

小样本的多模态遥感影像高层特征融合分类

贺琪¹, 李瑶¹, 宋巍¹, 黄冬梅^{1,2*}, 何盛琪¹, 杜艳玲¹

¹上海海洋大学信息学院, 上海 201306;

²上海电力大学, 上海 200090

摘要 在使用深度学习模型研究遥感影像地物分类问题时,某些地物的遥感影像可用于训练的样本很少。同时,多样化的遥感影像获取方式产生了大量不同空间分辨率的多模态遥感影像。融合这些多模态遥感影像,弥补样本量少导致分类精度低的缺陷,是小样本的遥感影像高精度分类领域中亟待解决的问题。针对上述问题,提出了考虑两种空间分辨率遥感影像相关关系的融合分类方法。首先,使用两个并行的深度学习网络分别提取两种空间分辨率影像的高层特征;其次,将提取到的高层特征通过融合方法进行融合;最后,得到融合后的高层特征作为输入,训练整个融合分类模型。实验表明,不同融合策略的分类精度不同,本文提出的基于高层特征级别的融合策略可以有效提高分类精度。

关键词 图像处理;深度学习;高层特征融合;多模态遥感影像;小样本

中图分类号 TP389.1

文献标识码 A

doi: 10.3788/LOP56.111001

Multimodal Remote Sensing Image Classification with Small Sample Size Based on High-Level Feature Fusion

He Qi¹, Li Yao¹, Song Wei¹, Huang Dongmei^{1,2*}, He Shengqi¹, Du Yanling¹

¹ College of Information Technology, Shanghai Ocean University, Shanghai 201306, China;

² Shanghai University of Electric Power, Shanghai 200090, China

Abstract The training sample size for some objects on the ground is quite small when applying a deep learning model to study the classification of remote sensing images. Meanwhile, diversified remote sensing image acquisition methods generate numerous multimodal remote sensing images with different spatial resolutions. Fusing these multi-modal remote sensing images to remedy the small sample size defect and achieve a highly precise classification of remote sensing images is an urgent problem to be solved. To this end, the present study proposes a fusion method for image classification based on the correlation of two spatial resolutions. A deep learning network is utilized to extract the high-level features of the remote sensing images in two spatial resolutions. Two types of high-level features are integrated via the proposed fusion strategy and further used as the input to train the whole network model. The experimental results demonstrate that the proposed fusion algorithm can achieve high classification accuracy. Further, because different fusion rules have different classification accuracies, a suitable selection can improve the classification accuracy.

Key words image processing; deep learning; high-level feature fusion; multimodal remote sensing image; small sample size

OCIS codes 100.4996; 100.3008; 100.4992

1 引言

遥感分类技术已在城市环境、精准农业、交通道

路设施、林业测量、军事目标识别和灾害评估等领域存在广泛应用。现今流行的分类方法包括决策树^[1]、支持向量机^[2]、条件随机场^[3]、马尔可夫随机

收稿日期: 2018-11-07; 修回日期: 2018-12-05; 录用日期: 2018-12-25

基金项目: 国家自然科学基金(41671431)、海洋大数据分析预报技术研发基金项目(2016YFC1401902)、上海市科委部分地方院校能力建设项目(17050501900)

* E-mail: dmhuang@shou.edu.cn

场^[4]和深度卷积神经网络(CNN)。其中,CNN在ILSVRC(large scale visual recognition challenge)上的应用效果显著^[5],因此备受关注。

CNN在遥感影像分类领域涌现了大量的研究成果。文献[6]中针对高光谱影像分类进行研究,使用加权的K近邻算法提取特征空间,利用特征空间组成三维张量的谱空联合信息训练三维卷积神经网络。文献[7]中采用八层卷积神经网络对白化后的影像进行分类。文献[8]中针对高光谱影像具有数据纬度高、标记样本少和计算大等问题,提出了谱空联合分类方法,该方法有效减少了只有少量样本时的错分概率,但在近邻协同过程中会引入异类地物,干扰实验精度。使用CNN方法研究具体地物的遥感影像分类问题时,可用于训练的样本总是呈现只包含几百张遥感影像的现象。面对CNN方法中需要训练的参数规模庞大的问题,小样本的特性将导致分类准确率低。

