

基于 U-Net 和特征金字塔网络的秸秆覆盖率计算方法

马 钦^{1,2} 万传峰¹ 卫 建¹ 汪玮韬¹ 吴才聪^{1,2}

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 农业农村部农机作业监测与大数据应用重点实验室, 北京 100083)

摘要: 针对田间秸秆覆盖分散、秸秆形态多样, 细碎秸秆识别困难, 传统图像识别方法易受光照、阴影等因素干扰等问题, 本文以黑龙江省齐齐哈尔市龙江县为研究地点, 构建田间秸秆图像数据集; 对图像进行裁剪、标注后, 构建了以 U-Net 为基础网络的秸秆检测模型。将编码阶段的网络结构换成 ResNet34 的前 4 层作为特征提取器, 增加模型的复杂度, 增强秸秆特征的提取; 为增强秸秆边缘识别, 在最高语义信息层对深层特征图使用多分支非对称空洞卷积块 (Multibranch asymmetric dilated convolutional block, MADC Block) 提取多尺度的图像特征; 为增加细碎秸秆的检测能力, 在跳跃连接阶段使用密集特征图金字塔网络 (Dense feature pyramid networks, DFPN) 进行低层特征图和高层特征图的信息融合, 利用特征图对应秸秆图像中感受野的不同, 解决秸秆形态多样的问题; 为避免秸秆特征图在上采样时的无效计算, 解码阶段使用快速上卷积块 (Fast up-convolution block, FUC Block) 进行上采样, 避免秸秆特征图在上采样时的无效计算。试验表明, 本文算法在车载相机采集到的秸秆图像数据集上平均交并比为 84.78%, 相比 U-Net 提高 2.59 个百分点, 该网络对于 640 像素 × 480 像素的图像平均处理时间低于 3 ms, 符合作业检测时的时间复杂度要求, 算法在一定程度上改善了阴影区域秸秆的识别问题, 提高了细碎秸秆的识别能力。

关键词: 秸秆检测; 计算机视觉; 非对称空洞卷积; 特征图金字塔网络

中图分类号: S23; TP391

文献标识码: A

文章编号: 1000-1298(2023)01-0224-11

OSID:



Calculation Method of Straw Coverage Based on U-Net Network and Feature Pyramid Network

MA Qin^{1,2} WAN Chuanfeng¹ WEI Jian¹ WANG Weitao¹ WU Caicong^{1,2}
(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. Key Laboratory of Agricultural Machinery Monitoring and Big Data Application, Ministry of Agriculture and Rural Affairs, Beijing 100083, China)

Abstract: In view of the scattered straw mulching in the field, the various straw shapes, the difficulty in identifying the fine straw, and the traditional image recognition methods are disturbed by factors such as light and shadow easily. Taking Longjiang County, Qiqihar City, Heilongjiang Province as the research site, a field straw image dataset was constructed. After cropping and labeling the image, a straw detection model based on U-Net network was constructed. Changing the network structure of the coding stage to the first four layers of ResNet34 as the feature extractor, the complexity of the model was increased and the extraction of straw features was enhanced. In order to enhance the detailed identification of straw edges, the multibranch asymmetric dilated convolutional block (MADC Block) was used to extract multi-scale image features on the deep feature map at the highest semantic information layer. In order to increase the detection ability of fine straws, dense feature pyramid networks (DFPN) were used in the skip connection stage to perform information fusion of low-level feature maps and high-level feature maps. Using the feature map to correspond to the difference of the receptive fields in the straw image, the problem of variety of straw shapes was solved. In order to avoid the invalid calculation of straw feature map during upsampling, the decoding stage used fast up-convolution block (FUC Block) was used for upsampling. Experiments result showed that the average intersection ratio of the algorithm on the straw image dataset collected by the vehicle camera was 84.78%, which was 2.59 percentage points

收稿日期: 2022-03-05 修回日期: 2022-03-25

基金项目: 国家发展改革委员会综合数据服务系统基础平台建设项目 (JZNYYY001)

作者简介: 马钦 (1978—), 女, 副教授, 博士, 主要从事计算机视觉研究, E-mail: sockline@163.com

通信作者: 吴才聪 (1975—), 男, 教授, 博士, 主要从事农机无人驾驶与自主作业和农机大数据研究, E-mail: wuzz@cau.edu.cn

higher than that of U-Net. The average processing time of the network for images with a size of 640 pixels × 480 pixels was less than 3 ms. Compared with manual measurement, the error was less than 5%, which met the time complexity requirements of operation detection. The algorithm can improve the identification of straw in the shadow area to a certain extent, and improve the identification ability of fine straw.

Key words: straw detection; computer vision; asymmetric dilated convolution; feature pyramid network

0 引言

保护性耕作是东北黑土地保护利用的有效途径^[1],其中包括秸秆深埋还田、覆盖还田等方式,统称为秸秆还田^[2]。秸秆还田是一种保护水土肥力的重要措施,不但可以减少秸秆焚烧带来的环境污染以及资源的浪费,对于改善土壤结构、保留土壤水分、增加土壤养分起到重要作用^[3]。

秸秆覆盖率是评价免耕还田程度的一个重要指标,但是传统秸秆覆盖人工统计方法如拉绳法、采样法、目测法需要消耗更多的人力物力且效率低,更容易受主观因素的影响。图像处理方法和机器学习算法可以实现秸秆检测及其覆盖率计算。文献[4]提出一种基于多阈值的多目标秸秆图像自动分割优化算法(DE-AS-MOGWO),提高了图像分割的准确性和处理速度。文献[5]采用神经网络和纹理特征结合的方式提取秸秆,方法耗时较长,无法满足实际作业需求。文献[6]提出了一种随机决策森林算法对土壤中的残留物进行分类,使用简单的像素比较和具有训练偏移量的邻近像素对单个像素进行分类。文献[7]将快速傅里叶变换和 SVM 相结合对秸秆图像进行识别,设计多尺度占比滤波器,对图像中的噪声和空洞进行修补,算法适用于多种情况。文献[8]提出一种 Sauvola 和 Otsu 算法相结合的方式进行图像的阈值分割来检测秸秆覆盖率,对于干扰因素较多的区域检测效果不理想。文献[9]通过 K-means 聚类 and 秸秆图像分区寻优方法结合进行图像分割,为秸秆覆盖率在线计算提供一种新方法。文献[10]依据图像行平均灰度标准差对图像进行分类,针对有秸秆图像进行分块,使用 Otsu 阈值分割计算秸秆覆盖率,提高检测方法的适用性。但是传统阈值分割方法受背景因素、光线因素、阴影等噪声的影响较大,晒干的秸秆颜色和土壤相近,难于区分秸秆和土壤。农田秸秆形态多样,对于分散的细碎秸秆识别困难。随着深度学习的发展,图像语义分割技术已经成为当下的研究热点,语义分割是图像在像素级别的分类操作。目前,语义分割技术主要应用在遥感图像^[11-12]、自动驾驶^[13-14]、医学影像^[15]等领域。文献[16-17]利用遥感图像进行秸秆检测,标准误差分别为

4.846% 和 8.46%。文献[18]利用低空无人机获得现场图像,用改进的 U-Net(ResNet18-U-Net)建立语义分割算法进行覆盖率检测。文献[19-20]基于语义分割的 DSRA-UNet 算法检测秸秆覆盖率,一定程度上能解决图像中的阴影问题。但是无人机航拍图像数据需要单独采集,无法和农机播种的其他作业环节相结合。

