

文章编号: 1004—5589 (2025) 02—0299—12

徐梦霞, 王明常, 王凤艳, 等. 基于多源数据融合与面向对象影像分析的吉林德惠侵蚀沟提取 [J]. 世界地质, 2025, 44 (2): 299–310. DOI: 10.3969/j.issn.1004–5589.2025.02.009.

XU M X, WANG M C, WANG F Y, et al. Extraction of erosion gully based on multi-source data fusion and object-based image analysis in Dehui of Jilin [J]. World Geology, 2025, 44 (2): 299–310. DOI: 10.3969/j.issn.1004–5589.2025.02.009.

基于多源数据融合与面向对象影像分析的 吉林德惠侵蚀沟提取

徐梦霞, 王明常*, 王凤艳, 纪雪, 刘星男

吉林大学 地球探测科学与技术学院, 长春 130026

摘要: 为提高基于遥感数据提取侵蚀沟信息的精度, 提出了基于多源数据融合和面向对象影像分析的侵蚀沟提取模型, 用以提取吉林省德惠市侵蚀沟的分布信息, 以期对侵蚀沟的产生起到预警作用。研究融合了 Sentinel-1 雷达数据、Sentinel-2 光学遥感影像和研究区地形数据, 采用多尺度分割方法对研究区地物进行分割。基于多尺度分割结果, 提取强度、纹理、光谱、指数、地形和几何特征用以构建多维特征集, 利用随机森林基尼指数对特征贡献度进行评估和筛选, 获取最优特征集。构建基于面向对象影像分析的随机森林提取模型对研究区的侵蚀沟信息进行提取, 采用生成随机点并对其目视解译的方法对提取精度进行评估。为进一步探究多源数据融合对侵蚀沟提取精度的影响, 基于相同的样本数据集, 针对不同数据源对研究区侵蚀沟进行提取并比较其精度。研究结果表明, 基于多源数据融合的侵蚀沟提取总体精度为 94.03%, Kappa 系数为 0.839, 表明该模型具有较好的性能。相较于单一的数据源, 多源数据融合具有更高的提取精度, 可为随机森林模型提供大量有效的特征值。结果表明, 德惠市侵蚀沟主要分布在中部和西南部, 侵蚀沟整体呈东北–西南带状展布, 多位于居民地周边并与一定的水系相连, 对其侵蚀灾害需加以重视。

关键词: 侵蚀沟; 随机森林; 多源数据; 面向对象影像分析; 特征优选

中图分类号: P237; X87 **文献标识码:** A **DOI:** 10.3969/j.issn.1004–5589.2025.02.009

Extraction of erosion gully based on multi-source data fusion and object-based image analysis in Dehui of Jilin

XU Mengxia, WANG Mingchang*, WANG Fengyan, JI Xue, LIU Xingnan

College of Geo-Exploration Science and Technology, Jilin University, Changchun 130026, China

Abstract: In order to improve the accuracy of extracting erosion gully information based on remote sensing data, the erosion gully extraction model based on multi-source data fusion and object-based image analysis (OBIA)

收稿日期/Received: 2024–10–10; 修订日期/Revised: 2024–12–05; 出版日期/Published: 2025–05–25

基金项目: 国家自然科学基金重点项目 (42330607); 吉林省教育厅科学技术研究项目 (JKH20231181KJ)

第一作者: 徐梦霞 (1999—), 女, 硕士研究生, 主要从事遥感数据处理与应用方面研究。E-mail: xmx22@mails.jlu.edu.cn

* 通信作者: 王明常 (1975—), 男, 教授, 博士生导师, 主要从事遥感与地理信息系统研究。E-mail: wangmc@jlu.edu.cn

© Editorial Office of World Geology. This is an open access article under the CC BY–NC–ND 4.0 license.

is proposed to extract the distribution information of erosion gully in Dehui City, Jilin Province, with a view to play an early warning role for the generation of erosion gully. The study fuses Sentinel-1 radar data, Sentinel-2 optical remote sensing images, and topographic data of the study area, and uses the multiresolution segmentation method to segment the features in the study area. Based on the results of multiresolution segmentation, intensity, texture, spectral, index, topographic and geometric features are extracted for constructing multi-dimensional sets, and the contribution of features is evaluated and optimized using the random forest Gini index to obtain the optimal set of features. The random forest extraction model based on OBIA was constructed to extract erosion gully information in the study area, and the extraction accuracy was evaluated using the method of generating random points and visually deciphering them. To further investigate the effect of multi-source data fusion on the accuracy of erosion gully extraction, based on the same sample dataset, the erosion gully in the study area was extracted for different data sources and its accuracy was compared. The results show that the overall accuracy of erosion gully extraction based on multi-source data fusion is 94.03%, and the Kappa coefficient is 0.839, indicating that the model has a good performance. Compared with a single data source, multi-source data fusion has higher extraction accuracy, and provides a large number of effective feature values for the random forest model. Erosion gullies in Dehui City are mainly distributed in the central and southwestern part of the city, and the erosion as a whole shows a northeast-southwest belt direction, mostly distributed around the settlement and connected with certain water systems, and its erosion situation needs to be emphasized.

Keywords: erosion gully; random forest; multi-source data; object-based image analysis; feature optimization

0 引言

侵蚀沟是一种由暂时性流水形成的线性槽形凹地,是东北地区土壤侵蚀问题的集中表现。侵蚀沟导致耕地破碎化和土壤有机质的严重流失,它的出现标志着土地退化已发展到一个需引起高度重视的极端阶段^[1-2]。有效识别侵蚀沟是土壤侵蚀监测中的关键环节,快速获取其相关信息对于耕地资源保护、农业生产及区域生态的可持续发展具有重要意义^[3]。

侵蚀沟的传统调查方法如填土法、测针法等已无法满足当前的研究需要^[4]。遥感技术于20世纪60年代兴起,因具有快速获取大范围地表覆盖信息的优点,在侵蚀沟遥感识别领域被广泛应用^[5]。目前,基于遥感影像对侵蚀沟信息进行提取主要有目视解译法、基于地形地貌特征的提取方法、面向对象的影像分析法以及基于深度学习的方法等。

目视解译是基于影像的光谱差异与专家知识,对遥感影像进行目视判读的方法。目视解译最早被用于侵蚀沟的制图,该方法依赖先验知识,效率较低,人工绘制的边界不确定性较大,但定性准确度高,至今仍作为侵蚀沟提取的一种主要方法被广泛

