

# 基于自反馈最优子类挖掘的视频异常检测算法



侯春萍 赵春月 王致芃

天津大学电气与信息工程学院 天津 300072

(hcp@tju.edu.cn)

**摘要** 视频异常检测算法是视频处理领域的研究热点之一,用于检测视频中是否包含异常事件。然而,由于没有异常样本参与训练过程,且异常样本与正常样本之间存在一定程度的相似性,因此很难设计出一种有辨识力的异常检测模型。为了解决上述问题,文中首先提出了一种基于相似度保持和样本恢复的特征选择方法,该方法能够保留正常样本的相似关系,进而可以学习到能够准确描述正常事件的特征。其次,将异常事件检测任务转化为分类任务,并提出了一种自反馈最优子类挖掘方法来获得最优分类器。如果一个测试样本被所有分类器判断为异常,则该样本最终将被判定为异常。在公共视频数据集(Avenue数据集、UCSD Ped2数据集)上进行的大量实验的结果表明,所提异常事件检测算法可以取得很好的结果。

**关键词:** 视频异常事件检测;特征选择;自反馈;最优子类挖掘;一类支持向量机

中图法分类号 TN911.73

## Video Abnormal Event Detection Algorithm Based on Self-feedback Optimal Subclass Mining

HOU Chun-ping, ZHAO Chun-yue and WANG Zhi-peng

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

**Abstract** Video anomaly detection algorithm is one of the hot issues in the field of video processing, and it is used to detect whether an abnormal event is contained in the video. However, since abnormal samples are not involved in the training process, and there is a certain degree of similarity between abnormal samples and normal samples, it is difficult to design an abnormal detection model with discrimination. In order to solve the above problems, firstly, this paper proposes a feature selection method based on similarity preservation and sample recovery. This method can retain the similarity of normal samples, and then learn features that can accurately describe normal events. Secondly, it formalizes the abnormal event detection as classification problem, and proposes a self-feedback optimal subclass mining method to find optimal classifier. The sample will be labeled as anomaly if all classifiers label it as anomaly. Extensive experiments on public video surveillance datasets (i. e. Avenue Dataset and UCSD Ped2 Dataset) demonstrate that the proposed abnormal event detection method can achieve good results.

**Keywords** Video abnormal event detection, Feature selection, Self-feedback, Optimal subclass mining, One-class SVM

## 1 引言

近年来,视频分析与处理受到越来越多的关注,其中异常事件检测(一种自动检测视频中是否包含异常事件帧的方法)在视频分析与处理领域占据了越来越重要的地位,成为了科学研究的重要方向<sup>[1-5]</sup>。然而,在多数情况下,异常事件检测都面临着极大的挑战。首先,视频数据通常是高维的大数据,经常包含过多的冗余信息,这意味着视频往往是不易处理的。其次,传统方法通常基于经验来确定 K-means 算法的聚类簇数量,而聚类簇数量的变化往往会极大地影响异常检测的准确性。为了解决上述问题,本文提出了适合于异常检测的特

征选择算法并提出了自反馈机制以确定最优的聚类簇数量。

在视频异常检测领域,还没有专门用于异常检测的特征选择算法。大多数视频异常检测工作中的特征都是基于一般的降维方法获得的,如利用主成分分析(Principal Component Analysis, PCA)算法<sup>[6-7]</sup>。但是,此类方法无法保留原始数据的关键特性,只能减小原始特征的冗余度。而异常检测工作需要更加有效的特征,因此必须在原始视频数据上进行特征选择。本文提出了一种基于相似度保持和样本恢复的特征选择方法(Sample Prevention and Sample Recovery, SPSR),其不仅能够保留正常事件的全局一致性以及局部几何结构信息<sup>[8-9]</sup>,并且进一步保证了对原始数据的恢复。其中,局部的

到稿日期:2020-08-22 返修日期:2020-10-24 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国际合作与交流 NFSC 项目(61520106002);国家自然科学基金(61731003)

This work was supported by the projects of International Cooperation and Exchanges NFSC(61520106002) and National Natural Science Foundation of China(61731003).

通信作者:王致芃(zpwang@tju.edu.cn)

几何结构信息指每个正常事件与其邻近的样本所构成的特殊局部关系。全局一致性则指成对样本相似性,保留训练数据集中每两个样本间的相似性可以保存全局信息。

由于异常样本不易获得,大多数传统方法通常将异常事件检测任务转化为边界点检测任务。它们利用正常事件训练模型,然后将偏离该模型的样本视为异常<sup>[10-12]</sup>。例如,基于一类支持向量机构建异常事件检测模型,利用正常事件训练一类支持向量机,若向量机将样本标记为边界点,则将其判断为异常<sup>[11]</sup>。但是,由于异常事件的定义不明确,且在大多数情况下异常样本与正常样本有一定的相似性,这些模型的边界往往不够准确,这将影响异常判断的准确性。因此,本文基于  $K$ -means 算法与一类支持向量机算法将异常事件检测任务转化为分类任务。然而,  $K$ -means 聚类算法的  $k$  值,即聚类簇数量,对异常事件检测模型的效率有很大的影响。如果  $k$  (聚类数)很小,则在每个聚类簇上训练的一类支持向量机可能会包含太多异常样本;如果  $k$  太大,则已合适划分的聚类簇将被进一步划分,而过度分割将影响聚类的有效性。已有的工作通常根据经验确定  $k$  值<sup>[11]</sup>,但是这种行为缺乏普遍性。因此,本文提出了一种自反馈最优子类挖掘算法,即利用一类支持向量机模型的重叠变化来确定最优  $k$  值。其基本原理如下:基于  $K$ -means 聚类获得原始子簇,在每个原始聚类簇上训练 one-class SVM 模型,同时计算 one-class SVM 模型重叠区域的大小,并根据重叠区域大小确定  $k$  值,即聚类簇的最佳数量,从而获得最佳分类器。最后,通过多次分类任务来判断异常事件。

