



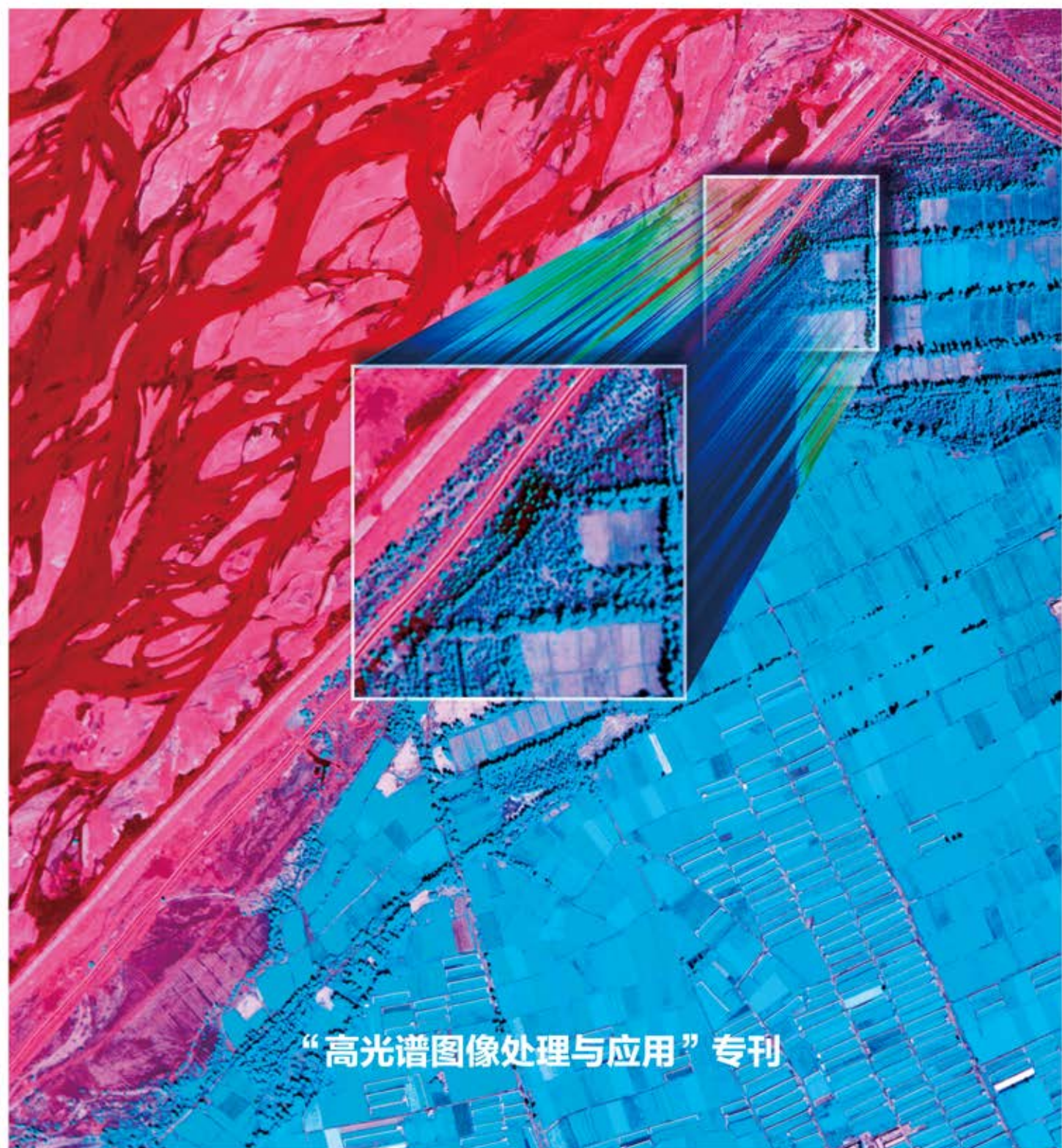
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
08
VOL.26

ISSN1006-8961
CN11-3758/TB



“高光谱图像处理与应用”专刊



第26卷第8期 (总第304期)

2021年8月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院空天信息创新研究院

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 吴一戎

编辑出版 《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@aircas.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China

Zip code 100190

E-mail jig@aircas.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

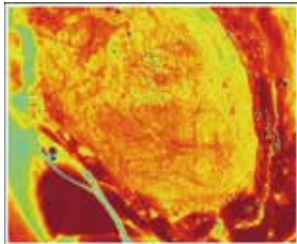
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元



高光谱图像在生物医学中的应用(第1764页)



HSRS-SC:面向遥感场景分类的高光谱图像数据集(第1809页)



多尺度超像素分割和奇异谱分析的高光谱影像分类(第1978页)

