Brunel University London

College of Engineering, Design and Physical Sciences
Department of Electronic and Electrical Engineering

Brunel University of London

School of Communications and Information Engineering

Chongqing University of Posts and Telecommunications

BEng Graduation Project

Antenna modelling based on convolutional neural network

(Anonymous Copy)

April 2025

Abstract

In the field of antenna design optimisation, traditional solutions mainly rely on full-wave electromagnetic simulation software such as HFSS (High Frequency Structure Simulator) and CST (Computer Simulation Technology), which are accurate but consume a lot of computational resources and have a long optimisation cycle. Especially for microstrip antennas with complex structure and large size, the simulation time increases exponentially, which seriously restricts the practical application. To solve this problem, this thesis proposes a microstrip antenna resonant frequency prediction method based on the combination of image convolutional neural network and efficient channel attention (ICNN-ECA) mechanism. In this thesis, an efficient channel attention (ECA) mechanism is introduced on the basis of the traditional image convolutional neural network (ICNN), which solves the limitation that the ICNN model is unable to differentiate the importance of different feature channels when dealing with the electromagnetic characteristics of microstrip antennas.The ECA module achieves the local dependence modelling among feature channels through lightweight one-dimensional convolution, and captures the inter-channel interactions without the need of dimensionality reduction, so that the network to adaptively strengthen the feature channels that contribute most to the resonant frequency prediction. The experimental results show that the prediction accuracy of the ICNN-ECA model is improved by 48.8% compared to the traditional ICNN model.

The main contribution of this thesis is that the ECA mechanism is applied to the field of microstrip antenna modelling for the first time, which significantly improves the model's ability to perceive the key electromagnetic features of microstrip antennas without adding too many parameters, and provides a highly efficient and reliable alternative for the rapid design and optimization of microstrip antennas.

Keywords: Microstrip antenna, Resonant frequency prediction, Convolutional neural network, Efficient channel attention mechanism

Acknowledgements

As I write these final words, I find myself reflecting on a journey that has been both challenging and profoundly rewarding. The completion of this thesis marks not just the culmination of research, but the end of four transformative years that have shaped my academic identity and personal growth.

I owe my deepest gratitude to my first advisor, whose exceptional mentorship has been the cornerstone of my academic development. Your unwavering support, intellectual rigor, and thoughtful guidance have influenced me far beyond this thesis. You challenged my thinking, refined my approach, and inspired a deeper appreciation for our field. These lessons will resonate throughout my career.

To my second advisor, thank you for your invaluable expertise and perspective. Your detailed feedback and innovative suggestions consistently elevated my work and pushed me to explore new dimensions of my research. The clarity and insight you brought to our discussions transformed what seemed like insurmountable challenges into exciting opportunities.

My undergraduate experience in telecommunications engineering has been extraordinary thanks to the outstanding faculty of the International College. What began as a tentative first step into an unfamiliar discipline evolved into a genuine passion under your collective guidance. Your dedication to teaching and commitment to excellence have given me both the technical foundation and intellectual curiosity essential for future success.

No words can adequately express my gratitude to my parents. Your unconditional support has been my constant through twenty years of education. You celebrated my successes, comforted me through difficulties, and never wavered in your belief that I could achieve my goals. This accomplishment is as much yours as it is mine. To my extended family, thank you for creating the secure foundation that allowed me to pursue my academic interests wholeheartedly. Your encouragement from afar has been felt and appreciated more than you know.

My friends and classmates have been an incredible source of inspiration,

collaboration, and sometimes much-needed distraction. Our late-night discussions, mutual encouragement, and shared challenges created a community that made this journey not just bearable but genuinely enjoyable. I especially want to thank my roommates, who have become family over these four years. Your friendship has been one of the most unexpected and treasured gifts of my university experience.

Finally, I extend my sincere appreciation to the committee members who generously contributed their time and expertise to evaluate this work. Your thoughtful assessment and constructive feedback have been invaluable in bringing this project to completion.

This thesis represents not just my work, but the collective investment of everyone mentioned here. Thank you all for being part of this chapter of my life.

Contents

List of Figures

List of Tables

List of Acronyms

Notation	Description	Page List
HFSS	High Frequency Structure Simulator	ii
CST	Computer Simulation Technology	ii
ECA	Efficient Channel Attention	ii
SVM	Support Vector Machine	2
GP	Gaussian Process	2
ANN	Artificial Neural Network	2
HKF	Hybrid Kernel Function	2
CNN	Convolutional Neural Network	2
ICNN	Image Convolutional Neural Network	2
LSTM	Long Short-Term Memory	2
ICNN-ECA	Image Convolutional Neural Network with Efficient Channel Attention	3
CNN-SENET	CNN-Squeeze-and-Excitation Network	5
RMSA	Rectangular Microstrip Antenna	10
MSE	Mean Square Error	26
MAE	Mean Absolute Error	26
RMSE	Root Mean Square Error	26
MAPE	Mean Absolute Percentage Error	26
SMAPE	Symmetric Mean Absolute Percentage Error	26

Chapter 1

Introduction

1.1 Background

In recent years, electromagnetic waves have been widely used in the field of communications. Antennas, as energy conversion devices between electromagnetic waves and guided waves, are the core components for achieving spatial radiation and reception in wireless communication systems. Microstrip antennas are widely used because they are easy to integrate with microstrip lines, have low manufacturing costs and are easy to mass produce [2]. The antenna is composed of a dielectric substrate, a metal patch and a ground plane. Its planar and layered geometric structure facilitates rapid modeling in simulation software. Compared with three-dimensional structures such as waveguides and horn antennas, its boundary conditions are clear, and the electromagnetic field distribution can be approximately analyzed through quasi-static analysis, providing a basis for theoretical verification of simulation results.

With the continuous development of computer technology, many electromagnetic problems can be solved by full-wave simulation software such as CST and HFSS. This type of software has high accuracy and plays a significant role in antenna design and optimization [3]. However, this traditional method has significant limitations. This type of software often requires tens of thousands of electromagnetic simulation calculations to obtain accurate results. Especially for microstrip antennas with complex structures and large sizes, the simulation time increases exponentially, which leads to huge consumption of computing resources and seriously restricts practical applications. Consequently, researchers globally are increasingly turning to machine learning methodologies to develop surrogate models that significantly reduce computational costs and optimization cycles while maintain-

ing reasonable accuracy, thus offering more efficient solutions for microstrip antenna design.

At present, the proxy models that have been applied to microstrip antenna modeling and the more common ones include support vector machine (SVM), Gaussian process (GP) and artificial neural network (ANN). Sun [4] proposed a method based on SVM and hybrid kernel function (HKF) to accurately model the resonant frequency of compact microstrip antennas (MSA). SVM is a small sample learning algorithm with good generalization ability. Through the characteristics of its decision function, it effectively avoids the ”dimensionality curse” problem. However, the disadvantage of this method is that the prediction results lack probabilistic meaning, and the determination of kernel parameters is relatively difficult [5]. Gao [6] proposed a collaborative training algorithm based on GP and SVM, which uses a small number of labeled samples and a large number of unlabeled samples for training, improving the prediction ability of the model.

However, all of the above machine learning-based algorithms have obvious shortcomings. GP and SVM rely on kernel functions and manual feature engineering, and ANN needs to explicitly encode spatial relationships and may ignore local structural information. In contrast, Convolutional neural network (CNN) is good at processing data with spatial structure and automatically capturing local features [7]. Cao [8] proposed a CNN-BiLSTM model based on attention mechanism, which predicts S_{11} parameters by using a 1D CNN. In addition to the above advantages, the unique structural design of CNN - including local receptive field mechanism, weight sharing strategy and sparse connection characteristics - also gives it robustness to translation and scale changes. These characteristics make CNN particularly outstanding in the field of image processing, and can efficiently extract hierarchical features from visual data, thus better adapting to various image analysis and recognition tasks [9]. Based on these advantages of CNN, Fu [10] proposed an image model-based CNN (ICNN) for modeling microstrip antennas. This thesis converts these 1D inputs into 2D image models and uses CNN’s image processing capabilities to improve the accuracy and generalization of modeling. Based on this research, Zhu[11] proposed an ICNN-LSTM model, which takes advantage of LSTM in processing long sequence data, thereby producing better accuracy in the prediction of S parameters.

However, when analyzing the image model of microstrip antennas, the contribution of electromagnetic characteristics information contained in different regions and channels to the resonant frequency varies significantly, and the standard CNN architecture cannot effectively identify and highlight

these key features. To address this problem, this thesis proposes an Image-based CNN-ECA (ICNN-ECA) antenna modeling method. Based on ICNN, this method introduces an efficient channel attention (ECA) mechanism to specifically enhance the feature channels that have the greatest influence on the prediction of antenna resonant frequency. In this architecture, the antenna parameters are first converted into a two-dimensional image matrix as input. After the convolutional layer extracts features, the ECA module calculates the local dependencies between feature channels and assigns weight coefficients to different channels, so that the network can more accurately capture the electromagnetic characteristics representation directly related to the resonant frequency in the image model. This targeted feature enhancement mechanism enables the model to more accurately understand the influence of antenna parameter changes on the resonant frequency, thereby improving the prediction accuracy.

1.2 Aim and Objectives

The main goal of this thesis is to develop a model based on ICNN to accurately predict the resonant frequency of microstrip antennas. Specific goals include:

1. Based on the existing 2D image representation method of microstrip antenna parameters, the prediction of the resonant frequency of microstrip antennas is realized.
2. Analyze and solve the limitation of traditional CNN in the inability to distinguish the importance of different feature channels in microstrip antenna modeling, and develop a CNN architecture that integrates ECA mechanism to enhance the model's perception of key electromagnetic characteristics of microstrip antennas.
3. Construct a complete ICNN-ECA network structure, optimize the network parameter configuration, improve the accuracy of resonant frequency prediction, and verify the performance advantages of the proposed method by comparing with existing models.

