利用机器学习识别无源光网络中的故障分支

KHOULOUD ABDELLI,[1,3] * CARSTEN TROPSCHUG,[2] HELMUT GRIESSER,[1] STEPHAN PACHNICKE[3]

[1]ADVA Optical Networking SE，弗劳恩霍夫大街 9a，82152 慕尼黑/马丁斯里德，德国[2] ADVA Optical Networking SE，Märzenquelle 1-3，98617 迈宁根，德国[3] 基尔克里斯蒂安-阿尔布雷希特大学，凯撒大街 2，24143 基尔，德国 *通讯作者：kabdelli@adva.com

收稿日期：XXXX年XX月；修订日期：X XXXX年X月X日；接受日期：XXXX年XX月；发布于：XXXX年XX月（文档编号 XXXXX）；出版日期：XXXX年XX月

无源光网络（PON）凭借其宽带宽、低成本部署与维护以及可扩展性，已成为一种颇具前景的宽带接入网络解决方案。为确保传输可靠并满足服务等级协议的要求，必须持续监测PON系统，以便快速识别并定位网络故障，从而降低维护成本、尽量缩短停机时间并提升服务质量。通常，PON系统中的服务中断主要由光纤断点以及光网络单元（ONU）发射器/接收器故障造成。当各光网络单元与光线路终端（OLT）的距离不同时，可通过分析记录的光时域反射仪（OTDR）曲线来识别故障光网络单元或分支。OTDR是一种常用于监测光纤链路的技术。然而，当来自两个或更多长度相近分支的反射重叠时，故障分支的隔离会变得极具挑战性；由于获得的是全局背向散射信号，因此很难区分故障分支。近年来，基于机器学习（ML）的方法在管理PON系统中的光故障方面展现出巨大潜力。当训练和测试数据均来自同一PON系统时，这类技术表现良好。但如果PON系统发生变化，例如增加分支，或改变两个相邻分支之间的长度差异，其性能可能会大幅下降。每次网络变更后都必须重新训练机器学习模型，这可能耗费大量时间。为解决上述问题，本文提出一种通用机器学习方法：该方法不依赖特定网络架构进行训练，可根据OTDR信号识别PON系统中长度接近的分支所对应的故障分支。该方法可应用于任意PON系统，无须因每次网络变更而重新训练。我们使用从PON系统获得的实验数据验证了所提出的方法。© 2022 Optica出版集团

http://dx.doi.org/10.1364/JOCN.99.099999

1. 引言

如今，无源光网络（PON）凭借其诸多优势，已成为全球新兴的宽带光纤接入网解决方案。这些优势包括超大带宽、成本效益高、可扩展性和灵活性等。无源光网络是光纤到户（FTTH）部署的主流技术，能够以经济高效的方式提供多种通信和多媒体服务，包括高质量的三网融合服务，涵盖数据、语音和视频，以及高速互联网接入 [1]。这得益于网络的无源特性和共享光纤基础设施。图 1 展示了.....的总体配置

无源光网络。PON系统由位于电信运营商中心局的光线路终端（OLT）、包含分光器等无源光器件的光分配网络（ODN），以及安装在用户驻地、提供服务的若干光网络单元（ONU）或光网络终端（ONT）组成。PON通常采用两种主流技术：以太网（以太网无源光网络）和吉比特无源光网络（GPON）。由于无需供电，也无需使用在室外设备中更容易发生故障的有源电子器件，PON系统可降低运行维护费用（运营支出）[2]。通过在此类系统中实施和部署有效的主动式或预测性故障管理维护方案，还可以进一步大幅降低运营支出。

PON系统容易发生多种故障，包括ODN光纤故障（例如，馈线光纤和配线光纤都可能受到光纤断裂的影响）、ONT发射器/接收器故障，以及跳线或连接器脏污、断裂或弯折。已部署的PON系统中，超过80%的故障发生在网络的配线/引入段[3]。美国联邦通信委员会（FCC）的报告指出，超过三分之一的服务中断是由光缆问题导致的 [4]，，此类中断可能给服务提供商或运营商造成巨大的经济损失[5]。要人工发现PON中的故障，需要具备相当丰富的专业知识，并花费大量时间进行探测，才能识别、定位并修复故障。这会导致运营支出大幅增加，并使客户满意度下降。因此，监控PON系统有助于自动、准确且自主地识别和定位网络故障，从而降低维护成本、缩短停机时间并提升服务质量。

光纤网络的监测主要通过光时域反射测量（OTDR）实现，这是一种基于瑞利后向散射的技术。OTDR广泛用于测量已敷设光纤链路的特性，例如光纤衰减、长度以及连接器和熔接损耗，也用于检测和定位光链路中的光纤故障或异常。OTDR的工作原理类似于光学雷达：它向待测光纤的一端注入一系列光脉冲。由于菲涅耳反射和瑞利散射效应，部分脉冲会反射回光源。反射信号的幅度会随距离记录下来。由此生成记录的OTDR轨迹（或波形），显示光纤上故障或事件的位置，并用于事件分析。OTDR已广泛用于监测点对点光链路。然而，将OTDR应用于PON系统颇具挑战，因为各分支的后向散射信号会叠加在一起，导致很难区分各分支的反向信号 [4]。对于分支终端等距的情况，事件分析更具挑战性，因为等长分支的反射信号会重叠并叠加，使分支识别变得十分困难。随着PON系统规模扩大，待监测的光纤链路数量随之增加，事件分析的复杂度也会上升，从而降低监测的可靠性 [6]。此外，远端节点处光分路器造成的较大分光损耗，会显著降低后向散射信号（例如，1:32光分路器会使测得的功率损失30 dB[6]）。因此，要准确检测和定位事件，需要ODTR具备较大的动态范围，并采用高灵敏度的故障检测方法。为克服上述挑战，有一种解决方案是使用可调谐OTDR并结合波分复用器，为每个分支配置专用监测波长[7-10]。然而，由于可调谐OTDR设备成本高昂，这种方案相当昂贵。另一种简单有效的方案是在每个分支末端使用光学参考反射器，通过检测参考反射器峰值的出现及其高度变化来提高故障可检测性 [4, 10, 11]。但这种方案要求各分支长度不同，因此不适用于已实际部署的网络。还有一种更简单的分支识别方案，是监测连接的光网络单元（ONU）接收或发送的光功率下降情况；发生光纤断点时，业务会中断。由于各种运营商商业模式下往往无法获取ONU性能数据，这类方案的适用性受到限制。

