

DOI: 10.19666/j.rlfed.202412263

基于 YOLOv8 优化注意力机制的汽轮机 转子滑动轴承振动幅值预测方法

李亚超¹, 刘昊煜², 徐皓康², 官宇晗², 祁湛桐², 顾煜炯²

(1. 国家能源集团花园电厂, 新疆 哈密 839000;

2. 华北电力大学能源动力与机械工程学院, 北京 102206)

[摘要] 滑动轴承的早期故障具有隐秘性, 为了准确预测其振动幅值, 提出了一种结合 YOLOv8 优化的 CBAM (convolutional block attention module) 的深度学习模型, 在 Backbone 和 Neck 之间嵌入 CBAM 模块以提升模型对重要振动特征的关注度, 同时采用改进的完全交并比损失函数, 提高目标检测精度。同时, 考虑到振动数据的非线性、非稳态特性, 在模型中添加经验模态分解 (empirical mode decomposition, EMD) 方法对振动状态数据进行预测, 以提高预测的准确性。结果表明: 该方法在 600 MW 汽轮机运行数据集上相较于传统 YOLOv8 和 YOLOv7, 在目标检测准确率上分别提升 2.85 个百分点和 8.50 个百分点, 均方根误差和平均绝对误差均有所减少; 此外, 在高噪声环境下, 该模型的误差波动较传统方法降低 30%, 表现出更强的泛化能力和稳定性。

[关键词] 注意力机制; 汽轮机振动; YOLO; 经验模态分解

[引用本文格式] 李亚超, 刘昊煜, 徐皓康, 等. 基于 YOLOv8 优化注意力机制的汽轮机转子滑动轴承振动幅值预测方法[J]. 热力发电, 2025, 54(5): 122-131. LI Yachao, LIU Haoyu, XU Haokang, et al. Vibration amplitude prediction method for turbine rotor sliding bearing based on YOLOv8 optimized attention mechanism[J]. Thermal Power Generation, 2025, 54(5): 122-131.

Vibration amplitude prediction method for turbine rotor sliding bearing based on YOLOv8 optimized attention mechanism

LI Yachao¹, LIU Haoyu², XU Haokang², GUAN Yuhan², QI Zhantong², GU Yujiong²

(1. Huayuan Power Plant, CHN Energy, Hami 839000, China;

2. Department of Energy, Power and Mechanical Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: The early faults of sliding bearings are highly concealed. To accurately predict their vibration amplitude, a deep learning model incorporating a YOLOv8-optimized CBAM attention mechanism is proposed. The CBAM module is embedded between the Backbone and Neck to enhance the model's focus on critical vibration features. Additionally, an improved complete intersection over union loss function is employed to enhance object detection accuracy. Considering the nonlinear and non-stationary characteristics of vibration data, the empirical mode decomposition (EMD) method is integrated into the model to improve the accuracy of vibration state prediction. The experimental results show that, on the 600 MW steam turbine operation dataset, this method improves the detection accuracy by 2.85 percentage points and 8.50 percentage points compared with that of the conventional YOLOv8 and YOLOv7, respectively. Moreover, the root mean square error (RMSE) is reduced, and the mean absolute error (MAE) decreases. Furthermore, in high-noise environments, the model's error fluctuation reduces by 30% compared with that of the conventional methods, demonstrating stronger generalization ability and stability.

Key words: attention mechanism; turbine vibration; YOLO; empirical mode decomposition

随着现代机械向复杂化、高精度和高速运行方向发展, 轴承的工作条件变得更加苛刻^[1]。由于载荷表

面之间的相对滑动, 滑动轴承在承受径向载荷时可能会出现磨损和退化, 这可能导致轴的偏心过大、液体

收稿日期: 2024-12-09 网络首发日期: 2025-02-20

第一作者简介: 李亚超 (1990), 男, 工程师, 主要研究方向为电厂热工技术, 20015594@ceic.com。

通信作者简介: 顾煜炯 (1968), 博士, 教授, 博士生导师, 主要研究方向为电站大型旋转机械振动监测与故障诊断, gyjncepur@163.com。

泄漏以及内部组件的干涉^[2-3]。当滑动轴承的磨损达到一定程度未及时更换,可能导致其部分或完全失效,从而影响整个设备的安全性和稳定性。而对于汽轮机轴系而言,滑动轴承的磨损失效可能带来难以估量的经济损失和灾难性后果。因此,汽轮机在运行过程中预测滑动轴承的振动幅值具有重要意义^[4]。

在滑动轴承故障诊断中,振动信号分析凭借实时性强、技术成熟和成本偏低的特点,成为广泛应用的检测手段。相较于油液分析和声发射信号分析,振动信号能够通过时域和频域分析提取关键特征信息,采集和处理更加便捷,尤其是在工况复杂和长期运行的条件下^[5]。然而,传统的振动信号分析方法往往依赖于人工经验和先验知识,效率较低且难以精准识别多种复杂故障^[6]。近年来深度学习技术的引入为振动信号的自动化处理和精确预测提供了全新的思路,卷积神经网络(CNN)因其在特征提取方面的优异性能,逐渐成为研究热点。注意力机制作为提升模型性能的重要工具,逐渐被集成到深度学习模型中。CBAM作为一种轻量化的注意力模块,通过结合通道和空间注意力,增强了特征表达能力。文献[7]针对复杂环境下的轴承故障类型和故障程度识别提出了一种浅层多尺度卷积神经网络结合多注意力机制的诊断方法,使用连续小波变换(CWT)将振动信号转换为时间频率图,并通过多尺度卷积层提取全局和局部特征,显著提升了模型对关键分类特征的关注,增强了弱信号标签的识别率。文献[8]对滚动轴承振动信号提出了一种结合注意力机制的CNN以增强特征提取能力,同时引入了梯度加权类激活映射(Grad-CAM)模块,为模型生成时间频率图的可视化热图,提升模型可解释性,实现了99.81%的故障分类准确率。吴静然等^[9]聚焦复杂噪声环境和多模态信号特点下的旋转机械故障识别,通过引入注意力机制对多尺度特征进行加权融合,自适应地为不同尺度特征分配权重,该方法在低信噪比环境下也表现出了较强的鲁棒性,结果表明其在分类精度上具有显著优势。

