

基于改进 LSTM 的蘑菇生长状态时空预测算法

杨淑珍¹ 黄杰¹ 苑进²

(1. 上海第二工业大学智能制造与控制工程学院, 上海 201209; 2. 山东农业大学机械与电子工程学院, 泰安 271018)

摘要: 密集蘑菇簇会严重影响蘑菇质量和自动采摘成功率。为避免形成超密集蘑菇簇, 提出一种蘑菇生长状态时空预测算法, 对蘑菇生长状态进行预测以指导提前疏蕾。该算法采用编码器-预测器框架, 将历史序列图像转换为 3D 张量序列作为模型的输入; 编码器网络中将卷积和长短时记忆 (Long short term memory, LSTM) 网络融合实现对蘑菇生长的时空相关性特征的提取; 在预测网络中加入扩散模型以解决预测图像的模糊问题; 此外, 在损失函数中增加了蘑菇面积差异损失函数来进一步减小预测蘑菇与实际蘑菇的形状和位置偏差。实验结果表明, 本文算法峰值信噪比可达 35.611 dB、多层次结构相似性为 0.927、蘑菇预测准确性高达 0.93, 有效提高了蘑菇生长状态图像预测质量和精度, 为食用菌生长预测提供了一种新思路。

关键词: 蘑菇; 生长状态预测; 长短时记忆网络; 扩散模型; 面积差异损失函数

中图分类号: TP183; S126 文献标识码: A 文章编号: 1000-1298(2024)03-0221-10

OSID:



Spatiotemporal Prediction Algorithm for Mushroom Growth Status Based on Improved LSTM

YANG Shuzhen¹ HUANG Jie¹ YUAN Jin²

(1. School of Intelligent Manufacturing and Control Engineering, Shanghai Polytechnic University, Shanghai 201209, China

2. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China)

Abstract: Dense mushroom clusters can significantly impact mushroom quality and the success rate of automated harvesting. To address this issue, a spatiotemporal prediction algorithm for mushroom growth status based on historical time series growth images was proposed, which can facilitate early bud thinning to prevent the formation of dense mushroom clusters. The algorithm employed a sequence-to-sequence structure, comprising an encoder and a predictor. In the input, historical image sequences were transformed into 3D tensor sequences and sent to encoder. Within the encoder network, a three-layer long short term memory (LSTM) model was utilized. Here, convolution was fused into LSTM cell to extract spatiotemporal correlation features of mushroom growth. Meanwhile, a diffusion model was introduced into the predictor to address the blurriness issue in predicting images. Furthermore, a mushroom area difference loss function was designed and incorporated into the loss function to further reduce the shape and positional deviations between the predicted and actual mushrooms. The experimental results indicated that the proposed spatiotemporal prediction algorithm for mushroom growth status achieved a peak signal-to-noise ratio of 35.611 dB, a multiscale structure similarity of 0.927, and a high mushroom mean intersection over union of 0.93, which represented improvements of 36%, 33% and 24%, respectively, over that of the ConvLSTM (Converlution LSTM) spatiotemporal prediction algorithm. This showed the proposed algorithm can effectively enhance the quality and accuracy of mushroom growth status image prediction, providing a approach for precise forecasting of edible mushroom growth.

Key words: mushroom; grow status prediction; LSTM; diffusion model; area difference loss function

收稿日期: 2023-11-07 修回日期: 2024-01-03

基金项目: 上海市科技计划项目 (21N21900600)、上海市科技兴农项目 (2019-02-08-00-10-F01123) 和山东省重点研发计划 (重大科技创新工程) 项目 (2022CXGC010609)

作者简介: 杨淑珍 (1978—), 女, 副教授, 博士, 主要从事采摘机器人和深度学习等研究, E-mail: szyang@sspu.edu.cn

0 引言

菌菇、猕猴桃、小番茄等易密集或成簇生长的果实,通常需要蔬果(蕾)避免果实过密导致的生长空间限制以及营养竞争而对果实的品质和产量造成影响^[1-2]。因此需要对蘑菇的密集程度、成熟期等生长状态特征进行预测,以指导提前疏蕾、避免产生超密集蘑菇簇,并对采摘任务进行合理安排,使蘑菇能适时被采摘、疏蕾,提高品质和产量。

作物生长一般通过其生长模型来进行预测^[3-4]。自 20 世纪 60 年代以来,已经建立了许多基于过程的作物模型并不断更新和改进^[5-8]。但是由于作物生长受到复杂的环境因素(如土壤成分和养分、温度、湿度等)、作物种植管理(如施肥策略、播种日期等)以及作物本身特性等诸多因素的非线性影响^[9-10],模型的预测结果往往不太准确^[11-12]。机器学习可较好地考虑输入变量和作物生长的复杂非线性^[13-14],是解决传统预测方法局限性的一种有效方法^[15],广泛用于作物生长的相关预测并取得了更高的精度^[16-19]。近年,深度学习开始被引入用于产量等农业预测并且效果优于机器学习^[20-22]。

食用菌生长预测技术发展类似。袁俊杰^[23]采用积差相关分析法研究了鸡腿菇子实体高度模型、直径模型和质量模型,并建立了基于温湿度的动态生长速率模型。鲍军鹏等^[24]将生长环境作为输入,采用模糊神经网络来拟合蘑菇的生长曲线,实现对木耳生长过程的预测。李毅志等^[25]采用非线性建模,建立了香菇菌盖直径动态模型和子实体鲜重动态模型,实现对香菇生长的预测。以上均采用建立生长模型的方法来对食用菌生长状态进行预测。将温度、湿度等主要生长环境因子作为模型的输入,研究环境因子对其生长的影响,然后建立相应子实体生长模型。但是蘑菇生长的随机性强、生长过程受影响的因素非常多,不仅仅受环境因子的影响,还会受到培养基料成分配比和发酵程度、覆土成分和厚度、浇水量及频率等影响,而且多因素间高度复杂非线性耦合,所以基于生长模型的预测精度通常不高。针对这个问题,LU 等^[26]提出了基于 YOLO V3 目标检测算法识别蘑菇大小,并通过识别的蘑菇大小数据计算蘑菇生长速度,由此预测蘑菇的收获时间。该文提出了基于图像监测下的生长预测方式,从当前蘑菇图像中提取生长状态特征,然后再对特征进行预测。该方法将机器视觉技术引入,直接基于蘑菇的实际生长大小进行序列预测。因为实际生长出来的蘑菇大小的历史趋势中蕴含了环境因子、培养

