



# 智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

## 深度自编码与自更新稀疏组合的异常事件检测算法

王倩倩, 苗夺谦, 张远健

引用本文:

王倩倩, 苗夺谦, 张远健. 深度自编码与自更新稀疏组合的异常事件检测算法[J]. 智能系统学报, 2020, 15(6): 1197–1203.

WANG Qianqian, MIAO Duoqian, ZHANG Yuanjian. Abnormal event detection method based on deep auto-encoder and self-updating sparse combination[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(6): 1197–1203.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202007003>

## 您可能感兴趣的其他文章

### 基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects

智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

### 加权PageRank改进地标表示的自编码谱聚类算法

An autoencoder spectral clustering algorithm for improving landmark representation by weighted PageRank

智能系统学报. 2020, 15(2): 302–309 <https://dx.doi.org/10.11992/tis.201904021>

### 多标记学习自编码网络无监督维数约简

Unsupervised dimensionality reduction of multi-label learning via autoencoder networks

智能系统学报. 2018, 13(5): 808–817 <https://dx.doi.org/10.11992/tis.201804051>

### 高斯核函数卷积神经网络跟踪算法

Convolutional neural network tracking algorithm accelerated by Gaussian kernel function

智能系统学报. 2018, 13(3): 388–394 <https://dx.doi.org/10.11992/tis.201612040>

### 基于深度学习的视频预测研究综述

Review of deep learning-based video prediction

智能系统学报. 2018, 13(1): 85–96 <https://dx.doi.org/10.11992/tis.201707032>

### 深度学习方法研究新进展

Progress report on new research in deep learning

智能系统学报. 2016, 11(5): 567–577 <https://dx.doi.org/10.11992/tis.201511028>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.202007003

# 深度自编码与自更新稀疏组合的异常事件检测算法

王倩倩, 苗夺谦, 张远健

(同济大学嵌入式系统与服务计算教育部重点实验室, 上海 201804)

**摘要:** 基于深度学习的异常检测算法输入通常为视频帧或光流图像, 检测精度和速度较低。针对上述问题, 提出了一种以运动前景块为中心的卷积自动编码器和自更新稀疏组合学习 (convolutional auto-encoders and self-updating sparse combination learning, CASSC) 算法。首先, 采用自适应混合高斯模型 (gaussian mixture model, GMM) 提取视频前景, 并以滑动窗口的方式根据前景像素点占比过滤噪声; 其次, 构建 3 个卷积自动编码器提取运动前景块的时空特征; 最后, 使用自更新稀疏组合学习对特征进行重构, 依据重构误差进行异常判断。实验结果表明, 与现有算法相比, 该方法不仅有效地提高了异常事件检测的准确性, 且可以满足实时检测需求。

**关键词:** 深度学习; 稀疏组合; 自动编码器; 自更新; 异常事件检测; 卷积神经网络; 无监督学习; 稀疏学习

**中图分类号:** TP391    **文献标志码:** A    **文章编号:** 1673-4785(2020)06-1197-07

**中文引用格式:** 王倩倩, 苗夺谦, 张远健. 深度自编码与自更新稀疏组合的异常事件检测算法 [J]. 智能系统学报, 2020, 15(6): 1197-1203.

**英文引用格式:** WANG Qianqian, MIAO Duoqian, ZHANG Yuanjian. Abnormal event detection method based on deep auto-encoder and self-updating sparse combination[J]. CAAI transactions on intelligent systems, 2020, 15(6): 1197-1203.

## Abnormal event detection method based on deep auto-encoder and self-updating sparse combination

WANG Qianqian, MIAO Duoqian, ZHANG Yuanjian

(Key Laboratory of Embedded System and Service Computing, Tongji University, Shanghai 201804, China)

**Abstract:** In the construction of a deep learning model for abnormal event detection, frames or optical flow are considered but the resulting accuracy and speed are not satisfactory. To address these problems, we present an algorithm based on convolutional auto-encoders and self-updating sparse combination learning, which is centered on the movement of foreground blocks. First, we use an adaptive Gaussian mixture model to extract the foreground. Using a sliding window, the foreground blocks that are moving, are filtered based on the number of foreground pixels. Three convolutional auto-encoders are then constructed to extract the temporal and spatial features of the moving foreground blocks. Lastly, self-updating sparse combination learning is applied to reconstruct the features and identify abnormal events based on the reconstruction error. The experimental results show that compared with existing algorithms, the proposed method improves the accuracy of abnormality detection and enables real-time detection.

