

基于残差网络的光学遥感图像场景分类算法

汪鹏^{**}, 刘瑞^{*}, 辛雪静, 刘沛东

河北工业大学人工智能与数据科学学院, 天津 300100

摘要 提出一种基于卷积神经网络中残差网络的遥感图像场景分类方法。本文方法在原网络模型中嵌入了跳跃连接和协方差池化两个模块,用于连接多分辨率特征映射和融合不同层次的多分辨率特征信息,并在 3 个公开的经典遥感数据集上进行了实验。结果证明,本文方法不仅可以将残差网络中不同层次的多分辨率特征信息融合在一起,还可以利用高阶信息来实现更具代表性的特征学习。与已有的分类方法相比,本文方法在场景分类问题上拥有更高的分类精度。

关键词 图像处理; 遥感图像; 卷积神经网络; 场景分类; 残差网络

中图分类号 TP753

文献标志码 A

doi: 10.3788/LOP202158.0210001

Scene Classification of Optical Remote Sensing Images Based on Residual Networks

Wang Peng^{**}, Liu Rui^{*}, Xin Xuejing, Liu Peidong

School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300100, China

Abstract This paper proposes a method for the scene classification of optical remote sensing images based on the residual network of convolutional neural networks. In the proposed method, two modules, i. e., jump connection and covariance pooling, are embedded in the original network model to achieve multiresolution feature mapping and combine different levels of multiresolution feature information. Experiments are conducted on three open classical remote sensing datasets. Results show that the proposed method can fuse the multiresolution feature information of different levels in the residual network and use higher-order information to achieve more representative feature learning. The proposed method exhibits higher classification accuracy in the scene classification problem compared with the existing classification methods.

Key words image processing; remote sensing image; convolutional neural network; scene classification; residual network

OCIS codes 100.5010; 100.3008; 100.2000

1 引言

随着遥感技术的飞速发展,光学遥感图像的采集方法不断地更新进步,遥感图像场景的分类问题引起了人们广泛关注。遥感图像场景的识别和分类能够广泛应用于城市规划、土地利用、全球环境污染和监测以及军事领域的目标检测等方面,这些都是

关系到国计民生的重要领域。所以,遥感图像的场景分类研究具有重要的理论意义和实际价值。

传统的遥感图像场景分类方法,例如 BP 神经网络、模糊神经网络、自组织特征映射网络等都存在收敛速度慢、限于局部收敛等缺点,严重依赖于低级特征描述,并且不能完全捕获场景中包含的高级语义信息。近些年,以卷积神经网络(CNN)为代表的

收稿日期: 2020-06-16; 修回日期: 2020-06-27; 录用日期: 2020-07-01

基金项目: 天津市自然科学基金重点项目(19JCZDJC40000)、天津市科技计划项目(18YFCZCC00060,18ZXZNGX00100)

* E-mail: hebutwangpeng2019@163.com; ** E-mail: hebutliurui2020@163.com

深度学习方法在图像识别与分类领域取得了重大突破。例如,朱明明等^[1]利用深度学习对机场进行识别检测。姜晓佳等^[2]基于卷积神经网络搭建 Hair-Net 模型,对毛发显微图像进行了自动分类。马俊成等^[3]利用深度卷积神经网络(DCNN)进行飞机目标分类,解决了分类精度低的问题。王燕妮等^[4]利用卷积神经网络算法对多光谱图像进行处理,实现了对古壁画颜料的分类。Castelluccio 等^[5]将 GoogleNet 用于遥感图像场景分类,结果表明预训练的 CNN 模型大大优于传统的基于特征的手工方法。Cheng 等^[6]使用 BoVW 模型编码单个卷积层。Chaib 等^[7]将 CNN 模型的最后两个全连接层融合在一起表示图像信息。Li 等^[8]提出了一种多尺度改进的 Fisher 核编码方法,并将不同层次的特征图进行了组合。这些研究均取得了比较好的实验效果。

但是卷积神经网络也存在一些缺点,CNN 模型在提取图像特征时,这些特征信息会因层而异,其中浅层网络提取图像的局部低层特征,深层网络提取图像的全局高层语义特征。以往的研究中往往利用传输到最后全连接层的信息来代表输入图像,这样做会忽略各层之间的互补信息。其次,在一些常见

的公共遥感数据集中存在不同类别相似性较大和同一类别差异性较大的问题,而这些问题也是影响分类准确率的重要因素^[9]。

为了解决以上问题,本文提出了一种基于残差网络(ResNet)的遥感图像场景分类方法。本文的方法不同于仅仅更改全连接层的常规传输方法,为了独立捕获中间层特征,在 ResNet 网络模型中嵌入了跳跃连接和协方差池化两个模块,以连接多分辨率特征映射和融合不同层次的多分辨率特征信息。

