

高宏亮, 黄承曦, 刘颖, 等. 基于对比型卷积神经网络的异常检测研究[J]. 智能计算机与应用, 2026, 16(3): 154-159. DOI: 10.20169/j.issn.2095-2163.24070303

# 基于对比型卷积神经网络的异常检测研究

高宏亮<sup>1</sup>, 黄承曦<sup>2,3</sup>, 刘颖<sup>1</sup>, 霍光磊<sup>2,3</sup>

(1 西安航天化学动力有限公司, 西安 710025; 2 福建(泉州)先进制造技术研究院, 福建 泉州 362000;

3 福建省智能运维机器人技术重点实验室, 福建 泉州 362000)

**摘要:** 针对人机实时交互场景中交互对象存在模糊性和时空不完整性等问题, 提出对比型卷积神经网络对机器人视角下的异常检测问题进行研究。首先, 根据格式塔原则利用高斯卷积构建图像序列间的差分地图和能量地图, 并基于 Kulback-Liebler 散度建立一致性直方图; 然后, 采用 2 个并行的卷积神经网络对一致性直方图进行特征提取, 并根据对数似然分布进行差异统计; 最后, 利用整体分布概率对 2 个网络的时空关系进行表征, 并通过对比和串接操作完成网络搭建。在交互行为数据库上进行了实验, 所提出的方法平均识别率为 79.3%, 对比于其他方法, 具有更高的准确性。

**关键词:** 异常检测; 卷积神经网络; 时空一致性表征; 对比型卷积神经网络

中图分类号: TP183

文献标志码: A

文章编号: 2095-2163(2026)03-0154-06

## Anomaly detection based on contrast convolutional neural network

GAO Hongliang<sup>1</sup>, HUANG Chengxi<sup>2,3</sup>, LIU Ying<sup>1</sup>, HUO Guanglei<sup>2,3</sup>

(1 Xi'an Aerospace Chemical Propulsion Co., Ltd., Xi'an 710025, China;

2 Fujian(Quanzhou) Institute of Advanced Manufacturing Technology, Quanzhou 362000, Fujian, China;

3 Fujian Key Laboratory of Intelligent Operation and Maintenance Robot Technology, Quanzhou 362000, Fujian, China)

**Abstract:** Aiming at the problems of fuzzy and spatio-temporal incompleteness of interactive objects in human-computer real-time interaction scenes, a contrast convolutional neural network is proposed to study anomaly detection from the robot perspective. Firstly, the difference map and energy map between image sequences are constructed by using Gaussian convolution according to Gestalt principle, and the consistency histogram is established based on Kulback-Liebler divergence. Then, two parallel convolutional neural networks are used to extract the features of the consistent histogram, and the difference statistics are carried out according to the log-likelihood distribution. Finally, the spatial and temporal relationship of the two networks is characterized by the global distribution probability, and the network is constructed by comparison and serial operation. Experiments are carried out on interactive behavior database, and the average recognition rate of the proposed method is 79.3%, which has higher accuracy than other methods.

**Key words:** anomaly detection; convolutional neural network; spatio-temporal consistency characterization; contrast convolutional neural networks

## 0 引言

基于视觉的异常检测是保障机器人实现安全、友好人机交互的关键功能之一, 是机器人及机器视觉领域的研究热点和难点。在自然人机交互过程中, 场景中的交互对象在时空上存在极大复杂性, 机器人往往无法获得交互对象完整、明确的信息。尽管国内外学者对异常检测进行了大量研究<sup>[1-5]</sup>, 机

器人视角下异常检测的模糊性和时空不完整性问题仍未很好地得到解决。

异常行为在时空上具有关联性, 依赖异常行为的静态姿态往往不能实现准确检测。因此, 本文将人机交互中的异常行为检测限定为异常视频的分类, 并重点对图像序列的相互关系进行时空表征, 将异常检测问题视为通过融合时空上下文信息而非在静态状态下进行空间异常检测来识别异常行为。