为了解决以上小样本导致的分类准确率低的问题,本文提出了一种小样本的多模态遥感影像高层特征融合分类方法——MRSIC。该方法将深度学习算法作为两种分辨率遥感影像的特征提取器,该特征提取器具有强大的抽象能力,使得提取到的高层特征具有高质量的特征表达。提取到高层特征后,通过该方法提出融合策略进行融合,最后,联合信息熵算法和梯度下降算法,使用融合后的综合特征作为输入,训练整个融合分类模型。为了验证该方法的有效性,使用两种不同分辨率模态的4个类别影像进行实验,检验分类方法的稳健性和扩展性。为了验证该方法中的两大优点:1)两种不同模态数据互补,可有效改善训练样本少导致的分类准确率低的问题;2)特征级别融合策略使模态之间的干扰性大大降低。实验与LeCun等^[9]提出的经典单模态网络、Simonyan等^[10]提出的单模态网络结构以及基于多模态的像素级融合策略^[11]进行对比,以验证MRSIC的有效性和合理性。

2 相关工作

2.1 基于小样本的分类方法

解决遥感影像训练数据少这一缺陷的方法分为两大类:基于迁移学习的影像分类和基于多模态融合的影像分类。文献[12]中提出了基于深度神经网络的迁移学习算法,使用英语数据集训练得到源领域模型,将其隐藏层作为特征提取器,模型的前几层网络可以促使不同的语言相互转移,实现一个仅用

小样本量就可以进行高质量语言翻译的模型。文献[13]中将卷积神经网络的迁移学习用于图像分类,采用贝叶斯正则化算法对网络模型中的超参数进行调优,得到基于小样本量的图像分类模型。文献[14]中提出了基于卷积神经网络的分级前反馈模型,在没有监督的情况下使用数据自动构建一个伪任务,在迁移训练过程中将有用的先验知识用于更新网络参数。使用迁移学习研究影像分类问题,可以有效降低模型训练的时间复杂度,但现实中很难寻找到与目标领域数据源高度吻合的源领域数据源,而且在计算特征空间和特征之间的距离时,特征变动会在很大程度上影响分类模型的精准度。

利用多模态遥感影像研究分类问题越来越受到关注,文献[15]中构造了交叉内核函数的分类方法,将不同时间分辨率的遥感数据集映射到同一个特征空间,该方法有效提高了分类模型在不同模态之间遥感影像分类的通用性。文献[16]中利用贝叶斯非参数的字典学习与改进的稀疏编码技术,构造出高分辨率的遥感影像,联合使用构造的遥感影像和原始影像进行组织实验。文献[17]中提出了基于马尔可夫方法的关系模型,将影像数据的几何结构信息关联起来,该方法综合考虑了模态本身特征及模态之间的相关性。

2.2 多模态融合分类方法

多种模态信息互补融合的方法在RGB-D(depth map)数据出现时具有很好的效果,文献[18]中综述了三维(3D)影像数据的特征,阐述了深度影像和RGB影像互补的优势,在对影像特征提取与分类应用中取得了显著的成绩。文献[19]中采用池化层与反卷积层两种模态信息融合的方式增加影像自身特征,减少了池化过程中自身特征的丢失。文献[20]中提出了一种融合光谱特征和几何特征的建筑物提取算法,实现了较高精度的建筑物提取。

图1中cov为卷积运算,pool为下采样池化运算,ful为全连接运算,softmax为分类器。常用的多模态融合策略分为基于像素级别、亚像素级别、特征级别和决策级别4种类别,采用多模态融合策略能够使不同模态数据之间的信息互补,以达到高分类精度的目标。文献[21]中采用如图1(a)所示的基于像素级别的融合策略,将两种模态的遥感影像预处理为一种模态的影像,然后将预处理的影像作为深度学习网络的训练集,该网络结构忽略了不同模态遥感影像在训练网络参数时模态信息相互干扰的风险;文献[22]中采用如图1(b)所示的基于特征级别的融合策略,使用两路深度学习网络分别训练

