

# SAT-CNN:基于卷积神经网络的遥感图像分类算法

张德园<sup>1</sup>,常云翔<sup>1</sup>,张利国<sup>2</sup>,石祥滨<sup>1</sup>

<sup>1</sup>(沈阳航空航天大学 计算机学院,沈阳 110136)

<sup>2</sup>(辽宁省通用航空研究院 人因工程研究室,沈阳 110136)

E-mail :dyzhang@sau.edu.cn

**摘 要:** 遥感图像空间分辨率较低,如何提取遥感图像特征是提升遥感图像分类性能的重要问题. 提出 SAT-CNN,一个基于卷积神经网络的遥感图像分类框架用于提取遥感图像特征. 设计了四个构件块并逐层堆叠构成 SAT-CNN,其中两个构件块用于提取遥感图像局部区域的底层特征,其他两个构件块用于提取遥感图像局部区域间的空间位置关系. 对 SAT-CNN 采用三种不同的参数配置进行训练,通过分析分类精度、SAT-CNN 特征的迁移性以及图像通道来研究 SAT-CNN 特征提取的准确性和通用性. 实验结果表明, SAT-CNN 在 SAT-4 以及 SAT-6 数据集上达到了较高的分类精度,能够提取适合遥感图像分类的图像特征.

**关 键 词:** 遥感图像分类;深度学习;卷积神经网络;特征迁移

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2018)04-0859-06

## SAT-CNN: Convolutional Neural Network Framework for Remote Sensing Image Classification

ZHANG De-yuan<sup>1</sup>, CHANG Yun-xiang<sup>1</sup>, ZHANG Li-guo<sup>2</sup>, SHI Xiang-bin<sup>1</sup>

<sup>1</sup>(School of Computer, Shenyang Aerospace University, Shenyang 110136, China)

<sup>2</sup>(Human Factors Engineering Laboratory, Liaoning General Aviation Academy, Shenyang 110136, China)

**Abstract:** Remote sensing image classification is an important approach to monitor the earth's surface. Due to the low spatial resolution of remote sensing image, feature extraction from remote sensing image is an important issue to improve the performance of remote sensing image classification. In this paper, we propose SAT-CNN network, a framework for remote sensing image classification based on convolutional neural network. We design four blocks and describe the SAT-CNN network with stacks of blocks. Two blocks are designed to extract local features of remote sensing image, and the other two components are used to extract spatial location relationships between different locations of remote sensing image. The SAT-CNN is trained with three different configurations. The accuracy and generalization of SAT-CNN features are studied by analyzing the classification accuracy, the transferability of SAT-CNN features and different channels of image. The experiment results on SAT-4 and SAT-6 dataset show that the SAT-CNN performs high accuracy and can extract better feature from remote sensing images.

**Key words:** remote sensing images classification; deep learning; convolutional neural network; feature transfer

## 1 引 言

遥感图像分类是根据卫星等遥感设备拍摄到的图像来识别地面目标,常用于监测地球的地表状况. 遥感图像受到传感器类型、波长、拍摄区域等因素的影响,使得采集到的图像空间分辨率较低、图像细节较模糊,因此遥感图像分类问题具有很大的挑战性<sup>[1]</sup>. 鉴于遥感图像标注样本数量较少,传统的遥感图像分类方法主要研究适合遥感图像的特征提取算法<sup>[2-4]</sup>,以及小样本分类算法<sup>[5]</sup>. 为了适应不同应用的特征普适性,人们提出了一些方法来获取更加灵活和泛化能力更强的特征,比如:稀疏编码<sup>[6]</sup>、fisher 特征<sup>[7]</sup>等.

近年来深度神经网络由于其优异的性能在计算机视觉的各个领域得到广泛的研究<sup>[8,9]</sup>. 其主要的成功在于能够通过

端到端的方式从大量标注的训练数据中学习到具有极强表达能力和泛化能力的特征. 深度神经网络在遥感图像领域的应用正在引起研究者的关注<sup>[10]</sup>.

然而在遥感图像分类任务上,采用遥感图像数据集训练端到端神经网络尚未得到深入的研究. 主要原因在于深度神经网络需要大量的训练数据,而遥感图像数据集中数据量较少,无法满足这种庞大的数据量需求<sup>[11]</sup>. 大规模的高分辨率遥感图像数据集直到 2015 年才出现, Basu<sup>[12]</sup> 通过对图像提取 150 种不同类型的特征,采用 DeepSAT 深度神经网络框架分类取得了较优的性能,然而采用端到端的神经网络训练模型是否适合遥感图像分类仍然未得到深入的研究.

