doi:10.6041/j.issn.1000-1298.2026.17.024

基于特征重校准与多尺度融合的麦穗计数方法

鲍文霞　杨磊磊　程志友　胡根生　梁　栋
（安徽大学农业生态大数据分析与应用技术国家地方联合工程研究中心，合肥 230601）

摘要：在现代化精准农业中，无人机遥感技术凭借高效、低成本等优势，已成为麦穗计数任务的重要手段。然而，无人机影像下麦穗目标尺寸较小，且飞行高度变化导致麦穗尺度不一，同时麦穗在田间分布密集，这些都增加了麦穗计数的难度。针对以上问题，本文提出了一种基于特征重校准与多尺度融合的小麦麦穗计数网络（Feature recalibration and fusion network，FReCalNet），提升高密度分布与多尺度变化下的麦穗计数精度。具体而言，首先设计了多尺度特征金字塔注意力融合（Feature pyramid attention fusion，FPAttnFusion），加强特征层之间的信息交互，缓解因麦穗尺度变化导致的特征不匹配问题；其次，结合提出的特征重校准与融合模块（Feature recalibration and fusion，FReCal），进一步抑制因麦穗分布密集引起的特征干扰与漏检；最后，构建了多维特征对齐模块（Multi-dimensional feature alignment，MDFA），通过多尺度信息聚合提升模型在不同场景下的鲁棒性和泛化能力。大量实验结果表明，本文方法的平均相对误差（Mean relative error，MRE）、平均绝对误差（Mean absolute error，MAE）和均方根误差（Root mean square Error，RMSE）分别达到 12.30%、14.64 和 17.26，相比性能最优的主流算法 TPH－YOLO 分别降低了 4.5%、4.4% 和 7.4%；相比常用检测方法 FamNet 则分别降低了 10.0%、10.8% 和 11.6%。本研究为高密度麦穗场景下的精准计数提供了有效的技术路径，同时为无人机智能监测与农业精准管理奠定了基础。

关键词：小麦；无人机遥感；麦穗计数；特征重校准；多尺度融合
中图分类号：S127　　**文献标识码**：A　　**文章编号**：1000-1298(2026)17-0254-12　　**OSID**：

Wheat Ear Counting Method Based on Feature Recalibration and Multi-scale Fusion

Bao Wenxia　Yang Leilei　Cheng Zhiyou　Hu Gensheng　Liang Dong
(*National and Local Joint Engineering Research Center for Agricultural Ecological Big Data Analysis and Application Technology*, *Anhui University*, *Hefei* 230601, *China*)

Abstract: In modern precision agriculture, unmanned aerial vehicle (UAV) remote sensing technology has become an essential tool for wheat ear counting due to its high efficiency and low cost. However, UAV imagery presents several challenges: wheat ears appear small in size, their scales vary with flight altitude, and they are densely distributed in the field, all of which increase the difficulty of accurate counting. To address these challenges, a novel wheat ear counting network, feature recalibration and fusion network (FReCalNet) was proposed, which enhanced counting accuracy under conditions of high-density distribution and multi-scale variation. Specifically, a feature pyramid attention fusion (FPAttnFusion) module was designed to strengthen information interaction across feature layers, mitigating feature misalignment caused by scale variations of wheat ears. In addition, a feature recalibration and fusion (FReCal) module was introduced to suppress feature interference and missed detections arising from densely packed wheat ears. Finally, a multi-dimensional feature alignment (MDFA) module was constructed to aggregate multi-scale information and improve the model's robustness and generalization across diverse scenarios. Extensive experimental results demonstrated that the proposed

收稿日期：2025－05－09　修回日期：2025－06－20
基金项目：合肥市自然科学基金项目（HZR2411）、安徽省教育厅高校科研计划项目（2023AH050084）、安徽省自然科学基金项目（2208085MC60）和国家自然科学基金项目（62273001、32372632）
作者简介：鲍文霞（1980—），女，教授，博士，主要从事农业与生态视觉分析与模式识别研究，E-mail：bwxia@ahu.edu.cn
通信作者：程志友（1972—），男，教授，博士，主要从事农业生态大数据分析与应用研究，E-mail：czy@ahu.edu.cn

method achieved a mean relative error (MRE) of 12.30%, a mean absolute error (MAE) of 14.64, and a root mean square error (RMSE) of 17.26, representing reductions of 4.5%, 4.4%, and 7.4%, respectively, compared with that of the best-performing baseline method TPH - YOLO. Moreover, compared with the commonly used detection method FamNet, the proposed method achieved reductions of 10.0%, 10.8%, and 11.6%, respectively. The research result can provide an effective technical approach for accurate wheat ear counting in high-density scenarios and lay a solid foundation for intelligent UAV-based monitoring and precision agricultural management.

Key words: wheat; UAV remote sensing; ear counting; feature recalibration; multi-scale fusion

0　引言

在全球人口持续增长与粮食需求不断上升的背景下,提升农业生产效率已成为保障粮食安全的关键途径[1]。小麦作为主要粮食作物,其产量直接关系到国家粮食安全和农业经济效益[2]。其中,麦穗数量作为反映小麦生长状况与预测产量的关键参数,具有重要的研究价值。传统麦穗计数方法主要依赖人工标注,存在效率低下与稳定性差等问题。人工计数不仅耗时费力,还易受到操作者主观因素的影响,在实际应用中难以满足对高效率与高精度的要求。为提升计数任务的自动化程度与稳定性,研究人员逐步引入了机器学习方法[3],用以替代传统的人工手段,典型方法包括基于支持向量机(SVM)[4]、随机森林(RF)[5]和 K 近邻(KNN)[6]等分类器的麦穗检测,以及基于回归模型的麦穗数量预测。这些方法在一定程度上减少了人工干预,但是由于小麦生长环境较为复杂使得麦穗形态多变,传统机器学习方法难以全面捕捉麦穗的多尺度信息和深层语义特征,进而导致麦穗计数的精确度较低,难以应用在实际的农业系统之中。

随着深度学习技术的兴起,复杂的神经网络基于其强大的特征提取能力进而更有效地应对各种环境下的麦穗检测与计数任务。基于深度学习的方法不再依赖人工特征设计,而是通过训练数据自动学习有效的特征表示,显著提升了复杂场景下的识别与计数精度[7]。尤其是卷积神经网络[8-12](Convolutional neural network,CNN)的大范围应用,为麦穗计数提供了强有力的技术支持,为密集农田区域的麦穗计数奠定了理论基础。

尽管基于深度学习的方法在地面图像的麦穗计数任务中取得了显著进展,但图像采集与标注的成本依然较高。尤其在全监督学习框架下,模型训练效果严重依赖大量高质量的人工标注数据,而麦田场景中麦穗分布密集、尺度变化显著,手动采集与标注不仅耗时耗力,还容易因目标数量密集、边界模糊等问题导致标注精度下降,进而影响模型性能。为提升数据采集效率并拓展计数模型的应用范围,研究人员将无人机遥感技术(UAV remote sensing)[13-19]引入麦穗计数任务。相比地面拍摄,无人机具备灵活部署、覆盖面积大和图像分辨率高等优势,能够在短时间内高效采集大规模农田影像,显著降低数据获取成本。然而,针对无人机影像中的麦穗计数任务,仍面临诸多困难。一方面,麦穗目标在高密度分布和复杂背景下极易出现尺度变化大、遮挡严重、边界模糊等问题,导致传统基于边界框或点标注的方法效率低下且易出错;另一方面,不同田块之间麦穗排列方式和密集程度存在差异,使得模型在不同场景下的适应性受限。因此,如何在无人机高分辨率影像中有效感知密集麦穗区域,提升模型对尺度变化与遮挡问题的鲁棒性,成为当前麦穗计数领域亟需突破的关键问题。

为了应对麦田图像中麦穗目标密集分布、尺度变化大与遮挡严重等问题,研究者提出了多种创新方法[20-22]以提升计数的准确性与鲁棒性。这些方法在处理密集、遮挡严重、尺度多变的目标分布场景中展现出良好的潜力,但面对麦穗计数中特有的高度相似、局部特征微弱等挑战时,仍存在特征提取不足、匹配干扰等问题。

