

文章编号: 1006-3080(2025)03-0353-10

DOI: 10.14135/j.cnki.1006-3080.20240712001

基于数据驱动与深度神经网络的智能配电网拓扑辨识

杜文凯, 陈心怡, 薛 栋

(华东理工大学能源化工过程智能制造教育部重点实验室, 上海 200237)

摘要: 网架结构的拓扑辨识是配电系统优化与控制的基础。随着风能和太阳能等可再生能源发电的高比例接入, 配电网的拓扑结构变得更加复杂且变化频繁, 显著增加了拓扑辨识的难度。为了提高拓扑辨识的准确率, 本文结合配电网的结构和运行特点, 提出了一种基于自组织映射(SOM)和卷积神经网络(CNN)深度学习框架的配电网拓扑辨识方法。考虑到配电网数据的高维特性, 该方法首先利用主成分分析(PCA)对高维电压幅值和有功功率数据进行降维, 进而使用SOM提取数据特征, 将其转换为二维特征图, 并通过CNN学习输入特征与拓扑标签之间的映射关系, 从而实现配电网拓扑结构的精准辨识。通过在33、69、123节点配电网算例上进行仿真实验, 验证了所提方法的有效性, 并且相较于其他方法, 该方法在辨识准确率和鲁棒性等性能上具有明显优势。

关键词: 配电网系统; 拓扑结构辨识; 数据驱动; 深度学习; 卷积神经网络

中图分类号: TP309

文献标志码: A

配电网作为电力网络中能源传输和分配的主要载体之一, 其安全稳定运行是保证供电可靠性的基础^[1]。由于分布式发电的配电网能节省成本、降低能耗, 并具有电力系统可靠性和灵活性等优势, 已成为21世纪电力工业建设的关键推动力^[2]。同时, 风电和光伏发电等可再生能源的快速发展和广泛应用极大地改变了传统配电网的规划, 而这些能源的不确定性和波动性也给配电网的优化和控制带来了重大的挑战^[3-4]。

了解拓扑结构是配电网系统实时分析、优化和控制的基础。因此, 在配电网系统的研究中, 准确进行拓扑辨识具有重要意义。输电网系统通常配备齐全的传感设备, 拥有完整的测量数据和较强的可观测性, 从而能够实现拓扑结构的精准辨识, 因此在输电网中, 拓扑辨识主要解决的问题是纠正现有拓扑中的测量错误^[5]。与输电网相比, 配电网具有复杂的结构, 其中涵盖了大量的馈线和节点, 设备种类繁多且结构复杂。配电网由于经济和技术的限制, 导

致内部监测设备和通信链路数量极为有限, 这限制了系统对运行状态的全面监测能力, 难以实时获取线路和开关状态的详细信息。此外, 配电网大规模接入了分布式可再生能源, 如风力发电和光伏发电, 这些能源的接入使得配电网的拓扑结构变化更为频繁和复杂, 从而显著增加了拓扑辨识的复杂性。这些特性共同作用, 进一步提高了配电网拓扑辨识的挑战和难度。因此, 提升配电网拓扑辨识的准确性和处理速度显得尤为迫切和重要。

配电网拓扑结构的辨识方法主要分为基于优化和数据驱动的方法。基于优化的方法依赖潮流分析和状态估计进行拓扑辨识。Gao等^[6]提出了一种基于可选匹配环功率的配电网拓扑辨识方法, 通过在每种可能的拓扑结构中运行状态估计器, 并选择匹配度最高的估计结果作为正确的拓扑结构。Arghandeh等^[7]通过潮流方程计算构建可能的拓扑结构库, 将测量数据与这些计算结果进行比较, 以确定误差最小的拓扑结构。Li等^[8]提出了一种基于潮流匹配和生

收稿日期: 2024-07-12

基金项目: 国家自然科学基金(62173147); 上海市浦江人才计划(20PJ1403000)

作者简介: 杜文凯(1998—), 男, 河南周口人, 硕士生, 主要研究方向为智能电网的数据驱动方法。E-mail: Y30210955@mail.ecust.edu.cn

通信联系人: 薛 栋, E-mail: dong.xue@ecust.edu.cn

引用本文: 杜文凯, 陈心怡, 薛 栋. 基于数据驱动与深度神经网络的智能配电网拓扑辨识[J]. 华东理工大学学报(自然科学版), 2025, 51(3): 353-362.

Citation: DU Wenkai, CHEN Xinyi, XUE Dong. Topology Identification for Smart Distribution Networks Based on Data-Driven and Deep Neural Networks[J]. Journal of East China University of Science and Technology, 2025, 51(3): 353-362.

成树算法的配电网拓扑辨识方法。Tian 等^[9]提出了一种基于混合整数二次规划的配电网拓扑辨识模型,适用于辐射状配电网运行的场景。Poudel 等^[10]提出了一种基于图论的状态估计模型选择方法,该方法利用图论枚举候选拓扑,并通过状态估计进行模型选择,从而完成配电网的拓扑辨识。基于潮流分析和状态估计的配电网拓扑辨识方法往往面临计算资源消耗量大和耗时长的问题。随着分布式能源比例的增加,以及大型配电网中海量数据处理需求的递增,这个问题变得格外严峻。同时,这也进一步加大了此类方法在实际配电系统中应用的复杂性。

数据驱动的配电网拓扑辨识方法不依赖于网络配置和线路参数等系统结构的先验知识,而主要依赖于测量数据。García 等^[11]提出了一种基于贝叶斯推理的数据驱动方法,用于配电网的相位拓扑辨识。Luan 等^[12]利用智能电表的电压时间序列测量曲线进行相关性分析,以验证相邻节点的连通性,并据此修正地理信息系统中配电网拓扑记录的错误。Yu 等^[13]提出了一种基于最大似然估计框架的变量误差模型,以促进配电网线路拓扑和参数的联合估计。Liao 等^[14]采用概率图模型来捕捉相邻节点电压之间的统计依赖关系,从而实现了配电网的拓扑结构辨识。Gavraro 等^[15]将配电网拓扑辨识问题建模为一个统计推断问题,并将观测到的电压值视为底层电网拓扑结构的函数,并由此提出了一种利用最大似然估计和最大后验概率方法的拓扑辨识方案。Kekatos 等^[16]设计了一种基于能源价格矩阵分解的数据驱动配电网拓扑辨识方法。Flynn 等^[17]提出了一种基于改进图学习算法的配电网拓扑估计方法。该方法首先将电压灵敏度矩阵转换为相应的距离矩阵,然后通过递归分组图学习算法从距离矩阵中估计出完整的网络拓扑。上述方法可以通过大量数据训练来完成配电网的拓扑辨识任务,但是其处理大量数据的能力和速度有限,这限制了拓扑辨识的准确性和效率。

