 dropout
FC, 512 neurons, Batch Normalization, GELU, 0.263 dropout
FC, 1024 neurons, Batch Normalization, GELU, 0.263 dropout
FC, 1024 neurons, Batch Normalization, GELU, 0.263 dropout
FC, 10 neurons

For the microstrip antenna classification problem, a MLP model was selected. During the training phase, multiple model variants were evaluated using Bayesian optimization. Table 2 presents the solution architecture which achieved the best results. A notable advantage of this solution is its relatively small number of model parameters.

The presented model architecture comprises many neurons in the input and output layers, with fewer neurons in the hidden layers. The architecture obtained through Bayesian optimization exhibits a non-monotonic distribution of neurons across layers. Rather than following a simple funnel-shaped pattern, the number of neurons increases and decreases between layers. Such a configuration can be interpreted as the optimizer balancing feature extraction and representation capacity at different stages of the network. Layers with more neurons may capture complex or higher-dimensional interactions in the input data, while narrower layers may enforce a form of dimensionality reduction or feature selection. This flexibility lets the model allocate resources where most beneficial for classification, rather than following a fixed monotonic structure. The resulting network, despite its unconventional shape, achieves high accuracy while maintaining a relatively moderate number of parameters.

Various activation functions were tested in the hidden layers, but the best results were obtained with Gaussian Error Linear Unit (GELU) function. Each hidden layer was added a dropout, which prevented overfitting by deactivating random neurons

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

during training, and batch normalization, which ensured proper fit and scaling of the activations. The final layer had no activation function. This is dictated by the fact that the loss function was calculated based on logits, not on the model's predictions.

The model was trained in Python 3.12.10 with PyTorch 2.9.0+cu128, using an NVIDIA GeForce RTX 4060 graphics card and CUDA 13.0 driver.

4 Bayesian Optimization

Bayesian optimization played a key role in both the training process and the development of the proposed solution. It is a highly effective technique, especially for optimizing black-box functions [41]. In such cases, the actual form of the function is unknown and therefore we are unable to independently select the optimal model parameters. However, we do know that it accepts certain input parameters and returns a specific output, for example, an assigned label.

Bayesian optimization has several key advantages. It uses probabilistic models to find optimal parameters. Therefore, it is more efficient at determining which parameters should be adjusted to align the desired objective function with the actual function. Consequently, the hyperparameter combinations selected are well-optimized, enhancing the model's outcomes. Furthermore, this solution is highly flexible, as it can be used in various machine learning domains and on any network architecture.

Bayesian optimization involves two key components. The first is a Bayesian statistical model that approximates the objective function $f(x)$ based on input samples x. Constructed using a Gaussian process (GP), this model enables the estimation of the unknown black-box function. The GP is defined by the mean function $m(x)$ and the covariance function $k(x,x')$.

$$f(x) \sim GP(m(x), k(x,x')) \tag{1}$$

where $m(x)$ is zero if no prior knowledge is available and $k(x,x')$ is kernel function which defines the covariance between any two points in the input space.

The second component is the acquisition function, which enables the correct selection of subsequent samples, which are evaluated using the objective function. The most popular acquisition functions include Expected Improvement (EI) and Upper Confidence Bound (UCB).

$$EI(x) = E[max(f(x) - f(x^+), 0)] \tag{2}$$

$$UCB(x) = \mu(x) + \kappa\sigma(x) \tag{3}$$

where $f(x^+)$ is current best observed value of f, $\mu(x)$ and $\sigma(x)$ are mean and standard deviation of the GP prediction at point x and κ is a parameter balancing exploration and exploitation.

The Bayesian optimization algorithm proceeds as follows:

1. Specify the number of iterations for Bayesian optimization.
2. Sample the objective function $f(x)$ at several points.
3. Create a surrogate model using Gaussian Process (GP).

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

4. Select the acquisition function based on the GP created.
5. Maximize the acquisition function to obtain the next sampling point *x*.
6. Point *x* is evaluated using the objective function *f(x)*, and the result is added to the training data of the surrogate model.
7. Repeat the steps: GP model update → acquisition function maximization → evaluation of *f(x)* until the iteration limit is reached or the stopping criterion is met.

