doi:10.6041/j.issn.1000-1298.2025.12.051

融合多尺度特征与多列卷积网络的水稻苗无人机遥感计数方法

潘　义[1]　苏浩然[1]　秦琳琳[1]　吴　刚[1]　石　春[1]　王宜坤[2]　华　胜[2]
(1.中国科学技术大学自动化系,合肥 230026;2.安徽喜洋洋农业科技有限公司,合肥 231521)

摘要:水稻早苗期种植密度与产量估算、田间管理密切相关。随着计算机视觉技术的发展,利用无人机图像识别技术自动统计水稻数量正逐步取代传统的人工采样和计数方法。然而,无人机拍摄的图像存在水稻苗形状不规则、水稻苗重叠、尺度变化和背景遮挡等问题,导致特征学习出现偏差,进而影响了水稻密度的准确估算和计算预测的效率。水稻苗之间显著的相对大小差异也成为影响计数精度的另一重要因素。针对这些问题,提出了一种基于改进的多列卷积神经网络的水稻苗计数算法(Rice counting network, RCN)。该算法根据目标的大小和形状,提出了改进高斯密度核,结合水稻图像与中心点 RGB 图像的相似度信息,使密度核更加贴合水稻苗纹理。对不同大小和形状的稻穗数量进行量化,并对尺度进行精细划分,计算出需要融合的不同特征感受野大小,采用融合膨胀卷积模块,增加了感受野大小,增强了模型对多尺度环境中的处理效果。在此基础上,对不同尺度的通路添加合适的注意力机制,增强不同尺度下对水稻苗合理信息的关注,在复杂背景、多尺度和密集遮挡的场景中体现出了明显的有效性。在高度 5 m 捕获的 RSD2023 数据集上,RCN 的平均绝对误差为 3.10 穴/m^2,均方误差为 4.03 穴^2/m^4,优于 MCNN、CSRNet、SCAR 和 TasselNetV2 等主流算法。

关键词:水稻苗;计数;密度图;多列神经网络;无人机遥感;多尺度特征
中图分类号:S252　　**文献标识码**:A　　**文章编号**:1000-1298(2025)12-0560-08　　**OSID**:

Rice Seedling Counting Method Based on Multi-column Convolutional Network Integrating Multi-scale Features

PAN Yi[1]　SU Haoran[1]　QIN Linlin[1]　WU Gang[1]　SHI Chun[1]　WANG Yikun[2]　HUA Sheng[2]
(1. *Department of Automation, University of Science and Technology of China, Hefei* 230026, *China*
2. *Anhui Xiyangyang Agricultural Technology Co., Ltd., Hefei* 231521, *China*)

Abstract: The planting density of rice in the early seedling stage is closely related to yield estimation and field management. With the development of computer vision technology, the use of drone image recognition technology to automatically count rice is gradually replacing traditional manual sampling and counting methods. However, images taken by drones have problems such as irregular rice seedling shapes, overlapping rice seedlings, scale changes, and background occlusion, which lead to deviations in feature learning, which in turn affects the accurate estimation of rice density and the efficiency of computational prediction. Significant relative size differences among rice seedlings will also become another important factor affecting counting accuracy. In response to these problems, a rice seedling counting algorithm (rice counting network, RCN) was proposed based on an improved multi-column convolutional neural network. Based on the size and shape of the target, an improved Gaussian density kernel was proposed, which combined the similarity information of the rice image and the RGB image of the center point to make the density kernel more suitable for the texture of the rice seedlings. The number of rice ears of different sizes and shapes was quantified, and the scales were divided finely to calculate the receptive field sizes of different features that needed to be fused. The fusion expansion convolution module was used to increase the receptive field size and enhance the model's multi-scale capabilities. On

收稿日期:2024-08-12　修回日期:2024-10-14
基金项目:喜洋洋公司农业科研发展基金项目(112100000066)
作者简介:潘义(2000—),男,硕士生,主要从事农业信息化技术研究,E-mail:alexpanyi@163.com
通信作者:秦琳琳(1975—),女,高级工程师,博士,主要从事人工环境建模与控制研究,E-mail:qinll@ustc.edu.cn

this basis, appropriate attention mechanisms were added to the pathways at different scales to enhance attention to reasonable information about rice seedlings at different scales. It showed obvious effectiveness in scenes with complex backgrounds, multi-scales and dense occlusions. On the RSD2023 data set captured at a height of 5 m, RCN's mean absolute error MAE = 3.10 holes/m^2 and mean square error MSE = 4.03 $holes^2/m^4$, which were better than that of the mainstream algorithms such as MCNN, CSRNet, SCAR and TasselNetV2.

Key words: rice seedlings; counting; density map; multi-column neural network; UAV remote sensing; multi-scale feature

0　引言

我国是最大的水稻生产国,水稻产量占我国粮食总产量的 30%[1]。在水稻种植过程中,种植密度是影响水稻生长和产量的重要因素[2]。水稻苗期的植株密度对预测产量、最终成穗密度及生长状况具有决定性作用,显著影响作物生长的多个方面及田间管理措施,是实现籽粒高产的重要农艺因素[3]。

因此,水稻苗期种植密度的计算至关重要。目前,水稻苗期植株计数仍然主要依赖于实地调查的人工采样和计数,并且经常需要人工处理每个样本,过程繁琐、低效且不准确[4]。

随着无人机遥感技术的不断进步,无人机平台在农业中的应用越来越广泛。无人机作为一种低成本的替代检测和数据分析技术[5],在农业信息采集中具有灵活、经济、周期短等特点[6]。在水稻监测方面,无人机具有快速获取植被信息和实时动态监测的优势。