近年来,以深度学习为代表的人工智能技术得到了迅猛发展,并在包括电力系统在内的多个领域中得到了广泛应用^[18]。深度学习在配电网拓扑辨识方面能够通过其强大的特征学习能力、适应大规模数据的能力以及优秀的泛化性能,有效提升拓扑辨识的准确性和效率,并已被广泛应用于配电网拓扑辨识。Li 等^[19]提出了一种基于无监督机器学习算法和图论的配电系统拓扑辨识方法。该方法仅利用智能电表的测量信息即可完成拓扑辨识,并能够生成简单的拓扑结构图。Jafarian 等^[20]为分布式能源资

源管理系统开发了一种基于深度神经网络的配电网拓扑辨识方法,该方法仅依赖于系统可用的测量值,有效减少了对大量数据的依赖问题。Ni 等^[21]提出了一种基于一维卷积神经网络的有源配电网拓扑辨识方法,用以应对配电网中测量设备不足的问题。

然而,面对可再生能源发电影响下配电网拓扑结构种类繁多、变化频繁以及辨识难度增加的问题,大多数现有的拓扑识别算法的辨识准确率无法得到保证。为了解决这些问题,本文提出了一种基于自组织映射(SOM)和卷积神经网络(CNN)深度学习框架的数据驱动方法,以实现复杂配电网拓扑结构的准确辨识。该方法首先通过主成分分析(PCA)对配电网的高维电压幅值和有功功率数据进行降维,在保留主要特征的同时,有效降低了数据的复杂性和计算负担。进而,利用 SOM 进一步提取降维后的数据特征,将其转化为二维特征图,从而充分捕捉了数据中各特征之间的关联性。最后,使用 CNN 对生成的二维特征图进行训练和学习,建立输入特征与拓扑标签之间的映射关系,从而实现对配电网拓扑结构的准确辨识。SOM-CNN 方法巧妙结合了 SOM 的特征提取功能和 CNN 的特征学习功能,能够高效处理配电网的高维数据,同时显著提高拓扑辨识任务的准确率。另外,该方法在处理具有高噪声和缺失值的数据时,也能保持高准确性,从而为配电网拓扑辨识提供了一种准确且鲁棒的数据驱动方案。

1 拓扑辨识模型和原理

1.1 配电网模型

本文使用无向图 $\mathcal{G} = (\mathcal{N}, \epsilon)$ 来表示一个包含 N 个节点的配电网系统,其中 $\mathcal{N} = \{1, 2, \dots, N\}$, 而 $\epsilon \in \mathcal{N} \times \mathcal{N}$ 表示节点之间的边集合。

潮流方程是配电网稳定运行和潮流计算的基础,描述了配电网电压(电压幅值和电压相角)和功率(有功功率和无功功率)的映射关系,具体如下:

$$P_i := P_i^g - P_i^d = V_i \sum_{j \in \mathcal{N}} V_j (G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij})$$

$$Q_i := Q_i^g - Q_i^d = V_i \sum_{j \in \mathcal{N}} V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij})$$

其中, P_i 和 Q_i 分别是节点 i 处的有功和无功注入功率,由发电功率减去负载功率确定;上标 g、d 分别表示发电和负载。 V_i 和 θ_i 分别表示节点电压的幅值和相角,并用 $\theta_{ij} = \theta_i - \theta_j$ 。 G_{ij} 和 B_{ij} 分别表示节点 i 和 j 之间的电导和电纳。

对于给定的网络拓扑结构,当 4 组变量中至少

有 $2N$ 个变量已知时, 通过求解潮流方程即可确定配电网的运行状态。不同拓扑结构的配电网通常具有不同的导纳矩阵, 这将导致系统变量具有不同的分布特性, 这种特性是数据驱动拓扑辨识方法的基础。结合配电网的实际运行特性和测量配置, 本文选择电压幅值向量 V 和注入有功功率 P 作为系统的原始特征集。

1.2 深度神经网络

1.2.1 自组织映射 SOM 是一种由 Kohonen^[22]提出的无监督人工神经网络 (ANN), 它能够进行数据降维、可视化、特征提取和聚类, 已在众多任务中得到广泛应用, 并被证明是有效的。典型的 SOM 网络如图 1 所示, 由输入层和竞争层 (也称为输出层) 组成^[23], 可以观察到, 所有输入向量都通过权重向量完全连接到每个输出神经元。

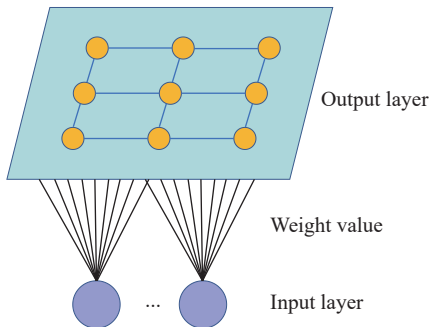


图 1 SOM 结构示意图
Fig. 1 Structure diagram of SOM

SOM 的工作原理是基于竞争学习和拓扑保持的机制, 通过在高维输入空间中形成低维拓扑结构的映射, 实现数据的非线性降维和特征提取, 使得相似的输入向量在映射空间中聚类在相邻区域。SOM 的详细训练过程见算法 1。

算法 1 的步骤如下: (1) 初始化竞争层神经元的权重。(2) 随机选择一个输入向量 d 。(3) 计算输入向量 d 与每个神经元权重向量之间的欧几里德距离, 并将距离最小的神经元确定为最佳匹配神经元

u 。(4) 对于每个神经元 v , 根据以下公式更新其权重向量:

$$w_v(s+1) = w_v(s) + \theta(u, v, s)\alpha(s)(d - w_v(s)) \quad (1)$$

其中, s 是当前迭代次数; v 是图中神经元的索引; u 是最佳匹配神经元的索引; w_v 是神经元 v 的权重向量; θ 是邻域函数; $\alpha(s)$ 是学习率, 随着迭代次数 s 递减。(5) 重复步骤 (2)~(4), 直到达到设定的迭代次数。

