doi:10.6041/j.issn.1000-1298.2020.09.018

基于田间图像的局部遮挡小尺寸稻穗检测和计数方法

姜海燕[1,2]　徐　灿[1]　陈　尧[1]　成永康[1]

(1.南京农业大学信息科技学院,南京 210095;2.南京农业大学国家信息农业工程技术中心,南京 210095)

摘要:大田水稻生长环境复杂,稻穗尺寸相对较小,且与叶片之间贴合并被遮挡严重,准确识别复杂田间场景中的水稻稻穗并自动统计穗数具有重要意义。为了实现对局部被叶片遮挡的小尺寸稻穗的计数,设计了一种基于生成特征金字塔的稻穗检测(Generative feature pyramid for panicle detection,GFP-PD)方法。首先,针对小尺寸稻穗在特征学习时的特征损失问题,量化分析稻穗尺寸与感受野大小的关系,通过选择合适的特征学习网络减少稻穗信息损失;其次,通过构造并融合多尺度特征金字塔来增强稻穗特征。针对稻穗特征中因叶片遮挡产生的噪声,基于生成对抗网络设计遮挡样品修复模块(Occlusion sample inpainting module,OSIM),将遮挡噪声修复为真实稻穗特征,优化遮挡稻穗的特征质量。对南粳 46 水稻的田间图像进行模型训练与测试,GFP-PD 方法对稻穗计数的平均查全率和识别正确率为 90.82% 和 99.05%,较 Faster R-CNN 算法计数结果分别提高了 16.69、5.15 个百分点。仅对 Faster R-CNN 算法构造特征金字塔,基于 VGG16 网络的平均查全率和识别正确率分别为 87.10% 和 93.87%,较 ZF 网络分别提高 3.75、1.20 个百分点;进一步使用 OSIM 修复模型、优化稻穗特征,识别正确率由 93.87% 上升为 99.05%。结果表明,选择适合特征学习网络和构建特征金字塔能够显著提高田间小尺寸稻穗的计数查全率;OSIM 能够有效去除稻穗特征中的叶片噪声,有利于提升局部被叶片遮挡的稻穗的识别正确率。

关键词:水稻;稻穗数;深度学习;对抗网络;特征金字塔

中图分类号:S126　　**文献标识码**:A　　**文章编号**:1000-1298(2020)09-0152-11　　**OSID**:

Detecting and Counting Method for Small-sized and Occluded Rice Panicles Based on In-field Images

JIANG Haiyan[1,2]　XU Can[1]　CHEN Yao[1]　CHENG Yongkang[1]

(1. *College of Information Science and Technology*, *Nanjing Agricultural University*, *Nanjing* 210095, *China*
2. *National Engineering and Technology Center for Information Agriculture*, *Nanjing Agricultural University*, *Nanjing* 210095, *China*)

Abstract: How to assess the number of rice panicles had been one of the key ways to realize high-throughput rice breeding in the modern smart farming, for that the panicle can reflect rice yield directly. In practical in-field scenarios of rice growing, the size of panicles was relatively small while the panicles were always occluded by the leaf seriously. So, it was a challenging task to accurately identify the rice panicle in the complex field scene and automatically count the number of panicles. In order to count the small-sized rice panicles locally occluded by leaves, an automatic counting method was designed which called generative feature pyramid for panicle detection (GFP-PD) based on the feature pyramid and the generative adversarial networks. To solve the problem of feature loss in feature learning of small size rice panicles, firstly, the relationship between the size of rice panicle and receptive field was analyzed quantitatively, and then the appropriate feature learning network was selected to reduce the information loss of rice panicles; secondly, the multi-scale feature pyramid was constructed and integrated to enhance the panicle features. For the noise in the panicle feature which caused by the leaves occlusion, a feature repairing network which called occlusion sample inpainting module (OSIM) was designed to optimize the quality of features containing leaves noise by restoring the noise to the real feature of rice panicles. The model was trained and tested by the in-field rice images taken from the variety of Nanjing 46. The average panicle counting accuracy and the average panicle recognition accuracy of GFP-PD were 90.82% and 99.05%, respectively, which were 16.69 percentage points and 5.15 percentage points higher than the

收稿日期:2019-11-26　修回日期:2020-01-08

基金项目:国家自然科学基金面上项目(31872847)和江苏省重点研发计划(现代农业)项目(BE2019383)

作者简介:姜海燕(1967—),女,教授,博士,主要从事农业人工智能与计算机视觉在表型检测上的应用研究,E-mail:jianghy@njau.edu.cn

results of Faster R－CNN. When constructing the feature pyramid for Faster R－CNN, the average counting accuracy and recognition accuracy based on VGG16 network were 87.10% and 93.87%, respectively, which were 3.75 percentage points and 1.20 percentage points higher than ZF network. After the OSIM repairing model was further used to optimize the panicle feature, the recognition accuracy was increased from 93.87% to 99.05%. The results showed that selecting the appropriate feature learning network and constructing the feature pyramid could significantly improve the count and recognition accuracy of small-size rice panicles in the field. The OSIM can remove the leaf noise in the feature of rice panicle effectively, which was useful to improving the recognition accuracy of the panicles partially covered by the rice leaves.

Key words: rice; panicle number; deep learning; adversarial networks; feature pyramid

0　引言

自动、准确地预估产量是水稻科学生产的重要环节。研究表明[1-2]，单位面积穗粒数是与产量最密切相关的农学性状。在水稻栽培或育种研究中，穗粒数一般通过结实率和穗数估算获得，因此，快速、准确地获取水稻穗数对智能测产意义重大。在实际应用中，水稻穗数的获取主要依靠人工统计，费时、费力。近年来，随着农业信息技术的发展，图像分析技术在植物叶片[3-6]和水果果实[7-9]的自动计数上取得成功，被认为是人工计数的有效替代方法。

田间水稻稻穗自动计数的关键是复杂场景下的稻穗识别。现有方法主要包括基于颜色或纹理的分割计数法和基于候选区域的分类计数法两类。基于颜色或纹理的分割计数法主要通过抽取穗的颜色特征[10-12]或纹理特征[13]，实现穗的分割并计数。ZHOU 等[14]融合小麦的颜色、纹理和边缘特征，并训练双支持向量机模型，分割小麦麦穗；FERNANDEZ-GALLEGO 等[15]使用滤波器去除土壤、叶片等无关噪声，计算局部峰值，实现麦穗的分割。穗分割方法计算简单、处理速度快，但是对光照敏感，分割效果和计数精度也容易受场景复杂度的影响。基于候选区域的分类计数法[16-18]首先生成穗的候选区域，然后对候选区域进行分类（穗或背景）并计数。常用的候选区域生成方法基于颜色或纹理特征，如 LI 等[19]基于 Laws 纹理能量生成穗部候选区域，训练 BP 神经网络，对候选区域分类并计算穗数。此外，一些研究尝试在生成候选区域时更充分利用物体的边界信息，将相似的像素点聚类生成超像素，进一步提高候选区域的精度。XIONG 等[20]使用简单线性迭代聚类生成超像素，为田间水稻图像生成稻穗候选区域；OLSEN 等[21]基于田间高粱图像生成不同尺度的超像素，并训练线性回归模型，实现计数。候选区域的生成是候选区域分类法的关键，当叶片贴近或粘连稻穗时，很难生成只包含穗或叶的高质量候选区域，含有噪声的候选区域影响分类器的分类性能，抑制最终的计数准确率。

