

基于多尺度分割的图像识别残差网络研究

袁单飞^{1,2},陈慈发^{1,2},董方敏¹

(1.三峡大学 计算机与信息学院,湖北 宜昌 443000; 2.湖北省建筑质量检测装备工程技术研究中心,湖北 宜昌 443000)

摘要:深度卷积神经网络能够解决复杂的计算机视觉问题,被广泛应用于图像识别任务中。在基于深度卷积神经网络的图像识别过程中,增加网络的深度和宽度能够产生丰富的特征信息,使用多尺度分割方法能够有效减少冗余的特征信息。然而,增加网络的深度和进行多尺度分割都会影响识别速度。如何在保证精度的同时提高识别速度,成为设计高效网络的关键问题。通过增加网络宽度的方法对ResNet残差网络进行改进,在保证精度的基础上提升识别速度。使用ResNet-D中的残差结构并减少网络长度,得到长度只有7层的残差网络,同时对HS-ResNet中的多尺度分割方法进行优化,只保留最后一次连接合并操作,得到图像识别残差网络SSRNet。在CIFAR 10和CIFAR 100数据集上的实验结果显示,SSRNet速度最高较ResNet网络提升7倍多,同时错误率最高下降8.81%,表明缩短网络长度可大幅加快图像识别速度,同时结合多尺度分割方法能够有效提升识别精度。

关键词:多尺度分割;残差网络;卷积神经网络;图像识别;图像分类

开放科学(资源服务)标志码(OSID):



中文引用格式:袁单飞,陈慈发,董方敏.基于多尺度分割的图像识别残差网络研究[J].计算机工程,2022,48(5):258-262,271.

英文引用格式:YUAN D F, CHEN C F, DONG F M. Research on residual network of image recognition based on multiscale split[J].Computer Engineering, 2022, 48(5): 258-262, 271.

Research on Residual Network of Image Recognition Based on Multiscale Split

YUAN Danfei^{1,2}, CHEN Cifa^{1,2}, DONG Fangmin¹

(1.College of Computer and Information Technology, China Three Gorges University, Yichang, Hubei 443000, China;

2.Hubei Province Engineering Technology Research Center for Construction Quality Testing Equipment, Yichang, Hubei 443000, China)

[Abstract] The emergence of deep Convolutional Neural Network(CNN) has contributed significantly to solving complex computer vision problems, and they have been widely used in image recognition tasks. Designing an efficient network has become the key to improving the performance of deep CNN. In the process of image recognition based on a deep CNN, increasing the depth and width of the network can produce rich feature information, whereas the use of a multi-scale segmentation method can effectively reduce redundant feature information. However, increasing the depth of the network in multi-scale segmentation affects the recognition speed. Thus, improving recognition speed while ensuring accuracy has become an important goal in designing efficient networks. To solve this problem, the network width is increased using ResNet to improve the recognition speed and ensure accuracy. Using the residual structure in ResNet-D and reducing the network length, a residual network with only seven layers is obtained. Concurrently, the multi-scale segmentation method in HS-ResNet is optimized, and only the last connection and merging operation are retained to obtain SSRNet. The experimental results on the CIFAR 10 and CIFAR 100 datasets show that the maximum speed of SSRNet is more than seven times higher than that of ResNet, and the error rate can be reduced by 8.81%. This demonstrates that shortening the length of the network can significantly accelerate the speed of image recognition, whereby the recognition accuracy is effectively improved in combination with the multi-scale segmentation method.

[Key words] multiscale split; residual network; Convolutional Neural Network(CNN); image recognition; image classification
DOI: 10. 19678/j. issn. 1000-3428. 0061392

0 概述

深度卷积神经网络已被广泛应用于图像分类、目标检测、实例分割等领域。在此背景下,如何设计高效

的网络模型成为提高网络性能的关键。早期的卷积神经网络(如 AlexNet^[1]和 VGGNet^[2])包含了简单的激活函数和卷积结构,使得多尺度特征的数据驱动学习成为可能。ResNet^[3]中的残差结构使得更深的网络能