1.2.2 卷积神经网络 CNN 是一种专门用于处理视觉数据 (如二维图像) 的 ANN。与传统的全连接层 ANN 不同, CNN 通常由 3 种主要层组成: 卷积层、池化层和全连接层。通过交替进行卷积和池化操作, CNN 能够高效地从原始图像数据中提取特征, 并在全连接层中执行分类任务。CNN 能够捕捉图像中的空间局部性特征, 从而提升图像处理和效率的效率和准确性。图 2 展示了包含两个卷积层和两个池化层的 CNN 架构示意图。

卷积层通过卷积操作从输入图像中提取特征。在每个卷积层中, 使用多个卷积核, 每个输入图像与每个卷积核进行卷积, 从而生成与卷积核数量相等的输出图像 (也称为特征图)。卷积层的可学习参数包括卷积核内的权重和偏置。这些权重和偏置是共享的, 使得卷积核的大小远小于输入图像的大小。这种设计不仅确保卷积层能够提取重要特征, 还显著减少了可学习参数的数量, 从而有助于构建更深的网络。

池化层在进一步减小图像尺寸方面起着关键作用。通过使用固定大小的池化窗口, 池化层将一组神经元聚合为输出图像中的一个单一神经元, 而无需进行训练。常用的池化技术包括最大池化和平均池化。在最大池化中, 保留的是输出图像中每个局部神经元群的最大值; 而在平均池化中, 计算并保留的是该局部神经元群的平均值。

通过卷积和池化操作, CNN 能够从原始输入数据中提取高级特征。然后, 这些特征传递给全连接

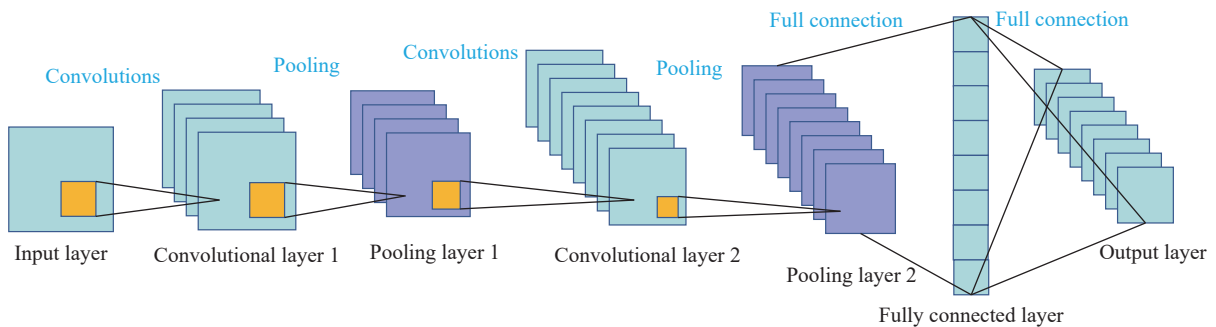


图 2 CNN 结构示意图

Fig. 2 Structure diagram of CNN

层,用于最终的分类、回归或其他预测任务。在全连接层中,每个神经元与前一层的所有神经元相连接,从而进行综合处理和决策。

通过上述操作,CNN通过组合卷积层、池化层和全连接层,有效地从输入图像中提取和处理特征,最终实现精确的分类和预测。

2 拓扑辨识的深度学习框架

2.1 数据标准化

为了确保初始特征集 F 中的所有特征对分析的贡献相同,本文首先采用最小-最大标准化方法将原始特征数据转换为区间 $[0,1]$ 内的值。最小-最大标准化的转换函数定义如下:

$$\hat{f}_i = \frac{f_i - f_i^{\min}}{f_i - f_i^{\max}} \quad (2)$$

其中, f_i 为特征 i 的初始值; f_i^{\min} 和 f_i^{\max} 分别为特征 i 的最小值和最大值; \hat{f}_i 为特征 i 的标准值。

2.2 主成分分析

PCA 是一种基于主成分概念的数据降维技术。它通过线性组合原始特征,生成能够保留最大方差的新特征,从而实现数据的降维和简化。

在拓扑辨识中,同时使用功率注入和电压幅值使特征数量是配电网系统节点数量的两倍。这样,在处理大规模网络时,后续 SOM 和 CNN 部分的计算时间会显著增加,从而减缓整个辨识过程。此外,配电网中一些相邻节点在所有拓扑结构中表现出相似的特征,这进一步强调了特征合并的必要性。引入 PCA 不仅能有效减少特征数量,简化后续训练过程,还可以提高整个深度学习框架的效率。具体而言,PCA 通过选择数据集中较大特征值所对应的特征向量,构建了主成分矩阵。随后,通过主成分矩阵对原始特征进行线性映射,实现了特征合并和数据降维,从而提升模型训练效率。同时,为了在不影响算法准确性的前提下提升效率,通过实验验证了主成分的选择,以在准确性和效率之间实现合理的平衡。

2.3 基于 SOM 的特征提取

在所提出的辨识框架中,SOM 的主要功能是将包含电压和有功功率信息的高维输入特征映射到二维空间。这个过程不仅用于从高维数据中提取特征,还作为后续 CNN 处理的预处理步骤,将配电网输入数据转换为符合 CNN 处理要求的二维形式。通常情况下,SOM 的期望输出是输入数据在输出映射中最佳匹配神经元的索引^[24]。然而,在本框架中,

关键技术是利用 SOM 网络权重和输入数据的内积,有效地将输入数据提取到二维空间,这种处理不仅有助于特征提取,还为后续的卷积神经网络处理奠定了基础^[25]。

在一个完全训练的 SOM 网络中,相邻神经元通常具有相似的权重,在本文提出的 SOM-CNN 框架中,这种相似权重特征会导致信息分布不均,从而出现某些区域的信息冗余,而其他区域则可能缺乏足够信息的情况。因此,SOM 网络在本框架中不进行传统意义上的训练。具体来说,在训练过程开始时,SOM 网络的权重随机初始化,并且在后续的训练阶段中,这些权重保持不变。这种做法有两个主要优点:(1)可以显著降低整个框架的计算复杂度;(2)有助于增加 SOM 特征图中许多部分的熵,从而提升框架的整体性能。

2.4 CNN 拓扑学习模型

CNN 具有较高的准确率和强大的泛化能力,是分析配电网数据的理想选择。CNN 已广泛应用于电力系统中的各种预测任务,包括负荷预测、价格预测和可再生能源发电预测等,显示出其在处理复杂数据和模式识别方面的显著优势。