1.3 Project Approach and Contributions

This thesis proposes a microstrip antenna resonant frequency prediction method based on ICNN-ECA. This method uses the existing two-dimensional

image representation scheme of microstrip antenna parameters, innovatively introduces the ECA module based on the standard CNN architecture, realizes the local dependency modeling between feature channels through lightweight 1D convolution, and adaptively weights different feature channels according to their importance to the prediction task. The main innovation of this thesis is that the ECA mechanism is applied to the field of microstrip antenna modeling for the first time, which solves the problem that the traditional CNN architecture cannot distinguish the importance of different features when processing the electromagnetic characteristics of antennas. Through the cross-channel interaction design without dimensionality reduction, the model's perception of key electromagnetic features is significantly improved without increasing too many parameters. In addition, this method establishes a network architecture with alternating nested convolutional layers and ECA modules, realizing seamless integration from feature extraction to feature enhancement. The proposed ICNN-ECA model can accurately capture the impact of microstrip antenna parameter changes on the resonant frequency, providing an efficient and reliable alternative to traditional full-wave electromagnetic simulation for the rapid design and optimization of microstrip antennas.

1.4 Project Plan

1.4.1 Gantt Chart

Figure 1.1 illustrates the project schedule and timeline. The chart shows the planned sequence of activities, including literature review, data collection, model development, experimental evaluation, and documentation. The timeline spans six months, with each task allocated an appropriate duration based on its complexity and importance to the overall project.

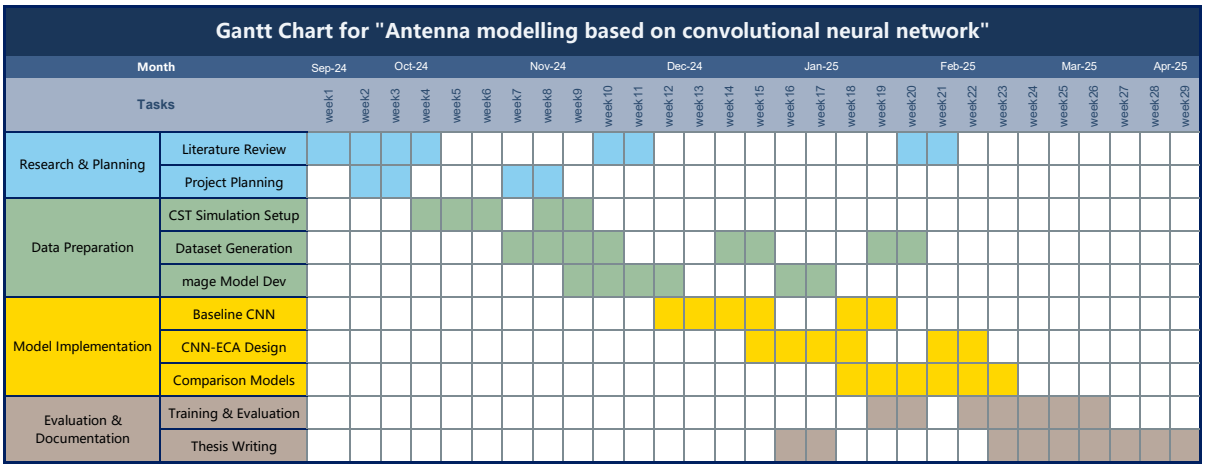

Figure 1.1: Gantt chart of project plan

1.4.2 Deliverables

The main deliverables of this project include:

1. A complete code implementation of the Pytorch-based ICNN-ECA model and its training weights.

2. A microstrip antenna parameter dataset with 500 sets of samples, and the corresponding CST full-wave simulation results.

3. A complete performance evaluation report of the model, including comparisons with the traditional CNN, CNN-SENET, and CNN-LSTM models Analysis.

4. Detailed configuration documents of the model training and testing environment to ensure the reproducibility of the experimental results.

5. An image model conversion tool to support the rapid conversion of microstrip antenna parameters to 2D image representations.

6. This thesis, “Antenna Modelling Based on Convolutional Neural Networks”, documents in detail the background of the research, theoretical basis, model construction and experimental results.

1.4.3 Other Considerations

This paper used the Claude 3.7 Sonnet Thinking generative large language model as an auxiliary tool in the writing process, mainly for English expression polishing and format standardization. It should be made clear

that all research contents of the paper, including theoretical analysis, model design, data processing, experimental design, result analysis and academic contribution, were completed independently by myself. The large language model is only used as a writing auxiliary tool and does not participate in any substantive research decision-making or content creation. All academic views and research results reflect the author's original ideas and professional level.

1.5 Dissertation Outline

This thesis is divided into five chapters, which systematically expounds the resonant frequency prediction method of microstrip antennas based on ICNN-ECA. Chapter 1 is an introduction, which introduces the research background, objectives and methodology; Chapter 2 comprehensively reviews the current research status, focusing on the development and application of convolutional neural networks and attention mechanisms; Chapter 3 describes the model construction process in detail, including data collection and preprocessing, image model construction and ICNN-ECA network architecture design; Chapter 4 shows the experimental setup and result analysis, and verifies the model performance through multiple evaluation indicators and comparative experiments; Chapter 5 summarizes the research results, points out the research limitations and proposes future work directions. The structural arrangement of this thesis ensures the systematicness and completeness of the research content, and provides readers with a clear research context.

Chapter 2

Current State of Research

2.1 Image model

Deep learning methods have made progress in the field of electromagnetic imaging and microwave modelling.Shao and Du [12] developed a deep learning network system to convert microwave signals acquired by a 24×24 antenna array at 4GHz into 128×128 high resolution images. They used a two-stage training method, first compressing the image into 256×1 vectors using a self-encoder, and then training the neural network to map the microwave signals to the compressed features, which reduces the training difficulty. Experiments demonstrate that this method outperforms the traditional distortion Born iterative method (DBIM) and phase confocal method (PCM) in reconstructing the position and shape of an object, and can be applied to imaging objects with dielectric constants in the range of 2-6.Jacobs [13] proposed a CNN regression-based method for predicting the resonance frequency of a dual-band pixelated microstrip antenna, which uses the two-dimensional layout of the antenna surface as an input directly, and through a Bayesian optimisation to determine the network hyperparameters, which obtained average relative errors of 0.13% and 0.22% on the test predictions, outperforming conventional feed-forward neural networks.

2.2 Convolutional neural network (CNN)

Convolutional neural network (CNN) is a type of deep feedforward neural network with convolution calculation, which is widely used in image feature extraction. Its development began with the concept of ”receptive field” proposed by Hubel and Wisesel in 1959[14], which laid the foundation for the hierarchical structure of visual processing. In the 1980s, Kunihiko Fukushima

proposed the ”neocognitron”[15], which partially realized the functions of convolution layer and pooling layer, and was regarded as the pioneering research of CNN. In 1989, LeCun constructed the first version of CNN in the field of computer vision[16]. The model contained two convolution layers and two fully connected layers. It used ”convolution” to describe the network structure for the first time and adopted the stochastic gradient descent training method that is still used today. In 1998, LeCun improved LeNet-5[17] and successfully applied it to handwritten digit recognition. After Hinton proposed the deep learning theory[18] in 2006, CNN developed rapidly. In 2012, AlexNet of Krizhevsky’s team won the ImageNet competition[19], which attracted widespread attention. Subsequently, network architectures such as VGGNet[20], GoogleNet[21], and ResNet[22] came out one after another, demonstrating excellent performance in the field of image processing.

2.3 CNN in antenna design

Wei et al [23] proposed a hybrid CNN-LSTM network for automatic parametric modelling of antenna structures, which achieves a fast conversion from antenna images to modelling codes and significantly improves the efficiency of data acquisition.Zhang et al [24] combined the particle swarm optimization algorithm (PSO) with CNN in a PSO-CNN method, which was successfully applied to the optimal design of fragmented antenna structures, and the prediction results were highly consistent with the HFSS simulation is highly consistent.Wu et al [25] developed a machine learning-assisted optimization (MLAO) method combining CNN and Gaussian process regression technique, which is applied to the geometric optimization design of antenna, and realizes the synergistic design between parameter optimization and geometric optimization.Karahan et al [26] applied CNN to the modeling and inverse design of multiport antenna, and predicted the S-parameters through neural network with genetic algorithm to achieve automatic synthesis of antenna structures.Singh and Panda [27] proposed a deep learning-enabled multi-objective antenna design method using CNN-assisted differential evolutionary algorithm, which significantly reduces the computational cost and improves the optimisation efficiency.

2.4 Attention mechanism

The selective perception mechanism displayed by human visual systems when processing massive information provides important implications for deep learning. As a deep learning technology that cleverly draws on the characteristics of human visual systems, the attention mechanism has been widely used in various deep learning tasks due to its excellent performance and convenience of modular design. The core idea of the attention mechanism is to obtain differentiated representations of the importance of feature maps through specific algorithms, allocate the computing resources of the neural network more to key areas, and use task results to reversely guide feature weight updates, thereby completing the corresponding tasks efficiently [28]. In convolutional neural networks, the attention mechanism is mainly divided into two research directions: single-channel attention and multi-channel attention. Single-channel attention such as SE-Net obtains the difference in importance between channels through extrusion excitation operations [29], while ECA-Net obtains more precise attention distribution through non-dimensionality-reduced cross-channel interaction [30]. Multi-channel attention such as SK-Net uses adaptive selection based on convolution kernels to achieve feature enhancement [31], while CBAM innovatively combines space and channel attention modules to comprehensively extract multi-dimensional information of the feature map [32]. Extensive experimental evidence demonstrates that attention mechanisms significantly enhance model performance. In radiation source recognition tasks, the optimal attention module improved recognition rates from 83.52% to 93.23%, representing a nearly 10 percentage point improvement. Future research directions include exploring hierarchical attention information in feature maps, developing more precise attention distribution representations, optimizing structures to reduce computational complexity, and investigating attention mechanisms' potential in cross-modal and self-supervised learning [33].