本文提出了 SAT-CNN,一个用于遥感图像分类任务的端到端的深度神经网络框架,该框架有四个构件块构成,其中两

个构件块用于提取遥感图像局部区域的底层特征,其他两个构件块用于提取遥感图像局部区域的空间位置关系.采用三种不同网络配置进行训练,实验结果表明 SAT-CNN 能够提取适合遥感图像分类的图像特征.对 SAT-CNN 的特征迁移性以及受图像通道的影响进行了详细的分析,验证了 SAT-CNN 框架得到特征的泛化性和鲁棒性.

## 2 相关工作

遥感图像分类任务过去的工作主要集中于设计复合滤波器,研究提取具有良好表达能力的特征来描述图像. Camps-Valls 等人<sup>[13]</sup>通过对同一区域的光谱和空间特征进行加权求和,设计了一个复合的滤波器.文献[14]中提出一种基于使用贝叶斯架构(bayes framework)的目标检测方法并使用 Latent Dirichlet Allocation 对目标进行后续的聚类.这种方法对检测高度抽象的目标群体(比如,停车场)有较好的效果.

过去几年,计算机视觉领域发生巨大变革并推进到采取一种快速的、可扩展的、端到端的学习架构:卷积神经网络(convolutional neural network, CNN).基于深度卷积神经网络的方法在分类、定位、语义分割和动作识别等任务上比其他方法得到远超出传统方法的精度.在大规模识别任务中学习到卷积神经网络特征已经成功迁移到了不同的视觉任务上,比如目标检测、语义分割和动作检测.虽然深度学习在计算机视觉应领域的所有方法中精度最高,但在遥感图像 1 分类任务上研究起步较晚,尚未得到充分的研究.文献[14]中介绍了一种深度学习算法,结合空间和光谱信息对高分辨率图像(VHR)进行有监督分类. Minh<sup>[15]</sup>等人研究了深度网络对航拍图像分类中类别标签噪声的鲁棒性.近几年, Romero<sup>[16]</sup>提出了一种深度卷积分层的架构,在 AVIRS Indiana's Indian 数据集上实验和测试分类的精度. Tokarczyk<sup>[17]</sup>中改进了特征提取技术和分类器提出了一种改进的多特征学习策略.由于训练数据受到明显的限制,他们的方法相比于随机森林没有过多的提升.文献[12]中提出了一种基于深度网络的遥感图像的分类架构 DeepSat,并构建了大规模高分辨率的遥感图像数据集 SAT-4 和 SAT-6. DeepSat 方法从图像中提取特征、进行归一化,然后采用深度置信网络对这些归一化后的特征进行分类.当有足够多的训练数据时,DeepSat 能够得到一个很高的分类精度.在大量标注训练样本的情况下,采用端到端的神经网络是否能够进行遥感图像特征提取是一个尚未充分研究的问题.本文我们采用端到端的方式直接使用深度卷积神经网络在未经理处理的遥感图像数据上学习和分类,分析深度学习在遥感图像分类任务上的性能.

## 3 卷积神经网络介绍

近年来卷积神经网络在图像分类任务上大幅提升传统方法的性能.卷积神经网络的发展主要得益于海量数据以及多种隐藏层神经元的提出,主要的隐藏层神经元操作描述如下:

1) 卷积操作 conv(convolutional): 卷积操作使用滑动窗口分割原始图像,在每个滑窗内使用相同的参数进行线性计算得到输出图像.在一张图像上使用一个卷积核(kernel)的进行卷积运算的计算如公式(1)所示:

$$\begin{aligned} y_{h'',w''} &= b + \sum_{h'=1}^{H'} \sum_{w'=1}^{W'} \sum_{c=1}^C f_{h',w',c} \times x_{i,j,c} \\ i &= S_h \times (h''-1) + h' - P_h^- \\ j &= S_w \times (w''-1) + w' - P_w^- \end{aligned} \quad (1)$$

其中,  $x \in R^{W \times H \times C}$  为输入图像;  $y \in R^{W'' \times H''}$  为输出图像;  $f \in R^{W' \times H' \times C}$  为卷积核,即在滑窗内使用的线性加权求和系数;  $(P_h^-, P_h^+, P_w^-, P_w^+)$  依次为在图像的上下左右四个方向上的零填充数量;  $(S_h, S_w)$  为滑窗在原始图像上每次滑动的距离.通常卷积神经网络的一个卷积层中可以使用多个卷积核进行多核卷积操作得到多张输出图像.

