引文格式：季民, 许建然, 张立国, 等. 高分影像盐田养殖场提取的大卷积核网络算法[J]. 地理空间信息, 2023, 21(4): 5-9. doi:10. 3969/j. issn. 1672-4623.2023.04.002

高分影像盐田养殖场提取的大卷积核网络算法

季　民[1]，许建然[1*]，张立国[2]，王春晓[3]
（1. 山东科技大学 测绘与空间信息学院，山东 青岛 266590；2. 山东省国土测绘院，山东 济南 250013；
3. 自然资源部海南基础地理信息中心，海南 海口 570203）

摘　要：提出一种基于有效感受野理论的大卷积核网络（LCKNet）的盐田养殖场提取模型。LCKNet模型基于编码器-解码器结构通过大卷积核模块（LCM）和全局感知模块（GRM）扩大有效感受野，并增强了对河流、滩涂、水库坑塘等易混地物的区分能力。在盐田养殖场数据集上进行对比测试，结果表明LCKNet能有效识别盐田养殖场与易混地物，实现了复杂区域的盐田养殖场精准自动识别。

关键词：遥感；语义分割；盐田养殖场；神经网络
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2023）04-0005-05

Large Convolution Kernel Network Algorithm for Extracting Salt Field Farms from High-resolution Images

JI Min[1], XU Jianran[1], ZHANG Liguo[2], WANG Chunxiao[3]
(1. College of Geodesyand Geomatics, Shandong University of Science and Technology, Qingdao 266590, China;
2. Shandong Provincial Land Surveying and Mapping Institute, Jinan 250013, China;
3. Hainan Basic Geographic Information Center, Ministry of Natural Resources, Haikou 570203, China)

Abstract: A salt farm extraction model based on large convolution kernel network (LCKNet) with effective receptive field theory is presented. Based on the encoder-decoder structure, the LCKNet model enlarges the effective sensing field through the large convolution core module (LCM) and global perception module (GRM) and enhances the ability to distinguish fragile objects such as rivers, beaches, reservoir pits and ponds. The results show that LCKNet can effectively identify salt farms and easily mixed land, and can accurately and automatically identify salt farms in complex areas.

Key words: remote sensing, semantic segmentation, salt farm, neural network

目前高分辨率遥感影像的盐田养殖场提取算法主要有两大类：面向对象的多尺度分割的盐田养殖场提取算法和面向像素的深度学习盐田养殖场分割算法。据此，本文提出一种基于有效感受野理论的大卷积核网络的盐田养殖场提取的语义分割模型，实现基于高分影像的盐田养殖场的有效特征信息提取和提高区分易混地物的能力。该算法通过大卷积核模块（LCM）与全局感知模块（GRM），实现了扩大卷积核的有效感受野和盐田养殖场语义特征信息的有效提取。并对2018年北方某地高分二号盐田养殖场数据集进行算法对比测试，为高精度自动化提取盐田养殖场提供参考[1-8]。

1　方法与数据集

1.1　有效感受野理论

在针对提取地物混杂的遥感对象的深度学习提取算法设计中，感受野起到确定卷积核提取遥感地物特征的覆盖范围的作用。感受野的计算方法如式（1）所示。

$$l_k = l_{k-1} + ((f_k - 1) \times \prod_{i=1}^{k-1} s_i \tag{1}$$

式中，l_k 为当前卷积层的感受野大小；l_{k-1} 为上一层的卷积层的感受野大小；f_k 为当前卷积层的卷积核大小；s_i 为当前卷积的步长。可以看出，如果要扩大卷积模块的感受野，需要连续串联多层卷积层的数量提

收稿日期：2022-11-01；修回日期：2023-02-01。
项目来源：省级自然资源监测监管大数据应用服务平台建设（2019JZZY020103）。
第一作者简介：季民（1970—），博士研究生，教授，研究方向为空间数据组织和GIS系统集，E-mail：1047499978@qq.com。
通信作者：许建然（1996—），硕士研究生，研究方向为国土资源与湿地监测，E-mail：18855057720@163.com。

