

基于自监督学习的变电站缺陷检测^①



刘华锋¹, 韩 翊², 茅耀斌¹

¹(南京理工大学 自动化学院, 南京 210094)

²(浙江华云信息科技有限公司, 杭州 310030)

通信作者: 茅耀斌, E-mail: myb_nust@126.com

摘要: 为了在使用少量标注样本情况下提升变电站设备缺陷检测精度, 提出一种基于自监督模型 SimSiam 的改进缺陷检测算法. 不同于原始 SimSiam, 改进后的算法无需使用标志性图像 (iconic images), 如 ImageNet 数据集, 而是直接利用非标志性图像 (non-iconic images) 如 COCO 数据集进行对比学习, 并在下游的缺陷检测任务上获得可媲美有监督方法的性能. 通过在投影层 (projection head) 和预测层 (prediction head) 中使用全卷积网络和空间注意力模块来代替 MLP, 保留高维特征的空间结构及局部信息; 同时在计算相似度前先对特征图进行均值池化以得到特征向量, 并对特征向量进行归一化以计算欧氏距离, 从而改进了自监督对比学习的损失函数. 实验结果表明该算法能充分利用非标志性图像进行对比学习, 并在只标注少量样本的条件下提升变电站设备缺陷检测的精度, 在表盘破损、挂空悬浮物、鸟巢、呼吸器硅胶变色及箱门闭合异常等 5 类缺陷检测任务上 mAP 达到 83.84%.

关键词: 变电站缺陷检测; 自监督学习; 非标志性图像; 投影层; 预测层

引用格式: 刘华锋, 韩翊, 茅耀斌. 基于自监督学习的变电站缺陷检测. 计算机系统应用, 2023, 32(5): 112-122. <http://www.c-s-a.org.cn/1003-3254/9088.html>

Substation Defect Detection Based on Self-supervised Learning

LIU Hua-Feng¹, HAN Yi², MAO Yao-Bin¹

¹(School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)

²(Zhejiang Huayun Information Technology Co. Ltd., Hangzhou 310030, China)

Abstract: To improve the precision of equipment defect detection of substations on the premise of only a small number of labeled samples used, this study proposes an improved defect detection algorithm based on the self-supervised model SimSiam. Unlike the original SimSiam, the improved algorithm directly utilizes non-iconic images, such as those in the dataset COCO, rather than using iconic images, like those in the dataset ImageNet, for contrastive learning and achieves performance comparable to that of any supervised methods in downstream defect detection tasks. By replacing multi-layer perception (MLP) networks with fully convolutional networks and spatial attention modules in the projection and prediction heads, the proposed algorithm preserves the spatial structure and local information of high-dimensional features. Furthermore, the output feature map is mean-pooled before similarity is calculated to obtain the eigenvector, which is then normalized to calculate the Euclidean distance and further modify the loss function of self-supervised contrastive learning. The experimental results show that the improved algorithm can make full use of non-iconic images for contrastive learning and improve the precision of equipment defect detection of substations on the premise of labeling only a small number of samples. Its mean average precision (mAP) reaches 83.84% in the detection of five types of defects, namely, broken meter dials, hanging suspended substances, nests, respirator silicone discoloration, and abnormal closure of the box door.

Key words: substation defect detection; self-supervised learning; non-iconic images; projection head; prediction head

① 收稿时间: 2022-10-04; 修改时间: 2022-11-04; 采用时间: 2022-12-23; csa 在线出版时间: 2023-03-24

CNKI 网络首发时间: 2023-03-27

随着电力系统的发展以及人们对电能需求量的增大,变电站在当下的日常生活和工业生产中变得愈发重要.由于变电站设备大多处于露天环境,可能出现诸如表盘破损、呼吸器硅胶变色等缺陷以及存在鸟巢、挂空悬浮物等异物,若未能及时发现此类问题,可能会造成不同等级的电力事故^[1],因此,对变电站设备进行定期巡检维护尤为重要.但是随着变电站规模扩大,设备增多,传统的人工巡检方式由于劳动强度大、人工成本高、效率低,难以达到正常巡检要求^[2].近年来,随着计算机视觉技术和网络技术的发展,变电站巡检质量和效率大大提高,电力维护人员可以通过监控相机、机器人、无人机等装备进行远程巡检,并进行缺陷的计算机自动识别.

传统的缺陷检测算法一般是通过使用 Canny 边缘检测、Hough 检测、直线检测以及小波变换等图像处理技术提取出巡检图片的纹理、颜色及形状等特征,然后再采用模板匹配、支持向量机 (support vector machine, SVM) 等机器学习算法对缺陷进行检测.文献 [3-7] 用 Gabor 滤波器进行纹理特征提取,以进行表面缺陷的检测.文献 [8] 利用 Gabor 滤波器具有频率选择和方向选择的特性,对带钢表面缺陷的纹理特征进行提取,得到了分割后的缺陷图像.文献 [9-11] 利用小波变换方法提取图像的纹理特征以进行表面缺陷的检测.

在电力系统中,传统算法一般利用颜色及形状特征来进行缺陷的检测.陈文浩等^[12]提出一种基于无人机电力巡检的玻璃绝缘子自爆、破损情况检测方法,对巡检所得图像,首先将 RGB 颜色空间转换为 LAB 颜色空间来避免光照影响;然后使用改进后的最大类间方差进行阈值分割;最后通过判断绝缘子对应位置上像素个数来准确定位绝缘子自爆或破损点的坐标,并在巡检图像中标识出来.Zhao 等^[13]提出了一种基于方位角检测和二值形状先验知识 (OAD-BSPK) 的复杂航空图像中不同角度多个绝缘子的定位方法.Kaftandjian 等^[14]研究设计了不同形态学的顶帽 (Top-Hat) 算子,该算子通过对不同类型的缺陷区域进行特征提取,再求得缺陷特征参数并建立缺陷模板,最后利用模板对缺陷进行分类.

随着深度学习技术的迅猛发展,人工特征提取方式逐渐被基于卷积神经网络的自动特征提取方式所代替.2018年,Cha 等^[15]最早将 Faster R-CNN 直接应用于桥梁表面缺陷定位,其主干网络被替换为 ZE-Net.