基、覆土等多因子耦合的生长影响,所以该方法相对仅基于生长环境因子的生长模型的预测要准确。但是这种方法只在时间这一个维度上进行预测,没有考虑周边的蘑菇所处位置、菌丝分布等空间分布信息对蘑菇生长的影响。然而,由于蘑菇是天然密集成簇生长的,相邻蘑菇之间的生长会形成营养竞争和空间竞争关系,蘑菇的密集程度及周围蘑菇的形态和分布都会对其生长产生较大影响。所以,要实现蘑菇高精度的生长预测,复杂空间维度的影响不可忽略。SAKURAI 等^[27]利用卷积网络对叶片的图像进行卷积操作提取到关键特征,然后采用 LSTM 与编码器-解码器模型结合来预测叶片的生长。该方法采用 ConvLSTM 模型基于历史图像序列对未来生长图像进行时空预测,通过 LSTM 提取生长的时间相关性,通过卷积提取生长图像中的空间相关性,实现了对植物生长的时间相关性和空间相关性这 2 个维度的深度特征融合。该方法首次验证了 ConvLSTM 用于农作物生长时空预测的可行性,但是其预测精度不高,仅为 72.8%。AIGNER 等^[28]利用新兴的预测方法生成对抗网络(Generative adversarial network, GAN)^[29-32]来训练 FutureGAN 网络去学习作物历史生长趋势,并分别应用于植物叶片和根系生长过程的预测。但该模型基于作物的掩码进行预测,作物的纹理和颜色信息被忽略。为了克服该问题,王春颖等^[33]基于滤除背景噪声的拟南芥 RGB 连续生长发育图,提出了利用时空长短时记忆模型 ST-LSTM(Spatiotemporal long short-term memory)来预测拟南芥的生长发育,预测结构的相似性达 0.87。但是该方法随着预测时长的增加,结果愈加模糊,并未对叶片的实际位置和形状的一致性进行评估。且仅对提取出的叶片本身进行预测,未将与生长相关的其它空间实物影响因素考虑进去。然而对于蘑菇,其生长背景中的菌丝分布对生长的影响较大,若只对子实体进行预测,预测精度会受到影响。

因此,本文提出基于蘑菇生长的历史图像序列 $\{\hat{X}_{t-M+1}, \hat{X}_{t-M+2}, \dots, \hat{X}_t\}$ 直接预测其未来的生长状态图像 $\{\tilde{Y}_{t+1}, \tilde{Y}_{t+2}, \dots, \tilde{Y}_{t+L}\}$,然后再从预测出的生长图像中提取出生长特征。相比从图像中先提取特征,然后按照特征的历史值序列进行预测,在直接对生长状态图像进行预测的方法中,影响生长关系的更多细节例如菌丝等均在图中一起参与预测,而且蘑菇与周边蘑菇及其所处位置等空间信息也一起参与预测,这种基于时空两个维度的子实体生长状态预测效果能更加接近实际状态。

1 蘑菇生长状态时空预测模型建立

1.1 时空预测模型

端对端蘑菇生长状态时空预测可视为时空序列图像预测问题,即根据观测到的前 M 次(包括当前的第 M 次)生长状态图像,预测得到未来时刻的生长状态图像,公式为

$$\begin{aligned} \hat{Y}_{t+1}, \hat{Y}_{t+2}, \dots, \hat{Y}_{t+L} = \\ \arg \max_{X_{t+1}, X_{t+2}, \dots, X_{t+L}} p(\hat{X}_{t+1}, \hat{X}_{t+2}, \dots, \hat{X}_{t+L} | \hat{X}_{t-M+1}, \\ \hat{X}_{t-M+2}, \dots, \hat{X}_t) \end{aligned} \quad (1)$$

式中 $\hat{X}_t$ ——实际图像

$\hat{Y}_t$ ——预测图像

p ——贝叶斯概率

由于蘑菇生长状态图像具有时间序列特征,而且历史图像信息对未来时刻的生长状态图像的预测有较大影响,因此可以选用具有较好序列关系长期依赖的 LSTM 来进行预测。LSTM 网络通过“门”结构在训练过程中学习该保存或遗忘来实现选择性记忆前序数据,通过“单元状态”来保证序列链中的相关信息,即使较早时刻的信息也能传送到较后时刻的单元中,克服了短时记忆的不足,从而实现信息的远距离传递^[34]。但是基于时间序列的 LSTM 预测并未考虑蘑菇生长状态图像像素间的复杂相关性等空间结构信息。而卷积运算可以较好地提取图像中的空间特征,所以首先将卷积与 LSTM 相结合,采用 LSTM 来学习生长状态图像数据中的时间相关性,通过卷积从历史图像中有效地获得空间相关特征,从而捕获到随着时间变化的空间相关性特征即时空特征。卷积 LSTM 单元结构如图 1 所示。

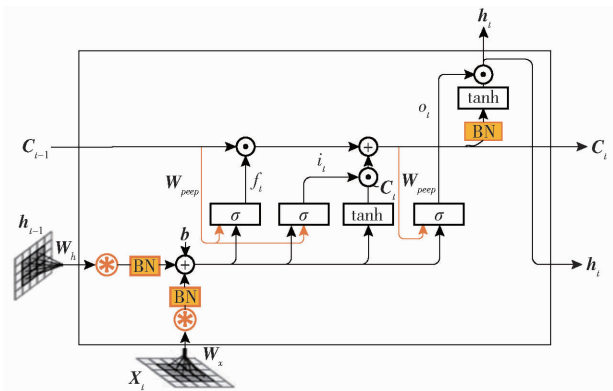


图 1 卷积 LSTM 单元结构示意图

Fig. 1 Structural diagram of ConvLSTM cell

卷积 LSTM 仍然包含输入门、遗忘门、状态门和输出门,在基础 LSTM 的输入门、遗忘门和输出门这 3 个门增加了 peepole 连接,在门限的计算中不仅

仅考虑 X_t 和 h_{t-1} ,还引入了 C_{t-1} 使得整体信息的流动更通畅、门限的计算更准确^[35]。更关键的是,在 LSTM 中,输入的 X_t 和 h_{t-1} 到各门间分别都是全连接层,无法对时空信息提取特征,而卷积 LSTM 则将 LSTM 的 input-to-gate 和 gate-to-gate 两个方向从全连接变为卷积,经过这样的变换,可以实现多维数据中进行卷积运算来捕获空间相关特征。其具体计算公式为

$$\begin{cases} i_t = \sigma(W_{xi} * X_t + W_{hi} * h_{t-1} + W_{ci} \odot C_{t-1} + b_i) \\ f_t = \sigma(W_{xf} * X_t + W_{hf} * h_{t-1} + W_{cf} \odot C_{t-1} + b_f) \\ C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c) \\ o_t = \sigma(W_{xo} * X_t + W_{ho} * h_{t-1} + W_{co} \odot C_t + b_o) \\ h_t = o_t \odot \tanh(C_t) \end{cases} \quad (2)$$

式中 X_t —— t 时刻单元输入

C_t —— t 时刻单元状态

h_t —— t 时刻隐含层输出

i_t ——输出门状态保留概率

f_t ——遗忘门状态保留概率

o_t —— t 时刻输出门输出概率

W_i, b_i ——输入门权重和阈值

W_f, b_f ——遗忘门权重和阈值

W_c, b_c ——状态门权重和阈值

W_o, b_o ——输出门权重和阈值

$*$ ——卷积运算符

$\odot$ ——哈德曼乘积

基于此所构建的蘑菇生长状态图像预测模型的详细架构如图 2 所示。整体预测采用 Sequence-to-Sequence 结构,包含编码器和预测器。首先,为了有效地对图像进行卷积以提取时空相关性特征来进行预测,将输入序列中观测到的生长状态图像转换为 3 维张量,用 $X_t \in \mathbf{R}^{P \times J \times K}$ 来表示,其中, J 表示该 3 维张量网格行数, K 表示列数, P 表示网格中的每个单元格内随时间变化的测量值。