序言 何明一 I

编者按 II

综述

高光谱图像空谱特征提取综述

叶珍, 白璘, 何明一 1737

高光谱图像在生物医学中的应用

李伟, 吕蒙, 陈天虹, 楚照耀, 陶然 1764

高光谱成像技术新进展

近红外高光谱图像数据预测技术

梅少辉, 张博威, 马明阳, 贾森 1786

低成本便携式多光谱成像系统的研发及优化

朱豪男, 胡孟晗, 张健, 李庆利 1796

数据集论文

HSRS-SC:面向遥感场景分类的高光谱图像数据集

徐科杰, 邓培芳, 黄鸿 1809

高光谱医学诊断

膜性肾病诊断的高光谱图像张量嵌入分析

吕蒙, 陈天虹, 李伟, 杨悦, 涂天琪, 李文歌 1823

傅里叶变换通道注意力网络的胆管癌高光谱图像分割

郑少佳, 邱崧, 李庆利, 周梅, 胡孟晗, 于观贞 1836

高光谱图像目标检测与识别

局部梯度轮廓变换的高光谱异常检测

胡静, 张钰婧, 赵明华, 李鹏, 李云松 1847

结合孪生网络和像素配对的高光谱图像异常检测

王德港, 饶伟强, 孙旭, 渠瀛, 刘雪梅, 高连如 1860

粗定位和协同表示的高光谱图像异常检测

胡静, 赵明华, 李鹏, 李云松 1871

高光谱图像深度模糊核聚类的洋底锰结核识别

张启忠, 郑恩迪, 王叶剑, 高发荣 1886

高光谱图像融合与恢复

波段自适应细节注入的高分五号与Sentinel-2遥感影像空谱融合

王海荣, 郭擎, 李安 1896

基于频率加权张量核范数的高光谱图像复原

刘盛, 曾海金, 孔文凤, 张鹏丹 1910

高光谱图像分类

面向高光谱图像分类的内容引导卷积深度网络并行实现

刘启超, 肖亮, 杨劲翔 1926

增强类可分性的高光谱图像分类

方帅, 张坤, 张晶, 曹洋, 石为开 1940

高光谱图像分类的自适应决策融合方法

叶珍, 董睿, 陈浩鑫, 白璘 1952

跨数据集评估的高光谱图像分类

潘尔婷, 马泳, 黄珺, 樊凡, 李峰, 马佳义 1969

多尺度超像素分割和奇异谱分析的高光谱影像分类

付航, 孙根云, 赵云华, 潘兆杰, 胡光, 张爱竹 1978

结合倒置特征金字塔和U-Net的高光谱图像分类

程嵩阳, 熊玉洁, 姚瑶, 李庆利 1994

高光谱图像小样本分类的卷积神经网络方法

吴鸿昊, 王立国, 石瑶 2009

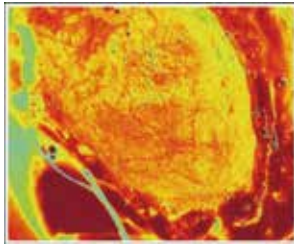
3D卷积自编码器高光谱图像分类模型

石延新, 何进荣, 李照奎, 曾志高 2021

封面素材由重庆大学光电工程学院黄鸿教授提供

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Application of a hyperspectral image in medical field: a review (P1764)



HSRS-SC: a hyperspectral image dataset for remote sensing scene classification (P1809)



Combining multiscale superpixel segmentation and singular spectral analysis for hyperspectral image classification (P1978)

Review

- Review of spatial-spectral feature extraction for hyperspectral image
Ye Zhen, Bai Lin, He Mingyi 1737
- Application of a hyperspectral image in medical field: a review
Li Wei, Lyu Meng, Chen Tianhong, Chu Zhaoyao, Tao Ran 1764

Advances in Hyperspectral Imaging

- Predicting near-infrared hyperspectral images from visible hyperspectral images
Mei Shaohui, Zhang Bowei, Ma Mingyang, Jia Sen 1786
- Development and optimization of a low-cost and portable multispectral imaging system
Zhu Haonan, Hu Menghan, Zhang Jian, Li Qingli 1796

Dataset

- HSRS-SC: a hyperspectral image dataset for remote sensing scene classification
Xu Kejie, Deng Peifang, Huang Hong 1809

Medical Hyperspectral Imagery

- Tensor-based graph embedding for discriminant analysis of membranous nephropathy hyperspectral data
Lyu Meng, Chen Tianhong, Li Wei, Yang Yue, Tu Tianqi, Li Wen'ge 1823
- Fourier transform channel attention network for cholangiocarcinoma hyperspectral image segmentation
Zheng Shaojia, Qiu Song, Li Qingli, Zhou Mei, Hu Menghan, Yu Guanzhen 1836

Hyperspectral Target Detection and Recognition

- Transformation of local gradient profiles for hyperspectral anomaly detection
Hu Jing, Zhang Yujing, Zhao Minghua, Li Peng, Li Yunsong 1847
- Siamese network with pixel-pair for hyperspectral image anomaly detection
Wang Degang, Rao Weiqiang, Sun Xu, Qu Ying, Liu Xuemei, Gao Lianru 1860
- Rough location and collaborative representation for hyperspectral image anomaly detection
Hu Jing, Zhao Minghua, Li Peng, Li Yunsong 1871
- Recognition of ocean floor manganese nodules by deep kernel fuzzy C-means clustering of hyperspectral images
Zhang Qizhong, Zheng Endi, Wang Yejian, Gao Farong 1886

Hyperspectral Image Fusion and Restoration

- Spatial-spectral fusion based on band-adaptive detail injection for GF-5 and Sentinel-2 remote sensing images
Wang Hairong, Guo Qing, Li An 1896
- Hyperspectral image restoration based on frequency-weighted tensor nuclear norm
Liu Sheng, Zeng Haijin, Kong Wenfeng, Zhang Pengdan 1910

Hyperspectral Image Classification

- Parallel implementation of content-guided deep convolutional network for hyperspectral image classification
Liu Qichao, Xiao Liang, Yang Jinxiang 1926
- Hyperspectral image classification with enhanced class separability
Fang Shuai, Zhang Kun, Zhang Jing, Cao Yang, Shi Weikai 1940
- Adjustive decision fusion approaches for hyperspectral image classification
Ye Zhen, Dong Rui, Chen Haoxin, Bai Lin 1952
- Hyperspectral image classification evaluated across different datasets
Pan Erting, Ma Yong, Huang Jun, Fan Fan, Li Hao, Ma Jiayi 1969
- Combining multiscale superpixel segmentation and singular spectral analysis for hyperspectral image classification
Fu Hang, Sun Genyun, Zhao Yunhua, Pan Zhaojie, Hu Guang, Zhang Aizhu 1978
- Hyperspectral image classification using an inverted feature pyramid network with U-Net
Cheng Songyang, Xiong Yujie, Yao Yao, Li Qingli 1994
- Convolution neural network method for small-sample classification of hyperspectral images
Wu Honghao, Wang Liguang, Shi Yao 2009
- Hyperspectral image classification model based on 3D convolutional auto-encoder
Shi Yanxin, He Jinrong, Li Zhaokui, Zeng Zhigao 2021

中图法分类号: TP751.1 文献标识码: A 文章编号: 1006-8961(2021)08-1969-09

论文引用格式: Pan E T, Ma Y, Huang J, Fan F, Li H and Ma J Y. 2021. Hyperspectral image classification evaluated across different datasets. Journal of Image and Graphics, 26(08): 1969-1977 (潘尔婷, 马泳, 黄珺, 樊凡, 李焱, 马佳义. 2021. 跨数据集评估的高光谱图像分类. 中国图象图形学报, 26(08): 1969-1977) [DOI:10.11834/jig.210123]

跨数据集评估的高光谱图像分类

潘尔婷¹, 马泳^{1,2}, 黄珺^{1,2*}, 樊凡^{1,2}, 李焱³, 马佳义^{1,2}

1. 武汉大学电子信息学院, 武汉 430072; 2. 武汉大学宇航科学与技术研究院, 武汉 430079;

