

基于 WSSA-VMD 与改进 Swin Transformer 的风机叶片声纹故障检测^{*}

张荣荣 王金明 李忠虎 张 飞

(内蒙古科技大学自动化与电气工程学院 包头 014010)

摘 要: 针对传统故障检测方法在部署难度、环境抗干扰能力及早期故障敏感性方面的不足,提出一种基于 WSSA-VMD 与改进 Swin Transformer 的声纹故障检测模型。首先,提出一种融合鲸鱼优化算法的麻雀搜索算法(WSSA),用于自适应优化变分模态分解(VMD)的关键参数,以提升对原始声纹信号的降噪性能;其次,选取优化 VMD 分解后的有效模态分量重构信号,并将其转换为对数梅尔谱图,作为高质量特征输入;最后,构建一种改进的 Swin Transformer 模型,通过引入卷积注意力模块(CBAM)并采用均方根归一化方法替换原归一化层,以增强模型的特征提取能力与训练稳定性。实验结果表明,所提模型的总体精准率达到 97.93%,不仅能够有效识别早期叶片故障,也展现出良好的泛化性能。

关键词: 风机叶片;参数优化;VMD;对数梅尔谱图;故障检测

中图分类号: TP391;TN911.7 **文献标识码:** A **国家标准学科分类代码:** 510.4030

Research on fault detection of acoustic signals of wind turbine blades based on WSSA-VMD and improved Swin Transformer

Zhang Rongrong Wang Jinming Li Zhonghu Zhang Fei

(School of Automation and Electrical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China)

Abstract: In response to the shortcomings of traditional fault detection methods in terms of deployment difficulty, environmental anti-interference capability, and early fault sensitivity, this paper proposes a voiceprint fault detection model based on WSSA-VMD and the improved Swin Transformer. First, a sparrow search algorithm integrated with the whale optimization algorithm (WSSA) is introduced to adaptively optimize the key parameters of variational mode decomposition (VMD), thereby enhancing the denoising performance of the original acoustic signals. Next, the effective intrinsic mode functions obtained from the optimized VMD are selected to reconstruct the signal, which is then converted into Log-Mel spectrograms to serve as high-quality feature inputs. Finally, an improved Swin Transformer model is constructed by incorporating the convolutional block attention module (CBAM) and replacing the original normalization layer with a root mean square normalization method, so as to strengthen the model's feature extraction capability and training stability. Experimental results demonstrate that the proposed model achieves an overall accuracy of 97.93%, effectively identifying incipient blade faults while exhibiting strong generalization performance.

Keywords: wind turbine blades; parameter optimization; VMD; Log-Mel spectrogram; fault detection

0 引言

风能作为清洁能源的重要组成部分,其规模化开发利用对实现“双碳”目标具有重要意义。风力发电机作为风能转换的核心设备,长期运行于复杂多变的恶劣环境中,叶片作为其关键部件极易产生裂纹、断裂等损伤,一旦

发生故障可能导致严重停机 and 经济损失^[1-3]。因此,开展风机叶片状态的早期智能故障诊断研究,对保障风电机组安全稳定运行、降低运维成本具有重大的工程实际价值。

目前,传统风电叶片故障检测主要采用振动分析^[4]、声发射^[5]、红外热成像^[6]及机器视觉^[7]等方法。尽管这些方法在实验中表现出一定效果,但仍存在诸多局限性。刘瑞

收稿日期:2025-11-14

^{*} 基金项目:国家自然科学基金(62161042)、内蒙古自然科学基金(2024LHMS06002)、内蒙古自治区科技计划项目(2025YFHH0061)资助

刚^[8]通过将双轴振动传感器安装在距离叶根 1/3 叶片总长的位置,采集叶片振动信号,并提出了结合时域与频域的分析方法来定位故障。然而振动信号的信噪比较低,实际部署时需要大量传感器,且风速、转速、温度、环境运行条件等都会显著影响叶片的动态特性。Bejger 等^[9]提出了基于声发射信号逆变换与相干函数分析的风电叶片损伤诊断方法,该方法通过提取波谱中的特征度量,有效区分了叶片无损与损伤状态,但声发射技术通常需在叶片上安装传感器,不仅实施难度大,还可能破坏叶片原有结构。刘毅等^[10]通过红外热成像技术对叶片腐蚀进行检测研究。而红外热成像的检测性能易受环境温度波动的影响,稳定性不足。机器视觉技术虽在叶片检测领域有所应用,如汤占军等^[11]通过对无人机拍摄的图像进行处理,并提出一种用改进的萤火虫算法优化支持向量机(support vector machine, SVM)的检测模型,实现了对叶片微小裂纹的检测。但该类方法受环境光照及天气条件制约显著,在机组运行状态下难以获取高质量图像,影响故障判断。

近年来,以 Transformer 及其注意力机制为核心的深度学习方,因其强大的长程依赖建模与特征自适应聚焦能力,在机械故障诊断领域取得了突破性进展,并迅速成为研究热点。在通用机械故障诊断方面,最新研究正致力于提升模型在强噪声与小样本下的鲁棒性:陈志刚等^[12]提出一种基于时间卷积网络(temporal convolutional network, TCN)-Transformer 的端到端自适应诊断模型。该模型通过 TCN 提取局部时频特征,并利用结合注意力机制的 Transformer 增强跨通道全局依赖建模,在行星齿轮箱多传感器诊断任务中实现了较高的识别准确率。Wang 等^[13]提出一种基于加权平均算法-特征模态分解与自适应多尺度注意力-Swin Transformer 的轴承故障诊断方法,通过优化特征分解过程、强化特征提取能力,在轴承声信号诊断中取得显著降噪效果与高精度识别性能,展现出显著优势。