2 基本原理

2.1 相关工作

在使用卷积神经网络来提取图像特征时,这些特征信息会因层而异,较浅的层反映比较具体的信息,较深的层反映比较抽象的信息,这些信息再通过全连接层进行整合,而在这个过程中可能会丢失各层的互补信息,从而导致分类精度降低。同时,在实际应用或者一些经典的公共数据集中,必不可少地会存在不同类别的图像具有很大的相似性或者同一类别的图像具有很大的差异性的问题,以在遥感图像领域极具挑战性的 NWPU-RESISC45 数据集为例,如图 1 和图 2 所示。

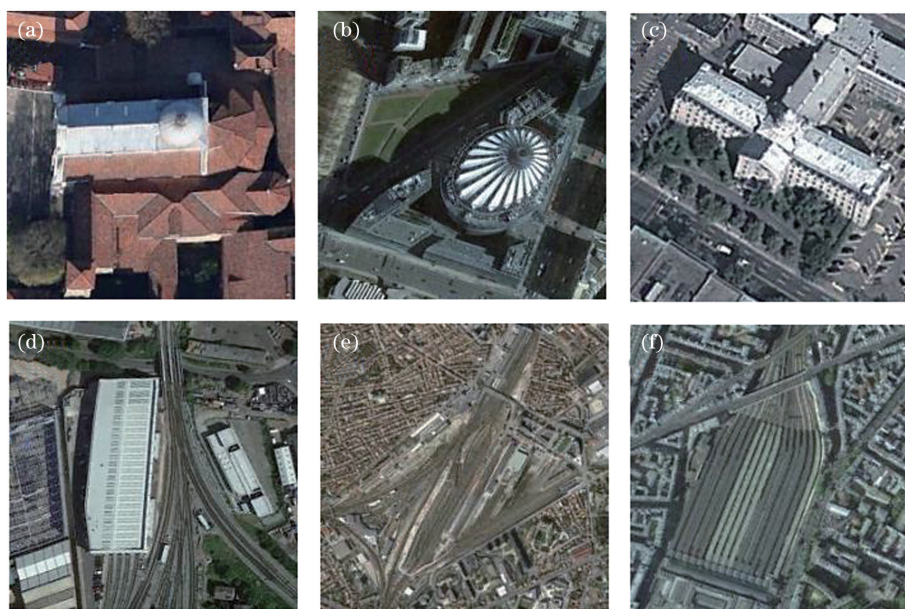


图 1 同一类的差异性。(a)(b)(c)教堂类别;(d)(e)(f)火车站类别

Fig. 1 In-class diversity. (a) (b) (c) Church category; (d) (e) (f) railway station category

随着卷积层深度的增加,生成的特征信息更加抽象化和多样化,这使得网络具有更强的学习能力。因此,CNN 在计算机视觉方面取得了重大突破。然而与单目标图像相比,卷积神经网络在复杂的空间关系中包含了多个目标,图像底层的细节特征被削

弱,这使得 CNN 在进行场景分类时不再具有区分性,从而无法得到令人满意的场景分类结果。因此,充分利用网络层的特征数据进行场景图像分类是非常必要的。所以,为了解决这一问题,本研究使用预先训练的 ResNet 网络,利用其中有快捷连接的残



图 2 不同类的相似性。(a)(b)高速公路与跑道类别;(c)(d)工业区与火车站类别;
Fig. 2 Between-class similarity. (a) (b) freeway versus runway; (c) (d) industrial area versus railway station;
(e) (f) stadium versus train station

差块从网络中的不同层提取特征映射,通过融合不同残差块提取的特征信息,达到显著增加深度而不过度拟合的效果,从而获得更高的分类精度。为了充分利用 ResNet 的优良性能,提出一种新的基于 ResNet 的结构,该结构可以将预先训练好的模型中的参数进行传递,其有别于只改变全连接层的传统方法,在 ResNet 网络模型中嵌入一个跳跃连接模块,以连接多分辨率特征映射进行分类。此外,还利用协方差池化方法对不同层次的多分辨率特征映射进行融合。与仅使用一阶统计量来集成特征映射的传统池化策略相比,协方差池化是一种使用二阶统计量信息来池化特征映射的方法。因此,这种方法可以提取更具代表性的特征,从而提高了 ResNet 在场景图像分类上的准确性。

ResNet 网络,又名残差网络,由 He 等^[10]提出,在 2015 年 ImageNet 比赛中排名第一。它主要设计了一种带有快捷连接的残差块,相当于在网络中增加了一条直连通道,使网络具有更强的恒等映射能力,从而在不过度拟合的情况下拓展了的网络深度,提高了网络性能。其中,残差块的结构如图 3 所示。