基金项目:国家自然科学基金新疆联合基金重点项目(U1703261)。

作者简介:袁单飞(1993—),男,硕士研究生,主研方向为计算机视觉、人工智能、嵌入式系统;陈慈发(通信作者),教授、研究员;董方敏,教授、博士生导师。

收稿日期:2021-04-24 修回日期:2021-05-31 E-mail:442130165@qq.com

够进行有效学习。WRN^[4]通过增加残差网络的网络宽度来提高识别精度。Res2Net^[5]使用了对残差结构进行多尺度分割的方法来提高识别精度。HS-ResNet^[6]对多尺度分割方法进行改进,提出了参数量更小、识别速度更快的网络模型。

通过增加网络的深度和宽度能够产生丰富的特征信息,通过多尺度分割方法能够充分利用特征信息,减少冗余的特征信息。但在增加网络深度和进行多尺度分割的同时,也会减慢网络的识别速度。如何在保证识别精度的同时提高识别速度,成为设计高效网络模型的一个重要目标。本文对HS-ResNet提出的多尺度分割方法进行改进,提出一种简单的多尺度分割方法,并通过减少网络层数和增加宽度的方法,在保证网络的识别精度和参数量相近的情况下,加快识别速度。

本文结合ResNet-D^[7]的残差结构和WRN增加网络宽度的方法对ResNet进行改进。通过减少网络层数的方法,在CIFAR数据集上得到网络长度只有7层的网络模型,并对HS-ResNet的多尺度分割方法进行优化,得到基于多尺度分割的图像识别残差网络SSRNet。从识别精度和速度两个角度,将该网络与PyramidNet^[8]进行比较,以验证网络性能。

1 相关工作

残差网络的出现使得深度学习被广泛应用于生产、生活的各个方面。残差网络通过求解预测值与观测值之间的差值来对拟合函数进行学习,使深层网络能够进行有效学习。近年来,基于残差网络的改进网络大幅提高了网络的识别精度。

1.1 ResNet

在深度学习中,网络层数的增多一般会伴随计算资源的消耗,网络模型容易过拟合,并造成梯度消失、梯度爆炸等问题。随着网络层数的增加,网络模型会出现退化现象,即随着网络层数的增多,训练损失逐渐下降,然后趋于饱和;如果再增加网络的深度,训练损失反而上升。当网络模型出现退化时,浅层网络模型能达到比深层网络模型更好的训练效果,这时如果把底层特征信息传到高层,那么效果应该至少不比浅层网络模型效果差。基于这种使用直接映射来连接网络模型不同层的思想,残差网络ResNet应运而生。

当输出维度与输入维度不等时,ResNet需要对输入维度进行升维操作。在ResNet中直接使用步长为2的 1×1 卷积进行升维,当输入图像大小减半时,会造成特征信息丢失。ResNet-D以ResNet为基础,在残差结构升维操作之前加入步长为2的 2×2 均匀池化,然后使用步长为1的 1×1 卷积进行升维,使得网络识别精度大幅提升。

1.2 WRN

随着网络层数加深,训练深度卷积神经网络存在着梯度消失与梯度弥散等问题,实验结果也表明,越深的网络模型带来的识别精度提升并不明显,反而需要降低识别速度^[9]。针对网络是否越深越窄效果越好,以及是否只要保证网络参数量,训练一个更宽更浅的

网络即可的问题,WRN提出了基于扩展通道数学习机制的卷积神经网络,以期通过更浅的网络模型来获得与深度网络模型相近的识别精度,以及更快的识别速度。

WRN通过增加网络宽度能够提升精度,当参数量相同时,WRN的速度更快。但与残差网络一样,当参数量过大时,其存在梯度消失与梯度弥散等问题。

1.3 Res2Net

在多个尺度上表示特征对视觉任务非常重要。卷积神经网络展示出更强的多尺度表示能力,在广泛的应用中实现一致的性能提升^[10-16],如文献[10]提出的多尺度特征提取和多级别特征融合的显著性目标检测方法,文献[11]提出的结合超像素分割的多尺度特征融合图像语义分割算法,文献[13]提出的基于双向特征金字塔和深度学习的图像识别方法,文献[15]提出的轻量级多尺度融合的图像篡改检测算法。在Res2Net之前的方法大多以特征金字塔^[17]分层方式表示多尺度特征。Res2Net在原有的残差单元结构中通过增加小的残差块来增加每一层的感受野范围,以更细的粒度表示多尺度特征,使得残差单元结构的中间主卷积从单分支变为多分支。