3. 武汉轻工大学数学与计算机科学学院, 武汉 430023

摘要: 目的 随着高光谱成像技术的飞速发展, 高光谱数据的应用越来越广泛, 各场景高光谱图像的应用对高精度详细标注的需求也越来越旺盛。现有高光谱分类模型的发展大多集中于有监督学习, 大多数方法都在单个高光谱数据立方体中进行训练和评估。由于不同高光谱数据采集场景不同且地物类别不一致, 已训练好的模型并不能直接迁移至新的数据集得到可靠标注, 这也限制了高光谱图像分类模型的进一步发展。本文提出跨数据集对高光谱分类模型进行训练和评估的模式。**方法** 受零样本学习的启发, 本文引入高光谱类别标签的语义信息, 拟通过将不同数据集的原始数据及标签信息分别映射至同一特征空间以建立已知类别和未知类别的关联, 再通过将训练数据集的两部分特征映射至统一的嵌入空间学习高光谱图像视觉特征和类别标签语义特征的对应关系, 即可将该对应关系应用于测试数据集进行标签推理。**结果** 实验在一对同传感器采集的数据集上完成, 比较分析了语义—视觉特征映射和视觉—语义特征映射方向, 对比了 5 种基于零样本学习的特征映射方法, 在高光谱图像分类任务中实现了对分类模型在不同数据集上的训练和评估。**结论** 实验结果表明, 本文提出的基于零样本学习的高光谱分类模型可以实现跨数据集对分类模型进行训练和评估, 在高光谱图像分类任务中具有一定的发展潜力。

关键词: 高光谱图像分类; 深度学习; 特征提取; 零样本学习; 高光谱语义特征

Hyperspectral image classification evaluated across different datasets

Pan Erting¹, Ma Yong^{1,2}, Huang Jun^{1,2*}, Fan Fan^{1,2}, Li Hao³, Ma Jiayi^{1,2}

1. Electronic Information School, Wuhan University, Wuhan 430072, China; 2. Institute of Aerospace Science and Technology, Wuhan University, Wuhan 430079, China; 3. College of Mathematics and Computer Science,

Wuhan Polytechnic University, Wuhan 430023, China

Abstract: **Objective** Hyperspectral sensors are evolving towards miniaturization and portability with the rapid development of hyperspectral imaging technology. The acquisition of hyperspectral data has become easier and less expensive as a result of this breakthrough. The broad applications of hyperspectral images in various scenes have arisen an increasing demand for high-precision and detailed annotations. The growth of existing hyperspectral classification models mainly focuses on supervised learning, and many of them have reached an almost perfect performance. However, nearly all of these models are trained and evaluated in a single hyperspectral data cube. On this condition, the trained classification model cannot be directly transferred to a new hyperspectral dataset to obtain reliable annotations. The main reason is that different hyper-

收稿日期: 2021-02-26; 修回日期: 2021-05-17; 预印本日期: 2021-05-24

* 通信作者: 黄珺 junwong@whu.edu.cn

基金项目: 国家自然科学基金项目(61903279); 湖北省自然科学基金项目(2019CFB162, 2018CFA006)

Supported by: National Natural Science Foundation of China (61903279); Natural Science Foundation of Hubei Province, China (2019CFB162, 2018CFA006)

spectral datasets are collected in irrelevant scenes and have covered inconsistent object categories. Accordingly, existing hyperspectral classification models have poor generalization capacity. The further development of the hyperspectral image classification model is also constrained. Consequently, a hyperspectral classification model with a generalization capability and ability to adapt to new and unseen classes across the different datasets must be developed. In this study, we propose a specific unique paradigm, which is to train and evaluate a hyperspectral classification model across different datasets. As previously mentioned, some new and unseen categories might be encountered when evaluating the classification model on another hyperspectral dataset. We introduce zero-shot learning into hyperspectral image classification to address this problem. **Method** Zero-shot learning can distinguish data from unseen categories, except for identifying data categories that have appeared in the training set. Zero-shot learning is based on the principle of allowing the model learn to understand the semantics of categories at first. Specifically, this mechanism employs auxiliary knowledge (such as attributes) to embed category labels into the semantic space and uses the data in the training set to learn the mapping relationship from images to semantics. On the basis of zero-shot learning, we introduce hyperspectral category label semantics as side information to address the unseen categories in the classification across different datasets. The overall workflow of our model can be divided into three steps. The first step is feature extraction, including hyperspectral image visual feature extraction and label semantics extraction. Hyperspectral image feature extraction aims to obtain high-level hyperspectral features with great capability to distinguish, and we refer to them as visual features. Most existing hyperspectral classification models are designed to extract robust hyperspectral features and have fine classification performance. Hence, we directly fine-tune current models to embed hyperspectral images into visual feature space. In the label semantic extraction, word2vec can map each word to a vector, representing the relationship between words. We employ word2vec model trained on a large external corpus to obtain hyperspectral label vectors. This model embeds a hyperspectral category label into a label semantic space. The second step is feature mapping. According to the choice of embedding space, this mechanism can be divided into visual to semantic feature mapping and semantic to visual feature mapping. Feature mapping is used to map these two features into the same feature space, and the model learns and optimizes the mapping on the basis of the correspondence between the hyperspectral data and annotations in the training set. The learned mapping establishes the correspondence between the hyperspectral visual features and the label semantic features. The third step is to employ this learned mapping to perform label reasoning on the testing set. Specifically, the mapping is applied to the hyperspectral visual features and label semantic features of the testing data, and the corresponding labels for the test set data are inferred by measuring the similarity of the two features. **Result** We selected a pair of datasets collected by the same type of hyperspectral sensor for comparative experiments, namely, the Salinas and Indian Pines datasets. This step is conducted to avoid the issue of differences in the physical representation of the object spectra resulting from hyperspectral data collection by different hyperspectral sensors. The workflow of our method is divided into three steps; thus, we adopt different models in each step for comparative experiments. We compared two visual feature extraction models, different feature mapping models, including visual to semantic feature mapping models, and semantic to visual feature mapping models. The quantitative evaluation metric is top- k accuracy. We listed the results of top-1 accuracy and top-3 accuracy. Experimental results show that the method that employs spatial-spectral united network (SSUN) as the visual feature extraction model and relation network (RN) as the zero-shot learning model reaches the best performance. Comparative experiments in different visual feature extraction models demonstrate that SSUN can obtain more distinguishable features because it considered the features in the spatial and spectral domains. When compared with the results from all models related to feature mapping, the semantic to visual feature mapping models outperform the other approaches. This result indicates that choosing visual feature space as the embedding space is a preferable alternative. We also analyze the reason for the current unsatisfied classification performances in detail. **Conclusion** In this study, we propose a specific pattern to train and evaluate the classification model across different datasets to improve the poor migration ability of the existing classification model. We introduce the semantic description of category labels in hyperspectral classification to cope with new and unseen categories in a new hyperspectral dataset and establish the association between seen and new datasets. The experimental results show that the feature extraction model SSUN improves the performance, and the semantic to visual feature mapping model RN outperforms several approaches. In conclusion, the proposed hyperspectral classification model based on zero-shot learning can be trained and evaluated across different datasets. The experimental results indi-

cate that the proposed model has certain development potential in hyperspectral image classification tasks.