针对上述问题,本文结合深度学习与特征增强策略,设计一种适用于无人机图像的特征重校准和多尺度融合的小麦麦穗计数网络(Feature recalibration and fusion network,FReCalNet)。通过设计 FReCalNet 增强特征学习、跨尺度融合及目标特征强化能力,从而显著提升复杂背景下的麦穗计数表现。在此基础上设计出特征重校准(FReCal)引入通道重加权策略,在提升显著特征响应的同时有效抑制背景干扰,从而增强模型对关键目标的关注能力;而 FPAttnFusion 模块进一步结合特征金字塔结构[23]与注意力机制[24],使多尺度特征之间能够更充分地融合与表达,从而提高模型对不同尺度麦穗的适应性;与此同时,多维特征对齐(MDFA)模块则通过在空间和语义维度上的多方向特征对齐机制有效缓解麦穗在密集遮挡与形态变化等复杂情况

下的表征困难。

1 材料与方法

1.1 数据集构建

数据采集工作于 2024 年 5 月 1—15 日在 2 个小麦试验田地块的多个时段进行,拍摄时小麦正处于灌浆期。第 1 处试验田种植单一小麦品种 AK1009,拍摄高度分别为 3、4、5、6 m。拍摄过程中,相机视线与小麦冠层表面设置多个夹角,包括垂直视角(近 90°)以及一定范围内的倾斜视角(如 60°、75°等)。第 2 处试验田包含 AK1009、AK1401、AK1706 和 YKM222 4 个小麦品种,拍摄高度统一设置为 4 m。2 处试验田均使用配备 1/2 英寸 CMOS 传感器、等效焦距为 24 mm 的 DJI Mavic Air 2 型无人机进行图像采集,原始图像分辨率为 8 000 像素 × 6 000 像素。

所有采集图像均经过统一的预处理操作,包括图像裁剪、尺寸规范化以及数据增强。其中,图像尺寸的统一通过区域裁剪方式实现:即在原始 8 000 像素 ×6 000 像素分辨率图像上,按照固定区域从中心或覆盖麦穗区域裁剪出若干尺寸为 800 像素 × 600 像素的图像块,以确保图像中包含完整且清晰的麦穗区域。其中,数据增强方法采用颜色空间转换与对比度自适应增强相结合的策略。具体而言,首先将原始 RGB 彩色图像转换至 YCbCr[25] 颜色空间,对其中的亮度分量 Y 采用自适应对比度增强算法进行增强,以提升图像在复杂光照环境下的表现能力。随后将增强后的图像再转换回 RGB 空间,以保留色彩信息并增强图像的可识别性。原图和增强后的图像如图 1 所示。经过上述处理,最终获得分辨率为 800 像素 ×600 像素的图像共 2 060 幅。

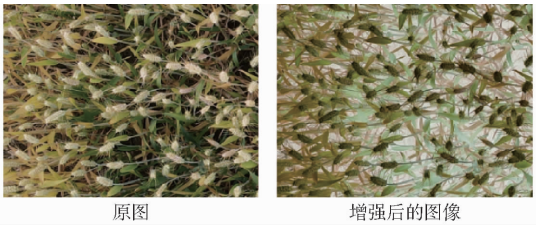


图 1 原图和增强后的图像对比

Fig. 1 Comparison between original image and enhanced image

为提升模型训练与评估的科学性,本文基于不同试验田的采集图像,构建了 2 个独立的数据集:Field1 - Height - Data 和 Field2 - Variety - Data。

(1) Field1 - Height - Data:该数据集来源于第 1 处试验田,包含 3、4、5、6 m 4 种不同高度的无人机图像,经过统一裁剪与预处理后共得到 1 560 幅图像(3 m:387 幅,4 m:393 幅,5 m:384 幅,6 m:396 幅)。

(2) Field2 - Variety - Data:该数据集来源于第 2 处试验田,专门针对 AK1009、AK1401、AK1706 和 YKM222 4 个小麦品种进行拍摄,图像经过裁剪与预处理后得到 500 幅图像,每个品种约 125 幅。

每个数据集均按比例 3∶1∶1 划分为训练集、验证集与测试集。不同高度和不同品种的小麦无人机示例图像分别如图 2、3 所示。

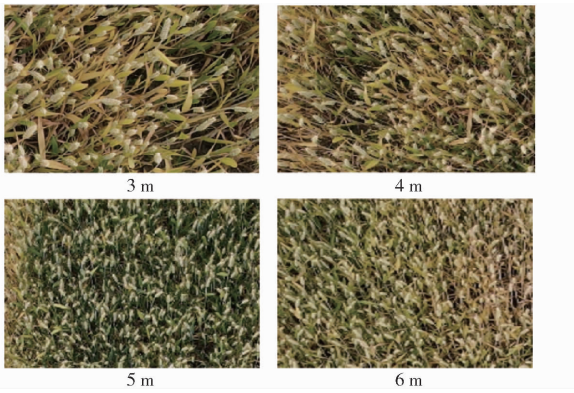


图 2 不同高度麦穗图像

Fig. 2 Images of wheat ears at different heights

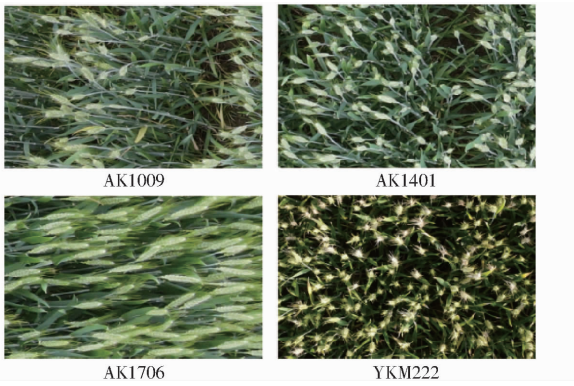


图 3 不同品种麦穗图像

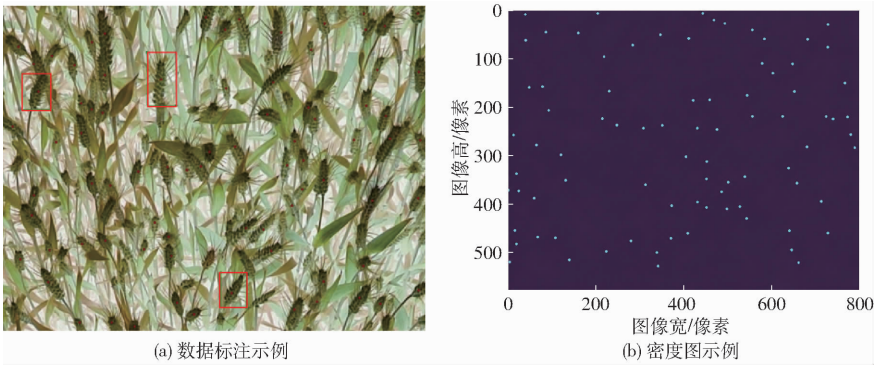
Fig. 3 Images of wheat ears of different varieties

1.2 数据标注

使用 LabelImg 标注软件,结合点标注与框标注 2 种方法进行标注。标注工作在数据经过裁剪和数据增强处理后进行。

首先采用点标注方式,对图像中所有麦穗进行逐一标记。标注点统一标注在每个麦穗的中心位置,即穗轴区域的中部位置,以作为密度图中高斯核函数的中心。这一策略有助于提高密度图的精度与稳定性。为保持数据标注的一致性,所有拍摄高度(3、4、5、6 m)和所有小麦品种(AK1009、AK1401、AK1706、YKM222)所对应图像均采用相同的点标注标准。整体数据集中共标注麦穗点 187 200 个,平均每幅图像标注 98 个点。此外,为增强模型在高密度场景中的特征学习能力,本文在每幅图像中人工选取 3 个典型麦穗区域进行框标注。典型麦穗的选

