doi:10.6041/j.issn.1000-1298.2025.12.047

基于 Gland - MSConNet 模型的棉花叶片色素腺体分割与表型量化分析

邵利敏[1,2]　耿玉红[1]　徐雅轩[1]　王国宁[3]　闫　庚[1]　张　怡[1]

(1.河北农业大学机电工程学院,保定 071001;2.河北省智慧农业装备技术创新中心,保定 071001;
3.河北农业大学农学院,保定 071001)

摘要:棉酚是棉属植物特有的有毒物质,在农业和制药等领域具有重要的研究价值。色素腺体是棉酚的主要载体,其表型特征与棉酚含量密切相关,能够为棉酚含量评估提供参考。目前,色素腺体检测主要依赖人工分析,耗时且效率低。为此,本文提出了一种可在复杂背景下对棉花叶片色素腺体进行精准语义分割的模型 Gland - MSConNet。该模型引入卷积双向 Mamba 模块和多尺度注意力模块,增强对多尺度特征的捕捉能力和复杂场景的适应性,并通过亚像素卷积实现高效上采样,避免局部细节丢失。结果表明,Gland - MSConNet 模型分割平均交并比为 89.83%,召回率为 95.03%,与 U - Net 模型相比,分别提高 4.66、5.78 个百分点,具有较高的分割精度和鲁棒性。Gland - MSConNet 模型为棉花叶片色素腺体分割与表型量化分析提供了有力的技术支持,为棉酚含量自动评估奠定了可靠基础。

关键词:棉花叶片;色素腺体;棉酚;语义分割

中图分类号:S24;S562　　**文献标识码**:A　　**文章编号**:1000-1298(2025)12-0510-12　　**OSID**:

Segmentation and Phenotypic Quantitative Analysis of Pigment Glands in Cotton Leaves Based on Gland - MSConNet Model

SHAO Limin[1,2]　GENG Yuhong[1]　XU Yaxuan[1]　WANG Guoning[3]　YAN Geng[1]　ZHANG Yi[1]

(1. *College of Mechanical and Electrical Engineering*, *Hebei Agricultural University*, *Baoding* 071001, *China*
2. *Technology Innovation Center of Intelligent Agricultural Equipment of Hebei Province*, *Baoding* 071001, *China*
3. *College of Agronomy*, *Hebei Agricultural University*, *Baoding* 071001, *China*)

Abstract: Gossypol, a toxic compound unique to the Gossypium species, holds significant research value in agriculture and pharmaceuticals. Pigment glands, the primary carriers of gossypol, exhibit phenotypic traits that strongly correlate with gossypol content, making them effective indicators for its evaluation. However, the detection of pigment glands primarily relies on manual analysis, which is both time-consuming and inefficient. To address these limitations, Gland - MSConNet, a novel model designed for the accurate segmentation of cotton leaf pigment glands in complex backgrounds was proposed. The model integrated the convolutional bidirectional Mamba (ConBimamba) and multi-scale attention aggregation (MSAA) modules, significantly enhancing its ability to capture multi-scale features and improving robustness in challenging scenarios. Additionally, the incorporation of sub-pixel convolution facilitated efficient upsampling, preserving fine-grained details. Experimental results demonstrated that Gland - MSConNet achieved a segmentation mIoU of 89.83% and a recall rate of 95.03%, representing improvements of 4.66 percentage points and 5.78 percentage points, respectively, over the U - Net model. These findings highlighted the model's high segmentation accuracy and robustness. The Gland - MSConNet model provided a strong technical foundation for the segmentation and phenotypic quantification of cotton leaf pigment glands, offering a reliable basis for the automated evaluation of gossypol content.

Key words: cotton leaf; pigment gland; gossypol; semantic segmentation

收稿日期:2025 - 05 - 26　修回日期:2025 - 07 - 24

基金项目:中央引导地方科技发展资金项目(246Z7403G)和华北作物改良与调控国家重点实验室自主研究课题(NCCIR2022ZZ - 15)

作者简介:邵利敏(1979—),男,教授,主要从事农业信息化研究,E-mail:shaolimin@hebau.edu.cn

0　引言

棉花是全球重要经济作物之一，提供优质天然纤维[1]、种子油和蛋白质[2]，具有产量高、成本低等优点。色素腺体在棉花叶片上呈棕色或深棕色斑点[3]，主要成分为棉酚及其衍生物[4]。棉酚广泛应用于农业、工业和医学等领域，具有抑制害虫[5-6]、抗氧化[7]、抗病毒[8]和抗肿瘤[9-10]等作用，但其毒性限制了棉籽作为食物资源的利用[11-12]。培育不含色素腺体的棉花品种，能够提高棉籽利用率，促进粮食安全和资源可持续利用[13]。因此，评价棉花植株中棉酚含量能够为棉花抗病虫育种提供依据，对提高棉花经济价值至关重要。

目前，棉酚测定常用方法包括高效液相色谱法[14]、毛细管电泳[15]和近红外光谱[16]等，这些方法损坏样品，且成本较高。色素腺体作为棉酚储存器官，其密度与棉酚含量显著正相关[17]，因此，色素腺体表型特征可作为棉酚含量评估的可靠依据[18]。

色素腺体数量是评估有酚棉中棉酚含量的重要指标。色素腺体直径为 100～400 μm，个体小且密集，人工计数难度大[19]。早期研究通过显微镜观察少量色素腺体或棉花叶片局部区域，效率低、误差大，且难以完全覆盖整个叶片[20]。OPITZ 等[21]结合相机与显微镜推测色素腺体密度，但样本量少、统计代表性差，尤其在腺体密度变化大的边缘区域表现不佳。传统方法难以满足大规模高效检测需求。

基于计算机视觉和深度学习的自动表型分析技术为作物育种提供了重要支持[22]。与传统方法相比，这类方法克服了费力、耗时等缺点，尤其是深度学习语义分割技术，已成为作物育种的重要工具。语义分割将图像像素划分为不同语义类别，有效处理大规模、高维数据，提高测量精度和效率[23]。主流的语义分割模型包括 U－Net[24]、DeepLabv3＋[25]、PSPNet[26]和 TransUnet[27]，广泛应用于植物表型的精准分割与定量分析[28]。凭借其快速、可靠、高通量的优势，语义分割模型逐步替代传统手动测量，大幅提高表型分析的效率和精度[29]。

尽管分割技术不断进步，但仍面临复杂背景下的误检或漏检问题，特别是小目标形态变化大，现有技术难以精准分割。如何平衡全局上下文信息与局部细节，仍是模型的瓶颈。近年来，针对这些挑战，通过改进深度学习和语义分割技术，结合不同架构和优化策略，有效提升了在遮挡、光照变化等复杂环境中的分割精度，推动了农业图像处理技术的发展[30-32]，为棉花叶片色素腺体的检测与表型分析提供了重要支持。

针对复杂背景下色素腺体精准分割和表型量化分析问题，本文提出 Gland－MSConNet 语义分割模型。该模型基于编码器-解码器架构进行改进，采用先插值后池化结构，避免下采样导致的特征信息丢失，提高小目标检测精度；引入卷积双向 Mamba 模块 ConBimamba 与多尺度注意力聚合（Multi－scale attention aggregation，MSAA）模块，提升特征融合效果和复杂背景下的检测能力；同时，使用亚像素卷积（Sub-pixel convolution）替代双线性插值上采样，避免细节丢失，以提升模型效率与精度。

