

引用格式:Cui Xianliang, Chen Lifu, Xing Xuemin, *et al.* Remote Sensing Image Scene Classification based on Frequency Band Feature Fusion and GL-CNN[J]. Remote Sensing Technology and Application, 2019, 34(4): 712-719.
[崔先亮, 陈立福, 邢学敏, 等. 基于频带特征融合的GL-CNN遥感图像场景分类[J]. 遥感技术与应用, 2019, 34(4): 712-719.]
doi:10.11873/j.issn.1004-0323.2019.4.0712

基于频带特征融合的GL-CNN遥感图像场景分类

崔先亮, 陈立福, 邢学敏, 袁志辉

(长沙理工大学电气与信息工程学院, 湖南 长沙 410114)

摘要: 高分辨率卫星遥感图像场景信息的分类对影像分析和解译具有重要意义, 传统的高分辨卫星遥感图像场景分类方法主要依赖于人工提取的中、低层特征且不能很好的利用图像丰富的场景信息, 针对这一问题, 提出一种基于频带特征融合与GL-CNN(Guided Learning Convolutional Neural Network, 指导学习卷积神经网络)的分类方法。首先通过NSWT(Non-Subsampled Wavelet Transform, 非下采样小波变换)提取出图像的高低频子带, 将高频子带进行频带特征融合得到融合高频子带, 然后联合频谱角向能量分布曲线的平稳区间分析实现融合高频子带与低频子带的样本融合, 最后指导卷积神经网络自动提取图像的高低频子带包含的高层特征来实现场景分类。通过对UCM_LandUse 21类数据进行试验表明, 本文方法的分类正确率达到94.52%, 相比以往算法有显著提高。

关键词: 非下采样小波变换; 频带特征融合; 指导学习; 样本融合; 场景分类

中图分类号: TP753 **文献标志码:** A **文章编号:** 1004-0323(2019)04-0712-08

1 引言

随着卫星技术的蓬勃发展, 遥感卫星获取的卫星图像分辨率越来越高, 标志着遥感领域已经进入了高分辨率时代^[1]。高分辨率遥感影像具有的空间与纹理特征也越来越丰富, 而在这些特征中包含了大量的场景语义信息。对场景信息的分类是语义分割和目标识别过程的分类基础, 提高场景分类的准确率能使目标识别和语义分割有所突破, 但由于图像的场景构成十分复杂, 一类场景中包含了多类目标, 如何准确地学习每类样本的主要特征成为极具挑战的课题, 已引起了遥感学术界的广泛关注^[2-3]。学者们先后提出了很多场景分类方法, 大致可以分为两类: Oliva等^[4]采用的基于低层次特征的

Gist方法和Sivic等^[5]基于中层语义特征建模的视觉词袋模型(BoVM)方法。但是中、低层人工特征不能很好地利用高分辨率卫星遥感图像丰富的场景信息。探究能够表达遥感图像高层次抽象特征的分类方法成为该领域的研究热点。而深度学习的出现有效地解决了这一问题, 深度学习使用卷积神经网络能够提取图像的高层次特征, 更好地表达遥感图像。

目前深度学习领域主要采用3种方法来进行场景分类^[6]: ① Castelluccio等^[7]采用预训练网络(OverFeat、GoogLeNet)直接对遥感图像进行分类, 由于预训练网络一般的训练样本为普通光学图像, 所以在遥感图像上的分类缺少针对性; ② Nagueira

收稿日期: 2018-05-23; 修订日期: 2018-10-06

基金项目: 国家自然科学基金青年基金项目“基于机载双天线InSAR系统三维地形实时获取算法研究”(41201468), 湖南省教育厅项目“高分辨率SAR图像复杂背景下高精度鲁棒的道路提取算法研究”(16B004), 国家自然科学基金青年基金项目“顾及流变参数的时序InSAR公路形变模型研究”(41701536), 国家自然科学基金青年基金项目“面向复杂地形的多通道干涉SAR高精度DEM稳健反演研究”(61701047)。

作者简介: 崔先亮(1994—), 男, 湖南常德人, 硕士研究生, 主要从事图像处理、深度学习、遥感图像解译研究。E-mail: 421428120@qq.com。

通讯作者: 陈立福(1979—), 男, 河北沧州人, 博士, 讲师, 主要从事遥感图像解译、深度学习研究。E-mail: lifu_chen@139.com。

等^[8]采用在预训练网络上使用部分遥感数据进行微调;然后再对遥感图像进行测试的方法,这种微调网络的方法虽然能使网络学习的特征倾向于遥感图像而使得分类的效果优于第一种方法,但是遥感图像的数据量难以微调整个庞大的预训练网络,因而图像特征并没有得到更深层次的学习;③Luus等^[9]针对遥感图像特点构建新的网络,再用遥感图像对网络进行训练和测试,但由于样本图像数量不足仅能使用浅层网络,同样导致了网络缺乏学习遥感图像高层次特征的能力。针对这些问题,本文提出通过学习图像高低频子带特征的方法来提高网络学习卫星图像高层特征的能力,采用融合原图像与高频子带特征的方法来解决许凤晖等^[10]在中网络训练高频子带时产生的过拟合问题,再将融合高频子带与低频子带的样本融合之后训练,能有效提升场景分类的精度。

