

编号		
所属学科	F22	计算机应用技术

B. 自然科学类

## 陕西省教育厅专项科研项目 申请书

项目名称：基于 AI 的红外与可见光融合算法研究

项目负责人：吴天

所在单位：西安电力高等专科学校

申请日期：2025-09-01

陕西省教育厅

二零一六年制

研究项目	名称	基于 AI 的红外与可见光融合算法研究					
	所属学科	F22	计算机应用技术				
	申请金额	2. 0000	起止年月	2025-09-01 — 2027-09-01			
	依托科研平台名称						
项目负责人	姓名	吴天	性别	女	出生日期	1998-03-25	
	学位	硕士	最后学位获得时间	2023-06-12	是否硕导、博导	否	
	专业技术职务	276	其它	定职时间			
	主要研究方向	红外光与可见光融合算法、激光SLAM、视觉SLAM、人工智能	联系电话	15503682801	所在学校	西安电力高等专科学校	

项目组成员						
姓名	年龄	专业技术职务	学位	所在单位	项目分工	本人签字
赵凯航	27	其它	硕士	西安电力高等专科学校	技术调研与需求分析、数据集构建	
王淼	47	高级讲师	硕士	西安电力高等专科学校	开发平台搭建	
负晓博	46	高级工程	学士	西安电力高等	数据集与	

		师		专科学校	算法优化	
陈婉莹	29	其它	硕士	西安电力高等 专科学校	系统优化 与完善	
崔新雨	27	其它	硕士	西安电力高等 专科学校	成果总结 与分析	
姜宇涵	27	其它	硕士	西安电力高等 专科学校	技术调研 与需求分 析	

主要研究内容	<p>本研究针对传统红外与可见光融合算法依赖手工规则设计、复杂环境下适配性差的核心痛点，以提升电力设备监测精度与运维效率为目标，聚焦基于生成对抗网络（GAN）的融合算法研发，确立 “损失函数优化 - 网络架构设计 - 场景适应性迭代” 的技术路线，旨在实现红外图像热辐射信息与可见光图像纹理细节的高效协同融合，同时结合国家电网变电站、特高压线路、电缆隧道等核心运维场景的特性开展定制化优化，确保算法在电力复杂场景中具备高鲁棒性与实用性，为电网智能化运维提供技术支撑。</p> <p>在算法核心机制研究中，首要突破 “热信息精准保留与结构细节清晰捕捉” 的平衡难题。其一，设计双目标导向的损失函数体系：生成器损失由对抗损失与内容损失构成，对抗损失通过生成器与判别器的动态博弈，驱动融合图像在纹理特征上逼近可见光图像（如变压器瓷套纹理、杆塔结构细节），确保结构缺陷可识别；内容损失采用弗罗贝尼乌斯范数约束，使融合图像与红外图像保持像素强度一致性，精准保留设备热辐射特征（如电缆接头过热区域、绝缘子温升异常），同时引入梯度算子约束匹配可见光图像的结构梯度，避免纹理失真，并通过可调节权重参数平衡热信息与结构信息的贡献占比，防止某类信息被压制。</p> <p>其二，优化网络架构以保障融合效率与精度：生成器采用五层卷积结构，全程避免下采样操作，减少图像细节丢失，前四层引入批归一化层与 Leaky ReLU 激活函数，有效缓解梯度消失问题，最后一层采用 tanh 激活函数输出归一化的融合图像；判别器同样为五层卷积架构，通过步长为 2 的卷积层替代传统池化层实现特征降维，在降低计算复杂度的同时保障分类精度，最终实现算法端到端运行，无需人工干预设计融合规则，大幅提升适配性。</p>
--------	---

研究内容摘要	本研究针对传统融合算法依赖手工规则、复杂场景适配性差的痛点，聚焦基于生成对抗网络的红外与可见光融合算法研发。通过优化双目标损失函数与网络架构，实现红外热辐射与可见光纹理高效融合。结合国家电网多场景特性优化算法，形成高鲁棒性基础算法框架，可支撑变电站、特高压线路等多项目,为电网智能化运维提供技术支撑。				
先进性、创新点	技术架构创新：突破传统算法手工规则依赖，构建基于 GAN 的端到端融合框架，生成器/判别器均采用五层卷积结构，避免下采样操作减少细节丢失，通过双目标损失函数平衡红外热辐射保留与可见光纹理捕捉，实现热信息与结构细节协同融合。 工程价值创新：算法兼具轻量化与高精度，可适配无人机、杆上终端等设备，支撑多项目智能化升级，且能扩展至多分辨率融合任务，为电网运维从 “被动检修” 向 “主动预防” 转型提供核心技术支撑。				
技术指标	视觉效果：融合图像清晰保留红外目标热辐射（高亮）与可见光纹理细节（背景、边缘），无重影、错位等配准异常。 指标表现：熵（EN）、空间频率（SF）高于源图像；融合图与源图的 SSIM、CC 值高；标准差（SD）、视觉信息保真度（VIF）优于传统方法。 工程指标：嵌入式平台 RV1126 上帧率 $\geq 25\text{fps}$ ，0-100 米距离配准成功率 $\geq 95\%$ 。				
预期成果	（1）软件成果类 实现基于 AI 的红外与可见光融合系统一套 （2）知识产权类 发表高水平论文一篇 登记软件著作权两项 （3）技术文件类 撰写完成：基于 AI 的红外与可见光融合系统研究工作报告				
新技术：	1	新工艺：		新材料：	
新装备：		新产品：		新品种：	
软件著作权：	2	集成电路布图：		药证：	
标准：		规范：		其他：	

论文：	1	著作：		研究报告：	1
专利：		培养博士：		培养硕士：	
资政、调研报告：					

工作进度安排(每项不超过 60 字)	
2025-09-01 — 2025-10-01	1.技术调研与需求分析 2.项目组组建与分工 3.计划制定与计划书撰写
2025-10-01 — 2026-06-01	1.任务拆解与详细方案制定 2.开发平台搭建 3.数据集构建 4.核心算法初研 5.基础模型与算法实现
2026-06-01 — 2026-12-01	1.数据集与算法优化 2.系统集成与测试 3.全流程算法研发与实现 4.阶段性总结与交流
2026-12-01 — 2027-08-31	1.系统优化与完善 2.多场景综合测试 3.成果总结与分析 4.成果提交与验收

项目负责人承担教育厅项目情况				
项目编号	项目名称	批准时间	结束时间	完成情况

# 承 诺 书

本人承诺严格遵守法律法规，服从有关文件要求和管理，确保所填报内容真实有效，数据准确客观，符合有关保密规定和学术道德规范要求，无知识产权等方面的争议。如存在违法违规违纪问题，自愿接受有关调查和处理。

承 诺 人：

年 月 日

## 单位审查与保证

1、系、所或依托平台审查意见（包括：审查项目的科学意义、实用价值、研究工作基础、经费预算合理性、申请者业务素质、研究能力和科学作风）

负责人（签字）：

单位（公章）

年 月 日

2、学校审查意见：

已对申请书进行了审核，同意上报。承诺在项目获得资助后做到以下几点：

- (1) 保证对研究计划所实施所需的人力、物力和工作时间等条件予以支持；
- (2) 严格遵守陕西省教育厅有关资助项目管理、财务管理等各项规定；
- (3) 督促项目负责人和本单位项目管理部门按教育厅的规定及时报送有关报表和材料；

需要说明的其他问题:

单位负责人 (签章)

单位（公章）

年 月 日



# 陕西省教育厅自然科学研究项目 可行性研究报告

项目名称：基于 AI 的红外与可见光融合算法研究  
项目负责人：吴天  
所在单位：西安电力高等专科学校  
申请日期：2025 年 8 月 25 日

陕西省教育厅

# 目 录

1	项目背景及意义 .....	1
2	国内外技术研究现状 .....	4
2.1	传统红外与可见光融合算法 .....	5
2.2	基于深度学习的红外与可见光融合算法 .....	8
2.3	融合系统在国家电网公司的应用及实践案例 .....	12
3	研究内容及方法 .....	14
3.1	研究方法概述 .....	15
3.2	损失函数设计 .....	16
3.3	网络架构设计 .....	17
3.4	算法评估指标 .....	18
3.5	算法优化方向 .....	20
3.6	算法预期结果 .....	21
4	项目进度安排 .....	21
5	经费预算 .....	22
6	预期成果 .....	23
7	推广应用前景分析 .....	24

# 1 项目背景及意义

当前全球电力系统正处于深度变革期，互联网技术与新能源技术的深度融合驱动“智能电网”成为能源转型的核心载体。对于国家电网公司而言，打造“坚强智能电网”不仅是支撑可再生能源消纳与低碳经济发展的战略举措，更需以安全性、可靠性、自适应性为核心突破制约其发展的关键瓶颈。中国作为拥有超 14 亿人口的电力消费大国，近年来伴随综合国力提升与民生需求升级，用电设备与负荷呈现指数级增长态势，这对电力系统的安全稳定运行提出了前所未有的挑战，也迫切需要先进技术手段为电网运维与管控提供全流程支撑。在此背景下，红外与可见光融合系统凭借红外图像的热辐射感知优势与可见光图像的细节表征能力，形成强互补特性，成为国家电网“坚强智能电网”建设中覆盖多核心环节的关键技术支撑，其应用深度贯穿设备监测、缺陷诊断、故障预警及安防管控等全流程，为电网智能化升级提供核心技术保障。

