

单位代码	10475
学号	104754211225
分类号	0213

河南大學

硕土学位论文

(专业学位)

基于卷积神经网络的月球撞击坑识别研究

Research on Lunar Impact Crater Identification Based on Convolutional
Neural Network

培养单位：数学与统计学院

专业学位领域：应用统计

专业学位类别：应用统计硕士

申请 人：王茗涛

指导教师：郑晨 教授

职业导师：张霄杰 高级统计师

二〇二三年七月

Research on Lunar Impact Crater Identification Based on Convolutional Neural Network

A Dissertation Submitted to
the Graduate School of Henan University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Applied Statistics

By
Wang Mingtao
Supervisor: Professor Zheng Chen
Date: July, 2023

摘要

月球撞击坑作为月球表面最普遍的地貌结构单元，记载了月球数十亿年间的撞击历史，对月球撞击坑的精准判读有助于推动月球的探索进程，构建人类对月球的深层次认知。近年来，随着深空探测技术的加速发展，高精度的月球测量数据为月球撞击坑的精准识别提供了研究基础。早期的撞击坑识别研究主要使用形态学、统计学等方法，通过人工构造撞击坑的几何特征对其进行识别，但是这些方法计算量大、操作复杂，且识别精度较低，难以适应当前研究需要。卷积神经网络因其强大的数据表征能力在图像识别领域异军突起，应用卷积神经网络对撞击坑进行精准识别已经成为当前该领域的主流研究方向。然而，由于月球撞击坑数量众多，且不同区域、大小、年代的撞击坑特征各异，常规的卷积神经网络难以对其进行有效区分，因此本论文在已有研究的基础上，对撞击坑识别应用的难点问题进行深入分析，一方面通过改进现有卷积神经网络架构以增强其对撞击坑的学习能力，另一方面针对撞击坑的直径特征提出分层策略来提高对训练数据的利用程度。本论文选取了美国 NASA 实验室提供的高精度月球数字高程模型作为实验数据，依据月球撞击坑专家标记库使用 Python 软件进行自动标注，分别以语义分割任务和目标检测任务为切入点展开实验，对改进网络和分层策略的有效性进行验证。本论文的主要研究内容如下：

(1) 在语义分割任务中，选用 UNet 架构作为撞击坑语义分割实验的基础框架，设计了基于 UNet 架构的 ResUNet、DARUNet 和 MobileUNet 网络，以特征提取能力更强的 ResNet 网络和运算更高效的 MobileNet 网络替换原始 UNet 网络的编码层，加强对撞击坑特征的提取能力和运行效率，同时设计了 DenseASPP 模块以增强网络的多尺度特征捕获能力，并在 ResUNet 的基础上提出了 DARUNet 网络。最后结合各网络的分割结果，设计了后处理模块以改善对撞击坑的整体分割效果。

(2) 在目标检测任务中，基于深度学习框架 MMdetection 重新构建实例分割模型 MaskRCNN 网络，优化模型的训练过程，增加数据增强操作，增强模型泛化性，并根据

月球撞击坑的直径范围提出分层策略，保留大型撞击坑特征，生成多尺度月球撞击坑数据集，提高网络对训练数据的利用程度。

实验结果显示，本论文构建的 ResUNet 网络的 F_1 分数最高达到 68.42%，DARUNet 网络的 F_1 分数达到 68.73%，MobileUNet 网络的 F_1 分数达到 68.45%，均超过经典方法 DeepMoon 的实验精度 66.40%，经过后处理模块后， F_2 分数达到 76.23%，有效提高了预测精度；MaskRCNN 模型在原始数据集的预测结果中 F_1 分数达到了 83.4%，在使用数据分层策略后，MaskRCNN 模型的 F_1 分数达到 84.3%。

关键词：月球撞击坑识别，UNet，分层策略，MaskRCNN

ABSTRACT

As one of the most common geomorphic structural units on the lunar surface, lunar impact craters record the history of billions of years of lunar impacts. Accurate identification of lunar impact craters can help to advance lunar exploration and build a deeper understanding of the moon for humanity. In recent years, with the accelerated development of deep space exploration technology, high-precision lunar measurement data has provided a research basis for the accurate identification of lunar impact craters. Early crater identification studies mainly used morphology, statistics, and other methods to identify them by constructing the geometric features of impact craters artificially. However, these methods have a high computational complexity, are operationally complex, and have low identification accuracy, making it difficult to meet current research needs. Convolutional Neural Networks are emerging in the field of image recognition due to powerful data representation capabilities, and the application of Convolutional Neural Networks for accurate identification of impact craters has become the mainstream research direction in this field. However, due to the large number of lunar impact craters and their different characteristics in terms of region, size, and age, conventional Convolutional Neural Networks are not effective in distinguishing them. Therefore, based on existing research, this thesis conducts an depth analysis of the challenging issues in the identification of impact craters and proposes solutions. On the one hand, by improving the existing Convolutional Neural Network architecture to enhance its learning the ability to impact craters. On the other hand, a hierarchical labeling strategy is proposed for the diameter feature of impact craters to improve the utilization of training data. In this thesis, the high-precision lunar digital elevation model provided by IV

NASA laboratory is selected as the experimental data, which is automatically annotated by Python software based on the expert marker library of lunar impact craters, and the semantic segmentation task and object detection task are respectively used as the starting point to carry out the experiment, to verify the effectiveness of the improved network and stratification strategy. The main research content of this thesis is as follows:

(1) In the semantic segmentation task, UNet architecture is selected as the basic framework for the semantic segmentation experiment of impact crater, and ResUNet, DARUNet and MobileUNet network based on UNet architecture are designed. ResNet network with stronger feature extraction ability and MobileNet network with more efficient operation are used to replace the coding layer of the original UNet network to enhance the feature extraction ability and operation efficiency of the impact crater. At the same time, the DenseASPP module is designed to enhance the multi-scale feature capture capability of the network, and the DARUNet network is proposed on the basis of ResUNet. Finally, the segmentation results of each network are combined, a post-processing module is designed to improve the overall segmentation effect of impact crater.

(2) In the object detection task, the instance segmentation model MaskRCNN network is reconstructed based on the deep learning framework MMdetection, and the training process of the model is optimized. Increase data enhancement operations and enhance model generalization. According to the diameter range of lunar impact craters, a stratification strategy is proposed, and a multi-scale lunar impact craters data set is generated to improve the utilization of training data.

The experimental results show, the F_1 score of ResUNet network constructed in this thesis reaches 68.42%, the F_1 score of DARUNet network reaches 68.73%, and the F_1 score of MobileUNet network

reaches 68.45%. After the post-processing module, the F_2 score reached 76.23%, which effectively improved the prediction accuracy. In the lunar impact crater target detection mission, the F_1 score of the MaskRCNN model reached 83.4% in the prediction results of the original dataset, and that of the MaskRCNN model reached 84.3% after using the data stratification strategy.

KEY WORDS: Lunar impact crater identification, UNet, Layering Strategy, MaskRCNN

目 录

第1章 绪论.....	1
1.1 研究背景和意义.....	1
1.1.1 研究背景.....	1
1.1.2 研究意义.....	2
1.2 撞击坑识别方法研究现状	2
1.2.1 早期研究方法.....	3
1.2.2 深度学习研究方法.....	4
1.3 撞击坑识别中的难点问题	6
1.4 研究创新点及技术路线.....	8
1.4.1 主要创新点.....	8
1.4.2 技术路线.....	9
1.5 论文结构安排.....	11
第2章 卷积神经网络理论概述	13
2.1 卷积神经网络介绍.....	13
2.1.1 卷积神经网络发展历史.....	13
2.1.2 卷积神经网络基本概念.....	13
2.2 深层卷积神经网络介绍.....	22
2.3 深度学习常用的评价指标介绍	24
2.4 本章小结.....	25
第3章 月球撞击坑数据介绍	27
3.1 研究数据说明.....	27
3.2 撞击坑目录介绍.....	28
3.3 数据集准备.....	29
3.3.1 语义分割任务数据集 Dataset1.....	29
3.3.2 目标检测任务数据集 Dataset2.....	31
3.4 本章小结.....	31
第4章 基于改进UNet网络的月球撞击坑识别.....	33
4.1 经典语义分割模型介绍.....	33
4.2 实验设计	36
4.2.1 ResUNet	36
4.2.2 DARUNet.....	37
4.2.3 MobileUNet.....	39
4.2.4 后处理模块.....	41
4.2.5 实验设备介绍及参数设置.....	42

4.3 实验结果分析.....	43
4.3.1 定量分析.....	43
4.3.2 定性分析.....	45
4.4 本章小结.....	47
第 5 章 基于分层策略下的 MaskRCNN 网络的月球撞击坑识别	49
5.1 经典目标检测模型介绍.....	49
5.2 实验设计.....	51
5.2.1 数据增强.....	52
5.2.2 分层策略.....	52
5.2.3 MaskRCNN.....	53
5.2.4 实验平台介绍及参数设置.....	54
5.3 实验结果分析.....	55
5.3.1 定量分析.....	55
5.3.2 定性分析.....	56
5.4 本章小结.....	57
第 6 章 结论及展望	59
6.1 研究结论.....	59
6.2 研究展望.....	60
参考文献.....	61

第1章 绪论

1.1 研究背景和意义

1.1.1 研究背景

作为距离地球最近的星体，自古以来月球就饱受人们的关注，无数神话传说更是给月球蒙上了一层神秘的面纱。随着科学技术的发展，月球的秘密逐渐被人们公开，人类对月球的认识越加深入，对于探索月球的渴望也愈加强烈。早在 17 世纪初，伽利略等科学家就通过望远镜实现对月观测，并简单绘制了月球表面图；到了 20 世纪下半叶，人类的探月活动达到高潮，美国成功发射月球空间探测器，月球探索正式进入空间探测阶段，苏联发射的月球系列探测器成功拍摄到月球背面照片，首次向人类展现出月球的清晰面貌，人类搭载阿波罗号飞船登陆月球，成功实现人类千百年来的登月梦想，随着科技的进步，月亮离我们的‘距离’也越来越近^[1]。

进入 21 世纪，月球探测再次进入活跃期，各航天大国开始了以重返月球为目标的新一轮的探月行动，向月球发射了许多探测卫星和探测器，对月球开展近距离勘测。已经发射的月球探测器包括欧盟的 SMART 探测器、美国的月球探测器、日本的曙光号探测器和中国的嫦娥系列探测器等，如图 1-1 所示。这些探测器收集到的数据不仅包含直观的月球遥感影像，还包括激光图像，地表高度等不同维度的信息，研究人员^[2]依据这些数据开展了大量关于月球探索的学术研究，极大的推动了人类对月探索进程。



中国嫦娥三号地形地貌相机

美国月球勘测轨道器

欧盟SMART-1探测器

图 1-1 各国月球探测器

1.1.2 研究意义

事实上，地外天体的形貌判读与识别一直是国内外学者研究的重要方向。在地外天体中，几乎都存在着相同的地貌特征——撞击坑，它是这些地外天体表面存在的最普遍、最典型的地貌特征之一，其记录了地外天体各时期的地质活动，通过研究这些撞击坑的形成年代、分布和形貌等特点，可以构建对地外天体的深层次认知，对人类探索未知的宇宙具有重要意义，因此对撞击坑的判读也是当前深空探测领域研究的热点问题之一。

月球撞击坑是月球表面通过陨石撞击而形成的环形凹坑，主要由环形山，中央峰，辐射纹等组成，是月球上主要的地貌特征^[1]。据不完全统计，月球上撞击坑的数量达到10万个以上，其记载了月球数十亿年的撞击历史，是月球最为典型的地貌单元和最基本的地貌形态。月球撞击坑的深入探索对于分析月球岩石构造^[3]、估算月表年龄^[4]、探测器软着陆地点的修正^[5]、月球地形地貌^[6]、恢复月球起源和演化历史等研究^[7]都具有重要研究价值。

面对当前月球形貌精细化研究趋势，精准识别月球撞击坑已成为月球地质研究的先决条件。但是月球撞击坑的数目众多，且由于不同区域、不同时代、不同直径的撞击坑形态特征都有所不同，传统的识别方法都存在一定局限，难以准确并高效的识别撞击坑，严重影响后续月球地貌研究。随着人工智能技术的发展，尤其是深度学习方法在计算机视觉领域的突破，不依靠目视解译，自动识别月球撞击坑成为现实，该项技术只需提供少量标签进行训练，就能自动‘学习’到撞击坑的普遍特征，引起研究人员的广泛关注。但是，目前的深度学习模型众多，在月球撞击坑的识别任务表现也各不相同，并且没有对撞击坑的特征进行深入研究，没有充分发掘深度学习方法的潜力。因此本论文针对月球撞击坑的独特特征，通过使用基于卷积神经网络的深度学习模型对其进行识别，并改进现有方法，充分发挥深度学习在月球撞击坑识别应用的优势，实现对月球撞击坑的精准识别，为月球地貌分析研究提供更精确的数据和方法。

1.2 撞击坑识别方法研究现状

受当时技术水平的限制，最早识别撞击坑的方法是借助望远镜对行星进行直接观测，通过专家手工绘制撞击坑的地貌特征。随着科技的发展，人类向宇宙中发射了许多观测

卫星和探测器，通过其携带的先进设备收集了各种类型的观测数据，这些数据的体量高达 TB 级。在数据爆炸的情况下，单纯依靠人工目视判读来识别撞击坑已难以充分利用深空探测数据资源。而随着计算机与人工智能计算的快速发展，借助机器学习方法和计算机的高效运算能力来智能识别撞击坑已成为当前研究的主要趋势。

1.2.1 早期研究方法

早期的撞击坑识别研究主要基于形态学、统计学等方法，根据遥感影像的几何形状、纹理特征、灰度差异等方面人工构造撞击坑特征，然后进行拟合和分类。主要方法有基于几何特征拟合的方法、基于模板匹配的方法、基于霍夫变换及其改进的方法、基于面向对象分类等。

(1) 基于几何特征拟合的识别方法

基于几何特征拟合的识别方法主要是通过提取月球遥感影像中撞击坑的几何形状，对要素特征做拓扑运算，实现对撞击坑的自动识别。

张锋等^[8]根据撞击坑呈环形构造这一形态特征来自动识别提取撞击坑。康志忠等^[9]从光学图像中提取几何特征来选择可能的候选陨石坑，并利用数字高程模型数据中的三维特征挑选最终撞击坑的识别结果；袁悦峰等^[10]通过对遥感影像的灰度梯度图做二值化操作，使用圆形构造拟合撞击坑；徐辛超等^[11]利用灰度梯度统计法统计光学图像的光照方向，使用双值分割方法对图像进行分区，以阴影区域和光照方向为种子点和约束条件，对区域撞击坑的边缘进行拟合。

(2) 基于模式分类的识别方法

基于模式分类的识别方法主要是利用计算机自动提取目标的特征，主要方法包括模板匹配识别、模糊模式识别，神经网络识别等。基于模式分类的识别方法通常是利用遥感影像中撞击坑的形状、纹理、阴影等特征，通过边缘检测、纹理分析等预处理手段，再进行面向对象、模板匹配、模糊变换等操作，最终分离出撞击坑的边界。

Bue 等^[12]利用坡度、纹理和剖面曲率等数字地形信息，应用机器学习方法来自动识别撞击坑；Bandeira 等^[13]人通过对火星撞击坑的自动识别验证了模板匹配算法的适用性，并结合图像处理技术和模式识别技术实现了对火星撞击坑的高效识别；Sawabe 等^[14]对月球遥感影像数据进行了边缘检测、边缘连接等处理，利用模糊霍夫变换识别撞击坑；

岳宗玉等^[15]运用面向对象的分类方法，通过对 Clementine 影像数据进行多尺度分割，构建特征空间从而识别撞击坑并提取撞击坑的边界；Vamshi^[16]使用基于对象的方法从地形数据中检测陨石坑，该方法使用高分辨率图像分割来创建对象，然后使用形状和形态测量标准从对象中提取陨石坑。

传统的撞击坑识别算法的识别效率和精度都具有一定局限性，其中人工构造的特征只适用于小部分具有相同特征的撞击坑识别，难以满足对更大范围撞击坑的识别。而通过计算机学习的特征对分布均匀，边缘清晰的撞击坑有较好的识别精度，但对密集分布的撞击坑识别效果极差。而且这些方法计算复杂，识别效率较低，使得这些方法更难以满足当前研究的实际需求。

1.2.2 深度学习研究方法

计算机视觉领域里包含着两大任务：目标检测(Object Detection)和图像分割(Image Segmentation)，目标检测即给定图片，判断图片中的物体属于什么类别，并定位目标所在的位置；图像分割属于端到端的预测任务，即对图像的每一个元素赋以具体的含义，并判断其属于什么类别，不仅需要检测到目标的类别和位置，还要描绘出目标的具体轮廓，根据任务目标不同分为语义分割(Semantic Segmentation)和实例分割(Instance Segmentation)，前者是对具有不同语义的图像部分进行分割，后者也被视为目标检测任务的延伸。在月球撞击坑识别任务中主要以语义分割和目标检测的需求为主，相关学者也基于这两种任务展开研究。

(1) 基于语义分割方法的研究

语义分割方法来源于全卷积神经网络端到端的思想，这一方法的精度已经在大型数据集上得到了验证，尤其是在其基础上提出的 UNet 网络在生物细胞的分割识别上的优异表现促使研究人员将语义分割方法应用在撞击坑识别领域上。

Silburt^[17]、Lee^[18]等一些学者通过使用简单的 UNet 网络架构分别对月球和火星撞击坑进行了判别；DeLatte 等^[19]探讨了卷积核尺寸和数量对撞击坑的识别影响，并对火星撞击坑进行了识别。这些方法的识别效率相较于人工判读方法有了很大提高，但都面临着相同的问题，既简单的网络架构不能有效地提取撞击坑的特征，难以精识别具有复杂结构的撞击坑。因此一些研究人员^[20-21]在他们工作的基础上对 UNet 网络进行改进，

通过加入残差结构来提高模型的网络容量，增强网络特征提取能力，使其能够学习到不同类型撞击坑的更多特征，以此达到对复杂撞击坑更精确的识别效果。

上述方法提高了识别精度，但大大增加了模型的体量，对计算能力有较高的要求，限制了模型的应用。因此，Yitian Wu 等^[22]通过对残差 UNet 网络进行优化，在保持精度的同时实现了对模型的精简，使基于深度学习的撞击坑识别算法更容易向资源受限的平台推广。此外，Yuqing Mao 等^[23]通过构建双路径卷积神经网络，分别提取月球光学影像与数字高程模型(Digital Elevation Model, DEM)影像的特征并将其融合，一定程度上解决了单一数据源可能会存在特征信息提取不充足的问题，加强了网络模型的泛化性和稳健性。这些识别方法对于直径较大的撞击坑有很好的效果，但是难以区分直径很小或者相互重叠的撞击坑，为了对这些撞击坑进行准确识别，Yutong Jia 等^[24]在基础网络基础上加入了注意力机制(Attention Mechanism)，增强模型对这一部分的撞击坑的学习能力。

(2) 基于目标检测方法

目标检测方法来源于区域卷积网络的思想。在撞击坑识别过程中，由于主要的目的是确定撞击坑的位置及轮廓，这与目标检测的目的是一致的；而且目标检测方法的检测速度相比分割任务要快很多，因此也有很多研究人员使用目标检测模型来识别撞击坑。常见的目标检测模型根据训练步骤可分为单阶段检测和两阶段检测，其中单阶段检测训练速度相对迅速，两阶段检测对于目标的预测较为准确，为保证识别精度，在撞击坑识别领域中两阶段检测算法的应用较多。