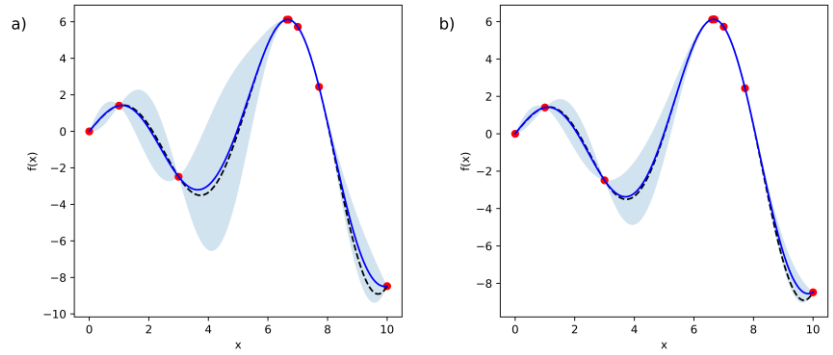

Fig. 2. Demonstration of Bayesian optimization process (blue – objective function, black dashed – black-box function, red points – samples): a) Results after *x* iteration; b) Results after *x + 1* iteration

Figure 2 illustrates the operation of the Bayesian optimization algorithm. The desired black-box function is marked with dashed lines. The blue line represents the objective function, which is estimated using GP. The blue areas are the uncertainty intervals of the GP model. As the number of iterations increases, these areas shrink, and the objective function more closely approximates the desired black-box function.

In this work, Bayesian Optimization is implemented using the Optuna library [42] with its default Tree-structured Parzen Estimator (TPE) sampler. Unlike classical formulations based on Gaussian Processes (GP), TPE does not rely on an explicit kernel function, but instead models the densities of promising and non-promising hyperparameter configurations using non-parametric estimators. This approach is particularly suitable for the considered problem, which involves a heterogeneous search space combining continuous, discrete, and categorical variables.

The number of optimization trials was set to 150 based on empirical convergence analysis, representing a trade-off between optimization quality and computational cost. During preliminary experiments, the validation performance stabilized well before reaching this limit, and increasing the number of trials did not lead to further improvements in the optimized model. Therefore, 150 trials were considered sufficient to ensure effective exploration of the hyperparameter space while maintaining reasonable computational cost. In each trial, a model was trained using hyperparameters proposed by the optimization procedure, and the objective was defined as the validation loss obtained after training. Parameter ranges are shown in Table 3.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

Table 3. Range of optimized parameters.

Learning parameter	Range
Number of layers	{2, 3, 4, 5, 6}
Number of neurons in each layer	{128, 256, 512, 1024, 2048, 4096}
Activation function	{ReLU, Leaky ReLU, GELU}
Dropout	(0.1, 0.4)
Initial learning rate	(1e-4, 1e-2)
Weight decay	(1e-6, 1e-3)
Scheduler patience	{5, 6, 7, 8, 9, 10, 11, 12, 13, 14}
Batch size	{32, 64, 128}

5 Results

Table 4 presents a detailed comparison of the proposed solution with the neural networks reported in related work [34, 35]. The evaluation includes both label-specific test accuracy for each of the ten frequency bands and the overall test accuracy, calculated as the macro-average of the per-label accuracies. Training and validation accuracies (best fold) are also reported. Additionally, for the Optimized NN, complementary metrics such as precision, recall, and F1 score are provided, giving a more complete picture of its classification performance.

The results show that the optimized neural network (NN), obtained through Bayesian optimization, consistently outperforms NN#1 [34], NN#2 [34] and CNN [35] across all evaluation metrics. Specifically, it achieves higher training accuracy (98.72% vs. 97.33%, 97.56%, and 95.97%) and validation accuracy (94.39% vs. 93.36%, 93.04%, and 93.90%). In terms of per-label test accuracy, the Optimized NN maintains superior performance across most frequency bands, resulting in an overall test accuracy of 94.29%, compared to 93.55% for NN#1, 93.31% for NN#2, and 93.87% for CNN. Notably, while the Optimized NN has a smaller number of parameters, it demonstrates better predictive performance, highlighting the effectiveness of Bayesian optimization for model design.

An additional observation from the per-label accuracies is that performance slightly decreases at higher frequency bands. For example, the lowest test accuracies correspond to the highest frequencies (above 6 GHz), indicating that classification becomes more challenging as frequency increases. Nevertheless, the optimized NN maintains higher accuracy than the reference networks even in these bands. This trend underscores the robustness of the proposed model and its suitability for applications requiring reliable performance across a wide frequency range. These results indicate that Bayesian optimization not only improves predictive performance but also leads to more compact model architectures compared to grid search-based approaches.

