

ResNet和EfficientNet遥感图像场景分类研究

江文

无锡商业职业技术学院, 江苏 无锡

收稿日期: 2022年4月13日; 录用日期: 2022年5月13日; 发布日期: 2022年5月20日

摘要

针对遥感图像分类问题, 首先采用了ResNet模型和EfficientNet模型进行训练, 其中, 前者采用18层网络结构, 并在测试集得到了59.8%的准确度; 后者采用了EfficientNet-B0和EfficientNet-B1模型, 最高分类精度为92.6%。在此基础上, 使用Ranger优化器替换了SGD方法, 虽然测试集精度与原方法近似, 但使用Ranger优化器训练EfficientNet模型具有更强的稳定性和更快的训练速度。最后, 对测试的结果进行了分析, 列举出每一类场景的分类精度, 分析并总结了改进方法。

关键词

遥感图像分类, ResNet模型, EfficientNet模型, Ranger优化器

Study on ResNet and EfficientNet Remote Sensing Image Scene Classification

Wen Jiang

Wuxi Vocational Institute of Commerce, Wuxi Jiangsu

Received: Apr. 13th, 2022; accepted: May 13th, 2022; published: May 20th, 2022

Abstract

The ResNet and EfficientNet are used to classify the remote sensing image. Among ResNet and EfficientNet, the former uses an 18-layer network structure and has a classification accuracy of 59.8% in the test set; the latter uses EfficientNet-B0 and EfficientNet-B1, with the highest classification accuracy of 92.6%. After that, we replace the SGD optimizer with Ranger optimizer. Although the new classification accuracy is similar to the original one, using the Ranger optimizer to train the EfficientNet model has stronger stability and faster training speed. Finally, we analyzed the test results, listed the classification accuracy of each type of scene, analyzed and summarized the improvement methods.

Keywords

Remote Sensing Image Classification, ResNet Model, EfficientNet Model, Ranger Optimizer

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 问题描述

遥感图像已被广泛应用于土地利用分类、地理图像检索和遥感场景分类等领域。遥感图像场景分类是指从多幅遥感图像中区分出具有相似场景特征的图像[1]，并对这些图像进行分类，为每一幅遥感图像赋予场景类别标签[2]，可以为城市规划和环境监测的后续工作提供一定的决策依据。遥感场景分类根据遥感图像的语义信息将遥感图像分为森林、城市、河流等多种类别[3]。遥感图像一般具有较高的空间分辨率[4]、高度复杂的几何结构和空间格局[5]。然而，不同类别的遥感图像之间有很高的相似性[6]，而同一场景类别中的图像可能有很大的差异。这就给遥感场景的分类带来了困难。

在本文中，针对如图 1 所示的包含典型的遥感图像场景对象，本文使用两种深度学习网络对指定的遥感图像进行场景分类，给出其预测结果。



Figure 1. Example of remote sensing image classification data set (part)

图 1. 遥感图像分类数据集示例(部分)

1.2. 研究思路

目前人工特征提取方法获取的分类结果在精度上往往不能令人满意。随着深度学习的发展，深度学习已经应用于许多研究领域[7] [8]。卷积神经网络作为最流行的深度学习技术之一[9]，自然地被引入到遥感场景分类中。一般来说，神经网络可以用大量的训练数据进行很好的训练[10]，而用少量的训练数据往往会造成过拟合和泛化能力差的现象。为了避免或缓解这种情况，一方面可以使用数据增强技术来扩展训练数据，另一方面，设计专门的网络也可以使训练更加高效。

一些经典的神经网络结构已经被应用于遥感场景分类，如 CaffeNet、GoogleNet 和 VGGNet。尽管基于这些网络的许多方法显著提高了分类精度，但它们仍然存在一些瓶颈。例如，用少量数据训练的网络依然有大量参数，这通常会导致卷积神经网络中训练结果的过拟合。这些网络一般比较浅，由于遥感图像的复杂性，为了更好地对遥感场景进行分类，分类网络需要很好地提取图像的高层语义特征。值得注意的是，更深层次的网络往往可以提取更抽象的语义信息。

如图 2 所示，本文首先基于 ResNet 模型，采用 Adam 优化器训练数据集图片，最终得到的测试集准确率为 59.8%；接着，本文使用了 EfficientNet 模型训练，最终得到的训练集和测试集准确率分别为 94.7% 和 92.6%。而后本文使用 Ranger 优化器训练 ResNet 模型与 EfficientNet 模型。其中，使用 Ranger 优化器训练的 EfficientNet-B1 模型，在测试集和训练集分别达到了 99.2% 和 92.4% 的准确率。