最新的研究将卷积神经网络（Convolutional neural network，CNN）应用于田间禾本科作物的穗识别与计数。张领先等[22]利用卷积神经网络训练冬小麦识别模型，实现对麦穗、叶片和阴影的准确识别。ALKHUDAYDI 等[23]基于全卷积网络实现了像素水平的田间麦穗识别与计数方法；MADEC 等[24]使用基于快速区域的卷积神经网络（Faster regions with convolutional neural networks，Faster R－CNN）方法估算小麦的穗密度，并取得 91% 的计数准确率。田间水稻的稻穗尺寸小，在卷积神经网络抽取的深层特征中，有限的稻穗信息损失严重；稻穗因自重下垂，高种植密度下叶片与稻穗贴合更紧密，叶片对稻穗的局部遮挡使得稻穗特征中混入叶片噪声，进一步制约了稻穗识别准确率和计数精度。目前，针对遮挡稻穗自动计数的研究鲜有报道。

针对存在大量叶片遮挡的成熟期田间水稻稻穗，提出一种基于生成特征金字塔的稻穗检测（Generative feature pyramid for panicle detection，GFP－PD）方法。首先，针对小尺寸稻穗在特征学习时的特征损失问题，量化分析稻穗尺寸与感受野大小的关系，通过选择合适的特征学习网络减少稻穗信息损失；其次，通过构造并融合多尺度特征金字塔来增强稻穗特征。然后，针对稻穗特征中叶片遮挡带来的噪声，基于生成对抗网络设计遮挡修复模块 OSIM，优化遮挡稻穗的特征质量。

1　GFP－PD 稻穗计数算法设计

1.1　Faster R－CNN 算法

本研究选择 Faster R－CNN[25]作为基础框架，其网络结构如图 1 所示。Faster R－CNN 结合相关检测算法[26-28]的优点，将特征学习、候选区域生成、目标分类及位置框回归 4 个关键环节集成于一个端到端网络。对于输入的原始图像，Faster R－CNN 首先利用卷积神经网络提取特征图（Feature map）。候选区域生成网络（Region proposal network，RPN）复用最深卷积层上提取的特征图，使用尺寸为 3×3 的滑动窗

口遍历该特征图,以每个锚点(滑动窗口中心)为中心,设计不同的长宽比生成多尺度候选区域。

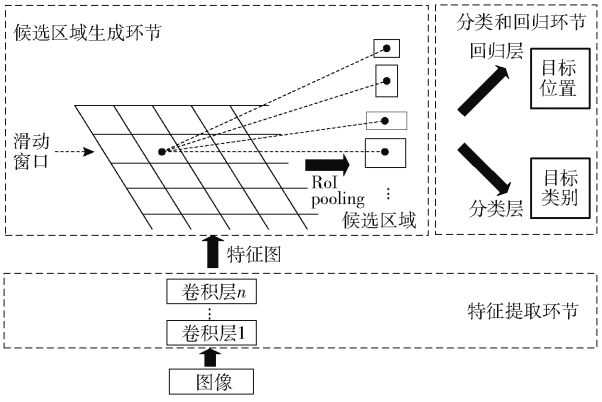


图 1 Faster R-CNN 算法结构示意图

Fig. 1 Structure diagram of Faster R-CNN algorithm

目标分类和位置框回归由一组级联的全卷积网络实现,输入生成的候选区域,计算待检测目标的类别概率实现分类,同时计算待检测目标最小外接矩形的中心点坐标、宽和高实现目标定位。Faster R-CNN 算法在具体应用时,检测精度易被影响。首先,若特征图的目标信息较少,RPN 生成的候选区域可能无法覆盖所有待检测目标,导致部分目标被漏检;其次,若某候选区域对应的特征中含有噪声,噪声将直接影响类别概率的计算结果,导致部分目标被错检。

对于本文研究对象,田间场景中含有大量无关的水稻叶片且部分稻穗被叶片遮挡。随机抽取 200 个稻穗进行统计,单个稻穗平均尺寸(长×宽)约为 260 像素×180 像素,长、宽均不足原始图像的十分之一,稻穗面积仅占原始图像的 0.38%。当直接应用 Faster R-CNN 算法检测稻穗时,存在以下两个主要问题:首先,小尺寸稻穗有限的原始信息在特征学习时不断损失,网络深层的特征图可能仅保留极少量目标信息,这导致映射特征图上的特征点所生成的候选区域并不能覆盖图像中的全部稻穗,大量小尺寸稻穗被漏检。其次,叶片遮挡不仅导致有限的稻穗原始信息进一步损失,而且使得稻穗的原始数据空间混入噪声。在特征学习的过程中,原始的局部叶片噪声将逐步传递给具有全局语义性的最终特征,这将直接导致稻穗特征质量下降,影响最终的稻穗识别和计数精度。

1.2 GFP-PD 算法设计

针对小尺寸稻穗的特征损失问题,首先,通过量化稻穗尺寸与网络感受野间的关系,选择适合的特征学习网络;其次,构建特征金字塔获取多尺度稻穗特征。针对稻穗特征含有的遮挡噪声,设计生成对抗式的修复模块为特征金字塔的每一层特征构建有效的优化映射,其目的是将含叶片噪声的稻穗特征拟合为真实的稻穗特征。GFP-PD 算法融合优化后的特征金字塔实现稻穗特征的有效增强,如图 2 所示,主要包括基于感受野分析的特征学习网络选择、基于生成对抗网络的特征修复、基于特征金字塔的特征融合和基于 Faster R-CNN 算法的检测 4 个主要环节。其中,Conv3_64 表示 64 通道且卷积核尺寸为3×3 的卷积计算,以此类推;pool/2 表示窗口尺寸为 2×2 的最大池化计算;Cls + Reg 表示计算类别和计算位置框。

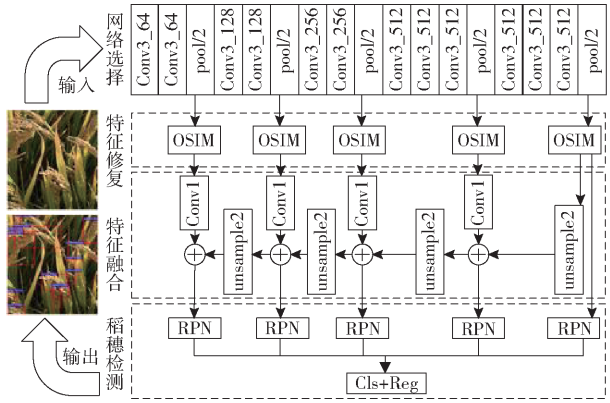


图 2 GFP-PD 算法框架示意图

Fig. 2 Structure diagram of GFP-PD algorithm

1.3 基于感受野分析的网络选择

特征学习网络将输入图像的原始信息进行迭代,抽象为特征图,同一网络不同卷积层对应特征图的语义信息和空间信息各不相同;而不同的特征学习网络因其结构不同,输出的特征图也存在差异。特征图上任一像素点所对应输入图像的区域被描述为感受野(Receptive field),卷积层感受野的计算公式为

$$S_{RF}(t) = (S_{RF}(t-1) - 1)N_s(t) + S_f(t) \quad (1)$$

式中 $S_{RF}(t)$——第 t 层卷积层感受野尺寸

$N_s(t)$——第 t 层卷积的步长

$S_f(t)$——第 t 层卷积滤波器尺寸

感受野区域内图像的信息直接影响学习到的特征质量。相较于待检测目标,感受野过大,特征中混入无关噪声;感受野过小,特征不能充分描述待检测目标。因此,总是希望特征学习网络的感受野与稻穗尺寸尽可能相近。理想情况下,GFP-PD 特征学习网络的感受野和待检测目标的尺寸关系为

$$\frac{S_{RF}(t)}{S_{obj}(h_{obj}, w_{obj})} \approx 1 \quad (2)$$

式中 $S_{obj}(h_{obj}, w_{obj})$——待检测目标尺寸

h_{obj}——待检测目标外接矩形长度

w_{obj}——待检测目标外接矩形宽度

根据式(1)计算卷积神经网络 Alex-Net[29]、ZF-Net[30]、VGG16-Net[31]和 Google-Net[32]的感

受野，不同网络的最后一层卷积层对应的感受野尺寸如表 1 所示。VGG16 - Net 最后一层特征对应的感受野为 212 像素 ×212 像素，相较于其他网络更接近稻穗尺寸，因此选择去除分类层的 VGG16 - Net 作为 GFP - PD 的特征学习网络。