本文的主要贡献如下:

(1)在研究了特征选择对异常检测的重要性的基础上,提出了一种基于相似性保持和样本恢复的特征选择方法。这意味着本文算法不仅可以充分利用原始数据的全局成对相似性和局部结构信息,还可以更好地保留样本的关键信息,以保证样本的恢复。

(2)提出了一种自反馈最优子类挖掘方法,并将异常事件检测转化为最优分类问题。该方法基于  $K$ -means 聚类和 one-class SVM 之间的自反馈机制来确定最优的聚类数量,从而获得最优的分类器。

## 2 算法描述

本文提出了基于相似性保持和样本恢复的特征选择方法 (SPSR)以及基于自反馈最优子类挖掘确定聚类数量的机制,如图 1 中所示。基于 SPSR 从维度为  $m(m' < m)$  的原始数据中提取出维度为  $m'$  的低维特征,如图 1 中的绿色框所示;自反馈最优子类挖掘过程如图 1 中的紫色框所示。获得低维特征后,基于  $K$ -means 聚类获得表示正常运动的子簇,然后在每个聚类簇上基于 one-class SVM 进行分类器的训练。同时,计算出一类支持向量机的重叠区域的大小,用于确定自反馈机制中聚类的数量,最后在确定  $k$  值之后,本文基于 one-class SVM 训练用于检测异常事件的分类器。

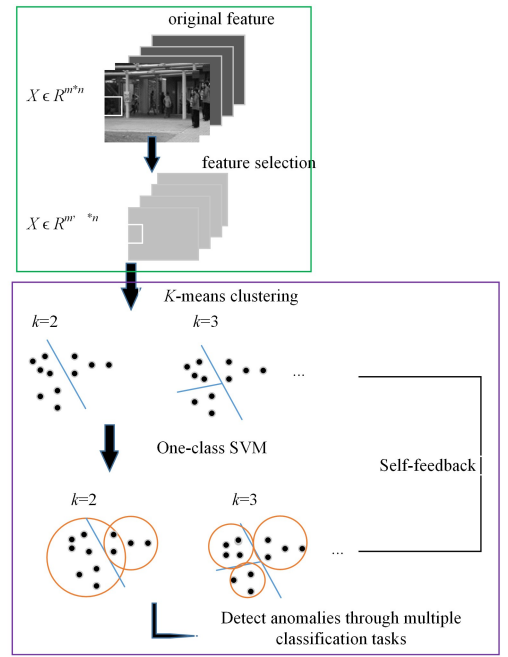


图 1 基于自反馈最优子类挖掘的异常检测框架 (电子版为彩色)

Fig. 1 Anomaly detection framework based on self-feedback optimal subclass mining

### 2.1 相似度保持和样本恢复特征选择算法

在 GLSPFS 方法中<sup>[13]</sup>,投影矩阵为  $\mathbf{W}$ ,将原始数据  $\mathbf{X}$  转换为低维数据  $\mathbf{XW}$ 。GLSPFS 使  $(\mathbf{XW})(\mathbf{XW})^T$  近似于成对相似性矩阵,以便使  $\mathbf{XW}$  可以表示全局信息。此外,使用  $\text{tr}(\mathbf{W}^T \mathbf{X}^T \mathbf{L} \mathbf{XW})$  ( $\text{tr}$  表示求取迹)捕获特征本身的几何属性,其中  $\mathbf{L}$  是表征数据局部几何结构的矩阵。GLSPFS 方法有效地保留了原始数据的全局信息和局部信息,但是忽略了对于原始数据的恢复,因此其提取的特征不适用于视频异常检测。

在充分利用原始数据(仅包含正常事件)的全局信息和局部信息的基础上,需要尽可能保证对原始数据的恢复,本文提出了基于相似性保留和样本恢复的特征选择方法的目标,如式(1)所示。

$$\begin{aligned} \min & \mu \text{tr}(\mathbf{W}^T \mathbf{X}^T \mathbf{L} \mathbf{XW}) + \lambda \|\mathbf{W}\|_F^2 + \beta \|\mathbf{E}\|_{2,1} \\ \text{s. t. } & \mathbf{X} = \mathbf{XW} \mathbf{P} + \mathbf{E}^1 \\ & \mathbf{V} = \mathbf{XW} + \mathbf{E}^2 \\ & \mathbf{E} = [\mathbf{E}^1; \mathbf{E}^2] \\ & \mathbf{P}^T \mathbf{P} = \mathbf{I} \end{aligned} \quad (1)$$

其中,  $\mathbf{W}$  为投影矩阵,利用  $\mathbf{W}$  将  $\mathbf{X}$  转换为低维特征  $\mathbf{XW}$ ;上标索引 T 表示矩阵的转置;  $\mathbf{V}$  是全局关系矩阵,  $\mathbf{V}$  可以通过将成对相似矩阵  $\mathbf{K}$  分解为  $\mathbf{K} = \mathbf{V} \mathbf{V}^T$  来获得。  $\mathbf{K}$  可以由高斯核关系式(2)计算得到。假设  $x$  和  $y$  是提取特征的随机样本,则它们之间的相似度为  $\mathbf{K}(x, y)$ 。

$$\mathbf{K}(x, y) = \begin{cases} \exp\left(\frac{\|x - y\|^2}{-2\sigma^2}\right), & x \neq y \\ 0, & \text{others} \end{cases} \quad (2)$$