It is important to note that the reported overall accuracy of 94.29% should be interpreted with caution due to the multi-label nature of the problem. In this setting, accuracy is dominated by true negatives, as each sample is associated with a 10-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

dimensional binary label vector in which most entries are zeros. Consequently, high accuracy can be achieved even when the model does not perfectly capture all active labels. This effect is reflected in the gap between accuracy (94.29%) and F1 score (78.63%), where the latter provides a more balanced evaluation by jointly considering precision and recall. The lower F1 score indicates that predicting all relevant labels for a given antenna remains a more challenging task, particularly for less frequent or overlapping frequency bands. Therefore, F1 score should be considered the primary performance indicator for this task, while accuracy serves as a complementary metric.

Table 4. Results of the experiments.

Model	NN#1 [34]	NN#2 [34]	CNN [35]	The Optimized NN
Training accuracy (best fold)	97.33%	97.56%	95.97%	98.72%
Validation accuracy (best fold)	93.36%	93.04%	93.90%	94.39%
Test label-specific Accuracy	[96.91, 94.77, 94.80, 95.51, 95.90, 93.50, 91.18, 91.39, 91.93, 89.65]%	[96.79, 95.04, 95.20, 95.75, 95.83, 93.45, 90.52, 90.39, 91.00, 89.15]%	[96.52, 95.26, 95.08, 95.54, 96.36, 93.83, 91.79, 92.11, 91.39, 90.84]%	[97.61, 96.29, 95.98, 96.39, 96.63, 93.79, 92.04, 91.45, 92.11, 90.65]%
Overall test accuracy	93.55%	93.31%	93.87%	94.29%
Precision	N/A	N/A	71.78%	80.67%
Recall	N/A	N/A	69.78%	76.69%
F1 score	N/A	N/A	70.15%	78.63%

6 Conclusions

The objective of this work was to apply BO to the training process of an MLP neural network for classifying microstrip antennas. This approach enabled the identification of antenna designs suitable for further optimization in an efficient and cost-effective manner compared to existing solutions reported in [34, 35]. By optimizing both training and architectural parameters, the proposed method achieved higher accuracy than previously reported models while maintaining a relatively moderate number of parameters, confirming the effectiveness of Bayesian optimization in this context.

Regarding limitations, the model was trained on a specific dataset of pseudo-random antenna designs, which may limit its generalization to other geometries or frequency ranges. The proposed model is intended for freeform antenna geometries with high design variability, and its generalization should therefore be interpreted within this specific domain. Furthermore, the dataset is restricted to pseudo-randomly generated designs and a predefined frequency range, which may limit applicability to other antenna types or operating conditions not considered in this study.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

In addition, only an MLP architecture was investigated. Other machine learning architectures could provide further insights, potentially improving performance.

Another limitation is that using Bayesian optimization increases the computational cost during the training stage due to multiple optimization trials. However, this cost is incurred offline and is of secondary importance, as the primary objective is to identify an effective model configuration. Moreover, Bayesian optimization enables more efficient exploration of the hyperparameter space compared to grid search and typically yields better results than random search.

Future work will follow three main directions. First, further optimization of the remaining model parameters will be conducted to improve classification performance. Second, alternative architectures, including convolutional neural networks, will be investigated and optimized using Bayesian approaches. Third, the proposed classifier will be integrated with generative models, such as conditional variational autoencoders, to support automated antenna design and validation pipelines.

