监控视频中采用深度支持向量数据描述的异常检测

李芳丽

(江西科技学院 信息工程学院，江西 南昌 330098)

摘　要：由于异常定义的模糊性,异常数据的稀少性,以及复杂的环境背景和人类行为,视频异常检测是计算机视觉领域中一大难题。现有基于深度学习的异常检测方法往往是利用训练好的网络提取特征或者是基于现有网络结构的,而并非针对于异常检测这个目标而设计网络的。提出一种基于深度支持向量数据描述 (Deep Support Vector Data Description, DSVDD) 的方法,通过学习一个深度神经网络,使得输入的正常样本空间能够映射到最小超球面。通过 DSVDD,不仅能找到最小尺寸的数据超球面以建立 SVDD,而且可以学习有用的数据特征表示以及正常模型。在测试时,映射在超球面内的样本被判别为正常,而映射在超球面外的样例判别为异常。提出的方法在 CUHK Avenue 和 ShanghaiTech Campus 数据集上分别取得了 87.4% 和 74.5% 的帧级 AUC,检测结果优于现有的最新方法。

关键词：视频监控；　异常检测；　深度支持向量数据描述；　深度学习

中图分类号：TP391.4　　**文献标志码：**A　　**DOI：**10.3788/IRLA20210094

Anomaly detection based on deep support vector data description under surveillance scenarios

Li Fangli

(School of Information Engineering, Jiangxi University of Technology, Nanchang 330098, China)

Abstract: Due to the ambiguity of anomaly definitions, the scarcity of anomalous data, as well as the complex environmental background and human behavior, video anomaly detection has always been a challenging problem in the field of computer vision. Existing anomaly detection methods based on deep learning often use a trained network to extract features or are trained based on the existing network structure, instead of designing a network for the goal of anomaly detection. In this paper, a new anomaly detection method—Deep Support Vector Data Description (DSVDD) was introduced, which was trained on an anomaly detection based objective. According to DSVDD, not only the smallest size data hypersphere coule be found to establish SVDD, but also useful data feature representations and normal models could be learned. Then, in the testing stage, the samples mapped inside the hypersphere were judged as normal, while the samples mapped outside the hypersphere were judged as abnormal. The method proposed in this paper achieves 87.4% and 74.5% frame-level AUC on the CUHK Avenue and ShanghaiTech Campus datasets, respectively, which outperforms existing state-of-the-art approaches.

Key words: video surveillance; anomaly detection; Deep Support Vector Data Description(DSVDD); deep learning

收稿日期:2021−02−08；　**修订日期:**2021−02−24

基金项目:江西省教育厅科学技术研究资助项目 (GJJ180975);江西科技学院质量工程项目 (江科发 [2018]48 号)

作者简介:李芳丽,讲师,硕士,研究方向为图像信息处理、计算机视觉等。

0 引 言

为了提升公共生活和资产的安全性,视频监视系统已经广泛部署于商场、医院、银行、街道、教育机构,城市行政办公室和智慧城市等公共场所中。在大多数情况下,如何及时、准确地检测视频异常事件是社会公共安全风险防控的主要目标。视频异常事件定义为视频中不符合正常模式的异常或不规则模式。这些事件往往包括打架、骚乱、违反交通规则、踩踏、持械以及遗弃行李等行为。近年来,异常事件检测已经逐渐成为计算机视觉和模式识别领域的研究热点,其主要难点是异常定义的模糊性,异常数据的稀少性,以及复杂的环境背景和人类行为。

