



# 基于改进深度残差网络的图像分类算法

储岳中<sup>1,2\*</sup>, 汪佳庆<sup>1</sup>, 张学锋<sup>1</sup>, 刘恒<sup>1</sup>

(1. 安徽工业大学计算机科学与技术学院 安徽 马鞍山 243032; 2. 安徽省工业互联网智能应用与安全工程实验室 安徽 马鞍山 243023)

**【摘要】** ResNet 的瓶颈结构存在一定的冗余输出, 面对类别较少的分类任务时, 冗余信息占比会非常大, 造成计算资源浪费。针对这个问题, 该文利用残差结构和连接操作设计了一个新的升维结构 (RC 结构) 用于改进瓶颈结构。RC 结构可以降低瓶颈结构的资源消耗, 并增强梯度传递。将 RC 与多种残差网络进行结合, 并在多个数据集上进行图像分类实验。实验结果表明, 基于 RC 的残差网络在面对类别较少的分类任务时, 能有效提高网络的效率和精度。

**关键词** 瓶颈结构; 升维操作; 冗余性; 残差网络

中图分类号 TP391 文献标志码 A doi:10.12178/1001-0548.2020314

## Image Classification Algorithm Based on Improved Deep Residual Network

CHU Yue-zhong<sup>1,2\*</sup>, WANG Jia-qing<sup>1</sup>, ZHANG Xue-feng<sup>1</sup>, and LIU Heng<sup>1</sup>

(1. School of Computer Science and Technology, Anhui University of Technology Maanshan Anhui 243032;

2. Anhui Engineering Laboratory for Intelligent Applications and Security of Industrial Internet Maanshan Anhui 243032)

**Abstract** Research shows that there is redundancy in the output feature maps of ResNet bottleneck structure. Such redundancy ensures a comprehensive understanding of the input data, but generates the redundant information which consumes additional computational resources. And the proportion of redundant information is very large when processing small category classification tasks. To solve this problem, a new dimension increasing structure is designed to improve the bottleneck structure by residual-like structure and concatenation operation. This structure is called residual concatenation (RC). The RC can not only reduce the amount of calculation and parameters of bottleneck structure, but also enhance the gradient transmission of back-propagation to improve the accuracy. In this work, the RC is combined with multiple residual networks and image classification experiments are performed on multiple datasets. The results show that the RC-based bottleneck structure can reduce the consumption and improve the accuracy of classification tasks while processing small category classification tasks.

**Key words** bottleneck; dimension up; redundancy; residual network

自从深度残差网络 ResNet 被提出来并在 ILSVRC2015 比赛中取得冠军后, 许多研究者把目光放在了 ResNet 的改进工作上, 由此出现了一大批优秀的基于残差结构的卷积神经网络。目前为止改进大体上可分为 4 类: 1) 是对输入模块的数据进行 split-transform-merge 操作, 例如使用分组卷积的 ResNext<sup>[1]</sup> 和使用层级残差结构的 Res2Net<sup>[2]</sup>, 它们能在降低消耗的同时增加网络的精度。2) 在残差结构里引入注意力机制, 例如加入通道自注意力机制的 SENet<sup>[3]</sup>、加入多尺度自适应调节感受的 SKnet<sup>[4]</sup>、使用分组注意力机制的 ResNeSt<sup>[5]</sup>。注意

力机制的使用可以让网络有偏向地选择重要的特征图进行特征提取, 从而提高精度。3) 优化模块拓扑结构, 例如使用预先激活方式的 ResNetV2<sup>[6]</sup>, 使网络易于优化并提高模型的正则化, 以及每个 stage 采用不同模块来优化拓扑结构的 IResNet<sup>[7]</sup>, 加强了信息在网络间的传播, 减少了信息损失。4) 对残差网络整体进行优化, 例如提出让网络变得更宽的 WRN<sup>[8]</sup> 和建立密集连接从而缓解梯度消失、加强特征传播的 DenseNet<sup>[9]</sup>。

上述网络大部分都是在 ResNet 的瓶颈结构基础上进行改进的。瓶颈结构由一条主干和一条残差

收稿日期: 2020-08-25; 修回日期: 2021-01-01

基金项目: 国家自然科学基金 (61971004)

