

文章编号: 1672-8785(2017)08-0037-07

基于深度学习的红外遥感 信息自动提取

陈睿敏^{1,2} 孙胜利¹

(1. 中国科学院上海技术物理研究所, 上海 200083 ;

2. 上海科技大学信息科学与技术学院, 上海 200120)

摘 要: 为了提高红外遥感图像地物信息自动提取的精确性, 同时避免人工提取遥感信息的低效性, 提出了一种基于 UNet 深度学习模型的遥感信息提取算法。该算法用于从红外遥感图像中分割出 5 类地物信息 (包括道路、建筑、树木、农田和水体)。首先, 对分辨率高但数量较少的训练数据进行小像幅的随机裁剪, 并对其进行相应的数据增强处理。然后搭建 UNet 深度学习模型, 并用它自动提取遥感图像的特征信息。采用交叉熵损失函数以及 Adam 反向传播优化算法对该模型进行训练, 并对测试样本中的 5 幅遥感图像进行精确的地物信息提取。最后, 运用 Jaccard 指数对测试结果进行精度评定。实验结果表明, 该方法对高分辨率红外遥感图像信息和可见光遥感图像信息进行了充分融合, 对于不同种类地物的定位和分类都取得了较高精度。

关键词: 深度学习; UNet ; 语义分割; 多光谱遥感

中图分类号: TP751 **文献标志码:** A **DOI:** 10.3969/j.issn.1672-8785.2017.08.008

Automatic Extraction of Infrared Remote Sensing Information Based on Deep Learning

CHEN Rui-min^{1,2}, SUN Sheng-li¹

(1. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

2. School of Information Science and Technology, ShanghaiTech University, Shanghai 200120, China)

Abstract: To improve the accuracy of automatic extraction of object information in infrared remote sensing images while avoiding the inefficiency of manual extraction of remote sensing information, a remote sensing information extraction algorithm based on the UNet deep learning model is proposed. The algorithm is used to segment five kinds of object feature information including road, building, tree, farmland and water in infrared remote sensing images. Firstly, a small number of high resolution training data are cropped randomly and corresponding data enhancement processing is implemented on them. Then, a UNet deep learning model is established and is used to extract the feature information in remote sensing images automatically. The model is trained by using the cross-entropy loss function and Adam optimization algorithm and is used to extract the object information in five remote sensing images accurately. Finally, the classification result is evaluated by using the Jaccard index. The experimental r-

收稿日期: 2017-04-01

基金项目: 中国科学院上海技术物理研究所 2015 年创新专项 (CX-63)

作者简介: 陈睿敏 (1993-), 男, 江西宜春人, 硕士研究生, 主要研究方向为计算机视觉与机器学习。

E-mail: chenrm@shanghaitech.edu.cn

results show that this method can fully fuse the high resolution infrared remote sensing image information with the visible remote sensing image information. It has higher accuracy in positioning and classification for various objects.

Key words: deep learning; UNet; semantic segmentation; multispectral remote sensing

0 引言

遥感图像分析在诸如现代化农业和城市规划之类的许多实际应用中具有不可或缺的重要作用。近年来,随着航天遥感技术的不断发展,可用卫星图像的数量大幅增加。法国的 Pleiades 卫星甚至在不到一天的时间内就可生成覆盖整个地球的高空间分辨率图像。然而,大量的红外和可见光遥感数据不仅会加大图像分析的难度,而且还会造成传统人力提取遥感地物成本的快速增加。

近年来,许多传统的遥感图像自动分类算法的思想是将每个像素的频谱分配给某一类。还有一些更先进的算法则通过组合来自几个相邻像素的信息来增强分类器的性能。这类方法通常被称为光谱空间分类。该方法在很大程度上依赖于不同类的单个像素或某些相邻像素的光谱之间的可分离性^[1]。然而,对于本文考虑的大像幅遥感图像分类问题,这些方法的精确性很难得到保证。而且,当前的大规模卫星图像通常不使用高光谱分辨率传感器,因而难以区分目标类别的光谱。另一方面,由于红外遥感图像覆盖的空间范围较大,类别内部存在相当大的变异性。这种差异进一步降低了类别的可分性^[2]。