其次,将转换为 3 维张量的生长状态图像序列数据输入到编码器。编码器单元采用将卷积神经网络和长短时记忆网络融合的 ConvLSTM,以有效地捕获生长状态图像预测序列中的时空特征。其中,为了较好地提取历史序列图像中的空间相关性,ConvLSTM 中对输入的 X_t 和 h_{t-1} 的卷积采用 3×3 的卷积核,经卷积、下采样、卷积、下采样以及全连接操作提取到相应特征。此外,为了进一步加深模型深度、提取更深层次的特征,提高预测准确性,又构建了 3 层 ConvLSTM 堆叠的结构,每层 ConvLSTM 的数量由输入图像序列长度决定。3 维张量的图像数

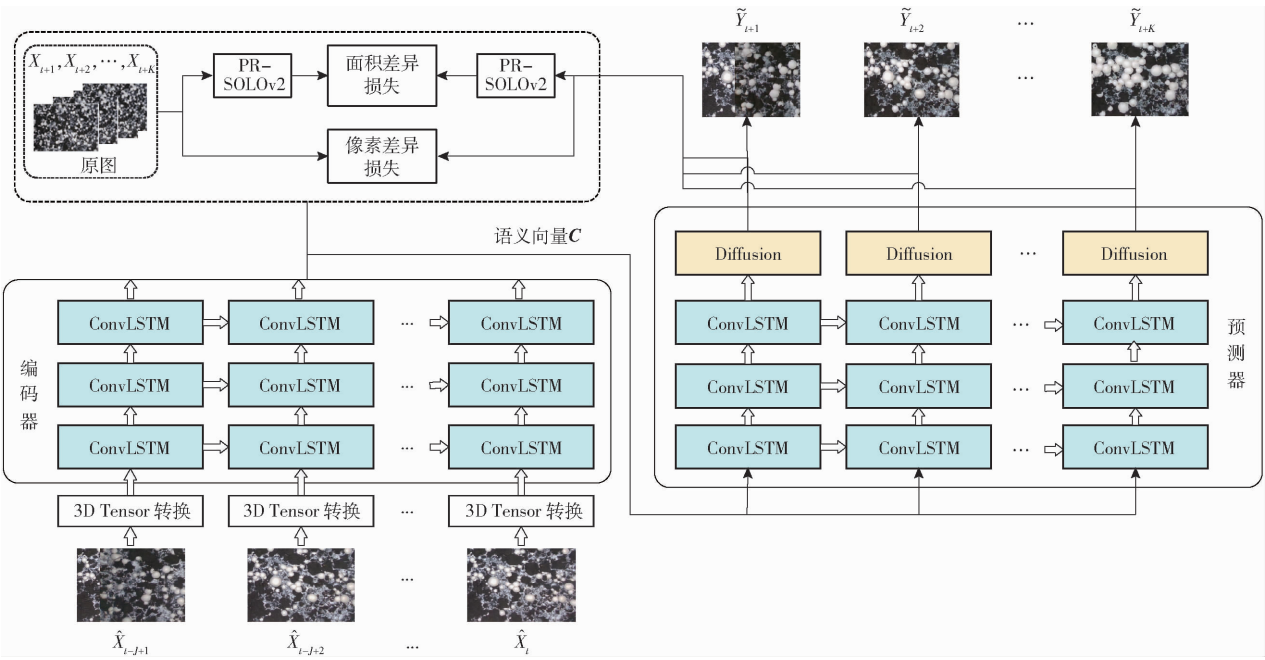


图2 蘑菇生长状态图像预测模型结构图

Fig.2 Structure diagram of mushroom growth status image prediction model

据输入到该多层 ConvLSTM 网络编码处理后得到语义向量 C 。

然后,将语义向量 C 输入预测器进行预测。预测器中也采用 3 层 ConvLSTM 网络,每层 ConvLSTM 的数目由输出序列长度决定。仅采用以上结构的网络进行预测,虽然能够融合卷积的特征提取能力和门控机制的时序记忆能力,使预测精度得到提高,但是由于图像中的预测主体蘑菇的生长过程是动态的,对于这种动态生长模式,不同的空间位置和时间戳会导致连续帧的局部相关结构不同。而普通卷积网络由于卷积核的作用位置固定,不随帧间对应关系的变化而改变。因此,使用位置不变的卷积网络来表示这种位置变量关系会出现模糊、蘑菇边缘不清晰、噪点等情况,导致预测图像的质量和精度降低。针对该问题,本文在预测器后加扩散(Diffusion)模型。利用扩散方程式计算 t 时刻梯度值 G 在 x 和 y 方向上变化的散度。在计算扩散过程中根据图像的不同区域来调整扩散系数,在边缘区域,扩散系数较小,以保持边缘的清晰度;而在平滑区域,扩散系数较大,以增加细节信息,从而增强和恢复图像的细节^[36-37],使得预测的图像质量尽量高、子实体轮廓尽可能清晰。

$$\partial G(x,y,t)/\partial t = \nabla \cdot (c(x,y,t) \nabla G(x,y,t)) \quad (3)$$

式中 $G(x,y,t)$ —— t 时刻点 (x,y) 处的梯度

∇ ——梯度算子

$c(x,y,t)$ ——扩散系数

1.2 损失函数设计

为了进一步提高模型预测图像的质量和精度,

特别是提高生长状态预测问题中的预测主体(各蘑菇)的形状和位置精度,本文对损失函数进行了改进,增加了能减小蘑菇个体差异的损失函数。改进后的损失函数由 2 部分组成:蘑菇差异损失函数和像素差异损失函数。

1.2.1 蘑菇差异损失函数

蘑菇生长状态预测的主体是蘑菇个体,通过预测蘑菇的长势情况,为疏蕾、收获期预测等提供数据基础。所以,需要保证每个蘑菇的形态和位置预测比较精准。因此,需要有一个专门对蘑菇的生长运动预测进行约束的损失函数。由于蘑菇的菌柄比较短,一般小于 60 mm,伸出培养基部分的更短。所以,在预测时长内,蘑菇姿态一般变化不大,主要是蘑菇大小会增长,因此为了降低计算复杂度,本文考虑只采用蘑菇菌盖面积差异来构建蘑菇差异损失函数 $L_{mr}(\hat{X}_t, \tilde{Y}_t)$, $\tilde{Y}_t$ 为预测模型预测的 t 时刻的生长状态图像, $\hat{X}_t$ 为 t 时刻对应的实际生长状态图像。由于一幅图像中往往包含多个蘑菇,因此取预测图像中各个蘑菇面积与实际图像中对应蘑菇面积均方误差(MSE)作为蘑菇差异损失函数,计算公式为

$$L_{mr}(\hat{X}_t, \tilde{Y}_t) = \text{MSE}(\hat{A}_t, \tilde{A}_t) \quad (4)$$

