

分类号: TM615

密 级: 公开

单位代码: 10433

学 号: 19504010487

# 山东理工大学

## 工程硕士学位论文

基于 GA-BP 神经网络的光伏阵列在线故障检测研究

Research on Online Fault Detection of Photovoltaic Array Based  
on GA-BP Neural Network

研 究 生: 刘 东

指 导 教 师: 李田泽教授

协 助 指 导 教 师:

学 位 类 别: 工 程 硕 士

专 业 领 域: 电 气 工 程

研 究 方 向: 分 布 式 电 源 并 网 技 术 及 应 用

论 文 完 成 日 期: 2022 年 6 月 8 日

## 独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得山东理工大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名：刘东

时间：2022年6月8日

## 关于论文使用授权的说明

本人完全了解山东理工大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件和磁盘，允许论文被查阅和借阅；学校可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后应遵守此协议)

研究生签名：刘东

时间：2022年6月8日

导师签名：

李曰峰

时间：2022年6月8日

# 《学位论文出版授权书》

本人完全同意《中国优秀博硕士学位论文全文数据库出版章程》(以下简称“章程”),愿意将本人的学位论文提交“中国学术期刊(光盘版)电子杂志社”在《中国优秀博硕士学位论文全文数据库》中全文发表。《中国优秀博硕士学位论文全文数据库》可以以电子、网络及其它数字媒体形式公开出版,并同意编入《中国知识资源总库》,在《中国博硕士学位论文评价数据库》中使用和在互联网上传播,同意按“章程”规定享受相关权益。

作者签名: 刘东

2022年6月8日

导师签名: 李国军

2022年6月8日

研究生学号	10433-19504010487	所在院系	电气与电子工程学院
-------	-------------------	------	-----------

## 摘要

太阳能作为最有应用前景的可再生能源，光伏发电技术已经引起了世界各国的高度重视，自 2014 年起，我国的光伏产业迎来了爆发性的增长，但光伏发电系统在实际运行时易受环境因素影响，导致光伏阵列发生故障，为减少发电量损失，提高光伏发电系统运行的可靠性，如何快速准确的完成光伏阵列故障的检测是目前亟需解决的问题，基于此，本文所作研究如下：

(1) 研究光伏电池工作原理及输出特性。通过分析光伏电池的数学模型及其等效电路结构，对其工作原理进行研究，在此基础上，研究电池单体在正常运行时其输出特性与光照和温度的关系，确定电池的典型参数的变化。

(2) 建立光伏阵列模型，模拟实际阵列运行中可能发生的故障类型，并收集故障数据。根据光伏电池的发电原理及阵列的连接结构的深入研究，结合在淄川新明珠光伏发电中心的实地考察，在 MATLAB/Simulink 中建立 4 行 3 列的仿真光伏阵列，在研究阵列故障产生原因的基础上，模拟故障运行状态，分析各故障状态下阵列输出特性与正常无故障时的差异，并获取实时故障运行数据。

(3) GA-BP 神经网络在光伏阵列故障检测中的可行性研究。针对 BP 神经网络初始权值阈值的随机选取对故障检测结果的干扰较大这一问题，本文在分析 BP 神经网络的结构组成及具体参数选择的基础上，采用遗传算法 (Genetic Algorithm, GA) 优选 BP 神经网络的初始权值阈值，建立基于 GA-BP 神经网络的光伏阵列故障检测模型，对光伏阵列常见四种故障状态进行检测。

(4) 为进一步提高 GA-BP 神经网络对光伏阵列故障类型的检测性能，提出一种融入鸽群优化算法 (Pigeon-Inspired Optimization Algorithm, PIO) 的混合遗传算法优化 BP 神经网络，分析两种算法的不同结合方式，对两个优化算法相互融合的可能性展开深入研究，并对鸽群优化算法的原始参数进一步改进。

(5) 研究经混合遗传算法优化后的 BP 神经网络在进行光伏阵列故障检测时与原始 GA-BP 神经网络的性能差异。对比分析实验结果表明：改进后 GA-BP 神经网络的迭代次数远少于原始 GA-BP 神经网络，有效缩短了检测时间，提高了故障检测模型的快速性，且改进后的 GA-BP 神经网络对多种故障状态的综合检测正确率达到 98.7%，同时对老化故障和阴影遮挡故障的检测也更为精准。

**关键词：**光伏发电；故障检测；BP 神经网络；遗传算法；鸽群优化算法

## Abstract

Solar energy is the most promising renewable energy, photovoltaic power generation technology has attracted the high attention of countries around the world, China's photovoltaic industry has ushered in explosive growth since 2014, but photovoltaic power generation system is vulnerable to environmental factors in actual operation and resulting in the failure of photovoltaic arrays. In order to reduce the loss of power generation and improve the reliability of photovoltaic power generation system operation, how to quickly and accurately complete the detection of photovoltaic array failure is the current urgent need to solve the problem, based on this, the research conducted in this paper is as follows.

(1) Study the working principle and output characteristics of photovoltaic cells. The working principle is studied by analyzing the mathematical model of PV cell and its equivalent circuit structure. On this basis, the relationship between the output characteristics of the cell monomer and light and temperature during normal operation is investigated, and the changes of typical parameters of the cell are determined.

(2) Establish a PV array model to simulate the types of faults that may occur in the operation of the actual array and collect fault data. According to the in-depth study of the power generation principle of PV cells and the connection structure of the array, combined with the fieldwork in Zichuan Xinmingzhu PV power generation center, a simulated PV array with four rows and three columns is established in MATLAB/Simulink, so as to simulate the common fault states of PV arrays under different fault conditions, to study the causes of array faults, to analyze the output characteristics of the array under each fault state compared with the normal fault-free condition, and real-time fault operation data are obtained.

(3) Feasibility study of GA-BP neural network in PV array fault detection. This paper uses Genetic Algorithm (GA) to select the initial weight threshold of BP neural network based on the analysis of the structural composition and specific parameter selection of BP neural network, and establishes the fault diagnosis model of PV array based on GA-BP neural network on account of the problem that the random selection of the initial weight threshold of BP neural network

interferes with the fault detection results. The model is used to detect four kinds of faults that often occur during the actual operation of PV arrays.

(4) In order to further improve the detection performance of GA-BP neural network for photovoltaic array fault types, a hybrid genetic algorithm integrated with Pigeon-Inspired Optimization Algorithm (PIO) was proposed to optimize BP neural network. The different combination methods of the two algorithms are analyzed, conducts in-depth research on the possibility of the integration of the two optimization algorithms, and analyzes the original parameters of the pigeon flock optimization algorithm were further improved.

(5) The performance difference between the BP neural network optimized by hybrid genetic algorithm and the original GA-BP neural network in photovoltaic array fault detection is studied. Comparative analysis of the experimental results shows that the number of iterations of the improved GA-BP neural network is much less than that of the original GA-BP neural network, which effectively shortens the detection time and improves the rapidity of the fault detection model, and the improved GA-BP neural network achieves a comprehensive detection accuracy of 98.7% for various fault states, and is also more accurate for the detection of aging faults and shadow obscuration faults.

**Key words: photovoltaic power generation; fault detection; BP neural network; Genetic Algorithm; Pigeon-Inspired Optimization Algorithm**

## 目录

摘要.....	I
ABSTRACT.....	II
目录.....	IV
<b>第一章 绪论.....</b>	<b>1</b>
1.1 研究背景与意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	2
1.2 光伏阵列故障检测技术国内外研究现状.....	3
1.2.1 红外图像检测方法.....	4
1.2.2 数学模型检测方法.....	4
1.2.3 智能网络检测方法.....	5
1.3 本文主要研究内容.....	6
<b>第二章 光伏阵列输出特性及故障分析.....</b>	<b>8</b>
2.1 光伏电池发电特性研究.....	8
2.1.1 光伏电池发电原理分析.....	8
2.1.2 光伏电池数学模型研究.....	9
2.1.3 光伏电池输出特性分析.....	11
2.2 光伏阵列连接结构及输出特性分析.....	14
2.2.1 光伏阵列的连接结构.....	14
2.2.2 光伏阵列正常运行时输出特性.....	16
2.3 光伏阵列不同故障状态下输出特性分析.....	17
2.3.1 光伏阵列故障形成原因分析.....	17
2.3.2 光伏阵列不同故障状态下输出特性分析.....	18
2.4 本章小结.....	22
<b>第三章 GA-BP 神经网络应用于光伏阵列故障检测研究.....</b>	<b>23</b>
3.1 神经网络分析.....	23
3.1.1 神经网络.....	23
3.1.2 神经网络特点分析.....	23
3.2 BP 神经网络结构及工作原理分析.....	24
3.2.1 BP 神经网络的结构分析.....	24
3.2.2 BP 神经网络工作原理分析.....	25

3.2.3 BP 神经网络的缺陷分析 .....	27
3.3 GA-BP 神经网络的构建.....	28
3.3.1 遗传算法理论分析.....	28
3.3.2 GA 优化 BP 神经网络思路分析.....	29
3.3.3 GA-BP 神经网络实现流程.....	30
3.4 GA-BP 神经网络应用于光伏阵列故障检测研究 .....	32
3.4.1 应用于光伏阵列故障检测的 BP 神经网络研究.....	32
3.4.2 光伏阵列故障数据预处理.....	36
3.4.3 GA-BP 神经网络应用于光伏阵列故障检测的验证分析 .....	38
3.5 本章小结 .....	42
<b>第四章 混合遗传算法优化 BP 神经网络的研究 .....</b>	<b>43</b>
4.1 鸽群算法原理.....	43
4.1.1 地图和指南针算子.....	43
4.1.2 地标算子.....	44
4.2 鸽群算法分析.....	46
4.2.1 鸽群算法局限性分析.....	46
4.2.2 鸽群算法实现流程分析 .....	46
4.2.3 鸽群算法参数的改进研究.....	48
4.3 混合遗传算法优化 BP 神经网络研究.....	49
4.3.1 遗传算法优化 BP 神经网络时存在的缺陷分析 .....	49
4.3.2 混合遗传算法的实现研究.....	50
4.3.3 混合遗传算法优化 BP 神经网络.....	51
4.4 本章小结 .....	53
<b>第五章 改进型 GA-BP 神经网络在线光伏阵列故障检测仿真验证分析.....</b>	<b>54</b>
5.1 改进型 GA-BP 神经网络实验参数设定 .....	54
5.1.1 输入输出变量的选取.....	54
5.1.2 改进型 GA-BP 神经网络各组成部分参数的选取.....	55
5.2 改进型 GA-BP 神经网络应用于光伏阵列故障检测实验结果与分析.....	56
5.2.1 适应度对比结果分析.....	56
5.2.2 改进型 GA-BP 神经网络模型训练.....	57
5.2.3 改进型 GA-BP 神经网络在光伏阵列故障检测中的实验结果与分析 .....	59
5.3 本章小结 .....	62
<b>第六章 总结与展望.....</b>	<b>63</b>
6.1 总结.....	63
6.2 展望.....	64



---

参考文献.....	65
附录.....	68
在读期间公开发表的论文和参与项目.....	74
致谢.....	75

# 第一章 绪论

## 1.1 研究背景与意义

### 1.1.1 研究背景

能源,是人类赖以生存发展的最主要物质资源,是经济社会发展不断前进的基础力量。能源问题,是当今世界人类社会发展进步必须首先解决的基础问题,目前世界各国所需的90%能源是由天然气、煤炭、石油等化石能源供应,而随着全球工业技术水平的发展与提升,人类对化石能源持续性增长的需求量与化石能源的不可再生性,及其所存总量的日益减少的矛盾越来越尖锐,并且化石能源在使用的过程中,也不可避免的增加了温室气体的排放量,在此消彼长的严峻形式下<sup>[1]</sup>,世界各国将目光投向了新能源的发掘和利用。

新能源一般是指结合新技术、新材料加以开发利用的可再生清洁能源。其中太阳能作为一种非常具有代表性的新能源,有着传统化石燃料发电技术不具备的优势:首先,太阳能光伏发电技术最大的优势在于太阳光几乎可以照射到地球表面的角角落落,地球上太阳能资源极为充足,哪怕用取之不竭用之不尽来比喻也不为过,因此利用太阳能进行发电完全无需担心一次能源缺乏的问题;其次,太阳能光伏发电更为安全环保,利用风能发电会产生巨大噪声,利用传统化石能源获取电力则会产生不可逆的环境污染,而光伏发电系统运行时极为安静并且十分清洁,在整个发电过程中几乎没有气体或者固体的产生。另外,光伏发电布局灵活多样,一般装设在建筑表面,构成光电建筑一体系统,大幅提高了土地资源的利用率,在边远地区或是特殊地域可以因地制宜设置分布式光伏电站,缓解上述地区的供电困难的问题。

中国,作为世界第二大经济体,在加快发展社会主义国民经济的今天,将重点着眼于新能源的进一步健康发展和使用,为保障能源供应的可持续性,环境的绿色无污染,我国高举发展绿色能源的旗帜,是促进全球绿色能源蓬勃发展的主要引领者和践行者,在当前为实现“碳达峰、碳中和”目标的新形势下,2020年9月22日,我国国家总书记习近平郑重向全世界人民表示,中国的二氧化碳排放力争于2030年前达到峰值(即碳达峰),努力争取2060年前实现碳中和<sup>[3]</sup>。在中国实现“双碳”目标的愿景下,更在要求我国的能源行业需要构建以新能源发电技术为主导的新型电力系统,实现能源体系的高效低能、洁净无污染<sup>[4-5]</sup>。在我国新能源的发现和利用技术中,当属太阳能光伏发电技术的发展速度如火如荼,其装机容量也是随年递增,我国在2008年至2014年迎来的光伏装机的爆发式增长,从2016年以来已基本呈现了线性上升的态势,光伏发电给电力系统的发展融入了全新的生命力,也使得我国的能源消费结构获得了更进一步的优化,据中华人民共和国2021年国民

经济和社会发展统计公报指出<sup>[6]</sup>，自 2016 年至 2021 年，我国太阳能光伏发电装机总量不断提高，具体数据如图 1.1 所示，且 2021 年全国并网太阳能发电装机容量达到 3259 亿千瓦时，同比增长 25.1%，占全社会用电量的 3.9%。

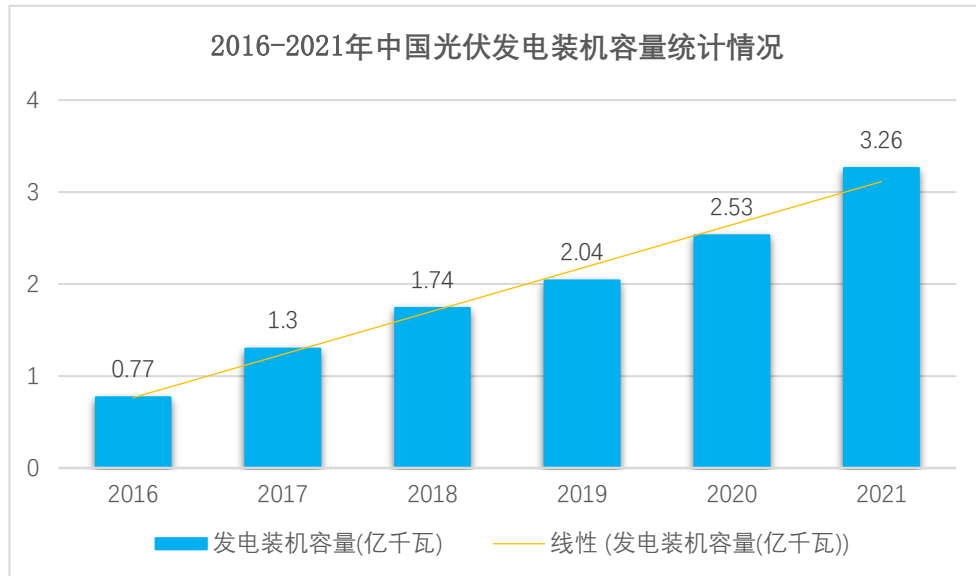


图 1.1 2016-2021 年中国光伏发电装机容量

Fig. 1.1 China's installed photovoltaic power generation capacity from 2016 to 2021

### 1.1.2 研究意义

世界各国不断增加光伏电厂的部署，新的挑战也随之而来。光伏电厂早先主要设置于太阳光照充足的戈壁，荒漠，山地等地区，不过如今光伏发电的概念早已深入人心，在农村的屋顶，路边矗立的太阳能路灯，屋外悬挂的太阳能空调外机，还有处于不断发展中的光伏汽车，早已覆盖了人类生活的方方面面，这也就意味着，如果光伏阵列突然发生故障，除了会给电力系统自身造成巨大的发电量损失之外，更有可能导致火灾等严重安全事故危及人们的生命安全，这也对光伏阵列在线故障检测技术的发展提出了新要求。

从发电量的角度来看，相对于水力发电和风力发电，囿于新型高效光伏电池材料研究进度缓慢的困境，光伏发电系统的能量转换率处于中等偏低的水平。在 Jiang L. L. 等学者论述的一篇文章指出，由于各种故障的出现，整个太阳能发电系统所承受的功率损耗接近 18.9%<sup>[7]</sup>，而由大量的光伏电池单体经过复杂的串并联而构成的光伏阵列作为整个太阳能发电系统的核心，能够健康平稳的工作，是保证光伏发电系统安全运行的关键。为了减弱故障给光伏发电系统带来的危害，造成不必要的发电量浪费，对光伏发电系统各装置实时情况的获取，并快速检测故障是否发生显得尤为重要。

另一方面，由光伏阵列发生故障引起的重大安全事故并不少见。据北极星太阳能光伏

网报道<sup>[8]</sup>, 2018年7月26日, 德国索林根当地一座住宅的太阳能板屋顶起火, 整个屋顶几乎全部遭到了烧毁, 这场灾难致使多个家庭的财产和生命安全受到了威胁, 现场照片如图1.2所示。而据索比光伏网报道, 在2020年, 国内因光伏阵列发生故障引起的事故就多达120起, 可见, 如何能够保证第一时间准确的获知光伏系统是否处于不正常的运行状态对维护我们的生命财产安全也有着重要意义。



图 1.2 德国太阳能电池板事故

Fig. 1.2 Accident of German photovoltaic array on the roof

从上述例子可以看出, 研究光伏阵列的故障检测技术, 在光伏电站设备的工作状况下实施在线检测, 及时准确的判断阵列有无出现故障, 可以提升整个太阳能发电体系的工作安全性、稳定性, 而且在一定程度上也可以减少不必要的能量损失, 弱化对化石能源的依赖性, 加快国家能源体系向可持续能源体系迈进的步伐, 加速绿色发电系统的建设进程。

## 1.2 光伏阵列故障检测技术国内外研究现状

光伏电厂一般建设在太阳能覆盖充裕的室外环境下, 受自然条件的影响较大, 因此光伏电厂中发电元件会经常发生各种故障, 如暴雪压塌部分光伏电池板造成部分线路断路, 大风天气下吹动电池板连接线造成短路, 雨水侵蚀电池板内部造成电池寿命大幅缩短, 提前老化等现象, 因此, 如何快速准确的检测阵列是否发生故障以及确定故障类型对于光伏供电系统健康运行和可靠供电来说显得尤为重要。

通过阅读文献, 总结整理国内外光伏阵列故障检测技术的方法, 大致可分为红外图像检测法, 数学模型检测方法, 智能网络检测方法<sup>[9]</sup>。

### 1.2.1 红外图像检测方法

在相同的发电环境下，光伏组件在正常工作与带故障运作时，会散发不同的热能，比如发电系统在带有被阴影遮挡的光伏组件工作时，故障电池会成为整个阵列中的负载，消耗阵列中其他组件产生的电能，造成自身温度的升高。红外图像检测方法，就是根据这一原理，利用红外线成像设备，将各种不可见的辐射能量转换为热图像，用不同的色彩表示各部分运行时的不同温度，再对热图像进行研究分析，提取光伏元件的温度特征，从而检测故障是否发生以及故障组件所在位置。红外图像检测法特别依赖于红外设备的优良，而且一般来说，越精密的红外设备所成图像分辨率越高，造价越为昂贵，对于光伏电站检测故障手段的选择，经济因素也是需要考量的一个重要条件。

2014年，Chine W 等人便是利用热成像仪来进行光伏阵列的故障检测<sup>[10]</sup>。他们利用红外线成像仪器测量光伏组件的工作温度，测算每个电池片的工作温度和整个光伏阵列的平均工作温度之间的差值，并依据温差值是否超过了提前设定好的阈值，来判断光伏模块有无出现故障。这种方法实现难度小，原理简单，但是一所大型光伏电站，光伏组件有成百上万个，对红外成像设备的精度要求十分严苛，成本过于昂贵。

2016年，Dotenco S 等学者设计了一种针对光伏阵列热斑故障的检测方法<sup>[11]</sup>，先利用热成像装置收集光伏电站的总体温度数据，然后对总体数据进行阈值分割和方差估计从而确定电站中的光伏阵列，其后，将光伏阵列再划分成光伏器件个体，最后获得元件中的温度中值，再利用格拉布斯检验和狄克逊检验判别出温度异常的光伏组件，此方法对于站内环境简单的光伏电站有着较好的效果，但是检测到的故障类型较为单一，且对于环境复杂的电站并不适用。

2017年，Alsafasfeh M 等学者提出了将 SLIC (Simple Linear Iterative Clustering) 算法用于对光伏电厂的红外检测图像的照片识别上来<sup>[12]</sup>，这个方案主要使用 SLIC 算法将红外图片细分为几个子区，然后通过提取几何、温度特征实现对热斑故障的识别，但是目前该方法仍停留在小型实验阶段，还没有很好的应用到实际中。

2021年，栾艳军等学者提出一种利用改进 Otsu\_ Iteration\_ Canny 算法与航拍红外图像相结合的方法来进行光伏阵列故障检测<sup>[13]</sup>，有效的将红外图像中的故障组件与周围环境相分割，提高了故障识别的准确性。

### 1.2.2 数学模型检测方法

数学模型检测方法的基本思想是通过分析光伏电站的发电规律，不同时刻、不同运行状态下的电压电流等参数，搭建一个理想化的数学模型，这种模型主要分为发电量预测模型、环境变化模型以及电气参数变化模型等。通过数学模型计算出理想输出电压、电流、功率等参数，之后与实际光伏阵列输出参数作对比来判断阵列是否发生故障。该方法比较

显著的特点便是检测效果的优劣特别依赖于数学模型构建的精准性，容错率比较低。

2014年，Silvestre S等学者提出了设置光伏阵列发生故障时各参数变化阈值的方法进行实时故障检测<sup>[14]</sup>。首先通过合理的分配传感器在阵列中的位置，检测阵列电压电流在不同工作状态下的电压与电流，建立了不同连接结构的光伏阵列电压电流变化数学模型，并从该模型中收集运行数据，最后与实际阵列运行数据进行对比，如果两者误差大于提前设定的阈值，则判断阵列发生故障。

2018年，Harrou F等学者将人工蜂群算法(ABC)引入到光伏阵列故障参数分类中去<sup>[15]</sup>。他们首先根据温度以及光照强度的变化，预测光伏阵列最大功率点电流、电压、功率的理论值，并利用ABC算法进行归类，然后建立了理想状态下的阵列参数变化模型，最后采集实际电厂电压电流及功率数值，与数学模型中输出的数据进行参数辨识并确定阵列是否处于故障状态，但是这种方法只能检测故障是否已经发生，并不能检测出具体的故障类型。

2020年，刘强等学者提出一种根据光伏阵列等效电路建立数学模型完成对故障电池检测识别的方法<sup>[16]</sup>。首先分析不同运行工况下的串联等效电阻的大小，然后将所得数据与所构建的等效电路模型进行对比，最终可以检测故障是否发生，并且该团队还通过在数学模型中引入 $\eta$ 因子用来估计故障严重程度。

### 1.2.3 智能网络检测方法

智能网络检测方法主要思想是将人工神经网络与智能寻优计算相结合，把从阵列中收集到的电压电流等数据交由神经网络来处理，通过采用智能算法优化网络结构或训练模型，可以自动建立故障数据特征值与故障类型间的映射关系，以便反应不同的故障类型。

2013年，王元章等人提出了利用BP神经网络来识别光伏阵列故障的方案<sup>[17]</sup>。他们针对不同运行状况下的光伏阵列输出特性曲线，以获取故障状态特征参数，从而构建了BP神经网络故障检测模型，但是由于并未对BP神经网络的结构和参数进行改进，导致该网络的检测性能不够理想。

同年，Massi Pavan等人提出利用贝叶斯神经网络进行光伏阵列故障检测的方法<sup>[18]</sup>。与王元章等人的网络模型不同的是，贝叶斯神经网络输入数据为光伏阵列故障前后的功率差，而且这种网络所用训练数据庞大，检测耗时较长，但是在检测准确性上有所提高。

2015年，乐治后等人将径向基神经网络应用于光伏阵列故障检测<sup>[19]</sup>。但是RBF神经网络缺少相应的理论依据，并且该网络也存在需要大量样本数据训练，克服网络陷入局部最优的弊端。同时，径向基神经网络结构远复杂于BP神经网络，并不能保障检测过程所要求的短时性。

2017年，Dhimsish M等人利用模糊逻辑分类器完成了对光伏阵列的故障诊断<sup>[20]</sup>。该方法首先通过实际运行数据计算每一时刻的电压比、电流比、功率比以及环境损失，接着将这些数据输入到经过正常数据训练的分类器中检测故障类型。

2019年,吴擎学者在卷积神经网络基础上引入经验模态分解的方法对光伏并网系统的逆变器进行分析研究,从而判断光伏阵列是否发生故障<sup>[21]</sup>。这种方法目前还停留在实验室仿真阶段,未能考虑实际电厂温度与环境的变化,但是该方法的优势在于可以分辨光伏发电系统直流侧六种不同类型的故障,对于复杂太阳能发电系统有很好的借鉴意义。

2021年,学者马辉提出了一种基于随机森林的光伏电站故障检测模型<sup>[22]</sup>,将采集到阵列数据通过随机森林算法进行分类,之后再输入到神经网络中进行训练,有效的降低了不平衡数据对于网络训练过程的影响,提高了检测的准确性。

除上述三类光伏阵列故障检测方法之外,还有一些更为典型的检测技术,在文献<sup>[23-26]</sup>中,学者们为了实现获取参数与经济效益的平衡,详细分析了在不同连接结构的光伏阵列中,对电压、电流传感器数量及位置的选取原则,其核心思想是使用相对较少的传感器收集可以描述光伏系统运行状况的必要运行参数以及周围环境参数,以便于确定光伏阵列有无出现故障。在文献<sup>[27-30]</sup>中,一些学者提出了应用于光伏阵列的故障检测的时域反射法和快速过采样主成分分析法。时域反射法是把一种特殊信号输入到传输线中,通过信号分析器得到该附加信号在传输线中的波形,分析波形的走向来判断阵列是否发生故障,这种方法的优势在于可以发现故障产生的大体位置,方便检修人员及时切除故障,但是该方法最大的缺点在于无法完成实时监测,只能离线检测。而快速过采样方法在不受辐照变化影响的情况下,通过监测各组串的电流信号,计算光伏串电流信号的异常变化来判断光伏组件是否有故障发生,可以完成对光伏电站在线检测的任务,但是该方法所能检测故障种类较少,不能很好的适用于实际。

从目前对光伏阵列故障检测的研究趋势来看,智能网络检测方法已经成为光伏阵列故障检测技术的主流。一方面是这种方法使用成本较低,并不需要大量工作人员对光伏电站进行周期性的检查,大大减少了人力物力的投入;另一方面在整个故障检测分类过程中,并不需要整个光伏电站停电,只需要读取每一部分光伏阵列外接传感器的电压电流等参数的实时数据就可以完成检测过程。

### 1.3 本文主要研究内容

光伏发电技术作为绿色发电技术的代表,在经过了二十几年的发展之后,现在的光伏产业已经具备了相当规模。为解决光伏发电系统在发电阶段不能及时发现故障的技术缺陷,从而导致巨大的安全隐患以及经济损失,本文通过对 GA-BP 神经网络在光伏阵列故障检测中的应用的深入研究,提出一种新的混合遗传算法优化 BP 神经网络从而迅速且准确检测和区分故障类型,本文的主要研究内容包括:

第一章主要分析课题的研究背景和研究光伏阵列故障检测技术的意义,并在此基础上总结目前常见的光伏阵列故障检测方法的研究现状。

第二章分析光伏电池运行原理,依据光伏电池发电原理及等效电路图,推导电池典型

参数的数学表达式。在此基础上,结合实际光伏电厂的现场考察经验,在 Matlab/Simulink 中建立 4 行 3 列的光伏阵列,详细分析光伏阵列在正常与故障情况下的输出特性,并对故障产生的原因进行研究,最后总结出光伏阵列发生不同故障时代表参数的变化趋势。

第三章对 GA-BP 神经网络在光伏阵列故障检测的应用可行性进行研究。通过分析 BP 神经网络结构、遗传算法的工作原理,建立 GA-BP 神经网络模型,将采集到的光伏阵列在不同故障状态时的数据集经过归一化处理作为网络的输入变量,验证 GA-BP 神经网络用于光伏阵列故障检测的有效性,并将检测结果与传统的 BP 神经网络进行对比分析。

第四章提出一种引入鸽群算法的混合遗传算法优化误差反馈神经网络。分析基本鸽群算法的原理以及在其解决问题时的优缺点,总结经遗传算法优化后的 BP 神经网络仍然存在的缺陷。通过结合鸽群算法的优势实现对遗传算法初始染色体种群数量优化,从而减少种群迭代时间并解决遗传算法容易陷入“早熟”状态的问题,并给出混合遗传算法优化 BP 神经网络的具体步骤和流程。

第五章将经过混合遗传算法优化的 BP 神经网络应用于光伏阵列故障检测。对比分析遗传算法和混合遗传算法优选的初始权值阈值的适应度大小以及算法迭代次数上的差异,比较改进前后的 GA-BP 神经网络在进行光伏阵列故障检测过程中,训练时间、训练误差上优劣,并对老化和阴影遮挡故障单独进行检测实验,进一步体现改进前后的性能差异。

第六章主要是对本论文的研究内容进行归纳总结,对下一步的研究工作进行展望。



## 第二章 光伏阵列输出特性及故障分析

### 2.1 光伏电池发电特性研究

#### 2.1.1 光伏电池发电原理分析

光伏电池利用太阳光进行电能转化的实质是“光生伏特效应”，该效应是由法国科学家 A.E. 贝克勒尔于 1839 年发现。这种光能转换为电能的现象广泛存在于固体物质与流体物质中，不过在一般物质中光电转换效率极低，几乎可以忽略，但在半导体材料中，光电转换效率可以达到 25%。正是凭借这一优势，半导体材料几乎垄断了光伏电池的制造原材料的选择，目前使用的多晶硅太阳能电池的最大理论转换效率大约为 31%。2020 年 4 月 13 日，在美国的《自然能源》杂志上刊登了一篇关于美国国家可再生能源实验室（National Renewable Energy Laboratory, NREL）研发出了目前世界上最高效的太阳能电池的文章，这种电池的最高能量转换效率超过了 47%。尽管这种光电转换效率的记录经常被打破，但至少从目前情况来看，如果这种 47.1% 转换效率的太阳能电池可以应用于商业，那必然会使光伏产业迎来新一场变革。

光生伏特效应包括了两个过程，第一个过程是由太阳光所提供的辐射能给 PN 结内的电子带来动能，一旦这种能量超过了半导体的禁带宽度，就会促使电子定向运动从而形成电子-空穴对，第二个过程是在 PN 结内部结垒式电场力的影响下，电子从浓度高的 P 区移动向了浓度低的 N 区，空穴由浓度高的 N 区流入了浓度低的 P 区，最后产生了 P 区正电荷堆积，N 区负电荷堆积的现象，如果 PN 结两极在外电路之间用导线接通，就会产生电流流通的回路<sup>[31-32]</sup>，光伏电池工作原理如图 2.1 所示。

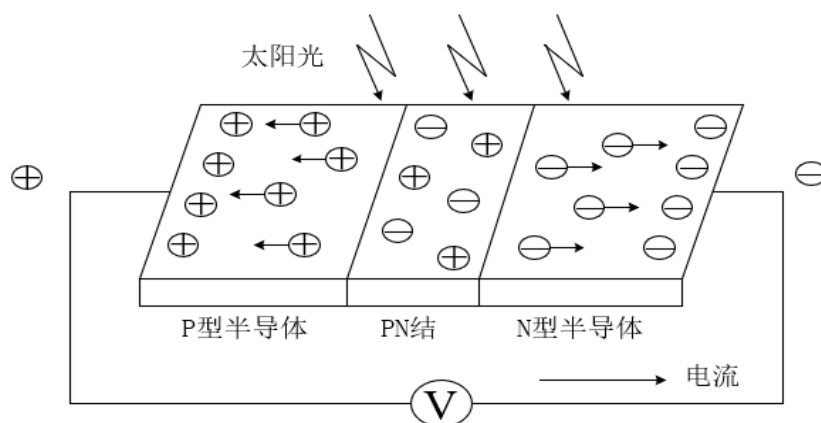


图 2.1 光伏电池工作原理

Fig. 2.1 The working principle of photovoltaic cells

### 2.1.2 光伏电池数学模型研究

在 Simulink 仿真软件中，比较常见的光伏电池等效模型有单二极管模型，双二极管模型，本文从经济型以及贴合实际性考虑：首先，在光伏阵列仿真程序中，单二极管模型精度已经达到所需条件，若使用双二极管模型，精度虽然会有提高，但会导致整个仿真过程收敛过慢。其次，由于目前在大部分的光伏电站中的光伏电池板都是采用单二极管并联，故选用单二极管模型光伏电池作为构成本文光伏阵列的基本单元。

光伏电池单体的单二极管模型，可以视作一个理想输出电流源与一个二极管并联组成，利用能量转换的形式产生电压，从而在闭合的外电路中形成电流。单二极管光伏电池的等效电路图如图 2.2 所示。图中，串联电阻  $R_s$  等效替代的是半导体材料的表电阻和两个电极间的固有电阻等，其电阻值非常小，一般情况下可以忽略。并联电阻  $R_{sh}$  为虚拟电阻，用来等效替代光伏电池本体的缺陷和漏电流泄露等原因造成的损耗<sup>[33]</sup>。

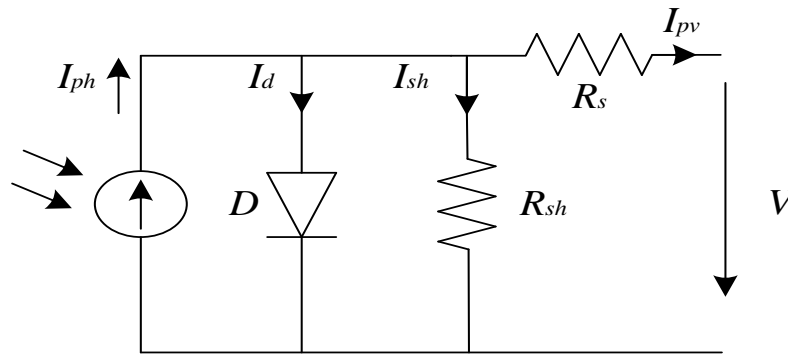


图 2.2 单二极管光伏电池等效电路图

Fig. 2.2 Equivalent Circuit Diagram of Single Diode Photovoltaic Cell

将图 2.2 结合基本电路原理分析，可以得知，当给予电池一定的光照和温度时，光伏电池所产生的电流满足以下关系<sup>[34-35]</sup>：

$$I_{pv} = I_{ph} - (I_d + I_{sh}) \quad (2.1)$$

其中， $I_{pv}$  为光伏电池输出电流； $I_{ph}$  为光伏电池经过能量转换所产生的初始电流； $I_d$  为流经并联二极管的分支电流； $I_{sh}$  为流经并联电阻的分支电流。

旁路二极管支路的分流电流  $I_d$  为：

$$I_d = I_0 \left( \exp \left[ \frac{q(V_{pv} + I_{pv} R_s)}{NkT} \right] - 1 \right) \quad (2.2)$$

其中， $I_0$  为并联二极管反向饱和电流； $q$  为电子电荷，其数值为  $1.6 \times 10^{-19} C$ ； $k$  为玻尔兹曼常数，其数值为  $1.38 \times 10^{-23} J/K$ ； $T$  为电池工作时的实际温度，单位为绝对温度， $25^\circ C$

时为 298K;  $N$  为二极管理想特性因子, 数值由二极管本身材料性质决定;  $R_s$  为光伏电池等效电路中串联电阻。

并联电阻  $R_{sh}$  与光伏电池电流电压的关系为:

$$V_{pv} = V_{sh} - I_{pv}R_s = I_{sh}R_{sh} - I_{pv}R_s \quad (2.3)$$

其中,  $V_{pv}$  为光伏电池输出电压;  $V_{sh}$  为并联旁漏电阻支路电压。

根据上述方程式, 总结单二极管光伏电池的电路等效关系可由公式 2.4 表示:

$$I_{pv} = I_{ph} - (I_d + I_{sh}) = I_{ph} - I_0 \left( \exp \left[ \frac{q(V_{pv} + I_{pv}R_s)}{NkT} \right] - 1 \right) - \frac{V_{pv} + I_{pv}R_s}{R_{sh}} \quad (2.4)$$

此外, 以下几个参数的变化情况可以描述光伏电池不同状态下的工作特性:

开路电压  $V_{oc}$ : 开路电压是在正常工作条件下, 当太阳能电池输出电流  $I_{pv}$  为零时, 输出的电压值。由式(2.4)和图 2.2 可知, 当输出电流  $I_{pv}$  为零时, 以下两式成立:

$$I_{ph} - (I_d + I_{sh}) = I_{ph} - I_0 \left( \exp \left[ \frac{qV_{oc}}{NkT} \right] - 1 \right) - \frac{V_{oc}}{R_{sh}} \quad (2.5)$$

$$V_{oc} = I_{sh}R_{sh} \quad (2.6)$$

短路电流  $I_{sc}$ : 短路电流是在正常工作条件下, 当光伏电池输出电压  $V_{pv}$  为零时, 输出的电流值。 $I_{sc}$  数值大小与光伏电池面积有着正相关的关系, 面积越大, 载流能力越强, 短路电流越大。

由公式(2.4)可知, 当  $V_{pv}$  为零时, 输出电流  $I_{pv}$  为:

$$I_{pv} = I_{ph} - (I_d + I_{sh}) = I_{ph} - I_0 \left( \exp \left[ \frac{I_{sc}R_s q}{NkT} \right] - 1 \right) - \frac{R_s}{R_{sh}} \quad (2.7)$$

当光伏电池工作在实际温度  $T_j$  时, 其短路电流  $I_{sc}$  为:

$$I_{sc} = I_{scr} - (1 + a_{sc}\Delta T) \quad (2.8)$$

$$\Delta T = T - T_{jr} \quad (2.9)$$

其中,  $I_{scr}$  为参考短路电流;  $a_{sc}$  为短路电流的温度系数;  $T_{jr}$  为参考环境温度, 一般取 25°C。

短路电流与光照强度  $S$  之间的关系由式(2.10)给出:

$$I_{sc} = I_{scr} \left( \frac{S}{S_r} \right) \quad (2.10)$$

其中,  $S_r$  为参考光照强度, 一般取  $1000\text{w/m}^2$ 。

从式(2.10)中可以看出, 光伏电池的短路电流与光照强度有着正比关系, 即短路电流会随着光照的增强而升高。

光伏阵列的最大功率点处的电压和电流也是能够反映其实际工作状况的典型参数。光伏电池在环境温度以及日光照射率均给定的情况下, 在正常无故障工作时, 其功率输出特性曲线是一条只有一个凸点的凸型曲线, 曲线最高的点为最大功率点  $P_m$ , 该点对应的电压为最大功率点电压  $V_{mpp}$ , 该点对应的电流为最大功率点电流  $I_{mpp}$ 。

### 2.1.3 光伏电池输出特性分析

本文在仿真软件中选用的光伏电池类别为 Zytech Engineering Technology ZT180S (下文简称 ZT180S 型), 在标准运行条件下 (电池所受太阳光辐射强度  $S$  为  $1000\text{W/m}^2$ 、电池工作环境温度  $T$  为  $25^\circ\text{C}$ ) 的有关技术参数如表 2.1 所示, 其内部电路如图 2.3 所示。

表 2.1 ZT180S 光伏电池模块参数

Tab. 2.1 Technical parameters of the ZT180S photovoltaic cell module

光伏电池参数	数值
开路电压 $V_{oc}$ (V)	44.71
短路电流 $I_{sc}$ (A)	5.21
开路电压温度系数	-0.36399
短路电流温度系数	0.102
最大功率点电压 $V_{mpp}$ (V)	37.58
最大功率点电流 $I_{mpp}$ (A)	4.79
串联等效电阻 $R_s$ ( $\Omega$ )	0.393
并联等效电阻 $R_{sh}$ ( $\Omega$ )	313.399

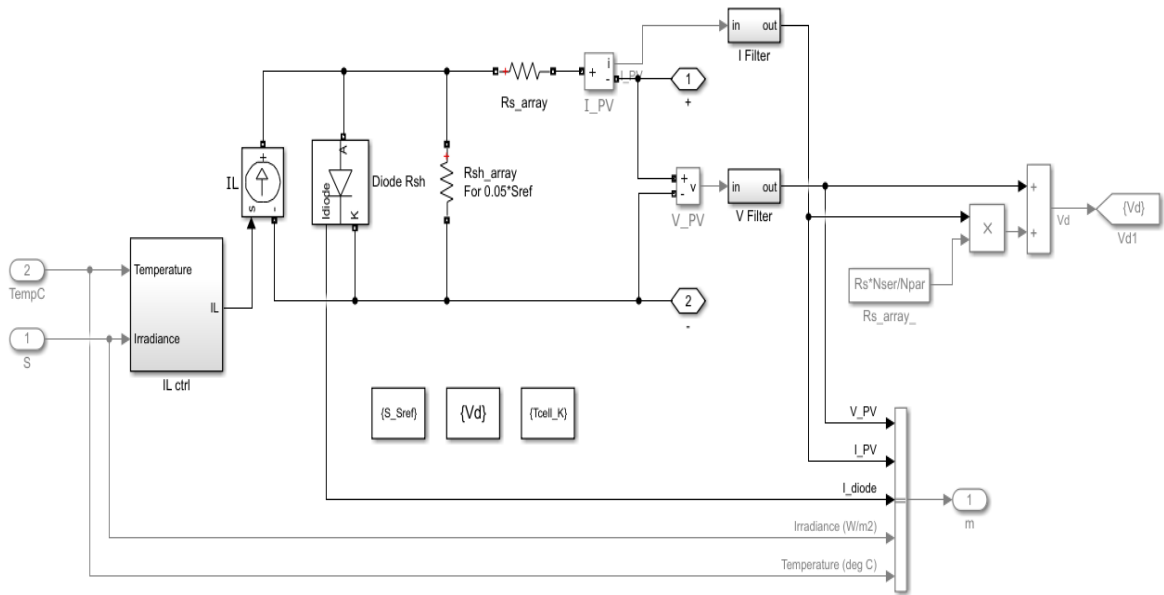
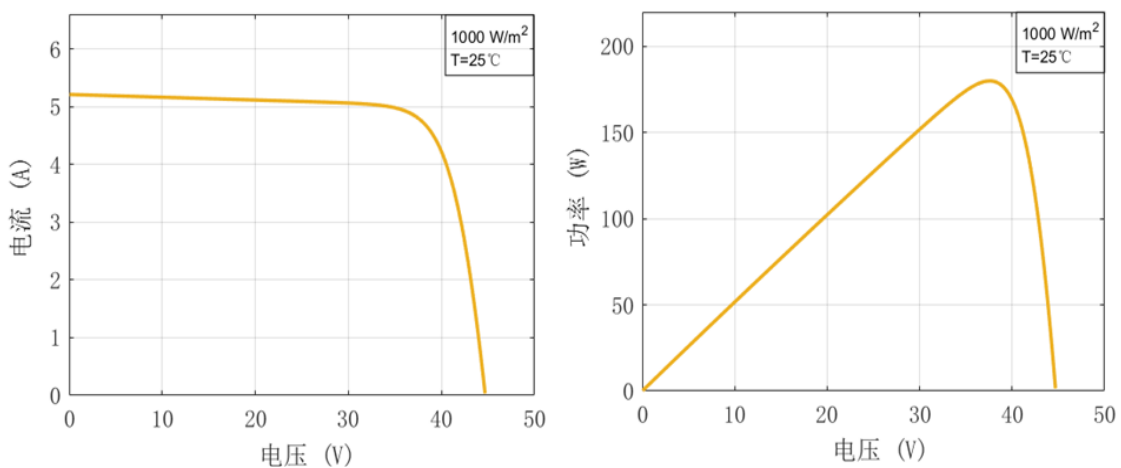


图 2.3 光伏电池内部电路图

Fig. 2.3 Internal circuit of photovoltaic cell

ZT180S 光伏电池在标准条件下的 I-V 输出特性曲线以及 P-V 输出特性曲线如图 2.4(a)、图 2.4(b)所示。在 ZT180S 光伏电池的 I-V 输出特性曲线中，电压为零时所对应的电流值为其短路电流大小，电流为零时所对应的电压值为其开路电压大小，而实际中光伏阵列中存在很多个电流电压传感器，利用这些传感器可以测定光伏电池的开路电压及短路电流大小。而 ZT180S 光伏电池的 P-V 曲线则是一条前段上升末尾下落的凸型曲线，该曲线的凸点为光伏电池的最大功率点，实际中通常是利用光伏最大功率点追踪技术得到光伏发电的功率曲线，完成对最大功率点处的电压电流值的测量。



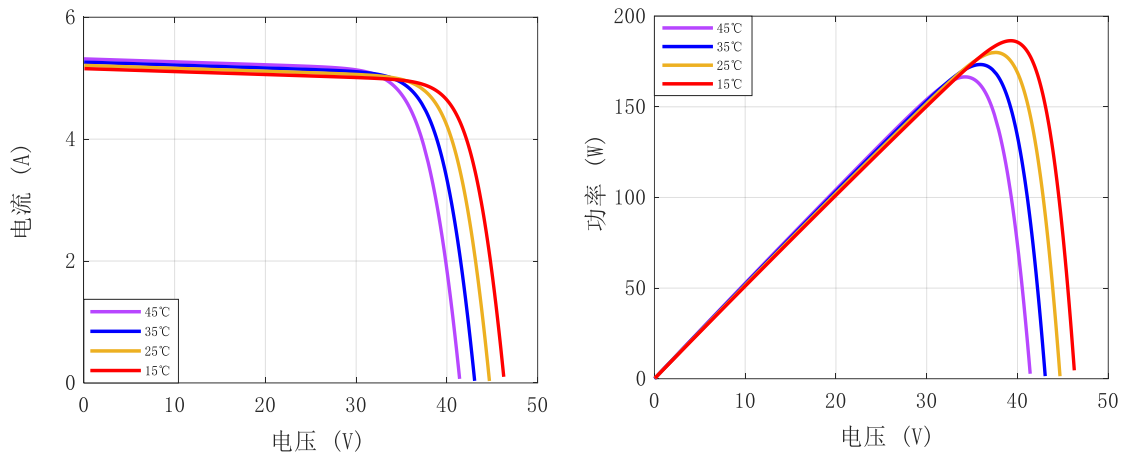
(a) 电池 I-V 曲线

(b) 电池 P-V 曲线

图 2.4 ZT180S 光伏电池 I-V 及 P-V 输出特性曲线

Fig. 2.4 I-V and P-V output characteristic curves of ZT180S photovoltaic cell

图 2.5 是 ZT180S 光伏电池单体, 在保持光照强度  $S=1000\text{w}/\text{m}^2$  条件下, 通过调整工作温度后得到的输出特性, 综合分析图 2.5(a)、(b), 可以得知: 随着温度的升高, 其开路电压降低程度较大, 短路电流则有着较小上升幅度, 最大功率点的输出功率略有减少, 因此, 可以认为其输出功率及开路电压与温度有着负相关的关系, 其短路电流与温度有着正相关的关系。



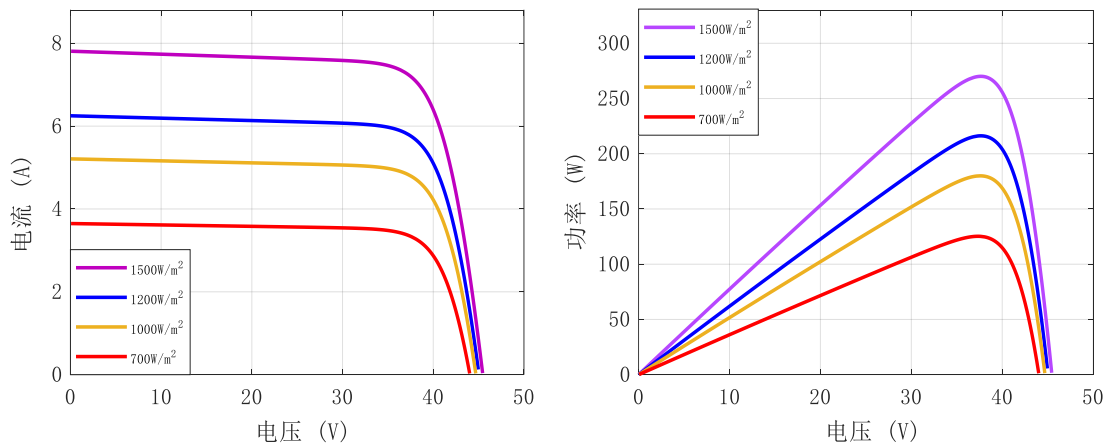
(a)不同温度下电池 I-V 曲线

(b)不同温度下电池 P-V 曲线

图 2.5 不同温度下光伏电池输出特性

Fig. 2.5 Output characteristics of photovoltaic cells at different temperature

图 2.6 中可以看出, 在保持环境温度  $T=25^\circ\text{C}$  不变时, 调整施加于光伏电池的光照强度, ZT180S 光伏电池的输出功率受到的影响较大。图 2.6(a)中的曲线说明了其开路电压随着光照强度的加强, 以较小幅度升高, 而其短路电流以较大的幅度升高, 图 2.6(b)中曲线的随着光照强度的增加, 电池最大功率点上移, 输出功率显著增加, 由此, 可以认为其输出功率同光伏电池的开路电压与短路电流光照强度有着正相关的关系。



(a)不同光照下电池 I-V 曲线

(b)不同光照下电池 P-V 曲线

图 2.6 不同光照下光伏电池输出特性

Fig. 2.6 Output characteristics of photovoltaic cells at different illumination intensity

## 2.2 光伏阵列连接结构及输出特性分析

### 2.2.1 光伏阵列的连接结构

光伏阵列是由若干个光伏电池单体经由导线连接，为了获得更大的输出功率而组成的一个大型光伏电池群，光伏电池之间最基本的连接方式是串联连接和并联连接，由光伏电池串联组成的光伏阵列，其通过的短路电流与每一个电池的最大短路电流相同，而与之相对的并联连接方式，阵列开路电压与电池单体相同，从这一点就可以看出，这两种连接方式非常简单，但是对于提高输出功率的作用十分有限，而且一旦发生故障，很容易造成整个阵列的崩溃，因此，在目前的光伏电厂中几乎看不到光伏电池的这两种连接方式。

串并联结构（Series-Parallel structure, SP）是现在大部分光伏发电厂所使用的一种连接结构<sup>[36]</sup>，这种结构是在串联与并联方式有机结合的基础上发展而来的，兼顾了串联和并联连接的优点，阵列一旦出现阴影遮挡或者开路故障，即使故障组件输出电流会减小，但是对太阳能发电系统整体的输出功率来说，其产生的影响要远远小于单独串联或是并联连接，大大缩小了故障范围。而相比于其他复杂连接结构，采用这种连接方式的光伏阵列的优势在于所需导线少，成本比较低，整体比较稳定，SP 结构的光伏阵列连接方式如图 2.7 所示。

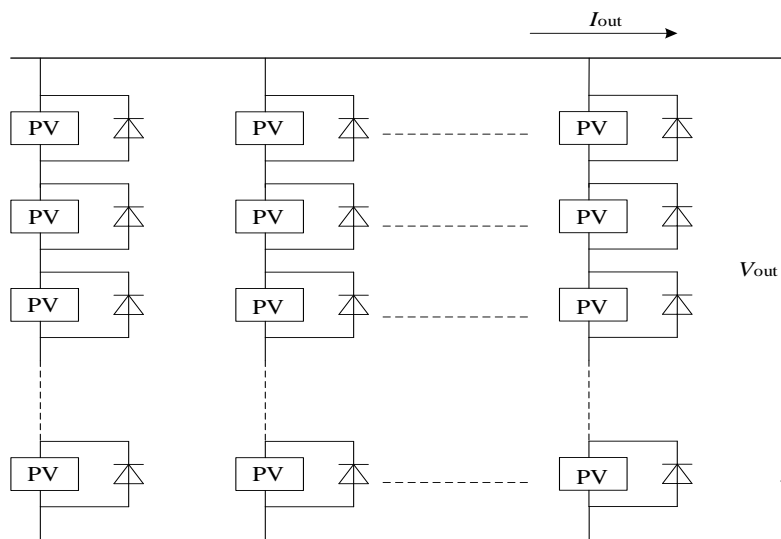


图 2.7 串并联连接结构

Fig. 2.7 Series-parallel structure

全连接结构（Total Cross Tied structure, TCT）一般见于比较重要的光伏电站中，其阵列连接结构如图 2.8 所示，当使用 TCT 结构进行光伏电池的连接时，由于各个电池单体之间的联系更加密切，当光伏阵列中的某个组件发生故障时，故障产生的影响会远弱于 SP 连接结构，阵列整体表现出接近于无故障发生时的输出特性，这一特点让 TCT 结构比 SP 结构更具可靠性<sup>[37]</sup>。TCT 结构的弊端也非常突出，一方面在于增加了对导线及电压电

流传感器的需求量，给设计阵列内部的走线带来了困难。另一方面是当电能在线路传递输送时，所产生的不必要消耗也急剧上升，同时，在阵列中的少量光伏电池发生故障或者损毁时，故障点的精确定位和快速排除也十分困难，对于综合考虑发电成本的光伏电厂而言并不经济。

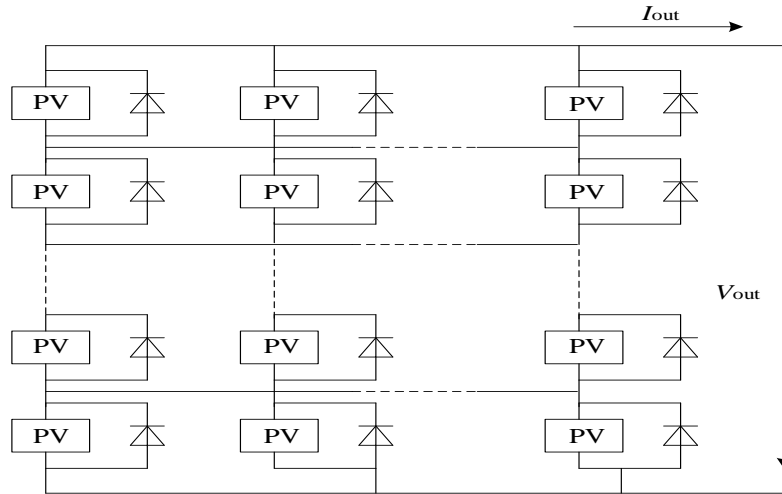


图 2.8 全连接结构

Fig. 2.8 Total Cross Tied structure

桥型连接结构（Bridge-Linked structure, BL）是指通过搭桥连接相邻的光伏组串，桥的数量一般不超过一串光伏电池数量的一半，目前来说应用的场合最少，一般是在实验室中分析单独分析阴影故障时采用这种连接方式，其连接方式如图 2.9。这种连接结构在 SP 连接结构上发展而来，主要是为了减少阵列故障引起的失稳影响，其连接复杂程度较 TCT 结构简单，在一定程度上提高了故障发生时传感器故障定位装置的灵敏性，但是对于大型光伏电站来说，仍是不够经济，这也是 BL 结构未能在实际中推广的重要原因。