因此,为了解决传统红外遥感图像地物分类中所出现的问题,本文采用一种基于 UNet 深度学习模型的机器学习方法,对在同一地区采集的高分辨率红外/可见光遥感图像进行了精确的自动分类。

1 深度学习模型

卷积神经网络是将深度学习用于图像领域的非常成功的一种模型^[3],它被广泛应用于图像分类、目标检测、语义分割和行为识别等场

合。卷积神经网络主要通过权值共享和 ReLU 激活函数等手段解决传统浅层神经网络中复杂度高、难收敛以及容易发生梯度消失现象等问题,从而构造了一种深层的神经网络。卷积神经网络通常由卷积层、池化层和全连接层组成。在此基础上,针对不同的应用场景还出现了很多更加复杂的网络结构,包括 FCN^[4]、ResNet^[5]、Faster RCNN^[6]、SSD^[7]和 GAN^[8]等等。与传统的图像特征提取算法相比,卷积神经网络的优势是可根据图像的局部信息,按照不同的层次逐层学习更深层次的特征,将纹理、颜色和边缘等浅层特征组合起来,并利用这些组合特征进行非常精确的分类。

1.1 全卷积神经网络

本文采用的 UNet 模型由 Ronneberger O 等人^[9]提出,主要应用于医学图像的语义分割。该模型是一种全卷积神经网络。它与传统的卷积神经网络最大的不同之处在于不包含全连接层。这种全卷积神经网络目前主要用于语义分割领域。然而随着近年来越来越多的模型采用这种结构,全卷积神经网络正在逐步取代包含全连接层的网络结构,进而成为了未来发展的趋势。简单说来,语义分割问题就是对一幅图像



图 1 语义分割效果

中的所有像素点进行分类;除了需要识别图像特定位置上的类别信息外,还要对不同物体之间的边界进行分割。从图 1 中可以很清晰地看到基于语义分割的效果。

Jonathan 等人提出的全卷积神经网络的整体结构主要可以分为两部分:前半部分为一个 AlexNet^[10],其作用是对图像的粗粒度特征和细粒度特征进行提取,以便于后续分类;而后半部分则与 AlexNet 有本质区别。在全卷积神经网络中,将 AlexNet 的最后三层全连接层替换为卷积层。这样,输出的结果就不是一个一维的分类信息,而是一个二维的特征图了。如此一来,对于输出的这个特征图,人们就可以实施接下来的分类步骤了。

在全卷积神经网络中还有两个比较重要的思想:一是上采样。在前半部分的全卷积神经网络中使用池化层来缩小输出图像的尺寸,而目标的输出结果则需要一个与原图像尺寸相同的分割图,所以要对最后一层进行上采样。二是跳跃结构。这种结构的目的是为了优化结果。直接对通过全卷积得到的特征图进行上采样所产生的结果通常很粗糙,而跳跃结构则先对不同池化层的结果进行上采样,然后结合这些结果对其进行修正,这样可以使获得的分割结果保持原有的分辨率。