使用^[6]。地形是影响侵蚀沟的关键因素^[7],基于地形地貌特征的提取方法是应用侵蚀沟在沟缘线处坡度的突然变化作为特征进行侵蚀沟提取,如利用双向地形阴影法可以从原理上分析影响侵蚀沟提取精度的干扰因素并提出相应解决方法,地学意义明显且无需训练样本^[8]。该方法的关键在于地形数据的选择,不同研究区侵蚀沟形态各异,东北地区侵蚀沟以切沟(宽1~2 m)为主^[9],在中低分辨率的地形数据中难以辨认,因此需要高精度的数字高程模型(digital elevation model, DEM)数据或激光雷达数据来提取侵蚀沟的长、宽、高等信息^[10]。但高精度地形数据获取成本较高,往往无法应用于在大区域尺度上的侵蚀沟提取^[11]。基于深度学习的侵蚀沟提取则是利用神经网络模型,U-Net、R2 U-Net和SegNet等网络已经被应用于绘制侵蚀沟影响区域^[12],但深度学习算法需要大量像素级或对象级训练数据才能获取较好的提取结果^[13]。

面向对象影像分析是以影像分割之后的图斑为基础单元,并在此基础上进行影像特征的计算,进而实现对目标信息的提取。面向对象是目前基于高空间分辨率遥感影像广泛应用的影像分析方法^[14]。侵蚀沟地貌特征复杂,光谱异质性较大,面向对象

的分析方法是以同质性较高的像元所组成的对象作为分析单元, 利用对象的光谱、几何和纹理等特征信息进行提取, 研究表明对侵蚀沟分割对象进行分析更适合处理侵蚀沟的光谱异质性^[15]。遥感数据的选择是基于面向对象提取侵蚀沟信息的重点之一。有研究将地形数据与多光谱遥感影像相结合, 在面向对象侵蚀沟提取中表现良好^[16-17]。早期面向对象的侵蚀沟提取方法多采用手动建立规则集, 根据纹理、地形等特征建立规则将侵蚀沟与其他地物分离^[18-19]。随着机器学习的发展, 支持向量机 (support vector machine, SVM)、随机森林 (random forest, RF) 等算法在侵蚀沟提取领域被广泛应用。目前, 基于面向对象影像分析的侵蚀沟提取研究多采用单一数据源或地形数据结合光学影像, 并未对多源数据融合对侵蚀沟提取精度的影响展开研究。雷达数据包含丰富的纹理信息, 能够综合体现地物的表面性质, 与光学影像和地形数据相融合作为侵蚀沟提取的数据源, 可以提供更加多维的特征信息, 理论上可以提高信息提取精度。

笔者以吉林省德惠市为研究区, 利用 Sentinel-1 雷达数据和 Sentinel-2 光学遥感影像, 结合研究区地形数据, 融合强度、纹理、光谱、指数、地形和几何特征等多种特征信息, 基于随机森林基尼指数对特征进行优选, 采用面向对象方法构建多特征随机森林模型, 用以提取研究区侵蚀沟的分布信息, 以期对侵蚀沟的产生起到预警作用, 也为国家在水土保持政策的制定上提供参考。

1 研究区域与数据

1.1 研究区概况

研究区位于吉林省德惠市 (125°14' ~ 126°24'E, 44°02' ~ 44°53'N), 地处松辽平原腹地, 位于吉林省中北部, 东临松花江, 西部与农安县接壤。研究区地形南高北低、东高西低, 海拔 161 ~ 230 m, 地势平缓。德惠市属温带大陆性季风气候区, 受季风气候影响, 年降雨量偏少, 平均气温较高, 四季分明。总面积约为 3 322.24 km², 其中耕地占 81.03%, 土壤耕层较厚, 土质肥沃。

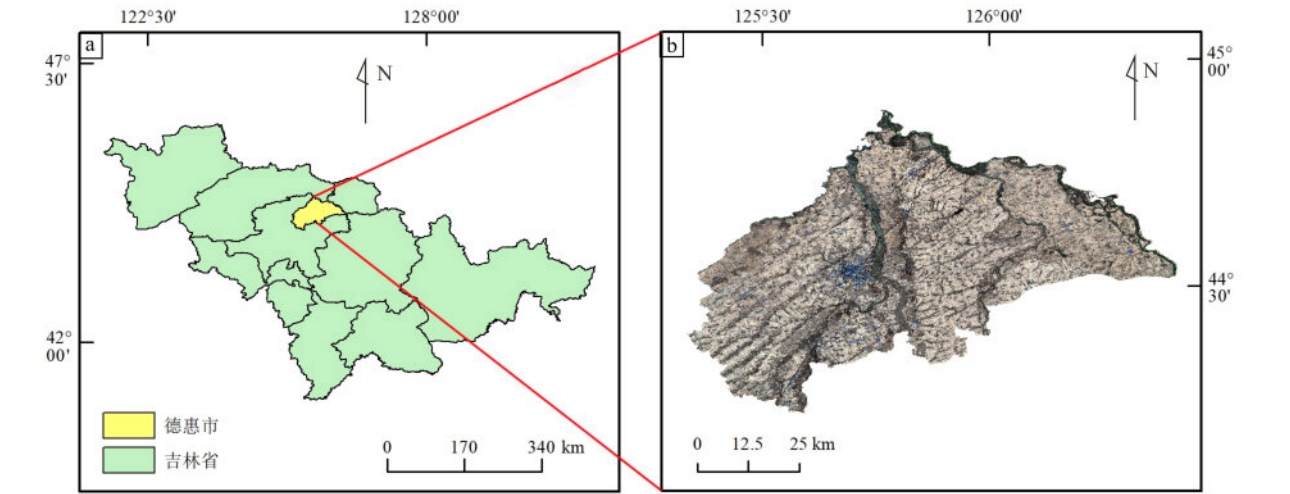


图 1 研究区 (a) 及数据源 (b)
Fig. 1 Study area (a) and data sources (b)

1.2 数据及预处理

1.2.1 数据源

哨兵一号 (Sentinel-1) 是哥白尼计划中的观测卫星, 包含两颗处于同一轨道平面的极轨卫星 A 星和 B 星。搭载 C 波段合成孔径雷达, 具有穿透力强、重访周期短、不间断监测等优点。因其不受天气的影响, 可以提供全天时全天候的图像, 可有效

