doi:10.6041/j.issn.1000-1298.2023.03.032

基于残差网络和特征融合的小麦图像修复模型

陶兆胜　官保国　李庆萍　赵　瑞　伍　毅　吴　浩
(安徽工业大学机械工程学院, 马鞍山 243032)

摘要: 针对基于生成对抗网络的多数图像修复算法所修复的图像纹理细节不清晰,不能充分融合神经网络提取的纹理细节信息和语义信息的问题,本文提出一种基于残差网络和特征融合的双阶段生成网络图像修复模型,通过修复训练集中被遮挡的图像,获取符合训练集整体分布的修复图像。首先,设计一种轻量型多尺度感受野残差模块,通过多个感受野不同的卷积核提取特征信息,提升粗化生成网络保留纹理信息的能力。其次构建一种双边精细修复网络结构,分别处理纹理细节信息和语义信息并进行聚合,实现图像的精细修复。最后基于 GWHD 数据集进行实验,验证本文算法的有效性。实验结果表明,本文模型较 CE、GL、PEN－Net、CA 算法,客观评价指标 L1－loss 降低 0.56～3.79 个百分点,PSNR 和 SSIM 提升 0.2～1.8 dB 和 0.02～0.08,并在人眼直观感受中实现了纹理结构清晰、语义特征合理的修复效果。相较于原 GWHD 数据集,在基于本文模型所扩充的小麦数据集中,运用 YOLO v5s 预测小麦麦穗的 mAP 提升 1.41 个百分点,准确率提升 3.65 个百分点,召回率提升 0.36 个百分点。

关键词: 小麦; 图像修复; 生成对抗网络; 残差网络; 特征融合
中图分类号: TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2023)03-0318-10　　**OSID**:

Wheat Image Inpainting Based on Residual Networks and Feature Fusion

TAO Zhaosheng　GONG Baoguo　LI Qingping　ZHAO Rui　WU Yi　WU Hao
(*College of Mechnical Engineering*, *Anhui University of Technology*, *Maanshan* 243032, *China*)

Abstract: Aiming at the problem that most image inpainting algorithms based on generative adversarial networks restore unclear image texture details and cannot fully integrate texture detail information and semantic information extracted by neural networks, a two-stage algorithm was proposed based on residual network and feature fusion. A network image inpainting model was generated, and inpainted images were obtained which conformed to the overall distribution of the training set by inpainting the occluded images in the training set. Firstly, a lightweight multi-scale receptive field residual module was designed, which extracted feature information through multiple convolution kernels with different receptive fields, and improved the ability of the coarsening generation network to retain texture information. Secondly, a bilateral fine inpainting network structure was constructed, which processed texture detail information and semantic information separately and aggregated them to realize fine inpainting of images. Finally, experiments were carried out on the GWHD dataset to verify the effectiveness of the algorithm. The experimental results showed that compared with the CE, GL, PEN－Net, and CA algorithms, the objective evaluation index L1－loss of this model was reduced by 0.56～3.79 percentage points, PSNR and SSIM were improved by 0.2～1.8 dB and 0.02～0.08, and the restoration effect with clear texture structure and reasonable semantic features was realized in the intuitive perception of human eyes. Compared with the original GWHD data set, in the wheat data set expanded based on the proposed algorithm herein, the average detection accuracy mAP of predicting wheat ears using YOLO v5s was increased by 1.41 percentage points, the accuracy rate was increased by 3.65 percentage points, and the recall rate was increased by 0.36 percentage points.

Key words: wheat; image inpainting; generative adversarial networks; residual network; feature fusion

收稿日期: 2022－04－22　修回日期: 2022－06－01
基金项目: 安徽省自然科学基金面上项目(2108085ME166)和安徽高校自然科学研究项目重点项目(KJ2021A0408)
作者简介: 陶兆胜(1970—),男,副教授,博士,主要从事机器视觉和机电一体化研究,E-mail: taozs@ahut.edu.cn

0　引言

小麦作为人类广泛种植的一种农作物，其生长情况一直是农业研究的重点领域之一，对此众多学者常对比不同时期的小麦图像以分析其生长情况。但由于保存措施不当和拍摄角度等问题，小麦图像常出现破损或遮蔽现象，进而影响小麦生长状况的分析准确性。为此如何修复破损或被遮蔽的小麦生长图像也成为众多学者研究的重点之一。

随着计算机视觉和深度学习的发展，许多学者将其应用于视觉检测领域特别是农产品的视觉预测领域[1]。视觉检测技术虽然提高了检测效率，但也存在图像数据量少、植物重叠遮蔽麦穗和局部破损等问题，从而降低了深度学习的训练模型精度，进而导致误检和漏检等缺陷，影响小麦生长情况的分析准确性。

图像修复算法运用图像已知区域对其破损区域进行合成、替换，生成在纹理细节和语义两方面都符合人类感知的替代图像。该算法分为两类：传统图像修复算法和基于深度学习的图像修复算法。传统图像修复算法有基于偏微分方程的算法和纹理合成算法。BERTALMIO 等[2]提出基于偏微分的 BSCB (Bertalmio, Sapiro, Caselles, Ballester) 图像修复算法，该算法以局部像素点为修复单位，沿等值线扩散局部颜色的光滑度，以保证边缘处的边界连续性，但该算法复杂度较高，易受噪声干扰，从而导致修复效果不稳定，且耗时较长。CRIMINISI 等[3]提出的纹理合成算法是以样本块为基本修复单位，在前景区域进行样本相似信息匹配，通过复制样本块对破损区域进行修复，但由于其需要对比整个图像的样本块，故修复效率较低。文献[4]针对单一修复模板易出现填充裂纹和错误像素的问题，提出基于边缘特征和像素结构相似度的图像修复算法。黄颖等[5]针对基于样本块的图像修复算法在图像修复过程中难以保持纹理结构连贯性的问题，提出了结合等照度线的曲率特征和高斯函数的图像修复改进算法。卢伟等[6]针对在根系表型检测中，因土壤遮挡而导致根系图像断裂的问题，提出一种基于腐蚀生长算法的玉米根系修复算法，检测作物根系表型。

基于深度学习的图像修复算法[7-8]通过学习未破损区域的图像分布，以使破损区域的修复效果尽可能地接近真实场景。PATHAK 等[9]提出一种基于编码-解码器结构和生成对抗网络的语义编解码算法(Context encoders, CE)，通过上下文信息修复破损图像。何凯等[10]提出一种级联残差模块，通过卷积层之间交叉级联增强特征复用。LIZUKA 等[11]将空洞卷积引入编码-解码器中，并设置全局和局部 2 个判别器以实现全局和局部的相关性，通过填充任意形状的破损区域，修复任意分辨率的破损图像。YU 等[12]在双阶段生成网络的解码器中，构建上下文注意力网络以捕获与破损区域相关性最高的区域，取得了较好的修复效果。杨昊等[13]构建了分层残差模块，并在跳跃连接中嵌入多尺度通道注意力模块，提高了模型捕获低级特征的能力。姜海燕等[14]为了实现对局部被叶片遮挡的小尺寸稻穗的计数，设计一种基于生成对抗网络的稻穗检测算法。

