



ScienceDirectMeasurement 期刊主页提供目录列表：[www.elsevier.com/locate/measurement](http://www.elsevier.com/locate/measurement)



# 用于机器智能故障诊断的少样本迁移学习

吴静耀、赵志斌、孙创 \*、严如强、陈雪峰

西安交通大学机械工程学院，陕西 西安 710049



## 文章信息

文章历史记录：2020 年 5 月 6 日收到  
2020 年 6 月 9 日以修订形式收到 2020  
年 7 月 5 日接受 2020 年 7 月 11 日在线  
提供

关键词：少样本学习  
智能诊断 迁移学习 元  
学习 旋转机械

## 抽象的

旋转机械大数据智能诊断已得到全面研究，但由于工况的多变性和故障样本的稀缺性，现有的诊断模型与实际应用之间还存在差距。为了解决这个问题，本文利用元学习构建了少样本迁移学习方法，用于可变条件下的少样本样本诊断。我们考虑旋转机械智能诊断的两种迁移情况，即条件迁移和人工到自然迁移，并构建了七种基于统一维卷积网络的少样本迁移学习方法，用于三个数据集的少样本诊断。提供不同样品容量和传输情况下的基线精度，以供全面比较和指导。此外，详细讨论了少样本场景下各种方法的数据依赖性、可迁移性和任务可塑性，数据分析结果表明元学习在相对简单的迁移任务上具有极少样本实例的机器故障诊断优势。我们的代码位于 <https://github.com/a1018680161/Few-shot-Transfer-Learning>。

<b1> </b1> Elsevier Ltd. 保留所有权利。

## 一、简介

### 1.1. 为什么少投转移很重要

航空发动机、高速列车电机、风力发电机等旋转机械 [1-3]，的诊断与预测，对其安全运行和高效工作起着核心作用。近年来，人们引入了基于稀疏分解、流形学习和最小熵反卷积的各种信号处理方法来从振动信号中提取故障关注特征并诊断故障模式 [4-6]。依靠深度学习强大的特征提取能力，智能诊断方法大大超越了传统的信号处理方法，并迅速引起了研究的兴趣 [7,8]。大多数著名的机器学习方法，例如支持向量机（SVM）[9]，深度玻尔兹曼机（DBM）[10]，卷积神经网络（CNN）[11]，自动编码器[12]，循环神经网络（RNN）[13]，生成对抗网络（GAN）[14] 应用于智能诊断领域。使用监督式深度学习方法和不可靠的方法进行智能故障诊断的综合开源基准测试

最近还提出了监督深度迁移学习方法 [15,16]。

然而，基于深度学习的智能诊断的成功依赖于安装在旋转机械上的传感器收集的大数据。尽管智能诊断方法的有效性已通过开源数据集上的大量实验得到证明，[17,18]，小样本学习仍然是旋转机械智能诊断的挑战。在少样本应用场景中，这些方法大多数都会失败，因为深度学习是数据贪婪的，而这一点在智能诊断领域尚未得到技术考虑。为了探索深度学习方法在有限数据下的数据敏感性，这里对轴承、正齿轮和斜齿轮数据集进行了不同样本量的实验结果。实验结果如图 1 所示。受限于极小的训练实例，在目标域上从头开始训练的准确性会随着训练实例数量的减少而显着下降。[19] 提出了一种基于深度非负约束稀疏自编码器的方法，用于有限标记数据的滚动轴承故障诊断，其中有限样本的精度下降也很明显。应该指出的是，虽然 Zhang 等人使用了 few-shot 训练策略。[20] 对于轴承故障诊断，其本质仍然是大实例训练，因为训练集和测试集之间没有类划分。

\* 通讯作者地址：陕西省西安市咸宁西路 28 号 710049 电子邮箱地址：ch.sun@xjtu.edu.cn (C.Sun)。

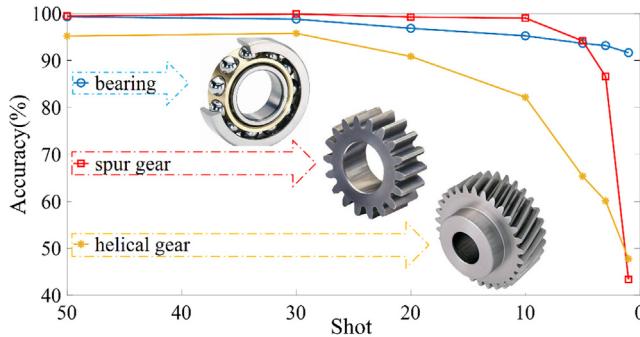


图 1. 诊断精度与射击次数的依赖关系。

## 1.2. 智能诊断如何转

为了解决小样本场景中的性能下降问题，利用相关任务中的知识或数据是一种直观的方法，也称为迁移学习。我们可以通过使用源域上的大数据（例如一种工作条件下的数据）训练分类网络并在目标域上（例如另一种工作条件下的数据）选择性地对其进行微调，将知识从源域转移到目标域。基于情景训练策略的元学习也在 [21] 中被提出，并在各个领域蓬勃发展 [22-24]。训练阶段是通过从源域数据中采样大量元任务在少样本场景中进行的。这意味着模型经过训练可以从少量实例中学。因此，在测试阶段，训练后的网络可以有效地适应目标域中的小样本任务。

根据故障诊断的应用背景和研究现状，不同工况之间的转换（条件转换）和从人工模拟故障到自然故障的转换（人工到自然转换）是可以有效解决样本不足问题的两种主要转换情况。

首先，实践中，众多旋转机械在不同的工况下运行 [25]。例如，机器的转速总是从 0 增加到固定或不断变化的值，机器的负载根据不同的要求而变化。机器的转速、负载和其他参数的变化将显著影响工作条件。为了获得故障检测器的良好推广，几乎每个工作条件都有足够的数据，但由于故障模拟实验的成本很高，这是不切实际的。为了解决这个问题，需要一种小样本学习策略，可以将知识从一个或多个条件转移到其他新条件。

其次，对于航空发动机等昂贵的大型设备，每一次故障都会造成巨大的经济成本和不可接受的人员伤亡，因此不可能获得足够的自然损伤数据来训练深度神经网络进行故障分类。在实践中，对旋转机械进行人工模拟故障，以代表现实中旋转机械损坏的故障 [26-28]。但人工模拟断层与自然断层之间存在较大差距。人工模拟故障通常采用放电加工（EDM）[29]，钻孔 [30]，线电极切割、手工电雕等方法。如图 2 所示，人工模拟故障的轴承有明显的加工痕迹，无论是外观还是机理都与自然故障不同，导致振动模式不同，因此人工模拟故障训练的模型无法有效利用用于自然断层的分类。虽然从人工轴承转移知识

自然轴承损坏的损坏已经在 [31] 中进行，有限样本的影响没有详细研究。为了用少量的自然损坏实例来训练故障检测器，同时利用大量数据进行人工模拟故障，开发用于旋转机械智能诊断的少样本迁移学习方法是必不可少的。

### 1.3. 少投转移中什么最重要

从 few-shot 的角度来看，在从源域到目标域的迁移过程中需要考虑不同方法的三个特点，以便在不同样本容量和不同任务难度下合理地采用 few-shot 学习策略。（1）。在 few-shot 场景下，各种方法的数据依赖性是什么？（2）。模型的微调能力如何影响小样本场景中方法的可迁移性？（3）。如何针对不同难度的迁移情况选择合适的模型？这些可以概括为数据依赖性、可转移性和任务可塑性。

在本文中，为了深入探索这三个不同的特征，基于统一维卷积网络的微调和元学习的七个模型在三个数据集中实现了少样本迁移学习。根据性能比较的结果，得出了数据依赖性、可转移性和任务可塑性三个定性结论。

本文的贡献可概括如下：

- 设计了基于一维卷积网络的统一少样本迁移学习框架，用于工作条件迁移和人工到自然迁移两种典型迁移情况下的机器故障诊断。
- 对设计的特征迁移方法、基于微调的方法和基于元学习的方法进行比较分析，以提供不同样本容量和迁移情况下的基线精度。
- 数据依赖性、可转移性和任务可塑性分析讨论了不同迁移任务下合理的方法选择，为进一步的研究提供指导。

## 2. 背景

### 2.1. 少样本学习

小样本学习最初是由人类学习过程激发的。人类只需一个或几次实例就可以学会识别新物体或掌握新概念，因此机器学习和深度学习模型也被寄予厚望，可以从几次实例中学习。当前研究中的少样本方法可以大致分为三个线程 [32]，即数据增强、数据 / 模型迁移学习和元学习。

不依赖于额外数据集的数据增强是在训练过程中执行的一种简单方法。近年来，生成对抗网络（GAN）因其卓越的性能而被广泛用于增强数据。例如，[33] 中提出了一种自动增强方法，以从设计的搜索空间中找到最佳的图像增强策略。然而，当数据极其稀缺时，基于深度学习的数据增强策略可能会遇到麻烦。

从相关数据中学习并将其转移到目标域是小样本任务最直观的解决方案。代表性方法包括特征转移方法和基于微调的方法。还有一些衍生方法，例如具有部分冻结层的基于微调的方法。此外，知识

