

基于改进深度残差网络的低功耗表情识别

杜 进 陈云华 张 灵 麦应潮

(广东工业大学计算机学院 广州 510000)

摘 要 为了提高表情识别率并降低表情识别的功耗,提出一种基于改进深度残差网络的表情识别方法。残差学习在解决深度卷积神经网络退化问题、使网络层次大幅加深的同时,进一步增加了网络的功耗。为此,引入具有生物真实性的激活函数来代替已有的整流线性单元(Rectified Linear Units,ReLU)函数,并将其作为卷积层激活函数对深度残差网络进行改进。该方法不仅提高了残差网络的精度,而且训练出的网络权重可直接作为与该深度残差网络具有相同结构的深度脉冲神经网络的权重。将该深度脉冲神经网络部署在类脑硬件上时,其能够以较高的识别率和较低的能耗进行表情识别。

关键词 表情识别,残差网络,Leaky Integrate and Fire(LIF)神经元,卷积神经网络

中图分类号 TP183 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.09.051

Energy-efficient Facial Expression Recognition Based on Improved Deep Residual Networks

DU Jin CHEN Yun-hua ZHANG Ling MAI Ying-chao

(School of Computers,Guangdong University of Technology,Guangzhou 510000,China)

Abstract To improve recognition rate and reduce power consumption of facial expression recognition systems,this paper proposed a facial expression recognition method using an improved deep residual networks (ResNets). Residual learning solves the degradation problem of the deep Convolutional Neural Networks(CNNs) to a certain degree and increases the network layers infinitely,but it makes deep CNNs face a more serious power consumption problem. To solve this problem,this paper introduced a new biologically-plausible activation function to improve ResNets and get a facial expression recognition method with both higher performance and lower power consumption. The Rectified Linear Units (ReLU) in the convolutional layers of ResNets are replaced with the new activation function Noisy Softplus. The obtained weights by using the improved ResNets can be directly applied to a deep Spiking Neural Networks(SNNs) architecture derived from the ResNets. The experimental results suggest that the proposed facial expression recognition method is able to achieve higher recognition rate and lower power consumption on a neuromorphic hardware.

Keywords Facial expression recognition, Residual networks, Leaky Integrate and Fire (LIF) neurons, Convolutional neural networks

1 引言

表情识别在人机交互、智能驾驶等领域有着广阔的应用前景,一直是图像理解、机器视觉、模式识别等领域的研究热点。表情识别的关键是获取有效的特征。在传统表情识别方法中,特征往往需要借助专业的领域知识对数据进行分析得到,如局部二值模式(Local Binary Pattern,LBP)^[1]、定向梯度直方图(Histogram of Oriented Gradients,HOG)^[2]、尺寸不变特征变换(Scale Invariant Feature Transform,SIFT)^[3]等。尽管如此,这些特征对头部姿态、环境光照等的要求较高,对现实场景下的表情识别效果有限。

近年来,随着深度学习方法的研究不断发展^[4]以及计算

机硬件性能的快速提升,以卷积神经网络(Convolutional Neural Networks,CNNs)为代表的人工神经网络(Artificial Neural Networks,ANNs)在表情识别应用中取得了显著效果^[5-7]。研究表明^[8-9],在一定程度上,神经网络层越深,网络的学习能力越好,分类能力越强。当网络层达到一定深度后,精度会达到饱和,此时网络层继续增加,精度反而会下降。He等^[10]将这种问题称为退化(Degradation),并提出了相应的解决方法——在深度神经网络中引入残差学习方法,构建深度残差网络(Residual Networks,ResNets),152层的残差网络在大型数据库ImageNet^[11]中,top-5的错误率仅为5.71%,大幅低于普通CNNs的错误率。ResNets虽然大幅提高了图像识别的精度,但由于网络层增加所导致的能耗也

到稿日期:2017-07-26 返修日期:2017-12-08 本文受广东省自然科学基金项目(2016A030313713,2014A030310169),广东省产学研合作专项项目(2014B090904080),广东省交通运输厅科技项目(科技-2016-02-030)资助。

杜进(1992-),男,硕士生,CCF会员,主要研究方向为深度学习、表情识别,E-mail:15598002907@163.com;陈云华(1977-),女,博士,副教授,主要研究方向为深度学习、神经形态计算等,E-mail:yhchen@gdut.edn.cn(通信作者);张灵(1968-),女,博士,教授,主要研究方向为模式识别、智能化信息处理、人工智能等;麦应潮(1994-),男,硕士生,主要研究方向为神经形态计算、机器学习。