在本文提出的配电网拓扑辨识框架中,首先使用 SOM 将配电网高维数据映射为二维特征图数据,进而使用 CNN 对从 SOM 中提取的二维特征图进行分类。

设计适合的 CNN 结构对于整体性能至关重要,特别是在拓扑学习任务中更为关键。尽管单个卷积层可能足以满足小型电网的特征提取需求,但如本文所讨论,大型配电网通常需要更复杂的结构。考虑到配电网数据的复杂性和规模,本文采用了由两个卷积层和两个池化层组成的 CNN(图 3)。这种设计确保了 CNN 能够充分提取和处理来自 SOM 的输入数据,并将其转化为适合分类任务的低维特征,从而实现对配电网拓扑的准确识别。通过这种结合使用 SOM 和 CNN 的方法,能够有效地处理和分析复杂的配电网数据,从而显著提升了拓扑辨识的准确性和效率。

2.5 算法流程与框架

本文设计的 SOM-CNN 拓扑辨识算法过程如图 4 所示。首先,通过最大-最小标准化和 PCA 对输入数据进行预处理,这一步旨在减少特征数量,从而提高深度学习框架的训练效率。然后,利用 SOM 对处理后的数据进行特征提取,生成二维特征图。最后,利用 CNN 训练生成的特征图和拓扑标签之间的映射关系。在训练完成后,使用训练好的网络进行配电

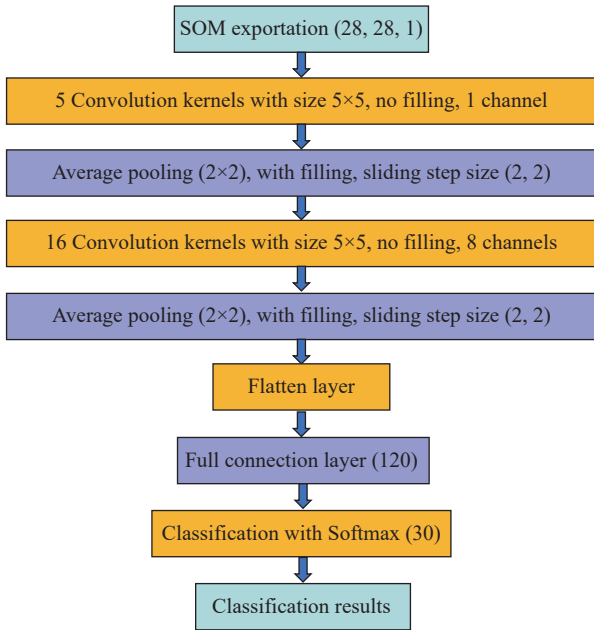


图 3 CNN 详细结构

Fig. 3 Detailed structure of CNN

网拓扑结构的辨识。

3 数值仿真

本节通过在 IEEE 33 节点 [26]、69 节点 [27] 和 123 节点 [28] 配电网算例上进行仿真实验, 对 SOM-CNN 配电网拓扑辨识框架的性能进行了测试和验证。案例研究中使用的所有数据均基于实际负载数据并通过 MATPOWER [29] 生成。所有与拓扑辨识方法相关的代码, 包括预处理、训练和测试, 均用 Python 语言编写。所有实验均在一台配备了 i7-8750H CPU、16 GB RAM 和 NVIDIA 1050 Ti GPU 的个人计算机上进行。

3.1 测试系统设置

对于每个配电网算例, 通过连接和断开不同支路构建了 30 个子拓扑。每个子拓扑生成 2000 组电压幅值和功率注入数据。将这些数据按 7 : 3 的比例划分为训练数据和测试数据, 其中 20% 的训练数据在训练过程中作为验证数据。具体而言, 训练集、验证集和测试集的数据数量分别为 33 600、8 400 和 18 000。

为了模拟分布式能源在配电网中高比例接入的情况, 将风力发电机和光伏发电接入原有电网, 其实际输出功率按小时变化 [30]。配电网初始拓扑和分布式能源位置如图 5~图 7 所示, WE 和 PV 分别表示风力发电机和光伏发电。此外, 在实际应用中, 由于微型相量测量单元 (PMU) 的测量误差, 电压和功

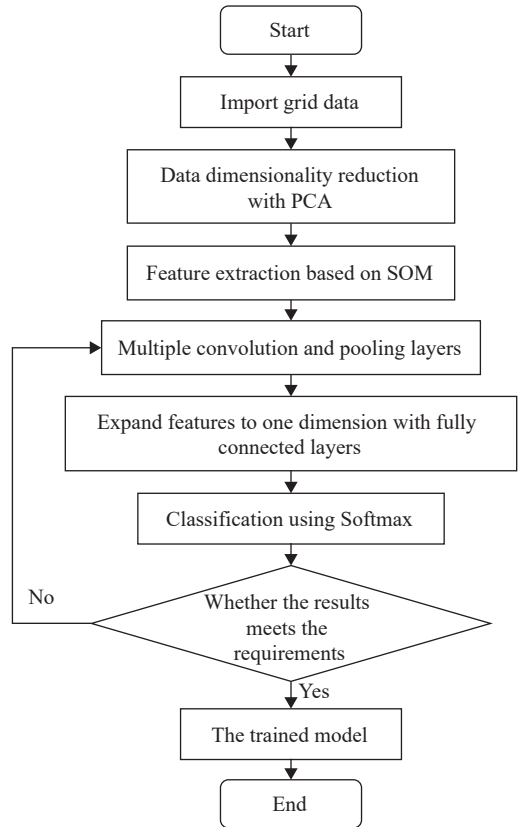


图 4 SOM-CNN 框架流程图

Fig. 4 Flow chart of the SOM-CNN framework

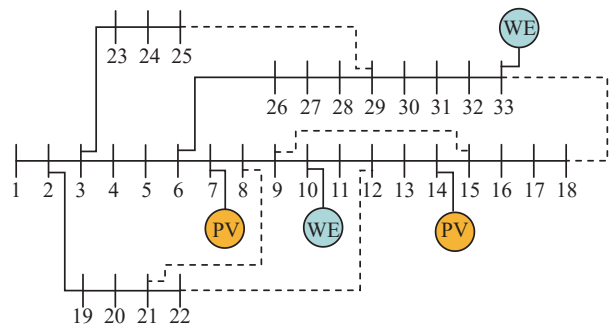


图 5 IEEE 33 节点配电网系统的基本拓扑

Fig. 5 Basic topology of IEEE 33-bus distribution network system

率数据很可能存在偏差。为了模拟这种偏差, 向原始数据中添加了零均值的高斯误差, 其 3σ 偏差与原始值的 0.01% 相匹配。

3.2 超参数优化

超参数是影响 CNN 性能的关键因素之一, 尽管近年来出现了大量的搜索和智能调优算法来处理这个问题, CNN 的超参数优化仍然面临诸多挑战: (1) 通常情况下, CNN 包含大量超参数 (通常从 100 到 1000 不等), 这使得调优过程对计算资源要求很高。(2) 每个 CNN 的超参优化算法会根据许多不同的因素而变化, 例如 CNN 的结构和输入数据的特征。