Acknowledgments. The authors declare that no funding was received for this research. Michał Wyciszkiewicz: Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. Bartosz Czaplewski: Conceptualization, Data curation, Project administration, Supervision, Writing – review & editing.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Volkov, A., Sobko, S., Sviridov, I., Baskakov, A., Fride, D.: Development of a Neural Network Module for Detecting Network Threats in Traffic Based on Convolutional and Recurrent Neural Networks. In: 2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon). pp. 28–31. IEEE, Saint Petersburg, Russian Federation (2024). https://doi.org/10.1109/ElCon61730.2024.10468509.
2. Jin, Y., Bao, J., Li, H., Yun, K., Yang, H.: Secure Communication Algorithm of Wireless Network Based on BP Neural Network. In: 2024 International Conference on Data Science and Network Security (ICDSNS). pp. 1–5. IEEE, Tiptur, India (2024). https://doi.org/10.1109/ICDSNS62112.2024.10691139.
3. M, B.B.P., M, A., J, D., S, S.: Byte Segmented Convolutional Neural Network for Network Traffic Classification. In: 2023 1st International Conference on Optimization Techniques for Learning (ICOTL). pp. 1–6. IEEE, Bengaluru, India (2023). https://doi.org/10.1109/ICOTL59758.2023.10434992.
4. Ahmed Naqvi, S.F., Malik, M.I., Mishra, A.K., Burse, K., Chagre, S., Rathor, G.P.: Hybrid Quantum AI and Classical Machine Learning Model for Telecommunication Network Traffic Analysis. In: 2025 International Conference on Engineering Innovations and Technologies (ICoEIT). pp. 1223–1228. IEEE, Bhopal, India (2025). https://doi.org/10.1109/ICoEIT63558.2025.11211783.
5. Ogbodo, E.U., Abu-Mahfouz, A.M., Kurien, A.A.: 5G RedCap Enhancement Towards Improved Cellular LPWAN/5G-IoT for Smart Cities and Industrial IoT Using Genetic Algorithm-Based Neural Network. Int. J. Sens. Wirel. Commun. Control. 15, 14–29 (2024). https://doi.org/10.2174/0122103279312219240520053431.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

6. Zang, Y., Zhang, Z., Li, S., Zhang, F., Chen, H.: Fiber neural networks for the intelligent optical fiber communication signal processing. iOptics. 1, 100009 (2025). https://doi.org/10.1016/j.iopt.2025.100009.
7. Luo, Y., Li, C., Yan, N., An, W., Ma, K.: Design of Multibeam Compressed High-Order Mode Antenna Based on Multitask Artificial Neural Network. IEEE Antennas Wirel. Propag. Lett. 25, 781–785 (2026). https://doi.org/10.1109/LAWP.2025.3639707.
8. Bod, M., Hassani, H.R., Taheri, M.M.S.: Compact UWB Printed Slot Antenna With Extra Bluetooth, GSM, and GPS Bands. IEEE Antennas Wirel. Propag. Lett. 11, 531–534 (2012). https://doi.org/10.1109/LAWP.2012.2197849.
9. Bekasiewicz, A., Kurgan, P., Koziel, S.: Numerically Efficient Miniaturization-Oriented Optimization of an Ultra-Wideband Spline-Parameterized Antenna. IEEE Access. 10, 21608–21618 (2022). https://doi.org/10.1109/ACCESS.2022.3152736.
10. Bekasiewicz, A., Dzwonkowski, M., Dhaene, T., Couckuyt, I.: Specification-Oriented Automatic Design of Topologically Agnostic Antenna Structure. In: Franco, L., De Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., and Sloot, P.M.A. (eds.) Computational Science – ICCS 2024. pp. 11–18. Springer Nature Switzerland, Cham (2024). https://doi.org/10.1007/978-3-031-63759-9_2.
11. Bekasiewicz, A., Askaripour, K.: Performance Comparison of Automatically Generated Topologically Agnostic Patch Antennas. In: 2024 25th International Microwave and Radar Conference (MIKON). pp. 143–146. IEEE, Wroclaw, Poland (2024). https://doi.org/10.23919/MIKON60251.2024.10633953.
12. Jacobs, J.P.: Accurate Modeling by Convolutional Neural-Network Regression of Resonant Frequencies of Dual-Band Pixelated Microstrip Antenna. IEEE Antennas Wirel. Propag. Lett. 20, 2417–2421 (2021). https://doi.org/10.1109/LAWP.2021.3113389.
13. Wu, G.-B., Zeng, Y.-S., Chan, K.F., Chen, B.-J., Qu, S.-W., Chan, C.H.: High-Gain Filtering Reflectarray Antenna for Millimeter-Wave Applications. IEEE Trans. Antennas Propag. 68, 805–812 (2020). https://doi.org/10.1109/TAP.2019.2943432.
14. Bekasiewicz, A., Koziel, S., Plotka, P., Zwolski, K.: EM-Driven Multi-Objective Optimization of a Generic Monopole Antenna by Means of a Nested Trust-Region Algorithm. Appl. Sci. 11, 3958 (2021). https://doi.org/10.3390/app11093958.
15. Whiting, E.B., Campbell, S.D., Mackertich-Sengerdy, G., Werner, D.H.: Dielectric Resonator Antenna Geometry-Dependent Performance Tradeoffs. IEEE Open J. Antennas Propag. 2, 14–21 (2021). https://doi.org/10.1109/OJAP.2020.3037826.
16. Lizzi, L., Viani, F., Azaro, R., Massa, A.: Optimization of a Spline-Shaped UWB Antenna by PSO. IEEE Antennas Wirel. Propag. Lett. 6, 182–185 (2007). https://doi.org/10.1109/LAWP.2007.894157.
17. Alroughani, H., McNamara, D.A.: The Shape Synthesis of Dielectric Resonator Antennas. IEEE Trans. Antennas Propag. 68, 5766–5777 (2020). https://doi.org/10.1109/TAP.2020.2988984.
18. Koziel, S., Bekasiewicz, A.: Multi-Objective Design of Antennas Using Surrogate Models. World Scientific (2017). https://doi.org/10.1142/q0043.
19. Koziel, S., Ogurtsov, S.: Multi-Objective Design of Antennas Using Variable-Fidelity Simulations and Surrogate Models. IEEE Trans. Antennas Propag. 61, 5931–5939 (2013). https://doi.org/10.1109/TAP.2013.2283599.
20. Koziel, S., Bekasiewicz, A.: Fast multi-objective design optimization of microwave and antenna structures using data-driven surrogates and domain segmentation. Eng. Comput. 37, 753–788 (2020). https://doi.org/10.1108/EC-01-2019-0004.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