在目标检测领域,YOLO系列模型凭借其高效性和精确性广泛应用于工业场景。其中,最新版本YOLOv8通过优化网络结构和多尺度特征提取能力,进一步提升了检测性能^[10]。文献[11]通过在YOLOv8模型中嵌入多种注意力机制,构建深度可变形卷积特征金字塔网络(TDSFPN),提高多尺度特征的融合和检测能力等方法对滚动轴承的表面缺陷检测进行多目标分类和定位,解决了多目标场

景中的高目标遗漏率问题,优化了特征融合与分类性能。文献[12]针对风力涡轮机叶片缺陷检测,提出了一种基于进化变分的YOLOv8模型(EV-YOLOv8),通过引入FasterNet模块和EfficCIoU(Efficient Complete IoU)损失函数,提升了模型的检测效率和资源使用率。文献[13]提出了改进的实时轴承缺陷检测算法CAC-YOLOv8,利用通道衰减网络(CAN)和复合池化空间金字塔快速模块,优化了特征提取和深度融合能力,在实时性和检测精度之间达到了良好平衡。文献[14]针对化工设备中的轴承缺陷检测,研究提出了YOLOv8-LMG模型,通过整合VanillaNet主干网络、Lion优化器、CFP-EVC模块和Shape-IoU损失函数,有效提高了复杂缺陷的检测能力,在保持低计算复杂度的同时显著提升了检测精度和效率。对于信号的降噪处理方面,王普等^[15]提出的小波阈值降噪方法在处理故障信号时有效减少了噪声,提升了信号的可检测性。文献[16]则通过小波变换方法进一步提高了故障诊断的准确性。

目前,将深度学习应用于滑动轴承故障诊断时,由于当前轴承故障相关数据相对匮乏,难以满足模型训练的要求。其次,现有研究多集中于单一对象、单一条件的故障诊断,缺乏能够适应多种运行条件、转速和多故障场景的综合诊断模型框架。同时,随着信号采集技术的进步,数据量快速增长,传统深度学习方法在处理大规模数据时可能效率低下,影响诊断效果^[17-19]。因此,本文提出了一种结合CBAM模块优化YOLOv8的深度学习模型。YOLOv8作为高效目标检测算法,通过优化后的CBAM注意力机制,可更精准地提取故障特征,增强了模型在复杂工况下的鲁棒性。该模型通过关注信号中关键区域,提升了特征提取的精度和诊断的稳定性,特别是在高噪声和变负载条件下,提高了诊断的准确性,提供了更加可靠的故障诊断方案。

1 预测方法

1.1 CBAM 模块

滑动轴承通常工作在复杂且恶劣的环境下,伴随着大量的噪声信号,使得微小故障难以通过直接观察发现。为了提高模型对这些故障的检测能力,本文在YOLOv8的Backbone和Neck之间引入了卷积块注意力模块CBAM。

在卷积神经网络中,不同的卷积核对应着不同的通道输出。同时,神经网络提取的特征各不相同,

特征图组中不同特征重要性也不尽相同^[20]。为了解决这一问题，将一个通道注意力模块引入卷积网络，以突出特征图组中各特征的重要性。考虑到特征图本身是一个二维图像，同时将空间注意力模块引入卷积网络，用于确定特征图值得关注部分。

1.1.1 通道注意力模块

该模块的特点是保持通道维度不变并压缩空间维度，重点关注样本被划分为不同类别的原因。通道注意力模块的机理^[21]为：

$$\begin{cases} A = \text{MLP}[\text{Avg P}(F)] \\ M = \text{MLP}[\text{MaxP}(F)] \\ M_c(F) = (A + M) \\ F' = F \otimes M_c(F) \end{cases} \quad (1)$$

式中： F' 为融合了通道注意力权重的特征矩阵； $M_c(F)$ 为通道注意力权重；Avg P(·)为平均池化操作；Max P(·)为最大池化操作；MLP[·]为多层感知器。

1.1.2 空间注意力模块

与前者相比，该模块的特点是保持空间维度不变并压缩通道维度。与通道注意力模块不同，空间注意力模块更关注信息的“位置”。空间注意力模块的机理为：

$$\begin{cases} M(F') = [\text{Avg P}(F'), \text{MaxP}(F')] \\ M_s(F') = \sigma(\text{conv}(M(F'))) \\ F'' = F' \otimes M_s(F') \end{cases} \quad (2)$$

式中： F'' 为融合了空间注意力权重的特征矩阵； $M_s(F')$ 为空间注意力权重；conv 为卷积操作。

注意力机制借鉴了人类感知行为中的资源分配方式，能够对重要信息给予更多关注^[22]，本文采用的 CBAM 结构将通道注意力模块直接与空间注意力模块串联，能够同时关注通道和空间信息，从而选择具有区分性的特征。

