

# 改进的卷积神经网络图片分类识别方法

闫河<sup>1,2</sup> 王鹏<sup>1\*</sup> 董莺艳<sup>1</sup> 罗成<sup>1</sup> 李焕<sup>1</sup>

<sup>1</sup>(重庆理工大学计算机科学与工程学院 重庆 401320)

<sup>2</sup>(重庆理工大学两江人工智能学院 重庆 401147)

**摘要** 在图像分类识别中,为了获得更高的分类精确度,需要对图片提取更精确和更能表现图片语义信息的特征,深度学习已成为特征提取最常用的方法。提出一种改进的深度卷积神经网络的图片分类模型。通过网络架构和内部结构两方面对经典的深度神经网络 AlexNet 的改进和优化,进一步提升特征的表达能力。通过在全连接层引入极限学习机,不仅提高了网络的分类能力和分类时间,而且使得该结构具有更优的数据处理能力。通过在两个标准数据集上的一系列对比实验,分析了不同的优化方法在不同情况下的作用,并证明了该网络结构的有效性。

**关键词** 卷积神经网络 极限学习机 图片识别 池化 特征提取

中图分类号 TP391.41

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2018.12.037

## IMAGE CLASSIFICATION AND IDENTIFICATION METHOD BASED ON IMPROVED CONVOLUTIONAL NEURAL NETWORK

Yan He<sup>1,2</sup> Wang Peng<sup>1\*</sup> Dong Yingyan<sup>1</sup> Luo Cheng<sup>1</sup> Li Huan<sup>1</sup>

<sup>1</sup>(College of Computer Science, Chongqing University of Technology, Chongqing 401320, China)

<sup>2</sup>(College of Liangjiang Artificial Intelligence, Chongqing University of Technology, Chongqing 401147, China)

**Abstract** In image classification and recognition, more accurate and better image semantic information features need to be extracted to obtain higher classification accuracy. Deep learning becomes the most common method for feature extraction. We proposed a picture classification model based on improved deep convolutional neural network in this paper. The classic deep neural network AlexNet was improved and optimized from two aspects: network architecture and internal structure, to further enhance the expressive ability of features. The extreme learning machine was introduced at the full connection layer, which gave the network a better classification capability and time, and made this structure more capable of data processing. Through comparison experiments on two standard data sets, it analyzes the effect of different optimization methods under different conditions and proves the effectiveness of the network structure.

**Keywords** Convolutional neural network Extreme learning machine Image recognition Pooling Feature extraction

## 0 引言

图像识别是通过计算机模拟人类对图片的分类理解,自动的把图片归为不同的语义类别。目前,图像识别在众多领域都有广泛的应用。对于图像识别目前有两个主要研究的方向:一是对特征的提取;二是对图片

分类算法的研究。目前基于支持向量机<sup>[1]</sup>的分类虽然取得较好的成绩,但是还存在挑战。由计算机表达出来的特征和人类对于大脑皮层所感知的信息还存在差距。这个问题也是我们面临的巨大挑战。如何用计算机构建一个对高层语义理解的方法是急需解决的关键问题。而随着卷积神经网络的提出,这个关键的问题也在逐渐的被解决。目前,深度学习技术在近几年取

