



计算机工程与应用  
Computer Engineering and Applications  
ISSN 1002-8331,CN 11-2127/TP

## 《计算机工程与应用》网络首发论文

题目: 基于运动筛选和 3D 卷积的视频早期烟雾检测  
作者: 高联欣, 魏维, 胡泳植, 冯宇浩  
网络首发日期: 2019-09-20  
引用格式: 高联欣, 魏维, 胡泳植, 冯宇浩. 基于运动筛选和 3D 卷积的视频早期烟雾检测. 计算机工程与应用.  
<http://kns.cnki.net/kcms/detail/11.2127.TP.20190920.0904.002.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于运动筛选和 3D 卷积的视频早期烟雾检测

高联欣, 魏 维, 胡泳植, 冯宇浩

成都信息工程大学 计算机学院, 成都 610225

**摘 要:** 针对基于视频的早期烟雾检测在复杂环境下的高误报和高漏检问题, 提出了一种基于运动筛选疑似区域的方法并设计了一个输入为 6 帧图片的多尺度 3D 卷积神经网络(6M3DC)来进行视频烟雾检测的算法。首先将视频帧进行均值滤波后通过背景差分模型获取运动区域并计算获得区域所在块, 然后通过颜色判断和均值 HASH 算法对运动块进行筛选提取疑似块并将不符合条件的块更新到背景图。最后通过组合连续 6 帧相同区域的疑似块输入多尺度 3D 卷积神经网络进行检测, 将检测为烟雾的块标记, 非烟雾块更新到背景图。实验结果表明, 算法对缓慢运动的烟雾有一定的适应性, 可以较好的在复杂环境下检测出烟雾。

**关键词:** 早期烟雾; 颜色判断; 均值 HASH; 多尺度 3D 卷积

**文献标志码:** A **中图分类号:** TP391.4 **doi:** 10.3778/j.issn.1002-8331.1906-0234

高联欣, 魏维, 胡泳植, 等. 基于运动筛选和 3D 卷积的视频早期烟雾检测. 计算机工程与应用

GAO Lianxin, WEI Wei, HU Yongzhi, et al. Video early smoke detection based on motion extraction and 3D convolution neural network. Computer Engineering and Applications

## Video Early Smoke Detection Based on Motion Extraction and 3D Convolution Neural Network

GAO Lianxin, WEI Wei, HU Yongzhi, FENG Yuhao

College of Computer Sciences, Chengdu University of information Technology, Chengdu 610225, China

**Abstract:** To solve the problem of high false alarm and high missed detection in the complex environment of early smoke detection based on video, a method based on motion extraction of suspected areas is proposed and a multi-scale 3D convolutional neural network with input of 6 frames(6M3DC) is designed for video smoke detection..Firstly, the motion regions are obtained through the background difference model after average filtering and the positions of the block in which the motion regions are located are recalculated, and then the motion blocks are extracted by color judgment and mean HASH algorithm and update the nonconforming blocks to the background image. Finally, by combining the suspected blocks of the same region of 6 consecutive frames as the input for the 3D convolutional neural network for detection, blocks detected as smoke are marked and non-smoke blocks are updated to the background image. The experimental results show that the algorithm is adaptive to slow moving smoke and can detect smoke in complex environment.

**作者简介:** 高联欣(1994-),男,硕士,研究领域为图形图像处理, E-mail: 503773453@qq.com; 魏维(1976-),男,博士,教授,研究领域为图形图像处理; 胡泳植(1998-),男,硕士,研究领域为图形图像处理; 冯宇浩(1996-),男,硕士,研究领域为图形图像处理。

**Key words:** early smoke; color judgment; mean HASH algorithm; multi-scale 3D convolutional neural network

## 1 引言

森林是地球生态系统的主体,在调节气候、涵养水源、防风固沙、改善土壤等方面起着重要的作用。而森林火灾则是突发性强、破坏性大、难以控制的自然灾害,一旦发生森林大火,就会造成大量的自然资源和人类财产的损失。在火灾发生前期会产生偏白色的烟雾,它与云、水雾等自然景象有很大的相似性,检测难度大,但是若能在早期检测出这一显著的视觉特征,就能提早预警,将火灾扼杀在摇篮,从而将损失降到最低。

