Trends of Microwave Devices Design Based on Artificial Neural Networks: A Review

Andrius Katkevičius 1, Darius Plonis 1, Robertas Damaševičius 2 and Rytis Maskeliūnas 3,*

1 Faculty of Electronics, Vilnius Gediminas Technical University, Naugarduko Str. 41, 03227 Vilnius, Lithuania; andrius.katkevicius@vgtu.lt (A.K.); darius.plonis@vgtu.lt (D.P.)
2 Faculty of Applied Mathematics, Silesian University of Technology, ul. Kaszubska 23, pok. 307A, 44-100 Gliwice, Poland; robertas.damasevicius@polsl.pl
3 Department of Multimedia Engineering, Kaunas University of Technology, Baršausko 59, 51423 Kaunas, Lithuania
* Correspondence: rytis.maskeliunas@ktu.lt

Abstract: The usage of techniques of the artificial neural networks (ANNs) in the field of microwave devices has recently increased. The advantages of ANNs in comparison with traditional full-wave methods are that the prediction speed when the traditional time-consuming iterative calculations are not required and also the complex mathematical model of the microwave device is no longer needed. Therefore, the design of microwave device could be repeated many times in real time. However, methods of artificial neural networks still lag behind traditional full-wave methods in terms of accuracy. The prediction accuracy depends on the structure of the selected neural network and also on the obtained dataset for the training of the network. Therefore, the paper presents a systematic review of the implementation of ANNs in the field of the design and analysis of microwave devices. The guidelines for the systematic literature review and the systematic mapping research procedure, as well as the Preferred Report Items for Systematic Reviews and Meta-Analysis statements (PRISMA) are used to conduct literature search and report the results. The goal of the paper is to summarize the application areas of usage of ANNs in the field of microwave devices, the type and structure of the used artificial neural networks, the type and size of the dataset, the interpolation and the augmentation of the training dataset, the training algorithm and training errors and also to discuss the future perspectives of the usage of ANNs in the field of microwave devices.

Keywords: artificial neural networks; microwave devices; analysis; syntheses; computer-based modeling

1. Introduction

The microwave industry is evolving rapidly in order to meet the modern needs of today [1–3]. There are presented many new structures of antennas [4–6], antenna arrays [7,8], filters [9,10], phase shifters [11,12], different resonators [13], delay lines [14], sensor arrays [15] and others. The main research directions are usually considered to be reconfigurability or tenability and miniaturization by searching for new constructions [16] or materials [17]. The miniaturization is especially important for the growing network of internet of things (IoT) devices [18,19].

The growing research in the field of microwave devices is also increasing the need to search for the new modeling techniques [20,21], which could be divided into the methods of synthesis [22,23] and analysis [24]. The conventional currently used methods are based on the usage of different forms of Maxwell’s equations. The most accurate results could be obtained by using analytical methods. The calculations are fast with analytical methods. On the other hand, it is usually quite difficult or impossible to create an accurate mathematical model of the desired microwave devices and the method solves only a partial case of a microwave device with certain exceptions [25]. Numerical methods are more universal.
Usually, the commercial software packages are used, which allows to create the 3D models of the microwave devices, but the calculations are time consuming. Calculations can take hours or even days when the structure of a microwave device is complex [26,27].

The possible migration from conventional methods to the methods based on the ANNs in the field of microwave devices was started to be investigated based on the disadvantages, which are mentioned above [28,29]. The usage of ANNs seems promising because ANNs are already successfully used in other fields, such as image processing [30], sound processing [31], and the forecasting of financial markets [32]. The advantage of artificial neural networks (ANNs) in this case is the speed of prediction. ANNs do not perform any calculations, but predict the results. The prediction is usually very fast. Therefore, the procedure of analysis or synthesis of microwave devices is performed significantly faster [33,34]. The prediction could be performed thousands of times faster than calculations by the conventional methods [35]. Therefore, methods based on the the ANNs could be used in real-time systems for solving complex problems. Additionally, these complex problems can be solved in embedded systems in real time [36,37].

On the other hand, it is also necessary to mention that before prediction, the ANN should firstly be trained [38]. The training could be with or without the supervisor and classified into supervised [39], semi-supervised [40] and unsupervised networks [41]. It depends on the type of the network and training dataset [42]. There are many different types of networks, starting from the simple one-layer perceptron network and continuing until convolutional networks [43]. All the ANNs address two challenges, which could be divided into the task of function approximation (regression) or pattern recognition (classification). The function approximation is used in the field of microwave devices [44].

The review of different types and structures of ANNs, transition algorithm from full-wave methods to the ANN-based methods, training datasets and applications areas of ANNs in the field of microwave devices is presented in this paper. Overall, the review consists of 113 articles. Forty-four articles substantiate the introduction. The principles of ANNs are discussed using three articles. The transition algorithm is discussed in all articles. The discussion of different types of ANNs, which could be used in the modeling of microwave devices, is included in 32 articles. The training types are discussed in 12 articles. Fuzzy logic is mentioned in 10 articles. The future perspectives are outlined in the end of the presented articles.

2. Research Methodology

The guidelines for systematic literature review and the systematic mapping research procedure, as well as the Preferred Report Items for Systematic Reviews and Meta-Analysis statements (PRISMA) [45], were used to conduct and report the reviews. This systematic review is based on a well-planned research method that assures a thorough and unbiased selection of all peer-reviewed publications connected to published research material. This methodology is used to collect relevant articles from trustworthy scientific sources, which are then sorted and mapped into numerous categories to indicate the current level of research in the use of failure detection technologies. This research map will be extremely helpful to practitioners and researchers in identifying cutting-edge areas and subjects for future study.

