

doi: 10.7690/bgzd.2025.12.017

基于 MTF-ResNet-CBAM 的兵工装备轴承故障诊断方法

李伟伟¹, 杨悦², 花国祥¹

(1. 无锡学院自动化学院, 江苏 无锡 214000; 2. 南京信息工程大学自动化学院, 南京 210044)

摘要: 针对传统方法在噪声干扰和复杂工况下性能有限的问题, 提出一种基于马尔可夫转移场 (Markov transition field, MTF)、残差网络 (residual network, ResNet) 与卷积注意力机制 (convolutional block attention module, CBAM) 的轴承故障诊断方法。将 1 维振动信号映射为 2 维 MTF 图像, 以保留时序依赖与动态特征; 利用 ResNet 进行深层特征提取, 通过 CBAM 在通道与空间维度自适应分配权重, 强化关键信息表达、抑制冗余干扰。在 4 类典型工况 (正常、内圈故障、外圈故障和滚动体故障) 下进行实验验证。结果表明: 该模型整体测试准确率达到 96.67%, 较 VGG、AlexNet 及 CNN 模型提升约 8%~15%, 该方法在兵工装备的复杂运行环境下能保持较高的诊断精度与稳定性。

关键词: 轴承故障诊断; 兵工装备; MTF; ResNet; CBAM

中图分类号: TP277.3 **文献标志码:** A

Fault Diagnosis Method for Military Equipment Bearings Based on MTF-ResNet-CBAM

Li Weiwei¹, Yang Yue², Hua Guoxiang¹

(1. School of Automation, Wuxi University, Wuxi 214000, China;

2. School of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: In order to solve the problem that the traditional method has limited performance under noise interference and complex working conditions, a method based on Markov transition field (MTF), residual network (ResNet) and convolutional block attention module (CBAM). The 1-D vibration signal is mapped to a 2-D MTF image to retain the temporal dependence and dynamic features. ResNet is used to extract the deep features, and CBAM is used to adaptively assign weights in the channel and spatial dimensions to enhance the expression of key information and suppress redundant interference. The experimental verification is carried out under four typical working conditions (normal, inner ring fault, outer ring fault and rolling element fault). The results show that the overall test accuracy of the model reaches 96.67%, which is about 8%~15% higher than that of VGG, AlexNet and CNN models, and the method can maintain high diagnostic accuracy and stability in the complex operating environment of ordnance equipment.

Keywords: bearing fault diagnosis; military equipment; MTF; ResNet; CBAM

0 引言

滚动轴承作为旋转机械中不可或缺的核心部件, 已在兵工装备、航空航天、能源动力以及交通运输等多个领域得到广泛应用, 其健康状态直接关系到系统运行的安全性与可靠性^[1]。在兵工装备长期服役过程中, 轴承易受高载荷、强冲击和复杂环境因素影响而产生疲劳、裂纹或磨损等故障, 若未能及时诊断, 不仅会降低装备的性能, 而且可能导致任务失效甚至严重事故^[2]; 因此, 发展适用于兵工领域的高精度轴承故障诊断方法, 对提升武器装备的作战可靠性和维护保障能力具有重要意义^[3-4]。

传统的轴承故障诊断方法多依赖时域、频域或

时频域特征提取, 并结合支持向量机 (support vector machines, SVM)、浅层神经网络等分类器进行模式识别^[5]。这类方法在一定程度上能够反映信号特征, 但高度依赖人工经验, 且在复杂噪声和变工况环境下易出现特征提取不足与诊断性能不稳定的问题^[6]。随着深度学习技术的发展, 卷积神经网络 (convolutional neural network, CNN) 因其端到端的自动特征学习能力, 被广泛应用于机械故障诊断^[7]; 但直接输入 1 维振动信号仍存在特征利用不足、类间可分性不强等局限^[8]。

为提升特征表达能力, 提出一种信号可视化的策略, 即将 1 维时序信号转化为 2 维图像, 再借助深度学习模型开展分类任务。例如, 格拉姆角场 (gramian angular field, GAF)、递归图 (recurrence

收稿日期: 2024-10-06; 修回日期: 2024-11-12

基金项目: 新能源电力系统全国重点实验室开放课题项目 (LAPS23014); 无锡学院引进人才科研启动专项 (2022r021)

第一作者: 李伟伟 (1982—), 男, 河南人, 博士。

plot, RP) 和马尔可夫转移场 (MTF) 等方法能够从不同角度揭示时间序列的动态特征。其中, MTF 通过统计信号状态间的转移概率, 将时间序列映射为 2 维结构化图像, 在保留时间相关性的同时有效表征动态演化规律, 在机械故障诊断任务中展现出良好的应用前景^[9-10]。

深层卷积网络在特征提取方面具备显著优势, 但随着深度增加, 模型容易出现梯度消失、训练困难等问题。残差网络 (ResNet) 通过引入跨层残差连接, 有效缓解了深度网络退化现象, 并具备更强的特征学习能力^[11]。同时, 注意力机制在近年来受到广泛关注, 其中卷积注意力模块 (CBAM) 能够在通道与空间维度自适应地分配权重, 突出关键信息并抑制冗余特征, 从而进一步提升模型的判别能力^[12]。

