

引文格式: 王雪,隋立春,钟棉卿,等.全卷积神经网络用于遥感影像水体提取[J].测绘通报,2018(6):41-45.DOI:10.13474/j.cnki.11-2246.2018.0173.

# 全卷积神经网络用于遥感影像水体提取

王 雪<sup>1</sup>,隋立春<sup>1,2</sup>,钟棉卿<sup>1</sup>,李顶萌<sup>3</sup>,党丽丽<sup>4</sup>

(1. 长安大学,陕西 西安 710054; 2. 地理国情监测国家测绘地理信息局工程技术研究中心,陕西 西安 710054; 3. 西安国际港务区规划局,陕西 西安 710026; 4. 咸阳职业技术学院,陕西 咸阳 712000)

**摘要:**提出了一种全卷积神经网络模型用于遥感影像的水体目标提取的方法,介绍了全卷积神经网络的基本原理及构建 3 种网络模型的过程。为了分析模型效果,首先搜集带有水体的影像数据,并将这些数据标注为水体和背景两类,然后利用构建的全卷积神经网络模型进行推理和学习获取先验模型,最后对测试影像进行水体提取试验。通过与传统的基于影像光谱特征的阈值法和基于图论的 GrabCut 算法提取结果进行比较,验证了本文方法的可行性。

**关键词:**遥感影像;水体提取;全卷积神经网络;阈值法;GrabCut 算法

中图分类号:P237 文献标识码:A 文章编号:0494-0911(2018)06-0041-05

## Fully Convolution Neural Networks for Water Extraction of Remote Sensing Images

WANG Xue<sup>1</sup>,SUI Lichun<sup>1,2</sup>,ZHONG Mianqing<sup>1</sup>,LI Dingmeng<sup>3</sup>,DANG Lili<sup>4</sup>

(1. College of Geology Engineering and Geomatics,Chang'an University,Xi'an 710054,China;  
2. Engineering Research Center,Geographical Conditions Monitoring National Administration of Surveying,Mapping and Geoinformation,Xi'an 710054,China;  
3. Planning Bureau,Xi'an International Trade and Logistics Park,Xi'an 710026,China; 4. Xianyang Vocational Technical College,Xianyang 712000,China)

**Abstract:** This paper presented fully convolution neural networks for extracting water from remote sensing images.The paper introduced the basic principle and method of three kinds of models of fully convolution neural networks.For extracting the water target,firstly,the image data were gathered and labeled two classes——water and background.By inference and learning in fully convolution neural networks,the trained models were obtained.Finally,water target of the test images was extracted.The feasibility of the proposed models was validated by comparing the result of extraction with that of the traditional thresholding method based on image spectral features and GrabCut algorithm based on graph theory.

**Key words:** remote sensing images;water extraction;fully convolution neural networks;thresholding method; GrabCut

随着遥感技术的发展,利用遥感影像提取水体信息是进行水资源调查和监测的重要手段和方法。水体提取的方法较多;基于影像光谱特征的提取方法,如毕海芸等<sup>[1]</sup>利用的单波段阈值法、基于阈值的多波段谱间关系法、基于阈值的水体指数法等;基于影像纹理特征的提取方法,如李小涛等<sup>[2]</sup>利用变差函数来描述影像的纹理特征,根据水体的纹理特征对水体进行提取;还有一类综合考虑影像的光谱特征、空间特征和纹理特征的面向对象的提取方法,如殷亚秋等<sup>[3]</sup>把影像分割成不同的对象,利用影像的光谱信息和空间信息,在对象层次上进行水体的提取。

近年来,随着人工智能、大数据、机器学习技术的发展,一些新型目标提取算法和技术不断出现,基

于全卷积神经网络(FCN)的遥感影像水体的提取就是其中之一。卷积神经网络(CNN)<sup>[4]</sup>的优点在于其多层结构能自动学习的特征,并且可以学习到多个层次的特征;较浅的卷积神经网络感知域较小,可以学习到一些局部区域的特征,较深的卷积层具有较大的感知域,能够学习到更加抽象的特征。目前,卷积神经网络不仅在全图式的分类上有所提高<sup>[5-7]</sup>,也在局部任务上,如目标检测上取得了进步<sup>[8-11]</sup>。