1.2 UNet 模型

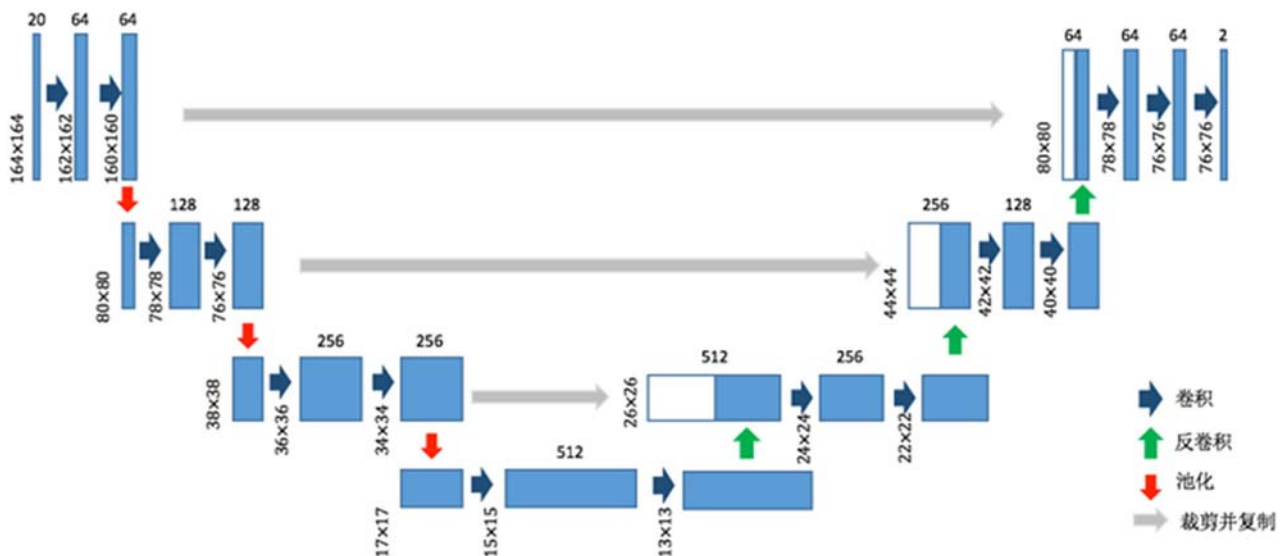


图 2 UNet 网络结构

UNet 模型的主要结构由卷积压缩部分和反卷积还原部分组成,如图 2 所示。在卷积压缩部分,它重复采用基于 2 个 3×3 卷积层和 1 个 2×2 最大池化层的结构,并利用池化层的降采样效果逐层提取更加高层的特征;在反卷积还原部分,先进行 2×2 的反卷积,使特征图的维数减半,然后接 2 个 3×3 卷积层,并重复这一结构;在输出层,用 1 个 1×1 卷积层将特征图映射成所需的类数。值得注意的是,池化层的输入必须是一个偶数宽/高尺寸的特征图。因此,网络设计一定要合理,以免出现参数错误。此外,在原有 UNet 结构的基础上,本文在卷积层之后添加了批标准化层并采用 Adam^[11] 优化算法对其进行了迭代收敛。其中,批标准化层既可避免在训练过程中发生“梯度弥散”现象,又可加快模型的收敛,进而提高模型精度。

UNet 模型的结构借鉴了全卷积神经网络的主要思想并对其进行了一定的优化。可以看出,UNet 是一种类似于 U 型对称的结构,前半部分是一个经典的 VGGNet-16^[12] 网络,而后半部分则是一个与前半部分相反的上采样过程。UNet 最重要的改进是在上采样部分,即加入了复制和裁剪通道,使网络能够将上下文信息从浅层传递到更高分辨率的层。这种结构使得 UNet 和全卷积神经网络同样得到一幅分割图,但是前者

的每个像素都包含了原图的上下文信息,使其对小目标分割具有更佳的效果。

2 实验过程及结果分析

2.1 数据集

本文采用文献 [13] 中所提供的遥感图像数据集。该数据集包含 25 幅 $1\text{ km} \times 1\text{ km}$ 大小的 3 通道和 16 通道遥感图像。这些图像都拍摄于同一个大区域。其中,3 通道遥感图像数据集包含可见光波段的 RGB 三通道图像,而 16 通道遥感图像数据集则包含 $400 \sim 1040\text{ nm}$ 波段的多光谱图像以及 $1195 \sim 2365\text{ nm}$ 波段的短波红外图像;全色波段的分辨率为 0.31 m ,多光谱波段的分辨率为 1.24 m ,红外波段的分辨率为 7.5 m (见图 3)。本文实验的目标就是将其中的 20 幅遥感图像作为训练集,以图像中包含的地物(水域、建筑、道路、树木和农田等五类)作为分类目标,对 UNet 深度学习模型进行训练,并对测试集中另外 5 幅图像所包含的地物进行识别和分类。



