

基于 CNN 模型的高分辨率遥感图像目标识别^{*}

曲景影 孙 显 高 鑫

(中国科学院电子学研究所空间信息处理与应用系统技术重点实验室 北京 100190)

摘 要: 遥感图像目标识别作为当前遥感图像应用领域中的主要研究内容,具有重要的理论意义和广泛的应用价值。近年来,深度学习成为机器学习领域的一个新兴研究方向,卷积神经网络(convolutional neural networks, CNN)是一种得到广泛研究与应用的深度学习模型。提出一种基于 CNN 模型的光学遥感图像目标识别方法,在传统 LeNet-5 网络结构的基础上,引入 ReLU 激活函数代替传统的 Sigmoid 函数和 tanh 函数,使用卷积展开技术将卷积运算转换为矩阵乘法,并对网络结构进行调整优化,提高目标识别的准确性和效率。利用 Quick Bird 上的 0.6 m 分辨率的遥感图像进行验证,实验结果表明,基于改进的 CNN 模型的方法可以取得较高的目标识别准确率和效率。

关键词: 遥感图像;卷积神经网络;激活函数;卷积展开;目标识别

中图分类号: TP391 TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Remote sensing image target recognition based on CNN

Qu Jingying Sun Xian Gao Xin

(Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Institute of Electronics,
Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Remote sensing image target recognition is of greatly theoretical significance and wide application value. In recent years, Deep Learning has become a new research direction in the field of machine learning, and Convolutional Neural Network (CNN) is a deep learning model of widely research and application. In this paper, a method of optical remote sensing image target recognition based on CNN model is proposed. Based on the traditional LeNet-5 network, ReLU activation function is used to replace the Sigmoid and tanh activation functions. Using convolutional layer unrolling technique, which makes the convolutional calculation be represented by matrix multiplication. The parameters and the structure of the CNN network are adjusted to make it optimal. The accuracy and effectiveness of the proposed method are verified by using Quick Bird 0.6m resolution on remote sensing image.

Keywords: remote sensing image; CNN; activation function; convolutional unrolling; target recognition

1 引 言

随着遥感技术的不断进步,遥感图像的空间分辨率也在稳步提高,人们可以获得越来越多的高分辨率遥感图像。因此,在大数据条件下的高分辨率遥感图像目标识别就显得尤为重要。遥感图像目标识别已被广泛研究,许多研究学者主要使用人工提取图像特征进行目标识别的方法。比如局部二值模式(LBP)、梯度方向直方图(HOG)^[1]、Gabor 等特征^[2],然后将这些特征以特征向量的形式输入到一个传统的分类器比如 SVM、AdaBoost、决策树等进行分

类,在目标识别任务中取得了很好的效果,比如飞机识别、舰船识别等^[3]。但是这种方法在人工特征提取方面、鲁棒性方面、位移、旋转等影响方面往往表现欠佳。

近年来,直接从数据中学习符合数据分布模型的方法受到广泛研究,比如线性自回归、神经网络^[4]、贝叶斯网络^[5]、基于统计学习的 SVM^[6]、模糊聚类等^[7]。相比于支持向量机等浅层分类器,含有多层隐含节点的神经网络能够更好地构建数据模型^[8]。然而多层神经网络由于存在梯度弥散、计算量大等问题,在实际应用中受到限制。以卷积神经网络为代表的深度学习方法,利用权值共享的技

收稿日期:2016-03

^{*} 基金项目:国家自然科学基金(41501485)项目资助

术,减少了网络中的参数,逐步克服神经网络中的计算量大等问题。LECUN Y 等^[9]设计并采用基于误差梯度的算法训练了 CNN 模型,并且在字符识别任务中展现出了相对于当时其他方法的领先性能。LAUER F 等^[10]将其应用在手写数字的识别上,目前,在欧洲许多国家已经使用 CNN 进行手写支票上的数字识别。LAWRENCE S 等^[11]使用 CNN 进行人脸识别。Ciresan 等^[12]将其应用于 CIFAR 等自然图像数据集上进行目标识别。TANG J 等^[13]将其应用于遥感图像中的舰船检测。鉴于 CNN 在目标识别及遥感图像目标检测上的应用,本文提出一种基于 CNN 模型的高分辨率遥感图像目标识别方法,对之前的目标识别方法存在的人工提取特征、鲁棒性差、计算量大等缺点进行改进,得到更好的遥感图像目标识别效果。

