doi:10.6041/j.issn.1000-1298.2025.07.064

基于 Fishmeal – Mask2Former 分割模型的掺假鱼粉显微图像识别方法

牛智有[1,2]　李武瑛妮[1,2]　孔宪锐[1,2]　耿　婕[1,2]　王伟霞[1,2]

(1.华中农业大学工学院, 武汉 430070; 2.农业农村部智慧养殖技术重点实验室, 武汉 430070)

摘要: 针对传统显微镜检法在鱼粉掺假识别中面临的肉眼辨别难度大、识别效率低的问题,提出一种改进 Masked-attention Mask Transformer(Mask2Former)的鱼粉显微图像掺假识别模型,以实现高分辨条件下、复杂鱼粉背景中对异状多目标掺假特征的自动化识别与分割。以鱼粉中掺杂动物源性肉粉为研究对象,构建了 Fishmeal 掺假鱼粉显微图像数据集,通过形态学分析将掺假特征细分为异常肌肉、骨骼、皮肤、血液和毛发 5 类组织;开发基于颜色阈值相似性的交互式标注软件,实现逐像素的提示性精确标注;改进 Mask2Former 分割模型架构,融合 ResNet50 骨干网络、多头注意力机制和多尺度特征处理机制,增强了鱼粉多样化特征的融合效果;引入重复加权双向特征金字塔网络(Bidirectional feature pyramid network, BiFPN)改进像素解码器,提升了小目标分割能力;引入可学习的尺度级嵌入和掩码注意力(Masked Attention)模块,通过限制交叉注意力关注范围提高了模型细节表现力;在 Multi-head Attention 层后加入 dropout 操作防止过拟合。实验结果表明,改进后的 Fishmeal – Mask2Former 在样本真伪判别阶段整体分类准确率达到 98.56%,在精细分割阶段对异常掺假特征定性识别平均精确率达到 80.52%、召回率为 76.01%、F1 分数为 78.86%,分割模型平均准确度提升至 82.08%,较传统方法具有显著优势。此外,设计分割结果可视化界面,为鱼粉品质检测中的掺假检测环节提供了一种直观、准确与高效的显微视觉自动化识别方法。

关键词: 鱼粉; 掺假; 语义分割; 显微图像; Mask2Former; 计算机辅助诊断

中图分类号: S816.17　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)07-0679-12　　**OSID**:

Detection of Adulterants in Fishmeal Microscopic Images Based on Segmentation Model of Fishmeal – Mask2Former

NIU Zhiyou[1,2]　LI Wuyingni[1,2]　KONG Xianrui[1,2]　GENG Jie[1,2]　WANG Weixia[1,2]

(1. *College of Engineering*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*

2. *Key Laboratory of Smart Farming for Agricultural Animals*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*)

Abstract: Fishmeal, a high-value and necessary protein feed raw material, confronts difficulties in assessing authenticity. The conventional identification process is rendered challenging due to the sophistication of visual discrimination and the inefficacy of microscopic testing techniques by inspector, particularly in the presence of subtle and multiple adulterants. These issues were tackled by proposing an end-to-end segmentation model based on the Masked-attention Mask Transformer (Mask2Former) architecture. Focused on identifying multi-target adulteration feature in complex fishmeal backgrounds under high-resolution conditions to realize automatic recognition and segmentation. Fishmeal was used and adulterated with animal-derived meat powder as an example, an adulterated fishmeal microscopic image dataset was created and the adulteration characteristics were classified into five types of tissues: abnormal muscle, bone, skin, blood, and hair based on morphological analysis. Then an interactive annotation program that utilized color threshold similarity to provide pixel-by-pixel suggestive annotations was developed. The updated Mask2Former model integrated ResNet50, multi-head Attention mechanisms, and multi-scale feature processing mechanisms to extract global information and multi-resolution features. Bidirectional feature pyramid network (BiFPN) enhanced the pixel decoder, thereby improving the model's performance in processing small-scale targets. The incorporation of the Masked

收稿日期: 2025-01-06　修回日期: 2025-02-19

基金项目: 国家自然科学基金项目(32172773)

作者简介: 牛智有(1965—),男,教授,博士,主要从事农产品加工技术与装备研究,E-mail: nzhy@mail.hzau.edu.cn

Attention module limited the scope of cross-attention calculations, enabling the model to more effectively focus on the target areas within the mask. The implementation of learnable scale-level embeddings and the reordering of self-attention and cross-attention within the decoder section bolstered the model's feature learning capabilities. Furthermore, by adding dropout operations after the multi-head Attention layer, the improved Fishmeal - Mask2Former model achieved the overall accuracy rate of 98.56% in distinguishing genuine or adulterated samples. In the fine segmentation stage, the mean average precision of abnormal adulteration feature recognition reached 80.52 %, the mean average recall reached 76.01%, the mean F1 score reached 78.86%, and the segmentation mean accuracy was improved to 82.08%. The method proposed had significant advantages over traditional methods. Finally, a visual strategy with its interface of the segmentation results was designed, aimed to provide an intuitive, accurate, and efficient microscopic visual automated identification method for the fishmeal adulteration in its quality detection.

Key words: fishmeal; adulteration; semantic segmentation; microscopic image; Masked-attention Mask Transformer; computer-aided diagnosis

0 引言

鱼粉是一种营养价值极高的动物性蛋白质饲料原料,广泛应用于水产与畜牧业养殖中,其品质直接影响着动物的生长发育和生产效能[1]。然而,在鱼粉需求量与价格持续攀升的背景下,不法商家通常在鱼粉中掺入非蛋白氮、低质动植物蛋白及同源劣质鱼粉等降本牟利。掺假鱼粉不仅导致商品饲料质量下降,还可能危害动物健康并降低养殖效益[2]。因此,开发准确、快速的掺假检测方法对保障饲料安全、维护市场秩序以及促进养殖业的健康发展具有重要意义。

传统的鱼粉掺假检测方法主要依赖人工感官或理化分析,存在主观性强、效率低下的缺陷[3];现代分析方法中的各类光谱分析、分子生物学检测技术等[4-6],虽能提升掺假检测的精度,但设备昂贵、检测成本高,难以大规模应用;显微检测技术[3]则通常借助显微元器件的放大能力,由经验丰富、专业能力强的检验员进行镜下特征的观察与图谱比对,具有检测流程简洁、耗时短、成本消耗低等优势。但传统镜检法高度依赖人工主观判断,难以应对日愈复杂的掺假特征。为提升鱼粉显微图像检测的客观性与自动化水平,GENG 等[7]采用轻量化设计建立了 YOLO v3 - Mobilenetv2 模型用于检测鱼粉样本可能存在的掺假目标(豆粕与肉骨粉),模型平均精度可达 78.49%,但对于复杂鱼粉背景的处理、小目标边缘的精确捕捉仍存在不足。

近年来,深度学习技术的快速发展为显微图像分析提供了全新的技术路径,也为复杂场景中的多目标识别与异状物分割提供了算法思路[7-13]。例如基于 Mask R - CNN 的作物复杂显微表型解析[8]、农业领域的病虫害识别[9]等;对于复杂自然环境中多尺度目标的捕捉难题,可采用深度语义分割网络捕捉交叠边界特征[10]、通过多分支解码结构改进 DeepLabv3 + 网络[11]等,提升复杂背景中的多类别目标分割准确率;在高分辨率遥感场景中,查询机制的有效引入可以提升 Mask2Former 模型对小目标特征的高敏感性[12-14]。由此可见,图像分割算法的发展为实现鱼粉显微图像在复杂背景中的掺假识别提供了技术支撑[15]。

