

基于深度字典学习的输电线路故障分类方法

张宇博¹, 郝治国¹, 林泽暄¹, 杨松浩¹, 刘志远², 于晓军²

(1. 西安交通大学 电气工程学院, 陕西 西安 710049; 2. 国网宁夏电力有限公司, 宁夏 银川 750001)

摘要:针对当前输电线路故障分类识别方法存在的阈值整定复杂、人工智能算法可解释性不足等问题,提出了一种基于深度字典学习的输电线路故障分类方法。该方法利用稀疏性约束驱动字典自动提取样本中的故障特征,同时深度字典结构使得所提取的故障特征具有较好的层次性和物理含义,符合人对故障的直观认识,一定程度上解决了数据驱动型方法可解释性不足的问题。最后,通过PSCAD/EMTDC仿真验证了所提方法的有效性。

关键词:输电线路;故障分类;稀疏表示;深度字典学习

中图分类号:TM 75

文献标志码:A

DOI:10.16081/j.epae.202204031

0 引言

电力系统中输电线路运行环境复杂,易受各种人为或自然因素影响而发生故障。在线路故障后快速、准确地识别故障类型,对于提高故障定位精度、缩短故障线路恢复运行时间、提高电力系统稳定性等具有重要意义。

已有的输电线路故障分类方法大致可分为基于物理模型的方法^[1-3]和基于数据驱动的方法^[4-10]。基于物理模型的输电线路故障分类方法通过构建输电线路的物理模型,对输电线路不同故障类型之间的差异进行解析,设计指标并选取阈值,实现故障类型的判别,具有物理意义明确、易于微机实现的优点,但也存在一定的局限性,例如:文献[1-2]基于输电线路的故障稳态模型判别故障类型,但故障前后系统参数的变化对方法的准确性影响较大;文献[3]基于输电线路故障暂态模型方法的准确性受故障条件的影响较大,且对采样率等硬件条件要求较高。基于数据驱动的输电线路故障分类方法借助大数据、人工智能等技术,从样本中自动提取故障特征,避免了复杂的物理建模过程和繁琐的阈值整定过程,例如:文献[4]利用滤波后的三相电流振幅作为故障电流特征,并将其输入支持向量机SVM(Support Vector Machine)实现故障分类;文献[5]将小波分析与经人工神经网络ANN(Artificial Neural Network)改良的粒子群优化PSO(Particle Swarm Optimization)算法结合来实现输电线路故障分类。这些基于传统人工智能的故障分类方法,设计流程相对简单,但在

故障特征提取、小波选型等方面,仍部分依赖人为参与^[11]。近年来,深度学习以其强大的非线性映射和特征提取能力得到了广泛的应用并且效果显著^[12]。目前已有研究将深度学习技术应用于输电线路故障分类问题并取得了一定成果。例如:文献[6]将自编码器SAE(Sparse AutoEncoder)与卷积神经网络CNN(Convolutional Neural Network)结合实现交流系统输电线路的故障分类问题;文献[7]利用并联卷积神经网络P-CNN(Parallel Convolutional Neural Network)快速检测直流输电线路的故障类别和故障支路。同样地,在小电流接地系统的故障选相问题中,深度学习的有效性也得以验证^[8-9]。但是深度学习模型是典型的黑箱算法,模型复杂,且所得特征缺乏可解释性。

稀疏表示与深度学习类似,都属于以数据驱动特征自动提取的方法。与深度学习不同的是,稀疏表示理论以字典原子的线性组合来重建信号,所提取的字典原子与信号本身是线性关系,具有较好的实际意义,符合人的直观认识。文献[10]对稀疏表示用于输电线路故障分类问题进行了初步探索,但是模型相对简单,且直接以故障样本组成字典,缺少了对故障样本特征的挖掘和学习过程,字典原子缺乏代表性。

本文基于稀疏表示理论,借鉴深度学习的逐层提取特征的思想,提出了一种可用于输电线路故障类型识别的深度字典学习模型。该模型利用稀疏性约束逐层学习故障数据中的典型结构特征,构成深度故障字典,最后根据故障样本的重建误差确定其所属类别。该方法无须人为提取故障特征,且模型学习到的特征符合人对故障的直观认识,一定程度上解决了故障特征的可解释性问题。大量仿真数据验证了本文所提方法具有较好的鲁棒性和泛化能力。

收稿日期:2021-08-16;修回日期:2021-11-17

在线出版日期:2022-04-12

基金项目:国家自然科学基金资助项目(52007143);中国博士后科学基金资助项目(2021M692526)

Project supported by the National Natural Science Foundation of China(52007143) and China Postdoctoral Science Foundation(2021M692526)

1 稀疏表示理论与深度字典学习模型

1.1 稀疏表示理论

稀疏表示就是使用少量基本信号的线性组合表示目标信号,通过“稀疏性”这一强制要求提取目标信号中的关键信息,从而实现信号的压缩和特征提取。其数学模型如式(1)所示。

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{r} \quad (1)$$

式中: $\mathbf{y}$ 为目标信号(样本), $\mathbf{y} \in \mathbf{R}^{n \times 1}$, n 为样本维度; $\mathbf{D} \in \mathbf{R}^{n \times M}$ 为基本信号矩阵, M 为基本信号数量; $\mathbf{x} \in \mathbf{R}^{M \times 1}$ 为目标信号 $\mathbf{y}$ 在基本信号矩阵 $\mathbf{D}$ 下分解的系数向量且要求其是稀疏的; $\mathbf{r} \in \mathbf{R}^{n \times 1}$ 为残差向量。

基本信号矩阵又被称为字典,每个基本信号都是一个字典原子,因此字典 $\mathbf{D}$ 可表示为:

$$\mathbf{D} = [\mathbf{d}_1 \ \mathbf{d}_2 \ \cdots \ \mathbf{d}_M] \quad (2)$$

式中: $\mathbf{d}_i \in \mathbf{R}^{n \times 1}$, $i=1,2,\dots,M$ 。

字典的构造是稀疏表示理论的关键,根据构造方法的不同可分为解析方法和学习方法2种^[13]。解析方法中字典是利用某种数学变换构造得到的,如离散余弦变换、小波变换等^[14],然后在预定义字典上寻求目标信号的稀疏表示。学习方法中字典是根据样本数据的特点自适应学习得到的,即字典原子可以根据样本和待求解问题的特点灵活调整。与解析方法相比,学习方法的字典原子形态更丰富,针对性更强,能更好地与样本内含的结构相匹配。因此,本文使用学习方法构造故障字典。

1.2 基于学习方法的字典构造

如1.1节所述,学习字典的构造依赖于研究的问题及样本数据的特点。具体地,结合式(1)可知,基于学习方法构造字典的实质就是求解特定的矩阵 $\mathbf{D}$ 使其在满足一定的稀疏性约束条件下使样本数据的表示误差最小。因此,学习方法中字典的构造可以转换为求解如式(3)所示的约束优化问题。

$$\begin{cases} \underset{\mathbf{D}, \mathbf{X}}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad \|\mathbf{x}_i\|_0 \leq s \quad \mathbf{x}_i \in \mathbf{X} \end{cases} \quad (3)$$