本文将 CNN 模型应用到遥感图像目标识别应用中,为了提高训练效率,引入 ReLU 激活函数代替传统的 Sigmoid 函数和 tanh 函数,减少梯度弥散带来的影

响,加速收敛。同时,使用卷积展开技术来实现优化,减少计算量。针对卷积神经网络所具有的层次化和局部感知区域的结构上的特点,选择最优的学习率、各层过滤器的数量和大小以提高目标识别的准确率。实验结果表明,本文提出的基于 CNN 模型的遥感图像目标识别方法在 Quick Bird 光学遥感图像数据集上取得了很好的准确率和效率。

2 CNN 模型简介

CNN 是一种包含卷积层的深度神经网络模型,已成为当前语音分析和图像识别领域的研究热点。它的权值共享网络结构使之更类似于生物神经网络,降低了网络模型的复杂度,减少了权值的数量。从结构上看,CNN 是一个多层的神经网络,主要由卷积层、子采样层和全连接层组成,每层由多个二维平面组成,而每个平面由多个独立神经元组成,经典的 LeNet-5^[14]结构如图 1 所示。

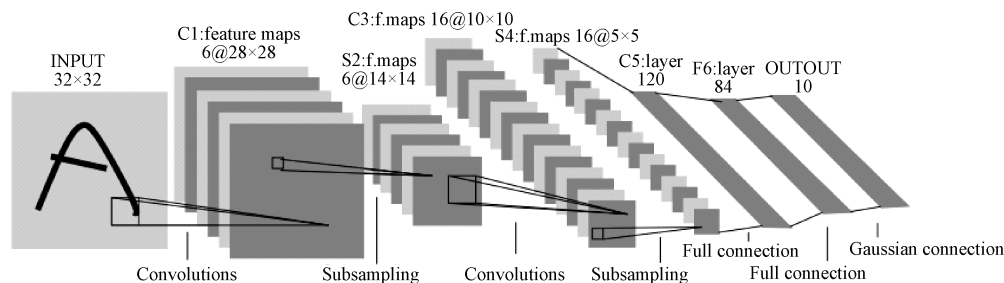


图 1 LeNet-5 结构

在典型的 CNN 模型里,卷积层和子采样层在起始的几层交替出现,而在最后靠近输出层的层次上,是一维的全连接网络。在这些层次中,二维的特征图已经被转化成一维网络的输入向量,此时在全连接层,会进行输入向量和权重向量的点积运算,再通过一次偏置运算将结果传递给激活函数,从而获得单元状态。

最后,欧式径向基函数单元 (euclidean radial basis function, ERBF) 组成 CNN 模型的输出层,每类样本具有一个单元,RBF 输出单元完成输入向量和参数向量之间欧氏距离的计算。当输入向量和参数向量距离越远,RBF 单元的输出越大。RBF 的输出被认为是作为衡量输入模式和与 RBF 关联类的模型的匹配程度的惩罚项。

3 基于 CNN 模型的遥感图像目标识别

3.1 方法流程

对于遥感图像目标识别任务,CNN 的处理流程如图 2 所示,首先将训练样本进行简单的规则化,调整为统一的尺寸,为防止训练样本中最开始几个样例可能出现劣质数据对训练过程产生不良影响,训练过程采用批处理的方式,即每次随机选取固定数目的训练样本作为一个小样本输入,使用 BP 算法对每个小样本进行一次权值更新,达到一定的迭代次数或者误差达到给定阈值时停止训练。