表 1　不同卷积神经网络的感受野尺寸

Tab. 1　Size of receptive field of different networks

名称	最深层卷积层感受野尺寸/（像素 ×像素）	S_{RF}/S_{obj}
ZF - Net	139 ×139	0.41
Alex - Net	195 ×195	0.81
VGG16 - Net	212 ×212	0.96
Google - Net	224 ×224	1.07

1.4　基于生成对抗网络的特征修复

一般来说，一个成功的检测器总是基于优秀的物体特征表示和足够规模的数据学习。叶片对稻穗的遮挡使得稻穗特征中混入叶片噪声，影响最终的稻穗识别准确率和计数精度。为此，GFP - PD 算法基于生成对抗网络设计 OSIM 来优化田间稻穗的特征表达，其核心思想是不断为遮挡噪声生成新的值并对抗式地使含有遮挡噪声的稻穗特征逐渐逼近真实稻穗特征。如图 3 所示，OSIM 的网络结构由生成器、局部判别器和全局判别器共同构成。生成器为遮挡区域生成新特征值，局部判别器比较修复后遮挡区域与对应真实区域的差异，而全局判别器比较修复后整幅图像与对应真实图像的差异。生成器与两个判别器进行迭代，直至两个判别器无法确定生成的新特征是否是真实稻穗特征为止。

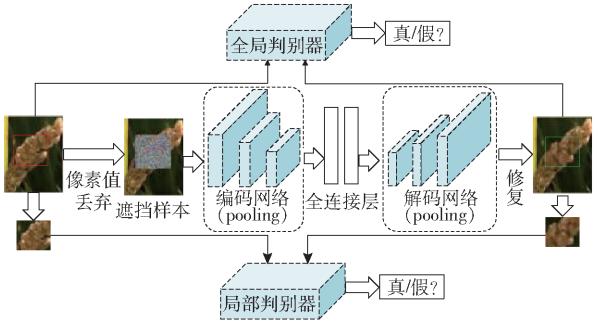


图 3　OSIM 网络结构示意图

Fig. 3　Structure diagram of OSIM

1.4.1　生成器

生成器被抽象为对遮挡特征的编码和解码过程。对于输入的含有叶片遮挡的稻穗图像，生成器通过编码网络学习图像特征并为其生成新特征。编码网络选择 VGG16 - Net 的 Conv1 层到 pool2 层作为基础结构，并在其后增加两个全连接层将填充后的新特征图传递到解码网络。为保证遮挡稻穗特征的修复精度，解码网络通过反卷积将生成的新特征解码为图像，送入判别器与真实图像进行比较。解码网络的结构与编码网络对称，不同之处仅在于使用解码网络 un-pooling 层替代编码网络中的 pooling 层。

1.4.2　局部与全局判别器

生成器可以保证生成特征和真实特征最终的值相差较小，但是生成特征并不一定能够有效地表示其对应的输入图像。不仅希望生成的稻穗特征与真实稻穗在内容上相似，而且能够有效地描述稻穗与图像中其他稻穗的结构关系。因此，分别为 OSIM 设计一个二分类的局部判别器和全局判别器。局部判别器与全局判别器采用相同的网络结构，选择 VGG16 - Net 的 Conv1 到 pool2 层作为基础结构并增加一个 sigmoid 分类层。局部判别器将生成器的注意力集中在遮挡区域的内部细节，引导修复后的遮挡区域在内容上与真实区域相似，而全局判别器则引导修复后的遮挡区域在结构上与其周围区域相近。

1.4.3　损失函数

为保证生成的新特征与真实稻穗特征在内容上相似，使用欧氏距离来度量生成特征与真实特征之间的差异，生成器损失函数 L_G 计算公式为

$$L_G = \frac{1}{2M}\sum_{i=1}^{M} \| x_i - x_i' \|_2^2 \tag{3}$$

式中　M——特征图上的像素点数量

x_i——稻穗的真实特征值

x_i'——遮挡处生成的新特征值

局部判别器和全局判别器拥有相同的损失函数，计算公式为

$$L_{LocalD} = L_{GlobalD} = \min_G \max_D E_{x\sim P_{data}(x)}(\lg D(x)) + E_{z\sim P_z(z)}(\lg(1 - D(G(z)))) \tag{4}$$

式中　L_{LocalD}——局部判别器损失函数

$L_{GlobalD}$——全局判别器损失函数

$E_{x\sim P_{data}(x)}$——真实数据分布

$E_{z\sim P_z(z)}$——遮挡噪声分布

$D(x)$——对生成数据的判别结果

$G(z)$——噪声生成的数据值

OSIM 的损失函数 L 由生成器和判别器共同构成，其计算公式为

$$L = L_G + L_{LocalD} + L_{GlobalD} \tag{5}$$

1.5　基于特征金字塔的特征融合

不同卷积层上输出特征图的语义信息和空间信息各不相同，低层卷积特征保留目标较多的空间和细节信息，而高层卷积特征具有较好的语义性。为进一步丰富稻穗特征，GFP - PD 算法参考特征金字塔网络（Feature pyramid networks，FPN）的结构[33]，抽取稻穗的多尺度特征并使用 OSIM 模块进行优化，按由深到浅的顺序融合高语义信息的深层特征

和高细节信息的浅层特征，使得稻穗特征在具有较好语义性的同时保留更多的空间信息。

特征融合时，对深层特征进行 2 倍上采样后与前一层特征进行横向连接（Lateral connections）。使用 RPN 分别对每一层融合后的新特征生成候选区域并进行检测。横向连接的计算公式为

$$\boldsymbol{F}_k = \boldsymbol{S}_{k-1} A_{1\times1} + f(\boldsymbol{S}_k) \quad (k = 1, 2, \cdots, n) \tag{6}$$

式中 $\boldsymbol{S}_{k-1}$——第 $k-1$ 层输出的特征矩阵

$\boldsymbol{S}_k$——第 k 层输出的特征矩阵

$A_{1\times1}$——1×1 卷积核

f——基于最邻近插值的 2 倍上采样函数

$\boldsymbol{F}_k$——第 $k-1$ 层和 k 层融合后的特征矩阵

1.6 基于 Faster R－CNN 算法的检测

GFP－PD 算法基于 Faster R－CNN 算法检测目标，实现目标分类和目标定位。对于 OSIM 优化的稻穗特征，GFP－PD 算法利用候选区域生成网络生成候选区域，使用尺寸为 3×3 的卷积核遍历（滑动步长为 1）优化后的特征图。设计 1∶1、2∶1、1∶2 共 3 种不同尺度的长宽比，以特征图上的每个像素点（锚点）为中心生成 9 个大小不同的预测框。检测模块的损失函数由类别损失和位置回归损失共同构成，损失函数为

$$L(\{P_i\},\{\boldsymbol{t}_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(P_i, P_i^*) + \lambda\frac{1}{N_{reg}}\sum_i P_i^* L_{reg}(\boldsymbol{t}_i, \boldsymbol{t}_i^*) \tag{7}$$

