

doi:10.3969/j.issn.1673-9833.2017.01.014

基于二值化卷积神经网络的手势分类方法研究

胡骏飞, 文志强, 谭海湖

(湖南工业大学 计算机学院, 湖南 株洲 412007)

摘要: 针对手势分类问题, 提出了一种基于二值化卷积神经网络的手势分类方法。根据神经网络在低精度化后仍能保持较高分类准确性和鲁棒性的特点, 结合传统高精度卷积网络手势分类方法与二值化方法提出一种网络结构。并通过实验研究了隐层参数对手势分类效果的影响, 并与常用的方法进行了分类性能和运行效率对比。实验结果表明, 所提出的方法在 $N=512$ 时的表现最佳, 与其他方法相比, 计算效率明显提升, 且错误率接近最好的结果。

关键词: 卷积神经网络; 深度学习; 二值化; 手势分类

中图分类号: TP391

文献标志码: A

文章编号: 1673-9833(2017)01-0075-06

Research on Hand Gesture Classification Based on Binary Convolutional Neural Networks

HU Junfei, WEN Zhiqiang, TAN Haihu

(School of Computer, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: A classification method based on binary convolutional neural networks has been proposed in view of some problems in current gesture classification. Based on the characteristics of neural networks, which can keep a relatively high degree of accuracy and robustness in classification even under a low precision, a proposal has been made of a new network structure with the traditional high-precision classification method of convolutional networks and the binary classification method combined together. In the process of the experiment, a research has been conducted on the effect of hidden layer parameters on the hand gesture classification, followed by a comparison between the classification performance and the operational efficiency of the conventional classification methods. The experimental results show that the proposed method has the best performance when $N=512$. Compared with other methods, its computational efficiency has been significantly improved, with its error rate close to the best result.

Keywords: convolutional neural networks; deep learning; binary; gesture classification

0 引言

随着时代变迁和科学技术发展, 尤其是虚拟现实、无人机等技术的发展, 人们对于人机交互的需求越来越多样化。在众多交互手段中, 作为人类除语言

外最为自然的交流方式之一的手势, 一定会获得更高的接受度。因此, 对手势识别的研究有着重要意义。

目前, 手势识别在手机交互、体感游戏操作等领域应用广泛。这些应用中通过使用距离检测装置(深度摄像头)提高识别精度, 达到了不错的识别效果。

收稿日期: 2016-12-23

基金项目: 国家自然科学基金资助项目(61170102)

作者简介: 胡骏飞(1990-), 男, 湖南邵阳人, 湖南工业大学硕士生, 主要研究方向为深度学习,

E-mail: hjunfei@foxmail.com

但同样制约了手势识别的更广泛应用，主要原因有：第一，市面上的摄像头不具备深度探测功能；第二，深度检测装置成本较高。因此，设计一种适应于普通摄像头的手势识别算法是存在市场需求。

近些年来，由于 Y. Lecun 等^[1]神经网络研究者的努力，卷积神经网络逐渐成为图像分类领域最热门的算法。一个重要原因是它的特征学习能力使得特征识别突破了需要人工设计的局限，使其具有了一定的通用性。但手势分类的算法通常会用在许多计算能力不高的设备上，卷积神经网络的大量计算消耗成为其更广泛应用的阻碍。因此，本文采用二值化近似策略对卷积网络手势分类方法进行优化，降低其计算消耗。采用二值化近似可有效减少权值的内存消耗，并简化计算中的大量高精度乘法运算，在对正确率影响不大的情况下，提高计算效率。本文的目的是使用二值化卷积神经网络 (binary convolution neural networks, BCNN) 对普通 webcam 摄像头手势识别系统中的分类模块进行优化，包括二值化卷积神经网络手势分类模型的建立和参数选择策略，以期获得更快的运行效率和更优的分类准确率。

