

❖ 特约专栏 ❖

深度学习的 MPCANet 火灾图像识别模型设计

张秀玲^{1,2}, 侯代标¹, 张逞逞¹, 周凯旋¹, 魏其珺¹

- (1. 燕山大学 河北省工业计算机控制工程重点实验室, 河北 秦皇岛 066004;
2. 燕山大学 国家冷轧板带装备及工艺工程技术研究中心, 河北 秦皇岛 066004)

摘要: 针对火灾发生时, 火灾图像背景复杂、人工特征提取过程繁琐、对火灾图像的识别泛化能力不强、容易出现精度不高、误报和漏报等问题, 提出了张量对象特征提取的多线性主成分分析 (Multilinear Principal Component Analysis, MPCA) 深度学习算法的火灾图像识别新方法。利用 MPCANet 建立火灾图像识别模型, 通过 MPCA 算法学习滤波器作为深度学习网络卷积层卷积核, 对张量对象的高维图像进行特征提取, 并把蜡烛图像和烟花图像作为干扰。通过仿真实验并与其他火灾图像识别方法对比得到提出的图像识别方法识别精度达到了 97.5%、误报率 1.5%、漏报率 1%。实验表明: 该方法可以有效解决火灾图像识别存在的问题。

关键词: 深度学习; MPCANet; 张量对象分析; 火灾图像识别

中图分类号: TP183 文献标志码: A DOI: 10.3788/IRLA201847.0203006

Design of MPCANet fire image recognition model for deep learning

Zhang Xiuling^{1,2}, Hou Daibiao¹, Zhang Chengcheng¹, Zhou Kaixuan¹, Wei Qijun¹

- (1. Key Laboratory of Industrial Computer Control Engineering of Hebei Province, Yanshan University, Qinhuangdao 066004, China;
2. National Engineering Research Center for Equipment and Technology of Cold Strip Rolling, Yanshan University, Qinhuangdao 066004, China)

Abstract: In view of the complicated background of the fire image, the complicated process of extracting the artificial feature, the poor generalization ability of the fire image, the low accuracy, false alarm rate, missing rate, the novel method for detecting fire images of multilinear principal component analysis (MPCA) was presented in the paper. The fire image recognition model was established by using MPCANet. Through the MPCA algorithm, the learning filter was used as the convolution kernel of deep learning network convolution layer, and the feature extraction of high dimensional images of tensor objects was taken, and candle images and fireworks images were taken as interference. Compared with other fire image recognition methods, the recognition accuracy of the proposed image recognition method

收稿日期: 2017-08-10; 修订日期: 2017-10-28

基金项目: 河北省自然科学基金-钢铁联合研究基金 (E2015203354); 河北省高校创新团队领军人才培养计划项目 (LJRC013); 河北省教育厅科学研究计划河北省高等学校自然科学研究重点项目 (ZD2016100); 2016 年燕山大学基础研究专项课题 (理工类) 培育课题 (16LGY015); 秦皇岛市科技局自筹项目 (201703A229)

作者简介: 张秀玲 (1968-), 女, 教授, 博士, 主要从事基于深度学习的智能识别与应用方面的研究。Email: zxlysu@ysu.edu.cn

reaches 97.5%, false alarm rate of 1.5%, missing rate of 1%. Experiments results show that this method could effectively solve the problems of fire image recognition.

Key words: deep learning; MPCANet; tensor object analysis; fire image recognition

0 引言

随着图像处理技术和计算机科学的不断发展, 图像处理的火灾识别监测技术迎来了迅猛的发展^[1]。该技术将机器学习的模式识别方法与数字化图像处理技术相结合^[2]。火灾图像包括烟雾图像和火焰图像, 尤其是火焰图像具有红外辐射特性、纹理特性、蔓延增长趋势等特征, 为早期火灾多物理量的探测和火灾的辨识奠定了基础。