取依据包括：穗形清晰、无遮挡、形态完整且具代表性。框标注区域用于作为支持区域（Support regions），引导模型在训练过程中关注关键麦穗特征，提升其在密集背景下的判别能力。整个数据集中共构建了 4 680 个框标注样本，涵盖不同高度与不同品种图像，保证了样本分布的均衡性与代表性，标注数据示例如图 4 所示。由此进一步提升了模型的泛化性能和鲁棒性。



(a) 数据标注示例　　(b) 密度图示例

图 4　标注数据示例

Fig. 4　Annotated data example

点标注不仅用于麦穗位置的标记，还进一步用于密度图的生成，以增强监督信号，提高计数精度。密度图通过高斯核密度[26]估计（Gaussian kernel density estimation，KDE）生成，并保持与原始图像相同的分辨率（800 像素 × 600 像素），用于模型训练和验证。密度图能够更细粒度地描述麦穗的分布情况，为模型提供更加丰富的监督信息。

2　研究方法

2.1　FReCalNet 模型

FReCalNet 整体采用端到端的结构设计，集成了多个功能模块，涵盖从特征提取到目标计数的完整流程，如图 5 所示。该方法在 FamNet 架构的基础上，保留其密度图预测模块的解码结构，同时针对麦穗密集、遮挡严重、背景复杂等实际问题，设计了更具表达力的特征重校准模块（Feature recalibration and fusion，FReCal）、多尺度特征金字塔注意力融合（Feature pyramid attention fusion，FPAttnFusion）以及多维特征对齐模块（Multi-dimensional feature alignment，MDFA）。通过三者的协同作用，FReCalNet 能够有效抑制复杂背景带来的干扰，强化特征表达与上下文理解能力，显著提升麦穗目标的识别与计数精度。

在模型结构中，输入图像首先通过主干网络 ResNet50 进行特征提取，随后送入 FPAttnFusion 融合策略中。该策略集成了卷积块注意力机制[27]（Convolutional block attention module，CBAM）与特征金字塔网络[28]（Feature pyramid network，FPN），通过通道注意力与空间注意力机制的协同作用，有效突出关键区域特征、抑制背景噪声干扰，从而提升特征表示的判别性。同时，FPN 结构通过逐层上采样与横向连接的方式，实现不同尺度特征的有效融合，使模型兼具强语义理解能力与空间分辨能力。随后，融合后的特征被输入至 FReCal 模块，该模块利用深度卷积实现通道级的动态加权，对特征进行显著性增强与信息重校准，能够有效突出麦穗区域的显著特征，同时抑制复杂背景的干扰，从而提升模型对关键目标的感知能力。此外，FReCal 还融合了局

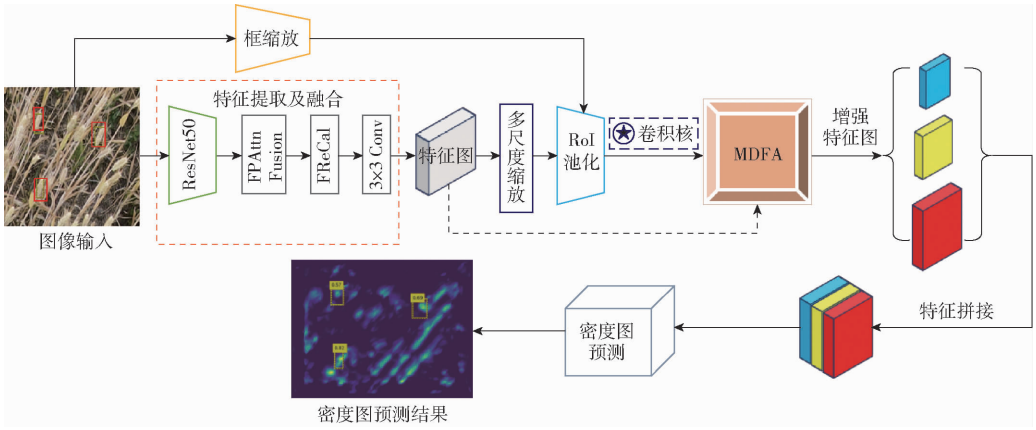


图 5　FReCalNet 整体结构图

Fig. 5　FReCalNet overall structure

部细节特征与全局上下文信息,优化特征分布,增强模型对密集区域中小目标的感知能力。在此基础上,增强后的特征图进一步通过一个 3×3 卷积层整合局部上下文信息与空间细节,有效平滑特征分布,为后续的密度图回归提供更加精细且稳定的特征表示。为了进一步增强模型对不同视角与姿态变化目标的感知能力,融合后的特征图经多尺度缩放和 RoI 池化被送入多维特征对齐模块(MDFA)进行处理。该模块通过对特征图进行水平翻转、垂直翻转及多角度旋转操作,生成多方向特征,以提升模型在非规则视角和复杂场景下的适应能力。经过统一尺度对齐与特征拼接后,多方向特征被输入卷积单元提取互信息,生成融合多方向信息的增强特征表示。在此基础上,增强特征被送入密度图预测模块,通过一系列卷积层逐步还原空间细粒度信息,最终输出密度图。密度图不仅能够准确描述目标在图像中的空间分布,还为后续的小麦穗数量估计提供了精确而可靠的依据,从而显著提升了整体计数性能。

2.2 小麦麦穗多尺度特征提取与重校准策略

2.2.1 多尺度特征融合策略

为了解决小麦麦穗计数中目标尺寸变化大、背景干扰强以及关键特征难以精准捕捉等问题,本文提出了一种多尺度特征融合策略(FPAttnFusion),旨在增强特征提取效果并优化多尺度信息融合过程。该策略的核心思想是在特征提取过程中引入注意力机制以突出关键区域、抑制田间背景干扰,同时结合特征金字塔结构融合不同尺度的语义与空间信息,从而提升模型对多尺度麦穗目标的感知能力。

图 6 展示了 FPAttnFusion 融合策略原理。首先,输入图像后送入 ResNet50 网络,依次通过 Conv1 至 Conv4 层进行层级化特征提取,其中 Conv3 和 Conv4 之后分别引入 CBAM,通过全局平均池化和最大池化计算通道权重,增强重要通道信息,同时利用卷积操作生成空间注意力权重,精准定位关键目标区域。随后,CBAM 模块输出的增强特征映射被分别记为 Lateral3 和 Lateral4,用于构建多尺度融合路径。二者作为中层和高层的代表性特征图,保留了不同语义层级的信息,并在注意力机制作用下聚焦于麦穗区域的显著特征。接着,Lateral3 和 Lateral4 被分别送入 FReCal 模块,以进一步进行特征变换与显著性增强操作,从而自适应地突出关键特征,抑制背景噪声干扰,为后续的密度图预测提供更具判别力的特征表示。在融合阶段,来自 Conv4 分支的高层次特征经过上采样匹配 Conv3 分支的尺度后,与其进行逐元素相加融合,使得低层的细节信息与高层的语义信息充分交互。融合后的特征进一步通过 FReCal 模块处理,强化特征表达能力,并最终经过 3×3 卷积提取精细特征,形成最终输出。FPAttnFusion 通过 CBAM 注意力机制赋予网络更强的特征感知能力,结合 FPN 结构促进跨层信息流动,并借助 FReCal 模块优化特征学习过程,从而在小目标密集、背景复杂的场景下展现出更强的鲁棒性和精度。特征融合过程,有助于提升模型对多尺度麦穗目标的表达能力,并增强其在不同光照和视角变化条件下的适应性。

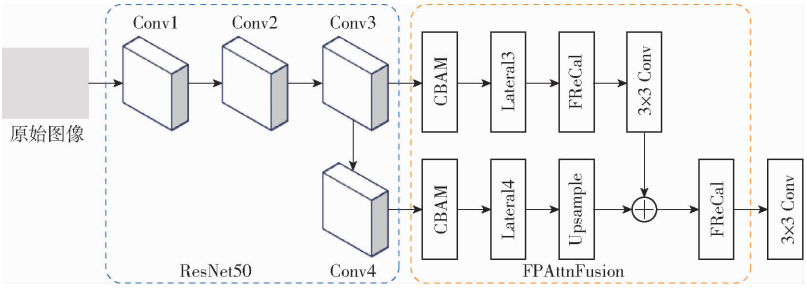


图 6 FPAttnFusion 融合策略结构图

Fig. 6 Structural diagram of FPAttnFusion fusion strategy

2.2.2 特征重校准模块

FReCal 模块通过引入特征分割与重排机制,增强特征间的结构关联性;利用深度卷积进行方向感知的特征增强,显著突出麦穗目标区域;同时通过门控机制对通道维度特征进行自适应加权融合,强化显著区域,抑制冗余背景干扰;最终借助残差连接保留原始信息,实现特征的高效重构与判别能力提升。

