基于编码-解码卷积神经网络的遥感图像语义分割

张哲晗[1,2], 方薇[1*], 杜丽丽[1], 乔延利[1], 张冬英[1], 丁国绅[1,2]

[1]中国科学院安徽光学精密机械研究所中国科学院通用光学定标与表征技术重点实验室, 安徽 合肥 230031;

[2]中国科学技术大学, 安徽 合肥 230026

摘要 农村地区遥感图像语义分割是进行城乡规划、植被以及农用地检测的基础。农村地区高分辨率遥感图像含有较为复杂的地物信息,对其进行语义分割难度较大。基于此,提出一种改进的对称编码-解码网络结构 SegProNet,利用池化索引与卷积融合语义信息及图像特征,通过 1×1 卷积构建 Bottleneck 层进一步提取细节、减少参数量,逐步加深过滤器深度以构建端到端的语义分割网络,改进激活函数进一步提升网络性能。实验结果表明,在 CCF 卫星数据集上,所提方法及经典语义分割网络 U-Net、SegNet 的准确率分别为 98.4%,80.3%,98.1%,所提方法较其他方法更优。

关键词 图像处理; 农用地检测; 遥感图像; 语义分割; 编码-解码网络; 深度学习

中图分类号 O436 **文献标志码** A **doi**: 10.3788/AOS202040.0310001

Semantic Segmentation of Remote Sensing Image Based on Encoder-Decoder Convolutional Neural Network

Zhang Zhehan[1,2], Fang Wei[1*], Du Lili[1], Qiao Yanli[1], Zhang Dongying[1], Ding Guoshen[1,2]

[1]*Key Laboratory of Optical Calibration and Characterization, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, Anhui* 230031, *China*;

[2]*University of Science and Technology of China, Hefei, Anhui* 230026, *China*

Abstract The remote sensing image semantic segmentation in rural areas is the basis for urban and rural planning, vegetation and agricultural land detection. Segmentation of a high-resolution remote sensing image of rural areas is difficult because of the complex image information. Herein, we designed a complete symmetric network structure that includes a pooled index and a convolution used to fuse semantic information and image features. The Bottleneck layer is constructed using 1×1 convolution and employed to extract the details and reduce the parameter quantity, deepen the filter depth to build an end-to-end semantic segmentation network, and improve the activation function to further enhance network performance. The experimental results show that the accuracies of the proposed method and the classical semantic segmentation networks U-Net and SegNet are 98.4%, 80.3%, and 98.1%, respectively on the CCF dataset. Thus, the proposed method achieves better performance than the other two methods.

Key words image processing; agricultural land detection; remote sensing images; semantic segmentation; encoder-decoder network; deep learning

OCIS codes 100.6890; 100.3010; 100.3008

1 引 言

农村地区遥感图像处理在农乡规划、农业生产、天气预测、农用地防护以及地震、泥石流类应急减灾等方面都有广泛应用[1-2]。农村地区遥感图像语义分割是基于像素级别将图像中的特定类别地物进行划分并标注,从而便于后续农村地区遥感图像的分析处理,为进一步进行农村地区遥感图像识别、量测以及分析等提供了基础[3]。

光照条件、图像采集角度和高度的不同导致遥感图像中语义分割场景存在差异,一定程度上加大了工作的难度[4]。如何对遥感图像进行准确分割一

收稿日期: 2019-09-25; **修回日期**: 2019-10-11; **录用日期**: 2019-10-21

基金项目: 国家自然科学基金(41601379)

* **E-mail**: fwei@aiofm.ac.cn

直是研究热点，目前主要的语义分割算法有全卷积网络（FCN）[5-7]、SegNet[8]、U-Net[9]等。王宏胜等[10]利用一种改进的种子区域生长算法对农村地区房屋进行提取，基于小区域种子提高了算法的抗噪能力，但需要人为进行多次阈值干预。欧攀等[11]将空间变换网络融入超快区域卷积神经网络，实现了对遥感图像中地物目标更高精度的检测，但稳健性较差，检测效果易受遥感图像背景复杂度影响，且无法实现像素空间信息定位。Kampffmeyer 等[12]分别通过基于块的模型、像素级到像素级模型及模型融合等方法对城市遥感图像中土地占比问题进行训练，但对细节地物分割精度较低。胡勇等[13]利用双边滤波器滤掉农村地区无人机遥感影像中的噪声，通过移动均值算法分割对农村地区房屋信息进行提取，但数据集量较少，且信息提取的自动化、智能化不高。杨建宇等[14]选取基于深度卷积神经网络的 SegNet 图像语义分割算法对遥感影像中的农村建筑物进行提取，但传统的 SegNet 方法分割精度较低，且训练耗时较长。

上述算法存在稳健度较低、分割精度不够理想、训练时效性较差，以及边界细节分割效果不明显的问题。首先，本文针对数据集量较小且标签较弱的问题，采用 Open CV 对数据进行增强，使小样本数据集得到扩充[15]；然后，建立基于编码器-解码器的网络框架 SegProNet，将上采样与池化索引结合，对农村地区遥感图像进行语义分割；最后，通过与经典语义分割网络 U-Net、SegNet 进行对比分析，验证了本文所设计网络在复杂地物类型中能实现高精度的分类效果和较好的分类速度。