将测试数据输入已经训练好的 CNN 模型中,通过前向传播最终得到分类的结果。

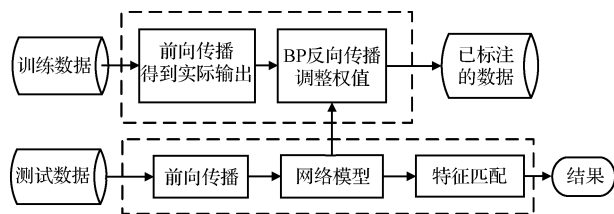


图 2 CNN 的处理流程

前向传播阶段主要包括卷积和采样两个过程。卷积和子采样的详细过程如图 3 所示。

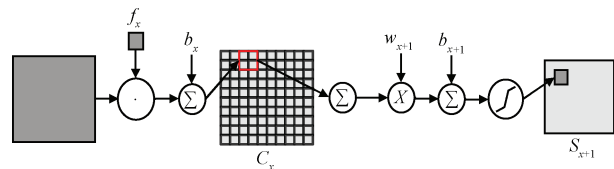


图 3 卷积和子采样过程

卷积过程:用一个可训练的滤波器 f_x 去卷积一个输入的图像,然后加一个偏置 b_x ,得到卷积层 C_x 。卷积层的形式如式(1)所示。其中, l 代表层数, k 为卷积核, M_l 表

示选择的输入特征图的集合,每个输出特征图会给一个额外的偏置 b 。

$$x_j^l = f(\sum_{i \in M_l} x_i^{l-1} \cdot k_{ij}^l + b_j^l) \quad (1)$$

采样过程:将卷积特征划分为数个 $n \times n$ 的不相交区域,通过 w_{l+1} 加权,再增加偏置 b_{x+1} ,然后通过一个激活函数,产生一个缩小 n 倍的特征映射图。子采样层的一般形式如式(2)所示。其中 $down(\cdot)$ 表示一个下采样函数。

$$x_j^l = f(w_{x+1} down(x_j^{l-1}) + b_{x+1}) \quad (2)$$

3.2 基于 ReLU 激活函数的 CNN 模型(RCNN)

在 CNN 模型中引入非线性函数作为激活函数,这样可以逼近任意函数,而不是输入的线性组合,实现深层网络的意义。常用的激活函数为 Sigmoid 函数或者 tanh 函数,然而,这两个函数都是饱和的。函数以及导数的曲线如图 4 所示,可以看出两个函数都是倒过来的碗状,即越接近目标,对应的导数越小,这种导数形式容易引起梯度弥散问题。

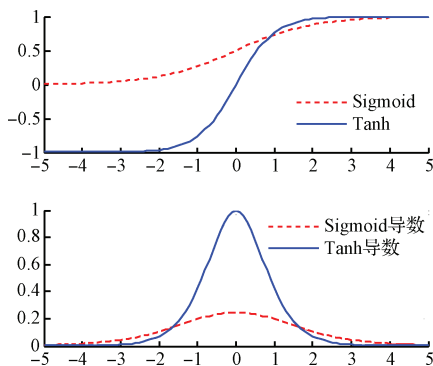


图 4 Sigmoid 函数、tanh 函数及其导数曲线

所谓梯度弥散问题具体是指,当使用反向传播方法计算导数的时候,随着网络深度的增加,反向传播即从输出层到网络的最初几层的梯度的幅度值会急剧地减小。结果就造成了整体的损失函数相对于最初几层的权重导数非常小。这样,当使用梯度下降法的时候,最初几层的权重变化非常缓慢,以至于它们不能从样本中进行有效的学习。Sigmoid 函数和 tanh 函数的另一个缺点是求导计算量很大,降低了实验的效率。

