

高分辨率遥感影像解译中的机器学习范式

周培诚^{1,2}, 程焱^{1,2}, 姚西文^{1,2}, 韩军伟²

1. 西北工业大学 深圳研究院, 深圳 518057;

2. 西北工业大学 自动化学院, 西安 710072

摘要: 高分辨率遥感影像解译是遥感信息处理领域的研究热点之一, 在遥感大数据知识挖掘与智能化分析中起着至关重要的作用, 具有重要的民用和军事应用价值。传统的高分辨率遥感影像解译通常采用人工目视解译方式, 费时费力且精度低。所以, 如何自动、高效地实现高分辨率遥感影像解译是亟待解决的问题。近年来, 随着人工智能技术的飞速发展, 采用机器学习方法实现高分辨率遥感影像解译已成为主流的研究方向。本文结合高分辨率遥感影像解译的典型任务, 如目标检测、场景分类、语义分割、高光谱图像分类等, 系统综述了5种代表性的机器学习范式。具体来说, 本文分别介绍了不同机器学习范式的定义、常用方法以及代表性应用, 包括全监督学习(如支持向量机、K-最近邻、决策树、随机森林、概率图模型)、半监督学习(如纯半监督学习、直推学习、主动学习)、弱监督学习(如多示例学习)、无监督学习(如聚类、主成分分析、稀疏表达)和深度学习(如堆栈自编码器、深度信念网络、卷积神经网络、生成对抗网络)。其次, 深入分析五种机器学习范式的优缺点, 并总结了它们在遥感影像解译中的典型应用。最后, 展望了高分辨率遥感影像解译的机器学习发展方向, 如小样本学习、无监督深度学习、强化学习等。

关键词: 遥感影像解译, 机器学习范式, 深度学习, 弱监督学习, 小样本学习, 强化学习

引用格式: 周培诚, 程焱, 姚西文, 韩军伟. 2021. 高分辨率遥感影像解译中的机器学习范式. 遥感学报, 25(1): 182-197

Zhou P C, Cheng G, Yao X W and Han J W. 2021. Machine learning paradigms in high-resolution remote sensing image interpretation. National Remote Sensing Bulletin, 25(1): 182-197 [DOI: 10.11834/jrs.20210164]

1 引言

随着遥感技术的迅速发展, 遥感影像数据呈现出高空间分辨率、高光谱分辨率、高时间分辨率、大数据的特点(Gong和Zhong, 2016; Tong, 2016)。现有的遥感影像信息处理与分析技术滞后于数据的获取能力, 知识转化和数据利用率低(Li等, 2014)。作为遥感信息处理的前沿研究领域, 遥感影像解译在遥感影像智能化分析中起着至关重要的作用, 广泛应用于自然资源调查(Tong等, 2016)、自然灾害监测(Fan等, 2016)、城市规划(Tong, 2016)、军事侦察(Luo等, 2020)等领域, 具有重要的民用和军事应用价值。

高分辨率遥感影像解译是指根据任务需求, 结合解译标志(如形状、大小、颜色、纹理等)

和相关知识与经验, 通过对高分辨率遥感影像中的地物信息进行建模、分析、推理与判断, 最终获取各类地物信息的过程(Lillesand等, 2015; Mei, 2001; Yang和Liu, 2004)。高分辨率遥感影像解译的典型任务包括遥感图像目标检测(Cheng和Han, 2016; Cheng等, 2016d, 2019; Han等, 2015; Li等, 2018a, 2020a; Xia等, 2018)、场景分类(Cheng等, 2017a, 2018b; Xia等, 2017)、高光谱图像分类(Cheng等, 2018a; Li等, 2019; Zhou等, 2019a)、语义分割(Yao等, 2016b)、变化检测(Mou等, 2019)、图像检索(Wang等, 2016; Xu等, 2020)、图像—文字转换(Lu等, 2018)、SAR图像目标识别(Wang等, 2017; Zhao等, 2016)、红外图像目标识别(Cheng等, 2013b)等。图1为高分辨率遥感影像解译任务示例。

收稿日期: 2020-05-18; 预印本: 2020-08-22

基金项目: 深圳市科技创新委员会基金(编号: JCYJ20180306171131643); 国家重点研发计划(编号: 2017YFB0502900); 国家自然科学基金(编号: 61772425, 61701415); 西北工业大学博士论文创新基金(编号: CX201622)

第一作者简介: 周培诚, 1988年生, 男, 博士研究生, 研究方向为计算机视觉与模式识别。E-mail: zhoupeicheng@mail.nwpu.edu.cn

通信作者简介: 韩军伟, 1977年生, 男, 教授, 研究方向为计算机视觉与模式识别。E-mail: jhan@nwpu.edu.cn

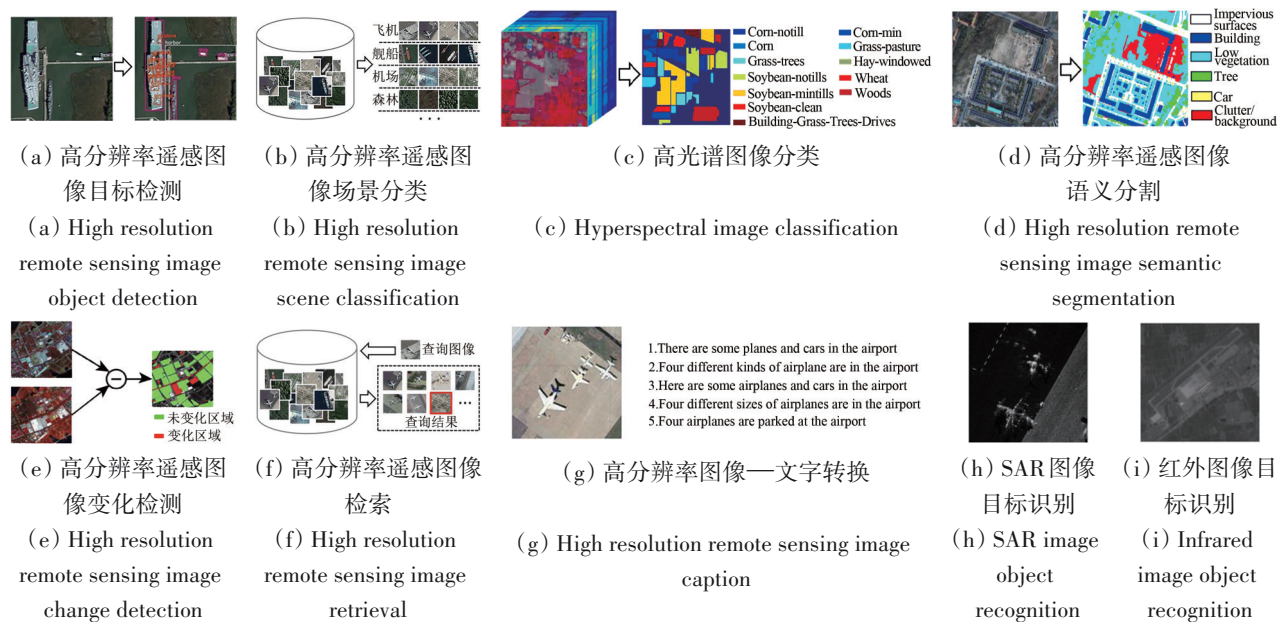


图1 高分辨率遥感影像解译任务示例

Fig. 1 Examples of high resolution remote sensing image interpretation

早期的高分辨率遥感影像解译是由具备专业知识（如地理、气象、农林等）的解译人员结合遥感影像特点和实际状况，通过经验进行人工解译（Lillesand等，2015；Mei，2001），但这种方式耗时、费力且精度不高。因此，如何自动、高效地实现高分辨率遥感影像解译是亟待解决的重要问题。近年来，随着人工智能技术的飞速发展，采用机器学习方法实现高分辨率遥感影像自动、快速、精确的解译已经成为主流的研究方向。

机器学习是对利用经验提高系统自身性能的计算机算法的研究（Zhou，2016）。如图2所示，机器学习的一般流程是通过学习算法，让计算机

从数据中训练模型，并预测未知数据。大体来说，高分辨率遥感影像解译的机器学习范式主要包括全监督学习、半监督学习、弱监督学习、无监督学习和深度学习。图3为5种机器学习范式的示意图。

