

印刷图像信息处理专题

基于卷积神经网络的光学遥感图像船只检测

欧阳颖卉, 林翬, 李树涛
(湖南大学, 长沙 410082)

摘要: **目的** 研究无需进行复杂的图像预处理和人工特征提取, 就能提高光学遥感图像的船只检测准确率和实现船只类型精细分类。**方法** 对输入的检测图像, 采用选择性搜索的方法产生船只候选区域, 用已经标记好的训练样本对卷积神经网络进行监督训练, 得到网络参数, 然后使用经过监督训练的卷积神经网络提取抽象特征, 并对候选区域进行分类, 根据船只候选区域的分类概率同时确定船只的位置以及类型。**结果** 与现有的 2 种检测方法进行对比, 实验结果表明卷积神经网络能有效提高船只检测准确率, 平均检测准确率达到 93.3%。**结论** 该检测方法无需进行复杂的预处理, 能同时对船只进行检测和分类, 并能有效提高船只检测准确率。

关键词: 卷积神经网络; 光学遥感图像; 船只检测

中图分类号: P407.8; TP391 文献标识码: A 文章编号: 1001-3563(2016)15-0001-06

Convolutional Neural Network Based on Ship Detection in Optical Remote Sensing Image

OUYANG Ying-hui, LIN Hui, LI Shu-tao
(Hunan University, Changsha 410082, China)

ABSTRACT: This papers aims to improve the ship detection precision in optical remote sensing images and realize sophisticated classification of ships without need to conduct complicated pre-processing step and manual feature extraction. For input detection images, selective search method was used to generate candidate ship regions. The parameters of convolutional neural network were obtained by supervised training with labeled training data. Then the trained convolutional network was utilized to classify the candidate regions. The classification results of ship candidate regions were used to locate the ship and identify the category of ship. Compared with two existing detection methods, the experimental results indicated that the proposed convolutional neural network based method could efficiently improve the detection precision and achieve an average detection precision of 93.3%. In conclusion, the method can detect and classify ships simultaneously without complicated pre-processing step. And it also can improve the efficiency of ship detection precision as well.

KEY WORDS: convolutional neural network; optical remote sensing image; ship detection

船只的检测是遥感图像研究的热点领域, 因其能广泛应用于水上交通监管、渔业管理以及船只救援等重要方面。船只的检测对于海上船只补给物资的包装也有重要作用, 远洋运输公司需根据船只的

类型确定物资包装的规格, 舰队需根据舰船的类型和位置确定补给物资的包装要求以及包装特点^[1]。过去常使用全天时、受天气影响较小的合成孔径雷达 (SAR) 图像对船只进行检测^[2]。随着近年来光

收稿日期: 2016-05-11

作者简介: 欧阳颖卉 (1992—), 女, 湖南祁东人, 湖南大学硕士生, 主攻图像检测。

通讯作者: 林翬 (1965—), 男, 湖南益阳人, 湖南大学副教授, 主要研究方向为图像识别。

学遥感图像分辨率的不断提高,相比于 SAR 图像,光学遥感图像能够提供更为丰富的视觉信息,从而有利于船只的定位以及识别^[3]。

目前的光学遥感图像检测方法主要有 2 个步骤:首先对识别的图像进行特征提取,再使用分类器对其进行分类。Zhu 等人提取了局部二值模式特征(LBP)以及形状和灰度分布特征,采用支持向量机(SVM)作为分类器来检测船只^[4]。Yang 等人利用海面特征和线性分类器来检测船只^[5]。Qi 等人首先提取船只检测候选区域,再使用船只的方向梯度直方图特征进行辨别^[6]。这些方法仅提取了低层特征,对于海浪,云层以及岛屿等干扰没有良好的鲁棒性;人工选取特征是非常耗时的,且强烈依赖于专业知识和数据本身的特征,对于目前数据量巨大的光学遥感数据不能实现高效的实时检测^[7];这些方法仅能定位船只的位置,没有利用光学遥感图像丰富详细的视觉信息对船只类型进行精细分类,而对船只类型进行分类有利于船只的监管以及确定所需物资包装的规格和包装特点。

