

轻量化类火星地貌快速识别分类方法研究

李建龙¹, 张立华^{2,3}, 何梨梨¹, 刘欣鑫¹, 杨彬¹, 周继时^{2,3}

(1. 湖南大学 电气与信息工程学院, 长沙 410082; 2. 国家航天局 探月与航天工程中心, 北京 100190; 3. 深空探测实验室, 北京 100190)

摘要: 类火星地貌识别分类旨在通过分析类火星地貌图像, 模拟研究火星环境, 对火星形成演化及潜在宜居环境探索等科学问题具有重要研究意义。针对目前火星地貌分类算法难以平衡模型分类效果和轻量化程度的问题, 提出一种轻量化类火星地貌快速识别分类方法 (Lightweight Rapid Recognition and Classification for Mars Analog Terrain, LWNet), 构建双分支教师-学生网络, 利用知识蒸馏技术减少模型参数量与计算量, 并嵌入注意力机制提高对地貌类型的分类识别能力, 实现分类模型的高精度和轻量化。为验证所提方法的分类性能表现, 通过采集地球上相似地貌, 构建了包括悬崖断面、沙漠、河道及雅丹在内的4种典型类火星地貌数据集, 每种地貌类型各获取800张图像, 用于LWNet开展快速识别分类实验。结果表明, LWNet总体分类精度达到97.81%, 相较于精度最高的Swin-Transformer仅下降了1.25%, 而其参数量和计算量却只有Swin-Transformer的1.3%和4.8%, 验证了所提轻量化类火星地貌快速识别分类方法的有效性和优越性。

关键词: 类火星地貌; 轻量化; 知识蒸馏; 图像分类

中图分类号: P185.3

文献标识码: A

文章编号: 2096-9287(2025)01-0077-09

DOI: 10.15982/j.issn.2096-9287.2025.20240055

引用格式: 李建龙, 张立华, 何梨梨, 等. 轻量化类火星地貌快速识别分类方法研究[J]. 深空探测学报(中英文), 2025, 12(1): 77-85.

Reference format: LI J L, ZHANG L H, HE L L, et al. Lightweight rapid recognition and classification for Mars analog terrain[J]. Journal of Deep Space Exploration, 2025, 12(1): 77-85.

引言

随着遥感技术和航天技术的飞速发展, 人类对深空探测的发展脚步显著加快。火星作为太阳系中与地球最为相似和距离最近的行星, 是当前及未来很长一段时间内深空探测的主要研究对象。目前对火星地貌图像的研究解译需要将其从卫星传感器传送到地表再进行处理, 应用成本大且处理效率低下。人工智能芯片和算法的发展, 为实现星上数据的在线处理和信息提取提供了可能, 通过在遥感卫星上搭载人工智能微系统以及轻量化的智能识别算法, 可以实现高精度的火星地貌图像识别。然而, 由于距离遥远, 即使在地面建设了相应的大口径数据接收天线(中国最大口径数据接收天线直径已达70 m), 仍然很难满足智能识别算法研究的海量高质量图像数据需求, 因此火星地貌图像作为开展火星科学研究最基础、最直观和最可靠的数据, 具有重要的科学价值。科学家们已经开始在地球上寻找与火星地貌相似的区域, 试图利用类火星地貌模拟研究火星环境。场景分类是一种根据图像特征信息对图像进行解译、分析、处理的前沿技术, 利用场景分类技术研究解译类火星地貌图像对揭示火

星地表覆盖与变化、研究火星大气与气候演变等深层次问题具有重要科学意义^[1]。

早期行星地貌图像实现分类的算法主要基于地表高程、地形起伏及坡度等可用于地表形态参数化的特征, 利用聚类法对地形地貌实现分类。例如, Bue和Stepinski根据海拔、坡度等6种数字地形参数, 使用自组织映射法和Ward聚类法将火星地貌划分为高地、陨石坑、低地、高起伏及河道几种类别地形^[2]。Wang等^[3]在地形参数中加入地形起伏, 并使用ISO聚类取代Ward聚类, 成功对月球地貌进行划分, 并将分类精度和Kappa系数分别提高至83.34%和0.77。尽管这些方法在早期行星地貌分类任务上表现出了不俗的性能, 在应对复杂场景时, 有限的特征提取和表征能力严重限制了它们的性能表现, 实际难以达到预期效果。

为解决上述问题, 充分挖掘图像数据信息, 增强分类算法的特征提取和表征能力, KNN^[4]、Decision Tree^[5]、SVM^[6]、RF^[7]等机器学习方法被引入到火星地貌分类领域中。Shang和Barnes对模糊粗糙特征选择(Fuzzy-rough feature selection, FRFS)进行了调整^[8], 然后将其与SVM结合构建图像分类器对单幅火星图像中的地貌类型进行分类, 同时Shang还研究了基于

KNN和Decision Tree的图像分类器,均取得了不错的效果。然而机器学习方法受参数选择影响大,参数选取的好坏直接影响到模型的精度,实际应用时泛化性能差。近年来,由于深度学习在自然图像处理领域获取的巨大成功,研究人员开始尝试将其作为一种强有力的技术手段来弥补上述缺陷,得益于神经网络强大的语义提取和表征能力,目前已有许多基于深度学习的算法被提出用于实现火星地貌分类。**Rothrock**等^[9]提出了一种基于深度卷积神经网络的算法SPOC,可识别沙地、基岩和陡坡等地形地貌,该算法成功应用于2020年火星车着陆区的地形识别。**Jiang**等^[10]提出一种端到端的鲁棒深度学习框架用于检测火星地貌,通过可旋转的锚箱机制同时定位和识别不同类型的火星地貌,并引入基于自动编码器的无监督预训练提升训练和推理性能。为增强全局特征和局部特征的差异性,**Sun**等^[11]提出一种新的基于CNN的分类网络GLNET,设计全局和局部双编码器提取图像中的判别性特征,并在网络训练时嵌入类间离散和类内紧凑的约束,进一步提升了分类模型的精度。这些方法在图像分类任务上性能表现卓越,然而由于其庞大的参数量与计算量,算法性能受到一定限制。

