



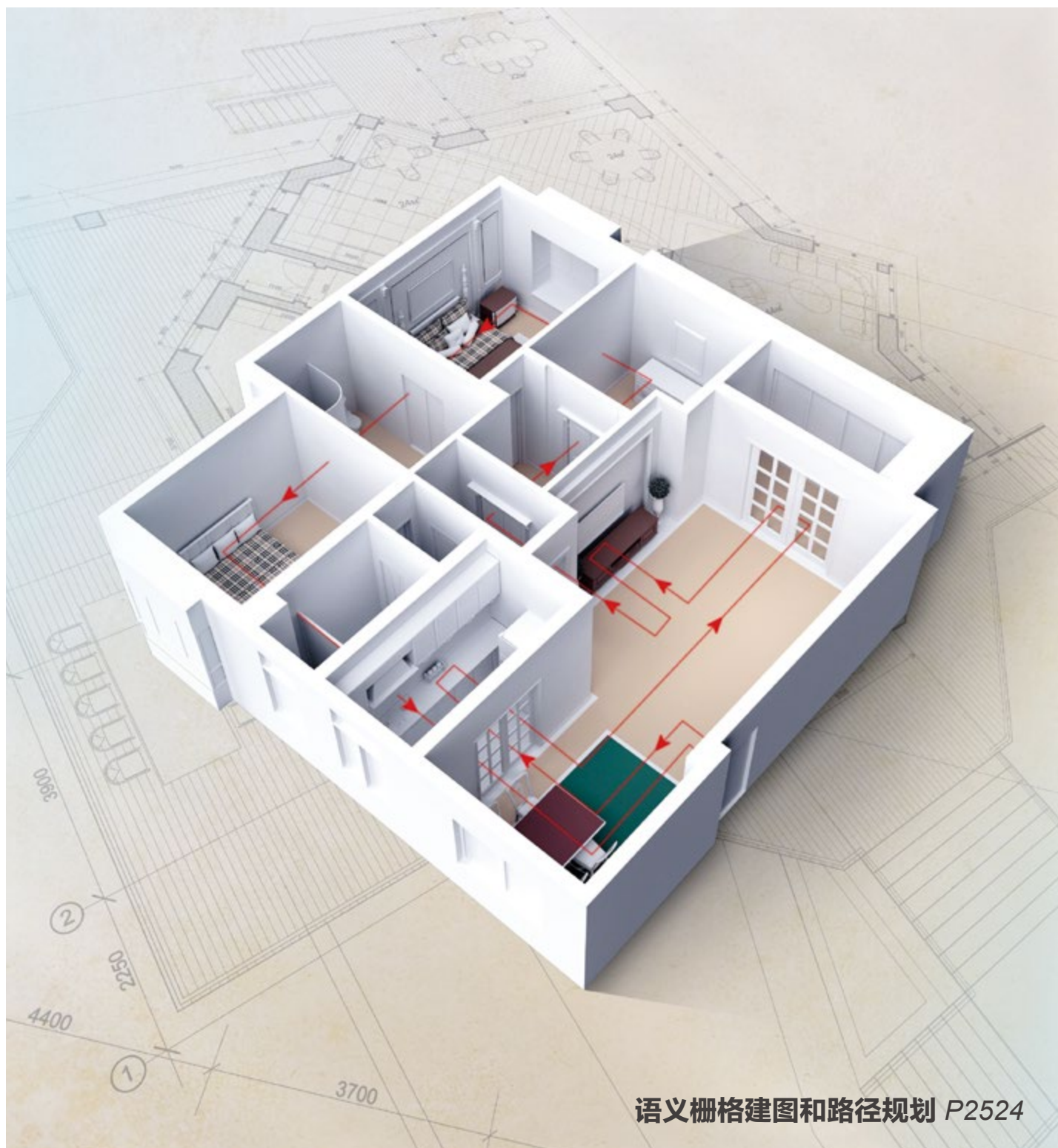
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
10
VOL.26

ISSN1006-8961
CN11-3758/TB



语义栅格建图和路径规划 P2524



第26卷第10期（总第306期）
2021年10月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@aircas.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

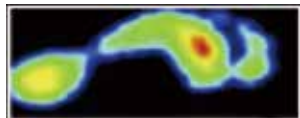
Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



穿鞋足迹序列的足迹能量图组表达与识别(第2357页)



时空图卷积网络与注意机制的视频目标分割(第2376页)



双光流网络指导的视频目标检测(第2473页)

综述

图像分类的深度卷积神经网络模型综述

张珂, 冯晓晗, 郭玉荣, 苏昱坤, 赵凯, 赵振兵, 马占宇, 丁巧林 2305

中国图像工程25年

章毓晋 2326

数据集论文

FGSC-23: 面向深度学习精细识别的高分辨率光学遥感图像舰船目标数据集

姚力波, 张筱晗, 吕亚飞, 孙炜玮, 李孟洋 2337

图像处理和编码

并行生成网络的红外—可见光图像转换

余佩伦, 施佳, 王晗 2346

图像分析和识别

穿鞋足迹序列的足迹能量图组表达与识别

王新年, 于丹, 张涛 2357

时空图卷积网络与注意机制的视频目标分割

姚睿, 夏士雄, 周勇, 赵佳琦, 胡伏原 2376

多支路协同的RGB-T图像显著性目标检测

蒋亭亭, 刘昱, 马欣, 孙景林 2388

改进R-FCN模型的小尺度行人检测

刘万军, 董利兵, 曲海成 2400

边缘与区域不一致性引导下的图像拼接篡改检测网络

蒋小玉, 刘春晓 2411

图像理解和计算机视觉

各向异性导向滤波的红外与可见光图像融合

刘明威, 王任华, 李静, 焦映臻 2421

卷积稀疏与细节显著图解析的图像融合

杨培, 高雷阜, 訾玲玲 2433

计算机图形学

利用距离与内能极小平滑链接Bézier曲线

李军成, 刘成志, 刘珊君 2450

NCIG 2020

结合置信度加权融合与视觉注意机制的前景检测

成科扬, 孙爽, 王文杉, 师文喜, 李鹏, 詹永照 2462

双光流网络指导的视频目标检测

尉婉青, 禹晶, 史薪琪, 肖创柏 2473

生成对抗网络驱动的图像隐写与水印模型

郑钢, 胡东辉, 戈辉, 郑淑丽 2485

融合语义—表观特征的无监督前景分割

李熹, 马惠敏, 马洪兵, 王弈冬 2503

场景视点偏移的激光雷达点云分割

郑阳, 林春雨, 廖康, 赵耀, 薛松 2514

激光—相机系统语义栅格建图和路径规划

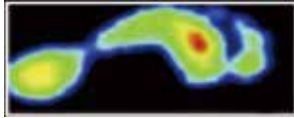
丁梦远, 郭迟, 黄凯 2524

融合词袋连通图的图像检索特征选择

李国祥, 王继军, 马文斌 2533

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Shoeprint sequence representation and recognition using shoeprint energy map set (P2357)



Spatial-temporal video object segmentation with graph convolutional network and attention mechanism(P2376)



Dual optical flow network-guided video object detection(P2473)

Review

Overview of deep convolutional neural networks for image classification

Zhang Ke, Feng Xiaohan, Guo Yurong, Su Yukun, Zhao Kai, Zhao Zhenbing, Ma Zhanyu, Ding Qiaolin 2305

Twenty-five years of image engineering in China

Zhang Yujin 2326

Dataset

FGSC-23: a large-scale dataset of high-resolution optical remote sensing image for deep learning-based fine-grained ship recognition

Yao Libo, Zhang Xiaohan, Lyu Yafei, Sun Weiwei, Li Mengyang 2337

Image Processing and Coding

Infrared-to-visible image translation based on parallel generator network

Yu Peilun, Shi Quan, Wang Han 2346

Image Analysis and Recognition

Shoeprint sequence representation and recognition using shoeprint energy map set

Wang Xinnian, Yu Dan, Zhang Tao 2357

Spatial-temporal video object segmentation with graph convolutional network and attention mechanism

Yao Rui, Xia Shixiong, Zhou Yong, Zhao Jiaqi, Hu Fuyuan 2376

Multi-path collaborative salient object detection based on RGB-T images

Jiang Tingting, Liu Yu, Ma Xin, Sun Jinglin 2388

Small-scale pedestrian detection based on improved R-FCN model

Liu Wanjun, Dong Libing, Qu Haicheng 2400

Edge and region inconsistency-guided image splicing tamper detection network

Jiang Xiaoyu, Liu Chunxiao 2411

Image Understanding and Computer Vision

Infrared and visible image fusion with multi-scale anisotropic guided filtering

Liu Mingwei, Wang Renhua, Li Jing, Jiao Yingzhen 2421

Image fusion method of convolution sparsity and detail saliency map analysis

Yang Pei, Gao Leifu, Zi Lingling 2433

Computer Graphics

Smoothing linked Bézier curves by distance and internal energy minimization

Li Juncheng, Liu Chengzhi, Liu Shanjun 2450

NCIG 2020

Foreground detection via fusing confidence by weight and visual attention mechanism

Cheng Keyang, Sun Shuang, Wang Wenshan, Shi Wenxi, Li Peng, Zhan Yongzhao 2462

Dual optical flow network-guided video object detection

Yu Wanjing, Yu Jing, Shi Xinqi, Xiao Chuangbai 2473

End-to-end image steganography and watermarking driven by generative adversarial networks

Zheng Gang, Hu Donghui, Ge Hui, Zheng Shuli 2485

Semantic-apparent feature-fusion-based unsupervised foreground segmentation method

Li Xi, Ma Huimin, Ma Hongbing, Wang Yidong 2503

LiDAR point cloud segmentation through scene viewpoint offset

Zheng Yang, Lin Chunyu, Liao Kang, Zhao Yao, Xue Song 2514

Semantic grid mapping and path planning combined with laser-camera system

Ding Mengyuan, Guo Chi, Huang Kai 2524

Feature selection method for image retrieval based on connected graphs and bag of words

Li Guoxiang, Wang Jijun, Ma Wenbin 2533

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2021)10-2305-21

论文引用格式: Zhang K, Feng X H, Guo Y R, Su Y K, Zhao K, Zhao Z B, Ma Z Y and Ding Q L. 2021. Overview of deep convolutional neural networks for image classification. Journal of Image and Graphics, 26(10): 2305-2325 (张珂, 冯晓晗, 郭玉荣, 苏昱坤, 赵凯, 赵振兵, 马占宇, 丁巧林. 2021. 图像分类的深度卷积神经网络模型综述. 中国图象图形学报, 26(10): 2305-2325) [DOI:10.11834/jig.200302]

图像分类的深度卷积神经网络模型综述

张珂^{1,2*}, 冯晓晗¹, 郭玉荣², 苏昱坤¹, 赵凯¹, 赵振兵¹, 马占宇², 丁巧林¹

1. 华北电力大学电子与通信工程系, 保定 071000; 2. 北京邮电大学人工智能学院, 北京 100086

摘要: 图像分类是计算机视觉中的一项重要任务, 传统的图像分类方法具有一定的局限性。随着人工智能技术的发展, 深度学习技术越来越成熟, 利用深度卷积神经网络对图像进行分类成为研究热点, 图像分类的深度卷积神经网络结构越来越多样, 其性能远远好于传统的图像分类方法。本文立足于图像分类的深度卷积神经网络模型结构, 根据模型发展和模型优化的历程, 将深度卷积神经网络分为经典深度卷积神经网络模型、注意力机制深度卷积神经网络模型、轻量级深度卷积神经网络模型和神经网络架构搜索模型等4类, 并对各类深度卷积神经网络模型结构的构造方法和特点进行了全面综述, 对各类分类模型的性能进行了对比与分析。虽然深度卷积神经网络模型的结构设计越来越精妙, 模型优化的方法越来越强大, 图像分类准确率在不断刷新的同时, 模型的参数量也在逐渐降低, 训练和推理速度不断加快。然而深度卷积神经网络模型仍有一定的局限性, 本文给出了存在的问题和未来可能的研究方向, 即深度卷积神经网络模型主要以有监督学习方式对图像进行分类, 受到数据集质量和规模的限制, 无监督式学习和半监督学习方式的深度卷积神经网络模型将是未来的重点研究方向之一; 深度卷积神经网络模型的速度和资源消耗仍不尽人意, 应用于移动设备具有一定的挑战性; 模型的优化方法以及衡量模型优劣的度量方法有待深入研究; 人工设计深度卷积神经网络结构耗时耗力, 神经网络架构搜索方法将是未来深度卷积神经网络模型设计的发展方向。

关键词: 深度学习; 图像分类(IC); 深度卷积神经网络(DCNN); 模型结构; 模型优化

Overview of deep convolutional neural networks for image classification

Zhang Ke^{1,2*}, Feng Xiaohan¹, Guo Yurong², Su Yukun¹, Zhao Kai¹,
Zhao Zhenbing¹, Ma Zhanyu², Ding Qiaolin¹

1. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071000, China;

2. Institute of Artificial Intelligence, Beijing University of Posts and Telecommunication, Beijing 100086, China

Abstract: Image classification (IC) is one of important tasks in support of computer vision. Traditional image classification methods have limitations on the aspect of computer vision. Deep learning technology has become more mature than before based on deep convolutional neural network (DCNN) with the development of artificial intelligence (AI) recently. The performance of image classification has been upgraded based on the maturation of the deep convolutional neural network

