

学校代码：10289
分 类 号：TN99
密 级：公 开
学 号：211210301106



江苏科技大学
硕士学位论文
(全日制专业学位)

图模型卷积神经网络自编码器
在天线设计上的应用

研究生姓名	符 浩	导师姓名	田 雨 波
申请学位类别	电 子 信 息 硕 士	学位授予单位	江 苏 科 技 大 学
学 科 专 业	电子信息	论文提交日期	2024 年 04 月 11 日
研 究 方 向	现代通信理论与系统	论文答辩日期	2024 年 05 月 23 日
答辩委员会主席	张友文教授	评 阅 人	盲审
			盲审

2024 年 05 月 23 日

分类号:	TN99
密 级:	公开
学 号:	211210301106

电子信息硕士学位论文

图模型卷积神经网络自编码器 在天线设计上的应用

学生姓名	符浩
指导教师	田雨波

江苏科技大学
二〇二四年五月

A Thesis Submitted in Fulfillment of the Requirements

for the Degree of Master of Electronic and Information Engineering

Image-based Convolutional Neural Network Autoencoder
Application in Antenna Design

Submitted by

Hao Fu

Supervised by

Tian Yu-Bo

Jiangsu University of Science and Technology

May, 2024

摘要

随着计算机技术和无线通信技术的发展,一些微波器件的优化设计问题变得越来越重要,例如接收和发射电磁波装置的天线,在地面通信、遥感遥测、雷达微波系统等应用领域发挥了重要的作用,因此如何高效快速的完成微波器件的优化设计逐渐成了一个重要的研究方向。目前较为流行的微波器件优化设计方法是通过全波电磁仿真软件结合全局优化算法进行设计。当使用三维电磁场仿真软件(CST Studio Suite, CST)、高频电磁仿真软件 (High Frequency Structure Simulator, HFSS)等电磁仿真软件对微波器件进行优化设计时,对于复杂的微波器件,其仿真是一项非常耗时且占据计算资源较多的工作。因此,基于机器学习(Machine Learning, ML)技术的代理模型辅助微波器件优化设计是一项非常有研究意义的工作,通过将微波器件的尺寸参数等作为输入并对代理模型进行训练,可以准确、快速的预测微波器件的电磁特性,大大减少了运算时间。

相较于传统的 ML 方法,深度学习(Deep learning, DL)理论被提出且得到了快速的发展,卷积神经网络(Convolutional Neural Network, CNN)作为 DL 的代表性成果也得到了极大的更新。由于在 CNN 模型中增加了局部感受野、参数的共享以及稀疏权重,相比其他的神经网络, CNN 具有平移和尺度不变性的特点,该特点使得网络能有效地对特定的特征进行识别,更适合于图像数据的处理。在本文中,采用像素化天线参数方式,将模型输入由一维表示为二维,类似于黑白图像,输出是天线的电气性能,实现基于卷积神经网络的精确建模。主要研究内容如下:

(1) 简要介绍图模型构建原理, CNN 和自编码器(Autoencoder, AE)建模理论,以及和贝叶斯优化(Bayesian Optimization, BO)的数学过程,最后给出 MATLAB 联合 HFSS 仿真方法流程。

(2) 通过引入图模型将天线数据像素化,映射到一个由 1 和 0 组成的二维矩阵,结合 CNN,成为图模型卷积神经网络(Image-based Convolutional Neural Network, IMCNN),以图模型构建和谐振频率预测回归为重点。通过矩形、圆形微带天线的谐振频率实验,验证了模型在预测天线谐振频率上的有效性和准确性。

(3) 研究基于图模型卷积神经网络自编码器(Image-based Convolutional Neural Network Autoencoder, IMAECNN)的增强模型。首先对算法结构进行阐述,研究数据增强技术并结合 IMCNN 网络,将其用于单频 PIFA 天线和 E 形微带天线的优化设计,得到较好的优化设计结果,并在 HFSS 上进行验证,证明在天线设计上的可行性。

(4) 研究了贝叶斯优化的图模型卷积神经网络(Image-based Convolutional Neural

Network Autoencoder based on Bayesian Optimization, BOIMAECCN), 论述了 IMAECCN 与 BO 算法结合的实现过程, 包括这两种先进计算方法相结合的理论基础、算法调整和实践研究, 探讨了 IMAECCN 的深度学习能力与 BO 的策略优化能力的协同作用, 为解决预测和优化问题提供了一种新的精准有效的方法, 通过双频微带天线和双频单极子天线实例验证了 BOIMAECCN 的建模能力。

关键词 图模型; 卷积神经网络; 自编码器; 微波器件; 优化设计

Abstract

As computer technology and wireless communication have advanced, the optimizing issue of microwave devices design has become increasingly important. For example, antennas that receive and transmit electromagnetic waves play a critical role in fields such as terrestrial communication, remote sensing, telemetry, and radar microwave systems. Consequently, the efficient and rapid optimization of microwave device design has gradually become a significant research direction. At present, a popular method for the optimization design of microwave devices involves the full-wave electromagnetic simulation software combined with global optimization algorithms. When using 3-D electromagnetic field simulation software such as CST Studio Suite or High Frequency Structure Simulator (HFSS) for the optimization design of microwave devices, the simulation process for complex devices can be extremely time-consuming and consume considerable computational resources. Therefore, the use of machine learning (ML) technology as surrogate model to assist in the optimization design of microwave devices is a research endeavor of significant merit. By using the dimensional parameters of microwave devices as inputs and training the surrogate model, the electromagnetic characteristics of the devices can be predicted accurately and quickly, substantially reducing the computational time.

Compared to traditional ML methods, the theory of deep learning (DL) has been proposed and has experienced rapid development. As a representative achievement of DL, Convolutional Neural Network (CNN) has undergone significant advancements. Owing to the incorporation of local receptive fields, shared parameters, and sparse connectivity in CNN models, they possess translational and scale invariance characteristics unlike other neural networks. These characteristics enable the network to effectively recognize specific features, making them more suitable for processing image data. In this study, antenna parameters are pixelated, transforming the model input from one-dimensional to two-dimensional representation, akin to black and white images. The output is the electrical performance of the antenna, thus achieving precise modeling based on CNN. The main research contents are as follows:

(1) This section succinctly elucidates the foundational principles of image model construction, the theoretical underpinnings of CNN, and Autoencoder (AE) modeling, in conjunction with the mathematical procedures of Bayesian Optimization (BO). Subsequently,

a procedural workflow is delineated that integrates MATLAB with HFSS simulation methodologies.

(2) The integration of an image model facilitates the pixelization of antenna data, mapping it onto a two-dimensional matrix composed of binary elements. In harmony with CNN, this synthesis yields the Image-based Convolutional Neural Network (IMCNN). Through empirical analysis of the resonant frequencies of rectangular and circular microstrip antennas, the efficacy and accuracy of the model in predicting the resonant frequencies of antennas have been substantiated.

(3) The enhancement model of Image-autoencoder Convolutional Neural Network (IMAECNN) based on Convolutional Neural Network model is studied. Firstly, the structure of the algorithm is described, the data enhancement technology is studied and combined with the IMCNN network, which is applied to the optimal design of single-frequency PIFA antenna and E-shaped microstrip antenna. The feasibility of the antenna design is verified by HFSS.

(4) This research delves into the Bayesian Optimization-enhanced Image-autoencoder Convolutional Neural Network (BOIMAECNN), initially detailing the integration process of IMAECNN with the BO algorithm. It discusses the realization process of the combination of IMAECNN and BO algorithm, including the theoretical basis, algorithm adjustment and practical research of the combination of these two advanced computing methods. The synergy between IMAECNN's deep learning ability and BO's strategy optimization ability is discussed, which provides a new accurate and effective method for solving the prediction and optimization problems. Furthermore, the modeling ability of the BOIMAECNN is verified by the examples of dual-frequency microstrip antenna and dual-frequency monopole antenna.

Keywords Image model; Convolutional Neural Network; Autoencoder; Microwave devices; Optimal design

目录

摘要.....	I
Abstract	III
第 1 章 绪论.....	1
1.1 研究的目的及意义	1
1.2 国内外研究现状	2
1.3 主要研究内容及结构安排	5
第 2 章 理论基础.....	7
2.1 图模型算法简介	7
2.1.1 图模型基础概念	7
2.1.2 图模型构建特征	7
2.2 卷积神经网络简介	8
2.2.1 卷积神经网络基本知识	8
2.2.2 卷积神经网络建立评估	10
2.3 自编码器网络简介	11
2.3.1 自编码器基本知识	11
2.3.2 自编码器基本分类	12
2.4 贝叶斯优化算法简介	13
2.4.1 贝叶斯优化算法基本知识	13
2.4.2 贝叶斯优化算法基本分类	14
2.5 HFSS 电磁仿真软件及其调用方法	15
第 3 章 基于图模型卷积神经网络在微带天线谐振频率建模	17
3.1 引言	17
3.2 图模型卷积神经网络系统	17
3.2.1 图模型卷积神经网络基本原理	17
3.2.2 图模型卷积神经网络基本流程	19
3.3 图模型卷积神经网络系统对矩形微带天线谐振频率预测	20
3.3.1 矩形微带天线天线基本介绍	20
3.3.2 矩形微带天线谐振频率预测	21
3.4 图模型卷积神经网络系统对圆形微带天线谐振频率预测	24
3.4.1 圆形微带天线天线基本介绍	24
3.4.2 圆形微带天线谐振频率预测	25
3.5 本章小结	27
第 4 章 基于图模型卷积神经网络自编码器的微带天线优化设计	29
4.1 引言	29
4.2 图模型卷积神经网络自编码器系统	29
4.2.1 图模型卷积神经网络自编码器基本原理	29
4.2.2 图模型卷积神经网络自编码器基本流程	30
4.3 图模型卷积神经网络自编码器对单频 PIFA 天线的优化设计	32
4.3.1 单频 PIFA 天线基本介绍	32
4.3.2 单频 PIFA 天线优化设计	33
4.4 图模型卷积神经网络自编码器对 E 形微带天线的优化设计	35

4.4.1 E 形微带天线基本介绍.....	35
4.4.2 E 形微带天线优化设计.....	36
4.5 本章小结	38
第 5 章 基于贝叶斯优化图模型卷积神经网络自编码器的天线建模.....	39
5.1 引言	39
5.2 贝叶斯优化图模型卷积神经网络自编码器系统	39
5.2.1 贝叶斯优化图模型卷积神经网络自编码器基本原理	39
5.2.2 贝叶斯优化图模型卷积神经网络自编码器基本流程	41
5.3 贝叶斯优化图模型卷积神经网络自编码器对双频微带天线建模	43
5.3.1 双频微带天线模型	43
5.3.2 双频微带天线的回波损耗建模	44
5.4 贝叶斯优化图模型卷积神经网络自编码器对双频单极子天线建模	47
5.4.1 双频单极子天线模型	47
5.4.2 双频单极子天线回波损耗建模	47
5.5 本章小结	51
总结与展望	53
参考文献	55
攻读硕士学位期间学术成果	61
致谢	63

第1章 绪论

1.1 研究的目的及意义

在天线工程领域,经常需要对一些具有复杂结构或复杂媒质组成的复杂多目标天线进行优化设计,比如多频带或超宽带天线、可重构或异形天线等。由于这些天线一般都存在大量的不连续性,在进行优化设计时,很难用一个显式来表达设计变量与多目标函数之间的关系,即很难找到对应的麦克斯韦方程组和边界条件进行精确求解,通常需要借助数值仿真计算或电磁仿真软件来建立这种关系。当采用基于数值计算方法的全波电磁软件进行仿真计算时,其计算结果精确,和实际情况吻合,可以得到精确的多目标电气特性,但其计算量、存储量与准确性成正比,若需得到精确的计算结果,则需要耗费大量的计算时间。所以,对于复杂天线,可以考虑以电磁场理论为基础,以高性能计算技术为手段,运用数据处理技术和人工智能技术等提供的各种方法,建立起天线参数和电气性能之间高精度代理模型,实现优化设计的目的。

迄今为止,对基于天线参数建立图模型鲜有研究,主要原因是机器学习从二十世纪发展至今已有成熟的算法解决对天线原始参数的预测优化问题,比如线性回归算法、贝叶斯算法、高斯过程等,但是随着 CNN 在图像领域愈发凸显的卓越性能,使认识到天线参数与 CNN 另一种结合方式的可能,所以在本次课题研究中首次采用天线图模型结合 CNN。

21 世纪以来,深度学习理论被提出且得到了快速的发展,同时 CNN 算法得到了极大的更新。由于在 CNN 模型中增加了局部感受野、参数的共享以及稀疏权重^[1],相比其他的神经网络, CNN 具有平移和尺度不变性的特点,该特点使得 CNN 能有效地对特定的特征进行识别,更适合于图像数据的处理^[2],并将 CNN 与其他模型相结合以解决微波器件应用中不同的问题。

AE 是深度学习中的一种非常重要的无监督学习方法,能够从大量无标签的数据中自动学习,得到蕴含在数据中的有效特征。因此,自编码方法近年来受到了广泛的关注,已成功应用于很多领域,例如数据分类、模式识别、异常检测、数据生成等。

神经网络在训练过程中容易陷入局部最优问题,这一挑战已成为深度学习领域的一个核心问题。局部最优意味着优化算法可能会在损失函数的一个非全局最优优点停止,从而影响模型的性能和泛化能力。为了解决这个问题,研究学者们投入了大量的努力,探索和发展了多种优化策略,以增强神经网络训练过程中的搜索能力,其中包括传统的,遗传算法^[3](Genetic Algorithm, GA)、粒子群^[4](Particle Swarm Optimization, PSO)等算法,但这类方式通常需要进行大量的迭代,且可能陷入局部最优解。鉴于此,针对

复杂天线的高精度建模和优化设计问题, 本研究以贝叶斯算法优化求解模型并将其应用在复杂天线的设计和 optimization 问题上^{[5][6][7]}: 以 BO 算法框架的概率代理模型, 根据天线的设计指标构建目标函数, 通过设计新的采集函数在拟定的设计空间内进行迭代并采样, 得到符合设计目标的最佳样本点, 最后对这些样本点进行权威的电磁仿真软件的验证实验, 验证算法是否完成拟定的天线设计指标。

本课题基于 IMCNN 和自编码器对微波器件的分析问题, 十分新颖且富有挑战性的研究方向; 研究成果将为天线高精度建模和快速优化设计提供可行策略和技术支持, 促进天线工程、深度学习等领域的发展, 因此具有广阔的应用前景。

1.2 国内外研究现状

1.2.1 图模型研究现状

Wenyi Shao^[8]提出研究了一种基于深度学习的微波图像重建方法。该神经网络能够将从 4GHz 的 24×24 天线阵列获得的测量微波信号转换为 128×128 图像。为了降低训练难度, 使用自动编码器, 用 256×1 个向量表示高分辨率图像(128×128); 使用神经网络, 将微波信号映射到压缩特征(256×1)向量。然后将两个神经网络组合成一个完整的网络进行重建, 两阶段训练方法降低了深度学习网络逆重构训练的难度。Jan Pieter Jacobs^[9]研究了一种基于 CNN 回归的双波段像素微带天线的两个谐振频率的精确建模方法, 该方法以天线的整个像素表面的表示为输入。比较了 CNN 与浅层和深层传统前馈神经网络结构的预测性能; 采用粒子群优化算法 (Particle Swarm Optimization, PSO) 确定结构和学习算法超参数。CNN 在所有网络中获得了最好的预测结果, 对两个共振频率的测试预测的平均相对误差分别为 0.13% 和 0.22%。

1.2.2 卷积神经网络研究现状

张新宇讨论了一种将 PSO 与 CNN 相结合的优化算法^[10], 用于设计高维、紧凑的段天线, 将 PSO-CNN 应用于分段式天线的优化设计, 与传统仿真软件的结果进行比较, 得到了快速优化的结果。Sruthy Skaria^[11]在文章中使用微型雷达传感器捕获 14 种不同手势的多普勒特征, 并训练深度卷积神经网络 (Deep Convolutional Neural Networks, DCNN) 对这些捕获的手势进行分类, 该体系结构的分类结果表明, 手势分类准确率超过 95%。Jiaxin Cai^[12]研究了一个基于深度学习的天线选择的问题, 将生成训练的标签天线系统通过最大化信道容量, 然后在通道矩阵上采用 DCNN 来明确地挖掘大量潜在衰减系数的提示, 最后采用的 CNN 来分配类标签来选择最优天线子集。Bin Wang^[13]在文中提到对于雷达和电子战系统中进行循环扫描的天线, 基于阈值测量的方法是无效的; 为了克服这一缺点, 讨论了一种 CNN 在雷达和电子战系统的天线都是环形扫描

的情况下识别雷达的天线扫描周期方法,用于快速准确分类的 CNN 模型,利用大量接收到的三个独立天线扫描周期的信号时间功率图像,在不同的实验条件下对所建立的模型进行训练和测试,数值实验结果和性能对比表明,在雷达和电子战系统天线均为圆形扫描的情况下,该方法具有较高的分类精度和有效性,当信噪比不小于 30dB 时,雷达天线扫描周期的平均识别精度可达 90%以上。Ricardo Lovato^[14]提出一种基于 CNN 的相控阵波束形成方法,给定期望的二维辐射方向图,神经网络计算合成该方向图所需的阵列相位,并且能够合成复杂的辐射方向图,利用 8×8 微带贴片天线阵列对神经网络进行训练,并通过仿真验证了神经网络的准确性。Mingxiang Stephen Li^[15]研究了一种基于机器学习的方法的分频天线,利用 CNN 来估计太赫兹通信频段的到达方向(Direction of Arrival, DOA),天线被故意打破对称性,产生准随机的辐射图,而 CNN 则捕捉辐射图与其各自到达角度之间的关系。实验表明,DOA 估计结果收敛到最小验证均方误差(MSE)为 3.9°,均方根误差(RMSE)为 1.9°。

1.2.3 自编码器研究现状

Aya Mostafa Ahmed^[16]研究了一种鲁棒的天线阵列缺陷相干源 DOA 估计算法。利用当前深度学习的进展来克服最先进的 DOA 算法面临的两个最常见的问题(即相干源和阵列缺陷)。研究了一种深度自动编码器,该模型能够在不需要空间平滑的情况下正确地解析相干源。与文献中使用的方法相比,结果显示有显著改善。Jiyun Tao^[17]在文中提到由于模拟移频阵列的硬件限制,数字和模拟波束形成器的联合设计是一个非凸优化问题。针对这一问题,通过将混合波束形成系统映射到自编码器(Auto Encoder, AE)神经网络,研究了一种智能混合波束形成设计方法,基于 AE 神经网络。这种方法将原本的非凸优化问题转化为神经网络的训练任务,通过训练神经网络,自动地设计出波束形成器和组合器。Yifei Tong^[18]在该文中提到为了实现对阵列天线装配精度的准确、高效预测,研究了一种基于 AE 和带 boost 的在线序列核极值学习机的预测方法。实验表明 Boosting-OSKELM 和人工神经网络(Artificial Neural Network, ANN)的平均 MSE 分别为 0.061 和 0.12,耗时分别为 0.85s 和 15s。Kwi Seob Um^[19]研究了一种基于人工神经网络的多波段天线优化设计方法。该网络是基于自动编码器的代理模型,由普通神经网络和自动编码器两部分组成。前神经网络首先对编码后的天线特性进行预测,然后对预测数据进行解码,得到天线特性。在训练编码器获得特征签名向量后,前端神经网络只对特征签名向量进行回归,降低了神经网络的复杂度和参数数量。这不仅减少了训练时间,而且大大减少了所需的训练数据数量。Guangfei Qi^[20]在文中研究了粒子群优化是一种基于群体协作的进化算法,可以有效地处理任意优化问题。因此,粒子群算法在天线设计中得到了广泛的应用。传统的天线设计方法直接使用粒子群算法和电磁模拟器,在每次优化迭代中都需要进行大量的电磁模拟,这非常耗时。针对