**Keywords:** deep learning; sparse combination; auto-encoder; self-updating; abnormal event detection; convolution neural network; unsupervised learning; sparse representation

异常事件检测是指通过图像处理、模式识别和计算机视觉等技术, 分析视频中的有效信息, 判断异常事件。作为智能视频监控系统的核心应用之一, 异常事件检测受到了国内外学者的广泛关注。

因光照、背景和视角等因素的影响以及缺少异常数据, 异常事件检测仍是一项具有挑战性的任务<sup>[1-2]</sup>。

异常事件检测通常包含特征提取和建立检测模型。特征分为底层和深度学习特征。底层特征主要有方向梯度直方图<sup>[3]</sup>、三维时空梯度<sup>[4]</sup>、光流直方图<sup>[5]</sup>等。近年来, 部分学者提出基于深度学习的检测算法<sup>[6-9]</sup>。Zhou 等<sup>[8]</sup>设计了一个特征提

收稿日期: 2020-07-01.

基金项目: 国家自然科学基金项目 (61976158, 61673301).

通信作者: 苗夺谦. E-mail: dqmiao@tongji.edu.cn.

取网络,包含运动混合和特征迁移两个模块。检测模型主要有分类和聚类<sup>[10-11]</sup>、误差值<sup>[12-13]</sup>和稀疏表示<sup>[14-17]</sup>。Lu等<sup>[16]</sup>取消稀疏约束,提出稀疏组合学习算法,检测速度可达150 f/s。文献<sup>[17]</sup>提出将稀疏参数映射到层叠循环神经网络中,提高了检测速度。

本文针对异常检测输入通常为视频帧或图像块的局限性,利用混合高斯算法提取运动前景像素,以滑动窗口方式根据前景像素点占比获取运动前景块,并设计3个自动编码器,分别以灰度图和梯度作为输入,学习前景块的时空特征。最后,采用改进的稀疏组合算法对特征建模,进行异常检测。

## 1 相关工作

### 1.1 背景建模

视频中的运动前景检测方法主要有:光流法、帧间差分法和背景差分法3种<sup>[18]</sup>。自适应混合高斯算法是常用的前景提取方法,可自适应地确定高斯数量,减少光照、噪声等带来的影响。本文采用Zivkovic<sup>[19]</sup>提出的自适应混合高斯算法,提取视频的运动前景。图1为运动前景提取结果图。可以看到,即使光照不均、树叶轻微浮动,运动前景区域分割效果依然很好。

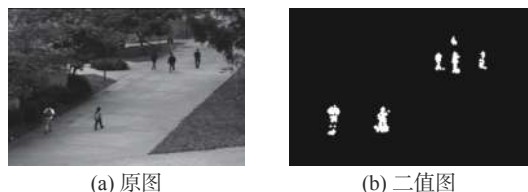


图1 自适应混合高斯背景建模算法提取运动前景  
Fig. 1 Extraction of moving foregrounds by adaptive GMM

### 1.2 异常事件检测中的特征学习

特征学习是异常检测的基础和关键。Lee等<sup>[10]</sup>考虑速度、方向和群体效应3种因素,采用Farneback算法<sup>[20]</sup>提取光流特征,并构建区域运动影响矩阵。

近年来,深度学习在计算机视觉领域取得了巨大成功。Xu等<sup>[6]</sup>首次将深度学习应用在异常检测任务,采用叠加去噪编码器提取特征。Hasan等<sup>[12]</sup>设计了两个编码器,分别为全连接自动编码器和三维全卷积前馈网络。此外,一些方法使用预训练网络提取表观特征。如文献<sup>[21]</sup>采用ImageNet<sup>[22]</sup>数据集预训练的卷积网络提取特征。该类方法受限于相关任务类型和模型跨领域问题,检测准确率偏低。

与上述方法不同的是,本文提出采用卷积自

动编码器学习运动前景块的时空特征,过滤冗余背景信息。

### 1.3 基于稀疏编码的异常检测

基于稀疏编码的异常检测<sup>[13-16]</sup>在训练阶段从正常特征数据 $X = \{x_1, x_2, \dots, x_n\}$ 中学习稀疏约束的字典 $D$ 。在测试阶段,给定测试样本 $x$ ,计算稀疏重构误差值:

$$\min_{\beta} \|x - D\beta\|_2^2, \text{ s.t. } \|\beta\|_0 \leq s$$

式中: $\beta$ 为稀疏编码; $s$ 用于控制稀疏度。稀疏编码的优化较为耗时,文献<sup>[14]</sup>减小字典大小来提升检测效率。Zhao等<sup>[15]</sup>改用更快速的求解器,但单帧检测仍需数秒时间。文献<sup>[16]</sup>提出取消稀疏约束学习多个小字典,作者称之为稀疏组合学习,可达到150 f/s的检测效率。本文在此基础上提出了一种自更新的稀疏组合学习方法,以权值的方式加入先验知识,并在线更新,提高稀疏组合的灵活性。