高 l_{k-1} 的数值，或者采用增加卷积层的步长的方式也就是池化的形式提高 s_i 的数值。

但并非感受野中的所有像素均对输出的特征图有影响，Luo[9] 等研究了深度卷积神经网络的感受野问题，提出了有效感受野理论即感受野的有效范围服从高斯分布，且从感受野的中心快速衰减。有效感受野的计算方式如式（2）所示。

$$\sqrt{\mathrm{Var}[S_n]}=\sqrt{\frac{n(k^2-1)}{12}}=O(k\sqrt{n}) \tag{2}$$

式中，S_n 为有效感受野大小；k 为卷积核的大小；n 为卷积模块中串联的次数。从有效感受野计算可见，卷积核尺寸的增长较串联层数的增长，更有利于有效感受野的提升。而大卷积核是尺寸大于5×5的卷积核，相较于小卷积核而言（3×3，5×5）参数量更大，不适合多层串联堆叠，但根据有效感受野的计算方法，可以避开这种设计风险。

1.2 大卷积核网络语义分割算法

1.2.1 网络整体结构

本文提出了基于编码器—解码器结构的大卷积核网络语义分割算法LCKNet。LCKNet主要由编码器、全局感知模块、解码器组成，编码器主要用于提取盐田养殖场的多尺度语义信息；全局感知模块主要用于提升编码器提取的深层特征信息的上下文聚合能力和网络对全局信息感知能力；解码器主要用于将全局感知模块中提取的语义信息与编码器提取的多尺度语义信息逐步融合，生成盐田养殖场的分类结果图。

1.2.2 大卷积核模块

大卷积核指卷积核尺寸大于5的卷积。基于有效感受野理论可知，如果要扩大有效感受野，则需要扩大卷积核尺寸并减少串联的卷积层数。基于此，大卷积核模块由卷积核大小为5以上尺寸的深度卷积（逐通道卷积），串联一个卷积核尺寸为1×1的卷积（逐点卷积），参考resnet[10]的残差学习，通过并联一个用来调整通道数的逐点卷积后将首尾进行连接。但根据卷积核的参数量（PQ）计算方法，算法如式（3）所示：

$$PQ=C_{\mathrm{input}}\times C_{\mathrm{output}}\times K^2 \tag{3}$$

式中，C_{input} 为输入特征图的通道数；C_{output} 为输出的通道数；K 为卷积核的尺寸。为了降低大卷积核带来的参数增长，参考Mobilenetv3[11]中的深度可分离卷积的设计，在大卷积核所在卷积设计为逐通道卷积，利用一个逐点卷积实现通道数的扩张，能有效减少参数的增长。

为了测试针对盐田养殖场特征提取的最佳卷积核参数，本文选取交并比（IoU）、精度（P）、召回率（R）、$F1$–Score（$F1$）4个指标对不同卷积核大小进行测试。其中交并比也称重叠度，用于评价对象定位的精准度；精度为所有预测为盐田养殖场样本中正确分类的盐田养殖场样本所占比例；召回率为正确分类中的盐田养殖场的样本比例；$F1$–Score为召回率和精度综合指标，$F1$–Score较高说明方法较为有效。4种指标的计算公式如下：

$$\mathrm{IoU}=\frac{TP}{TP+FN+FP} \tag{4}$$

$$P=\frac{TP}{TP+FP} \tag{5}$$

$$R=\frac{TP}{TP+FN} \tag{6}$$

$$F1=\frac{2P\times R}{P+R} \tag{7}$$

式中，TP 为分类正确的盐田养殖场的像素数；FN 为分类错误的其他类型的像素数；FP 为分类错误盐田养殖场的像素数。

测试结果如表1所示，从测试结果可知，随着卷积核尺寸的增大，交并比、$F1$–Score逐步提高，但在增加卷积核尺寸为13后，随着尺寸的增大，各项指标在逐步降低，故认为识别盐田养殖场特征的最佳卷积核尺寸为13，后续的实验均以此作为实验参数。

表1 卷积核尺寸测试

卷积核尺寸	交并比/%	精度/%	召回率/%	$F1$–Score/%
7	83.23	91.77	89.95	90.85
9	83.90	93.42	89.17	91.25
11	84.23	89.27	93.71	91.44
13	85.05	88.63	95.47	91.92
15	83.05	90.38	91.11	90.74
17	83.16	90.10	92.11	90.40

1.2.3 全局感知模块

为了提高全局感知能力，卷积核尺寸比特征图的尺寸要大一点，故卷积核尺寸计算方式如式（8）所示：

$$CS_{GRM}=\max(H_{ftm},W_{ftm})+1 \tag{8}$$

式中，CS_{GRM} 为全局感知模块的卷积核的尺寸；H_{ftm} 为特征图的高；W_{ftm} 为特征图的宽。通过比编码器部分最后一层提取的特征图大一点的设计，可以实现对深层特征图每一个像素的感知即将特征图上的所有像素同时在一个卷积核上进行计算，达到近似于transformer[12]中的全局自注意力机制的效果。此外同大卷积核模块设计，为了减少大卷积核带来的参数量，全局感知模块采用逐通道卷积与逐点卷积串联的形式。