尽管上述深度学习模型取得了较好的修复效果，但都没有将图像的纹理细节信息和语义信息进行充分融合以获得符合人类感知的图像。另外由于农作物图像往往包含野草、土地、庄稼等边缘纹理细节较丰富的目标，对运用图像修复算法修复破损或被遮蔽的小麦图像造成了一定的困难。为此本文从小麦图像的特点出发，首先以残差网络为生成器主体，构建粗化生成网络获取粗略修复结果，然后融合双边修复网络提取纹理细节特征和语义特征，生成精细修复结果，最后在小麦图像修复任务中，获得边缘纹理清晰、与周边区域连贯性强的修复图像。

1　图像修复原理

1.1　生成对抗网络

生成对抗网络是一种无监督深度学习模型，其通过生成器 G (Generator) 和判别器 D (Discriminator) 之间的互相博弈，生成与真实样本相近的图像。目前，大多数基于深度学习的图像修复算法是以 GAN 为基础网络架构，如图 1 所示。

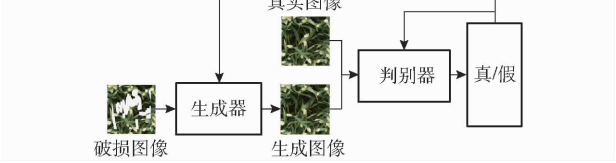


图 1　生成对抗网络示意图

Fig. 1　Schematic of generative adversarial network

生成对抗网络首先将破损图像输入生成器，提取修复图像后和真实图像输入判别器，然后将判别结果反馈给生成器，通过训练不断提高生成器修复图像破损区域的能力。

生成对抗网络通过减小交叉熵代价函数 $V(D,G)$ 交替训练生成器和判别器，不断降低破损图像和真实图像之间的分布差异，以便生成器输出结构纹理清晰的修复图像。

$$\min_G \max_D V(D,G) = E_{x\sim P_{\text{mask}}(x)}(\text{lb}(1-D(G(m)))) + E_{x\sim P_{\text{data}}(x)}(\text{lb}D(x)) \quad (1)$$

式中 $E(\cdot)$——分布函数期望

$P_{data}(x)$——真实样本分布

$P_{mask}(x)$——待修复图像分布

m——待修复图像

1.2 特征融合方法

在图像修复中，网络浅层提取的低级信息确定破损区域的纹理细节，而深层次的网络层提取的语义信息则决定图像修复效果的整体合理性，但网络层数过深将导致部分低级纹理细节信息丢失。因此，如何通过特征融合在深层神经网络中保留低级纹理细节信息和高级语义信息，对于提高图像破损区域的修复效果至关重要。

为了解决上述问题，郭启帆等[15]提出了基于特征金字塔的多尺度特征融合网络模型，设计了混合特征金字塔模块，将浅层的细节信息输入深层网络。YU 等[16]提出一种双边分割网络（BiSeNet V2），一方面通过层数较浅的纹理细节分支减少下采样次数，以获得丰富的细节特征；另一方面则通过卷积层数较多的语义分支提取图像的语义信息，从而更有效地聚合了纹理细节特征与语义特征。

1.3 注意力机制

注意力机制学习输入特征的空间维度或通道维度之间的关联程度，并根据不同的权重衡量不同特征的重要性，以便神经网络学习破损区域的关键信息。在图像修复任务中，LIU 等[17]使用连贯语义注意力机制学习破损区域像素之间的相关性，提升图像纹理质量。WANG 等[18]提出了一种轻量型通道注意力模块（Efficient channel attention，ECA），通过局部跨通道交互信息获取有效的通道维度权重。

2 双阶段生成网络

本文设计的双阶段生成网络，包括残差粗化修复网络、双边精细修复网络、判别器（图 2）。残差粗化修复网络构建 U 型结构以较快的速度和少量参数获取初步修复结果，引入多尺度感受野残差模块解决模型退化问题，然后通过跨通道连接（add）对特征进行融合。双边精细修复网络采用空间和通道组件之间的特征聚合以解决纹理细节信息缺失和感受野过小的缺陷。判别器区分真实图像和生成器 G 修复后的图像，并应用对抗损失函数将结果反馈给生成器，以监督生成器的训练，确保获取语义和纹理细节都较为合理的修复图像。本文判别器借鉴马尔可夫判别器[19]，将输入的修复结果映射成通道数为 1 的特征图，特征图的每个像素代表原图的一块区域，判别器输出特征图的均值。

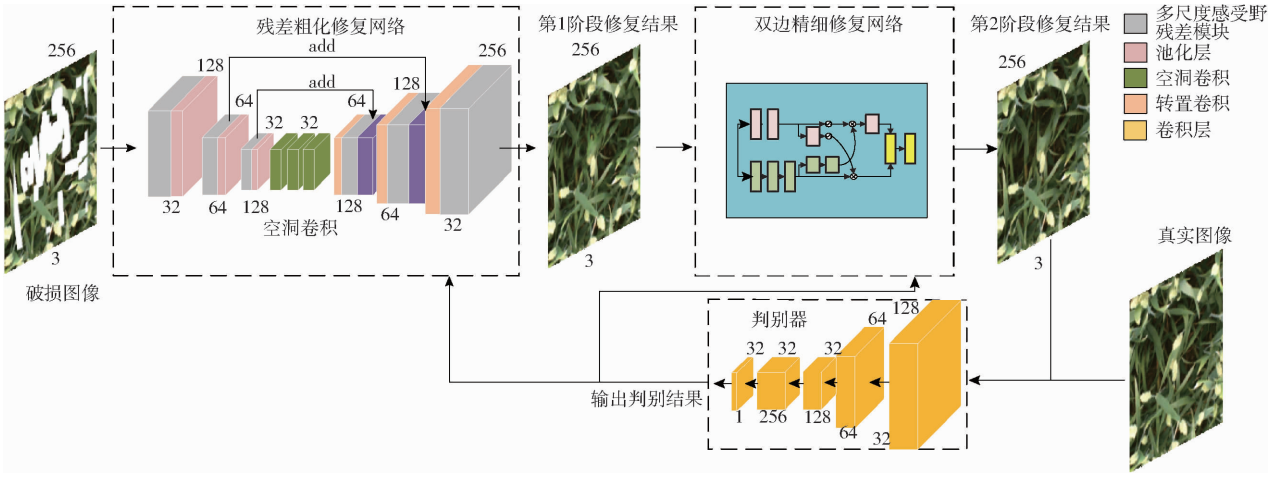


图 2 本文生成网络结构图

Fig. 2 Network structure of proposed completion

2.1 残差粗化修复网络

粗化生成网络是由 6 个多尺度感受野残差模块组成的编解码器结构。在每个残差模块后，首先运用一个步长为 2 的最大池化层进行下采样，以降低图像分辨率；接着编码-解码器的中间层采用一组 3×3 卷积核、空洞率为[1,2,5]的空洞卷积以增强感受野；最后通过卷积核为 3×3，步长为 2 的转置卷积进行上采样，并应用跳跃连接融合编码器所提取的特征信息与上采样所采样的特征信息，实现特征信息的跨通道流动。