## 2 基于运动前景快的特征表示

### 2.1 构建运动前景块

采用Zivkovic<sup>[19]</sup>方法,获取视频运动前景区域。以滑动窗口方式遍历视频帧,前景像素个数占比超过阈值的窗口块为运动前景块。由于前景检测结果中存在噪声点,通过该方式可过滤掉部分噪声,还可保留前景像素点的空间相对位置关系。

具体方法为,假定窗口内的前景像素个数为 $b$ ,计算前景像素点的占比:

$$p = \frac{b}{w \times w}$$

式中: $w$ 为窗口大小,本文为64。若 $p$ 大于阈值则当前窗口为运动前景块,否则为背景或噪声块。

### 2.2 基于卷积自动编码器的特征提取

本文构建的3个卷积自动编码器提取时空特征,如图2所示。位于中间的自动编码器以灰度图作为输入,学习表观特征。另外两个的输入均为梯度,分别学习前后运动特征。3个自动编码器采用相同的结构,其中编码器由3个卷积层和3个最大池化层构成,解码器为3个上采样层和3个卷积层,以及输出层,具体参数如表1所示。卷积自动编码器的输入输出均为 $64 \times 64 \times 1$ 。卷积层采用 $3 \times 3$ 滤波器,除输出层外均使用ReLU激活函数。编码器的卷积层分别设有32、16、8个滤波器,池化层的步长为2。解码器的第一个卷积层为8个滤波器,随后两层设为16和32,上采样层将输入大小扩大一倍。

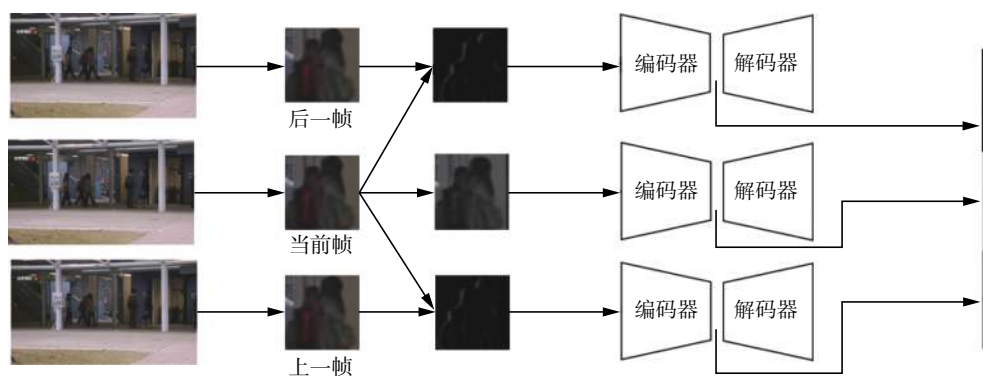


图 2 基于运动前景块的卷积自动编码器结构

Fig. 2 Structure of convolutional automatic encoder

表 1 卷积自动编码器的具体结构

Table 1 Structure specifications of the convolutional auto-encoder

| 层名称           | 描述   | 参数                              |
|---------------|------|---------------------------------|
| Input         | 输入层  | 64×64                           |
| Encoder-Conv1 | 卷积层  | 32个3×3的卷积核, stride=1, padding=1 |
| Encoder-Pool1 | 池化层  | kernel_size=2, stride=2         |
| Encoder-Conv2 | 卷积层  | 16个3×3的卷积核, stride=1, padding=1 |
| Encoder-Pool2 | 池化层  | kernel_size=2, stride=2         |
| Encoder-Conv3 | 卷积层  | 8个3×3的卷积核, stride=1, padding=1  |
| Encoder-Pool3 | 池化层  | kernel_size=2, stride=2         |
| Decoder-Pool1 | 上采样层 | kernel_size=2, stride=2         |
| Decoder-Conv1 | 卷积层  | 8个3×3的卷积核, stride=1, padding=1  |
| Decoder-Pool2 | 上采样层 | kernel_size=2, stride=2         |
| Decoder-Conv2 | 卷积层  | 16个3×3的卷积核, stride=1, padding=1 |
| Decoder-Pool2 | 上采样层 | kernel_size=2, stride=2         |
| Decoder-Conv3 | 卷积层  | 32个3×3的卷积核, stride=1, padding=1 |
| Conv          | 输出层  | 1个3×3的卷积核, stride=1, padding=1  |

