

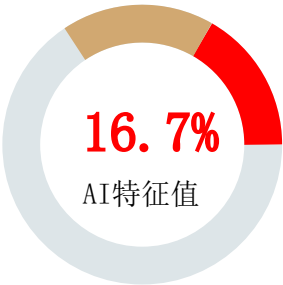
AIGC检测 · 全文报告单

NO:CNKIAIGC2026SJ_20260524207007

检测时间:2026-05-14 22:15:18

篇名：基于I-V特性曲线的分布式光伏故障诊断系统设计
作者：席润泽
单位：
文件名：

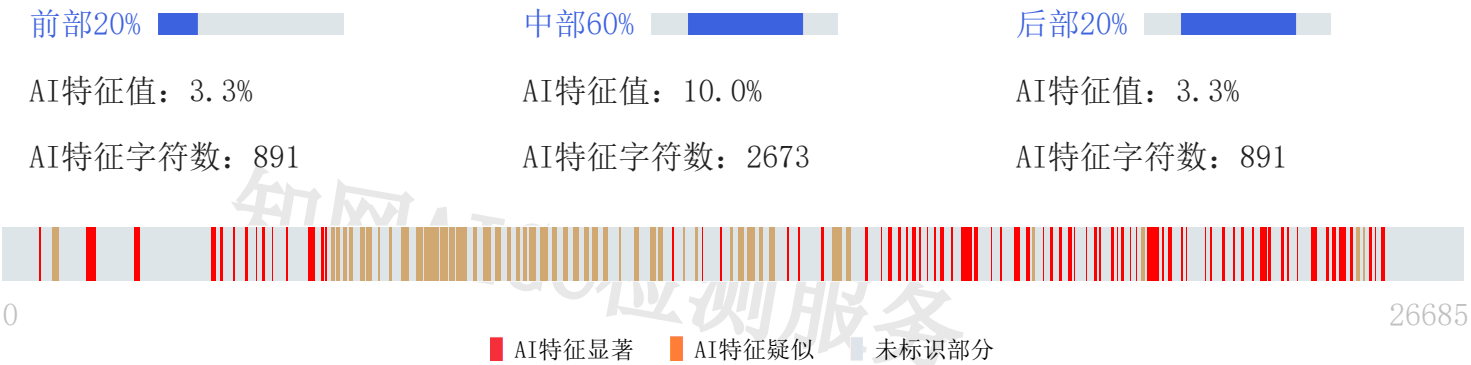
全文检测结果



AI特征值：16.7%
AI特征字符数：4456
总字符数：26685

- AI特征显著（计入AI特征字符数）
- AI特征疑似（未计入AI特征字符数）
- 未标识部分

AIGC片段分布图



分段检测结果

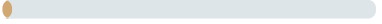
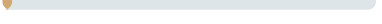
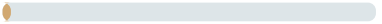



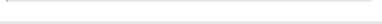
序号	AI特征值	AI特征字符数/章节(部分)字符数	章节(部分)名称
1	16.7%	4456 / 26685	论文全文



































1. 论文全文
AI特征值：16.7% AI特征字符数 / 章节(部分)字符数：4456 / 26685































片段指标列表

序号	片段名称	字符数	
1	片段1	38	0.1%

2	片段2	125		0. 5%
3	片段3	179		0. 7%
4	片段4	119		0. 4%
5	片段5	92		0. 3%
6	片段6	64		0. 2%
7	片段7	33		0. 1%
8	片段8	51		0. 2%
9	片段9	30		0. 1%
10	片段10	56		0. 2%
11	片段11	35		0. 1%
12	片段12	40		0. 1%
13	片段13	146		0. 5%
14	片段14	65		0. 2%
15	片段15	30		0. 1%
16	片段16	84		0. 3%
17	片段17	74		0. 3%
18	片段18	88		0. 3%
19	片段19	81		0. 3%
20	片段20	101		0. 4%
21	片段21	123		0. 5%
22	片段22	43		0. 2%
23	片段23	65		0. 2%
24	片段24	72		0. 3%
25	片段25	87		0. 3%
26	片段26	122		0. 5%
27	片段27	28		0. 1%
28	片段28	97		0. 4%
29	片段29	39		0. 1%
30	片段30	92		0. 3%
31	片段31	23		0. 1%
32	片段32	38		0. 1%
33	片段33	38		0. 1%
34	片段34	67		0. 3%
35	片段35	20		0. 1%

36	片段36	118		0.4%
37	片段37	113		0.4%
38	片段38	96		0.4%
39	片段39	87		0.3%
40	片段40	134		0.5%
41	片段41	104		0.4%
42	片段42	81		0.3%
43	片段43	74		0.3%
44	片段44	76		0.3%
45	片段45	113		0.4%
46	片段46	148		0.6%
47	片段47	99		0.4%
48	片段48	82		0.3%
49	片段49	106		0.4%
50	片段50	94		0.4%
51	片段51	126		0.5%
52	片段52	105		0.4%
53	片段53	34		0.1%
54	片段54	94		0.4%
55	片段55	125		0.5%
56	片段56	87		0.3%
57	片段57	25		0.1%
58	片段58	38		0.1%
59	片段59	56		0.2%
60	片段60	23		0.1%
61	片段61	28		0.1%
62	片段62	65		0.2%
63	片段63	110		0.4%
64	片段64	135		0.5%
65	片段65	69		0.3%
66	片段66	116		0.4%
67	片段67	41		0.2%
68	片段68	33		0.1%
69	片段69	47		0.2%

70	片段70	173		0.6%
71	片段71	109		0.4%
72	片段72	73		0.3%
73	片段73	36		0.1%
74	片段74	57		0.2%
75	片段75	54		0.2%
76	片段76	34		0.1%
77	片段77	74		0.3%
78	片段78	30		0.1%
79	片段79	35		0.1%
80	片段80	16		0.1%
81	片段81	39		0.1%
82	片段82	51		0.2%
83	片段83	31		0.1%
84	片段84	43		0.2%
85	片段85	201		0.8%
86	片段86	83		0.3%
87	片段87	30		0.1%
88	片段88	39		0.1%
89	片段89	65		0.2%
90	片段90	54		0.2%
91	片段91	63		0.2%
92	片段92	62		0.2%
93	片段93	29		0.1%
94	片段94	45		0.2%
95	片段95	47		0.2%
96	片段96	74		0.3%
97	片段97	26		0.1%
98	片段98	30		0.1%
99	片段99	65		0.2%
100	片段100	37		0.1%
101	片段101	60		0.2%
102	片段102	27		0.1%
103	片段103	56		0.2%

104	片段104	30		<div></div>	0.1%
105	片段105	23		<div></div>	0.1%
106	片段106	66		<div></div>	0.2%
107	片段107	232		<div></div>	0.9%
108	片段108	54		<div></div>	0.2%
109	片段109	71		<div></div>	0.3%
110	片段110	41		<div></div>	0.2%
111	片段111	26		<div></div>	0.1%
112	片段112	22		<div></div>	0.1%
113	片段113	47		<div></div>	0.2%
114	片段114	56		<div></div>	0.2%
115	片段115	46		<div></div>	0.2%
116	片段116	57		<div></div>	0.2%
117	片段117	41		<div></div>	0.2%
118	片段118	131		<div></div>	0.5%
119	片段119	70		<div></div>	0.3%
120	片段120	63		<div></div>	0.2%
121	片段121	46		<div></div>	0.2%
122	片段122	23		<div></div>	0.1%
123	片段123	125		<div></div>	0.5%
124	片段124	67		<div></div>	0.3%
125	片段125	65		<div></div>	0.2%
126	片段126	85		<div></div>	0.3%
127	片段127	44		<div></div>	0.2%
128	片段128	68		<div></div>	0.3%
129	片段129	75		<div></div>	0.3%
130	片段130	52		<div></div>	0.2%
131	片段131	61		<div></div>	0.2%
132	片段132	24		<div></div>	0.1%
133	片段133	68		<div></div>	0.3%

原文内容

毕 业 设 计（论 文）
基于I-V特性曲线的分布式光伏故障诊断
系统设计

学院名称： 电气工程与自动化学院

专业： 测控技术与仪器

班级： 测控技术221

学生姓名： 席润泽

学号： 2204332139

指导教师姓名： 李婕

指导教师职称： 讲师

2026年6月10日

毕业设计（论文）原创性声明和使用授权说明

原创性声明

本人郑重承诺：所呈交的毕业设计（论文），是我个人在指导教师的指导下进行的研究工作及取得的成果。尽我所知，除文中特别加以标注和致谢的地方外，不包含其他人或组织已经发表或公布过的研究成果，也不包含我为获得河南工学院及其它教育机构的学位或学历而使用过的材料。对本研究提供过帮助和做出过贡献的个人或集体，均已在文中作了明确的说明并表示了谢意。

作者签名： 日期：

指导教师签名： 日期：

使用授权说明

本人完全了解河南工学院关于收集、保存、使用毕业设计（论文）的规定，即：按照学校要求提交毕业设计（论文）的印刷本和电子版本；学校有权保留毕业设计（论文）的印刷本和电子版，并提供目录检索与阅览服务；学校可以采用影印、缩印、数字化或其它复制手段保存论文；在不以营利为目的前提下，学校可以公布论文的部分或全部内容。

