

文章编号: 2096-1618(2018)01-0001-07

# 基于社会力异常检测改进算法的人群行为模型

卢丽<sup>1</sup>, 许源平<sup>1</sup>, 卢军<sup>1</sup>, 黄健<sup>1</sup>, 张朝龙<sup>1</sup>, 王晶<sup>2</sup>

(1. 成都信息工程大学软件工程学院, 四川 成都 610225; 2. 英国哈德斯菲尔德大学, 西约克郡, HD1 3DH)

**摘要:** 社会力异常检测算法(SAFM)是检测人群异常行为(人群聚集和恐慌逃散等)的一种核心算法,提取的底层特征不能完整地描述人群的运动状态,导致人群异常行为的识别率低。为此,提出一种改进的社会力异常检测算法(SFDE)解决此问题。算法引入人群运动的轨迹避免底层特征的丢失,通过无监督方法将轨迹进行聚类,再通过轨迹和人群相互作用力建立人群行为模型(人群运动强度、人群方向熵和聚簇中心距离势能的作用力)。为证明算法有效性,应用改进的SFDE算法结合深度学习模型来识别不同的人群异常行为。通过UMN数据对算法进行验证,结果表明SFDE算法的准确率比传统的SAFM算法提高了18%,并且执行时间提高了2.2s。

**关键词:** 社会力; 人群异常检测; 无监督算法; 轨迹聚类; 群体行为; 深度学习; 视频监控

中图分类号: TP317.4

文献标志码: A

doi: 10.16836/j.cnki.jcui.2018.01.001

## 0 引言

随着世界经济的发展,在人民群众生活水平不断提高的同时,也越来越重视自身的安全,某些严重危害社会公共安全的重大事件,给人民群众的生命和财产造成了巨大的损失。因此,政府和相关部门投入了大量的人力和物力,并采取相应的措施防止此类群体突发事件的发生。其中,一项关键措施就是大范围的广场、火车站和人群汇集的街道等公共场所设置视频监控设备(closed circuit television, CCTV)。但是,传统的监控设备智能度不高,需要安保人员不间断地监控视频,因此需要投入大量人力资源。随着监控的规模的增大,安保人员的工作负担也急剧增加,极易造成群体突发事件(异常事件)的漏检和相应报警的延误。针对这一问题,如何设计一套智能、高效的视频群体异常行为识别算法,通过该算法实时、自动地分析视频中的场景信息,全方位地掌握人群的运动状态,并将异常的人群行为及时预警,成为安防领域中一个越来越重要的研究课题。

近几年人群行为检测有两种方法,一种是“基于对象”的分析方法,此方法将人群看作是多个个体的集合<sup>[1-2]</sup>。但在复杂的场景下,人与人之间相互遮挡,很难获取完整的个人信息,并且场景中每个人的运动存在着随机和无序性,无法提取有序和规律的运动特征。另一种是“基于整体”的分析方法,此方法将人群

当作一个整体研究,常用于中、高密度的结构化场景<sup>[3]</sup>。许多研究者常用此方法研究群体对象,例如Solmaz等<sup>[4]</sup>使用光流和线性近似算法对人群行为进行分类。Zhang Y等<sup>[5]</sup>使用人群之间的社会力属性判定人群的运动行为。黄鲜萍等<sup>[6-7]</sup>使用相关主题模型(CTM)提取人群行为特征,解决了相同场景下人群行为之间的相关性。但是这些方法对噪声敏感,容易丢失一些特征。因此Xu D等<sup>[8]</sup>对此进行了改进,提出多尺度光流法获得了更多的上下文信息。陈乔<sup>[9]</sup>提出Split-Merge分组技术,将视频中的人群划分成多个小群体,并在每个小群体的行为特征,然后使用贪婪算法对人群行为进行分类。虽然这些方法能够减少人群特征的丢失,但在新的场景下,该算法需要重新训练和测试,不仅增加了算法的复杂度,而且需要耗费大量时间进行预处理。为此,Fradi H等<sup>[10]</sup>引入了表现能力更强的跟踪方法,通过追踪人群运动的轨迹检测人群异常行为,但此方法没有考虑人群的运动强度。为了解决此问题,Zhou B等<sup>[11]</sup>将整个人群当作Agents,通过光流来跟踪人群的时空兴趣点检测异常行为,此方法达到了较好的效果。

综上所述,即要考虑算法的准确性,又要考虑算法的合理性,一种可行的方法是“基于整体”的方法研究人群行为,为此,提出一种改进的社会力检测算法(SFDE)提高人群异常行为的识别率。算法首先检测出人群运动的前景区域;然后使用光流法追踪前景区域的时空兴趣点得到人群运动的轨迹;并使用无监督方法对人群运行的轨迹进行聚类;最后通过轨迹和人群相互作用力建立模型。并将人群行为的轨迹图输入