用于提取地表信息。本研究采用的 Sentinel-1 GRD 级数据 (包含聚焦的合成孔径雷达图像, 且进行过多视处理) 来源于美国航空航天局 (<https://search.asf.alaska.edu/>), 成像时间为 2022 年 10 月 30 日。哨兵二号 (Sentinel-2) 卫星是高分辨率多光谱成像卫星, 携带一枚多光谱成像仪, 成像幅宽为 290 km, 主要用于全球环境监测。本研究采用

的 Sentinel-2 L2A 级遥感影像来源于 Google Earth Engine 平台。Google Earth Engine 是由谷歌推出的专门用于地球数据运算的遥感云平台，提供全球近几十年的 Sentinel、Landsat TM/OLI 等多源多尺度的卫星影像，还提供诸如气象、土地分类等多方面遥感数据集 (<https://code.earthengine.google.com/>)。选取成像时间为 2022 年 10 月 30 日的 Sentinel-2 遥感影像，覆盖 13 个光谱波段，包含 10 m、20 m 和 60 m 的 3 种空间分辨率，在监测地表信息方面具有显著优势。

先进陆地观测卫星 (advanced land observation satellite, ALOS) 是日本发射的陆地观测卫星，其搭载的全色遥感立体测绘仪分别用于星下点、前视和后视观测，观测数据用于生成 80°N 和 80°E 之间的高精度全球数字高程模型^[20]。选取两景覆盖研究区范围的 ALOS DEM 数据 (<https://search.asf.alaska.edu/>) 作为侵蚀沟提取的地形数据，空间分辨率为 12.5 m，能够较为准确地体现出研究区真实的地形地貌。

1.2.2 数据预处理

研究采用的 Sentinel-1 GRD 级产品为未经处理的原始影像，该数据是经过 WGS-84 椭球投影至地面的聚焦数据，需要对其进行预处理。预处理采用欧空局提供的开源软件 SNAP，更新 Sentinel-1 数据的轨道信息并对其进行地形校正，去除数据中的热噪声和相干斑噪声，并将振幅和相位信息转化为后向散射系数^[21-22]。Sentinel-2 L2A 级数据已经过几何精校正、UTM 投影、大气校正等处理，因此仅需对其重采样。Sentinel-2 卫星影像包含 10 m、20 m 和 60 m 3 种空间分辨率，重采样将 13 个波段的光谱分辨率统一为 10 m，为波段计算提供基础。对两景 DEM 数据裁剪镶嵌并滤波处理，过滤掉过高的高频信息，减少噪声对后续研究的影响。

2 基于面向对象的侵蚀沟提取方法

采用基于多源数据的侵蚀沟面向对象提取模型对德惠市侵蚀沟信息进行提取。主要步骤包括：①影像分割，对遥感影像进行多尺度分割，基于像元的同质性创建对象；②特征提取，基于 Sentinel-1 雷达数据提取强度和纹理特征，基于 Sentinel-2 遥感影像提取光谱和指数特征，基于研究区 DEM 提取地形特征，基于多尺度分割结果提取分割对象的

几何特征，构建对象特征集；③基于随机森林的特征优选和侵蚀沟提取，根据基尼指数对特征集进行特征优选，剔除多余特征，构建最佳特征集并对研究区侵蚀沟进行随机森林提取；④精度评价，利用混淆矩阵评估侵蚀沟提取精度，选取基于对象的总体精度、用户精度、生产者精度和 Kappa 系数作为评价标准。研究框架如图 2 所示。

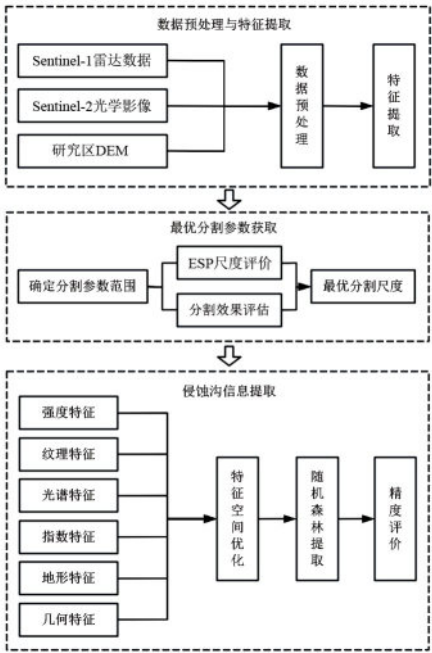


图 2 研究框架

Fig. 2 Research framework

2.1 多尺度分割

面向对象是目前基于高空间分辨率遥感影像广泛应用的影像分析方法。其核心思想是将同质性较高的像元组合为一个对象，并以此对象为分析单元，根据对象的特征进行影像分析。该方法的关键在于通过多个输入特征将影像分割成相对均匀的对象，其中涉及对影像分割方法的选择和分割参数的确定^[23-24]。

多尺度分割 (multiresolution segmentation) 是一种通过合并异质性增量最小的相邻对象实现对象增长，并以异质性增量最小值超过给定阈值为标志结束分割的方法，常用于复杂地区的分割或者是被分割的目标较为复杂的情况^[25]。多尺度分割的主要分割参数包括：波段权重、形状因子、紧致度因子和分割尺度。将 Sentinel-1 和 Sentinel-2 卫星影

像分别进行主成分分析, 选取获得的第一主成分和研究区 DEM 数据作为分割的参考层, 进行多次多尺度分割, 确定适用于研究区的形状因子和紧致度因子, 并利用 ESP 工具获取研究区的最佳分割尺度。

尺度参数估计工具 (estimation of scale parameter, ESP) 是由 Dragut 等^[26]提出的计算最优分割尺度参数的工具。通过计算在不同的分割尺度下所得到的对象同质性的局部变化, 并以此作为所分割的对象层的标准差, 从而用于判断分割效果的好坏。每一个形状因子与紧致度因子的组合对应一个最优分割尺度参数, 当局部变化的变化率值 (ROC-LV) 达到最大值时, 该点所对应的尺度参数为最优分割尺度参数。

2.2 特征提取与优选

2.2.1 特征提取

特征是描述地物信息的指标性表达, 可通过对象中各像元的灰度值与空间拓扑关系获得。对象特征的选择及准确描述, 是提取地物信息的关键环节^[27]。本文基于多尺度分割得到最优分割对象, 构建对象特征集。特征集包括基于 Sentinel-1 遥感影像的强度特征和灰度共生矩阵 (grey-level co-occurrence matrix, GLCM) 的纹理特征、基于 Sentinel-2 遥感影像的光谱特征和指数特征、基于研究区 DEM 数据的地形特征以及基于多尺度分割对象的几何特征 (表 1)。

