

# 控制与决策

Control and Decision

基于脉冲卷积神经网络稀疏表征的高分辨率遥感图像场景分类方法

张哲益, 曹卫华, 朱蕊, 胡文凯, 吴敏

引用本文:

张哲益, 曹卫华, 朱蕊, 胡文凯, 吴敏. 基于脉冲卷积神经网络稀疏表征的高分辨率遥感图像场景分类方法[J]. *控制与决策*, 2022, 37(9): 2305–2313.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0279>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [自适应感受野网络的行人重识别](#)

Adaptive receptive network for person re-identification

控制与决策. 2022, 37(1): 119–126 <https://doi.org/10.13195/j.kzyjc.2020.0505>

#### [复杂背景下全景视频运动小目标检测算法](#)

Panoramic video motion small target detection algorithm in complex background

控制与决策. 2021, 36(1): 249–256 <https://doi.org/10.13195/j.kzyjc.2019.0686>

#### [基于稀疏化神经网络的浮选泡沫图像特征选择](#)

Selection method for froth image characters based on sparse neural network

控制与决策. 2021, 36(7): 1627–1636 <https://doi.org/10.13195/j.kzyjc.2019.1788>

#### [一种基于多层语义特征的图像理解方法](#)

An image understanding method based on multi-level semantic features

控制与决策. 2021, 36(12): 2881–2890 <https://doi.org/10.13195/j.kzyjc.2020.0927>

#### [基于改进卷积神经网络的动力下肢假肢运动意图识别](#)

Intent recognition of power lower-limb prosthesis based on improved convolutional neural network

控制与决策. 2021, 36(12): 3031–3038 <https://doi.org/10.13195/j.kzyjc.2020.0326>

# 基于脉冲卷积神经网络稀疏表征的高分辨率 遥感图像场景分类方法

张哲益, 曹卫华<sup>†</sup>, 朱蕊, 胡文凯, 吴敏

(1. 中国地质大学(武汉)自动化学院, 武汉 430074; 2. 复杂系统先进控制与智能自动化湖北省重点实验室, 武汉 430074; 3. 地球探测智能化技术教育部工程研究中心, 武汉 430074)

**摘要:** 遥感图像场景分类对土地资源管理具有重要意义,然而高分辨率遥感图像中地物分布复杂,图像中存在着与当前场景无关的冗余信息,会对场景的精确分类造成影响. 对此,提出一种基于脉冲卷积神经网络(SCNN)稀疏表征的场景分类方法. 从稀疏表征出发,利用脉冲神经元的稀疏脉冲输出特性,设计脉冲卷积神经网络,去除遥感图像中与场景无关的冗余信息,实现对图像的稀疏表征;提出基于脉冲输出交叉熵损失函数的反向传播算法,在该算法的基础上利用梯度下降训练脉冲卷积神经网络,优化网络参数,实现遥感图像场景分类;通过实验验证方法的有效性,将所提出方法应用于 Google 和 UCM 两个遥感图像数据集,并与传统的卷积神经网络(CNN)进行对比. 实验结果表明,所提出方法可以对遥感图像进行稀疏表征,实现场景分类;相对于卷积神经网络,所提出方法在遥感图像场景分类任务上更具有优势.

**关键词:** 高分辨率遥感图像; 场景分类; 稀疏表征; 脉冲卷积神经网络

中图分类号: TP18

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0279

**引用格式:** 张哲益,曹卫华,朱蕊,等. 基于脉冲卷积神经网络稀疏表征的高分辨率遥感图像场景分类方法[J]. 控制与决策, 2022, 37(9): 2305-2313.

## Sparse representation with spike convolutional neural networks for scene classification of remote sensing images of high resolution

ZHANG Zhe-yi, CAO Wei-hua<sup>†</sup>, ZHU Rui, HU Wen-kai, WU Min

(1. School of Automation, China University of Geosciences, Wuhan 430074, China; 2. Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China; 3. Engineering Research Center of Intelligent Technology for Geo-Exploration, Ministry of Education, Wuhan 430074, China)

**Abstract:** Remote sensing imagery scene classification is of great significance to land resource management. However, the distribution of ground objects in remote sensing images of high spatial resolution (HSR) is complex, and there is redundant information irrelevant to the current scene in the images, which will affect the accurate classification of the scene. To solve this problem, a scene classification method based on the spike convolutional neural network (SCNN) is proposed. From the perspective of sparse representation, the SCNN is designed based on the sparse spike output characteristics of spike neurons to remove the redundant information irrelevant to the scene in remote sensing images and realize sparse representation of images. A backpropagation algorithm based on the spike output cross entropy loss function is proposed. Based on this algorithm, the SCNN is trained by gradient descent, and the network parameters are optimized to realize scene classification of remote sensing images. The validity of the proposed method is verified by experiments, where the proposed method is applied to two remote sensing imagery datasets, namely, Google and UCM, and compared with the traditional convolutional neural network (CNN). Experimental results show that the proposed method is able to perform sparse representation of remote sensing images and realize scene classification; and compared with CNN, the proposed method shows better performance in the remote sensing imagery scene classification task.

**Keywords:** HSR remote sensing images; scene classification; sparse representation; spike convolutional neural network

收稿日期: 2021-02-13; 录用日期: 2021-07-05.

基金项目: 湖北省自然科学基金创新群体项目(2020CFA031).

责任编辑: 曹进德.

<sup>†</sup>通讯作者. E-mail: weihuacao@cug.edu.cn.

## 0 引言

遥感技术的快速发展,使包含丰富信息的遥感图像数据日益增多,遥感图像数据解译成为了一个热点问题.场景分类是遥感图像数据解译的主要任务之一,它是指根据遥感图像中包含的信息,判别图像属于哪一类场景(如森林、农业用地、居住区、公路等).在地区区域划分、区域规划、土地利用统计等方面,遥感图像场景分类具有很大的应用价值.

针对高分辨率遥感图像的地物识别问题,一些研究人员通过提取图像的光谱特征、空间特征、纹理特征等低层特征<sup>[1-4]</sup>,实现了对图像中树木、建筑物、道路等地物的识别.但是这些方法只能识别遥感图像中的地物,无法对整张遥感图像所表示的场景进行判别,比如识别出了遥感图像中的建筑物,却无法判别整张图像中的建筑物表示的场景是住宅区还是商业区<sup>[5-6]</sup>.