1 相关工作

当前，手势分类算法主要有以下几种方式：对图像直接利用 k 近邻 (k -nearest neighbor, KNN) 进行手势模板匹配分类^[2]；利用人手的几何特征进行手势分类，通过凸缺陷检测分析指尖来分类手势；还有一种非常流行的方法是采用从图像中提取经过精心设计的通用特征，然后将这些特征输入分类器达到对手势进行分类识别的目的。这些算法的大量工作都放在了特征的选择和设计上，常用的特征包括：Haar 小波^[3]、尺度不变特征变换 (scale-invariant feature transform, SIFT)^[4]、方向梯度直方图 (histogram of oriented gradient, HOG)^[5]、hu 矩^[6]。

深度学习的兴起导致各分类识别问题都开始使用卷积神经网络 (convolution neural networks, CNN) 进行检测分类。相对之前的各种算法，卷积神经网络里输入最后部分的分类器中的特征均从数据中学习得到。早在 90 年代，卷积网络就已被应用到手势识别领域中^[7]。2011 年，J. Nagi 等^[8]在手势识别中采用了最大池化的卷积神经网络 (max pooling convolution neural networks, MPCNN) 方式对人机交互中手势进行分类识别，准确率达 96.87%。2014 年，该研究组研究了 CNN 与支持向量机 (support vector machine, SVM) 相结合的算法 CNNSVM^[9]，进一步

提升了正确率。常见的手势识别技术见表 1。

表 1 常见的手势识别技术

Table 1 Common types of gesture recognition

序号	缩写方法	特征提取方法	分类方式
1	KNN		KNN
2	HOG-SVM	HOG	SVM
3	MPCNN	CNN	NN
4	CNNSVM	CNN	SVM

对二值化神经网络的研究，源于著名的 Bengio 团队对卷积神经网络训练过程中权值精度变化影响的研究^[10]。他们发现在训练过程中对卷积权值的精度进行降低 (将浮点数改为定点数及动态定点数) 对网络的正确率影响非常小，但可显著降低程序的内存占用。随后的研究^[11]中，他们进一步将权值精度降为二值，依然取得了与最高正确率相近的结果。不仅如此，这种二值的连接方式是一种类似于 dropout (一种通过随机阻断传播，减轻网络过拟合问题的方法) 的效果，降低了噪声对网络的影响。此外，还有一些简化神经网络计算的方法被提出，如：boolean network^[12] 在训练中使用高精度权值，在测试中使用二值权值；V. P. Plagianakos 等^[13]采用 3-bit 的权值进行训练，也取得了不错的效果。

2 基于二值化卷积神经网络的手势分类

2.1 卷积神经网络的结构设计

一个卷积神经网络中包含许多操作层 (layer)。在网络的最前端，手势图像的像素点被拉成一个一维向量作为输入层 (input layer) 输入网络，之后的每一层都会将上一层的输出作为输入进行一些简单计算，然后把结果作为下一层的输入，直到数据流在最后一层输出结果，该过程称正向传播。每一层的计算都需一些参数，这些参数是控制输出结果的关键，卷积神经网络一般采用监督学习的方式对这些参数进行训练。网络在进行训练的过程中，通过计算输出层的代价函数进行反向传播，并在反向传播中对每一层的参数进行更新。

神经网络除了输入层和输出层，中间的层称作隐层 (hidden layer)，目前使用的卷积神经网络都拥有非常多的隐层。考虑到手势识别的样本复杂度，本文参考 MPCNN 网络架构结构设计。网络按照区块 (block) 进行划分，每个区块包含若干个操作层，网络区块分为卷积块、全连接块和 Softmax 块。

1) 卷积块。其主要对输入数据进行卷积，包括卷积层、BatchNormal 层、ReLU 层及最大池化层。

卷积层，用一个固定大小的卷积核对输入图像进

行卷积操作, 得到的卷积图像作为输出。卷积层是一种部分联通网络, 相比全联通网络需要的链接数要少, 每个输出值只与一个小区域链接, 且同一幅卷积图像的每个区域共享权值。这意味着卷积层需要学习的参数值只包含卷积核的参数及每个卷积核相应的偏置值。卷积层赋予了网络对图像的理解能力, 实际上卷积的操作是提取图像特定的特征, 人类视觉系统的处理过程也是对图像逐层提取特征逐步抽象。卷积层对参数的训练就是训练特定的特征提取器。