针对田间秸秆细碎分散、形态多样、秸秆和土壤受光照影响难以区分等难点,本文使用车载相机采集秸秆图像数据,对复杂秸秆图像进行图像预处理,筛选有效识别区域,以 U-Net 网络和特征图金字塔网络作为基础结构进行改进,以提高细碎分散秸秆和阴影区域秸秆的识别能力。

1 整体设计

1.1 秸秆检测流程

秸秆覆盖率检测主要采用计算机视觉的方法,通过车载相机设备采集农田秸秆图像数据集。结合农田秸秆形状多样的特点,针对农田场景有逆光或光照遮挡导致的阴影问题,先对秸秆图像进行图像裁剪,获取相机拍摄的正前方有效区域,去除农田外的干扰背景后进行图像投影变换,减少相机拍摄时产生的透视畸变。使用 Labelme 标注得到真值,使用数据增强技术进行数据集扩充。在模型设计阶段为了提高秸秆在复杂场景下的提取能力,使用 ResNet34 加深模型的网络层数;在模型层数增加的情况下为了避免产生网络梯度消失现象,在模型中加入残差结构;同时使用改进的特征金字塔网络和多分支非对称空洞卷积块,利用不同的感受野,增加不同形状尺寸秸秆和细碎秸秆的提取能力;最后用快速上卷积进行特征图上采样,避免网络上采样过程中出现过多的无效数据。由网络模型输出得到预测图,结合手动标注的真值图像,使用平均交并比、像素准确率和运行时间等指标进行模型评估。检测流程如图 1 所示。

1.2 图像采集

由于免耕播种有作业补贴,需要进行作业面积监测,同时无人机采集数据要单独作业,难以区分农田的边界和识别地块归属,所以选择在播种机作业时,使用播种监控设备上的摄像头进行秸秆图像数

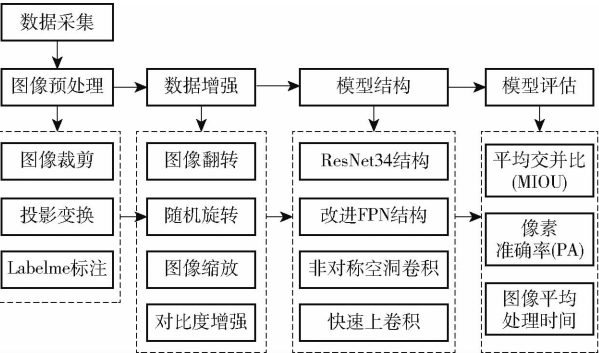


图1 秸秆检测流程图

Fig.1 Flow chart of straw detection

据的采集。本文所用秸秆图像数据为2020年5月采集于黑龙江省齐齐哈尔市龙江县,由北京德邦大为科技股份有限公司的车载监控设备采集得到,所用农机为DEBONT2BMG-2型播种机,监控设备包括显控终端、北斗定位终端、红外补光摄像头等,使用前后双摄像头拍摄,后摄像头主要监控作业情况,需要拍摄到农具,前摄像头拍摄范围为车前20 m以上,图像分辨率为640像素×480像素。从13块作业田块中选取200幅有效图像进行试验,训练集和验证集共150幅,测试集图像50幅。通过图像翻转、缩放、对比度增强等数据增强技术将训练集和验证集图像共扩增到1 250幅图像,其中训练集图像为1 000幅,验证集图像250幅。图2所示为不同覆盖等级下的秸秆图像数据,由图2e可知,田间的细小秸秆较多时,秸秆之间会有空隙(放大图中的椭圆框区域),这些像素点对应的是土壤,在进行秸秆

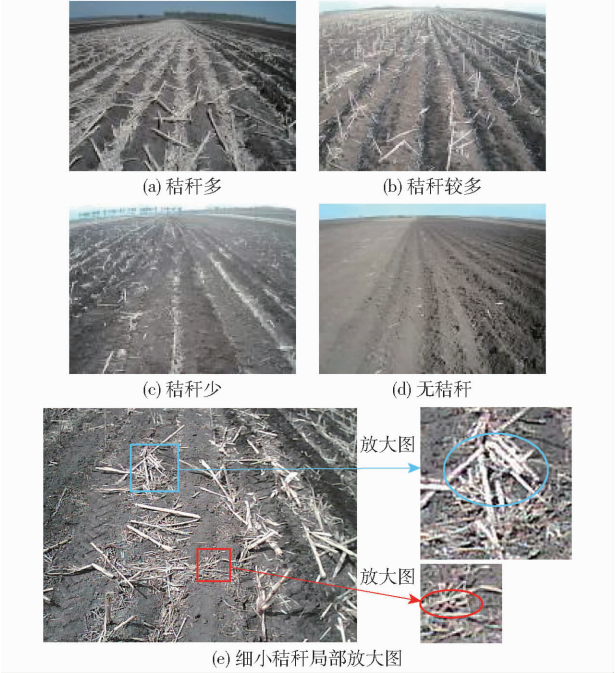


图2 不同覆盖等级的秸秆拍摄图像

Fig.2 Shooting images of straws with different mulching levels

检测时容易识别成秸秆区域,给秸秆覆盖度检测带来困难。

1.3 图像变换

由于不同农机相机安装位置、角度不同,导致图像拍摄视角产生差异,会有天空、房屋、树林等其他干扰背景存在,在图像预处理阶段,人工判定田块区域占整幅图像区域3/4以上为有效图像,并以此标准进行图像裁剪,去除天空区域。先使用图像行像素和的方式查找有效区域,天空区域对应的像素灰度大部分为230~250,田内土壤的像素灰度大部分在150以下,两者差异较大。根据图像的行像素和变化规律可知,一般情况下,水平分界线从天空区域到田内的过程中,行像素灰度总和逐渐减少,当有树林等其他干扰因素时,行像素和局部上升,临近分界线时像素和快速减少,因此图像的像素行数在[0, 120]范围内,判断田内和田外分界处的一个条件就是当行像素灰度和连续下降且下降的差值最大时,对应的水平线位置就是分界处,结果如图3a~3c黄线所示,分界线识别流程如图4所示。

由图3e可知分界线位置(红色虚线),计算分界线下区域的高与图像高的比例,对图像的宽进行等比例裁剪见图3h,最后将裁剪图缩到原尺寸得到结果图。

车载相机在采集图像数据时,秸秆图像存在近大远小的透视畸变,为了改善这种情况,文中采用图像透视变换^[21]进行处理,如图5所示,可以在一定程度上减小误差,投影变换公式为

$$\begin{cases} T = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \\ \begin{bmatrix} x' & y' & z' \end{bmatrix} = \begin{bmatrix} u & v & 1 \end{bmatrix} T \\ x = x'/z' \\ y = y'/z' \end{cases} \quad (1)$$

式中 T ——转换矩阵
(u,v)——图像像素坐标
(x,y)——转换之后归一化的图像坐标
(x',y',z')——转换后未归一化的图像坐标

1.4 图像标注

图像行灰度和检测到的有效区域图像,进行人工标注,生成标注图像作为真值,来评估语义分割模型的性能。使用Labelme标注工具,将图像分为土壤、秸秆两类,标签图中黑色表示土壤,红色表示秸秆。将图像中的物体按照不同的类别分别圈选并做好标记,用Python脚本语言生成标注图,结果如图6所示。

