

研究论文 ORIGINAL PAPERS

基于雪地足迹影像的物种分类识别方法 ——以东北地区六种哺乳动物为例

艾昊橙^{1#}, 戴文锐^{2#}, 马光凯², 姜广顺^{1*}

(1. 东北林业大学野生动物与自然保护地学院, 国家林业和草原局猫科动物研究中心, 哈尔滨 150040;
2. 东北林业大学计算机与控制工程学院, 哈尔滨 150040)

稿件运行过程

收稿日期: 2024-05-09

修回日期: 2024-05-11



关键词: 物种识别;
雪地足迹识别;
深度学习;
实例分割;
分类网络

Keywords: Species identification;
Snow footprint recognition;
Deep learning;
Instance segmentation;
Classification network

中图分类号: Q958.1; TP391.4

文献标志码: A

文章编号:

2310-1490(2025)-03-0501-13

DOI: 10.12375/ysdwxb.20250304

摘要

东北地区在我国生物资源上占据重要地位,受寒冷漫长的冬季以及丰厚降雪量的影响,野外工作者在冬季通常优先选择雪地足迹来完成野生动物的识别和监测,但通过足迹识别物种对人员资质要求较高,且对相似雪地足迹肉眼识别仍存在难度。因此,简单、高效且准确地通过雪地足迹影像来识别野生动物成为东北地区冬季野生动物监测的关键。为此,提出了基于雪地足迹和深度学习识别野生动物种类的方法,以东北地区6种哺乳动物为研究对象,通过野外人工采集并处理足迹影像制作雪地足迹数据集,选用分割模型SOLOv2作为分割网络完成对雪地足迹轮廓特征的自动提取,探究各物种的实例分割精度和效果。选用分类模型Swin-Transformer-Tiny作为分类网络实现对雪地足迹影像轮廓特征的自动分类识别,探究各物种的分类识别准确率。结果表明:本研究提出的雪地足迹识别分类方法具有较好的可信度;在分类网络中,通过人工标注雪地足迹轮廓完成物种自动分类可以达到89.9%的分类识别准确率;在雪地足迹识别网络中,通过其分割网络的结果对分类网络进行训练,最终完成整体的自动化识别,达到85.3%的分类识别准确率。总体上,通过雪地足迹影像以及深度学习技术,利用分割-分类网络构建野生动物雪地足迹识别网络,实现了基于动物雪地足迹影像的自动化物种分类识别,为我国北方冬季野生动物监测和保护提供了一种简单高效的方法。

基金项目: 国家重点研发计划项目(2023YFF1305000)

第一作者简介: 艾昊橙(1999—),男,硕士研究生;主要从事动物生态和深度学习研究。E-mail:1175999711@qq.com

*共同第一作者: 艾昊橙; 戴文锐

*通信作者: 姜广顺, E-mail:jgshun@126.com

A Species Classification and Identification Method Based on Snow Footprint Images: a Case Study of Six Mammalian Species in Northeast China

AI Haocheng^{1#}, DAI Wenrui^{2#}, MA Guangkai², JIANG Guangshun^{1*}

(1. College of Wildlife and Protected Area, Northeast Forestry University,

National Forestry and Grassland Administration Feline Research Center, Harbin 150040, China;

2. College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China)

Abstract: Northeast China holds a significant position in the country's biological resources. Influenced by its long, cold winters and abundant snowfall, field researchers often prioritize using snow footprints for wildlife identification and monitoring during the winter months. However, identifying species based on footprints requires highly qualified personnel and remains challenging when distinguishing visually similar footprints. Therefore, developing a simple, efficient, and accurate method to identify wildlife species through snow footprint images has become crucial for wildlife monitoring in this region. This study proposed a method for identifying wild animal species based on snow footprints and deep learning. Six mammalian species in northeast China were studied, and a snow footprint dataset was created by manually collecting and processing footprint images in the field. The segmentation model SOLOv2 was selected as the segmentation network to automatically extract the contour features of snow footprints, exploring the instance segmentation accuracy and effectiveness of each species. Selected the Swin-Transformer-Tiny classification model as the classification network to achieve automatic classification and recognition of snow footprint image contour features, and explored the classification and recognition accuracy of various species. The results indicate that the snow footprint recognition and classification method proposed in this study has good reliability; In classification networks, automatic species classification can achieve an accuracy of 89.9% by manually annotating snow footprint contours; In the snow footprint recognition network, the classification network is trained based on the results of its segmentation network, and the overall automated recognition is completed, achieving a classification recognition accuracy of 85.3%. Overall, by using snow footprint images and deep learning techniques, a snow footprint recognition network is constructed using segmentation classification networks, achieving automated species classification and recognition based on animal snow footprint images, and it has potential in terms of application prospects in winter, providing a simple and efficient innovative method for monitoring and protecting wildlife in northern China.

野生动物种类的监测是保护野生动物及其栖息地的重要环节^[1],其中物种识别在监测生物多样性和保护工作中发挥着关键作用。通过识别物种并了解其分布、数量和行为,不仅能够全面评估生态系统的健康状况,还有助于及早发现濒危物种的生存威胁,并为制定保护策略、划定管理保护区以及研究野生动物的迁徙行为提供科学依据。传统监测动物种类主要以人工为主,方法通常包括足迹、粪便、栖卧痕迹、毛发、叫声、食物残骸和架设的自动红外监测相机拍到的照片等数据来识别野生动物种类^[2]。

东北地区的自然保护区在我国生物资源上占据重要地位,其为许多珍稀濒危物种提供了栖息条件,

同时还是我国重要的生物多样性保护区域之一。受寒冷漫长的冬季以及丰厚降雪量的影响,该地区的野生动物在冬季活动时留下清晰且持久的雪地足迹链。借此特性,野外工作者在冬季通常优先选择雪地足迹来完成对野生动物的识别和监测。在众多物种识别方法中,足迹识别因其独特的非损伤性的识别方式,不要求被监测的动物被看见、捕获或处理^[3]。然而在实际野外监测中,通过足迹识别物种需要相关人员利用大量时间和实践来了解野生动物的足迹识别特征,且野生动物如狍(*Capreolus pygargus*)、梅花鹿(*Cervus nippon*)和野猪(*Sus scrofa*)等,其雪地足迹特征具有相似性,即使是清晰的雪地足迹,

仅依靠肉眼识别仍存在难度。因此,简单、高效且准确地通过雪地足迹影像来识别野生动物成为东北地区野生动物监测的关键。

足迹识别技术(footprint identification technique, FIT)可以实现基于足迹上标记关键点之间的距离、角度和面积测量来识别动物的个体及性别^[4-6],该技术已逐渐成为野生动物保护领域中一种有效且低成本的工具。这种“非损伤性”技术最初是为黑犀牛(*Diceros bicornis*)的识别而开发^[4]。如今,它也已成功地应用于猎豹(*Acinonyx jubatus*)^[7]、白犀牛(*Ceratotherium simum*)^[8]、东北虎(*Panthera tigris altaica*)^[9]、美洲狮(*Puma concolor*)^[3]和大熊猫(*Ailuropoda melanoleuca*)^[10]等濒危物种的识别。在东北虎足迹识别的相关研究中,已实现基于雪地足迹的关键点信息精准识别个体及性别,其性别判别准确率达到97.50%,个体识别准确率达到93.75%^[9]。相比于基于足迹的个体识别,通过雪地足迹实现物种识别的研究相对较少,且东北地区冬季降雪量大,野生动物活动时留下的雪地足迹常常只能得到足迹的轮廓特征而无法得到足迹的关键点特征。因此,通过足迹影像的轮廓以及短足迹链的特征对冬季东北地区不同种类动物进行物种分类识别具有重要的研究价值。