Key words: hyperspectral image classification; deep learning; feature extraction; zero-shot learning; semantic features of hyperspectral data

0 引言

高光谱图像相比传统遥感图像具有更高的光谱分辨率,通常以数据立方体的形式呈现。丰富的空间和光谱信息极大地增强了图像的感知能力,使高光谱遥感技术广泛应用于各个领域,如环境监测、资源勘探、军事侦察和灾难评估等(童庆禧等,2016;张兵,2016;Li等,2019)。

高光谱图像分类是指根据图像中的光谱和空间信息对每个像素点进行类别标记,是高光谱数据分析中最关键的技术之一(Li等,2016;冉琼等,2018;Jiang等,2019)。与传统高光谱图像分类方法不同,深度学习方法凭借其具有可挖掘高光谱图像深层次特征的优势,已经成为高光谱图像分类的有力工具。但是,由于目前高光谱图像的公开数据集少,有标注样本少并且人工标注成本昂贵,几乎所有现有的高光谱图像分类方法都在单个数据集上进行训练和测试(Mou等,2017;Mei等,2019;李冠东等,2019)。这导致高光谱分类模型的可靠性和精度存在被过高估计的可能(Paoletti等,2019),也使得现有模型泛化能力差,不能应用于新增数据,难以应对高光谱图像获取成本降低但高光谱精细标注需求激增的现状。

在计算机视觉领域,现有图像分类模型的发展大多集中于有监督学习,对有标注数据的数量和质量有越来越高的要求。而实际应用中更常面对的是样本质量不高、有标签样本数量少的问题,这也激发了小样本学习甚至零样本学习技术的发展(Xian等,2019)。迁移学习是这类问题的主流解决方案,零样本学习是迁移学习的一种极端情况,指依据现有已知类别的信息推断未见过的类别,这里的“零”指的是测试的样本类别在训练集中没有出现过。为了推断未见过的类别信息,零样本学习引入了自然语言处理中的语义信息来生成对类别的高维语义描述,其核心思想是通过在已知类别中建立视觉特征和语义特征之间的联系来实现新类别的识别。

在高光谱图像数据集中,不同场景采集到的数据通常覆盖了很多不同类别的地物。主流的高光谱图像分类方法在单个数据集上进行训练和测试,无法使分类模型学习到泛化能力。受零样本学习的启发,本文将地物类别的语义信息引入高光谱图像分类中,通过建立高光谱图像中的光谱及空间信息以及地物类别语义信息之间的联系,从而实现跨数据集对分类模型进行训练和测试,降低算法对数据集的依赖和标注的压力,提升模型的泛化性能。

本文工作引入了源自外部语料库中训练的自然语言处理模型,对比了两种高光谱图像视觉特征提取算法,在高光谱图像分类任务中分析了5种零样本学习算法,实现了对高光谱图像分类模型进行跨数据集的训练和测试。实验结果表明,零样本学习在该任务中有一定的应用潜力,跨数据集进行训练和测试的模式可以成为未来高光谱分类的发展趋势之一。

1 算法

高光谱图像跨数据集分类模型中涉及两个不同场景的数据集,其具体实现分为3个步骤:1)特征提取,包括高光谱数据的视觉特征提取和类别标签的语义特征提取;2)特征映射,用于在训练集中建立视觉特征和语义特征之间的联系,学习两种特征之间的映射关系;3)标签推理,即将训练得到的映射关系应用于测试集数据中,将测试集高光谱数据的视觉特征与测试集类别标签的语义特征对应,推理得到最终标签。

1.1 特征提取

对于高光谱图像训练数据集 $\mathbf{x}^{\text{tr}}$ 和高光谱图像测试数据集 $\mathbf{x}^{\text{te}}$,本文通过特征提取将高光谱图像数据映射到视觉特征空间,将地物类别名称映射到类别语义特征空间。如图1所示, $\mathbf{x}^{\text{tr}}$ 、 $\mathbf{x}^{\text{te}}$ 分别对应训练集和测试集的高光谱数据立方体, $\mathbf{y}^{\text{tr}}$ 、 $\mathbf{y}^{\text{te}}$ 分别对应训练集和测试集的类别名称列表。图像特征提取模型 $\phi(x)$ 将训练集和测试集的数据 $\mathbf{x}^{\text{tr}}$ 、 $\mathbf{x}^{\text{te}}$ 编码至视觉特征空间,得到对应的视觉特征 $\phi(\mathbf{x}^{\text{tr}})$ 、 $\phi(\mathbf{x}^{\text{te}})$ 。

而类别语义描述模型 $\psi(y)$ 则将训练集和测试集中的类别名称 y^{tr} , y^{te} 编码为类别语义特征 $\psi(y^{\text{tr}})$, $\psi(y^{\text{te}})$ 。

图像视觉特征提取旨在将高光谱图像中的目标像素相关的光谱和空间信息编码至视觉特征空间中,为训练集和测试集图像提取鲁棒的高光谱视觉特征。高光谱图像具有丰富的空间信息和光谱信息。其中,空间信息依赖于空间分辨率的高低,表征了地物的空间分布,而光谱信息通常表征为一条连续且精细的特征曲线,可以直接反映地物的特征。

在高光谱图像分类领域已经提出了多种性能优越的基于深度神经网络的分类模型,这类分类模型大多数是首先进行特征提取,再将提取到的特征输入分类器来进行分类。结合神经网络层次化特征提取的特性,可以灵活使用分类模型的中间层激活来输出较高层次的视觉特征。经特征提取后得到的特征能更准确细致地描述地物特性,本文中选取了两个不同的视觉特征提取模型进行实验,利用特征提取模型将训练集和测试集的数据嵌入至同一视觉特征空间,建立两个数据集之间的联系。

