

范圣哲, 贡亮, 杨智宇, 等. 面向水稻穗上谷粒原位计数与遮挡还原的轻量级 I2I 深度学习方法 [J]. 华南农业大学学报, 2023, 44(1): 74-83.
FAN Shengzhe, GONG Liang, YANG Zhiyu, et al. A lightweight I2I deep learning method for on-panicle grain in-situ counting and occluded grains restoration[J]. Journal of South China Agricultural University, 2023, 44(1): 74-83.

面向水稻穗上谷粒原位计数与遮挡还原的 轻量级 I2I 深度学习方法

范圣哲✉, 贡亮✉, 杨智宇, 王文杰, 刘成良

(上海交通大学 机械与动力工程学院, 上海 200240)

摘要:【目的】为解决传统水稻考种机谷粒表型分析算法在功能和效率上的局限性, 针对穗上谷粒原位计数和被遮挡谷粒几何特征还原设计一种基于深度学习的轻量级通用算法框架。【方法】将穗上谷粒原位计数与被遮挡谷粒还原这 2 个复杂任务分别拆解为 2 个阶段, 将其核心阶段建模为 I2I 问题。基于 MobileNet V3 设计 1 种能够解决 I2I 问题的轻量级网络架构, 并针对 2 个任务的特点分别设计了数据集图像制作方法, 选择合适的优化策略和超参数对其进行训练。训练结束后, 使用 TensorFlow Lite runtime 解释器将模型部署在考种机的树莓派 4B 开发板上, 并进行测试。【结果】该算法在穗上谷粒计数任务中具有良好的准确性、快速性, 且具有一定的泛化性能。在被遮挡谷粒的形状还原任务中, 该算法所还原的谷粒图像在面积、周长、长度、宽度和颜色分数评价指标中准确率均达到 97% 以上。【结论】该算法能够有效地完成穗上谷粒计数和被遮挡谷粒的还原任务, 且具有轻量级的优点。

关键词: 深度学习; 轻量级; 谷粒计数; 遮挡还原; 通用网络

中图分类号: S237; S511

文献标志码: A

文章编号: 1001-411X(2023)01-0074-10

A lightweight I2I deep learning method for on-panicle grain in-situ counting and occluded grains restoration

FAN Shengzhe✉, GONG Liang✉, YANG Zhiyu, WANG Wenjie, LIU Chengliang
(School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: 【Objective】To address the functional and efficiency limitations of the conventional grain phenotype analysis algorithm of seed analyzers, a deep learning based lightweight general algorithmic framework was designed for two tasks: In-situ counting of on-panicle grains and restoration of occluded grains. 【Method】Two complex tasks of on-panicle grains in-situ counting and restoration of occluded grains were decomposed into two stages, and their core stages were modeled as I2I problems. A lightweight network architecture capable of solving the I2I problem was designed based on MobileNet V3, and the data set generation method was designed according to the characteristics of these two tasks. Then the network was trained with appropriate optimization strategies and hyperparameters. After training, the model was deployed and tested with TensorFlow Lite runtime on Raspberry Pi 4B development board. 【Result】The algorithm had

收稿日期: 2022-02-12 网络首发时间: 2022-10-20 11:03:48

首发网址: <https://kns.cnki.net/kcms/detail/44.1110.S.20221020.0838.002.html>

作者简介: 范圣哲, 硕士研究生, 主要从事机器视觉和农业机器人研究, E-mail: fanshengzhe@sjtu.edu.cn; 通信作者: 贡亮, 副研究员, 博士, 主要从事农业机器人设计与控制、生物特征图像识别、嵌入式智能计算等研究, E-mail: gongliang_mi@sjtu.edu.cn

基金项目: 国家自然科学基金 (51775333)

good accuracy, rapidity and some generalizable performance in the task of on-panicle grain counting. In the task of occluded grains shape restoration, the evaluation accuracy of the restored images in the metrics of area, perimeter, length, width and color score were all over 97%. **【Conclusion】** The algorithm proposed in this paper can complete the task of on-panicle grain counting and occluded grains restoration effectively, and also has the advantage of being lightweight.