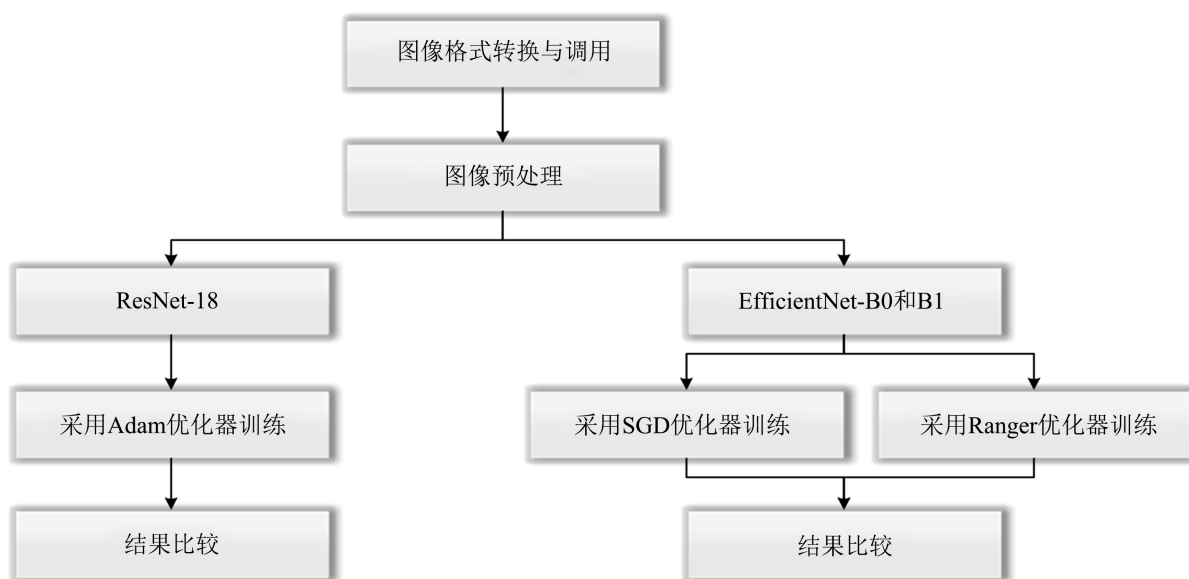


Figure 2. Flow chart of algorithm idea

图 2. 算法思路流程图

在本文的最后，本文对每个场景分类精度进行了分析，对算法进行了总结和评价，并针对遥感图像分类问题以后的研究方向进行了展望。

2. 算法原理及流程

2.1. 基于残差神经网络的遥感图像分类

1) 原理介绍

随着神经网络的深度的加深，网络的表达能力随之加强，但网络深度会带来梯度消失的问题，且随

着深度加深，梯度消失问题也越来越严峻。为了解决网络深度增加带来的梯度消失问题，本文采用了深度残差神经网络。

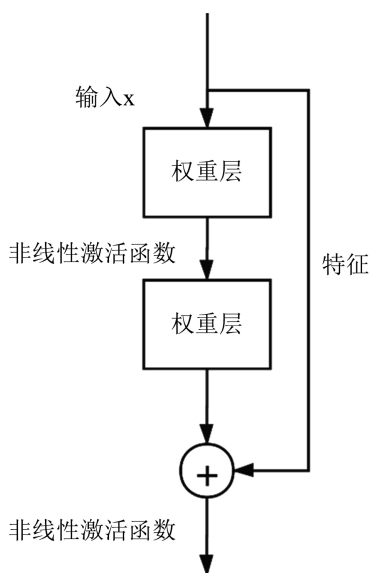


Figure 3. Residual module

图 3. 残差模块

如果直接去增加网络深度，这使得深度学习系统难以被优化[11]。假设现在有一个浅层网络，那么也应该存在一个深层网络。这个网络是以浅层网络为基础堆叠了多个 $x \rightarrow x$ (恒等映射)的映射组成，那么该深层神经网络的性能应该优于浅层网络的性能。然而，实验证明，我们无法找到这样的一个深层网络。这种现象说明，直接用叠加深度这样的方法来拟合这样的 $x \rightarrow x$ 恒等映射是非常困难的。在浅层网络达到饱和后，可以在其后面添加一个恒等映射层，这样即避免了增加网络的深度也能增强模型的性能，并且随着深度的增加，系统误差不会变大。

文献[12]中引入了深度残差学习框架来解决网络退化问题。假设原始神经网络的一个残差单元要学习的目标映射为 $H(x)$ ，这个目标映射可能很难学习。残差神经网络让残差单元不直接学习目标映射，而是学习一个残差 $F(x) = H(x) - x$ 。这样原始的映射变成了 $F(x) + x$ 。原始残差单元可以看作是由两部分构成，一个线性的直接映射 $x \rightarrow x$ 和一个非线性映射 $F(x)$ 。特别地，如果 $x \rightarrow x$ 是最优的学习策略，那么相当于把非线性映射 $F(x)$ 的权重参数设置为 0。恒等映射使得非线性映射 $F(x)$ 学习线性的 $x \rightarrow x$ 映射变得容易很多。

残差单元是深度残差网络的基本组成单元，一般由卷积层(Conv)、批处理归一化层(Batchnorm)、非线性激活函数(ReLu)共同构成残差单元。图 3 给出了残差单元的示意图，若令第 l 个残差单元的输入为 x_l ，那么下一层的输出 x_{l+1} 为：

$$x_{l+1} = f(x_l + F(x_l, W_l))$$

其中， $F(x_l, W_l)$ 是残差函数， W_l 是该残差函数对应的权重参数； $f(*)$ 是非线性激活函数 ReLu。

使用残差网络进行图像分类的一般流程为：首先使用多个残差学习单元构成深度残差网络。然后给定输入的图像数据，这些数据将依次通过卷积层、非线性激活函数层、批处理归一化层；之后，处理的结果将进一步被送入到多个残差单元；最后经过批处理归一化层和多个全连接层就能得到分类结果。

