

基于卷积神经网络的景象匹配区 适配性筛选方法

李邦杰¹, 武 健¹, 张大巧¹, 朱宏伟¹, 胡超帝², 白双豪¹

(1. 火箭军工程大学, 陕西 西安 710025;

2. 96728部队, 青海 格尔木 816099)

摘 要: 针对基于图像特征指标的景象匹配区适配性筛选方法效率低、景象匹配区规划周期长的问题, 提出一种基于卷积神经网络(Convolutional Neural Network, CNN)的适配性筛选算法。首先以典型的城市、荒漠、农田等6类景象为对象, 构建包括灰度均方差、高程方差、独立像元数和最大峰值比等的图像特征指标, 形成相关数据集; 设计基于CNN的景象匹配区适配性筛选模型, 并分析了迭代次数、训练样本数量、像元数等对模型计算精度的影响; 最后选取合适的迭代次数和全连接层神经元数, 利用数据集验证了本文算法。结果表明: 所提方法可以有效提升大规模典型景象的匹配适配性筛选效率。

关键词: 景象匹配; 卷积神经网络; 图像特征; 适配性筛选

中图分类号: TJ410

文献标志码: A

文章编号: 2097-2741(2024)03-045-007

A CNN-based Scene Matching Adaptability Screening Method

LI Bangjie¹, WU Jian¹, ZHANG Daqiao¹, ZHU Hongwei¹,

HU Chaodi², BAI Shuanghao¹

(1. Rocket Force University of Engineering, Xi'an 710025, Shaanxi;

2. Troops No. 96728, Golmud 816099, Qinghai)

ABSTRACT: In order to solve problems of low efficiency and long planning period of image characteristic index-based scene matching area adaptability screening method, an adaptability screening algorithm based on convolutional neural network (CNN) was proposed. By taking six types of typical scenes, such as cities, deserts, and farmlands as objects, image characteristic indexes including gray mean square deviation, elevation variance, number of independent pixels, and maximum peak ratio were constructed to form relevant datasets. Then, a scene matching area adaptability screening model of classified CNN was designed to analyze the influence of iteration times, training sample number and pixel number on the accuracy of model computation. Accordingly, by selecting proper iteration times and numbers of fully connected layer neurons, the proposed algorithm was verified on datasets. Results showed that the proposed method can effectively improve the efficiency of matching and adaptability screening of typical large-scale scenes.

收稿日期: 2024-03-27

第一作者: 李邦杰 (1978—), 男, 教授, 主要从事飞行器总体设计与应用研究。E-mail: ro.soldier@outlook.com

引用格式: 李邦杰, 武健, 张大巧, 等. 基于卷积神经网络的景象匹配区适配性筛选方法[J]. 火箭军工程大学学报, 2024, 38 (3): 45-50, 59.

KEYWORDS: scene matching; CNN network; image characteristic; adaptability screening

景象匹配作为一种高精度的辅助导航技术,在导航系统中扮演着至关重要的角色。然而,现有文献在景象匹配区的适配性筛选方面存在一定局限性。传统筛选方法主要依靠基于灰度或特征的匹配算法,如灰度均方差^[1]、高程方差^[2]、线特征^[3]、梯度特征^[4]等指标,这些方法往往与特定的匹配算法紧密相关,缺乏普适性和灵活性。近年来,人工智能的引入为景象匹配区的筛选带来了新的可能性^[5],如基于多尺度共生矩阵^[6]、分块二维主成分分析^[7]及 ResNet 神经网络^[8]等方法。尽管这些方法在提高准确性方面取得了显著成果,但大多侧重于提升适配性筛选的准确性,而对筛选效率的关注不足^[9]。

针对以上问题,本文旨在提高批量选择景象匹配区时的适配性筛选效率。提出了效率优先的景象匹配区适配性筛选模型来提高筛选速度;结合深度学习和模式识别技术,设计了一套适用于不同场景和复杂度的通用筛选框架,以增强筛选方法的普适性。