概括地说,当前有关视频异常检测的大多数研究工作可以分为两个步骤,例如特征提取和正常模型训练[1]。特征提取可以通过手工技术或自动特征提取技术 (表示学习或基于深度学习的特征) 来实现。在正常模型训练则是采用正常样本进行学习,然后将不符合所学习模型的样本判定为异常事件。那么,按照特征进行分类可以分为 3 类不同的方法[2]: (1) 基于轨迹的方法[3]: 这类方法通过对目标进行跟踪以获取轨迹特征,但是在密集场景对目标跟踪是一大难题; (2) 基于全局特征的方法[4−5]: 这类方法将视频帧作为一个整体,提取一些底层或者中层特征如时空梯度、光流等,在中等拥挤和密集环境中均有效; (3) 基于网格特征的方法[6]: 这类方法往往通过密集采样将视频帧划分为多个小网格,然后对单个网格提取底层特征,因为每个网格都可以单独被评价。按照采用不同的正常模型训练方式,也可以分为 3 种不同的方法: (1) 基于聚类的方法[7]: 这类方法往往基于一个假设,正常样本属于一个类别或者距离聚类中心较近,而异常样本则不属于任何类别或者远离聚类中心,然后针对正常样本进行聚类以建立模型; (2) 基于稀疏重构的方法[8−9]: 这类方法假设是,正常模式的稀疏线性组合能够以最小的重构误差表示正常活动,而由于训练数据集中不存在异常活动,因此能够以较大的重构误差表示异常模式; (3) 基于概率模型的方法[10]: 这种方法认为正常样本符合某个概率分布,而异常样本则不符合该分布。

近期,深度学习的最新进展证明了基于深度学习的方法在许多计算机视觉应用中的明显优势[11]。作为计算机视觉中的任务之一,视频异常检测也不例外。与传统基于手工特征方法不同的是,深度学习方法往往采用预训练网络对视频进行高层的特征提取,或者直接根据正常模型采用现有网络结构建立端对端的异常检测模型。对于前一种思路[12−13],和传统的异常事件检测两个步骤没有太多差别。而对于后一种思路[14−19],特征提取和模型建立两个步骤往往是在一个深度网络中联合优化的,因为能够实现二者最优,这也是深度学习方法的一大优势。这些端对端的深度网络包括深度自编码 (Auto-Encoder, AE)、深度孪生网络 (Deep Siamese Network, DSN)、生成对抗网络 (Generative Adversarial Nets, GAN)[15,20−24]。然而,这些网络模型往往是针对其他任务如生成模型、压缩等,不是针对异常检测任务而单独设计的。

文中在深度学习的框架下,面向异常检测任务,基于深度支持向量数据描述 (Deep Support Vector Data Description, DSVDD),提出一种新的异常检测方法。通过学习 DSVDD,能够找到建立 SVDD 的最小数据超球面,并且以获得输入数据的特征表示并且学习得到正常模型。为此,DSVDD 采用了经过联合训练的深度神经网络,将正常样本数据映射到最小体积的超球中。那么在测试时,映射在超球面内的样本被判别为正常,而映射在超球面外的样例判别为异常。文中方法将 RGB 图和光流图组成一个 6 通道数据直接输入到一个 DSVDD 模型中,即能同时检测外观异常和运动异常。在 Avenue[9] 和 ShanghaiTec[21] 两个公开数据集上的实验结果表明,文中提出的方法检测效果优异,超过现有技术发展水平。

1 算法原理

文中所提出的方法总流程如图 1 所示。在训练阶段,训练样本的 RGB 图像和光流图被密集采样,然后合并成一个 6 通道的数据,训练深度支持向量数据描述模型; 在测试阶段,同样获得待测视频帧的 RGB 图像和光流图组成的 6 通道数据,输入到学习好的深度支持向量数据描述模型后,判定该区域是否异常。在本节中,首先简要介绍支持向量数据描述原理,然后在此基础上重点描述基于深度支持向量数据描述的视频异常事件训练和测试过程。