因此,针对人工显微镜检法在鱼粉掺假识别中面临的主观性强、辨别难度大、对困难特征的判读敏感性不足以及识别效率低等问题,本文以鱼粉中掺杂动物源性肉粉为研究对象,采用超景深数码显微镜获取实验样品图像、建立掺假鱼粉显微图像数据集;基于 Mask2Former 框架建立鱼粉显微图像掺假识别模型,通过集成重复加权 BiFPN 网络改进像素解码器、引入 Masked Attention 模块限制交叉注意力计算的关注范围、引入可学习的尺度级嵌入及使用合成少数类过采样算法(Synthetic minority over-sampling technique,SMOTE)扩展数据集,并在 Multi-head Attention 层后加入 dropout 等操作对模型进行改进,以实现 Fishmeal - Mask2Former 模型对鱼粉显微图像中复杂背景下的异状多目标掺假特征进行有效识别与分割;模型预测结果以伪彩色掩膜叠加的可视化策略,直观显示出异常掺假特征的轮廓位置,以期实现复杂场景下掺假鱼粉样本的快速判别与异常显微特征自动分割,提升鱼粉掺假检测的客观性与效率。

1 材料与方法

1.1 样品采集与制备

以鱼粉中掺入特征易混淆的动物源性蛋白原料肉粉为实验样品。根据 GB/T 14699—2023《饲料采样》采集具有代表性的纯物质原料样品:进口鱼粉(秘鲁鱼粉、泰国鱼粉)与肉粉(鸡肉粉、猪肉粉),所采集的样品均符合鱼粉品质检疫标准。

根据 GB/T 20195—2006《动物饲料 试样的制

备》与随机区组实验设计方法，分别制备含有猪肉粉、鸡肉粉的单组分掺假鱼粉样品与多组分掺假混合样品。掺假试样按质量分数为 1% ~5%（梯度 1%）、6% ~20%（梯度 2%）配制[16]，混合均匀后装袋、密封并标记，共计 520 份掺假试样。

为确保粉料样品在特定放大倍数视野中的显微目标粒度保持基本一致，使用分样板对制备好的每类掺假试样按照 GB/T 14698—2017《饲料原料显微镜检查方法》进行筛分前处理，得到 3 类不同粒度范围的镜检样品：分别为 40 目（孔径为 0.42 mm）筛上物、60 目（孔径为 0.25 mm）筛上物和筛底盘物，分别称取 2.5 ~3.0 g 筛分试料均匀置于培养皿（直径 100 mm）中。

1.2　图像采集与显微特征分析

本文采用基恩士 VHX－6000 型数码显微系统进行鱼粉显微图像的采集。光学配置为：VH－Z20R 型变焦镜头，放大倍数设定为 30 ~50 倍，系统帧率为 30 f/s、快门速度为 1/60 s。观测载物台使用白色衬板，光源采用散射环形 LED 光源。移动 XY 载物平台捕获分辨率为 1 600 像素 ×1 200 像素的显微图像，以.JPG 的格式存储于磁盘。本文共采集到 500 幅无掺假图像与 1 500 幅有效掺假图像作为鱼粉掺假识别的原始图像。显微图像采集平台与采集示例如图 1 所示。

根据国家标准饲料镜检图谱[17]、行业镜检图谱以及相关学术理论研究[18－20]，对采集到的镜下图像进行形态学识别与显微特征分析。按照组织学理论将掺入鱼粉中的异常肉粉特征划分为 5 个类别：异常肌肉组织、异常骨骼组织、异常皮肤组织、血液组织和毛发；真实的鱼粉特征则归类为背景。

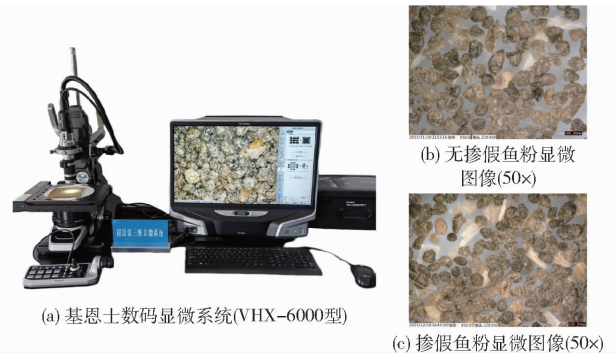

图 1　基恩士 VHX－6000 型显微图像采集平台与采集示例

Fig. 1　VHX－6000 microscopy systerm platform and image acquisition example

本实验所采集的鱼粉及其掺假物的部分显微特征如图 2 所示，其中，红色框为鱼粉中掺入肉粉的显微特征（视作异常掺假目标）；黑色框为纯品鱼粉显微特征（视作正常背景）。其中，同为动物源性蛋白原料的肉粉（猪肉粉、鸡肉粉）显微特征与鱼粉基质存在高度相似性，具体表现为动物源性成分中肌纤维的形态趋同、骨质碎片的颜色重叠、特征细节的高度相似等，肉眼难以完全辨别，呈现出组织结构的趋同性；此外，肉粉中的毛发、皮肤组织则可作为鉴定鱼粉中存在掺假的特异性判据。

1.3　数据集建立与掺假特征标注

通过 MySQL 数据库构建一个包含多维属性数据的 Fishmeal 鱼粉掺假识别显微图像数据集，存储信息包括样本来源、组分构成、前处理方法、图像元数据、显微图像的采集参数等，存储方式采用服务器本地存储。

图像标注阶段，为辅助标注者快速识别显微图

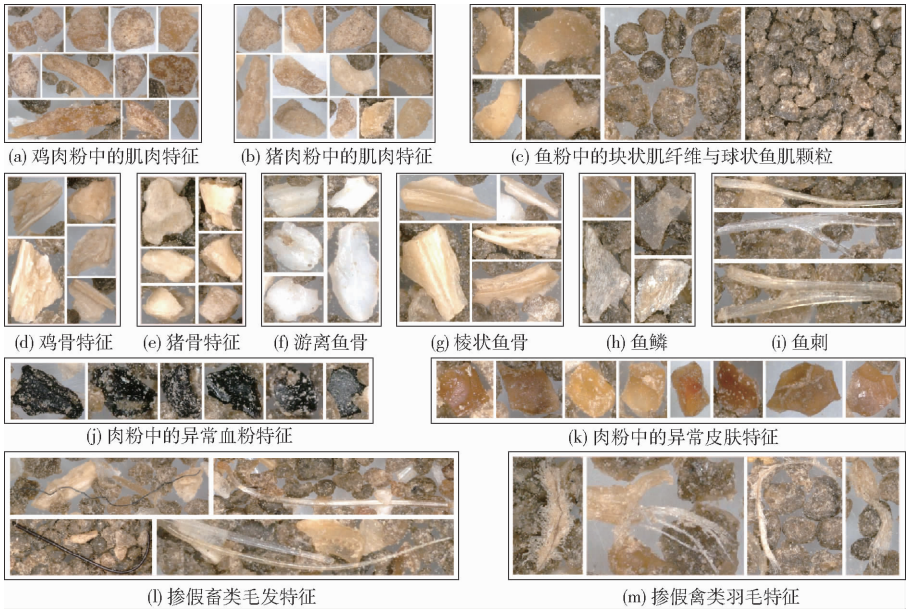

图 2　自建数据集中鱼粉及其掺假物的部分显微特征

Fig. 2　Some microscopic characteristics of fishmeal and its adulterants in self-built dataset

像中的异常不规则区域、尽可能避免人为漏标情形的出现，本文结合 SAM[21]分割技术建立近似阈值区域的边界查询[22]，设计了基于 EISeg（Efficient interactive segmentation）的动态提示性标注软件。对于显微图像中色度相近的复杂区域，可通过高斯核加权迭代优化边界像素归属、查询颜色相似的像素区域以建立边界，进而快速生成精确有效的特征轮廓提示性标注。标注文件以 JSON 格式存储，包含图像所含全部类别标签以及每个类别的颜色编码、图像索引和位置坐标，导出 PNG 格式的掩码图像作为真实标签。

标注完成后，数据集按照比例 7∶3 随机划分为训练集和测试集。选取训练集进行镜像、随机旋转、随机缩放大小、随机裁剪与随机光度变换等图像增强处理[23]，扩充训练数据至 4 000 幅以提高模型在实际应用中的泛化能力。测试集图像则保持原始状态，不进行任何预处理。显微图像数据集准备与特征标注流程图如图 3 所示。