近年来,随着计算机算力的发展以及深度学习理论的突破。基于深度学习的野生动物识别技术相比传统方法展现出了显著优势,其能够高效、准确地实现图像特征提取,且无须人工干预。基于深度学习的图像检测与识别技术已经广泛应用到动物物种识别中^[11],例如,王文成等^[12]对10种鱼类进行分类识别,其使用的网络框架为ResNet50,最终识别成功率达到93.33%;Norouzzadeh *et al.*^[13]利用Snapshot Serengeti dataset(目前最大的已标注野生动物数据集)和深度学习分类算法,对比了9种深度学习模型在48种野生动物识别任务中的表现,其中ResNet152架构以93.8%的准确率表现最优;Guo *et al.*^[14]基于Tri-AI技术对41种灵长类动物进行识别,识别准确率高达94.1%;齐建东等^[15]基于改进的BS-ResNeXt-50模型对密云地区野生动物图像进行物种识别,最高类别准确率达98.6%。

本研究将动物雪地足迹影像与深度学习方法结合,通过实例分割模型SOLOv2^[16]以及分类模型Swin-Transformer-Tiny^[17],构建分割-分类的雪地足迹物种识别技术,通过整合的分割模型与分类模型,

形成一体化的基于雪地足迹的物种识别模型。这为东北地区冬季野生动物的识别、监测和保护提供了一种简单高效的方法。

1 研究区概况

研究区覆盖东北3处国家级自然保护区和1个猫科动物饲养繁育中心。其中黑龙江小北湖国家级自然保护区地处张广才岭中段(44°03′—44°18′ N, 128°33′—128°45′ E),处于温带大陆性季风气候区,四季气候变化明显,自然地形复杂,生境丰富多样的特点为野生动物的活动和繁衍提供了理想条件^[18];黑龙江太平沟国家级自然保护区境内属小兴安岭北坡低山丘陵地带(48°02′—48°20′ N,130°31′—130°50′ E),处于温带大陆性季风气候区,地势东低西高,保护区内分布有大面积的森林、灌丛、草甸、沼泽及河流水域,为众多野生动物提供了良好的栖息环境和充足的食物^[19];吉林汪清国家级自然保护区位于吉林省东部延边朝鲜族自治州汪清县和琿春市境内(43°05′—43°30′ N, 130°23′—131°03′ E),属于中温带湿润温凉季风气候区,有国家一级重点保护野生动物5种,包括东北虎、东北豹(*Panthera pardus orientalis*)等^[20];中国横道河子猫科动物饲养繁育中心位于黑龙江省东部海林市(44°47′ N,129°06′ E),是目前全球最大的东北虎人工饲养繁育基地,处于温带大陆性气候带,四周崇山峻岭,全年日照充裕,降水量丰富,成为东北虎栖息地的首要选择,为开展东北虎保护研究项目提供了基础^[9]。

2 数据采集与处理

2.1 数据采集

2021—2023年冬季,在3个自然保护区区域内通过人工采集获得野生动物雪地足迹影像。由于野生东北虎数量稀少,野外跟踪时很难发现其不同个体的雪地足迹,因此在横道河子猫科动物饲养繁育中心人工采集不同个体的东北虎雪地足迹进行数据补充。采集对象包括东北虎、东北兔(*Lepus mandshuricus*)、黄喉貂(*Martes flavigula*)、狍、梅花鹿和野猪。使用iPhone12后置摄像头拍摄,在野外拍摄时不破坏原有足迹,考虑到物种个体以及四肢足迹的差异性,采集人员对同一物种的同一足迹链跟踪拍摄足迹6~10张,平均每种动物足迹拍摄照片500~600张。拍摄时手机镜头平行于足迹,距离地面50~60 cm,使用二倍焦距拍摄。在图片格式上备注种

类、地点和时间。小型动物的脚印小且密集,难以拍摄单一脚印,拍摄时可拍摄单个足迹或体现足迹特征的短足迹链。若遇到动物前后足重合的足迹,则选择完全重合的清晰足印拍照。由于动物行走方

式、雪地足迹存在时长和雪的松软都会影响足迹特征的清晰完整程度,因此采用人工方式,根据研究对象足迹特征的明显程度筛选出具备明显足迹特征的动物雪地足迹数据(图1)。

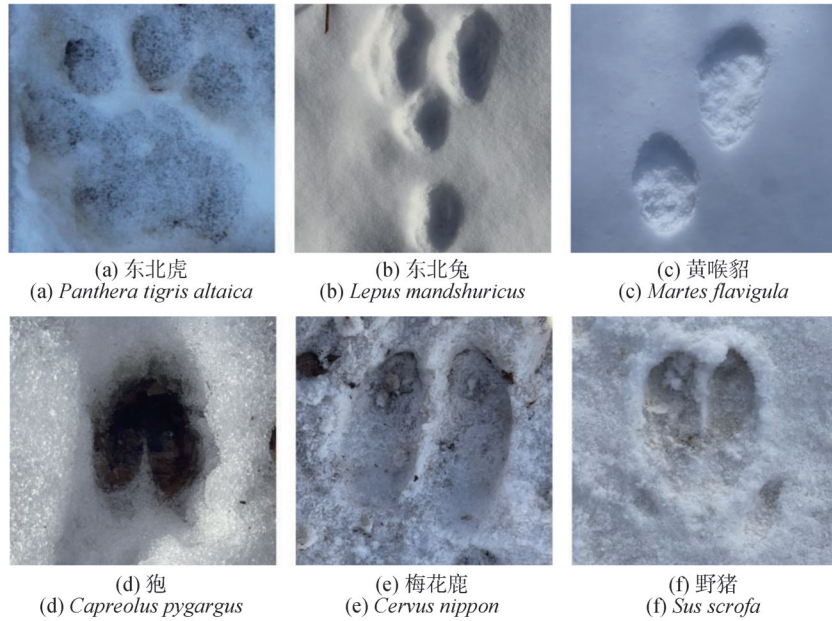


图1 具备明显足迹特征的动物雪地足迹

Figure 1 Animal snow footprint with distinct footprint characteristics

2.2 数据筛选

比对每张影像,确定足迹对应的动物种类,共筛选出具备明显足迹特征的影像3 538张。按照7:2:1构建雪地足迹数据集,其中2 476张作为训练集,707张作为验证集,355张作为测试集。使用图像标注工具

LabelMe^[21]标注雪地足迹影像的轮廓,并标记所属类别。标记时统一标准,对单个足迹进行连续标记,在断片处进行适当过渡连接;对于存在多个足迹的短足迹链影像,则对每个足迹进行单独标注。标记影像如图2所示。