在图像修复中，深层网络层虽然通过提取待修复图像的局部特征，较好地修复破损区域面积较大的图像，但也存在网络难以收敛的问题。而残差结构能较好地抑制网络层数过深所导致的梯度消失，且增大感受野以提高网络捕获破损区域与其距离较远位置之间特征相关性的能力，获取全局特征。为了抑制残差模块运算时的图像信息丢失和模型梯度混淆问题，ZHOU 等[20]提出了沙漏型残差模块

SandGlass，利用 2 个 1×1 卷积外层的深度卷积，提升网络宽度。TAN 等[21]通过卷积核为 3、5、7 的深度可分离卷积分别提取不同感受野的特征，在不增加存储代价的情况下，提高网络特征提取能力。

受此启发，本文设计的粗化生成网络构建了具有短跳跃连接的多尺度感受野残差模块（图 3），使用多个 3×3 深度卷积和 1×1 卷积的级联替代文献[21]的大卷积核。在保持感受野不变的前提下，减少参数量和计算量，并将第 2 条支路和第 3 条支路设计为沙漏型结构，避免模型梯度混淆。

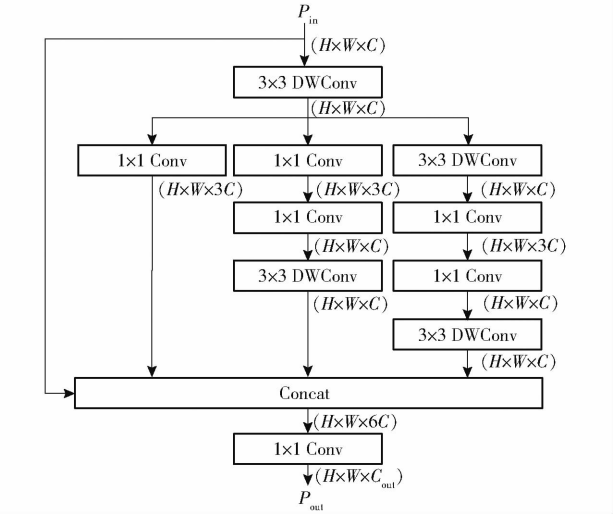


图 3　本文残差模块

Fig. 3　Proposed residual block

首先将特征图 P_{in} 输入至步长为 1，卷积核大小为 3×3 的深度卷积网络层以获取图像特征信息；接着将图像特征信息分别输送至 3 条支路。第 1 条支路运用 1×1 卷积将通道数升维至原通道数的 3 倍，感受野相较于 P_{in} 为 3。在第 2 条支路中，特征经 1×1 卷积升维，通过级联的 1×1 和 3×3 深度卷积提取特征信息，感受野相较于 P_{in} 为 5。第 3 条支路增加 1×1 和 3×3 的 2 个深度卷积级联结构，进一步提高感受野，感受野相较于 P_{in} 为 7。由于卷积层感受野决定网络层所提取的特征（包含全局信息）的完整程度，为此本文构建 3 种不同尺度的感受野以提取不同尺度的图像特征信息，并进行通道维度拼接（Concat），有利于增强破损区域和前景区域之间的语义相关性和纹理细节完整性。为了提取不同通道维度级联所获得的多样化特征信息，本文运用 1×1 卷积压缩特征的通道维度，输出特征图 P_{out}

$$P_{out} = C_{conv1}(C_{concat}(G_j(C_{conv3}(P_{in})) + P_{in})) \quad (2)$$

式中　$C_{conv3}(P_{in})$——将 P_{in} 送入 3×3 卷积层操作

G_j——3 种感受野尺度支路计算，$j = 1, 2, 3$

C_{concat}——通道数合并操作

C_{conv1}——1×1 卷积

2.2　双边精细修复网络

为了更好地融合低级纹理特征和高级语义特征，本文提出一种由细节分支和语义分支组成的双边精细修复网络（图 4）。

细节分支通过提高通道维度，增大细节信息依赖性，获取纹理特征。首先为了提取丰富的纹理细节信息，在不改变特征图的空间维度的基础上，运用 1×1 和 3×3 卷积层增加特征图通道数，通过 ReLU 激活函数增强模型的非线性，并分支输出；接着为了与语义分支结果进行融合，采用步长为 2 的 3×3 卷积和池化，将特征图调整为输入特征图的 1/4，并将其与经 Sigmoid 激活函数归一化的语义分支的输出特征图通过矩阵对应位置相乘进行聚合。

语义分支首先将特征图输入 Stemblock[22] 模块中，运用步长为 2 的卷积层提取特征，同时采用最大值池化和步长为 2 的卷积层进行下采样，减小特征的空间尺度，并通过 Concat 融合 2 种不同的下采样方式所获取的语义信息，以提升模型的特征表达能力；接着将特征图输入通道注意力模块 ECA 调整通

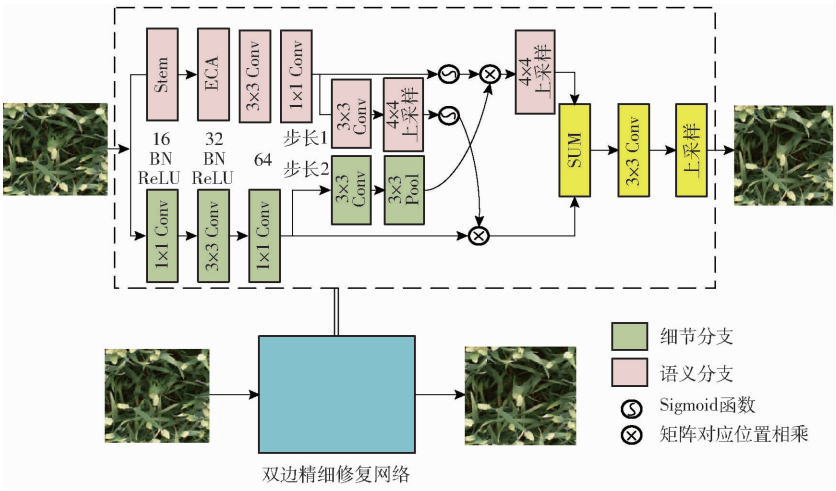


图 4　双边精细修复网络结构图

Fig. 4　Bilateral fine repair network

道权重,以提取通道维度信息;之后利用 3×3 和 3×3 卷积扩充特征图的通道维度,并分支;最终在空间维度上采样为原特征图的 4 倍,通过 Sigmoid 函数归一化,与细节分支的输出特征图通过矩阵对应位置相乘进行特征聚合。