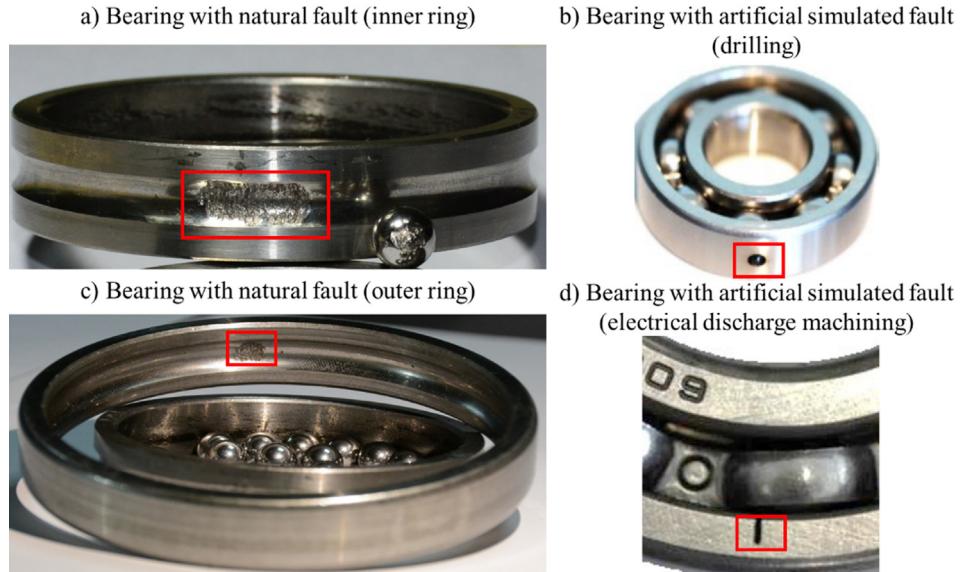


图 2. 具有自然故障的轴承和具有人工模拟故障的轴承之间的差异。

蒸馏是在 [34] 中提出的，采用教师和学生的学习模式。对于特征转移方法，在其他数据集上训练的模型直接用作特征编码器，随后使用冻结特征编码器在少镜头实例上训练新的分类器。对于特征转移方法，可以在相关数据集上对模型进行预训练，并在目标数据上进行微调。

基于元学习的方法的核心是“学会学习”。通过模仿人类的学习过程，元学习试图提高网络对更高级别任务而不是纯粹的分类任务的学习能力。与模型无关的元学习 [35]（MAML）及其衍生物学习如何快速适应少镜头实例。原型网络 [36] 及其衍生物学习评估支持实例和实例之间的相似性。测试实例。

上述 few-shot 学习方法已广泛应用于计算机视觉、强化学习等领域。还进行了各种基线数据集 [37] 和基准研究 [38]，以便对这些方法在计算机视觉方面进行有意义的比较和改进。此外，在医学诊断领域，少样本学习方法，如少样本决策树，被用于诊断超声乳腺肿瘤 [39]。提出了一个三头连体网络来诊断结肠、肺和乳腺组织 [40]，并使用基于匹配神经网络架构的预测模型来诊断青光眼 [41]。

与医学诊断领域类似，准备广泛且标记良好的数据集对于旋转机械智能诊断中的某些任务也是不可行的。为了强调少样本学习对于智能诊断的重要性并为进一步的研究提供指导，有必要对少样本旋转机械智能诊断进行基线研究，并探索不同方法的数据依赖性、可迁移性和任务可塑性。

## 2.2. 少样本迁移学习

基于源域和目标域具有不同分布的假设，迁移学习的主要目标是将源域学到的知识迁移到目标域。因此，迁移学习方法的策略重点是对齐两个领域之间的分布。根据要迁移的表达形式不同，迁移学习方法可以

分为四类 [42]，基于实例的方法、基于特征表示的方法、基于参数的方法和基于关系知识的方法。当支持实例的数量急剧减少时，迁移学习任务变成少样本迁移学习，有限实例的影响与不同领域的影响一样值得注意。虽然之前没有强调，但上面提到的大多数少样本学习方法都是基于迁移场景的。

## 3. 少样本迁移学习框架

### 3.1. 预处理

随着输入长度的增加，神经网络的参数大小会迅速增加，这意味着需要更多的训练数据来匹配模型容量。但输入太小会导致频率分辨率低，不利于有效的故障识别。所以权衡后，模型的输入长度统一设置为 1024。在案例研究过程中，三个数据集中的所有振动信号首先被截断为长度为 2048 的实例。尽管时域信号保留了原始信号的所有信息，但人们普遍认为频域信号比时域信号对故障更敏感。因此，本文将快速傅里叶变换（FFT）应用于每个实例，并将实例的长度减少到原始实例的一半。

### 3.2. 统一基础网络

为了比较的公平性，我们设计了一个统一的一维卷积网络，包括特征编码器网络  $f_\theta()$  和分类器网络  $g_j()$ ，如图 3(a) 和 (b) 所示。特征编码器网络由四个一维卷积块和一个一维自适应最大池层组成，适合处理一维序列信号。我们没有在图像处理中使用小卷积核来捕获图像中的边缘信息，而是为第一卷积层设计了内核大小为 10 的大卷积核，以提取振动信号丰富的浅层信息表示。因为振动信号在时域和频域上对整体的相关性更加敏感，并且第一层的内核尺寸太小会导致

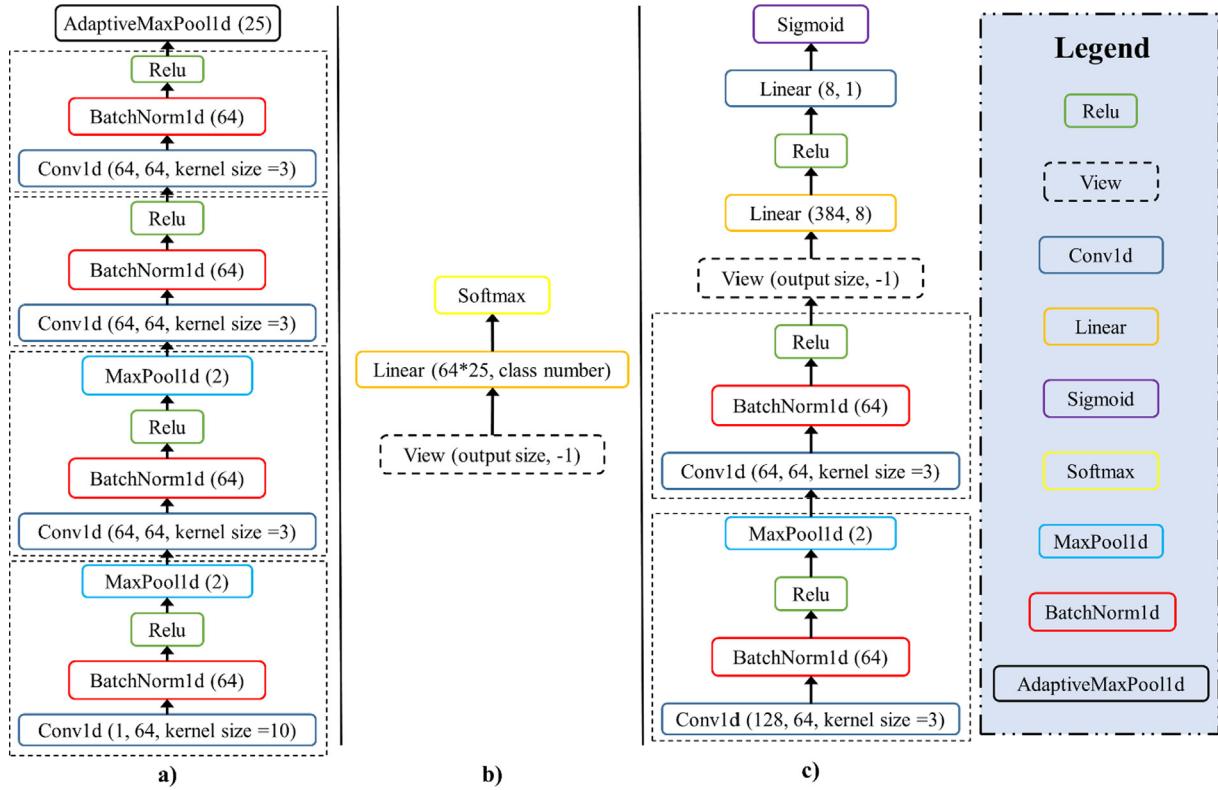


图 3. 一维卷积网络的网络结构 (a) 特征编码器, (b) 分类器, (c) 度量学习器。

丢失信号中的有用信息。特征编码器的输出是一个 1600 维特征张量，形状为 64 个通道 \* 25 个宽度。该特征编码器  $f_o()$  在图 4 中介绍的所有方法中实现。分类器网络由简单连接层和 softmax 激活函数组成。分类器网络  $g_j()$  在除基于元学习的方法之外的所有方法中实现。

的目标域。在目标域中，包含 N 种类型的类，每个类有 K 个实例。因此，为了简洁起见，在下文中，我们将目标域中的训练数据称为 N way \* K shot。在元学习中，来自源域或目标域的数据被分为支持集和查询集，支持集作为标记样本来生成模型的原型特征，查询集作为训练样本来更新模型。

### 3.3.1. 直接训练法

直接训练方法确实是一种直接对目标类别进行分类的简单方法，无需任何预训练或少样本策略。尽管将其与从源域学习特征的其他方法进行比较是不公平的，但它是

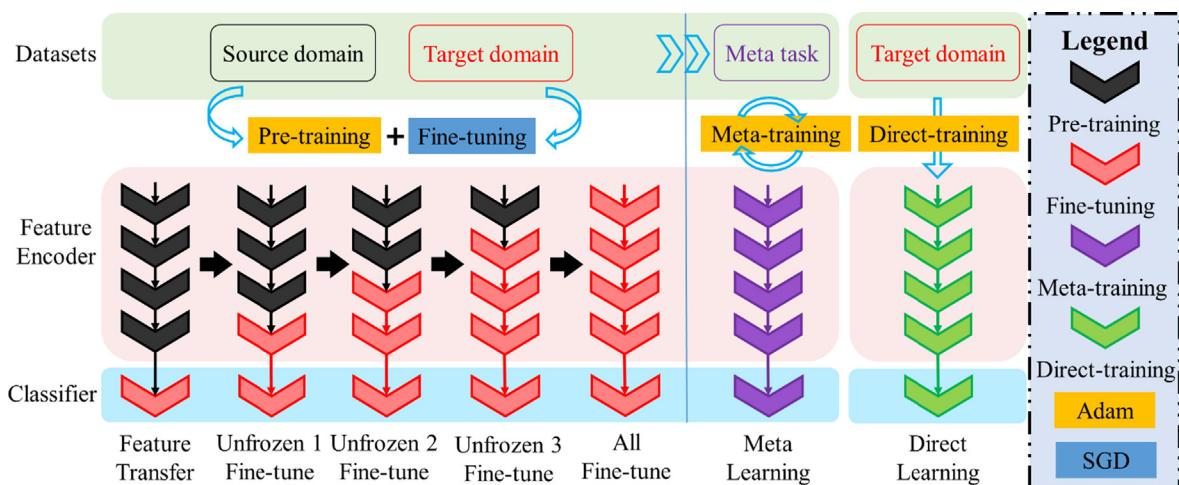


图 4. 所有方法的学习过程。

有必要证明这种简单方法提供下界 off-shot 学习的能力。

直接训练方法的学习过程是直接用目标域的数据训练特征编码器和分类器，如图 4 所示。使用学习率为 0.001 的 Adam 作为优化器。所有 N way \* K shot 数据被分为小批量并用 50 个 epoch 进行训练。最后 10 个 epoch 的平均准确率被视为该方法的最终准确率。下面将直接训练方法统一表示为直接训练网络 (DTN)。