国内外关于火灾图像的识别技术做了许多研究, 秦俊提出了基于飞行时间的深度图像变化率的火焰识别方法, 利用飞行时间法结合火焰深度图特征对火灾进行识别^[3]。

以上传统火灾图像识别方法是基于火灾图像一种或两种特征对火灾图像进行辨识, 但是却非常依赖于人工特征的选取^[4]。而且人工选取的图像特征基于特定数据集, 使识别算法对其他背景下的数据效果偏低, 泛化能力差。深度学习算法的出现, 为图像识别提供了新方法。王光耀使用机器学习的卷积神经网络(Convolutional Neural Network, CNN)对火灾火焰进行识别, 识别率达到了 97.1%^[5]。傅天驹利用深度学习 CNN 提出基于深度学习的林火图像识别的方法, 利用 CNN 自动提火灾图像特征, 避免了传统方法中人工提取特征的复杂性和盲目性^[6], 避免了传统识别算法复杂的特征提取和数据重构过程。

CNN 使用不同大小的卷积核, 使网络具有了不同的感受野, 可以提取到不同的图像特征, 随着卷积层数的加深, 提取更高级的图像特征, 这种表达更加抽象和概念化, 更容易识别。但是 CNN 的加深在训练中容易出现参数爆炸现象, 由于使用 SGD 算法优化网络, 梯度反向传播时容易出现梯度消失和梯度弥散的问题和训练小样本数据出现过拟合现象^[7]。针对以上问题, Chan 等^[8]提出了卷积网络的新变体, 即主成分分析网络(Principal Component Analysis Net

work, PCANet), 网络由卷积层中的 PCA 滤波器组, 非线性处理层中的二值化以及合并层中的块状直方图组成。它不利用 SGD 算法优化网络, 避免梯度反向传播时出现梯度消失和剃度弥散。整个网络结构简单, 训练的参数大大减少, 有效解决了参数爆炸问题。

CNN 和 PCANet 都是把二维图像输入网络中训练, 提取特征。当处理多维图像时, 特征提取可能表现不佳, 因为图像结构内的空间关系没有被充分利用。彩色火灾图像是具有列, 行和色彩模式的三维(三阶张量)对象^[9]。为解决存在的问题和对图像自身的特点文中提出了 MPCANet, 使用张量对象对高维火灾图像进行处理, 多线性特征提取方法, 多线性主成分分析(Multilinear Principal Component Analysis, MPCA)它将多维数据表示为张量不是矢量, 具有三个关键的优点: 保留的多维结构, 较低的计算需求以及较少的估计参数^[10]。国外学者研究发现, 张量更适合对多维图像进行识别分类^[11]。从高维火灾图像中提取高级语义特征, 提取的特征有效利用多维图像内的空间关系, 减少张量对象的类内不变形。文中设计了基于 MPCANet 的火灾图像识别模型, 提取的特征训练 SVM 分类器^[12]。并与基于 CNN 的火灾图像识别方法的识别结果进行对比, 评估提出方法的有效性。

1 MPCA 算法

MPCA 是 PCA 到多维的一个延伸^[13]。PCA 是投影向量到向量, MPCA 是投影张量到张量, 投影的结构相对简单, 另外运算在较低维度的空间进行, 因此处理高维度数据时有低运算量的优势 MPCA 可以捕获大部分原始张量输入变化的多线投影来进行特征提取, 通过将原始问题分解为一系列多投影子问题来进行, 因此适合处理高维火灾图像。

N 阶张量表示为 $A \in R^{I_1 \times I_2 \times \dots \times I_N}$, 共有 N 个 $I_n, n=1, 2, \dots, N$, 每个 I_n 都是张量 A 的 n 模式。图 1 为张量在模式 1 展开图。

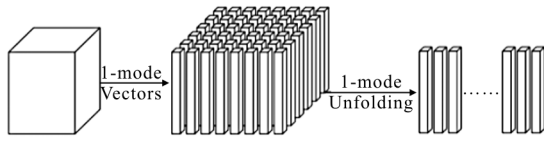


图 1 张量模式 1 展开图

Fig.1 Expansion pattern of tensor-model 1

A 与矩阵 $U \in R^{J_n \times I_n}$ n -mode 积定义为 $A \times_n U$:

$$(A \times_n U)(i_1, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N) = \sum_{i_n} A(i_1, \dots, i_n) \cdot U(j_n, i_n) \quad (1)$$