式中 $\hat{A}_t$ ——实际图像 $\hat{X}_t$ 中各蘑菇面积 $\{\hat{a}_1, \hat{a}_2, \dots, \hat{a}_i, \dots, \hat{a}_n\}$

$\tilde{A}_t$ ——预测图像 $\tilde{Y}_t$ 中各蘑菇面积 $\{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_i, \dots, \tilde{a}_n\}$

n ——图像中蘑菇数量,通过 PR-SOLOv2^[38] 高精度实例分割算法分割蘑菇获得

$\hat{a}_i$ ——实际图像中第 i 个蘑菇面积

$\tilde{a}_i$ ——预测图像中第 i 个蘑菇面积

蘑菇差异损失函数的建立可引导模型在训练的时候更多关注蘑菇的预测效果,从而提高蘑菇生长变化预测的精准性。

1.2.2 像素差异损失函数

结合使用图像的像素差异损失函数,以提高预测图像的整体质量。像素差异损失函数 $L_p(\hat{X}_t, \tilde{Y}_t)$ 定义为

$$L_p(\hat{X}_t, \tilde{Y}_t) = \text{MSE}(\hat{P}_t, \tilde{P}_t) \quad (5)$$

其中 $\hat{P}_t = \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_i, \dots, \hat{p}_m\}$

$$\tilde{P}_t = \{\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_i, \dots, \tilde{p}_m\}$$

式中 $\hat{P}_t$ ——实际图像 $\hat{X}_t$ 中每个像素点的 RGB 值

$\tilde{P}_t$ ——预测图像 $\tilde{Y}_t$ 中对应各像素点 RGB 值

$\hat{p}_i$ ——实际图像 $\hat{X}_t$ 中第 i 个像素点 RGB 值

$\tilde{p}_i$ ——预测图像 $\tilde{Y}_t$ 中第 i 个像素点 RGB 值

1.2.3 总损失函数

蘑菇生长状态图像预测模型的总损失函数 $L(\hat{X}_t, \tilde{Y}_t)$ 综合考虑蘑菇差异损失和图像像素差异损失,定义为

$$L(\hat{X}_t, \tilde{Y}_t) = \lambda_1 L_{mr}(\hat{X}_t, \tilde{Y}_t) + \lambda_2 L_p(\hat{X}_t, \tilde{Y}_t) \quad (6)$$

式中 λ_1 ——蘑菇差异损失系数

λ_2 ——像素差异损失系数

2 实验验证

2.1 数据集构建

分别对蘑菇的不同菇房的生长状态情况每隔 10 min 采集 1 次图像,每次采集 3 幅图像,并根据峰值信噪比(Peak signal-to-noise ratio, PSNR)选出失真度最小的一幅作为该时刻的生长状态图像,以此方式连续采集获取到蘑菇的生长周期图像数据。本文共收集 6 个周期蘑菇生长状态序列图像数据作为训练集,共 6 881 幅图像。

2.2 模型训练

模型训练环境为 Windows 10 操作系统、Intel(R) Core(TM) i9 - 10900 F2.81 GHz CPU、GeForce RTX3090 显卡,运行内存为 32 GB,配置 Python 3.7、PyTorch 1.8、CUDA 11.5、GPU 运行架构。

训练参数设置:神经网络优化算法采用随机梯度下降法(Stochastic gradient descent, SGD),训练批量大小为 8,初始学习率为 0.01,迭代次数为 5 000 次。

2.3 评价指标

对于生长状态的图像预测,要求所生成的图像

质量高、子实体失真小,故对生长状态图像预测算法的评价指标采用 2 个常用的生成图像质量评价指标:峰值信噪比(PSNR)、多层次结构相似性(Multiscale structure similarity, MS-SSIM),以及一个自定义的蘑菇预测准确性指标——蘑菇平均交并比(Mushroom mean intersection over union, MR-MIoU)来进行综合评估。

2.3.1 蘑菇平均交并比

本文虽然是对生长状态的图像进行预测,但是进行图像预测的目的是为了更加准确地提取出蘑菇的生长状态特征。所以,不仅需要评估所预测的生长状态图像的质量,更重要的是评估所预测出的生长状态图像中的主体(蘑菇)的准确性。由于所预测的生长状态图像中通常会包含多个蘑菇,所以,本文设计采用预测图像中各蘑菇个体与对应实际图像中各蘑菇的平均交并比(MR-MIoU)来进行评估。当蘑菇的实际区域和预测区域重合面积越大,MR-MIoU 的值也就越大,预测出的子实体区域就越准确。MR-MIoU 的计算公式为

$$\text{MR-MIoU}(\hat{X}_t, \tilde{Y}_t) = \frac{1}{k+1} \sum_{i=0}^k \frac{\hat{M}_i \cap \tilde{M}_i}{\hat{M}_i \cup \tilde{M}_i} \quad (7)$$

式中 $\hat{M}_i$ ——预测图像 $\hat{X}_t$ 中的第 i 个蘑菇掩膜

$\tilde{M}_i$ ——实际图像 $\tilde{Y}_t$ 中的第 i 个蘑菇掩膜

k ——图像中蘑菇数

2.3.2 峰值信噪比

PSNR 是一种常用的图像质量评价指标,用于衡量两幅图像之间差异、评估复原算法性能等^[39]。采用 PSNR 来计算本文算法预测出的图像与实际拍摄的图像之间的差异,PSNR 越高,表示预测的图像越准确。

2.3.3 多层次结构相似性

PSNR 计算简单,物理意义清晰。但是,这种基于均方误差的评价指标是基于误差敏感度的图像质量评价指标,使用线性变换来分解图像信号,不涉及相关性。而自然图像具有极高的结构性,表现在图像像素间存在着很强的相关性。多层次结构相似度(MS-SSIM)是基于不同分辨率的图像,并融合对比度、亮度、结构指标等要素,通过图像结构相关性来衡量图像之间相似度的图像质量评估方法。该指标基于人眼能够提取图像中结构化信息的假设,比传统方式更符合人眼视觉感知^[40]。MS-SSIM 越大,表示预测图像质量越好,为 1 时表示图像预测质量最好。

2.4 实验结果

2.4.1 历史数据步长选择

为了选择适合的图像时间间隔(即步长),本实

验从数据集中分别取不同步长的生长状态历史图像序列来预测相同预测时长的图像,以蘑菇的平均交并比(MR-MIoU)作为评价指标来比较其预测精度。

分别采用时间步长 10、30、60 min 且历史数据取 6 幅时序图像的数据集训练相应生长状态图像预测模型,对比预测时长分别为 60、120、180、360、480、1 000、1 750 min 时的预测精度。预测结果如表 1 所示,其趋势如图 3 所示。

