

基于改进 CBAM-DeepLab V3+ 的苹果种植面积提取

常 晗^{1,2} 郭树欣^{1,2} 张海洋^{1,2} 张 瑶^{1,2}

(1. 中国农业大学信息与电气工程学院, 北京 100083;

2. 中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083)

摘要: 为提高苹果种植区域的提取精度,提出了一种基于 Sentinel-2 和 MODIS 融合影像的 CBAM-DeepLab V3+ 模型。影响苹果种植区域提取精度的主要因素包括遥感影像的质量以及语义分割模型的性能。从影像质量角度来看,采用基于时序的时空融合算法 ESTARFM,通过融合 Sentinel-2 和 MODIS 的遥感影像数据,实现更高空间分辨率和时间分辨率数据的获取。与此同时,将训练样本从原始的 800 幅扩充至 2 400 幅,为后续语义分割模型提供更为充足的样本容量。在语义分割模型优化方面,为了进一步提高苹果种植面积的提取精度,以 DeepLab V3+ 网络结构模型为基础,引入基于通道和空间的 CBAM 注意力机制,进而发展出 CBAM-DeepLab V3+ 模型。与原始 DeepLab V3+ 模型相比,加入 CBAM 注意力机制的 CBAM-DeepLab V3+ 模型在拟合速度较慢、边缘目标分割不精确、大尺度目标分割内部不一致和存在孔洞等缺陷方面实现了突破,这些改进提高了模型的训练与预测性能。本研究采用原始 Sentinel-2 影像及时空融合后的影像数据集,结合烟台市牟平区王格庄镇的数据集和观水镇苹果数据集对 U-Net、FCN 以及 DeepLab V3+ 模型和 CBAM-DeepLab V3+ 模型进行对比,研究发现在苹果种植面积提取方面,CBAM-DeepLab V3+ 优化模型所取得的 MIoU 为 84.6%,苹果种植面积提取准确率达 90.4%。U-Net、FCN 和 DeepLab V3+ 模型的 MIoU 分别为 79.2%、75%、81.2%。此外,该模型预测的烟台市牟平区王格庄镇苹果种植面积为 3 433.33 hm²,与烟台市国民经济和社会发展统计公报公布的 3 666.66 hm² 相比,误差为 233.33 hm²,预测准确率高达 93.64%。

关键词: 苹果种植面积提取; 时空融合; 卷积神经网络; DeepLab V3+; 语义分割

中图分类号: TP75 **文献标识码:** A **文章编号:** 1000-1298(2023)S2-0206-08

Apple Planting Area Extraction Based on Improved DeepLab V3+

CHANG Han^{1,2} GUO Shuxin^{1,2} ZHANG Haiyang^{1,2} ZHANG Yao^{1,2}

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Key Laboratory for Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China)

Abstract: To improve the accuracy of apple cultivation area extraction, a CBAM-DeepLab V3+ model based on the fusion of Sentinel-2 and MODIS satellite images was proposed. The main factors affecting the accuracy of apple cultivation area extraction included the quality of remote sensing images and the performance of semantic segmentation models. From the perspective of image quality, a time-series spatiotemporal fusion algorithm called ESTARFM was employed to fuse Sentinel-2 and MODIS remote sensing image data, achieving higher spatial and temporal resolution data. Simultaneously, the training samples were increased from the original 800 to 2 400, providing more abundant sample capacity for the subsequent semantic segmentation model. In terms of optimizing the semantic segmentation model, in order to further improve the accuracy of apple cultivation area extraction, a CBAM attention mechanism based on channel and spatial information was introduced into the DeepLab V3+ network, resulting in the development of the CBAM-DeepLab V3+ model. Compared with the original DeepLab V3+ model, the CBAM-DeepLab V3+ model with the addition of CBAM attention mechanism achieved significant breakthroughs in terms of slower fitting speed, less accurate edge target segmentation, inconsistency in

收稿日期: 2023-06-20 修回日期: 2023-08-20

基金项目: 中央高校基本科研业务费专项资金项目(2023TC131)

作者简介: 常晗(1996—),女,硕士生,主要从事深度学习在农业智能检测方面的应用研究,E-mail: changhan@cau.edu.cn

通信作者: 张瑶(1988—),女,副教授,博士,主要从事农业定量遥感应用研究,E-mail: zhangyao@cau.edu.cn

segmenting large-scale targets, and existence of holes. These improvements enhanced the training and prediction performance of the model. The original Sentinel - 2 images and the spatiotemporal fusion images were used, combined with the datasets of Wanggezhuang Town in Muping District and the apple dataset of Guanshui Town to compare the U - Net, FCN, DeepLab V3 + models, and the CBAM - DeepLab V3 + model. The research findings indicated that in terms of apple cultivation area extraction, the overall accuracy (MIoU) achieved by the optimized CBAM - DeepLab V3 + model was 84.6%, and the accuracy of apple cultivation area extraction reached 90.4%. In comparison, the MIoU of U - Net, FCN, and DeepLab V3 + models were 79.2%, 75%, and 81.2%, respectively. Additionally, the predicted apple cultivation area of Wanggezhuang Town in Muping District was 3 433.33 hm², with only 233.33 hm² deviation compared with the data of 3 666.66 hm² published in the Yantai City National Economic and Social Development Statistics Report, resulting in a high prediction accuracy of 93.64%.