成几何倍数增长。如何在增加网络层数的同时降低执行能耗仍然是一个待解决的难题。

脉冲神经网络(Spiking Neural Networks, SNNs)^[12]作为类脑计算机对人脑进行模拟的载体,其神经元之间的通信机制是电压升降形成的脉冲(Spike),而非 ANNs 中的数值。SNNs 中的神经元并不像 ANNs 神经元会在每个迭代周期更新状态,而是仅当膜电位达到某一个特定的阈值时才发放脉冲。相对于 ANNs, SNNs 能以更低的功耗完成识别任务^[13],但目前很难基于脉冲规则学习算法来实现对深度 SNNs 的训练,因此 SNNs 进行模式识别的精度不及深度 ANNs;另外, ANNs 的识别功耗远高于 SNNs。为此,研究者通过研究生物神经元特性来建立具有生物真实性的 ANNs 神经元模型,使得可以通过训练深度 ANNs 来获得具有相同结构的深度 SNNs 的权重,从而提高 SNNs 的模式识别精度。LIU 等^[14]对 LIF 神经元特性进行研究,提出了一种能对 LIF 神经元的脉冲响应特性进行表达的 ANNs 激活函数模型 Noisy Softplus,以该模型为基础训练得到的深度 CNNs 的精度有所提高;同时,训练得到的权重可以直接作为具有相同结构的深度 SNNs 的网络权重,从而获得更高精度的 SNNs 网络,最终能以较高的精度和较低的功耗执行模式分类任务。他们以 Noisy Softplus 作为激活函数训练出 5 层 CNNs^[14],其已在手写数字集 MNIST^[15]上取得了 97.03% 的识别率,但在表情识别方面的效果尚待考证。由于表情识别比手写数字识别更加复杂,层数较少的神经网络很难对其进行有效的识别。因此,本文将 Noisy Softplus 函数引入网络层数不受限制的深度残差网络,对其进行改进,以获得能对表情进行有效识别的残差网络;同时,训练得到的网络权重可以直接用于具有相同结构的深度 SNNs,从而获得具有高精度、低功耗的表情识别方法。

2 残差学习

研究表明^[10],退化问题并不是由过拟合引起的,而是过多的层数及参数导致网络难以找到最优解。残差学习方法有助于解决退化问题。简单来说,就是两个或两个以上连续的神经网络层形成一个堆叠层,在这个堆叠层上添加一个快捷连接(Shortcut Connection),即输入跳过堆叠层直接连接到堆叠层的输出位置。实现快捷连接的过程叫恒等映射(Identity Mapping)。一个堆叠层和一个快捷连接一起形成了一个残差单元(Residual Unit)。具体地,将残差单元的输出表示为 $H(x)$,其中 x 表示残差单元的输入, $F(x) = H(x) - x$ 表示一个残差映射,本文对 $F(x)$ 进行拟合。相比直接对 $H(x)$ 进行拟合,残差映射 $F(x)$ 对输出变化更敏感,参数的调整幅度更大,从而加快了学习速度,提高了网络优化性能。

一个完整的残差单元结构如图 1 所示。正向传播过程可以定义为:

$$y = F(x, \{W_i\}) + x \quad (1)$$

其中, x 和 y 分别表示残差单元的输入向量和输出向量,函数 $F(x, \{W_i\})$ 表示待学习的残差映射。图 1 中的堆叠层只有两层,则 $F = W_2 \sigma(W_1 x)$, σ 表示 ReLU 激活函数。通过快捷连接进行 $F+x$ 运算,再进行一次 ReLU 激活,即 $\sigma(y)$ 。式(1)中的快捷连接既不引入额外的参数,也不增加计算的复杂度。以上说明,在相同参数、深度、宽度和计算成本条件下,深度