风电机组运行过程中产生的声学信号蕴含了丰富的状态信息,为叶片故障检测提供了一种非接触、低成本且易于部署的新途径。与依赖物理传感器接触安装的振动、声发射等方法相比,声纹检测可通过远程采集方式有效避免对机组运行的干扰,并具备更强的环境适应性。然而,现有基于声纹的故障诊断方法仍面临两大挑战:1)现场采集的声信号信噪比极低,背景噪声与干扰严重掩盖了微弱的早期故障特征;2)传统声学故障检测方法对叶片损伤的敏感度不足,难以建立高鲁棒性、高精度的识别模型。针对上述问题与不足,并在充分借鉴最新注意力模型研究进展的基础上,提出一种基于融合鲸鱼优化算法的麻雀搜索算法-变分模态分解(sparrow search algorithm integrated with whale optimization algorithm-variational mode decomposition, WSSA-VMD)与改进 Swin Transformer 的声纹故障诊断方法。首先,通过引入鲸鱼优化算法(whale optimization algorithm, WOA)、改进麻雀优化算法(sparrow search

algorithm, SSA),增强其全局寻优能力,并以此优化变分模态分解(variational mode decomposition, VMD)的模态数 k 与惩罚因子 α ,提升信号分解质量;其次,利用优化后的 VMD 对声纹信号进行降噪重构,并转换为对数梅尔谱图以增强故障特征的表达能力;最后,为克服现有 Swin Transformer 模型对高噪声声纹特征聚焦不足的缺陷,构建在深层网络中引入卷积注意力机制(convolutional block attention module, CBAM)并采用均方根归一化(root mean square normalization, RMSNorm)的改进 Swin Transformer 分类模型,提升对复杂特征的特征能力。为全面验证方法的有效性,本文设计了消融、对比以及泛化性实验,验证了该方法在风机叶片故障诊断中的有效性,为复杂噪声环境下的设备状态检测提供新的解决方案。

1 理论与方法

1.1 VMD

VMD^[14]是一种自适应信号分解方法,它的核心思想是通过构建并求解约束性变分问题,实现信号频域内自适应分解。具体步骤如下:

1)构建变分问题。假设原始声纹信号 $f(t)$ 可分解为 k 个具有有限带宽的模态函数 $\{u_k\}$ 及其对应中心频率 $\{\omega_k\}$,为使得各模态在频域上相互分离、重叠最小,并保证分解后的信号能够无损重构,建立如下约束变分模型:

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s. t. } \sum_k u_k = f(t) \end{cases} \quad (1)$$

式中: ∂_t 为时间 t 的偏导数; $\delta(t)$ 为狄拉克函数; $f(t)$ 是原始信号。

2)求解变分问题。为求解上述约束问题,引入二次惩罚因子和拉格朗日乘法算子,将其转化为非约束变分问题,表达式如式(2)所示。

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \\ & \| f(t) - \sum_k u_k(t) \|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \end{aligned} \quad (2)$$

式中: α 为二次惩罚因子; $\lambda(t)$ 为拉格朗日乘法算子。

1.2 SSA 与 WOA

SSA^[15]是一种模拟麻雀觅食与避险行为的群体智能优化算法。在该算法中,麻雀群体被划分为 3 种角色:发现者、加入者和警戒者,每种角色遵循不同的位置更新策略。

1)发现者负责引导种群搜索方向,其位置更新公式为:

$$\mathbf{X}_{i,j}^{t+1} = \begin{cases} \mathbf{X}_{i,j}^t \exp\left(-\frac{i}{aM}\right), & R_2 < ST \\ \mathbf{X}_{i,j}^t + q\mathbf{L}, & R_2 \geq ST \end{cases} \quad (3)$$

式中: t 为当前迭代次数; M 为最大迭代次数; a 为 $(0, 1]$ 内的随机数; i 为麻雀的次序数; j 为维数; $\mathbf{L}$ 为元素均为 1 的 $1 \times D$ 的矩阵; q 为服从标准正态分布的随机数; R_2 为 $[0,$

1]间的预警阈值;ST 为[0.5,1]间的安全值。当 $R_2 < ST$ 时,表明周围没有捕食者,反之表示周围有捕食者。

2)在觅食过程中,加入者会持续监视发现者,一旦发现者找到了更好的食物源,加入者便会主动发起争夺。其位置更新规则如式(4)所示。

$$\mathbf{X}_{i,j}^{t+1} = \begin{cases} q \exp\left(\frac{\mathbf{X}_{\text{worst}}^t - \mathbf{X}_{i,j}^t}{i^2}\right), & i > n/2 \\ \mathbf{X}_p^{t+1} + |\mathbf{X}_{i,j}^t - \mathbf{X}_p^{t+1}| \mathbf{A}^+ \mathbf{L}, & i \leq n/2 \end{cases} \quad (4)$$

式中: $\mathbf{X}_p$ 为发现者的最优位置; $\mathbf{X}_{\text{worst}}$ 为种群中最差位置; $\mathbf{A}^+$ 为 $1 \times D$ 的矩阵,其元素随机取值为 -1 或 1。当 $i > n/2$ 时,表示适应度低的麻雀处于饥饿状态,它需要飞到其他区域觅食。

3)当出现危险时,警戒者会立刻发出警示信号,促使其他麻雀飞向安全地点。其位置更新规则如式(5)所示。

$$\mathbf{X}_{i,j}^{t+1} = \begin{cases} \mathbf{X}_{\text{best}}^t + \beta |\mathbf{X}_{i,j}^t - \mathbf{X}_{\text{best}}^t|, & f_i \neq f_g \\ \mathbf{X}_{i,j}^t + K \left(\frac{|\mathbf{X}_{i,j}^t - \mathbf{X}_{\text{worst}}^t|}{(|f_i - f_w| + e)} \right), & f_i = f_g \end{cases} \quad (5)$$

式中: β 为服从标准正态分布的步长; $\mathbf{X}_{\text{best}}$ 为麻雀种群中的最优位置; K 为[-1,1]间的随机数; f_i 、 f_g 、 f_w 分别为当前个体、全局最优和最差适应度; e 设为常量。当 $f_i \neq f_g$ 时,表明个体处于种群边缘,易受攻击,需调整位置以趋近安全区域。

WOA^[16] 是一种受座头鲸捕食行为启发而提出的智能优化算法。该算法主要模拟了座头鲸在猎食过程中的 3 种行为:

1)围猎行为。鲸鱼通过收缩包围不断调整位置以接近猎物。其位置更新公式为:

$$\mathbf{X}(t+1) = \mathbf{X}^*(t) - \mathbf{A} \cdot \mathbf{D} \quad (6)$$

$$\mathbf{D} = |\mathbf{C} \cdot \mathbf{X}^*(t) - \mathbf{X}(t)| \quad (7)$$

式中: $\mathbf{X}^*(t)$ 为当前最优解,即猎物位置; $\mathbf{X}(t)$ 为当前鲸鱼的位置; $\mathbf{D}$ 为当前鲸鱼位置与猎物位置的距离; $\mathbf{A}$ 和 $\mathbf{C}$ 为控制包围行为的随机权重系数。

2)气泡网捕食行为。鲸鱼采用螺旋方式逼近猎物,模拟气泡网捕食行为:

$$\mathbf{X}(t+1) = \mathbf{D}' e^{bi} \cos(2\pi l) + \mathbf{X}^*(t) \quad (8)$$

$$\mathbf{D}' = |\mathbf{X}^*(t) - \mathbf{X}(t)| \quad (9)$$

式中: b 为螺旋形状常数; l 为[-1,1]的随机数。

3)搜索猎物行为。为增强全局搜索能力,鲸鱼以随机个体为参考进行搜索:

$$\mathbf{D} = |\mathbf{C} \times \mathbf{X}_{\text{rand}} - \mathbf{X}(t)| \quad (10)$$

$$\mathbf{X}(t+1) = \mathbf{X}_{\text{rand}} - \mathbf{A} \times \mathbf{D} \quad (11)$$

式中: $\mathbf{X}_{\text{rand}}$ 为随机选择的鲸鱼位置。

WOA 通过融合收缩包围与螺旋搜索机制,有效平衡了全局探索与局部开发能力,具有较强的鲁棒性和收敛性能。

1.3 Swin Transformer

Swin Transformer^[17] 采用层次化的网络结构设计,依次包含 4 个特征提取阶段。首先,进行输入图像的块划分,经过 Patch Partition 层,图像被划分为不重叠的 4×4 的图像块,特征维度随之由 $[H, W, 3]$ 转换为 $[H/4, W/4, 48]$ 。随后,在 Stage 1 中,Linear Embedding 层将每个块的通道数进行线性映射至 C ,生成维度为 $[H/4, W/4, C]$ 的特征图。进而,网络通过后续多个阶段处理,依次使用 Patch Merging 层对特征图进行下采样,并堆叠 Swin Transformer Block 以构建强大的特征表示。Swin Transformer 的整体结构如图 1 所示。

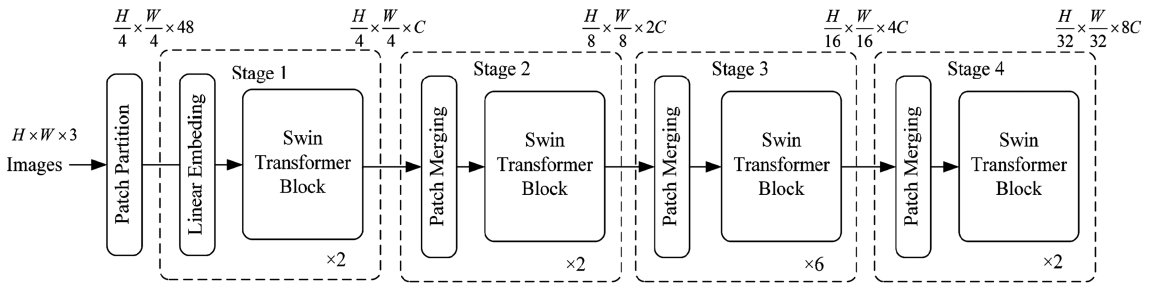


图 1 Swin Transformer 结构

Fig. 1 Swin Transformer structure

该模型的核心在于引入窗口多头自注意力(window multi-head self-attention, W-MSA)机制和移位窗口多头自注意力(shifted window multi-head self-attention, SW-MSA)机制。二者共同构成了如图 2 所示的 Swin Transformer 模块。该模型首先将输入特征图划分为多个不重叠的窗口,左侧路径中,W-MSA 模块依次通过 Layer Norm 层、W-MSA 单元、多层感知机(multi-layer perceptron, MLP)层并结合跳跃连接进行注意力计算;为促进跨窗口的信息交互,右侧的 SW-MSA 模块首先将不

同的窗口进行拼接,随后经 Layer Norm 层、SW-MSA 单元与 MLP 层并同样借助跳跃连接完成全局上下文融合。图 2 中 $\hat{Z}^l$ 与 Z^l 分别为第 l 个窗口多头自注意力模块和多层感知器模块输出的特征。

2 声纹故障诊断模型

2.1 基于 WSSA-VMD 的声纹信号降噪

风机叶片长期暴露于自然环境,不仅受到雨雪、风沙等环境侵蚀,还持续承受风电机组运行所产生的复杂动态载

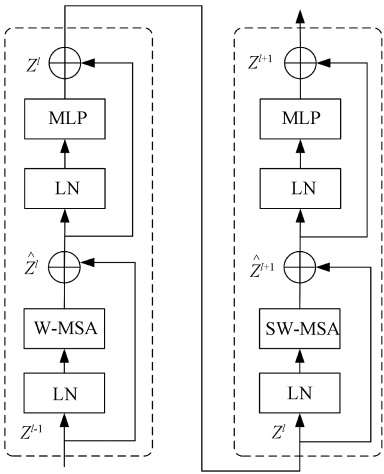


图 2 Swin Transformer 模块
Fig. 2 Swin Transformer block

荷。这些复合因素导致采集到的声纹信号中含有大量背景噪声,严重干扰有效信息的提取^[18]。因此,对原始信号进行降噪预处理是后续分析的关键步骤。本文采用 VMD 进行信号分解与去噪。声纹信号降噪流程如图 3 所示。