这种情况,提出一种基于深度卷积自编码器(Deep Convolutional Autoencoder, DC-AE)辅助粒子群优化的天线设计方法。利用电磁求解器采集的天线数据训练 DC-AE,建立代理模型。该代理模型可以快速准确地预测天线未知物理参数的响应,从而加快了粒子群算法的迭代过程。最后以天线设计实例对该方法进行了验证。Seyed Pouya Shojaei^[21]在文中比较了基于编码器的波束形成器与 CNN 和最小方差无失真响应方法在信噪比方面的性能。结果表明,基于编码器的方法的平均信噪比为 25.82dB,而 CNN 方法的平均信噪比为 22.40dB,MVDR 方法的平均信噪比为 17.64dB。基于编码器的方法的性能优于 CNN 方法,但远优于最小方差无失真响应方法。基于编码器的方法比 CNN 方法平均高出 3.42dB,比最小方差无失真响应方法平均高出 8.18dB。Mehmet Saygin Seyfioglu^[22]在文中工作表明,利用无监督预训练和卷积滤波支持的局部空间特征的卷积 AE 架构优于其他深度架构,如 CNN,AE 以及采用预定义特征的多类支持向量机。与其他深度学习架构相比,CNN 和 AE 的性能提升分别为 4.1%和 10.1%。

1.2.4 贝叶斯优化研究现状

贝叶斯优化(Bayesian Optimization, BO)是一种高效的全局优化方法,它核心依赖于概率代理模型和采集函数。这种方法的目标是在给定的设计空间中,通过迭代过程中不断采集新的样本观测值,从而快速地逼近并找到目标函数的最大值。此过程通过精心选择样本点来更新概率模型,优化采集函数,以此来指导下一步的样本选择,有效地平衡探索与利用,确保以最小的迭代次数达到最优解。近些年来,研究改进 BO 是一个热门方向,对于 BO 的改进离不开新的代理模型建立^[23]、代理模型的优化^[24]和采集函数的设计^[25]这三个方面。

构建合适的代理模型能够提升 BO 的性能,如使用随机森林为代理模型^[26],通过构造多棵决策树来提高计算效率和精度;也有学者构建深度神经网络作为代理模型^[27],深度神经网络能够很好的处理大规模复杂数据,由此能够弥补 BO 对复杂样本的处理能力。Snoek 等人^[28]研究了动态调整代理模型域边界的方案用以解决高维度数据处理困难的问题;David 等人^[29]设计了新的 BORING 算法应用在处理数据的降维,将高维度、复杂度高的数据样本进行了数据的降低维度处理,从而提高了算法处理高维度的数据样本能力。对采集函数的改进设计也广受关注,如通过并行 BO^[30]中的采集函数能够实现同时采集到多个样本点从而提高 BO 的采样效率,同时评估多组样本点对应的目标函数,提升算法的求最优解能力。Ginsbourg 等人^[31]研究蒙特卡洛对采集函数进行模拟从而一次产生多个样本点;也有学者采用多模态方法^[32]求解多个候选样本点,以减少模拟方式的耗时问题。2018 年,Tracey 等人^[33]实现了 STP 和期望改进函数结合的 BO,并将其应用在气动结构设计问题上;2020 年,Clare 等人^[34]分别使用改进的期望后悔最小化函数和置信界限最小化函数与 STP 结合组成新的 BO,从而增强了 BO

的能力。

对于 BO 在电磁优化问题上的应用的研究进展,国内外的研究人员展开了广泛的理论研究,分别从 BO 的结构上^[35]、采样策略^[36]、多目标优化^[37]上对电磁优化问题开展研究。如:有研究学者研究了一种两阶段 BO^[38]应用于集成电路的时钟偏移最小化和集成电压调节器的效率最大化的优化设计上,与传统的 BO 相比,两阶段 BO 优化框架通过减少所需要的迭代次数,从而缩短优化过程的时间,由此提升了优化设计的效率和准确性。有文献研究了一种并行 BO^[39]的采集函数改进策略,其采集函数在同一次迭代中能够采集到多组样本点,并行计算每组样本点出的输出,由此快速的完成设计区间内的样本点采集,改善整个优化过程的收敛性,能够使得算法更迅速的找到潜在的最优解,并将所设计的算法应用在超宽带微带天线设计上。在文献[40]中,研究人员使用了 BO 优化方法来设计和优化一个工作在 60GHz 频段的硅基阶梯形喇叭天线,该天线采用介质集成波导进行馈电,并且电磁波通过一个过渡结构向上传播,这个过渡结构可以用于实现硅衬底和空气之间的阻抗匹配;为了提高阻抗匹配和辐射增益,在过渡层上安装了一个由多层的硅层组成的阶梯型喇叭,并且通过在过孔中加入空气腔来改善性能;通过进行 90 次迭代,最终得到了一个阻抗带宽为 10.4%的优化结果。

有研究学者提出解决未知向量值函数的多目标优化问题^[41],通过所提算法计算出最坏情况下的遗憾边界解决了多任务内核建模任务之间相互依赖的问题。文献[42]提出将三个不同的高斯过程(Gaussian Process, GP)堆叠在一起进行 BO 多目标任务的扩展,并应用在热力学有限元模型的优化上。

1.3 主要研究内容及结构安排

本文主要研究 IMCNN,希望通过对其模型改进得到精度较高的回归优化模型。将 AE、BO 的思想引入 IMCNN 中,优化模型精度并将其应用于微波器件的优化设计问题中,促进 IMCNN 在微波器件优化设计的电磁领域中的发展应用,具体安排如下:

第一章绪论。首先介绍了本课题的研究目的与现实价值意义,从处理数据类型的角度,引入图模型系统解决如何精确对中、小型电磁参数数据进行回归预测的问题;介绍了图模型, CNN, AE 以及 BO,并展开研究其在微波器件优化设计等电磁领域的国、内外研究现状。

第二章基本原理。简要介绍图模型构建原理, CNN 和 AE 以及 BO 建模理论,最后给出 MATLAB 联合 HFSS 仿真方法流程。

第三章介绍了通过引入图模型将天线数据像素化,映射到一个由 1 和 0 组成的二维矩阵,这与具有标准天线的谐振频率建模形成对比,并结合 CNN,成为 IMCNN,

将其用于矩形微带天线，圆形微带天线的谐振频率预测实验。

第四章研究基于 IMAECNN 的模型。详尽地描述了提出的算法改进及其详细架构，强调了其在天线设计优化中的应用潜力和创新性。特别地，该改进算法被应用于两种不同类型的天线设计：单频 PIFA 天线和 E 形微带天线，旨在通过精细的优化过程，改善它们的性能指标。随后，为了证实所提出的改进算法在实际天线设计中的有效性和可行性，将优化后的设计结果提交给了业界标准的全波电磁场仿真软件 HFSS 进行了严格的验证。

第五章研究了贝叶斯优化原理，将之应用于 IMAECNN，该算法依据预定的设计指标，分别对双频微带天线(Dual-Frequency Microstrip Antenna, DFMA)和双频单极子微带天线(Dual-Frequency Monopole Antenna, DFMOA)进行了性能预测。为了验证这些预测结果的准确性，将它们在全波电磁仿真软件中进行了实验验证。通过这种方法，算法对于 DFMA 和 DFMOA 的预测结果得到了精确的验证，进一步证明了该改进算法在微波器件设计领域上的可靠性。这不仅展示了算法在预测复杂天线性能方面的有效性，也突出了其在精确设计和优化微波器件方面的应用潜力。

最后对本文进行了总结和展望，并进行致谢。对全文工作和取得成果进行总结、并对课题研究做出进一步展望。

第2章 理论基础

2.1 图模型算法简介

2.1.1 图模型基础概念

在传统的一维信号处理任务中，如天线谐振频率信号分析，数据通常以一维数组的形式表示。这些一维数据可以直接输入到一维 CNN 中进行处理。一维 CNN 专门设计用于处理时间序列数据或任何形式的序列化信号数据，它通过在数据的一维序列上滑动卷积核来提取特征。可以考虑将这些数据转换为二维形式以便使用二维 CNN，在此过程中需要对数据进行重新编码或转置操作。这种转换的目的是将原始的一维信号数据映射到一个二维平面上，从而形成一个二维图像。

2.1.2 图模型构建特征

(1) 完备性：完备性是衡量一个系统、模型或数据结构能够多么全面地表示其目标问题空间的属性。在讨论二维二值图形与一维向量的关系时，完备性特指这样一个能力：系统内的表示方法能够覆盖问题空间中的所有可能情况。也就是说，对于问题空间内的每一个存在的一维向量，图模型空间中必须有至少一个二维二值图形与之相对应。这个属性确保了系统的全面性，意味着在应用模型来分析或解决问题时，不会有遗漏的情况，从而可以全面地理解 and 处理问题空间中的所有可能性。

(2) 健全性：健全性是模型或数据结构的准确性和有效性。具体到二维二值图形与一维向量的映射关系中，健全性意味着图模型空间中的每一个图形都必须正确无误地表示或对应到问题空间中的一个特定一维向量。这一属性保证了系统输出的可靠性和有效性，确保了模型不会产生误导信息或错误结果。

(3) 非冗余性：非冗余性强调的是模型或数据结构的效率和简洁性。在二维二值图形与一维向量的映射关系中，非冗余性确保了一一对应的关系，每一个问题空间中的一维向量都唯一对应一个图模型空间中的二维二值图形，反之亦然。这意味着系统中不存在多余的、重复的表示，每个实体或情况都有一个且只有一个清晰、准确的表示。非冗余性的好处在于它提高了模型的效率，减少了处理时间和资源消耗，同时也减少了理解和分析模型的复杂性。

总体而言，这三个属性共同构成了模型设计的核心特性。它们确保了模型不仅能全面覆盖并准确表示问题空间，而且能以高效且简洁的方式实现这一目标。在实际应用中，平衡这些属性往往需要精心设计和调整，以确保模型既能满足实际需求，又能

保持高效性和可用性。

2.2 卷积神经网络简介

2.2.1 卷积神经网络基本知识

CNN 利用卷积核执行卷积运算，目的是从输入图像中提取隐含的特征。这一过程通过在图像的各个部分滑动卷积核并执行逐元素乘法累加操作来完成，从而捕捉到图像中的局部模式，如边缘、颜色变化等。随后，CNN 将这些提取出的特征组合在一起，形成更加复杂和抽象的特征表示。

紧接着，所提取的特征将经过池化（或下采样）处理，这一步骤旨在减少数据的维度和复杂性，同时保留最重要的信息。池化操作通常包括最大池化或平均池化，它们分别保留了特征图中的最大值或平均值，从而提高了模型的泛化能力，并减少对位置的敏感性。

输入的图像可能会经过多个这样的卷积和池化层的处理，每一层都旨在从前一层的输出中提取更加复杂和高级的特征。通过一系列的卷积和池化操作，CNN 能够从简单的边缘检测逐步过渡到识别图像中的复杂对象和模式。

最终，经过一次或多次的卷积和池化处理后，图像的特征被送入一层或多层全连接的神经网络中。这些全连接层的输出神经元负责根据图像的特征进行分类或其他形式的特征学习，输出模型的最终决策。这个过程使得 CNN 不仅能够在图像处理和识别任务中表现出卓越的性能，也能够在保持高效计算的同时，学习到图像的深层次特征表示。如图 2.1 所示。

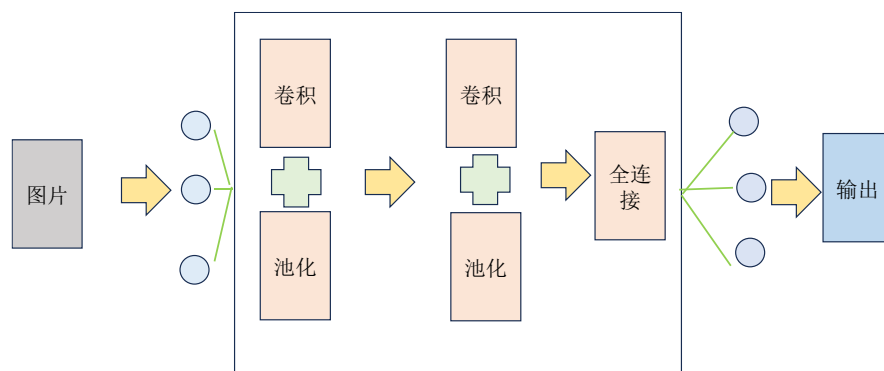


图 2.1 CNN 基本结构

Fig.2.1 Basic structure of CNN

构建 CNN 模型的过程中，不同的层如卷积层、池化层、全连接层和输出层被按照一定的规则排列组合。模型在连续训练的过程中，通过逐步调整这些层中的权重参数，逐渐优化到最理想的状态。CNN 模型的操作主要依赖于两个基本的流程：前馈传播和反向传播。在前馈传播过程中，输入图像被送入模型，并经过每一层的处理，最终模

型输出对应的预测结果。反向传播则是根据预测结果与实际结果之间的差异，从输出层向前逐层调整权重，通过这种方式不断迭代，直至模型性能稳定。

(1) 卷积层

卷积层是 CNN 特征提取很重要的一部分，其原理是作为一个特征提取器，提取数据集中的特征数据。卷积层的基本单元是一个多输入单输出的神经元，其基本原理可如图 2.2 所示。

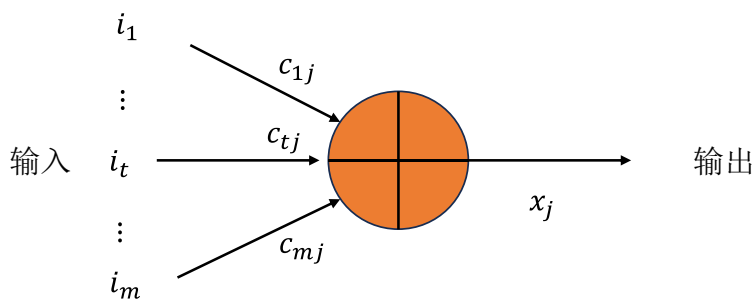


图 2.2 卷积神经元的基本原理

Fig.2.2 Basic fundamentals of convolutional neurons

图 2.2 中 i_m 表示样本输入数据， m 个样本数据同时输入到神经元 j ， c_{pj} 表示卷积层中设置的卷积核数据，也就是连接神经元的权重值，其中 $p=1,2,\dots,m$ 。那么，上述卷积神经元的整体过程可以由公式(2.1)表示。

$$x_j = f\left(\sum_{i=1}^m (i_m * c_{pj}) + b_j\right) \quad (2.1)$$

其中 $f(\bullet)$ 为激活函数，比较常见的是使用 *Sigmoid* 函数。其表达式如公式(2.2)所示；CNN 的训练就是对权重 c_{pj} 和偏置 b_j 进行训练。

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

例：若样本数据为 3×3 的矩阵数据集，设置卷积核为 2×3 ，其输出则为一个 2×1 的矩阵数据集。计算如公式(2.3)所示，其中 \otimes 表示卷积符号，公式(2.4)表示。

$$\begin{bmatrix} i_{11} & i_{12} & i_{13} \\ i_{21} & i_{22} & i_{23} \\ i_{31} & i_{32} & i_{33} \end{bmatrix} \otimes \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} \quad (2.3)$$

$$x_{11} = i_{11} \times c_{11} + i_{12} \times c_{12} + i_{13} \times c_{13} + i_{21} \times c_{21} + i_{22} \times c_{22} + i_{23} \times c_{23} \quad (2.4)$$

也就是说，卷积操作给输入的每个数据添加了不同的权重。因此通过合理的调整卷积核中的权重参数值，可以更高效的提取数据特征。

对输入图像进行特征提取时，卷积层展现了其独特的特点，包括局部感知能力、权值共享以及采用多个卷积核。这与全连接层相比，卷积层的这些特性为图像处理提供了优势。

(2) 池化层

池化层通过其降维、增强不变性、特征选择和层次性增强的功能，不仅提高了 CNN 的效率和效果，还提高了模型对于输入数据变化的鲁棒性。

考虑到微波器件的数据大部分是包含其尺寸参数和其对应的回波损耗扫频值的中、小型数据集，故在本文选择平均池化来构建特征组。平均池化是对空间信息求平均，不需要新的参数，因此对于数据空间变化的鲁棒性更强。平均池化由式(2.5)表示。

$$y_i = \frac{1}{s} \sum_s^{vs} x^{l(u,k)} \quad (2.5)$$

其中， y_i 表示第 i 个神经元平均池化后的结果， s 为池化卷积核的大小尺寸， $l(u,k)$ 表示第 l 层的第 u 通道的第 k 个输出神经元信息。 vs 表示卷积核的尺寸长度，例如图 2.3 所示。

(3) 全连接层

在 CNN 中，全连接层起着至关重要的作用，它的特点是层中的每个神经元都与上一层的每个神经元连接，由此得到了“全连接”的名称。其主要功能是基于从卷积层和池化层提取的高阶特征，对这些特征进行最终的分类或决策，从而将特征映射到样本标签或其他连续输出空间上。

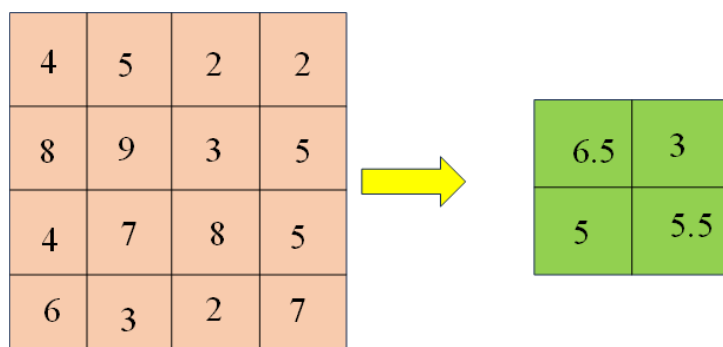


图 2.3 平均池化示意图

Fig.2.3 Schematic diagram of average pooling

2.2.2 卷积神经网络建立评估

CNN 在处理数据时，归一化处理的样本数据经提取得到特征数据，利用标签数据与样本输入数据计算伪逆矩阵来获得输入样本与输出标签数据之间的连接矩阵，并根据连接矩阵对相关样本数据进行预测。

预测后的数据需要进行评估，评估一个回归系统的预测性能有很多指标，比如平均绝对误差(Mean Absolute Error, MAE)、平均绝对百分误差(Mean Absolute Percentage Error, MAPE)、均方误差(Mean square Error, MSE)、均方根误差 (Root-Mean-square Error, RMSE)。这些评估系统的误差函数反应的是模型的误差而不会太多依赖系统输

入的样本数量。在本次课题的实验中评估指标选择 RMSE，其在 MSE 平方函数基础上加入了开方计算，避免了数据异常值引起的不稳定；在标签变量数据量纲较大的时候选择 MAPE，因为其对相对误差敏感，不会因为量纲大而影响全局效果，MAPE 是百分比值，一般认为 MAPE 小于 10% 时，表示预测精度较高。

2.3 自编码器网络简介

AE 通过学习输入数据的高效表示（即编码）来重构其自身的输出，其目的是让输出尽可能接近输入。在这个过程中，AE 能够捕捉到数据中的重要特征，并在此基础上进行数据的重构或其他任务。

2.3.1 自编码器基本知识

AE 算法属于无监督学习或是半监督范畴，在监督学习中神经网络的功能如式(2.6)所示。

$$\bar{x} = f(x); x: \mathbf{R}^{d_{in}}; \bar{x} = \mathbf{R}^{d_{out}} \quad (2.6)$$

d_{in} 是输入的特征向量长度， d_{out} 是网络输出的向量长度。网络模型通过把长度为 d_{in} 输入特征向量 x 变换到长度为 d_{out} 输出向量 \bar{x} ，这个过程可以看成是特征降维的过程，把原始的高维输入向量 x 变换到低维的变量 \bar{x} 。其原理图如图(2.4)所示。

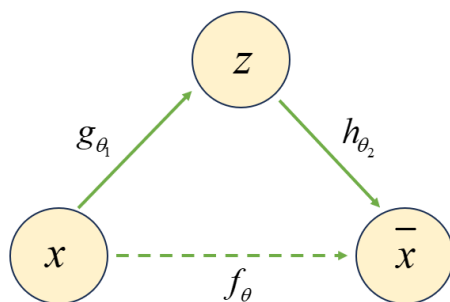


图 2.4 自编码器原理图

Fig.2.4 Schematic diagram of autoencoder

AE 将网络 f_{θ} 分为两个部分：编码器和解码器，编码器部分负责将高维度的输入数据 x 编码为低维度的隐变量，这一过程尝试学习映射关系： $g_{\theta_1}: x \rightarrow z$ ，以实现数据压缩表示。而解码器部分则负责将这些低维度的隐变量解码回原始的高维度数据 \bar{x} ，通过学习映射关系 $h_{\theta_2}: z \rightarrow \bar{x}$ 实现。

输入数据的编码和解码过程由编码器和解码器共同完成，这样的网络模型被称为 AE。当这个 AE 使用深层神经网络来参数化编码器和解码器时，它就被称为深度自编码器，其结构如图 2.5 所示。

