

可见光和 SAR 图像特征提取的挑战、方法和机遇

徐礼珺¹, 余晓刚², 何浩辰²

1. 航天工程大学 研究生院, 北京 101400;

2. 北京市遥感信息研究所, 北京 100192

摘要: 可见光图像和合成孔径雷达 SAR (Synthetic Aperture Radar) 图像是卫星遥感在城市规划、防灾减灾与国家安全等领域中的主要数据源。特征提取作为连接遥感图像与高层应用的关键桥梁, 其质量直接影响遥感智能解译在复杂场景下的效能。特征提取方法经历了从手工设计到数据驱动的深刻变革, 尤其在深度学习推动下取得显著突破。然而, 遥感数据标注成本高、专业性强, 往前依赖数据驱动的方法仍面临泛化能力不足、可解释局限等挑战。为此, 已有研究提出“数据+知识”的双驱动理论框架, 通过数据驱动的感知能力和知识驱动的认知能力实现更高级别的智能, 以突破单一数据驱动范式在泛化性和解释性上的瓶颈。为揭示双驱动框架在可见光和 SAR 图像特征提取中的技术逻辑与核心价值, 本文首先系统回顾了传统特征提取方法与基于深度学习的特征提取方法, 归纳这两类方法在特征提取中面临的挑战; 进而深入阐述知识驱动的特征提取方法, 将用于引导网络的知识归纳为视觉知识、地理空间知识和物理知识 3 类, 并探讨知识引导网络的训练范式。最后, 展望特征提取的未来发展趋势, 指出遥感大模型构建、数据与知识双驱动发展和前沿技术交叉融合式发展将成为重点研究方向。

关键词: 遥感智能解译, 可见光图像, SAR 图像, 手工特征提取, 深度特征提取, 数据和知识双驱动

中图分类号: TP75/P2

引用格式: 徐礼珺, 余晓刚, 何浩辰. 2026. 可见光和 SAR 图像特征提取的挑战、方法和机遇. 遥感学报, 30(6): 1564–1595

Xu L J, Yu X G and He H C. 2026. Challenges, methods, and opportunities in feature extraction of visible light and SAR images. National Remote Sensing Bulletin, 30(6): 1564–1595 [DOI: 10.11834/jrs.20265172]

1 引言

早在 20 世纪 60 年代研究人员已开始研究如何利用计算机从遥感图像中提取有用信息, 受限于当时的技术和手段, 主要依靠人工目视判读来获取特征并生成信息。随着数字图像处理技术的发展, “特征提取”作为一个明确的概念被正式提出 (Fukunaga, 1990)。特征提取是指从高维原始数据中提取能够反映其内在本质和规律的低维表示的数学过程, 特征提取的质量, 直接影响了方法在后续任务中的效能, 因此作为底层核心环节, 特征提取是连接遥感图像和应用之间的重要“桥梁”。

在特征提取的早期研究中, 主要采用基于物理先验的手工设计方法, 如尺度不变特征变换 SIFT (Scale-Invariant Feature Transform) (Lowe,

2004)、方向梯度直方图 HOG (Histogram of Oriented Gradients) (Dalal 和 Triggs, 2005) 等经典算法。这类人工特征提取虽然在受限场景下表现出了局部敏感性优势, 但是其本质缺陷体现在 3 个方面: 特征空间的泛化能力受制于设计者的经验边界; 多尺度语义关联的建模缺失; 复杂非线性关系的表征不足, 这些缺陷共同限制了传统遥感图像解译 (即主要依赖人工目视和特征工程的解译范式) 的应用效能。

2012 年, 卷积神经网络 CNN (Convolutional Neural Network) (Krizhevsky 等, 2012) 在 ImageNet 挑战赛中的突破性表现, 推动了以深度神经网络为代表的驱动范式在遥感领域的广泛应用, 并在一系列下游任务中证明了其强大的特征提取与表征能力。例如: 在场景分类任务中, 基于 CNN 的模型具备自主学习复杂地物结构与上下文

收稿日期: 2025-05-20; 预印本: 2026-03-04

第一作者简介: 徐礼珺, 研究方向为遥感目标检测和卫星信息智能处理。E-mail: LiJun_Xu@hgd.edu.cn

通信作者简介: 余晓刚, 研究方向为卫星信息智能处理。E-mail: 13366351343@189.cn

特征的能力，提升了分类精度（Bai等，2022；李智等，2024）；在目标检测任务中，CNN通过端端的学习方式提升了地物目标的检测性能和定位精度（Ma等，2024a；Yang和Mohamed，2025）；在语义分割任务中，全卷积网络等架构实现了像素级分类精度，为精细化分割提供强大支撑（Chen等，2024b；Zheng等，2023b）。上述广泛的应用成果标志着遥感图像处理技术从传统解译向“遥感图像智能解译”（指基于深度学习等人工智能技术，能够从数据中自动学习并提取语义特征的现代解译体系）范式的转换（Zhang等，2016）。深度学习方法通过层级化的非线性变换架构，实现了从像素级特征到高层语义特征的端到端自适应学习，在减少人工干预的同时，显著增强了模型对多尺度特征的协同表达能力。

然而，该类方法对标注数据规模的高度依赖，与遥感数据获取成本高、标注专业性强等现实约束相互制约，限制了遥感图像数据的高效应用，特别是近年来，随着高分辨率对地观测卫星迈入亚米级时代、载荷类型不断拓展、时间分辨率达到小时级。不同于常规图像，遥感图像不仅是像素的集合，也承载了地表地物演化动力学规律等深层信息，还带来了载荷成像机理特性导致的特异性挑战。例如：可见光图像易受大气湍流、光照异变等时变因素干扰；合成孔径雷达SAR（Synthetic Aperture Radar）图像需应对多极化散射机制、入射角效应与相干斑噪声的耦合等影响。

在此背景下，传统方法与深度学习范式的固有局限催生了遥感图像特征提取的“第三范式”——知识驱动的特征提取，即通过嵌入物理机理模型与领域知识图谱，在保持数据驱动优势的同时，提升特征鲁棒性、可解释性与场景适应性。

可见光和SAR图像是遥感图像中发展速度快、应用范围广、具有代表性的图像类型。前者以亚米级分辨率直观反应地物细节，后者实现穿透云雨的全天候观测，是减灾救灾、应急响应、国家安全等领域的主要发展方向，因此，本文以可见光图像和SAR图像的特征提取作为研究对象。

为系统梳理该领域的技术脉络与发展趋势，本文按照下述3个标准进行文献选取：（1）具有奠基性，包括具有开创性的和普遍认可的基础理论和模型文献，如SIFT（Scale-Invariant Feature Transform）、CNN、Transformer等文献；（2）强调

前沿性与相关性，文献选取重点集中于2019年—2025年发表在IEEE Journal、ISPRS Journal、CVPR、ICCV等顶级期刊与会议的高水平研究成果，且所选文献均可直接关联至可见光图像和SAR图像特征提取任务，并根据技术范式覆盖了传统方法、数据驱动的深度学习方法 and 知识驱动的深度学习方法；（3）补充性或展望性文献，包含早期研究成果（2019年之前）以及未来方法探讨的文献。图1详细展示了本文的参考文献类型及数量。

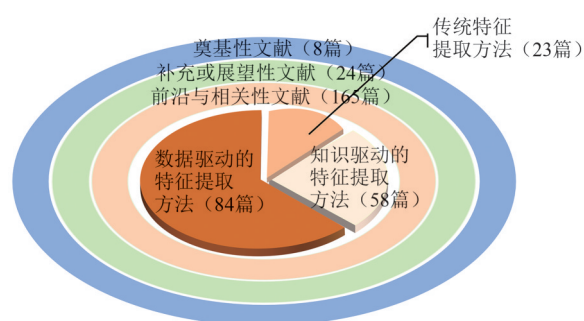


图1 文献选取标准与构成

Fig. 1 Literature selection criteria and composition

2 基于图像处理的特征提取

基于图像的特征提取是指从图像中自动识别并抽取出关键视觉信息（如边缘、纹理、形状等），其目标是获取图像中蕴含的“本质模式”——即能够唯一界定不同地物类别的内在与不变特征，该过程的核心在于摒弃原始图像中存在的冗余与噪声，从而凝练出能够代表地物属性、紧凑且有判别性的数据表示，便于后续分析或计算机学习。根据特征提取的发展历程，可将基于图像处理的特征提取方法归纳为两大类，传统特征提取方法和基于深度学习的特征提取方法。

2.1 传统特征提取方法

传统特征提取方法主要指在深度学习兴起之前，基于图像处理理论和人工设计特征算子的方法体系，其核心思想是通过数学模型或物理模型，从图像中提取能够表征地物视觉属性的低维特征，如边缘、纹理、颜色和结构等。这类方法不依赖大规模标注数据，具有设计透明、可解释性强、计算效率高等特点，但其性能受限于特征设计的经验性与场景适应性。

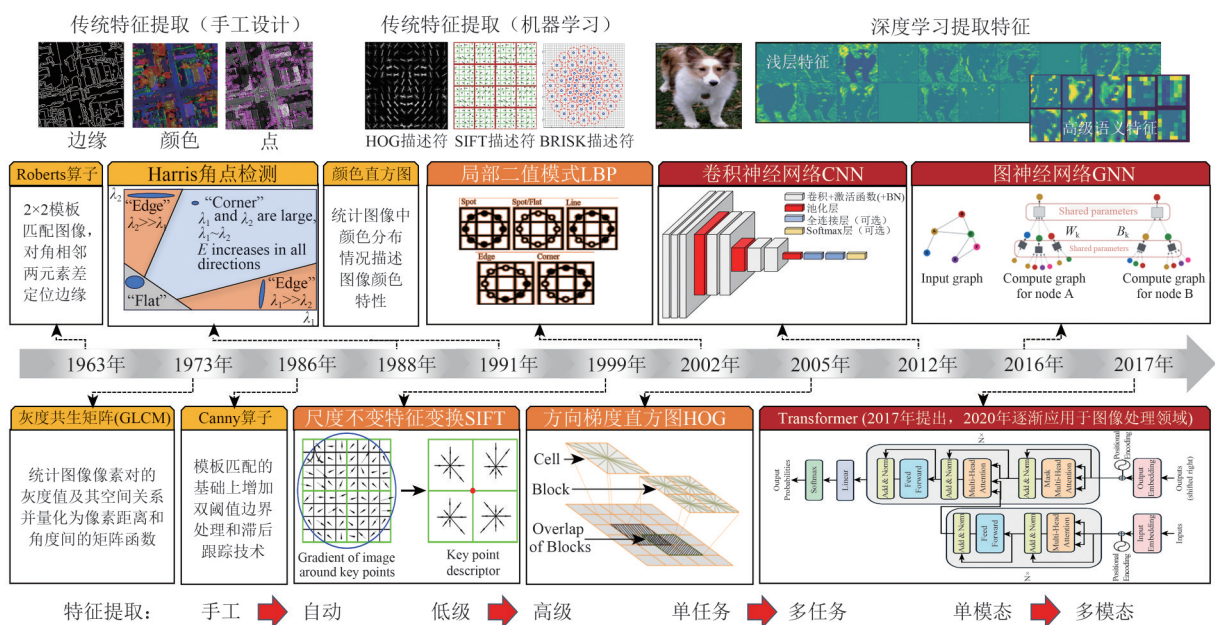


图2 基于图像处理的特征提取技术的发展历程

Fig. 2 Development history of feature extraction technology based on image processing

如图3所示，根据特征类型和提取原理，传统特征提取方法主要分为两条技术路线：基于底层视觉属性建模的路线和基于显著性结构检测的路线。

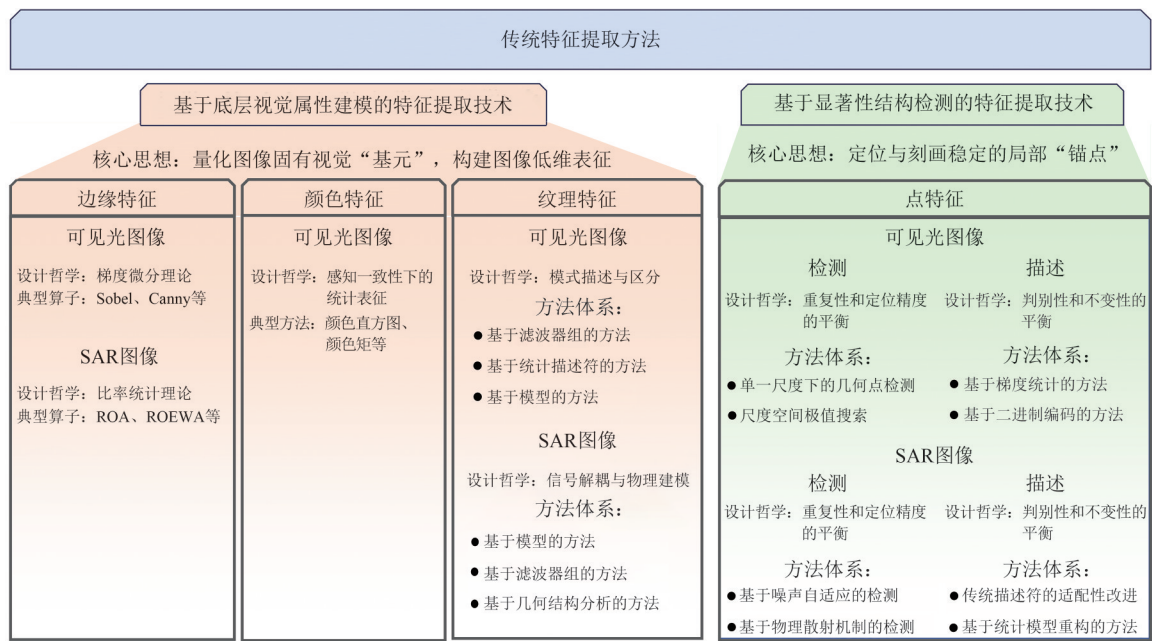


图3 传统特征提取方法的方法论框架

Fig. 3 Methodological framework of traditional feature extraction methods

2.1.1 基于底层视觉属性建模

这类方法遵循“描述—分析”的范式，其核心思想是将图像视为由一系列基础的、可量化的视觉属性构成的集合，通过手工设计特定的数学算子或统计模型，直接对上述属性进行测量和编

码，从而形成对图像内容的低维表征。人类视觉系统在理解场景时往往优先感知场景的底层属性，如物体的轮廓、表面粗糙度、色彩等，受此启发，底层视觉属性建模旨在通过计算手段模拟这一过程，主要围绕边缘（轮廓）、纹理、颜色3个维度

展开，其中边缘定义了“形”、纹理填充了“质”、颜色赋予了“色”，上述3类特征共同构成对图像底层外观的描述。这类方法为后续的语义理解提供了坚实的物理基础，适用于需要全面描述地物外观特征的任务如土地覆盖分类、地物识别等。

(1) 边缘特征。边缘是图像灰度变化显著的区域，蕴含方向、阶跃属性和目标形状等内在信息，是刻画地物轮廓与结构边界的基础视觉线索，在目标检测（Ma等，2024a）、图像分类（Wei等，2019）、立体匹配（Shi和Zhao，2023）等任务中广泛应用。边缘提取的核心是设计对强度变化敏感且能有效抑制噪声的差分或比率算子。

1) 可见光图像边缘提取。可见光图像边缘提取主要基于梯度微分理论实现，通过计算像素邻域内灰度空间变化来定位边缘，经典算子包括一阶微分算子Robert、Sobel、Prewitt，二阶微分算子LoG及以Canny算子为代表的最优边缘检测器。一阶算子计算简单但对方向敏感，刘丽霞等（2019）、Shi和Zhao（2023）通过丰富模板方向提升算子适用性；二阶算子则能检测任意方向但计算量较大；Canny算子在满足低错误率、高定位性和单边缘响应条件下通过双阈值策略提升边缘提取质量，常作为先验知识引导网络学习（Ding等，2019；Pan等，2021a；Zhou等，2024；Song等，2024b），改进方法如Shi和Zhao（2023）通过自适应阈值算法提高效率。

2) SAR图像边缘提取。SAR图像边缘提取算子需首要解决的问题是相干斑噪声的干扰，因此其算子的设计主要基于比率统计原理，通过同质区域比值恒定的特性来抑制乘性噪声。典型算子

包括ROA和ROEWA，前者通过计算相邻区域灰度均值比实现边缘检测，对斑点噪声具备一定鲁棒性（付凯城等，2016），后者引入指数加权函数计算区域均值比，在抑制边缘断裂和提取连续性边缘方面更具优势，但对边缘方向敏感性较弱（Luo等，2020；刘夯等，2017）。SAR图像边缘提取算子多服务于异源图像配准等需要精确几何对齐的任务（王丽娜，2021；谢志华，2022；韩雅鑫，2023）。

表1对上述边缘算子的特点进行总结，各算子可视化对比结果如图4。可见光图像来源于UC Merced数据集，像素分辨率为1英尺，所有算子参数采用原始论文中的默认参数；SAR图像为某合成孔径雷达成像，分辨率为10 m/像素，ROA算子的阈值设置为1.2，ROEWA算子的滑动窗口设置为5、指数衰减参数设置为0.5、阈值设置为1.4。

(2) 纹理特征。纹理特征描述了局部区域像素强度或相位分布的周期性、方向性、粗糙度等信息，在灾害监测（Shao等，2023）、多源图像匹配（Liu等，2021b）、图像分割（赵佳和安道祥，2024）等任务中广泛应用。由于可见光图像和SAR图像成像机理等的差异，对二者的纹理特征提取遵循不同的设计哲学和技术路线。

1) 可见光图像纹理提取。可见光图像的纹理主要源于地表反射率的空间差异，表现为受光照、材质、几何结构等影响的确定性强度分布，因此通过设计能够精确捕获和量化视觉空间规则的方法以实现地物空间模式的描述与区分，从而提取纹理特征，主要可分为基于滤波器组、基于统计描述符和基于模型3大类方法。

表1 典型边缘检测算子的性能对比与场景应用分析
Table 1 Performance comparison and scene application analysis of typical edge detection operators

图像	类型	方法	边缘宽度	边缘定位精度	边缘连续性	处理速度	鲁棒性	适合图像类型
可见光图像	一阶微分	Roberts算子	较细	差	较差,间断点多	较快	对噪声敏感	低噪声陡峭图像
		Prewitt算子	较粗	差	较差,边缘信息丢失	较快	一定抗噪性	强噪声和灰度渐变图像
		Sobel算子	较粗	较差,伪边缘、双边缘严重	较连续	较快	抗噪性较强	强噪声灰度图像
	二阶微分	LoG算子	较细	较高	较连续	较慢	对噪声敏感	强对比度图像
	多级检测	Canny算子	较细	高	连续	慢	抗噪性较强,但对椒盐噪声敏感	低对比度陡峭图像
SAR图像	—	ROA算子	较宽	较差	较差,边缘信息丢失	较慢	一定抗噪性	强噪声或具有丰富阶梯形边缘图像
		ROEWA算子	较细	高	较连续	较快	抗噪性较强	丰富线状目标图像



图4 不同边缘检测算子在可见光图像和SAR图像上的可视化对比

Fig. 4 Visual comparison of different edge detection operators on visible light images and SAR images

基于滤波器组的方法。该类方法认为不同纹理在特定的空间频率或方向上具有能量集中，因此可以通过利用不同的频率、方向等参数提取特征，其优势在于能够通过滤波器参数显式捕捉特定的纹理模式，适用于具有周期性或结构规律的纹理（Dash 和 Senapati, 2020; Cao 等, 2022; Liu 等, 2021b; Liu 和 Yang, 2023）。其中最常用的滤波器为 Gabor 滤波器，因其擅长提取多尺度、多方向的微观纹理结构在图像融合（Zhong 等, 2023）、图像匹配（Nunes 和 Pádua, 2024）等任务中广泛

使用。该类方法提取的特征具有明确的物理意义，特征判别性高，但仅能实现局部统计，且特征质量与滤波器数量和参数密切相关（Asokan, 2023; Wang 等, 2019）。

基于统计描述符的方法。该类方法直接计算局部区域的统计量来表征纹理的空间分布特性，其优势在于计算相对简单，在纹理特征明显且规律的场景中表现出色，常见方法有局部二值模式（Al Saidi 等, 2023）、灰度共生矩阵（Iqbal 等, 2021）等。该类方法能全面描述纹理的统计规律