2　网络设计

作为深度神经网络的一种，卷积神经网络（CNN）被广泛用于图像分类工作上。CNN 通过逐层提取图像数据中的特征实现从原始图像到新特征空间的映射，并进行分类，但在特征提取的过程中丢失了图像中用于识别目标位置和形态等的语义信息[16-18]。考虑农村地区遥感图像具有分辨率高、分割目标尺度不一且边缘细节复杂的特点，本文基于 SegNet 框架设计了一种端到端的编码器-解码器网络结构 SegProNet，在保留图像像素的空间语义信息的同时进行地物识别分割。

与 SegNet 相比，SegProNet 仅在前 3 层编码结构中进行最大池化并对应进行索引操作，从而达到像素级定位标准，最后 2 层去除池化仅作卷积处理，在保证了空间分辨率的同时通过对边界轮廓信息进行模糊感知实现了像素分割，进一步提升网络感知细节信息的能力，保证了像素定位预测的准确性。输入原始图像经过池化索引与上采样产生稀疏特征图，再与卷积结合生成稠密特征，通过选择性丢弃池化降低了空间信息的损耗。同时，避免了原始网络中由重复上采样导致的信息丢失，在保证语义和位置信息的同时平衡了感受野数量，使得边界划分得到改进，并大大减少了端到端训练的参数量，从而保证网络能够学习到较抽象的特征表示。本文将原始 SegNet 卷积层数由 26 层扩充至 27 层，加入 Bottleneck，并分别引入 ReLU、ELU 激活函数增加网络的非线性，使得网络可以表达更复杂的特征，并一定程度上降低了参数量。SegProNet 与 U-Net、SegNet 的网络结构对比如图 1 所示。

通过去除池化、相对深化卷积，使得网络在提取更丰富的细节图像特征的同时保证了空间语义特征的准确性，并针对去除池化导致的参数量上升，设置 Bottleneck 层，在拓展网络深度的同时提升训练效率。SegProNet 网络结构如图 2 所示。

