

浅析ResNet50

苏庆华*, 于淼淼*, 张一晨, 杨翼臣, 刘智源, 阿依加木·赛培

北京物资学院, 北京

收稿日期: 2022年9月2日; 录用日期: 2022年10月2日; 发布日期: 2022年10月11日

摘要

针对深层次的神经网络训练问题, 提出了简化对更深网络的训练, 提升速度和精度的残差网络(ResNet), ResNet作为卷积神经网络的一个里程碑式的模型。ResNet50为ResNet的一个代表, 其残余网络为了解决梯度消失, 深层网络训练的问题, 使用了残差连接提起残差特征, 作为该层输入中学习特征的减法, 使得其在各个领域的应用更为广泛。

关键词

图像处理, 模型分类, ResNet50

Brief Analysis of ResNet50

Qinghua Su*, Miaomiao Yu*, Yichen Zhang, Yichen Yang, Zhiyuan Liu, Ayijiamu Sanpei

Beijing Wuzi University, Beijing

Received: Sep. 2nd, 2022; accepted: Oct. 2nd, 2022; published: Oct. 11th, 2022

Abstract

Aiming at the problem of deep neural network training, this paper proposes a residual network (ResNet), which simplifies the training of deeper networks and improves the speed and accuracy. ResNet is a landmark model of convolutional neural networks. ResNet50 is a representative of ResNet. In order to solve the problems of gradient disappearance and deep network training, its residual network uses residual connection to lift residual features as the subtraction method of learning features in the input of this layer, which makes it more widely used in various fields.

Keywords

Image Processing, Model Classification, ResNet50

*通讯作者。



1. 引言

随着 CNN 的不断发展,为了获取深层次的特征,卷积的层数也越来越多。一开始 Le Net 网络只有 5 层,接着 AlexNet 为 8 层,后来 VggNet 网络包含了 19 层,GoogleNet 已经有了 22 层。但通过增加网络层数的方法来增强网络的学习能力的方法并不总是可行的,因为网络层数到达一定的深度之后,再增加网络层数,那么网络就会出现随机梯度消失的问题,也会导致网络的准确率下降。为了解决这一问题,传统的方法是采用数据初始化和正则化的方法,这解决了梯度消失的问题,但是网络准确率的问题并没有改善。而 ResNet 残差网络的出现可以解决梯度问题,而网络层数的增加也使其表达的特征也更好,相应的检测或分类的性能更强,再加上残差中使用了 1×1 的卷积[1],这样可以减少参数量,也能在一定程度上减少计算量。

本文针对 ResNet 特点及应用现状,首先介绍 ResNet 残差网络,接着对 ResNet50 进行了描述,重点对 ResNet 的应用现状进行分析,最后是对 ResNet 模型发展及应用的总结。

2. ResNet 模型

2.1. ResNet 残差网络

残差网络(ResNet)是在简单网络的基础上,其特点通过插入快捷链接将其转化为对应的残差,其不直接拟合目标,而是拟合残差。

ResNet 引入“shortcut connection”,其中包含一个“shortcut connection”的多层网络被称为一个残差块(shortcut connection)。若模型是在函数 $F(X)$ 的空间寻参,那么残差网络就是在 $x + f(x)$ 的空间寻参。其中,输入 X ,分为两路, X 为恒等映射, $F(X)$ 为残差映射。

其中, $F(X)$ 与 X 相加时,格式必须相同,运用批量归一化(也称为批量规范、批标准化)是一种用于通过重新居中和重新缩放对层的输入进行归一化来使人工神经网络更快、更稳定的方法。

ResNet 网络具有以下优点:

1) 将靠前若干层的某一层数据输出直接跳过多层引入到后边的输出部分。意味着后面的特征层的内容会有一部分由前面的某一层贡献。

2) ResNet 提高了深度加深时的效率和准确度。网络层数与处理速度、精度都具有很大关系,层数越大的计算量越大,精度越高,被广泛应用的是 ResNet34、ResNet50 和 ResNet101。

而 ResNet101 层数过多,过度在意细节而忽略整体,因此 ResNet50 的处理速度和精度都是最合适的。

2.2. ResNet50

ResNet50 使用 3×3 卷积,批标准化,将分散的数据统一。若 $F(X)$ 的数据维数变化(如 stride > 1 降维),则 X 也需要进行相应的变化(如对 X 做 1×1 的卷积)。求 $F(X)$ 残差的卷积均使用 3×3 conv,下采样维数降了一半。由于恒等映射 X 的存在,反向传播时,梯度可以从深层直接给到浅层,避免了梯度消失与爆炸。