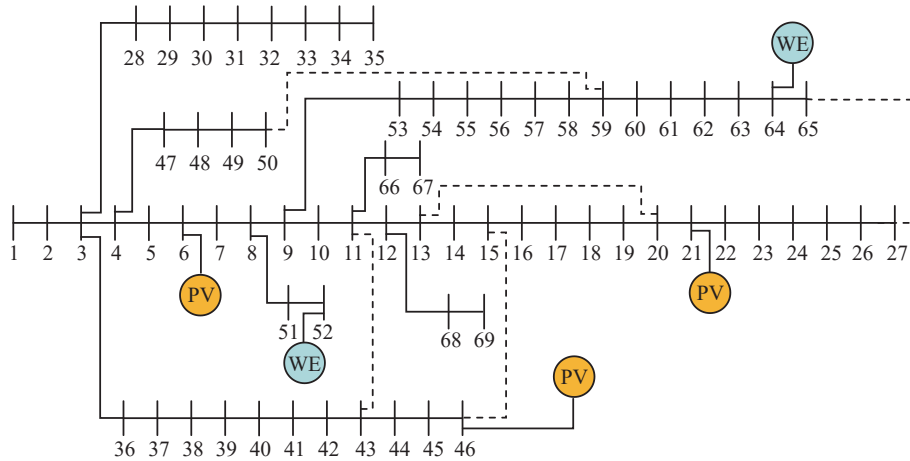


图 6 IEEE 69 节点配电网系统的基本拓扑

Fig. 6 Basic topology of IEEE 69-bus distribution network system

近年来,许多搜索和优化算法在 CNN 等人工神经网络的超参数优化领域得到了广泛应用,并取得了显著效果,如 Nelder-Mead 方法、粒子群优化 (PSO)、Sobol 序列、网格搜索和灰狼优化 (GWO)。Nelder-Mead 是一种启发式搜索方法,通常能够收敛到非平稳点; PSO 算法模拟了鸟群觅食行为,用于优化问题的全局搜索; Sobol 序列是一种准随机低差异序列,相较于随机生成的序列,更均匀地覆盖了整个区域;网格搜索是一种穷举搜索算法,用于遍历目标区域以寻找最优解; GWO 是一种受狼群行为启发的优化算法。

在提出的框架中,测试了本文提到的 5 种算法,不同超参数优化方法的结果如表 1 所示。可以看出,

PSO 在学习率的超参数优化方面表现最为出色,最适合用于 SOM-CNN 方法的超参数优化。因此,本文使用 PSO 来优化包括学习率、批次大小和全连接层神经元数量在内的 CNN 超参数。其他超参数,如 CNN 的结构,则通过手动调整。33 节点配电网拓扑辨识框架中 CNN 的超参数如表 2 所示。

3.3 拓扑辨识结果

为了进一步验证 SOM-CNN 配电网拓扑辨识框架的有效性,本文还采用了其他几种方法进行对比,包括支持向量机 (SVM)^[31]、轻量级梯度提升机 (LightGBM)^[32]、深度神经网络 (DNN)^[20] 和一维卷积神经网络 (1DCNN)^[21]。在这些方法中,所有用于训练和测试的数据均经过最小-最大标准化处理。

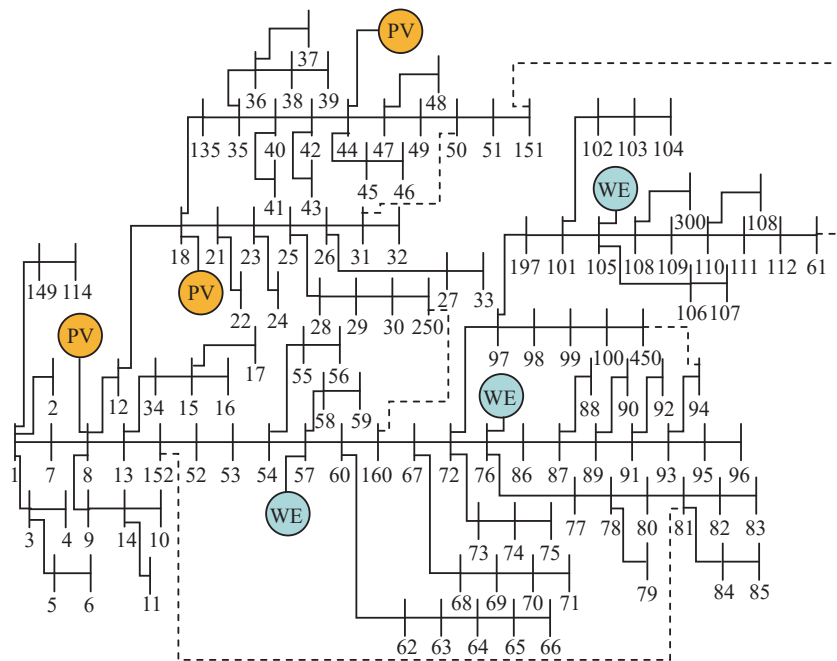


图 7 IEEE 123 节点配电网系统的基本拓扑

Fig. 7 Basic topology of IEEE 123-bus distribution network system

表 1 不同超参数优化方法的结果

Table 1 Hyperparameter optimization results of different methods

Method	Learning rate	Accuracy
Nelder-mead	0.001 076	0.996 93
PSO	0.001 207	0.997 10
Sobol	0.000 780	0.996 84
Grid Search	0.001 132	0.997 03
GWO	0.000 924	0.996 96

在 33 节点配电网算例中, DNN 的结构为 3 层, 神经元数量分别为 20、50 和 100; 1DCNN 的结构包括 3 个卷积层, 卷积核数量依次为 16、32 和 32。

表 2 33 节点配电网中使用的 CNN 超参数

Table 2 Hyperparameters of CNN used in 33-bus distribution network

Number of convolution kernel in convolution layer		Convolution kernel size	Number of neuron in the fully connected layer	Output dimension	Maximum training batch	Batch size	Learning rate	Optimizer	Activation function
1	2								
8	90	(5,5)	120	30	130	56	0.001 2	Adam	Tanh