He 等^[16]提出了基于 Faster R-CNN 的带钢表面缺陷检测网络,该网络的改进在于将主干网络中多级特征图组合为一个多尺度特征图.尹子会等^[17]针对变电设备典型视觉缺陷检测中快速定位及缺陷检测的自动化程度和效率较低问题,提出了一种融合边缘计算和改进的 Faster R-CNN 变电设备典型视觉缺陷的检测系统.白涛^[18]针对电力设备中的绝缘子缺陷检测问题,提出了“先定位、后识别”的方案,通过改进现有目标检测框架 Faster R-CNN 和基于深度学习的方法提出了一种新的级联式目标检测模型.

然而,以上涉及的基于深度学习的缺陷检测方法极度依赖于大量的有标注数据,一般称之为有监督学习.这种方式的缺陷一是需要消耗大量的人力去采集和标注数据;二是变电站缺陷样本非常少,难以采集到足够数量的数据,因此,使用有监督的目标检测方法往往达不到很好的效果.基于这一认识,越来越多的研究者寻求如何利用大量的无标注数据来提升深度学习的效能.近年来,自监督学习以其良好的数据利用效率和泛化能力引起了人们的广泛关注,自监督学习主要利用辅助任务从大规模的无标注数据中挖掘自身的监督信息,通过这种构造的监督信息对网络进行训练,从而学到对下游任务有用的表征.在自监督学习领域,对比学习被广泛应用于目标检测等以分类为基础的下游任务中,并能获得良好的性能.如 CMC^[19]、SwAV^[20]、MoCo^[21]、SimCLR^[22]、BYOL^[23]、SimSiam^[24] 等对比式自监督学习方法,通过在 ImageNet 数据集上进行预训练,然后迁移到下游目标检测任务上,其性能可以接近甚至优于有监督方法.

已有的对比式自监督学习框架依赖于类似 ImageNet 等数据集的潜在特征:每一幅图像内容明确,且均是该单个对象为中心.这类图像被称之为标志性图像 (iconic images),这使得同一图像的不同视图和裁剪区域均对应于同一目标物^[25],如图 1 所示.然而在实际场景中,几乎都是包含多个目标物的图像(如 COCO、Pascal VOC 等数据集图像),它们称之为非标志性图像 (non-iconic images),对此类图像进行随机裁剪或者其他数据增强操作很可能导致增强后的图像对应于不同的目标物,这类图像并不适用于已有的对比式自监督学习框架,如图 2 所示,两次随机裁剪出来的目标不一致且相差较大.

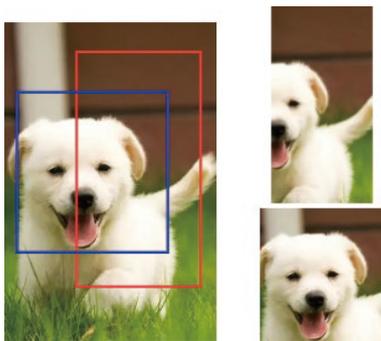


图1 标志性图像的随机裁剪

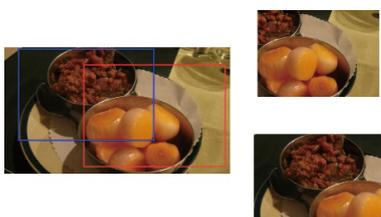


图2 非标志性图像的随机裁剪

针对上述问题,本文基于对比式自监督学习算法 SimSiam 提出了一种改进算法,改进后的算法能直接使用非标志性图像进行对比式学习.本文验证了该算法在通用目标检测任务上的性能,然后将其应用于变电站设备缺陷检测任务上.为了证明改进后算法对目标检测任务的性能提升效果,本文对比了由 detectron2 官方提供的基于 ImageNet 分类任务的有监督预训练模型和使用 SimSiam 算法在 COCO 数据集上进行自监督预训练获得的模型在下游目标检测任务上的性能.

实验结果表明,在 Pascal VOC 目标检测任务上,改进后的算法可达到 77.89% 的 mAP 值,而 SimSiam 算法和有监督模型分别为 75.50%、78.68%.在变电站设备缺陷检测任务上,改进后算法可达到 83.84% 的 mAP 值,不仅超过 SimSiam 算法的 81.44%,也优于有监督模型的 83.52%.

1 SimSiam 对比式自监督学习算法

在对比式自监督学习方法中,孪生网络是最常用的网络结构,孪生网络可以比较一幅图像的两个不同数据增强视图的相似性.在 BYOL 框架出现以前,大部分人认为对比式自监督模型避免坍塌的关键在于负样本的存在,然而随着研究的深入,文献 [23] 首次指出了即使不使用负样本,也可以避免模型的坍塌.

文献 [24] 表明了孪生神经网络表征学习避免坍塌的关键在于“prediction head”(预测层)和“stop-gradient”(梯度停止策略),其中 prediction head 是一种投影变换层,而 stop-gradient 则是在孪生网络的一个分支取消梯度传播策略.文献 [24] 还指出在不使用负样本对、大的训练批次以及动量编码器的情况下简单孪生网络也能学习到有用的特征表示. SimSiam 的网络结构如图 3 所示,首先,在对输入图像进行编码前会对原图像进行两种不同方式的随机数据增强(如随机裁剪、随机灰度化、随机高斯模糊、随机翻转等),然后将增强后的图像输入孪生网络中.在孪生网络中,编码器 $f(\theta)$ 使用 ResNet 网络,输出经特征图均值池化的一个 2048 维的特征向量;与编码器级联的是一个投影层 (projection head),它包括 3 个全连接层,其中最后一个全连接层不使用激活函数.在孪生网络的一个分支中增加了预测层,它也是由全连接层组成.需要注意的是,在网络的反向传播过程中,只对拥有预测层的分支进行参数更新,这一点和 MoCo、BYOL 等自监督框架保持一致.不同之处在于,在 SimSiam 算法中,并未使用动量的方式对孪生网络的另一分支进行参数更新,而是直接采用了参数共享的方式.如图 3 所示,增强视图 x_1 经编码器 $f(\theta)$ 得到特征向量 h_i ,再经投影层 $g(\theta)$ 得到特征向量 z_i ,最后经预测层 $q(\theta)$ 得到特征向量 v_i ;增强视图 x_2 只经过编码器和投影层分别得到特征向量 h_j 和 z_j .损失函数由 v_i 和 z_j 的负余弦相似度值所表示:

$$L = \frac{1}{2}D(v_1, z_2) + \frac{1}{2}D(v_2, z_1) \quad (1)$$

$$D(v_i, z_j) = -\frac{v_i \cdot z_j}{\|v_i\|_2 \cdot \|z_j\|_2} \quad (2)$$

其中, L 为总损失函数; $D(v_i, z_j)$ 为特征向量之间的负余弦相似度值; $D(v_1, z_2)$ 和 $D(v_2, z_1)$ 为对称损失,表示交换增强视图输入的孪生网络分支后得到的损失;针对两种不同的增强视图, v_1 和 v_2 为孪生网络中不含预测层分支的特征向量输出值; z_1 和 z_2 则为另一分支的特征向量输出值.