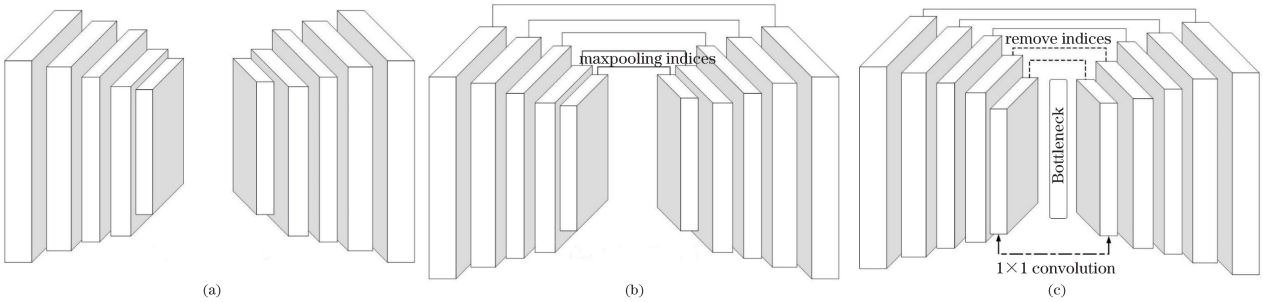


图 1　网络结构对比。(a) U-Net；(b) SegNet；(c) SegProNet

Fig. 1　Architecture comparison. (a) U-Net; (b) SegNet; (c) SegProNet

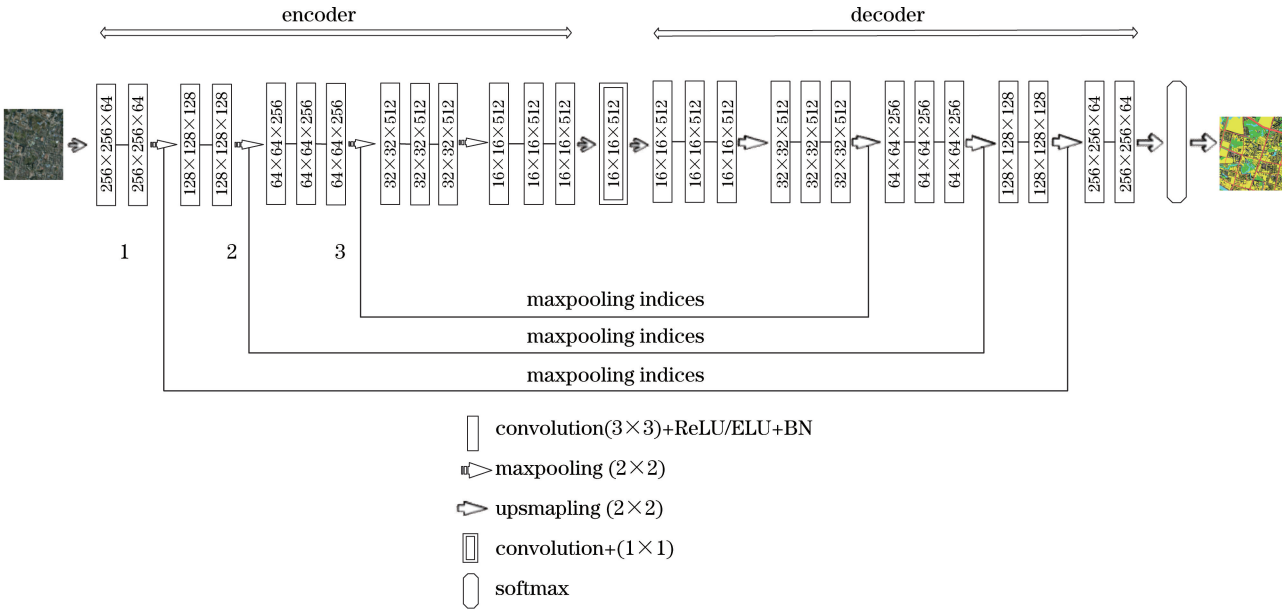


图 2　SegProNet 遥感图像语义分割网络结构

Fig. 2　Remote sensing image semantic segmentation network structure of SegProNet

网络包含：5 层编码结构，由 13 个 3×3 卷积核、5 个批归一化层(BN)、4 个 2×2 窗口、步长为 2 的最大池化(maxpooling)层组成；5 层解码结构，与编码器完全对称，由 13 个 3×3 卷积核、5 个批归一化层、4 个上采样层组成。编码结构与解码结构的连接处设置 1×1 卷积层，进一步对编码后的图像数据进行非线性特征提取，在加深网络深度的同时减少了参数量。

2.1　编码器结构

将图像规范到 256 pixel×256 pixel 大小作为网络的输入，通过卷积层和池化层组成的编码器对图像特征进行提取。SegProNet 利用 same 卷积保证图像长宽尺寸信息大小不变，通过最大池化索引(maxpooling indices)传递给解码器进行输入特征图的非线性上采样，以获取在编码过程中丢失的图像空间语义信息。在每一层最大池化后加入批归一化层，非线性函数映射后特征值的取值区间趋近饱和区，从而对输出特征值的分布进行标准化处理，避免了梯度消失。批归一化主要包括归一化和转换重构。

1) 归一化

网络中输入样本数据相关参数归一化的表达式为

$$\mu_{\mathrm{B}}=\frac{1}{m}\sum_{i=1}^{m}x_i,\tag{1}$$

$$\sigma_{\mathrm{B}}^2=\frac{1}{m}\sum_{i=1}^{m}(x_i-\mu_{\mathrm{B}})^2,\tag{2}$$

$$\bar{x}_i=\frac{x_i-\mu_{\mathrm{B}}}{\sqrt{\sigma_{\mathrm{B}}^2+\varepsilon}},\tag{3}$$

式中：m 为输入的批尺寸；μ_{B} 为输入的均值；σ_{B}^2 为偏方差；ε 为维持参数稳定所设常量；x_i 为输入的第 i 个样本值；$\bar{x}_i$ 为输入的第 i 个样本值对应的归一化值。经过这一步骤后，网络所学到的特征参数发生变化，则需经过重构恢复特征分布。

2) 特征分布重构

特征分布重构的表达式为

$$y_i=\gamma\bar{x}_i+\beta,\tag{4}$$

式中：γ 和 β 分别为权重值和偏置量。批归一化层能有效防止模型训练中的过拟合，并加快学习速度。因此，本文并没有采用 L2 正则和 Dropout 来对训练过拟合进行控制。

与 SegNet 不同，SegProNet 分别采用修正线性单元(ReLU)和指数线性单元(ELU)分析不同激活函数对网络语义分割性能的影响。ReLU 是一种常用的激活函数，它将所有负值变为 0，正值保留输出，使得神经元被稀疏激活。ReLU 具有更强的特征挖掘能力，利用 ReLU 对训练数据进行拟合，并进一步防止梯度消失。ReLU 函数式为

$$y=\begin{cases}0, & x\leqslant 0\\ x, & x>0\end{cases},\tag{5}$$

式中：x 为神经元输入；y 为神经元输出。

ELU 也属于修正类激活函数的一种，其为负值输入添加了一个非零输出。与 ReLU 不同的是，ELU 激活函数包括负指数项，函数表达式为

$$y=\begin{cases}a[\exp(x)-1], & x<0\\ x, & x\geqslant 0\end{cases},\tag{6}$$

式中：a 为常数参量定值。

2.2 解码器结构

解码器由上采样层、卷积层、池化层组成，上采样可恢复最大池化丢失的部分信息，并根据池化索引恢复像素位置信息。在解码器部分同样使用了 same 卷积。上采样操作过程如图 3 所示。