2) 批规范化操作 bnorm(batch normalization)<sup>[18]</sup>: 批规范化操作通过对神经网络隐藏层的输出进行归一化处理防止反向传播过程中的梯度弥散,加速神经网络在训练时的收敛速度.其计算如公式(2)所示:

$$\begin{aligned} y_{h,w,c,d} &= \omega_c \frac{x_{h,w,c,d} - \mu_c}{\sqrt{\sigma_c^2 + \varepsilon}} + b_c \\ \mu_c &= \frac{1}{HWD} \sum_{h=1}^H \sum_{w=1}^W \sum_{d=1}^D x_{h,w,c,d} \\ \sigma_c^2 &= \frac{1}{HWD} \sum_{h=1}^H \sum_{w=1}^W \sum_{d=1}^D (x_{h,w,c,d} - \mu_c)^2 \end{aligned} \quad (2)$$

其中,  $x \in R^{H \times W \times C \times D}$  为输入图像;  $y \in R^{H \times W \times C \times D}$  为输出图像;  $\omega, b \in R^C$  为线性变换系数;  $\mu_c$  和  $\sigma_c^2$  为均值和方差.

3) 修正线性操作 ReLU(rectified linear unit): 修正线性操作是网络中使用到的激活函数,其计算如公式(3)所示:

$$y_{h,w,c} = \max\{0, x_{h,w,c}\} \quad (3)$$

其中,  $x \in R^{H \times W \times C}$  为输入图像;  $y \in R^{H \times W \times C}$  为输出图像.

4) 最大池化 mp(maxpooling): 最大池化操作采用滑动窗口分割图像,在每个滑窗内取最大值作为输出,其计算如公式(4)所示:

$$y_{h'',w'',c} = \max_{1 \leq h' \leq H', 1 \leq w' \leq W'} x_{h''+h'-1, w''+w'-1, c} \quad (4)$$

其中,  $x \in R^{H \times W \times C}$  为输入图像;  $y \in R^{H'' \times W'' \times C}$  为输出图像;  $f \in R^{H' \times W' \times C}$  为滑窗大小,窗口每次滑动一个窗口大小的步长.

5) 全连接操作 fc(fully connected): 全连接操作对输入图像的所有数据进行加权求和的线性运算,可以通过卷积操作的方式来实现.公式(1)中窗口大小  $f \in R^{W' \times H' \times C}$  修改为  $f \in R^{W \times H \times C}$  即可将卷积操作变为全连接操作.

## 4 SAT-CNN 框架

### 4.1 SAT-CNN 网络设计

本文针对遥感图像数据集的特点提出了用于遥感图像分类的 SAT-CNN 卷积神经网络框架,主要考虑了两方面因素:第一,遥感图像空间分辨率较低导致图像的细节较为模糊,识别遥感图像主要依赖于遥感图像局部区域间的整体关系,不需要过多考虑遥感图像局部区域内存在复杂的空间位置关系;第二,遥感图像数据集中训练数据量相对较少,采用更大、更深的网络结构将导致 SAT-CNN 中有更多的参数需要训练,需要更大的训练数据量.为了便于实现和分析,我们依据模块化的思想提出了具有不同功能的构件块用来描述 SAT-CNN 的网络结构. SAT-CNN 由四个构件块(block)通过线性

堆叠的方式组成,每个构件块实现一个基本的功能:局部特征提取、全局特征提取和标签预测.如图 1 所示,构件块 1 和构件块 2 用于从遥感图像的局部区域内提取底层特征;构件块 3 用于整合不同底层特征间的非线性空间位置关系得到整张图像的全局特征;构件块 4 根据整张图像的特征对图像在每种类别上的可能性做出预测,得出预测标签.