变电站作为国家电网电力传输与分配的核心枢纽，其内部变压器、隔离开关、断路器等高压设备的实时热状态（如局部异常发热）与外观完整性（如机械损伤、瓷件污秽）直接决定电网整体安全水平，因此变电站电力设备监测是“坚强智能电网”稳定运行的基础支撑。传统单一模态监测技术难以满足“坚强智能电网”对设备状态感知的高精度需求：仅依赖红外技术虽可捕捉设备热辐射信号，实现夜间或雾、雨等恶劣天气下的持续监测，但受限于图像分辨率与纹理表征能力，无法清晰呈现设备结构细节，导致热异常对应的具体部件定位困难；而单一可见光技术虽能提供高分辨率的设备轮廓与纹理信息，却高度依赖光照条件，在夜间、低光或恶劣天气下易出现目标丢失与画面模糊，难以发现设备隐性热故障。红外与可见光融合系统通过双传感器协同采集、空间配准与多模态融合算法，有效整合红外图像的热目标信息与可见光图像的细节特征，弥补单一模态缺陷。其生成的融合图像可同时清晰呈现设备热分布与结构细节，为国家电网变电站无人值守模式提供可靠数据支撑——运维人员可通过融合图像精准定位故障点（如变压器高压侧热异常与对应套管瓷件污秽），实现设备状态的实时、全面感知，切实保障高压设备稳定运行，契合“坚强智能电网”对运维智能化的核心需求。

在设备状态精准感知的基础上，电力设备缺陷智能诊断是国家电网提升“坚强智能电网”运维效率的核心环节，其核心需求在于基于监测数据实现缺陷类型识别、等级判定与处理建议生成（如隔离开关接触不良、变压器过热、避雷器性能异常等场景），这直接关系缺陷处理的时效性与准确性，对降低电网故障风

险至关重要。传统缺陷诊断方式存在显著局限：人工排查依赖运维人员经验，不仅效率低下且易受主观因素影响产生误判，难以适配国家电网设备数量激增后的运维需求；单一模态诊断技术更存在功能短板——仅依托红外图像虽可获取温度数据，却难以准确区分设备类别（如电流互感器与电压互感器），导致缺陷归属模糊；仅依赖可见光图像虽能识别设备类型，却无法捕捉设备内部热状态，难以发现隐性热缺陷。红外与可见光融合系统通过分层协同处理机制实现智能化诊断：针对可见光图像，采用轻量级特征提取与多尺度检测技术，实现对多类型电力设备的精准识别；针对红外图像，通过增强热特征提取能力与边界框优化算法，精准捕捉设备温度分布；再借助决策级融合技术整合双模态检测结果，结合《带电设备红外诊断应用规范》等行业标准自动判定缺陷等级（紧急、严重、一般缺陷）。同时，系统配备图形化操作界面，降低运维人员操作门槛，无需专业编程技能即可实现双光谱监测，自动输出缺陷诊断结果与针对性处理建议（如“隔离开关接触不良，建议测量接触电阻”）。这种“数据驱动”的诊断模式彻底革新传统人工排查方式，大幅提升国家电网缺陷处理效率与准确性，为“坚强智能电网”运维智能化提供核心技术支持。

电缆接头作为国家电网地下电缆系统的故障高发环节，据统计其故障占比达电缆总故障的 44.88%，且多处于潮湿、封闭、低光照的电缆井环境中，长期受环境侵蚀易出现内部发热（如接触电阻增大导致局部过热）与机械损伤（如安装扭转破损），若未及时预警易引发短路甚至火灾，因此电缆接头故障监测预警是保障“坚强智能电网”地下输电通道安全的重点场景。传统监测手段难以适配电缆井的特殊环境：单一红外测温技术因电缆井空间狭窄导致测量距离波动，易产生温度误差，且无法表征接头机械损伤细节；单一可见光技术受井内黑暗环境限制，无法捕捉接头内部热信号，即便辅以补光，潮湿空气仍会导致画面模糊，难以全面掌握接头真实状态。红外与可见光融合系统针对电缆井场景特性进行定制化设计：硬件选型上采用小型化设备，适配井口狭窄的安装空间；针对红外图像，通过对比度增强与去噪处理抑制潮湿环境干扰，清晰呈现接头热分布；结合可见光图像捕捉接头外观细节（如机械扭转痕迹、绝缘层破损），再通过融合算法生成兼具热分布与外观特征的融合图像。为进一步提升测温精度，系统采用多项式拟合建立温度修正函数，将电缆接头测温误差从  $\pm 5^{\circ}\text{C}$  降至  $\pm 1.5^{\circ}\text{C}$ 。当接头温度超过预设阈值时，系统通过图像二值化、膨胀处理自动定位异常发热区域，同步触发生光报警并联动井盖锁闭装置，防止风险扩大。该系统实现对电缆接头“热状态 + 机械状态”的全方位监测，有效避免地下电缆故障扩大化，为国家电网“坚强智能电网”地下输电通道的安全运行提供坚实保障。

野外输电线路作为国家电网远距离电力传输的核心载体，其沿线杆塔长期暴露于野外环境，面临人员活动（如采药、放牧人员靠近，非法人员盗割线路）与动物活动（如牛、羊碰撞杆塔，鸟类、松鼠在绝缘子筑巢）的双重风险，这些行为易引发线路触电、杆塔倾斜、绝缘子短路等故障，直接威胁电力传输稳定性，因此野外输电线路沿线杆塔监测是“坚强智能电网”保障远距离传输安全的重要环节。传统监测技术难以适应野外复杂环境：单一红外技术虽可突破光照限制捕捉活体目标热辐射，但无法区分目标类型（如人员与动物）及人员携带工具情况，易引发误报警；单一可见光技术受野外光照变化（如黎明、黄昏）与恶劣天气（雨雪、雾霾）影响，画面清晰度波动大，难以识别动物筑巢、人员携带盗割工具等关键细节，无法精准判定风险等级。红外与可见光融合系统采用户外型耐候设备，适配野外复杂环境：通过红外传感器持续捕捉活体目标热辐射，不受光照与天气限制，即便在茂密树林遮挡下仍可定位隐藏目标；借助可见光传感器清晰采集杆塔细节信息，实现目标类型区分（人员 / 动物）、人员携带工具识别及绝缘子筑巢判断；系统预设杆塔周边警戒范围（如拉线 3 米内、绝缘子下方 2 米内），一旦识别目标进入警戒区，立即将包含目标类型、位置及实时融合图像的报警信息推送至运维终端。这种监测方式实现对野外杆塔的全天候、高精度监控，助力运维人员快速排查风险，大幅降低输电线路故障发生率，保障国家电网远距离电力传输的稳定可靠，契合“坚强智能电网”对远距离传输管控的需求。

变电站周界及高压设备区作为国家电网人员触电与设备碰撞风险的高发区域，无关人员（如周边居民误闯、施工人员违规靠近）或流浪动物（如猫、狗）进入易引发人员伤亡、设备跳闸等安全事故，因此该区域的安防管控是“坚强智能电网”保障人员与设备安全的基础防线。传统安防手段存在显著短板：单一红外安防系统虽可识别活体目标，但无法区分人员与动物，易将流浪动物误判为人员引发误报警；单一可见光安防系统在夜间或低光环境下失效，且受设备阴影、树木遮挡影响，难以精准识别目标是否进入危险区域。红外与可见光融合系统依托双传感器协同架构，通过红外传感器捕捉目标热辐射信号，在夜间无补光或雾天环境下仍可精准标记活体目标位置，规避光照与遮挡干扰；借助可见光传感器同步采集现场细节，实现人员与动物的清晰区分，及人员着装（如工装识别）、携带物品的判断，减少误判；系统实时处理融合数据，一旦识别目标进入预设警戒区域（如主变压器 5 米范围），立即触发现场声光报警，同步将融合图像与报警信息推送至运维终端，并联动门禁、围栏等安防装置限制目标移动。这种安防模式实现对变电站关键区域的全天候精准监控，有效预防人员触电与设备碰撞事故，为国家电网变电站的安全运行筑牢防线，支撑“坚强智能电网”的安全管

控目标。

电力仓库作为国家电网运维物资存储的核心场所，其内存放的电缆、绝缘材料、互感器等关键物资若遭遇非法盗窃或野生动物（如野兔、黄鼠狼）啃咬绝缘部件，不仅会影响电网运维工作开展，更可能因物资性能受损引发后续设备故障，因此电力仓库及物资存放区监控是“坚强智能电网”保障运维物资安全的重要支撑。传统监控方法存在局限：单一红外监控虽可在夜间识别活体目标，但无法区分人员与动物及人员携带搬运工具情况，难以精准识别盗窃行为；单一可见光监控在夜间需依赖补光设备，不仅能耗高且易产生光污染，且黑暗环境下画面清晰度差，无法及时发现动物钻缝进入仓库的行为。红外与可见光融合系统结合电力仓库场景需求进行针对性设计：通过红外传感器捕捉仓库周边及出入口的活体热信号，即便夜间仓库熄灯仍可精准识别“人员翻越围墙”“动物钻缝进入”等异常行为；借助可见光传感器采集目标细节，实现人员携带搬运工具判断与破坏性啮齿类动物识别；同时对双模态图像进行空间配准处理，确保目标位置对齐，避免图像偏移导致的误判。当系统检测到异常时，立即触发仓库声光报警，联动门禁系统锁闭出入口防止目标逃离或进一步破坏，并留存融合图像作为后续溯源证据。这种监控方式实现对电力仓库的全天候、低能耗安全防护，有效保障运维物资安全，为国家电网“坚强智能电网”运维工作的顺利开展提供物资保障。

