

简化型残差结构和快速深度残差网络

杨怀江^{1,2}, 王二帅^{1,3}, 隋永新^{1,2}, 闫 丰^{1,2}, 周 跃^{1,2}

(1. 中国科学院 长春光学精密机械与物理研究所, 长春 130033; 2. 长春国科精密光学技术有限公司, 长春 130033; 3. 中国科学院大学 大珩学院, 北京 100049)

摘 要:为解决当前深度残差网络模型训练缓慢的问题,设计了一种新型的残差结构。与典型的残差结构相比,该结构仅含有一个 Batch Normalization 和 ReLU 模块,通过减少网络训练过程的计算量降低了耗时,提升了模型训练速度。在常用的 CIFAR10/100 图像分类数据库上进行了对比实验分析,以该方法构建的深度为 110 层的网络 CIFAR10 分类错误率为 5.29%, CIFAR100 分类错误率为 24.80%,典型的 110 层深度残差网络分类错误率分别为 5.75% 和 26.02%;在训练耗时方面,该方法平均周期耗时为 133.47 s,典型的残差网络平均周期耗时为 208.26 s,提升了 35.91%;结果表明,该网络结构在保证分类性能的基础上极大地提升了训练速度,具有较好的实用价值。

关键词:图像处理;图像识别;图像分类;卷积神经网络;深度残差网络

中图分类号:TP183 **文献标志码:**A **文章编号:**1671-5497(2022)06-1413-09

DOI:10.13229/j.cnki.jdxbgxb20210027

Simplified residual structure and fast deep residual networks

YANG Huai-jiang^{1,2}, WANG Er-shuai^{1,3}, SUI Yong-xin^{1,2}, YAN Feng^{1,2}, ZHOU Yue^{1,2}

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;
2. Changchun National Extreme Precision Optics Co., Ltd., Changchun 130033, China; 3. Daheng College, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In order to address the problem of slow training of the current deep ResNets model, a novel residual structure is designed. Compared with the typical residual structure, the structure only contains a Batch Normalization and ReLU, which reduces training time and improves the training speed by reducing the amount of calculation in the network training process. The comparative experiments are carried out on the CIFAR10/100 image classification database. The classification error rate of 110 layers networks constructed by this method on CIFAR10 and CIFAR100 is 5.29% and 24.80%, respectively. The classification error rate of 110-ResNet is 5.75% and 26.02%, respectively. Training the network takes 133.47 (this method) and 208.26 (ResNet) seconds per epoch, increased by 35.91%. The results show that the network structure greatly improves the training speed while ensuring the classification performance, and has better practical value.

收稿日期:2021-01-13.

基金项目:国家重点研发计划项目(2018YFF01011900).

作者简介:杨怀江(1966-),男,研究员,博士.研究方向:光学信息融合,深紫外光刻技术.

E-mail:994301018@qq.com

Key words: image processing; image recognition; image classification; convolutional neural networks; deep residual networks

0 引言

自2012年以来,卷积神经网络方法在图像处理领域得到广泛应用,取得了远超其他方法的优异成绩,是当前图像分类、识别、检测领域的研究热点。2015年由He等^[1]提出的深度残差网络(Deep residual networks, ResNets)是卷积神经网络中最为重要的一种模型,该模型取得了ImageNet和COCO2015图像竞赛中分类、识别和检测等多个项目的第一名,引起学者们对该模型的广泛关注,随后人们开展了关于ResNets的大量研究工作。2016年,He等^[2]又提出了Pre-activation ResNets模型,对ResNets模型残差结构中的激活层和卷积层顺序做出调整,进一步提高了ResNets模型的收敛速度和性能。Targ等^[3]提出了Resnet in Resnet,将ResNets推广为多个残差分支结构。Zagoruyko等^[4]提出了Wide-ResNets,通过大幅增加残差结构中的卷积核数量,提升了网络的分类性能。Veit等^[5]和Wu等^[6]深入分析了ResNets,认为该模型表现为大量较浅网络的集合体,其性能主要依赖深度为10~34层的部分网络。国内方面有大量研究人员将ResNets模型用于交通信号灯^[7,8]、鱼类图像^[9]、医学图像^[10,11]、人脸识别^[12-14]、手势识别^[15]、昆虫识别^[16,17]等多种特定图像分类任务中。这些关于ResNets的研究主要探讨了ResNets的原理,推广了ResNets模型在图像分类识别领域的应用。

然而,当前训练ResNets模型仍面临着耗时长的问题,ResNets训练耗时过长消耗了大量的计算资源和时间成本,限制了ResNets在图像分类、图像识别领域的应用。例如,在ImageNet数据库采用多块高性能的GPU同时训练深度152层的ResNets也要花费数周时间。目前,减少训练用时的方法主要是微调法和Stochastic depth法,微调法即以其他大型数据库预先训练的公开模型为基础,利用任务数据进一步训练,对模型参数做微调。当任务数据与公开模型数据库差异较大时,其分类效果往往不太理想。Stochastic depth是Huang等^[18]提出的,在模型训练过程中随机跨过若干残差结构,通过控制每个残差结构的

保存概率变相缩减ResNets的深度,以此达到提升网络训练速度的目的;在测试时采用所有残差块,相当于构成一组具有不同深度的ResNets网络,保证了模型的分类性能。