Key words: apple planting area extraction; spatiotemporal fusion; convolutional neural network; DeepLab V3 + ; semantic segmentation

0 引言

中国是世界上最大的苹果生产国,同时也是世界上苹果消费和出口量最大的国家^[1-2]。苹果种植产业在我国农产品发展中占有重要地位,对农业发展、市场稳定等方面影响巨大^[3-4]。烟台市作为中国苹果种植的发源地,拥有最广泛的种植面积和最丰富的苹果产量。遥感技术与深度学习图像语义分割模型的有效结合^[5],为研究烟台市苹果种植面积提供了新的方法。准确把握苹果果园种植面积的变化趋势和分布情况,对于指导农业生产、制定农业政策、市场监测和规划等具有重要意义。

多源遥感数据融合是当前遥感领域中的热点研究方向,也是提升深度学习模型样本容量的有效手段。针对当前遥感数据无法同时获得高时间分辨率和高空间分辨率影像的问题,研究者相继提出了多种基于时空多源遥感数据的高时空分辨融合方法。GAO 等在 2006 年首次将 Landsat 遥感数据和 MODIS 遥感数据相融合,提出了一种时空自适应反射率融合模型 (Spatial and temporal adaptive reflectance fusion model, STARFM)^[6]。当地表均质性较高时,STARFM 算法可以精确地预测高分辨率遥感影像的反射率,但是当地表异质度较高时,其预测精度就会降低,并且很难精准预测地表突然变化的现象,为此 ZHU 等^[7]在 2010 年提出了一种增强型时空自适应反射率融合模型 (Enhanced spatial and temporal adaptive reflectance fusion model, ESTARFM),通过引入转换系数和线性光谱解混理论,提高了算法的时空一致性,因此对于多时相影像的融合效果更好。融合不同源的遥感数据可以充分利用各种传感器的优势,弥补单一传感器数据的缺陷,从而得到更全面、更准确的信息^[8]。

近年来,深度学习在图像识别、图像分类、自然语言处理、图像分割等方面表现出色,它能够从图像

数据中提取丰富的语义信息。LONG 等^[9]提出全卷积网络 (FCN),使用全卷积层实现对任意尺寸图像的处理能力,但 FCN 在细节边界处理、类别不平衡和空间信息利用等方面存在一些限制,因此提出一种边界预测能力强的语义分割模型,U - Net 由 RONNEBERGER 等^[10]提出,它得名于其网络结构呈 U 形,广泛应用于医学图像分割,特别擅长处理边界清晰、样本不平衡的图像分割问题,但是 U - Net 在处理复杂场景时收敛速度较慢,场景资源受限。为了改善收敛速度较慢和计算复杂度低等问题,DeepLab 使用空洞空间金字塔池化 (Atrous spatial pyramid pooling, ASPP) 层^[11]来融合不同尺度的特征,从而更好地处理不同大小的目标物体。如何进一步提升训练速度,提高边缘目标分割精度,有效扩充数据集,改善网络结构,增强深度学习模型的检测精度和泛化能力值得进一步分析与探究。

本文以烟台市牟平区观水镇为研究区,采用 ESTARFM 算法对 Sentinel - 2 和 MODIS 影像进行时空融合,以获得更大容量的样本数据集和更全面更准确的信息。选用基于编码器 - 解码器架构的 DeepLab V3 + 网络模型进行优化,引入基于通道和空间的 CBAM 注意力机制,更加充分地考虑遥感影像中苹果边界像元分类的问题,根据高阶和低阶特征图的特性,进而发展出 CBAM - DeepLab V3 + 模型,以实现苹果种植面积更为精准的提取。

1 研究区及数据源

1.1 研究区概况

研究区位于山东省烟台市牟平区观水镇 (图 1),覆盖范围处于 37°12'23.32" ~ 37°14'20.17"N 和 121°12'32.61" ~ 121°16'17.34"E 之间,全镇总面积 228 km²,下辖 80 个行政村。观水镇年平均气温为 11.6℃,年平均降雨量 709 mm,境内地貌类型属剥蚀丘陵,气候属温带季风型大陆性气候,独特的地形

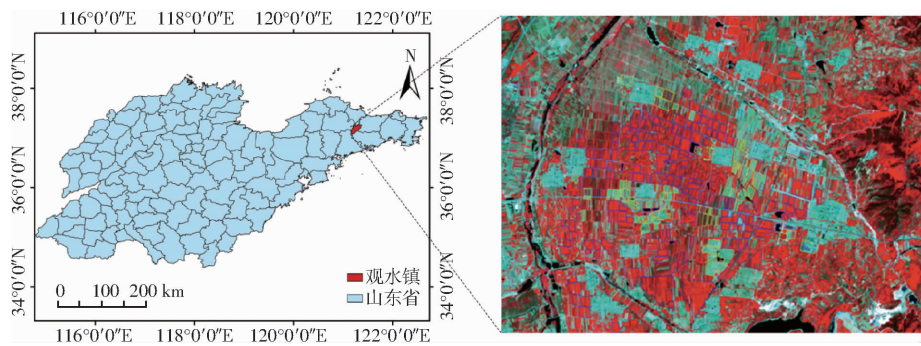


图 1 观水镇假彩影像

Fig. 1 False color image of Guanshui Town

和气候特点对苹果种植产生了积极影响。研究观水镇的苹果种植面积有助于揭示地形、气候等因素对苹果种植的影响,为未来果树种植规划提供指导。

1.2 遥感数据获取及预处理

1.2.1 Sentinel-2 影像数据及预处理

Sentinel-2 是欧洲空间局发布的一系列地球观测卫星,为研究地球表面特征(如植被^[12]、土壤^[13]、水体^[14]等)提供了丰富信息。本研究采集了烟台市观水镇苹果种植区域的 Sentinel-2 影像数据 3 景,成像日期分别为 2022 年 5 月 3 日、2022 年 6 月 2 日、2022 年 7 月 3 日。并对原始遥感影像进行预处理,包括辐射校正、大气校正、几何校正等操作,以提高图像质量和准确性。春季末尾至初夏时期,正是烟台地区苹果生长的关键时期,此时植被生长旺盛,叶绿素含量高,光谱特征明显,植被指数(如 NDVI^[15]、LAI^[16]等)对苹果树的响应较为敏感,能够更好地利用这些指数提取苹果种植面积。