近年来，机器学习（ML）已成为光通信系统中进行故障管理（例如故障识别、故障定位等）的高效工具 [12-16]。尤其是，基于机器学习的方法在解决 PON 中的故障管理问题方面展现出巨大潜力。在这方面，Usman 等人 [17] 提出了一种用于 PON系统故障监测的数据驱动方法，将光纤传感器与机器学习相结合：采用具有不同特征光栅的光纤布拉格光栅传感器采集监测数据，每个分支对应一种特征光栅，并利用机器学习技术进行故障识别。我们提出了一种基于门控循环单元的自编码器模型，可自动识别 PON 中光纤故障的类型并对其进行全面表征，无需受过培训的人员介入，也无需在网络基础设施上安装额外设备 [18]。然而，该模型无法区分 PON 中发生在长度相近的分支上的故障。基于机器学习技术的特定网络方法可用于识别 PON系统中的故障分支，并且在使用涵盖正常与故障分支状态的特定网络数据进行测试时表现良好。然而，如果将该机器学习模型应用于网络发生变更或扩展后从同一网络获取的新数据（例如增加更多分支、改变相邻分支之间的长度差异、移除部分分支等），该方法的性能就会不佳。因此，网络每次发生变更后，都需要重新训练机器学习模型以提升其性能，这是一项繁琐且耗时的工作 [19]。

针对这一问题，我们提出一种通用机器学习方法，无需依赖网络架构进行训练，即可识别无源光网络系统中分支长度相近（即两个相邻分支的长度差在 1 m 至 3 m 之间）或长度不同情况下的故障分支。该方法利用安装在各分支末端的反射器所获取的监测数据。我们首先将从无源光网络系统中使用传统光时域反射仪记录的光时域反射仪轨迹划分为短的定长序列；在该系统中，每根光纤均由一个光反射器进行监测，且每个序列最多包含两次反射。随后，我们将机器学习模型应用于每个序列，以输出故障分支类别。该机器学习方法经过训练，能够学习不同情况下正常或故障状态所呈现的不同模式，这些模式可能表现为两个相近的反射或单个脉冲，从而准确识别不同的故障分支场景，包括单个分支故障或多个故障。请注意，在本研究中，所提出的机器学习模型仅输出故障分支的标识符，而不输出故障在故障分支上的位置；故障位置可由采用集总反射器的测量方法确定。我们的方法可应用于任意无源光网络系统，无需针对每次网络变更重新训练。我们使用不同条件下从无源光网络系统获取的实验数据验证了所提出的方法。

本研究的主要贡献概述如下：

- 即使在支路长度接近的情况下（例如长度差为 2 m），所提出的模型也能有效识别无源光网络系统中的故障分支。
- 对于 PON 系统中的每项变更，所提出的方法都无需重新训练，因此可节省重新训练机器学习模型和收集新数据所需的时间与精力。
- 所提出的无源光网络监测与监控方法简单快捷；无需在网络基础设施上安装额外且昂贵或复杂的硬件（在每个分支末端放置一个光反射器，并使用传统 OTDR 设备记录 OTDR 轨迹，再将轨迹拆分后输入机器学习模型，以识别故障分支），而且记录 OTDR 测量所需的时间（2ms – 2s）非常短，可加快故障诊断。

本文其余部分的结构如下：第 2 节介绍构建所提出机器学习方法所涉及的机器学习算法背景，并讨论网络专用方法。第 3 节介绍所提出的通用模型。第 4 节总结全文。

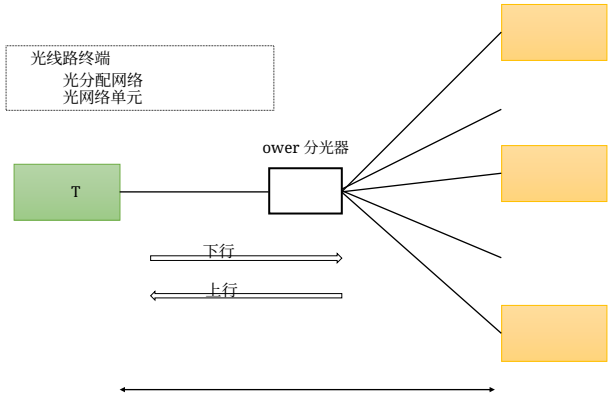


图 1. 无源光网络的总体架构。

2. 背景

本节简要介绍用于开发所提出方法的机器学习技术的基础知识。

2.1 长短期记忆

长短期记忆网络 [20] 是一种特殊的循环神经网络（RNN），能够处理序列数据，并在包括语音识别和自然语言处理在内的各种序列分类应用中实现最先进的性能。

记忆单元（也称为块记忆）是长短期记忆网络的基本计算单元，由权重和三个门组成，用于调节流向细胞状态的信息流。遗忘门决定从细胞状态中丢弃哪些信息。输入门决定将哪些新信息存储到细胞状态中，而输出门决定输出什么。如图 2 所示，先前的细胞状态 $\boldsymbol{c}_{t-1}$ 与先前的细胞输出 $\boldsymbol{h}_{t-1}$ 以及当前输入 $\boldsymbol{x}_t$ 相互作用，以确定内部状态向量中的哪些元素应当更新、保留或丢弃。长短期记忆网络单元通过应用以下方程进行更新：

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_{xf}\boldsymbol{x}_t + \boldsymbol{W}_{hf}\,\boldsymbol{h}_{t-1} + \boldsymbol{b}_f\right) \quad (1)$$
$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_{xi}\boldsymbol{x}_t + \boldsymbol{W}_{hi}\boldsymbol{h}_{t-1} + \boldsymbol{b}_i) \quad (2)$$
$$\tilde{\boldsymbol{c}}_t = \tanh(\boldsymbol{W}_{xc}\boldsymbol{x}_t + \boldsymbol{W}_{hc}\boldsymbol{h}_{t-1} + \boldsymbol{b}_c) \quad (3)$$
$$\boldsymbol{c}_t = \boldsymbol{f}_t \circ \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \circ \tilde{\boldsymbol{c}}_t \quad (4)$$
$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{xo}\boldsymbol{x}_t + \boldsymbol{W}_{ho}\,\boldsymbol{h}_{t-1} + \boldsymbol{b}_o) \quad (5)$$
$$\boldsymbol{h}_t = \boldsymbol{o}_t \circ \text{双曲正切}(\boldsymbol{c}_t) \quad (6)$$

其中，σ 是逻辑 sigmoid 函数，$\boldsymbol{f}$、$\boldsymbol{i}$、$\boldsymbol{c}$ 和 $\boldsymbol{o}$ 分别表示遗忘门、输入门、细胞激活和输出门向量。“∘”表示 Hadamard 乘积算子，所有 $\boldsymbol{b}$ 均为学习得到的偏置向量，所有 $\boldsymbol{W}$ 均为可训练的权重矩阵，而 $\tilde{\boldsymbol{c}}_t$ 是候选细胞值。