本文在Pre-activation ResNets的基础上细致地探讨了残差结构中激活层的功能和必要性,提出了仅含有一层Batch Normalization和ReLU单元的简化型残差结构,并以此构建深度残差网络。在CIFAR数据库上测试了新型残差结构的性能,结果表明,该网络结构能够在保证分类性能与传统ResNets相当的同时,大幅降低训练耗时。

1 ResNets模型

ResNets模型由多个残差结构堆叠而成,其中残差结构可以用式(1)表示:

$$Y = X_i + F(X_i, W_i) \quad (1)$$

式中: Y 和 X_i 分别为第 $i(i=1, 2, \dots, n)$ 个残差块的输出和输入; $F(\cdot)$ 和 W_i 分别为第 i 个残差结构的残差函数和相关参数。

1.1 残差结构

残差结构主要有基本型和瓶颈型两种:基本型包含两个依次相连的激活层-卷积层结构;瓶颈型包含3个依次相连的激活层-卷积层结构。其中,激活层包括批归整化(Batch normalization, BN)部分和线性整流单元(Rectified linear units, ReLU)部分,如图1所示。

与基本型残差结构相比,瓶颈型残差结构包含更多的卷积层和激活层,以此构建的深度残差网络耗时更多,因此本文仅考虑基本型残差结构。基本型残差函数可以用式(2)表示:

$$F(X_i, W_i) = W_{i,2}(f_2(W_{i,1}f_1(X_i))) \quad (2)$$

式中: $W_{i,1}$ 、 $W_{i,2}$ 均为残差结构中的卷积层参数; $f_i(\cdot)$ 为残差结构中的激活函数。

卷积层通过对特征图层各像素点邻域内的若干特征值做加权计算,在训练过程中按误差反向传播,计算各参数的偏导数,据此更新参数,最终达到特征提取的目的。以常用的 3×3 卷积核为例,对应的卷积计算和参数更新如式(3)所示:

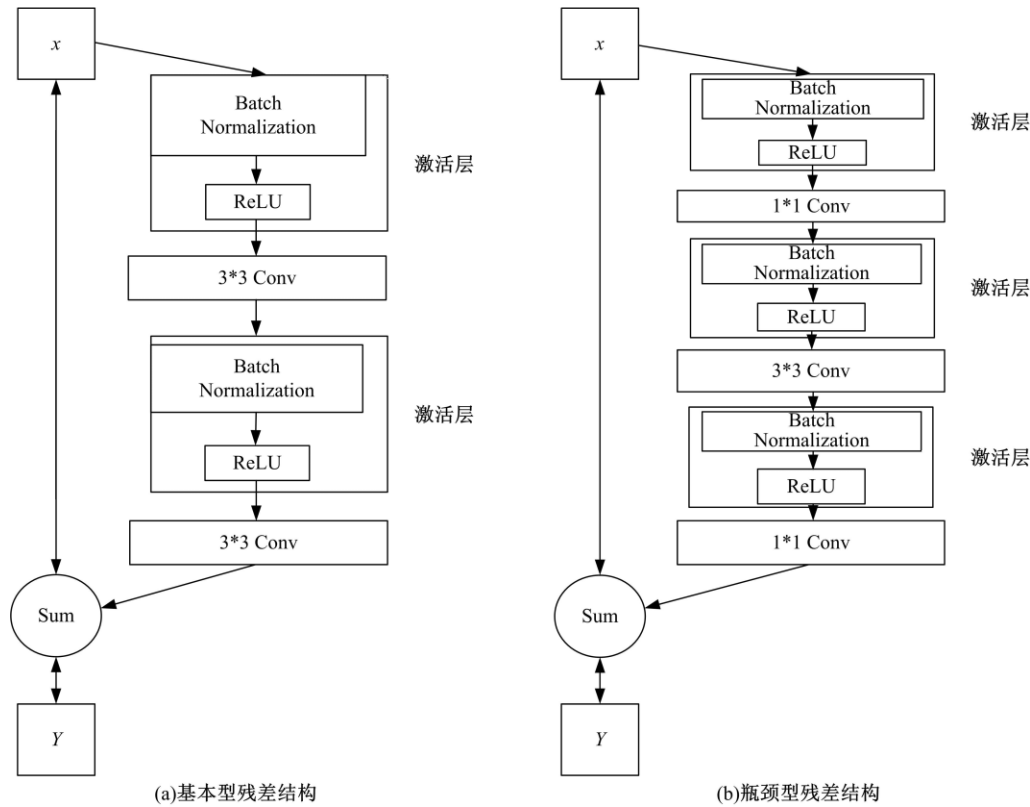


图1 基本型与瓶颈型残差结构

Fig. 1 Basic and bottleneck residual structure

$$\begin{cases} y_{c,h,w} = \sum_{d=1}^N \sum_{i=-1}^1 \sum_{j=-1}^1 w_{c,d,i,j} x_{d,h+i,w+j} + b_c \\ \Delta w_{c,d,i,j} = \sum_{h,w} \Delta y_{c,h,w} x_{d,h+i,w+j} \\ \Delta b_c = \sum_{h,w} \Delta y_{c,h,w} \\ \Delta x_{d,h,w} = \sum_{d=1}^N \sum_{i=-1}^1 \sum_{j=-1}^1 \Delta y_{d,h-i,w-j} w_{c,d,i,j} \end{cases} \quad (3)$$