由于在 SimSiam 中使用了梯度停止策略,因此损失函数可以进一步表示为:

$$L = \frac{1}{2}D(v_1, sg(z_2)) + \frac{1}{2}D(v_2, sg(z_1)) \quad (3)$$

其中, sg 为梯度停止策略,与文献 [24] 中 $stopgrad$ 定义一致,表示孪生网络其中一个分支不进行梯度更新.

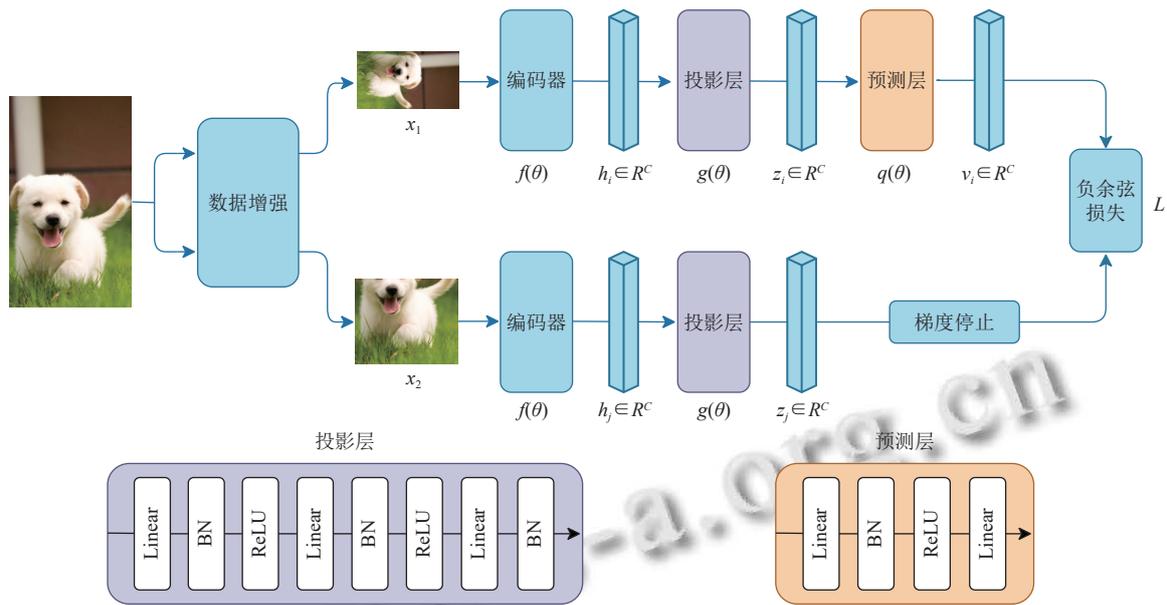


图3 SimSiam 网络结构

2 基于 SimSiam 的改进算法

由于 SimSiam 算法在使用非标志性图像进行对比式学习时无法学到很好的表征, 因此在下游任务中性能会下降, 而在变电站设备缺陷图像中往往会出现多种类别的缺陷且背景复杂, 不适合做 SimSiam 的预训练数据集. 为了合理利用大量无标注变电站缺陷数据, 本文对 SimSiam 算法进行改进. 主要包括: (1) 使用全卷积网络代替 MLP 重新构建投影层和预测层, 并在投影层中嵌入空间注意力模块; (2) 将损失函数由负余弦相似度修改为欧氏距离平方. 改进后算法能直接利用非标志性图像数据如 COCO 或者变电站设备缺陷图像进行对比学习, 而无需使用标志性图像. 实验结果表明, 当使用非标志性图像作为预训练数据集时, 改进后算法在下游目标检测任务上的性能不仅超越了 SimSiam, 且接近有监督模型. 论文源码地址为 <https://github.com/L-iiffer/Substation-Defect-Detection-based-on-Self-supervised-Learning/tree/master>.

2.1 投影层的改进

多层感知机 (multi-layer perceptron, MLP) 也叫人工神经网络, 由单层感知机演变而来. 单层感知机只有一个全连接层, 拥有简单的非线性变换能力, 而多层感知机拥有多个全连接层, 具有更强的非线性变换能力, 可以对复杂的数据特征进行拟合. 在 SimSiam 模型中, 投影层和预测层就是由多层感知机组成, 除了对编码

网络输出的特征进行非线性变换拟合之外, MLP 在自监督学习中被广泛用于提升模型对数据的表征能力^[26]. 另外, 值得注意的是, 在只有正样本的对比式自监督学习框架中, 如 BYOL、SimSiam 等, 预测层是避免模型坍塌的重要原因之一, 它的存在避免了直接拉近正样本对, 而是让模型去不断学习孪生网络中预测层分支到另一分支的映射.

在文献 [24] 中, 作者使用的预训练数据集为 ImageNet 中的大规模视觉识别挑战 (ImageNet large scale visual recognition challenge, ILSVRC) 2012 数据集, 用于图像分类和目标定位任务, 包含 1 000 个类别共 1 281 167 张图像, 该数据集绝大多数为标志性图像. 然而, 就数据收集的难易程度而言, 非标志性图像的来源更广泛, 更容易收集. 为了利用大量的无标注变电站设备缺陷数据, 本文基于 SimSiam 算法进行改进, 使得改进后算法能直接使用非标志性图像 (如 COCO 和变电站设备缺陷图像) 进行自监督对比式学习.

