

引用本文: 毛仕龙, 赵赞善, 王皓宇, 等. 基于深度交叉网络的多任务学习 OLT 设备故障识别算法[J]. 光通信技术, 2025, 49(3): 27-33.

基于深度交叉网络的多任务学习 OLT 设备故障识别算法

毛仕龙¹, 赵赞善^{1,2,3}, 王皓宇¹, 高冠军^{1*}

(1. 北京邮电大学 电子工程学院, 北京 100876; 2. 中国科学院声学研究所南海研究站, 海口 570105;

3. 陵水海洋信息海南省野外科学观测研究站, 海南 陵水 572423)

摘要: 针对光网络数据集不均衡导致的人工智能(AI)模型偏差及特征学习不足问题, 提出一种基于深度交叉网络(DCN)的多任务学习光线路终端(OLT)设备故障识别算法。首先, 通过标准化均值-曼哈顿距离评估潜在故障, 标记高相似度样本为质差数据; 其次, 构建 DCN-多任务学习(MTL)模型, 引入高阶特征交叉增强学习能力, 并以质差检测为辅助任务优化主任务训练。实验结果表明: 与传统的深度神经网络相比, 该算法在相同数据量下准确率提升 1.15%、召回率提升 11.83%、F1 分数提升 6.39%、曲线下面积(AUC)提升 5.91%, 各项指标均突破 0.95, 验证了算法在故障数据稀缺场景下的强检测能力。

关键词: 光接入网; 故障检测; 数据集不均衡; 深度交叉网络; 多任务学习

中图分类号: TN91 **文献标志码:** A **文章编号:** 1002-5561(2025)03-0027-07

DOI: 10.13921/j.cnki.issn1002-5561.2025.03.005

开放科学(资源服务)标识码(OSID):



Fault identification algorithm for OLT equipment based on deep cross network and multi-task learning

MAO Shilong¹, ZHAO Zanshan^{1,2,3}, WANG Haoyu¹, GAO Guanjun^{1*}

(1. School of Electronic, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Hainan Acoustics Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Haikou 570105, China;

3. Lingshui Marine Information Hainan Field Scientific Observation and Research Station, Lingshui Hainan 572423, China)

Abstract: To address the issues of artificial intelligence (AI) model bias and insufficient feature learning caused by imbalanced optical network datasets, this paper proposes a fault identification algorithm for optical line terminal (OLT) equipment based on deep cross network(DCN) and multi-task learning(MTL). First, potential faults are assessed using standardized mean Manhattan distance, with high-similarity samples labeled as poor-quality data. Subsequently, a DCN-MTL model is constructed, incorporating high-order feature interactions to enhance learning capability, while utilizing poor-quality detection as an auxiliary task to optimize the training of the primary fault detection task. Experimental results demonstrate that, compared to traditional deep neural networks, the proposed algorithm achieves improvements of 1.15% in accuracy, 11.83% in recall, 6.39% in F1-score, and 5.91% in area under the curve(AUC) under the same data volume, with all metrics exceeding 0.95. This validates the algorithm's strong detection capability in scenarios with scarce fault data.

Key words: optical access networks, fault detection, dataset imbalance, deep cross network, multi-task learning

收稿日期: 2025-03-17。

基金项目: 国家重点研发计划资助项目(2022YFB2903300)资助; 国家自然科学基金资助项目(62371064)资助。

作者简介: 毛仕龙(2000—), 男, 山东菏泽人, 硕士研究生, 现就读于北京邮电大学电子工程学院光电信息工程专业, 参与多项校企合作项目和一项国家重点研发计划项目。主要研究方向为机器学习辅助的光网络流量预测、故障检测及性能评估。

***通信作者:** 高冠军(1984—), 男, 博士, 教授, 主要研究方向为高速光传输、水下蓝绿激光通信与探测、海底光缆与海洋信息网络等。



0 引言

光网络故障主要分为硬故障^[1]和软故障^[2], 硬故障通常由硬件损坏引发, 例如光纤断裂、接口损坏等, 这类故障可通过光功率计、可视故障定位仪及光时域反射仪等专用工具进行快速定位; 而软故障则通常由设备性能劣化、配置错误等原因导致, 其诊断往往需要依赖领域专家或运营商运维团队的现场分析, 结合历史经验进行综合判断。随着人工智能(AI)技术的发展, 机器学习在光网络软故障检测中的应用逐渐受到关注。但是, 作为数据驱动技术, 机器学习算法的性能