图 7 为特征重校准模块 FReCal 模块的结构。首先,给定初始输入特征 $x_{\text{init}} \in \mathbf{R}^{B\times C\times H\times W}$(其中 B、C、H、W 分别表示特征图的批次大小、通道数、高度和宽度),经过 Flatten 操作转换为一维特征映射,并通过特征分割(Split)划分为两部分:x_1(可处理特征)和x_2(未处理特征)。随后,这两部分特征分别经过 Reshape 调整,使其更适配后续计算。x_1 进入部分卷积(Pconv)进行细粒度信息提取,并在 Flatten 后输入线性变换(Linear1),以增强全局特征表达能力。经过线性变换后的特征再次进行特征分割(Split),部分特征送入深度卷积(DWConv),实现通

道独立的特征提取，从而提升局部信息的捕捉能力。DWConv 处理后的特征与另一部分特征进行融合，并通过门控机制（Gate mechanism）进行自适应加权，确保关键特征得到增强，同时抑制冗余信息。

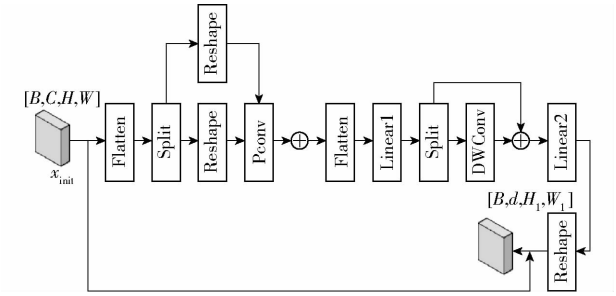


图 7　FReCal 模块结构图

Fig. 7　FReCal module structure diagram

在输出阶段，FReCal 模块通过 Reshape 调整特征尺寸，使其与原始输入特征（x_{init}）对齐，并通过残差连接相加，最终生成优化后的特征图。整个流程不仅优化了特征表达，提高了对小目标的关注度，还有效增强了模型的稳定性和泛化能力。实验结果表明，FReCal 模块能够显著提升小麦麦穗计数任务的准确性，并增强模型在复杂背景下的鲁棒性。

FReCal 的整体处理流程如图 8 所示，包含特征展平、分割、卷积增强、特征融合与残差连接等关键步骤，旨在对初始特征进行重校准并提升其对目标区域的响应能力。

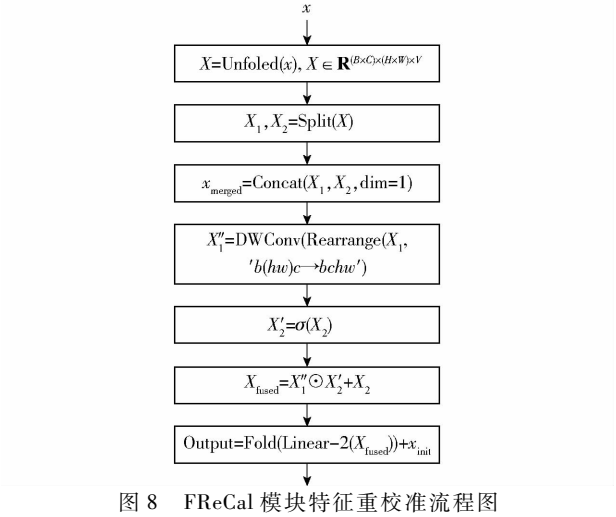


图 8　FReCal 模块特征重校准流程图

Fig. 8　Feature recalibration flowchart of FReCal module

2.2.3　多维特征对齐模块

多维特征对齐模块作为一种针对目标方向变化设计的特征增强机制，旨在提升模型在不同视角、小样本场景及复杂背景下的目标识别能力。该模块通过多方向变换、自适应尺度校准和特征融合，增强特征表达的丰富性，使模型能够更加鲁棒地应对目标方向、尺度和背景变化带来的挑战，特别适用于无人机影像、遥感图像和小样本目标检测等任务。

如图 9 所示，MDFA 首先采用一系列方向变换（水平翻转、垂直翻转、旋转 90°/180°/270°），模拟目标在真实场景中的不同朝向，以减少视角偏差对识别精度的影响，提高模型对多视角目标的感知能力。为了确保变换后的特征与原始特征保持一致，MDFA 进一步引入尺度自适应校准机制，统一调整不同方向变换后的特征尺度，避免因尺度不匹配导致的信息丢失，从而保证特征融合时的一致性。

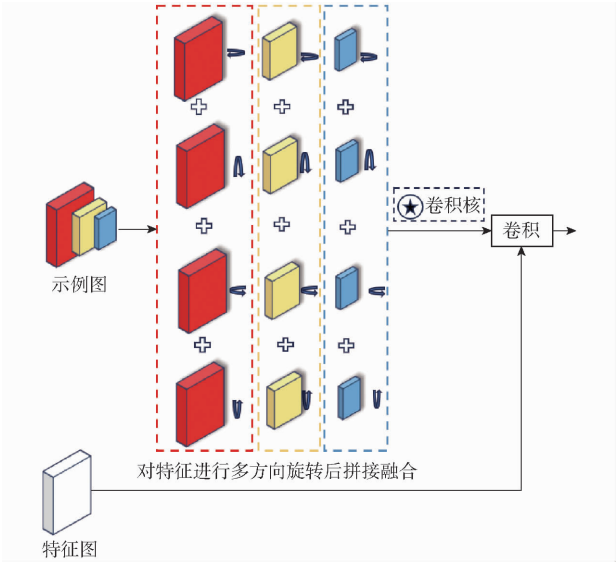


图 9　MDFA 模块结构图

Fig. 9　MDFA module structure diagram

在特征融合过程中，MDFA 结合通道注意力（Channel attention）和空间注意力（Spatial attention）机制，使不同方向变换后的特征能够自适应地分配权重，突出最具判别力的方向特征，同时有效抑制冗余信息和背景干扰。这一过程动态增强目标区域的特征表示，提高检测鲁棒性。随后，MDFA 采用特征拼接与 3×3 卷积融合，进一步提取跨方向的特征互信息，确保不同视角的特征能够高效融合，同时保持特征表达的紧凑性和判别性，使得模型即便在小样本学习场景下，仍能获得高质量的特征表示。

此外，MDFA 可与融合通道和空间注意力的特征金字塔网络（FPAttnFusion）互补使用，以增强模型在不同尺度上的目标检测能力。FPAttnFusion 结构能够进一步整合 MDFA 处理后的特征，使其与高层次的语义信息和低层次的细节信息充分融合，从而确保不同尺度的目标都能被高效检测，尤其在小目标密集场景下表现卓越。

2.3　模型优化目标与密度图预测模块设计

为实现对无人机图像中麦穗数量的准确估计，本文采用密度图回归（Density map regression）作为主要的计数策略。整个网络的优化目标是通过学习将输入图像映射为对应的密度图，使得密度图中所有

像素的积分值尽可能接近图像中麦穗的真实数量。

2.3.1 优化目标

在训练过程中,设输入图像为 I,其对应的真实密度图为 D^{gt},网络预测的密度图为 D^{pred},则优化目标函数为均方误差损失(Mean squared error, MSE)L_{MSE}

$$L_{MSE}=\frac{1}{N}\sum_{i=1}^{N}\| D_i^{pred}-D_i^{gt}\|_2^2$$

其中 N 为训练图像数量,D_i^{gt} 和 D_i^{pred} 分别表示第 i 幅图像的真实密度图和预测密度图。该损失函数能够有效约束网络输出与真实分布之间的差异,优化目标即为最小化该损失函数。

