

## 基于改进的ResNeXt网络结构的遥感图像分类

杨星, 宋玲玲, 王时绘

引用本文

杨星, 宋玲玲, 王时绘. 基于改进的ResNeXt网络结构的遥感图像分类[J]. 计算机科学, 2023, 50(6A): 220100158-6.

YANG Xing, SONG Lingling, WANG Shihui. Remote Sensing Image Classification Based on Improved ResNeXt Network Structure [J]. Computer Science, 2023, 50(6A): 220100158-6.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于图像数据耦合识别的输电线路安全风险评估方法](#)

Image Recognition Method of Transmission Line Safety Risk Assessment Based on Multidimensional Data Coupling

计算机科学, 2023, 50(6A): 220500032-6. <https://doi.org/10.11896/jsjcx.220500032>

### [基于动态时空神经网络的城市交通流量预测方法](#)

City Traffic Flow Prediction Method Based on Dynamic Spatio-Temporal Neural Network

计算机科学, 2023, 50(6A): 220600266-7. <https://doi.org/10.11896/jsjcx.220600266>

### [基于改进CNN-BP的多波束声纳高程数据预测研究](#)

Study on Multibeam Sonar Elevation Data Prediction Based on Improved CNN-BP

计算机科学, 2023, 50(6A): 220100161-4. <https://doi.org/10.11896/jsjcx.220100161>

### [面向交通流量预测的时空Graph-CoordAttention网络](#)

Spatial-Temporal Graph-CoordAttention Network for Traffic Forecasting

计算机科学, 2023, 50(6A): 220200042-7. <https://doi.org/10.11896/jsjcx.220200042>

### [基于多模态特征融合的时间序列异常检测](#)

Anomaly Detection of Time-series Based on Multi-modal Feature Fusion

计算机科学, 2023, 50(6A): 220700094-7. <https://doi.org/10.11896/jsjcx.220700094>

# 基于改进的 ResNeXt 网络结构的遥感图像分类

杨 星<sup>1</sup> 宋玲玲<sup>1</sup> 王时绘<sup>1,2</sup>

1 湖北大学计算机与信息工程学院 武汉 430062

2 湖北省教育信息化工程技术研究中心 武汉 430062

(yxjasson@163.com)

**摘 要** 遥感图像分类是遥感图像信息处理的关键方向之一,其分类精准率很大程度上限制了遥感技术整体的发展。对于遥感图像,传统机器学习算法与模型结构存在不能快速提取特征图,且分类结果不够准确的缺陷。针对这一问题,提出了一种改进的基于 ResNeXt 网络模型结合注意力机制,以优化后 SVM(支持向量机)算法替换全连接层的模型。首先引入计算机视觉中的注意力机制,对不同特征赋予不同的权重,提高对图像中用于分类部分有效信息的提取能力,然后结合 ResNeXt 网络,最后以优化后的 SVM 算法替换卷积神经网络末端的全连接层用于提升分类效果,同时在模型整体不增加超参数的情况下优化了网络性能。该网络模型在数据集 AID 上的实验结果表明,改进后的网络模型对深层特征的提取能力有显著提高,且优化后的网络模型对于多分类任务具有较优的分类效果。

**关键词:** 遥感图像;卷积神经网络;ResNeXt;注意力机制;场景分类;SVM

**中图法分类号** TP751

## Remote Sensing Image Classification Based on Improved ResNeXt Network Structure

YANG Xing<sup>1</sup>, SONG Lingling<sup>1</sup> and WANG Shihui<sup>1,2</sup>

1 School of Computer and Information Engineering, Hubei University, Wuhan 430062, China

2 Hubei Province Educational Information Engineering Technology Research Center, Wuhan 430062, China

**Abstract** Remote sensing image classification is one of the key directions of remote sensing image information processing, and its classification accuracy greatly limits the overall development of remote sensing technology. Traditional machine learning algorithms and model structures have the disadvantages that they cannot quickly extract feature maps from remote sensing images, and the classification results are not accurate enough. Aiming at this problem, an improved model based on the ResNeXt network model combined with the attention mechanism is proposed to replace the fully connected layer model with the optimized SVM (support vector machine) algorithm. Firstly, it introduces the attention mechanism in computer vision, assigns different weights to different features, improves the ability to extract effective information for the classification part of the image, then combines the ResNeXt network, and finally replaces the end of the convolutional neural network with the optimized SVM algorithm. The fully connected layer is used to improve the classification effect, and at the same time optimize the network performance without increasing the hyperparameters of the model as a whole. Experimental results of the network model on the data set AID show that the improved network model has a significant improvement in the ability to extract deep features, and the optimized network model has a better classification effect for multi-classification tasks.