随着计算机视觉技术的发展,相比于传统使用烟雾报警器、瞭望台、人工巡逻的方式来发现火灾烟雾,基于视频监控的方法更加的高效和节约成本。Chen 等人<sup>[1]</sup>通过实验发现烟雾的 R,G,B 三个通道的值非常接近,灰度值的范围在 80 到 220 之间,并提出了基于帧差法的运动分割后进行像素级分类烟雾的方法。Tian 等人<sup>[2]</sup>将视频每一帧分块后,利用大气散射模型推导出烟雾图像的形成模型,然后提出一种双重完备字典的方法,将烟雾检测转为了凸优化求解的过程。Ma 等人<sup>[3]</sup>对视频序列进行线性增强后,利用低秩矩阵恢复的方法获取烟雾的显著性,然后使用块匹配算法消除干扰后输入支持向量机进行烟雾识别。Russo 等人<sup>[4]</sup>利用背景减法从视频帧中提取前景,然后基于形状获取感兴趣区域并计算局部二值模式的值和直方图,形成特征向量,输入 Bhattacharyya 系数和支持向量机相结合的分类器中进行烟雾检测。Yuan 等人<sup>[5]</sup>根据烟雾颜色设计了基于模糊逻辑的检测规则,然后利用扩展卡尔曼滤波器对检测规则的输入和输出进行重构。Wang 等人<sup>[6]</sup>通过手工设计烟雾的颜色、模糊、轮廓、主运动方

向等特征输入支持向量机进行检测。Vijayalakshmi<sup>[7]</sup>使用模糊 c 均值的方法进行烟区定位检测。

然而,基于传统计算机视觉的烟雾检测非常依赖人工设计的特征,这可能导致烟雾图像的内在特征被忽略掉。随着深度学习的发展,卷积神经网络在图像识别上获得了巨大的成功。它的最大特点是不需要手工设计健壮的特征,而是从输入数据中自动的进行特征学习。在众多的图像比赛中,如 ImageNet,卷积神经网络获取的特征对图像的识别率要远远高于人工设计的特征。

Hu 等人<sup>[8]</sup>提出了一种基于时空卷积神经网络的烟雾检测方法,将相邻帧的烟雾光流图输入时间流网络获取烟雾运动信息,将单帧烟雾图片输入空间网络获取静态特征,然后融合获取的两部分特征输入支持向量机进行分类,该方法基于相邻帧提取光流,若烟雾运动缓慢则光流效果较差。Luo 等人<sup>[9]</sup>通过背景动态更新和暗通道先验算法检测可疑区域,再将可疑区域输入五层神经网络进行识别。Yuan 等人<sup>[10]</sup>受 GoogleNet Inception 模块的启发,提出了 DMCNN 网络来检测烟雾,相比于 ZF-Net,VGG16,Inception-v3,获取了更好的效果。然而 Luo 和 Yuan 的方法则没有考虑烟雾的时间信息。Yin 等人<sup>[11]</sup>设计了独立的两个网络分别学习烟雾空间特征和运动特征,之后输入 RNN 网络以循环的方式融合来进行检测识别烟雾,该方法网络过多,计算量大。Hu 等人<sup>[12]</sup>将卷积神经网络和循环神经网络相融合,将烟雾和其对应的光流图输入网络进行检测,该方法进行了光流的计算,开销较大。Zeng 等人<sup>[13]</sup>通过对现有目标检测算法 Faster-RCNN,

SSD, R-FCN 的特征提取部分进行修改从而适应烟雾的检测。该方法针对中后期拥有显著颜色特征的烟雾效果较好,但是对早期颜色特征不显著的烟雾效果不好。

针对上述问题,本文提出了基于运动筛选获取疑似烟雾块,并使用单独一个多尺度 3D 卷积神经网络同时获取烟雾的空间信息和时间信息的早期烟雾检测方法。该方法能够较好地检测出早期烟雾,极大降低误报率。

