

# 多模态遥感数据融合中的深度学习技术分析

方正楠, 赵春廷, 刘琦, 丁明董, 郭晋宏  
(青岛理工大学, 山东 266520)

摘要: 阐述深度学习 (DL) 作为一项智能技术, 其在数据表示和重建方面具有特别的能力, 在众多计算机视觉任务中取得显著突破。探讨基于 DL 的多模态遥感 (RS) 数据融合技术的应用。

关键词: 智能技术, 数据融合, 深度学习, 多模态, 遥感。

中图分类号: TP183, TP391.41

文章编号: 1000-0755(2023)07-0049-03

文献引用格式: 方正楠, 赵春廷, 刘琦, 丁明董, 郭晋宏. 多模态遥感数据融合中的深度学习技术分析[J]. 电子技术, 2023, 52(07): 49-51.

## Analysis of Deep Learning in Multimodal Remote Sensing Data Fusion

FANG Zhengnan, ZHAO Chuntin, LIU Qi, DING Mingdong, GUO Jinhong.  
(Qingdao University of Technology, Shangdong 266520, China.)

**Abstract** — This paper describes deep learning (DL) as an intelligent technology with special capabilities in data representation and reconstruction, which has made significant breakthroughs in numerous computer vision tasks. It explores the application of multimodal remote sensing (RS) data fusion technology based on DL.

**Index Terms** — intelligent technology, data fusion, deep learning, multimodality, remote sensing.

### 0 引言

遥感 (RS) 技术飞速发展, 让我们可以获取到相当复杂的异质性地球观测 (EO) 数据, 这为研究人员提供了以全新方式解决当前地球科学应用的机会。近年来, 随着地球观测数据的联合利用, 多模态遥感数据融合研究取得了巨大进展, 但由于缺乏对这些异构数据的综合分析和解释能力。因此, 这种不可忽视的限制进一步引起了对具有强大处理能力的替代工具的强烈需求。

### 1 研究背景

由于在观测地球环境方面的优势, 遥感在各种地球观测任务中发挥着越来越重要的作用<sup>[1]</sup>。随着多模态RS数据的可用性不断提高, 研究人员可以方便地获取适合当前应用的数据。尽管大量的多模态数据变得唾手可得, 但每一种模态都很难捕捉到一个或几个特定的属性, 因此无法完整地描述观察到的场景, 这对后续应用造成了很大的限制。所以说, 多模态数据融合是一种可以解决单模态数据困境的可行方法。

### 2 多模态遥感数据融合方法

#### 2.1 同质融合

同质融合包括泛锐化 (pansharpening、HS pansharpening, HS-MS融合) 和时空融合, 主要致

力于解决光学图像中空间分辨率与光谱分辨率以及空间分辨率与时间分辨率的权衡问题。

#### 2.1.1 全色锐化

全色锐化是指将多光谱 (MS) 和 全色图像 (Pan) 融合以生成高空间分辨率的Pan图像。一般来说, 自编码器 (AE)、卷积神经网络 (CNN) 和生成对抗网络 (GAN) 是基于 DL的全色锐化的常用网络架构。

(1) 监督方法。众所周知, 监督学习是通过已知的结果, 已知数据样本不断学习和训练, 对原始数据进行空间下采样来模拟原始数据。许多将全色锐化视为图像超分辨率问题的方法被相继提出<sup>[2]</sup>。CNN被广泛使用, 可以分为三大类, 即单分支、多分支和混合网络。

第一种方法只是将输入Pan和采样MS或其预处理版本连接成新的分量作为网络的输入。但是简单地将预插值的MS与Pan作为网络的输入, 不仅会忽略单个特征, 还会增加额外的计算负担。因此, 多分支网络应用不同的子网络分别提取模态特定的特征而不是平等地对待两个模态。

(2) 无监督方法。在有监督的方法中可能会出现与尺度相关的问题, 因为它们通常以较低的分辨率进行训练。然而, 无监督方法可以使用原始尺

作者简介: 方正楠, 青岛理工大学; 研究方向: 遥感技术、信息技术应用。

收稿日期: 2023-03-23; 修回日期: 2023-07-12。

度实现训练和测试过程，而不需要模拟参考。因此，无监督方法的关键是通过设计适当的损失函数来精确建立输入数据和融合之间的关系，退化的融合结果应该分别在空间域和光谱域上与输入的Pan和MS相同。