综上，红外与可见光融合系统通过在国家电网“坚强智能电网”建设中覆盖变电站设备监测、电力设备缺陷诊断、电缆接头预警、野外杆塔监测、变电站安防及电力仓库管控等多核心场景，形成从设备状态感知到风险防控的全流程技术支撑。该系统不仅弥补传统单一模态技术的不足，提升电网运维的智能化水平与安全可靠，更助力国家电网突破传统技术瓶颈，为“坚强智能电网”的战略落地与持续优化提供关键技术保障，成为国家电网推进能源转型与电网智能化升级的重要技术载体

## 2 国内外技术研究现状

在电力系统智能化转型进程中，红外与可见光融合系统通过整合红外图像的热辐射信息（可抗光照不足、雾雪等恶劣环境，精准突出设备热目标）与可见光图像的高分辨率纹理细节（完整保留设备结构特征，更符合人眼视觉习惯），为智能电网建设提供了关键技术支撑。这种跨模态信息融合机制，有效突破了单一图像源在电力设备状态监测、故障诊断及运行环境感知中的局限性，成为构建高可靠性、强自愈能力智能电网的重要基础。从技术发展脉络来看，红外与可见光

融合算法可分为传统红外与可见光融合算法、基于深度学习的红外与可见光融合算法两大类，国内外研究围绕“提升融合精度、适配电力工程场景”展开，形成了各具特色的技术路径，以下结合技术原理与研究进展展开系统阐述。

## 2.1 传统红外与可见光融合算法

传统红外与可见光融合算法诞生于 20 世纪 80 年代，核心逻辑遵循“图像分解 - 特征 / 子带融合 - 逆变换重建”，依托手工设计的数学变换与融合规则实现信息整合，无需大规模标注数据，在早期算力有限的电力场景（如传统变电站人工巡检辅助）中广泛应用，主要包括基于多尺度变换、基于稀疏表示、基于显著性检测及基于混合模型四类方法。

基于多尺度变换的融合方法（Multi-Scale Transformation, MST）是传统融合领域应用最广泛的技术分支，其核心思想是利用数学变换将红外与可见光图像分解为不同尺度、不同频率的子带图像——低频子带主要表征图像全局轮廓与整体亮度，高频子带则表征边缘、纹理、热梯度等局部细节，针对不同子带的特征特性设计差异化融合规则，再通过逆变换将融合子带重构为完整融合图像，以此平衡全局信息与局部细节的保留效果。国外是该方法的发源地，早期以金字塔类变换为核心，拉普拉斯金字塔（Laplacian Pyramid, LP）首次通过迭代低通滤波与下采样构建多层金字塔，实现图像多尺度分层融合，但存在“块效应”，子带融合后设备边缘过渡不自然；低通金字塔（Ratio Pyramid, RP）通过比率运算优化低频子带融合权重，在夜间电力设备监测中提升了背景与热目标的协调性，却难以充分保留细节信息。20 世纪 90 年代后，小波变换（Wavelet Transformation, WT）逐步取代金字塔变换，其中离散小波变换（Discrete Wavelet Transform, DWT）通过多尺度、多方向（水平、垂直、对角三个方向）分析解决了块效应问题，且具备更优的信噪比，在变压器套管热异常监测中表现出良好的细节保留能力，但二维小波基仅支持三个方向分解，难以完整捕捉电力设备复杂边缘（如电缆接头多曲率轮廓）。国内学者围绕电力场景需求展开针对性改进，国内学者师颖在 2018 年的研究中提出两种优化算法：一种是“基于 IHS + 小波变换”，先对可见光图像进行 IHS 色彩空间变换（IHS Transformation）提取亮度分量，再将亮度分量与红外图像进行小波分解，低频子带采用“区域平均梯度取大”规则以保留变压器全局轮廓清晰度，高频子带采用“区域能量”规则突出套管热异常边缘，实验表明该算法较单一小波算法提升图像清晰度 12%；另一种是“基于 IHS+Contourlet 变换”，用 Contourlet 变换替换小波变换，利用其支持 16 个方向分解的多方向特性，更好捕捉电力设备边缘纹理，在实验中该算法信息熵达

8.476、平均梯度达 6.279，均优于其他传统算法，成为当时变电站设备监测的优选方法。国内学者赵庆典在 2024 年的研究中进一步优化多尺度分解策略，提出“图像增强 + 二次 NSCT 分解”方法：先通过引导滤波增强可见光图像的电缆接头纹理细节，再对红外图像与增强后可见光图像进行首次非下采样轮廓波变换（Non-Subsampled Contourlet Transform, NSCT）分解融合，将融合系数二次分解为高频与低频子带后，与可见光图像高频子带再次融合，最终逆变换生成融合图像，该方法在电缆接头图像融合中边缘保持度提升 15% 以上；同时，他还提出“滚动引导滤波多尺度分解”，将图像通过滚动引导滤波分解为基础层（全局轮廓）与细节层（局部纹理），结合显著图加权融合，解决传统变换在电力设备监测中对比度低、边缘模糊的问题，使变压器过热区域识别率提升 20%。这类方法的优势在于可通过分层处理保留图像信息，能通过规则调整适配电力设备热目标与结构细节并重的需求，但手工设计的融合规则存在主观性，如小波高频“取大”规则易丢失可见光纹理，且分解层数过高时易出现伪影，在强电磁干扰下的变电站等复杂场景中适应性较差。

基于稀疏表示的融合方法（Sparse Representation, SR）依托“信号稀疏性原理”实现图像表征与融合，其核心逻辑是自然图像可通过少量“原子”（基础特征向量）的线性组合表示，这些原子构成包含图像所有可能模式的过完备字典；将红外与可见光图像通过优化算法稀疏分解为“少数非零系数”（仅表征红外热目标、可见光边缘等关键特征），按设计规则融合稀疏系数后，再利用过完备字典重构融合图像。该方法通过选择性保留稀疏系数，可有效抑制噪声干扰，在电缆井潮湿噪声环境的电力设备监测中表现优异。过完备字典的构建是该方法的核心，国外早期主要依赖数学变换，基于奇异值分解（Singular Value Decomposition, SVD）的字典通过分解电力设备标准图像矩阵提取原子，适用于变电站墙壁背景等纹理简单的场景；基于主成分分析（Principal Component Analysis, PCA）的字典通过提取数据主成分优化稀疏性，在变压器轮廓表征中表现较好，但对绝缘子污秽等复杂纹理适配性不足。国内学者围绕电力场景适配性展开优化，国内学者刘雯雯提出“卷积稀疏表示”，通过局部卷积约束使字典原子更精准捕捉电缆接头的纹理与热梯度特征，稀疏系数的非零占比降低 30%，同时保留关键热信息；国内学者陈艳菲引入注意力机制优化字典学习，对红外热目标区域的稀疏系数赋予更高权重，在变压器过热监测中使目标定位误差缩小至  $\pm 2$  像素。融合规则设计方面，早期多采用“加权平均”“阈值筛选”等简单策略，赵庆典提出“基于区域能量的稀疏系数融合”，计算稀疏系数对应的区域能量（能量越高表征特征越关键），对高能量系数优先保留，在电缆接头图像融合中信息熵提升 0.3~0.5。



这类方法的优势在于噪声抑制能力强，可突出红外热目标，但字典构建计算成本高，如 SVD 方法处理  $512 \times 512$  图像需耗时 10~15 秒，且在罕见电力设备故障等小样本场景下字典泛化性差，难以满足实时监测需求。

基于显著性的融合方法 (Saliency-Based Fusion, SBF) 以 “人类视觉注意力机制” 为核心，其核心逻辑是图像中不同区域的信息价值存在差异，通过显著性检测算法识别源图像中最具信息价值的区域 (如红外热目标、可见光关键纹理)，生成量化区域信息价值的显著性图，将显著性图作为权重因子指导融合过程，确保显著性区域在融合图像中突出显示，同时平衡红外与可见光的信息贡献，避免单一模态信息过度支配，在电力设备热异常定位等目标导向型融合中表现突出。显著性检测是该方法的基础，国外早期通过频域分析 (如傅里叶变换) 或区域对比度计算生成显著性图，但对夜间变电站等低对比度场景适应性差，易出现 “伪显著性区域”。国内学者结合电力场景优化检测算法，赵庆典提出 “频率调谐滤波 + 显著图生成”，先通过频率调谐增强红外热目标与背景的对比度，再计算区域梯度生成显著性图，在电缆接头过热监测中显著性区域识别准确率达 90%；师颖将显著性检测与多尺度变换结合，在 IHS+Contourlet 融合过程中，用显著性图调整高频子带融合权重，使变压器套管热异常区域在融合图像中更突出，视觉辨识度提升 30%。融合权重生成方面，传统方法多采用 “显著性值直接映射”，易导致权重失衡，赵庆典提出 “自适应权重调整”，对显著性图进行归一化处理，将红外热目标等高显著性区域权重设为 0.6~0.8，背景等低显著性区域权重设为 0.2~0.4，在杆塔鸟巢监测中平衡热目标与结构细节，避免鸟巢区域被背景掩盖。这类方法的优势在于符合人眼视觉习惯，目标突出性强，但显著性检测易受噪声干扰，在强电磁环境下显著性图易出现伪目标，且对变电站屋顶等无明显目标的场景融合效果不佳。