1　材料与方法

1.1　图像采集与预处理

棉花叶片色素腺体图像分割与表型分析流程如图 1 所示。本文图像采集于河北农业大学创新实验园（38.85°N，115.30°E），参照 WANG 等[33]的种植方法，在温度 22～24℃（夜间）、26～28℃（日间）温室条件下种植棉花，采用 50% 的霍格兰营养液进行养护，4 d 更换一次[34]。以 16 种陆地棉为研究对象，待棉花植株生长至 7～8 片真叶期，取倒二叶，利用爱普生 V39 型扫描仪采集棉花叶片原始图像，分辨率为 9 528 像素×13 475 像素。

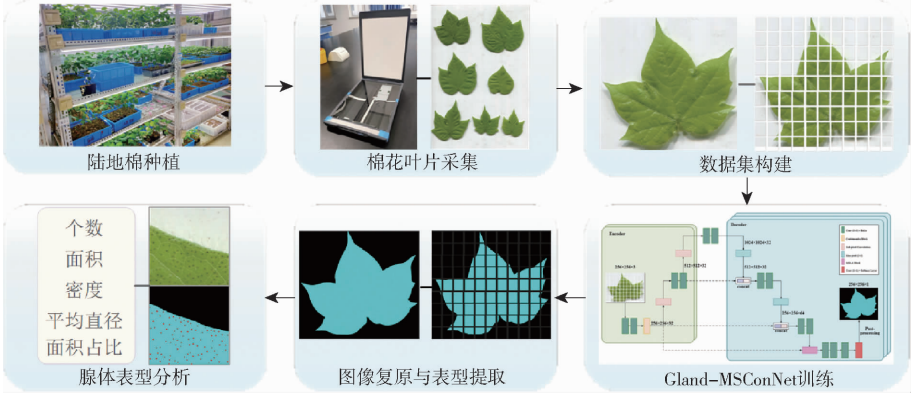


图 1　棉花叶片色素腺体图像分割与表型分析流程

Fig. 1　Workflow of image segmentation and phenotypic analysis of cotton leaf pigment glands

采用 LabelMe 工具对图像中背景、叶片和色素腺体 3 类区域进行标注，将标签分别命名为 background、leaf 和 gland。标注分 2 次完成：第 1 次标注背景和色素腺体区域；第 2 次标注背景和叶片区域。每次标注完成后，将标注数据以 JSON 文件格式保存，然后利用图像合并算法将 JSON 文件转换为二值图像，再将 2 幅二值图像合并为 1 幅。生成的图像中背景、色素腺体和棉花叶片分别对应像素值为 0、1 和 2，以 PNG 格式保存，便于后续模型训练与测试使用。

为了评价分割结果精度，基于扫描仪的固定扫描范围（210 mm × 297 mm）和图像像素分辨率，通过计算得出一个像素面积为 $4.857\,86 \times 10^{-4}$ mm^2。以此为依据，将图像中的像素面积转换为实际的棉花色素腺体面积，从而进一步评估色素腺体数量、面积、密度和平均直径，以及色素腺体面积与棉花叶片面积比值等表型参数。

1.2 数据集构建

由于原始图像视野较广，对显存消耗较大，直接处理存在困难，因此需要对棉花叶片的原始图像进行裁剪，得到 2 132 幅尺寸为 256 像素 × 256 像素的子图像。PNG 格式的标注图像也采用相同方式进行处理。裁剪后的数据按比例 8∶2 划分为训练集和测试集。训练集包含 1 206 幅裁剪后的子图像和 2 500 幅通过数据增强操作生成的图像，数据增强方式包括水平翻转、垂直翻转和镜像倒转，以丰富数据集并提升模型的泛化能力。测试集包含 926 幅裁剪后的子图像，为模型的训练和性能评估提供了充分且高质量的数据支撑。

1.3 Gland - MSConNet 语义分割模型

色素腺体与棉花叶片的面积占比极不平衡，且色素腺体体积小、分布密集，再加上叶片上叶脉阴影与色素腺体形态相似，导致分割结果出现偏差，获取色素腺体数量、面积、密度、平均直径、面积占比等关键表型参数变得异常困难。

本文提出 Gland - MSConNet 模型（如图 2 所示），用于棉花叶片色素腺体分割与表型分析，以有效应对上述挑战。该模型优化了经典的编码器-解码器架构，通过引入 ConBimamba、MSAA 和亚像素卷积模块显著提升模型的识别能力与泛化能力。优化之后的编码器和解码器结构，采用先插值后池化的方法，有效避免卷积导致的目标特征信息丢失。编码器部分通过标准卷积层提取特征，在浅层加入 ConBimamba 模块。解码器部分采用亚像素卷积实现高效上采样，降低计算开销。模型融合了多层次特征，采用跳跃连接结合不同深度的特征层，并借助 MSAA 模块实现多尺度特征的自适应聚合。

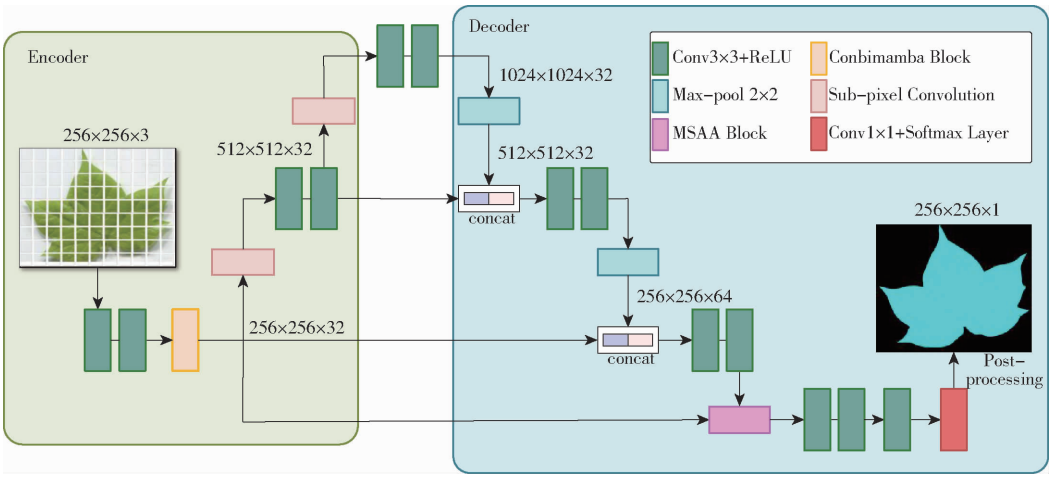


图 2 Gland - MSConNet 模型架构图

Fig. 2 Gland - MSConNet model architecture diagram

1.3.1 卷积双向 Mamba 模块

卷积双向 Mamba 模块 ConBimamba[35] 充分借鉴了 Mamba 在语言建模中的成功经验[36]，将其强大的序列建模能力巧妙地迁移到视觉领域。在编码器部分引入 ConBimamba 模块是 Gland - MSConNet 模型的关键改进点，通过结合多头注意力机制和扩展卷积结构，显著提升了特征表达的精细度和复杂场景中的表现能力。

引入 ConBimamba 模块，通过正向和反向动态交互的双向序列建模，有效捕获复杂纹理和特征的顺序关系，非常适用于腺体与背景交织的分割任务。ConBimamba 模块如图 3 所示，包含 2 个前馈模块（Feed - Forward），中间夹有 Mamba 和 Convolution。这种三明治的灵感来自 Macaron - Net[37]，将 Transformer 中的 Feed - Forward 替换为 2 个半步 Feed - Forward，也就是 Feed - Forward 的输出乘以 1/2，