图 3 文献 [13] 中的 RGB 遥感图像示例

2.2 数据预处理和数据增强

首先对原始输入图像进行预处理。经过观察,本文对原始图像进行了白平衡修正^[14]。白平衡修正是根据不同的预设校正基准从原始偏色图像中得到图像的偏色信息,然后根据算得的偏色信息对图像进行校正和恢复。

然后对数据进行增强处理。数据增强是深度学习中十分关键的一步。少量数据难以使复杂的深度模型收敛,容易产生欠拟合现象,而数据之间最好不要产生过多的线性关系,否则就容易造成模型在局部最优值上收敛。而且,在本文实验中,需分类的地物在遥感图像中所占的面积比例较小(见表 1),给模型训练带来了困难。基于以上原因,我们先对遥感图像进行 164×164 的随机采样,并且在训练时不将所有的采样样本作为输入,而是取其中一半包含目标地物的作为正例,将另一部分不包含目标地物的样本作为负例进行训练。这样一来,就可保证一次训练的正负样本比例相对均匀,有利于模型收敛。然后对采样样本进行随机的翻转、裁剪、去中心化和归一化等操作,最终产生了大量的训练样本。

表 1 地物在遥感图像中所占的面积

地物种类	训练集中所占的比例
建筑	5.39%
道路	3.81%
水体	2.61%
树木	11.74%
农田	30.90%

数据集提供了分类标签,给出了标注有水域、建筑、道路、树木和农田这五类地物的多边形地理坐标,并将其归一化为 $x \in [0, 1]$ 和 $y \in [-1, 0]$ 。于是需按照式 (1) ~ 式 (4) 进行地理坐标和图像坐标的转换:

$$W' = W \cdot \frac{W}{W+1} \quad (1)$$

$$x' = \frac{x}{x_{\max}} \cdot W' \quad (2)$$

$$H' = \frac{H}{H+1} \quad (3)$$

$$y' = \frac{y}{y_{\min}} \cdot H' \quad (4)$$

2.3 实验过程

实验所用的得分函数采用基于像素的 Soft-Max 得分,其定义为

$$p_k(x) = \frac{\exp(a_k(x))}{\sum_{k'=1}^K \exp(a_{k'}(x))} \quad (5)$$

式中, $a_k(x)$ 为对应第 k 类的输出图像中像素 x 处的值; $a'_k(x)$ 为其他输出图像中像素 x 处的值; K 为总类数。因此, 当 $p_k(x) \approx 1$ 时, $a_k(x)$ 取极大值, 像素 x 的分类结果也就是最好的。

于是, 实验所用的交叉熵损失函数为

$$E = \sum_{x \in \Omega} w(x) \log(p_{l(x)}(x)) \quad (6)$$

式中, $w(x)$ 为每一类的权值; 通过调节 $w(x)$ 可以调节整体分类的趋向性; Ω 为所有像素的集合。

在深度学习中, 样本数量不均匀会导致模型很难收敛, 而在本文实验中也出现了类似问题。在遥感图像中, 水域、建筑、道路、树木和农田这五类地物的面积大小很不平衡, 道路和水域的面积较少。因此, 本文不是训练一个完整的 6 类分类器 (包括背景), 而是采用 5 个二分类器分别对 5 类地物进行分类, 并对分类结果进行融合。

此外, 本文实验还进行了 5 折交叉验证, 即将训练集分为 5 份, 每次将其中 1 份取出作为测试集, 并将其他 4 份作为训练集。这种方法是机器学习中用来避免出现拟合现象的常用手段。图 4 和图 5 所示分别为训练过程中精确度和 loss 值的走势图。

