

UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA

## MASTER THESIS



学科专业	电子信息
学 号	202222280517
作者姓名	王本浩
指导老师	周雪 教授
学 院	电子科技大学（深圳）高等研究院

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_

UDC 注 1 \_\_\_\_\_

# 学 位 论 文

面向长尾分布数据的自监督预训练故障诊断方法研究

(题名和副题名)

王本浩

(作者姓名)

指导老师

周雪 教授

电子科技大学 成都

(姓名、职称、单位名称)

申请学位级别 硕士 学科专业 电子信息

提交论文日期 \_\_\_\_\_ 论文答辩日期 \_\_\_\_\_

学位授予单位和日期 电子科技大学 年 月

答辩委员会主席 \_\_\_\_\_

评阅人 \_\_\_\_\_

注 1: 注明《国际十进分类法 UDC》的类号。

A Master Thesis Submitted to  
University of Electronic Science and Technology of China

Discipline:	<b>Electronic and Information Engineering</b>
Author:	<b>Wang Benhao</b>
Student ID:	<b>202222280517</b>
Supervisor:	<b>Dr. Xue Zhou</b>
School:	<b>Shenzhen Institute for Advanced Study</b>

## 摘 要

为了适应日益增长的宽带信号和非线性系统的工程应用，用于分析瞬态电磁散射问题的时域积分方程方法研究日趋活跃。本文以时域积分方程时间步进算法及其快速算法为研究课题，重点研究了时间步进算法的数值实现技术、后时稳定性问题以及两层平面波算法加速计算等，主要研究内容分为四部分。

.....

**关键词：**时域电磁散射，时域积分方程，时间步进算法，后时不稳定性，时域平面波算法



## ABSTRACT

With the widespread engineering applications ranging from broadband signals and non-linear systems, time-domain integral equations (TDIE) methods for analyzing transient electromagnetic scattering problems are becoming widely used nowadays. TDIE-based marching-on-in-time (MOT) scheme and its fast algorithm are researched in this dissertation, including the numerical techniques of MOT scheme, late-time stability of MOT scheme, and two-level PWTD-enhanced MOT scheme. The contents are divided into four parts shown as follows.

**Keywords:** Time-domain Electromagnetic Scattering, Time-domain Integral Equation, Marching-on In-time (MOT) Scheme, Late-time Instability, Plane Wave Time-domain (PWTD) Algorithm



# 目 录

第一章 绪 论 .....	1
1.1 研究工作的背景与意义 .....	1
1.2 长尾学习方法的国内外研究历史与现状 .....	2
1.2.1 长尾学习的发展历程和研究现状 .....	2
1.2.2 自监督预训练的发展历程和研究现状 .....	3
1.2.3 对比自监督学习的发展与孪生网络的应用 .....	4
1.3 本文的主要贡献与创新 .....	5
1.4 本论文的结构安排 .....	5
第二章 长尾学习及相关方法的理论基础与轴承故障数据集 .....	6
2.1 长尾学习相关理论 .....	6
2.2 自监督学习相关理论 .....	7
2.3 孪生网络与对比学习相关理论 .....	7
2.4 相关数据集介绍 .....	9
2.4.1 CWRU 数据集介绍 .....	9
第三章 基于孪生网络对比学习的自监督预训练方法研究 .....	11
3.1 模型整体架构及其模块设计 .....	11
3.1.1 简单暹罗孪生网络 .....	11
3.1.2 简单暹罗孪生故障诊断网络 .....	12
3.1.3 基于 CMA-ES 优化器的最优数据增强策略搜索算法 .....	16
3.2 实验与分析 .....	20
3.2.1 简单暹罗孪生网络实现细节 .....	20
3.2.2 基于 CMA-ES 优化器的数据增强策略搜索 .....	25
3.2.3 在 CWRU 上的实验结果 .....	25
3.2.4 对 SimSiam 网络模块的探讨 .....	28
3.2.5 对数据增强模块的探讨 .....	30
3.3 本章小结 .....	30
第四章 基于半监督学习和分类面调整算法的微调方法研究 .....	33
4.1 模型架构 .....	33
4.1.1 半监督学习模型架构 .....	33
4.2 实验与分析 .....	36



第五章 其他长尾学习方法分析 (CB LOSS 等) .....	38
5.1 .....	38
5.2 本章小结 .....	38
第六章 全文总结与展望 .....	39
6.1 全文总结 .....	39
6.2 后续工作展望 .....	39
致 谢 .....	40
附录 A 中心极限定理的证明 .....	41
A.1 高斯分布和伯努利实验 .....	41
参考文献 .....	42
攻读硕士学位期间取得的成果 .....	45

## 第一章 绪论

### 1.1 研究工作的背景与意义

旋转机械作为现代工业中的关键机械设备，长期运行在高温、疲劳、重载的复杂环境中。产生的故障可能会造成严重的事故，造成巨大的经济损失和人员伤亡。智能诊断是预测健康管理（PHM）的关键组成部分，PHM 是直升机、航空发动机、风力涡轮机、高速列车等各种旋转机械中最重要的系统之一，旨在有效检测故障。传统的智能诊断方法主要包括使用信号处理方法进行特征提取和使用机器学习方法进行故障分类，这些方法已经取得了相当大的进展。然而，面对异构海量数据，由专家设计和选择的特征提取方法和从信号到条件的映射能力在很大程度上取决于先验知识，既费时又需要经验。因此，如何更精确、更有效地进行诊断仍然是一个具有挑战性的问题。

现实生活中的故障检测数据通常服从长尾分布<sup>[1]</sup>长尾分布数据是一种偏态分布，即头部类包含了大部分正常数据，相反地，尾部类包含的故障数据量比较少。用长尾数据训练神经网络时，神经网络的性能易被头部类影响而在尾部类上表现较差，对尾部类样本的错分往往会带来更大的损失，对尾部类样本的研究更具有价值意义。因此如何基于长尾分布的故障检测数据训练故障检测模型是一个具有现实意义的难点问题<sup>[2]</sup>。目前，在视觉识别长尾学习领域有诸多前沿的长尾学习方法，如变分自编码器<sup>[3]</sup>、生成对抗网络<sup>[4]</sup>、CB(Class-Balanced) Loss<sup>[5]</sup>、自监督学习<sup>[6]</sup>、MARC 决策面调整算法<sup>[7]</sup>，而在故障诊断长尾学习领域常见的方案为半监督学习方法<sup>[1]</sup>、集成学习<sup>[8]</sup>、样本加权<sup>[2]</sup>，少有前者应用前沿的长尾学习方法在故障诊断长尾学习领域。因此，基于前沿的长尾学习方法提出一个适用于故障诊断领域的长尾学习模型是一个具有现实意义的挑战。

目前，大部分神经网络的训练仍然使用的是有监督范式，需要耗费大量的标注数据，标注这些数据是非常耗时费力的。而自监督的提出就是为了打破对人工标注的依赖，即使在没有标注数据的情况下，也可以高效地训练网络。不同于传统的非监督学习，自监督学习通过人工构造输入数据的代理标签，然后依赖于代理标签设计神经网络的监督训练任务。神经网络从而学习到特征表示。文献 [9] 指出额外的自监督预训练能够使常规的长尾学习算法获得更高的性能，且验证了自监督预训练在 CIFAR-LT、ImageNet-LT 等视觉识别长尾数据集上的性能提升。目前，在故障诊断长尾学习领域常见的方案为半监督学习方法、集成学习、样本加权，而少有前者应用前沿自监督预训练方法在故障诊断长尾学习领域。因此，基

于自监督预训练的长尾数据故障诊断是一个具有现实意义的挑战。

## 1.2 长尾学习方法的国内外研究历史与现状

### 1.2.1 长尾学习的发展歷程和研究现状

国内外处理长尾问题通常从三个方面入手。一、样本层面，减少样本数量的欠采样如随机欠采样、NearMiss<sup>[10]</sup>、ENN<sup>[11]</sup>；增加样本数量的过采样，如随机过采样、SMOTE<sup>[12]</sup>、变分自编码器<sup>[3]</sup>、生成对抗网络<sup>[4]</sup>。二、损失函数层面，使用类平衡的损失函数如以数据频率的逆加权<sup>[13]</sup>，OHEM(Online Hard Example Mining)，Focal Loss<sup>[14]</sup>以及最新的 CB(Class-Balanced) Loss<sup>[5]</sup>。三、模型层面，如重复组合少数类样本与抽样的同样数量的多数类样本进行集成学习、半监督和自监督学习<sup>[6]</sup>、MARC 决策面调整算法<sup>[7]</sup>。

目前，基于长尾分布问题，国内外做出了许多研究。例如：Liu J 等人<sup>[15]</sup>，介绍了一种基于特征云的数据增强手段，从方差较大的数据分布宽广的头部类中学习类内多样性并转移到尾部类，并应用于人脸识别。Yang Y 等人<sup>[6]</sup>，介绍了不平衡样本的标签在模型训练中的价值，并提出了用半监督的学习方法使用原模型识别无标签的数据，并以获得伪标签的数据扩充原数据集，同时证明了自监督的模型预训练的有效性，且样本维度越高对性能提升越有效。吴磊等人<sup>[16]</sup>提出了一种面向长尾图像的个性化专家识别算法。该算法在残差网络的基础上构建多专家学习网络，并加入个性化学习模块、信息融合模块与个性化信息增强模块，以两阶段的学习方式提升长尾图像整体的识别准确率与中、尾部类别的识别效果。吴亮等人<sup>[8]</sup>构建了多级分组分类器，提升尾部类性能的同时，避免头部类性能损失。然后设计了基于多头注意力机制的分组特征重融合模块，为多级分类器输入更精细的特征。最后，基于多级分类器提出了 Logit 联合调整的方法，以缓解组间的不平衡。Wang Y 等人<sup>[7]</sup>介绍了一种对长尾视觉识别学习有效的 MARC 决策面调整算法，对原模型输出的预测分数额外做一次训练——只用三行代码。Cui Y 等人<sup>[5]</sup>，介绍了 CB(Class-Balanced) 损失函数，在原损失函数上乘以与“独特”样本数量有关的因子从而缓解类别不平衡。CB 因子与超参数  $\beta$  有关，平滑了由 1 到最常见的不平衡因子  $1/n$ 。综上，文献 [6] 为自监督预训练模型应用在故障诊断长尾学习领域提供了良好的启示，但半监督学习方法在现实的长尾学习模型可能不适用。文献 [8, 16] 介绍的集成学习方法削弱样本不平衡效应是常规方法，但单个模型在训练时若融合自监督预训练的方法，性能可能会得到提升。文献 [7] 介绍的决策面调整算法对一般的长尾学习模型都适用，但在故障诊断模型上是否可行有待商榷，且该算法对模型性能提升的理论性解释有待完善。文献 [5] 介绍的损失函数

有着新颖的思路,但超参数  $\beta$  需要人工选择在应用上可能存在难点。以上大部分为学者在视觉长尾学习领域的研究,此类算法在 CIFAR-LT、ImageNet-LT 等视觉识别数据集取得了较好的效果,而故障诊断长尾学习领域少有前沿的长尾学习算法的应用。因此,前沿的长尾学习算法在故障诊断长尾学习领域的应用具有很高的研究价值。

### 1.2.2 自监督预训练的发展历程和研究现状

2019 年, MoCo<sup>[17]</sup> 的提出掀起了视觉自监督学习的热潮,随后 SimCLR<sup>[18]</sup>、BYOL<sup>[19]</sup>、SwAV<sup>[20]</sup> 等主流自监督学习算法相继问世,使该领域呈现出百花齐放、百家争鸣的繁荣景象。2021 年末, MAE<sup>[21]</sup> 的提出进一步推动了自监督学习的发展,使其达到了一个前所未有的新高度。然而,在这背后,自监督学习经历了长期的迭代与发展。

目前,国内外学者针对自监督预训练开展了大量研究。例如,在计算机视觉领域, Yang 等人<sup>[6]</sup> 研究了不平衡样本的标签信息在模型训练中的价值,并提出了一种半监督学习方法,该方法利用原模型识别无标签数据并生成伪标签,从而扩充原始数据集,并验证了自监督预训练的有效性。此外,该研究发现,样本维度越高,性能提升越显著,这为自监督预训练在故障诊断长尾学习领域的应用提供了有力支持。Doersch 等人<sup>[22]</sup> 提出了基于局部图像块位置预测的自监督任务,即随机抽取图像中的两个块,并训练模型预测一个块相对于另一个块的位置。Zhang 等人<sup>[23]</sup> 构建了基于图像上色的自监督任务,将图像转换至 CIE Lab 颜色空间,并利用 L 通道作为输入,让模型预测 a 和 b 通道的颜色信息。Gidaris 等人<sup>[24]</sup> 通过人为旋转图像不同角度,并将旋转角度作为监督信号训练模型,以此实现图像语义特征的自监督学习。

近年来,自监督预训练方法也逐步应用于故障诊断领域,并受到广泛关注。例如, Zhang 等人<sup>[25]</sup> 提出了基于先验知识的自监督任务,并利用卷积自编码器进行训练,在小样本学习任务上取得了良好表现。Senanayaka 等人<sup>[26]</sup> 采用 One-Class SVM 输出的标签作为代理标签,并使用自监督预训练的卷积神经网络进行故障特征提取。W. Zhang 等人<sup>[9]</sup> 提出了一种基于信号块交换的自监督任务,即通过打乱时域信号块顺序构造“伪数据”,并训练模型区分原始数据和伪数据。实验结果表明,该方法在特征提取任务中表现优异。

综上,尽管<sup>[22-24]</sup> 等研究在计算机视觉领域取得了重要进展,但其方法在故障诊断任务中的直接应用仍存在一定困难。然而,这些研究的思路可为故障诊断领域的自监督预训练任务设计提供启示。例如,<sup>[26]</sup> 提出的 One-Class SVM 生成的

标签可结合聚类算法,使其适用于多分类任务; [9] 提出的信号块交换方法与 [22] 提出的图像块位置预测任务在思路具有一定相似性,可探索其结合方式。此外, [9,25,26] 提出的自监督预训练方法在任务设计上仍存在较大创新空间,且在长尾学习任务中的应用尚不充分。因此,故障诊断领域的自监督预训练仍处于发展初期,借鉴计算机视觉领域的前沿自监督任务,在现有基础上提出更具创新性、更高性能的自监督预训练任务,仍具有重要的研究价值和挑战。

