

# 基于深度聚类的人群运动行为检测

杨玉成<sup>1,2</sup> 张 乾<sup>1,2</sup> 岳诗琴<sup>1</sup> 邵定琴<sup>1</sup>

<sup>1</sup>(贵州民族大学数据科学与信息工程学院 贵州 贵阳 550025)

<sup>2</sup>(贵州省模式识别与智能系统重点实验室 贵州 贵阳 550025)

**摘 要** 针对人群运动行为检测中难以有效去除异常点的问题,提出一种基于深度聚类的人群运动行为检测方法。构造一种运动特征描述符,通过 K-medoids 算法对运动特征描述符聚类;采用 SqueezeNet 轻量级正向加速网络提取视频帧图像特征,通过 K-medoids 算法对图像特征聚类;将图像特征聚类指数作为运动特征描述符聚类指数的约束条件去除异常点;将去除异常点的聚类指数映射至每一帧图像实现人群检测。在国际公开数据集上进行实验,实验结果表明该算法有效提升了异常点去除率和聚类纯度。

**关键词** 行为检测 人群运动 深度聚类 K-medoids SqueezeNet

中图分类号 TP3

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2020.11.019

## CROWD MOTION BEHAVIOR DETECTION BASED ON DEEP CLUSTERING

Yang Yucheng<sup>1,2</sup> Zhang Qian<sup>1,2</sup> Yue Shiqin<sup>1</sup> Shao Dingqin<sup>1</sup>

<sup>1</sup>(School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, Guizhou, China)

<sup>2</sup>(Pattern Recognition and Intelligent Systems Key Laboratory of Guizhou, Guiyang 550025, Guizhou, China)

**Abstract** Aiming at the problem that it is difficult to effectively remove abnormal points in crowd motion behavior detection, we propose a crowd motion behavior detection method based on deep clustering. A motion descriptor was constructed, and K-medoids algorithm was used to cluster motion descriptors; the SqueezeNet lightweight forward acceleration network was used to extract video frame image features, and the K-medoids algorithm was used to cluster image features; the image feature clustering index was used as the constraint condition of motion descriptor clustering index to remove abnormal points; the clustering index of removing abnormal points was mapped to each frame image to realize crowd detection. The experimental results on international open datasets show that our algorithm can effectively improve the removal rate of abnormal points and the clustering purity.

**Keywords** Behavior detection Crowd motion Depth clustering K-medoids SqueezeNet

## 0 引 言

人群运动是公共场所中一种极为普遍的行为现象,随着人口数量的不断增加,公共场所人群密集程度也随之增加,这给公共场所人群管理与安全防控带来巨大难题<sup>[1]</sup>,因此人群运动行为检测成为公共安全领域的热点研究课题。人群运动行为检测涉及模式识别、计算机视觉、机器学习和人工智能等领域理论知

识。经过多年发展,人群运动虽然在理论体系和检测技术上取得诸多成绩,但仍然存在一些技术瓶颈尚未得到很好的解决,例如人群密集的场所中人群个体间存在严重的遮挡、其他非行人混入人群中、人群中个体运动方向和速度不一致性等。这些因素给人群运动行为检测带来巨大挑战。

场景人群聚类实质是对场景中的运动人群进行检测<sup>[2]</sup>,目前人群检测算法主要分为三大类:(1)基于轨迹点的算法;(2)基于视图的算法;(3)基于深度学习