**Keywords** Remote sensing image, CNN, ResNeXt, Attention mechanism, Scene classification, SVM

## 1 引言

遥感图像场景分类,是通过一定的算法提取出遥感图像的关键信息,从而判断该图像所属的场景种类。这一工作早期依靠有大量专业知识和丰富实践经验的专家利用遥感图像的相关特征如颜色、形状、纹理、光谱信息等来进行分类。遥感图像的场景分类一般用于辅助抢险救灾、军事行动、地理测绘以及农业生产<sup>[1]</sup>等。随着遥感图像数量的爆炸增长和图像信息的丰富,需要分类效率更好、结果更精确的方式来代替

传统的人工方法,深度学习的出现很好地解决了这个问题。过去的几十年间,世界各地的学者对遥感图像的分类问题提出了各种解决方法。按照训练时模型样本的标签信息是否参与可大致分为无监督分类、半监督分类、监督分类。下面介绍这几种分类方法在遥感图像场景分类领域的应用情况。

## 2 遥感图像场景分类研究现状

### 2.1 无监督分类

由于没有标注的遥感图像数量远大于有标签信息的遥感

基金项目:国家自然科学基金(61902114);湖北省重点实验室开放基金(2020SDSJ06)

This work was supported by the National Natural Science Foundation of China(61902114) and Hubei Provincial Key Laboratory Open Fund(2020SDSJ06).

通信作者:王时绘(wsh@hubu.edu.cn)

图像,近年来有很多无监督学习方法用于遥感图像分类。比较有代表性的包括  $K$  均值聚类 ( $K$ -means)、主成分分析法 (PCA)、稀疏编码、自编码器 (Auto Encoder, AE)、深度置信网络 (deep belief networks, DBN)。

$K$  均值聚类 ( $K$ -means) 算法的优势是比较简单,算法复杂度很低,容易理解。但是该算法在执行前必须指定一个  $K$  值,且算法的效果与  $K$  值的选取挂钩,目前尚未有一种明确的理论指导确定真实  $K$  值的方法<sup>[2]</sup>。

主成分分析法 (Principal Components Analysis, PCA) 本质上是一种基础的数学分析方法,基本思想是减少模型中重复的变量,降低耦合而尽可能保持原有的信息。通过该方法,可以保留图像信息中的主要成分<sup>[3]</sup>,忽略对分类作用不大的信息,从而起到降低噪声和维度的作用。虽然分类效果尚可,但是自动化程度不高,效率较低<sup>[4]</sup>。

稀疏编码 (Sparse Coding, SC) 的基本思想是构建一组非常完整的基向量,也称为字典,从而使得输入向量能够表示为这些基向量的线性组合<sup>[5]</sup>。基于稀疏编码的分类方法大多应用这种思想进行图像特征提取或者去噪,进而进行分类。

自编码器 (AE) 的基本模型包含了编码器和解码器,编码器把输入向量进行降维操作,把高维向量转化为低维度的数据,再通过激活函数 (也称为解码器) 把低维数据恢复成原始维度的信息,自编码器依靠编码和解码两个网络结构的训练得到最佳权重,这是一种典型的无监督分类算法<sup>[6]</sup>。Guo 等通过结合基于自编码器的离群点检索算法和基于集成学习的离群点检测算法,提出了一种以自编码器为基检测器的集成学习离群点检测算法,解决了传统方法未对基检测器进行优化选择而导致检测精度低的问题<sup>[7]</sup>。

深度置信网络 (Deep Belief Networks, DBN) 由多个受限玻尔兹曼机 (Restricted Boltzmann Machine, RBM) 组成,整体网络结构在 DBN 的最后一层设置 BP 网络,用于接受 RBM 的输出特征向量并将其作为它的输入向量,以有监督地训练实体关系分类器<sup>[8]</sup>。

## 2.2 半监督分类

近年来由于遥感图像的数量呈爆炸式增长,且大多是无标签信息,同时有标签和无标签信息的半监督算法越来越受到研究者的关注。目前大多半监督分类器是基于 SVM 算法和 GAN 算法的改进。Zhao 等把高光谱图像进行分割得到高纯度的同质分割斑块,获取训练样本后引入迁移学习,提出了一种基于同质区和迁移学习的新型半监督方法<sup>[9]</sup>。