### 1.2.3 对比自监督学习的发展与孪生网络的应用

自监督学习的核心目标是在无人工标注的情况下自动提取有效特征 [27]。近年来,对比自监督学习 (contrastive self-supervised learning) 作为无监督学习的一个重要分支,逐渐成为研究热点,并在图像识别任务中取得了最先进的 (SOTA) 性能。许多相关算法相继被提出,例如动量对比学习 (Momentum Contrast, MoCo) [17]、对比预测编码 (Contrastive Predictive Coding, CPC) [28] 以及用于视觉表征学习的简单对比学习框架 (SimCLR) [18]。

这些方法的核心思想是通过拉近相同样本的不同视图 (正样本对) 并推远不同样本 (负样本对) 来学习有意义的特征表示。这类对比学习方法通常依赖存储在记忆库 (memory bank) 中的大量负样本,以提高对比学习效果 [29]。此外,一些其他对比学习方法不依赖负样本,而是计算两个正样本对之间的相似性,例如 SwAV (Swapping Assignments between multiple Views) [20]、BYOL (Bootstrap Your Own Latent) [19] 以及 SDCT (Self-supervised Deep Correlation Tracking) [30]。

在对比自监督学习中,孪生网络 (Siamese Network) 因其适用于比较不同样本的网络结构,具有独特优势。Laine 和 Aila [31] 提出了基于孪生网络的  $\pi$ -model,用于训练深度神经网络,该方法在半监督学习场景下依赖不同的正则化规则与数据增强策略。Zheng 和 Yang [32] 提出了一种“记忆正则化”机制,其核心思想是让主模型自身 (而非外部模块) 学习域内知识,以实现无监督场景自适应。Zhao 等人 [33] 提出了基于相互学习 (mutual learning) 和知识蒸馏 (knowledge distillation) 的训练方法,以提高视觉目标跟踪任务的性能。

上述研究均在各自领域中引入孪生网络的思想,以解决特定问题。本文受到 SimSiam [34] 的启发,SimSiam 采用简化的孪生网络结构,并在计算机视觉 (CV) 任务中展现出卓越性能。基于此,我们提出了一种新颖的轴承故障诊断框架,该框架可适用于由不同神经网络层构建的多种模型。

与其他主流方法相比,SimSiam 具有以下三大优势:

1. 无需构造负样本对;

2. 无需采用大批量训练;
3. 无需使用动量编码器 (momentum encoder)。

### 1.3 本文的主要贡献与创新

本论文以时域积分方程时间步进算法的数值实现技术、后时稳定性问题以及两层平面波加速算法为重点研究内容, 主要创新点与贡献如下:

### 1.4 本论文的结构安排

本文的章节结构安排如下:

## 第二章 长尾学习及相关方法的理论基础与轴承故障数据集

### 2.1 长尾学习相关理论

不平衡数据在现实世界中无处不在，大规模数据集通常表现为长尾分布。长尾分布的学习亦是故障诊断领域中常见挑战，例如正常工作状态多，故障工作状态少，或是常见故障状态和罕见故障状态的样本个数差异。现实生活中的故障检测数据通常服从长尾分布，长尾分布数据是一种偏态分布，如图2-1所示，在现实生活中，大规模的训练样本呈现典型的长尾分布，即头部类别的样本个数较多，而尾部类别的样本个数较少，样本个数逐渐降低的分布。用长尾数据训练神经网络时，神经网络的性能易被头部类影响而在尾部类上表现较差，对尾部类样本的错分往往会带来更大的损失，对尾部类样本的研究更具有价值意义。因此如何基于长尾分布的故障检测数据训练故障检测模型是一个具有现实意义的难点问题。

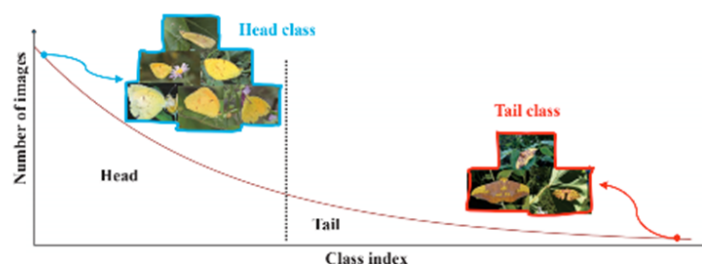


图 2-1 长尾分布示意图

长尾分布的样本具有非常显著的样本不均衡问题，导致训练好的模型易偏向于头部类，从而损害了对尾部类的识别效果。如图2-2(a)所示，样本个数均衡时，各个类别之间具有清晰的区分边界，各类都占据了宽广的特征空间。而当某些类的样本个数减少呈现样本分布失衡的长尾分布时，如图2-2(b)所示，尾部类在特征空间的分布十分狭窄，并依附在头部类附近，从而扭曲了特征空间，损害了类间的多样性和区分性。这对现代深度学习框架提出了一个重大挑战，即使使用数据重采样方法或类平衡损失等专门技术，在极端类不平衡的情况下，仍然会出现显著的性能下降。因此，为了进一步应对这一挑战，了解类不平衡学习所带来的不同特征至关重要。然而，与平衡数据不同的是，不平衡学习背景下的标签扮演着令人惊讶的有争议的角色，这导致了标签价值的持续困境：一方面，有标签监督的学习算法通常比无监督的学习算法产生更准确的分类器，这表明了标签的积极价值；然而，另一方面，不平衡的标签在学习过程中自然会产生“标签偏差”，其中大多数类别可以显著地驱动决策边界，表明标签的负面影响。因此，不平衡的

标签似乎是一把双刃剑。

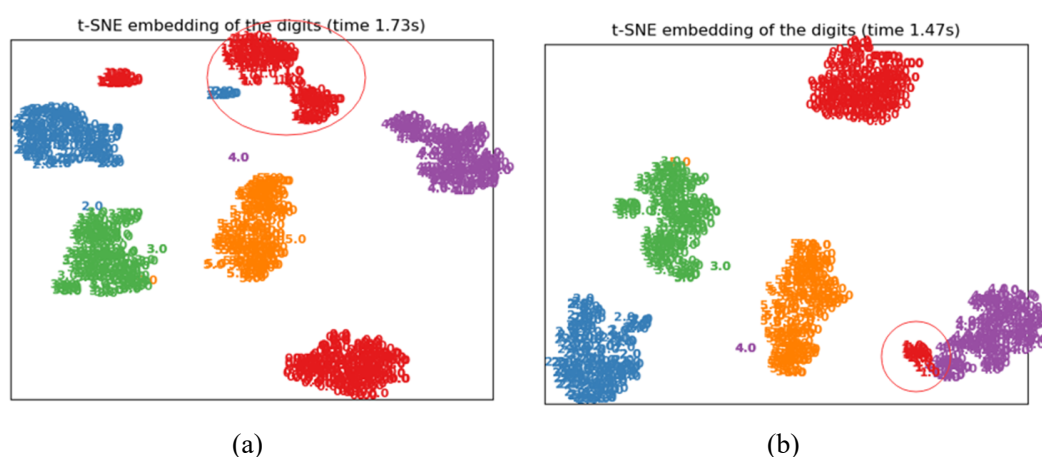


图 2-2 均匀/长尾分布数据的 tsne 图 (a) 样本均匀分布时的 T-SNE 分布图；  
(b) 尾部类“1”类样本个为 30，头部类样本个数为 180 个时的 T-SNE 分布图

## 2.2 自监督学习相关理论

自监督学习是一种范式，利用“代理任务” (pretext) 从大规模的无标签数据本身的内在结构挖掘通用的表征信息，通过这种构造的监督信息对网络进行训练，从而可以学习到对下游任务有价值的表征。如图2-3所示，自监督学习分为两个阶段：(1) 自监督预训练阶段。完全放弃标签信息，专注于数据本身，通过自监督学习进行预训练。从不平衡数据中学习标签不可知的特征表示，减少类别偏差对初始化的影响。(2) 微调阶段。使用第一阶段中通过自监督预训练的网络权重作为初始化。在此基础上，可以使用任何标准的不平衡学习技术进行进一步训练，学习最终的分类模型。由于预训练阶段和正常训练阶段是独立的，自监督学习可以与现有的不平衡学习方法无缝结合，增强模型性能，且自监督预训练阶段不依赖标签，使得网络能学习到更通用、更鲁棒的特征表示，从而避免了类别不平衡对特征学习的负面影响。

(todo: 1. 补充公式理论 Rethinking the Value of Labels for Improving Class-Imbalanced Learning)

## 2.3 孪生网络与对比学习相关理论

对比学习是一种学习方法，侧重于通过对比正反两方面的实例来提取有意义的表征。它利用的假设是，在学习到的嵌入空间中，相似的实例应靠得更近，而



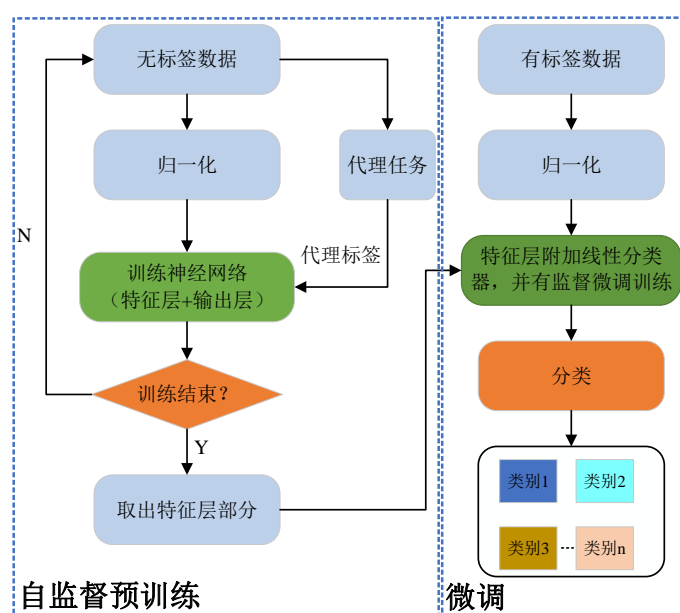


图 2-3 自监督训练范式流程

不相似的实例应离得更远。通过将学习作为一项辨别任务，对比学习允许模型捕捉数据中的相关特征和相似性。对比学习主要分为监督对比学习和自监督对比学习两大类。

监督对比学习是对比学习的一个分支，它利用标记数据来明确训练模型以区分相似和不相似的实例。在监督对比学习中，模型在数据点对及其标签上进行训练，指示数据点是否相似或不相似。目标是学习一个表示空间，其中相似的实例聚集得更近，而不同的实例则被推开。一种流行的方法是信息噪声对比估计（InfoNCE）损失函数。InfoNCE 损失在学习的表示空间中最大化正样本之间的一致性并最小化负样本之间的一致性。通过优化此目标，模型学会区分相似和不相似的实例，从而提高下游任务的性能。

自监督对比学习采用不同的方法，从未标记的数据中学习表示，而不依赖于显式标签。自监督对比学习利用借口任务的设计，从未标记的数据创建正负对。这些借口任务经过精心设计，旨在鼓励模型捕获数据中有意义的特征和相似性。自监督对比学习中常用的借口任务之一是生成增强视图。这涉及创建同一实例的多个增强版本并将它们视为正对，而来自不同样本的实例则被视为负对。通过训练模型来区分这些对，它可以学习捕获更高级别的语义信息并很好地推广到下游任务。(todo: 补充监督对比学习和自监督对比学习框图)

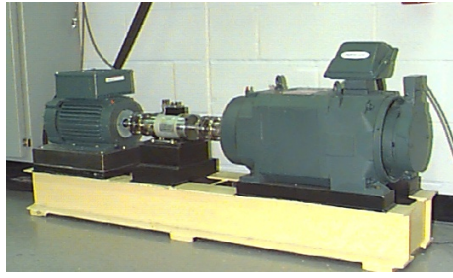


图 2-4 CWRU 轴承测试设备

## 2.4 相关数据集介绍

### 2.4.1 CWRU 数据集介绍

数据集由 CWRU (Case Western Reserve University) 提供。图 2-4 显示了测试设备，该设备包含一台 2 马力电机、一个扭矩编码器、一个测功机以及一些控制电路。测试使用的轴承为 6205-2RS JEM 深沟球轴承，分别安装在电机的驱动端和风扇端。

轴承故障是通过单点电火花放电加工 (Electro-discharge Machining) 制造的。根据故障位置，故障类型可分为以下四种：1) 外圈故障 (Outer Raceway Fault, OF); 2) 内圈故障 (Inner Raceway Fault, IF); 3) 滚动体故障 (Roller Faults, RFs); 4) 正常状态 (Normal Condition, NC)。轴承故障的严重程度可由故障直径 (Fault Diameter) 描述，故障直径指轴承或其他机械部件上出现的故障或损伤的直径尺寸。故障直径通常用来描述故障的大小和程度，对于故障诊断和预测维护非常重要。

本研究采集了安装在驱动端的轴承数据，采样频率为 12 kHz，故障直径分别为 0.007 英寸，0.014 英寸和 0.021 英寸的内圈故障、外圈故障、滚动体故障。不同故障类型的波形如图 2-5 所示。有关该数据集的更多详细信息，可访问 CWRU 轴承数据中心网站<sup>[35]</sup>。

