

引文格式: 王宏民, 覃才, 林蔚, 等. 基于深度学习及声电信号跨模态融合的电弧增材制造异常检测方法[J]. 航空制造技术, 2026, 69(6): 25010135.

WANG Hongmin, QIN Cai, LIN Wei, et al. Anomaly detection method for wire arc additive manufacturing based on deep learning and cross-modal fusion of acoustic and electrical signals[J]. Aeronautical Manufacturing Technology, 2026, 69(6): 25010135.

基于深度学习及声电信号跨模态融合的电弧增材制造异常检测方法

王宏民¹, 覃才², 林蔚², 郭慧林², 潘增喜^{2,3}, 宋莹莹⁴

(1. 五邑大学应急技术与管理学院, 江门 529020;

2. 五邑大学机械与自动化工程学院, 江门 529020;

3. 伍伦贡大学机械、材料、机电与生物医学工程学院, 伍伦贡 2522;

4. 五邑大学电子与信息工程学院, 江门 529020)

[摘要] 电弧增材制造(Wire arc additive manufacturing, WAAM)具有快速成形与适于轻量化设计的优点,在航天器关键零件制造上具有巨大潜力。针对WAAM在生产过程中可能产生各种缺陷的问题,对WAAM过程异常检测进行研究,提出了一种改进残差深度时序卷积网络(Improved residual deep temporal convolutional network, IRD-TCN)结合基于残差增强的轻量注意力机制(Residual-enhanced lightweight attention, RELA)模块的声电信号跨模态融合无监督异常检测方法。由于WAAM工作条件复杂,单一信号源的检测能力较为有限,因此采用基于跨模态关系分析的声电信号小波特征融合方法,结合IRD-TCN与RELA分析两种传感数据之间的关联变化,实现理想的检测效果。最终试验所得出的精确率、召回率、F1-score分别达到了98.37%、97.73%和98.10%,解决了传统数据融合方式在WAAM过程异常识别中准确性与鲁棒性不足的问题。

关键词: 电弧增材制造; 跨模态融合; 深度学习; 异常检测; 时序卷积; 注意力机制; 小波变换

中图分类号: V26; TG115.28

文献标志码: A

DOI: 10.16080/j.issn1671-833x.25010135



王宏民

教授, 博士, 研究方向为电弧增材制造、工业机器人、医疗机器人。

电弧增材制造(Wire arc additive manufacturing, WAAM)是一种新兴的金属加工技术,基于简单的硬件和材料即可实现高自由度、复杂的零件设计,已成为航空航天领域不可或缺的核心技术^[1]。WAAM以电弧为热源,将金属丝材熔化并逐层堆积成形,实现复杂零件的制造和修复^[2]。然而,在实际生产中,WAAM高温、高速的工作特性与多变的工艺参数

易引发多种缺陷,如塌陷、飞溅、驼峰等。这些缺陷会影响整个制造过程后续的沉积层,造成最终产品的力学性能差及成本增加等问题。因此,开发有效的WAAM缺陷检测技术,对于提升WAAM技术的生产质量具有重要的现实意义^[3]。目前,结合深度学习的WAAM过程缺陷检测已有一定的研究基础,常用的检测源有图像、超声波、红外热成像及电流电压

收稿日期: 2025-09-10; 退修日期: 2025-10-27; 录用日期: 2025-11-28

基金项目: 国家重点研发计划(2020YFB1313300); 国家自然科学基金(62073274); 江门市科技计划项目(江科(2025)172号)

通信作者: 宋莹莹, 高级实验师, 研究方向为电子信息。

数据等, 主流的检测算法有 YOLO、Transformer、长短期记忆网络等。即使目前能够对多种类型的数据结合不同的深度学习模型进行检测, 但均存在一定的局限性。WAAM 缺陷数据的模拟较为困难并具有随机性, 因此图像类数据采集效率低, 难以反映出焊道内部的缺陷, 且超声波穿透性易受高温环境影响, 单一的电信号所包含的信息较少, 难以支撑全面、准确的缺陷检测。此外, 尽管当前主流的算法模型已被广泛用于 WAAM 过程的缺陷检测, 但大多基于有监督学习, 对数据量要求较高, 面对长时序、采集难度大的监测数据时, 不仅难以捕捉长时间依赖关系特征, 还会面临数据量不足等问题。因此, 开发一种更先进的缺陷检测方法已成为提高 WAAM 过程可靠性, 拓展其应用前景的迫切需求。

WAAM 缺陷检测借助神经网络模型对海量数据迭代学习的优势, 大幅提升了检测精度与效率, 但受限于 WAAM 复杂的工作机制, 模型在实际训练与检测过程中仍存在以下问题: (1) 缺陷数据模拟难度高, 数据采集困难^[4]; (2) 虽然多数据融合已有一定研究基础, 但不同数据间的联合检测策略判断仍需改善, 例如两种不同类型数据的检测结果不相同时, 误检率较高; (3) 在多数据融合, 尤其是包含图像类数据时, 神经网络模型分析数据的运算量大, 检测的精度与效率难以兼顾。

目前, 众多学者针对 WAAM 缺陷检测已开展大量的研究工作。Mu 等^[5] 使用电流电压数据、图像数据和点云数据, 结合 3 种机器学习算法完成决策融合, 实现了 WAAM 过程的缺陷检测, 其中多层感知器分类器与 YOLOv5 分别用于时间和空间尺度的检测, 变分自编码器负责决策层的融合任务。Jirapattanaporn 等^[6] 针对烧穿类型的缺陷, 通过麦克风阵列获取声音数据后, 从频谱图中分析

不同焊缝类型的声音差异, 并使用递归神经网络、长短期记忆网络和单类支持向量机网络模型进行缺陷检测工作, 验证了 3 种网络用于声音缺陷检测方面的有效性。Mattera 等^[7] 利用快速傅里叶变换和离散小波变换方法进行频域分析分类, 以识别焊接信号数据中的异常, 试验表明, 快速傅里叶变换和离散小波变换二者的结合可以有效提高焊接缺陷的检测精度。Zhang 等^[8] 提出了一种基于声学信号的电弧定向能量沉积缺陷识别方法, 使用小波时频图和卷积神经网络识别不连续性和孔隙缺陷。然而, 以上研究大多停留在初步理论及试验阶段, 着重于验证神经网络模型的有效性以及对精度的追求, 未考虑模型运算效率与数据采集的困难性问题, 实际应用时的合理性与检测能力仍有待评估。上述研究为 WAAM 缺陷检测的进一步发展提供了一定的理论支撑。

本文提出了一种适用于 WAAM 的无监督异常检测方法。通过将 IRD-TCN 结合 RELA 对声电信号进行建模, 利用局部感知残差分支, 在强化短程依赖和微结构特征捕捉能力的同时, 有效控制模型的参数量和计算开销。IRD-TCN 通过学习正常类型声音信号的小波系数矩阵特征, 并将其拟合并映射至正常工况下的电信号数据, 实现声电信号间的跨模态关联建模, 根据所拟合的结果对异常数据进行识别。模型的训练仅依靠正常类型的数据即可完成, 模型性能验证与测试仅需少量异常数据, 可有效解决 WAAM 过程中异常数据采集困难、数据量不足的问题。