**基金项目:** 福建省自然科学基金(2022J05107); 泉州市科技计划项目(2022C004L)。

**作者简介:** 高宏亮(1985—), 男, 硕士研究生, 主要研究方向: 智能制造工程技术应用。

**通信作者:** 霍光磊(1982—), 男, 博士, 高级工程师, 主要研究方向: 机器人感知技术。Email: hg10417@163.com。

收稿日期: 2024-07-03

哈尔滨工业大学主办 ◆ 专题设计与应用

随着人工智能技术的不断突破,机器学习算法被应用到机器人异常检测中。Park 等学者<sup>[6]</sup>通过融合力、声音和运动信号的多模态信息提出一种基于隐藏马尔可夫模型-高斯过程的方法,用于监视 PR2 机器人上的异常行为。为监视工业机器人的异常动作,彭煜祺等学者<sup>[7]</sup>介绍了一种基于图像哈希技术的工业机器人姿态编码方法,通过完成图像分割并结合序列分析实现机器人的异常动作识别。赵廉等学者<sup>[8]</sup>采用人体关键点作为特征,结合机器学习中的分类器完成对工厂中工人睡觉和跌倒两种异常行为的检测。

近年来,深度学习算法成为异常检测的主流技术,卷积神经网络因其良好的图像特征提取和泛化能力被广泛采用。魏麟等学者<sup>[9]</sup>对基于卷积神经网络的 YOLOv5 框架进行了改进,将坐标注意力机制融入到原模型中,并通过优化损失函数实现模型训练,完成驾驶舱内飞行员的异常行为识别。余胤翔等学者<sup>[10]</sup>基于深度学习中的 SlowFast 网络对人体时空动作进行建模,并结合 FasterR-CNN 对社区安防异常行为进行检测。基于传统机器学习和基于深度学习的算法虽然取得了较好的异常检测效果,但忽略了行为序列间的时空语义特征。

为对异常行为的时空一致性进行表征,本文提出一种对比型卷积神经网络对行为的时空语义关系进行提取。从特征选择的角度,考虑图像序列的差分特征作为行为的局部差异,即像素级别上的差异。由于视觉决策过程不仅涉及差异信息,还与图像对中每个图像提供的信息有关,因此,考虑差分特征与原始图像的相关性,以便将感知差异性和相关特征结合到一致性表征中。通过前向和后向差分表征,结合对数似然分布对 2 个深度卷积图像进行差异统计,根据分布概率建立深度对比层,实现卷积神经网络深度特征间的语义表征,并通过对比和串接操作完成网络搭建。构建的对比层权重由基于深度特征的联合分布整体概率、2 个网络的各自分布概率以及成对特征的概率权重进行加权,最后通过 Softmax 回归,实现行为视频中的异常检测。本文主要贡献为:

- (1) 提出一种描述行为序列间的局部差分特征的算法。
- (2) 提出一种考虑分布概率的直方图软分配算法,用于行为时空语义提取。
- (3) 构建一种对比型卷积神经网络,用于学习行为的时空特征关系。

## 1 行为差异表征

### 1.1 差异地图和能量地图

基于图像差分,为了捕捉行为序列之间的细微差异,考虑到像素的局部结构相似,将局部支持窗口视为单位,以缓解由结构相似性引起的图像模糊问题。定义  $p_1$  和  $p_2$  表示图像  $I_1$  和  $I_2$  中位置相同的 2 个像素点,并定义其对应的矩形支持窗口为  $S(p_1)$  和  $S(p_2)$ ,其中像素点的位置对应于支持窗口的中心点;根据相似性和邻近性的格式塔原理, $p_1$  和  $p_2$  之间的差异为:

$$D(p_1, p_2) = \frac{\sum_{q_1 \in S(p_1), q_2 \in S(p_2)} w(p_1, q_1) w(p_2, q_2) e(q_1, q_2)}{\sum_{q_1 \in S(p_1), q_2 \in S(p_2)} w(p_1, q_1) w(p_2, q_2)} \quad (1)$$

其中,格式塔原理强调元素间的整体性,即认为各个子体具有相似性和连续性关系。像素级的匹配代价  $e(q_1, q_2)$  是通过每个颜色通道的绝对差异求和来计算的,而  $w(p, q)$  表示支持权重,定义如下:

$$w(p, q) = k \cdot \exp\left(-\left(\frac{\Delta c(p, q)}{\delta c} + \frac{\Delta l(p, q)}{\delta l}\right)\right) \quad (2)$$

其中,  $\Delta c(p, q)$  表示 CIE Lab 空间的色相相似性;  $\Delta l(p, q)$  表示欧几里得空间的位置距离;  $k, \delta c$  和  $\delta l$  均为可调节的常量。