图 3 中, x 既表示残差块的输入也表示上一层输出的特征映射, $F(x)$ 表示 x 经过两层权重映射后的结果,ReLU 为修正单元,残差块的输出为 $F(x)+x$ 。使用这种结构解决了传统卷积层或全连接层在进行信息传递时存在的丢失、损耗等问题。通过直接将信息从输入绕道传输到输出,一定程度

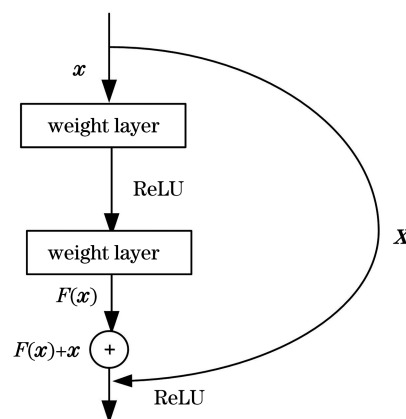


图 3 残差块结构

Fig. 3 Shortcut connection of resnet

上保护了信息的完整性。同时,由于学习的目标是残差,简化了学习的难度,缩短了消耗的时间。

2.2 实验原理与方法

本文提出的用于遥感图像场景分类的网络结构如图 4 所示。该模型是以 ResNet 网络为基础,主要包括以下几个步骤:1)预训练 ResNet 网络模型,提取残差块的特征信息;2)利用全局平均池化进行批量规范化,得到特征向量;3)利用跳跃连接和协方差池化的方法融合特征信息;4)通过 Softmax 层得到分类结果。下面,将详细地描述以上每个步骤。

2.2.1 预训练网络提取残差块之间的特征

预先训练的 18 层 ResNet 包含 8 个残差块,从相邻残差块提取的特征信息具有很大的相似性,且

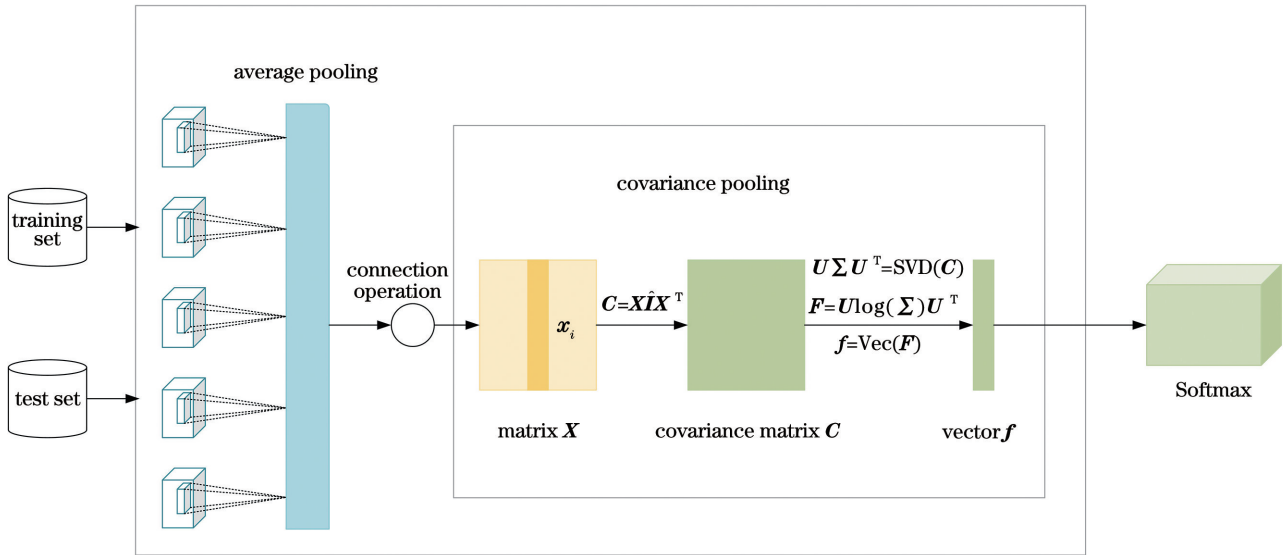


图 4 网络结构图

Fig. 4 Network structure diagram

从每个块中提取特征会增加网络的复杂度和计算时间,因此,本文在残差块之间提取特征,并且只执行 5 次特征提取。随着网络深度的增加,特征图的维数也随之增加。如果依次对特征映射进行卷积和合并,则维数会不断增加,故本研究利用全局平均池化操作(GAP)来降低特征映射的维数。在 GAP 之前,特征映射通过批处理规范化(BN)^[11]和激活函数(ReLu)进行预处理。对于具有 d 维 $\mathbf{X} = (\mathbf{x}_{(1)}, \mathbf{x}_{(2)}, \dots, \mathbf{x}_{(d)})$, 每个维度通过

$$\hat{\mathbf{x}}_{(k)} = \frac{\mathbf{x}_{(k)} - E[\mathbf{x}_{(k)}]}{\sqrt{\text{Var}[\mathbf{x}_{(k)}]}}, \quad (1)$$