表 1 不同历史序列步长预测的损失值对比

Tab. 1 Comparison of loss values predicted with different step sizes							
时间步	预测时长/min						
长/min	60	120	180	360	480	1 000	1 750
10	0.927	0.924	0.918	0.911	0.903	0.892	0.877
30	0.929	0.926	0.922	0.917	0.910	0.900	0.887
60	0.933	0.929	0.925	0.921	0.915	0.907	0.896

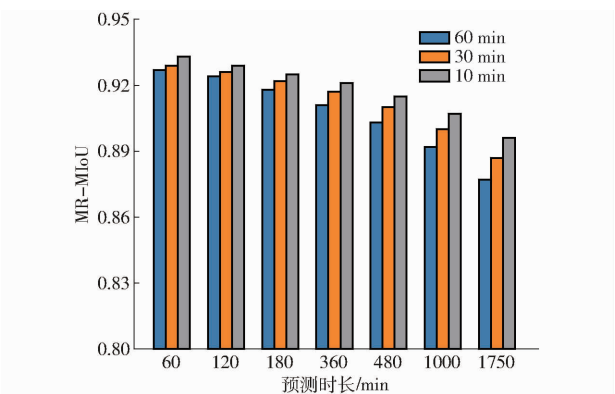


图 3 不同步长预测的 MR-MIoU 趋势图

Fig.3 MR-MIoU trend chart predicted by different step sizes

从图 3 可以看到,对于同一步长,随着预测时长的增加,预测精度逐步降低。这是因为预测时长增加,预测模型对预测距离远的时刻的状态敏感度降低;对于同一预测时长,时间步长 10 min 的预测精度低于 30 min,30 min 的又均低于 60 min 的,而且随着预测时长的增加,该趋势越来越明显,但是预测时长 180 min 和 360 min 的预测精度差别不大,均在 91% 以上。

由于蘑菇在成熟期的生长速度较快,如图 4 所示,一般每 10 h 能增长 1 倍左右。因此,对于天然密集生长的蘑菇,蘑菇之间存在空间竞争,间隔时间越长,竞争越激烈,会使得菌盖被挤压倾斜生长或变形,导致蘑菇的品级下降并增加自动采摘难度。此外,考虑到采收效率和开伞后品级下降问题,间隔 0.5 h 采集数据的管理策略对采收效益更有利。所以,综合考虑预测精度,以及采摘、疏蕾、开伞等蘑菇高质量的生产调控要求,本文选用 30 min 作为历史数据步长,以实现更精细化、精准化管理。

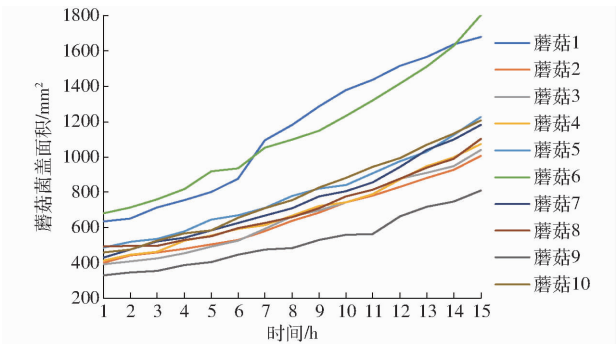


图 4 成熟期不同大小蘑菇生长趋势

Fig.4 Growth trend of mature mushrooms

2.4.2 图像预测模型预测结果和分析

为了测试本节所提出草腐菌生长状态图像预测模型性能,基于构建的蘑菇生长状态预测图像数据集,分别采用 ConvLSTM、SSIM-ConvLSTM 和本文所提的扩散 ConvLSTM (Diffusion ConvLSTM, Diffu-ConvLSTM)3 种预测模型进行训练,并比较验证以上 3 种模型的预测效果,其中 SSIM-ConvLSTM 是在 ConvLSTM 的基础上将损失函数改为损失函数 SSIM 的模型。

2.4.2.1 损失值对比

3 种预测模型损失值变化曲线如图 5 所示。ConvLSTM、SSIM-ConvLSTM 和 Diffu-ConvLSTM 损失值分别为 0.187 1、0.149 4 和 0.073 8,本文改进的图像预测算法 Diffu-ConvLSTM 的损失值最低,比其它两种算法降低 50%。

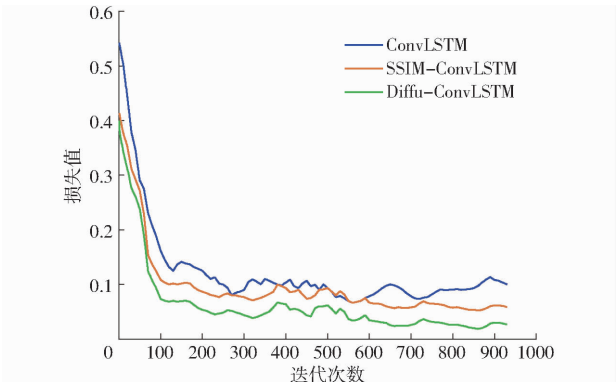


图 5 损失值变化曲线

Fig.5 Change curves of loss

由图 5 还可以看到,Diffu-ConvLSTM 的整个训练过程的损失值也均远底于前两者。3 种模型的损失值收敛速度差别不大,均收敛较快,但是 ConvLSTM 模型出现了过拟合,SSIM-ConvLSTM 和 Diffu-ConvLSTM 模型没有发生过拟合现象,并已训练到最优性能。

2.4.2.2 预测图像质量对比

采用相同的蘑菇生长状态历史图像序列,分别利用 3 种预测模型预测的生长状态图像如图 6

所示。由该图可见,由 ConvLSTM 模型预测的蘑菇生长状态图像存在图像模糊、黄色噪点、子实体轮廓不太清晰等问题,图像质量偏差;由 SSIM - ConvLSTM 预测的图像清晰度有所提高、噪点问题也得到明显改善,图像质量得到提高;由 Diffu - ConvLSTM 预测的生长状态图像比 SSIM - ConvLSTM 更清晰。

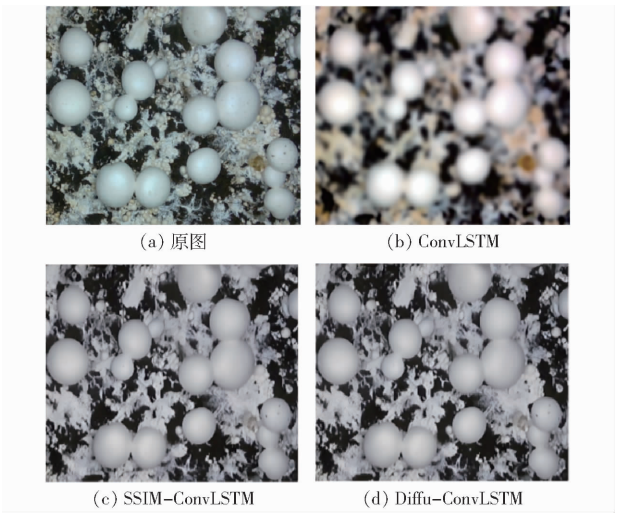


图6 不同模型预测图像模糊程度对比

Fig. 6 Images predicted by three different models

为了进一步对比3种模型的预测图像质量,分别采用3种模型随机预测的处于蘑菇不同生长阶段的20幅6h后的生长状态图像,比较图像质量评价指标PSNR和MS-SSIM。各模型预测图像质量评价指标如表2所示。ConvLSTM、SSIM-ConvLSTM和Diffu-ConvLSTM预测图像的平均PSNR分别为26.178、30.153、35.611 dB。Diffu-ConvLSTM的PSNR值最高,是SSIM-ConvLSTM的1.15倍、ConvLSTM的1.36倍。ConvLSTM、SSIM-ConvLSTM和Diffu-ConvLSTM预测图像的平均MS-SSIM分别为0.695、0.857和0.927。Diffu-ConvLSTM的MS-SSIM值最高,是SSIM-ConvLSTM的1.08倍、ConvLSTM的1.33倍。