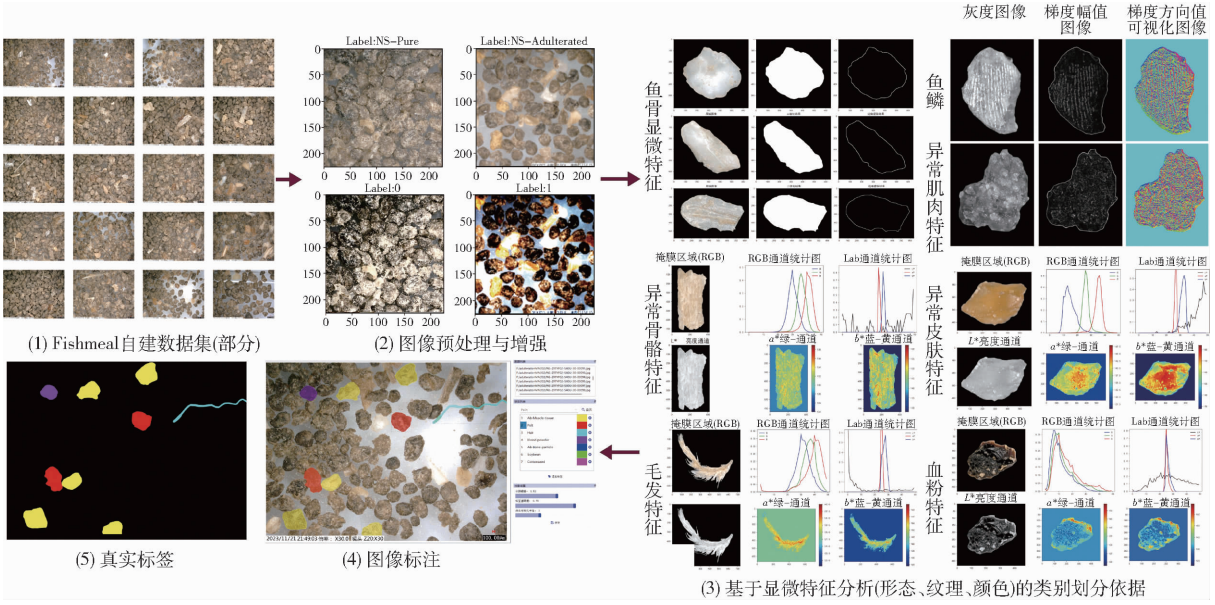

图 3 掺假鱼粉显微图像数据集准备与掺假特征标注流程

Fig. 3 Preparation of microscopic fishmeal dataset and abnormal feature labeling process

2 鱼粉掺假识别模型构建

2.1 Mask2Former 网络架构

Mask2Former 作为通用图像分割框架，在 MaskFormer[24]的掩膜分类范式基础上继承了 DETR（Detection transformer）框架的端到端型集合预测范式，具有基于 Masked Attention 的掩码注意力特征处理机制，能无缝地将逐像素分类模型转换为掩膜分类任务。对于语义和全景分割任务，Mask2Former 都受到相同的逐像素二进制掩膜损失和每个掩膜单一分类损失的监督[25]。Mask2Former 网络由 3 个核心模块组成：像素级模块（Pixel-level module）、Transformer 模块、分割预测模块（Segmentation module）。像素级模块通常负责生成图像中的逐像素嵌入，主要包含一个主干网络以提取低分辨率特征和一个像素解码器以融合局部到全局特征，用于实现局部到全局的信息融合；Transformer 模块可进行特征提取与融合，其将低分辨率特征上采样，生成高分辨率特征并处理对象查询，输出模型的掩膜和类别预测；分割预测模块则输出图像异常特征的轮廓与位置、所属类别及置信度。

2.2 Fishmeal－Mask2Former 算法改进

基于 Mask2Former 网络在算法层面的主要改进为：架构层面引入 Masked Attention 模块，通过局部特征强化机制提升异常组织的表征能力；改进像素解码器，集成重复加权 BiFPN 网络以优化多尺度特征融合效率，有效增强微小目标的捕捉精度；构建可学习特征查询机制，通过动态权重调整生成高置信度的掩码提议。改进后的 Fishmeal－Mask2Former 鱼粉显微图像掺假检测模型算法流程如图 4 所示。

2.2.1 像素级模块改进

使用预训练的 ResNet50 主干网络接收具有 3 个颜色通道（RGB）的高分辨率输入图像（3×1 600×1 200），进行多尺度特征的提取。构建 4 级特征金字塔$\{P_2,P_3,P_4,P_5\}$，通过深度可分离卷积对各级特征进行通道归一化，构建初始特征集并将特征张量送入像素解码器，如图 5 所示。

像素解码器使用重复加权 BiFPN 模块[26]双向整合高层级的强语义信息 P_5 与低层级的空间细节

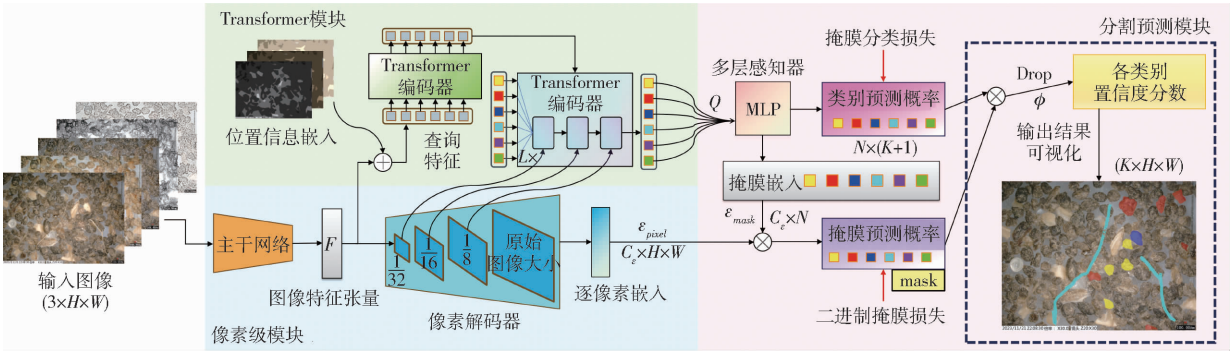

图 4　基于 Fishmeal－Mask2Former 的鱼粉显微图像掺假检测模型算法流程

Fig. 4　Algorithm flow of fishmeal microscopic image adulteration detection model based on Fishmeal－Mask2Former

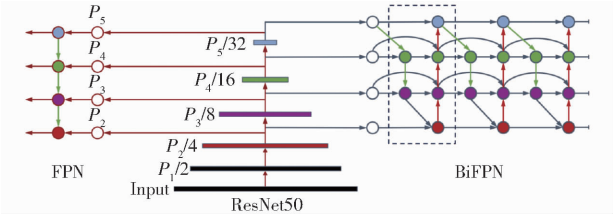

图 5　改进像素解码器中的 BiFPN 网络结构图

Fig. 5　Bidirectional feature pyramid network structure of improved pixel decoder

信息 P_2，替代传统的轻量级特征金字塔网络（Feature pyramid network，FPN）[27]，通过双向跨层连接和加权特征融合机制实现多级特征交互，以此优化多尺度特征的融合效果并确保重要特征和小目标特征得到有效关注。

2.2.2　Transformer 模块改进

（1）编码器特征提取

采用标准 Transformer 编码器处理从 ResNet50 骨干网络提取的多尺度特征。编码器核心改进为位置感知多头注意力的增加与残差特征的增强[28]。通过并行处理多个信息子空间的方式捕捉不同类型的特征关系，进一步增强模型的表达能力，而后将每层编码器输出通过残差连接与层归一化处理。Transformer 中的编码器与改进解码器结构如图 6 所示。