式中: $y_{c,h,w}$ 为输出, 下标 c 代表输出第 c 通道, h 代表特征图第 h 行, w 代表特征图第 w 列; $w_{c,d,i,j}$ 为卷积核权值, d 代表输入第 d 通道, i 和 j 代表卷积核在行和列的偏移量; $x_{d,h,w}$ 为输入; b_c 为偏差; Δ 代表微分。

1.2 残差结构中的激活单元

残差结构中激活层由 BN 和 ReLU 组成, 下面分别讨论这两种结构元素的功能。

1.2.1 ReLU

关于神经系统科学的研究^[19]指出, 皮质神经元的激活机制能够用整流器近似, 这一想法在卷积神经网络中得以应用。Krizhevsky 等^[20]首次在卷积神经网络中应用了 ReLU 激活函数(见图2), 具体如式(4)所示:

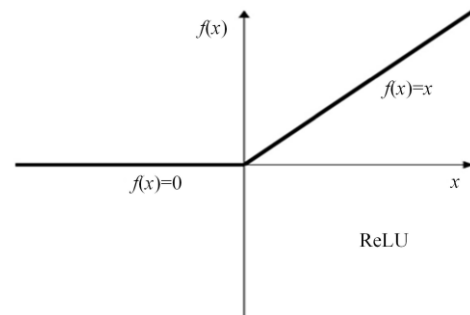


图2 ReLU 示意图

Fig. 2 Schematic diagram of ReLU

$$f(x) = \max(0, x) \quad (4)$$

特征数据经过 ReLU 单元, 产生负值的部分被强制输出为 0, 正值部分保持不变。采用 ReLU 激活函数时, 网络模型的收敛速度比采用 sigmoid 和 tanh 等激活方法时快很多。ReLU 由于其实现简单, 收敛快的特点被 NiN^[21]、AllCNN^[22]、Inception^[23]、Resnet 等诸多网络模型采用。在经典的残差结构当中, 每个激活部分都包含一个 ReLU 层, 其目的是提升卷积神经网络的非线性能力。然而, 如文献[2]所提到, 在残差结构尾部附加 ReLU, 会导致残差函数仅能产生非负偏置, 影响残差网络的性能, 严重时会使网络无法收敛。

文献[2]指出 ReLU 置于卷积层之前能有效提高残差结构提取非线性特征的能力,从而促进残差网络的收敛,改善其分类性能。

1.2.2 BN

卷积神经网络中 BN 结构是 Ioffe 等^[24]在 2015 年提出的,该结构将卷积层的输入规正化为近似高斯分布,从而使得训练更为容易,在参数初始化、学习率设置方面也更为宽松,提高了网络的性能。随后 BN 在 VGG^[25]、Inception、ResNet 等模型中得到广泛应用。

记 X_i 和 Y_i 为 BN 操作对应的输入和输出,则 BN 可以表示为:

$$\begin{cases} \hat{X}_i = \frac{X_i - E(X_i)}{\sqrt{\text{Var}(X_i)}} \\ Y_i = \text{BN}_{\gamma, \beta}(X_i) = \gamma \hat{X}_i + \beta \end{cases} \quad (5)$$

式中: E 、 Var 、 γ 、 β 分别为期望、方差、尺度缩放和偏置参数,在参数初始化时 γ 、 β 被设置为 1 和 0。 $\text{BN}_{\gamma, \beta}$ 代表参数为 γ 、 β 的 Batch-Normalization 操作。

网络训练期间, BN 需要多次重复计算输入图层的期望和方差,并据此调节 γ 和 β , 最终达到稳定模型的功能。研究表明^[26], BN 方法成功的根本原因在于, BN 是网络模型参数优化问题下的参数重新确定过程,该过程使得参数解空间更为平滑和稳定。

当前的卷积神经网络模型中 BN 与卷积层一般是组合使用,然而在正向特征传递过程的每个批次, BN 都要重新计算特征图层输入的期望和方差,在反向梯度传输过程的每个批次, BN 也要计算各参数的偏导,带来了巨大的额外计算量需求,使得模型训练的单个周期耗时明显提升。

2 简化型残差结构

2.1 两种简化型残差结构

ReLU 和 BN 的主要作用是提升网络的非线性能力和网络的稳定性。然而,第 1 激活层的输入是残差网络中上一个残差结构的复合输出,它受此前残差网络的影响,已经具有较强的非线性能力,分布也比较稳定。本文认为第 1 激活层对于一个残差结构分支来说几乎没有增强其线性能力的作用,尤其是对于深度较深的网络;另一方面,第 1 激活层中引入了 BN,带来了额外的计算量,必然降低训练速度。第 2 激活层的输入是残差结构中第 1 卷积层的输出,其分布会受到残差结构中第 1 卷积层参数更新的直接影响,因此有必要采用 BN 和 ReLU 提高其稳定性和非线性能力。