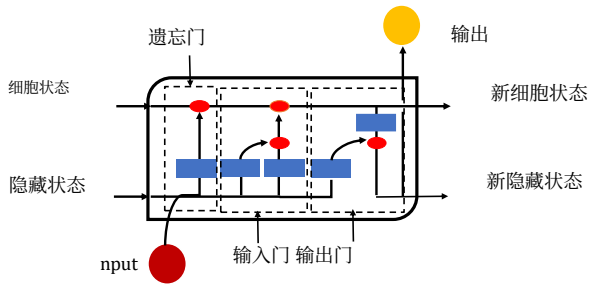


图2：长短期记忆（LSTM）单元的结构。

2.2 特定网络方法

采用基于机器学习技术的特定网络方法，可以解决无源光网络系统中的故障分支识别问题。在这种方法中，机器学习模型使用从无源光网络系统获取的数据进行训练，具体而言，使用较长的 OTDR 序列（即光分路器之后的 OTDR轨迹部分），其中包含来自所有分支的反射。通过学习不同故障场景的模式，包括单个分支发生故障或多条光纤发生多个故障的情况，机器学习方法能够识别故障分支。图 3 展示了机器学习模型的总体架构。机器学习方法以包含所有分支反射的 OTDR 序列作为输入，并输出故障分支标识符（$\boldsymbol{class_{id}}$）。输入 OTDR 序列的长度取决于无源光网络拓扑，尤其是光分路器之后的网络部分（例如，光网络单元的数量、光网络单元之间的长度差异、脉冲宽度以及基于 OTDR 的无源光网络监测所采用的采样等）。光网络单元数量和分支间距离增加时，序列长度也会变长。需要考察的类别数量 $\boldsymbol{l}$（即故障分支场景，包括单个分支发生故障、多个分支发生故障、所有分支均发生故障等情况）取决于分支数量 $\boldsymbol{m}$。$\boldsymbol{l}$ 随 $\boldsymbol{m}$（$\boldsymbol{l} \geq \boldsymbol{2^m}$）而增加。鉴于机器学习模型的输入由序列数据（OTDR 序列）组成，长短期记忆网络（LSTM）或门控循环单元（GRU）等循环神经网络是解决此类问题的理想选择，因为它们擅长处理序列数据并捕捉长期依赖关系。

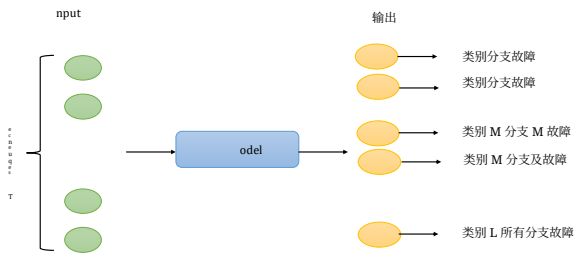


图3：用于识别无源光网络系统中故障分支的、基于机器学习的特定网络方法的总体架构。OTDR序列表示光分路器之后的OTDR轨迹部分，其中包含所有分支的反射。

机器学习模型的输出是故障分支标识符。

在学习阶段，给定来源于用于训练机器学习模型的同一网络或经过轻微修改的网络（例如，在保持与生成训练数据时相同的分支数量的同时，使用更长的馈线光纤或更低的分光比等）的数据，这种机器学习方法能够取得良好的效果

。但是，如果网络发生变化（例如，改变光网络单元之间的长度差异，增加或移除部分光网络单元等），机器学习模型的性能可能会严重下降，因为支撑整个序列中所有分支识别的独特反射模式发生了改变。因此，对于相关的网络变化，必须重新训练机器学习模型。机器学习方法的架构必须根据所作的变更进行调整（例如，增加或缩短包含所有分支所产生反射的输入序列的长度，根据新的分支数量调整类别数量等）。然而，针对网络拓扑的每次变化都重新训练机器学习模型，是一个耗时且低效的过程（尤其要考虑收集包含所研究类别（不同故障分支可能性）的所有模式的新数据、标注新数据、重新训练模型以及部署模型等步骤所需的时间；特别是当分支数量增加时）。因此，最好采用一种通用机器学习方法，使其能够独立于特定网络拓扑进行训练，并且能够适用于任何网络架构，而无需在网络发生变化时重新训练。

3　提出的通用方法

图4展示了所提出的、基于机器学习的无源光网络系统故障分支识别方法的工作流程。为实现基于OTDR的无源光网络监测，在每根光纤末端放置一个光反射器。假设在部署或安装期间已知无源光网络系统的网络拓扑或架构。在安装无源光网络时，或网络拓扑发生变化时，会进行一次OTDR测量。记录的轨迹将作为参考（预先测量的参考OTDR轨迹）存储。无源光网络系统部署后，会定期采集OTDR轨迹并存储在集中式数据库中。随后，将记录的OTDR信号拆分为定长序列。请注意，每条OTDR轨迹拆分后的起始点均位于光分路器位置之后。综合考虑OTDR测量所用的脉冲宽度以及两条间距较近的相邻分支之间的长度差，序列长度的设定方式可使一个序列内出现两个以上反射的可能性极低。随后，将生成的定长序列输入机器学习模型，以预测待研究事件的类型（C_1：两个反射，其中第一个反射为故障反射；C_2：两个反射，其中第二个反射为故障反射；C_3：两个反射，且均为故障反射；C_5：一个故障反射）以及正常状态（C_0：两个正常反射；C_4：一个正常反射；C_6：无反射）。然后，将机器学习模型的输出与无源光网络系统正常运行时的预期输出（参考轨迹）进行比较。根据预测输出与预期输出之间的差异，即可识别故障分支。例如，如果某个序列的机器学习模型输出为C_2，且根据参考轨迹，该序列对应两条分支a和b，则可判定b为故障分支。请注意，该机器学习方法通过输入最多包含两个反射的短序列进行训练，以学习所有可能故障场景的模式，包括同时发生两个故障的情况。因此，机器学习模型能够识别多条故障分支的情况，例如光纤中断导致网络节点之间的多根光纤同时受损。因为在这种情况下，对记录的OTDR轨迹进行拆分，特别是对包含所有分支反射的部分进行拆分后，会生成短序列；根据分支之间的长度差，这些序列中最多包含两个故障，或仅包含一个故障，而机器学习模型能够准确识别这些情况。

例如，由于网络节点之间包含多根光纤的光纤发生中断，在这种情况下记录的OTDR轨迹经过拆分后，尤其是包含所有分支反射的部分，将生成短序列；根据分支之间的长度差，这些序列中最多包含两个故障，或仅包含一个故障，而机器学习模型能够准确识别这些情况。