为了解决上述问题,研究者们提出许多方法实现遥感图像场景分类,可以分为视觉词袋<sup>[7-9]</sup>、局部检测器<sup>[10-11]</sup>和神经网络<sup>[12-14]</sup>这几类方法,其中神经网络这类方法的本质是对遥感图像进行表征学习,再将学习到的特征送入分类器,进而训练分类器实现场景分类.文献[12]提出一种深度无监督表征学习框架学习遥感图像的低层特征,并利用线性支持向量机实现场景分类;而遥感图像的低层特征没有语义上的意义,这会造成场景分类准确度较低.对此,文献[13]通过提出一种基于自动编码器的表征学习方法学习图像的中级视觉特征,该特征相比于低层特征包含更多的语义信息,从而提高了分类准确度.为了在场景分类时更好地利用遥感图像的局部信息,文献[14]提出一种基于卷积神经网络的无监督表征学习方法获取图像的局部信息.虽然上述方法实现了遥感图像场景分类,但由于没有考虑遥感图像数据具有地物分布复杂、存在与当前场景无关的冗余信息的特点,导致上述方法从冗余信息中学习到的特征会引起后续分类器对当前场景的误判,造成最终分类准确度不高.通过对遥感图像进行稀疏表征<sup>[15]</sup>,可以去除图像中的冗余信息,促进其有效分类;与组成卷积神经网络的传统神经元相比,脉冲神经元<sup>[16-17]</sup>具有稀疏脉冲输出特性,可以组成脉冲卷积神经网络以实现遥感图像的稀疏表征.

通过以上分析,本文提出一种基于脉冲卷积神经网络稀疏表征的高分辨率遥感图像场景分类方法.首先,利用具有稀疏脉冲输出特性的脉冲神经元设计脉冲卷积神经网络,实现对高分辨率遥感图像的

稀疏表征,减少图像的冗余信息;然后,针对遥感图像场景分类问题,根据脉冲神经元的输出特性提出脉冲输出交叉熵损失函数,并推导基于该损失函数的反向传播算法;最后,在Google和UCM两个遥感图像数据集上验证本文方法的有效性以及其相比于卷积神经网络的优势.

## 1 高分辨率遥感图像在场景分类任务中的数据特点分析

在场景分类任务中,高分辨率遥感图像数据具有地物分布复杂、存在与当前场景无关的冗余信息的特点.下面用一个具体的例子说明这个特点,如图1(a)所示,该遥感图像来自于Google数据集,所属的场景标签为“池塘”.从图1(b)中可以看出,该遥感图像的地物分布除了池塘之外,还有农田、建筑物、道路、树林,并且农田、建筑物占据图像的比例很大,甚至超过池塘所占的比例,这些信息的存在将影响对场景标签“池塘”的正确判别.由此可以看出遥感图像数据在场景分类任务中具有地物分布复杂及存在与当前场景无关的冗余信息的特点,冗余信息经过表征学习会产生冗余的特征,将引起分类器对场景的误判,最终影响分类准确度.

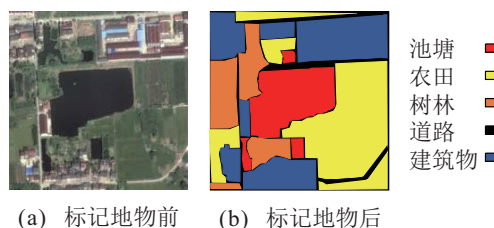


图1 场景标签为“池塘”的遥感图像

针对该问题,本文利用由脉冲神经元组成的脉冲卷积神经网络实现稀疏表征.脉冲神经元的输出为离散的脉冲序列,具有稀疏性.遥感图像输入脉冲卷积神经网络后,图像中的信息可以触发网络中的脉冲神经元发放脉冲,在此基础上利用场景标签进行有监督学习调节网络权重,使图像中只含有场景标签的信息才能触发脉冲神经元发放脉冲,以此去除图像中与场景无关的冗余信息,达到稀疏表征的目的,并且得到的脉冲(稀疏特征)会与场景标签建立映射,实现场景分类.

## 2 模型与方法

本节将介绍利用脉冲卷积神经网络对遥感图像进行稀疏表征和场景分类的方法,给出用于组成脉冲卷积神经网络的脉冲神经元模型,并提出了用于训练网络的基于脉冲输出交叉熵损失函数的反向传播算法.

### 2.1 用于脉冲卷积神经网络稀疏表征的脉冲神经元模型

在脉冲神经网络中,常用的脉冲神经元模型为漏积分点火(leaky integrate-and-fire, LIF)模型

$$\tau \frac{du(t)}{dt} = -u(t) + I(t). \quad (1)$$

其中:  $u(t)$  为神经元在  $t$  时刻的膜电压;  $\tau$  为膜电压时间常数;  $I(t)$  为该神经元的外部输入,大小由突触前神经元的脉冲与神经元之间的权重决定. 当膜电压  $u$  达到一个给定的阈值电压  $u_t$ , 神经元就会发放一个脉冲(神经元输出为1,反之为0),并且膜电压  $u$  会恢复至静息电压  $u_r$ ,这一发放脉冲的过程相当于传统神经元模型(图2(a))中的激活函数. 由图2(b)可以看出,与传统神经元相比,脉冲神经元的输出为脉冲序列,具有稀疏性,可以利用其组成脉冲卷积神经网络对遥感图像进行稀疏表征.

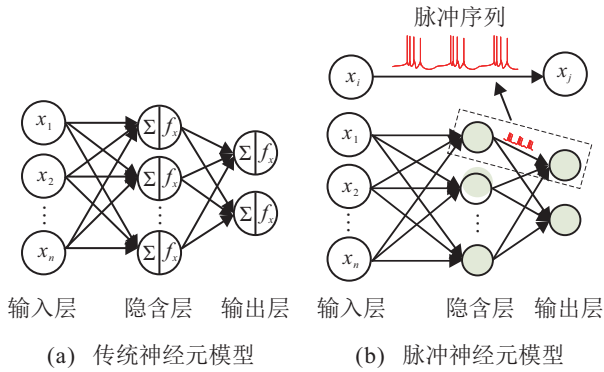


图2 两种人工神经元模型比较

由式(1)可知,LIF模型本质上为微分方程,相比于传统神经元模型,LIF模型中的膜电压  $u$  解的表达式十分复杂,其在空间上(网络的层与层之间)和时间上(膜电压  $u$  在不同时刻有不同的值)都有前向传播,并且模型的输出为不连续且无法求导的脉冲,这些问题导致不能直接利用反向传播算法和梯度下降训练由LIF模型组成的脉冲卷积神经网络,进而无法对遥感图像进行稀疏表征.