2) 使用残差网络进行遥感图像分类的具体步骤

步骤一：图像预处理

在这一步骤，本文对图像进行了 100×100 的随机裁剪，一方面考虑到计算机性能因素，即过高像素的图片会极大地延缓运行时间，不利于训练的正常进行；另一方面，随机裁剪相当于建立每个因子特征与相应类别的权重关系，减弱背景(或噪音)因子的权重，且使模型面对缺失值不敏感，也就可以产生更好的学习效果，增加模型稳定性。

步骤二：构建 ResNet 模型

考虑到显卡计算能力，本文使用 ResNet-18 模型进行遥感图像分类，

步骤三：使用 Adam 优化器对网络参数进行训练

Adam 优化器是训练模块，通过训练输出损失值误差的反向传播，结合 Adam 算法，训练网络。这里将最大迭代步骤设为 60 次，得到了最终的训练结果。

3) 结果分析

通过训练数据集中的 177,408 张图片，本文得到如下的损失值和准确度曲线：

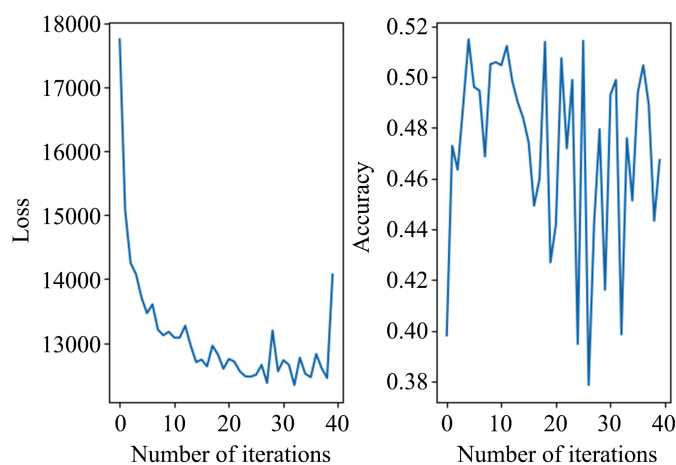


Figure 4. ResNet18 test set loss and accuracy (step size: 0.01)

图 4. ResNet18 测试集损失值及准确度(步长为 0.01)

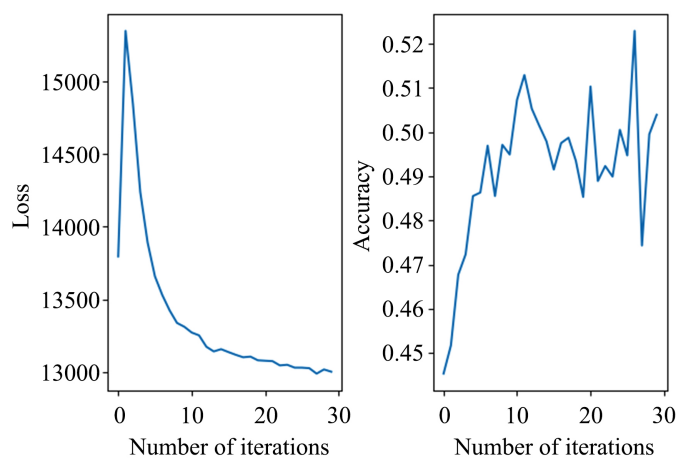


Figure 5. ResNet18 test set loss and accuracy (step size: 0.001)

图 5. ResNet18 测试集损失值及准确度(步长为 0.001)

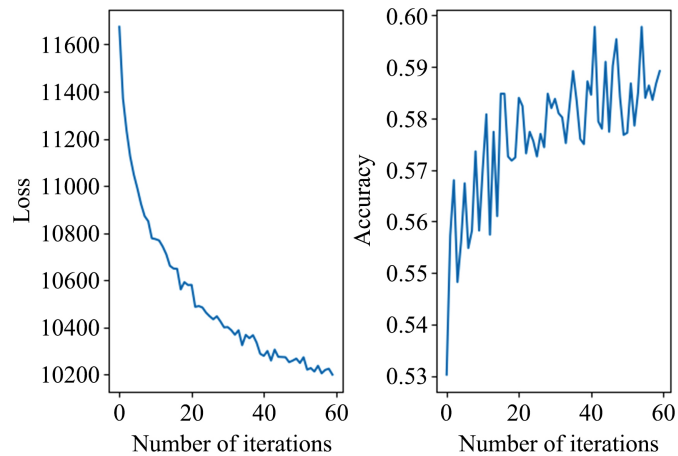


Figure 6. ResNet18 test set loss and accuracy (step size: 0.0001)

图 6. ResNet18 测试集损失值及准确度(步长为 0.0001)

通过图 4~6 我们可以发现,初始步长的设定对最终训练结果有较大的影响,并且当迭代次数较多时,训练的结果会出现过拟合的情况,即测试集的准确率逐渐下降,会低于训练集的准确率,准确度最高达到 59.8%。

2.2. 基于 EfficientNet 的遥感图像分类

1) 原理介绍

基于 ResNet 的图像分类主要通过加深网络层数的方法使函数拟合的结果更加精确,在过去大多数深度学习方法中,通常只放缩深度、宽度和图像大小三个维度中的一个。尽管我们可以较为轻松地做到缩放其中的任意两个维度甚至三个维度,但是任意缩放需要手动调整网络参数,并且经常只会产生次优的精度和效率。