1 WAAM 异常检测方法

WAAM 异常检测方法整体分为数据处理和异常检测两部分。数据处理包括对齐分割、特征提取及融合分析 3 个步骤, 用于对原始声电信号进行预处理和特征构建, 为后续分析

提供有效输入; 异常检测部分则依据 RELA 与 IRD-TCN 算法的结合, 对经过处理的数据特征进行建模与判别, 最终完成异常数据的识别与定位。

1.1 数据处理

数据处理的目的是从时间序列信号中提取能够反映原始数据特征的有效信息, 从而对数据进行对齐分割、特征提取及融合分析等一系列的操作。其中, 数据对齐针对采样率不一致的声音信号与电信号, 为实现统一输入所做的基础预处理; 数据分割是将每个焊道的长时间数据窗口分割成连续的短时段数据, 有利于后续检测的连续性分析。

数据进行对齐分割操作后, 进一步使用连续小波变换 (Continuous wavelet transform, CWT) 对数据进行时频分析并提取关键特征。相较于傅里叶变换, 连续小波变换可以同时时间和频率上对信号进行局部化分析, 对后续数据的进一步分析提供了有利条件^[9]。

WAAM 工作过程的声音信号与电信号在时间分布上均存在一定的规律^[10]。本文采用跨模态关系分析的方式实现数据融合, 依据声电信号间存在的规律进行数据的融合分析。

1.1.1 数据对齐与分割

数据对齐与分割预处理的目的是为后续特征提取及异常检测提供高质量、标准化的输入数据^[11]。本文以起弧瞬间声电信号形成有效波形的时刻为对齐起点, 以灭弧时刻前的整秒时刻为终点, 完成数据对齐; 随后采用 1 s 时长滑动窗口对声电数据进行分割, 其中声音数据采样率为 96 kS/s, 单个窗口包含 96000 个数据点, 电信号采样率为 2.5 kS/s, 单个窗口包含 2500 个数据点; 最后通过时间维度的最大池化操作, 实现声电信号采样点的进一步对齐。

1.1.2 小波特征提取

传统的时序分析方法通过傅里叶变换将信号分解为全局正弦波分

量,虽能表征信号的频域特性,但不利于分析时间相关性强的时序信号。为此,本文通过 Morlet 小波作为小波基函数,利用连续小波变换对声电数据进行特征提取,将信号分解为不同尺度和时间的小波系数组合,以适配本文 IRD-TCN 模型进行特征学习。小波特征提取过程如图 1 所示。

连续小波变换通过将信号与一系列缩放和平移的小波基函数进行卷积,生成时频域多尺度表征,即

$$W(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi^* \left(\frac{t-\tau}{s} \right) dt \quad (1)$$

式中, $x(t)$ 为预处理后的声电信号; s 为尺度系数; τ 为平移参数; t 为连续时间变量; $\psi(t)$ 为 Morlet 小波基函数,所生成的多尺度特征将构建小波系数矩阵作为异常检测的输入,其表达式为

$$\psi(t) = \pi^{-\frac{1}{4}} e^{i\omega_0 t} e^{-\frac{t^2}{2}} \quad (2)$$

式中, ω_0 为中心角频率。

1.1.3 跨模态关系融合分析

声音信号与电信号作为两种不同物理来源的感知数据,分别反映了工作过程中的声学行为与电气状态^[12]。由于两类信号均源于同一工作过程,其在时间演化趋势上有内在关联性,具备进行跨模态数据融合分析的基础。本文以声信号为主导模态,利用从时频维度提取的小波特征作为表示基准,通过神经网络向电信号进行映射拟合,从而建立声电信号间的关联度量模型,过程如图 2 所示。通过以上方式训练得到的 IRD-TCN 模型能够关注到两种数据在时间维度上的关联性,依据这种关联是否被打破进行异常识别。

跨模态关系融合之所以具备有效性,是基于声电信号在物理机制层面的互补特性。电信号直接表征电弧通道的电学特性,能够稳定地反映系统的能量输入,熔池的导电性以及工艺过程的整体连续性,对于声学环

境等外部干扰表现出较强的稳定性与鲁棒性。声音信号则源于电弧放电所激发的等离子体周期性振荡,金属熔滴过渡撞击以及保护气体流动等多物理过程的耦合,对塌陷、飞溅等瞬态异常,以及由层高变化引起的声学环境改变极为敏感。由于二者在物理生成机制上存在本质差异,电信号提供了工艺过程的宏观能量态势,声信号则揭示了其微观动态扰动行为。因此,两种信号的互补内容主要有以下两点。(1) 敏感性互补。电信号能稳定监测系统内部参数变化,声信号对外部过程状态变化表现敏感。(2) 鲁棒性互补。电信号受外部噪声影响小,声音信号能弥补电信号在复杂物理信息识别困难条件下的局限性。

基于这种互补性,可进一步从时间维度刻画二者的耦合规律。正常焊接过程所产生的声音信号 $A(\eta)$ 与电信号 $E(\eta)$ 在时间上是高度耦合的,满足以下非线性关系。

$$E(\eta) = f_\theta(A(\eta)) + \varepsilon \quad (3)$$

式中, f_θ 是通过神经网络建模的映射函数; ε 是系统扰动项。

模型训练的目标是最小化映射误差,即

$$Y_{\text{train}} = \frac{1}{T} \sum_{\eta=1}^T \|E(\eta) - \hat{E}(\eta)\|_2^2 \quad (4)$$

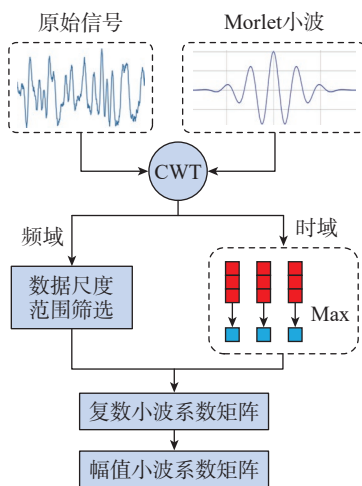


图 1 小波特征提取过程
Fig.1 Wavelet feature extraction process

式中, T 为时间序列的长度; $\hat{E}(\eta)$ 为目标信号模型的预测值。

在异常情况下,声电模态间的物理耦合被破坏,映射函数 f_θ 无法准确重建目标信号,从而导致预测误差显著上升。基于此,可以定义异常分数为

$$S(\eta) = |E(\eta) - \hat{E}(\eta)| \quad (5)$$

当 $S(\eta)$ 超过设定阈值时,即判断该时段焊接过程存在异常,实质上是通过检测模态间映射偏差实现异常捕捉。

跨模态关系融合构建了模态间的结构化关联建模机制,赋予模型判别“模态一致性”的能力。当成形过程出现异常时,即便此类异常在单一信号中未呈现明显特征,也会破坏声电信号间固有的耦合关系,模型可捕捉到这种物理关联的失配,从而精准感知异常的发生。相比之下,单一模态模型因缺乏跨域参照而无法检测此类异常;简单的特征拼接不仅难以有效发挥各自模态的独特优势,反而会因为模态间的量纲、信噪比及动态范围差异,引入大量干扰,并且浅层的数据堆砌完全忽略了声、电信号之间由底层物理过程所决定的深层耦合关系,不易于从本质上区分“正常耦合”与“失配异常”。本文通过对声、电信号间的结构性关联进行建模,将过程中的先验物理知识融入学习框架中。这种方法不仅增强了模型对异常的鲁棒性,还提高了其对异常信号的敏感性,从而构成了多模态融合策略的核心优势。