其中

$$L_{cls}(P_i, P_i^*) = -\lg(P_i^* P_i + (1-P_i^*)(1-P_i)) \tag{8}$$

$$L_{reg}(\boldsymbol{t}_i, \boldsymbol{t}_i^*) = \begin{cases} 0.5(\boldsymbol{t}_i - \boldsymbol{t}_i^*)^2 & (|\boldsymbol{t}_i - \boldsymbol{t}_i^*| < 1) \\ |\boldsymbol{t}_i - \boldsymbol{t}_i^*| - 0.5 & (|\boldsymbol{t}_i - \boldsymbol{t}_i^*| \geqslant 1) \end{cases} \tag{9}$$

式中 L——检测模块的整体损失函数

L_{cls}——用于分类的对数损失函数

L_{reg}——用于定位的鲁棒回归损失函数

i——锚点编号

P_i——锚点 i 被预测为目标的概率

P_i^*——锚点 i 表示正样本时，P_i^* 值为 1，反之为 0

$\boldsymbol{t}_i$——生成的预测框左上顶点和右下顶点坐标组成的 4 维向量

$\boldsymbol{t}_i^*$——原始的人工标注框左上顶点和右下顶点坐标组成的 4 维向量

N_{cls}——训练时的批尺寸（Mini-batch size）

N_{reg}——产生的候选区域数量

λ——分类损失和回归损失的平衡系数

2 基于 GFP－PD 算法的稻穗检测与计数

基于 GFP－PD 算法训练稻穗检测模型，流程如图 4 所示，主要包括 5 个步骤：原始图像采集；训练数据集构建；OSIM 修复模型训练；稻穗检测模型训

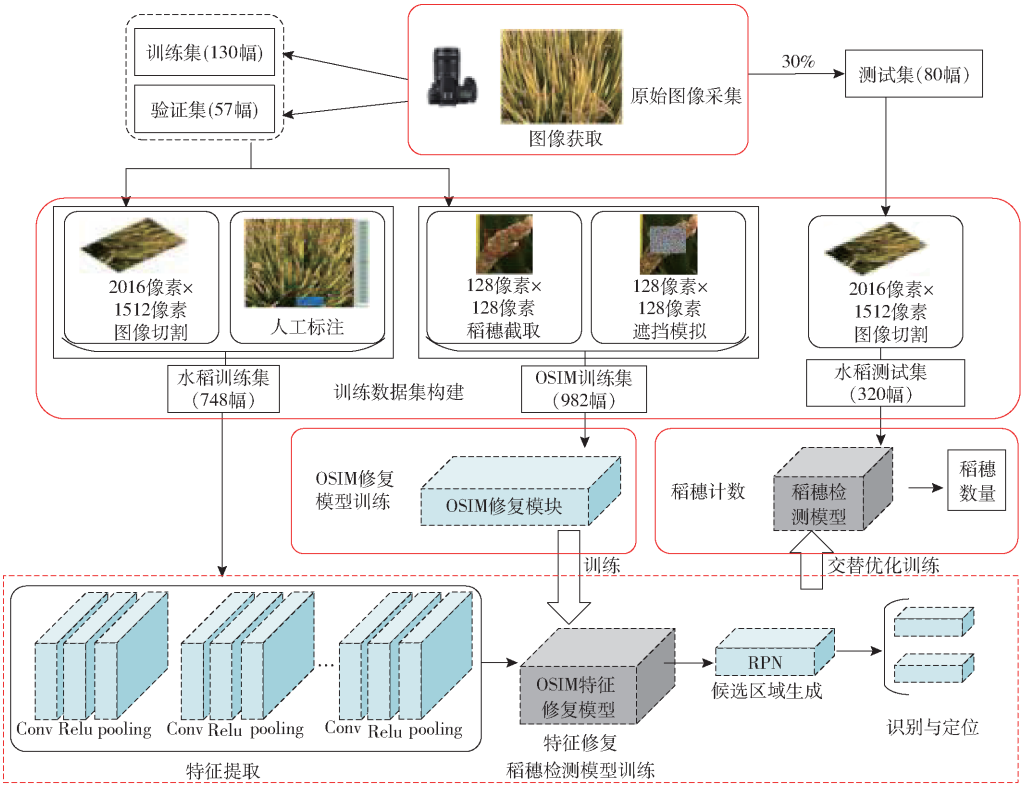


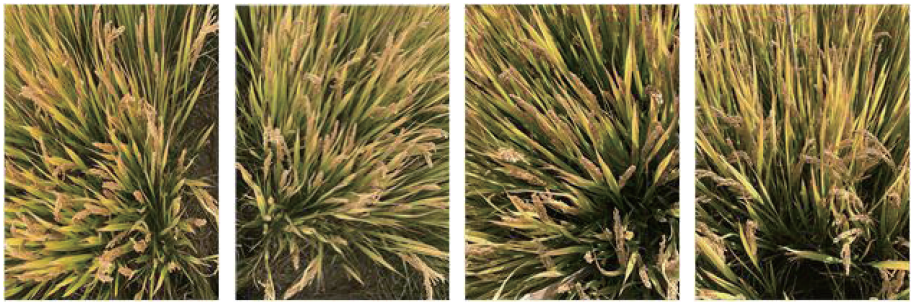
图 4 基于 GFP－PD 算法的稻穗检测与计数流程图

Fig. 4 Flow chart of panicle detection and counting based on GFP－PD

练；稻穗计数。

2.1　原始图像采集

水稻图像于 2017 年 10—11 月采集于江苏省南京市六合区艾津水稻生产基地（32° 16′ 44″ N，118°51′10″E，海拔 4 m），供试水稻品种为南粳 46，每穴 3 ~ 5 苗，株间距 30 cm，行间距 12 cm。使用佳能 EOS 70D 型数码相机在水稻植株侧上方 60 cm 处，进行随意角度拍摄，共采集图像 267 幅，图像分辨率均为 4 032 像素 ×3 024 像素。如图 5 所示，单幅图像中包含大量水稻稻穗与叶片，统计后稻穗平均尺寸约 260 像素 ×180 像素，绝大部分稻穗局部被叶片遮挡。