的算法。基于轨迹点的算法主要是根据人群运动轨迹的位置信息和时间信息进行检测。Takahashi 等<sup>[3]</sup>利用关键轨迹特征识别运动人群动作;Zou 等<sup>[4]</sup>利用关联轨迹点检测运动人群;Zhao 等<sup>[5]</sup>提出一种基于轨迹点光流直方图的算法检测异常人群,虽然基于人群运动轨迹点的算法能够快速检测异常人群,但是人群的轨迹信息与人群实际运动存在偏差<sup>[2]</sup>,所以该类算法检测结果可信度不高。基于视图的算法是指从不同的视频角度对运动人群进行检测。Peng 等<sup>[6]</sup>提出一种多视图贝叶斯模型对行人进行检测;Tian 等<sup>[7]</sup>将三维光谱图像切分为三个视图进行聚类;Qian 等<sup>[8]</sup>提出一种基于嵌入图的多视图聚类算法,虽然基于多视图的算法能够较为准确地检测人群,但是该类算法需要先验知识确定阈值,而且算法鲁棒性较差;Wang 等<sup>[2]</sup>提出一种自加权多视图人群聚类算法框架,虽然该算法不依赖于任何阈值参数并获得较强鲁棒性,但是该算法计算复杂度较高,在一定程度上增加了计算成本。深度学习算法在图像识别、姿态估计和语音识别等领域表现出优越的性能<sup>[9]</sup>,同时也被广泛应用于人群运动行为检测领域。Shuaibu 等<sup>[10]</sup>提出一种三维 CNN (Convolutional Neural Networks) 用于人群场景理解。Keçeli 等<sup>[11]</sup>在迁移学习的基础上提出一种人群暴力行为检测算法。Zhang 等<sup>[12]</sup>提出一种基于 SSD 框架的拥挤场景人群检测算法。Tzelepi 等<sup>[13]</sup>在 CNN 的基础上提出一种无人机自动人群检测算法,虽然深度学习算法具有强大的特征学习能力,但网络学习到的特征包含大量异常点信息给计算增加了负担并影响效率。

针对上述问题,本文提出一种基于深度聚类的人群运动行为检测算法。首先将人群运动轨迹的位置和速度组合得到运动特征描述符,通过 K-medoids 算法对运动特征描述符进行聚类得到聚类指数  $U_1$ ,从而将人群场景划分为不同的运动区域;然后通过 SqueezeNet 轻量级正向加速网络提取视频帧图像特征,再次采用 K-medoids 算法对图像特征进行聚类得到聚类指数  $U_2$  并将其作为  $U_1$  的约束条件去除异常点;最后将去除异常点的聚类指数映射至每一帧图像,从而实现人群检测。

## 1 SqueezeNet 正向加速网络

由于传统 CNN 网络存在训练周期长、参数多、计算内存需求大等缺陷,因此 Iandola 等<sup>[14]</sup>提出一种轻

量级 CNN 网络 SqueezeNet,其卷积核参数仅为传统 CNN 网络的 1/9;为减少网络输入参数,在传统 CNN 网络基础上增加压缩层,用于压缩输入图像的通道数;为减少信息损失将池化层置于卷积层后。

SqueezeNet 网络结构主要由输入层、卷积层、Fire 模块、池化层、输出层构成。Fire 模块的结构如图 1 所示,该模块由压缩卷积层和扩展卷积层组合而成,其中扩展卷积层由  $1 \times 1$  卷积核和  $3 \times 3$  卷积核构成。Fire 模块是 SqueezeNet 网络的核心组成部分,主要作用是减少图像输入通道数和卷积参数,其输出结果被作为整个网络的中间参数。在图像特征学习过程中,首先对输入图像进行卷积操作,然后经过 8 个 Fire 模块连续参数压缩,最后通过一层卷积操作得到  $1 \times 1\ 000$  维的特征向量。SqueezeNet 的网络结构如图 2 所示。