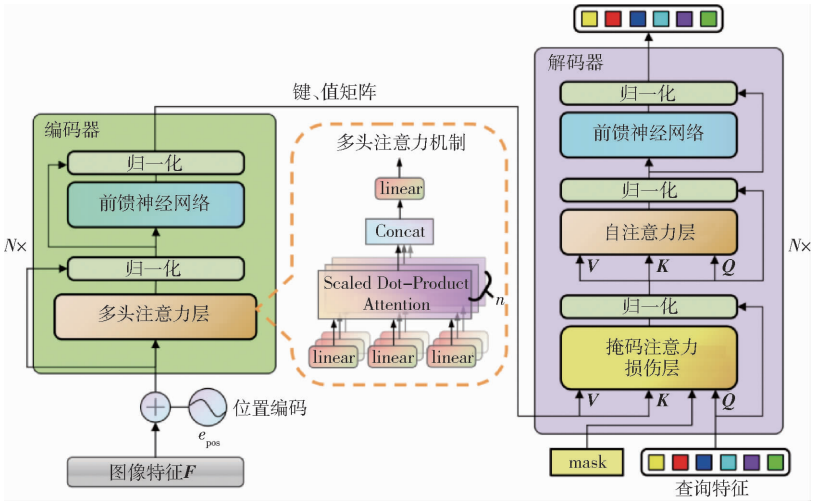

图 6　Transformer 模块中的编码器与改进解码器结构图

Fig. 6　Structure of encoder and optimized decoder in transformer module

序列化的图像映射记作输入长度为 N 的 $\boldsymbol{X}=(x_1,x_2,\cdots,x_N)$，维度在[257,768]范围内。为保留序列中各个 Patch 在原始图像中的空间位置信息，在编码器阶段引入基于正余弦编码的位置信息嵌入 $\boldsymbol{e}_{\text{pos}}$。输出的位置编码 $\boldsymbol{X}'=\boldsymbol{X}+\boldsymbol{e}_{\text{pos}}$，作为编码器输入向量参与自注意力模型的计算。多头注意力模型通过多组线性变化层（W_i^Q,W_i^K,W_i^V）将 $\boldsymbol{X}'$ 映射到 h 个不同的子空间，得到每个头查询矩阵 $\boldsymbol{Q}_i$、查询矩阵 $\boldsymbol{K}_i$ 和值矩阵 $\boldsymbol{V}_i$。

每个头的结果由前馈层拼接，经线性变换 W^O 得到最终输出，计算式为

$$\begin{cases}A_i=\text{softmax}\left(\dfrac{\boldsymbol{Q}_i\boldsymbol{K}_i^{\mathrm{T}}}{\sqrt{d_k}}\right)\\ \boldsymbol{C}_i=\boldsymbol{A}_i\boldsymbol{V}_i\end{cases}\tag{1}$$

$$\text{MultiHead}(\boldsymbol{Q}_{n\times d_k},\boldsymbol{K}_{n\times d_k},\boldsymbol{V}_{n\times d_v})=\text{Concat}(C_1,C_2,\cdots,C_h)\tag{2}$$

式中　h——注意力头数量，取 8

d_k——键向量维度

$\boldsymbol{A}_i$——每个注意力头的注意力权重矩阵

$\boldsymbol{C}_i$——加权分数

(2)解码器特征生成

使用集成掩码注意力机制与多尺度特征融合技术改进 Transformer 解码器。Transformer 解码器采用 3 层循环结构,通过掩码注意力实现多尺度特征融合。多个下采样版本特征分辨率为

$$\{F_l\}_{l=1}^{L}=\{F_{1/32},F_{1/16},F_{1/8}\} \tag{3}$$

式中 L——解码器循环层数,取 3

l——层索引

以连续映射的方式将像素解码器中第 l 层的多尺度特征 F_l 交替馈送到 Transformer 解码器的对应层级,每层融合位置编码 $\boldsymbol{e}_{\text{pos}}$和可学习尺度编码 e_{lvl}。引入的可学习尺度嵌入 e_{lvl}类似于一个区域提议网络(Region proposal network,RPN),在训练过程中允许模型根据不同特征图的分辨率动态调整特征尺度、自动学习并优化特征表示。该集成方法具备生成掩膜提议的能力,使模型能精确定位图像中不同尺度和形状的复杂目标对象并生成初步的掩膜区域。使用 $\boldsymbol{X}_l\in\boldsymbol{R}^{N\times C}$表达第 l 层的 N 个 C 维查询特征,原始表达式为

$$\boldsymbol{X}_l=\text{softmax}(\boldsymbol{Q}_l\boldsymbol{K}_l^{\mathrm{T}})\boldsymbol{V}_l+\boldsymbol{X}_{l-1} \tag{4}$$

$\boldsymbol{Q}_l$、$\boldsymbol{K}_l$、$\boldsymbol{V}_l$ 分别为第 l 层的查询、键、值矩阵,引入掩膜注意力操作符 Masked Attention。第 l 层解码器接收查询特征 $\boldsymbol{Q}_l$ 和掩码矩阵 $\boldsymbol{M}_{l-1}\in\{0,1\}^{N\times H_lW_l}$后,执行调制注意力计算,使用掩码矩阵 $\boldsymbol{M}$ 将注意力限制在以预测片段为中心的局部特征上,使模型专注于提议区域内的精确轮廓提取和像素级分割。查询特征 X_0 在输入到 Transformer 解码器之前是零初始化的,并与可学习的位置嵌入相关联。$\boldsymbol{M}_{l-1}$是特征位置(x,y)处的注意力掩膜表达式,由前 $l-1$ 个 Transformer 解码器层的二值化掩膜预测输出(阈值为 0.5),将其调整为与 $\boldsymbol{K}_l$ 相同的分辨率。由于传递给第 1 层自注意力的查询特征与图像无关,没有来自图像的信号,因此 $\boldsymbol{M}_0$ 在查询特征送入 Transformer 解码器之前直接从 X_0 处获得二进制掩膜预测。调制后注意力矩阵计算式为

$$\boldsymbol{X}_l'=\text{softmax}(\boldsymbol{M}_{l-1}+\boldsymbol{Q}_l\boldsymbol{K}_l^{\mathrm{T}})\boldsymbol{V}_l+\boldsymbol{X}_{l-1} \tag{5}$$

$$\boldsymbol{M}_{l-1}(x,y)=\begin{cases}0 & (\boldsymbol{M}_{l-1}(x,y)=1)\\ -\infty & (\text{其他})\end{cases} \tag{6}$$

解码器的输出与多层感知机(Multilayer perceptron,MLP)相连,最终由 MLP 输出一个用于生成每个像素类别概率分布的卷积层,表达式为

$$P(x)=\text{softmax}(W_f\text{MLP}(F)) \tag{7}$$

式中 $P(x)$——像素 x 类别概率分布

W_f——MLP 权重

F——从 Transformer 解码器输出的特征

2.2.3 分割预测模块

分割预测模块使用 Mask RCNN 的两阶段架构[29],引入 ROI Align 层实现语义分割的像素精准特征对齐,消除量化误差。通过 MLP 并行生成每个像素的类别预测概率和相应的掩膜嵌入向量;同时,利用交叉注意力层将掩膜嵌入向量与像素解码器输出的逐像素特征进行点积操作,经 sigmoid 激活函数激活生成二进制掩膜预测。类别预测表示模型对每个像素属于特定类别的置信度,掩膜预测则是对每个输入特征的空间分布编码。最终的分割结果通过类别预测概率与二进制掩膜预测的点积结果确定,应用阈值选择输出每个像素的类别标签与分割掩膜预测置信度。

2.3 模型评价指标

将鱼粉显微图像的掺假识别过程分为真伪判别阶段与精细分割阶段。依据图像分割任务的有监督评价法则[30],真伪判别阶段的评价指标为整体分类准确率;精细分割阶段的评价指标选取平均精确率、平均召回率、平均 F1 分数、平均准确度、平均交并比和平均 Dice 系数 6 个评估指标。