毛仕龙,赵赞善,王皓宇,等:基于深度交叉网络的多任务学习 OLT 设备故障识别算法

高度依赖于数据集质量^[3]。当前光网络设备监控数据集存在严重的类别不平衡问题,这一数据特性极大制约了 AI 算法的实际应用效果。

当前,针对数据类别不平衡问题的主流解决方案是通过过采样或欠采样技术实现数据平衡。代表性方法包括合成少数类过采样(SMOTE)及其改进算法自适应综合过采样(ADASYN)^[4]技术。然而,这类方法对异常值较为敏感,可能生成噪声样本,导致模型可解释性与分类准确率下降。为此,部分研究者采用自编码器结合异常检测技术^[5-6],仅利用正常样本或少量故障数据即可实现优异的检测性能,为解决数据不平衡问题提供了新思路。此外,一些研究者通过专家知识训练决策树模型^[7],将模型生成的专家知识标签作为附加特征引入故障检测模型,显著提升了模型性能。但该方法存在 2 个主要局限:一是决策树模型的训练仍依赖于专家标注的故障数据,在故障样本不足时可能引入噪声标签;二是泛化能力有限,在不同厂商设备场景下需要大量专家先验知识进行适配。当前基于机器学习的故障检测方法^[8-9]普遍存在 2 个不足:一是仅使用低阶特征作为输入,未能充分挖掘特征间的高阶交互关系,限制了模型的学习能力;二是大多采用单任务学习框架,没有利用相关任务的协同训练效应来提升模型性能和泛化能力。此外,现有研究主要集中在光传输网(OTN)的故障管理,得益于其环网拓扑结构的高容错性,而对接入网/城域网等线性拓扑场景的研究相对不足。当前,光接入网中的光线路终端(OLT)作为网络边缘的关键设备,其星型拓扑结构和缺乏冗余保护的特性使其成为网络可靠性的瓶颈环节。与具有环网冗余的 OTN 相比,OLT 设备一旦发生故障将直接导致大规模用户业务中断,严重影响服务质量。

针对这一挑战,本文提出一种基于深度交叉网络(DCN)的多任务学习 OLT 设备故障识别算法(下文简称 DCN-MMoE 算法),以缓解接入网设备性能数据集的类不平衡问题而导致的模型偏差、性能下降以及传统神经网络提取高阶交叉特征困难等问题,并增强模型的训练过程。

1 模型与方法

1.1 总体框架

本文提出的 DCN-MMoE 算法主要解决家庭接入网络中 OLT 设备故障智能识别的技术难题,其模型工作原理图如图 1 所示。在多域光网络架构中,内容网

负责海量数据的存储(如视频、网页、云服务等)与分发;传输网作用是大容量、长距离数据的承载,实现跨区域数据调度;接入网用于为终端用户提供连接入口;控制平面负责对接入网设备的工作状态进行实时监控,并具备数据采集与存储的作用。在家庭接入网络框架结构中,上层骨干网络设备,包括核心路由器(CR)、宽带远程接入服务器(BRAS)及核心交换机等关键设备均部署在专业机房环境中,并配置了主备冗余保护机制。同时,OTN 通常采用环形拓扑结构,通过物理层保护倒换机制确保高可靠性。相比之下,处于网络边缘的 OLT 设备通过分光器以星型拓扑直接连接终端用户,这种架构特性使其成为整个接入网络的单点故障关键点。

在 OLT 故障智能分析模型架构中,OLT 设备均具备运行状态监测功能,可采集接收光功率、发射光功率等工作状态数据。在获取 OLT 设备运维性能数据后,需先进行数据预处理,再输入至潜在故障度量模块。该模块的核心作用是从光模块性能数据集中识别出与已知故障数据相似度高、存在潜在故障风险的数据样本,并将其标注为质差数据。假设光模块的性能数据集包含 M 个正常数据和 P 个故障数据($M > P$),分别标记为 S 和 F ,第 i 个正常样本 S_i 与第 j 个故障样本 F_j 的光模块性能数据分别为 $[s_1^i, s_2^i, \dots, s_r^i]$ 和 $[f_1^j, f_2^j, \dots, f_r^j]$,其中 s 为正常样本, f 为故障样本, r 为模型输入特征数。在数据预处理阶段,首先对所有特征进行标准化处理以消除量纲和数值范围差异,标准化后的数据分别记为 S' 和 F' 。潜在故障度量模块通过计算标准化正常数据与各故障样本之间的相似度,从正常数据中筛选出与故障数据最相似的若干样本组成质差数据集 H ;最终,将正常数据集、质差数据集和故障数据集共同输入至共享专家训练模块和门控网络训练模块进行后续处理。