1.3　数据集介绍与标签制作

实验数据源为中国北方某地区的高分二号影像，空间分辨率为0.8 m，可见光光谱（RGB三通道）。借助ArcGIS软件在此遥感影像上进行目视解译并勾画出盐田养殖场的地物矢量文件，并赋予盐田养殖场和其他地物的属性值，通过矢量转栅格，生成标签文件。图1为应用数据增广方法后的效果图，增广后小样本数据集有3 800张影像，将其按7∶2∶1比例划分为训练数据集、验证数据集和测试数据集。

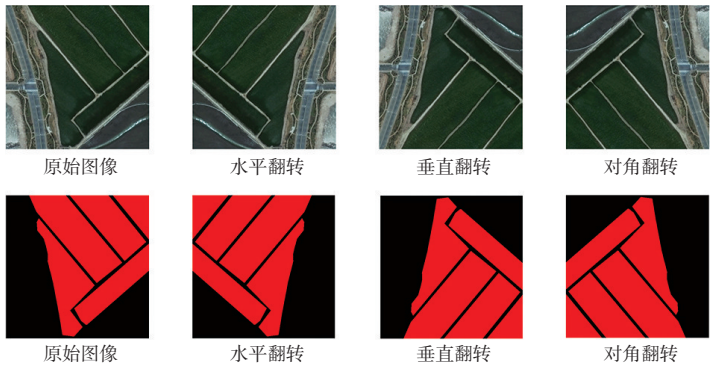


图1　数据增广效果图

2　实验与结果分析

为了证明LCKNet对盐田养殖场与易混地物（河流、坑塘、滩涂）的区分能力，选取Deeplabv3+[13]、pspnet[14]、Mobilenetv3[11]、hrnet[15]等有多个小卷积核串联设计的卷积神经网络作为易混地物提取效果的横向对比实验对象，并设计NCKNet，该网络采用三层串联的卷积核大小为3×3的小卷积核作为代替大卷积核模块并去除全局感知模块与LCKNet做纵向对比分析（消融实验）。实验以测试易混地物区分能力为目的，设计叠加分析实验作为定性分析并测试交并比、精度、召回率、$F1$–Score的数值作为定量分析。

2.1　实验参数设置

实验以tensorflow2.6.0为深度学习框架，计算机硬件为一张显存为6 G的2060显卡，设定学习率为0.000 1，跑100个纪元（epoch），运用Adam优化器，损失函数选用多分类交叉熵损失函数。为了防止过拟合，采用早停法优化训练，即连续3个训练纪元，验证样本误差不下降，则学习率减半；连续10个训练纪元，验证样本误差不下降，则停止训练。

2.2　横向对比分析

1）河流地物叠加对比分析。为了证明LCKNet在沿海遥感影像中对盐田养殖场与河流地物具有有效的区分能力，将多种深度学习语义分割方法提取的盐田养殖场与人工目视解译获得的标签的盐田养殖场做叠加分析，黄色的方框为提取错误的易混地物，蓝色部分为提取结果，红色部分为Ground Truth。

2）坑塘地物叠加对比分析。图2为盐田养殖场与

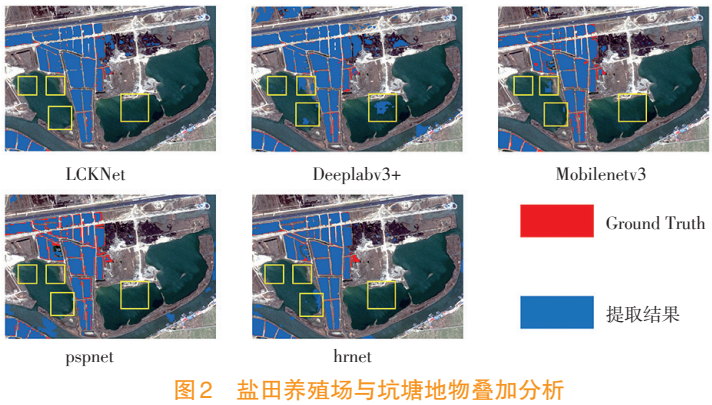


图2　盐田养殖场与坑塘地物叠加分析

坑塘地物叠加分析的结果，由图2可见，LCKNet仅有1处有少许错分地物稍强于pspnet，且错分区域与错分像素上要明显少于其余神经网络。综上，在区分坑塘地物的能力上LCKNet>pspnet>Mobilenetv3>hrnet>Deeplabv3+，说明在与坑塘这种纹理、光谱特征较为接近的地物混杂时，仍能较好区分两种地物。