如周期性、对比度、相关性等, 特征判别性高, 但对图像分辨率敏感, 且计算量与描述符质量密切相关, MLD-CBP (Kas 等, 2020)、旋转不变描述符 (Garg 和 Dhiman, 2021) 等变体被提出以寻找描述符质量与复杂度间的平衡。

基于模型的方法。该类方法假设纹理生成服从某种参数化的随机过程或物理模型, 通过估计模型参数来表征纹理, 其优势在于通过建模纹理生成过程能刻画复杂的纹理结构, 常用的有马尔可夫随机场模型, 该模型通过定义邻域系统和势函数建模纹理的空间相关性, 强调局部随机性与整体规律性的统一, 适用于建模具有空间相关性的平稳纹理 (Yao 等, 2022)。该类方法能捕获具有潜在物理意义的特征表示, 但其有效性依赖于模型假设与纹理实际生成机制间的吻合度, 因此泛化性有限。

2) SAR 图像纹理提取。SAR 图像的纹理主要源于地表粗糙度对雷达波的相干散射干涉, 由于原始 SAR 图像受严重的相干斑噪声干扰, 因此 SAR 图像的纹理特征提取方法首要解决的是如何从随机的斑点噪声中稳健估计出由真实地物结构决定的散射空间, 实现在统计物理模型约束下的信号解耦。主要可分为基于模型、基于滤波和基于结构张量的方法。

基于模型的方法。该类方法在 SAR 图像纹理分析中具有重要地位, 因为其能同时建模纹理的结构性和斑点噪声的随机性, 马尔可夫随机场模型和分形理论模型是常用的模型, 前者通过引入合适的能量函数有效平滑斑点噪声, 同时保持区域边界的清晰度 (Yin 等, 2020; 赵佳和安道祥, 2024), 后者基于自相似性的假设利用分形维数等参数描述纹理的粗糙度或复杂程度, 提取的参数具有明确物理含义, 能直接关联至地物的散射均匀性或表面粗糙度, 实现从数据到物理属性的逆向映射 (Yang 等, 2019)。

基于滤波器组的方法。该类方法原理与可见光图像中基于滤波器组的纹理特征提取方法相同, 但由于 SAR 图像中具有明显斑点噪声的干扰, 因此需要将 SAR 数据中的精确统计分布如 Gamma 分布、K 分布等融入滤波器设计中 (Mahapatra 和 Roy, 2019; Sharma 和 Panigrahi, 2021), 或直接设计基于比率操作的滤波器 (Ma 等, 2024b; Tang 等, 2023), 通过在特征提取前端抑制噪声方差,

提升特征的类内一致性。

基于几何结构分析的方法。该类方法基于“斑点噪声在局部是各向同性的、而真实地物产生的散射变化是具有明确方向性”的先验, 通过分析图像梯度场来估计每个像素点的主方向和各向异性程度, 在此基础上使用扩散滤波实现纹理方向的图像平滑, 从而增强图像的主导几何结构, 擅长提取由线性结构如道路、田地等引起的方向性纹理 (Pham, 2018)。

3) 颜色特征。颜色特征通过量化像素的色彩分布与属性, 能够帮助区分地物目标 (Rabie 等, 2024), 但其对环境光照变化敏感, 且缺乏空间结构信息, 故常作为辅助特征与其他特征结合使用 (Vimina 和 Jacob, 2019; Liu 等, 2021c; Xu 等, 2022)。由于 SAR 图像本质为单通道的强度图像, 不包含色彩信息, 因此仅对可见光图像进行分析。

在可见光图像中, 颜色特征通常在特定的颜色空间中定义和计算, 常见的颜色空间有 RGB 空间、HSV 空间和 Lab 空间, 为排除不同场景内容的干扰从而聚焦颜色空间转换本身的视觉效果, 图 5 展示了同一场景在不同颜色空间下的色彩表示对比, 以直观比较不同颜色空间下颜色特征表示的差异。

RGB 空间以红、绿、蓝三色加权叠加表示特征, 关注目标的真实颜色 (Hu 等, 2024), 但对环境敏感, 不适于亮度频繁变化的场景; HSV 空间将颜色解耦为色调 H、饱和度 S 和明度 V 等 3 个独立分量进行表征, 这种分离机制使该空间对颜色变化高度敏感, 因而能有效区分视觉上相近的颜色 (Anitha 等, 2024); Lab 空间亮度与色调的独立表示与人眼视觉感知最为相符, 在光照变化下仍能高质量表示特征 (Miao 等, 2025), 广泛用于数据比较和图像去雾等场景。

2.1.2 基于显著性结构检测

这类方法遵循“检测—定位—描述”的范式, 其核心思想是将图像视为由一系列离散的、具有特殊几何意义的关键点构成的集合, 这些关键点通常是图像中梯度变化剧烈、曲率显著的点, 具有独特性高、稳定性好、定位性优的特点, 通过设计特定的检测算法来提取关键点, 并为其构建判别性强的数学描述, 从而将高维信息浓缩为具有强表征能力的局部特征向量。这类方法适用于

图像配准、三维重建等需要对图像间几何对应关系进行精确建模的任务。在实现关键点提取时主要围绕两个核心阶段展开：检测阶段和描述阶段，

前者关注关键点的重复性与定位精度，后者则为检测到的关键点生成一个兼具判别力与不变性的数学描述符。

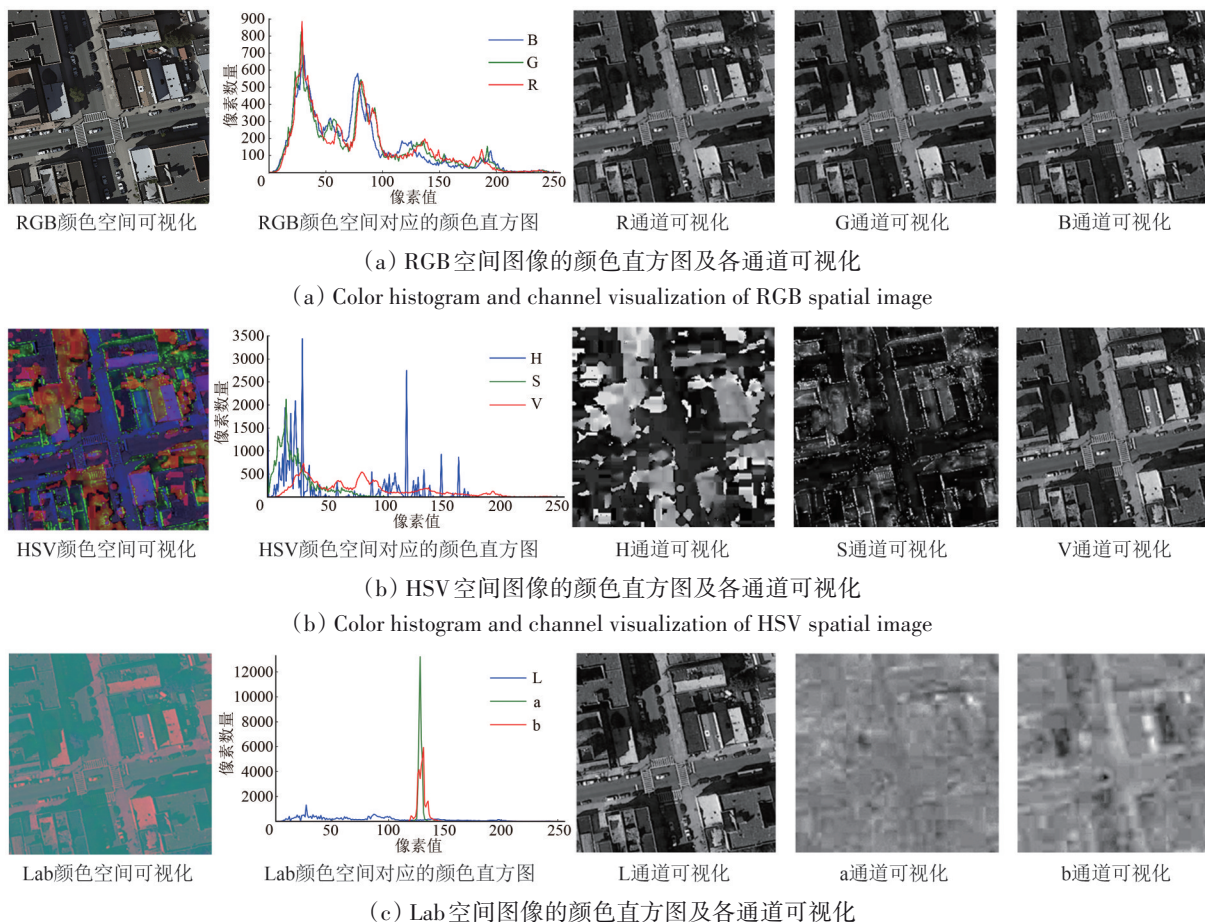


图5 不同颜色空间的色彩表示对比

Fig. 5 Comparison of color representation in different color spaces

(1) 可见光图像关键点提取。可见光图像呈现为规则的强度网格，其边缘清晰、角点明确，相关方法发展较为成熟，在图像融合（王丽娜等，2022）、图像配准（Pang等，2023）等任务中广泛使用。

可见光图像的关键点检测从单一尺度下对具有明确几何意义的点检测逐渐演进为在连续尺度空间中对具有稳定极值响应的显著性结构的检测。单一尺度下的检测以Harris角点检测为代表，通过计算图像中每个点的自相关矩阵来选择在两个正交方向上均存在显著灰度变化的关键点，这类方法提取的关键点定位精度较高，擅长处理有明确几何结构的地物，适用于地形地貌勘测、建筑轮廓识别等场景，但存在聚簇现象，且对尺度和光

照变化敏感，改进方法如空间域-频率域转换通过忽略空间因素影响提取更多均匀分布的关键点（Paul和Pati，2020；孙明超等，2021；王丽娜等，2022）、自适应阈值调整提高空间鲁棒性（刘艳和李一桐，2024）等。基于尺度空间检测的代表方法有SIFT、加速稳健特征SURF（Speeded Up Robust Features）等算法，通过构建高斯尺度空间并在该空间中搜索极值点同时确定关键点的空间位置和特征尺度，这类方法赋予了关键点固有的尺度不变性，提取的关键点对图像缩放、旋转及仿射变化表现出强大的鲁棒性，广泛应用于图像配准任务中（Wu等，2020；Chang等，2019；朱志文等，2011）。

可见光图像关键点描述符经历了从精细统计

邻域梯度分布以追求高判别性和鲁棒性向二进制编码以平衡描述符质量和计算复杂度的范式转变,前者生成的描述符维数较高,具备较强的特征判别力以及尺度不变性、旋转不变性和仿射变换鲁棒性,但计算开销很大,为了实现资源与描述符质量间的平衡,以二进制稳健独立基本特征 BRIEF (Binary Robust Independent Elementary Features)、定向 Fast 和旋转 BRIEF (Oriented FAST and Rotated BRIEF, ORB) 等为代表的二进制描述符向轻量化和实时性方向转变,此外,如精炼生成、优化表示等改进方式被提出 (Mao 和 Peng, 2019; Zhong 和 Li, 2019),以较低的存储需求实现对关键点的特征表示,但是以牺牲描述符鲁棒性为代价。

(2) SAR 图像关键点提取。SAR 图像由地物后向散射信号构成,受斑点噪声和几何畸变影响严重,在关键点提取中面临更严峻的挑战。相较于可见光图像提取角点结构作为关键点, SAR 图像中将稳定的物理散射中心作为关键点,在图像配准等异源任务中发挥重要作用 (Wang 等, 2022b; Pang 等, 2023)。

SAR 图像的关键点检测方法从改进可见光检测器以增强其对斑点噪声的鲁棒性,逐渐发展为直接根据 SAR 图像的物理散射特性来定义和提取关键点。基于改进传统检测器的方法,通过在梯度计算阶段采用噪声抑制策略或阈值响应设置中引入局部统计的自适应机制 (Zhang, 2025; Yu 等, 2019),在抑制斑点噪声引起的虚假响应的同时保持了对图像中真实角点检测的能力,这类方法保留了传统角点检测的几何意义,但其性能仍受限于 SAR 图像中角点结构不清晰问题。基于物理散射特性的检测直接面向 SAR 成像机理,其核心在于将关键点定义为具有特定物理意义的稳定散射中心,这类方法的代表是显著散射点检测,常通过基于统计分布的恒虚警率思想从杂波中分离出强散射点中心,在目标检测与分类等任务中广泛应用 (He 等, 2017; Soldal 等, 2019),但这类方法主要对人造目标或特定的强反射地物敏感,对于缺乏强散射中心的均匀自然区域如森林、水体等检测能力有限。

SAR 图像关键点描述符的发展经历了从改进传统描述符向设计描述符的转变,描述符的改进

涉及尺度空间构建、主方向分配、匹配策略等多个环节 (Dellinger 等, 2015; Xiang 等, 2017),然而,改进后的描述符本质上仍是基于梯度统计的,在强噪声下仍面临严峻挑战。基于 SAR 统计模型的描述符通过估计关键点邻域强度所服从的特定分布参数形成特征向量 (Cordeiro 等, 2019; Rey 等, 2021),这类描述符具有明确的物理意义和内在抗斑点噪声的鲁棒性。

传统方法多采用先检测后描述的流程,通过局部检测逐步优化描述符,保证其能充分反映相应区域重要信息并具有高度独特性,该过程实现简单、可解释性强且无需数据驱动,但随着应用需求的转变和数据时代的来临,存在对图像有选择性、提取局部特征且类型单一、缺少学习能力、效率低等问题。为适应现阶段遥感任务需求,多特征融合和与深度学习结合是主流的两大发展方向,前者通过互补的多类特征或局部与全局特征的融合形成更具判别性的特征表示,解决单类特征表示范围有限、信息不全面等问题,表 2 总结了部分方法及使用场景;后者将传统特征作为先验知识引入网络,以提高网络性能与可解释性,这初步体现了“知识驱动”的思想。

2.2 基于深度学习的特征提取方法

随着遥感数据规模与复杂性的增长,传统特征提取方法在泛化能力、非线性关系建模、多尺度语义关联等方面的固有缺陷日益凸显,在此背景下,基于深度学习的特征提取方法应运而生,深度学习通过数据驱动的方式从海量遥感图像中学习层次化、高判别的特征表示,然而,可见光和 SAR 图像处理面临着独特挑战:图像覆盖范围广、地物尺度变化差异大、成像条件多变等,这对深度特征提取提出了特定要求:(1)对规则/非规则空间结构的建模能力;(2)对多尺度与上下文信息的捕获能力;(3)对全局语义的理解能力。为应对上述挑战,深度学习在遥感领域的应用已从早期简单的网络架构移植发展出针对遥感特性定制的特征提取范式,本节从可见光和 SAR 图像处理的本质需求出发,系统梳理深度学习特征提取技术的发展脉络和演进逻辑,主要围绕局部结构建模和全局关系理解两个维度展开描述。

表2 多特征融合策略及适用场景分析
Table 2 Multi-feature fusion strategy and applicable scenario analysis

融合方式	文献	融合类型	表现	适用场景
多类型特征融合	Tao 等, 2021	边缘、纹理、形状	进行特征融合与降维, 基于 SVM 建模遍历最优参数提高图像分类性能	(1) 特征多样性需求 (2) 跨领域应用 (3) 强噪声场景
	Shen 等, 2023	边缘、纹理、颜色、形状	图像分割中能完整提取地物且大幅降低过分割现象, 提高纹理信息丰富、形状细长地物的分割精度	
	Zhou 等, 2023b	边缘、点	借助边缘提供的同质信息去除伪特征点干扰, 提高 SAR 和光学图像配准速度与精度	
	Qiao 等, 2024	纹理、颜色	复合特征显著提高叶面积指数估计精度, 为精准农业实施提供指导	
	Tang 等, 2024	边缘、点	复合特征有效优化单类特征表示, 提高 SAR 图像配准精度	
全局与局部特征融合	Wu 等, 2021		融合特征上下文信息及特征间的空间信息和互补信息, 解决 SAR 图像斑点噪声造成的图像分类难题	(1) 层次化需求
	Yan 等, 2023		融合图像全局信息熵图和局部复合特征描述符, 实现多模态高相似度图像的配准, 同样适用于非均质图像	(2) 精细定位与识别

2.2.1 面向局部空间结构

这类技术旨在模拟人类视觉系统对局部细节的感知能力, 其核心任务是从规则的像素网格中提取具有判别性的局部特征, 并有效建模地物内部的上下文关系。卷积神经网络凭借其固有的局部连接、权值共享和平移不变性等归纳偏置, 成为局部特征提取的理想架构。

早期研究中将自然图像领域的经典 CNN 架构如视觉几何组 VGG (Visual Geometry Group)、残差网络 ResNet (Residual Network) 等直接迁移至遥感任务, 利用其层次化的卷积-池化操作实现从底层边缘纹理到高层语义特征的自动提取, 然而, 可见光图像和 SAR 图像与自然图像在目标尺度分布、空间结构规律性、背景复杂性等方面存在显著差异, 直接迁移的方法在泛化性和特征判别力上均面临挑战。

为应对上述挑战, CNN 架构在下述 3 个方面进行了针对性改进, 以增强面向图像特性的局部建模能力: (1) 多尺度特征金字塔构建。针对地物尺度变化差异大问题, 以特征金字塔 FPN (Feature Pyramid Network) 为代表的多尺度融合架构被广泛使用 (Lin 等, 2017), 这类方法通过构建自上向下的上采样路径与横向的跳跃连接, 将深层语义信息与浅层细节特征进行融合, 使网络能够同时响应大尺度场景布局与小尺度地物细节 (Sun 等, 2023; Liu 等, 2024e), 解决了单一尺度

感受野与图像多尺度目标间的固有矛盾。(2) 自适应空间上下文建模。为更好捕捉遥感场景中具有规律性的空间结构, 如可变形卷积、空洞卷积等一系列卷积被引入以动态调整卷积核感受野的形状与范围, 从而更精确对齐不规则地物边界 (Hou 等, 2024; Hui 等, 2024; Feng 等, 2022); 此外, 各类通道与空间注意力机制如压缩与激励网络 SENet (Squeeze-and-Excitation Network)、卷积块注意力模块 CBAM (Convolutional Block Attention Module) 等增强了网络自适应加权重要特征通道与空间区域的能力 (Shi 等, 2025b; Xu 等, 2025), 帮助网络在复杂场景下有效聚焦地物目标。(3) 轻量化与效率优化设计。面向星上处理、边缘计算等资源受限场景, 以 MobileNet、ShuffleNet 为代表的轻量级 CNN 通过深度可分离卷积、组卷积、通道混洗等技术几乎不损失精度的情况下大幅降低参数量和计算成本 (Bai 等, 2022; Wang 等, 2024a), 为实时遥感应用提供了可能。

CNN 及其变体在提取局部空间结构特征方面取得了显著成功, 为后续高层语义任务提供了判别性强且具有几何一致性的基础特征表示, 适用于细节丰富、空间结构复杂及需像素级分类的场景, 然而, 其以规则网格为核心的建模范式, 在处理非欧式空间关系和超长程全局依赖关系时存在固有局限, 这催生了新的技术范式的出现。

2.2.2 面向全局语义关系

遥感解译任务不仅需要关注局部邻域信息,更需要理解场景的宏观语义、捕获地物间长程依赖关系以及建模复杂的非规则空间关系,仅依赖CNN范式难以实现上述需求,因此催生了面向全局语义关系的特征提取技术,这类技术旨在突破规则网格的限制,直接建模像素间的全局语义关联和拓扑结构,其发展主要沿基于图神经网络的关系建模方法和基于Transformer的全局上下文建模方法两条路径展开。

(1) 基于图神经网络的关系建模方法。图神经网络GNN(Graph Neural Network)为遥感图像分析提供了一种结构化关系推理的新范式,其将图像抽象为图结构,其中节点表示图像中的实体,边表示实体间的关系,通过消息传递机制,每个节点能聚合其邻域节点的信息并更新自身特征,从而在特征学习中显式编码实体间的结构化关系。这类方法中从图像空间到关系空间映射的图构建质量直接决定了最终的性能,遥感领域的图构建策略主要分为3类:1) 基于几何空间的图构建。基于图像中的空间位置关系进行图结构的构建,如将超像素分割结果中的每个超像素作为节点,若两个超像素共享边界则建立边,从而形成邻接图(Peng等,2022),这类方法保留了地物的空间连续性,适用于语义分割、土地分类等任务;2) 基于特征空间的图构建。在高层特征空间中,根据特征的相似性构建图,如通过对提取的深度特征图中采样一组特征向量作为节点,计算节点间余弦相似度或欧氏距离,对每个节点选择 k 个最相似邻居建立边(Song等,2024a),这类方法能捕捉语义层面的相似性,不受空间距离限制,在异源图像融合任务中广泛使用;3) 基于先验知识的图构建。利用领域信息显式定义关系,如根据地理信息数据中的地点、事件、观察等建立节点,节点间的关系被表示为边(Zhu,2024)。

