

时空域深度学习火灾烟雾检测

吴凡, 王慧琴, 王可

引用本文:

吴凡, 王慧琴, 王可. 时空域深度学习火灾烟雾检测[J]. 液晶与显示, 2021, 36(8): 1186–1195.

WU Fan, WANG Hui-qin, WANG Ke. Spatio-temporal deep learning fire smoke detection[J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(8): 1186–1195.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2020-0239>

您可能感兴趣的其他文章

Articles you may be interested in

一种高精度的卷积神经网络安全帽检测方法

High precision detection method of safety helmet based on convolution neural network

液晶与显示. 2021, 36(7): 1018–1026 <https://doi.org/10.37188/CJLCD.2020-0309>

基于改进Faster R-CNN的交通标志检测算法

Traffic sign detection algorithm based on improved Faster R-CNN

液晶与显示. 2021, 36(3): 484–492 <https://doi.org/10.37188/CJLCD.2020-0195>

基于注意力掩模融合的目标检测算法

Object detection algorithm based on attention mask fusion

液晶与显示. 2019, 34(8): 825–833 <https://doi.org/10.3788/YJYXS20193408.0825>

基于深度学习的航空对地小目标检测

Detection of small target in aerial photography based on deep learning

液晶与显示. 2018, 33(9): 793–800 <https://doi.org/10.3788/YJYXS20183309.0793>

基于深度学习的乳液泵缺陷检测算法

Defect detection algorithm of lotion pump based on deep learning

液晶与显示. 2019, 34(1): 81–89 <https://doi.org/10.3788/YJYXS20193401.0081>

文章编号:1007-2780(2021)08-1186-10

时空域深度学习火灾烟雾检测

吴 凡, 王慧琴*, 王 可

(西安建筑科技大学 信息与控制工程学院, 陕西 西安 710055)

摘要:烟雾是火灾早期检测的重要特征。传统机器学习及二维卷积神经网络烟雾检测算法对烟雾特征的提取局限于空间领域,无法提取时域信息。现有的三维卷积神经网络检测算法则存在计算成本高、检测时效低的问题,导致检测准确率和虚警率不理想。针对上述问题,本文提出一种基于时空域深度学习的烟雾视频检测方法。利用分块运动目标检测方法提取烟雾视频的运动目标,过滤非烟雾目标;同时将三维卷积神经网络拆分,形成一种二加一维时空域网络模块,提取时空域特征,提高检测时效。为抑制无关特征,引入注意力机制,增加压缩和激励网络重新标定特征通道权重,提升烟雾检测准确率。研究结果表明,本文所用算法的平均准确率为 97.12%,平均正确率为 97.06%,平均虚警率为 2.74%,平均检测帧率为 10.49 帧/s,满足火灾烟雾探测需求,检测时效得到明显提高。

关键词:烟雾检测;深度学习;时空域;运动目标检测;注意力机制

中图分类号:TP391.4;TH691.9 **文献标识码:**A **doi:**10.37188/CJLCD.2020-0239

Spatio-temporal deep learning fire smoke detection

WU Fan, WANG Hui-qin*, WANG Ke

(College of Information and Control Engineering,
Xian University of Architecture and Technology, Xian 710055, China)

Abstract: Smoke is an important feature of early fire detection. The extraction of smoke features by traditional machine learning and two-dimensional convolutional neural network smoke detection algorithms are limited to the spatial domain, and cannot extract temporal information. The existing three-dimensional convolutional neural network detection algorithm has the problems of high calculation cost and low detection time efficiency, which leads to unsatisfactory detection accuracy and false alarm rate. To solve the above problems, a smoke video detection method based on deep learning in spatio-temporal domain is proposed. The block moving target detection method is used to extract the moving targets of the smoke video and filter the non-smoke targets. At the same time, the three-dimensional convolutional neural network is split to form a two-plus-one-dimensional spatio-temporal network module, which extracts the characteristics of the spatio-temporal domain and improves the detection

收稿日期:2020-09-13; **修订日期:**2020-12-19.

基金项目:陕西省科技厅国际科技合作计划项目(No.2020KW-012);陕西省教育厅重点项目高端智库(No.18JT006);西安市科技局项目(No.GXYD10.1)

Supported by Shanxi Provincial Department of Science and Technology International Science and Technology Cooperation Program Project(No.2020KW-012); High-end Think Tank for Key Projects of Shaanxi Provincial Department of Education(No.18JT006); Xi'an Science and Technology Bureau Project(No. GXYD10.1)

* 通信联系人, E-mail: hqwang@xauat.edu.cn

time efficiency. In order to suppress irrelevant features, an attention mechanism is introduced to increase the compression and incentive network to recalibrate the weight of feature channels to improve the accuracy of smoke detection. The research results show that the average accuracy rate of the algorithm used in this paper is 97.12%, the average correct rate is 97.06%, the average false alarm rate is 2.74%, and the average detection frame rate is 10.49 frame/s. The needs of fire smoke detection is met, and the detection timeliness is improved significantly.