共享专家训练模块由多个专家模型组成,可以被视为一组神经网络,每个专家模型有自己独立的网络参数,用于在同一个数据集上提取不同维度的特征,旨在处理输入数据并生成多样化的特征表示。共享专家训练模块接受潜在故障度量模块输入的数据,然后利用不同专家模型从不同特征空间对输入数据进行特征提取和生成,每个专家模型针对某个任务或多个任务进行优化,在保留不同任务特征个性的基础上,着重提取不同任务特征间的共性,然后将这些特征进一步用于塔模型训练模块的任务预测。

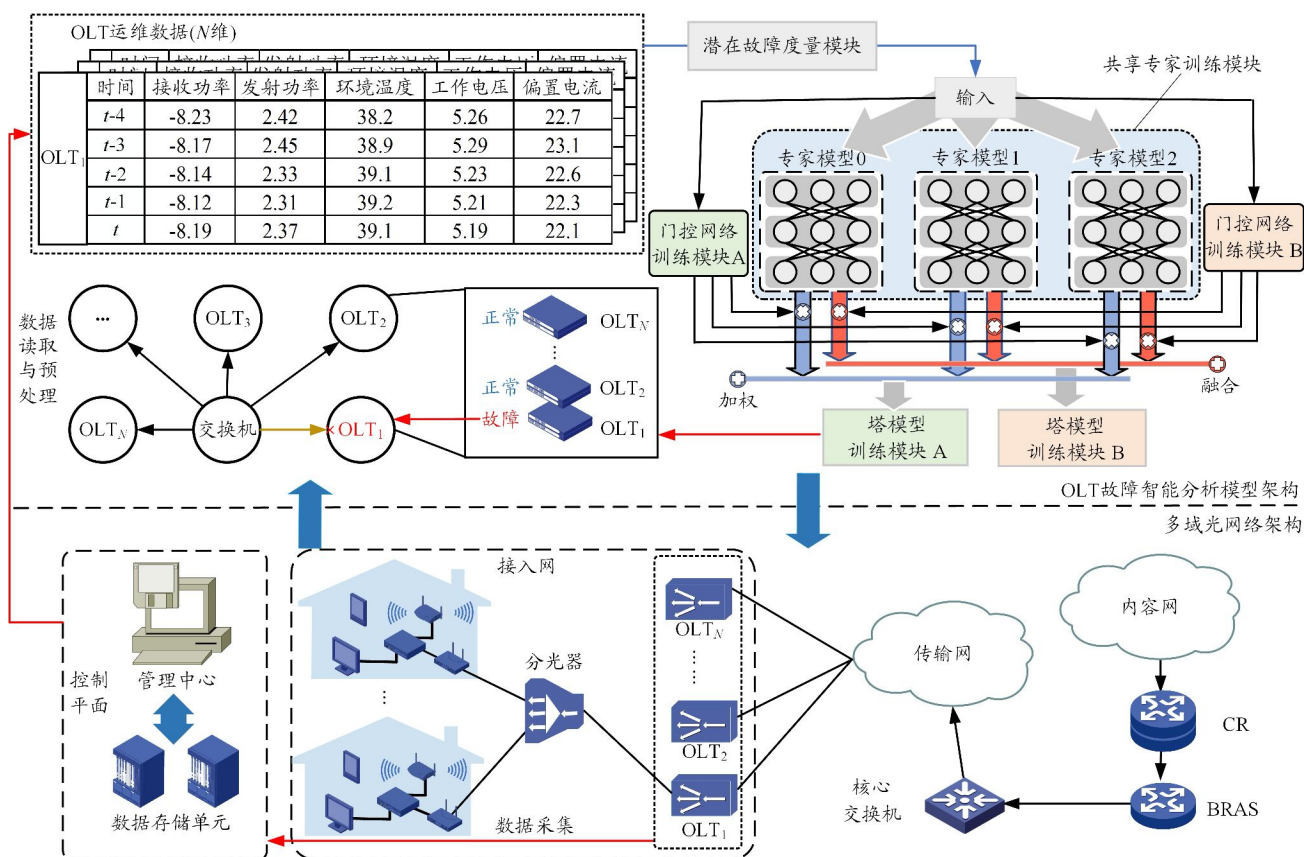


图1 DCN-MMoE 模型工作原理图

门控网络训练模块由多个门控网络组成,每个任务对应一个门控网络,通过门控机制来动态选择不同的专家,在不同任务上做出不同决策。门控网络训练模块接受潜在故障度量模块输入的所有数据,然后根据输入的特征使用软选择机制计算出每个专家网络的权重,进而决策出每个专家模型发挥作用的大小,将专家模型的输出加权组合,最终生成任务相关特征,输入到塔模型训练模块。

塔模型训练模块由多个神经网络塔组成,每个任务都有一个独立的塔,旨在基于专家模型生成的共享特征和门控网络决策的加权系数,进行特定任务的预测。不同任务的塔模型以共享专家模型的输出(经过门控网络加权)作为输入,进行各自任务的学习与优化,最后通过独立的输出层预测每个任务的判别结果。