## 2 算法设计与实现

为了能适应复杂环境下的早期烟雾检测,算法流程如图 1 所示。

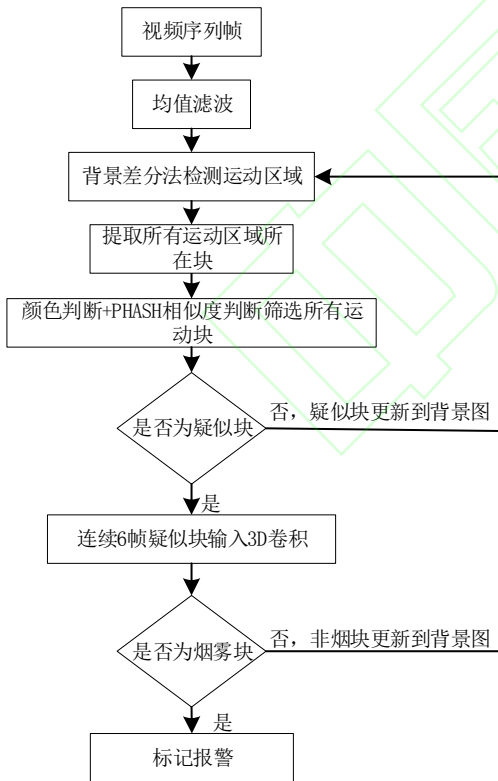


图 1 基于运动筛选和多尺度 3D 卷积的视频早期烟雾检测算法流程图

### 2.1 运动筛选

#### 2.1.1 视频序列预处理

算法首先将视频序列标准化为  $320 \times 240$  的尺寸,然后对视频帧进行了均值滤波的操作来减轻视频帧中的噪声对后续背景差分法的影响。均值滤波算法公式如公式 1 所示。

$$f(x,y) = \frac{1}{mn} \sum_{(x,y) \in S_{xy}} g(c,r) \quad (1)$$

其中  $g(c,r)$  表示原始图像,  $f(x,y)$  表示均值滤波后得到的图像,并令  $S_{xy}$  表示中心点在  $(x,y)$  处,大小为  $m \times n$  的滤波窗口。

#### 2.1.2 背景差分检测运动区域

常用的运动检测方法有帧差法,混合高斯模型和背景差分法三种算法。帧差法适用于具有一定运动速度的物体,由于烟雾可能存在运动缓慢的状态,帧差法将很难捕捉到该条件下烟雾的运动;混合高斯模型需要对背景像素进行建模,容易将缓慢运动的烟雾判为背景;背景差分法则可以通过对视频第一帧或者前几帧进行背景帧的建模,将之后获取到的视频帧与该背景帧进行比较,从而很好地检测到细微的运动。算法将输入视频进行滤波后的首帧作为背景图,之后对其进行动态更新。背景差分算法如公式 2,3 所示。

$$a(x,y) = |f_c(x,y) - f_b(x,y)| \quad (2)$$

其中  $f_c(x,y)$ ,  $f_b(x,y)$  分别代表当前帧和背景帧,  $a(x,y)$  代表当前帧和背景帧对应像素的差值的绝对值的差值图。

$$r(x,y) = \begin{cases} 255, & a(x,y) \geq T \\ 0, & a(x,y) < T \end{cases} \quad (3)$$

由于背景像素会存在微小变化,并且还存在着噪声的影响,故对差值图  $a(x,y)$  进行二值化操作得到  $r(x,y)$ ,将大于等于阈值  $T$  的像素值设置为 255 作为运动像素,小于  $T$  的设置 0 代表背景,在实验当中将  $T$  取值 10 可以得到较好的结果。

背景差分法能够检测到细微的运动但无法随着时间动态的更新背景,在光照变化剧烈或存在较多



运动物体的场景下会检测出大量的干扰目标,为了减少后续算法的运算量,还需要对背景差分所检测出的所有运动块进行筛选,选择烟雾可能性较大的块进行检测,同时将不符合条件的块更新到背景帧,从而实现动态的背景更新。算法将视频帧划分为 $10 \times 10$  共计 100 个区域块,并按照 2.2 节方法计算背景差分获取的运动区域所对应的块。基于以上分析,本文提出了基于 RGB 颜色空间和均值 HASH 算法的两步运动区域筛选的算法。

### 2.1.3 RGB 颜色判断与均值 HASH 算法

通过实验发现,当烟雾慢慢出现,烟雾所在块的 RGB 图像中 B 通道的像素值呈现逐渐上升的趋势。故对运动块取 B 通道的像素值与背景块的 B 通道像素值相减后求和,若大于 0 则送入下一步检测,若小于 0 则将这一个运动块判为背景更新到背景图。

然后采用均值 HASH 算法对基于 RGB 颜色空间筛选后的运动块进行二次筛选,其目的是为了筛选出与背景区分度较大的运动块。

均值 HASH 算法首先将  $32 \times 24$  的运动块缩放到  $8 \times 8$  的大小,以去除图像的高频和细节信息,然后将图像转化为灰度图,计算所有 64 个像素的灰度平均值,之后将图像的每一个灰度值和平均值相比较,大于等于平均值则记作 1,小于平均值则记作 0,这样就得到了 64 位由 0 和 1 组成的能代表这张图象的特征的一串 hash 编码,其流程如图 2 所示。