1.2 YOLOv8 网络

YOLOv8 模型结构如图 1 所示。

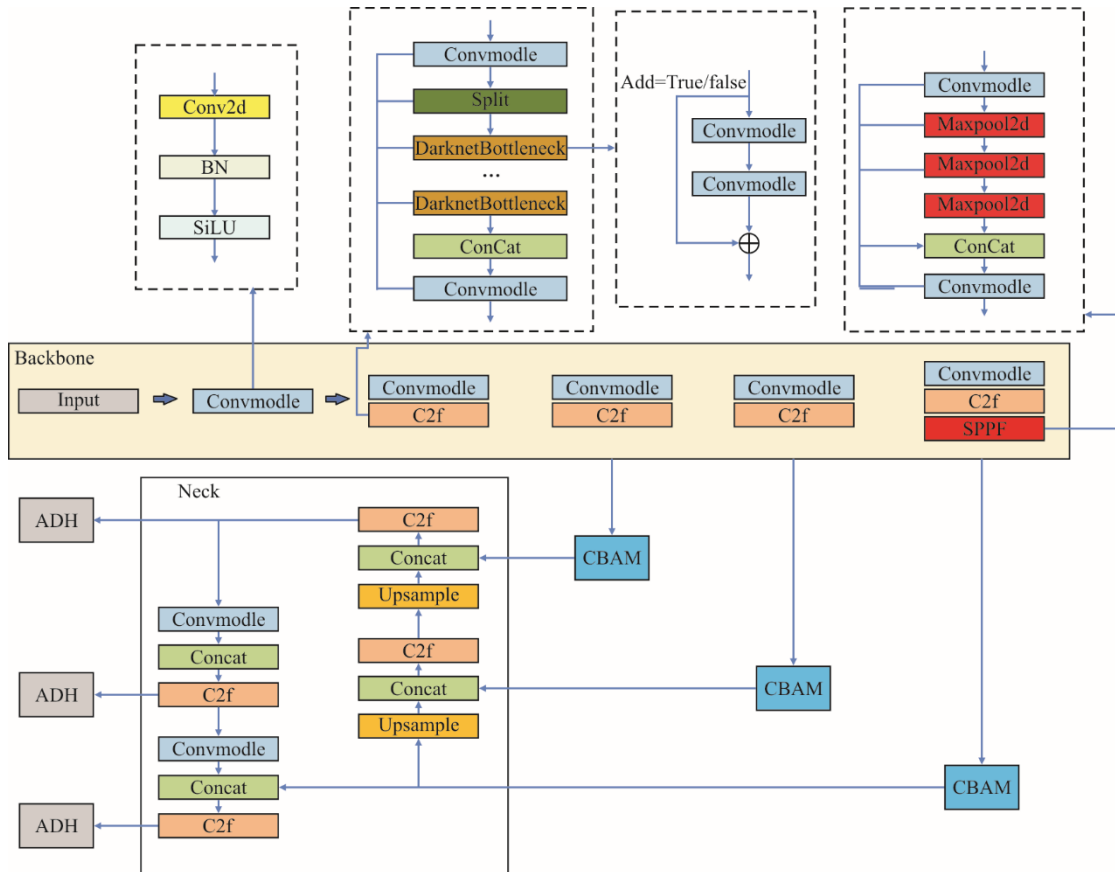


图 1 YOLOv8 模型结构
Fig.1 Structural diagram of the YOLOv8 model

YOLOv8 相较于之前的 YOLO 版本，主要改进了主干网络、特征融合与解耦头的架构设计，首先

延续了 CSP (cross stage partial) 设计理念，优化特征提取过程；使用 C2f 模块替代 YOLOv5 中的 C3

模块,实现模型轻量化并优化梯度流,并集成 SPPF (spatial pyramid pooling fast) 模块,有效融合局部和全局特征。同时,采用基于 PANet (path aggregation network) 和 FPN (feature pyramid network) 的特征融合方法,提升了多尺度特征的检测能力。在损失函数方面,选择 Varifocal Loss 作为分类损失,提高了分类性能,并且回归损失使用 CIoU 和 DFL,增强边界框回归的精度。此外,通过无锚点 (anchor-free) 设计,提高了检测速度和模型精度,在 COCO 数据集上的表现优于之前的 YOLO 版本^[23-24]。

本文提出了一些针对 YOLOv8 网络的改进,以提高其在滑动轴承故障诊断中的性能。在 Neck 和 Backbone 之间引入了 CBAM 注意力模块,增强了模型对重要区域的关注,提高了特征提取的精度。此外,将原有的检测头替换为非对称多级通道压缩解耦头 (ADH),将分类任务和回归任务分离处理,使模型在处理复杂目标时更加灵活和高效。为了进一步优化模型,本文还将回归分支的损失函数替换为改进的完全交并比 (ECIOU) 损失函数替换为改进的完全交并比 (ECIOU) 损失函数,该方法优化了边缘收敛情况,提高了检测精度。

1.3 快速傅里叶变换

快速傅里叶变换 (FFT) 基于离散傅里叶变换 (DFT),通过利用信号的对称性和周期性特性,采用分治法将 DFT 分解为多个较小的 DFT,从而将计算复杂度从 $O(N^2)$ 降低为 $O(M\log_{10}N)$,大幅提高了计算效率。快速傅里叶变换的显著特点在于其高效性和准确性,尤其在处理大规模数据和实时信号处理时表现优异。通过频谱图,可以直观地看到信号中各频率成分的强度分布,帮助工程师和研究人员分析信号特征,识别和诊断潜在问题。

一组长度为 N 的序列 $x(n_1)$, N 为 2 的幂次数,其经过 DFT,变换为:

$$X(K) = \sum_{n_1=0}^{N-1} x(n_1) W_N^{n_1 k_2} \quad (3)$$

式中: $W_N = e^{-j\frac{2\pi}{N}}$ 为旋转因子; j 为虚数符号; $X(K)$ 为 $x(n_1)$ 对应的 N 个离散频率点的相对幅值; k_2 为频率索引。

因 DFT 的计算量与 N^2 成正比,故当序列较长时,为了减少 DFT 的计算量,可考虑将长序列 $x(n_1)$ 的 DFT 分解成若干短序列的 DFT 之和。考虑到式(3)中 W_N^k 的周期性、对称性和可约性,将序列

$x(n_1)$ 按照式(3)分成 2 组,则可推出由 2 个序列长度为 $N/2$ 的 DFT 表示的式(1),结果为:

$$\begin{cases} W_N^{-n_1 k_2} = W_N^{(N-n_1)k_2} = W_N^{n_1(N-k_2)} \\ W_N^{k_2 + \frac{N}{2}} = -W_N^{k_2} \\ W_N^{\frac{N}{2}} = -1 \\ W_N^0 = 0 \end{cases} \quad (4)$$

$$\begin{cases} x_1(u) = x(2u) \\ x_2(u) = x(2u+1) \end{cases} \quad (5)$$

式中: $u=0, 1, \dots, \frac{N}{2}-1$ 。

$$\begin{aligned} X(K) &= \text{DFT}[x(n_1)] = \text{DFT}[x(2u)] + \text{DFT}[x(2u+1)] = \\ &= \sum_{u=0}^{\frac{N}{2}-1} x_1(u) W_{\frac{N}{2}}^{uk_2} + W_N^{k_2} \sum_{u=0}^{\frac{N}{2}-1} x_2(u) W_{\frac{N}{2}}^{uk_2} \end{aligned} \quad (6)$$

