



基于深度多特征融合的自适应CNN图像分类算法

李 伟¹, 黄鹤鸣¹, 武风英², 张会云¹

(1. 青海师范大学 计算机学院, 西宁 810008; 2. 青海省基础测绘院, 西宁 810001)

摘 要: 为更好地提取图像内容信息, 提高图像分类精度, 提出一种自适应卷积神经网络(CNN)图像分类算法。通过融合图像的主颜色特征, 利用CNN提取空间位置特征, 且针对多特征融合权重值的设定问题, 运用改进的差分演化算法优化各特征权重, 提高固定权重分类精确度。实验结果表明, 该算法分类精度相比CNN算法提升了9.2个百分点, 在图像分类中具有较好的分类效果。

关键词: 卷积神经网络; 自适应权重; 数据融合; 差分演化算法; 图像分类

开放科学(资源服务)标志码(OSID):



中文引用格式: 李伟, 黄鹤鸣, 武风英, 等. 基于深度多特征融合的自适应CNN图像分类算法[J]. 计算机工程, 2021, 47(9): 235-239, 251.

英文引用格式: LI W, HUANG H M, WU F Y, et al. Adaptive CNN-based image classification algorithm based on deep fusion of multi-feature[J]. Computer Engineering, 2021, 47(9): 235-239, 251.

Adaptive CNN-based Image Classification Algorithm Based on Deep Fusion of Multi-Feature

LI Wei¹, HUANG Heming¹, WU Fengying², ZHANG Huiyun¹

(1. School of Computer Science and Technology, Qinghai Normal University, Xining 810008, China;
2. Qinghai Basic Surveying and Mapping Institute, Xining 810001, China)

[Abstract] To improve the accuracy of image information extraction and image classification, an adaptive Convolutional Neural Network (CNN)-based algorithm is proposed for image classification. The algorithm effectively integrates the main color features of the image, and the spatial position features are extracted by using CNN. For the setting of the weight value of multi-features fusion, an improved differential evolution algorithm is presented to optimize the feature weight, so the accuracy of classification using fixed weight is improved. Experimental results show that the algorithm provides excellent image classification results, and its classification accuracy is 9.2% higher than that of CNN.

[Key words] Convolutional Neural Network (CNN); adaptive weight; data fusion; differential evolution algorithm; image classification

DOI: 10.19678/j.issn.1000-3428.0058387

0 概述

随着图像采集技术和处理技术的飞速发展, 图像已成为人们获取信息的主要方式。人们从浩瀚的图像数据库中搜索图像时, 需要对数据库中的图像进行分析处理, 而分析处理的前提是对这些海量图像进行分类。

随着大数据技术和硬件计算能力的发展, 神经网络作为一种新兴的深度学习的方法, 在各个领域取得了突破性进展^[1]。其中, 卷积神经网络(Convolutional

Neural Network, CNN)作为一种带有卷积计算的深层前馈型神经网络^[2-4], 能更好地获取图像的空间位置和形状信息, 有利于图像分类。2005年, STEINKRAUS等^[5]提出基于GPU的高效CNN训练方法, 大幅提高了CNN的运算能力; 2012年, KRIZHEVSKY等^[6]提出AlexNet网络, 采用ReLU激活函数并使用GPU分组卷积方式进行并行训练^[5]; GoogLeNet^[7]由SZEGEDY等于2014年提出; 2015年, LOFFE等^[8]将批标准化应用于神经网络, 保证了网络中各层的输出分布基本稳定, 大幅降低了网络对

基金项目: 国家自然科学基金(61462072, 61662062); 青海省自然科学基金(2016-ZJ-904)。

作者简介: 李 伟(1986—), 女, 博士研究生, 主研方向为图像处理、模式识别、智能系统; 黄鹤鸣, 教授、博士; 武风英, 工程师; 张会云, 博士研究生。

收稿日期: 2020-05-21 **修回日期:** 2020-07-29 **E-mail:** 980779950@qq.com

初始参数的依赖,提升了网络性能。

CNN具有良好的扩展性和鲁棒性,利用CNN进行图像分类时,能够快速且全面地提取空间位置和形状等相关特征,但是CNN对颜色特征不敏感。在进行图像分类时,CNN通过RGB三颜色通道读取图像内容进行分类的效果与它读取二值化图像信息进行分类的效果差别不大或者无差别。由于颜色特征具有很强的鲁棒性和灵活性,但不能体现图像空间分布的特点。为此,本文提出将主颜色特征与空间位置和形状相关特征进行深度融合的混合特征提取方法,并运用改进的差分演化算法优化多种特征之间的权值,以解决固定权值不能充分考虑图像自身特性的问题。