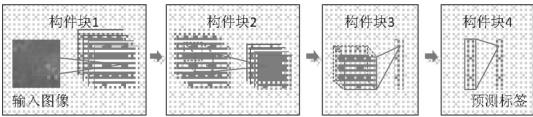


图 1 SAT-CNN 网络结构图  
Fig. 1 SAT-CNN network structure

SAT-CNN 中根据 LeNet、AlexNet 和 VggNet 三种经典的卷积神经网络的结构共性,设计了四种不同类型的构件块,分别命名为:CM、CCM、FC 和 Pre.其中,CM 和 CCM 类型构建块用于提取局部特征,FC 类型构建块用于提取全局特征,Pre 类型构建块用于预测类别标签.每个构件块由线性变换(conv)、批规范化(bnorm)、非线性变换(ReLU 或 pooling)等基本操作组成.四种不同类型构建块的内部结构如下:

- 1) CM:该构件块由一组卷积(conv)-批规范化(bnorm)-最大池化(mp)的三层结构构成.该构件块能够提取较简单的区域内像素关系.
- 2) CCM:该构建块由一组卷积(conv)-批规范化(bnorm)-激活函数(ReLU)-卷积(conv)-批规范化(bnorm)-最大池化(mp)组成的六层结构.其中两个卷积操作使用相同大小、相同个数的卷积操作.与 CMBLOCK 相比,该构件块提取较复杂的区域内像素关系.
- 3) FC:网络中使用到的一种全连接构件,由一组全连接(fc)-批规范化(bnorm)-激活函数(ReLU)组成的三层结构.该构件块能够提取不同区域特征之间的非线性空间关系.常用于将网络中所有的局部特征组合得到表示整张图像的高层特征.
- 4) Pre:网络中使用到的另一种全连接构件块,仅由一个全连接(fc)操作组成.与 FCBLOCK 不同,该构件块能够提取不同区域特征之间的线性空间关系.常用于将网络中提取到的特征输出成对每个类别标签的预测值.

SAT-CNN 网络由以上四种类型的构件块线性堆叠组成,并分别借鉴图像分类中常用的 LeNet、AlexNet 和 VggNet 网络结构,设置了三种不同参数的网络结构,分别命名为 SAT-LeNet、SAT-AlexNet 和 SAT-VggNet.三种不同网络结构的具体配置如表 1 所示.

表 1 中采用构件块的方式描述网络结构,其中 K 为构件块中卷积操作使用的卷积核大小,D 为使用的卷积核个数,P 为填充像素数量,S 为每次滑动的步长.

SAT-LeNet 和 SAT-AlexNet 的网络由两个 CM 类型的构建块、一个 FC 类型的构件块和一个 Pre 类型的构件块组成. SAT-LeNet 的前两个构件块用  $5 \times 5$  的卷积核“感知”图像上的局部信息. SAT-AlexNet 中采用了比 SAT-LeNet 更大的卷积核来提取更大区域的浅层局部特征,增加网络中的参数数量,分别使用  $11 \times 11$  和  $7 \times 7$  的卷积核. SAT-VggNet 中使

用到的卷积构件块为 CCM 类型的构建块,能够在图像的局部区域提取更复杂的特征、增加网络的深度.网络的前两个构件块中采用  $3 \times 3$  的卷积核.三种网络结构的后两个构件块中

表 1 SAT-CNN 网络配置  
Table 1 SAT-CNN network configuration

NetType	Block1	Block2	Block3	Block4
SAT-LeNet	CM	CM	FC	Pre
	K: $5 \times 5$	K: $5 \times 5$	K: $4 \times 4$	K: $1 \times 1$
	D: 32	D: 64	D: 128	D: 4or6
	P: 0, S: 1	P: 0, S: 1	P: 0, S: 1	P: 0, S: 1
SAT-AlexNet	CM	CM	FC	Pre
	K: $11 \times 11$	K: $7 \times 7$	K: $2 \times 2$	K: $1 \times 1$
	D: 32	D: 64	D: 128	D: 4or6
	P: 1, S: 1	P: 0, S: 1	P: 0, S: 1	P: 0, S: 1
SAT-VggNet	CCM	CCM	FC	Pre
	K: $3 \times 3$	K: $3 \times 3$	K: $4 \times 4$	K: $1 \times 1$
	D: 32	D: 64	D: 128	D: 4or6
	P: 0, S: 1	P: 0, S: 1	P: 0, S: 1	P: 0, S: 1

的线性操作为全连接操作,使用与前一层输出大小相同的卷积核来“感知”整张图像的信息.第一个全连接层输出的 128 维向量为神经网络从图像上学习到用来表示整张图像的全局特征;最后一个全连接层为预测构建块,输出 4 或 6 维的向量,对应于我们使用的数据集上的标签数量.输出的标签预测向量通过 softmax 层的计算得到一个可以用来表示图像在每个类别可能性的后验概率.