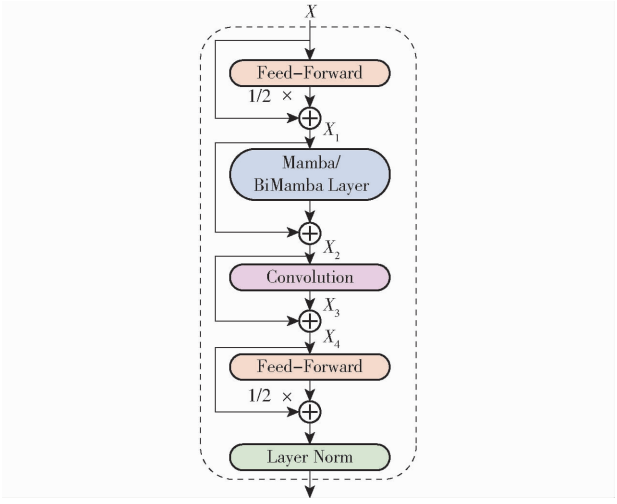


图 3　ConBimamba 模块

Fig. 3　ConBimamba module

然后再做残差连接。一个半步 Feed－Forward 在注意力层前面，另一个在注意力层后面，执行步骤如下：

（1）给定输入 X，通过第 1 个半步 Feed－Forward，输出 X_1。

（2）执行双向 Mamba。将 X_1 输入到正向 Mamba 模块中，输出 U_1；将 $\overline{X}_1$（X_1 的反向表示）输入到反向 Mamba 模块中，输出 U_2；再将 U_1、$\overline{U}_2$（U_2 的反向表示）和 X_1 在通道上进行拼接，并通过线性层恢复为和输入相同的通道数，然后添加残差连接，输出结果 X_2。

（3）执行 Conv1 × 1 卷积操作，输出 X_3；再通过 DWConv 以及门控线性单元等操作，输出 X_4。

（4）执行第 2 个半步 Feed－Forward，再进行归一化操作（Layer Norm），输出结果 Y，作为 ConBimamba 模块的最终输出。

在注意力机制的前后分别加入前馈“半步”结构，进一步增强了全局信息与局部特征之间的交互能力。同时，卷积操作弥补了传统序列建模对局部细节表达的不足，结合门控线性单元动态调控特征表达，使色素腺体区域和背景区域的特征分离更加精准。模块中的残差连接与归一化设计进一步增强了网络的稳定性和泛化能力，能够更好地应对色素腺体分布密集的问题。

ConBimamba 模块不仅优化了模型特征提取过程，而且显著改善了色素腺体在复杂场景中的分割效果。尤其是在色素腺体区域被叶脉阴影干扰或背景复杂变化时，ConBimamba 模块能够对色素腺体表型的关键细节进行精细化捕捉，保证了分割结果的稳定性和可靠性。

1.3.2　亚像素卷积

为实现高效上采样，Gland－MSConNet 模型在解码器部分采用亚像素卷积，其为一种高效的超分辨率（Super-resolution）方法[38]，结构简单，高效且易于实现。与传统卷积不同，亚像素卷积通过通道重排的方式提升图像分辨率。使用卷积操作增加通道维度，然后通过 PixelShuffle 操作将新增的通道数据重新组织到空间维度，从而实现图像放大，具体操作如图 4 所示，图像经过一系列卷积后，得到尺寸为 $H \times W \times C^2$ 的特征图，其中，H 表示图像高度，W 表示图像宽度，C 表示图像通道数。Sub-pixel 将特征图上每个像素的 C^2 个通道重新排列成一个 $C \times C$ 的区域，对应高分辨率图像中的一个 $C \times C$ 大小的子块，使得尺寸为 $H \times W \times C^2$ 的特征图被重新排列成 $CH \times CW \times 1$ 的高分辨率图像。作为一种创新的基于卷积神经网络（Convolutional neural network，CNN）的技术，亚像素卷积通过“像素重排”有效避免了传统方法中因下采样导致的信息丢失，不仅提升了图像质量，而且降低了计算复杂度。

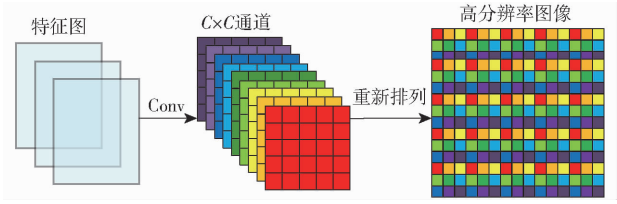


图 4　Sub-pixel Convolution 模块

Fig. 4　Sub-pixel Convolution module

1.3.3　多尺度注意力模块

MSAA 模块[39]专为聚合空间与通道维度信息设计，通过多层次的特征优化增强分割效果。本文在 Gland－MSConNet 模型中引入 MSAA 模块，以进一步提升模型对多尺度特征的捕捉能力和复杂场景的适应性。

MSAA 模块为 2 分支结构，由空间分支和通道分支组成，如图 5 所示，其中 N 为输入该模块特征图，K 为模块输出特征图。空间分支通过多尺度卷积和空间注意力机制生成权重，显著提升模型对复杂背景区域的感知能力；通道分支则通过捕获通道间的相关性，强化模型对重要特征的识别能力。空间分支和通道分支的特征融合通过逐元素交互完成，并通过元素乘法进行融合，保持特征信息的完整性和传递性。

给定输入特征图 $N \in R^{C \times H \times W}$，首先通过 Conv 1 × 1 将通道数降维得到

$$Z \in R^{D \times H \times W} = \mathrm{Conv}(N) \tag{1}$$

式中　D——经 Conv 1 × 1 降维后特征图的通道数，且 D 小于拼接前单个特征的通道数 C

将 Z 通过多个具有不同卷积核大小的卷积层，以相加的方式进行多尺度融合，计算式为

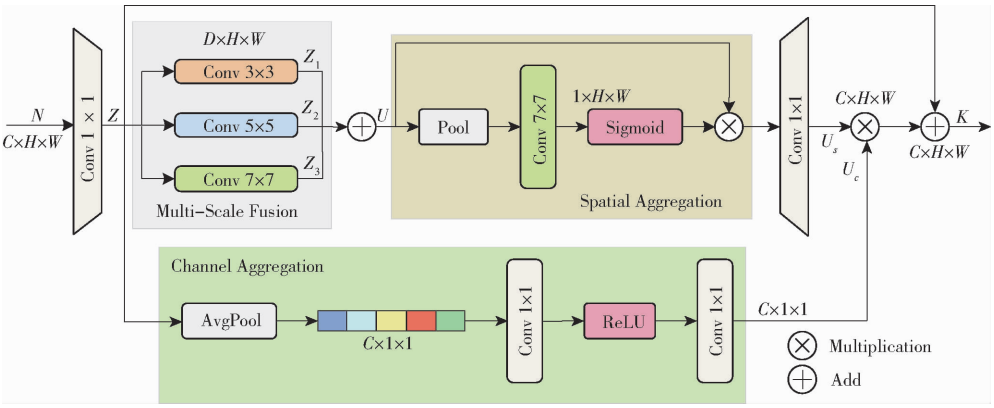


图 5 MSAA 模块

Fig. 5 MSAA module

$$U \in R^{D\times H\times W} = Z_1 + Z_2 + Z_3 \tag{2}$$

多尺度融合的输出被馈入到空间注意力中，首先在通道方向上执行平均池化和最大池化，然后拼接并降维，通过 Sigmoid 函数生成空间权重，即

$$W_s \in R^{1\times H\times W} = \mathrm{Sigmoid}(\mathrm{Conv}(\mathrm{Pool}(U))) \tag{3}$$