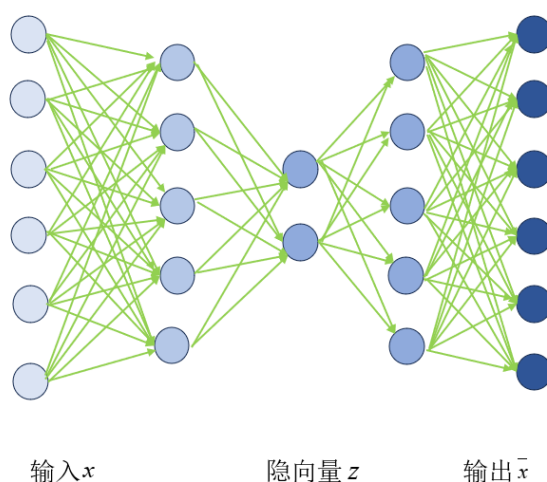


图 2.5 AE 结构图

Fig.2.5 Autoencoder structure diagram

AE 能够将输入变换到隐藏向量 z ，并通过解码器重建出 \hat{x} 。解码器的输出能够完美地或者近似恢复出原来的输入，即 $x \approx \hat{x}$ ，AE 的优化目标：

$$\text{Minimizer} = \text{dist}(x, \hat{x}), \hat{x} = h_{\theta_2}(g_{\theta_1}(x)) \quad (2.7)$$

$\text{dist}(x, \hat{x})$ 表示 x 和 \hat{x} 的距离度量，称为重建误差函数。常见的度量方法有欧氏距离的平方，计算方法如下：

$$\Phi = \sum (x_i - \hat{x}_i)^2 \quad (2.8)$$

AE 与传统神经网络在本质上并无差异，其区别仅在于训练过程中的监督信号从外部标签变为了 y 数据本身 x 。通过利用深层神经网络的强大非线性特征提取功能，AE 能够实现更优秀的数据表征。与线性技术如主成分分析法相比，AE 的表现更佳，能够更完美地重构输入数据。从原理上讲，它与均方误差是等同的。

2.3.2 自编码器基本分类

AE^[43] 是一类通过无监督学习来学习数据表示的神经网络。根据它们的设计和目的，AE 可以被分类为多种类型。每种类型的 AE 都旨在解决特定的问题或优化特定的任务，如特征提取、数据降维、去噪等。以下是一些主要的 AE 分类及其详细说明：

标准自编码器(Standard Autoencoders, SAE)^[44] 由一个编码器和一个解码器组成，旨在学习输入数据的压缩表示，然后尽可能准确地重构原始数据。这种类型的 AE 主要用于数据降维和特征学习。

去噪自编码器(Denoising Autoencoders, DAE)^[45] 的目标是从损坏的输入数据中恢复出干净的数据。在训练过程中，输入数据首先被人为添加噪声，然后 DAE 学习如何恢复原始的、未损坏的数据。这种类型的 AE 用于去噪和数据恢复。

稀疏自编码器(Sparse Autoencoders, SPAE)^[46]通过在损失函数中添加一个稀疏性惩罚项来鼓励模型学习更稀疏的数据表示。这种方法有助于提高模型的泛化能力,并且可以用于特征选择。

变分自编码器(Variational Autoencoders, VAE)^[47]是 AE 的一种扩展,它们通过引入概率分布来生成编码,从而使模型能够生成新的、类似于训练数据的数据点。VAE 在生成模型和数据生成任务中特别有用。

卷积自编码器(Convolutional Autoencoders, CAE)^[48]利用卷积层来构建编码器和解码器,使得它们特别适合处理图像数据。这种类型的 AE 可以有效捕捉图像数据的空间层次结构。

循环自编码器(Recurrent Autoencoders)^[49]使用循环神经网络(如 LSTM 或 GRU)来构建编码器和解码器,非常适合处理序列数据(如时间序列数据或文本)。这种 AE 可以捕捉数据中的时间依赖性。

序列到序列自编码器(Sequence-to-Sequence Autoencoders)^[50]是 RAE 的一种特殊形式,设计用于将输入序列转换为相同或不同长度的输出序列。它们通常用于机器翻译、文本摘要等任务。

对抗自编码器(Adversarial Autoencoders, AAE)^[51]结合了 AE 和生成对抗网络(Generative Adversarial Network, GAN)的特点。通过训练一个额外的判别器网络来区分潜在空间中的样本是否符合预期分布,AAE 旨在学习更加平滑且连续的数据表示。

2.4 贝叶斯优化算法简介

BO 作为高效的非参数核全局优化方法,在实际应用中具有广泛的用途,它已被用来优化机器人的控制策略和行为、优化系统的性能 and 设计参数、寻找最优策略以及优化各种参数,如超参数^[52]、神经网络的结构^[53]等。

2.4.1 贝叶斯优化算法基本知识

BO 是一种高效的机器学习策略,属于主动学习算法的一种,旨在通过迭代过程中的主动探索来寻找目标函数的最优解。该算法通常采用 GP 对目标函数进行建模,并使用采集函数,如改进概率、预期改进或置信上限等函数,这些函数的目的是在探索和利用之间寻求平衡。通过这种方式,贝叶斯优化能够在设计空间中有效地选择下一数据样本进行采样,进而在较少的迭代次数内逼近目标函数的全局最优解,特别适用于那些计算代价高昂或数据样本稀缺的优化问题。

其核心思想基于贝叶斯统计学,这是一种通过观测数据更新关于不确定性的知识的方法。这种思想被应用于对目标函数的探索和利用过程。通过采用 GP 作为目标函

数的先验分布，贝叶斯优化能够基于已观察到的数据点，更新其对目标函数全局行为的认知，即所谓的后验分布。高斯过程能够为任何一点提供一个预测的均值和方差，均值表示对目标函数的最佳估计，而方差则代表预测的不确定性。

(1) 高斯过程与后验更新：高斯过程是一种定义在连续域上的随机过程，其中任何有限集合的点的联合分布都是高斯的。这使得 GP 成为描述目标函数不确定性的理想工具，因为它既可以估计函数值（通过均值），也可以量化估计的不确定性（通过标准差）。随着新数据点的加入，GP 模型可以被更新，从而更准确地反映目标函数的行为。这一特性是贝叶斯优化能够有效地指导搜索过程的关键。

(2) 获取函数：获取函数在贝叶斯优化中起到至关重要的作用。它基于 GP 模型的预测和不确定性来评估在参数空间中各个点进行采样的潜在价值。通过优化获取函数，算法能够识别出那些既有望提升目标函数值，又或者能显著减少预测不确定性的点。常见的获取函数包括期望改进、概率改进和知识梯度。这些获取函数通过不同的策略平衡探索（寻找新区域以减少不确定性）和利用（在已知信息基础上优化目标函数）之间的关系。

(3) 贝叶斯优化的迭代过程：贝叶斯优化通过一个迭代的过程寻找目标函数的最优解。每一次迭代包含以下步骤：首先，基于当前 GP 模型，计算获取函数，并确定下一个评估点；其次，对选择的点进行目标函数评估，并将新获得的数据点加入到数据集中；最后，使用所有现有数据点更新 GP 模型。这一过程反复执行，直到达到预定的迭代次数或其他终止条件。

2.4.2 贝叶斯优化算法基本分类

BO 是一种可以用来优化任何没有显式公式的函数的黑盒优化算法，包括复杂的非凸函数和高噪声函数。BO 通常包含以下几个组成部分：先验模型、采样策略、目标函数评估和后验模型更新。根据采样策略的不同，BO 可以分为以下几种类型：

BO 的高斯过程回归(Gaussian process regression tuned by Bayesian optimization, GP-BO)^[54]中高斯过程（GP）作为一种概率模型，被应用于对目标函数的建模以及预测接下来最优点的位置，这是因为 GP 能够估计任何连续函数的不确定性。GP-BO 的主要优点包括模型的简洁性和易于理解的特性，以及能够依据先验知识来调整参数。

树形结构贝叶斯优化（Tree-structured Parzen Estimator Bayesian Optimization, TPE-BO）^[55]：TPE-BO 是一种基于树形结构 Parzen 估计器的 BO。TPE-BO 通过使用两个概率密度函数来对目标函数进行建模，分别对应于“好”的样本和“坏”的样本。这种方法允许 TPE-BO 定位下一最优点的策略是通过对比两个关键的密度函数。它的显著优点包括算法流程的简明性和执行的高效性，加之能够依据先验知识调节参数，增强了其灵活性和应用广度。借助于这种机制，TPE-BO 通过对设计空间进行反复的迭

代探索，能够高效解决目标优化问题。这种方法不仅优化过程简便直接，而且通过引入先验知识，为参数调整提供了理论依据，从而在复杂的优化任务中实现了更为精确和有效的搜索

蒙特卡罗树搜索贝叶斯优化（Monte Carlo Tree Search Bayesian Optimization, MCTS-BO）^[56]：MCTS-BO 是一种将 BO 与蒙特卡罗树搜索算法相结合的方法。MCTS-BO 结合了蒙特卡罗树搜索算法和 BO，其中 BO 负责选择下一个潜在的最优点，而蒙特卡罗树搜索则负责寻找最优解。这种结合的方法之所以有优势，是因为它能够克服 BO 在处理高维度问题时的限制，使 BO 能够适用于解决高维数据样本的问题。

2.5 HFSS 电磁仿真软件及其调用方法

电磁仿真计算的主要方法包括有限元法、矩量法和时域有限差分法，而 HFSS 软件是在进行复杂的天线和微波器件仿真计算时广泛使用的一种计算机软件。它能够为天线设计提供高精度的回波损耗参数、增益，以及为功分器设计提供插入损耗和隔离度参数，因此在微波天线、滤波器和功分器等电磁器件的设计领域中得到了广泛应用。为了验证本课题的实验，将采用 HFSS 软件结合 MATLAB 进行联合电磁仿真。

本文中部分实验将利用部分正交实验^[57]来得到实验数据，部分正交实验是研究多水平、多因素变量时为保证数据选取的科学性，同时利用尽量少的数据来表征全部特征的一种有效方法，依据正交性，可以挑选出具有代表性、均匀分散的点作为实验数据。本课题利用 MATLAB 联合仿真 HFSS，通过 MATLAB-HFSS-API 脚本结合部分正交实验得到训练和测试样本数据。

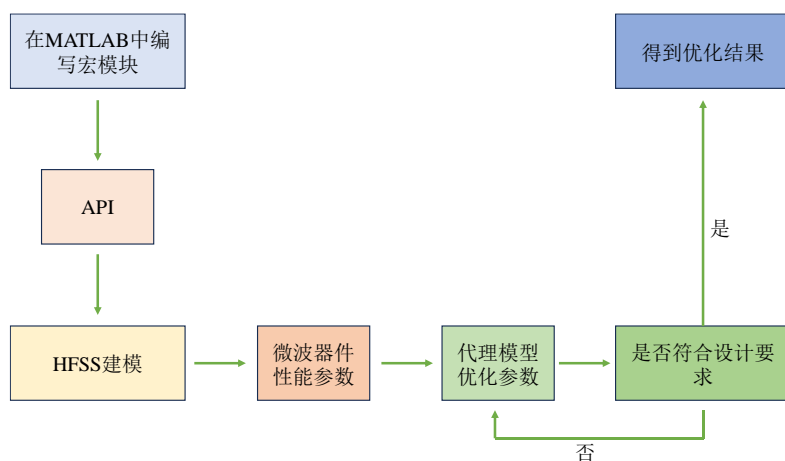


图 2.6 MATLAB 联合仿真 HFSS 流程

Fig.2.6 Diagram of MATLAB co-simulation HFSS

通过 HFSS 软件与 MATLAB 的联合电磁仿真，可以很方便的在脚本中编写天线或滤波器的尺寸信息来控制 HFSS 软件实现机器主动绘画天线的功能。通过在 MATLAB

只中的 `HFSS_Execute_Script` 接口函数中编写控制 HFSS 软件的仿真进行 3D 建模^[58]的程序，只需在 HFSS 中设置相应的扫频范围与扫频间隔，就可以来获得与微波器件优化设计指标相关的信息。在 `Script` 脚本文件中，依据相关设置可得到的输出结果如回波损耗曲线、方向图、插入损耗或者相应的谐振频率等，输出的结果可供建模分析微波器件的性能，并以此来判断微波器件是否符合优化设计目标。

联合仿真的过程如图 2.6 所示。利用 MATLAB 写宏命令，之后使用 `fopen` 生成 HFSS 可读脚本文件，包括构造天线模型、中心频率、扫频、设置边界条件和激励、设置优化变量参数。使用 `HFSS_Execute_Script` 自动调用 HFSS 仿真窗口，使用 `HFSS_Export_Network_Data` 导出结果。

第3章 基于图模型卷积神经网络在微带天线谐振频率建模

3.1 引言

在传统的机器学习领域，虽然在多个领域内展现出了它们的强大能力和灵活性，但在特定的应用场景，特别是涉及到复杂图像处理任务时，这些方法可能无法充分发挥潜力。在天线工程领域，传统的机器学习方法处理天线的尺寸参数和其它复杂数据时，往往无法直接应用或需要复杂的预处理步骤，这限制了它们的效率和效果。

针对这一问题，本章讨论了一种创新的方法，即通过将微带天线的尺寸参数进行图模型化处理，将这些参数转换成一个由 1 和 0 组成的二维（2-D）矩阵。这一转换过程类似于将天线参数编码成一张黑白图像，其中 1 和 0 代表不同的参数状态。这种方法不仅使得天线的尺寸参数能够被 CNN 以图像的形式处理，从而充分发挥 CNN 在图像处理方面的优势，同时也为天线设计和分析提供了一种全新的视角和工具。

在实施这一方法之后，本研究进一步将通过这种图模型化处理后的数据输入到 CNN 中，以进行模型训练和验证。为了评估这种方法的有效性，将 CNN 模型的预测性能与那些使用传统机器学习方法得到的结果进行了比较。这种比较不仅涵盖了模型的准确性，还包括了模型在处理复杂数据时的灵活性和效率。

通过这项研究，期望证明，通过将天线参数的图模型化处理与 CNN 的强大图像处理能力相结合，不仅可以提高模型对天线性能预测的准确性，还能够扩展 CNN 在非传统图像处理领域的应用范围，为天线设计和分析带来新的方法论和工具。这种方法的成功应用，将有助于推动天线工程领域以及其他需要处理复杂参数数据的领域的技术进步和创新。

3.2 图模型卷积神经网络系统

3.2.1 图模型卷积神经网络基本原理

图模型就是将待建模天线转换为二进制黑白图模型的形式，具体来讲，就是对上面提到的一维 CNN 的输入 x 进行转置操作，即将 x^T 作为二维 CNN 的输入，并令 x_i 以非 0 即 1 的二进制串的形式表示，则其取值表示为：

$$x_i = (x_i)_{\min} + \frac{\sum_{k=0}^{m-1} \alpha_k 2^k}{2^m - 1} \left((x_i)_{\max} - (x_i)_{\min} \right) \quad (3.1)$$

式中， $(x_i)_{\min}$ 代表 x_i 取值范围内的最小值， $(x_i)_{\max}$ 代表 x_i 取值范围内的最大值， m 为

二进制串的长度, a_k 为 1 或者 0, 代表该位置存在与否。

上式中如果 $\sum_{k=0}^{m-1} a_k 2^k$ 看成一个整体 A, 那么可以得出:

$$A = \frac{(x_i) - (x_i)_{\min}}{((x_i)_{\max} - (x_i)_{\min})} \times (2^m - 1) \quad (3.2)$$

式中 x_i 、 $(x_i)_{\min}$ 、 $(x_i)_{\max}$ 、 m 表示的含义与式(3.1)相同, 由此可得 A 是一个具体的十进制数值, 在程序编码过程中将 A 转为二进制数值。具体过程如图 3.1 所示。

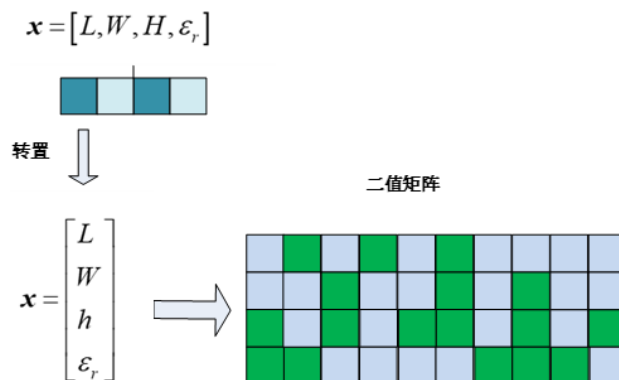


图 3.1 天线参数二维建模

Fig.3.1 Two-dimensional modeling of antenna parameters

这个图像在二维 CNN 中被视为输入, 其中每个像素的值可以是基于原始信号数据的某种表示, IMCNN 结构如图 3.2 所示。

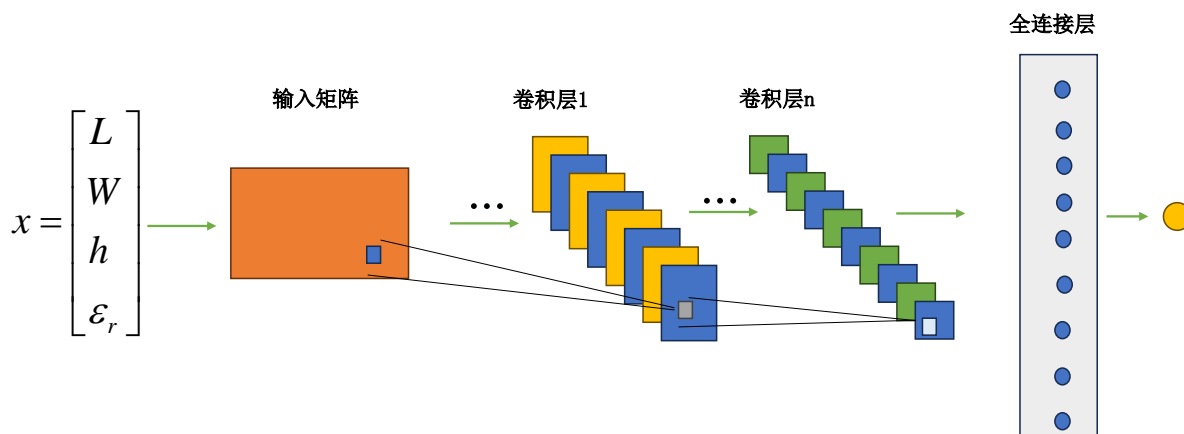


图 3.2 IMCNN 示意图

Fig.3.2 Schematic diagram of IMCNN

下面是对图模型 CNN 模型 (Image-based CNN, IMCNN) 超参数的简要介绍。

(1) 激活函数: 在神经网络中使用激活函数引入非线性变换以增强解决复杂问题的能力^[59]。对于 IMCNN, 激活函数使用 Relu 函数。

(2) 损失函数: 损失函数保证了 IMCNN 在训练数据集上的最佳性能, 需要最小化。回归损失主要包括均方误差(MSE)、平均绝对误差(MAE)和平均偏差误差(MBE)。

(3) 特征优化器：特征优化器用于为损失函数找到最优值。自适应矩估计(Adaptive Moment Estimation)通过计算梯度的一阶矩估计和二阶矩估计而为不同的参数设计独立的自适应性学习率^[60]。随机梯度下降法(Stochastic Gradient Descent, SGD)是一种基于梯度的优化算法^[61]，可以帮助找到损失函数最小的参数。RMSProp 法使用了超参来平滑动量项和梯度平方，在目标函数的最优解更为平缓的方向，会取得更大的进步，能够减小迭代过程中梯度摆幅过大的问题^[62]。

(4) Epoch 的数量和批处理大小：Epoch 是指整个数据集通过神经网络向前和向后传递一次。如果整个数据集不能一次传递到算法中，则必须将其分成几个小批。批次大小(Batch size)是指单个批次中存在的训练样本的总数。

(5) L2 正则化系数：为了防止过拟合，损失函数中添加一个正则化项来实现。这个正则化项是模型权重的 L2 范数的一个函数。L2 正则化有时也被称为岭回归(Ridge Regression)或权重衰减。L2 正则化系数，通常用符号 λ 表示，是一个超参数，用于控制正则化项的影响强度。在损失函数中加入 L2 正则化项后，新的损失函数 L 可表示为：

$$L = L_0 + \lambda \sum_{i=1}^n w_i^2 \quad (3.3)$$

(6) Dropout：在每次训练迭代中，Dropout 会随机选择一部分神经元（以及其对应的连接），并将它们的输出设置为零，选择的比例由 Dropout 率(通常表示为 p)来决定，表示每个神经元在每次训练迭代中被丢弃的概率，通过丢弃一部分神经元，减少模型对于训练样本的依赖，迫使模型学习到更加鲁棒的特征表示。这相当于每次迭代训练时都在训练一个“缩小”的网络，增加了网络结构的多样性，减少了复杂模型的过拟合问题。