## 2.3 监督分类

基于深度学习的遥感图像场景分类大多可以归类于全监督分类。卷积神经网络 (Convolutional Neural Network, CNN) 是深度学习的代表算法之一,它是一种前馈神经网络,经典模型结构通常由输入层、卷积层、池化层、全连接层组成。

卷积神经网络经过多年的发展,拥有了庞大体系和分类,经典的网络模型包括 AlexNet, VGGNet, GoogleNet, ResNet, MobileNet 等。这些网络模型通过训练大量的已知标签信息的数据来获得网络最佳的权重值,并通过激活函数或者其他的分类器获得分类结果,再对测试机进行预测提升泛化能力。全监督分类的缺点是需要大规模、海量的数据集进行网络训练,且对有标签信息的图像信息依赖较大。近年来,对于不同类型图像的处理问题,主流方向逐渐转移到使用多种不同

卷积神经网络模型去解决相关分类问题<sup>[10]</sup>。Li 等优化传统卷积神经网络,通过设计可变长度的个体编码策略,构建基于教与学优化 (TLBO) 的可变卷积自编码器 (CAE),并将其堆叠成一个 CNN 结构,可有效应用于医学图像分类<sup>[11]</sup>。Jiang 等将注意力机制结合 Inception-ResnetV2 网络结构,构建高度定制化的网络模型,在钢板表面缺陷的数据集中表现优异,具有很高的工程应用价值<sup>[12]</sup>。Cheng 等通过使用 CBDNet 去噪网络在图像预处理阶段提高井下图像质量,改进 ResNet50 网络模型,并提出了一种增大间距、降低拟合度的“soft-center loss”复杂损失函数,有效提高了煤矸分类及识别能力<sup>[13]</sup>。卷积神经网络的应用已涉及众多的学科和行业,对于遥感图像而言,深度学习目前被广泛应用于遥感图像领域的语义分割<sup>[14]</sup>、目标检测<sup>[15]</sup>等方向。

## 3 实验方法

### 3.1 网络结构和原理

#### 3.1.1 ResNeXt 网络模型深度特征提取

ResNeXt 是 ResNet 改进后的模型,相比原来的 ResNet 网络结构它增加了 cardinality 的数量,提高了模型分类的准确率,且可以在不增加超参数的情况下让模型更深、更宽、更有效。ResNet 和 ResNeXt 的 block 结构对比如图 1 所示。

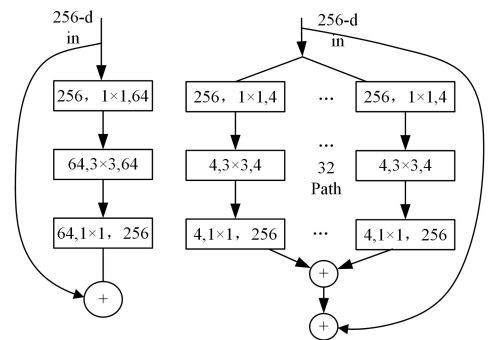


图 1 ResNet 和 ResNeXt 的 block 结构对比图

Fig. 1 Block structure comparison diagram of ResNet and ResNeXt

观察图 1 可以发现,ResNeXt 的 block 是在 ResNet 的基础上进行拓宽,借鉴了 Inception 系列的“split-transform-merge”结构多分支的思想,主动进行分组,整体结构类似,但是 ResNeXt 每个分组的结构完全一致,两者简单的模型对比如图 2 所示。

不同的是 Inception 中的不同分支的拓扑结构是不一样的,更定制化,如果需要进行调整,往往需要改动大量的超参数。而 ResNeXt 的每个分支的拓扑结构是完全一致的,再结合能解决深层网络中梯度弥散和精度下降问题的残差网络,也就是在后面添加一条 short-cut,这样就得到了 ResNeXt 网络结构,这样的结构与 Inception V4 高度相似,不同之处在于 ResNeXt 的每条分支拓扑结构完全一样,Inception V4 需要人工设计,而且 ResNeXt 是先进行卷积操作再执行单位加,Inception V4 是先进行拼接再执行卷积操作。

ResNeXt 的本质是分组卷积,与 AlexNet 受硬件条件限制而被迫进行分组卷积不同,ResNeXt 网络模型借鉴了 Inception 系列多分支的思想,并结合残差网络,并让每条分支的拓扑结构完全一致,在简化网络结构的同时提升了性能。