2.4 实验环境和超参数配置

实验环境基于 Linux 系统,硬件设备包括 AMD Ryzen 5 7600X 6 - Core 3Processor 处理器和 NVIDIA GeForce RTX 4060 显卡。深度学习框架采用 PyTorch 2.0.1、CUDA 11.6,在 MM Segmentation 基础上构建改进算法。实验在配备 8 GB 显存的工作站进行 GPU 加速训练,操作系统为 Ubuntu 18.04。

网络训练的主要超参数配置:优化器选择 AdamW[31],动量设置为 0.900;所有模型的主干网络初始学习率设置为 0.000 1,权重衰减系数为 0.05;实验中每个模型训练设置迭代次数(Epoch)为 40 000 次,批处理尺寸(Batch_size)均设置为 8;所有模型的骨干网络均在 ImageNet - 1K 数据集上进行权重参数的预训练;训练初期冻结模型主干,保持特征提取网络不变仅对网络进行微调,训练后期解冻网络,进行全网络参数的更新;学习率按周期性余弦函数自适应调整;为解决样本不平衡问题,引入 Focal Loss;为减少全局优化过程的网络损失震荡,设置线程数为 2。

2.5 模型训练

在 Transformer 解码器的每个中间层以及在解码器之前的查询特征上添加辅助损失,利用每个掩膜的分类损失和逐像素二进制掩膜损失进行最小化损失函数的监督训练。L_{BCE}为二元交叉熵损失,L_{Dice}

为 Dice 损失；$\boldsymbol{P}$ 为预测结果矩阵，$\boldsymbol{T}$ 为标签矩阵，M 为特征图通道数；整体损失函数 L_{total} 的混合加权计算式为

$$L_{total}=\alpha L_{cls}+\beta L_{mask}=\alpha\lambda_{cls}L_{cls}+\beta(\lambda_{ce}L_{ce}+\lambda_{dice}L_{dice}) \tag{8}$$

其中

$$L_{ce}=\frac{1}{M}\sum_{C=1}^{M}[\boldsymbol{T}\ln\boldsymbol{P}+(1-\boldsymbol{T})\ln(1-\boldsymbol{P})] \tag{9}$$

$$L_{dice}=1-2\sum_{C=1}^{M}\boldsymbol{TP}\Big/\left(\sum_{C=1}^{M}\boldsymbol{T}+\boldsymbol{P}\right) \tag{10}$$

式中　L_{mask}——掩膜损失

L_{cls}——分类损失

α——分类损失权重，取 0.6

β——二进制掩膜损失权重，取 0.4

λ——不同损失项可调权重

其中，λ_{BCE} 为 5.0，λ_{Dice} 为 5.0，λ_{cls} 则根据预测是否与真实标签匹配设置为 2.0 或 0.1。

3　结果与分析

3.1　不同识别模型对比

为测试图像分割算法在鱼粉掺假识别任务中的适用性与有效性，采用 K－Net、PSPNet、DeepLabv3＋、FastSCNN、SegFormer、U－Net 和改进的 Mask2Former 模型对自建 Fishmeal 数据集中的测试集进行分割结果的对比实验。首先，随机选取一批特征分布状态不同的 Fishmeal 测试集输入对比模型，并将 7 种语义分割算法在同一测试集上的分割效果进行图像细节输出。从左往右依次为原始图像，真实标签，以及 K－Net、Mask2Former、PSPNet、DeepLabv3＋、FastSCNN、SegFormer、U－Net 网络的分割效果，从上往下则是随机抽取的 30～50 放大倍数下的稀疏、密集、平铺状态下的显微图像样本。细节对比采用突出显示进行表达，如图 7 所示。

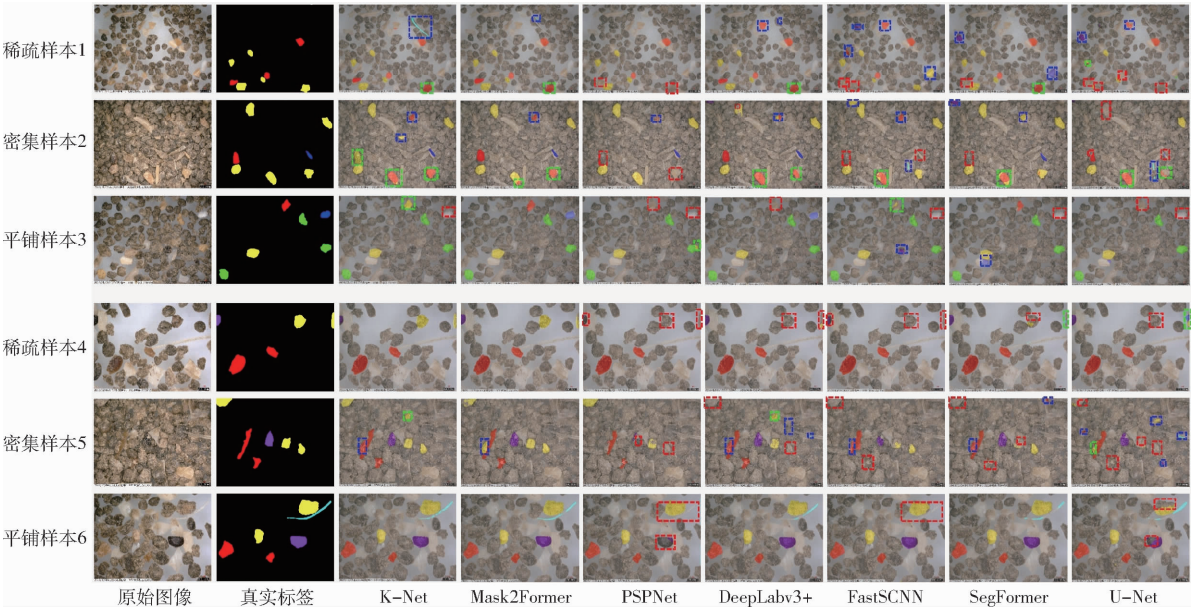

图 7　不同语义分割模型的鱼粉掺假检测分割效果对比

Fig. 7　Comparison of segmentation effects of different semantic segmentation models

图中，红色框代表预测样本区域漏检异常特征的情况（假阴性），绿色框代表预测的异常对象被检出但定性分类错误（假阳性），蓝色框代表不该被检出的正常鱼粉区域被错误检出并错误定性（假阳性）。绿色框和蓝色框统称为误检。

从图 7 可以看出，自建的 Fishmeal 数据集在各主流分割模型上均呈现出相对稳定的分割效果，表明 30～50 倍范围内的数码显微镜放大倍率能满足肉眼与计算机的视觉分辨条件，且该显微倍率下的特征视野与尺度不会对分割效果产生负面影响。另外，镜下视野中的目标特征密度排布状态不同可能会引发模型的性能表现的差异。其中，密集样本 2、5 在所测试的 7 个分割模型中均出现了明显的漏检、误检情况，而稀疏样本 1、4 与平铺样本 3、6 在分割模型中的表现普遍较好。这一现象可归因于稀疏或平铺状态下的试料，其显微视野中呈现出的单颗粒独立特征与清晰轮廓更易于被模型识别，进而可降低边界误判的概率。相反，在特征密集的样本图像中，视域内多颗粒之间存在轮廓的相互遮挡、重叠，导致各特征之间的边界不清晰，由此会增加模型的识别难度。进一步统计测试结果，发现密集样本在真伪判别阶段，模型的整体分类准确率普遍低至 84%，而精细分割阶段，模型的平均准确度甚至低于 72%，密集样本中通常仅能检测出颜色差异较大的异常区域，对色差较小的细微异常特征则容易忽略。由此表明样本排布密集可能会引发识别模型产生较