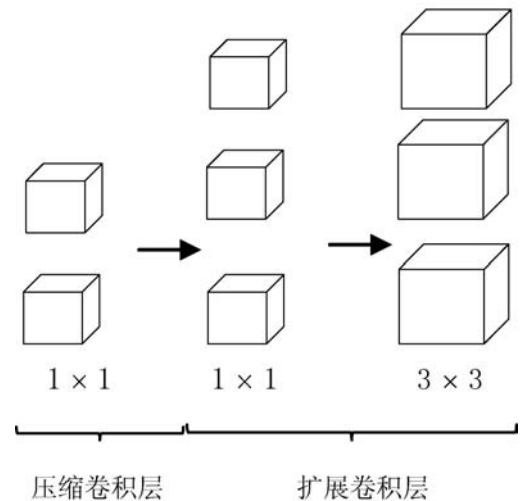


图 1 Fire 模块

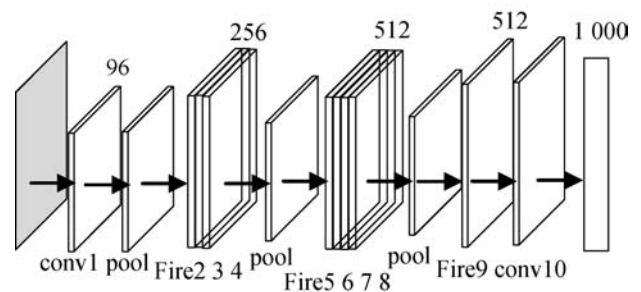


图 2 SqueezeNet 网络

虽然 SqueezeNet 网络与传统 CNN 网络相比性能有很大提升,但是网络计算速度仍然较慢。在 CNN 网络中,层与层之间、卷积核之间、输出特征图等多为并行结构。鉴于此,Motamedi 等<sup>[15]</sup>提出一种基于 FPGA 的加速器框架,该框架充分利用网络的并行结构对 SqueezeNet 正向网络加速。

FPGA 加速器框架主要由乘法器和加法器组成,具体结构如图 3 所示。第一层为乘法器层,由  $K$  个乘

子构成;第二层为加法器层,该层的主要作用是对乘法器层的结果进行累积。将乘法器层中的乘法子集与加法器中对应位置的加法子集进行组合得到 PCE(Parallel Convolution Engine) 模块;在网络输出层将 PCE 模块与加法子集组合构成并行输出模块。为使网络达到最佳性能,需要计算输入特征映射  $n$ 、乘子  $K$  和并行输出模块  $m$  的最优值。

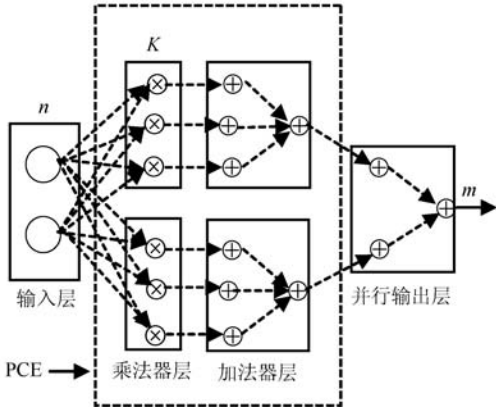


图3 FPGA 加速器结构

首先定义一个周期函数  $T$  如下:

$$T = \frac{f_p \times S}{m \times n \times W_p} \times \frac{W_p + K \times l}{K} \quad (1)$$

该函数即为乘法运算和加法运算次数。式中:  $f_p$  为特征总数;  $W_p$  为输出特征;  $S$  为卷积核的大小;  $l$  为信息传递损失量。然后定义一个网络运算总数计算函数:

$$N = 2 \times f_p \times S \quad (2)$$

最后定义一个峰值计算函数:

$$R = \frac{2 \times m \times n \times W_p \times K}{W_p + K \times l} \quad (3)$$

在式(1) - 式(3)中  $f_p$ 、 $W_p$ 、 $S$  为固定值,由此可以计算得到  $m$ 、 $n$ 、 $K$  的最优值。

FPGA 加速器强大的加速性能使 SqueezeNet 正向网络运行速度达到最佳状态。SqueezeNet 正向加速网络提取一幅图像的特征仅需  $22 \pm 1$  s,因此本文采用 SqueezeNet 正向加速网络提取图像特征。

## 2 K-medoids 算法

机器学习算法分为监督学习和无监督学习两大类。监督学习是对带有标签信息的数据进行学习从而得到数据的类别数;无监督学习是对无标签数据进行学习从而得到数据的聚类簇。

K-medoids 算法是无监督学习算法中的经典算法,

由于在处理聚类问题时表现出优越的性能,因此被广泛应用于人群检测领域。K-medoids 算法的核心思想:首先,选取  $K$  个数据点;然后,计算每个数据点  $\mathbf{X} = (x_1, x_2, \dots, x_k)$  与  $K$  的距离,将与  $K$  距离最近的点划分为同一聚类簇,同时更新  $K$  个数据点;最后,聚类函数收敛,得到总体数据的聚类簇个数<sup>[16]</sup>  $\mathbf{C} = (c_1, c_2, \dots, c_k)$ 。聚类函数定义为<sup>[17]</sup>:

$$D_{\min} = \sum_{i=1}^K d(\mathbf{x}_i, \mathbf{o}_i) \quad (4)$$

式中:  $\mathbf{o}_i$  为  $K$  个数据点的中心;  $d(\mathbf{x}_i, \mathbf{o}_i)$  是数据点与中心的距离。

$$d(\mathbf{x}_i, \mathbf{o}_i) = \|\mathbf{x}_i - \mathbf{o}_i\|_2^2 \quad (5)$$