Key words: Deep learning; Lightweight; Grain counting; Occlusion restoration; General network

水稻是世界上最重要的农作物之一^[1-2]。在水稻生产和育种过程中,使用适当方法对水稻性状进行表型学分析具有重要的指导价值。现代化表型学分析主要由自动考种系统完成^[3-4]。其中穗上谷粒原位计数是一件基础、重要而又困难的工作。由于穗上谷粒数量巨大,且图片中通常夹杂穗秆等杂物,存在严重的遮挡现象,因此是一件极具挑战性的工作。

遮挡谷粒的还原对于精准计算谷粒的几何特征具有重要意义。在现有的视觉算法中,物体遮挡会极大影响算法准确性^[5-6],且缺乏有效的解决方案。

近年来,计算机视觉已在多个领域取得落地成果:如作业目标检测^[7]、叶绿素含量预测^[8]、植物病害检测^[9]等。这些成果得益于深度学习技术的发展。在视觉领域中,深度学习的主要工具是卷积神经网络,其所能完成的任务主要可分为图像分类^[10]、目标检测^[11]、语义分割^[12]3 大类。深度学习技术在上述 3 大领域的出色表现使其在复杂的图像中对穗上谷粒计数成为可能。近年来,已有部分学者对使用深度学习算法进行目标计数进行了初步探索。

目前主流的目标计数方法可以分为“基于目标检测”和“基于语义分割”2 大类。前者需要利用网络在图像中准确标出每个目标的位置和尺寸,故广泛应用于交通场景的检测和计数^[13-14]。但现阶段在密集、模糊的场景下对各类目标进行精确检测仍是计算机视觉领域前沿的一大难题,且目标检测网络通常结构复杂,难以轻量化部署。因此,使用目标检测网络完成穗上谷粒的计数并非理想的方法。相比之下,使用语义分割网络解决谷粒计数问题仅需将图像逐像素预测为属于谷粒或背景,是一类由图像到图像 (Image to image, I2I) 的二分类问题,且语义分割网络相对简单,更便于轻量化部署。目前已有多种利用语义分割算法完成颗粒计数的成功案例,其计数对象包括人类胚胎细胞^[15]、植物叶片上的蚜虫^[16]及高粱^[17]等农作物。由此可见,使用语义分割算法实现复杂穗上场景的颗粒计数问题更具有可行性。

而被遮挡谷粒的还原问题类似于图像的语义

分割,本质上也可归为 I2I 的映射问题。近年来,一系列学者也将语义分割的相关算法应用于 RGB^[18]和深度图^[19]的修复任务。因此,将图像分割网络应用于遮挡谷粒的还原任务也具有可行性。

除此之外,深度模型对存储空间和算力的过高要求一直是其落地的一大阻碍。早期的大型网络如 AlexNet^[20]、VGG-16^[21]、ResNet-152^[10]无法在移动设备有限的算力条件下部署。近年来出现了一系列针对移动端的轻量级深度学习网络和计算框架。网络架构方面,如 MobileNet^[22-24]使用深度可分离卷积替代传统卷积,大大降低了计算量和参数量。其 V3 版本结合神经架构搜索 (Neural architecture search, NAS)^[25]和元学习 (Meta learning)^[26]技术,进一步优化了模型性能、降低了计算量和参数量。除此之外,MobileNet V3 与 DeepLab 系列的 ASPP^[27]技术结合,便可在轻量级设备上完成语义分割。计算框架方面,在 TensorFlow 框架中,开发人员针对边缘计算设备设计了 TensorFlow Lite runtime (TF Lite) 解释器,以便在边缘设备上完成轻量级模型的调用。

基于上述背景和技术条件,本文提出了一种轻量级 I2I 穗上谷粒表型分析方法,针对稻穗上谷粒原位计数与被遮挡谷粒还原这 2 个复杂任务进行阶段分解,将其核心阶段建模为一类 I2I 问题。在此基础上基于现有的考种机设备,为穗上谷粒计数和遮挡谷粒还原任务分别设计了一套训练集图像增广算法,基于 MobileNet V3 和 ASPP 模块建立了一套解决 I2I 问题的语义分割网络,将其分别应用于上述 2 个任务。使用 TF Lite 解释器将算法部署在考种机的树莓派设备上,并为穗上谷粒计数和遮挡谷粒还原任务分别设计了一套模型性能的评估方案。