2 算法原理与流程

2.1 NSWT 图像子带分解

小波变换作为图像处理中的基础方法主要应用于图像去噪、压缩以及分解。基础小波变换会导致分解之后的子带尺寸发生变化,一次分解之后会变为原图的四分之一大小,为了充分保证原图像素信息的完整性,采用 NSWT (Non-Subsampled Wavelet Transform, 非下采样小波变换) 将图像分解为一个低频子带和多个方向的高频子带,得到的子带图像不仅能够保持与原图像的尺寸一致而且能减少子带图像的成像噪声,因为原图像是多通道彩色图像,所以在变换之前将图像映射到 YCbCr 颜色空间,分别对单通道图像进行分解,再将变换后的单通道子带图像重构为多通道彩色图像。NSWT 的图像分解结果如图 1 所示。

非下采样小波变换分解得到的低频子带体现了图像的整体轮廓且保留了原图像大部分信息可以直接用于训练;而高频子带仅保留了原图像的边缘特征与细节信息,许凤晖等^[10]直接将高频子带用于训练出现了网络过拟合与泛化能力减弱的问题,从而导致高频子带的分类正确率低于原图像。本文采用图像融合的方法来解决高频子带训练的问题,因为水平高频分量所含的边缘特征较为明显,所以选择对其进行频带特征融合。

2.2 图像频带特征融合

图像融合 (Image Fusion) 是指将多个采集源信

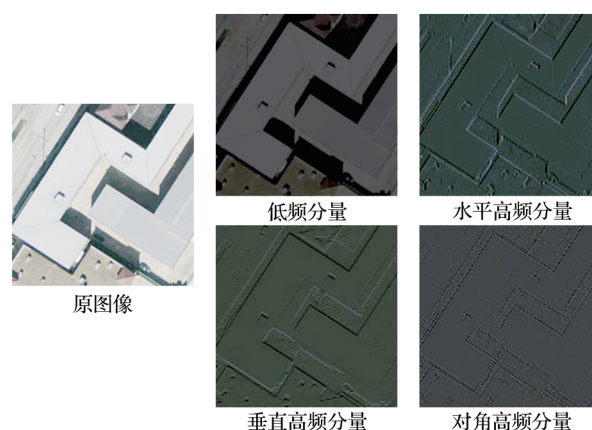


图1 NSWT分解结果图

Fig.1 NSWT decomposition result graph

道所采集到的关于同一目标的图像数据经过图像处理等技术最大限度的提取不同信道中的有利信息,最后综合成高质量的图像,以提高图像信息的利用率^[11]。因此,为解决上一节通过非下采样小波变换得到的高频子带在参与网络训练时由于特征信息太少容易产生过拟合的问题,本文采用 PCA (Principal Components Analysis, 主成分分析) 图像融合算法将高频子带与原图像进行特征融合^[12]来增强图像特征信息,融合算法的流程如图 2 所示。

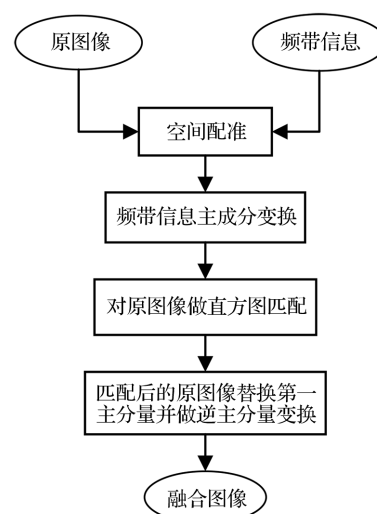


图2 PCA 图像融合流程图

Fig.2 PCA image fusion flowchart

根据流程图对参与融合的图片进行配准。以频带图像为例,将 X 定义为高频子带数据矩阵。

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

对高频子带数据矩阵进行标准化处理:

$$x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{\sqrt{\text{var}(x_j)}} \quad (2)$$

其中: x_{ij} 表示图像的像素值, \bar{x}_j 表示每列像素的均值, $\text{var}(x_j)$ 表示每列像素的方差, 相关代数式的计算如下:

$$\begin{cases} \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \\ \text{var}(x_j) = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \end{cases} \quad (3)$$

标准化后的数据矩阵用 X^* 表示, 计算 X^* 的相关系数矩阵 R , 及其特征值与对应的特征向量。

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \quad (4)$$

$$r_{ij} = \frac{1}{n-1} \sum_{i=1}^n x_{ii} x_{ij}$$

采用雅可比方法计算 R 的特征值为:

$$(\lambda_1, \lambda_2, \cdots, \lambda_p) \quad (5)$$

上式求得特征值对应的特征向量为:

$$A_i = (a_{i1}, a_{i2}, \cdots, a_{ip}); i = 1, 2, \cdots, p \quad (6)$$

各主分量 F 与贡献率 η 按如下方式计算:

$$\begin{cases} F_j = \sum_{k=1}^p x_k a_{jk}; j = 1, 2, \cdots, p \\ \eta = \lambda_i / \sum_{i=1}^p \lambda_i \end{cases} \quad (7)$$

