

基于语义分割的复杂场景下的秸秆检测

刘媛媛, 张硕, 于海业, 王跃勇, 王佳木

引用本文:

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 Yuan-yuan, ZHANG Shuo, YU Hai-ye, et al. Straw detection algorithm based on semantic segmentation in complex farm scenarios[J]. *Optics and Precision Engineering*, 2020, 28(1): 200–211.

在线阅读 View online: <https://doi.org/10.3788/OPE.20202801.0200>

您可能感兴趣的其他文章

Articles you may be interested in

利用卷积神经网络的自动驾驶场景语义分割

Autonomous driving semantic segmentation with convolution neural networks

光学精密工程. 2019, 27(11): 2429–2438 <https://doi.org/10.3788/OPE.20192711.2429>

基于DeepLabV3+与超像素优化的语义分割

Semantic segmentation based on DeepLabV3+and superpixel optimization

光学精密工程. 2019, 27(12): 2722–2729 <https://doi.org/10.3788/OPE.20192712.2722>

基于谱残差视觉显著性的带钢表面缺陷检测

Surface defect detection of steel strip based on spectral residual visual saliency

光学精密工程. 2016, 34(10): 2572–2580 <https://doi.org/10.3788/OPE.20162410.2572>

基于改进YOLOv2的快速安全帽佩戴情况检测

Fast helmet-wearing-condition detection based on improved YOLOv2

光学精密工程. 2019, 27(5): 1196–1205 <https://doi.org/10.3788/OPE.20192705.1196>

基于快速递推模糊2-划分熵图割的红外图像分割

Infrared image segmentation based on graph cut of fast recursive fuzzy 2-partition entropy

光学精密工程. 2016, 24(3): 668–680 <https://doi.org/10.3788/OPE.20162403.0668>

文章编号 1004-924X(2020)01-0200-12

基于语义分割的复杂场景下的秸秆检测

刘媛媛¹, 张 硕¹, 于海业², 王跃勇^{3*}, 王佳木¹

(1. 吉林农业大学 信息技术学院, 吉林 长春 130118;

2. 吉林大学 工程仿生教育部重点实验室, 吉林 长春 130025;

3. 吉林农业大学 工程技术学院, 吉林 长春 130118)

摘要: 基于阈值或纹理分割的秸秆覆盖率检测算法, 存在准确性低、复杂度高、运行耗时长等问题, 且对含有大量干扰因素的复杂农田场景分割效果不佳。本文提出了一种检测准确度高、训练参数少且运行速度快的语义分割算法(DSRA-UNet)。该算法结合 UNet 的对称编-解码架构, 在浅层特征图使用标准卷积, 深层采用深度可分离卷积, 并在每一层增加残差结构来加大网络深度, 以降低参数数量的同时提高精度。此外, 在跳级连接过程增加全局最大池化注意力机制, 进一步提高网络的分割精度。将算法在秸秆数据集上进行验证, 实验结果表明本文所提算法平均交并比达到 94.3%, 训练参数量仅为 0.76 M, 单张图片测试时间在 0.05 s 以下。该算法可以精准分割出秸秆和土壤, 并可在复杂环境下将干扰信息分割出, 可在一定程度上解决图像中的阴影问题。

关键词: 秸秆检测; 语义分割; 深度可分离卷积; 注意力机制; 残差结构

中图分类号: S24; TP751 **文献标识码:** A **doi:** 10.3788/OPE.20202801.0200

Straw detection algorithm based on semantic segmentation in complex farm scenarios

LIU Yuan-yuan¹, ZHANG Shuo¹, YU Hai-ye², WANG Yue-yong^{3*}, WANG Jia-mu¹

(1. College of Information Technology, Jilin Agriculture University, Changchun 130118, China;

2. Key Laboratory of Bionic Engineering, Ministry of Education,
Jilin University, Changchun 130025, China;

3. College of Engineering and Technology, Jilin Agricultural University,
Changchun 130118, China)

* Corresponding author, E-mail: yueyong10@mails.jlu.edu.cn

Abstract: The traditional segmentation algorithms for straw coverage detection basing on thresholds or texture features were difficult to get rid of the disadvantages of low accuracy, high complexity and time-consuming, and the effect of segmentation on complex farmland scenes containing a lot of interference factors was not good. Therefore, this paper proposed a semantic segmentation algorithm (DSRA-UNet) with high accuracy, a small mount of training parameters and high running speed.

收稿日期: 2019-08-27; **修订日期:** 2019-10-08.

基金项目: 国家自然科学基金资助(No. 31801753); 吉林省科技厅重点科技项目资助(No. 20180201014NY); 吉林省教育厅科学技术项目资助(No. JJKH20190927KJ); 吉林省发改委创新资金项目资助(No. 2019C054); 吉林大学工程仿生教育部重点实验室开放基金项目资助(No. K201706)

Combined with UNet's symmetric codec architecture, this algorithm used standard convolution in shallow feature maps, and depthwise separable convolution in deep ones. Residual structure was built in each layer to increase the network depth, which can reduce the number of parameters and improve the accuracy at the same time. In addition, the global maximum pooling attention mechanism was added during the skip connection process to further improve the segmentation accuracy of the network. The algorithm was verified on the straw datasets, and the experiment results showed that the mean of intersection over union reached to 94.3% in the proposed algorithm of this paper. The number of training parameters of the algorithm was only 0.76 M, and the test time of single picture was less than 0.05 s. The algorithm could accurately segment the straw and soil, and separate the interference information in the complex environment, especially solving the shadow problem in image.