较早将目标检测方法应用到撞击坑识别的是 Emami 等^[25]，该工作通过使用经典目标检测模型 FasterRCNN 网络对月球撞击坑直接进行识别，取得了较好的表现。但是 FasterRCNN 网络在提取特征时，只利用最后的深层特征进行检测，这使得很多细节信息被忽略，并且导致了较小尺寸的撞击坑检测性能较差。在目标检测中这种多尺度问题很常见，Tsung-Yi Lin 等^[26]为解决该问题提出了一种可同时保留深层特征和浅层特征的模型——特征金字塔网络(Feature Pyramid Networks, FPN)，在多尺度目标检测中取得了较好的效果。Chia-Yu Hsu 等^[27]利用 FPN 网络提取了不同尺度火星撞击坑的多层次语义信息，并生成了相应的特征图，同时利用霍夫变换将空间分布等先验知识用来初步定位撞击坑的所在位置，最后采用比例尺寸感知分类器(The scale-aware classifier)代替 FasterRCNN 中的分类器，使得检测模型在学习特征时更加关注撞击坑的尺寸，不仅有

助于解决撞击坑多尺度带来的问题，并且有效避免了将非撞击坑识别为撞击坑的现象。而之后何恺明等^[28]提出的实例分割模型 MaskRCNN 网络在 FasterRCNN 网络的基础上，加入 FPN 模块使模型能够保留更多细节信息，提高了对小型撞击坑的检测效果。一些研究人员^[29-30]使用该模型在识别月球撞击坑上均取得了不错表现。

基于单阶段目标检测模型的研究应用如 YOLO 模型及其优化模型，目前已经多个版本，研究人员^[31-32]也将其应用到了撞击坑的自动识别领域，达到了较为精确的结果。Xuxin Lin 等^[33]将单阶段检测与两阶段检测方法的代表模型 (FasterRCNN、FasterRCNN+FPN、CascadeRCNN、SSD、RetinaNet、YOLOv3、FoveaBox、FCOS、RepPoints) 在相同条件下进行实验，验证了两阶段模型识别效果相对更准确。

深度学习方法相比于传统算法具有更强大的特征提取能力，能够充分挖掘数据中的深层次信息，并且由于计算机硬件水平的提升，计算机的计算能力足够满足深层网络的运算需求，因此基于深度学习模型的撞击坑识别精度和速度已经大幅超过传统模型。

1.3 撞击坑识别中的难点问题

尽管深度学习在影像识别领域中已经取得了较好的应用，但在撞击坑识别中依然存在着很多问题需要解决，特别是在数据方面、网络架构和泛化性等方面。考虑到不同类型撞击坑识别具有一定程度的类似性，具体探析深度学习在撞击坑判读领域需要解决的难题。

(1) 数据问题

在基于深度学习方法的撞击坑识别中，最常用的数据类型是光学影像和数字高程模型。相较于一般影像，这两类深空影像数据都有着自身的特点。其中光学影像分辨率高，并且能够保留目标的形态特征，是使用频率最高的研究数据。然而，深空光学影像中由于没有差异性显著的地表附着物，其地表光谱特征，即外观是高度相似的，而光照形成的阴影或光亮区域往往成为撞击坑识别的典型特征，但由于在拍摄过程中光照角度的影响，其所提取的特征对于撞击坑与背景信息的可区分度不足，直接应用深层卷积神经网络进行识别的效果不佳。数字高程模型是对光学影像数据的有力补充，其包含大量的地形特征信息，在近几年的深度学习方法中得到了研究人员的关注。

深度学习在深空数据判读中面临的另一个问题是标签数据。目前，大部分研究人员在进行数据集的制作中倾向于选择早期发布的专家撞击坑目录，比如月球撞击坑的 Head、Povilaitis 目录等。但是在这些目录中依然存在着不少遗漏或者错判的撞击坑，不完善的数据集会给网络训练带来难度甚至误导。而较为完整的 Robbins 目录可靠性暂时难以得到验证，并且其广泛的标注范围也大大增加了研究人员的工作量。因此，目前尚缺乏一个完整的撞击坑标签数据库。如何提高撞击坑标签数据集的完整性、可靠性、可用性是制约撞击坑识别精度的关键因素。

(2) 特征提取问题

深度学习中通过多层的神经网络来提取目标的特征，使得目标对象与背景信息在提取或学习的特征中有较为显著的差异，亦即所谓的可分性。对于撞击坑来说，它的特征是复杂而多样的，天体不同区域、不同年代的撞击坑的特征可能完全不同，这对深度学习模型的特征学习带来了极大的考验。深度学习中不同的网络架构对于目标的特征表示各不相同，因此怎样设计适合深空遥感影像的深度学习网络架构，来判读多尺度和不同空间分布的撞击坑特征是需要考虑的。

具体而言，不同尺度撞击坑的形貌特征存在差异。撞击坑的直径可从几百米一直到几百公里，在撞击坑的形貌特征分类研究中^[34]，根据撞击坑的直径范围可分为简单撞击坑、复杂撞击坑以及撞击坑盆地，其中简单撞击坑的直径一般在 15 公里以内，呈碗状、底部平坦、边缘平滑；复杂撞击坑直径在 15 至 100 公里之间，边缘呈锯齿状，中心区域出现突出的中央峰；撞击坑盆地的直径则一般都大于 100 公里，出现峰环，且随着直径的增大，盆地内部会出现多个峰环。单从影像中看，大直径的撞击坑与小直径的撞击坑是具有显著区分度的。如何对多尺度不同类型的撞击坑进行准确区分对深度学习方法提出了很高要求。

不同区域的撞击坑其空间分布也具有很大差异，撞击坑主要是由于小行星或陨石随机撞击而产生的，其分布应为随机的。但由于月球面向地球的一面发生过大规模的火山活动，大多数的撞击坑包括撞击盆地因熔岩而被抹平，形成了月海，因此撞击坑呈现数量少且稀疏的特点。而月球背面的撞击坑多被保留下来，形成月陆，呈现撞击坑数量多且密集的特点。同时撞击坑之间也存在着不同的空间关系，最多的是离散关系，即独立互不影响的撞击坑，这些撞击坑基本都为小型简单撞击坑；其次是相交关系，即多个撞

击坑在不同程度上相邻甚至相交；最后是包含关系，即在大型撞击坑内部存在着多个小型撞击坑，这通常是由于二次撞击或多次撞击造成的。这些空间分布特征也为深度学习模型的识别造成极大的挑战。

(3) 泛化性问题

在深度学习模型中最令人关注的问题之一就是模型的泛化能力，即模型在训练集中取得的表现能否在测试集中达到同样精度，尤其在撞击坑识别中，我们更关注的是全月球的撞击坑分布识别。但我们只能用部分区域对模型进行训练，并希望训练好的模型能对全月的撞击坑进行自动提取。因此如何使得模型克服不同区域的差异是一大难题，同时在月球上的识别方法能否同步迁移到其他行星的撞击坑识别，也是研究人员着重关注的问题。

1.4 研究创新点及技术路线

1.4.1 主要创新点

本论文的研究目标为应用深度学习方法实现自动提取撞击坑，验证深度学习模型在撞击坑识别领域的准确性和高效性，在已有研究的基础上改进相关模型，提高其识别精度和效率；同时对月球撞击坑特征进行分析，根据撞击坑直径提出分层策略，提高对撞击坑多尺度特征的提取能力。整篇论文共分为背景阐述、实验基础和设计介绍、结论与展望三部分。本论文的主要创新点体现在以下两个方面：

(1) 改进 UNet 网络

本论文在经典语义分割模型 UNet 网络的基础上进行改进，借助 ResNet 架构的强特征提取能力和 MobiliNet 架构的轻量化设置，提出了 ResUNet 网络和 MobileUNet 网络；同时设计了 DenseASPP 模块以增强网络的多尺度特征捕获能力，并在 ResUNet 的基础上提出了 DARUNet 网络。

(2) 分层策略

本论文结合月球高程数据的特点，针对不同直径范围的月球撞击坑进行分析，提出数据分层标记策略，将撞击坑按直径进行分层标记，并通过下采样方式将其归一为同一范围，方便模型的训练，进一步加强模型对不同尺度撞击坑的学习能力。

具体工作流程如图 1-2 所示：

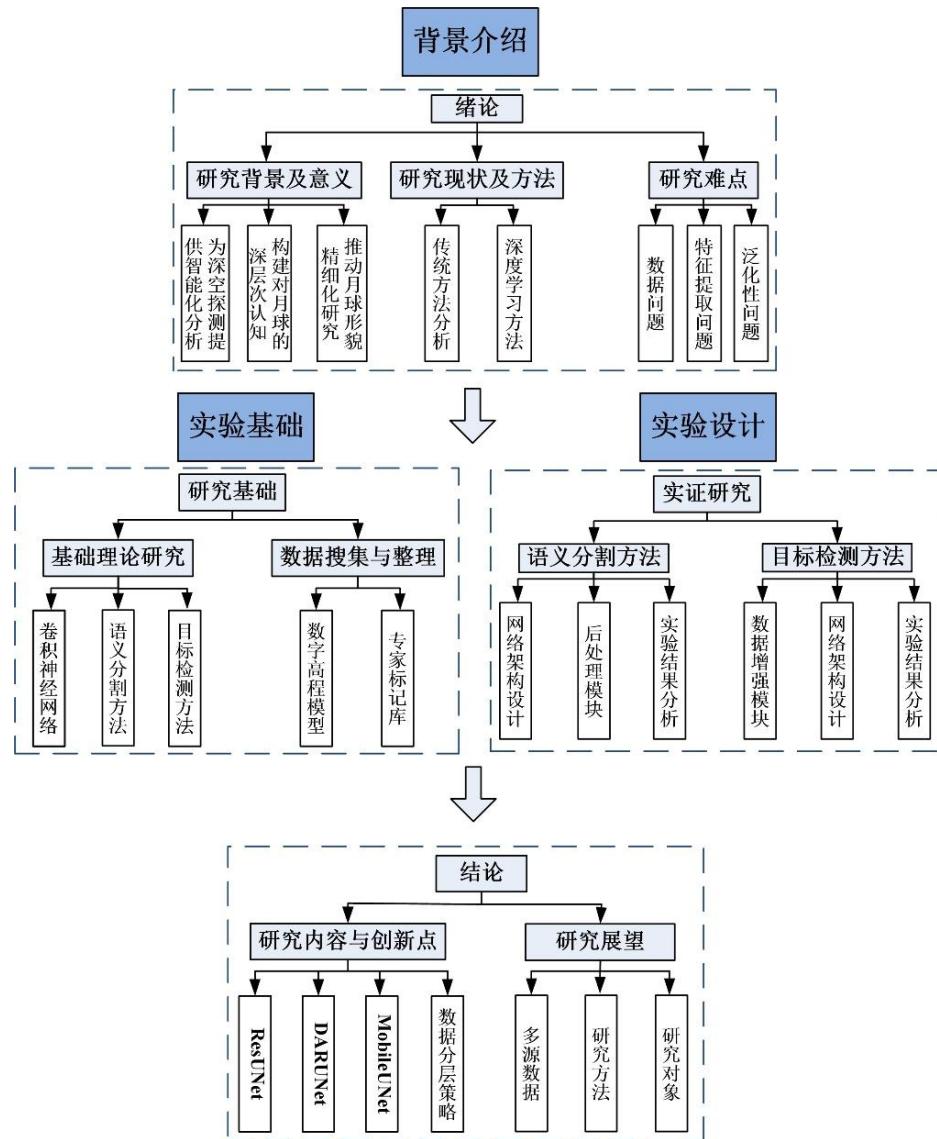


图 1-2 工作流程图

1.4.2 技术路线

本论文针对月球撞击坑自动识别任务进行实验，基于深度学习方法构建撞击坑智能识别模型，并对全月撞击坑进行识别。实验流程主要分为数据获取与处理、模型改进及应用、精度评价和结果分析部分，实验具体步骤如下所示：

(1) 月球数字高程模型数据获取及预处理

针对月球撞击坑数据问题，通过数据优势分析，选取月球数字高程模型影像数据作为实验数据，同时根据公布的专家撞击坑目录使用 Python 软件对撞击坑进行自动标注，生成用于语义分割任务和目标检测任务的数据集。

(2) 基于深度学习方法的撞击坑识别方法应用

在经典的语义分割模型 UNet 网络的基础上分别引入 ResNet 架构和 MobileNet 架构，加强网络的特征提取能力和训练效率，并结合空间空洞卷积金字塔池化和 DenseNet 网络中的密集连接的思想，提出 DenseASPP 模块以增强网络的学习能力，最后设计了后处理模块，对分割结果进行融合，提高模型的泛化能力；

使用 MMdetection 平台构建基于 MaskRCNN 网络的月球撞击坑实例分割模型，依据月球撞击坑数据直径特征提出训练数据分层策略，提高模型的泛化性能，增强对月球撞击坑的识别效果。

(3) 撞击坑检测模型精度评价和结果分析

本论文使用深度学习方法常用的指标(召回率、准确率、 F_1 分数和 F_2 分数)对模型的预测结果进行评价，同时使用模型计算复杂度和模型参数对模型的性能进行评估，并对分割和检测结果做可视化展示，将深度学习方法在月球撞击坑识别任务中存在的问题展开进一步分析。

实验的技术路线如图 1-3 所示：

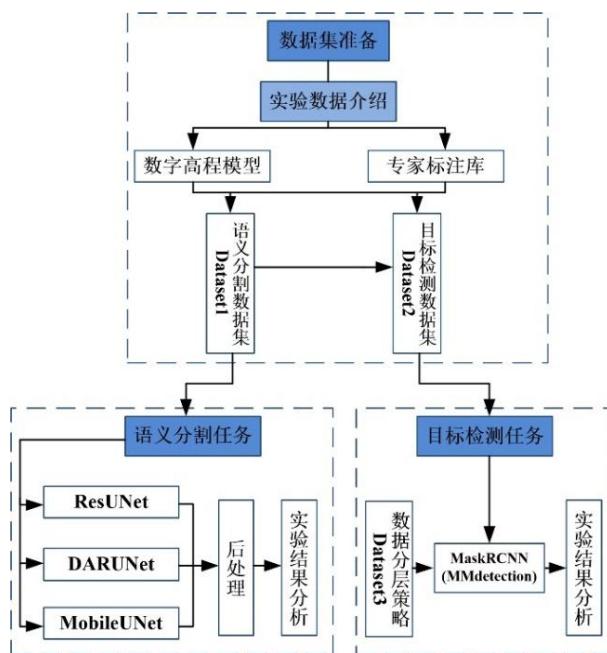


图 1-3 技术路线图

1.5 论文结构安排

本论文以月球数字高程模型为实验数据，使用深度学习方法对其存在的撞击坑进行检测识别。论文各章节主要内容如下：

第一章，绪论。主要介绍论文的研究背景和研究意义，并从国内外的研究现状归纳整理了当前研究存在的难点问题，简略介绍了本论文的研究内容和技术路线。

第二章，卷积神经网络理论阐述。主要介绍卷积神经网络的发展历史，对其中的基本概念做简单概述，并对深层卷积网络的构成进行详细介绍。

第三章，数据介绍。收集整理实验所需数据，对获取的数据做预处理操作，并制作成深度学习模型实验所需要的数据集格式。

第四章，UNet 网络的改进及在月球撞击坑的应用。首先在 UNet 网络基础上引入 ResNet 和 MobileNet 架构，构建了 ResUNet 网络和 MobileUNet 网络，并基于 DenseNet 和空洞空间卷积金字塔池化思想设计了 DenseASPP 模块，构建了 DARUNet 网络，然后通过后处理模块对网络的分割结果进行进一步整合，增强网络对撞击坑的识别能力，提高应用效果，最后通过量化指标对模型检测结果进行评估，并将分割结果可视化展示。

第五章，基于分层策略的月球撞击坑识别。基于撞击坑直径特征将撞击坑分为多个层次，按层次对撞击坑进行标记和下采样，提高数据的利用程度，并基于 MMdetection 平台构建实例分割网络 MaskRCNN 模型对月球撞击坑进行识别，最后通过量化指标对模型检测结果进行评估，并将检测结果可视化展示，进一步总结深度学习方法在月球撞击坑识别领域存在的优点和不足之处。

第六章，结论与展望。总结了本论文的具体工作和主要创新点，对研究工作中存在的不足做出说明，并在现有工作的基础上，提出未来研究设想。

第 2 章 卷积神经网络理论概述

2.1 卷积神经网络介绍

2.1.1 卷积神经网络发展历史

卷积神经网络(Convolutional Neural Network, CNN)是 Hubel 等^[35]从猫的神经结构研究中得到灵感所提出的一种神经网络的架构。1989 年 LeCun 等^[36]将反向传播算法应用于影像分类的卷积神经网络，在手写数字体识别数据集中取得当时最好的识别效果，经过算法的不断改进，LeCun^[37]在论文中构建了七层卷积神经网络模型 LeNet-5，这也标志着具有完整体系的卷积神经网络正式面世。但受限于当时硬件的条件和支持向量机 SVM^[38]的兴起等原因，上述结果并没有得到很好的推广应用。

近年来，随着人工智能技术的飞速发展，卷积神经网络又再次迎来发展机遇，2012 年 Krizhevsky 等^[39]在 ImageNet 竞赛中设计的 AlexNet 网络以绝对优势取得冠军，在计算机视觉领域引发极大轰动，成为深度学习发展的历史转折点，开启了深度学习爆发时代。2013 年，深度学习被 MIT Technology Review 认为是本年度最具有突破意义的十大技术之一。2015 年 LeCun 等^[40]在 Nature 杂志上发表了一篇关于深度学习的综述文章，系统地总结了深度学习的发展进程。随着深度卷积神经网络在大规模数据集的分类任务的优异表现，基于卷积神经网络的深度学习方法已经成为当前研究的热点。

2.1.2 卷积神经网络基本概念

卷积神经网络主要由输入层、隐层和输出层三部分组成。训练数据经过输入层，在隐层中通过卷积、激活和池化等操作提取影像特征，之后将得到的特征图拉伸展开并输入到分类层，通过构建全连接网络作为分类器，在输出层给出分类结果。在深度学习的模型训练阶段，会计算输出结果中每一类别的概率，然后通过与真实标签比较，计算损

失函数值，通过反向传播算法^[41]迭代更新参数，并以此确定最终的模型参数。具体架构如图 2-1 所示：

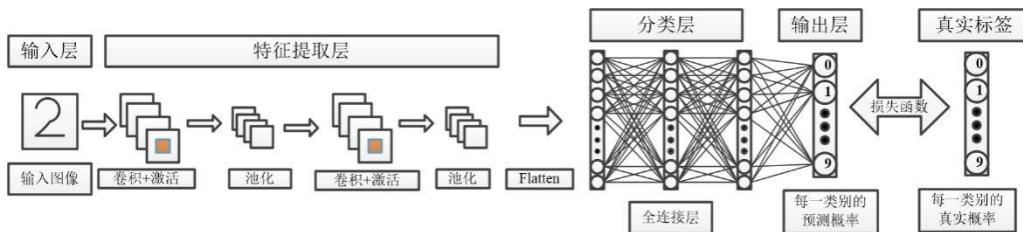


图 2-1 卷积神经网络架构

为深入了解卷积神经网络的工作原理，下面我们具体介绍卷积神经网络中各个模块的作用和具体运算过程。

(1) 卷积层

由于 CNN 主要处理图像数据，故 CNN 中的卷积运算其实是二维和三维的离散卷积，二维离散卷积是在输入图像上用二维卷积核从头到尾将其遍历一遍，然后对应位置相乘再相加，我们通过图像形式具体展示卷积的操作过程，图 2-2 即为卷积核大小为 3、步长为 1 的卷积操作：