作者签名： 日期：

基于I-V特性曲线的分布式光伏故障诊断

摘要：分布式光伏系统通常安装在房顶、山地等区域，布设点位分散，运行环境复杂。组件在长期服役中受到光照波动、局部阴影、积尘、线路老化与连接不良等多元因素作用，容易出现功率衰减及运行状态偏移。由于系统涉及组件数量多、空间分布广，单纯依靠人工巡检难以在故障发生初期及时识别异常；同时，仅凭电压、电流或功率这类单一电气量的变化进行判别，也容易受环境条件波动的干扰。光伏阵列的I-V特性曲线能够刻画不同工作点下的电压、电流与功率响应，其整体形态及特征参数中蕴含着较充分的故障信息。

针对分布式光伏阵列运行状态识别需求，构建了一套基于I-V特性曲线的故障诊断流程。建模部分以光伏电池单二极管等效电路为基础建立组件数学模型，并结合串并联拓扑关系构建2×2光伏阵列模型，用于分析辐照度与温度变化对I-V特性曲线和P-V特性曲线的影响规律。故障特征分析部分设定正常、局部遮挡、老化故障、开路故障和短路故障五类典型工况，对比不同状态下的曲线形态差异，并进一步考察开路电压、短路电流、最大功率点电压、最大功率点电流及填充因子等关键参数的变化趋势。在此基础上，从I-V特性曲线及相关电气参数中提取101维诊断特征，依次完成缺失值处理、稳健标准化、特征筛选和PCA降维，再将处理后的特征输入支持向量机，实现五类运行状态的分类训练与测试。为验证诊断流程，搭建了由光伏阵列、DC-DC电压电流采集板、STM32F407ZG主控、4G通信模块、云服务器和AIRIOT云平台组成的实验系统，完成I-V数据采集、远程上传、服务器端诊断及平台结果显示。在本文设定的2×2阵列和五类工况条件下，实验结果表明，该方法能够较稳定地区分不同运行状态，初步验证了基于I-V特性曲线开展分布式光伏故障诊断的可行性。

关键词：分布式光伏；I-V特性曲线；故障诊断；多模型融合；AIRIOT

Abstract:Distributed photovoltaic systems are usually deployed on rooftops, factory buildings, and other load-side locations, where operating conditions are complex and scattered. During long-term operation, photovoltaic modules are affected by irradiance fluctuation, local shading, dust accumulation, line aging, and poor electrical connections, which may cause power attenuation and abnormal operating states. Owing to the large number and wide distribution of modules, early faults are difficult to detect through manual inspection alone. In addition, diagnosis based only on voltage, current, or power is easily affected by environmental variations. Compared with single electrical indicators, the I-V characteristic curve reflects the variation of voltage, current, and power at different operating points and contains richer information for fault diagnosis.

This study proposes a fault diagnosis scheme for distributed photovoltaic arrays based on the I-V characteristic curve. A photovoltaic module model is established from the single-diode equivalent circuit, and a 2×2 photovoltaic array model is further constructed to analyze the effects of irradiance and temperature on the I-V and P-V characteristic curves. Five operating conditions, including normal state, local shading, aging fault, open-circuit fault, and short-circuit fault, are compared through curve shape and key parameters. Then, 101-dimensional features are extracted and processed by missing value handling, robust scaling, feature selection, and PCA dimensionality reduction, and a support vector machine is used for classification. An experimental system integrating a photovoltaic array, a DC-DC acquisition board, an STM32F407ZG controller, a 4G communication module, a cloud server, and the AIRIOT cloud platform is built for data acquisition, remote transmission, diagnosis, and result display. Results under the 2×2 array and five preset conditions show that the method can distinguish different operating states and preliminarily verify the feasibility of I-V characteristic curve-based fault diagnosis for distributed photovoltaic systems.

Key words:Distributed photovoltaic power; generation I-V characteristic curve;fault diagnosis;multi-model fusion;AIRIOT

河南工学院本科毕业设计（论文）

第1章 绪论

1.1 课题研究背景

分布式光伏阵列可直接并入负载侧，已在工业区、公共区域、居民房顶及农业设施等场景中得到应用，在减轻局部用电压力、降低输配电过程中的能量损耗以及提升可再生能源消纳能力方面具有较好的适配性。随着装机容量和接入点数量不断增加，分布式光伏系统的运行状态不再只是单个组件或单个逆变器的问题，而是直接关系到发电效率、供电稳定性和后期运维成本[1]。

光伏组件长期处于户外工况下，常受到温度波动、光照变化、灰尘附着、鸟粪遮挡、树木或建筑物遮蔽、雨雪侵袭以及线路连接老化等因素影响。这些因素可能引起组件输出能力下降，也可能进一步发展为局部遮挡、热斑、老化、开路、短路等典型故障。对于分布范围广、安装位置分散的光伏系统而言，故障往往具有隐蔽性和渐进性，早期异常不一定会立即造成系统停机，却会持续降低发电量，严重时还可能带来组件损坏和安全风险。及时发现异常并判断故障类型，已经成为分布式光伏系统运维中需要重点解决的问题[2][3]。

经典运维方式主要依靠人工巡检、定期维护以及少量电气参数监测来开展。人工巡检能够发现部分外观缺陷和明显设备异常，但受巡检周期、现场条件和人员经验影响较大，难以对大量分散组件进行连续、精细化监测。在环境扰动较强或多种异常因素同时存在时，参数变化与故障类型之间并不总是一一对应，容易出现误判或漏判。光伏系统规模扩大后，依靠人工经验进行事后排査的方式已经难以满足高效率、低成本和实时化运维需求，面向运行数据的故障诊断方法逐渐成为研究和应用的重要方向[4]。

光伏故障诊断的关键，在于从单块光伏或矩阵的输出特征中识别异常工况，并为后续检修与运行管理提供判断依据。通过对运行数据进行采集、分析和判断，可以在故障扩大之前发现发电性能下降的原因，减少无效发电损失，提升系统运行可靠性。在分布式光伏场景下，若能够结合数据采集、通信传输和在线分析技术建立诊断流程，就可以降低对人工巡检的依赖，使运维人员更快定位异常区域和故障类型[5]。

1.2 国内外研究现状

针对光伏系统故障诊断问题，国内外学者已开展了较为广泛的研究。相关研究可分为四类：基于电气参数的诊断、基于I-V或P-V曲线的诊断、基于机器学习的识别方法，以及面向在线监测的系统研究。不同研究路径的适用条件和局限性存在差异[6][7][8]。

早期光伏故障研究通常以电压、电流、功率和填充因子等电气参数为基础，借助阈值判定、参数辨识或模型对比完成故障识别。部分学者围绕参数辨识建立组件故障诊断模型，通过等效参数变化反映组件退化或失配情况。但在复杂工况下，仅依靠少量特征量往往难以准确区分不同异常状态，参数估计过程也容易受到测量误差和工况变化影响，模型适用范围与实际运行条件密切相关。

对于I-V特性曲线或P-V特性曲线开展故障诊断的方法，正逐步受到关注。与单一参数相比，完整曲线包含更多输出状态信息，更有利于观察故障引起的形态变化。国内有研究围绕晶硅光伏组件的I-V特性开展故障分析，讨论了遮挡、老化、开路、短路等工况下曲线的变化特点。部分研究者从曲线距离、斜率变化、关键拐点位置等角度提取特征，用于区分不同运行状态。这类方法在状态表征上较单点参数更充分，但对曲线采集精度、采样方式和特征设计也提出了更高要求[9][10][11]。当环境扰动较强或采样质量不足时，曲线特征的稳定性会受到影响，诊断效果也会随之波动。

在数据驱动方法不断引入的背景下，机器学习已逐步成为光伏故障诊断研究中的重要分支。相关研究通常先从I-V曲线、P-V曲线或运行参数中提取多维特征，再采用支持向量机、随机森林、神经网络等模型进行分类[12][13]。与经验阈值法相比，这类方法在处理非线性特征关系和多类状态识别方面更具灵活性。国内研究团队围绕特征提取与机器学习方法开展光伏组件复合故障诊断研究，结果表明，在合理构造特征的前提下，常见分类模型能够较好地区分若干典型故障状态[15][16][17]。国外研究中，深度学习方法也被用于光伏故障检测与诊断，通过卷积网络、循环网络或组合模型直接学习数据特征，以减轻人工特征设计的负担[18][19]。不过，这类方法通常依赖较大的样本规模和较稳定的数据分布，而真实电站中的故障样本往往较少，训练数据又容易受运行工况影响，因此模型泛化能力和跨场景适应性仍是需要进一步解决的问题。

但从现有研究看，仍有若干问题值得进一步关注：一是部分方法对单一特征或单一模型依赖较强，在复杂工况下的适应性仍受限制；二是一些研究主要聚焦算法设计，对采集、传输、存储和结果展示等工程环节涉及较少；三是仿真数据与实际运行数据之间客观存在差异，实验条件下得到的模型结果未必能够直接迁移到工程应用中。基于此，本文将I-V特性曲线作为主要分析对象，在光伏阵列建模和故障仿真的基础上，结合分布式系统设计进行实验平台验证，以期为分布式光伏故障诊断提供一种较为完整的实现思路。