然而,VMD 方法需人为预设模态分量 k 和惩罚系数 α ,若参数取值不当将显著影响分解效果。传统研究多通过观察中心频率来确定 k 值,并采用默认的 α 值,但该方法主观性较强,缺乏普适性与自适应性。为此,本文引入 WSSA 算法以自适应地确定 VMD 的最优参数组合 $[k, \alpha]$ 。SSA 在优化过程中容易过度依赖当前最优解,导致的早熟收敛。为改善这一局限,WSSA 在 SSA 的初始探索阶段引入 WOA,通过 WOA 的随机搜索行为增强种群多样

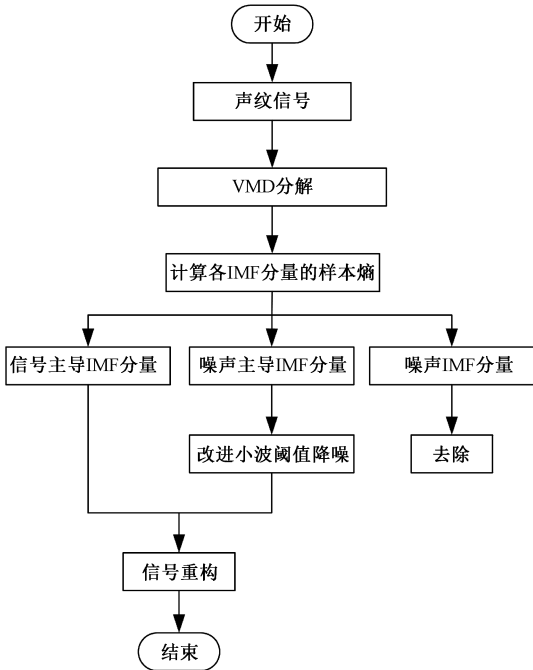


图 3 声纹信号降噪流程
Fig. 3 Voiceprint signal denoising flowchart

性,避免算法过早陷入局部最优,从而在保留 SSA 快速收敛特性的同时,融合了 WOA 强大的全局寻优能力,为复杂优化问题提供更为有效的解决方案。为验证 WSSA 的改进效果,将 WSSA 与单一 SSA 和 WOA 进行对比实验。实验选取 6 个通用的标准测试函数,标准测试函数如表 1 所示。设置种群数量为 30,最大迭代次数为 500,进行 30 次单独重复实验,测试结果如表 2 所示。

表 1 标准测试函数
Table 1 Standard test functions

标准函数	维度	搜索空间	最小值
$F_1(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$	30	$[-100, 100]$	0
$F_2(x) = \max x_i \{ x_i , 1 \leq i \leq n \}$	30	$[-100, 100]$	0
$F_3(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left(\frac{1}{n} \sum_{i=1}^n \cos 2\pi x_i \right) + 20 + e$	30	$[-32, 32]$	0
$F_4(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos \left(\frac{x_i}{\sqrt{i}} \right) + 1$	30	$[-600, 600]$	0
$F_5(x) = \frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6}$	30	$[-65.536, 65.536]$	1
$F_6(x) = \sum_{i=1}^n x_i \sin(x_i) + 0.1 x_i $	30	$[-50, 50]$	0

从表 2 可以看出,对于所有的标准测试函数,WSSA 算法在绝大多数情况下都表现出了比 SSA 和 WOA 更优的性能。具体而言,在函数 $F_1(x)$ 、 $F_2(x)$ 和 $F_6(x)$ 上,

WSSA 的平均值和标准差均显著低于其他两种算法,体现出 WSSA 在优化方面具有极强的收敛精度和稳定性。在 $F_3(x)$ 和 $F_4(x)$ 这两个函数上,WSSA 同样表现出色,其中

表 2 基于标准测试函数 8 种算法测试对比结果

Table 2 Comparison results of tests of eight algorithms based on standard test functions

标准函数	度量指标	WSSA	SSA	WOA
$F_1(x)$	平均值	7.55×10^{-97}	2.25×10^{-23}	4.18×10^3
	标准差	4.06×10^{-96}	1.17×10^{-22}	1.04×10^4
$F_2(x)$	平均值	7.75×10^{-27}	6.73×10^{-14}	6.88×10^{-11}
	标准差	4.17×10^{-26}	2.56×10^{-13}	2.68×10^{-10}
$F_3(x)$	平均值	4.44×10^{-16}	2.73×10^{-14}	2.34×10^{-15}
	标准差	0	5.99×10^{-14}	2.20×10^{-15}
$F_4(x)$	平均值	0	8.88×10^{-17}	0.01
	标准差	0	2.11×10^{-16}	0.06
$F_5(x)$	平均值	3.22	4.30	3.85
	标准差	2.76	3.66	3.47
$F_6(x)$	平均值	3.47×10^{-36}	4.14×10^{-6}	5.61×10^{-6}
	标准差	1.43×10^{-35}	1.48×10^{-5}	3.02×10^{-5}

$F_4(x)$ 的平均值和标准差均为 0,表明其在多次运行中均能稳定收敛至理论最优解;而在 $F_5(x)$,WSSA 的平均值和标准差虽只略低于 WOA 和 SSA,但也显示出其良好的全局探索能力。总体而言,WSSA 在收敛精度和稳定性方面优于对比算法,验证了其在处理不同类型优化问题时的有效性与优越性。