BatchNormal层, 该层对输入数据的各维度进行批量规范化, 加入该层可有效提升网络学习速率^[14]。

ReLU (rectified linear units) 层, 这是一个非线性激活单元, 通常作为激活函数用在全连接层之后。该层的表达式如式(1)所示, 即

$$L(X)=\max(X, 0), \quad (1)$$

式中: $L(X)$ 为网络层的函数;

X 为该层输入的数据张量。

此处是取 X 与 0 中的较大值。ReLU 层中不存在需要学习的参数。ReLU 是经常被用在现代神经网络中的激活函数, 相较于过去经常使用的 sigmoid 函数和 tanh 函数, 它有着以下优点: 第一, 相对于传统的 sigmoid 和 tanh, ReLU 的计算比较简单, 极大地减少了训练中的计算开销; 第二, ReLU 是一种单端饱和的激活函数, 可有效减少 sigmoid 这种双端饱和函数在梯度反向传播时造成的衰减现象, 使得网络能从这些在别的激活函数中被衰减掉的梯度中进行学习, 加快了训练的收敛速度; 第三, ReLU 作为一种稀疏性的激活函数, 能有效减少干扰特征被激活造成的影响。然而由于 ReLU 对于负数输入的梯度也为零, 会造成某些情况下无法继续学习。

最大池化层, 通过对每个 2×2 的区域取出一个值来减小输入图片的尺寸。最大池化层取值的策略是取出各个区域中 4 个像素值中的最大值, 最终输出图片的宽和高是输入图片的一半。最大池化层使得图像网络在向前传播的过程中抽象到更高的层级, 之后的卷积层可以在不同的尺度对图像进行特征提取。

2) 全连接块。这一个区块主要包含全连接层、BatchNormal层以及 ReLU 层。

全连接层, 输入数据会进行一次仿射变换, 也就是输入数据 X 与参数 w 进行点积运算再加上一个偏置 b 。一个有 n 个输入和 m 个输出的全连接层的表达式如式(2)所示:

$$L(X)=wX+b. \quad (2)$$

在全连接层中, 输出取决于参数矩阵 w 和偏置 b , 其中 $w \in \mathbb{R}^{(n \times m)}$, $b \in \mathbb{R}^m$, 这些参数在训练过程中需要学

习。全连接是神经网络中最基本的链接形式。

3) Softmax 块。放在最后的一个区块, 由 Softmax 层组成。

通过 Softmax 层计算出的数值可以输入代价函数优化网络, 也可以作为测试的结果输出。这一层同样不具备学习参数。Softmax 层输出数值代表着网络对输入对象属于哪种类型的概率评估, 因此这些数值相加之和为 1。该层的表达式如式(3)所示, 即

$$L(x)_i = \frac{e^{x_i}}{\sum_{k=1}^K e^{x_k}}, \quad (3)$$

式中: $L(x)_i$ 为输入对象属于第 i 类的概率;

x_i 为第 i 类的输入值;

K 为输入的总类数。

Softmax 是逻辑回归在多分类问题上的推广, 分类标签可以取得两个以上的值, 在手势识别中需要对两个以上的手势进行分类。

2.2 传统卷积网络的二值化方法

目前通行的深度卷积神经网络算法都有一个共同缺陷, 就是计算消耗巨大。主要体现在: 1) 权值占用大量内存, 这是由于权值采用双精度浮点数, 而一个深度神经网络往往拥有上百万个权值, 一个 vgg 网络(牛津大学 Visual Geometry Group 推出的网络, 简称为 vgg 网络)权值可以占到 1 GB, AlexNet 网络占用约 500 MB; 2) 计算中存在大量的乘法运算, 高精度的数乘运算对计算资源的消耗巨大。因此, 对网络计算消耗的优化也主要围绕这两个方面展开。

研究组在 MPCNN 手势分类方法的基础上, 提出一种基于二值化的卷积神经网络手势分类方法, 采用二值化近似的策略对神经网络进行改进, 降低其对计算资源的消耗。二值化网络降低计算资源消耗的方式主要有两个: 第一, 采用二值化近似的权值表示原有的双精度权值, 降低了网络在计算中的内存占用; 第二, 对每一层中计算消耗最大的乘法计算中的输入和权值采用二值化近似的值进行替代, 这样乘法计算就可以简化成加减法甚至位运算。