Res2Net模块结构简单,性能优秀,其在卷积神经网络的3个维度(深度、宽度和基数)之外,揭示了一个新的尺度维度。Res2Net模块可以很容易地与其他模块结合,特征提取能力更强大,且不增加计算负载。

1.4 HS-ResNet

HS-ResNet改进了对特征图多尺度分割的卷积和连接,在提升识别精度的同时,也提高了识别速度。HS-ResNet主要考虑以下3个问题:1)如何避免在特征图中产生冗余信息;2)如何在不增加计算复杂度前提下,使网络学习到更强的特征表达;3)如何在得到更高识别精度的同时,保持较快的识别速度。基于这3个问题,本文设计HS-Block模块来生成多尺度特征。

当通道数量较大时,Res2Net方法计算复杂度也随之增加。GhostNet^[18]实验结果显示,一部分特征图可以通过已有的特征图生成。HS-ResNet借助这一思想,在HS-Block内将 S_2 组卷积得到的特征图部分连接到 S_3 组,实现了特征图复用,降低了计算复杂度。

HS-ResNet提出了HS-Block模块,其可高效提取多尺度特征,在多个视觉任务(如图像分类、目标检测和实例分割)上取得了优秀的识别性能。HS-Block具有即插即用特性,可以轻易嵌入到现有网络中并提升识别性能。

2 HS-ResNet方法

HS-ResNet的多尺度分割方法使不同组的特征信息享受不同尺度的感受野。在前面连接进来的特

征信息中,卷积次数较少,感受野较小,更关注细节信息;在后面连接进来的特征信息中,卷积次数较多,感受野较大,更关注全局信息。通过不同大小的感受野增加特征信息的丰富性。本文提出的多尺度分割方法,对HS-ResNet的多尺度分割方法中多次连接合并的操作进行优化,只保留了最后一次连接合并操作,使得识别速度大大提升。

WRN提出基于扩展通道数学习机制的卷积神经网络,以期通过更浅的网络来获得与深度网络相近的精度,以及更快的识别速度。本文结合WRN的增加通道数和对网络长度进行缩短的方法,构建长度只有7层的网络模型,在保证识别精度的同时,进一步提升识别速度。

2.1 网络多尺度分割模块

本文提出3种多尺度分割网络模型SSRNet,用新的多尺度分割模块SS-Block代替ResNet中的3×3卷积,如图1所示。当特征信息输入时,执行以下步骤:

步骤1 把1×1卷积输入的特征信息按通道数平均分割成相等的2个部分(如果通道数为奇数,则向下取整)。

步骤2 一半特征信息直接送到最后进行合并,另一半特征信息进行卷积。

步骤3 重复上面步骤,直到最后一个特征信息。

步骤4 把最后一个特征信息与前面得到的前半特征信息进行合并,一起输出给1×1卷积。SSRNet-a的多尺度分割模块与ResNet的瓶颈模块结构相似,只是把ResNet瓶颈模块中3×3卷积用SS-Block替代。其中,SS-Block的卷积表示为3×3卷积+批量正则化+Relu激活函数。

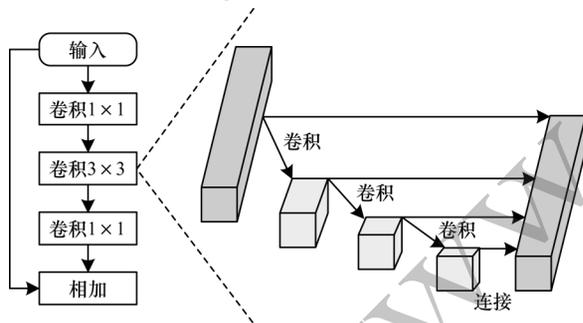


图1 SSRNet-a的SS-Block示意图

Fig.1 Schematic diagram of SS-Block in SSRNet-a

3种多尺度分割网络模型在CIFAR数据集^[19]上的详细结构如表1所示。其中:模型a表示只替换了ResNet瓶颈模块中的3×3卷积;模型b表示在模型a的基础上去掉了ResNet瓶颈模块中2个1×1降维和升维卷积;模型c表示在模型b的基础上使每次下采样的特征通道数等于第1组的特征通道数。模型a与模型b相比,识别精度相似,参数量更小,速度稍慢;模型c与模型b相比,识别精度较低,参数量更小,速度更快。