Key words: smoke detection; deep learning; spatio-temporal; moving target detection; attention mechanism

1 引言

火灾是自然和社会灾害中发生概率最高的灾害之一,对人类的生活和生命安全构成严重威胁^[1]。因此,快速而准确地检测火灾发生至关重要。

当前的火灾探测方式仍以各种火灾传感器为主,虽然其低价简单,但极易受到粉尘、气流以及高度等因素的干扰^[2]。通常情况下,明火在与空气充分接触之后才会产生,而在燃烧初期烟雾即已出现,因此准确检测烟雾能够比火焰检测更早发出火灾报警^[3]。

目前烟雾检测研究流程大致可分为疑似烟雾目标提取、烟雾特征提取以及烟雾检测几个重要部分。Barmoutis^[4]等基于 HSV (Hue, Saturation, Value) 颜色模型,使用结合烟雾颜色特征的背景差法提取帧内疑似烟雾区域。Park^[5]等结合帧差法和非参数颜色模型检测疑似烟雾目标,基于此算法设计了一种烟雾检测随机森林分类器。该算法可以加快烟雾检测速度,但检测烟区中易存在空洞,且用于复杂环境时烟雾检测虚警率高。在以上研究基础上,一些学者为增强分类器的性能,加入了机器视觉方法。李红娣^[6]等使用金字塔分解算法提取烟雾的金字塔纹理和边缘特征,并通过支持向量机(Support Vector Machine)进行训练和检测烟雾。Zhao^[7]等利用烟雾的颜色等特性,基于 CS Adaboost 算法对烟雾进行检测,该算法可有效地分辨浓雾和烟雾。上述方法的烟雾特征设计与提取多数由手工完成,需基于经验设定阈值作为识别烟雾的判断依据,未必能够反映烟雾的本质特征,其合理性会因烟雾本身和环境变

化受到影响。

近年来,基于深度学习的视频检测方法发展迅猛,作为一种性能强、适用性广的方法逐渐在火灾探测中得到应用^[8]。Kim^[9]等提出用卷积神经网络(Convolutional Neural Networks, CNN)检测视频中的烟雾。该研究利用输入视频图像的颜色信息提取烟雾候选区域,然后利用预训练的 CNN 进行烟雾检测。该方法较之前的传统检测方法误检率和漏检率得到改善,性能有所提高。陈俊周^[10]等融合烟雾的动静态纹理信息,提出基于级联卷积神经网络的烟雾纹理识别框架,提高烟雾检测的准确率,然而其将静态动态纹理信息分别处理导致算法复杂度增加,影响烟雾实时检测。孙颖^[11]提出了一种基于 3D 残差密集网络的烟雾检测算法,将残差网络和密集连接网络进行整合,形成 3D Residual Dense Block 网络模块,以提取烟雾的时空特征。神经网络算法虽然性能较优,但二维 CNN 无法提取时域特征,丢失了时间维度的帧间运动信息,因而检测效果不佳;三维 CNN 则因特征维度提升,计算成本大幅增加,影响检测时效和准确率。

本文提出一种基于时空域深度学习的火灾烟雾视频检测算法,利用分块的运动目标检测算法过滤非烟雾区域,再输入经预训练的二加一维时空域网络模型,提取烟雾的时空域特征,抑制无关特征,最后将烟雾区域分块标记,提高了检测准确率和时效。

2 相关算法概述

卷积神经网络是深度学习(Deep learning)的代表算法之一,模仿生物的视觉感知机制构建,可

进行监督学习和非监督学习。三维卷积神经网络是一种视频检测网络,在物体识别、动作检测等方面有着巨大优势。

一些研究使用二维神经网络对视频进行检测,研究对象大多数是视频的帧截图,忽略了帧间运动信息的时间序列,因此时间维度上的帧间运动信息会丢失。如果将二维卷积神经网络拓展到三维,便可同时提取时间和空间维度特征,增加时间维度的特征信息,滤波器的内核维度也因此会增加,在学习单个图片邻近像素的同时学习时间上接近的像素,即学习时空特征,其公式如式(1)所示:

$$V_{ij}^{xyz} = \tanh(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}), \quad (1)$$

(x, y, z) 是第 i 层第 j 个特征图像素点, P_i 、 Q_i 、 R_i 是第 i 层三维卷积核时空维度的大小, w_{ijm}^{pqr} 是前一层第 m 个特征图连接的权重, V_{ij}^{xyz} 是特征图像素点的结果值, m 是第 $i-1$ 层中与当前特征图相连的特征图个数, w_{ijm}^{pqr} 表示与 $i-1$ 层相连的第 m 个特征图的权重, $v_{(i-1)m}^{(x+p)(y+q)(z+r)}$ 表示第 $i-1$ 层中第 m 个特征图在 $(x+p, y+q, z+r)$ 处的值。三维卷积的原理如图 1 所示, H 、 W 、 L 分别为输入数据的高度、宽度与帧长度,使用空间尺寸为 $k \times k$ 、时间尺寸为 d 的卷积核进行卷积,得到后面的输出。

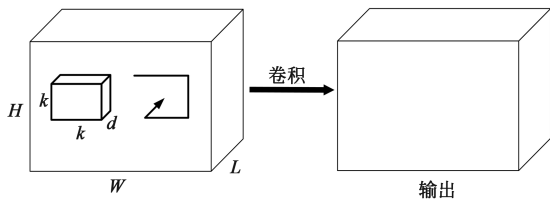


图 1 三维卷积原理图

Fig.1 Schematic diagram of three-dimensional convolution

3 时空域深度学习烟雾检测

3.1 疑似烟雾目标运动检测

由于烟雾属于运动目标,从视频图像中提取运动目标可滤除大量非烟雾区域,再对分割出的疑似烟雾区域进行预处理,输入网络模型进行烟

雾识别,可有效提高准确率和时效。为了使运动目标检测部分达到理想的效果,本研究将采用高斯混合模型和四帧差分法相结合的运动检测算法提取疑似烟雾目标,准备输入网络模型。