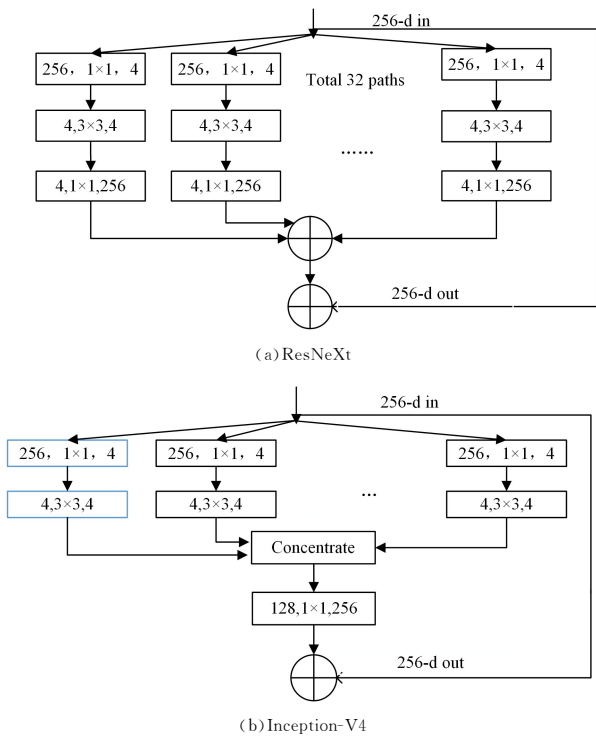


图 2 ResNeXt 与 Inception-V4 的结构对比

Fig. 2 ResNeXt and Inception-V4 structure comparison

### 3.1.2 注意力机制

遥感图像与其他种类图像有所不同,一般只有某个部分的特征对判别图像所属类别起到至关重要的作用,其他部分可能属于另一个类别,对于图像正确分类没有意义,甚至会影响分类的效果。计算机视觉中的注意力机制可以模拟人类的注意力,从冗余的信息中筛选出与当前任务相关的内容,并过滤与当前任务无关的图像信息,从而提高分类准确率<sup>[16]</sup>。本文的注意力模块(CBAM)由通道注意力和空间注意力模块串联而成,特征图先经过卷积层进行特征提取,通过通道、空间注意力模块后与输入层结果相加,然后通过分类器。

注意力模块示意图如图 3 所示。

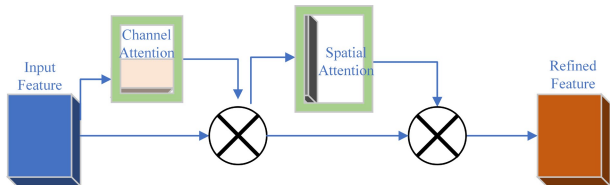


图 3 注意力模块示意图

Fig. 3 Schematic diagram of attention module

通道方向的 Attention 建模的是特征的重要性,如图 4 所示。通道注意力模型的作用方式是在正常卷积操作后,首先进行 Squeeze 操作,把空间维度进行特征压缩,使得每个二维特征图能够用一个实数表示,相当于具有全局感受野的池化操作,需要注意的是这里采用了最大池化和平均池化两种池化操作,特征通道数没有发生变化。然后进行 Excitation 操作作为每个特征通道生成权重,将得到的特征进行合并,通过归一化函数如 Sigmoid 函数对所得结果进行处理可得权重系数,将这个系数与输入的特征相乘可得缩放后的新特征,表达式如下:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \otimes F \quad (1)$$

其中,  $F$  指输入特征。

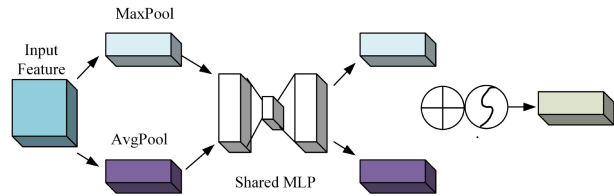


图 4 通道注意力模块结构

Fig. 4 Channel attention module structure

空间方向的 Attention 机制建模的是空间位置的重要性,用于关注有意义的特征。其模块结构如图 5 所示。输入特征图后,分别进行通道维度的平均池化和最大池化的操作,并将结果拼接在一起,然后经过卷积层,归一化的激活函数 Sigmoid 得到权重系数  $M_s$ ,再将其与输入的特征相乘可得到缩放后的新特征,表达式如下:

$$M_s(F) = \sigma(f([AvgPool(F), MaxPool(F)])) \otimes F \quad (2)$$

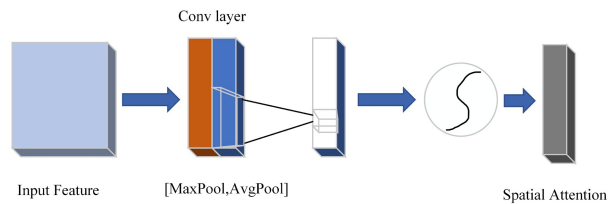


图 5 空间注意力模块结构

Fig. 5 Spatial attention module structure

注意力机制目前被广泛应用于遥感目标检测、场景分类、遥感分割<sup>[17]</sup>等任务。Qiao 等以残差网络 ResNet 为基础模型,分别在通道和空间两个维度建立注意力模块,解决了传统机器学习模型无法快速有效提取图像特征造成分类不准确的问题<sup>[18]</sup>。Li 等在卷积神经网络 ResNet, Inception-V3 的基础上通过线性融合特征得到对象整体信息,再通过注意力机制进一步提取细粒度的特征信息表达<sup>[19]</sup>。

### 3.1.3 分类器

类似于泰勒公式中用多项式函数去拟合光滑函数的思想,全连接层通过矩阵向量乘积  $y = Wx$  的卷积操作,将学到的分布式特征表示映射到样本标记空间,其本质是多项式形式的线性映射,分类能力有限,且全连接层的参数太多容易导致过拟合问题。本文使用优化后的 SVM 支持向量机算法,采用“one-against-one”方法,通过组合多个二分类器来实现多分类器的构造,从而完成遥感图像场景的多分类任务。即假设有  $M$  类遥感图像数据,每两类构建一个二分类 SVM,共应该有  $M(M-1)/2$  个二分类器,然后采用投票策略(Voting Strategy),即得票多的预测结果是对该数据的分类预测,若出现平票则选择索引较小的作为对其分类的预测结果。

## 3.2 实验设置

### 3.2.1 实验环境

本实验使用 SGD 优化器,动量 momentum 设置为 0.99,对每个参数使用相同的学习率,并且使用 Cosine 函数调整学习率,使得训练中的学习率能以余弦曲线的方式衰减从而达到收敛。实验中学习率如果太大则会出现超调现象,loss 随着迭代次数的增加没有减小的趋势;如果设置得太小则会

影响收敛的效率,大大增加训练的时长。本实验将初始学习率设置为 0.001。

表 1 实验参数配置表

Table 1 Experimental parameter configuration

实验参数	配置
操作系统	Windows10
GPU	Quadro RTX 5000 16GB
IDE	Pycharm2017
语言版本	Python3.7
框架	Pytorch1.6

### 3.2.2 数据集

为了验证模型的有效性,本文选取数据集 AID(Aerial



图 6 AID 数据集各类图像的示例

Fig. 6 Examples of various images in AID dataset

### 3.3 实验过程

实验首先把数据集中 40% 的数据划分为训练集,剩余的作为测试集用于检验模型的泛化能力,并在实验最后用不同的数据集检验模型的鲁棒性。训练轮次设置为 300 个周期,在每个训练周期内用 Adam 优化器进行反向传播训练,以自适应调整学习率。模型将输入数据提取出深层特征后,输入到用于替代全连接层的 SVM 网络结构,选取高斯核函数将数据映射到高维空间,用于使数据可以线性可分,且根据多分类的任务类型,将 SVM 以“one-against-one”方式进行调整,然后对测试集中的图像数据类别进行预测。Kappa 系数是一个用于一致性检验的指标,可用于衡量分类模型的效果。其计算是基于混淆矩阵(Confusion Matrix, CM)的,取值范围为  $[-1, 1]$ ,通常大于 0。分类准确率(Accuracy, Acc)和混淆矩阵(Confusion Matrix, CM)也常用于评价分类模型的效果。Acc 正确的样本数量占总测试集数量的百分比,其计算式如下:

$$Acc = \frac{T}{N} \times 100\% \quad (3)$$

其中,  $T$  表示分类正确的样本数量,  $N$  表示测试集的样本总数。混淆矩阵能快速观测到每个类的误分类情况,帮助分析分类器的分类效果。对角线上的元素表示分类正确的结果,第  $i$  行和第  $j$  列的数据表示第  $i$  类被误分为第  $j$  类的错误结果,通过混淆矩阵能直观地看出相近类别和容易混淆的类别之间分类的具体情况。