(1) 强度特征

Sentinel-1 遥感影像经预处理后得到 VH 和 VV 极化通道的后向散射系数。后向散射系数是 SAR 地表覆被识别的重要特征值, VV/VH 极化的雷达后向散射系数通常受土壤和植被的水分含量、植被类型、植被的覆盖程度及地表粗糙度和介电常数等因素的影响而呈现出不同的数值, 因此不同地物在 Sentinel-1 遥感影像中后向散射系数之间的特征差异是进行地物信息提取的重要参考依据^[28]。研究区内建筑物、水体、耕地等地物的地表粗糙度相较于侵蚀沟有显著区别, 后向散射系数存在明显差异, 因此将不同极化通道 (VV 和 VH) 的后向散射系数作为强度特征参与到侵蚀沟提取^[29-30]。

(2) 纹理特征

纹理特征描述了遥感影像中地物的空间灰度信息, 能够综合地体现遥感影像的表面性质, 可作为

表 1 特征提取
Table 1 Feature extraction

特征类型	特征名称	简称
强度特征	同极化后向散射系数	VV
	交叉极化后向散射系数	VH
纹理特征	均值	GL_M
	方差	GL_V
	熵	GL_E
	同质性	GL_H
	对比度	GL_Con
	非相似性	GL_D
	角二阶矩	GL_A
	相关性	GL_Cor
光谱特征	波段均值	Mean
	标准差	Sd
	波段最大差异	MaxDiff
	亮度	Brightness
指数特征	归一化植被指数	NDVI
	差值植被指数	DVI
	比值植被指数	RVi
	土壤调节植被指数	SAVI
	增强型植被指数	EVI
	水体指数	MNDWI
	土壤亮度指数	BI2
地形特征	坡度	Slope
	坡向	Aspect
	剖面曲率	PC
	山体阴影	HS
几何特征	面积	Area
	长宽比	LW
	长度	Length
	宽度	Width
	形状指数	SI
	不对称性	Asymmetry
	圆度	Roundness
	紧致度	Compactness
	矩形拟合	RecF

遥感影像分类的参考依据, 而 SAR 影像中就含有丰富的纹理信息。灰度共生矩阵是一种常用的提取影像纹理特征的方法, 能够较好地反映灰度值的分布特征和纹理的粗细程度^[31]。本文基于 Sentinel-1

VH、VV 双极化数据,选择 3×3 窗口大小,计算均值 (mean)、方差 (variance)、同质性 (homogeneity)、对比度 (contrast)、相异性 (dissimilarity)、熵 (entropy)、角二阶矩 (angular second moment) 和相关性 (correlation) 8 种常用的纹理特征,共提取 16 个纹理特征。

(3) 光谱特征

光谱特征是多光谱遥感数据进行地表信息提取的最基本特征之一。不同的地物由于自身理化性质不同,存在不同的反射率,这是光谱特征识别地物的基础。选取经过处理后的德惠市 Sentinel-2 遥感影像,用以分析侵蚀沟与研究区其他地物在影像光谱反射率上的差异和不同波段的可分离度。提取 Sentinel-2 影像波段中的 blue、green、red、vegetation red edge-1、vegetation red edge-2、vegetation red edge-3、NIR、vegetation red edge-4、water vapor、SWIR-1 和 SWIR-2 的波段均值、标准差、波段最

大差异和亮度作为光谱特征。

(4) 指数特征

通过 Sentinel-2 影像不同波段之间进行线性或非线性组合计算可以得到若干指数特征。指数特征可以突显出遥感影像中目标地物的数量、质量、时空分布和生长状况等信息,被广泛应用于土地状况监测、地表信息提取等方面^[32]。Sentinel-2 遥感影像共有 13 个波段,具有丰富的光谱信息,因此可以用于计算指数特征。本研究选取归一化植被指数 (normalized difference vegetation index, NDVI)、差值植被指数 (difference vegetation index, DVI)、比值植被指数 (ratio vegetation index, RVI)、土壤调节植被指数 (soil-adjusted vegetation index, SAVI)、增强型植被指数 (enhanced vegetation index, EVI)、水体指数 (modified normalized difference water index, MNDWI) 和土壤亮度指数 (the second brightness index, BI2) 共 7 个指数特征。

表 2 指数特征
Table 2 Index features

指数特征	计算公式	相关描述
NDVI	$\frac{\rho_{\text{NIR}} - \rho_{\text{Red}}}{\rho_{\text{NIR}} + \rho_{\text{Red}}}$	植被覆盖度和生长状态
DVI	$\rho_{\text{NIR}} - \rho_{\text{Red}}$	土壤背景敏感
RVI	$\frac{\rho_{\text{NIR}}}{\rho_{\text{Red}}}$	区分绿色与非绿色植被
SAVI	$\frac{\rho_{\text{NIR}} - \rho_{\text{Red}}}{\rho_{\text{NIR}} + \rho_{\text{Red}} + L} * (1 + L)$	减小土壤背景影响
EVI	$\frac{\rho_{\text{NIR}} - \rho_{\text{Red}}}{\rho_{\text{NIR}} + 6\rho_{\text{Red}} - 7.5\rho_{\text{Blue}} + 1} * 2.5$	提高对植被的敏感度
MNDWI	$\frac{\rho_{\text{Green}} - \rho_{\text{SWIR}}}{\rho_{\text{Green}} + \rho_{\text{SWIR}}}$	区分水体与其他地物
BI2	$\sqrt{\frac{(\rho_{\text{Red}} * \rho_{\text{Red}}) + (\rho_{\text{Green}} * \rho_{\text{Green}}) + (\rho_{\text{NIR}} * \rho_{\text{NIR}})}{3}}$	对土壤的亮度敏感

(5) 地形特征

侵蚀沟由沟头、沟沿、沟底等部分组成,其形态、分布等特征通常与该地区的地形有较强的相关性。地表由于坡度坡向的不同,传感器接收的太阳辐射能量也会有所差异。本研究基于研究区 DEM 数据计算得出坡度、坡向、剖面曲率和山体阴影等特征,并对计算得到的地形属性指标进行归一化,将其作为提取侵蚀沟的地形特征。

(6) 几何特征

侵蚀沟为线形伸展的槽形凹地,通常与地表的水文网或其他侵蚀沟相联,形成树枝状分布的系统。因其在卫星影像上具有独特的形态特点,提取侵蚀沟的几何特征可增加侵蚀沟识别的准确性。本研究基于多尺度分割结果,提取面积、长宽比、长度、宽度、形状指数、不对称性、圆度、紧致度和矩形拟合共计 9 个几何特征参与侵蚀沟提取。