Chapter 3

Model construction

3.1 Data collection and preprocessing

In order to more conveniently model the microstrip antenna, a rectangular microstrip antenna (RMSA) is selected for detailed description and edge feeding is adopted. Figure 3.1 shows the structure of the rectangular microstrip antenna, where L is the length of the patch, W is the width of the patch, h is the height of the substrate, ε_r is the substrate dielectric constant. The resonant frequency of the RMSA is primarily determined by these physical parameters [1]. Therefore, the input parameters can be expressed as $\boldsymbol{X} = [L, W, h, \varepsilon_r]$, and the output parameters can be expressed as $\boldsymbol{y} = f_r$, where f_r is the resonant frequency of the antenna.

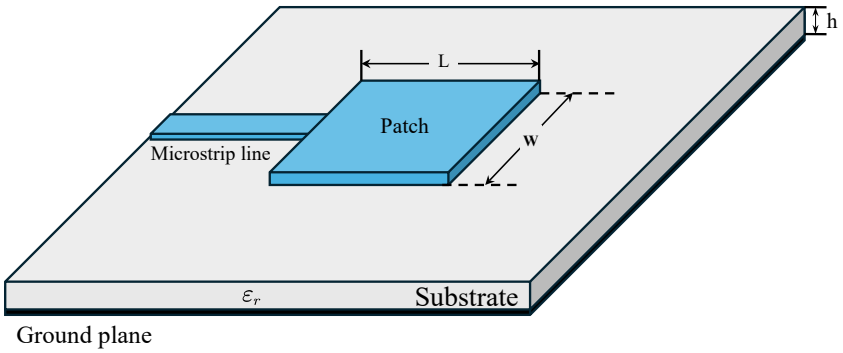

Figure 3.1: Rectangular microstrip antenna (RMSA) structure

3.1.1 Theoretical analysis of resonant frequency

Resonance frequency is the main electromagnetic characteristic parameter of microstrip antennas, defined as the frequency point at which the antenna resonates under a specific structure and medium, when the antenna

is able to efficiently convert RF signals into electromagnetic wave radiation [34]. It determines the centre operating frequency of the antenna and affects the match between the antenna and the communication system, the bandwidth characteristics and the data transmission capability, gain and directionality. In multi-band or broadband antenna design, the control of multiple resonant frequencies is the technical basis for realising multifunctional antennas. Considering the cavity model, the resonant frequency is given by:

$$(f_r)_{mnp} = \frac{1}{2\pi\sqrt{\mu\varepsilon}}\sqrt{\left(\frac{m\pi}{L}\right)^2 + \left(\frac{n\pi}{W}\right)^2 + \left(\frac{p\pi}{h}\right)^2} \tag{3.1}$$

For analytical purposes, the dominant mode TM_{010} (transverse magnetic wave mode) is typically employed as the fundamental basis for resonant frequency characterization. Considering the firing effects of the microstrip antenna, the resonant frequency can be expressed as:

$$f_r = \frac{c}{2L_{eff}\sqrt{\varepsilon_{reff}}} \tag{3.2}$$

where c is the speed of light in free space, L_{eff} is the effective length of the patch, and ε_{reff} is the effective dielectric constant. The effective length can be expressed as:

$$L_{eff} = L + 2\Delta L \tag{3.3}$$

where ΔL is the length extension due to the fringing effect.

$$\Delta L = 0.412h\frac{(\varepsilon_{reff} + 0.3)(W/h + 0.264)}{(\varepsilon_{reff} - 0.258)(W/h + 0.8)} \tag{3.4}$$

The effective dielectric constant can be expressed as:

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-1/2} \tag{3.5}$$

The width of the patch can be expressed as:

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{\varepsilon_r + 1}} \tag{3.6}$$

Using these equations, when the resonant frequency, dielectric constant and substrate height are given, the length and width of the microstrip antenna can be calculated.

3.1.2 CST simulation data extraction

The resonant frequency is usually characterised and measured by the S_{11} parameter, i.e., the reflection coefficient.The S_{11} parameter describes the ratio of the electromagnetic energy incident on the antenna to the energy reflected back, measured in decibels (dB). At the resonant frequency, the absorption of incident energy by the antenna reaches a local maximum, and accordingly, the S_{11} curve exhibits a clear local minimum.The relationship between the S_{11} parameter and the resonant frequency is shown in Figure3.2, from which it can be analysed that the parameter reaches a local minimum at 9.84 GHz, and thus the resonant frequency of the antenna is $f_r = 9.84GHz$.

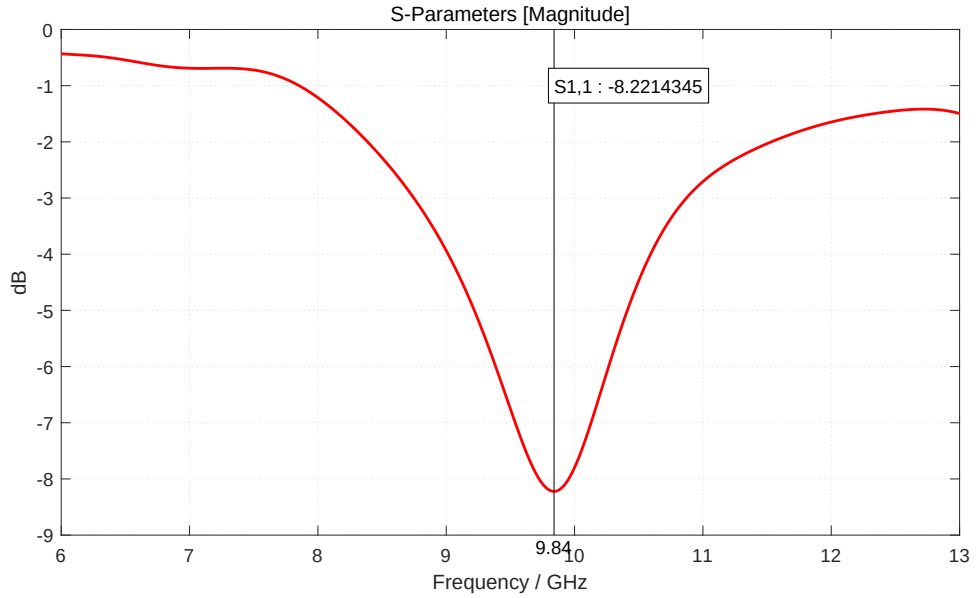

Figure 3.2: S_{11} parameter curve of microstrip antenna

In the field of numerical analysis of electromagnetic field, the main calculation methods can be summarised into three paradigms, namely, finite element method, method of moments and time-domain finite difference method, and CST software, as an important tool for electromagnetic simulation, is widely used in the numerical simulation of complex antenna structures and microwave devices. The software can provide high-precision parameter calculations required for antenna design, including return loss coefficients and radiation gain; meanwhile, in terms of power divider design, it can accurately analyse key performance indexes such as insertion loss and port isolation. Based on the above advantages, CST has been widely used in the design and optimisation of electromagnetic devices such as microwave antenna systems, frequency selective filters and power distribution

networks. In order to verify the theoretical model and experimental results of this study, this thesis adopts CST and MATLAB to carry out joint electromagnetic simulation and analysis, and the process is shown in Figure 3.3. Using the joint simulation method of MATLAB and CST, the CST software can be controlled by writing the dimension information of the antenna or filter in the script to achieve automated modelling and solving. The specific process of joint simulation in this thesis is shown as follows:

1. MATLAB loads the file "input_parameters.mat" containing the key parameters of the patch antenna.

2. Set the global simulation parameters such as feedline width, substrate dimensions, etc.

3. Establish the communication interface between MATLAB and CST through ActiveX server technology.

4. Create a new microwave studio project Configure the environment such as unit system, frequency range, boundary conditions, etc.

5. Automatically create the dielectric material according to the parameters and construct the geometric model including backplane, substrate, radiating patch and feeder.

6. Configure the wave port and excitation mode, call the CST solver to perform the electromagnetic field calculation, and finally save the simulation results to the specified location.

3.1.3 Data collection

The operating frequency band of microstrip antennas is usually distributed in the range of 1 GHz$\sim$ 30 GHz [35]. The low frequency band is mainly used in daily data communication systems, while the high frequency band is widely used in satellite communications. In order to ensure that the resonant frequency is effectively simulated within this reasonable range, this thesis uses equations 3.2 $\sim$ 3.6 to calculate the length and width parameters of the patch under the conditions of known resonant frequency, substrate dielectric constant and substrate height, and then transfers the data to CST for accurate simulation. Table 3.1 lists the dielectric constants and substrate height data of several typical materials, which provides an important reference for the design process.