1 深度多特征融合的CNN网络模型

为更好地读取图像内容信息,本文构建基于深度多特征融合的CNN网络模型总体框架。该框架共有4个卷积层:第1个卷积层称为数据特征融合层,可以充分融合去噪深度卷积特征和主颜色特征,随后是一个池化层,池化层的作用是降低网络参数,加快融合速度;第2个卷积层称为深度特征融合层,进一步进行特征融合,同第一个卷积层,随后也是一个池化层;第3个卷积层称为特征抽象表示层,该层中的卷积核大小由前2个卷积层中的 5×5 改为 3×3 ,即有助于消除特征中的噪声并提高特征的抽象表示,随后也是一个池化层;第4个卷积层称为特征高级表示层,有助于消除冗余特征并提高代表性特征,后面是一个池化层,随后是三层全连接层,用于特征分类以及反向传播过程中的参数优化。在网络的最后,采用Softmax层进行分类。Softmax是应对多分类问题的有监督学习方法^[9-11],为分类提供重要的置信度。

卷积层、非线性激活变换和池化层是CNN的3个基本组成部分^[2]。通过叠加多个具有非线性运算的卷积层和池化层,可以形成一个深度CNN,而分层提取的输入特征具有不变性和鲁棒性^[12]。具有非线性运算的卷积层^[13]如式(1)所示:

$$x_j^l = f\left(\sum_{i=1}^M x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (1)$$

其中:矩阵 x_i^{l-1} 是 $l-1$ 层的第 i 个特征图; x_j^l 是当前 l 层的第 j 个特征图; M 是输入特征图的数量; k_{ij}^l 和 b_j^l 随机初始化并分别设置为零,然后通过反向传播进行精确微调; $f(\cdot)$ 是一个非线性激活函数; $*$ 是卷积运算。使用特定的体系结构,如局部连接和共享权重,CNN往往具有良好的泛化功能。

基于深度多特征融合的CNN网络模型总体框架如图1所示。图1中的去噪深度卷积特征称为 f_c ,是经过3层卷积神经网络以后的去噪空间位置和形状关系信息,如图2所示。

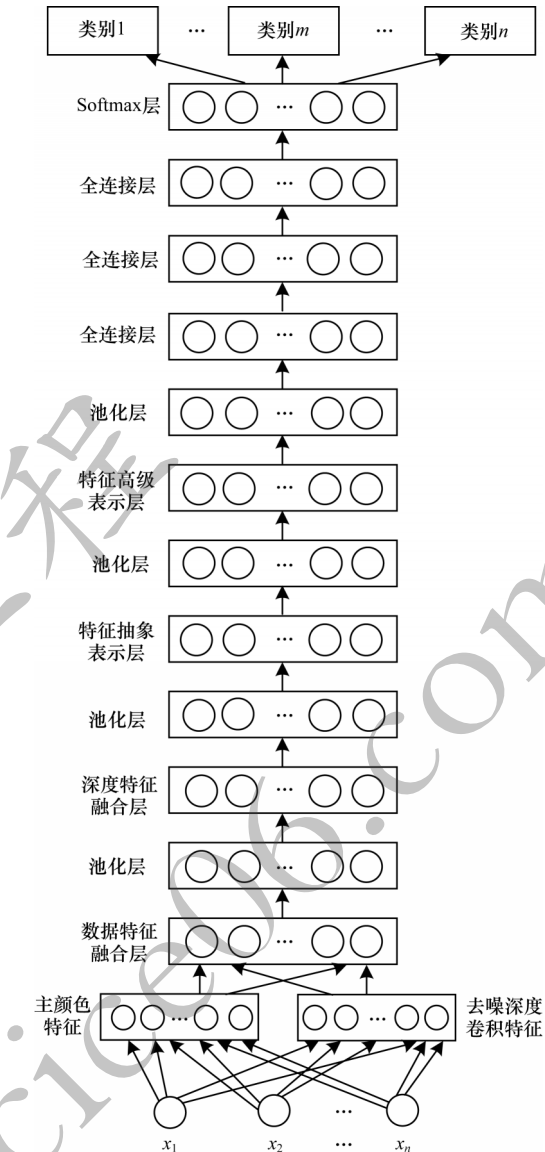


图1 基于深度多特征融合的CNN网络模型总体框架

Fig.1 Overall framework of CNN networks model based on deep fusion of multi-features