为实现网络模型的轻量化,刘猛等^[12]利用剪枝在模型压缩方面的优势,提出基于迭代剪枝VGGNet的火星图像分类方法,通过迭代剪枝舍弃部分不重要的神经元,减少模型参数量与计算量。**Shi**等^[13]结合深度可分离卷积与常规卷积通过两条分支网络提取图像特征,并设计一种双分支特征融合结构BFF将提取到的特征信息融合在一起,大大降低了模型复杂度。虽然VGGNet和LCNN-BFF有效缓解了模型臃肿的问题,但

在一定程度上牺牲了模型的性能,因而如何平衡模型分类效果和轻量化程度,在保持分类精度高的同时,尽可能降低模型的参数量与计算量,实现轻量化模型构建,仍然是一个亟需解决的问题。

因此,本文提出一种可实现遥感卫星搭载的轻量化类火星地貌快速识别分类方法,构建双分支教师-学生网络,利用知识蒸馏技术将教师网络的语义表征能力迁移到学生网络,并嵌入通道-局部空间注意力机制,在保证分类模型高精度的基础上,减少模型的参数量与计算量,实现模型轻量化。同时,通过采集地球上相似地貌,构建了包括悬崖断面、沙漠、河道及雅丹在内的4种典型类火星地貌数据集,便于后续开展快速识别分类试验。实验结果表明,所提模型具有更优越的性能。

1 类火星地貌分类识别方法

1.1 方法概述

知识蒸馏^[14]旨在令性能稍次但模型更轻量的学生网络学习模型臃肿但性能更优越的教师网络的特征知识并逼近教师网络输出,是一种有效压缩模型,减少计算资源和存储容量的方法。本文针对目前火星地貌分类难以平衡模型分类效果和轻量化程度的问题,提出了一种轻量化的类火星地貌快速识别分类方法,利用知识蒸馏技术构建双分支教师-学生网络,融合局部注意力改进传统的通道-空间注意力模块,进一步提升了模型的性能表现,图1为所提分类模型的整体结构。如图所示,教师网络将Swin-Transformer作为骨干网络进行特征提取,Swin-Transformer在图像分类任务上性能表现卓越,但由于其臃肿的网络模型,不便于硬件

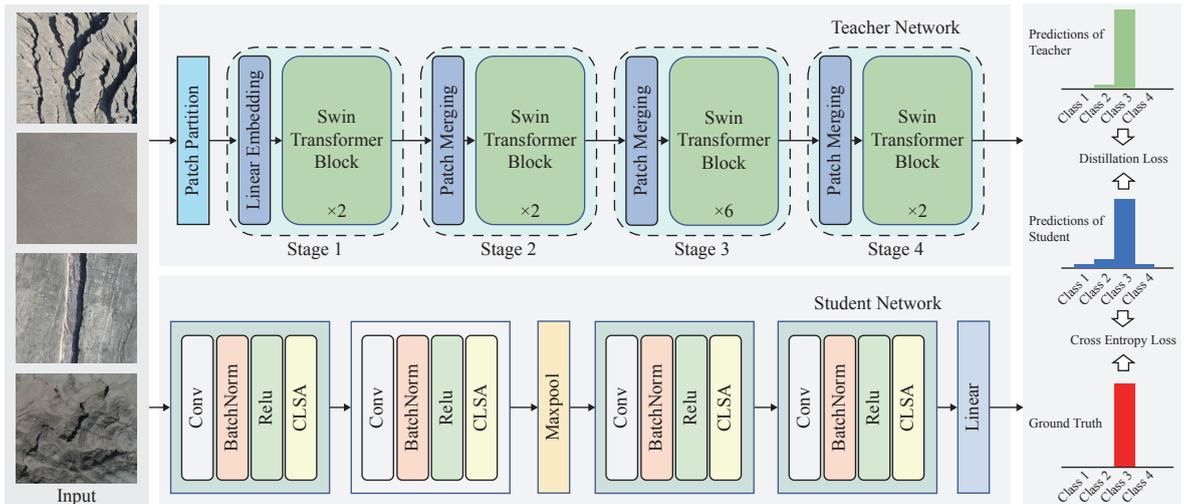


图1 所提类火星地貌分类识别模型

Fig. 1 The proposed classification model for Mars analog terrain

部署, 将其作为教师网络可为学生网络设定学习目标, 起到逐步优化学生网络的作用; 学生网络基于CNN架构设计搭建模型并嵌入通道-局部空间注意力(Channel-Local Spatial Attention Module, CLSA)模块, CNN相较于Swin-Transformer, 模型参数数量和计算复杂度大大降低, 利用知识蒸馏技术将Swin-Transformer的语义表征能力迁移到CNN, 最大程度上优化CNN的性能表现, 使其与教师网络逼近, 在保证模型高精度的同时实现轻量化。具体流程可分为两个阶段: ①教师网络训练, 利用构建的类火星地貌图像数据集, 对教师网络进行单独训练, 训练完成后保存模型; ②学生网络训练, 此时教师网络不参与训练只进行预测, 利用教师网络的推理预测结果以及Ground Truth对学生网络进行监督训练优化。推理时, 只加载训练好的学生网络模型, 输入图像并给出相应的分类识别结果, 从而实现轻量化模型构建。

1.2 教师网络训练

教师网络相对学生网络模型结构更复杂, 网络容量更大, 在训练时可学习到更丰富的特征表示和更平滑的决策边界, 进而引导学生网络在学习过程中调整学习目标, 并逐步提升特征提取与分类能力。本文将Swin-Transformer作为教师网络模型, Swin-Transformer是一种基于Transformer架构的深度学习模型^[15], 由微软研究院于2021年提出, 目前已被广泛应用于图像分类、目标检测、语义分割等计算机视觉任务。Swin-Transformer的具体结构如图1所示, 主要分为4个Stage, 在每个Stage中, 网络通过不同的操作来处理输入数据, 并逐步提取出更高级的语义特征。