3.2.2 图模型卷积神经网络基本流程

天线参数的图模型处理，以及将模型输入映射到一个由 1 和 0 组成的二维(2-D)矩阵的方法，实质上是一种将传统的天线设计参数转化为可由 CNN 处理的图像形式的技术。这种方法涉及到几个关键步骤，下面详细介绍这一过程。

1. 天线参数的选择和准备

首先，需要确定哪些天线参数是对性能预测最为关键的。这些参数可能包括天线的长度、宽度、高度、馈电点位置、基板材料的电磁属性等。一旦选择了相关参数，接下来的任务是准备这些参数的数据，以便进行后续的图模型化处理。

2. 参数的二维化表示

将选择的天线参数转换为二维矩阵的形式，是图模型化处理的核心步骤。这一转换过程涉及到将每个参数的值映射到一个二维平面上，类似于创建一张黑白图片。在这张图片中，参数的不同值将被编码为 1(表示黑色)和 0(表示白色)组成的向量。将每

一个输入都转为向量再合并便是图模型矩阵了。

3.矩阵的创建和优化

创建这样的矩阵可能需要一些预处理步骤，如归一化参数值、确定矩阵的分辨率(即矩阵的大小)，以及可能的参数值映射策略。此外，还需要考虑如何以最有效的方式表示天线的整体结构和特性，以便 CNN 可以从这些二维图像中学习到的最有用的特征。

4.利用 CNN 处理映射后的数据

一旦天线参数被成功映射到了由 1 和 0 组成的 2-D 矩阵，就可以将这些矩阵作为输入提供给 CNN。CNN 通过其卷积层、池化层和全连接层来处理这些输入，自动提取和学习天线参数与性能之间的复杂关系。这种处理方式利用了 CNN 在图像识别和特征提取方面的强大能力，为天线设计和性能预测提供了一种新的方法。

5.模型训练和验证

最后，通过训练数据集来训练 CNN 模型，并使用验证数据集来评估模型的性能。这不仅包括模型预测的准确性，还包括其泛化能力，即模型对未见过的数据的处理能力。

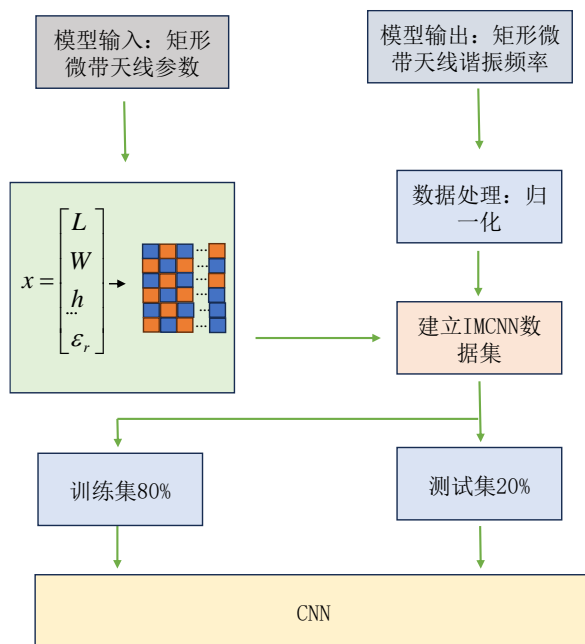


图 3.3 IMCNN 模型具体流程

Fig.3.3 Specific process of IMCNN model

3.3 图模型卷积神经网络系统对矩形微带天线谐振频率预测

3.3.1 矩形微带天线基本介绍

为了验证 IMCNN 模型的预测能力，本小节使用矩形微带天线（Rectangular Microstrip Antenna, RMSA）进行实验。矩形微带天线主要有三个组成部分，辐射源、

介质层和参考地，其结构示意图如图 3.4 所示，其中， L 和 W 表示 RMSA 俯视图的长度和宽度， ε_r 表示相对介电常数，介质基板的厚度用 h 表示，根据 Mehmet Kara 测得的 33 组数据^[63]作为数据集，具体内容为矩形微带天线不同尺寸数据在 TM_{10} 模式下的谐振频率数据在表 3.1 展示。

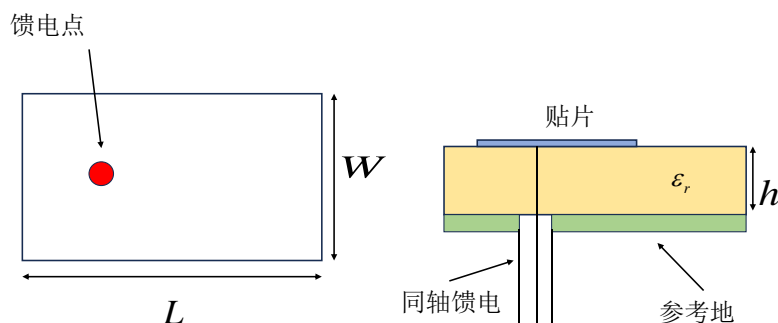


图 3.4 矩形微带天线结构示意图

Fig.3.4 Structure diagram of rectangular microstrip antenna

3.3.2 矩形微带天线谐振频率预测

谐振频率预测实验具体流程方式为：

(1) 确定 IMCNN 的输入与输出：以 RMSA 的 W, L, h 和相对介电常数 ε_r 为因变量，以其对应的在 TM_{10} 模式下的谐振频率 f_{meas} 为输出标签。输入记作 $\mathbf{X}=[W, L, h, \varepsilon_r]$ ，输出记作 $Y=[f_{meas}]$ 。为了确保实验的科学性，使用 MATLAB 中的 randperm 函数将 33 组数据随机分开，也即使用 27 组数据用作训练集，剩下 6 组用作测试集进行实验，输出标签数据则为样本数据对应的谐振频率值。

(2) IMCNN 模型具体参数设置：根据本实验的具体情况和一定的先验知识，在训练集上进行了一定数量的计算机实验，并探索了以下参数范围：Epoch(1-800)和 Batch size(1-10)。实验还测试了 IMCNN 的结构配置。超参数为功能优化器:SGD, Adam 和 RMSProp；网络结构：三层卷积层，卷积核分别为 3×5 , 3×5 和 2×2 ，两层全连接层；批大小为 3，最大 Epoch 数为 200，Dropout 分别取 0.01, 0.05, 0.1, 0.2, 0.3, 0.4；Padding 为 1。

(3) 训练模型：把输入、输出数据输入到 IMCNN 进行预测。因为输入尺寸维度较小，使用图模型将 1×4 的样本数据经过图模型转换后变成 $4 \times 10 \times 1$ 。经过三层卷积层，两层全连接层，模型参数如表 3.2 所示。

(4) 预测 RMSA 的谐振频率：IMCNN 本质上是一个回归模型，那么经过 IMCNN 模型输入后，RMSA 谐振频率输出只有一个谐振频率输出值。

实验结果：为了探究 IMCNN 模型的综合性能，将 IMCNN 与经典 BP、GP 这 2 种模型实验结果进行了比较。使用经典模型的输出和 IMCNN 模型的输出如表 3.3 所示。

从表 3.3 可以看出, IMCNN 模型的预测精度和拟合优度总体上优于经典方法。

表 3.1 RMSA 在 TM_{10} 模式下的谐振频率点

Table 3.1 Resonant frequency points in TM_{10} mode of RMSA

Number	W (cm)	L (cm)	h (cm)	ε_r	$f_{meas}/(\text{MHz})$
1	0.850	1.290	0.017	2.22	7740
2	0.790	1.185	0.017	2.22	8450
3	2.000	2.500	0.079	2.22	3970
4	1.063	1.183	0.079	2.25	7730
5	0.910	1.000	0.127	10.20	4600
6	1.720	1.860	0.157	2.33	5060
7	1.810	1.960	0.157	2.33	4805
8	1.270	1.350	0.163	2.55	6560
9	1.500	1.621	0.163	2.55	5600
10	1.337	1.412	0.200	2.55	6200
11	1.120	1.200	0.242	2.55	7050
12	1.403	1.485	0.252	2.55	5800
13	1.530	1.630	0.300	2.50	5270
14	0.905	1.018	0.300	2.50	7990
15	1.170	1.280	0.300	2.50	6570
18	0.790	1.255	0.400	2.55	7134
19	0.987	1.450	0.450	2.55	6070
20	1.000	1.520	0.476	2.55	5820
21	0.814	1.440	0.476	2.55	6380
22	0.790	1.620	0.550	2.55	5990
23	1.200	1.970	0.626	2.55	4660
24	0.783	2.300	0.854	2.55	4600
25	1.256	2.756	0.952	2.55	3580
26	0.974	2.620	0.952	2.55	3980
27	1.020	2.640	0.952	2.55	3900
28	0.883	2.676	1.000	2.55	3980
29	0.777	2.835	1.100	2.55	3900
30	0.920	3.130	1.200	2.55	3470
31	1.030	3.380	1.281	2.55	3200
32	1.265	3.500	1.281	2.55	2980
33	1.080	3.400	1.281	2.55	3150

本次实验以 RMSE 和 APE 来衡量回归输出值与现实目标值之间的偏差。在实验

中 IMCNN 模型设置了三种 CNN 的优化器，Dropout 的值设置为从 0.01 到 0.4。表 3.4 统计了 IMCNN 在给定环境中的回归性能。图 3.5 显示了预测样本的真实值和预测结果对比，拟合程度很好。

表 3.2 IMCNN 模型参数
Table 3.2 IMCNN model parameters

	输入大小	输出大小	权值矩阵大小
卷积层 1	4×10×1	4×8×2	3×5×2
卷积层 2	4×8×1	4×6×4	3×5×4
卷积层 3	4×6×4	3×5×8	2×2×8
全连接层 1	3×5×8	1×1×100	100×120
全连接层 2	1×1×100	1×1×20	20×100
输出层	1×1×20	1	1×20

表 3.3 不同方法建模的预测结果
Table 3.3 Prediction results of different modeling methods

f/MHz	BP	GP	IMCNN
3200	3178	3196.581	3149.43
3580	3644.6	3582.141	3580.81
4805	4703.3	4830.272	4789.17
5100	5291.4	5145.394	5092.45
6200	6147.2	6198.311	6162.057
8450	8233.1	8251.843	8504.42
APE	1.939	0.6400	0.4821

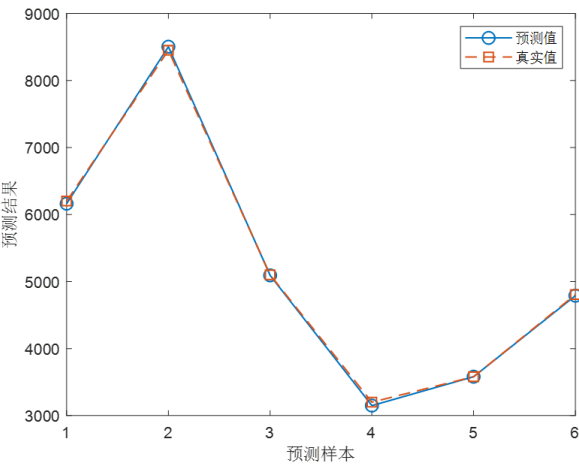


图 3.5 测试集预测结果

Fig.3.5 Test set prediction results

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - Y_i)^2} \quad (3.4)$$

$$APE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - Y_i|}{|y_i|} * 100 \quad (3.5)$$

表 3.4 IMCNN 回归 RMSE 和 APE

Table 3.4 IMCNN regression RMSE and APE

Optimizer	dropout	Mean_RMSE	Mean_APE
sgdm	0.01	107.1638	1.2109
adam	0.01	57.9005	0.63047
rmsprop	0.01	109.6913	1.3618
sgdm	0.05	57.5775	0.7728
adam	0.05	86.2207	0.9865
rmsprop	0.05	112.6033	1.242
sgdm	0.1	54.4365	0.48210
adam	0.1	121.5955	1.2851
rmsprop	0.1	97.1525	1.3187
sgdm	0.2	74.8965	0.6820
adam	0.2	85.7685	1.2163
rmsprop	0.2	100.0663	1.3297
sgdm	0.3	121.9117	1.4585
adam	0.3	135.7668	1.7366
rmsprop	0.3	172.4042	2.1876
sgdm	0.4	220.0528	2.8103
adam	0.4	182.524	2.1548
rmsprop	0.4	139.9485	1.5058

值得注意的是当优化方法为 SGD，Dropout 为 0.1 时 RMSE 为 54.4365，APE 为 0.48210，由此可见模型的回归拟合表现优秀。图 3.5 显示了预测样本的真实值和预测结果对比，拟合程度很好。

3.4 图模型卷积神经网络系统对圆形微带天线谐振频率预测

3.4.1 圆形微带天线基本介绍

本小节的研究对象为圆形微带天线（Circular Microstrip Antenna, CMSA），其结构

与矩形微带天线有所不同,采用了圆形贴片代替了矩形贴片^{[64][65]}。圆形微带天线相对于矩形微带天线,在性能指标方面具有更好的天线阻抗匹配、方向图、效率和带宽等特性。CMSA 和 RMSA 的结构是相似的,将矩形贴片换成圆形贴片即可,其侧面结构图如图 3.6 所示。圆形微带天线在 TM_{11} 模式下的谐振频率如表 3.5 所示。

表 3.5 圆形微带天线在 TM_{11} 模式下的谐振频率数据Table 3.5 Resonant frequency of CMSA in TM_{11}

贴片序号	a (cm)	h (cm)	ϵ_r	$f_{meas}/(\text{MHz})$
1	6.800	0.08000	2.32	835
2	6.800	0.15900	2.32	829
3	6.800	0.31800	2.32	815
4	5.000	0.15900	2.32	1128
5	3.800	0.15240	2.49	1443
6	4.850	0.31800	2.52	1099
7	3.493	0.15880	2.50	1570
8	1.270	0.07940	2.59	4070
9	3.493	0.31750	2.50	1510
10	4.950	0.23500	4.55	825
11	3.975	0.23500	4.55	1030
12	2.990	0.23500	4.55	1360
13	2.000	0.23500	4.55	2003
14	1.040	0.23500	4.55	3750
15	0.770	0.23500	4.55	4945
16	1.150	0.15875	2.65	4425
17	1.070	0.15875	2.65	4723
18	0.960	0.15875	2.65	5524
19	0.740	0.15875	2.65	6634
20	0.820	0.15875	2.65	6074

3.4.2 圆形微带天线谐振频率预测

与 RMSA 相同,以圆形贴片的半径 a 、介质基板的厚度 h 和相对介电常数 ϵ_r 为系统的输入变量,根据国外学者 Singh^[66]和 Seref Sagiroglu^[67]测量的数据,在数据集的划分方面,选取了 14 组实验数据作为训练数据,将 6 组数据作为测试数据,输出为对应的谐振频率。通过对比不同模型在预测结果上的表现,可以评估不同模型的准确性和泛化能力,为天线优化设计提供参考。

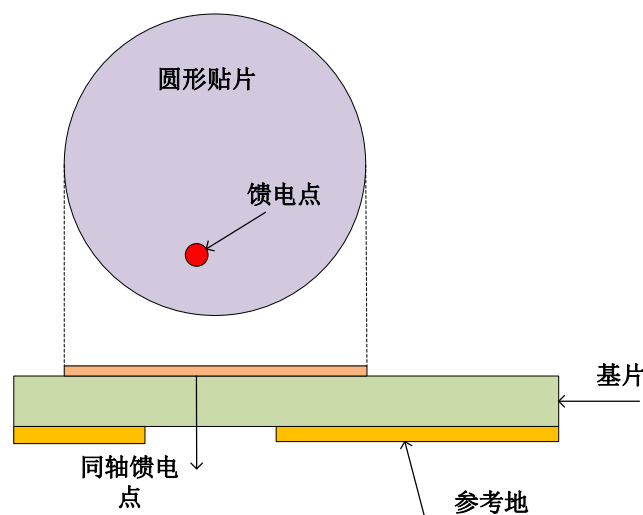


图 3.6 圆形微带天线的结构图
Fig.3.6 Structure diagram of CMSA

CMSA 的谐振频率预测实验具体流程如下。

(1) 确定 IMCNN 的输入与输出：以 CMSA 的 a, h 和相对介电常数 ϵ_r 为因变量，以其对应的在 TM_{11} 模式下的谐振频率 f_{meas} 为输出标签。输入记作 $\mathbf{X}=[a, h, \epsilon_r]$ ，输出记作 $Y=[f_{meas}]$ 。为了确保实验的科学性，使用 MATLAB 中的 `randperm` 函数将组数据随机分开。也即使用 14 组数据用作训练集，剩下 6 组用作测试集进行实验，输出标签数据则为样本数据对应的谐振频率值。

(2) 设定 IMCNN 模型具体参数设置：根据本实验的具体情况和一定的先验知识，在训练集上进行了一定数量的计算机实验，并确定以下超参数范围：功能优化器为 SGD；网络结构：三层卷积层，卷积核分别为 2×3 ， 2×3 和 2×2 ，两层全连接层；批大小为 3，最大 epoch 数为 200，Dropout 分别取 0.1；Padding 为 1。

(3) 训练模型：把输入、输出数据输入到 IMCNN 进行预测。因为输入尺寸维度较小，使用图模型将 1×3 的样本数据经过图模型转换后变成 $3 \times 10 \times 1$ 。经过三层卷积层，两层全连接层。

(4) 预测 RMSA 的谐振频率：IMCNN 本质上是一个回归模型，那么经过 IMCNN 模型输入后，CMSA 谐振频率输出只有一个谐振频率输出值，输出值如表 3.6 所示。

实验结果：本次研究采用了 RMSE 和 APE 这两个重要指标来量化回归模型输出值与实际目标值之间的差异，以此来评估回归模型的性能。RMSE 衡量的是预测值与实际值差异的平方的平均值的平方根，它提供了预测误差的绝对量度；而 APE 关注的是预测值偏离实际值的百分比，更多地反映了相对误差的大小。通过这两个指标，可以全面了解模型预测性能的优劣。由表 3.6 表示 IMCNN 实验 6 组测试集数据的预测结果 f_{pre} 与其测试结果 f_{meas} 的对比，实验的 RMSE 为 70.29，MAPE 结果为 4.37%，证明 IMCNN 有较好的回归能力。

表 3.6 圆形贴片天线的预测结果
Table 3.6 Predicted results of CMSA

$a(\text{cm})$	$h(\text{cm})$	ε_r	f_{meas}	f_{pre}
0.820	0.15875	2.65	6074	6165
0.770	0.23500	2.55	4945	4892
1.270	0.07940	2.59	4070	4163
6.800	0.31800	2.32	815	886
3.975	0.23500	4.55	1030	970
0.960	0.15875	2.65	5524	5486

3.5 本章小结

本小节提出一种全新的图模型概念，将传统的一维数据转换为更加复杂的二维数据结构，为数据分析和特征提取打开了新的视角。这种转换不仅增加了数据的维度，而且更重要的是，它使得数据之间的关联和内在结构能被更好地表示和利用。在这一基础上，利用 CNN 的强大特征提取能力，对二维图模型进行有效的卷积运算，从而提取关键特征，并构建了 IMCNN 回归模型，该模型的有效性通过在微带天线谐振频率的预测上进行验证。与其他建模方法相比，IMCNN 模型展现出了显著的性能优势。这一结果不仅证实了图模型结合 CNN 在数据处理上的高效性，同时也为未来的天线设计和优化提供了一个强有力的工具。

第4章 基于图模型卷积神经网络自编码器的微带天线优化设计

4.1 引言

在微波器件优化过程中,一般缺少大型数据集,这会导致一系列问题,包括过拟合、泛化能力差和模型评估不准确等。为了更好地应用数据,解决实际问题,避免数据资源浪费,数据增强技术提供了一种解决方案。

数据增强通过生成额外的训练样本来扩大数据集的规模,有效减少过拟合的风险,从而提升模型的泛化能力。这是通过对现有数据应用一系列变化实现的,增加了数据的多样性,使模型能够学习到更广泛的特征表示。其次,数据增强可以显著降低获取大规模标注数据的成本和代价,这在很多实际应用中是一个重要的考量,特别是在标注代价高昂或难以获得的领域。此外,还增强了模型对输入数据小变动的鲁棒性,提高了模型在面对真实世界数据时的性能和可靠性。总之,数据增强不仅提高了模型的准确性和泛化能力,还通过有效利用有限数据资源,增加了数据集的规模,也增强了数据的多样性,有助于模型学习到更加泛化的特征,从而在模型性能上表现得更好。