$\mathbf{o}_i$  的更新准则定义为<sup>[17]</sup>:

$$\mathbf{o}_i = \frac{1}{|c_i|} \sum_{j=1}^{c_i} c_{i,j} \quad (6)$$

$D_{\min}$  的值表示聚类簇内各数据点接近中心的程度,  $D_{\min}$  的值越小,表明各数据点与中心的接近程度越高,各数据点间的相似度也就越高。本文采用 K-medoids 算法对运动特征描述符和视频帧图像特征进行聚类。

## 3 深度网络人群运动聚类

公共场所人流量较大,人群密集度较高,通常混合了不同运动状态的人群目标,例如:移动方向、速度、轨迹等不同运动状态,同时还混合了其他非人群目标,例如:汽车、树木等目标。本文的研究重点是对不同运动状态的人群进行聚类,即检测不同运动状态的人群。

人群遮挡和光照等因素使人群区域内出现大量倒影,由于人群运动过程中倒影也随之运动,因此产生大量异常点数据,其次 K-medoids 的初始聚类中心是随机选取的,因此在聚类时易出现局部最优值<sup>[17]</sup>。上述因素直接影响聚类效果。鉴于此,本文融合 SqueezeNet 正向加速网络与 K-medoids 算法提出一种深度聚类人群运动行为检测算法。

### 3.1 运动区域划分

运动轨迹和运动速度是人群运动的重要特征,本文通过人群运动轨迹的位置信息和运动速度,定义一个运动特征描述符:

$$\mathbf{A}_{n \times 4} = [\mathbf{V}_{n \times 2}, \mathbf{P}_{n \times 2}] \quad (7)$$

式中:  $n$  为轨迹点数,  $\mathbf{V}$  是  $n \times 2$  维的速度矩阵,  $\mathbf{P}$  是  $n \times 2$  维的轨迹矩阵。引入运动特征描述符之后,采用 K-me-

doids 算法将人群场景划分为四个不同的运动区域,被划分在同一区域内的人群个体的运动状态是相同的,即运动速度和方向相同。具体划分过程为:(1) 采用 K-medoids 算法对运动特征描述符  $A_{n \times 4}$  进行聚类得到聚类指数  $U_1 = (u_{1,1}, u_{1,2}, \dots, u_{1,n}), u_{1,i} \in [1, 4]$ ; (2) 将  $\mu_{(u_{1,i}, V_n)} \rightarrow M, M$  表示图像。划分结果如图 4 所示。

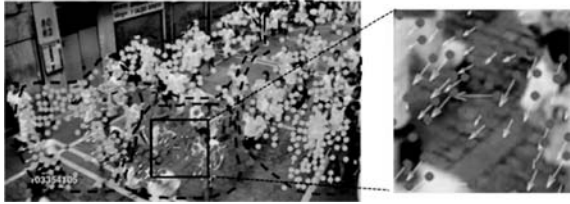


图 4 人群运动区域

图 4 所示为人群场景运动区域划分结果,不同颜色(图中用灰度描述)所覆盖的区域表示不同的运动区域,属于同一运动区域的人群运动状态相同。

### 3.2 深度人群聚类

本文将 SqueezeNet 正向加速网络与 K-medoids 算法融合,提出一种深度聚类算法用于人群聚类。人群场景图像能够直观地反映人群的运动状况,深度学习算法可完整地表达图像特征信息,因此本文将 SqueezeNet 正向加速网络用于图像特征提取,采用 K-medoids 算法对特征进行聚类。

首先采用 SqueezeNet 正向加速网络提取视频帧图像特征得到一个  $1 \times 1\,000$  维的特征向量。然后采用 K-medoids 算法对所提取的特征向量进行聚类得到聚类指数  $U_2 = (u_{2,1}, u_{2,2}, \dots, u_{2,1\,000}), u_{2,i} \in [1, 4]$ , 场景中的人群较为密集,在光照等因素影响下,人群之间会产生许多噪声,即  $U_1$  和  $U_2$  中存在异常特征点,这些异常特征点直接影响聚类结果,因此本文定义一个约束条件如下:

$$U'_1 = \begin{cases} U_1 & |U_1| \leq |U_2| \text{ 且 } u_{1,i} < u_{2,i} \\ 0 & \text{其他} \end{cases} \quad (8)$$

$$u'_{1,i} = \begin{cases} u_{1,i} - 1 & u_{1,i} > u_{2,i} \\ 0 & \text{其他} \end{cases} \quad (9)$$