### 3.3.2. 特征转移法

一种迁移方法是在源域中学习良好的特征表示，并将学习到的特征编码器视为固定特征提取器 [43]。特征迁移方法的学习过程可以分为两个阶段，如图 4 所示。首先，Adam 优化器在源域中学习特征编码器，学习率为 0.001，统一训练周期为 200。然后固定之前学习的特征编码器，并使用动量加速随机梯度下降 (SGD) 优化器在目标域中训练新的分类器。[44] 学习率为 0.01。所有 N way \* K shot 数据被分为小批量并用 50 个 epoch 进行训练。最后 10 个 epoch 的平均准确度被视为该方法的最终准确度。当源域和目标域足够相似时，可以利用大量源域知识来学习适合目标域的统一表示。因此在目标分类器的训练过程中，可以大大减少对大实例的依赖。下面将特征传输方法统一表示为特征传输网络 (FTN)。

### 3.3.3. 基于微调的方法

转移从源域学到的知识的另一种方法是训练特征编码器作为良好的初始化，并使用目标域中的数据微调预训练的特征编码器 [45]。基于微调的方法的有效性已经通过 Transformers 双向编码器表示 (BERT) [46] 在自然语言处理中的成功得到了证明。因此，在故障诊断领域，基于振动信号的预训练模型也可能在一定程度上有意义。在本文中，考虑了与特征编码器的四个卷积块相对应的四个不同级别的微调。如图 4 所示，解冻层的数量从 1 层逐渐增加到 4 层。为了简单起见，这四个模型被表示为 Unfrozen m Fine-tuning Net (UmFN)。基于微调的方法的学习过程也可以分为两个阶段。首先，Adam 优化器在源域中以 0.001 的学习率和 200 的统一训练周期学习特征编码器。之后，使用学习率为 0.01 的 SGD 优化器在目标域中对新的分类器和固定了一些层的特征编码器进行微调。所有 N way \* K shot 数据被分为小批量并用 50 个 epoch 进行训练。最后 10 个 epoch 的平均准确度被视为这些方法的最终准确度。

### 3.3.4. 基于元学习的方法

3.3.4.1. 元关系网。[21], Relation Net [47] 中基于 Match Net 的提出，用度量学习器代替传统的距离度量函数（如欧氏距离），可以根据特征编码器编码的特征的特性自适应地学习距离度量函数。Relation Net 在图像分类的各种少样本和零样本任务中取得了最先进的记录。维卷积元关系网络是从关系网络修改而来的，由一个特征编码器组成

图 3(a) 中的  $f_\theta$  和图 3(c) 中的度量学习器  $r_\phi$  具有 2 个卷积块和 2 个全连接层。如图 5 所示，支持集和查询集的两个实例都被输入到特征编码器中，并投影到统一的特征空间中。然后，双分支中提取的特征将连接成关系特征并输入度量学习器。最后，可以计算双分支中的实例之间的相似分数，并将查询实例分类为相似分数最大的类别。

3.3.4.2. 情景训练策略。在传统的分类策略中，所有 C 类的数据被分为训练集  $D^{\text{train}}\{x\}$  和测试集  $D^{\text{test}}\{x_n\}$ 。目标是学习信号  $x \in D^{\text{train}}$  和类标签  $c \in C$  之间的最佳投影函数  $f_\theta(x)$ ，其可以表述如下：

$$\kappa^*, \theta^* = \underset{\theta}{\operatorname{argmin}} \left( \sum \text{Loss}(g_\kappa(f_\theta(x)), c) \right); c \in C, x \in D^{\text{train}} \quad (1)$$

其中，始终选择交叉熵损失作为损失函数，并在信号  $x \in D^{\text{test}}$  上测试  $g_j f(\theta(x))$ 。

与传统的分类策略不同，基于元学习的情景训练策略可以建模为任务学习过程。首先，将来自源域的训练数据划分为训练支持集  $D_s^{\text{source}}\{x\}$  (包括每个训练类的 K 个实例) 和训练查询集  $D_q^{\text{source}}\{x\}$ 。类似地，来自目标域的测试数据被分为测试支持集  $D_s^{\text{target}}\{x_n\}$ ，包括每个测试类的 K 个实例，以及测试查询集  $D_q^{\text{target}}\{x_n\}$ 。目标是学习信号  $x$  和类标签  $c$  之间的投影函数  $r_\phi(f_\theta(x))$ ，其公式如下：

$$\mathbf{r}^*, \theta^* = \underset{\theta}{\operatorname{argmin}} \left( \sum \text{Loss}(f_\theta(x), c) \right); x \in D_s^{\text{source}} \quad (2)$$

在训练阶段，从源域中随机 N 类中选择 K 个实例作为支持集，并选择这 N 类的其余数据作为查询集来更新网络。在测试阶段，选择目标域中 N 类的 K 个实例作为支持集，并选择这 N 类的其余数据作为查询集来测试模型的分类准确性。

采用情景训练策略的元关系网络的学习过程如图 6 所示，包括准备、元学习阶段和迁移阶段。该方法的详细伪代码如表 1 所示。在准备过程中，根据不同任务的需求设置超参数并组织源域和目标域中的数据。在元学习阶段，特征编码器和度量学习器在源域中使用情景训练策略进行训练。使用均方误差 (MSE) 作为损失函数来计算预测的关系分数和标签之间的距离。在传输阶段，查询实例与目标域中的 N 路 \* K 镜头支持实例一起输入学习网络，以验证该方法的测试准确性。使用学习率为 0.001 的 Adam 作为优化器。100 集的平均准确率被视为最终准确率。下面将一维卷积元关系网络统一表示为 MRN。

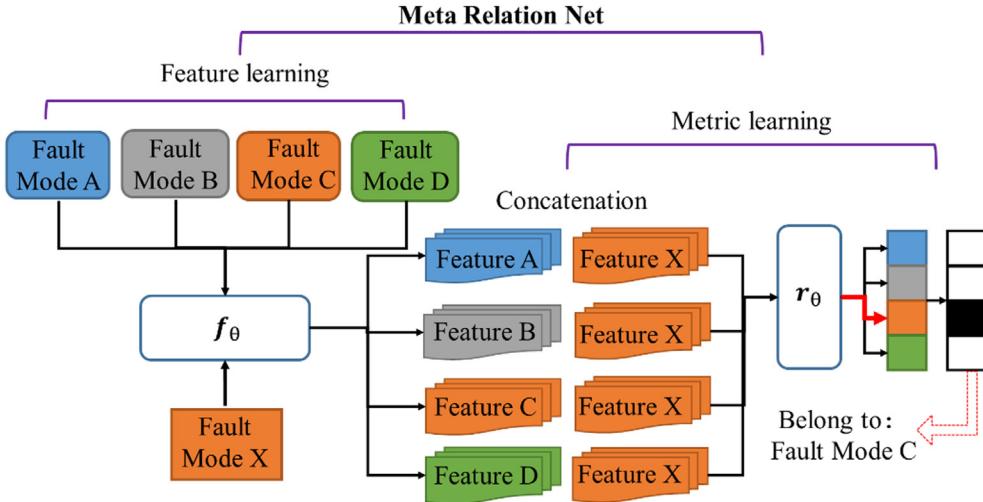


图 5. 具有双输入分支的一维元关系网络。

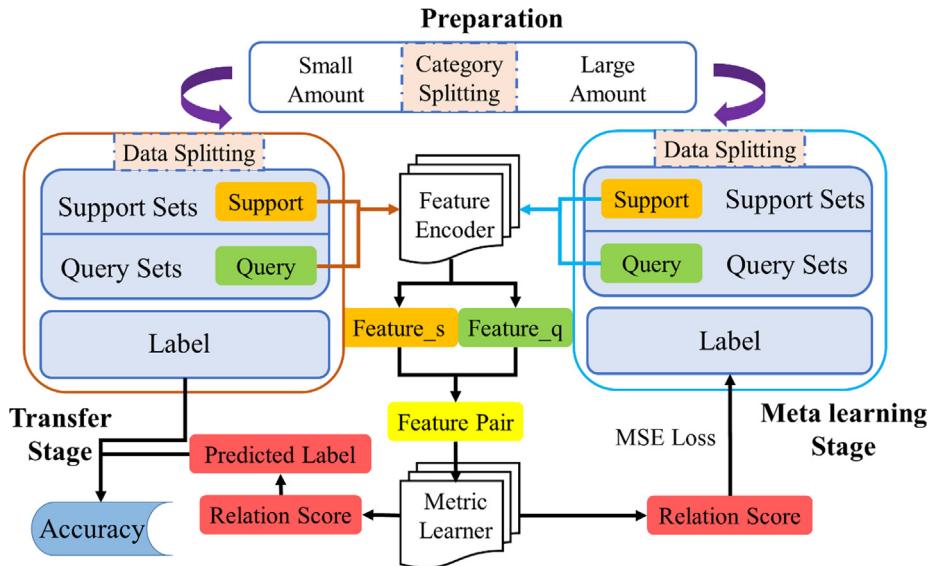


图 6. 采用情景训练策略的元关系网络的学习和迁移过程。

#### 4. 案例研究

##### 4.1. 数据集

###### 4.1.1. PU 轴承故障数据集

PU 轴承数据集 [48] 由 Lessmeier 等人提供，用于轴承状态监测和诊断。32 个 6203 型轴承，包括 12 个人为损坏轴承、14 个通过加速寿命测试的自然故障轴承和 6 个健康轴承。所有这些都在表 2 列出的 4 种不同的工况下进行了测试。同时，故障严重程度根据损坏长度分为两个级别，其中 1 级和 2 级分别代表损坏长度小于和大于 2mm。人工模拟故障考虑了三种处理模式。具有自然故障的轴承是通过加速运行至失效测试生产的。之后，将所有轴承安装在模块化试验台上进行统一测试。实验期间收集了电机电流和振动信号以及三个参数（径向力、负载扭矩和油温）。