收稿日期: 2020-06-16; 修回日期: 2020-09-21; 预印本日期: 2020-09-28

* 通信作者: 张珂 zhangkeit@ncepu.edu.cn

基金项目: 国家自然科学基金项目(62076093, 61871182, 61922015, 61773071, 61302163); 河北省自然科学基金项目(F2020502009, F2015502062, F2016502062); 北京市自然科学基金项目(4192055); 中央高校基本科研业务费专项资金资助(2020YJ006, 2020MS099)

Supported by: National Natural Science Foundation of China (62076093, 61871182, 61922015, 61773071, 61302163); Natural Science Foundation of Hebei Province, China (F2020502009, F2015502062, F2016502062); Beijing Municipal Natural Science Foundation (4192055); Fundamental Research Funds for the Central Universities (2020YJ006, 2020MS099)

model. This research has mainly focused on a comprehensive overview of image classification in DCNN via the deep convolutional neural network model structure of image classification. Firstly, the modeling methodology has been analyzed and summarized. The DCNN analysis has been formulated into four categories listed below: 1) classic deep convolutional neural networks; 2) deep convolutional neural networks based on the attention mechanism; 3) lightweight networks; 4) the neural architecture search method. DCNN has high optimization capability using convolution to extract effective features of the images and learn feature expression from a large number of samples automatically. DCNN achieves better performance on image classification due to the effective features based on the deeper DCNN research and development. DCNN has been encountering lots of difficulties such as overfitting, vanishing gradient and huge model parameters. Hence, DCNN has become more and more difficult to optimize. The researchers in the context of IC have illustrated different DCNN models for different problems. Researchers have been making the network deeper than before via AlexNet. Subsequently, the classified analyses such as network in network (NIN), Overfeat, ZFNet, Visual Geometry Group (VGGNet), GoogLeNet have been persisted on. The problem of vanishing gradient has been more intensified via the deepening of the network. The optimization of the network becomes more complicated. Researchers have proposed residual network (ResNet) to ease gradient vanishing to improve the performance of image classification greatly. To further improve the performance of ResNet, researchers have issued a series of ResNet variants which can be divided into three categories in terms of different solutions via ResNet variants based on very deep ResNet optimization, ResNet variants based on increasing width and the new dimensions in ResNet variants. The ResNet has been attributed to the use of shortcut connections maximization. Densely connected convolutional network (DenseNet) have been demonstrated and the information flow in DenseNet between each layer has been maximized. To further promote the information flow between layers, the DenseNet variants have been illustrated via DPN (dual path network) and CliqueNet. DCNN based on the attention mechanism has focused on the regions of interest based on the classic DCNN models and channel attention mechanism, spatial attention mechanism and layer attention mechanism can be categorized. DCNN need higher accuracy and a small amount of parameters and fast model calculation speed. The researchers have proposed the lightweight networks such as the ShuffleNet series and MobileNet series. The NAS (neural architecture search) methods using neural networks to automatically design neural networks have been concerned. The NAS methods can be divided into three categories: design search space, model optimization and others. Secondly, The image classification datasets have been commonly presented in common including MNIST (modified NIST (MNIST)) dataset, ImageNet dataset, CIFAR dataset and SVHN (street view house number (SVHN)) dataset. The comparative performance and analysis of experimental results of various models were conducted as well. The accuracy, parameter and FLOPs (floating point operations) analyses to measure the results of classification have been mentioned. The capability of model optimization has been upgraded gradually via the accuracy improvement of image classification, the decreasing amount of parameters of the model and increasing speed of training and inference. Finally, the DCNN model has been constrained some factors. The DCNN model has been mainly used to supervise deep learning for image classification in constraint of the quality and scale of the datasets. The speed and resource consuming of the DCNN model have been upgraded in mobile devices. The measurement and optimization in analyzing the advantages and disadvantages of the DCNN model need to be studied further. The neural architecture search method will be the development direction of future deep convolutional neural network model designs. The DCNN models of image classification have been reviewed and the experimental results of the DCNNs have been demonstrated.

Key words: deep learning; image classification (IC); deep convolutional neural networks (DCNN); model structure; model optimization

0 引言

图像分类是计算机视觉基本任务之一。顾名思义,图像分类即给定一幅图像,计算机利用算法找出其所属的类别标签。图像分类的过程主要包括图像

的预处理、图像的特征提取以及使用分类器对图像进行分类,其中图像的特征提取是至关重要的一步。传统的图像分类算法提取图像的色彩、纹理和空间等特征(孙君顶和赵珊,2009),其在简单的图像分类任务中表现较好,但在复杂图像分类任务中表现不尽人意。

随着智能信息时代的来临,深度学习(LeCun等,2015)应运而生。深度学习作为机器学习的一个分支,旨在模拟人类的神经网络系统构建深度人工神经网络,对输入的数据进行分析和解释,将数据的底层特征组合成抽象的高层特征,其在计算机视觉、自然语言处理等人工智能领域发挥了不可替代的作用。作为深度学习的典型代表,深度卷积神经网络(deep convolutional neural network,DCNN)在计算机视觉任务中大放异彩,与人工提取特征的传统图像分类算法相比,卷积神经网络使用卷积操作对输入图像进行特征提取,有效地从大量样本中学习特征表达,模型泛化能力更强。

2012年,AlexNet(Krizhevsky等,2012)大幅提升了图像分类等任务的性能,并取得了ImageNet(Deng等,2009)挑战赛冠军,使深度卷积神经网络逐渐成为计算机视觉领域研究的热点。随着DCNN网络结构越来越深,性能越来越强,过拟合、梯度消失、模型参数量巨大和难以优化等问题也越来越突出,科研人员针对不同的问题提出了不同结构的DCNN模型。

现有深度卷积神经网络综述的内容较为宽泛,不聚焦某一特殊任务,往往围绕深度卷积神经网络模型不同方面的优化方法与应用进行综述(林景栋等,2020;Gu等,2018;Khan等,2019;张顺等,2019)。本文以图像分类任务为载体,从经典深度卷积神经网络模型、注意力机制卷积神经网络模型、轻量级卷积神经网络模型以及神经网络架构搜索模型4个方面,综述深度卷积神经网络基础模型近年来的研究进展,分析和对比主要的深度卷积神经网络模型,并对深度卷积神经网络模型设计未来可能的发展方向进行展望。

1 深度卷积神经网络模型

本文根据近年来基于DCNN的图像分类研究发展过程 and 方向,将深度卷积神经网络模型分为以下4类:经典的深度卷积神经网络;基于注意力机制的深度卷积神经网络模型;轻量级深度卷积神经网络模型;基于神经架构搜索(neural architecture search,NAS)的深度卷积神经网络模型。经典的深度卷积神经网络模型主要是指以增加网络深度、提升网络性能为目标的DCNN模型;基于注意力机制

的深度卷积神经网络模型则是建立在经典DCNN模型基础上,采用注意力机制使网络模型更关注感兴趣的区域;随着DCNN模型越来越深,网络结构越来越复杂,网络参数量大的问题越来越突出,轻量级深度卷积神经网络通过改进模型结构降低网络参数量以适应嵌入式、移动式设备的需求;基于神经架构搜索的深度卷积神经网络模型采用神经网络自动设计DCNN模型结构,与人工设计DCNN相比更省时省力,是未来DCNN的发展趋势之一。

1.1 经典的深度卷积神经网络模型

1.1.1 从LeNet到GoogLeNet

LeNet模型(LeCun等,1998)是最早提出的卷积神经网络模型,主要用于MNIST(modified NIST)数据集中手写数字识别,模型结构如图1所示。其包含3个卷积层、2个池化层和2个全连接层,每个卷积层和全连接层均有可训练的参数,为深度卷积神经网络的发展奠定了基础。

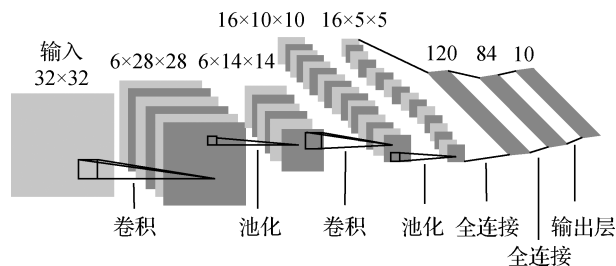


图1 LeNet模型结构(LeCun等,1998)

Fig. 1 Architecture of LeNet(LeCun et al.,1998)

尽管LeNet在小规模MNIST数据集上取得了不错的效果,但复杂的图像分类任务则需要大规模数据集以及学习能力更强的网络模型。2012年,Krizhevsky等人(2012)提出了AlexNet,网络结构如图2所示,该网络包含5个卷积层和3个全连接层,输入图像经过卷积操作和全连接层的操作,最后输入具有1000个节点的Softmax分类器完成图像分类。该网络通过使用线性整流函数(rectified linear unit,ReLU)作为激活函数,引入局部响应归一化(local response normalization,LRN)缓解梯度消失问题;使用数据增强和Dropout技术大大缓解了过拟合问题;并采用两个GPU并行计算的方式训练,提高了训练速度。AlexNet在2012年ImageNet比赛中以远超当时亚军的优势获得分类任务冠军。

早期的深度卷积神经网络,例如AlexNet,其结构首先为卷积层,随后为全连接层,最后为Softmax

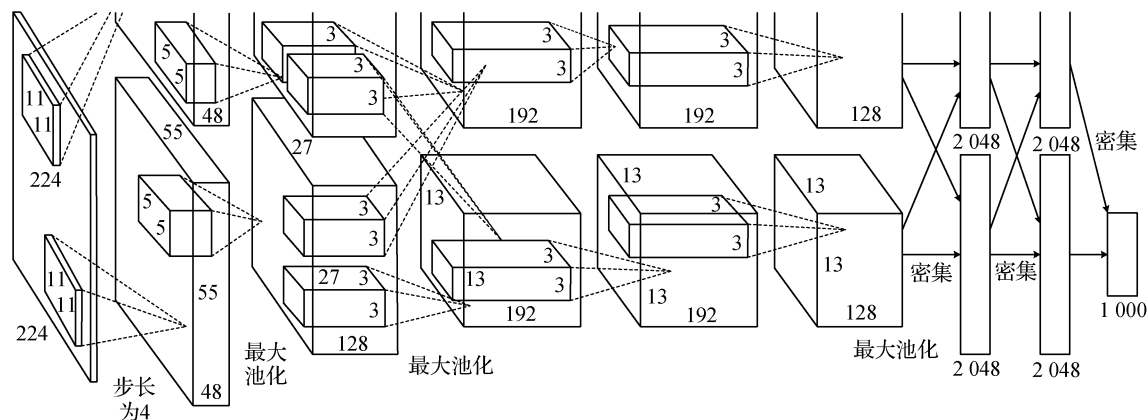


图2 AlexNet 网络结构(Krizhevsky 等,2012)

Fig.2 Architecture of AlexNet(Krizhevsky et al. ,2012)

分类器,网络的参数量很大。Lin 等人(2013)对传统的卷积神经网络结构进行改进,提出了 NIN(network in network),将卷积层替换为多层感知器卷积,并将全连接层替换为全局平均池化层。NIN 能学习到更复杂和有用的特征,而参数量仅为 AlexNet 的十分之一,在 CIFAR-10/100(Krizhevsky, 2009)、SVHN(street view house number)数据集上(Netzer 等,2011)获得了比 AlexNet 更高的准确率。

尽管 AlexNet 取得了进步,但对模型如何工作无从知晓,Zeiler 和 Fergus(2014)对 AlexNet 进行微调提出了 ZFNet。与 AlexNet 相比,ZFNet 使用了反卷积可视化特征图,前部的层使用了更小的卷积核和步长,通过研究遮挡局部图像对图像分类结果的影响获得了决定图像类别的关键部位。

Sermanet 等人(2014)提出的 OverFeat 网络实现了利用一种卷积神经网络架构完成图像分类、定位和检测任务,OverFeat 网络的前几层采用卷积神经网络在图像的多个位置和多个尺度上以滑动窗口的方式提取特征,改变 OverFeat 网络的后几层,实现定位、检测任务。

AlexNet 初始几层使用较大尺寸卷积核进行卷积,导致参数量较大,Simonyan 和 Zisserman(2015)提出了 VGGNet(Visual Geometry Group),继承了 AlexNet 和 LeNet 的框架,主要贡献在于通过堆叠采用 3×3 小卷积核的卷积层,增加了网络深度,提升了网络性能。VGGNet 包含 5 种结构,其中最常用的是 VGGNet-19 和 VGGNet-16。

一般来说,提升网络性能最直接的方法是增加网络深度,但随着网络深度的增加,参数量加大,网

络更易产生过拟合,同时对计算资源的需求也显著增加。Szegedy 等人(2015)提出的 GoogLeNet 采用了 Inception-v1 模块,该模块采用稀疏连接降低模型参数量的同时,保证了计算资源的使用效率,在深度达到 22 层的情况下提升了网络的性能。Inception-v1 模块结构如图 3 所示,该结构包含 4 条并行的支路, 3×3 卷积、 5×5 卷积之前的 1×1 卷积和 3×3 最大池化之后的 1×1 卷积用来减少参数量,输入图像进入该模块进行卷积或池化操作,得到 4 种尺寸相同的特征图,随后在通道维度上相并得到该模块的输出。Inception-v2 模块(Ioffe 和 Szegedy,2015)在 Inception-v1 模块基础上增加了批量归一化(batch normalization,BN)层和卷积分解,BN 算法减少了内部协方差转移,加速了网络的训练且缓解了梯度消失问题,卷积分解将 5×5 卷积分解成两个 3×3 卷积,从而减少了参数量。Inception-v3 模块(Szegedy 等,2016a)在 Inception-v2 模块基础上进行非对称卷积分解,如将 $n \times n$ 大小的卷积分解成