### 2.1.2 HS全色锐化

与全色锐化类似，HS全色锐化旨在将HS中的光谱信息与Pan中的空间信息相结合，以生成具有高空间分辨率的HS图像。

**监督方法。**监督方法旨在学习从输入数据到目标数据的转换，而这些转换在现实世界中是不存在的，因此通常会进行仿真实验。具体而言，通过分别在低空间分辨率HS和低光谱分辨率MS分别通过空间和光谱降解观察到的HS生成。通过这样做，两个模拟图像被视为网络的输入，原始HS作为参考。

**无监督方法。**与全色锐化相比，无监督的HS全色锐化研究很少。一个可能的原因是输入的Pan和MS具有相似的光谱覆盖范围，而Pan和HS在光谱范围上存在很大差异，这导致难以保留空间信息。利用梯度和高频损失对空间关系进行建模，其中首先通过比率估计策略生成初始化图像。

### 2.1.3 HS-MS数据融合

全色锐化相关工作可以看作是HS-MS数据融合的特例，旨在通过融合成对的HS-MS图像来获得具有高空间分辨率的HS产品。因此，许多基于DL的全色锐化方法通过必要的修改解决HS-MS融合问题。在此之后，根据全色锐化中的相同分类介绍典型方法。

**监督方法。**受监督的HS-MS融合遵循与HS全色锐化相同的方案，将输入Pan替换为MS。早期基于DL的HS-MS融合方法提出了经典结构如3-D CNN、残差网络、密集连接的网络和三分量网络等。

**无监督方法。**无监督HS-MS融合方法只需要一对HS-MS图像作为网络的输入，当网络优化完成后即可得到融合后的HS。这些方法大致包括两类，即基于编码解码的方法和基于生成约束的方法。前一类假设目标图像可以用两个矩阵相乘来表示，每个矩阵代表明确的物理意义。AE通常用于对此类过程进行建模。另一类方法旨在通过生成器以初始化图像作为输入直接生成目标图像。

### 2.1.4 时空融合

除了空间-光谱分辨率的权衡之外，还存在时空域的矛盾，即现有卫星平台在同一区域获取的高空间分辨率图像，往往时间间隔较长。反之亦然，这极大地阻碍了变化检测等实际应用。因此，时空融合旨在通过融合一对或多对粗略/精细图像和预测时间的粗略空间分辨率图像来生成具有精细空间分辨率的时间密集产品。

## 2.2 异构融合

与基于像素级融合的旨在生成具有高光谱、空间或时间分辨率的结果的同构融合不同，异构融合主要是指在激光雷达-光学、合成孔径雷达-光学、RS-GBD等方面的集成。这些数据的成像机制完全不同，特征级和决策级融合被广泛采用。

### 2.2.1 激光雷达-光学融合

激光雷达-光学融合可以应用于许多任务，例如配准、全色锐化、目标提取、森林生物量估计<sup>[3]</sup>。HS数据凭借其丰富的光谱信息被广泛应用于分类任务，但在光谱信息不足以区分目标的情况下，性能不可避免地遇到瓶颈。幸运的是，LiDAR系统能够获取三维空间几何图像，这弥补了HS的不足，因此联合利用HS和LiDAR数据识别材料成为近年来的热点。

### 2.2.2 合成孔径雷达-光学数据融合

与光学图像不同，SAR系统旨在采集地面物体的后向散射信号，不仅能反映雷达(RADAR)系统参数信息，还能体现观测场景的物理和几何特征。虽然SAR数据可以为光学图像提供补充知识，但它很容易产生斑点噪声，这可能会严重限制其实际应用潜力。

根据融合层次的不同，可以将合成孔径雷达-光学数据融合分为像素级、特征级和决策级三类。尽管合成孔径雷达与光学数据在成像机制上存在较大差距，但借助SAR图像通过像素级融合合成具有丰富纹理和结构信息的光学产品是可行的。在这种情况下，配准变得极其重要，并提出了基于深度学习的合成孔径雷达(SAR)和光学数据的配准方法，例如孪生CNN和自学习和可迁移网络。在获得一对共配准的SAR光学数据后，将许多传统的全色锐化方法扩展到合成孔径雷达-光学的像素级融合中。在不考虑信息冗余的情况下，直接特征提取策略的效率仍然有限。

因此，设计一种不受稀疏约束的新的BN技术，以减少不必要的特征并使网络更好地泛化。除了上述任务外，合成孔径雷达-光学数据融合还被应用于变化检测<sup>[4]</sup>、生物量估算<sup>[5]</sup>等。

