引文格式：覃现,郝容,邢鹏,等. 基于深度学习的国土资源遥感影像分类模型[J]. 地理空间信息, 2024, 22(10): 34-37. doi:10. 3969/j.issn. 1672-4623.2024.10.009

基于深度学习的国土资源遥感影像分类模型

覃　现[1]，郝　容[2]，邢　鹏[2*]，李玉堂[2]
（1. 广西壮族自治区地理信息测绘院，广西 柳州 545006；2. 二十一世纪空间技术应用股份有限公司，北京 100096）

摘　要：针对国土资源遥感影像分类中，遥感影像分类鲁棒性以及精确性不高的问题，提出结合深度学习的国土资源遥感影像分类模型。通过改进的UNet3+深度学习网络进行国土资源遥感影像分类检测，该改进UNet3+深度学习网络在保留UNet3+深度学习网络全尺度跳跃连接以及深度监督机制的前提下，引入注意力机制对深度学习网络的参数以及特征提取做出进一步优化，避免过拟合以及运算量较大的问题。通过实验验证，基于深度学习的国土资源遥感影像分类模型在不同尺度的遥感影像观测上均表现出优异的检测精度及分类性能。其中，当观测尺度为500 m时，边界分割误差可以控制在-2~1 m范围内。相比其他深度学习网络，改进模型在分类检测精度上更具优势。

关键词：国土资源；遥感影像；改进UNet3+深度学习网络；全尺度跳跃连接；深度监督
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）10-0034-04

Remote Sensing Image Classification Model for Territorial Resource Based on Deep Learning

QIN Xian[1], HAO Rong[2], XING Peng[2], LI Yutang[2]
(1. Geographical Information Surveying and Mapping Institute of Guangxi Zhuang Autonomous Region, Liuzhou 545006, China;
2. 21st Century Space Technology Application Co., Ltd., Beijing 100096, China)

Abstract: Addressing the issues of low robustness and accuracy in the remote sensing image classification for territorial resource, we proposed a classification model based on deep learning. The model utilizes an improved UNet3+ deep learning network to classify and detect the remote sensing images for territorial resource. While preserving the full-scale skip connection and deep supervision mechanism of the original UNet3+ network, the improved UNet3+ introduces an attention mechanism to further optimize the parameters and feature extraction of deep learning network, thereby avoiding overfitting and excessive computational complexity. Practical experiments have demonstrated that the remote sensing image classification model for territorial resource based on deep learning exhibits excellent detection accuracy and classification performance across different scales of remote sensing images. Specifically, when the observation scale is 500 m, the boundary segmentation error can be controlled within the range of -2 m to 1 m. Compared with other deep learning networks, the improved model demonstrates superior performance in terms of classification and detection accuracy.

Keywords: territorial resource, remote sensing image, improved UNet3+ deep learning network, full-scale skip connection, deep supervision

近几年随着深度学习技术的发展，其在分类任务以及遥感影像方面的应用逐渐广泛[1-3]，同时也表现出了优越的性能。U-Net深度学习网络作为一种成熟的深度学习网络，已在医学影像分割任务中体现出明显优势。但是对于国土资源遥感影像的应用仍然存在一定的局限性[4]。相比之下UNet3+深度学习网络更加适用于国土资源遥感影像分类的应用。UNet3+深度学习网络通过深度监督以及全尺度跳跃连接，显著提高了边界分割的精确度以及多尺度特征的提取精度。

基于此，通过注意力机制改进UNet3+深度学习网络，以减少模型运算量，避免过拟合问题。在UNet3+深度学习网络的基础上，通过增加全尺度跳跃连接的方式保证网络对更加细粒度特征的捕捉，保证解码的精确性；在解码器层引入了深度监督机制，计算深度学习网络每个阶段的损失，以保证训练过程具有更快的收敛速度以及位置感知的精确性；同时通过注意力机制避免特征提取过程中的过拟合现象。通过对UNet3+深度学习网络的改进，从而提高遥感影像分类的检测精度。