$\mathbf{L}$  是表征数据局部几何结构的拉普拉斯矩阵,由式(3)计算得到。其中,  $\mathbf{D}$  是对角矩阵,其对角元素为  $\mathbf{V}$  的行总和。

$$\mathbf{L}=\mathbf{D}-\mathbf{V} \quad (3)$$

$\mathbf{V}=\mathbf{X}\mathbf{W}+\mathbf{E}^2$  这一项用来保持全局结构信息。这意味着此时的特征选择更多地集中在全局信息的保存上,例如识别物体的总体形状。 $\mathbf{V}$  保持了原始特征的全局成对相似性。本文希望降维后的数据  $\mathbf{X}\mathbf{W}$  仍保留原始数据的全局特性,使用  $\mathbf{E}^2$  来衡量它们之间的误差,迭代目标是使  $\mathbf{E}^2$  最小。

目标项  $\mu(\mathbf{W}^T\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{W})$  被用于提取原始数据的局部结构。这是因为不仅全局成对相似度对于特征选择很重要,局部结构信息也很重要,局部信息的保留进一步增强了特征的鲁棒性。为了提取局部几何结构信息,很多文献也利用了本文的目标项形式<sup>[14-17]</sup>,如在线无监督多视图特征选择(OM-VFS)<sup>[14]</sup>、线性保留投影(LPP)<sup>[17]</sup>等。 $\mu$  用于调整局部特征的重要性。

$\mathbf{X}=\mathbf{X}\mathbf{W}\mathbf{P}+\mathbf{E}^1$  这一项是为了确保低维数据  $\mathbf{X}\mathbf{W}$  能够尽可能地恢复原始特征  $\mathbf{X}$ ,以进一步增强特征的鲁棒性,降低特征冗余度。投影矩阵  $\mathbf{W}$  可以减少冗余并在恢复矩阵  $\mathbf{P}$  的帮助下确保对原始特征  $\mathbf{X}$  的良好重构。本文使用  $\mathbf{E}^1$  来测量原始数据  $\mathbf{X}$  和恢复数据  $\mathbf{X}\mathbf{W}\mathbf{P}$  之间的误差,迭代目标是使  $\mathbf{E}^1$  最小。此外,本文还添加了正交约束,如  $\mathbf{P}^T\mathbf{P}=\mathbf{I}$ 。

引入  $\|\mathbf{W}\|_F^2$  来提高计算效率。其中, $\lambda$  是用于调整计算效率重要性的参数; $\|\cdot\|_F$  是 Frobenius 范数; $\mathbf{E}=[\mathbf{E}^1;\mathbf{E}^2]$  这一项是误差矩阵,将  $\mathbf{E}^1$  和  $\mathbf{E}^2$  的列进行垂直连接可以得到  $\mathbf{E}$ ,此时最小化  $\mathbf{E}$  是迭代目标, $\beta$  用于调整在特征选择算法中将误差矩阵最小化的重要性的参数; $\|\cdot\|_{2,1}$  是  $L_{2,1}$  范数, $L_{2,1}$  范数强制列为零<sup>[18]</sup>,即  $\|\mathbf{W}\|_{2,1}=\sum_{j=1}^M\sqrt{\sum_{i=1}^N[\mathbf{W}_{ij}]^2}$ ,用于调整稀疏性以减少特征的冗余。

本文利用 ADMM 方法<sup>[18-19]</sup> 解决了目标函数(即式(1))的优化问题。增广拉格朗日函数由式(4)给出。

$$\begin{aligned} L(\mathbf{W}, \mathbf{E}, \mathbf{P}, \mathbf{Y}^1, \mathbf{Y}^2) &= \mu \text{tr}(\mathbf{W}^T\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{W}) + \lambda \|\mathbf{W}\|_F^2 + \\ &\beta \|\mathbf{E}\|_{2,1} + \phi(\mathbf{Y}^1, \mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{E}^1) + \\ &\phi(\mathbf{Y}^2, \mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2) \end{aligned}$$

$$\text{s. t. } \mathbf{E}=[\mathbf{E}^1;\mathbf{E}^2]$$

$$\mathbf{P}^T\mathbf{P}=\mathbf{I} \quad (4)$$

其中, $\phi(Y, Q)=(\alpha/2)\|y\|_F^2+\langle Y, Q\rangle$ , $\langle \cdot \rangle$  表示内积, $\alpha$  是正的惩罚标量, $\mathbf{Y}^1$  和  $\mathbf{Y}^2$  是拉格朗日乘数。接下来将给出每个变量的解决方案。

(1) $\mathbf{W}$  子问题。在固定所有其他变量的情况下,此算法通过以下方式更新  $\mathbf{W}$ :

$$\begin{aligned} \min \mu \text{tr}(\mathbf{W}^T\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{W}) + \lambda \|\mathbf{W}\|_F^2 + \phi(\mathbf{Y}^1, \mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}- \\ \mathbf{E}^1) + \phi(\mathbf{Y}^2, \mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2) \\ \Leftrightarrow \min \mu \text{tr}(\mathbf{W}^T\mathbf{X}^T\mathbf{L}\mathbf{X}\mathbf{W}) + \lambda \|\mathbf{W}\|_F^2 + (\alpha/2) \|\mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}- \\ \mathbf{E}^1+\mathbf{Y}^1/\alpha\|_F^2 + (\alpha/2) \|\mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2+\mathbf{Y}^2/\alpha\|_F^2 \end{aligned} \quad (5)$$