式中 W_s——用于调整输入特征，并通过 Conv 层进行线性变换，输出U_s

通过通道注意力调整输入特征的通道重要性。从降维特征 $Z \in R^{D\times H\times W}$通过全局平均池化得到通道特征表示，然后通过 2 个 Conv 1×1 学习通道间的相关性，并通过 Sigmoid 函数生成通道权重，即

$$W_c = R^{D\times 1\times 1} = \mathrm{Sigmoid}(\mathrm{Conv}(\mathrm{Conv}(\mathrm{Pool}(Z)))) \tag{4}$$

通过生成的通道权重W_c对输入特征进行调整，输出U_c。空间注意力的输入和通道注意力的输入通过逐元素点积进行交互，并通过一个 Conv 层恢复至和输入特征相同的通道，最后添加一个残差连接得到最终的输出

$$K \in R^{C\times H\times W} = \mathrm{Conv}(U_s \odot U_c) + Z \tag{5}$$

MSAA 模块被嵌入在解码器部分，配合跳跃连接和池化操作，有效合并不同层次特征。MSAA 模块的引入优化了模型对小目标色素腺体的检测性能，有效解决了棉花叶片叶脉阴影干扰以及色素腺体分布密集等问题。通过整合多尺度信息，MSAA 模块显著提升了色素腺体表型特征的检测精度，为实现色素腺体特征的细粒度量化分析提供了技术保障。

1.3.4 损失函数

考虑到棉花叶片中色素腺体与背景之间存在显著的类别不平衡问题，腺体目标通常具有体积小、数量密集、边缘模糊等特征，且在整体图像中所占面积远小于背景区域，这使得交叉熵（Cross-entropy，CE）等传统损失函数在优化小目标区域识别方面存在明显局限。为有效应对上述挑战，本文引入 Dice Loss 与 Focal Loss 的组合策略，以同时兼顾小目标分割精度与模型对复杂背景的鲁棒性。

Dice Loss 目标是最小化预测分割结果与真实分割结果之间的 Dice 系数差异，对目标体积较小且类别高度不平衡的数据表现出良好的适应能力，适用于本文中密集分布的色素腺体检测任务。Dice 系数越高，表示预测结果与真实标签之间的重叠程度越好，计算式为

$$L_{\mathrm{Dice}} = 1 - \frac{2|A\cap B|}{|A| + |B|} \tag{6}$$

式中 A、B——模型预测区域和真实区域

$|A\cap B|$——A 和 B 交集

$|A|$、$|B|$——A、B 面积

在交叉熵损失的基础上，Focal Loss 引入调制因子和类别平衡参数，有效降低易分类样本对总损失的影响，从而增强模型对难分类样本的关注能力，计算式为

$$L_{\mathrm{Focal}} = -\alpha(1-p_t)^{\gamma}\lg p_t \tag{7}$$

式中 p_t——正确类别概率

α——类别平衡因子

γ——控制难易样本关注程度

为进一步融合二者优势，将 Dice Loss 和 Focal Loss 进行加权求和构建组合损失函数

$$L_{\mathrm{Total}} = \omega_1 L_{\mathrm{Dice}} + \omega_2 L_{\mathrm{Focal}} \tag{8}$$

式中 ω_1、ω_2——可调权重参数，用于平衡两项损失对总训练目标的影响

Dice Loss 和 Focal Loss 的集成提升了模型在密集目标检测中的精度，增强了模型训练稳定性和效率。

2 实验与结果分析

2.1 实验环境

硬件配置使用 NVIDIA GeForce RTX 4060 Laptop GPU 和英特尔 13th Gen Intel(R) Core(TM)

i9－13900H CPU，操作系统为 Windows 11；软件配置方面，采用 CUDA 11.8 和 CUDNN 8.2.1 进行深度学习计算加速，深度学习框架选择 TensorFlow。为确保试验结果的一致性，对网络模型参数进行了标准化处理。优化器选用 Adam，初始学习率 0.000 1，批量大小为 4，总训练轮次（epoch）为 100。同时，添加权重衰减系数为 0.000 5 以防止过拟合，增强模型泛化能力。

2.2　评价指标

为评价网络性能，使用平均交并比（Mean intersection over union，mIoU）、准确率（Accuracy）、精确率（Precision）、召回率（Recall）和 F1 分数（F1 score）等作为模型评价指标。

2.3　消融实验

为深入探讨模型各组成部分对整体性能的影响，设计了一系列消融实验。针对模型训练过程中关键的损失函数，分析其对分割效果的作用；在确定最优损失函数的基础上，进一步评估模型中关键创新模块的有效性，以验证各模块在提升色素腺体分割性能中的具体贡献。通过系统的实验分析，力求全面揭示各因素对模型表现的影响机制。

为评估不同损失函数对模型性能的影响，对比 4 种损失函数配置下的分割效果，结果如表 1 所示。与传统交叉熵损失相比，Dice Loss 更加关注预测结果与真实标签的区域重叠，能够有效提升对体积小、分布密集的色素腺体目标的分割效果，尤其 mIoU 表现突出。Focal Loss 通过抑制易分类样本的影响、强化对困难样本的关注，有助于提升模型在应对叶脉阴影和边缘模糊等复杂背景时的鲁棒性。二者结合使用后，模型在前景识别与背景抑制之间实现更优平衡，显著增强了类别不平衡条件下的训练稳定性和整体分割性能。相比单一损失函数，该组合方式进一步缓解了过拟合风险，提高了模型在复杂场景下的稳定性和泛化能力。结果表明，Dice + Focal 的组合策略在 mIoU 和 F1 分数两项指标均达最佳效果，验证了其在棉花叶片色素腺体分割任务中的有效性与推广潜力。

表 1　不同损失函数效果对比

Tab.1　Comparison of effects of different loss functions

%

损失函数	平均交并比	F1 分数
CE Loss	86.42	91.25
Dice Loss	88.17	92.84
Focal Loss	87.03	92.07
Dice + Focal Loss	89.83	94.18

为进一步验证改进网络中不同改进方案的有效性，进行了消融实验。以 Gland－MSConNet 模型中去除 ConBimamba 模块、MSAA 模块以及亚像素卷积模块后得到的模型作为基础模型，通过逐一添加模块测试方案性能，分析各模块对整体性能提升的作用，共设计了 8 组实验，结果如表 2 所示。所有实验在相同的数据集上进行，学习率、训练轮次等模型超参数保持一致。

表 2　消融实验结果

Tab.2　Ablation experiment results

%

ConBimamba	亚像素卷积	MSAA	平均交并比	准确率	召回率
×	×	×	82.43	99.20	85.01
√	×	×	86.90	99.43	90.75
×	√	×	85.51	99.29	89.88
×	×	√	86.30	99.35	90.15
√	√	×	87.95	99.52	92.23
×	√	√	87.72	99.50	92.10
√	×	√	88.80	99.59	94.12
√	√	√	89.83	99.66	95.03