为了解决上述问题,可以将LIF模型转换为另一种形式.通过给式(1)一个初始条件  $u(t)|_{t=t_{i-1}} = u(t_{i-1})$ ,可以求出  $u(t)$  的解

$$u(t) = u(t_{i-1})e^{-\frac{t-t_{i-1}}{\tau}} + \hat{I}(t). \quad (2)$$

由式(2)可以看出,  $u(t)$  由上一时刻的  $u(t_{i-1})$  与衰减系数的乘积和神经元外部输入  $\hat{I}(t)$  共同决定,这可以被看作  $u(t)$  迭代更新的规则.

基于式(2),并设置静息电压  $u_r$  为0,可以将LIF模型(式(1))转换为与传统神经元模型类似的迭代LIF模型<sup>[18]</sup>

$$\begin{aligned} x_i^{t+1,n} &= \sum_{j=1}^{l(n-1)} w_{ij}^n o_j^{t+1,n-1}, \\ u_i^{t+1,n} &= u_i^{t,n} f(o_i^{t,n}) + x_i^{t+1,n} + b_i^n, \\ o_i^{t+1,n} &= g(u_i^{t+1,n}). \end{aligned} \quad (3)$$

其中

$$f(o_i^{t,n}) = \tau e^{-\frac{t}{\tau}} (1 - o_i^{t,n}); \quad (4)$$

$$g(u) = \begin{cases} 1, & u > u_t, u = u_r; \\ 0, & u < u_t. \end{cases} \quad (5)$$

$w_{ij}^n$  为第  $n$  层的神经元  $i$  与神经元  $j$  间的连接权重,  $o_j^{t+1,n-1}$  为第  $n-1$  层的神经元  $j$  在  $t+1$  时刻发放的脉冲,  $x_i^{t+1,n}$  为第  $n$  层的神经元  $i$  在  $t+1$  时刻的外部输入,  $b_i^n$  为偏置项,  $u_i^{t,n}$  为第  $n$  层的神经元  $i$  在  $t$  时刻的膜电压,  $T$  为网络模拟时间,  $l(n-1)$  为第  $n-1$  层的神经元个数.

迭代LIF模型把原始LIF模型在空间和时间上的前向传播进行了简化,并且其形式与传统神经元模型类似,可以仿照反向传播算法和梯度下降对由迭代LIF模型组成的脉冲卷积神经网络进行训练,实现对遥感图像的稀疏表征及场景分类.

### 2.2 基于脉冲输出交叉熵损失函数的反向传播算法

文献[18]通过定义如下均方误差损失函数推导脉冲神经网络的参数更新公式:

$$L = \frac{1}{2S} \sum_{s=1}^S \left\| y_s - \frac{1}{T} \sum_{t=1}^T o_s^{t,N} \right\|_2^2. \quad (6)$$

其中:  $y_s$  为第  $s$  个训练样本的标签向量,  $o_s^{t,N}$  为最后一层(第  $N$  层)神经元的输出向量. 但因为遥感图像场景分类是一个多分类问题,所以损失函数使用交叉熵损失函数会更加有效<sup>[19]</sup>. 下面将推导基于交叉熵损失函数的脉冲神经网络参数的更新公式.

在脉冲神经网络中,由于神经元的输出为脉冲序列,需要将交叉熵损失函数转换为脉冲输出交叉熵损失函数,即

$$L = -\frac{1}{T} \sum_{t=1}^T o_{y+1}^{T,N} + \log \left( \sum_{y=0}^Y e^{\frac{1}{T} \sum_{t=1}^T o_{y+1}^{T,N}} \right). \quad (7)$$

其中:  $[0, \dots, Y]$  为多分类任务中的标签,  $o_{y+1}^{T,N}$  为网络最后一层的第  $y+1$  个神经元在  $T$  时刻发放的脉冲.

下面利用迭代LIF模型(式(3)~(5))和脉冲输出交叉熵损失函数(式(7))推导脉冲神经网络在反向传播中参数  $w$  和  $b$  的更新公式.

通过简单地推导可以得到参数  $w$  和  $b$  的初步梯度更新公式

$$\frac{\partial L}{\partial W^n} = \sum_{t=1}^T \frac{\partial L}{\partial u^{t,n}} \frac{\partial u^{t,n}}{\partial W^n} = \sum_{t=1}^T \frac{\partial L}{\partial u^{t,n}} o^{t,n-1}, \quad (8)$$

$$\frac{\partial L}{\partial b^n} = \sum_{t=1}^T \frac{\partial L}{\partial u^{t,n}} \frac{\partial u^{t,n}}{\partial b^n} = \sum_{t=1}^T \frac{\partial L}{\partial u^{t,n}}. \quad (9)$$

由式(8)和(9)可以看出,要得到参数 $w$ 和 $b$ 的最终梯度更新公式,需要计算出 $\partial L/\partial u^{t,n}$ .

对于 $\partial L/\partial u^{t,n}$ 的求解可以借鉴反向传播算法的链式求导法则<sup>[20]</sup>,并且由于迭代LIF模型在空间和时间上都有前向传播,求解 $\partial L/\partial u^{t,n}$ 需要从时间和空间两方面综合考虑.

假设损失函数 $L$ (式(7))关于 $o_i^{t,n}$ 的梯度为

$$\frac{\partial L}{\partial o_i^{t,n}} = \delta_i^{t,n}. \quad (10)$$

首先,当时间 $t = T$ ,并且在网络最后一层 $n = N$ 时,可以直接由式(7)推导得出

$$\frac{\partial L}{\partial o_i^{T,N}} = -\frac{1}{T} + \frac{\frac{1}{T} e^{\frac{1}{T} \sum_{t=1}^T o_i^{T,N}}}{\log\left(\sum_{y=0}^Y e^{\frac{1}{T} \sum_{t=1}^T o_{y+1}^{T,N}}\right)}. \quad (11)$$

此时关于 $u_i^{T,N}$ 的梯度可以由 $o_i^{T,N}$ 推导出,有

$$\frac{\partial L}{\partial u_i^{T,N}} = \frac{\partial L}{\partial o_i^{T,N}} \frac{\partial o_i^{T,N}}{\partial u_i^{T,N}} = \delta_i^{T,N} \frac{\partial g(u_i^{T,N})}{\partial u_i^{T,N}}. \quad (12)$$