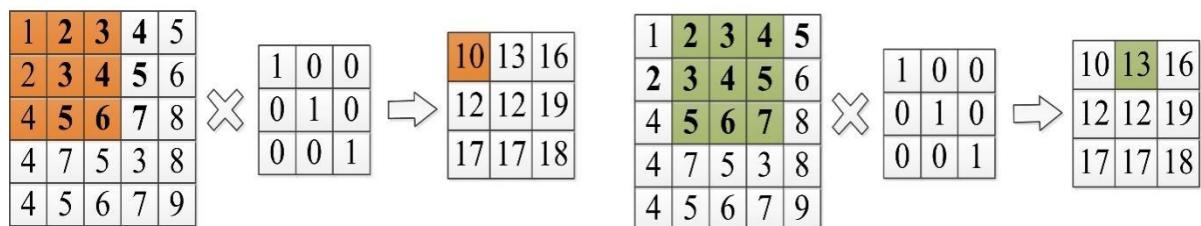


图 2-2 二维离散卷积计算过程

上述卷积操作会使特征图越卷变小，而且由于卷积操作是从输入图像与卷积核重叠的地方开始，对于边缘信息来说提取的不够充分，在经过多次卷积操作之后甚至会造成部分边缘信息丢失。为解决此问题，卷积操作可以通过使用外部用像素 0 进行填充的方式来保证对边缘信息的保护，同时还可以控制输出特征图的尺寸，在 CNN 中被广泛应用。图 2-3 即为填充为 1 的卷积操作：

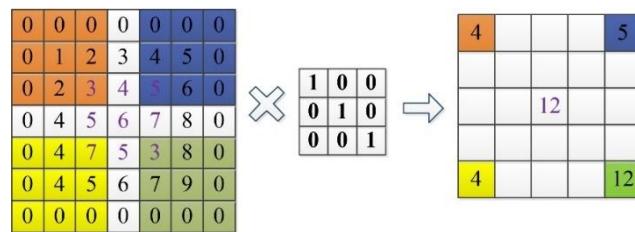


图 2-3 填充为 1、卷积核大小为 3、步长为 1 的卷积

同理也可推广得到多维离散卷积的运算过程，当输入数据具有多个维度时，需要用多个卷积核去遍历输入图像的每一维特征图，其中每一个卷积核需要与不同维度的特征图对应位置相乘再求和，之后再把该卷积核在不同维度的特征图上相应位置的求和结果再相加，最后每一个卷积核对应一个输出结果，该过程如图 2-4 所示：

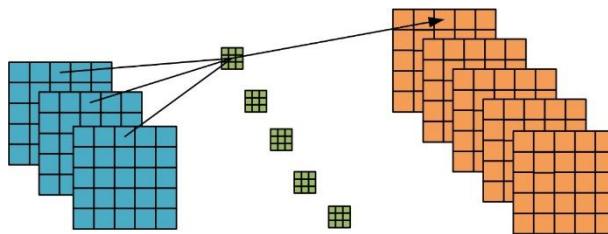


图 2-4 多维数据的离散卷积计算过程

在卷积操作中，卷积核的大小，步长以及是否填充这些参数需要提前给出，但是在卷积神经网络中由于全连接层的限制，输入输出特征图的大小必须固定，而不同的参数设置决定了输出特征图的尺寸，因此我们需要根据卷积计算公式来设置这些参数，保证特征图是我们需要的大小，卷积后的尺寸计算公式见式(2-1)：

$$N = \frac{H-K+2P}{S} + 1, \quad (2-1)$$

其中， N 为卷积后的尺寸， H 为卷积前的尺寸， K 为卷积核尺寸， P 为填充， S 为卷积步长。

除了上述卷积外，还存在空洞卷积，相比于标准卷积操作，增加了一个膨胀率参数 (dilation)，可视作将卷积核的感受野扩张 N 倍，在膨胀之后的特征图上做卷积操作，增大卷积核的感受野，在保持参数个数不变的情况下，使卷积操作可以获得更大范围的信息。通常在空洞卷积操作中，将膨胀率参数和填充参数数值保持一致，这样可以使卷积后特征图的尺寸不变。具体操作如图 2-5 所示：

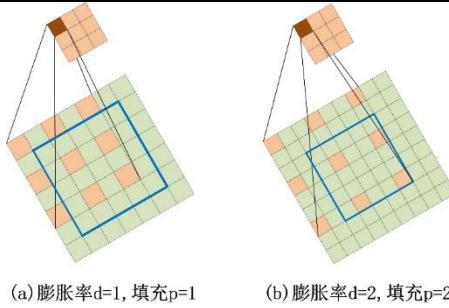


图 2-5 空洞卷积操作

空洞卷积的卷积核真实感受野计算公式见式(2-2):

$$K' = K + (K - 1) \times (d - 1), \quad (2-2)$$

其中, K 为卷积核尺寸, K' 为感受野尺寸, d 为膨胀率。

(2) 池化层

池化(Pooling)又被称为下采样, 其操作是将输入特征图的每一个局部区域(一般为2x2或者3x3的网格)用平均值、最大值等简单统计量代替, 使输出特征图的大小变为原来的四分之一或九分之一。池化通常在卷积操作之后进行, 在不改变目标的整体结构信息的条件下, 通过扩大感受野, 使网络能够捕获到图像更本质更宏观的深层特征, 同时可以降低特征图的尺寸达到减少网络参数量的目的, 防止发生过拟合现象。常见的池化操作有最大池化(Max Pooling)和平均池化(Average Pooling):最大池化使用最大值代替周边信息, 优点是可以让网络学习到图像的边缘和纹理信息; 平均池化使用周边元素的均值作为代表, 优点为保留图像的背景信息, 上述两种池化的具体操作示意如图 2-6 所示:

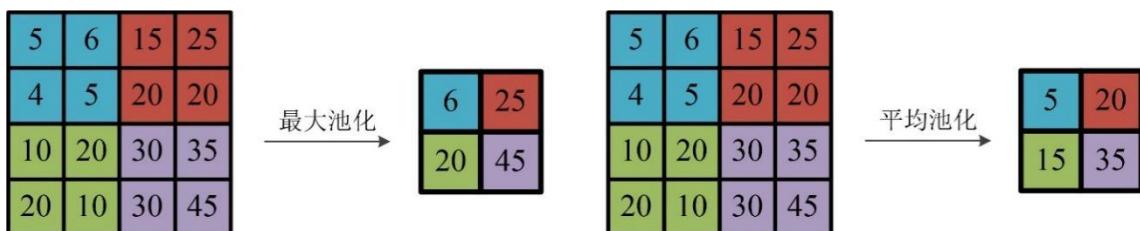


图 2-6 最大池化和平均池化操作示意图

(3) 激活函数

在卷积神经网络中, 如果只进行卷积和池化的操作, 经过多次运算之后也不过是单纯的线性组合, 并不能有效提取影像复杂的特征, 因此需要在卷积和池化操作中间引入非线性的激活函数来增强网络的表达能力。常见的激活函数有 Sigmoid 函数和 ReLU 函数等。

Sigmoid 函数及导函数公式见式(2-3)、式(2-4):

$$f(x) = \frac{1}{1+e^{-x}}, \quad (2-3)$$

$$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = f(x)(1-f(x)). \quad (2-4)$$

Sigmoid 函数又称为 Logistic 函数, 因其具有处处连续且单调可微的优良性质, 被广泛应用于早期的神经网络模型中。但其一阶导数有界, 最大值为 0.25, 随着输入值过大或过小, 导数值会趋近于 0, 当导数值过小时, 会影响模型参数的更新和学习效果。具体而言, 当导数较小时, 基于链式法则更新权重的反向传播算法, 在接近输出层的几层网络还可以正常更新, 但越接近输入层时, 其梯度在多次与小于 0.25 的导数值相乘已经无限趋近于 0。此时, 神经元的权重因为梯度过小而无法得到有效更新, 会出现梯度消失(Vanishing Gradients)的现象, 因此 Sigmoid 函数逐渐被其他激活函数所替代。

ReLU 函数及导函数公式见式(2-5)、式(2-6):

$$f(x) = \max(0, x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}, \quad (2-5)$$

$$f'(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}. \quad (2-6)$$

ReLU 函数又称为修正线性单元(Rectified Linear Unit), 是一种分段线性函数, 当输入为负时, 输出为 0, 梯度也恒为 0; 当输入为正时为恒等映射, 梯度恒为 1。ReLU 函数解决了 Sigmoid 函数中存在的梯度消失问题, 并且其求导简单, 计算量小, 是当前深度神经网络模型中使用最广泛的激活函数之一。但是在少数情况下, 当某些神经元的输入为负时, 其梯度为 0, 在反向传播过程中这些神经元将不会被再次激活, 权重不再得到更新, ReLU 函数将不起作用, 出现了 Dead ReLU 问题。针对该问题, 基于 ReLU 函数的一些变形函数被相继提出, 比较典型的一种变形是 Leaky ReLU 函数, 该函数及导函数公式见式(2-7)、式(2-8):

$$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases}, \quad (2-7)$$

$$f'(x) = \begin{cases} \alpha, & x < 0 \\ 1, & x \geq 0 \end{cases}. \quad (2-8)$$

通过定义可以看出, Leaky ReLU 函数通过给输入中存在的负值赋以一个小的权重 α (通常 $\alpha = 0.01$), 将负值的梯度变为 α , 解决了 Dead ReLU 问题, 在应用中取得了较好

的效果。诸如此类的函数还有限制 ReLU 函数输出值的 ReLU6 函数，该函数可将 ReLU 函数的输出值限制在 0 到 6 之内，其计算公式见式(2-9):

$$f(x) = \min(6, \max(0, x)). \quad (2-9)$$

还存在着一些基于 Sigmoid 函数和 ReLU 函数改进的激活函数，这些激活函数通常使用在具有特殊条件的神经网络中，比如 Swish 激活函数，其公式见式(2-10)、式(2-11):

$$f(x) = x * \text{Sigmoid}(\beta x), \quad (2-10)$$

$$f'(x) = \beta * f(x) + \sigma(\beta x) * (1 - \beta * f(x)). \quad (2-11)$$

当 $\beta = 0$ 时，Swish 函数为线性函数，当 $\beta = \infty$ 时，Swish 函数近似为 ReLU 函数，因此 Swish 激活函数可看作是介于线性函数与 ReLU 函数之间的平滑函数，由于该函数没有上界、存在下界，并且具有平滑、非单调的特性，可使神经网络具有更丰富的表现能力。Hswish 激活函数是将 Swish 激活函数的 Sigmoid 函数替换为 ReLU6 函数，在对深度学习模型效果影响不大的前提下，显著降低了计算量，被广泛应用于各种轻量级网络中，其计算公式见式(2-12):

$$f(x) = x * \frac{\text{ReLU6}(x+3)}{6}. \quad (2-12)$$

(4) 全连接层

在特征提取层完成对图像特征的提取后，输出特征图被展开成一维特征向量，并将其输入进全连接层中。在全连接层中输入神经元与输出神经元需要全部连接起来，并且可以通过设置最后一层神经元的个数等于类别数，起到分类的作用，其结构如图 2-7 所示:

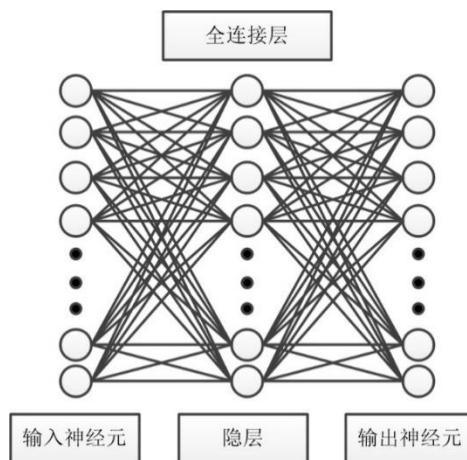


图 2-7 全连接层

(5) 损失函数

损失函数是用来度量真实值 Y 与模型预测值 $f(x)$ 之间的差异程度，通常用 $L(Y, f(x))$ 表示。损失函数越小，代表模型的预测效果越佳。在模型的训练阶段，输入数据通过网络模型的前向传播计算输出预测值，然后损失函数计算出预测值和真实值之间的损失值，再以降低损失函数为目标，经过反向传播的链式法则更新优化网络模型参数，提高模型的训练精度。常用的损失函数有交叉熵损失函数(Cross Entropy Loss)、二分类交叉熵损失函数(Binary Cross Entropy Loss)、 $SmoothL_1$ 损失函数等，以下是各个损失函数的具体计算公式。

交叉熵源于信息论，主要用来刻画两个概率分布间的差异性信息。在 CNN 模型中，交叉熵损失函数用来刻画模型预测值与真实值两个分布之间的相似程度，交叉熵的值越小，两个概率分布就越接近，交叉熵损失函数是目前卷积神经网络最常用的一种损失函数，尤其在离散分类问题中的应用，其中在二分类任务中使用二分类交叉熵损失函数(Binary Cross Entropy Loss)。交叉熵损失函数计算公式见式(2-13)、式(2-14):

$$CELoss = L(Y, f(x)) = - \sum_{i=1}^N Y_i \log f(x_i), \quad (2-13)$$

$$BCELoss = L(Y, f(x)) = - \frac{1}{N} \sum_{i=1}^N [Y_i \log f(x_i) + (1 - Y_i) \log (1 - f(x_i))]. \quad (2-14)$$

在二分类任务中，需要将预测结果固定在 0-1 之间，该损失函数将二分类损失函数与 Sigmoid 函数结合，使其在计算过程中数值更稳定，`BCEWithLogitsLoss` 函数计算公式见式(2-15):

$$L(Y, f(x)) = \text{Sigmoid} * BCELOSS. \quad (2-15)$$

$SmoothL_1$ 损失函数是对 L_1 损失函数和 L_2 损失函数的优化整合，其对异常值的处理比 L_2 损失函数更好，同时减少了 0 附近梯度的计算，在目标检测任务中应用较为广泛。

$SmoothL_1$ 损失函数计算公式见式(2-16):

$$L(Y, f(x)) = \begin{cases} \frac{1}{2}(Y - f(x))^2, & |Y - f(x)| < 1 \\ |Y - f(x)| - \frac{1}{2}, & |Y - f(x)| \geq 1 \end{cases}. \quad (2-16)$$

(6) Dropout 层

当在小型数据集上使用深层神经网络来训练时，过拟合情况很容易出现。如果模型的参数太多，而训练样本又太少，训练出来的模型很容易出现过拟合，具体表现在：模

型在训练数据上损失函数值小，预测准确率高，但在验证数据上损失函数值大，预测准确率低。基于此问题，Hinton 等^[42]提出了 Dropout 的策略，在前向传播的过程中，随机断开神经网络部分神经元之间的连接，通过牺牲一定的运行速度使模型具有更强的鲁棒性，提高神经网络的泛化性，具体操作步骤如下：

- 1、随机删掉一定比例的神经网络隐层神经元，输入输出神经元保持不变；
- 2、通过修改后的网络前向传播，通过反向传播更新未被删除的神经元的权重；
- 3、恢复删掉的神经元，权重保持不变。

(7) 批归一化

在深层神经网络中各层神经元的参数会在训练过程中不断得到更新，随着参数的不断变化，各层的输入分布也随之产生差异，会出现内部协变量偏移现象(internal covariate shift)，导致梯度消失、梯度下降收敛速度变慢等现象，为解决此问题，谷歌的研究人员^[43]在 2015 年提出一种批归一化(Batch Normalization, BN)的训练技巧，将输入数据归一化，使每层神经层的输入分布都为标准正态分布，该操作一般在激活函数之前进行，能够有效缓解深层网络中的梯度消失问题，使训练过程更稳定，具体操作步骤如表 2-1 所示：

表 2-1 Batch Normalization 算法步骤

输入：一个批次的数据 $B = \{x_1, x_2, \dots, x_n\}$ ，待学习的参数 γ, β ；

- 1) 计算一个训练批次数据的样本均值 $\mu_B = \frac{1}{n} \sum_{i=1}^n x_i$ ；
- 2) 计算每一个训练批次数据的样本方差 $\sigma_B^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_B)^2$ ；
- 3) 对该批次的训练数据做归一化 $\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ ， ϵ 是防止出现除数为 0 而设置的较小正数；
- 4) 对标准化的输入 x_i 再次进行缩放和平移， $y_i = \gamma \hat{x}_i + \beta$ ；

输出： $y_i = \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i)$ 。

(8) 上采样

在卷积神经网络中，经过多次卷积池化操作，输出特征图的尺寸要小于输入影像，为了实现端到端的精准预测，需要将提取的特征图尺寸上采样恢复至输入影像的尺寸。

在深度学习中，上采样一般有三种方法：插值法、反卷积法、反池化法，其中插值法和反卷积法应用最为广泛。

插值法是根据原有像素值和空间关系确定所要补充的空白像素值，将输入影像放大到需要的尺寸，常用方法有最近邻插值法(nearest interpolation)、单线性插值、双线性插值法(bilinear interpolation)、双三次插值法(bicubic interpolation)等。最近邻插值是用与空白像素点最近的像素值填充，线性插值是用像素之间空间位置关系确定像素值，具体操作可见图 2-8 和图 2-9：

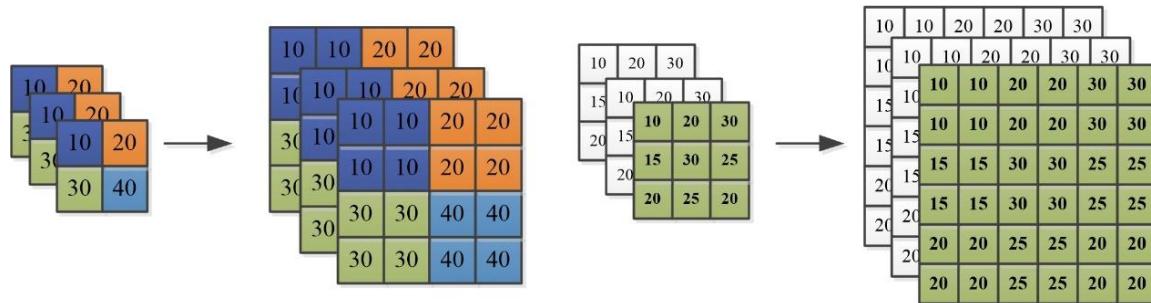


图 2-8 最近邻插值

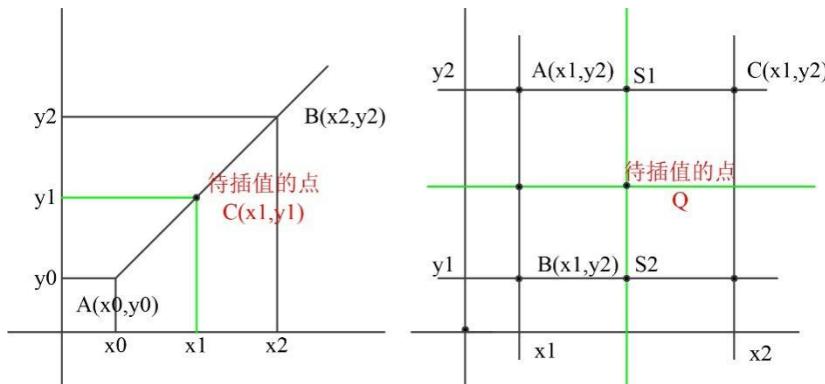


图 2-9 线性插值与双线性插值

反卷积又称为转置卷积，顾名思义就是卷积的逆运算。反卷积运算首先将特征图补零，然后用一定大小的卷积核通过卷积运算将特征图放大。反池化法和反卷积法类似，在最大池化过程中记下最大元素的位置，反池化上采样就将该元素放入放大后的位置中，其余位置补零，起到放大影像的效果。示例可见图 2-10：