以卷积块为例, 对卷积神经网络进行二值化近似改造的具体方式如下。

第一步, 在正向传播过程中, 对卷积网络的权值矩阵 w 中每个元素按照式(4)进行二值化得到 w^b , 并保留原本的 w , 即

$$w^b = \text{sign}(w) = \begin{cases} +1, & w \geq 0; \\ -1, & w < 0. \end{cases} \quad (4)$$

式中: w 为原权值矩阵 w 中的权值, 其中 $w \in \mathbb{R}^{(c_i * w_i * h_i)}$;

w^b 代表二值化近似后得到矩阵 w^b 中的权值, 其中 $w^b \in \{+1, -1\}^{(c_f * w_f * h_f)}$, c_f, w_f, h_f 为卷积核的数量、宽度和高度。标准的 sign 函数中, 当 $w=0$ 时, 会取得 $\text{sign}(w)=0$, 而这里为了达到二值化的效果, 不允许有第 3 个值存在, 所以规定当 $w=0$ 时取 $\text{sign}(w)=1$ 。

第二步, 在每层的前一层之前增加一个二值化激活层得到节点值 X^b , 替代原本的 ReLU 激活层, 如式 (5) 所示, 即

$$X_i^b = L(X_{(i-1)}) = \text{sign}(X_{(i-1)}), \quad (5)$$

式中: X_i^b 为二值化网络的第 i 层的输入值; $L(X_{(i-1)})$ 为第 i 层二值化激活层得到的值, 其中 $X_i^b \in \{+1, -1\}^{(c * w * h)}$, c, w, h 分别为输入图像的通道数、宽度和高度。 sign 的函数与式 (5) 保持一致。

最后, 将得到的 X^b 与权值 w^b 在二值化卷积层进行卷积操作, 如式 (6) 所示, 即

$$L^b(X^b) = X^b \otimes w^b, \quad (6)$$

式中: $L^b(X^b)$ 为二值化的网络层函数;

\otimes 为卷积操作。

X^b, w^b 分别经过式 (4) 和式 (5) 得到。

对于卷积块, 其结构也需要一定的调整。将归一化处理 BatchNormal 层和二值化激活层放在卷积操作之前, 这个是为了防止二值化激活层的结果在经过最大池化层时出现结果大多数为 1 的情况。具体的网络改变对比如图 1 所示。

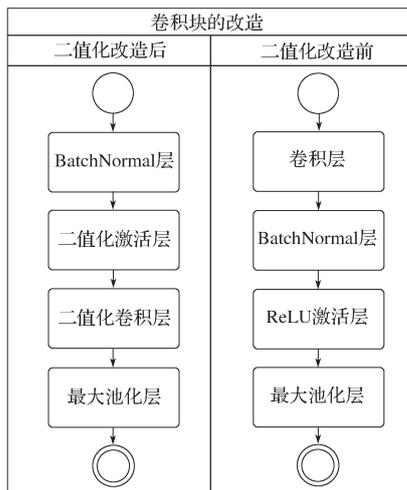


图 1 二值网络和非二值网络卷积块结构的区别

Fig. 1 Differences of the convolutional block structures between binary and non-binary networks

训练的反向传播过程如下: 最后一层算出梯度, 倒数第二层到第一层逐层反向传播算出节点的梯度以及权值的梯度, 再对二值化前保留的 w 进行更新得到 w^u 并进行如式 (7) 的 clip 操作, 即

$$\sigma(w^u) = \text{clip}\left(\frac{w^u + 1}{2}, 0, 1\right), \quad (7)$$

式中: w^u 为在正向传播过程中保留的浮点数值更新后的值;

clip 操作为将 $\frac{w^u + 1}{2}$ 的值在 0 和 1 处截断, 也就是把大于 1 的数量为 1, 把小于 0 的数量为 0。

3 实验结果及分析

3.1 实验环境与数据集

实验在 linux ubuntu14.04 环境下进行, 实验计算机配置采用英特尔 (Intel) i5 处理器、内存 8 G, 用于 CUDA 加速的 GPU 为 NVIDIA GTX950。图像采集以及预处理使用普通的 webcamera+opencv, 深度神经网络的建立训练在 theano 框架下进行。图 2 为实验采用的数据集。