表2 不同模型预测图像效果

Tab. 2 Effect of different models on predicting images			
评价指标	ConvLSTM	SSIM - ConvLSTM	Diffu - ConvLSTM
PSNR/dB	26.178	30.153	35.611
MS - SSIM	0.695	0.857	0.927
MR - MIoU	0.75	0.86	0.93

3种模型预测图像的图像质量评价指标变化曲线如图7所示。由图7可见,Diffu-ConvLSTM模型对应的PSNR和MS-SSIM的波动程度明显小于ConvLSTM模型。Diffu-ConvLSTM模型对预测图

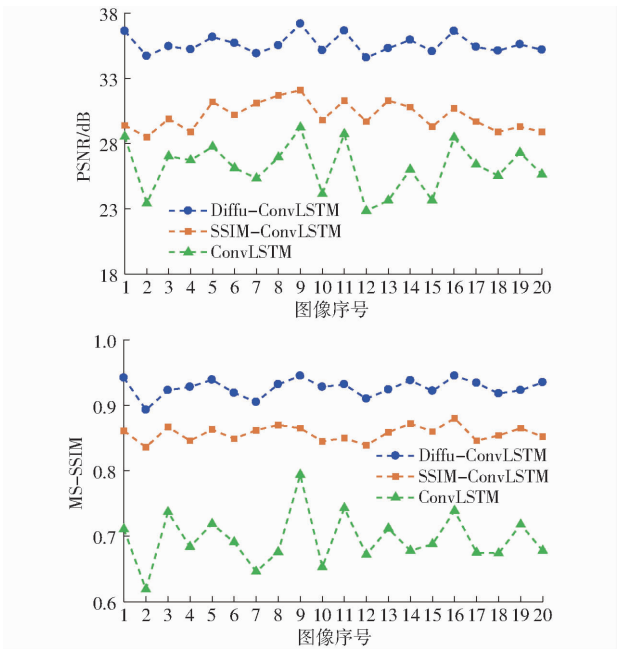


图7 不同模型预测图像的图像质量评价指标变化曲线

Fig. 7 Quality evaluation value curves of images predicted by different models

像的整体质量一致性更优。

通过对预测图像的PSNR和MS-SSIM进行对比分析可知,Diffu-ConvLSTM预测的图像质量比ConvLSTM和SSIM-ConvLSTM更优,这说明Diffusion对图像模糊及噪点的处理效果较好。

2.4.2.3 蘑菇个体预测准确性对比

对于蘑菇的生长状态图像预测,预测图像的质量要求是基本要求,更重要的是还要求对预测的主体(蘑菇)的形态和位置进行精准预测。

本实验采用2.4.2.2节的实验方案,对获得的20幅预测图像计算的蘑菇预测准确性指标(MR-MIoU)进行分析对比。各模型预测图像的MR-MIoU结果如表2所示。Diffu-ConvLSTM的平均MR-MIoU达0.93,预测子实体和实际子实体之间的重叠度非常高,是SSIM-ConvLSTM的1.08倍、ConvLSTM的1.24倍。这说明子实体差异损失函数对减少子实体差异具有明显作用。

由图8可观察到,Diffu-ConvLSTM模型对应的MR-MIoU的波动程度也明显小于ConvLSTM模型。这说明Diffu-ConvLSTM模型对每个子实体的形态和位置的预测一致性更高。

2.4.2.4 单个蘑菇预测效果对比

由于本研究中单个蘑菇预测的准确性也非常重要,因此本实验随机选取12幅预测图像中的共计513个蘑菇,分别计算其实际面积与预测面积,并绘制散点分布图,如图9所示。由该图可知,预测蘑菇面积与实际蘑菇面积的相关系数R达0.94,这说明预测蘑菇与实际蘑菇的相似程度较高。

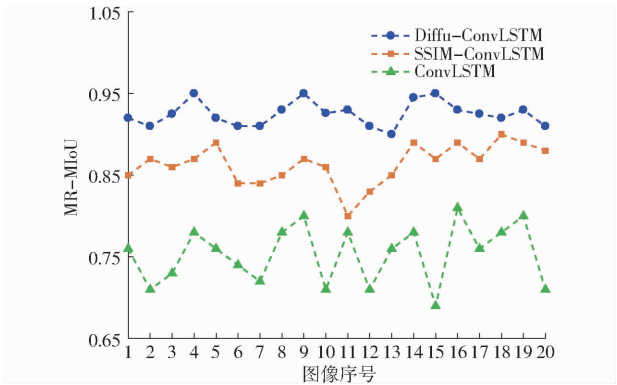


图 8 不同预测模型预测图像 MR-MIoU 变化曲线
Fig. 8 MR-MIoU curves of images predicted by different models

2.4.2.5 预测图像序列

本文模型预测出的蘑菇生长图像序列如图 10a 所示,与对应时刻的实际图像相比,相似度较高。

3 讨论

由实验可知,ConvLSTM 由于采用的是像素 MSE 损失函数,使用线性变换来分解图像信号,不涉及图像中极高的结构相关性,所以预测的图像质量相对偏低,出现了比较明显的模糊和噪声问题;采用像素 SSIM 损失函数,由于其通过图像结构相关性来衡量图像之间相似度来进行训练,考虑了图像像素的相关性,使得预测图像质量得到了提升;而