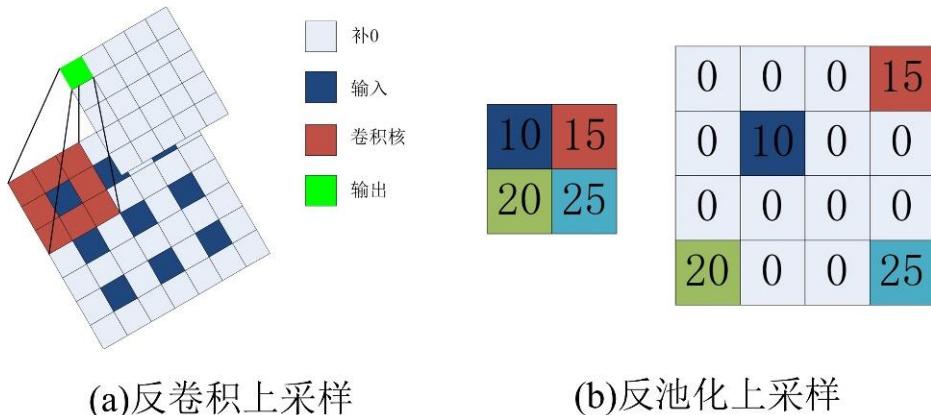


图 2-10 反卷积上采样与反池化上采样示意图

2.2 深层卷积神经网络介绍

随着人工智能技术的快速发展，越来越多基于卷积神经网络的深度学习模型被陆续提出。尽管模型架构千变万化，但其都依托于深度学习的基础模型，下面我们通过介绍几个深度学习模型中较为典型的网络结构，以此来探寻其发展情况。

(1) AlexNet

在 2012 年的 ImageNet 竞赛中，深层卷积神经网络 AlexNet 在识别精度上打破常规方法记录，以绝对优势摘得比赛桂冠。相比于传统的卷积神经网络，AlexNet 在网络深度上有了提高，在不同层中使用不同大小的卷积核提取特征，以 ReLU 函数作为激活函数，采用最大池化层来扩大感受野，最后使用全连接网络进行分类，在训练过程中通过 Dropout 策略避免模型出现过拟合，利用双 GPU 运算减少训练时间，最终的网络架构包含 5 个卷积层和 3 个全连接层，并通过实验证明了网络中每一层的重要性。

(2) VGG 网络

在 AlexNet 取得成功之后，研究人员开始对神经网络的深度有了更高的要求，随着网络层数的加深，模型的学习能力会越强、识别的效果越好已经逐渐成为共识。Simonyan 等^[44]提出了深度卷积神经网络 VGG 网络，在 AlexNet 的基础上加深了网络层数，该论文中根据卷积层数分为多个版本，其中 VGG16 网络应用最为广泛。

(3) ResNet

传统的 CNN 在传递信息时，由于卷积池化操作不可避免地会出现信息丢失等问题，如 VGG 等网络，在随着网络层数的加深，模型的识别精度在趋于稳定后会迅速下降，

出现网络退化的现象。因此在 VGG 网络的基础上, 何恺明等^[45]通过提出残差学习的思想, 在每层之间添加“短路链接(Shortcut Connection)”将原始输入信息映射到输出, 使得网络只用学习残差部分(输入与输出的差值), 通过这种特殊的映射能使网络在出现退化时能提前终止训练, 在一定程度上解决了网络层数加深所带来的问题, 其中残差学习模块结构如图 2-11 所示:

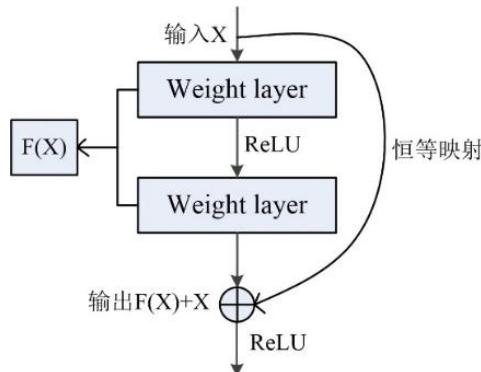


图 2-11 残差学习模块(Residual learning module)

在残差结构中, x 为上一层的输出, 通过卷积层得到残差块的输出结果 $F(x)$, 再通过“短路”链接将输入 x 映射到输出, 再将 x 与 $F(x)$ 相加, 最后使用 ReLU 激活函数对相加结果进行激活, 得到残差结构的最终输出结果。残差模块有两种类型: 图 2-12(a)是常规的残差模块, 由两个 3×3 的卷积层组成, 一般应用在层数比较浅的网络模型中; 图 2-12(b)又称瓶颈残差模块, 由 1×1 , 3×3 , 1×1 三个卷积依次堆叠而成, 其中第一个 1×1 卷积层的作用是通过减少通道数使 3×3 的卷积层能以较低维度的输入进行运算, 减少参数数量, 提高运算效率, 第二个 1×1 卷积层用来恢复通道数, 这种残差模块一般应用在较深的网络模型中。

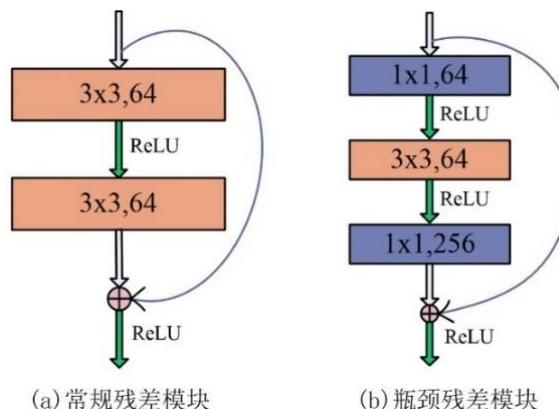


图 2-12 常规残差模块(building block)、瓶颈残差模块(bottleneck building block)

ResNet 网络就是以这两种残差模块为主要构成单位, 加上池化层和全连接层组成的, 具体网络结构如表 2-2 所示:

表 2-2 ResNet 网络结构

Feature map	ResNet18	ResNet34	ResNet50	ResNet101	ResNet152
112x112			Kernel=7, number=64, stride=2, padding=0		
56x56	$\{3 \times 3, 64\} \times 2$	$\{3 \times 3, 64\} \times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{cases} \times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{cases} \times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{cases} \times 3$
28x28	$\{3 \times 3, 128\}$ $\times 2$	$\{3 \times 3, 128\}$ $\times 4$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 512 \end{cases} \times 4$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 512 \end{cases} \times 4$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 512 \end{cases} \times 8$
14x14	$\{3 \times 3, 256\}$ $\times 2$	$\{3 \times 3, 256\}$ $\times 6$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 1024 \end{cases} \times 6$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 1024 \end{cases} \times 23$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 1024 \end{cases} \times 36$
7x7	$\{3 \times 3, 512\}$ $\times 2$	$\{3 \times 3, 512\}$ $\times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 2048 \end{cases} \times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 2048 \end{cases} \times 3$	$\begin{cases} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 2048 \end{cases} \times 3$
1x1x1000			平均池化、全连接层、softmax		

其中 ResNet 层数并不固定, 如 ResNet18、ResNet34、ResNet50、ResNet101、ResNet152, 其中的数字代表的是全部卷积层和全连接层的层数, 后续很多深度学习模型都把残差网络作为基础骨架使用。

2.3 深度学习常用的评价指标介绍

如何评价深度学习模型效果的好坏, 这就需要引入定量指标来进行度量。本节我们主要介绍几种在语义分割和目标检测中常用的几种评价指标。在给出评价指标之前, 需要先介绍几个基本的量化指标, 这些指标是基于二分类(正负样本)的任务, 也可以应用于多分类问题:

TP(True Positive): 真实类别为正样本, 预测类别也为正样本的数目;

FN(Flase Negative): 真实类别为正样本, 预测类别为负样本的数目;

FP(Flase Positive): 真实类别为负样本, 预测类别为正样本的数目;

TN(True Negative): 真实类别为负样本, 预测类别也为负样本的数目。

(1) 准确率(Accuracy)

准确率代表所有预测正确的样本占所有样本的比例，是评估深度学习模型全局准确程度的最直观指标之一，但由于其包含的信息较少，难以有效而全面地评价模型的性能，其计算公式见式(2-17):

$$Accuracy = \frac{TP+TN}{TP+FN+FP+FN}. \quad (2-17)$$

(2) 精确率(Precision)

精确度代表预测结果与真实结果都为正类的部分占所有预测结果为正类的比例，其计算公式见式(2-18):

$$Precision = \frac{TP}{TP+FP}. \quad (2-18)$$

(3) 召回率(Recall)

召回率代表预测结果与真实结果都为正类的部分占所有真实结果为正类的比例，其计算公式见式(2-19):

$$Recall = \frac{TP}{TP+FN}. \quad (2-19)$$

(4) F分数

F_1 分数可以同时衡量召回率和精确率，是影像领域最常用的评价指标之一。但在实际情况中，我们可能会根据实际问题有侧重地考虑召回率和准确率。由于 F_1 分数中召回率和准确率的作用同等重要，在需要考虑两者差异性的场合，难以区分度量两个指标的不同。为此，通过在 F 分数中设定权重 β ，可以对精确率和召回率赋以不同的重视程度，当精确率更重要时， $\beta < 1$ ；当召回率更重要时， $\beta > 1$ ；当同等重要时， $\beta = 1$ ，即为 F_1 分数。 F_1 分数和 F_β 分数计算公式见式(2-20)、式(2-21):

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall}, \quad (2-20)$$

$$F_\beta = (1 + \beta^2) \times \frac{Precision \times Recall}{\beta^2 \times Precision + Recall}. \quad (2-21)$$

2.4 本章小结

本章主要介绍了深度学习方法的发展历史和基本概念，为实现深度学习在月球撞击坑的应用提供理论基础，最后介绍了深度学习最常用的量化指标，并选取其中具有代表性的指标作为检验网络效果的定量指标。

第3章 月球撞击坑数据介绍

3.1 研究数据说明

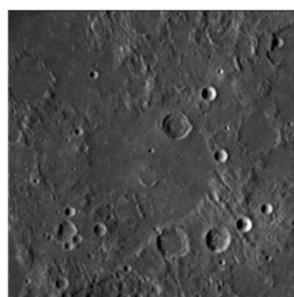
在近几十年以来，各国纷纷开展深空探测任务，发射了许多探测卫星和探测器，通过携带的各种先进设备对其进行勘测，而最直观展现行星面貌的数据就是光学影像，即通过 CCD 相机等近距离拍摄的星球表面影像。随着传感器技术的进步，获取影像的分辨率越来越高，星球表面地貌的细节信息、几何纹理等信息得以彰显；而立体相机、高度测量仪器，射线测量仪等其他载荷设备能够更全面监测行星的地形特征、物质结构等信息。在月球撞击坑的识别中，应用较为广泛的数据有 LROC 的光学影像和以 LOLA 为基础制作的数字高程模型。这些数据通常是经过处理后生成的全月球数据，根据需求的不同可以选择适合分辨率的产品。其中光学影像和数字高程模型如图 3-1 所示。

(1) 光学数据

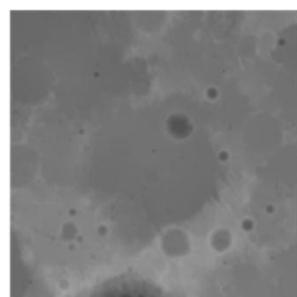
月球侦查轨道器照相机(Lunar Reconnaissance Orbiter Camera, LROC)是该轨道飞行器上携带的照相机，包括一个广角照相机(WAC)和一个窄角照相机(NAC)，通过这两个照相机可以完成对月球表面的多方位拍摄任务，提供了较高空间分辨率的月球光学影像。

(2) 数字高程模型

月球轨道飞行器激光测高仪(Lunar Orbiter Laser Altimeter, LOLA)是对整个月球的地形进行测量，通过该数据制作的月球高程数字模型 DEM 保留了月球表面大量的地形信息。



月球光学影像



月球数字高程模型

图 3-1 月球数据展示

3.2 撞击坑目录介绍

对于深度学习方法来说，标签数据，亦即训练数据，起着至关重要的作用。模型训练的好坏很大程度上取决于标签数据的数量及其准确程度，而由于人类对于撞击坑的认知只来源于经验，对于大部分撞击坑是无法准确判别出来，所以有关撞击坑的标签数据是非常稀少的。但是在天文学家的努力下，能够通过观测所获得的数据，利用统计方法来对特征相对明显的撞击坑进行识别，这就初步形成了较为可靠的撞击坑目录。这些目录都包含了撞击坑的位置和直径信息，大部分研究人员都是基于撞击坑公开目录作为自己研究方法的训练数据。

在月球撞击坑目录中，Head 等^[46]采用的方法较为保守，其识别的撞击坑大部分为直径较大且特征比较明显的，识别撞击坑的数目较少但可信度较高，但是目录中没有小直径的撞击坑，这对于深度学习的应用极为不利。于是在他们工作的基础上，Povilaitis 等^[47]主要关注具有较小直径的撞击坑的特征，通过对其进行分析识别出了一定数量的小型撞击坑，填补了 Head 等人未涉及到的撞击坑，完善了目录。因为上述两个目录之间存在的关系，研究人员通常把这两个目录合在一起共同作为标记数据集。但是在这两种月球撞击坑目录之外，还存在相当一部分撞击坑还未被标识出来^[48]。为了满足深度学习的需求，Robbins 等^[49]提出了一个包含 200 万月球撞击坑的目录，虽然其准确性相较其他目录有所下降，但是完整程度相比于其他目录有了很大提高，在面对样本量少的情况下，该目录是一个很好的选择。国际天文学联合会(International Astronomical Union)也公布了目前已被探测到的月球撞击坑目录。

当然，不同目录之间由于识别方法不同，识别结果也存在很大差异，因此在研究过程中应该根据任务和目标来确定合适的撞击坑目录，表 3-1 列出了在月球撞击坑识别领域中深度学习方法使用较多的撞击坑目录：

表 3-1 月球撞击坑目录

目录	直径范围(km)	撞击坑个数	选用数据	发表时间
Head	大于 20km	5185	DTM	2010
Povilaitis	5-20km	22746	WAC DTM	2018

3.3 数据集准备

由于光学数据受光照、阴影、噪声的影响较大，直接应用在深度学习模型中的效果并不佳，而数字高程数据受这些因素的影响小，能够较好满足实验要求，在当前研究中被广泛使用，因此本论文主要选用月球数字高程模型作为实验数据，以高分辨率的光学图像作为可视化展示。

本章的实验数据为美国 NASA 实验室提供的月球数字高程模型 SLDEM2015 数据，分辨率选取 118m (256 像素/度)，全图大小为 92160×30720 ，该数据是由 Barker 等^[50]根据月球勘测轨道器上搭载的轨道器激光测高仪(LOLA)搜集的月球表面高度测量值制作而成的，该数据的在线获取地址为：<https://pgda.gsfc.nasa.gov/products/54>，原始数据如图 3-2 所示：

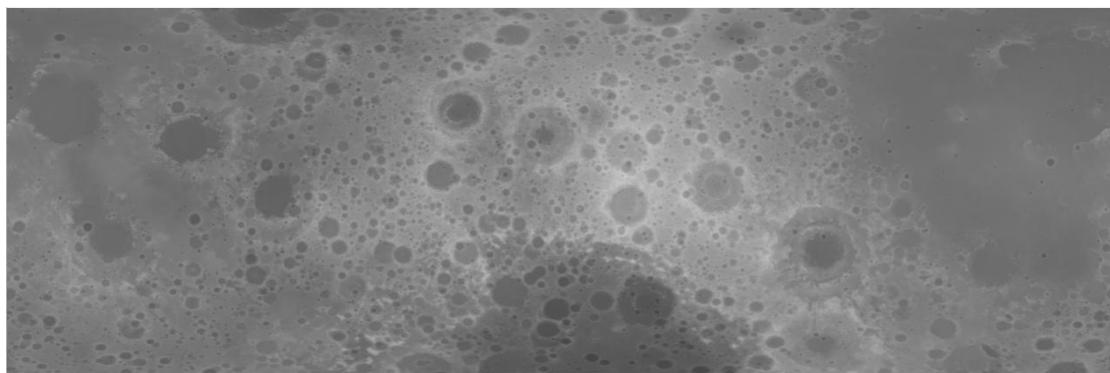


图 3-2 全月数字高程影像($60^{\circ}\text{S}-60^{\circ}\text{N}$, $0^{\circ}\text{E}-360^{\circ}\text{E}$)

3.3.1 语义分割任务数据集 Dataset1

在获得月球 DEM 数据后，由于数据大小、影像格式等各种因素的限制，不能直接使用原始数据进行网络训练，需要对其进行前期的预处理，将数据转化为满足实验要求的数据集格式。并将该数据集命名为 Dataset1，具体步骤如下：

(1) 影像格式转化

数字高程模型的数据格式为 JPEG2000，字节位深为 16bit/pixel，为了方便训练，在实验前期将数据格式统一转化为 PNG、24bit/pixel。

(2) 制作标签数据

在文献^[17]中，根据 Head-Povilaitis 目录提供的撞击坑经纬度和直径信息，采用宽度为 1 像素的圆环圈出每一个撞击坑的范围，并以此作为实验的 Targets。但由于语义分割需要逐像素地识别撞击坑区域，因此，在本节的实验中选择用实心圆来对撞击坑进行填充，并以此生成标签数据。此外，Head 目录中存在直径特大的撞击坑，在数据标注时存在大型撞击坑会覆盖掉小型撞击坑的情况，不利于网络进行训练。为更有效地利用深度学习方法，实验只对直径较小的撞击坑进行识别，即单独使用 Povilaitis 目录进行标注，需要注意的是，Povilaitis 目录提供的撞击坑的经度信息是[-180°E,180°E]。在训练过程中，标签数据通常采用 mask 格式，将背景设置为 0，撞击坑所在的像素值设为 1，将撞击坑识别作为二分类问题进行训练。

(3) 裁剪数据

因为训练数据的大小对于实验设备具有一定要求，过大的影像尺寸无法进行训练，而过小的尺寸会导致撞击坑无法完整出现在训练影像中，出现信息缺失的问题，制约网络模型的准确性。而撞击坑的直径像素范围为 42-169，在综合考虑这两种情况下，本节将数字高程模型的训练数据大小设置为 512×512，按顺序进行裁剪，这样既能保证实验的顺利进行，同时由于撞击坑的像素直径都小于分割图的 $\frac{1}{3}$ ，也可以保证大部分撞击坑都能够较为完整的出现在裁剪影像中，避免出现主要信息丢失问题。基于上述过程，数字高程模型集共分别得到了 10800 张影像，其中以 0°E-120°E 区域的数据作为验证集，120°E-240°E 区域的数据作为训练集，240°E-360°E 区域的数据作为测试集。数据集标签可视化结果如图 3-3、图 3-4 所示：

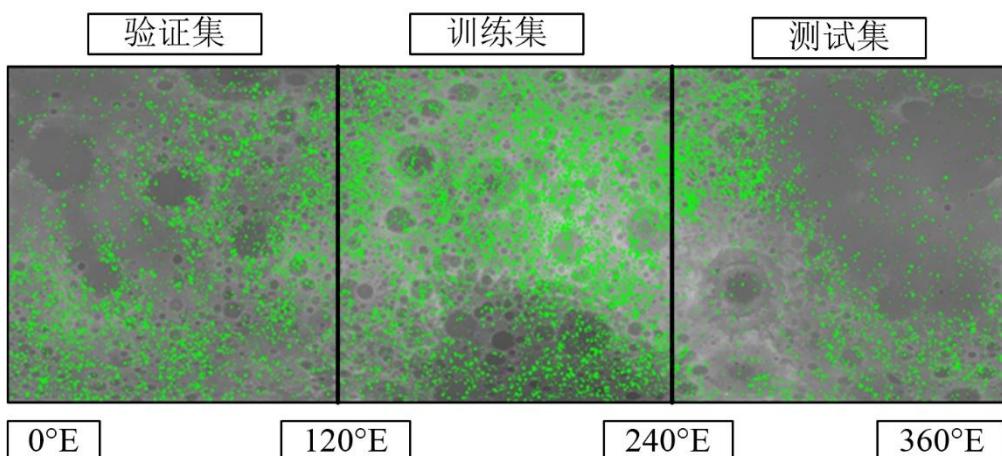


图 3-3 全月撞击坑分布图(Povilaitis 目录撞击坑分布)