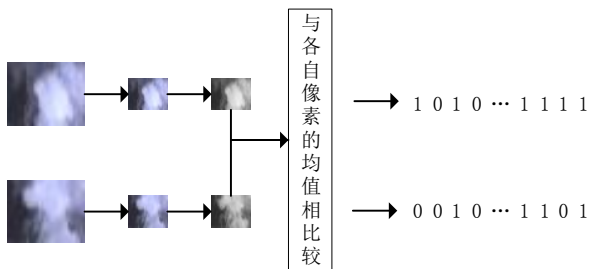


图2 均值 HASH 算法流程图

通过比较背景块和运动块的 64 位 hash 编码的汉明距离是否大于阈值 D 就可以进行相似度的判断,从而对运动区域进行筛选,不符合条件的块更新到背景图,实验中 D 设置为 3。

RGB 颜色判断和均值 HASH 算法组合筛选运动块的效果如图 3 所示。其中红色方框代表颜色判断所筛选的运动区域,黄色横线代表均值 HASH 算法在颜色判断的基础上进一步筛选后的结果。

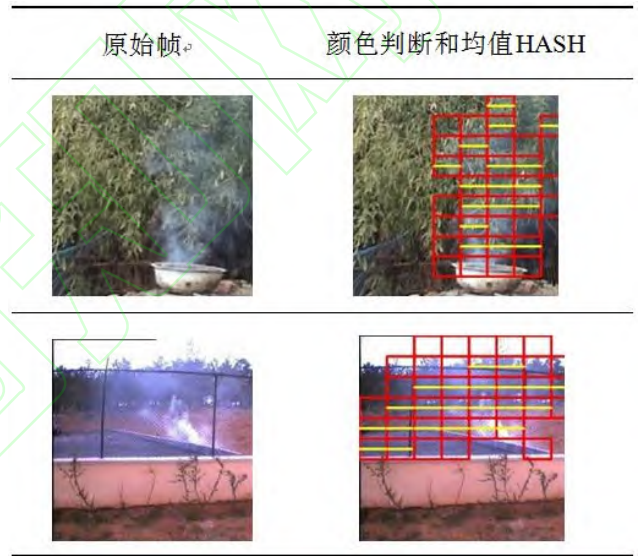


图3 颜色判断和均值 HASH 算法筛选运动区域

实验结果显示,经过提出的两步筛选算法,能够很好的将无关块去掉,从而得到疑似烟雾块。

## 2.2 背景动态更新

本文在 RGB 颜色判断,均值 HASH 算法中对非疑似块以及下文的 3D 卷积中判断为非烟雾块的区域进行背景动态更新的算法如算法 1 所示。

算法 1

1.  $x_l = \lfloor x/32 \rfloor, y_u = \lfloor y/24 \rfloor$

$x_t = (x_l + 1) * 32 - x, y_t = (y_u + 1) * 24 - y$

2. if  $xt \geq w$ :

$$xr = xl + 1$$

else:

$$xr = xl + \lfloor (w-xt)/32 \rfloor + 2$$

if  $yt \geq h$ :

$$yd = yu + 1$$

else:

$$yd = yu + \lfloor (h-yt)/24 \rfloor + 2$$

3. 区域 ( $yutoyd$ ,  $xltoxr$ ) 是否满足算法要求, 不满足则进入 4。

4. for  $i = yu$  to  $yd$  do:

for  $j = xl$  to  $xr$  do:

$$fb(yutoyd, xltoxr) = fc(yu to yd, xl to xr)$$

其中  $x, y, w, h$  为待更新区域的起始  $x, y$  轴坐标与宽  $w$  和高  $h$ , 通过算法获得的  $xl, xr, yu, yd$  为待更新区域所对应块分别在  $x$  轴和  $y$  轴的起始块和结束块。最后判断块是否满足外部算法 (颜色判断/均值 HASH/3D 卷积) 条件, 若不满足则将当前帧所对应的块替换掉背景帧的相应块, 从而实现背景的动态更新。

### 2.3 多尺度 3D 卷积神经网络

3D 卷积神经网络的输入是多帧图片组合成的块, 每一个块中包含了所对应目标的空间特征和该时间段的运动时间特征, 2D 卷积神经网络的输入则只是单帧的图片, 如图 4 所示。不同于 2D 卷积只能提取数据的空间特征, 3D 卷积可以从多帧图片中同时提取数据的时间和空间特征, 例如文献[14]使用 3D 卷积来捕捉视频流的空间和运动信息, 从而实现动作识别。