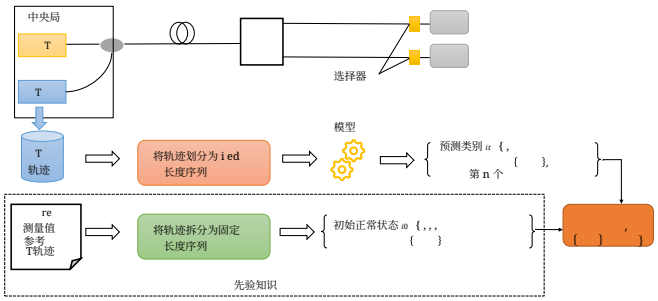


图4：利用OTDR监测数据识别无源光网络系统中故障分支的拟议方法流程图。参考轨迹表示在部署无源光网络时，或刚完成网络架构修改后进行的OTDR测量。

3.1 实验装置

为验证所提出的机器学习方法，按照图5所示搭建实验装置，记录包含故障分支和正常运行模式的OTDR轨迹，并在不同条件下进行测试。通过采用级联光分路器，搭建出分光比为1:128的真实无源光网络。每个分支末端均放置一个反射器，以产生反射。使用具有不同衰减设置（3 dB ... 8 dB）的光衰减器，调节部分分支的反射高度，从而模拟故障分支状态。所用衰减器为带有斜面物理接触（APC）连接器的固定衰减器，因为它们的性能优于带有物理接触（PC）连接器的可变光衰减器；后者可能会导致额外反射。请注意，衰减器的位置经过精心选择，以生成用于模拟不同待研究类别（$C_i \; \forall i \in \{1 \ldots 6\}$）的模式，如图6所示。不同衰减器的衰减值在3 dB至8 dB之间变化。为生成包含类别C_0模式的样本，采用正常状态下记录的OTDR信号（即使用该实验装置并移除所有衰减器后生成的OTDR轨迹）。两个相邻分支的长度差在2 m至6 m之间变化，以生成还包含彼此非常接近甚至重叠的反射脉冲的样本（具有挑战性的情况）。OTDR配置参数，即脉冲宽度、波长和采样时间，分别设为10 ns、1650 nm和2 ns。10 ns的脉冲宽度对应约2 m的空间分辨率。激光功率在0至16 dBm之间调整，以影响信噪比。

收集并平均多次OTDR测量后，记录一条OTDR轨迹所需的总测量时间在2 ms至2 s之间变化。需要注意的是，生成供机器学习模型分析的OTDR信号所需的总测量时间非常短，这有助于加快无源光网络系统中故障的诊断。请注意，即使输入的是短时间内记录的含噪OTDR信号，机器学习模型仍能良好运行，因为它能够从含噪数据中学习并提取知识，而噪声甚至可能有助于

避免机器学习模型过拟合。这与 传统的基于OTDR的事件分析方法形成对比，后者需要对OTDR信号进行大量平均以降低噪声，从而获得良好表现。图7展示了不同衰减设置下各分支的正常和故障行为。

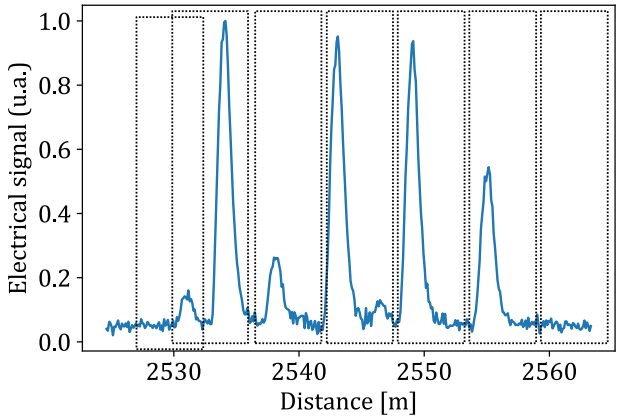


图6：生成的光时域反射仪信号示例，其中包含通用机器学习模型需要识别的各类别模式。每个虚线矩形（即从 OTDR轨迹中提取的长度为 60 的序列）代表一个类别 C_i, $i \in \{ \dots \}$ 的模式。通过改变分割索引的起始位置，第一个和第二个虚线矩形呈现出两种不同的模式。第一个虚线矩形包含 C_5 的模式，其中只有一个故障反射。第二个矩形呈现 C_1 的模式，其中第一个反射是故障反射。

3.2 数据预处理

生成的光时域反射仪信号，尤其是1:8分光器之后的轨迹部分，被分割成长度为60的短序列。请注意，我们将序列长度设置为60，是为了使每个序列中最多包含两次反射；因为在我们的研究中，对于所选的脉冲宽度，单次反射可以用长度为30的序列表示。尽管衰减器的设置有所变化，但所考虑的3 dB至8 dB衰减值范围较窄，由此产生的数据不够多样，无法构建稳健可靠的机器学习模型，并可能导致机器学习模型过拟合。更改衰减器值是一个耗时且需要手动操作的过程。因此，序列中被视为故障的分支的反射高度会被人为降低。 反射高度的降低比例所考虑的故障分支会在$\frac{1}{50}$到$\frac{1}{2}$之间随机变化，以模拟由于衰减设置范围较宽而导致的反射潜在变化，包括衰减值极高、可能使反射

完全消失（在高分光比或馈线光纤较长的情况下，远端节点之后发生光纤断点时，也会观察到类似效果）。 对于由序列中的两个反射构成的故障场景，在保持另一反射处于正常运行条件下的状态不变的同时，人为改变由故障分支引起的反射高度，或者改变由两个反射产生的两个

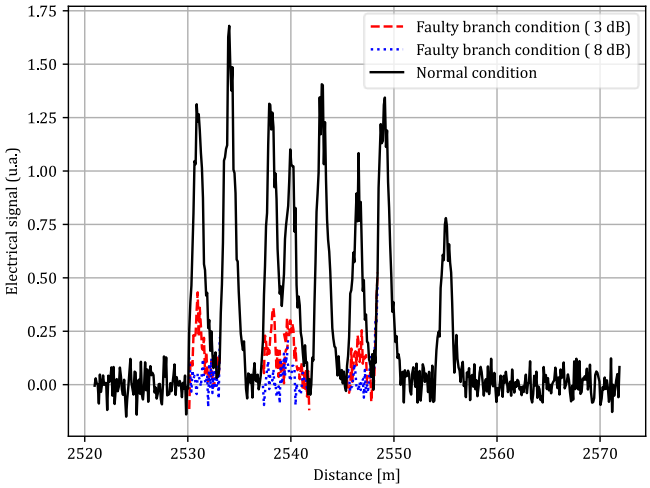


图7：不同衰减设置下的正常分支和故障分支模式。

同时、任意且以不同方式降低故障分支，以生成多样化的数据，其中包含不同的反射级别，这些级别可能由各种衰减设置造成，包括衰减级别极高、导致一个或两个反射完全消失的情况。针对不同故障情形，生成了包含各种模式的多样化序列。正常样本取自未添加任何衰减器时进行的实验装置。鉴于生成的数据存在一定噪声，对每个序列添加加性白高斯噪声，使峰值信噪比（PSNR）介于 5 至 30 dB 之间（噪声作为一种正则化手段，可增强模型的鲁棒性）。PSNR 的定义如下：