深度学习在图像分类、目标检测和图像分割方面展现出了优异的性能[7],并已应用于植物计数[8-9]。然而,标注图像是一项非常耗费人力和时间的任务,这通常导致全面注释的数据集包含较少的图像,而具有更多图像的数据集则注释不够详细[10]。此外,遮挡和重叠会导致模型的检测和计数结果出现严重偏差[11]。

密度图回归是一种用于估计图像中物体密度或数量的技术。它通过将检测和回归方法相结合,在提高召回率的同时输出物体的位置信息,该方法在物体计数领域有着广泛的应用[12-18]。

然而,基于多列卷积神经网络(Multi-column convolutional network, MCNN)的植物计数研究相对较少,大多数研究是基于简单背景下的单一尺度物体计数[19-20]。

当水稻苗的分辨率和绝对尺寸增加时,MCNN 的最大感受野不足以有效提取目标特征,导致感受野与多尺度物体大小之间不匹配,TasselNetV2[16]模型通过使用上下文信息来挖掘对象与像素之间的关系,从而扩展感受野,但这会显著增加网络的计算开销。因此,在 MCNN 模型的基础上,如何调整感受野以适应实际的多尺度水稻苗,是将 MCNN 应用于水稻苗计数需解决的关键问题。

农业领域中的计数对象是植物,其形状往往不是规则的圆形,导致由高斯核生成的密度图与实际植被分布不完全一致。文献[21]采用基于协方差矩阵改进的高斯核来模拟稻穗等狭长物体的位置分布,但这种方法仍存在局限性,即高斯核的固定性使其无法根据植被具体形状变化灵活调整。

水稻田还具有复杂背景及密集遮挡等问题。文献[22]发现,视觉注意机制在识别相似和复杂背景中的物体方面表现良好,为解决农田场景中水田反光、水稻苗遮挡、尺度差异等视觉问题提供了思路。受人眼主动动态感知信息模型的启发,结合区域注意力机制的应用,可以提高对密集场景、复杂背景以及尺度多样化场景下水稻苗检测的精度。

综上,针对高密度水稻苗计数的问题,本文通过改进多列卷积神经网络,采用相似度改进密度图,并使用膨胀卷积改进卷积核以增加感受野,采用合适的注意力机制对不同尺度分支进行细节和整体信息的增强,建立水稻苗计数网络(Rice counting network, RCN)。

1　材料与方法

1.1　研究区概况

实验在安徽省合肥市庐江县实验基地(31°47′N, 117°23′E)开展,如图 1 所示。实验区属于亚热带季风气候区,四季分明,气温变化明显,主要作物为小麦和水稻。全县年平均气温为 16℃,年平均降水量

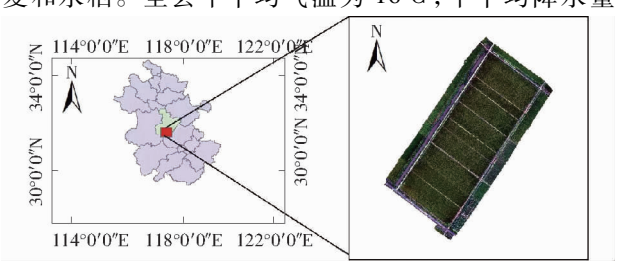

图 1　研究区域地理位置及正射影像

Fig. 1　Geographical location of study area and orthophoto

为 1 170.5 mm。降水主要集中在 4—8 月,7 月达到峰值。本次实验所用的水稻品种为常规晚粳稻品种银香 38。

1.2 数据获取及标注

实验中使用大疆 M100 型无人机,在高度 5 m 处以 2 m/s 的速度采集数据,水平重叠率为 80%,纵向重叠率为 80%。无人机图像于 2023 年 8 月 8 日前后水稻的幼苗期拍摄。数据收集时间主要在当地 10:00—13:00,在有利的光照条件下进行。使用大疆 GS PRO 软件进行航路规划,确保每次数据采集的高度和区域保持一致。

大疆 M100 型无人机搭载 ZENMUSE X3 型相机,分辨率为 1 920 像素×1 080 像素,在 5 m 空中的分辨率约为 0.3 cm,采用 1 s 间隔拍摄,并保存为 jpg 格式。最终剔除相似度过高或清晰度较差的图像,并使用 Photoscan 软件进行对齐,生成相应的正射影像。

1.3 数据标注与制作

由于水稻苗是具有中心特征的植物,本文采用点标注的方法,对水稻苗的中心进行标注,因为点标注只需要对其中心进行标注,故标注速度快,节省人力成本。使用 Python 编写程序来标注水稻植株的中心,并以左上角坐标为原点记录每株水稻的中心坐标,然后生成相应的.mat 文件。数据集的详细信息如表 1 所示,命名为 RSD2023(Rice Seeding Dataset 2023)。

为了防止光照强弱、拍摄角度等对训练产生影响,对图像进行旋转、裁剪、随机改变亮度和顺序打乱处理,增加了数据的鲁棒性。

表 1 数据集参数

Tab.1 Dataset

数据集	图像数量	水稻苗数/株	平均每幅图像水稻苗数/株
训练集	1 000	173 241	173.24
测试集	300	50 517	168.39
合计	1 300	223 758	172.12

2 水稻种植密度模型建立

本文以多列卷积神经网络(MCNN)为主干框架建立 RCN。MCNN 在高密度、高遮挡的人群计数场景中得到了较好的密度图和计数精度。本研究基于水稻株种植的场景对 MCNN 网络做了以下改进:

(1)在密度图生成过程中,引入了相似度增强的高斯核,通过将水稻苗周围像素点与中心点的相似度,与高斯卷积核进行点对点相乘,解决了由于水稻苗的不规则形状而导致的高斯核圆形不匹配问题。

(2)增加了膨胀卷积融合模块,与原始特征层融合,构建了多尺度的特征层,以解决多尺度稻田环境下 MCNN 感受野范围过小的问题。

(3)针对 MCNN 的 3 个感受野分支,分别引入了不同的注意力机制,增强对复杂稻田环境下复杂背景和遮挡问题的信息关注度。

2.1 模型概述

RCN 结构如图 2 所示,由 3 个模块组成:密度图生成模块(Density map generation module, DMFM)、改进的 MCNN 主干(Improved multi-column convolutional backbone, IMCB)、后处理模块(Post-processing module, PM)。

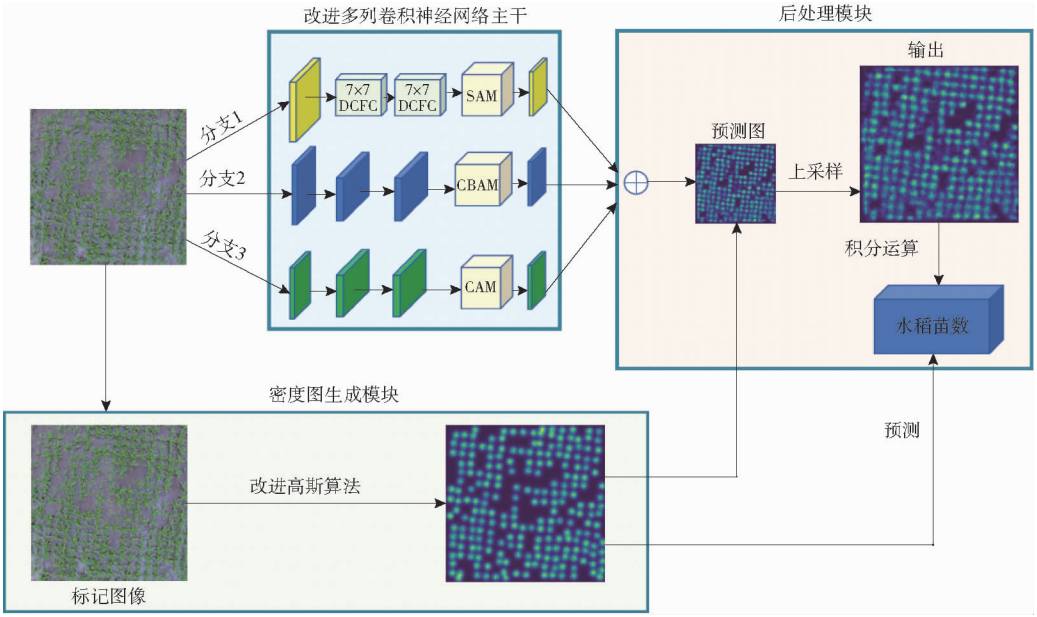

图 2 水稻苗计数网络结构图

Fig.2 Structure diagram of rice counting network

其中改进的 MCNN 主干结构如图 3 所示，共包含膨胀卷积、池化层、普通卷积和注意力模块 4 个模块。为了估计图像中的水稻苗数量，首先 IMCB 中通过不同尺度的卷积获取不同尺度特征，接着采用膨胀卷积扩展网络的感受野，然后使用注意力机制关注不同尺度中更加关键的信息，最后通过 1×1 卷积将这些特征融合，生成相应的密度图，通过 PM 模块进行分辨率的恢复并回归完成水稻苗计数。

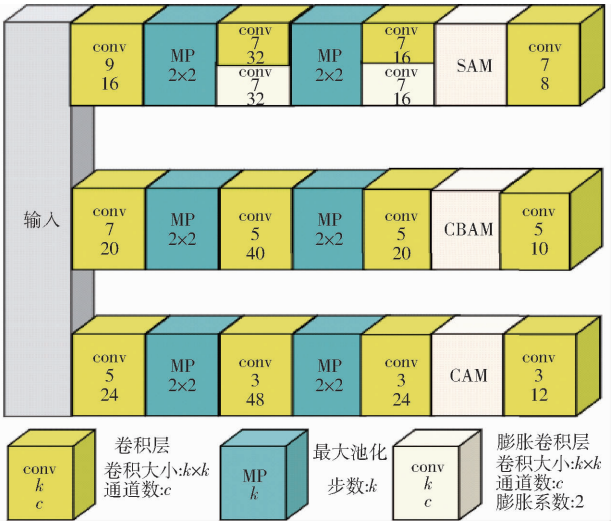

图 3　水稻苗计数 MCNN 主干结构图

Fig. 3　Specific structure of MCNN backbone for rice counting network

2.2　相似度改进密度图生成

在模型中的 DMFM 模块，首先按照类似 MCNN 的方法，根据标注的水稻株中心坐标进行高斯核卷积，生成从中心点向周围逐渐减小的概率密度图，即地面真值（Ground truth，GD）。高斯核卷积核公式为

$$H(x,y)=\frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{(\Delta x)^2+(\Delta y)^2}{2\sigma^2}} \tag{1}$$