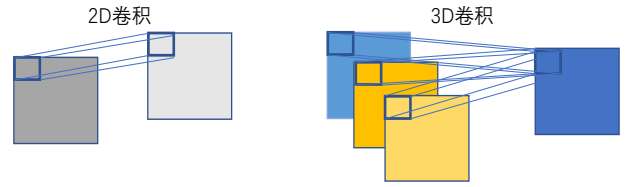


图 4 2D 卷积与 3D 卷积

在视频序列中, 由于烟雾存在运动状态, 有随时间变化的颜色, 纹理和形状等特征, 为了捕捉连续视频帧中的这些特征信息并让特征多样化, 使用 3D 卷积从不同尺度去提取烟雾的特征是非常适合的, 于是设计了如图 5 所示的多尺度 3D 卷积结构。

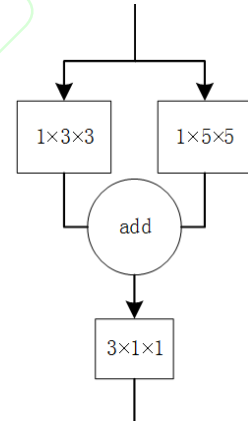


图 5 多尺度 3D 卷积结构

其中  $1 \times 3 \times 3$  和  $1 \times 5 \times 5$  卷积用于提取烟雾图像的不同尺度的空间信息, 之后将其融合在一起, 保持多尺度特征的共存, 最后使用  $3 \times 1 \times 1$  的卷积进行多帧之间的时间采样。最终将这样的 3D 模块进行堆叠, 就得到了多尺度 3D 卷积网络 (6M3DC)。结构图如图 6 所示。

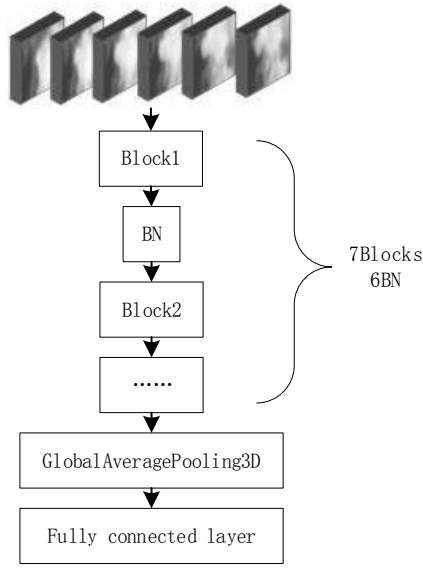


图 6 6M3DC 网络结构图

6M3DC 网络的每一个输入为连续 6 帧图像所组成的块，每一个块里面包含了某一时刻目标对应的运动状态和空间特征，网络的目的是通过学习每一个块中目标空间特征和运动特征，从而获得目标的空间与时间特征。网络的每一个 Block 结构都如图 3 所示，一共 7 个 Block。除了第一个 Block，每一个 Block 前面都接上了 BN 层（Batch Normalization）在提高训练速度的同时可以防止随着网络加深而产生的梯度扩散的问题，最后使用 Global Average Pooling 3D 来增强特征图内的响应并且不会增加网络的参数量，最后接上两个神经元的 Fully Connected Layer 进行分类。若 6M3DC 判断块为非烟雾，则按照 2.2 节方法将块更新到背景图；为烟雾，则进行标记预警。各模型网络参数如表 1 所示。

表 1 各模型网络参数

Model	Layer	Output	Parameters
6M3DC	Block1	$32 \times 24$	Filter:16; Activation: Relu
	Block2	$32 \times 24$	Filter:16; Activation: Relu

算法 1	Block3	$32 \times 24$	Filter:32; Activation: Relu
	Block4	$32 \times 24$	Filter:32; Activation: Relu
	Block5	$32 \times 24$	Filter:64; Activation: Relu
	Block6	$32 \times 24$	Filter:64; Activation: Relu
	Block7	$32 \times 24$	Filter:128; Activation: Relu
	Pooling Layer	128	GlobalAveragePooling3D
	Classification Layer	2	2Dfullyconnectedlayer Activation: softmax
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:16; Activation: Relu
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:16; Activation: Relu
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:32; Activation: Relu
算法 2	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:32; Activation: Relu
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:64; Activation: Relu
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:64; Activation: Relu
	Conv2D ( $3 \times 3$ add $5 \times 5$ )	$32 \times 24$	Filter:128; Activation: Relu
	Pooling Layer	128	GlobalAveragePooling2D
	Classification Layer	2	2Dfullyconnectedlayer Activation: softmax
	Conv3D ( $3 \times 3 \times 3$ )	$32 \times 24$	Filter:16; Activation: Relu
	Conv3D	$32 \times 24$	Filter:16; Activation: Relu