采用 Adam 方法<sup>[21]</sup>对网络的参数进行优化, 损失函数为

$$L(I, O) = \frac{1}{h, w} \sum_{i=1}^h \sum_{j=1}^w (I_{ij} - O_{ij})^2$$

式中:  $I$  和  $O$  分别为输入输出运动前景块;  $h$ 、 $w$  为运动前景块的高宽。训练数据为正常视频, 自动编码器学习正常运动目标的时空特征。将自动编码器的隐特征拼接作为运动前景块的时空特征。因每个卷积编码器的隐特征为  $8 \times 8 \times 8$ , 故运动前景块的时空特征为 1536 维 ( $8 \times 8 \times 8 \times 3$ )。

### 3 基于自更新稀疏组合异常检测算法

自更新稀疏组合检测算法流程图如图 3 所示, 从训练数据中学习稀疏组合集合  $S = \{S_1, S_2, \dots, S_k\}$ , 设定初始权值。给定测试数据, 计算加权最小二乘误差, 选取最小值作为结果输出。根

据输出值大小判断是否为异常, 同时选择置信度较高的测试样本对权值进行更新。

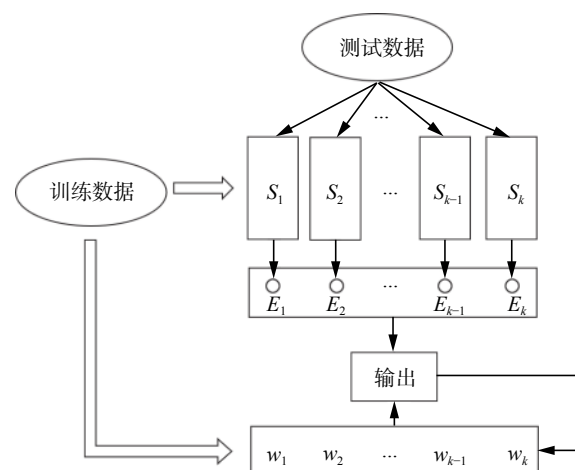


图 3 基于自更新稀疏组合学习检测模型

Fig. 3 Self-updating sparse combination learning detection model

### 3.1 模型构造

稀疏组合学习假定训练样本均可在稀疏组合集合中找出能够以较小误差进行线性重构的组合, 表示为

$$\begin{aligned} t = \min_{S, \gamma, \beta} \sum_{j=1}^n \sum_{i=1}^k \gamma_j^i \|x_j - S_i \beta_j^i\|_2^2 \\ \text{s.t.} \quad \sum_{i=1}^k \gamma_j^i = 1, \gamma_j^i \in \{0, 1\} \end{aligned} \quad (1)$$

式中:  $S_i$  为稀疏组合;  $x_j$  为运动前景块的特征;  $n$  为训练样本的总数;  $k$  为稀疏组合的个数;  $\gamma = \{\gamma_1, \gamma_2, \dots, \gamma_n\}$ ,  $\gamma_j = \{\gamma_j^1, \gamma_j^2, \dots, \gamma_j^k\}$ ,  $\gamma_j^i$  表示  $S_i$  是否为特征  $x_j$  的最优稀疏组合;  $\beta_j^i$  是稀疏组合  $S_i$  用于表示  $x_j$  的系数。

### 3.2 模型参数求解

为了提高检测效率,  $k$  值应尽可能小。引入超参  $\lambda$ , 当训练数据的重构误差小于该值时, 训练结束。式 (1) 更新为

$$\begin{aligned} t_j = \sum_{i=1}^k \gamma_j^i \{ \|x_j - S_i \beta_j^i\|_2^2 - \lambda \} \leq 0 \\ \text{s.t.} \quad \sum_{i=1}^k \gamma_j^i = 1, \gamma_j^i \in \{0, 1\} \quad \forall j \in \{1, 2, \dots, n\} \end{aligned} \quad (2)$$

采用迭代方式求解式 (2), 可快速找到表征大多数正常特征的稀疏组合。具体方法为, 对于第  $i$  轮, 更新  $S_i$ 。式 (2) 转化为

$$\begin{aligned} \min_{S_i, \gamma, \beta} \sum_{j \in \Omega_c} \gamma_j^i (\|x_j - S_i \beta_j^i\|_2^2 - \lambda) \\ \text{s.t.} \quad \sum_{i=1}^k \gamma_j^i = 1, \gamma_j^i \in \{0, 1\} \end{aligned} \quad (3)$$

式中  $\Omega_c$  为当前训练数据  $X_c \in X$  的下标集合。同样采用迭代方式求解式 (3), 分两步进行:

1) 固定  $\gamma$ , 更新  $S_i$  和  $\beta$ 。式 (3) 简化为二次方程:

$$L(\beta, S_i) = \sum_{j \in \Omega_c} \gamma_j^i \|x_j - S_i \beta_j^i\|_2^2 \quad (4)$$