此外,标准函数的收敛曲线对比如图 4 所示。

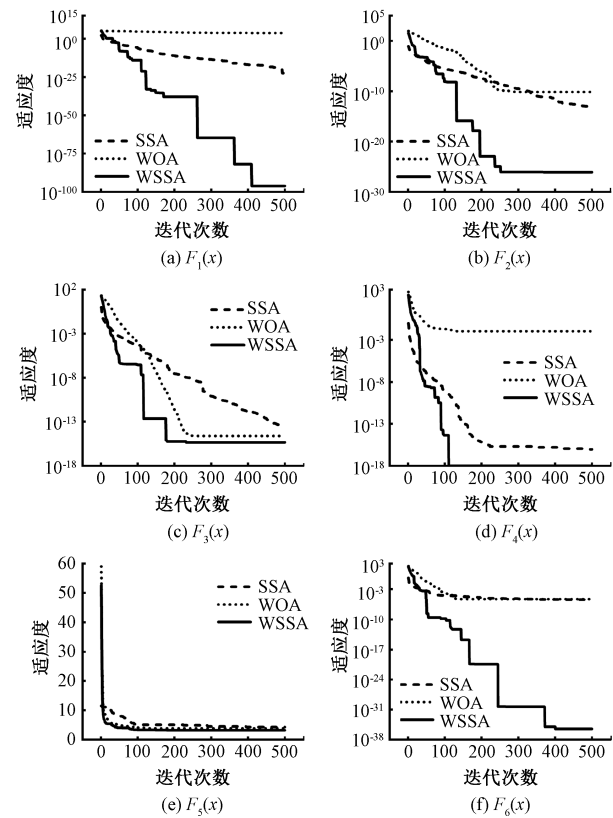


图 4 收敛曲线对比

Fig.4 Comparison of convergence curves

由图 4 可以看出,WSSA 在迭代初期适应度下降更快,且在后期能够稳定收敛至远低于 SSA 和 WOA 的数值,进一步说明其在收敛速度与寻优精度上的综合优势。

在 VMD 参数优化中,本文以信号包络熵最小化为适应度函数,包络熵越小,包络分布越集中、周期性越强、故障特征越明显。当包络熵最小时,表明 IMF 分量中的噪声含量最低,特征信息最丰富,分解效果达到最优。包络熵公式为:

$$\begin{cases} q_i = \frac{h(i)}{\sum_{i=1}^I h(i)} \\ E_q = -\sum_{i=1}^I q_i \lg q_i \end{cases} \quad (12)$$

式中: q_i 为概率分布系列; $h(i)$ 为调节后的包络信号; E_q 为信号的包络熵。

经迭代寻优,最终确定最优参数组合为 $[k, \alpha] = [8, 2\ 500]$ 。在获取最优参数后,对原始声纹信号进行 VMD 分解,分解结果如图 5 所示。

为评估各 IMF 分量的信号特性,计算了各分量的样本熵值,以区分信号主导分量、噪声主导分量及噪声分量。如图 6 所示,各 IMF 分量的样本熵分布清晰地展现了不同分量的噪声含量差异。

基于样本熵分析结果,样本熵值越小表明其包含的有效信号信息越丰富。首先,根据熵值分析结果剔除了被识别为纯噪声的 IMF1~3 分量。对于被判定为噪声主导的 IMF4 分量,采用改进的小波阈值降噪方法进行针对性处理,该方法对噪声成分处理方法进行了优化,能够在有效保留有用信号特征的同时去除噪声干扰。最后,将降噪处理后的 IMF4 分量与信号主导的 IMF5~8 分量进行重构,得到最终的去噪声纹信号。降噪前后的波形对比如图 7 所示。

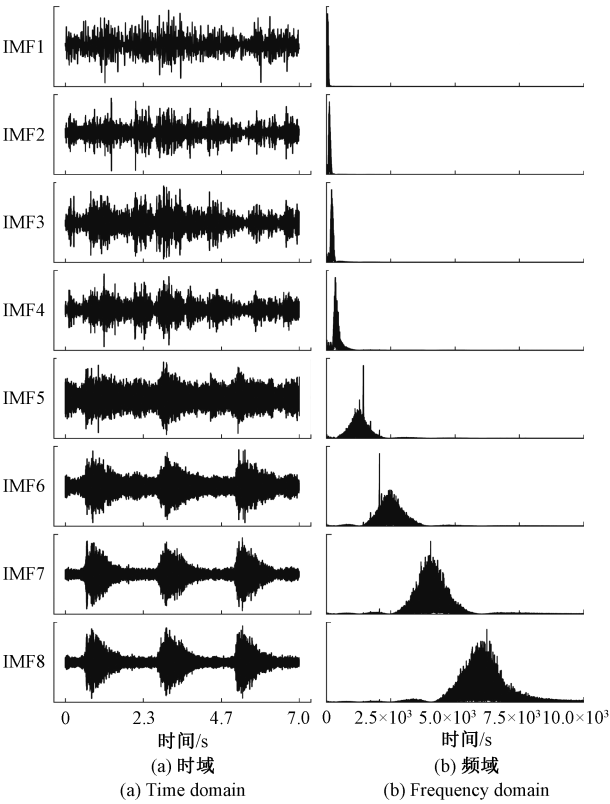


图 5 VMD 分解结果

Fig. 5 VMD decomposition results

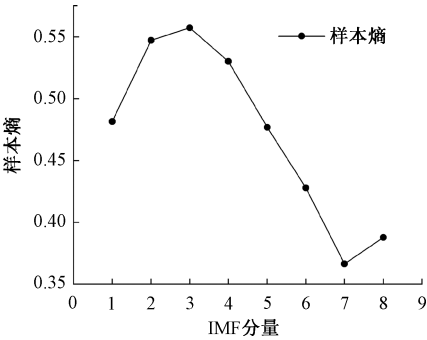


图 6 各 IMF 分量样本熵

Fig. 6 Sample entropy of each IMF component

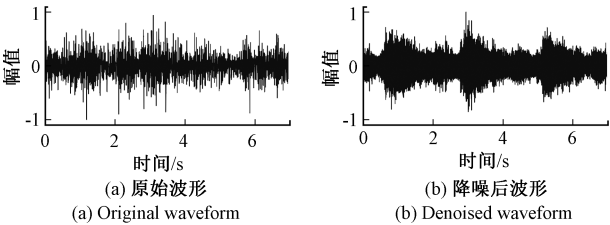


图 7 波形对比

Fig. 7 Waveform comparison

2.2 对数梅尔谱图转换

为将降噪后的时域声纹信号转换为更适合模型处理

的特征表示,需要选取有效的声纹特征。为此,本文系统评估了不同时频特征图对故障信息的保留能力,分别从风机叶片的降噪声纹信号中提取了小波图、语谱图特征及对数梅尔谱图特征,并在同一 SVM 分类器框架下进行故障识别性能的比较。