本文提出的全卷积神经网络算法用于水体目标提取,可以从抽象的特征中恢复出每个像素所属的类别,使从图像级别的分类进一步延伸到像素级别的分类。因此,本文提出的算法相对于传统的目标提取算法有更高的可信度和稳健性。

收稿日期: 2017-10-11; 修回日期: 2018-03-27  
基金项目: 国家自然科学基金(41372330); 国家自然科学基金青年基金(41601345)  
作者简介: 王 雪(1983—),女,博士生,讲师,主要研究方向为遥感数字图像处理与影像分析。E-mail:wangxue-2001@ 163.com

1 全卷积神经网络的基本原理

卷积神经网络是包含卷积层的深度神经网络模型。卷积神经网络模型较多,经典的模型有 2012 年 ImageNet 比赛冠军的模型 AlexNet<sup>[5]</sup>、2014 年 ImageNet 竞赛的双雄 VGG<sup>[6]</sup> 模型和 GoogleNet 模型<sup>[7]</sup>。这些典型的 CNN 模型中,开始几层都是卷积层 (CONV) 和池化层 (pool, 即下采样) 的交替,然后在最后一些层 (靠近输出层) 则是全连接 (FC) 的一维网络,这样将卷积层产生的特征图就映射成一个固定长度的特征向量。这些典型 CNN 结构适合于图像级的分类和回归任务。由于最后得到的是整个输入图像的一个数值描述的概率值,它会将原来二

维的矩阵 (图像) 压扁成一维的,从而使后半段没有空间信息。那么,如何从抽象的特征中恢复出每个像素的所属类别,即从图像级别的分类进一步延伸到像素级别的分类,Jonathan Long 等提出了全卷积神经网络 (FCN)<sup>[12]</sup>。

FCN 模型将传统的 CNN 模型的全连接层转化成一个卷积层,这样所有的层都是卷积层,故称为全卷积网络。FCN 的构建包括卷积化、上采样 (反卷积) 和跳层结构。

