

基于深度学习的异常事件检测

闻 佳^{1,2,3}, 王宏君^{1,2,3}, 邓 佳^{1,2,3}, 刘鹏飞^{1,2,3}

(1. 燕山大学信息科学与工程学院, 河北秦皇岛 066004;

2. 河北省计算机虚拟技术与系统集成重点实验室, 河北秦皇岛 066004;

3. 河北省软件工程国家重点实验室, 河北秦皇岛 066004)

摘 要: 面对复杂场景下异常事件检测的准确率偏低的情况, 本文提出一种基于深度学习的异常事件检测方法, 并将此方法扩展为异常事件分类方法. 利用神经网络模型提取特征, 将群体发散聚集事件, 群体密集聚集事件, 群体逃散事件和追赶事件这4种异常事件进行检测和分类. 通过 PKU-SVD-B 测试集对训练出来的模型进行测试实验, 并在 UMN 数据集上与几种方法做了对比实验, 验证了本文提出的基于深度学习的异常事件检测算法, 在适应多种不同场景的前提下, 对多种异常事件检测的准确率很高, 表明训练出来的模型对异常事件检测具有极强的泛化能力.

关键词: 异常检测; 异常分类; 深度学习; 图像处理; 卷积神经网络; 特征提取

中图分类号: TP389.1 **文献标识码:** A **文章编号:** 0372-2112 (2020)02-0308-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2020.02.013

Abnormal Event Detection Based on Deep Learning

WEN Jia^{1,2,3}, WANG Hong-jun^{1,2,3}, DENG Jia^{1,2,3}, LIU Peng-fei^{1,2,3}

(1. School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004, China;

2. The Key Laboratory for Computer Virtual Technology and System Integration of Hebei Province, Qinhuangdao, Hebei 066004, China;

3. State Key Laboratory of Software Engineering of Hebei Province, Qinhuangdao, Hebei 066004, China)

Abstract: Faced with low accuracy of abnormal event detection in complex scenarios, this paper proposes an abnormal event detection based on deep learning in various public scenes and multiple types of anomalies, and the method has been extended to an abnormal event classification method. The neural network model is used to extract features, and the four kinds of abnormal events, such as group divergence aggregation events, group intensive aggregation events, group escape events and catch-up events, are detected and classified. Test the trained model with PKU-SVD-B test set, compared with various methods on the UMN datasets, and verify the algorithm of abnormal event detection based on deep learning proposed in this paper. Under the premise of adapting to different scenarios, various abnormal events are detected. The high accuracy rate indicates that the trained model has strong ability to generalize abnormal event detection.

Key words: abnormal detection; abnormal classification; deep learning; image processing; convolutional neural network; feature extraction

1 引言

异常事件检测一直是计算机视觉领域的研究热点之一, 并且面临巨大的挑战^[1]. 许多学者对此做了研究^[2,3]. 异常事件检测算法主要有两个分支: 基于传统方法的异常事件检测算法^[4~6]和基于神经网络的异常事件检测算法^[7].

(1) 基于传统方法的异常事件检测算法

Si Wu 等人提出了一种基于贝叶斯框架的逃逸检

测方法^[8], 该算法通过构造光流的类条件概率密度函数, 结合贝叶斯框架, 检测异常人群逃逸事件. 孙慧等人提出一种隐马尔可夫(HMM)监控视频下的异常事件检测分析算法^[9], 该算法采用背景减除法进行前景提取, 用HMM建立正常事件模型, 用正常事件模型去检测样本集. Kumar D 通过目标跟踪和检测的方法进行异常事件检测^[10], 该方法通过提取前景目标的移动轨迹, 来获得场景中的移动目标团, 用获得的移动目标团来判定