1.2.2 MODIS 影像数据及预处理

MODIS 是一款中分辨率成像光谱仪,借助于 Terra 和 Aqua 两颗卫星的协同配合,MODIS 实现了几乎每天的全球覆盖,具备了较高的时间分辨率,这使其能够成为观察和追踪地球系统季节性变化和动态过程的理想数据源,支持多元化地球观测应用,例如气候变化监测^[17]、植被指数计算^[18]、陆地水分监测^[19]、海洋和冰雪研究。MODIS 遥感数据通过 NASA 的地球观测系统数据信息系统免费获取,对其进行转换、投影、重采样等操作,使之与 Sentinel-2 数据有着相同的像元尺寸。

1.2.3 Google Earth Engine 遥感大数据分析云平台

Google Earth Engine(GEE)是由谷歌推出的一种强大的地理空间数据分析和空间分析云平台。该平台旨在通过集成全球尺度长时间序列的遥感数据,如 Landsat、MODIS、Sentinel 等开源数据集,以及谷歌地球的卫星影像和其他地理空间数据,为用户 提供高效、快速的数据分析和处理能力。GEE 提供了一个功能强大的 JavaScript API 和 Python API,允

许研究人员、开发人员和 GIS 专业人员在云端对大量地理空间数据进行处理和分析。

由于 GEE 对 Python API 的支持不够完善,所以实验引入 geemap,geemap 是一个基于 GEE Python API 的开源 Python 库,旨在简化 GEE 在 Jupyter 环境中的地图创建、数据处理和可视化。geemap 提供一系列实用的功能,如交互式地图创建、数据叠加、地图导出等,为用户在 Jupyter 环境中使用 GEE 提供了便利。

1.3 样本数据

为了保证足够的数量和质量用于分类及精度验证,表 1 样本数据的采集方式是野外调查实地采样和利用目视解译标定二者相结合。使用奥维地图进行 GPS 定标,在苹果种植区域的边界和关键位置采集与地物类型相关的 GPS 定位点,将采集到的 GPS 定位点与实地考察记录相结合,得到有效样本点 496 个,其中苹果 211 个,林地 197 个,水体 28 个,村落 20 个,其他 40 个。

表 1 外业调查数据样本点

Tab.1 Field survey data sample points

地物名称	面积/m ²	点数	面数
其他	71 335	40	10
苹果	696 516	211	43
水体	83 879	28	7
林地	649 130	197	40
村落	402 118	20	4

1.4 数据增强

在本研究中,借助 GEE、geemap、QGIS 和 ENVI 等工具,对遥感影像进行预处理、裁剪、筛选和标注,为卷积神经网络模型训练提供数据支持。首先,通过 QGIS 选取苹果种植区域并生成矢量数据,接着对获得的 6 幅 10 240 像素×10 240 像素的遥感影像进行裁剪,6 幅影像分别为 2022 年 6 月 2 日原始影像的 RGB 影像及伪彩色影像,以及通过时空融合获得的 2022 年 5 月 28 日和 2022 年 6 月 16 日的合成影像的真彩色和假彩色影像,并转换为 jpg 格式。

之后,对影像进行去噪、增强和标注处理,使用 ENVI 和 Labelme 工具,将标注结果保存为 json 格式。同时,对图像进行数据增强,如翻转、旋转和缩放等。最后,将影像均匀裁剪为 2 400 幅 512 像素 × 512 像素的数据集。这套高质量数据集有助于提升卷积神经网络模型的泛化能力和准确性。

在获取基本数据集之后,为了对苹果种植区域进行精确标注,结合了 Google Earth 的高分辨率影像和先前收集的实地调查数据。在此过程中,采用 Labelme 工具进行图像标注处理。首先,使用多边形工具框选中图像的目标区域,并将其命名为“apple”。对于建筑物、水系和森林等其他地物,采用目视解译方法,分别将其标记为“building”、“water”和“forest”。未选中的区域将自动标记为“background”。完成标注后,生成了包含标注信息的 json 文件。接着,利用 Labelme 官方提供的 json 转换函数生成了对应的 png 标签文件。这些精确的标注信息为深度学习算法提供了准确的样本标签,为后续的模型训练提供了必要的支持,同时添加时空融合数据作为训练数据的补充。

接下来对遥感影像进行数据扩充操作,提高模型精度、防止过拟合。所采用的数据扩充方法包括水平反转、竖直反转以及对角反转,这些操作在维持原始影像特征的基础上,增加了数据的多样性和泛化能力。在完成数据扩充操作之后,本研究共获得样本 8 096 个。这些样本按照比例 6:2:2 被分配到训练集、验证集和测试集,以便进行后续的模型训练和评估。在此过程中,采用了交叉验证技术来训练和测试模型,以验证模型的稳定性和准确性。