其中　$\Delta x=x-x_i$　　$\Delta y=y-y_i$

式中　$H(x,y)$——对应(x,y)坐标高斯核的值

(x,y)——标注点坐标

(x_i,y_i)——标注点周围点坐标

σ——标准差

从图 4 中可以看出，水稻株呈现出不规则形状，而非标准的圆形，与 MCNN 模型中的人群计数特征不同。针对这一问题，在高斯核生成地面真值后，计算原始图像中心点与周围点的相似度，并将其与高斯卷积核进行点对点相乘生成地面真值，以增强密度图对细节信息的提取，确保地面真值沿水稻苗纹理从中心向外逐渐减小。生成相似度增强的高斯核（Similarity-enhanced Gaussian kernel，SGK）的公式为

$$G(x,y)=\frac{1}{1+\sqrt{m(a,b)-m(x,y)}}\times H(x,y) \tag{2}$$

式中　$G(x,y)$——对应(x,y)的真实密度图

$m(x,y)$——对应图像像素灰度

(a,b)——标注的水稻株的中心点坐标

通过上述处理，在不规则的水稻株中，能更好地提取密度图的细节信息，其效果如图 4 所示。

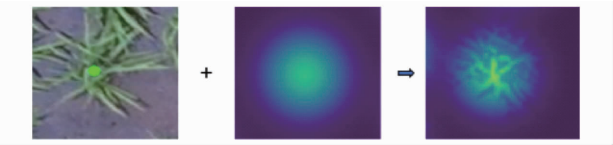

图 4　相似度改进的高斯核处理效果

Fig. 4　Similarity-enhanced Gaussian kernel

2.3　膨胀卷积融合模块

在模型中的 IMCB 模块中，为了解决多尺度稻田环境下 MCNN 感受野范围过小的问题，引入了一种基于膨胀卷积融合的模块（Dilated convolution fusion conv，DCFC）。以 3×3 卷积模型为例，具体方法如图 5 所示。

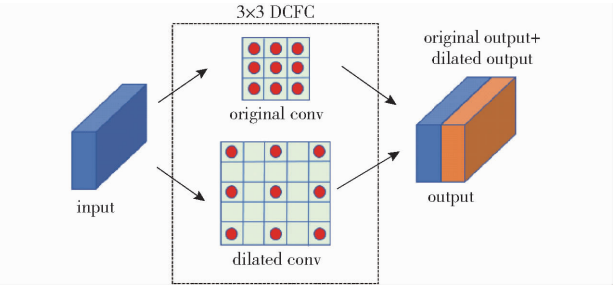

图 5　3×3 膨胀卷积融合模块

Fig. 5　3×3 dilated convolution fusion Conv

分别应用 3×3 的标准卷积和膨胀率为 2 的 3×3 膨胀卷积对输入进行卷积处理，然后将它们的输出进行通道叠加，从图 5 中可以看出 3×3 的 DCFC 公式为对应卷积的感受野大小分别为 3×3 和 5×5，对于其他大小为 $n\times n$ 的卷积核，膨胀率为 2 后的感受野大小分别为 $n\times n$ 和 $(2n-1)\times(2n-1)$，因此，膨胀卷积后的感受野从 $n\times n$ 扩展到了 $(2n-1)\times(2n-1)$。

为了保持输出图像尺寸不变，需要对图像周围进行填充，填充大小的计算公式为

$$p=d(i-1)/2 \tag{3}$$

式中　p——填充大小

d——膨胀率

i——输入图像尺寸

2.4　基于尺度注意力机制模块

在模型中的 IMCB 模块中，传统的 MCNN 模型通过 3 种不同尺度的分支来获取不同尺度的目标特

征,但是对于不同尺度信息的重视程度是相同的。然而,在实际的水稻株特征提取过程中,不同尺度的信息关注点是不同的。为了提高模型在计数精度上的表现,本文引入了 3 种不同的注意力机制,其中融合空间注意力和通道注意力机制的卷积块注意力模块(Convolutional block attention module, CBAM)[23]适用于中尺度范围。为了减少模型的资源消耗和运行时间,将空间注意力和通道注意力分别应用于更加适用的尺度上。在小尺度分支中,引入了通道注意力机制(Channel attention module, CAM),CAM 能够有效应对高度拥挤的场景,其中前景纹理与一些背景区域(如树木、建筑等)非常相似,CAM 的应用能够有效减小估计误差[24]。同时,在大尺度分支中,加入了空间注意力机制(Spatial attention module, SAM),SAM 能够在空间维度上捕捉广泛的依赖关系,特别适用于发现具有相似局部图案和纹理特征的场景,因此将其应用于尺度变化范围较大的场景中。

2.5 网络评价指标

本文使用平均绝对误差(Mean absolute error, MAE)、均方误差(Mean square error, MSE)、相对平均绝对误差(Root mean absolute error, rMAE)来评价本模型的计数性能。

3 结果与讨论

3.1 实验环境和训练参数

本实验在 64 位 Windows 11 操作系统的台式机上进行,运行内存为 16 GB,采用 RTX 3060 GPU 进行训练,模型使用 PyTorch 框架,Python 版本为 3.11,完成了水稻苗技术网络的构建。

训练参数设置如下:周期数为 2 000,学习率为 0.000 01,批处理大小为 500,动量(Momentum)设置为 0.9,优化算法为亚当(Adam)优化算法。

3.2 相似度改进密度图的有效性

SGK 生成的主要参数有:高斯核标准差 σ、高斯核大小 $s\times s$、相似度的权重 α。对于高度 5 m 拍摄的水稻苗图像进行统计,其在截取的 1 000 像素 × 1 000 像素分辨率图像中,水稻苗的平均尺度大概在 20 像素 ×20 像素到 40 像素 ×40 像素之间,平均大小约为 900 像素。