取  $\mathbf{W}$  的导数并将导数设置为零,可以获得:

$$\begin{aligned} \mathbf{W}=(2\alpha\mathbf{X}^T\mathbf{X}+\mu\mathbf{X}^T\mathbf{L}\mathbf{X}+\mu\mathbf{X}^T\mathbf{L}^T\mathbf{X}+2\lambda)^{-1}(\alpha\mathbf{X}^T\mathbf{V}- \\ \alpha\mathbf{X}^T\mathbf{E}^2+\mathbf{X}^T\mathbf{Y}^2+\alpha\mathbf{X}^T\mathbf{X}\mathbf{P}^T-\alpha\mathbf{X}^T\mathbf{E}^1\mathbf{P}^T+\mathbf{X}^T\mathbf{Y}^1\mathbf{P}^T) \end{aligned} \quad (6)$$

(2) $\mathbf{E}$  子问题。固定所有其他变量,此算法通过以下方式更新  $\mathbf{E}$ :

$$\min \beta \|\mathbf{E}\|_{2,1} + \phi(\mathbf{Y}^1, \mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{E}^1) + \phi(\mathbf{Y}^2, \mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2)$$

$$\begin{aligned} \Leftrightarrow \min \beta \|\mathbf{E}\|_{2,1} + (\alpha/2) \|\mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{E}^1+\mathbf{Y}^1/\alpha\|_F^2 + \\ (\alpha/2) \|\mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2+\mathbf{Y}^2/\alpha\|_F^2 \end{aligned} \quad (7)$$

重新组织以上公式,可以得到:

$$\begin{aligned} \min \beta \|\mathbf{E}\|_{2,1} + (\alpha/2) \|\mathbf{E}^1-\mathbf{X}+\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{Y}^1/\alpha\|_F^2 + (\alpha/ \\ 2) \|\mathbf{E}^2-\mathbf{V}+\mathbf{X}\mathbf{W}-\mathbf{Y}^2/\alpha\|_F^2 \\ \Leftrightarrow \min \alpha((\beta/\alpha) \|\mathbf{E}\|_{2,1} + (1/2) \|\mathbf{E}-\mathbf{G}\|_F^2) \end{aligned} \quad (8)$$

其中, $[\mathbf{G}=[\mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}+\mathbf{Y}^1/\alpha;\mathbf{V}-\mathbf{X}\mathbf{W}+\mathbf{Y}^2/\alpha]]$ ,表示  $\mathbf{G}$  通过将矩阵  $[\mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}+\mathbf{Y}^1/\alpha]$  和矩阵  $[\mathbf{V}-\mathbf{X}\mathbf{W}+\mathbf{Y}^2/\alpha]$  垂直联结得到。利用文献[20]中的方法可以获得  $\mathbf{E}$  的解,即:

$$[\mathbf{E}]_{:,i}=\begin{cases} \frac{\|G_{:,i}\|_2-\beta/\alpha}{\|G_{:,i}\|_2}G_{:,i}, & \|G_{:,i}\|_2>(\beta/\alpha) \\ 0, & \text{others} \end{cases} \quad (9)$$

(3) $\mathbf{P}$  子问题。关于  $\mathbf{P}$  的相关优化问题如下:

$$\min \phi(\mathbf{Y}^1, \mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{E}^1) \Leftrightarrow \min (\alpha/2) \|\mathbf{X}-\mathbf{E}^1+ \\ \mathbf{Y}^1/\alpha-\mathbf{X}\mathbf{W}\mathbf{P}\|_F^2 \quad (10)$$

$\mathbf{P}$  的解是  $\mathbf{P}=\mathbf{U}\mathbf{V}^T$ ,其中  $\mathbf{V}$  和  $\mathbf{U}$  是  $((\mathbf{X}\mathbf{W})^T(\mathbf{X}-\mathbf{E}^1+\mathbf{Y}^1/\alpha))$  的 SVD 左右奇异值<sup>[21]</sup>。

(4)更新乘数。乘数可以通过式(11)来更新,所有变量和乘数将迭代更新,直至收敛。

$$\begin{aligned} \mathbf{Y}^1 &= \mathbf{Y}^1 + \alpha(\mathbf{X}-\mathbf{X}\mathbf{W}\mathbf{P}-\mathbf{E}^1) \\ \mathbf{Y}^2 &= \mathbf{Y}^2 + \alpha(\mathbf{V}-\mathbf{X}\mathbf{W}-\mathbf{E}^2) \end{aligned} \quad (11)$$

## 2.2 自反馈最优子类挖掘方法

本文提出了一种介于  $K$ -means 聚类和 one-class SVM 间的自反馈机制,用于获得最优分类器。在特征选择过程中保持特征相似性的基础上,聚类方法可以更好地挖掘不同类型的正常运动。从仅包含正常事件的训练视频中选择特征后,本文通过利用  $K$ -means 聚类获得代表不同类型正常运动的子簇。但是,使用  $K$ -means 聚类的前提是,必须预先确定聚类数量。因此,本文提出了一种自反馈机制来确定  $k$  的最佳值, $k$  值最佳时分类器效果达到最优。