2 研究方法

2.1 ESTARFM 时空融合模型

ESTARFM 是一种改进型时空融合模型,ESTARFM 增加了一种基于多尺度和多波段的相似性权重计算方法,以提高融合结果的准确性,同时 ESTARFM 还引入了时间滤波器来处理影像的时间变化,提高了算法的时空一致性,ESTARFM 还在算法中引入了数据估计和修复机制,可以有效处理影像中的遮挡和云等问题,提高了数据完整性,因此对于多时相影像的融合效果更好。

如图 2 所示,原始数据采用 2022 年 5 月 3 日的 Sentinel-2 原始影像与 2022 年 5 月 3 日的 MODIS 影像,通过 ESTARFM 算法与 2022 年 5 月 28 日 MODIS 影像融合,得到 2022 年 5 月 28 日高时空分辨率影像。采用同样的方法,将 2022 年 7 月 3 日 Sentinel-2 影像与 2022 年 7 月 3 日 MODIS 影像,

通过 ESTARFM 算法与 2022 年 6 月 16 日的 MODIS 影像融合,得到 2022 年 6 月 16 日高时空分辨率影像。

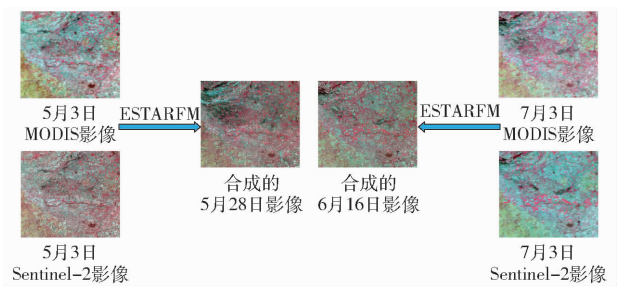


图 2 时空融合影像
Fig. 2 Spatio-temporal fusion image

2.2 DeepLab V3 + 网络模型

DeepLab V3 + [20] 是一种基于卷积神经网络的语义分割模型,通过引入空洞卷积和空洞空间金字塔池化(ASPP)等技术,旨在解决图像分割任务中的细节捕捉和多尺度特征提取问题。其网络结构主要由两部分组成:编码器(Encoder)和解码器(Decoder),图 3 为 DeepLab V3 + 模型结构。

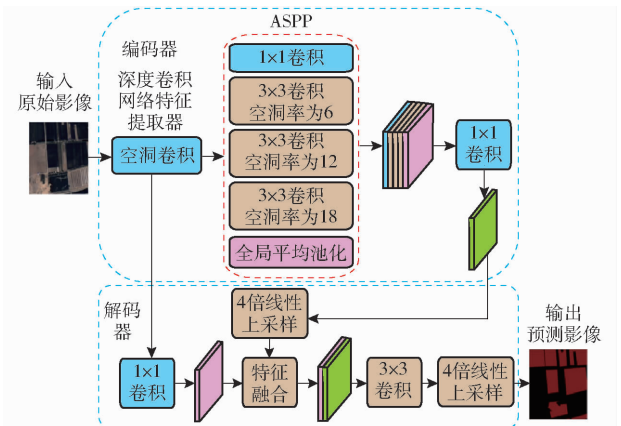


图 3 DeepLab V3 + 模型结构
Fig. 3 DeepLab V3 + model structure

编码器部分采用 Xception[21] 作为主干网络,用于提取图像特征。编码器中的 ASPP 模块通过多尺度空洞卷积捕捉不同尺度的上下文信息,提高语义分割的性能。

解码器部分负责逐步恢复特征图的尺寸,将编码器中的低级特征和上采样的高级特征进行融合,接着利用 3 × 3 卷积对融合后的特征图进行通道数变换,最后再通过 4 倍双线性插值上采样操作将特征图的分辨率恢复到与输入图像一致,并生成具有较高精度和细节表现的语义分割结果。

2.3 改进的 DeepLab V3 + 网络模型

注意力机制是一种模拟人类视觉注意力的方法,其目的是使模型在处理图像时更加关注与任务相关的区域。引入注意力机制可以帮助模型更好地捕捉地物特征,从而提高分类精度。CBAM 是一种

用于计算机视觉任务的注意力机制^[22],由通道注意力模块(Channel attention module)和空间注意力模块(Spatial attention module)组成,它旨在通过同时关注通道和空间信息来提高卷积神经网络的性能,如图4所示。

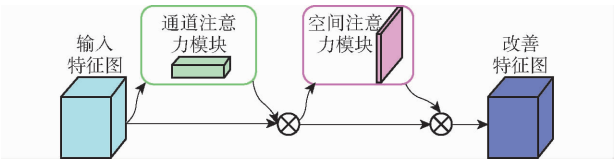


图4 CBAM 注意力机制结构
Fig. 4 CBAM attention mechanism structure

通道注意力模块如图5所示,对输入特征图 F 分别进行最大池化和平均池化,以捕获每个通道中的最大值和平均值信息,然后将得到的2个特征图传递到多层感知器,其中第1个全连接层进行降维,以减少计算复杂度,第2个全连接层进行升维,这2个层通过学习通道之间的关系来确定通道权重。经过多层感知器后的2个特征进行相加,通过 Sigmoid 激活函数产生通道权重,然后将这些权重与输入特征相乘,以调整每个通道的特征响应,生成空间注意力模块需要的输入特征。