图 3 为本文提出的两种简化残差结构,其达到了提升训练速度的目的,其中图 3(a)为删除第 1 激活层所得,图 3(b)为删除第 2 激活层所得。

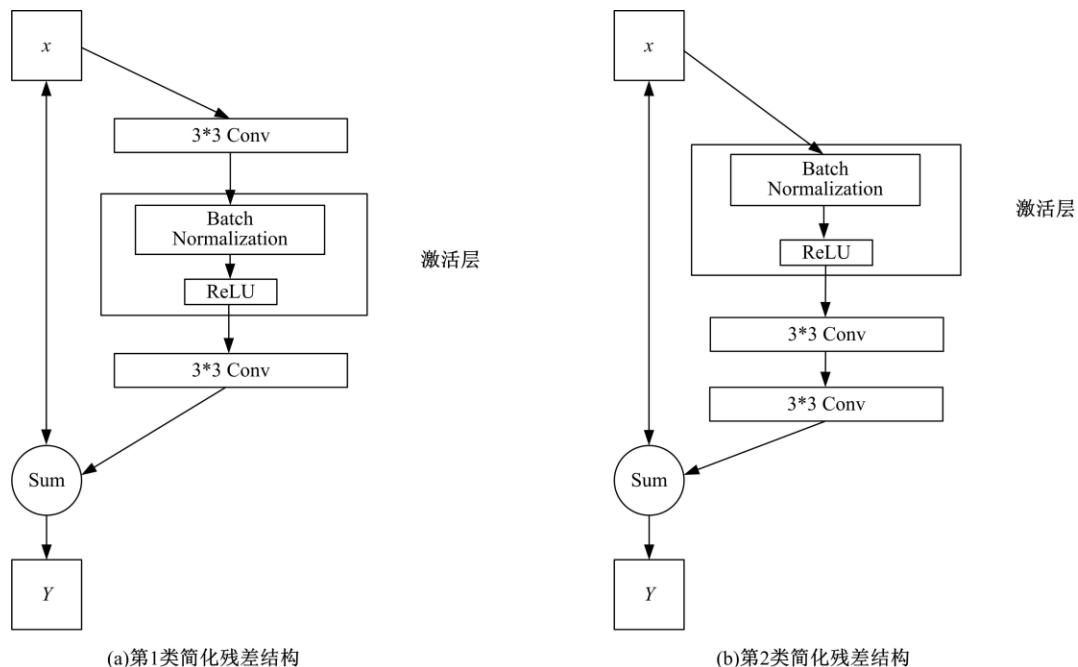


图 3 两种简化型残差结构

Fig. 3 Two kinds of simplified residual structure

与传统的残差结构相比,这两种残差结构减少了残差分支中激活层的数量。这种改动不仅减少了近一半数量的激活层数据迁移,节省了训练过程中的内存开支;还减少了残差分支中BN层对输入期望和方差的计算量,降低了模型训练耗时。

2.2 原理分析

基本型残差结构与简化残差结构的区别在于激活函数。基本型的激活函数可以看作BN与ReLU的耦合;简化残差结构中被删减的激活层可以用同等变换代替,其他部分与基本型完全一致。

记残差结构输入为 X ,第1卷积层输入为 X_1 ,第2卷积层输入为 X_2 ,输出为 Y ,对应的卷积操作和卷积参数为 $\text{Conv}(X, w, b)$,结合式(1)~(5),有基本型残差结构函数如式(6)所示,第1类简化残差结构和第2类简化残差结构如式(7)(8)所示。

$$\begin{cases} X_1 = \max(0, \text{BN}(X)) \\ X_2 = \max(0, \text{BN}(\text{Conv}(X_1, w_1, b_1))) \\ Y_{\text{base}} = X + \text{Conv}(X_2, w_2, b_2) \end{cases} \quad (6)$$

$$\begin{cases} X_1 = X \\ X_2 = \max(0, \text{BN}(\text{Conv}(X_1, w_1, b_1))) \\ Y_{1\text{st}} = X + \text{Conv}(X_2, w_2, b_2) \end{cases} \quad (7)$$

$$\begin{cases} X_1 = \max(0, \text{BN}(X)) \\ X_2 = \text{Conv}(X_1, w_1, b_1) \\ Y_{2\text{nd}} = X + \text{Conv}(X_2, w_2, b_2) \end{cases} \quad (8)$$

式中: Y_{base} 为基本型残差结构; $Y_{1\text{st}}$ 为第1类简化型残差结构; $Y_{2\text{nd}}$ 为第2类简化型残差结构; b_i 为残差结构中第 i 卷积层的偏差, $i=1,2$; w_i 为第 i 卷积层的权值, $i=1,2$ 。

随着训练周期的增加,模型越来越稳定,残差结构输出的特征值大多数为正值,此时基本型残差结构中第1激活层对输入的影响可以近似为仿射变换,第1激活层和第1卷积层的耦合作用可以等效为1个卷积层,如式(9)所示:

$$\begin{cases} X_1 = \max(0, \text{BN}(X)) \gg \gamma X + \beta \\ \text{Conv}(X_1, w_1, b_1) = \text{Conv}(X, \gamma w_1, \gamma b_1 + \beta) \end{cases} \quad (9)$$

结合式(6)(9),得到基本型残差结构如式(10)所示:

$$\begin{cases} X_2 = \max(0, \text{BN}(\text{Conv}(X, \gamma_1 w_1, \gamma_1 b_1 + \beta_1))) \\ Y_{\text{基本型}} = X + \text{Conv}(X_2, w_2, b_2) \end{cases} \quad (10)$$

对比式(7)和式(10)发现两式具有相似的形

式,两者差异仅为卷积层参数,当 γ, β 分别为1和0时两式完全相等,即第1类简化残差结构能够取代基本型残差结构。

结合式(8)(9)得到第2类简化残差结构如式(11)所示:

$$\begin{cases} X_2 = \text{Conv}(X, \gamma_1 w_1, \gamma_1 b_1 + \beta_1) \\ Y_{\text{第2类}} = X + \text{Conv}(X_2, w_2, b_2) \end{cases} \quad (11)$$