得了突破性的进展,特别是在视觉方面。卷积神经网络在图像分类<sup>[2]</sup>、图像检索<sup>[3]</sup>、图片识别<sup>[4]</sup>等方面都取得较好的效果。文献[5]在1999年实现了对手写体数字的识别。AlexNet<sup>[2]</sup>网络在2012年的ImageNet大赛上取得了冠军。VGG<sup>[3]</sup>网络进一步提高了识别精度,同时把网络层数增加到19层,并在2014年的比赛中取得了第一名的好成绩。文献[4]提出了更深层的网络结构,把网络层数进一步提高,并达到了22层,同时获得了2014年视觉挑战赛的冠军。文献[6]研究了线性整流函数,通过和GoogleNet的对比,把准确率提高了26%。通过以上大赛中应用卷积神经网络的对比实验可以发现。神经网络的深度直接影响计算机视觉任务的效果。但是,如果单纯地增加网络深度又会引发网络梯度消失问题。如果发生梯度消失,不仅不会让网络的性能更好,还会让网络的精度下降,甚至消失。针对以上的问题,文献[7]提出了残差网络结构,不仅解决了网络深度的问题,同时还有效解决了梯度消失问题,文献[8]提出了一种全新的方案解决梯度消失问题,同时让层之间的信息保持最大化的传送。通过对以上网络结构的对比发现,在解决问题的时候都是向网络的深度发展。但是随着网络深度的加深,需要训练的参数也会逐渐增多,对于电脑的存储空间也提出了新的要求,同时也需要性能更好的计算设备。对于实际应用情况会出现种种困难,如今已有学者在研究如何减低计算开销<sup>[9-10]</sup>,目前有效的方法是用特定的数据在预训练网络上训练特定的目标网络,简单讲就是迁移学习。也有研究<sup>[11]</sup>用改进的网络结构计算类别间的相似度实现图片的分类。还有研究<sup>[12]</sup>对于在图片分类中有多个标签的情况。由此设计出了HCP网络,并取得了较好的效果,卷积神经网络提取特征时具有平移、缩放和旋转不变性<sup>[13]</sup>等优点。

卷积神经网络的最后一层一般可以看成是一个线性分类器,所以卷积神经网络虽然能够提取特征,但不是最优的分类器。2016年,赵亮等<sup>[14]</sup>利用卷积神经网络结合HOG和HSV特征构造图像特征,通过用支持向量机对图片进行分类。支持向量机(SVM)是最广泛的分类器之一,但随着分类器的进一步发展,Huang等<sup>[15]</sup>于2004年提出单隐层前向神经网络(SLFNs)的ELM算法,并于2006年对ELM进行了进一步的研究,给出了一些理论证明和应用<sup>[15-16]</sup>。文献[17]提出了P-ELM<sup>[17]</sup>,根据ELM算法并利用统计学原理,对网络分类效果影响较弱的隐层进行裁剪,并优化网络结构,取得了较好的分类效果。同时也有众多学者将极限学习机应用到了实际应用,并都取得了很

好的效果。张建明等<sup>[18]</sup>提出了极限学习机应用于人脸识别,文武等<sup>[19]</sup>提出了极限学习机应用于车牌识别等。与其他分类器相比较,极限学习机的参数不需要迭代训练,直接通过计算所得,所以具有训练速度快,不需要选择学习率等参数,避免了寻优过程,应用图片分类识别具有较好的效果。但是极限学习机对特征较高的要求,这正好运用卷积神经网络来弥补。以上文献充分证明了卷积神经网络对于特征提取的有效性,同时也证明了极限学习机在分类识别上比目前先进的分类算法表现得更优。但是,卷积神经网络是最好的特征提取方法,却不一定是最好的分类方法,且卷积神经网络的训练时间长,而极限学习机恰恰可以弥补卷积神经网络的不足。同时卷积神经网络对于特征提取还可以进一步优化和改进。

针对卷积神经网络分类耗时长,误差率有待提高和提取更丰富的特征语义,本文提出了一种AlexNet卷积神经网络和ELM结合的图片分类方法,采用有监督的方法学习不同层次的图像特征,用改进网络结构提取更具表达能力的图片特征,并在全连接层和极限学习机结合。相比于其他的算法本文具有以下的特点:

(1) 提出一种高效的网络框架,不仅能够提高图片识别率,同时也能降低计算开销。

(2) 本文在池化层做了改进,用最大-均值的池化方式,不但扩大了局部感受野,同时还保留了图片更加精确的特征,满足了极限学习机对于特征的要求。

(3) 通过全连层和极限学习机的结合,提高了网络结构的分类能力。

通过实验结果证明,本文对卷积神经网络提出的优化和改进,应用于图片的分类识别取得了较好的效果,且性能优于目前最先进的方法。

## 1 卷积网络和极限学习机

### 1.1 卷积神经网络

卷积神经网络<sup>[20]</sup>是提取高级语义特征表现最好的算法,通过对图片的逐层卷积提取更加抽象的特征,越抽象的特征越能够表现图片的语义信息,并在图片的分类中变现出更好的性能。

AlexNet是一种常见的卷积神经网络结构,此网络由五层卷积层、三层池化层和三层全连接组成。网络结构的全连接层和池化层的主要目的是对特征的提取,全连接层把特征向量压缩成为一维的向量。此网络的训练的过程主要分为两个步骤。

(1) 前向传播阶段,计算公式如下:

$$y^{(l)} = f\left(\sum_{i \in m} W_i^{(l)} \otimes x_i^{(l-1)} + b^{(l)}\right) \quad (1)$$

式中: $y^{(l)}$ 为第  $l$  个卷积层的输出, $x^{(i)}$ 是输入向量, $\otimes$ 为卷积运算, $b^{(i)}$ 为偏置, $w^{(i)}$ 为该层的权值, $m$ 代表输入特征的集合, $f(x)$ 为激活函数。

(2) 反向传播阶段,对于有  $m$  个样本的数据集,会通过上述的前向传播输出预测值,根据预测值与网络期望的输出定义整体的目标函数。

$$E(W) = \min \sum_{i=1}^M L(Z_i) + \lambda \|w\|^2 \quad (2)$$

式中: $L(Z_i)$ 是对应网络的损失函数。通过最小化损失函数降低网络分类误差, $z_i$ 为反向传播的输入, $w$ 代表权值, $\lambda$ 代表归一项所代表的比重, $L(Z_i)$ 需要根据实际的分类情况具体决定。

### 1.2 极限学习机结构

极限学习机最初是为了解决单隐层前馈神经网络 SLFNs(Single-hidden layer feed-forward neural networks) 而提出,网络结构图如图 1 所示。

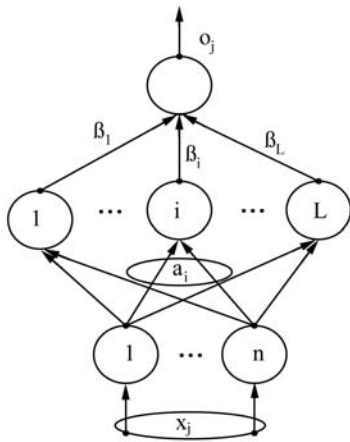


图 1 极限学习机的网络结构

假设有  $N$  个任意的不同样本  $(x_i, t_i), i = 1, 2, \dots, N$  其中,  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}] \in R^n, t_i = [t_{i1}, t_{i2}, \dots, t_{im}] \in R^m$ 。

则极限学习机的目标函数为:

$$H\beta = T \quad (3)$$

式中: $H$ 为神经网络隐层的输出矩阵; $T$ 为期望输出<sup>[12]</sup>。

$$H = \begin{bmatrix} h(x_1) \\ h(x_2) \\ \vdots \\ h(x_N) \end{bmatrix} \quad T = [t_1, t_2, \dots, t_N]^T \quad (4)$$

$$h(x) = g(W \cdot x + b) \quad (5)$$

式(4)和式(5)中: $W \in R^{L \times n}$ 为输入层到隐含层的权值; $b \in R^L$ 为隐含偏置量; $g(\cdot)$ 为激活函数。基本思想