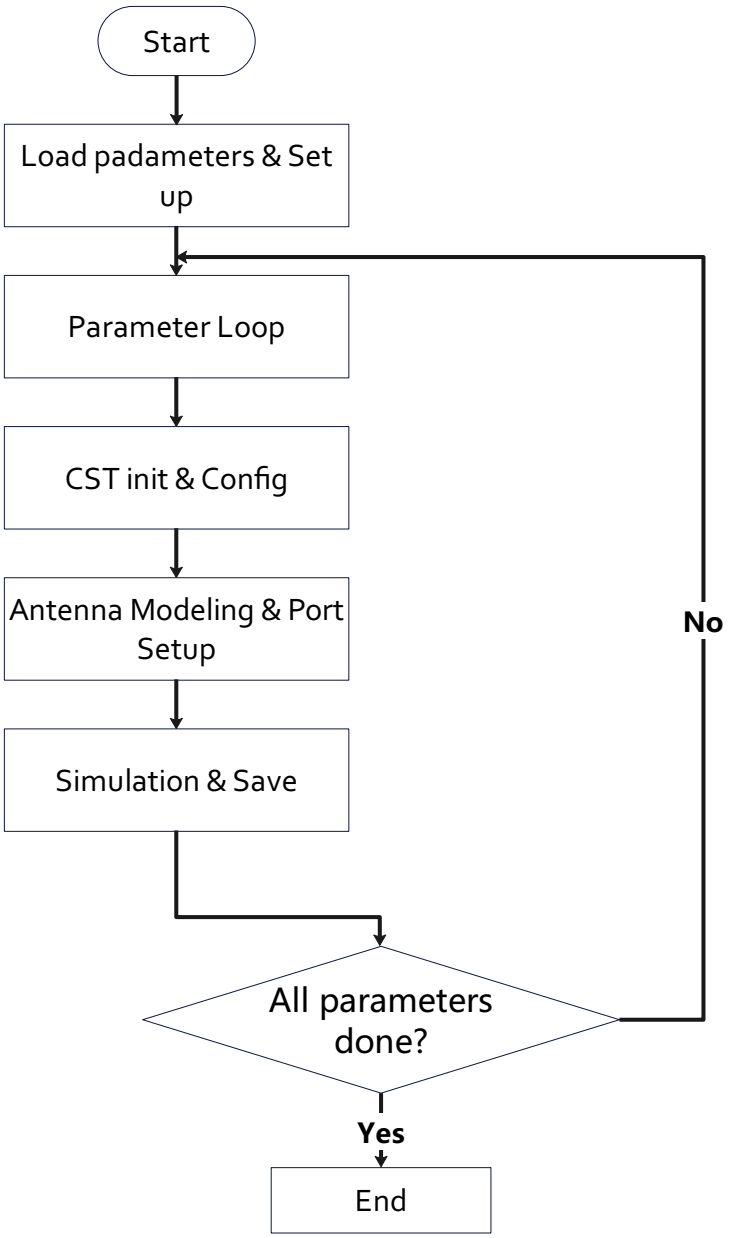

Figure 3.3: Flowchart of data extraction from CST simulation

The specific data collection process for the training and testing sets is described as follows:

1. Using the dominant mode TM_{010} as the operating mode of the microstrip antenna, the Monte Carlo method was applied to generate random resonant frequencies in the range of 1∼30 GHz, with random dielectric constants selected in the range of 2.26∼15, and random substrate heights in the range of 0.5∼2.0 mm.

2. Based on the theoretical model discussed previously, specifically equations 3.2 $\sim$ 3.6, the corresponding patch length H and width W were calculated, forming a complete input data vector $\boldsymbol{X} = [L, W, h, \varepsilon_r]$.

Table 3.1: Common microstrip antenna substrate material parameters [1]

Substrate	**Thickness (mm)**	ε_{r}
Duroid® 5880	0.127	2.20
RO 3003	1.575	3.00
	0.168	
RO 4350	0.508	3.48
	1.524	
DiClad 870	0.091	2.33
Polyguide	0.102	2.32
NH 9320	3.175	3.20
RF-60A	0.102	6.15

3. The above two steps were repeated, ultimately constructing a training dataset containing 500 samples and an independent testing dataset with 20 samples.

4. To obtain high-quality label data, each set of input data was imported into CST Microwave Studio software for full-wave electromagnetic field analysis. Precise resonant frequency values were extracted through the S_{11} parameter curve, serving as the reference standard for neural network training and evaluation.

3.2 Image model construction

This section elaborates on how antenna structural parameters are converted into two-dimensional image representations suitable for CNN processing. Traditional one-dimensional CNN has limitations in dealing with antenna optimization problems and fails to fully utilize the advantages of convolutional neural networks in image recognition. To overcome this problem, this thesis adopts the construction method of two-dimensional image model in [10].

Considering the key structural parameter set of RMSA, $\boldsymbol{X} = [L, W, h, \varepsilon_r]$, which includes length, width, substrate height and dielectric constant. These continuous physical quantities need to be converted into discrete digital representations through appropriate encoding methods. A binary encoding strategy is adopted to map each parameter to a binary sequence consisting of 0 and 1. For accuracy considerations, a 10-bit binary representation ($n = 10$) is selected to provide 1024 discrete values, which can meet the accuracy requirements of antenna design while ensuring computational ef-

ficiency. For the ith parameter $X_i \in \boldsymbol{X}$, its encoding process follows the following equation:

$$X_i = (X_i)_{min} + \frac{\sum_{j=0}^{n-1} b_j \cdot 2^j}{2^n - 1} \cdot ((X_i)_{max} - (X_i)_{min}) \tag{3.7}$$

where $(X_i)_{min}$ and $(X_i)_{max}$ represent the lower and upper limits of the parameter X_i in the design space, n represents the number of binary code bits, $n = 10$ is taken in this thesis, and $b_j \in \{0, 1\}$ represents the value of the j-th bit in the encoding sequence. Conversely, given the actual value of the parameter X_i, its corresponding binary representation can be obtained. Let $B = \sum_{j=0}^{n-1} b_j \cdot 2^j$, then we can deduce:

$$B = \left\lfloor \frac{X_i - (X_i)_{min}}{(X_i)_{max} - (X_i)_{min}} \cdot (2^n - 1) \right\rfloor \tag{3.8}$$

The rounding operation here ensures that B is an integer, which facilitates the subsequent generation of binary strings. Through the above normalization operation, a bidirectional mapping relationship is established between the antenna parameter space and the binary feature space. The corresponding model processing flow is shown in Figure 3.4.

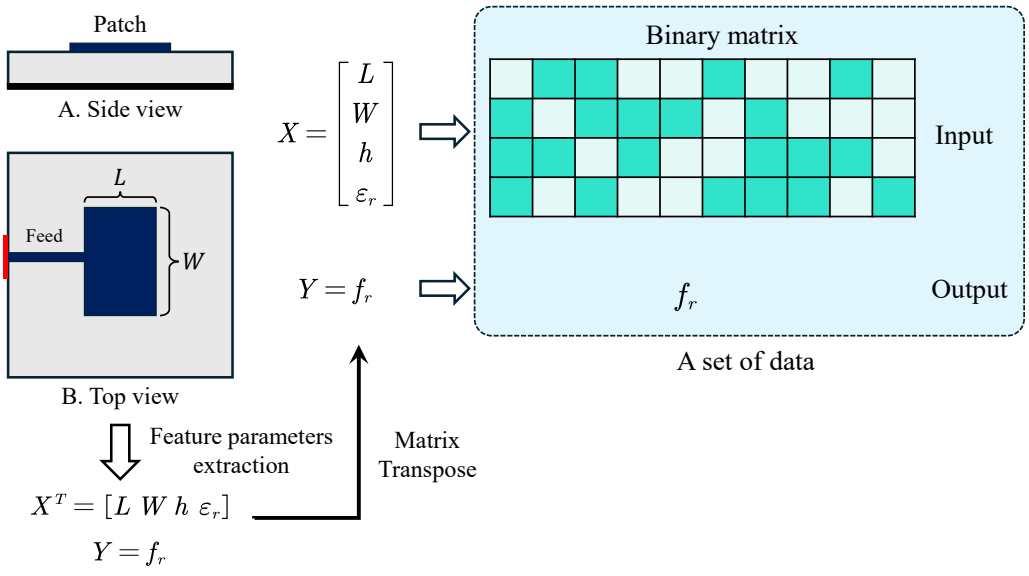

Figure 3.4: Image model of rectangular microstrip antenna

Through this encoding scheme, continuous antenna parameters are converted into 2D image representation, allowing CNN to fully exploit the advantages of spatial feature extraction. Compared with 1D representation, 2D image model more effectively utilizes the local receptive field and weight sharing characteristics of CNN, improving learning ability and generalization performance.

3.3 ICNN-ECA network architecture design

The electromagnetic characteristics of microstrip antennas usually present complex spatial distribution characteristics in image representation, where the information carried by different regions and channels contributes differently to the prediction of the resonant frequency. Although traditional CNN can extract spatial features, it lacks a perception mechanism for the importance differences between channels, making it difficult to accurately capture electromagnetic features that are highly correlated with the resonant frequency. To address this problem, the ICNN-ECA architecture proposed in this thesis achieves adaptive enhancement of key feature channels by introducing an ECA mechanism, thereby significantly improving the model performance.

3.3.1 Design Principle

The ICNN-ECA network architecture proposed in this thesis is based on two key findings: firstly, the electromagnetic field distribution of microstrip antennas exhibits complex spatial non-uniformity, with structural parameters having significantly different degrees of influence on resonant frequency; secondly, this non-uniform contribution characteristic is also reflected in the feature channels of neural networks, requiring specialized mechanisms for differentiated processing. Traditional neural networks have notable limitations when applied to microstrip antenna modeling. While fully connected neural networks can process numerical parameter inputs, they fail to capture spatial relationships between parameters; standard convolutional neural networks preserve spatial structural information but assign equal weights to all feature channels, failing to highlight the differentiated contributions of key feature channels. These limitations have driven research toward architectural design solutions capable of adaptively distinguishing channel importance.