前景是否有异常行为。

(2) 基于神经网络的异常事件检测算法

张伟峰中提出一种基于卷积神经网络(CNN)的面部遮挡的异常事件检测算法^[11]。刘顺华提出一种基于卷积神经网络的异常事件检测方法^[12],该算法结合滑动窗口和 PCA 算法,提取时空特征,用卷积神经网络进行异常检测的训练。

根据上述研究现状的分析,基于传统方法的异常事件检测算法检测模型简单,运算量小,数据集少等原因,很难应用在复杂多变的实际场合中,而目前基于神经网络的异常事件检测方法研究技术还不是很成熟,还需要继续研究与探索。本文正是利用深度神经网络作为异常事件的检测算法,针对实际的复杂场景进行研究,现实快速准确的异常事件检测。

2 基于深度学习的异常事件检测

在现实场景中,异常事件检测要适应不同复杂的场景,本文针对这些难点,提出一种新的基于深度学习的网络训练方法。本文将深度学习中的 GoogLeNet 引入到异常事件的检测中,通过 GoogLeNet 基本网络结构,结合一些网络训练技巧,对网络进行修正,训练出一种适合异常事件检测的模型,接下来对本文的异常事件检测算法进行介绍。

本文异常事件检测分成三部分,数据前期处理,模型训练及测试,异常事件检测。其中用 PKU-SVD-B 视频集进行数据前期处理,并将数据集以 7:3 的比例作为模型的训练集和测试集。

(1) 数据预处理

图像预处理包括获取图像 ROI (Region Of Interest)、图像下采样、添加光流信息、数据筛选、数据增强。

(2) 初始化网络权

本文采用 Xavier 初始化方法进行权重的初始化,首先,如式(1)所示。

$$Y = W_1 X_1 + W_2 X_2 + \cdots + W_n X_n \quad (1)$$

其中, Y 表示某层的输出矩阵, X 表示某层的输入矩阵, W 表示输入节点连接输出节点的权重矩阵。

由于式(1)中各 X 与各 W 均互相独立,所以 X 与 W 之间的方差式如(2)所示。

$$\begin{aligned} \text{Var}(W_i X_i) &= E[X_i]^2 \text{Var}(W_i) \\ &\quad + E[W_i]^2 \text{Var}(X_i) \\ &\quad + \text{Var}(W_i) \text{Var}(X_i) \end{aligned} \quad (2)$$

其中, $\text{Var}(W_i)$ 是权重矩阵的方差, $\text{Var}(X_i)$ 是输入矩阵的方差, $E[X_i]$ 与 $E[W_i]$ 是权重和输入的均值。

在实际网络参数设置时,期待 $E[X_i]$ 与 $E[W_i]$ 均为 0,所以在初始化权重时,将式(2)进行简化,简化结果如式(3)所示。

$$\text{Var}(W_i X_i) = \text{Var}(W_i) \text{Var}(X_i) \quad (3)$$

由于 W 与 Y 相互独立,则根据式(1)推出式(4)。

$$\begin{aligned} \text{Var}(Y) &= \text{Var}(W_1 X_1 + W_2 X_2 + \cdots + W_n X_n) \\ &= n \text{Var}(W_i) \text{Var}(X_i) \end{aligned} \quad (4)$$

其中, $\text{Var}(Y)$ 表示输出方差, n 表示输入的数量。如果想使得输入方差等于输出方差,则表示 $n \text{Var}(W_i) = 1$,可以推出式(5)。