(a) 正前方　(b) 左前方　(c) 右前方　(d) 随机角度

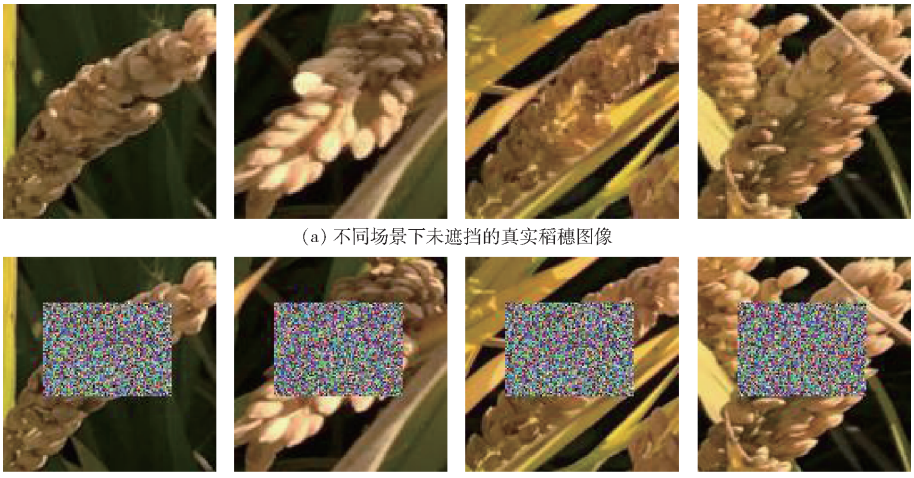
图 5　不同角度下获取的田间水稻图像

Fig. 5　Rice images taken from different angles

2.2　训练数据集构建

对于原始的田间水稻图像，随机选取 130 幅作为训练集，57 幅作为验证集，将剩余 80 幅作为测试集。训练集、验证集与测试集在图像样本空间上没有交集，基于划分后的数据集分别构造 OSIM 训练数据集、水稻训练集和水稻测试集。

（1）OSIM 训练数据集构建：基于 OSIM 的特征修复模型的损失在含有叶片遮挡的稻穗图像和未遮挡时的稻穗图像间传递，实际数据采集过程中，获得同一场景的含有叶片遮挡的稻穗图像和其对应未遮挡时的图像是困难的。因此本研究中，人工从划分后的水稻训练集和测试集图像中筛选未遮挡稻穗图像 982 幅，图像尺寸为 128 像素 ×128 像素。通过随机丢弃部分像素来模拟叶片遮挡，构造的 OSIM 训练数据集部分样本如图 6 所示。



(a) 不同场景下未遮挡的真实稻穗图像

(b) 不同场景下叶片遮挡模拟图像

图 6　OSIM 训练数据集示例

Fig. 6　Examples of train data for OSIM

（2）水稻训练集构建：本研究中单幅图像尺寸为 4 032 像素 ×3 024 像素，直接使用原图训练模型时导致硬件负载过高且耗时增加，因此按长、宽分别二等分切割训练集和验证集图像。切割后图像尺寸为 2 016 像素 ×1 512 像素，平均稻穗尺寸为 260 像素 ×180 像素，单个稻穗约占切割图像的 1.54%，稻穗相对于图像依然为小目标。利用开源标注工具 LabelImg 对切割后的训练集和验证集图像进行人工标注，将样本中的稻穗通过最小外接矩形标出，记录最小外接矩形的左上、右下顶点坐标，标注最小外接矩形框的类别标签为稻穗。切割后的图像和对应稻穗标注文件共同组成水稻训练集。

（3）水稻测试集构建：将测试集图像按长、宽分别二等分切割组成水稻测试集。

2.3 OSIM 特征修复模型训练

OSIM 被设计为一个独立的插件嵌入到 GFP-PD 中,为含有遮挡噪声的特征和真实稻穗特征构建一个有效映射。因此,OSIM 特征优化模型被独立训练。OSIM 的训练采用课程策略[33](Curriculum strategy)和随机梯度下降法(Stochastic gradient descent,SGD),整个过程包括 3 个阶段。首先,利用 L_G 训练遮挡区域并获得初步模糊内容;其次,利用 L_{LocalD} 提高修复区域内容的真实度;最后,利用 $L_{GlobalD}$ 调整修复区域与待检测目标的结构。随机生成特征修复模型的初始化参数,为了提高训练的有效性和效率,训练中后一阶段的模型总是以前一阶段的模型为基础微调(Fine turning)获得。

2.4 稻穗检测模型训练与稻穗计数

使用水稻训练集作为稻穗检测模型的输入数据,将通过特征学习网络获得的稻穗特征首先输入 2.3 节训练的 OSIM 特征修复模型,修复后的特征被输入 RPN 生成候选区域并进行识别和定位,使用随机梯度下降法和交替优化法(Alternating optimization)训练稻穗检测模型。稻穗计数时,将测试图像输入稻穗检测模型,在稻穗识别和定位的基础上统计稻穗数量。

3 实验与结果分析

3.1 模型训练软硬件平台

处理平台为 AMAX 的 PSC-HB1X 深度学习工作站,处理器为 Inter E5-2600 v3,主频为 2.1 GHz,内存为 128 GB,硬盘容量为 1 TB,GeForce GTX Titan X 显卡。运行环境为 Ubuntu 16.0.4、Caffe 和 Python 2.7。

3.2 评价指标

选用查全率 P_c、正确率 P_t 作为模型性能的评价指标,其中查全率是指检测正确的穗数与实际穗数的比值;正确率是指检测正确的穗数与识别为稻穗的目标数量的比值。计算公式为

$$P_c = \frac{N_{cor}}{N_{real}} \times 100\% \tag{10}$$

$$P_t = \left(1 - \frac{N_{err}}{N_{dect}}\right) \times 100\% \tag{11}$$

式中 N_{cor}——模型检测正确的穗数

N_{err}——模型检测错误的穗数

N_{real}——测试图像中实际含有穗数

N_{dect}——检测结果为稻穗的目标数量

3.3 GFP-PD 算法与其他算法的比较

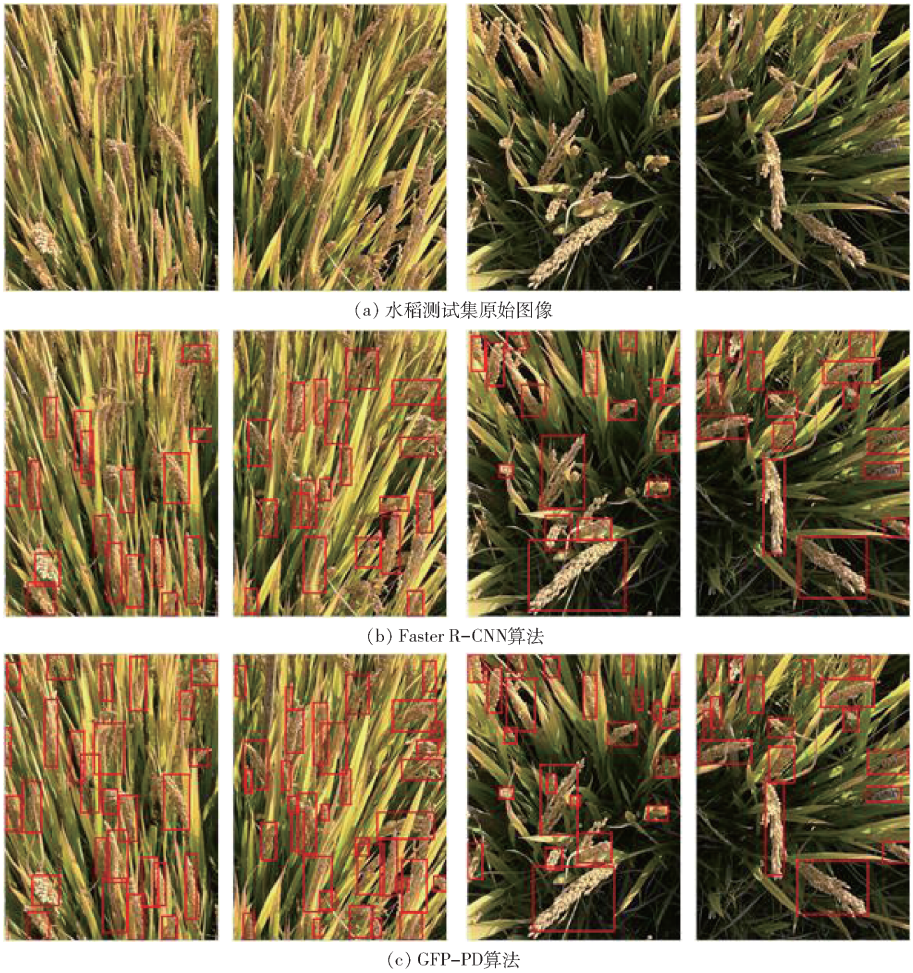
为了比较 GFP-PD 算法和其他算法对田间稻穗的计数精度,使用 2.2 节构建的水稻测试集作为测试数据,选择 LMM[15]、Panicle-SEG[20] 和 Faster R-CNN[24] 作为比较对象,4 种方法的平均查全率和正确率如表 2 所示。由表 2 可知,对于水稻测试数据集,本文方法性能最优,平均查全率和平均正确率分别为 90.82% 和 99.05%,较 Faster R-CNN 算法提高了 16.69、5.15 个百分点。Faster R-CNN 算法和 GFP-PD 算法对田间水稻的检测效果如图 7 所示,检测结果表明,相同测试场景下 GFP-PD 明显优于 Faster R-CNN 对稻穗的查全率。对于图 7 中第 3 列的测试图像,GFP-PD 算法能够准确地检测到图像中下部的一些微小稻穗,而 Faster R-CNN 却无法感知,这直接验证了 2.1 节中关于应用 Faster R-CNN 算法检测尺寸小数量多稻穗效果不佳的假设。