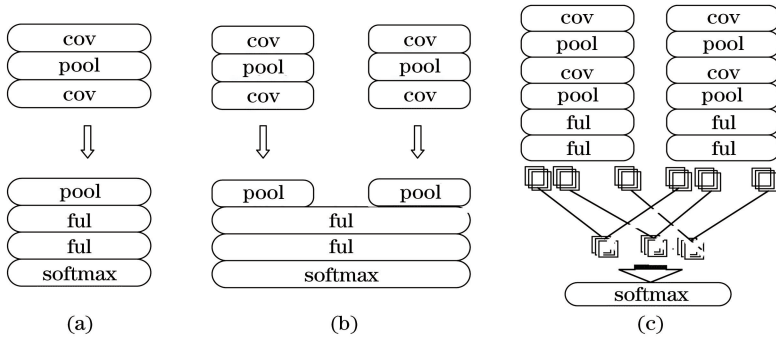


图 1 多模态深度学习网络。(a)像素级融合策略;(b)特征级融合策略;(c)高层特征级融合策略

Fig. 1 Multimodal deep learning network. (a) Pixel-level fusion strategy; (b) feature-level fusion strategy; (c) high-level feature-level fusion strategy

两种模态的遥感影像,将网络中的全连接层作为两种模态遥感影像融合层,但该网络结构没有充分考虑两个模态之间的相关性;为了克服以上难题,本文提出了如图 1(c)所示的 MRSIC 分类方法。

3 MRSIC

遥感影像可用于训练的数据少,导致训练集呈小样本的特性,而遥感影像的多模态特征给小样本分类问题的解决开辟了新的研究方向。为了减少多模态融合时的干扰性,提出了基于高层特征级融合的 MRSIC 方法,基本思想是将两种空间分辨率遥感影像的高层特征通过训练成熟的深度网络模型提取出来,然后基于该高层特征进行融合,实现信息互补,从而提高分类精度。

3.1 高层特征提取器

为获得高质量的多模态遥感影像高层特征,分别针对高分辨率数据集 $X = \{x_1, x_2, \dots, x_{1000}\}$ 和 low 分辨率数据集 $Y = \{y_1, y_2, \dots, y_{1000}\}$ 构建两个

平行的深度网络模型:高分辨率特征提取器(HRFE)和低分辨率特征提取器(LRFE)。高分辨率是指在遥感图像上能够区分的最小单元的尺寸小,可以更好地表征影像分辨地面目标细节。高分辨率影像具有明显的地物纹理特征,低分辨率影像有明显的地物轮廓特征,合理有效地融合两种模态可获得既有纹理信息又具有轮廓信息的高层特征。该模型的网络结构信息如表 1 所示。表中:ReLU 为激活函数,该函数前向传播计算简单、求解偏导数也简单,最重要的是不容易发生梯度发散问题,提高了模型的学习能力;* 为卷积运算, K_{cov1_k} 为第 k 个卷积核, f 为池化方法, W_1, W_2 为两种模态相关性的权重值, $b_{1,2,3,\dots,7}$ 为网络偏置, \pm 为将两种模态的高层特征映射到同一个特征空间;全连接层特征映射图 X'_{pool2_w} 为 X'_{pool2} 的第 w 行; X'_{ful1} 为融合前的高分辨率遥感影像的高层特征, Y'_{ful1} 为融合前的低分辨率遥感影像的高层特征; K_{ful1} 为第二层全连接的映射参数。

表 1 特征提取器的网络结构

Table 1 Network structure of feature extractor

Network structure	Formula	Output	Feature size
cov1	$X'_{cov1_k} = \text{ReLU}(K_{cov1_k} * x_i + b_1)$	$\{X'_{cov1_k}\}_{k=1}^{16}$	208×208
pool1	$X'_{pool1} = f(X'_{cov1_k} + b_2)$	X'_{pool1}	104×104
cov2	$X'_{cov2_k} = \text{ReLU}(K_{cov2_k} * X'_{pool1} + b_3)$	$\{X'_{cov2_k}\}_{k=1}^{16}$	104×104
pool2	$X'_{pool2} = f(X'_{cov2_k} + b_4)$	X'_{pool2}	104×104
ful1	$X'_{ful1} = X'_{pool2_1} \pm X'_{pool2_2} \pm \dots \pm X'_{pool2_{104}} + b_5$	X'_{ful1}	1×173056
Z'_{ful1}	$Z'_{ful1} = W_1 X'_{ful1} \pm W_2 Y'_{ful1} + b_6$	Z'_{ful1}	128×128
ful2	$Z'_{ful2} = Z'_{ful1} \times K_{ful1} + b_7$	Z'_{ful2}	1×128