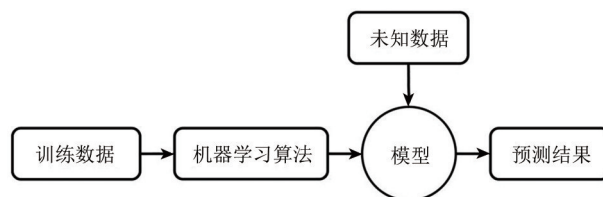


图2 机器学习的一般流程

Fig. 2 General flowchart of machine learning

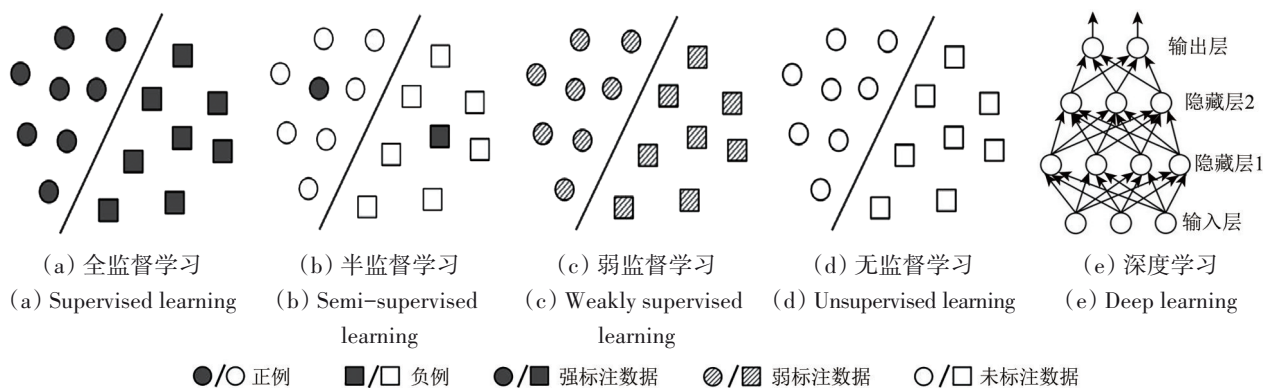


图3 5种机器学习范式

Fig. 3 Five kinds of machine learning paradigms

本文将高分辨率遥感影像解译中的5种机器学习范式进行系统综述（包括全监督学习、半

监督学习、弱监督学习、无监督学习和深度学习）。归纳并分析不同机器学习范式的优缺点，同

时介绍它们在遥感影像解译中的典型应用。最后展望机器学习在高分辨率遥感影像解译中的发展方向。

2 全监督学习

全监督学习是高分辨率遥感影像解译最常用的机器学习方法。全监督学习是指在全部训练数据都给定准确标注信息的情况下,进行特征表示和分类器训练的机器学习方法。图3(a)为全监督学习的示意图。在全监督学习中,良好的特征表示对训练高性能的分类器非常重要,特征表示经历了从手工设计特征到深度学习特征的发展过程。分类器的选择与设计也是非常关键的一个环节。常见的分类器包括支持向量机、K-最近邻、决策树、随机森林、概率图模型等。

2.1 支持向量机

支持向量机 SVM (Support Vector Machine) (Cortes 和 Vapnik, 1995) 的基本思想是在样本空间中寻找具有最大间隔的决策超平面。图4为线性 SVM 的示意图。对于线性不可分的问题,一种有效的解决方法是采用“软间隔”的思想,容许出现一些错分的样本。另外,也可以在 SVM 中引入核函数来解决线性不可分问题,通过将样本原始特征映射到一个更高维度的特征空间,使得样本在该高维特征空间内线性可分。

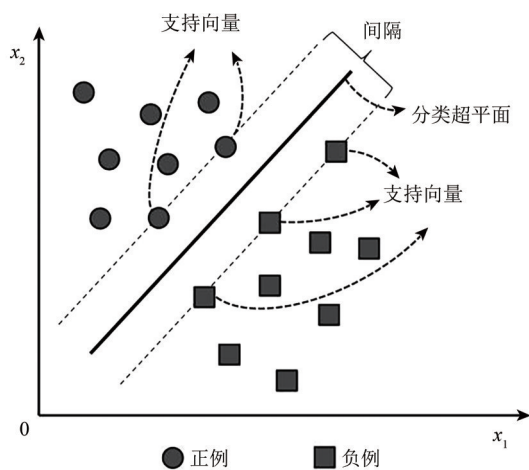


图4 线性支持向量机

Fig. 4 Linear support vector machine

在深度学习之前, SVM 是最流行和有效的分类器之一,在高分辨率遥感影像解译中被广泛应用。例如,在高分辨率遥感图像目标检测中,通

常将目标检测问题转化为分类问题来处理 (Cheng 等, 2013b, 2014)。另外, SVM 也应用于高分辨率遥感图像场景分类 (Cheng 等, 2015a, 2015b) 和高光谱图像分类 (Fauvel 等, 2008)。有关 SVM 在高分辨率遥感影像解译任务中更详细的介绍参见文献 (Cheng 和 Han, 2016; Mountrakis 等, 2011)。

2.2 K-最近邻

K-最近邻 (Cover 和 Hart, 1967) 的基本思想是采用多数投票策略,将与未知样本最近邻的 K 个样本中出现频次最多的类别作为该未知样本的类别。近年来, K-最近邻作为一种分类方法,已经被广泛应用于高分辨率遥感图像目标检测 (Zhu 等, 2005)、高分辨率遥感图像场景分类 (Cheng 等, 2013a) 和高光谱图像分类 (Ma 等, 2010)。

2.3 决策树和随机森林

决策树是通过将数据划分为具有相似值的子集而构建的一个树型结构分类器 (Zhou, 2016)。决策树包含一个根节点、若干非叶子节点和若干叶子节点。根节点是整个样本集,非叶子节点是特征属性,叶子节点对应输出结果。随机森林是由多个互不关联的决策树以随机方式构建的分类器,其分类结果是通过多个决策树的结果投票决定的。在高分辨率遥感影像解译中,可以通过随机森林编码遥感图像目标的不同旋转方向实现高分辨率遥感图像目标检测 (Lei 等, 2012),也可以通过随机森林有效地利用空间上下文信息实现遥感图像语义分割 (Lei 等, 2011)。

2.4 概率图模型

概率图模型是用图来表示变量相关关系的概率模型。该模型可以根据已有信息对感兴趣的未知信息进行估计和推测 (Zhou, 2016)。概率图模型的代表性方法包括条件随机场、马尔可夫随机场等。概率图模型在高分辨率遥感影像解译中得到广泛的应用。例如,通过引入条件随机场实现空间上下文信息建模,用于高分辨率遥感图像目标检测 (Yao 等, 2015; Zhong 和 Wang, 2007); 利用概率图模型获取场景图片之间的隐性关系,用于实现遥感图像场景分类 (Wang 等, 2018); 通过概率图模型建立空间邻域像素之间的关系实现遥感图像语义分割 (Zheng 等, 2017) 和高光谱图像分类 (Li 等, 2012; Zhang 等, 2011) 等。

除了以上4种典型的分类器，朴素贝叶斯 (Tao 和 Nevatia, 2001)、高斯混合模型 (An 和 Aksoy, 2014)、逻辑回归 (Zhang 和 Couloigner, 2006)、线性判别分析 (Bandos 等, 2009; Eikvil 等, 2009) 等方法也被用于高分辨率遥感影像解译。

3 半监督学习

通常情况下，全监督学习需要大量的标注数据，而标注数据数量不足会带来模型训练不充分和性能退化等问题。在高分辨率遥感影像解译中，数据标注需要消耗大量的人力物力，样本获取成本较大，而大量未标注数据却很容易获得。如何利用少量的有标注数据和大量的未标注数据来有效地训练模型，提升模型的预测性能，已经成为高分辨率遥感影像解译的一个研究热点问题。

半监督学习是解决这一问题的一种常见方法。半监督学习是在有标注数据数据量不足的情况下，通过引入大量未标注数据使得模型性能达到与全监督学习接近甚至更好的机器学习方法 (Zhou, 2013, 2016)。图3 (b) 为半监督学习的示意图。半监督学习主要包括纯半监督学习、直推学习和主动学习等3种类型。

3.1 纯半监督学习

纯半监督学习基于“开放世界”假设，假定训练数据中的未标注数据并非待预测的数据，在没有人工干预的情况下，由计算机自动地利用大量未标注数据辅助少量有标注数据进行学习，用于预测未参与训练的测试数据 (Zhou, 2013, 2016)。图5中a是纯半监督学习框架。

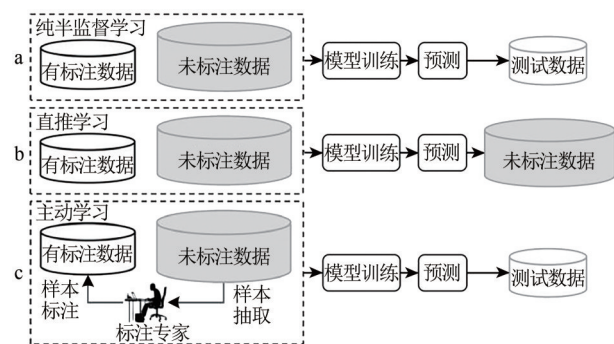


图5 纯半监督学习、直推学习和主动学习框架 (Zhou, 2018)

Fig. 5 Frameworks of pure semi-supervised learning, transductive learning and active learning (Zhou, 2018)

3.2 直推学习

不同于纯半监督学习，直推学习基于“封闭世界”假设，假定训练数据中的未标注数据就是待预测数据，目的是使得学习到的模型在未标注数据中获得最优泛化性能 (Zhou, 2013, 2016)。图5中b是直推学习框架。

3.3 主动学习

主动学习是一种需要专家参与的半监督学习。主动学习首先通过某种查询策略在未标注数据中自动抽取一些样本并由专家进行标注，获得其真实标注信息，然后将这些新的标注样本加入训练集进行模型训练，以提升模型泛化性能 (Zhou, 2013, 2016)。图5中c是主动学习框架。

在遥感影像解译中，半监督学习常用于高光谱图像分类。由于高光谱图像数据维度高且标注样本有限，两者之间的不平衡性使得基于全监督学习的高光谱图像分类比较困难。半监督学习可以利用有限的有标注数据从大量未标注数据中获取新的标注数据。近年来，涌现出大量基于半监督学习的高光谱图像分类方法 (Deng 等, 2019; Dópido 等, 2013; Persello 和 Bruzzone, 2014; Wang 等, 2015)。此外，在标注数据有限的情况下，使用半监督学习方法实现高分辨率遥感图像目标检测 (Bai 等, 2014)；使用半监督学习方法学习高层特征并用于场景分类 (Yang 等, 2015)；使用半监督聚类方法实现SAR图像语义分割 (Yao 等, 2016a)。

4 弱监督学习

在全监督学习和半监督学习中，标注数据都具有强监督信息。而在高分辨率遥感影像解译任务中，人工标注训练数据需要专业知识，数据标注耗时、代价高，且噪声、人为失误等因素会造成标注不准确。为了在弱标注条件下训练性能较好的预测模型，弱监督机器学习应运而生。