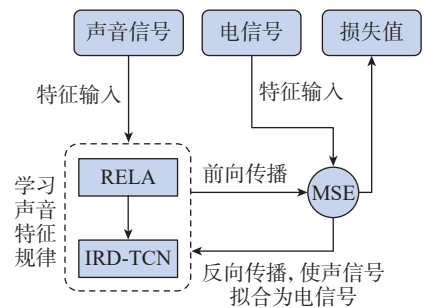


图 2 跨模态关系融合过程
Fig.2 Cross-modal relationship fusion process

1.2 基于 RELA 与 IRD-TCN 的异常检测

本文提出的 WAAM 异常检测模型基于两个部分: RELA 和 IRD-TCN。RELA 负责捕捉时间序列中的有效信息; IRD-TCN 由 3 个子模块组成, 分别是嵌入模块、主干模块、预测输出模块, 专注于从时间序列数据中提取多层次的时域特征并进行异常检测^[13]。二者的结合能够高效地从多层次的时域特征中识别异常模式, 整体检测过程如图 3 所示。

1.2.1 RELA 模块

RELA 是一种基于轻量注意力结构改进得到的注意力模块, 可进一步提升模型在长序列建模中的局部依赖感知能力。该模块以传统轻量注意力结构为基础, 结合线性注意力思想进行改进, 并引入一个并行的局部感知残差分支, 以此补偿长距离注意

力在局部细节建模方面的不足^[14]。

在标准注意力中, 输入序列 $X = \mathbb{R}^{B \times N \times D}$ (B 为 Batch size 的大小; N 为序列长度; D 为通道数) 被映射为查询 (Query, Q)、键 (Key, K) 和值 (Value, V) 3 种张量。

$$Q = XW^Q, K = XW^K, V = XW^V \quad (6)$$

注意力输出计算为

$$Y = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (7)$$

式中, d 为缩放因子。

传统的注意力机制尽管在长距离依赖建模方面表现优异, 但其计算复杂度为 $O(N^2d')$, 其中 d' 为特征维度, 难以适用于长序列或资源受限场景。RELA 采用线性注意力思想, 通过特征映射函数 $\phi(n)$ 将点积内积运算替换为核函数映射, 并将注意力计算重新排列, 传统注意力与线性注意力运行机制如图 4 所示。线性

注意力的运算过程为

$$Y_{\text{attn}} = \frac{\phi(Q) \cdot (\phi(K)^T V)}{\phi(Q) \cdot (\phi(K)^T 1_N)} \quad (8)$$

式中, $\phi(Q)$ 和 $\phi(K)$ 分别表示对查询和键的核特征映射。

传统多头注意力主要有两部分参数来源, 分别是 Q 、 K 、 V 与输出投影, 若输入通道数为 d_{model} , 则参数量为

$$P = 3 \times d_{\text{model}}^2 + d_{\text{model}}^2 \quad (9)$$

本文在线性注意力思想的基础上采用了低维投影以降低参数规模与计算复杂度, 将原有的维度数切分为多个注意力头 H 、压缩每头维度 d_{head} , 维持 $H \times d_{\text{head}}$ 的参数量为 $0.25d_{\text{model}}$, 则最终的参数量计算如下。

$$P = 3 \times (H \times d_{\text{head}})^2 + (H \times d_{\text{head}})^2 \quad (10)$$

因此, RELA 的参数量较传统多头注意力减少了 75%, 有效降低了计算复杂度与运算时间。

尽管轻量注意力机制在处理全

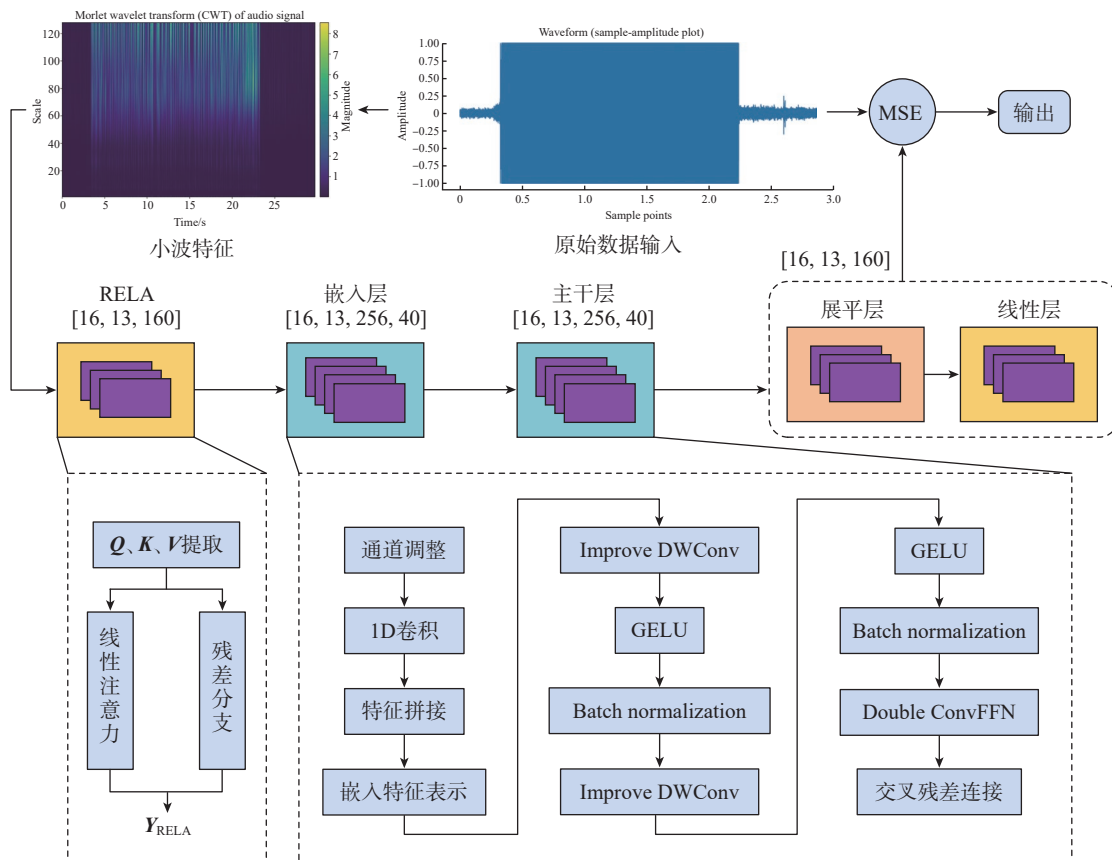


图 3 异常检测过程

Fig.3 Anomaly detection process

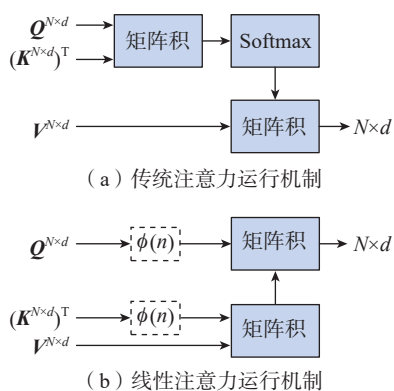


图4 注意力运行机制对比

Fig.4 Comparison of attention operation mechanisms

局依赖时效率较高,但其缺乏局部归纳偏置,不擅长捕捉短程依赖和局部纹理特征。为解决该问题,RELA中引入局部感知残差分支,该分支对张量 $V = \mathbb{R}^{B \times H \times N \times d_{\text{head}}}$ 按注意力头维度施加二维卷积,该卷积在时间轴上滑动以提取窗口长度为 k 的局部上下文,而在特征维上设置卷积核宽度为 1,保证对同一通道的逐时刻滤波。卷积分支的局部特征提取如式(11)所示。