文献[13]中系统地研究了模型缩放问题,主要包括如何调整网络深度、宽度、分辨率以获得更好的性能。基于这一问题,文献[13]提出了一种新的缩放方法,该方法使用简单高效的复合系数统一缩放深度、宽度、分辨率三个维度的大小。

将卷积层 i 定义为一个函数: $Y_i = \mathcal{F}_i(X_i)$, 其中 \mathcal{F}_i 是算子, Y_i 是一个输出张量, X_i 是一个输入张量且形状为 $\langle H_i, W_i, C_i \rangle$ 。

一个卷积神经网络 \mathcal{N} 可以表示为:

$$\mathcal{N} = \mathcal{F}_k \circ \dots \circ \mathcal{F}_2 \circ \mathcal{F}_1(X_1) = \bigcirc_{j=1 \dots k} \mathcal{F}_j(X_1)$$

在实际应用中,卷积神经网络一般可以被划分为多个具有相同体系结构的阶段,因此,一个卷积神经网络可以被定义为:

$$\mathcal{N} = \bigcirc_{i=1 \dots s} \mathcal{F}_i^{L_i}(X_{H_i, W_i, C_i})$$

其中 $\mathcal{F}_i^{L_i}$ 表示在第 i 个阶段将卷积层 \mathcal{F}_i 重复 L_i 次。为了更好地描述网络性能的优劣,文献中将该模型转换为一个带约束条件的优化问题:

$$\begin{aligned} & \max_{d, w, r} \text{Accuracy}(\mathcal{N}(d, w, r)) \\ & \text{s.t. } \mathcal{N}(d, w, r) = \bigcirc_{i=1 \dots s} \hat{\mathcal{F}}_i^{d \cdot \hat{L}_i}(X_{r \cdot \hat{H}_i, r \cdot \hat{W}_i, w \cdot \hat{C}_i}) \\ & \text{Memory}(\mathcal{N}) \leq \text{target_memory} \end{aligned}$$

$$FLOPS(\mathcal{N}) \leq target_flops$$

其中 d, w, r 是网络扩张系数的深度、宽度和分辨率; $\hat{\mathcal{F}}_i, \hat{\mathcal{L}}_i, \hat{\mathcal{H}}_i, \hat{\mathcal{W}}_i, \hat{\mathcal{C}}_i$ 是基础网络中预设定的参数。

深度: 在卷积神经网络的应用中, 拓展网络层数一直是最常用的方法。更深层次的卷积神经网络可以捕获更丰富、更复杂的特征。然而, 由于消失梯度问题, 更深的网络也更难训练。虽然一些技术, 如 dropout 和批量归一化等方法缓解了训练中存在的问题, 但非常深的网络的精度增益增加并不明显, 例如 ResNet-1000 具有与 ResNet-101 相似的精度, 即使它有更多的层。

宽度: 网络宽度的缩放通常应用于小规模网络模型, 宽度更大的网络往往能够捕获更加精细的特征, 并且更容易训练。然而, 极宽但很浅的网络往往难以捕捉更高层次的特征。在网络深度和分辨率确定的情况下, 拓展宽度维度很快会达到饱和从而无法继续提高精度。

分辨率: 使用更高分辨率的输入图像, 卷积网络可以获取更加精细的图案特征。但和其它两个维度一样, 分辨率过高对精度增益效果不明显, 容易达到饱和。

文献[13]中提出了一种新的复合缩放方法, 它使用复合系数 ϕ 统一缩放网络的宽度、深度和分辨率:

$$\begin{aligned} \text{depth} : d &= \alpha^\phi \\ \text{width} : w &= \beta^\phi \\ \text{resolution} : r &= \gamma^\phi \\ \text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\ \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned}$$

其中 α, β, γ 是通过很小范围搜索的方法决定的常量。 ϕ 是由用户指定的系数(控制用于模型缩放的额外资源的数量), α, β, γ 则能够决定额外资源到网络的宽度、深度和分辨率上的分配。在实际的深度学习的应用中, EfficientNet 充分体现了其高效性和准确性。

从基线网络 EfficientNet-B0 开始, 本文应用复合缩放方法, 分两步将其缩放:

步骤一: 本文首先固定 $\phi = 1$, 假设可获得的资源增加两倍, 对 α, β, γ 进行小范围搜索。特别地, 本文发现在 $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$ 的约束下, EfficientNet-B0 的最佳值为 $\alpha = 1.2$, $\beta = 1.1$, $\gamma = 1.15$ 。

步骤二: 然后, 本文将 α, β, γ 固定为常数, 并放大具有不同 ϕ 的基线网络, 从而获得到 EfficientNet-B1 到 B7。

2) 迁移学习

为了能加快模型的学习速度, 本次实验使用迁移学习, 将已经在 ImageNet 数据集上训练好的 EfficientNet 参数作为特征提取模型的初始参数, 从而继承模型从 ImageNet 数据集上学习到的特征提取能力, 再在此基础上进行微调。ImageNet 数据集包含了很多的分类类型, 数据量十分庞大, 可以赋予模型极强的特征提取能力以及泛化能力。该方法不仅可以节省大量的训练时间和计算资源, 更可以一定程度上削弱过拟合现象, 使模型拥有比随机初始化参数更好的表现。事实上, 从实验结果表明, 使用迁移学习后网络模型在训练前期就能得到相对较好的结果。