为严重的错检,导致模型的整体分割性能有所下降。

因此,排除密集样本 2、5,对各模型的分割效果进行比较。首先,K - Net 模型的整体分割效果较好,但在处理相似特征时存在一定程度的假阳性误检:样本 1 中 K - Net 将正常的鱼刺特征误判为毛发特征,样本 2、5 中将正常的鱼粉特征误判为异常肌肉组织,说明 K - Net 基于局部阈值分割的方式更加关注图像的细节特征,但对于背景较为复杂的鱼粉样本,掺假识别过程中存在灵敏度过高而导致误检极多的弊端。另外,样本 1、2、5 中发生的误检情况在其他模型中也普遍存在,均出现在异常肌肉特征和异常皮肤组织之间,这是由于掺假物肉粉的肌肉组织与皮肤组织本身在显微镜下所呈现的形态与颜色特征极为相似。PSPNet 模型对于各类别异常特征的定性识别具有不错效果:在处理异常皮肤组织的边界信息时尤其表现良好,分割对象的边缘平滑度和形状表达与个体标签较为接近,但是 PSPNet 在识别全局图像的掺假区域时,对小尺度的异常细粒度特征几乎无法准确识别,容易造成较多细节信息的遗漏而导致极多的漏检情况,因此,PSPNet 也不适合用于背景较为复杂、类内特征较为多样的鱼粉掺假检测任务。Mask2Former 在分割结果上与真实标签最为接近,该模型在样本 3、4、6 中均未出现任何漏检情况,说明 Mask2Former 能够均衡地关注多目标特征和全局信息。Deeplabv3 + 与 SegFormer 检出的毛发特征表现出最为明显的空洞现象与分割边缘的较差连续性。对于边界清晰的稀疏样本 1、4,FastSCNN 与 U - Net 呈现出较多的错误分割结果,表明二者并不能较好地建立鱼粉显微图像的全局信息。

为了更客观地评估各个模型的性能,进一步采用整体分类准确率来评价真伪判别阶段的模型表现,采用平均精确率、平均召回率、平均 F1 分数、平均准确度、平均交并比和平均 Dice 系数来比较各分割模型在掺假鱼粉特征精细分割阶段的表现。测试结果如表 1 所示。

表 1 主流分割模型在 Fishmeal 数据集上的性能比较

Tab. 1 Performance comparison of segmentation models on fishmeal dataset %

模型	整体分类准确率	平均精确率	平均召回率	平均 F1 分数	平均准确度	平均交并比	平均 Dice 系数
K - Net	98.74	83.43	71.02	77.64	77.11	65.76	77.64
Mask2Former	98.13	79.68	75.21	78.02	81.30	66.35	78.02
PSPNet	98.11	85.14	56.13	63.22	62.22	54.80	63.26
DeepLabv3 +	98.02	80.83	64.17	70.89	70.26	60.20	70.87
FastSCNN	97.97	82.27	60.36	74.11	66.45	54.92	64.39
SegFormer	98.09	79.71	67.11	72.69	73.20	60.96	72.69
U - Net	97.55	73.59	57.03	64.40	63.12	52.65	64.40

从表 1 可以看出,所有模型在真伪判别阶段的整体分类准确率均超过 97%,表明语义分割模型能有效捕捉掺假正样本的特征,并对单幅图像进行第一阶段的“正常样本”与“疑似异常样本”的高准度分类。

Mask2Former 在平均召回率(75.21%)、平均 F1 分数(78.02%)指标上表现最优,说明模型在保持正负样本判别高准确率的同时,能有效地分割识别出大部分的异常区域,可实现精确率和召回率之间的良好平衡,符合 Mask2Former 在复杂场景分割中的优势。另外,Mask2Former 在平均准确度(81.30%)、平均交并比(66.35%)和平均 Dice 系数(78.02%)上也取得了最高分,表明模型在掺假鱼粉特征精细分割判别阶段具有显著优势,能够较为准确地识别和定位异常特征。Mask2Former 模型优异表现可能与其交叉注意力机制对目标区域的聚焦能力有关,这与现有研究中对 Transformer 架构分割性能的结论一致。

而 K - Net 的整体分类准确率(98.74%)虽高,但其在平均召回率(71.02%)、平均交并比(65.76%)等细粒度指标上表现欠佳,表明其全局特征提取机制虽能有效区分正常与异常样本,但对局部异常区域的分割存在类别间性能差异,可能与动态卷积核的适应性限制有关。PSPNet 平均精确率(85.14%)显著优于其他模型,但其平均召回率(56.13%)为最低值,表明其金字塔池化模块虽能有效抑制背景误判(低假阳性),但对真实异常区域的漏检率较高(高假阴性),这与其对全局上下文信息的强依赖性和局部细节感知不足的特性相关。U - Net 的平均精确率(73.59%)与平均交并比(52.65%)显著低于其他模型,反映其跳跃连接结构在复杂鱼粉场景下的特征融合效率不足,模型均衡性差且无法适应复杂鱼粉背景中的细小特征提取任务。

Mask2Former 在平均召回率(75.21%)、平均 F1 分数(78.02%)和平均准确度(81.30%)3 项核心指

标上体现出最优的综合性能；K－Net 虽在平均精确率和整体分类准确率取得较高值，但平均召回率与平均交并比显著落后于 Mask2Former，存在明显的精度-召回失衡；U－Net 在所有对比模型中表现最弱，尤其在平均精确率和平均交并比上与其他模型差距显著。因此，Mask2Former 在对比实验中展现出了漏检率极低、误检情况较少的均衡优势，说明本文搭建的 Mask2Former 模型对解决鱼粉显微图像异常区域的分割识别任务具有独特优势，适应于高精度和高召回率的复杂场景。

3.2　Fishmeal－Mask2Former 模型改进

在 Mask2Former 模型验证阶段，对各类掺假特征在模型中的分割表现进行分析并改进，绘制各类别掺假特征识别精确率（Precision）随迭代次数的变化曲线，如图 8 所示。

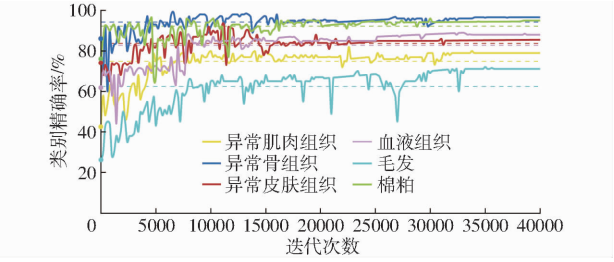

图 8　各类别异常特征在验证阶段的精确率变化曲线

Fig. 8　Precision of abnormal feature classes in validation phase

随着迭代次数的增加，各类别精确率均呈持续上升趋势并且在 35 000 次附近趋于平稳。说明模型在训练过程中可学习到数据中的有效特征并且模型稳定收敛。其中，各类别特征精确率均值分别为异常肌肉组织 74.94%、异常骨组织 94.25%、异常皮肤组织 83.75%、血液组织 82.83%、毛发 62.46%、植物性异常特征 92.19%，各组精确率均值表明 Fishmeal－Mask2Former 模型分别对上述 6 种易混淆掺假物质的识别能力，该方法的准确度与有效性均显著优于传统人工镜检法。Fishmeal－Mask2Former 模型对于镜下异常特征较明显的简单掺假情形来说，以鱼粉中掺入植物性原料为例的单类别识别均值达到 92.19%，这为模型的多类别泛化扩充提供了可能性。但是，模型在毛发、异常骨骼组织中呈现出了较大的类间精确性差异，这可能是各类别特征数量分布不平衡引发的性能波动。由此统计参与 Mask2Former 模型训练与测试的各类别特征数量，如图 9 所示。