在 SimSiam 模型中, 投影层和预测层均由 MLP 构成, MLP 的非线性变换能力来自全连接层后面的 ReLU 激活层, 将多个全连接层进行堆叠后, 可以大大增强 MLP 的非线性变换能力, 从而拟合出表征能力更强的特征向量. 对标志性图像而言, 增强前后其中心目标物不发生改变, 因此使用特征向量便可对原始图像进行表征; 而对于非标志性图像, 图中同时具有多个目标,

增强前后其中心目标物极有可能不一致,甚至存在多个目标物,此时再利用特征向量进行表征则会丢失图像的空间结构和局部信息,使得计算出来的相似度不够准确.这一点在自监督的预训练过程中得以体现:SimSiam模型在使用非标志性图像进行对比学习时,最终的损失值会比使用标志性图像更低,但是在下游目标检测任务上的性能却不如后者,原因就在于使用特征向量对非标志性图像进行表征时,将所含目标物差异较大的视图认为相似.

为了解决这一问题,本论文使用全卷积网络代替MLP,并适当地嵌入空间注意力模块(spatial attention module, SAM)来加强局部特征之间的联系.在投影层和预测层中,用特征图代替特征向量,可以始终保留特征的空间结构和局部信息;嵌入空间注意力模块的主要作用是在特征图中基于通道方向获取关键区域信息,如图4所示.在投影层的第1个卷积层和第2个卷积层后面嵌入空间注意力模块,空间注意力模块的卷积核大小为 3×3 ,如图5所示.在前文提到,预测层是避免模型坍塌的重要原因,为了和SimSiam模型保持一致而不破坏其结构,仅将全连接层替换为全卷积层,而不嵌入空间注意力模块,如图6所示.

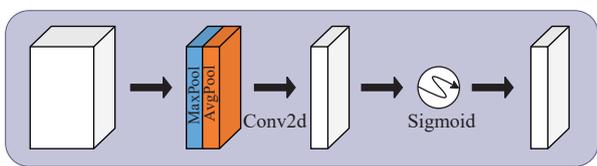


图4 空间注意力模块

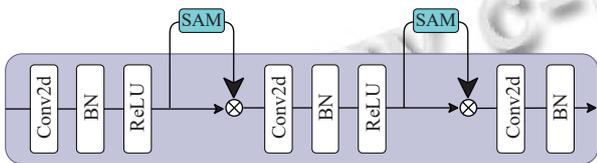


图5 投影层网络结构

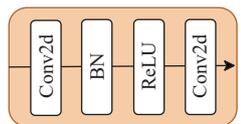


图6 预测层网络结构

由于全卷积网络输出的是 $C \times H \times W$ (C 、 H 、 W 分别表示特征图的深度、高度和宽度)大小的特征图,因

此在编码网络 $f(\theta)$ 中取消均值池化操作,保证编码网络的输出为 $C \times H \times W$ 的特征图.改进后的投影层和预测层不仅拥有MLP的非线性变换能力,还保留了特征图的空间结构以及局部信息,可以对非标志性图像进行良好的表征.

2.2 损失函数的改进

在图像分类领域,通常利用余弦相似度来度量图像间的相似性,例如在人脸识别领域^[27],使用余弦相似度作为衡量人脸相似的标准.余弦相似度通过测量两个向量内积空间的余弦值来度量它们之间的相似性,两个向量之间的角度余弦值可以用于确定两个向量是否大致指向相同的方向.因此,在对比式自监督学习中,也常采用余弦相似度作为度量图像相似程度的手段.

在SimSiam算法中,作者利用余弦相似度方法来计算投影层和预测层输出的特征向量的相似值,并以此作为孪生网络输入图像的增强视图相似程度的衡量.如式(1)和式(2)所示,采用负余弦相似度值作为损失函数,损失函数由对称损失组成,对称损失表示交换增强视图输入的孪生网络分支后所获取的负余弦相似度值.

尽管余弦相似度在衡量图像相似性上有一定优势,但欧氏距离度量对特征向量的具体数值差异更加敏感.为了验证这一点,本文使用欧氏距离对特征向量之间的差异进行度量,并将欧氏距离平方作为损失函数.实验结果表明,采用欧氏距离平方作为损失函数确实能进一步提升模型的性能.具体描述如下:

$$L = \frac{1}{2}d^2(v_1, z_2) + \frac{1}{2}d^2(v_2, z_1) \quad (4)$$

$$d(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (5)$$

其中, L 为使用欧氏距离度量后的损失函数; $d(a, b)$ 为欧氏距离度量; a 和 b 为 n 维特征向量, $a = (a_1, a_2, \dots, a_n)$, $b = (b_1, b_2, \dots, b_n)$; $d^2(v_1, z_2)$ 和 $d^2(v_2, z_1)$ 为两个特征向量的欧氏距离平方; v_1 、 v_2 、 z_1 、 z_2 和式(1)中定义一致.

在投影层和预测层中,输出均为 $C \times H \times W$ 的特征图,为方便计算欧氏距离,首先在通道 C 方向对特征图进行均值池化,获得 C 维特征向量;然后进行L2归一化处理,归一化后的特征向量其欧氏距离平方等价于

余弦相似度 (具体证明见附录 A). 也就是说, 当采用欧式距离平方作为损失函数值时, 不仅可以在特征向量方向上进行度量, 在数值上也能进行度量. 最终的损失函数为:

$$L = 2 + D(v_1, sg(z_2)) + D(v_2, sg(z_1)) \quad (6)$$

其中, $D(v_1, sg(z_2))$ 和 $D(v_2, sg(z_1))$ 与式 (3) 一致.

基于 SimSiam 改进后算法如图 7 所示, 虚线框中部分为对原 SimSiam 的改动.