对所有  $\gamma_j^i \neq 0$ , 可求得最优  $\beta$ 。 $\beta$  的闭式解为

$$\beta_j^i = (S_i^T S_i)^{-1} S_i^T x_j \quad (5)$$

通过块坐标下降法求得  $S_i$  的最优解为

$$S_i = \Pi[S_i - \delta_i \nabla_{S_i} L(\beta, S_i)] \quad (6)$$

式中  $\delta_i$  设置为 0.0001。

2) 根据  $\{S_i, \beta\}$ , 更新  $\gamma$ 。对  $x_j$  求解函数为

$$\min_{\gamma_j^i} \gamma_j^i \|x_j - S_i \beta_j^i\|_2^2 - \lambda \gamma_j^i, \quad \text{s.t.} \quad \gamma_j^i \in \{0, 1\} \quad (7)$$

易得  $\gamma_j^i$  的闭式解为

$$\gamma_j^i = \begin{cases} 1, & \|x_j - S_i \beta_j^i\|_2^2 < \lambda \\ 0, & \text{其他} \end{cases} \quad (8)$$

重复上述两个步骤, 直至式 (3) 收敛, 则第  $i$  轮结束。 $x_c$  为空时, 整个训练结束。

### 3.3 权值初始化和自更新

出现频率越高的稀疏组合所表示事件为正常的概率越高。为每个稀疏组合设定权值, 值越高重要性越大。在 3.2 节每轮训练结束后, 统计  $\gamma_j^i = 1$  的总数, 记为  $\tau_i$ 。对  $S_i$  的权值进行初始化:

$$w_i^0 = \frac{\tau_i^0}{n}$$

测试阶段, 选择重构误差值小于阈值  $\mu$  的样本对权值进行更新。给定测试特征  $x \in \mathbf{R}^p$ , 计算稀疏组合的加权重构误差:

$$e_i = (1 - w_i) \times E_i$$

式中  $1 \leq i \leq k$ 。若  $\min(e_1, e_2, \dots, e_k) \leq \mu$ , 则对权值进行更新, 公式为

$$\begin{aligned} \tau_i^{t+1} = \begin{cases} \tau_i^t + 1, & e_i = \min(e_1, e_2, \dots, e_k) \\ \tau_i^t, & \text{其他} \end{cases} \\ w_i^{t+1} = \frac{\tau_i^{t+1}}{\sum_{j=1}^k \tau_j^{t+1}} \end{aligned} \quad (9)$$

### 3.4 自更新稀疏组合检测

给定测试数据  $x$ , 计算稀疏组合的最小二乘误差:

$$\min_{\beta} \|x - S_i \beta^i\|_2^2$$

式中:  $1 \leq i \leq k$ ;  $\beta$  的闭式解为

$$\hat{\beta}^i = (S_i^T S_i)^{-1} S_i^T x$$

则  $x$  与  $S_i$  的加权重构误差为

$$(1 - w_i) \|x - S_i \hat{\beta}^i\|_2^2 = (1 - w_i) \left\| \left( S_i (S_i^T S_i)^{-1} S_i^T - I_p \right) x \right\|_2^2$$

式中  $I_p$  为  $p \times p$  的单位矩阵。为每个稀疏组合定义一个辅助矩阵  $H_i$ :

$$H_i = S_i (S_i^T S_i)^{-1} S_i^T - I_p$$

最终只需对测试数据  $x$ , 按照式 (10) 即可计算出其与每个稀疏组合的加权重构误差:

$$\begin{cases} E_i = \|H_i x\|_2^2 \\ e_i = (1 - w_i) E_i \end{cases} \quad (10)$$

若存在  $e_i < T$  ( $1 \leq i \leq k$ ),  $x$  为正常, 否则为异常。如果  $e_i < \mu$ , 则根据式 (9) 对权值进行更新, 完成权值的在线更新。

## 4 实验与分析

### 4.1 数据集

本文在数据集 Avenue<sup>[15]</sup> 和 UCSD<sup>[23]</sup> 上进行了大量实验。Avenue 数据集中共 16 个训练视频, 21 个测试视频。训练视频仅为正常事件, 测

试视频包含异常和正常事件。异常事件包含快速奔跑、抛掷物品和蹦跳等, 如图 4 所示。



图 4 Avenue 和 UCSD 数据集异常图像示例

Fig. 4 Abnormal images from the Avenue and UCSD datasets

UCSD 数据集分为 2 个场景 Ped1 和 Ped2。训练数据为正常事件, 测试数据涉及多种异常行为, 如行人道上骑车、出现货车、轮滑等, 如图 4 所示。Ped1 有 34 段训练视频和 36 段测试视频; Ped2 视频包含 16 段训练视频和 12 段测试视频。

