第 13 卷 第 5 期
2024 年 9 月
集 成 技 术
JOURNAL OF INTEGRATION TECHNOLOGY
Vol. 13 No. 5
Sep. 2024

引文格式：
冯天任, 陈世峰. 基于深度度量学习的强泛化开关仪表识别算法 [J]. 集成技术, 2024, 13(5): 30-39.
Feng TR, Chen SF. Strong generalization switchgear instrument recognition algorithm based on deep metric learning [J]. Journal of Integration Technology, 2024, 13(5): 30-39.

基于深度度量学习的强泛化开关仪表识别算法

冯天任[1,2] 陈世峰[2*]

[1](中国原子能科学研究院 北京 102413)
[2](中国科学院深圳先进技术研究院 深圳 518055)

摘 要 针对电厂开关检测方法无法应对现实开集环境，对稀有类别识别准确率低的现状，将目标识别问题转化为相似性度量问题，并提出新算法。新算法基于深度度量学习的三元组网络，利用加入 SE Block 的 ResNet-18 提取特征，并利用跨批次挖掘增强学习效果。为评估算法性能，创建了一个包含 3 300 张开关图片的数据集，并使用新算法在该数据集上进行了闭集测试、开集测试、小样本测试。结果表明：新算法在闭集状态下具有良好的区分能力，不仅能准确识别训练集中的类别，还能有效区分训练时未遇到的及出现频率较低的状态。由此表明，该算法不仅适用于现实世界的开集环境，而且能显著提升对小样本数据的识别精度。

关键词 深度度量学习；三元组网络；注意力机制；开关状态识别
中图分类号 TP399 **文献标志码** A doi: 10.12146/j.issn.2095-3135.20240205001

Strong Generalization Switchgear Instrument Recognition Algorithm Based on Deep Metric Learning

FENG Tianren[1,2] CHEN Shifeng[2*]

[1](*China Institute of Atomic Energy*, *Beijing* 102413, *China*)
[2](*Shenzhen Institute of Advanced Technology*, *Chinese Academy of Sciences*, *Shenzhen* 518055, *China*)
[*]*Corresponding Author*: *shifeng.chen@siat.ac.cn*

Abstract In response to the current power plant switch detection methods that are unable to cope with real-world open-set environments and the low accuracy in recognizing rare categories, the target recognition problem is transformed into a similarity measurement issue, and a new algorithm is proposed. The new algorithm is based on the triplet network of deep metric learning, using a ResNet-18 with an added SE Block to extract features, and enhances learning effects by cross-batch mining. To evaluate the performance of the algorithm,

收稿日期：2024-02-05 **修回日期**：2024-03-28
基金项目：深圳市科技创新委员会基础研究重点项目(JCYJ20200109114835623)；深圳市科技创新委员会技术攻关重点项目(JSGG20220831105002004)
作者简介：冯天任，硕士研究生，研究方向为计算机视觉；陈世峰(通讯作者)，副研究员，研究方向为计算机视觉，E-mail：shifeng.chen@siat.ac.cn。

a dataset with 3 300 switch images was created. The algorithm was tested on the self-built dataset for closed-set testing, open-set testing, and few-shot testing. The experimental results show that the algorithm demonstrates excellent discrimination ability in the closed-set state. It can not only accurately identify the categories in the training set but also effectively distinguish states that were not encountered during training and those with lower occurrence frequencies. This capability indicates that the algorithm is not only suitable for real-world open-set environments but also significantly improves the recognition accuracy for small-sample data.

Keywords　deep metric learning; triplet network; attention mechanism; switch state recognition

Funding　This work is supported by Shenzhen Science and Technology Innovation Commission (JCYJ20200109114835623, JSGG20220831105002004)