1.3 课题主要研究内容

本文以分布式光伏系统故障诊断为研究对象，基于光伏阵列输出特性分析，结合I-V曲线数据完成故障特征提取、诊断模型训练及实验平台验证。全文共分为六章，各章主要内容如下。

第1章对课题研究背景、国内外研究现状及论文主要内容进行概述。分析光伏组件在长期运行中可能出现的遮挡、老化、开路、短路等问题。梳理现有光伏故障诊断方法及其不足，明确本文的研究方向和章节安排。

第2章主要讨论光伏电池建模及阵列基础分析。该章先阐述光伏电池的工作原理，再采用单二极管等效模型建立数学模型，并结合数值迭代方法求得组件的I-V特性曲线和P-V特性曲线。分析辐照度和温度对光伏输出特性的影响，建立光伏阵列模型。

第3章对光伏阵列故障输出特性进行分析。该章以光伏阵列为研究对象，设置正常、局部遮挡、老化、开路和短路五种工况，用于模拟不同故障对阵列输出特性的影响。对各工况下的I-V曲线、P-V曲线及关键电气参数进行对比，明确不同故障的表现差异。

第4章构建光伏故障诊断模型并完成训练。该章整理五类故障样本数据，并在数据预处理的基础上，结合特征筛选与PCA降维方法压缩冗余信息，随后将处理后的特征输入SVM分类模型进行训练。最后通过测试集评估模型对五类光伏运行状态的识别效果。

第5章完成光伏故障诊断系统的实现与实验验证。该章搭建了分布式光伏故障诊断实验平台，平台由光伏阵列、采集板、主控板、4G通信模块、云服务器和AIRIOT云平台组成。采集板负责采集电压、电流数据并生成I-V曲线，主控板通过4G模块将数据上传至服务器，由服务器完成特征提取和模型诊断，最终在云平台上显示运行状态及故障识别结果。通过五类实验工况对系统进行验证，初步验证了该系统的可行性。

第6章对全文研究工作进行总结，并结合现有研究中的不足提出后续展望。该章归纳主要研究成果，同时围绕阵列规模扩展、真实运行数据引入、复合故障样本补充以及特征提取与诊断方法优化等方面，对后续研究方向进行说明。

1.4 本章小结

本章对课题研究背景、国内外研究现状及论文主要内容进行了概述。结合分布式光伏系统运行中的故障诊断需求，阐明了采用I-V特性曲线进行状态识别的必要性；在梳理现有研究思路的基础上，总结了参数法、曲线法、机器学习方法以及在线监测系统研究的主要特点与不足。最后，对本文的研究内容和章节安排作了说明。上述内容为后续开展光伏阵列建模、故障特性分析、诊断模型研究以及系统设计与测试提供了依据。

第2章 光伏电池建模及阵列基础分析

2.1 光伏电池工作原理与数学模型

2.1.1 光伏电池工作原理

光伏电池是光伏发电系统实现光电转换的基本单元，其工作机理基于半导体材料的光生伏特效应。晶体硅光伏电池通常由P型半导体和N型半导体构成PN结。当两种半导体材料相互接触后，受载流子浓度差影响，电子与空穴会发生扩散运动，结区附近由此形成空间电荷区，并建立由N区指向P区的内建电场。在无光照条件下，虽然PN结内部已经存在内建电场和电势差，但器件整体仍处于热平衡状态，因此无法持续向外输出电能。

在光照条件下，当入射光子的能量高于半导体材料的禁带宽度时，价带中的电子吸收光子能量并跃迁至导带，从而在半导体内部形成电子—空穴对。在PN结内建电场的作用下，光生载流子发生分离，其中电子向N区移动，空穴向P区移动，进而使光伏电池两端形成光生电动势。接入外部负载后，电子经外电路流向负载并返回电池内部，与空穴复合，最终形成持续的电流输出。图2.1所示为光伏电池光生伏特效应原理图。

图2.1 光伏电池光生伏特效应原理图

2.1.2 单二极管等效模型

在光照条件下，光伏电池可将太阳能转化为电能，其输出电流与输出电压之间表现出明显的非线性关系。。为准确表征这种输出特性，有必要建立相应的数学模型。

目前，光伏电池等效电路中应用较多的模型主要有单二极管模型和双二极管模型。双二极管模型通过增加一个二极管来描述载流子复合过程，仿真精度相对较高，但参数求取和计算过程也更为复杂。单二极管模型的精度虽略低于双二极管模型，但其计算量较小，参数物理意义明确，且多数参数可依据组件数据手册确定，因此在工程建模中更便于使用。综合本文故障诊断系统对建模精度与计算效率的要求，最终采用单二极管等效电路模型建立光伏电池数学模型。

该模型由光生电流源、二极管、串联电阻和并联电阻四部分组成。其中，光生电流源用于表征光照条件下产生的光生电流；二极管用于描述PN结的非线性伏安特性；串联电阻用于反映电池材料、电极以及连接导线引起的欧姆损耗；并联电阻则用于表征电池内部漏电流的影响。单二极管模型的等效电路如图2.2所示。

图2.2 单二极管模型图

根据单二极管等效电路，光伏电池的输出电流满足如下关系式：

式中， I 表示光伏电池输出电流， V 表示输出电压， I_{ph} 表示光生电流， I_0 表示二极管反向饱和电流， R_S 表示串联电阻， R_{sh} 表示并联电阻， N_S 表示串联电池片数量， V_t 表示考虑二极管理想因子后的热电压。

式中， A 为二极管理想因子， K 为玻尔兹曼常数， T 为电池绝对温度， q 为电子电荷量。

为获取光伏电池在不同工作点下的输出特性，本文在单二极管模型基础上对输出电压进行离散扫描。由于输出功率由电压和电流共同决定，在求得某一电压点对应的输出电流后，可按下式计算该点的输出功率：

先根据当前环境条件估算开路电压，其计算公式为：

式中， β_{voc} 为开路电压温度系数。随后在0至 V_{oc} 的范围内生成离散电压序列：

对于每一个电压点 V_i ，采用牛顿迭代法求解对应的输出电流 I_i ，并计算该点功率：

由全部离散点的电压、电流和功率数据，可得到光伏电池的 I - V 特性曲线和 P - V 特性曲线。最大功率点通过比较功率数组中的最大值确定：

其对应的最大功率点电压和电流分别表示为：

在不同辐照度 G 和温度 T_c 条件下，光伏电池的输出特性会随之变化。一般而言，在温度保持不变时，辐照度增大将引起光生电流增加，进而使输出电流和最大输出功率上升；在辐照度保持不变时，温度升高则会导致开路电压下降，从而使最大输出功率减小。

2.3 光伏电池输出特性分析

本文借助Python构建了光伏仿真平台，用于检验单二极管模型的输出特性，并生成后续光伏特性分析所需的仿真数据。该平台实现了光伏电池参数配置、环境参数输入、非线性方程数值求解、输出功率计算、最大功率点搜索以及特性曲线绘制等功能。

在平台实现层面，Python中的SingleDiodePVModel类负责单二极管模型的主要数值计算任务，包括光生电流计算、反向饱和电流计算、输出电流迭代求解、输出功率计算以及最大功率点计算。程序采用NumPy完成数值运算，借助Matplotlib绘制 I - V 特性曲线和 P - V 特性曲线，并通过PyQt6搭建可视化界面，实现参数输入、曲线预览和图像导出功能。

2.3.1 I - V 与 P - V 特性分析

在标准测试条件下，光伏电池的输出特性通常可由 I - V 特性曲线和 P - V 特性曲线表征。 I - V 曲线描述端电压变化时输出电流的响应规律， P - V 曲线则用于反映对应工作点的功率分布情况。。图2.4所示为单组件在正常工况下的典型输出曲线。由图可见， I - V 曲线在低电压区域内电流变化较小，表现出近似恒流特性；当端电压逐渐接近开路电压时，电流快速下降，曲线非线性特征明显。

图2.4 光伏电池 I - V 与 P - V 特性曲线

光伏电池的 I - V 特性曲线及发电能力通常由 I_{sc} 、 V_{oc} 、 V_{mpp} 、 I_{mpp} 和 P_{mpp} 等参数共同表征。其中， I_{sc} 对应端电压为零时的短路电流， V_{oc} 对应输出电流降至零时的开路电压； V_{mpp} 和 I_{mpp} 分别表示最大功率点处的电压和电流，两者的乘积即最大输出功率 P_{mpp} 。为评价组件 I - V 曲线的近矩形程度，通常引入填充因子 FF 作为辅助指标。 P - V 特性曲线一般呈单峰分布，其峰值位置对应电池的最优工作区间。

2.3.2 辐照度与温度影响分析

光伏电池的输出响应对辐照度和温度变化较为敏感，这两项环境因素直接影响其输出特性。由单二极管模型可知，辐照度升高会使光生电流增加，因此短路电流和最大输出功率通常呈上升趋势；相比之下，开路电压的变化幅度相对较小。图2.5(a)显示，在温度保持不变时，随着辐照度由低到高变化， I - V 曲线整体向上抬升，说明辐照度对电流分量的影响更为直接。