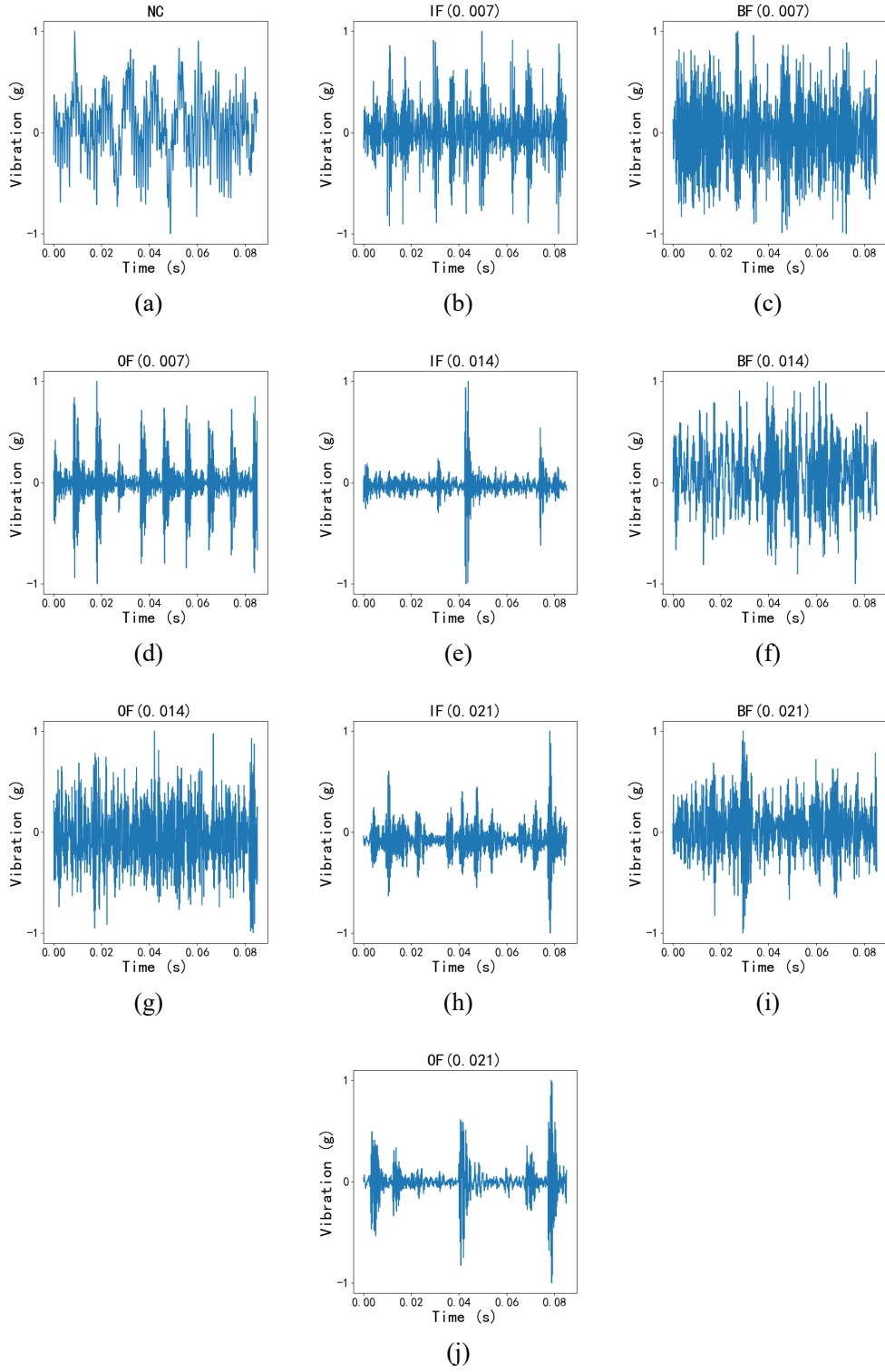


图 2-5 每个类别的样本信号: **(a)** NC; **(b)** IF(0.007); **(c)** BF(0.007); **(d)** OF(0.007); **(e)** IF(0.014); **(f)** BF(0.014); **(g)** OF(0.014); **(h)** IF(0.021); **(i)** BF(0.021); **(j)** OF(0.021)。

## 第三章 基于孪生网络对比学习的自监督预训练方法研究

### 3.1 模型整体架构及其模块设计

#### 3.1.1 简单暹罗孪生网络

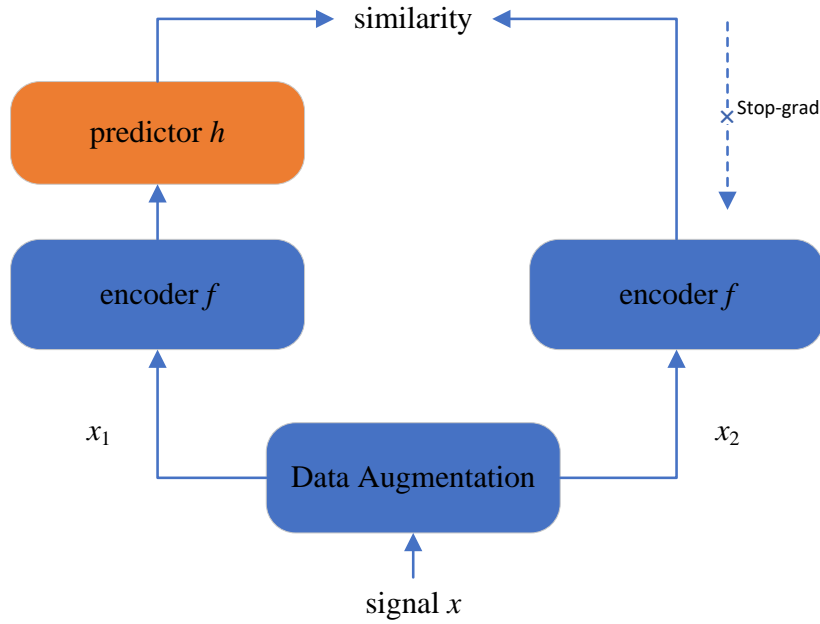


图 3-1 简单暹罗孪生网络

简单暹罗孪生网络（SimSiam——Simple Siamese）<sup>[34]</sup> 架构（图3-1）将信号  $x$  中的两个随机增强视图  $x_1$  和  $x_2$  作为输入。这两个视图由编码器网络  $f$  处理，该网络由骨干网络（例如 ResNet<sup>[36]</sup>）和投影 MLP 头部组成。编码器  $f$  在两个视图之间共享权重。预测 MLP 头部  $h$  转换一个视图的输出并将其与另一个视图进行匹配。将两个输出向量表示为  $p_1 \triangleq h(f(x_1))$  和  $p_2 \triangleq f(x_2)$ ，我们最小化它们的负余弦相似度：

$$\mathcal{D}(p_1, z_2) = -\frac{p_1}{\|p_1\|_2} \cdot \frac{z_2}{\|z_2\|_2} \quad (3-1)$$

其中， $\|\cdot\|_2$  是  $\ell_2$  范数。定义对称损失函数

$$\mathcal{L} = \frac{1}{2}\mathcal{D}(p_1, z_2) + \frac{1}{2}\mathcal{D}(p_2, z_1) \quad (3-2)$$

该损失函数作用于单段信号，总损失值在计算所有信号的损失后取平均值。损失

表 3-1 简单暹罗孪生网络的伪代码，用 Pytorch 描述

**简单暹罗孪生网络 Pytorch 伪代码**

```
#f: 骨干网络 + 投影 MLP
#h: 预测 MLP
for x in loader: # 加载一个包含 n 个样本的小批量数据 x
    x1, x2 = aug(x), aug(x) # 随机数据增强
    z1, z2 = f(x1), f(x2) # 投影, 形状为 n-by-d
    p1, p2 = h(z1), h(z2) # 预测, 形状为 n-by-d
    L = D(p1, z2)/2 + D(p2, z1)/2 # 损失函数

    L.backward() # 反向传播
    update(f, h) # SGD 更新

def D(p, z): # 负余弦相似度
    z = z.detach() # 停止梯度
    p = normalize(p, dim=1) # 对 p 进行 L2 归一化
    z = normalize(z, dim=1) # 对 z 进行 L2 归一化
    return -(p * z).sum(dim=1).mean() # 计算负余弦相似度
```

函数的最小值为-1。

方法中一个重要的组件是梯度停止（stop-grad）操作（图 1）。我们通过修改公式 (3-1) 来实现它：

$$D(p_1, \text{stopgrad}(z_2)) \quad (3-3)$$

这意味着在这一项中， $z_2$  被视为常数。类似地，公式 (3-2) 的实现形式为：

$$\mathcal{L} = \frac{1}{2}D(p_1, \text{stopgrad}(z_2)) + \frac{1}{2}D(p_2, \text{stopgrad}(z_1)) \quad (3-4)$$

第一项中  $x_2$  的编码器不会从  $z_2$  接收梯度，但在第二项中会从  $p_2$  接收梯度（反之亦然，对于  $x_1$  也是如此）。即在这一项中， $z_2$  被视为常数。

简单暹罗孪生网络的伪代码如表3-1所示。

### 3.1.2 简单暹罗孪生故障诊断网络

提出的用于故障诊断的简单暹罗孪生网络如图3-2所示，以下是架构的介绍。

1) **数据增强 (Data Augmentations):** 在对比表示学习算法中，数据增强 (DA) 起着至关重要的作用。成功地从振动信号中提取可区分的故障特征，依赖于为同一样本生成不同视图的质量。我们知道，面向图像的最新 (SOTA) 对比表示学习算法广泛使用多种数据增强方法，而针对序列数据的增强方法则相对较少。此外，用于对比学习生成正样本对的增强方法必须满足一个条件，即加入样本的噪声不

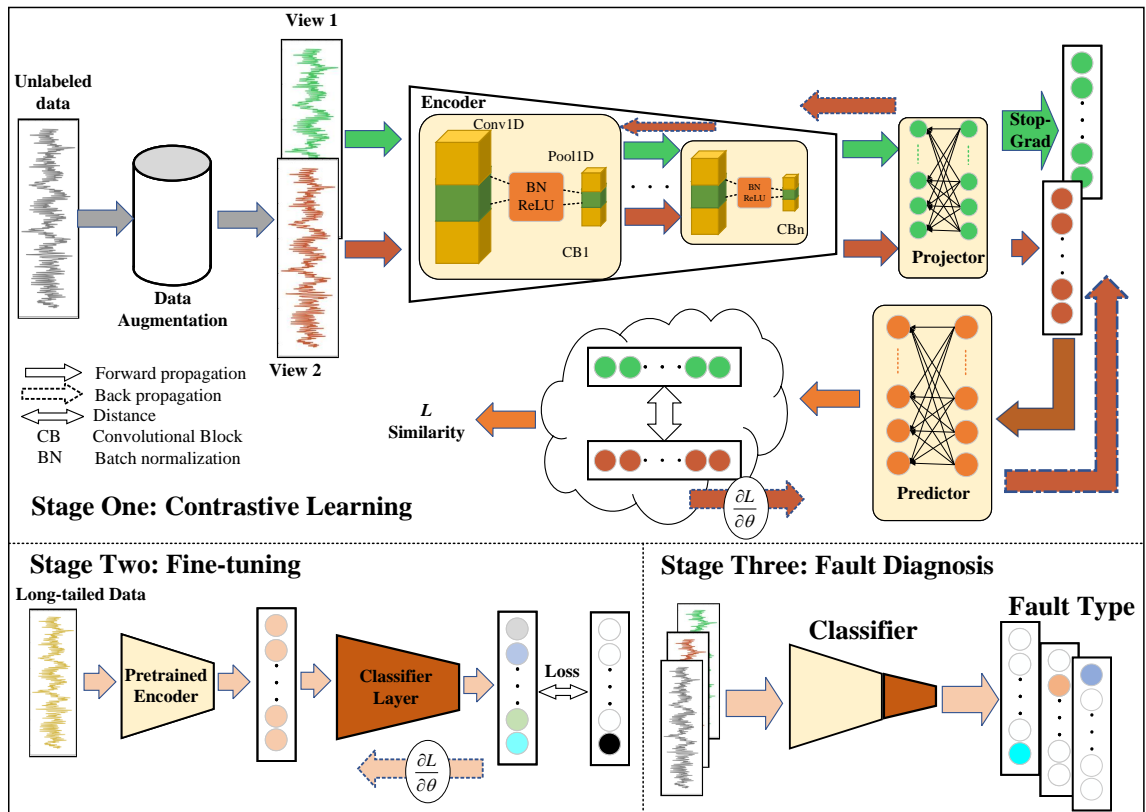


图 3-2 简单孪生故障诊断网络

能改变其语义含义。因此，我们选择了以下九种数据增强方法：

- **掩码 (Random Masked)**: 用于遮盖输入数据的一部分，即随机挑选信号的某些部分用 0 替代，通常是在序列数据或图像数据中随机选择部分区域进行掩码处理。这有助于模型在训练时学会忽略一些无关信息，从而提高其鲁棒性。
- **添加高斯噪声 (Adding Gaussian Noise)**: 将高斯噪声添加到数据中的方法，常用于增强模型的鲁棒性。通过这种方式，模型在训练过程中能够适应噪声，并提升其在实际应用中的性能，尤其是在存在噪声的环境中。
- **相位扰动 (Phase Perturbation)**: 修改信号的频率域中的相位信息，而保持幅度不变来生成新的数据样本。这种方法保留了信号的整体结构，但引入了细微的扰动，用于提高模型的泛化能力。
- **块打乱 (RandomChunkShuffle)**: 将时间序列数据分割成多个块，然后随机打乱这些块的顺序。这样可以创建不同的序列变体，增加模型对数据变异的适应能力，同时保持整体的语义不变。
- **随机缩放 (Random Scaled)**: 对数据进行随机幅度缩放来增强数据的方法。

通过改变数据的尺度，模型能够学习到不同幅度下的数据模式，从而提升其泛化能力。

- **随机绝对值 (Random Abs):** 对数据应用绝对值操作，将负值转换为正值或去掉负号。这个方法帮助模型处理包含负值的情形，增强其鲁棒性。
- **竖直翻转 (Random Vertical Flip):** 随机地将数据进行竖直翻转来创建新的样本。这对于一些对竖直方向变化不敏感的任务非常有效。
- **水平翻转 (Random Horizontal Flip):** 随机将图像或数据进行水平翻转，生成新的样本。这种方法对于对称性较强的数据非常有效，能增强模型的鲁棒性。
- **时移 (Time Shift):** 将数据在时间轴上平移一定的时间步长来生成新的样本。这种方法可以帮助模型学会适应时间序列数据中事件的变化位置，提升其对时间依赖的理解能力。