1　引　言

开关仪表是工业控制系统中的重要组成部分，早期采用人工巡检的方式监测其状态变化，并进行记录，大多数指标(如表计读数、指示灯状态、开关及旋钮状态)完全由操作员现场记录。数字化改造推进了电厂设备更新，引入了数字化仪表系统，使得数据可直接传输至控制室，操作员可进行远程监控，显著降低了人工的工作量。但当前的数字化改造无法覆盖所有的边缘设备开关，因此，为保证系统运行的可靠性和安全性，目前仍需人工进行场地巡检，以验证开关和断路器等设备的状态。然而，电厂场地设施复杂，很多元件并不方便人工巡检。从安全性、工作效率及可靠性的角度来讲，使用计算机视觉技术辅助人工巡检将是一种更好的选择。目前主流的开关识别检测方案主要采用基于深度学习的目标检测算法，包括多阶段检测算法 Faster R-CNN[1]及单阶段检测算法 YOLO[2]。刘超等[3]在 YOLOv4 的基础上加入通道注意力，以辅助识别变电站断路器的分合状态。史梦安等[4]在 YOLO-tiny 模型中加入 Receptive Field Block 模块，以进行多尺度特征融合，增强对旋钮状态的泛化识别能力。Fu 等[5]对 Faster R-CNN 进行改进，将其中两个全连接层替换为卷积层，并加入 ReLU 层，使改进后的 Faster R-CNN 能有效识别任意尺寸的高压柜开关。针对获取图像清晰度不足的问题，Dong 等[6]、Chen 等[7]先使用 SRCNN 对图像进行超分辨率重建，再使用 YOLOv3 对元件进行识别。在识别训练已知类别的图像时，上述方法可取得良好的效果，但在实际的工业应用中，目标检测框架在进行状态识别时，普遍会遇到以下两个问题。

(1)开集问题。在实际环境中，故障或人员误操作可能会导致开关状态与常规类别存在较大差异。例如，开关可能被错误地卡在中间位置。此外，现实中可能会有新型号的开关仪表加入，而新型号开关的特征与已知类别不同。传统的目标检测算法假设训练阶段包含所有可能出现的类别，模型的主要任务是将输入的样本准确地分类到上述预定义的类别，但不处理这种开集场景，因此只能将新类别加入训练集，并重新训练模型。为适应上述环境，需要更适用于实际工业场景的强泛化性识别算法，不仅能识别已知类别，对未知类别也应有足够的判断能力。

(2)训练集数据不均衡。用于开关状态识别的训练数据往往直接在生产环境中收集。由于设备在大部分时间都属于正常运行状态，因此，正常开关状态容易采集。相对地，异常开关状态往往出现在发生故障或突发情况时，由于这些情况

的发生频率较低，因此，异常开关样本的收集数量有限，会导致训练时不同类别的样本数量不平衡。不平衡的数据分布会使模型对稀有状态识别的准确性降低，因此，缓解样本数据分布不均衡对提升开关状态识别算法泛化性至关重要。采用数据增强方法可在一定程度上缓解训练集不平衡的问题，但现实中的异常状态往往无法通过数据增强方法获取。

针对上述情况，本文转换思路，将场景定义从分类识别问题转化为相似性度量问题。后者侧重于分析和比较样本之间的相似性，以此判断它们是否属于同一类别。与传统分类算法不同，相似性度量的适用范围不局限于训练阶段所接触到的类别。常用的度量学习方式包括对比学习和三元组网络[8]。在数据量有限的情况下，三元组网络可更快、更有效地学习类内及类间关系，因此，本文利用深度度量学习中的三元组网络(triplet network)，通过最小化同类距离，并最大化不同类别的距离，使被训练出的模型具有辨识开关状态之间的相似程度的能力。为进一步提升模型的泛化能力，并缓解数据不均衡带来的问题，本文选择 SE-ResNet[9]作为特征抽取网络，并对模型使用 L2 正则化。一个合适的特征抽取网络能有效捕捉开关状态的关键特征，而 L2 正则化有助于阻止模型过拟合，从而在面对新颖或者稀有的样本类别时，保持其识别的准确性和鲁棒性。通过上述设计，可实现一个具有强泛化性的开关仪表识别算法。

2 相关工作

2.1 开集识别

在传统的识别任务中，通常假设所有可能遇到的图像类别在训练阶段都是已知的，模型的设计目标是将输入样本正确分类到已知的类别中。然而，实际应用中可能会遇到一些训练时未接触过的类别，要区分已知类别，并排除未知类别，即为开集识别问题。

在面对开集图像时，传统分类算法与度量学习算法的区别如图 1 所示。在图 1a 的分类模型中，模型会从训练过程中出现过的类别中选择最相似的类别，这可能会导致面对新类别图像时分类错误。相比之下，右侧的相似性度量方法并不局限于已学过的类别，可以评估任何两幅图像之间的相似程度。