式中: $\mathbf{Y}$ 为样本矩阵, $\mathbf{Y} \in \mathbf{R}^{n \times N}$, N 为样本数量; $\mathbf{X}$ 为系数矩阵, $\mathbf{X} \in \mathbf{R}^{M \times N}$; $\mathbf{x}_i \in \mathbf{R}^{M \times 1}$ 为第 i 个样本对应的系数向量; s 为稀疏度指标, $s \in \mathbf{Z}^+$; $\|\cdot\|_{\mathrm{F}}$ 为Frobenius范数; $\|\cdot\|_0$ 为0范数,即向量中非零元素个数。

由于存在2个未定矩阵变量 $\mathbf{D}$ 和 $\mathbf{X}$,直接求解式(3)是一个NP-hard问题^[15],因此一般采用交替迭代进行求解:先固定字典 $\mathbf{D}$ 求解系数矩阵 $\mathbf{X}$ (此时 $\mathbf{D}$ 为已知量),然后利用求解得到的稀疏矩阵 $\mathbf{X}$ 对字典 $\mathbf{D}$ 进行更新(此时 $\mathbf{X}$ 为已知量),重复上述操作直至误差满足要求或达到最大迭代次数。

已知字典 $\mathbf{D}$ 求解系数矩阵 $\mathbf{X}$ 是一个标准的编码问题,利用正交匹配追踪OMP(Orthogonal Matching

Pursuit)^[16]算法可以快速求解。OMP算法的核心思想是:以贪婪迭代的方法选择字典 $\mathbf{D}$ 中的原子,使得在每次迭代的过程中所选择的字典原子与当前残差向量相关性最大,从原始信号向量中减去相关部分并反复迭代,直到迭代次数达到预设稀疏度。算法实现如附录A图A1所示。

固定系数矩阵 $\mathbf{X}$ 更新字典 $\mathbf{D}$ 是学习字典模型中的关键步骤,其中应用较为广泛的有最优方向法MOD(Method of Optimal Directions)^[17]、K-奇异值分解法K-SVD(K-Singular Value Decomposition)^[18]、序列泛化K均值SGK(Sequential Generalization of K-means)^[19]等算法。考虑到SGK算法能够逐原子更新,与SVD等其他算法相比复杂度更低,计算高效且对硬件资源占用更少,本文利用SGK算法实现字典 $\mathbf{D}$ 的更新,具体如式(4)所示。

$$\mathbf{d}_k^{\text{new}} = \mathbf{E}_k \mathbf{X}_{\text{row}k}^{\mathrm{T}} (\mathbf{X}_{\text{row}k} \mathbf{X}_{\text{row}k}^{\mathrm{T}})^{-1} \quad (4)$$

式中: $\mathbf{d}_k^{\text{new}}$ 为更新后的第 k 个字典原子; $\mathbf{X}_{\text{row}k}$ 为系数矩阵 $\mathbf{X}$ 的第 k 行。 $\mathbf{E}_k$ 由式(5)计算得到。

$$\mathbf{E}_k = \mathbf{Y} - \sum_{j \neq k} \mathbf{d}_j^{\text{old}} \mathbf{X}_{\text{row}j} \quad (5)$$

式中: $\mathbf{d}_j^{\text{old}}$ 为更新前的第 j 个字典原子; $\mathbf{X}_{\text{row}j}$ 为系数矩阵的第 j 行。

1.3 深度字典学习模型

传统的基于学习方法构造的字典是单层结构且以残差为驱动,当稀疏性指标大于1时,学习到的字典原子将由反映样本的轮廓特征向反映样本的细节特征转变,如附录A图A2所示。由式(4)所示的字典原子更新算法可知,这种单层字典结构使得反映样本不同特征尺度的字典原子之间相互影响,削弱了字典原子的泛化性能。进一步地,当字典中存在多个样本非常相似且稀疏性指标大于1时,OMP算法可能会得到错误的重构信号,从而影响方法的有效性和准确性^[20]。

考虑到传统学习方法的局限性,本文借鉴深度学习理论中的逐层特征提取思想,采用字典层数替代稀疏性指标,构建深度字典,即将单层多稀疏度结构转换为多层单稀疏度结构。具体地,构建一个 s 层的字典(s 对应传统学习方法中的稀疏度指标),各层字典的稀疏度 s' 均为1,以信号的表示残差逐层驱动字典学习,其原理如附录A图A3所示。采用多层字典结构保证了不同特征尺度的原子彼此独立、互不影响,同时单稀疏度结构避免了OMP算法可能出现的重构错误问题。

2 基于深度字典学习模型的故障分类算法

输电线路故障类型判别本质上是一个样本分类问题,利用深度字典学习模型解决分类问题时,主要由以下2个步骤构成:

1)深度故障字典的构建,即首先生成带类别标

签的样本库,然后利用1.3节所述的深度字典学习模型按类构建深度故障字典;

2)故障样本分类的实现,即将各类故障字典的同层合并,然后基于合并后的深度故障字典对待分类样本进行稀疏表示,利用稀疏系数对样本按类重构,重构误差最小的类别就是样本所属类别。

2.1 输电线路故障样本库

本文以电流信号作为故障类型判别的关键信息,具体地,利用故障前后一段时间内的各相电流和零序电流值按顺序拼接并归一化形成故障样本。考虑到一般断路器的动作时间为0.06~0.15 s^[21],为保证方法的有效性,取故障前1个工频周期和故障后3个工频周期的数据为有效数据,归一化方法如式(6)所示,某A相接地故障样本示例如附录A图A4所示。

$$\mathbf{i}^* = \frac{\mathbf{i}}{\max(|\mathbf{i}|)} \quad (6)$$

式中: $\mathbf{i}$ 为原始故障样本; $\mathbf{i}^*$ 为归一化故障样本。

输电线路故障类型有单相接地故障(Ag, Bg, Cg)、相间短路故障(AB, AC, BC)、两相短路接地故障(ABg, ACg, BCg)和三相短路故障(ABC),共计10种故障类型。则样本矩阵按上述分类顺序可表示为:

$$\mathbf{Y} = [\mathbf{Y}_1 \cdots \mathbf{Y}_i \cdots \mathbf{Y}_{10}] \quad (7)$$

式中: $\mathbf{Y}_i$ 为属于类别*i*($i=1, 2, \dots, 10$)的样本集, $\mathbf{Y}_i \in \mathbf{R}^{n \times N_i}$, N_i 为属于类别*i*的样本数量, $\sum_{i=1}^{10} N_i = N$ 。

2.2 按类构建深度故障字典

1.3节给出了深度字典学习模型的求解原理和一般形式,现将其扩展到求解多类问题上并给出其详细的求解过程。

由于各故障类型对应的样本库相互独立,故可对故障类别进行解耦,即由2.1节构建的各类样本库 $\mathbf{Y}_i$ 独立导出各类别对应的深度故障字典 $\mathbf{D}_i$ 。由于各子类深度故障字典的求解过程完全相同,下面以类别*i*为例给出深度故障字典的具体求解过程。

基于图A3所示的求解原理可知,各子层字典的求解过程彼此独立,以残差作为驱动,可采用串行递推的思路逐层进行求解。具体地,参照式(3),第*j*层的求解过程可表示为如式(8)所示的数学模型。

$$\begin{cases} \arg\min_{\mathbf{D}_i^j, \mathbf{X}_i^j} \|\Delta_i^{j-1} - \mathbf{D}_i^j \mathbf{X}_i^j\|_F^2 \\ \text{s.t. } \|(\mathbf{x}_m)_i\|_0 = 1 \quad m=1, 2, \dots, N_i \end{cases} \quad (8)$$