该方法显著减少了计算量,使其 $M\log_{10}N$ 与成正比,而非原始 DFT 的 N^2 。这种分解方式称为按时间抽取的 FFT 算法,通过逐步将输入序列按偶数和奇数分解,最终计算出每个频率分量的幅度和相位。

在实际应用中,FFT 通过将输出复数数组的索引映射到对应频率,得到各频率分量的实际频率:

$$f_k = \frac{K_3 f_T}{N} \quad (7)$$

式中: f_k 为频率成分的实际频率; f_T 为采样频率; K_3 为 FFT 输出数组的索引。

利用频谱图查看不同频率的强度,分析 FFT 输出,找到主要频率成分,从而计算各序列的周期。图 2 为由 FFT 得到的汽轮机振幅频域与时域。

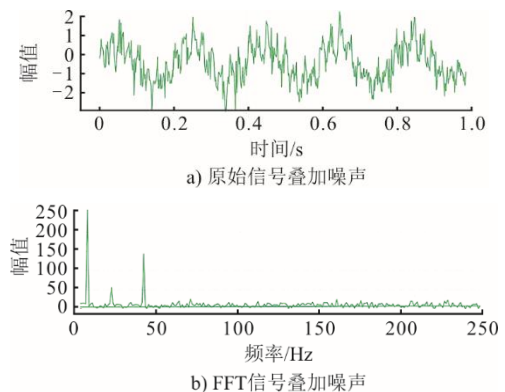


图 2 汽轮机振幅频域与时域

Fig.2 Frequency domain and time domain of the turbine vibration amplitude

1.4 经验模态分解

经验模态分解主要适用于分析非线性、非平稳信号,特别是在处理多分量信号和进行时间频率分

析时表现出色。它无需假设信号是线性和平稳的,因此在机械故障诊断领域被广泛应用^[25]。EMD算法的原理是将原始信号分解为多个本征模态函数进行处理,算法步骤如下。

1) 通过三次样条函数对原始信号 $X(t)$ 进行分析。首先,确定其极大值和极小值;然后,通过这些极值点拟合得到信号的上包络线 $X_{\text{MAX}}(t)$ 和下包络线 $X_{\text{MIN}}(t)$,从而有效提取信号特征,辅助故障诊断和状态评估。

2) 对上、下包络线求平均值。

$$M_1(t) = \frac{X_{\text{MAX}}(t) + X_{\text{MIN}}(t)}{2} \quad (8)$$

3) 将 $X(t)$ 与 $M_1(t)$ 相减得到余下信号 $d_1(t)$ 。

$$d_1(t) = X(t) - M_1(t) \quad (9)$$

4) 对余下信号 $d_1(t)$ 重复上述过程,直到满足模态分量的条件 SD 小于门限值才停止,得到第 1 个 IMF 分量余下信号 $c_1(t)$ 。SD 条件为:

$$\text{SD} = \sum_{i=0}^T \frac{|d_{k-1}(t) + d_k(t)|^2}{d_{k-1}^2(t)} \quad (10)$$

5) 对信号 $X(t)$ 与 $c_1(t)$ 求差,得到第 1 阶残差量 $r_1(t)$,将 $r_1(t)$ 代替原始信号 $X(t)$ 进行以上步骤的操作,重复 n 次后可以得到第 n 阶模态分量 $c_n(t)$ 和 n 阶的残差 $r_n(t)$ 。

$$\begin{cases} r_1(t) - c_2(t) = r_2(t) \\ \dots \\ r_{n-1}(t) - c_n(t) = r_n(t) \end{cases} \quad (11)$$

因此原始信号 EMD 的分解表达式为:

$$X(t) = \sum_{i=1}^N c_i(t) + r_n(t) \quad (12)$$

1.5 评价指标

在评估深度学习模型性能时,主要使用 4 个关键的指标。

1) 平均绝对百分比误差 δ_{MAPE} ,通过计算预测误差相对于实际值的百分比,提供了误差的相对大小。它以百分比形式表示,便于理解和比较,不同数据规模下都能使用。

2) 均方根误差 δ_{RMSE} ,其与模型的预测精度成反比关系,用于衡量预测值与实际值之间的差异。 δ_{RMSE} 对异常值比较敏感,因为它会放大较大的误差,因此在异常值存在的情况下, δ_{RMSE} 会较大。

3) 平均绝对误差 δ_{MAE} ,是预测值与实际值之间差异的绝对值的平均值。它对异常值的敏感度较低,因此是一种更加稳健的误差度量方式,数值越小表示模型的预测精度越高。

4) 决定系数 R^2 ,其表示模型对数据中变量关系的解释程度,其值范围为 0~1。数据与模型的拟合程度越高,其值越接近 1。

本文旨在用不同模型之间的比较,选择出最佳模型应用于实际场景。通过使用这些指标,可以确保模型在复杂工况下仍然具有良好的预测性能和可靠性,为工程应用提供有力支持。

$$\delta_{\text{MAPE}} = \frac{1}{N} \sum_{k=1}^N \frac{|P_k - O_k|}{O_k} \times 100\% \quad (13)$$

$$\delta_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{k=1}^N (P_k - O_k)^2} \quad (14)$$

$$\delta_{\text{MAE}} = \frac{1}{N} \sum_{k=1}^N |P_k - O_k| \quad (15)$$

$$R^2 = 1 - \frac{\sum_{k=1}^n (P_k - O_k)^2}{\sum_{k=1}^n (P_k - \bar{O}_k)^2} \quad (16)$$

式中: N 、 P_k 、 O_k 、 \bar{O}_k 分别为数据总数、第 k 个预测值、第 k 个实验值和所有实验值的平均值。

2 实验结果和评价

2.1 数据集选取

为验证所提模型的预测能力,选用中国西北某 600 MW 电厂 2023 年 12 月 15 日至 18 日的主要旋转设备实际运行数据进行测试。该数据集以 30 s 为 1 个采样间隔,共收集 10 000 组历史数据。在这些数据中,前 8 000 组数据作为训练集来训练深度学习模型,剩余 2 000 组数据则作为测试集,用以评估模型的预测性能。