$$V_{b,h,n,c}^{\text{loc}} = \sum_{t=-\gamma}^{\gamma} \omega_{h,i} V_{b,h,n+t,c} \quad (11)$$

式中, $V_{b,h,n,c}^{\text{loc}}$ 表示在相应位置的局部张量; b 为批次编号; h 为注意力头编号; n 为数据时间步长; c 为通道索引; γ 为局部卷积感知范围,即 $\gamma=0.5(k-1)$ 。得到局部特征后经过 ReLU 层与归一化 (Batch normalization, BN) 处理得到:

$$Y_{\text{res}} = \text{BN}(\text{ReLU}(\text{DWConv}(V^{\text{loc}}))) \quad (12)$$

最终两个分支的输出进行融合以构成 RELA 的输出:

$$Y_{\text{RELA}} = Y_{\text{attn}} + \lambda Y_{\text{res}} \quad (13)$$

式中, λ 为可学习的标量参数,用于平衡注意力分支与局部感知分支的贡献。

由于采用了 DWConv,分支新增参数量仅为 $H \times k$,计算量约为 $B \times H \times N \times d_{\text{head}} \times k$,相较于于支的主要计算项可忽略不计,极大程度提高了局部细节的建模能力。

综上分析,RELA 大幅降低了计算的复杂度,同时通过引入局部感知残差分支,有效补充了轻量注意力在局部依赖建模上的不足,增强了模型在捕捉短时动态和微小异常方面的能力。通过全局与局部特征的融合,RELA 显著提升了模型在长序列场景下的表达能力与鲁棒性,尤其适用于时间序列识别和异常检测等对细节敏感的任务。

1.2.2 IRD-TCN 初始特征嵌入模块

由于时间序列数据的各特征变量间直接相关性较弱,且不同时段的数据蕴含多尺度动态变化特征,若直接将多个特征变量合并为单一向量,难以捕捉数据在时间维度上的复杂映射规律^[15],因此 IRD-TCN 使用分块独立嵌入的方法进行研究。设输入的时间序列数据为 $A_{\text{in}} \in M \times L$,其中 M 和 L 分别代表变量个数和时间长度。整个分块独立嵌入的过程通过 1D 卷积来实现,每个小块的长度由卷积核尺寸决定,从一个输入通道映射到 D 个输出通道。通过将输入数据调整为 $A_{\text{comb}} \in \mathbb{R}^{M \times 1 \times L}$,1D 卷积层将每个变量的时间序列分解成 N 个小块并对每个小块中的多个变量进行独立处理,再把所有的小块合并到一个 D 维的向量中,即

$$A_{\text{com}} = \text{Embedding}(A_{\text{in}}) \quad (14)$$

$$A_{\text{comb}} \in \mathbb{R}^{M \times D \times N} \quad (15)$$

将长时序数据划分为多个数据块,可使模型更高效地处理整体数据,同时更好地捕捉数据的局部特征。此外,该过程对各变量的时间序列采用独立处理方式,因此多维度的变量信息得到完整保留。

1.2.3 IRD-TCN 改进型主干模块

传统 TCN 中常用的标准卷积感受野有限,难以捕捉序列中的长程依赖^[16]。因此,IRD-TCN 的主干模块在传统 TCN 的基础上进行了系统性创新改进,引入了改进型深度可分离卷积 (Improve DWConv)。与传统 DWConv 不同,该方法采用了更大的

卷积核,可以显著扩大感受野,同时在每一模块内部堆叠两层 Improve DWConv,增强局部特征逐步细化与整合的能力。每次 Improve DWConv 操作后,均接入 GELU 激活与 Batch normalization,进一步提升了非线性表达能力和训练稳定性。

在特征非线性变换模块中,IRD-TCN 采用了双卷积前馈网络 (Double ConvFFN) 以增强特征表达能力。传统的前馈网络(如 MLP)主要依赖全连接层,难以充分利用序列局部结构,而 IRD-TCN 能够通过两次 1D 卷积分别沿时间维度和通道维度进行建模。

在特征流动与梯度传递机制方面,IRD-TCN 还引入了交叉残差连接 (Cross residual connection) 进一步缓解深层网络中信息衰减的问题。不同于传统残差连接仅在模块内部进行特征叠加,交叉残差连接设计允许当前模块的输出,同时融合上一模块的输出特征,具体定义为

$$Y = H' + \alpha Z + \beta Z_{\text{prev}} \quad (16)$$

式中, Z 为本模块输入特征; H' 为当前模块的非线性变换输出; Z_{prev} 为前一模块的输入特征; α 与 β 为可学习参数。这种设计带来两方面的优势:一方面增强了多层次特征的信息交互与融合,使低层次细粒度特征和高层次抽象特征可同时用于预测;另一方面,有助于缓解深度模型训练中的梯度消失问题,提高训练稳定性与最终精度。

通过以上改进,IRD-TCN 的主干模块在感受野扩展、特征细粒度提取、变量交互建模与跨层信息流动 4 个方面实现了系统性增强,为后续异常检测任务提供了坚实的特征基础。

1.2.4 IRD-TCN 预测输出模块

预测输出模块的主要任务是完成从高维特征空间到低维输出空间的转换,同时最大程度保留判别性信息,具体流程包括特征整合、展平、线性变换以及输出层激活 4 个步骤,各

环节密切配合,确保最终输出结果具备良好的可训练性和推理速度。其中,特征整合阶段通过展平操作将二维特征张量转换为一维向量,展平操作的数学形式可表示为

$$y = \text{Flatten}(\mathbb{R}^{C \times T}) \quad (17)$$

式中, C' 为主干模块输出的特征通道数; T' 为经过下采样后的时间步长。展平操作不仅简化了后续线性变换的操作复杂度,而且能将跨通道与跨时间步的信息充分融合在一起,为最终的判别提供有利条件;其次,线性变换层对展平后的特征向量 y 施加仿射映射,实现特征空间的降维,输出预测向量 \hat{y} ,即

$$\hat{y} = W_y + b \quad (18)$$

式中, $W \in \mathbb{R}^{K' \times (C' \times T')}$ 为线性层的权重矩阵; $b \in \mathbb{R}^{K'}$ 为偏置向量, K' 为输出类别数。这种线性映射的作用:将高维复杂特征向量映射到目标预测空间,提取出与最终任务最相关的特征组合,从而完成特征向量到预测目标的有效对齐。