收稿日期: 2017-09-04

基金项目: 国家自然科学基金资助项目(61203172、61202250); 四川省科技厅资助项目(2017JY0011、2014GZ0007); 深圳重大国际合作资助项目(GJHZ20160301164521358)

两层卷积神经网络来识别不同的行为模式<sup>[12]</sup>。实验结果表明 SFDE 算法在识别率和平均精度方面都有提高,具有很好的应用前景。

## 1 人群行为模式的分析

### 1.1 社会力算法简介

社会力是研究人群行为的一种算法。算法将人看作具有一定质量的粒子并结合牛顿第二定律研究人群<sup>[13]</sup>。社会力的计算公式如下:

$$\begin{cases} m_b \frac{du_b}{dt} = Fa = Fp + F_{int} \\ Fp = \frac{1}{\tau}(u_b^p - u_b) \end{cases} \quad (1)$$

社会力受 3 个作用力。(1) 行人  $b$  以一定的质量  $m_b$ 、时间  $t$  和速度  $u_b$  运动所受到的力  $F_a$ 。(2) 行人自身的驱动力  $F_p$ 。(3) 外界环境对人的影响力以及行人之间的相互作用力  $F_{int}$ 。 $\tau$  表示松弛系数。 $u_b^p$  和  $u_b$  分别表示理想速度和实际速度。

通常出现异常人群行为时,行人的理想速度用  $u_b^q$  代替  $u_b^p$ ,公式如下:

$$\begin{aligned} u_b^q &= (1 - p_b) u_b^p + p_b u_b^m \\ m_b \frac{du_b}{dt} &= Fa = \frac{1}{\tau}(u_b^q - u_b) + F_{int} \end{aligned} \quad (2)$$

当权重参数  $p_b$  接近 1 时,行人的速度  $u_b^q$  等于周围人的平均速度  $u_b^m$ 。当  $p_b$  接近 0 时,行人的速度  $u_b^q$  等于之前的速度  $u_b^p$ 。

在异常情景下,假设行人的质量为 1,行人之间的相互作用力为

$$F_{int} = F_a = m_b \frac{du_b}{dt} - \frac{1}{\tau}(u_b^q - u_b) \quad (3)$$

### 1.2 社会力异常检测算法(SAFM)

SAFM 算法通过行人间的交互力  $F_{int}$ 、场景尺度  $W_{ij}$ 、冲突性  $Con_{ij}$  和拥挤性  $S_{ij}$  3 个社会属性判定人群的运动状态。 $F_{social}$  大于阈值  $T_{social}$  时为正常状态,否则为异常状态,异常检测算法计算公式为

$$F_{social} = W_{ij}(C_{ij} + S_{ij}) F_{int} \quad (4)$$

#### 1.2.1 行人间的冲突性

图像像素点的光流方向离散化到圆周  $[0^\circ, 360^\circ]$ ,并划分为 8 个方向角( $j$ )。冲突性  $Con_{ij}$  表示人群的混乱程度,人群的速度越大混乱程度越大。冲突性表达式为

$$\begin{cases} \varphi_{ij} = Hist_{ij}\{On\}, n \in 1 \cdots 8 \\ A_{ij} = \text{sgn}(\text{std}(\varphi_{ij}) - \text{std}(\varphi_r)) \\ Con_{ij} = A_{ij} \exp(\text{std}(\varphi_{ij}) - \text{std}(\varphi_r)) \end{cases} \quad (5)$$

其中  $\varphi_{ij}$  是网格块的方向直方图。 $\varphi_r$  是方向直方图的一个阈值。 $\text{std}$  代表直方图的标准差。 $A_{ij}$  表示符号函数,  $\text{std} > 0$  时值为 1,  $\text{std} < 0$  时值为 0。

#### 1.2.2 行人间的拥挤性

行人间的拥挤性  $S_{ij}$  表示人群的拥挤程度,人群的摩擦系数  $k_{ij}$  越高,拥挤程度越大。其中  $dens_{ij}$  是网格块的密度,  $\theta_r$  是人群密度的一个阈值。拥挤性表达式为

$$\begin{cases} S_{ij} = \text{std}(Hist(V_{ij})) \\ B_{ij} = \text{sgn}(dens_{ij} - dens_r) \\ C_{ij} = K_{ij} B_{ij} (dens_{ij} - dens_r) \end{cases} \quad (6)$$