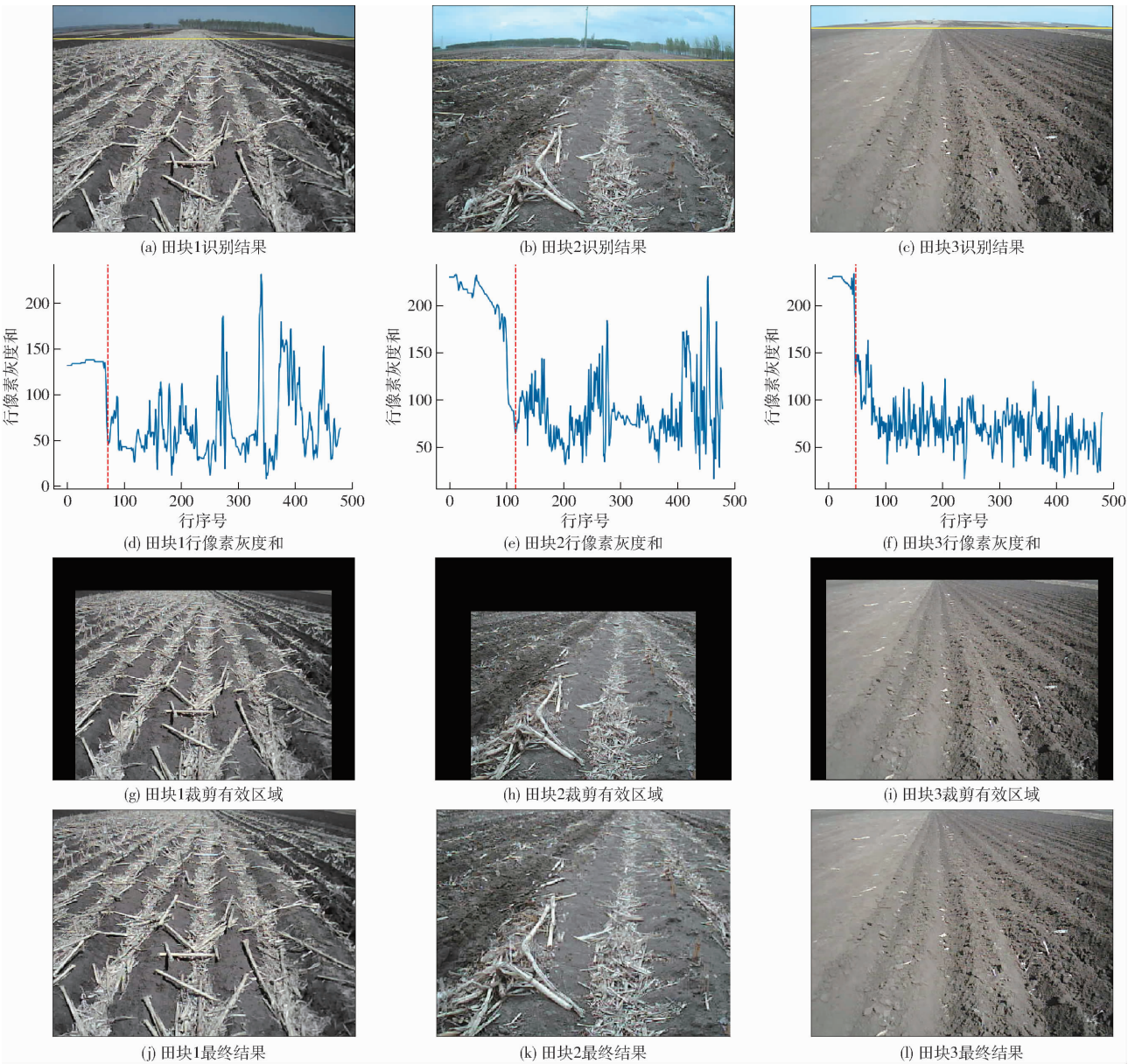


图 3 图像裁剪过程

Fig. 3 Process of image cropping

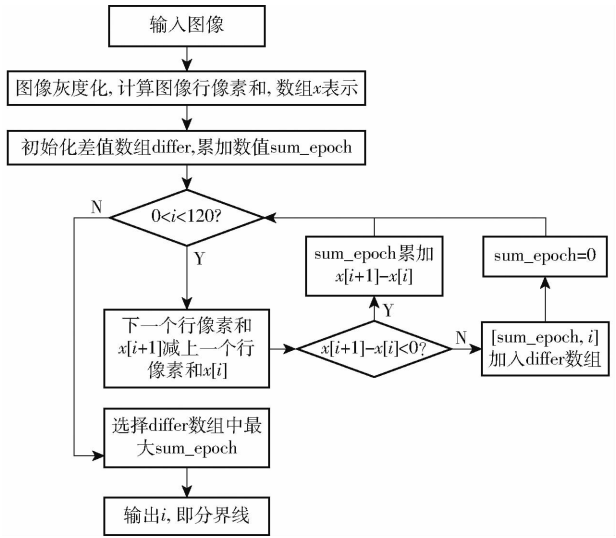


图 4 分界线识别流程图

Fig. 4 Flow chart of dividing line recognition

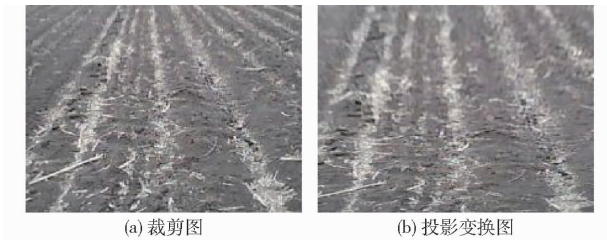


图 5 图像投影变换结果

Fig. 5 Result of image projective transformation

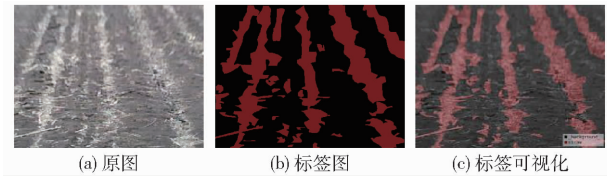


图 6 图像标注结果

Fig. 6 Result of image annotation

2 模型框架

2.1 模型概述

U-Net 网络是一种有编码器和解码器结构的

网络,通过跳跃连接将编码器阶段的特征图和解码器阶段的特征图信息进行融合。本文基于 U-Net 和 FPN 提出新的网络结构,如图 7 所示,其中特征图下方标注了该特征的尺寸和通道数。