由于异常行为包含许多与梯度相关的特征,通过考虑梯度上的相似性和邻近性对式(1)扩展,在  $S(p_1)$  和  $S(p_2)$  内,定义 3 个像素级匹配代价函数为:

$$e_j(q_1, q_2) = G_j(q_1) - G_j(q_2), j \in \{L, A, B\} \quad (3)$$

其中,  $G_j(q_1)$  表示通道  $j$  上像素点  $q_1$  处的梯度幅度。相应地,定义支持权重为  $p$  和  $q$  处梯度向量之间的欧氏距离:

$$w_j(p, q) = k_j \cdot \exp\left(-\left(\frac{\Delta c_j(p, q)}{\delta c_j} + \frac{\Delta l_j(p, q)}{\delta l_j}\right)\right), \quad j \in \{L, A, B\} \quad (4)$$

其中,  $\Delta l_j(p, q)$  表示  $p$  和  $q$  之间的欧氏距离,定义  $\Delta c_j(p, q)$  如下:

$$\Delta c_j(p, q) = \sqrt{(G_x_j(p) - G_x_j(q))^2 + (G_y_j(p) - G_y_j(q))^2} \quad (5)$$

基于格式塔原则,可分别利用  $w_j(p, q)$  和  $e(p, q)$  代替  $w(p, q)$  和  $e(p, q)$ ,并通过遍历图像  $I_1$  和  $I_2$  进行遍历,最后定义差异地图为  $\{D(I_1; I_2), D_L(I_1; I_2), D_A(I_1; I_2), D_B(I_1; I_2)\}$ 。

为了生成梯度级别的功率图,首先对第  $j$  通道

应用标准差为  $\sigma$  的高斯核  $G(\sigma)$  进行卷积操作,然后在第  $j$  通道图像和滤波后的图像之间产生一个残差堆叠,表示为:

$$R(I_j) = I_j - I_j \otimes G(\sigma) \quad (6)$$

在此基础上,建立梯度功率图如下:

$$P_j(I) = R(I_j) \otimes G(\sigma) \quad (7)$$

### 1.2 一致性表征

基于直方图对差异地图和能量地图进行一致性表征,直方图集合由 4 个直方图向量组成。考虑 2 个直方图  $H_1 = \{(\alpha_\mu, \Gamma_\mu) : \mu = 1, \dots, m\}$  和  $H_2 = \{(\beta_\nu, \Lambda_\nu) : \nu = 1, \dots, n\}$ , 其中  $\alpha$  和  $\beta$  表示索引值,  $\Gamma$  和  $\Lambda$  表示相应的频率。如果  $m = n$  并且  $\alpha_\mu$  与  $\beta_\mu$  对齐,则可以使用 Kulback-Liebler(KL) 距离作为 2 个直方图之间相似度的度量,即:

$$D_{KL}(H_1, H_2) = \sum_{\mu=1}^m \Gamma_\mu \log \frac{\Gamma_\mu}{\Lambda_\nu} \quad (8)$$

为了表示相似度,定义 2 个直方图之间的一致性算子为:

$$C((\alpha_\mu, \Gamma_\mu)(\beta_\nu, \Lambda_\nu)) = \Gamma_\mu \log \frac{\Gamma_\mu}{\Lambda_\nu} \quad (9)$$

然而,在差异直方图和能量直方图的索引数通常不相等。为了表示具有任意数量箱的 2 个任意直方图的相关性,考虑通过估算从  $H_1$  的索引值到  $H_2$

的 Earth-Mover(EM) 距离来对其进行估计,即:

$$F = [f_{\mu\nu}]^{m \times n} \quad (10)$$

此后,通过最小化  $\sum_{\mu} \sum_{\nu} f_{\mu\nu} d_{\mu\nu}$  来实现,其中  $d_{\mu\nu}$  表示 2 个索引值之间的  $L2$  范数。对于  $H_1$  上的每个  $\alpha_\mu$ , 存在  $n$  个对应的  $\{\beta_\nu : \nu = 1, \dots, n\}$  与之匹配,匹配概率为  $\{f_{\mu\nu} : \nu = 1, \dots, n\}$ , 而  $\sum_{\nu=1}^n f_{\mu\nu} = 1$ 。一致性算子计算点对点的对应关系可以被视为一个硬分配问题,对于  $H_1$  中的每个点,通过改进了  $H_1$  相对于  $H_2$  的分布相关性  $C(H_1 \parallel H_2)$ , 利用  $\Gamma_\mu \log \frac{\Gamma_\mu}{\Lambda_\nu}$  将  $H_1$  映射到  $H_2$ , 通过定义以下方程可实现基于概率的软分配:

$$C((\alpha_\mu, \Gamma_\mu)(\beta_\nu, \Lambda_\nu)) = \sum_{\nu=1}^n f_{\mu\nu} \cdot \Gamma_\mu \log \frac{\Gamma_\mu}{\Lambda_\nu} \quad (11)$$