### 1.3 改进的社会力异常检测算法(social force detection-SFDE)

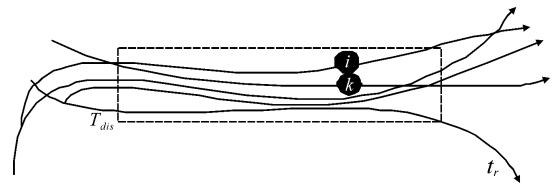


图1 人群运动轨迹聚类

基于之前的分析,在社会力异常检测算法的基础上,针对没有完整提取人群特征问题,下面进一步改进社会力异常检测算法。首先通过 Vibe 算法获得人群的前景区域<sup>[14]</sup>。然后使用高精度光流法追踪前景区域的时空兴趣点得到人群运动的轨迹<sup>[15]</sup>。如图 1 所示,第  $z$  帧人群运动的轨迹  $tr = \{(x_1, y_1, v_1) \cdots (x_n, y_n, v_n)\}$ 。特征点  $l$  和  $k$  之间的距离满足公式(7),并且两个特征点都在同一个方向角时,表明两个特征点能够聚集在同一个聚簇中。特征点  $l$  和  $k$  之间的距离 ( $D_{trlrk}$ ) 公式为

$$\begin{cases} D_{trlrk} = \sqrt{(tr_{kx} - tr_{lx})^2 + (tr_{ky} - tr_{ly})^2} \\ C_{lk} = w_1 \times \frac{f_{num} - D_{trlrk}}{f_{num}} + w_2 \times \frac{15 - |W_l - W_k|}{15} \end{cases} \quad (7)$$

$C_{lk}$  的值小于阈值 ( $T_{dis} = 0.9$ ) 时,表明  $l$  和  $k$  点能聚集在同一个聚簇中。 $W_l$  和  $W_k$  表示  $l$  和  $k$  点和方向角。 $w_1$  (0.8)、 $w_2$  (0.2) 分别是距离和方向的权重。 $f_{num}$  表示视频帧的长度(光流一次能追踪视频的帧数  $f_{num} = 15$ )。

人群轨迹的相邻矩阵  $B = \{B_1, B_2, \cdots, B_n\}$ ,  $B_i = \{tr_1, tr_2, \cdots, tr_n\}$ ,  $B_i$  的聚类中心  $C = (C_1, C_2, \cdots, C_m)$ 。

聚簇中心  $C_i(C_{ix}, C_{iy})$  计算公式为

$$C_{ix} = \frac{\sum_{j=1}^n tr_{jx}}{n}, C_{iy} = \frac{\sum_{j=1}^n tr_{jy}}{n} \quad (8)$$

无监督聚类轨迹算法流程如下:

输入: 聚簇的阈值 ( $T_{dis}$ ), 第  $z$  帧人群运动的轨迹

$tr_z = \{(x_1, y_1, v_1) \cdots (x_n, y_n, v_n)\}$

(1) 给矩阵  $S$  赋值, 标记相邻矩阵是否已聚类

for  $i=1$  to  $N-1$  do //  $S$  矩阵赋值

for  $j=i+1$  to  $N$  do

if ( $i=j$ )  $S[i][j]=0$  // 对角赋值为 0

else if ( $D_{i,j} \leq T_{dis}$ )  $S[i][j]=1$  // 未聚类

else if ( $D_{i,j} > T_{dis}$ )  $S[i][j]=-1$  // 已聚类

end for

end for

(2) 将每一帧视频进行聚类, 得到多个聚簇

while (矩阵  $S$  中不是所有元素都已聚类)

if (每行已全部聚类)

该行为一个聚簇, 更新矩阵  $S$

else

for  $i=0$  to  $N-1$  do

在一行中未聚类的最大个数, 更新矩阵  $S$

end for

end if

合并所有聚簇, 计算聚簇中心

end while

return  $S$

输出: 聚类中心  $C=(C_1, C_2, \cdots, C_m)$

## 1.4 改进算法的行为模式判定

### 1.4.1 人群的运动强度

第  $z+1$  帧特征点集  $ntr=\{(nx_1, ny_1, nv_1) \cdots (nx_n, ny_n, nv_n)\}$ ,  $n$  表示像素点总数。特征点  $l(x_l, y_l, v_l)$  在第  $z$  帧的光流速度  $v_l$  和方向  $\theta_l$  计算公式如下:

$$\begin{cases} v_l = \sqrt{(nx_l - x_l)^2 + (ny_l - y_l)^2} \\ \theta_l = \arctan\left(\frac{|ny_l - y_l|}{|nx_l - x_l|}\right) \end{cases} \quad (9)$$