细节分支和语义分支所提取的特征经过通道维度对应位置相加进行融合(SUM),之后输入 3×3 的卷积层以调整维度,并上采样为输入原图尺寸。双边精细修复网络通过双向信息聚合方式,将具有宽通道和浅网络层的细节分支得到的空间信息与具有窄通道和宽网络层的语义分支提取的高级语义进行融合,以提高网络的特征表达能力。

2.3 损失函数

为了确保修复后的图像纹理细节清晰和修复效果的语义合理性,本文使用一种结合像素重建损失、感知损失、对抗损失和总变分(Total variation,TV)损失的损失函数 L,计算公式为

$$L = \lambda_{pix} L_{pix} + \lambda_{per} L_{per} + \lambda_{adv} L_{adv} + \lambda_{TV} L_{TV} \tag{3}$$

式中 L_{pix}、λ_{pix}——像素重建损失与其权重

L_{per}、λ_{per}——感知损失与其权重

L_{adv}、λ_{adv}——对抗损失与其权重

L_{TV}、λ_{TV}——TV 损失与其权重

像素重建损失 L_{pix}通过计算真实图像 I^{gt}与生成器输出的修复图像 I^{out}的 L_1 范数以衡量二者之间的差异。L_{pix}计算公式为

$$L_{pix} = \| I^{gt} - I^{out} \|_1 \tag{4}$$

感知损失 L_{per}将 I^{gt}和 I^{out}输入卷积神经网络中,计算两者第 j 层网络中特征图在高维空间的距离,使得 I^{gt}和 I^{out}在高维的语义特征和纹理特征上具有相似性。L_{per}的计算公式为

$$L_{per} = \frac{1}{C_j H_j W_j} \| \phi_j(I^{gt}) - \phi_j(I^{out}) \|_2 \tag{5}$$

式中 $\phi_j(I)$——I 经过 VGG-19 网络提取的第 j 层特征图

C_j、H_j、W_j——第 j 层特征图的通道数、高度、宽度

对抗损失 L_{adv}监督生成对抗网络以保证 G 生成欺骗 D 的图像修复结果。L_{adv}的计算公式为

$$L_{adv} = \mathrm{lb}(D(I^{gt})) + \mathrm{lb}(1 - D(G(I^{out}))) \tag{6}$$

L_{TV}抑制总变分以消减噪声,减小相邻像素间的差别,实现修复结果的平滑性。L_{TV}的计算公式为

$$L_{TV} = \frac{1}{CHW} \sum_{i,l,k} \sqrt{(I^{out}_{i,l+1,k} - I^{out}_{i,l,k})^2 + (I^{out}_{i+1,l,k} - I^{out}_{i,l,k})^2} \tag{7}$$

式中 i、l、k——I^{out}的空间维度和通道维度索引

C、H、W——图像通道数、高度和宽度

3 实验验证

为了验证所提出的图像修复算法性能,本文采用中心矩形掩膜和随机生成不规则形状掩膜作为图像的破坏形式,并通过消融实验、定性评价、定量评价以及麦穗检测数据增强实验客观且全面地验证模型性能。

3.1 实验数据集

分析小麦的生产情况对于提高小麦产量、减少麦田病害有着重要意义。此外利用视觉技术获取小麦检测图像时,常出现目标被遮挡、图像破损、数据集样本数量较少等问题。因此,通过图像修复技术提高数据集的质量,在防止小麦检测的过拟合和优化泛化性能方面有着至关重要的作用。本文另外使用全球麦穗检测数据集(Global wheat head detection, GWHD)[23]验证本文算法对不同场景下小麦图像信息缺失的修复能力。GWHD 数据集共有 4 700 幅小麦头 RGB 图像,其中 4 000 幅作为训练集,700 幅作为测试集。

3.2 实验参数设置

本文所做的验证实验均基于 Pytorch 1.7.1 框架,计算平台使用的 CPU 为 Intel Core i7-11700K,GPU 为 NVIDIA RTX 3070,软件环境为 Python 3.7。

首先预处理训练集的图像,即运用立方卷积插值调整图像尺度为 256 像素×256 像素,针对小麦植株图像中个体出现位置的不确定性,本文使用中心矩形掩膜和随机生成不规则形状掩膜的方式模拟实际遮挡对模型进行训练和测试。算法的训练迭代次数(epoch)设置为 100 个,批处理样本数量(batchsize)设置为 2,初始学习率为 0.002,生成器采用 Adam 优化算法,判别器运用 SGD 优化算法。各损失函数权重分别设置为 $\lambda_{pix}=0.98$,$\lambda_{per}=0.01$,$\lambda_{adv}=0.009$,$\lambda_{TV}=0.001$,权重设置参照文献[11]。

3.3 实验结果及分析

3.3.1 模块有效性分析

为了验证残差模块和双边精细修复网络的有效性,本文基于 GWHD 数据集运用峰值信噪比(Peak signal to noise ratio,PSNR)进行消融实验。PSNR 数值越大,修复效果越好。

$$R_{PSNR} = 10\lg\left(\frac{255^2}{L_{MSE}}\right) \tag{8}$$

式中 L_{MSE}——真实图像与修复图像的均方误差

R_{PSNR}——真实图像和待修复图像的峰值信噪比

本文算法使用 CE 作为参照模型,CE+RB 为 CE 引入多尺度感受野残差块的修复算法,CE+双

边网络为 CE 融合双边精细修复网络的修复算法，CE + RB + 双边网络为 CE、多尺度感受野残差块和双边精细修复网络相结合的修复算法。BL 为本文所构建的修复算法的基础模型，BL + RB 为 BL 融入多尺度感受野残差块的修复算法，BL + 双边网络为 BL 融合双边精细修复网络的修复算法，本文模型（BL + RB + 双边网络）为 BL、多尺度感受野残差块和双边精细修复网络相结合的修复算法。对比实验的训练迭代次数设置为 100，其余实验参数与 3.2 节的参数相同。