## 4 实验结果与分析

### 4.1 注意力模块对不同网络模型的影响

随着卷积神经网络深度的增加,网络对深层特征的提取能力会获得一定的提升,但是当达到一定层数后模型的提取

能力反而会大幅度下降。本文的特征提取模型 CBAM\_ResNeXt 除了实现 ResNet 网络系列“加深网络层数提升模型准确率”的思想外,还引入了计算机视觉中的注意力机制,可以更关注图像中的有效信息。

Image Dataset)和 NMPU-RESISC45 Dataset(NWPU45)进行实验验证。其中 AID 数据集主要用于网络模型的特征学习和验证测试,NWPU45 主要用于验证模型的泛化能力以及测试鲁棒性,两者皆是在遥感图像场景分类领域中应用得最为广泛的、大数据量的数据集。NMPU45 是由西北工业大学创建的远程传感图像分类的公共可用基准,该数据集包含 31500 张遥感图像,涵盖了机场、棒球场、海滩等共计 45 个类别的场景,每个类别 700 张遥感图像<sup>[20]</sup>。AID 数据集<sup>[21]</sup>由武汉大学和华中科技大学于 2016 年发布,包含 30 个场景类别,每个类别约有 220~420 张图像,共有 10000 张像素大小为  $600 \times 600$  的遥感图像。AID 数据集图像的每类示例如图 6 所示。

能力反而会大幅度下降。本文的特征提取模型 CBAM\_ResNeXt 除了实现 ResNet 网络系列“加深网络层数提升模型准确率”的思想外,还引入了计算机视觉中的注意力机制,可以更关注图像中的有效信息。

为验证通道注意力模块和空间注意力模块对图像识别分类的有效性,实验设计其他没有添加注意力机制的网络模型和添加注意力机制后的网络模型对 AID 数据集的分类精确率进行对比,实验结果如表 2 所列。

表 2 各模型的分类准确率

Table 2 Classification accuracy of each model

(单位:%)

网络模型	分类准确率(Acc)	
	加入注意力机制前	加入注意力机制后
VGG16	65.10	67.65
MobileNet	82.45	83.39
ResNet50	81.34	84.56

通过表 2 所列的实验结果可以看出,应用于图像分类领域常见的网络模型,通过引入注意力机制后在数据集 AID 上的平均准确率 Acc 都有了较明显的提升,说明 CBAM 模块对于图像分类具有提取有效信息的正向作用,能够提升深层卷积网络对深度特征提取的表征能力。且从实验结果来看,不同网络对于遥感图像的处理能力存在差异,可见数据集的选定对于模型的学习能力和泛化能力也有重要的影响。

### 4.2 不同网络模型特征提取准确率对比

为验证 CBAM\_ResNeXt 网络模型对于遥感图像的分类处理能力,本文在数据集 AID 上设计不同网络模型(包括已知在图像分类领域效果非常不错的 SENet, ResNet50, VGG16)的实验,300 个训练周期内分类准确率 Acc 随着训练

周期 epoch 的迭代变化情况如图 7 所示。

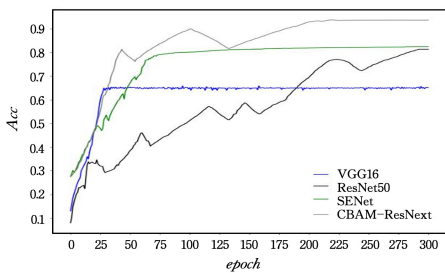


图 7 不同网络在数据集 AID 上的性能对比

Fig. 7 Performance comparison of different networks on AID data set

混淆矩阵如图 8 所示。

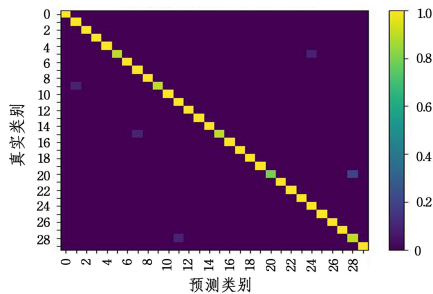


图 8 AID 数据集分类混淆矩阵

Fig. 8 AID data set classification confusion matrix

由图 8 所示的数据可以得出,本文所提出的网络模型在 AID 数据集上表现优异,且 Kappa 系数达到 0.95,说明模型的性能优越,能有效提取分类问题中的有效信息,明显提升卷积神经网络对遥感图像的分类能力。