是:随机产生  $W$  和  $b$ ,然后据此计算  $H$  和  $\beta$ ,有很多种方法,其中一种简单的方法为<sup>[21]</sup>:

$$\beta = H^+ T \quad (6)$$

## 2 基于深度神经网络的图像分类

### 2.1 AlexNet 网络框架的优化

AlexNet 网络结构如图 2 所示,此模型通过输入相应的图像和标签,输出图片类别的概率。本文将从两个方面优化和改进网络结构,首先对网络框架的优化,主要有两点:(1) 基于最大值池化的缺点,提出了最大值和均值结合的池化方法,扩大局部感受野。(2) 在全连接层后面接上极限学习机进行分类学习。其次,对于网络内部的优化主要是在卷积后,对数据规范化处理。

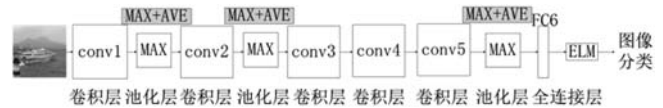


图 2 AlexNet 网络结构图及优化

### 2.2 最大值和均值结合的图像局部感受野

目前,深度学习中大部分的网络结构和识别的算法都用到了池化。为了能够提取局部感受野中精确的图像特征,本文研究了最大值池化和均值池化的特点,发现最大值池化可以提取图片局部的特征信息,通过局部的特征信息可以组合成为全局的特征信息。但是在特征的细节处理上效果表现欠佳,因而本文提出一种把两者结合的池化方式,既保留了轮廓的完整性,又在细节的处理上更精确。通过结合两者的优点,提取出更加精确的特征信息。

均值池化对细节的特征表现出更好的鲁棒性,而最大值池化则表现出平移不变形。通过公式<sup>[22]</sup>定义两个池化:

$$f(v) = \frac{1}{T} \sum_m^T v_m \quad (7)$$

$$f(v) = \max_{1 \leq m \leq T} v_m \quad (8)$$

式中: $v_m$ 表示图像中  $T$  个像素点的第  $m$  个像素点, $m$  表示该点在滑动窗口中的方位,池化即把  $v_m$  映射到相应的统计值。

本文结合了两者的优势,重新提出了最大-平均值池化:

$$f(v) = \left(\frac{1}{T} \sum_{m=1}^T v_m\right) + \max_{1 \leq m \leq T} v_m \quad (9)$$

在 AlexNet 网络结构中用于替换原有的方法,如图 2 灰色所表示的方法。

### 3 实验结果与分析

本文用 TensorFlow 开源框架实现图 2 所示的经过改进和优化的网络结构。实验室配置有 Intel® Core™ i7-4300U CPU, 主频 4.00 GHz, 内存 16 GB, Nvidia GTX 1080, 64 GB 内存的硬件平台上进行试验。

本文验证实验在两个公共的标准数据集: CIFAR-100<sup>[23]</sup> 和 ImageNet2017<sup>[24]</sup> 上进行。每个数据集都进行了如下两种神经网络优化方法的对比实验: (1) 在池化层中用最大值-均值池化替代最大值池化; (2) 在全连接层连接极限学习机即如图 2 所示的网络结构, 也是本文提出的深度卷积神经网络最重要优化方法。

定义分类性能的指标如下:

$$\text{误差率} = \frac{\text{被错误分类的图像}}{\text{被分类的图片总数}} \times 100\% \quad (10)$$

#### 3.1 数据集

CIFAR-100<sup>[23]</sup> 数据集包含了 100 类, 所有图像都是 32 × 32 的像素大小, 每一类图片包括 600 张彩色图片, 分别为 500 张训练图片和 100 张测试图片。

ImageNet<sup>[24]</sup> 数据集内有超过 1 500 万张、共 22 000 多个带标签的高分辨率图像。本实验采用 ILSVRC 数据集, ILSVRC-2017 是 ImageNet 的一部分, 选用其中的 1 000 个类别做对比实验。