3.1.1 高斯混合模型

在高斯混合模型^[12] (Gaussian mixture model) 中,每一个背景图片像素点的描述都由 N 个高斯分布组成,这几个高斯分布的加权组合形成背景。

首先是构建初始背景模型。读取视频序列前 N 帧图片,通过预处理将其转化成灰度图片序列 $f_i (i = 1, 2, 3, \dots, N)$, 将图片中坐标为 (x, y) 的像素值 $f_i(x, y)$ 分别向这 N 个高斯分布的均值 $u_{i,l}(x, y)$ 赋值,随后用一个相对较大的方差将每个高斯分布初始化,至此高斯混合模型建立完成,表达式如下:

$$P(f_i) = \sum_{l=1}^N \omega_{i,l} \eta_l(f_i, u_{i,l}, \sigma_{i,l}), \quad (2)$$

其中, N 是高斯模型的数量, $\eta_l(f_i, u_{i,l}, \sigma_{i,l})$ 代表第 l 个高斯分布, $\omega_{i,l}$ 表示第 l 个高斯分布的权重,通常初始值设为 $1/M$ 。

第二步是前景目标检测。读取下一帧图片,经预处理后记做 f_{i+1} ,依照 ω/σ 的值从大到小的顺序将 $f_{i+1}(x, y)$ 对应的高斯模型排序,选择比值较大的前 B 个高斯分布生成该图片的背景, B 的表达式如下:

$$B = \arg \min_B (\sum_{l=1}^B \omega_{i,l} > T), \quad (3)$$

其中, T 为使用者设定的阈值,亦称门限参数,随后依照下面的表达式(4)进行匹配,若 $f_{i+1}(x, y)$ 与其背景模型中序号为 k ($k \leq B$) 的高斯分布匹配,则该点可以认定为背景,否则为前景目标。 λ 是前景阈值,一般设为 2.5。

$$|f_{i+1}(x, y) - u_{i,k}| \leq \lambda \sigma_{i,k}. \quad (4)$$

第三步是背景模型更新。新的一帧图片完成检测后,若 $f_{i+1}(x, y)$ 被判定来自于前景,即说明前 B 个高斯分布模型没有一个能与该点对应的模型相匹配,需要一个新的高斯分布函数代替 ω/σ 值最小的那个分布函数,其均值为 $f_{i+1}(x, y)$,给定的权值较小,方差较大。

与 $f_{i+1}(x, y)$ 匹配的高斯分布表达式按下面的表达式更新:

$$\omega_{t+1,l}(x,y) = (1-\alpha)\omega_{t,l}(x,y) + \alpha, \quad (5)$$

$$u_{t+1,l}(x,y) = (1-\rho)u_{t,l}(x,y) + \rho f_{t+1}(x,y), \quad (6)$$

$$\sigma_{t+1,l}(x,y) = \sqrt{(1-\rho)\sigma_{t,l}(x,y)^2 + \rho(f_{t+1}(x,y) - u'_{t,l}(x,y))^2}, \quad (7)$$

不与 $f_{t+1}(x,y)$ 匹配的高斯分布表达式则按照表达式(8)更新:

$$\omega_{t+1,l}(x,y) = (1-\alpha)\omega_{t,l}(x,y), \quad (8)$$

最终将权值归一化,式中 α, ρ 是通常由经验设定的学习速率。

3.1.2 四帧差分法

帧间差分法(Frame difference method)是通过获得相邻连续帧的差别以进行运动目标检测,具有算法简单、时间复杂度低的优点,对动态背景适应性好,不需要提取和更新背景。然而常用的二帧差分法无法解决双影与空洞问题,本研究采用性能较好的四帧差分法^[13]以消除目标检测过程中存在的空洞和双影现象。具体步骤如下:

(1)读取视频序列连续4帧图片,处理为灰度图片后进行中值滤波去噪,获得连续4帧预处理过的图片,并将其设为 $I_k(x,y)$, $I_{k+1}(x,y)$, $I_{k+2}(x,y)$, $I_{k+3}(x,y)$ 。

(2)将4帧图片进行间隔差分,即第1帧与第3帧、第2帧与第4帧差分,使用阈值进行分割,获得二值化图片,表达式如下:

$$d_{1k} = \begin{cases} 1 & |I_k(x,y) - I_{k+2}(x,y)| > Z \\ 0 & |I_k(x,y) - I_{k+2}(x,y)| \leq Z \end{cases}, \quad (9)$$

$$d_{2k} = \begin{cases} 1 & |I_{k+1}(x,y) - I_{k+3}(x,y)| > Z \\ 0 & |I_{k+1}(x,y) - I_{k+3}(x,y)| \leq Z \end{cases}, \quad (10)$$