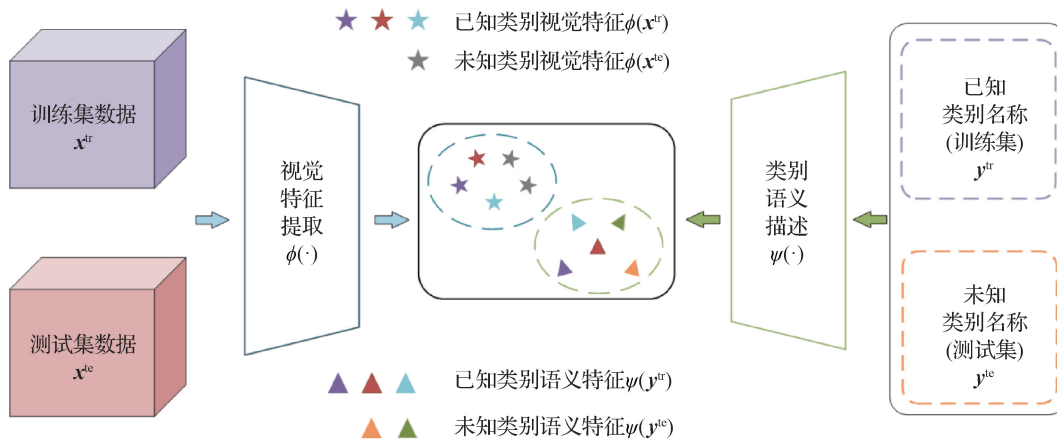


图1 特征提取图示

Fig. 1 Illustration of feature extraction

类别语义空间的构建可以将训练集和测试集中的类别映射到同一语义特征空间。现有的零样本学习方法中主要使用两种类型的类别语义描述,即属性和词向量。属性描述的是不同类别是否具有某一特性,通常是人工标注的,而高光谱数据中缺少这一信息来源。词向量是把类别名称的文本转换为向量空间中的向量,使向量空间中不同向量表示的相似度代表文本内容的相似度。词向量的转换是借助于在外部语料库中训练的自然语言处理模型完成的,该模型通常在一个大型无标注的文本语料库中预训练,通过学习单词之间的语义关联,将每个单词表示为一个固定维度的向量。本文使用词向量作为语义特征描述,以将高光谱不同数据集的地物类别名称通过词向量模型映射至同一特征空间,得到不同类别的语义描述。这里需要强调的是,特征提取的过程中,不涉及标签与数据的一一对应关系,本文仅使用训练集和测试集数据的类别名称。通过词向量转换模型,建立训练集和测试集类别名称之间的关联。

1.2 特征映射

在特征提取阶段,训练集数据和测试集数据通过特征提取模型建立了关联,利用视觉特征提取模型将训练集和测试集高光谱数据嵌入了同一视觉特征空间,已知类别和未见类别之间的语义关联也已经通过词向量转换模型建立,嵌入了同一语义特征空间。但得到的高光谱视觉特征和类别标签语义特征仍属于不同的特征空间,因此需要学习特征映射将视觉特征和语义特征关联起来,即

$$f = W[\phi(x^{\text{tr}}), \psi(y^{\text{tr}})] \quad (1)$$

在零样本学习中,根据嵌入空间的选择,特征映射有两种选择,如图2所示。一个是在语义—视觉特征映射 ($V \rightarrow S$),指将语义特征投影至视觉特征空间,另一种是视觉—语义特征映射 ($S \rightarrow V$),即将视觉特征投影至语义特征空间。目标函数的设计则是针对如何衡量映射后两部分特征的相似性。

语义—视觉特征映射方式的典型方法是 De-Vise (deep visual embedding) (Frome 等, 2013) 和 ALE (attribute label embedding) (Lampert 等,

2014), 都选择了语义特征空间作为嵌入空间。DeVise 方法中设计了基于排名损失的双线性相容函数, 而 ALE 学习的是一种基于铰链排名损失优化的映射。视觉—语义特征映射方式的代表方法有 DEM (deep embedding model) (Zhang 等, 2017) 和 RN (relation network) (Sung 等, 2018), 这两种方法都设计了一个子网络将视觉特征映射到语义空间。

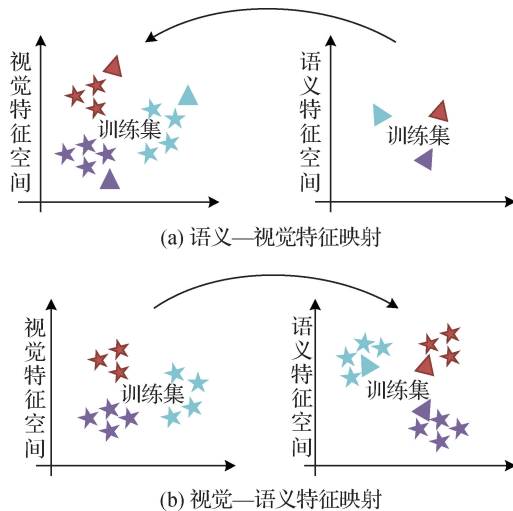


图2 特征映射图示

Fig. 2 Illustration of feature mapping ((a) semantic-visual feature mapping; (b) visual-semantic feature mapping)

1.3 标签推理

在训练集数据上学习到特征映射后, 建立了高光谱图像视觉特征和地物类别标签的语义特征的对应关系, 下一步是利用这一映射对测试集进行标签推理。如图3所示, 标签推理步骤即将该映射应用在测试集数据的视觉特征和类别语义特征中, 再通过度量两种特征的相似性来为测试集数据推理出对应的标签。

最常用的度量学习方法是使用最近邻分类器, 该方法依据两种特征的最近距离进行分类。上文介绍的典型算法中, DeVise 和 ALE 的优化目标是训练过程中基于边界的排名损失函数。另外还有采用线性回归来获得投影矩阵 W 的 SAE (semantic auto-encoder) (Kodirov 等, 2017), 通过计算投影后两特征之间的余弦距离, 再使用 K 均值聚类来推理标签。

DEM 中的标签推理是通过简单的距离度量来实现的, 在视觉特征空间中使用 K 最近邻搜索, 将

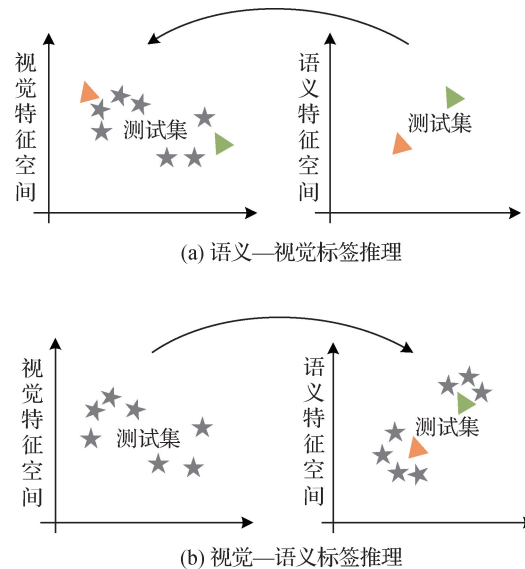


图3 标签推理图示

Fig. 3 Illustration of label reasoning ((a) semantic-visual label reasoning; (b) visual-semantic label reasoning)