表1 3种SSRNet模型的结构

Table 1 Structure of three SSRNet models

组名	输出大小/像素	模型a	模型b	模型c
组1	32×32	[3×3,128]×2	[3×3,256]×2	[3×3,256]×2
组2	16×16	[3×3,256]×2	[3×3,512]×2	[3×3,256]×2
组3	8×8	[3×3,512]×2	[3×3,1024]×2	[3×3,256]×2
平均池化	1×1	[8×8]	[8×8]	[8×8]

2.2 网络长度与分割尺度

从AlexNet开始,深度卷积神经网络通过增加网络长度对多尺度特征信息进行学习来提高网络精度。然而在增加网络长度的同时,会产生减少特征重用的问题。WRN通过增加网络宽度而不是网络长度来解决这个问题,使得识别速度更快,且网络长度越短,识别速度越快。随着网络长度的增加,卷积神经网络的感受野增大,但并不是全部感受野都对输出特征的贡献相同,感受野的中心区域对输出特征影响更大^[20]。笔者通过实验发现,当网络长度设置为感受野大小与输入图像大小比例的1.5~2倍时,网络的识别精度较高,同时识别速度较快,在精度和速度之间取得了较好的平衡。第k层感受野的计算公式如下:

$$l_k = l_{k-1} + (f_k - 1) \times \sum_{i=1}^{k-1} s_i \quad (1)$$

其中: l_{k-1} 是第k-1层的感受野大小; f_k 是当前层的卷积核大小; s_i 是第i层的步长。

根据式(1)计算可得SSRNet-a最后输出感受野大小为51×51像素,经过3次尺度分割后感受野为67×67像素,恰好为输入图像大小32×32像素的1.5~2倍。该网络模型的长度能在识别精度和识别速度之间取得较好的平衡。

2.3 网络宽度与下采样率

通过增加网络宽度(即特征通道数),可以在保持网络识别速度的同时提高识别精度。但增加网络宽度的同时会增加参数量,所以,网络宽度需要根据具体情况进行设置。当计算能力充足时,可以设置每组的网络宽度相同,不同组的网络宽度依次递增(例如模型a和模型b);当计算能力不足时,可以设置各组的网络宽度都相同,通过牺牲精度来减少网络模型的参数量(例如模型c)。

当输入图像过大时,通过下采样可以减少模型计算量。输入图像下采样到小尺寸(6×6像素,7×7像素,8×8像素)时比较合适^[21]。本文提出的多尺度分割网络模型在CIFAR数据集上,通过2次2倍下采样把输入图像尺寸从32×32像素缩小到8×8像素,然后使用均匀池化,把输出特征尺寸变成1×1像素后,进行图像识别。

3 实验与结果分析

在网络模型训练上,本文算法使用飞桨深度学习框架进行实现。实验数据集为CIFAR 10和CIFAR 100,

每幅图像都通过随机边缘填充裁剪 4 个像素,并通过随机水平翻转进行数据增强,最后进行均值归一化。训练环境为百度的 AI Studio, GPU 为 Tesla V100。

3.1 实验数据集

CIFAR 10 和 CIFAR 100 被标记为 8 000 万个微小图像数据集的子集,由 Alex Krizhevsky、Vinod Nair 和 Geoffrey Hinton 收集。

CIFAR 10 数据集由 10 个类的 60 000 个 32×32 像素的彩色图像组成,每个类有 6 000 幅图像,共 50 000 幅训练图像和 10 000 幅测试图像。10 个类别分别为飞机、汽车、鸟、猫、鹿、狗、青蛙、马、船和卡车。