即经过一段时间的训练模型较为稳定时,第2类简化残差结构其残差分支近似为两个连续的卷积层。连续两个小尺寸卷积核的作用可以等效为一个尺寸大的卷积核作用,两个连续的 3×3 卷积核等效为一个 5×5 的卷积核的参数替换方法如式(12)所示:

$$\begin{cases} X_{2,h,w} = \sum_{i_1,j_1=-1}^1 w_{1,i_1,j_1} X_{1,h+i_1,w+j_1} + b_1 \\ Y = \sum_{i_2,j_2=-1}^1 w_{2,i_2,j_2} X_{2,h+i_2,w+j_2} + b_2 \\ b_3 = \sum_{i_2,j_2=-1}^1 w_{2,i_2,j_2} b_1 + b_2 \\ w_{3,i_3,j_3} = \sum_{i_1+i_2=i_3,j_1+j_2=j_3} w_{1,i_1,j_1} w_{2,i_2,j_2} \\ Y = \sum_{i_3,j_3=-2}^2 w_{3,i_3,j_3} X_{1,h+i_3,w+j_3} + b_3 \end{cases} \quad (12)$$

式中: i, j 为卷积核在行和列的偏移量; w_1, w_2 为尺寸为 3×3 的卷积核; w_3 是尺寸为 5×5 的卷积核; b 为对应的偏差。

结合式(11)(12),可得第2类简化残差结构为:

$$Y_{\text{第2类}} = X + \text{Conv}(X, w_3, b_3) \quad (13)$$

式(13)表明此时第2类简化残差结构退化成残差分支仅包括一个卷积核尺寸为 5×5 的卷积层,如图4所示。

由式(3)可知,单一的卷积层近似为仿射变换,其输出是对应位置输入的加权和,当偏移量参数 b 较小时,卷积操作可以看作是线性变换。即第2类简化残差结构因缺乏卷积层间的激活单元,其非线性能力受到影响,在模型较为稳定时第2类简化残差结构几乎可以看作线性变换,无法通过训练进一步提升提取非线性特征的能力,从而影响模型的性能。

综上所述,本文认为利用第1类简化残差结构训练深度残差网络,能够在保持网络分类性能的前提下,减少训练过程中的计算量,降低内存开

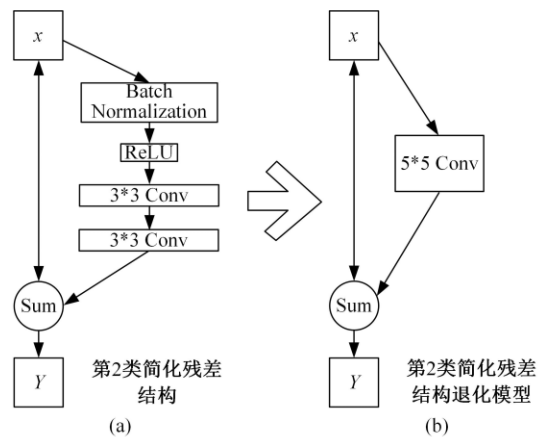


图 4 第 2 类简化型残差结构退化模型

Fig. 4 Degradation model of the second simplified residual structure

销、提高网络训练速度;第 2 类简化残差结构会造成网络分类性能下降,在其他方面与第 1 类简化残差结构一致。

3 实验分析

本文利用 CIFAR10/100 数据集分别训练了几组分类模型,详细对比了基本型 ResNet 与基于简化残差结构的快速深度残差网络的准确率、训练周期耗时等性能。

3.1 CIFAR 数据集

CIFAR10/100 图像数据集由尺寸为 32×32 的彩色图像组成。CIFAR10 包括飞机、汽车、鸟、猫、鹿、狗、青蛙、马、轮船、卡车等 10 个常见类别,每个类别含有 5000 张训练图像和 1000 张测试图像,最终构成了共包括 50 000 张图像的训练样本库和 10 000 张图像的测试样本库;CIFAR100 包括更为精细的 100 个图像类别,每个类别含有 500 张训练图像和 100 张测试图像。经过严格的人工筛选和标记,每张图像都带有唯一正确的标签。图像数据增强是卷积神经网络模型降低过拟合现象最常用的手段,本文在图像数据增强方面采用常规的图像裁剪和随机水平翻转;在图像数据预处理方面,本文利用训练图像的各通道像素均值和方差归一化整个数据库。

3.2 快速深度残差网络及实验环境

由于简化型残差结构与基本型残差结构相比,其训练计算量较少,使得训练用时得以缩减,因此本文将利用简化型残差结构设计的深度残差网络称为快速深度残差网络。该模型由低级特征提取层、中间层和分类器 3 部分组成。其中,低级

特征提取层即为一个卷积层;中间层分为 3 个阶段,每个阶段由相同数量的残差块组成;分类器结构为 BN-ReLU-全局平均池化层-全连接层。记中间层每个阶段的残差块数量为 n ,因为残差块含有两个卷积层,低级特征提取层含有一层卷积层,分类器中的全连接层也算作一层卷积层,则网络模型的总深度(仅计算卷积深度) $D=6n+2$ 。CIFAR 数据集图像尺寸较小,采用的卷积核尺寸均为 3×3 。低级特征提取层的卷积核数量与第 1 阶层中卷积核数量都为 16;第 2 阶层卷积核数量为 32;第 3 阶层卷积核数量为 64。模型训练的批尺寸为 128;权重衰减为 10^{-4} ;参数更新方法为 Nesterov momentum, momentum 权重为 0.9;训练周期数为 160;学习率前 80 个周期为 0.1,81~120 周期为 0.01,121~160 周期为 0.001。以上参数设置与 ResNet 完全一致,本文在 CIFAR10 上训练的深度为 110 层的快速深度残差网络模型示意图如图 5 所示。