式中, d_{1k}, d_{2k} 是差分处理过的图片, Z 是预设的二值化固定阈值。

(3)将第2步得到的二值化图片中运动目标的轮廓进行填充,由于背景为纯黑色,故选用易于分辨的纯白色对运动目标区域进行填充。

(4)为了减少双影现象,对第3步结果进行逻辑“与”操作,见下式:

$$D_k = d_{1k} \cap d_{2k}, \quad (11)$$

$$D_k = \begin{cases} 1 & d_k = 1 \\ 0 & d_k = 0 \end{cases}. \quad (12)$$

3.1.3 视频分块检测

烟雾部分明显特征是形状不确定,运动方向

和速度不规律,为了更便捷地标记运动区域,以及将疑似烟雾目标区域输入后续神经网络中进行学习,本文将原始视频数据的每一帧分成相同大小互不重合的小方块,进行如图2的分块运动检测,分块公式如下所示:

$$n_c = \left\lceil \frac{h_x}{h_x} \right\rceil$$

$$n_r = \left\lceil \frac{w_y}{w_y} \right\rceil, \quad (13)$$

其中, h_x, h_y 分别代表原始图像的宽和高, h_x, h_y 分别代表图像每个分块的宽和高, n_c, n_r 分别代表视频帧被分割的行数与列数。

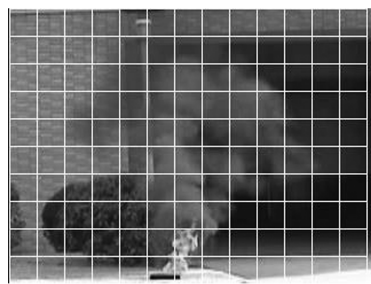


图2 烟雾图像分块处理示例

Fig.2 Example of block processing of smoke image

3.1.4 高斯混合模型与四帧差分法混合运动目标检测

本文采取四帧差分法与高斯混合模型混合算法进行运动目标检测。首先将读取的视频序列进行预处理,随后分别送入四帧差分改进算法和高斯混合模型当中进行前景目标提取,再将两个算法分割的前景目标进行“与”逻辑操作,通过连通性分析,形态学处理,得到疑似烟雾目标区域,最后使用分块将其标记并储存。具体算法流程如图3所示。

3.2 时空域深度学习火灾烟雾检测

烟雾本身外观特征复杂,也会环境变化增大特征差异。传统人工设计处理的特征难以描述烟雾的全部本质,极易受到相似目标的干扰,因此检测准确率不高。三维神经网络在视频检测,动作分类等领域优势巨大,而残差网络则在特征提取

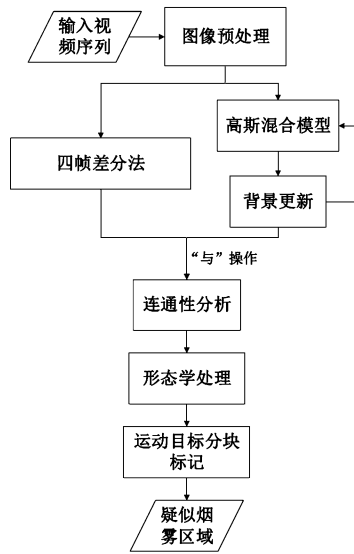


图 3 运动目标检测算法流程图

Fig.3 Moving target detection algorithm flow chart

方面拥有出色表现,本文结合二者优势,提出一种改进的二加一维时空域网络。将一般的三维卷积神经网络结构分解为二维加一维卷积网络层,先后提取视频图像的空域和时域特征,并加入注意力机制,利用该机制的小型子网络运算重新标定特征通道权重,过滤无关特征,最后通过全连接层得到最终检测结果。

3.2.1 二加一维时空域网络结构

三维卷积神经网络因其能同时提取时空域特征在视频检测领域得到应用,但是其存在参数量多、计算量高的问题。以最常见的三维网络模型 C3D(Convolutional 3D Network)为例,在使用数据集 Sports-1M 训练时,模型大小超过 300 MB。视频烟雾检测系统多用于嵌入式设备,这就要求烟雾时空域特征提取所用到的神经网络计算成本低,模型参数少。

二加一维神经网络模块是指将三维网络结构拆分为独立的二维空域卷积结构和一维时域网络结构,使得非线性数量增加一倍,同时分解交织的时间与空间信息,使得网络利于优化,从而在保证三维网络性能的前提下降低了计算成本。设输入特征图尺寸为 $l \times w \times h \times f$, 其中 l 为视频帧数, w 为视频宽度, h 为视频高度, f 为输入特征维度。原三维卷积核尺寸为 $t \times k \times k$, 将其拆分为 $1 \times k \times k$ 的空域卷积核与 $t \times 1 \times 1$ 的时域卷积核后,计算过程由 $t \times k \times k$ 变为 $t + k \times k$, 计算量得

到明显减少,然而参数量的锐减对模型的复杂性和表达能力有明显影响。由此需要在充分利用二加一维网络的优点的同时保留足够的参数量。

本文参考文献[14]的方法,通过一个超参数 M 更改中间特征通道数量,将二加一维参数量恢复到原三维网络的水平,图 4 为采用超参数 M 的二加一维结构。 M 的计算公式以及计算参数对比如下:

$$M_i = \left\lceil \frac{tk^2 N_{i-1} N_i}{k^2 N_{i-1} + t N_i} \right\rceil, \quad (14)$$

$$3D: N_{i-1} \times N_i \times t \times k \times k; (2+1)D: N_{i-1} \times M \times 1 \times k \times k + M \times N_i \times t \times 1 \times 1. \quad (15)$$