算法 3	(3 × 3 × 3)		
	Conv3D (3 × 3 × 3)	32 × 24	Filter:32; Activation: Relu
	Conv3D (3 × 3 × 3)	32 × 24	Filter:32; Activation: Relu
	Conv3D (3 × 3 × 3)	32 × 24	Filter:64; Activation: Relu
	Conv3D (3 × 3 × 3)	32 × 24	Filter:64; Activation: Relu
	Conv3D (3 × 3 × 3)	32 × 24	Filter:128; Activation: Relu
	Pooling Layer	128	GlobalAveragePooling3D
	Classification Layer	2	2Dfullyconnectedlayer Activation: softmax
	Conv3D (5 × 5 × 5)	32 × 24	Filter:16; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:16; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:32; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:32; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:64; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:64; Activation: Relu
	Conv3D (5 × 5 × 5)	32 × 24	Filter:128; Activation: Relu
	Pooling Layer	128	GlobalAveragePooling3D
	Classification Layer	2	2Dfullyconnectedlayer

			Activation: softmax
--	--	--	---------------------

### 3 实验分析与结论

实验基于 Keras 来构建和训练所提出的 3D 卷积网络，所有实验均在配备 Inter(R) Core(TM) i5-8400 CPU @ 2.80GHz 2.81GHz 和 NVIDIA Tesla M40 GPU 的 PC 上运行的 Win10 和 Ubuntu16.04 操作系统中进行的。

由于烟雾视频没有统一完善的公开数据集，本文的烟雾数据集来自土耳其比尔肯大学，内华达大学和中国科学技术大学的公开烟雾数据库。训练集从部分烟雾视频中进行裁剪，并对其进行水平翻转的数据增强后得到，如图 7 所示。

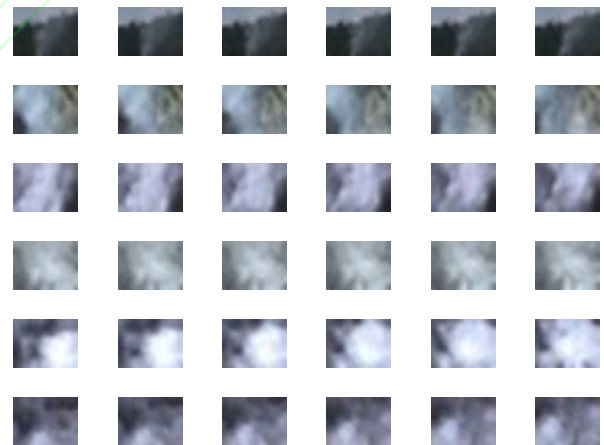


图 7 部分烟雾训练数据集

训练集包含正样本 10296 组，负样本 10068 组的 6×32×24×3 的烟雾块。测试视频如图 8 所示，其中 Video1~Video3 为中近距离快速运动的烟雾，Video4~Video6 为远距离缓慢运动烟雾。

Video1

Video2

Video3





图8 烟雾测试视频

### 3.1 评定标准

视频烟雾的检测主要从精确率，召回率和 F-Measure 三个方面来进行算法的评测，精确率是指预测为正的样本中有多少是对的，包括把烟雾预测为烟雾（TP）和把非烟雾预测为烟雾（FP）；召回率是指样本中的正例有多少被预测正确了，包括把烟雾预测为烟雾和把烟雾预测为了非烟雾（FN）；为了防止精确率和召回率出现矛盾的情况，使用 F-Measure 来综合考虑。公式如式 4，式 5 和式 6 所示。

$$\text{精确率}(P) = \frac{TP}{TP+FP} \quad (4)$$

$$\text{召回率}(R) = \frac{TP}{TP+FN} \quad (5)$$

$$F - \text{Measure}(F1) = \frac{2 \cdot P \cdot R}{P+R} \quad (6)$$

### 3.2 背景建模有效性验证

为了验证背景建模的有效性，将本文算法分为是否更新背景模型的两组对照试验，实验结果如表 2 所示。

表2 背景建模对照试验

视频编号		V1	V2	V3	V4	V5	V6
总帧数		2886	1200	1501	1393	5960	3000
烟雾帧数		2886	995	1000	527	1210	2300
未进行背景更新	P	94.5%	100%	57.8%	91.4%	64.3%	67.3%
	R	99.8%	99.2%	99.7%	94.9%	85.2%	97.6%
	F1	0.971	0.996	0.732	0.931	0.733	0.797
进行背景更新	P	93.4%	98.2%	90.9%	91.4%	94.7%	98.1%
	R	99.8%	96.9%	97.1%	94.5%	76.4%	66.4%
	F1	0.965	0.975	0.939	0.929	0.846	0.792