As a result, it is critical to stress that the goal of this review is to understand not only the use cases or applications of ANNs in the field of the microwave devices, but also the limitations and constraints of applying suitable methodologies. Furthermore, the most recent trends in technological techniques, processes, and concepts employed in the execution of these methods are discussed.

2.1. Research Design

The present research needs are provided in this subsection by identifying the preliminary research results based on the research question and keywords connected to the research topic.
2.1.1. Literature Review Questions

Machine-learning-based solutions for the efficient and dynamic design of microwave devices have taken a long time to develop. Over the years, several procedures, approaches, and strategies have been created to characterize the aspects involved in the growth of microwave devices. As a result, the following issues are addressed in this research:

**RQ1** What types of neural networks can be applied to the design of microwave devices?

**RQ2** What are the applicable transition algorithms from full-wave methods to the neural networks methods?

**RQ3** What are the applications and direction in microwave design using machine learning?

2.1.2. Research Process

Several bibliographic databases were used to look for and acquire publications for this investigation. These sources were chosen based on their track record of success. Table 1 shows the origins of the publications used as references in this study. This database can give the most significant papers and conferences pertinent to acoustical failure detection, as well as their entire text.

**Table 1. Online databases.**

<table>
<thead>
<tr>
<th>Database</th>
<th>URL</th>
</tr>
</thead>
</table>

2.1.3. Search Terms

After executing the initial search phase in the research databases by inputting keywords, an extra scanning step was performed to confirm the correctness of the research process and that the selection of studies related to the present research topic and task fit the requirements. Search engines were also used in this study to aid in the search for related research.

1. “Artificial neural network microwave devices” OR “computer based modelling microwave devices”;
2. “Design microwave devices”;
3. “Modeling microwave devices”;
4. “Application microwave devices”;
5. “Synthesis microwave devices”;
6. “Microwave devices”.

2.2. Review Conduction

This section explains the methods used to create the systematic literature review procedure. The SLR search process is influenced by the guidelines and frameworks used to create this article.

2.2.1. Selection of Relevant Papers

After acquiring exploratory research studies relevant to the study objectives, the obtained articles should be appraised for relevancy. As a consequence, a second evaluation was conducted in order to establish the relevance of the chosen first study. Furthermore, following the first screening, a random systematic review of the selected papers was performed to confirm the consistency of the inclusion and exclusion criteria. The research selection approach for the current systematic review is depicted in Figure 1, following the
Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology [45].

The following steps were taken to identify relevant research studies:
1. Use the provided terms to search the database and locate prior works linked to the research.
2. Ignore documents that do not meet the supplied search parameters.
3. Exclude papers with no evident link between title and abstract.
4. Read the articles in their entirety before evaluating them.
5. Assess the bibliography.
6. Perform the preliminary research.

2.2.2. Inclusion and Exclusion Criteria

Exclusion criteria include research publications that are unrelated to machine-learning-based microwave design or are based on traditional design methodologies, and so fall beyond the scope of this research study. This study focuses on SLR research publications pertinent to this topic. Furthermore, the study excluded similar studies on the same issue. As a consequence, the inclusion and exclusion criteria used in writing the SLR are shown in Table 2.

Table 2. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusivity criteria</th>
<th>Exclusivity criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Peer-reviewed original articles</td>
<td>1 Articles that are not written in English</td>
</tr>
<tr>
<td>2 Articles proposing an neural network based microwave design</td>
<td>2 Studies with invalidated techniques and algorithms</td>
</tr>
<tr>
<td>3 Articles that utilize other machine learning based microwave design methods</td>
<td>3 Articles that utilize neural network design on other purposes</td>
</tr>
<tr>
<td>4 Articles that present application of machine learning based microwave designs</td>
<td>4 Articles that do not utilize microwave design</td>
</tr>
<tr>
<td>5 Recency of articles in case of multiple repeated studies</td>
<td>5 Articles that do not clearly mention microwave in the title</td>
</tr>
<tr>
<td></td>
<td>6 Articles providing unclear results or findings</td>
</tr>
<tr>
<td></td>
<td>7 Duplicated studies</td>
</tr>
</tbody>
</table>

Figure 1. Procedure of research selection for the present schematic review based on PRISMA methodology [45].
2.2.3. Data Extraction

During the data extraction procedure, relevant information was collected from the articles and entered into a database. This database is made up of the elements in Table 3 [46].

<table>
<thead>
<tr>
<th>Data Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Article title</td>
</tr>
<tr>
<td>Year</td>
<td>Year of publication</td>
</tr>
<tr>
<td>Author(s)</td>
<td>The article author(s)</td>
</tr>
<tr>
<td>Publication type</td>
<td>Journal, Proceeding, etc.</td>
</tr>
<tr>
<td>Country</td>
<td>Researchers’ affiliation country</td>
</tr>
<tr>
<td>Contribution</td>
<td>The major contribution of the article</td>
</tr>
<tr>
<td>Summary</td>
<td>Summary of the article from our perspective</td>
</tr>
</tbody>
</table>

3. Transition from Full-Wave Methods to the Neural Networks Based Methods

To start with, artificial neural networks are the computer models of biological neural networks in the human brain. An ANN is a collection of linked artificial neurons that may impact each other’s behavior. The detailed information about the structure of the single neuron, weights between the single neurons, possible activation functions is provided in [47]. There are several types of activation functions that may be used to adapt to diverse nonlinear real-world tasks. This is significant since the majority of real-world input is nonlinear and there is a need for neurons to learn nonlinear representations [48].