### 2.2.3 RS-GBD融合

GBD包含来自社交媒体、地理信息系统、手机等广泛资源，这极大地有助于了解我们的生活环境。更具体地说，RS表现出从全球视角捕捉大尺度地球表面物理属性的强大能力。另一方面，GBD提供的信息与人类行为高度相关，作为RS的补充，GBD给出了丰富的社会经济描述。

POI指的是可以抽象成一个点的对象，如剧院、公交车站、房屋等。与RS数据不同，每个POI一般包含名称、坐标等一些地理信息，这些信息可以很容易地通过电子地图(如OpenStreetMap)获

取。由于每个POI的属性与功能设施密切相关,因此POI和RS的融合为城市功能区分类任务提供了新的机会。

### 3 结语

不断增长的多模态 RS 数据对 EO 任务既是挑战也是机遇。通过共同利用它们的互补特性,近年来取得了重大突破。特别是,人工智能相关技术由于其在特征提取方面的优越性,已经证明了其优于传统方法的优势。在上述 RS 大数据和前沿技术的推动下,基于 DL 的多模态 RS 数据融合成为遥感界的重要课题。具体来说,我们关注不同融合子领域中的基于 DL 的方法,并在使用的模型、任务和数据类型方面进行了详细研究。最后,令人鼓舞的是,DL 已应用于多模态 RS 数据融合的各个角落,并在近年来取得了巨大而有前途的成就,这为研究人员在未来进行深入研究提供了更多信心。

### 参考文献

- [1] Hong D., He W., Yokoya N., Yao J., Gao L., Zhang L., Chanussot J., Zhu X. Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing[J]. IEEE Geoscience and Remote Sensing Magazine 2021(9):52-87.
- [2] Azarang A., Ghassemian H. A new pansharpening method using multi resolution analysis framework and deep neural networks[C]. 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA), IEEE, 2017:1-6.
- [3] Zhang J., Lin X. Advances in fusion of optical imagery and lidar point cloud applied to photogrammetry and remote sensing[J]. International Journal of Image and Data Fusion, 2017(8):1-31.
- [4] Li X., Du Z., Huang Y., Tan Z. A deep translation (gan) based change detection network for optical and sar remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing 2021, 179:14-34.
- [5] Shao Z., Zhang L., Wang L. Stacked sparse autoencoder modeling using the synergy of airborne lidar and satellite optical and sar data to map forest above-ground biomass[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017(10):5569-5582.
- [6] Ying Sun, Xinchang Zhang, Xiaoyang Zhao, Qinchuan Xin. Extracting Building Boundaries from High Resolution Optical Images and LiDAR Data by Integrating the Convolutional Neural Network and the Active Contour Model[J]. Remote Sensing, 2018.
- [7] Lin Li, Jian Liang, Min Weng, Haihong Zhu. A Multiple-Feature Reuse Network to Extract Buildings from Remote Sensing Imagery[J]. Remote Sensing, 2018.
- [8] Prima Riza Kadavi, Chang-Wook Lee. Land cover classification analysis of volcanic island in Aleutian Arc using an artificial neural network(ANN) and a support vector machine(SVM) from Landsat imagery[J]. Geosciences Journal, 2018.
- [9] Yongyang Xu, Zhong Xie, Yaxing Feng, Zhanlong Chen. Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning[J]. Remote Sensing, 2018.
- [10] Feng Xiao, Baotong Liu, Runa Li. Pedestrian object detection with fusion of visual attention mechanism and semantic computation[J]. Multimedia Tools and Applications, 2019.
- [11] Chen Wang, Xiao Bai, Shuai Wang, Jun Zhou, Peng Ren. Multiscale Visual Attention Networks for Object Detection in VHR Remote Sensing Images[J]. IEEE Geoscience and Remote Sensing Letters, 2019.
- [12] Guoxiong Hu, Zhong Yang, Jiaming Han, Li Huang, Jun Gong, Naixue Xiong. Aircraft detection in remote sensing images based on saliency and convolution neural network[J]. EURASIP Journal on Wireless Communications and Networking, 2018.
- [13] Hui Yang, Penghai Wu, Xuedong Yao, Yanlan Wu, Biao Wang, Yongyang Xu. Building Extraction in Very High Resolution Imagery by Dense-Attention Networks[J]. Remote Sensing, 2018.
- [14] Nicolas Audebert, Bertrand Le Saux. Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018.