4.2 训练方法

采用计算预测值与实际标签之间的对数损失(softmax log-loss)交叉熵作为优化目标函数,使用基于动量的随机梯度下降的方法进行优化.将训练数据集中的数据每 500 个样本划分成一个数据块(mini batch),在全部数据块上调整参数为一轮迭代,设置迭代 200 轮寻找最优的网络参数.网络权重采用正态分布的随机采样来初始化,偏置系数初始化为 0,学习率设置为 0.001,动量值设置为 0.9.在训练时,输入 SAT 数据集中  $28 \times 28$  大小的四通道图像,输出为图像在每个类别上的线性预测值,根据预测值的大小判断网络对图像预测的类别.我们采用原始的图像作为输入数据,没有做任何的数据增强和图像预处理.文中使用到的三种网络都是在 SAT 数据集上重新训练得到的.

5 实验

5.1 实验设置

实验在 SAT-4 和 SAT-6 数据集的训练集上训练本文提出的三种不同结构的卷积神经网络模型,为每种类型的模型确定最优的参数并在测试集上测试分类精度.我们的实验结果将与 SAT-4 和 SAT-6 数据集上其他方法的结果进行比较.

SAT-4 和 SAT-6 是一组公开的高分辨率遥感图像数据集.这两个数据集中的图像从 National Agriculture Imagery Program (NAIP) 截取得到.该图像的地面空间分辨率为 1 米.图像包含 4 个通道,红(R)、绿(G)、蓝(B)以及近红外



(NIR)通道. SAT-4 包含共 50 万张图片,分为贫瘠土地 (barren)、树木 (trees)、草地 (grass) 以及除了以上三种类别的其他类别 (none) 共 4 个类别. SAT-6 包含 6 个土地覆盖类别: 贫瘠的土地 (barren)、树木 (trees)、草地 (grass)、道路 (road)、建筑物 (building) 和水体 (water). SAT 数据集中每个图像大小均为  $28 \times 28$  像素,训练和测试数据集使用伪随机数生成器随机化.

实验中使用 SAT-4 和 SAT-6 数据集中的 80% 数据作为训练数据来训练网络模型,20% 作为测试数据用来测试模型的性能. SAT-4 和 SAT-6 数据集的具体配置如表 2 所示.

表 2 SAT 数据集实验配置  
Table 2 SAT dataset configuration

	训练集	验证集	测试集
SAT-4	320000	80000	100000
SAT-6	259200	64800	81000

实验使用 MatConvNet19 构建文中提出的三种不同类型的卷积神经网络模型. 在一块带有 6GB 显存的 Nvidia GeForce GTX980Ti 显卡上使用 Nvidia 提供的 CuDNN 加速网络的训练过程. 训练过程周期性地地在训练集合的子集上评估交叉熵目标函数,调整网络参数,直到训练交叉熵稳定为止. 对于该部分算法,我们没有在验证集合上观察到过拟合.

5.2 SAT-CNN 的分类性能比较与分析

对于文中提出的三种模型,我们分别在 SAT-4 和 SAT-6 的训练集上训练卷积神经网络. 选择每种模型在验证集上性能最优的模型测试分类精度. SAT-CNN 与 DeepSat<sup>[12]</sup> 方法和 DBN 方法的分类精度如表 3 所示.

表 3 SAT-CNN 在 SAT-4 与 SAT-6 上的分类精度  
Table 3 Accuracy of SAT-CNN on SAT-4 and SAT-6

	SAT-4 数据集上 分类精度 (%)	SAT-6 数据集 上分类精度 (%)
SAT-LeNet	99.37	99.63
SAT-AlexNet	98.88	99.18
SAT-VggNet	99.98	99.97
DeepSat[12]	97.95	93.92
DBN[12]	81.78	76.47

本文提出的 SAT-CNN 获得的精度均达到了 98% 以上,比 DeepSat 方法有更优秀的表现. 三种不同类型的卷积神经网络在 SAT-4 和 SAT-6 数据集上都达到很高的分类精度,其中性能最佳的 SAT-VggNet 网络在 SAT-4 和 SAT-6 数据集上的分类精度分别为 99.98% 和 99.97%. 可以看出,卷积神经网络可以在不对数据进行任何预处理的情况下从原始图像数据中进行学习. 相较于传统的提取特征-选择分类器的遥感图像分类方法,SAT-CNN 在图像分类精度上有极大的提升.