本文在每个聚类簇上训练一类支持向量机模型,以将异常检测任务转化为分类任务。但是,在每个子簇上训练的一类支持向量机模型会有重叠,并且当  $k$  改变时,重叠大小也会改变,如图 2 所示。如果重叠太大,意味着每种正常运动实际上并未被准确分割出。在这种情况下,与每个正常运动相对应的分类器不能准确地将不同于当前运动的其他运动检测为异常事件。因此,本文希望基于重叠区域得到最佳  $k$  值,进而得到分类效果最准确的分类器。本文  $k$  值的取值范围为 3~12,在该取值范围内计算训练的一类支持向量机模型的重叠区域,通过重叠区域大小的变化情况来可以确定  $k$  的最佳值。我们认为,当重叠区域最小时,聚类效果最佳。

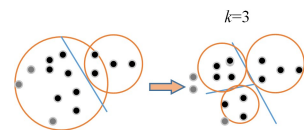
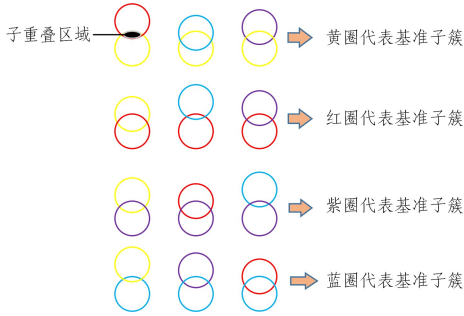


图 2  $k$  值变化时一类支持向量机模型的重叠效果图  
Fig. 2 Overlapping effect diagram of one-class SVM model as  $k$  changes

重叠大小的计算方法如下。假设  $k$  的值为 4, $K$ -means 聚类后将生成 4 个子簇。将第  $j$  个( $j \in [1, 4]$ )子簇视为基准

子簇,将其余 3 个子集群作为检查子簇。然后,在基准子簇上训练 one-class SVM 模型,计算每个检查簇与上述 one-class SVM 模型之间的子重叠。由于我们希望每个 one-class SVM 模型是没有重叠的,此时如果检查子簇的样本被基准子簇的一类支持向量机判断为正常,则说明基准子簇的一类支持向量机分类器存在重叠,即分类不准确。定义  $B_e(e \in [1, 3])$  为检查子簇中处于子重叠部分的样本(图 3 中的重叠部分为黑色区域)占检查子簇全部样本的百分比。假设每个检查子簇中的样本数为  $n_e$ ,子重叠区域中的样本数为  $n_{oe}$ 。第  $j$  个子簇作为基准子簇,通过对 3 个子重叠进行平均来获得此时的重叠大小,将其定义为  $F_j$ ,计算式如式(12)所示。如图 3 所示,有 4 个基准子簇,因此可以获得 4 组这样的重叠。最后,上述 4 组重叠的平均值为此  $k$  值对应的“一类支持模型的重叠”,将其定义为  $F$ ,用于确定  $k$  的值。

$$B_e = \frac{n_{oe}}{n_e}, F_j = \frac{\sum_{e=1}^3 B_e}{3}, F = \frac{\sum_{j=1}^4 F_j}{4} \quad (12)$$



注:每组中下面的圆是基准子簇,上面的圆是检查子簇

图 3 一类支持向量机的重叠(电子版为彩色)

Fig. 3 Overlapping of one-class SVM model

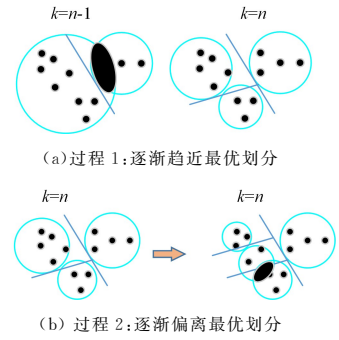
随着  $k$  值的增加,重叠区域会随之变化。本文提出  $k$  值变化理论: $k$  增大,重叠区域先减小后增大,最后不规则地变化。首先,当  $k$  值增加时,重叠区域减小,如图 4(a)所示。在这种情况下,随着子簇数量的增加,聚类结果将变得越来越准确。因此,one-class SVM 模型的重叠区域将变小。此后,当  $k$  值继续变大时,重叠区域也将变大,如图 4(b)所示。因为增加子簇的数量意味着将之前已合理划分的子簇进行过度分割,此时重叠区域的大小增加。假设当  $k$  等于 5 时,子簇的数量是合适的。随着子簇的数量从 5 增加到 6,第  $i$  个(其中  $i \in [1, 5]$ )子簇将被分割。在这种情况下,one-class SVM 的重叠区域将变大。最后,随着子簇数量的不断增加,重叠区域的大小将不规则地波动,因为当聚类数量太大时,聚类过程将变得毫无意义。因此,one-class SVM 模型的重叠区域变大之前,聚类效果是最佳的。本文提出的自反馈机制是子簇的数量由 one-class SVM 模型的重叠来判断。上述理论在第 3 节中得到了验证。

确定  $k$  的值后,利用多次的单分类任务进行视频中的异常事件工作。对于每个子簇,按照式(12)计算分类器  $g_m$  ( $m \in [1, k]$ )。

$$g_m(x) = \sum_{i=1}^m p_i \cdot x_i + b \quad (13)$$

其中, $x$  是测试样本( $u$  是样本数量), $b$  是偏差项, $p$  是权重向量。高斯滤波器对最终结果进行平滑处理。如果所有分类器

都将测试样本视为异常,那么我们认为测试样本是异常事件。



注:红色区域是重叠区域

图 4 当  $k$  值增加时,在每个子簇上训练的一类支持向量机模型的重叠情况(电子版为彩色)

Fig. 4 Overlapping of one-class SVM model trained on each sub-cluster as  $k$  increases