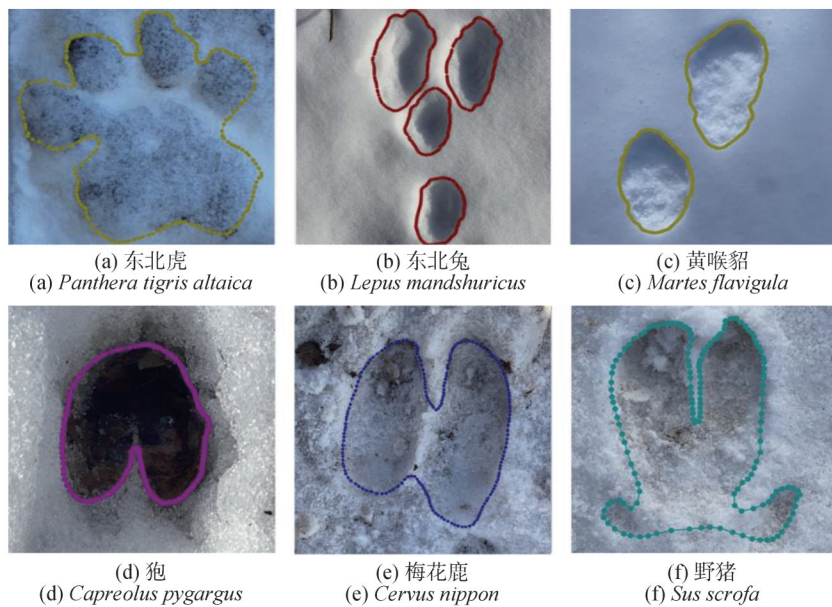


图2 人工标注后的动物雪地足迹

Figure 2 Manually annotated animal tracks in snow post-processing

2.3 数据预处理

针对人工标注的雪地足迹影像,一方面,利用数据集构建足迹轮廓分割模型,从原始影像中提取足迹区域;另一方面,将标注影像进行二值化处理以突出足迹特征,并通过归一化优化特征提取效果,生成 Handle 数据集,用于足迹识别模型训练。同时,利用分割模型对原始影像进行自动分割,并对分割结果

进行二值化和归一化处理,生成 Auto 数据集,用于分类模型训练。将 Handle 和 Auto 数据集结合构建 Handle + Auto 数据集,对分类网络进行训练,并通过不同训练集方案开展足迹识别网络的性能分析。最终,结合分割模型与分类模型,构建雪地足迹的物种识别网络,实现从影像处理到足迹识别的完整流程(图3)。

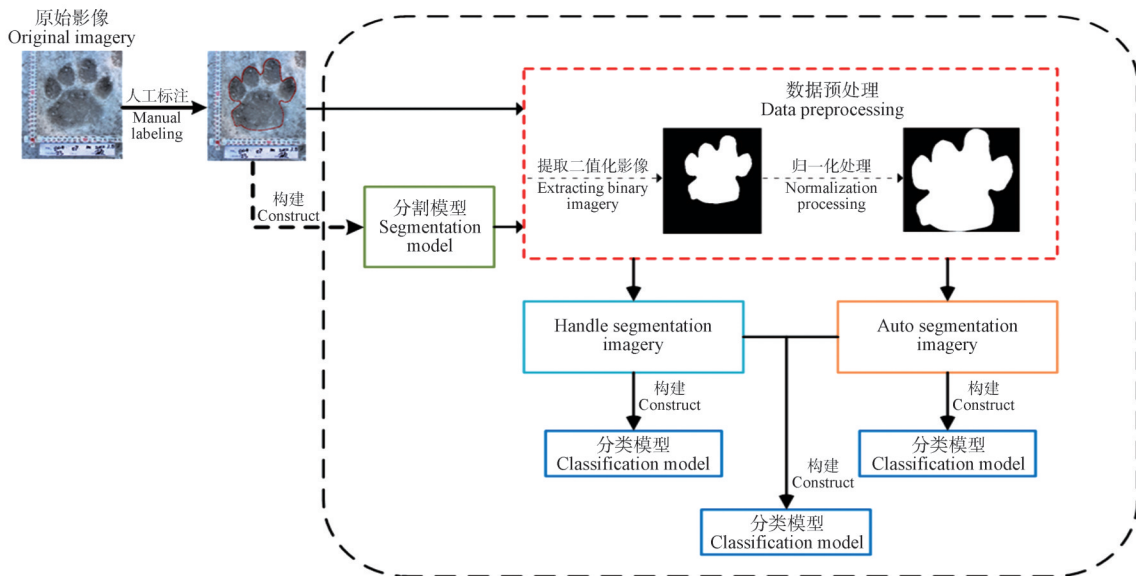


图3 雪地足迹影像预处理及模型构建流程

Figure 3 Snow footprint image preprocessing and model construction

3 分析方法

3.1 基于雪地足迹影像的物种识别网络

基于动物雪地足迹的物种识别网络由雪地足迹分割网络和足迹识别网络两部分组成。雪地足迹的原始影像通过足迹分割网络提取足迹轮廓,再经过二值化和归一化处理后传入分类网络。分类网络对预处理后的足迹影像进行特征提取,再经过分类器对其进行分类,从而实现基于雪地足迹的野生动物物种识别。雪地足迹分割-分类一体化的识别模型如图4所示,其中足迹分割网络由特征提取网络、特征融合网络以及检测头组成。特征提取网络通过 ResNet50 提取多尺度的足迹特征,特征融合网络利用 FPN (feature pyramid network)^[22] 将不同尺度的特征图进行整合,以增强分割精度。检测头对整合后的特征图生成足迹的分割掩码,用于足迹轮廓的提取。足迹识别网络包括特征提取模块和分类模块。特征提取模块从分割后处理的二值化足迹影像中提取高维特征,分类模块则根据提取的特征完成对足迹

所属物种的分类,从而实现物种识别的最终目标。

3.2 雪地足迹影像分割网络

采用深度学习分割算法 SOLOv2 构建雪地足迹分割模型。相比依赖检测框的 Mask R-CNN^[23], SOLOv2 通过将图像划分为固定网格,直接为每个网格分配类别和分割掩码,进而提升分割速度和精度^[16]。此外,SOLOv2 引入动态卷积机制 (dynamic convolution)^[24],通过学习动态卷积核权重进一步增强模型的表达能力与分割效果,其网络架构如图5所示。

SOLOv2 在 SOLO 基础上将掩码分支解耦为卷积核学习分支和特征学习分支。卷积核学习分支从特征图中生成每个实例的动态卷积核权重,而特征学习分支提取适用于生成掩码的实例特征图,二者联动以生成实例掩码。如图5所示,在分割流程中,雪地足迹的原始影像首先通过 ResNet 特征提取网络和 FPN 生成多层特征图 I。分类分支通过卷积操作对 I 中各层特征图进行处理,生成网格单元的实例类别概率分布。在掩码分支中,特征图 I 经卷积核学习分支生成动态卷积核权重 (大小为 $S \times S \times D$, 其中 S