1.1 卷积化

卷积化即将传统的 CNN 全连接层转化成一个卷积层,输出低分辨率的分割图像。如图 1 所示,以 AlexNet 模型为例。

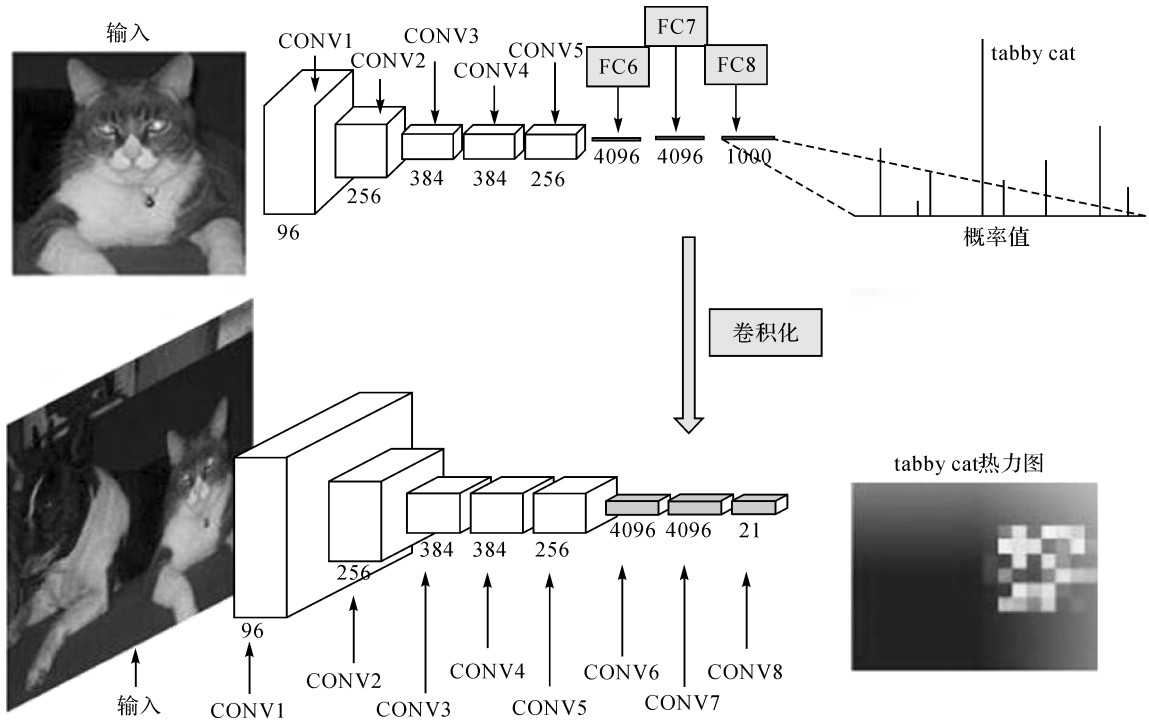


图 1 卷积神经网络构建全卷积神经网络示意图

图 1 中, CONV1—CONV5 为卷积层, 同时每层下采样, 图像尺寸每层降 2 倍, 全连接层 FC6、FC7、FC8 分别进行了 1×1 的卷积, 卷积化后的核尺寸 (通道数、宽、高) 分别变为 (4096, 1, 1)、(4096, 1, 1) 和 (类别数 N, 1, 1)。

在卷积神经网络中, 卷积计算需要矩阵化<sup>[13]</sup>, 若卷积参数输入为 4×4 图像, 输出为 2×2 图像, 卷积核尺寸为 3×3, 边界填充补零为 0, 步长为 1。则

矩阵乘就是将输入展开为 16 维向量  $\mathbf{x}$ , 输出展开为 4 维向量  $\mathbf{y}$ , 则卷积计算公式如下

$$\mathbf{y} = \mathbf{C}\mathbf{x} \tag{1}$$

变换后的稀疏阵  $\mathbf{C}$  由卷积核  $\mathbf{C}'$  变换而来,  $\mathbf{C}$  和  $\mathbf{C}'$  的表达式如下

$$\mathbf{C}' = \begin{bmatrix} w_{0,0} & w_{0,1} & w_{0,2} \\ w_{1,0} & w_{1,2} & w_{1,3} \\ w_{2,0} & w_{2,2} & w_{2,3} \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} w_{0,0} & w_{0,1} & w_{0,2} & 0 & w_{1,0} & w_{1,1} & w_{1,2} & 0 & w_{2,0} & w_{2,1} & w_{2,2} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{0,0} & w_{0,1} & w_{0,2} & 0 & w_{1,0} & w_{1,1} & w_{1,2} & 0 & w_{2,0} & w_{2,1} & w_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{0,0} & w_{0,1} & w_{0,2} & 0 & w_{1,0} & w_{1,1} & w_{1,2} & 0 & w_{2,0} & w_{2,1} & w_{2,2} & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{0,0} & w_{0,1} & w_{0,2} & 0 & w_{1,0} & w_{1,1} & w_{1,2} & 0 & w_{2,0} & w_{2,1} & w_{2,2} \end{bmatrix}$$

通过式(1)即可以进行卷积计算。

1.2 上采样(反卷积)