基于混合模型的融合方法 (Hybrid Model-Based Fusion, HMBF) 旨在突破单一传统方法的局限性，其核心逻辑是结合多尺度变换、稀疏表示、显著性检测等多种传统技术的优势，通过 “特征增强 - 多模态分解 - 协同融合” 的流程，同时优化红外热目标保留与可见光细节保留，适配多设备重叠、强噪声干扰等复杂电力场景。国外早期混合模型以 “目标增强 + 多尺度变换” 为核心，有学者提出先通过阈值分割增强红外热目标 (如电缆接头过热区域)，再采用拉普拉斯金字塔分层融合，在电力设备监测中目标识别率提升 25%，但存在目标边缘过渡不自然的问题。国内学者针对电力场景优化混合策略，师颖提出 “图像增强 + IHS+Contourlet” 混合模型，先对可见光图像进行三段线性变换增强以提升变压器纹理清晰度，再通过 IHS 变换提取亮度分量，与红外图像进行

Contourlet 分解融合，在变电站设备融合中信息熵达 8.476，较单一多尺度变换提升 0.5~0.8；赵庆典提出“引导滤波增强 + 滚动引导滤波多尺度分解 + 显著图融合”混合模型，先增强可见光细节，再将图像分解为基础层与细节层，结合显著图加权融合，在电缆接头故障监测中边缘保持度达 0.5978，较传统混合模型提升 10%~15%。这类方法的优势在于通过多技术协同提升融合质量，可适配复杂电力场景，但模型逻辑复杂，需多步骤参数调试，且泛化性差，某一环节参数调整需重新优化整体流程，工程落地成本较高。

综上，传统融合方法虽在特定电力场景中具备应用价值，但普遍依赖手工设计的变换与规则，难以充分捕捉红外与可见光图像的复杂特征（如电力设备的非均匀热分布与复杂纹理），且复杂场景适应性、实时性不足，这些局限为基于深度学习的融合方法发展提供了需求导向。

## 2.2 基于深度学习的红外与可见光融合算法

随着算力提升与电力场景标注数据（如变电站设备故障数据集、电缆接头热异常数据集）的积累，基于深度学习的红外与可见光融合算法在 2016 年后成为研究主流。这类算法以“数据驱动”为核心，通过神经网络自动学习红外与可见光的深层特征及融合策略，无需手工设计规则，在特征提取精度、复杂场景适配性上显著超越传统方法，根据网络架构差异可分为基于卷积神经网络、基于生成对抗网络、基于扩散模型、基于自编码器及基于 Transformer 的五类方法，国内学者围绕电力设备监测的“实时性、测温精度、目标识别适配性”展开大量工程化优化研究。

基于卷积神经网络的融合方法（Convolutional Neural Network, CNN）是深度学习融合领域的基础架构，其核心逻辑是通过“端到端特征提取 - 融合 - 重建”实现融合：卷积层凭借局部感受野高效提取红外与可见光的局部特征（如设备边缘、热梯度），池化层实现特征降维并保留关键特征，融合层通过注意力机制或加权策略整合特征，最后由反卷积层将融合特征重构为融合图像。该方法避免了手工规则设计，在电力设备细节保留方面表现突出。国外早期 CNN 融合方法以“通用架构适配”为核心，有学者提出通用端到端融合网络（Image Fusion CNN, IFCNN），通过结构化卷积层隐式处理特征提取与融合，在含电力设备样本的 TNO 数据集上信息熵突破 7.0，较传统算法提升 15%~20%；另有学者提出密集连接卷积融合网络（DenseFuse），通过密集连接（每层特征与所有前层特征连接）强化特征复用，在变压器套管图像融合中边缘细节保留率提升 25%。

国内学者针对电力场景优化网络结构，国内学者姚新雨提出“跨尺度注意力 CNN”，在网络中嵌入空间注意力模块（突出电缆接头热目标）与通道注意力模块（强化绝缘子纹理特征），在电缆接头故障监测中使测温误差缩小至  $\pm 1.5^{\circ}\text{C}$ ；国内学者王传洋将 CNN 融合与设备识别结合，在融合网络后添加分类分支，实现“变压器缺陷融合 - 定位 - 识别”一体化，缺陷诊断准确率达 91%，较“单独融合 + 单独分类”的两阶段方法提升 10%。为平衡精度与实时性，有学者通过网络架构搜索（Neural Architecture Search, NAS）自动生成适配嵌入式设备的轻量化 CNN，移除冗余卷积层并采用  $1 \times 1$  卷积核降维，在电力巡检机器人上帧率达 25fps，满足实时监测需求；师颖的后续研究还将 CNN 与传统 Contourlet 变换结合，用 CNN 优化高频子带融合规则，在变电站设备融合中信息熵与实时性均优于单一方法。这类方法的优势在于局部特征提取精度高、工程适配性强，但局限性在于卷积层感受野有限，难以捕捉图像长距离依赖关系（如杆塔与鸟巢的全局关联），在复杂背景下易丢失全局结构信息。

基于生成对抗网络的融合方法（Generative Adversarial Network, GAN）依托“博弈论”构建对抗机制，其核心逻辑是通过生成器与判别器的迭代对抗实现融合：生成器通常采用编码器 - 解码器结构（如 U-Net），提取红外与可见光的深层特征并整合重构为融合图像；判别器采用多层卷积分类网络，区分融合图像与源图像（红外 / 可见光），为生成器提供优化反馈，最终使生成器生成兼具热辐射信息与纹理细节的融合图像。该方法在模态信息平衡、复杂场景鲁棒性上表现优异，成为电力设备复杂环境监测的主流方法。国外早期 GAN 融合以“单一判别器”为核心，有学者提出首个用于红外与可见光融合的 GAN 模型（FusionGAN），生成器采用 U-Net 架构，判别器区分融合图像与可见光图像，在夜间变电站监测中融合图像纹理细节较传统方法提升 40%，但存在红外热信息丢失问题；为解决“模态失衡”，有学者提出双判别器条件生成对抗网络（Dual-Discriminator Conditional GAN, DDcGAN），引入红外判别器（Dis-IR）约束融合图像热辐射信息、可见光判别器（Dis-VIS）约束纹理信息，在变压器过热监测中热信息保留率提升 30%。国内学者针对电力场景优化 GAN 结构，国内学者宋宝平提出 STAFusion，将 Swin Transformer 嵌入 GAN 生成器，通过移动窗口机制捕捉全局依赖，同时用双判别器平衡模态，在 TNO 电力设备数据集上峰值信噪比（PSNR）达 63.8317、结构相似性（SSIM）达 0.9498，较传统 GAN 提升 15%~20%；国内学者梁洪源提出“温度约束 GAN”，在生成器损失函数中加入温度误差项（融合图像测温值与实际温度的均方误差），通过反向传播优化生成器，在电缆接头故障监测中将测温误差从  $\pm 5^{\circ}\text{C}$  降至  $\pm 1.2^{\circ}\text{C}$ ，故障

预警准确率提升至 95%。为提升抗噪声能力，有学者在 GAN 中引入注意力机制，对红外热目标区域的特征进行强化，在强电磁干扰的变电站场景中融合图像信噪比提升 20%；国内学者郭美青还将 GAN 与多源数据融合结合，输入红外、可见光及局部放电数据，生成“热 - 纹理 - 电气特征”融合图像，在变压器缺陷诊断中误判率下降 35%。这类方法的优势在于模态信息平衡、复杂场景鲁棒性强，但局限性在于训练不稳定，易出现模式崩溃，且计算成本高，GAN 训练需消耗传统 CNN 2~3 倍的算力。

基于扩散模型的融合方法（Diffusion Model, DM）是近年来新兴的生成式融合技术，其核心逻辑是通过“逐步去噪”过程生成融合图像：依托去噪扩散概率模型（Denoising Diffusion Probabilistic Model, DDPM）构建马尔可夫链，从随机噪声出发，迭代去除噪声并恢复图像细节，最终生成融合图像，该方法有效解决了 GAN 训练不稳定、模式崩溃的问题，在图像生成多样性、细节保留上表现突出，为电力遥感图像融合提供了新路径。DDPM 作为 DM 的核心组件，于 2020 年被提出，通过迭代去噪生成高质量图像，在电力线路全景融合中表现出优异的稳定性。国内学者将 DM 适配电力场景，有学者将三通道可见光与单通道红外图像合并为四通道输入，以该四通道图像为地面真值训练自监督扩散模型，在输电线路遥感融合中杆塔与导线的细节保留率提升 25%；另有学者将电力设备融合转化为“条件后验抽样”，通过似然校正约束生成图像与源图像的相似度，在变电站全景融合中全局结构一致性提升 30%。国内学者黎瑞虹进一步优化 DM 的实时性，提出“轻量化 DDPM”，通过将扩散步数从 1000 步降至 200 步，同时采用深度可分离卷积简化去噪网络，在遥感电力图像融合中帧率提升至 10fps，虽低于 CNN 但显著优于传统 GAN；同时，他还将 DM 与 Transformer 结合，用 Transformer 捕捉遥感图像的长距离依赖，在跨塔线路融合中全局关联信息保留率提升 20%。这类方法的优势在于生成稳定性高、细节丰富，但局限性在于计算成本高，完整 DDPM 训练需 GPU 显存不低于 16GB，实时性不足，暂未大规模应用于实时电力巡检。