求出的主分量有 p 个, 根据贡献率 η 选取贡献率大于 85% 的主分量, 按照降序排列可以求得第一主分量 F_1 。

将原图像和第一主分量图像进行直方图匹配, 然后将第一主分量用原图像替换。对选取的主分量一起做逆主分量变换得到融合图像。融合效果如图 3 所示。

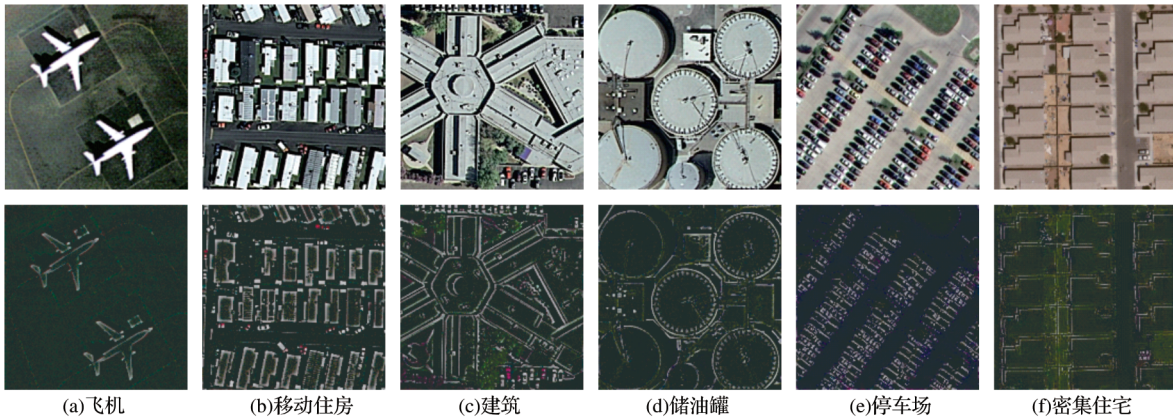


图 3 高频子带特征融合效果图

Fig.3 High frequency sub-band feature fusion effect chart

由图 3 可以看出, 进行图像特征融合之后的高频子带图像保留了部分色彩与轮廓特征, 且边缘特征与细节信息得到很好的凸显, 特征信息得到增强。在减少图像冗余信息的基础上, 保证网络在学习到高频特征的同时不会产生网络泛化与抗干扰能力减弱的问题。

2.3 卷积神经网络场景分类模型

场景分类的关键在于图像特征的提取与学习, 传统的分类模型使用人工选取的特征。由于人工的局限性, 在选取特征的同时需要大量的经验累积, 所以必须由专业人员来操作, 降低了特征选取的可操作性。以 Cabor、LBP、SIFT 等人工设计的特征为例, 这些特征在特定范围的识别中具有相当不错的分类效果, 但是根据样本集的不同, 这些特

征的应用范围受到很大的限制^[13]。区别于人工选取的特征, 卷积神经网络通过有监督学习来提取更加高层次的图像特征, 使网络模型的应用范围得到很大的扩展。基本的卷积神经网络主要由卷积层、池化层、激活层等多种类型的层结构堆砌而成。网络层数的选择与样本类别数以及样本总数量相关, 卷积神经网络的基本层类型介绍如下:

(1) 卷积层。卷积层是卷积神经网络的核心, 该层的功能是对前层输入进行特征提取^[14], 假设输入图像为二维矩阵 X , 大小为 $m \times n$, 卷积核组数量为 K 且大小为 $k \times k$, 进行卷积运算之后得到的输出 Y 大小为 $((m-k)/s+1) \times ((n-k)/s+1)$ 。相关公式如下:

$$y_i = b_i + \sum_j w_{ij} \otimes x_i \quad (8)$$

其中: x_i 表示卷积层的输入; w_{ij} 为权重; b_i 表示偏置;

s 表示卷积核的步长,卷积核组的数量 K 为经过卷积层处理之后产生的特征图的数量。

(2)激活层。因为线性模型的表达能力不够,所以使用激活函数加入非线性因素,激活之后的特征图为 F 。

$$f_s = \tau(y_i) = \tau(b_i + \sum_i w_{ij} \otimes x_i) \quad (9)$$

其中: τ 为激活函数,常用的激活函数有Sigmoid、Tanh、ReLU等,其中ReLU能够使部分神经元的输出为0,这样就增大了网络的稀疏表达能力,防止网络出现过拟合,而且很好地解决了网络层数增加造成的梯度消散问题,加快SGD的收敛速度,因此大部分的神经网络都使用ReLU(线性修正单元)作为