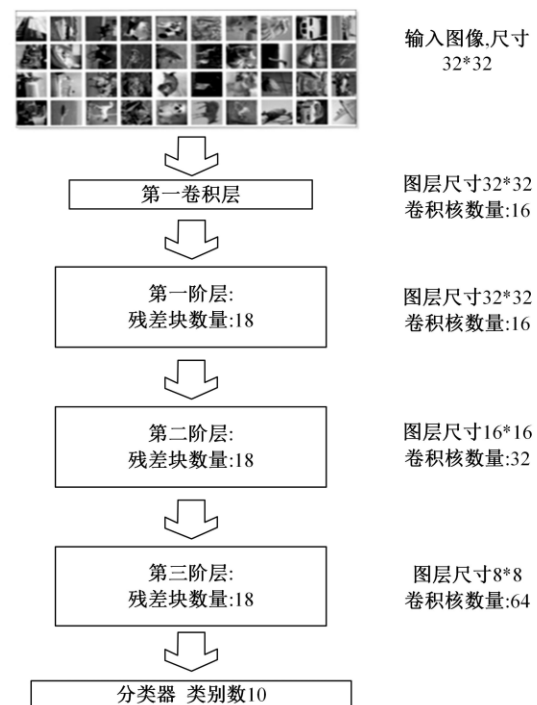


图 5 110 层快速深度残差网络模型示意图

Fig. 5 Schematic diagram of fast residual network with 110 layers

实验采用的操作系统为 Windows 7;CPU 为 intel core i3-8350k, 4.0 GHz, 16 GB 内存;GPU 为 NVIDIA GTX 1070, 8 GB 显存;卷积神经网络开发框架采用基于 Matlab 的 matconvnet-1.0-beta25。

3.3 两种简化型残差结构实验结果

为考察两种简化型残差结构的实际性能,本文分别利用基本型残差结构、第1类简化型残差结构、第2类简化型残差结构组建了 $n=3$ 的残差网络,即网络深度 $D=20$ 层。

本文对3种网络模型分别训练了5次,以减小网络参数初始化对网络分类性能的影响。然后,考察每种网络的平均错误率和平均周期用时,结果如表1所示。从每种网络模型中选取具有最小测试错误率的一个,绘制了如图6所示的测试分类错误率曲线。

表 1 三种残差结构的分类错误率和周期耗时对比

Table 1 Comparison of classification error rate and training time per epoch of three residual structures

残差网络	分类错误率/%	周期耗时/s
基本型	7.62 ± 0.15	30.07
第1类残差结构	7.68 ± 0.20	22.08
第2类残差结构	10.04 ± 0.18	22.10

从表1可以看出:第1类简化型残差结构错误率为7.68%与典型残差结构的错误率(7.62%)相当;周期耗时为22.08 s,比典型残差结构的周期耗时(30.07 s)缩减了26.68%,大幅提高了训练速度。第2类简化型残差结构周期耗时为22.10 s,与第1类简化型残差结构相当,但其错误率为10.04%,表现较差。

从图6可以看出:随着训练周期的增加,3种网络模型分类错误率持续波动下降并逐渐趋于稳定,因学习率的缩小使得网络参数微调步长变短,分类错误率在第81周期和第121周期出现了明显下降。第1类简化型残差结构与基本型残差结

构曲线基本重合,在前80周期因初始学习率较大,分类错误率曲线有小范围起伏,后80个周期起伏很小,稳定在7.6%附近;第2类简化型残差结构稳定性较差,分类错误率在前80个训练周期有很大起伏,121周期之后稳定在10%附近,明显高于前两者。

以上实验结果与前文分析一致,表明第1类简化型残差结构能够替代基本型残差结构,在保持网络分类性能的前提下大幅缩减网络训练耗时。

3.4 快速深度残差网络实验结果

为进一步验证第1类简化型残差结构组建的快速深度残差网络随深度加深的表现,在CIFAR10和CIFAR100数据集上设计了卷积深度分别56层和110层的快速深度残差网络,即每个阶层残差块数量 $n=9$ 和 $n=18$ 。由于训练用时随着网络深度的增加大幅增加,为节省时间,这些网络模型仅训练一次。在CIFAR10和CIFAR100数据集上分别训练160个周期后,考察了两类模型分类错误率和周期耗时,如表2所示。绘制了110层快速深度残差网络和ResNet的测试分类错误率曲线,如图7和图8所示。

表 2 基本型深度残差网络与快速深度残差网络对比

Table 2 Comparison of basic ResNet and fast deep residual networks(FastResNet)

残差网络	CIFAR10 错误率/%	CIFAR100 错误率/%	周期耗时/s
56层 ResNet	6.31	27.93	84.22
56层 FastResNet	5.81	26.91	58.84
110层 ResNet	5.75	26.02	208.26
110层 FastResNet	5.29	24.80	133.47