同一输入信号分别经过上述数据增强方法后的视图如图3-3所示。

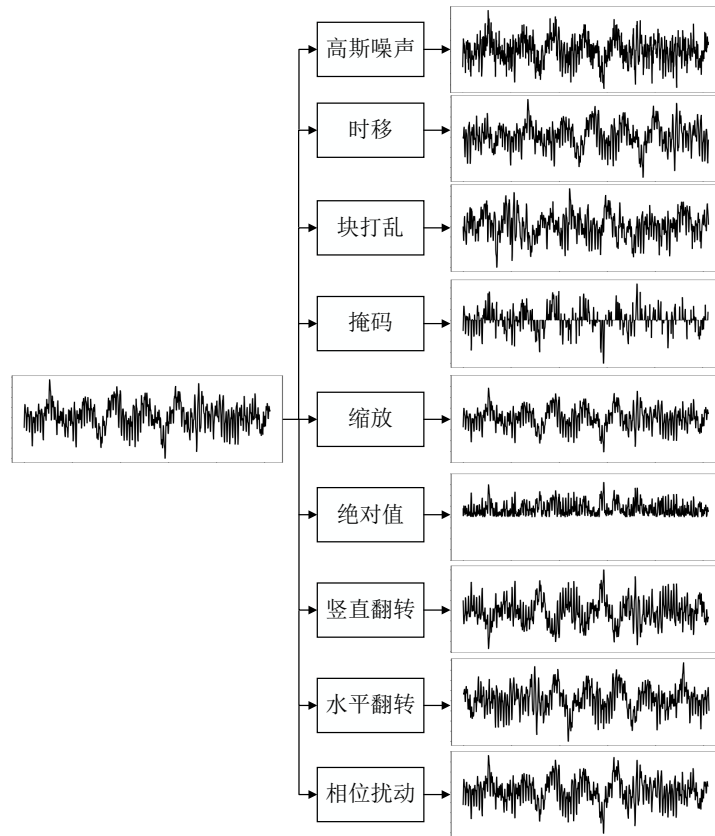


图 3-3 数据增强示意图

2) **编码器 (Encoder), 预测器 (Predictor) 和分类器 (Classifier):** 编码器是该算法中故障诊断的关键部分。我们主要使用其潜在编码来完成分类任务。由

于简单孪生网络是一个框架，我们可以根据需要构建编码器模型和预测器，其中预测器是一个相对简单的非线性函数。在第二阶段，需要在预训练的编码器后添加一个分类器层用于故障诊断，如图3-2所示。为了方便起见，我们使用多个常规的卷积神经网络（CNN）块来构建编码器，并使用全连接层来构建预测器和分类器。每个卷积块（CB）包括四个层：一个 1-D 卷积层  $f_{\text{BN}}$ 、一个激活函数层  $f_{\text{ReLU}}$ ，以及一个池化层  $f_{\text{Pool}}$ 。方程 (3-5) 展示了 1-D 卷积层的操作，其中输入的形狀为  $(N, C_{\text{in}}, L)$ ，输出的形狀为  $(N, C_{\text{out}}, L_{\text{out}})$ ，其中  $N$  是样本的数量， $C$  表示通道数， $L$  表示数据的长度。

$$\text{out}(N_i, C_{\text{out}_j}) = \text{bias}(C_{\text{out}_j}) + \sum_{k=0}^{C_{\text{in}}-1} \text{weight}(C_{\text{out}_j}, k) \otimes \text{input}(N_i, k). \quad (3-5)$$

BN 层  $f_{\text{BN}}$  用于加速训练过程并减少由于层输入分布不同而导致的内部协变量偏移（ICS）的影响<sup>[37]</sup>，当模型较深时，这种影响尤为严重。公式 (3-6) 展示了批归一化的过程。

$$f_{\text{BN}}(x_i, \mathbf{x}) = \gamma \frac{x_i - \mathbb{E}(\mathbf{x})}{\sqrt{\sigma(\mathbf{x})^2 + \varepsilon}} + \beta \quad (3-6)$$

$$\mathbb{E}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N x_i \quad (3-7)$$

$$\sigma(\mathbf{x}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mathbb{E}(\mathbf{x}))^2} \quad (3-8)$$

其中  $\mathbf{x}$  是整个小批量（mini batch）的向量，批量大小为  $N$ ， $x_i$  表示小批量中的第  $i$  个样本。 $\mathbb{E}(\mathbf{x})$  和  $\sigma(\mathbf{x})$  分别表示向量批量的均值和方差，如公式 (3-7) 和 (3-8) 所示。 $\varepsilon$  是一个超参数，通常设置为  $10^{-5}$ ，而  $\gamma$  和  $\beta$  分别是缩放因子和偏移因子，它们是可学习的参数，初始值分别设置为 1 和 0。

激活层  $f_{\text{ReLU}}$  可以表示为公式 (3-9)：

$$f_{\text{ReLU}} = \max(0, x) \quad (3-9)$$

$x_i^{(n)}$  表示通过第  $n$  个卷积块（CB）的第  $i$  个输出向量，可以用公式 (3-10) 表示：

$$x_i^{(n)} = f_{\text{Pool}}\{f_{\text{ReLU}}[f_{\text{BN}}(f_{\text{conv}}(\mathbf{x}^{(n-1)}), f_{\text{conv}}(x_i^{(n-1)}))]\} \quad (3-10)$$

**3) 优化器 (Optimizer)：**模型的训练不需要使用大批量优化器，例如分层自适应速率缩放（LARS）<sup>[38]</sup>，因为所提出的方法可以在典型批量大小下工作，而不



依赖于大批量训练。我们使用 SGD 优化器训练模型。网络参数通过公式 (3-11) 更新。

$$\theta_{l+1} = \theta_l - \eta_l \cdot \nabla_{\theta} \mathcal{L}(\mathbf{x}; \theta_l) \quad (3-11)$$

其中  $\theta_l$  表示时间  $t$  时的可学习参数,  $\eta_l$  表示时间  $t$  时的学习率,  $L(\cdot)$  表示损失函数。

4) **微调 (Fine-tuning)**: 在对比学习阶段完成后, 我们从模型中取出训练好的编码器  $f$ , 并将其与一个多层感知器 (MLP) 模块连接, 我们用带标签的数据微调构建一个完整的分类器用于故障诊断任务, 从而使模型获得分类能力。以准确率为性能指标, 如公式 (3-12):

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FP + FN} \quad (3-12)$$

其中,  $TN$  表示真阴性, 即被正确预测为阴性的阴性样本数量。真阳性 ( $TP$ ) 表示被模型正确判断为阳性的阳性样本数量。假阳性 ( $FP$ ) 表示被错误分类为阳性的阴性样本数量。假阴性 ( $FN$ ) 表示被错误分类为阴性的阳性样本数量。

当验证集的样本均匀采样后 (本研究采用此方法), 准确率等于宏平均召回率 (Macro-Averaged Recall)。宏平均召回率是对每个类别单独计算召回率, 然后取平均值:

$$\text{Macro-Recall} = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i} \quad (3-13)$$

其中,  $C$  是类别的总数,  $TP_i$  和  $FN_i$  分别是第  $i$  类的真正例和假负例。

5) **故障诊断 (Fault Diagnosis)**: 在完成微调后, 即可将模型应用于真实输入信号的故障诊断, 以预测故障类别。简单暹罗孪生故障诊断网络的训练过程如图3-4所示。首先, 我们需要对所有未标记的振动样本进行归一化预处理, 并使用数据增强方法生成输入样本的两个不同视图作为正样本对。我们将正样本对输入模型, 然后按照算法表3-1所示的流程训练模型。训练过程完成后, 我们取出编码器, 在其最后一层后附加一个线性分类层, 并使用标记样本对模型进行微调。线性分类层也可以用其他分类器 (如 K 近邻 (KNN) 和支持向量机 (SVM)) 替换。

### 3.1.3 基于 CMA-ES 优化器的最优数据增强策略搜索算法

简单暹罗孪生网络的一个难以接受的平凡解是所有输出都“坍塌”为一个常数。简单暹罗孪生网络利用数据增强生成两个视图, 并直接最大化同一输入的两个视图的相似性, 即不使用负样本对。因此简单暹罗孪生网络极大程度上依赖于数据增强的效果, 而人工筛选数据增强十分耗时耗力且难以获得最优解。该节提

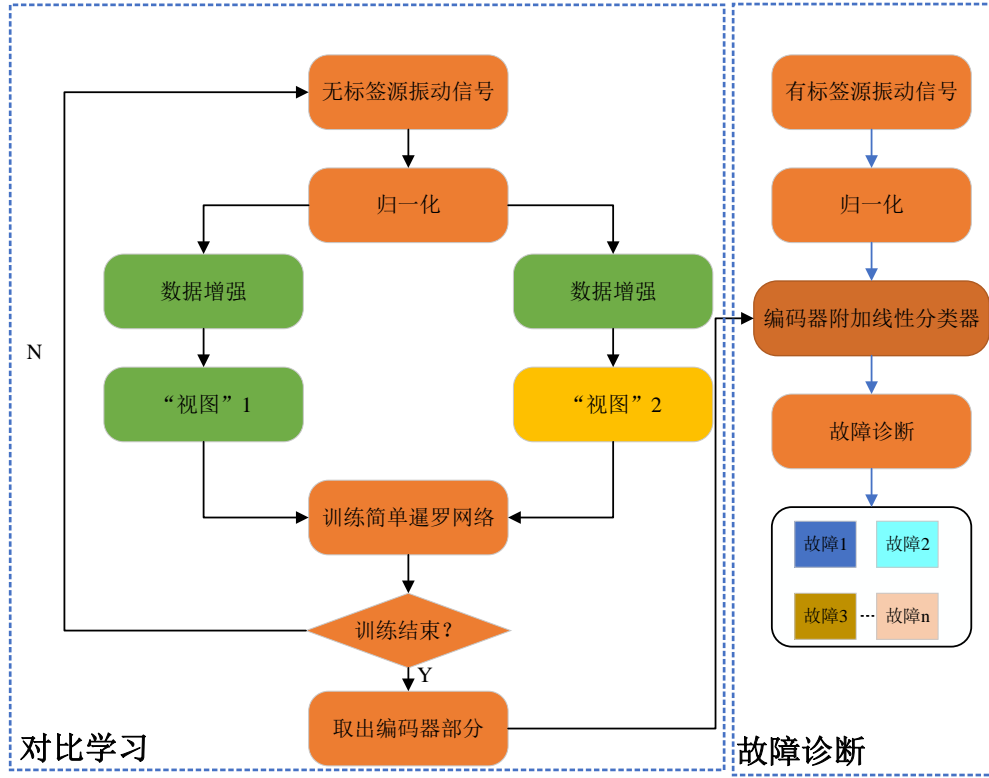


图 3-4 简单孪生故障诊断网络的训练流程图

出了基于 CMA-ES 优化器的最优数据增强策略搜索算法，CMA-ES（Covariance Matrix Adaptation Evolution Strategy，协方差矩阵自适应进化策略）是一种用于连续优化问题的进化算法。它是一种基于种群的优化方法，通过自适应调整协方差矩阵来指导搜索方向，从而在高维空间中高效地找到全局最优解。CMA-ES 在解决非线性、非凸、高维优化问题时表现出色，广泛应用于机器学习、工程优化和科学研究中。CMA-ES 的更新规则如下：

1. 采样新解：

$$\mathbf{x}_k^{(g+1)} = \mathbf{m}^{(g)} + \sigma^{(g)} \cdot \mathcal{N}(0, \mathbf{C}^{(g)}) \quad (3-14)$$

其中：

- $\mathbf{x}_k^{(g+1)}$  是第  $g+1$  代中的第  $k$  个候选解，
- $\mathbf{m}^{(g)}$  是第  $g$  代的均值向量，
- $\sigma^{(g)}$  是第  $g$  代的步长，
- $\mathbf{C}^{(g)}$  是第  $g$  代的协方差矩阵，
- $\mathcal{N}(0, \mathbf{C}^{(g)})$  是从多元正态分布中采样的随机向量。

2. 更新均值:

$$\mathbf{m}^{(g+1)} = \sum_{i=1}^{\mu} w_i \mathbf{x}_{i:\lambda}^{(g+1)} \quad (3-15)$$

其中:

- $\mu$  是选择的父代数量,
- $w_i$  是权重系数,
- $\mathbf{x}_{i:\lambda}^{(g+1)}$  是第  $g+1$  代中适应度排名前  $\mu$  的候选解。

3. 更新协方差矩阵:

$$\mathbf{C}^{(g+1)} = (1 - c_1 - c_\mu) \mathbf{C}^{(g)} + c_1 \mathbf{p}_c^{(g+1)} (\mathbf{p}_c^{(g+1)})^\top + c_\mu \sum_{i=1}^{\mu} w_i \mathbf{y}_{i:\lambda}^{(g+1)} (\mathbf{y}_{i:\lambda}^{(g+1)})^\top \quad (3-16)$$

其中:

- $c_1$  和  $c_\mu$  是学习率,
- $\mathbf{p}_c^{(g+1)}$  是进化路径,
- $\mathbf{y}_{i:\lambda}^{(g+1)} = (\mathbf{x}_{i:\lambda}^{(g+1)} - \mathbf{m}^{(g)}) / \sigma^{(g)}$ 。

4. 更新步长:

$$\sigma^{(g+1)} = \sigma^{(g)} \exp \left( \frac{c_\sigma}{d_\sigma} \left( \frac{\|\mathbf{p}_\sigma^{(g+1)}\|}{\mathbb{E}[\|\mathcal{N}(0, \mathbf{I})\|]} - 1 \right) \right) \quad (3-17)$$

其中:

- $c_\sigma$  是步长学习率,
- $d_\sigma$  是阻尼系数,
- $\mathbf{p}_\sigma^{(g+1)}$  是步长进化路径,
- $\mathbb{E}[\|\mathcal{N}(0, \mathbf{I})\|]$  是标准正态分布向量的期望范数。

我们将寻找最佳数据增强策略的问题形式化为一个搜索问题（见图 3-5）。我们的方法由两个组件组成：搜索算法和搜索空间。在高层次上，搜索算法（实现为 CMA-ES）采样一个数据增强策略  $S$ ，该策略包含有关使用哪种图像处理操作、每批次中使用该操作的概率以及操作幅度的信息。我们方法的关键在于，策略  $S$  将用于训练具有固定架构的神经网络，其验证准确率  $R$  将返回以评估适应度。搜索空间细节：在我们的搜索空间中，一个策略由 8 个信号数据增强子策略组成，如图 3-3 所示。此外，每个操作还关联两个超参数：