开集识别的目标是对已知类别的样本与未知

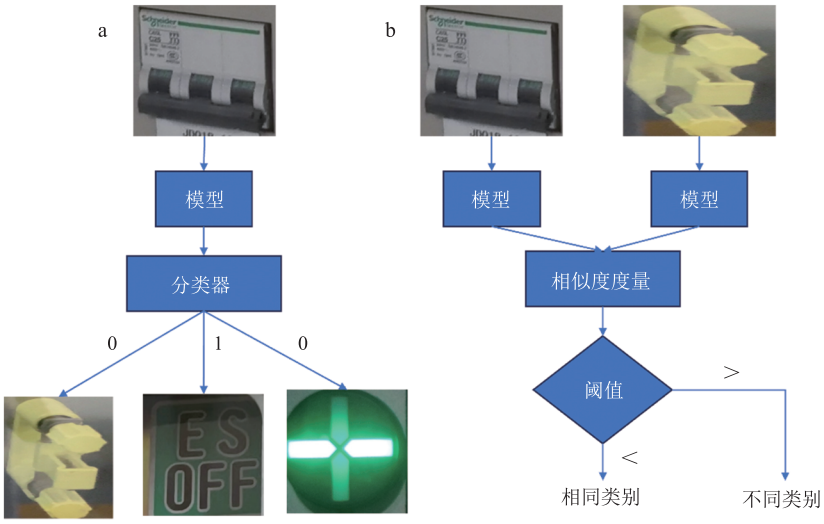


图 1 分类模型(a)和相似性度量模型(b)的区别

Fig. 1 The difference between classification models (a) and similarity measurement models (b)

类别的样本进行有效区分，关键挑战是在训练过程中建立一个可推广到未知类别的模型。一般解决方法是，建立一个代表已知类别分布的正常行为模型，并利用此模型鉴别那些与已知分布不一致的样本。常见的方法有两种，一种是利用深度度量学习检测未知类别的样本[10]，另一种是利用对抗网络[11–13]或变分自编码器[14–16]等生成模型进行无监督学习。

2.2　样本不均衡

在大多数情况下，电厂运行稳定，在进行数据采集时，得到的通常是正常运行状态下的监测数据。由于异常数据收集耗时长且成本高，因此，异常状态的数据相对较少，导致训练数据中的正常样本数量远大于异常样本数量。上述现象造成深度学习模型在识别异常状态时不稳定。针对该问题，已有不少研究，如调整权重的 Focal Loss[17]、gradient harmonizing mechanism[18]，以及基于多尺度度量学习的小样本学习[19]等。这些方法旨在改进深度学习模型对数据不平衡的适应能力，以提高对异常状态的识别稳定性和准确率。

3　方　法

3.1　三元组网络

新算法利用三元组网络[8]应对未知类别样本的识别问题，允许网络学习样本之间的相似性关系，并在特征空间中优化样本分布，使同类样本聚集在一起，异类样本远离，可使模型对不同类别的样本有良好的区分度，从而有效识别出未见过的样本状态。三元组网络的训练流程如图 2 所示，每次训练会选择 3 个图像输入到一个共享权重的模型中。其中，这 3 个图像分别为锚点图像、正样本图像及负样本图像。锚点图像和正样本图像属于同一类别，而负样本图像与锚点图像属于不同类别。通过模型可获得每个图像的特征向量，然后计算锚点图像和正样本图像之间的特征向量距离，以及锚点图像和负样本图像之间的特征向量距离。通过最小化同类样本距离、最大化异类样本距离，可让相同类别的样本聚集，不同类别的样本远离。

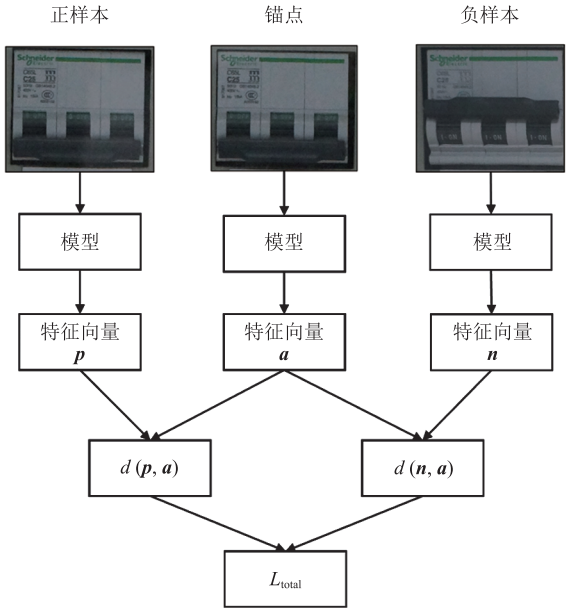


图 2　模型训练

Fig. 2　Training model

三元组网络的损失函数如式(1)所示。

$$L_{\text{triplet}}=\sum_{i=1}^{N}\max\left(0,d\left(f\left(x_i^{\text{anchor}}\right),f\left(x_i^{\text{pos}}\right)\right)-d\left(f\left(x_i^{\text{anchor}}\right),f\left(x_i^{\text{neg}}\right)\right)+\alpha\right) \tag{1}$$