1.2 潜在故障度量

潜在故障度量模块从正常光模块数据中找出与故障光模块数据相似度最高的一些光模块数据,将其标注为质差。本文利用均值距离标准化分析来度量正常数据与故障数据之间的相似度,该方法可以消除不同参数间量纲的影响,确保不同参数之间的度量具有一致性,具体步骤为:

步骤一:数据标准化处理^[10]。 R 表示光网络数据集,它由正常数据集 S 和故障数据集 F 组成,即 $R=[S, F]^T$, R' 表示 R 经过无量纲化处理后的数据矩阵。矩阵 $\mu=[\mu_1, \mu_2, \dots, \mu_r]^T$ 与矩阵 $\sigma=[\sigma_1, \sigma_2, \dots, \sigma_r]^T$ 分别表示数据集内每个特征数据均值与方差。

为去除光模块数据集中量纲和变化程度对数据的影响。本文采用标准化的方式对 R 进行无量纲处理,使光模块数据样本的方差为1,均值为0。处理方法如式(1)所示。

$$R' = \frac{R - \mu}{\sigma} = \begin{bmatrix} '1 & '1 & \cdots & '1 \\ s_1 & s_2 & \cdots & s_r \\ '2 & '2 & \cdots & '2 \\ s_1 & s_2 & \cdots & s_r \\ \vdots & \vdots & & \vdots \\ 'n & 'n & \cdots & 'n \\ f_1' & f_2' & \cdots & f_r' \end{bmatrix} \quad (1)$$

其中, n 为样本数。

步骤二:计算正常光模块样本与故障光模块样本之间的相似度。若计算正常光模块的第 z 个样本与故障数据之间的相似度,则选取样本 z 标准化后的数据 $S'_z=[s'_1, s'_2, \dots, s'_r]$ 作为参考数列;故障数据集中所有样

毛仕龙, 赵赞善, 王皓宇, 等: 基于深度交叉网络的多任务学习 OLT 设备故障识别算法

本标准化后的数据作为比较数列; 针对光模块的每个性能参数, 单独求出特征相似度, 然后求和再取平均, 作为最终相似度。即以均值-曼哈顿距离衡量二者相似度大小, 数学表达式如式(2)所示。

$$C_z = \sum_{i=1}^P \frac{|F'_i - S'_z|}{P \times r} = \sum_{i=1}^P \sum_{j=1}^r \frac{|f'_i - S'_z|}{P \times r} \quad (2)$$

其中, C_z 为第 z 个正常样本与故障样本的相似度, P 为故障数据个数。

步骤三: 从正常光模块样本中选取相似度较高的样本, 标注为质差数据。以 Label1 列为标签, 标注每个光模块的样本是否故障。在此基础上, 以步骤二计算的相似度结果为依据, 从中选取相似度较高的部分样本作为质差样本, 并对全部样本添加 Label2 列来表示是否质差。

1.3 共享专家网络

本文选取深度神经网络(DNN)和DCN 作为共享专家网络, 重点对比了引入高阶显式交叉特征后, DCN 相较于 DNN 在各项性能指标上的提升效果。

从理论角度看, DNN 虽然通过增加层数能够学习任意复杂的特征组合, 被视为通用的函数逼近器, 但近期研究发现其在实际应用中存在明显局限。当前研究通常直接将低阶、无交叉特征输入神经网络进行故障判别等任务, 然而实验数据表明, 即使是 DNN 也可能难以有效近似二阶或三阶特征交叉, 说明其在学习高阶交叉特征方面存在不足。为解决这一局限性, 研究人员提出了 DCN。DCN 的核心架构包含 2 个关键模块: DNN 模块和交叉神经网络模块。其架构设计旨在增强模型对高阶特征交叉的学习能力, 集成方式包括串行和并行 2 种模式。串行集成模式将交叉网络与深度网络顺序堆叠, 以交叉网络的输出作为深度网络的输入, 先通过交叉网络完成显式特征交叉, 再将结果输入 DNN 进行隐式高阶特征提取; 并行集成模式采用双路并行模式, 交叉网络与 DNN 独立并行处理输入特征, 最终将 2 路输出拼接以达成显式与隐式特征交叉的融合。

在 DNN 的基础上, DCN 通过引入交叉网络来提取高阶显式交叉特征, 并结合了残差连接的思想以缓解梯度消失问题, 从而使更深层网络的训练成为可能。其数学表达式如式(3)所示。

$$X_{l+1} = X_0 X_l^T W_l + b_l + X_l = f(X_l, W_l, b_l) + X_l \quad (3)$$

其中, X_l 和 X_{l+1} 分别为交叉网络的第 l 和 $l+1$ 层的输出, X_l^T 为 X_l 向量的转置, W_l 和 b_l 分别为该层的网络参