由表2可以看出:训练相同深度的快速深度残差网络和基本型深度残差网络,快速深度残差网络耗时大幅降低。其中,深度为56层的快速深度残差网络周期耗时为58.84 s,基本型深度残差网络周期耗时为84.22 s,相对减少了30.04%;深度为110层的快速残差网络周期耗时133.47 s,基本型周期耗时208.26 s,相对减少35.91%。在分类错误率方面,快速深度残差网络也取得比基本型深度残差网络略好的效果,110层快速深度残差网络分类错误率为5.29%,基本型分类错误率为5.75%,相对提升了8%。

在CIFAR10数据集上,本文还训练了深度达到668层的快速残差网络和基本型深度残差网

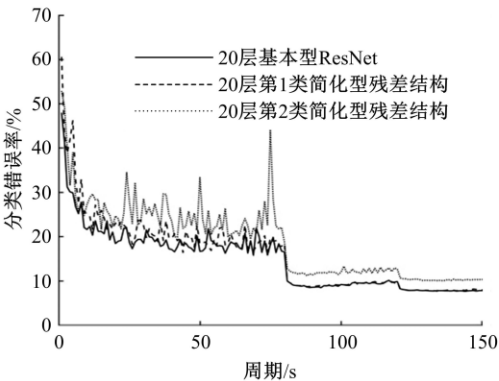


图 6 三种残差结构在 CIFAR10 上的测试分类错误率曲线

Fig. 6 Test classification error cure of the three residual structures on CIFAR10

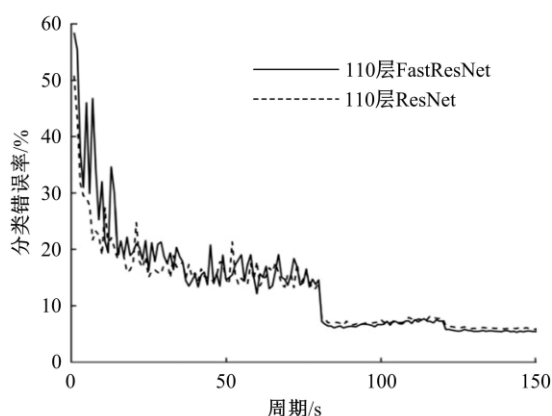


图7 110层 ResNet与 FastResNet在 CIFAR10上的分类错误率曲线

Fig. 7 Classification error curve of 110-ResNet and 110-FastResNet on CIFAR10

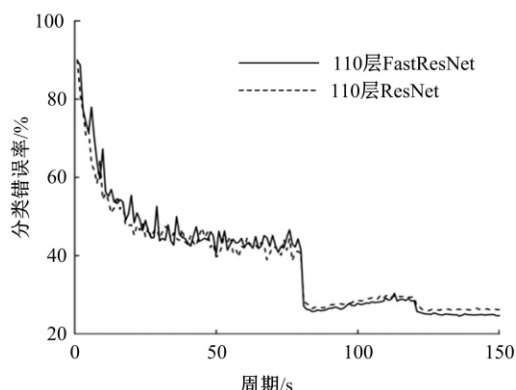


图8 110层 ResNet与 FastResNet在 CIFAR100上的分类错误率曲线

Fig. 8 Classification error curve of 110-ResNet and 110-FastResNet on CIFAR100

网络,其中快速深度残差网络分类错误率为5.09%,周期耗时为2020.92 s,基本型深度残差分类错误率为6.03%,网络周期耗时为3953.02 s;网络训练速度相对提升了48.88%。结果表明,网络深度越深,快速深度残差网络训练耗时提升的幅度越明显,其分类错误率也越低。此外,668层基本型深度残差网络分类错误率高于110层的5.75%,分类性能出现退化,即过拟合现象严重。本文认为超深的基本型深度残差模型包含过多的激活层,归一化参数繁多,导致出现较为严重的过拟合现象;快速深度残差网络减少了激活层的数量,一定程度上避免了这一问题。

4 结束语

本文提出了移除基本型残差结构中部分激活层的简化型残差结构,以此构建了快速深度残差

网络,并从理论分析和实验论证两方面验证了简化型残差结构的合理性。实验表明:删除残差结构第1激活层不会降低网络的分类性能,删除第2激活层会使得网络模型退化,导致分类错误率上升;训练快速深度残差网络比基本型深度残差网络周期耗时少,这一优势随着网络深度的加深更为突出;其分类性能也比基本型残差网络要好,网络深度越深,快速深度残差网络分类错误率越低,一致性越好。本文提出的快速深度残差网络方法在保证分类性能的前提下能够大幅降低残差网络训练时间成本,具有一定的实际价值。

参考文献:

- [1] He Kai-ming, Zhang Xiang-yu, Ren Shao-qing, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, LasVegas, NV, USA, 2016: 770-778.
- [2] He Kai-ming, Zhang Xiang-yu, Ren Shao-qing, et al. Identity mappings in deep residual networks[C]// European Conference on Computer Vision, Cham, 2016: 630-645.
- [3] Targ S, Almeid A D, Lyman K. Resnet in resnet: Generalizing residual architectures[EB/OL]. [2016-03-25]. <https://arxiv.org/pdf/1603.08029>
- [4] Zagoruyko S, Komodakis N. Wide residual networks[EB/OL]. [2017-06-14]. <https://arxiv.org/pdf/1605.07146>
- [5] Veit A, Wilber M, Belongie S. Residual networks behave like ensembles of relatively shallow networks [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Newyork: Curran Associates Inc, 2016:550-558.
- [6] Wu Z F, Shen C H, van den Hengel A. Wider or deeper: revisiting the resnet model for visual recognition[EB/OL]. [2016-11-30]. <https://arxiv.org/pdf/1611.10080>
- [7] 王若瑜. 基于 Resnet-50 的智能驾驶红绿灯分类研究[J]. 电子测试, 2019, 25(增刊1):143-145.
Wang Ruo-yu. Research on classification of intelligent driving traffic light based on Resnet-50[J]. Electronic Test, 2019, 25(Sup. 1):143-145.
- [8] 宋倩, 黄昶, 余慧瑶. 基于 TensorFlow 的交通标志形状识别[J]. 信息通信, 2017, 32(12): 286-288.
Song Qian, Huang Chang, Yu Hui-yao. Traffic sign shape recognition based on tensorflow[J]. Information & Communications, 2017, 32(12): 286-288.