对给定的类火星地貌图像, 将其表示为 $I \in R^{H \times W \times C}$, H 和 W 分别表示高和宽, C 表示图像的通道数, 图像对应的标签用 $Y \in \{0, 1, 2, 3\}$ 表示。首先, 输入样本 I 经过Patch Partition操作被划分成多个小patch, 每个patch在后续操作时作为一个独立单元进行处理。然后在Stage 1中, 使用一个Linear Embedding层将这些patch映射到一个高维空间, 从而使其携带更多的特征信息, 值得注意的是, 上述Patch Partition和Linear Embedding操作都是通过一次卷积操作完成的。Swin Transformer Block是Swin-Transformer的核心组件, 在每个Stage中, 都会重复堆叠多个Swin Transformer Block来构建网络, 随着Stage的深入, 特征图的大小逐渐减小, 而每个patch所携带的信息量逐渐增多, 使模型可在不同尺度上捕获到图像的不同特征。

1.3 学生网络训练

学生网络通过学习教师网络的输出特征和决策逻

辑, 旨在保持高预测准确率的同时, 大幅降低模型的参数量与计算复杂度。本文基于CNN架构搭建学生网络模型进行特征提取, 设计学生网络具体结构如图1所示, 包括3个主要部分: 4个特征提取模块、一个最大池化层和一个Linear线性层。每个特征提取模块由一个卷积层、一个归一化层、一个Relu激活层以及一个通道-局部空间注意力模块CLSA组成, 通过4个特征提取模块对输入的类火星地貌图像 I 进行逐级处理, 逐步提取更具判别性的高级语义特征。

为使模型着重关注样本中的关键区域, 充分挖掘图像的局部特征信息, 进一步提升模型对不同地貌的鉴别能力, 本文融合局部注意力改进了传统的通道-空间注意力模块并将其嵌入到特征提取模块中, 整体结构如图2所示。CLSA由通道注意力(Channel Attention, CA)和局部空间注意力(Local Spatial Attention, LSA)两部分组成。输入样本 $I \in R^{H \times W \times C}$, CLSA依次处理得到通道注意力图 $M_C(F) \in R^{1 \times 1 \times C}$ 和局部空间注意力图 $M_S(F') \in R^{H \times W \times 1}$, 整个注意力加权过程可表示为

$$F' = M_C \otimes F \quad (1)$$

$$\bar{F} = M_S \otimes F' \quad (2)$$

F' 表示经通道注意力加权后的特征图, \bar{F} 表示经局部空间注意力加权的特征图, 也即CLSA的最终输出。

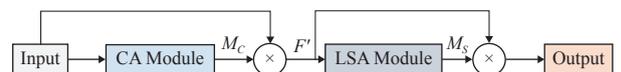


图2 通道-局部空间注意力模块CLSA
Fig. 2 Channel-Local Spatial Attention module (CLSA)

1) 通道注意力CA: 通道注意力可显式建模不同特征通道之间的依赖关系, 通过学习每个特征通道的重要性权重, 根据不同任务选择强调或抑制不同通道。CA模块如图3所示, 以 F 为输入, 首先, 通过最大池化和平均池化操作聚合 F 的全局空间信息, 得到两个不同的特征描述子 F_{\max} 和 F_{avg} , 将池化结果馈送至一个共享权重的多层感知机(MultiLayer Perception, MLP), 然后将映射结果相加, 最终通过Sigmoid得到通道注意力图 M_C , 上述流程可表示为

$$M_C(F) = \sigma(M_{\text{MLP}}(\text{Maxpool}(F)) + M_{\text{MLP}}(\text{Avgpool}(F))) \quad (3)$$

其中: Maxpool 和 Avgpool 分别表示最大池化和平均池化; σ 表示Sigmoid激活函数。

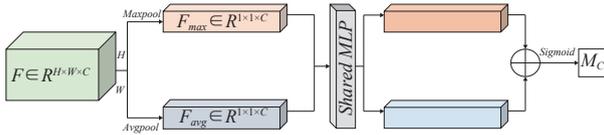


图3 通道注意力CA
Fig. 3 Channel Attention (CA)

2) 局部空间注意力LSA: 空间注意力可根据输入数据的上下文关系, 动态分配关键区域, 从而提高模型的特征选择能力和对复杂模式的识别能力。传统空间注意力机制通常聚焦于全局上下文信息, 而忽略了局部信息的提取, 难以捕获图像的细节特征。因此, 本文融合局部注意力改进了传统的空间注意力模块, 局部空间注意力模块LSA结构如图4所示, 以 $F' \in R^{H \times W \times C}$ 为输入, 首先, 在空间维度并行使用两个 2×2 的局部最大池化和局部平均池化, 得到局部特征 F'_{max} 、 $F'_{avg} \in R^{H/2 \times W/2 \times C}$, 然后在通道维度分别使用全局最大池化和全局平均池化聚合 F'_{max} 、 F'_{avg} 的特征信息, 池化结果分别用 F''_{max} 和 F''_{avg} 表示, 将其拼接到一起后使用一层卷积压缩通道数为1, 最后使用最近邻插值上采样到原始图像大小, 并通过Sigmoid激活函数得到局部空间注意力图 M_S , 上述流程可表示为

$$M_S(F') = \sigma\left(\text{up}\left(\text{conv}\left[F''_{max}; F''_{avg}\right]\right)\right) \quad (4)$$

$$F''_{max} = \text{Maxpool}_{\text{glo}}(\text{Maxpool}_{\text{loc}}(F')) \quad (5)$$

$$F''_{avg} = \text{Avgpool}_{\text{glo}}(\text{Avgpool}_{\text{loc}}(F')) \quad (6)$$

其中: conv 表示卷积操作; up 表示最近邻插值上采样。

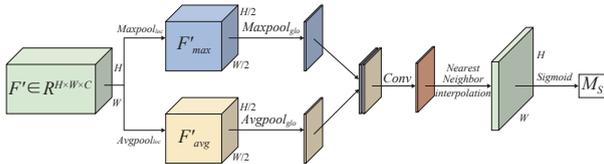


图4 局部空间注意力LSA
Fig. 4 Local Spatial Attention (LSA)