2.2.2 特征优选

遥感信息提取过程中,特征过多会提高椒盐现象发生的概率,对提取精度造成影响,因此需要对特征进行优选。随机森林可以对每个特征变量的重要程度进行评估和排序,以此达到特征优选的目的。其主要思想是比较特征变量每颗树贡献度的平均值,其中贡献度采用基尼指数(Gini index, GI)作为衡量标准^[33]。

采用基尼指数评价特征过程如下:假设样本集 S 中共有 M 个特征和 N 个类别,决策树共有 K 棵,则第 k 棵树节点 q 的基尼指数 GI 为:

$$GI_q^{(k)} = 1 - \sum_{n=1}^N (P_{qn})^2 \quad (1)$$

式中: P_{qn} 表示类别 n 在节点 q 中所占比例;节点 q 分枝后生成两个新节点,两个新节点的基尼指数分别为 $GI_q^{(k)}$ 和 $GI_r^{(k)}$,则分枝前后基尼指数变化量 $V^{(k)}$ 为:

$$V^{(k)} = GI_q^{(k)} - GI_L^{(k)} - GI_r^{(k)} \quad (2)$$

将每棵树每个节点的变化量求和,所得数值即为该特征的重要性评分。将所有特征的重要性评分归一化处理,即可获得各个特征的贡献度。

2.3 随机森林

随机森林(random forest, RF)是由 Breiman^[34]提出的一种由多个决策树组成的分类器,是对传统决策树的改进。目前,随机森林算法广泛用于信息提取、回归和降维等领域,其基本过程如下:从原始训练样本集中以 Bootstrap 的方式有放回的抽取 N 次,共得到 N 个训练样本集,未抽到的部分则为袋外数据(out of bag, OOB),用以分类后误差估计,称为袋外误差^[35];假设每个训练样本集有 M 个特征,根据 N 个训练样本集构建决策树模型,在每个节点处从总特征中随机选择 m 个特征($m \leq M$),对决策树节点采用完全分裂法进行划分,共得到 N 个决策树模型;基于集成学习理论,将 N 个决策树模型以相同的权重线性组合,未分类样本的类别由 N 个决策树分类结果以多数表决的方式投票决定^[36-37]。

相较于传统的决策树分类方法,随机森林算法具有如下优势:①数据挖掘能力更强,准确率更

高;②无需对数据进行规范化,对数据要求较为宽泛;③抗噪声能力强,过拟合现象产生的概率较低。

2.4 精度评价

利用混淆矩阵评估侵蚀沟提取结果的准确性,以多尺度分割对象作为最小样本,采用总体精度(OA)、生产者精度(PA)、用户精度(UA)和 Kappa 系数 4 个评价指标。

总体精度表示所有类别中正确提取的像元数量与像元总数的百分比,体现总体的正确程度:

$$OA = \frac{\sum_{i=1}^n x_{ii}}{N} \quad (3)$$

生产者精度表示某类别中正确提取的像元数量与该类别验证样本的像元总数的百分比,表示验证样本被正确识别的概率:

$$PA = \frac{x_{ii}}{x_{+i}} \quad (4)$$

用户精度表示某一类中正确提取的像元数量与该类别像元总数的百分比,可判断提取的可靠性:

$$UA = \frac{x_{ii}}{x_{i+}} \quad (5)$$

Kappa 系数是以所有像元为参考判断提取精度,表示提取结果与实际类别间的匹配程度;

$$Kappa = \frac{N \sum_{i=1}^n x_{ii} \sum_{i=1}^n (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^n (x_{i+} x_{+i})} \quad (6)$$

式(3)~(6)中: x_{ii} 为混淆矩阵中第 i 行、第 i 列的元素,即正确分类为 i 类的样本数; x_{+i} 和 x_{i+} 分别为混淆矩阵中第 i 列和第 i 行元素的总和; n 为总类别数(个); N 为用于参考的样本数(个)。

3 研究结果

3.1 多尺度分割结果

分割时尺度参数大小不同,得到的对象大小以及数量也不相同,分割尺度参数设置的值越小,分割得到的对象数目则越多,边界也就越破碎,反之

亦然。多尺度分割方法分割地物时不可避免会产生过分割或欠分割的情况，考虑到研究区侵蚀沟宽度普遍较窄，相较之下过分割比欠分割更适合提取目标信息^[38]。

确定研究区的最佳分割尺度的前提是判断适合目标地类的分割尺度区间。经过多次重复实验，综合分割地物的视觉效果和特征边界的实际情况，各类地物分割表现均较好的分割尺度的区间为 100 ~ 150，此时形状参数为 0.5，紧致度参数为 0.7。将该尺度参数范围输入 ESP2 工具，计算出该区间内分割对象同质性的局部变化。当 ROC-LV 取最大值时，该点对应分割尺度为 125，即在分割尺度 100 ~ 150 的区间内，该遥感影像的最佳分割尺度为 125。

3.2 特征优选结果

采用随机森林进行特征重要性评估，主要思想是比较特征变量每棵树贡献度的平均值。为避免特征冗余对侵蚀沟提取结果造成影响，基于随机森林基尼指数对由强度、纹理、光谱、指数、地形和几何特征构建的特征集进行特征贡献度排序，选取排名前 30 的特征作为最佳特征组合输入随机森林提取模型。优选结果如表 3 所示，即 VV、VH 等共计 30 个特征。

表 3 特征优选结果

Table 3 Feature optimization results

特征类别	特征名称
强度特征	VV、VH
纹理特征	GL_M_VV、GL_M_VH、GL_V_VV、GL_V_VH、GL_Con_VV、GL_Con_VH、GL_Cor_VV
光谱特征	Mean_Red、Mean_Blue、Mean_Green、Mean_NIR、Mean_VRE4、MaxDiff、Brightness
指数特征	NDVI、SAVI、MNDWI、BI2、EVI、RVI
地形特征	Slope、Aspect、PC
几何特征	LW、SI、Roundness、Compactness、RecF

3.3 侵蚀沟提取结果

利用目视解译法，结合研究区实地调查数据选取样本数据集。选取样本的过程中采用随机选取的方式，选取面积较大、形态特征较为明显的侵蚀沟分割对象作为样本，以保证训练样本准确可靠。随机森林模型的控制参数主要有树的个数（nTree）和每次分裂变量个数（nTry），设置 nTree 和 nTry