数与偏置项。

为了方便进行原理证明, 将偏置项 b 设置为零向量, 假设输入特征仅有二维, 即 $X_0 = [x_{0,1} \ x_{0,2}]^T$, 则

$$\begin{aligned} X_1 &= X_0 X_0^T W_0 + X_0 = \\ &[x_{0,1} \ x_{0,2}]^T [x_{0,1} \ x_{0,2}] [w_{0,1} \ w_{0,2}] + [x_{0,1} \ x_{0,2}]^T = \\ &[w_{0,1}^2 x_{0,1}^2 + w_{0,2}^2 x_{0,1} x_{0,2} + x_{0,1} w_{0,1} x_{0,2} x_{0,1} + w_{0,2}^2 x_{0,2}^2 + x_{0,2}] \end{aligned} \quad (4)$$

由式(4)可知, 当交叉网络叠加到 l 层, 特征交叉的最高阶数可达到 $l+1$ 阶, 并且包含了所有阶的交叉组合。

1.4 门控神经网络与塔神经网络

门控神经网络与塔神经网络都是由 DNN 组成, 它们的数量与任务数量相同, 即每个任务对应一个门控神经网络和一个塔神经网络。

门控神经网络的主要功能是在不同任务间进行决策, 具体方法是为每个专家网络的输出分配相应的权重。这些加权后的输出将被输入到对应的塔神经网络中, 以执行特定任务的判别预测。因此, 门控神经网络输出层的尺寸必须与专家网络的数量相匹配, 并且需要通过 softmax 函数(用 O 表示)进行归一化处理, 以确保各权重值位于 0 和 1 之间且总和为 1。这样, 每个专家网络的输出便根据其重要性得到了适当的加权。随后, 塔神经网络接收上述经过加权处理的专家模型输出作为其输入, 来进行最终的任务预测。判别预测过程可以描述为

$$E = [f_1(x), f_2(x), \dots, f_m(x)] \quad (5)$$

$$G^k = O(h^k(x)) = [g_1^k(x), g_2^k(x), \dots, g_m^k(x)] \quad (6)$$

$$I^k = G^k E = \sum_{i=1}^m g_i^k(x) f_i(x) \quad (7)$$

$$\sum_{i=1}^m g_i^k(x) = 1 \quad (8)$$

其中, m 为共享专家网络个数, k 为任务数, E 为专家网络的输出集合, $f_i(x)$ 为第 i 个专家网络的输出, $h^k(x)$ 为第 k 个门控神经网络的输出, G^k 和 $g_i^k(x)$ 为第 k 个任务门控神经网络的输出及该任务的第 i 个共享专家网络的权重, I^k 为对应第 k 个任务塔网络的输入。

1.5 故障识别判断标准

故障识别的最终目标是准确判断设备是否存在故障, 本质上这是一个二分类问题。在二分类问题中, 其性能可通过混淆矩阵(如表 1 所示)进行评估。混淆矩

表 1 二分类混淆矩阵

类别	正例(预测结果)	反例(预测结果)
正例(真实结果)	TP	FN
反例(真实结果)	FP	TN

阵由4个关键指标组成:真正例(TP),表示模型正确识别出的故障样本数量;真负例(TN),表示模型正确识别的正常样本数量;假正例(FP),表示模型将正常样本误判为故障的数量(误报);假负例(FN),表示模型将故障样本误判为正常的数量(漏报)。

通过混淆矩阵可以计算出准确率、召回率、F1 分数等指标,用以评估分类模型的性能。

2 实验结果

2.1 数据集与实验过程

为了验证所提算法的有效性,本文与国内某运营商合作,收集了 2023 年实际物理网络中 OLT 设备产生的性能数据作为实验数据集。该数据集共包含 51 425 条记录,其中正常数据样本 46 505 条,故障数据样本 4 920 条。对于故障数据的界定,我们通过与相关领域专家讨论及实地测试确定,以性能相关的告警(如光功率、电压和温度等)总持续时间作为评判标准。如果某一告警的持续时间超过 40 000 s,则该实例将被标记为故障状态,并且对应的样本被视为故障样本。