基于上述考虑, 笔者提出一种结合 MTF、ResNet 与 CBAM 的轴承故障诊断方法。实验结果表明, 该方法在多种工况下均能实现较高的识别精度与鲁棒性, 优于多种对比模型, 为复杂环境下的轴承故障诊断提供了新的技术思路, 也为武器系统健康监测与可靠运行提供了技术参考。

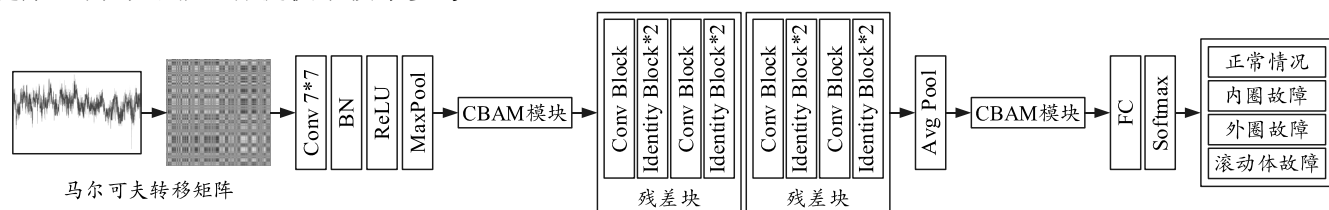


图 1 MTF-Resnet-CBAM 模型结构

2 基本理论

2.1 MTF

MTF 是一种基于马尔可夫过程的时间序列可视化方法, 能够将 1 维信号映射为 2 维图像表示, 以便捕捉其动态特性与潜在模式。该方法首先将时间序列信号离散化为若干状态并构建状态空间, 然后通过统计不同状态间的转移次数获得转移概率, 进而形成马尔可夫转移矩阵。该矩阵不仅描述了序列在不同状态间的演化规律, 还能够在 2 维图像中保留时间依赖关系, 从而在特征提取过程中兼顾时域信息与空间分布特征。

依据图像编码的基本原理, 假设 $X=\{x_1, x_2, \dots, x_n\}$ 为给定的时间序列, 其中 $x_i(i=1, 2, \dots, n)$ 表示在时刻 i 时振动信号的幅值。出于状态空间的便捷构建考虑, 依据序列幅值的分布将其划分为 Q 个分位区间, 然后使用分位数 $q_j(q_j \in [1, Q])$ 对时间序列进行

1 模型介绍

笔者提出的故障诊断模型融合了 MTF、ResNet 与 CBAM, 整体框架包括信号预处理、特征图生成、深层特征提取与注意力增强分类等环节。

在数据处理阶段, 振动信号经过归一化操作后由 MTF 方法映射为 2 维特征图, 该过程能够刻画状态间的转移概率, 保留时序依赖关系与动态演化特征, 相较于传统时频分析方法在复杂工况下具有更强的鲁棒性。

深度特征提取环节采用 ResNet 结构。残差单元通过跨层连接缓解了深层网络中的梯度消失问题, 确保模型在多层结构下仍能有效收敛, 并能够捕捉多层次的语义特征。

在特征优化环节引入 CBAM 模块。该模块对卷积特征在通道维度和空间维度进行加权分配, 通道注意力强化不同特征通道的重要性, 空间注意力聚焦于关键区域分布, 从而提升模型对故障特征的敏感度并抑制冗余干扰。

最终, 通过全连接层与 Softmax 分类器完成故障类别的判别。完整的网络模型结构如图 1 所示。

量化, 将每个离散后的信号点 x_i 映射到对应的分位区间 q_i 中。在完成量化处理后, 基于一阶马尔可夫过程对时间序列进行建模, 通过统计相邻分位区间内的转换次数来推导其转移概率。考虑到时间序列本身的顺序特征, 该概率能够有效反映相邻采样点之间的动态演化规律。由此可形成一个规模为 $Q \times Q$ 的马尔可夫转移矩阵 W , 其具体定义如式(1)和(2)^[13]:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1Q} \\ w_{21} & w_{22} & \cdots & w_{2Q} \\ \vdots & \vdots & & \vdots \\ w_{Q1} & w_{Q2} & \cdots & w_{QQ} \end{bmatrix}; \quad (1)$$

$$w_{ij} = p\{x_{t+1} \in q_j | x_t \in q_i\}. \quad (2)$$

式中 w_{ij} 为信号在时刻 t 处属于分位区间 q_i 时, 经一步转移后进入分位区间 q_j 的概率, 该转移概率满足

$\sum_{j=1}^Q w_{ij} = 1$ 的归一化条件。

传统马尔可夫链所构建的状态转移矩阵具有典型的“无记忆性”，即下一时刻的状态转移概率仅与当前状态相关，而与更早的历史序列无关。这种特性在一定程度上削弱了 1 维时序信号 X 的时间依赖性描述，若直接采用转移矩阵 W ，将不可避免地丢失部分时序信息。为弥补这一缺陷，MTF 方法在建模时引入时间步长与分位区间之间的耦合关系，并在时间维度上对转移概率进行有序展开，从而将状态转移矩阵 W 拓展为跨时间尺度的 MTF 矩阵 M ，更全面地保留了序列中的时序依赖特征。其数学定义如式(3)和(4)所示：