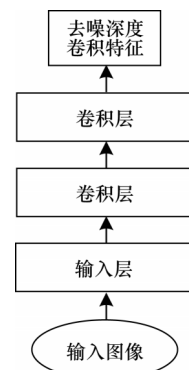


图2 去噪卷积神经网络结构

Fig.2 Structure of denoising convolutional neural networks

CNN的卷积层内部包含多个卷积核,每个卷积核能够提取空间位置和形状相关特征,卷积层内每

个神经元均与前一位置邻近区域,也被称为“感受野”^[3]的多个神经元相连,从而依赖于网络分层学习上下文不变特征^[13];形状和空间位置特征信息,这对于图像分类特别有用;另外,CNN还具有对输入图像特征的去噪功能。综合CNN的以上特点,图2所示去噪卷积神经网络输出的是去噪深度卷积特征。

2 自适应CNN图像分类算法

2.1 主颜色特征提取

不同于k-means算法随机选取数据集中的 k 个点作为聚类中心,“硬聚类”算法k-means++^[14]随机选取数据集中的点作为聚类中心,假设已经选取了 n 个初始聚类中心($0 < n < k$),则距离当前 n 个聚类中心越远的点被选为第 $n+1$ 个聚类中心的概率越高。

2.2 深度多特征融合算法

按照式(2)级联去噪卷积神经网络输出的去噪深度卷积特征 f_c 和使用k-means++聚类算法提取的主颜色特征 f_i ,得到深度数据融合特征 f_m :

$$f_m = \alpha f_i + \beta f_c \quad (2)$$

由于特征维度比较高且训练样本有限,可能会产生过拟合现象。本文研究使用Dropout方法解决过拟合问题。Dropout方法在不同的训练阶段,通过随机丢弃不同的神经元形成不同的神经网络,从而有效防止复杂的共适应,学习到更加正确的特征。

为增强网络的鲁棒性,本文研究还使用了ReLU激活函数^[15]。如果神经元的输出为正,则该函数接受该神经元的输出;如果输出为负,则返回0。ReLU函数的表达式如式(3)所示:

$$h(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3)$$

通过使用Dropout方法和ReLU激活函数,使多数神经元的输出变为0,再利用多层ReLU和Dropout实现基于稀疏的深度网络正则化,从而解决深度多特征融合图像分类的过拟合问题。

2.3 自适应权值图像分类算法

融合多特征进行图像分类时,为使分类结果尽可能更优,引入改进的差分演化算法寻找最优特征权重值的方法。差分演化算法^[16-18]是一种高效的全局优化算法,具有结构简单、易实现、优化能力强等特点。

融合主颜色与空间位置和形状相关特征时,为不影响实验的准确度,本文研究未对主颜色特征执行批标准化处理。经过多次实验证明,空间位置和形状相关特征在深度多特征融合时起主导作用,而主颜色特征起辅助作用,所以设去噪深度卷积特征权重值 $\beta=1$,主颜色特征权重值为 α 。个体 α 的适应度值由5次交叉验证的平均分类结果和Top10图像的分类错误数决定,如式(4)所示:

$$\text{fit}(\alpha) = Z_{\text{ZVAL}} - E_{\text{Top10}}^{\text{ERR}} \quad (4)$$

$$Z_{\text{ZVAL}} = \sum_{i=1}^5 V_i^{\text{VAL}} \quad (5)$$

$$V_i^{\text{VAL}} = \frac{\sum_{n=1}^{\text{epoch}} \text{val}_n}{\text{epoch}} \quad (6)$$

其中: n 是epoch的个数; VAL_i 是1次交叉验证的平均分类结果; Z_{ZVAL} 是5次交叉验证的平均分类结果; $E_{\text{Top10}}^{\text{ERR}}$ 是图像库中前10个图像的分类错误数。

由于待优参数只有1个,因此舍弃杂交操作,对标准差分演化算法^[19]中的变异和选择操作进行改进。得到基于差分演化的自适应权值CNN图像分类算法,算法详细步骤如下:

步骤1 在区间 $[10^{-1}, 10^{-10}]$ 内,初始化种群 P ,产生规模为NP的个体 α 。

步骤2 按照式(4)计算初始群体 P 中所有个体的适应度值。

步骤3 如果 $t < \text{最大迭代次数}$,则转步骤4,否则终止进化,将得到的最佳个体作为最优解输出。

步骤4 按照式(7)对每个个体 α_i 按照公式进行差分变异,产生 trial_{α_i} 。

$$\text{trial}_{\alpha_i} = \alpha_{\text{best}} + F(\alpha_{r_1} - \alpha_{r_2}) \quad (7)$$