1 景象匹配区筛选方法及主要评价指标

图像本身特征是传统景象匹配区筛选的基础。评价图像特征的主要指标有:灰度方差、相关长度、独立像元数和最高峰值比。

1.1 主要评价指标

(1) 相关长度。相关长度^[10]是度量图像粗糙程度的一个参数,任何图像像元之间都存在着某种关联。假设一幅图像的灰度值为 $A(i, j)$ ($i = 0, 1, 2, \dots, W-1; j = 0, 1, 2, \dots, H-1$), 则这幅图像的自相关函数定义式^[11]为

$$\rho(i, j) =$$

$$\frac{\sum_{i=1}^W \sum_{j=1}^H [A(i, j) - \bar{A}] [A(i + \Delta x, j + \Delta y) - \bar{A}]}{\sum_{i=1}^W \sum_{j=1}^H [A(i, j) - \bar{A}]^2}$$

式中: Δx 为水平方向位移; Δy 为垂直方向位移; $\bar{A}$ 为图像的平均灰度值; W 为图像宽度; H 为图像高度。相关长度定义为: 当 $\rho(i, j) \leq 0.368$ 时的位移。

(2) 独立像元数。独立像元数用于度量灰度信息源, 表示信息量的大小。独立像元数越多,

所代表的信息量越大。对于二维图像, 设 L_x 、 L_y 分别为水平和垂直方向的相关长度, 则独立像元数定义为

$$N = \frac{W}{L_x} \times \frac{H}{L_y}$$

(3) 灰度方差。灰度方差 δ 是灰度图像灰度值在空间分布上的度量, 表示灰度起伏。通常灰度方差愈大, 信噪比、匹配概率、匹配精度亦更高^[11]。其定义式为

$$\delta = \sqrt{\frac{1}{W \times H} \sum_i^W \sum_j^H (A(i, j) - \bar{A})^2}$$

(4) 最高峰值比。最高峰值比是衡量图像相关峰特征的一种系数, 可基于像素值进行计算^[12]。最高峰值比 S 定义式为

$$S = \frac{V_{\max}}{V_{\text{sub}}}$$

式中: $V_{\max}$ 为最高相关系数值; V_{sub} 为次最高相关系数值。相关系数的计算公式为

$$V = \text{cov}(X, Y) / (\text{std}(X) \times \text{std}(Y))$$

式中: $\text{cov}(X, Y)$ 为矩阵 X 和矩阵 Y 的协方差; $\text{std}(X)$ 和 $\text{std}(Y)$ 分别为矩阵 X 和 Y 的标准差。

1.2 景象匹配区筛选方法

有效的景象匹配区域应具备明显的景物特征、丰富的图像区域信息、良好的唯一性, 以及满足基准图要求的大小特征^[13]。在景象匹配区域的筛选过程中, 主要依赖计算机进行大规模数据处理。具体流程包括将输入的遥感卫星影像分割成多个与景象匹配区域大小相匹配的图像块, 并计算其图像特征值。

传统筛选方法通常采用多种特征指标, 如独立像元数、图像方差、相关峰特征及纹理频谱分布等。基于此, 本文选取上述指标作为景象匹配区域筛选的依据。

2 景象匹配区筛选训练集生成

在基于可见光遥感影像的景象匹配区研究中: 首先, 采用特定算法对原始图像进行预处理, 进而生成多个与预定景象匹配区大小相等的子图像; 随后, 通过对这些子图像进行一系列条件筛选, 最终构建训练集, 该训练集包含适用于景象匹配的阳性区域以及不适宜作为匹配区的阴性区域。

具体步骤包括以下几方面：

- (1) 将预处理后的灰度图像分割为 500 像素 × 500 像素的子图片，作为景象匹配区域的代表；
- (2) 对每一子图片进行特征提取，包括灰度方差、独立像元数、最高峰值比等关键参数；
- (3) 在特征参数解算完成后，依据预设的阈值对上述度量特征进行筛选，以确保所选样本的可靠性和有效性。