其中, $\text{Var}(\cdot)$ 代表样本方差,则输出的 $\mathbf{y}_{(x)}$ 写为

$$\mathbf{y}_{(k)} = \gamma_{(k)} \hat{\mathbf{x}} + \beta_{(k)}, \quad (2)$$

$$\gamma_{(k)} = \sqrt{\text{Var}[\mathbf{x}_{(k)}]}, \beta_{(k)} = E[\mathbf{x}_{(k)}]. \quad (3)$$

校正线性单位(ReLU)为

$$\Phi(\mathbf{y}_{(k)}) = \max(0, |\mathbf{y}_{(k)}|), \quad (4)$$

将其作为激活函数,为神经网络增加非线性。

从残差块提取到的特征通过跳跃连接操作融合在一起。所得到的多分辨率特征映射用 \mathbf{X} 表示,并将 \mathbf{X} 的体积沿特征映射的通道维数重塑为矩阵。然后,利用协方差池化层对得到的多分辨率特征映射进行聚合。下面将详细介绍新添加的模块,即跳跃连接和协方差池。

2.2.2 跳跃连接

不同残差块可以提取到具有不同空间分辨率的图像特征映射,这些特征映射通过平均池化操作降低维数,得到 3 组空间分辨率相同的特征映射,即 $\mathbf{X}_1 \in \mathbf{R}^{H \times W \times D_1}$ 、 $\mathbf{X}_2 \in \mathbf{R}^{H \times W \times D_2}$ 和 $\mathbf{X}_3 \in \mathbf{R}^{H \times W \times D_3}$ 。

在这种情况下,可以通过跳跃连接操作来获得聚合的多分辨率特征映射 \mathbf{X} ,跳跃连接操作的表达式为

$$\mathbf{X} = [\mathbf{X}_1; \mathbf{X}_2; \mathbf{X}_3] \in \mathbf{R}^{H \times W \times (D_1 + D_2 + D_3)}, \quad (5)$$

其中, $[\cdot; \cdot; \cdot]$ 表示沿第三维度的连接操作。图 5 给出了 3 个特征映射的跳跃连接方法的图示。使用跳跃连接来聚合多层特征映射的原因有两个。1) CNN 模型可以通过分层的方法自然地提取金字塔形状的特征图,解决了分类和目标检测任务中的尺度方差问题。2) 利用卷积神经网络不同的网络层可以提取不同的特征,其中浅层网络用于提取图像的局部底层特征,但提取时也会存在背景杂乱和语义模糊的问题,深层网络用于提取图像的全局高层语义特征,用于整合不同场景之间的差异,但往往忽略细节信息。所以这些来自不同层的特征映射会包含互补信息^[12-14],表现出各种互补特征,因此提供了可用于跳跃连接操作的补充信息。

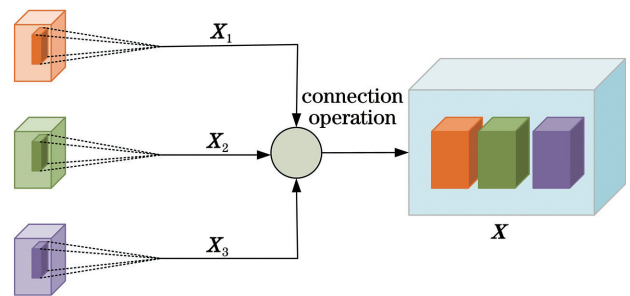


图 5 跳跃连接的图形示例

Fig. 5 Graphic example of jump connection

另外,为了连接具有不同空间分辨率的特征映射,采用平均池化操作来减少每一组特征映射的通

道数,从而得到空间分辨率相同的特征映射,然后利用跳跃连接操作将它们连接起来。

平均池化的数学定义如下。给定一个三维特征映射张量 $\mathbf{y} = [\mathbf{Y}_1; \mathbf{Y}_2; \mathbf{Y}_3] \in \mathbf{R}^{H \times W \times L}$, 其中 $\mathbf{Y}_i \in \mathbf{R}^{H \times W}$ 表示单一特征映射。假设步长为 k , 池化合并公式为

$$\mathbf{Z}_j = \frac{1}{k} \sum_{i=(j-1) \times k+1}^{j \times k} \mathbf{Y}_i, j=1, 2, \dots, L/k, \quad (6)$$

结果得到输出特征映射张量 $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_{L/k}] \in \mathbf{R}^{H \times W \times (L/k)}$ 。在实际中, 本文选择可以整除 L 的 k 值。

2.2.3 协方差池化

协方差池化的具体作用为融合多分辨率特征映射。经过跳跃连接, 可以得到以矩阵形式表示的多层特征映射, 利用协方差池化操作对矩阵进行奇异值分解得到对称矩阵, 然后进行矢量化得到特征向量。