激活函数^[15]。

(3)池化层:当输入层的图像尺寸较大时,经过卷积层运算之后产生的特征图仍然具有很高的特征维度,为了解决数据冗余的问题,使用池化层将特征图进行下采样降低特征维度,网络的运算速度也得到了提升,所以池化层也叫下采样层。一般在前层特征图的 2×2 的像素范围内采用MAX(最大值)、AVE(均值)等策略进行下采样^[16]。

本文的卷积神经网络结构如图4所示,包括1个输入层、4个卷积层、4个池化层、4个局部响应归一化层、3个全连接层、2个Dropout层以及1个Softmax层输出分类结果。

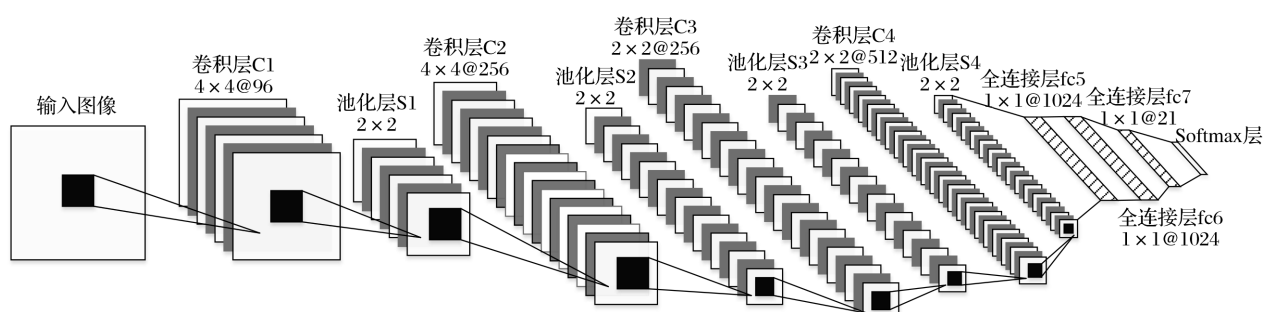


图4 本文构建的卷积神经网络结构及参数图

Fig.4 The convolutional neural network structure and parameter map constructed in this paper

指导学习的概念来源于心理学,是指有系统指导的按规定程序进行学习的一种学习方式。本文提出的GL-CNN(Guided Learning Convolutional Neural Network,指导学习卷积神经网络)是指采用频谱分析来实现高低频样本融合,以此来指导卷积神经网络有针对性的学习样本图像的高低频特征的分类模型,其基本流程为:

(1)对样本进行频谱分析,结合频谱角向能量分布曲线得出样本高低频分量的比重。

(2)分别训练融合高频子带与低频子带得出每类的分类结果,再对比频谱分析结果得出样本对高低频子带的敏感程度。

(3)根据样本类别对高低频子带的敏感程度进行样本融合再通过网络训练得出分类结果。

2.4 算法整体流程

(1)对源图像进行NSWT提取出图像的高低频子带。

(2)将高频子带与原图像进行频带特征融合得到融合高频子带。

(3)采用频谱分析的方法分析样本图像的高低频分量,然后通过频谱分析结果实现融合高频子带

与低频子带的样本融合,最后使用融合之后的样本指导卷积神经网络学习图像的高低频特征来实现场景分类,算法整体流程图如图5所示。

3 卫星图像场景分类试验

3.1 试验数据集

试验数据集UCM_LandUse为美国加州土地使用的高分辨率卫星遥感图像数据集^[17],其中包括飞机、河流、森林和住宅区等21个类别,图片大小为 256×256 ,每类图片为100张。本次试验中每类随机选取80张作为训练样本,20张作为测试样本。

3.2 试验过程与分析

为了更加直观地了解样本图像对高低频分量的敏感程度,采用频谱分析的方式对样本图像的高低频分量进行定性分析,通过对样本图像频谱角向能量分布曲线平稳区间的分析得出灌木丛、密集住宅、移动住房、停车场和中等密度住宅的高频分量相对突出,而高速公路、十字路口、天桥、跑道和稀疏住宅低频分量相对突出,其他比如飞机、港口和储油罐等类的高频与低频分量比较接近。可以认为高低频子带单独训练时的分类正确率与图像高