表 2 GFP-PD 与其他算法对测试图像的计数性能比较

Tab.2 Performance comparison of GFP-PD and other approaches on testing images %

算法	查全率 P_c				正确率 P_t			
	重复 1	重复 2	重复 3	均值±标准差	重复 1	重复 2	重复 3	均值±标准差
GFP-PD	90.41	90.88	91.17	90.82±0.39	99.27	99.01	98.86	99.05±0.20
LMM	82.82	81.46	82.20	82.16±0.68	95.58	95.09	94.86	95.18±0.36
Panicle-SEG	83.64	82.73	81.82	82.73±0.91	95.83	95.79	94.74	95.45±0.62
Faster R-CNN	74.29	73.92	74.17	74.13±0.19	93.61	93.73	94.35	93.90±0.31

LMM 与 Panicle-SEG 的计数精度相近,平均查全率均能够达到 82% 以上,和平均正确率 95% 以上,但较 GFP-PD 算法的查全率和正确率均分别降低约 8、4 个百分点。原因可能是对于基于颜色特征的分割算法 LMM,水稻稻穗与叶片较小麦贴合的更为紧密,这可能抑制 LMM 算法中滤波器对叶片等无关背景的去除性能,也提高了通过计算局部峰值来分割稻穗的难度。而对于 Panicle-SEG 算法,一方面可能是训练数据分布不同,Panicle-SEG 算法训练数据覆盖了中国主要的水稻品种,但与本文水稻品种相似的建模数据可能相对不足;另一方面,Panicle-SEG 算法构建模型的图像从小区正上方获取,而本文应用场景中叶片遮挡严重,这可能对 Panicle-SEG 算法中聚类生成候选区域的精度产生巨大干扰。



(a) 水稻测试集原始图像

(b) Faster R-CNN算法

(c) GFP-PD算法

图 7　Faster R－CNN 算法与 GFP－PD 算法检测效果对比

Fig. 7　Comparison of detection effects between Faster R－CNN and GFP－PD algorithms

3.4　基于 FPN 的特征增强有效性分析

为进一步量化基于感受野分析的特征学习网络选择和基于特征金字塔的特征增强对稻穗计数精度的提升，选择感受野差异显著的 ZF 网络(Zeilerfergus net)和 VGG16 网络(Visual geometry group net)并构造特征金字塔，分别训练稻穗检测模型。使用 2.2 节划分的水稻测试集作为测试数据，结果如表 3 所示。

表 3　网络选择和特征金字塔构造的性能分析

Tab. 3　Performance analysis of selecting network and constructing feature pyramid on test dataset　%

算法	查全率 P_c	正确率 P_t
Faster R－CNN＋ZF	71.03±0.80	92.29±0.86
Faster R－CNN＋VGG16	74.12±0.19	93.85±0.31
Faster R－CNN＋FPN＋ZF	83.35±0.43	92.67±0.41
Faster R－CNN＋FPN＋VGG16	87.10±0.70	93.87±0.37

由表 3 可知，基于 ZF 网络的 Faster R－CNN 的平均查全率和正确率分别为 71.03% 和 92.29%，相较于基于 ZF 网络的 VGG16 网络分别降低了 3.09、1.56 个百分点。类似的，对于构造特征金字塔后的 Faster R－CNN 算法，VGG16 网络的平均查全率和正确率较 ZF 网络也分别提高了 3.75、1.20 个百分点。使用 VGG16 作为特征学习网络的平均检测结果均优于 ZF，这说明据式(1)选择特征学习网络能够提升计数结果。分析感受野和稻穗尺寸选择 VGG16 网络后，构造特征金字塔的 Faster R－CNN 算法的平均正确率与 Faster R－CNN 算法相近，但平均查全率为 87.10%，较 Faster R－CNN 算法提高 12.98 个百分点，这表明构造特征金字塔提取并融合稻穗的多尺度特征，能够有效增强小尺寸稻穗的特征，从而进一步显著提升稻穗计数的查全率。

3.5　OSIM 特征优化有效性分析

为了验证 OSIM 模块对遮挡场景处理的有效性，对于图 6b 所示的遮挡模拟图像，使用反卷积将优化后的特征上采样到输入图像的相同尺寸，修复后的图像如图 8 所示。视觉结果显示，通过训练 OSIM 特征修复模型，可以有效降低遮挡噪声对目标特征的影响；同时，修复后的噪声一定程度上补充了损失的稻穗原始信息，更为准确地描述和表示输入的遮挡稻穗场景。

GFP－PD 算法对遮挡稻穗的检测效果如图 9

所示,绿框表示图像中真实存在的遮挡稻穗,红框表示 GFP－PD 算法检测到的稻穗。结果表明,GFP－PD 算法能够有效地识别局部被叶片遮挡的稻穗,但是稻穗位置的预测仍然存在一定偏差。为了进一步定量分析 OSIM 对 GFP－PD 算法计数精度的提升程度,组合 Faster R－CNN 算法和 OSIM 特征修复模块分别训练稻穗检测计数模型,对于水稻测试集,实验结果如表 4 所示。