视觉特征与语义特征的投影进行匹配。与前面方法中的固定度量或浅层学习度量不同, RN 选择通过深度学习与嵌入相结合的非线性相似度量来识别视觉特征和语义特征是否匹配, 回归得到相关程度的分数, 最大得分为 1, 不匹配则得分为 0。

2 实验

2.1 实验数据

本文选择了 Salinas 数据集 (SA) 和 Indian Pines 数据集 (IP) 进行对比实验, 其伪彩图和对应的参考标签图如图4和图5所示。选择这两个公开数据集有以下原因: 1) SA 和 IP 数据集都是由 AVIRIS 传感器采集的, 光谱波段范围为 400 ~ 2 500 nm; 2) 两个数据集均采集自农作物区域, 包含不同农作物类别。各类名称和对应的有标签样本数目如表1所示。另外, Salinas 数据集大小为 512 × 217 × 204, Indian Pines 数据集大小为 145 × 145 × 200。

2.2 实验设置

本文实验进行跨数据集的训练和测试, 设计了两组实验, 一是以 SA 为训练集, IP 为测试集, 记为 SA-IP, 二是以 IP 为训练集, SA 为测试集, 记为 IP-SA。如表1所示, IP 中有几个类别的有标签样本过少, 容易导致类别不平衡问题, 因此本文中暂不考虑样本数少于200的类别, 因此参与训练和测试的是

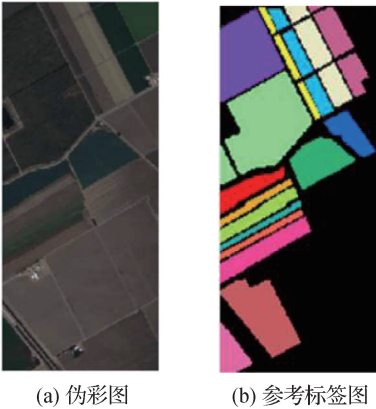


图4 Salinas 数据集的伪彩图和参考标签图
Fig. 4 The false color image and the corresponding ground-truth map of Salinas dataset((a) false-color image; (b) ground truth)

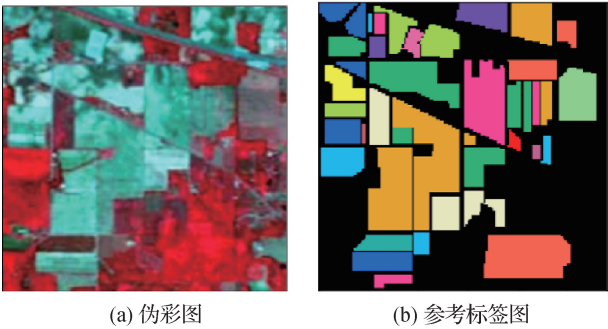


图5 Indian Pines 数据集的伪彩图和参考标签图
Fig. 5 The false color image and the corresponding ground-truth map of Indian Pines dataset ((a) false-color image; (b) ground-truth)

表1 Salinas(SA)和 Indian Pines(IP) 数据集的类别名称及有标签样本数量表
Table 1 Categories and the corresponding number of labeled samples of Salinas and Indian Pines datasets

编号	SA			IP		
	标签名称	数量	修正标签	标签名称	数量	修正标签
1	Brocoli_green_weeds_1	2 009	Broccoli green weeds 1	Alfalfa	46	—
2	Brocoli_green_weeds_2	3 726	Broccoli green weeds 2	Corn-notill	1 428	Corn no till
3	Fallow	1 976	Fallow	Corn-mintill	830	Corn min till
4	Fallow_rough_plow	1 394	Fallow rough plow	Corn	237	Corn
5	Fallow_smooth	2 678	Fallow smooth	Grass-pasture	483	Grass pasture
6	Stubble	3 959	Stubble	Grass-trees	730	Grass trees
7	Celery	3 579	Celery	Grass-pasture-mowed	28	—
8	Grapes_unstrained	11 271	Grapes unstrained	Hay-windrowed	478	Hay windrow
9	Soil_vinyard_develop	6 203	Soil Vinyard develop	Oats	20	—
10	Corn_seneseed_green_weeds	3 278	Corn green weeds	Soybean-notill	972	Soybean no till
11	Lettuce_romaine_4wk	1 068	Lettuce romaine 4	Soybean-mintill	2 455	Soybean min till
12	Lettuce_romaine_5wk	1 927	Lettuce romaine 5	Soybean-clean	593	Soybean clean
13	Lettuce_romaine_6wk	916	Lettuce romaine 6	Wheats	205	Wheats
14	Lettuce_romaine_7wk	1 070	Lettuce romaine 7	Woods	1 265	Woods
15	Vinyard_untrained	7 268	Vinyard untrained	Building-Grass-Trees-Drives	386	Building Grass Trees Drives
16	Vinyard_vertical_trellis	1 807	Vinyard vertical trellis	Stone-Steel-Towers	93	—

注：“—”表示该类别由于有标签样本数过少(低于100),不参与训练和测试。

IP 中的 12 个类别和 SA 中的 16 个类别。

本文实验选择 top-*k* 的准确率作为性能评价指标。其他实验设置细节如下：

1) 视觉特征提取模型。在原始的高光谱数据立方逐像素的提取光谱—空间特征描述,本文沿用

已提出的高光谱图像分类模型来进行高光谱图像的视觉特征提取,分离原模型中的最后一个全连接层,得到对应的视觉特征。本文中选择了两种视觉特征提取模型进行实验,一个是使用光谱特征的 RNN (recurrent neural network) 模型(Mou 等,2017),该网

络利用光谱的序列特性使用循环神经网络建模, 本文利用该网络提取的视觉特征维度为 128 维; 另一个是使用光谱—空间特征的 SSUN (spectral-spatial united network) 模型 (Xu 等, 2018), 该网络利用两个分支网络分别提取空间特征和光谱特征, 并建立联合网络进一步提取空谱联合特征, 本文利用该网络提取的视觉特征为 512 维。