3）滩涂地物叠加对比分析。图3为盐田养殖场与滩涂地物叠加分析的结果。图3表明LCKNet右边边缘处有少许错分像素，强于Deeplabv3+、pspnet，远好于Mobilenetv3、hrnet。说明LCKNet在盐田养殖场与沿海滩涂地物地理空间位置相近时，错分像素较少。

4）测试集横向定量分析，结果如表2所示。

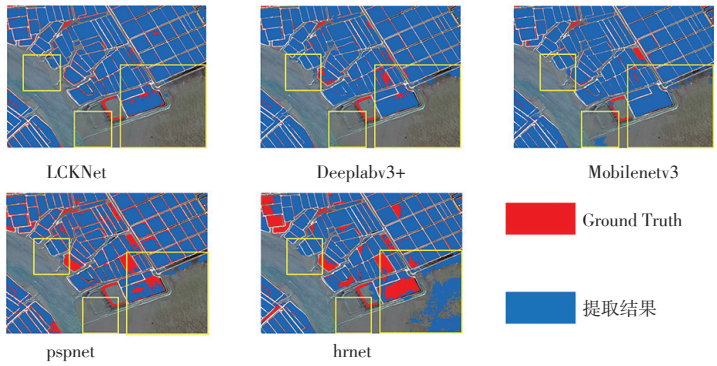


图3　盐田养殖场与滩涂地物叠加分析

表2　测试集横向定量分析

方法	交并比/%	精度/%	召回率/%	*F*1-score/%	训练时间/min
LCKNet	85.05	88.63	95.47	91.92	140.2
Deeplabv3+	78.88	89.55	86.87	88.19	305.7
Mobilenetv3	83.25	89.84	91.91	90.86	282.2
pspnet	73.48	78.47	92.04	84.72	328.6
hrnet	65.67	86.09	73.47	79.28	219.3

2.3　纵向对比分析（消融实验）

LCKNet通过LCM、GRM两个模块扩大了有效感受野，提升了有效特征的感知能力，为了证明LCM、GRM两个模块确对网络产生了正面的影响并增强了提取效果，设计了消融实验。如图4所示为LCKNet、NCKNet及添加了LCM模块的NCKNet-LCM和添加了GRM模块的NCKNet-GRM的纵向对比实验。实验一至三行依次为河流、坑塘、滩涂区域的提取结果，图中黄色方框为对盐田养殖场提取效果的局部对比。对比第一列与第二列可知，在同时去除LCM、GRM模块后，方框1、3、4、6、7区域出现严重的漏分，方框2、5区域边缘分割精度效果较差。对比第二三列可得，向NCKNet中加入LCM模块后，方框1、4、7区域的漏分情况得到有效的改善，但方框