2 试验与分析

2.1 试验平台与数据采集

试验所使用的机器人为 ABB IRB 2600 工业机器人;声音信号采集传感器采用 Xiris 自动化公司的 WeldMicTM,频率响应范围为 0.05~20 kHz,能够覆盖电弧声波 10 kHz^[17] 以内的重要频段,可保证过程关键物理特征被无损捕获;电信号采集传感器采用 National Instruments 公司的 NI USB-6210 系列数据采集卡,最高支持采样率为 250 kS/s。数据采集过程中声音信号与电信号的采样率分别为 96 kS/s 与 2.5 kS/s。试验平台如图 5 所示。数据采集的全程保持以下参数:焊枪口到沉积层间隔距离 15 mm、保护气流量 30 L/min,声音信号采集传感器固定于焊枪左侧且偏离于焊枪口正上方 15 cm 处,声音信号接收口指向焊接点,以确保声波主要沿直射路径到达传感器,电信

号数据采集卡则记录每个采样点的电流和电压值。数据采集全程共向上堆叠 12 层,异常类型数据通过手动调整变量实现,具体数据形式如表 1 所示,各焊道类型如图 6 所示。

2.2 声电信号数据层间对比试验

为验证数据融合的必要性,需进行声电信号数据层间对比试验,分为两部分:层间损失标准化差异对比试验与层间损失对比试验,旨在揭示单一信号的局限性,从而为数据融合提供依据。

声音信号能够直接反映焊接过程的状态,是进行状态监测和异常诊

断的理想信息源。具体而言,声音信号源于电弧剧烈燃烧、焊丝末端金属熔滴过渡、熔池振荡及保护气体流动的共同作用。当电流通过电极与工件之间形成等离子通道时,局部气体受高温高压作用发生周期性膨胀和收缩,从而辐射出声波。当有异常情况发生时,如保护气流量不足、送丝速度过快等,电弧燃烧状态会随之改变,声压幅度与频谱结构会产生相应的结构扰动^[18]。相比之下,电信号对异常的敏感度则相对不足,其反映的是电弧回路的综合电学状态,而非直接的物理事件。电源系统的快速

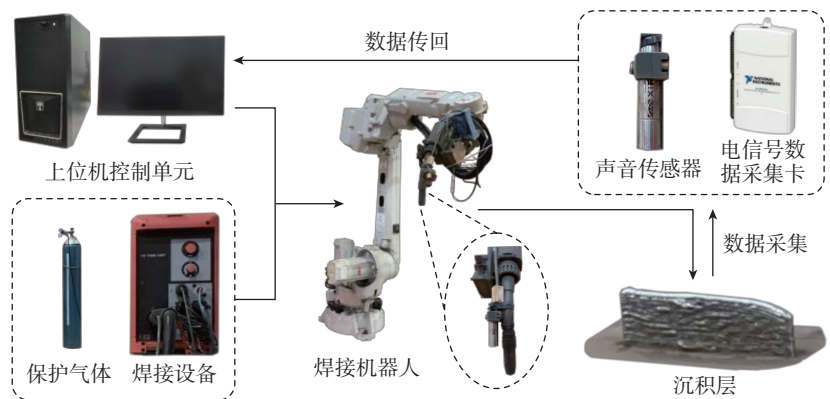


图 5 试验平台
Fig.5 Experimental platform

表 1 数据采集表
Table 1 Data collection form

层高	移动速度/ (mm/min)	送丝速度/ (m/min)	间隔距离/ mm	保护气流量/ (L/min)	类型
1	300	5	15	30	正常
2	300	6	15	30	正常
3	200	5	15	30	塌陷
4	300	5	30	10	飞溅
5	800	5	15	30	驼峰
6	300	3	15	30	正常
7	400	5	15	30	正常
8	200	5	15	30	正常
9	300	5	15	4	保护气不足
10	300	7	15	30	正常
11	300	8	15	30	正常
12	300	4	15	30	正常

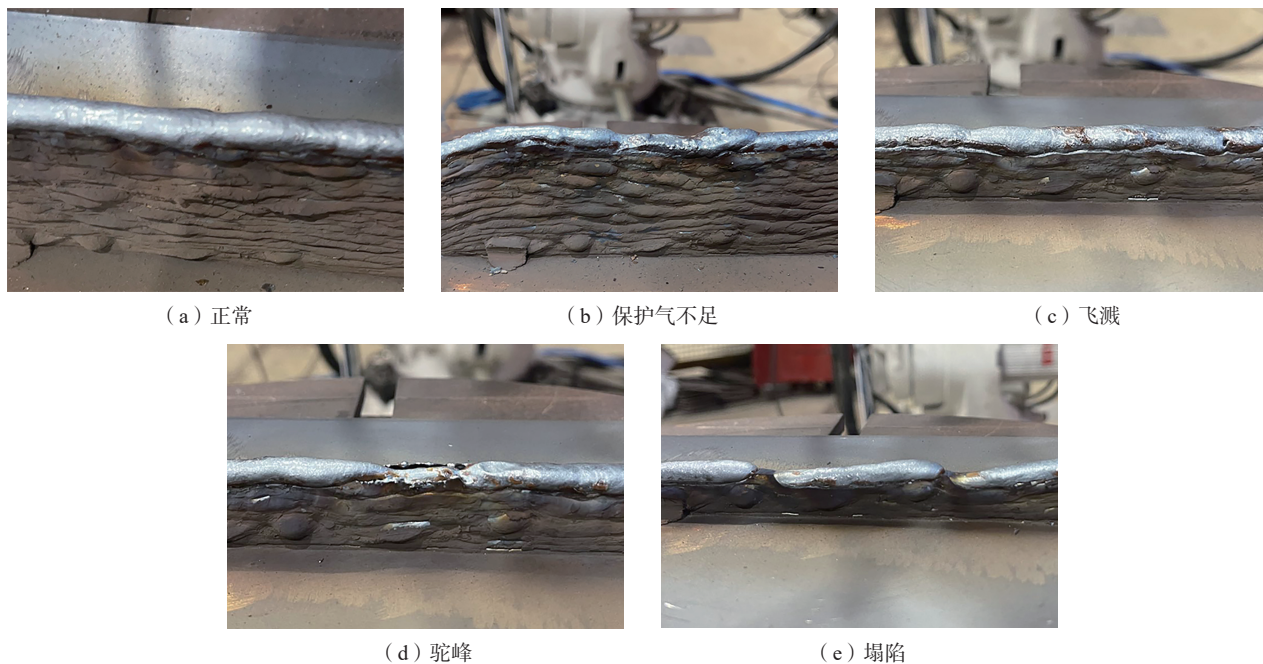


图6 焊道类型

Fig.6 Weld bead types

闭环控制会主动补偿外部扰动,力图维持电流电压的稳定,进而平滑甚至掩盖部分异常特征。基于此,通过层间损失标准化差异对比试验,以验证两种类型数据在正常层与异常层之间的差异性。