GNN在遥感领域的应用经历了从基础图卷积网络GCN(Graph Convolutional Network)到复杂关系推理网络的演进:1) GCN的基础应用。GCN通过在GNN中定义图卷积操作,实现节点特征的局部平滑与全局传播,能够有效利用地物间的上下文信息并实现信息间的交互(Tang等,2022; Sebt和Darvishnezhad,2023; Wang等,2024e; Liu等,

2024c);2) 图注意力网络的精细化建模。图注意力网络引入了注意力机制,为不同节点分配不同聚合权重,在处理关系重要性不均衡的遥感场景中具有优势(Zhou等,2023a);3) 动态GNN与时序建模。动态GNN能够建模图中节点和边随时间的变化,支持特征的动态更新,上述特性使其在建模复杂地理数据和处理多时相动态数据方面展现出独特优势(Yu等,2023; Wang等,2022c)。

为适配可见光和SAR数据的特点,研究者对基础GNN进行了多项创新性改进:1) 多尺度GNN。针对图像地物尺度差异大问题,构建了多尺度GNN,通过在不同分辨率层级上构建图并基于图池化和图上采样操作等实现多尺度信息的融合,实现对局部细节和全局信息的捕获(Song等,2023; Wang等,2025b)。2) 异构GNN。图像数据中通常包含多种类型的节点和关系,通过为不同类型的节点(如建筑的材质高度、道路的等级状态等)和边(如“连接”、“属于”等)设计不同的特征变换和聚合函数,能更好建模复杂场景(Zhu,2024)。3) 自构建GNN。区别于基础GNN自主学习节点和边信息,自构建GNN通过手动构建节点和边提取特征,常用于特定场景或数据结构,泛化性有限,适合小规模数据集(Zi等,2021; Liu等,2021a)。该方法需人工介入,常用于需深度挖掘实体联系或数据异质场景。

GNN的优势在于其关系建模的显式性、灵活性和可解释性,能够处理非欧式数据、建模复杂空间拓扑与语义关系,然而,其面临的挑战同样显著:一是图结构的敏感性,GNN性能高度依赖初始图的质量,而构建高质量最优图结构本身是一个难题;二是计算复杂度,GNN的计算复杂度与节点数量密切相关,限制了其在大范围图像中的应用。

(2) 基于Transformer的全局上下文建模方法。Transformer架构,特别是其核心组件自注意力机制,极大驱动了图像处理技术的发展,与CNN局部归纳偏置不同,自注意力机制允许模型直接计算图像中任意两个位置间的关联强度,从而实现对全局上下文信息的捕获与长程依赖关系的显式建模,这一特性使Transformer在处理大范围场景理解、多尺度目标交互和复杂背景下的目标聚焦等问题时具备优势(Zhang等,2024a,2024c; Dang等,2024)。

Transformer在视觉领域的成功应用始于ViT (Vision Transformer) (Dosovitskiy等, 2021), 其将图像分割为固定大小的非重叠图像块, 通过线性投影转换为序列嵌入, 并加入位置编码后输入编码器, 在图像分割、目标检测等需要精细像素级预测或复杂空间推理等任务中展现优势, 典型方法有SETR (Zheng等, 2021)、DETR (Carion等, 2020)等。Transformer在可见光和SAR图像任务中的演进围绕以下方面: 1) 特定的序列化策略。标准ViT的均匀分块策略在自然图像检测中适用, 但考虑到遥感图像中目标尺度差异大问题, 研究者提出多种改进策略, 如多尺度分块策略或重叠分块策略, 生成不同尺度大小的分块或具有重叠的分块, 从而保证小尺寸块的细节信息和大尺寸块的全局语义信息得到利用, 并有效缓解边界信息损失 (Liu等, 2024b; Shi等, 2024)。2) 地理感知的位置编码。可见光和SAR图像具有明确的地理参考信息, 传统ViT中可学习或正弦式的位置编码仅能提供序列内的相对顺序, 而忽略了绝对地理空间关系, 因此, 研究者设计了如地理坐标嵌入、相对位置编码等方式 (Yu等, 2024c; Gao等, 2025), 将地物的真实空间关系引入网络, 帮助网络理解几何畸变、空间邻近性原则等。

为适配可见光和SAR数据的特点, 研究者对基础Transformer进行了多种创新: 1) 高效注意力变体。标准自注意力机制的计算复杂度与序列长度的平方成正比, 这对高分辨率图像构成了严峻挑战, 为此, 如局部窗口注意力、可变形注意力、稀疏注意力等一系列高效注意力变体 (Wei等, 2023; Zhu等, 2023; Dai等, 2024; Lu等, 2023; Pan等, 2023, 2024b; Zhang等, 2024e) 和蒸馏、剪枝、对比学习等方式 (Fan等, 2023a; Zhang等, 2024c) 被提出, 实现了性能与资源间的平衡。2) 混合架构探索。Transformer的序列化特征提取方式使其在底层局部特征建模上存在固有不足, 为弥补上述缺点, 研究者将具有局部卷积优势的CNN架构与Transformer相结合, 具体有两种实现路径: 一是在Transformer编码器前或内部嵌入卷积网络, 利用卷积高效提取局部细节与结构化特征, 作为自注意力机制进行全局关系推理的基础 (Liu等, 2024f; Fan等, 2023b; Yuan和Xu, 2021; Zhao等, 2024); 二是设计并行双分支结构, 分别处理局部和全局信息, 再通过跨分

支注意力或特征融合模块实现交互 (Yan等, 2024; Lu等, 2024)。

Transformer以其强大的全局建模能力在大范围场景理解、长程关系建模等方面具有优势, 然而, 其仍面临下述挑战: 一是高度依赖数据集, Transformer通常需要大量数据才能充分训练, 而高质量标注数据集有限, 限制了其在小众领域的发展; 二是计算与内存开销, 尽管提出了各种注意力变体, 在处理大尺寸遥感图像时仍具备挑战; 三是归纳偏置的合理引入, 基础Transformer缺乏平移不变性、局部性等对视觉任务有益的归纳偏置, 如何将其以不损害网络表达能力的方式融入架构设计是关键, 也是挑战。

图6以图像分类任务为例展现深度模型处理遥感图像的通用流程, 骨干网络负责从输入图像中提取深度特征表示, 经特征聚合后由分类器输出预测结果, 以先验知识组成的监督信息在特征表示和模型训练两方面发挥作用, 在特征表示层面, 先验知识作为高层语义引导, 与深度特征进行聚合以塑造更具判别性的特征表示; 在模型训练层面, 先验知识被编码为具体的约束项与传统数据损失项组成损失函数实现对目标的共同优化, 迫使模型学习的特征符合先验知识遵循的领域规律。需要说明的是, 图6是以CNN架构为骨干网络展开的, Transformer因其内部结构间的差异不适用于图6展示的通用流程, 但基于Transformer的骨干网络同样也能承担特征提取的职能, 且“编码—提取”的宏观流程与图6是相似的。基于深度学习的方法利用神经网络重新定义和优化了传统特征提取任务, 例如, 在边缘特征提取中, 借助端到端网络融合多源异构信息以学习多维度的边缘信息, 并通过模块实现边缘信息的多尺度融合, 克服了传统算子对阈值敏感、易受噪声干扰的缺点 (Shi等, 2025a; Guo等, 2022); 在点特征提取中, 深度学习网络实现了从几何显著性到高层语义一致性的跨越, 具备提取定位精度更佳、语义内涵更丰富的点特征的能力 (Cao等, 2024; Xiao等, 2025b)。上述研究结果表明, 数据驱动的深度学习方法有效应对传统手工特征面临的3大缺陷: 首先, 模型从海量数据中端到端学习特征表示, 特征质量不再依赖于设计者的经验边界; 其次, 借助层次化的网络结构与注意力机制, 模型能够自主捕获与融合多尺度的上下文信息, 根本上解决

多尺度语义关联建模缺失问题；最后，深度模型强大的非线性拟合能力使其能够表征手工特征难以

以描述的复杂、高阶的数据模式，克服了传统方法在复杂非线性关系表征上的不足。

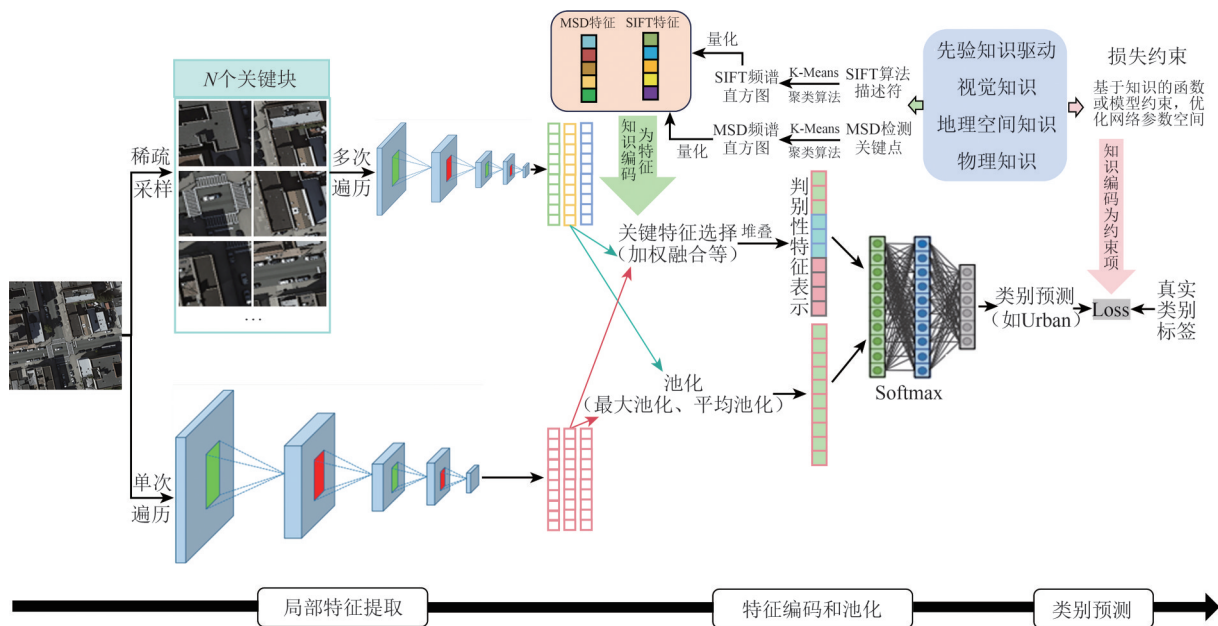


图6 基于深度学习的遥感图像模型处理的通用流程

Fig. 6 General process of remote sensing image model processing based on deep learning

尽管如此，实际使用过程中网络存在过度依赖数据集、网络退化、可解释性差等问题。针对数据集依赖，早期通过数据增强方式解决，目前更多采用小样本、跨域迁移、对比学习等学习方式，开启弱监督或无监督预训练时代；针对网络退化，提出一系列特征提取器（Pang等，2024；Sahu和Panda，2024；Yu等，2024a）、特征融合

模块（Lin等，2024；Shin等，2024）和注意力模块（Ai等，2024；Long等，2023），助力网络聚焦关键区域，生成高质量特征图，表3展示各模块在提升性能和应对挑战方面的作用；针对可解释性差，知识驱动的特征提取方法是解决方法之一，通过编码知识引导网络关注有意义特征，提高网络可解释性的同时降低对数据集的依赖。

表3 面向特征提取挑战的深度学习模块改进梳理

Table 3 Improvement and sorting of deep learning modules for feature extraction challenges

图像	文献	模块			解决的挑战										数据集	性能/%
		特征提取器	特征融合模块	注意力模块	遮挡	低对比度	语义相似	多尺度	多方向	密集分布	小目标	小样本	散射点关系	相干噪声		
可见光图像	Gao等, 2024		√	√		√	√	√			√				LoveDa	mIoU: 52.55
	Li和Ma, 2024	√	√	√				√	√	√	√				DOTA	mAP: 77.40
	Li和Qu, 2025		√	√				√		√	√				DIOR	mAP: 84.40
	Li等, 2024b	√	√	√				√		√	√				DIOR	mAP: 79.56
	Li等, 2024e		√	√				√		√	√				VisDrone	mAP: 48.80
	Pang等, 2024	√						√			√	√			WHU	IoU: 83.59
	Wan等, 2024			√			√	√							AID	OA: 93.53
	Wang等, 2024d	√	√	√			√	√			√				Pots-dam	mIoU: 87.20
	Wei等, 2024b		√	√			√		√						LEVIR-CD+	F1: 87.14
	Yu等, 2024a	√			√			√			√				DOTA	AP: 91.90
	Liu等, 2023	√	√					√			√	√			GF-2	F1: 90.79
	Wang等, 2023b	√	√				√	√	√		√				ACD	mAP: 49.40

续表																
图 像	文献	模块			解决的挑战										数据集	性能/%
		特征 提取器	特征融合 模块	注意力 模块	遮挡	低对 比度	语义 相似	多 尺度	多 方向	密集 分布	小 目标	小 样本	散射点 关系	相干 噪声		
S A R 图 像	Liu 等, 2024d	√	√				√				√			√	ALOS-2	mIoU:97.60
	Muzeau 等, 2024	√										√	√		Sentinel-1	mIoU:43.87
	Pan 等, 2024a	√	√				√	√	√	√			√	√	LSSDD	F1:80.89
	Shen 等, 2024		√	√			√	√	√	√			√	√	SSDD	AP:63.90
	Wang 等, 2025a	√	√	√		√	√				√			√	SV-1	mIoU:91.20
	Yasir 等, 2024	√		√			√	√	√	√			√	√	SSDD	F1:91.12
	Zhang 和 Wu, 2024	√	√	√			√		√	√				√	SSDD	mAP:97.30
	Zhao 等, 2023a	√	√	√			√	√		√			√	√	SSDD+	F1:94.30
	Ai 等, 2024		√	√				√				√	√		MSTAR	F1:95.80
	Cao 等, 2023		√	√			√	√		√				√	SSDD	F1:97.20

注：模块部分标记“√”表示文献对该部分模块进行改进；空白表示未进行改进，不是文献中不包含该部分模块。

3 基于图像的特征提取挑战

尽管前述特征提取方法在各自适用场景下取得了显著成效，但其大多局限于对图像本身进行处理，无论是传统方法还是基于深度学习的方法，特征提取操作都是针对图像表观数据进行处理。与自然图像相比，遥感图像因远距离成像而覆盖范围广，单幅图像包含多种地物目标，其形状、尺度、散射特性等各不相同。此外，可见光图像和 SAR 图像在成像机理、预处理方式、数据内容与表现形式等方面存在本质差异，进一步扩大了特征提取难度。遥感图像独特的成像机理与复杂的应用场景为基于图像处理的方法带来了内在挑战。本节分别总结可见光图像和 SAR 图像在特征提取方面所面临的具体挑战。

3.1 可见光图像面临的挑战

3.1.1 语义模糊性

可见光图像覆盖范围广，图像中包含复杂背景及多类目标，目标形状、尺度、方向等差异为特征提取带来难度，地物多样性使“同物异谱”和“同谱异物”的现象较为普遍，这要求特征提取具备语义层面的理解能力，能从有限样本中学习到的地物的内在与不变特征。背景的复杂性也增加了特征提取的难度，导致低对比度目标或伪装目标易融入背景中，使得其特征与背景噪声难以区分。目标自身属性带来的多样性使同类目标在外观上可能呈现较大差异，难以用单一的特征提取方法准确提取所有目标特征，要求特征提取具

备一定的灵活性和泛化性。

3.1.2 尺度极端化与成像随机性

遥感可见光图像同时存在大尺度场景的宏观格局与小尺度目标的微观细节，导致单一尺度的特征提取方法难以应对，要求特征提取架构是尺度自适应的，能充分关注不同尺度目标的重要信息；此外，尺度多样性导致高分辨率空间细节的保留与高层语义信息的获取间存在固有矛盾，要求特征提取架构能够协同优化局部细节与全局语义，生成兼具丰富细节和准确语义的特征图。同时，目标在图像中的成像具有随机性，表现在几何位置（如遮挡、分布密度）、亮度、方向等，上述因素会影响目标特征表示，要求特征提取具备鲁棒性。

3.1.3 成像条件敏感,特征稳定性退化

可见光图像是被动成像，受太阳高度、大气条件、传感器参数等因素共同作用，导致同一地物在不同成像时间下的特征可能发生剧烈变化，如光照条件中，强光可能导致地物过曝，丢失细节信息，弱光则会降低目标与背景的区分度；成像时间上，不同时间的光照角度会影响目标的阴影分布，不同季节下目标的外观、形状等可能发生巨大改变。上述因素都使目标特征产生非线性变化，直接影响了特征提取的稳定性，也带来了域偏移问题，制约了模型在跨季节变化检测、多源数据融合等应用中的实用性，需要特征提取具备“域自适应”的能力。同时，恶劣成像环境会降低图像质量，模糊目标特征，从而增加目标信息获取的难度。

3.2 SAR图像面临的挑战

3.2.1 散射特性建模困难

SAR作为微波相干成像系统,其图像本质上是地物散射点的干涉叠加结果,因此特征提取的核心难点在于建模散射点间的空间组合与结构关系,这要求特征提取模型具备强大的空间关系推理能力。此外,图像中的相干噪声可能掩盖目标细节信息或混淆特征,降低成像质量,要求特征提取算法能区分或抑制相干噪声。同时,SAR图像的侧视成像会引起几何失真现象,造成目标散射点空间结构的改变,复杂地形会加剧几何失真的干扰,为特征提取带来挑战。

3.2.2 特征度量的高度不稳定性

同一目标散射特性并不是固定的,气候、成像角度、成像时间、目标姿态等因素均会影响目标的散射特性,对于运动目标,散射点还会因多普勒效应产生模糊或位移,加剧提取难度;此外,不同尺度目标散射特征差异大,大尺度目标

散射特性明显,易于提取,小尺度目标则受相干噪声干扰严重,其散射点易与噪声相混淆,在特征提取时需要精确区分。

3.2.3 强杂波背景下特征分离困难

SAR图像中目标与背景杂波在散射机制上存在高度耦合和混淆,导致提取的特征缺乏判别性;此外,目标信号在成像中存在被背景杂波淹没的可能,导致特征空间中类间边界模糊,甚至减弱或吞没目标信号,最终使得特征的语义纯度显著降低。上述挑战要求特征提取架构具备噪声不变性学习能力以及低信噪比下特征筛选能力。

为直观展示上述挑战,图7和图8分别给出可见光图像和SAR图像特征提取中的典型挑战示例,其中,图7(a)和(b)体现了语义模糊性挑战,图7(c)和(d)体现了尺度极端化与成像随机性挑战,图7(e)和(f)体现了成像条件敏感挑战。图8(a)—(c)体现了散射特性建模挑战,图8(d)和(e)体现了特征度量不稳定挑战,图8(f)体现了强杂波背景下特征分离困难挑战。