注：√表示采用此模块，×表示未采用此模块。

由表 2 可知，未进行任何改进的模型的 mIoU、准确率和召回率分别为 82.43%、99.20% 和 85.01%，可作为后续对比的参考。在单模块实验中，引入 ConBimamba 模块后，mIoU 提升 4.47 个百分点，召回率提升 5.74 个百分点，对局部细节增强作用显著；引入亚像素卷积后，mIoU 提升 3.08 个百分点，表明该模块对高分辨率特征恢复具有一定贡献，但对复杂背景的适应性较弱；引入 MSAA 模块后，mIoU 提升 3.87 个百分点，对背景噪声的抑制效果较好。在双模块组合实验中，ConBimamba + 亚像素卷积组合的 mIoU 提升 5.52 个百分点，体现出对局部特征表达和边界细化的协同作用；ConBimamba + MSAA 组合表现最佳，mIoU 和召回率分别提升 6.37、9.11 个百分点，展现出 MSAA 模块对复杂背景的适应性和 ConBimamba 模块对细节的增强作用；亚像素卷积 + MSAA 组合的 mIoU 和召回率分别提高 5.29、7.09 个百分点，在高分辨率恢复和背景噪声处理上表现均衡，但略逊于 ConBimamba + MSAA 组合和 ConBimamba + 亚像素卷积组合。Gland－MSConNet 模型加入 3 个模块后，mIoU、准确率和召回率达到 89.83%、99.66% 和 95.03%，分别提高 7.4、0.46、10.02 个百分点。通过结合 ConBimamba 模块的局部细节增强、亚像素卷积的高分辨率细节复原和 MSAA 模块对复杂背景的适应性，提升了模型性能。

然而，消融实验也揭示了一些局限性，3 个模块全面组合后性能提升较小。相较 ConBimamba +

MSAA 组合,加入亚像素卷积后 mIoU 仅提升 1.03 个百分点,召回率提升 0.91 个百分点,提升幅度不太明显。这表明,ConBimamba 和 MSAA 组合已较好地增强了局部细节与背景抑制,加入亚像素卷积使部分功能重叠,导致其贡献相对有限。此外,尽管 Recall 指标整体较高,但不同实验组间差异较小,模型对更复杂场景的适应性仍需加强;同时,3 个模块的叠加增加了计算复杂度,推理速度有所下降,在一定程度上会影响在资源受限的场景中的应用。消融实验验证了各模块的独立有效性及协同作用,同时明确了进一步优化的方向。

2.4 对比实验

为更全面对比 Gland-MSConNet 模型的整体性能,验证其在棉花叶片色素腺体分割任务上的有效性,本文与其他模型进行了对比,主要包括 U-Net、PSPNet(Pyramid scene parsing network)、DeepLabv3+ 和 TransUnet。

2.4.1 定量评估

对比结果如表 3 所示。由表 3 可以看出,Gland-MSConNet 在各项评估指标上均表现较优,尤其是 mIoU 和 F1 分数,分别为 89.83% 和 94.18%,显著高于其他模型。Gland-MSConNet 的 mIoU 比 PSPNet、DeepLabv3+、TransUnet 和 U-Net 分别高 25、17.33、7.47、4.66个百分点,体现出较高的分割准确性。F1 分数也明显领先,较 U-Net 高 3.36 个百分点,并且精确率和召回率分别高 0.84、5.78 个百分点,同样表现优越,表明该模型在分割精度和召回率方面具有良好平衡,验证了 Gland-MSConNet 模型在棉花叶片色素腺体分割任务中的优越性。

表 3 对比实验结果

Tab.3 Results of Comparative experiment %

模型	平均交并比	精确率	召回率	F1 分数
PSPNet	64.83	65.53	65.94	65.73
DeepLabv3+	72.50	80.33	77.09	78.68
TransUnet	82.36	88.00	84.55	86.24
U-Net	85.17	92.48	89.25	90.82
Gland-MSConNet	89.83	93.32	95.03	94.18

2.4.2 定性评估

图 6 为不同模型下色素腺体分割效果对比。由图 6 可知,Gland-MSConNet 在细粒度表型特征的提取上表现最为出色,能够精准提取色素腺体的关键表型特征,实现精细化分割。其他模型(如 U-Net、PSPNet、DeepLabv3+ 和 TransUnet)在棉花叶片色素腺体的分割任务中普遍存在性能下降的问题,分割效果明显低于 Gland-MSConNet。DeepLabv3+ 和 PSPNet 在特征提取过程中会造成色素腺体的特征丢失,导致分割后的图像存在色素腺体被误分类为背景的问题,分割效果很不理想。虽然 U-Net 和 TransUnet 在一些情况下能够较为精准地分割色素腺体,但在处理含有叶脉和叶片边缘的图像时,叶脉阴影容易被误识别为色素腺体,叶片边缘的分割也较为模糊,导致分割结果与真实值差距较大。Gland-MSConNet 在面对复杂背景和形态变化的色素腺体时,展现了更强的适应性,能够有效减少叶脉阴影和叶片边缘的影响,确保分割结果的稳定性,特别是在密集分布和形态变异的情境下,表现出比其他模型更为精确的分割能力。

通过定量与定性综合评估,本文模型性能优于其他主流语义分割模型。Gland-MSConNet 凭借其优化的特征提取过程和细粒度表型特征捕捉能力,显著提升了棉花叶片色素腺体分割效果,并在复杂场景下表现出较强的稳定性和可靠性。

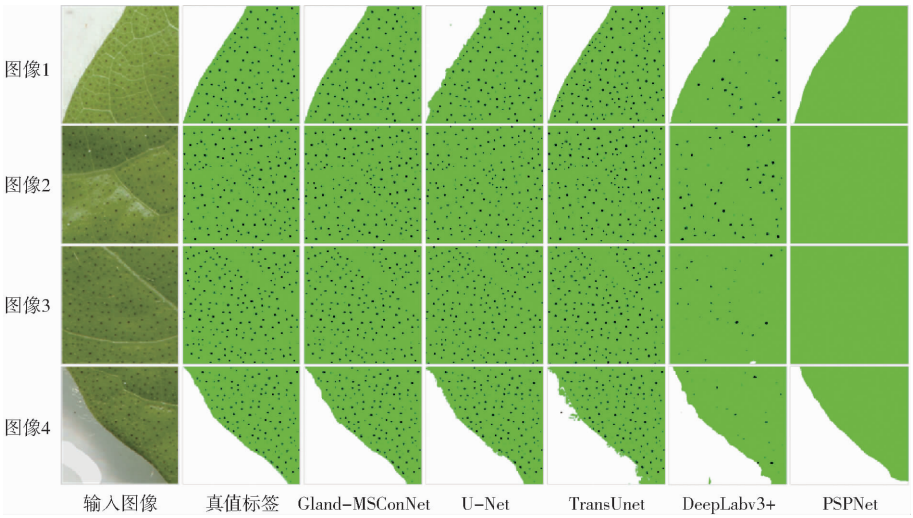


图 6 不同模型下色素腺体分割效果对比

Fig.6 Comparison of pigment gland segmentation across different models

2.5　模型复杂度与可部署性分析

为全面评估本文模型在实际应用中的部署可行性与计算效率，采用参数量、浮点运算量、推理时间、帧率、模型内存占用量及 ONNX 文件大小作为评价指标，与主流语义分割模型（PSPNet、DeepLabv3＋、TransUnet 和 U－Net）在多个关键性能指标对比结果见表 4。