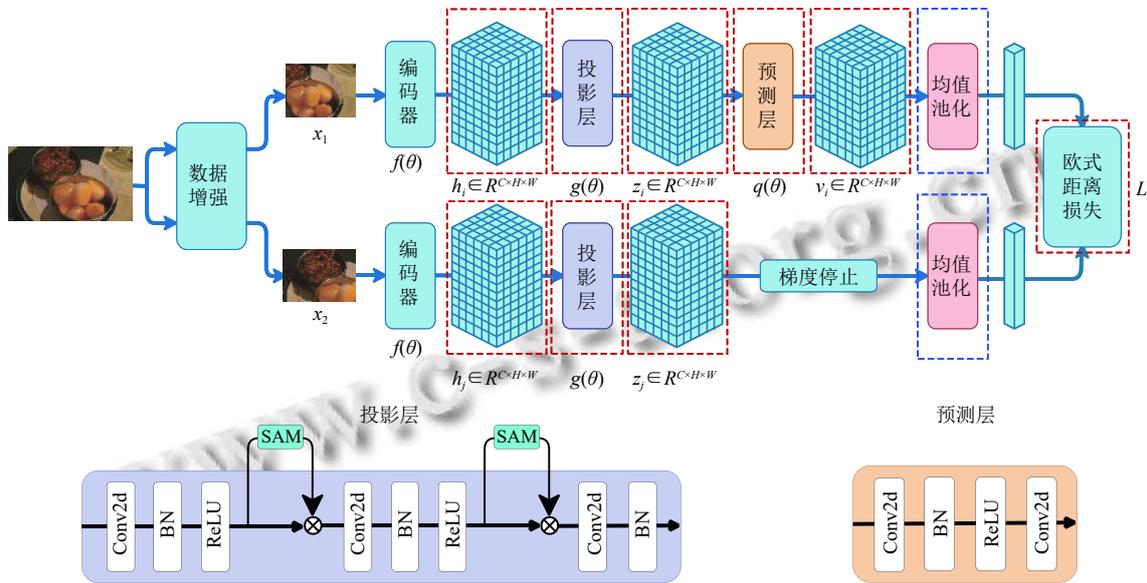


图 7 基于 SimSiam 模型改进算法的网络结构

3 实验数据集及评价指标

3.1 自监督学习数据集

本文的自监督学习算法所使用的数据集由两部分组成.

1) COCO2017 数据集. 包含 118287 张训练集图像、5000 张验证集图像以及 40670 张测试集图像, 实验中仅使用训练集图像进行自监督学习预训练.

2) 变电站设备缺陷数据集. 包含表盘破损、表盘模糊、绝缘子破裂、呼吸器硅胶变色等 25 类缺陷, 合计 36136 张图像, 全部用于自监督学习预训练.

3.2 下游任务数据集

在自监督学习领域, 需要通过下游任务去评估自监督学习的效果, 下游任务的实现仍然需要使用少量有标注数据, 实现方式是将自监督学习到的预训练模型在下游任务的有标注数据上进行迁移训练并进行性能评估. 本文中, 为了验证改进后算法是否能在非标志性图像数据集上获得更好的表征, 从而提升模型的检测性能, 首先将预训练模型迁移到 Pascal VOC 目标检测任务上. 在 VOC2007 数据集中, 包含训练集+验证集图像 5011 张, 测试集图像 4952 张; 在

VOC2012 数据集中, 包含训练集+验证集图像 11540 张. 混合 VOC2007 数据集和 VOC2012 数据集的训练集+验证集图像后得到 5011+11540=16551 张图像, 作为目标检测任务的训练集, 并使用 VOC2007 数据集的 4952 张测试集图像作为目标检测任务的测试集. Pascal VOC 目标检测任务中训练集和测试集各类别实例分布如图 8 所示.

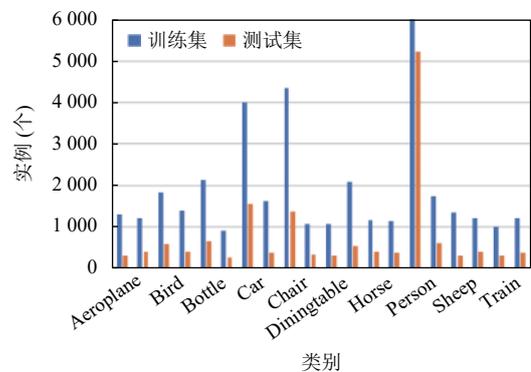


图 8 Pascal VOC 数据集实例分布

针对变电站设备缺陷检测任务, 同样进行了相关实验. 通过变电站现场监控相机抓拍和互联网图像抓

取方式, 本文获取了 9671 张变电站设备缺陷图像, 包括表计表盘破损、挂空悬浮物、鸟巢、呼吸器硅胶变色以及箱门闭合异常等 5 类缺陷. 随机挑选其中 8032 张图像作为训练集、其余 1605 张图像作为测试集, 所有图像均使用 LabelImg 工具进行标注. 变电站设备缺陷检测任务中训练集和测试集各类别实例分布如图 9 所示.

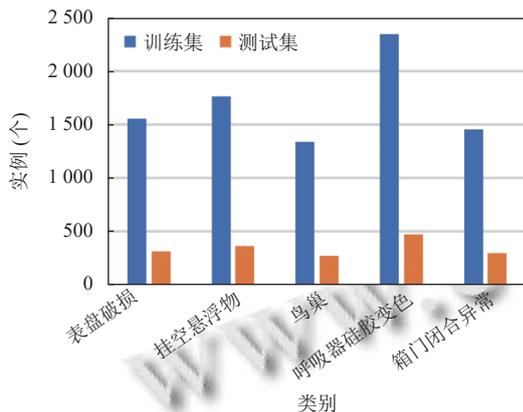


图 9 变电站设备缺陷数据集实例分布

3.3 评价指标

在文献 [24] 中, 为了评估预训练模型在下游目标检测任务上的性能, 作者使用了两种不同的评价指标进行评估: 在 Pascal VOC 目标检测任务中使用 VOC 评价指标; 在 COCO 目标检测任务中使用 COCO 评价指标. 因此, 对于变电站设备缺陷检测任务, 本文统一采用 COCO 评价指标进行模型性能评估; Pascal VOC 目标检测任务则使用 VOC 评价指标. 采用此种评估方法的另一优势在于: 对于不同的评价指标, 若模型的性能均有提升, 则更具说服力.

查准率 P 、查全率 R 计算公式如下:

$$P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN} \quad (7)$$

其中, TP 表示实际为某类正样本, 预测为该类别且预测框与真实框交并比 (intersection over union, IoU) 大于 0.5 的框的数量; FP 表示预测类别错误或交并比小于 0.5 的框的数量; FN 表示实际为负样本但预测为正样本的框的数量.

对不同置信度阈值求取不同的 P 、 R 组合, 即可得到 P - R 曲线, 该曲线与坐标轴围成的面积表示为 average precision (AP), 它用来衡量对某个类别的检测精度, AP 越大, 精度越高. 对所有类别的 AP 求平均值, 可得

到各类别的平均检测精度 mAP, 它用于衡量模型的整体检测精度.