3.3.1 准确率 图 8 展示了 5 种方法在 IEEE 33、69、123 节点配电网算例上的辨识准确率测试结果, 从辨识准确率的角度来看, SOM-CNN 方法在所有辨识方法中表现最佳。同时, 通过与 1DCNN 方法的对比, 验证了 SOM 在所提方法中的必要性和有效性。SOM-CNN 方法通过 SOM 从高维特征数据中提取二维特征图, 然后使用 CNN 进行拓扑辨识学习, 有效克服了 1DCNN 在捕捉序列中较长距离依赖关系的不足, 提高了拓扑辨识准确率。随着配电网规模的增大, SOM-CNN 相较于 1DCNN 的性能提升显著增加, 在 33 节点算例中, 性能提升了 1.1%, 而在 69 节点和 123 节点算例中, 性能分别提升了 1.3% 和 2.1%, 进一步证明了 SOM-CNN 在处理复杂拓扑结构时的显著优势。

为了进一步验证 SOM-CNN 方法的有效性, 计算了不同算例中 SOM-CNN 辨识结果的混淆矩阵, 并将这些矩阵缩放到 0~1 的范围, 结果如图 9 所示。混淆矩阵展示了每个类别被正确分类和误分类的情况, 从图中可以看出, 混淆矩阵的对角线元素接近于 1, 而非对角线元素接近于 0, 这表明所提出的方法在不同规模的配电网算例中均展现出了较高的辨识精度和灵敏度。

3.3.2 训练和预测时间 在实际应用中, 配电网拓扑辨识方法需要具备快速辨识的能力, 以便及时进行必要的调整。对于未见过的拓扑结构, 期望该方法

DNN 和 1DCNN 的批次大小均设为 64, 学习率为 0.001, 优化器和激活函数与所提方法保持一致。SVM 的惩罚系数和高斯核参数分别设置为 1 和 0.01, 而 LightGBM 的学习率、树的数量和最大深度分别设定为 0.01、1000 和 10。在处理更大规模的配电网算例时, 对比方法的参数将通过上述设置进行手动调整和测试, 以确保对比实验的可靠性。

在所有数据驱动方法中, 数据在训练和测试前均经过了最小-最大标准化处理。为了全面评估各方法在拓扑辨识任务中的表现, 本文选择了准确率、训练和预测时间、鲁棒性以及处理缺失值数据能力等多个指标进行比较分析。

能在有限时间内重新训练自身。这不仅要求模型在处理大量实时数据时具备高效的计算能力, 还需在面对突发情况时快速响应, 适应新的数据模式。

为了测试这两个指标, 可以通过衡量每组数据的识别时间和模型训练时间来评估方法的效率。在案例研究中, 对每种方法的训练和测试耗时进行了测试, 结果如图 10 所示。可以看出, 与传统机器学习方法 LightGBM 和 SVM 相比, SOM-CNN 模型的训练速度较慢, 但相较于深度学习 DNN 和 1DCNN, 其效率更高, 这验证了所提方法中基于 PCA 和 SOM 的数据降维和特征提取在提升效率方面的优势。在辨识速度上, 尽管 SOM-CNN 步骤较多, 仍可以在 0.0006 s 内完成预测, 能够满足工业应用的需求。