由表 4 可知，Gland－MSConNet 在保持较高分割精度的同时，参数量为 5.860×10^7，浮点运算量为 1.929×10^9，推理时间为 31.75 ms，帧率达到 31.50 f/s，模型内存占用量约为 320.24 MB。整体资源消耗处于传统轻量级模型（如 PSPNet）与复杂模型 Transformer（如 TransUnet）之间，表现出良好的性能与计算效率平衡。与 PSPNet 和 DeepLabv3＋相比，Gland－MSConNet 的 mIoU 分别提高 25、17.33 个百分点，而在模型复杂度上仅略高于二者；相比 U－Net，mIoU 提升 4.66 个百分点，浮点运算量增幅仅为 4.44%，同时具备更优的推理速度与更高的帧率。由此可见，该模型在显著提升小目标分割精度的同时，并未显著增加部署成本。

表 4　不同网络模型性能测试结果

Tab.4　Performance test results of different network models

模型	参数量	浮点运算量	推理时间/ms	帧率/(f·s^{-1})	模型内存占用量/MB	ONNX 文件大小/MB	平均交并比/%
PSPNet	4.672×10^7	1.185×10^9	16.61	60.21	270.19	178.08	64.83
DeepLabv3+	5.471×10^7	1.670×10^9	27.13	36.86	291.10	208.36	72.50
TransUnet	9.702×10^7	2.401×10^9	40.00	25.00	478.17	216.83	82.36
U－Net	4.393×10^7	1.847×10^9	20.49	48.80	260.72	167.50	85.17
Gland－MSConNet	5.860×10^7	1.929×10^9	31.75	31.50	320.24	183.79	89.83

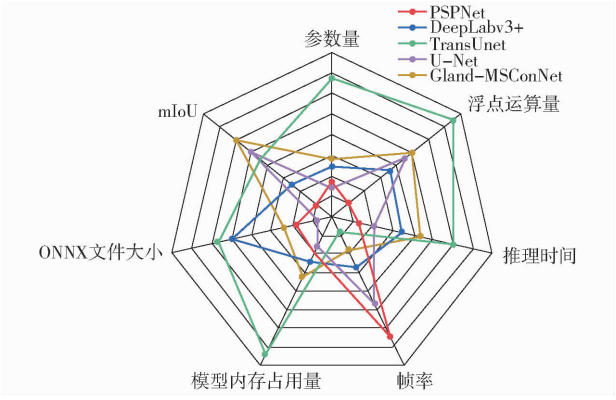


图 7　不同网络模型的综合性能对比

Fig.7　Comprehensive performance comparison of different network models

此外，Gland－MSConNet 模型设计中所引入的 ConBimamba 与 MSAA 模块仅在关键位置嵌入，避免了盲目堆叠注意力模块导致的资源浪费；解码器采用亚像素卷积替代双线性插值，不仅减少了运算量，还提升了细节还原能力，加上“先插值后池化”的结构优化策略，使得整个模型在不牺牲精度的前提下，实现了高效推理与良好的部署灵活性。为直观展示各模型的性能与资源消耗差异，图 7 为综合性能雷达图，进一步突出了 Gland－MSConNet 在精度与部署效率方面的均衡优势。

2.6　色素腺体表型量化分析结果

图 8 为 Gland－MSConNet 模型对棉花叶片色素腺体的表型量化分析结果，该叶片共有 6 278 个色素腺体，总面积为 113 847 像素，平均直径约 3.786 6 像素。

为了将像素级测量结果转换为实际物理单位，采用了像素面积换算系数表示每个像素对应的实际面积，pixel_area_mm^2 为 $4.857\,86\times10^{-4}$ mm^2。基于此换算关系，色素腺体总面积计算式为

$$S=113\,847\times4.857\,86\times10^{-4}\times10^{-2}=0.553\,1\ \text{cm}^2 \tag{9}$$

同理，叶片总面积为 16.133 6 cm^2。色素腺体面积与叶片总面积之比为 3.428%，色素腺体密度达 389.126 9 个/cm^2。Gland－MSConNet 模型能够精准分割棉花叶片上的色素腺体。

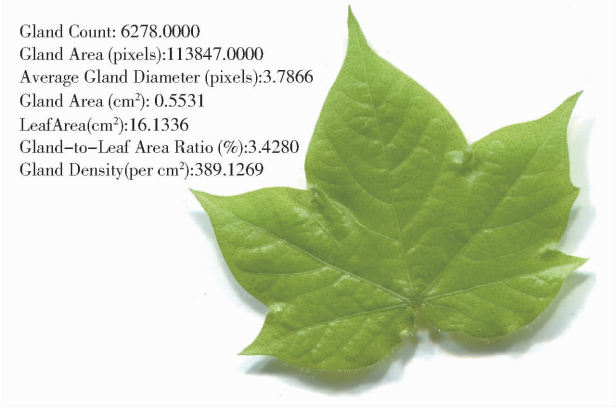


图 8　色素腺体表型量化分析结果

Fig.8　Results of quantitative analysis of pigment gland phenotype

3　讨论

3.1　模型改进

棉酚在农业和制药等领域具有重要研究价值，

色素腺体是棉酚的主要载体,其表型性状与棉酚含量密切相关,可用于棉酚含量评估。目前,色素腺体的检测主要依赖人工分析,工作量大且耗时。SHE 等[40]提出了一种棉花叶片色素腺体自动检测与计数的语义分割模型 IPP Net,可以实现对棉花叶片中色素腺体的精准分割,分割准确率为 96.7%,mIoU 为 81.81%,F1 分数为 80.04%,可以代替人工进行棉花色素腺体计数。XU 等[34]提出了一种棉花叶片色素腺体面积检测语义分割模型 GlandSegNet,可以实现对棉花叶片色素腺体的分割和面积测量,mIoU 达到 89.63%,F1 分数为 82.47%。这些研究仅完成了色素腺体数量和面积的检测,对于棉酚含量的评估工作,还需进一步提取色素腺体其他表型信息,如色素腺体密度、平均直径以及色素腺体面积与棉花叶片面积比值等关键表型信息。相比之下,本文所提用于快速检测棉花叶片中色素腺体表型的语义分割模型(Gland - MSConNet),可实现色素腺体在复杂背景下的精准分割和腺体表型特征的全面量化分析,准确率达到 99.66%,mIoU 为 89.83%,召回率为 95.03%,F1 分数为 94.18%。

Gland - MSConNet 模型针对密集分布的小目标和复杂背景设计了更具表现力的语义分割结构,通过引入 ConBimamba、MSAA 和亚像素卷积模块显著增强了模型对目标的识别能力与泛化能力。编码器部分通过标准卷积层提取特征,并在浅层加入 ConBimamba 模块,该模块结合多头注意力机制和扩展卷积,提升对视觉空间的敏感性,强化局部和全局特征的表达能力,满足全局上下文理解的需求。解码器部分采用 Sub-pixel 卷积实现高效上采样,与传统方法相比,提高了空间分辨率的还原精度。此外,模型融合了多层次特征,采用跳跃连接结合不同尺度的特征层,增强了通道信息表达能力,借助 MSAA 模块实现多尺度特征的自适应聚合,增强在不同尺度和复杂背景下的检测能力,显著提高了对小目标和形态多样性的色素腺体检测能力。对比实验和消融实验结果均表明,Gland - MSConNet 模型具有更好的分割精度和鲁棒性,分割效果显著优于 U - Net、PSPNet、DeepLabv3+ 和 TransUnet 等主流语义分割模型。