其次,当时间 $t = T$ ,并且在网络层数 $n < N$ 时,可以借鉴反向传播在空间上的链式求导规则求出

$$\begin{aligned} \frac{\partial L}{\partial o_i^{T,n}} &= \sum_{j=1}^{l(n+1)} \delta_i^{T,n+1} \frac{\partial o_i^{T,n+1}}{\partial o_i^{T,n}} = \\ &\sum_{j=1}^{l(n+1)} \delta_i^{T,n+1} \frac{\partial o_i^{T,n+1}}{\partial u_i^{T,n+1}} \frac{\partial u_i^{T,n+1}}{\partial o_i^{T,n}} = \\ &\sum_{j=1}^{l(n+1)} \delta_i^{T,n+1} \frac{\partial g(u_i^{T,n+1})}{\partial u_i^{T,n+1}} w_{ij}^{n+1}. \end{aligned} \quad (13)$$

此时,可以参照上一步求出

$$\frac{\partial L}{\partial u_i^{T,n}} = \frac{\partial L}{\partial o_i^{T,n}} \frac{\partial o_i^{T,n}}{\partial u_i^{T,n}} = \delta_i^{T,n} \frac{\partial g(u_i^{T,n})}{\partial u_i^{T,n}}. \quad (14)$$

然后,当时间 $t < T$ ,而网络层数为最后一层 $n = N$ 时,可以将网络的最后一层在时间上展开,然后在时间 $t$ 上进行反向传播,求出

$$\begin{aligned} \frac{\partial L}{\partial o_i^{t,N}} &= \frac{\partial L}{\partial o_i^{t+1,N}} \frac{\partial o_i^{t+1,N}}{\partial o_i^{t,N}} = \\ &\delta_i^{t+1,N} \frac{\partial o_i^{t+1,N}}{\partial u_i^{t+1,N}} \frac{\partial u_i^{t+1,N}}{\partial o_i^{t,N}} = \\ &\delta_i^{t+1,N} \frac{\partial g(u_i^{t+1,N})}{\partial u_i^{t+1,N}} u_i^{t,N} \frac{\partial f(o_i^{t,N})}{\partial o_i^{t,N}}. \end{aligned} \quad (15)$$

$$\frac{\partial L}{\partial u_i^{t,N}} = \frac{\partial L}{\partial o_i^{t+1,N}} \frac{\partial o_i^{t+1,N}}{\partial u_i^{t+1,N}} \frac{\partial u_i^{t+1,N}}{\partial u_i^{t,N}} =$$

$$\delta_i^{t+1,N} \frac{\partial g(u_i^{t+1,N})}{\partial u_i^{t+1,N}} f(o_i^{t,N}). \quad (16)$$

最后,当时间 $t < T$ ,且网络层数 $n < N$ 时,可以结合 $t = T, n < N$ 和 $t < T, n = N$ 这两种情况得到

$$\begin{aligned} \frac{\partial L}{\partial o_i^{t,n}} &= \\ &\sum_{j=1}^{l(n+1)} \delta_i^{t,n+1} \frac{\partial o_i^{t,n+1}}{\partial o_i^{t,n}} + \frac{\partial L}{\partial o_i^{t+1,n}} \frac{\partial o_i^{t+1,n}}{\partial o_i^{t,n}} = \\ &\sum_{j=1}^{l(n+1)} \delta_i^{t,n+1} \frac{\partial g(u_i^{t,n+1})}{\partial u_i^{t,n+1}} w_{ij}^{n+1} + \\ &\delta_i^{t+1,n} \frac{\partial g(u_i^{t+1,n})}{\partial u_i^{t+1,n}} u_i^{t,n} \frac{\partial f(o_i^{t,n})}{\partial o_i^{t,n}}. \end{aligned} \quad (17)$$

$$\begin{aligned} \frac{\partial L}{\partial u_i^{t,n}} &= \\ &\frac{\partial L}{\partial o_i^{t,n}} \frac{\partial o_i^{t,n}}{\partial u_i^{t,n}} + \frac{\partial L}{\partial o_i^{t+1,n}} \frac{\partial o_i^{t+1,n}}{\partial u_i^{t+1,n}} \frac{\partial u_i^{t+1,n}}{\partial u_i^{t,n}} = \\ &\delta_i^{t,n} \frac{\partial g(u_i^{t,n})}{\partial u_i^{t,n}} + \delta_i^{t+1,n} \frac{\partial g(u_i^{t+1,n})}{\partial u_i^{t+1,n}} f(o_i^{t,n}). \end{aligned} \quad (18)$$

通过式(18)的推导可以得到 $\partial L/\partial u_i^{t,n}$ 的梯度更新公式,但公式中存在一项 $\partial g(u)/\partial u$ ,由式(5)可知,函数 $g(u)$ 是阶跃函数,所以 $\partial g(u)/\partial u$ 是冲激函数,这样 $\partial L/\partial u_i^{t,n}$ 的梯度更新公式中就存在一项冲激函数.在利用梯度下降优化网络参数时, $\partial g(u)/\partial u$ 的存在会导致梯度爆炸或者消失.可以利用矩形函数<sup>[18]</sup>近似 $\partial g(u)/\partial u$ 解决这个问题,即

$$h(u) = \frac{1}{a} \text{sign}\left(|u - u_{th}| < \frac{a}{2}\right), \quad (19)$$

$$\lim_{a \rightarrow 0^+} h(u) = \frac{\partial g(u)}{\partial u}. \quad (20)$$

至此,在脉冲输出交叉熵损失函数下(式(7))网络参数 $w$ 和 $b$ 的反向传播梯度更新公式推导完成,只需将式(11)~(18)代入式(8)和(9)即可,在此基础上可利用梯度下降优化网络的参数.

### 3 实验结果与分析

本节将介绍用于场景分类实验的遥感图像数据集及实验相关设置,并对实验结果进行展示与分析.

#### 3.1 数据集介绍

本文利用两个遥感图像数据集进行实验,验证方法的有效性.下面对数据集进行介绍.

第1个数据集为武汉大学RS-IDEA团队设计的Google数据集<sup>[21]</sup>.该数据集的数据采集自谷歌地球卫星地图,一共包含了农业用地、商业用地、港口、闲置土地、工业用地、草地、立交桥、公园、池塘、住宅区、河流、水域这12种场景,每种场景有200张图像,每张图像的大小为 $200 \times 200$ 像素,空间分辨率为2 m.