COCO 评价指标中, 在计算 P - R 曲线与坐标轴围成的面积时, 使用的是 101 点插值法, 相较于 VOC 评价指标中使用的 11 点插值法或者平滑 P - R 曲线面积法, 计算方式更加严格, 计算结果更加精确^[28].

4 实验结果

实验过程分为对比式自监督学习、下游 Pascal VOC 目标检测性能评估、下游变电站设备缺陷检测性能评估 3 部分. 实验流程图如图 10 所示.

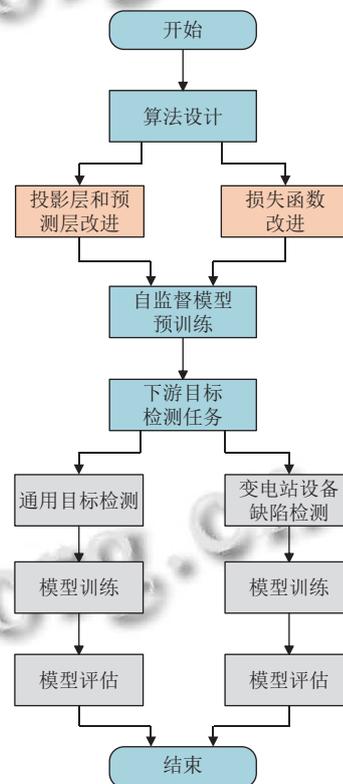


图 10 实验流程图

4.1 对比式自监督学习

在自监督模型训练中, 实验平台为运行 Ubuntu 16.04.1 操作系统的服务器, 它内含两张 NVIDIA GeForce GTX TITAN X 显卡, 每张显卡运行内存均为 12 GB. 深度学习软件环境为 CUDA 10.1、Python 3.7.10、PyTorch 1.7.1.

- 1) 数据集. 无标签的 COCO 2017 训练集, 包含图像 118287 张.
- 2) 数据增强方式. 和文献 [24] 中保持一致, 包括随

机裁剪、随机灰度化、随机高斯模糊、随机翻转等。数据增强用于孪生网络的不同分支进行对比学习,原始图像数量保持不变,COCO数据集图像数量为118287,变电站设备缺陷数据集图像数量为36136。

3) 网络结构。编码器采用ResNet-50作为主干网络,网络输入的大小为 224×224 ;在本文改进后的算法中,取消ResNet-50最后的均值池化层和全连接层。

4) 优化器。使用权重衰减为 $1E-4$ 的SGD优化器,初始学习率大小为0.05,按照BatchSize/256策略进行学习率的调整,并在训练过程中使用余弦衰减策略调整学习率大小。在数据集上训练模型400个epoch,并使用两个GPU进行分布式训练,由于运行内存的限制,每批次的大小为96个样本。

4.2 Pascal VOC 目标检测性能评估

在下游任务实验中,实验平台为运行Ubuntu 18.04.5操作系统的PC机,其显卡为NVIDIA GeForce RTX 3090,运行内存为24GB。使用深度学习平台detectron2搭建Faster R-CNN C4目标检测框架并实现下游任务。

为了验证本文提出的算法在非标志性图像上获得的表征优于SimSiam算法,分别使用SimSiam算法和本文算法按照第4.1节设置进行了自监督实验,然后将所得预训练模型用于Pascal VOC目标检测任务。为了更好地对比自监督学习和有监督学习在目标检测任务上的性能差距,使用detectron2官方提供的有监督预训练模型(ImageNet数据集进行有监督分类任务,主干网络为ResNet-50)用于Pascal VOC目标检测任务。使用第3.2节中所述Pascal VOC训练集进行精调,并在测试集上评估模型的性能。由于设备内存限制,将批次大小设置为4,初始学习率为0.005,优化器为SGD,总的迭代次数为72000,在第48000和64000次迭代学习率会下降为原来的0.1,并在前100个迭代内采用学习率预热策略。最终评估后的性能如表1所示,SimSiam算法所得预训练模型在下游任务上的检测精度最低,mAP值为75.50%;而本文算法所得预训练模型的mAP值为77.89%,不仅优于SimSiam算法,且接近有监督预训练模型的mAP值78.68%。此外,为了进一步证明改进后算法的优势,本文参考了文献[25]中BYOL算法在COCO数据集上所得预训练模型的性能,其在Pascal VOC测试集上的mAP值为75.80%。

表1 预训练模型在Pascal VOC测试集上的检测性能(%)

预训练模型	预训练数据	AP	AP ₅₀	AP ₇₅
有监督	ImageNet	49.55	78.68	52.92
SimSiam	COCO	48.40	75.50	52.31
BYOL	COCO	50.00	75.80	54.80
本文算法	COCO	49.49	77.89	53.09

为了验证单独应用第2.1节、第2.2节改进方法对自监督算法性能的提升效果,本文进行了消融实验。具体为:分别将第2.1节改进和第2.2节改进加入SimSiam算法中,然后进行自监督预训练,将所得预训练模型仍然应用于Pascal VOC目标检测任务,评估模型在测试集上的检测性能,实验结果如表2所示。从表2中可知,当仅使用全卷积网络替代MLP时,模型的检测性能就得到了提升,这说明全卷积网络保留了特征的空间结构和局部信息从而加强了模型对非标志性图像的表征能力,能够在图片中获取更多的信息;而单独应用欧式距离平方作为损失函数同样提升了模型的检测性能。这是因为对于自监督学习而言,当人工设定的代理任务越难时,意味着模型获取的图像语义特征更加丰富,因此在对比式学习中,更加严格的相似度度量方式有助于提升模型在下游任务中的性能,尤其当下游任务是以分类为基础的目标检测任务。

表2 消融实验结果(%)

预训练模型	预训练数据	AP	AP ₅₀	AP ₇₅
SimSiam	COCO	48.40	75.50	52.31
SimSiam+全卷积	COCO	49.18	76.65	52.31
SimSiam+欧氏距离	COCO	49.23	76.77	52.56