2.4 实验结果及精度评定

图 6 所示为建筑、树木、农田和道路的实验分割结果。由于样本采集于城市地区, 水体的实

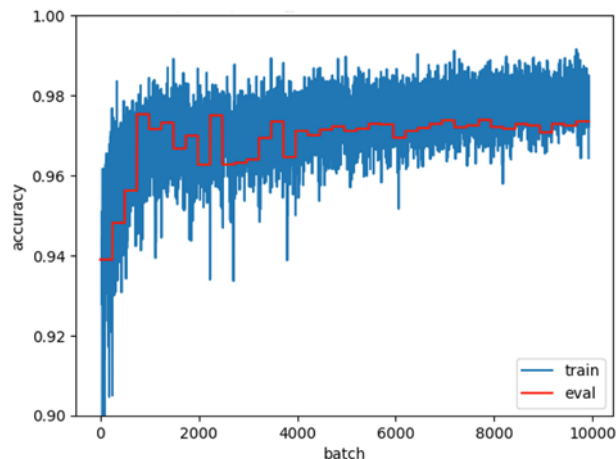


图 4 训练过程中精确度的走势图

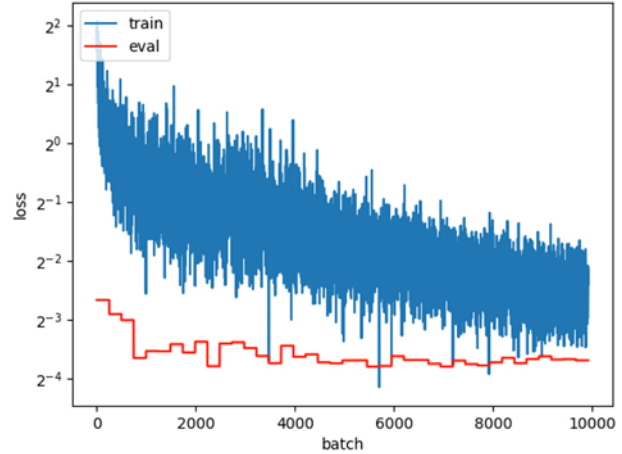


图 5 训练过程中 loss 值的走势图

验分割结果不明显。实验利用式 (7) 和式 (8) 将图像坐标还原成地理坐标, 从而形成由多个多边形组成的分类结果, 并最终采用 Jaccard 指数和分类精度对实验结果进行了精度评定。Jaccard 指数和分类精度的定义分别见式 (9) 和式 (10)。

$$x = x' \cdot \frac{x_{max}}{W'} \quad (7)$$

$$y = y' \cdot \frac{y_{max}}{H'} \quad (8)$$

$$Jaccard = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (9)$$

$$Precision = \frac{\text{正确分类的像素点个数}}{\text{被分类的像素点总个数}} \quad (10)$$

式中, A 是标签为第 k 类的总多边形区域; B 是被识别为第 k 类的总多边形区域。Jaccard 指数能够直观地表示地物被正确分类的程度。表 2 和表 3 分别列出了每一类地物的 Jaccard 指数结果和分类精度结果。