Individual neurons are then linked to neural networks, which can have a variable number of layers and neurons in each layer, starting from the single-layer perceptron network without the back propagation until the complex networks, such as convolution or deep neural networks, which all could be used in the field of modeling of microwave devices [49].

The transition procedure of the modeling of microwave devices from calculations using the full-wave methods to the prediction using ANNs methods consists overall of four stages. First of all, it is necessary to investigate the microwave devices using the full-wave methods in the first stage. The analytical methods or the numerical methods could be used in the first stage. The analytical methods require to make the mathematical model of the microwave device and to make the calculation by yourself. It requires a lot of experience and time in order to make the mathematical model every time the microwave device is modified. Numerical methods are usually used with commercial software packages such as Sonnet©, CST Microwave Studio©, HFSS© and others. Numerical methods usually require a lot of time and computer resources.

The review of the newest articles showed that usually the 3D model with a specific set of geometrical and physical parameters is selected in the first stage (1.1 block). The model is drawn in the commercial software package (1.2 block). The S parameters are obtained for the analysis (1.3 block) (Figure 2) [50].

The modeling is repeated many times in order to obtain a significant amount of data for training with the different sets of parameters.

The obtained data are divided into the testing dataset (2.1 block) and training dataset (2.2 block) in the second stage. It is very important to separate the data correctly. Usually about 70% of data is used for training and only 30% of data is used for the testing. It is very important to use unique unused data for the testing in order to obtain correct and accurate information about the training of the neural network. It is also very important that the values of important parameters from the training and test datasets are distributed throughout the relevant range of values. Neural networks have the property of getting lost when receiving datasets from an unseen range of values. It is also very important to select the correct structure of the neural network in order to have accurate predictions (2.3 block).
The training should be performed in the third stage of the transition algorithm. The mean squared error (MSE) is calculated during the training procedure in order to know the quality of training (3.1 block). The trained surrogate model is obtained when the network is trained and the MSE is less than the desired threshold (3.2 block).

The model of the trained neural network is tested with the testing dataset in the fourth stage of the transition algorithm (3.3 block). The model of ANN is considered final and can be used for the prediction of the parameters of the microwave device if accuracy with
the testing dataset is appropriate. If the MSE is higher than the desired threshold, it is necessary to go back to the second or even first stage and modify the structure of the neural network or to collect additional data for training [51].

4. Different Fields of Usage

ANNs are used for the design and analysis of different microwave devices: antennas [52], antennas arrays [53], filters [54], phase shifters [55], resonators [56], microwave circuits [57], traveling wave tubes [58], delay lines and others [59]. The application areas of ANNs are discussed in this section. The discussion is presented by taking into account the following factors: the model of the microwave device, the application area, the type and structure of ANN, the type and size of the training dataset, the interpolation and augmentation of the dataset, the training algorithm and training errors, and the accuracy of training and prediction.

4.1. Antennas

Antennas are one of the largest and rapidly evolving groups of microwave devices. For example, the substrate-integrated waveguide patch antenna, which works in the 12–18 GHz frequency range, is presented in [60]. The resonant frequency of the antenna is equal to 16.10 GHz. The return losses are less than $-10 \, \text{dB}$ to $-19 \, \text{dB}$ in the working frequency range. The overall dimensions of the antenna are equal to $24 \times 15.4 \, \text{mm}^2$. The thickness of the substrate is equal to 0.95 mm. The dielectric with $\varepsilon_r = 3.2$ is used for the substrate with a $\delta = 0.0018$ tangent loss.

Vilovich et al. [61] sought to investigate neural network capabilities in designing the rectangular patch (width and length) of a microstrip antenna. Their comparison analysis revealed that the neural-network-generated antennas performed better.

The multilayer perceptron (MLP) network with the back-propagation algorithm was used for the optimization of the dimensions of the presented patch antenna. The structure of the MLP network consists of three layers. The input layer has two neurons for the input of parameters of return loss $S_{11}$ and resonant frequency $f_0$. The network has predicted dimensions of the width $W$ of microstrip line, diameter $D$ of the patch, inner radius $R_1$ and outer radius $R_2$. The hidden layer consists of 15 neurons. The neural network allowed optimizing the initial set of dimensions of the patch antenna.

Another example of the microstrip patch antenna is presented in [62]. This time, the antenna had an elliptical form. The dimensions of the three ellipses were constructed in order to resonate at 2.4 GHz. The overall dimensions of the substrate of the elliptical antenna were equal to $100 \times 60 \, \text{mm}^2$. The similar feed-forward back-propagation ANN (FFBP-ANN) along with the Levenberg–Marquardt optimization algorithm was used to model the antenna design. The structure of the network was equal to 3-10-2. The network predicted the return loss $S_{11}$ and gain of the antenna while changing the radius geometrical parameters in the input. The overall 160 combinations of the input parameter with the corresponding return loss and gain were collected using the CST©EM simulation. A total of 100 combinations were used for the training. Then, 60 combinations were left for the testing. Results showed good agreement between the results obtained with CST© and the results predicted with the MLP network.