弱监督学习是指在监督信息较弱的情况下训练预测模型的机器学习方法。图3 (c) 为弱监督学习的示意图。弱监督信息是指粗粒度的标注信息。例如，在高分辨率遥感图像目标检测任务中，图像级标注信息 (是否含有某类地物目标) 相对于目标级标注信息 (目标边界框) 是弱标注信息；而在高分辨率遥感图像语义分割中，图像级或目

标级标注信息相对于像素级标注信息（目标分割轮廓）均是弱标注信息（Zhou 等，2019b）。图6是不同级别标注信息示例图。弱监督学习可以转化为多示例学习来完成高分辨率遥感影像解译任务。

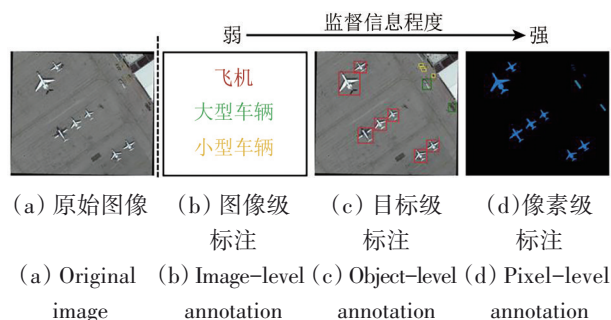


图6 不同级别标注信息示例

Fig. 6 Examples of annotation information of different levels

4.1 多示例学习

多示例学习是以示例包为单元的机器学习方法。多示例学习的训练集由示例包组成，每个示例包有若干未标注的样本。如果示例包中至少有一个样本是正示例，则该示例包是正示例包；如果示例包中所有样本都是负样本，则该示例包是负示例包。多示例学习是通过训练具有包级标注信息的多示例包，来实现对未知示例包的预测。图7是多示例学习框架。

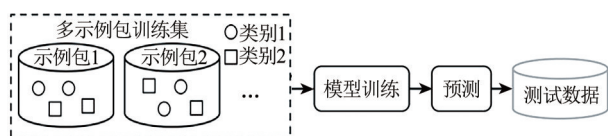


图7 多示例学习框架

Fig. 7 Framework of multiple instance learning

在高分辨率遥感影像解译中，目前已经有一些研究人员提出了基于弱监督学习的遥感图像目标检测方法（Han 等，2015；Zhang 等，2015；Zhou 等，2016），这些方法将弱监督目标检测问题转化为基于多示例学习的图像分类问题。但这些方法只能用于飞机、车辆、机场等单类别目标检测，而无法处理多类别遥感图像目标检测任务。针对这个问题，一种解决方法是采用渐进式示例精炼的方法实现弱监督遥感图像目标检测（Feng 等，2020），主要包括双重上下文示例精炼和候选框渐进自精炼两大模块；另一种解决方法是借鉴

由易到难的学习模式，采用动态课程学习方法实现弱监督遥感图像目标检测（Yao 等，2020）。此外，一些弱监督学习方法采用迁移学习来处理弱监督遥感图像目标检测问题（Li 等，2018b；Zhang 等，2016a）。其基本思路是将大规模遥感图像场景分类数据（如NWPU-RESISC45（Cheng 等，2017a））中的图像或训练的模型作为辅助知识，迁移到遥感图像目标检测数据（如NWPU VHR-10（Cheng 等，2014））中实现遥感图像目标检测。另外，弱监督学习方法也应用于遥感图像场景分类（Li 等，2020b）和语义分割（Yao 等，2016b）。

5 无监督学习

无监督学习是指在样本没有任何标注信息的情况下训练模型的一种机器学习方法。图3（d）为无监督学习的示意图。在高分辨率遥感影像解译中，无监督学习主要作为数据降维、特征选择和特征学习的手段，用于提取遥感影像的特征，来辅助完成高分辨率遥感影像解译任务。常用的无监督学习方法包括聚类、主成分分析、稀疏表达等。

5.1 聚类

聚类是一种应用广泛的无监督学习方法，其基本思想是根据特征将数据划分为若干不相交的子集。聚类既能单独使用，用于寻找并揭示数据的内在性质和规律，也可作为其他数据分析方法的基础（Zhou，2016）。在高分辨率遥感影像解译中，聚类方法可以用于完成高光谱图像聚类（Li 等，2013；Wang 等，2019b）和高光谱图像分类（Jia 和 Richards，2002）任务。

5.2 主成分分析

主成分分析PCA（Principal Component Analysis）是一种无监督数据降维方法，其主要思想是提取数据的主要特征向量（即方差最大的特征向量）来构成线性变换矩阵实现数据降维，使得在数据降维的同时尽可能多的保留数据的主要信息。PCA作为一种有效的特征降维技术，已经被广泛应用于高分辨率遥感影像解译任务。如在高光谱图像分类中，PCA通常用于降低高光谱特征维度（Kang 等，2017）。

5.3 稀疏表达

稀疏表达研究始于 Olshausen (Olshausen 和 Field, 1996) 等于 1996 年提出的神经元稀疏编码学说, 并随着压缩感知理论快速发展。稀疏表达的目的是利用过完备字典中原子的线性组合来表示输入信息, 使得在非零组合系数尽可能少的情况下, 重构信息与原始输入信息的误差尽可能小。稀疏表达作为一种特征表示方法, 在高分辨率遥感图像目标检测 (Han 等, 2014) 和高光谱图像分类 (Yu 等, 2017a, 2020; Zhang 等, 2016b) 任务中都有一定的应用。

除了以上几种无监督学习方法之外, 也有一些基于深度学习的无监督学习方法, 如自编码器、受限玻尔兹曼机和生成对抗网络等, 这些无监督学习方法将在深度学习部分详细介绍。

6 深度学习

随着大规模标注数据、高性能计算能力和先进的机器学习算法的出现, 深度学习在图像分析与理解等众多领域中取得了里程碑式的进展, 成为目前非常流行的一种机器学习方法。深度学习的实质, 是通过构建具有很多隐层的机器学习模型和海量的训练数据, 来学习更加抽象的含有语义信息的高层特征, 从而最终提升分类或预测的准确性。图 3 (e) 为深度学习的示意图。所以“深度模型”是手段, “特征学习”是目的。

典型的深度学习方法包括堆栈自编码器、深度信念网络、卷积神经网络、循环神经网络和生成对抗网络。下面简要介绍几种典型的深度学习方法及其在高分辨率遥感影像解译中的应用情况。

6.1 堆栈自编码器

堆栈自编码器 SAE (Stacked Autoencoder) 是由多个自编码器堆叠构成的深度学习模型。自编码器 (Vincent 等, 2010) 是一种重要的无监督特征学习方法。如图 8 (a) 所示, 它是一种三层神经网络模型, 包含输入层、隐藏层和重构层, 可以分为编码器和解码器两部分。自编码器通过最小化编码器输入数据与解码器重构数据的重构误差来学习隐藏层的特征。SAE 通常采用“预训练+微调”的方式训练。图 8 (b) 是堆栈自编码器的网络结构。

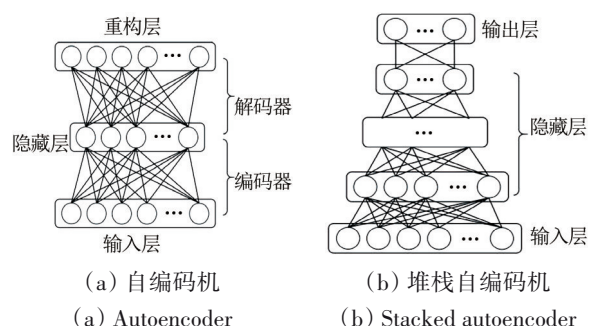


图 8 自编码器和堆栈自编码器

Fig. 8 Autoencoder and stacked autoencoder

在高分辨率遥感影像解译中, 堆栈自编码器通常作为特征学习或特征增强方法, 用以提取更加有效的特征, 提升高分辨率遥感影像解译能力, 已经被广泛用于高分辨率遥感图像场景分类 (Cheng 等, 2015a, 2015b)、高光谱图像分类 (Chen 等, 2014; Zhou 等, 2019a) 和高分辨率遥感图像语义分割 (Yao 等, 2016b)。

6.2 深度信念网络

深度信念网络 DBN (Deep Belief Network) 是由 Hinton 和 Salakhutdinov (2006) 首次提出。如图 9 所示, DBN 是由多个受限玻尔兹曼机 RBM (Restricted Boltzmann Machine) 堆叠构成的, 用于提取深度层级特征, 比单个 RBM 具有更强的特征表达能力。RBM 是一种状态随机生成网络, 包含可视层和隐藏层两层神经元, 且神经元只有激活和未激活两种状态, 可以通过输入数据学习概率分布。DBN 同样采用“预训练+微调”的方式训练。

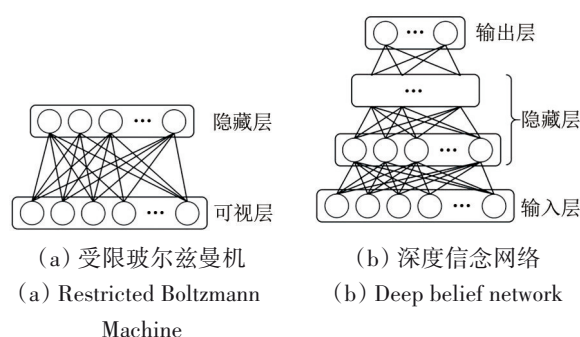


图 9 受限玻尔兹曼机和深度信念网络

Fig. 9 Restricted Boltzmann Machine and deep belief network

类似于堆栈自编码器, 在高分辨率遥感影像解译中, 深度信念网络也常作为特征学习或特征

增强方法, 提高对遥感目标的特征表达能力, 被应用于高分辨率遥感图像目标检测 (Han 等, 2015) 和高光谱图像分类 (Chen 等, 2015; Zhong 等, 2017)。