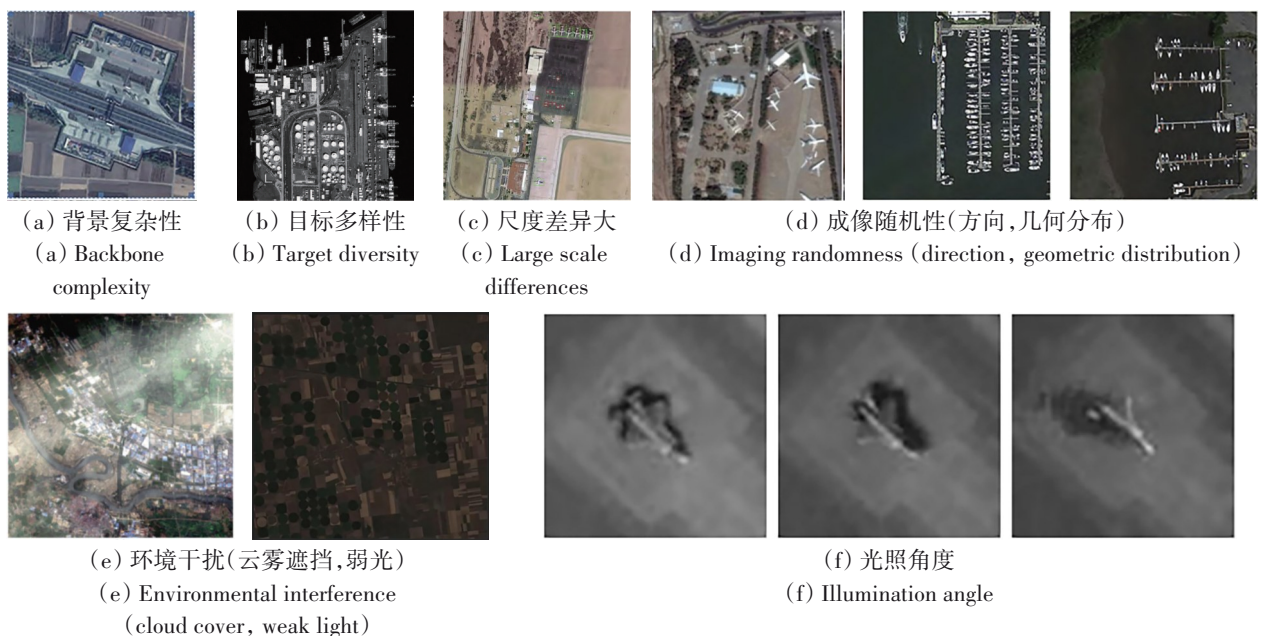
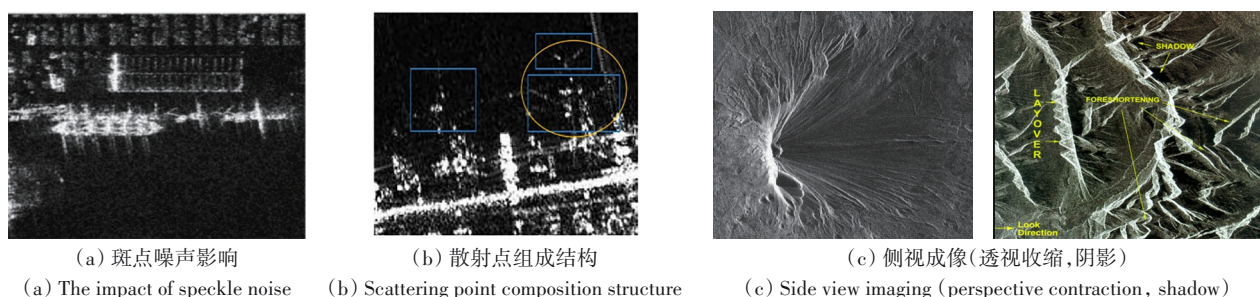


图7 可见光图像特征提取的典型挑战示例

Fig. 7 Typical challenge examples for feature extraction of visible light images



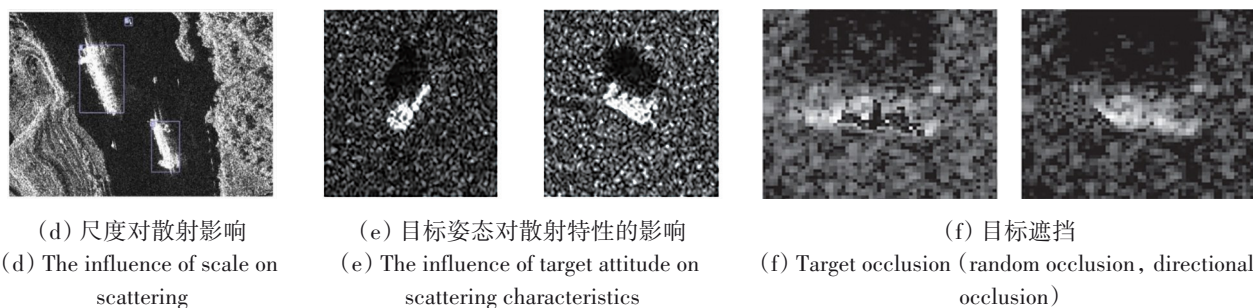


图8 SAR图像特征提取的典型挑战示例
Fig. 8 Typical challenge examples for feature extraction of SAR images

4 知识驱动的特征提取

单纯依赖图像表观数据难以捕捉遥感成像背后复杂的地学规律和物理机理。因此,突破现有特征提取范式瓶颈的关键在于为数据驱动模型注入领域知识的“认知”能力,形成“数据驱动感知”和“知识驱动推理”相结合的双驱动新范式。知识驱动的特征提取,是在数据驱动模型的架构设计、特征表示与优化过程中,显式地、结构化地将领域先验认知转化为可计算、可微分的形式,系统性融入深度神经网络的训练与推理过程,用以引导或约束网络的特征学习,其目标不仅是提高模型在特定任务上的性能,更是确保模型学习到的特征甚至推理结果与人类对该领域的认知规律相一致,从而实现从“感知”到“认知”的跨越。

在遥感信息领域,数据仍以密集型为主,知识则涵盖广泛,包括物理模型、形态学、地学等领域专业知识及目标形状、颜色等场景先验知识。单纯数据驱动存在数据获取难、细节信息丢失、虚假信息多等问题,单纯知识引导存在复杂问题难建模、病态问题难求解等问题,基于知识驱动的特征提取帮助解决单范式下存在的缺陷。本章首先系统归纳并详细阐述可用于驱动特征提取的知识体系(4.1节),进而梳理知识驱动下的网络训练范式(4.2节),最后对知识驱动范式的局限性和未来发展进行总结展望(4.3节)。

4.1 面向特征提取的知识体系及其应用

知识驱动的有效性首先取决于知识的系统性,为了理解不同知识类型在认知层级中所扮演的差异化角色,将引导特征提取的知识体系归纳为3类,并阐释其与认知过程的对应关系。

4.1.1 视觉知识

(1) 知识内涵。视觉知识是指通过人类先验或传统图像处理方法从图像中提炼出来的、对目标特性和场景结构的描述,这类知识具备明确的语义内涵,为模型提供了局部结构化约束,其核心价值在于,将传统手工特征提取器所编码的、关于“何为显著结构”的专家经验,转化为神经网络的显式监督信号或注意力引导,如将边缘先验信息作为辅助损失,实质上是在高层语义特征学习中锚定了几何不变性,缓解了纯粹数据驱动下因背景混淆或纹理干扰导致的边界模糊问题。可见光图像中的视觉知识包括目标的边缘、颜色、纹理等结构化特征(Chen等, 2024b; Hu等, 2024); SAR图像中的视觉知识则更侧重于描述电磁波散射特性,如由算法确定的强散射点、散射结构、以及振幅、相位等信息(Ren等, 2024; Pan等, 2024a; Zheng等, 2023a),提供关于地物更丰富的属性信息。

与数据驱动范式下网络从海量数据中自行学习特征不同,视觉知识通过将已提炼的、可解释的显式特征作为引导信号驱动网络训练,视觉知识的引入为模型提供了丰富的细节信息和语义线索,有助于提升网络在复杂背景或数据缺失场景下的特征提取稳定性。视觉知识来源广泛,除了依靠传统特征提取方法获取的手工特征外,受人类视觉系统启发的机制(如视觉显著性、多尺度分析等)也为构建视觉先验提供了重要思路。通过特定的融合机制将蕴含目标丰富语义信息的视觉先验知识注入网络,在降低网络对标注数据依赖性的同时提高特征表示的质量和鲁棒性。

(2) 典型应用。视觉知识源于对图像内容本身的先验理解,其核心价值在于为模型提供明确的、可解释的视觉引导信号,帮助网络聚焦关键

区域,并增强对地物结构和边界的感知能力,可有效应对语义模糊性与复杂背景干扰的挑战。

通常来说,目标的边缘是最稳定、最本质的视觉特征之一,将Canny、Sobel等算子提取的边缘信息作为显式监督信号引入网络中,可以强制网络在特征学习过程中强化边界响应,解决目标边缘语义模糊、预测不连续等问题(He等,2024; Lin等,2024; Chen等,2022; Ren等,2024),如Chen等(2024b)针对可见光图像语义分割任务提出GCIFNet,该方法利用Canny算子提取的边缘图进行显式监督,并通过一个并行的边缘增强分支将边缘信息与编码器特征进行融合,同时在训练中引入边缘监督损失,从而提升边界分割的准确性和连续性。类似的,在SAR图像中利用比率算子提取的边缘或显著散射点作为先验有助于引导网络在强噪声背景下清晰分离不同散射特性区域:Pan等(2024a)提出SRT-Net,该方法首先提取强散射点作为图节点,并通过神经网络学习其拓扑特征,再与主干网络特征融合,从而有效解决近岸场景下船舶与人工设施区分难问题。

视觉知识常通过辅助监督损失(如边缘损失、角点定位损失等)、与深度特征融合或注意力引导的方式融入网络(Liu等,2025; Ye等,2025),其优势在于实现直接、可解释性强,能有效提升网络在细节感知和结构还原方面的能力。

4.1.2 地理空间知识

(1) 知识内涵。地理空间知识是指地表自然与人文要素在时空演变过程中呈现的规律性、关联性和约束性信息,包括地表自然演变呈现的规律性时空先验信息以及地物间呈现的常理性空间先验信息。可见光图像和SAR图像本质上是对一定范围内地理空间内容的映射,地理空间知识,包括位置、拓扑、时序关系等,为特征学习注入了空间关联与上下文推理能力,其核心价值在于,通过构建一个超越像素的“关系场”,将离散的地物特征组织成具有语义连贯性的空间结构,即将地理学基本定律转化为结构化归纳偏置,这有助于增强网络对地理空间内容的感知能力,促进网络在复杂场景下的信息提取与推理能力。

在遥感可视化任务中,地理空间知识可进一步细分为时序知识、度量知识、空间结构知识和语义知识。时序知识是指地物在时间维度上表现

出的动态变化规律及其与空间特征的耦合关系,根据变化频率,可将其划分为短时间尺度(如昼夜、季节变化等)和长时间尺度(如年际、地质演变等)。短时间尺度变化通常表现为高频特征,适用于短期预测与实施决策(Lu等,2025; Wieland等,2025);长时间尺度变化则呈现出累积性的特点,主要用于气候变化、海岸线变迁等长期研究(Alzamili等,2025)。

度量知识指对地物及其空间关系的定量化描述,包括几何度量、地形度量、空间关系度量等。几何度量知识是对地物基本几何属性的量化,如长度、面积等,可为网络提供地物几何属性信息作为弱监督信号,也可作为约束条件优化网络输出;地形度量知识提供了地物形态特征如高度、坡度、坡向等,高度信息如数字高程模型DEM(Digital Elevation Model)数据等有利于提升分类和识别精度(Anitha等,2024; Xiao等,2025a),地形信息则帮助提高复杂地貌场景下的地物分割精度(Huang等,2025);空间关系度量知识将地物空间关系进行编码以提供地物间距离、密度、边界等信息,通过分析空间关系变化情况提供推理信息(Pawar和Sharma,2025)。

空间结构知识指地物在空间分布、排列组合及相互作用中形成的空间模式及关联规则,包括拓扑结构知识和层次结构知识。拓扑结构知识描述了地物间空间关系不变性,如连接性、邻接性、包含性等不受几何变换影响的不变关系,为空间分析、推理和决策提供结构化框架或拓扑空间约束(Geng等,2020);层次结构知识提供对象间的等级体系信息,支持网络实现分层特征提取细化语义粒度,增强对宏观—微观结构的感知能力,利于提高分类精度(Lin和Guo,2024)。

语义知识是对地理空间对象类别、属性、功能、上下文关联等的抽象描述,赋予对象明确的语义内涵,构建“特征—关系—场景”三维体系提升网络语义理解能力,可应用于上下文推理、多义性地物分类、小样本或弱监督等应用中(Wei等,2025; Qian等,2024)。

(2) 典型应用。地理空间知识描述了地物在真实世界中的时空分布、拓扑关系与演变规律,此类知识的引入使特征提取过程超越了单纯的像素分析,能够从地理学逻辑层面理解场景,从而引导网络学习超越像素表观、更具语义判别性和

时空泛化能力的特征表示,有效应对可见光图像和SAR图像面临的独特挑战。

可见光图像中普遍存在“同物异谱”或“同谱异物”的语义模糊现象,根源在于单纯依赖光谱和纹理特征难以捕捉高层语义上下文信息,地理空间知识中的空间关系和拓扑关系,通过将图像理解从孤立的实体理解转化为对场景内地理实体间结构化关系的推理,为语义理解提供了强有力的消歧依据:Li等(2022)提出一种地理知识驱动的表达学习框架,该方法首先利用地物空间分布模式先验构建结构化知识表示,并通过图注意力机制将其与深度特征相融合,同时设计空间一致性约束与对比损失来校正特征学习,确保网络学到符合地理规律的本质语义。

当成像时间不同时,同一场景下的地物由于太阳高度角、大气状况或季节更替等因素可能呈现出差异巨大的表现特征,造成严重的特征域偏离,然而,尽管地表实体的视觉外观会改变,其地理位置及与其他实体间的关系是固有的,这为应对成像条件敏感挑战给出了解决思路:Ayush等(2021)提出一个视觉特征提取器,其利用“同一地理位置在不同时间具有语义一致性”的先验构建正样本对,通过将同一坐标点、不同时间采样的图像作为正样本进行训练,迫使网络学习跨时间的特征不变性,从而提取本质的地理语义特征,提升模型在跨时相中的泛化能力;Zhao等(2023b)针对变化检测任务提出GeSAnet,设计了地理空间位置校正模块配准偏差,并利用地理空间内容推理模块构建像素级空间关系图使每个像素能聚合邻域变化信息,从而依据“地物变化在空间上不孤立”的先验知识有效识别并抑制孤立的伪变化点,并增强网络发现在空间上形成一致变化模式的小范围变化区域的能力。

可见光和SAR图像中地物尺度差异大,且在复杂地区存在背景干扰或几何畸变等问题,对特征提取的尺度一致性和地形适应性提出了高要求,数字高程模型DEM及其衍生的地形度量知识如高程、坡度等为此提供了关键的第三维定量信息。Li等(2023)提出一种地形知识耦合的深度学习识别框架,该框架利用水文分析方法从DEM中提取沟道网络、地形湿度指数等地形功能特征,并与图像外观特征进行空间对齐与融合,实现语义互补,还设计了基于水文连通性的空间约束规则与

地形一致性损失,监督网络捕获符合水力学原理的淤地坝区位特征,有效克服因尺度失真和背景干扰导致的误识别。

实际地物间存在错综复杂的时空关系与相互联系,这为“地理空间”向“图像空间”的转换带来挑战,如何将地物间的关系准确全面的表述与利用成为一大难点,早期主要通过坐标嵌入、空间金字塔池化等方式提供感兴趣对象的地理编码信息,然而,地理空间中的实体关系往往具有多层次、非线性特点,简单的地理信息编码难以全面刻画复杂的地理语义信息,为了更好地利用实体间的多种属性关系,基于图结构的知识表示方式逐渐成为主流,其核心思想是将地物视为节点并构建“实体—属性—关系”三元组进行表示,图结构的方式展现了地物间丰富的显性关系并支持计算、推理等方式深度挖掘实体间隐性关系(Zhu, 2024),如建筑物节点关联“高度”、“用途”等属性,并与道路、绿地等节点形成“相邻”、“包含”等关系,还能通过图结构推理挖掘隐性关系,如高密度建筑物节点可能对应商业区、住宅区等。Wang等(2024c)还将遥感智能解译与地理信息集成化,开发了遥感视觉任务与地理空间信息的多尺度、多阶段综合技术及相应系统,帮助全面感知复杂的地理空间。

4.1.3 物理知识

(1) 知识内涵。物理知识是通过特定物理事实推理、科学手段重复观察和实验得出的为科学界一致接受的结论,包括抽象的物理模型、基于物理规律的优化公式等用于反映特定物理现象或物理规律的数学描述。物理知识揭示了数据生成的根本机理,其核心价值在于将人类对真实世界的物理认知转化为算法约束,这意味着特征表示空间将被限制在物理定律所允许的流形上,通过可微分物理建模或物理一致性损失,网络学习的不再是简单的像素到标签的映射,而是具有明确物理意义的参数或中间表示,这有助于提升网络在低信噪比、严峻成像条件下的外推能力和可解释性。

遥感可视化任务中的物理知识可进一步细分为成像物理知识、环境物理知识和地物物理知识。成像物理知识指传感器与地物相互作用过程中遵循的物理规律,为图像解译提供成像机理知识,

可见光图像中主要包括光学畸变模型（径向/切向畸变校正等）、光子成像退化模型、光学成像约束模型等；SAR图像包括电磁散射模型、雷达方程、辐射传输模型等（Li等，2024d；Wei等，2024a；Wang等，2023c）。成像物理知识将物理规律转化为数学约束或特征表示，解决成像机理、数据低质等带来的挑战。

环境物理知识指环境系统与遥感成像间相互作用的物理规律，强调环境对观测的反馈，如大气、降水等环境因素对地物成像的影响，通过将上述反馈进行量化建立从原始信号到真实地物特性的逆向映射模型，从而消除环境干扰，如大气校正模型消除辐射失真，解决大气散射、云雾干扰等引起的图像质量下降问题、流体动力学模型耦合多物理场，提高网络动态预测能力，避免纯数据驱动的物理不合理性（Khan等，2025；Lihe等，2024），多用于基于环境因素的图像重建任务中。

地物物理知识描述了地物固有属性及其与传感器信号相互作用规律，呈现了地物本质特性（介电常数、粗糙度等）与遥感信号间的确定性映射关系，可应用于高精度识别与分类、分割、隐蔽/伪装目标检测等任务中，如介电常数模型反演掩埋目标深度及分类、斯托克斯矢量参数图像反演伪装目标偏振特异性等（Fluhrer等，2022；Wang等，2022a；Shen等，2021）。

（2）典型应用。物理知识揭示了可见光和SAR图像形成的数学物理规律，是从观测数据反演地物本征属性的桥梁，通过物理模型驱动深度学习，能从信号源头施加约束，有效应对可见光和SAR图像面临的独特挑战。

SAR图像的物理本质是地物对微波的相干成像，将电磁散射模型、大气物理模型引入网络，可以提升特征的可解释性和低信噪比下的鲁棒性，有效应对可见光图像特征退化、SAR图像散射建模困难、杂波干扰的挑战：Wang等（2023c）提出一种将麦克斯韦方程组直接嵌入网络前向传播的方法，通过构建可微分的物理仿真层并通过最小化预测与真实散射场差异进行损失约束，使网络直接利用确定的电磁规律进行推理，显著降低对标注数据的依赖并提升泛化能力；Lihe等（2024）针对图像去雾任务提出PhDnet，该方法将大气散射物理模型作为可微计算层嵌入网络，从而将病

态的去雾问题转化为物理参数估计问题，同时对解空间施加物理约束，使网络输出兼具物理合理性与视觉自然度的去雾图像。

为应对特征度量的高度不稳定性，一种可行的方法是利用物理模型生成不同条件下的仿真数据作为增广，通过引导网络学习对几何变化不敏感的本质散射特性提升特征度量的稳定性：Li等（2024d）构建一种基于自注意力的网络，该网络将SAR目标表现为稀疏强散射点的成像物理机制作为先验嵌入图像重建过程，同时在网络目标函数中引入由CNN参数化的散射特征先验项，使重建过程优先收敛至包含目标信号的解，并利用自注意力模块建模散射点间的长程依赖关系重建出全局结构，最终生成符合目标散射本质特征的高质量重建图像。