为了验证本文采用相似性和高斯核结合的密度图生成方法的有效性,使用原始 MCNN 模型。对比原始最佳参数与 SGK 在水稻苗密度图生成中的精度。最终结果如图 6 所示,可以得出以下结论:本文采用原始标准卷积核训练 MCNN,通过对比实验选择最符合图像中水稻苗大小的 σ 和 s,其中变量$\sigma\in$

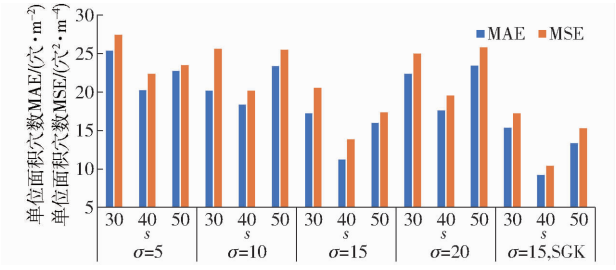

图 6 不同高斯核生成地面真值的模型性能

Fig. 6 Model performance of ground truth generation using different Gaussian kernel

$\{5,10,15,20\}$,$s\times s\in\{30\times30,40\times40,50\times50\}$。对比实验可以发现,当高斯核和水稻苗大小比较接近时,在不同的 σ 取值中预测效果都较好。通过对比实验,取 $\sigma=15$,$s=40$ 时,MAE 最小降至 11.17 穴/m^2,MSE 最小降至 13.83 穴^2/m^4。因此选择 $\sigma=15$,$s=40$ 作为卷积核的参数。

在对比取得的高斯核卷积参数的基础上,采用 SGK 进一步改进高斯核。由于水稻苗的非标准圆形形状,通过 SGK 改进后,改进高斯核生成的密度图的真值可以沿着水稻苗的纹理向外扩散。从图 6 可以看出,采用 SGK 后,不同尺寸的高斯核在 MCNN 中的表现都有所提升,当 $\sigma=15$,$s=40$ 时,MSE 取得最小值 9.17 穴^2/m^4。在后续实验中,始终采用$\sigma=15$,$s=40$ 的 SGK 生成地面真值。

3.3 DCFC 在不同层上的有效性

本文在 MCNN 模型的基础上添加了 DCFC 模块,增加了其感受野而不增加额外计算量。为了确定 DCFC 在网络结构中的合适位置,采用对比方法选择最佳安放位置。MCNN 模型共有 3 个分支,大尺度分支 Branch 1、中尺度分支 Branch 2 和小尺度分支 Branch 3,进入每个分支之后,均经过 3 个卷积核进行特征提取。通过将不同位置的卷积核替换成 DCFC 来进行对比实验,每个分支中对应的数字表示进入该分支后第几个卷积核替换成 DCFC,实验结果如图 7 所示。在 Branch 1 中使用 DCFC 时,与原始模型相比,MAE 都减少了,且 MSE 的降低幅度

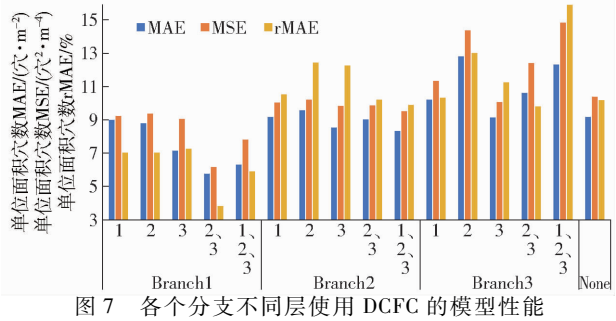

图 7 各个分支不同层使用 DCFC 的模型性能

Fig. 7 Model performance of using DCFC in different branches and layers

相对 MAE 和 rMAE 的减少幅度更大，当 2、3 层使用 DCFC 时，模型的精度最高。

在 Branch 2 和 Branch 3 中使用 DCFC 时，MAE、MSE 和 rMAE 都没有显著的下降，甚至出现上升的情况，说明在中尺度和小尺度中使用 DCFC 的效果不好，对水稻苗计数的预测效果不好。

当选取上述实验中效果最好的组合，即 Branch 1 中的第 2、3 层卷积替换成 DCFC，计算了各个分支中每一层的感受野大小，如表 2 所示。根据表 2 中的感受野可以看出，感受野最大达到 41 ×41。从上文中得知水稻苗的尺度大概在 20 像素 ×20 像素到 40 像素 ×40 像素之间，可以看出 DCFC 模块加强了模型对大水稻苗的感受，但是当 Branch 1 中 3 层都使用 DCFC 的时候，通过计算，最大感受野达到了 47 ×47，超出了本实验中最大的水稻苗尺寸，造成了信息的冗余。对 Branch 2 和 Branch 3 中的每一层添加 DCFC 时，可以看到相对于不添加 DCFC 模块，某些组合的 MSE 减小，但是基本上所有的 rMAE 都增加。从中可以得知，DCFC 可以增加每个卷积的感受野，但对每个感受野的预测精度下降。

表 2　膨胀卷积融合后不同分支和层的感受野大小

Tab. 2　Receptive field size of different branches and layers after dilated convolution