给定特征矩阵 $\mathbf{X} \in \mathbf{R}^{D \times N}$ (如图 4 中的矩阵 \mathbf{X}), 每一行用 l_2 范数规范化, 其中 $D = D_1 + D_2 + D_3$ 是特征的维数, $N = H \times W$ 是特征的个数, 协方差池化传播如下。首先, 计算协方差矩阵 \mathbf{C} , 表达式为

$$\mathbf{X} \mapsto \mathbf{C}, \mathbf{C} = \mathbf{X} \hat{\mathbf{I}} \mathbf{X}^T, \quad (7)$$

式中: $\mathbf{X} \mapsto \mathbf{C}$ 代表将协方差矩阵 \mathbf{X} 转换成对称矩阵 \mathbf{C} ; $\hat{\mathbf{I}} = (1/N) [\mathbf{I} - (1/N) \mathbf{I} \mathbf{I}^T]$, $\hat{\mathbf{I}}$ 是一个 $N \times N$ 的恒等式矩阵, \mathbf{I} 是一个 N 维列向量, 所有项都设置为 1。然后, 使用矩阵对数将协方差矩阵从流形空间变换到欧氏空间, 以获得集合特征 \mathbf{F} , 表达式为

$$\mathbf{C} \mapsto \mathbf{F}, \mathbf{F} = \mathbf{U}(\log \mathbf{\Sigma}) \mathbf{U}^T, \quad (8)$$

其中 $\mathbf{C} \mapsto \mathbf{F}$ 代表将协方差矩阵 \mathbf{C} 转换成对称矩阵 \mathbf{F} ; $\mathbf{C} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^T$, \mathbf{U} 和 $\mathbf{\Sigma}$ 表示 \mathbf{C} 的特征向量矩阵和特征值矩阵。在图 4 中, \mathbf{f} 是 \mathbf{F} 的矢量化。其中, \mathbf{F} 是对称矩阵, 因此, 只需要对 \mathbf{F} 的上三角形中的项进行矢量化, 即向量 \mathbf{f} 的维数为 $D(D+1)/2$ 。

3 实验结果分析

3.1 实验数据集

为了评估该方法的性能, 在 3 个经典的遥感场景图像数据集上进行实验。

UC Merced Land Use 数据集 (如图 6 所示为部分类别图像) 是从美国地质调查局的国家地图城市图像中手工获取而来, 包含 2100 幅图像, 分为 21 个场景类。每个类由 100 幅图像组成。

Google of SIRI-WHU 数据集 (如图 7 所示为部分类别图像) 是从谷歌地图中获取, 主要覆盖中国的



图 6 UC Merced Land Use 遥感图像数据集。(a) 海滩类别; (b) 棒球场类别; (c) 立交桥类别

Fig. 6 UC Merced Land Use remote sensing image dataset.

(a) Beach; (b) baseball field; (c) overpass



图 7 Google of SIRI-WHU 遥感图像数据集。

(a) 河流类别; (b) 池塘类别; (c) 港口类别

Fig. 7 Google of SIRI-WHU sensing image dataset.

(a) River; (b) pond; (c) harbor

城市地区, 由武汉大学的 RS_IDEA 团队设计建成。该数据集包含 12 个类别, 每类场景包含 200 幅图像。

NWPU-RESISC45 数据集^[15] (如图 8 所示为部分类别图像) 是由西北工业大学 (NWPU) 创建的遥感图像场景分类的公开可用数据集, 包含 31500 幅图像, 分为 45 个场景类, 每个类包含 700 幅图像。这个数据集在场景类数量和图像总数方面都是比较大的。因此, 它与本文考虑的其他数据集相比, 包含更丰富的图像变化、更大的内部差异和更高的类间相似性。



图 8 NWPU-RESISC45 遥感图像数据集。(a) 森林类别; (b) 圆形农田类别; (c) 河流类别

Fig. 8 NWPU-RESISC45 sensing image dataset.

(a) Forest; (b) circular farmland; (c) river

3.2 环境与参数设置

表 1 为本文在训练网络时的实验环境, 采用的网络模型为 ResNet, 使用的损失函数为交叉熵 Cross Entropy Loss, 优化器为 Adam。初始学习率为 0.01, 每迭代 40 次学习率减小到原来的十分之一, 实验过程中随机划分训练集和验证集, 划分比例为 8:2。

表 1 实验环境介绍

Table 1 Introduction of experimental environment

Experimental environment	Environment configuration
Operating system	Ubuntu 16.04
Software environment	Python 2.7, pytorch 0.4.1
CPU	Xeon(R), W-2123
Internal memory	DDR4, 32G