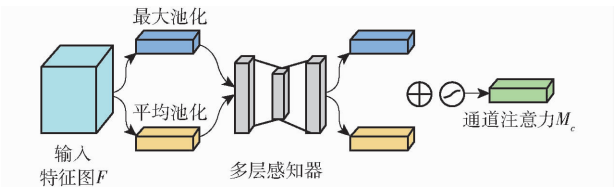


图5 通道注意力模块
Fig. 5 Channel attention module

空间注意力模块如图6所示,将通道注意力模块输出的特征图 F_1 作为本模块的输入特征图,经过最大池化和平均池化之后,将两者的结果联结起来形成一个具有双通道的特征图,通过卷积操作将组合后的特征图转换为单通道的特征图,经过 Sigmoid 激活函数后产生空间权重,这些权重将与输入特征图相乘,以加权每个空间位置的特征响应,得到最后的特征图。

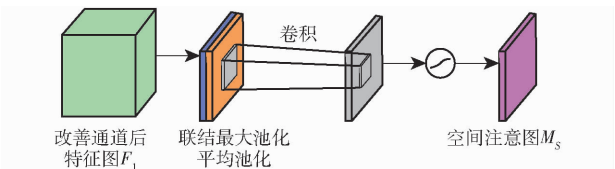


图6 空间注意力模块
Fig. 6 Spatial attention module

将 CBAM 双注意力机制添加到 DeepLab V3 + 模型中,如图7所示,模型将更好地关注与苹果种植面积提取任务相关的重要信息,从而提高网络对目标的定位精度。

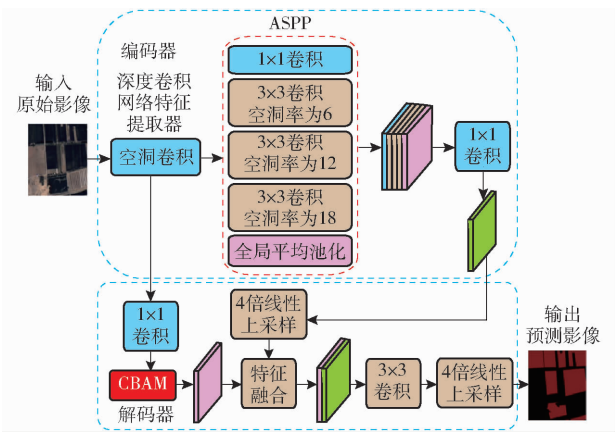


图7 添加 CBAM 注意力机制的 DeepLab V3 + 结构
Fig. 7 DeepLab V3 + structure of with added CBAM attention mechanism

在编码阶段,输入图像先通过具有多个通道的深度可分离卷积层进行主干特征提取网络的下采样,从而完成图像特征的捕获。随后将主干特征网络产生的高维特征图输入到 ASPP 模块,将这些处理后的输出叠加并拼接,并通过 1×1 卷积将通道数减少至256,降低计算复杂度,最后通过 CBAM 注意力机制进行加权操作,实现多尺度信息整合的高阶特征提取。

在解码阶段,首先对输入图像执行4倍下采样,以捕获具有丰富细节的低阶特征,利用 1×1 卷积调整通道数至48,并应用 CBAM 注意力机制处理低阶特征,过滤背景信息并凸显目标区域。接着将编码阶段获得的高阶特征执行4倍上采样,将 32×32 的特征图扩展至 128×128 ,使其尺寸与低阶特征图匹配。将低阶与高阶特征叠加,形成尺寸为 128×128 、通道数为304的特征图,之后通过 3×3 卷积调整通道数至256,实现特征融合,最后进行4倍上采样以恢复空间信息,并生成与输入图像尺寸一致的分类结果图。

3 实验结果与分析

3.1 ESTARFM 时空融合模型结果与精度评价

精度评价法主要是通过比较融合结果和实际地面信息的差异来评估融合的质量。在本研究中,针对2022年5月28日和2022年6月16日合成的遥感影像以及2022年6月2日的原始影像进行了定量评估。首先,提取这些影像的灰度,并计算均方根误差(Root mean square error, RMSE)、平均绝对误差(Mean absolute error, MAE)和决定系数(Coefficient of determination, R^2)。表2为计算结果。

对时空融合后影像与 Sentinel-2 原始影像进

表 2 融合影像及原始影像各项指标

Tab.2 Various indicators of fusion image and original image

影像	类型	MAE	RMSE	R ²
2022 年 6 月 2 日原始影像	RGB	0.007	0.012	0.88
	假彩色	0.011	0.016	0.80
2022 年 6 月 16 日合成影像	RGB	0.039	0.055	0.79
	假彩色	0.062	0.080	0.54
2022 年 5 月 28 日合成影像	RGB	0.014	0.023	0.87
	假彩色	0.024	0.035	0.77