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1Q} \\ m_{21} & m_{22} & \cdots & m_{2Q} \\ \vdots & \vdots & & \vdots \\ m_{Q1} & m_{Q2} & \cdots & m_{QQ} \end{bmatrix}; \quad (3)$$

$$m_{ij} = p\{w_{ij} | x_i \in q_i, x_j \in q_j\}. \quad (4)$$

式中 m_{ij} 为从分位数 q_i 转换为分位数 q_j 的概率，同时 $q_i, q_j \in [1, Q]$ 。

经由上述映射，1 维振动信号被转化为能够表征时序特性的 2 维特征表示，深度学习模型能够在该图像空间中对像素信息进行逐层遍历与提取，由此可提取深层特征并保持时间序列的内在依赖关系，从而降低关键信息的损失风险。通过结合分位数划分与马尔可夫转移概率，所得特征表示在信号耦合与噪声干扰条件下展现出更强的稳定性与鲁棒性。同时，将转移概率限定在 $[0, 1]$ 区间内，有助于缓解梯度爆炸问题，提升网络训练的稳定性并加速收敛。

MTF 以紧凑的形式捕捉时间序列的状态演化特征，在智能诊断与深度学习相关领域表现出较高应用价值。其核心思想是将序列的动态规律映射为 2 维结构化图像，从而为时序数据建模与特征学习开辟新的研究路径^[14]。

2.2 ResNet

随着网络深度的持续增加，模型易出现过拟合、梯度爆炸与梯度消失等现象，从而导致性能退化，不仅提升了训练复杂度，也使分类精度逐渐趋于饱和。为解决这一困境，提出 ResNet，其核心由多个残差块构成。如图 2 所示，设 X 为残差块的输入， $F(X)$ 为非线性映射函数， $H(X)$ 为残差块输出。残差

结构在卷积层输入与输出之间建立跨层连接，不再直接学习输入到输出的映射关系，而是学习二者之间的残差差异。这种设计有效降低了特征学习的复杂度，使深层网络更易优化，并显著加快收敛速度。

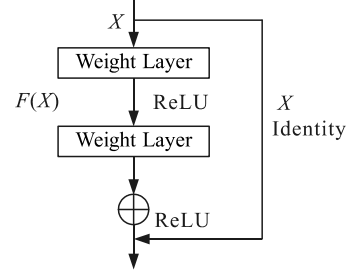


图 2 残差块基本结构

ResNet 采用跳跃连接，将学习目标更改为：

$$F(x) = H(x) - x. \quad (5)$$

式中： x 为 CNN 的输入； $H(x)$ 为处理后的输出； $F(x)$ 为最终输出。

在反向传播计算中，根据链式求导原理，损失函数 ϵ 对变量 x_l 的梯度可写为：

$$\begin{aligned} \frac{\partial \epsilon}{\partial x_l} &= \frac{\partial \epsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \epsilon}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i) \right) = \\ &= \frac{\partial \epsilon}{\partial x_L} + \frac{\partial \epsilon}{\partial x_L} \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i). \end{aligned} \quad (6)$$

在训练过程中， $\frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i)$ 不可能始终为

-1，因此在残差网络中不会出现梯度完全消失的情况。换言之， $\partial \epsilon / \partial x_l$ 表示的梯度能够在反向传播过程中由第 L 层直接传递至任意较浅的第 l 层，从而保证了深层网络的有效训练。

关于参数 w 的梯度为：

$$\hat{y}' = g(f(x))' = \frac{\partial \hat{y}}{\partial y} \frac{\partial y}{\partial w}. \quad (7)$$

式中 $\partial \hat{y} / \partial y$ 为顶部层结果和真实结果的偏差。在大批量训练的情况下，该值往往趋近于 0，即梯度幅值极小，导致特征难以有效更新，尤其当网络顶部结构为全连接层时，这一问题对模型性能的影响尤为显著。

ResNet 以高效的方式将前一层网络的输出与当前层网络的输出相结合，并将其作为下一层网络的输入，从而有效构建了深层神经网络模型。

$$y_{\text{residual}} = f(2x) + g(2f(2x)). \quad (8)$$

其梯度为：

$$y'_{\text{residual}} = \frac{\partial y}{\partial w} + \frac{\partial y}{\partial y} \frac{\partial y}{\partial w}. \quad (9)$$

ResNet 通过引入跳跃连接,将原始输入与非线性变换后的输出共同传递,从而实现残差映射的高效学习并获得更具判别力的特征表示。在其演化模型 ResNet50 中,网络深度进一步加深,可提取更高层次的特征图,从而提升图像表示能力与特征表达效率通过引入残差学习策略,ResNet50 利用跨层跳跃连接实现对输入与输出残差的建模,从而缓解了深层网络训练中的梯度消失难题,并在整体性能与收敛速度上取得了显著提升^[15]。

2.3 CBAM 注意力模块

注意力机制 (attention mechanism, AM) 是深度学习中的一项关键技术,其核心思想是使模型在信息处理过程中能够自动突出关键特征并削弱次要特征。该理念源于人类视觉系统的认知特点,即在感知复杂场景时,人类往往会将注意力集中于某些显著区域,而忽略无关细节。将这一机制引入到深度学习模型中,可以显著提升模型在各种任务上的性能,包括图像识别、语音识别、自然语言处理等。