The rectangular microstrip patch antenna for the terahertz applications is presented in [63]. The gallium arsenide $\varepsilon_r = 12.9$ material was used for the substrate. The overall dimensions of the substrate were equal to $0.06 \times 0.035 \, \text{mm}^2$. The structure of the feed-forward MLP network was equal to 3-20-2. The parameters of frequency $f$ (0.3–3 THz), dielectric constant $\varepsilon$ (10.2–12.9) and height $h$ (0.005–0.3 mm) were submitted to the input of the network. The return loss $S_{11}$ and bandwidth $W$ were predicted in the output of the network. The Levenberg–Marquardt training algorithm was used for the training. The 136 samples were collected using the HFSS©EM simulation.

All three earlier mentioned examples used the MLP network. The first example solved the synthesis optimization task when the desirable electrical parameters were known and
it was necessary to optimize the geometrical dimensions of the antenna starting from the initial selected dimensions. The second and third examples solved the analysis task when it was necessary to find the $S$ parameters for the description of the antenna.

The optimization procedure of the aperture coupled single and multilayer microstrip planar antennas is described in [64]. The calculation of microstrip planar antenna parameters using traditional full-wave methods is time consuming. Therefore, the optimization technique using ANNs is presented. Several different structures of the feed-forward neural network were used for the investigation. The input-hidden-output (3-20-2) network topology was used for the single layer planar antenna. The inputs of the network were of a dielectric constant $\varepsilon_r$, height of substrate $h$ and resonant frequency $f_r$. The parameters of the scattering matrix were predicted in the output layer. The input-hidden-hidden-output (7-20-20-2) network topology was used for the multilayer planar antenna. The parameters of frequency of operation, slot length, slot width, stub length, patch offset, patch width and feed width were given to the input of the network in order to predict the same parameter of the scattering matrix. The two hidden layers were selected because of the bigger non-linearity between the input and output parameters. The nonlinear sigmoid transfer function was used for the hidden neurons. The linear function was chosen for the input and output neurons. The error back-propagation algorithm was selected for the training. The number of neurons in the hidden layers was selected iteratively by increasing the number of neurons from a small number until the acceptable error was received.

Datasets for training were generated using the full-wave simulators. The presented model of neural network was able to predict the resonant length of the aperture coupled antenna in approximately 33 ms, while the computations with full-wave methods took about 20 s.

The main goal of the paper [65] was to improve the slip-ups in the radiation case of the antenna using ANN framework estimations. The T-shaped planar antenna for the 5G faraway application at 38 GHz was used in the investigation. This presented antenna is suitable for the millimeter wave rehash. The feed-forward neural network with back propagation was used in the investigation.

The main goal in [66] was to increase the efficiency of the modeling of two models of the multi-band PIFA planar antennas. Both models had a similar structure when the two folded inverted-L monopoles were printed on the plastic substrate $\varepsilon_r = 2.9$. The initial data for the training of the neural network were obtained with the HFSS© software package. The feed-forward multi-layer perceptron neural network (2-32-32-1) was used for the prediction. The frequency and the scattering $|S_{11}|$ values of the standalone antenna were considered input parameters. The $|S_{11}|$ value of the antenna in real mobile was considered the desired output. The prediction with MLP neural network allowed to reduce the processing time from several hours to 1.5 s. The prediction errors were equal to 3.43% and to 4.7%, respectively. The presented comparison of the $S_{11}$ was in the range of 0–2.5 GHz frequency.

The radial basis function (RBF) neural network was chosen in order to accurately predict the resonant frequency of the adjustable circular microstrip antenna while varying the patch radius, relative permittivity of the upper dielectric substrate, the air gap separation and thickness of the upper dielectric substrate [67]. The radial basis activation functions with a Gaussian form was used in every single neuron of the hidden layer. The overall $4 \times 1200$ samples dataset was collected for the neural network training and validation. The results, which were predicted with RBF and calculated with traditional full-wave methods, do not vary more than 100 MHz.

Ali et al. [68] presented a closed-loop antenna with an interdigital capacitor to enhance the electric field fringing at the patch core as part of wearable devices with little back radiation toward the human body. To maximize the suggested antenna performance in the desired frequency band, a neural network was used. The error percentage attained was fewer than 4.4 percent.

A completely different solution is presented in [69]. The parametric modeling of the ultra-wideband notched antenna was performed using the model of the convolutional
neural network (CNN). The main difference in this case is that the inputs to the neural network were not the geometrical parameters of the antenna but images with different topologies of the antenna. The overall dimensions of the used antenna were equal to $30 \times 35 \text{ mm}^2$. The working frequency range of the presented antenna was adjusted by changing the overall four parameters of the size and the position of the coupling strip. The performance of the antenna was investigated in the 2–12 GHz frequency range. The presented CNN consisted of four convolutional layers and two fully connected layers. Every convolutional layer consisted of convolutional and detector stages. The rectified linear unit (ReLU) was used in the detector stage, and batch normalization was employed at each layer. The inputs of the CNN were the cross-sectional binary images. Each value of the separated pixel represented different material. The substrate was represented by 0, while the coupling strip was represented by 1. The image with the coupling strip was partitioned into $415 \times 214$ grids. Overall, there were collected 64 training samples and 49 samples for testing purposes. The samples were collected using the commercial HFSS© software package. The initial smaller number of samples was increased using the data augmentation.

4.2. Antennas Arrays

Antenna arrays is another field of microwave devices in which the usage of intelligent methods grows rapidly. The ANNs are usually used in order to solve the tasks of beamforming, finding the angle of arrival (AoA), real-time analysis while saving the computational resource, defective elements determination, and fault finding in antenna arrays.