式中:下标*i*表示故障类别;上标*j*($j=1, 2, \dots, s$)表示第*j*层; $(\mathbf{x}_m)_i^j$ 为系数矩阵 $\mathbf{X}_i^j$ 的第*m*列。

残差矩阵 Δ_i^j 的定义如式(9)所示。

$$\begin{cases} \Delta_i^0 = \mathbf{Y}_i \\ \Delta_i^j = \Delta_i^{j-1} - \mathbf{D}_i^{j-1} \mathbf{X}_i^{j-1} \end{cases} \quad (9)$$

利用分步迭代求解的思想,首先固定字典 $\mathbf{D}_i^j$,对式(8)利用OMP算法求解系数矩阵 $\mathbf{X}_i^j$;然后固定系数矩阵 $\mathbf{X}_i^j$,更新字典 $\mathbf{D}_i^j$,此时系数矩阵 $\mathbf{X}_i^j$ 和残差矩阵 Δ_i^{j-1} 均是常矩阵,式(8)所示的约束优化问题简化为式(10)所示的无约束优化问题。

$$\arg\min_{\mathbf{D}_i^j} \|\Delta_i^{j-1} - \mathbf{D}_i^j \mathbf{X}_i^j\|_F^2 \quad (10)$$

利用SGK算法对式(10)所示的无约束优化问题进行求解,字典 $\mathbf{D}_i^j$ 按式(4)进行更新,只需利用式(8)与式(3)的对应关系将对应的变量进行替换即可。

迭代结束后,将更新后的 $\mathbf{D}_i^j$ 和 $\mathbf{X}_i^j$ 代入式(9)中求解下一层残差矩阵 Δ_i^j ,再进行下一层字典和系数矩阵的求解。重复上述操作,直到字典层数等于设定的总层数*s*时,类别*i*对应的深度故障字典构建完成。由于各类别的深度故障字典求解过程完全相同,因此对不同类别的样本库重复上述求解过程即可,图1给出了按类别构建深度故障字典的流程图。

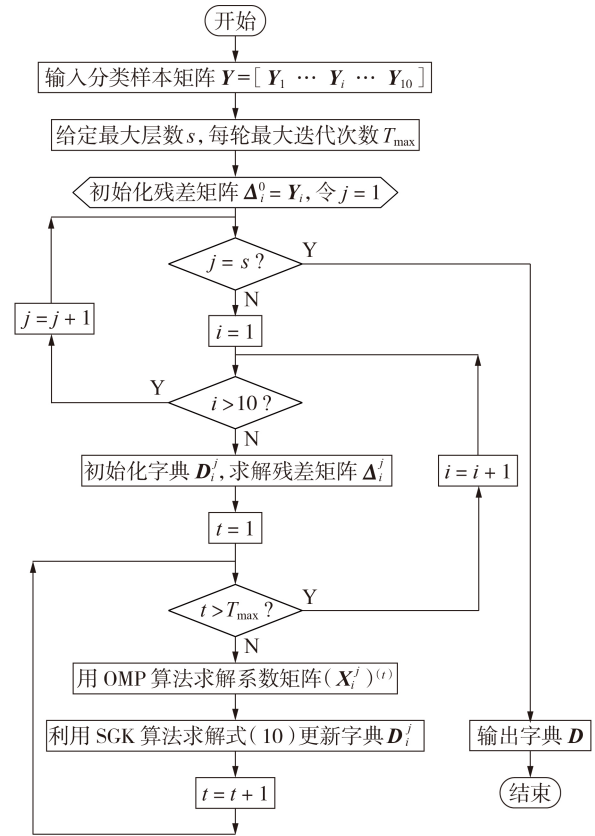


图1 按类别构建深度故障字典流程图

Fig.1 Flowchart of constructing multiclass deep fault dictionary

2.3 故障分类

将上述按类别学习得到的深度故障字典用于故障样本的分类时,其基本思想就是在满足一定稀疏性约束的条件下,利用该含有类别信息的深度故障

字典中原子的线性组合来表示待分类样本 $\mathbf{y}$ 。其数学实质就是在已知字典时,求解待分类样本的深度稀疏表示,根据稀疏表示系数在类间的分布信息判定样本所属类别。样本深度稀疏表示的数学模型为:

$$\begin{cases} \operatorname{argmin}_{\mathbf{x}^j} \|\mathbf{r}^{j-1} - \mathbf{D}^j \mathbf{x}^j\|_2^2 \\ \text{s.t. } \|\mathbf{x}^j\|_0 = 1 \end{cases} \quad (11)$$

式中: $\mathbf{D}^j$ 由各类别对应的深度故障字典中的第 j 层构成,即 $\mathbf{D}^j = [\mathbf{D}_1^j \cdots \mathbf{D}_i^j \cdots \mathbf{D}_{10}^j]$, 对应的系数向量为 $\mathbf{x}^j = [(\mathbf{x}_1^j)^T \cdots (\mathbf{x}_i^j)^T \cdots (\mathbf{x}_{10}^j)^T]^T$; 残差向量的定义如式(12)所示。

$$\begin{cases} \mathbf{r}^0 = \mathbf{y} \\ \mathbf{r}^j = \mathbf{r}^{j-1} - \mathbf{D}^j \mathbf{x}^j \end{cases} \quad (12)$$

根据前文所述,在字典已知的情况下,利用OMP算法可以快速求解式(11)所示的约束优化问题,即以残差为驱动,通过OMP算法可串行递推待分类样本在各层字典上的系数向量,实现对待分类样本的深度稀疏表示。具体地,由式(11)的约束条件可知每层字典的稀疏度为1,利用图A1所示的OMP算法求解稀疏度为1的样本稀疏系数向量即为求解与残差向量最为接近的字典原子及其表示系数。各层稀疏系数的求解是彼此独立的,即重复求解稀疏度为1的OMP算法,以残差为驱动串联各层字典,上一层的残差作为下一层的目标信号,详细求解过程见附录A图A5。

最后,根据系数向量按类别进行信号重构,重构误差最小的类别就是该样本所属的类别,判别方法如式(13)所示。

$$F_{\text{class}}(\mathbf{y}) = \operatorname{argmin}_i \frac{\left\| \mathbf{y} - \sum_{j=1}^s \mathbf{D}_i^j \mathbf{x}_i^j \right\|_2^2}{\|\mathbf{y}\|_2^2} \quad (13)$$

式中: $F_{\text{class}}(\mathbf{y})$ 为本文方法判定样本 $\mathbf{y}$ 所属的故障类别; $\left\| \mathbf{y} - \sum_{j=1}^s \mathbf{D}_i^j \mathbf{x}_i^j \right\|_2^2 / \|\mathbf{y}\|_2^2$ 为第 i 类字典的重构误差。

3 仿真验证

3.1 故障样本库的生成

本文利用PSCAD/EMTDC软件搭建双端电源模型进行仿真获取大量故障样本,故障仿真模型如图2所示,系统电压为220 kV,频率为50 Hz,输电线路长度为80 km,采样频率设置为2 kHz,线路参数参照文献[6]进行设置。

为尽可能多地获取故障仿真数据,提高模型的准确性,本文将各种可能的故障条件进行组合,故障条件设置如表1所示,共计得到5280份故障仿真数据,每种故障类型包含528份样本数据,按2.1节中