在分类问题中,数据集的均衡性取决于不同类别样本的数量比例。若各类别样本数量相近(比例接近 1:1),则数据集是均衡的;反之,则存在数据不均衡问题。在实际光网络中,故障数据通常较为稀缺。例如,本文所采集的现网 OLT 设备数据集中,正常样本与故障样本的比例高达 9.45:1,表现出严重的不均衡性。为缓解这一问题,本文采用潜在故障度量方法,从原始数据中筛选出与故障模式相似的样本,并标注为质差数据(共 5 080 条)。经调整后,正常数据与异常数据(含故障和质差数据)的比例降至 4.65:1,显著提升了数据分布的均衡性,有助于提高模型的泛化能力。每条数据记录包含以下关键设备监测指标:设备 ID、性能时间、光模块偏置电流(OBC)、光模块接收功率(ORP)、光模块发射功率(OTP)、光模块电压(OV)、光模块温度(OT)等设备信息。

实验的参数配置如表 2 所示,采用双路协同学习架构实现高效特征提取。DCN 主要负责高阶特征交叉建模,而 DNN 则专注于原始输入特征的非线性表征学习,二者输出的特征通过拼接方式融合形成联合表

表 2 模型参数设置

参数名称	值
DNN 输入层	5×64×256×64
DCN 输入层	5×5×5
门控神经网络输入层	5×4
塔神经网络输入层	69×16×1
激活函数	Relu
标准化	BatchNorm1d
优化器	Adam
Batch_size	1 024
损失函数	BCE(Binary CrossEntropy)

征,该表征同时包含交叉特征与原始特征信息。基于该联合表征,多个专家网络并行执行计算任务,每个网络学习不同的特征模式。门控神经网络根据当前任务需求动态调整各专家网络的权重系数,通过加权求和生成适配特定任务的表征向量。最终,该任务特定表征作为输入传递给下游任务塔,实现多任务场景下的差异化建模。

2.2 模型收敛分析

准确率是评估分类模型性能的基本指标,其定义为模型正确预测的样本数占总样本数的比例,数学表达式如式(9)所示。

$$A = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (9)$$

式中, A 为准确率, N_{TP} 为 TP 数量, N_{TN} 为 TN 数量, N_{FP} 为 FP 数量, N_{FN} 为 FN 数量。

实验过程中,本文对 DNN、DCN、DNN-MMoE 以及 DCN-MMoE 这 4 种模型在同一数据集上的性能进行了比较。不同模型的准确率随训练迭代次数的变化趋势如图 2 所示。可以看出,当迭代次数达到 200 次后,4 种模型的准确率均趋于收敛并稳定在 98% 以上。DCN 和 DCN-MMoE 模型的收敛速度明显慢于 DNN

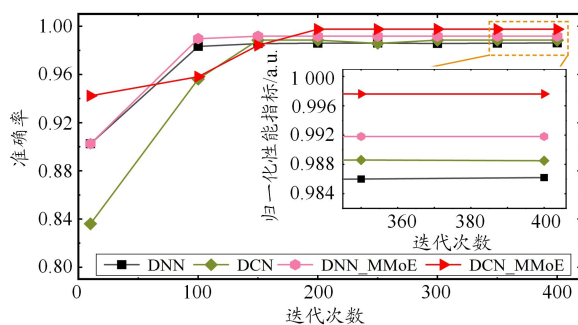


图 2 不同模型在测试集上准确率随迭代次数的变化趋势

毛仕龙, 赵赞善, 王皓宇, 等: 基于深度交叉网络的多任务学习 OLT 设备故障识别算法

和 DNN-MMoE 模型, DCN-MMoE 的收敛速度最慢但最终准确率最高。这是由于 DCN 网络通过显式的高阶特征交叉建模捕获深层特征关系, 这种复杂的交互模式需要更精细的参数调整, 导致训练初期需要更多迭代次数才能学习到有效表征; DCN-MMoE 在 DCN 基础上引入多专家架构, 每个任务需要独立优化其专属的专家网络, 并通过门控机制动态融合专家输出。这种设计虽然提升了模型容量, 但也带来了额外的梯度计算和参数协调开销, 进一步延缓了收敛速度; DNN 和 DNN-MMoE 主要依赖低阶特征组合, 模型结构相对简单, 因而能快速收敛。但这种简化的特征处理方式牺牲了对高阶交互信息的捕捉能力, 导致最终准确率低于 DCN 系列模型。

2.3 模型综合评估

针对数据集中故障样本(“1”标签)占比极低的情况, 仅依赖准确率评估模型性能存在明显局限。例如, 在故障样本仅占 5% 的数据集中, 若模型将所有样本预测为正常(“0”标签), 仍可获得 95% 的准确率, 但这完全掩盖了模型对故障样本的识别能力。因此, 为全面评估所提算法在不平衡数据集上的性能, 本文在考量准确率的同时, 还将召回率、F1 分数和曲线下面积(AUC)纳入评估体系。这些指标能够更准确地反映模型对少数类样本的识别效果, 从而客观验证算法的有效性。