实验结果如表 3 所示。对数梅尔谱图特征的分类准确率最高。这表明对数梅尔谱图在故障信息保留方面具有明显优势。其原因主要可归结为两点:1)该特征完整保留了信号的时频结构信息,具备与语谱图相当的时频分辨率;2)通过梅尔尺度滤波器组模拟人耳的非线性听觉感知机制,使其能更敏锐地捕捉与故障相关的细微声学特征。

表 3 不同特征识别性能

Table 3 Recognition performance of different features

特征图	SVM 准确率 %
小波图	81.80
语谱图	81.36
对数梅尔谱图	84.16

基于上述比较结果,本研究最终选择对数梅尔谱图作为声纹输入特征。对数梅尔谱图提取流程如图 8 所示。

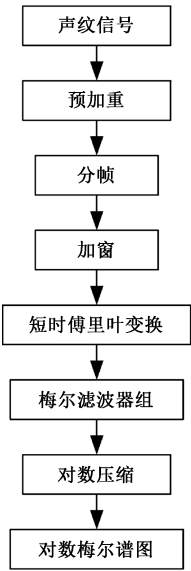


图 8 对数梅尔谱图流程

Fig. 8 Log-Mel spectrogram flowchart

2.3 基于改进的 Swin Transformer 的故障识别

为解决复杂工况下风机叶片故障的精准识别问题,在完成信号降噪与对数梅尔谱图转换的基础上,本文引入一种改进的 Swin Transformer 模型作为核心故障检测模型,用于实现故障的准确识别。

该模型以前文所得的对数梅尔谱图作为输入,其核心

创新主要体现在以下两个方面:

首先,模型整合了 CBAM 至分层特征提取流程中,通过其通道与空间双重注意力机制,在不同的层级上增强对

声纹特征的关注,有效抑制无关背景干扰,从而显著提升模型对细微故障特征的感知灵敏度。改进后的 Swin Transformer 模型如图 9 所示。

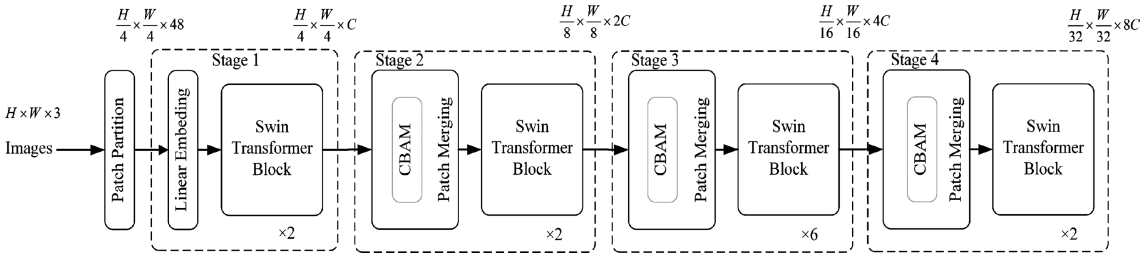


图 9 改进的 Swin Transformer
Fig. 9 Improved Swin Transformer

其次,将标准 Swin Transformer 模块中的 LayerNorm 层替换为 RMSNorm,该替换在有效保持训练稳定性的同时,简化了归一化计算过程,降低了模型复杂度,从而进一步提升了模型的训练与推理效率。改进后的 Swin Transformer Block 如图 10 所示。

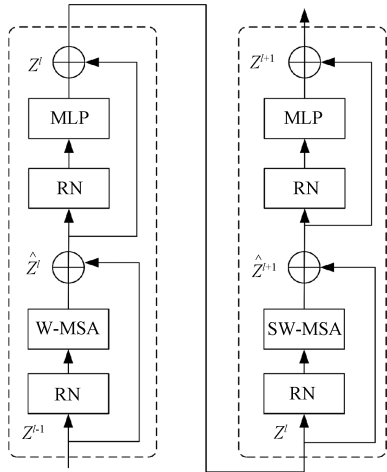


图 10 改进的 Swin Transformer 模块
Fig. 10 Improved Swin Transformer block

上述改进共同增强了模型在复杂工况下的特征鉴别能力与运行效能,为最终实现高可靠性的故障诊断奠定了坚实基础。

3 实验结果与分析

3.1 数据集

本研究通过声纹采集设备共获取 2 918 个风电机组叶片声纹样本,其中包含 2 378 个正常叶片样本与 540 个故障叶片样本(涵盖裂纹与异物附着类型)。为增强模型的泛化能力和训练稳定性,采用时间缩放与音调调节技术对原始数据进行增强,增强后的数据集如表 4 所示。随后,将增强后的数据集按 7:2:1 的固定比例划分为训练集、验证集和测试集。

表 4 增强后数据集

Table 4 Enhanced dataset

样本类型	增强后样本个数
正常样本	8 560
故障样本	2 870

3.2 实验环境与参数配置

实验硬件配置采用 NVIDIA GeForce RTX 3080 Ti GPU 和 Intel(R) Xeon(R) Silver 4210 CPU,操作系统为 Ubuntu 20. 04. 6 LTS,编程语言使用 Python3. 9,深度学习框架为 Pytorch2. 8. 0,CUDA 版本为 12. 0。模型训练参数设置如表 5 所示。

表 5 模型训练参数

Table 5 Model training parameters

训练参数	数值
初始学习率	0. 001
迭代次数	200
权重衰减	1×10^{-4}
批量大小	32
优化器	Adam

3.3 模型评价指标

为全面评估基于 WSSA-VMD 与改进 Swin Transformer 模型的性能优势,本研究采用准确率(accuracy, A)、精准率(precision, P)、召回率(recall, R)和 F1 分数(F-1 score, $F1$)作为性能评价指标,以系统衡量模型在检测任务中的表现。各评价指标定义为:

$$A = \frac{TP + TN}{TP + FP + FN + TN}$$
 (13)

$$P = \frac{TP}{TP + FP}$$
 (14)

$$R = \frac{TP}{TP + FN}$$
 (15)

$$F1 = \frac{2 \times P \times R}{P + R}$$
 (16)

式中:TP 指标签为正类且被正确预测的样本数量;FN 指标签为正类但被错误预测为负类的样本数量;FP 指标签为负类但被错误预测为正类的样本数量;TN 指标签为负类且被正确预测的样本数量。