MPCA 目的找到投影矩阵 $U^{(n)\top} \in R^{P_n \times I_n}$ 把张量, $m=1, 2, \dots, M$ 从原始张量空间 $R^{I_1} \otimes R^{I_2} \otimes \dots \otimes R^{I_N}$ 映射为张量子空间 $R^{P_1} \otimes R^{P_2} \otimes \dots \otimes R^{P_N}$ 的张量 $y_m \in R^{P_1 \times P_2 \times \dots \times P_N}$, $m=1, 2, \dots, M, P_n < I_n, n=1, 2, \dots, N$, 所以 y_m 可以用 A_m 表示为:

$$A_m \in R^{I_1 \times I_2 \times \dots \times I_N} y_m = A_m \times_1 U^{(1)\top} \times_2 U^{(2)\top} \times_1 \dots \times_N U^{(N)\top} \quad (2)$$

也可以用矩阵的形式表示为:

$$A_{m(n)} = U^{(n)} \cdot Y_{m(n)} \cdot (U^{(n+1)} \otimes U^{(n+2)} \otimes \dots \otimes U^{(N)} \otimes U^{(1)} \otimes U^{(2)} \otimes \dots \otimes U^{(n-1)}) \quad (3)$$

式中: \otimes 为 kronecker 积; $A_{m(n)}$ 为张量 A_m 在 n -mode 的展开矩阵; $Y_{m(n)}$ 为张量 Y_m 的 n -mode 展开矩阵。

MPCA 目标是确定最大化总张量散射 Ψ_y 的 N 个投影矩阵 $U^{(n)\top} \in R^{P_n \times I_n}, n=1, 2, \dots, N$ 。

$$\left\{ \bar{U}^{(n)}, n=1, 2, \dots, N \right\} = \arg \max_{\bar{U}^{(1)}, \bar{U}^{(2)}, \dots, \bar{U}^{(N)}} \Psi_y \quad (4)$$

总散射张量 Ψ_y :

$$\Psi_y = \sum_{m=1}^M \left\| \bar{y}_m - \bar{y} \right\|_F^2 \quad (5)$$

式中: \bar{y} 为平均张量 ($\bar{y} = (1/M) \sum_{m=1}^M y_m$)。

做为公式(4)的解, $\bar{U}^{(1)}, \bar{U}^{(2)}, \dots, \bar{U}^{(n-1)}, \bar{U}^{(n)}, n=1, 2, \dots, N$ $\bar{U}^{(n+1)}, \bar{U}^{(n+2)}, \dots, \bar{U}^{(N)}$ 为其他投影矩阵, 矩阵 $\bar{U}^{(n)}$ 由 P_n 特征向量组成。对应于以下矩阵的最大的 P_n 特征值。

$$\Phi^{(n)} = \sum_{m=1}^M (A_{m(n)} - \bar{A}_{(n)}) \cdot \bar{U}_{\Phi_{(n)}} \cdot \bar{U}_{\Phi_{(n)}}^T (A_{m(n)} - \bar{A}_{(n)})^T \quad (6)$$

$$\bar{U}_{\Phi_{(n)}} = (\bar{U}^{(n+1)} \otimes \bar{U}^{(n+2)} \otimes \dots \otimes$$

$$\bar{U}^{(N)} \otimes \bar{U}^{(1)} \otimes \bar{U}^{(2)} \otimes \dots \otimes \bar{U}^{(n-1)}) \quad (7)$$

式中: $\bar{A}_{(n)} = (1/M) \sum_{m=1}^M A_{m(n)}, A_{m(n)}$ 由公式(3)给出。

最终得到了张量子空间投影的特征向量 $\bar{U}^{(n)}, n=1, 2, \dots, N$ 和投影后的张量 y_m 。

2 MPCANet 火灾图像识别模型设计

MPCANet 由 MPCA 学习的滤波器组、二值化层和直方图层代替深度学习 CNN 的卷积层、非线性处理层和全连接层, 构成的两层深度学习网络模型, 火灾图像识别模型如图 2 所示。