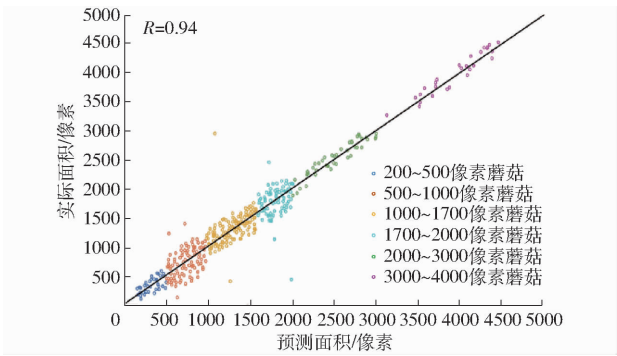
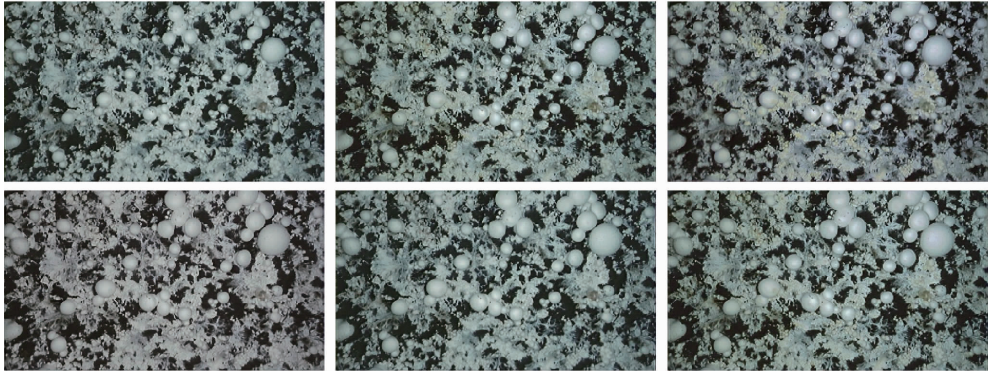
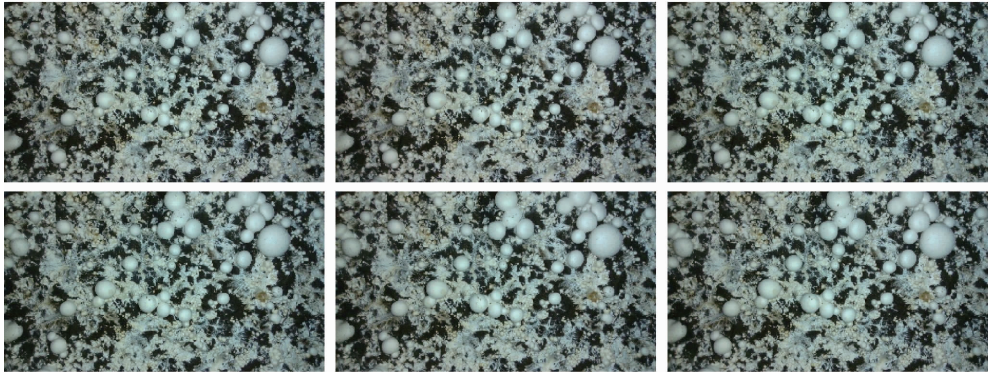


图 9 单个蘑菇实际面积与预测面积散点图
Fig. 9 Scatter plot of actual and predicted area of single mushroom

Diffusion 则通过区域散度来增强和恢复图像的细节,使得图像分辨力能达到更高,所以,其预测图像质量最好。此外,由 MR-MIoU 对比实验可知,SSIM-ConvLSTM 的 MR-MIoU 为 0.85 左右,而 Diffu-ConvLSTM 则能达到 0.93,这说明由于图像质量提高,其图像及子实体边缘更清晰、噪点也较少,可以一定程度上提高子实体的形态和位置预测精度,但是由于这种提高图像质量的方法关注的是图像的整体,并未对子实体生长状态预测的主体(蘑菇)进行特别关注。所以,其对蘑菇的预测精度的提高能力有限。而 Diffu-ConvLSTM 模型由于在损失函数中增加了蘑菇差异损失函数,能更好地引导模型更加关注蘑菇的生长运动,得到更高的蘑菇



(a) 本文模型预测图



(b) 原图

图 10 蘑菇生长图像序列原图与预测图对比

Fig. 10 Comparison between original and predicted image sequences of mushroom growth

形态和位置预测精度,这同时也证明了本文增加的蘑菇差异损失函数的有效性。

4 结束语

提出了一种基于生长时序图像序列的蘑菇生长状态时空预测算法。构建了蘑菇生长状态时空预测数据集;设计了 MR - MIoU 生长状态图像预测评价指标以准确地评价作物生长图像预测中作物形位预测的精度;提出了蘑菇生长状态时空预测算法,采用 Sequence - to - Sequence 框架,由编码器和预测器组成。编码器中将卷积与 LSTM 进行融合、在预测网

络中加入扩散模型,并在像素差异损失函数基础上增加了蘑菇差异损失函数来减小预测蘑菇与实际蘑菇的形态和位置偏差,解决了预测图像的模糊问题和预测蘑菇与实际蘑菇的形位一致性问题。本文模型的 PSNR 为 35.611 dB,MS - SSIM 高达 0.927,分别比 ConvLSTM 提高 36% 和 33%;蘑菇预测精度也最高,MR - MIoU 达 0.93,预测子实体和实际子实体之间的重叠度和相似度都非常高。本文方法是在食用菌领域的首次探索,在预测中融合了蘑菇生长的时空相关性,提高了蘑菇生长状态预测的深度、广度和精确度,能对蘑菇生长状态进行更精准的预测。

参 考 文 献

[1] LIU L, WANG L, LI X, et al. Effects of different bud thinning methods on nutritional quality and antioxidant activities of fruiting bodies of *Pleurotus eryngii*[J]. *Frontiers in Plant Science*, 2022,13:917010.

[2] 高鹏波,孙红红,戚佳攀. 猕猴桃疏果及果实膨大规律的研究与分析[J]. *安徽农学通报*, 2017, 23(1):48,113. GAO Pengbo, SUN Honghong, QI Jiapan, et al. Research and analysis on the laws of the kiwi fruit enlargement and the effects of fruit thinning[J]. *Anhui Agricultural Science Bulletin*, 2017, 23(1):48,113. (in Chinese)

[3] BASSO B, LIU L. Seasonal crop yield forecast: methods, applications and accuracies[J]. *Advances in Agronomy*, 2019,154:201 - 255.

[4] SAVIN I Y, OVECHKIN S V, ALEKSANDROVA E V, et al. The WOFOST simulation model of crop growth and its application for the analysis of land resources[J]. *Interperiodica Eurasian Soil Science*, 1997, 30(7):758 - 765.

[5] HOLZWORTH D P, HUTH N I, DEVOIL P G, et al. APSIM-evolution towards a new generation of agricultural systems simulation[J]. *Environmental Modelling and Software*, 2014,62:327 - 350.

[6] TOVIHOUDI P G, AKPONIKPE P B I, AGBOSSOU E K, et al. Using the DSSAT model to support decision making regarding fertilizer microdosing for maize production in the sub-humid region of Benin[J]. *Frontiers in Environmental Science*, 2019,7:1 - 15.

[7] KUIJPERS W J P, MOLENGRAFT M J G V D, MOURIK S V, et al. Model selection with a common structure: tomato crop growth models[J]. *Biosystems Engineering*, 2019,187:247 - 257.

[8] BOETTCHER U, WEYMANN W, PULLENS J, et al. Development and evaluation of HUME - OSR: a dynamic crop growth model for winter oilseed rape[J]. *Field Crops Research*, 2020, 246:107679.

[9] ASAREE, SCARISBRICK D H. Rate of nitrogen and sulphur fertilizers on yield, yield components and seed quality of oilseed rape (*Brassica napus* L.)[J]. *Field Crops Research*, 1995,44(1):41 - 46.