通过以上步骤，形成训练样本集，该样本集包含按照可用性分类的城市、荒漠、农田、热带雨林、山地、森林等 6 种景象匹配区专题图，还包含相应的阴性区域，以及每个子图片的详细特征参数。这些数据将为后续的景象匹配算法研究提供实验基础，并有助于提高遥感图像处理的准确性和效率。

3 景象匹配区快速筛选网络模型设计

针对景象匹配区的快速筛选问题，采用卷积神经网络（Convolutional Neural Network, CNN）模型，并利用深度学习框架 Keras 构建网络。在模型训练阶段，通过反向传播算法优化网络权重，旨在提升模型的准确性与泛化能力。所设计的筛选模型结合了对称与非对称模块，包括卷积层、池化层、跳跃连接、辍学层及全连接层，同时采用批处理规范化技术优化激活输入。损失函数选用 Softmax 函数。

模型选择上，以 Inception V3 作为基础，引入 ImageNet 的预训练参数。Inception V3 通过多分支处理特征图并进行多维融合，引入批处理规范化以加速收敛。此外，模型采用多层小卷积核替代大卷积核，减少参数量，并通过辅助分类器辅助更新网络浅层参数。模型架构见表 1。

使用的 Inception V3 预训练模型基于大规模标记图像集 ImageNet 进行训练，该数据集含逾千万标记图像及百万带边界框图像，为对象提供精确定位。ImageNet 数据集共 1 331 167 张图像，分为训练集（1 281 167 张）和评估集（50 000 张），确保了模型训练与评估的独立性。

4 景象匹配区快速筛选模型分析优化

在构建测试集时，精选多类图像样本进行评估。通过比较模型预测的特定类别图像数与实际

总数的比值，计算模型对各图像类别的预测准确率，探讨了参数配置、样本特性及 CNN 结构对模型性能的影响，并采用控制变量法，逐一分析了这些因素在模型训练中的作用。

表 1 Inception V3 的网络架构
Tab. 1 Architectures of Inception V3 networks

类型	内部尺寸	输入尺寸
Conv	(3, 3, 2)	(299, 299, 3)
Conv	(3, 3, 1)	(149, 149, 32)
Padded	(3, 3, 1)	(147, 147, 32)
Pool	(3, 3, 2)	(147, 147, 64)
Conv	(3, 3, 1)	(73, 73, 64)
Conv	(3, 3, 2)	(71, 71, 80)
Conv	(3, 3, 1)	(35, 35, 192)
Pool	(8, 8)	(8, 8, 2 048)

4.1 训练集样本数量影响

使用城市专题图像构建训练集、验证集及测试集。所有图像样本统一调整为 500 像素 × 500 像素的分辨率，训练迭代设定为 40 次，并在全连接层配置 1 024 个神经元。测试集包含 1 000 张不适用于景象匹配的城市图像。图 1 对比了不同规模训练集上训练的模型在相同测试集上的性能。

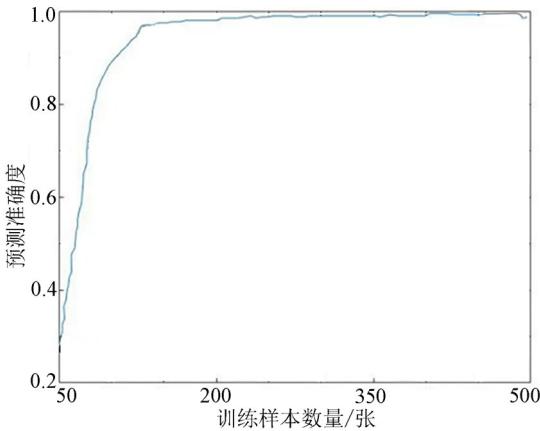


图 1 训练样本数量与模型预测准确率关系曲线

Fig. 1 Curve of the relationship between the number of training samples and the model prediction accuracy

由图 1 可以看出，随着训练集规模的增加，模型的泛化能力显著提升，进而提高了对测试集的预测准确率。在训练样本数量达到 100 张之前，模型泛化能力与样本数量呈线性正相关。