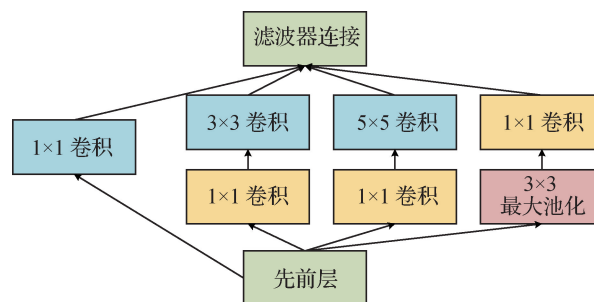


图3 Inception-v1 模块结构(Szegedy 等,2015)

Fig.3 Architecture of Inception-v1 module
(Szegedy et al. , 2015)

$1 \times n$ 卷积和 $n \times 1$ 卷积的串联, 且 n 越大, 参数量减少得越多。得益于内存优化, Inception-v4 (Szegedy 等, 2016b) 相比 Inception-v3 具有更统一的简化架构、更多的 Inception 模块, Inception-v4 提出的同时, Szegedy 等人 (2016b) 还将 Inception 模块与残差模块相结合, 提出了 Inception-ResNet 系列结构, 其中 Inception-ResNet-v1 网络在每个 Inception-ResNet 模块的最后加入了 1×1 卷积, Inception-ResNet-v2 网络对残差模块的输出进行了整形。

深层网络的优化十分困难, Srivastava 等人 (2015) 提出了 HN (highway network), 该网络使用门控单元调节信息流, 使信息可在多层之间畅通无阻流动, 从而有效抑制网络的梯度消失问题。

为使网络深度进一步加深, Larsson 等人 (2017) 提出了分形网络 (FractalNet), 该网络包含不同长度的交互子路径, 但不包含残差连接, 其通过重复使用简单的扩展规则可以生成极深的网络, 并取得了不错的性能。

1.1.2 ResNet 家族

深度卷积神经网络不断在图像分类任务上取得突破, 网络深度的增加提升了其特征提取能力。然而随着网络深度的增加, 梯度消失的问题越来越严重, 网络的优化越来越困难。据此, He 等人 (2016a) 提出了残差卷积神经网络 (residual networks, ResNet), 进一步加深网络的同时提升了图像分类任务的性能。

ResNet 由堆叠的残差块组成, 残差块结构如图 4 所示, 残差块除了包含权重层, 还通过越层连接将输入 x 直接连到输出上, $F(x)$ 为残差映射, $H(x)$ 为原始映射, 残差网络令堆叠的权重层拟合残差映射

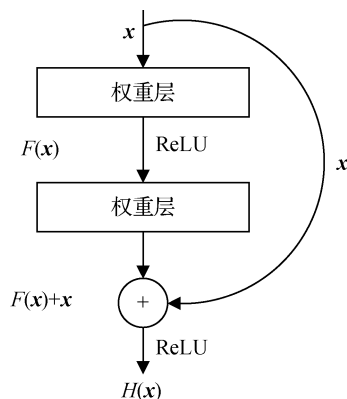


图4 残差块 (He 等, 2016a)

Fig. 4 Residual block (He et al., 2016a)

$F(x)$ 而不是原始映射 $H(x)$, 则 $F(x) = H(x) - x$, 而学习残差映射较学习原始映射简单。另外, 越层连接使得不同层的特征可以互相传递, 一定程度上缓解了梯度消失问题。

ResNet 通过堆叠残差块使网络深度达到 152 层, 残差网络在图像分类任务中获得了较大的成功。但随着网络的继续加深, 梯度消失问题仍然存在, 网络的优化越来越困难, 为进一步提升残差网络的性能, 研究者们提出了一系列残差网络的变体, 本文根据这些变体基本思路的不同, 将其分为 4 类: 基于深度残差网络优化的残差网络变体、采用新的训练方法的残差网络变体、基于增加宽度的残差网络变体和采用新维度的残差网络变体。基于深度残差网络优化的变体有 Pre-ResNet、加权残差网络 (weighted residual network, WResNet)、金字塔残差网络 (pyramidal residual network, PyramidalNet)、多级残差卷积神经网络 (residual networks of residual networks, RoR)、金字塔多级残差卷积神经网络 (pyramidal RoR, PRoR) 等; 采用新的训练方法有随机深度 (stochastic depth, SD) 网络、Swapout 和卷积残差记忆网络 (convolutional residual memory networks, CRMN) 等; 基于增加宽度的残差网络变体包括 ResNet in ResNet、宽残差网络 (wide residual networks, WRN) 和多残差网络 (multi-ResNet) 等, 其中的宽度包括特征图中的通道数、残差块中残差函数的数量等; 最后, 一些研究者在残差网络的改进中, 提出了新的维度, 例如基数、尺度和结构多样性等。

1) 基于深度残差网络优化的残差网络变体。He 等人 (2016b) 通过研究残差块中信息的传播, 提出了一种新的残差单元, 该残差单元去除了加法操作之后的 ReLU 激活函数, 而将残差支路的操作改为 BN-ReLU-conv-BN-ReLU-conv, 因此残差单元之间的信息、整个网络中的信息可以直接传播, 使得网络更易训练, 性能更好。

针对 ReLU 激活函数和逐元素相加之间的不兼容性, 以及深度网络很难使用 MSRA 初始化器 (He 等, 2015) 使深度网络收敛等问题, Shen 等人 (2016) 提出了加权残差网络 (weighted residual network, WResNet), 该网络中所有残差权重都初始化为零, 并以很小的学习率 (0.001) 进行优化, 这使得所有残差信号逐渐加到直接通路上, 这样 1192 层的残差网络比 100 层的残差网络收敛更快, 且从

100 + 层网络增长到 1 000 + 层网络,在准确率上会获得平稳的提升;最终残差权重分布在 $[-0.5, 0.5]$ 间,缓解了 ReLU 激活函数和逐元素相加之间的不兼容性。

在 ResNet 中,随着特征图通道数的增加而提取到的高级特征对分类任务更有效。Veit 等人 (2016) 研究发现,在 ResNet 特征图通道数加倍的情况下,删除残差单元中带有下采样的构建块会极大降低模型分类准确率,使用随机深度训练残差网络会缓解以上情况。受此启发,Han 等人 (2017) 提出了金字塔残差网络,该网络中每个残差单元的输出通道数都逐步增加,以将受下采样影响而集中分布在单个残差单元上的压力分布在所有残差单元上。Yamada 等人 (2016) 将 SD-ResNet 和 PyramidalNet 合并,提出了基于分离随机深度算法的深度金字塔残差网络模型 (deep pyramidal residual networks with separated stochastic depth, PyramidalSepDrop), 该网络的残差映射 $F(x)$ 被分成上下两部分,上部分用来增加通道数,下部分与输入 x 具有相同的通道数,两部分均使用随机深度的随机下降机制。

Zhang 等人 (2018a) 假设残差映射易于优化,则残差映射的残差映射更易优化,并据此在 ResNet 基础上逐级加入越层连接,构建了多级残差卷积神经网络,使得高层特征可以向低层传递,进一步抑制了梯度消失问题。RoR (residual networks for residual networks) 结构如图 5 (Zhang 等, 2018a) 所示,首先在所有残差块外添加一个越层连接,称为一级越层连接;然后,根据卷积核滤波器的种类将残差块分为若干组,在每组残差块外添加越层连接,称为二级越层连接;随后可将每组残差块再平分,添加越层连接;最后,原始残差块中的越层连接称为末级越层连接。

RoR 网络中每组残差块的特征图尺寸和通道数保持不变,下一组残差块开始时,特征图尺寸减半、通道数加倍,这使得网络中特征信息传递不连贯,会损失一些与预测相关的有用信息,限制了网络的分类性能。针对此问题,Zhang 等人 (2019a) 提出了金字塔多级残差卷积神经网络,结构如图 6 (Zhang 等, 2019a) 所示,该网络通过线性逐步增加每个残差块的输出通道数,保证高级属性多样性的同时也保证了信息的连续性。

2) 采用新训练方法的残差网络变体。模型深

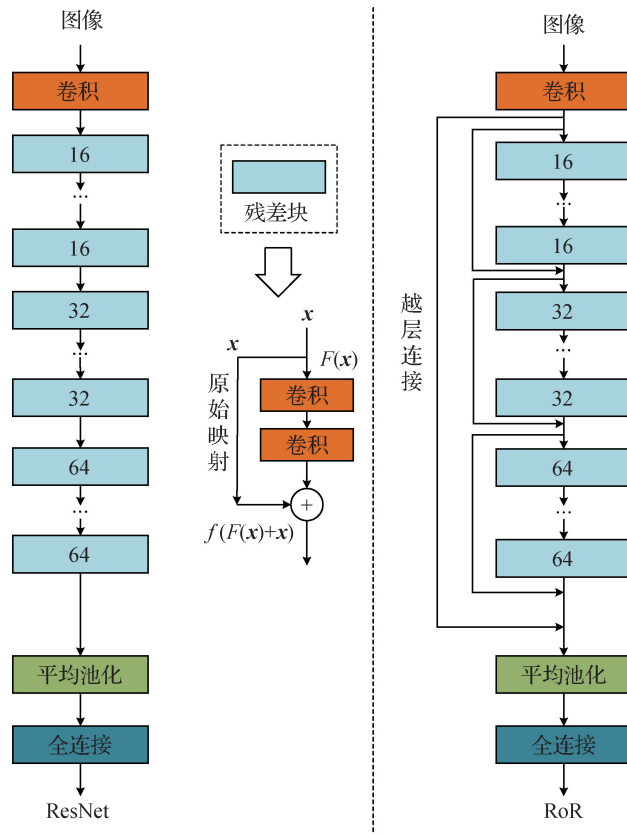


图5 RoR 模型结构 (Zhang 等人, 2018a)

Fig. 5 The architecture of RoR (Zhang et al., 2018a)

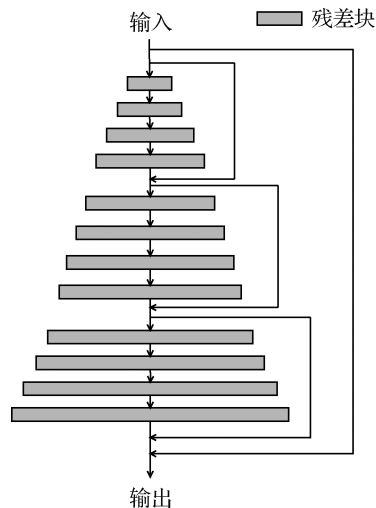


图6 金字塔多级残差卷积神经网络 (Zhang 等, 2019a)

Fig. 6 Pyramidal RoR (Zhang et al., 2019a)

度是模型表达能力的重要决定因素,然而非常深的网络面临着巨大的挑战:反向传播中的梯度消失问题、前向传播中的特征重用以及网络训练耗时长等问题。针对这些问题,研究者们从训练方法角度提出了一系列残差网络的变体。Huang 等人 (2016) 提出了一种新的随机深度残差网络训练方法,该方法

采用了集成学习思想,在模型训练时以一定的概率随机丢弃不同的残差块,在模型训练时每次迭代训练较浅的子网络,网络测试时则采用完整的深网络,模型训练时间缩短的同时获得分类准确率的提升。Singh 等人(2016)提出了一种新颖的随机训练方案——Swapout,其将 Dropout、随机深度等训练方法结合,从丰富的残差体系结构中取样,提高了残差网络的性能,较宽但较浅的 Swapout 网络可达到深度残差网络的性能。

Moniz 和 Pal(2016)提出了卷积残差记忆网络,该网络采用深度残差网络作为基础网络,使用长短期记忆(long short-term memory, LSTM)内存操作和算法架构的内存接口对网络进行训练,在 CIFAR-100 数据集上获得了当时最好的性能。

3) 基于增加宽度的残差网络变体。如上所述,为了提高残差网络的模型精度,研究者们主要致力于使 ResNet 深度更深或者深度残差网络的优化问题。而有些研究者另辟蹊径,他们提出的模型旨在使网络更宽,而不是更深,例如 ResNet in ResNet (RiR), WRN 和 Multi-ResNet。

ResNet 中的恒等映射导致不同特征的混合连接,然而在深度网络中,前部层学习的特征可能不再对后部层提供有用的信息。针对此问题,Targ 等人(2016)提出了一种广义的残差结构,该广义残差结构的模块单元是由残差流和瞬态流组成的并行结构,如图 7 所示,其中残差流包含越层连接且与原始残差块相似,瞬态流则是标准的卷积层,另外每个广义残差块中还有额外的两个卷积核滤波器来传递信息。两个连接的广义残差块称为 ResNet Init,将原始残差块的两个卷积层用 ResNet Init 代替,组成的新的结构称为 RiR 构建块,如图 8(Targ 等,2016)所示。RiR 网络在 CIFAR-10 数据集上取得了具有竞争力的结果。