从图 9 可知，参与模型训练与测试的各类别特征数量的确存在分布不平衡的情况。其中，毛发与骨骼组织的特征数量明显低于其他类别，这是由于

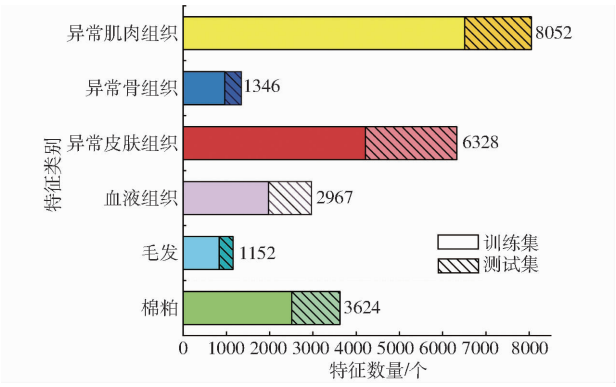

图 9　参与模型训练与测试的各类别特征数量

Fig. 9　Feature numbers in model training and testing

本文制备的掺假试样本身组分的质量不均衡（自然状态下，肉粉中所含的肌肉、皮肤、血液等组分超过 80%，而骨骼颗粒与羽发特征低于 15%），由此导致毛发类数据特征出现轻微欠拟合、骨骼类特征出现轻微过拟合的性能波动。因此，使用 SMOTE 算法[32]扩展前述两个不平衡数据集，并在 Multi-head Attention 层后加入 dropout 操作规避拟合现象，模型改进后训练损失对比如图 10 所示。

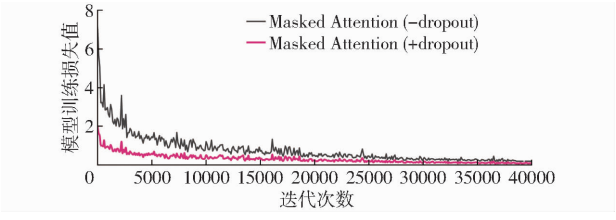

图 10　模型改进后训练损失对比

Fig. 10　Comparison diagram of training loss value

更换注意力以验证掩码注意力的有效性，测试结果如表 2 所示。与没有 dropout 操作的 Mask2Former 相比，Fishmeal－Mask2Former 模型的各项指标平均提升 0.72 个百分点，其中平均精确率和平均 F1 分数均增长 0.84 个百分点，可见加入 dropout 操作可帮助模型更快收敛并提高性能。

3.3　分割结果可视化

Fishmeal－Mask2Former 模型训练完成后，需部署到实际的应用场景中辅助检验人员快速评估鱼粉质量。为分别呈现模型真伪判别阶段与精细分割阶段的掺假预测输出结果，设计一种直观有效的可视化界面，如图 11 所示。左部分显示待检样本的原始信息，包括样品来源、前处理方式、图像采集设备等；右部分对应显示该样本的检测结果，即单幅图像的预测类别（“疑似掺假”或“无掺假”）及真伪判别置信度分数。被模型初步判别为“疑似掺假”的显微图像（阈值为 0.8）则继续输入至精细分割阶段，并以颜色映射、掩膜叠加的方式将模型对掺假区域实施异常类别识别的结果直接覆盖显示于原图。创建

表 2 更换注意力模块与加入 dropout 操作的模型测试结果

Tab. 2 Performance comparison of replacing attention module and adding dropout operation %

模型	改动模块	整体分类准确率	平均精确率	平均召回率	平均 F1 分数	平均准确度	平均交并比	平均 Dice 系数
Mask2Former	cross-attention[33]	98.02	77.20	72.63	76.68	78.15	64.05	77.64
	SMCA[34]	98.08	78.31	73.22	77.74	79.22	65.20	77.80
	mask pooling[35]	98.11	78.55	74.38	77.89	80.28	65.80	77.98
	Masked Attention (－dropout)	98.13	79.68	75.21	78.02	81.30	66.30	78.02
Fishmeal－Mask2Former	Masked Attention (＋dropout)	98.56	80.52	76.01	78.86	82.08	66.92	78.75



图 11 鱼粉掺假识别结果的可视化界面

Fig. 11 Visual interface of fishmeal adulteration detection results

显式颜色表为每个类分别指定唯一的颜色图例并生成对应的语义图，使用 cv2. addWeighted 函数将语义图和原图合并，利用 cv2. drawContours 函数绘制边缘以突出显示边界，最后在保留原始图像底层结构的基础上，以伪彩色掩膜可视化的方式清晰显示单幅图像的异常预测区域及各掺假类别预测概率分数。此外，设置“人在环机器学习”辅助环节，支持用户对模型的预测输出进行人为反馈和干预，实现鱼粉复杂掺假情形的计算机辅助诊断。

4 结论

(1)明确了鱼粉掺假显微视觉检测的图像采集条件，即两层筛法筛分前处理(40 目/60 目)后，分别取其筛上物定量平铺，在放大倍数为 30～50 的视野范围内采集排布状态相对稀疏的数码显微图像。该条件下模型输入数据可满足肉眼与计算机的视觉分辨条件且不会对分割效果产生负面影响。

(2)针对鱼粉中掺入动物源性肉粉的掺假情况，搭建的 Mask2Former 分割模型对于单幅鱼粉显微图像的真伪判别准确率可达 98. 13%。在具体异常特征的定性识别阶段，Mask2Former 模型对各类别特征的精确率均值分别为异常肌肉组织 74. 94%、异常骨组织 94. 25%、异常皮肤组织 83. 75%、血液组织 82. 83%、毛发 62. 46%，此外，对于异常特征较易区分的植物性掺假类别，其掺假物检出精度均值可达 92. 19%。较传统的主观肉眼判读来说，Mask2Former 分割模型具有显著优势。

(3)进一步使用 SMOTE 算法解决关于各类别特征数量分布的不平衡问题，扩展数据集并在 Multi-head Attention 层后加入 dropout 操作，改进后 Fishmeal－Mask2Former 掺假鱼粉显微图像分割模型的各项指标平均提升 0. 72 个百分点，真伪样本判别阶段的准确率达到 98. 56%，精细分割阶段对异常掺假特征识别的平均精确率达到 80. 52%、召回率为 76. 01%、F1 分数为 78. 86%，其中平均精确率和平均 F1 分数均增长 0. 84 个百分点，分割模型平均准确度提升至 82. 08%。

(4)设计模型推理阶段输出结果的可视化界面，直接输出单幅图像的真伪类别预测以及模型对掺假区域进行异常类别识别的分割结果，以高视觉区分度的伪彩色图例效果直观呈现。同时支持用户对模型的预测输出进行人为反馈和干预。

参 考 文 献

[1] 钱妤，蒋科技，王希挺，等. 中国鱼粉鱼油来源及利用[J]. 渔业信息与战略，2020，35(2)：91－100.
QIAN Yu，JIANG Keji，WANG Xiting，et al. Sources and utilization of fish meal and fish oil in China[J]. Fishery Information and Strategy，2020，35(2)：91－100.（in Chinese）

[2] 龙凡，肖俊，陈文静，等. 动物性蛋白源在水产饲料中替代鱼粉的研究进展[J]. 中国饲料，2023，68(11)：1－10.
LONG Fan，XIAO Jun，CHEN Wenjing，et al. Research progress on replacement of fish meal by animal protein sources in aquatic feed[J]. China Feed，2023，68(11)：1－10.（in Chinese）

[3] 葛敏，田蕴，王益军，等. 掺假鱼粉的识别方法[J]. 当代畜牧，2016(36)：24－25.

[4] 杨增玲，韩鲁佳，刘贤，等. 鱼粉中肉骨粉的可见-近红外光谱快速定性判别方法[J]. 农业工程学报，2009，25(7)：308－311.
YANG Zengling，HAN Lujia，LIU Xian，et al. Rapidly qualitative discrimination of meat and bone meal in fishmeal by visible and near infrared reflectance spectroscopy[J]. Transactions of the CSAE，2009，25(7)：308－311.（in Chinese）

[5] 朱瀛，高冰，史卓林，等. 基于机器学习的鸡、牛骨颗粒 Micro－CT 原位可视化鉴别[J]. 农业机械学报，2023，54(4)：394－398，438.
ZHU Ying，GAO Bing，SHI Zhuolin，et al. Machine learning based 3D in situ visual discriminant analysis of mammalian and non-mammalian bone meals by Micro－CT[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(4)：394－398，438.（in Chinese）

[6] 姚玉梅，韩鲁佳，杨增玲，等. 基于显微图像处理的不同种属肉骨粉鉴别分析[J]. 农业机械学报，2016，47(增刊)：347－352.
YAO Yumei，HAN Lujia，YANG Zengling，et al. Discriminant analysis of different species of meat and bone meal based on microscopic image processing[J]. Transactions of the Chinese Society for Agricultural Machinery，2016，47(Supp.)：347－352.（in Chinese）

[7] GENG J，LIU J，KONG X，et al. The fishmeal adulteration identification based on microscopic image and deep learning[J]. Computers and Electronics in Agriculture，2022，198：106974.