分别为 1 000 和 6，将训练样本和最佳特征集输入随机森林模型，对研究区侵蚀沟进行提取。图 3 为研究区内侵蚀沟信息的提取结果，目视解译可见研究区中部和西南部的侵蚀沟提取效果较好，在边界处则有较多误分地物。

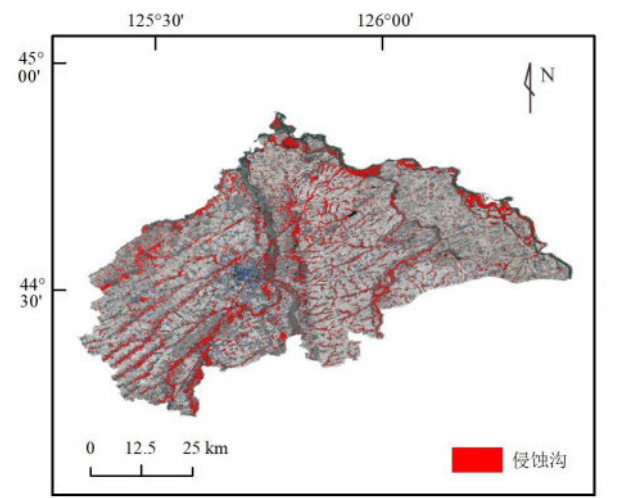


图 3 侵蚀沟提取结果
Fig. 3 Extraction results of erosion gully

3.4 侵蚀沟提取精度

研究采用随机样本点对提取结果进行精度分析。样本点以随机抽样为原则，考虑到整个研究区的面积大小，随机均匀抽取共 1 000 个样本点，样本点分布如图 4 所示。对每一个随机样本点进行目视解译判读，将样本点手动分为侵蚀沟与非侵蚀

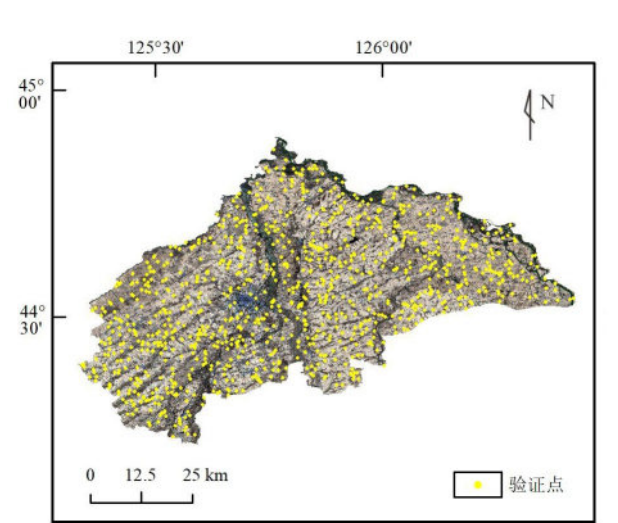


图 4 验证点分布图
Fig. 4 Distribution of validation points

沟，输入混淆矩阵进行精度分析。精度分析结果如表 4 所示，整个区域的侵蚀沟提取的总体精度为 94.03%，Kappa 系数为 0.839。其中，侵蚀沟地类的用户精度为 95.63%，生产者精度为 81.24%，提取精度较高。

表 4 侵蚀沟提取精度

Table 4 Extraction accuracy of erosion gully				
类别	PA/%	UA/%	OA/%	Kappa 系数
侵蚀沟	95.63	81.24	94.03	0.839
非侵蚀沟	93.56	98.66		

4 讨论

4.1 对比分析

基于多源数据提取目标信息可为随机森林模型提供大量用作参考的特征值，相较于单一数据源可提供更多特征信息。为了进一步探究多源数据融合对侵蚀沟提取精度的影响，基于相同的样本数据集，对不同数据源的侵蚀沟提取精度进行对比（表 5）。由表 5 可见，基于多源数据融合的侵蚀沟提取精度有着明显提高，证明了其优越性。

表 5 不同数据源提取精度对比

Table 5 Comparison of extraction accuracy of different data sources

数据源	OA/%	Kappa 系数
Sentinel-2	88.75	0.751
Sentinel-2 + DEM	89.97	0.804
Sentinel-2 + Sentinel-1	92.89	0.822
Sentinel-2 + Sentinel-1 + DEM	94.03	0.839

4.2 侵蚀沟空间分布分析

侵蚀沟由沟头、沟沿、沟底等部分组成，并常表现为沟壁坍塌，沟底下切等形式，多数侵蚀沟分布在耕地中，致使耕地支离破碎^[39-40]。吉林省德惠市为侵蚀沟多发地区，提取结果表明，侵蚀沟主要发生在研究区中部和西南部，侵蚀整体呈东北—西南带状走向，多分布于居民地附近并与一定的水系相连。研究区侵蚀沟多为发育程度较低的细沟、浅沟，整体侵蚀程度较轻，但研究区属温带大陆性

季风气候区，在平均气温较高且降雨量偏少的自然环境下，细沟和浅沟易发育成相对难以治理的切沟。切沟通常切破耕地犁底层，深入成土母质甚至疏松基岩，对耕地危害极大，因此对研究区的侵蚀情况仍需加以重视。

5 结论

（1）多源数据融合在一定程度上能够提高侵蚀沟信息提取的精度。基于 Sentinel-1 雷达数据、Sentinel-2 光学数据和地形数据融合，结合强度、纹理、光谱、指数、地形和几何特征等多种特征，采用基于面向对象影像分析的随机森林算法，以吉林省德惠市为研究区对侵蚀沟信息进行提取，总体精度为 94.03%，Kappa 系数为 0.839。相较于单一数据源，研究区侵蚀沟提取的总体精度和 Kappa 系数均有明显提升，说明多源数据可为随机森林模型提供大量特征值，可以显著提高侵蚀沟提取的准确度。

（2）德惠市中部的土壤侵蚀较为严重，侵蚀沟多为发育程度较低的细沟、浅沟。侵蚀整体呈东北—西南走向，多分布于居民地附近并与一定的水系相连，提取结果可为研究侵蚀沟产生机理和未来预测提供参考，后续可结合高分辨率遥感影像开展进一步的研究。

参考文献：

[1] 李全峰, 李吉程, 于明鹏, 等. 东北典型黑土区耕地侵蚀沟形态系统分类方法构建 [J]. 农业工程学报, 2024, 40 (14): 53-61.
LI Q F, LI J C, YU M P, et al. Systematic morphological classification method for cultivated land erosion gullies in the typical black soil regions of Northeast China [J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40 (14): 53-61.