残差网络参数调整的幅度会更大,优化效果更好,识别率更高。

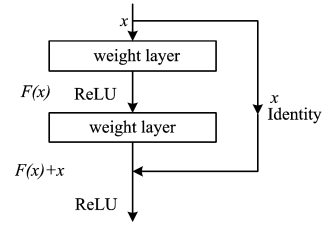


图 1 残差单元结构

Fig. 1 Residual unit structure

式(1)中的 x 和 F 的形状必须相等,如果不相等(比如输入或输出通道发生变化时),则可以通过快捷连接来执行线性投影 W_i ,进行形状匹配:

$$y = F(x, \{W_i\}) + W_i x \quad (2)$$

3 改进深度残差网络

3.1 LIF 神经元模型

ANNs 发展到现在已经显示出了强大的学习能力,并且在越来越多的认知领域超越了人脑,但是超高的功耗也限制了其应用。与人工神经元相比,生物神经元之间的通信机制是膜电位升降的脉冲,这种通信方式在功耗上具有得天独厚的优势。常用的生物神经元模型有 Hodgkin-Huxley(HH)模型^[16]、Izhikev-ich 模型^[17]、Adaptive Exponential Integrate and Fire(AdEx IF)模型^[18]、Leaky Integrate and Fire(LIF)模型^[19]等。其中,LIF 模型因在生物精确性和运算复杂度两个方面具有较好的综合性能,成为了最广泛使用的模型。LIF 神经元膜电位遵循以下动力学公式:

$$\tau_m \frac{dV}{dt} = V_{rest} - V + R_m I(t) \quad (3)$$

静止时,膜电位的值为 V_{rest} ,之后膜电位 V 随输入电流 I 的变化而变化,其中 t 表示时间,膜时间常数 $\tau_m = R_m C_m$, R_m 为膜电阻, C_m 为膜电容。如果给定一个恒定电流值 I ,则 LIF 神经元的响应函数即发放速率为:

$$\lambda_{out} = [t_{ref} - \tau_m \log(1 - \frac{V_{th} - V_{rest}}{IR_m})]^{-1} \quad (4)$$

其中, t_{ref} 表示绝对不应期,当 $IR_m > V_{th} - V_{rest}$ 时,会触发一个脉冲,否则膜电位达不到阈值 V_{th} ,输出发放率为 0。在实际情况下,输入的电流并不是恒定电流,而是带有噪声的电流。LIF 神经元对噪声电流的响应函数为:

$$\lambda_{out} = [t_{ref} + \tau_m \int_{\frac{V_{rest} - \mu \tau_m}{\sigma \sqrt{\tau_m}}}^{\frac{V_{th} - \mu \tau_m}{\sigma \sqrt{\tau_m}}} \sqrt{\pi} \exp(u^2) (1 + erf(u)) du]^{-1} \quad (5)$$

其中, μ 和 σ 分别表示电流的平均值和方差, $erf(u)$ 表示对输入电流 u 求高斯误差。图 2 给出了带有噪声的 LIF 神经元的响应曲线。其中实线表示的是输入恒定电流的 LIF 神经元的响应函数,这与整流线性单元(Rectified Linear Units, ReLU)^[20]函数的形状基本一致,随着噪声的增加,曲线水平上升,渐渐接近 Softplus 函数的形状。ReLU 和 Softplus 函数都是人工神经网络中常用的激活函数,如图 3 所示。其表达式分别为:

$$\text{ReLU}: y = \begin{cases} x, & x > 0 \\ 0, & x < 0 \end{cases} \quad (6)$$

4 实验以及结果分析

4.1 实验环境

本文以 pytorch 深度学习框架为基础,使用 Python 编程语言,在 ubuntu14.04 操作系统上进行相关实验。硬件平台为:第四代 Intel 酷睿 i7 6700k CPU, Nvidia Geforce GTX1080 GPU,显存 16 GB。

实验采用 The Extended Cohn-Kande Dataset (CK+) [22] 数据库、Karolinska 导向表情数据库 (Karolinska Directed Emotional Faces, KDEF) [23] 和 GENKI-4K [24] 数据库作为实验对象,来评估本文所提算法的性能。

CK+数据库包含了 123 个人的 593 个表情序列,其中带标签的有 309 个表情序列。本实验从带标签的每个表情序列中选取 8 张共计 2472 张表情图片。然后把选取的图片按照标签类别分成 7 个表情类(愤怒、蔑视、厌恶、恐惧、高兴、悲伤、惊讶),每个类别按照 4:1 的比例随机分成训练集和验证集。KDEF 数据库包含 70 个人的 7 种不同表情(愤怒、中性、厌恶、恐惧、高兴、悲伤、惊讶),每种表情从 5 个不同角度拍摄两次,一共有 4900 张照片。该数据库按照表情分为 7 类,随机选取每个人不同表情的一种姿态为验证集,其余图片作为训练集,刚好是 4:1 的比例。GENKI-4K 数据库包含 4000 张人脸照片,包含高兴和中性两个类,各 2000 张照片,涵盖了不同的人脸、光照、角度、成像条件以及头部姿态,两个类别按照 4:1 的比例分成训练集和验证集。根据 VGG-19 和 ResNet-19 的结构,需要把输入的表情图片尺寸调整为 227×227 像