#### 3.2 CIFAR-100 数据集实验结果分析

CIFAR-100 数据集是用于图片分类识别的数据集, 选择 100 张图片作为需要识别的图片。网络的输出定义为 100 类, 设定 batchsize 为 64, 初始学习率设定为 0.001, 同时用梯度下降的方法训练数据, 通过将数据分别在 Stochastic Pooling<sup>[25]</sup>、AlexNet + Fine - tuning<sup>[26]</sup>、AlexNet + Max - Ave - pooling<sup>[27]</sup> 等算法上做对比实验。对比结果如表 1 所示。

表 1 误差率在 CIFAR-100 数据集的比较结果

分类方法	误差率/%
Stochastic Pooling <sup>[25]</sup>	42
AlexNet + Fine - tuning <sup>[26]</sup>	32
AlexNet + FC + Max - Ave - pooling <sup>[27]</sup>	29
AlexNet + Max - Ave - pooling + ELM	26

从实验中可以得出如下结果, 不同的池化方法对于误差率有较明显的影响, 随机池化方式对特征的全局和细节提取不充分, 实验结果表现为误差率较高。同时, 通过 AlexNet + Fine - tuning<sup>[26]</sup> 调整规则后的池

化方式能够降低识别误差率, 相比于本文提出的最大值均值池化不仅考虑了特征的整体性, 同时兼顾了特征的细节。通过与极限学习机的结合, 并通过与目前性能最好的算法作对比, 实验分析可知, 最大值-均值池化解决了极限学习机对于特征的要求, 相对于噪声和其他干扰有更好的适应能力, 同时能够提取语义更加丰富和特征更精确的特征信息, 使得最终的误识率降到 26.67%, 充分证明了算法的有效性。

本文提出的最大值-均值和极限学习机结合的方法在误差率上明显降低。通过对比单独使用最大值池化的方法, 本文提出的方法能提取更加抽象的语义特征信息, 更有利于图片分类。相比于其他算法有效地降低了误差率, 通过与其他算法的对比实验, 在 CIFAR-100 数据集误差率降低到 26.67%, 证明了本文提出的方法在性能上明显优于其他算法。

#### 3.3 ImageNet 数据集结果分析

ImageNet 数据集内有超过 1 500 万张、共 22 000 多个带标签的高分辨率图像。本文选用 ILSVRC 数据集 (ImageNet Large-Scale Visual Recognition Challenge), ILSVRC-2017 是 ImageNet 的一部分, 选用其中的 1 000 个类别做对比实验, 网络的输出定义为 1 000 类, 设定 batchsize 为 64, 初始学习率设定为 0.001。通过将数据分别在 Inception-v4<sup>[28]</sup>、ResNet-200<sup>[29]</sup>、DenseNet-161<sup>[30]</sup>、SeNET<sup>[31]</sup> 和本文提出的算法作对比实验, 结果如表 2 所示。

表 2 误识率在 ILSVRC 上面的误差率的比较结果

分类方法	(top-1) 误差率/%	(top-5) 误差率/%
Inception-v4 <sup>[28]</sup>	21	5
ResNet-200 <sup>[29]</sup>	20	4
DenseNet-161 <sup>[30]</sup>	22	5
SENet <sup>[31]</sup>	17	3
AlexNet + Max - Ave - pooling + ELM	16	3

通过对比实验可以得出结论, 本文提出最大值-均值池化和极限学习机结合的方法, 改进了 AlexNet 深度卷积网络的结构, 在提取特征时运用均值和最大值结合的池化方式提取特征, 替代了原有的最大值池化方式, 有效地融合了特征提取方法对微小形变不变形和平移不变形的优势, 在特征提取中保留了更加重要的图片特征, 更加有利于极限学习机最后的分类识别。同时, 极限学习机在分类识别过程中通过广义逆计算所得, 在分类识别时耗时较短, 因而和目前最优的算法做了实验对比, 本文提出的算法有更好的性能表现。