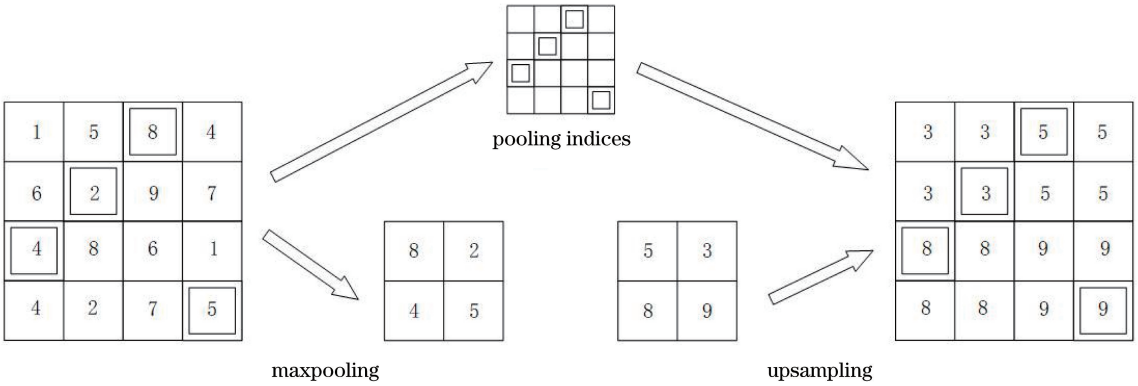


图 3　最大池化索引与上采样

Fig. 3　Maxpooling indices and upsampling

解码器使用每个 maxpooling 层存储的索引来对相应特征图进行上采样操作。为了将粗细纹理结合并防止参数冗余，在编码-解码结构连接处设置 Bottleneck 层。通过引入 1×1 卷积核层对编码器输出单元进行降维特征学习。编码器输出参数量可表示为

$$X=\sum_{i=1}^{n}F_{l_i}^2\cdot F_{c_i}\cdot h,\tag{7}$$

$$X'=h\cdot F+\sum_{i=1}^{n}F_{l_i}^2\cdot F_{c_i}\cdot F,\tag{8}$$

式中：X 为编码器输出参数量；X' 为加入 1×1 卷积核层后的输出参数量；F_{l_i} 为第 i 层卷积核大小；F_{c_i} 为第 i 层卷积核个数（过滤器深度）；l_i 为第 i 层神经元；c_i 为第 i 个卷积核；h 为输入特征图深度；F 为1×1 卷积核个数。一般情况下，$F<h$，因此加入 1×1 卷积核后计算的参数量减少，对经过编码后的每个像素在不同通道上进行线性组合，既保留了经过编码后的图像原有结构，也使得网络被拓深拓宽。

将在输出层前一层的全连接层替换为卷积层，使得网络前向传播的效率更高，将最后一层卷积输出的特征总和输入到 softmax 分类器，分类器最终输出 5 个概率，分别对应预测该样本属于每个类别的概率，即分类标签数量为 5。预测的分割对应于在每个像素处具有最大概率的类。

3　网络训练

实验分别训练了网络 U-Net、SegNet 与 SegProNet，并进行对比。作为一种全卷积神经网络，U-Net 通过结合下采样各层信息和上采样的输入信息来还原细节信息，并且逐步还原图像精度。与其不同的是，SegProNet 将池化索引和卷积相结合，并设置 1×1 卷积层加深网络，保证了像素定位的准确性，结合深层网络卷积层提高像素识别精确度，得到了更丰富的感受野，对植被、水体等地物的特征提取能力增强，更适应农村地区的遥感图像语义分割工作。

3.1 数据源与预处理

训练数据来自 2015 年中国南方某农村地区的高分辨率无人机遥感影像“CCF 数据集”，其分辨率为亚米级，光谱为可见光波段（RGB），将提供的训练样本分为 5 类：植被（标记 1）、建筑（标记 2）、水体（标记 3）、道路（标记 4）以及其他（标记 0），耕地、林地、草地均归为植被类。语义信息标注表如表 1 所示，其中 R、G、B 分别表示红、绿、蓝。

表 1　语义信息标注表

Table 1　Information of labels

Feature category	No.	Label color (R,G,B)
Background	0	(0,0,0)
Vegetation	1	(50,205,50)
Building	2	(245, 254, 0)
Water	3	(0, 255, 255)
Road	4	(255, 92, 75)

数据集为 5 幅大尺寸遥感高分辨率图像，尺寸过大且不统一，不可直接将其作为神经网络的输入，过小的切割步长可能导致子图标签分布不均衡影响训练准确度，因此将图像及其标注图像切

割成为 256×256 的子图。源数据集采用人工标注方式，标签真实值与地表真实值存在一定误差，标签较弱。本文采用扩充数据集的方式，通过 OpenCV 对切割后的原图和标注图像进行数据集扩张处理，主要有旋转（沿 y 轴镜像翻转）、加噪（加入椒盐噪声、高斯噪声）、光照调整（改变图像亮度、色彩饱和度、图像对比度），以及模糊等操作，最终得到 100000 张 256×256 的子图，通过样本扩充有效地解决了弱标签数据集问题。部分数据集展示如图 4 和图 5 所示。

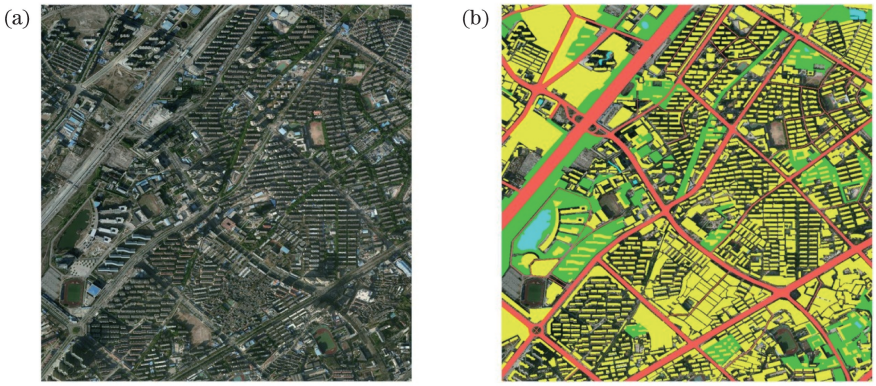


图 4　数据集展示。(a)训练集图像；(b)对应标签可视化图像

Fig. 4　Display of data sets. (a) Training set image; (b) corresponding label visualization image