## 2 对比型卷积神经网络搭建

为了学习一致性表征算子之间的关联特征,构建一个对比型 CNN: 利用 2 个并行的 GoogLeNet<sup>[11]</sup> 作为主干网络,进行深度特征提取,并延续到最后的池化层,如图 1 所示。

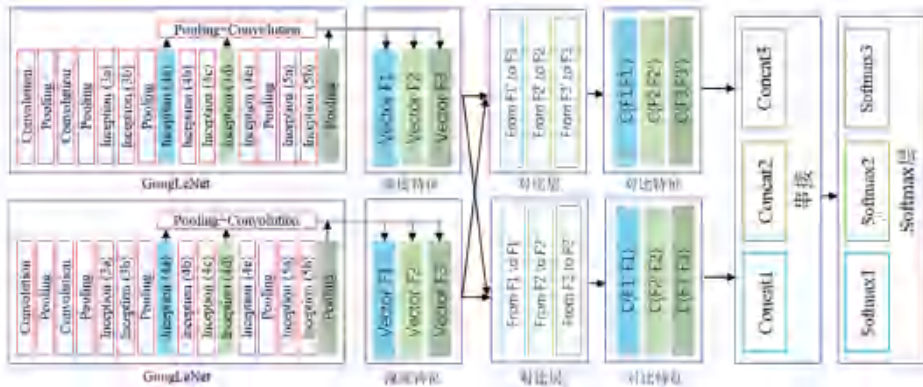


图 1 对比型卷积神经网络架构

Fig. 1 Architecture of contrast convolutional neural network

GoogLeNet 架构的 Softmax 回归有 3 个一维向量,根据原始的层名称,其中 2 个是从 inception4(a) 和 4(d) 层生成的,然后是一个池化、一个卷积和一个 dropout 层,另一个是从 inception5(b) 层生成的,接着是一个池化层。得到 3 对向量后,继续通过提出一个比较层来表示每对向量的比较关系。以任意一对向量为例,假设  $F = (f_1, f_2, \dots, f_n)$  和  $F' = (f'_1, f'_2, \dots, f'_n)$  表示这 2 个向量,其中  $n$  表示长度(对于 GoogLeNet 的 3 个向量为 1 024)。首先,将  $F$  和

$F'$  视为从 2 个并行 GoogLeNet 派生的 2 个分布,并假设存在一个联合分布  $C$  来确定彼此间的相互关系。考虑从  $f'_i$  到  $f_i$  的对数似然比  $\Lambda_i$ , 其中  $i \in [1, n]$ , 如果  $\Lambda_i > 0$ , 则表示对于第  $i$  个点,  $F$  分布模型比  $F'$  更好地拟合向量点。由于希望将这种关系更多地描述为与  $f_i$  相关,因此首先通过对  $f_i$  进行加权来定义评估因子  $e_i$ , 对此可表示为:

$$e_i = f_i \cdot \ln \frac{f_i}{f'_i} \quad (12)$$

因此,从  $F'$  到  $F$  的评估向量  $E(F, F')$  表示为  $(e_1, e_2, \dots, e_n)$ 。可知,得到的评估因子具有不对称性质,即从  $F'$  到  $F$  的因子不等于从  $F$  到  $F'$  的因子。假设  $\lambda$  表示从  $F$  中抽取  $C$  的总体概率,那么从  $F'$  中抽取  $C$  的概率为  $1 - \lambda$ ,在考虑总体分布之后,  $e_i$  被调整为  $e'_i$ ,可以计算如下:

$$e'_i = \lambda \cdot f_i \cdot \ln \frac{\lambda \cdot f_i}{(1 - \lambda) f'_i} \quad (13)$$