$$\text{Var}(W_i) = \frac{1}{n} = \frac{1}{n_{in}} \quad (5)$$

如果在反向传播有同样的操作,则有式(6)。

$$\text{Var}(W_i) = \frac{1}{n_{out}} \quad (6)$$

其中, n_{in} 表示输入神经元个数, n_{out} 表示输出神经元个数。在实际神经网络中,输入神经元数量与输出神经元数量大多不相等,所以推出式(7)。

$$\text{Var}(W_i) = \frac{2}{n_{in} + n_{out}} \quad (7)$$

均匀分布 $U(a, b)$ 的方差根据式(8)计算。

$$\text{Var}(U) = \frac{(b-a)^2}{12} \quad (8)$$

假设使其均值为零 $b = -a$, 则, 则式(8)简化为式(9)。

$$\text{Var}(U) = \frac{(2b)^2}{12} = \frac{2}{n_{in} + n_{out}} \quad (9)$$

可得 $b = \frac{\sqrt{6}}{\sqrt{n_{in} + n_{out}}}$, 所以最终 Xavier 初始化分布

$$\text{为} \left[-\frac{\sqrt{6}}{\sqrt{n_i + n_{i+1}}}, \frac{\sqrt{6}}{\sqrt{n_i + n_{i+1}}} \right].$$

(3) 模型训练过程

网络结构主要有数据层,卷积层,采样层,激活函数层,抛弃节点层和分类器层。网络的输入数据是具有三通道的待检测图像数据,网络的输出值是图像是否是异常事件的概率。

本文在中间层加入批归一化层还原数据分布,批归一化的算法原理如算法 1。

算法 1 批归一化算法

输入:批特征数据矩阵 X

输出:批归一化后的特征数据矩阵 Y

1. 计算上一层激活函数值的均值 $\mu = \frac{1}{m} \sum_{i=1}^m X_i$
2. 计算一层激活函数值的标准差 $\sigma^2 = \frac{1}{m} \sum_{i=1}^m (X_i - \mu)^2$
3. 将本层特征数据归一化处理 $\hat{X}_i = \frac{X_i + \mu}{\sqrt{\sigma^2 + \epsilon}}$
4. 重构层数据特征 $Y_i = \gamma \hat{X}_i + \beta$
5. 输出批归一化后的特征数据 Y

其中, γ 和 β 为两个训练网络时所需要调整的参数,第 3

步公式中的 ε 是一个接近 0 的常数.

(4) 模型测试过程

模型测试阶段的主要任务是测试训练模型的好坏,当测试集的准确率达到一定范围,便终止训练.测试网络跟训练网络一样,只是没有代价函数层,不能进行网络的反向传播,只是测试集在训练出的网络结构上复现一遍,最终的结果依据概率判断所属类别.

(5) 模型检测过程

异常事件的检测与模型测试过程相似,在不更改网络任何参数值的前提下,用检测集在训练好的模型上测试一遍,每张检测图像在检测模型上运行一遍,输出预测标签结果即可.

3 实验及性能分析

3.1 实验环境

本文中,采用 C++ 与 Python 双语言混合编程实现异常事件的检测与群体异常事件的分类检测.实现具体硬件配置详情,见表 1.

表 1 实验环境配置表

名称	配置
计算机	CPU: Intel(R) Core(TM) i7-7700
	内存: 8.00GB
	硬盘: 1TB
	OpenCV 版本: 2.4
服务器	Python 版本: 2.7
	CPU: Intel(R) Exon E5-2620 v3
	GPU: NVIDIA Tesla K40m
	内存: DDR3 16GB
	显存: 12GB
	硬盘: 1TB
	Caffe 版本: 1.0

3.2 异常事件检测的实验及分析

本部分主要通过大量实验找到了一种能够适应多种现实场景的深度学习框架的异常检测算法.通过对网络模型的训练,调整网络的输入格式、batch_size(批处理数)、learning_rate(学习率),达到最好的异常检测效果.

3.2.1 模型训练及分析

本文使用有 22 层网络结构的 GoogLeNet 作为原始的网络模型结构.主要是卷积和采样层的交替,以及 Inception 模块,网络模型的详细结构请参考^[13].损失函数采用逻辑回归损失函数,最后的输出层的神经元数根据算法功能设置成 2.对于深度学习网络模型的训练,网络的输入格式、batch_size(批处理数)、learning_rate

(学习率)都会影响模型的效果.