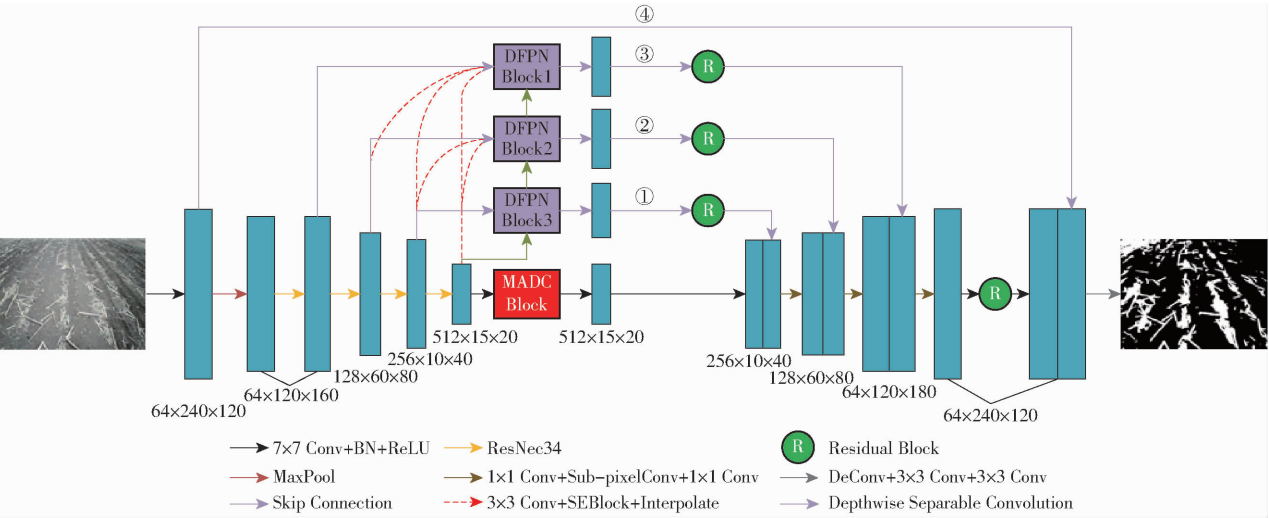


图 7 改进 U-Net + FPN 算法网络架构
Fig. 7 Network architecture of U-Net + FPN algorithm

网络模型包含特征编码器模块、最高语义信息层特征提取模块和特征解码器模块。编码阶段采用 ResNet34 的前 4 层网络结构,先进行卷积核尺寸为 7×7 的卷积和池化核为 2×2 的最大池化,然后通过多层串联的残差网络结构进行 4 次特征提取。最高语义信息层特征提取模块在密集空洞卷积^[22]的基础上提出多分支非对称空洞卷积块 (Multibranch asymmetric dilated convolutional block, MADC Block),使用多个分支的非对称空洞卷积对特征图进行不同尺度的信息提取。特征解码模块使用快速上卷积块代替普通上采样进行位置信息的恢复,同时用跳跃连接获取编码阶段同一层级的特征信息。跳跃连接中对编码阶段的特征信息通过密集特征金字塔网络 (Dense feature pyramid networks, DFPN) 模块获取不同尺度的特征图信息,前 3 次跳跃连接的 DFPN 模块之后使用标准残差块保持网络的收敛性,避免网络训练过程中的梯度消失现象。经过 4 次上采样得到的特征图通过残差块运算后与第 4 次跳跃连接得到的特征图进行融合,最后进行 1 次反卷积得到与输入图像相同尺寸的结果图。

2.2 编码-解码结构

FCN 网络模型^[23]提出了编码-解码结构,用卷积层替换传统卷积神经网络最后的全连接层,使用跳跃连接进行特征融合,摆脱了输入图像尺寸的限制。在 FCN 之后,同样使用编码-解码结构的 U-Net^[24]和 SegNet^[25]网络在此基础上进行了相应的改进。编码-解码结构主要是在网络中包含一个编码器模块和一个

解码器模块,编码器模块主要负责图像的特征提取,图像从空间信息转换到深层的语义类别信息,并记录每一个类别像素点的位置;解码模块对编码块输出的信息进一步优化,根据像素点的类别信息和位置信息进行复原,并将特征图放大到原输入图像的尺寸。

2.3 多分支非对称空洞卷积模块 (MADC Block)

MADC 模块根据空洞卷积^[26-27]思想,加入非对称空洞卷积 (Asymmetric dilated convolution, ADC) 进行秸秆信息识别,目的是为了扩大感受野同时,不降低图像分辨率且不引入额外参数量,非对称结构在不改变感受野的条件下,增强秸秆边缘关键点提取能力,提升对秸秆图像识别的鲁棒性。结合 Inception^[28-29]结构思想进行多尺度不同感受野的秸秆信息的获取,在增加网络宽度和深度的同时减少参数量,如图 8 所示,大感受野更好地提取大片覆盖秸秆信息并生成更抽象的特征,小感受野可以更好地提取小范围细碎秸秆信息。模块共有 5 个分支,第 1 个分支使用 3×3 卷积核,设置扩张率为 1;第 2 个分支使用扩张率为 3 的 3×3 非对称卷积,之后进行扩张率为 1 的 1×1 卷积;第 3 个分支先使用扩张率为 1 的 3×3 卷积,后跟扩张率为 3 的 3×3 非对称卷积,再进行扩展率为 1 的 1×1 卷积;第 4 个分支依次使用扩展率为 1 的 3×3 卷积、扩张率为 3 的 3×3 非对称卷积、扩张率为 1 的 5×5 非对称卷积和扩张率为 1 的 1×1 卷积;第 5 个分支使用原始输入特征图;MADC 模块中每次卷积之后都采取 ReLU 函数保持非线性。

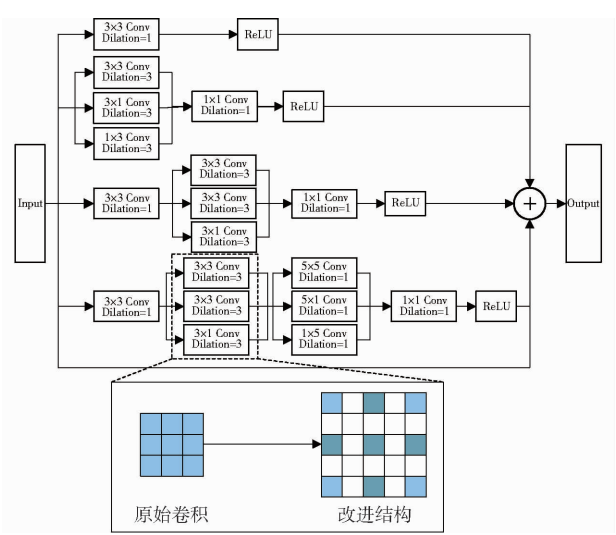


图 8 MADC 模块结构图

Fig. 8 Diagram of MADC module structure

2.4 残差块 (Residual Block)

在一定范围内,网络深度越深,拥有越好的性能,能够更好地提取秸秆特征,但是随着网络深度的增加,会导致梯度消失、梯度爆炸等问题出现^[30], ResNet34^[31]使用的残差网络可以缓解梯度消失现象,通过增加一条捷径学习融合原始特征信息,在增

加网络非线性的同时保持网络的原有特征,以增强表达能力。网络结构如图 9 所示。

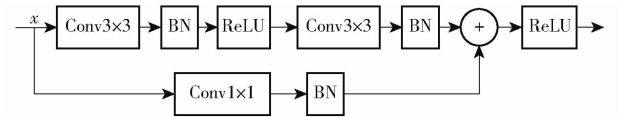


图 9 卷积残差块结构图

Fig. 9 Diagram of residual block

2.5 密集特征金字塔网络 (DFPN)

特征金字塔网络^[32] (Feature pyramid network, FPN)主要是为了解决目标尺寸差异较大时的检测缺陷,细小秸秆的像素信息较少,在网络下采样过程中容易丢失,在上采样过程中无法恢复,导致最后识别不成功。FPN 网络通过将语义信息较强的低分辨率特征图和空间信息更多的高分辨率特征图进行融合,构建多尺度的特征金字塔结构。在 FPN 的基础上,本文提出密集特征金字塔结构,如图 10 所示。网络结构通过稠密连接的方式,将下采样过程中的所有特征图都使用到特征图的上采样过程中,从而更充分地利用秸秆图像的细节信息。以第 2 次上采样过程为例,不同通道数的输入特征图(x_1, x_2, x_3)