在高层特征提取器中,卷积层采用的卷积核尺寸和深度为 3 pixel \times 3 pixel 和 3 pixel,采样间隔为 1 pixel,边界全 0 填充 1 pixel。池化层采用的下采样滤波器尺寸为 2 pixel \times 2 pixel,长和宽的步长均为 2 pixel。通过本小节的两路深度网络模型可以提

取到两种模态数据高质量的高层特征。

3.2 高层特征融合策略

融合两种模态数据时会出现信息干扰的问题,为了避免信息干扰导致的分类精度低的问题:1) 采用 3.1 节提出的网络模型分别提取两种模态遥感影

像的高层特征并融合,两种模态遥感影像的高层特征提取器中参数完全独立,极大地降低了模态间的干扰;2) 构建融合策略,建立两种模态间的相关性,从而实现模态间信息互补增强;3) 联合信息熵算法和梯度下降算法对整个分类模型的训练进行反馈调节。

MRSIC 在每次迭代循环结束后,依据信息熵和梯度下降算法的优化机制更新 MRSIC 参数。

1) 输入: 高分辨率数据集 X 和低分辨率数据集 Y 的遥感影像训练集。

2) 输出: 高层特征提取器的模型参数、分类精度最高的分类器。

3) 将两种分辨率的遥感影像 $Z(x, y \in Z)$ 读入到对应的高层特征提取器中,得到影像的高层特征,即 X'_{full} 和 Y'_{full} 。

4) 构建基于特征级别的融合策略: $Z'_{full} = W_1 X'_{full} \pm W_2 Y'_{full} + b_6$, 其中 $W_1 + W_2 = 1$ 为两种模态之间的相关性的权重参数。

5) 采用信息熵算法 $H(p, q) = - \sum_z p(z) \log q(z)$ 测量分类精度, 式中: $q(z)$ 为样本的真实概率分布; $p(z)$ 为样本的预测概率分布。该式将两个概率分布间距降至最小, 通过反向传播算法对网络

进行训练。

6) 使用梯度下降算法 $W^{(n+1)} = W^{(n)} - \delta \frac{\partial}{\partial W^{(n)}}$

$\text{softmax}(W^{(n)})$ 优化调节模型参数, 式中: δ 为设置的学习率; $W^{(n+1)}$ 为 softmax 分类器的参数经历过 $n+1$ 次迭代后的数值。

本文采用的网络结构优化分两个阶段, 第一个阶段通过前向传播算法计算得到预测值, 并通过信息熵算法计算预测值与真实值的概率分布距离。第二个阶段将信息熵算法作为目标函数, 使用反向传播算法计算每个参数的梯度, 再结合梯度、学习率和梯度下降更新每一个参数。

4 实验结果与分析

为了验证 MRSIC 的有效性、稳健性和扩展性, 进行 3 组对比实验。实验数据包括高分辨率的 NWPU VHR (northwestern poly-technique university and very-high-resolution remote sensing images)^[23] 和低分辨率的 UC (University of California) Merced Land Use^[24], 如图 2 所示, 共 2 种模态 4 种类别, 分别为棒球场 (bd)、高尔夫球场 (golf)、网球场 (te) 和仓库 (st)。

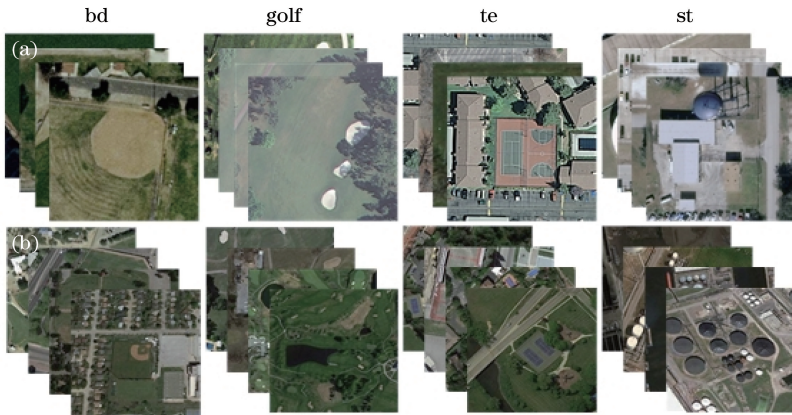


图 2 多模态遥感影像数据集。(a) 高分辨率遥感影像; (b) 低分辨率遥感影像

Fig. 2 Multimodal remote sensing image datasets. (a) High-resolution remote sensing images; (b) low-resolution remote sensing images