由于卷积化后输出的为低分辨率的图像,前五卷积层降采样下降 32 倍,为了解决低分辨率的问

题,将低分辨率图像进行上采样,输出同分辨率的分割图像。上采样的方法选用双线性插值反卷积<sup>[14]</sup>计算。图 2 为上采样示意图。

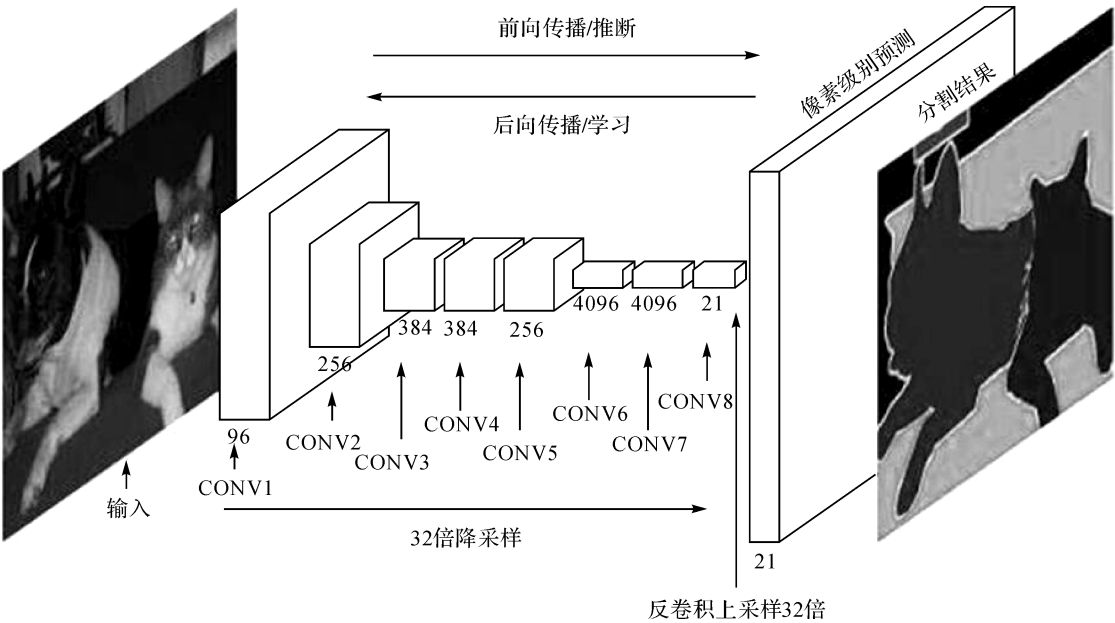


图 2 上采样示意图

与卷积化一样,反卷积操作也需要矩阵化,若卷积参数输入为 2×2 图像,输出为 4×4 图像,卷积核尺寸为 3×3,边界填充补零为 2,步长为 1,则矩阵乘就是输出展开为 4 维的向量  $y$ ,输入展开为 16 维的向量  $x$ ,反卷积计算公式如下

$$x=C^Ty$$
 (2)

1.3 跳层结构

由于从 32 倍下采样的图像直接上采样插值到原始图像大小,信息量损失较大,结果较粗糙,只能表示出对象的大致形状。为了解决这个问题,本文提出跳层结构,以精化分割提取的图像。

3 种模型的构建如图 3 所示。第 1 种模型是 FCN-32s,是在第 7 卷积层 (CONV7) 后直接采用上采样 32 倍的结果;第 2 种模型是 FCN-16s,是在第 4 池化层 (pool4) 后增一个 1×1 的卷积层,同时把第 7 层上采样 2 倍,再将二者融合。融合方法是直接进行矩阵对应位置相加,然后对融合的结果 Sum(1) 进行 16 倍的上采样,获得 FCN-16 模型;第 3 种模型是 FCN-8s,是在第 3 池化层 (pool3) 后增加一个 1×1 的卷积层,同时对 Sum(1) 进行 2 倍上采样,二者再进行融合得到 Sum(2),对 Sum(2) 再进行 8 倍上采样获得 FCN-8s 模型。

2 全卷积神经网络的模型构建和试验数据

本文选用 VGG16 为基本网络模型来构建 FCN 模型,使用 VGG16 作为初始网络,保留参数;舍弃最后的 2 个全连接层换为全卷积层,同时追加一个预测卷积层,按照跳层结构的原理,分别构建出 FCN-32s 网络模型、FCN-16s 网络模型和 FCN-8s 网络模型。

本文的试验数据包括训练数据和测试数据。训练数据是从遥感影像中裁剪出带有水体的 187 幅约为 500×500 像素大小的影像,分辨率约为 4 m,为了增加样本量,进行 90°、180°、270° 旋转,总共获取 748 幅影像;同时对这 748 幅影像的水体进行标注,除水体外其他标注为背景,共两类目标,作为标签数据;然后对 748 幅影像数据和标签数据的 70% 作为训练集,30% 作为验证集,引入 VGG16 构建的 FCN-32s、FCN-16s 和 FCN-8s 网络模型,进行推理和学习,通过参数微调,获取 3 种先验模型,迭代次数均为 50 000 次。

测试数据是在遥感地图上裁剪的 2 幅图像,这两幅图像与训练数据来自不同区域。一幅背景较简单,水体形状比较规则;另外一幅背景比较复杂,水体形状不规则,如图 4(a) 所示。