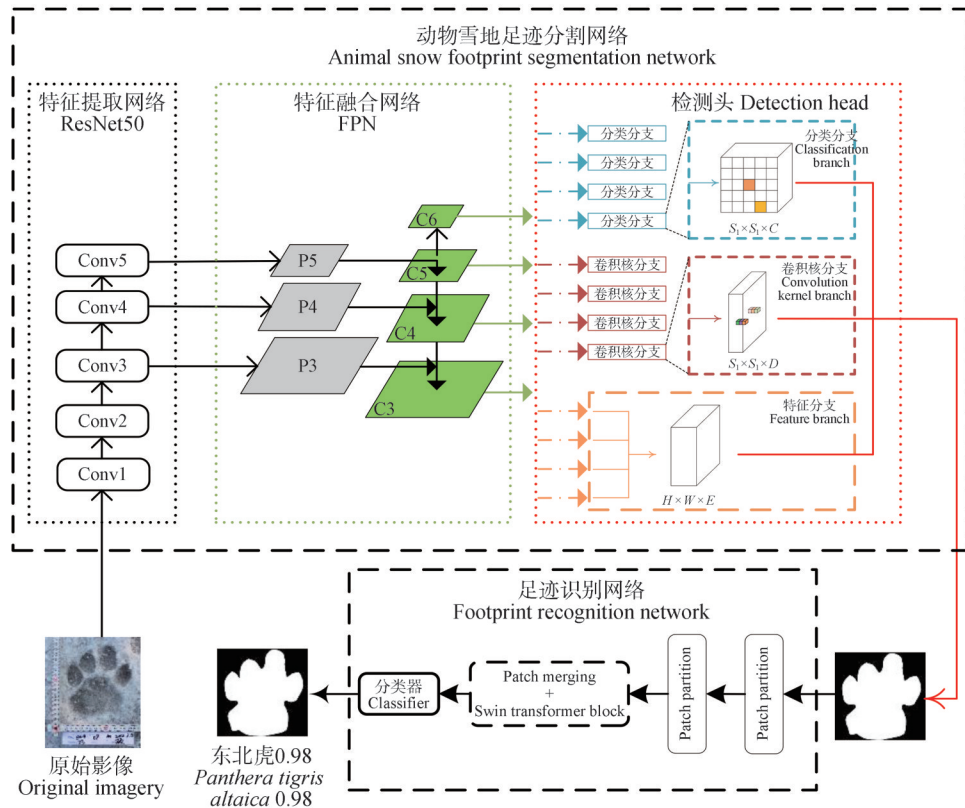


图4 基于雪地足迹物种识别网络
Figure 4 Species recognition network based on snow footprint

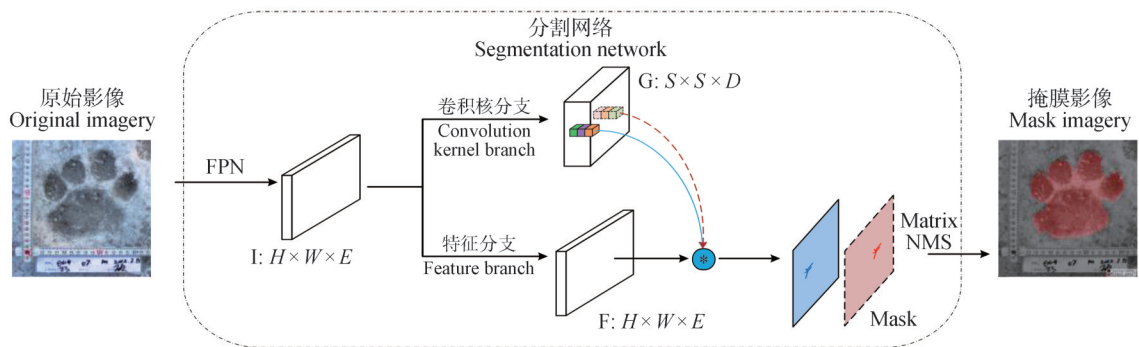


图5 SOLOv2结构
Figure 5 SOLOv2 structure

为网格数, D 为动态卷积核参数总数)。同时, 特征图 I 通过特征学习分支逐级上采样与加权融合, 生成高分辨率单一特征图 (分辨率为原始影像的 $1/4$)。最终, 特征学习分支的输出与动态卷积核结合, 生成对应实例的分割掩码^[25]。

3.3 雪地足迹影像分类网络

采用 Swin-Transformer-Tiny 构建雪地足迹影像的识别模型, Swin-Transformer-Tiny 网络结构如图 6 所示。首先, 将输入的 $H \times W \times 3$ 的图像分割为 N ($N = \frac{H}{P} \times \frac{W}{P}$) 个非重叠的小块 (patch), 每个 patch 的

大小为 $P \times P \times 3$ 。其次, 每个 patch 通过线性投影转化为高维特征, 记为 patch tokens, 再经线性嵌入层 (liner embedding) 转换为 C 维向量。最后将该向量送入 Swin-Transformer block 进行局部特征提取和跨窗口信息交互 (stage 1)。随后, 通过 patch merging 模块逐步合并 tokens, 减少特征分辨率并提升维度。整个网络由 4 个阶段 (stage 1 至 stage 4) 组成, 每一阶段在分辨率逐渐减小的同时提取更深层的语义信息, 最终形成多尺度表征^[17]。

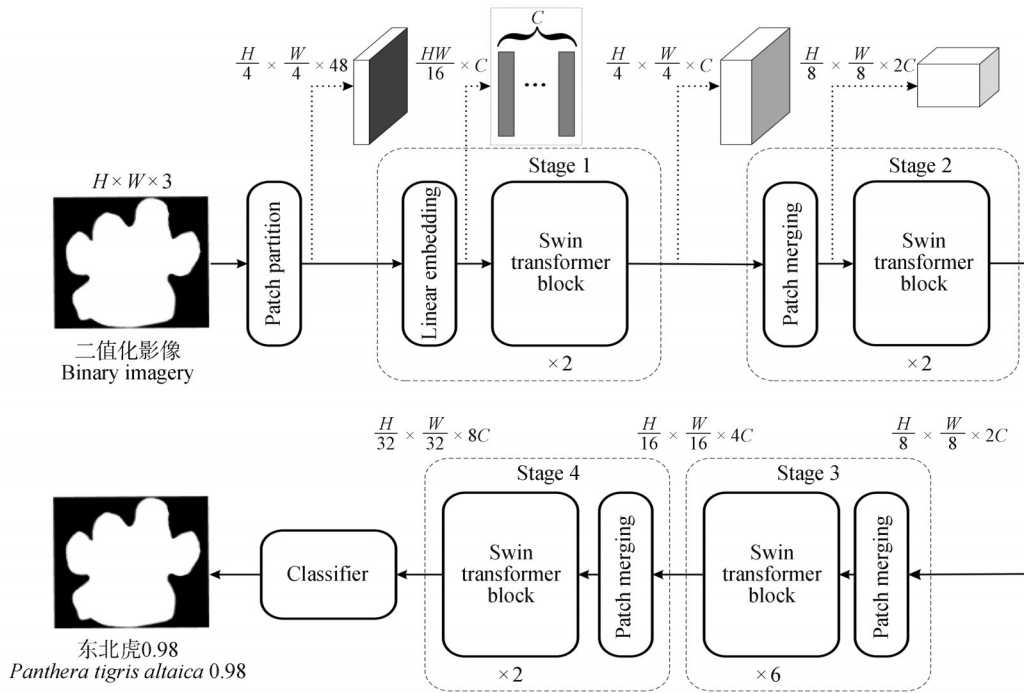


图6 Swin-Transformer 结构

Figure 6 Swin-Transformer structure

3.4 实验参数设置

实验平台硬件环境：内存 256 GB，显卡 4 张 NVIDIA3090，系统 Ubuntu 18.04，编程语言为 Python 3.8，深度学习框架为 PyTorch 1.10。分析使用人工标注后的动物雪地足迹数据集，以 7:2:1 比例随机划分为训练集、验证集和测试集。在后续对于足迹识别网络中分类网络的训练所使用的 Handle、Auto 以及 Handle + Auto 数据集均以 7:2:1 比例随机划分为训练集、验证集和测试集，且实验参数与分类实验参数一致。