火灾识别模型设计过程分为以下 4 步。

(1) MPCANet 输入层。输入图像为 M 个三阶张量火灾图像 $x_m \in R^{I_1 \times I_2 \times I_3}, m=1, 2, \dots, M$ 进行训练, 如图 3 所

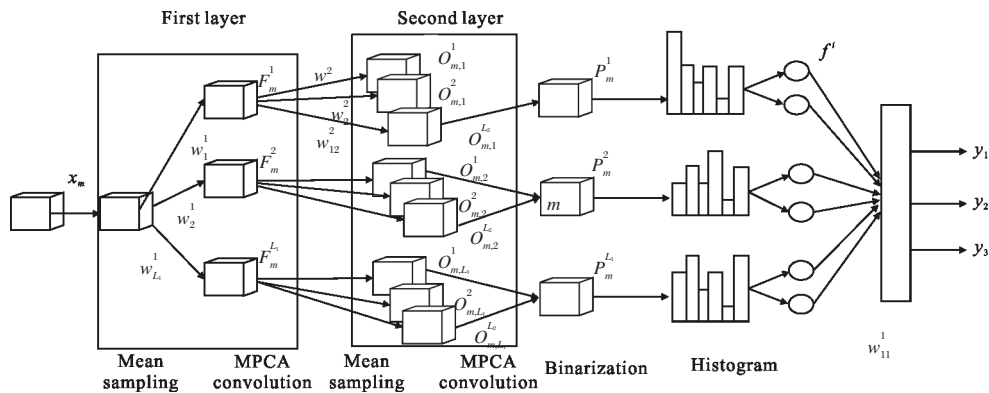


图 2 MPCANet 火灾图像识别模型结构

Fig.2 MPCANet fire image recognition model structure

示, I_1, I_2 为火灾图像的宽和高, I_3 表示火灾图像的 RGB 三通道。首先对输入张量 $x_m \in R^{I_1 \times I_2 \times I_3}$ 的像素进行分块, 分块大小 $k_1 \times k_2 \times k_3, 1 < k_i < I_i$ 。第 m 张张量图像的所有(重叠)分块为 $x_{m,i}^1 \in R^{k_1 \times k_2 \times k_3}, i=1,2,\dots,I_1 I_2 I_3$, 把所有分块按列排列 $X_m^1 \in \{x_{m,1}^1, x_{m,2}^1, \dots, x_{m,I_1 I_2 I_3}^1\}$ 。所有 M 张图像重复以上处理过程得到 $X^1 = \{X_1^1, X_2^1, \dots, X_M^1, m=1,2,\dots,M\}$, 上标 1 表示网络的第一层, 下标 1,2, m 表示输入图像索引。

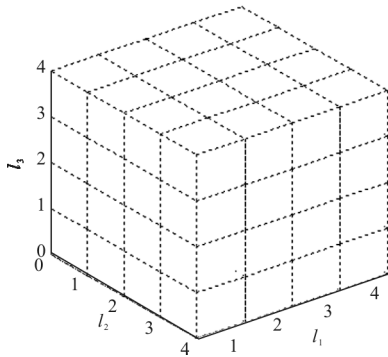


图 3 三阶张量火灾图像

Fig.3 Third-order tensor fire images

(2) MPCA 卷积层。对输入层得到的 $X^1 = \{X_1^1, X_2^1, \dots, X_M^1, m=1,2,\dots,M\}$ 使用第 1 章节中的 MPCA 算法进行处理, 得到投影矩阵 $U = \{U_1^n \in R^{k_n \times P_n}, n=1,2,3\}$, 每个分块张量用投影矩阵表示为:

$$S_{m,i}^1 = x_{m,i}^1 \times_1 U_1^{(1)\top} \times_2 U_1^{(2)\top} \times_3 U_1^{(3)\top} \in R^{P_1 \times P_2 \times P_3} \quad (8)$$