### 3 实验分析

本节在一些数据集上进行算法的有效性验证,并分析了本文算法与现有算法的对比结果。

#### 3.1 数据集

Avenue 数据集<sup>[22]</sup>包含 16 个训练视频和 21 个测试视频。UCSD 行人数据集<sup>[22]</sup>由 Ped1 和 Ped2 组成。许多现有的方法(如文献[22]中的方法)在实验中没有用到 Ped1,因为 Ped2 更具挑战性。因此,本文也仅在 Ped2 上进行实验。

#### 3.2 评估标准

本文以帧级水平的 ROC 曲线与 AUC 值为评价标准。ROC 曲线可以用真正率(True Positive Rate, TPR)和假正率(False Positive Rate, FPR)绘制。AUC 值可以通过计算 ROC 曲线与坐标轴形成的曲线面积得到。若视频帧包含至少一个异常像素,则该帧被视为异常。TPR 和 FPR 的计算方法如下:

$$TPR = \frac{\text{true positive frames' number}}{\text{positive frames' number}}$$

$$FPR = \frac{\text{false positive frames' number}}{\text{negative frames' number}}$$

#### 3.3 参数和实现细节

本文将训练视频和测试视频分为  $10 \times 10 \times 5$  的 cubes,即将训练视频和测试视频中的每个帧划分为大小为  $10 \times 10$  的块。然后,将 5 个连续的帧中相同位置的块连接起来,再提取特征。K-means 聚类算法是基于 LFeat<sup>[23]</sup> 工具包的 Lloyd<sup>[24]</sup> 方法,利用 K-means++<sup>[25]</sup> 进行初始化。一类支持向量机的正则化参数设置为 0.04,这意味着学习的模型会将 96% 的数据点视为正常,将另外 4% 的数据点视为异常<sup>[11]</sup>。在特征选择阶段,本文将  $\beta, \lambda$  和  $\mu$  分别设置为 2, 1, 0.5。

#### 3.4 Avenue 数据集上的结果

从 Avenue 数据集中提取 500 维的 3D 梯度特征。当  $k$  值在 3~12 之间变化时,使用 K-means 聚类得到代表不同类型正常运动的子聚类。然后,在每个子簇上训练一类支持向量机模型。最后,本文计算了当  $k$  从 3 变化到 12 时所学习的一类支持向量机模型的重叠与视频异常事件检测结果,如表 1 所列。

表1  $k$  值变化时一类支持向量机模型的重叠与视频异常事件检测结果Table 1 Overlapping of one-class SVM model and results of video abnormal events detection as  $k$  changes

$k$	overlapping/%	AUC value
3	39.652	0.782
4	33.155	0.782
5	31.740	0.781
6	27.849	0.782
7	25.204	0.782
8	27.645	0.782
9	26.149	0.780
10	26.470	0.780
11	24.785	0.780
12	23.389	0.780

可以看到,当 $k$ 的值介于3~7之间时,one-class SVM模型的重叠区域变小。当 $k$ 大于等于8时,重叠区域变大,这意味着某些划分良好的簇已被过度分割,使得重叠区域变大。当 $k$ 的值介于9~12之间时,重叠区域的大小不规则地波动。在这种情况下,子簇的数量过多,聚类效果下降。因此,当 $k$ 的值为7时,聚类效果最佳。但是,由于Avenue数据集的特征具有一定的稀疏性,当 $k$ 值发生变化时,Avenue数据集上的异常检测准确性变化不是很大。但是,当 $k$ 等于7时,异常检测的准确性仍然是最高的。当 $k$ 等于7时,AUC值为0.782。为了进一步验证 $k$ 值的变化理论,在UCSD数据集进行了同样的实验,并获得了较好的结果。

特征维度是非常重要的参数,它会影响异常检测的准确性,因此需要为本文的SPSR特征选择方法确定合适的特征维度。表2列出了AUC值随着特征维度的变化。显然,当特征维度等于150时,结果是最好的。在这种情况下,本文将设计对比实验来验证特征选择的有效性。

表2 特征维度变化时的AUC值

Table 2 AUC value changes with feature dimension

feature dimension	AUC value
100	0.819
120	0.816
150	0.821
180	0.816
200	0.812

为验证特征选择的有效性,也应用PCA获得维数为150的特征,目的是设计对比实验以证明本文的SPSR方法优于降维方法(即PCA方法)。图5给出了SPSR特征选择方法和PCA方法对应的ROC曲线,AUC值分别为0.821和0.807。从图5可以看出,与PCA方法相比,SPSR特征选择方法可以获得更好的结果。

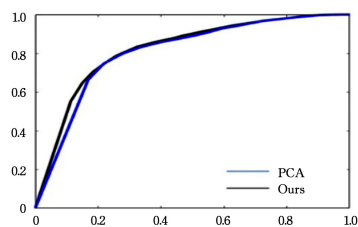


图5 利用SPSR特征选择方法和PCA方法降维的ROC曲线

Fig. 5 ROC curve using SPSR feature selection method and PCA method for dimensionality reduction

为了进一步验证本文提出的异常事件检测算法的性能,将本文的结果与以下的最新技术在帧级AUC水平上进行了比较,包括Lu等<sup>[26]</sup>、Hasan等<sup>[3]</sup>、Del Giorno等<sup>[27]</sup>、Ionescu等<sup>[4]</sup>、Luo等<sup>[28]</sup>和Chong等<sup>[29]</sup>提出的算法。其中,文献[3,30]中的方法是基于自编码器的深度学习算法,文献[28]中的方法是基于堆叠循环神经网络的深度学习算法,比较结果如表3所列。显然,本文算法的准确性高于一些深度学习算法与传统方法,这表明了其有效性。