实验环境搭建使用 Python3.8,代码使用深度学习框架 pytorch 编写。

2.2 实例分析

采用了一种深度学习框架来提升模型性能。首先,通过小波变化等数据预处理方法对电厂数据进行预处理,从基础数据集中划分出训练集和测试集,以评估模型的泛化能力。数据标注格式被调整为 YOLO 格式,并在训练前进行了数据增强处理。具体方法包括随机选择 4 张图片进行翻转、缩放和色彩调整等操作,然后合成为 1 张新的复合图像,有效降低了过拟合的风险。

图 3 为平滑处理后的电厂振动信号波形。由图 3 可见,图 3 显示了平滑处理后的电厂振动信号波形图。应用信号处理方法(如滤波或平滑算法)

有效去除了噪声成分，显著改善了信号波形的平稳性和清晰度。平滑后的信号保留了关键特征，为后续特征提取和分析提供了便利。这种处理显著提升了信号的可读性和分析准确性，为基于深度学习的故障诊断奠定了可靠的数据基础。

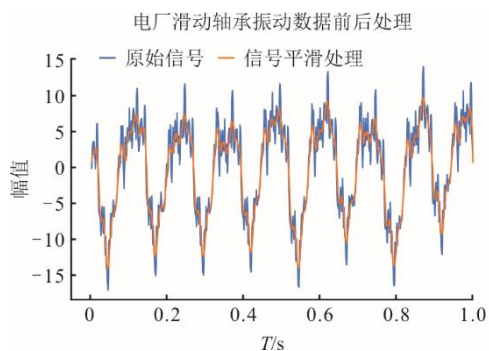


图 3 电厂信号平滑处理后波形
Fig.3 The smooth processed waveforms of power plant signal

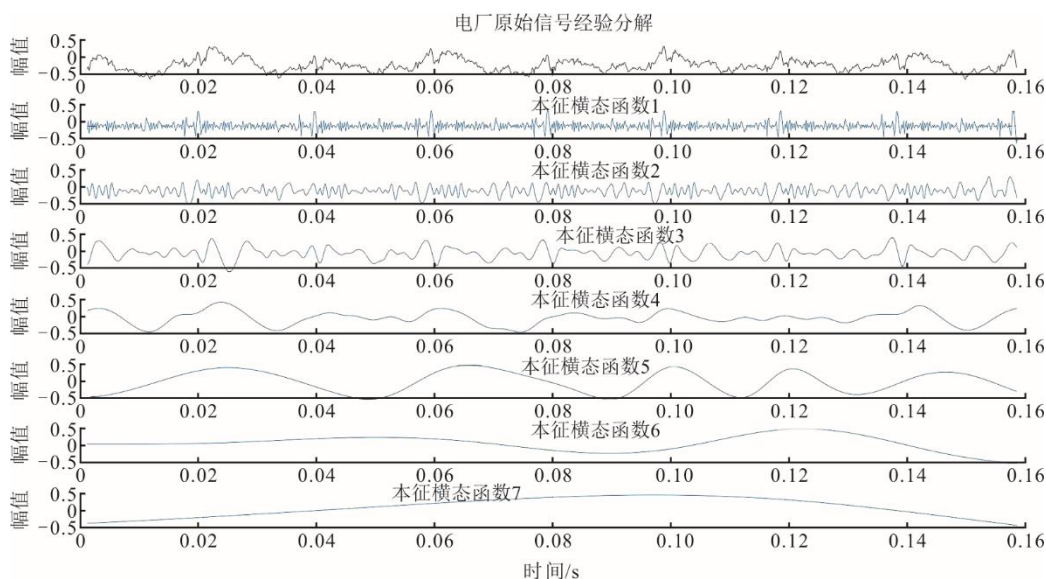


图 4 电厂信号经验模态分解结果
Fig.4 The empirical modal decomposition result of power plant signals

图 5 为振动信号频谱。振动信号的频谱揭示了主要频率成分及其幅值特征。通过采集电厂设备振动数据并进行快速傅里叶变换处理，可在频谱图中观察到显著频率峰值，这些峰值通常与设备的运行状态或潜在故障相关。在典型频谱图中，较高幅值的频率，如一倍频与二倍频可能对应设备的基本运行频率及其谐波成分，而背景噪声反映随机干扰的存在。频谱图分析可定位设备异常点，如频率偏移或异常峰值。图 6 为模型诊断流程，图 7 为模型训练收敛曲线。

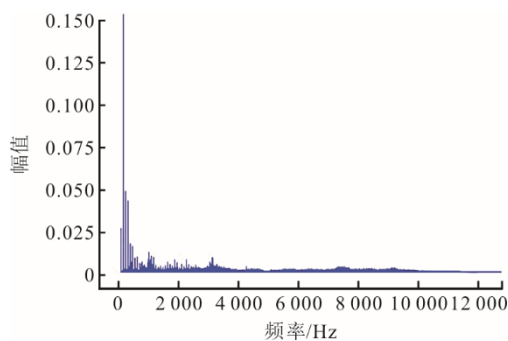


图 5 振动信号频谱
Fig.5 The spectrum of vibration signal

图 4 为电厂信号经验模态分解结果。电厂实地振动信号通过经验模态分解处理，可直观分析信号特征。原始信号包含汽轮机振动通道的复杂动态信息，包括多频率成分和随机噪声。经 EMD 分解后，信号被分解为一系列具有物理意义的本征模态函数 (IMFs)，这些模态按频率从高到低排列，每个模态捕获特定频率范围的振动特性。

高频 IMFs 通常反映噪声或快速振动变化，可能涉及机械冲击、润滑状态或轻微异常；低频 IMFs 则对应设备主要振动特征，如转轴的基本运行频率和长期趋势信息。该方法不依赖先验模型即可分离信号本质成分，为状态监测和故障诊断提供关键支持。同时，IMFs 的变化趋势直观反映了设备运行状态的动态特性及潜在异常。