之后把 $S_{m,i}^1$ 转化为向量 $f_{m,i}^1$, 假设 MPCANet 第 j 层的滤波器的个数为 L_j , 取 $f_{m,i}^1$ 的前 L_j 个元素生成新向量 $\tilde{f}_{m,i}^1$, 重复以上过程得到向量:

$$f_m = \left\{ \tilde{f}_{m,1}^1, \tilde{f}_{m,2}^1, \dots, \tilde{f}_{i,I_1 I_2 I_3}^1 \in R^{L_i I_1 I_2 I_3}, i=1,2,\dots \right\} \quad (9)$$

对 f_m 的每一行进行张量化 $F_m^{l_1} \in R^{I_1 \times L_1 \times I_3}, l_1=1,2,\dots,L_1$, 生成 x_m 在第一层的特征输出。第一层的输出作为第二卷积层的输入, 处理过程和第一卷积层相似。最后得到第二卷积层的输出。

(3) 网络输出层。对第二卷积层的输出的 $O_{m,l_1}^{l_2} \in R^{I_1 \times L_2 \times I_3}, l_2=1,2,\dots,L_2$ 特征张量使用 Heaviside 函数做二

值化(大于 0 映射为 1, 否则为 0), $\tilde{O}_{m,l_1}^{l_2} \in R^{I_1 \times L_2 \times I_3}, l_2=1,2,\dots,L_2$ 是二值化后的张量, 然后每个张量乘以一个加权值后相加, 得到张量 T_m^h :

$$T_m^h = \sum_{l_2=1}^{L_2} 2^{l_2-1} (\tilde{O}_{m,l_1}^{l_2}) \quad (10)$$

然后将张量分成 B 块, 得到分块的大小为 $k_1 k_2 k_3 \times B$, 像素取值范围 $[0, 2^{L_2-1}]$ 的整数, 计算分块张量的 T_m^h 直方图, 直方图矩阵大小为 $2^{L_2} \times B$, 将直方图向量化 $\text{Bhist}(T_m^h)$, 然后将 T_m^h 所有的向量化 $\text{Bhist}(T_m^h)$ 按列排列得到每张训练图片的特征向量 $f^m: f^m = [\text{Bhist}(T_m^1), \text{Bhist}(T_m^2), \dots, \text{Bhist}(T_m^{L_1})] \in R^{(2^{L_2})^{L_1} \times B}$ SVM 分类层。网络输出得到的特征向量 f^m 作为 SVM 分类器的输入, 训练分类器, 最后得到 3 类分类结果。

(4) 算法复杂度分析。算法复杂度主要有以下部分组成, 生成张量分块 $x_{m,i}^1 \in R^{k_1 \times k_2 \times k_3}$ 复杂度为 $k_1 k_2 k_3 + k_1 k_2 k_3 I_1 I_2 I_3$, 生成矩阵 Φ_n 公式(6)和特征分解过的复杂度为 $O(k_1 k_2 k_3 I_1 I_2 I_3)$ 和 $O((k_1 k_2 k_3)^3)$, MPCANet 第 i 层卷积的复杂度为 $L_i k_1 k_2 k_3 I_1 I_2 I_3$, 第二卷积层输出进行二值化的复杂度为 $2 L_2 I_1 I_2 I_3$, 对二值化的输出做直方图, 复杂度为 $O(I_1 I_2 I_3 B L_2 \log 2)$, 整个 MPCANet 的复杂度 $O(k_1 k_2 k_3 I_1 I_2 I_3 (L_1 + L_2) + (K_1 K_2 K_3)^3 + K_1^2 K_2 K_3 I_1 I_2 I_3)$ 。

3 实验结果与分析

利用 MPCANet 火灾图像识别模型对火灾数据集进行训练和测试, 数据集图片由百度图片搜索得到。数据集图片包括白天和夜晚各个火灾发生阶段的火灾图像, 另外蜡烛火焰和烟花图像作为干扰图像, 数据构成见表 1, 图 4 为训练样本、干扰训练图像、测试图像。

表 1 火灾图像数据集

Tab.1 Fire image data set

Classification	Fire	Suspected fire
Totality	1 800	1 400
Train	1 600	1 200
Val	100	100
Test	100	100