此外,为提升预测精度,后续实验还引入了平均绝对误差(MAE)、均方根误差(RMSE)和平均相对误差(MRE)作为评价指标,以综合衡量模型性能。

2.3.2 密度图预测模块

密度图预测模块位于整个网络结构的末端,主要用于将融合后的高维特征映射回归为与输入图像尺寸一致的密度图。该模块直接采用 FamNet[23] 中提出的解码结构,由多层卷积与上采样操作构成,能够有效适应目标密集、边界模糊等小麦麦穗计数场景中的困难。本文在该模块设计上未作改动,具体结构细节可参考原文[23]。

3 实验

本文的模型训练参数设置为:Adam 优化器,学习率为 0.000 01,训练轮数为 50,批量大小为 1。实验的硬件条件为:Intel(R) Xeon(R) Gold 6226 CPU @ 2.70 GHz,NVIDIA GeForce RTX 3080。

3.1 消融实验

为验证本文所提出模块在麦穗计数任务中的有效性,本文选取 FamNet[23] 作为基线网络进行对比实验。FamNet 作为一种在密集目标计数任务中具有良好基础性能的方法,能够在复杂场景中提取一定的目标特征并完成计数预测。消融实验结果如表 1 所示。FamNet 在未引入本文设计的模块时,其在麦穗计数任务中取得的平均绝对误差(MAE)为 16.42,均方根误差(RMSE)为 19.52,平均相对误差(MRE)为 13.81%。

表 1 消融实验结果

Tab.1 Results of ablation experiments

FReCal	FPAttnFusion	MDFA	MRE/%	MAE	RMSE
×	×	×	13.81	16.42	19.52
√	×	×	12.63	14.97	17.76
√	√	×	12.32	14.68	17.30
√	√	√	12.16	14.64	17.26

注:“√”表示添加该模块,“×”表示不添加该模块。

随着各模块的逐步加入,模型性能持续优化。引入 FReCal 模块后,得益于其在通道和空间维度对特征进行动态重标定的机制,模型对麦穗目标区域的感知能力显著增强,MAE 降至 14.97,RMSE 降至 17.76,MRE 也下降至 12.63%,表明 FReCal 能够有效提升特征表达能力。进一步添加 FPAttnFusion 模块后,模型融合了 FPN 的多尺度上下文信息与 CBAM 注意力机制,有效增强了特征融合能力与空间聚焦性。此阶段 MAE 和 RMSE 分别为 14.68 和 17.30,MRE 为 12.32%,虽略高于仅引入 FReCal 时的表现,但整体依然优于原始 FamNet,说明该模块在特定情境下可对特征感知形成补充。当 3 个模块全部集成后,模型在特征提取、注意力引导与跨尺度融合等方面协同增强,最终 MAE 显著下降至 14.64,RMSE 降至 17.26,MRE 进一步降至 12.16%,相较于 FamNet 分别降低了 47.1%、41.1% 和 11.9%。该结果充分验证了所提出模块在提升小样本条件下麦穗计数精度方面的显著效果。

从理论角度来看,FReCal 模块增强了模型对显著通道和空间区域的感知能力,FPAttnFusion 模块提升了多尺度特征的融合效率并引导模型聚焦关键区域,而 MDFA 模块则进一步加强了跨维度信息的表达与整合。三者相互协同,在提升模型感知能力与泛化性能方面发挥了重要作用。

3.2 不同计数网络性能对比结果

为了验证本文方法的有效性,将所提出的 FReCalNet 与现有主流的计数方法 FamNet、TPH-YOLO、BMNet、SAFECount、Sam[23] 进行了对比实验,结果如表 2 所示。从表 2 可以看出,FReCalNet 的 MAE、RMSE 和 MRE 3 项指标均优于其他对比方法,且在不同高度条件下均表现出较强的稳定性和泛化能力。具体而言,相较于 FamNet,FReCalNet 的 MAE 由 16.42 降至 14.64,下降了 10.84%;RMSE 从 19.52 降至 17.26,下降 11.55%;MRE 从 13.80% 降至 12.30%,下降 10.87%,体现出在多尺度输入条件下的显著精度提升。进一步地,FReCalNet 相比 BMNet 和 SAFECount,MAE 分别降低了 24.54% 和 24.38%,RMSE 分别降低了 31.78% 和 24.17%,MRE 分别下降 24.54% 和 24.37%,充分证明了本文方法在复杂场景下的强大特征表达与泛化能力。此外,从对 Sam 模型的结果分析来看,其 MRE、MAE 和 RMSE 指标均远高于其他方法,可能由于其在密集场景下的目标感知和特征分离能力较弱,导致预测误差较大。相比之下,FReCalNet 借助 FReCal、FPAttnFusion 和 MDFA 等模

块的协同作用，显著增强了模型对多尺度与密集区域目标的建模能力，使其在不同高度、不同视角下依然保持较高的计数准确性。实验结果表明，FReCalNet 在 Field1 – Height – Data 数据集的多高度条件下展现出优越的性能与稳定性，进一步验证了其在处理高密度目标和多尺度变化场景中的有效性。

表 2　不同模型在 Field1 – Height – Data 数据集上的实验比较

Tab. 2　Experimental comparison of different counting models on Field1 – Height – Data dataset

检测网络	MRE/%	MAE	RMSE
FamNet	13.80	16.42	19.52
TPH – YOLO	12.88	15.32	18.64
Sam	31.38	37.36	46.52
BMNet	16.30	19.40	25.31
SAFECount	16.27	19.36	22.76
FReCalNet	12.30	14.64	17.26

对 Sam 模型进一步分析表明，其 MAE 和 RMSE 远高于其他方法，可能由于在密集场景下特征提取能力有限，导致计数误差较大。相比之下，FReCalNet 通过结合 FReCal、FPAttnFusion 和 MDFA 等模块，提升了对密集目标和多尺度特征的建模能力，使得模型在不同目标密度和视角变化条件下依然能够保持较高的计数精度。此外，在 Field1 – Height – Data 数据集上的实验结果进一步验证了 FReCalNet 的有效性。相较于 FamNet，FReCalNet 的 MAE 和 RMSE 指标改进表明，FamNet 在多尺度特征提取和密集目标建模方面存在一定局限，导致其在复杂场景下的计数精度下降。与 BMNet 和 SAFECount 相比，FReCalNet 依托 FReCal、FPAttnFusion 和 MDFA 等模块，实现了更优的特征表达与融合能力，能够更好地适应不同目标密度和视角变化，从而在复杂背景下提供更准确的计数结果。

图 10 展示了不同计数网络在密度图预测结果可视化图示例，包括 FamNet、TPH – YOLO、Sam、BMNet、SAFECount 以及本文提出的 FReCalNet。图 10a 上图为对应的原始图像，可作为密度图预测结果的参考基准。图 10a 下图为目标区域的真实标注点图（蓝色点表示人工标注的麦穗中心）。从可视化结果来看，传统方法（如 FamNet 和 BMNet）在麦穗密集区域容易出现密度估计模糊或高估现象，特别是在背景较为复杂或麦穗重叠区域，预测的密度图出现较多虚假响应。Sam 模型的表现相对平衡，但在细粒度区域的响应仍不够集中，存在一定漏检现象。TPH – YOLO 在处理高密度和复杂背景区域的表现优于传统回归模型，其结构中的 Transformer 模块提升了对长距离依赖关系的建模能力，使得密度图预测更加稳定准确，尤其在边界区域具备更好的分离效果。然而，其在部分低密度区域的响应略显不足，存在一定低估趋势。相比之下，本文提出的 FReCalNet 在多个关键区域展现出更清晰、集中的响应分布，能够更准确地对麦穗进行密度估计。所生成的密度图在空间分布和目标数量上与真实标注图具有高度一致性，特别是在目标密集区域，能够准确区分相邻目标，有效抑制由复杂背景带来的虚假响应，体现出更强的特征感知与空间分布建模能力。此外，图 10 中黄色方框区域标注了典型难例区域，可直观看出 FReCalNet 对目标边界的识别更加清晰，预测结果更加贴合真实分布。