1.4 损失函数

损失函数用于量化模型预测结果和真实标签之间的差异程度, 通过损失函数设计, 模型可感知预测输出和真实输出之间的偏差, 进而在训练过程中逐步调整参数以最小化偏差。本文利用教师网络的推理结果以及Ground Truth对学生网络进行监督训练, 通过反向传播优化网络模型。设计总损失 L_{total} 如式 (7) 所示, 分为两部分: 蒸馏损失 L_{dis} 和交叉熵损失 L_{ce}

$$L_{\text{total}} = (1 - \lambda)L_{\text{dis}} + \lambda L_{\text{ce}} \quad (7)$$

其中: λ 为损失权重系数, 用以调整两部分损失函数的相对重要性, 帮助模型达到最优性能。

1) 蒸馏损失: 蒸馏损失基于教师网络的预测输出, 为学生网络提供一个相对平滑的目标分布。通过最小化蒸馏损失, 鼓励学生网络模仿教师网络的输出分布, 学习教师网络的高级语义特征和决策边界, 从而使学生网络在分类精度上尽可能逼近教师网络。本文采用KL散度作为蒸馏损失函数, KL散度可度量教师模型和学生模型两个概率分布之间的差异, 其优化目标如式 (8) 所示

$$L_{\text{dis}} = D_{\text{KL}}\left(\sigma\left(\frac{P_{\text{swin}}}{T}\right) \parallel \sigma(P_{\text{cnn}})\right) \quad (8)$$

其中: $D_{\text{KL}}(\cdot)$ 表示KL散度的计算, P_{swin} 和 P_{cnn} 分别表示Swin-Transformer和CNN的预测分类结果, T 为蒸馏温度, 通过调整 T 使学生网络尽可能多地从教师网络学习到更丰富和更连续的特征表示。

2) 交叉熵损失: 交叉熵损失用于计算学生网络预测和真实标签之间的差异, 度量预测类别概率分布 $q(x)$ 表示真实类别概率分布 $p(x)$ 的困难程度。利用交叉熵损失约束确保学生网络在模仿教师网络的同时, 可直接学习输入样本数据的真实标签, 降低过拟合风险, 进而提升模型泛化性能, 其优化目标如式 (9) 所示

$$L_{\text{ce}} = - \sum_{i=1}^n p(x_i) \log(q(x_i)) \quad (9)$$

其中: n 表示训练样本总数。

2 类火星地貌图像数据集

2.1 火星地貌特征

火星地貌复杂多样, 其形成与演化深受其内部地质构造及外部天体活动的影响。火星地貌的基本特征为南北二分性^[16], 北部大部分区域为广袤的平原, 平均海拔较低, 地势相对平缓; 南部为高原地带, 主要以火山高原和撞击高原为主; 中部则集中了大量的干涸河床, 是北部平原和南部高原的过渡地带。火星地貌依据其形成机制和地貌特征的不同大致可划分为峡谷、沙漠、干涸河道、雅丹、陨石坑及火山等地形地貌^[17]。多种不同类型的地貌相互交织分布, 共同构造形成了火星独特的地貌特征。

2.2 图像数据集

现有火星地貌识别和模拟演示的难点之一在于: 缺乏相应的类火星地貌图像分析数据库, 无法构建智能识别模型。为此, 本文基于火星上存在的悬崖断面

(峡谷的两侧)、沙漠、河道、雅丹4种典型地貌, 通过在地球上实地采集以上四种相似地貌, 构建了类火星地貌图像数据集一套。

新疆因其干旱极端的气候和独特的地质特征, 一直以来被考虑作为模拟研究火星环境的理想试验场所。例如, Tan等^[18]研究了南疆米兰戈壁风沙的沙通量密度和粒度分布, 对于火星表面风沙的模拟研究具有重要意义。Yue等^[18]利用HiRISE数据和数字高程模型对火星上的板状-脊状多边形地形(Platy-Ridged-Polygonized, PRP)进行了详细分析, 并研究了这些地形的分布和微地形特征, 同时Yue还指出在新疆塔里木盆地发现了类似的地貌特征, 表明该地区的地质过程可能与火星上的PRP地形形成过程相似。苟盛等^[19]对比了火星Evros Vallis和新疆塔里木盆地开都河流域的河网形态及水文特征, 指出Evros Vallis的流域形态与水文参数和地球上自然形成的流域十分相似。因此, 本文选定新疆作为类火星地貌数据集的实地采集点, 2024年5月, 在新疆克拉玛依开展无人机飞行实验进行数据采集, 实验所用两台无人机的平台参数如表1所示, 图5为悬崖断面、沙漠、河道及雅丹4种类火星地貌的图像示例。实际采集过程中, 为保证数据集的类内多样性, 对无人机的拍摄角度、环境(包括光照条件、拍摄背景)以及姿态等因素进行了充分调整, 确保了数据集的质量。

表1 无人机平台参数

Table 1 Parameters of the Unmanned Aerial Vehicle (UAV) Platform

型号	分辨率/pixel	实验飞行高/m
DJI Mavic 2	5 472 × 3 648	120
DJI Mavic 3	5 280 × 3 956	120



图5 类火星地貌图像示例

Fig. 5 Illustration of Mars analog terrain images

2.3 数据增强

数据增强^[20]是一种通过旋转、裁剪、翻转等方式改变输入数据, 以生成更多训练样本, 提高模型泛化能力和效果的技术。当训练数据过少时, 模型难以从有限的样本中学习到足够的信息来概括整个数据分布, 致使模型在训练集上的效果可能不错, 但在测试集上的泛化性能不佳, 而大规模的数据量有助于训练

模型学习更多的输入模式和规律, 降低模型过拟合的风险。本文通过无人机飞行实验采集悬崖断面地貌275张、沙漠地貌71张、河道地貌306张、雅丹地貌361张, 原始图像大小为5 472 × 3 648。后续通过翻转、裁剪等数据增强手段获取每种类别地貌图像各800张, 总计3 200张, 图像尺寸为2 048 × 2 048。所构建的类火星地貌图像数据集数据量充足, 数据类型典型完备。图6展示了各类地貌数据增强前后的数据量对比。