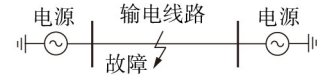


图2 双端电源系统故障仿真模型

Fig.2 Fault simulation model of dual-terminal power system

表1 故障条件设置

Table 1 Settings of fault conditions

参数	参数值
故障类型	Ag, Bg, Cg, AB, AC, BC, ABg, ACg, BCg, ABC
故障初相角 / (°)	0, 30, 60, ..., 330
故障电阻 / Ω	0.01, 0.1, 5, 10, 25, 50, 75, 100, 150, 200
故障距离 / km	20, 40, 60, 80

所述方法,生成故障样本库 $\mathbf{Y}$ 。

对于任一待分类样本 $\mathbf{y}$,其构建过程与2.1节中样本的构建过程相同。具体地,提取输电线路故障前一周期和故障后3个周期的三相电流和零序电流的采样值,按顺序拼接并进行归一化操作,即待分类样本 $\mathbf{y}$ 可表示为:

$$\mathbf{y} = [\mathbf{i}_A \ \mathbf{i}_B \ \mathbf{i}_C \ \mathbf{i}_0]^T \quad \mathbf{y} \in \mathbf{R}^{640 \times 1} \quad (14)$$

式中: $\mathbf{i}_A$ 、 $\mathbf{i}_B$ 、 $\mathbf{i}_C$ 为归一化的三相电流采样值, $\mathbf{i}_0$ 为归一化的零序电流采样值,且 $\mathbf{i}_A$ 、 $\mathbf{i}_B$ 、 $\mathbf{i}_C$ 、 $\mathbf{i}_0 \in \mathbf{R}^{1 \times 160}$ 。

3.2 构建深度故障字典

按类别生成故障样本库 $\mathbf{Y}$ 后,基于2.2节中的求解方法可构建用于故障分类的深度故障字典 $\mathbf{D}$ 。需要强调的是,求解过程中有一些关键的超参数需要确定,考虑到最终的残差矩阵 Δ_i^s 可以反映模型的误差,本文定义代价函数 J 如式(15)所示。

$$J = \sum_{i=1}^{10} \|\Delta_i^s\|_F \quad (15)$$

代价函数 J 的大小可以衡量模型的性能, J 值越小,表示模型误差越小,模型性能越好。下面以代价函数 J 为指标,讨论字典初始化方法、字典规模以及层数等关键超参数对模型性能的影响。

3.2.1 字典的初始化方法

求解子类 i 对应的深度字典 $\mathbf{D}_i$ 时,需要首先确定每一子层字典 $\mathbf{D}_i^j$ 的初值。本文比较了 K 均值初始化、直接初始化和随机初始化这3种不同初始化方法对模型性能的影响,结果如附录A图A6所示。 K 均值初始化方法是利用 K 均值算法^[22]得到 Δ_i^{j-1} 的 K 个聚类中心并将其作为第 j 层字典的初值;直接初始化方法是直接从 Δ_i^{j-1} 中选取一定量的样本作为字典的初值;随机初始化方法是随机生成字典的初值。需要说明的是,对于第1层字典 $\mathbf{D}_i^1$ 而言,其对应的上一层残差矩阵 Δ_i^0 即为类别 i 的样本矩阵 $\mathbf{Y}_i$,如式(9)所示。从图A6可以看出: K 均值初始化方法效果最好;随机初始化方法因在初始化时没有利用到样本本身的信息,导致初始点与最优解相差太远,效果较

差; K 均值初始化方法与直接初始化方法的效果相近,但 K 均值初始化方法所得结果的稳定性更好,这是由于 K 均值算法的内在逻辑保证了初始化的字典中不会有特别相近的原子,使得初始字典中包含了更多样本信息。因此,本文选择 K 均值初始化方法。

3.2.2 字典的规模

字典的规模也是影响模型性能的重要因素,应综合考虑模型性能和求解代价,选择合理的字典规模和层数,在保证模型性能的前提下尽可能缩短模型训练时间、减少对硬件资源的占用。

在 $[0.1, 0.5]$ 范围内,从 0.1 开始每隔 0.05 取 1 个值作为字典中原子总数占总样本总数的比例,并设置字典层数为 1—8 层,在上述设置下进行迭代求解,不同组合下模型最终的代价函数取值如附录 A 图 A7 所示。图中: r_s 为字典中原子总数占样本总数的比例。由图可见,当字典规模和层数增加到一定程度时,模型性能趋于稳定,此时继续增加字典规模和层数对模型性能的提升作用不大。因此本文最终选定字典层数为 5 层,即 $s=5$,字典中原子总数占样本总数的比例为 0.3,则每类别每层字典中有 31 个字典原子 ($5 \times 280 \times 0.3 / (10 \times 5) = 31$)。

3.3 分类效果及鲁棒性验证

由式(13)可知,根据待分类样本 $\mathbf{y}$ 在深度字典 $\mathbf{D}$ 中的分解系数对样本 $\mathbf{y}$ 按类进行重构,重构误差最小的类别就是本文方法判别的 $\mathbf{y}$ 所属的类别。因此,对比样本 $\mathbf{y}$ 在各个类别下的重构误差也可有效衡量方法的有效性,即样本 $\mathbf{y}$ 实际所属类别的重构误差越小、非所属类别的重构误差越大,则方法的性能越好。定义 2 个评价指标 e_{ii} 和 $e_{ik} (i \neq k)$ 以定量描述这一特性,如式(16)所示。

$$\left\{ \begin{array}{l} e_{ii} = \frac{\left\| \mathbf{y}^i - \sum_{j=1}^s \mathbf{D}_i^j \mathbf{x}_i^j \right\|_2^2}{\left\| \mathbf{y}^i \right\|_2^2} \\ e_{ik} = \min \frac{\left\| \mathbf{y}^i - \sum_{j=1}^s \mathbf{D}_k^j \mathbf{x}_k^j \right\|_2^2}{\left\| \mathbf{y}^i \right\|_2^2} \quad k \neq i \end{array} \right. \quad (16)$$

式中: $\mathbf{y}^i$ 表示样本 $\mathbf{y}$ 实际属于类别 i ; e_{ii} 为第 i 类字典的重构误差; e_{ik} 为其他类字典重构误差的最小值。若 e_{ii} 越接近 0、 e_{ik} 越接近 1,则表明样本 $\mathbf{y}$ 能够越好地用其实际所属类别的字典原子进行表达,分类效果越好。

重构误差指标 e_{ii} 和 e_{ik} 反映了单个样本的分类效果,进一步地,定义如式(17)所示的重构误差均值指标 $\bar{e}_{ii}$ 和 $\bar{e}_{ik}$ 来衡量所提方法在样本集上的整体表现。

$$\left\{ \begin{array}{l} \bar{e}_{ii} = \frac{1}{N} \sum e_{ii} \\ \bar{e}_{ik} = \frac{1}{N} \sum e_{ik} \end{array} \right. \quad (17)$$

基于 2.3 节给出的故障分类算法,利用故障样本库 $\mathbf{Y}$ 验证深度故障字典 $\mathbf{D}$ 的分类效果,结果表明算法能准确识别出全部样本所属的类别,分类正确率为 100%。同时,基于式(17)中的重构误差均值指标,得到本文所提方法在识别样本库 $\mathbf{Y}$ 中各类故障的性能表现如图 3 所示。