Key words: straw detection; semantic segmentation; depthwise separable convolution; attention mechanism; residual structure

1 引 言

秸秆还田作为当今世界范围内普遍应用的一项培肥地力的保护性耕作措施,在杜绝了秸秆焚烧所造成的大气污染的同时还有增肥增产的作用^[1-2],其对农业生产,生态环境有积极的影响。为了更好地限制秸秆焚烧,指导农民合理地进行秸秆还田,需要定量评估还田率。因此,田间秸秆覆盖率的检测就显得尤为重要。

为了准确、快速地计算田间秸秆覆盖率,利用近年来快速发展的计算机视觉技术来解决这个问题是非常有价值的。其具有成本低、效率高、经济效益大等特点。然而,该方法通常是由无人机搭载摄像头来完成的任务,这样就带来了诸多问题,如:由于田间环境一般较复杂,田地周围通常都会有植被、路、农机、房屋等干扰。尤其是周围的树木,其在烈日下投在田间的树影非常长。这些干扰因素均可能出现在无人机拍摄的图片中,导致处理难度增大,精度降低。为了解决复杂场景下秸秆检测难的问题,最关键的是要将这些物体分割出,这样可以大大降低干扰。

传统的分割方法以阈值分割最为常见,如文献[3]提出将 Sauvola 算法与 Otsu 算法相结合的方法来检测秸秆区域细节图像,但仅限于二分类,对多类别复杂田地检测效果差;文献[4]提出将纹理特征与 BP 神经网络相结合的方法来识别秸秆和土壤,但检测时间多达 10 s;因此,文献[5]提出基于多阈值图像分割算法的秸秆覆盖率检测,利

用 DE-GWO 算法对其进行多阈值图像分割,检测准确率达到 95%以上,并且单张图片测试时间缩短到 1 s 以内,但其算法稳定性较差。以上基于阈值的分割算法均存在同一个问题,即对于颜色相近的物体,会被归为一类。而基于纹理的算法,需要手动设计特征,大大增加了算法难度。此外,以上算法在面对复杂农田时效果均不好,尤其是由阴影引起的干扰因素。近年来,由于深度学习的发展,Hinton 等人在 2012 年 ILSVRC 比赛中,利用卷积神经网络将 ImageNet 图像分类的错误率大幅降低到 16%^[6]。之后卷积神经网络就被广泛地应用到图像识别^[7-8]、目标检测^[9-10]、图像分割^[11-12]等领域。其中利用全卷积网络解决语义分割问题由 Jonathan Long 等人提出^[13],开创了利用全卷积网络解决语义分割问题的先河。由于卷积神经网络可以通过对大量数据的学习,自动抽取每个类别的特征,所以对图像的分类效果显著,对于复杂场景下的分割效果要明显优于传统的基于纹理或阈值的分割。基于此,本文设计了一个新的分割网络来解决复杂场景下的秸秆覆盖率检测。整体架构采用了 UNet^[12] 对称编解码设计思想,设计了一个全残差网络^[14] (Residual Network, RN) 架构,并使用深度可分离卷积^[15] (Depthwise Separable Convolution, DSC) 降低参数量,和注意力机制^[16-19] (Attention Mechanism, AM) 来关注难分割的像素,使得网络在降低参数和深度较浅的情况下仍然可以达到较高的精度。所提算法在秸秆数据集上进行了测试,在不同的光照环境、摄像机拍摄高度、场景复杂度的

情况下,仍可以达到较好的分割效果。同时,该网络为端到端可训练,且参数量少,使得其模型体积小、运行速度快,大大降低了部署到嵌入式设备的难度,这非常有利于在农田这样电子设备较少的环境下进行实地实时检测。

2 模型框架

2.1 模型分析与概述

秸秆覆盖率检测基于计算机视觉技术,采用无人机高空拍摄的方式获取图片,然后将图片进行预处理后输入到卷积神经网络中,得到预测图。由于田间场景复杂,如图 1 所示,其中包括秸秆、土壤、公路、周边植被、农机、树影等干扰因素。其中田内若没有干扰,则田地上只含秸秆、草、土壤,由于草相对较少,且大部分被秸秆覆盖,因此将草和土壤归为一类。对于田内含有农机和树影等干扰的情况(从图中可以看出主要是树影影响范围较大),由于树影投在了秸秆上,导致该处难以分割,很容易将覆有树影的秸秆识别为土壤。为了准确分割出秸秆,本文利用卷积神经网络来提取秸秆特征,并提出了一种新的分割网络 DSRA-UNet。该网络借鉴 UNet 的对称编-解码语义分割模型,其中低水平特征层使用标准卷积,高水平特征层使用深度可分离卷积,两者相结合的方式来构造整个网络架构,以降低训练参数量。同时在每一层的两次卷积过程加入残差结构来加大网络深度和特征表达能力。最后在跳级连接(Skip Connection, SC)之前增加注意力机制获取更精准的信息。通过以上操作在降低网络参数和深度的情况下提升对秸秆的分割能力,以适应复杂的农田场景。

2.2 DSRA-UNet 算法

本文提出的网络架构如图 2 所示,其中特征图上方标注了该特征图的深度,圆圈标号为每一层特征提取模块的编号,虚线框内标注了每个符号的含义。下采样过程包含 5 次特征提取,上采样过程包含 4 次特征提取,下采样过程中的第 1~4 层与上采样中第 6~9 层通过跳级连接相连。编码阶段第 1 层使用标准卷积构造的残差块(Conv Residual Blocks, CRB)进行特征提取,第 2

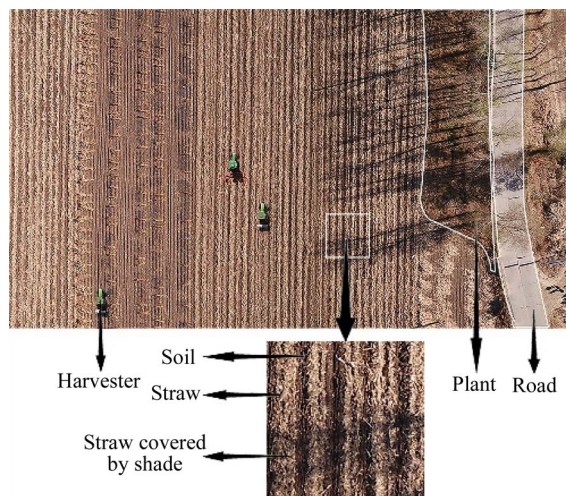


图 1 复杂农田场景示意图

Fig. 1 Schematic diagram of complex field scene

~5 层使用最大池化和深度可分离卷积构造的残差块(DSConv Residual Blocks)进行下采样特征提取,得到水平尺寸缩小 16 倍,深度扩大 16 倍的特征图。解码阶段在跳级连接之前加入了全局最大池化注意力机制(Global MaxPooling Attention Mechanism, GMPAM)来增加低层次特征图的语义信息,之后再经过一个 1×1 的卷积,并与经过反卷积的上一层高层次特征图进行拼接(Concatenate),以此结合同一尺度的下采样和上采样信息。拼接完成后经过深度分离卷积残差块进行特征融合,在第四次上采样后,将拼接的特征图经过标准卷积残差块,再通过 1×1 卷积和 softmax 函数得到与输入图片相同尺寸的预测图。