通过在多样化的数据上训练,模型能够更准确地预测微波器件的性能,提高设计的质量和創新性。最终,这种方法不仅优化了设计流程,还提高了设计成果的实用性和创新水平,为微波器件的研发领域带来了新的可能性。

4.2.1 图模型卷积神经网络自编码器基本原理

本节研究的是图模型卷积神经网络自编码器(Image-base Convolutional Neural Network Autoencoder, IMAECNN)。此模型分为三个部分:图模型模块、AE 模块以及 CNN 模块,如图 4.1 所示,其中 AE 是一种无监督的神经网络,学习方式是以自身输入为标签,通过编码器提取输入数据高级特征,然后通过解码器输出极大恢复原始数据的近似解,在此过程中编码器输出的高级特征和解码器输出近似解都可作为训练的增强数据。

在 IMAECNN 模型中通过自编码器获取额外数据,CNN 输入一部分数据来自图模型,一部分来自编码器获得的额外数据,通过数据增强技术以便卷积神经网络在图片提取特征过程中有更好的表现,此做法旨在提升天线或其他微波器件优化设计的效率,提高模型的稳定性。

AE 的性能和效率受到其结构配置、激活函数、损失函数和训练算法的共同影响。在研究中,采用了 Logistic Sigmoid 函数作为激活函数,均方误差(Mean Squared Error, MSE)作为损失函数,并选择了缩放共轭梯度(Scaled Conjugate Gradient, SCG)算法进行

网络训练。

(1) 激活函数：选用 Logistic Sigmoid 函数作为编码器和解码器的激活函数。该函数将输入映射到(0,1)区间，有助于模型捕获非线性特征，并为网络的每一层输出提供平滑的梯度。

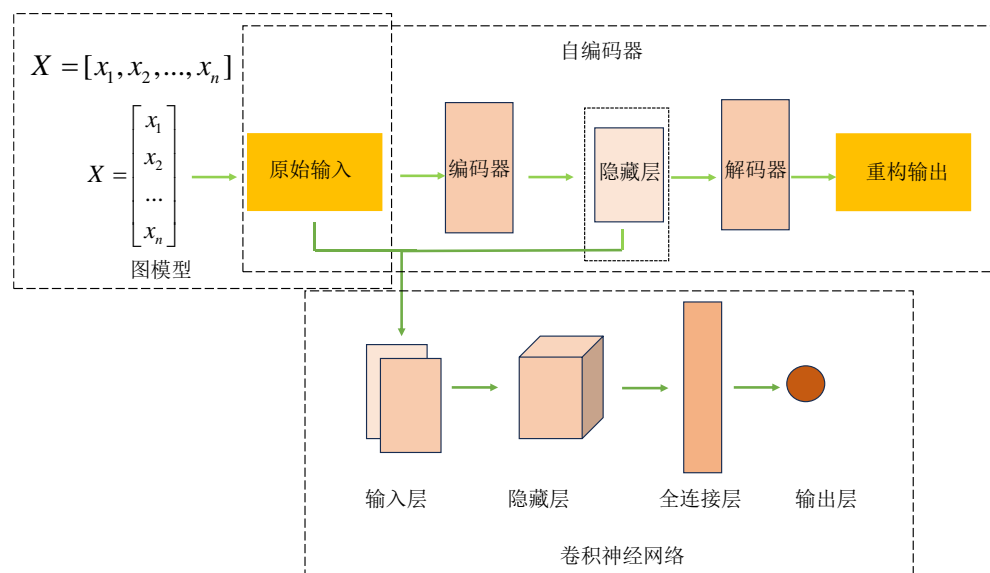


图 4.1 IMAECNN 模型图

Fig.4.1 IMAECNN model diagram

(2) 损失函数：采用 MSE 作为损失函数，以评估解码器输出与原始输入之间的误差。MSE 提供了一种量化自动编码器重建质量的直观方法，有助于指导网络训练的方向。

(3) 训练算法：使用缩放共轭梯度算法 SCG 算法训练自动编码器。SCG 是一种高效的优化方法，能够在不需要显式学习率设置的情况下快速收敛，特别适用于大规模数据集和复杂网络结构的训练。

Logistic Sigmoid 激活函数和 MSE 损失函数的组合，配合 SCG 训练算法，为自编码器的设计提供了一种有效的配置，不仅促进了网络的快速收敛，还确保了重建的准确性。在此次研究中，编码器与解码器的激活函数，损失函数以及训练算法均采用以上形式。

4.2.2 图模型卷积神经网络自编码器基本流程

本章研究的 IMAECNN 流程如图 4.2 所示。

(1) 数据准备阶段：在模型学习的初始阶段，把有标签数据进行图模型转换，这样的转换有助于捕捉数据的结构性质，提高学习的质量。与此同时，将图模型生成的数据使用 AE 神经网络学习提取高级特征。

(2) 数据融合阶段：在数据融合阶段，目的是整合不同来源的特征表示，形成一个全面的数据视图。自编码器隐藏层的高级特征是神经网络学习得出的，将图模型与提取的特征融合，构建一个新的特征矩阵。在这个过程中，改变矩阵的通道数，确保数据的不同维度被恰当地融合，这个过程对后续的 CNN 学习过程至关重要，因为它直接影响了模型能够学习到的特征类型和复杂性。

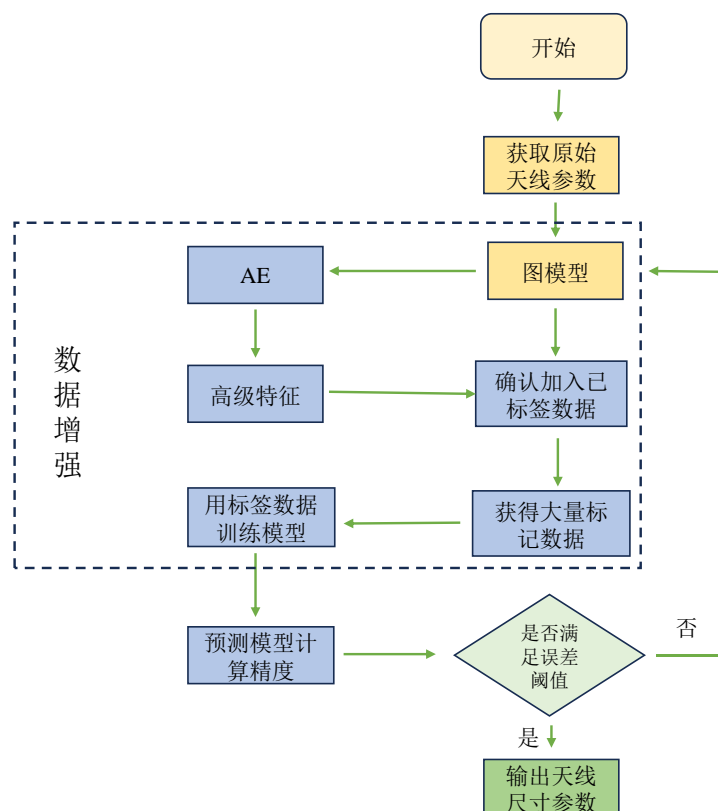


图 4.2 IMAECNN 流程图

Fig.4.2 IMAECNN flow chart

(3) 模型训练阶段：经过数据增强处理，融合后的矩阵被用作 CNN 的输入。CNN 具有强大的特征提取能力，能够从复杂的数据中自动学习到有用的信息。这种能力让 CNN 在面对数据增强后的融合矩阵时，可以有效地挖掘和识别出深层次的、有区分度的特征，进一步提升模型的性能。

(4) 反馈迭代阶段：在反馈迭代阶段，根据迭代条件来判断模型是否需要进一步的调整。如果模型的性能未达到预期的标准，或者还有进一步提升的空间，可能需要回到模型训练阶段调整参数。这一过程可能会循环进行，直到模型满足了设定的迭代条件。一旦满足条件，可以确认模型参数，并用它来预测例如天线尺寸等实际应用中的参数值。这个阶段是动态的，可能涉及到多轮的训练和验证，以确保模型的鲁棒性和准确性。

(5) 预测输出阶段：当模型训练完成后，测试数据集将被放入模型中进行预测。这一阶段是验证模型性能的重要环节，通过实际数据的预测输出，评估模型对于新数据

的泛化能力，同时也是模型调优的基础。

本此实验使用 RMSE 误差衡量标准来衡量系统训练结果是否满足预期效果，误差阈值设置为 7.5×10^{-2} ；也即系统迭代不断进行寻找，直到训练数据的预测误差低于 7.5×10^{-2} 后，返回保存系统结构数据并保存测试数据集。

4.3 图模型卷积神经网络自编码器对单频 PIFA 天线的优化设计

本小节将使用上面描述的 IMANCNN 对单频 PIFA 天线^[68](Planar Inverted-F Antenna, PIFA)进行优化设计，先根据获得的 PIFA 的相关回波损耗数据测试 IMANCNN 的回归能力；再利用系统输出最优尺寸，利用 HFSS 联合仿真，并输出其对应的回波损耗曲线；最后利用 IMAECNN 预测的 S_{11} 曲线与 HFSS 仿真的曲线进行拟合对比，验证该系统在天线优化设计方面的有效性。

4.3.1 单频 PIFA 天线基本介绍

根据 PIFA 的设计原则和经验公式^[69]，影响其性能的参数主要包括：贴片的长度 L_1 、宽度 W_1 以及基板 L_g 和 W_g 。本次实验选择这四个尺寸变量参数作为输入数据，并使用 IMAECNN 和 HFSS 联合仿真，通过对 PIFA 的这些尺寸参数和其对应的回波损耗扫频值进行建模来实现对其的优化设计，期望得到优化的天线具体尺寸参数与其对应的回波损耗曲线作为验证。本实验输入记作 $\mathbf{X}=[L_g, W_g, L_1, W_1]$ ，其中各变量的范围可根据表格 4.1 查看；其剩余尺寸参数设置为常量，见文献[69]，设置相对介电常数 ϵ_r 为 2.55，实验中根据四个天线参数的范围分别设置输入数据共有 130 组，获取数据流程如图 4.5 所示。

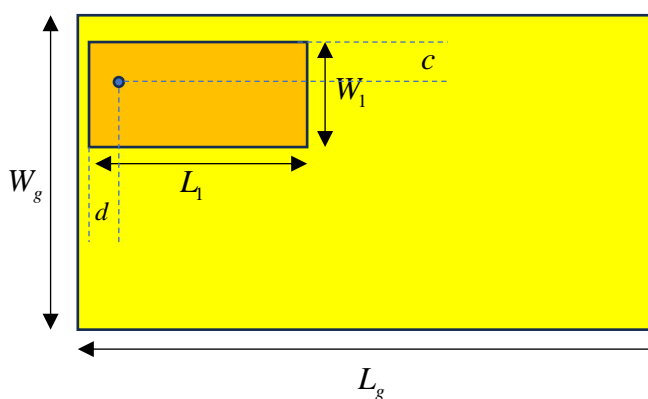


图 4.3 PIFA 的俯视结构图

Fig.4.3 Overhead structural diagram of PIFA

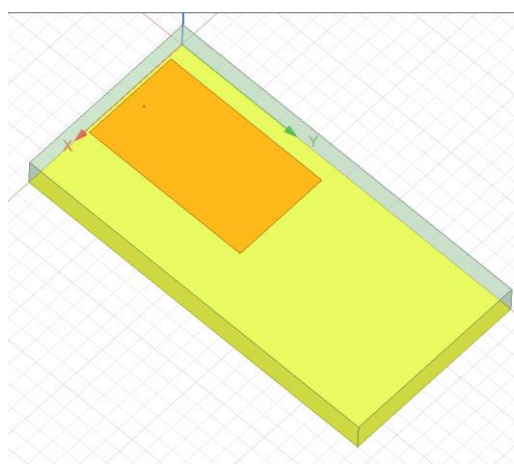


图 4.4 PIFA HFSS 建模图

Fig.4.4 HFSS modeling diagram of PIFA

4.3.2 单频 PIFA 天线优化设计

考虑到 PIFA 的设计指标是当天线工作在中心频率 0.85GHz 时，此刻天线回波损耗小于-10dB。所以在设置输出的时候，把扫频范围设置为 0.75GHz~1.0GHz，扫频步长为 0.01GHz，共有 26 个频率输出，即将输出标签数据记作 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{26}}]$ 。

表 4.1 PIFA 的相关变量及范围

Table 4.1 Variables and Ranges of PIFA

变量名	变量值范围/mm
L_g	118~122
W_g	58~62
L_1	58~62
W_1	30~34

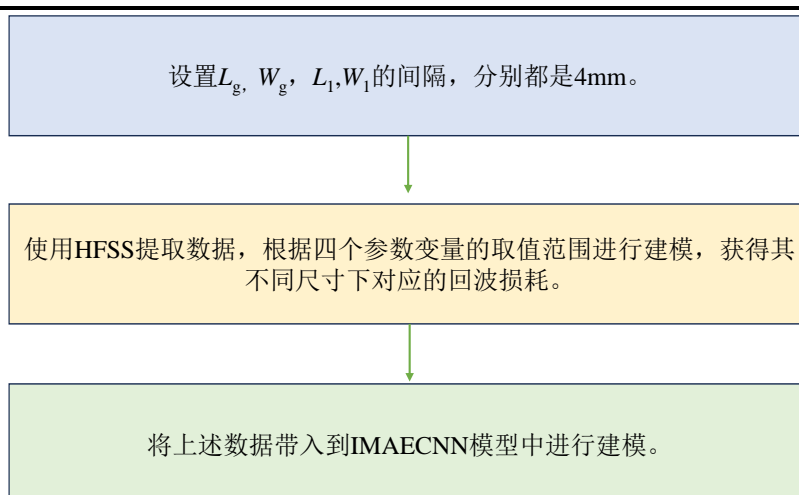


图 4.5 PIFA 数据获取流程

Fig.4.5 PIFA data acquisition process

本次实验设计为先利用数据验证 IMAECNN 系统的回归能力，之后利用系统寻优后的尺寸数据 PIFA 优化设计，具体流程如下所示。

(1) 确定 IMAECNN 的输入输出：根据上述介绍可以得到输入输出的具体数据，具体流程可由图 4.5 所示。输入数据 $\mathbf{X}=[L_g, W_g, L_1, W_1]$ ，将 130 组维度为 1×4 的样本数据经过图模型转换后变成 $4 \times 10 \times 1$ 。图模型使用 AE 进行神经网络学习，AE 提取的高级特征矩阵也是 $4 \times 10 \times 1$ ，再将图模型和高级特征矩阵融合成 $4 \times 10 \times 2$ 的矩阵，这是研究数据增强的核心，以此作为输入。将对应的输入输出数据以 4: 1 的比例分为训练样本集和测试样本集。

(2) 设定 IMCNN 模型具体参数设置：Epoch 数(1-800)和 Batch size(1-10)，网络结构：三层卷积层，卷积核分别为 3×5 ， 3×4 和 3×3 ，两层全连接层；批大小为 3，最大 Epoch 数为 200，Dropout 分别取 0.01，0.05，0.1，0.2，0.3，0.4；Padding 为 1。

(3) AE 模型具体参数设置：设置两个 AE，第一个隐藏层神经元节点设置 100，最

大 Epoch 为 300, 正则化 Regularization 为 0.004, 稀疏比例为 0.15; 第二个隐藏层神经元节点设置 40, 最大 Epoch 为 300, 正则化 Regularization 为 0.004, 稀疏比例为 0.15。

(4) 训练模型: IMAECNN 对输入数据进行预测。将图模型和高级特征矩阵融合成 $4 \times 10 \times 2$ 的矩阵作为 CNN 的回归预测任务的输入。系统最小误差 RMSE 阈值设置为 7.5×10^{-2} , 低于误差阈值时退出程序, 保存模型的最佳参数。

(5) 测试 IMAECNN 回归能力与优化设计: 利用剩余 26 组测试集测试网络的回归能力, 根据 PIFA 的设计要求其工作在 0.85GHz 时。系统将计算出最优尺寸参数, 并将此参数放入 MATLAB 中的 HFSS_Execute_Script 脚本进行 HFSS 仿真得到相应的回波损耗曲线值。

(6) 实验结果: 训练完成的模型参数如表 4.2 所示。为了进一步验证该系统的有效性, 根据表 4.3 所展示的优化尺寸数据进行 HFSS 仿真, 之后利用训练好 IMAECNN 预测拟合这组天线参数的 S_{11} , 并与 HFSS 软件仿真的 S_{11} 进行, 比较如图 4.6 所示。

表 4.2 模型参数

Table 4.2 IMAECNN model parameters

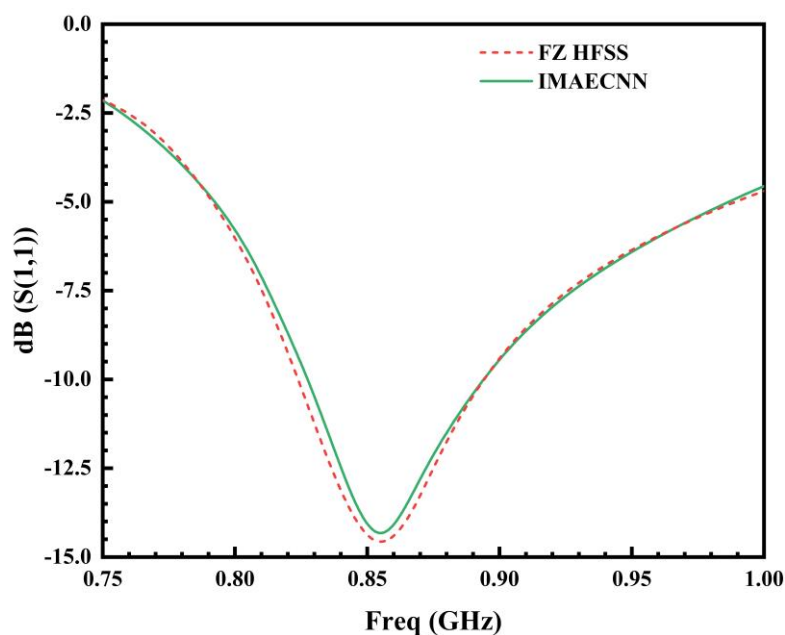
模型参数名	最优项
Batch size	3
功能优化器	sgdm
Dropout	0.1
CNN-epoch	200
AE-epoch	100

表 4.3 PIFA 优化尺寸

Table 4.3 Optimization Parameters of PIFA

变量参数名	优化尺寸参数值/mm
L_g	120
W_g	58
L_1	60
W_1	32

图 4.6 横轴为频率, 纵轴为对应的回波损耗值。红色虚线条 FZ HFSS 代表尺寸变量为[120,58,60,32]时在 HFSS 中的仿真结果。可以看出, 当频率在中心频率 0.85GHz 时, 回波损耗值可以达到-10dB, 绿色线条代表使用 IMAECNN 优化后的模型输出对该尺寸的预测结果。可以看出两条曲线基本一致, 说明本节 IMAECNN 的预测精度可以用来拟合回波损耗曲线, 帮助天线设计与优化来减少 HFSS 仿真时间。

图 4.6 PIFA 的 S_{11} 拟合图Fig.4.6 S_{11} fitting diagram of PIFA

4.4 图模型卷积神经网络自编码器对 E 形微带天线的优化设计

如前节, 利用已有的 E 形天线回波损耗数据, 测试 IMANCNN 模型在回归分析方面的性能; 使用 IMAECNN 系统产生的最优尺寸参数, 在 HFSS 软件中进行联合仿真, 绘制相应的回波损耗曲线; 最终, 将 IMAECNN 预测出的 S_{11} 曲线与 HFSS 仿真结果的曲线进行比较, 以此验证该系统在天线设计优化领域的应用效果。

4.4.1 E 形微带天线基本介绍

E 形微带天线 (E-Shaped Microstrip Antenna, ESMA) 是一种微带天线的变形, 通过在传统微带天线的辐射片上切割出特定形状的缺口 (在这种情况下是“E”形状) 来增加天线的带宽和增益。ESMA 的设计允许它在较宽的频带内工作, 同时保持较小的物理尺寸, 这使得它特别适用于需要高性能和紧凑尺寸的应用场景。相比于矩形和其他形状的微带贴片天线, ESMA 的带宽有所增加。而想要构建 ESMA, 只需调整槽的长度、宽度和位置。本节提出用 IMAECNN 算法对 ESMA 回波损耗建模, 在 ESMA 的尺寸参数和回波损耗之间建立起联系, 以便于之后的优化设计。