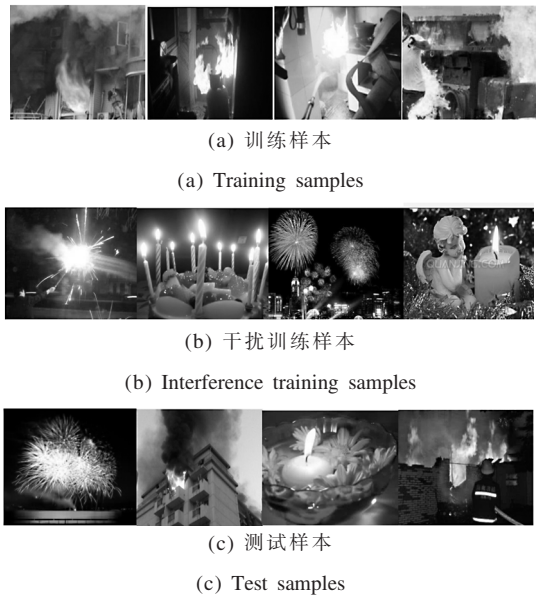


图 4 火灾图像数据集图像

Fig.4 Fire image database image

为了减少运算量,把图片大小统一处理为 $224 \times 224 \times 3$,三阶张量图像表示为 $\{x_m \in R^{224 \times 224 \times 3}, m=1,2,\dots, 3\ 200\}$ 。张量分块大小 $7 \times 7 \times 3, 9 \times 9 \times 3$,分块的重叠比为 0,实验对比分块组合实验效果。为对比该算法有效性,图 5 是 MPCANet、PCANet 的识别率曲线对比图,图中 MPCANet 为张量分块大小为 $7 \times 7 \times 3$ 的实验曲线,MPCANet1 代表张量分块 $9 \times 9 \times 3$ 的实验曲线。图 6 是传统深度图识别方法识别错误率曲线,错误率包括误报率和漏报率,图 7 为 CNN 识别方法优化曲线,包括识别精度和训练误差曲线,表 2 为算法识别结果对比表。

张量分块越小有助于增强非线性避免过拟合,加快运算速度,提高识别效果。由图 6 可知,网络选用 $7 \times 7 \times 3$ 比 $9 \times 9 \times 3$ 的张量分块的识别率更高。文中

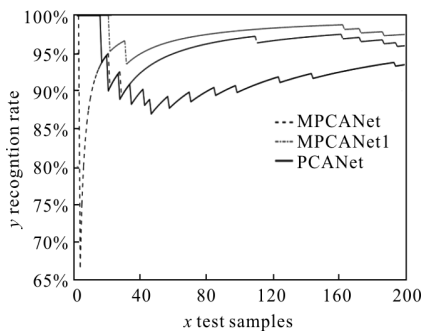


图 5 测试识别率曲线

Fig.5 Curve of test recognition rate

算法在训练完成的识别模型中,测试一张样本的时间为 0.8 s,满足在真实环境下的实用性要求。基于 matconvnet 的 CNN 网络结构,添加了 Dropout, BatchNormalization 层和 RELU 激活函数。综合图 6、7 的实验结果,给出了表 2 的实验结果对比。在存在干扰的环境中,文中的火灾图像识别算法可以在图像较少的情况下正确识别率、误报率和漏报率都高于于其他算法,证明文中提出的算法的有效性。