分割实验的参数：批大小(batch size)为 1，使用 SGD 作为优化器，学习率(learning rate)为 0.001，权重衰减(weight decay)为 0.0001，动量因子(momentum)为 0.9，仅使用随机水平翻转的数据增强方式。

为验证本研究所选用分割网络模型 SOLOv2 的性能，通过同一数据集对 SOLOv2、Mask R-CNN 以及 Queryinst^[26] 三种模型的分割精度以及分割效果进行比较分析。

分类实验的参数设置：批大小为 16，优化器选用 AdamW，初始学习率为 0.0001，训练轮次(epoch)为 100，数据增强采用随机水平翻转和随机垂直翻转以及颜色抖动(color jitter)，系数设定分别为 0.5、0.5

和 0.4，并采用余弦退火调整学习率。为验证本研究所使用的分类网络模型 Swin-Transformer-Tiny 性能，设置实验与其他分类模型进行对比，包括 ConvNeXt-Base^[27]、ResNet50、MobileViTV2^[28] 和 MobileNetV3^[29]。通过同一数据集对其分类准确度进行比较性分析。

3.5 评价指标

3.5.1 实例分割评价指标

采用 COCO 数据集中实例分割的评价指标平均精度(average precision, AP)和平均精度均值(mean average precision, mAP)^[30] 来评估雪地足迹分割的准确性，并使用帧率(frame rate)评估分割速率。AP 是在一定交并比(intersection over union, IoU)阈值下衡量模型的分割性能，反映精确率(precision, P)和召回率(recall, R)的平衡，其中 IoU 表示掩膜区域与模型预测结果的交并比。精确率表示预测对的正例数量占所有预测为正例数量的比例。召回率表示预测对的正例数量占整体正例的比例，即正例的检出比例。精确率、召回率定义为

$$P = \frac{T_p}{T_p + F_p}, \quad (1)$$

$$R = \frac{T_p}{T_p + F_N}. \quad (2)$$

式中: T_p 为被正确识别的目标个数, F_p 为被错误识别的目标个数, F_N 为没有被正确识别出来的目标个数。COCO数据集使用多个IoU阈值来评估模型性能,范围从0.50到0.95,步长为0.05,共10个IoU阈值。AP与mAP定义为

$$P_{APc(IoU_t)} = \int_0^1 P(R) dR, \quad (3)$$

$$P_{mAP} = \frac{1}{N_c} \sum_{c=1}^{N_c} \frac{1}{10} \sum_{t=1}^{10} P_{APc(IoU_t)} \quad (4)$$

式中: P 和 R 分别为精确率和召回率, N 为类别数, IoU_t 为第 t 个IoU阈值, $P_{APc(IoU_t)}$ 是类别 c 在特定 IoU_t 下的AP值。 P_{mAP} 是所有类别IoU在10个阈值中的平均值,其反映的是模型的总体分割性能。

3.5.2 分类评价指标

采用足迹识别的准确率(accuracy, Acc),即模型正确分类结果占总数据量的比例,来评估雪地足迹识别的准确性。准确率(A_{cc})的定义为

$$A_{cc} = \frac{T_p}{T_p + F_p + F_N} \quad (5)$$

式中: T_p 、 F_p 、 F_N 的含义同式(1)和式(2)。

4 结果

4.1 分割网络的有效性

4.1.1 不同模型对雪地足迹的分割性能

为验证所选用实例分割模型SOLOv2的性能,将SOLOv2、Mask R-CNN以及Queryinst三种模型进行

对比。由表1可见,SOLOv2分割速率略低于Mask R-CNN,但其实例分割精度最佳,其平均精度均值达到了64.6%,高于Mask R-CNN(58.2%)和Queryinst(58.3%)。

表1 不同模型对雪地足迹影像的分割性能

Table 1 Segmentation performance of different models on snow footprint imagery

模型 Model	分割平均精度均值 / % Segmentation mean average precision	帧率 / (帧·s ⁻¹) Frame rate
Mask R-CNN	58.2	15.8
Queryinst	58.3	12.8
SOLOv2	64.6	12.8

由于足迹识别网络的输入是二值化的足迹影像,分割的准确性极大地影响足迹识别的效果。图7为6种动物雪地足迹影像经实例分割后的二值化影像。从分割效果看,模型Mask R-CNN分割精准,但边缘处理不细致;模型Queryinst分割不够完整,在足迹的过渡连接处出现断片现象;模型SOLOv2能有效分割影像中的雪地足迹轮廓,且对边缘分割处理更细致、更完整。因此,精度更高的SOLOv2能够为足迹识别任务提供更可靠的分割影像,提升整体识别效果,尤其在复杂背景或足迹形状模糊的条件下,其优势更加突出。

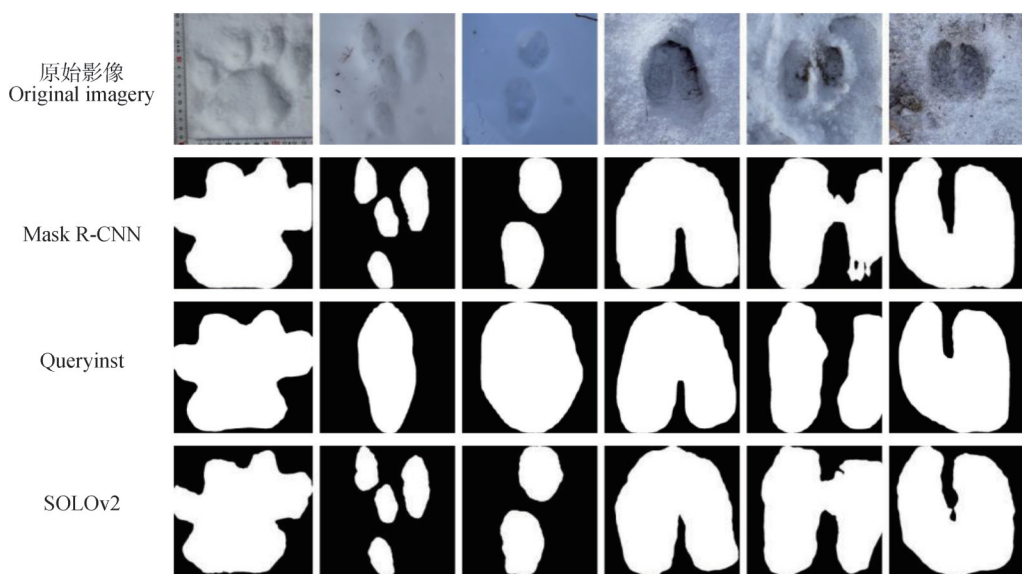


图7 不同实例分割模型分割效果

Figure 7 Segmentation performance of different instance segmentation models

4.1.2 分割网络对各物种的分割结果

表2为实例分割模型SOLOv2分别对6种哺乳动物的分割精度结果,有蹄类动物狍、梅花鹿以及野猪的分割精度偏低,其中野猪的分割精度最低,仅为56.0%。这说明野猪的原始影像以及标注后的雪地足迹数据集之间差异过大,也表明野猪的足迹形状存在一定的不确定性,进而使分割网络的分割结果之间出现精度差别。