为更直观展现不同方法在稠密分布和遮挡情况

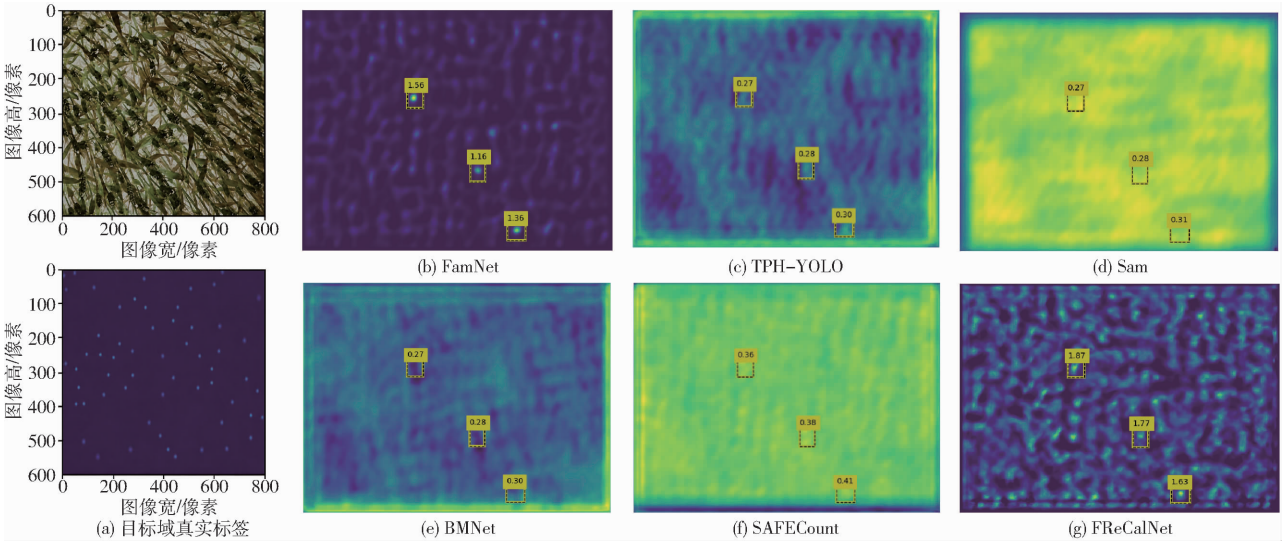


图 10　不同网络检测结果的可视化图像

Fig. 10　Visualization images of different network detection results

下麦穗数目预测细节与差异,图 11 在图 10 所示原始图像基础上,对各方法生成的密度图进行了叠加可视化展示。由图 11 可以看出,FamNet 与 SAFECount 在图像中呈现出较为均匀的大面积密度响应,尽管覆盖范围较广,但在麦穗存在部分重叠或密集区域易出现“过度预测”现象,导致误检数量明显增加。TPH-YOLO 和 BMNet 在目标密集区域的响应较弱,存在多个真实麦穗未能有效预测的情况,漏检现象较为突出,尤其在左下角和图像边缘区域更为明显。Sam 模型在响应聚集性方面表现一般,边界模糊、密度集中度不足,难以准确区分相邻麦穗。相比之下,FReCalNet 在密集区域内产生粒状、清晰且不重叠的高亮响应点,准确捕捉大部分麦穗目标,同时有效抑制背景干扰。尤其是在图像中央及右下方麦穗遮挡严重区域,FReCalNet 依然保持良好的结构感知能力,预测结果更贴近真实分布,展现出更强的遮挡鲁棒性与复杂背景下的泛化能力。

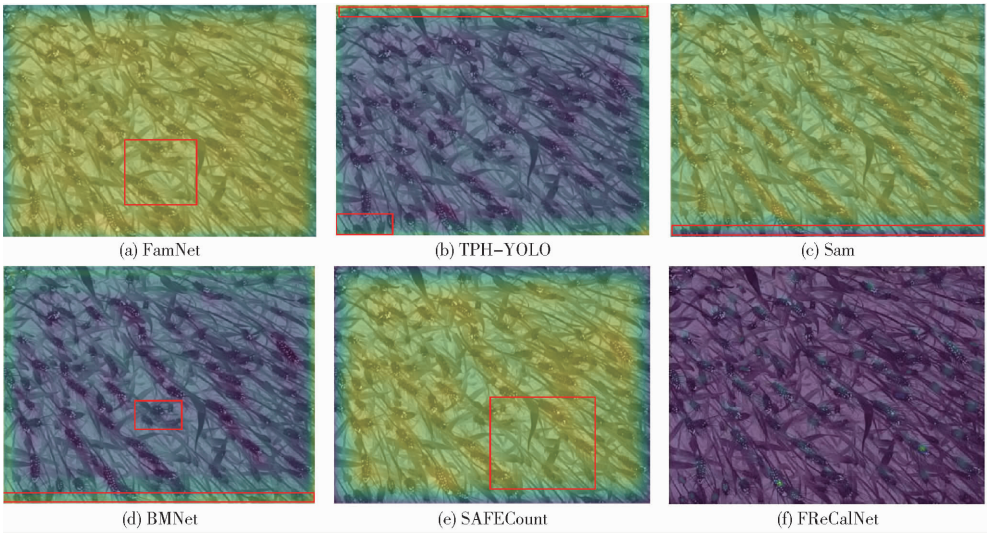


图 11 不同方法下预测密度图叠加可视化

Fig. 11 Comparison of predicted density map overlays across different methods

3.3 不同模型在 Field2-Variety-Data 数据集上的对比

为了评估不同模型泛化性能,对比了 6 种模型在混合品种麦穗图像 Field2-Variety-Data 数据集上的计数精度。实验结果如表 3 所示。

表 3 6 种模型在混合品种麦穗图像下计数结果分析

Tab. 3 Analysis of counting results of six models in mixed variety wheat ear images

检测网络	MRE/%	MAE	RMSE
FamNet	12.84	12.59	15.28
TPH-YOLO	16.29	16.14	19.38
Sam	25.70	28.19	30.58
SAFECount	14.35	15.72	17.08
BMNet	15.66	16.33	18.64
FReCalNet	9.66	8.68	11.50

从表 3 来看,Sam 模型的误差最高,MRE 达到 25.70%,MAE 和 RMSE 分别为 28.19 和 30.58,表明其在应对不同品种数据集时特征提取能力较弱,导致计数精度大幅下降。BMNet 和 SAFECount 的误差相对较低,其中 SAFECount 的 MRE 为 14.35%,MAE 为 15.72,RMSE 为 17.08,表现优于 BMNet(MRE 为 15.66%,MAE 为 16.33,RMSE 为 18.64),但两者的计数精度仍不够理想。TPH-YOLO 的 MAE 和 RMSE 分别为 16.14 和 19.38,其表现略逊于 FamNet 和 FReCalNet,说明其在处理多品种麦穗图像时存在一定的误检与低估问题,尤其在种类差异较大的样本中,其密度图响应存在不均匀现象。相比之下,FamNet 模型的 MRE 为 12.84%,MAE 为 12.59,RMSE 为 15.28,说明其在不同品种数据集上具有一定的泛化能力。然而,FReCalNet 模型在所有模型中表现最佳,其 MRE 显著降低至 9.66%,MAE 和 RMSE 分别为 8.68 和 11.50,远低于其他模型,表明其在面对品种差异和背景变化时依然能够保持较高的计数精度与稳定性。

在图像预测结果可视化分析方面,图 12 展示了多个样本在不同计数网络下的密度图预测效果对比。每行代表一个原始图像及其对应的真实密度图,多个模型(包括 FReCalNet、FamNet、Sam、BMNet、SAFECount 以及 TPH-YOLO)的预测结果。从可视化结果可以看出,FReCalNet 能够在大多数样本中生成与真实标签分布高度一致的密度图,其预测区域对麦穗分布响应准确,密度图边界清晰,且误检率较低,展现出较强的目标表征能力与计数精度。相比之下,FamNet 和 BMNet 在密集区域容易出