在视觉皮层的研究中,发现哺乳动物大脑皮层具有一定层次的深度结构。在视觉活动中,视觉信号在大脑皮层逐层传递时不断进行抽象。人们试图通过构建一系列机器学习公式对数据进行非线性变换,从而模仿大脑皮层的这一抽象过程,卷积神经网络就是受视觉皮层的深层结构启发而产生的。卷积神经网络是一种深度学习模型,它利用一系列的卷积层、池化层以及全连接层构建的一个多层网

络,来模仿人脑感知视觉信号的逐层处理机制,能捕捉输入与输出间建立的一种非线性映射关系。其具有的多层非线性映射的层结构,可以完成复杂的函数逼近^[8-9]。近年来,卷积神经网络(CNN)在图像分析领域不断取得突破性进展,尤其在字符识别、人脸检测,物体识别等领域,卷积神经网络在 ImageNet 的大规模视觉识别挑战赛(ILSVRC)中已连续刷新了图像分类和目标定位任务的识别率记录,说明卷积神经相比于传统的人工特征提取方法具有强大的图像表示和分类能力^[10]。卷积神经网络可以被看作是一个提取抽象、高阶特征的特征提取器和分类器,能够避免传统方法复杂的特征提取过程。文中提出了一种基于卷积神经网络的光学遥感图像船只检测方法,并设计卷积神经网络结构,能够同时实现船只的检测和分类,实现较高的平均检测准确率,并且能够精细区分货柜船、挖沙船以及游艇等船只的类型。

1 船只检测流程

文中提出的船只识别方法,首先将 1 张输入图像通过选择性搜索产生若干船只候选区域图像,再将产生的候选区域送入已经完成监督训练的卷积神经网络中进行特征提取和分类,得到候选区域属于每一类船只类型的概率。在所有船只候选区域图像中选择分类概率大于一定值的候选区域,该候选区域即是船只所在位置,分类概率最大的类,即是船只所属类型,详细的识别流程见图 1。

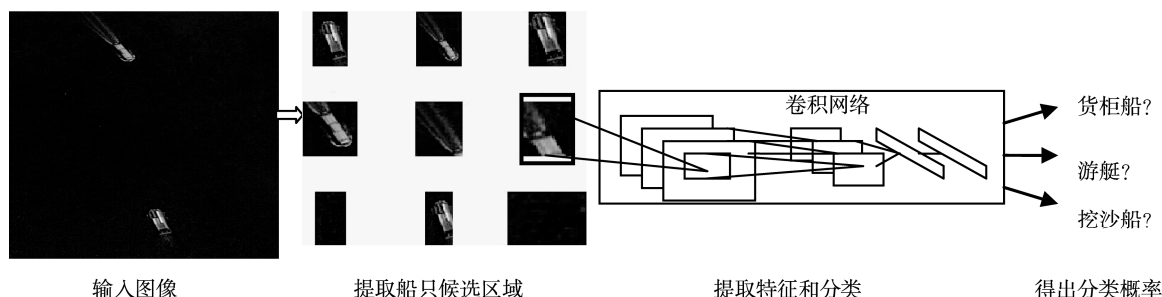


图 1 船只识别流程

Fig.1 Ship detection flow

2 选择性搜索

提取出可能包含有船只的候选区域是识别流程的第 1 步,候选区域的选择直接影响着检测准确率和计算效率。文中采用选择性搜索的方法来产生

船只候选区域,该方法先使用图形分割方法将原始图像分割成一个个小的区域 $R = \{r_1, \dots, r_n\}$,再计算相邻区域间的相似度 $s(r_i, r_j)$,将相似度最高的区域进行合并, $r_i = r_i \cup r_j$,对新合并的区域与邻域再进行相似度计算,直到整幅图像合并为整体^[11]。

在这一分阶聚合过程中,产生了若干张船只候选区域。不同于尺寸固定的滑动窗的方法,选择性搜索考虑到了不同物体比例的大小。选择性搜索常与卷积神经网络相连接,如同一个粗滤波器,显著减少了需要进行特征提取的区域数量。