本文将深度学习语义分割模型引入到遥感图像信息自动提取中, 弥补了大多数遥感图像信息提取算法只采用传统人工特征且精度不高

表 2 各类地物的 Jaccard 指数结果

地物种类	Jaccard 指数
建筑	0.729
道路	0.593
水体	0.853
树木	0.540
农田	0.844

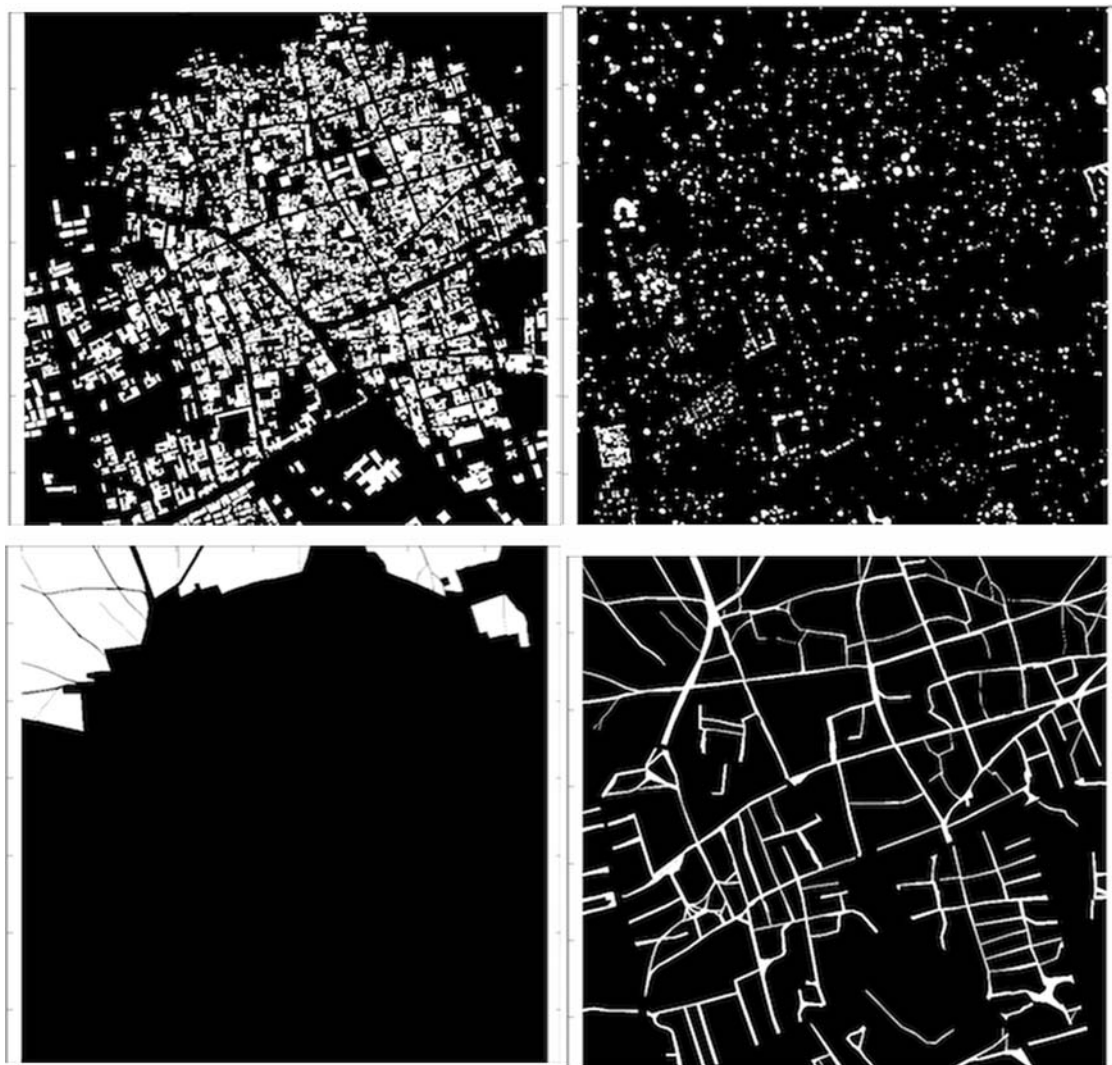


图 6 实验分割结果 (依次为建筑、树木、农田、道路)

表 3 各类地物的分类精度结果

地物种类	分类精度
建筑	0.963
道路	0.941
水体	0.960
树木	0.950
农田	0.967
平均	0.956
文献 [15] 算法	0.905
文献 [16] 算法	0.857

的缺陷。此前, UNet 模型通常用于医学影像分割, 而本文则将其应用于遥感图像的地物分类。另外, 本文模型还能够将红外遥感图像信息、多

光谱图像信息以及可见光图像信息结合起来并完成统一训练, 而传统方法却无法对这三类信息进行有效结合。经实验发现, 本文采用的 UNet 模型对于遥感图像地物分割具有很高精度, 在数据集^[13]上的平均识别精确度已经达到了 0.956。这比传统算法^[15]及以往的深度学习模型^[16]都有较大提升。