### 4.3 不同分类器对网络模型的影响

为验证进行优化后的 SVM 的分类能力,本次实验将前置网络结构 CBAM\_ResNeXt 网络所提取的深层特征作为输入,选取应用较为广泛的几种机器学习领域的分类器,分别是 K 近邻(K-Nearest Neighbor, KNN)<sup>[22]</sup>、朴素贝叶斯(Naive Bayes)<sup>[23]</sup>、决策树(Decision Tree)和全连接层(Fully Connected),以及本文所使用的应用于多分类任务的 SVM。分别选取上述几种分类器连接前置网络模型,并测试在数据集 AID 上的分类效果,取多次分类结果的平均准确率作为最终测试结果,多个分类器的分类效果如表 3 所列。

表 3 不同分类器的对比情况

Table 3 Comparison of different classifiers

分类器	平均准确率 / %
Fully Connected	91.38
KNN	87.13
Naive Bayes	89.63
DecisionTree	85.07
SVM	93.65

### 4.4 鲁棒性测试

为测试模型的鲁棒性<sup>[24]</sup>,实验测试本文模型在另一个遥感图像数据集 NMPU45,以及在遥感图像分类领域使用较为广泛的 RSC11 DataSet, SIRI-WHU, UC Merced Land-Use DataSet, RSSCN7 Dataset 这 4 个数据集上的性能,实验进行多次,取多次准确率的平均值作为模型在该数据集下的分类准确率,结果如表 4 所列。

表 4 模型在不同数据集上的表现

Table 4 Model performance on different datasets

数据集	图像大小	图像数量	场景类别数量	平均准确率 / %
RSC11 DataSet	512 * 512	1232	11	93.11
SIRI-WHU	200 * 200	2400	12	92.87
UC Merced Land-Use DataSet	256 * 256	2100	21	93.38
RSSCN7 Dataset	400 * 400	2800	7	93.26
NMPU45	256 * 256	31500	45	91.75

通过表 4 所列的实验数据可以得出结论:通过图像数据预处理使得输入数据保持一致后,本文所提出的网络结构在其他遥感图像数据集上表现优异,仍具有较高的平均准确率,具有相当的程序鲁棒性,可以适用于多数遥感图像数据集。

### 4.5 参数复杂度和时间复杂度分析

为进一步比较不同特征提取网络的性能,这里统计上述网络模型 SENet, ResNet50, VGG16 和本文所提出的 CBAM\_ResNeXt 结构的参数量、在数据集 AID 上训练每个周期的平均用时  $T$  和常用于评价模型整体复杂度的 GFLOPs (Giga Floating-point Operations Per Second, GFLOPs),统计结果如表 5 所列。

表 5 不同网络模型的性能对比

Table 5 Performance comparison of different network models

Network Model	$T/s$	GFLOPs	参数量
SENet	12.3	4	$35.7 \times 10^6$
ResNet50	13.6	3.9	$25.6 \times 10^6$
VGG16	15.7	15.5	$138 \times 10^6$
CBAM_ResNeXt	11.8	4.2	$25.3 \times 10^6$

通过表 5 所列数据结合图 7 的平均准确率分布情况可知,本文提出的 CBAM\_ResNeXt 模型在不增加超参数的情况下相比原 ResNet50 提高了近 12%,且由于增加了 cardinality 维度,使得卷积核学到的关系更加稀疏,在整体复杂度不变的情况下降低了过拟合的风险,并且训练每个周期的时间平均仅为 11.8 s,整体参数量较少,收敛速度更快。

**结束语** 本文提出了一种基于 ResNeXt 网络模型的改进网络结构,引入计算机视觉中的注意力机制,在空间和通道两个维度建立注意力模块,用于更加有效地提取图像的深层特征,并在网络结构的末尾使用多分类任务场景下的 SVM 代替了原先连接层进行最后的分类任务,实验结果显示分类效果显著。但是这样也使得模型的复杂度大大提高,与 ResNeXt 的不增加超参数提升性能的思想有所背离,下一步应该思考如何在降低性能的情况下降低模型复杂度。