图 5 为上述所有算法消融实验的 PSNR 对比结果。在训练迭代次数 100 内，相对于 CE，CE + RB、CE + 双边网络、CE + RB + 双边网络的 PSNR 平均值分别提升 0.22、0.26、0.34 dB；BL 的 PSNR 平均值相对于 CE 提高 1.4 dB；相较于 BL，BL + RB、BL + 双边网络、本文模型的 PSNR 平均值分别提高 0.14、0.16、0.24 dB。

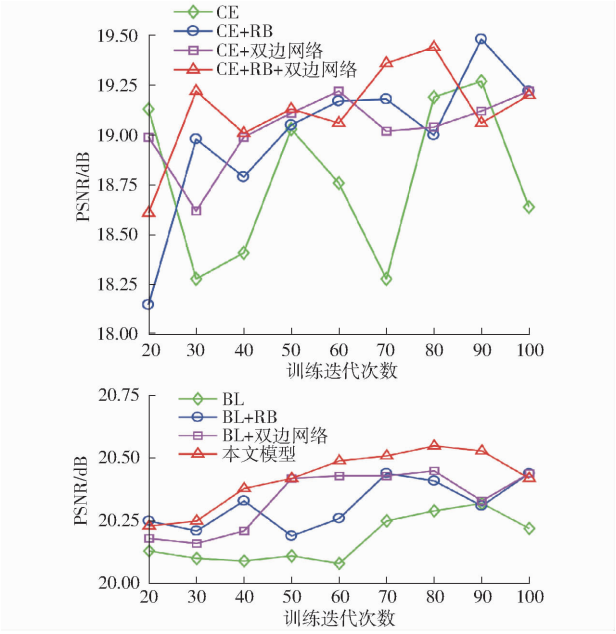


图 5　模型消融实验定量评价结果

Fig. 5　Quantitative evaluation results of model ablation experiment

修复效果对比如图 6 所示，图 6c、6g 细节修复较差，特别是在左下方麦穗区域，修复效果模糊。图 6d、6h 与图 6c、6g 对比可知，在上部麦穗区域，麦穗纹理细节较清晰，表明残差模块能通过多尺度感受野有效捕获低级纹理细节特征。从图 6e、6i 与图 6c、6g 相比可知，遮蔽区被修复的麦穗与叶片之间边界更为明显，自然地融入了周围环境，表明双边网络提取高级特征的能力强，能有效地捕捉图像的上下文信息，提高修复图像的合理性。由图 6f、6j 可知，相较于其它对比模型，本文采用的残差模块和双边网络能够有效地修复麦穗结构，特别是遮蔽区被修复的麦穗形状纹理细节结构较为真实，且虚化程度明显小于其它修复结果，感官上更符合人类认知。

3.3.2　定性实验

为了向小麦视觉检测算法提供更多的训练图像，本文以随机生成不规则形状掩膜和指定位置矩形掩膜遮挡小麦图像，修复结果如图 7 ~ 10 所示。图 7 较好地实现了麦穗的纹理结构修复，边缘纹理信息较完整。图 8 部分修复了叶片的被遮蔽纹理结构，但未能实现叶片的连接。图 9 虽然整体修复了破损的小麦麦穗，但红框内的小麦纹理结构略有缺失。图 10 较为完整地修复小麦的纹理结构。

为了对本文提出的修复算法进行定性评估，本文模型与 4 种深度学习图像修复算法分别在 GWHD 数据集进行对比实验，实验结果如图 11 所示。本文选取了生长状态不同、麦穗密集程度不同、掩膜尺寸不同的小麦图像作为测试图像。由图 11c 可知，图像修复结果出现纹理特征失真以及第 4 行图像麦穗发生结构纹理模糊现象。上述缺陷表明，当数据集数量较少时，CE 方法很难通过学习方式获

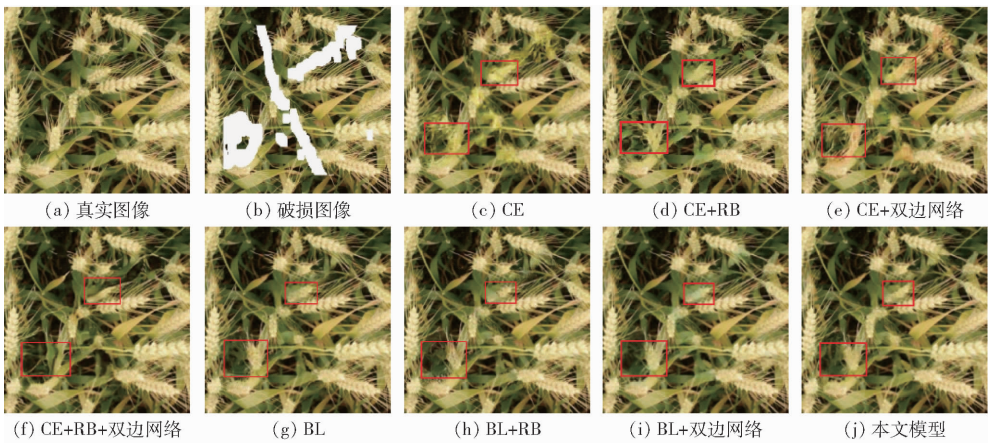


图 6　模型消融实验修复结果对比

Fig. 6　Comparison of inpainting results of model ablation experiment


(a) 破损图像 (b) 修复图像 (c) 真实图像

图 7 部分麦穗头区域修复结果

Fig. 7 Results of part wheat area inpainting


(a) 破损图像 (b) 修复图像 (c) 真实图像

图 8 叶片区域修复结果

Fig. 8 Results of leaf area inpainting


(a) 破损图像 (b) 修复图像 (c) 真实图像

图 9 不规则遮挡区域 1 修复结果

Fig. 9 Results of irregular area 1 inpainting

取合理的纹理特征信息,修复效果较差。由图 11d 可知,运用 GL(Globally and locally)算法修复的图像麦穗局部失真,即破损区域发生结构纹理扭曲杂乱、麦穗边缘纹理缺失和麦叶断裂等现象。从图 11e 可知,在 PEN - Net 算法[24]修复的图像中,存在结构纹理扭曲、被遮蔽区域纹理杂乱、麦秸杆部区域缺失等