The innovative approach of beamforming of the phased antenna array using a convolutional neural network is presented in [70]. The position of a beam is crucial for beam synthesis, and the neural network’s phase output must be sensitive to the spatial location of the desired beam in the input. Using the desired radiation pattern as input, the convolutional neural network is trained to compute phases of patch antennas in the $8 \times 8$ patch antenna array. The presented neural network consists of eight layers: four convolutional layers and four fully connected layers. The CNN beamformer receives a two-channel radiation pattern as input. The first and second channels show the actual gain in linear or dBi scales.

The leaky rectified linear unit (ReLU) was employed for the activation of all layers. There was no activation at the output layer because this was a regression problem rather than a classification one. The Adam optimization technique was utilized in order to update the network weights. The loss function was used in order to calculate the mean squared error (MSE). The flattened output of the final convolution layer had 82,800 neurons. ANSYS HFSS© software was used in order to collect the data for training and verification. The network was trained using 165,000 samples. The validations were made with 40,000 unique samples.

The beamforming design based on the deep learning neural network is presented in [71]. The presented beamforming neural network (BFNN) allowed to significantly improve the performance of beamforming in comparison with traditional beamforming algorithms. The improvement of beamforming was understood as the maximization of the spectral efficiency (SE), while having hardware limitation and imperfect channel state information (CSI). The authors also proposed the two-stage design approach to make the BFNN robust to imperfect CSI. The BFNN learned how to approach the ideal SE during the first offline training stage. The BFNN adapted itself to imperfect CSI and achieved robust performance to the channel estimation errors in the second online deployment stage.

For the 64 element linear antenna array, the authors used the deep neural network consisting of five layers: input layer, three dense layers and the final lambda layer. The self-defined Lambda layer was added at the end of the BFNN in order to ensure that the output of the BFNN was a complex-valued vector satisfying the constant modulus constraint. The ReLU activation function was used in the first to dense layers. The sigmoid activation function was used in the third dense layer.
The usage of the model of the multilayer direct distribution neural network for the beamforming and calibration applications in the linear omnidirectional antenna array is presented in [72]. The antennas were positioned at a half wavelength from each other without the interaction between the emitters. The main discussed arising issues were the problem of possible non-identical characteristics of the antenna channels and drift of characteristics over time.

The authors provided slightly different models of direct distribution neural network in order to solve the beamforming and calibration problems. The number of inputs of the first layer was equal to the number of antenna elements of the aperture in the beamforming example. The models of the digital signal from the ADC output were formed as input data to the network. The main important parameters were presented as relative location, frequency, waveform of the signal’s source, the interference source and the noise level. The authors used the network topology with one hidden layer. The digital beam signals were created in the output. The beamforming output spectrum components and pilot signals must also be fed into the input of neural network additionally in the case of the antenna array calibration problem. The supervised training was carried out during both cases.

The determination of the angle of arrival (AoA) is another use area for antenna arrays. The measuring of the direction from which a received signal is emitted is known as radio direction finding. The mapping of the connections between the properties of received signals and their incidence direction is the basis of radio direction finding. Radio direction finding is an inverse task of signal reception from a given direction. It is critical to be able to efficiently track the intended consumers. Traditional strategies are inefficient for achieving a super resolution AoA estimate in real time. As a result, machine learning approaches are typically used in order to increase the efficiency of direction finding, such as the processing speed, and the resolution of the angle [73].

Usually the problem appears to be working with practical antenna arrays, which have nonuniform elements. The arising problem could be solved with ANNs which approximate the mapping from received signals to AoAs with high estimation accuracy. An example of a practical antenna array with 30 elements is considered in [74].

The uniqueness of [75] is the nonlinear mapping of outputs of the receiving antennas with the associated direction of arrival (DoA) by using the combination of the detection network and the DoA estimation ANNs that allow to estimate the DoA. The detection network allowed to reduce the size of the training dataset and to individually train several deep neural networks that correspond to the different sectors of possible position. The detection network was used in order to divide the search area of the antenna array into different position sectors and to detect signals radiating from sources in each sector. The MLP-based network was used for both the detection and DoA estimation networks. The similar network optimization problem of antenna array is also presented in [76].

Artificial neural network can be used also for the specific tasks in the antenna arrays. The reconstruction of the complex excitation of antenna elements is presented in [77]. Complex excitation is one of most important parameter that describes radiation properties of a phased antenna. The properties of degradation of antenna array can be expressed by the deviation of practical element excitation’s from the ideal ones. The defects were simulated in antenna array in [77] with the hypotheses that 20% of total elements may contain defects. In comparison, the authors of [78] suggested an MTM antenna with a downsized profile MTM with circularly polarized increased gain performances, measuring the electromagnetic characteristics of the INP substrates and the printing method based on the KNN algorithm. The suggested antenna was determined to have adequate accuracy at the tested frequencies when bending effects were used to compare the results to the flat case.

4.3. Phase Shifters

According to Huang et al. [79], in order to optimize the throughput benefit of intelligent reflecting surface (IRS), the base station (BS) must acquire both the traditional direct channel
between the BS and the user equipment (UE) and the IRS reflected channel. The channel state information (CSI) of direct and IRS reflected connections is widely assumed to be completely understood at the BS. In practice, obtaining accurate CSI is challenging since the size of the IRS reflected channel grows with the number of reflecting parts.

IRS data rate deterioration is caused by both the large pilot overhead and the channel estimate mistakes. The authors of [80] presented a deep learning (DL)-based technique for determining the IRS phase shift while optimizing the data rate of IRS-aided communication systems. The advancement was made possible by using a deep neural network to establish a nonlinear link between the noisy estimated channel and the IRS phase changes.