分支	卷积/池化核(k)	步数	感受野大小
Branch 1	Conv(9)	1	9
	Pooling(2)	1	10
	DCFC(7/13)	1	16/22
	Pooling(2)	1	17/23
	DCFC(7/13)	1	23/29/30/35
	Conv(7)	1	29/35/36/41
Branch 2	Conv(7)	1	9
	Pooling(2)	1	10
	Conv(5)	1	14
	Pooling(2)	1	15
	Conv(5)	1	19
Branch 3	Conv(5)	1	5
	Pooling(2)	1	6
	Conv(5)	1	10
	Pooling(2)	1	11
	Conv(5)	1	15

因此 Branch 1 中采用 DCFC 很好地扩大了感受野，感受到了尺度较大的水稻苗信息。说明当感受野增加时，模型对极端误差的处理效果更好。

3.4　基于尺度注意力机制模块在不同分支上的有效性

对注意力模块进行对比实验，验证算法中注意力模块的有效性，实验结果如表 3 所示。

表 3　不同注意力模块下模型性能

Tab. 3　Performance under different attention model

模块	MAE/（穴·m^{-2}）	MSE/（穴^2·m^{-4}）	rMAE/%
None	5.74	6.15	3.80
SAM	4.20	6.71	3.60
CAM	5.12	6.25	3.52
CBAM	4.93	6.10	3.20
SAM + CAM + CBAM	3.10	4.03	2.90

同时使用 SAM、CAM 和 CBAM 后，模型的 MAE 达到了 3.10 穴/m^2，与各部分单独使用模块相比有显著降低，MSE 下降得更加明显，可以看出注意力机制对较大误差的部分处理效果较好。

为了验证各个注意力机制在视觉问题上的应用效果，在 RSD2023 中选择对应场景（密集遮挡、复杂背景、多尺度）图像各 50 幅来对比并验证在不同场景下通道注意力机制的有效性，示例如图 8 所示，对数据集中对应场景进行分类，并对每种场景进行单独实验，来验证不同分支中注意力机制的有效性，实验结果如表 4 所示，在不同分支中加入不同注意力机制后，都不同程度上提高了计数精度，CAM 在复杂背景和密集场景中的效果更加显著，SAM 在多尺度场景下注意力模块的下降程度更高。说明了 CAM 和 SAM 解决复杂背景、密集遮挡和多尺度问题的有效性。

3.5　RCN 与其他模型的对比

采用数据集 RSD2023，本文选取当前主流的目标计数方法 MCNN、CSRNet、SCAR、TasselNetV2 进行对比研究，具体的参数设置均按照源码中的默认设置，结果如表 5 所示。

通过对网络结果的分析，可以看出 RCN 的综合误差更低，与 TasselNetV2 相比，本文模型

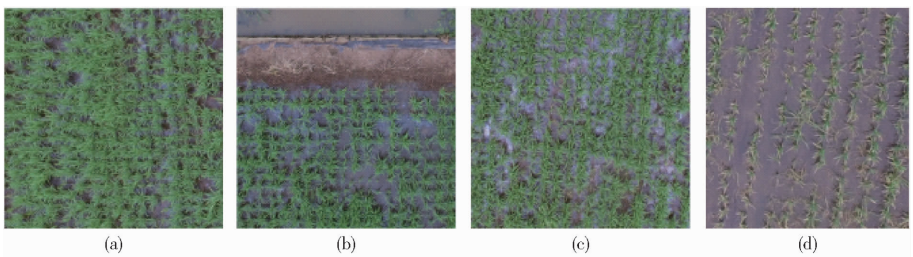

图 8　影响计数精度的不同场景示例

Fig. 8　Examples of different scenarios that affected counting accuracy

表 4 不同注意力模块在特定场景下的模型结果

Tab. 4 Model results based on different attention modules in specific scenarios

场景	模块	MAE/(穴·m^{-2})	MSE/(穴^2·m^{-4})	rMAE/%
复杂背景	CAM	5.15	7.16	3.45
	SAM	10.12	10.12	5.11
	None	10.01	13.58	12.15
多尺度	CAM	7.15	8.21	5.10
	SAM	4.91	5.15	3.66
	None	13.21	9.35	5.68
密集遮挡	CAM	6.15	8.14	7.15
	SAM	8.19	9.93	6.15
	None	11.19	13.93	10.15

表 5 几种经典密度图计数模型和 RCN 的模型结果

Tab. 5 Results of several classic density map counting models and RCN models

模型	MAE/(穴·m^{-2})	MSE/(穴^2·m^{-4})	rMAE/%
MCNN	11.17	13.83	12.54
CSRNet	10.25	11.24	10.22
SCAR	13.29	15.29	10.27
TasselNetv2	7.31	8.23	6.21
RCN(本文)	3.10	4.03	2.90