作者简介: 储岳中 (1971-), 男, 博士, 副教授, 主要从事模式识别与机器学习方面的研究。Email: mychu@126.com

支路组成, 主干上数据先由  $1 \times 1$  点卷积进行降维, 再经过  $3 \times 3$  卷积块进行特征提取, 最后由  $1 \times 1$  点卷积升维, 卷积过程中还使用 BN 操作进行归一化以及 ReLU 函数增加非线性。先降维最后升维的操作大大降低了  $3 \times 3$  卷积所需的计算量。传统的观点认为影响准确率的主要是  $3 \times 3$  卷积核, 所以多数工作都是对其进行改进, 而选择保留降维升维的  $1 \times 1$  卷积核。瓶颈结构比较简单, 但网络结构上也存在冗余, 因此后面出现的残差网络如 ResNext 才能在不过多增加计算量的前提下提高准确率。冗余的产生主要来自升维操作时生成的过多的类似特征图<sup>[10]</sup>。传统的改进方法偏向于增加升维时输入数据的特征图数量<sup>[8]</sup>以及复杂性<sup>[1-2, 4]</sup>。这两种操作都能有效减少瓶颈结构输出信息的冗余。但考虑到冗余性还和分类任务的类别数有关, 类别数越大, 需要的有效特征图数量也就越多。在面对如 10 分类这样的分类任务时, 传统的瓶颈结构及其改进还是不可避免地产生过多的冗余特征图。这些冗余虽然保证了网络对输入信息的充分理解, 但过多的冗余信息会导致计算资源的浪费。

针对这个问题, 本文对瓶颈结构里的升维操作进行了改进。利用残差结构建立了升维操作输入和输出之间的通路, 并使用连接操作代替残差结构中的加法。由于用到了残差结构和连接操作, 把这个结构命名为 RC。RC 结构不仅减少了升维操作的计算量和参数量, 而且提高了反向传播时的梯度传递, 在面对分类数较少的任务时能够有效提升网络精度。由于 RC 结构能和很多残差网络相结合, 在多个数据集上进行了多组对比实验, 结果显示基于 RC 的瓶颈结构在面对分类数较少的任务时, 效率和精度都有所提升。

## 1 研究方法

### 1.1 冗余性分析

图 1a 为 ResNet 原文<sup>[11]</sup>中提出的一个瓶颈结构案例。输入数据的特征图通道数为 256, 经过降维、特征提取后, 升维复原成 256 个通道, 并与支路传来的数据相加作为输出。比起直接使用一个  $3 \times 3$  卷积核进行特征提取, 这样做无疑大大降低了运算量, 但精度也会受到影响。图 1b 为保证输入输出特征通道数不变的情况下, 只使用一个  $3 \times 3$  卷积核进行特征提取的情况。在 cifar10 上做了个简单的对比实验, 相关配置和实验部分一致, 实验结果如表 1 所示。ResNet# 为将 ResNet29 中的瓶颈结构

替换为图 1b 时的情况, 结果显示使用瓶颈结构后计算量和参数量大大减少了(表格里 M 为 million, B 为 billion), 但在 SGD 和 Adam 这两种优化器下, 准确率都减少了 2%~3%。

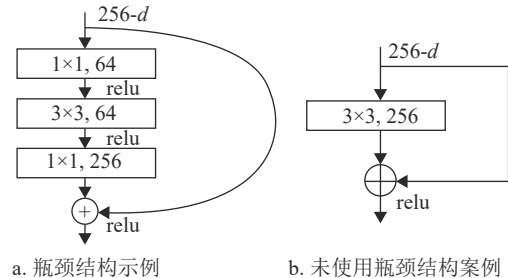


图 1 瓶颈结构案例和未使用瓶颈结构案例

表 1 ResNet29 和 ResNet# 实验对比

网络结构	运算量/B	参数量/M	准确率/%	
			SGD	Adam
ResNet29	0.76	4.92	91.30	91.52
ResNet#	4.47	31.52	94.62	93.64