(1) 模型输入图片格式的选择

模型训练时间是一项衡量模型的指标.本文实验了 700 × 700pixel、384 × 384pixel 与 224 × 224pixel 三种输入规格图像的训练,实验结果如表 2 所示.

表 2 图像大小对模型收敛时间的影响结果

图像大小/pixel	模型收敛所需时间/min
700 × 700	1727
384 × 384	805
224 × 224	697

从表 2 对比所示,当图像大小为 224 × 224pixel 时,模型达到收敛的时间最少.对于分辨率大的图像,网络需要处理的像素值多,所要消耗的时间必然多.分辨率过小的图像包含的信息量少,不利于图像的特征提取.

(2) 模型 batch_size 的选择

batch_size 的设置跟 GPU 与输入图像大小有关,如果输入图像过大,batch_size 就要适当的设置小一点,不然会出现内存溢出的错误提示,本文中输入图像大小是 224 × 224pixel,经过实验最后设置的 batch_size 是 25.

(3) 模型 learning_rate 的选择

对于深度学习网络模型的训练,学习率能够影响模型最终的性能.本文在同一网络结构下,分别取模型 learning_rate 为 0.1、0.01 以及 0.001 进行模型训练,经过测试得到如表 3 所示的结果.

表 3 学习率对异常事件检测模型的影响

学习率	准确率/%	召回率/%
0.1	87.5	76.2
0.01	91.2	87.8
0.001	95.9	90.3

从表 3 中看出,随着学习率减少,准确率与召回率均随之提升,设置大的学习率可能会使实验结果跳过最优点而取得局部最优点,使准确率偏低,设置小的学习率,虽然训练速度相对降低,但会取得更接近最优点的结果,所以准确率和召回率相对偏高.

最终训练模型的输入图像大小设置为 224 × 224pixel, batch_size 为 25,同时训练模型的 learning_rate 为 0.001.确定好网络结构及参数后,进行模型训练及测试,图 1(a)是训练过程中的代价值(Loss)的下降曲线图.

如图 1(a)所示,训练初期,随机初始化网络权值,网络稳定性弱,随着网络迭代次数的增加,网络稳定性增强,迭代到 5000 次时, Loss 值正常下降,当迭代到 10000 次时, Loss 值不再下降,从 10000 到 20000 期间,

网络有一点小波动,迭代到 20000 次以后, Loss 值几乎收敛于 0. 迭代到 40000 次后,模型稳定.

3.2.2 模型测试及分析

图 1(b) 是测试模型的 Accuracy 曲线图. 从图 1(b) 中看出,模型在迭代 20000 次之后,基本趋于稳定. 模型具有较高泛化能力. 正向传播输出一个检测正确率,通过输出的检测正确率,得到可视化结果,直观看出模型的效果.

本文在异常检测领域引入了深度学习框架,用 end-to-end 的方法训练并检测异常事件,更好的发挥了深度学习的优势.