因此,本文提出一种使用 ReLU(rectified linear units)函数作为激活函数,代替传统的 Sigmoid 函数和 tanh 函数的 RCNN 模型。ReLU 函数是一个非饱和的非线性函数,其函数和导数曲线如图 5 所示, $f(x) = \begin{cases} 0, & x < 0 \\ x, & x > 0 \end{cases}$, 导

数 $f'(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0 \end{cases}$, 求导非常容易。ReLU 函数会使一部分神经元的输出为 0,这样就实现了网络的稀疏性,并减少了参数的相互依存关系,缓解了过拟合问题。其导数对于大于 0 的部分恒为 1,于是在反向传播的时候能够将梯

度很好的传到较前面的网络,减小了梯度弥散问题带来的影响,加快了网络训练的收敛速度。

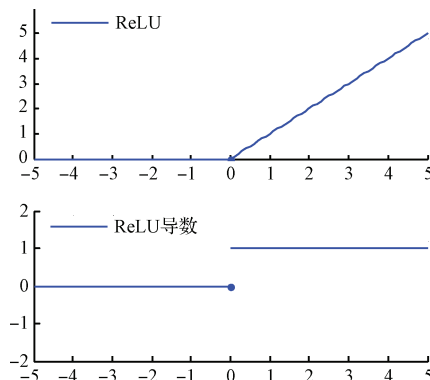


图 5 ReLU 函数及其导数曲线

3.3 基于矩阵乘法卷积展开技术的 CNN 模型(matrix multiple CNN, MMCNN)

当处理具有高分辨率的遥感图像时,尽管 CNN 模型具有权值共享的特点,但是由于其图像尺寸比常规手写字和自然图像大,对数据量大的样本完成卷积运算会消耗大量时间,这部分时间占全部网络运算时间的较高比例,从而导致了网络训练效率大幅降低^[15]。

所以,本文提出了一种基于矩阵乘法的卷积展开技术的优化 MMCNN 模型。此模型的核心思想是把卷积输入特征和卷积核分别展开,得到两个矩阵,通过矩阵的乘法完成卷积操作。输入特征展开得到的矩阵,其列数据是计算每个输出特性所需要的块状输入数据,卷积核在另一个矩阵展开,该矩阵的行数据由计算输出特性所需要的卷积核构成。如图 6 所示是一个简单的 MMCNN 模型实现的例子,简单起见,本例中输入特征图是通过 3 个卷积核做卷积,得到的 3 个输出特征图。通过对相同的数据进行 2 000 次循环计算,实验发现,矩阵乘法的方法比卷积方法快 5 倍左右。

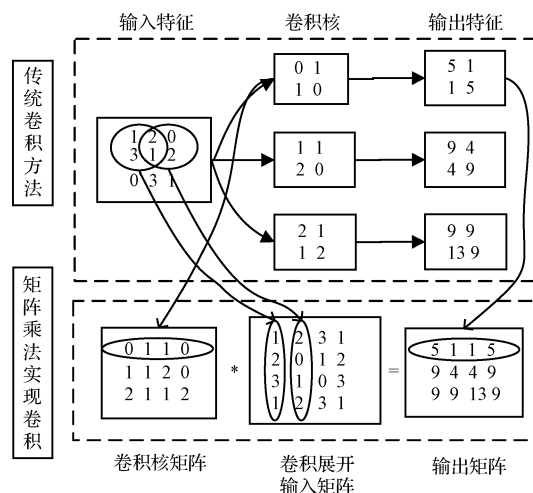


图 6 卷积展开例子

3.4 CNN 模型网络结构的优化

CNN 模型可以通过很多方法加以改进,比如将误差函数增加一个惩罚项来使得训练后的权值参数趋于稀疏。本文主要是针对卷积神经网络结构中学习率、网络层数、各层滤波器数量和大小对网络模型的训练过程和识别结果的影响。学习率的作用是控制每一步权值调整的幅度,学习率太小,收敛性容易得到保证,但收敛速度太慢;学习率太大,学习速度快,但可能导致震荡或发散。网络层数、各层滤波器数量和大小对网络的收敛情况和识别准确率都有影响。