[2] HUANG D, SU L, FAN H, et al. Identification of topographic factors for gully erosion susceptibility and their spatial modelling using machine learning in the black soil region of Northeast China [J]. Ecological Indicators, 2022, 143: 109376.

[3] 李珊珊, 王春梅, 唐杰, 等. 东北黑土区典型小流域浅沟向切沟发育及其地形临界研究 [J/OL]. 武汉大学学报 (信息科学版), 1-16 [2024-10-08]

- <https://doi.org/10.13203/j.whugis20230493>.
- LI S S, WANG C M, TANG J, et al. Study on the development of ephemeral to permanent gullies and the critical terrain threshold in typical watershed in the mollisol region of northeastern China [J]. *Geomatics and Information Science of Wuhan University*, 1–16 [2024–10–08] <https://doi.org/10.13203/j.whugis20230493>.
- [4] 李朋飞, 黄珂瑶, 胡晋飞, 等. 黄土丘陵沟壑区细沟发育形态的变化及其与侵蚀产沙的关系 [J]. *农业工程学报*, 2022, 38 (18): 92–102.
- LI P F, HUANG K Y, HU J F, et al. Morphological development of rills and its relationship with hillslope erosion in the hilly and gully Loess Plateau [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38 (18): 92–102.
- [5] WANG J, ZHANG Y, LI K, et al. Gully internal erosion triggered by a prolonged heavy rainfall event in the tableland region of China's Loess Plateau [J]. *International Soil and Water Conservation Research*, 2023, 11 (4): 610–621.
- [6] MCINNES J, VIGIAKI O, ROBERTS A M. Using Google Earth to map gully extent in the West Gippsland region (Victoria, Australia) [C] // *Proceedings of 19th International Congress on Modelling and Simulation*, Perth, Australia. 2011: 12–16.
- [7] GUAN Y, YANG S, ZHAO C, et al. Monitoring long-term gully erosion and topographic thresholds in the marginal zone of the Chinese Loess Plateau [J]. *Soil and Tillage Research*, 2021, 205: 104800.
- [8] 陈靖涛, 史明昌, 罗志东, 等. 基于双向地形阴影法的黄土侵蚀沟自动提取技术 [J]. *农业工程学报*, 2022, 38 (7): 127–135.
- CHEN J T, SHI M C, LUO Z D, et al. Automatic extraction technique for the erosion gully in the Loess Plateau based on bidirectional relief-shading method [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38 (7): 127–135.
- [9] 姜芸, 王军, 张莉. 东北典型黑土区侵蚀沟形态及分布特征 [J]. *农业工程学报*, 2020, 36 (7): 157–165.
- JIANG Y, WANG J, ZHANG L. Morphology and distribution characteristics of erosion gully in the typical black soil region of Northeast China [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2020, 36 (7): 157–165.
- [10] 李思进, 代文, 熊礼阳, 等. DEM分辨率对黄土侵蚀沟形态特征表达的不确定性分析 [J]. *地球信息科学学报*, 2020, 22 (3): 338–350.
- LI S J, DAI W, XIONG L Y, et al. Uncertainty of the morphological feature expression of loess erosional gully affected by DEM resolution [J]. *Journal of Geo-Information Science*, 2020, 22 (3): 338–350.
- [11] 董一帆, 聂勇, 熊东红. 基于 Google Earth 影像的横断山区沟蚀及侵蚀沟类型调查研究 [J]. *水土保持通报*, 2018, 38 (3): 151–157.
- DONG Y F, NIE Y, XIONG D H. Investigation of gully density and classification of Hengduan Mountainous area based on Google Earth images [J]. *Bulletin of Soil and Water Conservation*, 2018, 38 (3): 151–157.
- [12] LIU B, ZHANG B, FENG H, et al. Ephemeral gully recognition and accuracy evaluation using deep learning in the hilly and gully region of the Loess Plateau in China [J]. *International Soil and Water Conservation Research*, 2022, 10 (3): 371–381.
- [13] COELHO I M, COELHO V N, LUZ E J S, et al. A GPU deep learning meta heuristic based model for time series forecasting [J]. *Applied Energy*, 2017, 201: 412–418.
- [14] 吴艳双, 张晓丽. 结合多尺度纹理特征的高光谱影像面向对象树种分类 [J]. *北京林业大学学报*, 2020, 42 (6): 91–101.
- WU Y S, ZHANG X L. Object-oriented tree species classification with multi-scale texture features based on airborne hyperspectral images [J]. *Journal of Beijing Forestry University*, 2020, 42 (6): 91–101.
- [15] KARAMI A, KHOORANI A, NOOHEGAR A, et al. Gully erosion mapping using object-based and pixel-based image classification methods [J]. *Environmental & Engineering Geoscience*, 2015, 21 (2): 101–110.
- [16] D'OLEIRE-OLTMANN S, MARZOLFF I, TIEDE D, et al. Detection of gully-affected areas by applying object-based image analysis (OBIA) in the region of Taroudannt, Morocco [J]. *Remote Sensing*, 2014, 6 (9): 8287–8309.
- [17] 王文娟, 邓荣鑫, 张树文. 东北典型黑土区40年来沟蚀空间格局变化及地形分异规律 [J]. *地理与地理信息科学*, 2012, 28 (3): 68–71.
- WANG W J, DENG R X, ZHANG S W. Spatial pattern