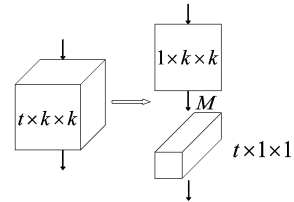


图 4 二加一维时空域网络结构

Fig.4 Spatio-temporal two-plus-one dimensional structure

在时域与空域卷积之后是批标准化层(Batch Normalization, BN)^[15], 其作用对象是每个隐层神经元,其输入分布在逐渐向非线性函数映射后,取值区间会向极限饱和区靠拢, BN 层可以将其强制拉回到均值为 0、方差为 1 的标准正态分布,使非线性变换函数的输入值落入对输入比较敏感的区域,有效解决不同层数据分布不一致和梯度消失问题。之后引入整流线性单元(Rectified Linear Unit, ReLU)激活函数层^[16],其作用是增加模型的非线性表达能力,防止过拟合现象,提升训练精度。

二加一维网络结构因为有 M 超参数的存在,使得时空域两个子卷积层之间增加了一个非线性操作,与原来同样参数量的三维卷积结构相比非线性操作数量翻倍,网络复杂度由此增加。第二个好处是时空域分解让优化的过程也分解开,三维时空卷积把空间信息和动态信息交织在一起,优化过程较为复杂。而二加一维卷积分别提取时空域特征,优化过程相对简单,可以使模型误差降低。

3.2.2 注意力机制层

为了提高检测效率,本文在二加一维网络结构中引入注意力机制。Hu^[17]等提出了一种压缩和激励网络(Squeeze-and-Excitation Network, SENet),在网络训练过程中可以自动重新标定特征,抑制对分类无用的特征,提高网络的分类识别能力。

该网络层包括以下 3 个处理步骤:

(1)压缩(Squeeze)操作:将大小为 $l \times w \times h \times f$ 的输入,使用一个全局池化层压缩输入张量中除特征通道维度 f 之外的所有维度,使其转化为一特征通道数大小的实数向量,大小为 $1 \times f$ 。

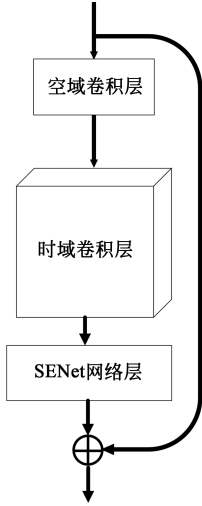


图 5 时空域注意力网络模块

Fig.5 Spatio-temporal attention network module

(2)激励(Excitation)操作:通过一个全连接层压缩转换的特征向量,使其维度降低到 f/r , 大小为 $1 \times f/r$,再经过 ReLU 函数激活后通过一个全连接层,得到一个输出维度与输入特征通道数相匹配的特征权重向量,大小为 $1 \times f$ 。

(3)权值重标定(Reweight)操作:使用 Sigmoid 函数将权重归一化,最后将 Excitation 操作得到的权重对特征通道进行加权,从而实现特征的重新标定。

将 SENet 和二加一维网络按图 5 的形式结合在一起,就形成了时空域注意力网络模块。

3.2.3 网络总体框架

本文提出的二加一维时空域网络采用时间卷积层和空间卷积层串联结构替代三维卷积结构提取时空特征,并引入注意力机制,过滤无关特征,提高了网络的检测性能。本文提出的网络结构整

体框架如图 6 所示。

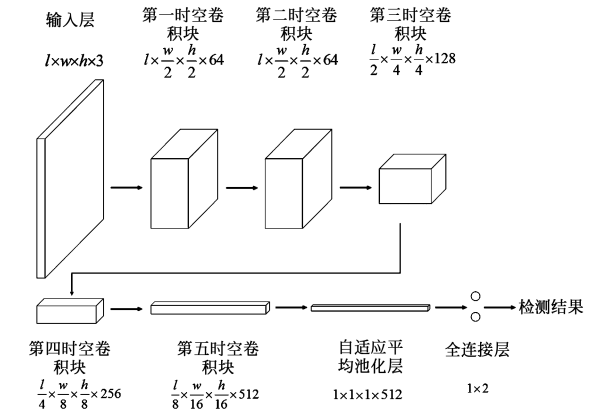


图 6 时空域深度卷积神经网络架构图

Fig.6 Graph of convolutional neural network in spatio-temporal

(1)输入层:输入为 l 帧连续 RGB 烟雾图像, 尺寸为 $l \times w \times h \times f$, 其中 l 为视频帧数, w 为视频宽度, h 为视频高度, f 为输入特征维度。本文输入为连续的 RGB 图像,故 f 取值为 3。

(2)第一时空卷积层:使用尺寸为 $3 \times 7 \times 7$ 的时空卷积核对输入进行卷积运算,输出特征图尺寸为 $l \times \frac{w}{2} \times \frac{h}{2}$, 通道数为 64, 并进行 BN 及 ReLU 函数处理。

(3)第一时空残差卷积块:使用尺寸为 $3 \times 3 \times 3$ 的时空卷积核对输入进行卷积运算,加入 SENet 网络层标定特征权重,输出特征图尺寸为 $l \times \frac{w}{2} \times \frac{h}{2}$, 不改变输出特征图尺寸和通道数, 并进行 BN 及 ReLU 函数处理。