其中，N 为三元组样本对的数量；x_i^{anchor} 为锚点(anchor)样本的输入；x_i^{neg} 为负样本的输入；x_i^{pos} 为正样本的输入；f 为特征提取模型，将输入映射到特征空间；d 为距离度量函数，如欧氏距离或余弦相似度；α 为边界值(margin)，用于控制正负类样本之间的距离。此外，为防止模型过拟合，增加了 L2 正则项，用以惩罚大的权重值。它通常会使得梯度变化相对平稳，有助于训练过程中的稳定性。L2 正则化如式(2)所示。

$$L_{\text{reg}}=\lambda\sum_{i=1}^{n}\left\|w_i\right\|_2^2 \tag{2}$$

其中，λ 为正则化系数；w_i 为模型中的权重参数；n 为权重参数的数量。因此，总损失函数为

$$L_{\text{total}}=L_{\text{triplet}}+L_{\text{reg}} \tag{3}$$

为缓解训练集中样本数量不均衡的问题，需选择恰当的三元组挖掘策略。合适的选择可提高模型性能，不当的挖掘策略则会让模型难以收敛。新模型采用难例样本挖掘(hard mining)，优先选择那些满足条件 $d(\boldsymbol{n},\boldsymbol{a})<d(\boldsymbol{p},\boldsymbol{a})$的三元组进行训练，即在特征空间中，负样本到锚点的距离小于正样本到锚点的距离。这种方法可更有效地提升模型区分易混淆样本的能力。此外，通过增加少数类别样本在训练中的抽样频率，可进一步提升模型对稀有类别的识别效果。

3.2 SE-ResNet

为三元组网络选取一个合适的特征提取网络很重要。一个具有较强表达能力的网络能稳定提取不同状态的特征。同时，还必须考虑现实中的硬件资源限制。残差网络(ResNet)是当前主流的特征提取网络之一[9]，通过引入残差结构，可有效避免梯度消失问题，使得深层网络的训练变得更加可行。ResNet-50 具有深层的网络结构，在处理复杂场景和大规模图像数据集时表现优秀。然而，在处理简单或小规模图像任务时，模型的高复杂度可能导致过拟合，并且会增加训练所需的时间和计算资源的消耗。因此本文选择模型更简单的 ResNet-18 作为特征提取网络。在防止过拟合，保持特征提取性能的同时，ResNet-18 可显著减少计算负担和资源需求，更适用于受资源限制的实际部署环境。此外，还引入 SE Block 模块[11]以增强 ResNet-18 的特征提取能力，不仅能让网络学习使用全局信息，而且能选择性地强调信息特征，抑制无用特征。SE Block 由压缩(squeeze)和激活(excitation)两部分组成：在压缩阶段，对特征图进行全局平均池化操作，以聚合全局空间信息；在激活阶段，将全局特征向量输入到一个全连接层中进行重新加权和激活，以生成调节特征通道的权重。通过这些权重与原始特征图的逐通道相乘操作，SE Block 可强化较为重要的特征，并抑制次要特征，从而获得增强的特征表示。将 SE Block 加入到 ResNet-18 后的模型整体结构如图 3 所示。

3.3 跨批次挖掘

在三元组网络的训练阶段需选取合适的样本对。常规做法是在一个批次的样本中寻找满足条件的样本对进行训练，样本对的数量严格受限于批次的大小。理论上，较大的批次可提供更多的样本对，模型的性能也会更好。但是受限于计算资源，大部分情况下无法使用大批次进行训练。为解决这个问题，Wang 等[20]提出跨批次挖掘方法(cross-batch memory，XBM)，利用训练过程中样本特征的慢漂移现象扩大训练样本选取范围。在训练初期，神经网络对同一样本的特征提取可能发生剧烈变化，但随着模型训练逐渐稳定，这些变化会逐渐稳定。因此，在训练较为稳定时，将连续多个批次的输出组合起来，并结合当前批次样本，在不增加批次大小的情况下扩展