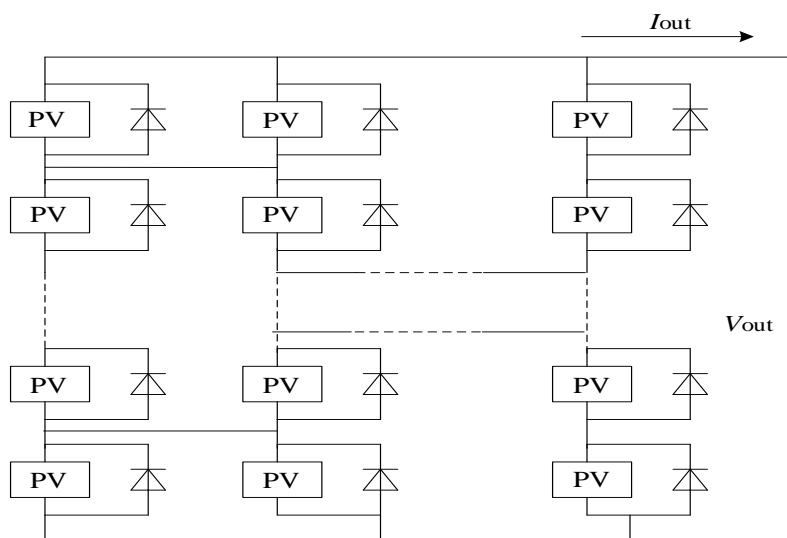


图 2.9 桥型连接结构

Fig. 2.9 Bridge-Linked structure



### 2.2.2 光伏阵列正常运行时输出特性

为了使本文所做研究更加具有实际意义，本文对淄博新明珠光伏发电中心进行了实地考察，参照了该光伏发电中心如图 2.10 所示的阵列连接方式，但是实际中其光伏电池单体非常多，在实验室中很难全部模拟，同时为了更好的模拟同一故障发生的不同方式，最终在仿真软件中搭建了如图 2.11 所示的 4 行 3 列的光伏阵列用于进行实验。



图 2.10 淄博新明珠光伏发电中心

Fig. 2.10 Zibo New Pearl Photovoltaic Power Generation Center

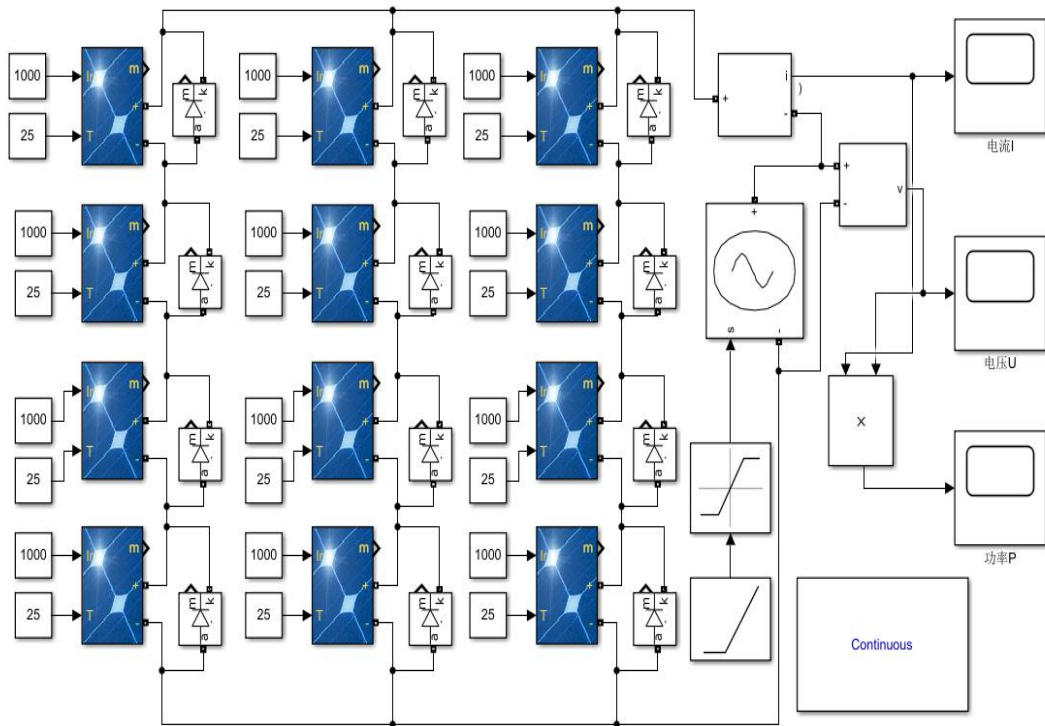


图 2.11 4×3 光伏阵列仿真模型

Fig. 2.11 4×3 Simulation model of PV array

光伏阵列正常运行时，其输出曲线如图 2.12 所示。理想状况下，阵列最大的输出功率是全部电池个体的输出功率之和，其中  $P_m: 12 \times 180.9 = 2160.2\text{W}$ ，滤除掉电池个体内部损耗以及阵列器件损耗，该阵列最大功率维持在 2010W 左右，短路电流为所有列电流之和，其中  $I_{sc} = 5.21 \times 3 = 15.63\text{A}$ ，开路电压为所有行输出电压之和，其中  $V_{oc} = 44.71 \times 4 = 178.84\text{V}$ 。

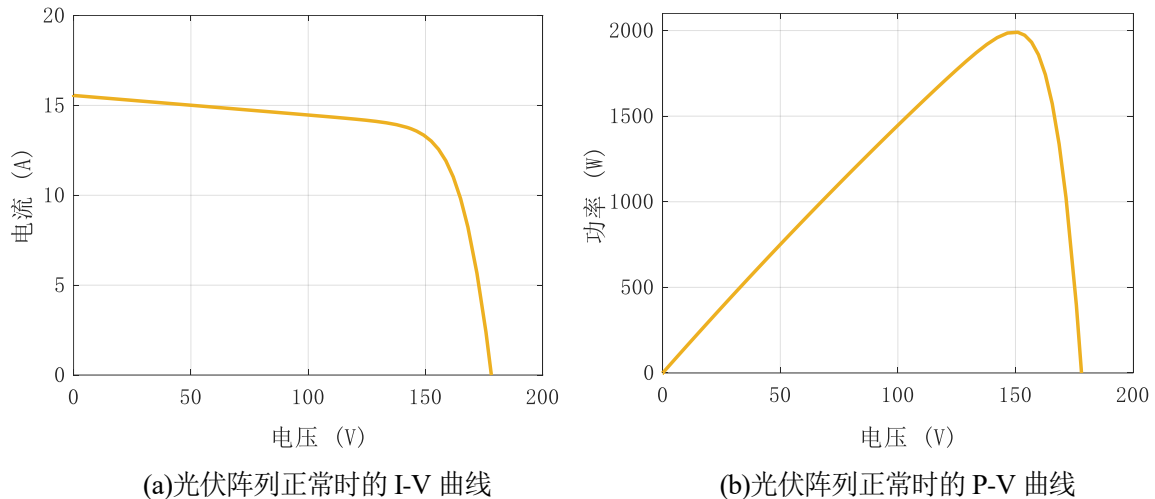


图 2.12 光伏阵列在正常情况下的输出特性

Fig. 2.12 Output characteristics of photovoltaic arrays under normal conditions

## 2.3 光伏阵列不同故障状态下输出特性分析

### 2.3.1 光伏阵列故障形成原因分析

在实际光伏电站中出现的故障多为老化，遮阳，开路，短路四种故障。

短路故障是光伏阵列发生次数最多的一种故障。光伏阵列一般暴露在条件比较恶劣的环境中，随着雨水的侵蚀或是风霜的冲刷，都有可能造成光伏组件表皮绝缘层损坏，导致线路正负极直接相接引起短路，短路故障的危害较大，易造成整个阵列的功率缺失，如果短路故障存在时间较长，则会不断的产生电弧并释放高额热量，引发光伏电站火灾，导致无法接受的后果。

开路故障也是光伏电站发生较为习见的一种故障形式，故障形成的因素大致有两种：一方面可能是在电池生产时，工艺把控不合格，在内部产生了损伤；另一方面，在作业人员进行光伏阵列架设或检修时，由于误操作而造成光伏电池内部接线断裂或是组串接线箱中接点不正常分离，使得光伏阵列在工作时故障组件与相邻组件之间不再存在电气联系，也由此导致故障所在串不能参与发电，造成整个阵列输出功率缺失。

阴影遮挡是各个光伏电站都会面临的一种特殊故障现象，而此类故障的出现，随机性占比大，例如建筑物、大树、云朵等的影子遮盖了光伏电站，又或者是鸟类粪便固化在阵

列表面，而长期的阴影遮挡也会造成被遮蔽组件无法正常散热，产生热斑现象，并在此基础上会烧坏线路导致短路或者开路故障的发生，对光伏电站的安全稳定运行产生不利影响。上述几种引起遮阴故障的原因中，除却云朵遮挡时间较短，且发生概率比较随机，暂时没有较好的预防方法，对于其他原因造成的遮阴现象，应该在平时加强对光伏电站的维护和清理，谨防遮阴现象的持续存在。

老化故障指的光伏组件随着运行时间的增加，由于制造工艺的局限，其外壳或者是内部电路发生脆化甚至损坏，一方面会造成整个阵列的输出功率的降低，另一方面加重了光伏电厂的安全隐患。正常来讲，现在国内各大光伏制造厂商承诺的光伏电池使用寿命大约在 25 年左右，但是光伏阵列受到环境的影响比较大，像是太阳光长时间的照射，恶劣天气引起阵列受到腐蚀都会致使光伏阵列的使用寿命大大缩减。

### 2.3.2 光伏阵列不同故障状态下输出特性分析

本文在前一小节已经分析并得出了光伏阵列在正常运行状态时的输出特性曲线，本节将使用前文搭建的 4 行 3 列光伏阵列仿真模型，在标准运行工况下进行开路、短路、阴影遮挡、老化四种故障的模拟，为方便描述故障具体情况，本节将原仿真模型进行简化，简化后的故障点设置如图 2.13 所示，并在此基础上，获得光伏阵列在不同故障状态下的输出特性曲线，并分析了各种故障状况下阵列参数的变化情况。

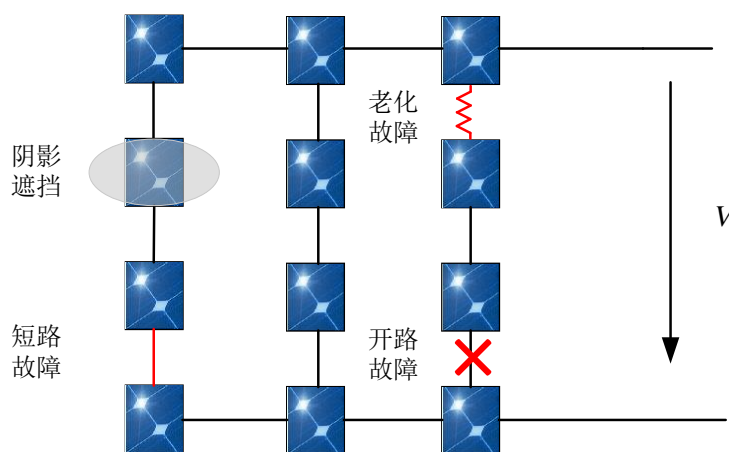


图 2.13 光伏阵列模拟故障设置

Fig. 2.13 PV Array Simulation Fault Settings

在光伏阵列仿真模型搭建完成之后，将光伏电池的正负极进行短接来模拟短路故障；开路故障通过切断故障电池与相邻光伏电池的连接导线来模拟；通过提高光伏电池的串联电阻来模拟老化故障<sup>[38]</sup>；光伏阵列被阴影遮挡则可以通过调整电池所受光照的强弱来模拟。考虑到在实际的光伏电厂中，故障出现位置的随机性与复杂性，比如遮阴故障发生时，由光伏电池上面的污秽物引起的遮阴故障一般只会影响个别片光伏电池，故设置 C1 类故障，

对于大面积树叶或云朵遮盖产生的阴影遮挡往往是一片较大面积的光伏电池受到影响，故设置 C2 类故障。同样，考虑某一串中的电池受环境影响，老化进程大致相同，故设置 D1 类故障，对于随年份增长而自然缩减寿命的光伏阵列，通过设置 D2 类故障进行模拟。具体故障类型设置如表 2.2 所示。

表 2.2 光伏阵列故障类型编号及说明

Tab. 2.2 PV array fault type number and description

故障编号	故障类型	故障内容
A1	短路故障	一个光伏电池短路
A2	短路故障	两个光伏电池短路
B1	开路故障	一串光伏电池开路
B2	开路故障	两串光伏电池开路
C1	阴影遮挡	同一串中一个光伏电池被遮挡
C2	阴影遮挡	同一串所有光伏电池被遮挡
D1	老化故障	一串光伏电池老化
D2	老化故障	两串光伏电池老化

故障编号 A1, A2 与正常状态时的输出特性曲线如图 2.14 所示。通过分析图 2.14(a)、2.14(b)可知，短路故障出现时，阵列的短路电流和最大功率点处电流变化并不显著，而阵列最大输出功率明显降低，阵列开路电压以及最大功率点电压数值下降明显，同时光伏阵列的最大输出功率会随着故障电池数量的增加，下降趋势更为显著。

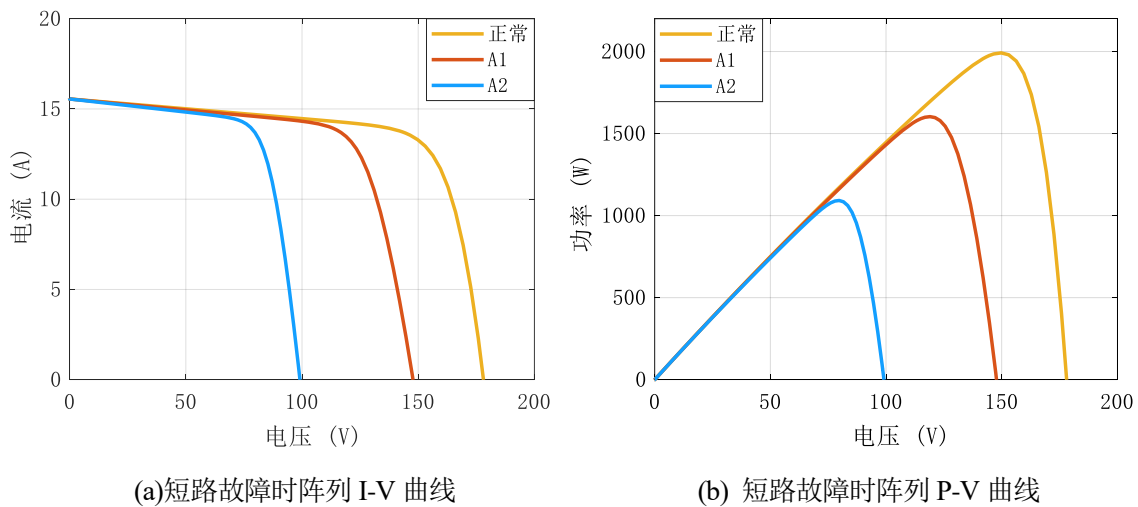


图 2.14 阵列发生短路故障时的输出特性曲线

Fig. 2.14 Array output characteristics at short-circuit fault

故障编号 B1, B2 与正常状态时的 I-V 曲线及 P-V 曲线如图 2.15(a)、(b)所示。当阵列中存在开路故障时, 该阵列的开路电压和最大功率点电压基本保持恒定, 阵列短路电流以及最大功率点处电流明显下降, 而且阵列的最大输出功率会随着开路组串数量的增加而下降。

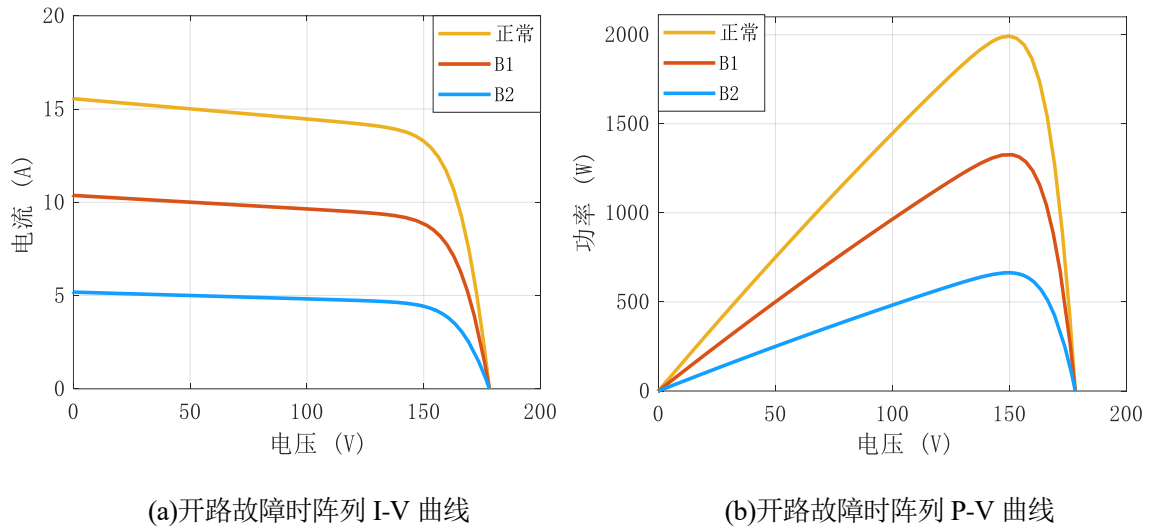


图 2.15 阵列发生开路故障时的输出特性

Fig. 2.15 Array Output Characteristics at Open Fault

故障编号 C1, C2 与正常状态时的输出特性曲线如图 2.16 所示。当光伏阵列中存在阴影遮挡现象时, 阵列的开路电压以及短路电流变化非常微弱, 在实际应用中可以忽略, 而阵列的 P-V 曲线以及 I-V 曲线则呈现多凸点的特征, 并且随着遮挡组串数量的增加, 其凸点个数也在不断增加, 故障时最大供电功率点处的最大功率, 最大电流以及最大电压均呈现非常明显的下降趋势且电流下降趋势显著大于电压的下降趋势。

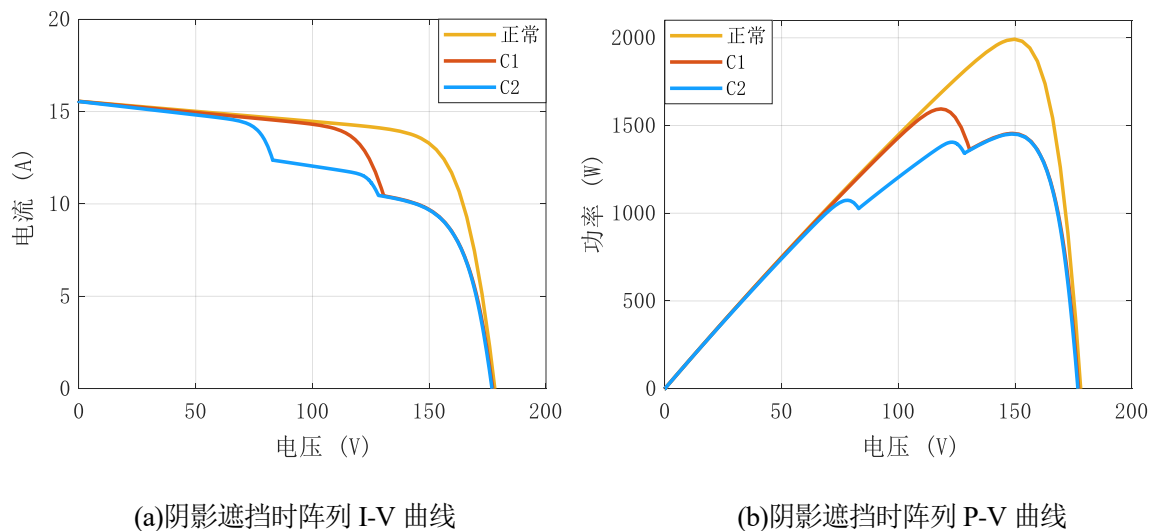


图 2.16 阴影遮挡时阵列的输出特性

Fig. 2.16 Array output characteristics with shadow occlusion

故障编号 D1, D2 与正常状态时的输出特性曲线如图 2.17 所示, 本文通过为故障组件串联  $5\Omega$  的电阻模仿每串光伏电池的老化现象<sup>[39]</sup>, 分析图 2.17(a)可知, 阵列中电池使用寿命即将到达终点, 出现老化现象时, 整个阵列的短路电流略微下降, 而开路电压一直维持不变, 从图 2.17(b)中可以看出, 阵列的输出功率随着老化电池数量的增加而明显下降, 而随着老化组串数的增多, 阵列最大功率点处的电压电流均有降低, 但下降程度略小于阵列发生阴影遮挡故障时, 且电流的下降趋势小于电压的下降趋势。

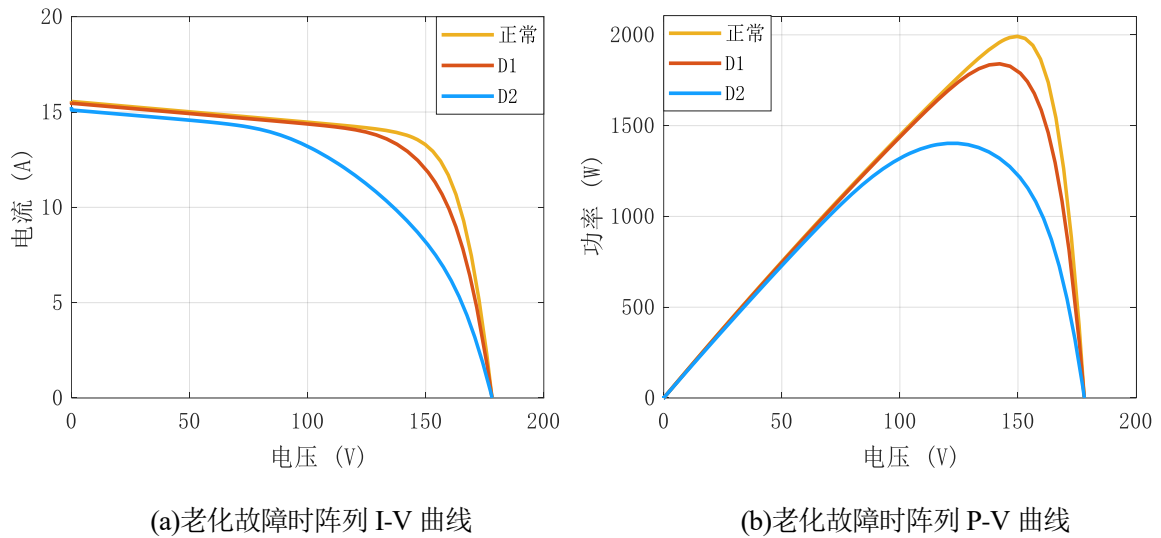


图 2.17 光伏电池老化时阵列输出特性

Fig. 2.17 Array output characteristics as photovoltaic cells age

通过以上对四种典型故障发生时阵列输出特性的分析可以发现, 当阵列中存在四种典型故障时, 阵列总是会有一个或者两个特征参数产生较为明显的变化, 据此, 将各个参数的变化趋势总结如表 2.3 所示。

表 2.3 不同故障类型时各参数变化趋势

Tab. 2.3 Variation trend of parameters under different fault types

故障编号	故障类型	开路电压	短路电流	最大输出	最大功率	最大功率
				功率	点处电压	点处电流
A1, A2	短路故障	下降	不变	下降	下降	不变
B1, B2	开路故障	不变	下降	下降	不变	下降
C1, C2	阴影遮挡	不变	不变	下降	下降	下降
D1, D2	老化故障	不变	下降	下降	下降	下降

## 2.4 本章小结

本章通过研究光伏电池的工作原理以及数学模型，分析光伏电池在不同连接方式下搭建的光伏阵列阵优缺点，经过对淄博新明珠光伏发电中心的实际考虑，在 MATLAB Simulink 程序中顺利搭建了 4 行 3 列的串并联形式的光伏阵列仿真模型，并通过设置各项故障条件，模拟生活中光伏阵列可能发生的典型故障，详细分析论证了光伏阵列在正常运行状态以及发生故障时的输出特性曲线变化趋势，及以阵列正常运行条件下为参考时的光伏阵列中不同参数在不同故障类型下的变化特征。

### 第三章 GA-BP 神经网络应用于光伏阵列故障检测研究

#### 3.1 神经网络分析

##### 3.1.1 神经网络

神经网络 (Artificial Neural Network, ANN) 是由美国心理学家 William James 教授于 1980 年, 在研究人体神经系统工作原理的基础上, 提出的一种模拟人脑功能的复杂数学计算系统。相对于生物大脑信息处理的最基本单元神经元而言, 神经网络的最基本处理单元称为节点, 网络由众多包含节点的层组成, 每一层之间的连接类似人脑的不同神经元之间的突触, 其模型如图 3.1 所示。从图中可以看出, 在 ANN 中, 相对于神经元中的树突和轴称之为边, 这些边的权重大小是通过神经网络输入层和隐藏层以及输出层执行某些特定的数学计算不断调整的<sup>[40]</sup>。信息处理层通过树突 (边) 接收来自输入层的信息, 这些信息通过不同的权重占比, 将信息输入到细胞体 (处理层), 再经过激活函数的筛选进行往下一层的传递, 这样就完成了一次信息的传递, 而在一个完整的神经网络中, 成千上万的神经元相互配合、协同工作, 进而实现了神经系统的基本功能。

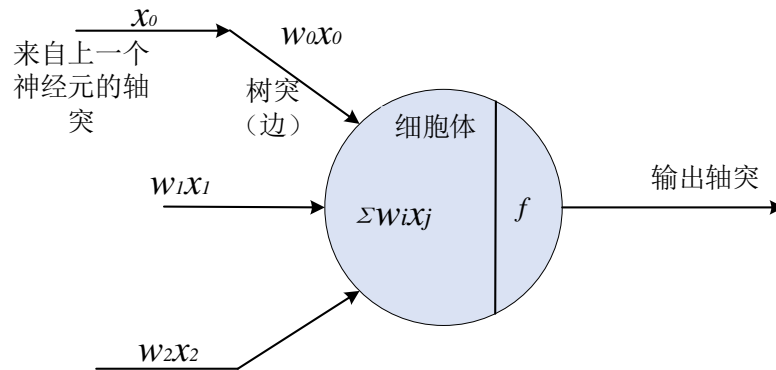


图 3.1 神经网络模型

Fig. 3.1 Artificial neural network model

##### 3.1.2 神经网络特点分析

神经网络实现了相对简单的数学表达式的结构化组合, 以产生更灵活和更具描述性的表达式, 以描述特定过程中的输入和输出关系。克服了传统问题分析时受限于计量复杂的缺点, 极大的推动了人工智能和信息处理技术的发展。总结 ANN 自一出世便受人青睐得以快速发展的原因, 以下几点可以说明:



第一, 具有强大自学习功能<sup>[41]</sup>。例如进行故障的检测识别时, 只需要把历史故障数据交由人工神经网络, 网络就会自我学习、自我训练, 慢慢学会识别类似的故障, 网络学习完毕后, 只需要输入想要判别的设备运行数据, 就能得到是否故障的结论。自学习功能是人工神经网络的灵魂所在, 而且现在的人工神经网络也凭借这一能力已经在自适应控制、市场调度等方面颇有建树。