第  $z$  帧人群运动的聚簇中心的运动强度  $E$  表示为

$$E = \frac{1}{2} m \sum_{l=1}^n (v_l)^2 \quad (m=1) \quad (10)$$

### 1.4.2 人群方向熵

特征点的光流对应方向角的值总数  $n_j$ , 例如特征点在  $(0 \sim 45^\circ)$ , 则  $n_1$  加 1。第  $z$  帧人群运动的方向熵

( $H$ ) 为

$$\begin{cases} p_j = \frac{n_j}{\sum_{k=1}^n n_k} \\ H = - \sum_{k=1}^n p_j \log_2 p_j \end{cases} \quad (11)$$

### 1.4.3 聚簇中心的作用力

第  $z$  帧聚簇中心的距离势能作用力  $F$ 。  $n$  是聚簇内的特征数,  $m$  是聚簇的个数。  $f_x, f_y$  分别是聚簇内的特征点到聚簇中心  $C_{xi}$  的横坐标和纵坐标的距离, 异常检测算法计算公式如下:

$$\begin{cases} f_x = \sum_{i=1}^m \sum_{j=1}^n (C_{xi} - tr_{xj})^2 \\ f_y = \sum_{i=1}^m \sum_{j=1}^n (C_{yi} - tr_{yj})^2 \end{cases} \quad (12)$$

$$F = \sqrt{f_x + f_y} \quad (13)$$

### 1.4.4 人群的相互作用力

人群运动强度、人群方向熵和聚簇中心的距离势能作用力组成第  $z$  帧人群的相互作用力  $F_{new}$ 。视频原始图像  $W_s$  作为卷积神经网络 CNN 的上层网络的输入, 特征向量  $F, E, H$  经过 CNN 学习后得到人群行为的分类。人群的相互作用力表达式为

$$F_{new} = F + E + H + W_s \quad (14)$$

### 1.4.5 人群行为模型

第  $z$  帧人群运动的轨迹组成的向量  $r = \{\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_i, \cdots, \vec{r}_n\}$ , 其中  $\vec{r}_i = (x_{i+1} - x_i, y_{i+1} - y_i)$ , 相邻两个向量之间的夹角 ( $\Phi$ ) 表示为

$$\Phi_{i,i+1} = \arccos \left[ \frac{\vec{r}_i \cdot \vec{r}_{i+1}}{|\vec{r}_i| |\vec{r}_{i+1}|} \right] \quad (15)$$

向量  $\vec{r}_{i,i+1}$  与向量  $\vec{r}_{i+1,i+2}$  之间的向量积为

$$U_{i,i+1} = (x_{i+1} - x_i)(y_{i+2} - y_{i+1}) - (x_{i+2} - x_{i+1})(y_{i+1} - y_i) \quad (16)$$

轨迹上相邻向量的夹角和 ( $\psi$ ) 为

$$\psi = \sum_{i=n}^{n-2} \theta_{i,i+1} \cdot (-1)^{\text{sign}(U_{i,i+1})} \quad (17)$$

人群异常的判定条件定义为

$$\begin{cases} (C \geq 2) \& (\psi < T_1) \& \\ (|\Phi_{s(i,i+1)} - \Phi_{e(i,i+1)}| < T_2) \& \Phi_{s(i,i+1)} \max \in [\pi/2, \pi] \end{cases} \quad (18)$$

$C$  是类簇的个数,  $T_1$  和  $T_2$  根据 UMN 数据库第 491(恐慌逃散) 和 1456(聚集斗殴) 帧图像用式 (17) 和式 (15) 算出, 得到  $T_1 = 30, T_2 = 80$ 。  $T_3$  和  $T_4$  的值根据 CNN 通过式 (14) 自动调节参数训练图像。

表 1 人群异常事件的判定条件

事件类型	判定条件( $T_3$ 、 $T_4$ 根据 CNN 自动调节参数)
恐慌逃散	满足式( 18) 且人群相互作用力( $F_{New}>T_4$ )
聚集斗殴	满足式( 18) 且人群的相互作用力( $T_3<F_{New}<T_4$ )

## 2 实验及分析

### 2.1 数据库

实验的软硬件环境为 Windows10 64 位的操作系统,并使用 VC++和 OpenCV 作为软件的开发平台。实验中使用明尼苏达州大学( University of Minnesota) 创建的一个 UMN 数据库,该数据集通过摄像机拍摄人群活动得到 7739 帧视频图像。其中,正常情况是场景下人只有行走。异常情况主要是行人恐慌逃散和聚集斗殴。是卷积神经网络的训练样本,其中,图 2( a)、图 2( b) 分别显示第 20、1040 帧人群正常行走的图像,图 2( c) 显示第 550 帧人群“恐慌逃散”的图像,图 2( d) 显示第 1500 帧人群“聚集斗殴”的图像。