为实现物理知识对深度模型的有效约束，关键是将物理方程转化为模型可用的约束，主要有两种通用路径，其构建流程如图9所示。

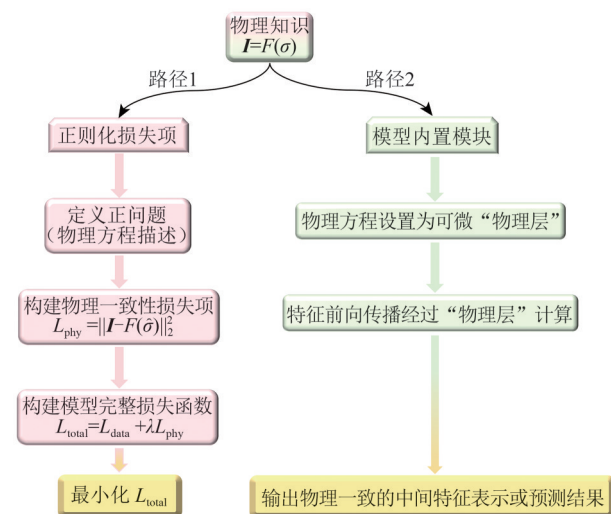


图9 物理约束的两种实现路径

Fig. 9 Two implementation paths of physical constraints

一种通用路径是将物理知识作为正则化损失项的形式约束模型训练，其核心是构建物理一致性损失，首先需要明确表示从物理参数 σ 到观测值 I 的正问题 $I = F(\sigma)$ ，通过将模型的预测输出 $\hat{\sigma}$ 带入上述方程中可得到模型对应的观测值 $F(\hat{\sigma})$ ，物理一致性损失定义为真实观测值和模型观测值间的差异 $L_{\text{phy}} = ||I - F(\hat{\sigma})||_2^2$ ，该正则化项与数据驱动损失相结合共同指导模型的训练。

另一种路径是将物理知识对应的数学公式作

为内置模块整合至模型架构中,成为前向传播的一部分。模型的特征提取模块负责从原始数据中提取特征,随后将特征输入至由物理模型定义的“物理层”中,该层被设计为可微分的,使得梯度可以反向传播,从而实现对前端特征的物理约束。“物理层”最终生成预测结果或具有明确物理含义的中间特征表示,保证了模型内部机制符合物理规律,增强了模型的可解释性。

物理知识主要用于解决纯数据驱动方法难以有效处理的成像机理、气候条件等带来的挑战(Lihe等, 2024; Li等, 2025b; Xu等, 2024),通常以物理模型或算法的形式引入网络,通过将物理模型数值化(Lihe等, 2024; Li等, 2024c)或采用优化算法、物理约束算法等作为损失函数(Chen等, 2024d; Pan等, 2021b),指导网络学习具有明确物理意义的特征,提高结果的可信度与可解释性。

总的来说,特征提取技术按照“传统手工设计—纯数据驱动—数据与知识双驱动”的演进路线发展,知识驱动方法不仅继承了纯数据驱动方法相对于传统方法的性能优势,更通过引入领域知识进一步提升模型在复杂场景下的可解释性和适应性,巩固了深度学习模型在解决传统方法3大缺陷方面取得的成果,表4列出了部分知识驱动的相关文献及其在各类任务中的性能增益(括号内数值表示较纯数据驱动方法的性能提升百分数)。此外,知识驱动的更深层意义在于对特征学习范式的根本性变革:(1)假设空间的显式约束。纯数据驱动模型在庞大的假设空间中搜索,易陷入局部最优或学习虚假关联,知识通过提供领域逻辑约束,大幅收缩了可行的假设空间,这降低了模型对海量标注数据的依赖,提升学习效率与泛化下限;(2)特征语义的可解释性注入。知识将难以理解的深度特征与人类可理解的概念关联起来,当网络的特征通道被显式引导去响应某种物理参数或地理属性时,其特征表示便自然具备了语义透明度,这使模型的决策过程由“黑箱”变为“灰箱”,为结果的可信度与追溯提供了基础;(3)归纳偏置的领域定制。知识驱动向网络注入了领域特定的归纳偏置,如SAR图像的斑点噪声统计特性、遥感图像的多尺度层级性等,这使网络能更敏锐捕捉到对任务至关重要的模式;(4)解决病态问题的能力。许多遥感反演任务如去雾、

图像重建等是病态的,存在多解,知识的引入将数据驱动的回归问题转变为基于物理模型的约束优化问题,为正则化解提供强有力的科学依据,从而导向更合理、更稳定的唯一解。

4.2 知识驱动的深度模型训练范式

知识驱动的关键在于实现领域知识与数据驱动模型的有效融合,这种融合并非简单叠加,而是需要设计特定的机制,将人类认知转化为算法可执行的约束。如图10所示,根据知识在模型中被表示和利用的根本形式,将其归纳为3种核心范式:知识即架构、知识即数据、知识即损失,其本质区别在于知识作用于模型的阶段与抽象层级,3种范式为解决纯数据驱动模型的固有缺陷提供了不同层级的技术路线。

(1)知识即架构。该范式将领域知识直接编码为模型前向传播图中的一个或多个可微分计算模块,这些模块通常具备明确的物理或数学含义,其功能和参数由先验知识定义而非从数据中学习。知识作为模型结构的一部分,引导信息流遵循特定的领域逻辑,如将大气散射模型构建为一个可微网络层(Lihe等, 2024)等。该范式的优势在于其可解释性强,知识的作用过程透明且能严格保证中间特征表示的物理合理性,适用于处理具有明确数学模型或规则体系的任务。然而,其局限性也十分显著:1)建模鸿沟。物理模型是对现实世界的简化表示,在运用时存在一定偏差;2)灵活性不足。固定架构难以适应知识不确定性或快速演变场景;3)工程复杂性高。将复杂的知识体系设计成高效且可微分的计算层具有挑战。

(2)知识即数据。该模式不改变模型架构,而是将知识转化为训练样本或特征表示与原始数据一同输入模型,知识通过数据增广、特征融合或构造特定样本对等方式隐式地引导模型学习,如将手工提取的边缘特征图与深度特征融合训练网络(Zheng等, 2023b)、将先验地理信息编码作为数据增广(Wang等, 2024b)等。该范式的核心在于将知识编码至数据的分布或关联中,其实现简单、与现有深度学习架构兼容性好,且能利用模型强大的拟合能力从“知识数据”中学习 to 更丰富的特征表示,但其缺点在于知识的引导作用变得间接和模糊,模型可能只能学习到知识数据表面的统计特征而非知识蕴含的深层规律,导致知识的“语义稀释”。

表 4 知识驱动特征提取方法的性能增益分析
Table 4 Performance gain analysis of knowledge driven feature extraction methods

类型	图像	文献	解决的挑战								数据集	指标	性能/%
			遮挡	低对比度	语义相似	多尺度	多方向	密集分布	小目标	小样本	散射点关系	相干噪声	
视觉知识	可见光图像	Anitha 等, 2024			√					√			WHDLD F1、mIoU 77.52(0.89 ↑) 63.84(0.60 ↑)
		Chen 等, 2024b			√	√			√				Yinchuan mAcc 88.40(2.37 ↑)
		He 等, 2024		√	√	√			√				Vaihingen mIoU、F1 71.07(1.77 ↑) 82.65(1.28 ↑)
		Hu 等, 2024			√	√	√		√				ISPRS Potsdam F1、mIoU 89.67(1.03 ↑) 91.74(1.22 ↑)
		Yu 等, 2024b	√			√		√	√				AI-TOD AR@0.5:0.95 AP@0.5 AP@0.5:0.95 65.50(1.10 ↑) 79.40(1.60 ↑) 62.50(1.60 ↑)
		Zheng 等, 2023b	√		√	√		√	√				iSAID mIoU 64.80(1.50 ↑)
		Chen 等, 2022	√	√	√	√		√	√				WHU-CD F1、IoU 91.71(0.60 ↑) 84.68(1.01 ↑)
	SSAR图像	Pan 等, 2024a				√	√	√	√		√	√	LSSDD F1、AP@0.5 80.89(2.28 ↑) 82.71(3.94 ↑)
		Ren 等, 2024			√				√		√	√	XIAN Image F1 81.80(—)
		孙珊珊 等, 2024					√	√	√			√	SSDD P、R 92.68(1.06 ↑) 79.87(2.60 ↑)
		梁烽 等, 2023				√					√		SARSea Land V1.0 mIoU 96.47(0.69 ↑)
		Zheng 等, 2023a			√							√	OpenSARShip FuSAR-Ship Acc 78.60(5.01 ↑) 87.23(3.88 ↑)
地理空间知识	可见光图像	李彦胜 等, 2024			√	√			√				UCM Deep Globe mIoU 75.13(12.1 ↑) 57.84(1.53 ↑)
		Lin 和 Guo, 2024				√		√	√				Sentinel-2 Acc 94.35(12.4 ↑)
		Wang 等, 2024b			√				√				SV-1 Image mIoU 91.20(1.51 ↑)
		Zhao 等, 2023b		√		√	√		√				LEVIR-CD CDD F1 90.05(0.91 ↑) 95.14(4.83 ↑)
	SSAR图像	Chen 等, 2024c			√			√	√		√	√	GF-3 Image — —
		Luo 等, 2024			√	√	√			√			SAR-RADD mAP 74.70(5.30 ↑)
	可见光图像	Chen 等, 2024d								√			LandSat Image NSE 94.00(24.0 ↑)
		Lihe 等, 2024	√	√		√							RSHaze-L PSNR(dB) SSIM 35.10(0.44 ↑) 93.48(0.32 ↑)
		Li 等, 2024c	√					√	√				Water-CD F1、IoU 91.71(1.22 ↑) 84.69(2.06 ↑)
		Xu 等, 2024	√	√									UBCSet PSNR(dB) SSIM 42.78(0.52 ↑) 96.40(0.51 ↑)
物理知识	SAR图像	Fan 等, 2024a						√	√			√	Sentinel-1 mAP、F1 94.00(7.40 ↑) 88.12(6.01 ↑)
		Li 等, 2025b	√							√	√		MSTAR PCC 99.46(0.28 ↑)
		Huang 等, 2022			√			√		√	√		Sentinel-1 OA 85.27(2.68 ↑)

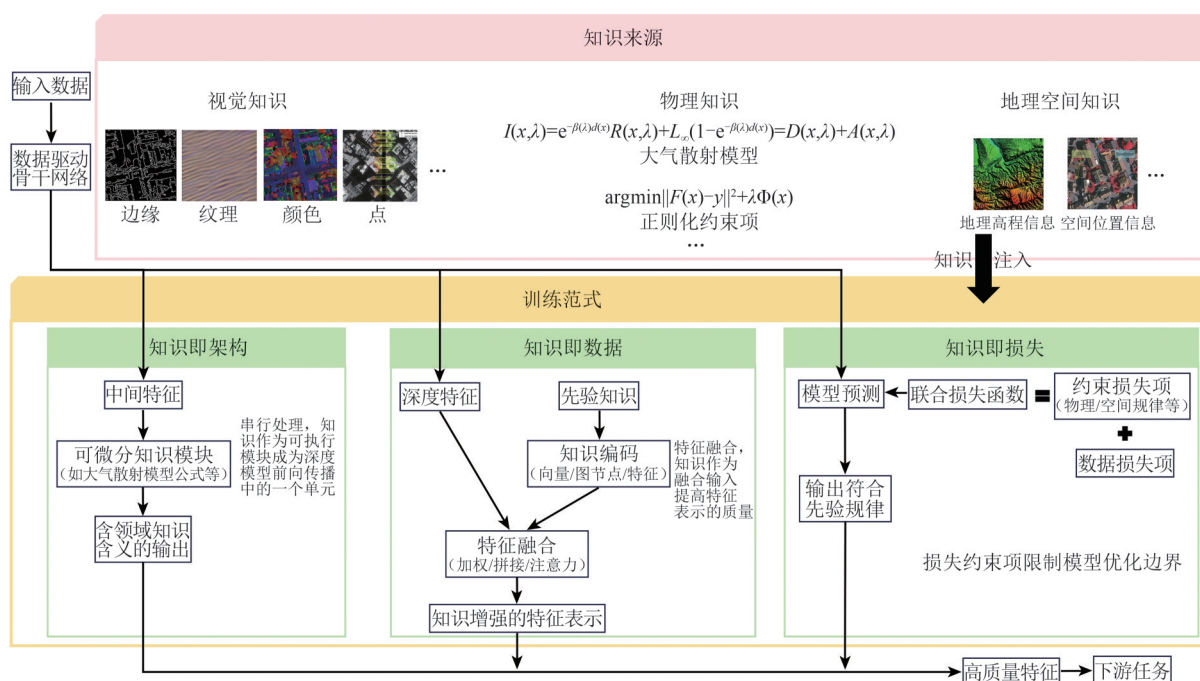


图10 知识驱动的遥感图像特征提取框架

Fig. 10 Knowledge driven remote sensing image feature extraction framework

(3) 知识即损失。该范式将领域知识表达为优化目标中的正则化项或约束项，通过损失函数直接塑造模型的特征空间与输出分布，知识不参与前向计算，而是在反向传播中惩罚违反领域规律的模型行为，如设计物理一致性损失 (Zhang 等, 2025) 或空间一致性损失 (Li 等, 2022) 等。该范式是最灵活的知识融合方式，具有很强的普适性，任何数学化的知识均可转化为损失项且能与任何模型架构结合。然而，其挑战在于需要设计数据损失项和知识损失项间的权重参数，且多约束冲突时优化可能不稳定。

4.3 知识驱动的限制性与未来发展

尽管知识驱动的特征提取范式通过3种训练范式在提升模型性能并应对可见光图像和SAR图像独特挑战中取得了显著进展，并实现了特征学习范式的根本性变革，但需清楚认识到，当前研究整体仍处于“知识应用”的初级阶段，远未达到“知识理解”和“知识创造”的程度。本节旨在系统性反思该范式的共性局限，并勾勒出其向更高级智能形态的发展。

4.3.1 共性局限与深层挑战

(1) 知识表征的“形式化”困境。尽管知识驱动方法在增强模型可解释性等方面展现出潜力，

但其实现前提是将领域知识转化为神经网络可计算的形式，这一过程要求将领域专家的认知转化为离散的数学方程或约束 (Chen 等, 2024b; Wang 等, 2023c)，而知识在转化过程中不可避免存在信息损失和语义偏差，该问题不仅损害了知识的完整表达，更限制了知识在某些难以被“形式化”的场景下的应用。

(2) 知识引导的静态性和孤立性。当前大多方法仍将知识视为预设的、固定的模块，不同类型的知识往往被独立编码并引入网络，缺乏跨知识类型的整合机制，更关键的是，一旦知识被引入网络即被“固定”，难以随着网络的优化进行自适应更新 (Lihe 等, 2024)。上述静态、孤立的知识引导方式，在面对未知场景或场景迁移时易导致模型的泛化性和鲁棒性下降。

4.3.2 未来发展

为克服上述局限，未来知识驱动范式的研究有必要向更深层、更具自适应性的方向发展：

(1) 发展具备动态更新能力的知识系统。未来的知识不应再保持静态性，而应具备动态更新优化的能力，这需要引入元学习或在线学习机制 (Zhang 等, 2024d; Miao 等, 2024)，通过不断更新内部知识表示提升模型在开放环境下的适应性。

(2) 构建因果知识驱动的特征学习方法。当前方法大多基于统计相关性建立知识约束，易受无关因素影响，通过构建因果推断理论有助于指导模型学习特征与地物间的关联模式及因果生成机制，增强模型识别虚假关联的能力以及对影响因素的鲁棒性（Zhang等，2024d；Li等，2024a）。

(3) 探索以大模型为通用知识基底的适配架构。近年来，以海量多模态数据预训练出的遥感基础大模型，展现出强大的通用表征能力，其内部蕴含丰富的隐式知识，在此基础上构建轻量化、模块化的专用任务适配器，通过将领域知识进行编码并协调大模型在特定任务上的特征生成与推理过程，实现知识的深度挖掘与利用（Li等，2025a）。这种“通用基底+专用适配”的架构，既能继承大模型的泛化性与表征能力，又能通过任务适配器快速适应专业场景。

5 展 望

随着应用需求不断改变及技术的快速发展，特征提取技术也向着多样化方向发展，本文从最新的研究进展出发展望未来特征提取技术可能的发展趋势。

(1) 遥感大模型。自遥感领域发展进入深度学习时代以来，遥感图像任务处理发生变革性变化，质量和精度有了质的提升，但目前深度模型多为专用模型，局限于特定的遥感任务或应用场景，训练后难以直接用于其他应用甚至相近应用中（如从目标分类到目标检测）。因此，实现高效、准确的通用特征提取算法仍然面临巨大挑战。目前深度学习在模型层面已经出现由专用模型向“大模型+专用模型”协同的发展趋势，这一趋势旨在通过在海量、多样化的遥感数据上预训练通用特征提取网络，再结合下游任务进行轻量级适配，以克服专用模型的局限性，有助于解决大数据存在的多类不确定性建模、复杂数据环境下多源信息融合、基于知识驱动的遥感大数据挖掘等挑战。当前，遥感大模型仍处于早期研究阶段，其发展路径、训练范式与应用生态仍在探索之中，仍面临如数据异构、模态对齐等核心挑战（张永军等，2024；秦其明，2025；张帅豪和潘志刚，2025）。尽管如此，遥感大模型仍具备克服多任务单独建模、各模态分别建模等挑战的潜力，基于遥感大模型打包部署可简化系统结构、降低维护

成本，形成多任务统一建模、多模态联合建模的通用模型。

(2) 数据与知识双驱动发展。双驱动范式中数据是基础，知识是关键，知识的获取能补充、指导与优化深度模型对数据的利用，目前知识的获取主要来源于两种途径：专业领域知识和跨领域知识的交叉融合。前者主要依赖专家经验或场景先验信息优化网络训练（Chen等，2024b；Lu等，2025；Wang等，2023c），后者通过借鉴或迁移其他领域中的成熟知识体系帮助解决本领域内难题，如基于传统尺度空间的启发提出特征金字塔网络FPN解决图像、目标多尺度问题（Lin等，2017）。此外，知识的突破与抽象对于构建新型神经网络具有推动作用，近几年部分学者致力于研究具有简单神经系统的龙虾脑工作机制，旨在挖掘其工作机制并从中建模出简单神经网络，这项工作若成功对于深度学习将是重大突破。知识对于网络的学习和构建有重要指导意义，挖掘基础理论、发展新型模型并将其与深度学习灵活结合是具有巨大发展潜力的方向之一。

(3) 前沿技术交叉融合式发展。随着深度学习在不同领域上的快速发展，涌现出一批批优秀的模型，如分段任意模型SAM（Segment Anything Model）（Chen等，2024a）、扩散概率模型（Fan等，2024b）、神经架构搜索（Wang等，2024f）等，这些模型在灵活性、自动化、生成能力等方面展露出优越性能，未来基于此类模型与现有技术的融合发展可为遥感智能解译提供新的动力和方向。

6 结 语

本文系统梳理了可见光与SAR图像特征提取的技术发展脉络，从传统手工特征到数据驱动的深度学习方法，再到融合领域知识的知识驱动范式，逐步揭示了特征提取从“人工设计”到“数据学习”再到“认知引导”的演进逻辑。主要结论如下：

首先，传统方法具有可解释性强、计算效率高等优势，但在泛化能力与非线性建模方面存在局限性；深度学习方法通过端到端学习显著提升了特征的表征能力，但对标注数据依赖性强，且在复杂遥感场景下面临语义模糊、尺度极端、成像条件敏感等挑战。

为应对上述挑战, 本文引入知识驱动的特征提取范式, 将引导网络学习的知识体系归纳为视觉知识、地理空间知识和物理知识3类。视觉知识通过边缘、纹理等先验增强网络对地物结构的感知; 地理空间知识利用空间拓扑、高程模型等约束提升特征的场景适应性; 物理知识则通过成像模型、散射机制等确保特征符合真实物理规律。知识通过“架构嵌入”“数据增强”“损失约束”等多种方式与深度学习模型融合, 形成“数据驱动感知+知识驱动推理”的双驱动动态, 不仅提升了模型性能, 也增强了特征的可解释性与跨域泛化能力。

展望未来, 特征提取技术将继续朝着多模态融合、大模型预训练、知识自适应演化等方向发展。遥感大模型的构建将推动通用特征提取框架的形成; 数据与知识的深度融合将进一步提升模型在少样本、跨域、弱监督等场景下的实用性; 而前沿技术如分段任意模型、扩散概率模型等与符号推理、因果推断、元学习等的交叉融合, 将为遥感智能解译注入新的活力。通过持续推动理论创新与技术整合, 特征提取将在遥感图像理解与应用中发挥更加核心与智能的作用。

参考文献 (References)