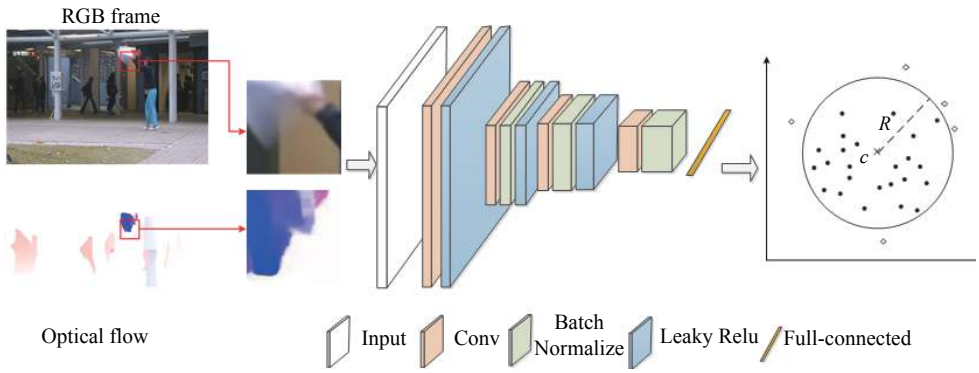


图 1 基于深度支持向量数据描述的异常检测流程

Fig.1 The flow chart of video anomaly detection based on DSVDD

1.1 支持向量数据描述

支持向量数据描述 (Support Vector Data Description, SVDD) 是一种基于边界数据 (支持向量) 的描述方法, 其目标是寻求一个包含所有或几乎所有的训练样本且体积最小的超球体 (中心为$\boldsymbol{c}\in F_k$和半径为$R>0$)。实际上, SVDD 优化问题可以转化为:

$$\begin{aligned}&\min_{R,\ W} R^2+\frac{1}{vn}\sum_i \xi_i\\&s.t.\ \|\varphi_k(x_i)-\boldsymbol{c}\|^2_{F_k}\leqslant R^2+\xi_i,\xi_i\geqslant 0\end{aligned}\tag{1}$$

式中: 松弛变量$\xi_i\geqslant 0$允许一个软边界和超参数$v\in(0,1]$控制着惩罚项与超球的体积边之间的平衡。因此, 落到超球面外的点, 例如$\|\phi_k(x_i)-c\|^2_{F_k}>R^2$, 则判断注定是异常的。SVDD 已被广泛应用于异常检测、人脸识别、语音识别、图像恢复和医学成像等领域[19]。

1.2 深度支持向量数据描述

DSVDD 通过学习一个权值为W的深度神经网络$\varphi(\cdot;W)$, 使得输入的正常样本空间$X\subseteq R^d$能够映射到一个中心为$\boldsymbol{c}$和半径为R最小超球面, 正常样例的映射在超球面内, 而异常样例的映射在超球面外。

具体来说, 对于样本区域输入空间$X\subseteq\mathbb{R}^d$和输出空间$F\subseteq\mathbb{R}^p$, 采用一个包含$L\in\mathbb{N}$个隐藏层的神经网络能够将输入空间投影到输出空间$X\to F$, 其中$W=\{W^1,W^2,\cdots,W^L\}$分别是隐藏层$\ell=\{1,2,\cdots,L\}$的权值。因此, $\phi(x;W)\in F$是输入样本$\boldsymbol{x}\in X$在神经网络ϕ下的特征表示, 那么 DSVDD 方法的目标就是联合优化网络权值W和输出空间符合中心为$\boldsymbol{c}$和半径为R 最小超球面约束。那么, 给定训练样本$D_n=\{x_1,x_2,\cdots,x_n\}$, DSVDD 的软边界目标函数为:

$$\begin{aligned}&\min_{R,\ W} R^2+\frac{1}{vn}\sum_{i=1}^{n}\max\left\{0,\|\varphi(x_i;W)-\boldsymbol{c}\|^2-R^2\right\}+\\&\frac{\lambda}{2}\sum_{\ell=1}^{L}\|W\|_F^2\end{aligned}\tag{2}$$

对于公式 (2) 来说, 在 SVDD 方法中, 最小化R^2表示最小化超球的体积。第二项是通过神经网络进行映射到超球外的惩罚项, 例如那些距离超球中心$\|\phi(x_i;W)-c\|$大于半径R的样本。超参数$v\in(0,1]$控制着超球的体积与边界的偏离之间的平衡, 即允许将某些点映射到球体外部。最后一项是网络参数权值W衰减的正则化项, 其中$\lambda>0$, 且$\|\cdot\|_F$表示 Frobenius 范数。

对于公式 (2) 的优化使得网络能够学习权值W, 使得数据点能够紧密地投影到超球中心$\boldsymbol{c}$附近。为此, 深度网络必须提取数据变化的共同因素。实际上, 正常样本往往能够紧密映射到超球中心$\boldsymbol{c}$附近, 而异常样本则被映射到更远离中心或超球面之外。通过这种方式获得了对正常模型的紧凑描述。