#### 4.2 实验设置

对于卷积自动编码器, 每个数据集都进行 200 轮训练, 前 100 轮学习率设为 0.001, 后 100 轮学习率设为 0.000 1, 采用小批量方法训练。式 (2) 中的  $\lambda$  设置为 0.04, 表示误差上限。实验均在 Ubuntu16.04、NVIDIA CUDA 8.0、PyTorch1.0、OpenCVs 3.0、Python 环境下完成。

#### 4.3 评价指标

采用异常检测常用的帧级 AUC 评价指标<sup>[24]</sup>。对于检测结果为异常的样本, 若其标注中存在异常像素, 则认为检测正确。AUC 值为接收者操作特性曲线 (receiver operating characteristic, ROC) 下包含的面积。

#### 4.4 结果与分析

##### 4.4.1 Avenue 数据集

图 5 为 Avenue 数据集测试视频 03 的异常得分。图中灰色部分为标注的异常帧, 横轴为帧序号, 纵轴为异常得分。可以看到灰色区间内的得分较高而其他区域则偏小。说明本文方法可以检测出这两个异常事件。

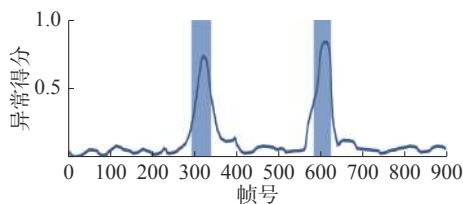


图 5 Avenue 数据集测试视频 03 重构误差值

Fig. 5 Reconstruction cost of Avenue dataset test video 03

表 2 列出了 Avenue 测试数据的检测结果, 由于篇幅所限只列出了部分结果, 所给结果中包含

最好与最差情况。最高检测率为视频 03, 包含两个快速奔跑异常。视频 07 的异常为行人蹦跳, 识别率为 90.2%, 是 21 个测试视频中的最低值。对比两段视频, 发现虽然两者都是速度上的异常, 但视频 03 中的运动状态变化远大于视频 07, 造成两者识别率差异较大。

表 2 本文方法在 Avenue 测试视频上的实验结果

Table 2 Results of proposed method on Avenue test videos

| 测试视频 | 总帧数  | 正确帧数 | 错误帧数 | 识别率/% |
|------|------|------|------|-------|
| 01   | 1424 | 1304 | 120  | 91.6  |
| 03   | 923  | 852  | 71   | 92.3  |
| 07   | 605  | 546  | 59   | 90.2  |

为了进一步验证本文算法的有效性, 与其他算法进行了比较, 如表 3 所示。可以看到, 相较于经典的稀疏组合算法<sup>[16]</sup>, 本文算法提升了 6.4%, 相较于最近的方法<sup>[25]</sup>, 也有 2.9% 的提升。

表 3 在 Avenue 数据集上的结果对比

Table 3 Comparison of performance on Avenue dataset with AUC

| 算法                      | AUC  |
|-------------------------|------|
| SC <sup>[15]</sup>      | 80.9 |
| Conv-AE <sup>[12]</sup> | 70.2 |
| C2ST <sup>[25]</sup>    | 84.4 |
| CASSC                   | 87.3 |

##### 4.4.2 UCSD 数据集

表 4 给出了本文算法与其他算法在 UCSD 数据集上的 AUC 值。本文算法 AUC 为 87.3%, 比 Conv-AE<sup>[12]</sup> 提高了 6.3%, 与 SC<sup>[15]</sup> 相比提升了 1.05%; 在数据集 Ped2 上, 本文算法比 SC<sup>[15]</sup> 提高了 11.01%。说明基于运动前景块的特征提取方法有效, 滤除了背景冗余信息。所构建的卷积自动编码器可学到更高语义层次的特征, 与手工提取特征相比, 鲁棒性更强。

表 4 在 UCSD 数据集上的 AUC 值对比

Table 4 Comparison of performance on UCSD dataset with AUC

| 算法                      | UCSD Ped1 | UCSD Ped2 | 平均    |
|-------------------------|-----------|-----------|-------|
| SF <sup>[26]</sup>      | 74.46     | 68.49     | 71.47 |
| MIM <sup>[10]</sup>     | 80.22     | 90.87     | 85.54 |
| SL-HOF <sup>[27]</sup>  | 79.94     | 89.52     | 84.73 |
| Conv-AE <sup>[12]</sup> | 81.00     | 90.00     | 85.50 |
| SC <sup>[15]</sup>      | 86.25     | 80.67     | 83.46 |
| CASSC                   | 87.30     | 91.38     | 89.34 |