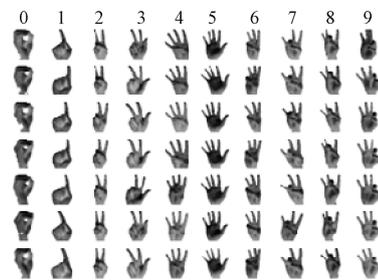
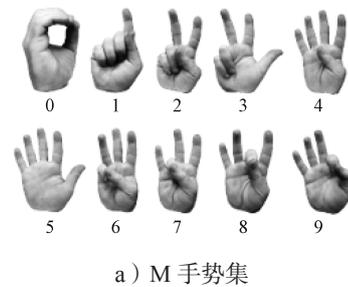


图 2 手势数据集

Fig. 2 Gesture database

实验采用的数据集由两部分组成, 一部分是根据美国标准手语 0~9 为标准使用 webcamera 采集的 10 种手势, 这里简称为 W 手势集。W 手势集在 webcamera 采集过程采取不同的光照以及不同的手势角度, 每种手势包含 2 200 个样本, 共 22 000 个样本。另一部分是 Massey 大学研究者制作的美国标准手语数据集^[15], 简称为 M 手势集。该数据集从 5 位志愿者的身上采集了包括代表数字 0~9 以及字母与 a~z 的手势总共 2 515 个手势样本, 每种手势取 65 个样本。

3.2 网络的训练与参数的选择

实验采用的网络结构由 2 个卷积块、2 个全连接

块和 1 个 Softmax 块组成。卷积块的设计相对固定, 所以参考文献 [8] 的设计, 核心大小均为 3×3 , 特征图的个数均设计为 20 个。全连接块 2 的节点数量由训练数据的类别决定, 全连接块 1 的节点数量 N 由实验中的测试结果决定。具体网络设置如表 2 所示。

表 2 卷积神经网络结构

Table 2 The architecture of the MPCNN

层数	类型	特征图大小	核心大小	滑动间隔
0	输入层	32×32		
1	卷积块 1	$28 \times 28 \times 20$	3×3	1
2	卷积块 2	$28 \times 28 \times 20$	3×3	1
3	全连接块 1	$1 \times 1 \times N$	1	1
4	全连接块 2	$1 \times 1 \times 10$	1	1
5	Softmax 块			

分别取 N 等于 64, 128, 300, 512 和 1 024 进行对比。由于 W 手势集和 M 手势集存在比较大的差异, 所以这里使用 W 数据集作为训练集, M 数据集作为测试集进行实验, 更能体现网络的泛化能力。网络的训练过程采用 Adam^[6] 的方式进行权值更新, 使用 dropout 进行规范化处理, 增强网络的鲁棒性。在不同节点数, 经过 200 epcho 的训练得到: 1) 取不同 N 条件下网络的收敛过程, 如图 3 所示; 2) 不同 N 条件下网络的错误率, 如图 4 所示。