在实际任务中, 往往假设训练样本D_n均为正常样本, 那么目标函数可以简化为一个单类分类问题如下:

$$\min_{W}\frac{1}{n}\sum_{i=1}^{n}\|\phi(x_i;W)-\boldsymbol{c}\|^2+\frac{\lambda}{2}\sum_{\ell=1}^{L}\|W\|_F^2\tag{3}$$

此时, DSVDD 简单地采用二次损失来惩罚每个深度网络表示$\phi(x_i;W)$到$\boldsymbol{c}$的距离。第二项是网络参数权值W衰减的正则化项, $\lambda>0$。公式 (3) 也可以看成是以中心$\boldsymbol{c}$为中心找到最小体积的超球面。但是和公式 (2) 采用软边界不同, 公式 (3) 通过最小化所有数据表示到中心的平均距离来收缩球体, 而不是通过直

接惩罚半径和落在球体之外的数据表示来收缩。同样, 为了将样本映射到尽可能靠近超球中心c的位置, 深度神经网络必须提取变化的公共因子。

可以通过常见的反向传播方法 (例如随机梯度下降法) 优化 DSVDD 中神经网络的权值W。由于网络权值W和超球半径 R 尺度不同, 采用一种学习率无法优化 DSVDD。因此, 需要以交替最小化/块坐标下降法交替优化网络权值W和超球半径 R。具体来说, 首先把超球半径 R 固定, 迭代k次训练神经网络权值W, 然后再通过线搜索求解超球半径 R, 这样接替迭代, 具体过程可以参考文献 [20]。

1.3 测 试

给定测试样本区域$\boldsymbol{x}' \in X$, 可以根据其通过神经网络$\phi(x';W^*)$后到超球中心$\boldsymbol{c}$的距离计算异常得分:

$$s(x') = \|\phi(x';W^*) - \boldsymbol{c}\|^2 \tag{4}$$

式中: W^*为已训练好的网络模型参数。值得注意的是, 网络参数能够使W^*完全表征 DSVDD 模型, 并且无需存储任何数据即可进行预测, 因此 DSVDD 具有非常低的存储复杂度, 测试时计算复杂度较小。

为了推断测试样本区域的是否为异常样本, 可以通过对$s(x')$设定阈值来进行判断:

$$s(x') \underset{\text{normal}}{\overset{\text{abnormal}}{\gtrless}} \theta \tag{5}$$

式中: θ为决定文中检测方法的敏感度的阈值。

2 实 验

2.1 数据集

文中在两个公开的数据集上评估了 DSVDD 方法的性能, 它们分别是: Avenue 数据集[9] 和 ShanghaiTech 数据集[21]。Avenue 数据集是用于视频异常检测的最广泛使用的基准之一。它包含 16 个训练视频片段和 21 个测试视频片段, 其中有 47 个在香港中文大学街道上发生的异常事件。每个视频长约 1 min, 分辨率为 640×360。正常事件是在街道上行走, 异常事件包括奔跑、游荡和投掷等。ShanghaiTech Campus 数据集[21] 是新提出的用于视频异常检测的最大数据集之一。与其他数据集不同, 在该数据集中视频剪辑来自 13 个不同的摄像机, 这些摄像机具有不同的照明条件和摄像机角度。它有 330 个训练视频片段和 107 个包含 130 个异常事件的测试视频片段。视频帧的分辨率为 856×480。该数据集中异常事件包括追逐和吵闹等。

2.2 评价指标

根据先前的工作[15], 文中计算帧级接收器工作特性 (ROC) 曲线, 并使用曲线下的面积 (AUC) 分数作为评估指标, 较高的 AUC 分数表示更好的异常检测性能。若视频帧中有一个区域判断为异常, 则该帧判断为异常。课题组首先获得所有视频帧的异常分数, 然后计算帧级 AUC 分数。