Among various attention mechanisms, ECA stands out due to its parameter efficiency and computational simplicity. Compared to SE (Squeeze-and-Excitation) modules that require numerous parameters, ECA implements local channel interaction modeling through lightweight one-dimensional convolution, maintaining sensitivity to differences in importance between channels while avoiding training difficulties and overfitting risks caused by overparameterization. This characteristic makes it an ideal choice for processing small-sample microstrip antenna datasets.

3.3.2 Efficient channel attention module

The working principle of the ECA module is to enhance the expressive capability of important feature channels through adaptive weighting. Its implementation process includes three main steps:

1. Perform Global Average Pooling (GAP) on the input feature map $X \in \mathbb{R}^{(W \times H \times C)}$, compressing each feature channel into a single descriptive value:

$$y = g(X) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{i,j} \tag{3.9}$$

   where H and W are the height and width of the feature map, respectively. This step effectively aggregates information from the spatial dimensions.

2. Unlike the dimensionality reduction strategy of the SE module, the ECA module directly captures local inter-channel dependencies through a one-dimensional convolution operation:

$$\omega = \sigma(C1D_k(y)) \tag{3.10}$$

   where $C1D_k$ represents a one-dimensional convolution with kernel size k, and σ is the sigmoid activation function. This design avoids information loss caused by dimensionality reduction operations, while only considering the interaction between each channel and its k neighboring channels, greatly reducing computational complexity. The convolution kernel size k is adaptively determined based on the number of channels C:

$$k = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{odd} \tag{3.11}$$

   where $|t|_{odd}$ denotes the nearest odd number to t, and γ and b are adjustable hyperparameters, typically set as $\gamma = 2$, $b = 1$.

3. Multiply the generated channel attention weights with the original feature map to complete channel recalibration:

$$\hat{X} = X \odot \omega \tag{3.12}$$

   where $\odot$ represents channel-wise multiplication.

Figure 3.5 illustrates the structure of the ECA module. This soft attention mechanism enables the model to automatically identify and enhance the

most influential feature channels for the current task, while suppressing irrelevant or redundant information. Compared to other attention mechanisms, the distinctive feature of ECA's design lies in its extremely high parameter efficiency. For a feature map with C channels, the ECA module introduces only k parameters from the one-dimensional convolution operation. This is significantly lower than the SE module which requires $\frac{2C^2}{r}$ parameters, where r represents the reduction ratio used in the SE design.

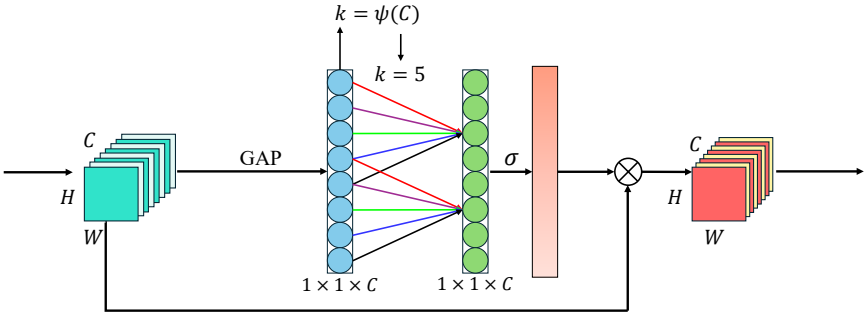

Figure 3.5: Efficient channel attention module

3.3.3 ICNN-ECA network structure

The ICNN-ECA network combines the feature extraction capability of CNN with the feature selection advantage of ECA to build a complete image regression model. The network takes a grayscale image as input and first gradually extracts image features through three convolutional layers. The first convolution layer uses a rectangular (1×3) convolution kernel to retain height information while reducing width, converting a single-channel image into a dual-channel feature map; the second convolution layer uses the same strategy (1×3) to further extract features and expand to four channels; the third layer uses a square (2×2) convolution kernel to reduce the spatial size of the feature map while expanding the number of channels to eight channels, completing the feature extraction stage. After each convolution layer, the network introduces the ECA module to achieve more precise regulation. The ECA module captures channel global information through global average pooling, then uses lightweight 1D convolution to model the relationship between channels, and finally generates weights through the sigmoid function and re-weights the feature channels, so that the network can adaptively focus on important feature channels. After feature extraction is completed, the network flattens the multi-dimensional feature map into a vector, and gradually reduces the dimension through a three-layer fully connected network to achieve the mapping from high-dimensional features

to a single prediction value. The entire structure of the ICNN-ECA network is shown in Figure 3.6.

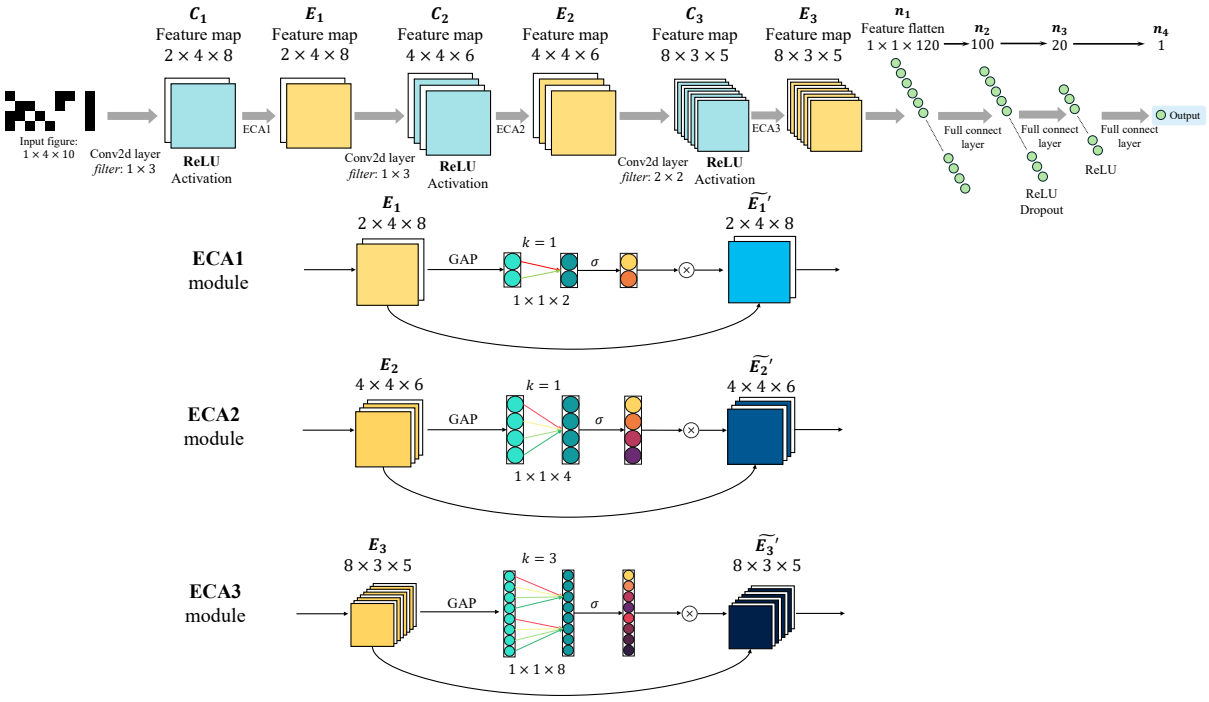

Figure 3.6: ICNN-ECA network structure

Chapter 4

Experiment and Results

4.1 Experimental setting

4.1.1 Experimental environment

This thesis was conducted on a high-performance computing platform equipped with an NVIDIA RTX 4090D graphics processor (24GB video memory), an Intel Xeon Platinum 8481C processor (16 cores), and 80GB of system memory. The software environment uses the PyTorch 2.5.1 deep learning framework, developed based on Python 3.12 (Ubuntu 22.04), with the CUDA 12.4 acceleration library. Code development and experimental implementation were completed using the Visual Studio Code integrated development environment.

4.1.2 Dataset construction

The experimental data comes from the CST Microwave Studio full-wave electromagnetic field simulation software, which generates the corresponding resonant frequency according to the microstrip antenna parameters. The resonant frequency range is controlled in 1∼30GHz to fully cover the coverage range of electromagnetic waves in communication, the substrate height range is in 0.5∼2.0mm, and the substrate vacuum dielectric constant is controlled between 2.2∼6.15.

The data set contains a total of 500 groups of samples, which are randomly divided into training sets of 400 groups and validation sets of 100 groups in an 8:2 ratio. In addition, 20 groups of samples are independently generated as test sets for the final model performance evaluation. Each group of samples contains four key parameters of the rectangular microstrip

antenna: length L, width W, substrate height h, dielectric constant ε_r and their corresponding resonant frequency f_r. All parameters are converted into 4×10 two-dimensional image representations through the 10-bit binary encoding method described above as the input of the ICNN-ECA model. Figure 4.1 shows the ICNN-ECA model flowchart using an example of $f_r = 9.8GHz$ and its corresponding four input parameters.