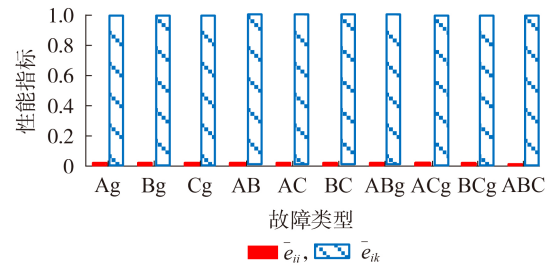


图 3 重构误差均值指标对比

Fig.3 Comparison of average indicators of reconstruction error

从图 3 可见,所提方法在保证正确分类的前提下,重构误差均值指标 $\bar{e}_{ii}$ 接近于 0、 $\bar{e}_{ik}$ 接近于 1,表明样本集中的样本可以很好得由其所属类别的字典原子表达,证明了所学习到的深度字典的有效性。同时,0 和 1 的对比表明所提分类指标具有较高的容错裕度和鲁棒性。

考虑到实际运行环境中存在噪声,本文对所提方法的抗噪性能进行了验证,利用 MATLAB 对样本库 $\mathbf{Y}$ 中全部 5280 份故障样本添加高斯白噪声,然后利用深度故障字典 $\mathbf{D}$ 对含有噪声的故障样本进行分类,不同噪声强度下本文所提方法和文献[10]所提方法的分类正确率和所有样本的重构误差均值指标对比结果如表 2 所示。由表可见,本文所提模型在各种噪声强度下均可以实现正确分类,且对样本的

表 2 2 种方法的抗噪性能对比

Table 2 Comparison of anti-noise performance between two methods

信噪比 / dB	方法	准确率 / %	$\bar{e}_{ii}$	$\bar{e}_{ik}$
20	本文	100	0.0353	0.9957
	文献[10]	100	0.3946	0.9128
18	本文	100	0.0417	0.9947
	文献[10]	100	0.3998	0.9111
16	本文	100	0.0528	0.9934
	文献[10]	99.98	0.3983	0.9113
14	本文	100	0.0598	0.9919
	文献[10]	99.94	0.3992	0.9109
12	本文	100	0.0702	0.9900
	文献[10]	99.78	0.4092	0.9076
10	本文	100	0.0826	0.9880
	文献[10]	99.09	0.4101	0.9064

重构误差均值受噪声强度的影响很小, $\bar{e}_{ii}$ 和 $\bar{e}_{ik}$ 之间差异显著, 证明了该模型具有较好的抗噪性能。而文献[10]所提方法的分类正确率随着噪声强度的增大而减小, 且重构误差均值指标 $\bar{e}_{ii}$ 和 $\bar{e}_{ik}$ 之间的差异显著小于本文所提方法。

在 A 相接地故障样本中添加 20 dB 噪声后, 本文和文献[10]所提方法的重构误差、分类效果对比如附录 A 图 A8 所示。由图可见, 本文和文献[10]所提方法均可正确分类, 但是本文所提方法的重构误差指标 e_{ii} 接近于 0, 明显小于文献[10]所提方法的指标值, 这表明本文所提方法的重构效果更好。文献[10]在噪声强度为 16 dB 时出现分类错误的情况如附录 A 图 A9 所示。由图可见, 根据式(13)所示的分类方法, 文献[10]所提方法将该 A 相接地故障样本错误分类为 AB 两相短路故障样本, 产生了明显的重构误差, 而本文所提方法可以进行正确分类。

3.4 泛化性能验证

为了验证本文所提方法的泛化性能, 本文直接使用基于图 2 的双端电源模型故障集的深度故障字典在附录 A 图 A10 所示的标准 IEEE 9 节点模型上进行验证。利用验证模型在 3 个不同位置进行了各种不同故障条件的仿真, 共计得到 2520 份故障仿真数据, 基于本文所提方法和文献[10]所提方法的分类结果及重构误差均值对比如表 3 所示。

表 3 2 种方法的泛化性能对比

Table 3 Comparison of generalization performance between two methods

故障位置	方法	准确率 / %	$\bar{e}_{ii}$	$\bar{e}_{ik}$
F_1	本文	100	0.1954	0.9673
	文献[10]	95.12	0.4744	0.8879
F_2	本文	100	0.1353	0.9771
	文献[10]	96.43	0.4249	0.8918
F_3	本文	100	0.1220	0.9783
	文献[10]	98.95	0.4039	0.9054

从表 3 可知, 本文所提基于双端电源系统的深度故障字典模型在标准 IEEE 9 节点模型的故障分类任务中取得了 100% 的正确率, 高于文献[10]所提方法的正确率, 表明本文所提深度字典模型的泛化性能要优于单层字典模型。

图 4 为本文所提方法在 e_{ii} 最大时的样本分类效果。由图可见, 即使在最坏的情况下, 样本实际所属类别字典的重构误差与其他类别字典的重构误差之间仍有显著区别, 本文所提方法仍可准确分类且分类结果具有较高的可靠性, 将基于双端电源系统的深度故障字典直接用于其他输电系统中线路故障分类任务的可行性证明了所提算法的泛化性能以及将其应用于实际工程的潜力。

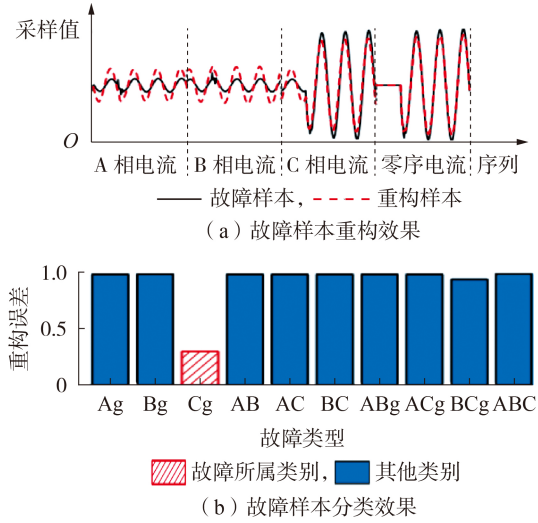


图 4 标准 IEEE 9 节点故障集中, 本文所提方法在 e_{ii} 最大时的样本分类效果

Fig.4 Classification effect of sample with largest reconstruction error in standard IEEE 9-bus fault set under proposed method

3.5 故障特征的可解释性

如前文所述, 深度学习等基于数据驱动的特征学习算法可解释性差, 是本文想要解决的主要问题之一。即本文欲构建一个基于数据的特征学习模型, 且所学习到的特征具有较好的物理含义。

高压输电线路故障后的暂态电流可表示为:

$$i_f = i_{\text{pow_AC}} + i_{\text{damp_DC}} + i_{\text{high}} \quad (18)$$

式中: i_f 为故障后暂态电流; $i_{\text{pow_AC}}$ 为工频电流分量; $i_{\text{damp_DC}}$ 为衰减直流分量; i_{high} 为其他因素产生的高频分量。 $i_{\text{pow_AC}}$ 和 $i_{\text{damp_DC}}$ 可视为故障电流中的低频分量, i_{high} 为故障电流中的高频分量。

图 5 为某故障样本及其基于深度故障字典的稀疏表示结果。图中: S_f 为故障样本; $L_j (j=1, 2, \dots, 5)$ 表示该故障样本在第 j 层对应的故障字典原子。由图 5 可见, 低层字典原子主要反映了故障样本的概貌特征, 即式(18)中的低频分量, 高层字典原子主要揭示了故障时刻的突变量等细节特征, 即式(18)中