图6给出了UCSD数据集上正确检测的异常帧和未检测出的异常帧。图中左侧为部分正确检出的异常帧,可以看到本文算法可将骑行、轮滑以及巡逻车检出;右侧为未检测出的异常帧,异常区域用红色圈标出,异常目标存在遮挡或与周围行人产生交互的情况,算法未能将其识别为异常。

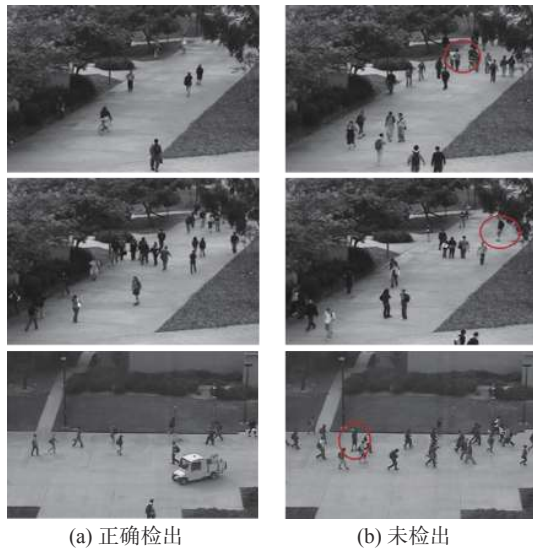


图6 UCSD数据集检测出的异常帧和未检出的异常帧  
Fig. 6 Comparisons of abnormal frames with detected and undetected abnormal frames on UCSD dataset

#### 4.4.3 运行时间

表5中所列为本文算法和经典稀疏组合学习算法<sup>[15]</sup>在数据集上每帧检测所需时间。本文算法虽慢于稀疏组合学习算法,但根据表3和表4的结果,本文算法检测效果更优。监控视频帧率通常为25 f/s,而本文算法平均能达到100 f/s,完全可以满足实时检测的要求。

表5 每帧处理时间  
Table 5 Processing times per frame s

| 算法                 | UCSD Ped1 | UCSD Ped2 | Avenue |
|--------------------|-----------|-----------|--------|
| SC <sup>[15]</sup> | 0.006     | 0.007     | 0.007  |
| CASSC              | 0.011     | 0.010     | 0.012  |

## 5 结束语

本文提出了一种以无监督方式学习运动前景块的深度特征方法和自更新稀疏组合异常检测算法。主要贡献包括4个方面:

1) 提出了以无监督方法基于自动编码器提取运动前景块的时空特征,过滤冗余背景信息,同时保留运动前景像素点的空间相对位置关系;

2) 针对经典的稀疏组合学习缺乏灵活性的问题,通过加入先验知识,提出了一种自更新稀疏组合学习算法;

3) 将深度学习与改进的稀疏组合学习进行结合,提出了一个新的异常检测框架;

4) 一系列监控场景数据的实验结果表明,本文提出的异常检测算法可以准确高效地检测视频中的异常事件。

## 参考文献:

- [1] POPOOLA O P, WANG Kejun. Video-based abnormal human behavior recognition—a review[J]. *IEEE transactions on systems, man, and cybernetics, part C (applications and reviews)*, 2012, 42(6): 865–878.
- [2] MABROUK A B, ZAGROUBA E. Abnormal behavior recognition for intelligent video surveillance systems: A review[J]. *Expert systems with applications*, 2018, 91: 480–491.
- [3] ROSHTKHARI M J, LEVINE M D. Online dominant and anomalous behavior detection in videos[C]//*Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition*. Portland, USA, 2013: 2611–2618.
- [4] ZAHARESCU A, WILDES R. Anomalous behaviour detection using spatiotemporal oriented energies, subset inclusion histogram comparison and event-driven processing[C]//*Proceedings of the 11th European Conference on Computer Vision*. Heraklion, Greece, 2010: 563–576.
- [5] WANG Tian, SNOUSSI H. Histograms of optical flow orientation for visual abnormal events detection[C]//*Proceedings of 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*. Beijing, China, 2012: 13–18.
- [6] XU Dan, YAN Yan, RICCI E, et al. Detecting anomalous events in videos by learning deep representations of appearance and motion[J]. *Computer vision and image understanding*, 2017, 156: 117–127.
- [7] 李俊杰, 刘成林, 朱明. 基于多任务 CNN 的监控视频中异常行人快速检测 [J]. *计算机系统应用*, 2018, 27(11): 78–83.  
LI Junjie, LIU Chenglin, ZHU Ming. Fast abnormal pedestrians detection based on multi-task CNN in surveillance video[J]. *Computer systems & applications*, 2018, 27(11): 78–83.
- [8] ZHOU J T, DU Jiawei, ZHU Hongyuan, et al. AnomalyNet: an anomaly detection network for video surveillance[J]. *IEEE transactions on information forensics and security*, 2019, 14(10): 2537–2550.
- [9] NGUYEN T N, MEUNIER J. Anomaly detection in video sequence with appearance-motion correspondence[C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South), 2019: 1273–1283.
- [10] LEE D G, SUK H I, LEE S W. Crowd behavior representation using motion influence matrix for anomaly detection