表3 与某些最新算法在帧级水平的AUC进行比较的结果

Table 3 Comparison of AUC with some advanced algorithms at frame level

Method	AUC value
Lu et al. <sup>[26]</sup>	0.809
Hasan et al. <sup>[3]</sup>	0.702
Del Giorno et al. <sup>[27]</sup>	0.783
Ionescu et al. <sup>[4]</sup>	0.806
Luo et al. <sup>[28]</sup>	0.817
Y. S. Chong et al. <sup>[30]</sup>	0.803
OURS	0.821

### 3.5 UCSD Ped2数据集上的结果

Avenue数据集场景相对来说不那么密集,因此3D梯度特征能很好地表示出对象的运动信息。然而,UCSD Ped2数据集的移动对象更小且数量更多,3D梯度特征无法很好地表示数据集。因此,本文方法针对UCSD Ped2数据集利用定向光流直方图(Histograms of Oriented Optical Flow, HOF)特征和定向梯度直方图(Histogram of Oriented Gradient, HOG)特征作为基本描述子。

对于HOG特征,将每个cube的3D梯度特征量化为具有30个bin的基于方向的直方图,直方图块均匀分布在 $0^\circ\sim 360^\circ$ 之间。因此,对于每个cube,可以获得维度为30的特征向量。同样,对于HOF特征,将每个cube的光流特征量化为具有30个bin的基于方向的直方图,直方图块均匀分布在 $0^\circ\sim 360^\circ$ 之间。对于每个立方体,可以获得维度为30的特征向量。最后,将HOF特征和HOG特征垂直合并,以便从每个高维数据集中获得尺寸为60的特征向量。由于此时特征尺寸较小,因此我们不做进一步的特征选择,而是直接通过合并后HOG和HOF的特征验证异常检测算法的有效性。

本文计算在UCSD Ped2数据集上学习的one-class SVM模型的重叠大小,以确定 $k$ 的值,结果如表4中的第二列所列。可以看到,当 $k$ 的值在3~7之间变化时,重叠区域减小。然后,当 $k$ 等于8时,下降趋势停止。因此,当 $k$ 为7时,聚类的效果最好。本文在 $k$ 值介于3~12时计算AUC值,如表4中的第三列所列。可以发现,当 $k$ 等于7时,我们的异常事件检测方法可以达到最佳结果。这意味着,当 $k$ 的值为7时,聚类结果是最准确的,即证明了本文的 $k$ 值变化理论。

为了进一步验证异常事件检测方法的有效性,本文将结果与其他最新方法在帧级AUC水平上进行了比较,如表5所列和图6所示,包括Unmasking<sup>[4]</sup>、Conv-AE<sup>[3]</sup>、MPPCA<sup>[31]</sup>、MPPC + SPT<sup>[30]</sup>、MDT<sup>[32]</sup>、Force Flow<sup>[29]</sup>。其中,Conv-AE是一种深度学习算法。可以看出,本文的异常事件检测方法是有效的,甚至可以与某些深度学习算法相提并论。

然而,目前深度学习还不能与我们的异常检测方法很好地结合,而且深度特征也不能提高此算法的异常检测的准确性。在未来的工作中,我们将尝试将深度学习应用于异常事件检测方法,并提取更有效的功能。

表4 一类支持向量机模型的重叠大小与 AUC 值变化情况

Table 4 Overlapping size of one-class SVM model changes with

AUC value		
$k$	overlapping/ %	AUC value
3	10.546	0.828
4	5.654	0.830
5	3.356	0.814
6	2.076	0.827
7	1.904	0.851
8	2.103	0.830
9	2.263	0.819
10	2.125	0.826
11	1.930	0.819
12	2.082	0.817

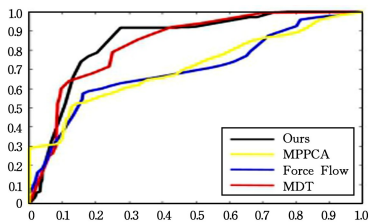


图6 UCSD Ped2 数据集中不同算法的 ROC 曲线

Fig. 6 ROC curves of different algorithms in UCSD Ped2 dataset

表5 本文算法与某些最新算法在帧级水平的 AUC 进行比较的结果

Table 5 Comparison of AUC between the proposed method and some advanced algorithms at frame level

Method	Frame AUC
MPPCA <sup>[31]</sup>	0.693
MPPC+SPF <sup>[32]</sup>	0.613
MDT <sup>[32]</sup>	0.829
Unmasking <sup>[4]</sup>	0.822
Conv-AE <sup>[3]</sup>	0.850
Force Flow <sup>[29]</sup>	0.556
Ours	0.851

**结束语** 本文提出了一种相似性保留和样本恢复特征的选择方法,用于保留全局成对样本相似性和数据的局部几何结构,还可以确保原始数据的恢复。此外,本文提出了一种自反馈最优子类挖掘方法,并将异常事件检测转化为分类问题。本文的自反馈机制将  $K$ -means 聚类和 one-class SVM 方法视为一个整体,以获得最优的分类器。对比许多传统方法,本文的异常事件检测方法取得了更好的结果,在 Avenue 数据集上的帧级 AUC 值为 0.8211,而在 UCSD Ped2 数据集上的帧级 AUC 值为 0.851。目前加入深度特征并不能提高异常检测的准确性,未来应考虑将本文算法与深度学习更好地结合。

## 参考文献

[1] SHI Y, YI Y, ZHANG Q, et al. Kernel null-space-based abnormal event detection using hybrid motion information[J]. Journal of Electronic Imaging, 2019, 28(2): 1-12.