2.3 标准卷积残差块

残差网络不仅可以解决在网络较深时导致的梯度消失和梯度爆炸问题,而且可以通过捷径路径将输入信息融合进来,使得网络更深,这样就可以在网络深度较浅的情况下加深网络,以增强特征表达能力。此模块将标准卷积与残差结构相结合,使用 3×3 卷积核、ReLU 激活函数构造残差块。主要用于高分辨率特征图的语义信息提取,因为高分辨率图一般深度较浅,大量信息存储在空间维度,而标准卷积相较于深度卷积可以更好地提取空间维度的信息。因此,本文将其用在特征图尺寸大、深度较浅的第 1 层和最后 1 层特征提取中。

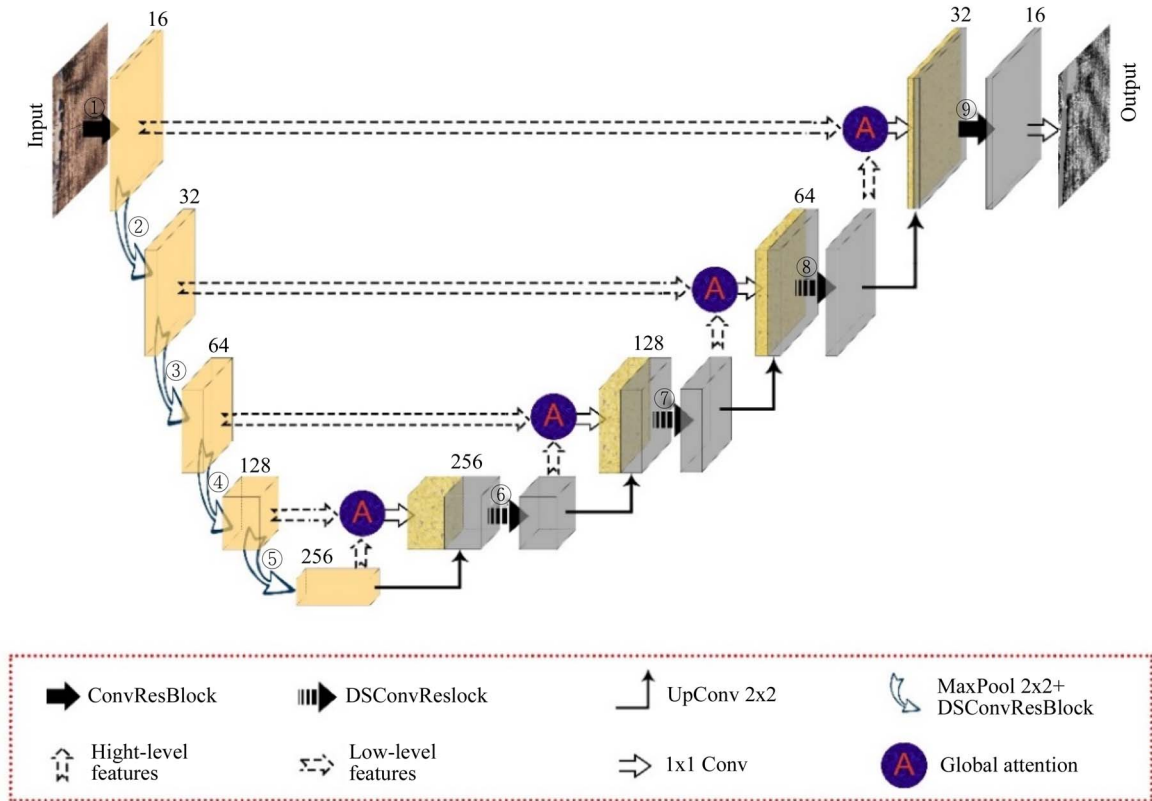


图 2 DSRA-UNet 算法网络结构

Fig. 2 Network architecture of DSRA-UNet algorithm

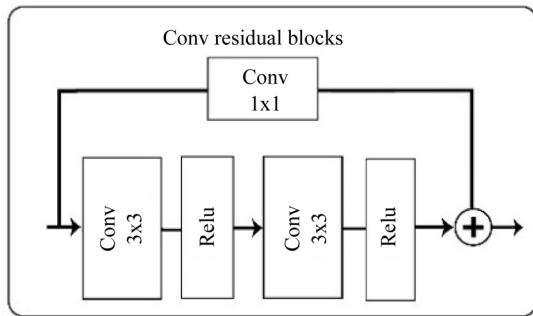


图 3 标准卷积残差块

Fig. 3 Normal convolution residual blocks

2.4 深度可分离卷积残差块

深度可分离卷积可看作先进行深度卷积 (Depthwise Convolution, DC), 再进行逐点卷积 (Pointwise Convolution, PC)。与标准卷积操作不同, 深度卷积的一个卷积核只负责一个通道, 得到与输入层通道数相同的特征图, 无法改变特征图深度。因此, 再使用逐点卷积即 1×1 卷积, 进行升维或降维。通过将标准卷积分解为两步可大大降低训练参数, 但也导致由于训练参数不足精

度下降的问题, 因此本文使用 5×5 的卷积核以抽取更多的特征信息, 同时增加了残差结构, 增强特征提取能力。由于深度卷积是对不同的深度进行单独抽取特征, 因此其获取深度信息的能力更高, 为了高效使用深度可分离卷积, 本文将其用在第 2~8 层的高层次特征图特征提取阶段。