1 问题建模与数据集制作

1.1 I2I 问题与方法

在制作数据集前首先需要对问题加以建模和定义。对自动化考种机而言,进行穗上稻粒原位计

数与被遮挡谷粒还原对高效、准确地完成水稻表型学分析具有重要意义。由于原位计数和遮挡还原这 2 种任务过程十分复杂、需要实现的目标相差较大,在传统的考种机算法框架下,解决这 2 种任务的方法很难统一。本文通过对任务的拆解,一方面将复杂的、难以实现的任务整体拆分为若干个简单的、容易实现的算法单元,另一方面将上述 2 种任务的核心阶段统一建模为 1 种 I2I 的映射问题。

具体而言,在本文考种机算法的框架下,完成穗上谷粒计数任务的算法被拆解为如下 2 个阶段: 1) 由谷粒稻穗图片得到谷粒分布概率图; 2) 将谷粒分布概率图先进行二值化,再将二值图进行连通域计数,即可得到穗上谷粒的数量。其中第 2 阶段是 1 个简单问题: 图像二值化可以使用大津法 (OTSU) 完成,连通域计数可以使用 openCV 中的 findContours 函数完成,二者在传统数字图像处理领域中都有十分成熟的解决方案。而第 1 阶段是 1 个复杂问题: 由复杂的穗上谷粒图像得到谷粒分布概率图目前尚无现成的解决方案。但这一问题本质上是由 RGB 图像到灰度图的映射,因此可以归纳为 I2I 问题。

类似地,被遮挡谷粒的还原问题也可被拆解为如下 2 个阶段: 1) 提取和移除“相互遮挡谷粒对”中的上层谷粒; 2) 将残缺的下层谷粒进行补全。其中第 1 阶段是 1 个简单问题: 由于“相互遮挡谷粒对”中的上层谷粒有清晰明显的轮廓边界,使用分水岭、小波等传统图像分割算法便可以得到非常理想的结果。而第 2 阶段是 1 个复杂问题: 被遮挡的下层谷粒轮廓残缺不全,部分颜色、纹理细节丢失,因此完成图像恢复是很困难的工作。由被遮挡的谷粒图像得到完整的谷粒图像,其本质上是 RGB 图像到 RGB 图像的映射问题,故也可以建模为一种 I2I 问题。

综上所述,这 2 种任务中最困难且最核心的阶段都可以建模为一类 I2I 的映射问题,本文将解决这一类问题的方法定义为 I2I 方法。由于原位计数和遮挡还原任务中的 I2I 问题都十分困难,本文为这 2 种任务设计 1 种基于卷积神经网络和深度学习的 I2I 方法。而由于 I2I 问题均具有“由图像到图像映射”的显著共性,因此,本文为解决原位计数和遮挡还原的深度学习 I2I 方法设计一种统一网络框架。

1.2 穗上谷粒原位计数数据集的制作

深度学习是一种基于统计学习的方法,其本质是寻找由输入分布至输出分布的一种映射关系,因

此使用深度学习方法解决 I2I 问题,首先需要制作足够的试验样本,以便后续针对数据集特点设计算法。应用于深度学习算法的数据集通常包括训练集和测试集 2 部分。

1.2.1 训练集制作 试验使用的水稻样品为籼稻品种‘JP69’和‘交源 5A’杂交所得的重组近交系群体,种植于上海市闵行区 (N 31.03°、E 121.45°) 的试验田中。所使用的试验样本同时包括乳熟期、蜡熟期、完熟期 3 个阶段。水稻图像由上海交通大学贡亮副研究员所在实验室自制的考种机于 2018 年 3 月采集。原始数据集的 RGB 图像 (如图 1a 所示) 共 45 张,图像中的谷穗经吸铁石压块固定,背景板为塑料材质且略有反光,上方由 LED 光源照明。