(a) 破损图像 (b) 修复图像 (c) 真实图像

图 10 不规则遮挡区域 2 修复结果

Fig. 10 Results of irregular area 2 inpainting

现象。由图 11f 可知,CA(Contextual attention)算法图像修复出现颜色异常、第 1 行和第 4 行的图像麦穗发生结构纹理修复模糊等缺陷。本文算法实现了第 1 行图像和第 3 行图像的小麦边缘结构纹理清晰和麦叶连接。虽然在第 2 行图像中,较好地修复麦穗结构纹理,但上部红框的麦穗未能实现完整修复。第 4 行图像较好地修复红框内的小麦麦穗的纹理结构,但未能完整地修复被遮蔽区域的麦穗及麦叶。通过定性实验结果对比,证明本文算法在纹理细节修复和语义合理性上的表现都优于其它模型。本文模型通过多尺度感受野残差模块和双边精细修复网络较为完整地修复了红色矩形框内的叶片,图像的整体纹理细节自然真实。

3.3.3 定量实验

由于通过人眼直观地定性实验分析并不能全面客观地反映算法修复效果,为此本文采用 L1 范数损失(L1 - loss)、R_{PSNR}、结构相似度[25](Structural similarity,SSIM)对修复效果进行定量实验。

L1 - loss 估算被修复的破损图像和真实图像的像素点之间的差异,其值越小,被修复的破损图像和

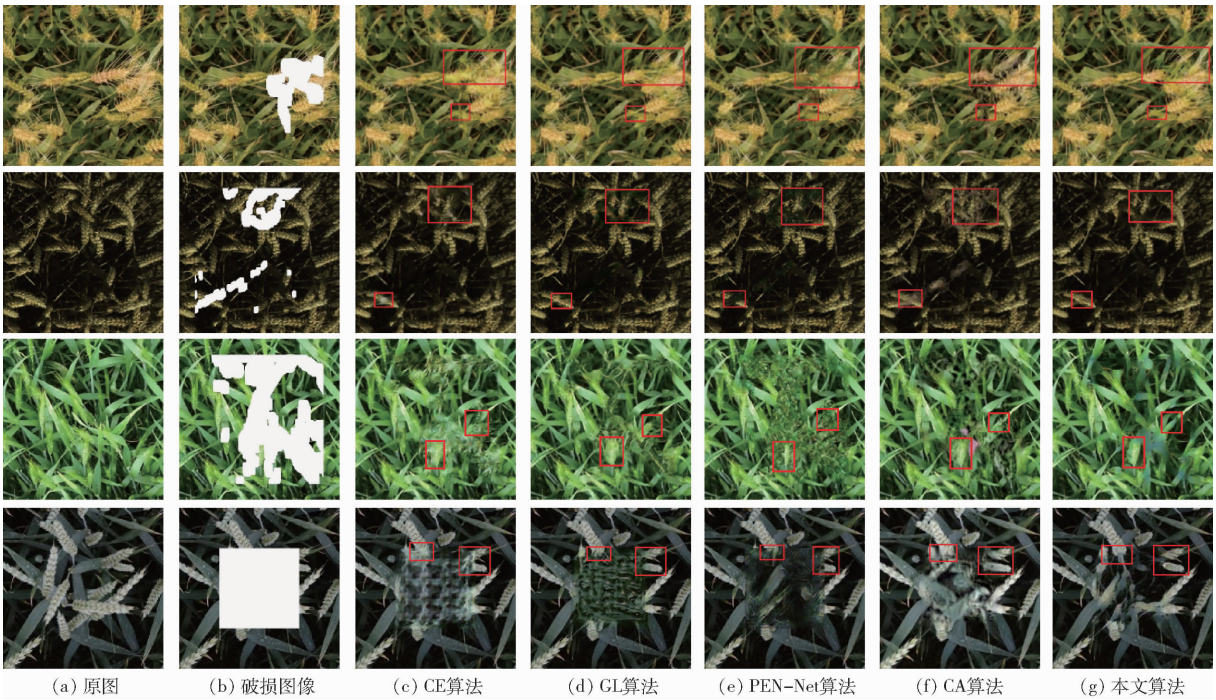

(a) 原图 (b) 破损图像 (c) CE算法 (d) GL算法 (e) PEN-Net算法 (f) CA算法 (g) 本文算法

图 11 GWHD 修复结果对比

Fig. 11 Comparison of inpainting results of GWHD

真实图像的差别越小。为了解决人眼对一个区域的感知易受周围环境影响的问题，本文引入 SSIM，从亮度、对比度、结构三方面度量图像相似性，其数值越高，修复效果越佳。

为了精确地评价本文算法的破损图像修复效果，分别比较了 CE 算法、GL 算法、PEN－Net 算法、CA 算法和本文模型的 L1－loss、PSNR 与 SSIM 3 个指标，结果如表 1 所示。

表 1　不同算法在 GWHD 数据集上的表现

Tab. 1　Performance of different algorithms on GWHD

算法	L1－loss/%	PSNR/dB	SSIM
CE	9.06	19.27	0.68
GL	7.87	19.87	0.72
PEN－Net	9.72	18.60	0.66
CA	11.10	20.20	0.67
本文	7.31	20.40	0.74

本文算法在客观评价指标 L1－loss，相较于 CE 算法降低 1.75 个百分点、相较于 GL 算法降低 0.56 个百分点、相较于 PEN－Net 算法降低 2.41 个百分点、相较于 CA 算法降低 3.79 个百分点。在客观评价指标 PSNR，相对于 CE 算法提高 1.13 dB、相对于 GL 算法提高 0.53 dB、相对于 PEN－Net 算法提高 1.8 dB、相对于 CA 算法提高 0.20 dB，在客观评价指标 SSIM，相较于 CE 算法提高 0.06、相较于 GL 算法提高 0.02、相较于 PEN－Net 算法提高 0.08、相较于 CA 算法提高 0.07。

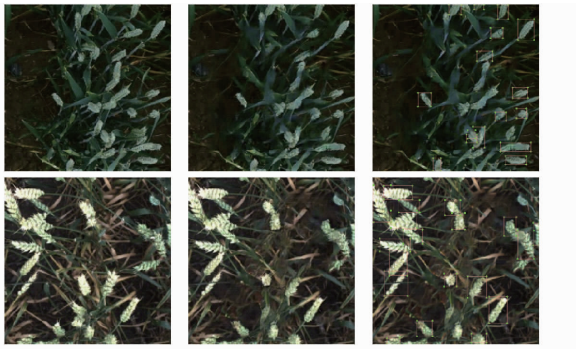
从表 1 可知，本文算法无论是在 L1－loss、PSNR，还是在 SSIM 均优于其他算法，表明本文算法捕捉纹理细节信息和提取语义信息的能力较强。

3.4　麦穗检测数据增强验证

为了验证本文算法是否能满足目标检测模型的数据增强任务要求，目标检测模型选择应用范围广泛的 YOLO v5s，训练平台硬件参数与 3.2 节相同，训练迭代次数设置为 300，批处理样本数量设置为 4，初始学习率设为 0.01。将 GWHD 中的 400 幅图像按照 2∶1∶1 划分为训练集、验证集、测试集。