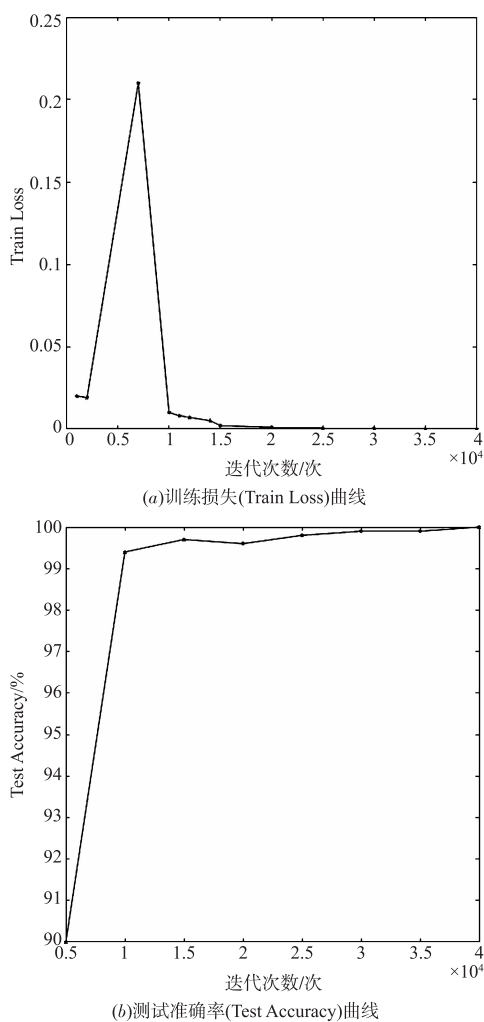


图1 检测模型训练损失曲线与测试准确率曲线

3.2.3 模型检测及分析

模型验证的过程,本文准备了 100 张正常图像以及 100 张异常图像,大小为 224×224 pixel. 验证用的网络结构与训练时用的网络架构大部分一样,网络没有损失输出,没有迭代修正网络参数的功能,最后多了一个 Softmax 概率层,根据输出的概率大小,得出预测结果. 图 2 是检测模型效果图. abnormal 标签表示此张图像模

型预测为异常事件,normal 标签表示此张图像模型预测为正常事件.



图2 异常事件检测效果图

3.2.4 算法对比及分析

为了更好的验证本文提出异常事件检测的方法,我们在 UMN 数据集^[14]上测试本文的方法. 本文将实验结果与传统光流法 (Optical Flow)、社会力模型 (Social Force Model)^[4]、纹理模型 (Streakline Model)^[5] 和交互势能模型 (Interaction Energy Potentials)^[6] 作了比较. 通过计算正检率 (True Positive Rate, TPR) 与误报率 (False Positive Rate, FPR) 来绘制受试者工作特征曲线 (ROC 曲线) 和曲线下的面积 (Area Under Curve, AUC) 来反应本方法的性能. 这些方法的 ROC 曲线如图 3 所示, AUC 在表 4 中所示.

通过实验结果可以看到本文的方法在 UMN 数据集上优于其他几种方法. 从 ROC 曲线上可以看到我们的方法在误报率很低的情况下正检率已经很高了,与之相比较的 3 种方法在误报率低于 20% 时正检率都低于 50%. 从整条曲线上看,我们的方法的正检率一直高于 90%,且通过 AUC 来看,本文提出的方法的 ROC 曲线下的面积最大,因此可以看出,本文的方法在处理多场景群体异常事件检测上相较于其他几种方法更加有效.

表 4 UMN 数据集上各种方法的 ROC 曲线面积

方法	Optical Flow	Streakline Model	Social Force Model
AUC	84%	90%	96%
方法	Interaction Energy Potentials	Ours	
AUC	98.5%	99%	

3.3 异常事件分类的实验及分析

本部分主要通过大量实验找到了一种能够适应多种现实场景的深度学习框架的异常事件分类算法. 通过对异常事件检测网络模型的修改,调整网络的输入方式,达到最好的异常分类效果.