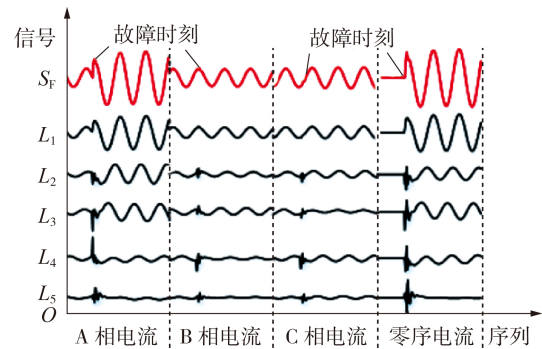


图 5 待分类样本及其稀疏表示字典原子

Fig.5 Sample to be classified and its dictionary atoms of sparse representation

的高频分量,且随着字典层数的增加,突变量等细节信息更为凸显。这种从由概貌到细节的层次化特征提取方式,符合人的认知过程,同时与实际输电线路故障信号的构成一致,具有较好的物理意义。

进一步地,本文利用 t-SNE 算法^[23]对学习到的深度故障字典逐层进行降维可视化,结果如附录 A 图 A11 所示。t-SNE 算法的基本思想是使高维空间相似的数据在低维空间(一般为 2 维或 3 维)的映射尽量接近,而差异较大的数据在低维空间的映射尽量远离。通过可视化降维后数据点在低维空间的分布,即可对原高维空间数据的相似程度有直观的感受。图 A11 所示的 t-SNE 可视化结果表明,各子层中同属一类的字典原子在降维后表现出明显的聚类特征,而不同类别原子映射到低维空间后距离较远。根据 t-SNE 算法的特点可知,高维空间的同类故障字典原子是相似的,而不同故障类别的字典原子差异是显著的,这表明深度字典模型有效提取了不同故障类型之间的差异信息,且保持了同类故障特征的一致性。同时,结合式(18)和图 5 可知,低层字典主要提取了故障电流中的低频故障特征,由于低频成分中包含一部分故障前的工频分量,这一分量是不受故障类型影响的,因此相当于低频故障特征中存在一定的“共模量”,具体表现为图 A11 中低层字典的类间差异相对较小。随着字典层数增加,高层字典主要提取故障电流中的高频故障特征,避免了工频分量的影响,因此不同类间的故障特征差异更加显著,具体表现为图 A11 中高层字典类间距离逐渐增大。上述可视化结果分析表明,深度故障字典有效提取了不同故障类型样本之间的从低频到高频部分的差异信息。

4 结论

本文基于稀疏表示算法同时借鉴深度学习的思想,提出了一种用于输电线路故障分类的深度字典学习方法,通过理论分析与仿真验证得到如下结论:

1) 深度字典学习模型逐层自动提取故障特征,所提取的故障特征具有较好的物理含义,低层字典包含故障的总体特征,高层字典揭示故障的细节特征,一定程度上解决了故障特征的可解释性问题;

2) 双端模型中学习到的深度故障字典在标准 IEEE 9 节点模型上的分类效果优异,说明该方法学习到的是故障的普遍特征,而非仅针对特定模型,具有较好的泛化能力,具有实际的应用潜力和价值。

相比现有的故障分类识别算法,本文提出的数据驱动的字典学习方法能充分利用电力系统自身的海量故障数据提取故障特征,自动生成辨识判据,大幅降低了技术实现难度。本文方法可用于电力输电线路故障的事后自动化分析,有助于故障准确测距

和故障的快速恢复。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 程志友,李小静,汤明金,等. 一种故障分量选相元件接地短路判据的修正[J]. 电力系统保护与控制,2018,46(16):19-24.
CHENG Zhiyou, LI Xiaoqing, TANG Mingjin, et al. A correction of ground short-circuit criterion of phase selection element based on faulty components[J]. Power System Protection and Control, 2018, 46(16): 19-24.
- [2] 张海,黄少锋. 利用电压辅助电流选相的同杆双回线单端电气量选相原理[J]. 中国电机工程学报,2013,33(7):139-148.
ZHANG Hai, HUANG Shaofeng. A fault phase selection scheme of currents with assistant voltages for common-tower double-circuit transmission lines using one-terminal electrical quantities[J]. Proceedings of the CSEE, 2013, 33(7): 139-148.
- [3] 林圣,梅俊涛,陈双,等. 基于暂态行波时频特征的输电线路故障检测与选相方法[J]. 电网技术,2012,36(7):48-53.
LIN Sheng, MEI Juntao, CHEN Shuang, et al. Fault detection and faulty phase determination of transmission lines based on time-frequency characteristics of transient travelling waves[J]. Power System Technology, 2012, 36(7): 48-53.
- [4] 费春国,霍洪双. 基于电流振幅与 SVM 的输电线路故障分类[J]. 电力系统及其自动化学报,2019,31(4):139-144.
FEI Chunguo, HUO Hongshuang. Fault classification for transmission lines based on current amplitude and SVM[J]. Proceedings of the CSU-EPSA, 2019, 31(4): 139-144.
- [5] 李达,薛卿,孔德健,等. 基于 PSO-BP 算法的高压输电线路故障分类[J]. 电气自动化,2018,40(6):42-44.
LI Da, XUE Qing, KONG Dejian, et al. Fault classification of high voltage transmission line based on PSO-BP algorithm[J]. Electrical Automation, 2018, 40(6): 42-44.
- [6] CHEN Kunjin, HU Jun, HE Jinliang. Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder[J]. IEEE Transactions on Smart Grid, 2018, 9(3): 1748-1758.
- [7] 王浩,杨东升,周博文,等. 基于并联卷积神经网络的多端直流输电线路故障诊断[J]. 电力系统自动化,2020,44(12):84-92.
WANG Hao, YANG Dongsheng, ZHOU Bowen, et al. Fault diagnosis of multi-terminal HVDC transmission line based on parallel convolutional neural network[J]. Automation of Electric Power Systems, 2020, 44(12): 84-92.
- [8] 殷浩然,苗世洪,郭舒毓,等. 基于 S 变换相关度和深度学习的配电网单相接地故障选线新方法[J]. 电力自动化设备,2021,41(7):88-96.
YIN Haoran, MIAO Shihong, GUO Shuyun, et al. Novel method for single-phase grounding fault line selection in distribution network based on S-transform correlation and deep learning[J]. Electric Power Automation Equipment, 2021, 41(7): 88-96.
- [9] 程文傲,徐明,高金峰. 小电流接地系统单相接地故障选线空间域图像生成及融合方法[J]. 电力自动化设备,2021,41(7):97-103.
CHENG Wenao, XU Ming, GAO Jinfeng. Spatial domain image generation and fusion method of single-phase grounding fault line selection for small current grounding system[J]. Electric Power Automation Equipment, 2021, 41(7): 97-103.
- [10] SHI Shenxing, ZHU Beier, MIRSAEIDI S, et al. Fault classification for transmission lines based on group sparse representation[J]. IEEE Transactions on Smart Grid, 2019, 10(4): 4673-4682.
- [11] 和敬涵,罗国敏,程梦晓,等. 新一代人工智能在电力系统故障