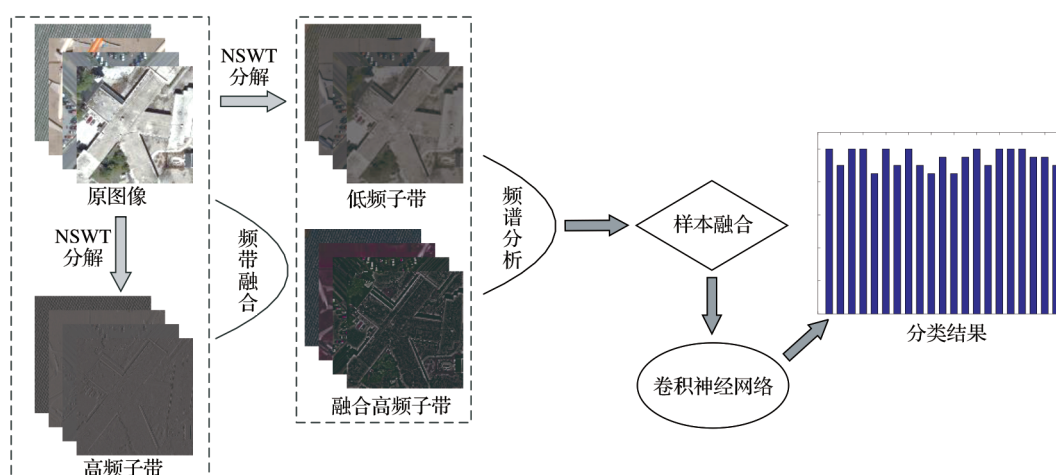


图 5 本文算法整体流程图

Fig.5 This article algorithm overall flow chart

低频分量的比重有很大的联系,频谱分析的结果能够给样本融合提供理论指导。

使用卷积神经网络模型对原图像、融合高频子带和低频子带单独进行训练得出的总体正确率分别为 87.4%、92.6% 和 88.8%,训练的每类正确率如图 6 所示。

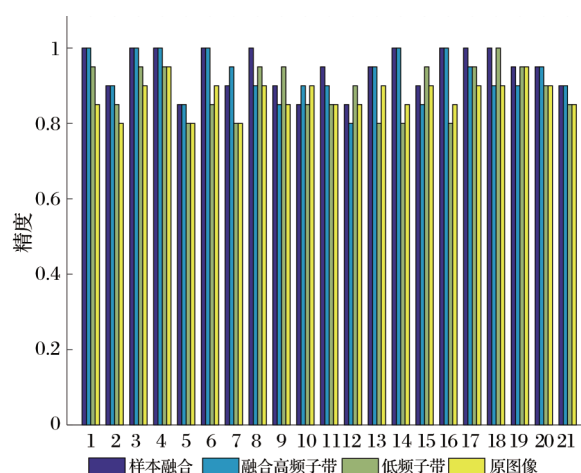


图 6 21 类样本分类正确率统计

Fig.6 21 class sample classification correct rate statistics

因为高低频信息的提取减少了原图像冗余信息,所以能得到更好的分类效果。联合频谱分析得到的各类别高低频分量的比重可以得出不同类别的样本图像对高低频子带敏感程度有很大区别,由图 6 可知高频分量较多的类别在融合高频子带上的分类正确率平均高于在低频子带 15% 左右,低频分量较多的类别在低频子带上的分类正确率平均高于在融合高频子带 10% 左右,高低频分量接近的类别在两个子带上的分类正确率相近,因此样本图像高低频分量的比重很大程度上决定了选取相应的子带进行分类。

对比高频与低频单独训练的分类结果,统计出融合高频子带分类结果的第 9、12、15 类(分别代表:高速公路、十字路口、天桥)的分类正确率未高于 85%,这 3 类样本都属于低频分量突出的样本,样本的频谱图与角向能量分布曲线示例如图 7 所示,频

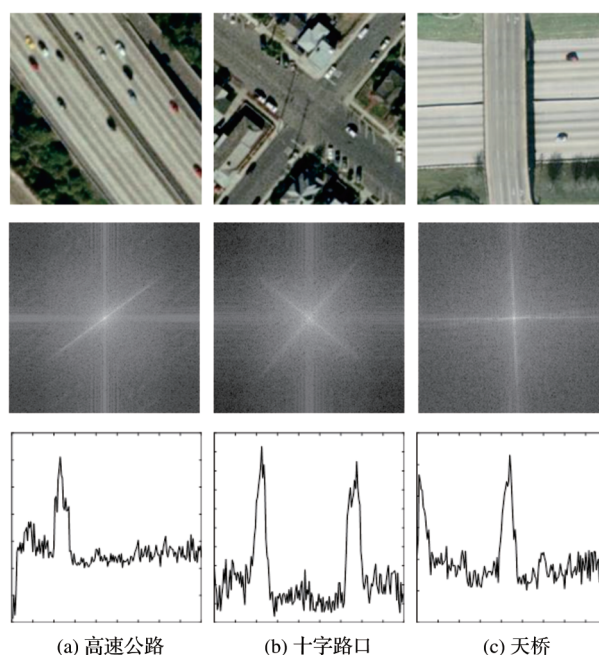


图 7 3 类样本频谱与角向能量分布曲线示例

Fig.7 Three types of sample spectrum and angular energy distribution curve examples

谱角向能量总体分布均匀、曲线的平稳区间较大。由低频子带的分类结果可知低频子带在这 3 类样本上的正确率均高于融合高频子带 10% 左右,因此将融合高频子带样本中分类正确率未高于 85% 的 3 类样本用低频子带代替来实现样本融合,样本融合效果如图 8 所示。