ESMA 是由矩形微带贴片天线插入两个平行槽所得, 它的结构如图 4.7 所示, 可以看出它的尺寸参数有贴片的长度 L 和宽度 W , 平行槽的长度 l_s 和宽度 w_s , 介质基底的厚度 h 和介电常数 ϵ_r , 每输入这一组尺寸参数集合便可得对应的回波损耗, 本次试验取 255 组。

4.4.2 E 形微带天线优化设计

根据上一小节对 ESMA 的讲解，设置输入数据为影响其工作性能的 6 个尺寸参数，包括贴片的长度 L 和宽度 W ，平行槽的长度 l_s 和宽度 w_s ，介质基板的厚度和相对介电常数，输入记作 $\mathbf{X}=[L, W, l_s, w_s, h, \epsilon_r]$ ，其中各变量的范围见表 4.4。其工作中心频率 3GHz，扫频值为 0.04GHz，故有 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{51}}]$ 。

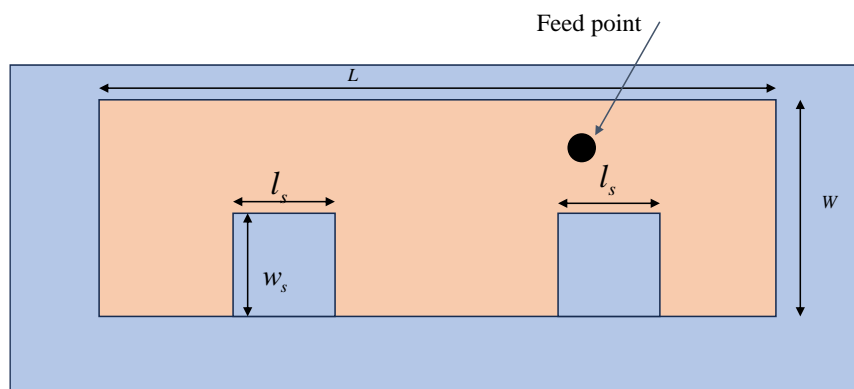


图 4.7 E 形微带贴片天线模型

Fig.4.7 E-shaped microstrip patch antenna model

表 4.4 ESMA 的相关变量及范围

Table 4.4 Variables and ranges of ESMA

变量参数名	变量范围/mm
L	25~40
W	20~30
l_s	2~9
w_s	4~26
h	1.57~3.17
ϵ_r	2.33~6.15

本次实验设计先利用数据再次验证 IMAECNN 的回归能力，之后利用系统返回的寻优数据在 HFSS 中进行仿真，最终得到满足 EMSE 设计要求的尺寸参数和 S_{11} 曲线，具体实验流程如下。

(1) 确定 IMAECNN 的输入输出：输入数据 $\mathbf{X}=[L, W, l_s, w_s, h, \epsilon_r]$ ，将 255 组维度为 1×6 的样本数据经过图模型转换后变成 $6 \times 10 \times 1$ 。通过图模型使用 AE 神经网络学习，将 AE 提取的高级特征矩阵也是 $6 \times 10 \times 1$ ，再将图模型和高级特征矩阵融合成 $6 \times 10 \times 2$ 的矩阵，以此作为输入，输出 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{51}}]$ 。将对应的输入输出数据集以 4: 1 的比例分为训练样本集和测试样本集。

(2) IMCNN 模型具体参数设置：Epoch 数(1-800)和 Batch size(1-10)，网络结构：

三层卷积层,卷积核分别为 4×5 , 4×5 和 2×2 ,两层全连接层;批大小为 3,最大 Epoch 数为 200,Dropout 分别取 0.01, 0.05, 0.1, 0.2, 0.3, 0.4; Padding 为 1。

(3) AE 模型具体参数设置:设置两层 AE,第一个隐藏层神经元节点设置 100,最大 Epoch 为 300,正则化 Regularization 为 0.004,稀疏比例为 0.15;第二个隐藏层神经元节点设置 60,最大 Epoch 为 300,正则化 Regularization 为 0.004,稀疏比例为 0.15。

(4) 训练模型:将图模型和高级特征矩阵融合成 $6 \times 10 \times 2$ 的矩阵作为 CNN 的回归预测任务的输入,模型最小误差 RMSE 阈值设置为 7.5×10^{-2} ,系统开始判断是否低于误差阈值时,退出程序,得到模型的最佳参数值。

(5) 测试模型回归能力与优化设计:利用剩余测试集测试网络的回归能力,根据 ESMA 的设计要求其工作在 3.0GHz 时,且其回波损耗小于 -15dB 。系统将计算出最优尺寸参数放入 MATLAB 中的 HFSS_Execute_Script 脚本进行 HFSS 仿真得到相应的回波损耗曲线值。

实验结果:训练完成的模型,此时改变的参数如表 4.5 所示。系统获得的最佳与目标拟合回波损耗曲线获得最优尺寸参数为 $[32.5, 25, 5, 12, 2.5, 2.33]$,具体见表 4.6,使用该尺寸在 HFSS 建模并仿真来验证优化效果。

表 4.5 IMAECNN 模型参数

Table 4.5 IMAECNN model parameters

变量参数名	优化尺寸参数值/mm
Batch size	3
功能优化器	adam
Dropout	0.15
CNN-epoch	150
AE-epoch	100

表 4.6 ESMA 优化尺寸

Table 4.6 Optimization parameters of ESMA

变量参数名	优化尺寸参数值/mm
L	32.5
W	25
l_s	5
W_s	12
h	2.5
ϵ_r	2.33

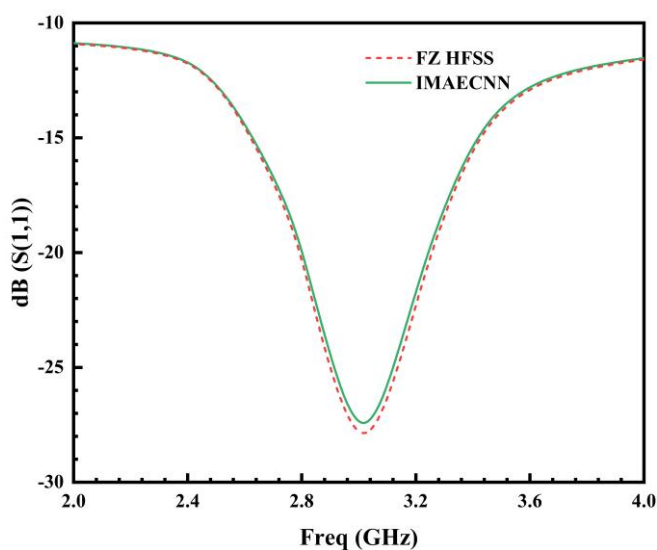
图 4.8 优化后的 ESMA 的 S_{11} 拟合图Fig.4.8 The S_{11} fitting diagram of the optimized ESMA

图 4.8 是根据返回尺寸数据在 HFSS 中仿真的回波损耗曲线与 IMAECNN 进行预测的回波损耗曲线拟合图，其中红色虚线条代表 HFSS 的仿真结果，横轴为频率区间 2GHz~4GHz，纵轴为回波损耗值。可以看出，当频率在中心频率 3.0GHz 时，回波损耗值可以达到 -25dB 以上，ESMA 要求在最深点的回波损耗值与中心频率偏差不超过 0.1GHz，故满足设计要求。绿色线条 IMAECNN 代表 IMAECNN 模型对该尺寸的预测结果，两条曲线基本一致，说明 IMAECNN 具备较强的回归预测能力。该模型也可用于其他微波器件场景的数据预测，帮助微波器件的设计与优化，减少 HFSS 仿真时间，提高设计效率。

4.5 本章小结

继第三章研究的 IMCNN 架构，本章研究了数据增强技术，通过使用 AE 神经网络进行无监督学习，提取高级特征，此过程中生成隐藏向量和近似解都可考虑作为增强数据使用。这种方法增加了数据特征的多样性。通过这种方式增强的数据帮助模型学习到更泛化的特征，使模型表现得更好。

在 IMAECNN 模型中，输入一部分是 AE 神经网络经过学习提取的高级特征，一部分是图模型特征，将这两组数据融合成到达数据增强的目的。为了验证模型的有效性，通过对单频 PIFA 天线和 E 型微带天线等进行优化设计，展示了利用 IMAECNN 模型对这些天线参数进行优化后的效果。为了验证效果，采用了 HFSS_Execute_Script 对优化后的尺寸参数进行了仿真验证，展现了该系统的广泛适用性和高效能力。

第5章 基于贝叶斯优化图模型卷积神经网络自编码器的天线建模

5.1 引言

第四章探讨了单频 PIFA 天线和 E 形微带天线的优化设计,验证了所提 IMAECNN 模型的有效性。然而,实验设置中采用的误差阈值主要关注于局部优化的实现,这意味着模型参数的调整很大程度上依赖于误差阈值的人为调整。这种方法虽然在某些情况下有效,但并不完全符合微波器件设计优化的初衷。本文研究设计意图更倾向于开发出一种能够精确回归并且具备高鲁棒性的模型,而不是仅仅依赖于局部参数的微调来实现性能的优化。

为了克服上述缺陷,本章节引入了贝叶斯优化方法。贝叶斯优化是一种基于贝叶斯统计原理的全局优化策略,结合目标函数的先验知识以及对已经获得的数据点进行概率推理,从而在优化过程中做出更加信息化的决策。BO 利用概率模型来预测目标函数的行为,根据预测结果指导搜索过程,以便更有效地探索参数空间,寻找到全局最优解。这种方法的优势在于它不仅可以处理目标函数形式未知或难以直接优化的情况,也特别适合于那些评估代价高昂的问题,如神经网络的训练过程。在文献[75]中使用改良后的 BO 对学生 T 过程进行参数优化,并对微带天线进行了有效的谐振频率建模和优化实验。与优化后的 BO 相比,全局优化的框架内提高模型的性能,减少人为调整误差阈值的困难。这不仅能够使模型在处理复杂微波器件设计问题时展现出更加精确的回归能力和更强的鲁棒性,而且还能够在其他需要优化的应用场景中发挥其潜力。由于使用神经网络缘故导致系统运行时间加长,在文献[75]中印刷偶极子天线的优化时间为 93s, E 型天线优化时间为 168s,在 BOIMAECNN 优化实验中的运行时间将在下文进行叙述。

5.2 贝叶斯优化的图模型卷积神经网络自编码器系统

5.2.1 贝叶斯优化的图模型卷积神经网络自编码器基本原理

本章节研究贝叶斯优化的图模型卷积神经网络 (Bayesian Optimized Image-based Convolutional Neural Network, BOIMCNN),分为 4 个部分:图模型模块、AE 模块、CNN 模块和 BO 模块。

在进行 BO 优化之前,步骤与第四章相同,下一步实验聚焦 BO 流程,需要设定一个优化目标函数,通过优化过程最大化或最小化的性能指标实现。一旦目标值确定,贝叶斯优化算法就会利用其内置的反馈机制来引导参数的优化过程。这种方法的关键

优势在于其能够基于先前的计算，连续地调整参数搜索策略，以期在可能的参数空间中找到最优解。

BO 是一个受欢迎的优化技术，用于有效地寻找黑盒目标函数 f 的全局最优解。它通过更新对目标函数 f 的先验知识并获得相应的观测值，进而更新一个更接近于 f 的后验分布，能够迅速地定位到最优解。

$$x^* = \arg \min_{x \in \mathbb{R}^d} f(x) \quad (5.1)$$

BO 由两个模块构成，一个是用于拟合目标函数的概率模型模块，另一个是用于获取新观测值的采集函数模块。GP 作为一种非参数的核学习方法，可以用公式(5.2)表示，其中包含一个均值函数 $\mu(x)$ 和一个协方差函数 $k(x, x')$ 。

$$f(x) \sim GP(\mu(x), k(x, x')) \quad (5.2)$$

$D_{1:n} = \{\{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_n, y_n\}\}$ 表示已观测集合，则可以得到协方差矩阵，此时协方差矩阵 \mathbf{K} 如(5.3)所示。

$$\mathbf{K} = \begin{bmatrix} k(x_1, x_1) & \cdots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \cdots & k(x_n, x_n) \end{bmatrix} \quad (5.3)$$

每输入新的观测值 x_{n+1} ，假定 GP 的均值为零，则输出 $y_{1:n}$ 和 y_{n+1} 的联合高斯分布可表示为：

$$\begin{bmatrix} y_{1:n} \\ y_{n+1} \end{bmatrix} \sim N(0, \begin{bmatrix} k & k \\ k^T & k(x_{n+1}, x_{n+1}) \end{bmatrix}) \quad (5.4)$$

k 可表示为：

$$k = [k(x_{n+1}, x_1) \bullet k(x_{n+1}, x_2) \bullet \dots \bullet k(x_{n+1}, x_n)] \quad (5.5)$$

计算得到 y_{n+1} 的后验概率：

$$p(y_{n+1} | D_{1:n}, x_{n+1}) = N(\mu(x_{n+1}), \sigma^2(x_{n+1})) \quad (5.6)$$

其中

$$\mu(x_{n+1}) = k^T \mathbf{K}^{-1} y_{1:n} \quad (5.7)$$

$$\sigma^2(x_{n+1}) = k(x_{n+1}, x_{n+1}) - k^T \mathbf{K}^{-1} k \quad (5.8)$$

经过上述公式的计算，GP 可以预测出 x_{n+1} 处的输出值，得到预测均值和预测方差，从而使用采集函数就可以去开采下一个可信度最高的观测值。

BO 方法常见的采集函数包括置信下限^[70](Lower confidence bound, LCB)、概率提高^[71](Probability of Improvement, PI)、熵搜索^[72](Entropy Search, ES)以及期望提高^[73](Expected Improvement, EI)，具体函数公式如表 5.1 所示。通过选择最佳的采集函

数能使寻求得的观测值对应的输出损失最小，损失 r_n 表示为：

$$r_n = |y^* - y_n| \quad (5.9)$$

其中 y^* 表示当前的最优观测值对应的输出预测值，表示 y_n 真实值。

表 5.1 常用采集函数

Table 5.1 Common collection functions

采集函数	公式
LCB	$a_{LCB}(x; \theta) = \mu(x) - \sqrt{\nu \beta_t s^2(x)}, \beta_t = 2 \ln(Dt^2 \pi^2 / (6)\delta)$
PI	$a_{PI}(x; \theta) = \Pr[f(x) > \theta_t] = 1 - \Phi(\gamma(x))$
EI	$a_{EI}(x; \theta_t) = [\phi(\gamma(x)) - \gamma(x)\mathbb{Q}(\gamma(x))]\sigma_{t-1}(x)$
ES	$a_{ES}((\theta, \varsigma); D_n) = E[\eta_n^* - \eta_{n+1}^* \theta_{n+1}, \varsigma_{n+1}] = (\theta, \varsigma)$

算法 5-1 贝叶斯优化算法

Algorithm5-1 Bayesian Optimization

算法 5-1: BO

输入：样本集 $D_{1:n} = \{\{\tilde{x}_1, \tilde{y}_1\}, \{\tilde{x}_2, \tilde{y}_2\}, \dots, \{\tilde{x}_n, \tilde{y}_n\}\}$ ；

训练参数 x 的上限和下限；

初始点 $\{x_1, x_2, x_3, \dots, x_N\}$

过程：

- 1: for $i=1,2,\dots,m$ do
- 2: 通过采集函数采集新的点 x^{opt}
- 3: $D_{1:n} \rightarrow$ 训练 GP , $x^{opt} \rightarrow$ 训练好的 GP 预测 f^{opt}
- 4: 对比保存最大/最小 f^{opt} 以及对应 x^{opt}
- 5: end for

BO 框架是一个迭代过程，主要包含 3 个步骤：

第 1 步，根据最大化采集函数来选择下一个最有“潜力”的评估点 x_n ；

第 2 步，根据选择的评估点 x_n 评估目标函数值 $y_n = f(x_n) + \varepsilon_n$ ；

第 3 步，把新得到的输入-观测值对 $\{x_n, y_n\}$ 添加到历史观测集 $D_{1:n-1}$ 中，并更新概

率代理模型，为下一次迭代做准备。

5.2.2 贝叶斯优化的图模型卷积神经网络自编码器基本流程

下面是贝叶斯优化神经网络的基本流程，如图 5.1 所示。

(1) 定义优化目标：第四章预先定义阈值 RMSE，使 IMAECNN 小于阈值就是满

足天线设计需求，本章节将 IMAECNN 的 RMSE 作为贝叶斯优化的目标函数，优化 IMAECNN 内部超参数，以此达到全局优化和轻量化模型这两个目的。

(2) 选择先验模型：根据目标函数的先验知识选择一个适当的概率模型是至关重要的一步。概率模型的作用不仅是为了预测不同参数配置下目标函数的可能输出，而且还要估计这些预测的不确定性。GP 是一种非常灵活的概率模型，能够通过考虑函数输出之间的相关性来捕捉目标函数的行为，并且能够自然地量化预测的不确定性。

(3) 确定获取函数：在不确定性高的区域探索和在已知性能好的区域进一步寻优之间找到平衡。给定当前已知的最优值，采用新参数进行实验所能带来改进的期望值，置信下限定义为

$$a_{LCB}(x; \theta) = \mu(x) - \sqrt{v\beta_t s^2(x)}, \beta_t = 2 \ln(Dt^2 \pi^2 / (6)\delta) \quad (5.10)$$

(4) 评估目标函数：利用新的参数配置来实际运行或测试目标函数，以便获得关于模型性能的新的观测数据。这些数据不仅提供了特定参数下目标函数输出的具体数值，而且这些新的观测结果将进一步用于更新贝叶斯优化过程中的概率模型。

(5) 更新模型：将新的观测结果加入到数据集中，并使用贝叶斯规则更新先验分布，从而得到新的后验分布。这一步骤通常涉及到重新计算高斯过程的参数。

(6) 迭代或终止：如果满足终止条件（如达到预设的迭代次数、性能改进不显著等），则停止优化过程；否则，返回步骤 4 继续迭代。

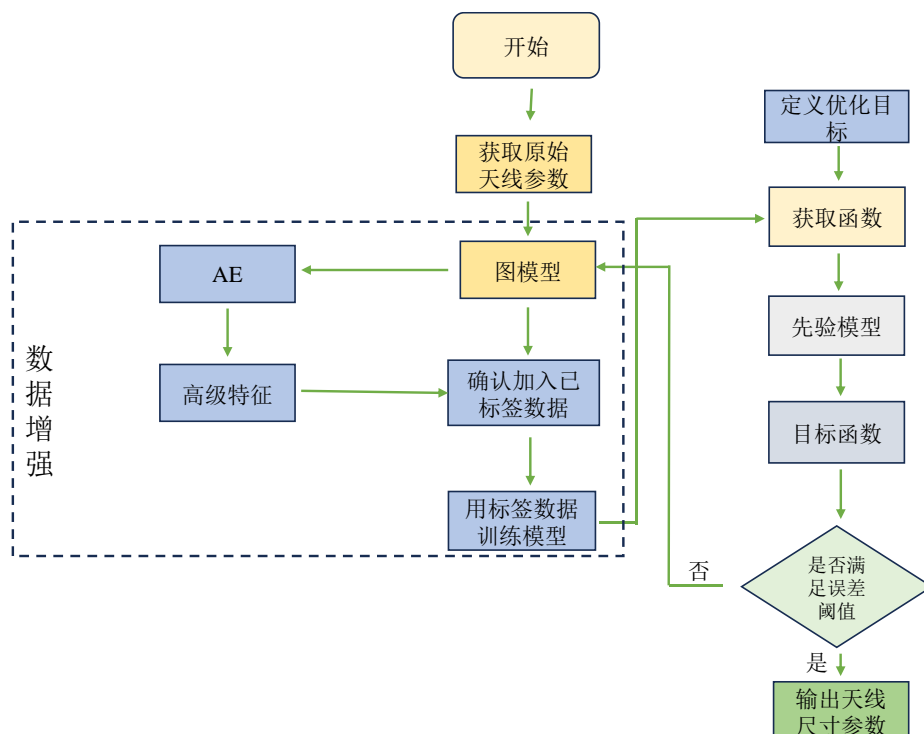


图 5.1 BOIMAECNN 的流程图

Fig.5.1 Flowchart of BOIMAECNN

5.3 贝叶斯优化图模型卷积神经网络自编码器对双频微带天线建模

本小节使用上面描述的 BOIMAECNN 对 DFMA^[74]进行回波损耗建模。在设计 DFMA 时,需要考虑的指标主要包括天线的尺寸参数以及模型的参数设置,对这些指标进行综合考虑和优化,可以设计出具有更加优秀性能的 DFMA。在本节的实验中,使用 BOIMAECNN 来进行优化设计。