图2.5(a) 不同辐照度下光伏电池 I - V 曲线

图2.5(b) 不同温度下光伏电池 P - V 曲线

温度变化对光伏电池输出特性的影响更集中地体现在电压侧。温度升高后，半导体内部载流子的热运动增强，反向饱和电流增大，开路电压随之下降，最大功率点也向低电压区域移动。图2.5(b)表明，在相同辐照度条件下，温度越高， P - V 特性曲线的峰值越低。由此可见，环境参数的变化本身即可引起曲线形态及关键特征量的偏移。

2.4 光伏阵列结构与建模

为考察串并联耦合对输出特性的影响，本文将 2×2 光伏阵列作为分析对象。该阵列由两条并联支路组成，每条支路包含两个串联光伏组件，属于典型的两串两并结构。相较于更大规模的阵列， 2×2 结构层次较为清晰，不仅能够体现串联升压、并联增流的基本规律，也便于后续设置不同工况并观察输出变化。

从拓扑关系来看，串联连接会使同一支路内各组件电流保持一致，支路端电压则为各组件端电压之和；并联连接则使各支路端电压保持一致，阵列总电流为各支路电流之和。图2.6给出了 2×2 光伏阵列的结构示意图。

图2.6 2×2 光伏阵列结构示意图

在单组件单二极管模型与 2×2 阵列拓扑关系分析的基础上，进一步构建了面向 2×2 光伏阵列的模块化仿真模型。该模型将单组件电气方程、阵列串并联关系以及故障注入过程整合到阵列数学核心模块中，使阵列输出特性能够在给定输入条件下逐点求解。不同于直接套用经验曲线，本文采用的处理方式从组件参数和电路约束出发，按照“输入—求解—输出”的路径计算 2×2 光伏阵列在不同辐照度、温度和故障条件下的输出响应。

进入模型计算的输入量包括阵列端电压扫描值、各组件辐照度矩阵、环境温度、故障模式、故障严重度以及故障作用的目标行列位置。其中，端电压扫描值用于确定当前计算点对应的阵列工作电压；辐照度矩阵和环境温度共同决定各组件的光生电流及温度修正量；故障模式、故障严重度和目标位置则用于确定需要修正的组件参数范围。对于 2×2 阵列而言，系统由两条并联支路组成，每条支路由两个组件串联构成，因此输入信息既要描述单组件状态，也要反映阵列拓扑关系。

阵列数学核心模块负责将单组件的输出特性映射为整个阵列的端口特性。对于单组件，系统先依据辐照度和温度对单二极管模型中的光生电流、反向饱和电流等参数进行更新，从而得到该组件在当前工况下的电压—电流关系。就故障注入而言，局部遮挡主要改变受影响组件的有效辐照度，老化故障对应组件输出能力下降，开路故障会限制所在支路的电流输出，短路故障则会改变对应组件或支路的电压分配。经过上述处理，不同故障工况便可体现为单组件输出特性或支路电气约束的变化。

在阵列连接关系层面，同一串联支路内的两个组件需满足电流一致，支路端电压则由两个组件端电压叠加得到；两条并联支路的端电压保持一致，阵列总电流则由两条支路电流相加得到。因此，在正常工况下， 2×2 阵列表现为串联升压和并联增流；当某一组件或某一支路发生故障时，其影响会通过串联电流约束和并联电流叠加关系传递到阵列输出端，进而形成不同的I-V特性曲线和P-V特性曲线。

为获得完整的输出特性曲线，仿真过程采用扫压法对阵列端电压进行连续扫描。具体而言，阵列端电压从0逐步增加至接近开路电压的范围，在每个电压采样点上由阵列数学核心模块计算对应电流，并进一步由 $P=V\times I$ 得到瞬时输出功率。当扫描电压较低时，曲线主要反映短路附近的电流输出能力；当电压接近最大功率点时，曲线拐点及功率峰值可表征阵列的有效输出能力；当电压继续升高并接近开路状态时，阵列电流逐渐减小至零。

模型输出端主要包括阵列I-V特性曲线、P-V特性曲线，以及由曲线提取的关键电气参数，如开路电压 V_{oc} 、短路电流 I_{sc} 、最大功率点电压 V_{mpp} 、最大功率点电流 I_{mpp} 和最大输出功率 P_{mpp} 。

图2.7 光伏阵列仿真流程图

第3章 光伏阵列故障输出特性分析

3.1 典型故障工况输出特性分析

本文以四块光伏组件构成的 2×2 阵列为分析对象，该阵列采用两串两并的连接方式，同时具备串联支路中的电压叠加特征和并联支路中的电流汇流特征。由于结构规模较小，既便于观察故障对曲线形态及输出参数的影响，也能够反映阵列内部失配、支路异常和功率重分配等典型现象。本文选取局部遮挡、老化、开路和短路四种故障作为研究对象。上述故障分别对应分布式光伏系统中的受光不均、性能衰减、导通路径中断和低阻旁路异常，具有较强的代表性。

为保证不同工况下曲线分析的可比性，仿真首先设置正常工况作为基准。正常工况下，四块组件的辐照度、温度及电气参数保持一致；随后在此基础上，分别调整局部组件的辐照度、组件等效参数或电气连接状态，以构建相应的故障工况。故障设置如表3.1所示。

表3.1 五种仿真工况参数设置表

工况	设置对象	仿真设置	主要作用
----	------	------	------

正常工况 四块组件 , $T=25^{\circ}\text{C}$ 对照基线

局部遮挡 M11组件 灯源, 遮挡后 等效 形成局部辐照不均

老化故障 第1并联支路 模拟线路或组件老化引起的串联损耗

开路故障 M11组件离线 所在支路等效开路 切断一条导通路径

短路故障 M11组件 使故障组件电压贡献降低

3.1.1 局部遮挡故障分析

在分布式光伏系统的运行过程中，局部遮挡是一类较为普遍的外部环境扰动。在实际运行过程中，光伏组件通常安装在屋顶、厂房顶部或空旷场地，容易受到周围环境影响。例如树叶、鸟巢、鸟粪、灰尘堆积、建筑物阴影以及附近设备投影等，都可能遮挡部分组件表面。由于遮挡通常只作用于某一块组件或某一局部区域，而不是整个阵列同时受影响，因此会造成阵列内部各组件接收辐照度不一致。被遮挡组件的有效光照强度降低后，其光生电流随之下降，在串联支路中会限制整条支路的电流输出，并进一步影响并联支路之间的电流分配。因此，局部遮挡故障往往会导致阵列输出电流和最大功率下降，严重时还会使曲线形态发生畸变。

在局部遮挡故障仿真中，外部光照基准设定为 1000 W/m^2 。随后在某一块组件表面引入遮挡条件，使该组件的实际等效辐照度低于未遮挡组件，其余组件仍保持 1000 W/m^2 、 25°C 的初始条件。在仿真中，遮挡程度通过遮挡系数折算为组件接收端的等效辐照度；本文将被遮挡组件等效辐照度设为600，其余组件仍为1000。图3.1为局部遮挡故障下阵列I-V曲线与正常工况曲线的对比。从曲线变化可以看出，局部遮挡后阵列输出电流能力明显下降，尤其在低电压区域，受遮挡组件所在支路不能提供与正常支路相同的电流，导致阵列总电流低于正常工况。

图3.1 局部遮挡故障I-V曲线对比图

图3.2对比了局部遮挡故障下阵列P-V特性曲线与正常工况曲线的差异。正常工况下，P-V特性曲线通常仅对应一个较为清晰的峰值，最大功率点位置也相对稳定；而在局部遮挡条件下，不同支路的最佳工作点不再一致，功率曲线峰值明显下降，最大功率点位置随之发生偏移。

图3.2 局部遮挡故障P-V曲线对比图

3.1.2 老化故障分析

老化故障主要与光伏组件长期运行过程中材料性能退化及电气连接能力下降有关。分布式光伏系统长期处于户外环境，组件持续受到高温、紫外线、雨水、湿热循环、风沙以及温度变化等因素作用。随着运行时间延长，组件封装材料可能发生老化，接线端子和焊带可能出现氧化、松动或接触不良，导线及连接部位的等效电阻也可能逐步增大。这类退化通常不会像开路故障或短路故障那样导致组件立即失效，而更多表现为输出能力持续减弱。本文采用增大串联电阻的方式对老化故障进行模拟，以反映老化后线路损耗增加、负载工作时电压降增大以及最大功率点电压和输出功率下降等现象。

老化故障用于描述组件长期运行后性能逐渐衰减的状态。在仿真中，老化故障采用串联电阻增大的方式设置，即在第1并联支路中引入额外老化电阻。与局部遮挡不同，老化故障通常不是外部遮挡造成的瞬时输出降低，而是组件材料、封装结构和电气连接状态退化后的综合表现。本文在设置老化工况时，保持阵列拓扑不变，在第1并联支路中串联 5Ω 老化电阻，用于模拟导线氧化、焊点接触不良和组件内部连接电阻增大造成的附加损耗。