2.3 补充细节

对于两个数据集, 每帧都被调整为大小 320×240, 光流图像由参考文献 [22] 中提供的 RAFT 光流法通过在 things 数据集上预训练的网络计算得到。将原始视频帧和计算获得的光流图合并成一个 6 通道的数据, 然后按照尺寸 20×20 的大小裁剪为 16×12 个网格图像, 输入到 DSVDD 中进行训练和预测。DSVDD 中深度神经网络部分按照 Conv (16, 3×3)-Leaky ReLU-ConvTran (32, 3×3)-BN-Leaky ReLU-ConvTran (64, 3×3)- BN-Leaky ReLU –FullyConnectd64 的结构。训练阶段批量大小设置为 128, 初始学习率为 0.000 3, weight decay 为 0.000 1, 并训练 1 000 次迭代。DSVDD 模型是在配备 Intel I7-9700K CPU, 16 GB RAM 和 NVIDIA 2080ti GPU 的计算机上使用 PyTorch 实现的。

2.4 实验结果

此节将提出的 DSVDD 与仅使用正例样本训练的几种最新的方法取得的结果进行了比较, 这些方法包括 Conv-AE[15]、Stacked RNN[21]、Unmasking[23]、Davide et al.[24]、Object-centric auto-encoder[25] 和 MemAE[26]、New Baseline[27]。表 1 中列出了这些方法在两个数据集上的帧级异常检测的评估结果。

在 CUHK Avenue 数据集上, 文中提出的 DSVDD 方法优于其他方法取得的结果, 其 AUC 得分为 87.4%, 比 2018 年提出的作为基线的方法[27] 高出 2.3%。就目前所知, 就此数据集中所有测试视频的帧级 AUC 得分而言, 文中提出的 DSVDD 取得了最好的结果。值得注意的是, Object-centric auto-encoder[25] 方法在他们的论文中取得了 89.3% 的帧级 AUC, 但是这个通过他们的论文中的不同指标计算得出的, 通过实际计算 Object-centric auto-encoder[25] 方法取得的帧级 AUC 得分应为 86.5%, 比文中提出的方法低 0.9%。

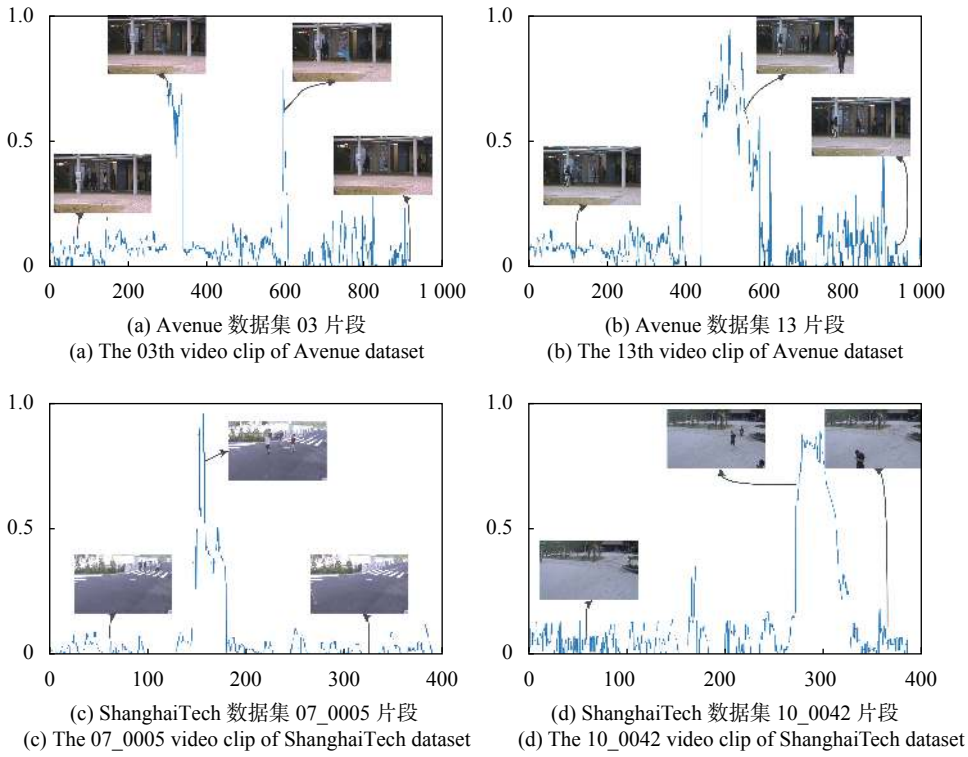
表 1 帧级异常检测结果 AUC 得分

Tab.1 AUC scores of the anomaly detection results

Method	Avenue	ShanghaiTech Campus
Conv-AE[15]	80.0%	60.9%
Stacked RNN[21]	81.7%	68.0%
Unmasking[23]	80.6%	—
Davide et al.[24]	—	72.8%
Object-centric auto-encoders[25]	86.5%	78.5%
MemAE[26]	83.3%	72.2%
New Baseline[27]	85.1%	72.8%
Ours	87.4%	74.5%