在下面的案例研究中，仅使用了 1 个健康轴承、8 个人为损坏轴承和 5 个自然故障轴承的数据。

N09\_M07\_F10 工作条件用于执行人工到自然的小样本迁移学习。这 14 个轴承的详细信息如表 3 所示。时域和频域信号的可视化如图 7 的 (1) 和 (4) 所示。

###### 4.1.2. SQI 齿轮故障数据集

在 SQI 驱动系统诊断综合试验台上进行了条件传递实验。如图 8 (a) 所示，试验台由电机、控制器、两级行星齿轮箱、滚动轴承支撑的两级平行轴齿轮箱、径向加载装置和磁粉制动器组成。它可以模拟直齿轮、斜齿轮、滚动轴承的故障及其耦合故障。

在本实验中，仅考虑平行轴变速箱第一级的齿轮故障，并以三种不同的旋转模式对一个正常的正齿轮和三个故障的正齿轮进行测试。第一模式为加减速循环，从 0rpm 转速加速到 1800rpm，再从 1800rpm 减速到 0rpm，简称“0-30-0Hz”。第二种模式以 1800 rpm 的恒速平稳运行，缩写为“30 Hz”。第三种模式顺利

表1 采用情景训练策略的元关系网络的详细伪代码。

Preparation
<b>1. Set up and initialize</b> feature encoder $f_\theta$ and metric learner $r_\phi$ ; K, N, M → way, shot, query instances number for each class
<b>2. Splitting support and query</b>
$D_s^{\text{source}} \{\mathbf{x}\}, C_s^{\text{source}} \{c\} \rightarrow D_q^{\text{source}} \{\mathbf{x}\}, C_q^{\text{source}} \{c\}$
$D_s^{\text{target}} \{\mathbf{x}\} \rightarrow D_s^{\text{target}} \{\widehat{\mathbf{x}}\}, D_q^{\text{target}} \{\widehat{\mathbf{x}}\}$
Meta-learning Stage
<b>3. Input:</b> $D_s^{\text{source}} \{\mathbf{x}\}, D_q^{\text{source}} \{\mathbf{x}\}, C_s^{\text{source}} \{c\}, C_q^{\text{source}} \{c\}$
<b>For i in train epochs</b>
(1) N*K instances $\{\mathbf{x}^s\}$ from $D_s^{\text{source}} \{\mathbf{x}\}$ M*K instances $\{\mathbf{x}^q\}$ from $D_q^{\text{source}} \{\mathbf{x}\}$
(2) $\{\mathbf{z}^s\}, \{\mathbf{z}^q\} = f_\theta(\{\mathbf{x}^s\}, \{\mathbf{x}^q\})$
(3) $\{\mathbf{z}\} = \text{concat}(\{\mathbf{z}^s\}, \{\mathbf{z}^q\})$
(4) $\{\mathbf{y}\} = \text{softmax}(r_\phi(\{\mathbf{z}\}))$
(5) $\{d\} = \text{Match}(C_s^{\text{source}} \{c\}, C_q^{\text{source}} \{c\})$
(6) <b>Backpropagation</b> Loss = MSE ( $\{\mathbf{y}\}, \{d\}$ )
<b>End For</b>
<b>Output:</b> $f_\theta, r_\phi$
Transfer Stage
<b>4. Input:</b> $D_s^{\text{target}} \{\mathbf{x}\}, D_q^{\text{target}} \{\mathbf{x}\}, C_s^{\text{target}} \{c\}$
<b>For i in test epochs</b>
(1) N*K instances $\{\mathbf{x}^s\}$ from $D_s^{\text{target}} \{\widehat{\mathbf{x}}\}$ M*K instances $\{\mathbf{x}^q\}$ from $D_q^{\text{target}} \{\widehat{\mathbf{x}}\}$
(2) $\{\mathbf{z}^s\}, \{\mathbf{z}^q\} = f_\theta(\{\mathbf{x}^s\}, \{\mathbf{x}^q\})$
(3) $\{\mathbf{z}\} = \text{concat}(\{\mathbf{z}^s\}, \{\mathbf{z}^q\})$
(4) $\{\mathbf{y}\} = \text{softmax}(r_\phi(\{\mathbf{z}\}))$
<b>End For</b>
<b>Output:</b> $\{\mathbf{y}\}$

表2 PU 轴承故障数据集中 4 种工况描述。

Setting name	Load torque [Nm]	Radial force [N]	Rotational speed [rpm]
N15_M07_F10	1500	0.7	1500
N09_M07_F10	900	0.7	900
N15_M01_F10	1500	0.1	1500
N15_M07_F04	1500	0.7	1500

表3 人工到自然迁移小样本学习中使用的轴承的详细信息。

Name	Fault location	Manufacturing way	Severity
KA01	Outer ring	electrical discharge machining	1
KA03	Outer ring	electric engraver	2
KA05	Outer ring	electric engraver	1
KA07	Outer ring	Drilling	1
KA08	Outer ring	Drilling	2
KI01	Inner ring	electrical discharge machining	1
KI03	Inner ring	electric engraver	1
KI05	Inner ring	electric engraver	2
Name	Fault location	Fault description	Severity
K001	Healthy		
KA04	Outer ring	fatigue: pitting	1
KB23	Outer + inner	fatigue: pitting	2
KB27	Outer + inner	Plastic deform.: Indentations	1
KI04	Inner ring	fatigue: pitting	1

以 2400 rpm 的恒定速度运行，缩写为“40 Hz”。因此，本实验包括恒定旋转和可变旋转。齿轮的详细信息如图 8b 所示，包括健康齿轮、缺齿故障、齿根裂纹故障和表面磨损故障。每个故障的位置都用红色矩形标记。振动信号由采样频率为 20480 Hz 的加速度传感器收集。信号在时域和频域的可视化如图 7 的 (2) 和 (5) 所示。

#### 4.1.3. 千鹏齿轮故障数据集

另外在千鹏诊断综合试验台上进行了条件转移实验。如图 9 (a) 所示，试验台由电机、齿轮箱故障模拟、动平衡模拟、轴承故障模拟、径向加载装置和磁粉制动器组成。它可以模拟齿轮箱、轴动平衡、轴承及其联轴器故障。

本实验仅考虑齿轮箱故障，并在 21 种不同转速模式下对 1 个正常斜齿轮和 6 个故障斜齿轮进行了测试。所有这些转速模式分别在 1000rpm、1020rpm、1040rpm、...、1400rpm 的恒速下平稳运行。齿轮的详细信息如图 9 (b) 所示，包括健康齿轮、缺齿故障、四个具有不同尺寸根部裂纹（1 mm、2 mm、3 mm 和 4 mm）的齿轮、一个具有一个齿面磨损故障的齿轮和一个具有双齿面磨损故障的齿轮。表面磨损厚度为 0.5 毫米。振动信号由采样频率为 20480 Hz 的加速度传感器采集。信号在时域和频域的可视化如图 7 的 (3) 和 (6) 所示。

#### 4.2. 条件转让

条件传递是旋转机械故障诊断的常用研究方法。对于旋转机械，工作条件（例如转速和负载）总是随时间变化的。传统的智能诊断方法往往侧重于单一工况或某些工况的故障分类。一旦机器在看不见的工作条件下运行，之前学习的模型就会失败。为了适应新条件下的少样本情况，应进行旋转机械各种条件之间的少样本传递研究。本案例研究在两个数据集（SQI 齿轮故障数据集和千鹏齿轮故障数据集）上进行了实验。利用直接训练网络（DTN）、特征传输网络（FTN）、Unfrozen 1 微调网络（U1FN）、Unfrozen 2 微调网络（U2FN）、Unfrozen 3 微调网络（U3FN）、Unfrozen 4 微调网络（U4FN）和元关系网络（MRN）等七种方法来进行这两个数据集的小样本传输。

在 SQI 齿轮故障数据集上考虑了代表不同转移难度的四种转移情况，其中包括 30 种转移情况。40（从“30 Hz”转换为“40 Hz”），40?30（从“40 Hz”转换为“30 Hz”），0-30-0?40（从“0-30-0 Hz”转换为“40 Hz”），以及 40?0-30-0（从“40 Hz”转换为“0-30-0 Hz”）。对于每种转移情况，考虑从 1 个镜头到 10 个镜头的 4 个镜头组。SQI 齿轮故障数据集的最终精度结果列于表 4 中，其中每种传递情况下的最佳方法以粗体显示，最差的方法以下划线显示。

结果表明，FTN 和 MRN 在所有传输情况下都取得了最佳性能。出乎意料的是，基于微调的方法在大多数情况下实现了最差的性能，甚至低于“低界”（DTN）。这可能是因为不合理地微调特征编码器会破坏学习到的特征表示并适得其反，特别是在极少样本的情况下。所有最好的精度都在 30?40, 40?30, 0-30-0?即使只有 1-shot 实例，40 个情况也接近或高于 98%，这表明两个工作条件之间的信号足够相似，而 40?0-30-0 中的结果都低于 90%，因为从单一工作条件转移到单一工作条件或在单一工作条件之间转移显然更困难。