从对比结果来看,Gland - MSConNet 在实现高精度分割的同时,展现出良好的计算资源控制能力和部署灵活性。整体资源消耗介于传统轻量模型与复杂 Transformer 架构之间,显著提升分割精度的同时,保持了适中的参数规模和推理成本,体现出较高的性能-效率平衡。本文模型性能的提升并不完全依赖于模型复杂度的增加,例如,尽管 TransUnet 的参数量高达 9.702×10^7,浮点运算量达 2.401×10^9,但其分割精度提升有限,说明简单堆叠模块增加了计算量,并不能带来成比例的性能收益。相比之下,Gland - MSConNet 通过合理设计网络结构,在有效控制模型复杂度的同时实现了显著的性能提升,体现出“小幅复杂度换取大幅性能”的优化策略。这种设计理念不仅提升了密集小目标分割的准确性,也增强了模型在资源受限环境下的部署可行性。

3.2 模型泛化能力评估

为进一步验证 Gland - MSConNet 模型在其他植物类型上的适应性,选取由克莱姆森大学采集的高粱穗数据集作为迁移测试样本。高粱穗籽粒排列密集且边界紧密,籽粒与穗轴在颜色和纹理上高度相似,导致目标易粘连且背景复杂,这些因素对模型边缘感知和小目标分割能力提出了较高要求。实验中,未对模型进行额外微调,直接将训练完成的 Gland - MSConNet 应用于该数据集,高粱穗原始图像与分割结果如图 9 所示。由图 9 可知,Gland - MSConNet 凭借其多尺度注意力聚合模块和卷积双向 Mamba 模块,能够精准分割高密度小目标,有效识别出每个籽粒,边界清晰且结构完整,展现出良好的泛化能力。



(a) 原始图像 (b) 分割结果

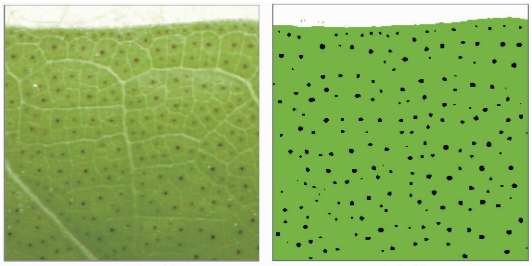
图 9 高粱穗原始图像与分割结果对比

Fig. 9 Comparison between original image and segmentation result of sorghum panicle

实验结果表明,Gland - MSConNet 不仅在棉花叶片色素腺体检测中表现优异,还具备适应复杂植物结构及密集小目标分布的能力,具备较强的跨作物泛化潜力,为后续在其他作物表型分析及无损检测任务中的应用提供了坚实基础。

3.3　讨论及展望

尽管 Gland－MSConNet 在整体分割性能上取得了较好表现，但在部分样本中仍存在色素腺体边缘不够光滑、轮廓不规则等问题，如图 10 所示，这在一定程度上影响了腺体形态特征（如直径、面积等）的准确提取。造成该现象的原因可能包括：①图像采集过程中存在轻微模糊或噪声，导致边缘细节缺失。②色素腺体目标尺寸较小，边缘像素在训练样本中占比较低，模型难以充分学习其边界特征。③当前网络结构尚未显式引入针对边缘区域的感知机制。



(a) 原始图像　　(b) 分割结果

图 10　色素腺体原始图像与分割结果对比

Fig. 10　Comparison of original and segmented pigment gland images

为进一步提升模型对色素腺体边缘区域的建模能力，未来可在数据预处理与模型训练、推理等阶段进行改进。在数据准备阶段，采用高分辨率的数据增强策略，例如对腺体边缘区域进行局部放大、锐化增强、边缘对比度提升等操作，以丰富模型对边缘细节的感知样本，提升模型对小目标边界的鲁棒性。在模型训练阶段，引入边缘感知损失函数（Edge-aware Loss），通过对边缘像素进行加权监督，引导模型更加准确地学习目标轮廓。在推理阶段，集成 CascadePSP 模块，对模型初步输出进行多尺度细化，提升分割结果边缘连续性与结构完整性。

在分割性能优化的基础上，后续还将进一步拟合色素腺体表型特征（如密度、平均直径和面积等）与棉酚含量之间的理化参数关系，建立非破坏性棉酚含量估测模型，并结合不同生育期图像构建基于时间序列的棉酚含量动态预测模型，最终开发集成分割与分析的无损检测平台，为棉花精准育种提供技术支撑。

4　结束语

探索了一种小目标且分布密集环境下进行棉花叶片色素腺体分割的方法。利用本文所提出的 Gland－MSConNet 模型，实现了在复杂背景下对色素腺体的精准分割和表型特征的量化分析，解决了色素腺体检测耗时费力等问题。为提高检测精度，在模型编码器部分引入 ConBimamba 模块，在模型解码器部分引入 MSAA 模块，上采样部分选用亚像素卷积。消融实验结果表明，改进模型的 mIoU、准确率和召回率分别达到 89.83%、99.66% 和 95.03%。此外，实验结果表明，与传统语义分割模型相比，本文模型能有效避免叶片叶脉阴影和叶片边缘造成的色素腺体丢失问题，具有更好的分割精度和鲁棒性，F1 分数达到 94.18%。本文模型在保持高精度的基础上，兼顾了模型复杂度与推理效率，参数量适中且计算资源消耗合理，展现出良好的实际部署潜力，不仅适用于实验室高精度分析，也具备面向边缘计算和实时检测的应用价值。

参　考　文　献

[1] KOMALA M, GANESAN N M, KUMAR M. Genetic variability, heritability and correlation analysis in F_2 populations of ratoon upland cotton hybrids[J]. International Journal of Agriculture Environment and Biotechnology, 2018, 11(6): 815－827.

[2] GERASIMIDIS K, TRIANTAFILLOU D, BABATZIMOPOULOU M, et al. Preparation of an edible cottonseed protein concentrate and evaluation of its functional properties[J]. International Journal of Food Sciences and Nutrition, 2007, 58(6): 486－490.

[3] ZHANG J, MENG Y Z, CHEN Z P. A small target detection method based on deep learning with considerate feature and effectively expanded sample size[J]. IEEE Access, 2021, 9: 96559－96572.

[4] GAO W, XU F, LONG L, et al. The gland localized CGP1 controls gland pigmentation and gossypol accumulation in cotton [J]. Plant Biotechnology Journal, 2020, 18(7): 1573－1584.

[5] GUO H, QU Y, SUN S, et al. Antiviral effect of the cotton plant-derived gossypol against tomato yellow leaf curl virus[J]. Journal of Pest Science, 2023, 96(2): 635－646.

[6] KONG G, DAUD M, ZHU S. Effects of pigment glands and gossypol on growth, development and insecticide-resistance of cotton bollworm (*Heliothis armigera* (Hubner))[J]. Crop Protection, 2010, 29(8): 813－819.

[7] MOHAMMED M A, AMER N M, ABDALLAH H M I, et al. A comprehensive tool in recycling plant-waste of *Gossypium barbadense* L agricultural and industrial waste extracts containing gossypin and gossypol: hepatoprotective, anti-inflammatory and antioxidant effects[J]. Plant Methods, 2024, 20(1): 54.