在图 1 中, 两种模块的输出都是 256 个通道, 但实验结果表明, 图 1b 结构有着更好的特征表达能力, 其输出的特征图所含信息更有效。图 1a 输出的特征图虽然也是 256 个通道, 但相较图 1b 的输出特征图存在一些冗余信息。设瓶颈结构中  $3 \times 3$  卷积核输出通道数为  $n$ ,  $1 \times 1$  卷积核升维所产生的通道数为  $m$ , 则存在一个比例  $s=m/n$ ,  $s$  越大代表特征图通道数上升比例越大。文献 [7] 设计了更宽的 ResNet, 即  $m$  不变的情况下  $n$  更大的 ResNet 网络, 实验表明增加  $n$  能有效增加 ResNet 模块的特征表达能力。反过来也能得到: ResNet 瓶颈结构中,  $m$  不变的情况下,  $n$  越小,  $s$  越大, 模块的特征表达能力越弱, 冗余性越高。因此与  $s=4$  的 ResNet 网络不同, 后面新出现的残差网络都选择了较小的  $s$  值。

瓶颈结构的冗余性除了和  $s$  值有关外, 还和任务的类别数有关, 以 cifar10 和 cifar100 为例, 进行了 3 组对比实验, 如表 2 所示。 $k$  为  $n$  不变的情况下  $s$  缩放的倍率,  $k$  越小, 瓶颈结构输出的特征图通道数  $m$  就越少。从表中可以看出 ResNet29( $k=1$ ) 在 cifar10 和 cifar100 上准确率相差很大。100 分类相较 10 分类, 对网络的精确度要求更高, 本文认为, 在其他条件不变的情况下, 分类任务的类别数越大, 维持精度所需的输出特征图就越多。而 ResNet29 的网络结构每层输出的特征图是固定的, 所以它在 10 分类问题上存在的冗余特征图可

能非常多。为了验证这个猜想, 对  $k=\{3/4, 1/2\}$  的情况进行了测试。表 2 显示, 在 cifar10 上,  $k=1$  和  $k=3/4$  的结果十分相似, 但  $k=3/4$  时, 瓶颈结构输出特征图的数量相较  $k=1$  减少了 25%, 由此可见 ResNet29 在 cifar10 上存在相当多的冗余, 而在 cifar100 里  $k<1$  时准确率明显下降, 这说明在处理分类数较多的任务时, 特征图的冗余性会有所降低, 削减输出特征图的通道数会使得图像特征表达不充分, 从而影响最终的准确率。

表 2 不同  $k$  值的 ResNet29 的对比实验

网络结构	准确率/%	
	cifar10	cifar100
ResNet29( $k=1$ )	91.30	68.86
ResNet29( $k=3/4$ )	91.27	67.86
ResNet29( $k=1/2$ )	90.66	66.38

## 1.2 RC 结构

上节分析了瓶颈结构冗余性产生的原因, 消耗资源生成冗余的特征图明显是不合适的, 为了尽量减少这种额外的消耗, 设计了新的升维操作, 即 RC 结构。RC 结构如图 2 所示, 把输出  $Y$  分为  $y$  和  $y'$  两部分,  $y$  直接由  $x$  得到,  $y'$  由  $x$  经过  $1 \times 1$  点卷积生成, 最后把两者连接起来作为输出。

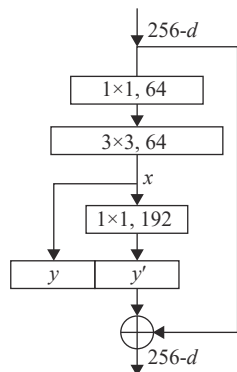


图 2 基于 RC 的瓶颈结构

RC 降低了计算量和参数, 假设点卷积的输入特征图的大小为  $n \times w \times h$ ,  $n$  为通道数, 输出特征图的通道数为  $m$ 。单样本情况下传统的升维操作运算量为:

$$l = w \times h \times n \times m$$

参数量为:

$$p = n \times m$$

RC 结构下的运算量为:

$$l' = w \times h \times n \times (m - n)$$

参数量为:

$$p' = n \times (m - n)$$

很明显  $l > l'$ 、 $p > p'$ , 升维操作的消耗大大减少。

RC 结构将  $x$  直接与输出连接起来, 反向传播时加强了这部分的梯度传递, 这使得 RC 结构在降低了消耗的同时还能增加网络精度。

### 1.3 RC 结构适用性分析

虽然 RC 模型输出的通道没有改变, 但经过  $1 \times 1$  点卷积生成的特征图减少了, 这有可能导致升维时输入的信息没有被完全表达。根据 1.2 节的假设, 传统瓶颈结构的升维操作里, 点卷积是权重形状为  $[m, n, 1, 1]$  的张量, 设其权重  $W = [w_1, w_2, \dots, w_m]$ , 其中  $w_i$  是  $[n, 1, 1]$  的张量, 则升维操作的运算可以简化为:

$$Y_1 = w_1 x \oplus w_2 x \oplus \dots \oplus w_m x = Wx \quad (1)$$

式中,  $\oplus$  表示张量在第一个维度上相互连接, 而非张量数值相加;  $w_i x$  为对  $x$  进行卷积操作, 权重为  $w_i$ 。在 RC 结构里, 点卷积权重的形状为  $[m-n, n, 1, 1]$ , 设其权重  $W' = [w'_1, w'_2, \dots, w'_{m-n}]$ , 则 RC 结构下升维操作的运算为:

$$Y_2 = w'_1 x \oplus w'_2 x \oplus \dots \oplus w'_{m-n} x \oplus x = W'x \oplus x$$

又根据卷积操作的特性, 存在一个张量  $W'' = [w''_1, w''_1, \dots, w''_n]$ , 使得:

$$x = w''_1 x \oplus w''_2 x \oplus \dots \oplus w''_n x = W''x$$

式中,  $w''_i$  为由  $n-1$  个“0”和一个“1”构造的形状为  $[n, 1, 1]$  的张量, 不同的  $w''_i$ 、1 的位置也不一样。则有:

$$Y_2 = W'x \oplus W''x = (W' \oplus W'')x \quad (2)$$

式中,  $W$  和  $W'$  是随着反向传播而不断变化的张量;  $W''$  则是固定的, 且  $(W' \oplus W'')$  与  $W$  的形状都是  $[m, n, 1, 1]$ , 因此  $(W' \oplus W'')$  是  $W$  的一种部分权重固定不变的特殊情况。当瓶颈结构存在大量冗余时, 通过  $W'$  输出的较少数量的特征图能够充分表达输入数据的特征, 引入 RC 结构后的瓶颈结构的特征表达能力就不会受到影响, 但当瓶颈结构的冗余性较低时,  $W$  比 RC 结构里的  $W' \oplus W''$  有更多可以训练的权重, 具有更好的特征表达能力, 于是就可能出现引入 RC 结构后瓶颈结构的特征表达能力不如原来的情况。因此 RC 结构更适合存在较多冗余信息的网络, 例如处理分类数较少的分类任务时。

## 2 实验与分析

实验部分首先进行了 RC 结构关于反向传播的

测试, 以验证 1.2 节的结论。而后分别在 cifar10、cifar100 和 TinyImageNet 这 3 个数据集上进行了 RC 结构的功能性验证, 实验选用的网络为 ResNet、WRN、ResNetV2、ResNext、SENet、Res2Net 和 SKNet, 最后将前面表现得优秀的网络在遥感影像数据集上进行测试以验证其迁移性。

## 2.1 RC 反向传播测试

为了验证 RC 结构对反向传播的影响, 在图 2 的结构上做了些修改, 如图 3 所示,  $x$  与一个单位阵  $E$  相乘从而得到  $y$ , 由于  $y = x \times E = x$ , 所以图 3 结构下的  $y$  与图 2 结构下的一致, 但在训练过程中利用代码中断  $y = x \times E$  这一过程里的梯度传递, 这种结构记为 RC\_S。

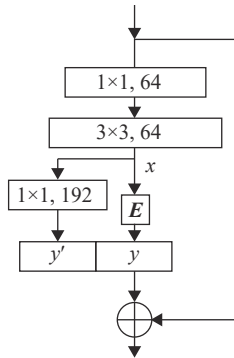


图 3 中断反向传播的 RC

如表 3 所示, 由于 ResNet 在 10 分类问题上存在大量冗余, 所以中断了  $y$  到  $x$  梯度传播的 ResNet29\_RC\_S 与 ResNet29 训练结果相似, 但引入了梯度传播后, ResNet29\_RC 的准确率比改进前提高了 1.28%。由此可见, 直接将  $x$  作为部分输出能带来精度提升, 这也是后面实验里引入 RC 结构的残差网络能比原版表现更优秀的原因。