ResNet50 网络结构主要包括以下几部分:

1) Identity Block: 输入和输出的维度相同,可以串联多个,可直接相加,维度不变(input shape = output shape)。

2) Conv Block: 输入和输出的维度是不一样的, 不能连续串联, 它的作用本来就是为了改变特征向量的维度不能连续串联, 以让维度相等然后相加改变维度(input shape != output shape)。

3) 残差: 观测值与估计值之间的差。需求解的映射为: $H(x)$ 求解网络的残差映射函数, 也就是 $F(x)$, 其中 $F(x) = H(x) - x$ 。这里 $H(x)$ 就是观测值, x 就是估计值(也就是上一层 ResNet 输出的特征映射)。

细化以上各部分 Resnet50 网络中包含 49 个卷积层、一个全连接层。如图 1 所示, Resnet50 网络结构可以分成七个部分, 第一部分不包含残差块, 主要对输入进行卷积、正则化、激活函数、最大池化的计算。第二、三、四、五部分结构都包含了残差块, 图 1 中的绿色图块不会改变残差块的尺寸, 只用于改变残差块的维度。在 Resnet50 网络结构中, 残差块都有三层卷积, 那网络总共有 $1 + 3 \times (3 + 4 + 6 + 3) = 49$ 个卷积层, 加上最后的全连接层总共是 50 层, 这也是 Resnet50 名称的由来。ResNet 网络[2]的输入为 $224 \times 224 \times 3$, 经过前五部分的卷积计算, 输出为 $7 \times 7 \times 2048$, 池化层会将其转化成成一个特征向量, 最后分类器会对这个特征向量进行计算并输出类别概率。

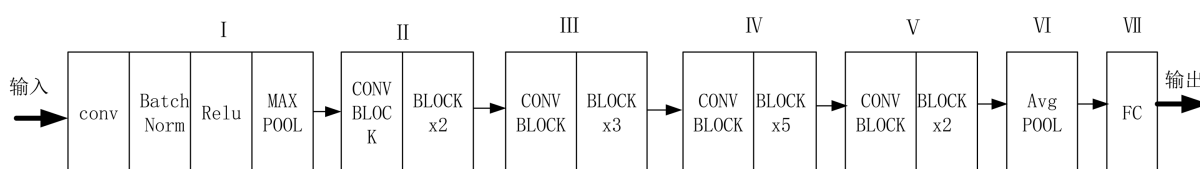


Figure 1. Structure diagram of ResNet50

图 1. ResNet50 结构图

3. ResNet 应用现状

随着近几年科技的发展中, “互联网+”模式的应用在各种研究方面已经成为现代科技的热点, 卷积神经网络在图像的处理中有很大优势[3]。

目前 ResNet 网络模型由于模型本身提高了深度的同时加深了模型本身的效率和准确度的优秀性质在检测、分割、识别等领域里得到了广泛的应用, 人工智能图像处理在医学影像和医疗智能决策中发挥着极其重要的作用[4][5], 互联网医疗的快速发展促进了科研和临床实践的有效结合, 使得医疗产业的发展通过互联网技术的应用而提升了准确率和速度被广泛应用于医学图像领域中。目前 ResNet 模型在良性骨折[6]、COVID-19 检测[7]、良恶性肿瘤、肺癌、乳腺癌、皮肤疾病、心脑血管疾病等重大疾病的临床辅助诊断方面取得良好效果。在医学图像领域, ResNet 网络模型结构对医学图像分类识别、分割、融合、检测、合成、超分辨率等方面有着突出的效果, 有助于医生做出正确的医学判断和更好的患者预后, 为临床计算机辅助诊断提供参考。期望为临床医生提供帮助, 为医疗行业提供更大的便利。

不只是医学领域, ResNet 模型也因其本身优秀的图像处理性能, 现如今被广泛应用于很多行业。ResNet 除了应用于医学中鉴别恶性肿瘤, 还可在房屋坍塌等灾害复杂环境下提供有效的帮助。在这种复杂多变的环境, 和紧急的救援时间中, 救援机器人需自动快速, 精准的对受伤人员进行动作识别与状态分析, 因此运用 ResNet 网络模型对复杂图像中的人体关节点和肢体进行检测有着重要的意义。