- Ai J Q, Qu Z, Zhao Z C, Zhang Y, Shi J and Yan H. 2024. An SAR target classification algorithm based on the central coordinate attention module. *IEEE Sensors Journal*, 24(2): 1941-1952 [DOI: 10.1109/JSEN.2023.3338218]
- Al Saidi I, Rzaia M and Debayle J. 2023. Completed homogeneous LBP for remote sensing image classification. *International Journal of Remote Sensing*, 44(12): 3815-3836 [DOI: 10.1080/01431161.2023.2227320]
- Alzamili M H E, Mahmood V and Rao P J. 2025. Assessment of urban heat island based on remote sensing and geo-spatial approach in Al-Kut Region, Iraq. *Journal of Earth System Science*, 134(2): 103 [DOI: 10.1007/s12040-025-02558-6]
- Anitha V, Manimegalai D and Kalaiselvi S. 2024. Downstream lingering attention transformer network (DsLATNet) for land use land cover classification: a bicolor deep learning framework. *Applied Soft Computing*, 165: 112074 [DOI: 10.1016/j.asoc.2024.112074]
- Asokan A. 2023. A self-adaptable Manta ray optimized Gabor filter for satellite image enhancement. *Earth Science Informatics*, 16(2): 1503-1517 [DOI: 10.1007/s12145-023-00963-3]
- Ayush K, Uzkenet B, Meng C L, Tanmay K, Burke M, Lobell D and Ermon S. 2021. Geography-aware self-supervised learning//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal: IEEE: 10161-10170 [DOI: 10.1109/ICCV48922.2021.01002]
- Bai L, Liu Q X, Li C L, Zhu C L, Ye Z and Xi M. 2022. A lightweight and multiscale network for remote sensing image scene classification. *IEEE Geoscience and Remote Sensing Letters*, 19: 8012605 [DOI: 10.1109/LGRS.2021.3078518]
- Cao J Y, You Y N, Li C and Liu J. 2024. TSK: a trustworthy semantic keypoint detector for remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5607120 [DOI: 10.1109/TGRS.2024.3352899]
- Cao M X, Lei J, Xie W Y, Zhang J Q and Li Y S. 2023. SAR-Net: multi-scale direction-aware SAR network via global information fusion. [2024-10-09]. <https://arxiv.org/abs/2312.16943v1>
- Cao W, Wu S Q, Liu Z Y and Agaian S. 2022. Scale-aware guided and structure-preserved texture filter. *IEEE Geoscience and Remote Sensing Letters*, 19: 8021305 [DOI: 10.1109/LGRS.2021.3113985]
- Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A and Zagoruyko S. 2020. End-to-end object detection with transformers//*Proceedings of the 16th European Conference on Computer Vision*. Glasgow: Springer: 213-229 [DOI: 10.1007/978-3-030-58452-8_13]
- Chang H H, Wu G L and Chiang M H. 2019. Remote sensing image registration based on modified SIFT and feature slope grouping. *IEEE Geoscience and Remote Sensing Letters*, 16(9): 1363-1367 [DOI: 10.1109/LGRS.2019.2899123]
- Chen K Y, Liu C Y, Chen H, Zhang H T, Li W Y, Zou Z X and Shi Z W. 2024a. RSPrompter: learning to prompt for remote sensing instance segmentation based on visual foundation model. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 4701117 [DOI: 10.1109/TGRS.2024.3356074]
- Chen L, Qu Z Y, Zhang Y, Liu J Y, Wang R W and Zhang D Z. 2024b. Edge-enhanced GCIFNet: a multiclass semantic segmentation network based on edge enhancement and multiscale attention mechanism. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 4450-4465 [DOI: 10.1109/JSTARS.2024.3357540]
- Chen L F, Fang Z H, Xing J and Cai X M. 2024c. How can geostatistics help us understand deep learning? An exploratory study in SAR-based aircraft detection. *International Journal of Applied Earth Observation and Geoinformation*, 134: 104185 [DOI: 10.1016/j.jag.2024.104185]
- Chen Y F, Dong Y R, Sun Y X and He C. 2024d. Typical lake area is accurately predicted and assessed based on deep learning algorithms and associated physical mechanisms. *Earth Science Informatics*, 17(3): 2317-2330 [DOI: 10.1007/s12145-024-01282-x]
- Chen Z L, Zhou Y, Wang B, Xu X W, He N, Jin S and Jin S R. 2022. EGDE-Net: a building change detection method for high-resolution remote sensing imagery based on edge guidance and differential enhancement. *ISPRS Journal of Photogrammetry and Remote Sensing*, 191: 203-222 [DOI: 10.1016/j.isprsjprs.2022.07.016]
- Cordeiro G M, Cintra R J, Rêgo L C and Nascimento A D C. 2019. The gamma generalized normal distribution: a descriptor of SAR imagery. *Journal of Computational and Applied Mathematics*, 347: 257-272 [DOI: 10.1016/j.cam.2018.07.045]

- Dai J G, Shi N N, Zhang T D and Xu W H. 2024. TCME: thin cloud removal network for optical remote sensing images based on multi-dimensional features enhancement. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5641716 [DOI: 10.1109/TGRS.2024.3464862]
- Dalal N and Triggs B. 2005. Histograms of oriented gradients for human detection//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. San Diego: IEEE, 1: 886-893 [DOI: 10.1109/CVPR.2005.177]
- Dang G P, Mao Z A, Zhang T Y, Liu T, Wang T, Li L Z, Gao Y, Tian R Q, Wang K and Han L. 2024. Joint superpixel and Transformer for high resolution remote sensing image classification. *Scientific Reports*, 14(1): 5054 [DOI: 10.1038/s41598-024-55482-y]
- Dash S and Senapati M R. 2020. Noise robust Laws' filters based on fuzzy filters for texture classification. *Egyptian Informatics Journal*, 21(1): 37-49 [DOI: 10.1016/j.eij.2019.10.003]
- Dellinger F, Delon J, Gousseau Y, Michel J and Tupin F. 2015. SAR-SIFT: a SIFT-like algorithm for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(1): 453-466 [DOI: 10.1109/TGRS.2014.2323552]
- Ding K M, Yang Z D, Wang Y Y and Liu Y M. 2019. An improved perceptual hash algorithm based on U-net for the authentication of high-resolution remote sensing image. *Applied Sciences*, 9(15): 2972 [DOI: 10.3390/app9152972]
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J and Houslsby N. 2021. An image is worth 16x16 words: transformers for image recognition at scale//*Proceedings of the 9th International Conference on Learning Representations*. [s.l.]: ICLR: 1-21
- Fan J C, Sui Z T and Wang X Z. 2024a. Multiphysical interpretable deep learning network for oil spill identification based on SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 4202015 [DOI: 10.1109/TGRS.2024.3357800]
- Fan J C, Zhou J L, Wang X Z and Wang J. 2023a. A self-supervised transformer with feature fusion for SAR image semantic segmentation in marine aquaculture monitoring. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 4207915 [DOI: 10.1109/TGRS.2023.3321595]
- Fan J F, Shi Z W, Ren Z P, Zhou Y K and Ji M. 2024b. DDPM-SegFormer: highly refined feature land use and land cover segmentation with a fused denoising diffusion probabilistic model and transformer. *International Journal of Applied Earth Observation and Geoinformation*, 133: 104093 [DOI: 10.1016/j.jag.2024.104093]
- Fan L L, Zhou Y, Liu H M, Li Y J and Cao D P. 2023b. Combining Swin transformer with UNet for remote sensing image semantic segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5530111 [DOI: 10.1109/TGRS.2023.3329152]
- Feng M Z, Sun X, Dong J Y and Zhao H R. 2022. Gaussian dynamic convolution for semantic segmentation in remote sensing images. *Remote Sensing*, 14(22): 5736 [DOI: 10.3390/rs14225736]
- Fluhrer A, Jagdhuber T, Tabatabaenejad A, Alemohammad H, Montzka C, Friedl P, Forootan E and Kunstmann H. 2022. Remote sensing of complex permittivity and penetration depth of soils using P-band SAR polarimetry. *Remote Sensing*, 14(12): 2755 [DOI: 10.3390/rs14122755]
- Fu K C, Zhang Z X and Qu Y. 2016. SAR image road detection based on Canny-ROA operator and Hough transform. *Modern Electronics Technique*, 39(23): 1-4 (付凯城, 张竹馨, 曲毅. 2016. 基于Canny-ROA算子和Hough变换的SAR图像道路检测. *现代电子技术*, 39(23): 1-4) [DOI: 10.16652/j.issn.1004-373x.2016.23.001]
- Fukunaga K. 1990. *Introduction to Statistical Pattern Recognition*. 2nd ed. Boston: Academic Press [DOI: 10.1016/C2009-0-27872-X]
- Gao T, Lan C Z, Zhou C X, Zhang Y X, Huang W J, Wang Y Q and Wang L H. 2025. Arctic sea ice motion retrieval from multisource SAR images using a keypoint-free feature tracking algorithm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 230: 258-274 [DOI: 10.1016/j.isprsjprs.2025.09.013]
- Gao Y P, Luo X L, Gao X J, Yan W H, Pan X and Fu X L. 2024. Semantic segmentation of remote sensing images based on multi-scale features and global information modeling. *Expert Systems with Applications*, 249: 123616 [DOI: 10.1016/j.eswa.2024.123616]
- Garg M and Dhiman G. 2021. A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants. *Neural Computing and Applications*, 33(4): 1311-1328 [DOI: 10.1007/s00521-020-05017-z]
- Geng K, Sun X, Yan Z Y, Diao W H and Gao X. 2020. Topological space knowledge distillation for compact road extraction in optical remote sensing images. *Remote Sensing*, 12(19): 3175 [DOI: 10.3390/rs12193175]
- Guo Q, Tong L Z, Yao X D, Wu Y W and Wan G T. 2022. CD_HIEFNet: cloud detection network using haze optimized transformation index and edge feature for optical remote sensing imagery. *Remote Sensing*, 14(15): 3701 [DOI: 10.3390/rs14153701]
- Han Y X. 2023. *Research on Optical and SAR Image Matching Based on OS-HOG*. Beijing: Chinese Academy of Surveying and Mapping (韩雅鑫. 2023. 基于OS-HOG的光学与SAR影像匹配研究. 北京: 中国测绘科学研究院) [DOI: 10.27481/d.cnki.gzcky.2023.000020]
- He C, Tu M X, Liu X L, Xiong D H and Liao M S. 2017. Mixture statistical distribution based multiple component model for target detection in high resolution SAR imagery. *ISPRS International Journal of Geo-Information*, 6(11): 336 [DOI: 10.3390/ijgi6110336]
- He X, Zhou Y, Liu B, Zhao J Q and Yao R. 2024. Remote sensing image semantic segmentation via class-guided structural interaction and boundary perception. *Expert Systems with Applications*, 252: 124019 [DOI: 10.1016/j.eswa.2024.124019]
- Hou J J, Zhou H K, Yu H M and Hu H J. 2024. RCSFN: a remote sensing image scene classification and recognition network based on rectangle convolutional self attention fusion. *Signal, Image and Video Processing*, 18(12): 8739-8756 [DOI: 10.1007/s11760-024-03511-8]
- Hu Q Q, Wang F T, Fang J T and Li Y. 2024. Semantic labeling of high-resolution images combining a self-cascaded multimodal fully convolution neural network with fully conditional random

- field. *Remote Sensing*, 16(17): 3300 [DOI: 10.3390/rs16173300]
- Huang T, Cao H Y and Xiong L Y. 2025. Deep learning delineates alluvial fans driven by topographic knowledge and imagery. *CATE-NA*, 256: 109089 [DOI: 10.1016/j.catena.2025.109089]
- Huang Z L, Yao X W, Liu Y, Dumitru C O, Datcu M and Han J W. 2022. Physically explainable CNN for SAR image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 190: 25-37 [DOI: 10.1016/j.isprsjprs.2022.05.008]
- Hui Y, You S J, Hu X H, Yang P P and Zhao J. 2024. SEB-YOLO: an improved YOLOv5 model for remote sensing small target detection. *Sensors*, 24(7): 2193 [DOI: 10.3390/s24072193]
- Iqbal N, Mumtaz R, Shafi U and Zaidi S M H. 2021. Gray level co-occurrence matrix (GLCM) texture based crop classification using low altitude remote sensing platforms. *PeerJ Computer Science*, 7: e536 [DOI: 10.7717/peerj-cs.536]
- Kas M, Khadiri I E, Merabet Y E, Ruichek Y and Messoussi R. 2020. Multi level directional cross binary patterns: new handcrafted descriptor for SVM-based texture classification. *Engineering Applications of Artificial Intelligence*, 94: 103743 [DOI: 10.1016/j.engappai.2020.103743]
- Khan J, Friedman A, Evans S, Wang R Z, Beins K, Andrews D and Bondi-Kelly E. 2025. FOCUS on contamination: a geospatial deep learning framework with a noise-aware loss for surface water PFAS prediction. [2025-05-08]. <https://arxiv.org/abs/2502.14894v1>
- Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//*Proceedings of the 26th International Conference on Neural Information Processing Systems*. Lake Tahoe: Curran Associates Inc.: 1097-1105
- Li D Z, Lu Y J, Wu J P, Zhou W L and Zeng G J. 2024a. Causal reinforcement learning for knowledge graph reasoning. *Applied Sciences*, 14(6): 2498 [DOI: 10.3390/app14062498]
- Li H C and Ma H. 2024. A rotated object detection strategy for remote sensing images using misaligned cross-fusion structures. *Applied Soft Computing*, 166: 112181 [DOI: 10.1016/j.asoc.2024.112181]
- Li H D and Qu H C. 2025. DASSF: dynamic-attention scale-sequence fusion for aerial object detection//*Proceedings of the 13th International Conference on Computational Visual Media*. Hong Kong: Springer: 212-227 [DOI: 10.1007/978-981-96-5809-1_12]
- Li H F, Li Q J, Yang C, Guo W, Li M Y, Yuan H Y, Shao R and Peng C L. 2025a. Task knowledge injection: training-free adaptation of multimodal large language models for remote sensing image understanding. *IEEE Geoscience and Remote Sensing Letters*, 22: 6010605 [DOI: 10.1109/LGRS.2025.3581558]
- Li J J, Tian P H, Song R, Xu H T, Li Y S and Du Q. 2024b. PCViT: a pyramid convolutional vision transformer detector for object detection in remote-sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5608115 [DOI: 10.1109/TGRS.2024.3360456]
- Li J S, Jin C H, Shen Y and Ye W X. 2024c. TRSAnet: a remote sensing deep learning model for water body change detection based on time-reversal semantic asymmetry. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 4210613 [DOI: 10.1109/TGRS.2024.3458951]
- Li M, Huo W B, Wu J J and Yang J Y. 2024d. SAR image reconstruction method for target detection using self-attention CNN-based deep prior learning. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5222714 [DOI: 10.1109/TGRS.2024.3461840]
- Li M Q, Dai W, Fan M T, Qian W, Yang X, Tao Y and Zhao C Y. 2023. Combining deep learning and hydrological analysis for identifying check dam systems from remote sensing images and DEMs in the Yellow River Basin. *International Journal of Environmental Research and Public Health*, 20(5): 4636 [DOI: 10.3390/ijerph20054636]
- Li M Z, Chen Y L, Zhang T and Huang W. 2024e. TA-YOLO: a lightweight small object detection model based on multi-dimensional trans-attention module for remote sensing images. *Complex and Intelligent Systems*, 10(4): 5459-5473 [DOI: 10.1007/s40747-024-01448-6]
- Li P, Hu X W, Feng C Q and Feng W K. 2025b. Physics-guided interpretable CNN for SAR target recognition. *Chinese Journal of Aeronautics*, 38(5): 103253 [DOI: 10.1016/j.cja.2024.09.029]
- Li W Y, Chen K Y, Chen H and Shi Z W. 2022. Geographical knowledge-driven representation learning for remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5405516 [DOI: 10.1109/TGRS.2021.3115569]
- Li Y S, Wu K, Ouyang S, Yang K, Li H P and Zhang Y J. 2024. Geographic knowledge graph guided semantic segmentation of remote sensing images. *National Remote Sensing Bulletin*, 28(2): 455-469 (李彦胜, 武康, 欧阳松, 杨坤, 李和平, 张永军. 2024. 地学知识图谱引导的遥感影像语义分割. *遥感学报*, 28(2): 455-469) [DOI: 10.11834/jrs.20231110]
- Li Z, Gao L R, Zheng K and Ni L. 2024. Research progress of high-resolution remote sensing image scene classification. *National Remote Sensing Bulletin*, 28(11): 2739-2760 (李智, 高连如, 郑珂, 倪丽. 2024. 高分辨率遥感图像场景分类研究进展. *遥感学报*, 28(11): 2739-2760) [DOI: 10.11834/jrs.20243519]
- Liang F, Zhang R X, Chai Y T, Chen J Y, Ru G B and Yang W. 2023. A sea-land segmentation method for SAR images using context-aware and edge attention based CNNs. *Geomatics and Information Science of Wuhan University*, 48(8): 1286-1295 (梁烽, 张瑞祥, 柴英特, 陈金勇, 茹国宝, 杨文. 2023. 一种结合上下文与边缘注意力的SAR图像海陆分割深度网络方法. *武汉大学学报(信息科学版)*, 48(8): 1286-1295) [DOI: 10.13203/j.whugis20210078]
- Lihe H Z Y, He J, Yuan Q Q, Jin X Y, Xiao Y and Zhang L P. 2024. PhDnet: a novel physic-aware dehazing network for remote sensing images. *Information Fusion*, 106: 102277102277 [DOI: 10.1016/j.inffus.2024.102277]
- Lin J J, Wang G J, Peng D F and Guan H Y. 2024. Edge-guided multi-scale foreground attention network for change detection in high resolution remote sensing images. *International Journal of Applied Earth Observation and Geoinformation*, 133: 104070 [DOI: 10.1016/j.jag.2024.104070]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017. Feature pyramid networks for object detection//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*