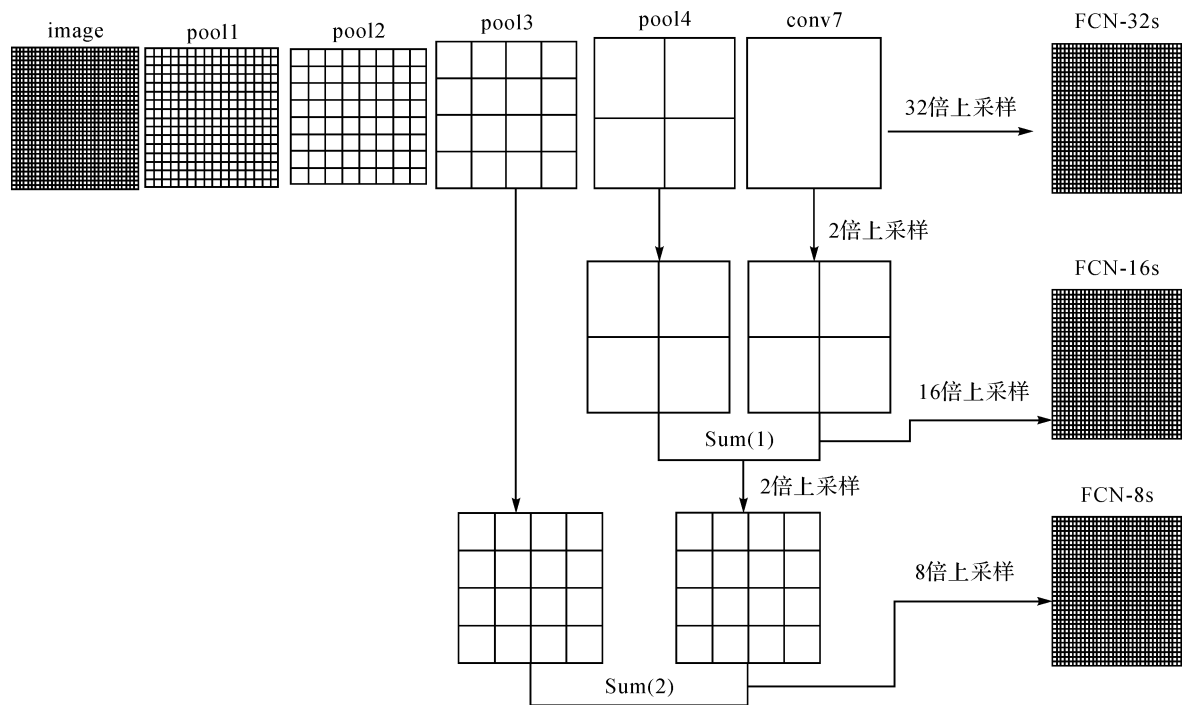


图 3 FCN 3 种模型的构建

3 试验结果及分析

全卷积神经网络由前文直接获取的 3 种先验模型进行测试;阈值法的阈值通过随机在水体区域选取分布均匀的 15 个点,确定出阈值的范围进行水体提取;Grabcut<sup>[15]</sup>通过人机交互界面粗略地画出前景和背景线,来自动构建高斯混合模型,进行水体提取。几种方法提取结果如图 4 所示。