召回率, 亦称查全率, 表示模型正确分类的正例样本数量与实际正类样本数量的比值, 其数学表达式如式(10)所示。随着召回率的提升, 模型对于真实正例样本的检测能力就越强, 从而在光网络设备发生故障时, 能够迅速并准确地进行识别并报警。

$$R_{\text{Recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \quad (10)$$

其中, R_{Recall} 表示召回率。

F1 分数是综合衡量模型精确率与召回率的重要指标, 其数学表达式如式(11)所示。在光网络故障检测任务中, 理想的模型应同时具备高故障检出率和低误报率。F1 分数值越接近 1, 表明模型在识别真实故障与避免误判正常样本上表现越优, 尤其适用于数据高度不平衡的场景。

$$F_1 = \frac{2P_{\text{Precision}}R_{\text{Recall}}}{P_{\text{Precision}} + R_{\text{Recall}}} \quad (11)$$

其中, F_1 表示 F1 分数值, $P_{\text{Precision}}$ 表示精确率。

AUC 是评估二分类模型性能的重要指标, 其取值范围为 0.5~1。AUC 值越接近于 1, 表示模型越具有完

美的分类能力, 能完全区分正负样本; 相反, 如果 AUC 接近 0.5, 则表示模型等同于随机猜测。

DNN、DCN、DNN-MMoE 及 DCN-MMoE 在分类验证集上的效果如图 3 所示。DCN-MMoE 相比最初的 DNN 的准确率提高了 1.15%, 召回率提升了 11.83%, F1 分数增长了 6.39%, AUC 上升了 5.91%。由图 3 可得以下结论:

1) 高阶特征交叉有效性。相比 DNN 模型, 引入 DCN 后的模型在 4 个评估指标上均出现了不同程度的提升。DCN-MMoE 相比 DNN-MMoE 的准确率提升 0.58%, 召回率提升 6.01%, F1 分数提升 3.12%, AUC 值提升 3%。这表明, 利用交叉网络提取高阶特征交叉能够丰富建模信息, 进一步增强模型的学习能力。具体而言, 交叉网络有助于捕捉数据中的复杂模式, 从而提高模型的预测性能。

2) 多任务学习有效性。相比单任务学习, 引入多任务学习后的 4 个模型的评估指标均有显著提升。DCN-MMoE 相比 DCN 在准确率、召回率、F1 分数和 AUC 方面分别提升了 0.91%、9.36%、5.03% 和 4.68%。由此可知, 通过引入如质差判别这样的辅助任务, 可以帮助主要任务(例如故障判别模型)更有效地学习光网络样本从正常状态到质量下降再到发生故障的全过程。即使在数据分布不均衡的情况下, 这种方法也能显著提高模型性能, 实现更高的分类精度, 不仅增强了模型的整体表现, 还提高了其在处理实际问题时的鲁棒性和适应性。

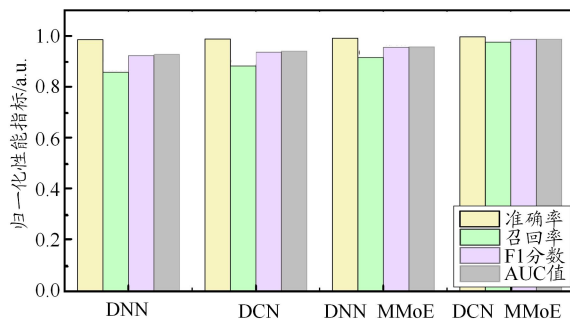


图3 不同模型方案在测试集上的分类结果

2.4 模型建模过程分析

本文以 DCN-MMoE 模型为例, 探究了专家网络数量对模型性能的影响规律, 结果如图 4 所示。可以看出, 随着专家网络数量的增加, 4 个性能指标均呈现先上升后平稳的趋势。在指标上升阶段, 由于数据集规模较大且样本特征丰富, 增加专家网络能显著增强模型的特征提取能力, 4 个评估指标均持续上升, 这表

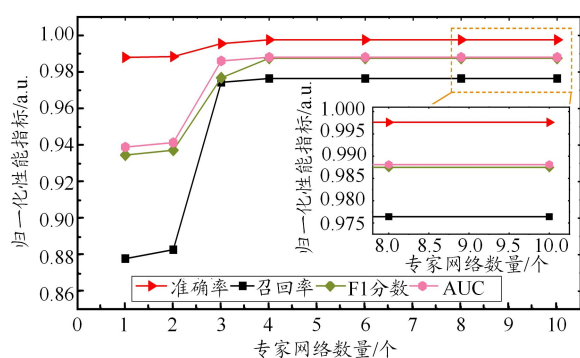


图4 DCN-MMoE 性能随专家网络数量的变化趋势

明在一定范围内扩大网络规模可有效提升模型对复杂特征的捕捉能力。当专家网络达到一定数量后,模型性能进入平稳期,此时继续增加网络规模不再带来明显提升,说明特征表达能力已趋于饱和,模型性能转而受限于数据集本身的质量和规模。此外,通过图中放大图可以看出,在性能收敛状态下,测试集所有模型指标均高于 0.975,说明了 DCN-MMoE 算法在数据稀缺情况下仍能取得优秀的模型综合性能。