此外,声音信号在传播过程中会受到传播路径长度与方向的显著影响。层高的增加改变声波的反射和边界条件,不仅会导致部分频率产生相位干涉变化或混响条件改变,还会使回声声波传播至传感器的路径延长^[19],最终导致不同层高采集的声音信号频谱结构存在明显差异。相比之下,电信号通过数据采集卡经闭合回路采集,对层高变化引发的声学环境改变等物理因素具有较强的鲁棒性,其信号特征能够保持相对稳定。基于此,本文设计层间损失对比试验,以验证相较于声音信号,电信号在层高变化影响下具备更高的稳定性。

两部分试验的数据集划分如下。层间损失标准化差异对比试验训练集为第1层数据,共95个样本;验证集与测试集为同一批样本,第4层与

第9层均为33个样本。层间损失对比试验训练集为第1层数据,共95个样本;验证集与测试集为同一批样本,第2层与第6层均为31个样本。所有样本声音信号数据与电信号数据一一对应。

开展层间损失标准化差异对比试验,通过对相邻两层的损失值进行归一化处理并将其映射至0~1之间,得到标准化损失差异值,以此验证声音信号与电信号对异常的敏感度差异。标准化损失差异值 z_{norm} 计算方式为

$$z_{\text{norm}} = \frac{|\alpha' - \beta'|}{|\alpha' + \beta'|} \quad (19)$$

式中, α' 与 β' 为两个层级各自得到的损失值。

试验以第1层(正常类型)数据训练得出的模型,对第4层(飞溅异常类型)和第9层(保护气不足异常类型)的数据进行损失标准化差异曲线比较,如图7所示。可以看出,两种信号所进行的30次迭代损失对比中,第1次迭代结果后的每个轮次声音信号的训练和验证损失值差异均比电信号的大,声音信号对异常

的敏感程度显著高于电信号。

层间损失对比试验以第1层数据训练得出的模型分别与第2层(正常类型)、第6层(正常类型)的数据进行损失对比,观察声音信号和电信号各自的损失曲线走向,如图8和9所示。试验数据显示,声音信号的第1层数据与第2层数据每次迭代的损失值相差极小,与第6层数据每次迭代的损失值相差较大;电信号无论是第2层数据还是第6层数据,与第1层数据每次迭代的损失值差异相近。由试验结果分析可知,声音信号受层高的影响较大,而电信号则不受层高的影响。

2.3 多方法对比试验

通过对比试验验证本文所提方法的有效性,试验从跨模态关系数据融合与传统特征层融合,单一声音或电信号与两者融合,是否引入RELA模块以及有部分当前主流的时间序列分析模型的性能差异多方面进行比较,所有试验均采用相同的数据集和参数设置。训练集8个正常类型层每层有95个样本,共760个样

本；验证集与测试集为同一批数据，8个正常类型层每层有31个样本，4个异常类型层每层有33个样本，总共380个样本，所有样本声音信号

数据与电信号数据一一对应。判断异常的阈值通过计算训练集的损失平均值来决定，评价指标包括精确率 (Precision)、召回率 (Recall) 和 $F1$ -

score, 分别用于衡量模型预测异常样本的准确性, 对真实异常样本的识别能力, 以及二者的综合平衡, 计算方式如下。

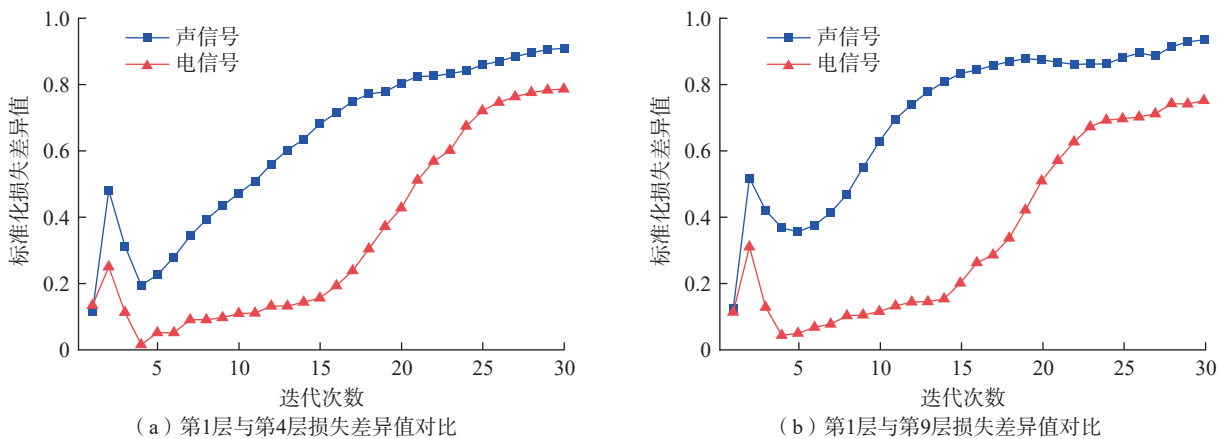


图7 声电信号数据不同层标准化损失差异值对比

Fig.7 Comparison of normalized loss differences across different layers for acoustic and electrical signal data