4 实验与分析

4.1 实验数据集

利用从 Quick Bird 上获取的 0.6 m 分辨率的光学遥感图像进行实验,包括棒球场、城市居民区、灌木、立交桥和停车场 5 种目标,部分样本如图 7 所示。随机从带有标签的数据集中每类选出 2 500 幅,共 12 500 幅作为训练样本,5 000 幅作为测试样本,训练样本与测试样本不重复,所有图像的尺寸归一化为 100×100 像素。



图 7 部分样本图像

4.2 CNN 参数选择

针对 CNN 架构中学习率、网络层数、各层滤波器数量和大小对网络模型的训练过程和分类结果的影响。通过对比实验,选出最优的学习率,在这基础上,通过调节网络层数、卷积核大小及数量,选择平均错误率最小的网络结构,以提高目标识别的准确率。

4.2.1 学习率的影响

选择不同的学习率值对数据集进行实验,如图 8 所示为不同学习率对损失函数产生的影响。可以看出,当学习率为 0.01 时收敛速度最快而且收敛时损失函数达到最小,所以本实验选用 0.01 为学习率的值。

4.2.2 网络层数、各层滤波器数量和大小的影响

在上一实验选定的学习率的基础上,通过不同的卷积

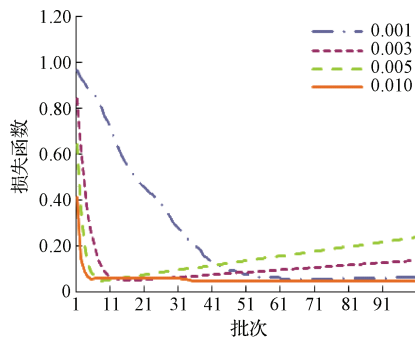


图 8 不同学习率对损失函数的影响

层、采样层层数、各层卷积核数量、卷积核大小对错误率影响的实验结果进行分析,输入图像大小均为 100×100 ,输出结果为 5 类。通过实验结果进行比较分析,选择最优的 CNN 网络结构进行高分率遥感图像数据集的目标识别,以提高目标识别的准确率。在学习率为 0.01 的条件下,通过选择卷积层和采样层共为 2 层、4 层、6 层和 8 层进行实验,每个卷积层后面接一个采样层,通过 200 次实验,计算平均错误率,实验结果如图 9 所示,通过实验结果可以看出,卷积层和采样层共 4 层时平均错误率最低。

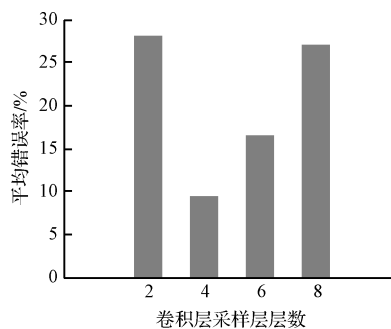


图 9 不同卷积层和采样层层数对错误率的影响

在保证卷积层和采样层共 4 层的结构基础上,通过调节卷积核的大小进行多组对比实验,实验结果如图 10 所示,分为 6 组实验,每组实验进行 10 次,横坐标为每组第一个卷积核尺寸,每组中不同颜色的系列表示第二个卷积核尺寸,纵坐标为平均识别准确率,可以看出卷积核大小分别为 21×21 和 17×17 时识别准确率最高。