(4)第二时空残差卷积块:使用尺寸为 $3 \times 3 \times 3$ 的时空卷积核对输入进行卷积运算,加入 SENet 网络层标定特征权重,输出特征图尺寸为 $\frac{l}{2} \times \frac{w}{4} \times \frac{h}{4}$, 通道数为 128, 并进行 BN 及 ReLU 函数处理。

(5)第三时空卷积块:使用尺寸为 $3 \times 3 \times 3$ 的时空卷积核对输入进行卷积运算,加入 SENet 网络层标定特征权重,输出特征图尺寸为 $\frac{l}{4} \times \frac{w}{8} \times \frac{h}{8}$, 通道数为 256, 并进行 BN 及 ReLU 函数处理。

(6)第四时空卷积块:使用尺寸为 $3 \times 3 \times 3$ 的

时空卷积核对输入进行卷积运算,加入 SENet 网络层标定特征权重,输出特征图尺寸为 $\frac{l}{8} \times \frac{w}{16} \times \frac{h}{16}$,通道数为 512,并进行 BN 及 ReLU 函数处理。

(7)输出层:将输入数据送入通道数为 512 的全局平均池化层,最后通过全连接层得到是否为烟雾的评估结果。

4 实验及结果分析

本研究实验环境为 Win10 64 位操作系统,内存为 16 GB,CPU 为 Intel Core i7-8700,图形处理器为 NVIDIA GeForce RTX2070 8G,深度学习框架为 Pytorch,在 Python 和 Microsoft Visual Studio Code 编程环境下实现。

4.1 实验设置

为验证本算法的有效性,数据集从课题组自行拍摄的烟雾视频以及网络烟雾视频集中截取,选取不同场景下的烟雾正样本及负样本进行火灾识别实验。场景分为室内与室外开阔空间两种环境,干扰物有与烟雾颜色相近的云朵、加湿器水雾、白色背景墙,树林等。总样本为 208 184 张,每 8 张连续图片为一个视频块,其中 188 792 张用于 train 数据集,9 792 张用于 val 数据集,9 600 张作为 test 数据集,每帧样本大小为 32×32 。

4.2 网络模型训练

本次研究使用 4.1 节的数据集进行训练,在参数设定环节中,我们选择随机梯度下降(Stochastic Gradient Descent,SGD)作为网络损耗的优化器,训练 Epoch 数量设置为 50,一次训练所选取的样本数(Batch_size)为 64,设置动量系数为 0.9 和权重衰减值为 0.000 5,初始学习率设置为 0.001,学习率每经过 10 个训练周期(Epoch)便衰减为原来的 1/10。

4.3 评价指标

为了测试本研究算法的性能,采用文献[18]的评价标准。计算得到准确率(ACC),正确率(TPR),虚警率(FPR),公式如下:

$$ACC = \frac{(TP+TN)}{N}, \tag{16}$$

$$TPR = \frac{TP}{TP+FN}, \tag{17}$$

$$FPR = \frac{FP}{TN+FP}, \tag{18}$$

式中,N 为总烟雾样本数,TP 为被检测为有烟雾的有烟雾样本数,TN 为被检测为无烟雾的无烟雾样本数,FP 为被检测为有烟雾的无烟雾样本数,FN 为被检测为无烟雾的有烟雾样本数。

4.4 实验结果与性能对比

本次研究共选择 10 段视频作为测试,具体描述说明如表 1 所示。

表 1 测试视频说明

Tab.1 Test video description

序号	是否为 烟雾	视频描述
1	是	黑色
2	是	黑色,有玻璃幕墙干扰
3	是	烟雾快速移动,变化剧烈
4	是	黑色背景,烟雾稀薄
5	是	有行人,烟雾稀薄
6	是	背景复杂,烟雾区域小
7	是	缓慢运动稀薄
8	否	黑暗环境,大型彩灯和车灯
9	否	白色墙背景,移动行人
10	否	阴天环境,植物轻微晃动

图 7 所示为部分视频烟雾检测效果,图中检测出的烟雾区域为白色方块标记的位置。实验结果表明,本文采用的烟雾检测算法在 10 段测试视频中都取得了良好的效果。对于不同颜色、不同浓度、不同速度的烟雾都能够取得理想的检测效果。在加湿器水雾、云等类烟运动物体、多云和白色背景墙的复杂背景干扰下也具有较好的鲁棒性。

为充分验证本算法的有效性,将去掉第三与第四时空残差卷积块,减少特征通道维度变换次数的模型称为算法 1,所有时空残差卷积块去掉 SENet 网络层的模型称为算法 2,共同与本文提出的算法进行对比,对比的数值为 10 个视频评价标准的平均数值,分别为平均准确率(AACC),平

均正确率(ATPR)和平均虚警率(AFPR),结果如表2所示。

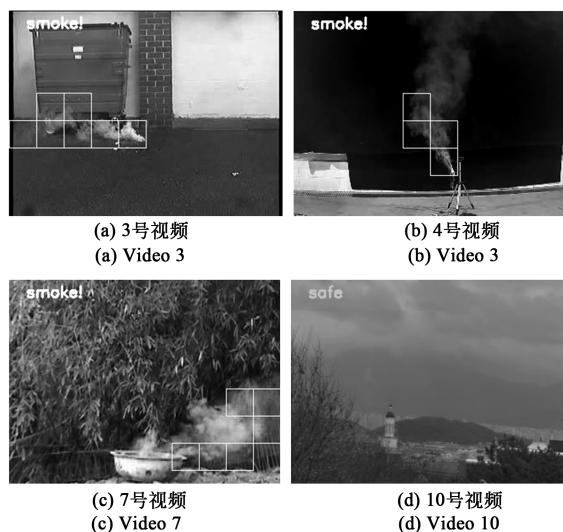


图7 部分检测结果。(a)、(b)、(c)为有烟雾视频,(d)为无烟雾视频。

Fig.7 Part of the test results. (a),(b),(c) Smoke videos; (d) Non-smoke video.