从表 3 的实验结果中可以看出,SAT-6 数据集上各种网络结构的分类精度与 SAT-4 数据集上的相比差距不大,甚至高于 SAT-4 数据集上各种网络结构的分类精度. SAT-6 数据集比 SAT-4 数据集中有更多更复杂的类别需要区分,从直觉上来说其类精度应低于 SAT-4 数据集上的分类精度. 我们对 SAT-VggNet 在两个数据集上的分类结果进行分析,其迷惑矩

阵分别如表 4 和表 5 所示. 表中第一列为样本的真实标签,第一行为样本的预测标签,每行的数值表示各个类别的样本经过 VggNet 预测后分类到的各个类别的样本数量.

表 4 中,SAT-4 数据集中共分错 20 个样本,barren 误分为 grass 的有 11 个;表 5 中,SAT-6 数据集中共分错 27 个样本,

表 4 SAT-VggNet 在 SAT-4 数据集上的迷惑矩阵  
Table 4 Confusion matrix of SAT-VggNet on SAT-4

	barren	trees	grass	none
Barren	26177	1	11	0
Trees	1	20230	0	0
Grass	1	1	17943	1
None	4	0	0	35630

barren 误分为 grass 的有 24 个. 由迷惑矩阵可以看出,在 SAT-4 和 SAT-6 数据集上最容易混淆的两类均为 barren 和 grass.

表 5 SAT-VggNet 在 SAT-6 数据集上的迷惑矩阵  
Table 5 Confusion matrix of SAT-VggNet on SAT-6

	building	barren	trees	grass	road	water
building	3714	0	0	0	0	0
barren	0	18352	1	14	0	0
tree	0	0	14185	0	0	0
grass	0	10	0	12586	0	0
road	2	0	0	0	2068	0
water	0	0	0	0	0	30068

由于 SAT-6 数据集中包含的 barren 和 grass 两个类别的样本比例相对于 SAT-4 数据集包含的样本比例要更少,所以 SAT-6 比 SAT-4 数据集要更加容易区分. 图 2 为数据集中标签为 barren 被分类成 grass 的图像.

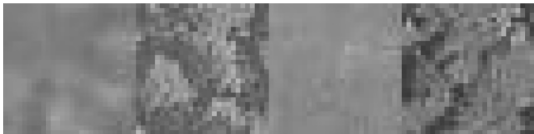


图 2 SAT-4 数据集中 barren 被误分类为 grass 的图像样例  
Fig. 2 Barren images misclassified as grass in SAT-4

可以看出,SAT 数据集中存在部分标签为 barren 实际更接近 grass 的图像. 主要原因在于 SAT 数据集在构建时采用一张完整的大小的遥感图像进行分割标注,有些小区域中可能包含两种或多种地面目标,所以数据集中可能存在部分类别之间有交叉的样本图像.

5.3 SAT-CNN 的特征迁移性分析

神经网络迁移是用训练后的神经网络模型在新数据集上提取特征进行分类. 实验中,使用 SAT-4 数据集上训练得到的三种网络模型和 SAT-6 上训练得到的网络模型分别迁移到 SAT-6 和 SAT-4 数据集上. 将迁移网络的最后一个预测构建块移除,余下的部分作为特征提取器在迁移数据集上提取特征. 然后,使用迁移数据集上网络结构中最后一个预测构件块作为线性分类器对特征进行分类. 简而言之就是将迁移网络最后一层的预测构件块替换为迁移数据集上相应的预测构

**件块**,重新调整预测构件块的参数至最优,并在测试集上测试分类精度.实验结果如表6所示.

三种 SAT-CNN 网络结构均有良好的迁移性能. SAT-4 网络模型向 SAT-6 数据集迁移的结果中,最高的分类精度为 96.31%;SAT-6 网络向 SAT-4 数据集迁移的结果中为 92.70%.虽然迁移后的方法性能有所下降,但相对于传统的方法这种迁移的结果仍有较好的精度.说明了卷积神经网络能够从大

表6 采用 SAT-CNN 的特征迁移的分类精度表  
Table 6 Accuracy of SAT-CNN's feature transfer

	NET-4 迁移到 SAT-6 数据集分类精度(%)	NET-6 迁移到 SAT-4 数据集上分类精度(%)
SAT-LeNet	94.04	92.70
SAT-AlexNet	96.31	86.85
SAT-VggNet	92.24	90.42

量的数据中通过端到端的结构学习到辨别性很强的高层抽象特征.这种特征能够采用线性的方式划分来对图像分类,并且能够很好地用来提升传统分类器的性能.