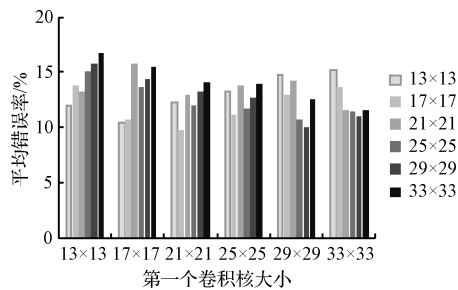


图 10 不同卷积核对识别错误率的影响

通过前3个实验,已经选出了最优的学习率、卷积层采样层层数和卷积核尺寸,在这些最优参数的基础上,通过调节卷积核的数量对结构进行继续优化,以提高目标识别的准确率。实验结果如图11所示,横坐标表示第一层卷积核数量,每组不同系列表示第二层卷积核数量,纵坐标为平均错误率,结果显示在其他参数确定的情况下,卷积核数量分别为6和12时错误率达到最低。

通过实验可以发现学习率为0.01,卷积层采样层层数为4,卷积核大小分别为 21×21 和 17×17 ,卷积核数量分别为6和12时,对于本文使用的高分辨率遥感图像数据集平均错误率最低,效果最好,所以选择这些参数的结构为本实验的CNN网络结构。

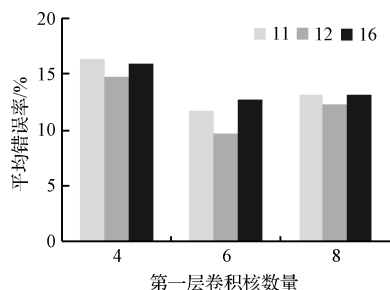


图11 不同卷积核数量对错误率的影响

4.3 识别结果

实验3.2.1和实验3.2.2已经初步确定了CNN模型的网络结构及参数,在此结构的基础上进行第2节所提出的改进方法,通过使用ReLU激活函数得到的RCNN模型、使用矩阵乘法(matrix multiple)实现卷积展开的MMCNN模型以及两个改进方法结合的RMMCNN。

表1所示为不同方法的识别准确率,其中HOG特征是用9个方向区域的VL_feat方法提取的特征,G为使用Matlab求梯度函数的方法提取的梯度特征,pixel(简称为p)为原图像的像素值,使用LibSVM方法进行分类。CNN方法利用10折交叉验证,其中训练集与测试集分别为12 500,5 000,共17 500个样本,测试集与训练集图像不重复。实验结果表明CNN的方法优于其他方法。

表1 不同方法的预测准确率

方法	SVM+HOG (%)	SVM+G (%)	SVM+p (%)	CNN (%)
1	88.01	66.25	62.12	89.12
2	87.23	64.38	63.58	91.45
3	86.41	65.45	62.64	92.24
4	87.75	65.35	64.05	89.97
5	89.96	66.75	62.84	91.48
6	88.46	65.46	63.02	92.81
7	87.92	64.35	62.55	89.62
8	89.74	65.64	61.79	90.98
9	87.36	64.67	61.88	91.43
10	89.53	64.78	63.23	92.86
平均准确率 (%)	88.237	65.31	62.77	91.196

为验证本文提出的改进方法的有效性,使用同一台计算机进行实验,处理器 Intel(R) Core(TM) i5-2400, CPU @3.1 GHZ,内存4 GB,通过采用ReLU激活函数和卷积展开技术使CNN的效率提高了10倍左右。

5 结论

本文提出一种基于CNN模型的高分辨率遥感图像目标识别方法,CNN模型避免了显式的特征提取,隐式地从训练数据中进行学习。这使得卷积神经网络明显有别于其他分类器,通过结构重组和减少权值将特征提取功能融合进多层感知器。它可以直接处理灰度图像,能够直接用于处理基于图像的分类,实现了目标的自动识别。

在CNN模型的基础上通过多组对比实验,选择了最优的学习率、网络层数、各层滤波器数量和尺寸,提高了目标识别的准确率。通过引入ReLU激活函数得到RCNN模型,使用矩阵乘法实现卷积展开技术得到MMCNN模型。实验表明,该方法可取得优于SVM的目标识别准确度,并且通过改进提高了CNN模型训练和测试的效率。在下一步工作中,将继续对模型进行改进,以进一步提高遥感图像目标识别的准确率。