表2 分割模型SOLOv2对各物种的分割结果

Table 2 Segmentation results of SOLOv2 model for different species

物种 Species	分割平均精度 /% Segmentation average precision
东北虎 <i>Panthera tigris altaica</i>	74.4
东北兔 <i>Lepus mandshuricus</i>	66.8
黄喉貂 <i>Martes flavigula</i>	68.6
狍 <i>Capreolus pygargus</i>	61.2
梅花鹿 <i>Cervus nippon</i>	60.7
野猪 <i>Sus scrofa</i>	56.0

4.2 分类网络的有效性

4.2.1 不同模型对雪地足迹影像的分类性能

为验证分类模型Swin-Transformer-Tiny的性能,依据准确率性能指标,对比ConvNeXt-Base、ResNet50、MobileViTV2和MobileNetV3的分类性能。由表3可见,Swin-Transformer-Tiny的表现最为出色,其分类识别准确率最高(89.9%)。这一结果表明,Swin-Transformer-Tiny在处理雪地足迹数据时,能够更好地捕捉到物种特征,从而保证识别的准确性和可靠性。

表3 不同模型对雪地足迹影像的分类性能

Table 3 Recognition accuracy of snow footprint imagery across different models

网络模型 Model	准确率 /% Accuracy
ConvNeXt-Base	88.7
ResNet50	88.1
MobileViTV2	88.7
MobileNetV3	88.1
Swin-Transformer-Tiny	89.9

4.2.2 分类网络对各物种的分类识别结果

模型Swin-Transformer-Tiny对6种哺乳动物的分类识别准确率结果见表4,从表4可以看到,该分类

网络在人工标注的数据集Handle训练下对东北虎和东北兔实现了准确的分类识别。在分类识别结果中,黄喉貂的准确率为94%、狍的准确率为90%、梅花鹿的准确率为83%、野猪的准确率为63%。由于狍、梅花鹿和野猪的分类识别准确率偏低,因此对分类网络结果的混淆矩阵进行分析。

表4 分类网络Swin-Transformer-Tiny对各物种的分类识别结果

Table 4 Classification results of Swin-Transformer-Tiny model for various species

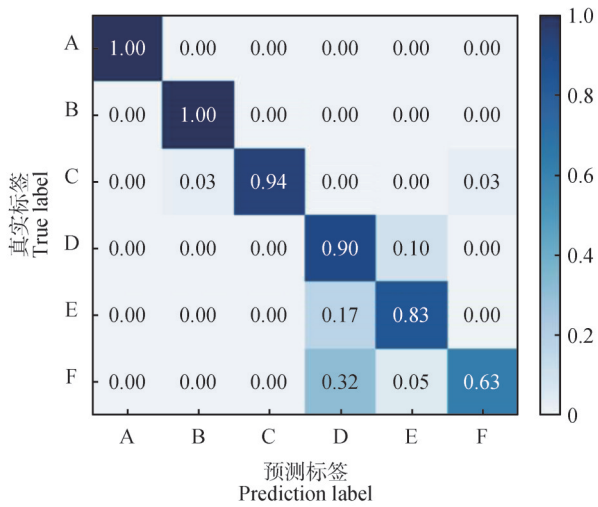
物种 Species	准确率 /% Accuracy
东北虎 <i>Panthera tigris altaica</i>	100
东北兔 <i>Lepus mandshuricus</i>	100
黄喉貂 <i>Martes flavigula</i>	94
狍 <i>Capreolus pygargus</i>	90
梅花鹿 <i>Cervus nippon</i>	83
野猪 <i>Sus scrofa</i>	63

4.2.3 分类网络结果混淆矩阵分析

混淆矩阵可用于观察Swin-Transformer-Tiny模型在各个物种类别上的表现。从结果的直观表现来看,Swin-Transformer-Tiny在东北虎和东北兔的分类识别结果中准确率极高,而有10%的梅花鹿足迹被分类识别为狍,有17%的狍足迹被分类识别为梅花鹿,野猪足迹中32%被分类识别为狍,5%被分类识别为梅花鹿(图8)。狍、梅花鹿和野猪皆为有蹄类动物,其雪地足迹特征具有相似性,表明不同物种可能存在相似的足迹特征,进而导致网络识别错误,而对于足迹特征明显的物种则产生较高的分类识别准确率。

4.2.4 分类网络结果的可视化

分类网络的热力图可用于直观展示模型对不同类别的置信度分布。热力图中的暖色区域表示模型对特定类别具有较高的预测可信度,即模型认为该区域属于该类别的概率较高;而冷色区域表示模型对特定类别的预测可信度较低,即模型认为该区域属于该类别的概率较低。图9为Swin-Transformer-Tiny在部分验证集分类结果的热力图,可见热力区域多集中在足迹的轮廓部分,表明模型在分类过程中主要依据足迹轮廓特征进行识别。



A. 东北虎; B. 东北兔; C. 黄喉貂; D. 狍; E. 梅花鹿; F. 野猪。
A. *Panthera tigris altaica*; B. *Lepus mandshuricus*; C. *Martes flavigula*; D. *Capreolus pygargus*; E. *Cervus nippon*; F. *Sus scrofa*.

图8 雪地足迹识别归一化混淆矩阵

Figure 8 Snow footprint recognition normalized confusion matrix

4.3 雪地足迹识别网络模型实验性能

为了探究采集到雪地足迹后,采用分割-分类一体化模型自动实现物种识别的可行性及其识别准确性,使用不同的数据集(Handle、Auto以及Handle + Auto)来训练和验证 Swin-Transformer-Tiny 分类网络,最终得到6项如表5所示的分类结果。结果表

表5 不同方案对雪地足迹识别网络的分类识别准确率

Table 5 Classification accuracy of the snow-terrain footprint identification network under identical schemes

分割方式/模型 Segmentation method /model	训练集 Training set	测试集 Test set	准确率/% Accuracy
Handle	Handle	Handle	89.9
Handle	Handle	Auto	81.6
SOLOv2	Auto	Handle	78.3
SOLOv2	Auto	Auto	82.1
SOLOv2	Handle + Auto	Handle	80.2
SOLOv2	Handle + Auto	Auto	85.3