在 ResNet 中,存在特征过度重用问题,寻求少量精确度的增加需要将网络层数加倍。针对该问题,Zagoruyko 和 Komodakis(2016)提出了宽残差网络,WRN 在原始残差块的基础上成倍地增加残差块中卷积核的个数,增加了网络的宽度,如图 9 所示,变量 k 代表宽网络卷积核较基准网络卷积核的倍数,该网络降低了网络深度,其性能远超相同层数的残差网络。

实验表明,深度残差网络难以通过在整个网络

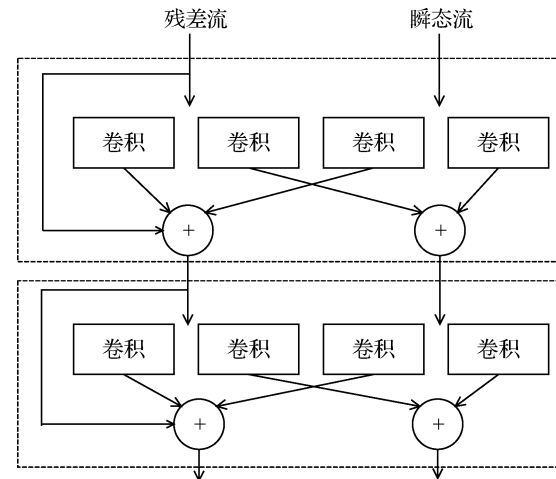


图 7 两个广义残差块(Targ 等,2016)

Fig. 7 Two generalized residual blocks
(ResNet Init)(Targ et al., 2016)

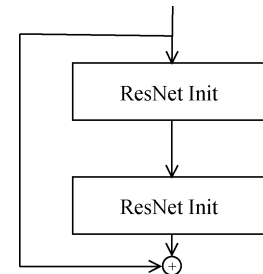


图 8 2 层的 RiR 构建块(Targ 等,2016)

Fig. 8 Two-layer RiR block(Targ et al., 2016)

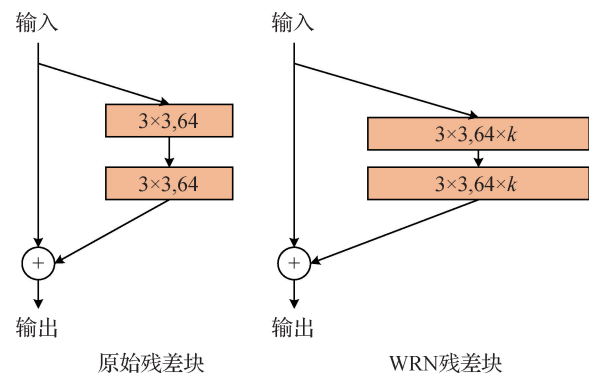


图 9 原始残差块与 WRN 残差块对比

(Zagoruyko 和 Komodakis, 2016)

Fig. 9 Comparison of basic residual block between
WRN residual block(Zagoruyko and Komodakis, 2016)

深度保持梯度来解决梯度消失问题,相反,可通过组合指数分布的、不同深度的网络来解决此问题,即增加网络的多重性。据此,Abdi 和 Nahavandi(2017)提出了多残差网络(Multi-ResNet),该网络通过增加

每个残差块中残差函数的数量,在保持深度不变的情况下,增加了网络的多样性。实验表明,增加残差块中残差函数数量比增加网络深度性能更好,与包含相同数量卷积层的深度残差网络相比,Multi-ResNet 能够取得更小的错误率。

4) 基于新维度的残差网络变体。基于增加网络深度或增加网络宽度的残差网络变体都取得了很好的效果,一些研究者对残差网络进行深入分析,提出了一系列基于新维度的残差网络变体。

Zhang 等人(2017)从结构多样性的维度出发,使用 Inception 模块代替残差单元,并通过多种形式的多项式组合推广了 inception residual 单元,构建了 PolyInception 模块,这种新的设计不仅增加了结构的多样性,也增强了残差组件的表达能力。

Xie 等人(2017)提出了一种简单、高度模块化的网络体系结构——ResNeXt,该模型通过重复聚合一系列具有相同拓扑结构的构建块来构建,该模型提供了一个新维度——基数,相比于增加深度和宽度,增加基数是一种更有效的获取精度提升的方法。

多尺度特征表示在分类任务中非常重要,从 AlexNet 到 ResNets,这些骨干卷积神经网络的进展不断展现出多尺度特征表示的重要作用,Gao 等人(2021)提出的 Res2Net 网络在更细粒度层次上提高了多尺度表征能力。该模块以瓶颈块为基础,应用了一个新的维度——尺度(scale),将 1×1 卷积的输出特征图在通道维度上均匀拆分为几个特征图子集,每个特征图子集进行卷积,并将得到的结果进行连接,再经过 1×1 卷积处理。Res2Net 在多个数据集上的图像分类任务中表现出更优的性能。

1.1.3 DenseNet 家族

Highway Network、残差网络以及随机深度训练方法都有一个共同点,即在网络中从高层到低层创建了越层连接路径。为确保网络中各层之间信息流最大化,Huang 等人(2017)提出了密集连接卷积神经网络(densely connected convolutional network, DenseNet),该网络使用了一种简单的连接模式,即将所有层直接相连。为进一步改善网络中的信息流,研究者们研究 DenseNet 路径连接的方式,提出了一系列 DenseNet 变体,例如双路径网络(dual path network, DPN)、CliqueNet 和 ConDenseNet 等。

DenseNet 由密集块组成,密集块结构如图 10 所示,密集块采用前馈的方式将所有层(具有相同输

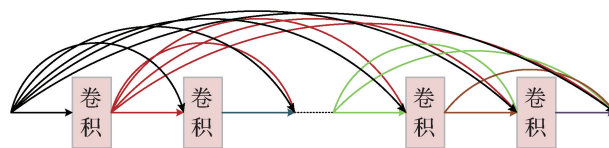


图 10 密集块示意图

Fig. 10 Dense block

出特征图大小)直接相连,每一层都从其前部所有层获得输入并将自己的输出特征图传递到后部层,这种方式增强了特征重用,并可缓解梯度消失问题。与残差网络不同的是,DenseNet 将特征图传递到下一层之前没有采用求和而是通过通道的合并来组合特征图。

对 ResNet 和 DenseNet 的拓扑的优点和缺陷进行研究,发现 ResNet 可增强特征重用,DenseNet 支持特征探索,为受益于两种拓扑结构,Chen 等人(2017)提出了双路径网络,该网络由类似瓶颈层的小模块组成,每个小模块中首先为 1×1 卷积层,紧接着为 3×3 卷积层,最后为 1×1 卷积层,最后 1×1 卷积层的输出分成两部分,第 1 部分逐元素相加到残差路径,第 2 部分连接到密集连接路径,这两部分的结果合并为下一个小模块的输入。DPN 具有较高的参数效率、更低的内存消耗,在图像分类和目标检测任务上获得了不错的性能。

残差网络和密集连接网络的成功归因于递归结构和迭代细化,然而残差路径和密集连接路径的迭代机制仅仅存在于相邻层或从高层到低层。Yang 等人(2018a)提出了 CliqueNet 网络,该网络将迭代机制发挥到极致,在同组任两层之间都存在正向和反向连接,在传播期间,将前几个层连接起来以更新下一层,然后将更新的层连接起来以重新更新先前层,这些层由此构成一个循环,交替更新,这也意味着 CliqueNet 比 DenseNet 连接更密集。CliqueNet 在多种图像分类数据集上以较少的参数量获得了优异的性能。

DenseNet 中每层都与其前部层相连以获得特征重用,但当前层的特征不被后部层所需要时,网络就产生了冗余。针对此,Huang 等人(2018)提出了 ConDenseNet,该模型将学习分组卷积与密集连接相结合,在训练时,学习分组卷积删除多余特征所对应层间的连接以使网络变得稀疏,在测试时,使用标准的分组卷积来实现。实验表明,CondenseNets 比当时最先进的紧凑型卷积网络更有效,可以应用在移

动设备中。

1.2 使用注意力机制的卷积神经网络

近年来,注意力机制是卷积神经网络研究又一焦点,人眼观看一幅图像,首先看全局,然后将注意力集中在某个细节,将注意力集中在有价值的部分,忽略价值不大的部分。使用注意力机制的卷积神经网络与人眼类似,强调目标中有用的部分,可以是某些空间、某些通道或某些层等,抑制价值不大的部分。主要的注意力机制可以分为空间域、通道域和层域注意力机制。

Wang 等人(2017)提出了使用注意力机制的卷积神经网络——残差注意力网络,该网络通过堆叠多个注意力模块来构建,在每个注意力模块中,掩码分支采用自底向上、自顶向下的前馈结构,将前馈和反馈注意过程展开为单个前馈过程,以使网络可进行端到端训练,同时生成特征的软权重。

Hu 等人(2018)专注于通道关系,提出了挤压激励模块(squeeze-and-excitation block, SE block),如图 11 所示,SE block 通过显式地建模通道之间的相互依赖性来重新校准通道的特征响应,即选择性地增强有用的通道特征,抑制无用的通道特征。该模块首先将全局空间信息挤压到通道描述符中,再将通道描述符通过门控机制,生成每个通道特征的权重,将权重与模块的输入相乘,完成通道特征重校准。SE block 的使用仅轻微增加模型的复杂度和计算量,却取得了较好的效果。Li 等人(2019)在 SE 模块基础上进行改进,提出了 SK(selective kernel)模块,与 SE 模块不同的是,SK 模块混合了两种不同大小的卷积核,获得了不同感受野的特征信息。

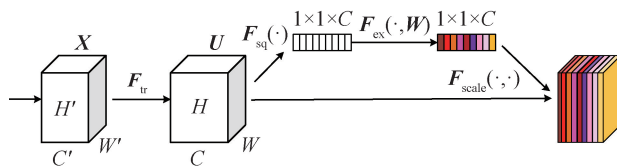


图 11 挤压激励模块(Hu 等,2018)

Fig. 11 SE block(Hu et al., 2018)

SE block 中,并没有考虑空间域特征的相关性。针对此问题,Woo 等人(2018)提出了卷积块注意力模块(convolutional block attention module, CBAM),该模块包括通道注意力模块和空间注意力模块两部分,输入特征图首先输入通道注意力模块,分别使用

平均池化和最大池化聚集空间信息生成两个空间内容描述符,随后两个空间内容描述符通过一个共享网络生成通道注意力图;通道注意力模块的输出特征图输入空间注意力模块,首先在通道维度上使用平均池化和最大池化操作,随后将两个操作的结果连接起来以生成特征描述符,特征描述符再经过卷积操作生成空间注意力图,空间注意力图与空间注意力模块的输入相乘,得到 CBAM 模块的输出,即同时考虑了通道域特征相关性和空间域特征相关性。

在网络末端插入全局二级池化(global second-order pooling, GSoP)以提高网络的非线性建模能力越来越受到关注,受 SE block 启发,Gao 等人(2019)提出了一种新的 GSoP 模块,并将其从低层引入到高层,其沿着通道维度或者空间维度捕获全局二阶统计信息,可方便地插入到现有网络中,在较小的开销下进一步提高其性能。

卷积操作是局部操作,重复局部操作具有计算效率低、优化困难和多跳依赖性建模困难等限制,针对此问题,Wang 等人(2018)提出了 Non-local block,利用非局部操作,捕获远程的依赖。该非局部操作将输入特征图的各个位置特征的加权和作为某一位置的响应。Non-local block 可以应用于多种神经网络架构。Cao 等人(2019)研究发现,Non-local 网络中不同查询位置的注意力图几乎相同,基于此发现,Cao 等人(2019)通过在所有查询位置使用相同的、独立于查询位置的注意力图,得到了简化的 Non-local 网络,使得网络计算量大大减小,随后,Cao 等人(2019)使用了与简化的 Non-local block 同样的上下文建模和融合方法、与 SE block 相同的转换方法,构建了全局上下文模块(global context block, GC block),GC block 在多种视觉识别任务中表现优于简化的 Non-local block 和 SE block。

Zhang 等人(2019b)在 DenseNet 上进行了注意力机制的探索,首先构建了通道特征重标定密集连接卷积神经网络(channel feature reweight DenseNet, CFR-DenseNet),使用挤压激励模块对通道特征重标定,随后构建了层间特征重标定密集连接卷积神经网络(inter-layer feature reweight DenseNet, ILFR-DenseNet),使用双挤压激励模块对层间特征进行重标定,最后将 CFR-DenseNet 和 ILFR-DenseNet 进行融合,构建了多级特征重标定密集连接卷积神经网络。

络模型 (multiple feature reweight DenseNet, MFR-DenseNet), MFR-DenseNet 网络结构如图 12 所示 (其中, SEM 为 squeeze-and-excitation module, DSEM 为 double squeeze-and-excitation module)。Zhang 等人(2019c)在 CFR-DenseNet 基础上添加了能够建模卷积特征点之间相互依赖关系的特征点重标定模块,构建了一种特征通道及特征点重标定密集连接