结合图 3 中 DCN 以及图 4 中 DCN-MMoE 的实验结果,可以了解到 DCN 的准确率为 98.85%,召回率为 88.28%,F1 分数为 93.72%,AUC 值为 94.13%,而 DCN-MMoE 在专家个数为 1 时,准确率为 98.8%,召回率为 87.78%,F1 分数为 93.45%,AUC 值为 93.89%。这表明 DCN-MMoE 在专家网络个数为 1 时,其模型能力甚至不如单个 DCN,这是因为在本文提出的多任务学习框架中,专家网络是共享的,单个专家网络的特征表示能力有限,同时又受到参数共享与任务间又会发生冲突,即一个专家在尝试优化一个任务的性能时,可能会损坏其它任务的性能。然而,随着专家网络数量的增加,这种现象得到了逐步缓解。

3 结束语

本文提出了一种基于 DCN 的多任务学习 OLT 设备故障识别算法,旨在解决数据类别不均衡导致的模型偏差及特征学习不充分问题。实验结果表明,引入高阶特征交叉后的 DCN 算法在准确率、召回率、F1 分数及 AUC 等关键指标上均显著优于传统 DNN 算法。进一步引入多任务学习框架后,DNN-MMoE 和 DCN-

MMoE 模型性能再获提升,其中 DCN-MMoE 表现最佳,与未引入交叉网络和多任务学习的 DNN 算法相比,准确率提高 1.15%,召回率提升 11.83%,F1 分数增长 6.39%,AUC 上升 5.91%。本研究为故障判别领域提供了新的研究方向和技术手段。

参考文献:

- [1] NYARKO O, ADEKOYA A F, WEYORI B A. Predicting the actual location of faults in underground optical networks using linear regression [J]. Engineering Reports, 2021, 3(3): 1-13.
- [2] LUN H, LIU X, CAI M, et al. GAN based soft failure detection and identification for long-haul coherent transmission systems[C]//2021 Optical Fiber Communications Conference and Exhibition (OFC), March 6-11, 2021, San Francisco, USA. Piscataway: IEEE, 2021: 1-3.
- [3] WANG D, ZHANG Z, ZHANG M, et al. The role of digital twin in optical communication: fault management, hardware configuration, and transmission simulation [J]. IEEE Communications Magazine, 2021, 59(1): 133-139.
- [4] 童莹萍, 冯伟, 宋怡佳, 等. 面向不平衡高光谱遥感分类的 SMOTE 和旋转森林动态集成算法[J]. 遥感学报, 2022, 26(11): 2369-2381.
- [5] LIU S L, WANG D, ZHANG C, et al. Semi-supervised anomaly detection with imbalanced data for failure detection in optical networks [C]//Optical Fiber Communication Conference (OFC), June 6-10, 2021, San Francisco, USA. Washington: Optica Publishing Group, 2021: Th1A-24-1-Th1A-24-3.
- [6] 曾创展, 蔡晨贾农, 姚耀, 等. 基于区块链的白盒化光网络故障检测 [J]. 光通信技术, 2024, 48(2): 72-77.
- [7] ZHANG C, SUN Z, YANG W, et al. Expertise-enhanced machine learning for failure detection on field-deployed optical modules [J]. Journal of Lightwave Technology, 2025, 43(1): 137-154.
- [8] ZHANG C, WANG D, WANG L, et al. Cause-aware failure detection using an interpretable XGBoost for optical networks [J]. Optics Express, 2021, 29(20): 31974-31992.
- [9] ZHANG C, WANG D, JIA J, et al. Potential failure cause identification for optical networks using deep learning with an attention mechanism [J]. Journal of Optical Communications and Networking, 2022, 14(2): A122-A133.
- [10] HAHN G J, SHAPIRO S S. Statistical models in engineering: chapter 3: data standardization and transformation [M]. New York: John Wiley & Sons, Inc., 1994: 45-66.
- [11] WANG R, SHIVANNA R, CHENG D, et al. DCN V2: improved deep & cross network and practical lessons for web-scale learning to rank systems [C]//The Web Conference 2021 (WWW '21), April 19-23, 2021, Ljubljana, Slovenia. New York: Association for Computing Machinery, 2021: 1785-1797.