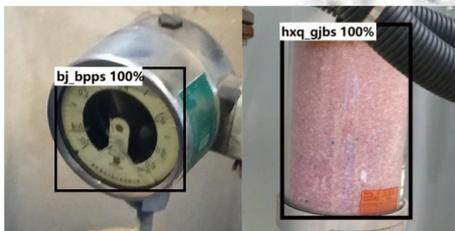
4.3 变电站设备缺陷检测性能评估

在变电站设备缺陷检测任务中,首先,仍然按照第4.1节的方案进行自监督实验,使用SimSiam算法和本文算法进行自监督预训练,然后在下游缺陷检测任务上评估预训练模型的性能。当自监督模型预训练数据集与下游任务训练集相同或者为相同域时,有助于提升预训练模型在下游任务上的性能^[29]。因此,在保证训练集样本数量充足的条件下,同时也为了充分利用大量的无标注变电站设备缺陷样本,本文尝试混合第3.1节中COCO数据和变电站设备缺陷数据作为新的自监督训练集,图像数量为154423,同样使用本文算法进行自监督学习,所有的实验结果如表3所示。根据表中数据可知,本文算法所得预训练模型在缺陷检测任务上的mAP值为83.67%,相较于SimSiam预训练

模型提升了 2.23 个百分点,且略优于有监督预训练模型的 mAP 值 83.52%。当使用混合数据集时,本文算法所得预训练模型的检测性能得到进一步提升, mAP 值达到了 83.84%。一些典型缺陷检测结果如图 11 所示。为准确衡量模型的推理速度,对 100 张图像的检测时间求平均值,结果表明:基于 CPU 的推理速度为 1 f/s;基于 GPU 的推理速度为 11.1 f/s。

表 3 预训练模型在缺陷测试集上的检测性能 (%)

预训练模型	预训练数据	AP	AP ₅₀	AP ₇₅
有监督	ImageNet	49.88	83.52	53.18
SimSiam	COCO	47.94	81.44	48.79
本文算法	COCO	50.54	83.67	53.19
本文算法	COCO+变电站缺陷数据	50.27	83.84	51.67



(a) 表盘破损和呼吸器硅胶变色



(b) 箱门闭合异常和挂空悬浮物



(c) 鸟巢和挂空悬浮物

图 11 缺陷检测结果

5 结论

为提高变电站设备缺陷检测精度,本文提出一种基于自监督学习算法 SimSiam 的改进算法,使用全卷

积网络取代 SimSiam 中的 MLP 模块,并在投影层中嵌入空间注意力模块,以尽可能地保留图像的空间结构和局部信息;同时,将损失函数修改为欧氏距离度量以获得更好的图像表征。本文提出的算法可以直接使用无标签的 COCO 数据集和变电站设备缺陷数据集进行自监督预训练。通过上述方式,将预训练模型应用在变电站设备缺陷检测任务中,实验表明,对 5 类缺陷检测的 mAP 达到了 83.84%,相较于有监督模型的检测性能,提升了 0.32%;同时相较于 SimSiam 算法预训练模型的检测性能提升了 2.4%。证明了基于 SimSiam 改进后的算法无需使用标志性图像进行对比学习,在非标志性图像上也能学到良好的表征,有利于提升变电站设备缺陷检测精度。鉴于目前自监督方法性能还无法完全超过有监督方法,而在缺陷检测任务上,本文改进后算法可以与有监督方法相持平,证明了改进后算法的优势。

参考文献

- 1 应樱. 基于改进 Faster R-CNN 的变电站设备缺陷检测算法研究 [硕士学位论文]. 杭州: 浙江大学, 2021. [doi: 10.27461/d.cnki.gzjdx.2021.000392]
- 2 伍艺佳, 华雄, 王丽蓉, 等. 基于注意力机制学习的变电站设备缺陷检测方法. 计算机与现代化, 2021, 36(2): 7-8. [doi: 10.3969/j.issn.1006-2475.2021.02.002.]
- 3 Tsa DM, Wu SK. Automated surface inspection using Gabor filters. The International Journal of Advanced Manufacturing Technology, 2020, 16(7): 474-482. [doi: 10.1007/s001700070055]
- 4 Bodnarova A, Bennamoun M, Latham S. Optimal Gabor filters for textile flaw detection. Pattern Recognition, 2002, 35(12): 2973-2991. [doi: 10.1016/S0031-3203(02)00017-1]
- 5 Kumar A, Pang GKH. Defect detection in textured materials using Gabor filters. IEEE Transactions on Industry Applications, 2002, 38(2): 425-440. [doi: 10.1109/28.993164]
- 6 Tsai DM, Lin CP, Huang KT. Defect detection in coloured texture surfaces using Gabor filters. The Imaging Science Journal, 2005, 53(1): 27-37. [doi: 10.1179/136821905X26935]
- 7 Hu GH. Optimal ring Gabor filter design for texture defect detection using a simulated annealing algorithm. Proceedings of 2014 International Conference on Information Science, Electronics and Electrical Engineering. Sapporo: IEEE, 2014. 860-864. [doi: 10.1109/InfoSEEE.2014.6947789]