图3.3展示了老化故障条件下阵列I-V特性曲线与正常工况曲线之间的对比结果。由仿真曲线可以看出，老化后I-V曲线整体向内收缩，低电压区电流输出能力下降，拐点附近变化更加平缓。由于老化电阻会在负载电流较大时产生额外压降，曲线拐点区域会变钝，填充因子减小，而开路点附近电流接近零，电阻压降较小，因此开路电压变化不明显。相比局部遮挡，老化故障一般不会形成特别明显的多段台阶，而更接近于输出能力的整体衰减。

图3.3 老化故障I-V曲线对比图

图3.4为老化故障下阵列P-V曲线对比图。老化工况下，最大功率点对应的功率值明显低于正常工况，功率峰值整体下降。串联电阻老化后，P-V曲线通常仍保持单峰形态，但峰值降低且最大功率点向较低电压区域偏移。

图3.4 老化故障P-V曲线对比图

3.1.3 开路故障分析

开路故障通常是由光伏组件或支路连接中断引起的。在实际分布式光伏系统中，组件接头松脱、连接器接触不良、导线断裂、接线端子脱落以及组件内部互联条断开等问题，都可能导致某一块光伏板离线。当光伏阵列中某一块组件离线时，其所在串联支路无法形成完整导通回路，该支路基本失去输出电流能力。由于另一条并联支路仍可继续工作，阵列输出电压不一定明显降低，但总输出电流会明显减小。因此，开路故障主要表现为电流侧参数下降，尤其是短路电流和最大功率点电流下降较为明显，最大输出功率也会随之降低。在仿真中开路故障按“11组件离线”处理，并将其所在支路等效为断路状态，使该支路无法形成正常电流通道。对于光伏阵列，如果一条并联支路发生开路，则该支路基本不再向外输出电流，阵列输出主要由剩余正常支路承担。

图3.5为开路故障下阵列I-V曲线与正常工况曲线的对比。开路故障对曲线最直接的影响体现在电流侧。正常工况下，两条并联支路共同提供输出电流，阵列短路电流接近两条支路电流之和；发生开路后，故障支路电流贡献消失或明显减小，阵列总电流随之下降。因此，开路工况下和通常低于正常工况，I-V曲线整体电流水平下移。

图3.5 开路故障I-V曲线对比图

图3.6为开路故障下阵列P-V曲线对比图。由于输出功率由输出电压和输出电流共同决定，开路故障造成的电流下降会直接导致最大功率降低。若一条支路完全开路，则阵列最大输出功率会接近剩余正常支路所能提供的功率，明显低于正常状态下两条支路共同输出的功率。此时最大功率点电流下降更为突出。

图3.6 开路故障P-V曲线对比图

3.1.4 短路故障分析

短路故障是指阵列中某一组件或局部线路被低阻通道旁路的异常状态，其常见诱因包括线路绝缘损坏、接线错误、接头进水以及组件内部局部击穿等。在仿真平台中，可通过将某一组件端口设置为低阻连接，使其两端电压接近零的方式模拟短路故障。与开路故障导致电流通道中断不同，短路故障通常不会使电流通路完全消失，而是使故障部分失去原有的电压贡献。

图3-7展示了短路故障下阵列I-V特性曲线与正常工况曲线的对比结果。相较于电流侧，短路故障对电压侧的影响更为明显。正常工况下，串联组件的端电压逐级叠加形成支路电压；当某一组件发生短路后，其端电压被显著压低，导致所在支路能够提供的最大电压减小，阵列输出电压范围也随之收缩。因此，在短路故障条件下，开路电压 V_{oc} 和最大功率点电压 V_{mpp} 通常低于正常工况，I-V特性曲线在电压方向上表现为左移或提前终止。

图3.7 短路故障I-V曲线对比图

图3.8为短路故障下阵列P-V曲线对比图。短路后，阵列在较高电压区域的输出能力受到限制，最大功率点通常向低电压区域移动，同时峰值功率下降。若短路仅发生在单个组件上，阵列仍可能保持一定电流输出，但由于有效串联电压减少，功率峰值明显降低。若短路范围扩大，P-V曲线的电压范围和功率峰值会进一步减小。

图3.8 短路故障P-V曲线对比图

3.2 五种工况输出特性对比分析

在完成四类故障曲线分析的基础上，进一步将其与正常工况结合，对五种工况之间的差异进行综合比较。本文主要从I-V特性曲线形态、P-V特性曲线形态、关键电气参数以及故障可辨识特征四个方面展开对比分析，图3.9为五种工况下I-V曲线对比图。正常工况曲线作为基准，整体变化平滑，电流和电压输出能力均处于较高水平。局部遮挡故障下，曲线可能出现分段下降或局部弯折，反映受遮挡支路与正常支路之间的输出失配。老化故障下，曲线一般表现为整体内缩，曲线形态较正常工况相似但输出能力下降。开路故障下，曲线在电流方向下降最为明显，说明支路导通路径减少后阵列电流输出能力受到限制。短路故障下，曲线在电压方向压缩更突出，表明故障组件无法提供正常电压贡献。

图3.9 五种工况曲线对比图

图3.10为五种工况下P-V曲线对比图。正常工况下功率曲线峰值最高且峰值位置稳定。局部遮挡工况下，P-V曲线峰值下降且位置发生偏移，其曲线形态与正常工况差异较大。老化故障下，功率峰值下降，但通常仍保持较平滑的单峰特征。开路故障下，由于电流能力下降，功率峰值明显降低，最大功率点电流对应值下降较为突出。短路故障下，最大功率点向低电压方向移

动，峰值功率随有效输出电压降低而下降。

图3.10 五种工况P-V曲线对比图

从关键电气参数角度看，五种工况之间的差异可以通过、、和进行归纳。正常工况下各参数作为基准值；局部遮挡故障对、和影响较明显，同时曲线峰值数和局部斜率变化具有诊断价值；老化故障主要表现为下降和填充因子减小，若老化程度较均匀，则和可能同步缓慢下降；开路故障对和最敏感，电流侧参数下降明显；短路故障对和更敏感，电压侧参数下降更突出。表3.2给出了五种工况关键参数变化趋势的对比。

表3.2 五种工况关键电气参数对比表

工况	功率下降/%						
正常工况	43.20	17.07	33.86	15.77	521.49	0.00	
局部遮挡	43.13	13.66	33.89	12.63	419.92	19.48	
老化故障	43.20	15.71	32.48	10.02	320.33	38.57	
开路故障	43.20	8.54	33.77	7.91	263.86	49.40	
短路故障	19.71	17.07	18.15	15.68	272.28	47.79	

由表3.2可以看出，虽然对所有故障都敏感，但它只能说明阵列输出能力下降，不能单独区分故障类型。若仅以最大功率下降作为判断依据，局部遮挡、老化、开路和短路都可能被归为异常状态，但无法进一步确定异常来源。因此，五种工况对比时应同时关注电流侧、电压侧和曲线形态特征。和更适合反映遮挡和开路等电流受限类故障；和更适合反映短路等电压受限类故障；P-V曲线峰值是否降低、最大功率点是否偏移也可用于辅助识别局部遮挡。

从曲线形态差异来看，局部遮挡在五种工况中的变化最为复杂，其电流受限现象和功率峰值偏移与阵列内部支路失配密切相关；老化故障的曲线变化相对平缓，容易与环境波动引起的功率下降相混淆，因此需要结合参数的长期变化及填充因子进行判断；开路故障主要表现为电流输出能力明显下降，曲线差异较为直观；短路故障则主要表现为电压输出范围收缩，最大功率点向低电压区域移动。通过上述对比，可将不同故障从“功率均下降”这一共同现象进一步区分为电流侧异常、电压侧异常、曲线形态异常和整体衰减异常等不同类型。

结合五种工况的对比结果可以看出，正常工况曲线可作为判断故障偏移的基准；局部遮挡主要对应支路失配及功率峰值下降；老化故障主要表现为输出能力平缓衰减；开路故障主要表现为电流输出能力降低；短路故障主要表现为电压输出范围压缩。上述差异可为后续故障诊断特征构建提供依据，但本章仅对曲线与参数变化规律进行分析，不涉及诊断模型训练过程。

表3.3 五种工况输出特征对比表

工况	I-V曲线特征	P-V曲线特征	敏感参数
正常工况	曲线平滑，电流与电压范围较完整	单峰且峰值最高	基准参数
局部遮挡	电流侧下降，曲线局部变形	峰值降低，峰值位置偏移	、、
老化故障	整体内缩，拐点变平缓	仍以单峰为主，峰值降低	、填充因子
开路故障	电流水平明显下移	功率峰值大幅降低	、
短路故障	电压范围被压缩	峰值向低电压区移动	、

3.3 本章小结

本章基于2×2光伏阵列仿真模型，设置了正常、局部遮挡、老化、开路和短路五种工况，并对各工况下的I-V曲线、P-V曲线及关键电气参数进行了对比分析。仿真结果表明，局部遮挡主要导致输出电流和最大功率下降，老化故障因串联电阻增大使最大功率点电压和功率明显降低，开路故障使离线组件所在支路失去输出能力，表现为电流侧参数大幅下降，短路故障则主要压缩阵列输出电压范围，使最大功率点向低电压区域偏移。通过对、、和等参数的比较，可以看出不同故障在电压侧、电流侧和功率侧具有不同的变化特征，这些特征可为后续故障诊断模型的特征构建提供依据。