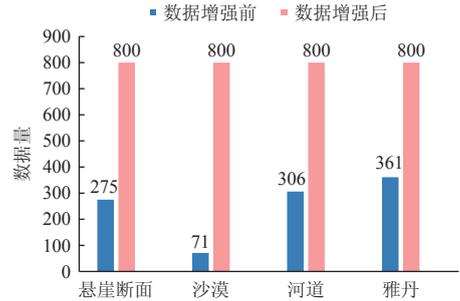


图6 数据增强前后各类地貌的数据量对比

Fig. 6 Comparison of data volume for various terrains before and after data augmentation

3 实验结果与分析

3.1 实验设置

1) 实验细节: 按8:1:1的比例将自建数据集划分为训练集、验证集与测试集, 网络输入图像尺寸为224 × 224。基于Pytorch1.11.0使用单卡NVIDIA GeForce RTX 3090Ti实现了所提轻量化类火星地貌分类识别方法, 优化算法采用Adam, 初始化学学习率为0.000 1, 权重衰减为0.000 1, Batch Size大小为16, 训练轮次为100, 超参数蒸馏温度T和损失权重λ分别设置为11和0.3。

2) 评价指标: 综合了OA、Kappa、F1、Precision及Recall共5种指标对模型性能进行评估, 如式10~15所示; 同时引入Params、FLOPs和FPS共3种指标评估模型轻量化程度和推理效率。

其中OA为总体分类精度, 其值 O_a 表示正确分类样本占所有样本的比例; Kappa系数用于衡量分类模型的准确性, 其值 K_a 越高代表分类效果越好; Precision为精确率, 其值 P_r 越高代表模型正确识别正类样本的能力越强; Recall为召回率, 其值 R_c 表示模型对正类样本的覆盖程度; F1分数为精确率和召回率的调和平均值F1; Params和FLOPs分别表示模型的参数量和计算量, 其值越小模型越轻量; FPS表示每秒处理的图像帧数, 其值越大表示模型推理效率越高。

$$O_a = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

$$K_a = \frac{O_a - P_e}{1 - P_e} \quad (11)$$

$$F_1 = \frac{2 \times P_r \times R_c}{P_r + R_c} \quad (12)$$

$$P_r = \frac{TP}{TP+FP} \quad (13)$$

$$R_c = \frac{TP}{TP+FN} \quad (14)$$

$$P_e = \frac{(TP+FP) \times (TP+FN) + (TN+FN) \times (TN+FP)}{(TP+FP+TN+FN)^2} \quad (15)$$

其中: OA为总体分类精度, 其值 O_a 表示正确分类样本占有所有样本的比例; Kappa系数用于衡量分类模型的准确性, 其值 K_a 越高代表分类效果越好; Precision为精确率, 其值 P_r 越高代表模型正确识别正类样本的能力越强; Recall为召回率, 其值 R_c 表示模型对正类样本的覆盖程度; F_1 分数为精确率和召回率的调和平均值; Params和FLOPs分别表示模型的参数量和计算量, 其值越小模型越轻量; FPS表示每秒处理的图像帧数, 其值越大表示模型推理效率越高。

3.2 对比实验

为验证所提类火星地貌分类识别模型LWNet的有效性与优越性, 在自建数据集上开展了广泛实验, 并将LWNet的实验结果与国际已有的其它11种先进算法进行了对比, 包括两种机器学习算法: KNN和

Decision Tree, KNN和Decision Tree作为机器学习最为经典和使用最广泛的两种算法, 在图像分类领域极具代表性; 9种深度学习算法: VGG16^[21]、GLNET、LCNN-BFF、Swin-Transformer、Res2Net50^[22]、EMTCAL^[23]、Xception^[24]、ShuffleNet^[25]、MobileNet^[26-27], 深度学习对比算法以模型分类精度和轻量化程度为出发点考虑选择, 其中VGG16、Swin-Transformer以及Res2Net50因其优越的性能在图像分类、目标检测等计算机视觉领域被广泛应用, LCNN-BFF、Xception、ShuffleNet和MobileNet共4种对比算法模型轻量, 作为对比可增强实验结果的可靠性。

图7对LWNet提取的4种类火星地貌中间特征进行了可视化展示, 表2为对比实验结果, 图8展示了各算法的分类混淆矩阵。如表2所示(加粗标注表示最优结果), 所提LWNet的OA、Kappa和 F_1 分数分别达到了97.81%、97.08%和97.81%, 分类效果优于多数对比算法, 仅次于Swin-Transformer和EMTCAL位列第三, 表明了LWNet卓越的分类性能; 且LWNet的参数量仅有1.11M, 在所有对比算法中位列第一, 尽管计算量相对ShuffleNet和MobileNet两种轻量化算法并不十分突出, 但LWNet的总体分类精度却分别提升了2.18%和3.12%。此外, LWNet的FPS达到了每秒79.97帧, 在10种深度学习算法中位列第二, KNN、Decision Tree以及MobileNet这3种算法的FPS均高于LWNet, 但它们在分类效果上的表现却不尽人意。因此, 综合所有指标进行评估, 所提基于知识蒸馏的轻量化类火星地貌分类识别模型性能表现最好。