使用 MRSIC 方法研究多模态影像分类时, 影响分类效果的因素主要有网络学习能力和融合策略两个部分。其中, 影响网络学习能力的因素有网络模型中卷积核的个数和影响梯度下降算法的训练批次大小; 影响融合策略的因素包括融合权值、多模态匹配规则以及融合规则。

4.1 模型参数分析

学习率和训练批次的大小影响网络模型是否可

以找到相对最优解。寻找最优解是一个热点问题。当以全数据集作为一个批次时, 调参的方向能够更好地代表样本总体, 从而更准确地朝极值所在的方向调节。但由于数据集海量增长, 内存受到限制, 一次性载入所有数据变得越来越困难。为此使用批梯度下降算法来获得相对最优解。

卷积核和训练批次的大小对分类精度以及训练时间的影响如表 2 所示, 当以卷积核数目作为因变

表2 MRSIC模型参数设置对分类精度的影响
Table 2 Influence of MRSIC model parameter setting on classification accuracy

Parameter	8 kernels	16 kernels	32 kernels	8/ batch	16/ batch	32/ batch
Time /s	1356.9	2456.2	4828.1	1431.4	2452.2	4418.32
Accuracy / %	95.3	90	92.6	50	90	82.2

量,其他参数不变时,采用8个卷积的分类精度达到95.3%,而16个卷积核的精度最低。深度学习算法没有最优解,关注目标是使用何种学习算法在数据生成分布上的效果最好,卷积核的数量与学习能力成正比,故在MRSIC中选用16个卷积核。当设置批次大小为因变量,其他条件不变时,每个批次为16个样本时的分类精度最高,而当设置8个样本时,分类模型发散。当设置的批次样本数偏大时,所需迭代全数据集的次数亦增大,否则影响分类精度。

考虑到迭代次数、数据处理速度和局部极值的获取以及卷积核个数对网络模型计算能力的影响,为使模型具有理想的局部最优值、时效性和复杂函数的学习能力,MRSIC训练批次的数据个数为16、设置16个卷积核以及大小为0.0001的学习率。

4.2 融合策略分析

为说明融合策略具有抗干扰性、小样本适用性以及模态间信息互补性,共给出5组对比实验,网络结构分别为THRSI (trained by high-resolution remote sensing image)-LeNet^[9]、TLRSI (trained by low-resolution remote sensing image)-LeNet、纳米神经网络(NanoNet)^[25]、VGG^[10] (visual geometry group)网络和MRSIC,如表3所示。其中,THRSI-LeNet、TLRSI-LeNet、NanoNet和VGG为单模态深度学习网络;NanoNet为基于迁

表3 7种分类模型分类精度

Table 3 Accuracy of seven classification models

Network structure	Dataset	Amount	Accuracy /%
THRSI-LeNet	NW	1000	83.2
THRSI-LeNet	NW	500	74.6
TLRSI-LeNet	UC	1000	77.3
TLRSI-LeNet	UC	500	67.2
H&L-LeNet	NW,UC	1000	82.5
HUL-LeNet	NW,UC	1000	72.0
NanoNet	UC	500	77.1
VGG	NW,UC	500	70.4
In Fig. 1(b)	NW,UC	1000	85.0
MRSIC	NW,UC	1000	90.5
MRSIC	NW,UC	500	90.0

移学习的在线分类服务,采用的深度模型结构为VGG。

对比单模态的网络模型之间的分类精度情况可知,当样本量增加一倍时,相同的网络模型分类精度提高8.6%~10.1%,与此同时,使用在线网络服务的迁移学习NanoNet(500张训练数据)比使用相同数据的TLRSI-LeNet提高了9.9%的分类精度。但使用MRSIC中的融合策略时,小样本量(500张训练数据)即可达到单模态网络模型(1000张训练数据)的分类精度,同时,在样本量增加的情况下,分类精度仅提高了0.5%~2.5%,相比增加了一倍样本量的单模态网络模型,所提出的两种融合分类方法具有更强、更稳定的分类能力。