3.3.3 不同噪声水平下的拓扑辨识 在基本测试系

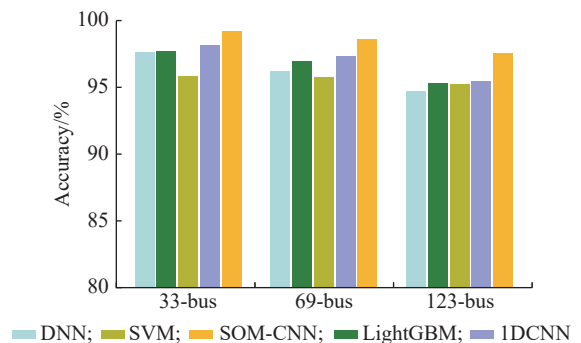


图 8 5 种不同方法的辨识结果

Fig. 8 Identification results of five different methods

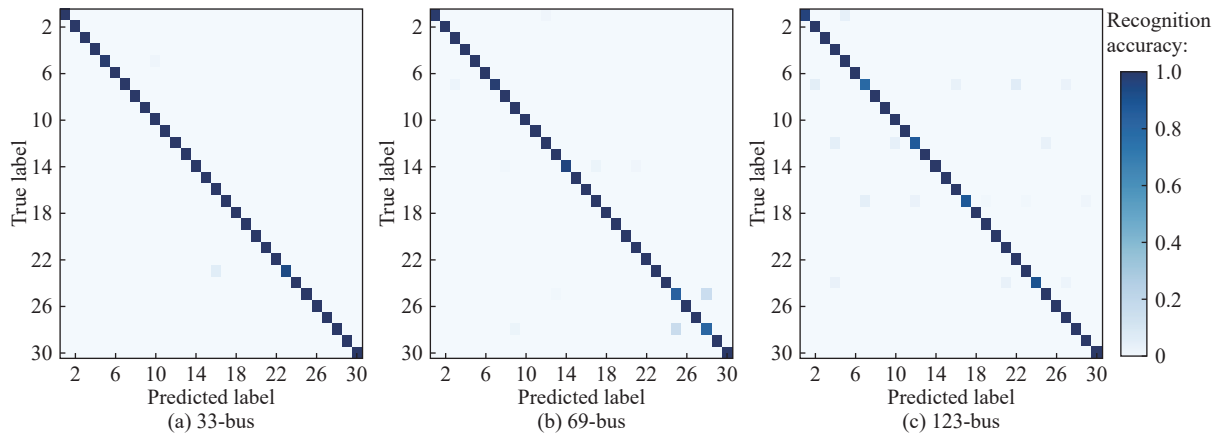


图 9 SOM-CNN 结果的混淆矩阵

Fig. 9 Confusion matrices of SOM-CNN results

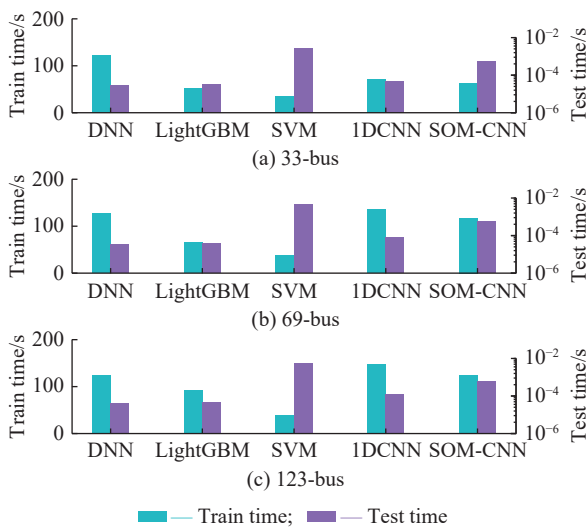


图 10 不同方法的时间消耗

Fig. 10 Time consumption of different methods

统中,假设所有节点都配备了微型 PMU(Phasor measurement unit)进行测量,因此噪声水平相对较低。然而,现实中并非所有网络都配备微型 PMU,有些仅使用 PMU,甚至有些网络仍依赖传统的测量方法。在数据偏差较大的情况下,采用合适且高效的方法进行拓扑辨识显得尤为重要。为了测试 SOM-CNN 方法在处理高偏差情况下的适应能力,在所有数据中添加了基本测试系统的 5 倍和 10 倍高斯噪声,以模拟未配备微型 PMU 测量的配电网系统。不同配电网案例的测试结果如图 11 所示,通过分析可知,相较于其他方法, SOM-CNN 在高噪声环境下仍表现出较高的稳定性和准确性,性能在噪声干扰时下降较少,这表明 SOM-CNN 方法在处理配电网数据中的噪声方面展现出较强的鲁棒性,进一步证明了其在实际应用中的可靠性和有效性。

3.3.4 缺值实验 为了评估 SOM-CNN 方法对缺失

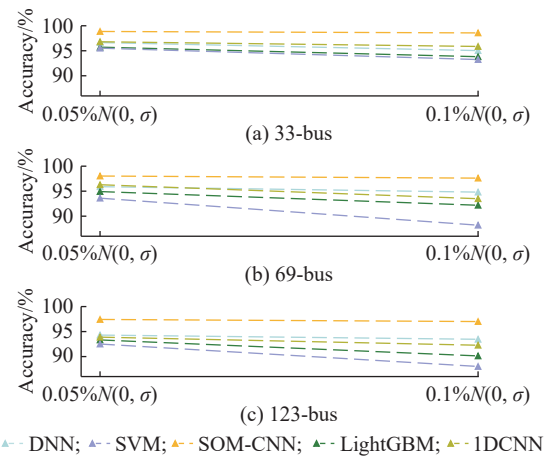


图 11 不同方法在高噪声电网数据上的准确率

Fig. 11 Accuracy of different methods on power grid data with higher noise

值数据的适应性,本文在测试组中随机删除了一定比例的数据,并采用文献 [33] 中的最小方差法对缺失值进行补全。表 3 示出了数据缺失率分别为 2% 和 5% 时, DNN、1DCNN 和 SOM-CNN 方法在 33 节点配电网系统中的拓扑辨识结果。结果显示,尽管缺失值会影响准确率,所提出的 SOM-CNN 方法仍能有效适应配电网数据中的缺失情况,保持可靠的拓扑辨识性能。同时,与 DNN 和 1DCNN 方法相比, SOM-CNN 的准确率下降幅度较小,表明该方法对缺失值的适应性更强。

表 3 不同缺失率下 33 节点配电网系统中 3 种方法的准确率
Table 3 Accuracy of three methods in a 33-bus distribution network system at different missing rates

Missing rate	Accuracy/%		
	DNN	1DCNN	SOM-CNN
2%	90.67	91.04	95.42
5%	81.36	80.27	88.68

4 结束语

本文提出了一种结合 SOM 与 CNN 的深度学习框架,用于配电网拓扑辨识。该框架利用 PCA 和 SOM 进行输入数据的降维和特征提取,将输入数据转换为二维特征图;通过 CNN 学习输入特征与拓扑标签之间的映射关系,从而实现对配电网拓扑结构的准确辨识;通过对 33、69、123 节点配电网算例的数值仿真结果进行分析,该深度学习框架不仅在准确率上优于其他方法,还展现出可靠的辨识效率。此外,该框架对测量系统的要求较低,因为其主要依赖于电压幅值和有功功率注入数据。更值得注意的是,该框架在处理具有高噪声和缺失值的数据时,依然保持了较高的准确性,具有出色的实用性和可靠性。这表明所提出的拓扑辨识框架在面对测量噪声时具有出色的鲁棒性,即使在缺乏微型 PMU 或 PMU 的网络中也能有效运行,这个特点显著提升了该框架在实际部署中的可行性和应用价值。

在未来的工作中,我们将致力于进一步优化数据预处理流程,以缩短训练和识别时间。此外,将 SOM-CNN 方法与时间序列分析相结合,可以有效应对智能配电网中的时变系统问题。这样的结合将提升系统的适应性和精确性,进一步增强其在实际应用中的有效性。

参考文献:

- [1] 刘奉奉,薛栋.基于分区选择的主动配电网分布式最优潮流分析[J].华东理工大学学报(自然科学版),2024,50(4):550-559.
- [2] MESKIN M, DOMIJAN A, GRINBERG I. Impact of distributed generation on the protection systems of distribution networks: analysis and remedies—review paper[J]. IET Generation, Transmission & Distribution, 2020, 14(24): 5944-5960.
- [3] 陆佳琳,方舒,杜松怀,等.高比例可再生能源新型农村电网的技术挑战与展望[J].电力需求侧管理,2022,24(6):38-43.
- [4] 刘中华,牛玉刚,贾廷纲.基于风-光-储联合优化的最优潮流[J].华东理工大学学报(自然科学版),2022,48(2):221-230.
- [5] ABUR A, EXPOSITO A G. Power System State Estimation: Theory and Implementation[M]. [s.l.]: CRC Press, 2004.
- [6] GAO Y, ZHANG Z, WU W, *et al.* A method for the topology identification of distribution system[C]//2013 IEEE Power & Energy Society General Meeting. Vancouver, BC: IEEE, 2013: 1-5.
- [7] ARGHANDEH R, GAHR M, VON MEIER A, *et al.* Topology detection in microgrids with micro-synchrophasors [C]//2015 IEEE Power & Energy Society General Meeting. Denver, CO, USA: IEEE, 2015: 1-5.
- [8] LI P, LI Y, XIONG Q. International journal of electrical power with energy systems[C]//Proceedings of the IEEE International Symposium on Industrial Electronics. Piscataway, NJ, USA: IEEE, 2014: 749-759.
- [9] TIAN Z, WU W, ZHANG B. A mixed integer quadratic programming model for topology identification in distribution network[J]. IEEE Transactions on Power Systems, 2015, 31(1): 823-824.
- [10] POUDEL S, RAMACHANDRAN T, VEERAMANY A, *et al.* Topology identification using graph theory informed state estimation-based model selection for power distribution systems[J]. IEEE Transactions on Industrial Informatics, 2023.
- [11] GARCÍA S, MORA-MERCHÁN J M, LARIOS D F, *et al.* Phase topology identification in low-voltage distribution networks: A Bayesian approach[J]. International Journal of Electrical Power & Energy Systems, 2023, 144: 108525.
- [12] LUAN W, PENG J, MARAS M, *et al.* Smart meter data analytics for distribution network connectivity verification[J]. IEEE Transactions on Smart Grid, 2015, 6(4): 1964-1971.
- [13] YU J, WENG Y, RAJAGOPAL R. PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids[J]. IEEE Transactions on Power Systems, 2017, 33(4): 4335-4347.
- [14] LIAO Y, WENG Y, LIU G, *et al.* Urban MV and LV distribution grid topology estimation via group lasso[J]. IEEE Transactions on Power Systems, 2018, 34(1): 12-27.
- [15] CAVRARO G, KEKATOS V, VEERAMACHANENI S. Voltage analytics for power distribution network topology verification[J]. IEEE Transactions on Smart Grid, 2017, 10(1): 1058-1067.
- [16] KEKATOS V, GIANNAKIS G B, BALDICK R. Online energy price matrix factorization for power grid topology tracking[J]. IEEE Transactions on Smart Grid, 2015, 7(3): 1239-1248.
- [17] FLYNN C, PENGWAH A B, RAZZAGHI R, *et al.* An improved algorithm for topology identification of distribution networks using smart meter data and its application for fault detection[J]. IEEE Transactions on Smart Grid, 2023, 14(5): 3850-3861.
- [18] 杨挺,赵黎媛,王成山.人工智能在电力系统及综合能源系统中的应用综述[J].电力系统自动化,2019,43(1):2-14.
- [19] LI H, LIANG W, LIANG Y, *et al.* Topology identification method for residential areas in low-voltage distribution networks based on unsupervised learning and graph theory[J]. Electric Power Systems Research, 2023, 215: 108969.
- [20] JAFARIAN M, SOROUDI A, KEANE A. Distribution system topology identification for DER management systems using deep neural networks[C]//2020 IEEE Power & Energy Society General Meeting (PESGM). Montreal, QC, Canada: IEEE, 2020: 1-5.

- [21] NI J, TANG Z, LIU J, *et al.* A topology identification method based on one-dimensional convolutional neural network for distribution network[J]. *Energy Reports*, 2023, 9: 355-362.
- [22] KOHONEN T. The self-organizing map[J]. *Proceedings of the IEEE*, 1990, 78(9): 1464-1480.
- [23] BARLETTA V S, CAIVANO D, NANNAVECCHIA A, *et al.* Intrusion detection for in-vehicle communication networks: An unsupervised Kohonen SOM approach[J]. *Future Internet*, 2020, 12(7): 119.
- [24] GAYATHRI M, ARIYARATNE A, KAHAWALA S, *et al.* Learning rule optimization and comparative evaluation of accelerated self-organizing maps for industrial applications[C]//IECON 2021—47th Annual Conference of the IEEE Industrial Electronics Society. [s.l.]: IEEE, 2021: 1-6.
- [25] SONG W, MAHALEC V, LONG J, *et al.* Modeling the hydrocracking process with deep neural networks[J]. *Industrial & Engineering Chemistry Research*, 2020, 59(7): 3077-3090.
- [26] BARAN M E, WU F F. Optimal capacitor placement on radial distribution systems[J]. *IEEE Transactions on Power Delivery*, 1989, 4(1): 725-734.
- [27] BARAN M E, WU F F. Network reconfiguration in distribution systems for loss reduction and load balancing[J]. *IEEE Transactions on Power Delivery*, 1989, 4(2): 1401-1407.
- [28] KERSTING W H. Radial distribution test feeders[C]//2001 IEEE Power Engineering Society Winter Meeting Conference Proceedings. Columbus, OH, USA: IEEE, 2001: 908-912.
- [29] ZIMMERMAN R D, MURILLO-SÁNCHEZ C E, THOMAS R J. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education[J]. *IEEE Transactions on Power Systems*, 2010, 26(1): 12-19.
- [30] DOROSTKAR-GHAMSARI M R, FOTUHI-FIRUZABAD M, LEHTONEN M, *et al.* Value of distribution network reconfiguration in presence of renewable energy resources[J]. *IEEE Transactions on Power Systems*, 2015, 31(3): 1879-1888.
- [31] BRERETON R G, LLOYD G R. Support vector machines for classification and regression[J]. *Analyst*, 2010, 135(2): 230-267.
- [32] KE G, MENG Q, FINLEY T, *et al.* Lightgbm: A highly efficient gradient boosting decision tree [C]//Advances in Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates, 2017: 3146-3154.
- [33] 裴宇婷, 秦超, 余贻鑫. 基于 LightGBM 和 DNN 的智能配电网在线拓扑辨识 [J]. *天津大学学报 (自然科学与工程技术版)*, 2020, 53(9): 939-950.

Topology Identification for Smart Distribution Networks Based on Data-Driven and Deep Neural Networks

DU Wenkai, CHEN Xinyi, XUE Dong

(Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Topology identification of the network structure is fundamental for the optimization and control of distribution systems. With the high penetration of renewable energy generation, such as wind and solar power, the topology of distribution networks has become more complex and changes frequently, significantly increasing the difficulty of topology identification. To improve the accuracy of topology identification, this paper proposes a distribution network topology identification method based on a deep learning framework that combines Self-Organizing Maps (SOM) and Convolutional Neural Networks (CNN), taking into account the structure and operational characteristics of distribution networks. This method first uses Principal Component Analysis (PCA) to reduce the dimensionality of high-dimensional voltage magnitude and active power data. It then employs SOM to extract data features and transform them into a two-dimensional feature map. Finally, CNN is used to learn the mapping between the input features and topology labels, enabling accurate identification of the distribution network topology. The effectiveness of the proposed method is validated through simulation experiments on 33-, 69-, and 123-bus distribution network cases. Compared to other methods, this approach demonstrates significant advantages in terms of identification accuracy and robustness.

Key words: distribution network system; topology identification; data driven; deep learning; convolutional neural network

(责任编辑: 李娟)