素。图 6 给出了一些本文实验数据库样例。

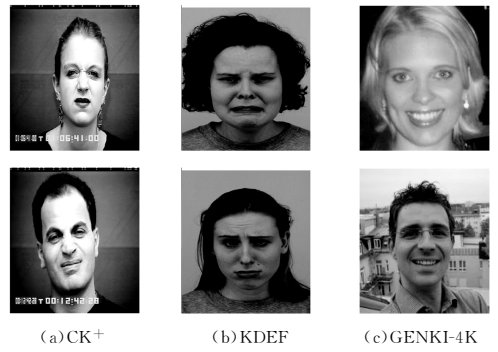


图 6 实验数据库样例

Fig. 6 Some examples of experimental database

需要注意的是,在对使用了 Noisy Softplus 函数的 ResNet-19 结构进行训练时,根据式(8)、式(9)可知, σ^2 须为正数,则数据在初始化过程中需要被归一化到 $[0, 1]$ 。作为对比,本实验同时进行 VGG-19 和 ResNet-19 的训练。神经网络的最后一层为全连接层,这与训练数据的类别有关,CK+ 和 KDEF 数据库最后的全连接层数都为 7,而 GENKI-4K 数据库为 2。

4.2 识别性能

本文按照设定好的网络结构对 3 种数据库进行了多次重复实验。表 1 列出了上述数据库在不同神经网络结构中的不同表情类的识别率以及平均识别率。

表 1 不同网络结构的识别率

Table 1 Recognition accuracy with different network structures

数据库	网络模型	使用的激活函数	恐惧	愤怒	藐视	厌恶	高兴	中性	悲伤	惊讶	平均
CK+	VGG-19	ReLU	94.55	96.45	95.67	94.87	94.68	—	94.23	95.07	95.07
	ResNet-19	ReLU	94.87	98.01	97.55	97.15	98.08	—	97.55	97.15	97.19
	ResNet-19(本文)	Noisy Softplus	96.77	98.32	98.34	98.24	98.08	—	98.14	98.15	98.05
KDEF	VGG-19	ReLU	92.8	83.72	—	95.15	93.55	87.32	85.05	84.8	88.91
	ResNet-19	ReLU	95.71	85.71	—	97.14	98.57	90	87.14	92.86	92.45
	ResNet-19(本文)	Noisy Softplus	95.71	92.86	—	98.57	98.57	92.86	87.14	95.71	94.49
GENKI-4K	VGG-19	ReLU	—	—	—	—	92.32	89.55	—	—	90.93
	ResNet-19	ReLU	—	—	—	—	95.82	92.59	—	—	94.05
	ResNet-19(本文)	Noisy Softplus	—	—	—	—	96.66	94.63	—	—	95.57

(单位: %)

3 种表情数据库的识别率随训练次数变化的情况如图 7 所示。使用 ReLU 函数的 ResNet-19 在 CK+, KDEF 和 GENKI-4K 上的平均识别率分别为 97.19%, 92.25% 和 94.05%, 比使用 VGG-19 结构的识别率分别提高了 2.12%, 3.54% 和 3.12%。ResNet-19 和 VGG-19 均为 19 层的网络结构,但 ResNet-19 比 VGG-19 多了两个卷积层,少了两个全连接层,这样前者通过多的卷积层可以学习到更高级的特征,

并且减少的全连接层降低了网络复杂度,提高了学习效率。而改进的 ResNet-19 仅仅把 ReLU 激活函数替换成 Noisy Softplus,在各数据库中的识别率已经分别达到了 98.05%, 94.49% 和 95.57%, 稍高于 ReLU 的识别率。