- dition (CVPR). Honolulu: IEEE: 936-944 [DOI: 10.1109/CVPR.2017.106]
- Lin Y and Guo J F. 2024. Fuzzy geospatial objects-based wetland remote sensing image Classification: a case study of Tianjin Binhai New area. *International Journal of Applied Earth Observation and Geoinformation*, 132: 104051 [DOI: 10.1016/j.jag.2024.104051]
- Liu B, Du S H, Bai L B, OuYang S, Wang H Y and Zhang X Y. 2023. Water extraction from optical high-resolution remote sensing imagery: a multi-scale feature extraction network with contrastive learning. *GIScience and Remote Sensing*, 60(1): 2166396 [DOI: 10.1080/15481603.2023.2166396]
- Liu B, Li B, Sreeram V and Li S F. 2024a. MBT-UNet: multi-branch transform combined with UNet for semantic segmentation of remote sensing images. *Remote Sensing*, 16(15): 2776 [DOI: 10.3390/rs16152776]
- Liu G H, Shan Z, Meng Y B, Akbar T A and Ye S T. 2024b. RDPGNet: a road extraction network with dual-view information perception based on GCN. *Journal of King Saud University - Computer and Information Sciences*, 36(3): 102009 [DOI: 10.1016/j.jksuci.2024.102009]
- Liu G H and Yang J Y. 2023. Exploiting deep textures for image retrieval. *International Journal of Machine Learning and Cybernetics*, 14(2): 483-494 [DOI: 10.1007/s13042-022-01645-0]
- Liu H, He Z W, Zhao Y B, Teng J and Wang Y. 2017. Improved ROEWA edge detector for SAR images. *Journal of Remote Sensing (in Chinese)*, 21(2): 273-279 (刘夯, 何政伟, 赵银兵, 滕靖, 王勇). 2017. SAR 图像 ROEWA 边缘检测器的改进. *遥感学报*, 21(2): 273-279 [DOI: 10.11834/jrs.20176045]
- Liu L X, Li B W, Wang Y P and Yang J Y. 2019. Remote sensing image segmentation based on improved Canny edge detection. *Computer Engineering and Applications*, 55(12): 54-58, 180 (刘丽霞, 李宝文, 王阳萍, 杨景玉). 2019. 改进 Canny 边缘检测的遥感影像分割. *计算机工程与应用*, 55(12): 54-58, 180 [DOI: 10.3778/j.issn.1002-8331.1811-0180]
- Liu L Y, Fu L, Zhang Y F, Ni W X, Wu B, Li Y, Shang C J and Shen Q. 2024c. CLFR-Det: cross-level feature refinement detector for tiny-ship detection in SAR images. *Knowledge-Based Systems*, 284: 111284 [DOI: 10.1016/j.knsys.2023.111284]
- Liu M, Hou B, Ren B, Jiao L C, Yang Z and Zhu Z W. 2024d. Subpixel feature pyramid network for multiscale ship detection in synthetic aperture radar remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 15583-15595 [DOI: 10.1109/JSTARS.2024.3452680]
- Liu P, Gao Y, Zheng X T, Wang H S, Zhao Y M, Wu X R, Lu Z H, Yue Z C, Xie Y T and Hao S F. 2025. Integrating attention mechanism and boundary detection for building segmentation from remote sensing images. *Frontiers in Neurorobotics*, 18: 1482051 [DOI: 10.3389/fnbot.2024.1482051]
- Liu Q H, Kampffmeyer M, Jenssen R and Salberg A B. 2021a. Self-constructing graph neural networks to model long-range pixel dependencies for semantic segmentation of remote sensing images. *International Journal of Remote Sensing*, 42(16): 6184-6208 [DOI: 10.1080/01431161.2021.1936267]
- Liu R, Ling J and Zhang H S. 2024e. SoftFormer: SAR-optical fusion transformer for urban land use and land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218: 277-293 [DOI: 10.1016/j.isprsjprs.2024.09.012]
- Liu X M, Li J B, Pan J S and Wang S. 2021b. An advanced gradient texture feature descriptor based on phase information for infrared and visible image matching. *Multimedia Tools and Applications*, 80(11): 16491-16511 [DOI: 10.1007/s11042-020-10213-z]
- Liu Y and Li Y T. 2024. Adaptive FAST corner detection optimization algorithm based on grayscale mean value. *Electronics Optics and Control*, 31(2): 65-71, 91 (刘艳, 李一桐). 2024. 基于灰度均值的自适应 FAST 角点检测优化算法. *光电与控制*, 31(2): 65-71, 91 [DOI: 10.3969/j.issn.1671-637X.2024.02.010]
- Liu Z Y, Wang K X, Dong H and Wang Y. 2021c. A cross-modal edge-guided salient object detection for RGB-D image. *Neurocomputing*, 454: 168-177 [DOI: 10.1016/j.neucom.2021.05.013]
- Long L H, Zhu Y T, Yan J W, Liu J J and Wang Z Y. 2023. New building extraction method based on semantic segmentation. *National Remote Sensing Bulletin*, 27(11): 2593-2602 (龙丽红, 朱宇霆, 闫敬文, 刘敬瑾, 王宗跃). 2023. 新型语义分割 D-UNet 的建筑物提取. *遥感学报*, 27(11): 2593-2602 [DOI: 10.11834/jrs.20211029]
- Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2): 91-110 [DOI: 10.1023/B:VISI.0000029664.99615.94]
- Lu C, Zhang X, Du K L, Xu H and Liu G C. 2024. CTCFNet: CNN-transformer complementary and fusion network for high-resolution remote sensing image semantic segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5408717 [DOI: 10.1109/TGRS.2024.3458446]
- Lu J, Li J, Fu H K, Zou W L, Kang J R, Yu H W and Lin X L. 2025. Estimation of rice yield using multi-source remote sensing data combined with crop growth model and deep learning algorithm. *Agricultural and Forest Meteorology*, 370: 110600 [DOI: 10.1016/j.agrformet.2025.110600]
- Lu Y T, Min L T, Wang B L, Zheng L, Wang X X, Zhao Y Q, Yang L and Long T. 2023. Cross-spatial pixel integration and cross-stage feature fusion-based transformer network for remote sensing image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5625616 [DOI: 10.1109/TGRS.2023.3334490]
- Luo R, He Q S, Zhao L J, Zhang S Q, Kuang G Y and Ji K F. 2024. Geospatial contextual prior-enabled knowledge reasoning framework for fine-grained aircraft detection in panoramic SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5226213 [DOI: 10.1109/TGRS.2024.3487780]
- Luo Y X, An D X, Wang W and Huang X T. 2020. Improved ROEWA SAR image edge detector based on curvilinear structures extraction. *IEEE Geoscience and Remote Sensing Letters*, 17(4): 631-635 [DOI: 10.1109/LGRS.2019.2926428]
- Ma X P, Zhang X K, Pun M O and Liu M. 2024a. A multilevel multimodal fusion transformer for remote sensing semantic segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5403215 [DOI: 10.1109/TGRS.2024.3373033]
- Ma X S, Li L and Wang G. 2024b. Blind edge-retention indicator for

- assessing the quality of filtered (Pol)SAR images based on a ratio gradient operator and confidence interval estimation. *Remote Sensing*, 16(11): 1992 [DOI: 10.3390/rs16111992]
- Mahapatra D K and Roy L P. 2019. Efficient modelling of SAR texture with a gamma-inverse gamma distribution for MAP-based speckle suppression. *IET Radar, Sonar and Navigation*, 13(6): 1018-1030 [DOI: 10.1049/iet-rsn.2018.5230]
- Mao W and Peng X W. 2019. WLIB-SIFT: a distinctive local image feature descriptor//Proceedings of 2019 IEEE 2nd International Conference on Information Communication and Signal Processing (ICICSP). Weihai: IEEE: 379-383 [DOI: 10.1109/ICICSP48821.2019.8958587]
- Miao M J, Huang H M, Huang K D and Wang S Q. 2025. A dual branch network combining detail information and color feature for remote sensing image dehazing. *International Journal of Machine Learning and Cybernetics*, 16(4): 2231-2247 [DOI: 10.1007/s13042-024-02388-w]
- Miao W, Huang K, Xu Z, Zhang J T, Geng J and Jiang W. 2024. Pseudo-label meta-learner in semi-supervised few-shot learning for remote sensing image scene classification. *Applied Intelligence*, 54(20): 9864-9880 [DOI: 10.1007/s10489-024-05670-0]
- Muzeau M, Frontera-Pons J, Ren C F and Ovarlez J P. 2024. SAFE: a SAR Feature Extractor based on self-supervised learning and masked Siamese ViTs. [2024-10-06]. <https://arxiv.org/abs/2407.00851>
- Nunes C F G and Pádua F L C. 2024. An orientation-robust local feature descriptor based on texture and phase congruency for visible - infrared image matching. *IEEE Geoscience and Remote Sensing Letters*, 21: 5002505 [DOI: 10.1109/LGRS.2024.3379091]
- Pan D C, Gao X, Dai W, Fu J M, Wang Z R, Sun X and Wu Y M. 2024a. SRT-Net: scattering region topology network for oriented ship detection in large-scale SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5202318 [DOI: 10.1109/TGRS.2024.3351366]
- Pan J H, He C, Huang W, Cao J D and Tong M. 2024b. Wavelet tree transformer: multihead attention with frequency-selective representation and interaction for remote sensing object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5637023 [DOI: 10.1109/TGRS.2024.3442575]
- Pan S M, Tao Y L, Nie C C and Chong Y W. 2021a. PEGNet: progressive edge guidance network for semantic segmentation of remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, 18(4): 637-641 [DOI: 10.1109/LGRS.2020.2983464]
- Pan X R, Ye T Z, Xia Z F, Song S J and Huang G. 2023. Slide-transformer: hierarchical vision transformer with local self-attention//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver: IEEE: 2082-2091 [DOI: 10.1109/CVPR52729.2023.00207]
- Pan Y, Meng Y H and Zhu L. 2021b. SAR image despeckling method based on improved Frost filtering. *Signal, Image and Video Processing*, 15(4): 843-850. [DOI: 10.1007/s11760-020-01805-1]
- Pang S Q, Ge J Y, Hu L, Guo K T, Zheng Y, Zheng C L, Zhang W and Liang J M. 2023. RTV-SIFT: harnessing structure information for robust optical and SAR image registration. *Remote Sensing*, 15(18): 4476 [DOI: 10.3390/rs15184476]
- Pang S Y, Hu H C, Zuo Z Q, Chen J and Hu X Y. 2024. Masked feature modeling for generative self-supervised representation learning of high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 8434-8449 [DOI: 10.1109/JSTARS.2024.3385420]
- Paul S and Pati U C. 2020. Automatic optical-to-SAR image registration using a structural descriptor. *IET Image Processing*, 14(1): 62-73 [DOI: 10.1049/iet-ipr.2019.0389]
- Pawar N S and Sharma K V. 2025. Comprehensive review of remote sensing integration with deep learning in landslide forecasting and future directions. *Natural Hazards*, 121(20): 23687-23721 [DOI: 10.1007/s11069-025-07171-w]
- Peng F F, Lu W, Tan W X, Qi K L, Zhang X K and Zhu Q S. 2022. Multi-output network combining GNN and CNN for remote sensing scene classification. *Remote Sensing*, 14(6): 1478 [DOI: 10.3390/rs14061478]
- Pham M T. 2018. Fusion of polarimetric features and structural gradient tensors for VHR PolSAR image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(10): 3732-3742 [DOI: 10.1109/JSTARS.2018.2868545]
- Qian Y T, Pu X Y, Jia H C, Wang H P and Xu F. 2024. ARNet: prior knowledge reasoning network for aircraft detection in remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5205214 [DOI: 10.1109/TGRS.2024.3359764]
- Qiao D, Yang J T, Bai B, Li G W, Wang J G, Li Z H, Liu J C and Liu J Y. 2024. Non-destructive monitoring of peanut leaf area index by combing UAV spectral and textural characteristics. *Remote Sensing*, 16(12): 2182 [DOI: 10.3390/rs16122182]
- Qin Q M. 2025. AlphaEarth foundations: the potential and challenges of remote sensing foundation models. *Journal of Geo-Information Science*, 27(10): 2593-2602 (秦其明. 2025. AlphaEarth foundations: 遥感基础大模型的潜力与挑战. *地球信息科学学报*, 27(11): 2593-2602) [DOI: 10.12082/dqxxkx.2025.250426]
- Rabie T, Baziyad M, Sani R, Bonny T and Fareh R. 2024. Color histogram contouring: a new training-less approach to object detection. *Electronics*, 13(13): 2522-2522 [DOI: 10.3390/electronics13132522]
- Ren J H, Shang R H, Chen J S, Zhang W T, Feng J, Liu M M, Wang C, Xu S H and Stolkin R. 2024. SAR image segmentation based on complicated region-sensitive adaptive superpixel generation and hybrid edge correction. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5229417 [DOI: 10.1109/TGRS.2024.3499374]
- Rey A A, Gambini J and Delrieux C A. 2021. Texture descriptors for robust SAR image segmentation. *Journal of Applied Remote Sensing*, 15(4): 046511 [DOI: 10.1117/1.JRS.15.046511]
- Sahu S R and Panda S. 2024. A novel depth-wise separable convolutional model for remote sensing scene classification. *Journal of the Indian Society of Remote Sensing*, 52(8): 1787-1803 [DOI: 10.1007/s12524-024-01904-3]
- Sebt M A and Darvishnezhad M. 2023. Feature fusion method based on local binary graph for PolSAR image classification. *IET Radar, Sonar and Navigation*, 17(6): 939-954 [DOI: 10.1049/rsn2.12389]

- Shao L, Wang X Q, Zhai W and Ding X. 2023. Research on building area extraction method of SAR image integrating ULBP and Gabor texture features. *Technology for Earthquake Disaster Prevention*, 18(1): 186-193 (邵乐, 王晓青, 翟玮, 丁香. 2023. 综合ULBP与Gabor纹理特征的SAR影像建筑区提取方法研究. *震灾防御技术*, 18(1): 186-193) [DOI: 10.11899/zzfy20230120]
- Sharma R and Panigrahi R K. 2021. Texture classification-based NLM PolSAR filter. *IEEE Geoscience and Remote Sensing Letters*, 18(8): 1396-1400 [DOI: 10.1109/LGRS.2020.2998959]
- Shen J M, Bai L, Zhang Y Q, Momi M C, Quan S W and Ye Z. 2024. ELLK-net: an efficient lightweight large kernel network for SAR ship detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5221514 [DOI: 10.1109/TGRS.2024.3451399]
- Shen X L, Guo Y Q and Cao J Z. 2023. Object-based multiscale segmentation incorporating texture and edge features of high-resolution remote sensing images. *PeerJ Computer Science*, 9: e1290 [DOI: 10.7717/peerj-cs.1290]
- Shen Y, Lin W F, Wang Z F, Li J, Sun X Q, Wu X Y, Wang S and Huang F. 2021. Rapid detection of camouflaged artificial target based on polarization imaging and deep learning. *IEEE Photonics Journal*, 13(4): 7800309 [DOI: 10.1109/JPHOT.2021.3103866]
- Shi C P, Shi K J, Zeng Z X and Zhu F. 2025a. Dynamic feature enhancement network guided by multi-dimensional collaborative edge information for remote sensing image compression. *Knowledge-Based Systems*, 310: 112996 [DOI: 10.1016/j.knosys.2025.112996]
- Shi C P, Shi K J, Zhu F, Zeng Z X, Ding M X and Jin Z. 2024. An enhanced global feature-guided network based on multiple filtering noise reduction for remote sensing image compression. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5646220 [DOI: 10.1109/TGRS.2024.3483871]
- Shi L and Zhao Y F. 2023. Edge detection of high-resolution remote sensing image based on multi-directional improved sobel operator. *IEEE Access*, 11: 135979-135993 [DOI: 10.1109/ACCESS.2023.3338355]
- Shi S H, Fang Q, Xu X and Dong D Z. 2025b. Multiscale Gaussian attention mechanism for tiny-object detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 5635216 [DOI: 10.1109/TGRS.2025.3591260]
- Shin Y, Shin H, Ok J, Back M, Youn J and Kim S. 2024. DCEF2-YOLO: aerial detection YOLO with deformable convolution - efficient feature fusion for small target detection. *Remote Sensing*, 16(6): 1071 [DOI: 10.3390/rs16061071]
- Soldal I H, Dierking W, Korosov A and Marino A. 2019. Automatic detection of small icebergs in fast ice using satellite wide-swath SAR images. *Remote Sensing*, 11(7): 806 [DOI: 10.3390/rs11070806]
- Song X Y, Hua Z and Li J J. 2023. GMTS: GNN-based multi-scale transformer Siamese network for remote sensing building change detection. *International Journal of Digital Earth*, 16(1): 1685-1706 [DOI: 10.1080/17538947.2023.2210311]
- Song X Y, Hua Z and Li J J. 2024a. Context spatial awareness remote sensing image change detection network based on graph and convolution interaction. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 3000316 [DOI: 10.1109/TGRS.2024.3357524]
- Song Y P, Li C L, Xiao S Y, Zhou Q L and Xiao H. 2024b. A parallel Canny edge detection algorithm based on OpenCL acceleration. *PLoS One*, 19(1): e0292345 [DOI: 10.1371/journal.pone.0292345]
- Sun H K, Chen Y X, Lu X B and Xiong S W. 2023. Decoupled feature pyramid learning for multi-scale object detection in low-altitude remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16: 6556-6567 [DOI: 10.1109/JSTARS.2023.3293395]
- Sun M C, Ma T X, Song Y M and Peng J Q. 2021. Automatic registration of optical and SAR remote sensing image based on phase feature. *Optics and Precision Engineering*, 29(3): 616-627 (孙明超, 马天翔, 宋悦铭, 彭佳琦. 2021. 基于相位特征的可见光和SAR遥感图像自动配准. *光学精密工程*, 29(3): 616-627) [DOI: 10.37188/OPE.20212903.0616]
- Sun S S, Zhang L J and Zhao H. 2024. SAR ship detection model based on edge enhancement and attention mechanism. *Electronics Optical and Control*, 31(8): 92-97, 110 (孙珊珊, 张丽娟, 赵辉. 2024. 基于边缘增强与注意力机制的SAR舰船检测模型. *电光与控制*, 31(8): 92-97, 110) [DOI: 10.3969/j.issn.1671-637X.2024.08.015]
- Tang F Y, Li Z F, Zhang Q J, Suo Z Y, Zhang Z X, Xing C and Guo H C. 2023. Polarimetric synthetic aperture radar speckle filter based on joint similarity measurement criterion. *Remote Sensing*, 15(21): 5224 [DOI: 10.3390/rs15215224]
- Tang G L, Wei Z H and Zhuang L. 2024. SAR image registration: the combination of nonlinear diffusion filtering, hessian features and edge points. *Sensors*, 24(14): 4568 [DOI: 10.3390/s24144568]
- Tang X, Zhang H Y, Mou L C, Liu F, Zhang X R, Zhu X X and Jiao L C. 2022. An unsupervised remote sensing change detection method based on multiscale graph convolutional network and metric learning. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5609715 [DOI: 10.1109/TGRS.2021.3106381]
- Tao J H, Bie Y X, Gu Y H, Wang H and Zhou W. 2021. Research on support vector machines image classification algorithm based on multi-feature fusion. *Aerospace Shanghai (Chinese and English)*, 38(S1): 98-102 (陶佳慧, 别雨轩, 顾约翰, 王辉, 周维. 2021. 基于多特征融合的SVM图像分类算法研究. *上海航天(中英文)*, 38(S1): 98-102) [DOI: 10.19328/j.cnki.2096-8655.2021.s1.017]
- Vimina E R and Jacob K P. 2019. Feature fusion method using BoVW framework for enhancing image retrieval. *IET Image Processing*, 13(11): 1979-1985 [DOI: 10.1049/iet-ipr.2018.5381]
- Wan Q, Xiao Z F, Yu Y, Liu Z Q, Wang K and Li D R. 2024. A hyperparameter-free attention module based on feature map mathematical calculation for remote-sensing image scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5600318 [DOI: 10.1109/TGRS.2023.3335627]
- Wang C, Zhu F Z, Zhu B, Zhang Q and Ma H B. 2024a. Reference-based super-resolution reconstruction of remote sensing images based on a coarse-to-fine feature matching transformer. *Engineering Applications of Artificial Intelligence*, 135: 108787 [DOI: 10.1016/j.engappai.2024.108787]
- Wang H, Ouyang S, Liu Q H, Liao K F and Zhou L J. 2022a. Deep-