图 5　部分训练集图像示例

Fig. 5　Some training set images

3.2　参数设置及实验平台

实验将数据集按 5∶1 比例进行划分，即 75000 张图像作为训练集，25000 张图像作为验证集。学习率为控制网络的学习速度的参量。实验中首先根据经验值设置学习率为 0.001[19]，此时模型收敛慢，同时出现过拟合；改进学习率为 0.05 后，训练曲线振荡较大，且模型无法达到最优值。经过多次调参，最终选取合适学习率为 0.01，批尺寸设置为

10，实验迭代次数设置为 30，优化器设置为随机梯度下降（SGD），SGD 随机选取某个样本进行梯度更新，按

$$W' = W - \alpha \nabla J_t(W), \tag{9}$$

更新权重。式中：W'为更新后权重；W 为原始权重；α 为学习率；$J_t(W)$为第 t 个样本损失。

设置卷积核个数由 64 逐步增加到 512，提升了网络预测性能。同时在激活函数的设置上，本文分别采用 ReLU 和 ELU 对网络进行训练，并对比其最终语义分割效果。为了保证对比验证结果的准确性，所有网络均采用统一的损失函数，即

$$l = -\sum_j \sum_N y_{N_j} \cdot \lg(y^*_{N_j}), \tag{10}$$

式中：$y^*_{N_j}$ 表示网络对样本 j 所属类别的预测；y_{N_j} 表示样本 j 所属真实类别。

实验在 Windows 7/64 操作系统上进行，采用 TensorFlow 1.5.0＋Keras2.14 深度学习模型，Python3.6.1 编辑器以及 Cuda7.0＋Cudnn9.0，并用 Visio studio2015＋OpenCV 进行数据处理。本次实验的硬件平台采用 NVIDA GeForce Quadro M4000 8 GB GPU 进行加速。

3.3　评价指标

IoU 指标为常见的模型预测指标，它反映了目标区域和真实源标签的交互比，计算公式为

$$I = \frac{\sum t_{\mathrm{p}}}{\sum f_{\mathrm{n}} + \sum t_{\mathrm{p}} + \sum f_{\mathrm{p}}}, \tag{11}$$

式中：t_{p}为预测结果中属于该类的实际也为该类的像元个数；f_{n}为预测属于该类而实际为其他类别的像元个数；f_{p}为预测结果中属于其他类别而实际为此种类别的像元个数。

一般情况下，IoU 值超过 0.5，则认为该网络具有较好图像分割性能。

精确率可表征预测结果精确率，代表模型中预测出的属于该类的像元中实际属于该类的像元所占比例，计算公式为

$$p = \frac{\sum t_{\mathrm{p}}}{\sum t_{\mathrm{p}} + \sum f_{\mathrm{p}}}。 \tag{12}$$

精确率又叫查准率，也是语义分割等工作中常见评价指标。除此之外，本文还采用召回率对不同网络的农村地区遥感图像语义分割效果作出了评估，计算公式为

$$r = \frac{\sum t_{\mathrm{p}}}{\sum t_{\mathrm{p}} + \sum f_{\mathrm{n}}}。 \tag{13}$$

召回率代表预测结果中属于该类的实际也为该类的像元个数占实际该类像元数的比值，也称查全率，反映了模型对数据集中所有正样本的分类性能。

4　实验结果分析

采用几种经典的语义分割网络对预测图像进行语义分割，并对比分割效果，根据数据语义类别进行不同颜色的标注。实验分别比较了 SegProNet 与 SegNet、U-Net 的训练耗时，如表 2 所示。

表 2　各网络训练及预测耗时对比

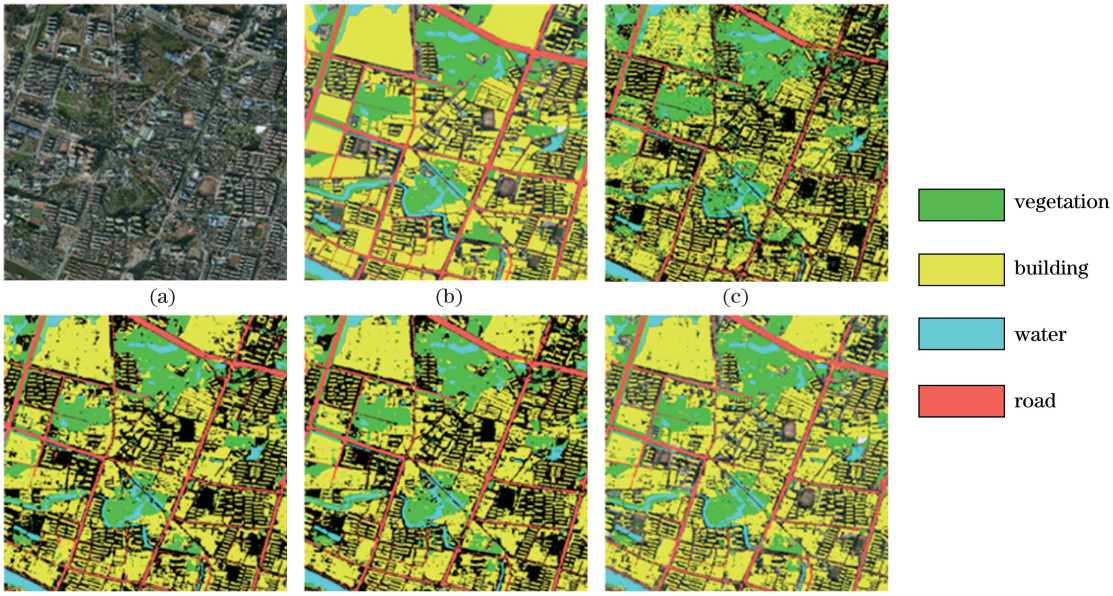
Table 2　Comparison of training and prediction time of each network