千鹏齿轮故障数据集中采集了 21 组转速数据。为了进行不同转移难度的少样本转移实验，四种训练 - 测试分裂

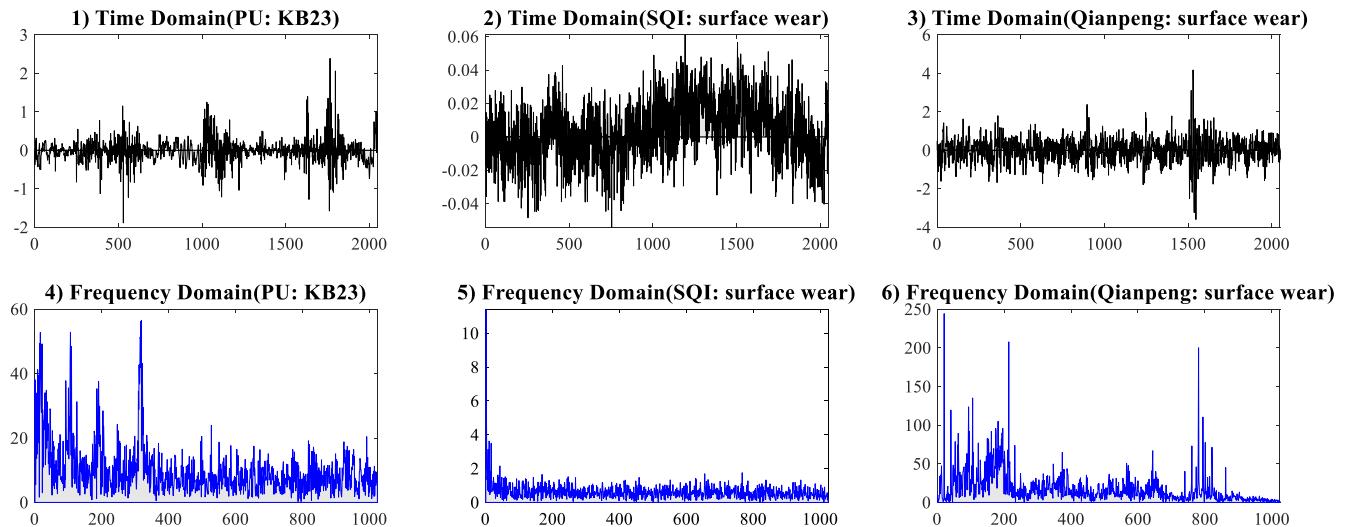


图 7. 来自三个数据集的信号在时域和频域中的可视化。

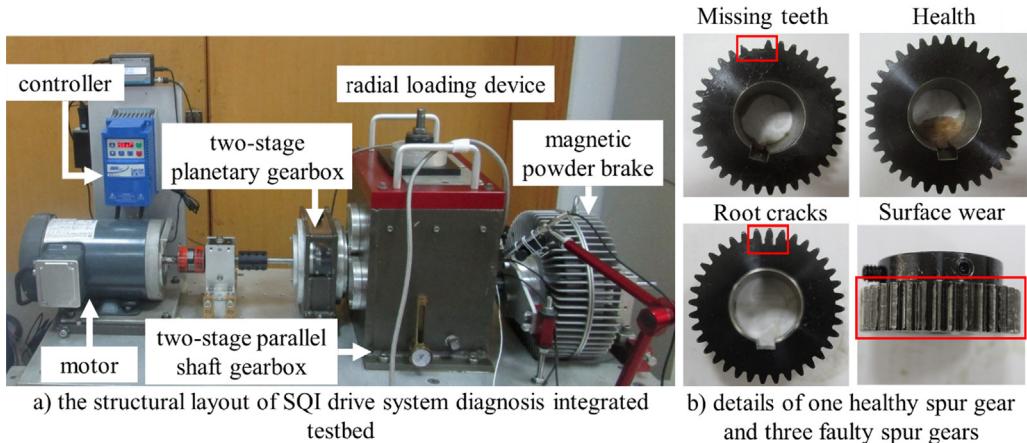


图 8.SQI 数据集。

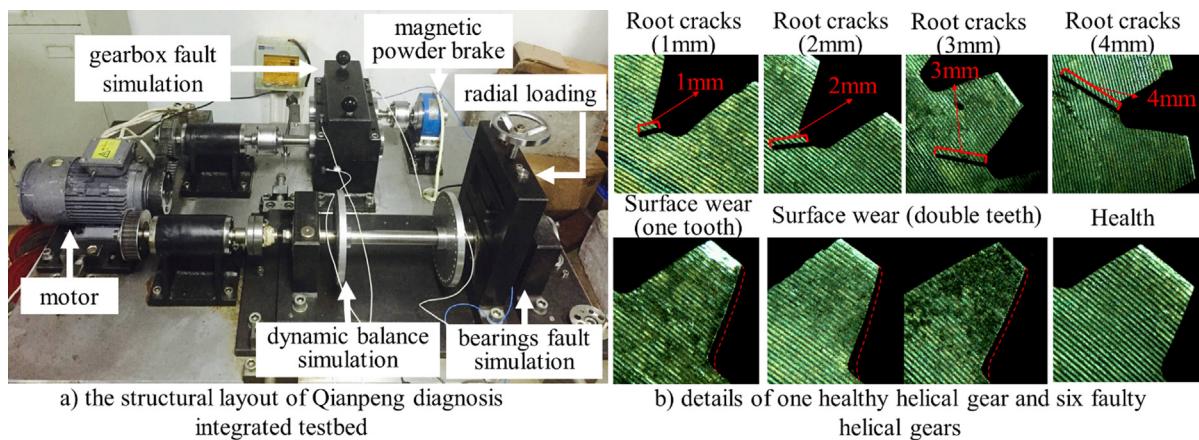


图 9.Qianpeng 数据集。

考虑的情况包括 1? 2-21（从速度 1000 rpm 转换到速度 1020, ..., -1400 rpm），1-5? 6-21（从速度 1000, ..., -1080 rpm 转换到速度 1100, ..., -1400 rpm），1-10? 11-21（从速度 1000 换乘，

...、-1180 rpm 至速度 1200、...、-1400 rpm）和 1-20? 21（从速度 1000, ..., -1380 rpm 转换到速度 1400 rpm）。对于每种转移情况，还考虑从 1 个镜头到 10 个镜头的四个镜头组。最终准确率结果

表 4 SQI 齿轮故障数据集的精度比较（百分比）。

Conditions	Shot	DTN	FTN	U1FN	U2FN	U3FN	U4FN	MRN
30 → 40	1-shot	75.31	90.85	91.75	93.12	72.75	<u>66.57</u>	<b>97.99</b>
	3-shot	97.73	98.93	98.75	77.83	92.35	<u>73.64</u>	<b>99.59</b>
	5-shot	94.25	<b>99.86</b>	97.11	96.25	<u>88.46</u>	94.21	99.44
	10-shot	99.13	<b>99.84</b>	98.55	96.34	97.29	<u>92.32</u>	99.69
	1-shot	<u>43.39</u>	87.51	68.60	73.35	55.41	57.20	<b>98.62</b>
40 → 30	3-shot	86.51	97.00	<u>78.84</u>	<b>99.68</b>	93.89	81.03	99.61
	5-shot	94.17	99.66	99.79	98.99	92.63	<u>85.42</u>	<b>99.91</b>
	10-shot	99.05	<b>100.00</b>	<u>90.52</u>	98.81	94.14	93.12	99.63
	1-shot	72.55	93.32	91.62	82.23	<u>56.28</u>	74.40	<b>98.72</b>
	3-shot	89.36	96.69	94.10	72.04	90.89	<u>69.57</u>	<b>98.91</b>
0-30-0 → 40	5-shot	92.84	98.68	97.90	95.82	<u>83.28</u>	86.97	<b>99.24</b>
	10-shot	94.88	<b>99.86</b>	98.44	94.48	<u>90.23</u>	94.51	99.10
	1-shot	51.66	67.78	47.68	45.94	44.84	<u>35.63</u>	<b>70.38</b>
	3-shot	60.22	<b>86.70</b>	84.27	58.12	69.78	<u>50.91</u>	82.92
	5-shot	81.92	<b>87.31</b>	69.93	74.24	56.57	<u>51.62</u>	85.07
40 → 0-30-0	10-shot	88.33	<b>88.56</b>	84.88	84.75	83.76	<u>70.57</u>	85.99

SQI 齿轮故障数据集如表 5 所示，其中每种传递情况下最好的方法用粗体表示，最差的方法用下划线表示。

与 SQI 数据集中一样，FTN 和 MRN 在所有转移情况下都取得了最佳性能，而基于微调的方法在大多数情况下都取得了最差的性能，甚至低于“低界”（DTN）。出乎意料的是，MRN 在 1?2-21 中实现了最差的精度，这意味着 MRN 在这种情况下根本不起作用。原因可能是一种工作条件转移到多种工作条件偏离了元学习的定义。从源域学到的知识不满足可迁移性的要求，换句话说，MRN 的性能增益取决于源域中丰富的信息。

#### 4.3. 人工到自然的转换

事实上，现有的机器故障诊断方法很大程度上依赖于实验室故障模拟产生的数据，这意味着研究与工业应用之间仍然存在差距。从人为故障转移到自然故障是旋转机械智能诊断应用中解决自然故障样本少的一个重要途径。换句话说，我们可以用成本相对较低的断层样本替代大量成本较高的自然断层样本，并获得相对较高的精度。受此启发，使用相同的七种方法对 PU 轴承故障数据集进行了案例研究，包括人工模拟故障数据和自然故障数据。