为说明融合算法具有抗干扰能力,表3共给出了5组多模态融合策略的对比实验,网络结构包括H&L (trained by high-resolution and low-resolution remote sensing image)-LeNet、HUL (trained by high-resolution U low-resolution remote sensing image)-LeNet、VGG、图1(b)中所示结构与MRSIC。其中,H&L-LeNet、HUL-LeNet使用网络结构如图1(a)所示,融合策略在数据预处理阶段进行,前者将两种3通道的模态数据叠加为6通道的模态数据,后者则将两种3通道的数据直接并接形成一种模态的数据;VGG使用与H&L-LeNet一样预处理的数据集。在使用融合策略的分类方法中MRSIC模型的融合策略分类精度最高,相比于其他策略分类精度提高了4.5%~18.5%。

为证明MRSIC方法的稳健性和扩展性,表4给出了3组对比实验,直接将MRSIC方法应用到不同分辨率的网球场和仓库数据集上。结果表明:该方法相比于使用单模态的THRSI-LeNet分类精度提高了13.2%;MRSIC应用在4种类别分类时,相较于THRSI-LeNet分类精度提高了11.3%。

表4 MRSIC稳健性验证实验

Table 4 Robustness verification of MRSIC

Category	MRSIC accuracy /%	THRSI-LeNet accuracy /%
bd, golf	90	74.6
te, st	87	73.8
bd, golf, te, st	75.3	64

综上,H&L-LeNet、HUL-LeNet以及VGG的融合策略由于未充分考虑两种模态间的相互关系,信息之间有互补性和互扰性,分类准确率不高。

相比于分类精度有较大提高的图 1(b)融合策略, MRSIC 中的融合策略具有更有效的抗干扰能力和更强、更稳定的分类能力。同时,表 4 说明 MRSIC 的稳健性和扩展性具有可信度。

5 结 论

提出了面向小样本的多模态遥感影像融合分类方法,其核心思想是构造融合算法,利用融合策略将两种分辨率的遥感影像高层特征融合,该融合策略反映的是两种模态的相关性。由于不同模态的遥感影像含有自身对地观测的特有信息,采用融合策略训练出来的分类器具有更强的学习能力,从而提高了目标分类精度。所提方法在 NWPU VHR 和 UC Merced Land Use 数据集上获得了 90.5% 的分类精度,有效实现了 CNN 模型在小样本量训练集上的应用,为多模态遥感影的分类提供了新的思路。后续研究方向将包括:

1)在融合算法中,当两种模态的数据量相差很大时,如何充分使用全部数据集来增强模型的分类能力,获得高精度的通用模型。

2)在深度学习的基础上,进一步研究由不同时间分辨率、不同空间分辨率、不同角度分辨率以及不同成像原理形成的多种模态信息对分类器的影响,以取得更好效果。

3)统筹使用迁移学习和多模态融合算法,共同致力于提高模型的分类精度。

参 考 文 献

- [1] Tanha J, van Someren M, Afsarmanesh H. Semi-supervised self-training for decision tree classifiers [J]. *International Journal of Machine Learning and Cybernetics*, 2017, 8(1): 355-370.
- [2] Guo Y, Jia X, Paull D. Sequential classifier training for rice mapping with multitemporal remote sensing imagery [J]. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, IV-4/W2: 161-165.
- [3] Xia M, Cao G, Wang G Y, *et al.* Remote sensing image classification based on deep learning and conditional random fields [J]. *Journal of Image and Graphics*, 2017, 22(9): 1289-1301.
夏梦, 曹国, 汪光亚, 等. 结合深度学习与条件随机场的遥感图像分类 [J]. *中国图象图形学报*, 2017, 22(9): 1289-1301.
- [4] Cao X Y, Xu L, Meng D Y, *et al.* Integration of 3-

dimensional discrete wavelet transform and Markov random field for hyperspectral image classification [J]. *Neurocomputing*, 2017, 226: 90-100.