基于自编码器的融合方法（Autoencoder, AE）是一种无监督学习网络，其核心逻辑是通过编码器（提取特征）与解码器（重构图像）实现融合：先在自然图像（如 KAIST 数据集）上预训练 AE，使编码器具备特征提取能力，再通过编码器分别提取红外与可见光的特征向量，按规则融合特征后由解码器重构融合图像。该方法无需大规模标注数据，在罕见电力设备故障等小样本场景中表现出优势。国外早期 AE 融合以“单尺度编解码”为核心，有学者首次将 AE 用于红外与可见光融合，通过残差连接避免梯度消失，在变压器图像融合中信息熵达

7.8; 后续又有学者提出嵌套连接自编码器 (RFN-Nest), 引入嵌套连接与注意力机制, 在电缆接头图像融合中边缘保持度达 0.58, 较单尺度 AE 提升 15%。国内学者针对电力小样本场景优化 AE, 国内学者 Ren 等人在 AE 中加入残差模块, 构建 “残差 - 卷积 AE”, 在仅含 50 组故障数据的电缆接头融合中泛化性提升 30%; 国内学者 Han 等人提出 “分类显著性 AE”, 通过设备缺陷分类任务引导 AE 学习关键特征, 在变压器缺陷融合中缺陷区域特征权重提升 40%, 诊断准确率达 88%。为提升特征融合的合理性, 有学者提出 “解离化表征 AE”, 通过解耦网络分离红外热特征与可见光纹理特征, 避免模态干扰, 在变电站周界安防中误报率下降 30%; 国内学者孙浩还将 AE 与传统稀疏表示结合, 用 AE 学习过完备字典, 稀疏表示优化融合系数, 在电力设备噪声图像融合中信噪比提升 20%。这类方法的优势在于无监督学习、小样本适配性强, 但局限性在于特征融合规则仍需手工调整 (如注意力权重), 深层特征的可解释性差, 难以追溯融合逻辑。

基于 Transformer 的融合方法依托 “自注意力机制” 捕捉图像全局依赖关系, 其核心逻辑是通过多头自注意力模块学习红外与可见光图像的长距离关联 (如电力杆塔与周围环境的全局结构), 突破 CNN 感受野限制, 在复杂场景融合中表现出巨大潜力, 是近年来的研究热点。国外早期 Transformer 融合以 “全局自注意力” 为核心, 有学者首次将 Swin Transformer 用于图像融合, 通过层次化特征表示突出全局语义信息, 在输电线路全景融合中全局结构一致性提升 35%; 另有学者提出 TGFuse, 结合 Transformer 与 GAN, 用 Transformer 捕捉全局关系, GAN 平衡模态信息, 在野外杆塔监测中目标识别率提升 40%。国内学者针对电力场景优化 Transformer 结构, 宋宝平提出 STAFusion, 采用 Swin Transformer 的移动窗口机制, 将自注意力计算限制在  $16 \times 16$  等局部窗口, 计算复杂度较传统 Transformer 下降 40%, 同时通过移位操作实现跨窗口信息交互, 在变电站设备融合中 SSIM 达 0.9498; 他后续还提出 KFSTFusion, 新增关键特征提取模块 (KFEM), 通过通道 - 空间注意力突出电力设备边缘与热目标, 再由 Transformer 模块 (STM) 构建全局关系, 在电缆接头图像融合中边缘保持度达 0.5812, 较 STAFusion 提升 5%。为平衡全局与局部特征, 有学者设计 “局部 - 全局双分支 Transformer”, 局部分支 (CNN) 保留绝缘子纹理等设备细节, 全局分支 (Transformer) 捕捉杆塔与鸟巢等长距离依赖, 在电力线路融合中细节与全局的协调性提升 30%; 国内学者张杰将 Transformer 与 CNN 结合, 用 CNN 提取局部特征, Transformer 整合全局信息, 在变电站全景融合中实时性与精度均优于单一 Transformer。这类方法的优势在于全局建模能力强、

复杂场景适配性好，但局限性在于计算成本高，处理  $1024 \times 1024$  图像需 GPU 显存不低于 24GB，且在小样本场景下泛化性差，需大量电力标注数据支撑。

综上，基于深度学习的融合算法虽在精度与鲁棒性上超越传统方法，但仍面临挑战：实时性与精度的平衡需进一步优化，如通过模型剪枝、量化实现 Transformer 的轻量化；电力场景小样本泛化性待提升，需结合半监督、无监督学习策略；下游任务适配性需加强，现有方法多关注视觉效果，需强化与缺陷诊断、目标跟踪等电力下游任务的协同。未来研究可结合电力设备的物理特性（如电缆接头热分布模型），构建“物理约束 - 数据驱动”融合模型，推动算法从“视觉优化”向“工程实用”深化，更好适配智能电网建设需求。

### 2.3 融合系统在国家电网公司的应用及实践案例

在国家电网公司推进“坚强智能电网”与“新型电力系统”建设的进程中，红外与可见光融合系统凭借“热状态监测 + 结构细节识别”的双重优势，已深度融入电力设备全生命周期管理，覆盖变电站设备运维、输电线路巡检、电缆故障预警等核心场景。国家电网公司及下属单位通过技术研发、专利转化与工程落地，形成了一系列适配电力场景的融合应用方案，有效提升了设备缺陷诊断准确率与运维效率。

在特高压场景中，红外与可见光融合系统聚焦复杂环境下的精准巡检与风险防控。2025 年 3 月，国网山东电力在 1000 千伏曹湖线、湖高线应用双光融合巡检系统，通过无人机搭载高清可见光与红外热成像双光谱设备，以“可见光查销钉级缺陷、红外光测设备温度”的双光云台设计，精准捕捉金具磨损、绝缘子破损等机械损伤及线夹发热等隐蔽缺陷，巡检效率较人工提升 4 倍，成功应对湖区段高湿度、多鸟类活动的复杂环境；该技术通过“问题发现 - 技术研发 - 成果应用”闭环机制，已在山东电网推广，并计划向更多复杂环境特高压线路扩展。2024 年，国网浙江湖州供电公司针对  $\pm 800$  千伏金塘线覆冰风险，研发“红外热成像 + 可见光”双光监测装置，依托人工智能透雾技术，在低能见度天气下视物距离达 20 米，解决覆冰监测难题，其专项技术团队攻关模式形成可复制经验，为特高压线路精准运维提供支撑。

超高压场景则侧重智能巡检模式创新与算法优化。2025 年 4 月，国网湖北直流公司在  $\pm 500$  千伏林枫换流站部署“空地协同智能巡检系统”，通过 WPAI 无线网络实现红外与可见光数据实时回传、同步分析，无人机按固定航线自动红外测温，突破人工巡检视角限制，还能自动生成报告，其“标准化航线 + 实时

分析 + 人机互补”模式，为超高压换流站巡检提供可借鉴方案。2025 年 3 月，国网浙江电科院研发的“红外大脑”智能识别系统上线，基于“最小化标注 + 阶梯式学习 + 干扰点屏蔽”技术路线，在复杂环境中红外缺陷检出率超 81%、误检率低于 1.7%，单日 12000 张图像 1 小时即可完成分析，已在浙江超高压线路发现 37 起缺陷；该系统经小型化形成边缘算力装置，后续将构建“云边协同”体系，实现“边飞边识别”，其算法实战优化与全流程智能化思路，为超高压线路高效诊断提供技术范式。

配电网场景聚焦轻量化、高效化的基层运维需求。2025 年 5 月，国网福建电力上线配网红外图像快速筛查模块，通过背景过滤、双光融合、温度自动解析技术，30 分钟可完成 1000 余张红外图筛查，单张处理时间低于 2 秒，日处理量达人工 60 倍，部署于配电机巡平台后推动“人工判图”向“AI 诊断”转型，其专用模块开发、人机协同与闭环管理经验，适配配网设备分散的特点。2023 年 9 月，国网宁波供电公司在慈城镇试点杆上双目红外测温仪，集成热成像与可见光芯片，实现 10 千伏变压器远程测温、告警及环境观测，计划在重点变压器推广，其“重点设备试点 + 兼容性设计 + 厂商合作”策略，平衡配网运维成本与效率。

变电站场景以立体巡检体系构建为核心。2025 年夏季，国网淮安供电公司打造无人机、机器人、视频监控“三剑客”体系，四足机器人搭载双光摄像头，在极端天气下替代人工红外测温，设备巡检替代率达 95%、效率提升 65%，无人机年均完成 3.7 万基铁塔巡检，视频系统发现 124 条缺陷；该公司通过“数据中枢 + 空中兵团”协同，还开发树障预警平台，精度达 92.3%，其多装备协同与数据整合思路，优化变电站运维流程。2024 年 7 月，国网清河县供电公司应用变电站红外巡检系统，双光图像就地融合实现“缺陷识别 + 温度测量”同步，高压线路巡视时间缩至人工 1/4、缺陷识别率 100%，通过规范巡检机制、加强人员培训与厂家合作，确保技术常态化应用。

输电线路场景侧重特殊环境适应性与无人化技术。2025 年 5 月，国网唐山供电公司在 110 千伏甄北隧道应用四足机器人，具备全地形自主导航能力，搭载双光设备与气体检测模块，2 小时完成人工 4 小时的巡检量，还捕捉到 0.3 毫米电缆绝缘缺陷，依托高压电缆智能管控平台构建“立体巡检 + 集中监控”模式，提升隐患处置速度。2024 年 12 月，国网新疆电力在沙漠地段改造无人机，全绝缘处理后搭载双光设备，实现绝缘子带电冲洗与异常发热监测，解决沙漠巡检难题，计划向其他特殊环境推广，其“环境适配改造 + 人机协同 + 安全管控”经验，为极端地区线路运维提供参考。