其中: α_{best} 为当前群体的最优个体; r_1, r_2 为群体中随机选择的2个个体,且 $r_1 \neq r_2$ 。

步骤5 利用式(4)评估每个个体 α_i 的适应度值。

步骤6 对每个个体 α_i 进行更新操作。

$$\alpha_i = \begin{cases} \text{trial}_{\alpha_i}, & \text{fit}(\text{trial}_{\alpha_i}) > \text{fit}(\alpha_i) \\ \alpha_i, & \text{其他} \end{cases} \quad (8)$$

步骤7 跳转至步骤3。

3 实验结果与分析

3.1 实验参数

本文研究所用图像库是Corel公司提供的标准图像素材库Corel-1000。该库包含了人物、建筑物、公交车、花卉、动物等10类共1000幅图像。选择其中的800幅作为训练集,而剩余的200幅作为测试集。为了方便网络读取,对图像库进行了预处理操作,将图像大小由原来的 $3 \times 256 \times 328$ 压缩成 $3 \times 48 \times 48$ 。

神经网络模型的Epoch设为50,每次迭代次数为8,批处理大小为100幅图像。第1个卷积层的卷积核大小是 5×5 ,滤波器32个;第2个卷积层的卷积核大小是 5×5 ,滤波器64个;第3个、第4个卷积层的卷积核大小是 3×3 ,滤波器是128个。卷积步长设为1,填充设为0。池化层步长设置为2,池化窗口大小为 2×2 。最后,Softmax层有10个神经单元,将图像分为10类。另外,演化代数设为50,种群规模为37,

缩放因子 F 为 0.5。

通过最小化损失函数来获得最优解,本文研究使用交叉熵误差函数^[20]作为损失函数,其表达式为:

$$Q = -\frac{1}{N} \sum_{n=1}^N (y_n \ln o_n + \tilde{y}_n \ln \tilde{o}_n) \quad (9)$$

其中: $\tilde{y}_n = 1 - y_n$; $\tilde{o}_n = 1 - o_n$; N 为批处理训练样本个数; y 为每个样本的真实标签值; o 为网络的实际输出值。

此外,实验采用 Adam 优化器^[21]优化损失函数。结合 AdaGrad 和 RMSProp 2 种优化算法的优点,综合考虑梯度的一阶矩估计(梯度的均值)和二阶矩估计(梯度的未中心化的方差),计算更新步长。

1) 计算时刻 t 的梯度:

$$g_t = \nabla_{\theta} J(\theta_{t-1}) \quad (10)$$

2) 计算梯度的指数移动平均数:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (11)$$

其中: m_0 初始值设置为 0; 系数 β_1 是指数衰减率,默认为 0.9。

3) 计算梯度平方的指数移动平均数:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (12)$$

其中: v_0 初始值设置为 0; β_2 是指数衰减率,控制之前梯度平方的影响情况。

4) 对梯度均值 m_t 进行偏差纠正:

$$\hat{m}_t = m_t / (1 - \beta_1^t) \quad (13)$$

5) 对梯度方差 v_t 进行偏差纠正:

$$\hat{v}_t = v_t / (1 - \beta_2^t) \quad (14)$$

6) 更新参数,默认学习率 $\alpha = 0.001$:

$$\theta_t = \theta_{t-1} - \alpha \times \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon) \quad (15)$$

其中: $\varepsilon = 10^{-8}$, 避免除数为 0。步长通过梯度均值和梯度均方根进行自适应调节。

3.2 实验结果

本文研究将主颜色特征与空间位置和形状相关特征进行深度多特征融合的卷积神经网络称为 MCNN; 将基于深度多特征融合的自适应卷积神经网络称为 AMCNN。当 $\alpha = 1e-5$, $\beta = 1$ 时, CNN 和 MCNN、AMCNN 在测试集上的实验结果如表 1 所示。

表 1 CNN、MCNN 和 AMCNN 在测试集的准确率

Table 1 Accuracy of CNN, MCNN and AMCNN on test set		
算法	准确率	提高率
CNN	82.0	—
MCNN	86.5	5.49
AMCNN	91.2	11.21