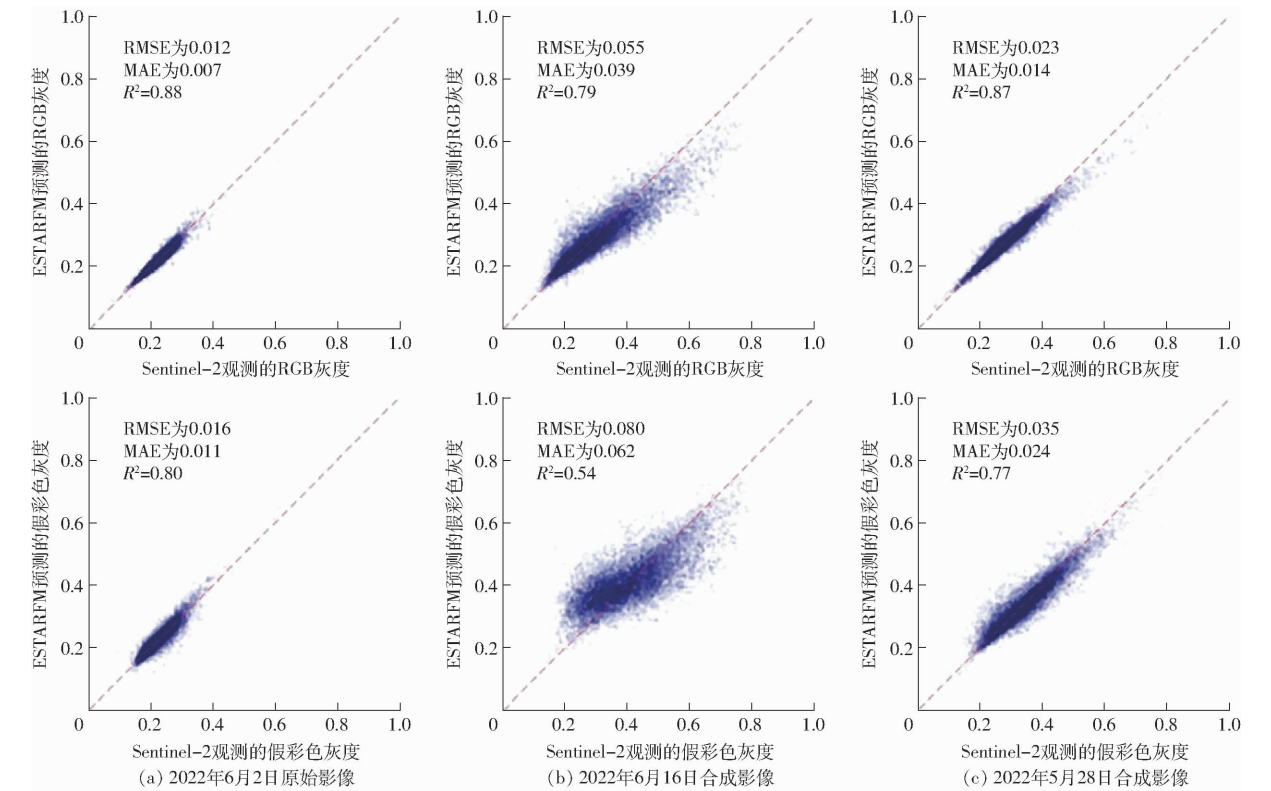


图 8 融合影像与原始影像对比

Fig.8 Comparing fused image with original image

硬件条件,选择使用 PyTorch 框架构建实验模型。

3.2.2 超参数设置

学习率是控制模型权重更新幅度的关键参数。较大的学习率可能导致模型无法收敛,而较小的学习率可能导致训练过程过慢,常见的初始学习率为 1×10^{-4} ,可以结合学习率衰减策略进行调整,本实验首先使用 1×10^{-4} ,每经过 5 个迭代周期,将学习率乘以 0.9。批量大小表示每次训练迭代使用的样本数量。较大的批量大小可以提高训练速度,但需要更多的显存,根据硬件限制,选择 16、32、64 作为批量大小。迭代次数是模型在整个训练集上进行训练的次数。过多的迭代次数可能导致过拟合,而过少的迭代次数可能导致欠拟合,本实验设定迭代次数为 100 个迭代周期。优化器是用于优化模型权重的算法。常见的优化器有 SGD (随机梯度下降)、Adam、RMSprop 等。在本实验中,选择使用 Adam 优

行对比,发现它们的 RMSE、MAE 和 R^2 已经达到可以作为训练数据使用的标准,在原始遥感影像数据不足的情况下,可以使用合成影像来辅助进行农作物种植面积的遥感提取,图 8 为融合影像与原始影像对比图。

3.2 改进算法有效性验证

3.2.1 实验环境

在本研究中,实验环境基于 Windows 10 操作系统,采用 NVIDIA A4000 显卡,综合考虑实验需求和

化器,其中 beta1 设置为 0.9,beta2 设置为 0.999,epsilon 设置为 1×10^{-8} 。损失函数用于衡量模型预测与实际标签之间的差异。对于本例中的语义分割任务,使用交叉熵损失 (Cross entropy loss) 作为损失函数。权重衰减是一种正则化方法,用于防止模型过拟合,设置权重衰减为 1×10^{-5} 。在 ASPP 模块中,设置空洞卷积膨胀率 (Dilation rates) 为 6、12、18。图 9 展示了初始学习率从 1×10^{-5} 上升到 1 过程中损失值与学习率的变化关系。

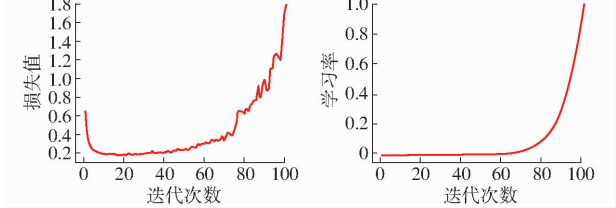


图 9 损失值及学习率变化曲线

Fig.9 Loss and learning rate variation curve

由图 9 可知,在第 21 轮时损失值最小为 0.175,此时学习率接近 1×10^{-4} ,因此初始学习率定为 1×10^{-4} ,表 3 为本实验的训练参数。