- 1) 应用操作的概率  $\in [0, 1]$ ,
- 2) 操作的幅度映射后  $\in [0, 10]$ 。

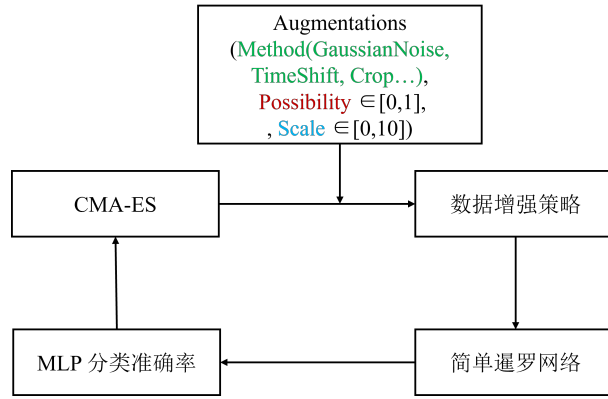


图 3-5 使用 CMA-ES 搜索最优的数据增强策略框架概览

映射为线性变换，如变量  $x \in [a, b]$  到  $y \in [0, 10]$  的变换为

$$y = \frac{10(x - a)}{b - a} \quad (3-18)$$

以下介绍各个数据增强子策略幅度的表示：

- **掩码 (Random Masked)**：随机选取 100 个不重叠的大小为  $s$  的子区间，将其置 0。 $s$  作为增强幅度由  $[1, 5]$  映射为  $[0, 10]$ 。
- **添加高斯噪声 (Adding Gaussian Noise)**：将高斯噪声的信噪比 SNR 作为增强幅度，由  $[2, 6]$  映射为  $[0, 10]$ 。
- **相位扰动 (Phase Perturbation)**：相位扰动服从均匀分布  $U(-perturb_{\max}, perturb_{\max})$ ， $perturb_{\max}$  作为增强幅度由  $[0.1, 0.5]$  映射为  $[0, 10]$ 。
- **块打乱 (RandomChunkShuffle)**：将时间序列数据分割成大小均匀的  $s$  个块，然后随机打乱这些块的顺序。 $s$  作为增强幅度由  $[10, 100]$  映射为  $[0, 10]$ 。
- **随机缩放 (Random Scaled)**：缩放的幅度服从均匀分布  $U(1.0 - scale, 1.0 + scale)$ 。 $scale$  作为增强幅度由  $[0.05, 0.6]$  映射为  $[0, 10]$ 。
- **随机绝对值 (Random Abs)**：无幅度值。
- **竖直翻转 (Random Vertical Flip)**：无幅度值。
- **水平翻转 (Random Horizontal Flip)**：无幅度值。

则 CMA-ES 优化器求解的问题可以描述为

$$\arg \max_{p, s} \text{Accuracy}(p, s, \text{SimSiam-Net}) \quad (3-19)$$

其中， $p$  和  $s$  为求解的参数，分别满足  $p_i \in [0, 1]$  和  $s_i \in [0, 10]$ ，Accuracy 为简单暹罗网络验证的准确率。

## 3.2 实验与分析

为了验证所提出方法的有效性，我们选择了三种不同的轴承故障数据进行实验。它们分别是 CWRU 数据集、SQ 数据集和 PB 数据集。此外，为了证明所提出方法的优越性，我们选择了几种流行的方法作为对比方法，这些方法可以分为两类。一类是传统的机器学习算法，另一类是当前基于深度学习的智能方法。对比方法如下所示。

1. **SVM [31]**: 多类 SVM 是一种强大且多功能的机器学习模型，能够执行线性或非线性分类任务。首先，我们将计算原始数据的 16 个时域指标（均值、平方根、偏度等），并将它们作为 SVM 分类器的输入。
2. **CNN [32]**: CNN 是一种监督学习方法。在这里，CNN 的网络架构与所提出方法的编码器  $f$  相同。
3. **LiftingNet [33]**: 一种监督学习方法，其中应用了分层特征学习，网络架构与原始文章保持一致。
4. **Multiscale [34]**: 一种监督学习方法，其中模型采用了多尺度特征提取单元。网络架构与原始文章保持一致。
5. **多感受野图卷积网络 (MRF-GCN) [35]**: 一种使用多感受野图卷积网络 (GCN) 的监督学习轴承故障诊断方法，采用两阶段训练策略。
6. **快速自注意力卷积生成对抗网络 (QSCGAN) [36]**: QSCGAN 是一种新颖的半监督生成对抗网络，用于故障诊断，采用两阶段训练策略。模型与原始文章保持一致。
7. **BYOL [27]**: BYOL 也是一种自监督表示学习方法，它依赖于两个相互交互和学习的神经网络块。编码器  $f$  和数据增强方法与所提出方法相同。

### 3.2.1 简单暹罗孪生网络实现细节

由于简单暹罗网络是一个框架，我们可以根据需要构建编码器模型和预测器，以下介绍本文构建的简单暹罗网络框架的细节。

- **Encoder**: 堆叠了十个 CNN 块作为编码器，其中卷积层的核大小设置为 3，填充设置为 1，步幅设置为 1，隐藏层采用 256 个核。
- **Predictor**: 我们使用两个线性全连接层构建预测器，分别具有  $256 \times 512$  和  $512 \times 256$  个神经元。我们在它们之间放置了一个批归一化层和一个修正线性单元 (ReLU) 层。
- **Classifier**: 至于分类器层，我们也部署了两个线性全连接层，分别具有  $256 \times 256$  和  $256 \times \text{num\_class}$  个神经元，其中  $\text{num\_class}$  表示故障类别的数

量。在两个线性层之间也放置了一个批归一化层和一个 ReLU 层。

- **Optimizer:** 优化器为 SGD 优化器，其学习率初始设置为  $0.05 \times \frac{\text{BatchSize}}{256}$ ，学习率采用余弦衰减调度。余弦衰减调度的公式为：

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min}) \left(1 + \cos\left(\frac{t\pi}{T}\right)\right)$$

其中：

- $\eta_t$  是第  $t$  步的学习率，
- $\eta_{\max}$  是初始学习率（最大学习率），
- $\eta_{\min}$  是最小学习率，
- $t$  是当前训练步数，
- $T$  是总训练步数（衰减周期），
- $\cos$  是余弦函数。

权重衰减为 0.0001，SGD 动量为 0.9。

- **损失函数:** 对比学习预训练阶段的损失函数为余弦相似度见3-1，微调阶段的损失函数为交叉熵损失函数（Cross-Entropy Loss Function）。对于多分类问题，交叉熵损失函数表示为

$$\mathcal{L}_{\text{CE}} = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (3-20)$$

其中：

- $C$  是类别数量，
- $y_i$  是真实标签的 one-hot 编码（第  $i$  类的真实概率），
- $\hat{y}_i$  是模型预测的第  $i$  类的概率。
- **长尾分布数据集模拟构造:** 模拟构建长尾分布数据集流程如图3-6所示。设定不平衡因子  $\beta = x_{\max}/x_{\min}$  为数据集中样本数量最多的类与样本数量最少的类的数量之比。帕累托分布的概率密度函数为  $p(x) = \frac{\alpha x_{\min}^{\alpha}}{x^{\alpha+1}}$ ，令  $x_{\min} = 1$ ，则  $p(x) = \frac{\alpha}{x^{\alpha+1}}$ 。以不平衡因子构建服从帕累托分布的长尾分布数据集，不同不平衡因子的帕累托分布概率密度函数如图3-7所示，可以看到当最大类与最小类样本数的比值  $\beta$  增大时，帕累托分布的形状参数  $\alpha$  也随之增大。这表明分布的头部类别占比提高，而尾部类别占比显著降低，长尾效应越发显著。下面将介绍形状参数  $\alpha$  的求解方法，以及各类别样本占比的计算步骤。已知帕累托分布的累积分布函数为：

$$F(x) = 1 - x^{-\alpha}, \quad x > 1 \quad (3-21)$$

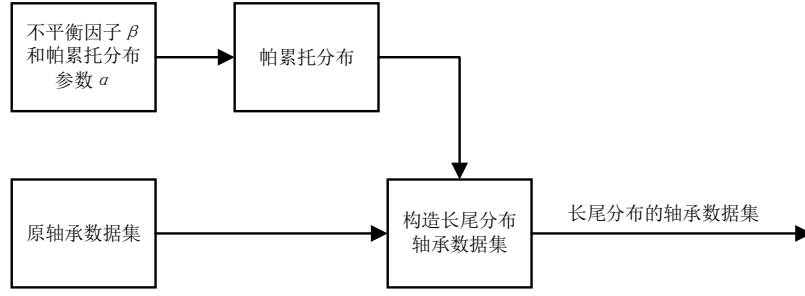


图 3-6 基于帕累托分布构造长尾分布的轴承数据集流程图

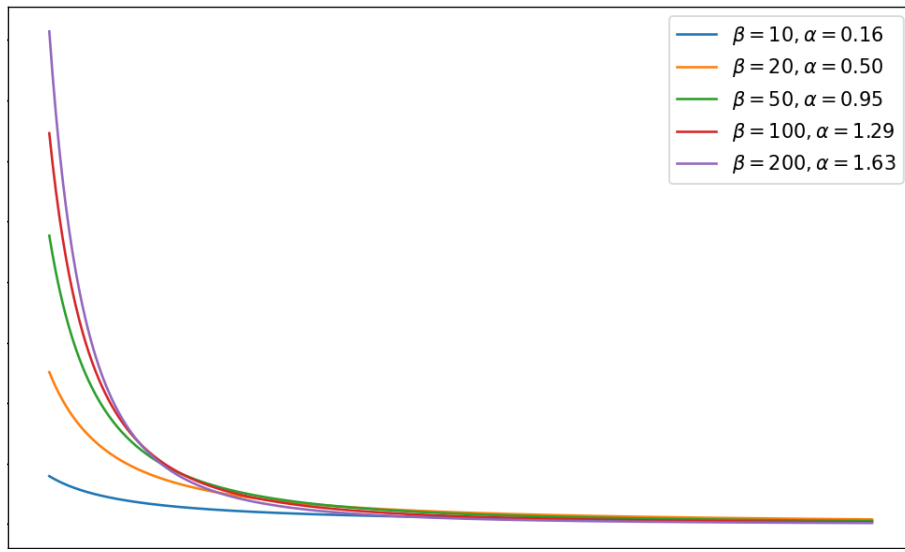


图 3-7 不同  $\beta$  的帕累托分布概率密度图

每类的概率定义为：

$$P(n \leq x < n+1) = F(n+1) - F(n) = n^{-\alpha} - (n+1)^{-\alpha} \quad (3-22)$$

设最大类与最小类的样本数比值为  $\beta$ ，则有以下关系式：

$$\beta = \frac{P(1 \leq x < 2)}{P(n \leq x < n+1)} = \frac{1 - 2^{-\alpha}}{n^{-\alpha} - (n+1)^{-\alpha}} \quad (3-23)$$

目标是：

1. 根据给定的  $\beta$  和类别总数  $n$ ，求解形状参数  $\alpha$ 。
2. 计算每类的样本占比  $P(n \leq x < n+1)$ 。
3. 将所有类别的概率归一化，使其和为 1。

具体步骤：

1. **求解  $\alpha$**

根据  $\beta$  的定义，解以下非线性方程以确定  $\alpha$ ：

$$\beta = \frac{1 - 2^{-\alpha}}{n^{-\alpha} - (n+1)^{-\alpha}} \quad (3-24)$$

该方程一般无解析解，可以通过数值方法（如 Newton-Raphson 或其他优化算法）求解。

## 2. 计算每类的概率

每类的概率可以通过以下公式计算：

$$P(n \leq x < n+1) = n^{-\alpha} - (n+1)^{-\alpha}, \quad n = 1, 2, \dots, N \quad (3-25)$$

## 3. 概率归一化

将所有类别的概率归一化，计算归一化后的概率：

$$P_{\text{norm}}(n \leq x < n+1) = \frac{P(n \leq x < n+1)}{\sum_{k=1}^N P(k \leq x < k+1)} \quad (3-26)$$

其中  $N$  为类别总数，归一化后各类别的样本占比之和为 1：

$$\sum_{n=1}^N P_{\text{norm}}(n \leq x < n+1) = 1 \quad (3-27)$$

- **实现细节：**每个样本包含 **1024** 个数据点。训练过程分为两个阶段：
  - 在**第一阶段**，我们随机选择每个故障类别的 **100** 个无标签样本进行模型训练，训练总计 **500** 个周期。
  - 在**第二阶段**，我们根据式（3-26）构造服从 **帕累托分布** 的长尾分布有标签样本，以此微调模型，微调总计 **50** 个周期。

在整个训练过程中，每个输入模型的 mini-batch 大小设定为 **64**。

- **性能评估：**测试集由每个故障类别均匀地随机选取 **100** 个样本组成。需要注意的是，微调过程中使用的子集与测试集完全独立。性能评估由以下几个指标组成：
  - **验证集准确率：**在验证集上评估模型的分类准确率，以衡量模型的泛化能力。该指标表示了模型对未知数据的分类能力。计算公式如式（3-12）。
  - **t-SNE 特征提取可视化：**使用 t-SNE（t-分布随机邻居嵌入，t-Distributed Stochastic Neighbor Embedding）方法将高维特征降维至二维空间进行可视化分析，直观展示不同类别之间的聚类情况。该方法有助于分析模型对不同故障类别的区分能力。t-SNE 通过以下三个步骤来进行降维：
    1. **计算高维相似度：**t-SNE 首先计算每对数据点之间的相似度。在高



维空间中，对于任意两点  $x_i$  和  $x_j$ ，其相似度通过条件概率  $p_{ij}$  表示，计算公式为：

$$p_{ij} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)}{\sum_{j \neq i} \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)} \quad (3-28)$$

其中， $\|x_i - x_j\|$  表示高维空间中两点之间的欧氏距离， $\sigma_i$  是与点  $x_i$  相关的局部尺度参数，反映了该点的邻域大小。通过这种方式，t-SNE 保证了高维空间中相邻的数据点具有较高的相似度。

2. **映射到低维空间：**接着，t-SNE 将数据映射到低维空间中，通常是二维或三维。在低维空间中，t-SNE 使用 t-分布（Student's t-distribution）来计算数据点之间的相似度，公式为：

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{j \neq i} \left(1 + \|y_i - y_j\|^2\right)^{-1}} \quad (3-29)$$

其中， $y_i$  和  $y_j$  是低维空间中的数据点， $\|y_i - y_j\|$  是它们之间的欧氏距离。t-分布的重尾特性使得它能够更加有效地处理聚类之间的距离，同时避免了在低维空间中过度拥挤的现象。