第二, 具有充分逼近任意复杂的非线性问题的能力。人工神经网络不仅能够迅速解决线性函数问题, 并且在进行非线性问题的求解时, 还能够利用自身处理系统的连续迭代计算无限的逼近非线性曲线, 这一特点正好与光伏阵列的非线性输出特性相契合, 因此非常适合用来检测光伏阵列出现的故障。

第三, 高速寻优能力出众。当寻求某一复杂问题的优化解时, 如果仅仅依靠人力求解, 往往需要花费大量的时间来进行繁琐的计算, 而使用一种专门根据某问题所设计的人工神经网络, 利用现代计算机的极速浮点运算能力, 就能够在很短的时间内找到最佳处理结果。

人工神经网络的这些优点引发了国内外学者浓厚的研究兴致, 在经过几十年的研究发展, 截止到现在, 已经提出的神经网络有自组织映射网络、振荡神经网络、BP 神经网络、前向神经网络、结合神经网络以及卷积神经网络等上百种网络模型, 在这之中, 当属 BP 神经网络研究热度最高, 应用范围最广。

## 3.2 BP 神经网络结构及工作原理分析

### 3.2.1 BP 神经网络的结构分析

BP 神经网络的全称是误差反馈传递型神经网络(Back Propagation Neural Network, BPNN)是一种利用误差反向传播学习算法形成的一种人工神经网络<sup>[42]</sup>, BP 神经网络是迄今为止人工神经网络发展中最为完善的分支, 其延续了 ANN 的良好的学习能力以及优异的数学特性, 其使用场景十分广阔, 通过查阅文献, 可以发现如今的 BP 神经网络已经广泛的应用于数据预测, 模式识别, 故障分类等方面的问题<sup>[43-45]</sup>。

BP 神经网络在架构组成上与传统 ANN 结构十分相似, 主要由输入层, 隐含层以及输出层构成。输入层的主要任务是接收外部信息, 并将信息传输到隐含层中; 隐含层的主要任务是接收来自输入层的信息, 并经过激活函数的处理后将信息传递到输出层; 输出层的主要工作是将从隐含层得到的信息与期望输出做比较, 如果两组数据满足一定的误差条件, 则输出信号; 如果不满足误差条件, 则会往前传递反馈信号。

作为一种多层映射网络, 对于大部分的信息处理问题, BP 神经网络的输入输出层层数一般设置为 1 层, 而隐含层需要根据数据的复杂程度以及处理精度要求, 在 1 到 3 层之间选取, 一般来说, 层数越多, 就越能发挥网络的自学习能力, 但是多隐含层面临的问题其一在于网络训练时间一般会呈现几何增长的态势, 二是对于训练算法的要求也更为严苛,

采用传统的梯度下降法很难完成多隐含层的误差要求,图 3.2 为单隐含层的 BP 神经网络简要结构图。

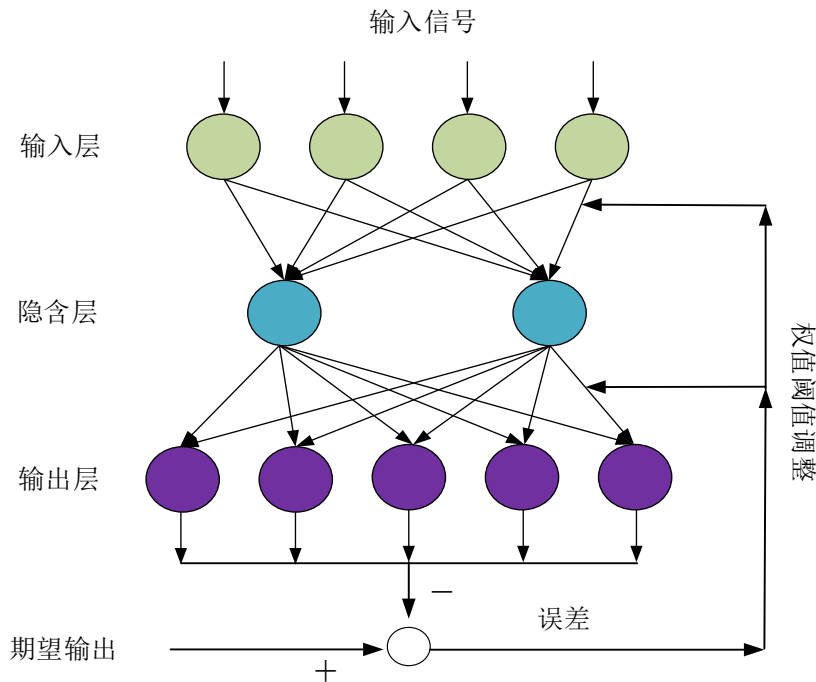


图 3.2 BP 神经网络结构

Fig. 3.2 Structure of BP neural network

### 3.2.2 BP 神经网络工作原理分析

BP 神经网络工作过程可分为信息正向传播过程与误差反向传播过程<sup>[46]</sup>,各过程工作原理如下。

#### (1) 信息正向传播过程

在一个输入层有  $m$  个节点,隐含层有  $q$  个节点,输出层有  $n$  个节点的 BP 神经网络中,输入信号  $X_i$  在正向传播过程中,网络的权值保持不变,经过隐含层的处理之后,在输出层得到实际输出  $Y$ ,并通过误差函数计算  $Y$  与期望输出  $\tilde{Y}$  间的误差,若误差过大则转入反向传播过程。

设  $F(x)$  为隐含层激励函数,此时先以 Sigmoid 函数为例,函数公式为(3.1),则神经网络正向传播时过程如下:

$$F(x) = \frac{1}{(1 + e^{-ax})} \tag{3.1}$$

设  $h_j$  为隐含层第  $j$  个节点的输入值,则有:

$$h_j = \sum_{i=1}^m w_{ij}x_i - \gamma_j \quad (3.2)$$

其中,  $w_{ij}$  为输入层第  $i$  个节点与隐含层第  $j$  个节点的连接权值;  $\gamma_j$  为隐含层第  $j$  个节点的阈值。

设  $H_j$  为隐含层第  $j$  个节点的输出值, 则有:

$$H_j = F\left(\sum_{i=1}^m w_{ij}x_i - \gamma_j\right) \quad (3.3)$$

设  $Y_k$  为输出层第  $k$  个节点的实际输出值, 则有:

$$Y_k = \sum_{j=1}^n H_j w_{jk} - \theta_k, k = 1, 2, \dots, n \quad (3.4)$$

其中,  $w_{jk}$  为隐含层第  $j$  个节点与输出层第  $k$  个节点的连接权值;  $\theta_k$  为输出层第  $k$  个节点的阈值。

### (2) 误差反向传播过程

误差信号以某种方式从输出层开始向前传播, 网络的权值和阈值会随着误差信号的反馈不断进行调整, 直至实际输出与期望输出保持在可接受的误差范围内, 这是 BP 神经网络的反向工作过程。

假设样本数量为  $C$ , 且第  $p$  个样本在第  $k$  个节点的期望输出为  $T_k^p$ , 那么根据误差准则函数可得到  $C$  个样本数据的误差为:

$$e_k = \frac{1}{2} \sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p)^2 \quad (3.5)$$

依据 BP 神经网络采用梯度下降法进行修正可知:

$$\Delta w_{ij} = -\eta \frac{\partial e}{\partial w_{ij}} \quad (3.6)$$

$$\Delta \gamma_j = -\eta \frac{\partial e}{\partial \gamma_j} \quad (3.7)$$

$$\Delta w_{jk} = -\eta \frac{\partial e}{\partial w_{jk}} \quad (3.8)$$

$$\Delta \theta_k = -\eta \frac{\partial e}{\partial \theta_k} \quad (3.9)$$

联立以上四个公式, 可以获得网络隐含层权值变化量为:

$$\Delta w_{ij} = -\eta \frac{\partial e}{\partial w_{ij}} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial w_{ij}} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial Y_k} \frac{\partial Y_k}{\partial w_{ij}} \quad (3.10)$$

网络隐含层阈值变化量为:

$$\Delta\gamma_j = -\eta \frac{\partial e}{\partial \gamma_j} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial \gamma_j} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial Y_k} \frac{\partial Y_k}{\partial \gamma_j} \quad (3.11)$$

网络输出层连接权值变化量:

$$\Delta w_{jk} = -\eta \frac{\partial e}{\partial w_{jk}} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial w_{jk}} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial Y_k} \frac{\partial Y_k}{\partial w_{jk}} \quad (3.12)$$

网络输出连接阈值的变化量为:

$$\Delta\theta_k = -\eta \frac{\partial e}{\partial \theta_k} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial \theta_k} = -\eta \frac{\partial e}{\partial h_j} \frac{\partial h_j}{\partial Y_k} \frac{\partial Y_k}{\partial \theta_k} \quad (3.13)$$

又因为:

$$\eta \frac{\partial e}{\partial h_j} = -\sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) \quad (3.14)$$

$$\frac{\partial h_j}{\partial w_{ij}} = y_j; \quad \frac{\partial h_j}{\partial \gamma_j} = 1; \quad \frac{\partial h_j}{\partial w_{jk}} = x_k; \quad \frac{\partial h_j}{\partial \theta_k} = 1$$

$$\eta \frac{\partial e}{\partial y_j} = -\sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) h_j' w_{ij} \quad (3.15)$$

$$\frac{\partial y_j}{\partial h_j} = h_j' \quad (3.16)$$

$$\frac{\partial Y_k}{\partial H_k} = H_k' \quad (3.17)$$

BP 神经网络各层的权值阈值的最终调整量如下:

$$\Delta w_{ij} = \eta \sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) H_k' w_{jk} h_j' x_k \quad (3.18)$$

$$\Delta\gamma_j = \eta \sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) h_j' \quad (3.19)$$

$$\Delta w_{jk} = \eta \sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) h_j' y_j \quad (3.20)$$

$$\Delta\theta_k = \eta \sum_{p=1}^C \sum_{k=1}^n (T_k^p - Y_k^p) H_k' w_{jk} h_j' \quad (3.21)$$

### 3.2.3 BP 神经网络的缺陷分析

BP 神经网络凭借其独有的误差反馈机制成为了迄今为止应用最为广泛,最为成功的人工神经网络,但是其本身仍然存在一些缺陷<sup>[47]</sup>,具体表现在以下几个方面:

(1) BP 神经网络全局收敛速率较慢。BP 神经网络在进行网络训练时,常常因为对精度要求过高或是样本量过大,而使得 BP 神经网络收敛速度在后期显著的慢于前期,甚至会出现无法收敛的现象。

(2) BP 神经网络容易陷入局部极值。BP 神经网络进行问题求解时,有时会因为数据样本间的差异性太大导致函数处理时出现多个峰值,容易陷入局部最优而无法跳出。

(3) 网络整体的性能特别依赖于初始权值与阈值。基于其最初权值阈值的随机性,导致网络在处理相同问题时会呈现出不同的性能,所以如何正确得到最适合网络工作的初始权值阈值,是改进 BP 神经网络的关键问题。

(4) 关于 BP 神经网络隐含层层数及层内节点数,目前未有明确的定义方式。隐含层的层数和层内节点数对于整个网络训练速度和识别精度有着决定性作用,而目前常用经验公式或者逐个试验来确定。

### 3.3 GA-BP 神经网络的构建

#### 3.3.1 遗传算法理论分析

遗传算法 (Genetic Algorithm, GA) 是 Holland 教授于 1975 年受达尔文生物进化论的启发而提出,该算法是模拟自然界生物的自然选择和遗传学机理的进化过程的计算模型,是一种通过模拟自然进化过程随机搜索最优解的方法<sup>[48]</sup>。

遗传算法基本思想是模拟自然界优胜劣汰的进化现象,把对最优解的搜索空间视作为生物的继承进化过程,该算法的目标主体只针对于待求解问题本体,不再考虑函数微分和函数是否离散的限制,以一种随机性的规则完成求解过程,遗传算法摒弃了传统算法所需硬性规则限制,这也就意味该算法有着传统求解算法并不具备的并行求解能力和全局寻优能力,而构成该算法的基本要素主要包括:

##### (1) 染色体编码

染色体编码操作是指通过将待求解问题所有的解进行某种固定形式的重新编写,形成一系列的符号串来表示种群中的染色体,完成从表现型到基因型的转变,通常采用的编码方式有二进制编码(其等位基因由 0 和 1 来表示),实数制编码(其等位基因由每一个实数表示)。

##### (2) 适应度函数

遗传算法中适应度函数是指在评估个体在染色体种群中优劣性所采取的函数规则。其选取方式灵活,一般会随着函数问题最优值取极大值或是极小值而改变,当所求函数的最优值为前者时,通常采用公式(3.22)作为适应度函数,当所求函数的最优值为后者时,则一般采用公式(3.23)作为评价规则。

$$fit(x) = g(x) \quad (3.22)$$

$$fit(x) = \frac{1}{g(x)} \quad (3.23)$$

### (3) 染色体的选择复制

染色体的选择复制操作象征的是以一定的概率选择父染色体和母染色体进行交配行为产生下一代个体，选择操作体现了遗传算法不断进化的思想。

### (4) 交叉和变异操作

父染色体和母染色体通过交换各自拥有的基因，形成一个新的个体的过程称为遗传算法的交叉操作，这一新个体顺势拥有了父母的基因特性，而在交叉的过程中，会有部分基因产生变异，使这一个新的个体拥有不同于父母所有遗传基因的特性。交叉和变异的概率决定了遗传算法种群的多样性是否丰富，也制订了遗传算法在解决优化问题时的随机选择规则。

### (5) 染色体解码

这一操作过程相当于染色体编码过程的逆运算，将不同的染色体字符串输出为待求解问题的真实值，从而完成由基因型到表现型的转变。对于实际中的不同问题所采取的解码方式也不尽相同，如解决分类问题时，一般采用 ONE HOT 解码方式，解决函数极值问题时，就会采用二进制或实数制的解码方式。

## 3.3.2 GA 优化 BP 神经网络思路分析

针对上一节中提到的 BP 神经网络目前仍然存在的缺陷，利用遗传算法优化 BP 神经网络的思想，主要从下面几个方面入手：

(1) 遗传算法优化 BP 神经网络的初始权值和阈值<sup>[49]</sup>。BP 神经网络中用于学习训练的常见数学方法是梯度下降法，但采用这种学习规则的网络性能与最初的第一组权值和阈值的优劣息息相关，网络本身的第一组权值阈值向量一般是在某个区间中概率捡取的，这也在一定程度上加重了网络收敛成功以及收敛速度的不确定性，不利于实际问题的分析。而利用遗传算法可以将初始权值阈值的选取作为一个求解最优值的函数问题，经过算法一系列的操作后选出最优初始值赋给 BP 神经网络进行学习训练，以此提高网络的性能。

(2) 遗传算法对 BP 神经网络拓扑结构选择的优化。BP 神经网络拓扑结构优劣对于一个网络的处理能力影响极大，输入输出层的选取取决于网络输入量和输出量的维数，选择方法较为固定，但是隐含层的选择没有统一的理论指导，过多的隐含层或者隐含层节点虽然会提高训练精度，但是带来的问题便是网络处理能力急剧变慢，而利用遗传算法选择处理问题最优网络结构也成为了优化思想的中的一环。

(3)利用遗传算法修改网络的学习规律。神经网络的学习速率选取尚未有统一的规范，通常是经过不断的调整实验来选定的一个固定值，但是如果凭靠固定选取的参数，轻则造成网络训练所耗时间长，重则导致整个系统不收敛。因此采用遗传算法对 BP 神经网络的学习规律加以优化，从而使得网络学习能力可以动态变化，从而适应更加难以处理的问题，这也是遗传算法优化 BP 神经网络的另一种优化思路。

本文经过综合考虑，认为第一种优化思路对网络的优化效果最好，优化所需时间最短，并且对于不同问题具有很好的泛化性。第二种方案对于 BP 神经网络结构的优化，完全可以通过测试数据集的多次实验来确定最佳结构，而考虑到 BP 神经网络的学习率调整远没有更好的网络初始权值阈值对于网络性能提升影响大，故选择利用遗传算法的全局寻优能力，经过继承进化的过程找到与所求问题最为契合的第一组权值阈值，减少神经网络陷入局部极值的几率，提高网络收敛成功的确定性，缩减网络求解问题全部所需时间。

### 3.3.3 GA-BP 神经网络实现流程

遗传算法的基本思想决定了其在实现函数求解的方式是采取递进式方法：从在解空间中随机选择一部分解作为初始解，通过选择、交叉、变异等步骤逐步接近最优解，在进化过程，使子代在继承父代优点的同时，仍然保留一部分独立性<sup>[50]</sup>，但从总体上确保子代优于父代，遗传算法优化 BP 神经网络的具体步骤如下：

步骤 1.种群(Population)初始化：BP 神经网络的权值阈值通常是在(0,1)这个区间内选择，所以 GA 的初始种群为(0,1)范围内的数据集，通常来说，个体的数量与种群的多样性之间存在一种正相关的关系。但过多的个体也会相应的造成运算量成几何式增长，延长寻优过程所需时间；而如果个体的数量设置过少，则会增加算法陷入“早熟”的几率。

步骤 2.染色体编码(Code)：本文中的染色体为 BP 神经网络的权值和阈值，编码操作就是将这些值按照权值在先，阈值在后，按结构顺序排列生成染色体的过程。本文采用的编码方式为实数编码，对应于神经网络结构的编码形式如图 3.3 所示。

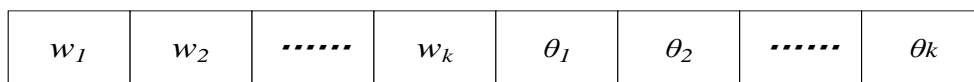


图 3.3 权值阈值编码形式

Fig. 3.3 Coding method of weight and threshold

步骤 3.适应度(Fitness)函数的设定：本文所建立的 GA-BP 神经网络的用途是完成故障的检测并给其分类，而 BP 神经网络的工作特点在于误差调整，所以适应度函数可以取均方误差的倒数函数，或者是实际与期望输出误差和的倒数函数。本文选用的适应度函数为网络期望误差与实际误差差值和的倒数，其数学表达式由公式(3.24)给出。

$$F = \frac{L}{\sum_{k=1}^n (T_k - Y_k)} \quad (3.24)$$

其中,  $L$  为常数;  $k$  为  $n$  个网络输出节点中的第  $k$  个。

**步骤 4.选择复制(Reproduction):** 按照适应度的高低进行顺序排列, 在这个进化过程中按照一定的规则, 将排序位于后面的个体剔除, 将排序在前的部分个体进行复制进化。目前经常使用的选择方法有轮盘赌法和精英保存法, 前者提高了复制过程的随机性, 后者则加强了快速收敛性, 基于对神经网络的缺陷考量, 本文使用了精英保存的方法来达到快速完成故障检测过程, 加强整个网络的实时性的目的。

**步骤 5.交叉(Crossover)和变异(Mutation):** 传统的遗传算法一般将交叉和变异的概率取为固定值, 这种取值方式束缚了种群在迭代过程的多样性, 无法有效的平衡种群进化前期与进化后期的稳定性。如果将交叉的概率设计过大, 很容易造成种群趋向单一化, 而如果将交叉的概率  $P_c$  设计的过小, 则会造成新个体产生的速度过慢; 变异概率  $P_m$  大小对于种群影响刚好和交叉操作造成的影响相反。因此, 为了保证种群在遗传过程中的多样性, 兼顾优良个体与新生个体在整个种群的比例, 本文选择了一种自适应改变交叉和变异的概率<sup>[51]</sup>, 使得染色体在遗传过程中交叉变异的取值根据适应度的大小不断改变, 数值调整公式如下:

$$P_c = \begin{cases} P_{cmax} - \frac{(P_{cmax} - P_{cmin})(f_{avg} - f')}{f_{avg} - f_{min}}, & f' \leq f_{avg} \\ P_{cmax}, & f' > f_{avg} \end{cases} \quad (3.25)$$

$$P_m = \begin{cases} P_{mmax} - \frac{(P_{mmax} - P_{mmin})(f_{avg} - f')}{f_{avg} - f_{min}}, & f' \leq f_{avg} \\ P_{mmax}, & f' > f_{avg} \end{cases} \quad (3.26)$$

其中,  $P_{cmax}$ 、 $P_{mmax}$  分别为交叉和变异最大取值;  $P_{cmin}$ 、 $P_{mmin}$  分别为交叉和变异最小取值;  $f'$  为当前个体适应度;  $f_{avg}$  为种群平均适应度。

**步骤 6.转向个体评价(Assess):** 通过计算个体的适应度, 来判断完成继承进化的个体是否要优于上一代平均适应度, 如果判断满足所设要求, 则继续进行步骤 7 完成解码, 若不满足则转向步骤 4 进行第二次循环。

**步骤 7.解码(Decoding):** 将最终优胜出的染色体按照对应于步骤 2 的排序及编码方式进行解码。

**步骤 8.赋值(Assign):** 由上一步可知, 染色体字符串到具体数值的转换已经完成, 现在将遗传算法优选出的一组权值阈值递送给 BP 神经网络。



上述所有步骤执行完成之后，遗传算法便成功完成了对 BP 神经网络的优化过程，优化后的 BP 神经网络可以使用这组权值阈值进行网络的训练学习，进行接下来对测试数据集的检测。GA-BP 神经网络模型的实现流程如图 3.4 所示。

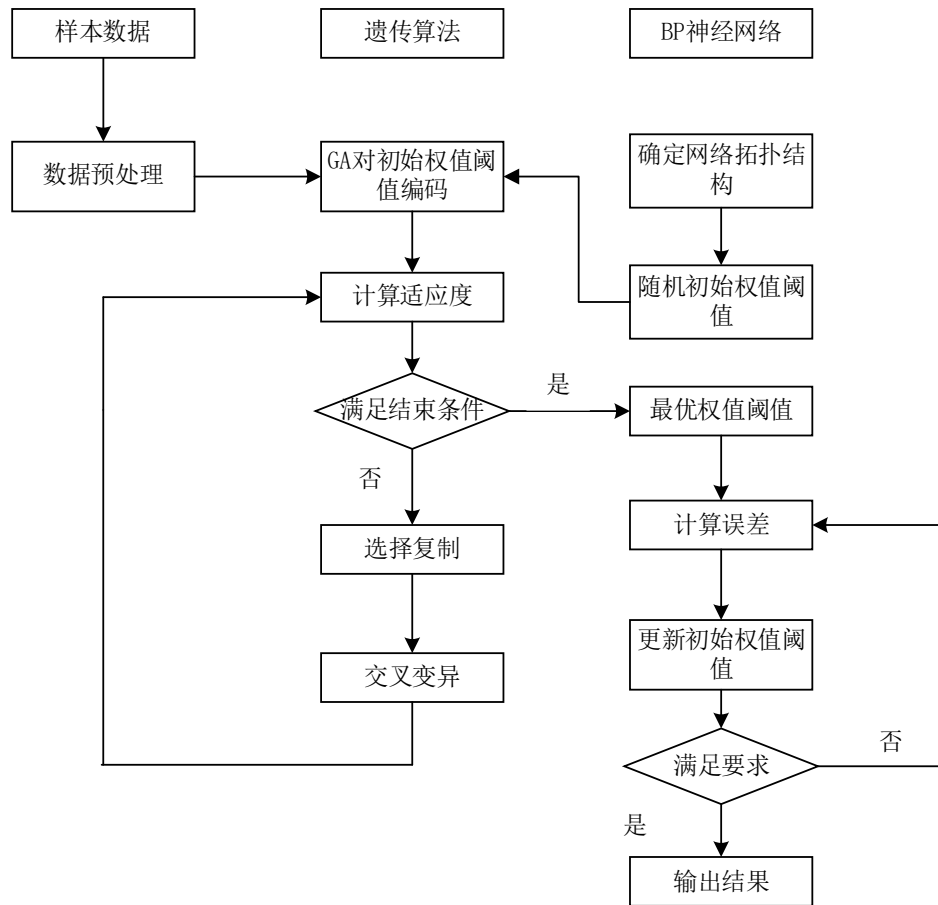


图 3.4 GA-BP 神经网络工作流程

Fig. 3.4 Flow chart of GA-BP neural network

### 3.4 GA-BP 神经网络应用于光伏阵列故障检测研究

#### 3.4.1 应用于光伏阵列故障检测的 BP 神经网络研究

BP 神经网络输入输出层的选择一般是通过数据集维度来选择，在本文的第二章中，已经详细分析了光伏阵列发生故障时特征参数的变化规律，所以在进行光伏阵列故障检测时的模型输入变量为其四个特征参数，分别是开路电压  $V_{oc}$ ，短路电流  $I_{sc}$ ，最大功率点处的电压  $V_{mmp}$ ，最大功率点处的电流  $I_{mmp}$ ，即输入层节点数  $m=4$ 。输出层要求表征五种故障状态，所以采用输出量在  $[0, 5]$  区间内取整的方法，即输出层节点数  $n=5$ 。

隐含层的任务是完成对输入数据的处理，隐含层越多，所能处理的信息更为全面<sup>[52]</sup>，

但是所需节点数也会相应增多，通常来说隐含层层数不超过三层，本文采集到的故障数据集与故障状态之间为非线性映射的关系，电压电流的取值只保留了两位小数且要求网络训练过程满足一定的快速性，故只采用一层隐含层。一般而言，隐含层的节点数目，在一定的数量范围内，节点数越多，检测准确度就越高，但随之带来的是网络训练速度会变慢很多，隐含层节点数常用以下公式来确定：

$$\sum_1^m C_q^i > b \tag{3.27}$$

$$q = \log_2 m \tag{3.28}$$

$$q = \sqrt{m + n} + a \tag{3.29}$$

其中， $\alpha$  为在[1,10]取到的整数； $b$  为输入数据的维数； $n$  为输出层节点数； $q$  和  $m$  分别为隐含层与输入层节点数；如果  $i > q$ ，规定  $C_q^i = 0$ 。

在实际试验中，因为数据量度相差比较大，数据复杂程度较大，常通过“试凑法”来选取隐含层节点数。具体实验方法为先通过公式(3.29)确定一个大致的区间，然后不断更改  $\alpha$  的取值，在 Matlab 软件中反复验证，本文中  $q = \text{sqrt}(4+5) + a$ ，故在[4,13]个节点数中选择，实验误差评价选择均方根误差  $mse$ ，计算出的隐含层节点数与误差大小以及迭代次数的关系如表 3.1 所示。

表 3.1 不同隐含层节点数对应数据集误差大小

Tab. 3.1 Error size of data set corresponding to the number of nodes in different hidden layers

隐含层节点数	均方误差大小	迭代次数
4	0.030434	132
5	0.009631	93
6	0.003802	121
7	0.003748	103
8	0.005407	97
9	0.002197	92
10	0.004166	88
11	0.006956	123
12	0.008907	105
13	0.034622	117

从上表中可以看出，BP 神经网络误差随着隐含层结点数的增加，基本保持着先降低后

升高的规律，在设置 9 个隐藏层节点时，网络迭代次数较少且误差最小，所以对于解决本文光伏阵列故障检测分类问题，选择 9 个隐含层节点数最为合适。

在网络的节点数已经确定好之后，接着是对激活函数的选择，激活函数用来处理各节点输入的信息，并将其映射到下一级的节点中<sup>[53]</sup>，本文采用的激活函数为最常用的 Sigmoid 函数，其数学表达式已经由公式 (3.1) 给出。

网络的结构确定好之后，最为重要的参数选择就是训练函数的选择，BP 神经网络除了最早提出的梯度下降训练法，近年来很多学者在此基础上提出了一些改进的训练算法，如 Trainlm 算法，Traingdx 算法等等。不同的算法对于网络训练的时间和精度差异比较大，为了选择更为适合光伏阵列故障分析的训练方法，从故障数据集中挑选了 100 组故障数据测试了各种训练函数迭代次数与所能达到精度，测试结果如下：

### (1) Traingd 训练函数

Traingd 算法是早提出的梯度下降训练法，其特点是训练速度较慢，误差保持线性下降。由于其提出的时间较早，该算法在面对复杂数据时，需要经过非常多次的迭代次数才能达到期望的误差要求，且容易陷入局部极值。在处理一些变化较大的数据时，甚至有可能无法收敛。图 3.5 为使用 Traingd 训练函数时，BP 神经网络的训练性能示意图。