3.3.1 模型训练及测试分析

分类模型训练过程同检测模型训练过程类似,输入图像大小设置为 224×224 pixel, batch_size 为 25,

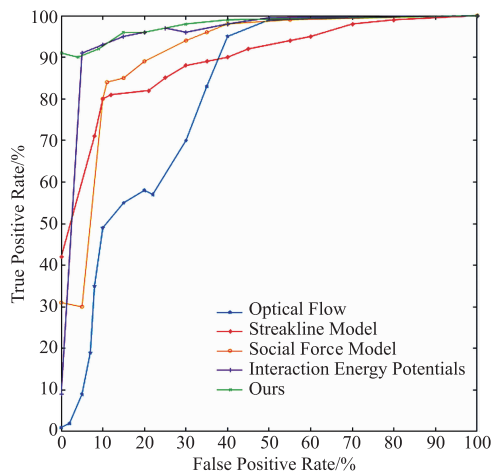


图3 各种异常检测方法的ROC曲线

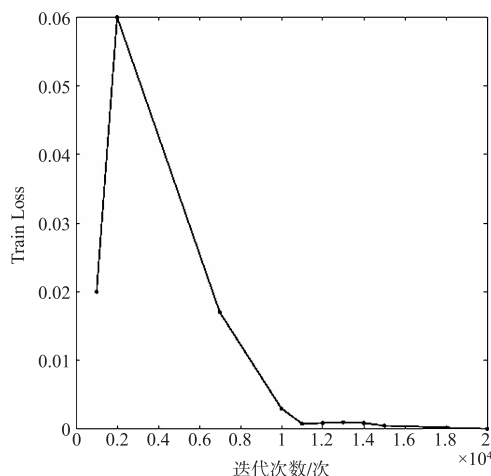
learning_rate 为 0.001. 网络中间层的结构不变,修改代价损失函数,将逻辑回归损失函数修改为 Softmax 交叉熵损失函数,并在最后输出层将输出数目与标签类别数设置成一样,所以最后一层神经元个数设置成 5,包括 4 类群体异常和正常事件. 修改好网络结构,重新调整数据分布,设置好训练集和测试集,以及设置好各网络参数之后,进行模型训练及测试,图 4(a) 是训练过程中 Loss 下降曲线.

如图 4(a) 所示,训练初期,随机初始化网络权值,网络不稳定, Loss 下降曲线出现动荡,随着网络训练迭代次数增加,网络稳定性增强,迭代到 2000 次时, Loss 值开始正常下降,当迭代到 10000 次时,开始缓慢下降,迭代到 15000 万次以后, Loss 值几乎收敛到 0,模型趋于稳定. 网络模型训练完成后,利用测试集对模型进行测试,如图 4(b) 是测试数据集的 Accuracy 曲线图,从曲线中可以看到模型对测试集的分类准确率达到 80% 左右,最后准确率大约稳定在 85% 左右,准确率相对偏高,虽然还没有达到能够实际应用的要求,但是本文提出的算法对于异常事件的分类检测具有一定的贡献.

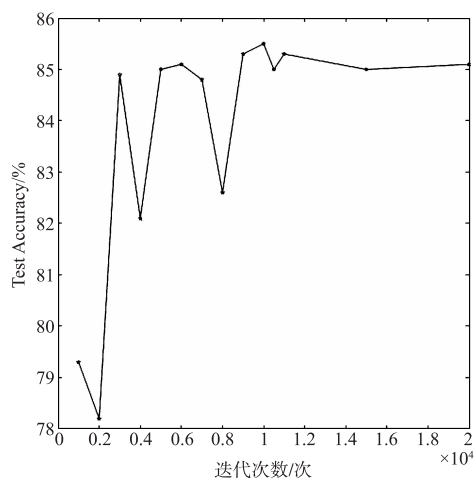
3.3.2 模型检测及分析

模型验证过程,本文准备了 100 张异常图像以及 50 张正常图像. Normal 代表正常事件、abnormal_a 代表群体发散聚集事件、abnormal_b 代表群体密集聚集事件、abnormal_c 代表群体逃散事件、abnormal_d 代表追赶事件. 部分检测效果图如图 5 所示.

本文提出的异常事件分类方法能够从正常事件中分类出群体发散聚集事件,群体密集聚集事件,群体逃散事件和追赶事件这 4 种异常事件,当遇到新的异常检测数据时,我们提出的异常分类算法可以判断出此类事件不属于正常事件,但是无法指出此类事件是哪类异常事件,此时如果想要正确判断异常事件的类别则需要重新调整网络结构,添加新的一类输出,并重新进



(a) 训练损失(Train Loss)曲线



(b) 测试准确率(Test Accuracy)曲线

图4 分类模型训练损失曲线与测试准确率曲线



图5 异常事件分类检测效果图

行训练,调整网络参数.