[10] FISHER R A. Definitions and determination of crop yield, yield gaps, and of rates of change[J]. *Field Crops Research*, 2015, 182:9 - 18.

[11] HAO S, RYU D, WESTERN A, et al. Performance of a wheat yield prediction model and factors influencing the performance: a review and meta-analysis[J]. *Agricultural Systems*, 2021, 194:103278.

[12] PAUDEL D, BOOGAARD H, WIT A D, et al. Machine learning for large-scale crop yield forecasting[J]. *Agricultural Systems*, 2021,187:103016.

[13] KIM N, PARK C W, HUH M, et al. An artificial intelligence approach to prediction of corn yields under extreme weather conditions using satellite and meteorological data[J]. *Applied Sciences*, 2020,10(11):3785.

[14] KHAN S N, KHAN A N, TARIQ A, et al. County-level corn yield prediction using supervised machine learning[J]. *European Journal of Remote Sensing*, 2023,56(1):2253985.

[15] ELBASI E, ZAKI C, TOPCU A E, et al. Crop prediction model using machine learning algorithms[J]. *Applied Sciences*, 2023,13(16):9288.

[16] 马战林,文枫,周颖杰,等. 基于作物生长模型与机器学习算法的区域冬小麦估产[J]. *农业机械学报*, 2023,54(6):136 - 147. MA Zhanlin, WEN Feng, ZHOU Yingjie, et al. Regional winter-wheat yield estimation based on coupling of machine learning algorithm and crop growth model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023,54(6):136 - 147. (in Chinese)

[17] SHARIFI A. Yield prediction with machine learning algorithms and satellite images[J]. *Journal of the Science of Food and Agriculture*, 2021,101(3):891 - 896.

[18] 王来刚,郑国清,郭燕,等. 融合多源时空数据的冬小麦产量预测模型研究[J]. *农业机械学报*, 2022,53(1):198 - 204,458. WANG Laigang, ZHENG Guoqing, GUO Yan, et al. Prediction of winter wheat yield based on fusing multi-source spatio-temporal data[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022,53(1):198 - 204, 458. (in Chinese)

[19] TAWSEEF A S, TABASUM R, FAISAL R L. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming[J]. Computers and Electronics in Agriculture, 2022, 198:107119.

[20] FERCHICHI A, ABBES A B, BARRA V, et al. Forecasting vegetation indices from spatio-temporal remotely sensed data using deep learning-based approaches; a systematic literature review[J]. Ecological Informatics; an International Journal on Ecoinformatics and Computational Ecology, 2022, 68:101552.

[21] MOYSIADIS V, TSAKOS K, SARIGIANNIDIS P, et al. A cloud computing web-based application for smart farming based on microservices architecture[C] // 11th International Conference on Modern Circuits and Systems Technologies (MOCAST). IEEE, 2022:1 – 5.

[22] CAO J, ZHANG Z, TAO F, et al. Integrating multi-source data for rice yield prediction across China using machine learning and deep learning approaches[J]. Agricultural and Forest Meteorology, 2021, 297:108275.

[23] 袁俊杰. 食用菌生长模型及栽培室环境控制系统研究[D]. 镇江:江苏大学, 2007.

YUAN Junjie. Study on growth model and chamber environment control system for edible fungus production[D]. Zhenjiang: Jiangsu University, 2007. (in Chinese)

[24] 鲍军鹏, 刘晓东, 章洁. 用模糊神经网络预测蘑菇生长过程[J]. 计算机工程与应用, 2002(20):221 – 224, 234.

BAO Junpeng, LIU Xiaodong, ZHANG Jie. Predicting mushroom growth process with fuzzy neural network[J]. Computer Engineering and Applications, 2002(20):221 – 224, 234. (in Chinese)

[25] 李毅志, 陆畅, 侯昭宇, 等. 基于温度和相对湿度的香菇子实体生长模型构建[J]. 食用菌学报, 2021, 28(6):87 – 97.

LI Yizhi, LU Chang, HOU Zhaoyu, et al. Establishment of growth models for lentinula edodes fruiting body development based on temperature and relative humidity[J]. Acta Edulis Fungi, 2021, 28(6):87 – 97. (in Chinese)

[26] LU C P, LIAW J J. A novel image measurement algorithm for common mushroom caps based on convolutional neural network[J]. Computers and Electronics in Agriculture, 2020, 171: 105336.

[27] SAKURAI S, UCHIYAMA H, SHIMADA A, et al. Plant growth prediction using convolutional LSTM[C] // 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP), Setúbal: SciTePress, 2019:105 – 113.

[28] AIGNER S, KONER M. FutureGAN: anticipating the future frames of video sequences using spatio-temporal 3d convolutions in progressively growing GANs[J]. arXiv: 10.01325, 2018.

[29] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. Advances in Neural Information Processing Systems, 2014, 27: 2672 – 2680.

[30] 闫保中, 苏邓军. 基于 GAN 网络的时间序列预测算法[J]. 应用科技, 2022, 49(2):114 – 118, 126.

YAN Baozhong, SU Dengjun. A time series prediction algorithm based on GAN network[J]. Applied Science and Technology, 2022, 49(2):114 – 118, 126. (in Chinese)

[31] CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks; an overview[J]. IEEE Signal Process Magazine, 2018, 35(1): 53 – 65.

[32] 马惠玲, 曹梦柯, 王栋, 等. 苹果货架期 GAN – BP – ANN 预测模型研究[J]. 农业机械学报, 2021, 52(11):367 – 375.

MA Huiling, CAO Mengke, WANG Dong, et al. Study on shelf-life prediction of apple with GAN – BP – ANN model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(11):367 – 375. (in Chinese)

[33] 王春颖, 泮玮婷, 李祥, 等. 基于 ST – LSTM 的植物生长发育预测模型[J]. 农业机械学报, 2022, 53(6):250 – 258.

WANG Chunying, PAN Weiting, LI Xiang, et al. Plant growth and development prediction model based on ST – LSTM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(6):250 – 258. (in Chinese)

[34] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735 – 1780.

[35] GERS F A, SCHMIDHUBERER J. Recurrent nets that time and count[C] // Proceedings of the International Joint Conference on Neural Networks. IEEE, 2000:189 – 194.

[36] LIANG Y, YANG Z, ZHANG K, et al. Single image super-resolution via a lightweight residual convolutional neural network[J]. arXiv: 1703.08173v2.

[37] ZHANG K, GOOL L V, TIMOFTE R. Deep unfolding network for image super-resolution[C] // IEEE CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020:10545 – 10554.

[38] 杨淑珍, 朱浩宇, 俞涛. 基于改进的 SOLOv2 的蘑菇簇轮廓分割和重构方法: 202211252987.7[P]. 2022 – 10 – 13.

[39] SARA U, AKTER M, UDDIN M S. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study[J]. Journal of Computer and Communications, 2019, 7:8 – 18.

[40] LI C, BOVIK A C. Content-partitioned structural similarity index for image quality assessment[J]. Signal Processing Image Communication, 2010, 25(7):517 – 526.