1　基于深度学习的遥感分类框架

1.1　遥感影像分类整体框架

基于深度学习的国土资源遥感影像分类，主要通过

收稿日期：2024-04-12。
第一作者简介：覃现（1979—），本科，高级工程师，从事地理信息系统应用、遥感应用工作，E-mail：1725155607@qq.com。
通信作者：邢鹏（1977—），本科，高级工程师，从事卫星遥感、地理信息系统工程应用研究工作，E-mail：howardok@126.com。

对国土资源目标进行遥感影像的识别与定位，完成国土资源遥感影像的检测与分类。检测的整体流程如图 1 所示。

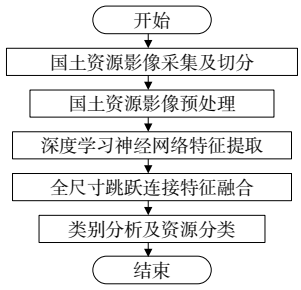


图 1　遥感影像分类检测整体流程

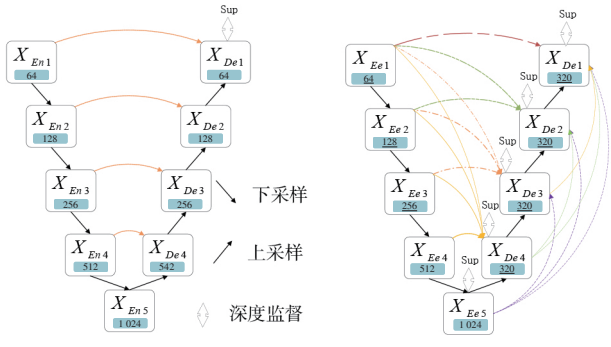
1.2　遥感影像分类检测

国土资源遥感影像的图像预处理，需要对遥感影像的样本数据进行标准化以及归一化，得到可以向深度学习神经网络中直接输入的样本数据。样本数据在基于 UNet3+改进的深度学习神经网络中，经过图像特征融合以及特征检测，得到关于国土资源的遥感影像分类结果。将分类结果与实际标签进行比对，通过误差分析进行分类正确与否的判断。通过多次迭代，得到国土资源遥感影像的最终分类结果。

2　深度学习网络及改进设计

2.1　UNet3+深度学习网络概述

UNet3+深度学习网络基于 U-Net 网络改进构建，通过全尺度跳跃连接[5-6]以及深度监督[7-8]对多尺度特征进行细化处理，保证分割边界的清晰以及位置感知的精确。UNet3+深度学习网络与 U-Net 深度学习网络的改进对比如图 2 所示。



a U-Net　　b UNet3+

图 2　深度学习网络结构对比

UNet3+深度学习网络中的全尺度跳跃连接弥补了 U-Net 深度学习网络中无法进行全尺度信息细化分析的缺陷，能够对国土资源遥感影像的位置以及边界进行细化处理。以图 2b 的解码器 X_{De3} 为例，其全尺度连接的特征图结构如图 3 所示。

图 3 中，X_{En1}、X_{En2} 为尺度较小的编码器，表示特征融合的细粒度语义。X_{En3} 为相同尺度的编码器，

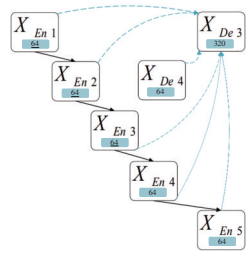


图 3　全尺度连接特征图结构

X_{En4}、X_{En5} 为尺度较大的编码器，表示特征融合的粗粒度语义。细粒度语义特征融合中，通过无重叠最大池化的方式实现下采样[9]，粗粒度语义特征融合时，通过双线性差值的方式进行上采样。通过细粒度语义以及粗粒度语义分别进行的下采样以及上采样操作，拼接融合 5 层编码器特征，通过 BN 以及 ReLU 操作，得到新的特征图。特征图 X_{Dem} 可表示为：

$$X_{\mathrm{Dem}}=\mathrm{H}\left[\underbrace{C(D(X_{\mathrm{Ena}}))_{a=1}^{m-1}}_{\mathrm{Scales:1^{th}-}m^{\mathrm{th}}}C(X_{\mathrm{Enm}}),\underbrace{C(U(X_{\mathrm{Dea}}))_{a=m+1}^{M}}_{\mathrm{Scales:}(m+1)^{\mathrm{th}}-M^{\mathrm{th}}}\right] \quad (1)$$