- learning-based method for estimating permittivity of ground-penetrating radar targets. *Remote Sensing*, 14(17): 4293 [DOI: 10.3390/rs14174293]
- Wang L N. 2021. Research on the registration technology of optical and SAR remote sensing images. Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences) (王丽娜. 2021. 可见光与 SAR 遥感图像配准技术研究. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所)) [DOI: 10.27522/d.cnki.gkcg.2021.000015]
- Wang L N, Liang H D, Wang Z S, Xu R and Shi G F. 2022. Feature point detection for optical and SAR remote sensing images registration. *Optics and Precision Engineering*, 30(14): 1738-1748 (王丽娜, 梁怀丹, 王中石, 徐瑞, 石广丰. 2022. 面向可见光和 SAR 影像配准的特征点检测. *光学精密工程*, 30(14): 1738-1748) [DOI: 10.37188/OPE.20223014.1738]
- Wang M M, Zhang J X, Deng K Z and Hua F F. 2022b. Combining optimized SAR-SIFT features and RD model for multisource SAR image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5206916 [DOI: 10.1109/TGRS.2021.3074630]
- Wang M W, Gao L, Huang X H, Jiang Y and Gao X J. 2019. A texture classification approach based on the integrated optimization for parameters and features of Gabor filter via hybrid ant lion optimizer. *Applied Sciences*, 9(11): 2173 [DOI: 10.3390/app9112173]
- Wang M Y, Li X, Tan K, Mango J, Pan C and Zhang D. 2024b. Position-aware graph-CNN fusion network: an integrated approach combining geospatial information and graph attention network for multiclass change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 4402016 [DOI: 10.1109/TGRS.2024.3350573]
- Wang P, Luo Y and Zhu Z L. 2025a. FDI-YOLO: feature disentanglement and interaction network based on YOLO for SAR object detection. *Expert Systems with Applications*, 260: 125442 [DOI: 10.1016/j.eswa.2024.125442]
- Wang Q, Feng W Q, Yao L F, Zhuang C, Liu B H and Chen L J. 2023a. TPH-YOLOv5-air: airport confusing object detection via adaptively spatial feature fusion. *Remote Sensing*, 15(15): 3883 [DOI: 10.3390/rs15153883]
- Wang R F, Wang L, Wei X H, Chen J W and Jiao L C. 2022c. Dynamic graph-level neural network for SAR image change detection. *IEEE Geoscience and Remote Sensing Letters*, 19: 4501005 [DOI: 10.1109/LGRS.2021.3132167]
- Wang S, Han W, Huang X H, Zhang X H, Wang L Z and Li J. 2024c. Trustworthy remote sensing interpretation: concepts, technologies, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 209: 150-172 [DOI: 10.1016/j.isprsjprs.2024.02.003]
- Wang S L, Hu Q W, Wang S H, Zhao P C, Li J Y and Ai M Y. 2024d. Category attention guided network for semantic segmentation of Fine-Resolution remote sensing images. *International Journal of Applied Earth Observation and Geoinformation*, 127: 103661 [DOI: 10.1016/j.jag.2024.103661]
- Wang W, Liu C, Liu G Q and Wang X. 2024e. CF-GCN: graph convolutional network for change detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5607013 [DOI: 10.1109/TGRS.2024.3357085]
- Wang W N, Ran L Y, Yin H L, Sun M J, Zhang X W and Zhang Y N. 2024f. Hierarchical shared architecture search for real-time semantic segmentation of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 1-13 [DOI: 10.1109/TGRS.2024.3373493]
- Wang Y, Shao Z F, Lu T, Huang X, Wang J M, Zhang Z Z and Zuo X L. 2025b. Lightweight remote sensing super-resolution with multi-scale graph attention network. *Pattern Recognition*, 160: 111178 [DOI: 10.1016/j.patcog.2024.111178]
- Wang Y, Zhao Y W, Wu L F, Yin X J, Zhou H G, Hu J and Nie Z P. 2023b. An early fusion deep learning framework for solving electromagnetic inverse scattering problems. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 2005914 [DOI: 10.1109/TGRS.2023.3337410]
- Wei C, Duke B, Jiang R W, Aarabi P, Taylor G W and Shkurti F. 2023. Sparsifiner: learning sparse instance-dependent attention for efficient vision transformers//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver: IEEE: 22680-22689 [DOI: 10.1109/CVPR52729.2023.02172]
- Wei H, Gao K, Wang J, Tang Q Y, Tang X X and Xu F J. 2024a. Optical imaging degradation simulation and transformer-based image restoration for remote sensing. *IEEE Geoscience and Remote Sensing Letters*, 21: 6006205 [DOI: 10.1109/LGRS.2024.3381581]
- Wei J J, Sun K M, Li W Z, Li W B, Gao S, Miao S X, Zhou Q H and Liu J Y. 2024b. Robust change detection for remote sensing images based on temporospatial interactive attention module. *International Journal of Applied Earth Observation and Geoinformation*, 128: 103767 [DOI: 10.1016/j.jag.2024.103767]
- Wei Q, Miao T Y, Yao H T, Jiao S M, Jin Y, Hu Z T and Peng Q L. 2025. Semantic segmentation for remote sensing images using multi-layer semantic Markov random field with semantic text information. *Digital Signal Processing*, 162: 105174 [DOI: 10.1016/j.dsp.2025.105174]
- Wei Q R, Wang Y K and Xie P Y. 2019. SAR edge detector with high localization accuracy//Proceedings of IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. Yokohama, Japan: IEEE, 3744-3747 [DOI: 10.1109/IGARSS.2019.8899205]
- Wieland M, Schmidt S, Resch B, Abecker A and Martinis S. 2025. Fusion of geospatial information from remote sensing and social media to prioritise rapid response actions in case of floods. *Natural Hazards*, 121(7): 8061-8088 [DOI: 10.1007/s11069-025-07120-7]
- Wu S L, Zeng W K and Chen H D. 2020. A sub-pixel image registration algorithm based on SURF and M-estimator sample consensus. *Pattern Recognition Letters*, 140: 261-266 [DOI: 10.1016/j.patrec.2020.09.031]
- Wu X Y, Wen X B, Xu H X, Yuan L M and Guo C L. 2021. Multi-feature fusion and adaptive kernel combination for SAR image classification. *Applied Sciences*, 11(4): 1603 [DOI: 10.3390/app11041603]
- Xiang Y M, Wang F, Wan L and You H J. 2017. An advanced rotation invariant descriptor for SAR image registration. *Remote Sensing*,

- 9(7): 686 [DOI: 10.3390/rs9070686]
- Xiao K, Zhao X Y, Ding Y, Huang C, Lin J Y, Mai Y X, Sun Y and Xin Q C. 2025a. Ultra-high spatial resolution mapping of urban forest canopy height with multimodal remote sensing data and deep learning method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 9865-9882 [DOI: 10.1109/JSTARS.2025.3545482]
- Xiao Y, Zhang C L, Jiang B, Chen Y and Tang J. 2025b. Multimodal remote sensing image registration via modality perception and self-supervised position estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 63: 5213414 [DOI: 10.1109/TGRS.2025.3576290]
- Xie Z H. 2022. Research on key technology of visible and SAR remote sensing image registration. University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences) (谢志华. 2022. 可见光与SAR遥感图像配准关键技术研究. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所)) [DOI: 10.27522/d.cnki.gkcg.2022.000043]
- Xu L Y, Li H F, Shen H F, Zhang C and Zhang L P. 2024. A thin cloud blind correction method coupling a physical model with unsupervised deep learning for remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 218: 246-259 [DOI: 10.1016/j.isprsjprs.2024.09.008]
- Xu Y J, Zhang Y J, Cui J R, Cui Z W and Yang Y T. 2022. A robust edge-guided and color-guided network for single image dehazing. *Computers and Electrical Engineering*, 103: 108301 [DOI: 10.1016/j.compeleceng.2022.108301]
- Xu Y T, Qi Q, He W, Zhang G Y, Chen S Y and Tu B. 2025. GSINet: gradual semantic interaction network for remote sensing object detection based on dual attention mechanism. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 11877-11891 [DOI: 10.1109/JSTARS.2025.3551551]
- Yan R X, Yan L Q, Geng G H, Cao Y F, Zhou P B and Meng Y L. 2024. ASNet: adaptive semantic network based on transformer - CNN for salient object detection in optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5608716 [DOI: 10.1109/TGRS.2024.3362836]
- Yan X T, Shi Z, Li P and Zhang Y F. 2023. IDCF: information distribution composite feature for multi-modal image registration. *International Journal of Remote Sensing*, 44(6): 1939-1975 [DOI: 10.1080/01431161.2023.2193300]
- Yang R, Xu X, Xu Z Z, Dong H, Gui R and Pu F L. 2019. Dynamic fractal texture analysis for PolSAR land cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8): 5991-6002 [DOI: 10.1109/TGRS.2019.2903794]
- Yang X and Mohamed A S A. 2025. Gaussian-based R-CNN with large selective kernel for rotated object detection in remote sensing images. *Neurocomputing*, 620: 129248 [DOI: 10.1016/j.neucom.2024.129248]
- Yao H T, Wang X P, Zhao L, Tian M, Jian Z N, Gong L and Li B W. 2022. An object-based Markov random field with partition-global alternately updated for semantic segmentation of high spatial resolution remote sensing image. *Remote Sensing*, 14(1): 127 [DOI: 10.3390/rs14010127]
- Yasir M, Liu S W, Xu M M, Wan J H, Nazir S, Ul Islam Q and Dang K B. 2024. SwinYOLOv7: robust ship detection in complex synthetic aperture radar images. *Applied Soft Computing*, 160: 111704 [DOI: 10.1016/j.asoc.2024.111704]
- Ye M L, Wang Q, Zhang C M, Qin S D and Yan S Y. 2025. A combined deep learning and morphology approach for DFS identification and parameter extraction. *Computers and Geosciences*, 196: 105856 [DOI: 10.1016/j.cageo.2025.105856]
- Yin J J, Liu X Y, Yang J, Chu C Y and Chang Y L. 2020. PolSAR image classification based on statistical distribution and MRF. *Remote Sensing*, 12(6): 1027 [DOI: 10.3390/rs12061027]
- Yu H F, Yao F L, Lu W X, Liu N Y, Li P G, You H J and Sun X. 2023. Text-image matching for cross-modal remote sensing image retrieval via graph neural network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16: 812-824 [DOI: 10.1109/JSTARS.2022.3231851]
- Yu L J, Zhi X Y, Hu J M, Zhang S Q, Niu R Z, Zhang W and Jiang S K. 2024a. Improved deformable convolution method for aircraft object detection in flight based on feature separation in remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 8313-8323 [DOI: 10.1109/JSTARS.2024.3386696]
- Yu L J, Zhi X Y, Zhang S Q, Jiang S K, Hu J M, Zhang W and Huang Y X. 2024b. A method for detecting aircraft small targets in remote sensing images by using CNNs fused with handcrafted features. *IEEE Geoscience and Remote Sensing Letters*, 21: 6010105 [DOI: 10.1109/LGRS.2024.3403548]
- Yu Q Z, Zhou S, Jiang Y X, Wu P and Xu Y. 2019. High-performance SAR image matching using improved SIFT framework based on rolling guidance filter and ROEWA-Powered feature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(3): 920-933 [DOI: 10.1109/JSTARS.2019.2897171]
- Yu W T, Zhuo L and Li J F. 2024c. GCFormer: global context-aware transformer for remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 4703212 [DOI: 10.1109/TGRS.2024.3381738]
- Yuan W and Xu W B. 2021. MSST-Net: a multi-scale adaptive network for building extraction from remote sensing images based on Swin transformer. *Remote Sensing*, 13(23): 4743 [DOI: 10.3390/rs13234743]
- Zhang H D and Wu Y Q. 2024. CSEF-Net: cross-scale SAR ship detection network based on efficient receptive field and Enhanced hierarchical fusion. *Remote Sensing*, 16(4): 622 [DOI: 10.3390/rs16040622]
- Zhang J H, Liu H Z, Zhang Y, Li M H and Zhan Z Q. 2024a. Faster transformer-DS: multiscale vehicle detection of remote-sensing images based on transformer and distance-scale loss. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 1961-1975 [DOI: 10.1109/JSTARS.2023.3335283]
- Zhang L P, Zhang L F and Du B. 2016. Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geosci-*

- ence and Remote Sensing Magazine, 4(2): 22-40 [DOI: 10.1109/MGRS.2016.2540798]
- Zhang S H and Pan Z G. 2025. Remote sensing large models: review and future prospects. *Remote Sensing Technology and Application*, 40(1): 1-13 (张帅豪, 潘志刚. 2025. 遥感大模型: 综述与未来设想. *遥感技术与应用*, 40(1): 1-13) [DOI: 10.11873/j.issn.1004-0323.2025.1.0001]
- Zhang W, Li T T, Zhang Y T, Pei G S, Jiang X R and Yao Y Z. 2024b. LTFormer: a light-weight transformer-based self-supervised matching network for heterogeneous remote sensing images. *Information Fusion*, 109: 102425 [DOI: 10.1016/j.inffus.2024.102425]
- Zhang W, Wang X S, Wang H Y and Cheng Y H. 2024c. Causal meta-reinforcement learning for multimodal remote sensing data classification. *Remote Sensing*, 16(6): 1055 [DOI: 10.3390/rs16061055]
- Zhang W N. 2025. Optical and SAR image registration based on feature constrained algorithm. *Scientific Reports*, 15(1): 36373 [DOI: 10.1038/s41598-025-20332-y]
- Zhang X D, Zhuang Y H, Guo Q, Yang H D, Qian X L, Cheng G, Han J W and Huang Z L. 2025. Φ -GAN: physics-inspired GAN for generating SAR images under limited data//*Proceedings of 2025 IEEE/CVF International Conference on Computer Vision*. Honolulu: IEEE: 29075-29085
- Zhang X Z, Xie F Y, Ding H D, Yan S C and Shi Z W. 2024d. Proxy and cross-stripes integration transformer for remote sensing image dehazing. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5640315 [DOI: 10.1109/TGRS.2024.3457868]
- Zhang Y J, Li Y S, Dang B, Wu K, Guo X, Wang J, Chen J D and Yang M. 2024. Multi-modal remote sensing large foundation models: current research status and future prospect. *Acta Geodaetica et Cartographica Sinica*, 53(10): 1942-1954 (张永军, 李彦胜, 党博, 武康, 郭昕, 王剑, 陈景东, 杨铭. 2024. 多模态遥感基础大模型: 研究现状与未来展望. *测绘学报*, 53(10): 1942-1954) [DOI: 10.11947/j.AGCS.2024.20240019]
- Zhao J and An D X. 2024. Improved Markov random field for building segmentation in SAR images. *Remote Sensing Technology and Application*, 39(2): 405-412 (赵佳, 安道祥. 2024. 改进马尔科夫随机场的SAR图像建筑物分割. *遥感技术与应用*, 39(2): 405-412) [DOI: 10.11873/j.issn.1004-0323.2024.2.0405]
- Zhao J X, Jiao L C, Wang C, Liu X, Liu F, Li L L, Ma M R and Yang S Y. 2024. Knowledge guided evolutionary transformer for remote sensing scene classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(10): 10368-10384 [DOI: 10.1109/TCSVT.2024.3407138]
- Zhao M, Zhang X and Kaup A. 2023a. Multitask learning for SAR ship detection with Gaussian-mask joint segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5214516 [DOI: 10.1109/tgrs.2023.3304847]
- Zhao X Y, Zhao K Y, Li S Y and Wang X H. 2023b. GeSAnet: geospatial-awareness network for VHR remote sensing image change detection. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5402814 [DOI: 10.1109/TGRS.2023.3272550]
- Zheng H, Hu Z G, Yang L, Xu A K, Zheng M G, Zhang C and Li K Q. 2023a. Multifeature collaborative fusion network with deep supervision for SAR ship classification. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 5212614 [DOI: 10.1109/TGRS.2023.3297648]
- Zheng J W, Shao A H, Yan Y D, Wu J and Zhang M Y. 2023b. Remote sensing semantic segmentation via boundary supervision-aided multiscale channelwise cross attention network. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 4405814 [DOI: 10.1109/TGRS.2023.3292112]
- Zheng S X, Lu J C, Zhao H H, Zhu X T, Luo Z K, Wang Y B, Fu Y W, Feng J F, Xiang T, Torr P H S and Zhang L. 2021. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE: 6877-6886 [DOI: 10.1109/CVPR46437.2021.00681]
- Zhong B and Li Y B. 2019. Image feature point matching based on improved SIFT algorithm//*Proceedings of 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC)*. Xiamen, China: IEEE: 489-493 [DOI: 10.1109/ICIVC47709.2019.8981329]
- Zhong R J, Fu Y, Song Y S and Han C X. 2023. A fusion approach to infrared and visible images with Gabor filter and sigmoid function. *Infrared Physics and Technology*, 131: 104696 [DOI: 10.1016/j.infrared.2023.104696]
- Zhou M, Zhou Y, Yang D W and Song K. 2024. Remote sensing image classification based on canny operator enhanced edge features. *Sensors*, 24(12): 3912 [DOI: 10.3390/s24123912]
- Zhou W J, Fan X M, Yan W Q, Shan S D, Jiang Q P and Hwang J N. 2023a. Graph attention guidance network with knowledge distillation for semantic segmentation of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 4506015 [DOI: 10.1109/TGRS.2023.3311480]
- Zhou Y, Han Z, Dou Z, Huang C B, Cong L, Lv N and Chen C. 2023b. Edge consistency feature extraction method for multi-source image registration. *Remote Sensing*, 15(20): 5051 [DOI: 10.3390/rs15205051]
- Zhu L, Wang X J, Ke Z H, Zhang W and Lau R. 2023. BiFormer: vision transformer with bi-level routing attention//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver: IEEE: 10323-10333 [DOI: 10.1109/CVPR52729.2023.00995]
- Zhu R. 2024. Geospatial knowledge graphs//Richardson D, Castree N, Goodchild M F, Kobayashi A L, Liu W D and Marston R A, eds. *International Encyclopedia of Geography: People, the Earth, Environment and Technology*. Chichester: John Wiley and Sons [DOI: 10.1002/9781118786352.wbieg2207]
- Zhu Z W, Shen Z F and Luo J C. 2011. Parallel remote sensing image registration based on improved SIFT point feature. *National Remote Sensing Bulletin*, 15(5): 1024-1039 (朱志文, 沈占锋, 骆剑承. 2011. 改进SIFT点特征的并行遥感影像配准. *遥感学报*, 15(5): 1024-1039) [DOI: 10.11834/jrs.20110216]
- Zi W J, Xiong W, Chen H, Li J and Jing N. 2021. SGA-Net: self-constructing graph attention neural network for semantic segmentation of remote sensing images. *Remote Sensing*, 13(21): 4201 [DOI: 10.3390/rs13214201]

Challenges, methods, and opportunities in feature extraction of visible light and SAR images

XU Lijun¹, YU Xiaogang², HE Haochen²

1. Graduate School, Aerospace Engineering University, Beijing 100400, China;

2. Beijing Institute of Remote Sensing Information, Beijing 100192, China

Abstract: Visible light and Synthetic Aperture Radar (SAR) images are two primary data sources in satellite remote sensing, widely used in fields such as urban planning, disaster prevention, and national security. Feature extraction serves as a key bridge connecting remote sensing images with high-level applications, directly influencing the effectiveness of intelligent interpretation in complex scenarios. Feature extraction methods have undergone a profound transformation from handcrafted design to data-driven learning, achieving significant breakthroughs, particularly with the advancement of deep learning. However, due to the high cost and strong domain specificity of remote sensing data annotation, previous data-driven methods still face challenges such as insufficient generalization and limited interpretability. To address these issues, existing researches have proposed a “data + knowledge” dual-driven framework, which leverages data-driven perception and knowledge-driven reasoning to achieve higher-level intelligence, aiming to break through the bottlenecks of generalization and interpretability inherent in single data-driven paradigms. To reveal the technical logic and core value of this dual-driven framework in feature extraction for visible light and SAR images, this paper first systematically reviews traditional feature extraction methods and deep learning-based methods, summarizing the main challenges faced by these two categories. It then elaborates on knowledge-driven feature extraction, categorizing the knowledge used to guide networks into visual knowledge, geospatial knowledge, and physical knowledge, and discusses the training paradigms for knowledge-guided networks. Finally, it outlines future trends, identifying the construction of remote sensing foundation models, the dual-driven development of data and knowledge, and the cross-integration of cutting-edge technologies as key research directions.

This study follows a structured review methodology, focusing on literature from authoritative sources such as IEEE TGRS, ISPRS Journal, CVPR, and ICCV between 2019 and 2025. First, traditional feature extraction methods—including edge detection (e.g., Sobel, Canny, ROEWA), texture analysis (e.g., Gabor filters, GLCM), color features, and keypoint extraction (e.g., SIFT, Harris)—are summarized and analyzed for their strengths and limitations. Second, deep learning-based methods are examined, covering convolutional neural networks for local spatial modeling, graph neural networks for relational reasoning, and vision transformers for global context understanding. Third, a knowledge-driven framework is proposed to address challenges such as semantic ambiguity, scale variation, imaging condition sensitivity, and speckle noise interference. This framework integrates three categories of knowledge—visual knowledge (e.g., edges, textures), geospatial knowledge (e.g., DEM, topological relations), and physical knowledge (e.g., scattering models, atmospheric physics)—into deep networks through mechanisms such as knowledge as architecture, knowledge as data, and knowledge as loss.

The analysis demonstrates that traditional methods offer high interpretability and computational efficiency but lack adaptability to complex scenes and nonlinear relationships. Deep learning methods significantly improve feature representation and task performance but are heavily dependent on large annotated datasets and suffer from poor generalization under domain shifts. The proposed knowledge-driven approach effectively mitigates these issues. For example, incorporating edge prior knowledge enhances boundary precision in semantic segmentation, embedding elevation data improves terrain-aware classification, and integrating physical models such as electromagnetic scattering models aids in SAR image interpretation and speckle suppression. Experimental results from cited studies show performance gains across various tasks, such as increased mIoU in segmentation and enhanced accuracy in object detection, confirming that knowledge-guided features are discriminative, stable, and interpretable.

Feature extraction for visible light and SAR images has evolved from manual design to data-driven learning and is now progressing toward a hybrid paradigm that combines data-driven perception with knowledge-driven reasoning. This study highlights that integrating visual, geospatial, and physical knowledge into deep learning models can address key challenges in remote sensing feature extraction, including semantic ambiguity, scale extremes, imaging variability, and noise interference. Future research should focus on developing large-scale remote sensing foundation models, deepening the synergy between data and knowledge, and promoting cross-disciplinary integration with emerging technologies like vision-language and diffusion models. Such advancements will further enhance the intelligence, reliability, and applicability of remote sensing image interpretation systems.

Key words: remote sensing intelligent interpretation, visible light images, SAR images, manual feature extraction, depth feature extraction, data and knowledge driven