例如,SENet 通过对通道间的依赖关系进行建模,让模型学会对不同通道赋予不同的注意力权重。CBAM 则是在此基础上引入了空间注意力机制,即不仅关注哪些通道更重要,还关注特征图中的哪些位置更为关键。

CBAM 是一种轻量且有效的注意力机制模块,专为加强 CNN 的特征提取能力设计。它通过有序地整合通道注意力和空间注意力 2 个子模块来关注重要的通道和空间区域,通道注意力致力于识别最具信息量的通道,而空间注意力则突出关键的空间位置。

经 CBAM 处理的特征图首先通过通道注意力模块,该模块利用最大池化与平均池化的相加操作来挑选重要通道。随后,特征图传递至空间注意力模块,在此通过堆叠最大池化和平均池化操作的结果来确定关键的空间区域。得到的通道和空间权重与原始特征图相乘,以此调节特征图的适应性,其模型结构如图 3 所示。CBAM 通过此种融合池化操作的策略,放弃了单一的池化方法,从而有效提升了模型对不同特征的关注度和表现力^[16]。

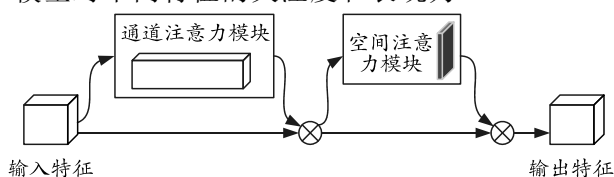


图 3 CBAM 注意力模块结构

$F \in R^{C \times H \times W}$ 是假设的输入特征图,通过 CBAM 注意力模块,依序生成 1 维的通道关注度映射图: $M_c = R^{C \times 1 \times 1}$, 2 维空间的注意图 $M_s = R^{1 \times H \times W}$ 。注意力整体过程可以描述为:

$$F' = M_c(F) \otimes F; \quad (10)$$

$$F'' = M_s(F') \otimes F'。 \quad (11)$$

式中: F' 为经过通道注意力操作获得的更新特征图; F'' 为紧接着经由空间注意力机制处理得到的最终特征图。

在通道注意力模块,输入特征图通过融合最大池化与平均池化的操作进行加权,从而产生通道注意力图。该模块通过压缩特征映射的空间尺寸来评估图像各通道的相对重要性,强调关键特征同时抑制次要信息。结合 2 种池化操作的方法,相较于单独采用任一池化方式,能够提供更丰富的表征能力。通道注意力的计算可以表达为:

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \\ &\quad \text{MLP}(\text{MaxPool}(F))) = \\ &\quad \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c)))。 \end{aligned} \quad (12)$$

式中: σ 为 Sigmoid 函数; W_0 与 W_1 共享 MLP 的权重, $W_0 \in R^{C/r \times C}$ 、 $W_1 \in R^{C \times C/r}$ 。

在空间注意力模块中,输入特征图首先通过最大池化与平均池化在特征空间上进行聚合,随后将二者的结果堆叠以生成空间注意力映射。该机制作为通道注意力的补充,能够识别在空间维度上承载关键信息的位置,并通过融合平均池化与最大池化的结果来增强特征表达。

这样就会形成 2 个 2 维的特征图 $F_{\text{avg}}^s \in R^{1 \times H \times W}$ 和 $F_{\text{max}}^s \in R^{1 \times H \times W}$,它们各自象征着通道的均值池化与最大池化特性。这 2 类特征图经由标准卷积操作融合,形成一个 2 维空间注意力映射。该模块通过在通道维度上的池化操作生成有效的特征描述,以实现空间注意力的建模,其具体计算过程可由式(13)加以说明:

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \\ &\quad \text{MaxPool}(F)])) = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])). \end{aligned} \quad (13)$$

式中: σ 为 Sigmoid 函数; $f^{7 \times 7}$ 为 7×7 大的卷积核^[17]。

3 数据获取及处理

3.1 数据集介绍

笔者所用实验平台主要由电机、传感器、联轴器、滚动轴承、齿轮箱及主轴组成,分别安装正常

轴承、内圈故障轴承、外圈故障轴承与滚动体故障轴承。在转速 600 r/min 的条件下，对 4 种状态(正常、内圈故障、外圈故障、滚动体故障)的振动信号进行了采集，轴承各状态如图 4 所示。