分类算法的提高率为:

$$C_{CA} = \frac{A_{MCNN}^{Accuracy} - A_{CNN}^{Accuracy}}{A_{CNN}^{Accuracy}} \times 100\% \quad (16)$$

$$A_{Accuracy} = r/a \quad (17)$$

其中: a 代表图像库中的图像总数; r 代表图像库中分类正确的图像数目。

从实验效果可以看到,进行深度多特征融合时由于加入了主颜色特征弥补了 CNN 对颜色特征不敏感的欠缺,因此 MCNN 比 CNN 分类准确率提高了 5.49 个百分点;而 AMCNN 算法在融合深度多特征的基础上进一步优化权重值,寻找出各特征的最优权值,从而更大程度地提升了分类准确率,比 CNN 分类准确率提高了 11.21 个百分点。

图 3 是 AMCNN、MCNN 和 CNN 在测试集上的曲线精度对比,可以看到:和 CNN 相比,AMCNN 和 MCNN ($\alpha = 1e-5$) 在有效提高分类精度的同时加快了收敛速度。此外,由于各网络模型都使用了去噪深度卷积特征,因此网络模型的性能相对稳定和健壮。在整个测试集的曲线精度对比上没有出现分类精度大幅度起伏的情况。

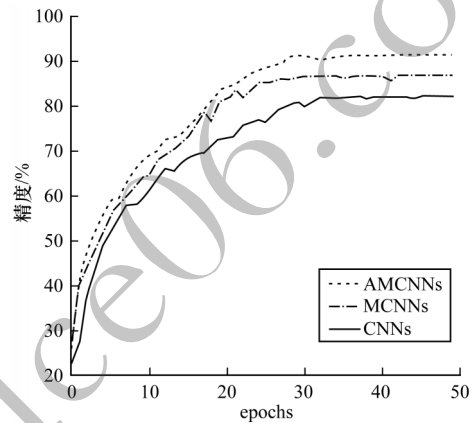


图 3 AMCNN、MCNN 和 CNN 在测试集上的曲线精度对比
Fig.3 Comparison of curve accuracy of AMCNN, MCNN and CNN on test set

使用 MCNN 网络模型对图像库进行分类时,不同的 α 、 β 值会产生不同的实验效果,故选取不同的 α 、 β 值进行实验,表 2 是其在测试集上的实验结果。

表 2 测试集上不同 α 、 β 值的实验结果

Table 2 Experimental results of different α and β values on test set

算法	α 值	β 值	准确率/%
MCNN	1e-5	1	86.5
MCNN	1e-6	1	89.0
AMCNN	2.5e-7	1	91.2

从表 2 可以看出:当 α 值设置为 $1e-5$, β 值设置为 1 时,准确率为 86.5%;而当 α 值被设置为 $1e-6$ 时,准确率为 89%,提升了 2.5 个百分点。从而可见,不同的 α 、 β 值对 MCNN 网络模型产生的实验效果不尽相同,当各特征权值达到最优时,AMCNN 模型的准

确率达到91.2%。因为实验使用的标准图像素材库Corel-1000中只有1 000幅10种类别的图像,所以当准确率达到91.2%时,是一个较好的分类结果。

AMCNN、MCNN和CNN在测试集上的准确率如图4所示。

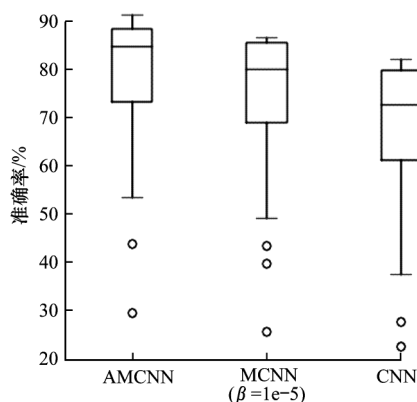


图4 AMCNN、MCNN和CNN在测试集上的准确率

Fig.4 Accuracy of AMCNN, MCNN and CNN on test set

从图4可以看出:CNN的最大、最小和中值准确率以及上四分位数和下四分位数均小于MCNN($\alpha=1e-5$)和AMCNN的相应数字特征。此外,还可以观察到:AMCNN的性能也相对稳定,抛出的异常值也是最优的。