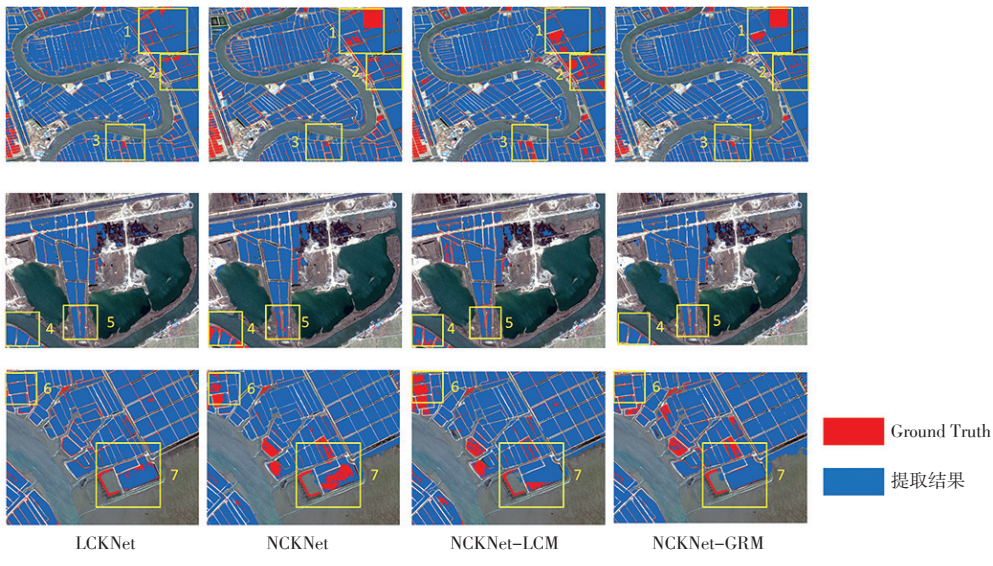


图4　消融实验叠加分析

2、3区域漏分情况依然较严重。对比第二至四列可知，向NCKNet中加入GRM模块后，方框2、3、5、6、7区域的漏分问题得到改善，但在方框1区域的漏分情况较严重。

从表3可见，LCKNet整体指标优于未加入LCM、GRM的NCKNet。

表3　消融实验定量分析

方法	交并比/%	精度/%	召回率/%	$F1$-score/%
LCKNet	85.05	88.63	95.47	91.92
NCKNet	79.27	83.82	83.59	84.01
NCKNet-LCM	83.37	89.51	92.40	90.93
NCKNet-GRM	84.26	89.07	91.71	90.37

2.4　实验结果分析

测试集横向定性定量分析说明了在盐田养殖场地物提取方面LCKNet相较于目前主流的语义分割网络，具有更好的遥感地物识别定位精度和更有效的语义分割性能。纵向对比（消融实验）分析证明了本文提出的LCM、GRM两个基于大卷积核与有效感受野理论设计的模块确实可以提升神经网络对盐田养殖场地物提取的性能，并证实了扩大有效感受野可以有效地增强盐田养殖场地物的提取效果。

3　结　语

本文针对海岸带的盐田养殖场与河流、坑塘、滩涂等沿海遥感地物自动解译易错分问题，提出了基于大卷积核和有效感受野理论的大卷积核网络语义分割算法LCKNet。该网络区别于目前主流的语义分割网络在于利用少数串联的大卷积核层有效地抑制了多层卷积层串联带来的有效感受野的快速衰减，增强了对遥感地物有效特征信息的感知与提取能力，实现盐田养殖场遥感地物识别提取性能的提升。通过定量与定性的对比实验和消融实验，表明本文提出的LCKNet在河流、坑塘、滩涂地物与盐田养殖场混杂区域具有良好的识别性能，错分在易混地物区域的像素较少，提出的LCM、GRM模块可以有效提高模型的提取能力，减少了海岸带盐田养殖场的错分误分。

参考文献

[1] 杨九新. 高台盐湖发展盐田卤水养殖初探[J]. 甘肃科技，2021,37(23):48-50

[2] 侯学会,李新华,徐洪彪,等. 基于时序特征的黄河三角洲土地覆盖动态及驱动力分析[J]. 水资源与水工程学报,2021,32(1):233-240

[3] 金明煜. 基于高分影像的盐田信息提取方法研究[D]. 北京:中国地质大学(北京),2018

[4] 蔚亮,任春颖,贾明明,等. 1986年以来6个时期杭州湾及其周边区域滨海养殖用地的分布及其变化[J]. 湿地科学，2021,19(5):534-542

[5] 陈磊士,高霞霞,廖玉芳,等. 基于CART决策树的高分二号洞庭湖区湿地分类方法[J]. 测绘通报,2021(6):12-15

[6] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2015:3 431-3 440

[7] 曾梦婷. 基于卷积神经网络的农业遥感图像语义分割及其应用研究[D]. 合肥:安徽农业大学,2021

[8] 邓泓,杨滢婷,刘兆朋. 基于深度学习的无人机水田图像语义分割方法[J]. 中国农机化学报,2021,42(10):165-172

[9] Luo W,Li Y,Urtasun R,et al. Understanding the Effective Receptive Field in Deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems,2016,29(2):4 898-4 906

[10] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770-778

[11] Howard A, Sandler M, Chu G, et al. Searching for Mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision,2019:1 314-1 324

[12] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[C]//International Conference on Learning Representations,2021

[13] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV),2018:801-818

[14] Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2017:2 881-2 890

[15] Sun K, Xiao B, Liu D, et al. Deep High-resolution Representation Learning for Human Pose Estimation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019:5 693-5 703