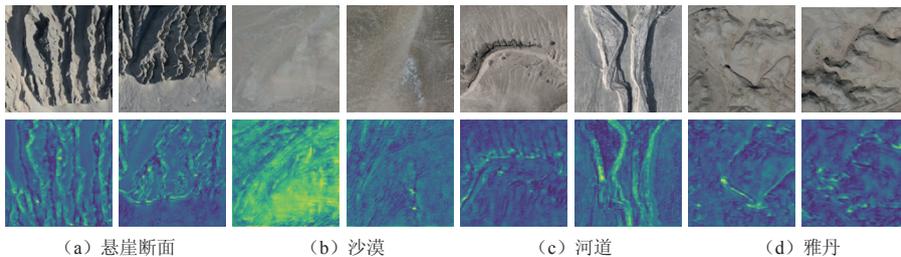


图7 特征可视化结果

Fig. 7 Visualization results of features

表2 对比实验结果

Table 2 Comparative experimental results

Method	OA	Kappa	F_1	Precision	Recall	Params/M	FLOPs/M	FPS
KNN	0.949 2	0.942 0	0.949 0	0.949 3	0.949 2	—	—	119.82
Decision Tree	0.937 5	0.928 9	0.937 4	0.937 4	0.937 5	—	—	134.57
VGG16	0.959 4	0.945 8	0.959 2	0.962 8	0.959 4	134.28	15 466.18	45.39
GLNET	0.853 1	0.804 2	0.850 5	0.857 4	0.853 1	269.57	92 800.46	25.18
LCNN-BFF	0.962 5	0.980 8	0.963 1	0.967 4	0.962 5	6.18	1 827.90	68.49
Swin-Transformer	0.990 6	0.987 5	0.990 6	0.990 8	0.990 6	86.75	15 168.81	31.12
Res2Net50	0.962 5	0.950 0	0.962 3	0.966 4	0.962 5	23.66	4 303.81	28.52

续表 2

Method	OA	Kappa	F1	Precision	Recall	Params/M	FLOPs/M	FPS
EMTCAL	0.984 3	0.979 2	0.984 4	0.984 8	0.984 4	27.80	4 233.92	47.02
Xception	0.978 1	0.970 8	0.977 9	0.979 1	0.978 1	20.82	4 597.28	43.20
ShuffleNet	0.956 3	0.941 7	0.955 6	0.957 1	0.956 2	1.29	154.70	76.82
MobileNet	0.946 9	0.929 2	0.946 2	0.949 7	0.946 9	3.22	587.89	161.92
Ours(LWNet)	0.978 1	0.970 8	0.978 1	0.979 0	0.978 1	1.11	731.46	79.97

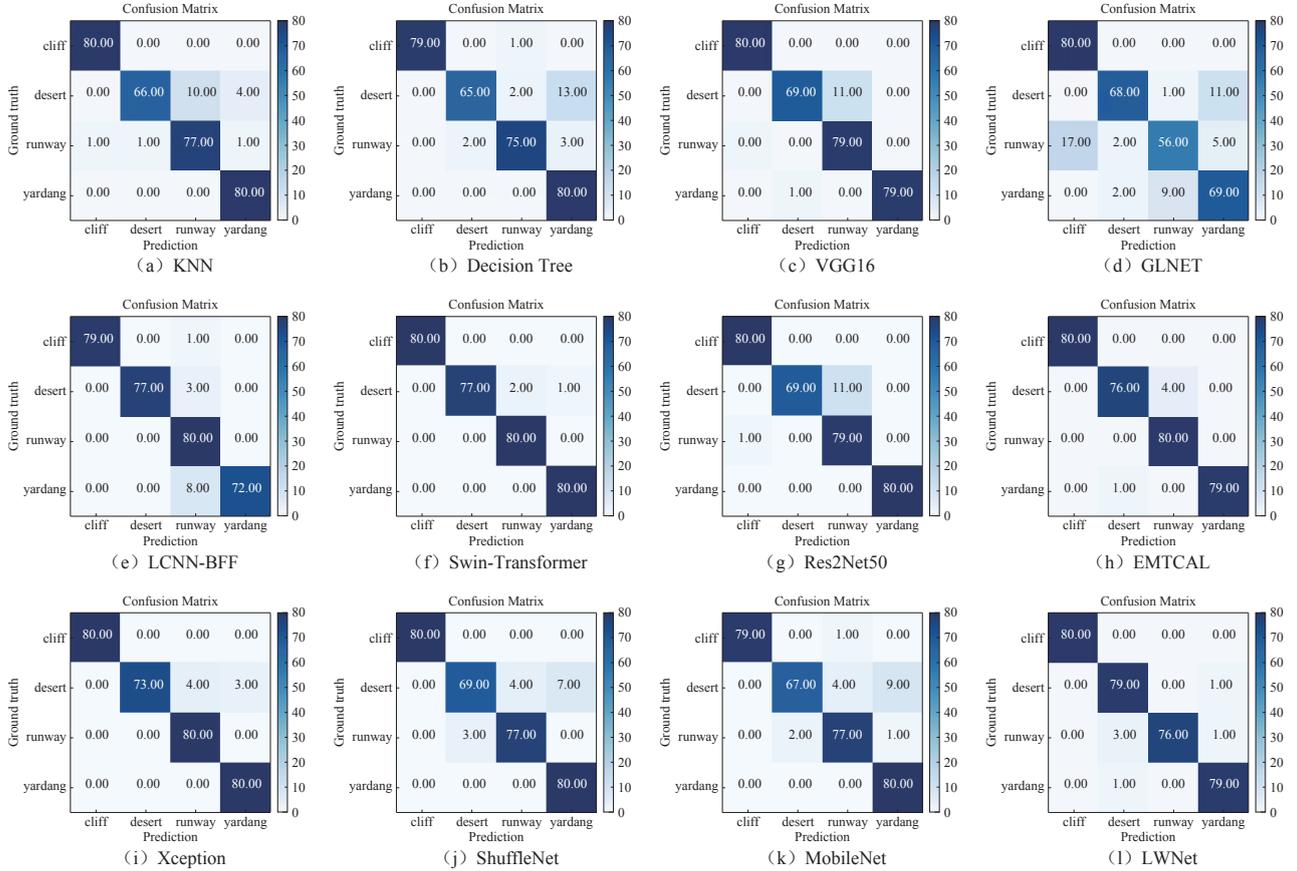


图 8 各算法的分类混淆矩阵

Fig. 8 Classification confusion matrices of the various algorithms

3.3 消融实验

为研究通道-局部空间注意力模块CLSA以及蒸馏操作对所提分类模型LWNet的影响, 同时探索超参数蒸馏温度 T 和损失权重 λ 的最优取值, 本文进一步设置了如下3组消融实验。