6.3 卷积神经网络

卷积神经网络 CNN (Convolutional Neural Network) (Krizhevsky 等, 2012) 是受生物学感受野机制启发而提出的一种前馈神经网络。CNN 采用卷积和权值共享训练策略, 极大地降低了模型复杂度, 节省训练开支。此外, CNN 模型的局部连接和空间下采样特点, 使得 CNN 具有一定的平移、尺度和扭曲不变性。图 10 为 CNN 的模型结构。CNN 模型通常由卷积层、池化层和全连接层组合而成。

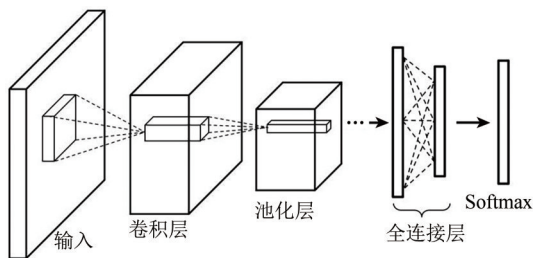


图 10 卷积神经网络模型结构

Fig. 10 Framework of convolutional neural network

(1) 卷积层: 卷积层是 CNN 的核心, 目的是对输入数据进行特征提取。在卷积层中, 使用一组滤波器 (即卷积核), 与前一层中位置邻接区域内的元素进行卷积操作, 并使用非线性激活函数 (如线性整流函数) 进行映射, 获取卷积特征图。

(2) 池化层: 在卷积层进行特征提取后, 采用空间池化方法对卷积特征图进行下采样, 用于特征选择和空间降维。池化层可以有效地降低模型复杂度, 节省计算成本。常用的池化方法包括极大值池化和均值池化。

(3) 全连接层: 全连接层是将当前网络层中的每个节点与上一层的所有节点相连, 可以整合卷积层或者池化层中提取到的特征。

常用的卷积神经网络结构有 AlexNet (Krizhevsky 等, 2012)、VGGNet (Simonyan 和 Zisserman, 2014)、GoogLeNet (Szegedy 等, 2015)、ResNet (He 等, 2016) 等。在高分辨率遥感影像解译中, 往往根据任务特点对常用卷积神经网络模型进行改进, 实现不同的高分辨率遥感影像解译任务。

CNN 在高分遥感图像目标检测任务中应用最为广泛。基于 CNN 的遥感图像目标检测算法, 大都是在计算机视觉领域的两阶段目标检测算法, 如 Faster RCNN (Ren 等, 2017) 或单阶段目标检测算法 (如 YOLO (Redmon 等, 2016) 的基础上改进而成的。高分辨率遥感图像中的目标存在方向任意性 (Cheng 等, 2016c, 2016d, 2019; Li 等, 2018a)、尺度多样性 (Cheng 等, 2020)、几何形变 (Xu 等, 2017)、密集排列 (Cheng 等, 2018c) 等特点。基于 CNN 的高分辨率遥感图像目标检测, 就是结合遥感图像目标的特点, 对深度网络模型进行改进, 使其能够有效地实现目标检测。

在高分辨率遥感图像场景分类中, 最简单的方法是采用“预训练+微调”的方式训练 CNN 模型, 提取特征并使用 SVM 分类 (Cheng 等, 2016a, 2017a)。随着高分辨率遥感图像数据规模、类别多样性和场景复杂度的不断提高, 这种简单的基于 CNN 的分类方法已经不能满足高分辨率遥感图像场景分类的需求。针对这些难点问题, 通过对 CNN 特征的再加工 (Cheng 等, 2017b)、目标函数的改进 (Cheng 等, 2018b, 2019) 和网络结构的改变 (Lu 等, 2019) 来实现高分辨率遥感图像场景分类。

此外, CNN 模型在高光谱图像分类 (Chen 等, 2016; Cheng 等, 2018a; Gao 等, 2020; Xu 等, 2018) 和高分辨率遥感图像语义分割 (Liu 等, 2020; Peng 等, 2019) 中也大量应用。

6.4 循环神经网络

循环神经网络 RNN (Recurrent Neural Network) 是一种用于处理序列数据 (如自然语言、语音等) 的神经网络。RNN 的隐藏单元之间有循环连接, 每个隐藏单元的输出由当前的输入状态和上一层的输出状态决定。

在高光谱图像分类中, 高光谱图像邻域像素之间存在高度相关性, 具有序列数据结构特点, 适合使用 RNN 处理 (Mou 等, 2017)。随着 RNN 的快速发展, 遥感图像—文字转换 (Lu 等, 2018) 受到越来越多的关注。另外, RNN 在其他高分辨率遥感影像解译中也得到了广泛应用, 例如, 将 RNN 与注意力机制结合提取关键区域, 实现高分辨率遥感图像场景分类 (Wang 等, 2019a); 使用 RNN 获取像素的语义上下文信息, 用于高分辨率

遥感图像语义分割 (You 等, 2020) 和 SAR 图像分类 (Geng 等, 2018); 使用 RNN 挖掘变化信息, 用于高分辨率遥感图像变化检测 (Chen 等, 2020)。

6.5 生成对抗网络

生成对抗网络 GAN (Generative Adversarial Network) 是一种自监督深度学习方法。如图 11 所示, GAN 模型包含两个模块: 生成器和判别器。生成器目标是生成尽可能真实的样本以骗过判别器, 而判别器的目标是不断提高判别样本真伪的能力。这两个模块一般都以深度学习模型为基础, 且激活函数必须可微。GAN 学习过程是生成器和判别器相互博弈的过程。

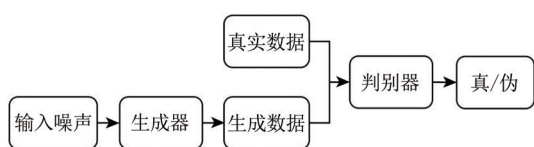


图 11 生成对抗网络结构

Fig. 11 Framework of generative adversarial network

在高分辨率遥感影像解译任务中, GAN 通常作为一种自监督图像生成方法, 用于解决数据不足的问题, 已经被应用于高分辨率遥感图像场景分类 (Ma 等, 2019) 和高光谱图像分类 (Zhu 等, 2018); 另外, GAN 也可以作为一种无监督迁移学习方法用于高分辨率遥感图像语义分割 (Liu 等, 2019)。

在高分辨率遥感影像解译中, 除了以上几种深度学习模型, 也出现了一些其他的神经网络学习方法, 如宽度学习 (Kong 等, 2019)、深度森林 (Boualleg 等, 2019)、极限学习机 (Tang 等, 2015) 等。

7 不同机器学习范式的优缺点及其典型应用

7.1 全监督学习

全监督学习属于强监督学习方法, 只利用有标注数据训练模型。全监督学习中常见的分类器包括支持向量机、K-最近邻、决策树、随机森林、概率图模型等。

优点: 全监督学习可以学习一个从输入变量到输出变量的映射函数, 且输出变量可以是离散变量 (即分类), 也可以是连续变量 (即回归)。

全监督学习方法理论相对完备, 算法简单, 应用范围广, 在数据标注质量高且数据量充足时, 效果较好。

缺点: 全监督学习需要强监督信息, 算法的性能多依赖于训练数据的标注质量, 而高质量的强监督标注数据难以获取或代价高昂, 且标注数据需求量大; 全监督学习很难应用于缺乏训练数据的任务。

典型应用: SVM 是常用的全监督分类方法, 在高分辨率遥感图像场景分类 (Cheng 等, 2015a, 2015b)、高光谱图像分类 (Fauvel 等, 2008; Yu 等, 2017b) 中大量应用, 而高分辨率遥感图像目标检测通常也转化为分类问题处理 (Cheng 等, 2013b, 2014)。SVM 理论相对完备, 可以解决非线性、高维数据分类的问题, 无需依赖整个数据, 泛化能力较高, 也可以用于少量样本学习。缺点是当训练样本很多时, 效率低下, 处理多类分类问题时需要利用二分类 SVM 构造多类分类器。

K-最近邻是最简单的一种全监督学习方法, 其学习过程主要是采用多数投票的策略, 对未知样本的类别进行判断, 无需估计参数, 简单直接, 缺点是算法性能依赖于所选实例, 当训练数据较多时, 计算量较大。在遥感影像解译中, K-最近邻作为常用的分类方法, 应用于高分辨率遥感图像目标检测 (Zhu 等, 2005)、场景分类 (Cheng 等, 2013a) 和高光谱图像分类 (Ma 等, 2010)。

决策树计算简单, 容易理解, 可解释性强, 对中间属性值缺失具有鲁棒性, 也可以处理不相关特征, 但模型复杂, 容易过拟合, 且忽略了数据之间的相关性。随机森林是由多个决策树以随机方式构建的, 采用多数投票预测结果, 相比于决策树, 能够缓解过拟合问题。在遥感影像解译中, 可以使用决策树或随机森林编码遥感地物目标的属性, 如目标的方向 (Lei 等, 2012)、空间上下文信息 (Lei 等, 2011), 并分别用于遥感图像目标检测和语义分割。

概率图模型可以根据变量之间的相关关系, 根据已有信息对未知信息进行估计与推测。在高分辨率遥感影像解译中主要应用于目标检测、场景分类、语义分割、高光谱图像分类。例如, 在目标检测中可以使用概率图模型预测上下文信息 (Yao 等, 2015; Zhong 和 Wang, 2007); 在场景分类中可以获取场景图片之间的隐性关系 (Wang