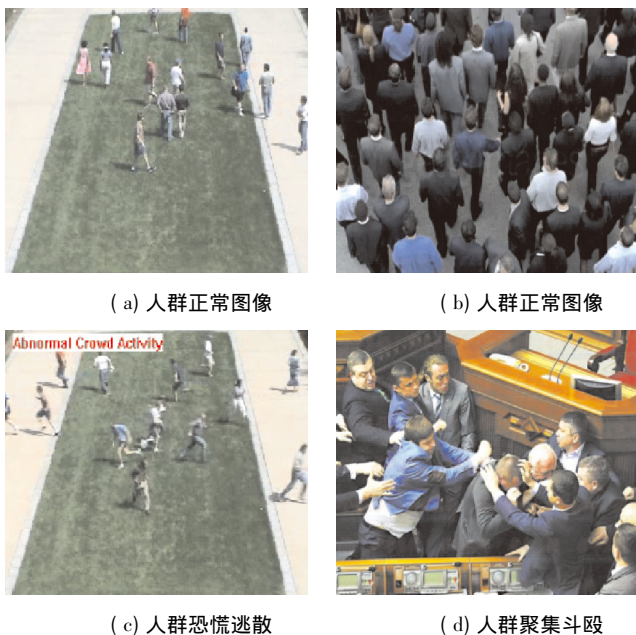


图 2 卷积神经网络训练的 UMN 训练样本

### 2.2 算法整体框架

图 3 是人群异常行为识别的整体框架,主要流程的如下:

数据集: 数据集包括 7739 帧 UMN 人群活动图。

活动类型判断: 每帧视频图片通过式( 14) 计算得到的特征向量,并通过标记人群运动的模式,例如广场步行( 1)、恐慌逃散( 2) 和聚集斗殴( 3)。

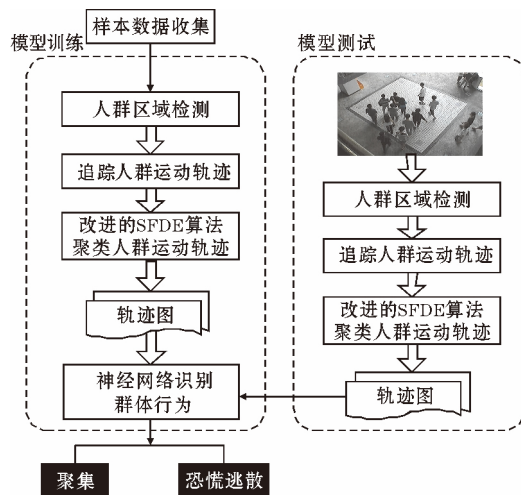


图 3 人群异常行为识别的整体框架

轨迹图: 将每帧特征向量转化为相应的轨迹图。

CNN 识别人群异常行为: 视频原始场景图像和被标记的轨迹图像分别输入卷积神经网络( Convolutional Neural Networks, CNN) 训练和测试,实验中通过 fine-tune 微调 CNN 得到不同场景下的群体行为。

在训练阶段: 被标记的 3000 帧轨迹图( 正常和异常) 输入到下层 CNN 中,并设置相应的参数( 学习率 0.001,卷积核大小和步长) 自动优化网络结构。卷积神经网络结构如图 4 所示,卷积神经网络共有两层神经网络结构,每层网络有 1 个数据层、5 个卷积层( Conv)、3 个池化层( Pool)、2 个归一化层( Norm)、3 个全连接层( FC)。上层 CNN 模型的初始化使用可使用一个预先训练的模型参数识别不同的场景<sup>[16]</sup>。下层网络分支初始化使用以下参数训练卷积神经网络, Conv( 96, 7, 2) -ReLU-Pool( 3, 2) -Norm( 5) -Conv( 256, 5, 2) -ReLU-Pool( 3, 2) -Norm( 5) -Conv( 384, 3, 1) -ReLU-Conv( 384, 3, 1) -ReLU-Conv( 256, 3, 1) -ReLU-Pool( 3, 2) -FC( 3)。