式中，M 为编码器总的个数；m 为编码方向上第 m 个下采样层；$H(\ )$为相应的特征聚合，其中包括一个卷积层，以及对应的 BN 和 ReLU 操作；$C(\ )$为卷积操作；$D(\ )$为上采样操作；$U(\ )$为下采样操作。

针对基于深度学习的国土资源遥感影像分类检测，全尺度跳跃连接保证了特征尺度的精细化，同时还需要通过全阶深度监督保证快速收敛以及深度监控。图 2b 中体现了在 UNet3+深度学习网络中的深度监督结构，在每一个编码器中都对应一个 sup，从而实现相应的双线性采样。通过双线性采样，可以将任意解码器进行扩展，保证每一个特征图的全分辨率。在实际的深度监督机制应用中，向每个编码器引入对应的监督损失函数，计算得到对应的损失函数之后，通过反向传播对损失函数进行更新，以加快网络收敛。同时可以表征边界信息的完整，利于国土资源遥感影像的特征检测。

结合 UNet3+深度学习网络的国土资源遥感影像分类检测模型中，通过分类图像去噪[10]避免由于背景噪声导致的过分割情况。对于任意解码器中的二维张量 X_{En5} 来说，通过卷积操作、Sigmoid 函数、Dropout 等，得到存在分割目标或者不存在风格目标对应的概率。在此过程中，进行国土资源遥感影像分类的判断，主要依据最深层二维张量实现[11-12]。同时结合上采样以及下采样操作，提高了国土资源遥感影像分类的精确性。

2.2　基于注意力机制的改进

尽管 UNet3+深度学习网络基于 U-Net 深度学习网

络改进明显，但依旧存在一定的冗余融合以及过拟合问题，同时参数量相对较多。

因此，采用CBAM模块进行UNet3+深度学习网络的改进，该模块可以在UNet3+深度学习网络中无缝衔接，提高特征图的数据特征表达，更加关注基本特征，实现国土资源遥感影像的精确监测。CBAM模块的具体实现过程如图4所示。

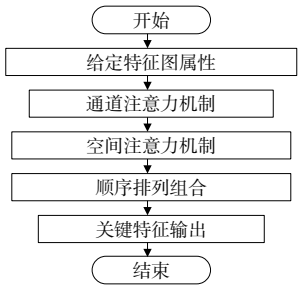


图4 CBAM模块具体实现过程

图4中，对于给定的国土资源遥感影像特征图，通过CBAM模块通过通道注意力机制[13]生成对应的一维通道图，通过空间注意力机制生成对应二维空间图，由不同维度之间的转换匹配进行顺序组合，输出对应的关键特征。

其中，对于通道注意力机制而言，往往通过双池化机制对国土资源遥感影像特征图层进行处理。即采用平均池化以及最大池化的方式得到通道注意力图层。其中，平均池化特征表示为：

$$M_{\text{avg}}(F)=\text{sig}\left(W_1\left(W_0\left(F^c_{\text{avg}}\right)\right)\right) \tag{2}$$

式中，$M_{\text{avg}}(F)$为平均池化特征；sig（）为sigmoid激活函数；W_1与W_0为多层感知机MLP的权重。同样地，最大池化特征表示为：

$$M_{\max}(F)=\text{sig}\left(W_1\left(W_0\left(F^c_{\max}\right)\right)\right) \tag{3}$$