表 5 千鹏齿轮故障数据集精度对比（百分比）。

Train-test splitting	Shot	DTN	FTN	U1FN	U2FN	U3FN	U4FN	MRN
1 → 2-21	1-shot	<u>34.57</u>	<b>52.23</b>	48.52	48.37	38.93	48.01	42.76
	3-shot	56.78	<b>67.31</b>	63.54	57.59	58.30	54.58	<u>43.64</u>
	5-shot	64.16	<b>77.20</b>	66.81	73.12	66.13	62.80	<u>43.45</u>
	10-shot	77.43	<b>82.82</b>	74.12	79.17	63.16	76.22	<u>55.69</u>
	1-shot	47.78	53.92	40.29	49.90	48.95	<u>39.01</u>	<b>67.28</b>
1-5 → 6-21	3-shot	<u>60.12</u>	71.99	65.08	70.12	66.74	69.16	<b>86.24</b>
	5-shot	65.35	82.50	71.57	68.45	62.13	<u>60.68</u>	<b>93.88</b>
	10-shot	82.16	91.16	80.91	76.26	<u>73.99</u>	78.16	<b>97.39</b>
	1-shot	<u>41.54</u>	59.50	58.78	55.36	54.45	50.92	<b>88.43</b>
	3-shot	<u>67.79</u>	73.58	78.90	70.62	68.71	68.49	<b>93.28</b>
1-10 → 11-21	5-shot	73.32	88.76	75.32	81.35	78.09	<u>70.83</u>	<b>96.57</b>
	10-shot	85.28	90.47	81.71	86.95	<u>76.93</u>	79.95	<b>94.95</b>
	1-shot	79.63	77.97	69.14	85.48	<u>58.14</u>	77.35	<b>97.84</b>
	3-shot	91.15	95.48	89.79	88.46	87.54	<u>80.77</u>	<b>99.91</b>
	5-shot	94.34	95.54	92.30	<u>81.43</u>	95.46	86.55	<b>99.38</b>
1-20 → 21	10-shot	96.23	97.64	95.57	<u>89.50</u>	92.21	90.22	<b>99.43</b>

表 6 PU 轴承故障数据集的精度比较（百分比）。

Shot	DTN	FTN	U1FN	U2FN	U3FN	U4FN	MRN
1-shot	91.67	87.80	<b>94.27</b>	93.86	91.20	<u>84.48</u>	94.10
3-shot	93.10	<b>98.62</b>	93.10	96.61	92.88	<u>89.57</u>	94.85
5-shot	93.67	<b>98.45</b>	96.84	98.36	94.03	<u>91.89</u>	95.69
10-shot	95.26	<b>97.79</b>	97.00	96.82	<u>92.04</u>	97.14	95.87

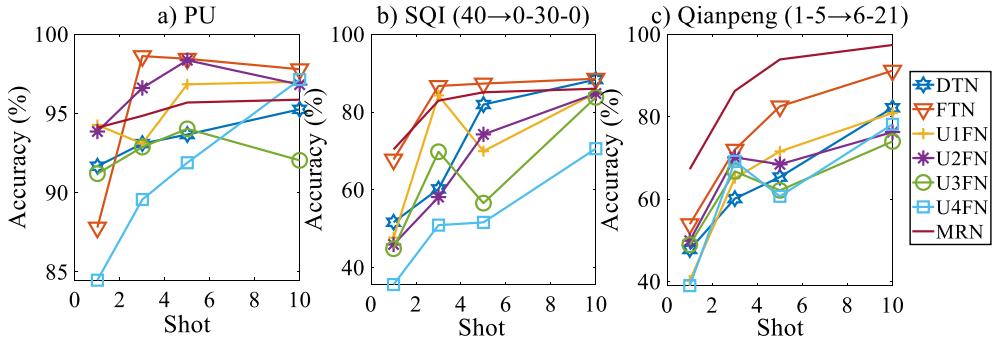


图 10. 七种方法对三个数据集的数据依赖性。

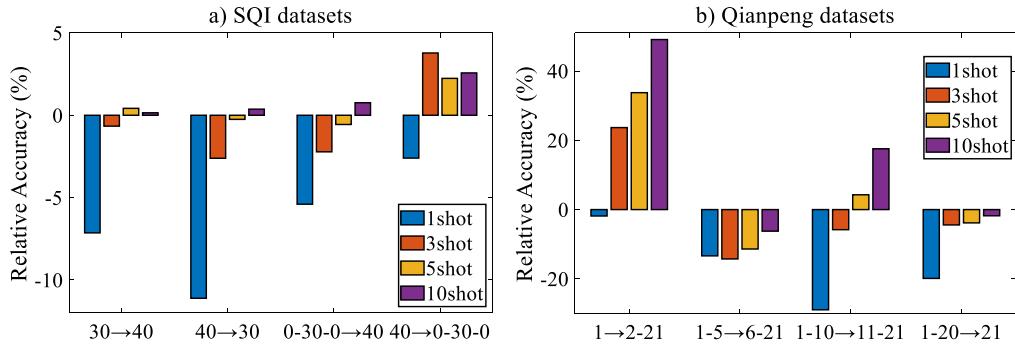


Fig. 11. Data dependency of FTN aSQI 和 Qianpeng 数据集上的 MRN. (相对精度是指减去 a

accuracy of FTN to the accuracy of MRN).

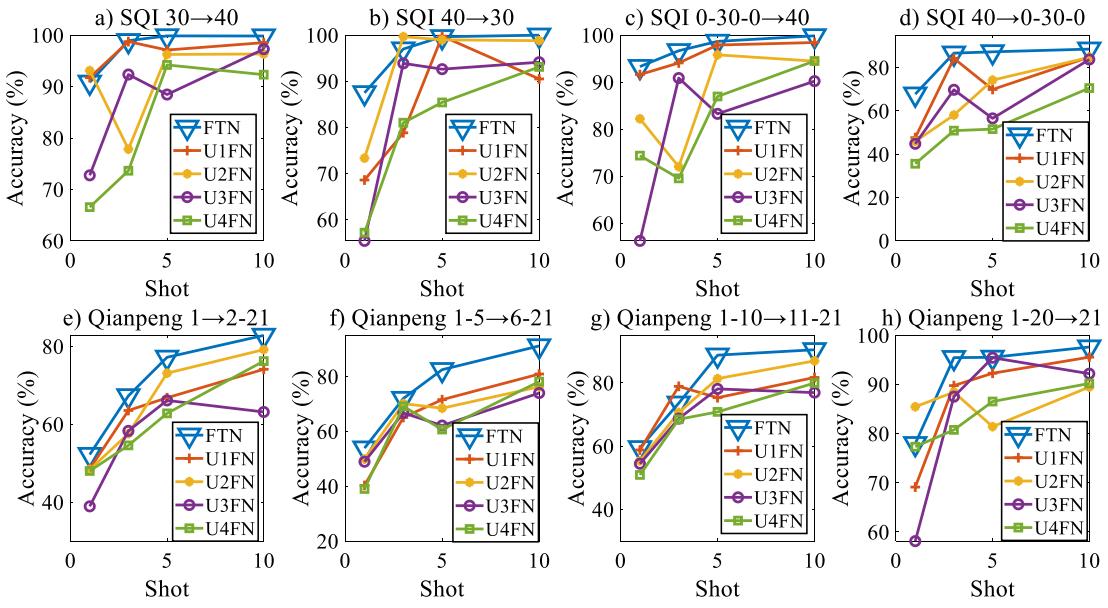


图 12. SQI 和 Qianpeng 数据集的解冻研究。

本案例研究中的转移情况设置为从八个具有人为故障的轴承转移到四个具有自然故障的轴承和一个健康轴承，如表 3 所示。每种转移情况还考虑了从 1 到 10 的四个射击组。最终准确率的结果列于表 6 中，其中每种转移情况下最好的方法用粗体显示，最差的方法用下划线显示。

结果表明，FTN 在 3-shot、5-shot 和 10-shot 情况下达到了最佳精度。而 U1FN 达到了最好的准确率。

在 1 发的情况下。同时，在 1-shot 情况下，MRN 的准确率接近第一。精度最差的也是 U3FN 和 U4FN。可以得出结论，从人工模拟断层数据集中学习的知识可以有效地转移到自然断层数据中，提高分类精度。这为从实验室数据集中学习的模型提高现实中的泛化性能提供了一种可行的方法。

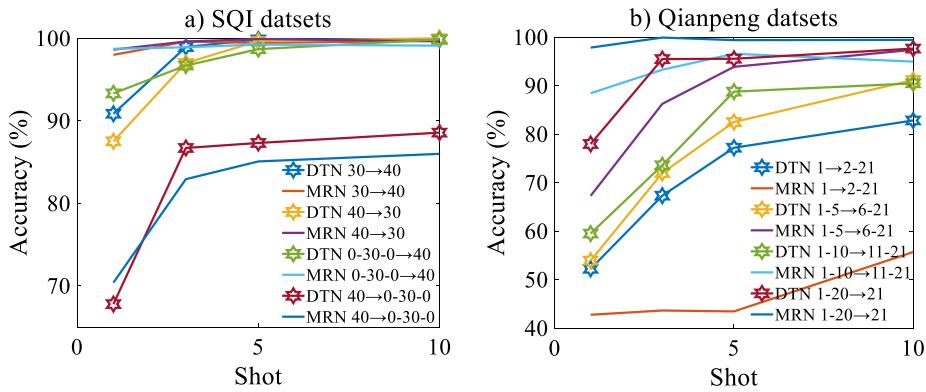


图 13. SQI 和 Qianpeng 数据集的传输难度研究。

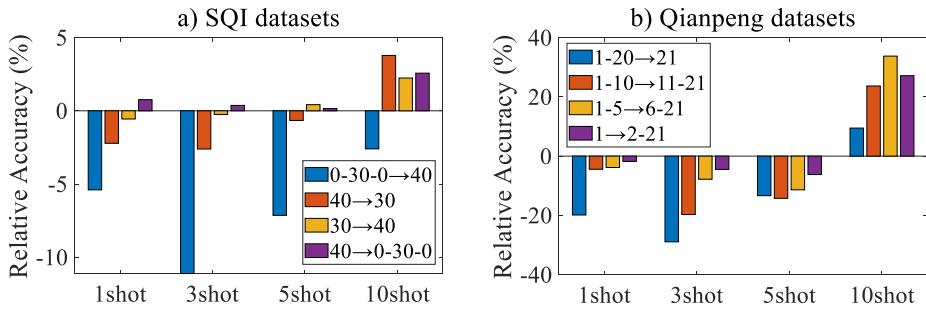


图 14. FTN 和 MRN 在 SQI 和 Qianpeng 数据集上的任务可塑性研究（相对精度是指用 MRN 的精度减去 FTN 的精度）。

## 5 结果分析

条件转移和人工到自然转移的案例研究表明，没有一种方法在所有转移情况下都占主导地位。因此需要针对不同的传输情况采用不同的方法。具体来说，提出数据依赖性分析来研究样本容量对诊断准确性的影响。可迁移性分析的目标是探索模型的微调能力如何影响小样本场景中方法的可迁移性。通过改变源域和目标域之间的关系，可以利用任务可塑性（难度）分析来指导各种迁移情况下方法的选择。