卷积神经网络模型(channel-wise and feature-points reweight DenseNet, CAPR-DenseNet)。针对 MFR-DenseNet 无法端到端训练、训练测试耗时长的问题,郭玉荣等人(2020)提出了可端到端学习的双通道特征重标定密集连接卷积神经网络(dual feature reweight DenseNet, DFR-DenseNet),同时实现了 DenseNet 的通道特征重标定和层间特征重标定。

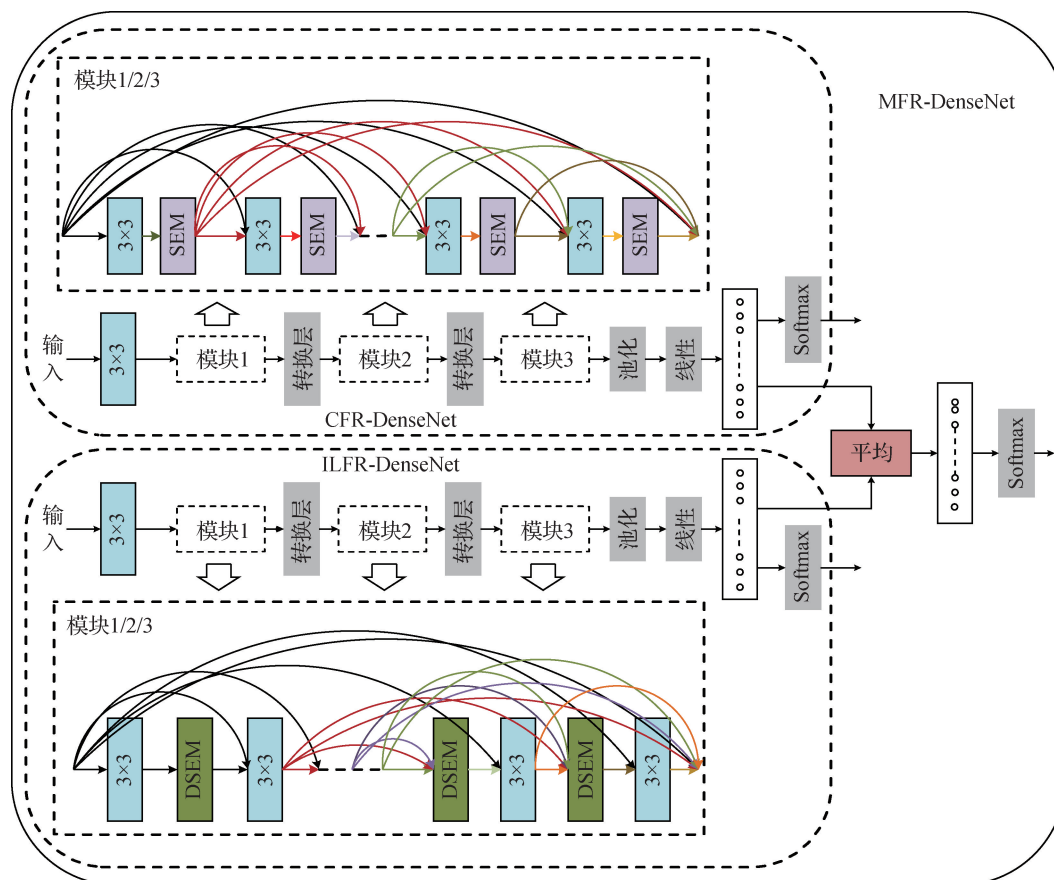


图 12 MFR-DenseNet 整体架构(Zhang 等,2019b)

Fig. 12 Structure of MFR-DenseNet(Zhang et al. , 2019b)

最近的多种计算机视觉任务都以 ResNet 为骨干网络,然而 ResNet 有限的感受野尺寸和跨通道交互的缺乏使得 ResNet 在分类以外的任务上表现较差,据此,Zhang 等人(2020)提出了 Split-Attention 模块,该模块沿着通道维度将特征图划分为几个组和更细粒度的分支,每个组的特征表示由其分支表示的加权组合表示,通过堆叠 Split-Attention 模块得到 ResNeSt 网络,该网络更易迁移到分类任务之外的其他任务。

1.3 轻量级网络

轻量级网络,顾名思义,即参数量小、计算复杂度低的网络。轻量级网络在训练期间需要较少的跨

服务器通信,传输到例如自动驾驶汽车设备上需要更少的传输带宽,在内存小的硬件上也可以存储。典型的轻量级网络有 SqueezeNet、Xception、MobileNet 系列和 ShuffleNet 系列等。

为了提高内存利用率和运行速度等,Iandola 等人(2016)提出了 SqueezeNet,该网络由 fire module 组成。fire module 分为挤压卷积层和扩展卷积层两部分,前者仅包含 1×1 卷积,卷积核滤波器个数为 $s_{1 \times 1}$,后者包含 1×1 卷积和 3×3 卷积操作,卷积核滤波器个数分别为 $e_{1 \times 1}$ 和 $e_{3 \times 3}$, $s_{1 \times 1}$ 远小于 $e_{1 \times 1}$ 与 $e_{3 \times 3}$ 的和。SqueezeNet 在达到与 AlexNet 同样精度的同时,参数量降低为 AlexNet 的五十分之一。

Inception 模块可解释为常规卷积和深度可分离卷积的中间步骤,受此启发,Chollet (2017) 提出了 Xception 模型,该模型将 Inception 模块替换为深度可分离卷积。深度可分离卷积首先在输入的每个通道上独立执行空间卷积,即深度卷积(depthwise convolution),随后将深度卷积输出的通道投影到新的通道空间,即逐点卷积(pointwise convolution)。Xception 模型比具有同样参数数量的 Inception v3 模型在大型数据集上表现更优异,其可以更有效地利用参数。

Howard 等人(2017)为移动和嵌入式应用提出了一种轻量级网络——MobileNetV1,该网络使用了深度可分离卷积,除此之外,还提出了两个超参数——宽度乘数 α 和决议乘数 ρ ,使得其可根据应用的不同选择不同的模型大小。Sandler 等人(2019)提出的 MobileNetV2 继续使用 MobileNetV1 中的深度可分离卷积,并在此基础上提出了使用倒置残差和线性瓶颈的模块。Howard 等人(2019)将神经架构搜索(neural architecture search, NAS)与网络结构的设计结合,提出了 MobileNets 系列的第3代模型——MobileNetV3,该网络除了使用 NAS 算法和 NetAdapt 算法来优化模型,还重新设计了瓶颈模块,并在网络的后半部分使用了新的激活函数 h-swish。

针对计算能力有限的移动设备,Zhang 等人(2018b)使用分组卷积和通道洗牌操作构建了 ShuffleNetV1 网络,在保持模型精度的同时大大减少了计算量。

评估以上轻量级网络的计算复杂度往往使用间接指标——浮点运算数(floating point operations, FLOPs),但使用间接指标与使用直接指标有差异, Ma 等人(2018)通过对使用间接指标和直接指标之间的差异进行实验与分析,提出了4条轻量级网络设计的指导方针:输入输出通道数相同使内存访问成本最小;过多的分组卷积增加内存访问成本;网络碎片降低并行度;应减少逐元素操作。根据4条指导方针, Ma 等人(2018)在 ShuffleNetV1 的基础上,引入通道分离操作,提出了 ShuffleNetV2,其不仅有效,而且准确。

1.4 NAS

在过去的几年里,虽然深度卷积神经网络在图像分类、语义分割等领域等取得巨大的成功,但是深

度卷积神经网络的结构设计仍需要许多专业知识和大量时间,近年来,神经架构搜索(neural architecture search, NAS)方法逐渐成为研究的热点之一,其采用神经网络自动设计神经网络结构。NAS 方法可分为3类:1)基于设计不同搜索空间的 NAS 方法;2)基于模型优化的 NAS 方法;3)其他改进的 NAS 方法。

1)基于设计不同搜索空间的 NAS 方法。卷积神经网络的结构和连通性可由可变长度的字符串来指定,据此,Zoph 和 Le(2016)提出使用循环网络作为控制器来生成这样的字符串,训练这些字符串对应的卷积神经网络得到验证集上的准确率,准确率作为奖励信号,使用策略梯度来优化控制器,以使控制器使用更合适的参数来生成更准确的模型。NAS 的搜索方法直接应用在大数据集上,计算量非常大。针对这个问题,Zoph 等人(2018)设计了新的搜索空间,搜索空间包括结构相同参数不同的卷积层,搜索最佳单元架构就转变成搜索最佳结构体,在小数据集上搜索网络构建块,然后迁移到大数据集上。搜索最佳结构体比搜索整个网络结构快得多,并且通过简单地改变卷积单元数量和卷积单元中滤波器的数量,可以创建不同计算需求的网络架构。

Bello 等人(2017)提出神经优化器搜索,搜索空间中为优化器各种因素的组合,得到两个新的更新规则 PowerSign 和 AddSign。Ramachandran 等人(2017)使用自动搜索技术发现了新的激活函数 Swish,在许多富有挑战性的数据集上,其比使用广泛的 ReLU 激活函数效果更好。

数据增强在图像分类器的准确性中起到至关重要的作用,但是数据增强是手工实现的,Cubuk 等人(2019)提出了 AutoAugment 过程,可以自动搜索改进的数据增强策略。

使用传统进化算法设计的分类器尚未达到人类专家设计的分类器的准确性,Real 等人(2019)对演进算法进行改进,首次开发出超越手工设计的分类器 AmoebaNet-A。改进有两点:1)为每个架构添加年龄属性,倾向于搜索年轻的模型;2)搜索空间为分类器,此搜索空间将卷积神经网络与有向图相关联,顶点表示隐藏状态,标签的边缘表示常见的网络操作,提出的变异规则仅通过将边缘的原点随机重新连接到不同的顶点,并通过随机重新标记边缘(覆盖整个搜索空间)来更改体系结构。

2) 基于模型优化的 NAS 方法。为提高神经架构搜索的有效性, 研究者们提出了一系列神经架构搜索的优化方法。

Liu 等人(2018)提出使用基于序列模型的优化策略, 在该策略中, 以复杂度递增的顺序搜索模型, 同时学习替代模型以指导在结构空间中进行搜索。

NAS 逐个训练每个子网络以收敛, 测量每个子网络的准确率时也丢弃了每个网络的权重, 因此 NAS 的计算量很昂贵。Pham 等人(2018)提出的 efficient neural architecture search (ENAS) 强制所有子网络共享权重, 以避免从头到尾地训练每个子网络, 与当时某些自动设计的模型相比, 使用 GPU 的时间少得多。

针对直接指标和间接指标有差异的问题, 以及特定平台网络模型难以迁移到其他平台的问题, Yang 等人(2018b)提出了 NetAdapt 优化算法, 贡献主要有两点: 该算法在优化循环中加入了直接度量, 直接度量是由从目标平台获得的经验度量来评估的, 这使得算法可以适应任何平台; 该算法为自动约束网络优化算法, 在满足约束条件时, 使精度最大化。

神经架构搜索效率低下的原因可归结为其为一个黑盒在离散的搜索空间中进行优化, Liu 等人(2019)提出了可微架构搜索 (differentiable architecture search, DARTS), DARTS 不再搜索候选的离散模型, 而将搜索空间设置为连续且可微的, 因此网络在验证集上可使用梯度下降来优化, 使用较少的资源能获得不错的性能。

为进一步设计出在移动设备应用的卷积神经网络, Tan 等人(2019a)提出了移动神经架构搜索 (mobile neural architecture search, MNAS), 其创新点有两点: 将模型的准确率和延迟作为奖励信号, 其中延迟通过在真实移动设备上执行模型来直接测量, 而不是使用不准确的间接指标, 例如 FLOPs; 提出了新的分解层次搜索空间, 避免了之前自动搜索方法中单元类型少的问题。

3) 其他 NAS 方法。在有资源预算的情况下, 扩展模型可获得更好的精度, Tan 等人(2019b)提出了复合扩展方法, 相比于之前的方法, 该方法用一组固定的比例系数来均匀地缩放网络的宽度、深度和分辨率。扩展模型的有效性取决于基础模型, Tan 等人(2019b)使用神经架构搜索设计出了一种新的基础模型, 使用复合扩展方法将该模型扩展, 得到一系

列模型, 即 EfficientNets。

尽管自动搜索技术很有效, 但得到的模型是单个的, 并有一定的巧合性, Radosavovic 等人(2020)旨在发现模型结构设计规则, 提出设计模型的设计空间, 其中设计空间是模型结构的参数化表示, 得到了由规则、简单、优秀网络组成的 RegNet 设计空间, RegNet 设计空间可以泛化为更大的计算状态、调度长度和网络块类型, 在相同训练设置下, RegNet 模型比 EfficientNets 模型快 5 倍。

神经架构搜索除了应用在图像分类任务上, 还应用在目标检测任务上, 例如 Ghiasi 等人(2019)提出的 NAS-FPN, Zoph 等人(2019)提出的基于神经架构搜索的数据增强策略。

1.5 小结

DCNN 极大地推进了图像分类任务的进展。DCNN 通过多层的非线性变换, 在大量图像训练数据中提取特征以代替手工提取的特征, 并且深层次的卷积神经网络具有极强的特征学习和表达能力, 能够学习到复杂的全局信息和上下文信息。此外, DCNN 采用一些手段可达到轻量化的效果, 方便应用于移动式设备和嵌入式设备中。还可使用神经网络自动设计和优化 DCNN, 得到性能较优异的模型的同时节省了人力和时间。