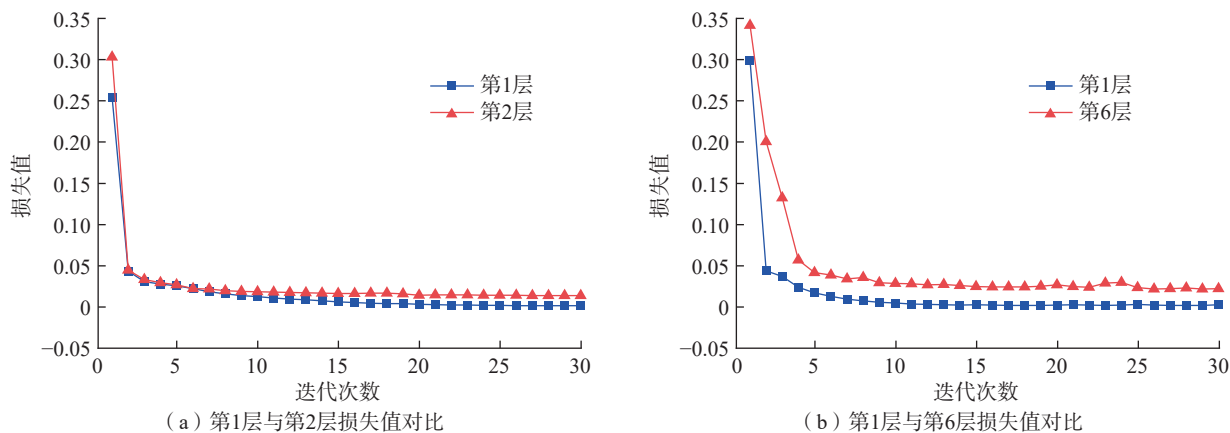


图8 声音信号数据不同层损失曲线走向

Fig.8 Trends of loss curves across different layers for acoustic signal data

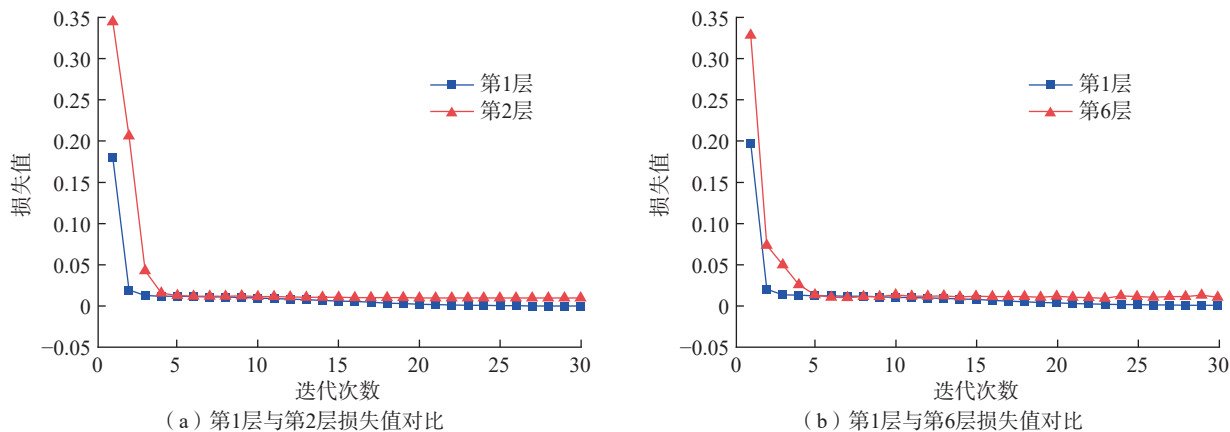


图9 电信号数据不同层损失曲线走向

Fig.9 Trends of loss curves across different layers for electrical signal data

$$\text{Precision} = \frac{TP}{TP + FP} \quad (20)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (21)$$

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (22)$$

式中, TP 代表实际异常且检测为异常的数据; FP 代表实际正常但检测为异常的数据; FN 代表实际为异常但检测为正常的数据。最终基于 IRD-TCN 模型的不同方法对比试验结果如表 2 所示,不同检测模型对比试验结果如表 3 所示。

表 2 结果显示,单一声音信号检测的 $F1\text{-score}$ 、精确率、召回率均达到 90% 以上且均比单一电信号检测的要高,声音信号所包含的更大的信息量对检测效果起到更优先的积极作用。特征层融合的 $F1\text{-score}$ 提升至 93.75% 左右,精确率与召回率相差约 5%,说明特征层融合虽能增强特征判别力,但检测的鲁棒性仍然不足。而跨模态声电信号融合机制是性能提

升的关键因素,声音信号对电弧外部扰动高度敏感,电信号则反映内部能量波动与电弧稳定性。跨模态关系融合检测使模型在声电信号间建立动态对应的时序关联,显著提升了检测鲁棒性,使 $F1\text{-score}$ 能进一步提升至 94.90% 左右,更能有效识别异常。加入 RELA 模块后, $F1\text{-score}$ 达到 98.10% 左右,精确率与召回率差异较小且均接近 98%。显然,本文方法拥有高检测精度,同时具备较为优秀的检测鲁棒性。

由表 3 对比不同模型的结果可知, CAE、GAN、LSTM、Transformer 模型的 $F1\text{-score}$ 均低于 93%,且精确率与召回率相差较大,差值分别约为 6.2%、5.6%、4%、3.1%,在复杂时序建模上存在局限。相比之下, IRD-TCN 能够在时间序列建模中兼顾局部瞬态特征与全局时序依赖,其多层深度可分离卷积能有效捕获电弧声电信号在短时间尺度下的突变模式,残差结构在特征传播中保持梯度稳定,避免信息衰减。RELA 模块则通

过引入轻量注意力机制,在时间维度上强化跨时序特征间的关联,使模型具备持续跟踪异常趋势的能力。最终本研究检测方法 IRD-TCN 模型表现最优, $F1\text{-score}$ 达到 98.10% 左右,兼具全局依赖建模与局部特征保持能力,在精度与鲁棒性上均优于其他模型。

3 结论

(1) WAAM 过程所产生的声音信号对焊道异常的发生更为敏感,通过模型分析后,其正常信号与异常信号的损失标准化差异值显著大于电信号的对应差异值。

(2) 声音信号虽然对异常情况的检测敏感度更高,但是易受沉积层堆叠起来后回声声波的路径距离与方向的改变所影响。通过第 1 层所采集到的数据训练得到的模型对第 2 层和第 6 层的数据进行测试试验,声音信号第 1 层损失值持续下降并最终趋近于 0.001,第 2 层损失值趋近于 0.013,第 6 层损失值趋近于

表 2 基于 IRD-TCN 模型的不同方法对比试验结果

Table 2 Comparison results of different methods based on IRD-TCN model

序号	$F1\text{-score}/\%$	精确率/ $\%$	召回率/ $\%$	融合类型	数据类型	RELA
1	90.49 ± 0.35	90.84 ± 0.70	90.15 ± 0.76	无融合	声音数据	有
2	89.73 ± 0.76	90.08 ± 0.76	89.39 ± 0.76	无融合	电数据	有
3	93.75 ± 0.42	95.38 ± 0.74	90.15 ± 0.76	特征层	声电数据	有
4	94.90 ± 0.78	96.69 ± 0.87	91.67 ± 0.76	跨模态	声电数据	无
5	98.10 ± 0.38	98.37 ± 0.11	97.73 ± 0.75	跨模态	声电数据	有

表 3 不同检测模型对比试验结果

Table 3 Comparison results of different detection models

模型	$F1\text{-score}/\%$	精确率/ $\%$	召回率/ $\%$	融合类型	数据类型
CAE	90.57 ± 0.49	87.70 ± 1.15	93.94 ± 0.76	跨模态	声电数据
GAN	92.91 ± 0.42	94.96 ± 0.24	89.39 ± 0.76	跨模态	声电数据
LSTM	90.62 ± 1.21	92.68 ± 0.87	88.64 ± 0.76	跨模态	声电数据
Transformer	92.86 ± 0.84	94.78 ± 1.74	91.67 ± 0.75	跨模态	声电数据
IRD-TCN	98.10 ± 0.38	98.37 ± 0.11	97.73 ± 0.75	跨模态	声电数据

0.022, 损失值随着层高而逐步增加; 电信号第 1 层损失值趋近于 0.0005, 第 2 层损失值与第 6 层损失值均趋近于 0.01。因此电信号相比于声音信号, 基本不受层高造成的回声环境改变影响, 应用于 WAAM 异常检测具有相对稳定的优势。由于声音信号和电信号在异常检测方面均存在各自的不足之处, 因此需要将二者进行融合实现更高效、更准确地检测效果。

(3) 针对单一声音信号或者电信号的模型训练, 即便在 IRD-TCN 的基础上加入了 RELA 模块, $F1$ -score 分别只有 90.49% 和 89.73% 左右; 特征层融合方式同样结合了 RELA 模块与 IRD-TCN 进行检测, $F1$ -score 提升至 93.75% 左右; 在跨模态数据融合的基础上, 以 IRD-TCN 是否加入 RELA 模块进行对比, 未结合 RELA 模块与结合 RELA 模块检测结果的 $F1$ -score 分别为 94.90% 和 98.10% 左右。当前其他主流的时序分析模型在本文试验中的 $F1$ -score 多集中在 90%~93%, 且精确率与召回率存在明显差异, 准确性与鲁棒性均存在不足。通过分析, 本研究方法在 WAAM 异常检测的应用中具有理想的检测效果。

(4) 本研究为 WAAM 领域提供了一套实时异常检测方法与技术方案, 该方法的应用可有效降低航空领域零部件制造中的不合格率。未来研究可进一步探索实时异常检测与实时控制系统的融合路径, 以实现 WAAM 过程的在线缺陷预警与自适应优化调控, 进而从源头上降低缺陷的产生概率。

参考文献

[1] PANT H, ARORA A, GOPAKUMAR G S, et al. Applications of wire arc additive manufacturing (WAAM) for aerospace component manufacturing[J]. *The International Journal of Advanced Manufacturing Technology*, 2023,

127(11): 4995–5011.