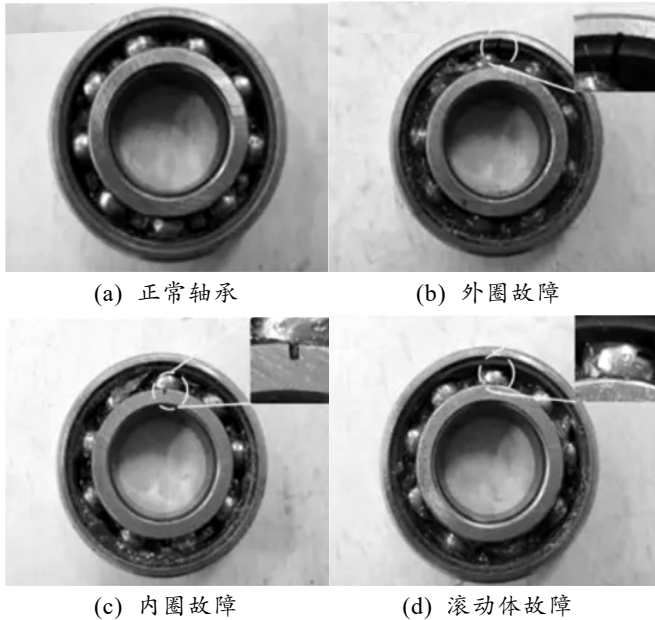


图 4 轴承各状态图片

数据集中样本长度设置为 2 048, 每种工况下取 150 个样本, 共计 600 个样本。数据集的诊断任务设置如表 1 所示。

表 1 数据集诊断任务

工况	标签	样本数
正常	0	150
内圈故障	1	150
外圈故障	2	150
滚动体故障	3	150

3.2 数据处理

利用马尔可夫转移矩阵, 将这 600 个 1 维振动信号样本映射为 2 维马尔可夫图像。在该图像中, 时间序列的动态演化特征被转化为像素强度或颜色分布, 从而使后续能够借助图像处理与深度学习方法开展分析与分类。图 5—8 展示了 1 维振动序列经过映射后生成的 2 维图像示例。

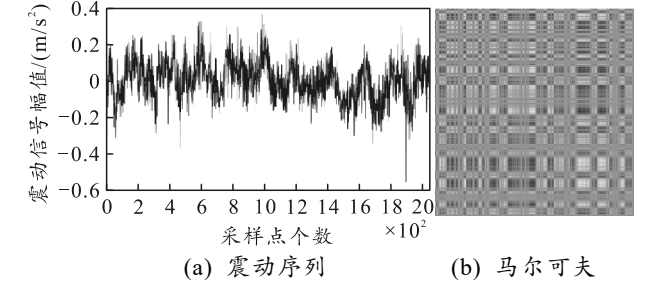


图 5 正常工况马尔可夫转移场转化结果

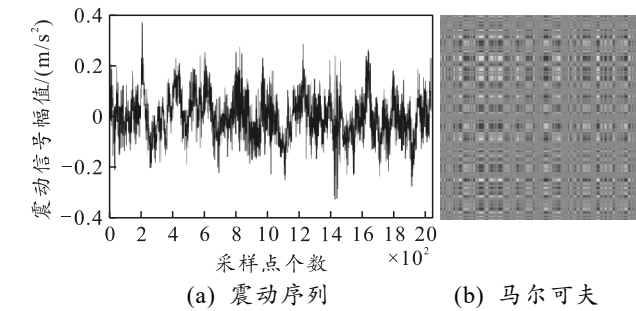


图 6 内圈故障马尔可夫转移场转化结果

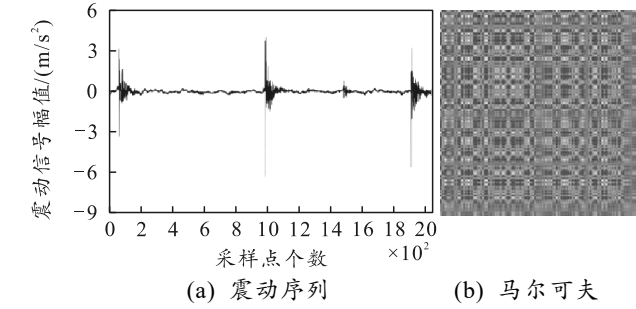


图 7 外圈故障马尔可夫转移场转化结果

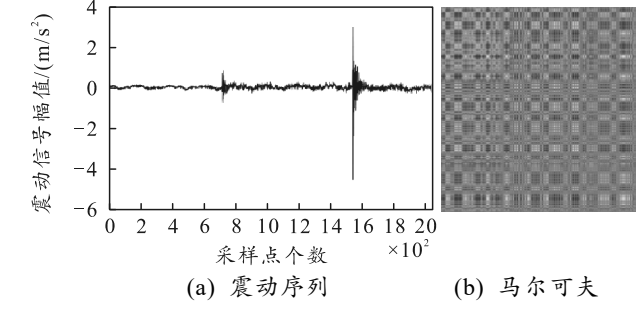


图 8 滚动体故障马尔可夫转移场转化结果

4 实验分析

4.1 参数配置及评估指标

实验在配备 NVIDIA RTX 4070 Laptop GPU 与 Intel i9-14900HX CPU 的计算平台上进行, 模型的构建与训练依托 Pytorch2.6-Python 3.7 环境完成。优化算法采用 Adam, 初始学习率设定为 0.001, 权重衰减系数为 0.005, 并结合 ReduceLROnPlateau 策略, 在验证集准确率连续 8 个周期未提升时, 将学习率衰减为原来的 0.5, 且最低可降至 $1e^{-6}$, 以增强模型的收敛稳定性并抑制过拟合。在批处理大小的设置上, 分别对 16、32、64 与 128 4 种参数进行对比实验, 综合考虑训练耗时与模型性能后, 确定 32 为最优批量规模。最大迭代次数设定为 100, 以在收敛性与训练效率之间取得平衡。

在实验设计中, 为检验所构建滚动轴承故障诊断模型的有效性, 将全部特征向量数据集按照 80% 与 20% 的比例随机划分为训练部分和测试部分。训

训练数据用于模型的参数优化与结构学习,而测试数据则用于在未见样本上进行验证,以考察模型的识别准确率与泛化水平。

评估指标为准确率、精确率、召回率和 F_1 分数,计算公式如下所示:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}; \quad (14)$$