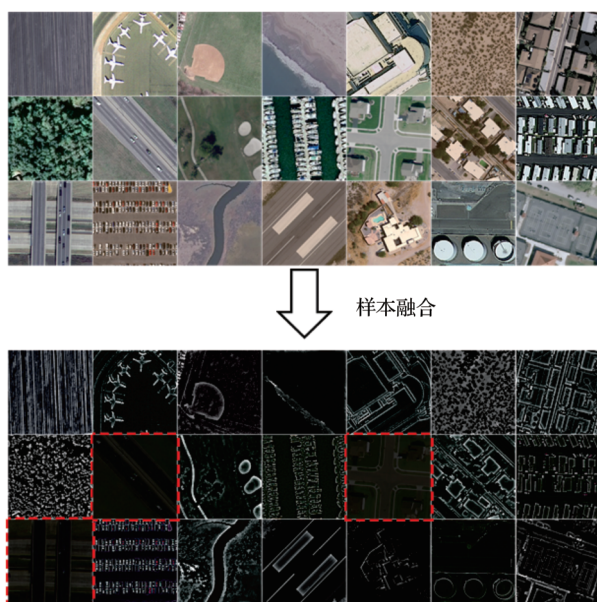


图8 样本融合效果示例

Fig.8 Sample fusion effect example

图8中的红色虚线框表示样本融合之后的低频子带示例,其余为融合高频子带示例。由于融合高频子带与低频子带的样本融合之后,卷积神经网络能够有针对性的学习图像的高低频特征,网络的分类效果得到再次提升,在调整的3类目标上分类正

确率平均提高6%,总体正确率达到94.52%,每类的分类结果如图9所示。

3.3 卫星图像场景分类算法分析对比

为了验证本文提出算法的优势,针对美国加州土地使用UCM_LandUse数据集对算法做分析和对比,并将结果列入表1~3。

表1 融合前后高频子带分类正确率比较

Table 1 Comparison of high frequency subband classification accuracy before and after fusion

	原图像	未融合	频带融合
正确率/%	87.4	78.6	92.6

表2 高低频分量显著的10类样本分类正确率分析

Table 2 Analysis of classification accuracy of 10 types of samples with high and low frequency components

类别	融合高频子带				
子类	灌木丛	密集住宅	移动住房	停车场	中密住宅
正确率/%	100	95	100	100	95
平均值/%	98				
类别	低频子带				
子类	高速公路	十字路口	天桥	跑道	稀疏住宅
正确率/%	95	90	95	100	95
平均值/%	95				

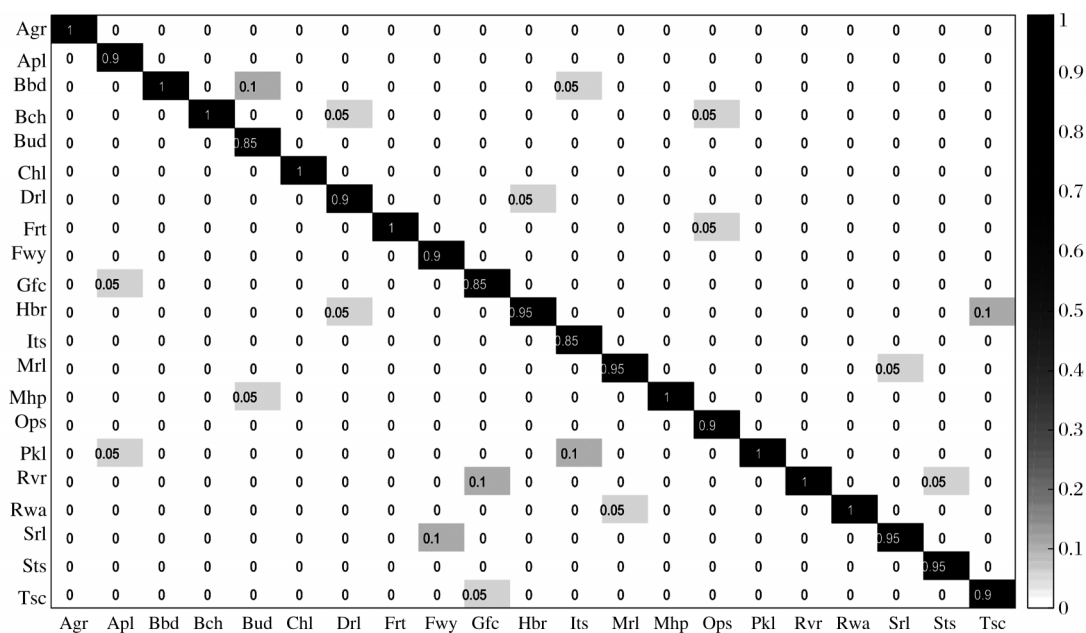


图9 样本融合后的分类混淆矩阵

Fig.9 Classification confusion matrix after sample fusion

对比未融合的高频子带与原图像的分类正确率可知由于原高频子带包含的场景信息过少导致网络过拟合不能充分学习高频子带特征,因而分类正确率低于原图像。融合之后的高频子带图像克