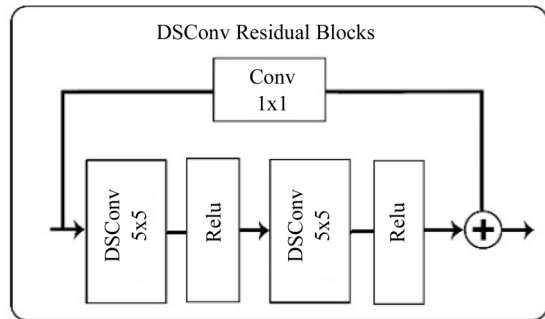


图 4 深度可分离卷积残差块

Fig. 4 Depthwise separable convolution residual blocks

2.5 全局最大池化注意力机制

注意力机制在自然语言处理、图像分类等领域已经有广泛应用^[16-17],近年来也被应用于语义分割并取得了不错的效果^[18]。受文献^[19]中所提注意力上采样模块启发,本文提出的注意力机制如图 5 所示,由于低层次特征图(Low-level Features)包含更多的位置信息,而高层次特征图(High-level Features)则包含丰富的类别信息,因此为了获得更精准的特征图,可以通过对低层次特征加权更丰富的类别信息而得到。所以本文利用高层次特征图中丰富的语义信息为低层次特征图进行特征选择,使低层次特征图增加更多细节信息。由于特征图在经过最大池化后可以获取当前感受野下更细节的信息,因此将高层次特征图经过全局最大池化得到 $1 \times 1 \times N$ 的特征向量以提供全局上下文信息,再经过 1×1 卷积、ReLU 降维至低水平特征图维度,然后与经过 1×1 卷积操作的低层次特征进行相乘得到加权的特征图。经过权值加权后可以对特征图起到选取的作用,从而使特征图获得更丰富的类别信息和更精准的位置信息。

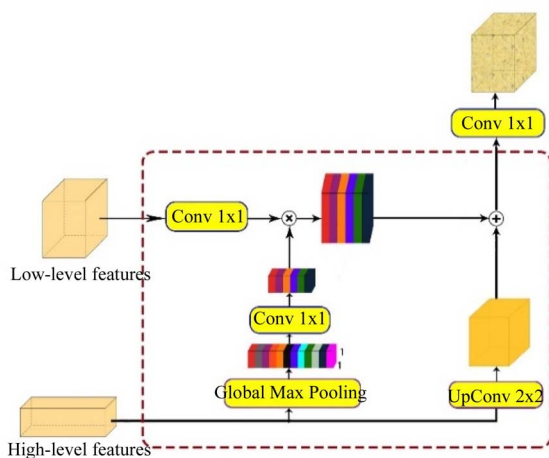


图 5 全局最大池化注意力机制

Fig. 5 Global maxpooling attention mechanism

3 训练

3.1 数据预处理

本文数据集由 DJI 大疆悟 2 代无人机于 2018 年 10 月在吉林省榆树市大岗乡拍摄。由于秸秆只有在秋季收割之后才会产生,因此样本数

量非常少,有效数据仅 120 张,其中 100 张是由 2 min 视频按每 1.2 s 截取下来的,但这 100 张图片仅仅是单一的秸秆田地图像,缺少农机、路、植物、树影等干扰因素。只有剩下 20 张含有这些复杂的干扰因素,如图 6 所示。因此需要进行数据增强,提高网络的鲁棒性,增强网络对复杂场景的分割。由于大部分图片不包含干扰信息,所以需要在 100 张仅含土壤和秸秆的图像基础上,通过增加干扰信息的方式来合成新的图片。首先合成含有路、植物、房子等干扰因素的图片,主要方法是将 20 张等有干扰信息的图片中的路、植物、农机、房子等裁剪出,然后将其加到不含干扰信息的图片上并做一些旋转缩放等操作,即可产生包含多类别的图片。对于树影,由于其数据集很少且树影有透明度,很难从原图中剪裁出进行图像合成。因此,本文利用 photoshop 工具生成了大量树影,再将其与 100 张基础图片进行合并。通过以上方法,得到了 500 张合成图,利用这 500 张合成图与 100 张仅含秸秆土地的原图像做为数据集。

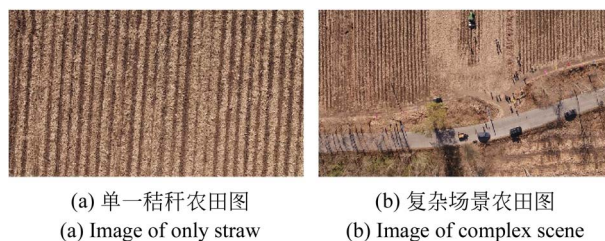


图 6 秸秆农田图的有效数据

Fig. 6 Effective data of image of straw scene

3.2 损失函数

多分类常使用交叉熵损失函数,但该损失函数在面对大量简单样本和少量困难样本时,由于交叉熵会更多的关注简单样本,而少量困难样本对模型的贡献较低,导致模型对难分样本的预测能力降低。因此本文使用改进的交叉熵^[20](Focal Loss),以使网络更多地关注较难分类的像素点,同时引入平衡因子 α ,让网络更加专注于秸秆的分割,公式定义如式(1)所示:

$$L = - \sum_{i=1}^n \alpha_i y_i (1 - \hat{y}_i)^{\gamma} \lg \hat{y}_i, \quad (1)$$

式中: L 为损失函数值; n 为类别数; y_i 为第 i 个

类别真实值,其中 $\sum_{i=1}^n y_i = 1$; \hat{y}_i 为第 i 个类别预测值,其中 $\sum_{i=1}^n \hat{y}_i = 1$; γ 取 2; α_i 为第 i 个类别权重。由于土地、农机、公路、植被、房子 5 个干扰类别在图像中所占比例相似,因此尝试 α 取相同值,而秸秆是该算法的关注对象,因此尝试 α 取干扰类别 α 值的两倍,经过实验测试干扰类别 α 取 0.15,秸秆 α 取 0.3 时效果最佳。通过该平衡后可以让网络更多地关注秸秆像素信息,并对被阴影遮挡的秸秆产生更好的分割效果。

3.3 训练超参数

(1) 优化器: Adam

(2) 学习率 (learning-rate): 初始学习率为 0.001, 衰减系数为 0.5, 最小为 $1e-6$ 。

(3) 批量大小 (batch size): 2

(4) 训练迭代周期 (epochs): 100

(5) 每周期步数 (steps per epoch): 600

4 实验与结果讨论

将上述算法在秸秆数据上进行定性及定量评估,所提算法在 Ubuntu18+python3.6+tensorflow10+Ker-as2 的环境下训练及测试,使用 1 块 Nvidia GTX 1080Ti 显卡进行训练,训练时间大约需要 3 h。