3. **最小化 KL 散度：**t-SNE 的目标是最小化高维空间和低维空间中数据点相似度分布的差异。为此，t-SNE 通过最小化 Kullback-Leibler (KL) 散度来优化模型的映射过程，KL 散度的计算公式为：

$$\text{KL}(P||Q) = \sum_{i,j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (3-30)$$

通过迭代优化，t-SNE 会调整低维空间中的数据点位置，直到高维和低维空间的相似度分布尽可能接近。

t-SNE 降维伪代码见表3-2。

- **验证集 KMeans 分类准确率：**在验证集上使用 KMeans 算法对 Encoder 提取的特征进行无监督分类，并计算分类准确率。通过与真实标签对比，评估模型在无监督情境下的表现，进一步验证其分类能力。KMeans 算法的基本步骤如下：

1. **初始化：**随机选择 K 个数据点作为初始簇的质心（centroids）。

2. **分配簇**: 将每个数据点分配给离它最近的簇的质心。通常使用欧氏距离度量:

$$\text{Distance}(x_i, c_j) = \|x_i - c_j\|$$

其中,  $x_i$  是数据点,  $c_j$  是簇的质心。

3. **更新质心**: 对于每个簇, 重新计算其质心, 即簇中所有数据点的均值:

$$c_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i$$

其中,  $S_j$  表示第  $j$  个簇中的数据点,  $|S_j|$  是该簇中数据点的数量。

4. **重复迭代**: 重复步骤 2 到 4, 直到质心不再发生变化或达到最大迭代次数。
5. **使用匈牙利算法找到最佳匹配**: 在完成迭代后, 使用匈牙利算法来优化簇标签的匹配, 以确保最合适的标签分配。匈牙利算法可以通过最小化簇间的匹配成本来实现最佳匹配。

KMeans 算法的伪代码见表3-3。

### 3.2.2 基于 CMA-ES 优化器的数据增强策略搜索

本研究采用 **CWRU 数据集** 作为对比学习预训练的无标签数据, 样本总数为 **500** 个。用于微调的有标签数据共 **100** 个, 涵盖 **10** 个类别, 每类包含 **10** 个样本。各类别信号的时域波形如图 2-5 所示。

在优化过程中, 我们基于 **协方差矩阵适应进化策略 (CMA-ES)** 进行最优数据增强策略的搜索。CMA-ES 的超参数设定如下: 变量维度为 **14** (即  $\text{len}(\mathbf{p}) + \text{len}(\mathbf{s})$ ), 种群规模设定为 **11**, 最大迭代次数设定为 **10** 轮。经过优化后, 所获得的最优数据增强策略如表 3-4 所示。

### 3.2.3 在 CWRU 上的实验结果

在本节中, 我们采用表3-4所示的数据增强参数, 并使用 **CWRU 数据集** 作为对比学习预训练的无标签数据集, 其中样本总数为 **500** 个。用于微调的有标签数据共计 **100** 个, 并通过 **帕累托分布** 生成不同不平衡因子的长尾分布标记样本。最终, 我们在样本分布均匀的测试集上计算分类准确率, 并进行 **5 次重复试验** 取平均值。与其他方法的准确率对比结果见表3-5, 图3-8 展示了 SimSiam、CNN 和 SVM 三种方法在均匀/长尾分布数据集上的 t-SNE 特征投影。从图中可以直观地看到不

表 3-2 t-SNE 算法伪代码

---

**t-SNE 算法伪代码**

---

```

# 输入：高维数据  $X = \{x_1, x_2, \dots, x_n\}$ 
# 输出：低维数据  $Y = \{y_1, y_2, \dots, y_n\}$ 

# Step 1: 计算高维空间中数据点之间的相似度  $p_{ij}$ 
for i = 1 to n: # 对每个数据点  $x_i$ 
    for j = 1 to n:
        # 计算  $x_i$  和  $x_j$  之间的欧氏距离，并转化为条件概率  $p_{ij}$ 
         $dist_{ij} = \text{norm}(x_i - x_j)$ 
         $p_{ij} = \frac{\exp(-dist_{ij}^2 / (2 * \sigma_i^2))}{\sum(\exp(-dist_{ij}^2 / (2 * \sigma_i^2)) \text{ for all } j)}$ 
         $P[i][j] = p_{ij}$  # 存储  $p_{ij}$ 

# Step 2: 初始化低维空间中的数据点  $Y$ 
 $Y = \text{random\_initialization}(n, d)$  # 随机初始化低维空间中的数据点

# Step 3: 迭代优化：最小化KL散度
for t = 1 to T: # 最大迭代次数 T
    # 计算低维空间中的相似度  $q_{ij}$ 
    for i = 1 to n:
        for j = 1 to n:
             $dist_{ij} = \text{norm}(y_i - y_j)$ 
             $q_{ij} = \frac{(1 + dist_{ij}^2)^{-1}}{\sum((1 + dist_{ij}^2)^{-1} \text{ for all } j)}$ 
             $Q[i][j] = q_{ij}$  # 存储  $q_{ij}$ 

    # 计算KL散度并计算梯度
     $KL = \sum(P[i][j] * \log(P[i][j] / Q[i][j]) \text{ for all } i \text{ and } j)$ 
     $gradients = \text{compute\_gradients}(P, Q, Y)$  # 计算梯度

# Step 4: 更新低维空间中的数据点  $Y$ 
 $Y = Y - \text{learning\_rate} * gradients$  # 梯度下降更新低维数据点

# Step 5: 终止条件判断
if converged(KL): # 如果KL散度收敛
    break

return Y # 返回降维后的低维数据  $Y$ 

```

---

表 3-3 KMeans 算法伪代码

**KMeans 算法伪代码**


---

```

# 输入：数据集  $X = \{x_1, x_2, \dots, x_n\}$ ，簇数  $K$ ，最大迭代次数  $T$ 
# 输出：簇标签  $\{y_1, y_2, \dots, y_n\}$ 

# 初始化：随机选择  $K$  个数据点作为初始质心  $C = \{c_1, c_2, \dots, c_K\}$ 
for t = 1 to T: # 最大迭代次数 T
    for i = 1 to n: # 对每个数据点  $x_i$ 
        # 计算每个质心的距离，分配数据点给最近的质心
        distances = [] # 初始化一个空列表
        for cj in C: # 对每个质心  $c_j$ 
            dist = norm( $x_i - c_j$ ) # 计算  $x_i$  到  $c_j$  的距离
            distances.append(dist) # 添加到距离列表
        yi = argmin(distances) # 分配数据点  $x_i$  到最近的质心
        # 更新数据点的簇标签
        labels[i] = yi

    for j = 1 to K: # 对每个簇
        # 更新质心为簇中所有数据点的均值
        cluster_points =
            [ $x_i$  for  $x_i, label$  in zip( $X, labels$ ) if  $label == j$ ]
        # 计算簇内数据点的均值作为新质心
        C[j] = mean(cluster_points)

# 使用匈牙利算法找到最佳匹配
optimal_labels = hungarian_algorithm(labels, C)

# 如果质心不再变化
if no_change_in_centroids(C): # 如果质心不再变化
    break # 退出

return optimal_labels # 返回优化后的簇标签  $\{y_1, y_2, \dots, y_n\}$ 

```

---

表 3-4 CMA-ES 搜索最优的数据增强策略参数

增强编号	增强名称	概率 $p$	幅度 $s$
DA0	高斯噪声	0.9761	9.6284
DA1	相位扰动	0.9899	3.4435
DA2	块打乱	0.9687	1.3436
DA3	掩码	0.7873	8.1961
DA4	缩放	0.5776	7.8206
DA5	绝对值	0.4020	—
DA6	竖直翻转	0.4392	—
DA7	水平翻转	0.4420	—

同类别的样本如何在低维空间中分布，从而评估各方法的特征提取效果。

表 3-5 SimSiam、CNN 与 SVM 在不同  $\beta$  值下的准确率

不平衡因子 $\beta$	SimSiam	CNN	SVM
1	0.9515	0.8979	0.7198
2	0.9517	0.9113	0.3490
5	0.9315	0.7373	0.3510
10	0.9402	0.8538	0.2990
50	0.9131	0.6804	0.3917
100	0.8450	0.5358	0.3115

### 3.2.4 对 SimSiam 网络模块的探讨

图3-9展示了“使用停止梯度（Stop-Grad）”与“不使用停止梯度”的比较。架构和所有超参数保持不变，唯一的区别是是否使用停止梯度或预测器。图3-9(a)显示了训练损失。不使用停止梯度时，优化器迅速找到一个退化解，并达到最小可能损失  $-1$ 。为了证明这种退化是由坍缩引起的，我们研究了  $\ell_2$  归一化输出  $z/\|z\|_2$  的标准差（std）。如果输出坍缩到一个常数向量，则所有样本的 std 在每个通道上应该为零。这可以从图 2（中）的红色曲线中观察到。

作为对比，如果输出  $z$  具有零均值各向同性高斯分布，我们可以证明  $\frac{z}{\|z\|_2}$  的标准差为  $\frac{1}{\sqrt{d}}$ 。下证：假设  $z \in \mathbb{R}^d$  是一个零均值各向同性高斯分布，即  $z \sim \mathcal{N}(0, I_d)$ ，其中  $I_d$  是  $d$  维单位矩阵，表示协方差矩阵为单位矩阵，意味着每个分量  $z_i$  独立且服从标准正态分布。我们需要证明的是， $\frac{z}{\|z\|_2}$  的标准差（即每个分量的标准差）为  $\frac{1}{\sqrt{d}}$ ，其中  $\|z\|_2 = \sqrt{\sum_{i=1}^d z_i^2}$  是  $z$  的  $L_2$  范数。

**证明步骤：**

1. **标准化操作：** 设  $z = (z_1, z_2, \dots, z_d)$ ，则  $z$  的每个分量  $z_i$  都是独立的标准正态

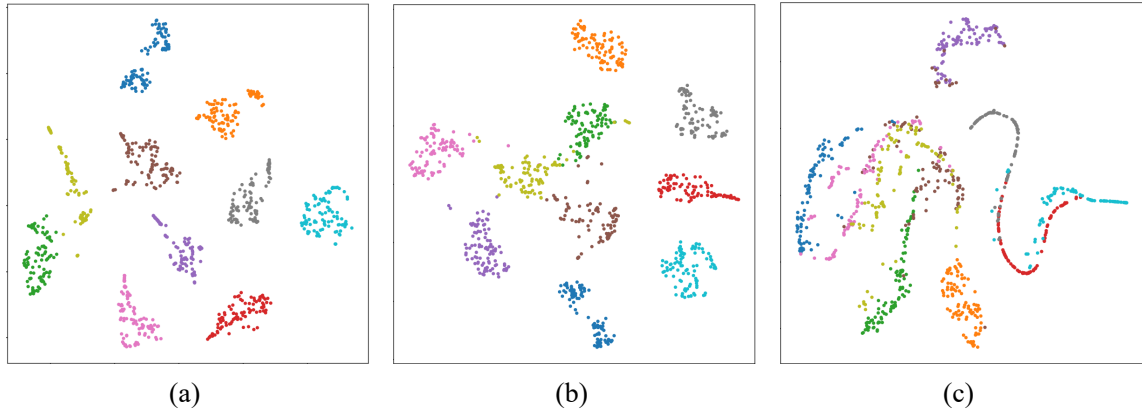


图 3-8 验证集数据经特征提取的 T-SNE 可视化: (a) SimSiam 的 T-SNE 分布图; (b) CNN 的 T-SNE 分布图; (c) SVM 的 T-SNE 分布图。

随机变量  $z_i \sim \mathcal{N}(0, 1)$ , 并且  $\|z\|_2 = \sqrt{\sum_{i=1}^d z_i^2}$  是  $z$  的欧几里得范数。我们关注的是向量  $\frac{z}{\|z\|_2}$  的每个分量。

2. **定义随机变量:** 令  $\hat{z} = \frac{z}{\|z\|_2}$ , 即  $\hat{z}$  是单位向量, 表示将  $z$  投影到单位球面上的结果。
3. **每个分量的分布:** 因为  $z$  是各向同性高斯分布, 其方向是均匀分布在单位球面上的。对于标准化后的向量  $\hat{z}$ , 每个分量  $\hat{z}_i = \frac{z_i}{\|z\|_2}$  的分布不再是标准正态分布, 而是受到范数约束的结果。
4. **期望值和方差:** 由于  $z$  是各向同性的,  $\hat{z}_i$  的分布是对称的, 所有分量  $\hat{z}_i$  的期望值为 0。因此, 接下来我们计算  $\hat{z}_i$  的方差。

对于  $\hat{z}_i$  的方差, 我们有:

$$\text{Var}(\hat{z}_i) = \mathbb{E}[\hat{z}_i^2] - (\mathbb{E}[\hat{z}_i])^2 \quad (3-31)$$

由于  $\mathbb{E}[\hat{z}_i] = 0$ , 我们只需要计算  $\mathbb{E}[\hat{z}_i^2]$ 。

$$\mathbb{E}[\hat{z}_i^2] = \mathbb{E}\left[\frac{z_i^2}{\|z\|_2^2}\right] = \frac{\mathbb{E}[z_i^2]}{\mathbb{E}[\|z\|_2^2]} = \frac{1}{d} \quad (3-32)$$

其中  $\mathbb{E}[z_i^2] = 1$  (因为  $z_i$  是标准正态分布), 而  $\mathbb{E}[\|z\|_2^2] = \mathbb{E}[\sum_{i=1}^d z_i^2] = d$ , 因为每个  $z_i^2$  的期望为 1。

5. **标准差:** 由于方差  $\text{Var}(\hat{z}_i) = \frac{1}{d}$ , 因此标准差是:

$$\text{std}(\hat{z}_i) = \frac{1}{\sqrt{d}} \quad (3-33)$$

图3-9(b)的红色曲线显示，使用停止梯度时，std 值接近  $\frac{1}{\sqrt{d}}$ 。这表明输出没有坍塌，而是分散在单位超球面上。图3-9(c)绘制了 KMeans 聚类的验证准确率。KMeans 聚类可以作为进展的监控工具。使用停止梯度时，监控显示出准确率稳步提高。

图3-9的绿色曲线显示模型在移除预测器（Predictor）后的表现，可以看出移除预测器后发生了与移除 Stop-Grad 类似的坍塌。事实上，如果使用对称损失（式(3-2)），这一现象是可以预期的。现在的损失函数是

$$\frac{1}{2}D(z_1, \text{stopgrad}(z_2)) + \frac{1}{2}D(z_2, \text{stopgrad}(z_1)). \quad (3-34)$$