21. Koziel, S., Bekasiewicz, A.: Domain segmentation for low-cost surrogate-assisted multi-objective design optimisation of antennas. IET Microw. Antennas Propag. 12, 1728–1735 (2018). https://doi.org/10.1049/iet-map.2017.0635.
22. Koziel, S., Bekasiewicz, A.: Comprehensive Comparison of Compact UWB Antenna Performance by Means of Multiobjective Optimization. IEEE Trans. Antennas Propag. 65, 3427–3436 (2017). https://doi.org/10.1109/TAP.2017.2700044.
23. Koziel, S., Bekasiewicz, A.: Rapid Simulation-Driven Multiobjective Design Optimization of Decomposable Compact Microwave Passives. IEEE Trans. Microw. Theory Tech. 64, 2454–2461 (2016). https://doi.org/10.1109/TMTT.2016.2583427.
24. Koziel, S., Bekasiewicz, A.: Multi-objective optimization of expensive electromagnetic simulation models. Appl. Soft Comput. 47, 332–342 (2016). https://doi.org/10.1016/j.asoc.2016.05.033.
25. Bengherbia, B., Bougheriza, M.E.A., Bouhedda, M., Tobbal, A., Toubal, A., Hentabeli, H.: Implementation of a 3D Indoor Positioning System Using MLP-based Distance Estimation for IoT Applications. In: 2022 International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT). pp. 209–214. IEEE, Jember, Indonesia (2022). https://doi.org/10.1109/ICEECIT55908.2022.10030667.
26. Efendi, R., Wahyono, T., Widiasari, I.R.: Hybrid DBSCAN-SMOTE with Attention-Augmented MLP for Detecting Low-Frequency DDoS Attacks in Imbalanced Network Traffic. In: 2025 13th International Conference on Awareness Science and Technology (iCAST). pp. 1–6. IEEE, Yogyakarta, Indonesia (2025). https://doi.org/10.1109/iCAST68191.2025.11300066.
27. Montassar, A., Delleji, T., Slimeni, F., Abid, M.A., Attia, R.: RF Classification of UAVs Using MLP Model and Software Defined Radio. In: 2024 IEEE/ACS 21st International Conference on Computer Systems and Applications (AICCSA). pp. 1–9. IEEE, Sousse, Tunisia (2024). https://doi.org/10.1109/AICCSA63423.2024.10912631.
28. Kang, P.: Malicious encrypted traffic detection based on Bert and one-dimensional CNN model. In: 2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI). pp. 282–287. IEEE, Changchun, China (2023). https://doi.org/10.1109/ICETCI57876.2023.10176917.
29. Badawadagi, S.S., Kumar. Cm, M.: SliceNet: A Hybrid CNN Model for Network Slicing in 5G Communications. In: 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT). pp. 1–5. IEEE, Trichirappalli, India (2024). https://doi.org/10.1109/ICEEICT61591.2024.10718550.
30. Ravindra L. Pardhi: Fault Identification in Optical Transport Network Using CNN. J. Electr. Syst. 20, 2151–2156 (2024). https://doi.org/10.52783/jes.3934.
31. A, A., S, K., Saranya, P.C., Arumugam, D.: Securing Smart Grid IoT System using A Robust RNN based Cyberattack Detection Framework. In: 2025 International Conference on Electronics and Renewable Systems (ICEARS). pp. 617–622. IEEE, Tuticorin, India (2025). https://doi.org/10.1109/ICEARS64219.2025.10940607.
32. Bartels, J., Hagihara, A., Minati, L., Tokgoz, K.K., Ito, H.: An Integer-Only Resource-Minimized RNN on FPGA for Low-Frequency Sensors in Edge-AI. IEEE Sens. J. 23, 17784–17793 (2023). https://doi.org/10.1109/JSEN.2023.3286580.
33. Dabool, H., Alashwal, H., Alnuaimi, H., Alhouqani, A., Alkaabi, S., Al Ahbabi, A.: Comparative Analysis of Hyperparameter Tuning Methods in Classification Models For Ensemble Learning. In: 2024 7th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI). pp. 1–5. IEEE, Guangzhou, China (2024). https://doi.org/10.1109/ACAI63924.2024.10899492.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9