5.3.1 双频微带天线模型

DFMA 采用了矩形贴片结构,基板选用了介电常数 ϵ_r 为 2.4 的 FR4 材料。天线的双频工作原理基于对贴片结构的几何尺寸和加载技术的精确控制,通过调整贴片的长度和宽度,以及馈电点的位置,可在不同的频率上激发共振,从而实现双频工作。DFMA 的 HFSS 的建模结构如图 5.2 所示,俯视图如图 5.3 所示。根据 DFMA 的设计原则,影响其性能的主要参数分别是馈电点在贴片中的位置 L_1, L_2 , 贴片的长度 L_0 , 宽 W_0 , 并使用 BOIMAECNN 和 HFSS 联合仿真对 DFMA 的这些尺寸参数和其对应的回波损耗扫频,得到的数据进行建模。

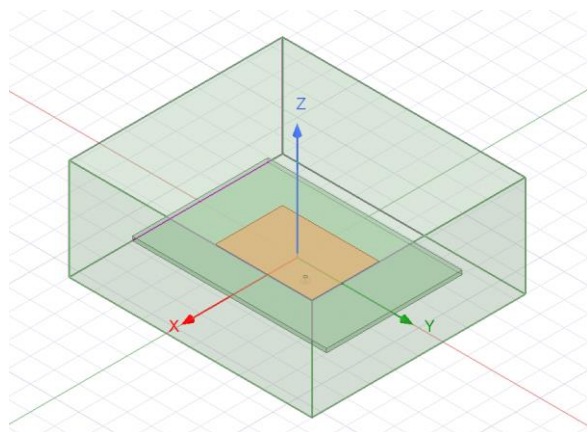


图 5.2 DFMA 建模图

Fig.5.2 DFMA modeling diagram

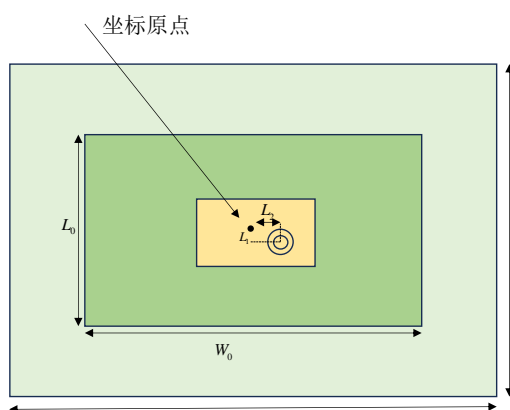


图 5.3 DFMA 俯视结构图

Fig.5.3 DFMA overlooking structure diagram

5.3.2 双频微带天线的回波损耗建模

根据 DFMA 的原理，设置输入数据为影响天线性能的 4 个尺寸参数，记作 $\mathbf{X}=[L_0, L_1, L_2, W_0]$ ，其中各变量的范围见表 5.2。实验中取 130 组输入数据， $H=1.6$ 。

考虑到 DFMA 的设计指标是当天线工作在中心频率 2.45GHz 时，覆盖 400MHz 带宽，且其回波损耗值小于 -10dB。所以在设置输出的时候，把扫频范围设置为 1.5GHz~2.75GHz，扫频步长为 0.05GHz，即输出标签数据记作 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{26}}]$ 。

表 5.2 DFMA 变量参数

Table 5.2 DFMA variable parameters

变量名	变量值范围/mm
L_0	25~30
L_1	5~7
L_2	9~11
W_0	38~40

(1) 确认 BOIMAECNN 的输入输出：采用 HFSS 对该天线进行仿真，获取天线原始数据，输入记作 $\mathbf{X}=[L_0, L_1, L_2, W_0]$ ，训练输出为 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{26}}]$ 。通过 HFSS 获取 130 组训练数据后，26 组用于测试。

(2) IMCNN 模型具体参数设置：优化器为 SGD；网络结构：三层卷积层，卷积核分别为 3×5 ， 3×5 和 2×2 ，两层全连接层；批大小为 3，最大 epoch 数为 200，Dropout 取 0.1；Padding 为 1。

(3) AE 模型具体参数设置：设置两个 AE，第一个隐藏层神经元节点设置 100，最大 epoch 为 100，正则化 regularization 为 0.004，稀疏比例为 0.15；第二个隐藏层神经元节点设置 40，最大 epoch 为 100，正则化 regularization 为 0.004，稀疏比例为 0.15。

(4) BO 模型具体参数设置：先验模型使用 GP，采集函数是 LCB，以 IMAECNN 预测天线回波损耗的 RMSE 作为目标函数，对网络的训练周期，正则化系数，Dropout 数和批处理大小进行寻优。

(5) 模型训练：训练网络过程中以 RMSE 为目标。为了达到目标最小，进行一系列的模型参数优化工作，通过迭代调整和优化这些参数，旨在找到一组最优参数配置。

(6) 数据测试：DFMA 的设计要求其工作在 2.45GHz 时，且其回波损耗小于 -10dB。系统将计算出最优尺寸参数，放入 MATLAB 中的 HFSS_Execute_Script 脚本进行 HFSS 仿真，得到相应的回波损耗曲线值。

训练结果：在表 5.3 中 BatchSize 表示的是模型的最小批处理数量；Epoch 表示的是最大循环周期，LearnRate 表示的是初始学习率；L2 表示的是正则化系数；Dropout 表示的是下采样系数。系统运行时间为 224.35s，在图 5.4 中展示的 BOIMAECNN 模

型的收敛过程中目标函数的输出值与预测值的走势图，很明显采用 GP 作为概率模型和 LCB 作为采集函数的 BO 模型的性能随着迭代次数的增加而更加精准，且在第七次训练就达到收敛。最佳尺寸参数如表 4 所示。

表 5.3 BOIMAECCN 模型参数

Table 5.3 BOIMAECCN model parameters

输出值	估计值	BatchSize	Epoch	LearnRate	L2	Dropout
0.058935	0.058935	9	15	0.0056048	4.8728e-05	0.078978
0.058935	0.059697	3	26	0.0039482	0.00037779	0.19198
0.0010928	0.0050012	2	44	0.0039893	3.3611e-05	0.03732
0.0010928	0.0011014	10	49	0.005657	7.1475e-05	0.16617
0.0010928	0.0011073	8	41	0.0016175	0.00096592	0.039145
0.0010928	0.0047242	9	46	0.0073296	1.0046e-05	0.1999
0.0010928	0.0010973	10	10	0.0028467	4.4126e-05	0.18413
0.0010928	0.0027527	10	50	0.0011862	3.7792e-05	0.19489
0.0010928	0.001099	9	49	0.0017092	0.00015021	0.052303
0.0010928	0.0011006	10	44	0.007054	7.5276e-05	0.073207
0.0010928	0.0010971	10	47	0.0060855	4.7933e-05	0.036044
0.0010928	0.0011027	10	14	0.0010842	0.00097697	0.13756
0.0010928	0.0010509	10	49	0.0028334	0.00031269	0.12626
0.0010928	0.001005	7	41	0.0024376	0.00010812	0.17045
0.0010928	0.001028	20	44	0.0064309	0.0004705	0.054108
0.0010928	0.0010196	20	43	0.0012699	1.7618e-05	0.017415
0.0010928	0.0010058	6	44	0.0011236	0.00050039	0.071717
0.0010928	0.0010973	20	50	0.0096337	1.5971e-05	0.066558
0.0010928	0.001076	20	48	0.0090427	8.9041e-05	0.098016
0.0010928	0.0010995	4	40	0.0045866	1.7439e-05	0.0034267

图 5.5 表示根据模型预测的最优天线尺寸在 HFSS 仿真得到的天线回波损耗曲线，横轴为频率，纵轴为对应的回波损耗值。红色虚线条 FZ HFSS 代表尺寸变量为 [27,2,9,39,28] 时在 HFSS 中的仿真结果，可以看出，当频率在中心频率 2.45GHz 时，回波损耗值可以达到 -10dB，在 1.5GHz~2.75GHz 频率区间有超过 400MHz 的带宽，故

满足设计要求；绿色线条表使用 BOIMAECCN 对该回波损耗的预测结果。从图中可以看出两条曲线基本一致，说明本章节所研究的 BOIMAECCN 在达到一定迭代次数后，模型具备良好的收敛性能，达到设计的最初目标。

表 5.4 BOIMAECCN 预测 DFMA 天线尺寸

Table 5.4 BOIMAECCN predicts DFMA antenna size

变量参数名	优化尺寸参数值/mm
L_0	27
L_1	2
L_2	9
W_0	39

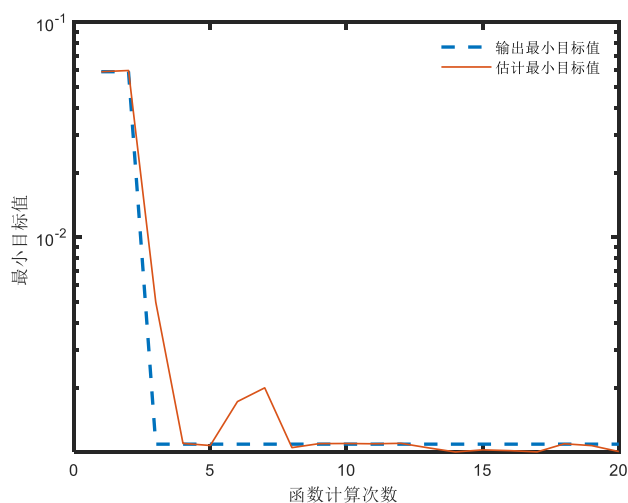


图 5.4 目标函数值

Fig.5.4 Value of objective function

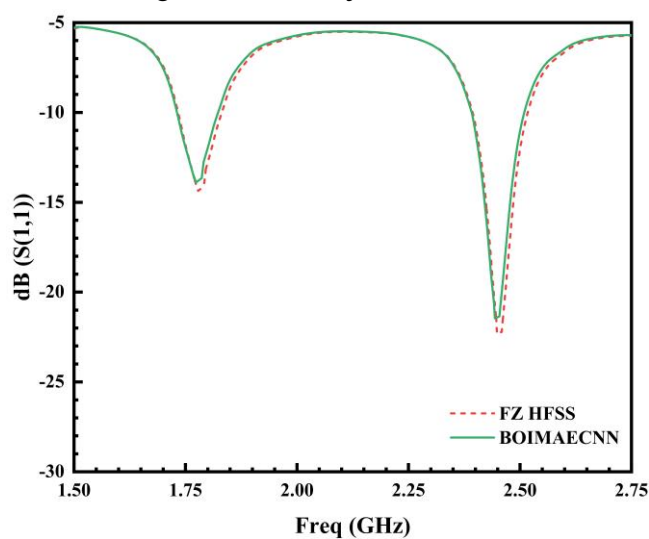


图 5.5 DFMA 的 S11 拟合图

Fig.5.5 S11 fitting diagram of DFMA

5.4 贝叶斯优化图模型卷积神经网络自编码器对双频单极子天线建模

为了进一步验证 BOIMAECNN 的性能,本节对 DFMOA(Dual-Frequency Monopole Antenna, DFMOA)进行建模。

5.4.1 双频单极子天线模型

DFMOA 是一种特殊设计的无线电天线,能够在两个不同的频率上工作。这种天线的设计允许它捕获来自两个独立信号,使其在多种通信系统中非常有用。DFMOA 的 HFSS 的建模结构如图 5.6 所示,尺寸图如图 5.7 所示。DFMOA 工作频率在 2.4GHz 和 5.2GHz,在本次研究中,主要集中在 DFMOA 中心频率在 2.4GHz 时的 S_{11} 建模。

根据 DFMOA 的设计原则,影响其性能的主要参数分别是贴片尺寸参数 $L, L_1, L_2, R_1, R_2, S, W$ 以及介质板的高度 H 。本次实验选择这 8 个尺寸变量参数作为输入数据,并使用 BOIMAECNN 和 HFSS 联合仿真,基于这些仿真数据进行建模。通过上述两种建模方式比较,验证所使用的 BOIMAECNN 精准建模能力。

5.4.2 双频单极子天线回波损耗建模

根据 DFMOA 的原理,设置输入数据为影响天线性能的 8 个尺寸参数,此时模型输入记作 $\mathbf{X}=[L, L_1, L_2, R_1, R_2, W, S, H]$,其中各变量的范围可根据表格 5.5 查看,设置相对介电常数 ϵ_r 为 2.4,取 205 组试验数据。

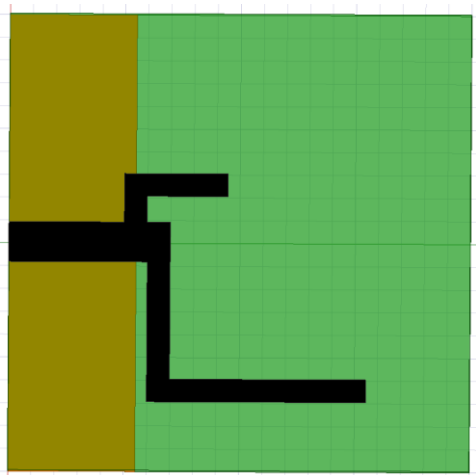


图 5.6 DFMOA 建模图

Fig.5.6 DFMOA modeling diagram

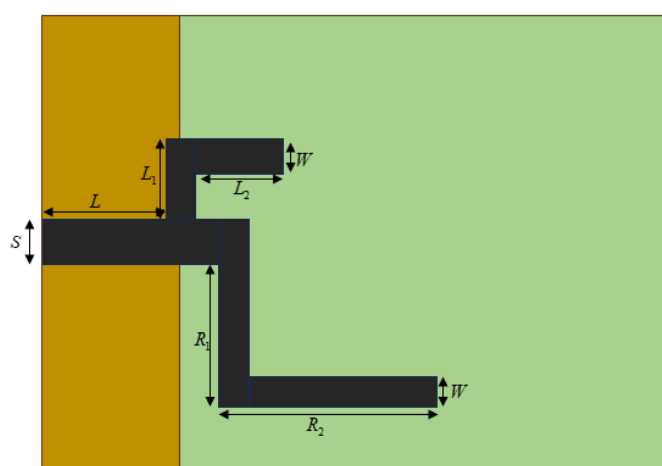


图 5.7 DFMOA 俯视结构图

Fig.5.7 DFMOA overlooking structure diagram

DFMOA 的设计指标是当天线工作在中心频率 2.4GHz 时,覆盖 400MHz 带宽,且其回波损耗值小于 -10dB。扫频范围设置为 2.2GHz~2.6GHz,扫频步长为 0.01GHz,输出标签数据记作 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{41}}]$ 。

表 5.5 DFMOA 变量参数
Table 5.5 Fixed size parameters of DFMOA

变量	Value/mm	变量	Value/mm
L	13~15	R_2	18~20
L_1	3~5	W	1~3
L_2	13~15	S	3~5
R_1	9~11	H	1~3

(1) 确认 BOIMAECNN 的输入输出：采用 HFSS 对该天线进行联合仿真获取天线原始数据，输入记作 $\mathbf{X}=[L, L_1, L_2, R_1, R_2, W, S, H]$ 。训练输出为天线的性能参数，记作 $\mathbf{Y}=[S_{11_1}, S_{11_2}, \dots, S_{11_{41}}]$ 。获取 256 组训练数据后，随机选取其中的 164 组用于训练，41 组用于的测试。

(2) IMCNN 模型具体初始化参数设置：功能优化器为 SGD；网络结构：三层卷积层，卷积核分别为 4×5 , 4×5 和 2×2 ，两层全连接层；批大小为 3，最大 Epoch 数为 200，Dropout 取 0.1；Padding 为 1。

(3) AE 模型具体初始化参数设置：设置两层 AE，第一个隐藏层神经元节点设置 150，最大 Epoch 为 100，正则化 Regularization 为 0.004，稀疏比例为 0.15；第二个隐藏层神经元节点设置 80，最大 Epoch 为 100，正则化 Regularization 为 0.004，稀疏比例为 0.15。

(4) BO 模型具体参数设置：先验模型使用 GP，采集函数是 LCB，使用天线回波损耗 RMSE 作为目标函数，对网络的训练周期、正则化系数、Dropout 数和批处理大小进行寻优。

(5) 模型训练：设定好目标函数，模型通过不断添加新的样本点来更新目标函数的后验分布，以此为基础进行寻优。为了达到这一目标，通过迭代调整和优化这些参数，旨在找到一组最优参数配置。

(6) 数据测试：对 DFMOA 测试数据进行验证。DFMOA 的设计要求其工作在 2.45GHz 时，且其回波损耗小于 -25dB。系统将计算出最优尺寸参数，放入 MATLAB 中的 HFSS_Execute_Script 脚本进行 HFSS 仿真，得到相应的回波损耗曲线值。

训练结果：在图 5.4 中展示的 BOIMAECNN 模型的收敛过程中目标函数的输出值与预测值的走势图，系统运行时间为 285.55s，很明显采用 GP 作为概率模型和 LCB 作为采集函数的 BO 模型的性能随着迭代次数的增加而更加精准，且在第六次训练就达到收敛。在表 5.6 中参数表示的含义同上节相同。模型在收敛时的最佳参数设置，如表 5.7 所示。

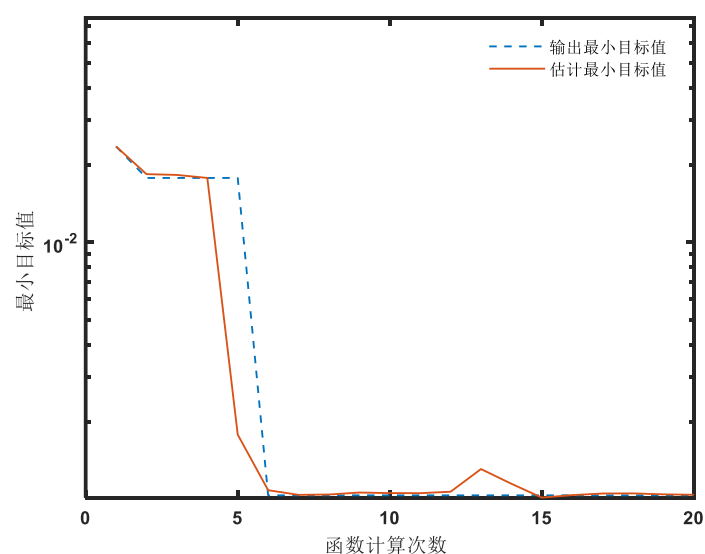


图 5.8 目标函数值

Fig.5.8 Value of objective function

表 5.6 BOIMAECCN 模型参数

Table 5.6 BOIMAECCN model parameters

观测值	估计值	BatchSize	Epoch	LearnRate	L2	Dropout
0.023669	0.023669	9	46	0.0087371	1.7226 e-5	0.16122
0.017841	0.018465	8	21	0.0015307	5.3053 e-5	0.19411
0.017841	0.018333	5	16	0.0058145	3.7291 e-5	0.061532
0.017841	0.017844	6	40	0.0013824	2.4609e-5	0.01398
0.017841	0.017842	6	35	0.0028337	1.0959e-5	0.19998
0.0010348	0.0010837	8	27	0.0010201	1.0592e-5	7.4024e-5
0.0010348	0.0010392	8	48	0.0016308	1.0004e-5	0.00656
0.0010348	0.0010424	8	29	0.0031674	53.285e-5	0.0002652
0.0010348	0.001461	8	26	0.0016161	6.0556e-5	0.10534
0.0010348	0.001456	10	19	0.0050487	19.919e-5	0.000265
0.0010348	0.001455	10	10	0.0015481	6.9116e-5	0.18512
0.0010348	0.001469	10	11	0.0013728	3.418e-5	0.15273
0.0010348	0.001431	2	11	0.0012122	1.4035e-5	0.001125
0.0010348	0.001848	7	19	0.0011985	17.843e-5	0.00043192
0.0010348	0.001414	9	31	0.0049753	1.6098e-5	0.00089626
0.0010348	0.001437	10	25	0.0017833	39.54e-5	0.00052738
0.0010348	0.001452	9	22	0.0051856	4.9044e-5	0.00061189
0.0010348	0.001453	10	11	0.0038376	2.7591e-5	0.19922
0.0010348	0.001044	10	24	0.0016497	1.5764e-5	0.016366
0.0010348	0.00104	10	11	0.0020453	16.867e-5	0.0033091