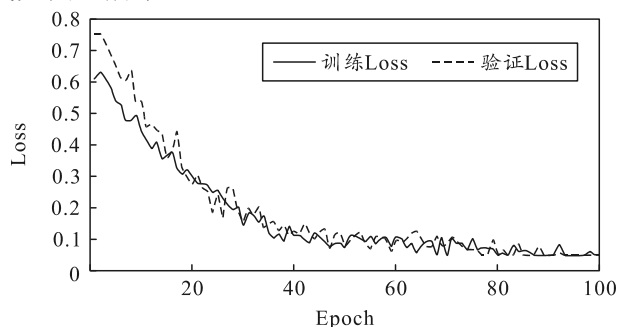
$$\text{Precision} = \frac{TP}{TP + FP}; \quad (15)$$

$$\text{Recall} = \frac{TP}{TP + FN}; \quad (16)$$

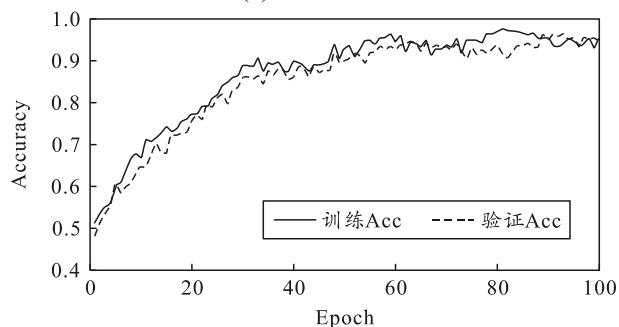
$$F_1\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}。 \quad (17)$$

4.2 结果分析

所有图像样本均被统一缩放至 224×224 像素后输入模型进行训练。模型训练准确率和损失值曲线如图 9 所示。



(a) 损失曲线



(b) 准确率曲线

图 9 模型训练准确率和损失值曲线

由图 9 可见: MTF-ResNet-CBAM 网络在初始阶段即表现出较高的识别精度,随着迭代次数的增加,训练准确率持续提升,并在约 20 次迭代后趋于稳定并保持在较高水平。同时,损失函数在前 10 次迭代内呈显著下降趋势,并在 40 次迭代后逐渐平稳接近于 0。这表明所设计的网络模型能够有效收敛,具备良好的拟合性能与训练效果。测试集上各类别的实验结果如表 2 所示。

表 2 测试集上各类别的实验结果

类别	Accuracy/%	Precision/%	Recall/%	F_1 -score
正常	100.00	96.77	100.00	0.983 6
内圈故障	96.67	100.00	96.67	0.983 1
外圈故障	93.33	96.55	93.33	0.949 2
滚动体故障	96.67	93.55	96.67	0.950 8

从表 2 可以看出: 笔者提出模型在 4 类工况下均取得了较高的分类性能。对于“正常”类别, Recall 达到 100%, 表明模型能够完整识别出所有正常样本, 未出现漏检情况; 其 Precision 为 96.77%, 说明虽然存在极少量误判, 但整体分类效果依然可靠。对于“内圈故障”类别, Precision 达到 100%, 表明所有被预测为内圈故障的样本均为真实内圈故障, 未出现误判; 而 Recall 相对较低 96.67%, 说明该类样本仍存在少量漏检。相比之下, “外圈故障”和“滚动体故障”的 Precision 与 Recall 均维持在 93%~97%, 二者较为均衡, F_1 -score 分别为 0.949 2 和 0.950 8, 体现了模型在这 2 类故障诊断任务中的稳健表现。

图 10 展示了测试集的混淆矩阵, 纵轴为实际故障类别, 横轴为模型识别结果。

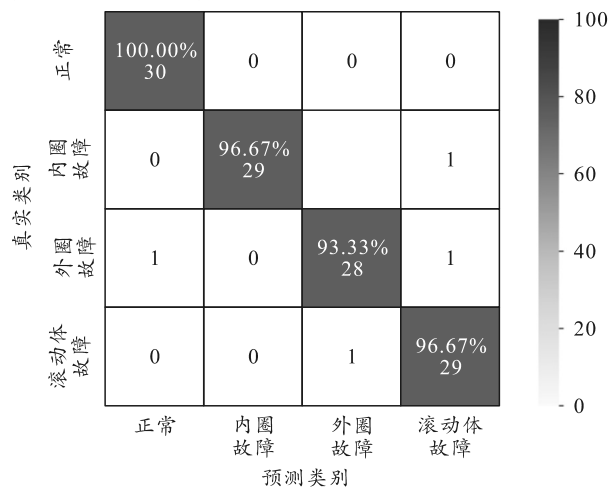


图 10 测试集混淆矩阵

从分类结果可以看出: 该模型在正常状态、内圈故障、外圈故障和滚动体故障 4 类工况下的识别准确率分别为 100%、96.67%、93.33% 和 96.67%, 整体测试集的综合准确率达到 96.67%。这一结果表明, 所提出的模型在多类故障诊断任务中不仅保持了较高的分类精度, 而且在不同类别之间具有较好的平衡性与鲁棒性, 能够有效区分不同故障类型, 验证了其在实际应用中具备较强的可靠性与实用价值。