3 卷积神经网络

3.1 卷积神经网络结构

为了实现船只检测准确率高的要求,设计了一种用于船只检测的卷积神经网络。文中的卷积神经网络由 9 层组成,其中包括 3 层卷积层(C1, C3, C5),以及每层对应连接的最大池化层(S2, S4, S6),一层全连接层(F7),对应货柜船、游艇和挖沙船 3 类的输出层,以及用来输出图像对应于每一类概率的 soft-max 层。文中所采用的卷积网络结构见图 2。

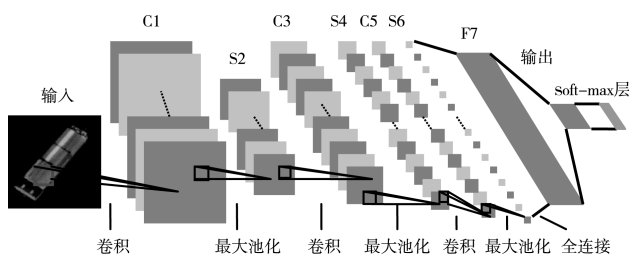


图 2 卷积神经网络结构

Fig.2 The structure of the convolutional neural network

卷积层是卷积神经网络的核心部分。卷积层将不同的卷积核分别与前一层的所有特征图进行卷积求和,加上偏置,再将结果通过修正线性单元(ReLU)激活函数的输出形成当前层的神经元。卷积操作的表达式为:

$$y^j = f(0, b^j + \sum_i k^{ij} * x^i) \quad (1)$$

式中: x^i 和 y^j 分别为第 i 层的输入和第 j 层的输出; k^{ij} 为 i 层和 j 层之间的卷积核; $*$ 表示卷积; b^j 为 j 层的偏置; f 为激励函数。这里选择 ReLU 作为非线性激励函数,而不是常用的 Sigmoid 函数或者双曲正切函数,因为 ReLU 函数能使数据稀疏,其表达式为:

$$y = \max(0, x) \quad (2)$$

池化层通常跟随在卷积层后面。随着网络的层数的递增,特征图的个数急剧增加,通过池化来降低卷积层输出的特征向量,使结果不易出现过拟合

现象。最大池化法的表达式:

$$y_{j,k}^i = \max_{0 \leq m, n < s} \{x_{j-s+m, k-s+n}^i\} \quad (3)$$

对第 i 层特征图的每一个神经元 x^i 使用 $s \times s$ 大小的窗口进行下采样产生输出 y^i 。文中采用最大池化法,用 2×2 大小的窗口对上一层的特征图进行下采样。最大池化法是计算窗口区域的最大值为池化后的值,该方法能选出最大不变量来加速网络收敛速度^[12]。最大池化法并不改变特征图的数目,只改变了特征图的大小,降低了特征图对于平移、缩放和旋转的敏感度。

最后一层隐含层是全连接层,全连接层将池化后的特征图转化为一维数组,每个神经元的输出表达式为:

$$y_j = \max(0, \sum_i x_i \cdot \omega_{i,j} + b_j) \quad (4)$$

式中: x_i 为池化后的神经元的值; $\omega_{i,j}$ 为权值; b_j 为偏置。激活函数仍采用 ReLU 函数。全连接层的神经元个数直接影响着网络的拟合效果和训练速度。

最后的输出层是用来预测每一类输出概率的 soft-max 层, soft-max 是一种多分类器,当给定输入样本后,能输出 0 到 1 之间的数值,代表样本属于该类的概率,其表达式:

$$p_i = \frac{\exp(y_i)}{\sum_{j=1}^n \exp(y_j)} \quad (5)$$

式中: y_i 为第 i 个神经元输出的结果; n 为神经元个数; $\sum_{j=1}^n \exp(y_j)$ 为总的神经元输出结果。

首先将输入图像归一化到 $3 \times 40 \times 40$ (通道 $\times$ 高 $\times$ 宽),再使用 32 个 5×5 的卷积核对输入图像进行卷积,得到特征图大小为 36×36 的卷积层(C1 层),之后对 C1 层的特征图进行下采样,池化层(S2 层)中的特征图的每个神经元与 C1 层中相对应的特征图的 2×2 邻域相连。根据池化层的特点, S2 层中的特征图数目不变,大小将缩小为 C1 层特征图大小的 $1/4$, 变成 18×18 。同样,卷积层(C3 层)使用 64 个 3×3 的卷积核卷积 S2 层中的特征图,进行下采样得到特征图大小为 8×8 的池化层(S4 层),卷积层(C5 层)使用 128 个 3×3 的卷积核卷积 S4 层中的特征图。再对 C5 层的特征图进行下采样,最终得到池化层(S6 层) 128 个为 3×3 的特征图。