参考文献

- [1] 张楚金,王耀南,卢笑,等. 基于假设验证和改进HOG特征的前车检测算法[J]. 电子测量与仪器学报, 2015, 29(2):165-171.
- [2] 张国刚,徐向辉. 基于加权纹理特征的SAR图像目标识别算法[J]. 国外电子测量技术, 2015, 34(9): 22-25.
- [3] LIU G, SUN X, FU K, et al. Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior[J]. Geoscience and Remote Sensing Letters, 2013, 10(3): 573-577.
- [4] 瞿继双,瞿松柏,王自杰. 基于特征的模糊神经网络遥感图像目标分类识别[J]. 遥感学报, 2009, 13(1): 67-74.
- [5] FRIEDMAN N, GEIGER D, GOLDSZMIDT M. Bayesian network classifiers[J]. Machine learning, 1997, 29(2-3): 131-163.
- [6] 吴哲君,赵忠华,唐雷. 基于SVM的行人步态实时分类方法[J]. 电子测量技术, 2015,38(7):41-44.
- [7] 赵春晖,齐滨. 基于模糊核加权C-均值聚类的高光谱图像分类[J]. 仪器仪表学报, 2012, 33(9): 2016-2021.
- [8] 高强,阳武,李倩. DBN层次趋势研究及其在航拍图像故障识别中的应用[J]. 仪器仪表学报, 2015, 36(6):1267-1274.
- [9] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Pro-

- ceedings of the IEEE, 1998, 86(11): 2278-2324.
- [10] LAUER F, SUEN C Y, BLOCH G. A trainable feature extractor for handwritten digit recognition[J]. Pattern Recognition, 2007, 40(6): 1816-1824.
- [11] LAWRENCE S, GILES C L, TSOI A C, et al. Face recognition: A convolutional neural-network approach[J]. Neural Networks, IEEE Transactions on, 1997, 8(1): 98-113.
- [12] CIRESAN D C, MEIER U, MASCI J, et al. Flexible, high performance convolutional neural networks for image classification[C]. IJCAI Proceedings-International Joint Conference on Artificial Intelligence, 2011, 22(1): 1237.
- [13] TANG J, DENG C, HUANG G B, et al. Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(3): 1174-1185.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]. Advances in neural information processing systems, 2012: 1097-1105.
- [15] 刘进锋. 一种简洁高效的加速卷积神经网络的方法[J]. 科学技术与工程, 2014(33): 240-244.

作者简介

曲景影, 1990年2月出生, 硕士研究生。主要研究方向为遥感图像解译、机器学习。
E-mail: qjy1990@163.com

(上接第44页)

- [13] 范思航, 时文娟, 黄艳芝, 等. 超声波多普勒流量计换能器的研究与应用[J]. 国外电子测量技术, 2014, 33(2): 84-88.
- [14] 薛敏彪, 巨欢, 张炳军, 等. 利用部分相关法扩展多普勒容限的方法[J]. 国外电子测量技术, 2014, 33(6): 41-47.
- [15] 凌祥, 王夷, 赵雷鸣, 等. 二维距离-多普勒海杂波信号模拟与实现方法[J]. 电子测量技术, 2014, 37(5): 108-111.
- [16] 郑志霞, 冯勇建. MEMS接触电容式高温压力传感器的温度效应[J]. 电子测量与仪器学报, 2013, 27(12): 1141-1147.
- [17] 郭兴军, 李朋伟, 张文栋, 等. 基于电容式 MEMS 器件的静电斥力驱动研究[J]. 仪器仪表学报, 2014, 35(8): 1738-1743.
- [18] 王少阳, 张瑞峰, 孙晓伟, 等. 基于 FPGA 的 MEMS 测量控制系统的设计[J]. 电子测量技术, 2014, 37(12): 58-61.

作者简介

朱赛健, 1993年出生, 男, 江苏启东人, 硕士, 主要研究方向武器系统与运用工程。
E-mail: zhusaijiannwpu@163.com