3) 算法与代码分析

步骤一: 本文首先对图片进行了缩放与剪裁, 将图片像素转换为 224×224 , 再将图片作仿射变换, 增强数据特征, 最后将图片标准化, 加速网络模型的收敛速率。

步骤二: 调用 EfficientNet 网络模型库文件, 并导入已有的 B0~B7 模型参数作为训练的预设值。

步骤三: 使用 pytorch 中的 SGD 优化器训练网络参数, 设置好超参数。这一部分的训练过程与 ResNet 训练类似, 区别在于采用了 SGD 优化器。

4) 结果分析

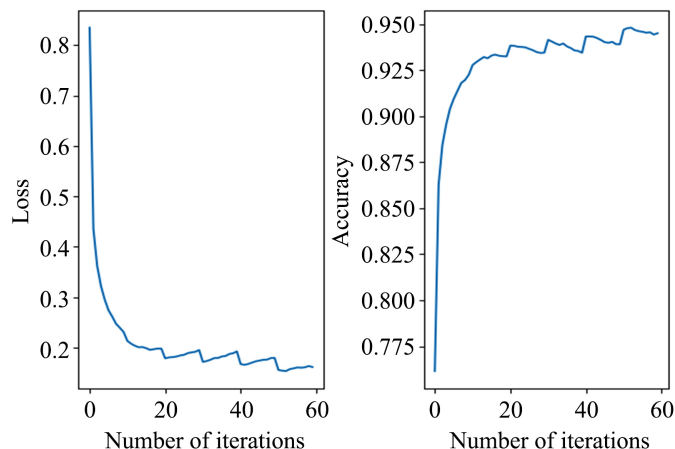


Figure 7. EfficientNet-B0 training set loss and accuracy curve (step size: 0.001)

图 7. EfficientNet-B0 训练集损失值及准确度曲线(步长为 0.001)

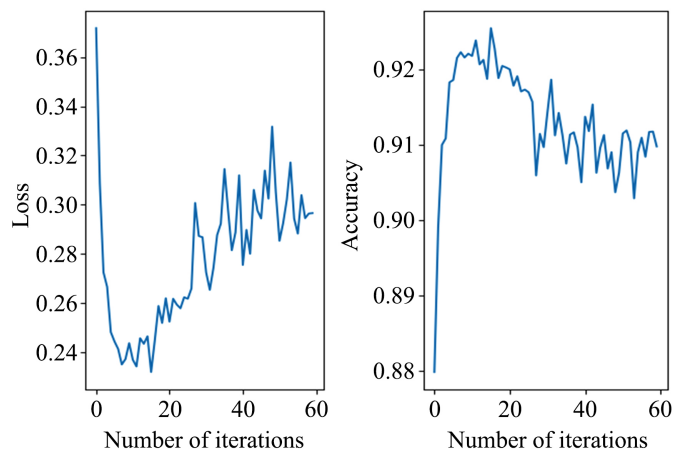


Figure 8. EfficientNet-B0 test set loss and accuracy curve (step size: 0.001)

图 8. EfficientNet-B0 测试集损失值及准确度曲线(步长为 0.001)

从图 7、图 8 可知,随着迭代步数的增加,训练时的准确度一直呈上升趋势,但测试集的准确度在 20 步之后开始逐渐下降,主要原因在于过拟合。因此,在 EfficientNet-B0 网络下,测试集的准确率最高为 92.6%。

3. 基于改进 EfficientNet 的遥感图像场景分类

3.1. Ranger 优化器

在以上两个网络的训练过程中,本文最初分别采用了 Adam 算法和 SGD 算法作为网络优化器。SGD 算法由于更新频繁,容易造成代价函数的震荡,且有可能只得到局部最优解;Adam 算法需要一个预热阶段,否则在训练刚刚启动的时候就很容易陷入不好的、可能有问题的局部最优。在训练刚开始时有非常大的变动,可以说是优化器还没见到足够多的数据,没法做出准确的自适应动量选择。因此,在原有优化器的基础上,本文使用文献[14]和[15]中提出的优化算法,采用了 Ranger 优化器对 EfficientNet 进行训练。

文献[14]中提出了 LookAhead 优化器,该优化器采用了一种全新的设计:维持两套权重,并在两者

之间进行内插,可以说是,它允许更快的那一组权重向前看(也就是探索),同时更慢的那一组权重可以留在后面,带来更好的长期稳定性。这种做法带来的效果就是降低了训练过程中的变差,以及大大降低了对次优的超参数的敏感性(从而减少了大量尝试超参数调节的需要);同时,它在许多种不同的深度学习任务中都可以达到更快的收敛速度。可以说这是一项飞跃式的进步。

文献[15]中提出了 RAdam 算法,该算法采用了一个整流器函数,它可以基于实际遇到的变差计算出一个预热启发值。然后这个整流器可以动态地对自适应动量进行开、关或者加阻尼,避免它以全速运动,直到来自数据的变差开始稳定为止。这样的做法就避免了用人工方法执行一个预热阶段,训练过程也就自动稳定下来了。