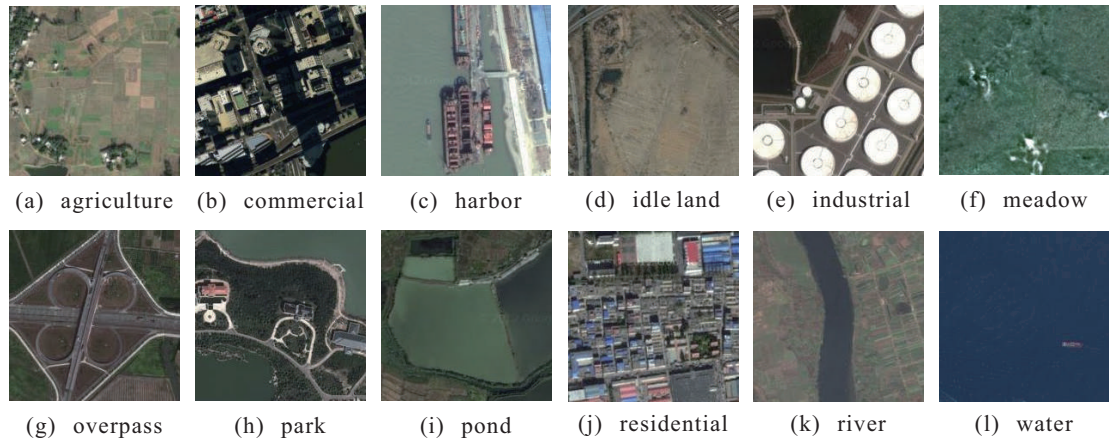


图3 Google数据集



图4 UCM数据集

Google数据集如图3所示。

第2个数据集为UCM数据集<sup>[22]</sup>。该数据集的数据是从美国地质勘探局收集的大型国家城市区域图像中手动提取出来的,一共包含了农业用地、停机坪、棒球场、海滩、建筑物、丛林、高密度住宅区、森林、高速公路、高尔夫球场、港口、十字路口、中密度住宅区、移动房车公园、立交桥、停车场、河流、飞机跑道、低密度住宅区、储油罐、网球场这21种场景,每种场景有100张图像,每张图像的大小为256×256像素,空间分辨率为1英尺。UCM数据集如图4所示。

### 3.2 实验设置

#### 3.2.1 图像增广

两个数据集中每个场景的遥感图像数量比较少,所以在图像预处理时加入了图像增广。首先将一张图像以50%的概率水平翻转,然后在图像周围填充10个像素,最后以实际的图像大小截取填充后的图像(如图5(a))。图像增广不仅扩大了数据集的规模,还

相当于加入了噪声,防止网络过拟合。

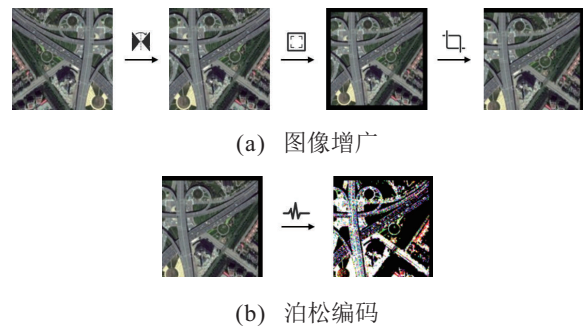


图5 图像预处理

#### 3.2.2 图像编码

由于本文的神经网络是由迭代LIF模型组成的脉冲卷积神经网络,在图像输入网络前,需要将图像编码为脉冲的形式,本文采用的脉冲编码方式为泊松编码<sup>[23]</sup>。首先随机生成一个与图像数据矩阵大小相同的矩阵,该随机矩阵的每个元素值在[0,1]之间,然后将图像数据矩阵的每个元素和与之对应位置的

随机矩阵元素进行比较,若图像数据矩阵的元素数值大,则该元素位置编码为1,代表发放一个脉冲;反之,编码为0,代表不发放脉冲.图像编码后的效果如图5(b)所示.

### 3.2.3 相关参数

本文中迭代LIF模型的参数值如表1所示,借鉴文献[18]利用迭代LIF模型进行MNIST手写数字集分类实验的参数.

表1 迭代LIF模型相关参数

参数名称	参数含义	参数值
$u_t$	阈值电压	0.5
$u_r$	静息电压	0
$\tau/ms$	神经元膜电压时间常数	0.2
$T/ms$	网络模拟时间	50
$a$	矩形函数参数	0.5

### 3.2.4 脉冲卷积神经网络结构

针对上述两个遥感图像大小不同的数据集,本文构造了两个脉冲卷积神经网络,整体上都是5个卷积层,3个全连接层(最后一层为输出层).两种脉冲卷积神经网络的结构参数如表2所示,其中池化层为最大池化层.针对两个遥感图像数据集的脉冲卷积神经

表2 针对两个数据集的脉冲卷积神经网络结构参数

数据集	输入数据大小 (channel×height×width)	网络结构参数
Google	3×200×200	6C5-P2-16C5-P2-32C21-P2-64C2-P2-128C2-P2-128-120-84-12
UCM	3×256×256	6C5-P2-16C5-P2-32C2-P2-64C2-P2-128C2-P2-128-120-84-21

网络结构如图6所示.

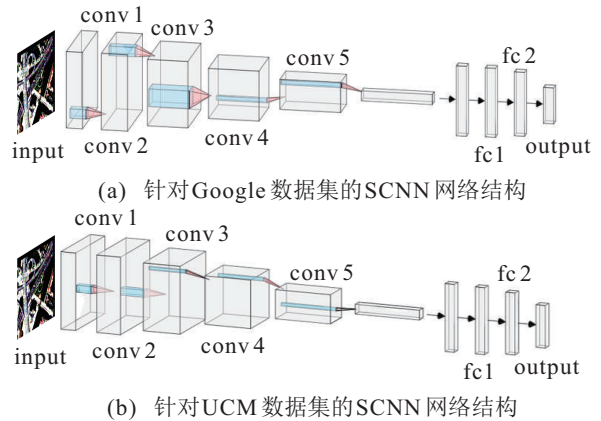


图6 针对两个数据集的脉冲卷积神经网络结构

### 3.3 稀疏表征结果

下面以Google数据集为例,说明SCNN对遥感图像稀疏表征的有效性及鲁棒性.在网络经过训练后,从Google数据集中随机选取一张遥感图像(场景标签为“立交桥”,标签序号为6),经过图像增广(相当于加入噪声)和泊松编码后将其输入网络(如图6(a)所示).在一个网络模拟时间后( $T = 50\text{ ms}$ ),将网络5个卷积层和3个全连接层(最后一层为输出层)中脉冲神经元发出的脉冲序列可视化,因为脉冲神经元输出的脉冲序列为0或1,所以将其可视化为黑白点图,如图7所示,括号里为图像数据大小(channel×height×width).从整体上看,脉冲卷积神经网络层数越深,网络对遥感图像的代表越稀疏,尤其第2个全连接层,仅仅用几十个脉冲表示一张遥感图像,并且在此基础上输出层序号为6的神经元发出的脉冲最多,表明网络对该遥感图像所属的场景分类