等, 2018); 在语义分割 (Zheng 等, 2017) 和高光谱图像分类 (Li 等, 2012; Zhang 等, 2011) 中, 主要使用概率图模型建立空间邻域像素的关系。

7.2 半监督学习

半监督学习介于全监督学习和无监督学习之间, 属于标注信息不完备的学习方法。其代表性方法包括纯半监督学习、直推学习、主动学习。

优点: 在标注数据难以获取或特定类别样本量少的情况下, 半监督学习能够有效地挖掘大量无标注数据隐藏的信息, 辅助少量有标注数据训练模型, 改善模型学习性能; 适于处理具有“有标注数据少, 未标注数据多”现象的应用任务; 能共同利用有标注数据和无标注数据学习模式。

缺点: 半监督学习的数据分布非常重要, 主要依据数据分布的连续性与一致性; 当不满足平滑假设、聚类假设、流行假设或模型假设不符合实际情况时, 利用无标注数据可能会使模型性能下降; 在模型训练过程中, 初始训练的模型对未标注数据的标注错误容易累积; 部分方法缺少严格的理论证明; 复杂数据下构建的半监督模型参数较多, 结果不稳定且缺乏参数选取的指导经验; 存在数据不平衡性和噪声问题。

典型应用: 半监督学习可以利用大量无标注数据辅助少量有标注数据训练模型, 在高分辨率遥感影像解译中, 典型的应用是高光谱图像分类 (Deng 等, 2019; Dópido 等, 2013; Persello 和 Bruzzone, 2014; Wang 等, 2015)。此外, 半监督学习也可以应用于目标检测 (Bai 等, 2014)、场景分类 (Yang 等, 2015) 和语义分割 (Yao 等, 2016a) 中。

7.3 弱监督学习

弱监督学习也介于全监督学习和无监督学习之间, 属于标注信息不完备的学习方法。不同于半监督学习方法, 弱监督学习使用粗粒度标注数据学习模型, 学习过程中不包含强监督标注数据。

优点: 标注信息相对强监督学习容易获得, 不依赖于强监督标注信息, 只需要获得相对粗粒度的标注信息, 标注成本较低。

缺点: 弱监督学习的标注信息往往不完整和不确切, 很容易造成模型训练不稳定, 需要精心

地进行初始化操作, 性能与全监督学习相比还存在较大差距。在多类分类问题和多标签学习中, 每个样本都可能会同时赋予多个标签, 分类任务情况会比较复杂。

典型应用: 在高分辨率遥感影像解译中, 弱监督学习通常用于高分辨率遥感图像目标检测。早期的方法主要针对单类别目标, 借鉴多示例学习实现目标检测。针对多类目标检测, 目前的方法是采用渐进式的、由易到难的学习模式, 通过迭代的方法, 逐步进行实例精炼和检测器学习 (Feng 等, 2020; Yao 等, 2020)。还有一些方法采用迁移学习的方法, 将大规模多类数据库中训练的模型迁移到目标检测数据库中, 用于目标检测器的初始化, 实现弱监督目标检测 (Li 等, 2018b; Zhang 等, 2016a)。

7.4 无监督学习

无监督学习在学习过程中不需要特定任务的监督信息, 目标是通过无标注数据学习来揭示数据的内在性质及规律, 为进一步的数据分析提供基础 (Zhou, 2016)。其代表性方法包括聚类、主成分分析、稀疏表达等。

优点: 无监督学习的数据集不需要先验信息, 可以用于探寻数据内在的分布结构及性质, 减少无关和冗余信息, 自主地学习数据的抽象形式, 获得数据的隐特征, 作为遥感影像解译任务的前驱过程。

缺点: 主要用于聚类分析或作为字典学习和特征学习的方法, 用于提取特征; 如果直接用于遥感影像解译, 准确性难以保证, 预测结果存在不确定性; 无监督学习方法属于数据驱动型方法, 利用数据的结构特征进行字典学习或特征学习, 所学习到的字典或特征与具体任务相关性不高。

典型应用: 在高分辨率遥感影像解译中, 利用无标注数据的特征分布, 使用聚类方法将相似样本聚为一类, 可以实现高光谱图像聚类 (Li 等, 2013; Wang 等, 2019b) 和高光谱图像分类 (Jia 和 Richards, 2002), 这种方法是像素级分类方法, 效率不高。另外也有一些基于对象的遥感影像分类方法, 主要是面向对象的图像分析方法 (Blaschke, 2010), 通常先做不同尺度的超像素分割, 然后再对每个对象进行分类。

无监督学习作为字典学习和特征学习的方法, 可以利用数据(不使用标注信息)学习出字典或特征表达, 然后再利用有监督的学习方法训练模型并预测未知数据。如PCA作为一种无监督特征降维方法, 主要用于降低高光谱图像的特征维度, 用于高光谱图像聚类或分类(Kang等, 2017)。稀疏表达和自编码器通过无监督方式学习字典和网络, 实现遥感影像特征提取。

7.5 深度学习

深度学习是近年来流行的一种新的更高级的学习范式, 通过多层表示对数据之间的复杂关系进行建模的算法, 可与全监督学习、半监督学习、弱监督学习和无监督学习方法结合使用。代表性方法包括堆栈自编码器、深度信念网络、卷积神经网络、循环神经网络、生成对抗网络等。

优点: 不需要人工设计特征, 具有很强的特征学习和表达能力, 可以直接从数据中学习更加抽象且含有语义信息的高层特征, 模型预测性能高, 具有较好的迁移学习能力。

缺点: 缺乏严格的理论支撑, 网络模型的可解释性不强; 网络模型大多依赖经验设计; 需要大量有标注的数据; 网络模型复杂度高, 容易出现局部最优解或过拟合问题; 训练耗时, 计算量大; 存储和计算资源要求高, 通常需要GPU平台; 模型的超参数选择多依赖于经验和技巧。

典型应用: 在高分辨率遥感影像解译中, CNN作为最常用的一种深度学习方法, 已经被广泛用于高分辨率遥感图像目标检测、场景分类、高光谱图像分类等任务。在高分辨率遥感图像目标检测中, 通常在计算机视觉的目标检测框架(如Faster R-CNN(Ren等, 2017), YOLO(Redmon等, 2016)等)基础上, 结合遥感图像中目标的特点, 对深度网络进行改进, 来有效地实现目标检测。例如, 针对遥感图像目标方向的不确定性, 可以提取旋转不变特征实现水平框目标检测(Cheng等, 2016b), 也可以设计能够准确识别目标角度的方法来实现有向边框目标检测(Ding等, 2019); 针对目标的尺度差异性, 采用多尺度方法对网络结构进行改进, 实现多尺度目标检测(Cheng等, 2020); 针对遥感目标的密集排列问题, 采用自适应方法, 实现密集目标检测(Cheng等, 2018c)。在高分辨率遥感图像场景分类中,

在原始“预训练+微调”的分类方法基础上, 结合遥感图像特点, 通过对目标函数(Cheng等, 2018b, 2019)和网络结构(Lu等, 2019)进行改进, 实现遥感图像场景分类。在高光谱图像分类中, CNN可以作为特征提取的一种方法用于高光谱图像分类(Cheng等, 2018a), 也可以将CNN结构改进为1D CNN(Hu等, 2015)和3D CNN(Chen等, 2016)提取高光谱图像特征以实现高光谱图像分类。

其他深度学习模型, SAE和DBN通常作为特征学习和特征增强的方法, 用于高分辨率遥感图像目标检测(Han等, 2015)、场景分类(Cheng等, 2015a, 2015b)和高光谱图像分类(Chen等, 2014, 2015; Zhong等, 2017; Zhou等, 2019a)。RNN可用于建立遥感影像中领域像素的相关性, 实现高光谱图像分类(Mou等, 2017)和高分辨率遥感图像语义分割(You等, 2020)。GAN作为图像生成方法用于数据扩增, 实现高分辨率遥感图像场景分类(Ma等, 2019)和高光谱图像分类(Zhu等, 2018)。

8 结 语

本文对高分辨率遥感影像解译中的机器学习范式进行了系统综述, 主要介绍了全监督学习、半监督学习、弱监督学习、无监督学习和深度学习五种代表性的机器学习范式, 分析并总结了不同机器学习范式的优缺点及其在高分辨率遥感影像解译中的典型应用。

随着遥感技术的不断发展, 遥感影像数据呈现出海量、复杂化的特点, 高分辨率遥感影像解译的机器学习方法将趋向于使用少量标注数据学习, 甚至直接从无标注数据中自动学习, 这将给高分辨率遥感影像解译带来极大便利, 推动遥感技术进一步发展。未来的研究方向有如下几种:

(1) 小样本学习算法研究。近年来, 深度学习在遥感影像解译中获得了广泛的使用。但是基于深度学习的方法通常依赖于大量高质量的人工标注数据, 在高分辨率遥感影像解译任务(如场景分类、目标检测、语义分割等)中, 由于数据标注的困难性(如需要大量领域专家进行数据标注)或数据本身的稀缺性, 大量有标注数据难以获取; 另外, 随着遥感数据的不断获取, 新的地物目标也会不断涌现, 这些问题限制了高分辨率

遥感影像解译的效果。如何在少量有标注数据条件下,实现高分辨率遥感影像解译,小样本学习提供了一种可行的解决方案。小样本学习是指在只给定少量训练数据的条件下,训练可以有效地识别目标的机器学习方法。小样本学习不仅可以解决高分辨率遥感影像解译中标注数据不足的问题,也可以对高分辨率遥感影像中的新类别或新概念进行挖掘与识别,具有重要的研究意义。