表 3 RC 的反向传播测试

网络结构	运算量/B	参数量/M	准确率/%
			cifar10
ResNet29	0.76	4.92	91.30
ResNet29_RC_S	0.84	4.78	91.39
ResNet29_RC	0.72	4.62	92.58

## 2.2 cifar 数据集实验

cifar 数据集包括 cifar10 和 cifar100, 分别对应 10 分类和 100 分类任务。由于 cifar 的图片大小为  $32 \times 32$ , 所以本文选用 ResNext29\_32 $\times$ 4d 作为基础框架, 并在其基础上搭建了需要测试的网络。训练批次为 120, 优化器选取的是 SGD, 初始学习率为 0.1, 并且采用每隔 30 轮学习率除以 10 的调整

策略, 对于数据预处理, 采用的是随机裁剪、随机反转和标准化。对实验结果采用的是去掉极小值后取平均值的统计方法, 每个数据都是 3 组及以上样本的平均值。

首先对 ResNet、WRN、ResNetV2(PreAct-ResNet)、ResNext、Res2Net、SKNet 和引入 SE 结构的 ResNet 进行测试, 实验结果如表 4 所示。除了 SKNet 外, 其余残差网络引入 RC 后的准确率在 cifar10 上都有所增长。而在 cifar100 上, Res2Net 准确率降低了, 一方面 RC 结构在处理分类数较多的问题时会出现特征表达不完全的情况, 另一方面 Res2Net 引入的多尺度信息可能会和 RC 结构产生一定冲突, 如果这种冲突直接作用在 Linear 层, 会对分类产生较大影响。于是将 Res2Net29\_RC 的最后一个瓶颈结构替换为原来的结构, 为了方便描述, 这种针对 Res2Net 的改进记为 RC\_res2。结果显示, Res2Net\_RC\_res2 在 cifar10 的准确率有了进一步提升, 而且在 cifar100 上也超过了原版。

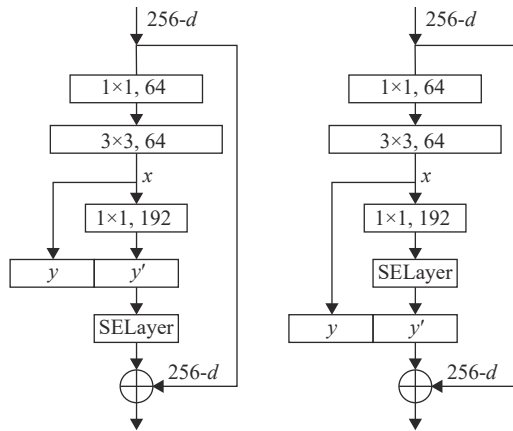
表 4 cifar 数据集的实验

网络结构	运算量/B	参数量/M	准确率/%	
			cifar10	cifar100
ResNet29	0.76	4.92	91.30	68.86
ResNet29_RC	0.72	4.66	<b>92.58</b>	<b>71.23</b>
WRN29( $d=2$ )	2.10	13.79	91.04	68.23
WRN29_RC( $d=2$ )	1.95	12.75	<b>93.26</b>	<b>73.60</b>
WRN29( $d=3$ )	4.13	27.29	92.08	70.10
WRN29_RC( $d=3$ )	3.78	24.96	<b>93.37</b>	<b>73.60</b>
ResNext29_32 $\times$ 4d	0.78	4.76	93.37	74.08
ResNext29_32 $\times$ 4d_RC	0.63	3.75	<b>93.61</b>	<b>75.19</b>
PreActResNet29	0.89	5.05	93.16	71.68
PreActResNet29_RC	0.84	4.78	<b>93.37</b>	<b>71.92</b>
Res2net29	0.99	6.28	93.84	76.59
Res2Net29_RC	0.84	5.19	94.29	74.88
Res2Net29_RC_res2	0.86	5.45	<b>94.59</b>	<b>76.64</b>
SKNet29	0.36	10.45	92.56	72.70
SKNet29_RC	0.32	9.68	92.17	72.38
SKNet29( $s=8$ )	/	/	91.05	/
SKNet29_RC( $s=8$ )	/	/	91.20	/
SEResNet29	0.76	7.00	91.03	65.45
SEResNet29_RC	0.72	6.73	93.08	66.85
SEResNet29_RC_se	0.72	5.83	<b>93.58</b>	<b>73.78</b>