[2] YU Y, SHEN W, HUANG H, et al. Abnormal event detection

in crowded scenes using two sparse dictionaries with saliency [J]. Journal of Electronic Imaging, 2017, 26(3): 033013.

[3] 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. 2016: 733-742.

[4] TUDOR I R, 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.

[5] RAVANBAKHSH M, NABI M, SANGINETO E, et al. Abnormal event detection in videos using generative adversarial nets [C] // 2017 IEEE International Conference on Image Processing (ICIP). 2017: 1577-1581.

[6] WANG S, ZHU E, YIN J, et al. Anomaly detection in crowded scenes by sl-hof descriptor and foreground classification[C] // 23rd International Conference on Pattern Recognition (ICPR). 2016.

[7] WANG S, ZHU E, YIN J, et al. Video anomaly detection and localization by local motion based joint video representation and oclm[J]. Neurocomputing, 2018, 277(FEB): 161-175.

[8] LEI Q, YE Y, L S, et al. Abnormal event detection based on multi-scale markov random field[C] // Chinese Conference on Computer Vision. 2015.

[9] CUI L, LI K, CHEN J, et al. Abnormal event detection in traffic video surveillance based on local features[C] // 2011 4th International Congress on Image and Signal Processing. 2011: 362-366.

[10] ZELNIK-MANOR L, ROSENBLUM K, ELDAR Y C. Dictionary optimization for block-sparse representations [J]. IEEE Transactions on Signal Processing, 2012, 60(5): 2386-2395.

[11] IONESCU R T, SMEUREANU S, POPESCU M, et al. Detecting abnormal events in video using narrowed normality clusters [C] // in 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). 2019: 1951-1960.

[12] AMRAE S, VAFAEI A, JAMSHIDI K, et al. Abnormal event detection in crowded scenes using one-class svm[J]. Signal, Image and Video Processing, 2018, 12(6): 1115-1123.

[13] LIU X, WANG L, ZHANG J, et al. Global and local structure preservation for feature selection [J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 25(6): 1083-1095.

[14] SHAO W, HE L, LU C T, et al. Online unsupervised multi-view feature selection[C] // 2016 IEEE 16th International Conference on Data Mining (ICDM). 2016: 1203-1208.

[15] ZHANG T, JIE Y, ZHAO D, et al. Linear local tangent space alignment and application to face recognition[J]. Neurocomputing, 2007, 70(7): 1547-1553.

[16] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.

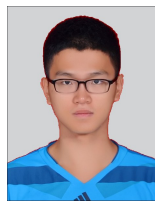
[17] HE X, NIYOGI P. Locality preserving projections [C] // Advances in Neural Information Processing Systems. 2004: 153-160.

[18] ZHOU T, ZHANG C, GONG C, et al. Multiview latent space

- learning with feature redundancy minimization[J]. IEEE Transactions on Cybernetics, 2018, 99: 1-14.
- [19] LIN Z C, LIU R, SU Z. Linearized alternating direction method with adaptive penalty for low-rank representation[J]. arXiv: 1109.0367, 2011.
- [20] LIU G, LIN Z, YAN S, et al. Robust recovery of subspace structures by low-rank representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 35(1): 171-184.
- [21] TANG C, ZHU X, LIU X, et al. Cross-view local structure preserved diversity and consensus learning for multi-view unsupervised feature selection[C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2019: 5101-5108.
- [22] IONESCU R T, KHAN F S, GEORGESCU M I, et al. Object-centric auto-encoders and dummy anomalies for abnormal event detection in video[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 7842-7851.
- [23] VEDALDI A, FULKERSON B. Vifeat: An open and portable library of computer vision algorithms[C] // Proceedings of the 18th International Conference on Multimedia 2010. Firenze, Italy, 2010.
- [24] DU Q, FABER V, GUNZBURGER M. Centroidal voronoi tessellations: Applications and algorithms[J]. SIAM review, 1999, 41(4): 637-676.
- [25] ARTHUR D, VASSILVITSKII S. K-means++: The advantages of careful seeding[C] // Eighteenth Acm-siam Symposium on Discrete Algorithms, 2007: 1027-1035.
- [26] LU C, SHI J, JIA J. Abnormal event detection at 150 fps in matlab[C] // Proceedings of the IEEE International Conference on Computer Vision, 2013: 2720-2727.
- [27] DEL GIORNO A, BAGNELL J A, HEBERT M. A discriminative framework for anomaly detection in large videos[C] // European Conference on Computer Vision, Springer, 2016: 334-349.
- [28] 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.
- [29] MEHRAN R, OYAMA A, SHAH M. Abnormal crowd behavior detection using social force model[C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009: 935-942.
- [30] CHONG Y S, TAY Y H. Abnormal event detection in videos using spatiotemporal autoencoder[C] // International Symposium on Neural Networks. Springer, 2017: 189-196.
- [31] KIM J K, GRAUMAN K. Observe locally, infer globally: A space-time mrf for detecting abnormal activities with incremental updates[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2009: 921-2928.
- [32] MAHADEVAN V, LI W, BHALODIA V, et al. Anomaly detection in crowded scenes[C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2010: 1975-1981.



**HOU Chun-ping**, born in 1957, Ph. D. professor, Ph. D supervisor. Her main research interests include wireless communications and 3D imaging.



**WANG Zhi-peng**, born in 1992, Ph. D. His main research interests include video and image anomaly detection.