[8] 翁海勇，李效彬，肖康松，等. 基于 Mask R－CNN 的柑橘主叶脉显微图像实例分割模型[J]. 农业机械学报，2023，54(7)：252－258.
WENG Haiyong，LI Xiaobin，XIAO Kangsong，et al. Instance segmentation model for microscopic image of citrus main leaf vein based on Mask R－CNN[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(7)：252－258.（in Chinese）

[9] 王金星，马博，王震，等. 基于改进 Mask R－CNN 的苹果园害虫识别方法[J]. 农业机械学报，2023，54(6)：253－263，360.
WANG Jinxing，MA Bo，WANG Zhen，et al. Pest identification method in apple orchard based on improved Mask R－CNN[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(6)：253－263，360.（in Chinese）

[10] 顾文娟，魏金，阴艳超，等. 基于改进 DeepLabv3＋的番茄图像多类别分割方法[J]. 农业机械学报，2023，54(12)：261－271.
GU Wenjuan，WEI Jin，YIN Yanchao，et al. Multi-category segmentation method of tomato image based on improved DeepLabv3＋[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(12)：261－271.（in Chinese）

[11] 熊俊涛，刘柏林，钟灼，等. 基于深度语义分割网络的荔枝花叶分割与识别[J]. 农业机械学报，2021，52(6)：252－258.
XIONG Juntao，LIU Bolin，ZHONG Zhuo，et al. Litchi flower and leaf segmentation and recognition based on deep semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(6)：252－258.（in Chinese）

[12] GUO S，YANG Q，XIANG S，et al. Mask2Former with improved query for semantic segmentation in remote-sensing images[J]. Mathematics，2024，12(5)：765.

[13] JING Z. ConvNeXt－Mask2Former：a semantic segmentation model for land classification in remote sensing images[C]//IEEE Image and Deep Learning（CVIDL），2024：676－682.

[14] CHENG B，MISRA I，SCHWING A G，et al. Masked-attention mask transformer for universal image segmentation[J]. arXiv e-prints，arXiv.2112.01527，2021.

[15] 田萱，王亮，丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报，2019，30(2)：440－468.
TIAN Xuan，WANG Liang，DING Qi. Review of image semantic segmentation based on deep learning[J]. Journal of Software，2019，30(2)：440－468.（in Chinese）

[16] 杨凯，何昱廷，李沃霖，等. 基于近红外光谱技术检测全蛋粉掺假[J]. 华中农业大学学报，2024，43(2)：264－272.
YANG Kai，HE Yuting，LI Wolin，et al. Detection of adulteration in whole egg powder based on near-infrared spectroscopy[J]. Journal of Huazhong Agricultural University，2024，43(2)：264－272.（in Chinese）

[17] GB/T 34269—2017　饲料原料显微镜检查图谱[S].

[18] 李桢桢，尹明雨，王红丽，等. 水产品肌肉组织微观结构变化及其检测方法研究进展[J]. 食品科学，2023，44(9)：278－286.
LI Zhenzhen，YIN Mingyu，WANG Hongli，et al. Microstructural change of muscle tissues of aquatic products and methods for its detection：a review[J]. Food Science，2023，44(9)：278－286.（in Chinese）

[19] YOON C H, BONES P J, MILLANE R P. Image analysis for electron microscopy of muscle fibers[C]// Digital Image Computing: Techniques & Applications, Dicta. IEEE, 2005.
[20] 唐嘉潞, 杨钟亮, 张淞, 等. 结合显微视觉和注意力机制的毛羽检测方法[J]. 智能系统学报, 2022, 17(6): 1209 - 1219.
TANG Jialu, YANG Zhongliang, ZHANG Song, et al. Detection of yarn hairiness combining microscopic vision and attention mechanism[J]. CAAI Transactions on Intelligent Systems, 2022, 17(6): 1209 - 1219. (in Chinese)
[21] 王淼, 黄智忠, 何晖光, 等. 分割一切模型 SAM 的潜力与展望:综述[J]. 中国图象图形学报, 2024, 29(6): 1479 - 1509.
WANG Miao, HUANG Zhizhong, HE Huiguang, et al. Potential and prospects of segment anything model: a survey[J]. Journal of Image and Graphics, 2024, 29(6): 1479 - 1509. (in Chinese)
[22] 张悦, 王晓懿, 李娟. 面向图像分割的数据标注系统[J]. 铁路通信信号工程技术, 2022, 19(11): 47 - 53.
ZHANG Yue, WANG Xiaoyi, LI Juan. Data annotation system for image segmentation[J]. Railway Signalling & Communication Engineering, 2022, 19(11): 47 - 53. (in Chinese)
[23] 张明华, 龙腾, 宋巍, 等. 基于改进 Cascade R - CNN 和图像增强的夜晚鱼类检测[J]. 农业机械学报, 2021, 52(9): 179 - 185.
ZHANG Minghua, LONG Teng, SONG Wei, et al. Object detection of underwater fish at night based on improved Cascade R - CNN and image enhancement[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 179 - 185. (in Chinese)
[24] CHENG B, SCHWING A G, KIRILLOV A. Per-pixel classification is not all you need for semantic segmentation[J]. Advances in Neural Information Processing Systems, 2021, 34: 17864 - 17875.
[25] YAO M, AICH A, SUH Y, et al. Efficient Transformer Encoders for Mask2Former-style models[J]. arXiv preprint, arXiv: 2404.15244, 2024.
[26] DOHERTY J, GARDINER B, KERR E, et al. BiFPN - YOLO: one-stage object detection integrating bi-directional feature pyramid networks[J]. Pattern Recognition, 2025, 160: 111209.
[27] LIN T, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 - 2125.
[28] 宋彦,杨帅,郑子秋,等.基于多头自注意力机制的茶叶采摘点语义分割算法[J].农业机械学报,2023,54(9):297 - 305.
SONG Yan, YANG Shuai, ZHENG Ziqiu, et al. Semantic segmentation algorithm based multi-headed self-attention for tea picking points[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 297 - 305. (in Chinese)
[29] 李旺, 陶洋. Mask Boundary R - CNN 实例分割边界精度算法[J]. 小型微型计算机系统, 2024, 45(7): 1702 - 1709.
LI Wang, TAO Yang. Mask Boundary R - CNN instance segmentation boundary accuracy algorithm[J]. Journal of Chinese Computer Systems, 2024, 45(7): 1702 - 1709. (in Chinese)
[30] 于营, 王春平, 付强, 等. 语义分割评价指标和评价方法综述[J]. 计算机工程与应用, 2023, 59(6): 57 - 69.
YU Ying, WANG Chunping, FU Qiang, et al. Survey of evaluation metrics and methods for semantic segmentation[J]. Computer Engineering and Applications, 2023, 59(6): 57 - 69. (in Chinese)
[31] LOSHCHILOV I, HUTTER F. Decoupled weight decay regularization[J]. arXiv:1711.05101, 2017.
[32] WANG S, DAI Y, SHEN J, et al. Research on expansion and classification of imbalanced data based on SMOTE algorithm [J]. Scientific Reports, 2021, 11(1): 24039.
[33] LIN H, CHENG X, WU X, et al. CAT: cross attention in vision transformer[J]. arXiv:2106.05786, 2021.
[34] CUI S, DENG H. PMG - DETR: fast convergence of DETR with position-sensitive multi-scale attention and grouped queries [J]. Pattern Analysis and Applications: PAA, 2024(2):27.
[35] ZHENG F, CAI T, WANG Y, et al. A mask-pooling model with local-level triplet loss for person re-identification[J]. IEEE Access, 2020, 8: 138191 - 138202.