3.3 实验结果

本文分别以未经改进的 ResNet 网络和经过本

文方法改进后的网络在上述 3 个公开的经典遥感数据集 (UC Merced Land Use 数据集、Google of SIRI-WHU 数据集和 NWPU-RESISC45 数据集) 中进行实验, 得到的实验结果如下。

3.3.1 类别的准确率对比

图 9~11 为未经改进的 ResNet 网络与本文方法在 UC Merced Land Use 数据集、Google of SIRI-WHU 数据集和 NWPU-RESISC45 数据集中每一个场景类的分类准确率对比柱状图。

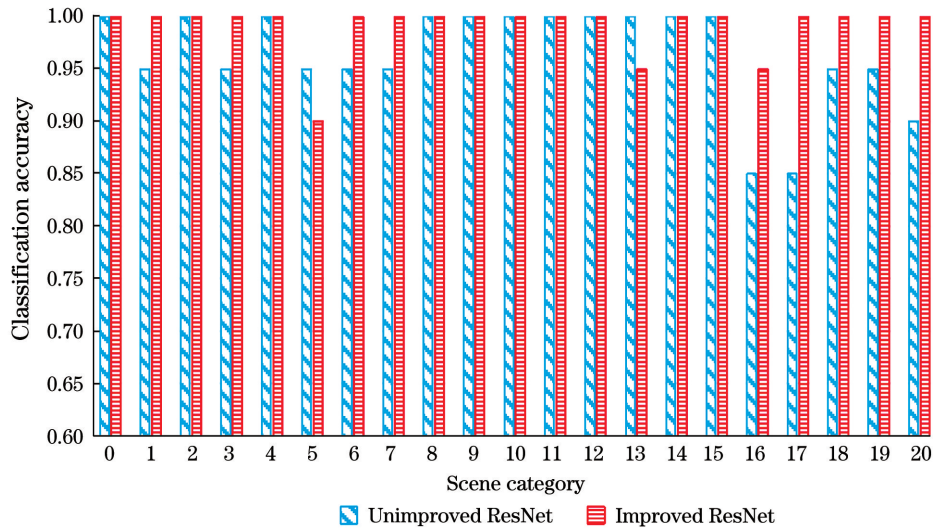


图 9 UC Merced Land Use 数据集的分类结果

Fig. 9 UC Merced Land Use data set classification results

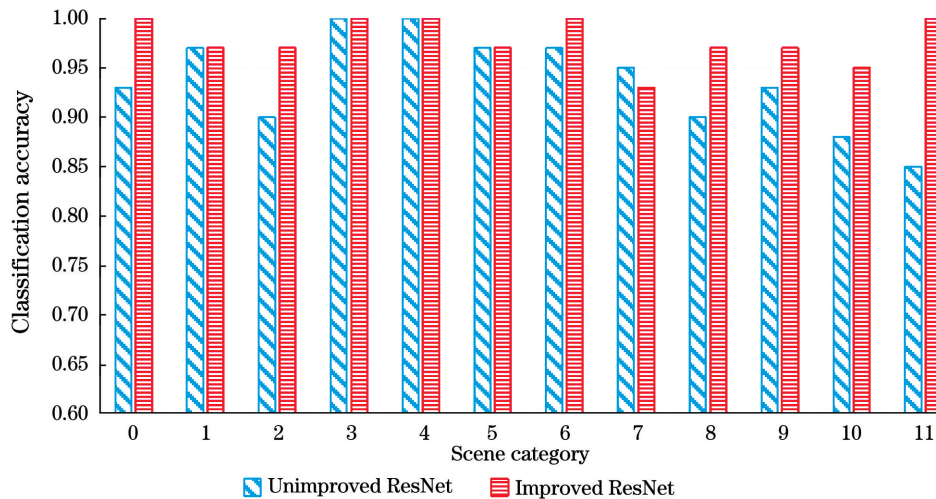


图 10 Google of SIRI-WHU 数据集的分类结果

Fig. 10 Google of SIRI-WHU data set classification results

从图 9~11 可以看出, 本文提出的改进方法明显地提高了数据集中一些类别的分类准确率。由图 8 可以得到, 在 UC Merced Land Use 数据集中所有类别的分类精度都达到了 90% 以上, 其中类别 16(储油罐)、17(建筑)和 20(稀疏住宅)的精度提高

明显, 分别提高了 10 个百分点、15 个百分点、10 个百分点。由图 9 可得, 在 Google of SIRI-WHU 数据集中, 类别 11(草地)的精度提高明显, 由 85% 提高到了 100%, 其他类别的精度也有着不同程度的提高。比较有挑战性的是 NWPU-RESISC45 数据