除了 Res2Net 外, 也对 SEResNet 做出了特殊改进, 如图 4 所示。RC 结构之所以能提高准确率, 是因为将输入升维操作的数据直接作为输出的一部分, 从而加强了这部分的梯度传递。而在与 SELayer 结合的时候,  $y$  参与了自注意力机制的情况下可能会影响到其梯度传递, 所以设计了图 4b 的结构。为方便区分, 这种结构记为 RC\_se。表 4 显



示, 在 cifar 数据集上, SEResNet29\_RC\_se 的准确率和效率比 SEResNet29\_RC 与 SEResNet29 更加出色。



a. 直接引入 RC 结构的网络结构 b. 改进后的网络结构

图 4 针对 SEResNet 做出的特殊改进

此外, 对于 SKNet 与 RC 结合的失败, 本文也做出了各种改进, 但准确率都达不到原版的效果, 引入 RC 结构后, SKNet 的特征表达能力受到很大的负面影响。SKNet 较小的  $s$  值 (SKNet 瓶颈结构的  $s=2$ ) 是影响因素之一, 根据 1.1 节的分析, 越大的  $s$  值代表网络的冗余性越多, 与 RC 结构的结合效果就越好。本文尝试增大 SKNet 的  $s$  值, 发现  $s$  值增大时, RC 结构起到了提高准确率的作用, 这说明 SKNet 也是遵行“冗余性越高, 与 RC 结构结合效果越好”这一结论。因此 SKNet 与 RC 结构结合失败的原因目前只能归结为: SKNet 瓶颈结构的输出特征图冗余性较低, 加入 RC 结构后会影响其特征表达能力。

### 2.3 TinyImageNet 实验

TinyImageNet 是 200 类的 ImageNet 子集, 图片大小为  $64 \times 64$ , 本文以 ResNet50 作为基础, 搭建需要测试的网络, 训练相关设置和 cifar 保持一致。训练结果如表 5 所示。正如 1.3 节分析的, RC 结构不适合分类数较大的任务, 在 2.2 节 cifar100 的实验里也出现 Res2Net-RC 的准确率有所下降的情况, 而在 200 类的 TinyImageNet 里, 这一特性更加明显, 除了 SEResNet50 外, 其余网络的准确率都出现了小幅下降。所以 RC 结构并不适合处理分类数较多的任务, 且不同残差网络对应的分类数临界值也会有所不同, 需要根据具体情况判断是否使用 RC 结构。

### 2.4 更多的数据集

为了验证 RC 结构的有效性, 针对遥感影像数

据做了额外的实验, 数据集选用了 12 分类的 SIRI-WHU Dataset<sup>[12]</sup> 与 21 分类的 UC Merced Land-Use DataSet<sup>[13]</sup>, 采用 4:1 的比例划分为训练集和测试集。与前面实验不同的是, 采用 Adam 作为训练优化器, 学习率初始值为 0.001, 每隔 30 个批次除以 2, 训练 120 个批次, 选用 ResNet50 为基础网络, 实验结果如表 6 所示。SIRI-WHU Dataset 的图片大小为  $224 \times 224$ , UC Merced Land-Use DataSet 的图片大小为  $256 \times 256$ , 面对图片大小在 200 以上的数据集时, RC 在 ResNetV2(PreActResNet) 上表现得并不理想, 但在其他网络上都有一定的提升。

表 5 TinyImageNet 的实验

网络结构	运算量/B	参数量/M	准确率/%
ResNext50	1.35	23.19	56.50
ResNext50_RC	1.08	18.14	56.05
ResNet50	1.30	23.71	52.87
ResNet50_RC	1.24	22.44	52.72
Res2Net50	1.37	23.85	60.01
Res2Net50_RC_res2	1.19	20.52	60.00
SEResNet50	1.32	33.99	51.81
SEResNet50_RC_se	1.24	28.81	<b>56.31</b>
PreActResNet50	0.75	4.91	57.51
PreActResNet50_RC	0.72	4.66	57.38