- 分析及定位中的研究综述[J]. 中国电机工程学报, 2020, 40(17):5506-5516.
- HE Jinghan, LUO Guomin, CHENG Mengxiao, et al. A research review on application of artificial intelligence in power system fault analysis and location[J]. Proceedings of the CSEE, 2020, 40(17):5506-5516.
- [12] SHRESTHA A, MAHMOOD A. Review of deep learning algorithms and architectures[J]. IEEE Access, 2019, 7:53040-53065.
- [13] 练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展[J]. 自动化学报, 2015, 41(2):240-260.
- LIAN Qiushen, SHI Baoshun, CHEN Shuzhen. Research advances on dictionary learning models, algorithms and applications[J]. Acta Automatica Sinica, 2015, 41(2):240-260.
- [14] SALLEE P, OLSHAUSEN B A. Learning sparse multiscale image representations[J]. Advances in Neural Information Processing Systems, 2002, 15:1351-1358.
- [15] NATARAJAN B K. Sparse approximate solutions to linear systems[J]. SIAM Journal on Computing, 1995, 24(2):227-234.
- [16] PATI Y C, REZAIIFAR R, KRISHNAPRASAD P S. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition[C]//Proceedings of 27th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA: IEEE, 1993:40-44.
- [17] ENGAN K, AASE S O, HUSOY J H. Method of optimal directions for frame design[C]//IEEE International Conference on Acoustics, Speech, and Signal Processing. Phoenix, AZ, America: 1999:2443-2446.
- [18] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11):4311-4322.
- [19] SAHOO S K, MAKUR A. Dictionary training for sparse representation as generalization of K-means clustering[J]. IEEE Signal Processing Letters, 2013, 20(6):587-590.
- [20] USMAN K. Introduction to orthogonal matching pursuit[EB/OL]. (2017-08-31)[2021-08-10]. <http://korediantousman.staff.telkomuniversity.ac.id>.
- [21] 张保会, 尹项根. 电力系统继电保护[M]. 北京: 中国电力出版社, 2010:10-11.
- [22] 陶莹, 杨锋, 刘洋, 等. K均值聚类算法的研究与优化[J]. 计算机技术与发展, 2018, 28(6):90-92.
- TAO Ying, YANG Feng, LIU Yang, et al. Research and optimization of K-means clustering algorithm[J]. Computer Technology and Development, 2018, 28(6):90-92.
- [23] MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9(86):2579-2605.

作者简介:



张宇博

张宇博(1997—),男,博士研究生,主要研究方向为人工智能在新能源电力系统中的应用(E-mail:zyb970305@stu.xjtu.edu.cn);

郝治国(1976—),男,教授,博士,主要研究方向为电力系统继电保护、新能源接入电力系统(E-mail:zhghao@mail.xjtu.edu.cn);

林泽暄(1998—),男,硕士研究生,主要研究方向为人工智能在电力系统故障诊断中的应用;

杨松浩(1989—),男,助理教授,博士,研究方向为电力系统暂态稳定分析与控制(E-mail:songhaoyang@xjtu.edu.cn)。

(编辑 任思思)

Transmission line fault classification method based on deep dictionary learning

ZHANG Yubo¹, HAO Zhiguo¹, LIN Zexuan¹, YANG Songhao¹, LIU Zhiyuan², YU Xiaojun²

(1. School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. State Grid Ningxia Electric Power Co., Ltd., Yinchuan 750001, China)

Abstract: In order to solve the problems of complex threshold setting and insufficient interpretability of artificial intelligence algorithms in current transmission line fault classification and recognition methods, a transmission line fault classification method based on deep dictionary learning is proposed. Driven by the constraint of sparsity, the fault features of the samples are automatically extracted. Meanwhile, the structure of a deep dictionary enables the extracted fault features to have definite hierarchy and physical meaning, which is consistent with human intuitive understanding. The proposed method solves the problem of insufficient interpretability of data-driven methods to a certain extent. Finally, the effectiveness of the proposed method is verified by the simulation in PSCAD/EMTDC.

Key words: transmission line; fault classification; sparse representation; deep dictionary learning

附录 A

输入：字典 D ，样本矩阵 Y ，稀疏性指标 s
输出：系数矩阵 X
流程：对样本矩阵中的每个样本 y_i 均执行如下操作
 1) 初始化误差向量 $r^{(0)} = y_i$ ，索引集 $A^{(0)} = O$ ， $k = 1$
 2) 求解字典 D 中与残差 $r^{(k-1)}$ 最相关的字典原子序号

$$\lambda = \arg \max_j \left| \left\langle r^{(k-1)}, d_j \right\rangle \right|, j = 1, 2, \dots, M$$

 3) 更新索引集 $A^{(k)}(k) = \lambda$ ，更新选中的字典原子构成的集合

$$D^{(k)} = D(:, A^{(k)})$$

 4) 利用最小二乘法求解有效系数向量

$$\tilde{x}^{(k)} = \arg \min_{\tilde{x}} \|y_i - D^{(k)} \tilde{x}\|_2$$

 5) 更新残差 $r^{(k)} = y_i - D^{(k)} \tilde{x}^{(k)}$ ， $k = k + 1$
 6) 如果 $k > s$ ，令 $X(A^{(s)}, i) = \tilde{x}^{(s)}$ ，结束；否则，返回步骤2)

图 A1 OMP 算法实现

Fig.A1 Implementation of OMP algorithm

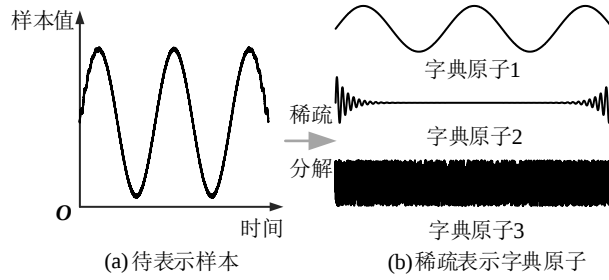
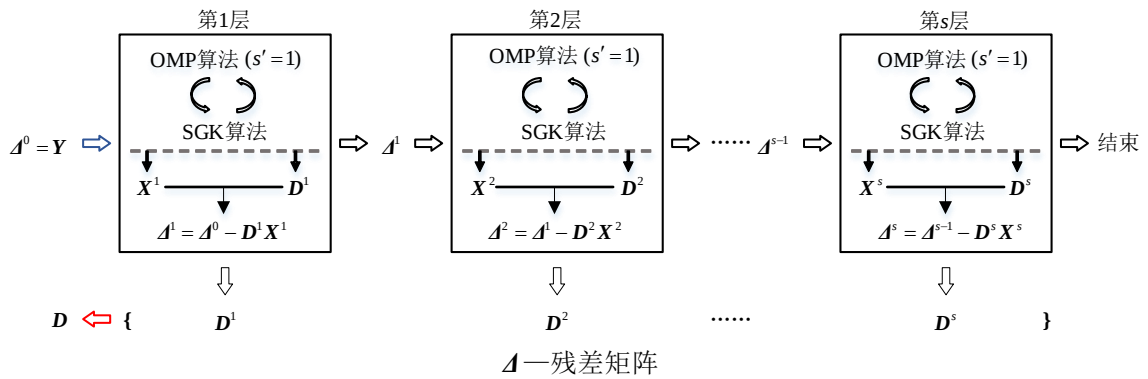