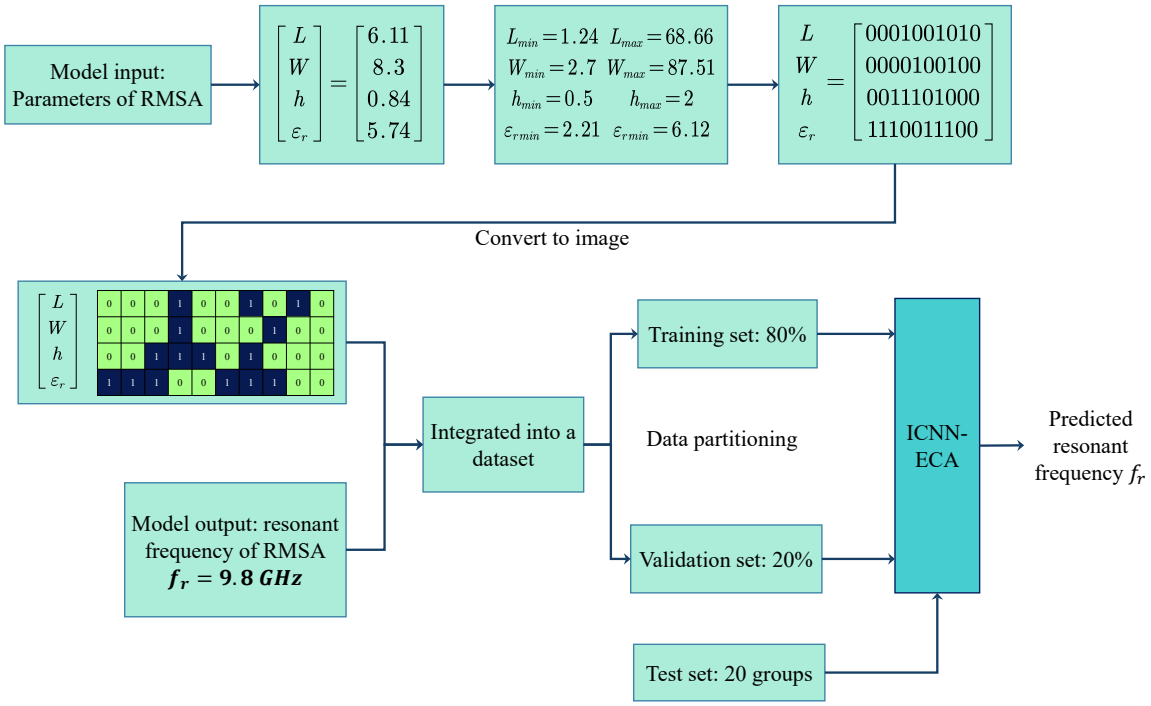

Figure 4.1: Flowchart of the proposed ICNN-ECA model

4.2 Network parameter settings

4.2.1 Hyperparameter configuration

This section elaborates on the parameter configuration scheme of the ICNN-ECA model, including hyperparameters, network structure parameters, and optimization methods, to provide theoretical basis and technical support for model implementation. Hyperparameters determine the basic characteristics of the model training process and have a significant impact on the final performance. Table 4.1 lists the hyperparameter configurations used in this thesis.

1. Batch size: Batch size refers to the number of samples used in each iteration. Reference [36] uses the PAC-Bayes theory to prove that there

Table 4.1: ICNN-ECA Model Hyperparameter Configuration

Hyperparameter Name	Value
batch_size	3
learning_rate	0.001
num_epochs	200
train_val_split	0.8
dropout_rate	0.1
normalization_mean	[0.5]
normalization_std	[0.5]
Optimizer	SGD

is a negative correlation between the ratio of batch size to learning rate and the generalization ability of deep neural networks, that is, when the ratio of batch size to learning rate is too large, the generalization ability is poor. Therefore, a batch size that is too large will lead to a decrease in generalization ability, while a batch size that is too small will increase the calculation time. In this thesis, this parameter is set to 3, which allows the model to maintain training stability and efficiently capture information in the data.

2. Learning rate: In the gradient descent algorithm, the learning rate controls the step size of each parameter update. A learning rate that is too large may cause the algorithm to fail to converge to the local optimal solution, while a learning rate that is too small will make the training process slow or even stagnant. Through experimental verification, a learning rate of 0.001 can ensure the smooth convergence of the model in this task.

3. Epochs: Epochs refers to the number of times the model traverses the entire training set. If the number of epochs is set too small, the model may not fully learn the data features; if it is set too large, it may face the risk of overfitting. This thesis sets epochs to 200 to prevent overfitting while ensuring that the model can converge.

4. Dropout: Dropout refers to the probability that neurons are randomly turned off during training, which prevents the model from overfitting. Dropout is an effective regularization technique that forces the network to learn more robust feature representations by randomly turning off some neurons during training. This thesis sets dropout to 0.1, which

can effectively control the risk of overfitting without excessively suppressing the model's expressiveness.

5. Image normalization: The image normalization parameter is used to normalize the input features to a standard interval. This thesis uses a mean of 0.5 and a standard deviation of 0.5 to map the original pixel values to the [-1,1] interval, which is conducive to rapid convergence and numerical stability of the model.

4.2.2 Network structure parameter settings

The ICNN-ECA network structure includes three functional modules: convolutional feature extraction, attention enhancement, and fully connected mapping. Table 4.2 lists the parameter configuration of each layer in detail.

Table 4.2: ICNN-ECA Model Architecture Parameters

Layer Name	**Input Size**	**Output Size**	**Kernel Size/ Number of Nodes**	**Activation Function**
Convolution Layer 1	4×10×1	4×8×2	(1,3)×2	ReLU
ECA Module 1	4×8×2	4×8×2	Adaptive	Sigmoid
Convolution Layer 2	4×8×2	4×6×4	(1,3)×4	ReLU
ECA Module 2	4×6×4	4×6×4	Adaptive	Sigmoid
Convolution Layer 3	4×6×4	3×5×8	(2,2)×8	ReLU
ECA Module 3	3×5×8	3×5×8	Adaptive	Sigmoid
Flatten Layer	3×5×8	120	-	ReLU
Fully Connected Layer 1	120	100	100	ReLU
Fully Connected Layer 2	100	20	20	ReLU
Fully Connected Layer 3	20	1	1	-

The ICNN-ECA network structure adopts an alternating nested design of CNN and ECA mechanisms. The convolution layer is responsible for extracting the spatial features of antenna parameters from the two-dimensional image representation, and a hierarchical feature expression is formed through three-layer convolution operations. The ECA module is directly attached to each convolution layer to achieve instant recalibration of the convolution output feature channel. The overall network structure adopts a combination of three-layer convolution, three-layer ECA modules and three-layer full connection. The parameter configuration of each layer is determined as the

current setting through experiments, realizing the mapping from image input to resonant frequency prediction.

4.2.3 Optimizer selection

This thesis compares the performance of three mainstream deep learning optimization algorithms in the task of predicting the resonant frequency of microstrip antennas. The stochastic gradient descent (SGD) algorithm is the most basic optimization method. It calculates the gradient of the loss function with respect to the parameters and updates the model parameters at a fixed learning rate. The Adam optimizer combines the momentum method with the adaptive learning rate mechanism. It not only stores the exponential moving average of the gradient, but also maintains the exponential moving average of the square of the gradient. It can automatically adjust the learning rate of each parameter. The RMSProp optimizer focuses on dealing with the swing problem in gradient descent. It adjusts the learning rate by performing an exponential moving average on the square of the gradient to balance the update step size in each direction. The experiment compares the performance of the three optimizers under the same conditions. Table 4.3 shows the change in loss of the three optimizers during training. Figure 4.2 shows the performance comparison of the three optimizers on the validation set.

Table 4.3: Performance Comparison of Different Optimizers

Optimizer	**Val. Loss (MSE)**	**Comp. Cost**	**Training Stability**
SGD	0.2061	Low	High
Adam	0.6967	High	High
RMSProp	0.1796	Medium	Low

4.2.4 Loss function selection

This thesis constructs a scientific and reasonable loss function system to evaluate the prediction performance of the model from multiple perspectives. Reference [37] points out that the loss function for deep learning regression problems mainly includes three categories: loss function based on absolute error, loss function based on relative error, and loss function based

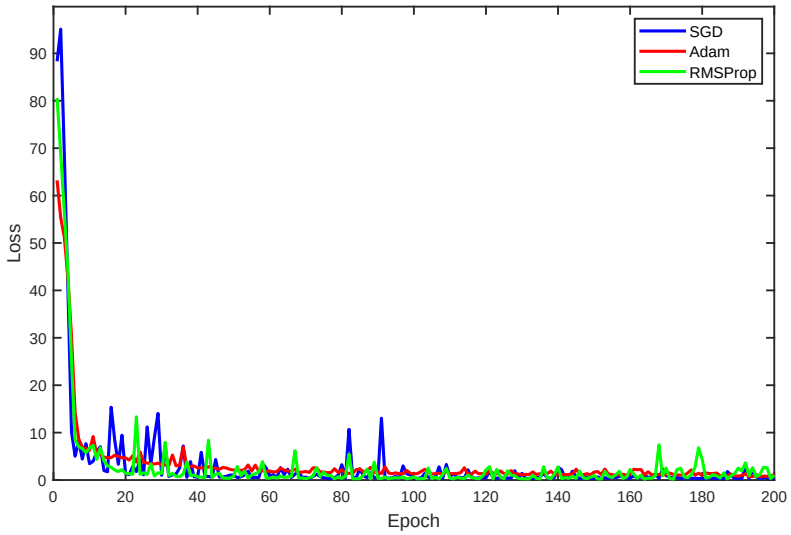

Figure 4.2: Loss comparison of different optimizers

on goodness of fit. Each type of loss function guides the model parameter update from different dimensions, and together constitutes a complete optimization framework.