为了得到更好的拟合效果,文中设置全连接层(F7层)的神经元个数为500。为了反映货柜船、挖沙船和游艇这3类船只的分类结果,设置输出神经元个数为3。

3.2 卷积神经网络参数训练

卷积神经网络作为一种监督学习算法,需要大量已标记的数据对神经网络节点进行训练,找出输入与输出之间的映射关系,从而使网络具有逐层抽取数据特征和深度挖掘数据信息的能力^[13]。卷积神经网络的参数训练分为前向传播和反向传播等2个阶段。在前向传播阶段,将训练样本送入网络,经过中间的隐含层,逐层变换、映射,直到输出层得出激励响应。在反向阶段,将激励响应同训练输入对应的目标输出求差,得到隐含层和输出层响应误差,根据误差进一步对网络模型的参数进行调优^[14]。

在参数训练阶段,使用 soft-max loss 层来计算反向传播时的损失函数,其表达式为:

$$E = \frac{1}{N} \sum_{n=1}^N \lg(p_i, l_n) \quad (6)$$

式中: N 为输出神经元个数; l 为输入对应的目标输出,即类别标签; p_i 为第 i 个神经元经 soft-max loss 层后输出的概率。

在反向传播过程中,卷积神经网络将得到的损失函数利用梯度下降法对权值和偏置值进行更新^[15]。其表达式为:

$$\Delta W_{ij}(k+1) = \eta \Delta W_{ij}(k) - \alpha(1-\eta) \frac{\partial E}{\partial W_{ij}(k)} \quad (7)$$

$$W_{ij}(k+1) = W_{ij}(k) - \Delta W_{ij}(k+1) \quad (8)$$

$$\Delta b_j(k+1) = \eta \Delta b_j(k) - \alpha(1-\eta) \frac{\partial E}{\partial b_j(k)} \quad (9)$$

$$b_j(k+1) = b_j(k) - \Delta b_j(k+1) \quad (10)$$

式中: k 为迭代次数; ΔW_{ij} 和 Δb_j 分别为权值和偏置的调整值; W_{ij} 和 b_i 为更新后的权值和偏置值; α 为学习效率; η 为动量因子; E 为式(6)所表示的损失函数。