除 DCNN 应用于图像分类外, 循环神经网络、图神经网络也可用于图像分类中。CNN 更加关注局部特征, 若标签之间具有很强的关联性, CNN 视野有限, 不会很好地处理这类任务。循环神经网络 (recurrent neural network, RNN) 添加了反馈单元, 将模型当前位置的输出反馈给模型, 从而帮助下一位置进行决策。LSTM 作为 RNN 的一种, 能够解决 RNN 无法处理长距离的依赖问题, 还能够缓解 RNN 梯度爆炸或消失问题, 在处理长序列数据上非常有效。图神经网络 (graph neural network, GNN) 是一个较新的研究领域, 其比 CNN 更擅于因果推理。图像也可转化成图, 对图的分析至关重要, GNN 是直接应用于图的神经网络, 对图级进行分类, 在图像分类中有一定的应用。

2 数据集

深度卷积神经网络模型在数据集上进行训练和测试, 依据在数据集上得到的分类结果来衡量模型

的优劣,因此使用好的数据集至关重要。在本文所综述的深度卷积神经网络模型中,只有 LeNet-5 使用了 MNIST 数据集,其他模型大多使用了 ImageNet 数据集、CIFAR-10/100 数据集和 SVHN 数据集,本节介绍 MNIST 数据集、CIFAR-10 数据集、CIFAR-100 数据集和 SVHN 数据集。

2.1 MNIST 数据集

MNIST(modified NIST)数据集用来识别手写数字,从 NIST 数据集的 SD-1(special dataset 1)和 SD-3(special dataset 3)构建的,其中包含手写数字的二进制图像。

NIST 数据集将 SD-3 作为训练集,将 SD-1 作为测试集,但 SD-3 比 SD-1 更易识别,原因在于 SD-3 来源于人口普查局雇员,SD-1 来源于高中生。以上表明分类结果要独立于完整样本集中训练集和测试集的选择,因此,通过混合 NIST 数据集来建立新的数据集很有必要。SD-1 有 58 537 幅图像,由 500 位作者书写,排列混乱,SD-3 中图像是顺序的。新的训练集共有 60 000 幅图像,一部分来源于 SD-1 中前 250 位作家所书写,剩余部分来源于 SD-3。新的测试集有 60 000 幅图像,一部分来源于 SD-1 剩余 250 位作家所书写,一部分来源于 SD-3。新数据集称为 MNIST 数据集。在 LeNet-5 实验中,训练集为完整的新训练集,共 60 000 幅图像,测试集为新测试集的子集,共 10 000 幅图像。

2.2 ImageNet 数据集

ImageNet 数据集是具有超过 1 500 万幅带标签的高分辨率图像的数据集,这些图像大约属于 22 000 个类别,这些图像从互联网收集并由人工使用亚马逊的机械土耳其众包工具贴上标签。从 2010 年开始,每年举行一次名为 ImageNet 大规模视觉识别挑战赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)的竞赛,ILSVRC 使用 ImageNet 的子集,共 1 000 个类别,每个类别大约包含 1 000 幅图像,总共大约有 120 万幅训练图像,50 000 幅验证图像和 150 000 幅测试图像。

深度卷积神经网络模型在 ImageNet 数据集上进行训练和测试,衡量模型优劣的指标为 top-5 错误率和 top-1 错误率。ImageNet 通常有 1 000 个类别,训练和测试时,对每幅图像同时预测 5 个标签类别,若预测的 5 个类别任意之一为该图像的正确标签,则

视为预测正确,那么预测错误的概率为 top-5 错误率;若对图像预测一次,预测错误的概率为 top-1 错误率。

2.3 CIFAR-10/100 数据集

CIFAR-10 数据集有 60 000 幅彩色图像,分辨率大小为 32×32 像素,共 10 个类别,每个类别包含 6 000 幅图像。训练集包含 50 000 幅彩色图像,测试集包含 10 000 幅彩色图像。测试集的图像取自 10 个类别,每个类别分别取 1 000 幅,剩余的图像构成训练集。

CIFAR-100 数据集与 CIFAR-10 数据集类似,不同的是 CIFAR-100 数据集有 100 个类别,每个类别包含 600 幅图像,每个类别有 500 幅训练图像和 100 幅测试图像。

2.4 SVHN 数据集

SVHN 数据集用来检测和识别街景图像中的门牌号,从大量街景图像的剪裁门牌号图像中收集,包含超过 600 000 幅小图像,这些图像以两种格式呈现:一种是完整的数字,即原始的、分辨率可变的、彩色的门牌号图像,每个图像包括检测到的数字的转录以及字符级边界框;一种是剪裁数字,图像尺寸被调整为固定的 32×32 像素。

SVHN 数据集分为 3 个子集,73 257 幅图像用于训练,26 032 幅图像用于测试,531 131 幅难度稍小的图像作为额外的训练数据。

3 实验结果分析

图像分类模型主要在 ImageNet、CIFAR-10、CIFAR-100 和 SVHN 数据集上进行训练和测试,本文整理了各个模型在以上 4 个数据集上的实验结果(表 1),其中 ImageNet 数据集为 2012 年 ILSVRC 使用的版本。所整理的实验结果来源于相应的参考文献,参考文献中没有相应实验结果的,以“-”表示。

实验结果的描述主要包括图像分类的错误率和模型的参数量。在卷积神经网络发展的初期,研究者们致力于提高网络的分类精度,深度卷积神经网络发展愈发成熟之际,为节约训练测试的时间和内存的使用,不仅要求深度卷积神经网络具有较高的图像分类精度,还要求其具有较小的参数量。衡量轻量级网络优劣的度量除准确率、参数量外,还有浮点运算数(FLOPs)。

表 1 基础网络模型在不同数据集上的错误率
Table 1 The error rate of the basic DCNNs on different datasets

模型	数据集				
	ImageNet		CIFAR-10	CIFAR-100	SVHN
	top-1 验证集(测试集)错误率	top-5 验证集(测试集)错误率			
AlexNet(Krizhevsky 等,2012)(单个网络)	40. 70(-)	18. 20(-)	-	-	-
AlexNet(Krizhevsky 等,2012)(5 个网络)	38. 10(-)	16. 40(16. 40)	-	-	-
OverFeat(Sermanet 等,2014)(单个网络)	35. 74(-)	14. 18(-)	-	-	-
OverFeat(Sermanet 等,2014)(7 个网络)	33. 96(-)	13. 24(13. 60)	-	-	-
ZFNet(Zeiler 和 Fergus,2014)(单个网络)	37. 50(-)	16. 00(16. 10)	-	-	-
ZFNet(Zeiler 和 Fergus,2014)(6 个网络)	36. 00(-)	14. 70(14. 80)	-	-	-
Network in Network(Lin 等,2013)	-	-	8. 81	35. 68	1. 94
GoogLeNet(Szegety 等,2015)(单个网络)	-	7. 90(7. 90)	-	-	-
GoogLeNet(Szegety 等,2015)(7 个网络)	-	6. 67(6. 67)	-	-	-
VGGNet(Simonyan 和 Zisserman,2015)(单个网络)	24. 40(-)	7. 10(7. 00)	-	-	-
VGGNet(Simonyan 和 Zisserman,2015)(2 个网络)	23. 70(-)	6. 80(6. 80)	-	-	-
Inception-v2(Ioffe 和 Szegedy,2015)(单个网络)	25. 20(-)	7. 82(-)	-	-	-
Inception-v2(Ioffe 和 Szegedy,2015)(模型融合)	20. 10(-)	4. 9(4. 82)	-	-	-
Inception-v3(Szegedy 等,2016a)(单个网络)	18. 77(-)	4. 20(-)	-	-	-
Inception-v3(Szegedy 等,2016a)(网络融合)	17. 20(-)	3. 46(3. 58)	-	-	-
Highway Network(Srivastava 等,2015)	-	-	7. 72	32. 39	-
Inception-v4(Szegedy 等,2016b)(单个网络)	17. 70(-)	3. 80(-)	-	-	-
Inception-resnet-v1(Szegedy 等,2016b)(单个网络)	18. 80(-)	4. 30(-)	-	-	-
Inception-resnet-v2(Szegedy 等,2016b)(单个网络)	17. 80(-)	3. 70(-)	-	-	-
FractalNet(Larsson 等,2017)	24. 12(-)	7. 39(-)	4. 59	23. 36	1. 87

注：“-”表示对应文献中无数据。

1) 如表 1 所示,基础网络模型中,从 LeNet 到 GoogLeNet,图像分类准确率越来越高,这是因为网络越深,提取的特征越有效。

2) 随着网络加深,梯度消失问题越来越严重,ResNet 利用越层连接,使网络高层特征向低层传递,缓解梯度消失问题的同时大大提高了分类任务的性能。为进一步提升残差网络的性能,研究者们提出了一系列残差网络的变体。例如,He 等人(2016b)改变了残差块中 BN、ReLU、Conv 的顺序,提出了 Pre-ResNet; Han 等人(2017)通过线性逐步增加残差块中特征图的通道数,构建了金字塔残差网络; Zhang 等人(2018a)在原始残差网络的基础上,逐级加入越层连接,构建了 RoR; Zhang 等人(2019a)在 RoR 的基础上,线性逐步增加每个残差块中特征图的通道数,构建了金字塔 RoR。这些方

法表明改善层与层之间信息的流动非常重要,同时表明了层与层之间加入短路径的重要性。此外,为使深度残差网络更易优化,研究者们提出了新的训练方法,例如, Huang 等人(2016)提出的随机深度以一定概率丢弃不同的残差块,缩短了训练时间,还有 Singh 等人(2016)提出的随机训练方案 Swapout,不难发现,这些训练方法旨在减少训练时模型的复杂度。此外, Zagoruyko 和 Komodakis(2016)在原始残差块的基础上成倍地增加残差块中卷积核的个数构建了 WRN; Abdi 和 Nahavandi(2017)通过增加每个残差块中残差函数的数量构建了 Multi-ResNet,这些方法都增加了网络的宽度,在保持深度不变的同时提升了残差网络的性能,ResNet 及其变体在 ImageNet, CIFAR, 以及 SVHN 数据集上的实验结果如表 2 和表 3 所示。

表 2 ResNet 家族在 ImageNet 数据集上的错误率和参数量

Table 2 The error rate and parameter amounts of the Residual families on ImageNet datasets

模型	ImageNet		参数量/M
	top-1 错误率/%	top-5 错误率/%	
ResNet-152(He 等,2016a)	19.38	4.49	—
Pre-ResNet-200(He 等,2016b)	20.10	4.80	64.50
Pyramidal Residual Networks-200(Han 等,2017)	19.20	4.70	116.40
RoR-3-152(Zhang 等,2018a)	20.55	5.14	—
SD-ResNet-152(Huang 等,2016)	21.78	—	—
WRN-50-2(Zagoruyko 和 Komodakis,2016)	21.90	6.03	68.90
multi-resnet-101(Abdi 和 Nahavandi,2017)	19.93	—	—
ResNeXt-101(Xie 等,2017)	19.10	4.40	—
PolyNet(极深)(Zhang 等,2017)	17.36	3.45	—
Res2 Net-101(Gao 等,2021)	20.81	5.57	—

注:仅最后一行 20.81 和 5.57 为 ImageNet 数据集上的测试错误率,其他为在 ImageNet 验证集上的错误率,“—”表示对应文献中无相关数据。

表 3 ResNet 家族在 CIFAR 数据集和 SVHN 数据集上的错误率和参数量

Table 3 The error rate and parameter amounts of the Residual families on CIFAR and SVHN

模型	CIFAR-10/%	CIFAR-100/%	SVHN/%	参数量/M
ResNet-110(He 等,2016a)	6.43	—	—	1.70
Pre-ResNet-1001(He 等,2016b)	4.62	22.71	—	10.20
Weighted ResNets-1192(Shen 等,2016)	4.70	—	—	19.3
Pyramidal Residual Networks-272(Han 等,2017)	3.31	16.35	—	26.00
PyramidalSepDrop-182(Yamada 等,2016)	3.31	16.18	—	—
RoR-3-WRN58-4 + SD(Zhang 等,2018a)	3.77	19.73	1.59	13.30
Pyramidal RoR + (146 层)(Zhang 等,2019a)	2.96	16.40	1.59	38.00
SD-ResNet-110(Huang 等,2016)	5.23	24.58	1.75	1.70
Swapout-32(Singh 等,2016)	4.76	22.72	—	7.43
CRMN-32(Moniz 和 Pal,2016)	4.16	19.79	1.68(16 层)	>40
WRN-28-10(Zagoruyko 和 Komodakis,2016)	4.00	19.25	—	36.50
18-layer + wide RiR(Targ 等,2016)	5.01	22.90	—	10.30
multi-resnet-200(Abdi 和 Nahavandi,2017)	4.35	20.42	—	10.20
ResNeXt-29(Xie 等,2017)	3.58	17.31	—	68.10
PolyNet(极深)(Zhang 等,2017)	—	—	—	—
Res2 NeXt-29(Gao 等,2021)	—	16.56	—	36.90

注:“—”表示对应文献中无相关数据。

3)Huang 等人(2017)提出的 DenseNet 进一步提升了深度卷积神经网络的性能,究其原因,在 DenseNet 的每个密集块中,每层的特征图都与其前部层的特征图直接相连,并将自己的特征图传递给所有的后部层,从