2) 语义空间构建。高光谱数据集中很多类别都是由多个单词组成的, 需要进行预处理操作。首先删除类别名称中各个单词中间的连接线, 然后利用训练好的模型, 在已知语料库中查询并加载出每个单词对应的词向量, 然后加和求平均值, 得到每一个类别名称所对应的词向量。但由于现有的开源语料库规模有限, 并不是所有的类别名称都能在语料库中找到对应的单词, 因此, 本文对某些类别的名称进行了修正和微调。比如“Brocoli”修正为“Broccoli”、“notill”以及“mintill”调整为“no till”和“mintill”、“Lettuce_romaine_4wk”中的“4wk”表示的是生长周期, 但是这并不是一个正常的单词, 因此本文将该类别调整为“Lettuce romain 4”等。本文实验中具体使用的修正后的类别名称列表如表 1 所示。本文使用了在谷歌新闻数据集上预训练的 word2vec 模型 (Mikolov 等, 2013), 该模型中包含 300 万单词和短语的词向量, 维度为 300 维。模型 word2vec 是用来产生词向量的经典模型, 使用浅层的双层网络将独热编码形式的词向量映射到分布式形式的词向量。通过 word2vec 模型, 可以将高光谱不同数据集的地物类别标签映射至同一特征空间, 得到每个类别名称的语义描述特征。

3) 特征映射。特征映射用于建立高光谱图像视觉特征以及地物类别语义特征之间的关联。本文中对视觉—语义特征映射及语义—视觉特征映射方法均进行了实验验证, 包括 DeVise, WLE (word 2vec label embedding), SAE, DEM 和 RN 这 5 种算法, 实验中所涉及各个对比算法的相关参数设置如表 2 所示。

2.3 实验结果

表 3 中列出了主流分类模式和本文分类模式下同一特征提取模型的性能对比。主流分类模式在单个数据集中进行训练—测试, 本文提出的分类模式是跨数据集进行训练—测试。实验中, 训练样本的选取为训练数据集每个类别有标签样本的 10%, 训

表 2 各模型相关参数设置表

Table 2 Parameter configuration of comparative methods

模型	参数
DeVise (Frome 等, 2013)	$Margin = 0.1; r = 0.0001$
WLE (Lampert 等, 2014)	$r = 0.001$
SAE (Kodirov 等, 2017)	$\lambda = 100000; r = 0.0001$
DEM (Zhang 等, 2017)	$\lambda = 0.0001; r = 0.0001$
RN (Sung 等, 2018)	$r = 0.0001$

注: $Margin$ 为损失函数的边界项, r 为学习率, SAE 中 λ 为投影矩阵的正则化系数, DEM 中 λ 为损失函数的正则化系数。

表 3 两种分类模式的对比实验结果

Table 3 Experimental results of two classification patterns

				/%
		RNN	SSUN	
训练—测试	SA-SA	SA-IP	SA-SA	SA-IP
总体精度	93.42	8.23	99.78	10.17
训练—测试	IP-IP	IP-SA	IP-IP	IP-SA
总体精度	94.97	6.03	99.97	6.67

练轮次为 200, 批次大小为 128。实验结果显示, 如 SSUN 模型, 在 Salinas 数据集中单独进行训练—测试时, 其总体分类精度可以达到 99.78%, 但将模型直接迁移至 Indian Pines 数据集中应用时, 精度仅为 10.17%。这说明在单个数据集中进行训练—测试的模式具有过拟合的问题, 不能适应新数据中地物类别以及空间分布上的差异, 这种分类模式下训练出来的模型难以泛化至新的数据集。

各对比算法的实验结果如表 4 所示。SA-IP 组测试集有 12 个类别, IP-SA 组有 16 个类别。首先, 两种不同特征提取算法 RNN 和 SUNN 的对比实验结果表明, 联合使用光谱和空间特征有更好的分类性能。这是由于 SA 和 IP 数据集中大都是农作物类别, 包含大量的同质区域, 空间特征的利用有较大作用, 而光谱特征又能够表征地物的光谱特性。因此, 使用联合提取空间和光谱特征的模型能得到更鲁棒的高光谱视觉特征。

对比视觉—语义特征映射 (visual→semantic, $V \rightarrow S$) 和语义—视觉特征映射 (semantic→visual, $S \rightarrow V$) 两种不同映射方式的实验结果, 可以看出, 语义—视觉特征映射的几种方法在实验中有更好的表现, 尤其是 DEM 和 RN 两种方法。因此, 更优的映

表 4 各对比算法的实验结果
Table 4 Experimental results of comparative methods

精度	模型	映射方向	/%			
			SA-IP		IP-SA	
			RNN	SSUN	RNN	SSUN
top-1	DeVise	V→S	13.46	15.02	10.95	12.57
	WLE	V→S	8.57	10.89	6.82	7.33
	SAE	V→S	10.14	11.03	6.04	8.46
		S→V	11.03	12.09	7.92	7.29
	DEM	S→V	16.65	20.87	10.63	9.05
	RN	S→V	19.87	13.75	13.26	14.11
top-3	DeVise	V→S	28.36	31.21	24.52	26.83
	WLE	V→S	22.49	23.76	16.14	20.04
	SAE	V→S	23.57	26.88	15.37	23.15
		S→V	27.69	25.41	16.94	22.97
	DEM	S→V	28.35	34.02	23.85	26.08
	RN	S→V	31.67	28.36	29.82	37.46

注:top-1 和 top-3 都是 top-*k* 精度,表示预测结果中概率最大的前 *k* 个结果包含正确标签的占比。加粗字体为每列最优值。

射方向选择是将语义特征投影嵌入至视觉特征空间中寻找对齐。此外,在所有特征映射的模型中,RN 这种使用子网络学习非线性映射的模型在分类精度上有更好的表现。

此外,由于这两组数据集中包含一些语义相近但光谱差距大的类别,如“Brocoli_green_weeds_1”和“Brocoli_green_weeds_2”,还有“Soybean-notill”,“Soybean-mintill”和“Soybean-clean”等,这些类别名称中含有共有的单词,在语义特征上较难区分。但由于其光谱差距较大,在数据集中定义成了不同的地物类别,导致标签推理结果中产生误分类。这部分相近类别在语义描述上的区分难度对整体模型的性能也有一定影响。

3 结 论

现有高光谱图像分类由于在单个数据集中进行训练和测试,这导致了高光谱分类模型迁移能力差,无法应对新数据集。本文提出了跨数据集对分类模型进行训练和评估的方式。由于高光谱数据采集的差异,跨数据集会带来新的没见过的类别。受零样