服了由非下采样小波变换得到的高频子带图像存在的信息量大幅度减少的问题,使网络更加容易学习到高频子带信息。许凤晖等^[10]直接使用高频子带训练的正确率为78.6%,由表1可知特征融合之

表 3 不同算法的分类正确率对比
Table 3 Comparison of classification accuracy
of different algorithms

分类算法	分类正确率/%
MNCC ^[18]	88.26
ConvNET ^[19]	89.79
SVM ^[20]	78.57
CCM-BOVW ^[21]	86.64
MS-DCNN ^[10]	91.34
PCA-CNN ^[1]	92.86
本文算法	94.52

后的高频子带图像在分类正确率上为 92.6%，相比于未融合的高频子带提高 14%，能够充分体现融合高频子带在分类上具有明显优势。

由表 2 可知，高频分量突出的样本图像在融合高频子带分类中正确率均不低于 95%，有的类甚至达到了 100%，低频分量突出的样本在低频子带分类中也取得了很高的正确率，由图 6 统计得出剩余的高低频分量接近的样本图像在融合高频子带分类中的平均正确率略高于低频子带 3% 左右，且均高于原图像，体现了低频子带与特征融合之后的高频子带在分类上的优越性。由此可见，卷积神经网络能很好的学习图像高低频子带所携带的特征信息，与其他分类方法对比结果如表 3 所示。

由表 3 可知本文提出的通过频谱角向能量分布曲线平稳区间的分析指导卷积神经网络交替学习融合高频子带和低频子带的方法能够有效的增强卷积神经网络的学习效果，通过对比也发现基于卷积神经网络的分类算法 (MNCC、ConvNet、MS-DCNN、PCA-CNN 以及本文的算法) 的分类效果要优于基于学习中、低层特征的分类算法 (SVM 和 BOVM)，进一步证明了卷积神经网络的优越性。

4 结 语

针对遥感图像的场景分类，本文提出的分类方法通过非下采样小波变换提取出图像的高低频子带，再对原高频子带进行图像特征融合得到新的高频子带，然后采用频谱分析的方法完成高低频子带的样本融合，最后通过卷积神经网络训练来实现场景分类。在标准的遥感图像公开数据集上的分类结果表明，本算法能够克服原高频子带训练产生的过拟合，能够使网络有针对性的学习图像的高低频特征，在场景分类中有明显的优势。但本算法还存在不足，在高低频分量接近的图像上还未能有比较

好的方法去处理，在卷积神经网络的参数和结构上还可以继续优化改进，以后的工作会针对这两点进行研究。

参考文献 (References):

- [1] He Xiaofei, Zou Zhengrong, Tao Chao, *et al.* Combined Saliency with Multi-Convolutional Neural Network for High Resolution Remote Sensing Scene Classification[J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(9): 1073-1080. [何小飞, 邹峥嵘, 陶超, 等. 联合显著性和多层卷积神经网络的高分影像场景分类[J]. *测绘学报*, 2016, 45(9): 1073-1080.]
- [2] Fan Min, Han Qi, Wang Fen, *et al.* Scene Image Categorization Algorithm based on Multi-level Features Representation [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2017, 46(6): 1909-1915. [范敏, 韩琪, 王芬, 等. 基于多层次特征表示的场景图像分类算法[J]. *吉林大学学报 (工学版)*, 2017, 46(6): 1909-1915.]
- [3] Liu Dawei, Han Ling, Han Xiaoyong. High Spatial Resolution Remote Sensing Image Classification based on Deep Learning [J]. *Acta Optica Sinica*, 2016, 36(4): 0428001. [刘大伟, 韩玲, 韩晓勇. 基于深度学习的高分辨率遥感影像分类研究[J]. *光学学报*, 2016, 36(4): 0428001.]
- [4] Oliva A, Torralba A. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope [J]. *International Journal of Computer Vision*, 2001, 42(3): 145-175.
- [5] Sivic J, Zisserman A. Video Google: A Text Retrieval Approach to Object Matching in Videos [C]// *Proceedings of the Ninth IEEE International Conference on Computer Vision*. Nice, 2003, 2: 1470-1477.
- [6] Xiao X Z, Tuia Devis, Li C M, *et al.* Deep Learning in Remote Sensing: A Review [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2017: 1-60.
- [7] Castelluccio M, Poggi G, Sansone C, *et al.* Verdoliva. Landuse Classification in Remote Sensing Images by Convolutional Neural Networks [J]. *arXiv:1508.00092*, 2015.
- [8] Nogueira K, Penatti O, Santos J. Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification [J]. *Pattern Recognition*, 2016, 61: 539-556.
- [9] Luus F, Salmon B, Bergh F, *et al.* Multiview Deep Learning for Land-use Classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(12): 2448-2452.
- [10] Xu Suhui, Mu Xiaodong, Zhao Peng, *et al.* Scene Classification of Remote Sensing Image based on Multi-scale Feature and Deep Neural Network [J]. *Acta Geodaetica et Cartographica Sinica*, 2016, 45(7): 834-840. [许夙晖, 慕晓东, 赵鹏, 等. 利用多尺度特征与深度网络对遥感影像进行场景分类[J]. *测绘学报*, 2016, 45(7): 834-840.]
- [11] Yilihamu Yaermaimaiti, Xie Lirong, Kong Jun. Remote Sensing Image Fusion based on PCA Transform and Wavelet Transform [J]. *Infrared and Laser Engineering*, 2014, 43(7): 2335-2340. [伊力哈木·亚尔买买提, 谢丽蓉, 孔军. 基于 PCA