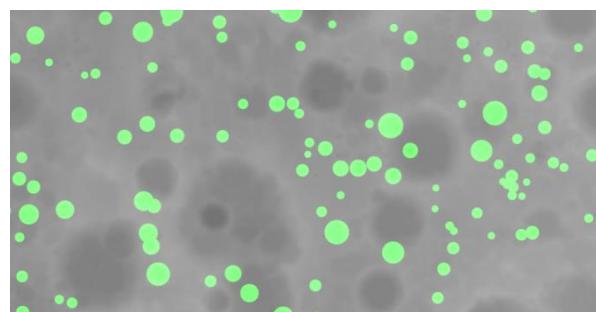


图3-4 局部区域展示(绿色实心圆为Povilaitis目录中记录的撞击坑)

3.3.2 目标检测任务数据集 Dataset2

相较于语义分割数据集，目标检测数据集格式略有不同，本节通过使用 Python 代码将 Dataset1 中的标签数据由掩膜图片格式转化为带有标签信息的 Json 文件，并按照 COCO 数据集的格式对数据集进行重新整理，将制作的用于目标检测模型的数据集命名为 Dataset2。

3.4 本章小结

本章首先介绍了当前月球撞击坑识别领域使用数据的来源，提供了可靠的月球撞击坑专家识别目录，并选取了数字高程模型作为本论文的实验数据，最后使用 Python 软件对专家目录中存在的撞击坑进行自动标注。同时根据语义分割任务和目标检测任务的需求，制作了两个撞击坑数据集，并对数据的格式进行相应调整，以满足深度学习模型训练的需要。

第 4 章 基于改进 UNet 网络的月球撞击坑识别

本章以月球撞击坑为识别对象，使用语义分割模型 UNet 网络对其进行自动检测，任务目标是将撞击坑与背景分割出来，并在 UNet 网络的基础上做出改进，使其更适合对撞击坑数据集进行识别。主要从增强特征提取能力和提高训练效率两方面出发，分别引入 ResNet 和 MobileNet 架构，构建了 ResUNet 网络和 MobileUNet 网络；同时设计了密集连接的空洞空间卷积金字塔池化模块，增强 ResUNet 网络对多尺度特征的捕获能力；最后提出了一个后处理模块，通过对上述网络进行集成，进一步提高对月球撞击坑的预测效果。本章主要内容包括经典语义分割模型介绍、实验设计以及预测结果的评价等。

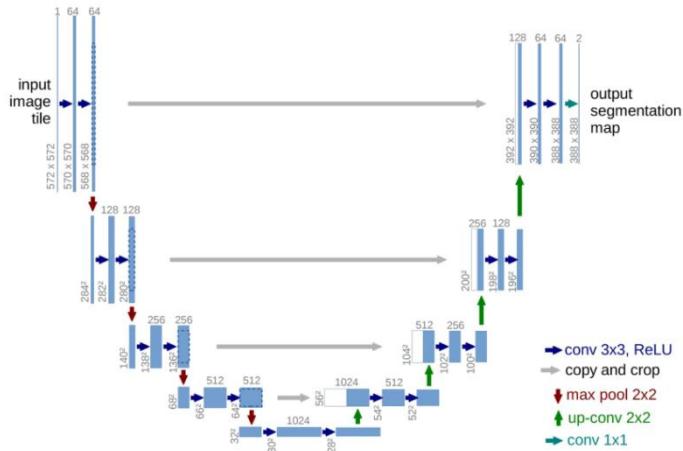
4.1 经典语义分割模型介绍

(1) 全卷积神经网络

首次将卷积神经网络应用到图像语义分割领域的是 Jonathan 等^[51]提出的全卷积神经网络(Fully Convolutional Networks, FCN)，FCN 是在 AlexNet、VGGNet、GoogleNet 的基础上使用反卷积层对卷积层的特征图进行上采样，使输出图像的尺寸恢复到输入图像的大小，最后在上采样的特征图上将最后的全连接层改为卷积层并进行逐像素分类，从而达到对每一个像素的预测，实现了对图像端到端的预测。

(2) UNet 网络

FCN 网络是处理图像语义分割的开山之作，但是在分割的预测结果上不够精确，对图片的细节划分比较模糊，而 Ronneberger 等^[52]提出的 UNet 网络在 FCN 的基础上对上采样过程中的跳跃连接层做出了改进，相比于 FCN 网络跳跃连接的方式，UNet 网络是通过保存下采样的特征图，并以拼接的方式将其与上采样后的特征图融合，进而保留更多图像信息，在图像分割的细节方面取得了更好效果，弥补了 FCN 的不足，其网络结构如图 4-1 所示：

图 4-1 UNet 网络架构示意图^[51]

UNet 网络是一个经典的编码器解码器结构，左边编码器部分是传统的卷积网络，由四个重复结构组成，每个结构包含两个 3×3 的无填充卷积层和一个步长为 2 的最大池化层，在下采样过程中，通过控制卷积核的个数使特征图的通道数增加；底层部分是两次卷积操作；右边解码器部分是上采样过程，通过线性插值的方式放大特征图，恢复在下采样过程中丢失的图像细节信息，然后与下采样得到的特征图拼接起来，需要注意的是，由于在下采样中需要拼接的特征图尺寸比上采样恢复的特征图大，故在拼接前需要将下采样得到的特征图进行裁剪，再通过跳跃连接层进行拼接，拼接后的特征图再进行 2 次 3×3 的卷积操作，依次进行四次；其中每次卷积操作之后都使用 ReLU 激活函数进行激活；最后通过 1×1 的卷积操作将 64 通道的特征图转化为特定类别数量的预测结果进行输出。

(3) MobileNet 网络

自残差网络解决深层卷积网络退化问题以来，深度学习模型都趋向于使用深层卷积网络，但随着网络层数的加深，模型的参数量随之增大，计算消耗更高，为减小模型参数量，Andrew 等^[53]提出了轻量化网络 MobileNet 模型，通过设计深度可分离卷积 (Depthwise Separable Convolution, DSC)，大幅减少卷积过程造成的计算量，同时以深度可分离卷积为基础构建的逆转残差模块(Inverted Residual Block)，能够在保证精度的前提下，显著提升模型的运算速度，而且在卷积过程中引入 SeNet 网络^[54]中的通道注意力机制(Squeeze-and-Excitation, SE)，增强网络对全局特征的学习能力。

深度可分离卷积操作减少了卷积所需的参数量，同时能够保持与标准卷积相近的精度，主要过程包含两个独立运算的卷积操作：深层卷积(Depthwise Convolution)和点卷积(Pointwise Convolution)。深度卷积是对输入特征图的每个通道分别使用一个卷积核遍历，然后将结果拼接到一起输出，通道之间保持相互独立；点卷积实质上是 1×1 卷积，主要目的是为了将输入特征图在深度方向上加权组合，以实现不同通道之间的融合，具体操作如图4-2所示：

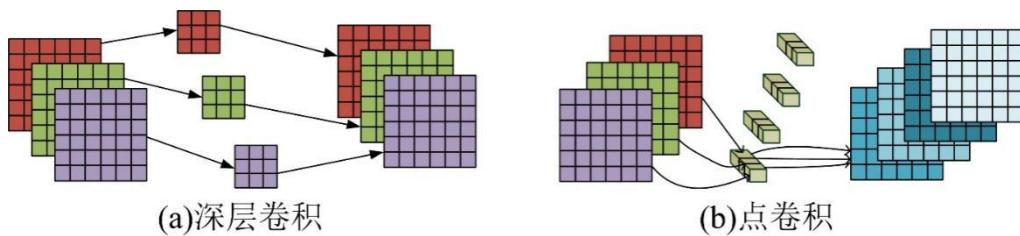


图4-2 深层卷积(左)、点卷积(右)运算示意图

常规逆转残差模块主要包含以下步骤：先使用 1×1 的点卷积增加通道数，再使用 3×3 的深层卷积，最后使用 1×1 的卷积将通道数降为原始维度，为避免网络退化，在第二次点卷积之后，将特征图与输入特征图数值相加；带有SE机制的逆转残差模块通过对深层卷积做全局平均池化和全连接操作，使网络在卷积过程中更关注全局信息。具体操作如图4-3和图4-4所示：

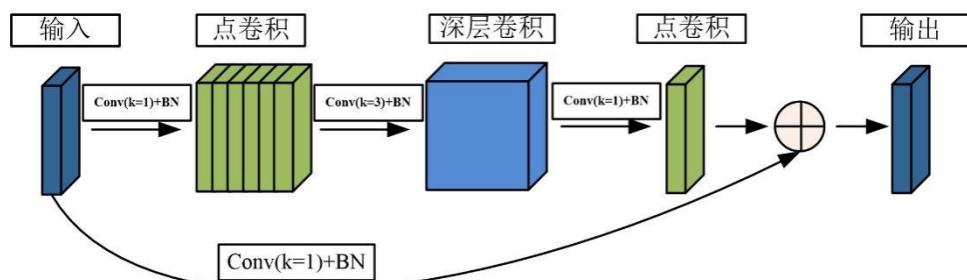


图4-3 常规逆转残差结构示意图

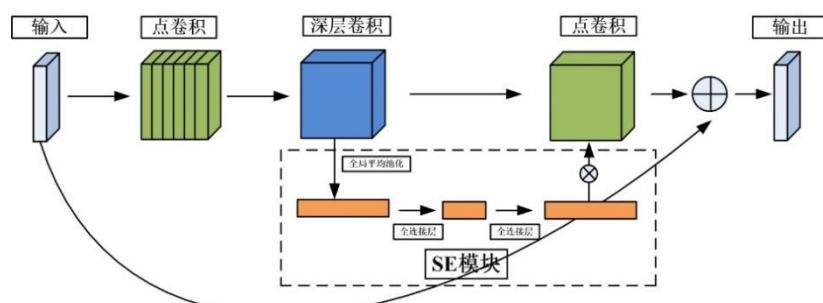


图4-4 带有SE模块的逆转残差模块

4.2 实验设计

针对撞击坑的特征提取难题，本论文对原始 UNet 网络的编码层部分做了调整，为提高网络特征提取能力和模型运算效率，设计了三个基于 UNet 架构的变种网络。

一、分别使用不同层数的 ResNet 网络替代原始 UNet 网络的编码层，对输入图像进行特征提取，通过增加卷积层数提高网络的特征提取能力，同时利用残差学习思想，在卷积计算过程中引入残差模块，通过连接卷积层的输入和输出，避免模型随着网络深度的增加出现的退化问题，以 ResUNet-34、ResUNet-50、ResUNet-101 作为该网络的简称。

二、空洞空间卷积金字塔池化 (Atrous Spatial Pyramid Pooling, ASPP) 是由 Liang-Chieh Chen 等^[55]提出的，借助空洞卷积扩大卷积核感受野的特点，可以将不同膨胀率的空洞卷积结合到一起获取目标多尺度特征信息，本论文引入该模块以增加卷积核的感受野，使网络能够对尺寸较大的撞击坑的特征做充分提取，同时为避免过大膨胀率出现的无效卷积操作，借鉴 Gao Huang 等^[56]提出的 DenseNet 网络中密集连接的思想，将每层空洞卷积组合起来，设计了 DenseASPP 模块，提高该模块的特征提取能力，最终将该模块应用于 ResUNet 网络的底层连接部分，将该网络命名为 DARUNet 网络。

三、MobileNet 网络依靠深度可分离卷积大量减少了网络中因卷积运算带来的计算量。因此本论文将 MobileNet 架构作为 UNet 网络的编码层实现特征提取功能，并通过注意力机制和深度可分离卷积的共同作用，在不大幅增加网络运算量的同时提高网络对撞击坑的关注程度，使模型在特征提取时更关注撞击坑的本质特征，将该网络命名为 MobileUNet 网络。

4.2.1 ResUNet

原始的 UNet 网络每一层的卷积采用的是两个 3×3 卷积操作，而后将特征图与上采样之后的特征图直接叠加，随着网络层数的加深，这种结构可能会出现网络退化问题，而 ResUNet 网络的残差结构能够有效解决深层网络出现的退化问题，通过在 UNet 网络的卷积模块中引入残差结构，提高网络的整体稳定性，网络架构如图 4-5 所示：

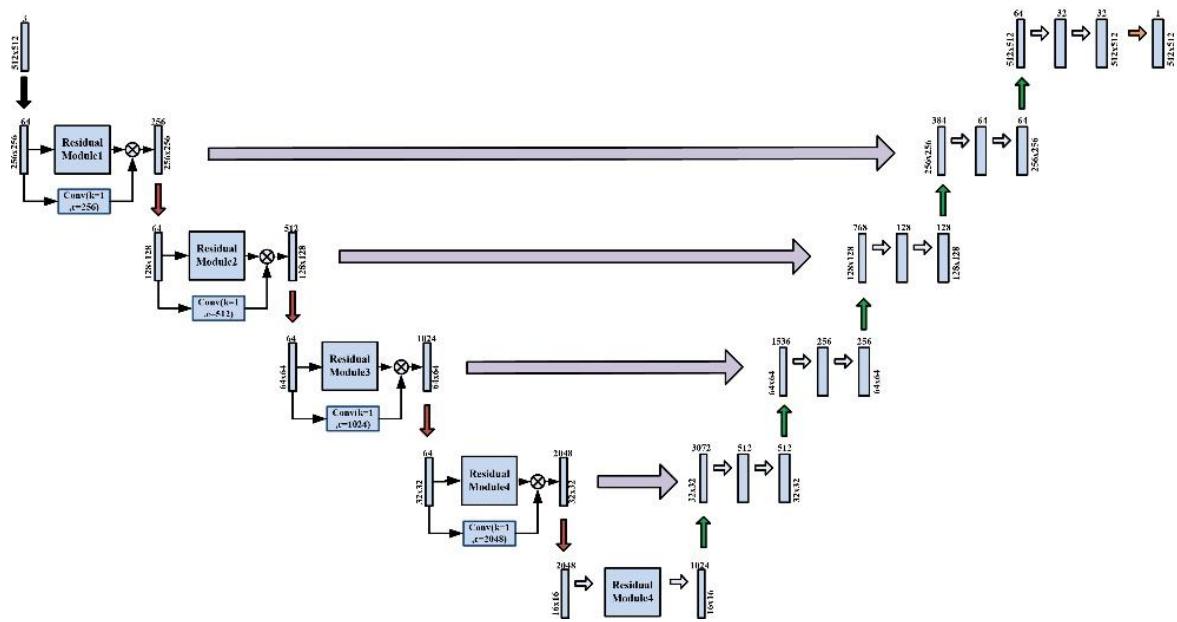


图 4-5 ResUNet 网络架构示意图

本节的 ResUNet 网络采用编码器解码器结构，网络深度为五层，左侧为编码器，右侧为解码器，以 ResUNet-101 网络为例，详细介绍该网络的具体架构。在编码层使用 ResNet101 结构，除第一层外每一层都进行三次卷积模块操作，第一个卷积模块的卷积核大小为 1×1 ，步长为 1，卷积核个数为 64；第二个卷积模块卷积核大小为 3×3 ，步长为 1，填充为 1，卷积核个数为 64，第三个卷积模块的卷积核大小为 1×1 ，步长为 1，卷积核逐层增加，激活函数为 ReLU 函数，同时为了加快网络收敛速度，在激活函数前加入批归一化方法。在编码器部分的最后一次卷积模块中卷积核个数分别设置为 256、512、1024、2048，解码器部分的个数设置为 512、256、128、64，其中最底部的卷积核个数为 1024。

与原始 UNet 网络不同之处在于，每一层使用 1×1 的卷积将输入特征图与经过三次卷积后的特征图结果相加，避免网络退化问题。在下采样过程中，使用步长为 2 的卷积操作来缩小特征图大小，减少网络参数量。在上采样阶段，使用双线性插值的方法恢复特征图大小，并与同一层编码器得到的特征图进行拼接，使网络能够保留原图更多特征信息。同时，使用二分类交叉熵损失函数计算损失，通过反向传播更新参数，最后使用 Sigmoid 函数将预测结果进行输出。

4.2.2 DARUNet

本节在 ASPP 的基础上进行改进, 以密集连接的方式设计了 DenseASPP 模块。首先使用空洞卷积对输入特征图提取特征, 然后将卷积后的特征图与输入特征图叠加, 作为下一部分的输入, 该步骤重复三次, 空洞卷积的膨胀率依次为 6、12、18, 最后使用 1×1 的卷积将通道数恢复为输入通道数, 通过密集连接操作, 可以对特征信息作进一步整合。相比于原始的 ASPP 模块, DenseASPP 在运算过程中减少了参数量, 并增强了对特征的提取能力, 具体结构如图 4-6 所示:

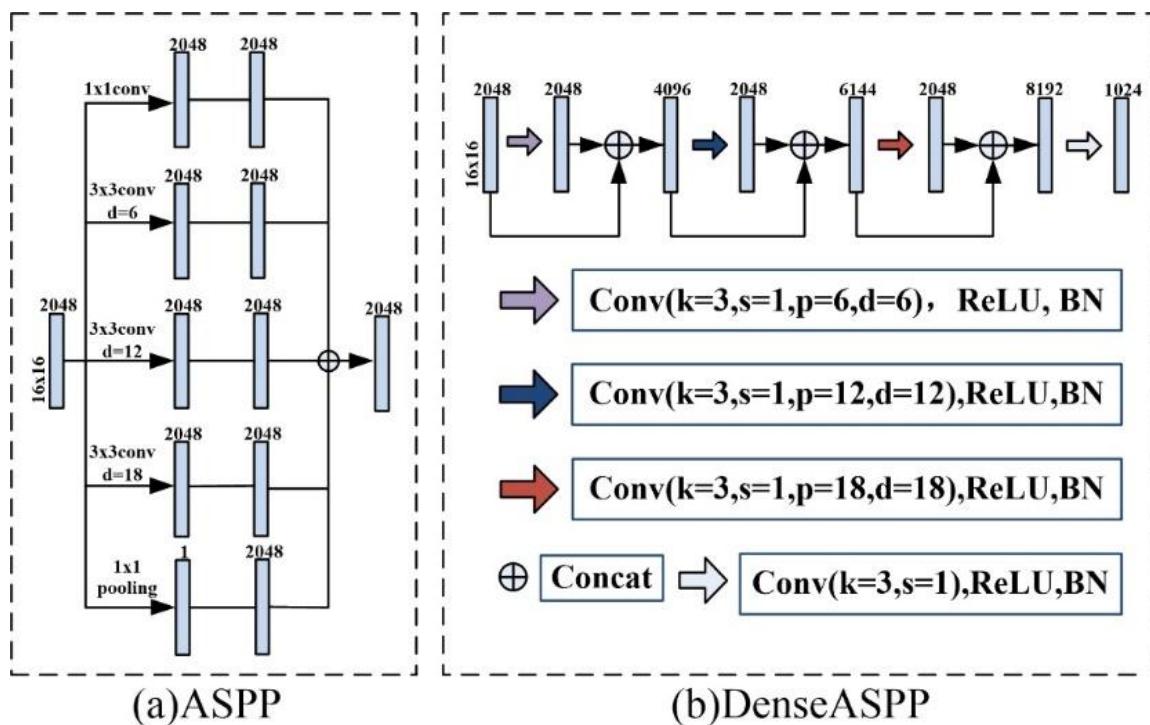


图 4-6 ASPP(左)、DenseASPP(右)示意图

本节的 DARUNet 网络的编码器结构与 4.2.1 节一致, 在编码器和解码器部分使用 ResUNet 网络进行特征提取, 在最底层的连接部分使用 DenseASPP 模块替代原有的普通卷积模块, 通过该模块将下采样的特征图进行再整合, 进一步增强该网络的分割效果, 网络架构如图 4-7 所示:

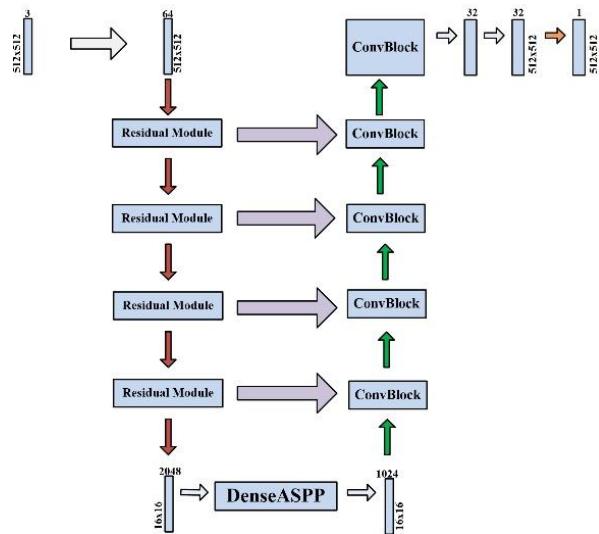


图 4-7 DARUNet 网络架构示意图

4.2.3 MobileUNet

本节的 MobileUNet 网络依然采用编码器解码器结构，共分为五层，主要改进之处在于使用深度可分离卷积代替常规卷积，以减少网络的参数，达到简化网络运算的效果，网络架构如图 4-8 所示：

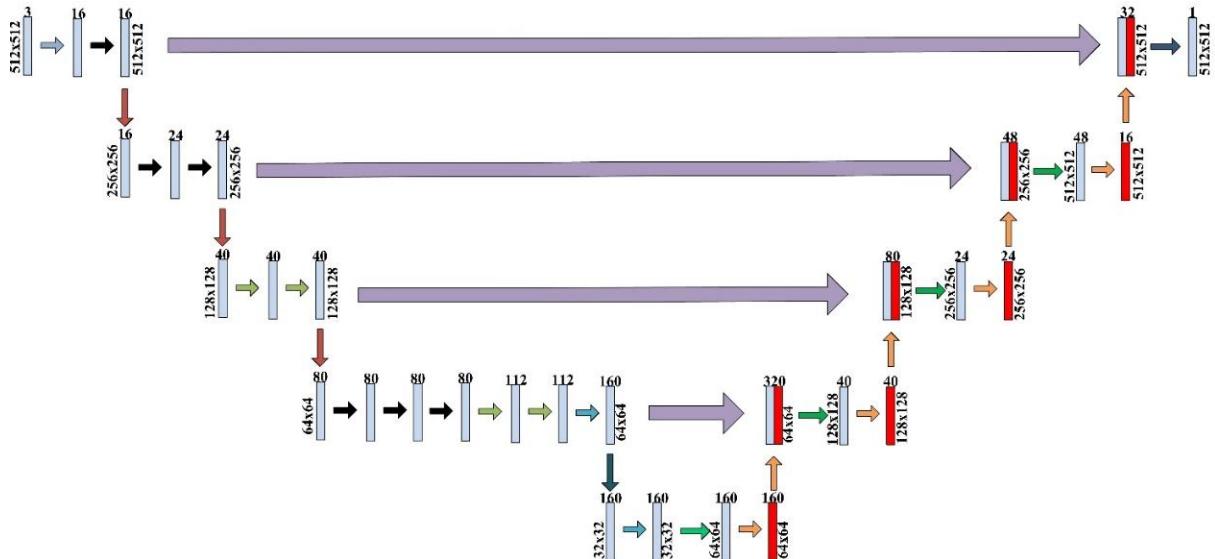


图 4-8 MobileUNet 网络架构示意图

右侧的编码器结构采用 MobileNetv3 网络的基础模块，前两层使用卷积核大小为 3，步长为 1 的常规逆转残差模块；第三层使用两个卷积核大小为 5，带有 SE 模块的逆转残差模块增强网络的特征提取能力；第四层先使用三个常规逆转残差模块，再使用三个带有 SE 模块的逆转残差模块，其中第六个逆转残差模块的卷积核大小为 5；第五层使

用一个带有 SE 模块的逆转残差模块，其中卷积核大小为 5。前三层的激活函数为 ReLU 函数，后两层的激活函数为 Hswish 函数，在激活前使用批归一化操作避免网络退化。下采样操作是通过步长为 2 的卷积操作使特征图的大小变为原来的一半。

左侧的解码器结构是通过双线性插值的方法使特征图恢复到原来尺寸的 2 倍，并在上采样之后使用一个普通卷积模块将其通道改变至与相同编码层的输出特征图通道数一致，然后与编码层的特征图进行拼接作为下一层的输入。最后使用一个 1×1 的卷积操作将特征图变为单通道的特征图，并使用 Sigmoid 函数进行分类，该网络同样使用二分类交叉熵损失函数计算损失。

其中编码层卷积模块的参数具体设置见表 4-1：

表 4-1 MobileUNet 网络编码层的卷积模块

	特征图尺寸	输入通道	中间通道	输出通道	卷积核尺寸	步长	SE	激活函数
Layer1	512	3	16	16	3	1	无	ReLU
	512	16	16	16	3	1	无	ReLU
Layer2	256	16	64	24	3	2	无	ReLU
	256	24	72	24	3	1	无	ReLU
Layer3	128	24	72	40	5	2	无	ReLU
	128	40	120	40	5	1	有	ReLU
Layer4	128	40	120	40	5	1	有	ReLU
	64	40	240	80	3	2	无	Hswish
Layer5	64	80	200	80	3	1	无	Hswish
	64	80	184	80	3	1	无	Hswish
Layer4	64	80	184	80	3	1	无	Hswish
	64	80	480	112	3	1	有	Hswish
Layer5	64	112	672	112	3	1	有	Hswish
	64	112	672	160	5	1	有	Hswish
Layer5	32	160	672	160	5	2	有	Hswish
	32	160	960	160	5	1	有	Hswish

4.2.4 后处理模块

本实验的后处理模块分为两部分，第一部分将三个模型在验证集的预测结果进行融合，提高预测结果的准确率；第二部分是运用 Python 代码将预测结果进行过滤，筛选符合标准的撞击坑作为候选结果，并对候选撞击坑区域以最小外接圆进行覆盖，提高预测结果的召回率；最后将分割结果进行输出，并根据图像位置复原撞击坑的经纬度和半径信息。

(1) 模型融合

模型融合的思想来源于机器学习^[57]，简单来说就是训练多个模型，使用一定策略将其组合起来，使集成后的模型具有更好的预测性能。本文采取投票法对三个模型进行融合，由于三个模型中 ResUNet 网络和 DARUNet 模型架构具有高度重合性，分割结果比较类似，而 MobileUNet 网络的特征提取结构与前两个模型具有一定差异，在对不同类型的撞击坑的识别上有不同的表现，因此在融合过程中，对前两个模型设置较小的权重，第三个模型设置较大的权重，即前两个基于 ResUNet 的模型中在验证集得分低的模型赋权重为 $\frac{1}{6}$ ，高的模型权重为 $\frac{2}{6}$ ，MobileUNet 的模型权重为 $\frac{3}{6}$ 。预测结果中撞击坑像素值设置为 255，背景类设为 0，当三个模型在测试集的预测结果中每一个像素点经加权计算后，若结果大于或等于 $\frac{255}{2}$ ，则认为该点像素值属于撞击坑，否则视为背景。

(2) 撞击坑轮廓处理

由于模型输出的预测结果为撞击坑与背景的分割图，撞击坑的分割结果并不准确，有些撞击坑区域属于漏判、错判，还存在一些撞击坑的分割结果并不完整，为了规范分割结果，本实验设计了撞击坑轮廓处理程序，对模型的预测结果进行清洗，程序处理步骤如下：

一、恢复全月图片：由于测试集是对分割后的月球图片进行预测，在分割图片的边缘可能存在着不完整的撞击坑，有些能被检测出来，有些会被遗漏，因此需要将分割图按照裁剪顺序恢复为完整的月球图片，方便后续的筛选和补充，测试集恢复后的图片分辨率为 30720×30720 ；

二、提取轮廓信息：将预测结果中所有预测为撞击坑的区域筛选出来，根据其轮廓找到该区域的最小外接圆的中心位置和半径，并输出每一个撞击坑的圆心坐标以及半径；

三、确定候选撞击坑：由于本数据集预测的为直径在 5km 到 20km 的撞击坑，该部分标记撞击坑的像素半径在 20 至 85 之间，为保证预测准确性，删除预测结果中像素半径小于 10 的撞击坑，并将像素半径超过 100 的撞击坑半径上限设为 100，筛选过后的区域即为候选撞击坑；

四、轮廓内部补充：根据轮廓最小外接圆的中心位置和半径对候选撞击坑区域画圆，并进行内部填充。

(3) 建立月球撞击坑目录

根据上述步骤获得的月球撞击坑的预测结果，分割图坐标与撞击坑真实经纬度和直径转化公式见式 4-1、式 4-2、式 4-3，通过转换公式可以恢复撞击坑真实经纬度和半径信息，其中经纬度范围为 [60°N, -60°N]、[0°E, 360°E]，半径单位为 km。

$$Lon = \frac{X}{92160} * (360), \quad (4-1)$$

$$Lat = \frac{Y}{30720} * (-120) + 60, \quad (4-2)$$

$$Radius = radius * 0.11845, \quad (4-3)$$

其中， X 、 Y 为撞击坑的像素位置， $radius$ 为识别撞击坑的像素半径。

4.2.5 实验设备介绍及参数设置

本实验的代码是基于 PyTorch 编写的，并使用 CUDA 进行加速计算，由于数据量比较大，对实验设备有一定要求，表 4-2 是本次实验的实验设备、环境配备要求：

表 4-2 实验设备、环境配置要求

实验设备	GPU	一块 NVIDIA RTX 3090 显卡(显存 24G)
	CPU	Intel(R) Xeon(R) Silver 4216(运行内存 32G)
环境配置	Cuda	11.1
	Cudnn	8.0.5
	Python	3.8
	PyTorch	1.8.0

由于深度学习方法中不同参数的设置往往会造成较大的影响，不同模型对于参数的需求也各不相同，比如不同的优化器会造成结果出现不同程度的偏差，而且在模型训练中存在一些需要手动调整的参数，这些参数无法使用常规的优化手段直接优化，通常凭借人们的过往经验进行调整，比如批处理量(Batch Size)、学习率(Learning rate)、训练轮次(Epoch)等，这些参数的选择也会不同程度影响模型的效果，本实验通过对改进模型进行多次实验，最终以验证集的最优测试结果作为各个模型选取参数的依据，表 4-3 是模型最终确定的参数：

表 4-3 实验参数设置

Model	Loss Function	Optimizer	Learning rate	Batch Size	Epoch
ResUNet	BCEWithLogitsLoss	SGD	0.0025	4	50
DARUNet	BCEWithLogitsLoss	SGD	0.0025	8	50
MobileUNet	BCEWithLogitsLoss	RMSprop	0.00001	8	50

4.3 实验结果分析

4.3.1 定量分析

本实验使用的数据集为 Dataset1，为了对比不同网络预测的月球撞击坑分割精度，选用召回率、精确度、 F_1 分数和 F_2 分数作为实验的定量评价指标，其中 TP 为正确识别的撞击坑像元数、 FP 为错误识别为撞击坑的像元数、 FN 为遗漏的撞击坑像元数，同时选用模型参数量(Parameters)、模型复杂度(Flops)和图片处理速度(Speed)作为检验网络性能的指标，选取的对比实验为 Silburt^[17]等人在 2019 年基于 UNet 网络设计的 DeepMoon 方法在月球撞击坑的识别结果。实验结果如表 4-4 和表 4-5 所示：

表 4.4 实验分割结果展示

Model	Val(±1.0%)			Test(±1.0%)			Mean		
	Precision	Recall	F_1	Precision	Recall	F_1	Precision	Recall	F_1
DeepMoon	81.00	56.00	66.22	80.00	57.00	66.57	80.50	56.50	66.40
ResUNet-34	63.54	50.95	56.55	70.33	56.55	62.69	66.94	53.75	59.62
ResUNet-50	82.48	54.67	65.76	80.79	60.26	69.03	81.64	57.47	67.40
ResUNet-101	77.85	59.16	67.23	74.60	65.25	69.61	76.23	62.21	68.42
DARUNet-50	79.38	58.48	67.35	78.10	63.60	70.11	78.74	61.04	68.73
DARUNet-101	79.05	57.62	66.65	75.96	64.23	69.61	77.51	60.93	68.13
MobileUNet	75.05	60.03	66.71	75.48	65.58	70.18	75.27	62.81	68.45

表 4-5 模型参数展示

Model	Parameters	Flops(GFlops)	Speed(img/s)
ResUNet-50	73793699	166.47	13.6
DARUNet-50	280378531	219.36	12.5
MobileUNet	1910309	13.26	22.5

本论文分别以 ResNet-34、ResNet-50、ResNet-101 网络作为 ResUNet 网络的特征提取部分，从表 4-4 的结果来看，发现随着网络层数的加深，特征提取的能力越强，其中以 ResUNet-101 的精度最高， F_1 分数达到 68.42%；在 ResUNet-50 的基础上加入 DenseASPP 模块后增强了网络的学习能力， F_1 分数为 68.73%，分割精度得到提升；MobileUNet 的 F_1 分数为 68.45%，介于 ResUNet-101 与 DARUNet-50 之间。最后与经典方法 DeepMoon 对比，本实验的改进模型在各项指标上均有所突破，再次证明了构建的模型及模块在月球撞击坑识别的有效性。表 4-5 的结果显示，MobileUNet 模型的参数量少，在保证实验精度的同时模型运行更快、占用内存更小，证明了轻量化模型 MobileUNet 网络在实际应用中具有超过 ResUNet 网络的优越性能。

在撞击坑识别实验中，由于我们对模型的召回率有更高的要求，希望其能更多更有效地识别专家撞击坑目录中存在的撞击坑，同时由于初始的分割效果较差，在逐像素识别的过程中，存在一些不规则的识别结果。基于此，本文提出的后处理模块希望能对模

型的预测结果进行修正，补充分割图的完整性，使其更好地满足任务需求，为了有效判别后处理模块的真实效果，该部分的指标我们主要选取 F_β 分数($\beta = 2$)作为评价指标，在保证精度的同时使指标结果更侧重于召回率，实验结果如表4-6所示：

表4-6 后处理模块实验结果

Modell	Precision	Recall	F_2
ResUNet	80.79	60.26	63.49
ADRUNet	78.10	63.60	66.05
MobileUNet	75.48	65.58	67.35
后处理	57.39	78.61	76.23

预测结果经过后处理模块后，在损失一部分精确率下召回率得到了很大提升， F_2 分数得到了显著提高，实验结果证明了本章提出的后处理模块在基于语义分割模型的撞击坑识别中能够对分割结果进行有效补充。

4.3.2 定性分析

(1) 分割结果可视化

为更清楚地展现模型的预测结果，本节将经过后处理后的融合模型的结果进行可视化展示，同时将预测结果覆盖在高分辨率的月球光学图像上，并对撞击坑密度较大的区域和密度较小的区域分别放大，以直观展示预测结果的准确度和不足之处，实验结果如图4-9、图4-10所示：

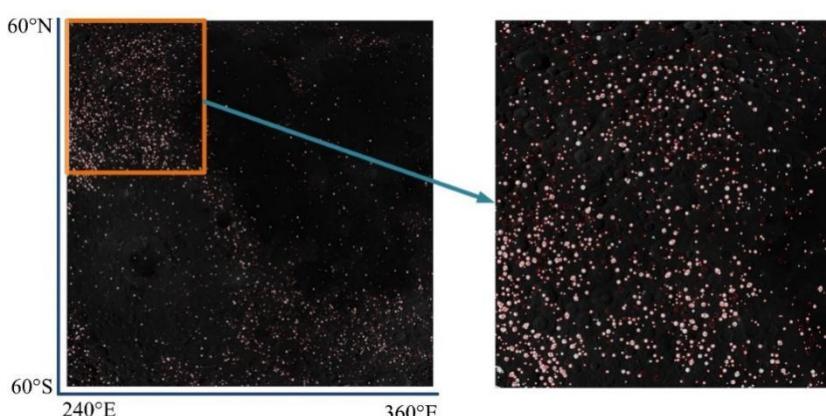


图4-9 月球撞击坑密集区域识别结果 (红色圆圈为专家标记撞击坑，白色为模型的预测结果)

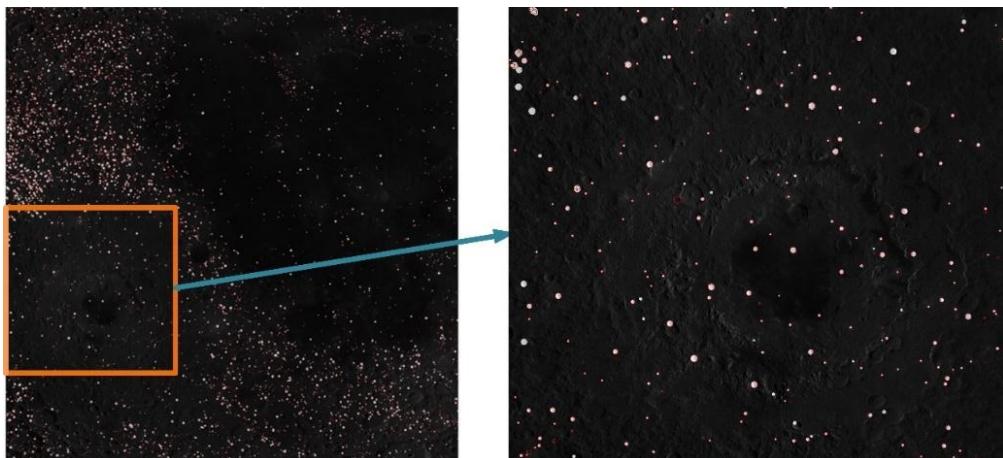


图 4-10 月球撞击坑稀疏区域识别结果(红色圆圈为专家标记撞击坑, 白色为模型的预测结果)

从分割结果中可以看出, 目录中包含的大部分撞击坑都能够被训练的模型识别出来, 识别精度可以达到初步的预期效果。但在具体的细节方面仍具有很多值得探讨的地方, 比如在撞击坑分布较为稀疏的区域, 撞击坑能够被很好的检测出来, 但是在分布较为集中的区域, 距离相近的撞击坑在训练过程中的学习效果不佳, 对于一些重叠在一起的撞击坑, 难以将其边界精准分割; 在地势复杂的区域, 尤其是较高纬度的区域, 有相当一部分撞击坑没有被识别出来, 通过图 4-9 可看到, 这部分撞击坑的周围存在着很多大型撞击坑, 推断其形成来源为陨石撞击时喷溅导致的凹坑, 撞击坑深度较浅, 在数字高程模型数据中与背景相似, 因此导致深度学习模型难以学习到其真实特征。从图 4-10 还可以看到深度学习模型识别到一些未被专家撞击坑目录记录到的撞击坑, 这部分撞击坑较为分散且独立存在, 常规的手段(比如专家识别)容易忽略。因此深度学习方法对于寻找一些人类尚未发现的撞击坑具有一定研究价值。

(2) 撞击坑识别分析

在月球[240°E-360°E,60°N-60°S]区域内, Povilaitis 目录中共存在 4056 个直径在 5km-20km 的撞击坑, 深度学习模型在测试集上的实验结果显示共识别到 8000 个直径在 0.1km-53.52km 的撞击坑, 其中直径小于 5km 的有 3489 个, 在 5km-20km 的有 4377 个, 大于 20km 的有 134 个, 在可视化结果中可以看到深度学习方法对小型撞击坑上具有较强的检测优势, 使用深度学习方法能够有效补充小型撞击坑的信息, 提高撞击坑目录的完整性。