电缆场景则探索智慧化与多技术融合应用。2024 年 7 月，国网河北电力在雄东电缆隧道应用 “5G + 数字孪生” 系统，机器人红外测温数据实时同步至数字孪生平台，巡检时长从 48 小时缩至 1 小时、效率提升 48 倍，缺陷识别准确率 80%，通过 5G 支撑与 AI 摄像头试点，推动隧道可视化巡检，技术已在河北推广。2025 年 5 月，国网冀北电力在甄北隧道应用同款四足机器人，同样实现微米级缺陷检测与巡检效率翻倍，依托智能管控平台构建协同运检模式，其针对隧道环境的机器人配置、多技术融合与数据分析机制，为电缆隧道无人化运维提供成熟方案。

### 3 研究内容及方法

本研究针对传统红外与可见光融合算法依赖手工规则设计、复杂环境下适配性差的核心痛点，以提升电力设备监测精度与运维效率为目标，聚焦基于生成对抗网络（GAN）的融合算法研发，确立 “损失函数优化 - 网络架构设计 - 场景适应性迭代” 的技术路线，旨在实现红外图像热辐射信息与可见光图像纹理细节的高效协同融合，同时结合国家电网变电站、特高压线路、电缆隧道等核心运维场景的特性开展定制化优化，确保算法在电力复杂场景中具备高鲁棒性与实用性，为电网智能化运维提供技术支撑。

在算法核心机制研究中，首要突破 “热信息精准保留与结构细节清晰捕捉” 的平衡难题。其一，设计双目标导向的损失函数体系：生成器损失由对抗损失与内容损失构成，对抗损失通过生成器与判别器的动态博弈，驱动融合图像在纹理特征上逼近可见光图像（如变压器瓷套纹理、杆塔结构细节），确保结构缺陷可识别；内容损失采用弗罗贝尼乌斯范数约束，使融合图像与红外图像保持像素强度一致性，精准保留设备热辐射特征（如电缆接头过热区域、绝缘子温升异常），同时引入梯度算子约束匹配可见光图像的结构梯度，避免纹理失真，并通过可调节权重参数平衡热信息与结构信息的贡献占比，防止某类信息被压制。其二，优化网络架构以保障融合效率与精度：生成器采用五层卷积结构，全程避免下采样操作，减少图像细节丢失，前四层引入批归一化层与 Leaky ReLU 激活函数，有效缓解梯度消失问题，最后一层采用 tanh 激活函数输出归一化的融合图像；判别器同样为五层卷积架构，通过步长为 2 的卷积层替代传统池化层实现特征降维，在降低计算复杂度的同时保障分类精度，最终实现算法端到端运行，无需人工干预设计融合规则，大幅提升适配性。

从应用潜力与价值来看，该算法可作为核心技术基础，支撑国家电网多类核



心项目的智能化升级：在变电站设备运维场景中，其优化后的“噪声抑制 - 双信息融合”能力，可输出兼具设备热异常（如变压器铁芯热点、断路器触头过热）与结构缺陷（如瓷套污秽、操作机构变形）的融合图像，为后续设备同步诊断系统开发提供高质量数据输入，解决传统人工两次作业的效率痛点；在特高压线路巡检场景中，轻量化后的算法可适配无人机边缘终端，凭借全天候环境适应性（突破光照、雾雪限制），为线路金具磨损、绝缘子异常、导线过热等缺陷的自动化识别提供算法支撑，助力巡检模式从“人工主导”向“智能自主”转型；在电缆接头预警场景中，精准的温度修正与图像配准技术，可捕捉接触不良导致的微小温升趋势，为地下电缆故障早期预警系统搭建提供关键的温度数据与图像对齐能力，降低故障扩大风险；在配网运维场景中，微型化算法可部署于台区小型监测终端，支持设备图像批量处理，为配网台区快速筛查提供高效算法工具，缓解运维人员现场排查压力。

### 3.1 研究方法概述

本研究采用一种基于生成对抗网络的新型红外与可见光图像融合方法，该方法能够同时保留红外图像中的热辐射信息和可见光图像中的纹理细节信息，并且可避免传统融合策略中手动设计复杂的融合规则的问题。本方法的融合结果类似于经过锐化的红外图像，既具有清晰突出的目标，又包含丰富的细节信息，这对基于图像融合的目标检测与识别系统十分有利。

为了同时保留红外图像的热辐射信息和可见光图像丰富的纹理信息，本模块将红外与可见光图像融合问题构建为一个对抗性问题，其示意图如图 1 所示。

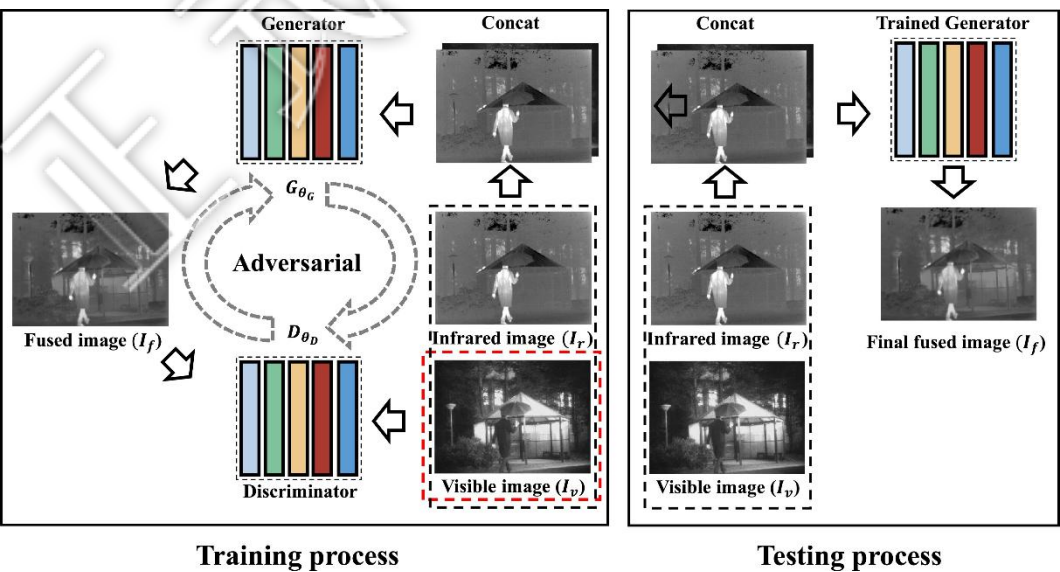


图 1 红外与可见光图像融合网络示意图

首先，将红外图像  $I_r$  和可见光图像  $I_v$  在通道维度上进行拼接。随后，拼接后的图像被输入至生成器  $G_{\theta_g}$  中，生成器的输出即为融合图像  $I_f$ 。由于所设计的生成器损失函数的特性，在没有判别器  $G_{\theta_d}$  的情况下， $I_f$  会倾向于保留红外图像  $I_r$  的热辐射信息以及可见光图像  $I_v$  的梯度信息。

接下来，将融合图像  $I_f$  和可见光图像  $I_v$  输入至判别器，判别器的目标是区分  $I_f$  和  $I_v$ 。所设计的红外图像与可见光图像融合模块在生成器和判别器之间建立了一种对抗博弈，在此过程中， $I_f$  会逐渐包含越来越多可见光图像  $I_v$  中的细节信息。

在训练阶段，一旦生成器生成的样本（即  $I_f$ ）无法被判别器区分，就能得到预期的融合图像  $I_f$ 。我们只需将  $I_f$  和  $I_v$  的拼接图像输入至训练好的生成器中，其输出便是我们最终的融合结果。

### 3.2 损失函数设计

红外与可见光图像融合网络的损失函数由两部分组成，即生成器的损失函数和判别器的损失函数。下面将分别对其进行介绍。

首先，生成器的损失函数包含两项：

$$L_G = V_{\text{FusionGAN}}(G) + \lambda L_{\text{content}} \quad (1)$$

其中， $L_G$  表示总损失，等式右侧第一项  $V_{\text{FusionGAN}}(G)$  表示生成器与判别器之间的对抗损失，其定义如下：

$$V_{\text{FusionGAN}}(G) = \frac{1}{N} \sum_{n=1}^N (D_{\theta_d}(I_f^n) - c)^2 \quad (2)$$

其中， $I_f^n$  表示由生成的融合图像， $N$  为融合图像的数量， $c$  为生成器希望判别器对伪造数据所持有的判断值。

第二项  $L_{\text{content}}$  表示内容损失， $\lambda$  用于平衡  $V_{\text{FusionGAN}}(G)$  与  $L_{\text{content}}$  之间的权重。由于红外图像的热辐射信息由其像素强度表征，而可见光图像的纹理细节信息可部分由其梯度表征，因此我们要求融合图像  $I_f$  与  $I_r$  具有相似的强度，且与  $I_v$  具有相似的梯度。具体地，定义如下：

$$L_{\text{content}} = \frac{1}{HW} (\|I_f - I_r\|_F^2 + \xi \|\nabla I_f - \nabla I_v\|_F^2) \quad (3)$$