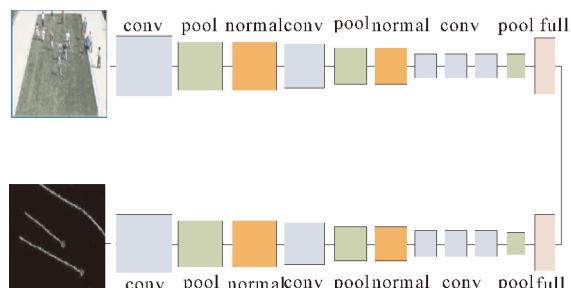


图 4 卷积神经网络提取特征

数据层。统一调整大小为  $224 \times 224 \times 3$  的图像。并计算所有训练图片的均值( mean)。在训练和测试阶段,分别将图像( M) 三通道( RGB) 减去均值文件的值。

$$M = RGB - mean \tag{19}$$

卷积层。CNN 网络结构的神经元使用卷积核从局部到整体提取不同的特征(例如颜色、纹理、轮廓)<sup>[17-18]</sup>。当前层第  $j$  个神经元与前一层神经元的权重( $W_{ij}$ )再加上偏置项,并使用 RELU 函数通过非线性映射生成第  $j$  个特征图 $c_j$ 。卷积操作的公式为:

$$c_j = \sigma \left( \sum_{i \in M_j} W_{ij} \times x_i + b_j \right) \tag{20}$$

池化层。通过卷积层学习后的特征输入采样层降低特征的维度,这些特征在窗口区域由于 max-pooling 函数选择窗口的最大值作为特征的输出容易造成过拟合现象,而 mean-pooling 函数选择窗口的平均值作为特征的输出容易造成欠拟合现象。为提高模型的识别率,文中采用自适应阈值的方法避免过拟合和欠拟合现象,其基本思想是在窗口区域计算一个适合当前区域的阈值,将小于阈值的值删除,大于阈值的值保留并求平均值 $u_{ij}$ 。激活值 $a_{ij}$ 通过池化区域 $y_{m,n}$ 得到激活值的概率 $p_{ij}$ ,并通过加权统计池化区域的得到该区域的阈值 $e_{m,n}$ <sup>[19]</sup>。计算公式如下:

$$\begin{cases} u_{ij} = \frac{\sum_{i,j \in y_{m,n}} \{a_{ij} - e_{m,n} \geq 0? 1:0\} \times a_{ij}}{\sum_{i,j \in y_{m,n}} \{a_{ij} - e_{m,n} \geq 0? 1:0\}} \\ e_{m,n} = \sum_{i,j \in y_{m,n}} p_{ij} \times a_{ij} \\ p_{ij} = \frac{a_{ij}}{\sum_{k,j \in y_{m,n}} a_{kj}} \end{cases} \tag{21}$$

全连接层。下采样的全部特征输入全连接层。将所有特征直接输入 softmax 分类器中对不同的人群行为(或场景)进行分类。 $N$  表示 3 个属性, $t_n$  表示所属的类别的目标标签(范围为 1~3), $o_n$  表示概率预测值。整个 CNN 网络的损失函数表示为:

$$E = -\frac{1}{N} \sum_{n=1}^N t_n \log o_n + (1 - t_n) \log(1 - o_n) \tag{22}$$

测试阶段:通过训练阶段得到的 CNN 模型测试 4739 帧视频图像,得到群体行为的分类信息(广场步行(1)、恐慌逃散(2)和聚集斗殴(3))。

2.3 实验结果

视频帧图像利用图 4 的 CNN 神经网络结构识别不同的人群的行为活动。图 5 是 UMN 人群活动的轨迹图,其中,图 5(a)和图 5(b)是通过 CNN 识别为人群正常行为轨迹图。图 5(c)是通过 CNN 识别为人群恐慌逃散的异常轨迹图。图 5(d)是通过 CNN 识别为

人群聚集斗殴的异常轨迹图。

2.4 实验对比

为验证深度学习(CNN)的识别率高于传统机器学习的方法(支持向量机(SVM)和随机森林(RF))。实验中通过测试 UMN 数据库比较分类的精确度,训练时间复杂度和测试时间复杂度。对比结果如表 2 所示,CNN 比机器学习(SVM 和 RF)的分类精度高 18.9%。在时间复杂度上,机器学习和 CNN 的时间复杂度(训练和测试)相差不大。