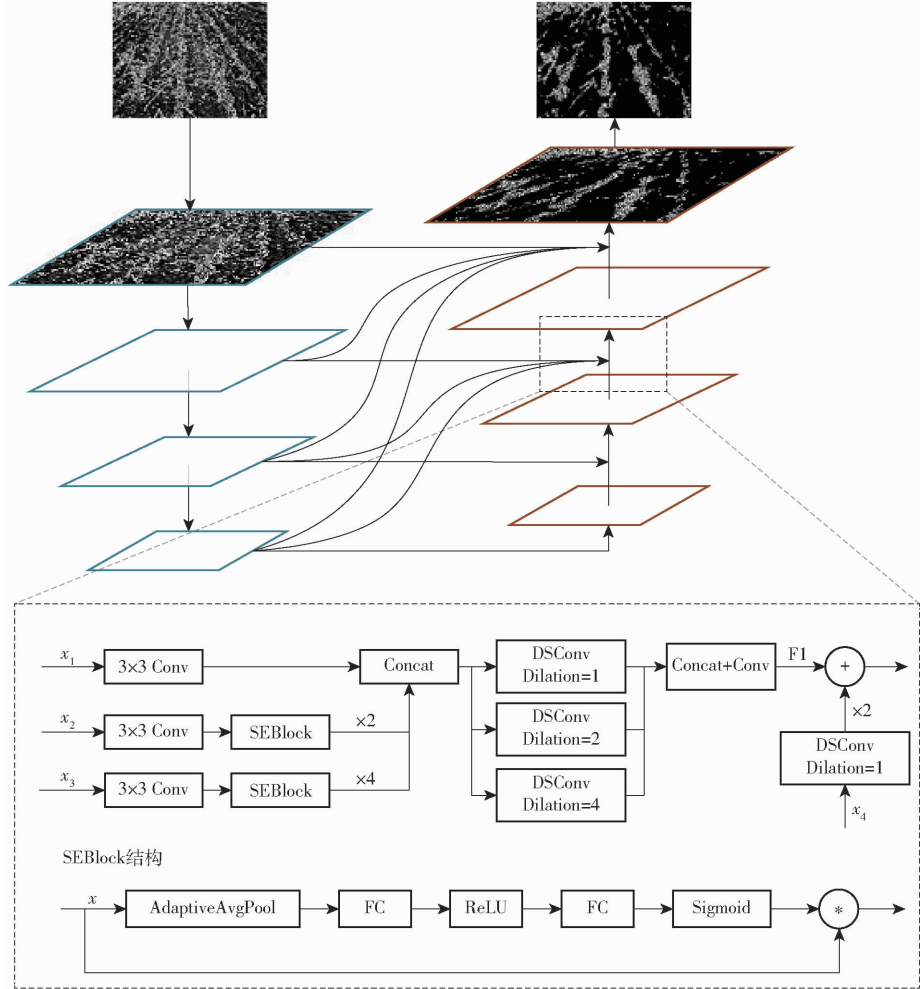


图 10 密集特征金字塔网络结构图

Fig. 10 Diagram of dense feature pyramid networks

使用 3×3 卷积转换到相同通道数。 x_2 、 x_3 卷积生成的低分辨率特征图使用 SE^[33] 模块 (Squeeze-and-Excitation block, SEBlock) 增强秸秆信息的特征,然后上采样到 x_1 特征图的相同尺寸。将得到的 3 个特征图进行拼接,再用多尺度的深度可分离卷积进行秸秆信息提取,以应对秸秆形状多样的问题,之后使用 1×1 卷积进行通道数转换生成特征图 F1。最后将前一个上采样过程中生成的特征图 x_4 依次进行深度可分离卷积改变通道数和上采样改变特征图尺寸,与生成特征图 F1 相加。

其中 SE 模块先在通道维度进行全局池化,得到通道的全局特征,然后使用全连接层和激活函数来获取通道之间的关系。第 1 个全连接层起降维的作用,降低模型的复杂度并且可以提升网络的泛化能力,后面跟 ReLU 激活函数,第 2 个全连接层负责将通道数恢复到原特征图的维度,后面使用 Sigmoid 函数激活,将各个通道的特征值缩到 0 ~ 1 之间,最后将各个通道的激活值乘以 x_2 、 x_3 卷积生成低分辨率特征图,提取每个通道中的主要信息。SE 模块基于通道注意力机制的方式,得到通道权重,来增强特征图中的秸秆特征,抑制其他干扰因素的特征。

2.6 快速上卷积块 (FUC Block)

常见的上采样方法有插值法、转置卷积和反池化。插值法在上采样时没有考虑图像边缘的信息,缩小了特征图不同特征边界处的差异,导致图像的边缘模糊。转置卷积会导致生成的图像中产生棋盘效应。反池化原理是特征图的下采样过程中,最大池化时记录最大值的索引位置,在上采样时将特征值填入记录的位置,并将其余位置补零,来增大特征图的尺寸,但是这种方法因为引入大量 0,在后续的卷积过程中大多计算是无效的。本文使用快速上卷积块^[34] (Fast up-convolution block, FUC Block) 进行特征图的上采样,结构如图 11 所示。对于原始特征

图,反池化之后会增加 75% 的 0 值,后续使用 3×3 卷积计算时,相当于对原始特征图进行 2×2 、 2×1 、 1×2 和 1×1 的有效参数计算。因此本文分别使用 2×2 、 2×1 、 1×2 和 1×1 的卷积核对原始特征图进行卷积,然后将生成的 4 个特征图拼接,并使用子像素卷积将拼接图的每个通道对应位置元素重新组合,生成尺寸比原始特征图增大一倍的特征图,来替代普通的上采样方法,在达到相同效果时,快速上卷积块比普通上采样方法计算量更小,加快了运算速度。

3 训练

3.1 损失函数

损失函数反映了模型预测值和真实值之间的差距,深度学习模型训练过程中使用损失函数进行模型评估和监测^[35-36],本文使用二分类的交叉熵损失函数,在进行梯度下降计算时可以避免网络出现梯度弥散的现象,加快网络的学习深度,定义为

$$L = -\frac{1}{n} \sum_i [y \ln a + (1 - y) \ln (1 - a)] \quad (2)$$

式中 n ——像素点总量
 i ——像素点序号
 y ——像素点真实值,若像素点为秸秆,取 1,否则取 0
 a ——像素点被预测为秸秆的概率,取 0 ~ 1

3.2 训练超参数

网络模型使用 Adam 优化器,在每一个 epoch 中计算损失函数的梯度,从而进行参数更新;学习率为 10^{-4} ;batch size 设置为 10;训练迭代周期为 200,激活函数使用 ReLU 函数。

4 试验结果

4.1 评价指标

为了检验模型识别秸秆的准确性,使用秸秆图像分割平均交并比 (Mean intersection over union, MIOU)、像素准确率 (Pixel accuracy, PA) 和单幅图像平均处理时间等来衡量算法性能。

4.2 改进网络结构评估

本文使用的模型预测平台 GPU 配置为 Tesla V100 32GB,训练时间约 5 h,为了更好地比较各个修改结构的效果,以 U-Net 为基础,使用控制变量法,逐步添加改进的结构,在训练集上进行模型训练,并通过验证集的指标准确率来比较模型的效果,如表 1 所示。