4.4 本章小结

本章节主要从语义分割任务需求出发，通过改进经典语义分割网络，并以月球撞击坑数据集进行验证，主要的工作有：

一、在UNet网络的基础上，分别从提高网络特征提取能力和提高网络运行效率两方面对编码层结构进行改进，引入ResNet网络、密集空洞空间卷积金字塔池化模块和MobileNet网络，设计了三个模型：ResUNet、DARUNet、MobileUNet，从实验结果来看，这些模型都明显提高了在月球撞击坑识别的分割精度，且在识别效率上得到显著增强。

二、设计了一个撞击坑轮廓处理程序，先使用投票法对模型实验结果进行融合，利用圆形拟合方法对撞击坑的初始分割结果进行完善，进一步优化模型的分割结果，提高预测结果在专家撞击坑目录的召回率，最后根据分割结果输出预测撞击坑的经纬度和半径信息。

本章节的实验是对月球撞击坑进行语义分割和识别。通过使用月球数字高程数据，借助深度学习中的语义分割网络方法来自动识别并提取月球撞击坑。实验结果显示在小型撞击坑的识别上，无需进行复杂的数据预处理变换就能够达到一个较好的精度，相较于经典统计学方法，精度更高，操作更简便，并且可以通过检测结果，对专家撞击坑目录做有效补充。

第 5 章 基于分层策略下的 MaskRCNN 网络的月球撞击坑识别

上一章节中我们使用语义分割模型在月球撞击坑数据集上进行了有效分割，但由于语义分割模型逐像素识别的任务特性使得模型的分割速度受限，而月球图像中的数据大部分为背景，对背景类逐像素分析是比较浪费时间的，而目标检测任务更关注感兴趣区域，只需要对可能存在目标的区域进行检索，对于撞击坑识别应用是更有效率的选择。因此本章节从目标检测方法的角度入手，通过 MMdetection 平台构建实例分割模型 MaskRCNN 网络完成对月球撞击坑的识别，同时引入 Head 撞击坑目录补充大型撞击坑数据，并根据撞击坑直径范围提出了撞击坑数据分层策略，以重采样方式弥补大型撞击坑因直径过大难以识别问题，提高数据利用程度。本章主要内容包括经典目标检测模型的介绍、实验设计以及实验结果分析。

5.1 经典目标检测模型介绍

(1) RCNN 网络

首次将深度学习方法引入目标检测领域的是 Girshick 等^[58]提出的区域卷积神经网络 (Regions with CNN Features, RCNN)，通过选择性搜索算法(Selective Search)找出图片中可能包含检测类别的区域，再利用卷积神经网络提取特征，然后使用支持向量机进行分类，最后构建线性模型预测目标所在的位置。

具体而言，RCNN 网络首先使用选择性搜索算法对输入图像生成约 2000 个候选区域，将图像分割成小区域，将具有相似特征的小区域进行合并，直到整张图像被合并到一个区域为止，候选区域即为在合并过程中出现的区域；之后将候选区域归一化，统一裁剪成相同尺寸，然后采用深度卷积神经网络作为基础网络骨架提取图像特征；最后使用支持向量机作为分类器进行判别，输入为 AlexNet 网络中全连接层输出的一维特征向量，输出为是否属于此类的结果，同时为了提升定位准确度，RCNN 网络用边界框回归 (Bounding-Box Regression)的方法对结果框进行了修正。

(2) Fast RCNN 网络

虽然 RCNN 在目标检测任务中有着优异表现,但仍存在很多问题制约着精度的提升,例如在生成候选区域的过程中会存在大量重叠情况,这在神经网络中会产生大量冗余计算,同时由于分类和回归两个任务是独立进行的,没有利用两者之间存在的关联关系,这些问题使 RCNN 的运行速度很慢,因此在 RCNN 的基础上 Girshick^[59]改进了对候选区域的处理方式,同时巧妙地将训练过程中简化为一个基于多任务损失的单阶段训练,在保持模型精度的同时,大大缩减运行时间。

首先对整幅影像进行了卷积和池化运算,形成初步的卷积特征图,在通过搜素性算法生成候选区域后,将候选区域和整张图片作为输入,送入深度卷积网络中提取特征,为了加快训练速度, Fast RCNN 采用了一种感兴趣区域池化(ROI Pooling)的方法,使每一个感兴趣区域被池化成相同大小的特征图,比如感兴趣区域为 150x150, 在提取特征后等操作后特征图变为 15x15, 然后通过最大池化,将特征图变为 3x3 的大小,这样可以使每一个感兴趣区域的特征图大小一致,最后利用全连接层对特征向量进行分类。具体操作及流程如图 5-1、图 5-2 所示:

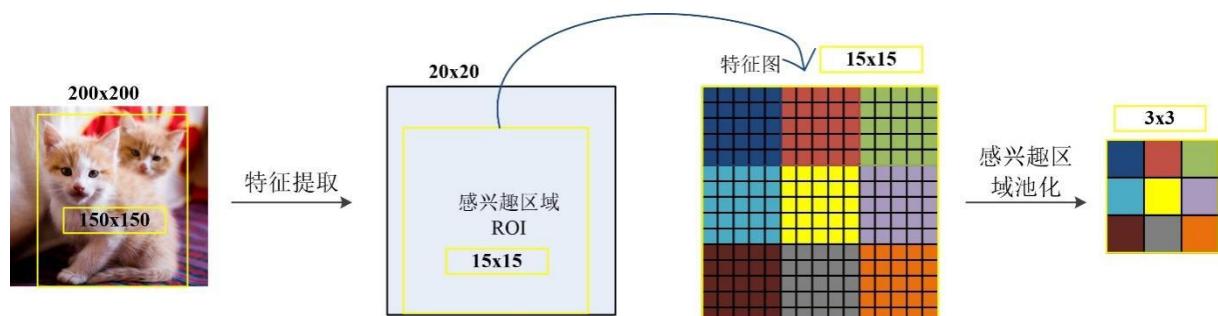


图 5-1 ROI Pooling 运算示意图

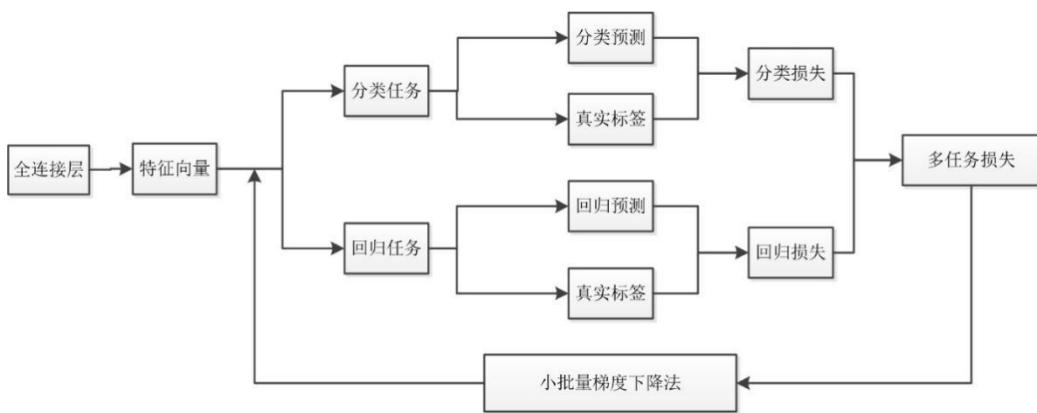


图 5-2 多任务损失函数计算流程图

(3) Faster RCNN 网络

Fast RCNN 网络相比于 RCNN 网络，在训练和测试阶段的速度已有较大提升，但在生成候选区域仍然是一个耗时耗力的工作，任少卿等^[60]在 Fast RCNN 的基础上用区域生成网络(Region Proposal Network, RPN)来代替选择性搜索算法，提出了 Faster RCNN 网络。Faster RCNN 主要由四个模块组成：首先使用深度卷积网络提取图像的特征图，之后用 RPN 网络生成提议区域(region proposals)，最后将 RPN 网络生成的提议区域映射到卷积得到的特征图上，经过 ROI Pooling 层生成区域特征图，然后输入进全连接层，最后用分类器预测分类结果，用边界框回归获得检测框最终的精确位置。

由于 Faster RCNN 网络通过引入 RPN 网络替代原有的选择性搜索，使得特征提取、候选区域生成、边界框回归、分类检测都整合在了一个网络模型中，不仅提高了模型的综合性能与各模块间的协同，更极大提升了目标检测的速度。

(4) MaskRCNN 网络

实例分割模型 MaskRCNN 网络是何恺明等^[28]在 Faster RCNN 网络的基础上做出的改进工作，通过改良 ROI Pooling 的计算方式，设计了更精准的池化方法 ROI Align，使用双线性插值的方法替代 ROI Pooling 直接取整的操作，减少误差，不仅提高了检测的精度，还支持同步输出目标的分割图。

MaskRCNN 网络首先对输入图像做特征提取，通常选择 ResNet50 或 FPN 等基础网络提取输入图像的特征，然后将其输出的特征图送入 RPN 网络中生成候选区域，再通过 ROI Align 进行感兴趣区域池化，最后增加了一个全连接网络分支，用来输出检测目标的类别检测框和分割图。

5.2 实验设计

由于撞击坑识别任务的核心是将撞击坑完整的提取出来，相比于常规目标检测模型，实例分割模型不仅识别精度更高，而且可以输出需要识别的目标轮廓，满足撞击坑识别任务的需求。因此本节实验选用 MaskRCNN 网络对月球撞击坑进行检测和分割，并且通过 MMdetection 框架对 MaskRCNN 网络的训练过程进行优化，提高网络泛化能力。同时基于月球撞击坑的直径特点，采用分层标记的方式重新构造月球撞击坑数据集，增强对数据的利用程度。

5.2.1 数据增强

在深度学习实验中，由于训练样本数据量小，通常需要对数据集在原有基础上进行扩充，提高训练样本量，增强模型的学习能力，不仅能避免模型出现过拟合情况，还能提高模型的性能，增强模型的鲁棒性，常用的数据增强方法有：

一、随机裁剪：将原始数据按照一定比例进行放大和缩小，改变图像中的表面特征，防止网络过度学习到训练样本的特定特征，避免模型出现过拟合情况，增加模型的稳定性，本实验将数据随机裁剪至原图的 0.8-1.25 倍。

二、水平翻转：将原始数据按照一定角度进行翻转，增强网络对不同形状的撞击坑的适应能力。本实验将训练集数据按照 50% 的比例对进行随机翻转。

三、数据标准化：对图像进行标准化处理，先计算训练样本的均值和方差，然后将所有图像的像素值减去均值再除以方差，移除图像的平均亮度值，使网络更关注目标本身。

四、加载预训练权重：调用在大型数据集 ImagNet 上经过训练的模型权重作为实验模型的初始权重，使模型更快收敛，并具有更强的泛化性。

5.2.2 分层策略

在语义分割实验中，只使用了 Povilaitis 月球撞击坑目录，该目录主要记录了小型撞击坑数据，但是在训练过程中由于大型撞击坑标签的缺失，会导致网络对一些不存在于 Povilaitis 目录中的大型撞击坑产生误判情况，制约实验的精度提升，为了弥补该缺点，本节引入 Head 撞击坑目录，增加标记撞击坑范围，提高撞击坑数据利用程度，但由于撞击坑直径范围过大，从几千米到几百千米，特征差距明显且会存在覆盖问题，直接对其进行标记会导致标记数据的不准确。通过对撞击坑分析，Head 目录中存在的大型撞击坑产生年代较为久远，大部分出现退化现象，其撞击坑的外表特征基本不明显，且由于大型撞击坑的高度更深，在数字高程模型的表现较为明显，因此为增强标签数据的可靠程度，避免网络对大型撞击坑进行误判，本文根据撞击坑直径大小设计了分层策略，根据撞击坑尺寸分为四个层次，以第一层次为基准，将其他三个层次的撞击坑下采样至

第一层次，使大型撞击坑的深层特征在数据中表现更明显，增强目标检测网络对撞击坑特征的学习能力。重新生成撞击坑数据集，并将该数据集命名为 Dataset3。

(1) 分层标记

Povilaitis 目录撞击坑直径范围为 5km-20km，Head 目录撞击坑范围为 20km 以上，根据撞击坑直径范围进行分层标记(标记方法与 3.2.1 节一致)，首先以直径在 5km-20km 的撞击坑为第一层标记，直径在 20km-80km 的撞击坑为第二层标记，直径在 80km-320km 的撞击坑为第三层标记，直径在 320km 以上的撞击坑为第四层标记，依次在全月高程模型数据上对专家撞击坑目录中存在的撞击坑进行标记，得到四张大小为 30720×92160 的标记图。为保持撞击坑直径大小分布一致，对第二、三、四层标记图分别进行下采样处理，将第二层标记图下采样 4 倍得到 7680×23040 的特征图、第三层标记图下采样 16 倍得到 1920×5760 的特征图，第四层标记图下采样 64 倍得到 480×480 的特征图。

(2) 数据裁剪

为方便模型的训练，需要保证输入图片的尺寸大小一致，将第一层标记图和经过下采样后的第二、三、四层特征图统一裁剪至 480×480 ，数据集划分不变，仍以 0° E- 120° E 区域的数据作为验证集， 120° E- 240° E 区域的数据作为训练集， 240° E- 360° E 区域的数据作为测试集，共得到 13107 张影像。

(3) 转换标签格式

按照上述操作制作的标签数据为掩膜数据，而目标检测模型通常所需的标签数据格式为 Json 文件，因此需要将其转化为训练所需格式，并按照 COCO 数据集的格式对数据集进行重新整理。

5.2.3 MaskRCNN

本章基于目标检测框架 MMdetection 构建的 MaskRCNN 模型共分为三个层次：特征提取部分、候选区生成部分和输出预测结果部分。首先在特征提取部分使用 ResNet50 网络作为骨干网络，提取输入图像的特征，之后再用 FPN 网络对 ResNet50 网络提取的特征再次优化，加强网络对多尺度特征的学习能力，其中 FPN 网络的每一层输入分别为 ResNet 网络的每一层输出；然后在候选区生成部分使用 RPN 网络，在特征图上生成撞击坑的候选区域，其中以交叉熵损失函数作为分类损失，用于计算锚框属于前景还是

背景的概率，并以 $SmoothL_1Loss$ 作为边界框损失，以修正目标提议框的位置；最后将 RPN 网络输出的特征图经过 ROI 并 Align 层送入模型输出层，在该阶段分两个分支：一个分支使用全连接层进行分类，预测前景或者背景，并输出撞击坑的预测框；另一个分支使用卷积网络输出撞击坑与背景的分割图。网络架构如图 5-3 所示：

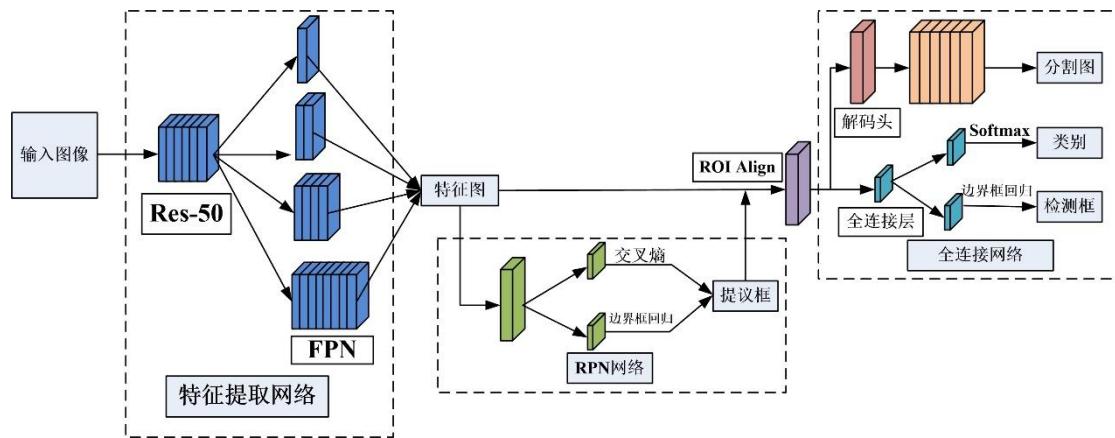


图 5-3 MaskRCNN 模型框架示意图

5.2.4 实验平台介绍及参数设置

实验选用的测试平台为商汤科技提供的开源项目——OpenMMLab，其中包含了众多基于 PyTorch 框架的工具包，如语义分割工具包 MM Segmentation、目标检测工具包 MMDetection、影像分类工具包 MMClassification 等，同时其内置有多种算法和预训练模型，给研究人员的研究工作带来了极大便利。本章节实验主要依赖 MMDetection 工具包进行，目前该项目已经在工业界和学术界被广泛应用，从 2019 年起，谷歌学术引用已超一千次，腾讯、阿里等多家互联网公司将其作为训练框架应用到相关产品中，其成熟性和可靠性已经得到充分认可，相关代码可参考

<https://github.com/open-mmlab/mmdetection>。

本次实验的实验设备及参数设置如表 5-1 所示：

表 5-1 实验设备、环境配置及超参数设置

实验设备	GPU	一块 NVIDIA RTX 3090 显卡(内存 24G)
	CUDA、CuDnn	11.1、8.0.5
环境配置	Python、PyTorch	3.8、1.8.0
	MMDetection、MMCV	2.28.1、1.3.8
	分类损失	<i>CrossEntropyLoss</i>
损失函数设置	回归损失	<i>SmoothL1Loss</i>
	Mask 分割损失	<i>CrossEntropyLoss</i>
	Batch Size	4
超参数设置	Optimizer	SGD(learning rate=0.0025, momentum=0.9)
	Dropout	0.10
	Epoch	40

5.3 实验结果分析

为了验证本章构建的 MaskRCNN 模型对月球撞击坑识别的有效性，分别从定性角度和定量角度两个方面对模型的检测结果进行评估。其中定量分析是通过使用深度学习常用的评价指标对模型的性能进行检验，以精准的数值对模型的真实性能进行综合评价；定性分析通过使用训练好的模型在验证集和测试集进行可视化展示，从直观的可视化结果上评估该模型的性能以及存在的问题。

5.3.1 定量分析

本实验的数据集为原始数据集 Dataset2 和使用分层策略的数据集 Dataset3，依然采用召回率，精准率和 F_1 分数作为定量评价指标对模型的检测结果进行分析，其中 TP 为正确认别的撞击坑的检测框个数、 FP 为错误识别为撞击坑的检测框个数、 FN 为遗漏识别的撞击坑检测框个数。选取的对比实验为 AliDib^[29] 等人在 2020 年使用原始 MaskRCNN 模型在月球撞击坑的识别结果。实验结果如表 5-2 所示：

表 5-2 实验结果展示

Model	Val($\pm 1.0\%$)			Test($\pm 1.0\%$)			Mean		
	Precision	Recall	F_1	Precision	Recall	F_1	Precision	Recall	F_1
(MaskRCNN)									
Dataset2	76.5	88.6	82.1	79.4	90.7	84.6	78.0	89.7	83.4
Dataset3	78.7	89.9	83.9	79.6	90.3	84.6	79.2	90.1	84.3
AliDib et al	--	--	--	40.2	85.1	54.6	40.2	85.1	54.6

从评价指标来看，本论文基于 MMdetection 框架构建的 MaskRCNN 模型的实验结果相比与原始的 MaskRCNN 网络有了很大提升，在月球撞击坑的识别结果中 F_1 分数已经达到 83.4%，在专家撞击坑目录的召回率也已达到 90%，这足以说明本章所应用的 MaskRCNN 模型对月球撞击坑的检测识别是有效的，并且在使用分层策略下，MaskRCNN 网络的指标均有所提升，证明了该策略是有效的。