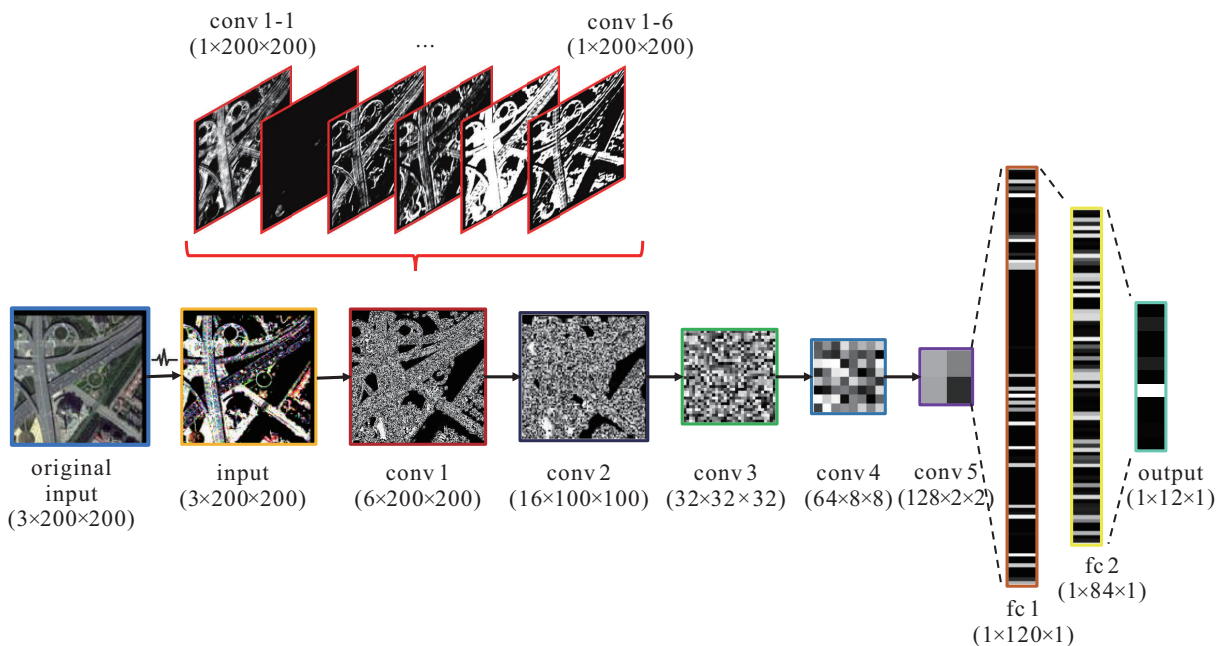


图7 Google数据集中一张场景标签为“立交桥”(标签序号为6)的遥感图像SCNN稀疏表征展示

正确,说明了网络的稀疏表征是有效的,并且对含有噪声的遥感图像具有鲁棒性;从局部上看,以第1个卷积层 conv1 为例,其对图像的特征由6个卷积核共同组成,由图7可以看出,conv1的每个卷积核(conv1-1至conv1-6)对图像的特征具有稀疏性,因此它们组合得到conv1对图像的特征也具有稀疏性(其他卷积层也是如此).由此可以看出,脉冲卷积神经网络能够对地物分布复杂且存在冗余信息的遥感图像进行有效且鲁棒的稀疏表征.

### 3.4 场景分类结果

经过训练和测试,可以得到脉冲卷积神经网络在两个数据集上进行遥感图像场景分类实验的性能.其中:SCNN在Google数据集上的分类准确率为90.00%,在UCM数据集上的分类准确率为81.90%.实验结果表明,SCNN经过训练可以实现遥感图像场景分类.SCNN在两个数据集上的测试混淆矩阵如图8(a)和图9(a)所示.