5.4 图像通道对 SAT-CNN 的影响分析

普通的图像分类任务中使用到的图像为可见光图像,即 RGB 三通道图像.遥感图像数据集 SAT-4 和 SAT-6 中的图像与 RGB 图像相比增加了一个近红外光线拍摄到的图像.为了研究遥感图像中图像通道对 SAT-CNN 的影响,我们分别使用 SAT 数据集中四通道图像的可见光(RGB)和近红外(nir)通道的数据进行分类.提取 SAT-4 和 SAT-6 数据集四通道图像中的前三个通道生成可见光图像,构建新的 SAT 可见光数据集 SAT-4-rgb 和 SAT-6-rgb.使用 SAT-4 和 SAT-6 数据集中图像的第四个通道生成近红外图像,构建 SAT 近红外数据集 SAT-4-nir 和 SAT-6-nir.

根据 5.2 节的实验结果,采用 SAT-VggNet 网络结构分别在 SAT-4-rgb、SAT-6-rgb、SAT-4-nir 和 SAT-6-nir 四个新生成的数据集上进行实验.修改 SAT-VggNet 网络中的输入层,使卷积核的维度与新构建的四个数据集中的数据的维度相同,保持其他网络配置不变.

重新在四个数据集上采用与 4.2 节实验中相同的配置训练网络.从 200 轮迭代中选取在验证集上误差最小的网络作为测试用网络模型.分别使用在 SAT-4-rgb、SAT-6-rgb、SAT-4-nir 和 SAT-6-nir 数据集上得到的网络模型在相应的测试数据集上进行分类测试.表 7 为使用 SAT-VggNet 在四种数据集上的测试精度.

表7 SAT-VggNet 采用不同通道数据的分类精度

Table 7 Accuracy of SAT-VggNet classify on different channels

数据集	分类精度(%)
SAT-4-rgb	99.97
SAT-6-rgb	99.97
SAT-4-nir	97.71
SAT-6-nir	98.88

在使用了可见光部分 SAT-4-rgb 和 SAT-6-rgb 进行的分类实验中,分类精度均达到了 99.97%.与原始的四通道数据

集上的测试精度相比,仅丢失了 0.01%~0.02% 的分类精度.使用了近红外部分的数据集 SAT-4-nir 和 SAT-6-nir 进行的实验中,分类精度分别为 97.71% 和 98.88%.与原始数据集相比,分别丢失了 2.28% 和 1.1% 的分类精度.在 SAT 数据集中,地面目标的类别较少且各类目标间差异较为显著,便于分类.同时,由于各个类别图像的颜色和纹理等可见光部分的特性较明显,采用可见光部分进行图像分类能够达到较高的分类精度.此外,不同的地面目标反射的近红外光线不同,近红外图像能够在一定程度上表现不同目标间的差异从而用来进行目标分类.卷积神经网络能够从不同的数据通道中学习到体现目标在该光谱下得特性,提取易于分辨的特征,得到较好的分类结果.

6 结 论

本文采用模块化的思想构建了面向遥感图像分类的 SAT-CNN 卷积神经网络框架,并设计了四种不同类型的构件块对网络结构进行描述和实现.根据遥感图像特点,SAT-CNN 堆叠两个卷积功能的构件块和两个全连接功能的构件块,分别用于建模图像局部区域内的像素关系以及区域间的像素关系,并给出三种不同的网络参数设置.实验分析说明了卷积神经网络可以通过端到端的方式从图像中学习到的显著性较强的特征,在 SAT 数据集上训练得到的 SAT-CNN 达到了较高的分类精度.针对 SAT-CNN 是否能够作为遥感图像的特征提取模块,对迁移性的分析表明 SAT-CNN 作为特征提取模块的可行性.此外,本文分析了 SAT-CNN 在不同图像通道上的性能.实验结果表明,使用可见光部分通道时性能几乎没有下降;在近红外通道仍然保持较高的分类精度,说明了 SAT 数据集中的可见光部分中包含了大部分与目标相关的特性.

References:

[1] Panchal S,Thakker R. Implementation and comparative quantitative assessment of different multispectral image pansharpening approaches [J]. Journal of Virology,2015,73(6):4622-4630.