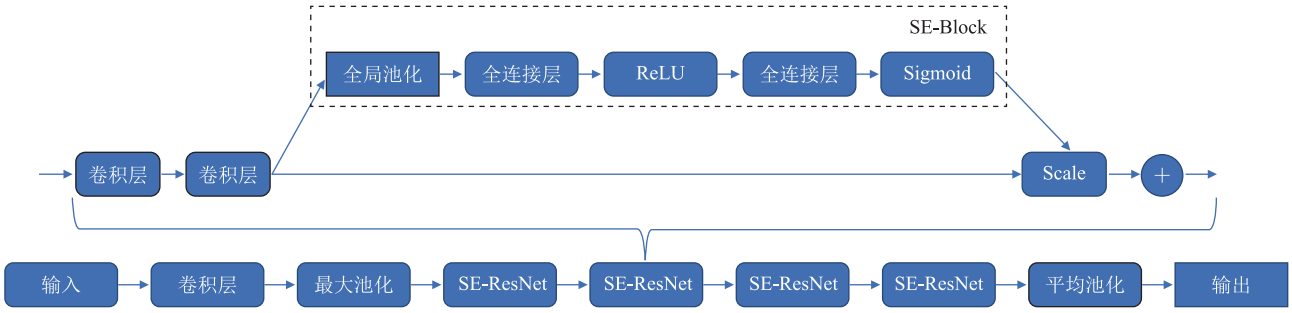


图 3 网络整体架构

Fig. 3 The overall architecture of the network

样本对的搜索范围，以提高训练效率。然而，XBM 的过早引入或批次与 XBM 队列的大小比例不合适可能会导致特征漂移的稳定性不足，甚至可能导致训练产生相反效果，因此，需要通过计算特征偏移判断使用 XBM 的时机。令输入为 x，迭代步长为 dt，在第 t 次迭代时的特征漂移定义如式(4)所示。

$$D(x,t;\Delta t):=\left\|f\left(x;\theta^{t}\right)-f\left(x;\theta^{t-\Delta t}\right)\right\|_2^2 \tag{4}$$

其中，θ 为 f 的参数。模型运行时，在训练的每个周期都会从每个类别中固定抽取一定数量的样本，用于计算特征漂移。同时设置一个超参数阈值，只有当所有类别中的特征偏移都小于该阈值时，才利用 XBM 扩大样本对的搜索范围。这种策略可确保在引入 XBM 时，样本特征在不同批次间具有足够的稳定性，避免因特征漂移导致的负面影响。

4　实　验

4.1　数据集

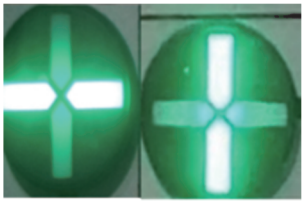
本文中，将显示“开/关”状态或其他状态信息的设备归类为开关类仪表。为验证算法性能，构建了一个包含 3 300 张开关类仪表图像的数据集，图像分辨率从 50×50 像素到 300×300 像素不等，如图 4 所示。数据集中的图像来源包括两部分，一部分是从互联网上收集的 2 000 张图像，涵盖不同环境下多种类型开关。另一部分是在电厂的实际生产环境中采集到的 1 300 张图像。整个数据集中的图像被分为 13 个大的类别，其中，普通开关、压板开关及分合指示灯的各种状态如图 5 所示。

图 4　数据集中全部开关种类展示

Fig. 4　Display all types of switches in the dataset

(a) 普通开关

(b) 压板开关



(c) 分合指示灯

图 5　几种开关的状态展示

Fig. 5　Status display of several types of switches

4.2　评价指标

为综合评价模型性能，选取评价指标时需考虑 3 个方面。

(1) 训练时出现类别的识别准确率。

(2) 训练时未出现类别的识别准确率。

(3) 模型的性能评价，即聚类效果。

首先，利用精度(precision)评判识别准确率。该指标反映在预测结果中排在第一位置的样本的准确度。精度 P 的定义如式(5)所示。

$$P=\frac{TP}{TP+FP} \tag{5}$$

其中，TP 为真阳性；FP 为假阳性。

其次，利用归一化互信息(normalized mutual information，NMI)及全类平均精度(mean average precision at R，mAP@R)评估模型聚类效果。mAP@R 在计算时引入距离作为权重，它结合了召回率(recall)的优点，并可以直接从特征嵌入空间进行计算，而不需要额外进行聚类。mAP@R 的定义如式(6)所示。

$$\text{mAP@R}=\frac{1}{R}\sum_{i=1}^{R}P(i) \tag{6}$$

其中，R 为与测试样本距离最近的样本数量；$P(i)$为第 i 个位置的精度。此外，由于两个不同种类仪表通常并不需要进行识别对比，因此还需要单独测试相同类别不同状态仪表的识别效果。图 6 至图 8 利用直方图表示样本距离，将特征向量归一化后计算欧氏距离。其中，蓝色为相同类别样本距离，橘黄色为不同类别样本距离。将模型的判别阈值设为 0.6，如果计算出的距离大于阈值，则表示识别为不同类别，反之则为同一类别。