The next several articles solve the problem of the complex control of dual active bridge (DAB). Challenges arise due to many non-linearly changing parameters which affect the operation of the DAB. Authors from [81] proposed the triple-phase shift control (TPSC) method, which allowed to significantly decrease the current amount that flows through the high frequency (HF) transformer. This was achieved by replacing the generally used lookup table with a neural network model. The main idea was that the look-up table stores the optimized modulation parameters, which are discrete. Neural networks allowed to estimate the nonlinear predictor function for the TPS associate. As a result, the efficiency was increased to 99%. The investigation was made in Simulink Matlab©. The feed-forward neural network was used for the investigation. The investigated optimal size of the hidden layers was equal to 10. The same problem of optimal power efficiency was also investigated in [82]. This time, the authors used the deep neural network, which allowed to figure out the connections between modulation parameters and power loss.

Another example is presented in [83]. This time, the convolution neural network is used in order to make the nonlinear phase shifter independent from the changing baud rate. Thus, it can be seen that the field of application of ANNs is wide, ranging from the synthesis task to model the phase shifter itself until the error detection and optimization.

4.4. Other Applications

The large application areas of the modeling of microwave devices using ANNs are discussed in detail. Of course, these are clearly not all application areas of microwave devices. A few more options are discussed. The transmission lines are another group of microwave devices. It could be also the traveling wave tubes, delay lines and waveguides. In this section, we try to discuss the more exotic examples, where more complicated types of networks or more complicated structures or materials of the microwave devices are used.

The [84] presents the example of the waveguides construction optimization task, which should work in the X-band frequencies. The optimal geometrical parameters were obtained using the neural network-biased genetic algorithm.

Hu et al. [85] described the use of physics informed neural networks (PINNs) to address rectangular waveguide difficulties. In order to discover solutions for electric and magnetic fields, partial differential equations (PDEs) were substituted by PINN. PDEs may be naturally encoded into the loss function using PINNs, whereas partial derivatives with respect to input variables can be generated using automated differentiation (AD) incorporated into recent deep learning packages.

The more exotic variant of the usage of ANNs for the analysis and synthesis of waveguides is presented in [86]. Instead of using the conventional dielectric materials, the rectangular waveguide with metamaterials was used. The investigated waveguide was designed to work in the infrared wavelength λ telecommunication spectrum (1200 nm to 1700 nm), where long propagation length and high confinement of the fields are desirable. The MLP network was used in order to calculate the main device propagation properties, such as the propagation length L and the penetration depth dp. The two hidden layers were used for the analysis task and one hidden layer for the task of synthesis. The best configurations were obtained using 28 and 14 neurons for L estimation, 42 and 28 neurons for dp estimation and 15 neurons for the task of synthesis. The activation functions in all
cases were tangent hyperbolic for the hidden layers and the linear one in the output layer, respectively. The training process was carried out with the Levenberg–Maquardt algorithm.

The usage of ANNs for the modeling of delay lines was presented in [87]. Delay lines are used in many different microwave circuits for signal synchronization. Authors wanted to increase the simulation speed while keeping the same accuracy. Therefore, the ANNs were used. The ANN allowed to speed up the analysis 2000 times while keeping high 99.5% accuracy. The MLP network was used for the prediction.

The time delay neural network (TDNN) is presented in [88]. This type of network is the modified multilayer perceptron neural network. Therefore, the comparison of both networks is presented in this paper. Authors advise to use TDNN when the nonlinear and dynamic effects become significant and when it is necessary to have the possibility to build the more general models without building proper equivalent circuit models.

GaAs metal-semiconductor-field-effect transistor (MESFET) and GaAs high-electron mobility transistor (HEMT) samples were used to validate the TDNN example. These two examples show that the suggested TDNN is a viable and efficient method for simulating many types of nonlinear microwave devices. For example the training and test errors of the MLP were equal to 31.13% and 33.68%. Respectively, TDNN results were 6.41% and 6.58% with one delay buffers and 1.49% and 1.86% with seven delay buffers. The number of neurons was the same and equal to 40 in both MLP and TDNN networks.

5. Neural Networks Classification According to the Learning Type

All previously presented ANNs examples are usually trained using the supervised training method. Supervised training allows to obtain usually desirable accuracy but also has its disadvantages. As already was told, the success of neural network training depends directly on the size and quality of the training dataset. Sometimes it is not available to collect a huge dataset of training samples. In addition, the collection of training samples using the conventional full-wave methods takes a lot of time. There are already investigations that analyze the ways of extracting the desirable features from huge wireless data streams [89] and also how to automate the collection of the training datasets [90] that is especially important for the training and analysis of microwave devices. Mishra et al. [91] stated in their review study that neural networks are excellent for DOA estimation and beam shaping, due to their nonlinear nature, capacity to manage enormous parallel operations, capability of universal approximation, accuracy, and speed. For smart antenna design, neural network approaches outperform print numerical techniques. The study finds that in the construction of smart antennas, a hybrid neural network model with an appropriate combination of different types of neural networks may be beneficial. The arising issues of the collection of the training datasets could be also solved by using semi-supervised and unsupervised learning (Figure 3).

![Figure 3. Neural networks classification by the learning type.](image)

The authors from [92] proposed the semi-supervised training method, which consists of the initial training and the self-training. Using this proposed method, only the small
initial training dataset is required, which is usually obtained using the full-wave simulation. After the first initial training, the presented model produces the unlabeled training dataset and the model trains itself until the testing accuracy is satisfied. In comparison with the usual training technique, which collects the training dataset from full-wave EM simulations, the suggested model only uses a limited number of labeled samples.