第4章 光伏故障诊断模型构建与训练

4.1 机器学习简介

机器学习是一类利用已有样本数据建立输入特征与目标标签之间映射关系的方法。在分类任务中，模型通过训练样本学习不同类别之间的判别规律，并将该规律用于新样本的类别预测。与传统阈值判断方法相比，机器学习方法能够同时利用多个特征信息，适合处理特征维度较高、类别数量较多以及类别边界不完全线性的诊断问题[20]。

在光伏阵列故障诊断中，不同运行状态会引起I-V曲线、P-V曲线以及开路电压、短路电流、最大功率点等关键参数的变化。例如，短路故障通常会影响到电压输出范围，开路故障会降低电流输出能力，遮挡故障容易造成曲线局部突降，老化状态则会表现为整体功率衰减。这些故障差异可以通过特征提取转化为机器学习模型的输入，从而实现由曲线数据到故障类别的自动判定。

常见的机器学习分类模型包括支持向量机、随机森林、梯度提升树、XGBoost、多层感知器和AdaBoost等。不同模型在分类边界表达能力、参数规模、训练复杂度和可解释性方面各有特点。本文选择SVM作为主要诊断模型。其他模型仅用于后续性能对比，以判断当前特征体系是否具有普适性，不作为本章算法原理分析的重点。

从本文任务特点看，光伏故障诊断具有多特征、多类别和非线性边界等特点。SVM通过最大化分类间隔提高模型稳定性，并能够借助核函数处理非线性分类问题，适合用于本文的五类光伏阵列运行状态识别。

4.1.1 SVM基本思想

支持向量机是一种典型的监督分类方法，其基本思想是在特征空间中寻找能够区分不同类别样本的分类超平面，并使距离分类边界最近的样本到该边界的间隔尽可能大。靠近分类边界并对边界位置起决定作用的样本称为支持向量。分类间隔越大，模型对局部扰动和样本噪声的敏感程度通常越低。

在实际故障诊断数据中，不同类别样本往往难以做到完全线性可分。为提高模型对样本波动的适应能力，SVM 引入软间隔思想，允许少量样本位于分类间隔内或出现分类误差，并通过惩罚因子控制误分类代价。该机制能够在训练精度和泛化能力之间取得平衡，避免模型过度依赖少量异常样本。

对于光伏阵列故障诊断问题，不同故障类别之间的边界通常并非简单线性关系。例如，遮挡和老化都可能导致最大输出功率下降，短路和开路也可能在部分电气参数上出现相近变化。因此，本文采用带RBF核函数的SVM进行分类，使模型能够在非线性特征空间中建立故障类别之间的判别边界。

SVM 的分类决策函数可简写为：

式中，为待分类样本，为训练样本，为类别标签，为核函数，和为训练得到的模型参数。该表达式表明，SVM的分类结果由训练样本、类别标签和核函数共同决定。

4.1.2 RBF核函数与多分类处理

本文采用径向基函数（Radial Basis Function, RBF）作为SVM的核函数，其形式为：

式中，为核函数参数，用于控制样本相似度随距离变化的敏感程度。当两个样本在特征空间中距离较近时，核函数值较大；当样本距离较远时，核函数值较小。光伏故障特征中既包含最大功率、开路电压和短路电流等电气参数，也包含斜率、曲率和局部突降等曲线形态特征。该类特征之间存在一定非线性关系，采用RBF核有助于增强模型对非线性分类边界的表达能力。

标准SVM本身属于二分类模型，而本文诊断对象包括正常、短路、开路、遮挡和老化五类状态。在多分类任务中，可以通过多个二分类器组合完成最终类别判定。考虑到本文样本规模为840组，属于中小规模数据集，SVM 训练复杂度和部署成本均处于可接受范围。与结构更复杂的神经网络模型相比，SVM 参数数量较少，模型输出也更便于结合特征重要性和混淆矩阵进行分析。

4.2 模型构建

如图4-1所示，本文光伏故障诊断流程首先以实测 I-V 曲线及其对应故障标签作为原始数据来源，对样本进行数据清洗和类别划分；随后从曲线中提取基础电气特征、曲线形态特征、故障专用特征和频域特征，形成 101 维诊断特征向量；在模型训练前，进一步通过特征筛选和 PCA 降维减少冗余信息，得到适合 SVM 输入的低维特征空间；最后完成 SVM 多类型故障分类训练与评价，并将训练后的模型保存，用于诊断调用。

图4.1 光伏故障诊断模型训练与调用流程图

4.2.1 训练数据来源与样本整理

本章训练数据来源于课题实验过程中自行采集和整理的光伏I-V曲线数据，数据采集时，通过光伏测试电路获得组件端电压、电流等输出信息，并结合实验设定的运行状态进行人工标注。每条样本对应一次I-V曲线采集结果，原始数据包含一组按扫描顺序记录的电压、电流点，并附带故障类别标签。各类别样本分布见表4.1。

表4.1 实测训练样本类别分布表

状态名称	标签代码	总样本数	占比
正常	normal	200	23.81%
短路	short	145	17.26%
开路	open	200	23.81%
遮挡	shelter	200	23.81%
老化	old	95	11.31%
合计	-	840	100%

4.2.2 特征提取与数据预处理

I-V曲线中包含了故障状态下电压、电流和功率变化的综合信息。第三章的分析表明，短路、开路、遮挡和老化等故障会在开路电压、短路电流、最大功率点以及曲线弯曲程度等方面表现出不同变化。因此，在模型训练前需要将原始曲线转换为能够反映这些差异的数值特征。提前流程图如图4.2所示。

图4.2 I-V曲线特征提取流程图

4.2.3 101维诊断特征构建

完成曲线清洗和平滑处理后，本文从I-V曲线和P-V曲线中提取101维诊断特征。特征构建并非简单叠加原始采样点，而是围绕故障机理提取能够反映输出能力、曲线形态和局部异常的综合特征。101维故障诊断特征既包含开路电压、短路电流、最大功率点等基础电气参数，也包含曲线斜率、曲率、面积以及统计量等形态信息。对于遮挡、老化等不易仅凭单点电气量区分的状态，还加入了故障专用特征和频域特征，以增强模型对曲线整体变化的识别能力。特征组成如表4-2所示。

表4.2 故障诊断特征组成表

特征类别	维度	主要内容	作用
基础电气特征	19	开路电压、短路电流、最大功率点电压、最大功率点电流、最大功率、填充因子等	反映组件整体输出能力和关键工作点变化
曲线形态特征	50	I-V曲线斜率、曲率、面积、分段变化率和局部形态参数等	描述曲线弯折、平台段和下降段变化
故障专用特征	12	针对短路、开路、遮挡、老化等状态构造的比例和差异特征	突出典型故障引起的局部异常
频域与统计特征	20	均值、方差、偏度、峰度及曲线序列变换特征等	补充曲线整体波动和分布信息
合计	101	-	构成模型输入特征向量

完成特征提取后，对数据进行了统一预处理。对于个别缺失或异常值，采用同类别或整体统计值进行修正，避免模型训练时受到无效记录干扰。由于电压、电流、功率及统计特征的量纲不同，本文采用稳健标准化方法对特征进行缩放，其形式为：

式中，为中位数，和分别为第一、第三四分位数。与普通标准化相比，该方法对个别极端值不太敏感，适合实测数据中存在少量波动的情况。随后根据随机森林特征贡献度进行筛选，将原始101维特征压缩为51维有效特征，并进一步采用PCA保留95%方差，最终得到5维主成分表示。PCA前5个主成分的累计方差贡献率约为96.37%，能够保留大部分原始特征信息，同时降低模型输入冗余。

4.2.4 特征筛选与PCA降维

101维诊断特征中既包含主要判别信息，也可能存在一定冗余。SVM使用RBF核进行分类时，需要计算样本之间的距离，若不同特征数值尺度差异较大，可能导致模型过度关注取值范围较大的特征。因此，在模型训练前，本文先采用中位数填充缺失值，再使用稳健标准化方法对特征尺度进行统一处理。

标准化处理后，本文利用随机森林特征重要度结果进行特征筛选，将101维特征压缩至 51 维。该过程主要用于去除作用较弱或冗余程度较高的输入项，并不改变SVM作为最终诊断模型的地位。随后采用PCA对筛选后的特征继续降维，最终保留5个主成分，累计解释方差为96.47%。该结果表明，少量主成分已经能够保留原始特征中的主要判别信息。

PCA 二维分布见图4.3。图中展示前两个主成分，用于观察不同故障样本在低维空间中的分布趋势。由于部分开路样本在比值和斜率特征上取值较大，二维可视化时会拉伸坐标尺度。为便于观察主体样本的分布情况，图4.3采用局部显示方式展示主要样本区域。需要说明的是，模型训练、测试和评价均基于完整样本完成，未改变原始数据集。