(2) 无监督深度学习算法研究。深度学习要获得好的性能,通常依赖于大量有标注的训练数据。在高分辨率遥感影像解译中,数据标注耗时费力。为了避免标注大规模遥感数据所耗费的大量成本,可以采用无监督深度学习方法,直接从未标注数据中自主地学习、发现、寻找数据的模式。作为无监督深度学习的一类方法,自监督学习近年来获得很大的关注度。自监督学习可以通过无标注数据本身的结构或特性,自动生成数据标签,然后采用类似监督学习的方法训练模型。这为大规模的高分辨率遥感影像解译提供了一种研究思路,具有重要的研究价值。

(3) 强化学习算法研究。强化学习是智能体(Agent)以“试错”的方式,通过最大化智能体与环境交互过程中的累积奖励,从当前环境状态中获得行动决策的最优策略的方法。在高分辨率遥感影像中,所采集到的数据存在不完整性和不确定性,这给遥感影像解译带来了困难和挑战。为了有效地利用这些数据进行推理与决策,考虑采用强化学习的方法,在只输入原始图像,不进行具体指导的情况下,通过交互学习的方式,实现信息不完备和不确定情况下的高分辨率遥感影像解译。

参考文献(References)

- Ari C and Aksoy S. 2014. Detection of compound structures using a gaussian mixture model with spectral and spatial constraints. *IEEE Transactions on Geoscience and Remote Sensing*, 52(10): 6627-6638 [DOI: 10.1109/TGRS.2014.2299540]
- Bai X, Zhang H G and Zhou J. 2014. VHR object detection based on structural feature extraction and query expansion. *IEEE Transactions on Geoscience and Remote Sensing*, 52(10): 6508-6520 [DOI: 10.1109/TGRS.2013.2296782]
- Bandos T V, Bruzzone L and Camps-Valls G. 2009. Classification of hyperspectral images with regularized linear discriminant analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3): 862-873 [DOI: 10.1109/TGRS.2008.2005729]
- Blaschke T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1): 2-16 [DOI: 10.1016/j.isprsjprs.2009.06.004]
- Boualleg Y, Farah M and Farah I R. 2019. Remote sensing scene classification using convolutional features and deep forest classifier. *IEEE Geoscience and Remote Sensing Letters*, 16(12): 1944-1948 [DOI: 10.1109/LGRS.2019.2911855]
- Chen H, Wu C, Du B, Zhang L P and Wang L. 2020. Change detection in multisource VHR images via deep siamese convolutional multiple-layers recurrent neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 58(4): 2848-2864 [DOI: 10.1109/TGRS.2019.2956756]
- Chen Y S, Jiang H L, Li C Y, Jia X P and Ghamisi P. 2016. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10): 6232-6251 [DOI: 10.1109/TGRS.2016.2584107]
- Chen Y S, Lin Z H, Zhao X, Wang G and Gu Y F. 2014. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6): 2094-2107 [DOI: 10.1109/JSTARS.2014.2329330]
- Chen Y S, Zhao X and Jia X P. 2015. Spectral - spatial classification of hyperspectral data based on deep belief network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6): 2381-2392 [DOI: 10.1109/JSTARS.2015.2388577]
- Cheng G, Guo L, Zhao T Y, Han J W, Li H H and Fang J. 2013a. Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA. *International Journal of Remote Sensing*, 34(1): 45-59 [DOI: 10.1080/01431161.2012.705443]
- Cheng G and Han J W. 2016. A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 117: 11-28 [DOI: 10.1016/j.isprsjprs.2016.03.014]
- Cheng G, Han J W, Guo L, Liu Z B, Bu S H and Ren J C. 2015a. Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8): 4238-4249 [DOI: 10.1109/TGRS.2015.2393857]
- Cheng G, Han J W, Guo L, Qian X L, Zhou P C, Yao X W and Hu X T. 2013b. Object detection in remote sensing imagery using a discriminatively trained mixture model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 85: 32-43 [DOI: 10.1016/j.isprsjprs.2013.08.001]
- Cheng G, Han J W and Lu X Q. 2017a. Remote sensing image scene classification: benchmark and state of the art. *Proceedings of the IEEE*, 105(10): 1865-1883 [DOI: 10.1109/JPROC.2017.2675998]
- Cheng G, Han J W, Zhou P C and Guo L. 2014. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 98: 119-132 [DOI: 10.1016/j.isprsjprs.2014.10.002]

- Cheng G, Han J W, Zhou P C and Xu D. 2019. Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection. *IEEE Transactions on Image Processing*, 28(1): 265-278 [DOI: 10.1109/TIP.2018.2867198]
- Cheng G, Li Z P, Han J W, Yao X W and Guo L. 2018a. Exploring hierarchical convolutional features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(11): 6712-6722 [DOI: 10.1109/TGRS.2018.2841823]
- Cheng G, Li Z P, Yao X W, Guo L and Wei Z L. 2017b. Remote sensing image scene classification using bag of convolutional features. *IEEE Geoscience and Remote Sensing Letters*, 14(10): 1735-1739 [DOI: 10.1109/LGRS.2017.2731997]
- Cheng G, Ma C C, Zhou P C, Yao X W and Han J W. 2016a. Scene classification of high resolution remote sensing images using convolutional neural networks//2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Beijing: IEEE: 767-770 [DOI: 10.1109/IGARSS.2016.7729193]
- Cheng G, Si Y J, Hong H L, Yao X W and Guo L. 2020. Cross-scale feature fusion for object detection in optical remote sensing images. *IEEE Geoscience and Remote Sensing Letters*: 1-5 [DOI: 10.1109/LGRS.2020.2975541]
- Cheng G, Yang C Y, Yao X W, Guo L and Han J W. 2018b. When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5): 2811-2821 [DOI: 10.1109/TGRS.2017.2783902]
- Cheng G, Zhou P C and Han J W. 2016b. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7405-7415 [DOI: 10.1109/TGRS.2016.2601622]
- Cheng G, Zhou P C and Han J W. 2016c. RIFD-CNN: rotation-invariant and fisher discriminative convolutional neural networks for object detection//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 2884-2893 [DOI: 10.1109/CVPR.2016.315]
- Cheng G, Zhou P C and Han J W. 2016d. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7405-7415 [DOI: 10.1109/TGRS.2016.2601622]
- Cheng G, Zhou P C, Han J W, Guo L and Han J G. 2015b. Auto-encoder-based shared mid-level visual dictionary learning for scene classification using very high resolution remote sensing images. *IET Computer Vision*, 9(5): 639-647 [DOI: 10.1049/iet-cvi.2014.0270]
- Cheng L, Liu X, Li L L, Jiao L C and Tang X. 2018c. Deep adaptive proposal network for object detection in optical remote sensing images. *arXiv preprint arXiv: 1807.07327*
- Cortes C and Vapnik V. 1995. Support-vector networks. *Machine Learning*, 20(3): 273-297 [DOI: 10.1007/BF00994018]
- Cover T and Hart P. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21-27 [DOI: 10.1109/TIT.1967.1053964]
- Deng C, Xue Y M, Liu X L, Li C and Tao D C. 2019. Active transfer learning network: a unified deep joint spectral - spatial feature learning model for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(3): 1741-1754 [DOI: 10.1109/TGRS.2018.2868851]
- Ding J, Xue N, Long Y, Xia G S and Lu Q K. 2019. Learning RoI transformer for oriented object detection in aerial images//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE: 2844-2853 [DOI: 10.1109/CVPR.2019.00296]
- Dópido I, Li J, Marpu P R, Plaza A, Dias J M B and Benediktsson J A. 2013. Semisupervised self-learning for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 51(7): 4032-4044 [DOI: 10.1109/TGRS.2012.2228275]
- Eikvil L, Aurdal L and Koren H. 2009. Classification-based vehicle detection in high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1): 65-72 [DOI: 10.1016/j.isprsjprs.2008.09.005]
- Fan Y D, Wu W, Wang W, Liu M and Wen Q. 2016. Research progress of disaster remote sensing in China. *Journal of Remote Sensing*, 20(5): 1170-1184 (范一大, 吴玮, 王薇, 刘明, 温奇. 2016. 中国灾害遥感研究进展. *遥感学报*, 20(5): 1170-1184) [DOI: 10.11834/jrs.20166171]
- Fauvel M, Benediktsson J A, Chanussot J and Sveinsson J R. 2008. Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11): 3804-3814 [DOI: 10.1109/TGRS.2008.922034]
- Feng X X, Han J W, Yao X W and Cheng G. 2020. Progressive contextual instance refinement for weakly supervised object detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*: 1-11 [DOI: 10.1109/TGRS.2020.2985989]
- Gao L R, Gu D X, Zhuang L N, Ren J C, Yang D and Zhang B. 2020. Combining t-distributed stochastic neighbor embedding with convolutional neural networks for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 17(8): 1368-1372 [DOI: 10.1109/LGRS.2019.2945122]
- Geng J, Wang H Y, Fan J C and Ma X R. 2018. SAR image classification via deep recurrent encoding neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4): 2255-2269 [DOI: 10.1109/TGRS.2017.2777868]
- Gong J Y and Zhong Y F. 2016. Survey of intelligent optical remote sensing image processing. *Journal of Remote Sensing*, 20(5): 733-747 (龚健雅, 钟燕飞. 2016. 光学遥感影像智能化处理研究进展. *遥感学报*, 20(5): 733-747) [DOI: 10.11834/jrs.20166205]
- Han J W, Zhang D W, Cheng G, Guo L and Ren J C. 2015. Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. *IEEE Transactions on Geoscience and Remote Sensing*, 53(6): 3325-3337 [DOI: 10.1109/TGRS.2014.2374218]
- Han J W, Zhou P C, Zhang D W, Cheng G, Guo L, Liu Z B, Bu S H