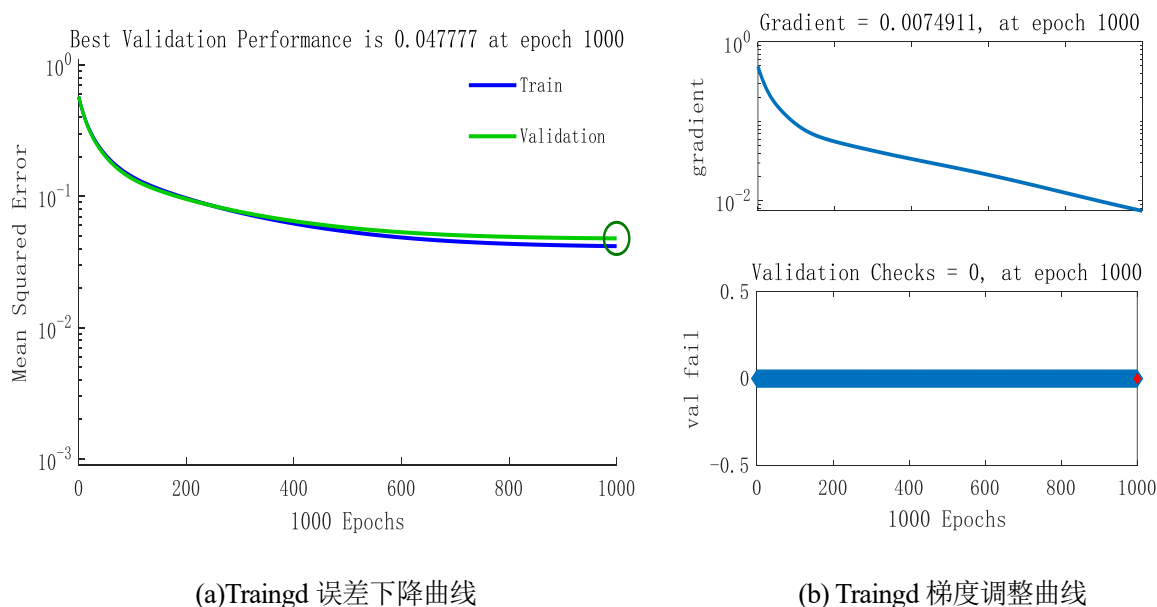


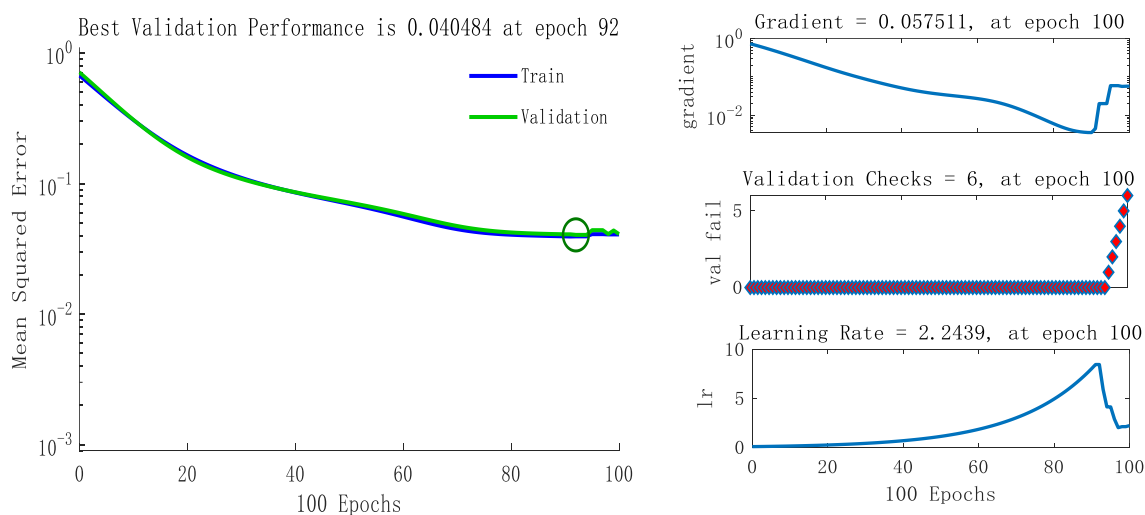
图 3.5 Traingd 函数训练性能图

Fig. 3.5 Traingd performance chart

在图 3.5(a)中可以看出，网络经过了 1000 次迭代时，其训练误差才刚刚接近 0.04，说明该训练函数对于光伏阵列的故障数据识别困难，而从图 3.5(b)中的曲线走向可以非常直观的看出，在 Traingd 训练函数参与的训练过程中，为了追寻较小的误差，其修正梯度是一直保持线性负调整，无论是从训练时间还是误差精度上看，该训练函数无法完成网络训练的任务。

### (2) Traingdx 训练函数

Traingdx 算法是在梯度下降法基础上改进而来的训练算法，该算法添加了动量因子，理论收敛速度远大于 Traingd 算法，并且对于大批量的数据处理性能更加优异，图 3.6 是采用 Traingdx 训练函数时，BP 神经网络的训练性能图。



(a)Traingdx 误差下降曲线

(b) Traingd 梯度调整曲线

图 3.6 Traingdx 函数训练性能图

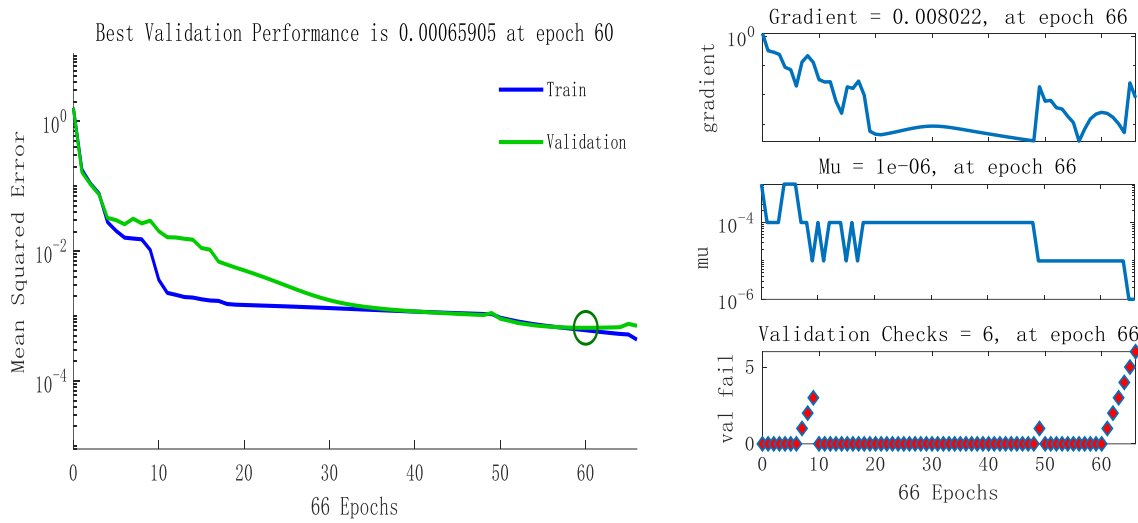
Fig. 3.6 Traingdx performance chart

分析图 3.6(a)中的曲线可知，Traingdx 训练函数参与的训练过程，网络在 100 次迭代时完成了收敛，结束了训练过程，在经过迭代 92 次时误差达到了最小值 0.040484，由图 3.6(b)中的信息可知，该训练函数的梯度调整曲线不再一直保持下降趋势，而是通过不断的调整学习率来决定梯度是否下降。由此可见，动量因子对于加强梯度的修正过程有着一定的积极意义，但是对于光伏故障数据集来说，其所能达到的误差精度还是不尽人意。

### (3) Trainlm 训练函数

Trainlm 算法全称为 Levenberg-Marquardt 算法，也是在梯度下降法的基础上融合牛顿法改良而来的一种最小二乘法，对于大中等数据集有着最快的收敛速度，图 3.7 是采用 Trainlm 训练函数时，BP 神经网络的训练性能图。

从图 3.7(a)中可以直观的看到在 60 次迭代时，训练误差就已经达到了 0.00065905，最终总共经过 66 次迭代完成了整个的训练过程。而从图 3.7(b)中可以看出，对于 100 组训练数据，采用该算法的神经网络其梯度调整频率更加频繁，说明该训练函数对于光伏故障有着更好的识别性。



(a)Trainlm 误差下降曲线

(b)Trainlm 梯度调整曲线

图 3.7 Trainlm 函数训练性能图

Fig. 3.7 Trainlm performance chart

从上述对三种训练函数的性能对比可以发现，当采用 `Traingd` 函数对光伏阵列故障数据集进行训练过程时，整个网络不能有效的收敛，而且在 1000 次训练时，其误差 0.047，远不能达到我们规定的误差要求；采用 `Traingdex` 函数训练时，网络可以有效的收敛，但是迭代次数仍是较大，且误差只能接近 0.04，对于精度上的要求，该函数还不能很好的满足，如果采用更多得到数据集进行训练，也无法满足快速性的要求。而当采用 `Trainlm` 函数时，网络所展现的训练性能均好于采用另外两种算法，故在网络其他结构以及参数已经确定的情况下，本文选用了 `Trainlm` 算法作为 BP 神经网络的训练函数。

### 3.4.2 光伏阵列故障数据预处理

数据预处理是将原始输入量按照比例转换为没有量纲的标么值，本文所搭建的网络所需输入变量为光伏阵列的 4 个典型参数，这些参数之间单位不同，也有着不同的数量级，如果不对这些数据集归一化处理而直接将生数据交由网络自学习，则可能会由于数据误差太大而引起整个网络崩溃，所以必须将这些典型参数数据集转换到不受其他单位等因素影响的标么值。本文采用的归一化方法为最大最小均值法，所用公式如下：

$$X' = \frac{X_i - X_{min}}{X_{max} - X_{min}} \tag{3.30}$$

其中： $X_{max}$  为样本数据中的最大值； $X_{min}$  为样本数据中的最小值； $X'$  为均一化之后的数据； $X_i$  为原始输入数据中的第  $i$  个数据。

从第二章搭建的光伏阵列模型中提取的部分样本输入数值如表 3.2 所示，得到的这些

样本数值归一化结果如表 3.3 所示。

表 3.2 部分故障数据  
Tab. 3.2 Partial fault data

故障编号	故障类别	开路电压	短路电流	最大功率点处电	最大功率点处电
		$V_{oc}(V)$	$I_{sc}(A)$	压 $V_{mmp}(V)$	流 $I_{mmp}(A)$
A1	短路故障	149.42	15.53	118.33	12.98
A1	短路故障	98.95	15.59	82.47	13.18
B1	开路故障	178.76	10.68	150.79	9.68
B2	开路故障	178.72	5.17	149.25	4.51
C1	阴影遮挡	176.67	15.3	120.38	14.89
C2	阴影遮挡	176.39	15.25	152.75	10.16
D1	老化故障	178.39	15.45	142.34	13.57
D2	老化故障	178.69	14.97	126.14	11.74
E	正常运行	178.79	15.64	150.74	13.49

表 3.3 经归一化后的部分故障数据  
Tab. 3.3 Normalized partial fault data

故障编号	故障类别	开路电压 $V_{oc}$	短路电流 $I_{sc}$	最大功率点处	最大功率点处
				电压 $V_{mmp}$	电流 $I_{mmp}$
A1	短路故障	0.25987	0.95925	0.04099	0.75576
A1	短路故障	-0.9768	0.99814	-0.9882	1
B1	开路故障	0.99501	0.00703	0.88331	0.06732
B2	开路故障	0.99675	-0.9851	0.96556	-0.9238
C1	阴影遮挡	0.94690	0.9203	0.07023	0.98076
C2	阴影遮挡	0.94166	0.93703	0.9983	0.06923
D1	老化故障	0.9965	0.91111	0.71002	0.75192
D2	老化故障	0.99401	0.8296	0.2659	0.36923
E	正常运行	0.9877	0.99629	0.92866	0.93569

### 3.4.3 GA-BP 神经网络应用于光伏阵列故障检测的验证分析

为了验证 GA-BP 神经网络模型应用于光伏阵列故障检测的可行性，本文采用自前文建立光伏阵列仿真模型中采集阵列实时运行数据的方案，在本章实验中，共采集到了 1000 组数据，其中，用于网络自学习样本 850 组，测试网络最终性能样本 150 组。

数据采集结束之后，需要对 GA-BP 神经网络的相关参数进行设定。此次实验中对遗传算法参数设置如下：初始染色体数量为 50；限制进化代数为 50 代；染色体发生交叉的最大最小概率分别为 0.3 和 0.1；染色体发生变异的最大最小概率分别为 0.1 和 0.01。对 BP 神经网络的参数设置如下：网络训练次数为 1000 次，学习速率为 0.1，限制最小目标误差为 0.0001，训练函数采用 Trainlm 函数。

程序运行完成后，得到如图 3.8 和图 3.9 所示的遗传算法优化前后的 BP 神经网络均方误差和网络迭代次数之间的关系曲线，从图 3.8 中可以看出，未经优化 BP 神经网络自第 94 次迭代之后，其均方误差不再呈现下降趋势，此时均方误差的精度为 0.0022961，说明此时网络已经完成训练；而从图 3.9 中可以看出经遗传算法优化后的神经网络在 65 次迭代之后其均方误差基本保持不变，而且精度达到了 0.00073651，对比以上优化前后数据，可以得知，经过遗传算法优化后的神经网络缩减了近乎 1/3 所需迭代次数，在保证短时性的前提下提高了网络所能达到的误差精度，这也显示出，GA-BP 神经网络的训练性能更为高效。

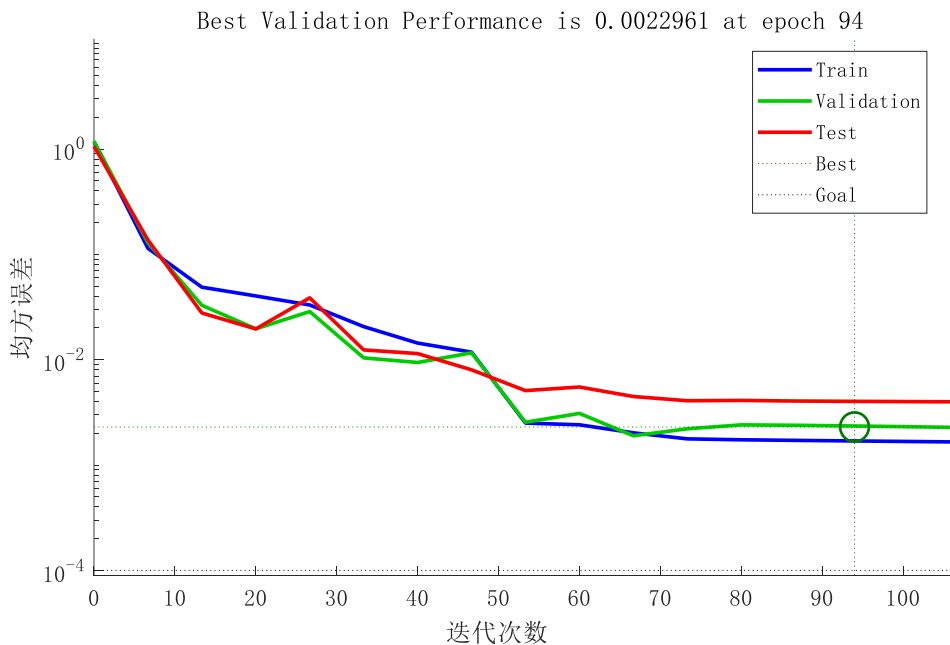


图 3.8 BP 神经网络训练误差曲线

Fig. 3.8 BP neural network training error curve

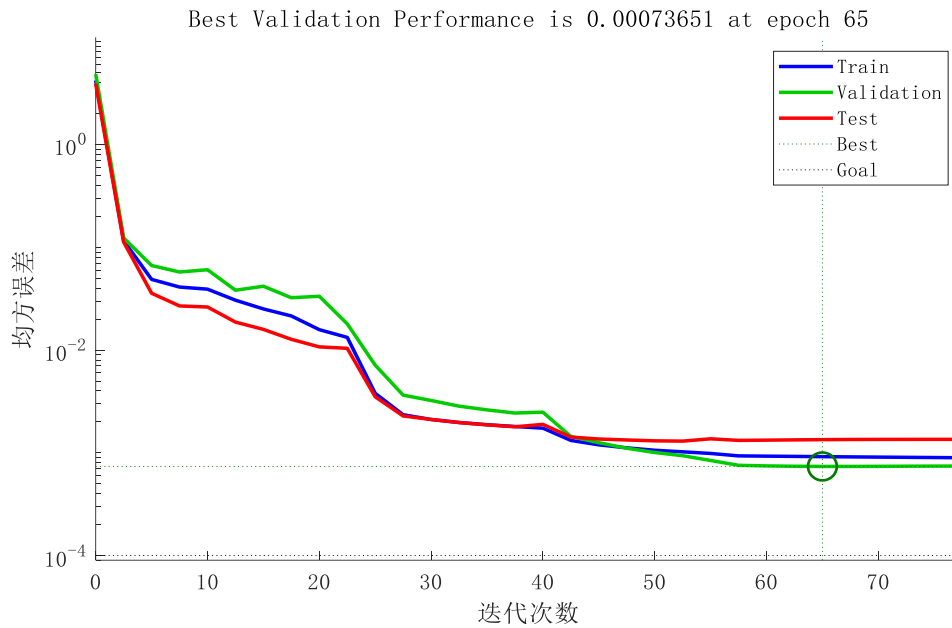


图 3.9 GA-BP 神经网络训练误差曲线

Fig. 3.9 GA-BP neural network training error curve

为研究遗传算法优化前后的 BP 神经网络的故障检测准确率差异，将完成训练的两组神经网络用于测试样本的故障检测，考虑到样本容量可能对检测结果产生影响，先将 60 组测试样本用于网络检测，BP 神经网络诊断结果如图 3.10 所示。分析网络对 60 组测试数据的检测结果可知，BP 神经网络有 5 组数据判别错误，其检测准确率为 91.7%。

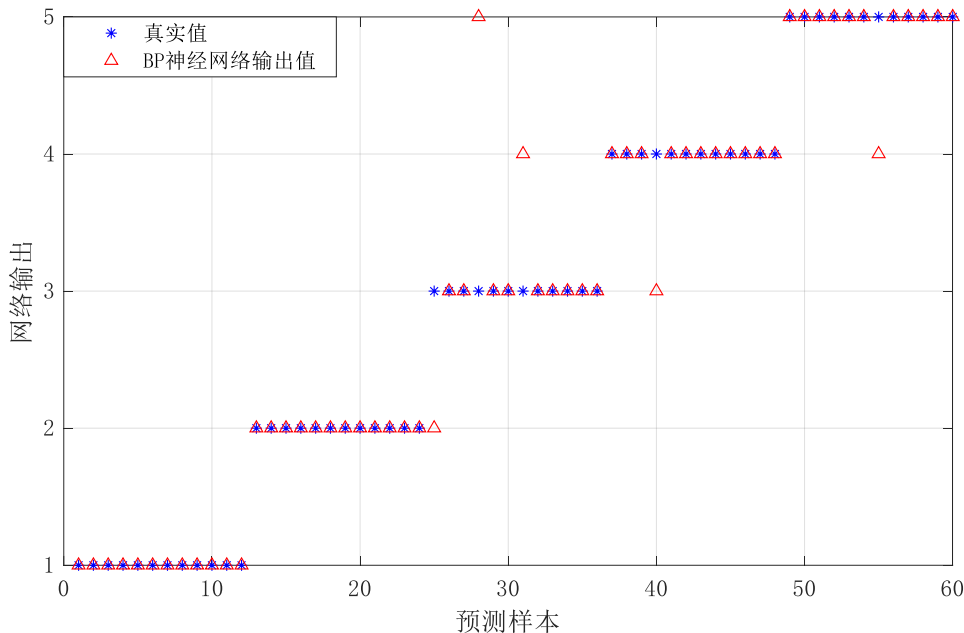


图 3.10 BP 神经网络对小容量样本的检测结果

Fig. 3.10 Detection results of small-capacity samples by BP neural network

GA-BP 神经网络的检测结果如图 3.11 所示，其中存在 3 组故障数据检测出错，剩余的



57 组测试数据均正确识别，其检测准确率为 95%。而且对于短路、开路故障数据的检测更为精准，不存在误判这两种故障的情况。而在面对数据区分度较低的老化和阴影遮挡故障，GA-BP 神经网络的检测精准度也有了一定程度的提高。

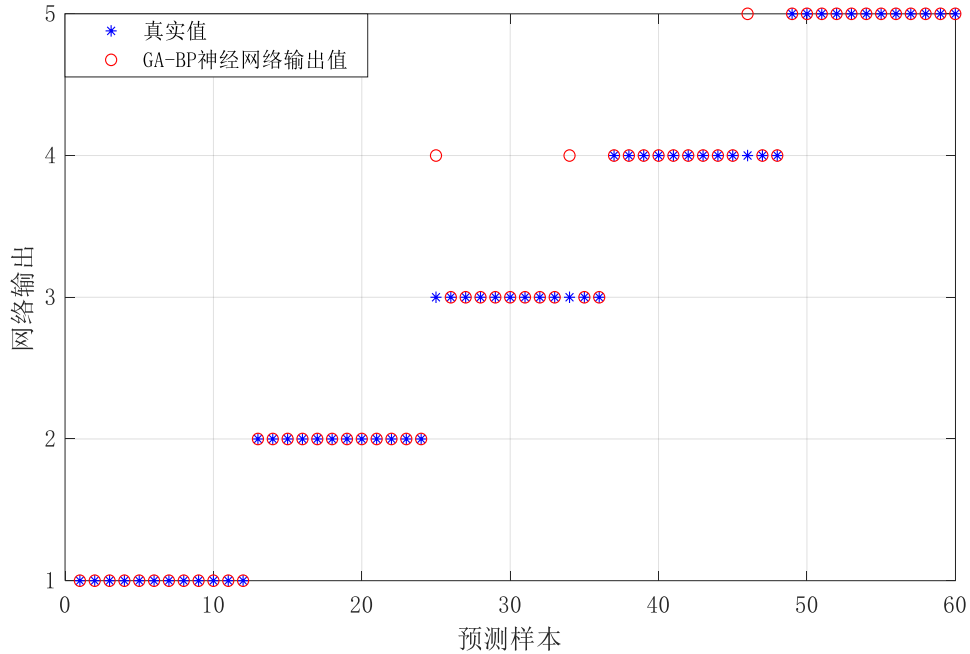


图 3.11 GA-BP 神经网络对小容量样本的检测结果

Fig. 3.11 Detection results of small-capacity samples by GA-BP neural network

由以上检测结果可知，在进行小容量样本的检测时，GA-BP 神经网络比原 BP 神经网络的准确率提高了 3.3%，可以更精准的判别故障类型。为了进一步验证经过遗传算法优化后的 BP 神经网络的可靠性，将 150 组数据测试样本作为大容量样本进行故障检测结果的验证，在这些数据集中，各种类型的故障数据与正常数据均比分配，即每种类型的数据各有 30 组。

图 3.12、3.13 是遗传算法优化前后的 BP 神经网络在进行大容量样本检测时的实验结果图，从图 3.12 中可以看出，随着样本数量的增加，原 BP 神经网络故障检测失败组数也相应的增加，统计故障检测结果，共有 10 组数据判别错误，其检测准确率约为 93.3%，较小样本的检测准确率稍有提升。在这些检测失败的数据中，老化与阴影遮挡数据仍是占比较大，尤其是将阴影遮挡故障数据判别为正常数据，这在实际应用中是绝对无法接受的结果，而对于短路和开路故障，BP 神经网络检测结果较为理想，这与两种类型的故障数据在整个数据空间中具有较高的辨识度也有一定的关系。

从图 3.13 可以看出，遗传算法参与优化的 BP 神经网络对于大容量光伏阵列故障数据集的检测精准率达到了 96%，存在 6 组判别错误的的数据，较之前的小容量样本识别准确率也有提升。而这些检测错误的的数据集中分布在老化和阴影遮挡故障中，对于短路和开路故

障的识别具有非常高的可靠性。

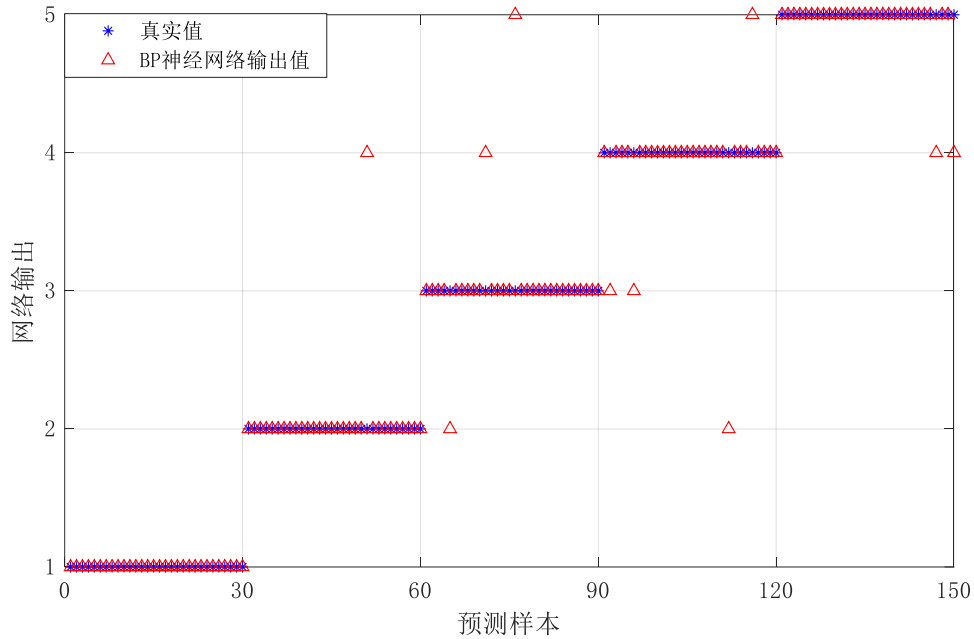


图 3.12 BP 神经网络对大容量样本的检测结果

Fig.3.12 Detection results of large-capacity samples by BP neural network

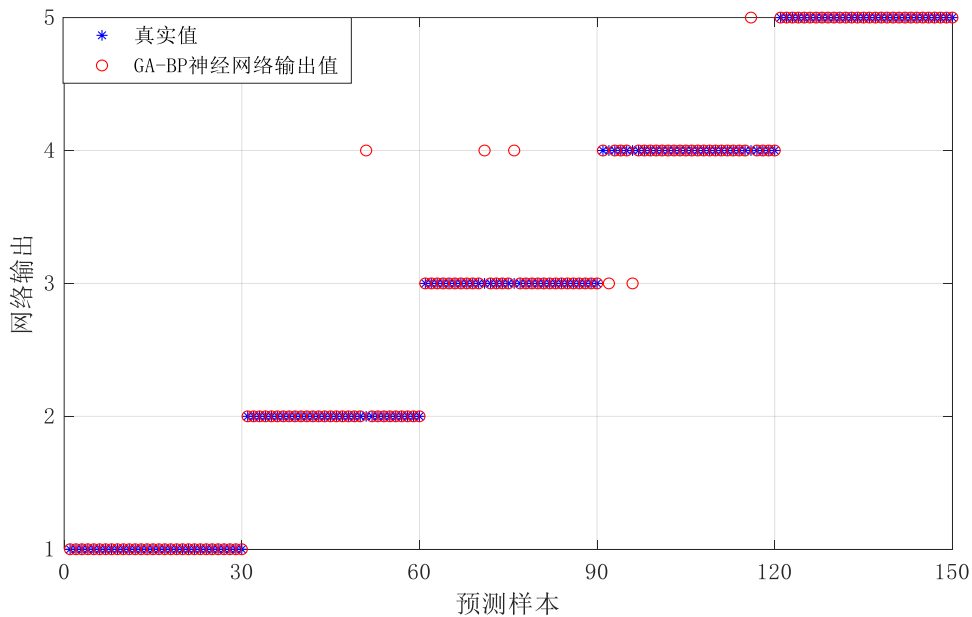


图 3.13 GA-BP 神经网络对大容量样本的检测结果

Fig.3.13 Detection results of large-capacity samples by GA-BP neural network

分析以上实验结果，可知经过遗传算法优化后的 BP 神经网络可以顺利完成对光伏阵列故障状态的检测，而且无论是对于大容量样本预测还是对于小容量样本预测的精度问题上，GA-BP 神经网络有着非常明显的优势，在网络训练速度这一方面，采用遗传算法优选