### 参考文献

- [1] HASHEMI B L, GEBREHIWOT A. Deep learning for remote sensing image classification for agriculture applications[J]. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, 44: 51-54.
- [2] LU D L, NING Q, ZANG J. Improved KNN algorithm based on BP neural network decision-making[J]. Computer Applications, 2017, 37(S2): 65-67, 88.
- [3] SPINELLA M. Improved principal component analysis and linear regression classification for face recognition[J]. Signal Processing, 2018, 145: 175-182.
- [4] FU W F, ZOU W B. Research progress of deep learning in re-

- remote sensing image classification[J]. Application Research of Computers, 2018, 326(12): 3521-3525.
- [5] BU X L, HUO H, FANG T. Rotating r-obust remote sensing image feature extraction based on sparse representation[J]. Computer Engineering, 2012, 38409(14): 124-127.
- [6] NIU X X, SUN A, WANG Y F, et al. Research on remote sensing image classification based on deep learning[J]. Laser Journal, 2021, 42(5): 10-14.
- [7] GUO Y Y, YU J, DU X S, et al. Outlier detection algorithm based on autoencoder and integrated learning [J]. Computer Applications, 2022, 42(1): 2018.
- [8] VADDI R, MANOHARAN P. Hyperspectral remote sensing image classification using combinatorial optimisation based unsupervised band selection and CNN[J]. IET Image Processing, 2020, 14(15): 3909-3919.
- [9] ZHAO C J, ZHOU S G, DING Q, et al. Semi-supervised classification of hyperspectral images based on homogeneity area and transfer learning[J]. World of Geographic Information, 2019, 26(5): 45-52.
- [10] WAN Y L, ZHONG X W, LIU H, et al. A review of applications of convolutional neural networks in hyperspectral image classification [J]. Computer Engineering and Applications, 2021, 57(4): 1-10.
- [11] LI W, FAN Y C, JIANG Q Y, et al. Medical image classification method based on variable convolutional autoencoder optimized for teaching and learning [J]. Computer Applications, 2022, 42(2): 592-598.
- [12] JIANG P Y, TAO Q C, AI M Q. Image Classification of Steel Surface Defects Based on Attention Mechanism and Deep Learning[J]. Computer Applications and Software, 2021, 38(9): 214-219.
- [13] CHENG D Q, WANG Y C, KOU Q Q, et al. Mine Image Classification Based on Improved Deep Residual Network[J]. Computer Application Research, 2021, 38(5): 1576-1580.
- [14] WANG X, ZHANG H Y, LING C. Semantic Segmentation of SAR Remote Sensing Image Based on U-Net Optimization[J]. Computer Science, 2021, 48(11A): 376-381.
- [15] YUAN X X, WU Q. Object Detection in Remote Sensing Images Based on Saliency Feature and Angle Information[J]. Computer Science, 2021, 48(4): 174-179.
- [16] LIU J W, LIU J W, LUO X L. Research progress of attention mechanism in deep learning[J]. Journal of Engineering Science, 2021, 43(11): 1499-1511.
- [17] ZHANG B, ZHANG X J, ZHAO B C, et al. Application of an improved DenseASPP network in remote sensing image segmentation[J]. Computer Applications and Software, 2021, 38(7): 46-52.
- [18] QIAO X X, SHI W Z, LIU X X, et al. Remote sensing image scene classification based on ResNet dual attention mechanism [J]. Computer System Applications, 2021, 30(8): 243-248.
- [19] LI W S, WANG Z X, LI S H, et al. Weakly supervised fine-grained image classification based on attention mechanism[J]. Computer Systems Applications, 2021, 30(10): 232-239.
- [20] 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.
- [21] XIA G S, HU J, HU F, et al. AID; A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 3965-3981.
- [22] ZHANG M L, ZHOU Z H. ML-KNN: A lazy learning approach to multi-label learning[J]. Pattern Recognition; The Journal of the Pattern Recognition Society, 2007, 40(7): 2038-2048.
- [23] NIR F, DAN G, MOISES G. Bayesian Network Classifiers. [J]. Machine Learning, 1997, 29(2/3): 131-163.
- [24] SUN H, CHEN J, LEI L, et al. Summarization of Anti-robustness Technology of Deep Convolutional Neural Network Image Recognition Model[J]. Journal of Radars, 2021, 10(4): 571-594.



**YANG Xing**, born in 1997, postgraduate. His main research interests include deep learning and remote sensing image processing.



**WANG Shihui**, born in 1965, Ph.D, professor, master supervisor, is a member of China Computer Federation. His main research interests include machine learning and software engineering.