Network category	Training time /h	Prediction time /h
U-Net	16.3	0.26
SegNet	27.6	0.61
SegProNet	23.2	0.52
SegProNet＋ELU	22.6	0.51

从训练结果来看：网络 U-Net 收敛速度最快，耗时 16.3 h；SegNet 训练耗时为 27.6 h；SegProNet 训练耗时为 23.2 h。设计的网络 SegProNet 在结构上进一步精简：利用批归一化层防止梯度爆炸，利用小批量上的均值和标准差调整神经网络中间输出，使得整个神经网络各层中间输出的数值更稳定，进而加速收敛；同时，采用 Bottleneck 层减少编码器提取的特征图参数量，在保证分割精度的同时取得了相对更好的分割效果，在内存和计算时间上保证了效率，且易于训练；而 ELU 激活函数存在负值，可以将激活单元的输出均值往零值推近，减少偏移效应进而使梯度接近于自然梯度，相较于 ReLU 能使网络更快拟合。

针对中国某南方农村地区遥感图像，对不同方法所得语义分割结果进行可视化，结果如图 6 所示。语义分割结果细节展示如图 7 所示。

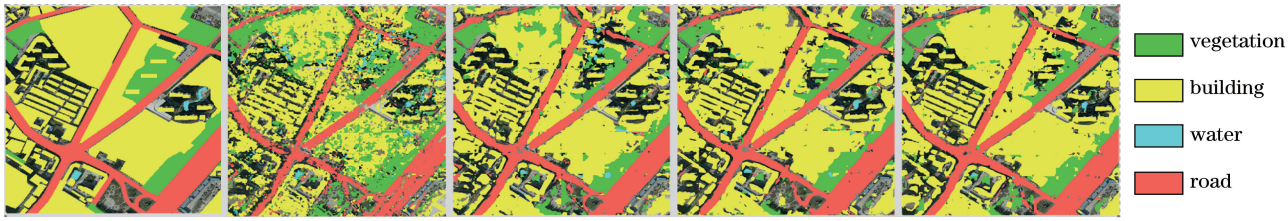
由图 6 和图 7 得出，改进后方法的分割结果将错分降到最小，和真实值最接近，且分割效果更直观，整体视觉感知最优。各分割方法所得评价指标具体数据如表 3 所示。



(a)　(b)　(c)　(d)　(e)　(f)

图 6　实验结果对比。(a)原始图像；(b)原始标签可视化结果图；(c) U-Net 分割结果；(d) SegNet 分割结果；(e) SegProNet＋ReLU 分割结果；(f) SegProNet＋ELU 分割结果

Fig. 6 Comparison of experimental results. (a) Original image; (b) original label visualization result; (c) U-Net segmentation result; (d) SegNet segmentation result; (e) segmentation result of SegProNet＋ReLU; (f) segmentation result of SegProNet＋ELU



(a)　(b)　(c)　(d)　(e)

图 7　实验结果细节对比。(a)原始标签可视化结果图；(b) U-Net 分割结果；(c) SegNet 分割结果；(d) SegProNet＋ReLU 分割结果；(e) SegProNet＋ELU 分割结果

Fig. 7 Comparison of experimental details. (a) Original label visualization result; (b) U-Net segmentation result; (c) SegNet segmentation result; (d) segmentation result of SegProNet＋ReLU; (e) segmentation result of SegProNet＋ELU

表 3　各方法指标评价

Table 3　Evaluation indicators of each method

Category	Evaluation	Vegetation	Building	Water	Road
U-Net	Precision	0.7565	0.5232	0.7368	0.6904
	Recall	0.6341	0.7847	0.8237	0.7571
	IoU	0.5266	0.4575	0.6364	0.5652
SegNet	Precision	0.8340	0.7254	0.8555	0.7576
	Recall	0.8164	0.8338	0.8541	0.8722
	IoU	0.7023	0.6339	0.7464	0.6819
SegProNet	Precision	0.8533	0.7508	0.8837	0.7815
	Recall	0.8239	0.8587	0.8653	0.8736
	IoU	0.7216	0.6682	0.7768	0.7021
SegProNet＋ELU	Precision	0.8531	0.8861	0.8792	0.8664
	Recall	0.8534	0.7528	0.8914	0.8164
	IoU	0.7441	0.6863	0.7942	0.7251