4.2 样本像素规模影响

在模型设计中，图像预处理（如缩放）是 CNN 训练前的关键步骤。图像尺寸调整可能导致

图像细节丢失,影响模型的学习和识别能力。为评估这些影响,使用城市专题图像构建了训练集、验证集和测试集。训练集包含可用作与不可用作景象匹配区的图像各 5 000 张,验证集包含两类图像各 1 500 张。模型训练设置 40 次迭代,全连接层包含 1 024 个神经元。测试集由 1 000 张不可用作景象匹配区的城市图像组成。图 2 展示了不同像素尺寸训练集对模型预测准确率的影响。结果表明,图像尺寸缩小程度过大会导致特征信息丢失,从而降低模型的泛化能力和预测稳定性。

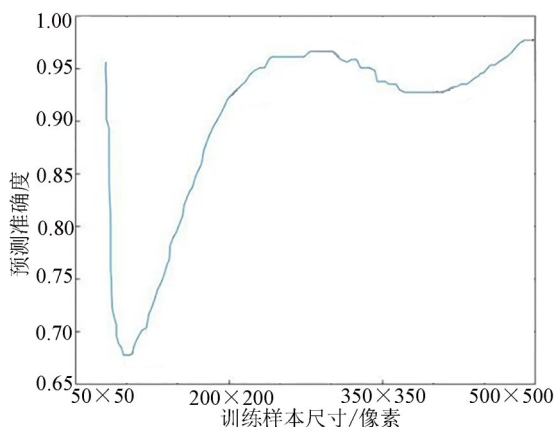


图 2 样本像素数与模型预测准确率关系曲线

Fig. 2 Curve of the relationship between sample pixel counts and the model prediction accuracy

4.3 迭代次数影响

虽然增加迭代次数通常能够提升模型性能,但也可能导致过拟合。因此,确定适当的迭代次数需要考虑任务特性和数据集特征。

使用城市专题图像分别构建了训练集、验证集和测试集。训练集由可用作与不可用作景象匹配区的图像各 5 000 张(分别定义为类 1 和类 2)构成,验证集包含上述两类图像各 1 500 张。图像分辨率统一设置为 500 像素×500 像素,全连接层含有 1 024 个神经元。测试集由 1 000 张城市专题图像组成,包括可用作与不可用作景象匹配区的图像。图 3 展示了不同迭代次数下模型预测准确率的变化曲线。

图 3 显示:迭代次数少于 40 次时,模型对类 1 和类 2 的预测准确率偏低,且两类图像的准确率差异较大;随着迭代次数超过 40 次,模型的预测准确率下降,且准确率差异逐渐增大,这可能指示过拟合现象;当迭代次数接近 40 次时,模型对两类图像的预测准确率最高,差异最小,表明此时模型具有最佳的预测性能和泛化能力。

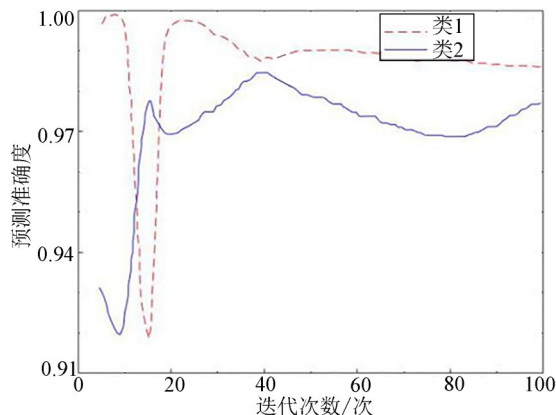


图 3 迭代次数与模型预测准确率关系曲线

Fig. 3 Curve of the relationship between iteration times and the model prediction accuracy

4.4 训练集类型影响

地域差异对模型训练具有显著影响。若训练集与测试集图像地域不一致,可能导致模型泛化能力下降。因此,构建 CNN 模型时,应优先选择与目标地域相匹配的图像作为训练样本。