### 3.4 极限学习机隐层神经元的个数对图像分类结果的影响

为了验证极限学习机隐层神经元的个数对图像分类的影响,分别在以上两个公共数据集上做了不同神经元个数的对比实验。第一个公共的数据集 CIFAR-100 共有 100 类,为了能够匹配网络,选用的神经元的个数分别为  $n = (80, 90, 100, 125, 130)$ 。第二个公共数据集 ILSVRC-2017 共有 1 000 类,为了更好地匹配网络,选用神经元的个数为  $n = (800, 900, 1 000, 1 250, 1 300)$ 。通过在本文提出的方法 AlexNet + Max - Ave - pooling + ELM 上进行实验。

从表 3 和表 4 可以得出极限学习机误差率会随着神经元的个数的变化而改变。当极限学习机隐层的神经元个数越接近分类个数时,错误分类率越低。由于极限学习机是通过求广义逆的方法直接计算出权值,当隐层神经元的个数越接近分类的目标时,误差率越小,同时通过计算得到的分类效果也越好。当我们在 CIFAR100 做实验时,由于分类的类别是 100,因而误识率在隐含层神经元接近 100 时效果达到最优。当在 ILSVRC-2017 做实验时,由于分类的类别是 1 000,当隐层神经元达到 1 000,ILSVRC-2017 数据集上表现的性能达到最优。

表 3 错误分类率在不用隐层神经元上的比较结果 %

数据集	80	90	100	125	130
CIFAR	66	43	24	42	48

表 4 不用隐层神经元上的错误分类率 (top-1) %

数据集	800	900	1 000	1 250	1 300
ILSVRC-2017	32	26	16	27	34

### 3.5 极限学习机对计算开销的提升

为了验证改进后的网络结构对计算开销的影响,通过本文算法和最新的深度学习算法做对比实验,并选用 ILSVRC-2017 作为数据集。Alexnet 深度神经网络结构在池化层后连接全连接层,并把所有特征信息延展为 4 096 维的特征向量,并通过 Softmax 分类器对图片分类。此结构在全连接层中需更新迭代的参数较多,迭代更新耗时长,且训练之后模型参数内存需 218 MB。改进后的模型一次前向和反向传播速率如表 5 所示,同时改进后的模型参数内存需要 21 MB,缩小为原 AlexNet 模型的 0.1 倍。通过对网络结构一次前向传播和反向传播耗时作为对比标准,比较不同算法的耗时比较,结果如表 5 所示。由于极限学习机对于特征分类采用求广义逆的计算方式,相比于传统的卷积

神经网络反向迭代更新权值,节约了大量的计算耗时,一次前向传播和反向传播总耗时降低到 54 ms,达到了优化的目的。通过实验结果表明了本文改进的网络结构有效地降低了计算开销。

表 5 一次前向传播和反向传播总耗时结果比较

算法	耗时/ms
Stochastic Pooling <sup>[25]</sup>	123
AlexNet + Fine - tuning <sup>[26]</sup>	124
AlexNet + Max - Ave - pooling <sup>[27]</sup>	126
Inception-v4 <sup>[28]</sup>	148
ResNet-200 <sup>[29]</sup>	164
DenseNet-161 <sup>[30]</sup>	165
SENet <sup>[31]</sup>	168
AlexNet + Max - Ave - pooling + ELM	54

## 4 结 语

本文提出了基于 AlexNet 深度卷积神经网络的分类框架,通过改进网络结构,提高特征的表达能力,进一步提高识别精度。本文尝试单独使用最大值池化和运用常规的分类方法作为对比方法,通过在标准数据集上的对比实验,证明本文提出的方法在准确率和计算开销上优于其他算法。未来的工作是进一步改进网络的结构,提取更具有区分度的特征和进一步改善计算的耗时,进一步提高识别率并应用于图片识别等具体应用领域。

## 参 考 文 献

- [1] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121 - 167.
- [2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097 - 1105.
- [3] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. 779 - 788.
- [4] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015. 1 - 9.
- [5] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learn-