权值阈值的 BP 神经网络可以以较少的迭代次数完成更为精准的误差训练，而整个检测过程中所需要时间主要由网络训练时间决定，所以这也大大减少了整个检测过程所需时间，而且其网络整体性能也得到了一定的提升，但是从最终实验结果来看，GA-BP 神经网络在样本学习过程最小误差精度还是不够精准，对于识别区分度较低的故障数据带来了一定的困难，而且面对大容量数据样本检测时，存在将故障数据误检测为正常数据以及将数据辨识度较高的开路故障判别为老化故障的问题，这一点也为日后将该网络投入到实际运行中的光伏电站造成了阻碍。

### 3.5 本章小结

本章首先分析了 BP 神经网络结构组成以及工作原理，然后对遗传算法的工作原理与重要参数研究，给出了遗传算法优化 BP 神经网络的具体方式和具体流程，经过多次实验确定了用于光伏阵列故障检测的 BP 神经网络的最佳结构以及最优参数，通过实验验证了本章所建立的 GA-BP 神经网络模型对容量大小不同的光伏阵列故障数据集的检测与识别的可行性，最后分析对比实验结果发现，GA-BP 神经网络在进行光伏阵列故障诊断时比传统的 BP 神经网络的性能更加优异，但是对于老化故障与阴影遮挡故障数据的识别还不够理想，如何进一步提高 GA-BP 神经网络的性能是接下来所需做的工作。

## 第四章 混合遗传算法优化 BP 神经网络的研究

在上一章节中，顺利的验证了 GA-BP 神经网络在光伏阵列故障检测中的可行性，但是从最终检测结果来看，经遗传算法优化的 BP 神经网络的训练迭代时间还是较长，不利于快速完成检测过程，故障识别准确率虽然得到了提高，但是对于阴影遮挡及老化故障的识别还不够理想，考虑到 GA-BP 神经网络受遗传算法影响较大，所以，本章提出了一种鸽群算法与遗传算法相结合的混合遗传算法寻找 BP 神经网络的最优初始权值阈值，以此来进一步提高所建立的 GA-BP 神经网络性能。

### 4.1 鸽群算法原理

鸽群优化算法 (Pigeon-inspired Optimization algorithm, PIO) 是由段海滨教授及其团队受自然界中鸽子回巢活动的启发，通过研究鸽群在回巢过程中的特殊导航行为，提出了一种新型仿生群体智能优化算法。该算法分为两个阶段，第一个阶段是鸽群在寻找巢穴的前期，会使用地图和指南针算子作为寻路工具<sup>[54-55]</sup>，确定当前位置信息，这个阶段本质上为鸽群在进行全局寻优的过程；另一个阶段是当鸽群靠近目的地时，鸽子无法再依靠磁场和太阳进行寻路，只能通过对一些如河流，十字路口，大型广告牌等参照物的记忆，进行对目标位置信息的搜索，这一阶段实际代表着鸽群局部寻求最优解的过程。

#### 4.1.1 地图和指南针算子

鸽子在回巢初始阶段，可以利用地磁场造成的磁偏角在大脑中建立位置地图<sup>[56]</sup>，并且可以把太阳高度作为指南针来调整飞行方向，这个阶段称之为鸽群算法中地图和指南针算子参与的第一个迭代过程阶段。当利用鸽群算法处理函数优化问题时，该过程相当于全局寻优的过程，先对每个可行解赋予鸽子个体  $i$  的含义，然后给予每个个体初始的位置和对应的速度，分别记为：

$$X(i) = [X_1(i), X_2(i) \dots X_D(i)], i = 1, 2 \dots N_p \quad (4.1)$$

$$V(i) = [V_1(i), V_2(i) \dots V_D(i)], i = 1, 2 \dots N_p \quad (4.2)$$

其中， $N_p$  为鸽群总数目； $D$  为需要求解问题的维度。

鸽子飞行中的每一个位置都代表着函数问题的一个解，鸽子个体在这一过程中不断更新自己所在位置，在第  $t+1$  次迭代时，其位置更新按照公式 (4.3) 进行计算。

$$X^{(t)}(i) = X^{(t-1)}(i) + V^{(t)}(i), i = 1, 2 \dots N_p \quad (4.3)$$

其中,  $X^{(t)}(i)$ 为鸽子在当前迭代过程中的位置;  $X^{(t-1)}(i)$ 为鸽子上一次迭代过程的位置;  $V^{(t)}(i)$ 为鸽子当前迭代过程中的速度。

鸽群中的个体会以不同的速度抵达下一时刻想要到达的位置, 其速度改变大小按照公式 (4.4) 更新。

$$V^{(t)}(i) = V^{(t-1)}(i)e^{-R} + rand(X_g - X^{(t-1)}(i)) \quad (4.4)$$

其中,  $R$  为地图与指南针算子, 一般取值 0.2;  $rand$  为一个 0 到 1 之间的随机数;  $X_g$  为当前全局最优位置。

鸽群个体经过不断的更新自己的位置与速度, 会得到当前位置距离目的地的远近, 将这个距离的远近大小描述为适应度, 以适应度的大小来评估当前位置所对应解的质量, 之后通过比较所有鸽子的位置并选出适应度最优的鸽群个体。该迭代过程将进行若干次直至满足结束条件后终止, 并把在该过程中得到的位置结果递送给地标算子以完成下一阶段的迭代。直观的地图与指南针算子参与的迭代阶段如图 4.1 所示:

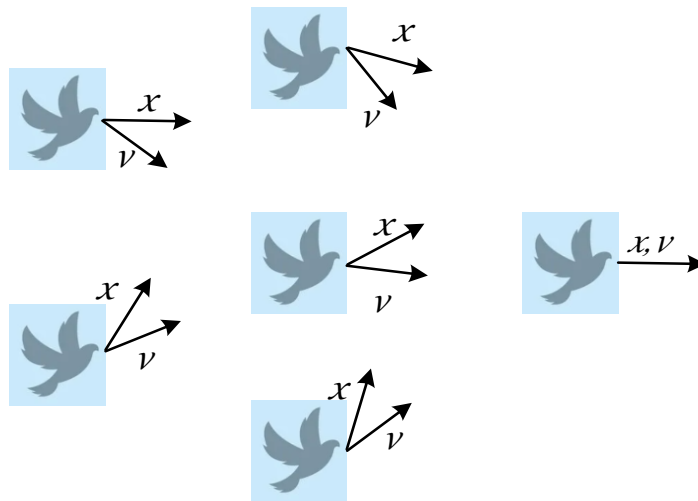


图 4.1 地图和指南针算子参与的迭代阶段

Fig. 4.1 The iterative stage of Map and compass operators

#### 4.1.2 地标算子

当鸽子飞近巢窠时, 所感受的磁偏角会缩小, 已经无法通过磁场完成接下来的飞行, 此时鸽群则会通过附近熟悉的地标来进行位置确认, 如果某些个体对附近地标十分熟悉, 那么这一部分的鸽子则会带领对周围的环境没有辨识度的剩余鸽群接近目的地, 而一些无法追随鸽群的个体则会被舍弃, 这个过程就是地标算子运算阶段, 实质上为鸽群算法局部寻优的阶段。该算子对算法的局部探索能力进行提升优化, 其参与运算的过程中, 鸽群数

量  $N_p$  每经过一次迭代都会减少一半, 仅保留距离巢窠最近一半个体, 并计算剩余鸽群的位置中心  $X^{(t)}(c)$ , 使鸽群朝向这一位置飞行, 此阶段所涉及的公式如下:

$$N_p(t) = \frac{N_p(t-1)}{2} \tag{4.5}$$

$$X^{(t)}(c) = \frac{\sum X^{(t)}(i) \text{fitness}(X^{(t)}(i))}{N_p(t) \sum \text{fitness}(X^{(t)}(i))} \tag{4.6}$$

$$X^{(t)}(i) = X^{(t-1)}(i) + \text{rand}(X^{(t)}(c) - X^{(t-1)}(i)) \tag{4.7}$$

其中,  $\text{fitness}(X^{(t)}(i))$  为鸽群当前位置适应度, 用来评价当前解的质量。

同遗传算法类似, 鸽群算法在求解函数最大最小值时,  $\text{fitness}(X^{(t)}(i))$  的公式也会随之调整, 当所求值为前者时, 一般采用公式(4.8); 当所求值为后者时, 常采用公式(4.9)。

$$\text{fitness}(X^{(t)}(i)) = f_{\max}(X^{(t)}(i)) \tag{4.8}$$

$$\text{fitness}(X^{(t)}(i)) = \frac{1}{f_{\min}(X^{(t)}(i)) + \varepsilon} \tag{4.9}$$

其中,  $f_{\min}$  与  $f_{\max}$  为鸽群中最小与最大适应度值;  $\varepsilon$  为无限接近于 0 的极小值。

由上述公式可以看出, 这一阶段个体自身速度不再参与运算, 在个体不被舍弃的前提下, 所有鸽子会逐渐的向位置中心飞行, 较上一个迭代过程, 鸽群可以更快地寻找到目的地。当循环迭代次数超过所规定的最大循环次数后停止迭代, 输出结果。地标算子参与的迭代阶段如图 4.2 所示。

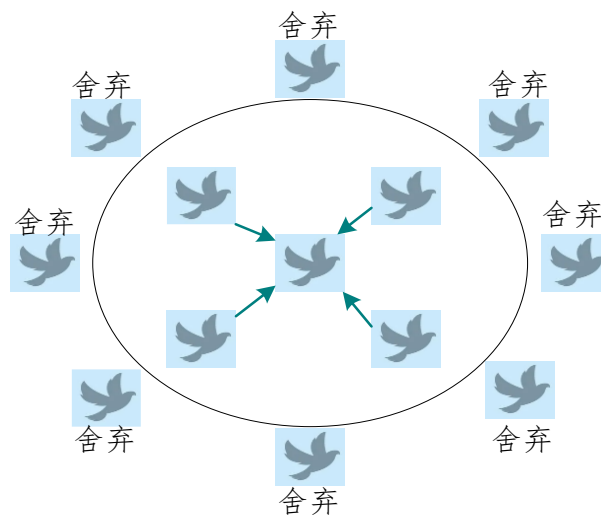


图 4.2 地标算子参与的迭代阶段

Fig. 4.2 The iterative stage of landmark operator

## 4.2 鸽群算法分析

### 4.2.1 鸽群算法局限性分析

通过对鸽群算法的原理分析,发现鸽群算法在进行寻优过程中的优点十分明显,具体有以下三点:

(1) 可扩展性强。为与其他智能算法的融合提供了有利条件,有利于发挥不同算法的优越性能。

(2) 全局搜索能力强。在地标算子阶段,利用鸽群之间的信息共享机制,适应度低的个体可以向适应度高的个体位置靠拢,对于复杂多极值问题,可根据其他鸽子搜索过的最优优点进行择优选取。

(3) 鸽群算法整个流程对比遗传算法来说需要设置的参数简单,鸽群整个寻优进化过程的相关操作也非常方便,能够以极快的速度寻找到最优值,缩短搜索时间。

而鸽群算法的缺点也是比较突出,具体表现在以下三个方面:

(1) 局部搜索能力不够灵活。鸽群算法一般采用的确定性规则完成对鸽群个体位置与速度的更新,这在一定程度上限制了种群的多样性,不利于跳出局部最优值。

(2) 鸽群算法是来源于对动物行为的模拟,由于提出时间较晚,所以算法理论还不够完备,而且现有的理论研究成果较少,从数理角度严格证明该算法的可行性和有效性还相对艰难。

(3) 目前对于鸽群优化算法的结构和参数选取一般采用试错法,通过不断的调整参数对比优化结果确定各关键因子,缺乏相应的实用性指导原则。

### 4.2.2 鸽群算法实现流程分析

研究分析鸽群算法原理,总结归纳了鸽群算法在解决函数优化问题时的流程,具体步骤如下:

步骤 1.环境初始化并设定适应度函数。

步骤 2.初始化鸽群优化算法并设定相关参数,包括指南针与地图算子最大迭代次数  $N_{c1max}$ , 地标算子最大迭代次数  $N_{c2max}$ , 初始种群总数  $N_p$ , 变量维度  $D$ , 指南针因子  $R$ , 问题上下界  $search\ range$ 。

步骤 3.为所有个体随机生成一个  $D$  维的初始位置  $X^0(i)$  和初始速度  $V^0(i)$ , 初始位置要在问题上下界之间, 初始速度在  $[0,1]$  范围内选取。根据选择的适应度函数将历史最优值初始化, 迭代次数  $N_c$  置为 1。

步骤 4.地图与指南针算子参与运算的迭代过程。在限制的循环次数  $N_{c1max}$  内, 利用公式(4.3)、(4.4)对每一个体进行速度与位置的更新, 评价种群中所有个体的位置优劣, 选取

处于较优位置的个体，再据此更新全局最优位置  $X_g$ 。

步骤 5.当循环迭代次数大于  $N_{c1max}$  时，进入地标算子迭代过程，如果鸽群没有达到最大循环次数  $N_{c1max}$ ，则返回步骤 4 继续进行当前迭代过程。

步骤 6.将经过指南针与地图算子迭代留下的种群作为地标迭代过程的初始种群，在限制的循环次数  $N_{c2max}$  内，继续迭代运算。

步骤 7.依据适应度值对鸽群进行优劣降序的排序，舍弃小于平均适应度的个体，更新种群数目并计算此时的位置中心，利用公式(4.6)、(4.7)调整个体位置和鸽群中心位置，令鸽群向中心位置靠拢。

步骤 8.当循环迭代次数超过  $N_{c2max}$  时，终止迭代过程并向外输出最终结果。否则，返回步骤 7 进行循环迭代。

图 4.3 为鸽群算法优化过程的具体流程图。

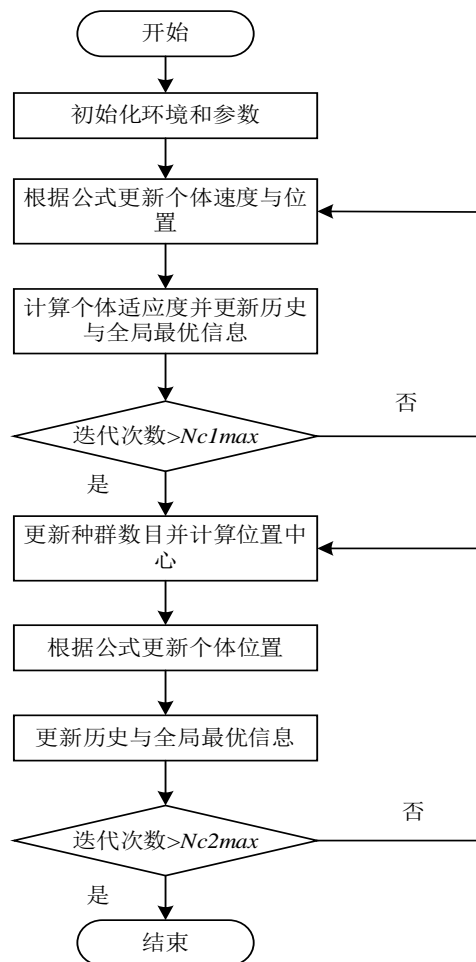


图 4.3 鸽群算法流程图

Fig. 4.3 Flow chart of pigeon-inspired optimization algorithm



### 4.2.3 鸽群算法参数的改进研究

鸽群算法在地图与指南针算子参与迭代的过程中只需要调整鸽群中鸽子的位置和速度这两个参数，实现过程较为简单，利于添加其他参数来增强 PIO 算法初始阶段的全局寻优能力，通过查阅资料阅读文献得知可以增加速度权重因子  $w$  调整不同个体的飞行速度<sup>[57]</sup>，使得适应度较低的飞行速度减缓，延长向最优位置靠拢时间，增加适应度较低的个体的飞行速度，缩短向最优位置的靠拢时间，从而在一定程度上提高鸽群再进行全局搜索最优位置时的多样性，减少陷入局部最优值的概率。惯性权重  $w$  因子的选取方法目前主要有线性调整策略与非线性调整策略。

线性调整的权重因子是以一个固定的步长进行调整，其调整方法如公式(4.10)所示。在地图和指南针算子参与运算的迭代过程前期，可以选取相对较大的权重提高鸽群的飞行速度，保证鸽子位置的多样性，增强全局搜索能力；在迭代后期，则可取较小的权重减缓鸽群飞行速度，保证迭代过程收敛精度，防止跳跃出最优解。但是，采用线性调整的策略在进行多峰问题寻优时，非常容易陷入局部极值，引发早熟现象。

$$w = w_{max} - \frac{N_c(w_{max} - w_{min})}{N_{c1max}} \quad (4.10)$$

其中， $w_{max}$  为最大权重因子，数值上等于起始权重值； $w_{min}$  为达到最大迭代次数时权重数值； $N_c$  为当前迭代次数。

非线性调整策略也是在前期赋予鸽群速度一个较大的权重，在迭代后期给与鸽群较小的速度，但是这种策略舍弃了一成不变的静态递减方式，而是根据其迭代过程动态调整权重因子的变化大小，从调整原理上来讲，采用动态调整的权重因子可以增加鸽群迭代过程的灵活性，常用的公式如(4.11)、(4.12)。

$$w = w_{min} \left( \frac{w_{max}}{w_{min}} \right)^{1/(1+10N_c/N_{c1max})} \quad (4.11)$$

$$w = w_{max} - (w_{max} - w_{min}) \left( \frac{1}{1 + e^{3.4 - 0.095N_c}} \right) \quad (4.12)$$

一般来讲，权重因子的范围一般设置在[0,1]的范围内选取，但是为了保证后期还有一定的跳出极值的能力，则在[0.4,0.9]范围内取值的权重因子对算法的优化效果最好，将第一个迭代过程最大迭代次数  $N_{c1max}$  设置为 70 时，利用以上三个调整公式改变权重因子数值大小的变化曲线如图 4.4 所示：

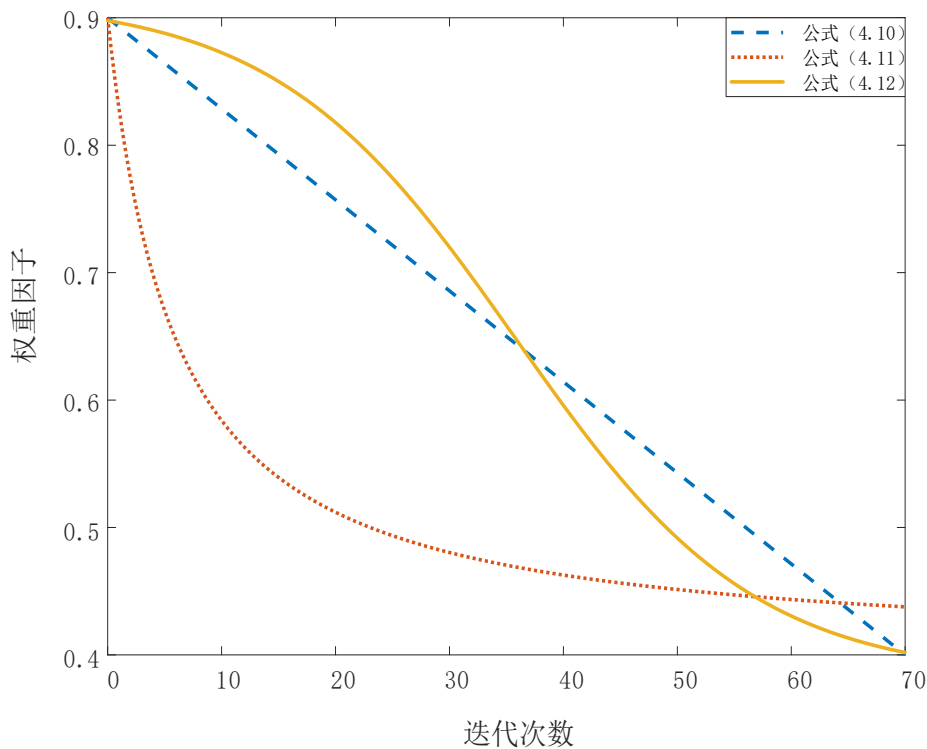


图 4.4 不同调整公式下的权重因子变化曲线

Fig. 4.4 Change curve of weighting factor under different adjustment formulas

从图中曲线可以看出，当使用公式(4.10)线性调整权重因子时，其下降速率不变；利用公式(4.11)实现对权重因子的调节时，在迭代前期下降速度相当迅速，但到了中后期逐渐趋向平稳，这种方法的优势在于，自迭代前中期开始，鸽群就开始寻找局部最优解，从而极大的提高了算法的局部寻优能力，但缺点是降低了前期全局搜索的效率；采用公式(4.12)的调整方法时，可以看出权重因子的调节呈反 S 型，前期保持较大的权重，确保全局鸽群的全局寻优能力，在中期权重因子迅速下降，快速度过中间过程，而在后期下降速度变缓，保证了鸽群的局部寻优能力。所以，本文采取了公式(4.12)作为惯性权重粒子的调节方式，添加惯性权重因子后的速度更新公式如下：

$$V^{(t)}(i) = wV^{(t-1)}(i)e^{-R} + rand(X_g - X^{(t-1)}(i)), i = 1, 2 \dots N_p \quad (4.13)$$

### 4.3 混合遗传算法优化 BP 神经网络研究

#### 4.3.1 遗传算法优化 BP 神经网络时存在的缺陷分析

在本文第三章中，顺利完成了遗传算法对于 BP 神经网络的优化，并且取得了较好的效果，但是从遗传算法本身来讲，还是有着以下几点缺陷：

(1) 极易“早熟”。对于遗传算法来讲,初始染色体种群的随机选取使得算法的种群个体可能全部散布在局部最优值周围,在处理非线性多极值问题时,早熟现象表现的尤为突出,使算法容易收敛到局部最优解。

(2) 实现过程复杂。在进行最优值寻找时,遗传算法首先要对初始种群进行编码操作,而在找到最优解之后还要完成解码操作才能输出。随着样本数据的增多,问题复杂程度直线上升,相比较其他智能优化算法,遗传算法计算效率要远低于粒子群算法、鸽群算法等。

(3) 遍历性差。遗传算法属于随机类算法,因此对于初始染色体种群规模和质量的选取有着一定的依赖性,就如何保证 BP 神经网络的初始权值阈值最适合进行网络学习这一问题而言,在最不利的情况下,遗传算法初始染色体种群全都落在了 $[-0.2,0.2]$ 的区间内,在后续的寻优过程中,该算法很难完成 $[-1,1]$ 整个区间的搜索。

(4) 遗传算法的全局搜索性能并不稳定。遗传算法在全局搜索最优解的阶段,依赖于交叉和变异算子的取值,但是通常情况下,交叉变异的概率仅为 1%到 8%左右,多次运算时,其最优值的结果也不稳定。

#### 4.3.2 混合遗传算法的实现研究

鸽群算法作为最新提出的一种仿真智能算法,其具有的优点非常明显。为了验证鸽群算法与遗传算法深度融合的可行性,着手从上文提及到的两种算法的优缺点进行分析讨论。

对比分析遗传算法与鸽群算法在处理函数问题时的优缺点可以发现,鸽群算法可扩展性为与遗传算法的融合提供了可能性,PIO 算法后期的个体的记忆能力提高了算法全局探索最优值的能力,且 PIO 算法需要的设置的参数简单,其算法实现过程所需时间远远小于遗传算法染色体进化过程。与之相反的,遗传算法的局部寻优能力十分优秀,但是受制于染色体种群数量及进化代数的困境,导致算法实现过程所需时间过长<sup>[53]</sup>,综上考虑,推断出鸽群算法与遗传算法的结合可以从以下两个方面进行:

##### (1) 鸽群算法与遗传算法深度融合

交叉和变异因子是决定遗传算法全局与局部寻优能力的关键,而在传统遗传算法的基础上引入鸽群算法中鸽子位置与速度的更新方法是两个算法深度融合的实现思想,利用鸽群两个迭代过程中的位置与速度更新方法,替换交叉和变异概率原有的固定取值,使其交叉变异概率的取值随着种群适应度不断地改变,这个方法的优势在于,摒弃了原有的固定取值模式,利用了鸽群中不同个体之间的信息交互功能完善遗传算法中染色体的信息传递与学习,这种融合方式可以在一定程度上增强了寻找局部最优的能力。

##### (2) 鸽群算法与遗传算法并行融合

遗传算法求解问题所需时间与性能与初始种群数量与所设定进化代数有着正相关的关

系,一般来说,数据样本越复杂需要设置的初始种群数量及所需进化代数就越多,求解问题所需时间就会更长。同时,针对非线性多峰值优化问题,遗传算法随机性选取种群的特点,增大了原始种群全部落在局部最优值附近的概率,不利于问题的顺利求解。基于以上分析,本文所提出的鸽群算法与遗传算法并行融合的思想在于利用鸽群算法全局寻优能力速度快的特性,先行为遗传算法划定初始种群范围,然后再由遗传算法在这个范围内进行优化过程的操作,这种算法融合的思想优势在于一方面充分发挥了 PIO 算法运行速度快,全局寻优能力强,另一方面结合了 GA 算法后期局部寻优能力灵活的特点,能够解决单一算法处理问题时的不足之处。

考虑到遗传算法交叉和变异因子的选取可以采用自适应调整公式进行调节,则第一种融合方式缺乏实施意义,并没有很好的发挥两种算法的不同优势,所以本文采用提出的第二种融合方法,形成一种新的混合遗传算法优化 BP 神经网络,最终实现新型网络训练迭代时间短,收敛速度快,增强网络跳出局部极值并达到较高精度的能力。

#### 4.3.3 混合遗传算法优化 BP 神经网络

鸽群算法与遗传算法并行融合后形成的混合遗传算法优化 BP 神经网络的实现思路是:先确定 BP 神经网络的第一组权值与阈值的选取范围,在这一范围中,首先选择一定数量的不同数值作为 PIO 算法的初始种群,然后依次参与两个算子的迭代过程,在经过鸽群算法若干次迭代之后,保留的个体将转向遗传算法的继承或进化过程,最后将适应度最优的染色体解码赋值给 BP 神经网络。

考虑到鸽群优化算法在地标算子参与迭代阶段,鸽群数量会大幅削减,容易导致传递给遗传算法的初始种群数量过少,导致算法崩溃,于是在本文提出的混合遗传算法中,舍弃了鸽群第二个迭代过程中利用公式(4.5)对于鸽群数量的控制方式,而是采用通过适应度排名的方法选择最终保留群体,从而使这种融合方式可以充分发挥鸽群算法全局寻优快以及算法迭代时间短的特点,在一定程度上缩减遗传算法的初始种群数,大大减少遗传算法寻优所用时间,混合遗传算法的主要步骤如下:

步骤 1.设定遗传算法与鸽群算法的初始参数。遗传算法与 PIO 算法主要参数在上文中均以提及,此处不再赘述。

步骤 2.鸽群个体的初始位置与速度初始化。

步骤 3.计算个体的适应度,更新全局最优适应度,进行地图和指南针算子参与迭代过程。

步骤 4.当最大迭代次数大于  $N_{c1max}$  时,将鸽子个体按照适应度由大到小顺序排列,执行步骤 5,如果小于最大迭代次数,返回步骤 3 循环迭代。

步骤 5.更新个体位置,并计算鸽群位置中心。

步骤 6.判断迭代次数是否大于  $N_{c2max}$ ,如不满足,转到步骤 6 循环运算过程,如果满

足，则进行步骤 7。

步骤 7.将这些鸽群个体按适应度排名,选择前  $N$  位的鸽子作为遗传算法的初始染色体,进行编码操作。

步骤 8.染色体进行选择、遗传、变异操作。

步骤 9.计算适应度大小,如果适应度满足要求,则进行步骤 10, 否则返回步骤 8 循环过程。

步骤 10.完成对染色体的解码,将最优权值阈值赋值给 BP 神经网络。

鸽群算法与遗传算法并行融合优化 BP 神经网络的流程如图 4.5 所示。

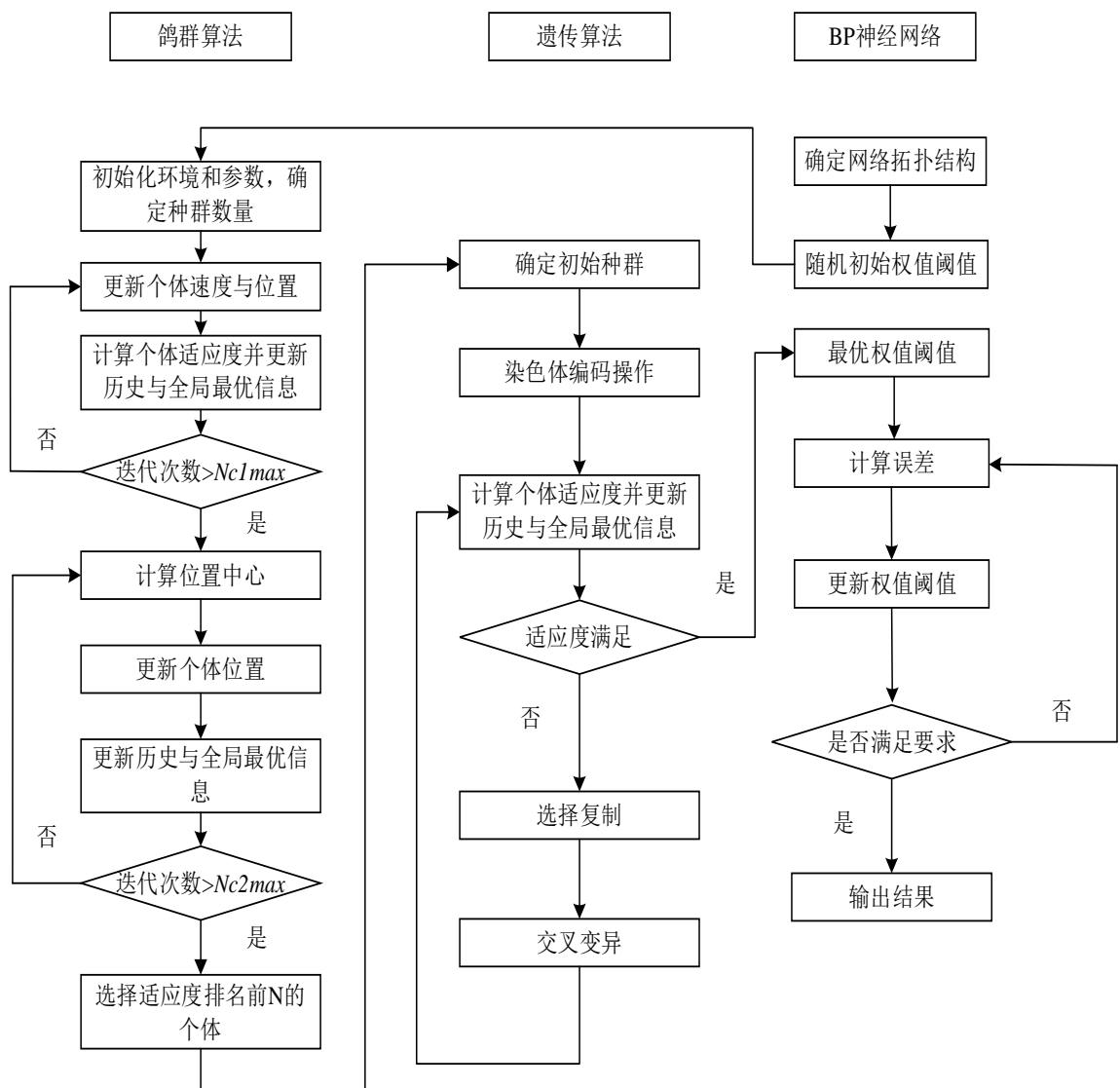


图 4.5 混合遗传算法优化 BP 神经网络流程图

Fig. 4.5 Flow chart of BP neural network optimized by hybrid genetic algorithm

#### 4.4 本章小结

针对能否引入一种新的算法来使 GA-BP 神经网络光伏阵列故障检测模型的性能变得更为出色这一问题, 本章创新性地提出将鸽群算法与遗传算法并行融合形成的混合遗传算法用于 BP 神经网络的优化。首先研究了鸽群算法原理及优化过程, 并对鸽群优化算法的速度调整公式进行了改进, 紧接着分析了鸽群算法与遗传算法的优缺点。在此基础上, 研究了鸽群优化算法与遗传算法的两种融合方式, 最后给出了混合遗传算法优化 BP 神经网络的具体实现步骤与流程。

## 第五章 改进型 GA-BP 神经网络在线光伏阵列故障检测

### 仿真验证分析

#### 5.1 改进型 GA-BP 神经网络实验参数设定

##### 5.1.1 输入输出变量的选取

在验证融入鸽群算法后的 GA-BP 神经网络光伏阵列故障检测性能实验中, 本章节设计改进后的 GA-BP 神经网络结构与前文使用的 GA-BP 神经网络结构一致, 为 4\*9\*5 的网络结构, 改进型 GA-BP 神经网络输入仍然选择光伏阵列开路电压  $V_{oc}$ , 短路电流  $I_{sc}$ , 最大功率点处电压  $V_{mmp}$ , 最大功率点处电流  $I_{mmp}$ 。

前文已经验证了 GA-BP 神经网络用于光伏阵列故障检测的可行性, 为了消除样本数量对于改进前后 GA-BP 神经网络训练误差的影响, 确保对比实验的科学性, 故在这一章的实验数据处理的选择上仍是使用光伏阵列在不同运行状况时的 850 组数据作为训练样本, 选择 150 组数据作为研究改进前后网络故障识别正确率的验证数据。

神经网络的不同输出状态代表着光伏阵列的不同运行情况, 如阵列产生短路故障、开路故障、阴影遮挡以及光伏电池单体老化等, 本章主要对改进前后的 GA-BP 神经网络光伏故障检测模型进行性能对比, 所以, 在两个模型的输出选择上要保持一致, 部分检测样本数据及对应的网络正确输出状态如表 5.1 所示。

表 5.1 部分样本数据及对应网络输出值

Tab. 5.1 Partial sample data and corresponding network output values

类别编号	故障类别	$V_{oc}(V)$	$I_{sc}(A)$	$V_{mmp}(V)$	$I_{mmp}(A)$	网络输出值
A1	短路故障	149.37	15.52	118.77	13.94	1
A2	短路故障	99.79	15.51	80.08	14.12	1
B1	开路故障	178.59	10.54	151.33	10.08	2
B2	开路故障	178.69	5.26	151.25	4.25	2
C1	阴影遮挡	176.79	15.48	118.51	14.61	3
C2	阴影遮挡	177.45	15.47	151.65	10.38	3
D1	老化故障	176.35	15.3	147.22	13.69	4
D2	老化故障	178.57	15.44	126.95	11.64	4
E	无故障	178.66	15.64	150.32	14.59	5

### 5.1.2 改进型 GA-BP 神经网络各组成部分参数的选取

改进型 GA-BP 神经网络模型由鸽群算法、遗传算法以及 BP 神经网络三个部分组成，同时需要对比改进前后的 GA-BP 神经网络在该实验中的准确性与快速性，所以整个实验的参数设定分为鸽群算法、遗传算法以及 BP 神经网络的设定三部分，BP 神经网络结构与实验参数均与上文相同，在此不在赘述，而采用混合遗传算法时，为增强改进前后的对比性，要保证 PIO 算法最终输出种群数要与 GA 算法的初始种群数保持一致，遗传算法与鸽群优化算法的具体参数设置如表 5.2、表 5.3 所示。

表 5.2 鸽群优化算法参数设定

Tab. 5.2 Parameter settings of PIO algorithm

算法参数	数值
初始鸽群数目 $N_p$	100
维度 $D$	95
地图指南针算子 $R$	0.2
初始权重因子 $W_{max}$	0.9
最终权重因子 $W_{min}$	0.4
地图和指南针算子最大迭代次数 $N_{c1max}$	70
地标算子最大迭代次数 $N_{c2max}$	30
适应度排名前 $N$ 位保留	50
搜索空间下限	-100
搜索空间上限	100

表 5.3 遗传算法参数设定

Tab. 5.3 Parameter settings of Genetic algorithm

算法参数	数值
染色体种群规模 $N$	50
交叉概率最小值 $P_{cmin}$	0.3
交叉概率最大值 $P_{cmax}$	0.1
变异概率最小值 $P_{mmin}$	0.01
变异概率最大值 $P_{mmax}$	0.1
进化代数	50



## 5.2 改进型 GA-BP 神经网络应用于光伏阵列故障检测实验结果与分析

### 5.2.1 适应度对比结果分析

分析改进 GA-BP 神经网络前后, BP 神经网络的最优权值阈值分别由遗传算法和混合遗传算法给予, 所以为了保证对比结果的科学性, 本节选用 BP 神经网络期望输出与实际输出差值之和的倒数作为两种算法适应度函数, 其数学表达式已经由公式(3.24)给出, 且适应度越高, 说明误差越小。

原遗传算法在进行 BP 神经网络最优初始权值阈值的寻优过程中权值阈值适应度值与种群进化代数的关系如图 5.1 所示。

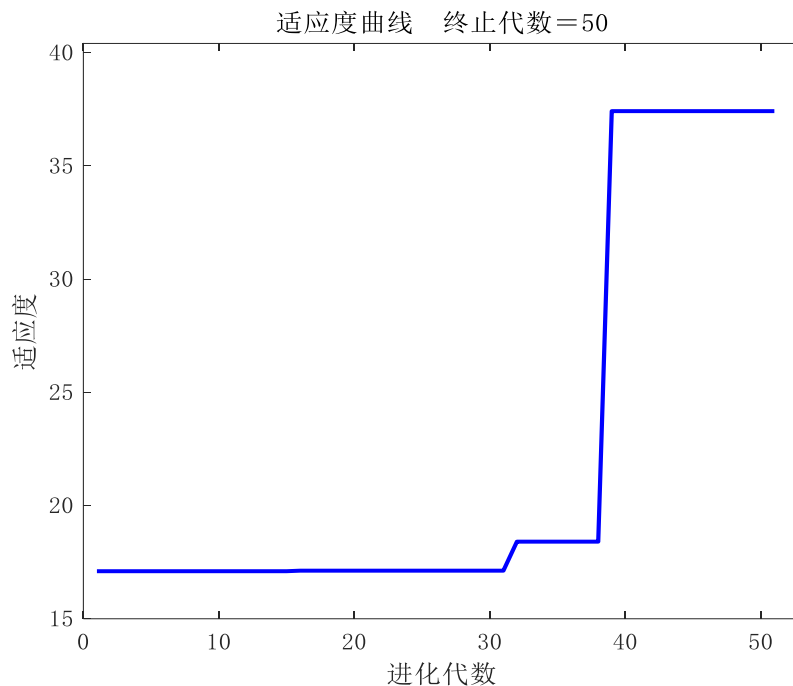


图 5.1 遗传算法适应度曲线

Fig. 5.1 Fitness curve of genetic algorithm

由图 5.1 可知, 遗传算法经过了 2 次适应度的变化, 在染色体进化 38 代时, 适应度到达最大, 数值为 37.42, 该适应度对应的期望输出与实际输出的误差和约为 0.027, 而且随着迭代次数的增加, 适应度值不再变化, 说明此时遗传算法已经找到网络的所需的一组最优初始解, 并将寻优获得的最优解作为 BP 神经网络的初始权值阈值完成对训练样本数据集的学习。

图 5.2 为混合遗传算法在进行 BP 神经网络最优权值阈值选取过程中, 染色体进化代数与适应度的变化曲线。

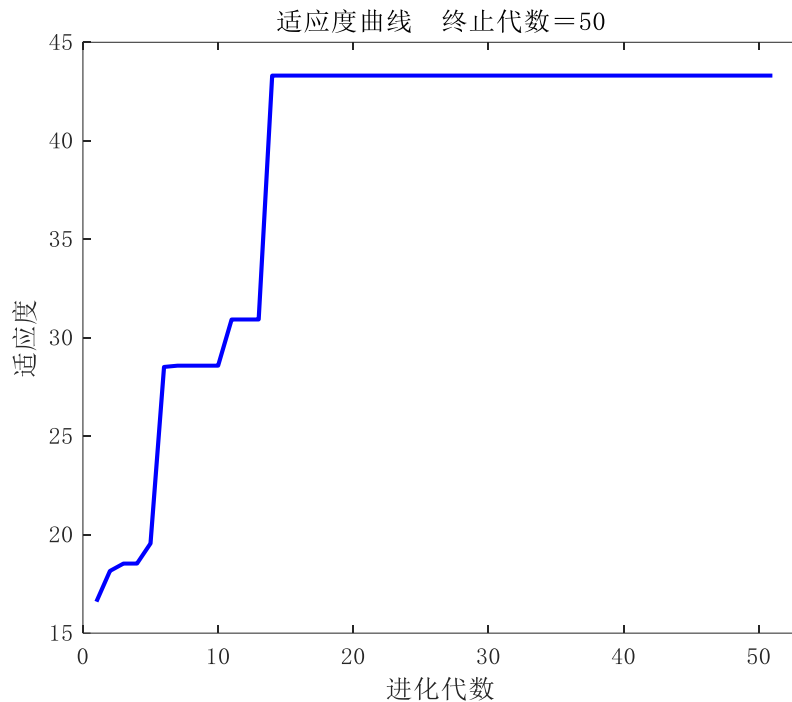


图 5.2 混合遗传算法的适应度曲线

Fig. 5.2 Fitness curve of Hybrid genetic algorithm

分析图 5.2 可知,混合遗传算法适应度经过了 6 次变化,染色体种群仅仅进化 17 次便达到最大值 43.31,此时对应的期望输出与实际输出间的误差和约为 0.023,从 18 代至第 50 代,染色体适应度不再变化,说明种群不在进化,已经完成了对 BP 神经网络最优权值阈值的筛选。

对比遗传算法与混合遗传算法个体适应度与进化代数的关系发现,结合 PIO 算法的混合遗传算法,在染色体进化过程中,遍历性得到提高,跳出局部极值的能力得到加强,其适应度值所对应的网络误差总和减少,而且所需进化代数更少,所以,在参数设置时,对于相同的初始种群数,完全可以通过减少其进化代数,缩减算法所需时间。

### 5.2.2 改进型 GA-BP 神经网络模型训练

在参数设定完成之后,将采集到的 850 组训练数据分别交给 GA-BP 神经网络和经过改进后的神经网络完成网络的训练过程,两个网络训练迭代次数与所能达到的最小均方误差的关系曲线如图 5.3 与图 5.4 所示。

从图 5.3 中可以看出,GA-BP 神经网络在进行了 61 次迭代之后,整个网络对于训练数据的均方误差达到了 0.00070766,且随着之后迭代次数的增加,误差减小的趋势基本保持不变,说明此时网络已经完成了训练过程。

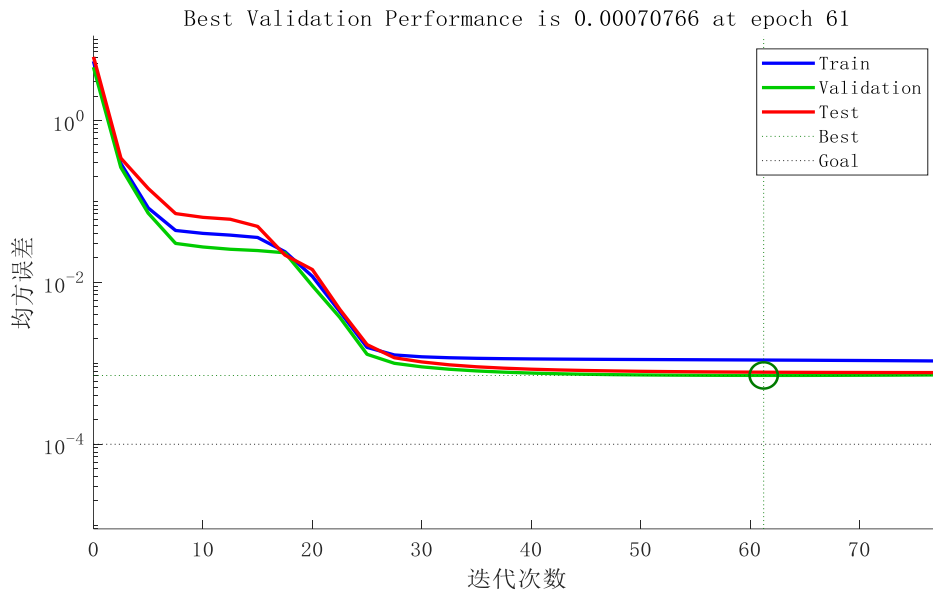


图 5.3 遗传算法优化的 BP 神经网络训练误差曲线

Fig. 5.3 BP neural network training error curve by genetic algorithm

而从图 5.4 可以看出，混合遗传算法优化的 BP 神经网络在网络训练 42 次之后，训练误差就达到了 0.00023121，且误差减小趋势基本不再变化，此时网络完成了对数据的训练过程，较未改进之前所需迭代次数缩短了约 1/3 左右，而较原 BP 神经网络所需迭代次数近乎缩减了一半，由此可以得出，改进后 GA-BP 神经网络在训练快速性方面胜于原始 GA-BP 神经网络。

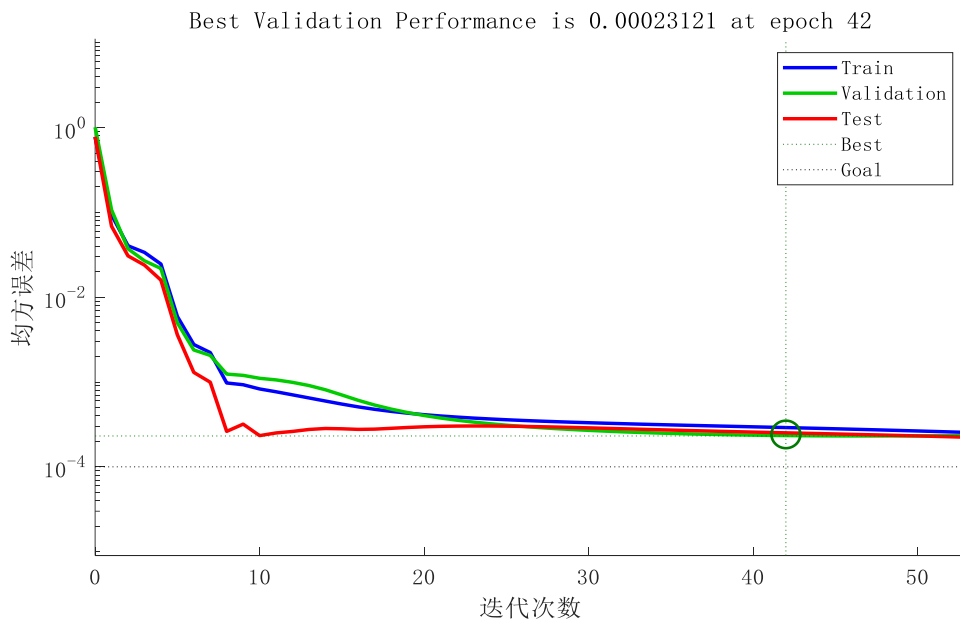


图 5.4 混合遗传算法优化 BP 神经网络训练误差曲线

Fig. 5.4 BP neural network training error curve optimized by hybrid genetic algorithm

### 5.2.3 改进型 GA-BP 神经网络在光伏阵列故障检测中的实验结果与分析

考虑到在实际应用中，光伏电站运行时的历史数据较多，这就意味着用于 BP 神经网络训练的数据集数量十分充足，而需要检测的数据一般为一天或两天内的实时数据，与训练数据集要少很多，且由第三章所作实验可知，在针对不同样本容量的数据集进行检测分析时，大容量样本的检测正确率略高于小容量样本，所以本文选取了 150 组不同运行状态下的实时数据用来验证改进后的 GA-BP 神经网络性能，在这 150 组数据中，为了更好的反映改进前后网络的性能差异，体现实验对比的公平性，将每种故障类型的数据占比设为五分之一，即每种类型的故障数据为 30 组，然后将与 BP 神经网络以及 GA-BP 神经网络进行检测时所使用的 150 组数据交由混合遗传算法优化 BP 神经网络进行检测，故障检测结果如图 5.5 所示。

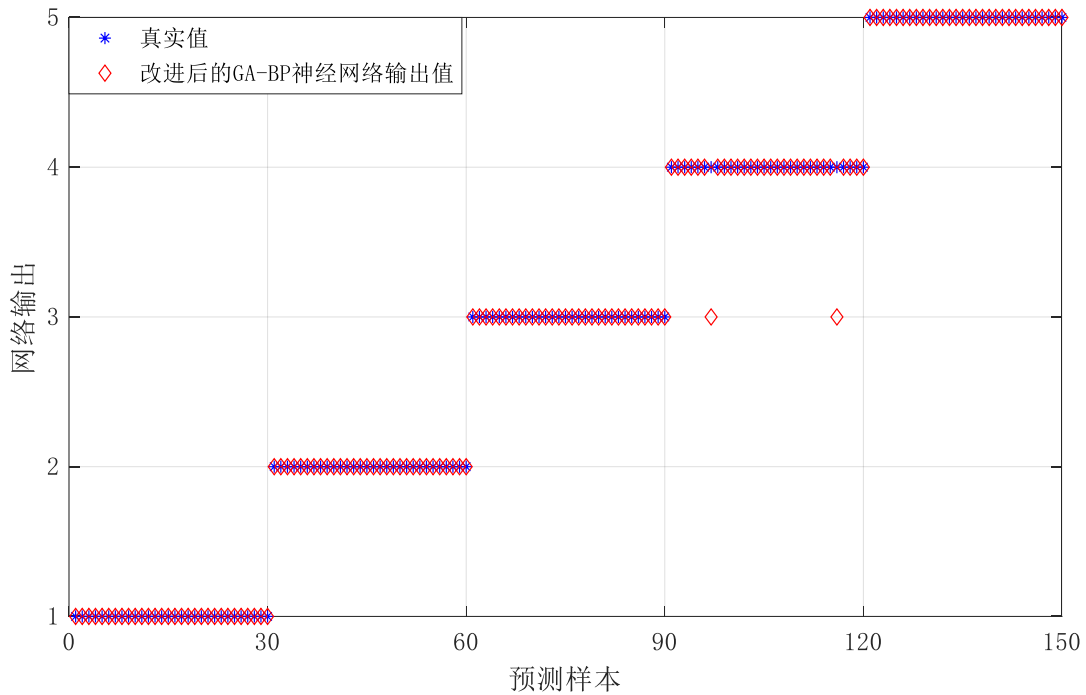


图 5.5 改进后的 GA-BP 神经网络故障检测结果

Fig. 5.5 Detection results of improved GA-BP network

分析图中数据可知，改进后的 BP 神经网络对于开路和短路类型的故障识别率非常高，这是由于无论是对于单组串开路或者多组串开路的开路故障，亦或是单电池短路或多电池短路的短路故障，其电压电流变化特征明显，具有非常鲜明的特性，所以两种网络识别结果特别理想，但是遮阴故障以及老化故障这两种类型的故障数据在发生单电池故障时，其特征参数变化量并不相同，此时对于网络的精度要求并不高，一旦发生多电池遮阴或者老化故障，阵列的特征参数会出现一定的重合区，对于网络的精度要求非常高。这也是存在对电池阴影遮挡数据检测错误的主要原因。

将改进前后的 GA-BP 神经网络故障检测结果进行统计,统计结果如表 5.4、表 5.5 所示,分析表中数据可知,混合遗传算法优化的 BP 神经网络正确完成了对 148 组故障数据的识别,只将 2 组阴影遮挡故障数据检测判别为老化故障,不存在对其他故障类型的误检,检测正确率达到了 98.3%,高于原始 GA-BP 神经网络的 96%故障识别率,说明经过混合遗传算法优化后的 BP 神经网络的误差精度更高,对于各种故障数据集的识别更为有效,更适合进行光伏阵列的故障检测。

表 5.4 遗传算法优化的 BP 神经网络检测结果统计

Tab. 5.4 Statistics of BP Neural Network Detection Results Optimized by Genetic Algorithm

故障状态	验证样本数量	正确检测组数	错误检测组数	正确率
短路故障	30	30	0	
开路故障	30	29	1	
阴影遮挡	30	28	2	96%
老化故障	30	27	3	
正常无故障	30	30	0	

表 5.5 混合遗传算法优化的 BP 神经网络检测结果统计

Tab. 5.5 Statistics of BP Neural Network Detection Results Optimized by Hybrid Genetic Algorithm

故障状态	验证样本数量	正确检测组数	错误检测组数	正确率
短路故障	30	30	0	
开路故障	30	30	0	
阴影遮挡	30	30	0	98.7%
老化故障	30	28	2	
正常无故障	30	30	0	

分析以上实验结果发现,改进后的 GA-BP 神经网络对于多种故障的检测有着更高的准确率,而两者检测精准度的差异主要体现在对阴影遮挡与老化故障的数据识别方面,在剩余两种故障类型和正常数据的检测识别中,有混合遗传算法参与的 BP 神经网络要优于未改进之前的网络,其检测可靠性得到了进一步的提高。

为了进一步研究改进前后的 GA-BP 神经网络在检测老化和阴影故障时的性能优劣,本文单独采集了 150 组故障数据作为验证数据集,其中老化故障数据 60 组,阴影遮挡故障

数据 60 组，正常无故障数据 30 组，考虑到实际检测过程中受限于时间的紧迫性及故障类型的不确定性，一般不会针对以上两种故障情况而单独进行网络的训练，所以采用上文已经完成训练的神经网络进行故障数据的检测，实验结果如图 5.6、图 5.7 所示。