图4.3 PCA降维二维分布图

从PCA图可以看出，正常、短路、遮挡和老化样本在前两个主成分平面上形成了较清晰的分布区域；开路样本由于故障特征较为特殊，主体样本集中于图中左侧附近。

4.3 SVM模型训练过程

4.3.1 SVM参数设置

本文使用SVC构建SVM分类器，主要参数见表4.3。

表4.3 SVM模型主要参数表

参数名称	设置值	使用说明
Kernel	rbf	使用径向基核函数，处理非线性分类边界
C	15	控制误分类惩罚强度，平衡训练精度和泛化能力
Gamma	true	按特征维度和方差自动确定RBF核参数
Probability	balanced	输出类别概率，便于后续显示诊断置信度
Random_state	42	根据类别数量调整权重，缓解样本不平衡

在上述参数中，RBF核函数是SVM实现非线性分类的关键设置。当前故障特征包含电气参数和曲线形态信息，不同类别之间的边界并非简单线性关系，因此采用RBF核更符合本文样本特征。参数 C 用于控制误分类惩罚强度，在训练精度和泛化能力之间取得平衡。参数 用于获得类别概率输出。

4.4.3 训练与模型保存

模型训练流程包括数据读取、标签映射、曲线预处理、特征提取、缺失值填充、标准化、特征筛选、PCA降维、SVM训练和测试集评价等步骤。训练完成后，预处理器、特征筛选器、PCA降维器和SVM分类器分别保存为joblib文件。

模型保存的目的在于保证后续系统部署阶段能够复用训练阶段形成的完整处理流程。服务器端接收到采集端上传的80对I-V数据后，需要按照与训练阶段一致的流程完成特征构建和降维处理，再输入已训练的SVM模型进行类别判定。若训练阶段与部署阶段的特征处理方式不一致，即使分类器本身参数正确，也可能导致诊断结果出现偏差。模型训练流程图如图4.4所示。

图4.4 模型训练流程图

4.4 SVM模型性能评估

4.4.1 SVM测试结果

在当前测试集上，SVM 的 Accuracy、Precision、Recall和Weighted F1均为1.0000。该结果表明，在当前样本划分条件下，经过特征提取、标准化、特征筛选和PCA降维处理后，五类样本能够被SVM有效区分。

测试集分类结果通过混淆矩阵热力图表示，如图4.5所示。五类样本均位于混淆矩阵主对角线位置，非对角线误判数量为0。

图4.5 SVM混淆矩阵热力图

具体来看，测试集中的40组正常样本、29组短路样本、40组开路样本、40组遮挡样本和19组老化样本均被正确识别。结合第三章故障曲线分析可知，短路、开路、遮挡和老化在电压范围、电流输出能力、局部突降和功率衰减等方面存在可观测差异。经过特征工程处理后，这些差异能够转化为SVM可利用的分类信息。

4.4.2 特征重要性分析

RBF核SVM不像线性模型那样直接给出各输入特征的线性系数。为分析主要输入特征对分类任务的作用，本文结合特征筛选阶段得到的随机森林重要度结果进行解释。主要特征的重要性热力图如图4.6所示。

图4.6 特征重要性热力图

由图4-6可知，重要特征包括、、、、和。其中，反映阵列最大输出能力，可用于判断故障造成的功率损失；、、和分别反映电压、电流及最大功率点附近的变化；和描述曲线中段形态；与开路状态下的电流平台变化相关。

这些特征共同对应第三章中分析的故障输出规律。SVM的输入并非依赖某一单独参数，而是综合利用输出能力、曲线形态和局部故障特征完成分类。这也说明当前特征集合在故障机理和模型输入之间具有一定一致性。

4.5 本章小结

本章围绕光伏I-V曲线故障诊断模型的构建与训练展开，整理了840条带标签I-V曲线数据，并按正常、短路、开路、遮挡和老化五类完成训练集与测试集划分；在特征构建方面，本文采用101维故障诊断特征，并经过缺失值处理、稳健标准化、特征筛选和PCA降维后输入分类模型；在模型训练方面，以SVM为核心模型作为诊断核心，结果表明，基于实测I-V曲线特征的模型能够较好地区分主要故障类别，优化集成模型在测试集上取得较高的准确率和F1值，为后续服务器端在线诊断调用提供了模型基础。

第5章 光伏故障诊断系统实现与实验验证

5.1 实验平台总体构成

为检验前文故障诊断模型能否进入实际运行流程，本文在实验阶段搭建了基于I-V特性曲线的分布式光伏故障诊断平台。平台的设计重点放在光伏输出数据采集、I-V数据传输、诊断模型调用和结果显示几个环节上，用于验证本文方法从算法模型到系统应用的落地过程。

实验平台由光伏阵列及模拟太阳装置、三合一DC-DC采集板、STM32F407ZG主控板、部署故障诊断模型的云服务器以及AIRIOT云平台组成。系统运行时，光伏阵列在不同工况下产生电压、电流输出，采集板获取光伏输出侧数据，主控板完成数据规整和上传，云服务器对I-V数据进行处理并调用模型，最后由云平台显示诊断结果。整体组成和数据流向如图5.1所示，实验平台实物图如图5.2所示。

图5.1 光伏故障诊断实验平台组成图

图5.2 光伏故障诊断实验平台实物图

5.2 实验故障工况设置

为检验系统对不同状态的识别能力，本文设置正常、局部遮挡、老化、开路和短路五类工况。正常工况作为对照组，用于检查系统能否识别无故障状态；其余四类工况对应分布式光伏运行中较常见的异常类型。

局部遮挡通过遮挡部分组件或降低局部模拟光照强度实现。该工况主要模拟树影、建筑物遮挡和积灰不均等情况。老化工况通过串联电阻降低组件等效输出实现，用于模拟组件长期运行后的性能下降。开路故障通过断开一路组件或支路连接实现，短路故障通过短接局部组件或负载实现。各类工况的设置方式见表5.1。

表5.1 典型故障工况设置表

工况类别	设置方式	验证目的
正常工况	光伏阵列处于正常照射和正常连接状态	检查系统对正常状态的识别情况
局部遮挡	遮挡部分组件	检查系统对遮挡类故障的识别情况
老化故障	串联电阻降低组件等效输出	检查系统对性能衰减类故障的识别情况
开路故障	断开一路组件或支路连接	检查系统对支路断开异常的识别情况
短路故障	短接局部组件或负载	检查系统对短路异常的识别情况

5.3 实验平台组成

5.3.1 光伏阵列及模拟太阳装置

光伏阵列及模拟太阳装置是实验平台的前端对象。实验中通过模拟太阳装置为光伏组件提供可控光照条件，使组件在正常状态和故障状态下产生对应的电压、电流输出。

正常工况下，光伏组件保持完整连接并接受稳定光照，输出电压、电流随工作点变化形成标准I-V曲线。局部遮挡工况通过对部分光伏板进行物理遮挡实现，用于模拟树影、建筑物阴影、积灰不均等实际场景。老化工况通过在支路中串联额外电阻降低组件等效输出能力实现，以此模拟光伏组件长期运行后的性能衰减。开路故障通过拔掉一块光伏板的连接线实现，对应实际运行中的支路断开、接线松脱问题；短路故障通过对支路负载进行短接实现，对应线路绝缘破损、接线错误引发的局部短接异常。

5.3.2 DC-DC电压电流采集板

电压电流采集是本系统中连接光伏阵列与后续诊断算法的重要环节，其采样结果直接影响I-V特征提取和故障识别的准确性。本文采集端采用三合一DC-DC采集板完成光伏侧输出数据获取与工作点调节。该采集板以STM32G474CE为控制核心，基于 ARM Cortex-M4 内核，具备电压电流采样、PWM 调节、数据转换与滤波处理能力，能够满足光伏输出端数据采集及工作状态调节的需求。

采集板采用四开关同步Buck-Boost拓扑结构，可在一定输入输出条件下实现升压和降压调节。其输入电压范围为12~48VDC，输入电流不超过8A，输出电压范围为1~48VDC，输出电流不超过8A，额定功率为150W。该采集板还具备输入过压、欠压、输出过压、过流、过温及电流采样零漂异常等运行状态检测与保护设计。

在故障诊断实验中，采集板的主要作用是获取光伏侧电压、电流数据，并配合DC-DC变换过程形成I-V曲线离散采样序列。采样过程中，光伏阵列接入采集板输入端，系统以输入端电压和输入端电流作为光伏侧输出数据。采集板在接收到I-V曲线采集指令后，通过改变DC-DC变换器的PWM调节量，使光伏阵列依次工作在不同运行点，并记录各工作点下对应的电压、电流值。

图5.3 电流电压采样原理图

5.3.3 主控模块

主控板位于三合一DC-DC采集板与诊断端之间，是系统数据传输与采样管理的中间环节。本文实验平台主控单元采用STM32F407ZG作为核心控制器，该芯片基于ARMCortex-M4内核，最高主频为168MHz，具有较为丰富的片上资源和外设接口，能够满足系统对串口通信、数据缓存、状态管理和通信调度等功能的需求。

主控板接收数据后，按照故障诊断模型的输入要求对采样点进行缓存和格式整理，形成电压、电流一一对应的 I-V 样本，对于采样点数量不足、字段缺失或通信中断导致的不完整样本，主控端通过异常标记或重新采集的方式进行处理，以保证进入诊断流程的数据具有较好的完整性和一致性。