- and Wu J. 2014. Efficient, simultaneous detection of multi-class geospatial targets based on visual saliency modeling and discriminative learning of sparse coding. *ISPRS Journal of Photogrammetry and Remote Sensing*, 89: 37-48 [DOI: 10.1016/j.isprsjprs.2013.12.011]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hinton G E and Salakhutdinov R R. 2006. Reducing the dimensionality of data with neural networks. *Science*, 313(5786): 504 [DOI: 10.1126/science.1127647]
- Hu W, Huang Y Y, Wei L, Zhang F and Li H C. 2015. Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015: 258619 [DOI: 10.1155/2015/258619]
- Jia X P and Richards J A. 2002. Cluster-space representation for hyperspectral data classification. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3): 593-598 [DOI: 10.1109/TGRS.2002.1000319]
- Kang X D, Xiang X L, Li S T and Benediktsson J A. 2017. PCA-based edge-preserving features for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12): 7140-7151 [DOI: 10.1109/TGRS.2017.2743102]
- Kong Y, Cheng Y H, Chen C L P and Wang X S. 2019. Hyperspectral image clustering based on unsupervised broad learning. *IEEE Geoscience and Remote Sensing Letters*, 16(11): 1741-1745 [DOI: 10.1109/LGRS.2019.2907598]
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Red Hook, United states: ACM: 1097-1105
- Lei Z, Fang T, Huo H and Li D R. 2012. Rotation-invariant object detection of remotely sensed images based on texton forest and hough voting. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4): 1206-1217 [DOI: 10.1109/TGRS.2011.2166966]
- Lei Z, Fang T and Li D R. 2011. Land cover classification for remote sensing imagery using conditional texton forest with historical land cover map. *IEEE Geoscience and Remote Sensing Letters*, 8(4): 720-724 [DOI: 10.1109/LGRS.2010.2103045]
- Li D R, Zhang L P and Xia G S. 2014. Automatic analysis and mining of remote sensing big data. *Acta Geodaetica et Cartographica Sinica*, 43(12): 1211-1216 (李德仁, 张良培, 夏桂松. 2014. 遥感大数据自动分析与数据挖掘. *测绘学报*, 43(12): 1211-1216) [DOI: 10.13485/j.cnki.11-2089.2014.0187]
- Li J, Bioucas-Dias J M and Plaza A. 2012. Spectral - spatial hyperspectral image segmentation using subspace multinomial logistic regression and markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3): 809-823 [DOI: 10.1109/TGRS.2011.2162649]
- Li K, Cheng G, Bu S H and You X. 2018a. Rotation-insensitive and context-augmented object detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4): 2337-2348 [DOI: 10.1109/TGRS.2017.2778300]
- Li K, Wan G, Cheng G, Meng L Q and Han J W. 2020a. Object detection in optical remote sensing images: a survey and a new benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159: 296-307 [DOI: 10.1016/j.isprsjprs.2019.11.023]
- Li S S, Zhang B, Li A, Jia X P, Gao L R and Peng M. 2013. Hyperspectral imagery clustering with neighborhood constraints. *IEEE Geoscience and Remote Sensing Letters*, 10(3): 588-592 [DOI: 10.1109/LGRS.2012.2215005]
- 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]
- Li Y S, Zhang Y J, Huang X and Yuille A L. 2018b. Deep networks under scene-level supervision for multi-class geospatial object detection from remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146: 182-196 [DOI: 10.1016/j.isprsjprs.2018.09.014]
- Li Z L, Xu K, Xie J F, Bi Q and Qin K. 2020b. Deep multiple instance convolutional neural networks for learning robust scene representations. *IEEE Transactions on Geoscience and Remote Sensing*, 58(5): 3685-3702 [DOI: 10.1109/TGRS.2019.2960889]
- Lillesand T M, Kiefer R W and Chipman J W. 2015. *Remote Sensing and Image Interpretation*. 7th ed. Hoboken: John Wiley and Sons
- Liu Q H, Kampffmeyer M, Jenssen R and Salberg A B. 2020. Dense dilated convolutions' merging network for land cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 58(9): 6309-6320 [DOI: 10.1109/TGRS.2020.2976658]
- Liu W, Su F L and Huang X. 2019. Unsupervised adversarial domain adaptation network for semantic segmentation. *IEEE Geoscience and Remote Sensing Letters*: 1-5 [DOI: 10.1109/LGRS.2019.2956490]
- Lu X Q, Sun H and Zheng X T. 2019. A feature aggregation convolutional neural network for remote sensing scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10): 7894-7906 [DOI: 10.1109/TGRS.2019.2917161]
- Lu X Q, Wang B Q, Zheng X T and Li X L. 2018. Exploring models and data for remote sensing image caption generation. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4): 2183-2195 [DOI: 10.1109/TGRS.2017.2776321]
- Luo R, Wang L, Xiao Y J, He Y and Zhao D F. 2020. Application of deep learning technology in military field. *Command Control and Simulation*, 42(1): 1-5 (罗荣, 王亮, 肖玉杰, 何翼, 赵东峰. 2020. 深度学习技术在军事领域应用. *指挥控制与仿真*, 42(1): 1-5) [DOI: 10.3969/j.issn.1673-3819.2020.01.001]
- Ma D G, Tang P and Zhao L J. 2019. SiftingGAN: generating and sifting labeled samples to improve the remote sensing image scene classification baseline in vitro. *IEEE Geoscience and Remote Sensing Letters*, 16(7): 1046-1050 [DOI: 10.1109/LGRS.2018.2890413]
- Ma L, Crawford M M and Tian J. 2010. Local manifold learning-based k -nearest-neighbor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48(11): 4099-4109 [DOI: 10.1109/TGRS.2010.2055876]

- Mei A X. 2001. *An Introduction to Remote Sensing*. Beijing: Higher Education Press (梅安新. 2001. 遥感导论. 北京: 高等教育出版社)
- Mou L C, Bruzzone L and Zhu X X. 2019. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 924-935 [DOI: 10.1109/TGRS.2018.2863224]
- 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]
- Mountrakis G, Im J and Ogole C. 2011. Support vector machines in remote sensing: a review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3): 247-259 [DOI: 10.1016/j.isprsjprs.2010.11.001]
- Olshausen B A and Field D J. 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583): 607-609 [DOI: 10.1038/381607a0]
- Peng C, Li Y Y, Jiao L C, Chen Y Q and Shang R H. 2019. Densely based multi-scale and multi-modal fully convolutional networks for high-resolution remote-sensing image semantic segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8): 2612-2626 [DOI: 10.1109/JSTARS.2019.2906387]
- Persello C and Bruzzone L. 2014. Active and semisupervised learning for the classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 52(11): 6937-6956 [DOI: 10.1109/TGRS.2014.2305805]
- Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]
- Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]
- Simonyan K and Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE: 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Tang J X, Deng C W, Huang G B and Zhao B J. 2015. Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3): 1174-1185 [DOI: 10.1109/TGRS.2014.2335751]
- Tao Z and Nevatia R. 2001. Car detection in low resolution aerial image// Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, BC, Canada: IEEE: 710-717 [DOI: 10.1109/ICCV.2001.937593]
- 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]
- Tong X D. 2016. Development of China high-resolution earth observation system. *Journal of Remote Sensing*, 20(5): 775-780 (童旭东. 2016. 中国高分辨率对地观测系统重大专项建设进展. 遥感学报, 20(5): 775-780) [DOI: 10.11834/jrs.20166302]
- Vincent P, Larochelle H, Lajoie I, Bengio Y and Manzagol P A. 2010. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 11(12): 3371-3408
- Wang Q, Liu S T, Chanussot J and Li X L. 2019a. Scene classification with recurrent attention of VHR remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2): 1155-1167 [DOI: 10.1109/TGRS.2018.2864987]
- Wang R, Nie F P, Wang Z, He F and Li X L. 2019b. Scalable graph-based clustering with nonnegative relaxation for large hyperspectral image. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10): 7352-7364 [DOI: 10.1109/TGRS.2019.2913004]
- Wang S G, Wang M, Yang S Y and Jiao L C. 2017. New hierarchical saliency filtering for fast ship detection in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 55(1): 351-362 [DOI: 10.1109/TGRS.2016.2606481]
- Wang Y B, Zhang L Q, Tong X H, Nie F P, Huang H Y and Mei J. 2018. LRAGE: learning latent relationships with adaptive graph embedding for aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2): 621-634 [DOI: 10.1109/TGRS.2017.2752217]
- Wang Y B, Zhang L Q, Tong X H, Zhang L, Zhang Z X, Liu H, Xing X Y and Mathiopoulos P T. 2016. A three-layered graph-based learning approach for remote sensing image retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10): 6020-6034 [DOI: 10.1109/TGRS.2016.2579648]
- Wang Z Y, Nasrabadi N M and Huang T S. 2015. Semisupervised hyperspectral classification using task-driven dictionary learning with laplacian regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 53(3): 1161-1173 [DOI: 10.1109/TGRS.2014.2335177]
- Xia G S, Bai X, Ding J, Zhu Z, Belongie S, Luo J B, Datcu M, Pelillo M and Zhang L P. 2018. DOTA: a large-scale dataset for object detection in aerial images//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, United States: IEEE: 3974-3983 [DOI: 10.1109/CVPR.2018.00418]
- Xia G S, Hu J W, Hu F, Shi B G, Bai X, Zhong Y F, Zhang L P and Lu X Q. 2017. AID: a benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7): 3965-3981 [DOI: 10.1109/TGRS.2017.2685945]
- Xu F, Yang W, Jiang T B, Lin S J, Luo H and Xia G S. 2020. Mental retrieval of remote sensing images via adversarial sketch-image fea-