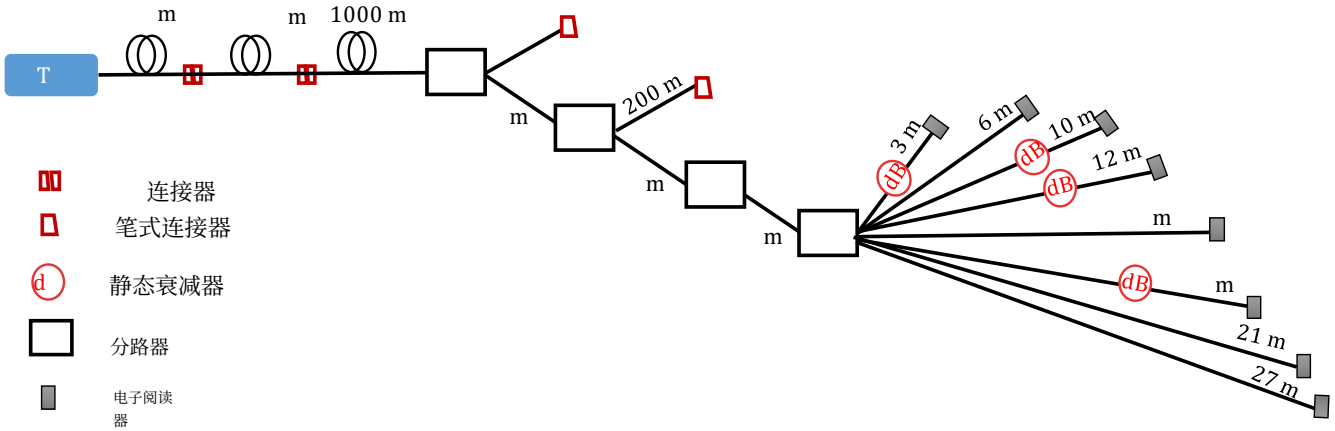


图 5：在无源光网络中生成故障分支数据的实验装置。

$$PSNR = \frac{a}{\sigma_{noise}} \quad (7)$$

其中， $\boldsymbol{a}$ 为反射高度， $\boldsymbol{\sigma_{noise}}$ 为噪声的标准差。

图 8 总结了用于生成多样化噪声 $\boldsymbol{C_1}$ 样本示例的流程：给定正常（无故障）信号，第一个分支（分支 1）出现故障。

总共构建了一个包含**167,020** 个样本的数据集（每个待研究类别包含 ~23860 个示例），并对其进行归一化，然后划分为训练集（60%）、验证集（20%）和测试数据集（20%）。

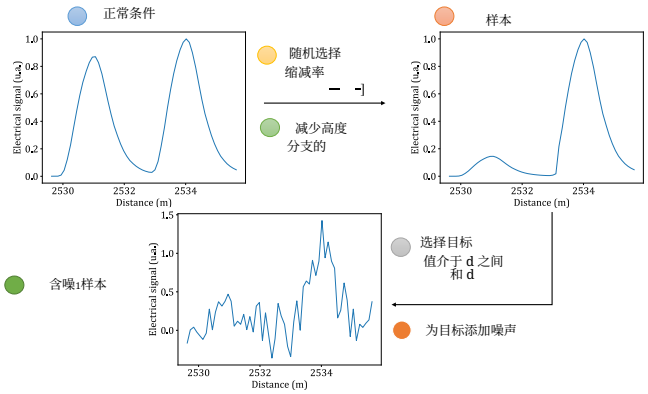


图 8：给定正常状态信号，生成多样化噪声 C_1 样本的流程（此处以分支 1 为例）。

3.3 模型架构

所提出的机器学习模型的架构如图 9 所示。该机器学习方法以长度为 60 的 OTDR轨迹序列作为输入，并输出类别类型（ $C_i \ \forall i \in \{ \dots \}$）。输入被送入一个由 32 个单元组成的长短期记忆网络层，以提取相关特征 $[h_1, h_2 \dots h_{32}]$。随后，提取的特征被传递至包含 16 个神经元的全连接层，再进入一个搭配 Softmax激活函数的输出层，其可以表示如下：

$$\text{Softmax}\,(\boldsymbol{z})_i = \frac{\exp(z_i)}{\sum_{j=1}^{k} \exp\,(z_j)} \quad \text{for } i = 1 \dots k \quad (8)$$

其中 $\boldsymbol{z} = (z_1 \dots z_k) \in \mathrm{R}^k$ 表示输入向量， z_i 表示输入$\boldsymbol{z}$的一个元素。
Softmax 函数输出不同类别（模型的不同结果）的概率分布。

模型通过采用自适应矩估计（Adam）优化器，最小化分类交叉熵（即代价函数或损失函数）进行训练。

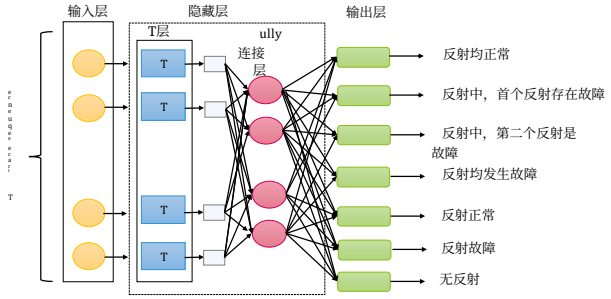


图 9：所提出的基于机器学习的通用事件识别方法。

3.4 性能评估

首先，使用未见过的含噪测试数据集评估机器学习方法的性能（该数据集是按照第 3.2 小节所述原则，从生成的数据中随机抽取的）。图 10 所示的混淆矩阵表明，机器学习方法能够较好地对不同类别进行分类
平均诊断准确率为 93%。机器学习模型可能会偶尔将类别 $\boldsymbol{C_1}$ 和 $\boldsymbol{C_2}$ （两个反射，其中第一个或第二个反射发生故障）误分类为类别 $\boldsymbol{C_4}$ （一个正常反射），这是由于这些类别相似，尤其是在故障反射（两个反射之一）逐渐消失或完全消失的情况下。类别 $\boldsymbol{C_5}$ 和 $\boldsymbol{C_3}$ 可能会被偶尔误分类，具体来说，在类别 $\boldsymbol{C_3}$ 的情形中，一个故障反射完全消失（在这种情况下， $\boldsymbol{C_3}$ 的模式与 $\boldsymbol{C_5}$非常相似）。在反射高度大幅降低的情况下，类别 $\boldsymbol{C_6}$ 可能会被误分类为 $\boldsymbol{C_5}$ 。模型产生的误分类率极低，这可以解释为：在低信噪比条件下，由于噪声主导序列，噪声可能会淹没信号，使所研究类别的模式看起来相似，因此难以区分不同类别。