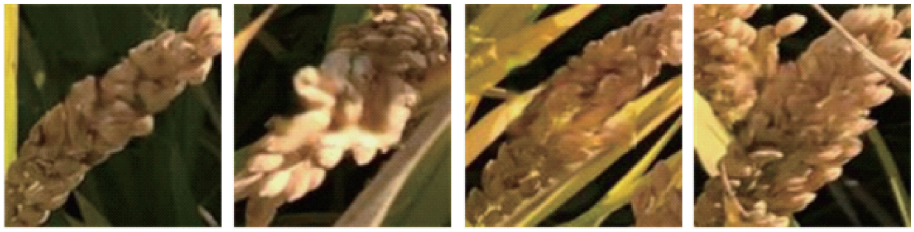


图 8 基于 OSIM 修复后的稻穗特征反卷积效果示意图

Fig. 8 Visual effect diagram of panicle feature repaired based on OSIM

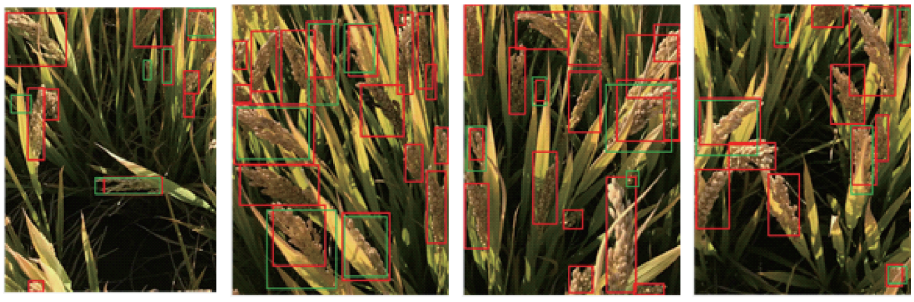


图 9 GFP－PD 算法对局部被叶片遮挡稻穗的检测效果

Fig. 9 Detecting effect of GFP－PD for panicles occluded by leaves locally

表 4 基于 OSIM 的特征修复对计数模型性能的提升分析

Tab. 4 Performance analysis of feature optimization based on OSIM on 2017 test dataset %

方法	查全率 P_c	正确率 P_t
Faster R－CNN + VGG16	74.12 ± 0.19	93.85 ± 0.31
Faster R－CNN + VGG16 + OSIM	74.30 ± 0.55	97.34 ± 0.26
Faster R－CNN + FPN + VGG16	87.10 ± 0.70	93.87 ± 0.37
Faster R－CNN + FPN + OSIM + VGG16(GFP－PD)	90.82 ± 0.39	99.05 ± 0.20

由表 4 可知,对于 VGG16 网络,基于 OSIM 训练的稻穗计数模型对稻穗的识别正确率达 97.34%,比 Faster R－CNN 提高了 3.49 个百分点。OSIM 修复模块优化稻穗特征后稻穗识别的正确率上升,这一趋势在构造特征金字塔后也同样明显。构造特征金字塔后的 Faster R－CNN 算法对水稻测试集的识别正确率仅为 93.87%,较 OSIM 特征优化后的正确率下降了 5.18 个百分点。正确率的变化验证了 1.2 节中含有噪声的稻穗特征将抑制稻穗识别正确率的假设,表明基于生成对抗网络的 OSIM 特征修复模块能够去除稻穗特征中的噪声,有效地提高了田间遮挡条件下成熟期水稻稻穗的特征质量,进而提升其计数精度。

3.6 GFP－PD 算法的影响因素分析

对田间水稻图像计数时,发现水稻测试集中有些子图中穗和叶清晰,而有些子图中穗和叶模糊且有部分穗人眼难以识别,称之为清晰子图和模糊子图。模型对于同一图像的不同局部的计数精度偏差明显,为了进一步探究算法和模型的应用边界,基于 GFP－PD 算法训练稻穗计数模型,对水稻测试集的清晰子图和模糊子图的计数结果如表 5 所示。

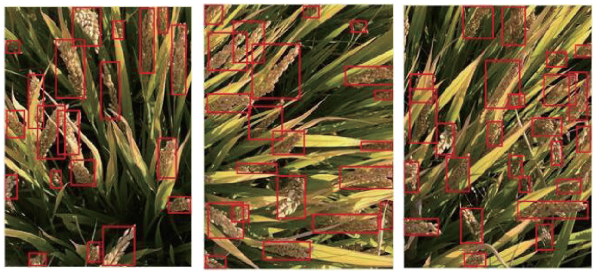
表 5 GFP－PD 算法对不同清晰度测试图像的计数结果比较

Tab. 5 Results of GFP－PD for images with different definitions %

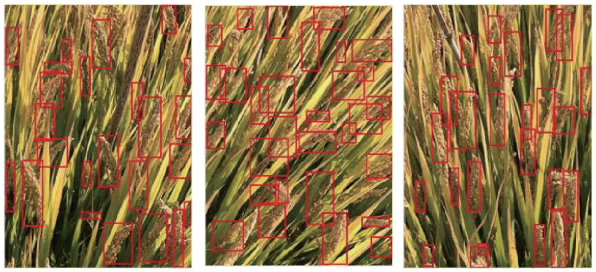
子图清晰度	查全率 P_c	正确率 P_t
清晰	96.99 ± 0.23	99.07 ± 0.61
模糊	85.45 ± 0.41	98.56 ± 0.33

由表 5 可知 GFP－PD 算法对清晰子图的查全率为 96.99%,比模糊子图高 11.54 个百分点;对清晰子图的稻穗识别正确率比模糊子图提高 0.51 个百分点。GFP－PD 算法对清晰子图的计数精度明显优于模糊子图,稻穗检测模型对模糊稻穗并不敏感。如图 10 所示,GFP－PD 算法能够准确地识别清晰子图中的绝大部分稻穗,但是对于模糊子图上半部边缘附近的稻穗存在明显漏检。由图 10 可知,清晰子图主要对应于原始图像的中下部分,在获取时处于相机焦点附近;而模糊子图主要对应原始未分割图像的上半部分边缘,在获取时处于相机焦点以外。模糊子图的稻穗原始信息进一步减少,虽然

构造特征金字塔一定程度上增强了稻穗特征，但基于融合后特征图生成的候选区域仍然不能有效覆盖图像中所有的稻穗，这影响了本文稻穗计数模型最终的计数查全率。



(a) 清晰子图



(b) 模糊子图

图 10　GFP－PD 算法对不同清晰度子图的检测效果

Fig. 10　Detecting effects of GFP－PD algorithm for panicles with different definitions

3.7　讨论

基于感受野分析选择特征学习网络和基于特征金字塔提取并融合多尺度稻穗特征，有利于提升田间小尺寸稻穗的检测和计数精度；基于 OSIM 的特征修复，有利于提升田间稻穗特征的质量，显著提高了稻穗识别的正确率；基于 GFP－PD 算法训练稻穗检测模型，可以有效识别局部被叶片遮挡的田间稻穗并自动预测稻穗的位置。进一步分析计数结果发现，抑制计数精度的另一个主要因素是存在部分区域重叠的稻穗。如图 11 所示，稻穗检测模型易将粘连的稻穗识别为一个稻穗或将重叠稻穗的部分误识为一个完整的稻穗。

对于 GFP－PD 算法而言，构建 OSIM 修复模型虽然能够提高稻穗特征质量，但是消耗额外的计算量。在实际的应用中，还需要进一步考虑水稻品种（亚种）、图像采集装置、光照条件、成像高度等因素变化的影响。对于形态差异明显的不同品种水稻稻穗，下一步将尝试对特征提取网络和候选区域建议网络进行优化，提取稻穗与叶更细粒度的可分性特征来提高分类和计数的精度。



图 11　GFP－PD 算法对粘连或重叠稻穗的检测效果

Fig. 11　Detecting effects of GFP－PD algorithm for overlapping panicles

4　结论

（1）GFP－PD 算法选择最佳特征学习网络并构建特征金字塔，显著提高了田间小尺寸稻穗的查全率。对于构造特征金字塔后的 Faster R－CNN 算法，基于 VGG16 网络的稻穗平均查全率和识别正确率分别为 87.10% 和 93.87%，比 ZF 网络分别提高了 3.75、1.20 个百分点；而对于 Faster R－CNN 算法，基于 VGG16 网络的平均查全率和识别正确率比 ZF 网络也分别提高了 3.09、1.56 百分点。