本文结合了以上两种优化器的优势,采用了 Ranger 优化器。该优化器的核心思想是将 LookAhead 中的 Adam 优化器替换成 RAdam。在训练的初始阶段, RAdam 能为优化器提供最棒的基础值。借助一个动态整流器, RAdam 可以根据变差大小来调整 Adam 优化器中的自适应动量,并且可以提供一个高效的自动预热过程; LookAhead 减少了超参数调节的工作量,同时许多不同的深度学习任务中都有更快的收敛速度、最小的计算开销。

3.2. 结果分析

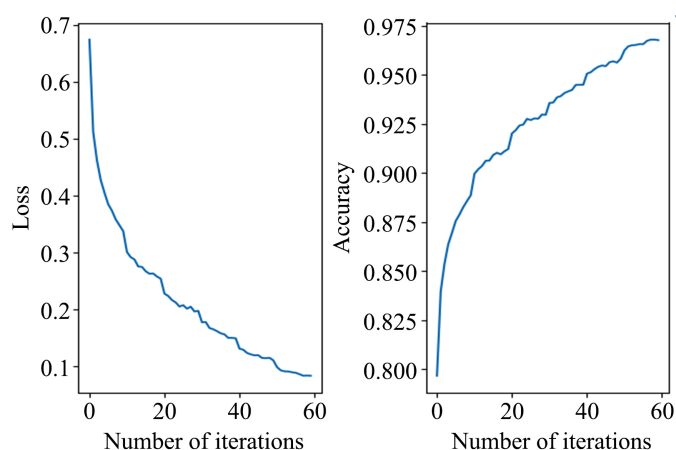


Figure 9. EfficientNet-B0 training set loss and accuracy curve based on Ranger optimizer (step size: 0.01)

图 9. 基于 Ranger 优化器的 EfficientNet-B0 训练集损失值及准确度曲线(步长为 0.01)

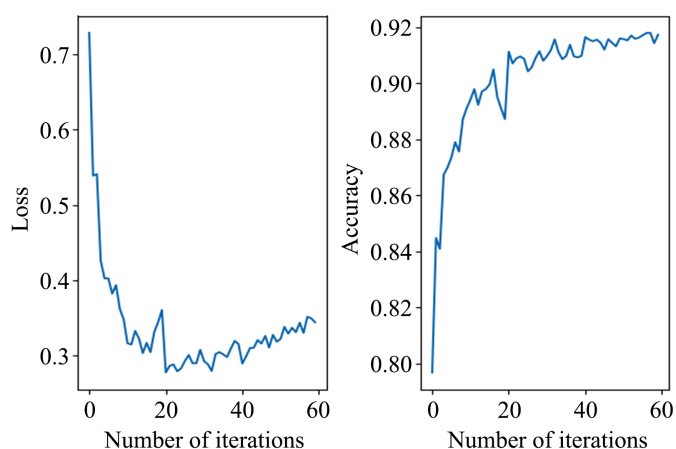


Figure 10. EfficientNet-B0 test set loss and accuracy curve based on Ranger optimizer (step size: 0.01)

图 10. 基于 Ranger 优化器的 EfficientNet-B0 测试集损失值及准确度曲线(步长为 0.01)

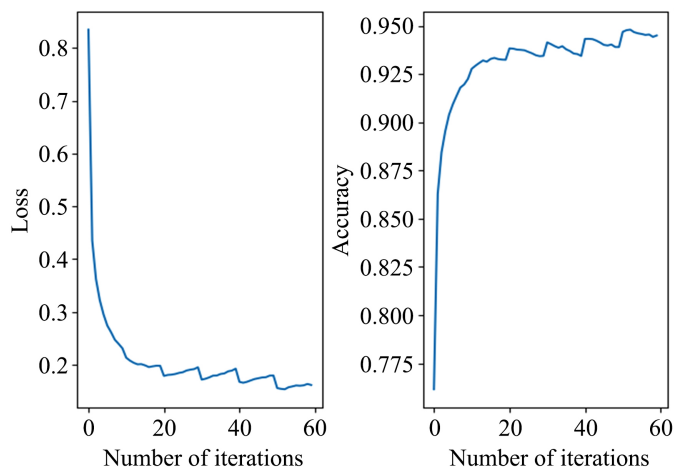


Figure 11. EfficientNet-B0 training set loss and accuracy curve based on Ranger optimizer (step size: 0.001)
图 11. 基于 Ranger 优化器的 EfficientNet-B0 训练集损失值及准确度曲线(步长为 0.001)

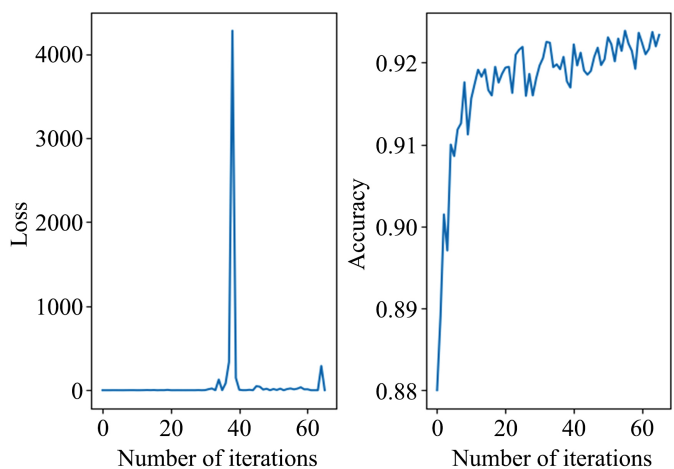


Figure 12. EfficientNet-B0 test set loss and accuracy curve based on Ranger optimizer (step size: 0.001)
图 12. 基于 Ranger 优化器的 EfficientNet-B0 测试集损失值及准确度曲线(步长为 0.001)