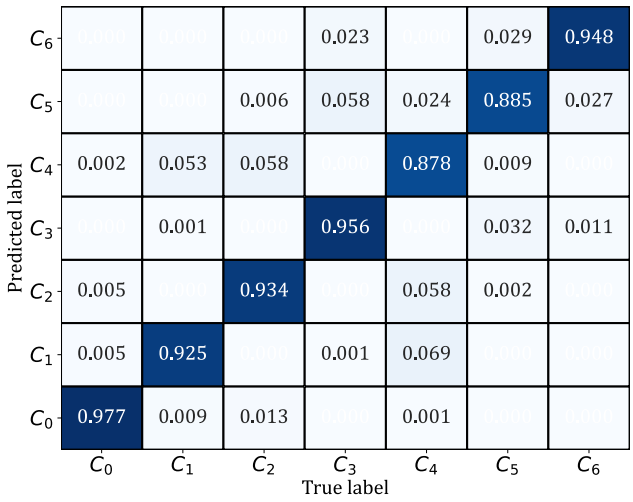


图 10：通用机器学习方法在含噪生成测试数据集上得到的混淆矩阵。

随后，使用不含任何人工噪声的实验数据评估机器学习模型的性能。

对用于生成通用模型训练数据的实验装置进行改造，以生成实验测试数据集并评估模型的鲁棒性。衰减器的位置有所调整：衰减器现在分别位于 6 m、10 m、12 m、18 m、21 m 和 27 m 的分支上。对于 10 m 和 12 m 的分支，所设定的衰减设置与训练机器学习模型时使用的设置不同。移除了 15 m 分支末端的终端匹配器，以模拟光纤断裂故障。

图 11 显示了使用调整后的实验装置生成的实验数据对机器学习方法进行测试后得到的混淆矩阵。由于生成的数据几乎没有噪声（与输入机器学习模型进行训练的数据相比，噪声更少），该机器学习方法能够以极高的准确率对不同类别进行分类，同时误分类率极低。

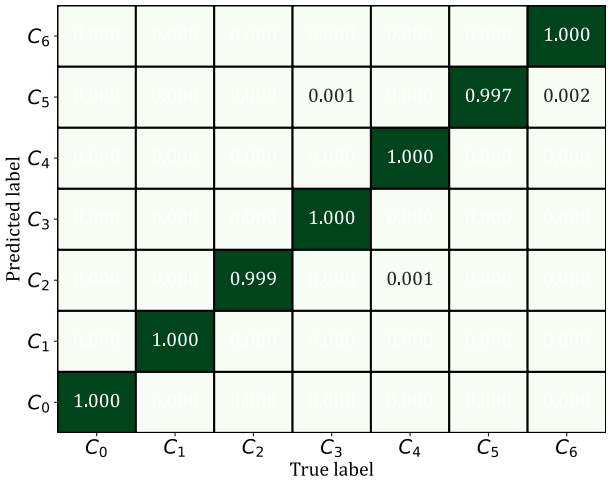


图 11：通用机器学习方法在未添加任何人工噪声的情况下，对未见过的实验数据集进行分类所得到的混淆矩阵。

3.5 故障分支识别

结合机器学习模型的输出结果以及对 PON 系统拓扑（现有分支）的了解，可用于识别故障分支。图 12 展示了一个光时域反射仪信号示例（位于 1:8分光器之后的一段 OTDR轨迹），该信号使用上一小节所述的调整后实验装置记录。首先对光时域反射仪信号进行归一化处理，然后将其拆分为长度为 60 的序列。将所得序列输入机器学习模型，以预测每个序列所属的类别。随后，根据机器学习模型对每个序列的输出结果，以及该序列中存在的分支标识符信息，即可识别故障分支，如表 1 所示。

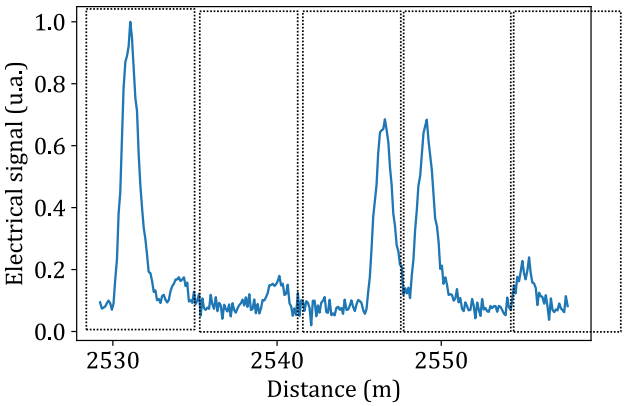


图 12：测试光时域反射仪信号被拆分为多个序列的示例，其中每个序列对应的机器学习模型输出以红色标出。

表 1 根据机器学习模型的输出和分支标识符信息识别故障分支

机器学习模型输出 / 模型	分支标识符	故障分支
C_2	1,2	2
C_3	3,4	3,4
C_3	5,6	5,6
C_5	7	7
C_5	8	8

3.6 所提出的机器学习模型的优化

在评估不同规模的网络架构后，选择性能最佳且复杂度适中的网络架构。研究了参数——LSTM隐藏层的数量和全连接层中神经元的数量——对所提出方法诊断准确率的影响。表2显示，单个LSTM隐藏层是以最低复杂度实现最高准确率的最佳选择。表3显示，包含16个神经元的全连接层是兼顾准确率与复杂度的理想选择。

表2 EFFECT 对LSTM隐藏层数量对PROPOSED 模型所达到的准确率和训练时间的影响。

数量 LSTM层	准确率（%）训练时间（秒）
	93 2522.9
2	92.7 4310
3	92.5 7844.6
4	92.4 14972

表3 全连接层中神经元数量对 PROPOSED 模型所达到的准确率和训练时间的影响。

神经元数量 神经元	准确率（%）训练时间（秒）
8	92.4 2413
16	93 2522.9
32	92.8 2990
64	92.9 7537.2

3.7 所提方法与其他机器学习算法的比较

我们将所提出的机器学习模型与其他机器学习分类器进行比较，包括决策树（DT）、朴素贝叶斯（NB）、随机森林（RF）、梯度提升（GB）和人工神经网络（ANN）。采用诊断准确率作为评估性能指标，使用未见过的测试数据集评估不同的机器学习算法。图13所示结果表明，我们的方法优于其他机器学习分类器，准确率最高，能够提供一个