4.1 评价指标

为了更好地评价秸秆的覆盖准确率,本文使用秸秆覆盖率 R_{SCP} 、秸秆平均覆盖率误差 E_{SCR} 和秸秆交并比 S_{IoU} 等来衡量算法性能。

4.1.1 秸秆覆盖率误差

预测秸秆覆盖率为:

$$R_{SCP} = \frac{P_p}{H \times W}. \quad (2)$$

实际秸秆覆盖率为:

$$R_{SCT} = \frac{P_t}{H \times W}. \quad (3)$$

由式(2)和式(3)得秸秆覆盖率误差:

$$E_{SCR} = \frac{P_p - P_t}{H \times W}, \quad (4)$$

式中: R_{SCT} 为实际秸秆覆盖率; H 为图片高度; W 为图片宽度; P_p 为预测的秸秆像素点个数; P_t 为真实的秸秆像素点个数。

4.1.2 秸秆交并比

$$S_{IoU} = \frac{P_i}{P_p + p_i - P_i}, \quad (5)$$

式中: P_i 为相交的像素点个数; P_p 为预测的秸秆像素点个数; P_t 为真实的秸秆像素点个数。

4.2 不同算法效果对比

为了验证本文所提算法的有效性,将该算法与其他算法进行比较,结果如表 1。由表可以看出 SegNet^[21] 与 Deeplab^[22] 在该数据集表现较差且它们的参数量均较高。UNet3 为全部使用 3×3 卷积核尺寸的 UNet, 由于其深度和卷积核尺寸均较小, 导致精度较低, 但也正因其网络简单, 所以 FPS 较高。UNet5 则采用 5×5 卷积核, 从而增大了特征提取能力, 使得 mIoU 提高到了 85.75%, 但参数量增加 2.5 倍左右。DE-GWO^[5] 为传统阈值分割算法, 其对含阴影等的复杂场景检测效果较差, mIoU 也较低。本文所提算法 mIoU 达到了 94.3%, 且参数量降低到 0.76M, FPS 也较高, 优于上述其他方法。

表 1 不同算法之间的比较

Tab. 1 Comparison of different algorithms

Algorithms	mIoU/%	Parameters/M	FPS
UNet3	80.21	1.9	28.6
SegNet	70.65	10	13.9
DeeplabV3	79.16	40.9	10
UNet5	85.75	5.0	18.5
DE-GWO	72.5	—	1.5
Ours	94.3	0.76	21

如图 8 为本文所提算法在不同场景下秸秆的检测结果, 以及同其他算法的对比效果。其中类别所对应的灰度级别如图 7 所示。图 8 从上到下

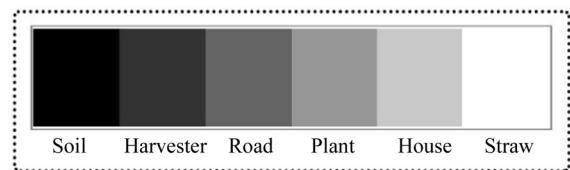


图 7 每个类别对应的灰度颜色

Fig. 7 Grayscale color corresponding to each category

依次为原图、真实值、及各种算法得出的预测结果,从左到右展示了不同场景下秸秆覆盖率检测效果。对于第 1 张最简单场景下秸秆的检测难度较低,但 SegNet 检测效果较差。第 2 张是含有树影的场景,由图可以看出 UNet3 对覆盖在秸秆上的树影检测效果较差,UNet5 则可以将部分秸秆

检测出,而本文提出的算法可以将大部分被阴影遮盖的秸秆分割出。对于样本 3 和样本 4,均是复杂农田场景,由图可以看出其他方法均出现了较大面积的错误,尤其是在路、植被和树影周围。而本文提出的算法,则可以更好地分割出这些干扰信息,进而提高了秸秆覆盖率检测精度。