其梯度的方向与  $D(z_1, z_2)$  的梯度相同，只不过幅度被缩放了  $\frac{1}{2}$ 。在这种情况下，使用 Stop-Grad 相当于移除 Stop-Grad 并将损失函数缩放  $\frac{1}{2}$ 。因此，观察到模型坍塌（图3-9）。图3-10可视化了这种坍塌带来的对特征提取的影响。表3-6也显示了移除 Stop-Grad 或预测器后的巨大性能损失。

表 3-6 SimSiam、SimSiamw/o SG 与 SimSiamw/o Pred 在不同  $\beta$  值下的准确率

不平衡因子 $\beta$	SimSiam	SimSiamw/o SG	SimSiamw/o Pred
1	0.9515	0.5338	0.5217
2	0.9517	0.5025	0.5575
5	0.9315	0.4496	0.4952
10	0.9402	0.4919	0.4604
50	0.9131	0.3579	0.3892
100	0.8450	0.3412	0.3704
平均值	0.9222	0.4462	0.4657

### 3.2.5 对数据增强模块的探讨

在3.2.2小节，我们在限定范围内求得了各个数据增强子策略的最优参数解，并在之后的实验中取得了良好的结果。本节讨论表3-4中的 8 个数据增强子策略对模型性能的影响。图3-11显示了移除单个子策略后模型训练过程中的损失函数及特征标准差。

## 3.3 本章小结

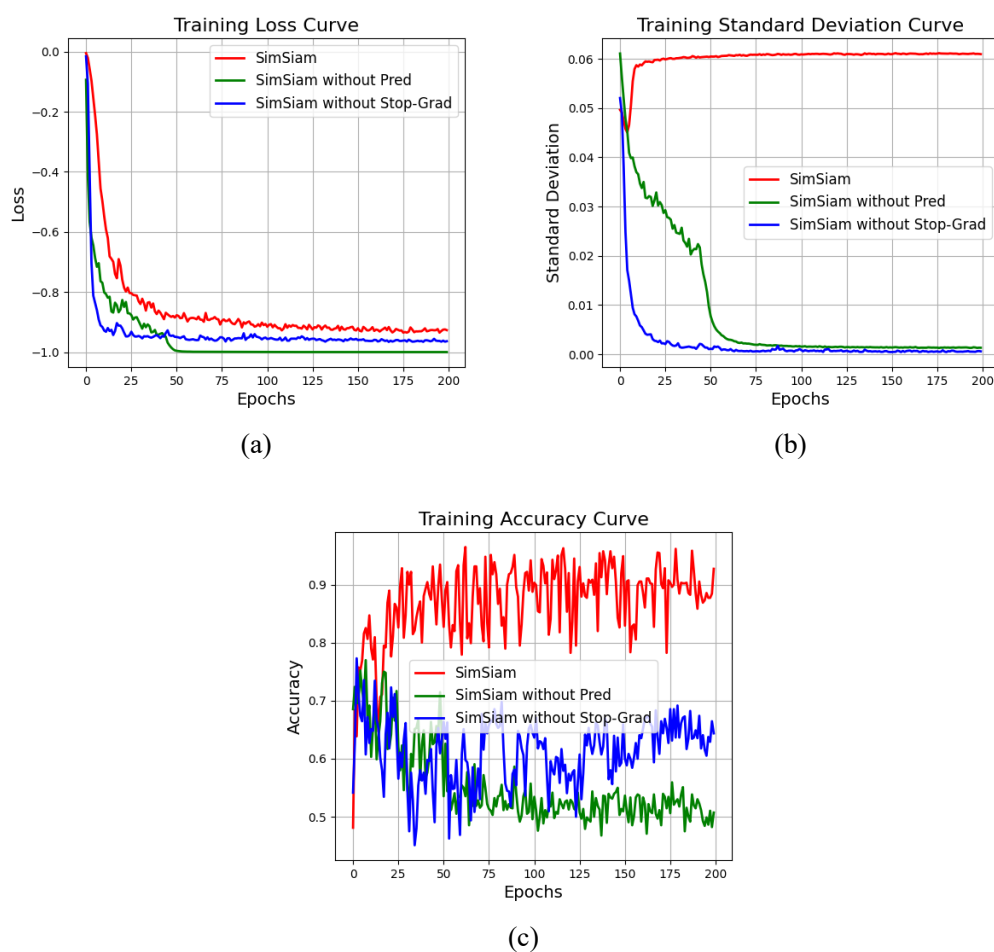


图 3-9 SimSiam, SimSiam(无 Stop-Grad 模块) 与 SimSiam(无预测器模块) 的训练过程对比 (a) 损失函数; (b) 提取特征的标准差; (c) K-Means 聚类准确率

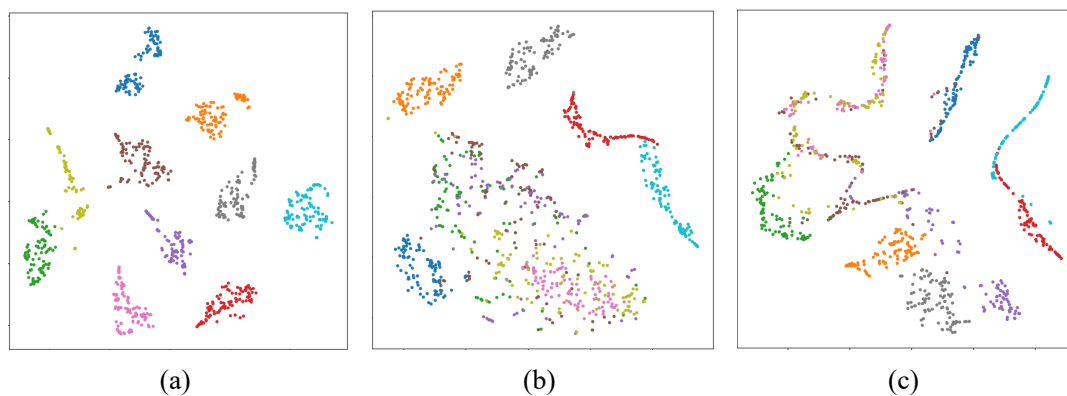


图 3-10 SimSiam, SimSiam(无 Stop-Grad 模块) 与 SimSiam(无预测器模块) 的 T-SNE 图 (a) SimSiam; (b) SimSiam without Stop-Grad; (c) SimSiam without Predictor



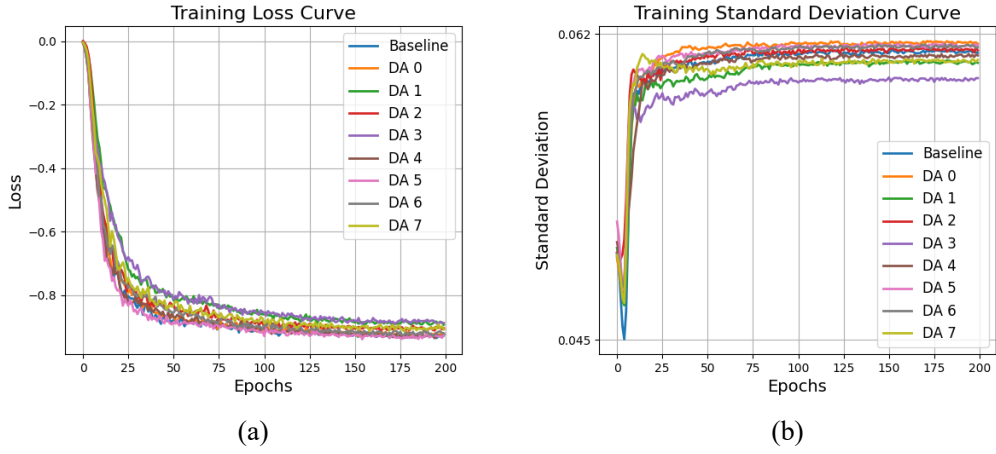


图 3-11 移除不同数据增强子策略的训练过程对比 (a) 损失函数；(b) 提取特征的标准差；

表 3-7 移除不同数据增强子策略在不同  $\beta$  值下的准确率

不平衡因子 $\beta$	DA0	DA1	DA2	DA3	DA4	DA5	DA6	DA7
1	0.9279	0.9127	0.9254	0.9204	0.9106	0.9298	0.9394	0.9338
2	0.9219	0.9121	0.9223	0.9275	0.9183	0.9333	0.9427	0.9456
5	0.9246	0.8969	0.8940	0.9217	0.8810	0.9088	0.9117	0.9310
10	0.9258	0.9348	0.8944	0.9298	0.8792	0.9358	0.9333	0.9442
50	0.8563	0.8379	0.8221	0.8467	0.7490	0.8165	0.8573	0.8387
100	0.7935	0.8169	0.7356	0.7194	0.7119	0.7912	0.8333	0.7869
平均值	0.8836	0.8852	0.8656	0.8776	0.8417	0.8859	0.9029	0.8967
相比 Baseline	-0.0386	-0.0370	-0.0566	-0.0446	-0.0805	-0.0363	-0.0193	-0.0255

## 第四章 基于半监督学习和分类面调整算法的微调方法研究

在上一章中，我们展示了自监督学习范式中的二阶段式训练（如图2-3所示）在故障诊断任务中的长尾学习场景下，显著提升了模型性能。然而，在微调阶段，样本分布的不平衡仍会影响模型，导致分类器的决策面发生偏移。如图4-1所示，决策面会朝尾部类的空间偏移，导致头部类在特征空间中占据更大的体积，从而降低了模型对尾部类的分类准确率。在本章节中，我们将深入探讨半监督学习与决策面调整算法，旨在减轻长尾分布对微调阶段决策面偏移的影响，从而提高尾部类的识别能力。

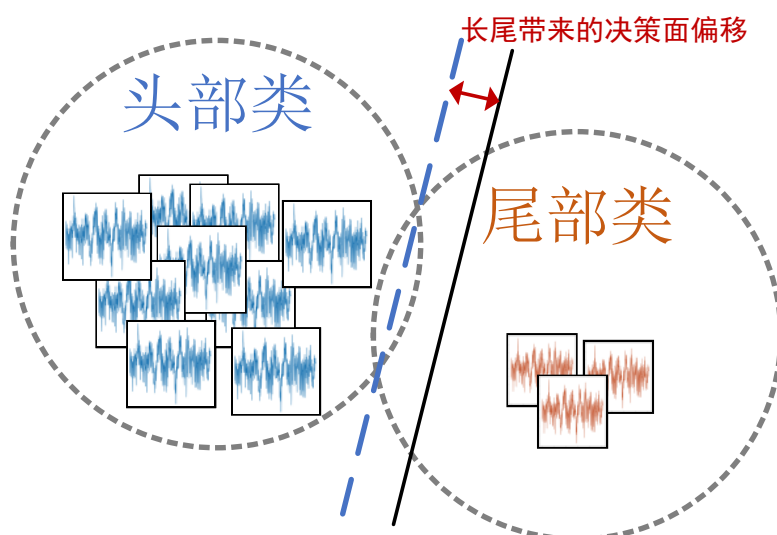


图 4-1 由长尾分布带来的决策面偏移示意图

### 4.1 模型架构

#### 4.1.1 半监督学习模型架构

本研究提出了一种结合半监督学习与决策边界调整的框架，以有效解决长尾分布和样本不平衡问题。受到文献 [6, 7] 的启发，研究借鉴了半监督学习的经典流程（如图4-2所示）。该流程分为两个阶段：

在第一阶段，使用原始的不平衡标记数据集  $\mathcal{D}_L$  训练一个初步分类器  $f_{\theta}$ 。随后，在第二阶段，利用该分类器对未标记数据集  $\mathcal{D}_U$  进行推断并生成伪标签  $\hat{y}$ 。伪标签

数据与标记数据结合，通过最小化以下损失函数优化模型：

$$\mathcal{L}(\mathcal{D}_L, \theta) + \omega \mathcal{L}(\mathcal{D}_U, \theta), \quad (4-1)$$

其中， $\omega$  为未标记数据的损失权重， $\mathcal{L}(\mathcal{D}_L, \theta)$  和  $\mathcal{L}(\mathcal{D}_U, \theta)$  分别代表基于标记数据和伪标签数据的损失项。通过整合这两部分损失，最终得到优化后的分类器  $f_{\hat{\theta}_f}$ ，以更好地建模数据分布，尤其是尾部类别（少数类别）。该过程通过充分利用未标记数据  $\mathcal{D}_U$ ，提升模型对少数类别的分类能力，从而显著优化其决策边界。

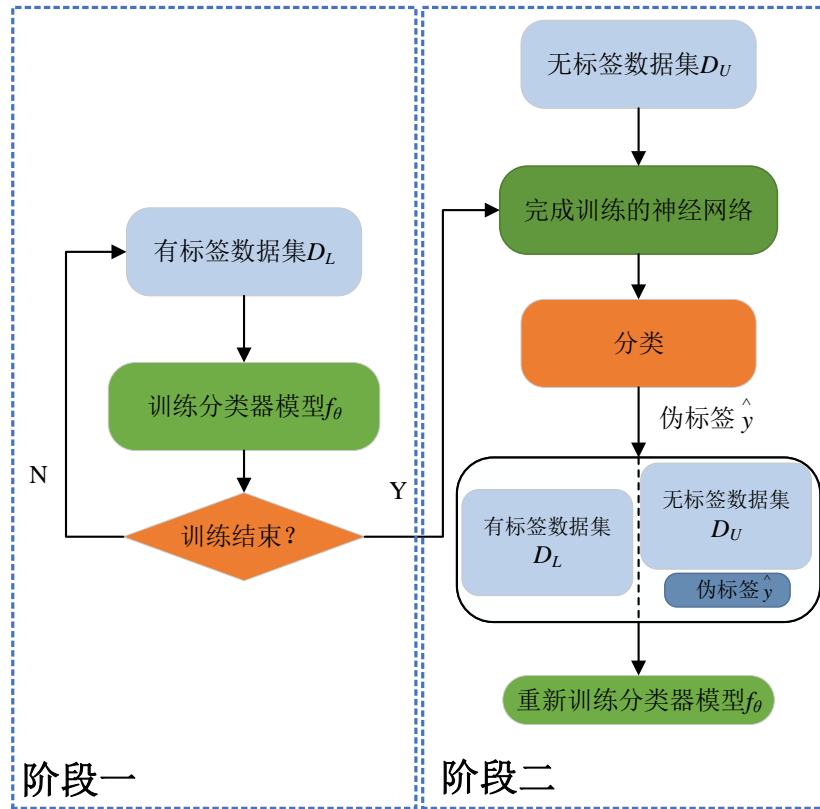


图 4-2 半监督学习流程图

与此同时，在处理长尾分布和样本不平衡问题时，决策面调整方法（MARC）提供了一种有效的方式来改善分类器的性能，特别是在少数类别（尾部类别）上。文献 [7] 介绍的决策面调整算法通过引入边界校准方法，调整分类器的决策面，使其更加均衡地处理各个类别的数据。