由式(8)和式(9)得到去除异常点之后的人群聚类指数  $U'_1 = (u'_{1,1}, u'_{1,2}, \dots, u'_{1,n})$ , 最后将  $\mu_{(u'_{1,i}, V_n)} \rightarrow M$ , 从而实现人群聚类。

### 3.3 算法流程

本文将 K-medoids 算法与 SqueezeNet 正向加速网络融合,提出一种深度聚类算法,算法流程如算法 1 所示。

#### 算法 1 本文算法

输入视频帧图像,人群运动轨迹

SqueezeNet 提取图像特征:  $f = 1 \times 1\,000$

计算运动特征描述符:  $A_{n \times 4} = [V_{n \times 2}, P_{n \times 2}]$

计算聚类指数:

K-medoids 对  $f$  聚类得  $U_1 = (u_{1,1}, u_{1,2}, \dots, u_{1,n})$

K-medoids 对  $A_{n \times 4}$  聚类得  $U_2 = (u_{2,1}, u_{2,2}, \dots, u_{2,1\,000})$

去除异常点:

If  $|U_1| \leq |U_2| \& |u_{1,i}| \leq |u_{2,i}|$

then  $U'_1 = U_1$

Else  $U'_1 = 0$

End if

If  $u_{1,i} > u_{2,i}$

$u'_{1,i} = u_{1,i} - 1$

Else  $u_{1,i} = 0$

End if

输出  $U'_1 = (u'_{1,1}, u'_{1,2}, \dots, u'_{1,n})$

映射至图像:  $\mu_{(u'_{1,i}, V_n)} \rightarrow M$

## 4 实验

针对本文提出的基于深度聚类的人群运动行为检测算法,通过在各种国际公开数据集上进行实验,结果表明本文算法能够准确检测场景中的不同人群,与其他人群运动行为检测算法相比,本文算法具有更高的异常点去除率 (Abnormal Point Removal Rate, APR) 和纯度值 (PU)。

本文所有验证实验平台为 Windows 7 64 位系统, CPU 2.5 GHz, RAM 4.0 GB, MATLAB R2017b。本文选取了 CCD (CUHK Crowd Dataset)<sup>[18]</sup>、CMD (Collective Motion Database)<sup>[19]</sup>、MPT (MPT-20 × 100)<sup>[20]</sup> 三个公开数据集进行实验,并将本文提出的算法与 MCC (Measuring Crowd Collectiveness)<sup>[19]</sup>、CF (Coherent Filtering)<sup>[21]</sup> 两种主流人群运动行为检测算法进行实验对比。为客观评价提出算法的性能,选择人群聚类的 PU 值和 APR 值作为综合评价指标。

### 4.1 人群聚类评价

图 5 所示为本文算法与其他人群聚类算法在三个国际公开数据集上的人群聚类结果,其中相同颜色(图中用灰度描述)点所覆盖的区域表示属于同一类的人群。图 6 所示为本文算法去除人群场景中异常点的结果,图中“+”符号表示被去除的异常点。实验结果表明,本文算法能够对场景中的不同人群进行有效

聚类,并且能够有效去除人群场景中由于遮挡和光照影响所产生的异常点数据,证明了本文算法在人群聚类和异常点去除问题上的优越性能。

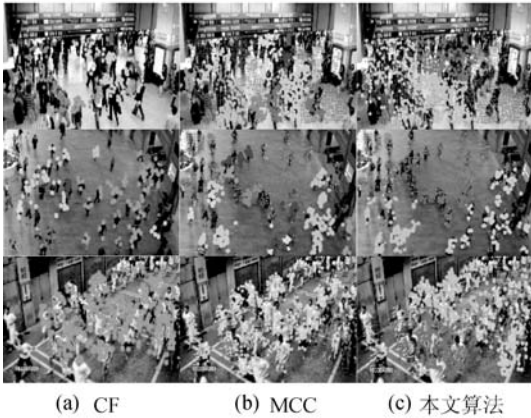


图5 各种算法的聚类结果



图6 异常点去除

为了进一步验证本文算法的优越性,表1统计出了本文算法与其他算法在三个不同数据集上去除的异常点数和APR值。在CMD数据集上,本文算法去除的异常点数比CF算法多170个点,比MCC算法多48个点;本文算法的APR值比CF算法高0.05,比MCC算法高0.08。在CCD和MPT数据集上,本文算法去除的异常点数和APR值均明显优于其他人群聚类算法。