### 5.2. 可转移性研究

基于微调的方法和 FTN 的结果如图 12 所示。可以看出，FTN 优于所有基于微调的方法。没有模型在基于微调的方法中占据主导地位，并且它们的性能依赖于不同的传输情况。显然，不合理的解冻策略会损害分类性能，甚至适得其反。这种现象表明，微调层的容量应该与镜头数量兼容。过多的未冻结层会导致小样本情况下微调的混乱。因此，在极少数镜头集中，仅仅解冻分类器对于这些数据集更有意义。

#### 5.1. 数据依赖性研究

不管随机性的影响如何，可以预期测试精度随着发射次数的增加而增加，如图 10 所示。同时，由于微调中未冻结层过多而导致不稳定性，基于微调的方法的结果存在一定程度的随机波动。

同样值得注意的是，FTN 和 MRN 都无法在所有镜头中占据主导地位。此外，从图 11 中可以看出，随着镜头数量的增加，FTN 的性能逐渐超过 MRN。具体来说，MRN 仅在极少数镜头的情况下占主导地位，当镜头数量达到一定水平时，FTN 将优于 MRN。原因是 FTN 是数据贪婪的，可以从数据增长中获得更多收益，而 MRN 是专门为少样本情况设计的，对数据增长不太敏感。因此，两种方法的数据依赖性导致性能交叉。

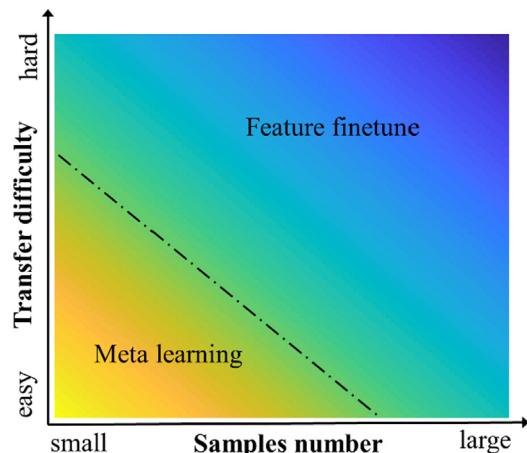


图 15. 小样本情况下 FTN 和 MRN 的任务可塑性和数据依赖性（背景中的黄色和蓝色分别表示 MRN 和 FTN 占主导地位）。

表 7 不同方法的优缺点。

Methods	Pros	Cons
Direct	training	Without source domain data required
Low Feature	accuracy transfer	Suitable for easy transfer task Suitable for medium-shot data
Require Fine-tuning	related source domain data Low computer resources required	Low accuracy
Meta-learning		Suitable for extremely low-shot data Suitable for difficult transfer task
Require	sufficient source domain data	

### 5.3. 任务可塑性（难度）研究

在案例研究中，不同的迁移情况代表不同的迁移难度级别。对于 SQI 数据集，转移难度的等级（从容易到困难）约为 30?40（单对单）= 40?30（单对单）≈ 0-30-0?40（变化到单个）<（单个到变化）。对于 Qianpeng 数据集，迁移难度的等级约为 1-20?21（多对单）< 1-10?11-21（多对多）< 1-5?6-21（多对多）< 1?2-21（单个到多个）。正如预期的那样，图 13 表明，传输情况越困难，DTN 和 MRN 的精度就越低。最简单的情况和最困难的情况之间的准确率差距很大，可以达到 50%。因此，在执行小样本迁移学习任务之前，有必要了解情况的困难程度。还值得注意的是，在旋转机械智能诊断领域，诸如从单一到变化、从单一到变化甚至跨域转移等困难转移的情况更为严重。

从图 14 可以看出，FTN 和 MRN 都不能在所有难度级别上占据主导地位。随着难度的增加，FTN 的性能逐渐超过 MRN。具体来说，MRN 仅在相对简单的情况下占主导地位，当难度达到一定程度时，FTN 将优于 MRN。这表明，当两个领域足够相似时，元学习策略可以更稳定地将源领域知识转移到目标领域，同时避免过度拟合。但随着两个域之间的差异变得更大，对目标域中的样本进行微调变得更加必要。

### 5.4. 综合分析

结合上面的分析，镜头数量和任务难度对 FTN 和 MRN 的综合影响可以更清晰地表达，如图 15 所示。MRN 可以在极少镜头和相对容易的迁移情况下占据主导地位，而特征迁移方法可以随着镜头数量或迁移难度的增加而占据主导地位，这将指导未来少镜头迁移学习的发展，利用这两种方法。此外，我们在表 7 中总结了本文构建的不同方法的优缺点。

## 六、结论

本文研究了考虑迁移的旋转机械智能诊断中的少样本迁移学习问题。

条件转移和人工到自然转移的情况。包括 DTN、FTN、U1FN、U2FN、U3FN、U4FN 和 MRN 在内的七个模型是通过专门为振动信号设计的统一卷积基础网络来实现的。对 PU 轴承故障数据集、SQI 齿轮故障数据集、千鹏齿轮故障数据集进行了案例研究。为旋转机械智能诊断中不同样本容量下的详细研究提供了基准精度。

少样本情况下迁移学习的结论总结为：（1）测试精度随着样本数量的增加而增加。元关系网络往往在极少的镜头 (<5) 中占据主导地位，并且随着镜头数量的增加，特征传输网络逐渐占据主导地位。（2）测试精度随着传输难度的降低而提高。元关系网更具体于简单的传输任务。当源域和目标域之间的相似性降低时，特征传输网络优于其他网络。（3）太多的解冻层数会导致微调的混乱甚至适得其反，削弱模型的可移植性，因此，在极少数样本集下，仅仅解冻分类器更有意义。这些结论将引起诊断准确性之外的更多关注，并为旋转机械智能诊断中的小样本迁移学习的进一步研究提供指导，例如。极少样本特定方法、跨域少样本迁移方法和混合模型少样本迁移方法。

### CRediT 作者贡献声明

Jingyao Wu: 概念化、方法论、可视化、写作原稿、写作评论和编辑、形式分析。赵志斌: 概念化、数据管理、调查。孙创: 概念化、资金获取、项目管理、写作 - 初稿、写作 - 审查和编辑。严如强: 资金获取、监督、数据管理。陈雪峰: 资金获取、监管、资源。

### 宣布竞争利益

作者声明，他们没有已知的可能影响本文报道的工作的相互竞争的财务利益或个人关系。

### 致谢

该工作得到国家重点研发计划 [No. 2018YFB1702400]，国家自然科学基金 [No. 51705398]，陕西省 2020 年自然科学基础研究计划 [No. 2020JQ-042]，和航空科学基金 [No. 2019ZB070001]。

### 参考

- [1] L. 崔小王, 徐勇, 姜红, 周建, 一种滚动轴承剩余寿命预测的切换无迹卡尔曼滤波器方法, Measurement 135 (2019) 678–684, <https://doi.org/10.1016/j.measurement.2018.12.028>.
- [2] L. 宋红, 王, 陈平, 基于症状提取和三价逻辑模糊诊断理论的设备逐步模糊诊断方法, IEEE Trans. 模糊系统 (2018), <https://doi.org/10.1109/TFUZZ.2018.2833820>.
- [3] L. 崔军黄, 张凤, 褚凤, HVS RMS 定位公式及定位律: 球轴承外圈故障定位诊断, 机械. 系统. 信号处理。 (2019), <https://doi.org/10.1016/j.ymssp.2018.09.043>.
- [4] Y. 李克, 丁克, 何刚, 焦晓, 基于稀疏分解和阶次跟踪的齿轮箱故障诊断非平稳振动特征提取方法, 测量 124 (2018) 453-469. <https://doi.org/10.1016/j.measurement.2018.04.063>.
- [5] X. 蒋, 成, 石, 黄, 沈, 朱, 一种新的 I0 范数嵌入式轴承损伤早期故障诊断方法, 测量 127 (2018) 414-424. <https://doi.org/10.1016/j.measurement.2018.06.016>.