表2 不同结构性能对比

Tab.2 Performance comparison of different network structures

算法	AACC%	ATPR%	AFPR%
算法1	94.03	93.58	4.92
算法2	96.19	95.47	3.44
本文算法	97.12	97.06	2.74

由表2可知,算法1由于减少时空残差卷积块导致网络的层数减少,模型深度不足,对烟雾特征提取不充分,故而本文算法比算法1平均准确率增加3.09%,虚警率降低2.18%,说明较深层数的网络模型能够有效提取烟雾特征。而对于算法2,本文算法增加了SENet层以重新标定特征权重,减少了冗余的无用特征,因此平均准确率提升0.97%,平均虚警率也有所改善,降低了0.7%,说明增加Senet层能够使网络模型对烟雾特征的注意力提高。

为进一步验证烟雾检测算法的有效性,本文还将采用LBP和LBPV检测烟雾的文献[19]方法,采用PCA主成分分析和Inception Resnet v2网络算法的文献[20]方法,采用VGG16和Res-

net50网络融合算法的文献[21]方法以及采用3D密集残差网络的文献[11]方法加入本文算法性能对比,结果如表3所示。

表3 与其他算法性能对比

Tab.3 Performance comparison with other algorithms

算法	AACC%	ATPR%	AFPR%
文献[19]	82.05	83.78	21.98
文献[20]	92.1	90.68	4.60
文献[21]	94.78	94.24	3.98
文献[11]	96.05	95.60	2.91
本文算法	97.12	97.06	2.74

由表3结果可知,与传统的手工设计提取烟雾特征以及机器视觉算法的检测方法相比,卷积神经网络能够实现端对端、自动选取烟雾的检测特征,且特征种类和规模更加丰富,因此文献[20],文献[21]、文献[11]的算法及本文算法平均准确率均可以达到90%以上,文献[21]的方法平均虚警率比文献[19]降低了18%。而对于各种基于二维卷积神经网络的算法,本文算法使用的网络模型能够对图像和运动信息建模,同时提取空域与时域特征,最终平均准确率提升2.34%,虚警率改善明显,降低了1.24%,说明提取时空域特征可以显著提升烟雾检测效果。对于文献[11]方法,由于本文算法增加了非线性因素且时空域分开运算,因而检测准确率提升1.07%,虚警率提升0.17%。

为了验证烟雾检测算法的时效,本文以每秒传输帧数(Frames Per Second,FPS)作为对比指标,将采用三维密集残差网络的文献[11]方法,以及去掉本文算法时空网络块中SENet层的算法3加入性能对比,结果如表4所示。

表4 检测效率对比

Tab.4 Comparison of detection efficiency

算法	FPS
文献[11]	9.25
算法3	10.67
本文算法	10.49

由表4结果可知,本文的二加一维网络算法

比标准的三维网络检测速率有明显提升,且未因 SENet 层的加入出现大幅降低的现象,对比算法 3,检测准确率有所提升,综合效果更好。

5 结 论

本文为了准确地从视频中检测烟雾,充分提取视频的时空域特征,同时改善三维网络模型的检测时效问题,提出了一种加入注意力机制的二加一维时空域深度学习检测算法。利用分块的运动目标

检测算法,过滤非烟雾目标,经预处理后输入到二加一维神经网络模型进行时空域特征提取。为抑制无关特征,使用注意力机制重新标定特征通道,经全连接层输出检测结果后将烟雾区域分块标定。在实验数据集测试得到的结果中,平均准确率为 97.12%,平均正确率为 97.06%,平均虚警率为 2.74%,平均 FPS 为 10.49 帧/s。实验数据表明,该算法可以有效减少复杂场景、类烟目标对烟雾特征的干扰,相比现有三维 CNN 算法提升了检测速率,取得了良好的检测效果。

参 考 文 献:

- [1] 史劲亭,袁非牛,夏雪.视频烟雾检测研究进展[J].中国图象图形学报,2018,23(3):303-322.
SHI J T, YUAN F N, XIA X. Video smoke detection: a literature survey [J]. *Journal of Image and Graphics*, 2018, 23(3): 303-322. (in Chinese)
- [2] 吴凡.基于深度学习的火灾检测算法研究与实现[D].杭州:杭州电子科技大学,2020.
WU F. Research and implementation of fire detection algorithm based on deep learning [D]. Hangzhou: Hangzhou Dianzi University, 2020. (in Chinese)
- [3] 袁非牛,夏雪,李钢,等.面向烟雾识别与纹理分类的 Gabor 网络[J].中国图象图形学报,2019,24(2):269-281.
YUAN F N, XIA X, LI G, *et al.* GaborNet for smoke recognition and texture classification [J]. *Journal of Image and Graphics*, 2019, 24(2): 269-281. (in Chinese)
- [4] BARMPOUTIS P, DIMITROPOULOS K, GRAMMALIDIS N. Smoke detection using spatio-temporal analysis, motion modeling and dynamic texture recognition [C]//*Proceedings of the 2014 22nd European Signal Processing Conference*. Lisbon: IEEE, 2014: 1078-1082.
- [5] PARK J O, KO B C, NAM J Y, *et al.* Wildfire smoke detection using spatiotemporal bag-of-features of smoke [C]//*Proceedings of 2013 IEEE Workshop on Applications of Computer Vision*. Clearwater Beach: IEEE, 2013: 200-205.
- [6] 李红娣,袁非牛.采用金字塔纹理和边缘特征的图像烟雾检测[J].中国图象图形学报,2015,20(6):772-780.
LI H D, YUAN F N. Image based smoke detection using pyramid texture and edge features [J]. *Journal of Image and Graphics*, 2015, 20(6): 772-780. (in Chinese)
- [7] ZHAO Y Q, LI Q J, GU Z. Early smoke detection of forest fire video using CS Adaboost algorithm [J]. *Optik-International Journal for Light and Electron Optics*, 2015, 126(19): 2121-2124.
- [8] 陈献明,王阿川,王春艳.基于深度学习的木材表面缺陷图像检测[J].液晶与显示,2019,34(9):879-887.
CHEN X M, WANG A C, WANG C Y. Image detection of wood surface defects based on deep learning [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(9): 879-887. (in Chinese)
- [9] KIM Y J, KIM E G. Image based fire detection using convolutional neural network [J]. *Journal of the Korea Institute of Information and Communication Engineering*, 2016, 20(9): 1649-1656.
- [10] 陈俊周,汪子杰,陈洪瀚,等.基于级联卷积神经网络的视频动态烟雾检测[J].电子科技大学学报,2016,46(6):992-996.
CHEN J Z, WANG Z J, CHEN H H, *et al.* Dynamic smoke detection using cascaded convolutional neural network for surveillance videos [J]. *Journal of University of Electronic Science and Technology of China*, 2016, 46(6): 992-996. (in Chinese)
- [11] 孙颖.基于 3D 残差密集网络的视频烟雾检测研究[D].长春:东北师范大学,2019.
SUN Y. Video smoke detection based on 3D residual dense network [D]. Changchun: Northeast Normal University, 2019. (in Chinese)

- [12] 程淑红,马继勇,张仕军,等.改进的混合高斯与YOLOv2融合烟雾检测算法[J].计量学报,2019,40(5):798-803.
CHENG S H, MA J Y, ZHANG S J, *et al.* Smoke detection algorithm combined with improved Gaussian mixture and YOLOv2 [J]. *Acta Metrologica Sinica*, 2019, 40(5): 798-803. (in Chinese)
- [13] 纪青华,禹素萍.基于Surendra背景减除法和四帧差分法的目标检测算法[J].计算机应用与软件,2014,31(12):242-244.
JI Q H, YU S P. Object detection algorithm based on surendra background subtraction and four-frame difference [J]. *Computer Applications and Software*, 2014, 31(12): 242-244. (in Chinese)
- [14] TRAN D, WANG H, TORRESANI L, *et al.* A closer look at spatiotemporal convolutions for action recognition [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 6450-6459.
- [15] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//*Proceedings of the 32nd International Conference on Machine Learning*. Lille: JMLR.org, 2015: 448-456.
- [16] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks [J]. *Journal of Machine Learning Research*, 2011, 15: 315-323.
- [17] HU J, SHEN L, ALBANIE S, *et al.* Squeeze-and-excitation networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [18] 张斌,魏维,高联欣,等.基于时空域深度神经网络的野火视频烟雾检测[J].计算机应用与软件,2019,36(9):236-242,259.
ZHANG B, WEI W, GAO L X, *et al.* Wildfire video smoke detection based on spatio-temporal deep neural network [J]. *Computer Applications and Software*, 2019, 36(9): 236-242, 259. (in Chinese)
- [19] YUAN F N. Video-based smoke detection with histogram sequence of LBP and LBPV pyramids [J]. *Fire Safety Journal*, 2011, 46(3): 132-139.
- [20] 王涛,宫宁生,蒋贵祥.基于深度学习的烟雾识别研究[J].电子技术应用,2018,44(10):131-135.
WANG T, GONG N S, JIANG G X. Smoke recognition based on the depth learning [J]. *Application of Electronic Technique*, 2018, 44(10): 131-135. (in Chinese)
- [21] 王洋,程江华,刘通,等.一种多网络模型融合的烟雾检测方法[J].计算机工程与科学,2019,41(10):1771-1776.
WANG Y, CHENG J H, LIU T, *et al.* A smoke detection method based on fusing multiple network models [J]. *Computer Engineering and Science*, 2019, 41(10): 1771-1776. (in Chinese)

作者简介:



吴凡(1995—),男,河北张家口人,硕士研究生,2018年于河北建筑工程学院获得学士学位,主要从事图像处理及深度学习方面的研究。E-mail: 13294068818@sina.cn



王慧琴(1970—),女,山西长治人,博士,教授,2002年于西安交通大学获得博士学位,主要从事智能信息处理、信息理论与应用、信息技术与管理、数字建筑等方面的研究。E-mail: hqwang@xauat.edu.cn