Authors from [92] presented two application examples of the usage of the semi-supervised training network. In the first example, the analysis of the microstrip-to-microstrip vertical transition model is presented. The input vector consists of six geometrical parameters of lengths and widths of different sectors of the microstrip. The output vector consists of real and imaginary parts of the $S_{11}$ parameter. The HHFS© software package was used for the collection of the initial training dataset. The investigated range was from 0 to 15 GHz. First of all, the initial training was performed with 15 randomly selected labeled samples when the self-training process started, which overall had 17 iterations. Despite the training period, the major advantage of this self-training model was that it required far less optimization time than the full-wave EM simulations.

Musumeci et al. [93] provided another example of the usage of semi-supervised neural networks. This time, machine learning (ML) was applied for autonomous failure diagnosis in microwave networks based on real-world data. Various problems contribute to connection outages. The authors had data on six types of failure in microwave networks. The authors employed autoencoder-like ANNs to blend knowledge from a small amount of manually labeled data with a big amount of unlabeled data.

Yang et al. [94] provided an example of a deep semi-supervised learning approach from the perspectives of model creation and unsupervised loss functions. Zhang et al. [95] described a technique for spectrum sensing in a real radio environment based on a semi-supervised deep neural network (SSDNN). SSDNN was used to extract characteristics of signal from a limited number of labeled samples. Unlabeled samples were utilized in the self-training procedure. The network was retrained using the expanded dataset. On a dataset of 124,800 samples, several tests were conducted. Only 18% of the data was initially labeled. The reached classification with SSDNN was higher than 90%.

The sampling approach also has a significant influence on the learning accuracy of ANNs. In reference to Xiao et al. [96], for example, a semi-supervised radial basis function neural network with a new sampling method was presented to decrease the nonuniform error distribution and sluggish convergence caused by the uncertainty of sample selection in the training process. This new sample technique enabled us to maintain the same level of training and testing accuracy over the whole sampling zone.

The main reasons for using the unsupervised learning are the model building and the dimensionality reduction [97]. As it is written in [98], the unsupervised learning is the key component in the 6G systems. The future wireless technologies heavily depend on the artificial intelligence that could be trained, unsupervised, because the amounts of important information increase rapidly. In particular, it will be popular in the optimization procedures [99].

The example of unsupervised learning in the field of microwave devices and communication was provided in [100]. The hybrid beamforming architecture with antenna selection allows the system to be flexible and have hardware efficiency. It is regarded as a critical technology for fifth-generation wireless communication networks (5G). The selection network was the ASNet, and the hybrid beamforming network was the BFNet. ResNet was used in both networks to extract features from the channel matrix. The unlabeled samples were chosen, using the authors’ suggested deep probabilistic subsampling method for ASNet and a specifically constructed quantization function for BFNet. The authors also introduced a configurable loss function, which allows a phased unsupervised training technique to efficiently train the combined network.

K-means, hidden Markov model (HMM), auto encoders (AEs), self-organizing maps (SOMs), fuzzy C-means and other unsupervised algorithms are examples of successful application in this area. According to Schmarje et al. [101], unsupervised machine learning
may be used to improve the performance of deep learning (DL) algorithms, such as convolutional neural networks (CNNs) and long short-term memory (LSTM) algorithms.

6. Fuzzy Logic

Fuzzy logic techniques are also used in the design and analysis of microwave devices [102]. Starting from the design of microwave transistors [103], continuing in the field of microwave circuits [104] and ending with completed microwave devices, such as waveguides [105], antennas [106], power amplifiers [107], power converters [108] and others [109]. The fuzzy logic techniques help not only in the design and analysis of microwave devices, but also in the procedure of the manufacture of microwave devices as it is presented in [110]. The control of technological process of induction soldering is significantly complicated, and the presented mathematical model together with intelligent methods, one of which is fuzzy logic, allowed to improve the quality of the induction heating process, which is an important part of the soldering process.

In general, fuzzy logic is a technique which allows to represent and manipulate with uncertain information. Fuzzy logic can be assigned to the group of low-level artificial intelligence. It is used in the lower-level machine control, so-called fuzzy controllers. Fuzzy logic is an extension of conventional logic in which a concept’s degree of truth can range from 0.0 to 1.0 [111].

For example, fuzzy logic was used in [112] in order to achieve better operation parameters of two-way symmetrical Doherty power amplifiers. Authors claimed that where was no solution to achieve more than 60% fractional bandwidth using the continuous-mode technique for the design of Doherty power amplifiers. Authors achieved a 15% increment in fractional bandwidth and reached up to 66.7% by using the technique, which was based on fuzzy logic. The fuzzy logic allowed to speed up and simplify the continuous-mode based technique. Authors used the combination of the K-means unsupervised learning clustering algorithm and a modeled continuous-mode technique. This combination allowed not only to extend the fractional bandwidth, but also to improve other important parameters, such as efficiency, output power and gain. The proposed technique also allowed automating the calculation of optimum characteristic impedances and electrical lengths of the transmission lines of the two-way symmetrical Doherty power amplifier.

Another example of the usage of fuzzy logic was presented in [113]. The authors presented the concept to evaluate the performance of the antenna, which is based on the fuzzy logic. The monopole, dipole, inverted f and helix antennas were designed in Matlab© using the antenna design toolbox. The directivity, VSWR, reflection coefficients parameters of each antenna were analyzed in order to evaluate the performance of each antenna. The conventional mathematical techniques could not describe the complex decision-making rules for the evaluation of the performance of the antennas with regards to these varying parameters. Fuzzy logic allowed to solve the task. The fuzzy decision system was developed using the fuzzy logic toolbox in Matlab©.