- [6] X. 丁强, 李强, 林丽, 何强, 邵勇, 旋转机械故障诊断中瞬态特征提取的快速时频流形学习及其重构, 测量 141 (2019) 380-395. <https://doi.org/10.1016/j.measurement.2019.04.030>. [7] H. Wang, S. Li, L. Song, L. Cui, 一种基于多振动信号图像融合的基于卷积神经网络的新型故障识别方法, Comput. Ind. (2019), <https://doi.org/10.1016/j.compind.2018.12.013>. [8] X. 郭, 陈丽, 沈春, 分层自适应深度卷积神经网络及其在轴承故障诊断中的应用, 测量 93 (2016) 490-502. <https://doi.org/10.1016/j.measurement.2016.07.054>. [9] A. 维多多, B.S. 杨, 支持向量机在机械状态监测与故障诊断中的应用, 机械工程学报, 2017, 系统. 签名. 过程. (2007), <https://doi.org/10.1016/j.ymssp.2006.12.007>. [10] G. 胡, 李, 夏, 罗, 深度玻尔兹曼机与多粒度扫描森林集成协同方法及其在工业故障诊断中的应用, 计算机. Ind. (2018), <https://doi.org/10.1016/j.compind.2018.04.002>. [11] X. 李, 张, 丁, J.Q. 孙, 滚动轴承故障诊断的多层次自适应方法, 信号处理. (2019), <https://doi.org/10.1016/j.sigpro.2018.12.005>. [12] H. 邵, 江, 赵, 王, 一种新型的旋转机械故障诊断深度自编码特征学习方法, 机械工程学报. 系统. 信号处理. (2017), <https://doi.org/10.1016/j.ymssp.2017.03.034>. [13] H. 刘, 周, 郑, 江, 张, 基于循环神经网络自动编码器的滚动轴承故障诊断, ISA Trans. (2018), <https://doi.org/10.1016/j.isatra.2018.04.005>. [14] X. Li, W.Zhang, Q.Ding, 使用深度生成神经网络的滚动轴承跨域故障诊断, IEEE Trans. 工业电子. (2019), <https://doi.org/10.1109/TIE.2018.2868023>. [15] Z. 赵等人, 用于智能故障诊断的无监督深度迁移学习: 一项开源和比较研究, 2019 年 12 月. [16] Z. 赵等人, 旋转机械智能诊断的深度学习算法: 开源基准研究, 2020 年 3 月. [17] Z. Chen, K. Gryllias, W. Li, 使用可转移卷积神经网络的旋转机械智能故障诊断, IEEE Trans. Ind. 信息学 16 (1) (2020) 339-349, <https://doi.org/10.1109/TII.2019.2917233>. [18] Z. 吴江华, 赵坤, 李晓, 一种自适应深度迁移学习轴承故障诊断方法, 测量 151(2020)107227, <https://doi.org/10.1016/j.measurement.2019.107227>. [19] X. Li, H. Jiang, K. Zhu, R. Wang, 一种用于滚动轴承故障诊断的深度传输非负约束稀疏自动编码器, 具有少量标记数据, IEEE Access 7 (2019) 91216-91224, <https://doi.org/10.1109/ACCESS.2019.2926234>. [20] A. 张, S. Li, Y. Cui, W. Yang, R. Dong, J. Hu, 基于少样本学习的有限数据滚动轴承故障诊断, IEEE Access 7 (2019) 110895-110904, <https://doi.org/10.1109/access.2019.2934233>. [21] O. Vinyals、C. Blundell、T. Lillicrap、K. Kavukcuoglu、D. Wierstra, 一次性学习的匹配网络, Adv. 神经信息. 过程. 系统. (2016) 3637-3645. [22] J. Nie, N. Xu, M. Zhou, G. Yan, Z. Wei, 基于少样本学习的 3D 模型分类, 神经计算 (2019), <https://doi.org/10.1016/j.neucom.2019.03.105>. [23] G.J. 阿吉亚尔, R.G. 曼托瓦尼, S.M. Mastelini, A.C.P.F.L. 德卡瓦略, G.F.C. 坎波斯, S.B. 初级, 一种用于选择图像分割算法的元学习方法, 模式识别. 莱特. 128 (2019 年 12 月) 480-487, <https://doi.org/10.1016/j.patrec.2019.10.018>. [24] Olier 等人, Meta-QSAR: 元学习在药物设计和发现中的大规模应用, Mach. 学习. 107 (1) (2018 年 1 月) 285-311, <https://doi.org/10.1007/s10994-017-5685-x>. [25] Y. 雷斌, 杨、X. 蒋、F. 贾、N. Li、A. K. Nandi, 机器学习在机器故障诊断中的应用: 回顾和路线图, Mech. 系统. 信号处理. 138 (一月) (2020), <https://doi.org/10.1016/j.ymssp.2019.106587>. [26] M. Blodt、P. Granjon、B. Raison、G. Rostaing, 使用定子电流监控的感应电机轴承损坏检测模型, IEEE Trans. 工业电子. 55 (4) (2008) 1813-1822, <https://doi.org/10.1109/TIE.2008.917108>. [27] S. 南迪, H.A. Toliyat, X. Li, 电动机的状态监测和故障诊断 - 评论, IEEE Trans. 能量转换器. 20 (4) (2005) 719-729, <https://doi.org/10.1109/TEC.2005.847955>. [28] J. Zarei, J. Poshtan, 用于轴承故障检测的先进 Park 向量方法, Tribol. 国际. 42 (2) (2009) 213-219, <https://doi.org/10.1016/j.triboint.2008.06.002>. [29] S.A. Niknam, T. Thomas, J. Wesley Hines, R. Sawhney, 不平衡轴承的声发射数据分析, Int. J. 普罗格. 愈合. 马纳格. 4 (特刊 2) (2013). [30] Y. Amirat, V. Choqueuse, M. Benbouzid, 使用发电机定子电流单极分量进行基于 EEMD 的风力涡轮机轴承故障检测, Mech. 系统. 信号处理. 41 (1-2) (2013) 667-678, <https://doi.org/10.1016/j.ymssp.2013.06.012>. [31] Y. Chen, G. Peng, C. Xie, W. Zhang, C. Li, S. Liu, ACDIN: 弥合轴承故障诊断中人工轴承损坏和实际轴承损坏之间的差距, 神经计算 294 (2018) 61-71, <https://doi.org/10.1016/j.neucom.2018.03.014>. [32] J. 舒, 徐, 孟德, 大数据时代的小样本学习, 2018, 第 1-76 页. [33] E.D. Cubuk、B. Zoph、D. Mane、V. Vasudevan、Q.V. Le, Autoaugment: 从数据中学习增强策略. 在: 过程. IEEE 计算. 苏克. 会议. 计算. 维斯. 模式识别, 卷. 2019 年 6 月, 没有. 第 3 节, 第 113-123 页, 2019 年. <https://doi.org/10.1109/CVPR.2019.00020>. [34] N. Frosst, G. rey Hinton, 将神经网络蒸馏成软决策树, 载于: CEUR 研讨会论文集, 卷. 2071, 2018. [35] C. Finn、P. Abbeel、S. Levine, 用于快速适应深度网络的模型不可知元学习, 见: 第 34 期 Int. 会议. 马赫. 学习. ICML 2017, 卷. 2017 年第 3 期, 第 1856-1868 页. [36] J. Snell、K. Swersky、R. Zemel, 用于小样本学习的原型网络, 见: Adv. 神经信息. 过程. 系统, 卷. 2017 年 12 月, 没有. Nips, 2017 年, 第 4078-4088 页. [37] E. Triantafillou 等人, 元数据集: 用于从少数示例中学习的数据集的数据集, 2019 年, 第 1-22 页. [38] Y. 郭, N.C.F. Codella、L. Karlinsky、J.R. Smith、T. Rosing、R. Feris, 跨域少样本学习评估的新基准, 2019 年. [39] Q. Huang, F. Zhang, X. Li, 使用 BI-RADS 特征诊断超声乳腺肿瘤的少样本决策树, Multimed. 工具应用 77 (22) (2018) 29905-29918, <https://doi.org/10.1007/s11042-018-6026-1>. [40] A. Medela 等人, 组织病理学图像中的很少的镜头学习: 减少生物数据集上标记数据的需求, 见: Proc. - 国际. 症状. 生物医学. 成像, 卷. 2019 年 -4 月, 没有. Isbi, 2019 年, 第 1860-1864 页. <https://doi.org/10.1109/ISBI.2019.8759182>. [41] M. Kim, J. Zuallaert, W. De Neve, 使用高分辨率眼底图像的小型数据集进行青光眼诊断的少样本学习, 见: MMHealth 2017 - Proc. 第二国际. 工作. 多媒体. 个人. 愈合. 愈合. Care, 与 MM 2017 同处, 2017 年, 第 89-92 页. <https://doi.org/10.1145/3132635.3132650>. [42] S.J. Pan, Q. Yang, 迁移学习调查, IEEE Trans. 知道. 数据工程 22 (10) (2010) 1345-1359, <https://doi.org/10.1109/TKDE.2009.191>. [43] R. Raina, A. Battle, H. Lee, B. Packer, A.Y. Ng, 自学学习: 从未标记数据的迁移学习, ACM Int. 会议. 过程. 序列. 227 (2007) 759-766, <https://doi.org/10.1145/1273496.1273592>. [44] Sutskever、J. Martens、G. Dahl、G. Hinton, 论深度学习中初始化和动量的重要性, 载于: 第 30 届机器学习国际会议, ICML 2013, 2013 年, 第 1 期. 第 3 部分, 第 2176-2184 页. [45] J. Yosinski、J. Clune、Y. Bengio、H. Lipson, 深度神经网络中的特征如何可转移?, Adv. 神经信息. 过程. 系统. 4 (一月) (2014) 3320-3328. [46] J. 德夫林, M.-W. Chang, K. Lee, K. Toutanova, BERT: 用于语言理解的深度双向转换器的预训练, 见: 2019 年北方会议论文集, 2018 年, 第 4171-4186 页. <https://doi.org/10.18653/v1/N19-1423>. [47] F. 宋元阳、张丽、项天祥、P.H.S. 托尔, T.M. Hospedales, 学习比较: 用于少样本学习的关系网络, 见: IEEE 计算机学会视觉和模式识别会议论文集, 2018 年, 第 1199-1208 页, <https://doi.org/10.1109/CVPR.2018.00131>. [48] C. 莱斯梅尔, J.K. Kimotho, D. Zimmer, W. Sextro, 使用电动机的电机电流信号对机电驱动系统中的轴承损坏进行状态监测: 数据驱动分类的基准数据集, 载于: 2016 年第三届欧洲预测与健康管理协会会议, 2016 年, 第 1 期. 厘米, 第 152-156 页.