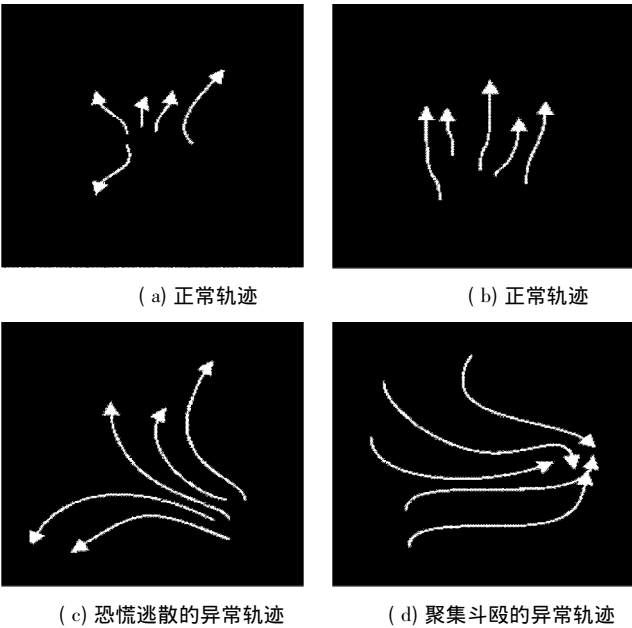


图 5 UMN 人群活动的轨迹图

表 2 CNN 与机器学习模型在分类精度上的比较

模型	卷积神经网络(CNN)	支持向量机(SVM)	随机森林(RF)
分类准确度/%	88.9	69.8	70.1
训练时间/min	80	73	78
测试时间/ms	105	98	110

为验证算法的有效性,文中的 SFDE 与 SAFM<sup>[4]</sup>, HOFO<sup>[14]</sup>和 Optical Flow<sup>[20]</sup>进行对比,算法分别使用 UMN 数据库在相同的实验条件下进行对比。由表 3 算法的平均精确度可知,文中的 SFDE 算法得到人群行为的平均精确度比 SAFM 高 18%。另一方面,HOFO 算法得到人群行为的平均精度为 0.74。Optical Flow 算法得到人群行为的平均精度为 0.66。这 3 种算法和文中算法都使用光流进行计算,但这些方法没有考虑到人群密集时人与人之间的相互遮挡影响人群行为的跟踪,从而丢失了某些底层特征。而本文通过轨迹聚类成多个小群体来研究人群行为,因此文中的算法比这三种算法结果更精确。

表3 算法的准确度比较

	SFDE 算法 / %	SAFM	HOFO	Optical Flow
广场步行	0.95	0.90	0.87	0.88
恐慌逃散	0.9	0.62	0.62	0.77
聚集斗殴	0.83	0.6	0.73	0.32
平均精度	0.89	0.71	0.74	0.66

表4显示SFDE算法与其他算法之间的比较。由表4可知,文中的AUC面积优于其他经典算法光流、社会力和HOFO。并且文中的执行时间最优,比SAFM提高了2.2 s。图6是所有算法的ROC曲线,由图可知SFDE算法的AUC的面积最大优于其他传统方法,即文中方案具有较高的准确率和较高的计算速度。

表4 SFDE算法与其他方法的时间比较

方法	AUC	耗时/s
Optical Flow	0.84	35
SAFM	0.945	23.5
HOFO	0.94	22
SFDE	0.97	21.3

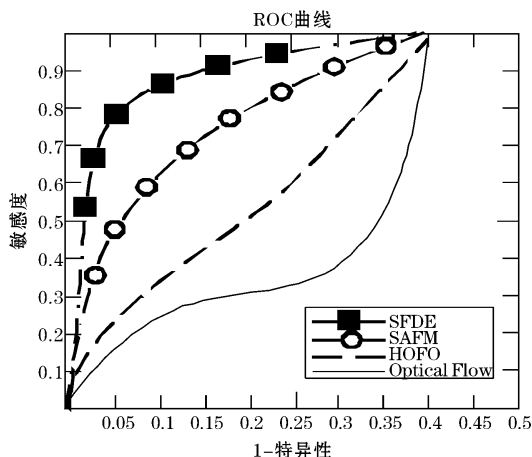


图6 UMN数据库的ROC曲线

### 3 结束语

提出一种改进的社会力检测算法(SFDE)来提高人群异常行为的识别率。算法充分利用人群运动的轨迹提取人群运动特征,并通过聚类的轨迹分析人群行为模型,进而将人群行为模型结合深度学习模型来识别不同场景下各种异常行为。实验使用UMN数据进行验证。结果表明,新算法的准确率比SAFM算法提高了18%,并且执行时间提高了2.2 s。由于实验人群行为还未关联语义,今后的工作重点是人群行为的时空轨迹特征与高级语义特征进行关联。