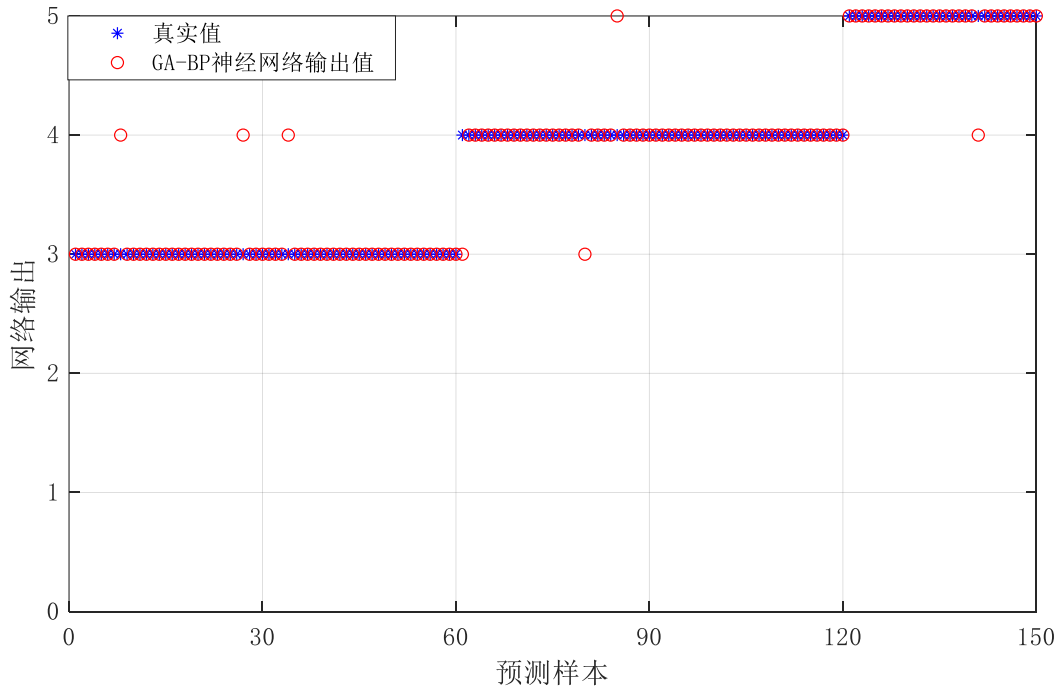


图 5.6 GA-BP 神经网络故障识别结果

Fig.5.6 Results of GA-BP neural network fault identification

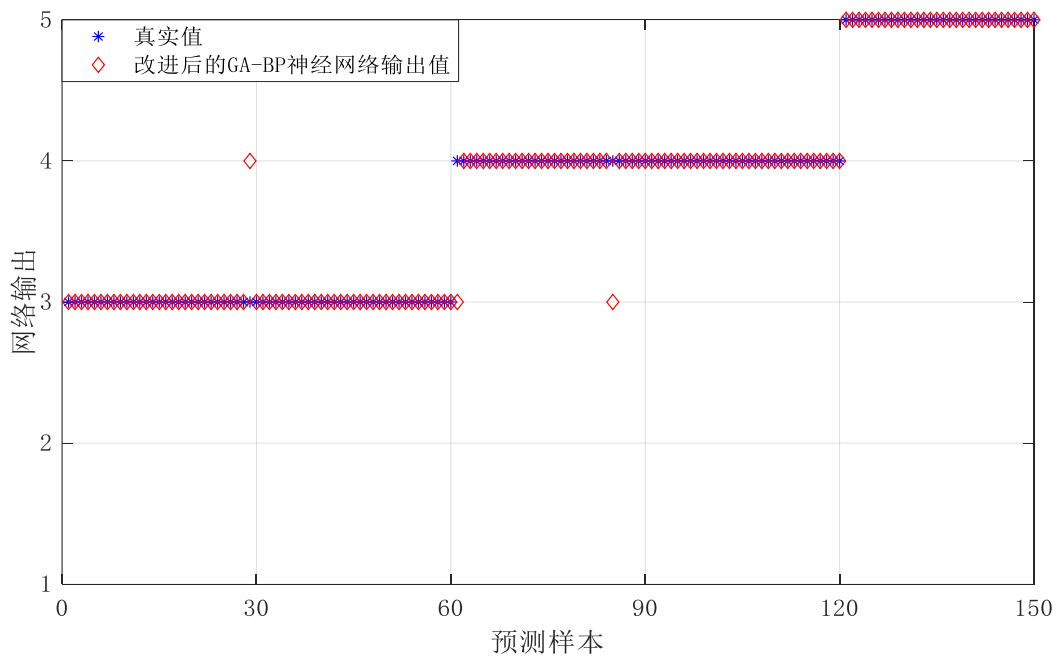


图 5.7 改进后的 GA-BP 神经网络故障识别结果

Fig.5.7 Results of improved GA-BP neural network fault identification

分析图 5.6 可知,原 GA-BP 神经网络在对老化故障和阴影遮挡故障进行检测时,共有 7 组数据检测错误,在这 7 组数据中,将 3 组老化故障误检为阴影遮挡,将 3 组阴影遮挡数据检测为老化故障,将 1 组阴影遮挡数据判别为正常数据,最终的检测准确率只有 95.3%,可见原 GA-BP 神经网络对于上述两种故障类型的检测结果并不理想。而从图 5.7 可以看出,经混合遗传算法优化后的 BP 神经网络对上述两种故障的检测性能有了显著的提升,只有 3 组数据检测错误,正确检测率达到了 98%,且没有将故障数据判别为正常数据,从可靠性而言,要优于原 GA-BP 神经网络。

根据以上实验分析,在 BP 神经网络中引入的混合遗传算法较于之前原始遗传算法,有着更快的染色体种群进化速度,从而以更短时间完成赋值操作,并且这一组权值阈值的使用,十分明显的缩减了神经网络的训练时间,且从对老化与阴影遮挡故障的检测实验中可以看出,网络的训练误差精度也得以提升,对于故障数据集的识别更加精准,有力的证明了本文提出的混合遗传算法优化 BP 神经网络模型用于光伏阵列故障检测的可行性与优异性。

### 5.3 本章小结

本章对经过混合遗传算法优化 BP 神经网络在光伏阵列故障检测中的性能进行了仿真实验分析,首先在混合遗传算法与原遗传算法在进行个体寻优阶段,对染色体适应度及所需进化代数进行了对比,发现混合遗传算法所需进化代数更少,个体适应度更高,更不容易陷入局部最优解,在一定程度上减少了整个检测过程所需时间,再与原 GA-BP 神经网络性能进行了对比,实验结果表明,无论是在神经网络训练收敛性能上,还是故障预测精度上,改进后的 GA-BP 神经网络模型均优于原始模型。

## 第六章 总结与展望

### 6.1 总结

光伏电站一般建立在环境较为恶劣的地方,在工作过程中非常容易发生故障,为了提高光伏发电系统运行稳定性,如何快速有效的实现光伏发电系统故障检测,成为了当前必须解决的问题,围绕这一问题,本文对光伏阵列故障检测技术做了深入研究,主要进行的工作以及获得的研究成果总结如下:

(1) 经过对目前光伏发电系统的所处背景深入了解,总结分析了目前国内外光电检测技术的研究现状,最终选择基于智能网络检测的方法进行本课题的研究。

(2) 通过研究光伏电池的工作原理,对光伏电站实际工作过程中故障形成的主要因素及故障状况下阵列输出特性的变化规律做了细致的分析,并结合对实际光伏电站的实地考察,完成了4行3列串并联结构的光伏阵列仿真模型的搭建。并且通过对搭建阵列设置故障条件,完成了实际光伏阵列运行时可能会出现短路、开路、阴影遮挡和老化故障的模拟,提取到了故障实时运行数据。

(3) 完成了GA-BP神经网络在光伏阵列故障检测中的应用研究。BP神经网络检测性能与其初始权值阈值的选取有很大的关系,为了改善这一缺陷,引入了遗传算法来为BP神经网络优选初始权值阈值,从而建立起应用于光伏发电系统的GA-BP神经网络故障检测模型。首先对应用于光伏阵列故障检测的BP神经网络结构做了深入研究,然后根据不同故障状态下阵列参数的变化,选取了检测模型的输入量。最后,从建立的光伏阵列中提取到了不同运行工况下的数据,交由故障检测模型进行检测识别,从而判断阵列故障状态,最终结果表明,无论是对于小容量样本还是大容量样本,经遗传算法优化后的BP神经网络,都可以顺利完成光伏阵列故障检测的任务,并且对于故障的检测识别率分别为95%和96%,均优于原始BP神经网络。

(4) 针对GA-BP神经网络检测速度、精度如何进一步提高这一问题,提出了一种混合遗传算法优化BP神经网络的解决方案。利用鸽群算法求解最优化问题时,全局搜索能力迅速这一优点,为遗传算法优选初始种群的选择范围,缩减染色体进化所需代数,从而更快速的将更适合的权值阈值提供给BP神经网络,顺利实现了算法的融合。

(5) 验证了混合遗传算法优化的BP神经网络在光伏阵列的故障检测识别中的应用的有效性。通过实验发现,改进后的GA-BP神经网络比原始网络训练过程加快了一倍左右,且对于故障的检测准确率得到了提高,达到了98.7%,而且在单独进行电池老化与阴影遮挡故障的检测实验中证明,本文所提出的混合遗传算法优化的BP神经网络在光伏阵列在线故障检测中的性能表现的更加优异,准确率达到98%。



## 6.2 展望

近几年光伏行业在我国呈现出爆发式增长的态势，也推动着我国能源消费结构的巨大转变，不过受限于配套技术不够完善，光伏发电系统的运行安全性、可靠性还有待提高，因此，光伏阵列在线故障检测技术的研究对光伏产业的健康发展显得尤为重要。本文所研究的课题在实验室中取得了较为理想的成果，但是由于自身知识储备以及客观实践条件等问题的限制，对于本课题的研究还不够完善，仍有一些欠缺之处值得进一步的研究：

(1) 本文的实验是在 MATLAB 软件中搭建的光伏阵列模型的基础上进行的，设置的环境条件较为理想，收集到的实验数据与现实电厂所能采集到的实时数据存在一定的偏差，接下来应该将所建立模型套用在实际光伏电厂中，分析本文所建立模型在实际应用的性能表现，进行下一步的改进。

(2) 本文只是对单一的故障类型的检测进行了仿真实验，实际系统中存在多重复杂故障出现的可能性，对于多重故障的检测识别也应该是今后进一步研究的方向。

## 参考文献

- [1] 吴刘星,边雪晴. 中国化石能源消耗与空气污染和经济发展的关联性研究[J].现代商业,2015(24):262-264.
- [2] 李德森. 探讨我国能源革命——从化石能源到新能源[J].科学家,2017,5(17):84-85.
- [3] 教煥宗. 基于虚拟同步机控制的双馈风电并网运行关键技术研究[D].浙江大学,2021.
- [4] 饶宏. 数字电网推动构建以新能源为主体的新型电力系统[J].电力设备管理,2021,60(08):21-22.
- [5] 本报评论部. 正确认识和把握碳达峰碳中和[N].人民日报,2022-02-11(005).
- [6] 国家统计局. 中华人民共和国 2020 年国民经济和社会发展统计公报(1)[N].人民日报,2021-03-01(010).
- [7] L.L.Jiang and D.L. Maskell, "Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods," 2015 International Joint Conference on Neural Networks (IJCNN),Killamey,2015,pp.1-8.
- [8] 本报评论部. 安全无小事! 2018 上半年发生的光伏事故汇总[N].北极星太阳能光伏网,2018-07-27(006).
- [9] 易威. 光伏阵列的参数辨识及故障诊断研究[D].浙江大学,2021.
- [10] Chne W. MelitA, ManPA, et al Fautd detecton method for end-conneded photovoltaicplant[J]. Renewable Energy,2014,66(3):99-110.
- [11] Dotenco S, Dalsass M, Winkler L, et al. Automatic detection and analysis of photovoltaic modules in acrialinfared imagery[C]. Applications of Computer Vision IEEE,2016:19.
- [12] Alsafasfeh M. Abdel-Qader L Bazuin B. Fault detecton in photovoltaic system using SLIC andther malimages[C] Intemational Conference on Infomation Technology,2017,672-676.
- [13] 栾艳军. 光伏组件航拍红外图像的故障智能识别[D].河北大学,2021.
- [14] Silvestre S, Silva M.A.D, Chouder A, et al. New procedure for fault detection in grid connected P Vsystems based on the evaluation of current and voltageindicators[J]. Energy Conversion & Management,2014,86(10):241-249.
- [15] Harrou F, Sun Y, Taghezouit B, et al. Reliable fault detection and diagnosis of photovoltaicsystems based on statistical monitoring approaches[J]. RenewableEnergy,2018,116(pt.a):22-37.
- [16] 刘强,郭珂,毛明轩,万怡浩,周思宇,吴昊. 一种基于串联等效电阻的光伏故障检测方法[J].太阳能学报,2020,41(10):119-126.
- [17] 王元章,李智华,吴春华. 基于 BP 神经网络的光伏组件在线故障诊断[J].电网技术,2013,37(8):2094-2100.

- [18] Massi Pavan A, Mellit A, De Pieri D, et al. A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants[J]. Applied Energy, 2013, 108(26):392-401.
- [19] 乐治后. 基于 RBF 神经网络的光伏电池故障诊断[J]. 上海电气技术, 2015, 8(03):28-32.
- [20] Dhimish M, Holmes V, Mehrdadi B, et al. Diagnostic method for photovoltaic systems based on signal layer detection algorithm[J]. Electric Power Systems Research, 2017, 5(151):26-39.
- [21] 吴擎. 基于深度学习的光伏发电系统故障诊断方法研究[D]. 山东大学, 2019.
- [22] 马辉. 基于随机森林的光伏电站结构故障诊断与分类研究[D]. 西安理工大学, 2021.
- [23] 李兵峰. 光伏阵列故障检测方法的研究[D]. 天津大学, 2010.
- [24] 程泽, 李兵峰, 刘力, 等. 一种新型结构的光伏阵列故障检测方法[J]. 电子测量与仪器学报, 2010, 24(02):131-136.
- [25] 胡义华, 陈昊, 徐瑞东, 等. 基于最优传感器配置的光伏阵列故障诊断[J]. 中国电机工程学报, 2011, 31(33):19-30.
- [26] 唐冀, 朱永强, 郝嘉诚. 基于传感器最优布置的光伏阵列阴影诊断与定位[J]. 太阳能学报, 2018, 39(02):513-519.
- [27] Chao K H, Ho S H, Wang M H. Modeling and fault diagnosis of a photovoltaic system[J]. Electric Power Systems Research, 2008, 78(01):97-105.
- [28] Takashima T, Yamaguchi J, Otani K, et al. Experimental studies of fault location in PV module strings[J]. Solar Energy Materials & Solar Cells, 2009, 93(06):1079-1082.
- [29] Alam M K, Khan F H, Johnson J, et al. PV faults: Overview, modeling, prevention and detection techniques[C]. Control and Modeling for Power Electronics, IEEE, 2013, 1-7.
- [30] 李元良, 丁坤, 陈富东等. 基于快速过采样主成分分析法的光伏阵列故障诊断[J]. 电网技术, 2019, 43(01):308-315.
- [31] 杜帅. 单相光伏并网逆变系统控制策略研究[D]. [硕士学位论文]. 郑州: 华北水利水电大学, 2019.
- [32] 何智成. 带储能装置的独立光伏发电系统研究[D]. [硕士学位论文]. 株洲: 湖南工业大学, 2018.
- [33] 张晓阳. 光伏阵列故障在线诊断的 SOM 神经网络算法研究[D]. 山东理工大学, 2021.
- [34] Azzouzi M, Popescu D, Bouchahdane M. Modeling of Electrical Characteristics of Photovoltaic Cell Considering Single-Diode Model[J]. Journal of Clean Energy Technologies, 2016, 4(6):414-420.
- [35] Ahmed M T, Gonçalves T, Tlemcani M. Single diode model parameters analysis of photovoltaic cell[C]. 2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA), Birmingham, 2016, 396-400.
- [36] Garoudja E, Chouder A, Kara K., et al. An enhanced machine learning based approach for failures detection and diagnosis of PV systems[J]. Energy Conversion & Management, 2017, 151(nov.):496-513.
- [37] 孟少飞. 基于 TCT 结构光伏阵列的静态重构方法研究[D]. 宁夏大学, 2021.

- [38] Tang J N, Zhu Y Q, Wang W S. Fault diagnosis method and simulation analysis for photovoltaic array[C]. International Conference on Electrical and Control Engineering. New York: Institute of Electrical and Electronics Engineers, 2011, 1569-1573.
- [39] 夏超浩. 光伏组件模型参数辨识及其故障识别研究[D]. 西安理工大学, 2021.
- [40] 毛健, 赵红东, 姚婧婧. 神经网络的发展及应用[J]. 电子设计工程, 2011, 19(24): 62-65.
- [41] 丛爽. 面向 MATLAB 工具箱的神经网络理论与应用[M]. 中国科学技术大学出版社, 1998.
- [42] Widrow B, Rumelhart D E, Lehr M A. Neural networks: applications in industry, business and science[J]. Communications of the Acm, 1994, 37(3): 93-105.
- [43] 谢雨诗, 何丽娜, 徐文燕, 明新国. 基于 AFOA-BP 神经网络的燃气轮机研制风险预测[J]. 热能动力工程, 2021, 36(10): 228-234.
- [44] Min-Rong Chen, et al. An adaptive fractional-order BP neural network based on extremal optimization for handwritten digits recognition[J]. Neurocomputing, 2020, 391(pp.): 260-272.
- [45] Xinhong Ni, Zeng Yang. Inverter fault diagnosis based on BP neural network[C]//Proceedings of 2021 6th International Conference on Automation, Control and Robotics Engineering(CACRE). 2021: 203-208.
- [46] 黄丽. BP 神经网络算法改进及应用研究[D]. 重庆师范大学, 2008.
- [47] 吕琼帅. BP 神经网络的优化与研究[D]. 河南: 郑州大学, 2011.
- [48] 李敏强, 寇纪淞. 遗传算法的基本理论与应用[M]. 北京: 科学出版社, 2002: 32-37.
- [49] 崔艺梦. 光伏阵列故障诊断方法研究[D]. 华北水利水电大学, 2020.
- [50] 墨蒙, 赵龙章, 龚媛雯等. 基于遗传算法优化的 BP 神经网络研究应用[J]. 现代电子技术, 2018, 41(09): 41-44.
- [51] 孙文娟. 自适应遗传算法的改进及其在爆炸冲击响应谱时域合成优化中的应用研究[D]. 中国科学技术大学, 2019.
- [52] 孙弋清. BP 神经网络优化算法研究及在故障诊断中的应用[D]. 秦皇岛: 燕山大学, 2018.
- [53] McClelland, J.L., Rumelhart, D.E. An interactive activation model of the effect of context in perception: I. An account of basic findings[J]. Psychological Review, 2012, 88(88): 375-407.
- [54] DUAN H B, QIAO P X. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning[J]. International Journal of Intelligent Computing and Cybernetics, 2014, 7(1): 24-37.
- [55] 段海滨, 叶飞. 鸽群优化算法研究进展[J]. 北京工业大学报, 2017, 43(01): 1-7.
- [56] GUILFORD T, ROBERT S, BIRO D, et al. Positional entropy during pigeon homing II: a navigational interpretation of Bayesian latent state models [J]. Journal of Theoretical Biology, 2004, 227(1): 25-38.
- [57] 朱凯. 基于混合粒子群优化算法的电力系统多目标无功优化研究[D]. 广西: 广西大学, 2015.

## 附录

本文所涉及的 MATLAB 核心代码如下：

### (1) BP 神经网络

```
%% 数据归一化
[inputn,inputps]=mapminmax(input_train);
inputn_test=mapminmax('apply',input_test,inputps);
outputn_train=ind2vec(output_train);
outputn_test=ind2vec(output_test);
%节点个数
inputnum=size(input_train,1);
hidddennum=9;
outputnum=size(outputn_train,1);
%% 标准的 BP 神经网络
net0=newff(inputn,outputn_train,hidddennum,{'tansig','purelin'},'trainlm');
net0.trainParam.epochs=1000;           % 训练次数
net0.trainParam.lr=0.1;                % 学习速率
net0.trainParam.goal=0.0001;           % 训练目标最小误差
net0.trainParam.show=25;               % 显示频率
net0.trainParam.mc=0.01;               % 动量因子
net0.trainParam.min_grad=1e-6;         % 最小性能梯度
net0.trainParam.max_fail=12;           % 最高失败次数
%训练
[net0,tr0]=train(net0,inputn,outputn_train);%开始训练
figure,plotperform(tr0);                %误差 MSE 下降线
an0=sim(net0,inputn_test);             %测试集的预测结果
Output_test0=vec2ind(an0);
```

### (2) 遗传算法

```
%% 种群初始化-----
```

```

individuals=struct('fitness',zeros(1,sizepop), 'chrom',[]);
avgfitness=0;
bestfitness=0;
bestchrom=0;
trace=[0 0];
for i=1:sizepop
    individuals.chrom(i,:)=Code(lenchrom,bound);
    [individuals.fitness(i),individuals.net{i,1},individuals.net{i,2}]=objfun(individuals.chrom(i,:),
inputnum,hiddennum,outputnum,net,inputn,outputn);
    for j=1:nvar
        v(i,j)=rand;
    end
end
%找最好的染色体
[bestfitness ,bestindex]=max(individuals.fitness);
bestchrom=individuals.chrom(bestindex,:);
bestnet=individuals.net{bestindex,1};
bestnet_tr=individuals.net{bestindex,2};
avgfitness=sum(individuals.fitness)/sizepop;
%% 迭代求解最佳初始阈值和权值
% 迭代开始
for i=1:maxgen
%选择
individuals=Select(individuals,sizepop);
%交叉
individuals.chrom=Cross_new(pc1,pc2,lenchrom,individuals.chrom,individuals.fitness,sizep
op,bound);
individuals.chrom=Cross(pcross,lenchrom,individuals.chrom,sizepop,bound);
% 变异
individuals.chrom=Mutation(pm1,pm2,lenchrom,individuals.chrom,sizepop,i,maxgen,bound
);
individuals.chrom=Mutation(bestchrom,individuals.chrom,sizepop,bound,v);
    end
% 计算适应度
for j=1:sizepop

```

```

[individuals.fitness(j),individuals.net{j,1},individuals.net{j,2}]=objfun(individuals.chrom(j,:),
inputnum,hiddennum,outputnum,net,inputn,outputn);
    end
%寻找最优和最劣染色体个体
    [newbestfitness,newbestindex]=max(individuals.fitness);
    [worestfitness,worestindex]=min(individuals.fitness);
% 更新最优解及其位置
    if bestfitness<newbestfitness
        bestfitness=newbestfitness;
        bestchrom=individuals.chrom(newbestindex,:);
        bestnet=individuals.net{newbestindex,1};
        bestnet_tr=individuals.net{newbestindex,2};
    end
% 利用最优的解置换最差的解
    individuals.chrom(worestindex,:)=bestchrom;
    individuals.fitness(worestindex)=bestfitness;
    avgfitness=sum(individuals.fitness)/sizepop;
    trace=[trace;avgfitness bestfitness];

```

### (3) 鸽群算法

```

%% 鸽群算法参数初始化
dim=length(lenchrom);
Nc1=round(Max_iter*0.7);
Nc2=Max_iter-Nc1;
if(max(size(ub))==1)
    ub=ub.*ones(1,dim);
    lb=lb.*ones(1,dim);
end
%% 鸽群初始化
pop=SearchAgents_no;
X0=initialization(pop,dim,ub,lb);
X=X0;
%计算鸽群初始适应度值
fitness=zeros(1,pop);

```

```

for i = 1:pop
    fitness(i) = fun(X(i,:),inputnum,hiddennum,outputnum,net,inputn,outputn);
end
[fitness, index]= sort(fitness);%排序
GBestF = fitness(1);%全局最优适应度值
for i = 1:pop
    X(i,:) = X0(index(i),:);
end
curve=zeros(1,Max_iter);
GBestX = X(1,:);%全局最优位置
X_new = X;
%初始速度
Vec = rand(pop,dim);
VecNew = Vec;
%鸽群第一阶段
for t = 1: Nc1
    Vec = VecNew;
    for i = 1:pop
        R = rand;
        %速度更新
        TempV=w*Vec(i,:) + rand.*(GBestX - X(i,:));
        %位置更新
        TempPosition = X(i,:).*(1-exp(-R*t)) + TempV;
        %边界检查
        for j = 1:dim
            if(TempPosition(j)<lb(j) || TempPosition(j)>ub(j))
                TempPosition(j) = lb(j) + rand.*(ub(j) - lb(j));
                empV(j) = rand;
            end
        end
        X_new(i,:) = TempPosition;
        VecNew(i,:)=TempV;
    end
    for j=1:pop
        fitness_new(j) = fun(X_new(j,:),inputnum,hiddennum,outputnum,net,inputn,outputn);
    end
end

```



```

end
X = X_new;
fitness = fitness_new;
%排序更新
[fitness, index]= sort(fitness);%排序
for j = 1:pop
    X(j,:) = X(index(j),:);
end
if(fitness(1) < GBestF)
    GBestF = fitness(1);
    GBestX = X(1,:);
end
curve(t) = GBestF;
end
%鸽群第二阶段
for t = 1:Nc2
    S = 0;
    for j = 1:round(pop/2)
        S = S + X(j,:).*fitness(j);
    end
    Xcenter = S./((pop/2)*sum(fitness(1:round(pop/2))));
%根据中心更新种群
    for i = 1:round(pop/2)
        for j = 1:dim
            Temp = X(i,j) + rand.*(Xcenter(j) - X(i,j));
            while Temp < lb(j) || Temp > ub(j)
                Temp = X(i,j) + rand.*(Xcenter(j) - X(i,j));
            end
        end
        X(i,:) = Temp;
    end
    for j=1:pop
        fitness(j) = fun(X(j,:),inputnum,hiddennum,outputnum,net,inputn,outputn);
    end
end
%排序更新

```

```
[fitness, index]= sort(fitness);%排序
for j = 1:pop
    X(j,:) = X(index(j),:);
end
if(fitness(1) < GBestF)
    GBestF = fitness(1);
    GBestX = X(1,:);
end
curve(Nc1 + t) = GBestF;
end
Best_pos =GBestX;
Best_score = curve(end);
%% 选用 PIO 算法优化结果,对遗传算法进行优化
save('data1.mat','X');
```

## 在读期间公开发表的论文和参与项目

### 发表论文:

- [1] 刘东, 李田泽, 刘开石, 张晓阳, 韩鸿雁. GA-BP 神经网络在光伏阵列故障检测中的应用研究[J]. 电源技术, 2021, 45 (03) :370-373.
- [2] 刘开石, 李田泽, 刘东, 王梦婕, 徐立宾. 基于 ABC-SVM 算法的光伏阵列故障诊断[J]. 电源技术, 2021, 45 (09) :1171-1174.

### 参与项目:

- [1] 参与淄博市校城合作项目研究(编号 2019ZBXC119): 太阳能光伏阵列故障在线检测系统及关键技术研究
- [2] 参与山东省高等学校科技计划项目研究: 太阳能光伏发电及关键技术研究
- [3] 参与山东省研究生建设项目研究(编号 SDYKC18067): 现代检测理论

### 发明专利:

太阳能光伏发电组件故障在线检测装置及其检测方法, 申请号: 202111383940.X

## 致谢

但见时光流似箭，眨眼之间，研究生三年的学习生活即将结束，在落笔致谢的瞬间，仿佛又回到了刚来理工大的那一天，回头想去，能够有幸来到理工大，能够有幸认识这么多优秀的老师和同学，能够有幸拥有这段经历，真的是我一生值得珍视的财富。借此机会，想对这一路上支持、帮助我的人表示衷心的感谢。

首先特别感谢我的研究生导师李田泽教授，李老师为人宽厚，对待学问非常严谨，我的论文能顺利完成离不开他的敦促和指导。在学术上，大到论文的选题到实验数据的处理分析，小到论文格式的修改，他都会不厌其烦的帮我检查答疑，这中间每个过程都离不开李老师对我的指导。在生活上，李老师积极乐观的生活态度和全始全中的做事态度给我树立了良好的榜样。在我的前半生中，能够遇到李老师，真的是我莫大的幸运，再次感谢老师的教诲之恩。

非常感谢我的舍友刘宽，栗义康，马玉路，在这三年的时间我们一起上课，一起学习，相互鼓励，相互支持，我会永远记得我们的点点滴滴，特别是我们边讨论边准备期末考试的日子，感谢你们对我三年来的照顾。

感谢张晓阳师哥，韩鸿雁师姐，同门刘开石，在你们的陪伴下一起做实验，一起讨论问题，在创作论文的过程中我们始终保持交流沟通，在遇到瓶颈的时候，在于他们的交流中我也找到了自己的方向。我们的友谊因此变得更加的醇厚，感谢他们的包容与帮助。让我感觉到学习的日子不再那么枯燥，感谢 302 实验室的师哥师姐们，幸运遇到你们。

无比感谢我的家人，感谢父母一直以来对我的支持，正因为有你们作为我坚强的后盾，我才能安心的学习并取得今天的成绩。请你们放心，无论遇到多大的困难，我都会坚持到底，我会努力让我们的家庭变得越来越好。

最后感谢在本次论文评阅和答辩的各位专家老师们，向你们表示深深的谢意！