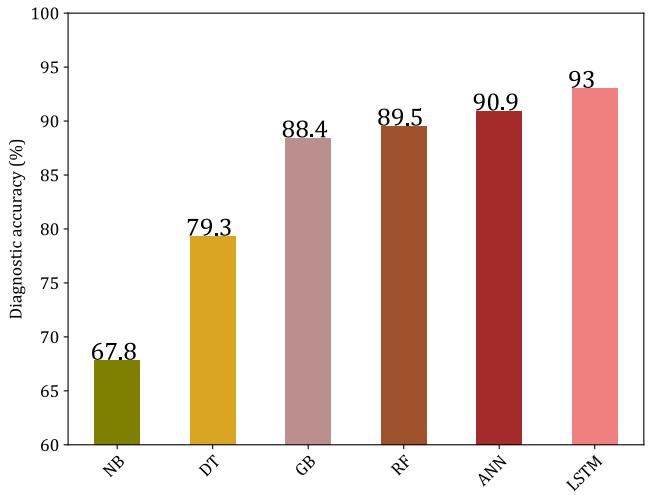
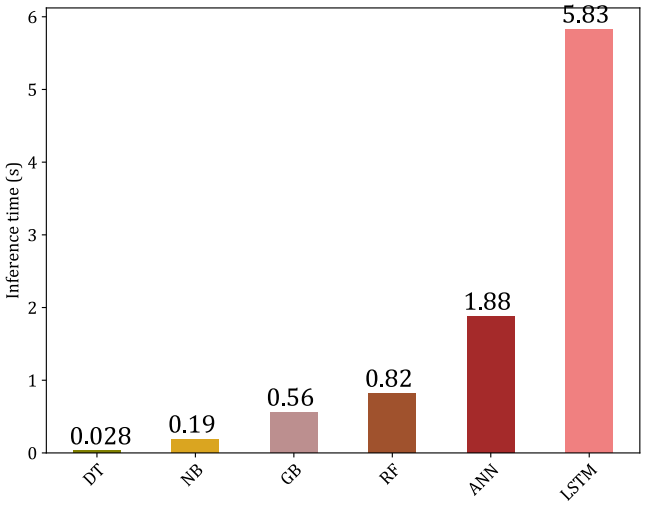


图 13：不同机器学习算法在诊断准确率方面的比较。

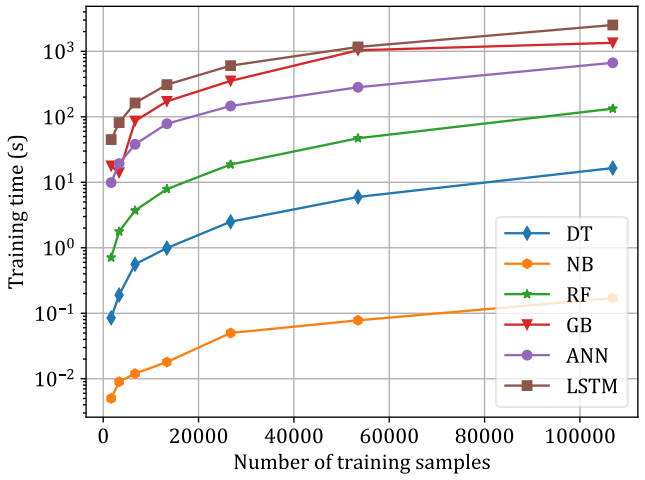
准确率提高了 2.1% 以上，因为 LSTM 在处理序列数据和捕捉长期依赖关系方面表现更好。

图 14(a) 展示了所提出的模型与其他机器学习方法在推理时间上的比较结果（处理包含 33,404 个样本的新测试数据并进行预测所需的时间）。由于 LSTM 网络的复杂度较高且具有并行性，所提出的方法在测试（推理）时耗时更长，而浅层机器学习技术 DT、NB、RF 和 GB 的测试耗时则显著更短。

图14(b)展示了在不同输入规模（训练样本数量不同）下训练各机器学习模型所需的计算训练时间。结果证实，所提出的方法速度较慢，因为其学习阶段需要更多时间；而其他较简单的机器学习模型，尤其是朴素贝叶斯（NB）和决策树（DT），速度更快。



(a)



(b)

图 14：不同机器学习模型的计算复杂度比较：(a) 不同模型的推理时间；(b) 各机器学习模型随输入规模（训练样本数量）变化的训练时长。

3.8 对不同网络的适用性

为评估机器学习方法的鲁棒性及其对不同无源光网络系统的适用性，搭建了图15所示的实验装置。使用一个1:4分光器重现无源光网络系统。采用可变光衰减器（VOA）改变反射的衰减，从而模拟故障分支条件。针对第三个分支（7.3 m）和第四个分支（9.6 m），在不同的VOA设置下记录了一些光时域反射仪轨迹。光时域反射仪参数（即脉冲宽度、激光功率、光时域反射仪测量的平均次数、采样时间等）与生成训练数据时所用的参数相同。利用生成的轨迹测试机器学习模型的性能，具体而言，测试的是分光器之后的光时域反射仪信号部分。图16展示了机器学习方法对不同VOA设置下第三个和第四个分支所生成的光时域反射仪信号示例的输出。如图所示，机器学习模型能够正确识别每个序列的类别，因此可以正确识别故障分支，而且无需使用从新无源光网络系统获得的新数据重新训练机器学习方法。

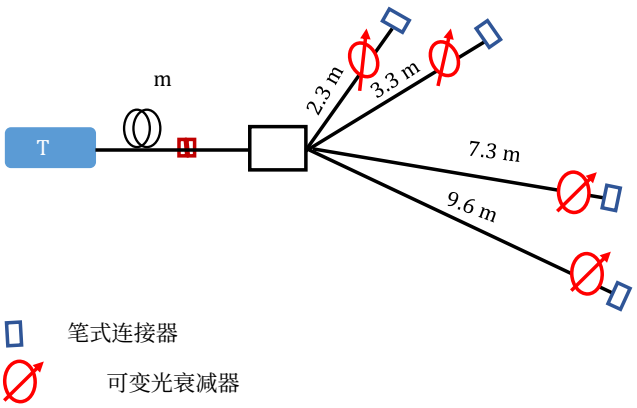
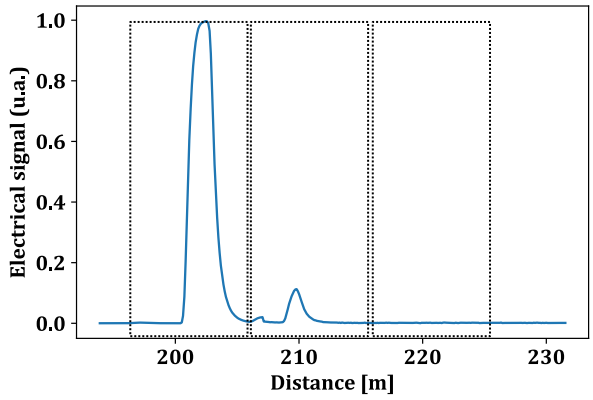
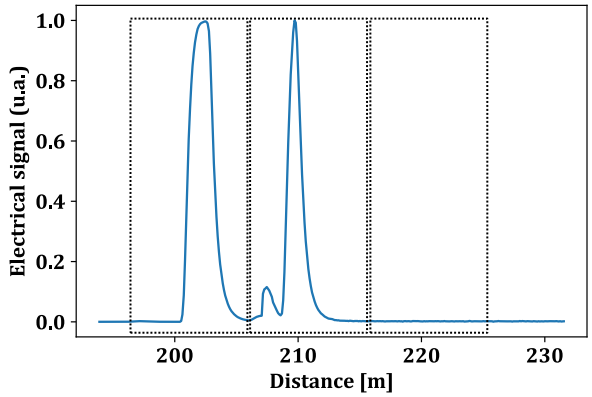