为证明模型在特征学习方面的优势, 笔者使用 T-SNE 技术对样本进行了 2 维可视化。图 11 展示

了模型处理前后的分布情况。原始样本在降维空间中呈现出无序和交叠，类别界限模糊；而在经过模型提取与映射后，各类样本逐渐聚合成独立簇群，类间隔离度显著提升，充分体现了模型在故障特征表达与分类判别上的优越性能。

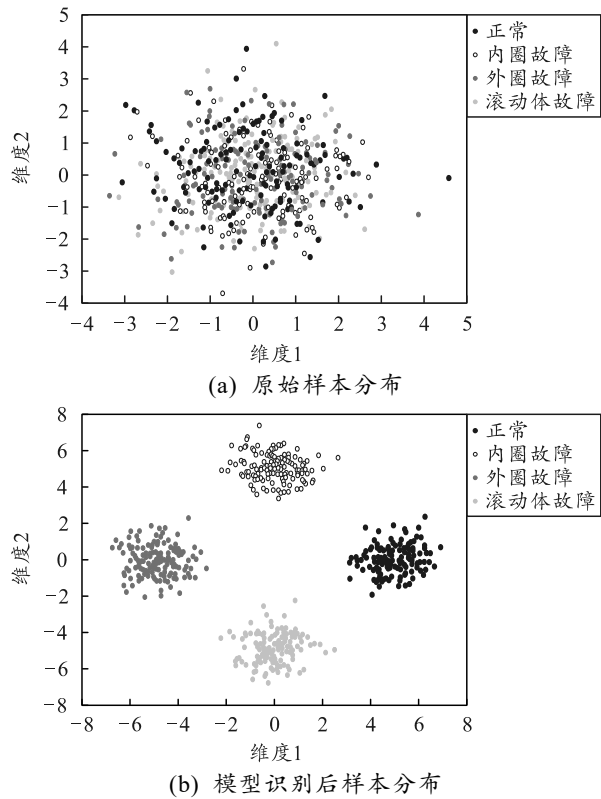


图 11 T-SNE 降维分布在模型识别前后的结果

4.2 对比实验

为验证 MTF-ResNet50-CBAM 模型的性能优势，笔者在相同数据集与实验条件下，将其与 VGG、AlexNet 及传统 CNN 等典型模型进行了对比。不同模型训练准确率对比如图 12 所示。

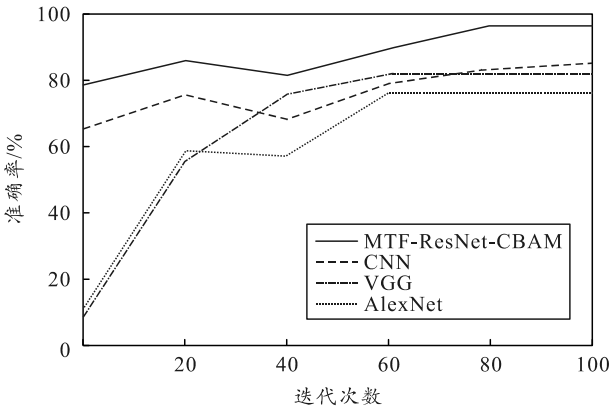


图 12 不同模型训练准确率对比

由图 12 可知：所提模型在分类准确率上表现最优，相较于 CNN 提升约 10%，体现出更为突出的

故障识别能力与稳定性。
对比 4 种模型在整体准确率以及不同工况下特征向量提取过程中的损失率，结果显示：MTF-ResNet-CBAM 模型不仅在分类准确率上表现最佳，同时在特征损失率方面也最低。该对比结果充分证明了所提方法的优势，使得滚动轴承故障的诊断更加精确与可靠。

4.3 消融实验

为验证所提模型各组成模块对整体性能的贡献，笔者进行了消融实验。实验依次移除或替换关键模块，构建以下对比模型：

- 1) ResNet：直接以原始振动信号训练 ResNet 模型，不进行 MTF 转换；
 - 2) MTF-ResNet：采用 MTF 将信号转化为 2 维图像，但不引入 CBAM 模块；
 - 3) MTF-ResNet-SE：在 MTF-ResNet 基础上，引入 SE 注意力模块作为对比；
 - 4) MTF-ResNet-CBAM(本文中方法)：完整模型，结合 MTF、ResNet 与 CBAM。
- 消融实验对比如表 3 所示。

表 3 消融实验对比

模型	Accuracy/%	Precision/%	Recall/%	F ₁ -score
ResNet	73.45	73.45	73.20	0.733 0
MTF-ResNet	80.68	80.70	80.60	0.806 5
MTF-ResNet-SE	86.45	86.50	86.40	0.864 5
MTF-ResNet-CBAM	96.67	96.72	96.67	0.966 7

实验结果表明：MTF 转换在提升模型特征表达能力方面具有显著效果，模型准确率由 73.45%提高至 80.68%，说明时序信号可视化能够有效增强故障识别的区分度。进一步，引入注意力机制能够强化模型对关键信息的捕捉，其中 SE 模块使准确率提升至 86.45%，验证了注意力机制在特征选择中的优势。综合来看，完整的 MTF-ResNet-CBAM 模型取得了最佳性能，准确率达到 96.67%，相比原始 ResNet 提升约 23%，充分证明了 MTF 与 CBAM 的结合在复杂工况下具有突出的协同优势与鲁棒性。