主控通信部分采用塔石4G模块实现远程上传，该模块支持4G网络传输和串口数据交互，能够满足分布式光伏场景下采集端与服务器端之间的数据连接需求。本文实验中的数据传输路径为“三合一DC-DC采集板—STM32F407ZG主控板—4G通信模块—服务器端诊断程序”。在该链路中，主控板一方面完成采集数据的整理与转发，另一方面根据工况切换和诊断端处理状态协调采样与上传节奏，减少重复或无效数据传输，从而为后续故障诊断提供规范的I-V数据输入。

图5.4 主控芯片原理图

5.3.4 诊断模型部署

本文故障诊断模型部署在服务器端。服务器接收主控板上传的I-V数据后，运行预处理和特征提取程序，并调用第四章训练完成的诊断模型。与采集端和主控端相比，服务器端更适合承担模型计算、结果存储和平台转发等任务，也便于后续接入更多采集节点。

诊断端接收到主控板上传的数据后，先检查数据格式，包括采样点数量、电压电流字段顺序和数值范围。通过检查的数据进入预处理环节，主要包括异常点修正、平滑处理和归一化处理。随后，程序提取与故障相关的电气特征和曲线特征，例如开路电压、短路电流、最大功率、最大功率点电压、最大功率点电流、填充因子以及曲线局部变化特征等。

图5.5 云服务器模型部署图

5.3.5 AIRIOT云平台

云平台显示模块用于展示系统运行状态和诊断结果。采用AIRIOT物联网平台实现数据展示、设备点数据查看和异常报警等功能。平台界面主要显示当前电压、电流数据、I-V曲线、采集状态、上传状态和故障识别结果。实验人员通过浏览器进入平台

后，可以查看当前样本对应的诊断类别，也可以查看历史实验记录。图5.6为云平台运维界面，图5.7为云平台诊断结果显示界面，图5.8为系统报警图，图5.9为单块诊断图。

图 5.6 云平台运维界面

图5.7云平台诊断结果显示界面

图5.8 单块运行诊断结果

5.4 本章小结

本章围绕分布式光伏故障诊断系统的实现与实验验证展开，搭建了由光伏阵列及模拟太阳装置、三合一DC-DC采集板、STM32F407ZG主控板、4G通信模块、服务器端诊断程序和AIRIOT云平台组成的实验平台，形成了从现场数据采集到诊断结果显示的完整流程。系统运行过程中，采集板完成光伏侧电压、电流获取和I-V样本形成，主控板与4G模块负责数据整理和远程上传，服务器端完成数据处理与模型诊断，云平台用于结果展示。通过正常、局部遮挡、老化、开路和短路五类工况测试，验证了系统完成 I-V 数据采集、上传和故障识别的可行性。

第6章 总结与展望

6.1 全文总结

本文围绕分布式光伏系统故障诊断问题，以光伏组件I-V特性曲线为研究对象，完成了光伏建模、故障特性分析、诊断模型构建及系统实验验证工作。相较于单一电气量数据，I-V特性曲线能更完整反映阵列不同运行状态的输出变化，适合作为故障识别的核心依据。主要工作包括：

(1) 基于电池等效电路模型分析输出特性，结合串并联关系建立光伏阵列模型，明确了光照、温度及电池参数对I-V、P-V曲线的影响规律。选取2×2光伏阵列为研究对象，其结构能体现组件串并联关系及局部异常的影响，满足本科阶段系统设计与验证的需求。

(2) 针对局部遮挡、老化、开路、短路四类典型故障，对比分析了正常与故障状态下I-V、P-V曲线的差异，明确了不同故障对短路电流、开路电压、最大功率点及曲线形态的影响规律。

(3) 从I-V曲线及相关参数中提取运行状态特征，采用SVM机器学习方法构建混合诊断模型。经样本构建、模型训练与测试，验证了基于I-V特性的故障诊断可行性。实验结果表明，在设定的样本范围和故障类别内，该模型可有效区分典型故障状态。

(4) 结合前述建模与诊断模型，完成了实验平台用分布式光伏故障诊断系统设计，实现了数据采集、处理、特征提取、模型诊断及结果显示的完整流程。实验验证表明，系统可完成典型故障数据的处理与识别，基本达到设计目标。

6.2 未来展望

后续可从贴近实际应用的角度进一步开展研究。研究以2×2光伏阵列为对象，而实际分布式光伏系统组件更多、组串结构更复杂，单个组件异常的影响易被削弱或叠加。后续可建立更大规模阵列模型，分析多组串、多支路下I-V曲线的变化规律，优化故障特征提取方法。主要改进方向包括：

(1) 样本主要来自实验平台，条件相对可控。实际运行中光照、温度、灰尘、组件差异及噪声等均会影响输出曲线。后续可引入真实电站运行数据，结合仿真与实验数据构建多样化样本集，增加不同季节、天气及运行年限的数据，提升模型在复杂环境中的适应性。

(2) 研究仅分析了四类典型单一故障，而实际运行中存在复合故障、渐变故障及早期轻微异常，其初期特征不明显，易与正常波动混淆。后续可补充相关故障样本，研究更精细的特征提取方法，提升系统对复杂状态的识别能力。

(3) 研究采用SVM方法验证了诊断思路的可行性。后续可优化特征选择、参数寻优及样本平衡方法，降低模型对特定实验条件的依赖；进一步结合物理机理与数据驱动方法，增强模型的可解释性；并通过更多测试验证模型在不同组件、阵列及环境下的通用性。

后续需围绕扩大阵列规模、引入真实数据、丰富故障类型、提升模型泛化能力及完善系统工程化实现展开研究，推动该诊断方法从实验验证走向现场应用。

参考文献

杨洪雷. 分布式光伏电站运维管理与发展趋势[J]. 上海节能, 2022, (09)

杨文, 倪志强. 光伏电站设备状态监测与故障诊断技术研究与应用[J]. 电力与能源, 2025, 46(01):16-19+48.

乔苏朋, 杨艳, 陈世群, 等. 光伏阵列故障检测方法综述[J]. 电气技术, 2021, 22(07):1-6.

唐克洪, 周慧, 徐俊. 分布式光伏电站智能化运维系统研究与应用[J]. 科技创新与应用, 2023, 13(19):193-196.

周盛龙, 常梦星, 杨鹤松, 等. 光伏电站在线监测及故障诊断系统设计[J]. 电气技术与经济, 2024, (04):95-97+102.

孙建民, 梁凌, 李庚达, 等. 光伏组件故障诊断技术综述[J]. 太阳能, 2022, (02):12-22.

魏红平, 俞东波, 马发明, 等. 光伏阵列故障智能诊断方法研究综述[J]. 机械研究与应用, 2025, 38(06):176-179.

Liu Y, Wu Y. Fault diagnosis of photovoltaic modules: A review[J]. Solar Energy, 2025, 293113489-113489.

余辉. 基于机器学习的光伏I-V输出特性建模与故障诊断研究[D]. 福州大学, 2021.

邱宇鹏. 基于I-V特性的光伏组件故障智能诊断研究[D]. 扬州大学, 2024.

王一鸣, 许颇, 杨珂, 等. 基于I-V曲线距离特征的光伏阵列故障诊断[J]. 电子设计工程, 2025, 33(16):51-54+59.

彭自然, 许怀顺, 肖仲平, 等. 深度优化的集成学习模型EKSSA-CatBoost:实现光伏阵列故障高精度智能诊断[J]. 仪器仪表学报, 2025, 46(05):324-338.

朱壮华, 张卫平, 王文彦, 等. 基于图神经网络的光伏组串故障诊断[J]. 分布式能源, 2024, 9(04):78-85.

李睿智, 俞东波, 陈道林, 等. 基于大数据分析平台的光伏阵列故障诊断系统[J]. 科技风, 2025, (36):1-3.

土俊鹏. 基于深度学习的光伏阵列故障诊断模型[J]. 灯与照明, 2025, 49(05):133-135.

曹岚, 周成功, 袁斌霞, 等. 基于梯度提升回归树-自训练贝叶斯优化支持向量机的光伏阵列故障诊断方法[J]. 太阳能学报, 2025, 46(06):289-297.

江颖, 关艳飞, 刘磊. 基于人工智能的光伏发电系统故障诊断与预测技术研究[J]. 中国设备工程, 2025, (14):156-158.

Zeng W, Cai Y, Fan J. A Mamba-Transformer-based deep learning framework for photovoltaic fault diagnosis in infrared images[J]. Advanced Engineering Informatics, 2026, 74(PA):104633-104633.

Shin W, Lee S J, Ju C Y, et al. CNN-based photovoltaic fault diagnosis using normalized I-V curves with Explainability analysis[J]. Solar Energy, 2025, 301113958-113958.

翟嘉琪, 杨希祥, 程玉强, 等. 机器学习在故障检测与诊断领域应用综述[J]. 计算机测量与控制, 2021, 29(03):1-9.

致 谢

说明:

- 1、支持中、英文内容检测;
- 2、AI特征值=AI特征字符数/总字符数;
- 3、红色代表AI特征显著部分, 计入AI特征字符数;
- 4、棕色代表AI特征疑似部分, 未计入AI特征字符数;
- 5、检测结果仅供参考, 最终判定是否存在学术不端行为时, 需结合人工复核、机构审查以及具体学术政策的综合应用进行审慎判断。



关注微信公众号