（2）设计并训练 OSIM 能够有效去除稻穗特征中的叶片噪声，显著提高了局部被叶片遮挡的稻穗的识别正确率。对于基于 VGG16 网络的 Faster R－CNN 方法，使用 OSIM 修复模型优化稻穗特征前后，对水稻测试集的识别正确率分别为 93.85% 和 97.34%；当构造特征金字塔时，使用 OSIM 修复模型优化稻穗特征后，识别正确率由 93.87% 提高为 99.05%，对遮挡稻穗的识别正确率提高了 5.18 个百分点。

（3）GFP－PD 算法能够有效检测局部被叶片遮挡的田间小尺寸稻穗，并对其计数。稻穗计数模型的平均查全率和识别正确率分别为 90.82% 和 99.05%，比 Faster R－CNN 算法结果分别提高了 16.69、5.15 个百分点。但 GFP－PD 算法对清晰稻穗的查全率较模糊稻穗提高 11.54 个百分点，样本清晰度是制约 GFP－PD 算法检测精度的主要因素。

参　考　文　献

[1] SLAFER G A, SAVIN R, SADRAS V O. Coarse and fine regulation of wheat yield components in response to genotype and environment[J]. Field Crops Research, 2014, 157:71－83.

[2] FERRANTE A, CARTELLE J, SAVIN R, et al. Yield determination, interplay between major components and yield stability in a traditional and a contemporary wheat across a wide range of environments[J]. Field Crops Research, 2017, 203: 114－127.

[3] GIUFFRIDA M V, MINERVINI M, TSAFTARIS S A. Learning to count leaves in rosette plants[C]// Proceedings of the Computer Vision Problems in Plant Phenotyping, 2015:7－10.

[4] DOBRESCU A, VALERIO G M, TSAFTARIS S A. Leveraging multiple datasets for deep leaf counting[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2072－2079.

[5] AICH S, STAVNESS I. Leaf counting with deep convolutional and deconvolutional networks[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2080 - 2089.
[6] BARRÉ P, STÖVER B C, MÜLLER K F, et al. LeafNet: a computer vision system for automatic plant species identification [J]. Ecological Informatics, 2017, 40: 50 - 56.
[7] MUSSADIQ Z, LASZLO B, HELYES L, et al. Evaluation and comparison of open source program solutions for automatic seed counting on digital images[J]. Computers and Electronics in Agriculture, 2015, 117: 194 - 199.
[8] MALDONADO W, JOSÉ C B. Automatic green fruit counting in orange trees using digital images[J]. Computers and Electronics in Agriculture, 2016, 127:572 - 581.
[9] STEIN M, BARGOTI S, UNDERWOOD J. Image-based mango fruit detection, localization and yield estimation using multiple view geometry[J]. Sensors, 2016, 16(11): 1915 - 1923.
[10] 陈含, 吕行军, 田凤珍,等. 基于 Sobel 算子边缘检测的麦穗图像分割[J]. 农机化研究, 2013, 35(3):33 - 36.
CHEN Han, LÜ Xingjun, TIAN Fengzhen, et al. Wheat panicle image segmentation based on Sobel operator-edge detection [J]. Journal of Agricultural Mechanization Research, 2013, 35(3):33 - 36. (in Chinese)
[11] 刘涛, 孙成明, 王力坚,等. 基于图像处理技术的大田麦穗计数[J/OL]. 农业机械学报, 2014, 45(2):282 - 290.
LIU Tao, SUN Chengming, WANG Lijian, et al. In-field wheat ear counting based on image processing technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(2):282 - 290. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20140247&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2014.02.047. (in Chinese)
[12] 赵锋, 王克俭, 苑迎春. 基于颜色特征和 AdaBoost 算法的麦穗识别的研究[J]. 作物杂志, 2014,30(1):141 - 144.
ZHAO Feng, WANG Kejian, YUAN Yingchun. Study on wheat spike identification based on color features and AdaBoost algorithm [J]. Crops, 2014,30(1):141 - 144. (in Chinese)
[13] COINTAULT F, GUERIN D, GUILLEMIN J P, et al. In-field triticumaestivum ear counting using colour-texture image analysis[J]. New Zealand Journal of Crop and Horticultural Science,2010,36(2):117 - 130.
[14] ZHOU C, LIANG D, YANG X, et al. Wheat ears counting in field conditions based on multi-feature optimization and TWSVM [J]. Frontiers in Plant Science, 2018, 9: 1024 - 1040.
[15] FERNANDEZ-GALLEGO J A, KEFAUVER S C, GUTIÉRREZ N A, et al. Wheat ear counting in-field conditions: high throughput and low-cost approach using RGB images[J]. Plant Methods, 2018, 14(1): 22 - 34.
[16] DUAN L, HUANG C, CHEN G, et al. Determination of rice panicle numbers during heading by multi-angle imaging[J]. The Crop Journal, 2015, 3(3): 211 - 219.
[17] ZHU Y, CAO Z, LU H, et al. In-field automatic observation of wheat heading stage using computer vision [J]. Biosystems Engineering, 2016, 143:28 - 41.
[18] LU H, CAO Z, XIAO Y, et al. Region-based colour modelling for joint crop and maize tassel segmentation[J]. Biosystems Engineering, 2016, 147: 139 - 150.
[19] LI Q, CAI J, BERGER B, et al. Detecting spikes of wheat plants using neural networks with Laws texture energy[J]. Plant Methods, 2017, 13(1): 83 - 96.
[20] XIONG X, DUAN L, LIU L, et al. Panicle-SEG: a robust image segmentation method for rice panicles in the field based on deep learning and superpixel optimization[J]. Plant Methods, 2017, 13(1): 104 - 119.
[21] OLSEN P A, RAMAMURTHY K N, RIBERA J, et al. Detecting and counting panicles in sorghum images[C]// 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 2018: 400 - 409.
[22] 张领先,陈运强,李云霞,等. 基于卷积神经网络的小麦麦穗检测计数系统[J/OL]. 农业机械学报,2019,50(3):144 - 150.
ZHANG Lingxian, CHEN Yunqiang, LI Yunxia, et al. Detection and counting system for winter wheat ears based on convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2019,50(3):144 - 150. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20190315&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2019.03.015. (in Chinese)
[23] ALKHUDAYDI T, ZHOU J. SpikeletFCN: counting spikelets from infield wheat crop images using fully convolutional networks [C]// International Conference on Artificial Intelligence and Soft Computing. Springer, Cham, 2019: 3 - 13.
[24] MADEC S, JIN X, LU H, et al. Ear density estimation from high resolution RGB imagery using deep learning technique[J]. Agricultural and Forest Meteorology, 2019, 264(1): 225 - 234.
[25] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137 - 1149.
[26] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 - 587.
[27] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904 - 1916.
[28] GIRSHICK R. Fast R - CNN[C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440 - 1448.
[29] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// Advances in Neural Information Processing Systems, 2012: 1097 - 1105.
[30] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]// European Conference on Computer Vision. Springer, Cham, 2014: 818 - 833.
[31] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// International Conference of Learning Representation, 2015.
[32] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1 - 9.
[33] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 - 2125.
[34] BENGIO Y, LOURADOUR J, COLLOBERT R, et al. Curriculum learning[C]// Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009: 41 - 48.