式中，$M_{\max}(F)$为最大池化特征，据此得到相应的通道注意力为：

$$M_c(F)=M_{\text{avg}}(F)+M_{\max}(F) \tag{4}$$

对于空间注意力机制而言，同样通过双池化的方式进行通道描述，而后通过特征拼接得到卷积层，规格为7×7，结合平均池化特征以及最大池化特征得到新的特征图特征：

$$M_s(F)=\text{sig}(f^{7\times7}([F^S_{\text{avg}};F^S_{\max}])) \tag{5}$$

据此将注意力机制模块与UNet3+深度学习网络进行融合，为进一步提高运算效率，降低运算复杂度，将图2b中的UNet3+深度学习网络对称删减一个层级。同时针对UNet3+深度学习网络中可能出现的过拟合情况，将卷积层用Bottleneck layer结构代替，保证网络参数更少，并进一步避免过拟合情况的产生。

3 实验结果及分析

3.1 实验设计

针对基于深度学习的国土资源遥感影像分类实验检测，基于Python3.7编译，运行系统为Windows11，CPU频率为3.69 Hz，运行内存为32 GB，显卡为RTX3060。以AID为国土资源遥感影像数据分析数据库。遥感影像数据库中包括1万张关于国土资源的遥感影像，依据国土资源的存在形式划分，可以分为六大类即土地资源、水资源、矿物资源、生物资源、海洋资源以及其他基础设施等。各类土地资源分布均匀，遥感影像数量在1 600~1 700的范围内。其具体数值如表1所示。

表1 遥感影像分类数量统计

国土资源类型	遥感影像数量/张
土地资源	1 636
水资源	1 695
矿物资源	1 627
生物资源	1 698
海洋资源	1 650
其他基础设施	1 694

表1中所有国土资源遥感影像包含不同季节下不同比例尺的遥感影像图片，以验证国土资源遥感影像分类的普适性。其中，训练集7 000张，测试集2 000张，验证集1 000张。遥感影像采集完成后首先对其进行切片处理，裁剪后的遥感影像均为256×256。

3.2 遥感影像目标检测分析

针对遥感影像目标检测的分析，主要分析不同比例尺情况下，遥感影像分类模型对国土资源遥感影像的检测情况。实际实验中，分别对比例尺30 km、10 km、500 m的不同情况，验证遥感影像检测分类性能，不同示意图如图5所示。



图5 不同比例尺遥感影像

实验中共选取不同比例尺遥感影像各200张，通过模型检测与实际影像标签对比，结合误差分析验证遥感影像检测的准确性。对比结果如图6所示。

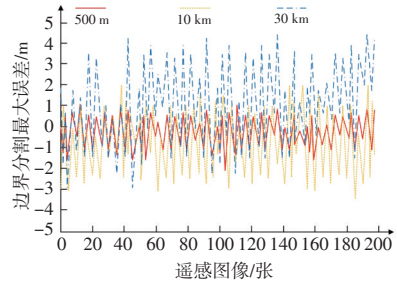


图6 遥感影像检测误差对比结果

图6中，当遥感影像比例尺为500 m时，依据基于深度学习的国土资源遥感影像检测模型与实际影像标签对比，边界分割的最大误差在−2~1 m的范围内；10 km时，边界分割的最大误差在−3~2 m的范围内；30 km时，边界分割的最大误差在−3~4.5 m范围内。随着遥感影像观测尺度的不断增大，遥感影像的检测精度略有缩减。整体来看，在不同观测尺度下，遥感影像观测误差可以控制在稳定水平，同时最大误差不超过±5 m，表明了基于深度学习的国土资源遥感影像检测的精确性以及鲁棒性。

3.3　国土资源分类精确度分析

国土资源分类的精确度实验，验证采用UNet3+深度学习网络的国土资源遥感影像分类模型对于遥感影像的细节分析以及整体性能。实验中，通过将改进的UNet3+深度学习网络与基本UNet3+深度学习网络、U−Net深度学习网络进行对比实验，验证所提国土资源遥感影像模型的精确性。精确度评价指标通过精度、召回率以及准确度[14]3个维度进行对比。3个维度的计算方式可以表示为公式（6）

$$\begin{cases} R=\dfrac{TP}{TP+FN} \\ P=\dfrac{TP}{TP+FP} \\ A=\dfrac{P_t}{M} \end{cases} \tag{6}$$