5.3.2 定性分析

为更清楚地展现 MaskRCNN 模型的识别效果，本节将数据集 Dataset3 的预测结果进行可视化展示，由于验证集和测试集的范围太大，我们根据撞击坑分布密度，选取密度较大的区域作为示例展开分析，可视化结果如图 5-4、图 5-5 所示：

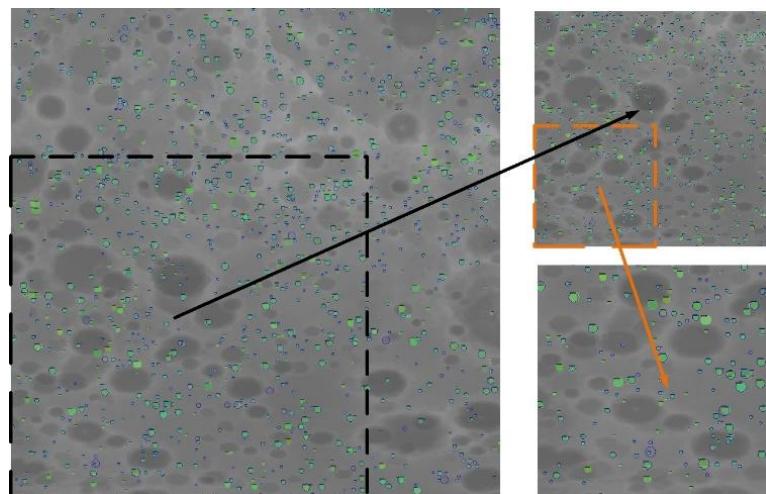


图 5-4 蓝色圆圈为专家识别撞击坑，绿色实心圆为 MaskRCNN 网络预测结果(区域分别为 [15°S,60°S,0°E,45°E]、[30°S,60°S,0°E,30°E]、[45°S,60°S,0°E,15°E])

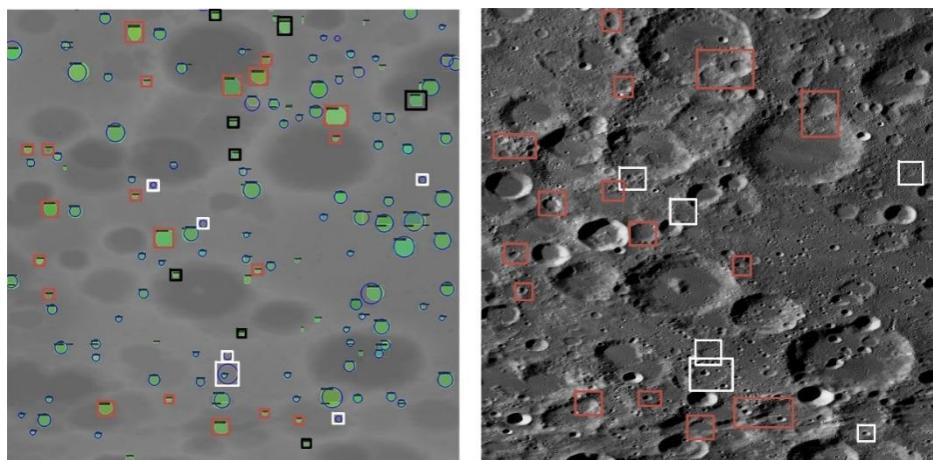


图 5-5 左侧为 MaskRCNN 网络在区域[45°S,60°S,0°E,15°E]的预测图, 右侧为该区域的高分辨率光学图像(橙色框为目视可见撞击坑, 黑色框为目视不可见撞击坑, 白色框为网络漏检的撞击坑)

从图 5-5 中我们可以看到, MaskRCNN 模型能够充分识别专家撞击坑目录, 漏检的撞击坑数目很少, 如白色方框所示, 在光学图像中能够很明显看到这部分撞击坑受光照影响较轻, 可以推断出这些撞击坑的深度较浅, 这是使用高程模型数据进行训练的不足之处; 此外, MaskRCNN 模型也识别到了相当一部分未被专家撞击坑目录所记录的撞击坑, 如橙色方框所示, 该部分的撞击坑主要处于大型撞击坑的边缘, 地貌状况较为复杂, 其在光学影像中有明显的撞击痕迹, 因此通过该网络检测的撞击坑可视为对专家撞击坑目录的补充; 同时 MaskRCNN 模型也识别到一些难以辨别的撞击坑, 如黑色方框所示, 这一部分的撞击坑直径普遍较小, 难以从光学影像中判断其是否属于撞击坑, 该部分撞击坑的识别仍需结合实际情况作进一步判断。

5.4 本章小结

本章节主要从月球撞击坑的直径特征出发, 设计分层策略解决难以对多尺度撞击坑实现精准识别的问题, 并从目标检测任务需求出发, 基于 MMdetection 框架优化 MaskRCNN 模型的训练过程, 提高对月球撞击坑的识别精度, 主要的工作有:

一、引入更大直径的撞击坑专家标记库对数据集进行重新标记, 根据对大型撞击坑的特征进行分析, 提出了分层策略, 通过将不同直径的撞击坑重采样至同一层次, 增加数据集中撞击坑的特征信息; 使用数据增强方法对月球撞击坑数据集进行数据扩充, 增

强模型的泛化性能，进而提高深度学习方法对月球数据集的利用程度；同时具体介绍了 MaskRCNN 模型的具体架构和实验过程参数的设置情况。

二、基于 MMdetection 框架重新构建 MaskRCNN 模型，并分别在原始数据集和使用分层策略的数据集上进行实验，实验结果证明了本章提出的分层策略的有效性，同时从定性和定量两方面分析 MaskRCNN 网络在月球撞击坑识别领域的优势和不足，以可视化方式将模型的预测结果清晰展示出来，进一步探析深度学习方法在对撞击坑识别中可以发挥优势的地方。

本章节的实验结果显示在小型撞击坑的识别上，只需要少量训练标签，MaskRCNN 网络就能够迅速且精确的自动识别出撞击坑，并在图像中将其分割出来，同时设计的分层策略能有效提高模型的整体检测效果。相较于语义分割方法，精度更高，消耗时间更短，对设备要求不高，操作简单。

第 6 章 结论及展望

6.1 研究结论

撞击坑识别算法发展到现在已经有四十多年的历史，从刚开始的人工识别到统计学方法识别、再到如今深度学习的应用，识别的数量从几百个到几万个，识别的范围从几百米到几十公里，识别区域从局部到全月球，一步步提高了人们对撞击坑的了解程度和天体形貌的整体认知。但是仍然存在很多未知的撞击坑等着我们去探索，如今基于深度学习的方法极大促进了撞击坑识别的发展，但是这些深度学习方法的侧重点各有不同，识别的效果也有差异。本论文根据月球撞击坑特征，选取月球数字高程模型数据，基于卷积神经网络对月球撞击坑进行高效准确识别，分别从语义分割和目标检测角度出发，主要研究成果和创新点如下：

一、针对语义分割任务，分别构建了基于 ResNet 和 MobileNet 为骨干网络的语义分割模型，并设计了 DenseASPP 模块进一步增加了 ResUNet 模型的分割精度，通过实验证明，本论文构建的模型提高了深度学习在该领域的识别精度和应用效率，最后设计了后处理模块，对分割结果进行优化，提升了模型的分割结果。

二、针对目标检测任务，基于 MMdetection 框架重新构建了 MaskRCNN 模型，在训练过程中对数据集使用多种数据增强方法提高模型训练效果，并设计了撞击坑数据分层策略，提高模型对多尺度撞击坑的识别能力。实验结果显示在月球撞击坑数据集具有高效性，且撞击坑分层策略有效提升了目标检测模型在月球撞击坑识别的预测精度。

三、分析和总结深度学习模型在月球撞击坑的预测结果，进一步补充月球撞击坑专家库的遗漏，发现深度学习方法在该领域存在的优势为对小型撞击坑具有更强的敏感性。同时本论文以较为简单的操作实现了对月球撞击坑的智能识别，拓宽了该领域的研究视野，为更好解决撞击坑识别难点提供新的思路。

6.2 研究展望

本研究还可以从以下几个方面对月球撞击坑展开进一步识别研究：

一、增加研究数据类型：现有的月球观测数据类型丰富，除了数字高程数据外，还有红外图像、高分辨率的光学图像、3D 点云数据等，不同类型的数据可以从不同角度刻画撞击坑的特征，单一的数据来源可能会因数据本身存在的问题而使深度学习模型难以识别到一些特殊的撞击坑，比如一些深度较浅的撞击坑在数字高程模型中表现不明显，容易被忽略，在未来的研究中，如何充分利用撞击坑的多源信息是研究的重点。

二、拓宽研究模型深度：本论文主要利用卷积网络在局部特征提取能力强的特点对撞击坑进行识别，但现有基于 Transformer 的深度学习模型从全局特征出发，在图像识别领域也具有很大优势，对于处于复杂地形的撞击坑和直径较大的撞击坑，该类型的模型可能会起到更有效的作用。其中，语义分割方法存在的不足之处主要由于训练数据的制约，在地形复杂的区域，难以对撞击坑进行精准分割，且无法有效对直径较大的撞击坑进行识别；目标检测方法中存在的不足之处在于难以识别复杂地形的撞击坑，且对退化的撞击坑识别效果不佳。

三、丰富研究对象范围：由于数据集的限制，本论文的主要研究对象为直径范围在 5-20km 以内的撞击坑，对于该范围之外的撞击坑的识别还需要进一步的研究。

同时，在深空探测领域中除了月球，仍存在很多地外星体值得人类探索，但由于距离遥远，高昂的探测成本制约了人类的探索进程，如果能将深度学习方法迁移至地外星体的前期研究中，能够以低成本的方式对地外星体构建初步认知，使人类可以针对性地对有价值的地外星体进行深度开发。

参考文献

- [1] 欧阳自远. 月球科学概论[M]. 北京: 中国宇航出版社, 2005, 1-2.
- [2] 李春来, 刘建军, 任鑫等. 嫦娥一号图像数据处理与全月球影像制图[J]. 中国科学: 地球科学, 2010, 40: 294-306.
- [3] Gan, F. P., Yu, Y. M., Yan B K. Primary Study of the Relationship Between the Lunar Surface Topography and Geological Informations[C]. Lunar and Planetary Science Conference, 2010.
- [4] 郭弟均, 刘建忠, 张莉等. 月球地质年代学研究方法及月面历史划分[J]. 地学前缘, 2014, 21(06): 45-61.
- [5] Yue, Z., Di, K., Liu, Z., et al. Lunar regolith thickness deduced from concentric craters in the CE-5 landing area[J]. Icarus, 329: 46-54.
- [6] 王庆龙. 月球雨海地区的地质演化[D]. 吉林大学, 2016.
- [7] OuYang, Z. Y. Scientific objectives of Chinese lunar exploration project and development strategy [J]. Advance in Earth Sciences, 2004, 19(03): 351-358.
- [8] 张锋, 邹永廖, 郑永春等. 月表撞击坑自动识别与提取的新方法及其应用[J]. 地学前缘, 2012, 19(06): 118-127.
- [9] Kang, Z., Luo, Z., Hu, T., et al. Automatic Extraction and Identification of Lunar Impact Craters Based on Optical Data and DEMs Acquired by the Chang'E Satellites[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(10): 4751-4761.
- [10] 袁悦锋, 朱培民, 赵娜等. 基于数学形态学的月海圆形撞击坑自动识别方法[J]. 中国科学: 物理学 力学 天文学, 2013, 43(03): 324-332.
- [11] 徐辛超, 徐爱功, 刘少创. 一种结合灰度特征的月面撞击坑提取方法[J]. 遥感信息, 2017, 32(01): 79-84.
- [12] 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(01): 265-274.
- [13] Bandeira, L., Saraiva, J., Pina, P. Impact Crater Recognition on Mars Based on a Probability Volume Created by Template Matching[J]. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(12): 4008-4015.
- [14] Sawabe, Y., Matsunaga, T., Rokugawa S. Automated detection and classification of lunar craters using multiple approaches[J]. Advances in Space Research, 2006, 37(01): 21-27.
- [15] 岳宗玉, 刘建忠, 吴淦国. 应用面向对象分类方法对月球撞击坑进行自动识别[J]. 科学通报, 2008, 22: 2809-2813.
- [16] Vamshi, G. T., Martha, T. R., Kumar, K. V., et al. An object-based classification method for automatic detection of lunar impact craters from topographic data[J]. Advances in Space Research, 2016, 57(09): 1978-1988.
- [17] Silburt, A., Ali-Dib, M., Zhu, C. C., et al. Lunar Crater Identification via Deep Learning[J]. Icarus, 2019, 317(06): 27-38.
- [18] Lee, C. Automated crater detection on Mars using deep learning[J]. Planetary and Space Science, 2019, 170: 16-28.
- [19] DeLatte, D. M., Crites, S. T., Guttenberg, N., et al. Segmentation Convolutional Neural Networks for Automatic Crater Detection on Mars[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(08): 2944-2957.
- [20] Lee, C., Hogan, J. Automated crater detection with human level performance[J]. Computers and Geosciences, 2021, 147: 104645-104663.
- [21] Wang, S., Fan, Z. Z., Li Z.M., et al. An Effective Lunar Crater Recognition Algorithm Based on Convolutional Neural Network[J]. Remote Sensing, 2020, 12(17): 2694-2712.
- [22] Wu, Y. T., Wan, G., Liu, L., et al. Intelligent Crater Detection on Planetary Surface Using Convolutional Neural Network[C]. IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference, 2021, 5: 1229-1234.
- [23] Mao, Y. Q., Yuan, R. G., Li, W., et al. Coupling Complementary Strategy to U-Net Based Convolution Neural Network for Detecting Lunar Impact Craters[J]. Remote

- Sensing, 2022, 14(03): 661-678.
- [24]Jia, Y. T., Liu, L., Zhang, C. Y. Moon Impact Crater Detection Using Nested Attention Mechanism Based UNet++[J]. IEEE Access, 2021, 9: 44107-44116.
- [25]Emami, E., Ahmad, T., Bebis, G., et al. Lunar Crater Detection via Region-based Convolutional Neural Networks[C]. 49th Lunar and Planetary Science Conference, 2018.
- [26]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), 2017, 936-944.
- [27]Hsu, C. Y., Li, W.W., Wang, S.Z. Knowledge-Driven GeoAI: Integration Spatial Knowledge into Multi-Scale Deep Learning for Mars Crater Detection[J]. Remote Sensing, 2021, 13(11): 2116-2135.
- [28]He, K. M., Gkioxari, G., Dollár P., et al. MaskR-CNN[C]. IEEE International Conference on Computer Vision, 2017, 2980-2988.
- [29]Ali-Dib, Mohamad, Menou, et al. Automated crater shape retrieval using weakly-supervised deep learning[J]. Icarus, 2020, 345-359.
- [30]Zang, S., Mu, L., Xian, L., et al. Semi-Supervised Deep Learning for Lunar Crater Detection Using CE-2 DOM[J]. Remote Sensing, 2021, 13: 2819-2840.
- [31]Benedix, G. K., Lagain, A., Chai, K., et al. Deriving surface ages on Mars using automated crater counting[J]. Earth and Space Science, 2019, 7(03): 1005-1020.
- [32]崔兴立, 丁恣, 王冠. 基于卷积神经网络的月球南极-艾特肯盆地撞击坑自动识别及中型撞击坑绝对模式年龄估算[J]. 南京大学学报(自然科学), 2021, 57(06): 905-915.
- [33]Lin, X. X., Zhu, Z. W., Yu, X. Y., et al. Lunar Crater Detection on Digital Elevation Model: A Complete Workflow Using Deep Learning and Its Application[J]. Remote Sensing, 2022, 14: 621-641.
- [34]Harald Hiesinger, James, W. Head. New Views of Lunar Geoscience: An Introduction and Overview[C]. Reviews in Mineralogy and Geochemistry, 2006, 60 (01): 1–81.
- [35]Hubel, D. H., Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex[J]. The Journal of Physiology, 1968, 195(01): 215-243.

- [36]Lecun, Y., Boser, B., Denker, J. S., et al. Backpropagation applied to handwritten zip code recognition[J]. *Neural Computation*, 1989, 1(04): 541-551.
- [37]Lecun, Y., Bottou, L., Bengio, Y., et al, Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [38]Platt, John. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines[J]. *Advances in Kernel Methods-Support Vector Learning*. 1998, 208-229.
- [39]Krizhevsky, A., Sutskever, I., Hinton, G. E. Imagenet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2012, 60: 84-90.
- [40]LeCun, Y., Bengio, Y., Hinton, G. E. Deep learning[J]. *Nature*, 2015, 521: 436–444.
- [41]Rumelhart, D. E., Hinton, G. E., Williams, R. J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533–536.
- [42]Hinton, G. E., Srivastava, N., Krizhevsky, A., et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. *Neural and Evolutionary Computing*, 2012, 3(04): 212-223.
- [43]Ioffe, S., Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]. *Proceedings of the 32nd International Conference on Machine Learning*, 2015, 37: 448-456.
- [44]Simonyan, K., Zisserman, A. Very deep convolutional networks for large-scale image recognition[C]. *International Conference on Learning Representations*, 2015, 1-14.
- [45]He, K. M., Zhang, X., Y., Ren, S.Q., et al. Deep Residual Learning for Image Recognition[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, 770-778.
- [46]Head, J. W., Fassett, C. I., Kadish, S. J., et al. Global distribution of large lunar craters: implications for resurfacing and impactor populations[J]. *Science*, 2010, 329(5998): 1504-1507.
- [47]Povilaitis, R. Z., Robinson, M. S., Bogert, C. H., et al. Crater density differences: Exploring regional resurfacing, secondary crater populations, and crater saturation

- equilibrium on the moon[J]. Planetary and Space Science, 2018, 162: 41-51.
- [48] Robbins, S. J., Antonenko, I., Kirchoff, M. R., et al. The variability of crater identification among expert and community crater analysts[J]. Icarus, 2014, 234: 109-131.
- [49] Robbins, S. J., Stuart, J. A New Global Database of Lunar Impact Craters >1–2 km: 1. Crater Locations and Sizes, Comparisons With Published Databases, and Global Analysis[J]. Journal of Geophysical Research: Planets, 2019, 124(04): 871-892.
- [50] Barker, M. K., Mazarico, E., Neumann, G. A., et al. Haruyama, J., Smith, D. E. A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera[J]. Icarus, 2016, 273: 346-355.
- [51] Long, J., Shelhamer, E., Darrell, T. Fully convolutional networks for semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015, 3431-3440.
- [52] Ronneberger, O., Fischer, P., Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. Medical Image Computing and Computer-Assisted Intervention, 2015, 9351: 234-241.
- [53] Andrew, H., Mark, S., Grace, C., et al. Searching for MobileNetV3[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019, 1314-1324.
- [54] Hu, J., Shen, L., Sun, G. Squeeze-and-Excitation Networks[J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, 7132-7141.
- [55] Chen, L.C., Papandreou, G., Kokkinos, I. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(04): 834-848.
- [56] Gao, H., Zhuang, L., Maaten, L. V. D., et al. Densely Connected Convolutional Networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, 2261-2269.
- [57] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 171-196.
- [58] Girshick, R., Donahue, J., Darrell, T., et al. Rich Feature Hierarchies for Accurate Object

- Detection and Semantic Segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014, 580-587.
- [59]Girshick, R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015, 1440-1448.
- [60]Ren, S.Q., He, K.M., Girshick, R., et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[C]. Proceedings of the 28th International Conference on Neural Information Processing Systems, 2015, 1: 91-99.

明止 德於 新至 民善