4 结语

本文以权衡检测精度和处理多场景多异常为目标,提出了一种新颖的异常事件检测算法,并且扩展出了一种异常事件分类检测方法. 本文中,先根据异常事件特征,进行数据前期预处理,然后将深度学习理论应用在异常事件检测和分类中,训练异常事件检测模型,

最后用测试集检测模型的性能,实验表明,本文提出的异常事件检测和分类算法对于多场景下,多异常事件的检测具有很高的准确率。

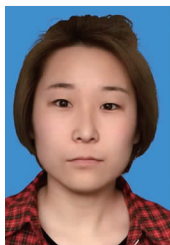
参考文献

- [1] 郭倩如. 监控视频中异常事件检测方法研究[D]. 杭州: 杭州电子科技大学, 2017. 3-7.
- [2] Ahmed, Sk. Arif, Dogra, Debi Prosad, et al. Surveillance scene representation and trajectory abnormality detection using aggregation of multiple concepts[J]. Expert Systems with Applications, 2018, 101(1): 43-55.
- [3] Gnouma M, Ejbali R, Zaied M. Abnormal events. detection in crowded scenes[J]. Multimedia Tools & Applications, 2018, (3): 1-22.
- [4] Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model[A]. IEEE/ACM Computer Vision and Pattern Recognition[C]. Miami, Florida, USA: IEEE, 2009. 20-25.
- [5] Mahadevan V, Li W, Bhalodia V, et al. Abnormally detection in crowded scenes[A]. IEEE/ACM Computer Vision and Pattern Recognition[C]. San Francisco, California, USA: IEEE, 2010. 1975-1981.
- [6] Cui X, Liu Q, Gao M, et al. Abnormal detection using interaction energy potentials[A]. IEEE/ACM Computer Vision and Pattern Recognition[C]. Colorado, USA: IEEE, 2011. 3161-3167.
- [7] 黄荷, 俞亚萍, 张之江. 基于神经网络的密集人群视频异常检测[J]. 电子测量技术, 2017, 40(11): 103-107.
- [8] Wu S, Wong H S, Yu Z. A Bayesian model for crowd escape behavior detection[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(1): 85-98.
- [9] 孙慧. HMM 监控视频下的异常事件检测分析[J]. 信息与电脑(理论版), 2017(20): 42-54.
- [10] Kumar D, Bezdek J C, Rajasegarar S, et al. A visual-numeric approach to clustering and anomaly detection for trajectory data[J]. Visual Computer International Journal of Computer Graphics, 2017, 33(3): 265-281.
- [11] 伟峰. 视频监控中异常事件实时检测方法研究与实现[D]. 北京: 中国科学技术大学, 2017. 45-48.
- [12] 刘顺华. 基于监控视频的人群异常事件检测[D]. 郑州: 郑州大学, 2017.
- [13] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[A]. IEEE/ACM Conference on Computer Vision and Pattern Recognition[C]. Boston, USA: IEEE, 2015. 1-9.
- [14] University of Minnesota. Unusual Crowd Activity Dataset [EB/OL]. <http://mha.cs.umn.edu/movies/crowdactivity-all.av>, 2019.

作者简介



闻 佳 女, 1981 年生于黑龙江齐齐哈尔. 现为燕山大学副教授、硕士生导师. 主要研究方向为模式识别和计算机视觉.
E-mail: wjyanyuan@aliyun.com



王宏君(通信作者) 女, 1994 年生于河北承德. 现为燕山大学信息科学与工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: 2642264392@qq.com



邓 佳 女, 1992 年生于湖南怀化. 现为燕山大学信息科学与工程学院硕士研究生. 主要研究方向为机器学习.
E-mail: 1351270663@qq.com



刘鹏飞 男, 1994 年生于河北保定. 现为燕山大学信息科学与工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: 2353865452@qq.com