为了充分提取 2 个向量的相互关系信息,进一步考虑每一个点对的整体特征,并使用点对因子来表示其关系。对于第  $i$  个点,基于  $e_i$  通过比较  $f_i$  与配对特征  $(f_i + f'_i)$  来定义比较因子  $c_i$ ,数学公式为:

$$c_i = f_i \cdot \ln \frac{f_i}{f_i + f'_i} \quad (14)$$

同理,通过考虑  $C$  的总体概率分布,可将  $c_i$  调整为:

$$c'_i = \lambda \cdot f_i \cdot \ln \frac{\lambda \cdot f_i}{\lambda \cdot f_i + (1 - \lambda) f'_i} \quad (15)$$

因此,可定义对比层的表达式为:

$$C(F, F') = \lambda \cdot E \cdot (F, \lambda F + (1 - \lambda) F') \quad (16)$$

### 3 实验

#### 3.1 数据库

考虑到跌倒、弯腰捡东西和举手请求帮助是人机交互中最常见的异常情况,为了使机器人能够区分异常和正常活动,构建一个包含 2 000 个视频的异常检测数据库对所提算法进行验证。数据库分为 4 个类别:跌倒、弯腰、举手和正常行为,其中正常行为包括步行、烹饪、倒水等。首先,基于现有的基准数据库和收集 222 个跌倒、247 个弯腰和 208 个举手的样本。为了扩展数据并使之更接近真实世界的场景,通过在网络上使用文本搜索实现异常视频的检索。为了确保数据集的质量,移除恶作剧视频、新闻视频和包含合辑的视频,并删除模糊和含糊不清的视频。通过对收集到的视频进行清理和剪辑,由具有计算机视觉背景的 10 名标注者进行标注。对于正常视频,从 Toyota Smart home 数据集<sup>[12]</sup>中选择 600 个样本,分为 14 类,用来验证方法的泛化能力,因此,总共获得了 1 000 个异常视频和 1 000 个正常视频,示例如图 2 所示。

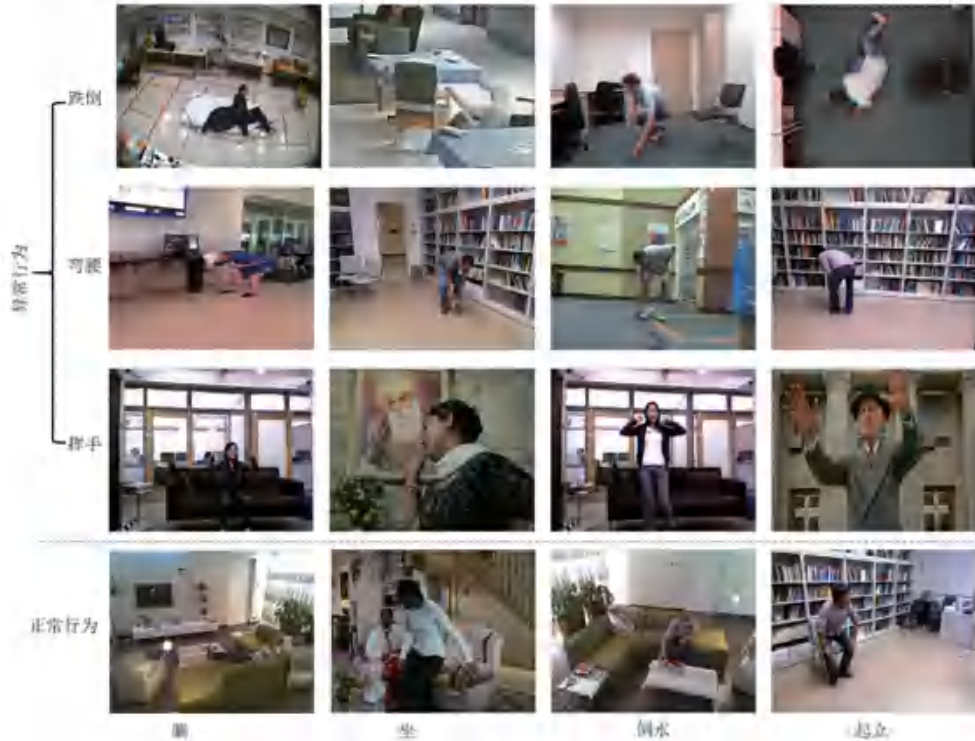


图 2 异常检测数据库样本示例

Fig. 2 Examples of anomaly detection database

#### 3.2 实验结果

在构建的数据集上使用所提方法进行异常检测的混淆矩阵结果见表 1。其中,混淆矩阵计算过程

为计算每个测试样本的识别结果,并对每个类别结果进行统计,将正确识别的结果与所有样本个数的比例作为矩阵对角线上的元素,将识别错误的识别

结果与总体样本个数比例作为矩阵非对角线上的元素。由结果可知,使用骨干网络 GoogLeNet 的 Loss1、Loss2 和 Loss3 得到的平均精度 (mean Average Precision, mAP) 分别为 76.0%、78.7% 和 79.4%,所提出的方法能够很好地区分“挥手”和“弯腰”异常活动,获得了 90.4% 和 86.7% 的 AP。