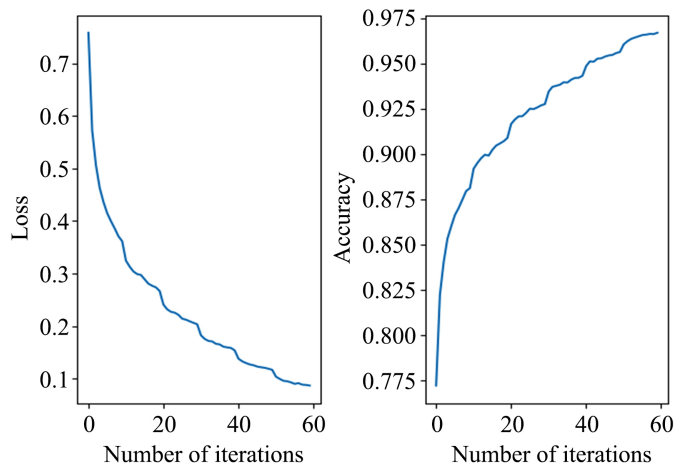


Figure 13. EfficientNet-B1 training set loss and accuracy curve based on Ranger optimizer (step size: 0.01)
图 13. 基于 Ranger 优化器的 EfficientNet-B1 训练集损失值及准确度曲线(步长为 0.01)

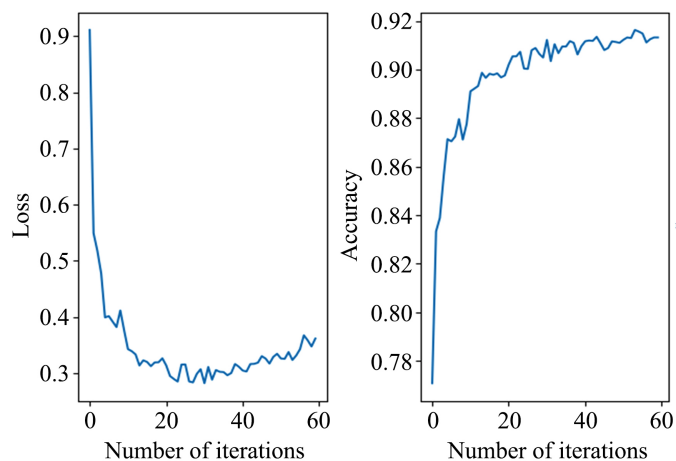


Figure 14. EfficientNet-B1 test set loss and accuracy curve based on Ranger optimizer (step size: 0.01)

图 14. 基于 Ranger 优化器的 EfficientNet-B1 测试集损失值及准确度曲线(步长为 0.01)

从图 9~14 可知, 使用 Ranger 优化器后, 尽管训练集准确度很高, 但过度学习了训练集的不必要细节, 最后测试集表现不分伯仲, 但后者表现的稳定性强, 训练速度更快。

4 结果分析

根据基于 Ranger 优化器的 EfficientNet-B1 模型分类结果, 本文对遥感图像分类结果中的每个场景的精确度进行了量化分析。由图 15 可知, 部分场景的分类精度较低, 包括立交桥(36.7%), 路边停车区(66.1%), 停车场(76.8%), 桥梁(78.5%)等六个场景精度在 80% 以下。

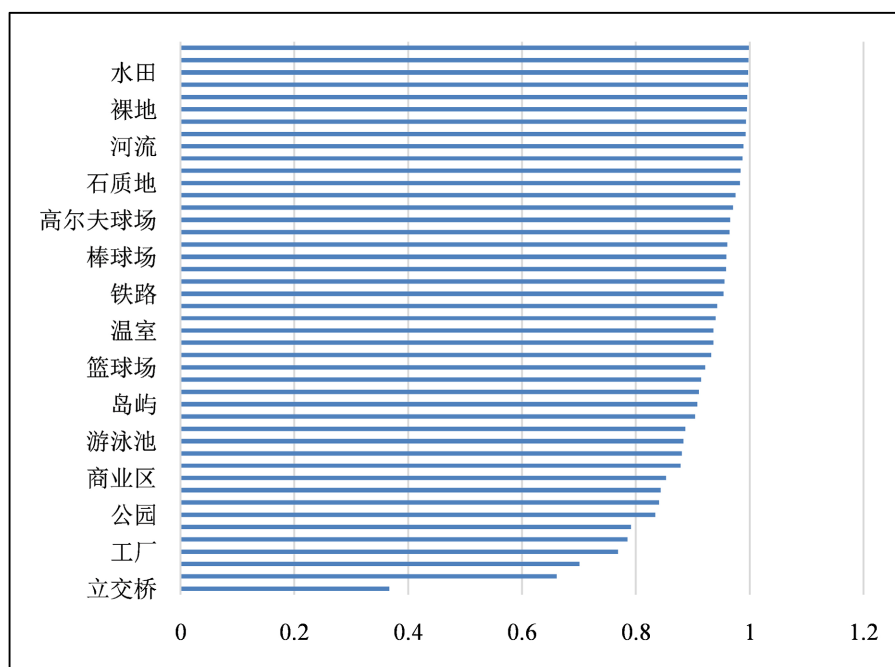


Figure 15. Accuracy of various types of remote sensing image classification (part)