[2] Xie Yong-hua,Yang Jing-yu. Aproach to remote-sensing image detection based on multi-direction two-scale wavelet transform [J]. Mini-Micro Systems,2005,26(6):1054-1056.

[3] Qiu Jian-bin,Li Shi-jin,Zhu Yue-long,et al. Fast hyperspectral image band selection method based on curve shape feature[J]. Journal of Chinese Computer Systems,2014,35(8):1906-1910.

[4] Mojaradi B,Abrahami-Moghaddam H,Zoej M J V,et al. Dimensionality reduction of hyperspectral data via spectral feature extraction[J]. IEEE Transactions on Geoscience & Remote Sensing, 2008,12(4):579-585.

[5] Liu Meng-meng,Liu Ya-lan,Sun Guo-qing,et al. SVM land cover classification based on spectral and textural features using stratified samples[J]. Remote Sensing Technology and Application,2014,29(2):315-323.

[6] Sun Xiao-xia,Qu Qing,Nasrabadi N M,et al. Structured priors for sparse-representation-based hyperspectral image classification[J]. IEEE Geoscience & Remote Sensing Letters,2014,11(7):1235-1239.

[7] Yang Guo-peng,Yu Xu-chu,Chen Wei,et al. Hyperspectral remote sensing image classification based on kernel fisher discriminant a-

- analysis[J]. Journal of Remote Sensing, 2008, 12(4): 579-585.
- [8] Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks[J]. In Advances in Neural Information Processing Systems, 2012: 1097-105.
- [9] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition[C]. International Conference on Learning Representations, 2015: 1-14.
- [10] Zhang Liang-pei, Shen Huan-feng. Progress and future of remote sensing data fusion[J]. Journal of Remote Sensing, 2016, 20(5): 1050-1061.
- [11] Zhang Liang-pei, Li Jia-yi. Development and prospect of sparse representation-based hyperspectral image processing and analysis[J]. Journal of Remote Sensing, 2016, 20(5): 1091-1101.
- [12] Basu S, Ganguly S, Mukhopadhyay S, et al. Deepsat: a learning framework for satellite imagery[C]. Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2015, 37.
- [13] Camps-Valls G, Gomez-Chova L, Munoz-Mari J, et al. Composite kernels for hyperspectral image classification[J]. Geoscience & Remote Sensing Letters IEEE, 2006, 3(1): 93-97.
- [14] Vaduva C, Gavat I, Datcu M. Deep learning in very high resolution remote sensing image information mining communication concept[C]. Signal Processing Conference, 2012: 2506-2510.
- [15] Mnih V, Hinton G. Learning to label aerial images from noisy data[C]. International Conference on Machine Learning, 2012.
- [16] Romero A, Gatta C, Camps-Valls G. Unsupervised deep feature extraction of hyperspectral images[C]. IEEE Workshop on Hyperspectral Image and Signal Processing, Whispers. IEEE, 2014.
- [17] Tokarczyk P, Montoya J, Schindler K. An evaluation of feature learning methods for high resolution image classification[J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, I-3: 389-394.
- [18] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]. Proceedings of the 32nd International Conference on Machine Learning (ICML-15), 2015: 448-456.

#### 附中文参考文献:

- [2] 谢永华, 杨静宇. 基于多方向二进小波变换的遥感图像特征提取方法[J]. 小型微型计算机系统, 2005, 26(6): 1054-1056.
- [3] 仇建斌, 李士进, 朱跃龙, 等. 基于曲线形状特征的快速高光谱图像波段选择[J]. 小型微型计算机系统, 2014, 35(8): 1906-1910.
- [5] 刘萌萌, 刘亚岚, 孙国庆, 等. 结合纹理特征的 SVM 样本分层土地覆盖分类[J]. 遥感技术与应用, 2014, 29(2): 315-323.
- [7] 杨国鹏, 余旭初, 陈伟, 等. 基于核 Fisher 判别分析的高光谱遥感影像分类[J]. 遥感学报, 2008, 12(4): 579-585.
- [10] 张良培, 沈焕锋. 遥感数据融合的进展与前瞻[J]. 遥感学报, 2016, 20(5): 1050-1061.
- [11] 张良培, 李家艺. 高光谱图像稀疏信息处理综述与展望[J]. 遥感学报, 2016, 20(5): 1091-1101.