MARC 算法包括两个关键参数： $\omega$  和  $\beta$ ，其中  $\omega$  是用于调整输出的尺度因子，而  $\beta$  则是用于调整输出的偏置项。这两个参数通过最小化损失函数来进行优化，从而获得适合当前任务的最优决策面。算法的具体实现步骤参考表 4-1。在 MARC 伪代码中， $\omega$  和  $\beta$  是通过 `torch.nn.Parameter` 初始化为可训练的参数，它们分

别表示边界的缩放因子和偏置项。使用 `torch.no_grad()` 关键字来避免计算梯度，确保在前向传播时不会影响到梯度计算。接下来，代码通过计算网络的权重范数以及模型在给定输入  $x$  上的输出结果，来调整输出的决策面，生成校准后的输出 `logit_after`。在实际应用中，MARC 算法有助于减小尾部类别的分类误差，尤其是在样本不平衡的情况下，通过有效地调整决策边界来改善模型的泛化能力和分类精度。

---

#### MARC 伪代码（用 Pytorch 描述）

---

```

1: 初始化边界校准方法：
    omega = torch.nn.Parameter(torch.ones(1, num_classes))
    beta = torch.nn.Parameter(torch.zeros(1, num_classes))

2: 输入：训练数据  $x$ ，标准的预训练神经网络模型。

3: with torch.no_grad():
4:     w_norm = torch.norm(model.fc.weight, dim=1)
5:     logit_before = model(x)
6:     logit_after = omega * logit_before + beta * w_norm
7: 计算损失并更新参数 omega 和 beta。

```

---

表 4-1 MARC 算法的实现步骤

基于上述半监督学习和 MARC 决策边界调整算法，本研究改进了上一章的 SimSiam 架构，提出了一种结合两者的 SimSiam 模型架构，称为 **基于半监督学习和 MARC 微调的 SimSiam 模型**。该模型通过三阶段的逐步优化，旨在解决长尾分布和样本不平衡问题，并显著提升分类器在尾部类别上的性能。其具体架构流程如图 4-3 所示。该模型架构包括以下三个阶段：

1. **阶段一（自监督学习）**：在第一阶段，基于 SimSiam 架构对模型进行预训练。具体而言，利用大量无标签样本和少量有标签的长尾样本，训练一个性能较优的特征提取网络和初步分类器。SimSiam 的自监督特性能够在数据量受限的情况下充分挖掘数据的潜在分布信息，为后续在半监督学习提供高质量的特征表示。
2. **阶段二（半监督学习）**：在第二阶段，基于阶段一训练得到的 SimSiam 模型，使用其生成无标签样本的伪标签（Pseudo Labels），将伪标签视为无标签数据的“真实标签”。随后，将补充有伪标签的无标签数据与原始有标签数据集相结合，通过最小化标记数据和伪标签数据的损失函数，重新优化分类器的全连接层参数。这一阶段的目标是进一步提高模型的泛化能力，尤其是在尾部类别上的表现。
3. **阶段三（MARC 决策边界调整）**：在第三阶段，利用经过半监督学习优化后

的分类器及其输出的 logits，对决策边界进行微调。具体而言，通过训练两个向量参数  $\omega$  和  $\beta$ ，分别控制分类器输出的尺度因子和偏置项，调整分类器的决策面，以提升模型在尾部类别上的分类准确率。MARC 的引入有效平衡了模型在头部类别和尾部类别上的性能差异。

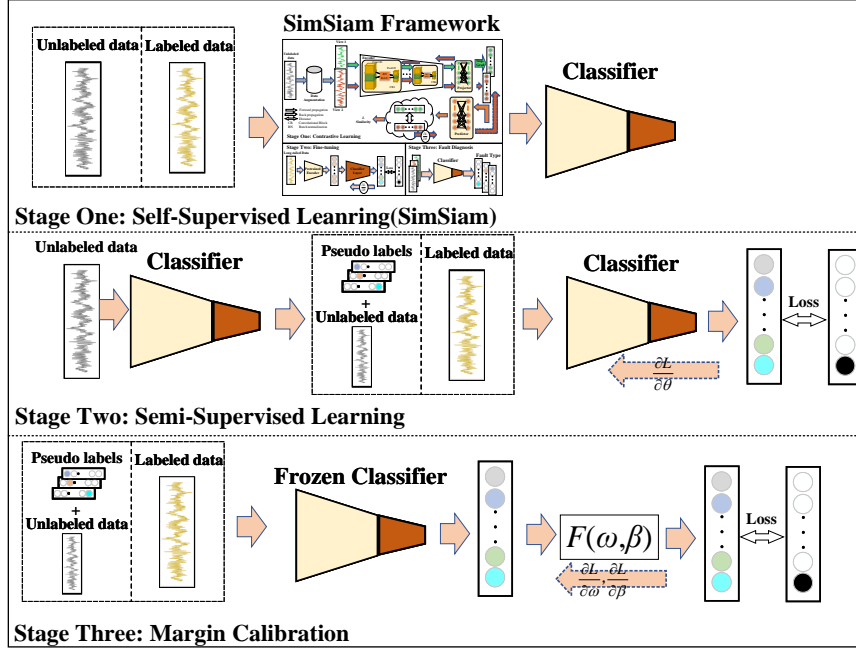


图 4-3 基于半监督学习和 MARC 微调的 SimSiam 模型流程框图

整体流程通过结合 SimSiam 的特征提取能力、半监督学习对伪标签的利用，以及 MARC 算法对决策边界的微调，形成了一个多阶段优化框架。该方法在处理长尾分布和样本不平衡问题时，展现出显著的优势，尤其是在提高尾部类别分类准确率方面效果显著。

## 4.2 实验与分析

采用 CWRU 数据集，数据集构造方法与 3.2.3 节相同。一阶段训练 150 个周期，二阶段重新训练 150 个周期。在不同不平衡因子下的准确率结果如表 4-2 所示。

表 4-2 SimSiam 与 Semi 在不同  $\beta$  值下的准确率

不平衡因子 $\beta$	SimSiam	Semi	Marc	Semi_Marc
1	0.9515	0.9813	0.94896	0.9781
2	0.9517	0.9656	0.95917	0.9698
5	0.9315	0.9406	0.94021	0.9615
10	0.9402	0.9573	0.9575	0.9615
50	0.9131	0.9344	0.91687	0.9271
100	0.8450	0.9031	0.87896	0.9406

## 第五章 其他长尾学习方法分析 (CB LOSS 等)

### 5.1

### 5.2 本章小结

## 第六章 全文总结与展望

### 6.1 全文总结

### 6.2 后续工作展望



## 致 谢

在攻读博士学位期间，首先衷心感谢我的导师 XXX 教授

## 附录 A 中心极限定理的证明

### A.1 高斯分布和伯努利实验

## 参考文献

- [1] 南京大学, 南京智谷人工智能研究院有限公司. 一种适用于长尾带噪分布数据的故障检测方法 [P]. 中国, 发明专利, CN202211435489.6, 2023 年 2 月 3 日.
- [2] 中国矿业大学. 一种面向长尾分布数据的工业过程故障诊断方法 [P]. 中国, 发明专利, CN202310182227.1, 2023 年 6 月 2 日.
- [3] Kingma D P. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013.
- [4] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [5] Cui Y, Jia M, Lin T-Y, et al. Class-balanced loss based on effective number of samples[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019: 9268-9277.
- [6] Yang Y, Xu Z. Rethinking the value of labels for improving class-imbalanced learning[J]. Advances in neural information processing systems, 2020, 33: 19290-19301.
- [7] Wang Y, Zhang B, Hou W, et al. Margin calibration for long-tailed visual recognition[C]. Asian Conference on Machine Learning, 2023: 1101-1116.
- [8] 吴亮, 梁振, 张焱鑫, et al. 基于多级学习的长尾分布下交通多目标检测 [J]. 计算机与数字工程, 2023, 51(3): 599-605.
- [9] Zhang W, Li X, Ma H, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision[J]. Knowledge-Based Systems, 2021, 213: 106679.
- [10] Shen X, Marks E. Near-miss information visualization tool in bim for construction safety[J]. Journal of construction engineering and management, 2016, 142(4): 04015100.
- [11] Wilson D L. Asymptotic properties of nearest neighbor rules using edited data[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1972, 408-421.
- [12] Chawla N V, Bowyer K W, Hall L O, et al. Smote: synthetic minority over-sampling technique[J]. Journal of artificial intelligence research, 2002, 16: 321-357.
- [13] Huang C, Li Y, Loy C C, et al. Learning deep representation for imbalanced classification[C]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 5375-5384.
- [14] Lin T. Focal loss for dense object detection[J]. arXiv preprint arXiv:1708.02002, 2017.

- [15] Liu J, Sun Y, Han C, et al. Deep representation learning on long-tailed data: A learnable embedding augmentation perspective[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020: 2970-2979.
- [16] 吴磊, 韩华, 黄丽, et al. 面向长尾图像的个性化专家识别算法 [J]. 电光与控制, 2023, 30(11): 62-66.
- [17] He K, Fan H, Wu Y, et al. Momentum contrast for unsupervised visual representation learning[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: .
- [18] Chen T, Kornblith S, Norouzi M, et al. A simple framework for contrastive learning of visual representations[C]. International conference on machine learning, 2020: 1597-1607.
- [19] Grill J-B, Strub F, Althé F, et al. Bootstrap your own latent-a new approach to self-supervised learning[J]. Advances in neural information processing systems, 2020, 33: 21271-21284.
- [20] Caron M, Misra I, Mairal J, et al. Unsupervised learning of visual features by contrasting cluster assignments[J]. Advances in neural information processing systems, 2020, 33: 9912-9924.
- [21] He K, Chen X, Xie S, et al. Masked autoencoders are scalable vision learners[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022: 16000-16009.
- [22] Doersch C, Gupta A, Efros A A. Unsupervised visual representation learning by context prediction[C]. Proceedings of the IEEE international conference on computer vision, 2015: 1422-1430.
- [23] Zhang R, Isola P, Efros A A. Colorful image colorization[C]. Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14, 2016: 649-666.
- [24] Gidaris S, Singh P, Komodakis N. Unsupervised representation learning by predicting image rotations[J]. arXiv preprint arXiv:1803.07728, 2018.
- [25] Zhang T, Chen J, He S, et al. Prior knowledge-augmented self-supervised feature learning for few-shot intelligent fault diagnosis of machines[J]. IEEE Transactions on Industrial Electronics, 2022, 69(10): 10573-10584.
- [26] Senanayaka J S L, Khang H V, Robbersmyr K G. Toward self-supervised feature learning for online diagnosis of multiple faults in electric powertrains[J]. IEEE Transactions on Industrial Informatics, 2020, 17(6): 3772-3781.
- [27] Min S, Dai Q, Xie H, et al. Cross-modal attention consistency for video-audio unsupervised learning[J]. arXiv preprint arXiv:2106.06939, 2021.

- [28] Oord Avan den, Li Y, Vinyals O. Representation learning with contrastive predictive coding[J]. arXiv preprint arXiv:1807.03748, 2018.
- [29] Wu Z, Xiong Y, Yu S X, et al. Unsupervised feature learning via non-parametric instance discrimination[C]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018: 3733-3742.
- [30] Yuan D, Chang X, Huang P-Y, et al. Self-supervised deep correlation tracking[J]. IEEE Transactions on Image Processing, 2020, 30: 976-985.
- [31] Laine S, Aila T. Temporal ensembling for semi-supervised learning[J]. arXiv preprint arXiv:1610.02242, 2016.
- [32] Zheng Z, Yang Y. Unsupervised scene adaptation with memory regularization in vivo[J]. arXiv preprint arXiv:1912.11164, 2019.
- [33] Zhao H, Yang G, Wang D, et al. Deep mutual learning for visual object tracking[J]. Pattern Recognition, 2021, 112: 107796.
- [34] Chen X, He K. Exploring simple siamese representation learning[C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021: 15750-15758.
- [35] Loparo K. Case western reserve university bearing data center[J]. Bearings Vibration Data Sets, Case Western Reserve University, 2012, 22-28.
- [36] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016: 770-778.
- [37] Ioffe S. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. arXiv preprint arXiv:1502.03167, 2015.
- [38] You Y, Gitman I, Ginsburg B. Large batch training of convolutional networks[J]. arXiv preprint arXiv:1708.03888, 2017.

## 攻读硕士学位期间取得的成果

- [1] Li J Y, **Zhang S**, Zhao Y W, et al. New memory method of impedance elements for marching-on-in-time solution of time-domain integral equation[J]. Electromagnetics, 2010, 30(5): 448-462.
- [2] 李四, **张三**. 时间步进算法中阻抗矩阵的高效存储新方法 [J]. 电波科学学报, 2010, 25(4): 624-631.
- [3] 李四, **张三**. 时域磁场积分方程时间步进算法稳定性研究 [J]. 物理学报, 2013, 62(9): 090206-1-090206-6.
- [4] **张三**, 李四. 时域磁场积分方程时间步进算法后时稳定性研究. 电子科技大学学报 [J] (已录用, 待刊).
- [5] **Zhang S**. Parameters discussion in two-level plane wave time-domain algorithm[C]. 2012 IEEE International Workshop on Electromagnetics, Chengdu, 2012: 38-39.
- [6] **张三**, 李四. 时域积分方程时间步进算法研究 [C]. 电子科技大学电子科学技术研究院第四届学术交流会, 成都, 2008: 164-168.
- [7] **张三** (4). 人工介质雷达罩技术研究. 国防科技进步二等奖, 2008 年.
- [8] XXX, XXX, XXX, XXX, 王升. XXX 的陶瓷研究. 四川省科技进步三等奖, 2003 年 12 月.