本文还与其他文献中相关表情数据库的实验结果进行了对比,如表 2 所列。与其他方法相比,本文提出的表情识别方法具有较高的识别率。

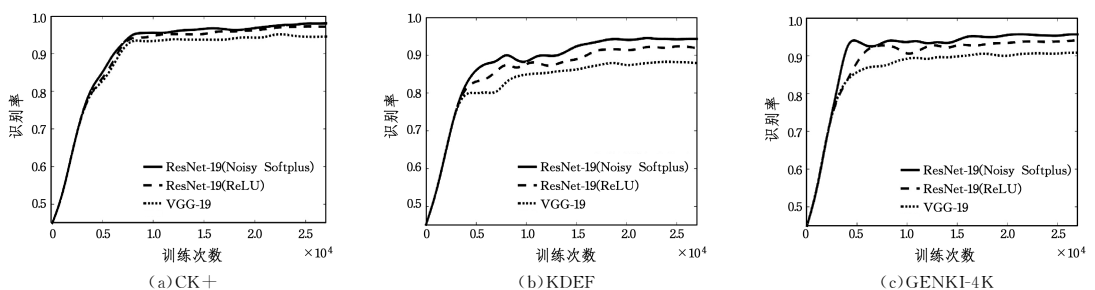


图 7 本文实验数据库的识别率随训练次数变化的情况

Fig. 7 Changes of experimental database recognition accuracy with number of training times

表 2 本文方法与其他文献方法的对比

Table 2 Comparison of this method and other methods in references

(单位:%)

CK+		KDEF		GENKI-4K	
方法	识别率	方法	识别率	方法	识别率
AlexNet ^[25]	92.20	AdaBoost ^[26]	87.20	AdaBoost ^[29]	79.22
3DCNN-DAP ^[5]	92.40	LDBP+SVM ^[27]	92.15	LBP+FAP ^[30]	92.00
LDBP+SVM ^[27]	94.98	SCAE ^[28]	92.52	PwDV+ELM ^[29]	93.42
ResNet-19(本文)	98.01	ResNet-19(本文)	94.49	ResNet-19(本文)	95.57

实验结果证明了本文提出的改进残差网络的表情识别方法不仅是可行的,而且在一定程度上具有生物真实性的 Noisy Softplus 激活函数比 ReLU 的性能更优。并且,本文实验的 3 种数据库的差异较为明显,其结果也验证了本文提出的改进方法具有很好的泛化能力。

4.3 能耗估计

使用 Noisy Softplus 训练得到的 ANNs 权重可以直接作为与 ANNs 具有相同结构的 SNNs 的网络权重。在类脑硬件平台 Systems of Neuromorphic Adaptive Plastic Scalable Electronics(SyNAPSE)^[31-32]上运行 SNNs 时,突触每传递一个脉冲约消耗 0.37 pJ 的能量,本文所设计的 ResNet-19 网络连接数为 1.15×10^8 。设脉冲的发放率为 100 Hz,则与 ResNet-19 具有相同结构的 SNNs 在 SyNAPSE 上运行时,每秒钟消耗的能量约为:

$$0.37 \times 10^{-12} \times 1.15 \times 10^8 \times 100 \text{ J} \approx 0.0043 \text{ J} \quad (11)$$

在 Field-Programmable Gate Array (FPGA)^[33]上实现具有 4.35×10^8 个神经元连接的 CNNs 时,能耗约为 5 J/s^[34-35]。而 ResNet-19 的连接数约为 1.15×10^8 ,则可以估算出其在 FPGA 上的运行时间,进而计算出每秒钟的能耗约为:

$$5 \times \frac{1.15 \times 10^8}{4.35 \times 10^8} \text{ J} \approx 1.32 \text{ J} \quad (12)$$

估算出的能耗对比如表 3 所列。因此,将本文方法训练得到的权重作为 SNNs 的权重,构建得到的 SNNs 在 SyNAPSE 上运行时,识别精度与 ResNet-19 相当,但其能耗只有 ResNet-19 的 $\frac{0.0043}{1.32} \approx \frac{1}{307}$ 。

表 3 能耗对比

Table 3 Comparison of energy consumption

名称	硬件平台	每秒能耗/J
ResNet-19	FPGA	1.3200
SNNs	SyNAPSE	0.0043

结束语 本文采用具有生物真实性的 Noisy Softplus 激活函数代替 ReLU 函数,对所设计的深度残差网络进行训练,不仅提高了深度残差网络的表情识别精度,更重要的是,基于 Noisy Softplus 激活函数训练得到的深度残差网络的权重可以直接作为与深度残差网络具有相同结构的深度 SNNs 的网络权重,使得将该深度 SNNs 部署到类脑硬件上时,能以深度残差网络的精度和 SNNs 的功耗来实现表情识别。下一步将针对面部表情识别的特点,对 LIF 神经元本身的模型参数(如脉冲发放阈值、膜时间常量等)进行研究与分析,进一步给出优化后的 LIF 神经元模型参数。