3 结束语

本文采用 UNet 深度学习模型分别提取了遥感图像的底层特征和高层特征, 并运用交叉熵损失函数和 Adam 优化算法进行了模型训练。采用该模型对遥感图像的 5 类地物信息 (包括建筑、道路、水体、树木和农田) 进行了分割, 取

得了良好的分类和识别效果。实验结果表明, 本文采用的 UNet 深度学习模型获得了显著优于传统算法的分割结果, 非常适用于处理遥感图像这类小目标的分割问题, 因而具有巨大的应用潜力。通过本文不难得出以下结论: 随着深度学习在图像和计算机视觉领域的快速进展, 其在基于图像的红外遥感信息处理领域的应用范围将会越来越广, 人们对遥感信息的提取也将变得越来越精确。

参考文献

- [1] Maggiori E, Tarabalka Y, Charpiat G, et al. Convolutional Neural Networks for Large-scale Remote-sensing Image Classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**(2): 645–657.
- [2] 董崧, 臧淑英, 吴长山, 等. SPOT-7 遥感图像融合技术对比研究 [J]. *测绘与空间地理信息*, 2017, **40**(1): 75–78.
- [3] 曹林林, 李海涛, 韩颜顺, 等. 卷积神经网络在高分遥感影像分类中的应用 [J]. *测绘科学*, 2016, **41**(9): 170–175.
- [4] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation [C]. Boston: 2015 IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [5] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]. Las Vegas: 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [6] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-time Object Detection with Region Proposal Networks [C]. Montreal: The Twenty-ninth Annual Conference on Neural Information Processing System, 2015.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot Multibox Detector [C]. Amsterdam: The 14th European Conference on Computer Vision, 2016.
- [8] Zhao J, Mathieu M, LeCun Y. Energy-based Generative Adversarial Network [J]. arXiv: 1609.03126, 2016.
- [9] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C]. Munich: The 18th International Conference on Medical Image Computing and Computer Assisted Intervention, 2015.
- [10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet Classification with Deep Convolutional Neural networks [C]. Lake Tahoe: The Twenty-sixth Annual Conference on Neural Information Processing Systems, 2012.
- [11] Kingma D, Ba J. Adam: A Method for Stochastic Optimization [J]. arXiv: 1412.6980, 2014.
- [12] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-scale Image Recognition [J]. arXiv: 1409.1556, 2014.
- [13] Kemker R, Kanan C. Deep Neural Networks for Semantic Segmentation of Multispectral Remote Sensing Imagery [J]. arXiv: 1703.06452, 2017.
- [14] Hung S B, Kandhadai A, Chiu A C. Automatic White Balance Method and Apparatus: US7728880 [P]. 2010–06–01.
- [15] 张建廷, 张立民, 徐涛. 语义网络引导的遥感图像分类 [J]. *遥感信息*, 2016, **31**(1): 38–42.
- [16] 程环环, 王润生. 利用贝叶斯网络融合空间上下文的高分辨遥感图像分类 [J]. *计算机工程与科学*, 2011, **33**(1): 70–76.
- [15] 丁亮. 隐身飞机尾焰红外探测仿真技术研究 [D]. 成都: 电子科技大学, 2016.
- [16] 简爱平. 红外目标图像建模与仿真 [D]. 西安: 西北工业大学, 2001.
- [17] 王杏涛. 飞行器红外辐射特征及发动机舱蒙皮降温规律研究 [D]. 南京: 南京航空航天大学, 2015.
- [18] 郑礼宝. 二元喷管对发动机性能影响的实验研究 [J]. *推进技术*, 1997, **18**(4): 58–61.

(上接第 7 页)

- [12] 张小英, 王先炜. 轴对称矢量喷管红外特性的数值计算研究 [J]. *航空动力学报*, 2000, **19**(5): 366–369.
- [13] 何建伟, 曹晨, 张昭. 红外系统对隐身飞机的探测距离分析 [J]. *激光与红外*, 2013, **43**(11): 1243–1247.
- [14] 史际刚. 飞机尾焰红外辐射特性研究 [D]. 西安: 西北工业大学, 2005.