3.4 消融实验

为了评估 C(CBAM)和 N(RMSNORM)两种方法对网络性能的影响,本文设计了消融实验,通过逐步添加和组合添加的方式,分析各改进方法对模型性能的贡献,具体实验结果如表 6 所示。

表 6 消融实验

模型	A/%	P/%	R/%	F1/%	FPS
S T	92.91	87.05	84.32	85.66	95
S T-C	96.50	92.73	93.38	93.05	90
S T-N	95.54	90.97	91.29	91.13	123
S T-C+N	99.21	97.93	98.95	98.44	115

消融实验结果表明,原始模型 A 为 92.91%。单独加入 CBAM 后,A 提升至 96.50%,F1 达到 93.05%,FPS 稍降;仅使用 RMSNorm 时,A 为 95.54%,F1 为 91.13%,FPS 有所提升。同时引入 CBAM 与 RMSNorm 后,模型达到最优性能:A 为 99.21%,F1 分数 98.44%,且保持较高 FPS,说明二者具有良好协同增强效果。

3.5 检测结果分析

本研究首先对声纹信号进行去噪,随后将其转换为对数梅尔谱图作为特征表示。将提取的特征输入改进的 Swin Transformer 模型进行训练后,得到的混淆矩阵如图 11 所示。模型性能评估结果表明,P 达到 97.93%,R 为 98.95%,显示出优异的故障检测能力。

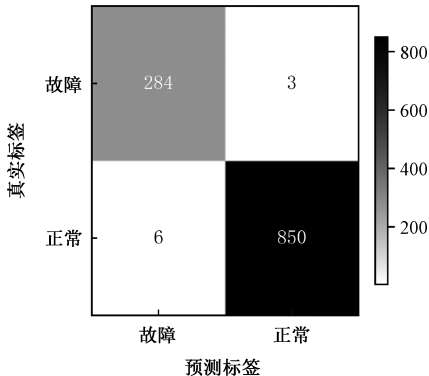


图 11 混淆矩阵

Fig. 11 Confusion matrix

为验证基于 WSSA-VMD 与改进 Swin Transformer 的风机叶片异常检测方法的优越性,本文与 SVM、卷积神经网络(convolutional neural network,CNN)、Transformer 模型以及标准的 Swin Transformer 模型进行了对比实验,实验结果如表 7 所示。

表 7 检测结果对比

模型	A	P	R	F1
SVM	84.16	64.40	82.58	72.37
CNN	87.23	69.42	87.80	77.54
Transformer	87.75	86.57	60.63	71.31
S T	92.91	87.05	84.32	85.66
本文方法	99.21	97.93	98.95	98.44

实验结果表明,本文所提出的方法在 4 种核心性能指标上均显著优于各对比模型。具体而言,本文方法 97.93%的高精准率显著降低了误报风险,减少了不必要的运维成本;98.95%的高召回率确保了故障检测的全面性,有效避免了漏检可能引发的安全事故。相较于 SVM、CNN、Transformer 模型以及标准 Swin Transformer 模型,均有提升。综合以上性能优势可知,基于 WSSA-VMD 与改进 Swin Transformer 的故障检测方法,在风机声纹故障检测任务中具有更优的适用性与有效性。

本研究设计了全面的验证实验来评估故障检测模型的泛化能力。实验数据按工况划分为两组:工况 1(包含正常样本与异物故障样本)和工况 2(包含正常样本与裂纹故障样本)。鉴于两类故障样本数量不平衡,实验以样本较少的故障类型数量为基准进行均衡采样。

首先,在同工况测试中建立性能基准:工况 1 内测试使用工况 1 数据训练和测试;工况 2 内测试使用工况 2 数据训练和测试。这两项实验反映了模型在数据分布一致条件下的最优性能。

随后,进行跨工况泛化测试:在工况 1→工况 2 任务中,模型使用工况 1 数据训练,在工况 2 数据上测试。此时,模型需适应新工况条件并识别此前未接触过的裂纹故障类型。为进行双向验证,同步进行工况 2→工况 1 跨工况测试,全面考察模型在不同工况间的识别能力。模型泛化性能验证结果如表 8 所示。

表 8 泛化性实验结果

测试场景	A	P	R	F1
工况 1 内测试	98.21	97.10	95.71	96.40
工况 2 内测试	97.86	95.71	95.71	95.71
工况 1→2	96.93	93.98	93.71	93.84
工况 2→1	96.86	93.47	94.00	93.73

实验结果表明,所提出的方法在不同工况之间具有稳定的泛化性能,即使在工作条件变化且出现新型故障时,仍能保持较高的检测准确率与故障召回能力,进一步验证了本文提出的基于 WSSA-VMD 与改进 Swin Transformer 的故障检测方法的适用性与有效性。

4 结 论

针对风机叶片声纹信号降噪与故障识别问题,本文提出了基于 WSSA-VMD 与改进 Swin Transformer 的智能诊断方法。研究首先通过融合 WOA 与 SSA 算法优点的 WSSA,自适应地确定了 VMD 的最佳分解参数(k 与 α),有效克服了传统 VMD 参数选择依赖经验的局限性,实现了在强噪声背景下对声纹信号的高质量分解与降噪。随后,将降噪信号转换为对数梅尔谱图,构建了能够同时保留声音时频特征与听觉特性的有效输入。

在识别阶段,本文对 Swin Transformer 模型进行了关键性改进:在整合了 CBAM 注意力机制至分层特征提取流程中,强化了对故障敏感特征的筛选与聚焦能力;采用 RMSNorm 替代 LayerNorm,在保证训练稳定性的同时提升了计算效率。最终模型通过其分层移位窗口机制,实现了从局部细微异常到全局模式演变的精准捕捉。

实验结果表明,本方法在准确率等关键指标上均显著优于 SVM、CNN 等传统模型,验证了其在风机声纹异常检测任务中的优越性、鲁棒性与工程应用潜力。本研究为复杂设备的状态监测与智能运维提供了一种行之有效的技术路径。