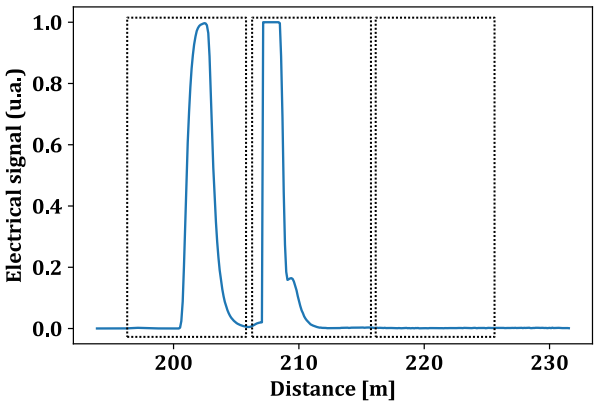
图15：用于测试机器学习方法鲁棒性的另一种PON系统实验装置。



(a)



(b)



(c)

图16：不同可变光衰减器设置下记录的光时域反射仪信号示例：(a) 第三和第四分支分别设置为7 dB和4 dB衰减；(b) 第三分支设置为10 dB衰减；(c) 第四分支设置为4 dB衰减。

4. 结论

本文提出了一种基于机器学习的方法，用于识别无源光网络系统中的故障分支。该方法的训练不依赖于特定的网络拓扑，并且无需因网络

变化。所提出方法的有效性已通过源自无源光网络系统的实验数据得到验证。所介绍的机器学习方法有助于改善无源光网络的监测，尤其是在反射脉冲重叠这一具有挑战性的场景中。请注意，所提出的方法使用包含所研究类别完整模式的 OTDR 轨迹序列进行训练和测试。然而，在实际应用中，机器学习方法应应用于任意分段的 OTDR 序列，这些序列可能包含事件的部分模式。因此，为了增强机器学习算法对这类可能出现的真实场景的适应能力，我们计划未来使用包含各研究类别部分模式和完整模式的数据集对其进行训练。未来，我们还计划扩展机器学习模型的架构，使其能够同时输出故障分支标识符和故障位置。此外，我们计划进一步增强

所提出机器学习方法的泛化能力，以应对具有等间距分支的无源光网络系统：通过使用包含等长分支场景的数据重新训练模型，或采用迁移学习或持续学习方法。

致谢。 感谢 Bernd Lebenwein 先生在实验装置搭建方面给予的支持。本研究在 CELTIC-NEXT 项目 AI-NET-PROTECT（项目编号 C2019/3-4）框架下开展，并部分获得德国联邦教育与研究部资助（FKZ16KIS1279K）。

参考文献

[1] H. Abbas、Huda 等人，《无源光网络的下一代：综述》，《网络与计算机应用杂志》，2016年。[2] K. Yuksel 等人，《无源光网络（PON）中的光层监测：综述》，2008年透明光网络国际会议十周年会议，2008年，第92-98页。[3] M. M. Rad 等人，《利用光学编码进行光纤故障 PON 监测：用户地理分布的影响》，载于《IEEE通信汇刊》，第58卷，第4期，第1172-1181页，2010年4月。[4] M. A. Esmail 等人，《时分复用无源光网络的物理层监测技术：综述》，载于《IEEE通信调查与教程》，第15卷，第2期，第943-958页，2013年第二季度。[5] B. sante 等人，《光纤切断的影响及其经济损失的科学分析》，. -51. 10.9790/2834- 1603023451，2021年。[6] M.- . ad 等人，《无源光网络监测：挑战与要求》，EEE ommunications Magazine，第49卷，第2期，第s45-S52页（2011年）。[7] M. Thollabandi 等人，《基于自注入锁定反射式半导体光放大器直接调制的可调谐 T T T，用于波分复用无源光网络线路监测》，roc. E，布鲁塞尔，比利时，2008年9月。[8] K. zawa 等人，《使用可调谐 T 对 s 进行业务中单线路监测的现场试验》，第届国际光纤传感器会议，威尼斯，2000年，第880-883页。[9] Villafani Caballero, Diego Rodrigo 等人，《用于波分复用监测的可调谐 T 测量》。SBMO/IEEE MTT-S 国际微波与光电子学会议（IMOC）（2013年）：1-5。[10] K. Yuksel 等人，《无源光网络（PON）中的光层监测：综述》，2008年透明光网络国际会议十周年会议，2008年，第92-98页。[11] Caviglia、rancesco 和 aleria i iase。《s 中的光维护》。第届欧洲光通信会议。ECOC '98（IEEE目录号98TH8398）1（1998年）：621-625，第1卷。[12] F. Musumeci 等人，《光网络故障管理中的机器学习教程》，载于《光波技术杂志》，第37卷，第16期，第4125-4139页，2019年8月15日。[13] L. Velasco 和 D. Rafique，《基于机器学习的故障管理》，2019年光纤通信会议暨展览会（OFC），2019年，第1-3页。[14] A. P. Vela 等人，《弹性光网络中的误码率劣化检测与故障识别》，载于《光波技术杂志》，第35卷，第21期，第4595-4604页，2017年11月1日。

《光波技术杂志》，第35卷，第21期，第4595-4604页，2017年11月1日。[15] T. Panayiotou 等人，《利用统计机器学习解决光网络中的故障定位问题》，载于 Journal of Optical Communications and Networking，第10卷，第3期，第162-173页，2018年3月。[16] K. Abdelli 等人，《光纤监测中的基于机器学习的异常检测》，载于 Journal of Optical Communications and Networking，第14卷，第5期，第365-375页，2022年5月，doi: 10.1364/JOCN.451289。[17] A. Usman 等人，《通过整合机器学习和光纤传感器对无源光网络进行故障监测》。International Journal of Communication Systems。第35卷，2022年。[18] K. Abdelli 等人，《基于门控循环单元的自编码器，用于无源光网络中的光链路故障诊断》，ECOC 2021。[19] Parisi、German gnacio 等人。《利用神经网络进行持续终身学习综述》。神经网络《International Neural Network Society 官方期刊》第113卷（2018年）：54-71页。. [20] S. ochreiter 等人，《长短-期记忆》，神经计算，9(8):1735–1780（1997年）。