构建了包含城市、荒漠、农田、热带雨林、山地、森林 6 类专题图像的训练集,每类各 10 000 张。验证集则包含每类 3 000 张图像。图像分辨率统一为 500 像素×500 像素,迭代 40 次,全连接层含 1 024 个神经元。测试集由每类 1 000 张不可用作景象匹配区的图像组成。图 4 展示了不同专题图像混合训练后 CNN 模型的预测准确率。

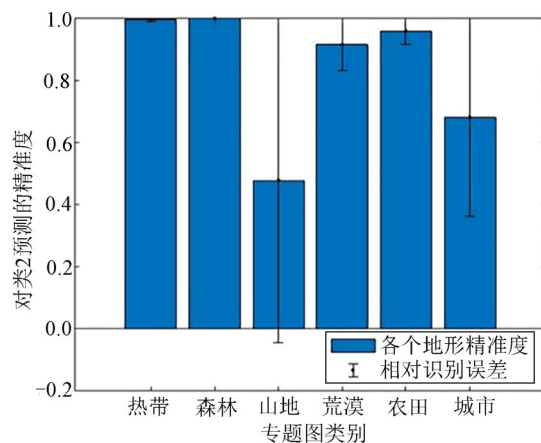


图 4 混合所有专题样本图片的模型预测准确率

Fig. 4 Model prediction accuracies of all thematic samples-mixed images

结果表明,混合专题图像训练的模型在不同区域的预测准确率存在显著差异,如森林图像准确率达到 1,而山地仅为 0.482,揭示了混合训练集对模型泛化能力的影响较大。

图5是利用两类专题图混合作为训练样本集、验证样本集,以各类1 000张图片作为单独测试集,分别得到卷积神经网络在6类专题图像中特征学习的差异。

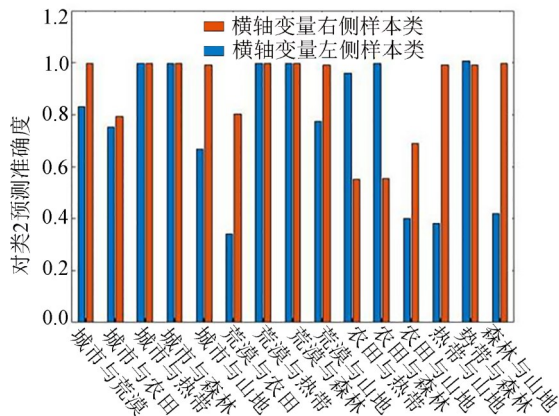


图5 混合任意两类专题图与模型预测准确率

Fig. 5 Model prediction accuracies of any two types thematic samples-mixed images

图5显示:模型在农田、荒漠和山地图像的特征提取上存在缺陷,专题图间特征点的差异影响了模型的提取效果,导致预测偏差,这降低了模型的普适性,尤其是在农田和山地类别上;特征相似的专题图(如城市与热带雨林、森林)之间的预测结果一致性较高;城市与农田图像由于共享某些特征(例如规整的绿化带),导致模型难以区分,从而降低了预测准确率。增加训练集的多样性可以提升模型对图片特征的识别能力。

4.5 全连接层优化

本研究采用Inception V3模型,添加2个全连接层,针对500像素×500像素的城市专题图像,构建了包含5 000张图片的训练集、1 500张图片的验证集和2 000张图片的测试集。经过40次迭代训练,结果如图6所示。

图6表明:全连接层神经元的数量对模型的预测准确率和稳定性有显著影响。当神经元数量增加到1 000时,模型的预测准确率显著提高并趋于稳定;然而,随着神经元数量的进一步增加,模型开始出现过度拟合,导致准确率下降。

5 模型评估

5.1 景象匹配区筛选准确率

图7展示了使用不同专题图像作为独立测试集的预测结果,这些图像未参与训练与验证过程。

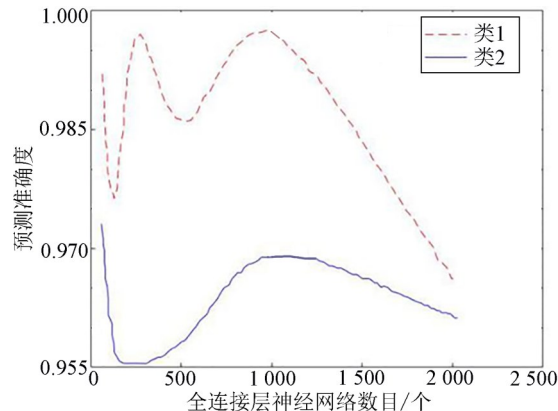


图6 全连接层神经元数目与模型预测准确率关系曲线

Fig. 6 Curves of the relationship between the number of fully connected layer neurons and model prediction accuracy