- ing applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [6] He K, Zhang X, Ren S, et al. Delving deep into rectifiers; surpassing human-level performance on imagenet classification[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. IEEE, 2015:1026-1034.
- [7] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016:770-778.
- [8] Huang G, Liu Z, Maaten L V D, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017:2261-2269.
- [9] Wang J, Kumar S, Chang S F. Semi-supervised hashing for large-scale search[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(12):2393.
- [10] Lei J B, Caruana R. Do deep nets really need to be deep? [C]//International Conference on Neural Information Processing Systems. MIT Press, 2014:2654-2662.
- [11] Qu Y, Lin L, Shen F, et al. Joint hierarchical category structure learning and large-scale image classification[J]. IEEE Transactions on Image Processing, 2017, 26(9):4331-4346.
- [12] Wei Y, Xia W, Lin M, et al. HCP: a flexible CNN framework for multi-label image classification[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(9):1901-1907.
- [13] Arel I, Rose D C, Karnowski T P. Research frontier: deep machine learning—a new frontier in artificial intelligence research[J]. IEEE Computational Intelligence Magazine, 2010, 5(4):13-18.
- [14] 赵亮, 王晓峰, 袁逸涛. 基于深度卷积神经网络的船舶识别方法研究[J]. 舰船科学技术, 2016, 38(15):119-123.
- [15] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks [C]//Proceedings of the International Joint Conference on Neural Net works (IJCNN 2004). Budapest, Hungary, July 2004:985-990.
- [16] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70(1):489-501.
- [17] Rong H J, Ong Y S, Tan A H, et al. A fast pruned-extreme learning machine for classification problem [J]. Neural computing, 2008, 72(1/3):359-366.
- [18] 张建明, 刘阳春, 吴宏林. 基于极限学习机与子空间追踪的人脸识别算法[J]. 计算机工程, 2016, 42(1):168-173.
- [19] 文武, 乔龙辉, 何鹏. 基于自适应差分进化极限学习机的车牌识别算法[J]. 电子技术应用, 2017, 43(1):133-136, 140.
- [20] 白琮, 黄玲, 陈佳楠, 等. 面向大规模图像分类的深度卷积神经网络优化[J]. 软件学报, 2018, 29(4):1029-1038.
- [21] 余丹, 吴小俊. 一种卷积神经网络和极限学习机相结合的人脸识别方法[J]. 数据采集与处理, 2016, 31(5):996-1003.
- [22] Wang H, Cai Y, Zhang Y, et al. Deep learning for image retrieval: what works and what doesn't[C]//IEEE International Conference on Data Mining Workshop. IEEE Computer Society, 2015:1576-1583.
- [23] Krizhevsky A. Learning multiple layers of features from tiny images[R/OL]. Computer Science Department, University of Toronto, 2009. <https://www.cs.toronto.edu/~kriz/cifar-10-binary.tar.gz>.
- [24] Zeiler M D, Fergus R. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks[C]//Proceedings of the International Conference on Learning Representation. 2013.
- [25] Snoek J, Larochelle H, Adams R P. Practical Bayesian optimization of machine learning algorithms [C]//International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:2951-2959.
- [26] Goodfellow I J, Wardefarley D, Mirza M, et al. Maxout networks[C]//Proceedings of the 30th International Conference on Machine Learning. 2013:1319-1327.
- [27] Yang H F, Lin K, Chen C S. Supervised learning of semantics-preserving hash via deep convolutional neural networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(2):437-451.
- [28] Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database [C]//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009:248-255.
- [29] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[C]//ICLR Workshop, 2016.
- [30] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks [C]//European Conference on Computer Vision. Springer, Cham, 2016:630-645.
- [31] Huang G, Liu Z, Maaten L V D, et al. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017:2261-2269.
- [32] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[EB]. eprint arXiv:1709.01507, 2017.