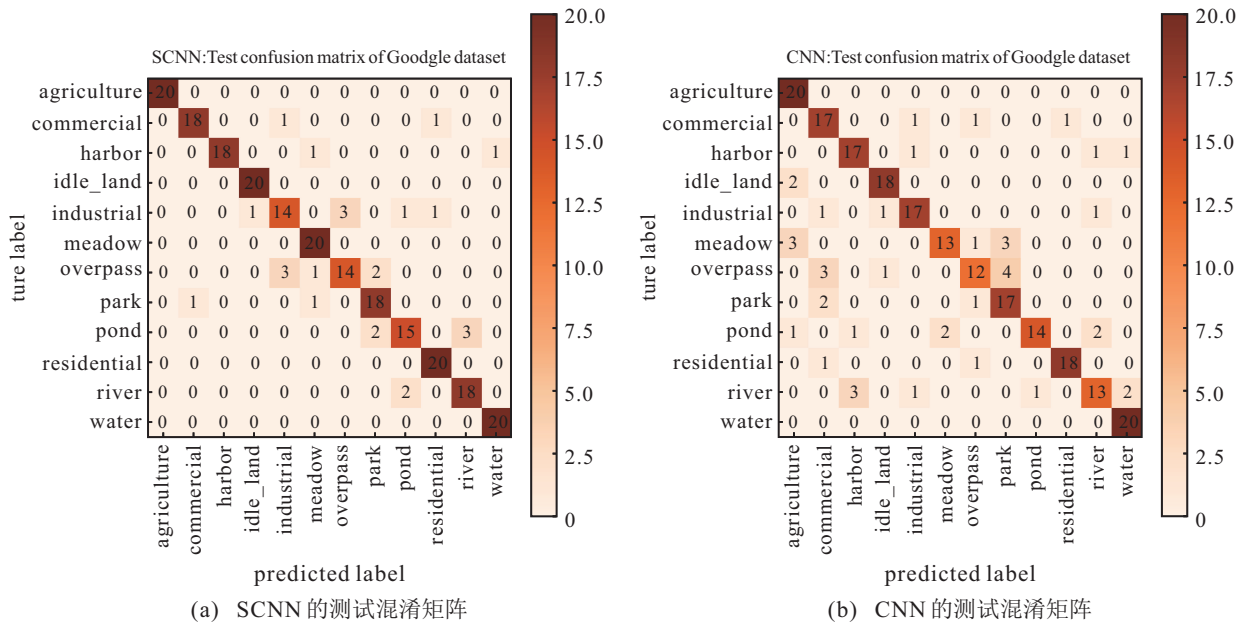


图8 SCNN与CNN在Google数据集上的测试混淆矩阵

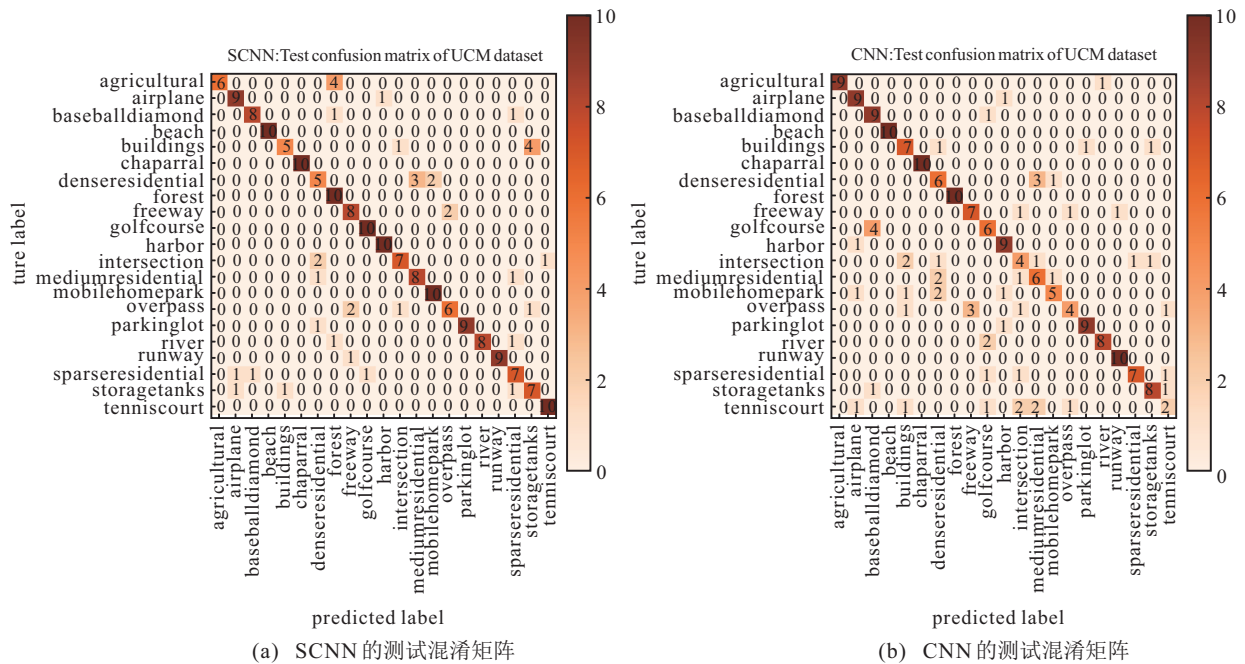


图9 SCNN与CNN在UCM数据集上的测试混淆矩阵

由两个测试混淆矩阵可以看出,在Google数据集有12个场景种类的情况下,SCNN仅仅在立交桥

和工业用地、池塘和河流之间有着轻微的混淆度. UCM数据集的场景种类有21个,其混淆矩阵表明,在



该数据集上, SCNN对农业用地和森林、建筑物和储油罐、高密度住宅区和中密度住宅区这几种场景的混淆度较大. 这说明随着场景的增多, 会出现一些相似的场景(比如高密度住宅区和中密度住宅区), 这时网络对这些相似场景的稀疏表征会出现混淆, 导致场景分类不准确.

为了验证SCNN相较于CNN在遥感图像场景分类任务上的优势, 本文进行了对比实验, 搭建了与图6(a)和图6(b)所示结构相同的两个CNN, 分别用于Google和UCM数据集进行场景分类实验. CNN在两个数据集上的测试混淆矩阵如图8(b)和图9(b)所示, 分类准确度如表3所示. 由对比实验结果可知, 由于脉冲神经元的稀疏脉冲输出特性, 由其组成的SCNN能对遥感图像进行稀疏表征, 去除图像中与场景无关的冗余信息, 进而达到更加准确的分类准确度. 因此, SCNN相较于CNN在遥感图像场景分类任务上更有优势. 除此之外, 本文方法与其他文献方法也进行了比较, 如表3所示. 其中文献[24-25]中的方法应用于Google数据集, 文献[12, 26]中的方法应用于UCM数据集. 由表3可以看出, 本文与其他文献方法相比具有一定的优越性, 并且本文是首次将脉冲卷积神经网络应用于遥感图像场景分类任务, 还具有较大的提升空间.

表3 SCNN在两个数据集上的分类准确度及与其他方法的比较

数据集	方法	分类准确度/%
Google数据集	PLSA <sup>[24]</sup>	89.60
	SPM <sup>[25]</sup>	77.69
	CNN(同本文图6(a)结构)	81.67
	<b>SCNN(本文方法)</b>	<b>90.00</b>
UCM数据集	UFL <sup>[12]</sup>	81.67
	SCK++ <sup>[26]</sup>	77.38
	CNN(同本文图6(b)结构)	73.81
	<b>SCNN(本文方法)</b>	<b>81.90</b>

本文考虑了遥感图像中地物分布复杂及存在冗余信息的数据特点, 利用脉冲神经元组成脉冲卷积神经网络对遥感图像进行稀疏表征, 去除图像中与场景无关的冗余信息, 进而在基于脉冲输出交叉熵损失函数的反向传播算法的基础上利用梯度下降训练该网络, 实现了遥感图像场景分类.

## 4 结论

本文利用脉冲神经元稀疏脉冲输出的特性, 针对遥感图像大小设计脉冲卷积神经网络, 对地物分布复

杂及含有冗余信息的遥感图像进行稀疏表征, 并在提出的基于脉冲输出交叉熵损失函数的反向传播算法的基础上利用梯度下降训练网络, 实现了高分辨率遥感图像场景分类. 实验结果表明: 本文提出的方法能够对遥感图像进行稀疏表征, 并能从复杂冗余的图像中获取简洁有效的信息, 实现场景分类, 相比于传统的卷积神经网络更具有优势. 但随着场景的增多, 会出现相似场景, 本文方法的性能会有所降低, 在后面的研究中可以将本文方法与遥感图像的光谱特征、空间特征、纹理特征等低层特征相结合, 从而提高相似场景的分类准确度; 在遥感图像场景分类任务中, 目前依然是对静态遥感图像进行场景分类, 如果对动态遥感图像进行场景分类, 则需要研究网络脉冲输出的时效性问题.