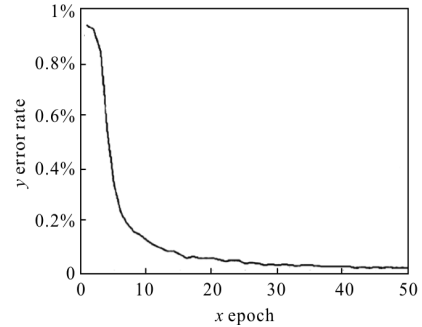


图 6 迭代次数与错误率关系图

Fig.6 Relationship between the number of iterations and the error rate

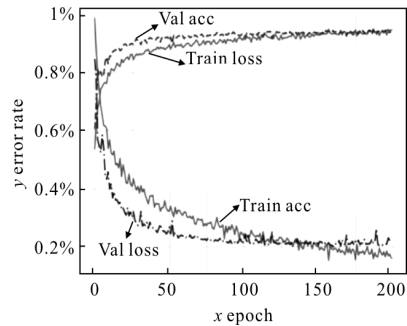


图 7 CNN 优化曲线

Fig.7 CNN optimized curve

表 2 识别率对比表

Tab.2 Comparison table of recognition rate

Algorithm	Recognition rate	False alarm rate	Omission rate	Interference
Depth-map	91.5%	3.8%	4.7%	y
CNN	93.6%	3.6%	2.8%	y
PCANet	94.9%	2%	1.8%	y
Text algorithm	97.5%	1.5%	1%	y

4 结论

文中 MPCANet 深度学习网络和 PCANet 相比,在 PCANet 的基础上进行改进,提出的 MPCANet 主

要使用张量对高维图像进行自动特征提取,充分利用图像空间结构特征,而且训练过程不需要误差反传,大量减少训练的参数,提高了训练效率。充分利用了图像的空间结构信息和高级语义特征,在少于普通 CNN 深度学习网络训练样本情况下比 CNN 构建的模型识别精度更高,而且火灾图像数据集中包含火灾发生过程中的各个阶段和干扰图像,提高了火灾图像识别的泛化能力,解决了火灾在较强干扰环境下误报和漏报问题。

参考文献:

- [1] Fan Yizhou, Ma Hongbing. Video based forest-fire smoke detection [J]. *J Tsinghua Univ (Sci& Technol)*, 2015 (2): 243-250. (in Chinese)
- [2] Xie Di, Tong Ruofeng, Tang Min, et al. Distinguishable method for video fire detection [J]. *Journal of Zhejiang University (Engineering Science)*, 2012, 46 (4): 698-704. (in Chinese)
- [3] Han Yuchong, Qin Jun, Ma Xingming, et al. Identification of fire flame based on variation rate of time-of-flight-depth-map method [J]. *Infrared and Laser Engineering*, 2014, 43(1): 338-344. (in Chinese)
- [4] Minchul Shin, Munsang Kim, Dong-Soo Kwon. Baseline CNN structure analysis for facial expression recognition[C]// IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 2016: 724-729.
- [5] Wang Guangyao. Research on fire detection methods based on machine learning [D]. Dalian: Dalian University of Technology, 2015. (in Chinese)
- [6] Fu Tianju, Zheng Change, Tian Ye, et al. Forest fire recognition based on deep convolutional neural network under complex background [J]. *Computer and Modernization*, 2016(3): 52-57. (in Chinese)
- [7] Chen Huiying, Wang Weibing, Wang Tingfeng, et al. Relationship between performance of stochastic parallel gradient descent algorithm and distribution rule of deformable mirror[J]. *Chinese Optics*, 2016, 9(4): 432-438.(in Chinese)
- [8] Chan T H, Jia K, Gao S, et al. PCANet: A simple deep learning baseline for image classification [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5017-5032.
- [9] Lathauwer L D, Moor B D, Vandewalle J. On the best rank-1 and rank-(R_1, R_2, \dots, R_n) approximation of higher order tensors [J]. *Siam Journal on Matrix Analysis and Applications*, 2000, 21(4): 1324-1342.
- [10] Song X, Meng L, Shi Q, et al. Learning tensor-based features for whole-brain fMRI classification[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention-Miccai, 2015: 613-620.
- [11] Hou J, Chau L P, Magnenat Thalmann N, et al. Scalable and compact representation for motion capture data using tensor decomposition [J]. *IEEE Signal Processing Letters*, 2014, 21(3): 255-259.
- [12] Pei Liran, Jiang Pingping, Yan Guozheng. Research on fall detection system based on support vector machine [J]. *Optics and Precision Engineering*, 2017, 25(1): 182-187. (in Chinese)
- [13] Lu H, Plataniotis K N, Venetsanopoulos A N. MPCA: Multilinear principal component analysis of tensor objects.[J]. *IEEE Transactions on Neural Networks*, 2008, 19(1): 18-39.