当MCNN网络模型进行深度多特征融合时,使用深度去噪卷积特征和未使用深度去噪卷积特征(NMCNN)在测试集上得到的实验结果如表3所示。

表3 MCNN和NMCNN在测试集的准确率

Table 3 Accuracy of MCNN and NMCNN on test set

算法	α 值	β 值	准确率/%
MCNN	$1e-5$	1	86.5
NMCNN	$1e-5$	1	83.5

从表3可以看出:当进行深度多特征融合时,深度去噪卷积特征可以提高图像分类的准确率;相对于NMCNN,MCNN的实验性能提升了3个百分点。同时,与表1中使用了深度去噪卷积特征的CNN实验效果相比,融合主颜色特征比只使用深度去噪卷积特征的实验性能提升了1.5个百分点。

4 结束语

本文提出一种基于深度多特征融合的自适应CNN图像分类算法。该算法利用卷积神经网络提取空间位置和形状特征,弥补CNN网络模型只读取图像空间位置信息而对颜色信息不太敏感的不足,并在此基础上运用改进的差分演化算法优化各特征权重值的方法。实验结果表明,与CNN算法相比,该算法的图像分类准确率有较大幅度的提升,而且模型性能更加稳定和健壮。下一步将尝试通过对抗网络增大图像库样本数量,使模型性能达到更优。

参考文献

- [1] 周飞燕,金林鹏,董军. 卷积神经网络研究综述[J]. 计算机学报,2017,40(6):1229-1251.
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network[J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251. (in Chinese)
- [2] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [3] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks[J]. Pattern Recognition, 2018, 77(5): 354-377.
- [4] 杨少戈. 基于深度学习的冠脉造影图像分割[D]. 北京: 北京邮电大学, 2019.
YANG S G. Coronary angiography segmentation based on deep learning[D]. Beijing: Beijing University of Posts and Telecommunications, 2019. (in Chinese)
- [5] 温煌璐. 基于卷积神经网络的图像分类算法研究[D]. 乌鲁木齐: 新疆大学, 2018.
WEN H L. Research on image classification algorithms based on convolutional neural networks[D]. Urumqi: Xinjiang University, 2018. (in Chinese)
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Proceedings of ACM NIPS'12. New York, USA: ACM Press, 2012: 1097-1105.
- [7] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 1-9.
- [8] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of ACM International Conference on Machine Learning. New York, USA: ACM Press, 2015: 448-456.
- [9] 田启川,王满丽. 深度学习算法研究进展[J]. 计算机工程与应用, 2019, 55(22): 25-33.
TIAN Q C, WANG M L. Research progress of deep learning algorithm [J]. Computer Engineering and Applications, 2019, 55(22): 25-33. (in Chinese)
- [10] KIRAN R, PRADEEP K, BHARAT B. A novel deep learning based hybrid recommender system [J]. Expert Systems with Applications, 2020, 144(3): 113-124.
- [11] PENG X, ZHANG X M, LI Y P, et al. Research on image feature extraction and retrieval algorithms based on convolutional neural network [J]. Journal of Visual Communication and Image Representation, 2019, 69(1): 102-115.
- [12] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: a review and new perspectives [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [13] CHEN Y S, LI C Y, GHAMISI P, et al. Deep fusion of remote sensing data for accurate classification [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(8): 1253-1257.

(下转第251页)

(上接第 239 页)

- [14] 刘畅. 基于大数据技术支撑的线损异常分析与线损预测研究[D]. 上海:上海交通大学,2018.
LIU C. Research on exception analysis and forecasting of line loss based on big data theory[D]. Shanghai: Shanghai Jiao Tong University, 2018. (in Chinese)
- [15] LAMBERT J, SENER O, SAVARESE S. Deep learning under privileged information using heteroscedastic dropout [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA: IEEE Press, 2018: 8886-8895.
- [16] YAO X, XU Y. Recent advances in evolutionary computation [J]. Journal of Computer Science and Technology, 2006, 21(1): 1-18.
- [17] STORN R, PRICE K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces[J]. Journal of Global Optimization, 1997, 11(4): 341-359.
- [18] LIU B, WANG L, JIN H Y. Advances in differential evolution[J]. Control and Decision, 2007, 22(7): 721-729.
- [19] ABBASS H A, SARKER R, NEWTON C. PDE: a pareto-frontier differential evolution approach for multiobjective optimization problems[C]//Proceedings of IEEE Conference on Evolutionary Computation. Washington D. C. , USA: IEEE Press, 2001: 971-978.
- [20] LIN M, CHEN Q, YAN S. Network in network[EB/OL]. [2020-04-02]. <https://arxiv.org/abs/1312.4400>.
- [21] KINGMA D, BA J. Adam: a method for stochastic optimization[EB/OL]. [2020-04-02]. <https://arxiv.org/pdf/1412.6980.pdf>.

编辑 索书志