本学习的启发,本文在高光谱分类中引入类别标签的语义描述信息,以建立已知数据集和新数据集的关联。

新模式下的高光谱分类模型工作流程可以分为 3 步:特征提取、特征映射以及标签推理。首先使用特征提取模型分别提取图像的视觉特征和标签的语义描述,然后通过特征映射学习视觉特征和语义特征之间的关联,最后将学习到的映射应用在测试集数据中推理得到标签。

为了保证光谱数据特征的一致性,本文选择使用同一传感器采集的两个数据集进行实验。实验结果表明,同时考虑光谱和空间特征的视觉特征提取模型能捕获类别可分性更强的特征;视觉—语义特征映射的模式能更好地学习到两种特征之间的联系。零样本学习以及标签语义描述的引入在高光谱图像分类任务中有一定的发展潜力。但是,目前的跨数据集评估的分类性能不尽如人意,还有很大的提升空间。这和高光谱数据集本身的类别精细度以及语义描述的领域相关性也有很大的关联。在下一步的研究中,将更多关注于这一问题,以提升模型的性能与可迁移能力。

参考文献 (References)

Frome A, Corrado G S, Shlens J, Bengio S, Dean J, Ranzato M A and Mikolov T. 2013. DeViSE: a deep visual-semantic embedding model// Proceedings of the 26th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: Curran Associates Inc. : 2121-2129

Jiang J J, Ma J Y, Wang Z, Chen C and Liu X M. 2019. Hyperspectral image classification in the presence of noisy labels. IEEE Transactions on Geoscience and Remote Sensing, 57(2): 851-865 [DOI: 10.1109/TGRS.2018.2861992]

Kodirov E, Xiang T and Gong S G. 2017. Semantic autoencoder for zero-shot learning//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 4447-4456 [DOI: 10.1109/CVPR.2017.473]

Lampert C H, Nickisch H and Harmeling S. 2014. Attribute-based classification for zero-shot visual object categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(3): 453-465 [DOI: 10.1109/TPAMI.2013.140]

Li C, Ma Y, Mei X G, Liu C Y and Ma J Y. 2016. Hyperspectral image classification with robust sparse representation. IEEE Geoscience and Remote Sensing Letters, 13(5): 641-645 [DOI: 10.1109/LGRS.2016.2532380]

- Li G D, Zhang C J, Gao F and Zhang X Y. 2019. Doubleconvpool-structured 3D-CNN for hyperspectral remote sensing image classification. *Journal of Image and Graphics*, 24(4): 639-654 (李冠东, 张春菊, 高飞, 张雪英. 2019. 双卷积池化结构的3D-CNN 高光谱遥感影像分类方法. *中国图象图形学报*, 24(4): 639-654) [DOI: 10.11834/jig.180422]
- Li S T, Song W W, Fang L Y, Chen Y S, Ghamisi P and Benediktsson J A. 2019. Deep learning for hyperspectral image classification: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9): 6690-6709 [DOI: 10.1109/TGRS.2019.2907932]
- Mei X G, Pan E T, Ma Y, Dai X B, Huang J, Fan F, Du Q L, Zheng H and Ma J Y. 2019. Spectral-spatial attention networks for hyperspectral image classification. *Remote Sensing*, 11(8): #963 [DOI: 10.3390/rs11080963]
- Mikolov T, Sutskever I, Chen K, Corrado G and Dean J. 2013. Distributed representations of words and phrases and their compositionality [EB/OL]. [2021-02-11]. <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>
- Mou L C, Ghamisi P and Zhu X X. 2017. Deep recurrent neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7): 3639-3655 [DOI: 10.1109/TGRS.2016.2636241]
- Paoletti M E, Haut J M, Plaza J and Plaza A. 2019. Deep learning classifiers for hyperspectral imaging: a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158: 279-317 [DOI: 10.1016/j.isprsjprs.2019.09.006]
- Ran Q, Yu H Y, Gao L R, Li W and Zhang B. 2018. Superpixel and subspace projection-based support vector machines for hyperspectral image classification. *Journal of Image and Graphics*, 23(1): 95-105 (冉琼, 于浩洋, 高连如, 李伟, 张兵. 2018. 结合超像元和子空间投影支持向量机的高光谱图像分类. *中国图象图形学报*, 23(1): 95-105) [DOI: 10.11834/jig.170201]
- Sung F, Yang Y X, Zhang L, Tao X, Torr P H S and Hospedales T M. 2018. Learning to compare: relation network for few-shot learning// *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA: IEEE: 1199-1208 [DOI: 10.1109/CVPR.2018.00131]
- Tong Q X, Zhang B and Zhang L F. 2016. Current progress of hyperspectral remote sensing in China. *Journal of Remote Sensing*, 20(5): 689-707 (童庆禧, 张兵, 张立福. 2016. 中国高光谱遥感的前沿进展. *遥感学报*, 20(5): 689-707) [DOI: 10.11834/jrs.20166264]
- Xian Y Q, Lampert C H, Schiele B and Akata Z. 2019. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(9): 2251-2265 [DOI: 10.1109/TPAMI.2018.2857768]
- Xu Y H, Zhang L P, Du B and Zhang F. 2018. Spectral-spatial unified networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10): 5893-5909 [DOI: 10.1109/TGRS.2018.2827407]
- Zhang B. 2016. Advancement of hyperspectral image processing and information extraction. *Journal of Remote Sensing*, 20(5): 1062-1090 (张兵. 2016. 高光谱图像处理与信息提取前沿. *遥感学报*, 20(5): 1062-1090) [DOI: 10.11834/jrs.20166179]
- Zhang L, Xiang T and Gong S G. 2017. Learning a deep embedding model for zero-shot learning// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 3010-3019 [DOI: 10.1109/CVPR.2017.321]

作者简介



潘尔婷, 1996年生, 女, 博士研究生, 主要研究方向为计算机视觉、遥感图像处理。

E-mail: panerting@whu.edu.cn



黄珺, 通信作者, 男, 副教授, 主要研究方向为红外成像及应用、图像处理、超光谱成像、嵌入式系统。

E-mail: junwong@whu.edu.cn

马泳, 男, 教授, 博士生导师, 主要研究方向为红外热成像、红外超光谱、机器视觉。E-mail: mayong@whu.edu.cn

樊凡, 男, 讲师, 主要研究方向为信号处理、红外图谱成像、红外与遥感信息处理。E-mail: fanfan@whu.edu.cn

李皞, 男, 讲师, 主要研究方向为红外光谱、信号处理、环境遥感。E-mail: lihao@whpu.edu.cn

马佳义, 男, 教授, 博士生导师, 主要研究方向为图像匹配、信息融合、图像超分辨率、红外与遥感图像处理。

E-mail: jyma2010@gmail.com