参 考 文 献

- [1] SHAN C, GONG S, MCOWAN P W. Facial expression recognition based on Local Binary Patterns: A comprehensive study [J]. Image & Vision Computing, 2009, 27(6): 803-816.
- [2] HU Y, ZENG Z, YIN L, et al. Multi-view facial expression recognition[C]//Proceedings of IEEE International Conference on Automatic Face & Gesture Recognition. Amsterdam: IEEE Press, 2008: 1-6.
- [3] TARIQ U, LIN K H, LI Z, et al. Emotion recognition from an ensemble of features[C]//Proceedings of IEEE International Conference on Automatic Face & Gesture Recognition and Workshops. Santa Barbara: IEEE Press, 2012: 872-877.
- [4] HINTON G E, OSINDERO S, TEH Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. Neural Computation, 2014, 18(7): 1527-1554.
- [5] LIU M, LI S, SHAN S, et al. Deeply Learning Deformable Facial Action Parts Model for Dynamic Expression Analysis[C]//Proceedings of Asian Conference on Computer Vision. Singapore: Springer, 2014: 143-157.
- [6] KAHOU S E, PAL C, BOUTHILLIER X, et al. Combining modality specific deep neural networks for emotion recognition in video[C]//Proceedings of ACM on International Conference on Multimodal Interaction. Sydney: ACM Press, 2013: 543-550.
- [7] KIM B K, LEE H, ROH J, et al. Hierarchical Committee of Deep CNNs with Exponentially-Weighted Decision Fusion for Static Facial Expression Recognition[C]//Proceedings of ACM on International Conference on Multimodal Interaction. New York: ACM Press, 2015: 427-434.
- [8] SUN S, CHEN W, WANG L, et al. On the Depth of Deep Neural Networks: A Theoretical View[C]//Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. Phoenix: AAAI Press, 2015: 2066-2072.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE Press, 2014: 1-9.
- [10] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Press, 2015: 770-778.
- [11] DENG J, DONG W, SOCHER R, et al. ImageNet: A large scale hierarchical image database[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE Press, 2009: 248-255.

chines: Recurrent Neural Networks for Analyzing Relations in Group Activity Recognition[C]// Computer Vision and Pattern Recognition. IEEE, 2016; 4772-4781.

- [9] HE C Y, WANG P, ZHANG X H, et al. Detection of abnormal behavior in small and medium population based on intelligent surveillance [J]. Computer Application, 2016, 36 (6): 1724-1729. (in Chinese)

何传阳, 王平, 张晓华, 等. 基于智能监控的中小人群异常行为检测[J]. 计算机应用, 2016, 36(6): 1724-1729.

- [10] SONG D N, WANG P, ZHANG X H, et al. Detection of abnormal behavior in small and medium groups based on video surveillance [J]. Computer Engineering and Design, 2016, 37 (9): 2507-2514. (in Chinese)

宋丹妮, 王平, 张晓华, 等. 基于视频监控的中小群体异常行为检测[J]. 计算机工程与设计, 2016, 37(9): 2507-2514.

- [11] LIN J, LIN L. Detection method of group disturbance behavior based on optical flow modulus change frequency [J]. Computer Science, 2016, 43(5): 283-287. (in Chinese)

林杰, 林拉. 基于光流模值变化频率的群体骚乱行为检测方法[J]. 计算机科学, 2016, 43(5): 283-287.

- [12] KRYJAK T, GORGON M. Real-time implementation of the ViBe foreground object segmentation algorithm[C]// Computer Science and Information Systems. IEEE, 2013; 591-596.

- [13] CEBECI Z, YI Z, LDI F. Comparison of K-means and fuzzy C-means algorithms on different cluster structures[J]. Journal of Agricultural Informatics, 2015, 6(3): 13-23.

(上接第 307 页)

- [12] IZHIKEVICH E M. Simple model of spiking neurons[J]. IEEE Transactions on Neural Networks, 2004, 14(6): 1569-1572.