- 8 丛家慧, 颜云辉, 董德威. Gabor 滤波器在带钢表面缺陷检测中的应用. 东北大学学报 (自然科学版), 2010, 31(2): 257–260. [doi: [10.3969/j.issn.1005-3026.2010.02.027](https://doi.org/10.3969/j.issn.1005-3026.2010.02.027)]
- 9 Lee CS, Choi CH, Choi JY, *et al.* Feature extraction algorithm based on adaptive wavelet packet for surface defect classification. Proceedings of the 3rd IEEE International Conference on Image Processing. Lausanne: IEEE, 1996. 673–676. [doi: [10.1109/ICIP.1996.560968](https://doi.org/10.1109/ICIP.1996.560968)]
- 10 周鹏, 徐科, 刘顺华. 基于剪切波和小波特征融合的金属表面缺陷识别方法. 机械工程学报, 2015, 51(6): 98–103. [doi: [10.3901/JME.2015.06.098](https://doi.org/10.3901/JME.2015.06.098)]
- 11 Singh A, Dutta MK, Parthasarathi M, *et al.* Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image. Computer Methods and Programs in Biomedicine, 2016, 124: 108–120. [doi: [10.1016/j.cmpb.2015.10.010](https://doi.org/10.1016/j.cmpb.2015.10.010)]
- 12 陈文浩, 姚利娜, 李丰哲. 无人机电网巡检中的绝缘子缺陷检测与定位. 计算机应用, 2019, 39(S1): 210–214.
- 13 Zhao ZB, Liu N, Wang L. Localization of multiple insulators by orientation angle detection and binary shape prior knowledge. IEEE Transactions on Dielectrics and Electrical Insulation, 2015, 22(6): 3421–3428. [doi: [10.1109/TDEI.2015.004741](https://doi.org/10.1109/TDEI.2015.004741)]
- 14 Kaftandjian V, Zhu YM, Dupuis O, *et al.* The combined use of the evidence theory and fuzzy logic for improving multimodal nondestructive testing systems. IEEE Transactions on Instrumentation and Measurement, 2005, 54(5): 1968–1977. [doi: [10.1109/TIM.2005.854255](https://doi.org/10.1109/TIM.2005.854255)]
- 15 Cha YJ, Choi W, Suh G, *et al.* Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. Computer-aided Civil and Infrastructure Engineering, 2018, 33(9): 731–747. [doi: [10.1111/mice.12334](https://doi.org/10.1111/mice.12334)]
- 16 He Y, Song KC, Meng QG, *et al.* An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1493–1504. [doi: [10.1109/TIM.2019.2915404](https://doi.org/10.1109/TIM.2019.2915404)]
- 17 尹子会, 孟荣, 范晓丹, 等. 融合边缘计算和改进 Faster R-CNN 的变电站设备典型视觉缺陷检测系统. 中国科技论文, 2021, 16(3): 343–348. [doi: [10.3969/j.issn.2095-2783.2021.03.018](https://doi.org/10.3969/j.issn.2095-2783.2021.03.018)]
- 18 白涛. 基于深度学习的电力绝缘子识别方法研究 [硕士学位论文]. 西安: 西安石油大学, 2021. [doi: [10.27400/d.cnki.gxasc.2021.000994](https://doi.org/10.27400/d.cnki.gxasc.2021.000994)]
- 19 Tian YL, Krishnan D, Isola P. Contrastive multiview coding. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 776–794. [doi: [10.1007/978-3-030-58621-8_45](https://doi.org/10.1007/978-3-030-58621-8_45)]
- 20 Caron M, Misra I, Mairal J, *et al.* Unsupervised learning of visual features by contrasting cluster assignments. Proceedings of the 34th Conference on Neural Information Processing Systems. Vancouver: NeurIPS, 2020. 1–13.
- 21 He KM, Fan HQ, Wu YX, *et al.* Momentum contrast for unsupervised visual representation learning. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 9726–9735. [doi: [10.1109/CVPR42600.2020.00975](https://doi.org/10.1109/CVPR42600.2020.00975)]
- 22 Chen T, Kornblith S, Norouzi M, *et al.* A simple framework for contrastive learning of visual representations. Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020. 1597–1607.
- 23 Grill JB, Strub F, Altché F, *et al.* Bootstrap your own latent: A new approach to self-supervised learning. Proceedings of the 34th Conference on Neural Information Processing Systems. Vancouver: NeurIPS, 2020. 1–14.
- 24 Chen XL, He KM. Exploring simple siamese representation learning. Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 15745–15753. [doi: [10.1109/CVPR46437.2021.01549](https://doi.org/10.1109/CVPR46437.2021.01549)]
- 25 Liu ST, Li ZM, Sun J. Self-EMD: Self-supervised object detection without ImageNet. arXiv:2011.13677v3, 2021.
- 26 Wang YZ, Tang SX, Zhu F, *et al.* Revisiting the transferability of supervised pretraining: An MLP perspective. Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 9173–9183. [doi: [10.1109/CVPR52688.2022.00897](https://doi.org/10.1109/CVPR52688.2022.00897)]
- 27 刘一玮, 杨韬, 刘瑾, 等. 基于余弦相似度的人脸识别系统的实现. 电子技术与软件工程, 2015, (9): 90.
- 28 Everingham M, van Gool L, Williams CKI, *et al.* The Pascal visual object classes (VOC) challenge. International Journal of Computer Vision, 2010, 88(2): 303–338. [doi: [10.1007/s11263-009-0275-4](https://doi.org/10.1007/s11263-009-0275-4)]
- 29 Cole E, Yang X, Wilber K, *et al.* When does contrastive visual representation learning work? Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022. 14735–14744. [doi: [10.1109/CVPR52688.2022.01434](https://doi.org/10.1109/CVPR52688.2022.01434)]

附录 A

向量 a 与 b 归一化欧氏距离与余弦距离之间的关系如下.

定义余弦距离:

$$\cos(a, b) = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}}$$

假设: $a = (a_1, a_2, \dots, a_n)$, $b = (b_1, b_2, \dots, b_n)$.

定义归一化欧氏距离:

$$d(a', b') = \sqrt{\sum_{i=1}^n (a'_i - b'_i)^2}$$

其中,

$$a' = \frac{1}{\sqrt{\sum_{i=1}^n a_i^2}} \cdot (a_1, a_2, \dots, a_n)$$

$$b' = \frac{1}{\sqrt{\sum_{i=1}^n b_i^2}} \cdot (b_1, b_2, \dots, b_n)$$

为归一化向量, 则:

$$\begin{aligned} d(a', b') &= \sqrt{\sum_{i=1}^n \left(\frac{a_i}{\sqrt{\sum_{i=1}^n a_i^2}} - \frac{b_i}{\sqrt{\sum_{i=1}^n b_i^2}} \right)^2} \\ &= \sqrt{\sum_{i=1}^n \left(\frac{a_i^2}{\sum_{i=1}^n a_i^2} + \frac{b_i^2}{\sum_{i=1}^n b_i^2} - \frac{2a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}} \right)} \\ &= \sqrt{\frac{\sum_{i=1}^n a_i^2}{\sum_{i=1}^n a_i^2} + \frac{\sum_{i=1}^n b_i^2}{\sum_{i=1}^n b_i^2} - \frac{2 \sum_{i=1}^n a_i b_i}{\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2}} \\ &= \sqrt{2 - 2 \left(\frac{\sum_{i=1}^n a_i b_i}{\sum_{i=1}^n a_i^2 \sum_{i=1}^n b_i^2} \right)} \\ &= \sqrt{2 - 2 \cos(a, b)} \end{aligned}$$

因此, 归一化欧氏距离与余弦距离之间的关系为:

$$d^2(a', b') = 2(1 - \cos(a, b)) = 2(1 - \cos(a', b'))$$

由正文式 (2) 可知:

$$D(a', b') = -\frac{a'}{\|a'\|_2} \cdot \frac{b'}{\|b'\|_2} = -\cos(a', b')$$

于是:

$$d^2(a', b') = 2(1 + D(a', b'))$$

(校对责编: 牛欣悦)