通过上述信号预处理方法，有助于减少电厂现场采集测点数据得到的不确定因素如噪声或不平衡，提高模型学习特征的有效性。

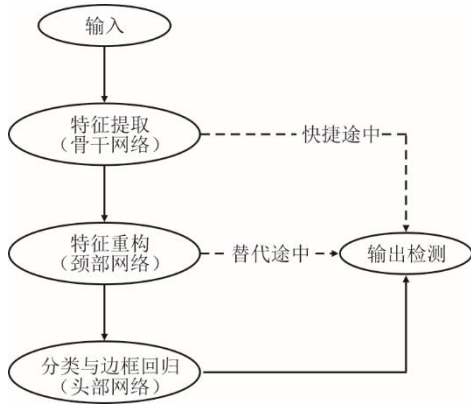


图 6 模型诊断流程
Fig.6 The model diagnostic process

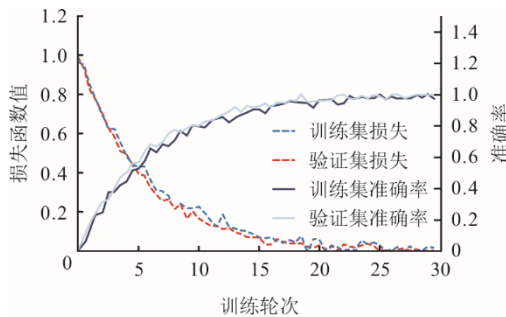


图 7 模型训练收敛曲线
Fig.7 Convergence curves of model training

由图 7 可见，训练和验证损失随训练进程逐渐下降，说明模型能有效学习并减少误差。同时，训练和验证精度逐步提升，展现出模型的学习能力与模式捕捉能力较强。在训练初期，损失值快速下降，精度显著提升；约第 7 次训练后，损失值和精度趋于平稳，表明模型收敛良好。训练损失约在第 7 次后下降至 0.05，验证损失稳定在 0.07，训练精度和验证精度分别达到 0.93 和 0.91，波动极小。

验证损失与验证精度在部分训练周期中略有波动，可能与数据噪声或验证集分布差异有关。此类波动反映模型的泛化能力。结果表明模型在训练中表现出良好的收敛性与鲁棒性，并在验证数据上展现出较强泛化能力。模型约在第 7 次训练后达到较高性能，显示出其高效性，适合快速部署。损失值随着训练推进逐渐下降，精度逐步上升，表明模型能有效拟合数据并提升性能。

模型在训练集与验证集上的表现趋势一致，表明其泛化能力良好。虽然验证损失与精度相对训练数据略有波动，但其波动反映了模型对数据子集的适应性。这表明模型在未见数据上的性能稳定。即使数据中加入随机噪声，模型仍能保持良好收敛性，同时保证训练过程的稳定。适度的噪声也增强了模

型在实际应用中的鲁棒性。

YOLOv8 是一种先进模型，在算法和结构上进行了优化，基于先前 YOLO 版本的成功经验，进一步引入了诸多新功能和改进，使其能够更快速、更准确地进行目标检测和跟踪。其设计目标是实现高效、准确且易于使用的特性，满足不同应用领域的需求。YOLOv8 不仅在目标检测和跟踪任务中表现优异，还适用于实例分割、图像分类和姿态估计等任务。图 8 为不同 YOLO 模型平均预测精度比较。图 8 中： x 轴表示每个模型的参数数量（以百万为单位）。通常，参数越多，模型结构越复杂。 y 轴表示在本实验数据集上的 mAP 值，这是一个评价模型检测准确度的标准。mAP 值越高，模型的检测准确度越好。

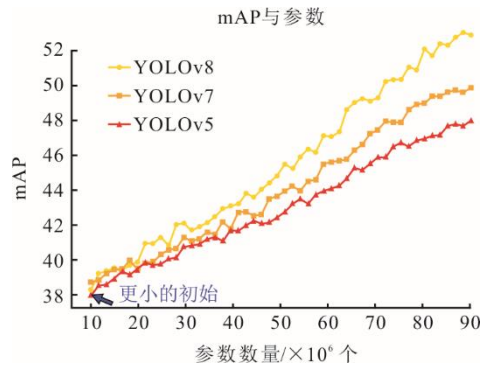


图 8 不同 YOLO 模型平均预测精度比较
Fig.8 The average prediction accuracy of different models

本实验评估了不同模型在半精度优化 (FP16) 条件下的精度和响应速度，重点处理图像。实验在 NVIDIA A100 GPU 和 TensorRT 平台上进行，通过加载不同复杂程度的模型，对其在本实验数据集上的 mAP (平均精度均值) 和处理延迟 (每张图像的平均处理时间) 进行了对比分析，结果如图 9 所示，实验评估了不同复杂度模型在高效性和精度间的平衡。

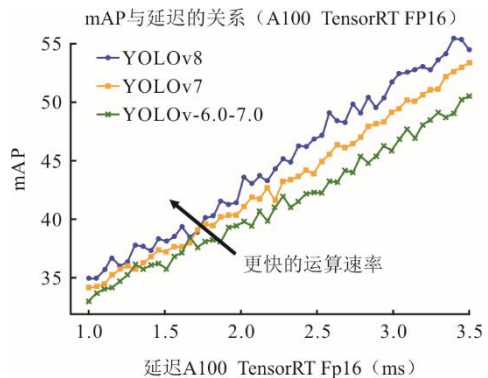


图 9 半精度优化下不同模型的精度比较
Fig.9 The accuracy of different models under half-precision optimization

x 轴表示在使用 NVIDIA A100 GPU 和 TensorRT 进行半精度 (FP16) 优化时, 模型处理每张图像的平均延迟 (以毫秒每张图像计)。延迟越低, 模型响应越快。y 轴表示在数据集上的 mAP (平均精度均值) 这是一个衡量模型准确性的指标。它是模型识别图像中物体并准确标记其边界框的能力的综合评分。分数越高, 模型的准确性越好。

在测试阶段, 利用测试集计算每个数据点的残差, 统计不同残差区间内的频次, 最终绘制直方图展示残差分布 (图 10)。直方图以 x 轴表示残差值 (单位与振动幅值一致), y 轴表示每个残差区间内的频率 (出现次数)。