在 ShanghaiTech Campus 数据集上，文中提出的方法 DSVDD 实现了 74.5% 的帧级 AUC 评分，比 2018 年提出的作为基线的方法[27] 高出 1.7%，仅次于 Object-centric auto-encoder[25] 方法取得的 78.5%。Object-centric auto-encoder[25] 方法采用的是基于对象检测的方法进行异常检测，其性能在很大程度上取决于其对象检测算法的输出。因此，基于检测的方法无法确定之前未出现的异常事件，而这在异常检测中经常出现。相似地，MemAE 方法[26] 需要在预训练的姿势估计器的帮助才能取得较好结果，因此仅限于检测与人有关的异常事件。相比之下，文中提出的 DSVDD 方法没有这种局限性，并且在应用于各种场景时都非常可靠。明显地，除了这两种特殊限定的方法外，文中提出的 DSVDD 方法在帧级 AUC 上至少领先其他方法 1.7%。

在图 2 中，展示了文中提出的方法中异常分数曲线的一些示例，并给出了一些具有正常或异常事件的关键帧。其中，横坐标为视频帧数，纵坐标异常分数已经归一化到 1。可以看出，在两个数据集中，文中提出的方法可以正确区分正常和异常事件。如果突发异常事件，如图 2(a) 中奔跑，异常分数将急剧增加，如果异常事件是缓慢发生的，如图 2(b) 中缓慢走向镜头，异常分数将逐渐增加。如果导致异常的对象在摄像头视野中消失，则异常得分会迅速降低到接近 0。



(a) Avenue 数据集 03 片段
(a) The 03th video clip of Avenue dataset

(b) Avenue 数据集 13 片段
(b) The 13th video clip of Avenue dataset

(c) ShanghaiTech 数据集 07_0005 片段
(c) The 07_0005 video clip of ShanghaiTech dataset

(d) ShanghaiTech 数据集 10_0042 片段
(d) The 10_0042 video clip of ShanghaiTech dataset

图 2 部分检测结果示例

Fig.2 Examples of the detection results

3 结 论

在文中提出了 DSVDD，这是一种基于深度学习的视频异常检测方法。DSVDD 可以看成是深度学习方法和支持向量数据描述方法的结合，采用了经过联合训练的深度神经网络，将正常样本数据映射到最小

体积的超球中。那么在测试时，映射在超球面内的样本被判别为正常，而映射在超球面外的样例判别为异常。在两个公共数据集上的大量实验结果表明，文中提出的方法明显优于现有方法，这证明了文中提出的异常检测方法的有效性。今后将在保证算法准确性的基础上，降低计算复杂度，重点是提高算法的实时性能，以更好地应用于实际场景。

参考文献：

[1] Luo W, Liu W, Lian D, et al. Video anomaly detection with sparse coding inspired deep neural networks [J]. ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, 2019, 27: 1-15.

[2] Nayak R, Pati U C, Das S K. A comprehensive review on deep learning-based methods for video anomaly detection [J]. ***Image and Vision Computing***, 2020, 106: 104078.

[3] Cosar S, Donatiello G, Bogorny V, et al. Toward abnormal trajectory and event detection in video surveillance [J]. ***IEEE Transactions on Circuits and Systems for Video Technology***, 2016, 27(3): 683-695.

[4] Xu K, Jiang X, Sun T. Anomaly detection based on stacked sparse coding with intraframe classification strategy [J]. ***IEEE Transactions on Multimedia***, 2018, 20(5): 1062-1074.