1) 注意力模块和蒸馏的有效性验证: 通过控制变量选择去除或保留CLSA模块和蒸馏操作设计消融实验, 具体实验结果如表3所示(其中D表示Distillation蒸馏操作, A为Attention注意力模块)。由表3可知, 当CLSA模块和蒸馏都被去除时, 总体分类精度OA为93.43%; 当选择保留蒸馏或CLSA模块时, OA分别提高了1.57%和2.82%; 当两个模块均被保留时, 模型分类效果最佳, 达到了97.81%。证明所提CLSA模块和蒸馏操作可在一定程度上有效提升模型分类精度。

表 3 注意力模块和蒸馏的有效性验证结果

Table 3 Validation of the effectiveness of attention module and distillation

D	A	OA	Kappa	F1	Precision	Recall
—	—	0.934 3	0.912 5	0.934 0	0.940 1	0.934 4
√	—	0.950 0	0.933 3	0.949 3	0.952 9	0.950 0
—	√	0.962 5	0.950 0	0.962 2	0.965 5	0.962 4
√	√	0.978 1	0.970 8	0.978 1	0.979 0	0.978 1

2) 蒸馏温度 T 的最优取值: 蒸馏温度 T 用于调整模型的输出概率分布, 提升模型的泛化性能。本文通过在1~100内对 T 进行取值设计消融实验, 具体实验结果如表4所示(加粗标注表示最优结果), 可知当 T 取11时, 总体分类精度最高, 即 T 的最优取值为11。

3) 损失权重 λ 的最优取值: 损失权重 λ 用于平衡教师模型和学生模型的预测输出, 通过调整 λ 的取值, 实

现更高效的知识迁移, 确保学生模型充分学习教师网络行为。通过在0.1~0.9范围内对 λ 进行取值设计消融实验, 具体实验结果如表5所示(加粗标注表示最优结果), 可知当 λ 取0.3时, 总体分类精度最高, 即 λ 的最优取值为0.3。

表4 蒸馏温度 T 的最优取值实验结果

Table 4 Experimental determination of the optimal distillation temperature (T)

T	OA	Kappa	F1	Precision	Recall
1	0.946 8	0.929 1	0.946 5	0.953 7	0.946 9
3	0.950 0	0.933 3	0.949 1	0.954 1	0.950 0
5	0.959 3	0.945 8	0.959 2	0.963 9	0.959 4
7	0.956 2	0.941 7	0.955 7	0.962 0	0.956 2
9	0.962 5	0.950 0	0.962 9	0.966 7	0.962 5
11	0.978 1	0.970 8	0.978 1	0.979 0	0.978 1
13	0.953 1	0.937 5	0.952 5	0.959 7	0.953 1
15	0.956 3	0.941 7	0.955 9	0.958 0	0.956 2
20	0.946 9	0.929 2	0.945 9	0.948 6	0.946 9
50	0.937 5	0.912 5	0.934 1	0.941 3	0.934 4
100	0.925 0	0.900 0	0.922 1	0.934 1	0.925 0

表5 损失权重 λ 的最优取值实验结果

Table 5 Experimental determination of the optimal loss weight (λ)

λ	OA	Kappa	F1	Precision	Recall
0.1	0.962 5	0.950 0	0.962 3	0.967 4	0.962 5
0.2	0.959 3	0.945 8	0.958 7	0.963 7	0.959 4
0.3	0.968 8	0.958 3	0.968 6	0.972 2	0.968 8
0.4	0.953 1	0.937 5	0.952 8	0.959 7	0.953 1
0.5	0.959 3	0.945 8	0.959 2	0.963 4	0.959 4
0.6	0.965 6	0.954 2	0.965 1	0.967 2	0.965 6
0.7	0.953 1	0.937 5	0.952 6	0.958 6	0.953 1
0.8	0.946 8	0.929 2	0.945 9	0.949 7	0.946 9
0.9	0.956 3	0.941 7	0.955 7	0.962 0	0.956 2

4 结 论

本文提出一种轻量化类火星地貌快速识别分类方法, 在保证模型分类高精度的同时实现了轻量化。具体来说, 构建双分支教师-学生网络模型, 利用知识蒸馏技术在模型压缩和迁移学习方面的优势, 将教师网络的语义提取和表征能力迁移到学生网络, 同时融合局部注意力改进传统通道-空间注意力机制并嵌入学生网络, 以增强模型对火星地貌的分类识别能力, 进一步提升所提模型的性能表现。同时, 建立了一个类火星地貌图像分析数据库, 模拟真实火星地形地貌; 所提轻量化类火星地貌快速识别分类方法在自建数据集上的实验结果验证表明了LWNet的有效性和优越性。

参 考 文 献

[1] ZHAO H, LIU S, TONG X, et al. MarsMapNet: a novel superpixel-

guided multiview feature fusion network for efficient martian landform mapping[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 1-16.