为了更具体地比较每个结构对模型改进的效果,进行了 5 组不同模型结构的试验,结果如表 2 所

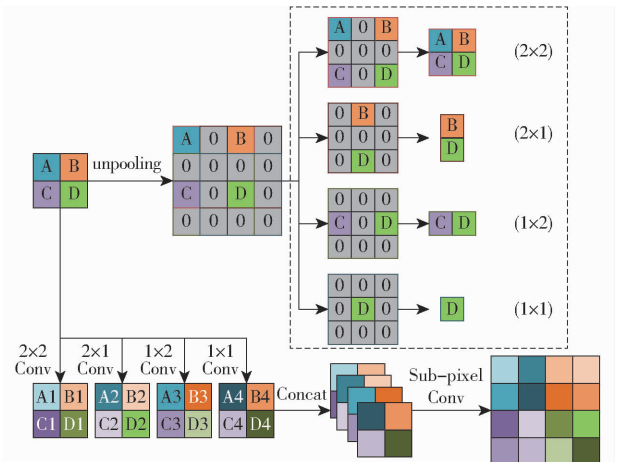


图 11 快速上卷积模块结构图

Fig. 11 Diagram of fast up-convolution module

表 1 编码阶段不同网络结构的准确率

Tab.1 Accuracy results of different network structures in encoding stage

特征提取网络结构	平均交并比	像素准确率/%
U-Net	82.19	94.64
U-Net + ResNet18	81.72	94.55
U-Net + ResNet34	82.25	94.81
U-Net + Vgg16	79.64	93.94

表 2 添加不同模块后效果对比

Tab.2 Comparison of effects after adding different modules

试验 序号	编码结构	多分支 非对称 空洞卷 积	密集特 征金 字塔	解码结构	平均 交并 比/%	像素 准确 率/%	处理 时间/ ms
	ResNet34			反卷 积 快速上 卷积			
1				√	82.19	94.65	1.40
2	√			√	82.25	94.81	1.75
3	√	√		√	83.78	95.29	1.81
4	√	√	√	√	84.59	95.63	2.44
5	√	√	√	√	84.78	95.65	2.53

注：“√”表示试验采用此模块结构。

示。试验 1 为初始网络 U-Net,试验 2 将 U-Net 编码阶段的网络结构替换为 ResNet34,试验 3 表示在试验 2 的基础上加入 MADC Block,试验 4 在试验 3 的基础上添加了 DFPN 模块,试验 5 在试验 4 的基础上,将上采样时的反卷积结构替换为 FUC Block。各网络结构在验证集上进行准确率比较。

由表 2 可知,使用 U-Net 网络进行农田秸秆分割,MIOU 为 82.19%,本文改进网络的 MIOU 提高 2.59 个百分点,达到 84.78%,由于网络结构中加入 MADC、DFPN 等模块,增加了训练参数,导致模型参数量和运行时间有所增加。对比试验 2、3 可以看出,MADC 模块可以更好地获取特征图中秸秆的空间信息,比未使用 MADC 模块时的 MIOU 提高 1.53 个百分点。对比试验 3、4 可以看出,DFPN 模

块可以将深层特征图语义信息和浅层的空间信息相结合,其中包含通道注意力机制用来减少干扰信息,进一步提高网络的特征提取能力,网络模型的 MIOU 提高 0.81 个百分点。对比试验 4、5 可以看出,加入快速上卷积模块比试验 4 的 MIOU 提高 0.19 个百分点,主要原因是 FUC Block 使用多个尺寸不同卷积核对特征图进行卷积,并使用子像素卷积进行特征图的信息融合,对于特征图的边界信息提取得更加准确。

对于 DFPN 结构中的 SEBlock,其中第 1 个全连接层起降维作用,降维系数 r 是一个超参数,为了找到合适的降维系数,以及在降维过程中避免信息丢失,本文将超参数取值 2、4、8、16 进行试验对比,结果如表 3 所示。由表 3 可知,当 r 取值 8 时,模型准确率最高,在提高模型泛化能力同时可以加快模型的运算速度,因此本文 SE 模块选择 $r = 8$ 作为超参数。

表 3 不同超参数取值的准确率

Tab.3 Hyperparameters of running results of different algorithms

算法超参数 r	平均交并比/%	像素准确率/%
2	84.35	95.44
4	84.33	95.48
8	84.55	95.61
16	84.46	95.48

为了更直观地展示网络模型改进效果,选择表 2 中的试验 1、3、4、5 共 4 组网络模型,分别代表 U-Net 网络,以及 U-Net 编码阶段替换为 ResNet34 后依次添加 MADC Block、DFPN Block、FUC Block 结构,选取不同环境下的田间图像进行预测,以 Labelme 工具标注得到的二值图像作为真值,结果如图 12 所示。

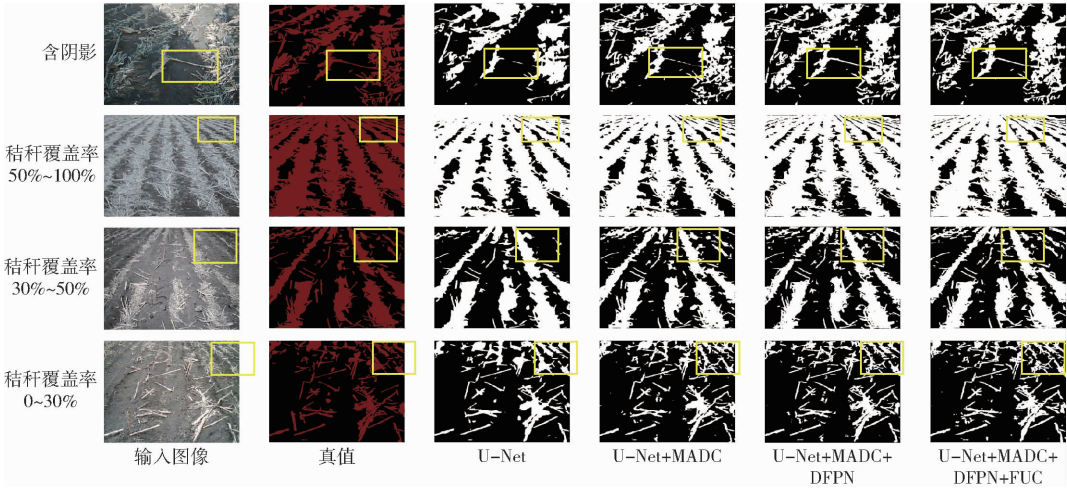


图 12 不同覆盖率下网络预测效果对比

Fig.12 Comparison of network prediction effects in different situations

由图 12 可知,在有阴影干扰情况下,原始 U-Net 网络的识别不够精细,阴影区域的秸秆识别不够完整,但是改进后的 U-Net 网络能够更精确地识别出秸秆位置,在一定程度上解决由阴影所带来的干扰。在秸秆覆盖率 50% ~ 100% 时,改进的 U-Net 网络和原始 U-Net 识别结果相差不大,主要区别在秸秆和土壤的边界处分割更加精细。在秸秆覆盖率处于 30% ~ 50% 时,改进网络对远处离散的秸秆区域识别的更加准确。当秸秆覆盖率为 0 ~ 30% 时,秸秆主要以细碎分散的秸秆为主,改进网络对于细碎秸秆的分割更接近真实情况,对秸秆覆盖区域中夹杂的土壤区域也能分离出来。由以上分析可以看出,本文改进的算法在大部分情况下都能比原始 U-Net 更好地完成分割任务。