4.3 训练细节

所有模型使用了相同的训练配置：主干模型是一个加入了 SE Block 的 ResNet-18 网络，并删去最后的全连接层，使其直接输出 512 维的特征向量。将训练的批量大小设为 128。为降低模型过拟合的风险，提高模型的泛化能力，在训练时，对所有模型使用相同的数据增强策略：首先将图片调整至 256×256 像素，然后从图片中心裁剪出 224×224 像素的区域。以 40% 的概率应用随机颜色调整(包括亮度、对比度、饱和度和色调变化)，30% 的概率进行随机仿射变换(包括少于 10° 的旋转、最多 5% 的平移和 5% 的缩放)，以及 30% 的概率做随机透视变换，之后对数据进行归一化处理。所有模型参数通过 Adam 优化器更新，初始学习率设为 1×10^{-4}。在三元组网络的损失函数中，将欧氏距离作为度量标准。在为提高模型表现而引入的 XBM 策略中，用于确定启动时间的特征漂移阈值设为 0.2。后续实验均使用同一个三元组网络模型进行测试。

5 结 果

5.1 已知类别识别测试

已知类别的识别测试即模型在正常情况下的识别能力。本文将多种损失函数进行比较，包括三元组网络损失(Triplet Loss)、CosFace[21]、Sub-center ArcFace[22]、Contrastive Loss[23]、ProxyNCA[24]、N-pair loss[25]。同时还选取 YOLOv5s 代表主流的目标检测框架。对每种模型均使用上述配置进行训练，结果如表 1 所示。可以看到，三元组网络在 Precision@1 及 NMI 指标上都取得了领先的分数，只有 mAP@R 比 CosFace 略低。识别准确率也高于主流方法 YOLOv5s。由于 YOLOv5s 为目标检测框架，因此无法直接使用 NMI 和 mAP@R 等聚类效果评价指标。

表 1 对已知类别的测试

Table 1 Known class recognition

损失函数/算法	Precision@1	NMI	mAP@R
Triplet Loss	92.30%	0.826	0.920
CosFace	92.10%	0.812	0.921
Sub-center ArcFace	89.20%	0.769	0.891
Contrastive Loss	85.40%	0.720	0.857
ProxyNCA	90.80%	0.829	0.897
N-pair loss	88.70%	0.772	0.854
YOLOv5s	92.15%	—	—

5.2　同类开关测试

由于正常情况下不需要将不同种类的开关放在一起进行识别，因此仅单独比较三元组网络模型对同种开关的识别效果。以普通开关为例，开关样式及状态如图 5(a)，将识别结果用样本之间距离的直方图表示，如图 6。可以看到，模型可将类内距离和类间距离完全分开，区分效果良好。

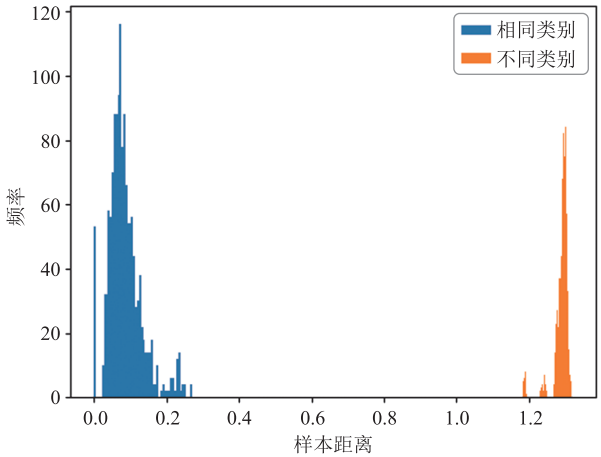


图 6　普通开关的距离直方图

Fig. 6　Histogram of distances for normal switches

5.3　开集测试

为验证模型的开集识别性能，用压板开关图像进行开集测试。数据集中的压板开关共有3种状态，其样式和状态见图 5(b)。选择前两种状态作为训练样本，最后一种用来模拟异常类别。计算样本之间的欧式距离，并将结果汇总成直方图进行分析，如图 7 所示。可以看出，距离分布较同类别测试结果更散乱，但使用 0.6 作为阈值仍然可以准确分开类内距离和类间距离，准确率为 100%。作为对比，YOLOv5s 会将新类别直接识别成已有的类别，且置信度不稳定，无法进行有效判断。