5 结论

为解决复杂工况及噪声干扰条件下轴承故障诊断精度不足的难题，笔者提出一种基于 MTF-ResNet-CBAM 的智能诊断方法。该方法利用 MTF 将 1 维振动信号映射为 2 维图像，有效保留了时间序列的动态转移特征；结合 ResNet 的深层特征提取能力和 CBAM 的特征增强机制，实现了对轴承故障

的高精度识别。

实验结果表明：该模型在 4 类典型工况下均表现出优异的分类性能，整体准确率达到 96.67%，显著优于传统 ResNet、MTF-ResNet 及 MTF-ResNet-SE 等对比模型。消融实验进一步验证了 MTF 转换在时序特征表征中的有效性，以及 CBAM 在特征选择与冗余抑制方面的优势。综合分析可知，该模型在 Precision、Recall 与 F_1 -score 等多项评价指标上均保持较高水平，体现了良好的泛化能力与鲁棒性。

综上，MTF-ResNet-CBAM 模型为轴承故障诊断提供了一种新颖而有效的技术路径，不仅具有重要的理论意义，而且在兵工装备的健康监测、寿命管理及战备保障等方面展现出广阔的应用前景。未来工作可进一步探索多源信号融合与轻量化模型设计，以推动该方法在兵工装备智能化维护与复杂战场环境下的实际应用。

参考文献：

- [1] 王钊玮, 辛博, 胡睿杰, 等. 基于小波降噪和经验模态分解的旋翼电机轴承故障特征提取方法[J/OL]. 机电工程, 1-10[2025-09-09]. <https://link.cnki.net/urlid/33.1088.th.20250908.1120.004>.
- [2] 罗建, 安会平, 李宁. 基于 LabVIEW 与 Matlab 的滚动轴承故障诊断[J]. 兵工自动化, 2023, 42(9): 56-58, 82.
- [3] 惠永永, 毛高磊, 赵小强, 等. 基于 MIEC 与 MEAViT 的滚动轴承故障诊断方法[J/OL]. 铁道科学与工程学报, 1-14[2025-09-09]. <https://doi.org/10.19713/j.cnki.43-1423/u.T20250743>.
- [4] 郑则君, 宋冬利, 朱朝全, 等. 改进 SGMD-SCA 协同优化的轴箱轴承故障特征提取与诊断[J/OL]. 实验室研究与探索, 1-6[2025-09-09]. 10.19927/j.cnki.ssytt.2025.09.018.
- [5] 王子, 段争光, 王兵, 等. 基于度量学习的收割机滚动轴承故障诊断[J]. 中国农机化学报, 2025, 46(10): 112-119.
- [6] 林金亮, 刘瞰东, 张馨月, 等. 改进 EBCLS 模型在滚动轴承故障预测中的应用[J/OL]. 福州大学学报(自然科学版), 1-7[2025-09-09]. <https://doi.org/10.7631/issn.1000-2243.25036>.
- [7] 贺思艳, 任利娟, 田新诚. 基于卷积神经网络的滚动轴承故障诊断[J]. 兵工自动化, 2019, 38(3): 39-41, 54.
- [8] 马新娜, 张策, 李豪, 等. 复杂工况下的轴承故障诊断方法综述 [J/OL]. 软件导刊, 1-10[2025-09-09]. <https://doi.org/10.11907/rjdk.241564>.
- [9] 田健, 王强, 张亮, 等. 应用时频分析的内燃机曲轴轴承故障特征识别方法研究[J]. 内燃机与配件, 2025(16): 49-51.
- [10] 张永昌. 基于 CNN 与 NGO-SVM 的滚动轴承故障诊断研究[J]. 自动化应用, 2025, 66(15): 278-281.
- [11] 张波. 基于 MTF 图像和深度学习的滚动轴承故障识别研究[D]. 大连: 大连交通大学, 2025.
- [12] 王鹏, 邢高举, 牛浩平. 基于 ResNet50-CBAM 模型的滚动轴承故障诊断研究[J]. 科技创新与应用, 2025, 15(19): 1-4.
- [13] 杨敏, 孙文磊, 刘志远, 等. 基于 MTF-DCGAN 的齿轮箱故障诊断方法研究[J]. 机床与液压, 2025, 53(12): 17-24.
- [14] 金岩, 缪成翔. 基于 MTF-DARCNN 的滚动轴承故障诊断方法[J/OL]. 机电工程, 1-13[2025-09-09]. <https://link.cnki.net/urlid/33.1088.TH.20250605.1557.008>.
- [15] 方学宽, 张才, 白植志, 等. 基于 3Cur-ResNet50-GCN 的齿轮箱故障诊断模型[J]. 机床与液压, 2025, 53(14): 15-23.
- [16] 张春光. 改进 YOLOv5 网络的带钢表面缺陷检测方法[D]. 阜新: 辽宁工程技术大学, 2023.
- [17] 李琳辉, 张鑫亮, 付一帆, 等. 基于 TC-YOLOv7 算法的可见光与红外后融合检测研究[J]. 汽车工程, 2023, 45(12): 2280-2290.