在卷积神经网络的训练过程中,前向传播和反向传播2个过程反复交替,直到网络收敛。

4 实验结果与分析

文中实验分为2步:使用训练样本对卷积神经

网络进行训练,得到卷积网络参数;用检测样本进行检测。

4.1 参数训练

训练使用3000张已标记好的船只样本(货柜船、游艇以及挖沙船各1000张,见图3)对卷积神经网络进行训练。

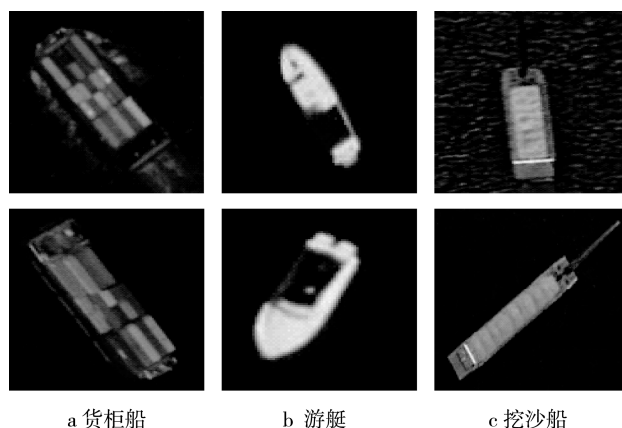


图3 训练样本

Fig.3 Sample images from the training dataset

训练参数设置:学习率为0.002;系数衰减率为0.0002。并用另外的300张船只图片作为测试样本对训练好的网络进行测试,测试结果见图4。

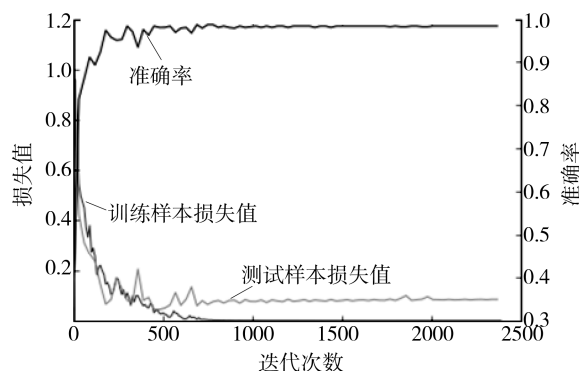


图4 损失和准确率曲线

Fig.4 The loss and accuracy curves

由图3可知训练样本和测试样本的损失值随迭代次数的增加而不断减小,卷积神经网络在1000次迭代时已经收敛到了一个极值,测试样本的分类准确率不断提高。经过2500次迭代后,测试准确率稳定在99.18%。由此说明经过监督训练后的网络参数能精细区分货柜船、游艇和挖沙船。

4.2 检测样本测试

检测样本由60幅像素大小超过500×500的

光学遥感船只图像构成,其中包含货柜船、游艇和挖沙船各 20 艘。为了验证方法的有效性,选择了其他 2 种方法进行对比:DPM(Deformable Parts Models)方法,DPM 是一种成功的物体检测方法,使用 HOG 作为特征提取器,SVM 作为分类器^[16];Lenet-5 网络,是具有 2 层卷积层的卷积神经网络,在字符识别方面具有较高的精确度。每幅图像的平均检测时间为 9.44 s,部分实验结果示例见图 5。对比了不同方法对于 3 种船只的检测分类结果,见表 1。

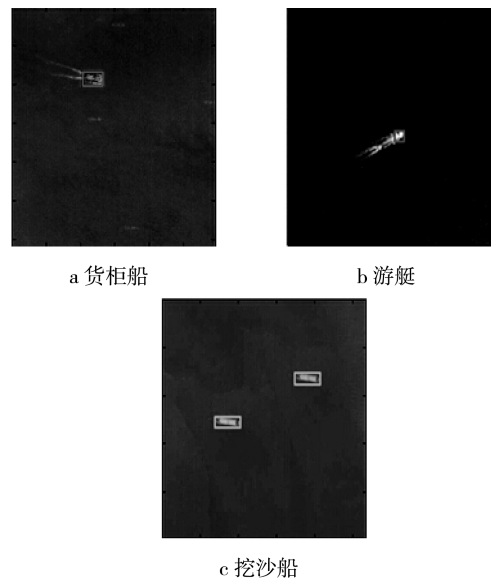


图 5 检测结果示例
Fig.5 The sample result of detection

表 1 船只定位分类结果
Tab.1 The results of detection and classification

检测方法	货柜船	游艇	挖沙船	平均检测 准确率/%
文中 CNN 方法	20	17	19	93.3
DPM	17	15	17	81.7
Lenet-5	18	15	18	85.0

从以上结果可以看出,文中提出的卷积神经网络船只检测方法尽管在游艇这一类上与其他 2 类在识别精度有一定差距(因为游艇的体积较小、颜色较为接近海面),但相比于传统的物体检测方法和结构较为简单的 LeNet-5 网络有更高的检测准确率,3 类的平均检测准确率达到 93.3%。卷积神经网络能够准确定位船只位置,并对船只类型正确分类,能够满足光学遥感图像船只检测以及分类的应用要求。

5 结语

提出了一种基于卷积神经网络的光学遥感图像识别方法,该方法先通过选择性搜索产生船只候选区域,再通过卷积神经网络进行特征提取和分类,解决了传统方法需要进行人工特征提取的问题,能有效定位光学遥感图像中船只的位置,并且能够对船只类型进行精细分类。实验验证了该方法具有较高的检测准确率。

参考文献:

[1] 王艳芳,王建民. 海军舰艇物资包装建设需求与对策研究[J]. 包装工程, 2016, 37(3): 76—80.
WANG Yan-fang, WANG Jian-min. Building Requirements and Countermeasures of Material Packaging on Navel Ship[J]. Packaging Engineering, 2016, 37(3): 76—80.