式中，R为精度；P为召回率；A为准确度；TP为将国土资源类型确定并预测为真；TN为将国土资源类型确定并预测为假；FP为将不确定国土资源类型确定为指定类型并预测为真；FN为将确定国土资源类型预测为不确定资源类型；P_t为分类检测为真的数量；M为遥感影像检测的总数。基于不同方式的检测结果对比如图7所示。

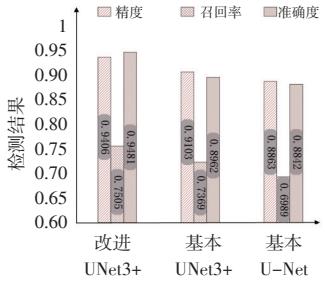


图7　不同方式遥感影像分类检测结果

图7中，改进的UNet3+深度学习网络与基本UNet3+深度学习网络以及基本U−Net深度学习网络相比，在精度、召回率以及准确度3个方面均表现优异。其中，基本UNet3+深度学习网络与改进UNet3+深度学习网络在各个维度的差别相对较小。但由于改进UNet3+深度学习网络引入了注意力机制，进一步简化了深度学习网络的参数，因此运算效率更高，整体性能更加稳定。

4　结　语

基于深度学习的国土资源遥感影像分类模型，以UNet3+深度学习网络为基础依据，通过CBMA注意力机制进行改进，进一步提升国土资源遥感影像分类的鲁棒性以及精确度。经过实验验证，国土资源遥感影像分类模型在边界分割精度以及分类准确度方面的效果提升明显。与基本UNet3+和U−Net网络相比，改进的UNet3+网络在精度、召回率和准确度3个维度上均表现更佳，同时运算参数更少。

参考文献

[1] 李森,陈海鹏,邱博. 一种融合面向对象与深度学习的地表覆盖监测成果质检技术[J]. 测绘通报,2022(增刊):174-178
[2] 刘立,董先敏,刘娟. 顾及地学特征的遥感影像语义分割模型性能评价方法[J]. 自然资源遥感,2023,35(3):80-87
[3] 徐永浩. 面向高光谱遥感影像分类的深度学习与对抗防御方法研究[J]. 武汉大学学报(信息科学版),2022,47(1):157-157
[4] 朱传海,陈学泓,陈晋,等. 基于分类后验概率空间的孪生Nested-UNet(SNU-PS)变化检测网络[J]. 遥感学报,2023,27(9):2 006-2 013
[5] 肖振久,郝明,曲海成,等. 融合注意力和扩张卷积的遥感影像道路信息提取方法[J]. 遥感信息,2024,39(1):18-25
[6] 李东旭,谢先明. 基于GAUNet的深度学习相位解缠方法[J]. 遥感信息,2023,38(3):77-84
[7] 田宪辉,谢先明,刘媛媛,等. 基于深度学习的相位解缠方法[J]. 遥感信息,2022,37(5):93-100
[8] 刘世琦,李旋,丁少鹏,等. 边缘监督与特征融合的遥感影像建筑物提取方法[J]. 测绘科学,2023,48(9):76-88
[9] 孔祥稳,王常颖,张世超,等. 改进的U-Net网络在遥感影像道路提取中的应用[J]. 遥感信息,2022,37(2):97-104
[10] 刘宇佳,谢诗哲,杜阳,等. 结合空间语义注意力的二段式遥感图像修复网络[J]. 自然资源遥感,2024,36(1):58-66
[11] 于航,谭炳香,沈明潭,等. 基于机器学习算法的机载高光谱图像优势树种识别[J]. 自然资源遥感,2024,36(1):118-127
[12] 刘晓亮,王志华,邢江河,等. 融合多源地理数据与高分辨率遥感影像的尾矿库识别与监测:以云南省个旧市为例[J]. 自然资源遥感,2024,36(1):103-109
[13] 高辰,陈芸芝,董琰,等. 加入混合池化和标签信息优化的油田地物提取模型[J]. 地球信息科学学报,2024,26(3):753-763
[14] 胡荣明,李鑫,竞霞,等. YOLOv7在探地雷达B-Scan图像解译中的应用[J]. 测绘通报,2023(8):29-33