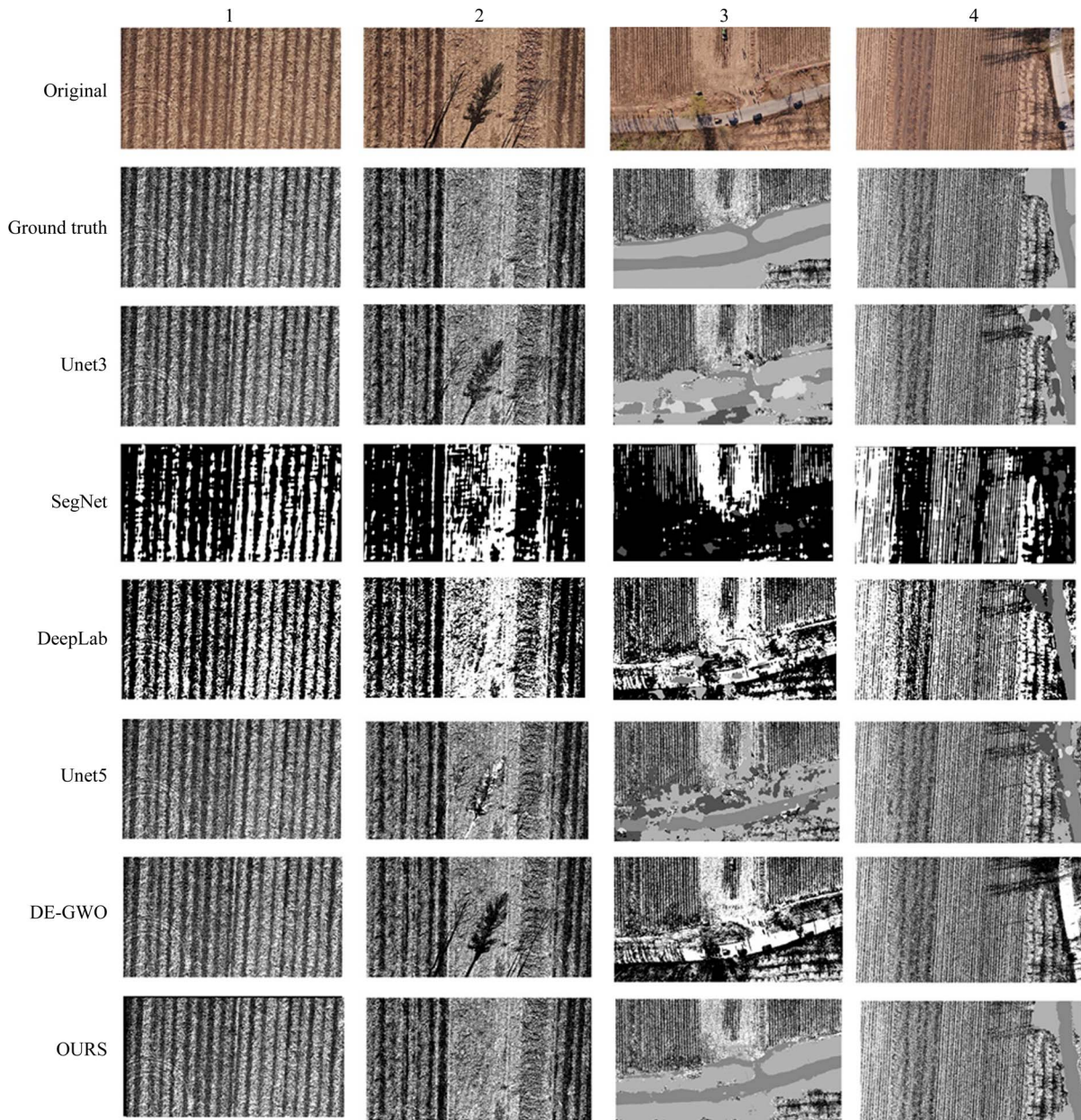


图 8 不同算法的实验结果

Fig. 8 Experimental results of different algorithms

4.3 基础架构对比

为了验证本文所提网络基础架构的有效性,设计了两组对比实验。第一组实验用来验证低层

次特征图使用标准卷积,高层次特征图使用深度可分离卷积的有效性。第二组实验用来验证用于低层次特征提取的标准卷积使用 3×3 的内核大

小,高层次特征提取的深度可分离卷积使用 5×5 内核大小的有效性。

4.3.1 卷积类型对比

由于高分辨率特征图(如输入的三通道图片,和最后一次上采样后得到的特征图),尺寸较大、深度较浅,其大部分信息存储在空间维度。因此本文提出,高分辨率低深度特征图使用标准卷积,低分辨率高深度特征图使用深度可分离卷积。实验结果如表 2,其中其他参数保持不变,仅改变卷积型。All-Conv 表示全部使用标准卷积,All-DSCConv 表示全部使用深度可分离卷积,Conv+DSCConv 表示浅层使用标准卷积,深层使用深度可分离卷积。由表可以看出当全部使用标准卷积时 mIoU 达到 93.4%,精度较高,但参数量是其他方法的 7 倍左右。当全部使用深度可分离卷积时 mIoU 为 91.73%,精度有所降低,但参数量仅为 0.75 M。当采用本文提出的结构时 mIoU 达到 94.3%,而参数量相较于上一种仅增加了 0.01 M 左右,但精度上升了 2.57%。由此可得,标准卷积可以更好地抽取空间信息,适合在高分辨率特征图中使用。而深度可分离卷积可以更好地抽取高深度特征图,因为深度卷积可以分别抽取每个通道的特征图,对深度信息的利用率高。因此本文所提结构可以充分利用空间信息和深度信息,从而达到参数量和精度的平衡。

表 2 不同卷积结果对比

Tab.2 Comparison of results from different convolution

Methods	mIoU/%	Parameters/M	PA/%
All-Conv	93.4	5.4	97.92
All-DSCConv	91.73	0.75	97.59
Conv+DSCConv	94.3	0.76	97.94

4.3.2 不同卷积核尺寸对比

由于标准卷积参数量大,深度可分离卷积参数量少,因此本文提出标准卷积采用 3×3 的卷积核尺寸,而深度可分离卷积采用 5×5 的卷积核尺寸,并保持其他结构不变,实验结果如表 3。

表 3 不同卷积核尺寸结果对比

Tab.3 Comparison of results from different convolution kernel sizes

Methods	mIoU /%	Parameters /M	PA /%
Conv3+DSCConv3	92.91	0.743	97.25
Conv5+DSCConv3	92.89	0.760	97.42
Conv5+DSCConv5	93.24	0.782	97.25
Conv3+DSCConv5	94.3	0.76	97.94

Conv 代表标准卷积,DSCConv 代表深度可分离卷积,3 和 5 分别代表 3×3 、 5×5 卷积核。由于均使用了标准卷积和深度可分离卷积相结合的方式,因此参数量变化较少。从实验结果可以看出 Conv 使用 5×5 卷积核的精度均低于使用 3×3 的,DSCConv 使用 3×3 卷积核的精度均低于使用 5×5 的。由此可得,在同时使用标准卷积和深度可分离卷积情况下, 3×3 卷积核可以更好地帮助标准卷积抽取更丰富的空间信息,而 5×5 卷积核可以更好地帮助深度可分离卷积抽取更丰富的深度信息,且参数量增加较少,从而使网络结构更加有效。

4.4 不同模块的对比

为了研究 DSRA-UNet 算法中各模块的作用,设计了 4 个不同的网络进行效果对比,实验结果如表 4。首先是骨干网络,由于中间层(即第二和第八层)使用 5×5 标准卷积,此部分参数量较高,mIoU 为 89.18%。第二部分是中间层改为深度可分离卷积,从而将模型参数大大降低,但 mIoU 仅降低了 0.49%。第三部分是在上采样和下采样过程中均加入残差结构,从表中可以看出引入该结构后精度有较大提升,mIoU 提高到了 92.13%。第四部分是引入了全局最大池化注意力机制,从表中可以看出引入该机制后精度有了进一步提升,并且可以很好地平衡参数量和精度,mIoU 达到了最高 94.3%,而参数量仅为 0.76 M。

表 4 不同模块结果对比

Tab. 4 Comparison of results from different modules

Methods	mIoU/%	Parameters/M
Bone	89.18	5.06
Bone+DS	88.69	0.415
Bone+DS+Res	92.13	0.416
Bone+DS+Res+Atten	94.3	0.76

4.5 实地测试结果

将本算法用于实地复杂场景农田的测试,得到了较好的实验结果。由于秸秆检测主要关注秸秆的覆盖率,因此本部分仅列出秸秆的检测结果(由于农机会对秸秆造成一定的干扰因此也将农机标出),以下给出四个不同场景下的秸秆覆盖率检测结果。