5 讨论

5.1 数据采集和处理的影响因素

原始雪地足迹影像采集于冬季野外,受雪的深度、松软程度、留存时长以及动物活动等条件的影

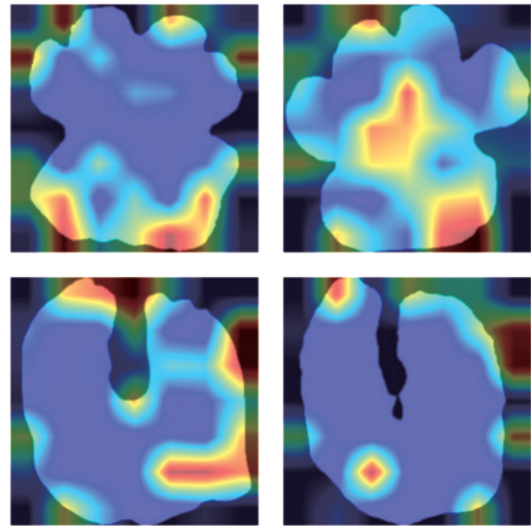


图9 部分雪地足迹影像分类结果热力图

Figure 9 Heat map of partial snow footprint image classification results

响,雪地足迹往往存在轮廓不清晰、无法辨认和无法采集等情况。因此,在数据集制作前,通过人工筛选去除无法体现足迹轮廓特征的影像,以保证数据集的质量。针对不同物种的足迹特点,设定相应的足迹标记方案:对于体型较小且单个足迹特征不明显

的物种(如黄喉貂和东北兔)采用短足迹链形式^[31],以便更好地体现其足迹特征。对于体型较大且单个足迹特征明显的物种(如东北虎和野猪),则采用单个足迹标记方案。同时,为了减少镜头畸变对实际雪地足迹轮廓的影响,采集时统一拍摄高度。在人工标注足迹轮廓时,针对单个足迹,采用封闭曲线进行连续标注,并对轮廓断裂处进行适当连接;对于短足迹链影像,对其中每个足迹进行单独标注,以减少背景信息对分类网络的影响。

5.2 分割结果的影响因素

本研究表明,SOLOv2对不同物种的分割精度差异较大(AP为56.0%~74.4%)。其中对东北虎足迹分割的精度最高,达74.4%,而野猪足迹分割精度最低,仅为56.0%,且整体上有蹄类动物(狍、梅花鹿和野猪)的足迹分割精度普遍低于非有蹄类动物(东北虎、东北兔和黄喉貂)。这一现象可能由以下因素共同造成。(1)足迹形态的复杂性:有蹄类动物独特的活动方式在深雪环境中会留下杂乱的足迹,即使对图像进行筛选,人工标记的轮廓与实际轮廓也存在误差;此外,部分有蹄类动物具有悬蹄特征的雪地足迹,增加了模型在特征提取时的不确定性,从而影响模型的分割性能。(2)雪地环境的干扰:由于雪质地松软,这一特性致使动物在雪地上的足迹清晰度难以保证。受此影响,模型的分割结果往往会出现偏差,具体表现为分割区域相比于实际区域要么偏小,要么偏大。

5.3 分类结果的影响因素

图8中的混淆矩阵结果表明,基于雪地足迹对有蹄类动物(狍、梅花鹿和野猪)的识别存在互相误判的情况。深入分析这一结果,可以发现东北虎和东北兔的识别准确率较高,这可能源于2个关键因素:(1)二者的足迹影像质量更高;(2)二者足迹轮廓具有明显的区分度。相比之下,有蹄类具有相似的足部结构导致足迹轮廓特征相似度较高,其中野猪的识别准确率最低(仅有63%),这与以下因素相关:(1)原始数据采集的影像质量较差;(2)野猪的特殊行走方式造成其在雪地上的足迹特征复杂多变。

5.4 本研究方法的意义和优势

足迹识别是一种有效的非损伤性物种识别手段,在多个濒危物种的相关研究中成效显著。但在中国东北地区仅有关于东北虎的足迹识别研究^[9],

且研究方法主要依赖其足迹的关键点信息的几何特征来识别个体及性别。我国东北地区独特的自然条件为冬季野生动物雪地足迹影像的采集提供了优势。尽管深度学习技术已广泛应用于野生动物保护领域,然而如何结合深度学习和雪地足迹影像高效、准确地识别和监测野生动物仍是个关键问题。为此,本研究提出了一种基于动物雪地足迹和深度学习的物种识别网络以应对此问题。

在本研究方法中,通过整合分割和分类模型实现基于动物雪地足迹所属物种的自动识别,并取得了85.3%的准确率。该方法为野生动物雪地监测提供了一种简单高效的解决方案。实验结果相较于利用足迹识别技术对东北虎的性别判别准确率(97.50%)以及个体识别准确率(93.75%)的结果^[9]仍有差距,但该方法重在研究东北地区野生动物的物种分类识别,是对现有物种识别方法的补充。针对雪地足迹,本研究模型能够自动提取足迹轮廓并对其分类,这一自动化的方法展现了其独特的创新性和应用潜力。

本研究采用SOLOv2实例分割模型与Swin-Transformer-Tiny分类模型,除能够自动识别雪地足迹对应的物种外,还对设备的算力要求不高,适用于手机、无人机等设备,便于野外工作和生态检测。该方法不仅提升了对野生动物活动范围、种群数量和行为习性的监测效率,还为野外工作者和相关专业的人员提供了便利和准确的雪地足迹物种识别工具,对自然保护和生态学研究具有重要意义。

6 结论

本研究将动物足迹识别与深度学习方法结合,对我国东北地区6种哺乳动物的雪地足迹进行影像分类识别。研究表明,SOLOv2分割模型在足迹轮廓分割上表现出色。在人工标记雪地足迹的数据集上,Swin-Transformer-Tiny分类模型达到89.9%的识别准确率。在自动分割的数据集上,Swin-Transformer-Tiny分类模型也达到85.3%的识别准确率。综合考虑模型的识别准确性和实用性,自动分割和自动识别相结合的雪地足迹识别网络相比于人工标注和识别,在处理大量的雪地足迹影像时能够保证高效性、便携性和准确性。

在今后的研究工作中将扩展雪地足迹的特征研究,并扩充雪地足迹数据集中的物种种类及数据量,

同时改进算法,以提高自动化识别的准确度。此外,还可将雪地足迹识别技术应用于手机、无人机等设备,进一步促进东北地区野生动物识别和监测。

参考文献:

- [1] 肖文宏,周青松,朱朝东,等. 野生动物监测技术和方法应用进展与展望[J]. 植物生态学报, 2020, 44(4): 409-417.
- XIAO W H, ZHOU Q S, ZHU C D, *et al.* Advances in techniques and methods of wildlife monitoring[J]. Chinese Journal of Plant Ecology, 2020, 44(4): 409-417.
- [2] 罗旭. 野生动物的痕迹识别[J]. 生命世界, 2019(6): 60-67.
- LUO X. Trace recognition of wild animals[J]. Life World, 2019(6): 60-67.
- [3] ALIBHAI S, JEWELL Z, EVANS J. The challenge of monitoring elusive large carnivores: an accurate and cost-effective tool to identify and sex pumas (*Puma concolor*) from footprints[J]. PLoS One, 2017, 12(3): e0172065.
- [4] JEWELL Z C, ALIBHAI S K, LAW P R. Censusing and monitoring black rhino (*Diceros bicornis*) using an objective spoor (footprint) identification technique[J]. Journal of Zoology, 2001, 254(1): 1-16.
- [5] ELBROCH L M, JANSEN B D, GRIGIONE M M, *et al.* Trailing hounds vs foot snares: comparing injuries to pumas *Puma concolor* captured in Chilean Patagonia[J]. Wildlife Biology, 2013, 19(2): 210-216.
- [6] PIMM S L, ALIBHAI S, BERGL R, *et al.* Emerging technologies to conserve biodiversity[J]. Trends in Ecology & Evolution, 2015, 30(11): 685-696.
- [7] JEWELL Z C, ALIBHAI S K, WEISE F, *et al.* Spotting cheetahs: identifying individuals by their footprints[J]. Journal of Visualized Experiments, 2016(111): e54034.
- [8] ALIBHAI S K, JEWELL Z C, LAW P R. A footprint technique to identify white rhino *Ceratotherium simum* at individual and species levels[J]. Endangered Species Research, 2008, 4: 205-218.
- [9] 顾佳音. 东北虎雪地足迹个体识别技术研究[D]. 哈尔滨: 东北林业大学, 2013.
- GU J Y. Research on Amur tiger (*Panthera tigris altaica*) individual identification from snow footprints[D]. Harbin: Northeast Forestry University, 2013.
- [10] LI B V, ALIBHAI S, JEWELL Z, *et al.* Using footprints to identify and sex giant pandas[J]. Biological Conservation, 2018, 218: 83-90.
- [11] ZUERL M, STOLL P, BREHM I, *et al.* Automated video-based analysis framework for behavior monitoring of individual animals in zoos using deep learning: a study on polar bears[J]. Animals, 2022, 12(6): 692.
- [12] 王文成,蒋慧,乔倩,等. 基于ResNet50网络的十种鱼类图像分类识别研究[J]. 农村经济与科技, 2019, 30(19): 60-62.
- WANG W C, JIANG H, QIAO Q, *et al.* Research on image classification and recognition of ten fish species based on ResNet50 network[J]. Rural Economy and Science-Technology, 2019, 30(19): 60-62.
- [13] NOROUZZADEH M S, NGUYEN A, KOSMALA M, *et al.* Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning[J]. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(25): E5716-E5725.
- [14] GUO S T, XU P F, MIAO Q G, *et al.* Automatic identification of individual primates with deep learning techniques[J]. iScience, 2020, 23(8): 101412.
- [15] 齐建东,马钟添,张德怀,等. 基于BS-ResNeXt-50的密云云地区野生动物图像识别[J]. 林业科学, 2023, 59(8): 112-122.
- QI J D, MA Z T, ZHANG D H, *et al.* Wildlife image recognition in Miyun district based on BS-ResNeXt-50[J]. Scientia Silvae Sinicae, 2023, 59(8): 112-122.
- [16] WANG X L, ZHANG R F, KONG T, *et al.* SOLov2: dynamic and fast instance segmentation [EB/OL]. arXiv (2020-10-23) [2024-04-09]. <https://arxiv.org/abs/2003.10152>.
- [17] LIU Z, LIN Y T, CAO Y, *et al.* Swin transformer: hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10-17, 2021. Montreal: IEEE, 2021: 10012-10022.
- [18] 孙继旭,项凤影,关艳辉,等. 利用红外相机调查黑龙江小北湖国家级自然保护区兽类及鸟类多样性[J]. 安徽农业科学, 2023, 51(6): 106-109.
- SUN J X, XIANG F Y, GUAN Y H, *et al.* Investigation on mammal and bird diversity in Heilongjiang Xiaobeihu National Nature Reserve by infrared camera[J]. Journal of Anhui Agricultural Sciences, 2023, 51(6): 106-109.
- [19] 杨飞飞. 太平沟保护区哺乳动物类群栖息地选择的互作研究[D]. 哈尔滨: 东北林业大学, 2021.
- YANG F F. Study on interaction of habitat selection of mammals in Taipinggou Nature Reserve[D]. Harbin: Northeast Forestry University, 2021.
- [20] 孔维尧,孙权,刘鑫鑫,等. 基于红外相机监测的汪清自然保护区东北豹种群动态[J]. 林业科学, 2019, 55(5): 188-196.
- KONG W Y, SUN Q, LIU X X, *et al.* Population dynamic of Far Eastern leopard (*Panthera pardus orientalis*) in Wangqing Nature Reserve based on infrared camera monitoring[J]. Scientia Silvae Sinicae, 2019, 55(5): 188-196.
- [21] TORRALBA A, RUSSELL B C, YUEN J. LabelMe: online image annotation and applications[J]. Proceedings of the IEEE, 2010, 98(8): 1467-1484.
- [22] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//2017 IEEE Conference on

- Computer Vision and Pattern Recognition (CVPR), July 21–26, 2017. Honolulu: IEEE, 2017: 2117–2125.
- [23] HE K M, GKIOXARI G, DOLLÁR P, *et al.* Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22–29, 2017. Venice: IEEE, 2017: 2961–2969.
- [24] CHEN Y P, DAI X Y, LIU M C, *et al.* Dynamic convolution: attention over convolution kernels [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13–19, 2020. Seattle: IEEE, 2020: 11030–11039.
- [25] 庄前伟, 王志明, 吴龙贻, 等. 基于改进SOLOv2的穴盘幼苗图像分割方法[J]. 南京农业大学学报, 2023, 46(1): 200–209.
- ZHUANG Q W, WANG Z M, WU L Y, *et al.* Image segmentation method of plug seedlings based on improved SOLOv2 [J]. Journal of Nanjing Agricultural University, 2023, 46(1): 200–209.
- [26] FANG Y X, YANG S S, WANG X G, *et al.* Instances as queries [C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV), October 10–17, 2021. Montreal: IEEE, 2021: 6910–6919.
- [27] LIU Z, MAO H Z, WU C Y, *et al.* A ConvNet for the 2020s [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13–24, 2022. New Orleans: IEEE, 2022: 11976–11986.
- [28] MEHTA S, RASTEGARI M. Separable self-attention for mobile vision transformers [EB/OL]. arXiv (2022-06-06) [2024-04-09]. <https://arxiv.org/abs/2206.02680v1>.
- [29] HOWARD A, SANDLER M, CHEN B, *et al.* Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27 to November 2, 2019. Seoul: IEEE, 2019: 1314–1324.
- [30] LIN T Y, MAIRE M, BELONGIE S, *et al.* Microsoft COCO: common objects in context [C]//FLEET D, PAJDLA T, SCHIELE B, *et al.* Computer vision: ECCV, 2014: lecture notes in computer science, Vol. 8693. Cham: Springer, 2014: 740–755.
- [31] 皮库诺夫, 米切尔, 杜尼中科, 等. 远东地区野生动物足迹指南[M]. 李冰, 编译. 哈尔滨: 东北林业大学出版社, 2008: 49.
- ПИКУНОВ Д Г, МИКЕЛЛ Д Г, ДУНИШЕНКО Ю М, *et al.* A field guide to animal tracks of the Far East [M]. LI B, compile. Harbin: Northeast Forestry University Press, 2008: 49.