4.3 与其他网络效果对比

本文通过和其他算法运行结果相比来评估本算法的性能。使用卷积核为 3 × 3 的 U-Net 算法、加入注意力机制的 Attention-U-Net 算法、CE-Net^[22] 算法、CPF-Net^[37] 算法进行了对比,结果如表 4 所示。

表 4 不同算法运行结果比较			
Tab.4 Comparison of running results of different algorithms			
算法	平均交并比/%	像素准确率/%	处理时间/ms
U-Net	82.19	94.65	1.40
Attention-U-Net	82.00	94.72	1.66
CE-Net	83.04	95.09	1.77
CPF-Net	83.69	95.20	2.09
Improved U-Net	84.78	95.64	2.53

由表 4 可知,本文算法 MIOU 和 PA 指标高于其他算法,在细碎秸秆识别上更加精细,网络通过增加非对称空洞卷积和特征图金字塔网络结构,增强了网络的特征提取能力。由于网络结构中参数多,模型复杂度高,网络对图像处理时间更长,是相对于其他网络,但整个模型对 640 像素 × 480 像素 的图像平均处理时间在 3 ms 以下,满足实际作业需求。

4.4 实际覆盖率计算对比

使用人工测量和本试验算法对 6 个不同田块秸秆的覆盖率进行计算,计算方式为识别结果图中秸秆像素点除以整个图像的像素点,对比结果如表 5 所示,误差率为真实覆盖率和预测覆盖率之间的差值绝对值。由表 5 可知,算法预测的覆盖率和人工测量的覆盖率相近,最大误差率为 3.8%。

试验另外随机选取验证集中的 136 幅秸秆图像

表 5 不同田块的覆盖率计算						
Tab.5 Calculation of coverage rate of different fields						
	田块 1	田块 2	田块 3	田块 4	田块 5	田块 6
人工估算	32.7	33.1	19.6	20.1	73.6	14.4
算法预测	32.3	36.3	22.5	23.9	75.8	17.2
误差率	0.4	3.2	2.9	3.8	2.2	2.8

数据,分别按覆盖率等级进行划分,划分为 0 ~ 10%、10% ~ 20%、20% ~ 30%、30% ~ 50%、50% ~ 100% 5 个集合,结果如图 13 所示。其中 0 ~ 10% 等级共 27 幅图像,10% ~ 20% 等级共 36 幅图像,20% ~ 30% 等级共 34 幅图像,30% ~ 50% 等级共 28 幅图像,50% ~ 100% 等级共 11 幅图像。根据结果可知,不同覆盖等级下,算法识别的准确率和人工标注测量的覆盖率大致相等,算法识别准确率的折线图在人工测量折线图上下波动,误差较小。

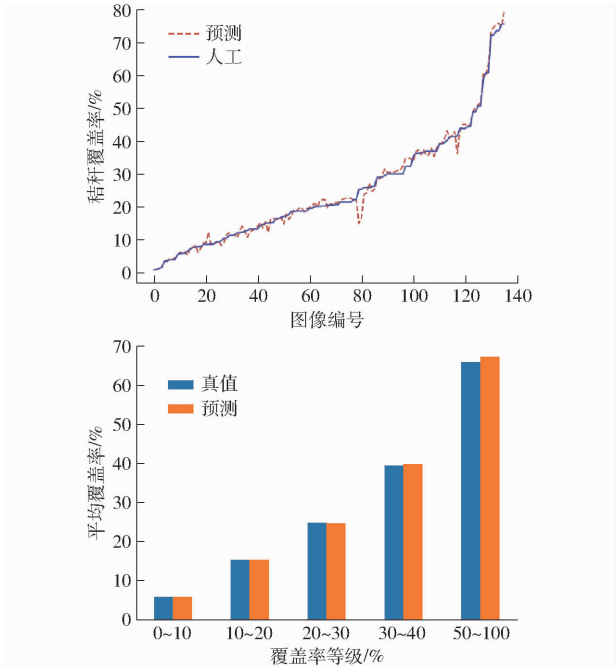


图 13 不同覆盖率等级效果对比
Fig. 13 Comparison of effects of different coverage levels

5 结论

- (1)将编码阶段网络结构替换为 ResNet34 结构,通过残差连接加深网络层数,避免网络训练中梯度消失问题,增加网络复杂度,提升秸秆特征提取能力。
- (2)在网络最高语义信息层使用多分支的非对称空洞卷积,空洞卷积提高卷积核的感受野,非对称结构增强了秸秆局部关键点的权重,提高了模型鲁棒性。
- (3)结合 FPN 网络,提出密集特征金字塔网络,结合高层特征图和底层特征图信息,对不同尺寸秸

秆的识别能力更强。

(4)使用快速上卷积模块替换普通上采样方式,避免网络训练过程中的无效参数运算。

(5)提出的改进 U-Net,结合改进的 FPN 网络等结构,提高了网络的分割精度,分割 MIOU 为 84.78%,相较于原 U-Net 提高 2.59 个百分点,平均处理时间低于 3 ms,符合农机正常作业时秸秆覆盖检测的要求。

参 考 文 献

[1] 孟凡杰,于晓芳,高聚林,等. 黑土地保护性耕作发展的制约瓶颈和突破路径[J]. 农业经济问题,2020(2):135-142.
MENG Fanjie, YU Xiaofang, GAO Julin, et al. Constraint bottleneck and breakthrough path of conservation tillage development in black soil[J]. Issues in Agricultural Economics, 2020(2):135-142. (in Chinese)

[2] 张妍茹,王欣然,徐继东. 黑土地保护利用有机肥施用-秸秆深埋-覆盖还田技术模式分析[J]. 农民致富之友,2021(4):221.
ZHANG Yanru, WANG Xinran, XU Jidong. Analysis of technical model of black soil conservation utilization of organic fertilizer application, straw deep burial-mulching and returning to the field[J]. Friends of Farmers Get Rich, 2021(4):221. (in Chinese)

[3] 贺美. 秸秆还田对黑土有机质变化的影响效应[D]. 北京:中国农业科学院,2016.
HE Mei. Effects of straw return on changes in organic matter in black soil[D]. Beijing: Chinese Academy of Agricultural Sciences, 2016. (in Chinese)

[4] 刘媛媛,孙嘉慧,张书杰,等. 用多阈值多目标无人机图像分割优化算法检测秸秆覆盖率[J]. 农业工程学报,2020,36(20):134-143.
LIU Yuanyuan, SUN Jiahui, ZHANG Shujie, et al. Detection of straw coverage using multi-threshold multi-target UAV image segmentation optimization algorithm[J]. Transactions of the CSAE, 2020, 36(20):134-143. (in Chinese)

[5] 李洪文,李慧,何进,等. 基于人工神经网络的田间秸秆覆盖率检测系统[J]. 农业机械学报,2009,40(6):58-62.
LI Hongwen, LI Hui, HE Jin, et al. Field straw coverage detection system based on artificial neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(6):58-62. (in Chinese)

[6] PETER R N, JOHANN P, THOMAS B, et al. A machine learning approach for pixel wise classification of residue and vegetation cover under field conditions[J]. Biosystems Engineering, 2018, 169:188-198.