34. Czaplewski, B.: Neural Network for Evaluating the Operational Range of Antennas with Randomly Generated Designs. In: Lees, M.H., Cai, W., Cheong, S.A., Su, Y., Abramson, D., Dongarra, J.J., and Sloot, P.M.A. (eds.) Computational Science – ICCS 2025. pp. 394–402. Springer Nature Switzerland, Cham (2025). https://doi.org/10.1007/978-3-031-97635-3_47.
35. Czaplewski, B.: A convolutional neural network for predicting the frequency range of antennas based on randomized designs. Eng. Appl. Artif. Intell. 165, 113561 (2026). https://doi.org/10.1016/j.engappai.2025.113561.
36. Huang, J., Li, W., He, Y., Zhang, L., Wong, S.-W.: Optimization of Antenna Design Using the Artificial Neural Network and the Simulated Annealing Algorithm. In: 2021 Computing, Communications and IoT Applications (ComComAp). pp. 119–122. IEEE, Shenzhen, China (2021). https://doi.org/10.1109/ComComAp53641.2021.9653055.
37. Kouhalvandi, L., Aygun, S., Ozoguz, S., Matekovits, L., Karamzadeh, S.: Surrogate Modeling with Hybrid CNN-RNN and GAN Architectures for Antenna Designs. In: 2025 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC). pp. 307–310. IEEE, Palermo, Italy (2025). https://doi.org/10.1109/APWC65665.2025.11190408.
38. Liu, X., Li, P., Zhu, Z.: Bayesian optimisation-Assisted Neural Network Training Technique for Radio Localisation. In: 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). pp. 1–5. IEEE, Helsinki, Finland (2022). https://doi.org/10.1109/VTC2022-Spring54318.2022.9860977.
39. Hu, P., Ukil, A., Nair, N.-K.C.: Photovoltaic Maximum Power Point Tracking Based on Bayesian Optimization Neural Network. In: IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society. pp. 1–6. IEEE, Singapore, Singapore (2023). https://doi.org/10.1109/IECON51785.2023.10312289.
40. Czaplewski, B.: 1Ddesign-to-label-1.8.3 - Antenna Designs with Operating Frequency Range Labels (1.8.3) [Data set]. Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics (2025). https://doi.org/10.5281/zenodo.14982975.
41. Pan, J., Fu, H., Qiu, K., Li, W., Xie, J.: Sensitivity-analysis guided Bayesian optimization for crystal plasticity parameter identification. Int. J. Mech. Sci. 311, 111205 (2026). https://doi.org/10.1016/j.ijmecsci.2026.111205.
42. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2623–2631. ACM, Anchorage AK USA (2019). https://doi.org/10.1145/3292500.3330701.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_9