[5] Cheng K W, Chen Y T, Fang W H. Video anomaly detection and localization using hierarchical feature representation and gaussian process regression[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 2909-2917.

[6] Leyva R, Sanchez V, Li C T. Video anomaly detection with compact feature sets for online performance [J]. ***IEEE Transactions on Image Processing***, 2017, 26(7): 3463-3478.

[7] Huang Lehong, Cao Lihua, Li Ning, et al. A state perception method for infrared dim and small targets with deep learning [J]. ***Chinese Optics***, 2020, 13(3): 527-536. (in Chinese)

[8] Xu D, Yan Y, Ricci E, Sebe N. Detecting anomalous events in videos by learning deep representations of appearance and motion [J]. ***Computer Vision and Image Understanding***, 2017, 156: 117-127.

[9] Lu C, Shi J, Jia J. Abnormal event detection at 150 FPS in MATLAB[C]//Proceedings of IEEE International Conference on Computer Vision, 2013: 2720-2727.

[10] Weixin L, Mahadevan V, Vasconcelos N. Anomaly detection and localization in crowded scenes [J]. ***IEEE Transactions on Pattern Analysis and Machine Intelligence***, 2014, 36(1): 18-32.

[11] Zhou Hongqiang, Huang Lingling, Wang Yongtian. Deep learning algorithm and its application in optics [J]. ***Infrared and Laser Engineering***, 2019, 48(12): 1226004. (in Chinese)

[12] Ravanbakhsh M, Nabi M, Mousavi H, et al. Plug-and-play CNN for crowd motion analysis: An application in abnormal event detection[C]//IEEE Winter Conference on Applications of Computer Vision, 2018: 1689-1698.

[13] Zhang Xiaorong, Hu Bingliang, Pan Zhibing, et al. Tensor representation based target detection for hyperspectral imagery [J]. ***Optical and Precision Engeering***, 2019, 27(2): 488-498. (in Chinese)

[14] Zhang Dongge, Fu Yutian. One class support vector machine used for blind pixel detection [J]. ***Infrared and Laser Engineering***, 2018, 47(4): 0404001. (in Chinese)

[15] Hasan M, Choi J, Neumanny J, et al. Learning temporal regularity in video sequences[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[16] Ramachandra B, Jones M J, Vatsavai R R. Learning a distance function with a Siamese network to localize anomalies in videos[C]//The IEEE Winter Conference on Applications of Computer Vision, 2020: 2598-2607.

[17] Ravanbakhsh M, Nabi M, Mousavi H, et al. Abnormal event detection in videos using generative adversarial nets[C]//IEEE International Conference on Image Processing, 2017.

[18] Li Angze, Wang Xianshuang, Xu Xiangjun, et al. Fast classification of tobacco based on laser-induced breakdown spectroscopy [J]. ***Chinese Optics***, 2019, 12(5): 1139-1146. (in Chinese)

[19] Yu Liandong, Chang Yaqi, Zhao Huining, et al. Method for improving positioning accuracy of robot based on support vector regression [J]. ***Optics and Precision Engineering***, 2020, 28(12): 2646-2654. (in Chinese)

[20] Luo W, Liu W, Gao S. A revisit of sparse coding based anomaly detection in stacked rnn framework[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 341–349.

[21] Ruff L, Vandermulen R A, Gornitz N, et al. Deep one-class classification[C]//Proceedings of the 35 th International Conference on Machine Learning, 2018.

[22] Ionescu R T, Smeureanu S, Alexe B, et al. Unmasking the abnormal events in video[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2895–2903.

[23] Teed Z, Deng J. RAFT: Recurrent all pairs field transforms for optical flow[C]//European Conference on Computer Vision,

2020: 402-419.

[24] Abati D, Porrello A, Calderara S, et al. Latent space autoregression for novelty detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 481-490.

[25] Ionescu R T, Khen, Geogrescu M I, et al. Object-centric auto-encoders and dummy anomalies for abnormal event detection in video[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2019: 7842-7851.

[26] Morais R, Le V, Tran T, et al. Learning regularity in skeleton trajectories for anomaly detection in videos[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 11996-12004.

[27] Liu W, Luo W, Lian D, et al. Future frame prediction for anomaly detection –a new baseline[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6536-6545.