现预测扩散现象，导致计数偏高；SAFECount 模型则表现出保守趋势，部分样本中存在明显的低估；Sam 模型在背景复杂或纹理模糊的场景下，其预测边界模糊，存在漏检风险；而 TPH－YOLO 在遮挡区域的响应保持一定稳定性，但在处理混合品种密度差异较大的区域时仍存在局部误判的问题，其密度图响应有时出现离散或欠集中的现象，影响整体预测准确性。尤其在具有显著遮挡、背景杂乱或视角倾斜等复杂场景中，FReCalNet 模型仍保持稳定的密度图生成效果，显示出良好的鲁棒性与泛化能力。

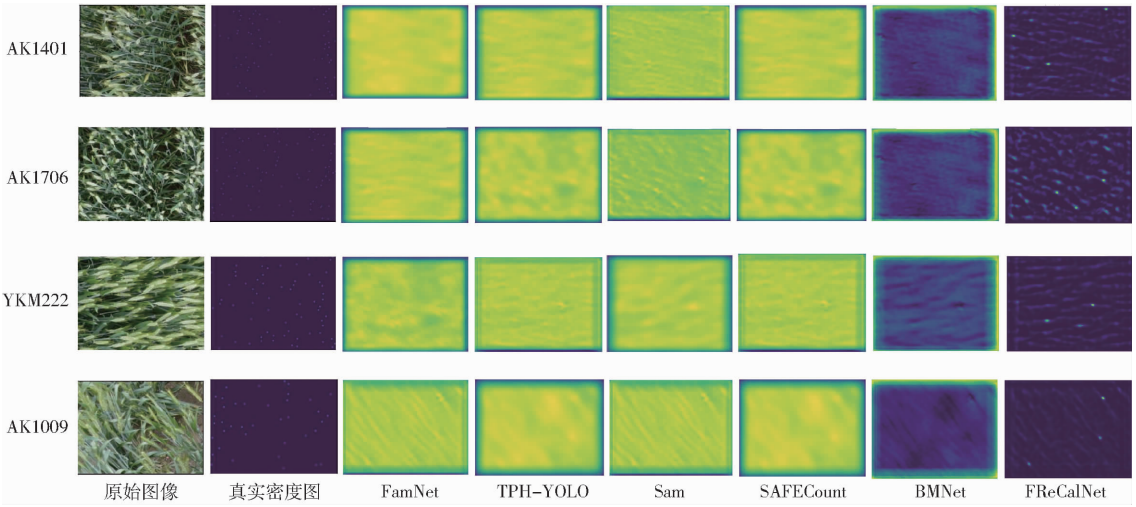


图 12　6 种模型在混合品种麦穗图像下计数结果可视化

Fig. 12　Visualization of counting results of six models in mixed variety wheat ear images

3.4　FReCalNet 在不同品种麦穗数据上的计数结果

为进一步分析 FReCalNet 在不同麦穗品种下的适应能力，在 Field2－Variety－Data 数据集的多个品种子集上对 FReCalNet 进行了评估，结果如表 4 所示。

表 4　FReCalNet 在不同品种麦穗上计数结果比较

Tab. 4　Comparison of FReCalNet counting results on different varieties of wheat ears

麦穗品种	MRE/%	MAE	RMSE
AK1009	9.49	8.54	9.61
AK1401	12.01	10.81	12.43
AK1706	10.70	9.63	11.54
YKM222	9.00	8.10	9.59

从表 4 可看出，模型在 WheatCount－Variety 数据集中不同品种上的表现存在一定差异。具体而言，AK1009 和 YKM222 的预测精度较高，MRE 分别为 9.49% 和 9.00%，MAE 分别为 8.54 和 8.10，RMSE 分别为 9.61 和 9.59，表明模型在这些品种上能够较为稳定地进行计数，误差较小。这主要得益于这 2 个品种的麦穗形态较为清晰，背景干扰较少，使得模型能够提取到更具代表性的稳定特征。相较之下，AK1401 的预测误差最大，MRE 高达 12.01%，MAE 为 10.81，RMSE 为 12.43，明显高于其他品种，说明模型在该品种上的计数任务存在一定困难。可能的原因是 AK1401 的麦穗形态复杂、间距密集，且样本数量相对较少，导致模型的泛化能力受限，从而影响计数精度。此外，AK1706 在测试中的表现处于中间水平，MRE 为 10.70%，MAE 为 9.63，RMSE 为 11.54，说明该品种的特征复杂度适中，模型的计数精度亦表现为中等。本实验仅选用 FReCalNet 作为分析对象，主要目的是探究其在复杂形态差异下的表现差异，而非进行横向对比。其他模型由于结构适配性或样本约束等问题，在子集实验中的复现存在一定困难，后续工作将进一步拓展至多模型分析，以验证方法的普适性与稳健性。

综合来看，尽管改进模型在大多数品种上取得了较好效果，但对于 AK1401 仍存在进一步优化空间。未来可通过增加该品种样本数量、进行品种特定的数据增强，或引入自适应调整机制来提升计数精度。

3.5　FReCalNet 在不同飞行高度下麦穗数据集上的计数结果

为探究模型在不同拍摄视角与尺度变化下的计数能力，基于 Field1－Height－Data 数据集对 FReCalNet 在不同拍摄高度时进行了分析，实验结果如表 5 所示。

表 5　FReCalNet 在不同飞行高度下的计数结果比较

Tab. 5　Comparison of FReCalNet's counting results at different flight altitudes

飞行高度/m	MRE/%	MAE	RMSE
3	10.11	8.59	9.89
4	9.46	8.27	9.05
5	10.20	9.21	9.72
6	10.73	9.50	13.13

实验结果表明,模型在不同高度拍摄条件下的计数精度呈现出一定趋势变化。4 m 高度的数据集表现最佳,MRE 为 9.46%,MAE 为 8.27,RMSE 为 9.05,说明该高度下的麦穗特征清晰,背景干扰较少,模型能够稳定提取有效特征。相比之下,3、5 m 高度的数据集表现略逊。3 m 条件下,MRE 为 10.11%,MAE 为 8.59,RMSE 为 9.89;而 5 m 条件下,MRE 为 10.20%,MAE 上升至 9.21,RMSE 为 9.72。这一结果表明,较近(3 m)或略远(5 m)的视角可能使得图像中的目标出现局部遮挡、角度变化或清晰度不足,影响模型的识别和计数能力。值得注意的是,在 6 m 高度条件下,模型的误差显著上升,MRE 达到 10.73%,MAE 增至 9.50,RMSE 更是上升至 13.13,成为所有高度中表现最差的情况。这可能是由于较远的飞行高度导致麦穗目标在图像中的尺寸显著减小,局部细节信息缺失,同时密集目标之间的重叠增强,进一步增加了特征提取与目标判别的难度,造成较大的计数误差。后续研究可以通过引入超分辨率图像重建方法、增强多尺度特征提取能力,或结合多高度联合训练策略,进一步提升模型在多尺度场景下的表现与稳定性。

4 结束语

针对无人机采集中因飞行高度变化引发的目标尺度差异及麦穗密集分布带来的检测困难,提出了一种基于特征重校准与多尺度融合的麦穗计数网络(FReCalNet)。该方法设计并结合了 FPAttnFusion、FReCal 和 MDFA 模块,通过引入多尺度特征提取与注意力机制,有效加强了特征层间的信息交互,抑制了因密集分布引起的特征干扰,从而显著缓解了目标尺度变化与密集遮挡造成的漏检问题。大量实验定量结果表明,所提 FReCalNet 模型的平均相对误差(MRE)、平均绝对误差(MAE)和均方根误差(RMSE)分别达到了 12.30%、14.64 和 17.26;与性能最优的主流算法 TPH - YOLO 相比,上述误差分别降低了 4.5%、4.4% 和 7.4%,与常用检测方法 FamNet 相比则分别降低了 10%、10.8% 和 11.6%,充分验证了其在高密度与多尺度变化条件下的有效性与优越性。

参 考 文 献

[1] 邓秀新. 现代农业与农业发展[J]. 华中农业大学学报:社会科学版, 2014(1):1 - 4.
Deng Xiuxin. Modern agriculture and agricultural development[J]. Huazhong Agricultural University Journal: Social Science Edition, 2014(1): 1 - 4. (in Chinese)

[2] 赵广才, 常旭虹, 王德梅, 等. 小麦生产概况及其发展[J]. 作物杂志, 2018, 34(4):1 - 7.
Zhao Guangcai, Chang Xuhong, Wang Demei, et al. General situation and development of wheat production[J]. Crops, 2018, 34(4): 1 - 7. (in Chinese)

[3] 石弘一. 机器学习综述[J]. 通讯世界, 2018, 25(10):253 - 254.