表1 不同聚类算法去除异常点数和APR值

数据集	CF		MCC		本文算法	
	异常点数	APR	异常点数	APR	异常点数	APR
CCD	60	0.25	709	0.59	<b>769</b>	<b>0.63</b>
CMD	179	0.53	301	0.50	<b>349</b>	<b>0.58</b>
MPT	124	0.32	189	0.32	<b>261</b>	<b>0.45</b>

## 4.2 与其他算法比较结果

为更加客观地证明本文提出算法的优越性能,本文在CCD数据集上进行实验,并比较本文提出算法与其他人群聚类算法的PU值。表2统计了本文算法与其他人群聚类算法在CCD数据集上的PU值,可以看出,本文算法的PU值比MCC算法高0.079,比CF算法高0.066,进一步证明了其优越性。

表2 不同聚类算法的各项指标值

评价指标	CF	MCC	本文算法
PU	0.647	0.634	<b>0.713</b>

为更加综合地证明本文算法的有效性,本文对各种人群聚类算法的聚类数进行对比。图7统计了本文算法与其他人群聚类算法的聚类数,本文算法所得聚类数与Groundtruth一致,CF算法比Groundtruth多2类,MCC算法比Groundtruth多10类,进一步证明了本文算法的优越聚类性能。

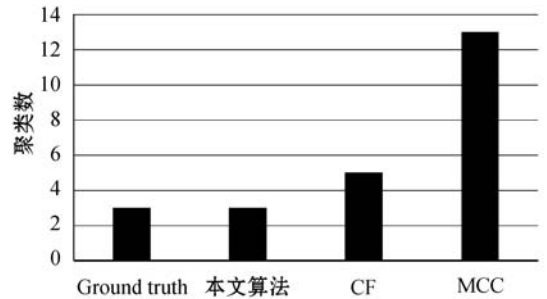


图7 各算法聚类数

## 5 结语

为解决现有算法不能有效去除异常点的问题,本文提出一种基于深度聚类的人群运动行为检测算法。将SqueezeNet正向加速网络与K-medoids算法融合去除人群异常点,通过在各种国际公开数据集进行实验对比,本文算法在准确率与目前主流算法相同的情况下,得到了更高的异常点去除率和更高的聚类纯度,实验结果证明了本文算法的有效性。未来将继续探索SqueezeNet正向加速网络的优化以及与其他聚类算法的结合,进一步提高人群检测算法的准确率和其他各项性能。

## 参 考 文 献

- [1] Xu M L, Li C X, Lv P, et al. An efficient method of crowd aggregation computation in public areas[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28(10): 2814 - 2825.
- [2] Wang Q, Chen M L, Nie F P, et al. Detecting coherent groups in crowd scenes by multiview clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(1): 46 - 58.
- [3] Takahashi M, Naemura M, Fujii M, et al. Human action recognition in crowded surveillance video sequences by using features taken from key-point trajectories[C]//Computer Vision and Pattern Recognition Workshops (CVPRW), 2011: 9 - 16.
- [4] Zou Y, Zhao X, Liu Y C. Detect coherent motions in crowd scenes based on tracklets association[C]//2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015: 4456 - 4460.