- [2] BUE B D, STEPINSKI T F. Machine detection of martian impact craters from digital topography data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(1): 265-274.
- [3] WANG J, CHENG W, ZHOU C, et al. Automatic mapping of lunar landforms using DEM-derived geomorphometric parameters[J]. *Journal of Geographical Sciences*, 2017, 27(11): 1413-1427.
- [4] COVER T M, HART P E. Nearest neighbor pattern classification[J]. *IEEE Trans. Inf. Theory*, 1967, 13: 21-27.
- [5] HUNT E B, MARIN J, STONE P J. Experiments in induction[J]. *The American Journal of Psychology*, 1966, 80(4): 651-653.
- [6] CORTES C, VAPNIK V. Support-vector networks[J]. *Machine Learning*, 1995, 20(3): 273-297.
- [7] BREIMAN L. Random Forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [8] SHANG C, BARNES D. Fuzzy-rough feature selection aided support vector machines for Mars image classification[J]. *Computer Vision and Image Understanding*, 2013, 117(3): 202-213.
- [9] ROTHROCK B, KENNEDY R, CUNNINGHAM C, et al. Spoc: deep learning-based terrain classification for mars rover missions[C]//Proceedings of AIAA SPACE 2016. [S. l.]: AIAA, 2016.
- [10] JIANG S, WU F, YUNG K L, et al. A robust end-to-end deep learning framework for detecting Martian landforms with arbitrary orientations[J]. *Knowledge-Based Systems*, 2021, 234: 107562.
- [11] SUN H, LIN Y, ZOU Q, et al. Convolutional neural networks based remote sensing scene classification under clear and cloudy environments[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). [S. l.]: ICCVW, 2021.
- [12] 刘猛, 刘劲, 尹李君, 等. 基于迭代剪枝VGGNet的火星图像分类[J]. *液晶与显示*, 2023, 38(4): 507.
- LIU M, LIU J, YI L J, et al. Martian image classification based on iterative pruning VGGNet[J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(4): 507.
- [13] SHI C, WANG T, WANG L. Branch feature fusion convolution network for remote sensing scene classification[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, 13: 5194-5210.
- [14] HINTON G. Distilling the knowledge in a neural network[J]. *ArXiv Preprint ArXiv*, 2015, 12: 02531.
- [15] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. [S. l.]: IEEE, 2021.
- [16] 王琤. 火星典型地貌特征知识的描述分类与识别[D]. 武汉: 武汉大学, 2009.
- WANG Z. Characterization classification and identification of Martian specific landform-Characters[D]. Wuhan: Wuhan University, 2009.
- [17] 刘洋, 吴兴, 刘正豪, 等. 火星的地质演化和宜居环境研究进展[J]. *地球与行星物理评论*, 2021, 52(4): 416-436.
- LIU Y, WU X, LIU Z H, ZOU Y L. Geological evolution and habitable environment of Mars: progress and prospects[J]. *Reviews of Geophysics and Planetary Physics*, 2021, 52(4): 416-436.
- [18] TAN L, ZHANG K, WANG H, et al. Vertical sand flux density and grain-size distributions for wind-blown sand over a gobi surface in milan, southern Xinjiang, China[J]. *Frontiers in Environmental Science*, 2022, 10: 1-10.

- [19] YUE Z, GOU S, MICHAEL G, et al. An Investigation of the hypotheses for formation of the platy-ridged-polygonized terrain in elysium planitia Mars[J]. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2017, XLII-3/W1: 205-211.
- [20] 苟盛, 岳宗玉, 邸凯昌, 等. 火星Evros Vallis与塔里木盆地开都河流域河网形态和水文特征比较[J]. *遥感学报*, 2018, 22(2): 313-323.
GOU S, YUE Z Y, DI K C, et al. Quantitative comparison of morphometric and hydrological characteristics of valley networks between Evros Vallis on Mars and Kaidu River in Tarim Basin as terrestrial analog[J]. *Journal of Remote Sensing*, 2018, 22(2): 313-323.
- [21] 孙书魁, 范菁, 孙中强, 等. 基于深度学习的图像数据增强研究综述[J]. *计算机科学*, 2024, 51(1): 150-167.
SUN S K, FAN J, SUN Z Q, et al. Survey of image data augmentation techniques based on deep learning[J]. *Computer Science*, 2024, 51(1): 150-167.
- [22] SIMONYAN K. Very deep convolutional networks for large-scale image recognition[J]. *ArXiv Preprint ArXiv: 1409.2014*, 12: 1556.
- [23] GAO S H, CHENG M M, ZHAO K, et al. Res2net: a new multi-scale backbone architecture[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 43(2): 652-662.
- [24] TANG X, LI M, MA J, et al. EMTCAL: Efficient multiscale transformer and cross-level attention learning for remote sensing scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-15.
- [25] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S. l.]: IEEE, 2017.
- [26] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S. l.]: IEEE, 2018.
- [27] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications (2017)[J]. *ArXiv Preprint ArXiv: 2017.126: 1704.04861*.

作者简介:

李建龙(2000-), 男, 硕士研究生, 主要研究方向: 遥感图像变化检测。

通信地址: 湖南大学电气与信息工程学院(410082)

E-mail: LJL@hnu.edu.cn

杨彬(1989-), 男, 副教授, 博士生导师, 主要研究方向: 图像处理、模式识别、人工智能等。**本文通信作者**。

通信地址: 湖南大学电气与信息工程学院(410082)

E-mail: binyang@hnu.edu.cn

Lightweight Rapid Recognition and Classification for Mars Analog Terrain

LI Jianlong¹, ZHANG Lihua^{2,3}, HE Lili¹, LIU Xinxin¹, YANG Bin¹, ZHOU Jishi^{2,3}

(1. School of Electrical and Information Engineering, HuNan University, Changsha 410082, China;

2. Lunar Exploration and Space Engineering Center, China National Space Administration, Beijing 100190, China;

3. Deep Space Exploration Lab, Beijing 100190, China)

Abstract: Recognition and classification of Mars analog terrain aim to simulate and study the Mars environment by analyzing Mars analog terrain images, which holds significant research value for exploring scientific questions such as formation, evolution, and potential habitability of Mars. In response to the challenge of balancing classification performance and model lightweighting in current Mars terrain classification algorithms, a lightweight, rapid recognition and classification method for Mars analog terrain is proposed (LWNet). This algorithm constructs a dual-branch teacher-student network, employs knowledge distillation to reduce the number of parameters and computational load of the model, and integrates attention mechanism to enhance the capability of terrain classification and recognition, achieving high accuracy and lightweight classification models. To verify the classification performance of the proposed method, a dataset of Mars analog terrain on Earth was established, including four typical Mars landforms: cliff, desert, channel, and yardang, with each type of terrain consisting of 800 images. The dataset was employed to conduct rapid recognition and classification experiments with LWNet. The results indicate the overall accuracy reaches 97.81%, which only decreases by 1.25% compared with Swin-Transformer, while its Parameters and FLOPs are only 1.3% and 4.8% of Swin-Transformer, respectively. Experimental results verify the effectiveness and superiority of the LWNet.

Keywords: Mars analog terrain; lightweight; knowledge distillation; image classification

Highlights:

- A lightweight algorithm for rapid recognition and classification of Mars analog terrain was proposed, achieving lightweight while ensuring high accuracy of the model.
- The traditional channel-spatial attention module was improved by integrating local attention to enhance the recognition capability of terrain category identification.
- A database was established for the analysis of Mars analog terrain, addressing the current gap.

[责任编辑: 高莎, 英文审校: 宋利辉]