[4] 范梦扬, 马钦, 刘峻明, 等. 基于机器视觉的大田环境小麦麦穗计数方法[J]. 农业机械学报, 2015, 46(增刊 1):234 - 239.
Fan Mengyang, Ma Qin, Liu Junming, et al. Counting method of wheat ear in field based on machine vision technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(S1): 234 - 239. (in Chinese)

[5] 李欣海. 随机森林模型在分类与回归分析中的应用[J]. 应用昆虫学报, 2013, 50(4):1190 - 1197.
Li Xinhai. Using random forest for classification and regression[J]. Chinese Journal of Applied Entomology, 2013, 50(4): 1190 - 1197. (in Chinese)

[6] Zhang S, Li X, Zong M, et al. Learning K for KNN classification[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2017, 8(3): 1 - 19.

[7] Alzubaidi L, Zhang J, Humaidi A J, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions[J]. Journal of Big Data, 2021, 8: 1 - 74.

[8] 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: 225 - 234.

[9] Zhang R, Jia Z W, Wang R B, et al. Phenotypic parameter extraction for wheat ears based on an improved Mask - RCNN algorithm[J]. INMATEH-Agricultural Engineering, 2022, 66(1):267 - 278.

[10] 丁国辉, 许昊, 温明星, 等. 基于经济型低空无人机对小麦重要产量表型性状的多生育时期获取和自动化分析[J]. 农业大数据学报, 2019, 1(2):19 - 31.
Ding Guohui, Xu Hao, Wen Mingxing, et al. Developing cost-effective and low-altitude UAV aerial phenotyping and automated phenotypic analysis to measure key yield-related traits for bread wheat[J]. Journal of Agricultural Big Data, 2019, 1(2): 19 - 31. (in Chinese)

[11] Yang B, Gao Z, Gao Y, et al. Rapid detection and counting of wheat ears in the field using YOLOv4 with attention module [J]. Agronomy, 2021, 11(6): 1202.

[12] 鲍文霞, 谢文杰, 胡根生, 等. 基于 TPH YOLO 的无人机图像麦穗计数方法[J]. 农业工程学报, 2023, 39(1): 155 – 161.
Bao Wenxia, Xie Wenjie, Hu Gensheng, et al. Wheat ear counting method in UAV images based on TPH – YOLO[J]. Transactions of the CSAE, 2023, 39(1): 155 – 161. (in Chinese)

[13] 李冰, 刘镕源, 刘素红, 等. 基于低空无人机遥感的冬小麦覆盖度变化监测[J]. 农业工程学报, 2012, 28(13): 160 – 165.
Li Bing, Liu Rongyuan, Liu Suhong, et al. Monitoring changes of winter wheat coverage based on low-altitude UAV remote sensing[J]. Transactions of the CSAE, 2012, 28(13): 160 – 165. (in Chinese)

[14] 王佳诗,崔晨曦,杜奥博,等. 基于 GA – DETR 的自然场景谷类作物穗部检测方法[J]. 农业机械学报,2026,57(1): 125 – 139.
Wang Jiashi, Cui Chenxi, Du Aobo, et al. Method for spike detection of cereal crops in natural scenes based on GA – DETR [J]. Transactions of the Chinese Society for Agricultural Machinery, 2026,57(1):125 – 139. (in Chinese)

[15] 汪沛, 罗锡文, 周志艳, 等. 基于微小型无人机的遥感信息获取关键技术综述[J]. 农业工程学报, 2014, 30(18): 1 – 12.
Wang Pei, Luo Xiwen, Zhou Zhiyan, et al. Key technology for remote sensing information acquisition based on micro UAV [J]. Transactions of the CSAE, 2014, 30(18): 1 – 12. (in Chinese)

[16] 田振坤, 傅莺莺, 刘素红, 等. 基于无人机低空遥感的农作物快速分类方法[J]. 农业工程学报, 2013, 29(7): 109 – 116.
Tian Zhenkun, Fu Yingying, Liu Suhong, et al. Rapid crops classification based on UAV low-altitude remote sensing[J]. Transactions of the CSAE, 2013, 29(7): 109 – 116. (in Chinese)

[17] 万亮, 岑海燕, 朱姜蓬, 等. 基于纹理特征与植被指数融合的水稻含水量无人机遥感监测[J]. 智慧农业, 2020, 2(1): 58 – 67.
Wan Liang, Cen Haiyan, Zhu Jiangpeng, et al. Using fusion of texture features and vegetation indices from water concentration in rice crop to UAV remote sensing monitor[J]. Smart Agriculture, 2020, 2(1): 58 – 67. (in Chinese)

[18] 翟肇裕,张梓涵,徐焕良,等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报,2024,55(11):1 – 20.
Zhai Zhaoyu, Zhang Zihan, Xu Huanliang, et al. Review of applying YOLO family algorithms to analyze animal and plant phenotype[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024,55(11):1 – 20. (in Chinese)

[19] 王猛, 隋学艳, 梁守真, 等. 利用无人机遥感技术提取农作物植被覆盖度方法研究[J]. 作物杂志, 2020, 36(3): 177 – 183.
Wang Meng, Sui Xueyan, Liang Shouzhen, et al. Research on the method of extracting crop vegetation coverage using UAV remote sensing technology[J]. Crops, 2020,36(3): 177 – 183. (in Chinese)

[20] Ranjan V, Sharma U, Nguyen T, et al. Learning to count everything[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 3394 – 3403.

[21] Cui W, Yan C, Yan Z, et al. BMNet: a new region-based metric learning method for early Alzheimer's disease identification with FDG – PET images[J]. Frontiers in Neuroscience, 2022, 16: 831533.

[22] Ma Zhen, Hong Xiaowei, Shangguan Qiaoying. Can SAM count anything? An empirical study on SAM counting[EB/OL]. (2023 – 04 – 19)[2026 – 05 – 14]. https://arxiv.org/abs/2304.10817.

[23] 林志洁, 罗壮, 赵磊, 等. 特征金字塔多尺度全卷积目标检测算法[J]. 浙江大学学报(工学版), 2019, 53(3): 533 – 540.
Lin Zhijie, Luo Zhuang, Zhao Lei, et al. Multi-scale fully convolutional target detection algorithm with feature pyramid[J]. Journal of Zhejiang University (Engineering Science), 2019, 53(3): 533 – 540. (in Chinese)

[24] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64 – 72.
Zhang Chenjia, Zhu Lei, Yu Lu. Review of attention mechanism in convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(20): 64 – 72. (in Chinese)

[25] Yang Y, Peng Y H, Liu Z G. A fast algorithm for YCbCr to RGB conversion[J]. IEEE Transactions on Consumer Electronics, 2007, 53(4): 1490 – 1493.

[26] Kristan M, Leonardis A, Skočaj D. Multivariate online kernel density estimation with Gaussian kernels[J]. Pattern recognition, 2011, 44(10 – 11): 2630 – 2642.

[27] Woo S, Park J, Lee J Y, et al. Cbam: convolutional block attention module[C]// Proceedings of the European conference on computer vision (ECCV), 2018: 3 – 19.

[28] 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.

[29] He J, Liu B, Cao F, et al. Few-shot object counting with dynamic similarity-aware in latent space[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1 – 14.