由表 3 可以看出，使用本文改进的网络，在数据集上对植被、道路、建筑、水体四种地物进行分割，取得的效果最优。SegProNet 选择性丢弃池化，减少下采样对数据参量的压缩，结合部分池化索引保留图像中的精确位置信息，采用 Bottleneck 层实现多个特征图的线性组合，提高了网络特征提取能力，同时，激活函数 ELU 的效果要好于 ReLU，因为 ELU 中的负指数项能够防止静默神经元出现，提高学习效率。

实验的损失和精度曲线如图 8 所示。

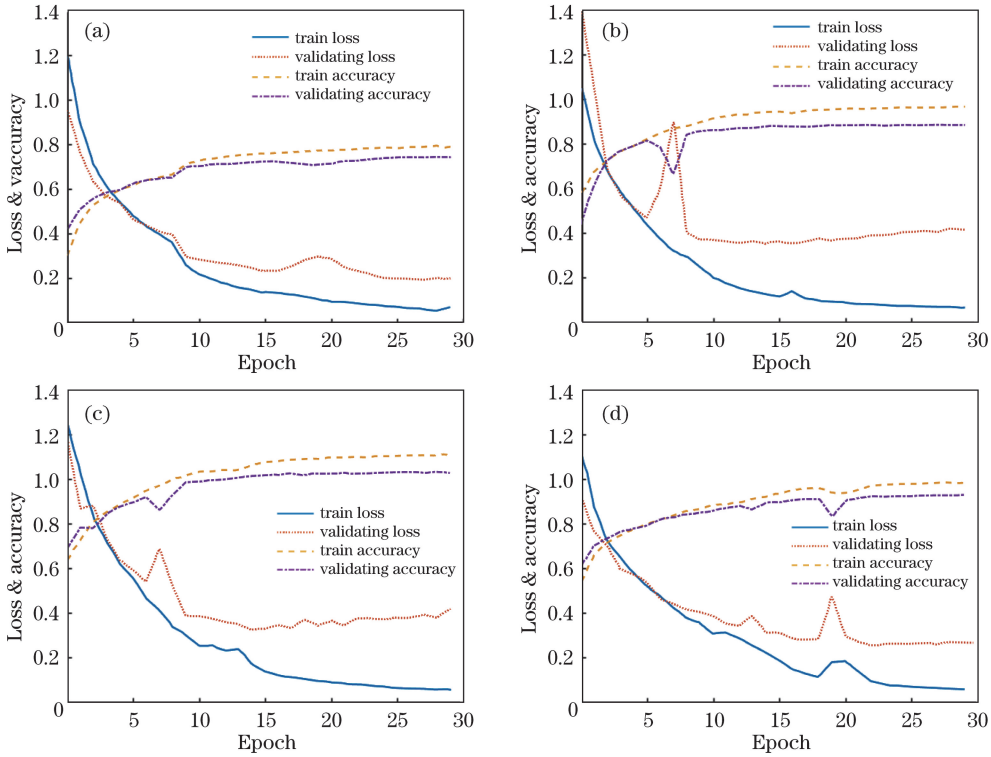


图 8　各网络的损失和精度曲线。(a) U-Net；(b) SegNet；(c) SegProNet；(d) SegProNet+ELU

Fig. 8　Each network loss and accuracy curves. (a) U-Net; (b) SegNet; (c) SegProNet; (d) SegProNet+ELU

通常情况下，损耗越低、精度越高，说明网络的性能越好。从图 8 可看出，结合训练振荡程度、网络分割效率、收敛速度等因素，网络 SegProNet 综合表现最优。通过池化索引实现像素定位并在深层网络卷积层中提取特征，减少细节信息丢失，同时引入 1×1 卷积核实现跨通道聚合，进一步降维，从而减少参数量，并使网络的训练达到 NIN(network-in-network)的效果，使得网络模型更加稳定，训练振荡频率降低，使得像素分割更加精确。

5　结　　论

改进设计了一种编码-解码卷积神经网络 SegProNet，通过实验验证了其相较于经典语义分割网络具有更好的分割效果及更快的收敛速度。采用池化索引和上采样保证了像素定位，结合深层卷积、丢弃池化产生稠密特征，对丢失的高频信息和像素位置信息进行了补充，逐步加深过滤器深度使网络的语义分割时效性增强。同时，与经典语义分割网络相比，本文网络的结构更为精简，收敛时间相对较短，具有良好的学习效率及实时性，实现了农村遥感图像中路网、水体、山地、植被以及人造建筑等地物提取的定位，对于明晰集体土地产权、监测农村居民用地、土地覆盖信息提取、农用地测算和乡间道路提取等方面具有研究意义，推动了农村可持续发展。实验结果表明：1）与 SegNet、U-Net 相比，网络 SegProNet 在 CCF 数据集上具有更好的语义分割性能；2）激活函数 ELU 的软饱和性对输入噪声具有更好的鲁棒性；3）加深网络深度可能会因数据集量不足而发生过拟合；4）优化器 SGD 在一次迭代后更新一次样本梯度，且没有冗余，较其他优化器能使网络更快收敛。

在后续研究中，拟考虑加入空洞卷积及模型融合集成学习方式进一步优化结构，实现更加精确的农村地区遥感图像语义分割。

参 考 文 献

[1] Wu Z C, Hu Z W, Zhang Q, et al. On combining spectral, textural and shape features for remote sensing image segmentation[J]. Acta Geodaetica et Cartographica Sinica, 2013, 42(1): 44-50.
巫兆聪, 胡忠文, 张谦, 等. 结合光谱、纹理与形状结构信息的遥感影像分割方法[J]. 测绘学报, 2013, 42(1): 44-50.