- [9] 王文成, 蒋慧, 乔倩, 等. 基于 ResNet50 网络的十种鱼类图像分类识别研究[J]. 农村经济与科技, 2019, 30(19): 60-62.
Wang Wen-cheng, Jiang Hui, Qiao Qian, et al. Research on classification and recognition of ten fish images based on Resnet-50[J]. Rural Economy and Science-Technology, 2019, 30(19): 60-62.
- [10] 常川. 基于 ResNet 深度网络的人类蛋白质图谱图像分类方法研究[J]. 医学信息学, 2019, 40(7): 45-49.
Chang Chuan. Study of image classification method of human protein atlas based on resnet deep network[J]. Journal of Medical Intelligence, 2019, 40(7): 45-49.
- [11] 王恒, 李霞, 刘晓芳, 等. 基于 ResNet50 网络的乳腺癌病理图像分类研究[J]. 中国计量大学学报, 2019, 30(1): 72-77.
Wang Heng, Li Xia, Liu Xiao-fang, et al. Classification of breast cancer histopathological images based on ResNet50 network[J]. Journal of China Jiliang University, 2019, 30(1): 72-77.
- [12] 张枫, 田联房, 杜启亮. 基于残差网络与中心损失的人脸识别[J]. 计算机工程与设计, 2019, 40(6): 1689-1695.
Zhang Feng, Tian Lian-fang, Du Qi-liang. Face recognition based on Resnet and center loss[J]. Computer Engineering and Design, 2019, 40(6): 1689-1695.
- [13] 刘媛. 基于深度学习的说话人识别研究[D]. 上海: 上海交通大学信息与通信工程学院, 2015.
Liu Yuan. Research on deep leaning based speaker recognition[D]. Shanghai: School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, 2015.
- [14] 朱超平, 杨艺. 基于 YOLO2 和 ResNet 算法的监控视频中的人脸检测与识别[J]. 重庆理工大学学报: 自然科学, 2018, 32(8): 170-175.
Zhu Chao-ping, Yang Yi. Face Detection and recognition in monitoring video based on YOLO2 and ResNet algorithm[J]. Journal of Chongqing University of Technology(Natural Science), 2018, 32(8): 170-175.
- [15] 熊才华, 巩言丽, 廉华, 等. 基于 ResNet-50 改进的 Faster R-CNN 手势识别算法[J]. 计算机时代, 2019, 36(9): 1-4.
Xiong Cai-hua, Gong Yan-li, Lian Hua, et al. An improved Faster R-CNN hand gesture recognition algorithm based on ResNet-50[J]. Computer Era, 2019, 36(9): 1-4.
- [16] 刘子毅. 基于图谱特征分析的农业虫害检测方法研究[D]. 杭州: 浙江大学生物系统工程与食品科学学院, 2017.
Liu Zi-yi. Detection of agricultural pest insects based on imaging and spectral feature analysis[D]. Hangzhou: College of Biosystems Engineering and Food Science, Zhejiang University, 2017.
- [17] 张银松, 赵银娣, 袁慕策. 基于改进 Faster-RCNN 模型的粘虫板图像昆虫识别与计数[J]. 中国农业大学学报, 2019, 24(5): 115-122.
Zhang Yin-song, Zhao Yin-di, Yuan Mu-ce. Insect identification and counting based on an improved Faster-RCNN model of the sticky board image[J]. Journal of China Agricultural University, 2019, 24(5): 115-122.
- [18] Huang Gao, Sun Yu, Liu Zhuang, et al. Deep networks with stochastic depth[C]//European Conference on Computer Vision, Cham, 2016: 646-661.
- [19] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks[C]//Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Canada, 2010: 249-256.
- [20] Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [21] Lin Min, Chen Qiang, Yan Shui-cheng. Network in network[EB/OL]. [2014-03-04]. <https://arxiv.org/pdf/1312.4400>
- [22] Springenberg J, Dosovitskiy A, Brox T, et al. Striving for simplicity: the all convolutional net[EB/OL]. [2015-04-13]. <https://arxiv.org/pdf/1412.6806>
- [23] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, 2015: 1-9.
- [24] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB/OL]. [2015-03-02]. <https://arxiv.org/pdf/1502.03167>
- [25] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2015-04-10]. <https://arxiv.org/pdf/1409.1556>
- [26] Santurkar S, Tsipras D, Ilyas A, et al. How does batch normalization help optimization? [EB/OL]. [2021-01-02]. <https://proceedings.neurips.cc/paper/2018/file/905056c1ac1dad141560467e0a99e1cf-Paper.pdf>