- [5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [6] Liu Y T, Li Z Q, Yang X L. Application of improved convolution neural network in remote sensing image classification [J]. *Journal of Computer Applications*, 2018, 38(4): 949-954.
刘雨桐, 李志清, 杨晓玲. 改进卷积神经网络在遥感图像分类中的应用 [J]. *计算机应用*, 2018, 38(4): 949-954.
- [7] Yang H, Li L Q, Yang H H, *et al.* Method of urban management cases' image classification based on convolutional neural network [J]. *Computer Engineering and Applications*, 2018, 54(10): 242-248, 266.
杨浩, 李灵巧, 杨辉华, 等. 基于卷积神经网络的城管案件图像分类方法 [J]. *计算机工程与应用*, 2018, 54(10): 242-248, 266.
- [8] Fu G Y, Gu H Y, Wang H Q. Spectral and spatial classification of hyperspectral images based on convolutional neural networks [J]. *Science Technology and Engineering*, 2017, 17(21): 268-274.
付光远, 辜弘炆, 汪洪桥. 基于卷积神经网络的高光谱图像谱-空联合分类 [J]. *科学技术与工程*, 2017, 17(21): 268-274.
- [9] LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [10] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2018-10-20]. <https://arxiv.org/abs/1409.1556>.
- [11] Eitel A, Springenberg J T, Spinello L, *et al.* Multimodal deep learning for robust RGB-D object recognition [C] // 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 28-October 2, 2015, Hamburg, Germany. New York: IEEE, 2015: 15666832.
- [12] Xu Y, Du J, Dai L R, *et al.* Cross-language transfer learning for deep neural network based speech enhancement [C] // The 9th International Symposium on Chinese Spoken Language Processing, September 12-14, 2014, Singapore, Singapore. New York:

- IEEE, 2014: 14700579.
- [13] Han D M, Liu Q G, Fan W G. A new image classification method using CNN transfer learning and web data augmentation [J]. *Expert Systems With Applications*, 2018, 95: 43-56.
- [14] Ahmed A, Yu K, Xu W, *et al.* Training hierarchical feed-forward visual recognition models using transfer learning from pseudo-tasks[M]//Forsyth D, Torr P, Zisserman A. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2008: 69-82.
- [15] Camps-Valls G, Gomez-Chova L, Munoz-Mari J, *et al.* Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(6): 1822-1835.
- [16] Ding X, Jiang Y, Huang Y, *et al.* Pan-sharpening with a Bayesian nonparametric dictionary learning model [C] // *Artificial Intelligence and Statistics*. [S.l.]: [s.n.], 2014: 176-184.
- [17] Hane C, Zach C, Cohen A, *et al.* Joint 3D scene reconstruction and class segmentation [C] // 2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. New York: IEEE, 2013: 13824292.
- [18] Fang X, Wang G H, Yang H C, *et al.* High resolution remote sensing image classification combining with mean-shift segmentation and fully convolution neural network [J]. *Laser & Optoelectronics Progress*, 2018, 55(2): 022802.
方旭, 王光辉, 杨化超, 等. 结合均值漂移分割与全卷积神经网络的高分辨遥感影像分类[J]. *激光与光电子学进展*, 2018, 55(2): 022802.
- [19] Tu S Q, Xue Y J, Liang Y, *et al.* Review on RGB-D image classification [J]. *Laser & Optoelectronics Progress*, 2016, 53(6): 060003.
涂淑琴, 薛月菊, 梁云, 等. RGB-D 图像分类方法研究综述[J]. *激光与光电子学进展*, 2016, 53(6): 060003.
- [20] He M Y, Cheng Y L, Liao X J, *et al.* Building extraction algorithm by fusing spectral and geometrical features [J]. *Laser & Optoelectronics Progress*, 2018, 55(4): 042803.
何曼芸, 程英蕾, 廖湘江, 等. 融合光谱特征和几何特征的建筑物提取算法[J]. *激光与光电子学进展*, 2018, 55(4): 042803.
- [21] Couprie C, Farabet C, Najman L, *et al.* Indoor semantic segmentation using depth information[EB/OL]. (2013-03-14) [2018-10-20]. <https://arxiv.org/abs/1301.3572>.
- [22] Wang W, Yang X Y, Ooi B C, *et al.* Effective deep learning-based multi-modal retrieval[J]. *The VLDB Journal*, 2016, 25(1): 79-101.
- [23] Cheng G, Han J W. A survey on object detection in optical remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 117: 11-28.
- [24] Cheng G, Zhou P C, Han J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(12): 7405-7415.
- [25] Zhou R T, Li Z P, Wu C, *et al.* Buddy routing: a routing paradigm for NanoNets based on physical layer network coding [C] // 2012 21st International Conference on Computer Communications and Networks (ICCCN), July 30-August 2, 2012, Munich, Germany. New York: IEEE, 2012: 12965366.