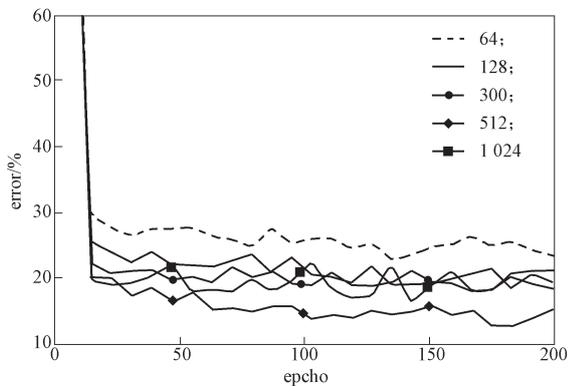


图 3 不同隐层节点数网络的收敛过程

Fig. 3 Convergence processes of the networks with different hidden nodes

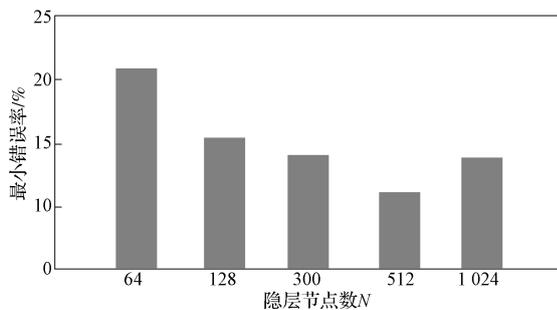


图 4 不同隐层节点网络的错误率对比

Fig. 4 A contrast between different error rates with different hidden nodes

从训练结果可看出 N 在取 512 时相对其他选择项收敛的结果更为稳定, 取得的正确率更高。因此, 取 512 作为手势分类网络的隐层节点数 N 的值。

3.3 实验比较

为了验证分类性能, 将提出的方法与以下 3 种分类方法进行对比: 1) 基于 k 近邻的手势分类方法^[2], 该方法直接使用图片的像素点作为输入, 分类器的参数取 $k=5$; 2) 采取手势的方向梯度直方图特征, 再使用 SVM 作为分类器 (HOG-SVM) 的手势分类方法^[5]; 3) 使用高精度的卷积神经网络的手势分类方法, 该方法采用与本文相同的网络架构, 即 2 个包含 20 个 3×3 卷积块链接、一个包含 512 个隐层节点的全连接块 (MPCNN)^[8]。提出方法采用的隐层节点数 $N=512$ (BCNN)。

对比实验共包含 3 组实验: 第一组, 在分别从 W 数据集每种手势中随机取出 2 000 个样本作为训练集, 另外的 200 个样本作为测试集; 第二组, 采用 M 数据集中的数字 0~9 的手势样本每组取 55 个作为训练集, 10 个作为测试集; 第三组, 使用 W 数据集作为训练集, M 数据集作为测试集。对 KNN、HOG-SVM、MPCNN 和提出方法 BCNN 分别进行实验, 实验的测试结果见表 3, 其中运行时间为处理 2 000 个样本的运行时间。

表 3 不同分类方法分类测试结果

Table 3 Test results of different classification algorithms

分类方法	错误率 / %			运行时间 / s
	第一组	第二组	第三组	
KNN	1.10	18.30	13.66	19.00
HOG-SVM	5.61	25.00	31.31	4.19
MPCNN	0.42	0.21	10.82	1.83
BCNN	0.98	0.45	10.71	1.13

通过实验结果可以看出, KNN 能从总体上比较完整地体现出数据集的空间分布正确率较高, 但是其计算效率低, 消耗内存较大, 尤其是在训练集比较大的情况下特别严重。HOG-SVM 虽然计算效率大幅提升, 但是正确率下滑严重。MPCNN 网络进行分类, 无论是正确率还是运行效率, 都要高于传统方法, 是一种十分优秀的分类模型。在 MPCNN 网络相同架构的情况下, BCNN 网络在第一组和第二组实验中的正确率仅有很小的降低, 在第三组实验中甚至还超过 MPCNN 网络, 且在运行效率上有了进一步提升。由于 theano 框架内在计算机机制会将二值化的权值也当做高精度进行计算, 其在高精度的神经网络和二值化的神经网络测试运行时间差保持在同一个数量级上, 加速效果还没有达到理论值, 因此下一步的研究将会放在前向传播的程序优化上。

4 结语

针对手势分类问题,对多种手势分类方法特别是卷积神经网络的方法进行了研究^[17],提出一种基于二值化卷积神经网络手势分类方法。对所提出方法的隐层节点数 N 在取不同参数情况下的分类准确率和训练收敛稳定性进行对比实验,结果表明该方法在 $N=512$ 的表现最佳。通过与其他 3 种分类算法的对比实验,证明了 MPCNN 比 KNN 和 HOG-SVM 模型具备更好的鲁棒性、准确度和运行效率,而二值化的卷积神经网络 BCNN 在分类准确性损失不大的情况下进一步提高了运行的效率。由于二值算法的计算效率高,可以将其运行在一些小微型设备上,用于开发手势识别系统。接下来的研究可以放在进一步提高算法的准确性以及对二值算法的移植上。

参考文献:

- [1] LECUN Y, KAVUKCUOGLU K, FARABET C. Convolutional Networks and Applications in Vision[C]// Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS). Paris: IEEE, 2010: 253-256.
- [2] STENGER B. Template-Based Hand Pose Recognition Using Multiple Cues[C]//Asian Conference on Computer Vision-ACCV 2006. Hyderabad: DBLP, 2006: 551-560.
- [3] 安涛,彭进业,吴静.基于 Haar 小波分解的实时手势识别[J].计算机工程,2011,37(24):138-140.
AN Tao, PENG Jinye, WU Jing. Real-Time Gesture Recognition Based on Haar Wavelet Decomposition[J]. Computer Engineering, 2011, 37(24): 138-140.
- [4] DARDAS N H, GEORGANAS N D. Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques[J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(11): 3592-3607.
- [5] 任彧,顾成成.基于 HOG 特征和 SVM 的手势识别[J].科技通报,2011,27(2):211-214.
REN Yu, GU Chengcheng. Hand Gesture Recognition Based on HOG Characters and SVM[J]. Bulletin of Science and Technology, 2011, 27(2): 211-214.
- [6] 翁汉良,战荫伟.基于视觉的多特征手势识别[J].计算机工程与科学,2012,34(2):123-127.
WENG Hanliang, ZHAN Yinwei. Vision-Based Hand Gesture Recognition with Multiple Cues[J]. Computer Engineering and Science, 2012, 34(2): 123-127.
- [7] NOWLAN S J, PLATT J C. A Convolutional Neural Network Hand Tracker[C]//Advances in Neural Information Processing Systems 8. Denver: NIPS, 1995: 901-908.
- [8] NAGI J, DUCATELLE F, CARO G A D, et al. Max-Pooling Convolutional Neural Networks for Vision-Based Hand Gesture Recognition[C]//2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). Kuala Lumpur: IEEE, 2011: 342-347.
- [9] NAGI J, CARO G A D, GIUSTI A, et al. Convolutional Neural Support Vector Machines: Hybrid Visual Pattern Classifiers for Multi-Robot Systems[C]//2012 11th International Conference on Machine Learning and Applications (ICMLA). Raton: IEEE, 2012: 27-32.
- [10] COURBARIAUX M, DAVID J P, BENGIO Y. Training Deep Neural Networks with Low Precision Multiplications[J/OL]. [2016-10-22]. <http://arxiv.org/abs/1412.7024>.
- [11] COURBARIAUX M, BENGIO Y, DAVID J P B. Binary Connect: Training Deep Neural Networks with Binary Weights During Propagations[J/OL]. [2015-11-12]. <http://arxiv.org/abs/1511.00363>.
- [12] KIM M, SMARAGDIS P. Bitwise Neural Networks[J/OL]. [2016-01-22]. <http://arxiv.org/abs/1601.06071>.
- [13] PLAGIANAKOS V P, VRAHATIS M N. Training Neural Networks with 3-Bit Integer Weights[C]// Proceedings of Genetic and Evolutionary Computation Conference (GECCO 1999). Orlando: IEEE, 1999: 910-915.
- [14] IOFFE S, SZEGEDY C. Batch Normalization Accelerating Deep Network Training by Reducing Internal Covariate Shift[J/OL]. [2016-12-02]. <http://arxiv.org/abs/1502.03167>.
- [15] BARCZAK A L C, REYES N H, ABASTILLAS M, et al. A New 2D Static Hand Gesture Colour Image Dataset for ASL Gestures[J]. Research Letters in the Information and Mathematical Sciences, 2011, 15: 12-20.
- [16] KINGMA D, BA J. Adam: A Method for Stochastic Optimization[J/OL]. [2016-12-20]. <http://arxiv.org/abs/1412.6980>.
- [17] 卢永乐,文志强,李建飞.基于改进模拟退火算法的 LUT 逆半调模板选择[J].湖南工业大学学报,2015,29(1):76-82.
LU Yongle, WEN Zhiqiang, LI Jianfei. Template Selection for LUT Inverse Halftoning Based on Improved Simulated Annealing Algorithm[J]. Journal of Hunan University of Technology, 2015, 29(1): 76-82.

(责任编辑:申剑)