其中  $H$  和  $W$  分别表示输入图像的高度和宽度， $\|\cdot\|_F$  表示矩阵的弗罗贝尼乌斯

范数， $\nabla$ 表示梯度算子。 $L_{\text{content}}$ 的第一项旨在使融合图像  $I_f$  保留红外图像  $I_r$  的热辐射信息，第二项旨在使融合图像  $I_f$  保留可见光图像  $I_v$  的梯度信息， $\xi$ 是一个正参数，用于控制这两项之间的权衡。

实际上，即使没有  $G_{\theta_b}$ ，我们也能得到一幅融合图像，该图像能够保留红外图像的热辐射信息和可见光图像的梯度信息。但这往往是不够的，因为可见光图像中的纹理细节不能仅通过梯度信息完全表征。因此，我们在生成器和判别器之间建立了一个对抗博弈，以基于可见光图像  $I_v$  对融合图像  $I_f$  进行调整。这能使  $I_f$  包含更多的纹理细节。

判别器的损失函数定义如下：

$$L_D = \frac{1}{N} \sum_{n=1}^N (D_{\theta_b}(I_v) - b)^2 + \frac{1}{N} \sum_{n=1}^N (D_{\theta_b}(I_f) - a)^2 \quad (4)$$

其中  $a$  和  $b$  分别表示融合图像  $I_f$  和可见光图像  $I_v$  的标签， $D_{\theta_b}(I_v)$  和  $D_{\theta_b}(I_f)$  分别表示对可见光图像和融合图像的分类结果。判别器旨在基于从图像中提取的特征来区分融合图像和可见光图像。我们采用最小二乘损失函数，该函数遵循最小化皮尔逊  $\chi^2$  散度的原则。这使得训练过程更加稳定，且判别器的损失函数能够快速收敛。

### 3.3 网络架构设计

本模块的网络架构由两部分组成，即生成器和判别器。两者的架构均基于卷积神经网络设计。

生成器的网络架构如图 2 所示，生成器是一个简单的五层卷积神经网络，其中第一层和第二层使用  $5 \times 5$  的滤波器，第三层和第四层使用  $3 \times 3$  的滤波器，最后一层使用  $1 \times 1$  的滤波器。每一层的步长均设置为 1，卷积过程中不进行填充操作。本模块的生成器输入是一幅未添加噪声的拼接图像。

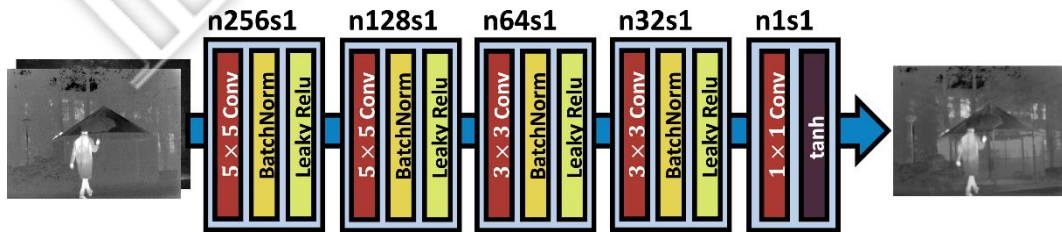


图 2 生成器的网络架构

为了提高生成图像的多样性，许多研究通常通过卷积层提取输入图像的特征图，然后通过转置卷积层将图像重构为与输入图像相同的尺寸。对于红外与可见

光图像融合而言，每一次下采样过程都会丢失源图像中的一些细节信息，而这些信息对融合至关重要（Yu et al. 2015）。因此，我们仅采用卷积层，不引入下采样操作。这也能保持输入和输出的尺寸一致，因此在我们的网络中无需使用转置卷积层。

此外，为了避免梯度消失问题，我们遵循深度卷积生成对抗网络（deep convolutional GAN）（Rao et al. 2023）中关于批归一化和激活函数的规则。为了降低对数据初始化的敏感性，我们在前四层中采用了批归一化操作，批归一化层能够使我们的模型更加稳定，还有助于梯度有效地反向传播到每一层。在激活函数方面，前四层使用 Leaky ReLU 激活函数，最后一层使用 tanh 激活函数。

我们所设计的判别器网络架构是一个简单的五层卷积神经网络，如图 3 所示。从第一层到第四层，卷积层均采用  $3 \times 3$  的滤波器，步长设为 2，且不进行填充操作。这一点与生成器网络不同，其根本原因在于判别器本质上是一个分类器，它首先从输入图像中提取特征图，然后对其进行分类。因此，通过将步长设为 2，可起到与池化层类似的作用。

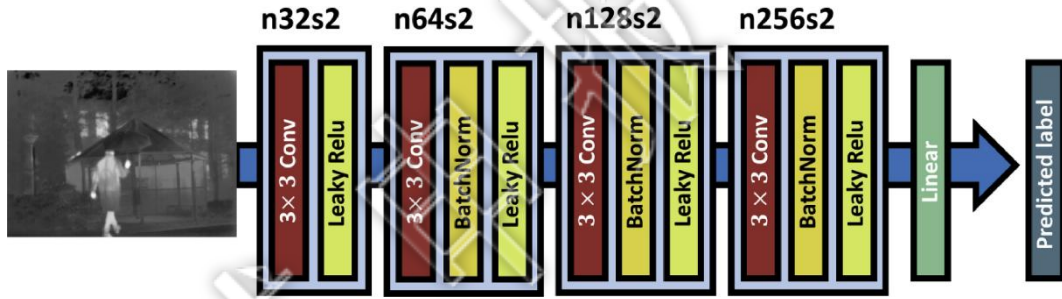


图 3 判别器网络架构

为避免在模型中引入噪声，我们仅在第一层对输入图像执行填充操作，其余三层卷积层均不进行填充。从第二层到第四层，我们使用批归一化层。此外，前四层均采用 Leaky ReLU 激活函数。最后一层是线性层，主要用于分类。

### 3.4 算法评估指标

仅通过主观评价难以准确评估融合性能，因此还需要融合指标进行客观评价。近年来，许多研究提出了各种融合指标，但似乎没有一种指标绝对优于其他指标。因此，有必要选择多种指标来评估不同的融合方法。我们使用六个指标对不同融合方法的性能进行定量评估，即熵（EN）、标准差（SD）、结构相似性指数（SSIM）、相关系数（CC）、空间频率（SF）和视觉信息保真度（VIF）。这六个指标的定义

如下：

熵（EN）基于信息论定义，用于衡量融合图像所包含的信息量。其数学定义如下：

$$EN = -\sum_{l=0}^{L-1} p_l \log_2 p_l \quad (5)$$

其中，L 表示灰度级数，在我们的实验中设为 256； $p_l$  是融合图像中对应灰度级的归一化直方图。熵值越大，表明融合图像包含的信息量越多，融合方法的性能越好。

标准差（SD）基于统计概念定义，反映图像中各个像素值与平均值的偏离程度。其数学定义如下：

$$SD = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (F(i, j) - \mu)^2} \quad (6)$$

其中，F 是尺寸为  $M \times N$  的融合图像， $\mu$  是融合图像 F 的平均值。对比度高的区域往往更能吸引人类的注意力，而对比度高的融合图像通常具有更大的标准差，这意味着该融合图像具有更好的视觉效果。

结构相似性指数（SSIM）用于模拟图像的损失和失真，衡量源图像与融合图像之间的结构相似性。SSIM 主要包含三个部分：相关性损失、亮度失真和对比度失真。这三个部分的乘积即为融合图像的评估结果，SSIM 的定义如下：

$$SSIM_{X,F} = \sum_{x,f} \frac{2\mu_x\mu_f + C_1}{\mu_x^2 + \mu_f^2 + C_1} \cdot \frac{2\sigma_x\sigma_f + C_2}{\sigma_x^2 + \sigma_f^2 + C_2} \cdot \frac{\sigma_{xf} + C_3}{\sigma_x\sigma_f + C_3} \quad (7)$$

$$SSIM = SSIM_{A,F} + SSIM_{B,F} \quad (8)$$

其中， $SSIM_{X,F}$  表示源图像 X 与融合图像 F 之间的结构相似性；x 和 f 分别表示源图像和融合图像在尺寸为  $M \times N$  的局部窗口中的图像块； $\sigma_x$  和  $\sigma_f$  表示标准差； $\sigma_{xf}$  表示源图像和融合图像的协方差； $\mu_x$  和  $\mu_f$  表示源图像和融合图像的平均值； $C_1$ 、 $C_2$  和  $C_3$  是使算法稳定的参数。 $SSIM_{A,F}$  和  $SSIM_{B,F}$  分别表示红外图像或者可见光图像与融合图像之间的结构相似性。SSIM 值越大，表明性能越好。

相关系数（CC）衡量融合图像与源图像之间的线性相关程度，其定义如下：

$$CC = \frac{(r_{AF} + r_{BF})}{2} \quad (9)$$

其中，

$$r_{XF} = \frac{\sum_{i=1}^M \sum_{j=1}^N (X(i,j) - \bar{X})(F(i,j) - \mu)}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N (X(i,j) - \bar{X})^2 \left( \sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \mu)^2 \right)}} \quad (10)$$

$\bar{X}$  表示源图像  $X$  的平均值。CC 值越大，表明融合图像与源图像越相似，融合性能越好。

空间频率（SF）用于衡量图像的梯度分布，其定义如下：

$$SF = \sqrt{RF^2 + CF^2} \quad (11)$$

其中，RF 是空间行频率，定义为

$$RF = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (F(i,j) - F(i,j-1))^2} \quad (12)$$