[8] WANG W, YANG H, WANG Y, et al. Gossypol detoxification in the rumen and *Helicoverpa armigera* larvae: a review[J]. Animal Nutrition, 2021, 7(4): 967 - 972.

[9] PAUNOVIC D, RAJKOVIC J, NOVAKOVIC R, et al. The potential roles of gossypol as anticancer agent: advances and future directions[J]. Chinses Medicine, 2023, 18(1): 163.

[10] ZUBAIR H, AZIM S, KHAN H Y, et al. Mobilization of intracellular copper by gossypol and apogossypolone leads to reactive oxygen species-mediated cell death: putative anticancer mechanism[J]. International Journal of Molecular Sciences, 2016, 17(6): 973.

[11] SUNIKUMAR G, CAMPBELL L M, PUCKHABER L, et al. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol[J]. Proceedings of the National Academy of Sciences, 2006, 103(48): 18054 - 18059.

[12] GADELHA I C N, FONSECA N B S, OLORIS S C S, et al. Gossypol toxicity from cottonseed products[J]. The Scientific World Journal, 2014, 2014(1): 231635.

[13] LIU Y, ZHAI Y, LI Y, et al. Multiple strategies to detoxify cottonseed as human food source[J]. Frontiers in Plant Science, 2022, 13: 1080407.

[14] 崔光红,陈家春,蔡大勇.高效液相色谱法测定棉根皮中棉酚含量[J].中国中药杂志,2002,27(3):173 - 175.
CUI Guanghong, CHEN Jiachun, CAI Dayong. Determination of gossypol in cotton root bark by HPLC[J]. China Journal of Chinese Materia Medica, 2002, 27(3):173 - 175. (in Chinese)

[15] VSHIVKOV S, PSHENICHNOV E, GOLUBENKO Z, et al. Capillary electrophoresis to quantitate gossypol enantiomers in cotton flower petals and seed[J]. Journal of Chromatography B, 2012, 908: 94 - 97.

[16] LI C, ZHAO T, LI C, et al. Determination of gossypol content in cottonseeds by near infrared spectroscopy based on Monte Carlo uninformative variable elimination and nonlinear calibration methods[J]. Food Chemistry, 2017, 221: 990 - 996.

[17] ZANG Y, XU C, XUAN L, et al. Identification and characteristics of a novel gland-forming gene in the cotton[J]. The Plant Journal, 2021, 108(3): 781 - 792.

[18] JAN M, LIU Z, GUO C, et al. An overview of cotton gland development and its transcriptional regulation[J]. International Journal of Molecular Sciences, 2022, 23(9): 4892.

[19] 钱玉源,张海娜,崔淑芳,等.棉花色素腺体的研究进展[J].作物研究,2017,31(2):205 - 210.
QIAN Yuyuan,ZHANG Haina, CUI Shufang, et al. Research progress of cotton pigment gland[J]. Crop Research, 2017, 31(2): 205 - 210. (in Chinese)

[20] 王玉红,蔡小彦,张振梅,等.四个栽培棉种叶片色素腺体分布比较[J].中国棉花,2018,45(12):16 - 19.
WANG Yuhong, CAI Xiaoyan, ZHANG Zhenmei, et al. Comparison of distribution of pigment glands in leaves of four cultivated cotton varieties[J]. China Cotton, 2018, 45(12): 16 - 19. (in Chinese)

[21] OPITZ S, GRIT K, JONATHAN G. Increased terpenoid accumulation in cotton (*Gossypium hirsutum*) foliage is a general wound response[J]. Journal of Chemical Ecology, 2008, 34(4): 508 - 522.

[22] POUND M P, ATKINSON J A, TOWNSEND A J, et al. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping[J]. GigaScience, 2017, 6: 1 - 10.

[23] HAO S, ZHOU Y, GUO Y. A brief survey on semantic segmentation with deep learning[J]. Neurocomputing, 2020, 406: 302 - 321.

[24] RONNEBERGER O, FISCHER P, BROX T. U - Net: convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-assisted Intervention (MICCAI 2015), 2015: 234 - 241.

[25] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-Decoder with atrous separable convolution for semantic image segmentation[C]// Computer Vision-ECCV 2018, 2018: 833 - 851.

[26] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018: 6230 - 6239.

[27] CHEN J, LU Y, YU Q, et al. TransUNet: transformers make strong encoders for medical image segmentation[J]. arXiv: 2102.04306, 2021.

[28] JIN C, ZHOU L, PU Y, et al. Application of deep learning for high-throughput phenotyping of seed: a review[J]. Artificial Intelligence Review, 2025, 58: 76.

[29] LI Z, GUO R, LI M, et al. A review of computer vision technologies for plant phenotyping[J]. Computers and Electronics in Agriciculture, 2020, 176: 105672.

[30] 苏童,王琳,班超,等.基于语义分割的中后期玉米行间路径导航线检测[J].农业机械学报,2024,55(10):275 - 285.
SU Tong, WANG Lin, BAN Chao, et al. Interrow path navigation line detection of maize in middle and late period based on

semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 275 – 285. (in Chinese)

[31] 陈燕,陆嘉豪,胡小春,等.基于 VGG－UNet 的食用菌菌丝体表型参数自动测量方法[J].农业机械学报, 2024, 55(1): 233 – 240.
CHEN Yan, LU Jiahao, HU Xiaochun, et al. Automated measurement method of phenotypic parameters of edible mushroom mycelium based on VGG－UNet[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 233 – 240. (in Chinese)

[32] 武锦龙,吴虹麒,李浩,等.基于改进 DeepLabV3 + 的荞麦苗期无人机遥感图像分割识别方法研究[J].农业机械学报, 2024, 55(5): 186 – 195.
WU Jinlong, WU Hongqi, LI Hao, et al. Segmentation of buckwheat by UAV based on improved lightweight DeepLabV3 + at seedling stage[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 186 – 195. (in Chinese)

[33] WANG G, WANG X, ZHANG Y, et al. Dynamic characteristics and functional analysis provide new insights into long non-coding RNA responsive to *Verticillium dahliae* infection in *Gossypium hirsutum*[J]. BMC Plant Biology, 2021, 21: 68.

[34] XU Y, WANG G, SHAO L, et al. GlandSegNet: semantic segmentation model and area detection method for cotton leaf pigment glands[J]. Computers and Electronics in Agriculture, 2023, 212: 108130.

[35] ZHANG X, ZHANG Q, LIU H, et al. Mamba in speech: towards an alternative to self-attention[J]. arXiv: 2405.12609, 2024.

[36] LI K, CHEN G, YANG R X, et al. SPMamba: state-space model is all you need in speech separation[J]. arXiv: 2404.02063, 2024.

[37] 沈逸文,孙俊.结合 Transformer 的轻量化中文语音识别[J].计算机应用研究, 2023, 40(2): 424 – 429.
SHEN Yiwen, SUN Jun. Lightweight Chinese speech recognition with Transformer[J]. Application Research Computers, 2023, 40(2): 424 – 429. (in Chinese)

[38] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 1874 – 1883.

[39] LIU M, DAN J, LU Z, et al. CM－UNet: hybrid CNN－Mamba UNet for remote sensing image semantic segmentation[J]. arXiv: 2405.10530, 2024.

[40] SHE L, WANG N, XU Y, et al. Detection and counting of pigment glands in cotton leaves using improved U－Net[J]. Frontiers in Plant Science, 2023, 13: 1075051.