### 参考文献:

- [1] Junior J C S J, Musse S R, Jung C R. Crowd Analysis Using Computer Vision Techniques [J]. IEEE Signal Processing Magazine, 2010, 27(5): 66-77.
- [2] Tu P, Sebastian T, Doretto G, et al. Unified Crowd Segmentation [J]. Lecture Notes in Computer Science, 2008, 53(5): 691-704.
- [3] Li T, Chang H, Wang M, et al. Crowded Scene Analysis: A Survey [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(3): 367-386.
- [4] Solmaz B, Moore B E, Shah M. Identifying Behaviors in Crowd Scenes Using Stability Analysis for Dynamical Systems [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(10): 2064-2070.
- [5] Zhang Y, Qin L, Yao H, et al. Abnormal crowd behavior detection based on social attribute-aware force model [J]. 2012: 2689-2692.
- [6] 黄鲜萍. 人群运动主题语义特征提取和行为分析研究[D]. 杭州: 浙江工业大学, 2015.
- [7] 徐戈, 王厚峰. 自然语言处理中主题模型的发展[J]. 计算机学报, 2011, 34(8): 1423-1436.
- [8] Xu D, Song R, Wu X, et al. Video anomaly detection based on a hierarchical activity discovery within spatio-temporal contexts [J]. Neurocomputing, 2014, 143(16): 144-152.
- [9] 陈乔. 多人交互行为的分组检测及语义特征提取[D]. 南京: 南京邮电大学, 2016.
- [10] Fradi H, Dugelay J L. Towards crowd density-aware video surveillance applications [J]. Information Fusion, 2015, 24(C): 3-15.
- [11] Zhou B, Tang X, Wang X. Learning Collective Crowd Behaviors with Dynamic Pedestrian-Agents [M]. Kluwer Academic Publishers, 2015: 50-68.
- [12] Vapnik V, Lerner A. Pattern recognition using generalized portrait method [J]. Automation & Remote Control, 2008, 24(24): 774-780.
- [13] Mehran R, Oyama A, Shah M. Abnormal crowd behavior detection using social force model [C]. Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009: 935-

- 942.
- [14] Barnich O ,Van D M. ViBe: a universal background subtraction algorithm for video sequences [J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society 2011 20( 6) : 1709.
- [15] Brox T ,Bruhn A ,Papenberg N ,et al. High Accuracy Optical Flow Estimation Based on a Theory for Warping [J]. 2004 3024( 10) : 25–36.
- [16] Ouyang W ,Luo P ,Zeng X ,et al. DeepID-Net: multi-stage and deformable deep convolutional neural networks for object detection [J]. Eprint Arxiv 2014.
- [17] Huang J ,Kumar S R ,Mitra M ,et al. Image Indexing Using Color Correlograms [C]. Conference on Computer Vision and Pattern Recognition ,1997: 762.
- [18] Dalal N ,Triggs B. Histograms of Oriented Gradients for Human Detection [C]. Computer Vision and Pattern Recognition 2005. CVPR 2005. IEEE Computer Society Conference on 2005: 886–893.
- [19] 詹智财. 基于卷积神经网络的视频语义概念分析 [D]. 镇江: 江苏大学 2016.
- [20] Mehran R ,Oyama A ,Shah M. Abnormal crowd behavior detection using social force model [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2009: 935–942.

## A Crowd Behavior Model based on an Improved Social Force Anomaly Detection Algorithm

LU Li<sup>1</sup> , XU Yuan-ping<sup>1</sup> , LU Jun<sup>1</sup> , HUANG Jian<sup>1</sup> , ZHANG Chao-long<sup>1</sup> , WANG Jing<sup>2</sup>

( 1.College of Software Engineering ,Chengdu University of Information Technology ,Chengdu 610225 ,China; School of Computing & Engineering , University of Hudderseld , Queensgate , Huddersheld , UK)

**Abstract:** Social Force anomaly detection algorithm ( SAFM) is a core algorithm for detecting abnormal crowd behaviors ( e.g. ,crowd aggregations and panic escapes ,etc.) .Some low-level features of the algorithm can't fully describe the movement states of the crowd ,so the classification recognition rate is very low. Thus ,an improved social force anomaly detection algorithm ( SFDE) is proposed in this research to solve this problem. This algorithm introduces the trajectories to avoid the loss of low-level features , and it also groups trajectories into clusters by applying an unsupervised algorithm. Finally a model of crowd behavior can be established through combination of trajectories and multiple crowd forces ( e.g. ,the kinetic energy of the crowd ,the entropy of motion direction and the force of cluster centers) .To test the validity and effectiveness of the proposed algorithm ,this paper presents how to apply SFDE together with the deep learning model to recognize various crowd behaviors. The SFDE has been tested and evaluated by using the UMN dataset. Experimental results show that the accuracy of SFDE is 18 % higher than SAFM , and the execution time is decreased by 2.2 s.

**Keywords:** social force; crowd anomaly detection; unsupervised algorithm; trajectory clustering; crowd behavior; deep learning; video monitoring