而比 ResNet 进一步改善了层与层之间信息流的传播,再次突出了短路径的重要性。随后,一系列 DenseNet 变体提出,继续提升了 DenseNet 的性能。DenseNet 家族在不同数据集的实验结果如表 4 所示。

表 4 DenseNet 家族在不同数据集上的错误率

Table 4 The error rate of the DenseNet families on different datasets

模型	ImageNet		CIFAR-10	CIFAR-100	SVHN
	top-1 错误率	top-5 错误率			
DenseNet (Huang 等, 2017)	20.85	5.30	3.46	17.18	1.59
DPN (Chen 等, 2017)	18.55	4.16	—	—	—
CliqueNet (Yang 等, 2018a)	24.01	7.15	5.06	23.14	1.51
CondenseNet (Huang 等, 2018)	26.20	8.30	3.46	17.55	—

注: “—”表示对应文献中无相关数据。

4) 采用注意力机制的深度卷积神经网络也取得不错的性能,除了准确率保持甚至超越经典的深度卷积神经网络,参数量也大大减少。使用注意力机制的深度卷积神经网络在 CIFAR 上的实验结果如表 5 所示。例如, Hu 等人(2018)提出的 SE 模块建模通道特征之间的依赖性, Wang 等人(2018)提出的 non-local block 利用非局部操作来构建,捕获远程的依赖, Zhang 等人(2019b)提出的 MFR-DenseNet 首

先进行通道特征重标定再进行层间特征重标定, Zhang 等人(2019c)构建的 CAPR-DenseNet 在 CFR-DenseNet 基础上添加了能够建模卷积特征点之间相互依赖关系的特征点重标定模块, Zhang 等人(2020)提出的 DFR-DenseNet 使通道特征重标定和层间特征重标定同时进行,并且大大减少了参数量。从以上方法可以看出,注意力机制实质强调有用的部分,忽略用处不大的部分,使模型提取的特征更有效。

表 5 使用注意力机制的深度卷积神经网络在 CIFAR 数据集上的错误率和参数量

Table 5 The error rate and parameter amounts of the DCNNs using attention mechanism on CIFAR

模型	ImageNet		参数量/M	CIFAR-10/%	CIFAR-100/%	参数量/M
	top-1 错误率/%	top-5 错误率/%				
残差注意力网络 (Wang 等, 2017)	19.50	4.80	51.30	3.90	20.45	8.60
SENet (Hu 等, 2018)	16.88	3.58	—	—	—	—
ResNeXt101 + CBAM (Woo 等, 2018)	21.07	5.59	48.96	—	—	—
MFR-DenseNet (Zhang 等, 2019b)	—	—	—	3.57	18.27	14.20
DFR-DenseNet (郭玉荣等, 2020)	—	—	—	4.29	21.86	2.86
CAPR-DenseNet (Zhang 等, 2019c)	—	—	—	3.86	20.16	7.11
skNet (Li 等, 2019)	18.40	—	48.90	—	—	—
ResNeSt (Zhang 等, 2020)	15.50	—	111.00	—	—	—
GCNet (Cao 等, 2019)	24.00	7.66	42.45	—	—	—
GSoP-Net2 (Gao 等, 2019)	21.19	5.64	58.70	—	18.85	3.60

注: “—”表示对应文献中无相关数据。

5) 轻量级网络在 ImageNet 数据集上的实验结果如表 6 所示,由表 6 可知,大多数轻量级网络在 ImageNet 数据集上进行图像分类实验, ShuffleNet 取得分类最好的结果,虽然大多数轻量级网络性能不及经典的深度卷积神经网络,但是参数量却大大减少。Chollet(2017)提出的 Xception 模型使用了深度可分离卷积; MobileNets 系列主要使用了深度可分离

卷积, MobileNetV1 还提出了两个超参数来衡量模型的大小, MobileNetV2 提出了倒置残差和线性瓶颈, MobileNetV3 使用了 NAS 和 NetAdapt 算法来优化模型; Zhang 等人(2018b)提出的 ShuffleNetV1 使用了分组卷积和通道洗牌操作, Ma 等人(2018)对描述模型复杂度的直接指标和间接指标之间的差异进行研究,提出了 ShuffleNetV2。从以上方法可以看出,使用深

度可分离卷积、分组卷积、通道洗牌操作,或者提出新的优化方法,都是为了使参数得到有效的利用。

表6 轻量级网络在 ImageNet 数据集上的错误率、参数量和 FLOPs

Table 6 The error rate, parameter amounts and FLOPs of the lightweight DCNNs on ImageNet

模型	ImageNet		参数量/M	FLOPs/M
	top-1 错误率/%	top-5 错误率/%		
SqueezeNet(Iandola 等,2016)	39.60	17.50	—	—
Xception(Chollet,2017)	21.00	5.50	22.86	—
MobileNetV1(Howard 等,2017)	29.40	—	4.20	569
MobileNetV2(1.4)(Sandler 等,2019)	28.00	—	3.40	300
MobileNetV3(V3-Large 1.0)(Howard 等,2019)	24.80	—	5.40	—
ShuffleNetV1(Zhang 等,2018b)	24.70	4.70	—	527
SE-ShuffleNet v2-164(Ma 等,2018)	18.56	—	—	12 700

注:“—”表示所引用文献中无相关数据。

6)利用自动搜索技术设计的深度卷积神经网络逐步超越了人工设计的深度卷积神经网络的性能。NASNet、PNASNet、EfficientNets 曾在一些数据集上获得当时最好的图像分类性能,Ramachandran 等人(2017)使用神经网络搜索设计出比 ReLU 性能更好的激活函数。神经架构搜索在不同数据集

上的结果如表7所示。分析神经架构搜索设计的思路,可以发现,模型的自动设计中,可以使用自动搜索技术来搜索模型的构建块、激活函数、分类器、数据增强策略和基础模型,也可使用自动搜索技术搜索模型的优化方法,以提高网络设计的速度。

表7 神经架构搜索在不同数据集上的错误率

Table 7 The error rate and parameter amounts of the NAS on different datasets

模型	ImageNet		CIFAR-10	CIFAR-100
	top-1 错误率	top-5 错误率		
NAS(Zoph 和 Le,2016)	—	—	3.65	—
NASNet(Zoph 等,2018)	17.30	3.80	2.40	—
AddSign-ld(Bello 等,2017)	26.20	8.40	4.50	—
PowerSign-cd(Bello 等,2017)	26.10	8.10	8.00	—
Inception-ResNet-v2(Swish)(Ramachandran 等,2017)	19.60	4.80	—	—
DenseNet(Swish)(Ramachandran 等,2017)	—	—	5.20	16.20
PNASNet(Liu 等,2018)	17.10	3.8	3.41	—
ENAS(Pham 等,2018)	—	—	2.89	—
AmoebaNet-A(Real 等,2019)	16.10	3.40	3.40	—
MobileNetV2(NetAdapt)(Yang 等,2018b)	29.10	—	—	—
AutoAugment(Cubuk 等,2019)	16.5	3.50	1.50	10.70
DARTs(Liu 等,2019)	26.70	8.70	2.76	—
MNASNet(Tan 等,2019a)	24.80	7.50	—	—
EfficientNets(Tan 等,2019b)	15.60	2.90	1.10	8.30
RegNet(Radosavovic 等,2020)	20.10	—	—	—

注:仅 DARTs 结果中的 26.70 和 8.70 为 ImageNet 数据集上的测试错误率,其他未指明是验证集还是测试集上的结果,“—”表示所引用文献中无相关数据。

4 结 语

人的神经系统由上亿个神经元构成,人类每天接收无数的视觉、听觉信息,都由神经系统处理并传出。深度卷积神经网络与人的神经系统类似,是一种信息处理系统。作为深度学习的重要方法,深度卷积神经网络广泛应用于计算机视觉任务,尤其图像分类任务中,其使用卷积自动提取特征,判断图像类别。图像分类模型愈发成熟,本文围绕着图像分类的深度卷积神经网络模型结构,首先介绍了从 LeNet 到 GoogLeNet 的基础模型、ResNet 家族、DenseNet 家族、使用注意力机制的深度卷积神经网络模型、轻量级网络和神经架构搜索六个部分,随后介绍了深度卷积神经网络进行图像分类实验常用的数据集,包括 ImageNet 数据集、CIFAR-10/100 数据集、SVHN 数据集和 MNIST 数据集,最后对所有的模型进行了实验结果的总结和分析。

深度卷积神经网络在图像分类任务上大放异彩,但其与人类神经系统不可比拟,远远未达到人类的认知水平,存在以下困难和挑战:

1) 深度卷积神经网络进行图像分类是监督式学习,即需要给每幅图像贴上标签,“告诉”模型图像类别,以学习此类别图像的特征,再用学习到的特征判断未知类别的图像。而人类大脑具有思考能力,更加灵活,更多的是无监督式学习,人类视觉系统判断一幅图像之前并不需要被告知几千上万图像类别。因此,如何使模型可以像人脑具有思考能力也许是未来的研究方向,非监督式学习是未来的发展前景。

2) 深度卷积神经网络模型应用于移动设备具有挑战性,这是因为需要保持模型准确率的同时还要考虑模型的大小以及速度。设计应用于移动设备的深度卷积神经网络模型有待进一步研究。

3) 近年来,提升模型的效率成为研究热点,但是多数研究都是基于优化间接度量来简化网络的,例如 MACC(memory accesses)值、FLOPs、权重数量,虽然间接指标易于运算且易于集成到优化框架,但间接指标的优化不一定使直接指标减少,例如模型速度和网络资源消耗,直接指标和间接指标的差异与平台也有关系。

4) 人工设计网络耗时耗力,近年来神经架构搜索非常流行,是未来网络设计的发展趋势,但是其训

练时间长,内存资源耗费非常多,对研究者的实验环境有一定的要求。

参考文献 (References)

- Abdi M and Nahavandi S. 2017. Multi-residual networks: improving the speed and accuracy of residual networks [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/1609.05672.pdf>
- Bello I, Zoph B, Vasudevan V and Le Q V. 2017. Neural optimizer search with reinforcement learning//Proceedings of the 34th International Conference on Machine Learning (ICML). Sydney, Australia; [s. n.]: 459-468
- Cao Y, Xu J R, Lin S, Wei F Y and Hu H. 2019. GCNet: non-local networks meet squeeze-excitation networks and beyond//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul, Korea (South): IEEE: 1971-1980 [DOI: 10.1109/ICCVW.2019.00246]
- Chen Y P, Li J N, Xiao H X, Jin X J, Yan S C and Feng J S. 2017. Dual path networks//Advances in Neural Information Processing Systems 30. Long Beach, USA; [s. n.]: 4467-4475
- Chollet F. 2017. Xception: deep learning with depthwise separable convolutions//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 1800-1807 [DOI: 10.1109/CVPR.2017.195]
- Cubuk E D, Zoph B, Mané D, Vasudevan V and Le Q V. 2019. Auto-Augment: learning augmentation strategies from data//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 113-123 [DOI: 10.1109/CVPR.2019.00020]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR.2009.5206848]
- Gao S H, Chen M M, Zhao K, Zhang X Y, Yang M H and Torr P. 2021. Res2Net: a new multi-scale backbone architecture. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(2): 652-662 [DOI: 10.1109/TPAMI.2019.2938758]
- Gao Z L, Xie J T, Wang Q L and Li P H. 2019. Global second-order pooling convolutional networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 3024-3033 [DOI: 10.1109/CVPR.2019.00314]
- Ghiasi G, Lin T Y and Le Q V. 2019. NAS-FPN: learning scalable feature pyramid architecture for object detection//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 7036-7045 [DOI: 10.1109/CVPR.2019.00720]
- Gu J X, Wang Z H, Kuen J, Ma L Y, Shahroudy A, Shuai B, Liu T,