The absolute error loss function measures the direct difference between the predicted value and the true value, mainly including mean square error (MSE), mean absolute error (MAE), and root mean square error (RMSE). MSE imposes higher weight on larger errors and serves as the main loss function to guide parameter optimization during training; MAE provides a linear measure of error, is insensitive to outliers, and is closer to intuitive understanding; RMSE, as the square root of MSE, is not directly used as a loss function, but maintains sensitivity to large errors. At the same time, its dimension is consistent with the original data, which is easy to explain.

The relative error loss function breaks through the absolute numerical limit and provides a proportional optimization target. This thesis explores two possible loss functions: mean absolute percentage error (MAPE) and symmetric mean absolute percentage error (SMAPE). MAPE calculates the percentage of the prediction error relative to the actual value, which is directly interpretable; SMAPE solves the instability problem that MAPE may appear when the actual value is close to zero by standardizing the error as the percentage of the mean of the predicted value and the actual value, and provides a more robust relative error optimization target.

The goodness-of-fit loss function evaluates the model's ability to explain data variation. The coefficient of determination (R^2) quantifies the proportion of the dependent variable variance explained by the model, and the ideal value is 1; this indicator usually ranges from 0 to 1, where 1 means

that the model perfectly explains all data variation and 0 means that the model fails to explain any variation. By minimizing $1-R^2$, the corresponding loss function can be constructed to guide the model to better capture the data regularity. A high R^2 value (> 0.99) indicates that the model can accurately describe the impact of parameter changes on the resonant frequency. However, R^2 is sensitive to extreme values, so it is often used in conjunction with other loss functions in practical applications. Comprehensively evaluate model performance. The corresponding equations are as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{4.1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{4.2}$$

$$RMSE = \sqrt{MSE} \tag{4.3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{4.4}$$

$$SMAPE = \frac{100\%}{n}\sum_{i=1}^{n}\frac{2|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i| + \epsilon} \tag{4.5}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.6}$$

The following algorithm 1 shows the pseudocode of the ICNN-ECA model for resonant frequency prediction.

4.3 Model performance evaluation

Table 4.4 shows the performance comparison results of the ICNN-ECA model and three comparison models on the test set. It can be clearly observed from the table that the ICNN-ECA model proposed in this thesis outperforms other comparison models in all evaluation indicators. In terms of absolute error indicators, the MSE of the ICNN-ECA model is 0.1971, which is 48.8% lower than 0.3848 of the baseline CNN model, 82.3% lower than 1.1159 of the CNN_SENET model, and 77.9% lower than 0.8901 of the CNN_LSTM model. This significant reduction shows that the ICNN-ECA model has a clear advantage in prediction accuracy.

In terms of relative error indicators, the ICNN-ECA model also shows superior performance. MAPE and SMAPE are 5.9613% and 5.6914% respectively,

Algorithm 1 ICNN-ECA Model for Resonant Frequency Prediction

function CNNRegressor(input)
    // Sequential convolution blocks with ECA modules
    $x \leftarrow \text{ReLU}(\text{Conv2D}(input, (1,3), channels = 2))$
    $x \leftarrow \text{ECAModule}(x, channel = 2)$
    $x \leftarrow \text{ReLU}(\text{Conv2D}(x, (1,3), channels = 4))$
    $x \leftarrow \text{ECAModule}(x, channel = 4)$
    $x \leftarrow \text{ReLU}(\text{Conv2D}(x, (2,2), channels = 8))$
    $x \leftarrow \text{ECAModule}(x, channel = 8)$
    // Fully connected layers
    $x \leftarrow \text{Flatten}(x) \rightarrow \text{FC}(120, 100) \rightarrow \text{Dropout}(0.1)$
    $x \leftarrow \text{FC}(100, 20) \rightarrow \text{FC}(20, 1)$
    return x
end function
function ECAModule(x, channel)
    $k \leftarrow \text{adaptive\_kernel\_size}(channel)$ ▷ Based on channel count
    $y \leftarrow \text{GlobalAvgPool}(x) \rightarrow \text{Conv1D}(k) \rightarrow \text{Sigmoid}()$
    return $x \times y$ ▷ Channel-wise attention
end function

which are significantly improved compared with other models. In particular, compared with the CNN_LSTM model, the MAPE of ICNN-ECA is reduced by 52.4% and the SMAPE is reduced by 58.6%, indicating that the ICNN-ECA model can maintain relatively stable prediction accuracy when processing antenna parameters of different scales.

The R^2 indicator reflects the model's ability to explain data variation. The ICNN-ECA model achieved the highest value of 0.9978, close to the ideal 1.0, indicating that the model can explain 99.78% of the variation in the resonant frequency. In contrast, the CNN_SENET model has the lowest R^2 value of 0.9877, while the standard CNN model and the CNN_LSTM model are 0.9958 and 0.9902, respectively. Although the R^2 values of all models are over 0.98, indicating that they all have strong fitting capabilities, the slight advantage of the ICNN-ECA model in this indicator still shows that it is better at capturing the complex nonlinear relationship between antenna parameters and resonant frequency.

In order to more intuitively observe the prediction accuracy of different models for the resonant frequency, Figure 4.3 shows the comparison results of ICNN-ECA and other comparison models with Ground truth on the test set. Table 4.5 shows the prediction results of different models on the test set, including the predicted values of the ICNN-ECA model and the corresponding relative errors.

Table 4.4: Performance comparison of different models

Model	MSE	MAE	RMSE	MAPE(%)	SMAPE(%)	R^2
CNN	0.3848	0.4994	0.6203	6.4769	6.6819	0.9958
CNN_SENET	1.1159	0.7252	1.0564	8.3061	8.815	0.9877
CNN_LSTM	0.8901	0.7674	0.9434	12.5327	13.7422	0.9902
CNN_ECA	0.1971	0.3374	0.444	5.9613	5.6914	0.9978

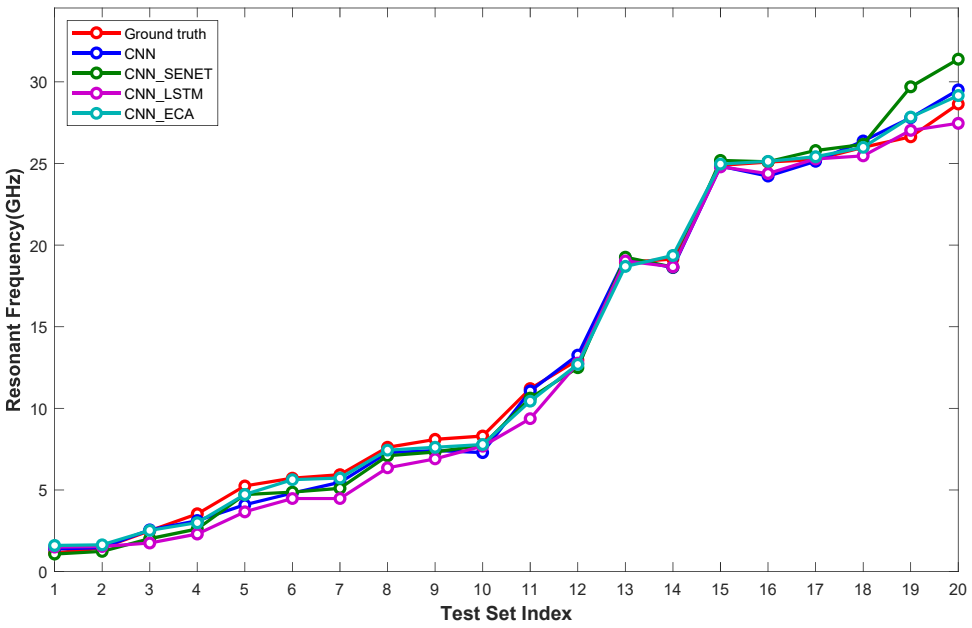

Figure 4.3: Prediction results of different models

Table 4.5: Prediction results of different models

$f(GHz)$	**CNN***	**CNN_SENET***	**CNN_LSTM***	**CNN_ECA***
1.14	1.39(+0.22)	1.07(-0.05)	1.50(+0.32)	1.60(+0.41)
1.44	1.47(+0.02)	1.23(-0.15)	1.54(+0.07)	1.64(+0.14)
2.50	2.56(+0.02)	2.02(-0.19)	1.75(-0.30)	2.52(+0.01)
3.54	3.12(-0.12)	2.61(-0.26)	2.30(-0.35)	3.00(-0.15)
5.25	4.09(-0.22)	4.72(-0.10)	3.66(-0.30)	4.72(-0.10)
5.72	4.81(-0.16)	4.87(-0.15)	4.48(-0.22)	5.63(-0.02)
5.94	5.47(-0.08)	5.10(-0.14)	4.47(-0.25)	5.73(-0.04)
7.61	7.29(-0.04)	7.09(-0.07)	6.36(-0.16)	7.44(-0.02)
8.10	7.42(-0.08)	7.33(-0.10)	6.90(-0.15)	7.62(-0.06)
8.31	7.29(-0.12)	7.74(-0.07)	7.68(-0.08)	7.78(-0.06)
11.21	11.07(-0.01)	10.63(-0.05)	9.37(-0.16)	10.44(-0.07)
12.96	13.26(+0.02)	12.49(-0.04)	12.77(-0.01)	12.69(-0.02)
18.94	19.22(+0.01)	19.27(+0.02)	19.03(+0.00)	18.70(-0.01)
19.15	18.62(-0.03)	18.65(-0.03)	18.67(-0.03)	19.36(+0.01)
24.89	24.83(-0.00)	25.19(+0.01)	24.80(-0.00)	24.97(+0.00)
25.09	24.23(-0.03)	25.11(+0.00)	24.39(-0.03)	25.14(+0.00)
25.24	25.13(-0.00)	25.80(+0.02)	25.28(+0.00)	25.41(+0.01)
26.00	26.37(+0.01)	26.17(+0.01)	25.47(-0.02)	25.99(-0.00)
26.63	27.80(+0.04)	29.70(+0.12)	27.03(+0.02)	27.84(+0.05)
28.65	29.51(+0.03)	31.39(+0.10)	27.47(-0.04)	29.17(+0.02)

*Values in parentheses show relative errors

Chapter 5

Conclusions and Future Work

This thesis proposes a method for predicting the resonant frequency of microstrip antennas based on ICNN-ECA. By converting the physical parameters of the antenna into a two-dimensional image representation and innovatively introducing the ECA module to enhance the network's perception of key electromagnetic features, the accurate mapping from antenna structural parameters to resonant frequency is successfully achieved. Experimental results show that the proposed ICNN-ECA model significantly outperforms the traditional CNN, CNN-SENET and CNN-LSTM models in multiple evaluation indicators, with the MSE reduced by 48.8-82.3% and the coefficient of determination R^2 reaching 0.9978, which verifies the effectiveness of this method in the field of antenna performance prediction. The main contribution of this thesis is that the ECA mechanism is applied to the field of microstrip antenna modeling for the first time, and the limitations of the traditional CNN architecture in dealing with differences in the importance of electromagnetic features are solved through lightweight design.