CIFAR 100 数据集与 CIFAR 10 数据集类似,其中包含 100 个类,每个类包含 600 个图像。每类各有 500 个训练图像和 100 个测试图像,共有 50 000 幅训练图像和 10 000 幅测试图像。CIFAR 100 中的 100 个类被分成 20 个超类。每个图像都带有一个“精细”标签(其所属类)和一个“粗糙”标签(其所属超类)。

3.2 训练策略

网络模型训练使用的是带动量的随机梯度下降法进行反向传播,在 CIFAR 10 和 CIFAR 100 训练集上共训练 300 轮。初始学习率设置为 0.000 01,进行 10 轮线性预热训练。在此基础上,以 0.1 的学习率开始 290 轮的余弦衰减训练。卷积核的权重参数使用 MSRA^[22]进行初始化,权重衰减系数设置为 0.000 05,动量系数设置为 0.9,每批数据大小设置为 128,同时使用了系数为 0.05 的标签平滑策略。

3.3 结果分析

从 CIFAR 10 数据集中随机抽取 4 幅图像,SSRNet 的识别结果如图 2 所示。原图为 32×32 像素的图像,左上角为识别结果对应的类别标签。

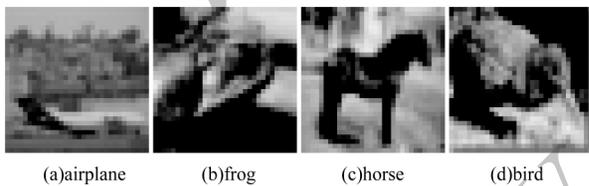


图 2 SSRNet 识别结果

Fig.2 Recognition result of SSRNet

SSRNet、ResNet 和 HS-ResNet 以相同训练参数,在 CIFAR 10 数据集上训练得到的识别速度与识别精度对比如图 3 所示,其中:圆形面积表示模型参数量的大小,识别速度在 Tesla V100 上批次大小为 1 时测试得到:左上角 3 个圆圈表示本文提出的多尺度分割网络模型 SSRNet,右下角 3 个圆圈表示残差网络模型 ResNet,中间圆圈表示 SSRNet 中的多尺度分割模块 SS-Block 使用 HS-Block 替代得到的网络模型 HS-ResNet。SSRNet-c 的识别精度与 ResNet-56 相近,速度快于 ResNet-20。使用缩短网络长度方法得到的 HS-ResNet 与 SSRNet-a 和 SSRNet-b 相比,识别精度略低,识别速度稍慢。由此可见,多尺度分割模块 SS-Block 与 HS-Block 相比,识别精度更高,识别速度更快。

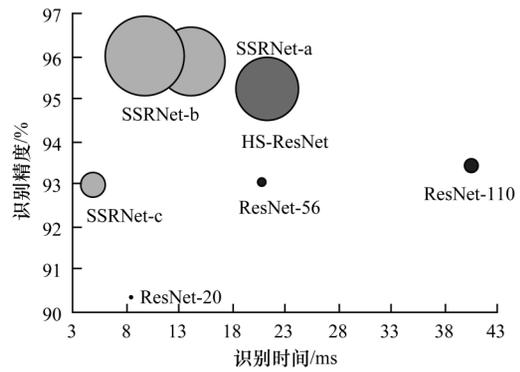


图 3 不同网络模型的识别速度与识别精度对比

Fig.3 Comparison of recognition latency and recognition accuracy of different network models

在本文的训练环境和训练策略下,不同网络模型在 CIFAR 数据集上的实验结果如表 2 所示,其中,加粗表示最优值。多尺度网络模型 SSRNet-c 的识别速度最快,是 ResNet-56 识别速度的 4 倍;SSRNet-a 和 SSRNet-b 的识别错误率相近,分别为 4.14% 和 4.04%。前者模型参数量更少,后者识别速度更快。当对识别速度要求较高时,使用 SSRNet-c 可以获得最快的识别速度。当对识别精度和模型参数量要求较高时,使用 SSRNet-a 可以获得较高的识别精度和较低的模型参数量。SSRNet-b 是 SSRNet-a 与 SSRNet-c 性能的折中,在识别精度、识别速度和模型参数量之间取得了较好的平衡。