实验结果显示，在精确率方面，未进行背景更新时，烟雾识别的平均准确率远小于进行背景更新时，并且在环境较为复杂的 V3 和 V5 视频中出现了大量的误检。在召回率方面，是否进行背景更新差异不大。综上所述，对背景进行建模，能够更加稳定地检测各种环境下的烟雾。

### 3.3 与 2D 卷积相比较

为了验证本文提出的 3D 卷积网络效果好于 2D 卷积网络。在相同条件下，将图 4 的 3D 卷积结构去掉了最后的  $3 \times 1 \times 1$  卷积，将  $1 \times 3 \times 3$  和  $1 \times 5 \times 5$  卷积修改为  $3 \times 3$  和  $5 \times 5$  的 2D 多尺度卷积，将其作为算法 1，并与同样为 2D 卷积的 Yuan<sup>[10]</sup> 的算法进行实验对比，如表 3 所示。

实验结果显示，2D 卷积在对烟雾视频进行检测时拥有较低精确率和较高的召回率，但非常不稳

定，对不同环境下的烟雾检测差异较大，究其原因是在检测中存在与烟雾空间特征相似的干扰物，从而产生了误报。而 3D 卷积除了从空间特征的角度，还会在时间角度上进行检测判断，从而大大提高了精确率和稳定性。

表 3 6M3DC 与 2D 卷积相比较

视频编号		V1	V2	V3	V4	V5	V6
总帧数		2886	1200	1501	1393	5960	3000
烟雾帧数		2886	995	1000	527	1210	2300
算法 1	P	65.8%	46.4%	40.1%	42.2	7.8%	49.4%
	R	100%	99.2%	92.2%	100%	35.5%	99.3%
	F1	0.794	0.633	0.559	0.594	0.128	0.659
Yuan <sup>[10]</sup>	P	52.5%	38.1%	34.8%	40.6%	99.4%	49.3%
	R	100%	63.6%	69.8%	92.6%	26.5%	99.3%
	F1	0.689	0.477	0.465	0.564	0.419	0.659
6M3DC	P	93.4%	98.2%	90.9%	91.4%	94.7%	98.1%
	R	99.8%	96.9%	97.1%	94.5%	76.4%	66.4%
	F1	0.965	0.975	0.939	0.929	0.846	0.792

### 3.4 与 3D 卷积相比较

为了验证本文提出的 3D 卷积的有效性，将图 4 的卷积结构分别替换为单一的  $3 \times 3 \times 3$  和  $5 \times 5 \times 5$  的 3D 卷积作为算法 2 和算法 3，并与同样为 3D 卷积的薛继光<sup>[15]</sup>的算法进行比较，如表 4 所示。

表 4 6M3DC 与其他 3D 卷积相比较

视频编号		V1	V2	V3	V4	V5	V6
总帧数		2886	1200	1501	1393	5960	3000
烟雾帧数		2886	995	1000	527	1210	2300
算法 2	P	67.7%	54.7%	85.9%	66.7%	97.6%	65.2%
	R	99.8%	99.2%	97.1%	17.8%	78.1%	65.2%
	F1	0.807	0.705	0.911	0.281	0.868	0.652
算法 3	P	55.9%	99.8%	74.4%	93.8%	77.7%	80.6%
	R	99.8%	100%	97.1%	23.0%	77.4%	47.7%
	F1	0.717	0.999	0.843	0.369	0.775	0.599
薛继光 <sup>[15]</sup>	P	93.9%	100%	94.5%	0	0	0
	R	99.2%	91.5%	90.8%	0	0	0
	F1	0.965	0.955	0.926	0	0	0
6M3DC	P	93.4%	98.2%	90.9%	91.4%	94.7%	98.1%
	R	99.8%	96.9%	97.1%	94.5%	76.4%	66.4%
	F1	0.965	0.975	0.939	0.929	0.846	0.792