如图 9,其中秸秆像素值为 255,农机为 50,其他背景为 0。样本 1 虚线框中是大量被阴影遮挡的秸秆,可以看出其他算法均出现较大误判。样本 2 为田间包含农机和人以及他们的影子,从中可以看出,其他算法对这些干扰信息的分割能力较弱,影子部分通常会被识别为土地。而本文算法,不仅可以将农机分割出,同时也可以将小目标的人以及影子正确分类为秸秆及土壤。样本 3 为包含树影、植被、房屋、公路的复杂农田场景,使用 UNet 和 DSRA-UNet 算法均可将植被、房屋、公路分割出,且 DSRA-UNet 算法可以大大提高被树影遮挡部分的秸秆分类准确性。样本 4 则为包含了大量植被和公路的复杂农田,由结果可以看出本文所提算法对其分割效果最好,仅有少数分割错误。

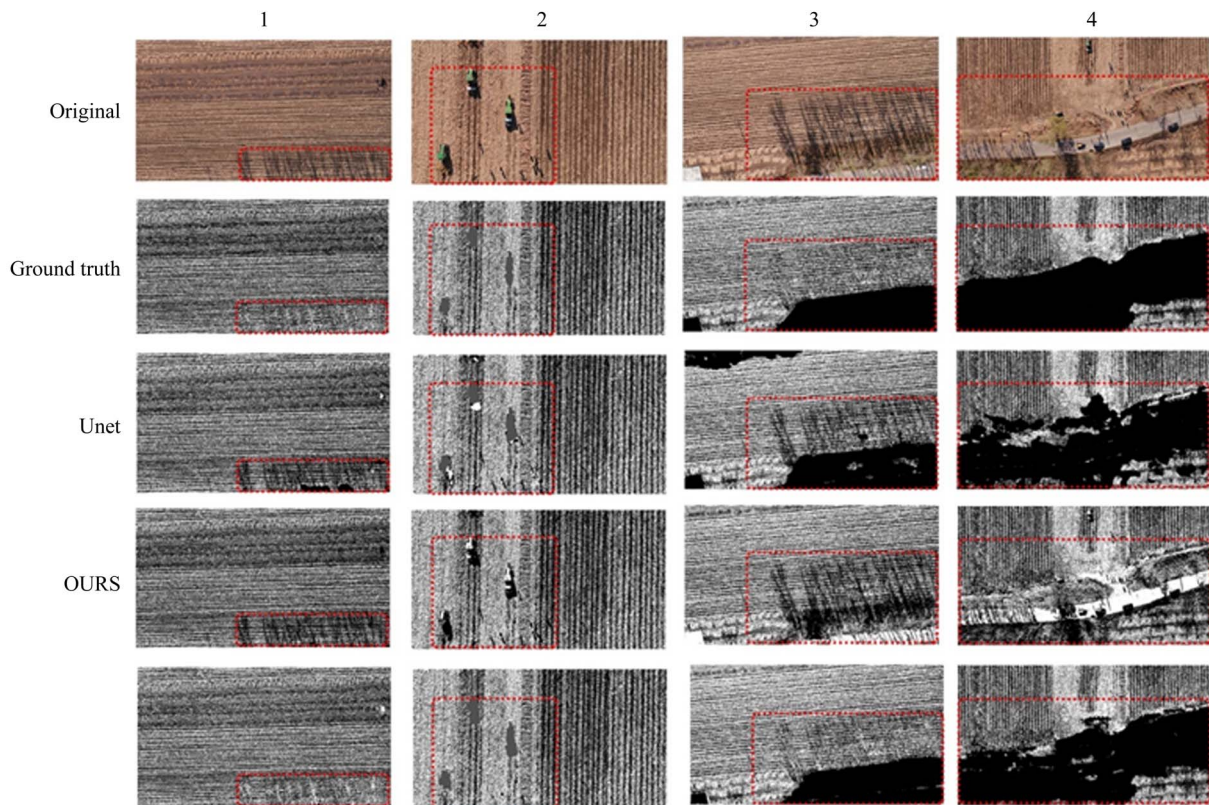


图 9 不同复杂度的农田检测结果

Fig. 9 Detection results from farm with different complexity

对上述四个复杂场景下的农田,利用三个不同的算法对其分割后得到的实验结果如表 5 所示,可以看出本文所提算法秸秆的 IoU 均为最

高。样本 1 由于改善了阴影部分的分割,因此秸秆 IoU 相较于 UNet 和传统阈值分别有 7.1% 和 11.1% 的性能提升,同时覆盖率误差也降低到了

1.3%。样本2由于本文算法可以很好的将农机分割出,并且消除人和人影的干扰,秸秆IoU达到了93.7%,秸秆覆盖率误差也达到了最低,其中经实验验证,分割出的农机所占面积的40%为秸秆。而传统方法无法将农机分离出,大部分都分割为秸秆,因此误差较大。样本3则为既包含

大量树影也包含公路、植被、房屋的复杂农田,本文所提算法对两者的处理均较好,秸秆IoU较UNet和传统阈值算法提升了9.7%和19.1%,覆盖率误差降低了一半。样本4为包含大量植被和公路的复杂农田,在这种大面积为干扰因素的情况。

表5 不同算法对秸秆检测的结果

Tab.5 Results of straw detection from different algorithms

Sample number	$S_{IoU}/\%$			$R_{SCR}/\%$			E_{SCR}			T/s		
	UNet	DE-GWO	OURS	UNet	DE-GWO	OURS	UNet	DE-GWO	OURS	UNet	DE-GWO	OURS
1	84.9	80.9	92.0	42.3	39.7	45.8	-0.022	-0.038	0.013	0.050	0.722 6	0.048
2	90.8	88.1	93.7	47.4	48.7	48.2	-0.006	0.007	0.002	0.053	0.728	0.048
3	81.1	71.7	90.8	43.4	50.2	48.3	-0.034	0.034	0.015	0.054	0.720 2	0.047
4	76.4	51.4	89.3	27.4	46.4	27.69	0.016	0.206	0.019	0.054	0.691 7	0.047

下本文所提算法秸秆IoU比UNet和传统阈值算法高出12.9%和39%,达到了较好的分割效果。综上所述,本文所提算法能很好地解决复杂场景下秸秆的检测和大量树影遮挡下的秸秆分割,且运行一张 960×528 尺寸的图像时间为0.048s,速度相较于传统算法有明显的优势,满足实际检测中对精度和速度两方面的要求。