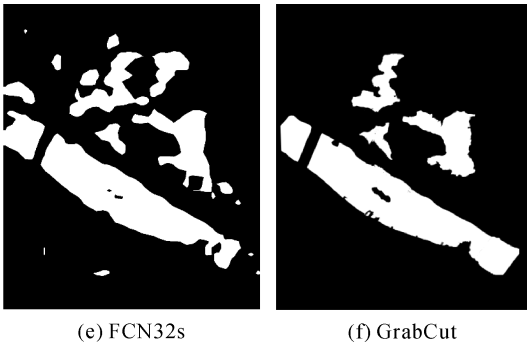


图 4 几种方法试验结果

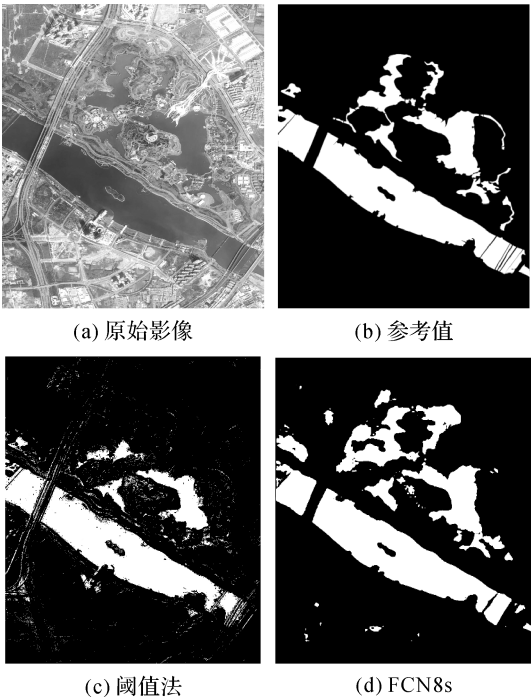


图 4 所列为试验结果,其中图 4(a)为原始影像,图 4(b)为人工判读的结果,作为理想的水体区域参考值,图 4(c)为阈值法结果,虽然能提取出水体,速度比较快,方法简单,但是噪声比较多,边缘也不是特别明显。另外不同的影像阈值是不固定的,每幅影像要重新选择阈值,通用性差。图 4(d)和(e)均为全卷积神经网络提取水体,选择 FCN8s 和 FCN32s 两种模型的测试结果,可以看出 FCN8s 提取的边界更清晰,提取得更精细一些,而 FCN32s 提取结果虽然位置正确,但是比较粗略。图 4(f)为 GrabCut 提取结果,GrabCut 对背景简单的提取效果比对背景复杂的要好,但是在提取时需要人工交互选取前景和背景,工作量大,泛化能力较差,精度也不是很高,漏掉的比较多,对于水体提取,该方法还需要改进。因此,全卷积神经网络尤其是 FCN8s 模



型提取水体自动化程度高,效果较好。

4 结论与展望

本文提出了一种全卷积神经网络模型用于水体目标提取的方法,通过 3 种模型 FCN32s、FCN16s 和 FCN8s,对遥感影像上的水体目标进行分割提取试验,通过与传统的阈值法和基于图论的 GrabCut 算法提取水体对比,全卷积神经网络的 FCN8s 模型提取结果稳健性较强,泛化能力较好。

随着人工智能和大数据的崛起,若能利用遥感影像的高光谱数据的特征,并增加基础地理信息的面状和线状地图信息,构成大数据进行训练,笔者相信深度卷积神经网络将在测绘遥感领域发挥出更大的优势。

参考文献:

[1] 毕海芸,王思远,曾江源,等.基于 TM 影像的几种常用水体提取方法的比较和分析[J].遥感信息,2012,27(5):77-82.

[2] 李小涛,黄诗峰,郭怀轩.基于纹理特征的 SPOT5 影像水体提取方法研究[J].人民黄河,2010,32(12):5-6.

[3] 殷亚秋,李家国,余涛,等.基于高分辨率遥感影像的面向对象水体提取方法研究[J].测绘通报,2015(1):81-85.

[4] 吴岸城.神经网络与深度学习[M].北京:电子工业出版社,2016.

[5] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks [C] // International Conference on Neural Information Processing Systems. Nevada: Curran Associates Inc,2012:1097-1105.

[6] SIMONYAN K,ZISSERMAN A.Very Deep Convolutional Networks for Large-scale Image Recognition [J]. Computer Science,2014(9):1409-1456.

[7] SZEGEDY C, LIU W, JIA Y, et al.Going Deeper with Convolutions [C] // IEEE Conference on Computer Vision and Pattern Recognition.Santiago:IEEE Computer Society,2015:1-9.

[8] GIRSHICK R,DONAHUE J,DARREL T, et al.Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [J]. Computer Science, 2013(6):580-587.

[9] GIRSHICK R.Fast R-CNN[C]//International Conference on Computer Vision.Santiago:IEEE,2015:1440-1448.

[10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.

[11] HE K,ZHANG X,REN S,et al.Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(9):1904-1916.

[12] LONG J,SHELHAMER E,DARREL T.Fully Convolutional Networks for Semantic Segmentation[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(4):640-651.

[13] DUMOULIN V, VISIN F. A Guide to Convolution Arithmetic for Deep Learning[J].Stat,2016,1050:23.

[14] ZEILER M D,FERGUS R.Visualizing and Understanding Convolutional Networks [C] // European Conference on Computer Vision.Zurich :Springer, 2014:818-833.

[15] ROTHER C,KOLMOGOROV V,BLAKE A.“GrabCut”: Interactive Foreground Extraction Using Iterated Graph Cuts[J].IEEE Transactions on Graphics,2004,23(3):309-314.

[16] GARCIAGARCIA A,ORTSESCOLANO S,OPREA S,et al. A Review on Deep Learning Techniques Applied to Semantic Segmentation [J]. Computer Science, 2017(4):1704.