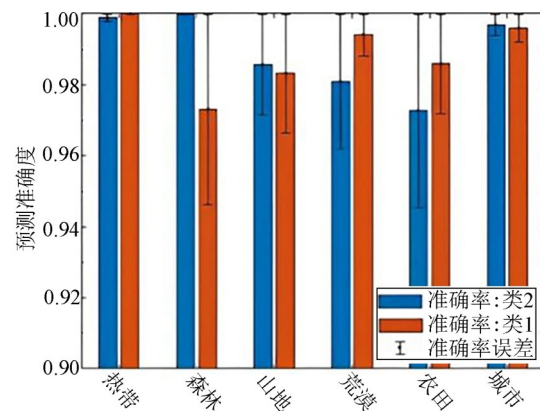


图7 各专题样本训练模型的准确率

Fig. 7 Accuracies of different thematic sample training models

对比图7与图4表明,与混合所有专题图像训练的CNN模型相比,单一专题图像训练的模型在专题预测上达到了更高的准确率。这进一步证实了CNN模型在图像预测任务中的可靠性和实用性,尤其是在使用单一专题图像进行训练时。

5.2 景象匹配区筛选耗时

对比传统基于特征匹配的景象匹配区筛选方法^[14]与CNN模型在不同数量灰度基准图输入下的耗时,结果如图8所示。图8表明:在输入图片数量小于10时,传统计算筛选具有时效优势;然而,随着样本数量的增加,传统计算的耗时呈现线性增长,显著高于CNN景象匹配区筛选模型的耗时。因此,在处理大规模数据时,CNN模型在适用性和效率上更有优势。这一优势主要源于CNN模型能够有效利用GPU的并行计算能力,实现高效的特征提取和匹配,从而显著提升筛选效率。

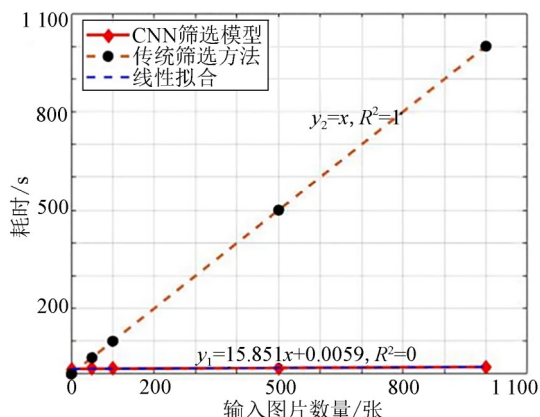


图 8 两类筛选技术耗时曲线

Fig. 8 Time consumption curves of two types of screening techniques

6 结论

构建了一种基于 CNN 的景象匹配区适配性筛选模型。通过对网络全连接层的优化及对不同景物类型的精细分类,本模型在保持较高适配性判断准确度的同时,有效发挥了深度网络在 GPU 并行计算方面的优势,显著提高了大规模景象匹配区适配性筛选的效率。试验结果显示,相较传统景象匹配区筛选算法,本模型的筛选效率更高。

参考文献:

- [1] 辛瑞,张霄力,彭侠夫,等.结合区域全局特征和 Ann-SIFT 的二阶段快速景象匹配算法[J].电子技术应用,2023,49(5):135-141.
XIN R, ZHANG X L, PENG X F, et al. Two-stage fast scene matching algorithm combining global descriptor and Ann-SIFT[J]. Application of Electronic Technique, 2023, 49(5): 135-141.
- [2] 蔡潇.基于相位特征的异源图像匹配算法[J].光学与光电技术,2021,19(1):48-53.
CAI X. Heterogeneous image matching algorithm based on phase features[J]. Optics & Optoelectronic Technology, 2021, 19(1): 48-53.
- [3] 曹菲,杨小冈,缪栋,等.景象匹配制导基准图选定准则研究[J].计算机应用研究,2005,22(5):137-139.
CAO F, YANG X G, MIAO D, et al. Study on reference image selection roles for scene matching guidance[J]. Application Research of Computers, 2005, 22(5): 137-139.
- [4] 沈孔怀.基于 SA-HRNet 的 SAR 景象适配区选取方法研究[D].武汉:华中科技大学,2020:21-35.

SHEN K H. Research on matching area selection of SAR image based on SA-HRNet[D]. Wuhan: Huazhong University of Science & Technology, 2020: 21-35.

- [5] 刘楠,岳亚洲,张晓冬,等.基于多尺度共生矩阵的景象匹配导航适配性分析[C]//新型惯性元件与先进导航技术研讨会文集.北京,中国惯性技术学会,2023:290-298.
- [6] 孙凯,林强,陈良洁.基于分块双向 2DPCA 及 ResNet 的景象区域适配性分析[J].网络安全技术与应用,2022(5):46-47.
- [7] 汪金花,张博,吴兵,等.基于贡献因子 BP 神经网络的地磁适配性研究[J].合肥工业大学学报(自然科学版),2020,43(12):1668-1675.
WANG J H, ZHANG B, WU B, et al. Research on adaptability evaluation of underground geomagnetic positioning based on BP neural network with contribution factor[J]. Journal of Hefei University of Technology (Natural Science), 2020, 43(12): 1668-1675.
- [8] 肖晶,齐晓慧,段修生,等.基于深度卷积神经网络的地磁导航方向适配性分析[J].工程科学学报,2017,39(10):1584-1590.
XIAO J, QI X H, DUAN X S, et al. Direction-matching-suitability analysis for geomagnetic navigation based on convolutional neural networks[J]. Chinese Journal of Engineering, 2017, 39(10): 1584-1590.
- [9] 杨娇.基于深度学习的适配区选取方法研究[D].武汉:华中科技大学,2019:48-53.
YANG J. Suitable matching area selection method based on deep learning[D]. Wuhan: Huazhong University of Science & Technology, 2019: 48-53.
- [10] 马龙,缪栋,沈绪榜.基于图像自匹配性的景象匹配区的选定准则[J].上海航天,2001,18(1):50-54.
MA L, MIAO D, SHEN X B. A selection criterion of scene matching field based on image self-matching[J]. Aerospace Shanghai, 2001, 18(1): 50-54.
- [11] 张磊,王海涛.基于灰度方差的图像分割算法在医学图像处理中的应用[J].中国医学影像技术,2011,27(2):196-198.
ZHANG L, WANG H T. Application of image segmentation algorithm based on gray variance in medical image processing[J]. Chinese Journal of Medical Imaging Technology, 2011, 27(2): 196-198.
- [12] 李瑞峰,张丽丽.基于最高峰值比的图像分割算法

- 在医学图像处理中的应用[J]. 中国医学影像技术, 2011, 27(6): 909-912.
- LI R F, ZHANG L L. Application of image segmentation algorithm based on maximum peak-to-trough ratio in medical image processing[J]. Chinese Journal of Medical Imaging Technology, 2011, 27(6): 909-912.
- [13] 刘中华, 王晖, 陈宝国. 景象匹配区选取方法研究[J]. 计算机技术与发展, 2013, 23(12): 128-133.
- LIU Z H, WANG H, CHEN B G. Research on selection for scene matching area[J]. Computer Technology and Development, 2013, 23(12): 128-133.
- [14] 李竹林. 景象匹配算法及其适应性分析[J]. 计算机工程与应用, 2014, 50(13): 182-186.
- LI Z L. Scene matching algorithm and its adaptability analysis[J]. Computer Engineering and Applications, 2014, 50(13): 182-186.