5.4　小样本测试

为检测三元组网络模型是否能应对训练样本不均衡的问题，选用分合指示灯进行小样本测试，指示灯样例见图 5(c)。分合指示灯共两个状态，每个状态只有 10 张图。在训练三元组模型时，在每个状态内随机选择 5 张进行训练，剩下 5 张进行测试，结果如图 8 所示。可以看到，当使用 0.6 作为距离阈值时，仍然可以准确分开类内距离和类间距离，识别准确率为 100%。作为对比，使用 YOLOv5s 进行同样的测试识别，准确率为 80%。

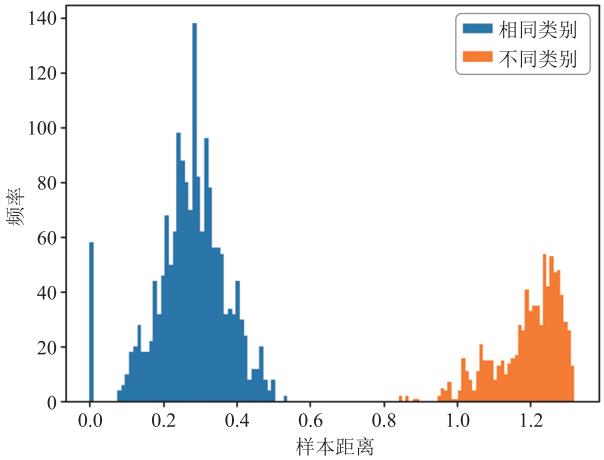


图 7　压板开关的距离直方图

Fig. 7　Histogram of distances for push-button switches

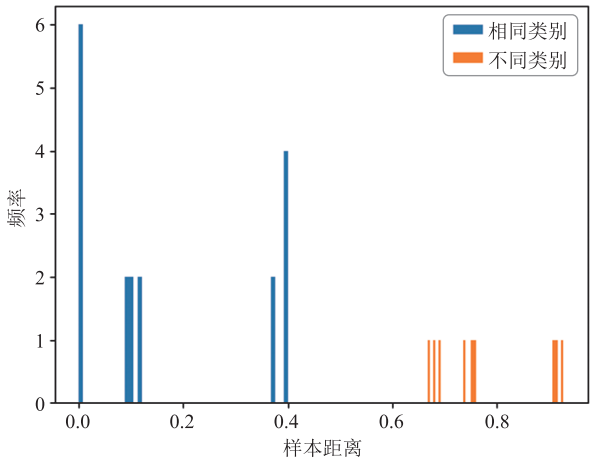


图 8　分合指示灯的距离直方图

Fig. 8　Histogram of distances for open/close indicator lights

6　讨论与分析

针对开关识别场景，为处理现实中存在的开集问题和样本数量不均衡问题，本文提出的新算

法将识别分类问题转化为相似性度量问题。通过计算样本之间的特征向量距离判断它们是否属于同一类别。新算法利用深度度量学习中的三元组网络，将 SE-ResNet 作为特征提取网络，利用 XBM 增加训练样本对。测试结果表明，新算法在已知类别和同类开关的识别测试中均表现出色，与其他度量学习方法及主流目标识别框架 YOLOv5s 相比，在识别精确度和聚类效果方面，也显示出一定优势。另外，开集和小样本的测试结果证实了新算法在未见过的状态及稀有状态下的区分能力，表明模型具有较强的泛化性，可应对开集环境，同时能在一定程度上缓解样本不均衡问题。目前，尽管数据规模和开关种类有限，通过扩充数据集有望进一步提高模型表现。

7 结 论

针对现实场景常见的开集问题及稀有状态的识别准确度问题，本文提出的新算法将状态识别问题转化为相似性度量问题，并应用深度度量学习方法识别工业开关仪表状态。新算法采用三元组网络模型，其中，三元组网络包含基于 SE-ResNet 的特征提取网络，并使用带有 L2 正则项的三元组损失进行训练，同时利用 XBM 提升训练效果。此外，构建了一个包含 13 种开关仪表的自定义数据集，用于训练和测试模型。最后，对新算法进行了闭集测试、开放集测试和小样本测试，并通过测试指标和距离直方图证明模型在开关仪表识别方面的高准确率，整体性能优于其他解决方案。新模型可有效区分已知类别和未知类别，具有强泛化识别能力。此外，轻量级网络适合现实场景的部署应用。

未来的工作主要集中在进一步提升泛化识别能力。具体方向包括扩充数据集，以涵盖更多开关种类和状态；使用不同的数据增强技术，改进损失函数和距离度量方法，以提升性能。由于三元组对样本对的敏感性，因此还可以进一步调整样本挖掘策略，如动态改变样本挖掘方式等。

参 考 文 献

[1] Ren SQ, He KM, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.