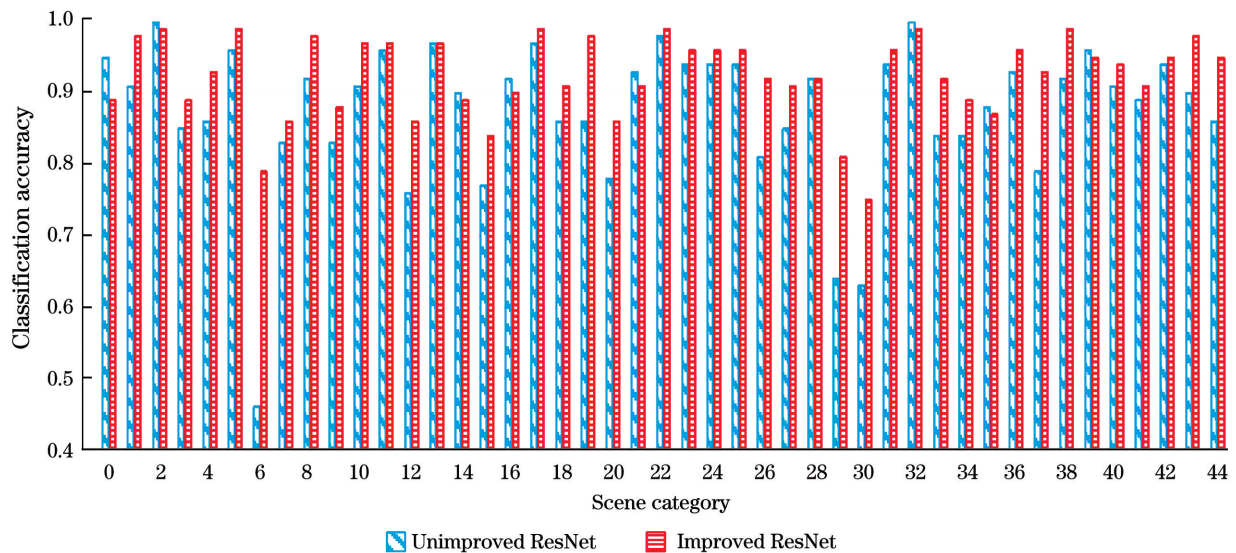


图 11 NWPU-RESISC45 数据集的分类结果

Fig. 11 NWPU-RESISC45 data set classification results

集,因为它的类别繁多、数量庞大,所以分类的难度也比较高。本文的研究方法在该数据集上也取得了不错的效果(如图 10 所示),与未改进的方法相比,类别 6(中等住宅)、19(梯田)、29(豪华住宅)、30(教堂)、37(火车站)的分类精度分别提高了 33 个百分点、12 个百分点、17 个百分点、12 个百分点和 14 个百分点。虽然在某些类别中本文提出的方法并未体现出优势甚至准确率有所降低,这主要因为这些类别中的图片相似性比较高,但并不影响整体效果。

3.3.2 类别间的错分率对比

在不同的数据集中不同类别间可能会出现错分,这是因为他们之间有着非常相似的特征信息。例如在 UC Merced Land Use 数据集中的公园类别和稠密住宅区类别、NWPU-RESISC45 数据集中的商业区类别和稠密住宅区类别、Google of SIRI-WHU 数据集中的住宅区类别和商业区类别等,它们都是描述居住区的,唯一的区别是建筑的密度。在 UC Merced Land Use 数据集中的农田类别和高尔夫球场类别、Google of SIRI-WHU 数据集中的池塘类别和港口类别等,这些都是描述比较空旷地区的,类间的相似性非常大。因此,在未经改进的网络中不同类别间很容易错分。根据实验结果,可知本文方法可以改善由类间相似性导致的错误分类,从而降低了类间的错分率。

3.3.3 与现有方法的对比

为了更好地突出本文方法的优势,表 2~4 分别列出了现有的几种研究方法与本文方法在以上 3 个数据集中的分类结果对比。

表 2 不同模型在 UC Merced Land Use 数据集的分类结果

Table 2 Comparison of the classification results obtained for the UC Merced Land Use dataset unit: %

Model	OA
DCA ^[7]	96.90
AlexNet+MSCP ^[14]	96.70
SCCov ^[16]	98.04
ResNet	96.70
Ours	99.76

表 3 不同模型在 Google of SIRI-WHU 数据集的分类精度

Table 3 Comparison of the classification results obtained for the Google of SIRI-WHU dataset unit: %

Model	OA
SRSCNN ^[17]	93.40
AlexNet+Softmax ^[18]	95.63
AlexNet+SVM ^[18]	95.83
ResNet	93.75
Ours	97.91

表 4 不同模型在 NWPU-RESISC45 数据集的分类精度

Table 4 Comparison of the classification results obtained for the NWPU-RESISC45 dataset unit: %

Model	OA
DCNN ^[9]	89.22
VGG+MSCP ^[14]	88.93
SCCov ^[16]	89.30
ResNet	87.61
Ours	92.45