- 变换与小波变换的遥感图像融合方法[J]. 红外与激光工程, 2014, 43(7): 2335-2340.]
- [12] Ji Feng, Li Zeren, Chang Xia, *et al.* Remote Sensing Image Fusion Method based on PCA and NSCT Transform[J]. Journal of Graphics, 2017, 38(2): 247-252. [纪峰, 李泽仁, 常霞等. 基于PCA和NSCT变换的遥感图像融合方法[J]. 图学学报, 2017, 38(2): 247-252.]
- [13] Bengio Y, Courville A, Vincent P. Representation Learning: A Review and New Perspectives[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(8): 1798-1828.
- [14] Zhao Juanping, Guo Weiwei, Liu Bin, *et al.* Convolutional Neural Network-based SAR Image Classification with Noisy Labels[J]. Journal of Radars, 2017, 6(5): 514-523. [赵娟萍, 郭炜炜, 柳彬, 等. 基于概率转移卷积神经网络的含噪标记SAR图像分类[J]. 雷达学报, 2017, 6(5): 514-523.]
- [15] Wang Siyu, Gao Xin, Sun Hao, *et al.* An Aircraft Detection Method based on Convolutional Neural Networks in High-Resolution SAR Images[J]. Journal of Radars, 2017, 6(2): 195-203. [王思雨, 高鑫, 孙皓, 等. 基于卷积神经网络的高分辨率SAR图像飞机目标检测方法[J]. 雷达学报, 2017, 6(2): 195-203.]
- [16] Xu Feng, Wang Haipeng, Jin Yaqiu. Deep Learning as Applied in SAR Target Recognition and Terrain Classification[J]. Journal of Radars, 2017, 6(2): 136-148. [徐丰, 王海鹏, 金亚秋. 深度学习在SAR目标识别与地物分类中的应用[J]. 雷达学报, 2017, 6(2): 136-148.]
- [17] Yang Y, Newsam S. Bag-of-visual-words and Spatial Extensions for Land-use Classification[C] // ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM GIS), 2010.
- [18] Liu Y, Fu Z Y, Zheng F B. Scene Classification of High Resolution Remote Sensing Image based on Multimedia Neural Cognitive Computing[J]. Systems Engineering and Electronics, 2015, 37(11): 2623-2633.
- [19] Nogueira K, Miranda W O, Santosj A D. Improving Spatial Feature Representation from Aerial Scenes by Using Convolutional Network[C] // Proceedings of the 2015 28th SIBGRAPI Conference on Graphics, Patterns and Images. Salvador: IEEE, 2015: 289-296.
- [20] Zheng X W, Sun X, Fu K, *et al.* Automatic Annotation of Satellite Images via Multifeature Joint Sparse Coding with Spatial Relation Constraint [J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(4): 652-656.
- [21] Zhao L J, Tang P, Huo L Z. Land-use Scene Classification Using a Concentric Circle-structured Multiscale Bag-of-visual-words Mode[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(12): 4613-4620.

Remote Sensing Image Scene Classification based on Frequency Band Feature Fusion and GL-CNN

Cui Xianliang, Chen Lifu, Xing Xuemin, Yuan Zhihui

(School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, 410114, China)

Abstract: The classification of high-resolution satellite remote sensing image scene information is of great significance for image analysis and interpretation. The traditional high-resolution satellite remote sensing image scene classification method mainly relies on the artificially extracted middle and low-level features and can not make good use of image-rich scenes. In response to this problem, a classification method based on band feature fusion and GL-CNN (Guided Learning Convolutional Neural Network) is proposed. Firstly, the high-low frequency sub-band of the image is extracted by NSWT (Non-Subsampled Wavelet Transform), and then the high-frequency sub-band is fused to obtain the fused high-frequency sub-band, and then the angular energy distribution curve is combined. The stationary interval analysis realizes the fusion of the fusion high-frequency sub-band and the low-frequency sub-band, and finally guides the convolutional neural network to automatically extract the high-level features contained in the high-low frequency sub-band of the image to realize the scene classification. Experiments on UCM_LandUse 21 data show that the classification accuracy of this method reaches 94.52%, which is significantly improved compared with previous algorithms.

Key words: Non-subsampled wavelet transform; Band feature fusion; Guided learning; Sample fusion; Scene classification