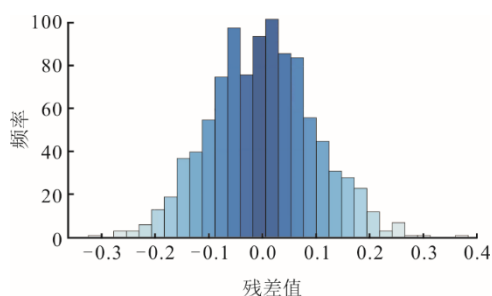


图 10 模型实验结果残差直方图

Fig.10 Residuals histogram of model experimental results

由图 10 可见, 改进后的模型在振动幅值预测方面表现优异。绝大部分残差集中于-0.1~0.1, 表明模型预测误差小, 整体精度高。残差分布的中心

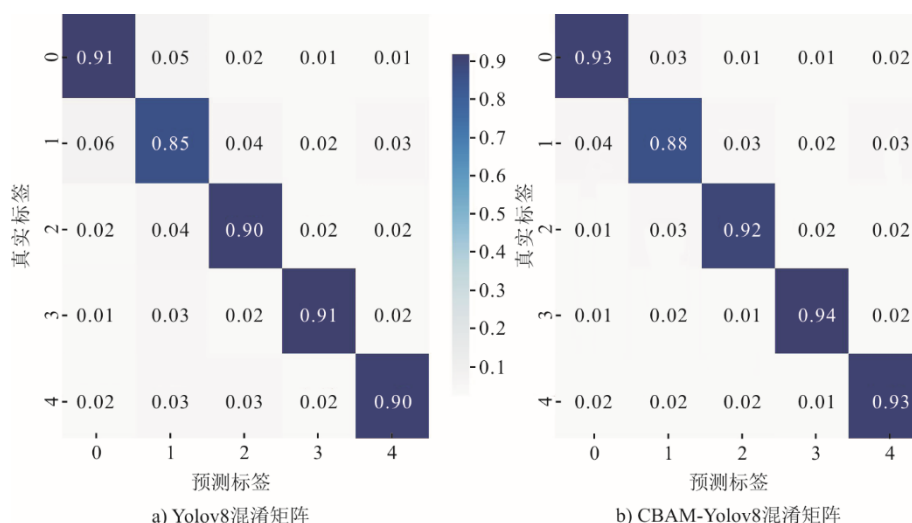


图 11 添加注意机制的混淆矩阵比较

Fig.11 The confusion matrices with attention mechanism

总体来看, CBAM-YOLOv8 模型在所有类别上的预测准确率显著提升, 误分类率大幅下降, 尤其是类别 1 与类别 2 之间的误分类明显减少。这表明,

区域呈现明显的正态分布特性, 进一步证明了模型预测结果的稳定性和无偏性。与传统模型相比, 改进模型在预测能力上更为出色, 表现为残差分布范围更窄, 误差更低。实验结果充分说明, 该模型在复杂工况和大规模数据处理中的准确性和鲁棒性显著提升, 为实际应用提供了可靠支持。

图 11 对比了引入 CBAM 注意力机制后的 CBAM-YOLOv8 模型与原始 YOLOv8 模型的混淆矩阵。

由图 11 可见, 原始 YOLOv8 模型在类别 1 上的预测准确率为 0.91, 误分类至类别 2 和类别 3 的概率分别为 0.05 和 0.02, 其余类别误分类率较低; 类别 2 的预测准确率为 0.85, 误分类至类别 1 的概率为 0.06, 其余类别误分类率较低; 类别 3 的预测准确率为 0.90, 误分类率较低; 类别 4 和类别 5 的预测准确率分别为 0.91 和 0.90, 误分类率均较低。

添加 CBAM 模块后, CBAM-YOLOv8 模型在类别 1 上的预测准确率提升至 0.93, 误分类至类别 2 的概率降至 0.03, 其余类别误分类率进一步减少; 类别 2 的预测准确率升至 0.88, 误分类至类别 1 的概率降至 0.04, 其余类别的误分类率进一步降低; 类别 3 和类别 4 的预测准确率分别提升至 0.92 和 0.94, 误分类率几乎为 0; 类别 5 的预测准确率为 0.93, 误分类率保持较低水平。

CBAM 模块能够有效增强模型对关键信息区域的关注, 从而提升多类别分类任务的准确性与鲁棒性。表 1 为各类模型准确性和 mAP50-95/% (平均精度

均值在不同的交并比阈值下的结果) 的比较结果。

表 1 各类模型准确性和 mAP50-95/% 比较
Tab.1 The accuracy and mAP50-95/% of various models

模型	Pre/%	mAP50-95/%
CBAM-YOLOv8	93.96	53.5
YOLOv8	91.11	55.0
YOLOv7	85.46	52.0
YOLOv5-6.0-7.0	89.00	50.0

3 结 语

本文探讨了利用基于注意力机制的 YOLOv8 模型对轴承故障进行诊断的方法。研究通过引入 CBAM 注意力模块, 增强了模型对重要特征的捕捉能力, 从而提高了故障诊断的精度。实验结果表明, 改进后的 CBAM-YOLOv8 模型在准确性上较原始模型提升了 2.85 百分点, 在各类轴承故障的预测中表现出更高的准确性。该方法有效降低了模型的过拟合风险, 并增强了其在复杂工况下的泛化能力。这一成果对工程应用具有重要的实际意义, 为未来在旋转机械故障诊断领域的研究提供了新思路。

[参 考 文 献]