RCN 的 MAE、MSE、rMAE 分别降低了 4.21 穴/m^2、4.20 穴^2/m^4、3.31 个百分点。

图 9 为随机选取的 5 幅图像，从图中可以看出，RCN 对于水稻苗的中心识别更加明显，同时通过密度图计算出的水稻苗穴数更接近真实值。

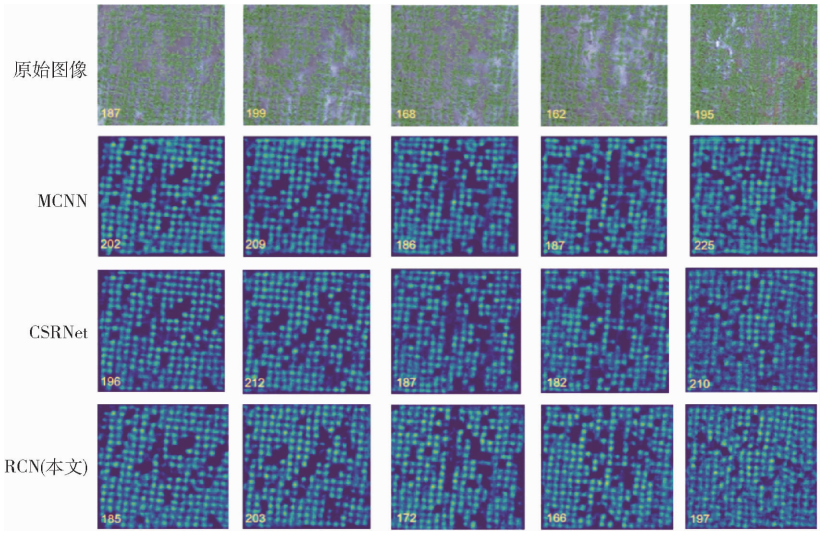

图 9 主流计数网络(MCNN、CSRNet)和 RCN 密度图以及计数结果

Fig. 9 Mainstream counting networks (MCNN, CSRNet) and RCN density maps and counting results

3.6 模型在大田中的应用

为了保证本文的 RCN 算法在整块大田中的适用性，本文采用对不同区域的图像进行识别，并根据图像的 GPS 定位对图像生成的密度值进行拼接，颜色越深代表种植越密集，最终效果图如图 10 所示，很好地展示了水稻种植密度的分布情况。

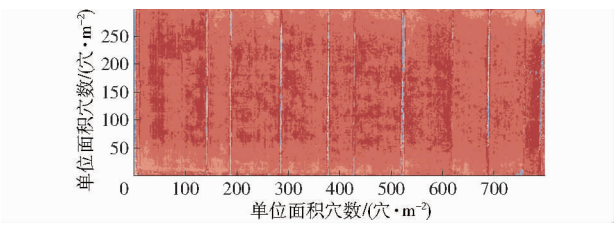

图 10 水稻整田的密度分布

Fig. 10 Density distribution of rice field preparation

对密度图进行后处理，计算出整田的水稻苗数量共为 15 035 穴。整田共 0.71 hm^2，机插时的行穴距为 25 cm × 16 cm，机插有效率为 0.95，可以计算出理论水稻苗的穴数为 15 841.25，与根据本实验计算出来的数量总体相近。

4 结论

(1)通过 SGK 生成的地面真值可以提高模型对水稻苗细节纹理的提取，使输出的图像更加贴合水稻苗的细节，可以使模型对水稻苗的提取更加准确。本文 SGK 主要基于相似度设计了一种精细化的高斯核。对于无人机在高度 5 m 上成像的水稻苗场景，采用大小为 40、σ 为 15 的改进高斯核其 MSE 能够降至 9.17 穴^2/m^4，与采用原始 MCNN 中的高斯核训练模型相比，MAE 降低了 16.21 穴/m^2。

(2)DCFC 可以增加模型的感受野，对不同尺度目标的鲁棒性更强，针对大田水稻苗期场景中，由于生长状况不一致，导致植物大小变化较多。为了增加对不同尺度的水稻苗的信息获取，本文根据 MCNN 原始架构，计算出了其对应各层各模块的感受野，通过 DCFC 模块使感受野的范围更符合水稻苗尺度范围，通过实验发现可以使得 MSE 从

9.17 穴^2/m^{4}降至 5.74 穴^2/m^4。

(3)针对背景复杂、密集遮挡、多尺度问题，在大尺度上使用通道注意力机制，提取更大尺度的水稻苗植株的上下文信息，并在小尺度上增加空间注意力机制，增加对较小水稻苗和复杂背景(水田反光、杂草、土壤)的识别精度。

(4)采用以上模块，田间水稻苗计数的精度得到了显著提高，其中精度最高 MAE 可以达到 3.10 穴/m^2，MSE 可以达到 4.03 穴^2/m^4，与 MCNN、CSRNet、SCAR、TasselNetv2 模型相比，MAE 分别降低 8.07、7.15、10.19、4.21 穴/m^2。MSE 分别降低了 9.80、7.21、11.26、4.20 穴^2/m^4。自制数据集的实验结果表明，本文提出的 RCN 模型在处理高密度水稻苗计数问题上的有效性。

参 考 文 献

[1] 国家统计局. 中国统计年鉴 2023[M]. 北京:中国统计出版社, 2023.

[2] 梁青铎. 施氮量和机插密度对水稻生长、产量和氮肥利用效率的影响[D]. 北京:中国农业科学院, 2020.
LIANG Qingduo. Effects of nitrogen application rate and planting density on growth, yield and nitrogen use efficiency of rice [D]. Beijing: Chinese Academy of Agricultural Sciences, 2020. (in Chinese)

[3] ALIPOUR F, MOBASSER H R. Effect of planting density on growth characteristics and grain yield increase in successive cultivations of two rice cultivars[J]. Agrosystems, Geosciences & Environment, 2021,4(4): 20213.

[4] PICKSON R, GUI P, CHEN A, et al. Climate change and food security nexus in Asia: a regional comparison[J]. Ecological Informatics, 2023,76:102038.

[5] ASLAN M, DURDU A, SABANCI K, et al. A comprehensive survey of the recent studies with UAV for precision agriculture in open fields and greenhouses[J]. Applied Sciences, 2022, 12(3):1047.

[6] ZHANG C, KOVACS J M. The application of small unmanned aerial systems for precision agriculture: a review[J]. Precision Agriculture, 2012, 13: 693 - 712.