[2] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[3] 刘超, 韩懈. 基于 SE-YOLOV4 的变电站断路器分合状态识别算法 [J]. 软件导刊, 2022, 21(9): 40-44.
Liu C, Han X. Substation circuit breaker opening and closing state recognition algorithm based on SE-YOLOV4 [J]. Software Guide, 2022, 21(9): 40-44.

[4] 史梦安, 陆振宇. 基于 YOLO-tiny-RFB 模型的电站旋钮开关状态识别 [J]. 计算机应用, 2020, 40(12): 3679-3686.
Shi MA, Lu ZY. Power station rotary switch status recognition based on YOLO-tiny-RFB model [J]. Journal of Computer Applications, 2020, 40(12): 3679-3686.

[5] Fu CZ, Si WR, Lu QY, et al. Study of a detection and recognition algorithm for high-voltage switch cabinet based on deep learning with an improved faster-RCNN [C] // Proceedings of the 2018 International Conference on Engineering Simulation and Intelligent Control (ESAIC), 2018: 87-91.

[6] Dong C, Loy CC, He KM, et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(2): 295-307.

[7] Chen HP, He ZT, Shi BW, et al. Research on recognition method of electrical components based on YOLO V3 [J]. IEEE Access, 2019, 7: 157818-157829.

[8] Schroff F, Kalenichenko D, Philbin J. Facenet: a unified embedding for face recognition and

clustering [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 815-823.

[9] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[10] Wang CJ, Xin C, Xu ZL. A novel deep metric learning model for imbalanced fault diagnosis and toward open-set classification [J]. Knowledge-Based Systems, 2021, 220: 106925.

[11] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks [J]. Communications of the ACM, 2020, 63(11): 139-144.

[12] Schlegl T, Seeböck P, Waldstein SM, et al. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery [C] // Proceedings of the International Conference on Information Processing in Medical Imaging, 2017: 146-157.

[13] Sabokrou M, Khalooei M, Fathy M, et al. Adversarially learned one-class classifier for novelty detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 3379-3388.

[14] Kingma DP, Welling M. Auto-encoding variational bayes [Z/OL]. arXiv Preprint, arXiv: 1312.6114, 2013.

[15] Cao A, Luo Y, Klabjan D. Open-set recognition with gaussian mixture variational autoencoders [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(8): 6877-6884.

[16] Buquicchio L, Gerych W, Alajaji A, et al. Variational open set recognition (VOSR) [C] // Proceedings of the 2021 IEEE International Conference on Big Data (Big Data), 2021: 994-1001.

[17] Lin TY, Goyal P, Girshick R, et al. Focal loss for dense object detection [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.

[18] Li BY, Liu Y, Wang XG. Gradient harmonized single-stage detector [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 8577-8584.

[19] Jiang W, Huang K, Geng J, et al. Multi-scale metric learning for few-shot learning [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(3): 1091-1102.

[20] Wang X, Zhang HZ, Huang WL, et al. Cross-batch memory for embedding learning [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 6388-6397.

[21] Wang H, Wang YT, Zhou Z, et al. CosFace: large margin cosine loss for deep face recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 5265-5274.

[22] Deng JK, Guo J, Liu TL, et al. Sub-center ArcFace: boosting face recognition by large-scale noisy web faces [C] // Proceedings of the European Conference on Computer Vision, 2020: 741-757.

[23] Hadsell R, Chopra S, LeCun Y. Dimensionality reduction by learning an invariant mapping [C] // Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006: 1735-1742.

[24] Teh EW, DeVries T, Taylor GW. Proxynca++: revisiting and revitalizing proxy neighborhood component analysis [C] // Proceedings of the European Conference on Computer Vision, 2020: 448-464.

[25] Sohn K. Improved deep metric learning with multi-class N-pair loss objective [C] // Proceedings of the 30th International Conference on Neural Information Processing Systems, 2016: 1857-1865.