图 A2 $s=3$ 时学习字典模型示意图

Fig.A2 Schematic diagram of learning dictionary model when $s = 3$



A —残差矩阵

图 A3 深度字典学习模型原理图

Fig.A3 Schematic diagram of deep dictionary learning model

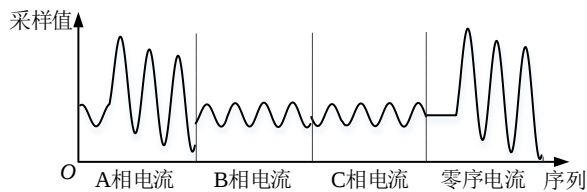


图 A4 某归一化 A 相接地故障样本

Fig.A4 Normalized sample of phase-A grounding fault

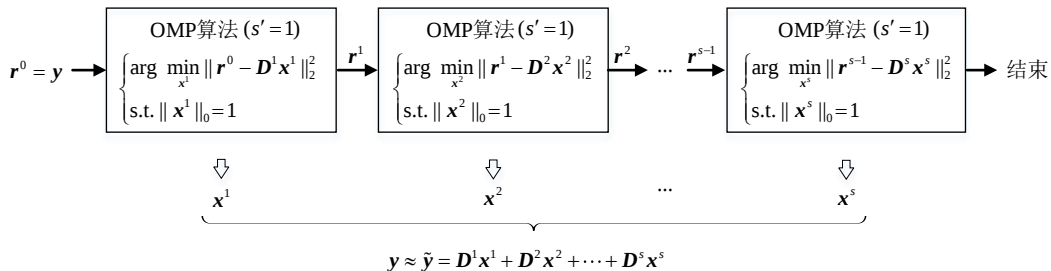


图 A5 待分类样本的深度稀疏表示示意图

Fig.A5 Schematic diagram of deep sparse representation of sample to be classified

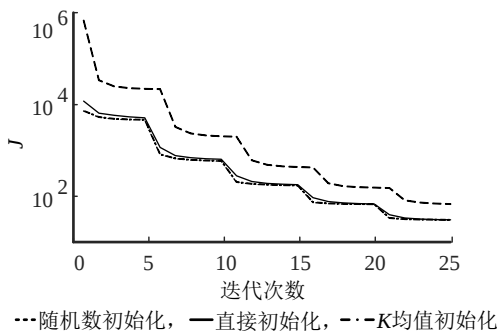


图 A6 不同初始化方法对模型性能的影响

Fig.A6 Influence of different initialization methods on model performance

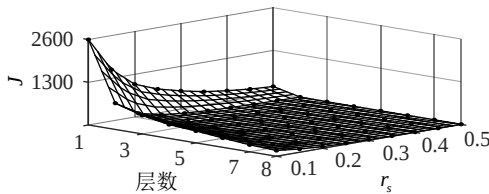


图 A7 字典规模及层数对模型性能的影响

Fig.A7 Impact of dictionary size and number of layers on model performance

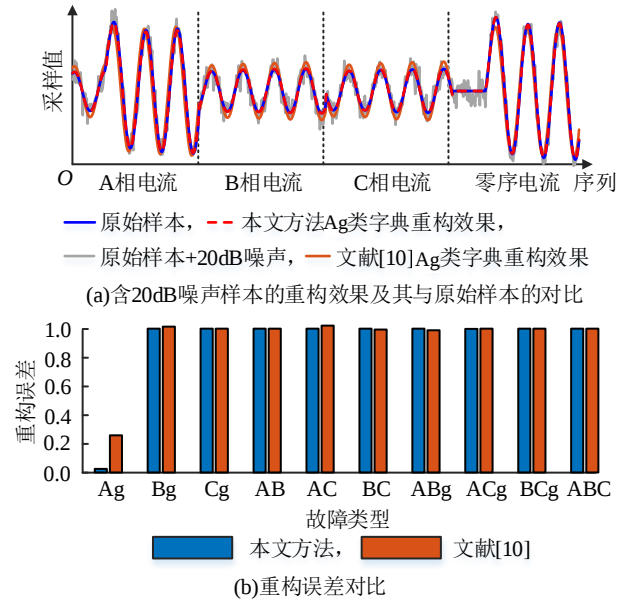


图 A8 某含 20dB 噪声样本的分类和重构效果对比

Fig.A8 Classification and reconstruction effect of sample with 20dB noise

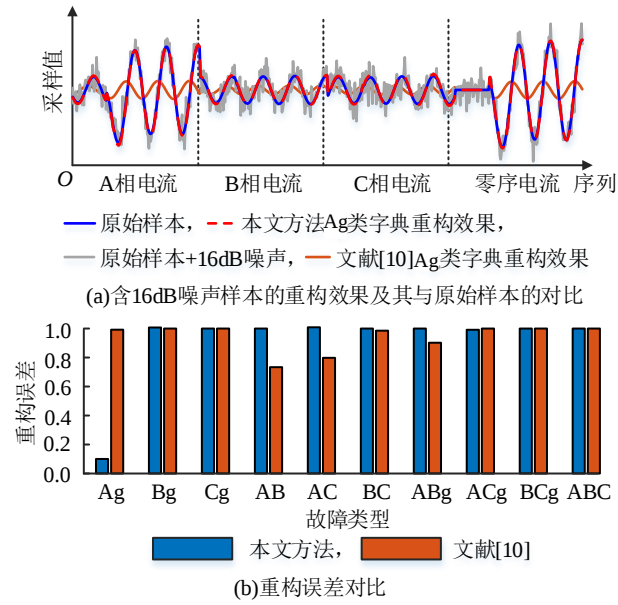


图 A9 某含 16dB 噪声样本的分类和重构效果对比

Fig.A9 Classification and reconstruction effect of sample with 16dB noise

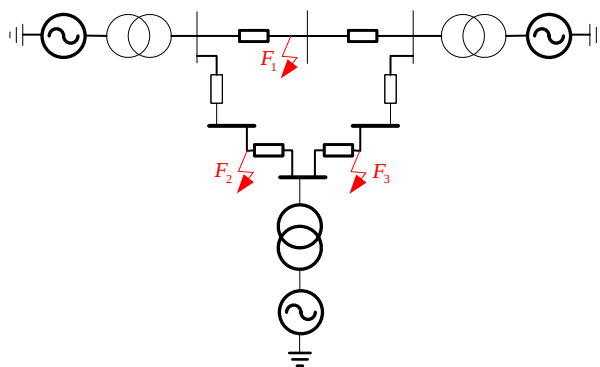


图 A10 IEEE 标准 9 节点故障仿真模型

Fig.A10 Fault simulation model based on standard IEEE 9-bus

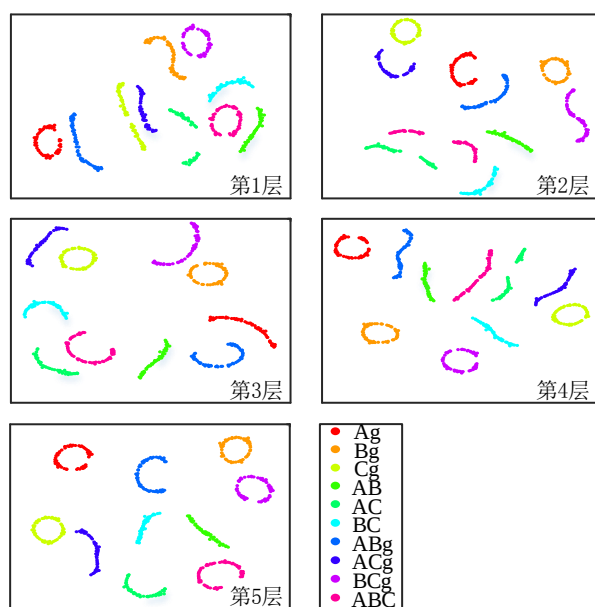


图 A11 各层字典的 t-SNE 可视化结果

Fig.A11 Visualization results of dictionary of each level based on t-SNE algorithm