表 2 CIFAR 数据集上不同网络模型的识别速度与错误率对比

Table 2 Comparison of recognition latency and recognition error rates of different networks models on CIFAR datasets

网络模型	参数量/ 10 ⁶	识别速度/ (frame·s ⁻¹)	错误率/%	
			CIFAR 10	CIFAR 100
ResNet-20	0.2	117	9.70	—
ResNet-56	0.9	47	6.99	—
ResNet-110	1.7	25	6.57	29.68
HS-ResNet-20	0.2	37	9.71	—
HS-ResNet-56	0.8	12	7.14	—
HS-ResNet-110	1.5	6	6.52	29.68
HS-ResNet-a	17.1	47	4.81	20.50
SSRNet-a	20.2	71	4.14	21.00
SSRNet-b	26.8	100	4.04	20.87
SSRNet-c	3.1	200	7.07	30.69

表 3 展示了不同的网络模型在 CIFAR 数据集上的实验结果,本文网络模型的识别错误率与其他网络模型相近时,网络层数最少。目前用于图像识别的神经网络大都是基于深度卷积神经网络,网络层数远大于本文的网络层数,网络层数越少,识别速度越快。当识别错误率相近时,本文网络模型的识别速度快于其他网络模型。

表3 CIFAR数据集上不同网络模型的错误率

Table 3 Comparison of error rates of different network models on CIFAR datasets

网络模型	参数量/ 10^6	层数	错误率/%	
			CIFAR 10	CIFAR 100
ResNet	1.7	110	6.57	29.68
WRN	36.5	28	4.17	20.50
PyramidNet	1.7	110	4.58	23.12
HS-ResNet	1.5	110	6.52	29.68
SSRNet-a	20.2	7	3.82	19.38
SSRNet-b	26.8	7	3.87	20.87
SSRNet-c	3.1	7	6.75	30.69

在本文的训练环境和训练策略下, CIFAR数据集上 SSRNet-b消融实验的结果如表4所示。使用多尺度分割模块 SS-Block 替换 ResNet 中的残差模块后, 错误率相近, 但识别速度从 25 frame/s 下降到 11 frame/s。使用缩短网络长度方法改进 ResNet 后, 错误率相近, 识别速度从 25 frame/s 提升到 192 frame/s, 提升 7 倍多。同时使用多尺度分割方法和缩短网络长度方法改进 ResNet 后, 得到的 SSRNet 的错误率大幅度下降, 在 CIFAR 10 和 CIFAR 100 数据集上分别下降 2.53% 和 8.81%, 识别速度从 25 frame/s 提升到 100 frame/s, 提升了 3 倍左右。由此可见, 缩短网络长度方法对识别速度有大幅提升, 同时结合多尺度分割方法也可以提高识别精度。

表4 CIFAR数据集上 SSRNet-b消融实验的识别速度与错误率对比

Table 4 Comparison of recognition latency and recognition error rates of SSRNet-b ablation experiment on CIFAR datasets

网络模型	识别速度/ (frame·s ⁻¹)	错误率/%	
		CIFAR 10	CIFAR 100
ResNet-110	25	6.57	29.68
SSRNet-110	11	6.84	30.54
ResNet-b	192	6.76	30.79
SSRNet-b	100	4.04	20.87

4 结束语

本文提出的多尺度分割方法在对多尺度特征信息进行表示时, 并没有增加很多模型参数量和计算量, 而缩短网络长度的方法在保持识别精度的同时, 大幅提高了识别速度。结合这 2 种方法得到的基于多尺度分割的图像识别残差网络 SSRNet, 与其他网络模型相比, 在识别精度相近时识别速度更快。下一步将把该网络模型应用到目标检测领域中, 并进一步加快识别速度。