实验结果显示，对于 V1~V3 的中近距离的烟雾，所有 3D 算法都拥有着较高的平均召回率，但是精确率存在着较大的差异，薛继光<sup>[15]</sup>和 6M3DC 的平均准确率远远高于算法 2 和算法 3；对于 V4~V6 远距离的烟雾，由于距离远，存在着运动不显著的问题，所以相比于中近距离烟雾，整体召回率有较大

的下降，但 6M3DC 的平均召回率和平均准确率都要远远优于其他算法。薛继光的算法无法对远距离烟雾进行检测的原因是其在运动处理时采用了帧差法，无法检测到缓慢运动的烟雾，致使其后续网络无法进行检测。

3.5 实验结果可视化

为了验证本文算法的有效性，将算法 1，算法 2，算法 3 和本文算法进行了可视化，如图 9 所示，可见本文算法在不同环境下能够较为精确的覆盖所有烟雾区域。

















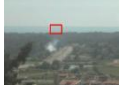







视频 编号	算法 1	算法 2	算法 3	本文 算法
V1				
V2				
V3				
V4				
V5				
V6				

图 9 实验结果可视化

4 结束语

本文设计的运动筛选和 6M3DC 网络相结合的视频早期烟雾检测算法，利用 RGB 颜色判断，均值 HASH 算法和多尺度 3D 卷积神经网络，对中近距离的烟雾的平均召回率达 97.9%，平均准确率达 94.2%；对远距离烟雾的平均召回率达 79.1%，平均准确率达 94.7%。能够较好地适应不同场景下的烟雾检测。未来的工作计划是对运动区域的筛选进行进一步地研究，并构建更为有效的卷积神经网络以适应缓慢运动的烟雾的检测。

参考文献：

[1]Chen T, Yin Y, Huang S, et al. The smoke detection for early fire-alarming system base on video processing[C]. 2006 International Conference on Intelligent Information Hiding and Multimedia, 2006: 427-430.

[2]Tian H, Li W, Ogunbona P O, et al. Detection and Separation of Smoke From Single Image Frames[J]. IEEE Transactions on Image Processing, 2018, 27(3): 1164-1177.

[3]Ma S, Zhang Y, Xin J, et al. An early forest fire detection method based on unmanned aerial vehicle vision[C]. 2018 Chinese Control And Decision Conference (CCDC), 2018: 6344-6349.

[4]Russo A U, Deb K, Tista S C, et al. Smoke Detection Method Based on LBP and SVM from Surveillance Camera[C]. 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018: 1-4.

[5]Yuan C, Liu Z, Zhang Y. Learning-Based Smoke Detection for Unmanned Aerial Vehicles Applied to Forest Fire Surveillance[J]. Journal of Intelligent & Robotic Systems, 2019, 93(1): 337-349.

[6]Wang L, Li A. Early fire recognition based on multi-feature fusion of video smoke[M]. 2017 36th Chinese Control Conference (CCC),2017:5318-5323.

[7]S.R V, Muruganand S. Smoke detection in video images using background subtraction method for early fire alarm system[M].2017 2nd International Conference on Communication and Electronics Systems (ICCES), 2017: 167-171.

[8]Hu Y, Lu X. Real-time video fire smoke detection by utilizing spatial-temporal ConvNet features[J]. Multimedia Tools and Applications, 2018, 77(22): 29283-29301.

[9]Luo Y, Zhao L, Liu P, et al. Fire smoke detection algorithm

based on motion characteristic and convolutional neural networks[J]. Multimedia Tools and Applications, 2018, 77(12): 15075-15092.

[10]Yuan F, Zhang L, Wan B, et al. Convolutional neural networks based on multi-scale additive merging layers for visual smoke recognition[J]. Machine Vision and Applications, 2019, 30(2): 345-358.

[11]Yin M, Lang C, Li Z, et al. Recurrent convolutional network for video-based smoke detection[J]. Multimedia Tools and Applications, 2019, 78(1): 237-256.

[12]Hu C, Tang P, Jin W, et al. Real-Time Fire Detection Based on Deep Convolutional Long-Recurrent Networks and Optical

Flow Method[C]. 2018 37th Chinese Control Conference (CCC), 2018: 9061-9066.

[13]Zeng J, Lin Z, Qi C, et al. An Improved Object Detection Method Based On Deep Convolution Neural Network For Smoke Detection[C]. 2018 International Conference on Machine Learning and Cybernetics (ICMLC), 2018: 184-189.

[14]Ji S, Xu W, Yang M, et al. 3D Convolutional Neural Networks for Human Action Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 221-231.

[15]薛继光. 基于时空域深度神经网络的视频烟雾检测研究[D]. 西南交通大学, 2018.