图 15. 遥感图像分类各类图像精度(部分)

通过对这些分类精度较低的场景进行对比分析, 本文发现其原因有两种: 第一种因为部分场景的相似度较高, 由图 16 所示, 立交桥和桥梁的特征相似点过多, 不易区分。因此, 未来可以使用常识推理与属性识别的方法, 加强对相似特征的分, 以期得到更好的效果; 第二种原因可以总结为同类图片的特征差异过大。如图 17 所示, 像停车场这样的同类图片的特征差别比较大, 针对这种情况, 还需要增强模型的泛化能力。



Figure 16. Bridge (left) and overpass (right)
图 16. 桥梁(左)与立交桥(右)



Figure 17. Parking lot (same scene classification)
图 17. 停车场(同一场景分类)

5. 总结与展望

5.1. 方法总结与评价

针对遥感图像分类, 本文对残差神经网络与 EfficientNet 两种模型进行训练, 在分别使用 Adam 与 SGD 优化器优化的情况下, 其测试准确度分别为 66% 与 89%。然后, 本文提出使用 Ranger 优化器训练 EfficientNet 模型, 使得测试结果达到 91%。实验表明, 相比于残差神经网络, EfficientNet 在遥感图像分

类有较好的优势，并且使用 Ranger 优化器能让深度学习网络的训练过程更快速，更平稳。

5.2. 展望

除了计算机性能方面的约束，在以下几个方面依然存在很大的改进或提升的空间：

1) 优化器参数的调整。本文采用的 Ranger 优化器中依然存在不少的超参数，需要人为地提前设定，例如步长参数 α ，参数的一阶和二阶动量的衰减率 β_0 和 β_1 等，这些超参数都会对训练的结果产生较大的影响。

2) 在文献[16]中提出了一些针对梯度消失或梯度爆炸的问题的解决方法，也应引入到遥感图像分类的训练过程中。

3) 可以考虑常识推理与属性识别，即机器能够从给定的例子中发现规律与规则，增强具有相似特征类间的图片以及有较大差异特征的类内图片的识别。

基金项目

国家自然科学基金项目(61873111)，教育部商业智能应用协同创新中心、江苏省商业智能应用工程技术开发中心阶段研究成果。

参考文献

- [1] 姬腾飞. 基于卷积神经网络的遥感图像场景分类研究[D]: [硕士学位论文]. 开封: 河南大学, 2019.
- [2] 刘源. 基于卷积神经网络的卫星图像场景分类技术研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2020.
- [3] 钱晓亮, 李佳, 程焱, 等. 特征提取策略对高分辨率遥感图像场景分类性能影响的评估[J]. 遥感学报, 2018, 22(5): 758-776.
- [4] 朱圣, 邓继忠, 张亚莉, 等. 基于无人机低空遥感的水稻田间杂草分布图研究[J]. 华南农业大学学报, 2020, 41(6): 67-74.
- [5] 鞠洪润, 左丽君, 张增祥, 等. 中国土地利用空间格局刻画方法研究[J]. 地理学报, 2020, 75(1): 143-159.
- [6] 彭晏飞, 宋晓男, 瞿玲玲, 等. 基于卷积神经网络和改进模糊 C 均值的遥感图像检索[J]. 激光与光电子学进展, 2018, 55(9): 173-182.
- [7] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
- [8] 奚雪峰, 周国栋. 面向自然语言处理的深度学习研究[J]. 自动化学报, 2016, 42(10): 1445-1465.
- [9] 余萍, 曹洁. 深度学习在故障诊断与预测中的应用[J]. 计算机工程与应用, 2020, 56(3): 1-18.
- [10] 李小薪, 梁荣华. 有遮挡人脸识别综述: 从子空间回归到深度学习[J]. 计算机学报, 2018, 41(1): 177-207.
- [11] 赵志成, 罗泽, 王鹏彦, 等. 基于深度残差网络图像分类算法研究综述[J]. 计算机系统应用, 2020, 29(1): 14-21.
- [12] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [13] Tan, M. and Le, Q.V. (2019) Efficient Net: Rethinking Model Scaling for Convolutional Neural Networks.
- [14] Zhang, M., Lucas, J., Ba, J., et al. (2019) Lookahead Optimizer: K Steps forward, 1 Step Back. *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, Vancouver, 8-14 December 2019, 9597-9608.
- [15] Liu, L., Jiang, H., He, P., et al. (2019) On the Variance of the Adaptive Learning Rate and Beyond.
- [16] Li, J., Fang, C. and Lin, Z. (2019) Lifted Proximal Operator Machines. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 4181-4188. <https://doi.org/10.1609/aaai.v33i01.33014181>