从上述训练集随机选取 100 幅图像对其遮挡，接着运用本文修复算法对其进行修复，最后应用 LabelImg 标注工具重新标注，添加至原训练集，旨在扩充数据集中图像多样性（图 12）。相较于 GWHD 数据集真实图像，虽然修复图像的麦穗数较原图像减少，但将其添加于 GWHD 数据集可增强数据集的多样性和差异性。扩充后的数据集包含 500 幅小麦图像，其中训练集、验证集、测试集的占比为 3∶1∶1。

为了验证扩充的数据集能够有效提升目标检测模型精度，本文在相同验证集上使用目标检测领域



(a) 真实图像　(b) 生成图像　(c) LabelImg标注

图 12　麦穗图像数据增强

Fig. 12　Wheat image data augmentation

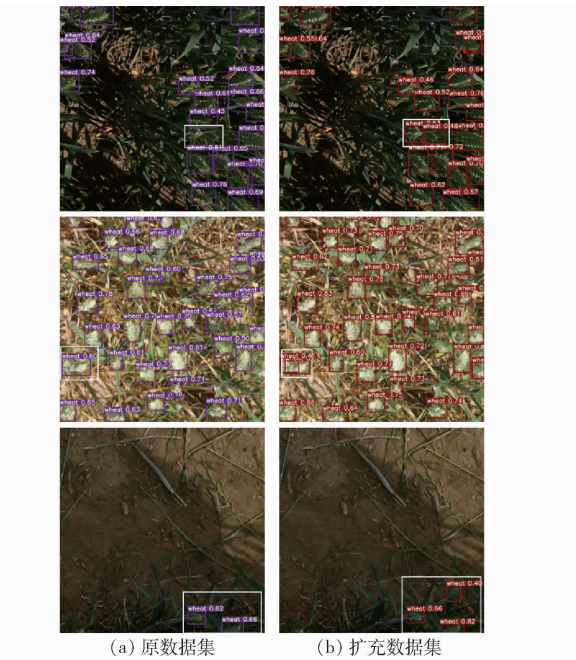
常用的 mAP、准确率和召回率作为评价指标，对比在 GWHD 数据集和本文扩充的 GWHD 数据集的测试结果，如表 2 所示。从表 2 可知，相对于未扩充前的数据集，YOLO v5s 的 mAP 提高 1.41 个百分点，准确率提升 3.65 个百分点，召回率提升 0.36 个百分点，检测精度明显提升。置信度阈值为 0.4 时的检测效果如图 13 所示，在小麦植株分布密集的上侧、中部和右下侧区域（图中白色框所示），图 13a 相较于图 13b 出现漏检现象，说明在数据集样本数量较少时，通过本文修复算法对数据集进行扩充能

表 2　目标检测评价指标测试结果

Tab. 2　Results of object detection evaluation index

%

数据集	mAP	准确率	召回率
GWHD	77.06	39.34	83.59
扩充的 GWHD	78.47	42.99	83.95



(a) 原数据集　(b) 扩充数据集

图 13　原数据集与扩充数据集 YOLO v5s 检测效果对比

Fig. 13　Comparison of YOLO v5s detection effects between original dataset and expanded dataset

有效地提高目标检测模型精度。

4 结束语

提出了一种基于生成对抗网络和残差结构的双生成阶段的图像修复算法,基于 GWHD 数据集进行了对比实验验证。结果表明,本文模型的 L1 - loss、峰值信噪比和 SSIM 分别为 7.31%、20.40 dB 和 0.74。此外,通过修复被人为施加掩膜遮挡的小麦破损图像的方式,以扩充仅有 400 幅小麦图像的小样本数据集,使得目标检测模型 YOLO v5s 的 mAP 提高 1.41 个百分点,准确率提升 3.65 个百分点,召回率提升 0.36 个百分点。故本文算法能够通过图像修复得到纹理结构清晰、语义特征合理的多样性小麦图像,因此可作为一种数据增强方法,为基于深度学习的农作物视觉检测模型提供训练样本数量增强的农作物数据集。

参 考 文 献

[1] 梁胤豪,陈全,董彩霞,等. 基于深度学习和无人机遥感技术的玉米雄穗检测研究[J]. 福建农业学报, 2020,35(4): 456 - 464.
LIANG Yinhao, CHEN Quan, DONG Caixia, et al. Application of deep-learning and UAV for field surveying corn tassel[J]. Fujian Journal of Agricultural Sciences, 2020,35(4): 456 - 464. (in Chinese)

[2] BERTALMIO M, SAPIRO G, CASELLES V, et al. Image inpainting[C]// Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM Press, 2000:417 - 424.

[3] CRIMINISI A, PEREZ P, TOYAMA K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200 - 1212.

[4] 陶兆胜, 张敬寒, 王磊,等. 基于边缘特征和像素结构相似度的图像修复算法[J]. 计算机辅助设计与图形学学报, 2019,31(10): 1768 - 1776.
TAO Zhaosheng, ZHANG Jinghan, WANG Lei, et al. Image inpainting algorithm based on edge feature and pixel structure similarity[J]. Journal of Computer-Aided Design & Computer Graphics, 2019,31(10): 1768 - 1776. (in Chinese)

[5] 黄颖, 李凯, 杨明. 基于样本块与曲率特征的图像修复改进算法[J]. 计算机应用研究,2018,35(4): 1274 - 1276.
HUANG Ying, LI Kai, YANG Ming. Improved algorithm for image inpainting based on sample block and curvature features [J]. Application Research of Computers,2018,35(4):1274 - 1276. (in Chinese)

[6] 卢伟,李也,王玲,等.基于腐蚀生长算法的不同活力玉米种子根系表型研究[J].农业机械学报,2020,51(4): 224 - 231.
LU Wei, LI Ye, WANG Ling, et al. Root phenotypic detection of different vigorous maize seeds based on corrosion growth algorithm of image[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(4):224 - 231. (in Chinese)

[7] 范春奇, 任坤, 孟丽莎,等. 基于深度学习的数字图像修复算法最新进展[J].信号处理,2020,36(1): 102 - 109.
FAN Chunqi, REN Kun, MENG Lisha, et al. Advances in digital image inpainting algorithms based on deep learning[J]. Journal of Signal Processing,2020,36(1):102 - 109. (in Chinese)

[8] 姜艺, 胥加洁, 柳絮,等. 边缘指导图像修复算法研究[J]. 计算机科学与探索,2022,16(3): 669 - 682.
JIANG Yi, XU Jiajie, LIU Xu, et al. Research on edge-guided image repair algorithm[J]. Journal of Frontiers of Computer Science and Technology,2022,16(3):669 - 682. (in Chinese)

[9] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context encoders: feature learning by inpainting[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2536 - 2544.