- Wang X X, Wang G, Cai J F and Chen T. 2018. Recent advances in convolutional neural networks. *Pattern Recognition*, 77: 354-377 [DOI: 10.1016/j.patcog.2017.10.013]
- Guo Y R, Zhang K, Wang X S, Yuan J S, Zhao Z B and Ma Z Y. 2020. Image classification method based on end-to-end dual feature reweight DenseNet. *Journal of Image and Graphics*, 25(3): 486-497 (郭玉荣, 张珂, 王新胜, 苑津莎, 赵振兵, 马占宇. 2020. 端到端双通道特征重标定 DenseNet 图像分类. 2020. 中国图象图形学报, 25(3): 486-497) [DOI: 10.11834/jig.190290]
- Han D, Kim J and Kim J. 2017. Deep pyramidal residual networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 6307-6315 [DOI: 10.1109/CVPR.2017.668]
- He K M, Zhang X Y, Ren S Q and Sun J. 2015. Delving deep into rectifiers: Surpassing human level performance on imagenet classification//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE: 1026-1034 [DOI: 10.1109/ICCV.2015.123]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016a. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016b. Identity mappings in deep residual networks//Proceedings of the 14th European Conference on Computer Vision (ECCV). Amsterdam, The Netherlands: Springer: 630-645 [DOI: 10.1007/978-3-319-46493-0_38]
- Howard A, Sandler M, Chen B, Wang W J, Chen L C, Tan M X, Chu G, Vasudevan V, Zhu Y K, Pang R M, Adam H and Le Q. 2019. Searching for MobileNetV3//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE: 1314-1324 [DOI: 10.1109/ICCV.2019.00140]
- Howard A G, Zhu M L, Chen B, Kalenichenko D, Wang W J, Weyand T, Andreetto M and Adam H. 2017. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/1704.04861.pdf>
- Hu J, Shen L and Sun G. 2018. Squeeze-and-excitation networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE: 7132-7141 [DOI: 10.1109/CVPR.2018.00745]
- Huang G, Liu S C, van der Maaten L and Weinberger K Q. 2018. CondenseNet: an efficient DenseNet using learned group convolutions//Proceedings of 2018 Conference on computer vision and pattern recognition (CVPR). Salt Lake City, USA: IEEE: 2752-2761 [DOI: 10.1109/cvpr.2018.00291]
- Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Huang G, Sun Y, Liu Z, Sedra D and Weinberger K Q. 2016. Deep networks with stochastic depth//Proceedings of the 14th European Conference on Computer Vision (ECCV). Amsterdam, The Netherlands: Springer: 646-661 [DOI: 10.1007/978-3-319-46493-0_39]
- Iandola F N, Han S, Moskewicz M W, Ashraf K, Dally W J and Keutzer K. 2016. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size [EB/OL]. [2020-05-22]. <http://arxiv.org/pdf/1602.07360.pdf>
- Ioffe S and Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift//Proceedings of the 32nd International Conference on Machine Learning (ICML). Lille, France: [s. n.]: 448-456
- Khan A, Sohail A, Zahoor U and Qureshi A S. 2019. A survey of the recent architectures of deep convolutional neural networks [EB/OL]. [2020-08-10]. <https://arxiv.org/pdf/1901.06032.pdf>
- Krizhevsky A. 2009. Learning Multiple Layers of Features from Tiny Images. Technical Report TR-2009, University of Toronto, Toronto
- Krizhevsky A, Sutskever I and Hinton G E. 2012. Imagenet classification with deep convolutional neural networks//Proceedings of the 26th Conference on Neural Information Processing Systems (NIPS). Lake Tahoe, USA: [s. n.]: 1106-1114
- Larsson G, Maire M and Shakhnarovich G. 2017. FractalNet: ultra-deep neural networks without residuals//Proceedings of the 5th International Conference on Learning Representations. Toulon, France: [s. n.]
- LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. *Nature*, 521(7553): 436-444 [DOI: 10.1038/nature14539]
- LeCun Y, Bottou L, Bengio Y and Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of IEEE*, 86(11): 2278-2324 [DOI: 10.1109/5.726791]
- Li X, Wang W H, Hu X L and Yang J. 2019. Selective kernel networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 510-519 [DOI: 10.1109/CVPR.2019.00060]
- Lin J D, Wu X Y, Chai Y and Yin H P. 2020. Structure optimization of convolutional neural networks: a survey. *Acta Automatica Sinica*, 46(1): 24-37 (林景栋, 吴欣怡, 柴毅, 尹宏鹏. 2020. 卷积神经网络结构优化综述. 自动化学报, 46(1): 24-37) [DOI: 10.16383/j.aas.c180275]
- Lin M, Chen Q and Yan S C. 2013. Network in network [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/1312.4400.pdf>
- Liu C X, Zoph B, Neumann M, Shlens J, Hua W, Li L J, Li F F, Yuille A, Huang J and Murphy K. 2018. Progressive neural architecture search//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer: 19-35 [DOI: 10.1007/978-3-030-01246-5_2]
- Liu H X, Simonyan K and Yang Y M. 2019. DARTS: differentiable architecture search//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: [s. n.]
- Ma N N, Zhang X Y, Zheng H T and Sun J. 2018. ShuffleNet V2:

- practical guidelines for efficient CNN architecture design//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer; 122-138 [DOI: 10.1007/978-3-030-01264-9_8]
- Moniz J and Pal C. 2016. Convolutional residual memory networks [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/1606.05262v1.pdf>
- Netzer Y, Wang T, Coates A, Bissacco A, Wu B and Ng A Y. 2011. Reading digits in natural images with unsupervised feature learning. NIPS Workshop on Deep Learning and Unsupervised Feature Learning. Granada; [s. n.]: 12-17
- Pham H, Guan M Y, Zoph B, Le Q V and Dean J. 2018. Efficient neural architecture search via parameter sharing//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden; [s. n.]: 4092-4101
- Radosavovic I, Kosaraju R P, Girshick R, He K M and Dollár P. 2020. Designing network design spaces//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA; IEEE; 10425-10433 [DOI: 10.1109/CVPR42600.2020.01044]
- Ramachandran P, Zoph B and Le Q V. 2017. Searching for activation functions//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada; [s. n.]
- Real E, Aggarwal A, Huang Y P and Le Q V. 2019. Regularized evolution for image classifier architecture search//Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 31st Innovative Applications of Artificial Intelligence Conference, 9th AAAI Symposium on Educational Advances in Artificial Intelligence. Honolulu, USA; AAAI Press; 4780-4789
- Sandler M, Howard A, Zhu M L, Zhmoginov A and Chen L C. 2019. MobileNetV2: inverted residuals and linear bottlenecks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE; 4510-4520 [DOI: 10.1109/CVPR.2018.00474]
- Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R and LeCun Y. 2014. OverFeat: integrated recognition, localization and detection using convolutional networks//Proceedings of the 2nd International Conference on Learning Representations. Banff, Canada; [s. n.]
- Shen F L, Gan R and Zeng G. 2016. Weighted residuals for very deep networks//Proceedings of the 3rd International Conference on Systems and Informatics (ICSAI). Shanghai, China; IEEE; 936-941 [DOI: 10.1109/ICSAI.2016.7811085]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA; [s. n.]
- Singh S, Hoiem D and Forsyth D A. 2016. Swapout: learning an ensemble of deep architectures//Advances in Neural Information Processing Systems 29. Barcelona, Spain; [s. n.]: 28-36
- Srivastava R K, Greff K and Schmidhuber J. 2015. Highway networks [EB/OL]. [2020-5-22]. <http://arxiv.org/pdf/1505.00387.pdf>
- Sun J D and Zhao S. 2009. Image Low-Level Feature Extraction and Retrieval Technology. Beijing: Electronic Industry Press (孙君顶, 赵珊. 2009. 图像低层特征提取与检索技术. 北京: 电子工业出版社)
- Szegedy C, Ioffe S, Vanhoucke V and Alemi A A. 2016b. Inception-v4, Inception-ResNet and the impact of residual connections on learning//Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA; AAAI; 4278-4284
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA; IEEE; 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z. 2016a. Rethinking the Inception Architecture for Computer Vision//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA; IEEE; 2818-2826 [DOI: 10.1109/CVPR.2016.308]
- Tan M X, Chen B, Pang R M, Vasudevan V, Sandler M, Howard A and Le Q V. 2019a. MnasNet: platform-aware neural architecture search for mobile//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA; IEEE; 2820-2828 [DOI: 10.1109/CVPR.2019.00293]
- Tan M X and Le Q V. 2019b. EfficientNet: rethinking model scaling for convolutional neural networks//Proceedings of the 36th International Conference on Machine Learning. Long Beach, USA; [s. n.]: 6105-6114
- Targ S, Almeida D and Lyman K. 2016. Resnet in Resnet: generalizing residual architectures [EB/OL]. [2020-05-22] <https://arxiv.org/pdf/1603.08029.pdf>
- Veit A, Wilber M J and Belongie S. 2016. Residual networks behave like ensembles of relatively shallow networks//Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS). Barcelona, Spain; ACM; 550-558 [DOI: 10.5555/3157096.3157158]
- Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, Wang X G and Tang X O. 2017. Residual attention network for image classification//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE; 6450-6458 [DOI: 10.1109/CVPR.2017.683]
- Wang X L, Girshick R, Gupta A and He K M. 2018. Non-local neural networks//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE; 7794-7803 [DOI: 10.1109/CVPR.2018.00813]
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: convolutional block attention module//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer; 3-19 [DOI: 10.1007/978-3-030-01234-2_1]
- Xie S N, Girshick R, Dollár P, Tu Z W and He K M. 2017. Aggregated residual transformations for deep neural networks//Proceedings of

- 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 5897-5995 [DOI: 10.1109/CVPR.2017.634]
- Yamada Y, Iwamura M and Kise K. 2016. Deep pyramidal residual networks with separated stochastic depth [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/1612.01230.pdf>
- Yang T J, Howard A, Chen B, Zhang X, Go A, Sandler M, Sze V and Adam H. 2018b. NetAdapt: platform-aware neural network adaptation for mobile applications//Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer: 289-304 [DOI: 10.1007/978-3-030-01249-6_18]
- Yang Y B, Zhong Z S, Shen T C and Lin Z C. 2018a. Convolutional neural networks with alternately updated clique//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE: 2413-2422 [DOI: 10.1109/CVPR.2018.00256]
- Zagoruyko S and Komodakis N. 2016. Wide residual networks//Proceedings of the British Machine Vision Conference (BMVC). York, UK; BMVA Press: 87.1-12 [DOI: 10.5244/C.30.87]
- Zeiler M D and Fergus R. 2014. Visualizing and understanding convolutional networks//Proceedings of the 13th European Conference on Computer Vision (ECCV). Zurich, Switzerland: Springer: 818-833 [DOI: 10.1007/978-3-319-10590-1_53]
- Zhang H, Wu C R, Zhang Z Y, Zhu Y, Lin H B, Zhang Z, Sun Y, He T, Mueller J, Manmatha R, Liu M and Smola A. 2020. ResNeSt: split-attention networks [EB/OL]. [2020-05-22]. <https://arxiv.org/pdf/2004.08955.pdf>
- Zhang K, Guo L R, Gao C and Zhao Z B. 2019a. Pyramidal RoR for image classification. Cluster Computing, 22(2): 5115-5125 [DOI: 10.1007/s10586-017-1443-x]
- Zhang K, Guo Y R, Wang X S, Yuan J S and Ding Q L. 2019b. Multiple feature reweight DenseNet for image classification. IEEE Access, 7: 9872-9880 [DOI: 10.1109/ACCESS.2018.2890127]
- Zhang K, Guo Y R, Wang X S, Yuan J S, Ma Z Y and Zhao Z B. 2019c. Channel-wise and feature-points reweights densenet for image classification//Proceedings of 2019 IEEE International Conference on Image Processing (ICIP). Taipei, China; IEEE: 410-414 [DOI: 10.1109/ICIP.2019.8802982]
- Zhang K, Sun M, Han T X, Yuan X F, Guo L R and Liu T. 2018a. Residual networks of residual networks: multilevel residual networks. IEEE Transactions on Circuits and Systems for Video Technology, 28(6): 1303-1314 [DOI: 10.1109/TCSVT.2017.2654543]
- Zhang S, Gong Y H and Wang J J. 2019. The development of deep convolution neural network and its applications on computer vision. Chinese Journal of Computers, 42(3): 453-482 (张顺, 龚怡宏, 王进军. 2019. 深度卷积神经网络的发展及其在计算机视觉领域的应用. 计算机学报, 42(3): 453-482) [DOI: 10.11897/SP.J.1016.2019.00453]
- Zhang X C, Li Z Z, Loy C C and Lin D H. 2017. PolyNet: a pursuit of structural diversity in very deep networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA; IEEE: 3900-3908 [DOI: 10.1109/CVPR.2017.415]
- Zhang X Y, Zhou X Y, Lin M X and Sun J. 2018b. ShuffleNet: an extremely efficient convolutional neural network for mobile devices//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE: 6848-6856 [DOI: 10.1109/CVPR.2018.00716]
- Zoph B and Le Q V. 2016. Neural architecture search with reinforcement learning//Proceedings of the 5th International Conference on Learning Representations. Toulon, France; [s. n.]
- Zoph B, Cubuk E D, Ghiasi G, Lin T Y, Shlens J and Le Q V. 2019. Learning data augmentation strategies for object detection//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK; Springer: 566-583 [DOI: 10.1007/978-3-030-58583-9_34]
- Zoph B, Vasudevan V, Shlens J and Le Q V. 2018. Learning transferable architectures for scalable image recognition//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA; IEEE: 8697-9710 [DOI: 10.1109/CVPR.2018.00907]

作者简介



张珂, 1980年生, 男, 教授, 主要研究方向为深度学习、计算机视觉、生物特征识别。

E-mail: zhangkeit@ncepu.edu.cn

冯晓晗, 女, 硕士研究生, 主要研究方向为深度学习和图像分类。E-mail: 2499128718@qq.com

郭玉荣, 女, 博士研究生, 主要研究方向为深度学习和图像分类。E-mail: 490133104@qq.com

苏昱坤, 男, 硕士研究生, 主要研究方向为深度学习和生物特征识别。E-mail: 842371910@qq.com

赵凯, 男, 硕士研究生, 主要研究方向为深度学习和电力视觉。E-mail: 2321834855@qq.com

赵振兵, 男, 教授, 主要研究方向为电力视觉。

E-mail: zhaozhenbing@ncepu.edu.cn

马占宇, 男, 教授, 主要研究方向为模式识别与机器学习基础理论与方法。E-mail: mazhanyu@bupt.edu.cn

丁巧林, 女, 副教授, 主要研究方向为电力智能信息处理。

E-mail: dingqiaolin@ncepu.edu.cn