表1 异常检测混淆矩阵 (Loss/Loss2/Loss3)

Table 1 Confusion matrix of anomaly detection (Loss/Loss2/Loss3)

检测内容	跌倒	弯腰	挥手	正常行为
跌倒	<b>68.9/71.8/7.0</b>	5.5/4.9/5.2	0.6/0.4/0.5	25.0/22.9/21.7
弯腰	4.0/2.7/2.3	<b>85.2/87.1/86.7</b>	1.3/1.0/0.9	9.5/9.2/10.1
挥手	1.1/0.8/0.8	1.2/0.7/0.6	<b>86.3/87.8/90.4</b>	11.4/10.7/8.2
正常行为	13.9/10.5/10.9	3.4/2.9/3.1	2.0/1.4/1.5	<b>80.7/85.2/84.5</b>
总体精度	76.0/78.7/79.4			

为了验证差异一致性表征方法的有效性,将其与4种用于时空特征提取的流行方法进行比较:图像差分<sup>[13]</sup>、光流<sup>[14]</sup>、密集自适应自相关 (Dense Adaptive Self Correlation, DASC)<sup>[15]</sup> 和 SIFT flow<sup>[11]</sup>, 对比结果见表2。为了与对比型 CNN 的输入对齐<sup>[16]</sup>,邻域中提取的特征按顺序堆叠成前向和后向,并互相不重叠。基于图像差分的方法在数据集上获得了 64.4% 的 mAP,与关注区域时空变化的光流相比,差距为 7.0%。在4种特征中,SIFT flow<sup>[11]</sup> 由于特征的稀疏性,异常检测的性能最差 (mAP = 61.7%)。DASC<sup>[15]</sup> 通过计算基于窗口的补丁之间的相似性来设计局部描述符,在数据库上获得了 70.3% 的 mAP。与捕获图像对应关键外观线索不同,所提方法考虑了时间差异和序列间的相关特性,并显示出其强大的特征表示能力。

表2 不同异常检测方法结果对比

Table 2 Comparison of results of different anomaly detection methods

方法	mAP/%
SIFT flow <sup>[11]</sup>	61.7
图像差分 <sup>[13]</sup>	64.4
光流 <sup>[14]</sup>	68.6
DASC <sup>[15]</sup>	70.1
Two-stream network <sup>[17]</sup>	74.8
I3D <sup>[18]</sup>	76.1
TPN <sup>[19]</sup>	77.8
TDN <sup>[20]</sup>	78.5
本文方法	79.4

同时,为了验证所提对比型 CNN 的有效性,在数据库上分别采用 Two-stream network<sup>[17]</sup>、I3D<sup>[18]</sup>、

然而,“跌倒”并没有被完全识别 (mAP = 72.6%)。此外,Loss1 在大多数情况下表现比 Loss2 和 Loss3 差,而 Loss2 在正常行为和“弯腰”的检测结果比 Loss3 好,大多数“跌倒”识别失败可以归因于“跌倒”和正常活动之间的相似性较大,例如“坐下”和“躺下”。

TPN<sup>[19]</sup> 和 TDN<sup>[20]</sup> 的默认设置进行实验,并与所提方法进行比较,具体见表2。由结果可知,Two-stream network<sup>[17]</sup> 通过基于 VGG 和 FasterRCNN 的卡尔曼滤波器提取数据历史图像和 RGB 帧的深度特征,获得了 74.9% 的 mAP。I3D<sup>[18]</sup> 得益于其优越的三维卷积特征提取能量,其 mAP 为 76.3%。通过在时间上构建金字塔或差分模块,TPN<sup>[19]</sup> (mAP = 77.8%) 和 TDN<sup>[20]</sup> (mAP = 78.5%) 克服了 Two-stream network<sup>[17]</sup> 和 I3D<sup>[18]</sup> 时空表征能力弱的缺点,异常检测效果较好,所提方法相对于其他方法精度更高。