- [ 5 ] Zhao K, Liu B, Li W H, et al. Anomaly detection and localization: A novel two-phase framework based on trajectory-level characteristics [ C ]//2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE, 2018:1-6.
- [ 6 ] Peng P X, Tian Y H, Wang Y W, et al. Robust multiple cameras pedestrian detection with multi-view Bayesian network [ J ]. Pattern Recognition, 2015, 48 ( 5 ): 1760-1772.
- [ 7 ] Tian L, Du Q. Parallel multi-view low-rank and sparse subspace clustering for unsupervised hyperspectral image classification [ C ]//2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). IEEE, 2018:618-621.
- [ 8 ] Qian B, Gu X G, Shu Z Q, et al. Constrained multi-view NMF with graph embedding for face clustering [ C ]//2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES). IEEE, 2018: 103-106.
- [ 9 ] Marsden M, McGuinness K, Little S, et al. ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification [ C ]//2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017:1-7.
- [ 10 ] Shuaibu A N, Malik A S, Faye I. Adaptive feature learning CNN for behavior recognition in crowd scene [ C ]//2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE, 2017:357-361.
- [ 11 ] Keçeli A S, Kaya A. Violent activity detection with transfer learning method [ J ]. Electronics Letters, 2017, 53 ( 15 ): 1047-1048.
- [ 12 ] Zhang W J, Tian L H, Li C, et al. A SSD-based crowded pedestrian detection method [ C ]//2018 International Conference on Control, Automation and Information Sciences (ICCAIS). IEEE, 2018:222-226.
- [ 13 ] Tzelepi M, Tefas A. Graph embedded convolutional neural networks in human crowd detection for drone flight safety [ J / OL ]. IEEE Transactions on Emerging Topics in Computational Intelligence. (2019-03-04) [ 2019-07-18 ]. DOI: 10.1109/TETCI.2019.2897815.
- [ 14 ] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size [ EB ]. arXiv:1602.07360, 2016.
- [ 15 ] Motamedi M, Gysel P, Akella V, et al. Design space exploration of FPGA-based deep convolutional neural networks [ C ]//2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC). IEEE, 2016:575-580.
- [ 16 ] 蒋华,季丰,王慧娇,等.改进 Kmeans 算法的海洋数据异常检测 [ J ]. 计算机工程与设计, 2018, 39 ( 10 ): 3132-3136.
- [ 17 ] 左进,陈泽茂.基于改进 K 均值聚类的异常检测算法 [ J ]. 计算机科学, 2016, 43 ( 8 ): 258-261.
- [ 18 ] Shao J, Loy C C, Wang X G. Scene-independent group profiling in crowd [ C ]//2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014:2219-2226.
- [ 19 ] Zhou B L, Tang X O, Zhang H P, et al. Measuring crowd collectiveness [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36 ( 8 ): 1586-1599.
- [ 20 ] Solera F, Calderara S, Cucchiara R. Socially constrained structural learning for groups detection in crowd [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38 ( 5 ): 995-1008.
- [ 21 ] Zhou B L, Tang X O, Wang X G. Coherent filtering: Detecting coherent motions from crowd clutters [ C ]//2012 European Conference on Computer Vision. Springer, 2012:857-871.
- ~~~~~
- (上接第 105 页)**
- [ 5 ] Viola P, Jones M J. Robust real-time face detection [ J ]. International Journal of Computer Vision, 2004, 57 ( 2 ): 137-154.
- [ 6 ] Ohn-bar E, Trivedi M M. To boost or not to boost? on the limits of boosted trees for object detection [ C ]//2016 23rd International Conference on Pattern Recognition, 2016.
- [ 7 ] 刘宏哲,杨少鹏,袁家政,等.基于单一神经网络的多尺度人脸检测 [ J ]. 电子与信息学报, 2018, 40 ( 11 ): 63-70.
- [ 8 ] Jiang H, Learned-miller E. Face detection with the faster R-CNN [ C ]//2017 12th International IEEE Conference on Automatic Face Gesture Recognition, 2016.
- [ 9 ] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39 ( 6 ): 1137-1149.
- [ 10 ] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [ C ]//European Conference on Computer Vision, 2016.
- [ 11 ] Hu P, Ramanan D. Finding tiny faces [ C ]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [ 12 ] Zhang S, Zhu X, Lei Z, et al. S3FD: single shot scale-invariant face detector [ C ]//2017 IEEE International Conference on Computer Vision, 2017.
- [ 13 ] Tang X, Du D K, He Z, et al. Pyramidbox: a context-assisted single shot face detector [ C ]//European Conference on Computer Vision, 2018.
- [ 14 ] Jain V, Learned-miller E. Fddb: a benchmark for face detection in unconstrained settings [ M ]. UMass Amherst Technical Report, 2010.
- [ 15 ] Yang S, Luo P, Chen C L, et al. WIDER FACE: a face detection benchmark [ C ]//2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.