表 6 遥感影像数据集的实验

网络结构	准确率/%	
	SIRI-WHU Dataset	UC Merced Land-Use
ResNet50	94.33	92.14
ResNe50_RC	<b>94.72</b>	<b>92.86</b>
ResNext50_32x4d	94.83	92.14
ResNext50_32x4d_RC	<b>96.17</b>	<b>93.01</b>
PreActResNet50	93.83	91.43
PreActResNet50_RC	93.33	91.19
Res2Net50	95.00	92.60
Res2Net50_RC_res2	<b>95.67</b>	<b>92.62</b>
SEResNet50	95.67	94.26
SEResNet50_RC	<b>96.50</b>	93.81
SEResNet50_RC_se	95.67	<b>95.48</b>
SERes2Net50	95.83	95.28
SERes2Net50_RC_res2	96.33	94.86
SERes2Net50_RC_res2_se	<b>96.50</b>	<b>95.48</b>

## 3 结束语

本文分析了残差网络的冗余性, 针对瓶颈结构进行了改进, 使用 RC 代替原来的升维操作。RC 使用残差结构和连接操作, 降低了升维操作的计算量和参数量, 并增强了梯度传播。利用 cifar 数据集和 tinyImagenet 数据集在多种残差网络上进行了测试, 验证了 RC 在较少类别分类任务上的有效性。

最后还在遥感影像识别领域做了额外的实验。本文的实验部分也存在一些问题没有得到解答,例如在 SKNet 上直接套用 RC 结构并没有发挥出理想的效果,根本原因是什么?能不能做出有效改进?两个瓶颈结构之间的升维降维操作是否存在冗余计算,可否进一步降低计算量等。RC 结构还有值得挖掘的地方,期待未来的研究能完善它。

### 参 考 文 献

- [1] XIE Sai-ning, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks[C]//2017 IEEE Conference on Computer Vision (CVPR). [S.l.]: IEEE, 2017: 5987-5995.
- [2] GAO Shang-hua, CHENG Ming-ming, ZHAO Kai, et al. Res2Net: A new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019(99): 1.
- [3] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE, 2018: 7132-7141.
- [4] LI Xiang, WANG Wen-hai, HU Xiao-lin. Selective kernel networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA: IEEE, 2019: 510-519.
- [5] ZHANG Hang, WU Chong-ruo, ZHANG Zhong-yue, et al. ResNeSt: Split-attention networks[EB/OL]. [2019-9-18]. <https://hangzhang.org/files/resnest.pdf>.
- [6] HE K, ZHANG X, REN S, et al. Identity mappings in deepresidual networks[J]. *Computer Vision and Pattern Recognition*, 2016, DOI: 10.1007/978-3-319-46493-0\_38.
- [7] DUTA I C, LIU Li, ZHU Fang, et al. Improved residual networks for image and video recognition[EB/OL]. [2019-11-20]. <https://arxiv.org/abs/2004.04989>.
- [8] ZAGORUYKO S, KOMODAKIS N. Wide residual networks[J]. *Computer Vision and Pattern Recognition*, 2016, DOI: 10.5244/C.30.87.
- [9] GAO Huang, ZHUANG Liu, LAURENS V D M, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2017: 2261-2269.
- [10] HAN Kai, WANG Yun-he, TIAN Qi, et al. GhostNet: More features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2020, DOI: 10.1109/CVPR42600.2020.00165.
- [11] HE Kai-ming, ZHANG Xiang-yu, REN Shao-qing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2016: 770-778.
- [12] ZHONG Yan-fei, ZHU Qi-qi, ZHANG Liang-pei. Scene classification based on the multifeature fusion probabilistic topic model for high spatial resolution remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015(99): 1-16.
- [13] YANG Y, NEWSAM S. Bag-of-visual-words and spatial extensions for land-use classification[C]//ACM International Symposium on Advances in Geographic Information Systems. [S.l.]: ACM, 2010: 270-279.

编辑 叶芳