## 4 结束语

面向人机交互中的异常检测问题,提出了一种融合差异一致性表征的对比型卷积神经网络,解决交互对象存在模糊性和时空不完整性的难点问题。对比型卷积神经网络构建了差异地图和能量地图对行为序列间的颜色和梯度特征进行提取,并通过结合 KL 距离和 EM 概率实现时空特征的一致性表示;在骨架网络的基础上,利用联合分布概率对深度特征进行语义描述,实现异常行为的全局特征表征。在现有数据库基础上整合了适用于人机交互的异常检测数据库,并对所提方法进行了验证和比较。实验结果显示对比型卷积神经网络 mAP 达到 79.4%,证明了其优越性。在后续研究中,将探索在更多应用场景下算法的有效性,并验证模型的泛化能力。

## 参考文献

- [1] 郭惠,耿宝光. 基于机器学习算法的人体步态异常状态检测与识别方法[J]. 智能计算机与应用, 2023, 13(10): 25-28.

- [2] NANDAKUMAR C S, MITCHELL D, ERDEN S M, et al. Anomaly detection methods in autonomous robotic missions [J]. *Sensors*, 2024, 24(4): 1330.
- [3] 石洋宇, 左景, 谢承杰, 等. 多尺度融合与FMB改进的YOLOv8异常行为检测方法[J]. *计算机工程与应用*, 2024, 60(9): 101-110.
- [4] AMINE E, ROHAN S, CHRISTOPHER A, et al. Semantic anomaly detection with large languagemodels [J]. *Autonomous Robots*, 2023, 47(8): 1035-1055.
- [5] 毛国君, 吴星臻, 邢树礼. 基于多尺度流模型的视觉异常检测研究[J]. *自动化学报*, 2024, 50(3): 640-648.
- [6] PARK D, KIM H, KEMP C C. Multimodal anomaly detection for assistive Robots[J]. *Autonomous Robots*, 2019, 43: 611-629.
- [7] 彭煜祺, 魏巍, 陈灯, 等. 基于视觉的工业机器人异常动作检测方法研究[J]. *武汉工程大学学报*, 2021, 43(4): 462-467.
- [8] 赵廉, 周雷, 郭育恒, 等. 工厂场景中的异常行为检测[J]. *软件导刊*, 2024, 23(1): 57-62.
- [9] 魏麟, 谭任翔, 何峻毅, 等. 基于改进YOLOv5的飞行员异常行为识别方法[J]. *航空计算技术*, 2023, 53(6): 20-24.
- [10] 余胤翔, 黄胤政, 李宁, 等. 基于深度学习的社区安防异常行为检测研究[J]. *智能计算机与应用*, 2023, 13(6): 189-192.
- [11] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2015: 1-9.
- [12] DAS S, DAI R, KOPERSKI M, et al. Toyota smart home: Real- world activities of daily living [ C ]// *Proceedings of International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2019: 838-842.
- [13] LIU Ce, YUEN J, TORRALBA A. SIFT Flow: Dense correspondence across scenes and its applications [ J ]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(5): 978-994.
- [14] 陈琳琳, 邓华军, 张殿喜, 等. 基于图像差分及邻域特性的航空磁环缺陷检测[J]. *科技创新与应用*, 2023, 13(23): 50-53.
- [15] 付婧祎, 余磊, 杨文, 等. 基于事件相机的连续光流估计[J]. *自动化学报*, 2023, 49(9): 1845-1856.
- [16] KIM S, MIN D, HAM B, et al. Dasc: Robust dense descriptor for multi-modal and multi-spectral correspondence estimation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(9): 1712-1729.
- [17] 王梓舟, 周新志, 严华. 基于双流结构的异常行为检测模型[J]. *计算机应用与软件*, 2022, 39(2): 188-193.
- [18] 程相贵, 刘钊, 郭放. 结合双流13D和注意力机制的视频异常事件检测[J]. *信息与电脑*, 2022, 34(24): 65-68.
- [19] YANG Ceyuan, XU Yinghao, SHI Jianping, et al. Temporal pyramid network for action recognition [ C ]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2020: 588-597.
- [20] WANG Limin, TONG Zhan, JI Bin, et al. TDN: Temporal difference networks for efficient action recognition [ C ]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2021: 1895-1904.