5 结 论

本文针对传统算法无法在复杂场景的农田下高效分割秸秆的问题,提出使用语义分割的方法

来解决。通过对UNet的改进设计了一种新的网络架构DSRA-UNet网络,使用标准卷积残差块和深度可分离卷积残差块构建整个网络以降低参数量和um提高网络分割精度,并在跳级连接中融入全局最大池化注意力机制进一步提高网络性能。训练中采用带有平衡因子的焦点损失函数,使网络更加关注难分类别和提高秸秆对网络的贡献率。同时还设计了基础架构、不同模块等的对比实验,进一步验证了本文所提算法的有效性。将算法用于秸秆分割mIoU达到94.3%,参数量仅为0.76M,且运行速度达到每秒21帧,效果优于基于阈值或纹理的传统分割和其他语义分割网络。

参考文献:

- [1] 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 M. *Effects of Straw Returning Methods On Black Soil Organic Matter*[D]. Beijing: Chinese Academy of Agricultural Sciences, 2016. (in Chinese)
- [2] 宫纯文,李建玉,李学彬,等. 秸秆燃烧和沙尘对气溶胶光学特性的影响[J]. 光学精密工程, 2015, 23(10):2768-2777.
GONG CH W, LI J Y, LI X B, *et al.*. Influence of straw burning and dust on aerosol optical properties [J]. *Opt. Precision Eng.*, 2015, 23(10): 2768-2777. (in Chinese)
- [3] 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 L L, XU L J, WEI SH, *et al.*. Straw coverage detection method based on Sauvola and Otsu segmentation algorithm[J]. *Agricultural Engineering*, 2017, 7(4):29-35. (in Chinese)
- [4] 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 H W, LI H, HE J, *et al.*. Field coverage detection system of straw based on artificial neural net-

- work[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2009, 40(6):58-62. (in Chinese)
- [5] 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 Y Y, WANG Y Y, YU H Y, *et al.*. Detection of straw coverage rate based on multi-threshold image segmentation algorithm[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2018, 49(12):27-35,55. (in Chinese)
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6):84-90.
- [7] 潘仙张,张石清,郭文平. 多模深度卷积神经网络应用于视频表情识别[J]. *光学 精密工程*, 2019, 27(4): 963-970.
PAN X ZH, ZHANG SH Q, GUO W P. Video-based facial expression recognition using multimodal deep convolutional neural networks[J]. *Opt. Precision Eng.*, 2019, 27(4): 963-970. (in Chinese)
- [8] 郭保青,王宁. 基于改进深度卷积网络的铁路入侵行人分类算法[J]. *光学 精密工程*, 2018, 26(12): 3040-3050.
GUO B Q, WANG N. Pedestrian intruding railway clearance classification algorithm based on improved deep convolutional network [J]. *Opt. Precision Eng.*, 2018, 26(12): 3040-3050. (in Chinese)
- [9] 吕晓琪,吴凉,谷宇,等. 基于三维卷积神经网络的低剂量 CT 肺结节检测[J]. *光学 精密工程*, 2018, 26(5): 1211-1218.
L? X Q, WU L, GU Y, *et al.*. Detection of low dose CT pulmonary nodules based on 3D convolution neural network [J]. *Opt. Precision Eng.*, 2018, 26(5): 1211-1218. (in Chinese)
- [10] 潘蓉,孙伟. 基于预分割和回归的深度学习目标检测[J]. *光学 精密工程*, 2017, 25(10s): 221-227.
PAN R, SUN W. Deep learning target detection based on pre-segmentation and regression [J]. *Opt. Precision Eng.*, 2017, 25(10s): 221-227. (in Chinese)
- [11] 董月,冯华君,徐之海,等. Attention Res-Unet: 一种高效阴影检测算法[J]. *浙江大学学报(工学版)*, 2019, 53(2):373-381+406.
DONG Y, FENG H J, XU ZH H, *et al.*. Attention Res-Unet: an efficient shadow detection algorithm [J]. *Journal of Zhejiang University (Engineering Science)*, 2019, 53(2): 373-381+406. (in Chinese)
- [12] 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*. Cham: Springer, 2015: 234-241.
- [13] JONATHAN L, EVAN S, TREVOR D. Fully Convolutional Networks for Semantic Segmentation[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431- 3440.
- [14] 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 (CVPR)*, 2016: 770- 778.
- [15] FRANCOIS C. Xception: Deep Learning with Depthwise Separable Convolutions[C]. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017:1251-1258.
- [16] VASWANI A, SHAZEER N, PARMAR N, *et al.*. Attention is All You Need[C]. *Advances in neural information processing systems*, 2017: 5998-6008.
- [17] ZHAO B, WU X, FENG J, *et al.*. Diversified visual attention networks for fine-grained object classification[J]. *IEEE Transactions on Multimedia*, 2017, 19(6): 1245-1256.
- [18] JO Y J, CHO H, LEE S Y, *et al.*. Quantitative phase imaging and artificial intelligence: a review [J]. *IEEE Journal of Selected Topics in Quantum Electronics*, 2018, 25(1): 1-14.
- [19] LI H, XIONG P, AN J, *et al.*. Pyramid attention network for semantic segmentation[J]. *arXiv preprint arXiv:1805.10180*, 2018.
- [20] LIN T Y, GOYAL P, GIRSHICK R, *et al.*. Focal Loss for Dense Object Detection[C]. *Proceedings of the IEEE international conference on computer vision*, 2017: 2980-2988.
- [21] 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.
- [22] CHEN L C, ZHU Y, PAPANDREOU G, *et al.*

Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation [C]. *Pro-*

ceedings of the European conference on computer vision (ECCV), 2018: 801-818.

作者简介:

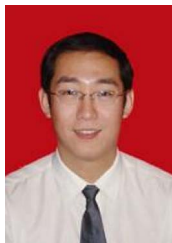


刘媛媛(1980—),女,吉林长春人,工学博士,讲师,硕士生导师,主要从事农业信息化图像和视频信号处理方面研究。E-mail:liuyuanyuan1980@126.com



于海业(1963—),男,吉林长春人,教授,博士生导师,1991年于吉林工业大学获得博士学位,现任吉林大学生物与农业工程学院院长,主要从事模式识别与智能系统、生物环境与能源工程研究。E-mail:haiye@jlu.edu.cn

通讯作者:



王跃勇(1977—),通信作者,男,吉林长春人,工学博士,副教授,硕士生导师,主要从事智能系统、生物环境与能源工程研究。E-mail: yueyong10@mails.jlu.edu.cn