[10] 何凯,刘坤,李宸,等.级联式生成对抗网络图像修复模型[J].天津大学学报(自然科学与工程技术版), 2021, 54(9): 917 - 924.
HE Kai, LIU Kun, LI Chen, et al. Image inpainting model using cascaded generative adversarial network[J]. Journal of Tianjin University (Science and Technology), 2021, 54(9):917 - 924. (in Chinese)

[11] LIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics, 2017, 36(4): 1 - 14.

[12] YU J H, LIN Z, YANG J M, et al. Generative image inpainting with contextual attention[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 5505 - 5514.

[13] 杨昊,余映. 利用通道注意力与分层残差网络的图像修复[J]. 计算机辅助设计与图形学学报, 2021,33(5): 671 - 681.
YANG Hao, YU Ying. Image inpainting using channel attention and hierarchical residual networks[J]. Journal of Computer-Aided Design & Computer Graphics, 2021,33(5): 671 - 681. (in Chinese)

[14] 姜海燕,徐灿,陈尧,等. 基于田间图像的局部遮挡小尺寸稻穗检测和计数方法[J]. 农业机械学报, 2020,51(9): 152 - 162.
JIANG Haiyan, XU Can, CHEN Yao, et al. Detecting and counting method for small-sized and occluded rice panicles based on in-field images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(9): 152 - 162. (in Chinese)

[15] 郭启帆,刘磊,张珹,等. 基于特征金字塔的多尺度特征融合网络[J]. 工程数学学报,2020,37(5): 521 - 530.
GUO Qifan, LIU Lei, ZHANG Cheng, et al. Muti-scale feature fusion network based on feature pyramid model[J]. Chinese Journal of Engineering Mathematics, 2020,37(5): 521 - 530. (in Chinese)

[16] YU C, GAO C, WANG J, et al. Bisenet v2: bilateral network with guided aggregation for real-time semantic segmentation

[J]. International Journal of Computer Vision, 2021, 129(11): 3051 – 3068.
[17] LIU H, JIANG B, XIAO Y, et al. Coherent semantic attention for image inpainting[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 4170 – 4179.
[18] WANG Q, WU B, ZHU P, et al. ECA – Net: efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020:11531 – 11539.
[19] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 2242 – 2251.
[20] ZHOU D, HOU Q, CHEN Y, et al. Rethinking bottleneck structure for efficient mobile network design[C]// European Conference on Computer Vision. Springer, Cham, 2020: 680 – 697.
[21] TAN M, LE Q V. Mixconv: mixed depthwise convolutional kernels[C]// Proceedings of the British Machine Vision Conference (BMVC), 2019:1 – 13.
[22] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception – v4, inception-resnet and the impact of residual connections on learning[C]// Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence, 2017:4278 – 4284.
[23] DAVID E, SEROUART M, SMITH D, et al. Global wheat head detection 2021:an improved dataset for benchmarking wheat head detection methods[J]. Plant Phenomics, 2021, 2021: 277 – 285
[24] ZENG Y, FU J, CHAO H, et al. Learning pyramid-context encoder network for high-quality image inpainting[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 1486 – 1494.
[25] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 – 612.

(上接第 265 页)

[12] 王李娟,孔钰如,杨小冬,等. 基于特征优选随机森林算法的农耕区土地利用分类[J]. 农业工程学报,2020,36(4): 244 – 250.
WANG Lijuan, KONG Yuru, YANG Xiaodong, et al. Classification of land use in farming areas based on feature optimization random forest algorithm[J]. Transactions of the CSAE, 2020, 36(4): 244 – 250. (in Chinese)
[13] 阜宁县地方志办公室.阜宁年鉴[J].南京:江苏人民出版社,2021.
[14] 陈子琪,陈亮,李建松. 基于 InSAR 相干系数的河湖岸线水事活动监测[J]. 人民黄河,2020,42(7):160 – 163.
CHEN Ziqi, CHEN Liang, LI Jiansong. Monitoring water matter activity of river and lake shoreline based on coherent coefficient[J]. Yellow River, 2020, 42(7): 160 – 163. (in Chinese)
[15] BOUMAN B A M. Crop modelling and remote sensing for yield prediction[J]. Egyptian Journal of Biology, 1995, 43(2): 143 – 161.
[16] TOAN L T, RIBBES F, WANG L F, et al. Rice crop mapping and monitoring using ERS – 1 data based on experiment and modeling results[J]. IEEE Transactions on Geoscience and Remote Sensing, 1997, 35(1): 41 – 56.
[17] BOUARABA A, BELHADJ-AISSA A, BORGHYS D, et al. InSAR phase filtering via joint subspace projection method: application in change detection[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1817 – 1820.
[18] SUN Y, JIANG L, LIU L, et al. Mapping glacier elevations and their changes in the western Qilian Mountains, northern Tibetan Plateau, by bistatic InSAR[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 11(1): 68 – 78.
[19] DASH M, LIU H. Feature selection for classification[J]. Intelligent Data Analysis,1997, 1(1 – 4): 131 – 156.
[20] HUANG J, LI G, HUANG Q, et al. Joint feature selection and classification for multilabel learning[J]. IEEE Transactions on Cybernetics, 2018, 48(3): 876 – 889.
[21] KONONENKO I, ŠIMEC E, ROBNIK-ŠIKONJA M. Overcoming the myopia of inductive learning algorithms with RELIEFF [J]. Applied Intelligence, 1997, 7(1): 39 – 55.
[22] ZHOU Y, ZHANG R, WANG S, et al. Feature selection method based on high-resolution remote sensing images and the effect of sensitive features on classification accuracy[J]. Sensors(Basel), 2018, 18(7): 2013.
[23] 朱梦豪,李国清,彭壮壮. 特征优选下的农作物遥感分类研究[J]. 测绘科学,2022,47(3):122 – 128.
ZHU Menghao, LI Guoqing, PENG Zhuangzhuang. Remote sensing crop classification method based on feature selection[J]. Science of Surveying and Mapping, 2022, 47(3): 122 – 128. (in Chinese)
[24] 王长耀,刘正军,颜春燕. 成像光谱数据特征选择及小麦品种识别实验研究[J]. 遥感学报,2006,10(2):249 – 255.
WANG Changyao, LIU Zhengjun, YAN Chunyan. A experimental study on imaging spectrometer data feature selection and wheat type identification[J]. Journal of Remote Sensing, 2006, 10(2): 249 – 255. (in Chinese)
[25] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5 – 32.
[26] RODRIGUEZ-GALIANO V F, GHIMIRE B, ROGAN J, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2012, 67: 93 – 104.