- [1] 李晓博, 舒进, 牛瑞杰, 等. 基于主动电磁控制的滑动轴承-转子系统自激振动抑制数值仿真[J]. 热力发电, 2020, 49(9): 80-86.
LI Xiaobo, SHU Jin, NIU Ruijie, et al. Numerical simulation of self-excited vibration suppression of oil-film bearing-rotor system based on active electro-magnetic control[J]. Thermal Power Generation, 2020, 49(9): 80-86.
- [2] DAI J, TIAN L, CHANG H. An intelligent diagnostic method for wear depth of sliding bearings based on MGCNN[J]. Machines, 2024, 12(4): 266.
- [3] LI Q, ZHANG W, CHEN F, et al. Fault diagnosis of nuclear power plant sliding bearing-rotor systems using deep convolutional generative adversarial networks[J]. Nuclear Engineering and Technology, 2024, 56(8): 2958-2973.
- [4] CHEN F, WANG X, ZHU Y, et al. Time-frequency transformer with shifted windows for journal bearing-rotor systems fault diagnosis under multiple working conditions[J]. Measurement Science and Technology, 2023, 34(8): 085121.
- [5] TANG H, LIAO Z, OZAKI Y, et al. Stepwise intelligent diagnosis method for rotor system with sliding bearing based on statistical filter and stacked auto-encoder[J]. Applied Sciences, 2020, 10(7): 2477.
- [6] SONG L, WANG H, CHEN P. Step-by-step fuzzy diagnosis method for equipment based on symptom extraction and trivalent logic fuzzy diagnosis theory[J]. Ieee Transactions on Fuzzy Systems, 2018, 26(6): 3467-3478.
- [7] HUANG T, FU S, FENG H, et al. Bearing fault diagnosis based on shallow multi-scale convolutional neural network with attention[J]. Energies, 2019, 12(20): 3937.
- [8] LUO J, ZHANG X. Convolutional neural network based on attention mechanism and Bi-LSTM for bearing remaining life prediction[J]. Applied Intelligence, 2022, 52(1): 1076-1091.
- [9] 吴静然, 丁恩杰, 崔冉, 等. 采用多尺度注意力机制的旋转机械故障诊断方法[J]. 西安交通大学学报, 2020, 54(2): 51-58.
WU Jingran, DING Enjie, CUI Ran, et al. A diagnostic approach for rotating machinery using multi-scale feature attention mechanism[J]. Journal of Xi'an Jiaotong University, 2020, 54(2): 51-58.
- [10] 武福, 蒋鹏民, 李忠学, 等. 基于轻量化 YOLOv8s 的轨道扣件状态检测方法[J]. 北京交通大学学报, 2024, 48(5): 59-68.
WU Fu, JIANG Pengmin, LI Zhongxue, et al. Light-weight detection method for track fastener status based on improved YOLOv8s[J]. Journal of Beijing Jiaotong University, 2024, 48(5): 59-68.
- [11] DING P, ZHAN H, YU J, et al. A bearing surface defect detection method based on multi-attention mechanism YOLOv8[J]. Measurement Science and Technology, 2024, 35(8): 086003.
- [12] WANG H, SHEN Q, DAI Q, et al. Evolutionary variational YOLOv8 network for fault detection in wind turbines[J]. Computers Materials & Continua, 2024, 80(1): 625-642.
- [13] LIU B, ZHAO Y, CHEN B, et al. CAC-YOLOv8: real-time bearing defect detection based on channel attenuation and expanded receptive field strategy[J]. Measurement Science and Technology, 2024, 35(9): 096004.
- [14] LIU M, ZHANG M, CHEN X, et al. YOLOv8-LMG: An improved bearing defect detection algorithm based on YOLOv8[J]. Processes, 2024, 12(5): 930.
- [15] 王普, 李天垚, 高学金, 等. 分层自适应小波阈值轴承故障信号降噪方法[J]. 振动工程学报, 2019, 32(3): 548-556.
WANG Pu, LI Tianyao, GAO Xuejin, et al. Bearing fault signal denoising method of hierarchical adaptive wavelet threshold function[J]. Journal of Vibration Engineering, 2019, 32(3): 548-556.
- [16] BAKIR T, BOUSSAID B, HAMDAOUI R, et al. Fault detection in wind turbine system using wavelet transform: Multi-resolution analysis[C]. 2015 IEEE 12th International Multi-Conference on Systems, Signals & Devices (SSD15), Mahdia, Tunisia, 16-19 March 2015.
- [17] HUANG Y, WANG K, DENG Z, et al. Synchronous averaging with sliding narrowband filtering for low-speed bearing fault diagnosis[J]. Journal of Sound and Vibration, 2024, 586: 118503.
- [18] XIN Y, LI S. Novel data-driven short-frequency mutual information entropy threshold filtering and its application to bearing fault diagnosis[J]. Measurement Science and Technology, 2019, 30(11): 115006.
- [19] CHEN J, PAN J, LI Z, et al. Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals[J]. Renewable Energy, 2016, 89: 80-92.
- [20] ZHANG Q, WEI X, WANG Y, et al. Convolutional neural network with attention mechanism and visual vibration signal analysis for bearing fault diagnosis[J].

- Sensors, 2024, 24(6): 1831.
- [21] QIN H, PAN J, LI J, et al. Fault diagnosis method of rolling bearing based on CBAM_ResNet and ACON activation function[J]. Applied Sciences-Basel, 2023, 13(13): 7593.
- [22] XU Z, WANG Y, XIONG W, et al. A novel attentional feature fusion with inception based on capsule network and application to the fault diagnosis of bearing with small data samples[J]. Machines, 2022, 10(9): 789.
- [23] 杨博, 胡珍珍. 基于 YOLOv8n 改进算法的自动驾驶目标检测 [J/OL]. 控制工程 [2025-02-05]. DOI:10.14107/j.cnki.kzgc.20240494.
YANG Bo, HU Zhenzhen. Automatic driving target detection based on YOLOv8n improved algorithm[J]. Control Engineering of China[2025-02-05]. DOI:10.14107/j.cnki.kzgc.20240494.
- [24] 钟帅, 王丽萍. 改进 YOLOv8 的无人机航拍图像目标检测方法 [J/OL]. 航空兵器, 1-11[2025-02-07]. <https://kns.cnki.net/kcms/detail/41.1228.TJ.20241218.1415.003.html>.
ZHONG Shuai, WANG Liping. Improved YOLOv8 object detection method for drone aerial images[J/OL]. Aero Weaponry, 1-11[2025-02-07]. <https://kns.cnki.net/kcms/detail/41.1228.TJ.20241218.1415.003.html>.
- [25] 王璐, 张爽. 基于 EMD 和 FFT 的自适应 X 射线脉冲星信号降噪方法 [J]. 电波科学学报, 2025, 40(2): 381-394.
WANG Lu, ZHANG Shuang. Adaptive denoising method of X-ray pulsar signal based on EMD and FFT[J]. Chinese Journal of Radio Science, 2025, 40(2): 381-394.

(责任编辑 杜亚勤)