- change and topographic differentia of gully erosion in the type black soil area of Northeast China during the past 40 years [J]. *Geography and Geo-Information Science*, 2012, 28 (3): 68-71.
- [18] HUANG X, XIONG L Y, JIANG Y H, et al. Mapping gully affected areas by using Sentinel 2 imagery and digital elevation model based on the Google Earth Engine [J]. *Catena*, 2023, 233: 107473.
- [19] SHRUTHI R B V, KERLE N, JETTEN V. Object-based gully feature extraction using high spatial resolution imagery [J]. *Geomorphology*, 2011, 134 (3/4): 260-268.
- [20] ZHANG K, GANN D, ROSS M, et al. Accuracy assessment of ASTER, SRTM, ALOS, and TDX DEMs for Hispaniola and implications for mapping vulnerability to coastal flooding [J]. *Remote sensing of Environment*, 2019, 225: 290-306.
- [21] 杜伟娜, 徐爱功, 宋耀鑫, 等. 新型 SAR 传感器一级地距产品绝对辐射定标方法 [J]. *国土资源遥感*, 2016, 28 (4): 30-34.
- DU W N, XU A G, SONG Y X, et al. Absolute radiometric calibration of level-1 detected ground range products of new SAR sensors [J]. *Remote Sensing for Land and Resources*, 2016, 28 (4): 30-34.
- [22] 郭山川, 杜培军, 蒙亚平, 等. 时序 Sentinel-1A 数据支持的长江中下游汛情动态监测 [J]. *遥感学报*, 2021, 25 (10): 2127-2141.
- GUO S C, DU P J, MENG Y P, et al. Dynamic monitoring on flooding situation in the middle and lower reaches of the Yangtze River Region using Sentinel-1A time series [J]. *National Remote Sensing Bulletin*, 2021, 25 (10): 2127-2141.
- [23] AVCI Z D U, KARAMAN M, OZELKAN E, et al. OBIA based hierarchical image classification for industrial lake water [J]. *Science of the Total Environment*, 2014, 487: 565-573.
- [24] ANDERSON C J, HEINS D, PELLETIER K C, et al. Improving machine learning classifications of *Phragmites australis* using object-based image analysis [J]. *Remote Sensing*, 2023, 15 (4): 989.
- [25] XING H, NIU J, FENG Y, et al. A coastal wetlands mapping approach of Yellow River Delta with a hierarchical classification and optimal feature selection framework [J]. *Catena*, 2023, 223: 106897.
- [26] DRAGUT L, TIEDE D, LEVICK S R. ESP: a tool to estimate scale parameter for multi-resolution image segmentation of remotely sensed data [J]. *International Journal of Geographical Information Science*, 2010, 24 (6): 859-871.
- [27] 于佩鑫, 周询, 刘素红, 等. 东北黑土区侵蚀沟遥感影像特征提取与识别 [J]. *遥感学报*, 2018, 22 (4): 611-620.
- YU P X, ZHOU X, LIU S H, et al. Feature extraction and recognition of erosion gully based on remote sensing image in the black soil region in Northeast China [J]. *Journal of Remote Sensing*, 2018, 22 (4): 611-620.
- [28] 徐国良, 卢凌莹, 杨璨, 等. 基于多时相 Sentinel-1A 的丘陵山区耕地非粮化特征识别及驱动机制 [J]. *农业工程学报*, 2023, 39 (22): 236-245.
- XU G L, LU L Y, YANG C, et al. Identification and driving mechanisms of non-grain cultivated land in hilly and mountainous areas based on multi-temporal Sentinel-1A images [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2023, 39 (22): 236-245.
- [29] 张廷龙, 韩晓乐, 包懿, 等. 基于 Sentinel-1/2 数据融合的县域农业大棚提取 [J]. *农业工程学报*, 2024, 40 (19): 135-145.
- ZHANG T L, HAN X L, BAO Y, et al. Extracting county-level agricultural greenhouses using Sentinel-1/2 data fusion [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2024, 40 (19): 135-145.
- [30] SAFARI M M, MALIAN A. Plant disease mapping in paddy growing stages using remotely sensed data [J]. *Environmental Earth Sciences*, 2025, 84 (1): 1-19.
- [31] 王镭, 赵红莉, 蒋云钟, 等. 月尺度农作物提取中 GF-1 WFV 纹理特征的应用及分析 [J]. *自然资源遥感*, 2021, 33 (3): 72-79.
- WANG R, ZHAO H L, JIANG Y Z, et al. Application and analyses of texture features based on GF-1 WFV images in monthly information extraction of crops [J]. *Remote Sensing for Natural Resources*, 2021, 33 (3): 72-79.
- [32] 朱钟正, 陈玉福, 朱文泉, 等. 适用于多目标遥感自动解译的最佳专题指数筛选 [J]. *遥感技术与应用*, 2017, 32 (3): 564-574.
- ZHU Z Z, CHEN Y F, ZHU W Q, et al. Selection of optimum thematic indices applied to the multi-objective and automatic remote sensing interpretation [J]. *Re-*

- Remote Sensing Technology and Application, 2017, 32 (3): 564-574.
- [33] 姚茂林, 江洪, 张丽玉. 基于 Sentinel-2 影像红边光谱指数与特征优选的竹林提取研究 [J]. 海南大学学报 (自然科学版), 2022, 40 (4): 373-381.
YAO M L, JIANG H, ZHANG L Y. Bamboo information extraction from Sentinel-2 image based on improved spectral indices and random forest Gini index [J]. Natural Science Journal of Hainan University, 2022, 40 (4): 373-381.
- [34] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45: 5-32.
- [35] 何云, 黄翀, 李贺, 等. 基于 Sentinel-2A 影像特征优选的随机森林土地覆盖分类 [J]. 资源科学, 2019, 41 (5): 992-1001.
HE Y, HUANG C, LI H, et al. Land-cover classification of random forest based on Sentinel-2A image feature optimization [J]. Resources Science, 2019, 41 (5): 992-1001.
- [36] CHEN X, ZHANG Y, WANG S, et al. Comparative study of machine learning methods for mapping forest fire areas using Sentinel-1B and 2A imagery [J]. Frontiers in Remote Sensing, 2024, 5: 1446641.
- [37] FU Z, WU B, ZENG H, et al. Mapping desert shrubs in Inner Mongolia using universal medium resolution satellite imagery: development of new spectral indices [J]. GIScience & Remote Sensing, 2025, 62 (1): 2428366.
- [38] 白韬, 杨国东, 王凤艳, 等. 一种面向对象的最优分割尺度计算模型 [J]. 吉林大学学报 (地球科学版), 2020, 50 (1): 304-312.
BAI T, YANG G D, WANG F Y, et al. Object-oriented optimal segmentation scale calculation model [J]. Journal of Jilin University (Earth Science Edition), 2020, 50 (1): 304-312.
- [39] 宋爽, 范昊明, 牛天一. 东北黑土区不同季节侵蚀沟形态发育比较分析 [J]. 水土保持学报, 2022, 36 (1): 18-23.
SONG S, FAN H M, NIU T Y. Comparative analysis of the development the geometric forms of gully in different seasons in black soil area of Northeast China [J]. Journal of Soil and Water Conservation, 2022, 36 (1): 18-23.
- [40] FANG H, GUO M. Aspect-induced differences in soil erosion intensity in a gullied hilly region on the Chinese Loess Plateau [J]. Environmental Earth Sciences, 2015, 74: 5677-5685.