[2] MATTERA G, POLDEN J, CAGGIANO A, et al. Anomaly detection of wire arc additively manufactured parts via surface tension transfer through unsupervised machine learning techniques[J]. *Procedia CIRP*, 2024, 126: 686–691.

[3] SUROVI N A, SOH G S. Acoustic feature based geometric defect identification in wire arc additive manufacturing[J]. *Virtual and Physical Prototyping*, 2023, 18(1): e2210553.

[4] 马明, 郭鑫鑫, 魏正英. 基于深度学习的钛合金电弧增材制造表面形貌缺陷分类识别与预测[J]. *材料导报*, 2025, 39(13): 234–240.

MA Ming, GUO Xinxin, WEI Zhengying. Classification and prediction of surface defects in titanium alloy wire arc additive manufacturing (WAAM) based on deep learning[J]. *Materials Reports*, 2025, 39(13): 234–240.

[5] MU H C, HE F Y, YUAN L, et al. A digital shadow approach for enhancing process monitoring in wire arc additive manufacturing using sensor fusion[J]. *Journal of Industrial Information Integration*, 2024, 40: 100609.

[6] JIRAPIATTANAPORN P, LAWANONT W. Development of anomaly detection model for welding classification using arc sound[C]//2022 14th International Conference on Knowledge and Smart Technology (KST). Piscataway, NJ: IEEE, 2022: 57–62.

[7] MATTERA G. Defect monitoring in wire arc additive manufacturing using frequency domain analysis[J]. *Materials Research Proceedings*, 2024: 50–59.

[8] ZHANG H, WU Q R, TANG W L, et al. Acoustic signal-based defect identification for directed energy deposition-arc using wavelet time-frequency diagrams[J]. *Sensors*, 2024, 24(13): 135–141.

[9] BEVANS B, RAMALHO A, SMOQI Z, et al. Monitoring and flaw detection during wire-based directed energy deposition using in-situ acoustic sensing and wavelet graph signal analysis[J]. *Materials & Design*, 2023, 225: 111480.

[10] GAJAH, LIU F. Defects monitoring of laser metal deposition using acoustic emission sensor[J]. *The International Journal of Advanced Manufacturing Technology*, 2017, 90(1): 561–574.

[11] 田根, 朱甫宏, 王文宇, 等. 基于机

器学习的传感器监测在金属激光增材制造中的应用[J]. *材料导报*, 2025, 39(2): 175–190.

TIAN Gen, ZHU Fuhong, WANG Wenyu, et al. Application of sensor monitoring based on machine learning in metal laser additive manufacturing[J]. *Materials Reports*, 2025, 39(2): 175–190.

[12] GUO Y B, ZHANG Y M, PAN Z X, et al. Recent progress of sensing and machine learning technologies for process monitoring and defects detection in wire arc additive manufacturing[J]. *Journal of Manufacturing Processes*, 2024, 125: 489–511.

[13] LUO D H, WANG X. ModernTCN: A modern pure convolution structure for general time series analysis[C]//International Conference on Learning Representations. New York, 2024

[14] CHEN X W, QIU P J, ZHU W H, et al. TimeMIL: Advancing multivariate time series classification via a time-aware multiple instance learning[EB/OL]. (2024-05-12)[2025-10-21]. <https://arxiv.org/abs/2405.03140>.

[15] 李嘉源, 王晓东, 何启学. CNN增强型Informer模型在工业时间序列预测中的应用及性能优化[J]. *计算机应用*, 2024, 44(S2): 79–83.

LI Jiayuan, WANG Xiaodong, HE Qixue. Application and performance optimization of CNN enhanced Informer model in industrial time series prediction[J]. *Journal of Computer Applications*, 2024, 44(S2): 79–83.

[16] HEWAGE P, BEHERA A, TROVATI M, et al. Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station[J]. *Soft Computing*, 2020, 24(21): 16453–16482.

[17] HAUSER T, REISCH R T, KAMPS T, et al. Acoustic emissions in directed energy deposition processes[J]. *The International Journal of Advanced Manufacturing Technology*, 2022, 119(5): 3517–3532.

[18] ZHAO S P, QIU X J, BURNETT I, et al. A lumped-parameter model for sound generation in gas metal arc welding[J]. *Mechanical Systems and Signal Processing*, 2021, 147: 107085.

[19] TAYFUR S, ALVER N, TAYFUR S, et al. Attenuation and frequency characteristics of acoustic waves in steel and synthetic fiber-reinforced concrete: 3D-PCT and unsupervised pattern recognition[J]. *Applied Sciences*, 2022, 12(24): 181–192.

Anomaly Detection Method for Wire Arc Additive Manufacturing Based on Deep Learning and Cross-Modal Fusion of Acoustic and Electrical Signals

WANG Hongmin¹, QIN Cai², LIN Wei², GUO Huilin², PAN Zengxi^{2,3}, SONG Yingying⁴

(1. School of Emergency Technology and Management, Wuyi University, Jiangmen 529020, China;

2. School of Mechanical and Automation Engineering, Wuyi University, Jiangmen 529020, China;

3. School of Mechanical, Materials, Mechatronics and Biomedical Engineering, University of Wollongong, Wollongong 2522, Australia;

4. School of Electronics and Information Engineering, Wuyi University, Jiangmen 529020, China)

[ABSTRACT] Wire arc additive manufacturing (WAAM) offers advantages such as rapid fabrication and compatibility with lightweight structural design, and thus exhibits great potential for manufacturing critical aerospace components. To address the various defects that may arise during the WAAM process, an unsupervised cross-modal anomaly detection method is introduced, integrating an improved residual deep temporal convolutional network (IRD-TCN) with a residual-enhanced lightweight attention (RELA) module for the fusion of acoustic and electrical signals. Owing to the complexity of WAAM operating conditions, the detection capability of a single sensing modality is limited. Therefore, wavelet-based features from acoustic signals are fused with electrical signal features through cross-modal relational analysis, enabling IRD-TCN and RELA to identify changes in the correlation between the two sensing signals and achieve high-quality detection. Experimental results indicate that the proposed approach attains a precision of 98.37%, a recall of 97.73%, and an $F1$ -score of 98.10%, effectively addressing the limitations of traditional data-fusion methods in terms of accuracy and robustness for anomaly identification in WAAM processes.

Keywords: Wire arc additive manufacturing (WAAM); Cross-modal fusion; Deep learning; Anomaly detection; Temporal convolution; Attention mechanism; Wavelet transform

(责编 七七)