参考文献

- [1] 盛昌文,姜永正,黄磊,等. 基于叶尖定时的风机叶片裂纹故障识别研究[J]. 仪器仪表学报, 2024, 45(4): 57-65.
SHENG C W, JIANG Y Z, HUANG L, et al. Research on fan blade crack fault identification based on tip timing [J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 57-65.
- [2] 张泽辰,姚顺春,梁骁俊,等. 基于小目标感知增强的风机叶片缺陷智能检测[J]. 仪器仪表报, 2025, 46(9): 159-172.
ZHANG Z C, YAO S C, LIANG X J, et al. Wind turbine blade intelligent defect detection based on small object perception enhancement [J]. Chinese Journal of Scientific Instrument, 2025, 46(9): 159-172.
- [3] YANG C C, DING S H, ZHOU G S. Wind turbine blade damage detection based on acoustic signals[J]. Scientific Reports, 2025, 15(1): 3930.
- [4] 黄凯,刘强,张墨. 基于小波分析的风机叶片振动实时监测方法[J]. 自动化应用, 2022(7): 41-43, 47.
HUANG K, LIU Q, ZHANG M. Real-time monitoring method of fan blade vibration based on wavelet analysis[J]. Automation Application, 2022(7): 41-43, 47.
- [5] YAN Q, CHE X C, LI S, et al. π -FBG fiber optic

- acoustic emission sensor for the crack detection of wind turbine blades[J]. Sensors, 2023, 23(18): 7821.
- [6] 尹玉,张永,王健,等. 基于热红外图像的风力机叶片损伤识别方法研究[J]. 太阳能学报, 2022, 43(2): 492-497.
YIN Y, ZHANG Y, WANG J, et al. Research on wind turbine blade damage identification method based on thermal infrared image[J]. Acta Energiæ Solaris Sinica, 2022, 43(2): 492-497.
 - [7] TAN X G, ZHANG G M. Research on surface defect detection technology of wind turbine blade based on UAV image[J]. Instrumentation, 2022, 9(1): 41-48.
 - [8] 刘瑞刚. 基于振动的风机叶片故障监测方法研究[J]. 中国新技术新产品, 2022(18): 69-71.
LIU R G. Research on vibration-based fault monitoring method of wind turbine blades[J]. New Technology & New Products of China, 2022(18): 69-71.
 - [9] BEJGER A, DRZEWIENIECKI J B, BARTOZZKO P, et al. The use of coherence functions of acoustic emission signals as a method for diagnosing wind turbine blades[J]. Energies, 2023, 16(22): 7474.
 - [10] 刘毅,储银贺,李波. 基于持续热激励红外检测的风机叶片腐蚀检测[J]. 激光与红外, 2023, 53(5): 716-722.
LIU Y, CHU Y H, LI B. Corrosion detection of fan blade based on continuous thermal excitation infrared detection[J]. Laser & Infrared, 2023, 53(5): 716-722.
 - [11] 汤占军,孙栋钦,李英娜,等. 改进 FA 优化 SVM 的风机叶片裂纹检测模型[J]. 可再生能源, 2022, 40(9): 1189-1194.
TANG Z J, SUN D Q, LI Y N, et al. Improved FA optimized SVM based crack detection model for fan blades [J]. Renewable Energy Resources, 2022, 40(9): 1189-1194.
 - [12] 陈志刚,陶子纯,王衍学,等. 基于混合注意力的 TCN-Transformer 行星齿轮箱故障诊断方法[J]. 振动与冲击, 2025, 44(20): 348-356.
CHEN Z G, TAO Z C, WANG Y X, et al. Fault diagnosis method of planetary gearbox with TCN-Transformer based on hybrid attention mechanism[J]. Journal of Vibration and Shock, 2025, 44(20): 348-356.
 - [13] WANG H D, WANG H K, XIE J Z, et al. Fault diagnosis of rolling bearing acoustic signal under strong noise based on WAA-FMD and LGAF-Swin Transformer[J]. Processes, 2025, 13(9): 2742.

[14]

万佳诚, 曾宪文, 李靖超. 基于参数优化 VMD 与宽卷积神经网络的齿轮箱故障诊断[J]. 电子测量技术, 2025, 48(10): 25-32.
WAN J C, ZENG X W, LI J C. Fault diagnosis for gearbox based on parameter optimization VMD combined with wide convolutional neural network[J]. Electronic Measurement Technology, 2025, 48(10): 25-32.

[15]

江兵, 李响, 巢一帆, 等. 基于 KPCA-CGSSA-KELM 的变压器故障识别方法[J]. 电子测量与仪器学报, 2024, 38(5): 139-147.
JIANG B, LI X, CHAO Y F, et al. Transformer fault recognition method based on KPCA-CGSSA-KELM[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(5): 139-147.

[16]

张炎亮, 李营. 基于多尺度排列熵和 IWOA-SVM 的滚动轴承故障诊断[J]. 电子测量技术, 2023, 46(19): 29-34.
ZHANG Y L, LI Y. Rolling bearing fault diagnosis based on multi-scale permutation entropy and IWOA-SVM[J]. Electronic Measurement Technology, 2023, 46(19): 29-34.

[17]

周舟, 陈捷, 吴明明. 基于 CWT 和优化 Swin Transformer 的风电齿轮箱故障诊断方法[J]. 振动与冲击, 2024, 43(15): 200-208.
ZHOU Z, CHEN J, WU M M. Fault diagnosis method for wind power gearbox based on wavelet transform and optimized Swin Transformer [J]. Journal of Vibration and Shock, 2024, 43(15): 200-208.

[18]

余洪伍, 汤占军, 马锦雄. 基于声纹特征融合的风机叶片异常识别方法[J]. 电子测量与仪器学报, 2024, 38(11): 99-108.
YU H W, TANG Z J, MA J X. Wind turbine blade anomaly recognition method based on sound feature fusion [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(11): 99-108.

作者简介

张荣荣, 硕士研究生, 主要研究方向为基于声纹识别的风电机组叶片故障检测。
E-mail: 3287705924@qq.com
王金明(通信作者), 副教授, 主要研究方向为风电机组检测。
E-mail: yb15661510224@163.com