表 5.7 BOIMAECNN 预测 DFMOA 天线尺寸
Table 5.7 BOIMAECNN predicts DFMOA antenna size

变量参数名	优化尺寸参数值/mm
L	13
L_1	3
L_2	14
R_1	9
R_2	20
W	2
S	3
H	1

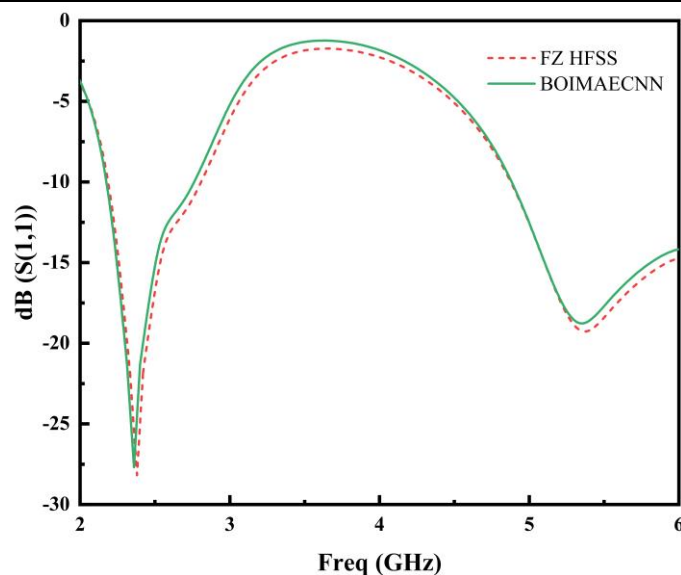


图 5.9 S_{11} 对比图

Fig.5.9 Comparison of S_{11}

图 5.9 横轴为频率，纵轴为对应的回波损耗值。红色线条 FZ HFSS 代表尺寸变量为[13,3,14,9,20,2,3,1]时在 HFSS 中的仿真结果；绿色线条代表 BOIMAECNN 模型的预测结果。可以看出，当频率在中心频率 2.4GHz 时，回波损耗值可以达到-25dB，在 2.2GHz~2.6GHz 频率区间，DFMOA 最深点与回波损耗值与中心频率相差不超过 0.1GHz，两条曲线基本一致，说明本章节所研究的 BOIMAECNN 的在达到一定迭代次数后，模型具备良好的收敛性能，预测精度可以用来拟合回波损耗曲线，而且具备自我更新参数的能力，故满足设计要求。

5.5 本章小结

在第四章基础上，本章节通过将 BO 方法集成到 IMAECNN 回归模型中，优化其超参数，进一步增强模型的性能。通过引入 BO，旨在利用其全局优化的优势，自动调整和优化 IMAECNN 模型的超参数，从而提升模型的预测精度和泛化能力。BO 能够以相对较少的迭代次数找到全局最优解。在 IMAECNN 模型中引入 BO，意味着能够更加智能和高效地调整那些影响模型预测性能的关键超参数，而无需人工干预。

为了充分验证 BOIMAECNN 模型的有效性和优势，选择 DFMA 和 DFMOA 天线的回波损耗建模。BOIMAECNN 模型展现出了优异的预测精度和泛化能力，在两种天线设计的回波损耗预测问题上均获得了显著的结果。与传统方法相比，BOIMAECNN 模型能够显著降低对实际物理测试的依赖，减少了天线设计和优化过程中的时间和成本开销。

总结与展望

利用构建代理模型替代电磁仿真软件进行建模优化成为了一大热点。代理模型通过学习输入与标签之间的映射关系,可以进行高精度预测,同时减少调用仿真软件的时间。本文选择使用图像领域愈发凸显的卓越性能的卷积神经网络,并对其进行改进。同时,大部分代理模型的训练建模需要使用可靠的标签样本数据,而这些基本来源于耗时的仿真软件,因此本文中使用了数据增强技术来进一步提高微波器件的优化效率。本文的主要内容如下:

首先,研究了基于 IMCNN 的深度学习系统。利用图模型的完备性、健全性以及非冗余性使得 CNN 提取特征的数据有效,经过卷积和池化提取特征后做预测回归任务,使用 RMSA 和 CMSA 的尺寸参数和谐振频率数据集进行模型训练与测试,验证了 IMCNN 预测回归的功能。

之后,研究了 IMAECNN。利用 AE 隐藏层提取的高级特征与原始图模型融合作为 CNN 的输入,对单频 PIFA 天线和 E 形微带天线进行优化设计实验,将寻优结果在 HFSS 中仿真验证。

最后,研究了 BOIMAECNN,以探讨并证实贝叶斯优化技术在处理复杂电磁设计挑战时的实用性和效果。实验通过引入一种综合了 BO、AE 和 CNN 的先进代理模型,对两种天线设计进行了精确的优化计算。利用 HFSS_Execute_Script 这一自动化脚本,验证设计的正确性和有效性。

本文所做工作仍存在一些不足与局限,需要更加深入研究与完善,可以从以下几个方面做出进一步的研究:

(1) 本文目的是构建有效代理模型来优化设计微波器件,微波器件目前用的较多的包括常用的线性互易元件包括:微波贴片天线、功分器、定向耦合器、滤波器、移相器。在本文中,实验中使用的天线模型居多,因此接下来可以增加一些其他微波器件进行实验验证本文提出系统在优化设计微波器件中的有效性。

(2) 本文在研究优化天线问题时,一些固定的尺寸数据是按照经验和其他已有实验结果设置的,所以之后的实验选择需要优化的变量,可以将之前的固定尺寸也作为优化目标进行训练,来进一步探讨其是否对优化结果有明显的影响。

由于作者水平有限,论文中难免存在错误与不足,恳请读者批评指正。

参考文献

- [1] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [2] Zhou D, Lu L, Zhao J, et al. A new learning algorithm based on strengthening boundary samples for convolutional neural networks[C]. MATEC Web of Conferences. EDP Sciences, 2020, 327: 02004.
- [3] Konak A, Coit D W, Smith A E. Multi-objective optimization using genetic algorithms: A tutorial[J]. Reliability engineering & system safety, 2006, 91(9): 992-1007.
- [4] 冯茜, 李擎, 全威, 等. 多目标粒子群优化算法研究综述[J]. 工程科学学报, 2021, 43(06): 745-753.
- [5] Kapetanakis T N, Vardiambasis I O, Ioannidou M P, et al. Neural network modeling for the solution of the inverse loop antenna radiation problem[J]. IEEE Transactions on Antennas and Propagation, 2018, 66(11): 6283-6290.
- [6] Massa A, Marcantonio D, Chen X, et al. DNNs as applied to electromagnetics, antennas, and propagation—A review[J]. IEEE Antennas and Wireless Propagation Letters, 2019, 18(11): 2225-2229.
- [7] Yao H M, Jiang L. Machine-learning-based PML for the FDTD method[J]. IEEE Antennas and Wireless Propagation Letters, 2018, 18(1): 192-196.
- [8] Shao W, Du Y. Microwave imaging by deep learning network: Feasibility and training method[J]. IEEE transactions on antennas and propagation, 2020, 68(7): 5626-5635.
- [9] Jacobs J P. Accurate modeling by convolutional neural-network regression of resonant frequencies of dual-band pixelated microstrip antenna[J]. IEEE Antennas and Wireless Propagation Letters, 2021, 20(12): 2417-2421.
- [10] Zhang X, Tian Y, Zheng X. Optimal design of fragment-type antenna structure based on PSO-CNN[C]. 2019 International Applied Computational Electromagnetics Society Symposium-China (ACES). IEEE, 2019, 1: 1-2.
- [11] Skaria S, Al-Hourani A, Lech M, et al. Hand-gesture recognition using two-antenna Doppler radar with deep convolutional neural networks[J]. IEEE Sensors Journal, 2019, 19(8): 3041-3048.
- [12] Cai J, Zhong R, Li Y. Antenna selection for multiple-input multiple-output systems based on deep convolutional neural networks[J]. Plos one, 2019, 14(5): e0215672.
- [13] Wang B, Wang S, Zeng D, et al. Convolutional Neural Network-Based Radar Antenna Scanning Period Recognition[J]. Electronics, 2022, 11(9): 1383.
- [14] Lovato R, Gong X. Phased antenna array beamforming using convolutional neural networks[C]//2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting. IEEE, 2019: 1247-1248.
- [15] Li M S, Abdullah M, He J, et al. Frequency-Diverse Antenna with Convolutional Neural Networks for Direction of Arrival Estimation in Terahertz Communications[J]. IEEE Transactions on Terahertz Science and Technology, 2024.
- [16] Ahmed A M, Eissa O, Sezgin A. Deep autoencoders for DOA estimation of coherent sources using imperfect antenna array[C]. 2020 Third International Workshop on Mobile Terahertz Systems (IWMTS). IEEE, 2020: 1-5.

- [17] Tao J, Chen J, Xing J, et al. Autoencoder neural network based intelligent hybrid beamforming design for mmWave massive MIMO systems[J]. IEEE Transactions on Cognitive Communications and Networking, 2020, 6(3): 1019-1030.
- [18] Tong Y, Wang M, Zhou T. Prediction of Array Antenna Assembly Accuracy Based on Auto-Encoder and Boosting-OSKELM[J]. Processes, 2023, 11(5): 1460.
- [19] Um K S, Kim N J, Heo S W. Surrogate - based model using auto - encoder for optimising multi - band antennas[J]. IET Microwaves, Antennas & Propagation, 2022, 16(11): 725-732.
- [20] Qi G, Bao S, Wei Y, et al. Accurate antenna design by deep auto-encoder surrogate model assisted particle swarm optimization[C]. 2022 IEEE 5th International Conference on Electronic Information and Communication Technology (ICEICT). IEEE, 2022: 875-880.
- [21] Shojaei S P, Soleimani H, Soleimani M. Supervised Autoencoder-Based Beamforming Approach for Satellite mmWave Communication[J]. International Journal of Antennas and Propagation, 2024.
- [22] Seyfioğlu M S, Özbayoğlu A M, Gürbüz S Z. Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities[J]. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(4): 1709-1723.
- [23] Shahriari B, Swersky K, Wang Z, et al. Taking the human out of the loop: A review of Bayesian optimization[J]. Proceedings of the IEEE, 2015, 104(1): 148-175.
- [24] Shilton A, Gupta S, Rana S, et al. Accelerated Bayesian Optimisation through Weight-Prior Tuning[C]. International Conference on Artificial Intelligence and Statistics. PMLR, 2020: 635-645.
- [25] Astudillo R, Frazier P. Bayesian optimization of composite functions[C]. International Conference on Machine Learning. PMLR, 2019: 354-363.
- [26] Noda F, Nishijima K, Furuya K. Hyperparameter Tuning of the Shunt-murmur Discrimination Algorithm Using Bayesian Optimization[C]. 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2020: 929-933.
- [27] Lim Y F, Ng C K, Vaitesswar U S, et al. Extrapolative Bayesian optimization with Gaussian process and neural network ensemble surrogate models[J]. Advanced Intelligent Systems, 2021, 3(11): 2100101.
- [28] Snoek J, Rippel O, Swersky K, et al. Scalable bayesian optimization using deep neural networks[C]. International conference on machine learning. PMLR, 2015: 2171-2180.
- [29] Yenicelek D. Parameter optimization using high-dimensional Bayesian optimization[J]. arxiv preprint arxiv:2010.03955, 2020.
- [30] Daulton S, Balandat M, Bakshy E. Parallel bayesian optimization of multiple noisy objectives with expected hypervolume improvement[J]. Advances in Neural Information Processing Systems, 2021, 34: 2187-2200.
- [31] Ginsbourger D, Le Riche R, Carraro L. Kriging is well-suited to parallelize optimization[M]. Computational intelligence in expensive optimization problems. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010: 131-162.
- [32] Ramachandram D, Lisicki M, Shields T J, et al. Bayesian optimization on graph-structured search spaces: Optimizing deep multimodal fusion architectures[J]. Neurocomputing, 2018, 298: 80-89.
- [33] Tracey B D, Wolpert D. Upgrading from gaussian processes to student's t processes[C]. 2018 AIAA Non-Deterministic Approaches Conference. 2018: 1659.
- [34] Clare C, Hawe G, McClean S. Expected Regret Minimization for Bayesian Optimization with Student's-t Processes[C]. Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition. 2020: 8-12.

- [35] Liu K, Cui Y, Ren J, et al. An improved particle swarm optimization algorithm for Bayesian network structure learning via local information constraint[J]. IEEE access, 2021, 9: 40963-40971.
- [36] Qian C, Xiong H, Xue K. Bayesian optimization using pseudo-points[J]. arxiv preprint arxiv:1910.05484, 2019.
- [37] Belakaria S, Deshwal A, Doppa J R. Max-value entropy search for multi-objective Bayesian optimization[J]. Advances in neural information processing systems, 2019, 32.
- [38] Zhang S, Yang F, Yan C, et al. An efficient batch-constrained bayesian optimization approach for analog circuit synthesis via multiobjective acquisition ensemble[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2021, 41(1): 1-14.
- [39] Zhan D, Qian J, Cheng Y. Pseudo expected improvement criterion for parallel EGO algorithm[J]. Journal of Global Optimization, 2017, 68: 641-662.
- [40] Yang A, Xue W, Tian C, et al. A design of silicon based horn antenna optimized using bayesian optimization[C]. 2020 IEEE 6th International Conference on Computer and Communications (ICCC). IEEE, 2020: 438-442.
- [41] Chowdhury S R, Gopalan A. No-regret algorithms for multi-task bayesian optimization[C]. International Conference on Artificial Intelligence and Statistics. PMLR, 2021: 1873-1881.
- [42] Tran A, Eldred M, McCann S, et al. srMO-BO-3GP: A sequential regularized multi-objective Bayesian optimization for constrained design applications using an uncertain Pareto classifier[J]. Journal of Mechanical Design, 2022, 144(3): 031705.
- [43] Coutinho M G F, Torquato M F, Fernandes M A C. Deep neural network hardware implementation based on stacked sparse autoencoder[J]. IEEE Access, 2019, 7: 40674-40694.
- [44] Dörner S, Rottacker S, Gauger M, et al. Bit-wise Autoencoder for Multiple Antenna Systems[C]. 2021 17th International Symposium on Wireless Communication Systems (ISWCS). IEEE, 2021: 1-5.
- [45] Zhang H H, Yao H M, Jiang L, et al. Fast full-wave electromagnetic forward solver based on deep conditional convolutional autoencoders[J]. IEEE Antennas and Wireless Propagation Letters, 2022, 22(4): 779-783.
- [46] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]. Proceedings of the 25th international conference on Machine learning. 2008: 1096-1103.
- [47] Kingma D P, Welling M. An introduction to variational autoencoders[J]. Foundations and Trends® in Machine Learning, 2019, 12(4): 307-392.
- [48] Liu Y, Ponce C, Brunton S L, et al. Multiresolution convolutional autoencoders[J]. Journal of Computational Physics, 2023, 474: 111801.
- [49] Yang C, Wang X, Mao S. Unsupervised detection of apnea using commodity RFID tags with a recurrent variational autoencoder[J]. IEEE access, 2019, 7: 67526-67538.
- [50] Amiriparian S, Freitag M, Cummins N, et al. Sequence to sequence autoencoders for unsupervised representation learning from audio[C]. DCASE. 2017: 17-21.
- [51] Creswell, Antonia, and Anil Anthony Bharath. "Denoising adversarial autoencoders[J]." IEEE transactions on neural networks and learning systems 30.4 (2018): 968-984.
- [52] Victoria A H, Maragatham G. Automatic tuning of hyperparameters using Bayesian optimization[J]. Evolving Systems, 2021, 12(1): 217-223.
- [53] Kolar D, Lisjak D, Pająk M, et al. Intelligent fault diagnosis of rotary machinery by convolutional neural network with automatic hyper-parameters tuning using bayesian optimization[J]. Sensors, 2021, 21(7): 2411.

- [54] Kopsiaftis G, Protopapadakis E, Voulodimos A, et al. Gaussian process regression tuned by bayesian optimization for seawater intrusion prediction[J]. Computational intelligence and neuroscience, 2019.
- [55] Lho D, Park J, Park H, et al. Bayesian optimization of high-speed channel for signal integrity analysis[C]. 2019 IEEE 28th Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS). IEEE, 2019: 1-3.
- [56] Mern J, Yildiz A, Sunberg Z, et al. Bayesian optimized monte carlo planning[C]. Proceedings of the AAAI conference on artificial intelligence. 2021, 35(13): 11880-11887.
- [57] 洪智超, 宗智, 王轶赓等. 基于正交试验方法的重叠网格技术不确定度研究[J]. 船舶力学, 2020, 24(7): 865-873.
- [58] 李婵娟, 邵特, 傅世强. 基于 MATLAB 和 HFSS 的叠层微带天线优化设计[J]. 信息技术, 2016, 2016(2):61-63.
- [59] Angiulli G, Versaci M. Resonant frequency evaluation of microstrip antennas using a neural-fuzzy approach[J]. IEEE Transactions on Magnetics, 2003, 39(3): 1333-1336.
- [60] Zhang X, Tian Y, Zheng X. Optimal design of fragment-type antenna structure based on PSO-CNN[C]. 2019 International Applied Computational Electromagnetics Society Symposium-China (ACES). IEEE, 2019, 1: 1-2.
- [61] Shah P, Moghaddam M. Super resolution for microwave imaging: A deep learning approach[C]. 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting. IEEE, 2017: 849-850.
- [62] Zou F, Shen L, Jie Z, et al. A sufficient condition for convergences of adam and rmsprop[C]. Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. 2019: 11127-11135.
- [63] Guney K, Sagiroglu S, Erler M. Generalized neural method to determine resonant frequencies of various microstrip antennas[J]. International Journal of RF and Microwave Computer - Aided Engineering: Co - sponsored by the Center for Advanced Manufacturing and Packaging of Microwave, Optical, and Digital Electronics (CAMPmode) at the University of Colorado at Boulder, 2002, 12(1): 131-139.
- [64] Singh J, Singh A P, Kamal T S. Design of circular microstrip antenna using artificial neural networks[C]. Proceedings of the World Congress on Engineering. 2011, 2(1): 9-12.
- [65] Sağıroğlu Ş, Güney K, Erler M. Resonant frequency calculation for circular microstrip antennas using artificial neural networks[J]. International Journal of RF and Microwave Computer - Aided Engineering: Co - sponsored by the Center for Advanced Manufacturing and Packaging of Microwave, Optical, and Digital Electronics (CAMPmode) at the University of Colorado at Boulder, 1998, 8(3): 270-277.
- [66] Badra N M, Allam A, El-Rafei A. WiFi antenna design and modeling using artificial neural networks[C]. 2019 International Conference on Innovative Trends in Computer Engineering (ITCE). IEEE, 2019: 270-274.
- [67] Kushwah V S, Tomar G S. Design and analysis of microstrip patch antennas using artificial neural network[J]. Trends in Research on Microstrip Antennas, 2017: 55.
- [68] Fan X H, Tian Y B, Zhao Y. Optimal design of Yagi microstrip antenna based on particle swarm optimization with fitness estimation[C]. 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama). IEEE, 2018: 653-660.
- [69] 商锋, 李晓娇. 一种宽带单频 PIFA 天线的设计[J]. 太赫兹科学与电子信息学报, 2019, 17(03): 445-456.

- [70] Pal A, Zhu L, Wang Y, et al. Constrained surrogate-based engine calibration using lower confidence bound[J]. IEEE/ASME Transactions on Mechatronics, 2021, 26(6): 3116-3127.
- [71] De Villiers D I L, Couckuyt I, Dhaene T. Multi-objective optimization of reflector antennas using kriging and probability of improvement[C]. 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting. IEEE, 2017: 985-986.
- [72] Takeno S, Fukuoka H, Tsukada Y, et al. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization[C]. International Conference on Machine Learning. PMLR, 2020: 9334-9345.
- [73] Zhan D, Xing H. Expected improvement for expensive optimization: a review[J]. Journal of Global Optimization, 2020, 78(3): 507-544.
- [74] Tari O, Belen A, Mahouti P, et al. Surrogate Based Design Optimization of Multi-Band Antenna[C]. 2021 International Applied Computational Electromagnetics Society Symposium (ACES). IEEE, 2021: 1-4.
- [75] 李清. 学生 T 过程结合贝叶斯优化算法及在天线问题上的应用 [D]. 江苏科技大学, 2023. DOI:10.27171/d.cnki.ghdcc.2023.000216.