[7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84 - 90.

[8] KITANO B T, MENDES C C, GEUS A R, et al. Corn plant counting using deep learning and UAV images[J]. IEEE Geoscience and Remote Sensing Letters, 2019(8):1 - 5.

[9] 王宏乐, 叶全洲, 王兴林,等. 基于 YOLOv7 的无人机影像稻穗计数方法研究[J]. 广东农业科学, 2023,50(7):74 - 82.
WANG Hongle, YE Quanzhou, WANG Xinglin, et al. Rice panicles counting method based on YOLOv7 using unmanned aerial vehicles images[J]. Guangdong Agricultural Sciences, 2023, 50(7): 74 - 82. (in Chinese)

[10] HANBURY A. A survey of methods for image annotation[J]. Journal of Visual Languages & Computing, 2008,19(5):617 - 627.

[11] WANG Y, QIN Y, CUI J, et al. Occlusion robust wheat ear counting algorithm based on deep learning[J]. Frontiers in Plant Science, 2021, 12:645899.

[12] LEMPITSKY V, ZISSERMAN A. Learning to count objects in images[J]. Advances in Neural Information Processing Systems (NIPS), 2010, 1:1324 - 1332.

[13] ZHANG Y, ZHOU D, CHEN S, et al. Single-image crowd counting via multi-column convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016: 589 - 597.

[14] LI Y, ZHANG X, CHEN D, et al. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 1091 - 1100.

[15] ZHANG J, ZHAO B, YANG C, et al. Rapeseed stand count estimation at leaf development stages with UAV imagery and convolutional neural networks[J]. Frontiers in Plant Science, 2020, 11: 617.

[16] WU J, YANG G, YANG X, et al. Automatic counting of in situ rice seedlings from UAV images based on a deep fully convolutional neural network[J]. Remote Sensing, 2019, 11(6): 691.

[17] XIONG H, CAO Z, LU H, et al. TasselNetv2: in-field counting of wheat spikes with context-augmented local regression networks[J]. Plant Methods, 2019, 15(1): 150.

[18] 饶雄飞,周龙宇,杨春雷, 等. 基于无人机多光谱影像和关键点检测的雪茄烟株数提取[J]. 农业机械学报, 2023, 54(3): 266 - 273.
RAO Xiongfei, ZHOU Longyu, YANG Chunlei, et al. Counting cigar tobacco plants from UAV multispectral images via key points detection approach[J], Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 266 - 273. (in Chinese)

[19] FENG A, LI H, LIU Z, et al. Research on a rice counting algorithm based on an improved MCNN and a density map[J]. Entropy, 2021, 23(6): 721.

(下转第 580 页)

HAN Yu, LIU Haoran, LIN Wenshu. Tree species identification based on improved YOLOv10 and UAV RGB imagery[J]. Forest Engineering, 2025,41(5): 922 - 935. (in Chinese)

[24] 王展, 吴浩东, 陈复扬. 基于改进 YOLO11 的太阳电池微缺陷检测方法[J]. 激光与光电子学进展,2025,62(20):1 - 16.
WANG Zhan, WU Haodong, CHEN Fuyang. Solar cell microdefect detection method based on improved YOLO11[J]. Laser and Optoelectronics Progress, 2025,62(20): 1 - 16. (in Chinese)

[25] TIAN Y, YE Q, DOERMANN D. YOLO v12: attention-centric real-time object detectors[J]. arXiv preprint arXiv: 2502.12524,2025.

[26] 邹荣, 李金炎, 王权, 等. 基于相机与激光雷达融合的番茄果实三维定位研究[J]. 农机化研究,2025,47(10):1 - 10,18.
ZOU Rong, LI Jinyan, WANG Quan, et al. Research on three-dimensional positioning of tomato fruit based on fusion of camera and LiDAR[J]. Journal of Agricultural Mechanization Research, 2025, 47(10): 1 - 10,18. (in Chinese)

[27] 闫英. 基于深度学习的樱桃番茄识别与定位技术研究[D]. 北京:北京信息科技大学,2024.
YAN Ying. Research on recognition and positioning technology of cherry tomatoes based on deep learning[D]. Beijing: Beijing Information Science & Technology University,2024. (in Chinese)

[28] 王元昊. 柑橘采摘机器人目标识别定位与末端执行器设计[D]. 重庆:西南大学,2024.
WANG Yuanhao. Target recognition and positioning for citrus harvesting robot and design of end-effector[D]. Chongqing: Southwest University,2024. (in Chinese)

[29] BISONG E. Matplotlib and seaborn[M]// Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners. Berkeley, CA: Apress,2019:151 - 165.

(上接第 567 页)

[20] YANG H, YUAN H, WANG Y, et al. Rape seedling density estimation in-field conditions based on improved multi-column convolutional neural network[J]. Agronomy Journal, 2024, 116(3): 810 - 825.

[21] CHEN Y, JIANG H, ZHANG X, et al. Refined feature fusion for in-field high-density and multi-scale rice panicle counting in UAV images[J]. Computers and Electronics in Agriculture, 2023, 211: 108032.

[22] LAN S, REN Z, WU Y, et al. SaccadeNet: a fast and accurate object detector[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10397 - 10406.

[23] WOO S, PARK J, LEE J Y, KWEON I S, Cbam: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.

[24] GAO J, WANG Q. SCAR: spatial-/channel-wise attention regression networks for crowd counting[J]. Neurocomputing, 2019, 363: 1 - 8.