[2] 吴樊,王超,张波,等. SAR 图像船只分类识别研究进展[J]. 遥感技术与应用, 2014, 29(1): 1—8.
WU Fan, WANG Chao, ZHANG Bo, et al. Study On Vessel Classification in SAR Imagery: A Survey[J]. Remote Sensing Technology and Application, 2014, 29(1): 1—8.

[3] BI Fu-kun, ZHU Bo-cheng, GAO Li-ning, et al. A Visual Search Inspired Computational Model for Ship Detection in Optical Satellite Image[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 9(4): 749—753.

[4] ZHU Chang-ren, ZHOU Hui, WANG Run-sheng, et al. A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features[J]. IEEE Transaction on Geoscience and Remote Sensing, 2010, 48(9): 3446—3456.

[5] YANG Guang, LI Bo, JI Shu-fan, et al. Ship Detection from Optical Satellite Images Based on Sea Surface Analysis[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(3): 641—645.

[6] QI Sheng-xiang, MA Jie, LIN Jin, et al. Unsupervised Ship Detection Based on Saliency and S-HOG Descriptor from Optical Satellite Images[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(7): 1451—1455.

[7] 高常鑫,桑农. 基于深度学习的高分辨率遥感影像目标检测[J]. 测绘通报, 2014(S0): 108—111.
GAO Chang-xin, SANG Nong. Deep Learning for Object Detection in Remote Sensing Image[J]. Bulletin of Surveying and Mapping, 2014(S0): 108—111.

[8] 孙志军,薛磊,许阳明,等. 深度学习研究综述[J]. 计算机应用研究, 2011, 29(8): 18—36.
SUN Zhi-jun, XUE Lei, XU Yang-ming, et al. Over-

- view of Deep Learning[J]. Application Research of Computers, 2011, 29(8): 18—36.
- [9] HINTON, GEOFFREY E, RUSLAN R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786): 504—507.
- [10] 卢官明, 何嘉利, 闫静杰, 等. 一种用于人脸表情识别的卷积神经网络[J]. 南京邮电大学学报: 自然科学版, 2016, 36(1): 16—22.
- LU Guan-ming, HE Jia-li, YAN Jing-jie, et al. Convolutional Neural Network for Facial Expression Recognition[J]. Journal of Nanjing University of Posts and Telecommunications: Natural Sciences, 2016, 36(1): 16—22.
- [11] UIJLINGS J, SANDE K, GEVERS T, et al. Selective Search for Object Recognition[J]. International Journal of Computer Vision, 2013, 104(2): 154—171.
- [12] NAGI J, DUCATELLE F, CARO D, et al. Max-pooling Convolution Neural Network for Vision-based Hand Gesture Recognition[C]// 2011 IEEE International Conference on Signal and Image Processing Application, 2011: 342—347.
- [13] 余勇维, 殷国富, 殷鹰, 等. 基于深度学习网络的射线图像缺陷识别方法[J]. 仪器仪表学报, 2014, 35(9): 2012—2019.
- YU Yong-wei, YIN Guo-fu, YIN Ying, et al. Defect Recognition for Radiographic Image Based on Deep Learning Network[J]. Chinese Journal Scientific Instrument, 2014, 35(9): 2012—2019.
- [14] 曹乐, 李润, 袁艳, 等. 基于 BP 神经网络的气调包装评价方法研究[J]. 包装工程, 2014, 35(1): 38—41.
- CAO Le, LI Run, YUAN YAN, et al. Study on the Method of MAP Evaluation Based on BP Network[J]. Packaging Engineering, 2015, 35(1): 38—41.
- [15] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based Learning Applied to Document Recognition[J]. IEEE Journal and Magazines, 1989, 86(11): 2278—2324.
- [16] FELZENSZWALB P F, MCALLESTER D, RAMANAN D. A Discriminatively Trained, Multiscale, Deformable Part Model[C]// In Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1—8.