表 3 训练参数
Tab.3 Training parameters

参数	数值
批量大小	4
初始学习率	1×10^{-4}
学习率衰减因子	0.9
训练次数	100

为了验证本文中提出的改进方法的有效性,在制作的苹果提取数据集上对 U-Net、FCN、DeepLab V3+ 和改进后的 DeepLab V3+ 进行训练。使用 Xception 作为主干网络来完成模型训练,为了科学地评价实验成果,利用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、MIoU 和 Kappa 系数对模型提取结果进行比较和评估。训练结果如表 4 所示。

表 4 实验结果
Tab.4 Experimental results

模型	准确率	精确率	召回率	MIoU	Kappa 系数
U-Net	84.6	88.2	79.4	79.2	76.6
FCN	86.0	86.2	86.0	75.0	78.7
DeepLab V3+	88.4	90.0	91.6	81.2	87.5
CBAM-DeepLab V3+	90.4	92.4	93.0	84.6	90.6

相比于 U-Net、FCN 和 DeepLab V3+,改进后的模型在精确率、准确率、召回率上都表现出更高的精度,说明这个模型能有效地解决误分和漏分的问题,适合苹果种植面积提取。相较于 DeepLab V3+ 模型,改进后模型准确率提高 2 个百分点,这表明整体的特征提取效果得到了显著改善。同时,精确率与召回率分别提升 2.4 个百分点和 1.4 个百分点。精确率的显著增长表明改进模型在对苹果像元分类精度方面取得了重大进展。尽管召回率仅略有提升,但原始模型的召回率已处于较高水平,这也证实了 DeepLab V3+ 模型对苹果特征提取具有良好的适应性。Kappa 系数提高 3.1 个百分点,表明本研究的改进优化模型不仅在苹果分类准确度上取得了进展,其他地物分类精度及误分漏分情况也得到了优化。

3.3 算法优化后苹果种植面积提取

根据对比实验结果可得,改进后的 DeepLab V3+ 模型具有最高的提取精度和最佳的提取效果。因此,本研究选择改进后的模型对烟台市牟平区王格庄镇的作物分布信息进行提取。为了确保准确

性,对牟平区王格庄镇进行了实地考察,并通过训练后的模型输出了该地区的局部作物分布提取结果。

图 10 为提取了烟台市牟平区王格庄镇作物种植分布信息。参与提取的遥感影像空间分辨率为 10 m,即每个像素点对应面积为 100 m²。通过对被分类为苹果的像素点数量进行统计,得到牟平区王格庄镇的苹果种植面积数据。如表 5 所示,研究区域的苹果种植面积为 3 433.33 hm²,与牟平区政府提供的 2022 年王格庄镇苹果种植面积 3 666.66 hm² 相比,误差为 233.33 hm²,相对误差为 6.36%。利用 QGIS 3.22.15 对提取结果进行叠加分析后,发现在改进后的 Deeplab v3+ 模型提取的苹果分布信息中,共有 78 个验证点被正确分类,而错误分类点有 6 个,分类精度达到 93.64%。这表明该方法在提取作物种植面积信息方面具有较高的准确性。

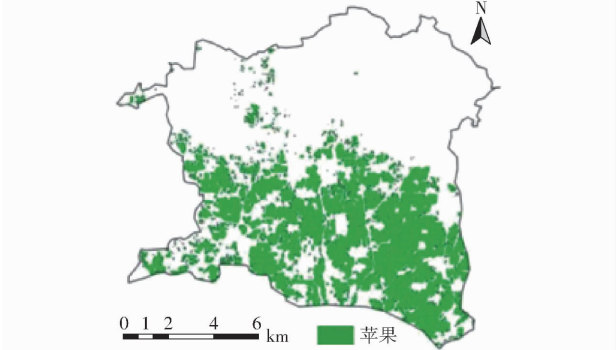


图 10 王格庄镇作物种植分布情况

Fig.10 Distribution of crop planting in Wanggezhuang Town

表 5 苹果种植面积提取模型对比
Tab.5 Comparison of apple planting area extraction models

模型	提取结果/hm ²	误差/hm ²	相对误差/%
FCN	3 273.33	393.33	10.72
U-Net	3 353.33	313.33	8.54
DeepLab V3+	3 393.33	273.33	7.45
CBAM-DeepLab V3+	3 433.33	233.33	6.36

4 结论

(1) 针对传统遥感方法在苹果种植面积提取上存在的精度不足和自动化水平较低的问题,本研究以 Sentinel-2 和 MODIS 融合影像为基础,提出了一种改进的 DeepLab V3+ 模型来提取苹果种植面积。通过将该方法与其他深度学习方法进行比较,证实了本研究所提出的改进模型在苹果种植面积提取方面具有更高的有效性。

(2) 基于 DeepLab V3+ 模型,采用 Xception 作为主干网络以提取特征,引入注意力机制以增强模

型关注重要特征的能力,并采用组归一化方法来减少模型复杂性,CBAM - DeepLab V3 + 优化模型所取得的 MIoU 为 84.6%,苹果提取准确率达 90.4%。

(3)应用改进方法对王格庄镇的苹果种植面积进行提取,根据提取结果计算得出的苹果种植面积为 3 433.33 hm²,而政府统计数据显示种植面积为 3 666.66 hm²,二者相差 233.33 hm²,预测准确率为 93.64%。