参考文献

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [2] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2021-06-11]. <https://arxiv.org/abs/1409.1556>.
- [3] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [4] ZAGORUYKO S, KOMODAKIS N. Wide residual networks [S. l.]; BMVC, 2016: 1-12.
- [5] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.
- [6] YUAN P C, LIN S F, CUI C, et al. HS-ResNet: hierarchical-split block on convolutional neural network[EB/OL]. [2021-06-11]. <https://arxiv.org/abs/2010.07621>.
- [7] HE T, ZHANG Z, ZHANG H, et al. Bag of tricks for image classification with convolutional neural networks[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2019: 558-567.
- [8] HAN D, KIM J, KIM J. Deep pyramidal residual networks[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 6307-6315.
- [9] 任长娥, 袁超, 孙彦丽, 等. 宽度学习系统研究进展[J]. 计算机应用研究, 2021, 38(8): 2258-2267.
- [10] REN C G, YUAN C, SUN Y L, et al. Research of broad learning system[J]. Application Research of Computers, 2021, 38(8): 2258-2267. (in Chinese)
- [11] 黎玲利, 孟令兵, 李金宝. 多尺度特征提取和多级别特征融合的显著性目标检测方法[J]. 工程科学与技术, 2021, 53(1): 170-177.
- [12] LI L L, MENG L B, LI J B. Salient object detection based on multi-scale feature extraction and multi-level feature fusion[J]. Advanced Engineering Sciences, 2021, 53(1): 170-177. (in Chinese)
- [13] 官申珂, 林晓, 郑晓妹, 等. 结合超像素分割的多尺度特征融合图像语义分割算法[J]. 图学学报, 2021, 42(3): 406-413.
- [14] GUAN S K, LIN X, ZHENG X M, et al. A semantic segmentation algorithm using multi-scale feature fusion with combination of superpixel segmentation[J]. Journal of Graphics, 2021, 42(3): 406-413. (in Chinese)
- [15] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6.
- [16] REN H, WANG X G. Review of attention mechanism[J]. Journal of Computer Applications, 2021, 41(S1): 1-6. (in Chinese)
- [17] 赵升, 赵黎. 基于双向特征金字塔和深度学习的图像识别方法[J]. 哈尔滨理工大学学报, 2021, 26(2): 44-50.
- [18] ZHAO S, ZHAO L. On image recognition using bidirectional feature pyramid and deep neural network[J]. Journal of Harbin University of Science and Technology, 2021(2): 44-50. (in Chinese)

(下转第 271 页)

(上接第 262 页)

- [14] 朱旭东,熊贇. 基于多层次注意力与图模型的图像多标签分类算法[J]. 计算机工程, 2022, 48(4): 173-178, 190. ZHU X D, XIONG Y. Multi-label image classification algorithm based on multi-scale attention and graph model[J]. Computer Engineering, 2022, 48(4): 173-178, 190. (in Chinese)
- [15] 吴旭,刘翔,赵静文. 一种轻量级多尺度融合的图像篡改检测算法[J]. 计算机工程, 2022, 48(2): 224-229, 236. WU X, LIU X, ZHAO J W. A lightweight multiscale fusion algorithm for image tampering detection[J]. Computer Engineering, 2022, 48(2): 224-229, 236. (in Chinese)
- [16] 王柳程,欧阳城添,梁文. 基于改进特征金字塔网络的人体姿态估计[J]. 计算机工程, 2021, 47(8): 251-259, 270. WANG L C, OUYANG C T, LIANG W. Human pose estimation based on improved pyramid feature network[J]. Computer Engineering, 2021, 47(8): 251-259, 270. (in Chinese)
- [17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA, USA: IEEE Press, 2017: 936-944.
- [18] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA, USA: IEEE Press, 2020: 1577-1586.
- [19] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images[J]. Handbook of Systemic Autoimmune Diseases, 2009, 1(4): 1-5.
- [20] LUO W G, LI Y J, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks[EB/OL]. [2021-06-11]. <https://arxiv.org/pdf/1701.04128v1.pdf>.
- [21] CAO X. A practical theory for designing very deep convolutional neural networks[EB/OL]. [2021-06-11]. <http://pdfs.semanticscholar.org/7922/2fad9f671be142bd7e42cd785a2cb06a1d30.pdf>.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA, USA: IEEE Press, 2015: 1026-1034.

编辑 金胡考