- ture learning. *IEEE Transactions on Geoscience and Remote Sensing*: 1-14 [DOI: 10.1109/TGRS.2020.2984316]
- Xu X D, Li W, Ran Q, Du Q, Gao L R and Zhang B. 2018. Multi-source remote sensing data classification based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 56(2): 937-949 [DOI: 10.1109/TGRS.2017.2756851]
- Xu Z Z, Xu X, Wang L, Yang R and Pu F L. 2017. Deformable convnet with aspect ratio constrained NMS for object detection in remote sensing imagery. *Remote Sensing*, 9(12): 1312 [DOI: 10.3390/rs9121312]
- Yang G and Liu X N. 2004. The present research condition and development trend of remotely sensed imagery interpretation. *Remote Sensing for Land and Resources*, 16(2): 7-10, 15 (杨枕, 刘湘南. 2004. 遥感影像解译的研究现状和发展趋势. *国土资源遥感*, 16(2): 7-10, 15) [DOI: 10.6046/gtzyyg.2004.02.02]
- Yang W, Yin X S and Xia G S. 2015. Learning high-level features for satellite image classification with limited labeled samples. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8): 4472-4482 [DOI: 10.1109/TGRS.2015.2400449]
- Yao W, Dumitru C O, Loffeld O and Datcu M. 2016a. Semi-supervised hierarchical clustering for semantic SAR image annotation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(5): 1993-2008 [DOI: 10.1109/JSTARS.2016.2537548]
- Yao X W, Feng X X, Han J W, Cheng G and Guo L. 2020. Automatic weakly supervised object detection from high spatial resolution remote sensing images via dynamic curriculum learning. *IEEE Transactions on Geoscience and Remote Sensing* [DOI: 10.1109/TGRS.2020.2991407]
- Yao X W, Han J W, Cheng G, Qian X M and Guo L. 2016b. Semantic annotation of high-resolution satellite images via weakly supervised learning. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6): 3660-3671 [DOI: 10.1109/TGRS.2016.2523563]
- Yao X W, Han J W, Guo L, Bu S H and Liu Z B. 2015. A coarse-to-fine model for airport detection from remote sensing images using target-oriented visual saliency and CRF. *Neurocomputing*, 164: 162-172 [DOI: 10.1016/j.neucom.2015.02.073]
- You H F, Tian S W, Yu L and Lv Y L. 2020. Pixel-level remote sensing image recognition based on bidirectional word vectors. *IEEE Transactions on Geoscience and Remote Sensing*, 58(2): 1281-1293 [DOI: 10.1109/TGRS.2019.2945591]
- Yu H Y, Gao L R, Li W, Du Q and Zhang B. 2017a. Locality sensitive discriminant analysis for group sparse representation-based hyperspectral imagery classification. *IEEE Geoscience and Remote Sensing Letters*, 14(8): 1358-1362 [DOI: 10.1109/LGRS.2017.2712200]
- Yu H Y, Gao L R, Liao W Z, Zhang B, Pižurica A and Philips W. 2017b. Multiscale superpixel-level subspace-based support vector machines for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, 14(11): 2142-2146 [DOI: 10.1109/LGRS.2017.2755061]
- Yu H Y, Shang X D, Zhang X, Gao L R, Song M P and Hu J C. 2020. Hyperspectral image classification based on adjacent constraint representation. *IEEE Geoscience and Remote Sensing Letters*: 1-5 [DOI: 10.1109/LGRS.2020.2982706]
- Zhang B, Li S S, Jia X P, Gao L R and Peng M. 2011. Adaptive markov random field approach for classification of hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 8(5): 973-977 [DOI: 10.1109/LGRS.2011.2145353]
- Zhang D W, Han J W, Cheng G, Liu Z B, Bu S H and Guo L. 2015. Weakly supervised learning for target detection in remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 12(4): 701-705 [DOI: 10.1109/LGRS.2014.2358994]
- Zhang F, Du B, Zhang L P and Xu M Z. 2016a. Weakly supervised learning based on coupled convolutional neural networks for aircraft detection. *IEEE Transactions on Geoscience and Remote Sensing*, 54(9): 5553-5563 [DOI: 10.1109/TGRS.2016.2569141]
- Zhang Q P and Couloigner I. 2006. Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery. *Pattern Recognition Letters*, 27(9): 937-946 [DOI: 10.1016/j.patrec.2005.12.003]
- Zhang X R, Song Q, Gao Z Y, Zheng Y G, Weng P and Jiao L C. 2016b. Spectral-spatial feature learning using cluster-based group sparse coding for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(9): 4142-4159 [DOI: 10.1109/JSTARS.2016.2593907]
- Zhao Z Q, Jiao L C, Liu F, Zhao J Q and Chen P H. 2016. Semisupervised discriminant feature learning for SAR image category via sparse ensemble. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6): 3532-3547 [DOI: 10.1109/TGRS.2016.2519910]
- Zheng C, Zhang Y and Wang L G. 2017. Semantic segmentation of remote sensing imagery using an object-based markov random field model with auxiliary label fields. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5): 3015-3028 [DOI: 10.1109/TGRS.2017.2658731]
- Zhong P, Gong Z Q, Li S T and Schönlieb C B. 2017. Learning to diversify deep belief networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(6): 3516-3530 [DOI: 10.1109/TGRS.2017.2675902]
- Zhong P and Wang R S. 2007. A multiple conditional random fields ensemble model for urban area detection in remote sensing optical images. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12): 3978-3988 [DOI: 10.1109/TGRS.2007.907109]
- Zhou P C, Cheng G, Liu Z B, Bu S H and Hu X T. 2016. Weakly supervised target detection in remote sensing images based on transferred deep features and negative bootstrapping. *Multidimensional Systems and Signal Processing*, 27(4): 925-944 [DOI: 10.1007/s11045-015-0370-3]
- Zhou P C, Han J W, Cheng G and Zhang B C. 2019a. Learning compact and discriminative stacked autoencoder for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7): 4823-4833 [DOI: 10.1109/TGRS.2019.2893180]

- Zhou X L, Chen X J and Chen S Y. 2019b. Weakly supervised learning-based object detection: a survey. *Computer Science*, 46(11): 49-57 (周小龙, 陈小佳, 陈胜勇, 雷帮军. 2019. 弱监督学习下的目标检测算法综述. *计算机科学*, 46(11): 49-57) [DOI: 10.11896/jsjcx.181001899]
- Zhou Z H. 2013. Disagreement-based semi-supervised learning. *Acta Automatica Sinica*, 39(11): 1871-1878 (周志华. 2013. 基于分歧的半监督学习. *自动化学报*, 39(11): 1871-1878) [DOI: 10.3724/SP.J.1004.2013.01871]
- Zhou Z H. 2016. *Machine Learning*. Beijing: Tsinghua University Press, 2016 (周志华. 2016. *机器学习*. 北京: 清华大学出版社)
- Zhou Z H. 2018. A brief introduction to weakly supervised learning. *National Science Review*, 5(1): 44-53 [DOI: 10.1093/nsr/nwx106]
- Zhu C, Shi W, Pesaresi M, Liu L, Chen X and King B. 2005. The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics. *International Journal of Remote Sensing*, 26(24): 5493-5508 [DOI: 10.1080/01431160500300354]
- Zhu L, Chen Y S, Ghamisi P and Benediktsson J A. 2018. Generative adversarial networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9): 5046-5063 [DOI: 10.1109/TGRS.2018.2805286]

Machine learning paradigms in high-resolution remote sensing image interpretation

ZHOU Peicheng^{1,2}, CHENG Gong^{1,2}, YAO Xiwen^{1,2}, HAN Junwei²

1. *Research and Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China;*

2. *School of Automation, Northwestern Polytechnical University, Xi'an 710072, China*

Abstract: High-resolution remote sensing image interpretation is a major topic in remote sensing information processing. It plays a vital role in the knowledge mining and intelligent analysis of remote sensing big data and has important application values in civil and military fields. The traditional methods of high-resolution remote sensing image interpretation generally use manual visual interpretation, which is time consuming and laborious and has low accuracy. Therefore, interpreting high-resolution remote sensing images automatically and efficiently is an urgent problem to be solved. The rapid development of artificial intelligence technology in recent years has made machine learning the mainstream research direction of high-resolution remote sensing image interpretation. In this study, we systematically review five kinds of representative machine learning paradigms on the basis of the typical tasks of high-resolution remote sensing image interpretation, such as object detection, scene classification, semantic segmentation, and hyperspectral image classification. Specifically, we introduce their definitions, typical methods, and applications. The representative machine learning paradigms include supervised learning (e.g., support vector machine, k-nearest neighbor, decision tree, random tree, and probabilistic graph model), semi-supervised learning (e.g., pure semi-supervised learning, transductive learning, and active learning), weakly supervised learning (e.g., multiple instance learning), unsupervised learning (e.g., clustering, principal component analysis, and sparse coding), and deep learning (e.g., stacked auto-encoder, deep belief network, convolutional neural network, and generative adversarial network). Then, we comprehensively analyze the strengths and limitations of the five kinds of machine learning paradigms and summarize their typical applications in remote sensing image interpretation. Finally, we summarize the development direction of high-resolution remote sensing image interpretation, such as few-shot learning, unsupervised deep learning, and reinforcement learning.

Key words: remote sensing image interpretation, machine learning paradigm, deep learning, weakly supervised learning, few-shot learning, reinforcement learning

Supported by Science Technology and Innovation Commission of Shenzhen Municipality (No. JCYJ20180306171131643); National Key Research and Development Program of China (No. 2017YFB0502900); National Natural Science Foundation of China (No. 61772425, 61701415); Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University (No. CX201622)