[2] Xu Y Y, Hu S, Wang B Y. Present status of unmanned aerial vehicles remote sensing for forest resources monitoring in China[J]. Forestry and Environmental Science, 2017, 33(1): 97-101.
徐誉远, 胡爽, 王本洋. 无人机遥感在我国森林资源监测中的应用动态[J]. 林业与环境科学, 2017, 33(1): 97-101.

[3] Rao Y H, Wu L X, Zou Z L, et al. A preliminary study on rural land use and remote sensing dynamic monitoring technology[J]. Bulletin of Agricultural Science and Technology, 2008(2): 56-57.
饶月辉, 吴连喜, 邹自力, 等. 农村土地利用与遥感动态监测技术初探[J]. 农业科技通讯, 2008(2): 56-57.

[4] Guo X Y, He D J. Application of artificial neural networks in village land use classification[J]. Journal of Agricultural Mechanization Research, 2011, 33(1): 190-194.
郭小英, 何东健. 人工神经网络在农村土地利用分类中的应用[J]. 农机化研究, 2011, 33(1): 190-194.

[5] Wu Z H, Gao Y M, Li L, et al. Fully convolutional network method of semantic segmentation of class imbalance remote sensing images[J]. Acta Optica Sinica, 2019, 39(4): 0428004.
吴止锷, 高永明, 李磊, 等. 类别非均衡遥感图像语义分割的全卷积网络方法[J]. 光学学报, 2019, 39(4): 0428004.

[6] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.

[7] Fu G, Liu C J, Zhou R, et al. Classification for high resolution remote sensing imagery using a fully convolutional network[J]. Remote Sensing, 2017, 9(5): 498.

[8] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.

[9] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.

[10] Wang H S, Li Y S, Xie J L. Rural house extraction based on an improved seeded region growing[J]. Geomatics & Spatial Information Technology, 2017, 40(10): 40-42, 47.
王宏胜, 李永树, 谢嘉丽. 一种基于改进种子区域生长的农村房屋提取[J]. 测绘与空间地理信息, 2017, 40(10): 40-42, 47.

[11] Ou P, Zhang Z, Lu K, et al. Object detection in of remote sensing images based on convolutional neural networks[J]. Laser & Optoelectronics Progress, 2019, 56(5): 051002.
欧攀, 张正, 路奎, 等. 基于卷积神经网络的遥感图像目标检测[J]. 激光与光电子学进展, 2019, 56(5): 051002.

[12] Kampffmeyer M, Salberg A B, Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 26-July 1, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 680-688.

[13] Hu Y, Zhang X C, Ma Z Z, et al. Rural residential area extraction from UAV remote sensing imagery [J]. Remote Sensing for Land & Resources, 2016, 28(3): 96-101.
胡勇, 张孝成, 马泽忠, 等. 无人机遥感影像中农村房屋信息快速提取[J]. 国土资源遥感, 2016, 28(3): 96-101.

[14] Yang J Y, Zhou Z X, Du Z R, et al. Rural construction land extraction from high spatial resolution remote sensing image based on SegNet semantic segmentation model[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(5): 251-258.
杨建宇, 周振旭, 杜贞容, 等. 基于SegNet语义模型的高分辨率遥感影像农村建设用地提取[J]. 农业工程学报, 2019, 35(5): 251-258.

[15] Kaiser P, Wegner J D, Lucchi A, et al. Learning aerial image segmentation from online maps[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(11): 6054-6068.

[16] Zhang X N, Zhong X, Zhu R F, et al. Scene classification of remote sensing images based on integrated convolutional neural networks[J]. Acta Optica Sinica, 2018, 38(11): 1128001.
张晓男, 钟兴, 朱瑞飞, 等. 基于集成卷积神经网络的遥感影像场景分类[J]. 光学学报, 2018, 38(11): 1128001.

[17] Tan G H, Hou J, Han Y P, et al. Low-parameter real-time image segmentation algorithm based on convolutional neural network[J]. Laser & Optoelectronics Progress, 2019, 56(9): 091003.
谭光鸿, 侯进, 韩雁鹏, 等. 基于卷积神经网络的低参数量实时图像分割算法[J]. 激光与光电子学进展, 2019, 56(9): 091003.

[18] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching[J]. Laser & Optoelectronics Progress, 2018, 55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. 激光与光电子学进展, 2018, 55(8): 081005.

[19] Kaoudi Z, Quiané-Ruiz J A, Thirumuruganathan S, et al. A cost-based optimizer for gradient descent optimization[C]//Proceedings of the 2017 ACM International Conference on Management of Data, May 14-19, 2017, Chicago, Illinois, USA. New York: ACM, 2017: 977-992.