CF 是列频率，定义为

$$CF = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (F(i,j) - F(i-1,j))^2} \quad (13)$$

SF 值越大，表明融合图像具有越丰富的边缘和纹理。

### 3.5 算法优化方向

本项目设计的红外与可见光图像融合方法，通过构建生成器与判别器的对抗博弈机制，实现了对红外图像热辐射信息和可见光图像纹理细节信息的有效保留。该方法作为端到端模型，规避了传统融合策略中手动设计复杂活动水平度量和融合规则的局限性，为红外与可见光图像融合提供了一种全新的解决方案。

从技术框架来看，该方案的核心优势在于：

- 1) 生成器通过内容损失函数确保融合图像保留红外图像的像素强度（热辐射特征）和可见光图像的梯度信息（基础纹理）；
- 2) 判别器与生成器的对抗过程进一步强化了融合图像对可见光图像细节纹理的捕捉，使融合结果兼具目标显著性和细节丰富性；
- 3) 网络架构设计中避免下采样操作，减少了源图像细节信息的丢失，保证



了输入与输出尺寸的一致性，无需额外的图像重构步骤。

- 4) 此外，该框架具有通用性，不仅适用于同分辨率红外与可见光图像融合，还可扩展至不同分辨率源图像的融合任务，为其他类似融合需求（如遥感图像全色锐化）提供了可借鉴的思路。

尽管该方案表现出显著优势，但其仍存在可优化空间：例如，对极端复杂场景（如强噪声、低对比度源图像）的融合效果有待进一步验证；此外，扩展至更广泛的融合任务（如多光谱与全色图像融合）时，网络架构可能需要针对特定数据特性进行调整。同时需要进一步引入目标距离估计模块，通过单目深度估计或多传感器融合实时获取场景中目标与成像设备的距离信息。首先设计距离自适应的配准策略：根据距离动态调整变换模型——远距离采用全局尺度+旋转校正，近距离引入局部非刚性变换，针对不同距离下的几何偏差特性优化配准参数。例如，当距离估计结果显示目标为“近距离”时，模型侧重校正边缘区域的透视变形；当为“远距离”时，优先保证全局尺度一致性。最终的配准模型应该可覆盖 0-100 米（或更广）距离范围，减少因距离差异导致的配准失效。综上，未来工作可围绕这些方向展开，以提升框架的鲁棒性和适用范围。

### 3.6 算法预期结果

结合文献与性能指标，本方案的预期融合效果如下：

- 1) 视觉效果：融合图像同时清晰保留红外目标的热辐射信息（高亮显示）与可见光的纹理细节（如背景纹理、边缘轮廓），无配准异常（如重影、错位）。
- 2) 指标表现：熵（EN）与空间频率（SF）显著高于单一源图像（信息更丰富、纹理更清晰）；结构相似性指数（SSIM）与相关系数（CC）在融合图像与两源图像间均保持较高值（结构与特征保留好）；标准差（SD）与视觉信息保真度（VIF）优于传统方法（对比度高、信息损失少）。
- 3) 工程指标：在嵌入式平台 RV1126 上满足实时帧率要求（如 $\geq 25\text{fps}$ ），不同距离（0-100 米）下配准成功率 $\geq 95\%$ 。

## 4 项目进度安排

阶段	开始日期	结束日期	计划完成内容
阶段一	2025 年 9 月	2025 年 10 月	1. 技术调研与需求分析 2. 项目组组建与分工 3. 计划制定与计划书撰写
阶段二	2025 年 10 月	2026 年 6 月	1. 任务拆解与详细方案制定 2. 开发平台搭建 3. 数据集构建 4. 核心算法初研 5. 基础模型与算法实现
阶段三	2026 年 6 月	2026 年 12 月	1. 数据集与算法优化 2. 系统集成与测试 3. 全流程算法研发与实现 4. 阶段性总结与交流
阶段四	2026 年 12 月	2027 年 8 月	1. 系统优化与完善 2. 多场景综合测试 3. 成果总结与分析 4. 成果提交与验收

5 经费预算

科目名称	预算金额	备注
(一) 直接费		
1. 人工费		
(1) 专职研究人员费		
(2) 临时性研究人员费		
2. 设备费		
(1) 仪器设备使用费		



(2) 软件使用费		
3. 业务费	1.2	
(1) 材料费		
(2) 资料费		
(3) 印刷出版费		
(4) 专利与知识产权事务费		
(5) 会议费		
(6) 差旅费		
(7) 培训费		
4. 场地使用费		
(1) 场地物业费		
(2) 场地使用租金		
5. 专家咨询费	0.8	
(二) 间接费		
(三) 外委支出		
1. 外委研究支出		
2. 仪器设备租赁费		
3. 外协测试试验与加工费		
(四) 税金		
合计	2	

## 6 预期成果

### (1) 软件成果类

实现基于 AI 的红外与可见光融合系统一套

### (2) 知识产权类

发表高水平论文一篇

登记软件著作权两项

### （3）技术文件类

撰写完成：基于 AI 的红外与可见光融合系统研究报告

## 7 推广应用前景分析

红外与可见光融合系统依托“热辐射信息 - 结构细节互补”的核心优势，已在国家电网多电压等级、多运维场景中验证其应用价值。随着智能电网建设的深入推进，该系统的推广前景呈现“场景覆盖广度扩展、技术融合深度加强、综合效益持续提升”的特征，未来将成为电网运维领域的核心技术工具。

从应用场景的拓展潜力来看，当前系统已覆盖特高压、超高压、配电网等电压等级，以及变电站、输电线路、电缆隧道等关键场景，未来将进一步向“全运维链条、全复杂环境”延伸。在特高压领域，除现有线路巡检外，应用范围将拓展至换流站阀厅、电抗器等结构更复杂的设备，结合数字孪生技术实现设备热状态与结构状态的实时对应，提升复杂设备的状态感知精度；在配电网领域，监测对象将从当前的变压器、环网柜，延伸至台区分支箱、表计接线等末梢设备，通过轻量化融合算法适配配网设备数量多、分布散的特点；在极端环境适配方面，针对高原冻土区的低温干扰、沿海盐雾区的设备腐蚀等问题，将通过优化传感器抗干扰性能与算法环境适应性，实现运维场景“无死角”覆盖。参考现有推广速度，预计未来 3-5 年，该系统可覆盖国家电网 90% 以上的核心运维场景。

从技术融合的深化方向来看，该系统将与人工智能、边缘计算、5G 等技术深度结合，推动电网运维从“被动监测”向“主动预防”转型。在多模态数据融合方面，将突破当前仅依赖红外与可见光图像的局限，整合设备声纹（如异常振动声音）、机械振动（如变压器铁芯振动信号）、局部放电数据等多源信息，构建“热 - 声 - 振 - 电”多维度诊断体系，使设备缺陷识别准确率进一步提升 10%-15%；在智能算法应用方面，电力设备缺陷大模型将与融合系统结合，实现“缺陷自动识别 - 故障原因分析 - 运维方案生成”的全流程自动化，例如通过模型可直接判断设备过热是源于接触不良还是绝缘老化，并输出针对性的消缺建议；在边缘计算部署方面，将复杂融合算法轻量化后部署于无人机、巡检机器人等终端设备，实现“数据采集 - 实时推理 - 现场预警”的同步完成，使现场响应时间从分钟级缩短至秒级，显著提升应急处置效率，这些技术融合方向将持续增强系统的核心竞争力。

从经济与社会效益的提升空间来看，系统规模化推广将为国家电网带来显著

的价值增量。经济效益层面，随着应用规模扩大，设备采购与部署成本预计降低 20%-30%，而运维效率提升带来的成本节约将更为突出 —— 根据现有试点数据测算，系统全面推广后，每年可减少人工巡检成本超 50 亿元，通过提前预警设备缺陷避免的停电损失可达百亿元级别；社会效益层面，系统将进一步降低运维人员的作业风险，在电缆隧道、高空杆塔等高危场景中，人工巡检替代率可提升至 98% 以上，减少人员暴露于危险环境的次数；同时，通过减少停电时长与优化能源利用效率，每年可间接减少数十万吨碳排放，为 “双碳” 目标实现提供支撑，此外，技术推广还将带动传感器制造、算法研发等上下游产业发展，形成协同发展的产业生态。

从推广落地的保障路径来看，国家电网已形成的 “问题导向研发、标准化建设、产学研协同” 等经验，将为系统全面推广提供基础支撑。未来将进一步完善技术标准体系，细化不同场景下传感器选型参数、算法性能指标、运维操作流程等规范，确保各省级电力公司应用的兼容性与一致性；在推广策略上，将延续 “试点验证 - 技术迭代 - 规模推广” 的闭环模式，优先在运维需求迫切、场景代表性强的区域（如东部用电负荷中心、西部复杂地形区域）扩大试点范围，降低大规模推广的风险；在人才培养方面，通过 “无人机飞手实训” “AI 算法应用培训” 等模式，培养兼具电力运维知识与技术应用能力的复合型人才，预计未来 2-3 年可培养专业技术人才超万名，为系统落地提供人力保障。

综上，红外与可见光融合系统在国家电网的推广应用，既具备技术可行性与场景适配性，又符合智能电网与新型电力系统建设的需求，未来将成为推动电网运维数字化、智能化转型的关键技术，在保障电网安全稳定运行、提升综合效益方面发挥重要作用。