在农业生产中, ResNet 模型也发挥着重要作用, 对于农作物的病虫害检测判断分类有着显著效果。例如渔业养殖方面, 对于水下目标精准识别并且进行指导养殖生产、辅助养殖决策具有十分重要的作用, 而目标识别精度和运行效率是影响识别技术深入应用的关键问题。因此依靠 ResNet 模型的检测, 分割, 识别能力解决问题, 以此提高农业生产效率。

在考古工作研究中, 运用 ResNet 网络有效地提升文物识别准确率。例如古文本因年代久远而导致的图像退化等特点, 进而导致识别难度大, 部分字符的标注数据不足造成基于深度学习的模型识别准确率

不高,泛化能力差。针对上述问题,基于 ResNet 的识别方法,快速有效分类提高准确率。

由于 ResNet 其本身优异的更快的速度和更准确的精度性质,ResNet 网络更多的被应用于我们的日常生活中,如垃圾分类[8],车辆识别,城市交通路面状态识别,电路故障分类研究,人脸识别[9]等,进行快速准确的识别图像进行分类。在人脸识别技术中,如高铁进站时的人脸识别,Res Net 在特征提取过程发挥了重要作用,因此广泛应用于目标检测等计算机视觉任务。

虽然残差神经网络在诸多的图像处理领域取得了突破性的进展,并且现如今已经发挥了实质性的作用,但是仍然面临离不开大规模的高质量精确标注数据集的支持这一问题。ResNet 网络学习需要大量的图像数据,但绝大多数图像数据缺乏有效的数据标注且数量有限,而 ResNet 模型对于稀疏标注或未标注数据进行学习,ResNet 模型的完成过程相对简单且模型的泛化能力较弱。因此优质的数据源和好的数据标注非常重要,如何实现在标注数据集不完善的情况下进行弱监督学习,在 ResNet 图像处理未来的研究中具有重要意义。

4. 总结

本文通过对 ResNet 模型的分析,尤其是对 ResNet50 的分析可知,ResNet 解决了实际应用过程中深度神经网络学习的时候的梯度消失或者爆炸问题,使得在检测、分割、识别等应用需求领域里得到了广泛的应用。

基金项目

国家自然科学基金青年基金“低计算量高精度半结构化环境视觉/惯性紧组合的 AGV 定位导航方法研究”(61803035)。

参考文献

- [1] Hossain, B., Hasan Sazzad Iqbal, S.M., Islam, M., et al. (2022) Transfer Learning with Fine-Tuned Deep CNN ResNet50 Model for Classifying COVID-19 from Chest X-Ray Images. *Informatics in Medicine Unlocked*, **30**, Article ID: 100916. <https://doi.org/10.1016/j.imu.2022.100916>
- [2] 朱伟鹏, 操超超, 刘金华. 基于 ResNet50 模型的乳液泵缺陷分类[J]. 工业控制计算机, 2022, 35(7): 59-61.
- [3] Lecun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. *Nature*, **512**, 436-444. <https://doi.org/10.1038/nature14539>
- [4] 陈弘扬, 高敬阳, 赵地, 汪红志, 宋红, 苏庆华. 深度学习与生物医学图像分析 2020 年综述[J]. 中国图像图形学报, 2021, 26(3): 475-486.
- [5] 周涛, 刘赞璨, 陆惠玲, 叶鑫宇, 常晓玉. ResNet 及其在医学图像处理领域的应用: 研究进展与挑战[J]. 电子与信息学报, 2022, 44(1): 149-167. <https://doi.org/10.11999/JEIT210914>
- [6] 刘珂, 王奇政, 陈永晔, 等. 基于 ResNet50 深度学习模型鉴别脊柱良恶性骨折[J]. 临床放射学杂志, 2021, 40(12): 2350-2355.
- [7] Bolhassani, M. (2021) Transfer Learning Approach to Classify the X-Ray Image that Corresponds to Corona Disease Using ResNet50 Pretrained by ChexNet. <https://arxiv.org/abs/2105.08382>
- [8] 张丽艳, 赵艺璇, 牟钰鹏, 等. 基于 ResNet50 的列车常见垃圾自动分类算法[J]. 大连交通大学学报, 2021, 42(4): 101-105.
- [9] 谢鑫. 基于 ResNet50 的人脸识别模型[J]. 科技资讯, 2020, 18(35): 9-11.