从表 2~4 可以看出,本文方法在 UC Merced Land Use 数据集上的分类精度为 99.76%,在 Google of SIRI-WHU 数据集上的分类精度为 97.91%,在 NWPU-RESISC45 数据集分类精度为 92.45%,相比于现有的几种优秀方法,分类精度都有所提高。基于以上实验结果,本文提出的方法在提高分类准确率方面十分有效。

4 结 论

本文提出了一种新的用于光学遥感图像场景分类的网络模型。新的网络在 ResNet 网络中引入跳跃连接和协方差池化两个新的模块,不仅可以将在 ResNet 模型中不同层次的多分辨率特征信息融合在一起,而且可以利用高阶信息来实现更具代表性的特征学习。本文在 3 个公开的遥感图像数据集上进行了综合实验,并与现有的方法进行比较,验证了新方法的有效性。

本文方法虽然在一定程度上提高了分类准确率,但是对于某些类别的分类效果还并不十分完美。所以,在后继的研究中,本人将利用高光谱遥感图像,通过提取更多的光谱信息进行分析,从而实现目标更加准确的识别与分类

参 考 文 献

- [1] Zhu M M, Xu Y L, Ma S P, et al. Airport detection method with improved region-based convolutional neural network [J]. *Acta Optica Sinica*, 2018, 38 (7): 0728001.
朱明明, 许悦雷, 马时平, 等. 改进区域卷积神经网络的机场检测方法 [J]. *光学学报*, 2018, 38 (7): 0728001.
- [2] Jiang X J, Gao S H. Automatic classification of microscopic hair images based on improved convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231002.
姜晓佳, 高树辉. 基于改进卷积神经网络的毛发显微图像自动分类 [J]. *激光与光电子学进展*, 2019, 56 (23): 231002.
- [3] Ma J C, Zhao H D, Yang D X, et al. Design and optimization of deep convolutional neural network for aircraft target classification [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231006.
马俊成, 赵红东, 杨东旭, 等. 飞机目标分类的深度卷积神经网络设计优化 [J]. *激光与光电子学进展*, 2019, 56(23): 231006.
- [4] Wang Y N, Zhu D N, Wang H Q, et al. Multispectral image classification of mural pigments based on convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(22): 221001.
王燕妮, 朱丹娜, 王慧琴, 等. 基于卷积神经网络的壁画颜料多光谱图像分类 [J]. *激光与光电子学进展*, 2019, 56(22): 221001.
- [5] Castelluccio M, Poggi G, Sansone C, et al. Land use classification in remote sensing images by convolutional neural networks [J]. *Acta Ecologica Sinica*, 2015, 28(2): 627-635.
- [6] Cheng G, Li Z P, Yao X W, et al. Remote sensing image scene classification using bag of convolutional features [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(10): 1735-1739.
- [7] Chaib S, Liu H, Gu Y F, et al. Deep feature fusion for VHR remote sensing scene classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(8): 4775-4784.
- [8] Li E Z, Xia J S, Du P J, et al. Integrating multilayer features of convolutional neural networks for remote sensing scene classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55 (10): 5653-5665.
- [9] Cheng G, Yang C Y, Yao X W, et al. When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(5): 2811-2821.
- [10] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [11] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. (2015-02-11) [2020-06-16]. <https://arxiv.org/abs/1502.03167>.
- [12] Song W W, Li S T, Fang L Y, et al. Hyperspectral image classification with deep feature fusion network [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(6): 3173-3184.
- [13] Hu F, Xia G S, Hu J W, et al. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery [J]. *Remote Sensing*, 2015, 7(11): 14680-14707.
- [14] He N J, Fang L Y, Li S T, et al. Remote sensing scene classification using multilayer stacked covariance pooling [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56 (12): 6899-6910.
- [15] Cheng G, Han J W, Lu X Q. Remote sensing image

- scene classification: benchmark and state of the art [J]. Proceedings of the IEEE, 2017, 105(10): 1865-1883.
- [16] He N J, Fang L Y, Li S T, et al. Skip-connected covariance network for remote sensing scene classification [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(5): 1461-1474.
- [17] Liu Y F, Zhong Y F, Fei F, et al. Scene semantic classification based on random-scale stretched convolutional neural network for high-spatial resolution remote sensing imagery [C] // 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), July 10-15, 2016, Beijing, China. New York: IEEE Press, 2016: 763-766.
- [18] Zhang K, Hei B Q, Li S Y, et al. Complex scene classification of remote sensing images based on CNN [J]. Remote Sensing for Land & Resources, 2018, 30(4): 49-55.
- 张康, 黑保琴, 李盛阳, 等. 基于 CNN 模型的遥感图像复杂场景分类 [J]. 国土资源遥感, 2018, 30(4): 49-55.