Future work will focus on the following aspects: expanding the scale of the data set and increasing the diversity of antenna structures, building a multi-task learning framework that can simultaneously predict multiple antenna performance parameters, exploring the joint application mechanism of spatial attention and channel attention, studying the transfer learning ability between different types of antennas, and building an end-to-end intelligent antenna design system. These extensions will further enhance the versatility and practicality of the model, provide more comprehensive support for the rapid design and optimization of microstrip antennas, and promote the development of wireless communication technology.

References

[1] C. A. Balanis, *Antenna Theory: Analysis and Design*. John Wiley & Sons, Dec. 2015.

[2] D. Xia, "Simulation analysis and optimization of microstrip patch antennas," Master's thesis, Huazhong University of Science and Technology, 2007.

[3] S. Koziel and A. Bekasiewicz, "Computationally-efficient multi-objective optimization of antenna structures using point-by-point pareto set identification and local approximation surrogates," in *2015 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*, Aug. 2015, pp. 1–3, https://ieeexplore.ieee.org/abstract/document/7414990.

[4] S. Fei-Yan, T. Yu-Bo, and R. Zuo-Lin, "Modeling the resonant frequency of compact microstrip antenna by the pso-based svm with the hybrid kernel function," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 29, no. 6, pp. 1129–1139, 2016, https://onlinelibrary.wiley.com/doi/abs/10.1002/jnm.2171.

[5] T. Zhang, "Application of gaussian process ensemble in microstrip antenna," Master's thesis, Jiangsu University of Science and Technology, 2021.

[6] J. Gao, Y. Tian, and X. Chen, "Antenna optimization based on co-training algorithm of gaussian process and support vector machine," *IEEE Access*, vol. 8, pp. 211 380–211 390, 2020, https://ieeexplore.ieee.org/document/9264163.

[7] J. Li, Q. Xu, X. He, Z. Liu, D. Zhang, R. Wang, R. Qu, and G. Qiu, "Cfformer: Cross cnn-transformer channel attention and spatial feature fusion for improved segmentation of low quality medical images," http://arxiv.org/abs/2501.03629, Jan. 2025.

[8] Y. Cao, X. Hu, and S. Yin, "Research on antenna inverse modeling method based on attention mechanism cnn-bilstm," in *The Proceedings of 2024 International Conference of Electrical, Electronic and Networked Energy Systems*, L. Jia, Y. Lv, Q. Yang, L. Xiong, D. Sun, and Y. Liu, Eds. Singapore: Springer Nature Singapore, 2025, vol. 1319, pp. 403–410, https://link.springer.com/10.1007/978-981-96-2050-0_42.

[9] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, Nov. 2017, https://ieeexplore.ieee.org/document/8103129/?arnumber=8103129.

[10] H. Fu, Y. Tian, F. Meng, Q. Li, and X. Ren, "Microstrip antenna modelling based on image-based convolutional neural network," *Electronics Letters*, vol. 59, no. 16, p. e12910, Aug. 2023, https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/ell2.12910.

[11] Z. Zhu, Y. Tian, and J. Sun, "Antenna modeling based on image-cnn-lstm," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 9, pp. 2738–2742, Sep. 2024, https://ieeexplore.ieee.org/document/10540172/.

[12] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 7, pp. 5626–5635, Jul. 2020, https://ieeexplore.ieee.org/document/9034483.

[13] J. P. Jacobs, "Accurate modeling by convolutional neural-network regression of resonant frequencies of dual-band pixelated microstrip antenna," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 12, pp. 2417–2421, Dec. 2021, https://ieeexplore.ieee.org/document/9541024.

[14] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *The Journal of Physiology*, vol. 148, no. 3, pp. 574–591, Oct. 1959, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1363130/.

[15] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, Apr. 1980, https://doi.org/10.1007/BF00344251.

[16] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, Dec. 1989, https://ieeexplore.ieee.org/abstract/document/6795724.

[17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, https://ieeexplore.ieee.org/abstract/document/726791.

[18] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006, https://www.science.org/doi/10.1126/science.1127647.

[19] r. K, u. S, and H. E, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, May 2017, https://dl.acm.org/doi/10.1145/3065386.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," http://arxiv.org/abs/1409.1556, Apr. 2015.

[21] C. Szegedy, e. L. W, a. J. Y, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE, Jun. 2015, pp. 1–9, http://ieeexplore.ieee.org/document/7298594/.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778, http://ieeexplore.ieee.org/document/7780459/.

[23] X. Zhang, Y. Tian, and X. Zheng, "Optimal design of fragment-type antenna structure based on pso-cnn," in *2019 International Applied Computational Electromagnetics Society Symposium - China (ACES)*, vol. 1, Aug. 2019, pp. 1–2, https://ieeexplore.ieee.org/document/9060460.

[24] Z. Wei, Z. Zhou, and M. Shen, "Fast and automatic parametric modeling of antenna structure using cnn-lstm network for efficient data generation," in *2023 IEEE Conference on Antenna Measurements and Applications (CAMA)*, Nov. 2023, pp. 61–64, https://ieeexplore.ieee.org/document/10352738.

[25] P. Singh, S. S. Panda, and R. S. Hegde, "Deep-learning empowered multi-objective antenna design: A polygon patch antenna case study," in *2024 National Conference on Communications (NCC)*, Feb. 2024, pp. 1–6, https://ieeexplore.ieee.org/document/10486033.

[26] E. A. Karahan, Z. Shao, and K. Sengupta, "Deep learning aided modelling and inverse design for multi-port antennas," in *2024 IEEE International Symposium on Antennas and Propagation and INC/USNC-URSI Radio Science Meeting (AP-S/INC-USNC-URSI)*, Jul. 2024, pp. 799–800, https://ieeexplore.ieee.org/document/10686871.

[27] Q. Wu, W. Chen, C. Yu, H. Wang, and W. Hong, "Machine-learning-assisted optimization for antenna geometry design," *IEEE Transactions on Antennas and Propagation*, vol. 72, no. 3, pp. 2083–2095, Mar. 2024, https://ieeexplore.ieee.org/document/10379000.

[28] XJ. Feng, ZW. Zhang, and J. Shi, "Text sentiment analysis based on convolutional neural networks and attention model," *Application research of computers*, vol. 35, no. 5, pp. 1434–1436, 2018.

[29] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, Jun. 2018, pp. 7132–7141, https://ieeexplore.ieee.org/document/8578843/.

[30] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, http://dx.doi.org/10.1109/CVPR42600.2020.01155.

[31] X. Li, W. Wang, X. Hu, and J. Yang, "Selective kernel networks," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, Jun. 2019, pp. 510–519, https://ieeexplore.ieee.org/document/8954149/.

[32] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Computer Vision – ECCV 2018*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham: Springer International Publishing, 2018, vol. 11211, pp. 3–19, https://link.springer.com/10.1007/978-3-030-01234-2_1.

[33] C. ZHANG, L. ZHU, and L. YU, “Review of attention mechanism in convolutional neural networks,” *Computer Engineering and Applications*, vol. 57, no. 20, pp. 64–72, 2021.

[34] A. Munir, I. Novianti, and B. Hasanah, “Experimental investigation of adm-based microstrip square patch antenna with resonant frequency lowering characteristic,” in *2020 International Workshop on Antenna Technology (iWAT)*, Feb. 2020, pp. 1–4, https://ieeexplore.ieee.org/abstract/document/9083912.

[35] P. F. Hu, Y. M. Pan, X. Y. Zhang, and S. Y. Zheng, “Broadband filtering dielectric resonator antenna with wide stopband,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 4, pp. 2079–2084, Apr. 2017, https://ieeexplore.ieee.org/abstract/document/7856943.

[36] F. He, T. Liu, and D. Tao, “Control batch size and learning rate to generalize well: Theoretical and empirical evidence,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, Eds., vol. 32. Curran Associates, Inc., 2019, https://proceedings.neurips.cc/paper_files/paper/2019/file/dc6a70712a252123c40d2adba6a11d84-Paper.pdf.

[37] V. Plevris, G. Solorzano, N. Bakas, and M. Ben Seghier, “Investigation of performance metrics in regression analysis and machine learning-based prediction models,” in *8th European Congress on Computational Methods in Applied Sciences and Engineering*. CIMNE, 2022, https://www.scipedia.com/public/Plevris_et_al_2022a.