- tion[C]//Proceedings of 2013 2nd IAPR Asian Conference on Pattern Recognition. Naha, Japan, 2013: 110–114.
- [11] 胡正平, 张乐, 尹艳华. 时空深度特征 AP 聚类的稀疏表示视频异常检测算法[J]. 信号处理, 2019, 35(3): 386–395.
- HU Zhengping, ZHANG Le, YIN Yanhua. Video anomaly detection by AP clustering sparse representation based on spatial-temporal deep feature model[J]. Journal of signal processing, 2019, 35(3): 386–395.
- [12] HASAN M, CHOI J, NEUMANN J, et al. Learning temporal regularity in video sequences[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 733–742.
- [13] WEI Hao, LI Kai, LI Haichang, et al. Detecting video anomaly with a stacked convolutional LSTM framework[C]//Proceedings of 12th International Conference on Computer Vision Systems. Thessaloniki, Greece, 2019: 330–342.
- [14] CONG Yang, YUAN Junsong, LIU Ji. Sparse reconstruction cost for abnormal event detection[C]//Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA, 2011: 3449–3456.
- [15] ZHAO Bin, LI Feifei, XING E P. Online detection of unusual events in videos via dynamic sparse coding[C]//Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA, 2011: 3313–3320.
- [16] LU Cewu, SHI Jianping, JIA Jiaya. Abnormal event detection at 150 FPS in MATLAB[C]//Proceedings of the IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 2720–2727.
- [17] LUO Weixin, LIU Wen, GAO Shenghua. A revisit of sparse coding based anomaly detection in stacked RNN framework[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 341–349.
- [18] SAJID H, CHEUNG S C S. Universal multimode background subtraction[J]. *IEEE transactions on image processing*, 2017, 26(7): 3249–3260.
- [19] ZIVKOVIC Z. Improved adaptive Gaussian mixture model for background subtraction[C]//Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK, 2004: 28–31.
- [20] FARNEBÄCK G. Two-frame motion estimation based on polynomial expansion[C]//Proceedings of the 13th Scandinavian Conference on Image Analysis. Halmstad, Sweden, 2003: 363–370.
- [21] SMEUREANU S, IONESCU R T, POPESCU M, et al. Deep appearance features for abnormal behavior detection in video[C]//Proceedings of the 19th International Conference on Image Analysis and Processing. Catania, Italy, 2017: 779–789.
- [22] RUSSAKOVSKY O, DENG Jia, SU Hao, et al. ImageNet et large scale visual recognition challenge[J]. *International journal of computer vision*, 2015, 115(3): 211–252.
- [23] MAHADEVAN V, LI Weixin, BHALODIA V, et al. Anomaly detection in crowded scenes[C]//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA, 2010: 1975–1981.
- [24] BRADLEY A P. The use of the area under the ROC curve in the evaluation of machine learning algorithms[J]. *Pattern recognition*, 1997, 30(7): 1145–1159.
- [25] LIU Yusha, LI Chunliang, PÓCZOS B. Classifier two-sample test for video anomaly detections[C]//Proceedings of the 29th British Machine Vision Conference. Newcastle, UK, 2018: 71.
- [26] MEHRAN R, OYAMA A, SHAH M. Abnormal crowd behavior detection using social force model[C]//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA, 2009: 935–942.
- [27] WANG Siqi, ZHU En, YIN Jianping, et al. Anomaly detection in crowded scenes by SL-HOF descriptor and foreground classification[C]//Proceedings of 2016 23rd International Conference on Pattern Recognition. Cancun, Mexico, 2016: 3398–3403.

#### 作者简介:



王倩倩, 硕士研究生, 主要研究方向为视频中的异常事件检测与行人重识别。



苗夺谦, 教授, 博士生导师, 主要研究方向为人工智能、机器学习、大数据分析、粒度计算。主持完成国家自然科学基金项目 6 项, 在研国家重点研发计划项目 1 项、公安部重点计划项目 1 项。荣获 CAAI 吴文俊人工智能自然科学奖二等奖、国家教学成果二等奖, 授权专利 12 项。发表学术论文 100 余篇, 出版教材和学术著作 10 部。



张远健, 博士研究生, 主要研究方向为粒度计算、不确定性。