According to these two examples above, it is possible to summarize that fuzzy logic allows to think not only about the two possible values, such as logical low and logical high, but also about the whole range of values between the logical low and logical high. All the steps are described in Figure 4.

First of all, it is necessary to convert the main crisp values of the parameters of the microwave devices into the fuzzy logic values using the fuzzifier. These results are described in terms of the degree of membership of the fuzzy sets. Secondly, the “If and Then” sets of rules will be used for the decision making. Third, the defuzzification procedure will convert data back to specific real values. The same stages were performed in the above papers.
7. Discussion and Future Perspectives

According to the discussed examples, it is clear that ANNs are widely used in the field of synthesis and analysis of microwave devices. The main application areas of ANNs are antennas, antennas arrays, phase shifters, filters, waveguides, traveling wave tubes and others. ANNs usually replace the traditional full-wave methods while performing the synthesis or analysis of microwave devices.

There are many different types and structures of ANNs starting from simple MLP network in order to solve the regression task till the more complex deep ANNs or convolutional ANNs, which could solve more complicated classification tasks. The learning algorithm and training dataset has also a very important role in order to have accurate prediction results with ANNs. The training dataset can even influence the selection of the type and structure of the neural network.

Due to their widespread use, the usage of ANNs is expected to grow significantly in the future. On the other hand, there are still many challenges to apply ANNs in practical engineering problems. One of the most important challenges is a very huge and still increasing amount of training samples, which cause difficulties in the practical training of ANNs. The field of microwave devices requires collecting big amount of samples for training in order to have high learning and test accuracy. The calculations with traditional full-wave methods take a lot of time. Additionally, the field of wireless communication in the microwave frequency range streams a large amount of data, which all should be analyzed and used for training. Therefore, the future perspective is to use the semi-supervised or unsupervised training of ANNs, which could itself extract the main features from obtained signals and learn during the operation.

When it is difficult and time consuming to collect the training dataset, a good future perspective could be the augmentation of the training samples. It is already widely used in other fields such as image processing and others. After collecting the small dataset of samples with traditional full-wave methods, the training data could be augmented by using intelligent and automated augmentation methods usually also based on ANNs.

Finally, there is the issue of data dependency, which cannot be completely resolved because neural networks learn from examples and researchers label the training data so that the network knows “what is what”. As it was already mentioned, unsupervised and
semi-supervised networks exist. However, these networks are more complex. In unlabeled data, an unsupervised network attempts to minimize ladder patterns, differences, and similarities. Semi-supervised networks act as a bridge between supervised and unsupervised training by training an initial model with labeled data and then applying iteratively to label unlabeled data with a set of tolerances. It appears that semi-supervised and unsupervised networks might enhance the process of collecting and training using datasets, although this is not a foolproof strategy. Because of its structure, unsupervised networks are employed when a vast amount of data is acquired. For example, there are fewer networks to search for abnormalities in the field of microwave devices. Semi-supervised neural networks are used to solve classification and regression issues. However, according to the researchers, this can assist to enhance the learning rate, but can also make it worse. Again, a tailored strategy is necessary. Adapting for the field of microwave devices is not always possible. Issues with summarizing training data cannot be resolved automatically. Every application requires its own network topology and training dataset. The difficulties arise from the time required to acquire data using full-wave approaches. When the researcher already has the training dataset, labeling it is required and it is a time consuming process. There is no way to bypass these phases.

8. Conclusions

The usage of artificial neural networks in the field of the synthesis and analysis of microwave devices is reviewed and discussed in this article. The study was conducted by discussing and summarizing individual examples from separated articles. Overall, the review consisted of 113 scientific articles.

The main focus while discussing every article was on the application areas of artificial neural networks in the field of microwave devices, the type and structure of the used artificial neural networks, the type and size of the dataset, the interpolation and automated augmentation of the training dataset and the training algorithm. The summary of the reviewed methods is presented. The future trends are also discussed. It is likely that more research will be conducted in order to automate the collection of training datasets by using different techniques of interpolation and data augmentation.

The situation in the field of microwave devices and artificial neural networks is changing significantly; therefore, we will aim to track research and publish an update to this review paper as the situation changes.

Author Contributions: Conceptualization, A.K. and D.P.; methodology, A.K.; formal analysis, D.P. and R.M.; investigation, A.K., R.M. and D.P.; resources, A.K. and D.P.; data curation, A.K. and R.M.; writing—original draft preparation, A.K. and D.P.; writing—review and editing, R.M. and R.D.; visualization, D.P. and R.M.; supervision, R.M. and R.D.; funding acquisition, R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: The authors would like to thank the editors and reviewers for many constructive suggestions and comments that helped improve the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


108. Shah, F.M.; Maqsood, S.; Damaševičius, R.; Blažauskas, T. Disturbance rejection and control design of mvdc converter with evaluation of power loss and efficiency comparison of sic and si based power devices. Electronics 2020, 9, 1878. [CrossRef]


110. Tynchenko, V.S.; Milov, A.V.; Bukhtoyarov, V.V.; Kukartsev, V.V.; Tynchenko, V.V.; Bashmur, K.A. Control of the technological process of the waveguide paths induction soldering in the framework of Industry 4.0. J. Phys. Conf. Ser. 2019, 1399, 044087. [CrossRef]