## 参考文献(References)

- [1] Hossain M D, Chen D M. Segmentation for object-based image analysis(OBIA): A review of algorithms and challenges from remote sensing perspective[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150: 115-134.
- [2] Abdullah A Y M, Masrur A, Adnan M S G, et al. Spatio-temporal patterns of land use/land cover change in the heterogeneous coastal region of Bangladesh between 1990 and 2017[J]. Remote Sensing, 2019, 11(7): 790.
- [3] Ferreira M P, Wagner F H, Aragão L E O C, et al. Tree species classification in tropical forests using visible to shortwave infrared worldview-3 images and texture analysis[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 149: 119-131.
- [4] Lu C Y, Ren C Y, Wang Z M, et al. Monitoring and assessment of wetland loss and fragmentation in the cross-boundary protected area: A case study of Wusuli river basin[J]. Remote Sensing, 2019, 11(21): 2581.
- [5] Tong G, Li Y, Zhang W, et al. Point set multi-level aggregation feature extraction based on multi-scale max pooling and LDA for point cloud classification[J]. Remote Sensing, 2019, 11(23): 2846.
- [6] Luo H, Wang C, Wen Y Q, et al. 3-D object classification in heterogeneous point clouds via bag-of-words and joint distribution adaptation[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(12): 1909-1913.
- [7] Ali N, Zafar B, Iqbal M K, et al. Modeling global geometric spatial information for rotation invariant classification of satellite images[J]. PLoS One, 2019, 14(7): e0219833.
- [8] Ni K, Wang P, Wu Y Q. High-order generalized orderless pooling networks for synthetic-aperture radar scene

- classification[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(11): 1716-1720.
- [9] Bouteldja S, Kourgli A, Aissa A B. Efficient local-region approach for high-resolution remote-sensing image retrieval and classification[J]. *Journal of Applied Remote Sensing*, 2019, 13: 016512.
- [10] Lv Y F, Zhang X H, Xiong W, et al. An end-to-end local-global-fusion feature extraction network for remote sensing image scene classification[J]. *Remote Sensing*, 2019, 11(24): 3006.
- [11] Tian Z Z, Zhan R H, Wang W, et al. Object detection in optical remote sensing images by integrating object-to-object relationships[J]. *Remote Sensing Letters*, 2020, 11(5): 416-425.
- [12] Cheriyyadat A M. Unsupervised feature learning for aerial scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 52(1): 439-451.
- [13] Cheng G, Zhou P C, Han J W, et al. Auto-encoder-based shared mid-level visual dictionary learning for scene classification using very high resolution remote sensing images[J]. *IET Computer Vision*, 2015, 9(5): 639-647.
- [14] Zhang F, Du B, Zhang L P. Saliency-guided unsupervised feature learning for scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(4): 2175-2184.
- [15] Liu Z, Wu X J, Shu Z Q. Sparsity augmented discriminative sparse representation for face recognition[J]. *Pattern Analysis and Applications*, 2019, 22(4): 1527-1535.
- [16] 程龙, 刘洋. 脉冲神经网络: 模型、学习算法与应用[J]. *控制与决策*, 2018, 33(5): 923-937.  
(Cheng L, Liu Y. Spiking neural networks: Model, learning algorithms and applications[J]. *Control and Decision*, 2018, 33(5): 923-937.)
- [17] 胡一凡, 李国齐, 吴郁杰, 等. 脉冲神经网络研究进展综述[J]. *控制与决策*, 2021, 36(1): 1-26.  
(Hu Y F, Li G Q, Wu Y J, et al. Spiking neural networks: A survey on recent advances and new directions[J]. *Control and Decision*, 2021, 36(1): 1-26.)
- [18] Wu Y J, Deng L, Li G Q, et al. Spatio-temporal backpropagation for training high-performance spiking neural networks[J]. *Frontiers in Neuroscience*, 2018, 12: 331.
- [19] McCaffrey J D. Why you should use cross-entropy error instead of classification error or mean squared error for neural network classifier training[EB/OL]. (2013-11-05)[2020-12-24]. <https://jamesmccaffrey.wordpress.com>.
- [20] Daryanavard S, Porr B. Closed-loop deep learning: Generating forward models with backpropagation[J]. *Neural Computation*, 2020, 32(11): 2122-2144.
- [21] Zhao B, Zhong Y F, Zhang L P, et al. The fisher kernel coding framework for high spatial resolution scene classification[J]. *Remote Sensing*, 2016, 8(2): 157.
- [22] Yang Y, Newsam S. Bag-of-visual-words and spatial extensions for land-use classification[C]. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. San Jose, 2010: 270-279.
- [23] Ferrari U, Deny S, Marre O, et al. A simple model for low variability in neural spike trains[J]. *Neural Computation*, 2018, 30(11): 3009-3036.
- [24] Bosch A, Zisserman A, Munoz X. Scene classification using a hybrid generative/discriminative approach[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(4): 712-727.
- [25] Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories[C]. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. New York, 2006: 2169-2178.
- [26] Yang Y, Newsam S. Spatial pyramid co-occurrence for image classification[C]. *2011 International Conference on Computer Vision*. Barcelona, 2011: 1465-1472.

## 作者简介

张哲益(1998—), 男, 博士生, 从事脉冲神经网络、机器学习的研究, E-mail: zzy2015@cug.edu.cn;

曹卫华(1972—), 男, 教授, 博士生导师, 从事过程控制、智能系统和机器人技术等研究, E-mail: weihuacao@cug.edu.cn;

朱蕊(1996—), 女, 硕士生, 从事脉冲神经网络、深度学习的研究, E-mail: 20141004116@cug.edu.cn;

胡文凯(1987—), 男, 教授, 博士生导师, 从事复杂工业系统数据挖掘、智能监控等研究, E-mail: wenkaihu@cug.edu.cn;

吴敏(1963—), 男, 教授, 博士生导师, 从事过程控制、鲁棒控制等研究, E-mail: wumin@cug.edu.cn.

(责任编辑: 孙艺红)