[7] 李佳,吕程序,苑严伟,等. 快速傅里叶变换结合 SVM 算法识别地表玉米秸秆覆盖率[J]. 农业工程学报,2019,35(20):194-201.
LI Jia, LÜ Chengxu, YUAN Yanwei, et al. Fast Fourier transform combined with SVM algorithm to identify surface maize straw coverage[J]. Transactions of the CSAE, 2019, 35(20):194-201. (in Chinese)

[8] 王丽丽,徐岚俊,魏舒,等. 基于 Sauvola 与 Otsu 算法的秸秆覆盖率图像检测方法[J]. 农业工程,2017,7(4):29-35.
WANG Lili, XU Lanjun, WEI Shu, et al. Image detection method of straw coverage based on Sauvola and Otsu algorithm[J]. Agricultural Engineering, 2017, 7(4):29-35. (in Chinese)

[9] 安晓飞,王培,罗长海,等. 基于 K-means 聚类 and 分区寻优的秸秆覆盖率计算方法[J]. 农业机械学报,2021,52(10):84-89.
AN Xiaofei, WANG Pei, LUO Changhai, et al. Straw coverage calculation method based on K-means clustering and zoning optimization[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10):84-89. (in Chinese)

[10] 刘现悦,刘卉,王侨,等. 农田秸秆覆盖率检测方法与试验[J]. 中国农机化学报,2021,42(2):203-208.
LIU Xian Yue, LIU Hui, WANG Qiao, et al. Method and test of straw coverage detection in farmland[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(2):203-208. (in Chinese)

[11] 张静,靳淇兆,王洪振,等. 多尺度信息融合的遥感图像语义分割模型[J]. 计算机辅助设计与图形学学报,2019,31(9):1509-1517.
ZHANG Jing, JIN Qizhao, WANG Hongzhen, et al. Semantic segmentation model of remote sensing image based on multi-scale information fusion[J]. Journal of Computer-Aided Design and Graphics, 2019, 31(9):1509-1517. (in Chinese)

[12] 杨建宇,周振旭,杜贞容,等. 基于 SegNet 语义模型的高分辨率遥感影像农村建设用地提取[J]. 农业工程学报,2019,35(5):251-258.
YANG Jianyu, ZHOU Zhenxu, DU Zhenrong, et al. Extraction of rural construction land from high-resolution remote sensing images based on SegNet semantic model[J]. Transactions of the CSAE, 2019, 35(5):251-258. (in Chinese)

[13] 罗鹏飞. 基于自动驾驶城市场景的语义分割研究[D]. 武汉:武汉大学,2019.
LUO Pengfei. Research on semantic segmentation based on autonomous driving urban scenario [D]. Wuhan: Wuhan University, 2019. (in Chinese)

[14] 陈必科. 车道线检测和语义分割模型在自动驾驶中的应用[D]. 南京:南京理工大学,2017.
CHEN Bike. Application of lane line detection and semantic segmentation model in autonomous driving[D]. Nanjing: Nanjing University of Science and Technology, 2017. (in Chinese)

[15] 吴玉超,林岚,王婧璇. 基于卷积神经网络的语义分割在医学图像中的应用[J]. 生物医学工程学杂志,2020,37(3):533-540.
WU Yuchao, LIN Lan, WANG Jingxuan. Application of semantic segmentation based on convolutional neural network in medical images[J]. Chinese Journal of Biomedical Engineering, 2020, 37(3):533-540. (in Chinese)

[16] CAI Wenting, ZHAO Shuhe, WANG Yamei, et al. Estimation of winter wheat residue coverage using optical and SAR remote

sensing images[J]. Remote Sensing,2019,11(10):1163.

- [17] MUHAMMAD S M,ZHOU Jun, SUN Chuanliang, et al. Assessment of wheat straw cover and yield performance in a rice-wheat cropping system by using Landsat satellite data[J]. Sustainability,2019,11(19):5369.
- [18] ZHOU Deyi,LI Mao,LI Yang,et al. Detection of ground straw coverage under conservation tillage based on deep learning[J]. Computers and Electronics in Agriculture,2020,172:105369.
- [19] 刘媛媛,张硕,于海业,等. 基于语义分割的复杂场景下的秸秆检测[J]. 光学精密工程,2020,28(1):200-211.
LIU Yuanyuan,ZHANG Shuo,YU Haiye, et al. Straw detection in complex scenarios based on semantic segmentation[J]. Optical Precision Engineering,2020,28(1):200-211. (in Chinese)
- [20] 刘媛媛,王跃勇,于海业,等. 基于多阈值图像分割算法的秸秆覆盖率检测[J]. 农业机械学报,2018,49(12):27-35,55.
LIU Yuanyuan,WANG Yueyong,YU Haiye, et al. Straw coverage detection based on multi-threshold image segmentation algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(12):27-35,55. (in Chinese)
- [21] 蔡英凤,张田田,王海,等. 基于实例分割和自适应透视变换算法的多车道线检测[J]. 东南大学学报(自然科学版),2020,50(4):775-781.
CAI Yingfeng,ZHANG Tiantian,WANG Hai, et al. Multi-lane line detection based on instance segmentation and adaptive perspective transformation algorithm[J]. Journal of Southeast University (Natural Science Edition),2020,50(4):775-781. (in Chinese)
- [22] GU Z, CHENG J, FU H, et al. CE-Net: context encoder network for 2D medical image segmentation[J]. IEEE Transactions on Medical Imaging,2019,38(10):2281-2292.
- [23] SHELHAMER E, LONG J, DARRELL T, et al. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):640-651.
- [24] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention—MICCAI 2015. Springer, Cham, 2015:234-241.
- [25] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017,39(12):2481-2495.
- [26] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]//ICLR, 2016.
- [27] YANG J, ZHU J, WANG H, et al. Dilated MultiResUNet: dilated multi-residual blocks network based on U-Net for biomedical image segmentation[J]. Biomedical Signal Processing and Control,2021,68:1-13.
- [28] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:1-19.
- [29] SURAYYA A B,SHRI K. Dense dilated inception network for medical image segmentation[J]. International Journal of Advanced Computer Science and Applications (IJACSA),2020,11(11):785-793.
- [30] 张杨忆,林泓,管钰华,等. 改进残差块和对抗损失的 GAN 图像超分辨率重建[J]. 哈尔滨工业大学学报,2019,51(11):128-137.
ZHANG Yangyi, LIN Hong, GUAN Yuhua, et al. Super-resolution reconstruction of GAN images with improved residual blocks and countermeasure loss[J]. Journal of Harbin Institute of Technology,2019,51(11):128-137. (in Chinese)
- [31] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770-778.
- [32] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.
- [33] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(8):1-13.
- [34] LAINA I, RUPPRECHT C, BELAGIANNIS V, et al. Deeper depth prediction with fully convolutional residual networks[C]//2016 Fourth International Conference on 3D Vision (3DV),2016:239-248.
- [35] HANG Z, ORAZIO G, IURI F, et al. Loss functions for image restoration with neural networks[J]. IEEE Transactions on Computational Imaging,2017,3(1):47-57.
- [36] BOSMAN A S, ENGELBRECHT A, HELBIG M. Visualising basins of attraction for the cross-entropy and the squared error neural network loss functions[J]. Neurocomputing, 2020,400(4):113-136.
- [37] FENG S, ZHAO H, SHI F, et al. CPFNet: context pyramid fusion network for medical image segmentation[J]. IEEE Transactions on Medical Imaging,2020(10):3008-3018.