参 考 文 献

[1] 张强强,霍学喜,刘军弟,等.世界苹果产销格局及市场动态预测分析[J].世界农业,2016(7):147-152.
ZHANG Qiangqiang, HUO Xuexi, LIU Jundi, et al. The analysis and prediction of world apple production pattern and market dynamics[J]. World Agriculture, 2016(7): 147-152. (in Chinese)

[2] 束怀瑞,陈修德.我国果树产业发展的时代任务[J].中国果树,2018(2):1-3.
SHU Huairui, CHEN Xiude. The epochal task of fruit tree industry development in our country[J]. China Fruits, 2018(2): 1-3. (in Chinese)

[3] 孟艳玲,汪景彦,康国栋,等.我国苹果生产现状分析[J].中国果树,2007(1):43-44.
MENG Yanling, WANG Jingyan, KANG Guodong, et al. Analysis of current situation of apple production in our country[J]. World Agriculture, 2007(1): 43-44. (in Chinese)

[4] 程存刚,赵德英.新形势下我国苹果产业的发展定位与趋势[J].中国果树,2019(1):1-7.
CHENG Cungang, ZHAO Deying. Development positioning and trends of China's apple industry under the new situation[J]. World Agriculture, 2019(1): 1-7. (in Chinese)

[5] 何雨岩.基于深度学习的图像语义分割综述[J].长江信息通信,2023,36(1):77-79.
HE Yuyan. A review of image semantic segmentation based on deep learning[J]. Changjiang Information and Communications, 2023,36(1): 77-79. (in Chinese)

[6] FENG G, JEFFREY M, MATHEW S, et al. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006,44(8): 2207-2218.

[7] ZHU X, CHEN J, GAO F, et al. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions[J]. Remote Sensing of Environment, 2010,114(11): 2610-2623.

[8] 牛海鹏,王占奇,肖东洋.基于时空数据融合的县域水稻种植面积提取[J].农业机械学报,2020,51(4):156-163.
NIU Haipeng, WANG Zhanqi, XIAO Dongyang. County-level rice planting area extraction based on spatiotemporal data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020,51(4): 156-163. (in Chinese)

[9] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[10] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical Image Computing and Computer-assisted Intervention, 2015.

[11] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015,37(9): 1904-1916.

[12] 祁向前,孙德浩,贾连星.哨兵2号多时相植被指数作物分类及监测[J].测绘工程,2022,31(6):47-53.
QI Xiangqian, SUN Dehao, JIA Lianxing. Sentinel-2 multi-temporal vegetation index crop classification and monitoring[J]. Engineering of Surveying and Mapping, 2022,31(6): 47-53. (in Chinese)

[13] 赵成,梁盈盈,冯浩,等.基于 Sentinel-2 遥感影像的黄土高原覆膜农田识别[J].农业机械学报,2023,54(8):180-192.
ZHAO Cheng, LIANG Yingying, FENG Hao, et al. Identification of plastic mulched farmland on the Loess Plateau based on Sentinel-2 satellite imagery[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 180-192. (in Chinese)

[14] 吴庆双,汪明秀,申茜,等. Sentinel-2 遥感图像的细小水体提取[J].遥感学报,2022,26(4):781-794.
WU Qingshuang, WANG Mingxiu, SHEN Qian, et al. Small water body extraction method based on Sentinel-2 satellite multi-spectral remote sensing image[J]. National Remote Sensing Bulletin, 2022,26(4): 781-794. (in Chinese)

[15] 白燕英,高聚林,张宝林.基于 NDVI 与 EVI 的作物长势监测研究[J].农业机械学报,2019,50(9):153-161.
BAI Yanying, GAO Julin, ZHANG Baolin. Monitoring of crops growth based on NDVI and EVI[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(9): 153-161. (in Chinese)

[16] 周西嘉,张悦,王鹏新,等.基于 Sentinel 的时间序列田块尺度 LAI 重建与冬小麦估产[J].农业机械学报,2022,53(8):173-185.
ZHOU Xijia, ZHANG Yue, WANG Pengxin, et al. Reconstruction of time series LAI and winter wheat yield estimation at field scales based on Sentinel satellites[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(8): 173-185. (in Chinese)

[17] 易扬,胡昕利,史明昌,等.基于 MODIS NDVI 的长江中游区域植被动态及与气候因子的关系[J].生态学报,2021,41(19):7796-7807.
YI Yang, HU Xinli, SHI Mingchang, et al. Vegetation dynamics and its relationship with climate factors in the middle reaches of the Yangtze River based on MODIS NDVI[J]. Acta Ecologica Sinica, 2021,41(19): 7796-7807. (in Chinese)

[18] 田苗,单捷,卢必慧,等.基于 MODIS - EVI 时间序列与物候特征的水稻面积提取[J].农业机械学报,2022,53(8):196-202.
TIAN Miao, SHAN Jie, LU Bihui, et al. Extraction of rice planting area based on MODIS - EVI time series and phenological characteristics[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022,53(8): 196-202. (in Chinese)

[19] ZHANG F, ZHANG L, SHI J, et al. Soil moisture monitoring based on land surface temperature - vegetation index space derived from MODIS data[J]. Pedosphere, 2014,24(4): 450-460.

[20] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//European Conference on Computer Vision, 2018:833-851.

[21] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1800-1807.

[22] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision, 2018:3-19.