- [13] PAUGAMMOISY H, BOHTE S. Computing with Spiking Neuron Networks[M]. The Netherlands: Springer, 2012; 335-376.

- [14] LIU Q, FURBER S. Noisy Softplus: A Biology Inspired Activation Function[C]// International Conference on Neural Information Processing. Kyoto: Springer, 2016; 405-412.

- [15] MNIST database[OL]. <http://yann.lecun.com/exdb/mnist>.

- [16] DAYAN P, ABBOTT L F. Theoretical Neuroscience[M]. Cambridge: MIT Press, 2001; 180-191.

- [17] IZHIKEVICH E M. Which model to use for cortical spiking neurons[J]. IEEE Transactions on Neural Networks, 2004, 15(5): 1063-1070.

- [18] BRETTE R, GERSTNER W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity [J]. Journal of Neurophysiology, 2005, 94(5): 36-37.

- [19] LIU Y H, WANG X J. Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron[J]. Journal of Computational Neuroscience, 2001, 10(1): 25-45.

- [20] GLOROT X, BORDES A, BENGIO Y. Deep Sparse Rectifier Neural Networks[C]// Proceedings of the International Conference on Artificial Intelligence and Statistics. Fort Lauderdale: MIT press 2011; 315-323.

- [21] HE K M, ZHANG X, REN S, et al. Identity Mappings in Deep Residual Networks[C]// Proceedings of European Conference on Computer Vision. Amsterdam: Springer, 2016; 630-645.

- [22] LUCEY P, COHN J F, KANADE T, et al. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. San Francisco: IEEE Press, 2010; 94-101.

- [23] KDEF database[OL]. <http://www.emotionlab.se/resources/kdef>.

- [24] GENKI-4K database[OL]. <http://mplab.ucsd.edu>.

- [25] MOLLAHOSSEINI A, Chan D, MAHOOR M H. Going deeper in facial expression recognition using deep neural networks[C]// Proceedings of Winter Conference on Applications of Computer

Vision. New York: IEEE Press, 2016; 1-10.

- [26] LIEW C F, YAIRI T. A comparison study of feature spaces and classification methods for facial expression recognition [C] // Proceedings of International Conference on Robotics and Biomimetics. Shenzhen: IEEE Press, 2013; 1294-1299.

- [27] SANTRA B, MAUKHERJEE D P. Local Dominant Binary Patterns for Recognition of Multiview Facial Expressions[C]// Proceedings of the Asian Conference on Computer Vision. Guwahati: ACM Press, 2016; 1-25.

- [28] RUIZ A, ELSHAW M, ALTAHHAN A, et al. Stacked deep convolutional auto-encoders for emotion recognition from facial expressions[C]// Proceedings of International Joint Conference on Neural Networks. Anchorage: IEEE Press, 2017; 1586-1593.

- [29] CUI D, HUANG G B, LIU T. Smile detection using Pair-wise Distance Vector and Extreme Learning Machine[C]// Proceedings of International Joint Conference on Neural Networks. Vancouver: IEEE Press, 2016; 2298-2305.

- [30] ZHANG L, TJONDRONEGORO D, CHANGRAN V, et al. Towards robust automatic affective classification of images using facial expressions for practical applications [J]. Multimedia Tools & Applications, 2016, 75(8): 1-27.

- [31] CRUZ-ALBRECHT J M, YUNG M W, SRINIVASA N. Energy-Efficient Neuron, Synapse and STDP Integrated Circuits[J]. IEEE Transactions on Biomedical Circuits & Systems, 2012, 6(3): 246-256.

- [32] MEROLLA P, ARTHUR J, AKOPYAN F, et al. A digital neurosynaptic core using embedded crossbar memory with 45pJ per spike in 45nm [C] // IEEE Custom Integrated Circuits Conference. San Jose: IEEE Press, 2011; 1-4.

- [33] TRIMBERGER S M. Field-Programmable Gate Array Technology[M]. New York: Springer, 1994; 68-87.

- [34] CAO Y, CHEN Y, KHOSLA D. Spiking Deep Convolutional Neural Networks for Energy-Efficient Object Recognition[J]. International Journal of Computer Vision, 2015, 113(1): 54-66.

- [35] FARABET C, MARTINI B, AKSELROD P, et al. Hardware accelerated convolutional neural networks for synthetic vision systems[C]// Proceedings of IEEE International Symposium on Circuits and Systems. Paris: IEEE Press, 2010; 257-260.