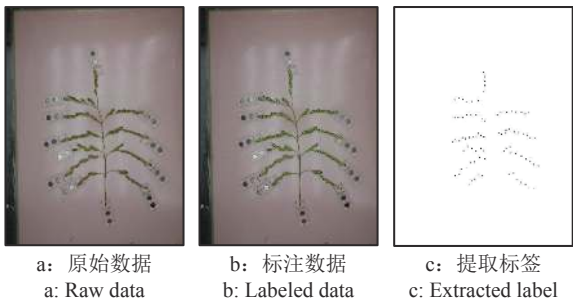


图 1 谷粒计数网络训练数据
Fig. 1 Training data of grain counting network

在深度监督学习任务中,需要一定数量的人工标记样本作为标签。标注过程中,考虑到完整标注谷粒形状的成本过于昂贵,且难以分辨黏连的谷粒,因此仅采用绿色圆点标记对谷粒进行大致标注。标记点的尺寸也需适中: 过大的标记易导致标记间相互重叠,从而影响分割算法的精度; 过小的标记会造成模型对于标记的位置过于敏感,且加剧样本不均问题。经试验测定,使用谷粒平均面积的 30%,即直径约 4 mm 的标记,效果为佳。标注后的训练图像如图 1b 所示,其中单独提取的标签图像如图 1c 所示。标签图像是单通道二值图像,属于谷粒的像素位置被标注为 1,其余位置被标注为 0。对于用于语义分割的神经网络,模型的输出应当与该标签图像保持尽量小的误差。

数据集中图像的像素分辨率高达 $2\,448 \times 3\,264$,不宜直接输入网络: 由于图像尺寸过大且需计数的谷粒小而多,若通过缩小尺寸的方式输入网络,将导致谷粒细节丢失、难以分辨,从而大大影响计数的精度; 而如果保持图像的尺寸不变但增加网络的输入尺寸,将引起网络的参数急剧增加,不仅容易发生过拟合,还会急剧提升所需的存储空间和计算

量,致使模型过度臃肿而难以在移动端部署。

为解决输入尺寸过大的问题,图像输入网络前将进行预切块操作。为保证样本切块的随机性和前景、背景像素数量的均衡性,切块策略被设计为 2 个交替执行的阶段:第 1 阶段使用完全随机的策略在原尺寸的图像中采集像素为 489×652 的图像块若干,第 2 阶段将丢弃谷粒标注面积占比小于全图面积 0.01% 的图像块。经过第 2 阶段的筛选,不含谷粒的图像块将被过滤,这将保证训练集中的前景、背景像素所占比例与实际情况相差不大。2 个阶段将反复执行,直至每张原始图像中均已采集 20 张不含纯背景的图像块,最终得到 900 张图像块用于深度网络的训练。

为使训练集中的图像样本能够尽可能逼近复杂的真实情况,同时提升深度模型的泛化能力以及降低过拟合的风险,需要对已生成的训练集图像进行样本增强:采用 TensorFlow 提供的 tf.data 模块将图像通过随机翻转、随机放缩、随机亮度、随机对比度、随机饱和度等方式进行增广。除此之外为方便运算,将所有图像块尺寸取整缩放至 490×650。得到的训练图像块样本如图 2 所示。

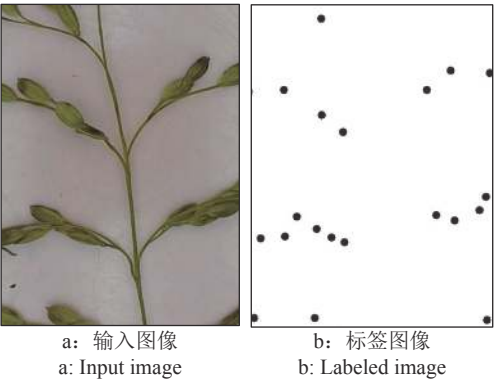


图 2 经分块后的图像块样本
Fig. 2 Image block samples after slicing

1.2.2 测试集制作 测试集用来测试模型训练后的性能,因此在模型预测时需保证图像各分块的相对位置不变,故图像块不再随机采样,而是将图像按顺序切块。测试的图像块尺寸与训练集保持一致,即 490 像素×650 像素。当图像块的尺寸不足 490×650 时,直接缩放至该尺寸。模型推理完成后,将各个结果的图像块按顺序恢复至原大小和位置,即可得到模型预测的完整谷粒分布概率图。

1.3 谷粒去遮挡任务数据集的制作

在制作被遮挡谷粒还原任务的数据集时,若使用考种机采集真实的“相互遮挡谷粒对”图像,将存在下列问题:1) 需要大量的试验操作,数据采集

工作量大、周期长;2) 在由“相互遮挡谷粒对”采集下层被遮挡谷粒图像时,需要手动移除上层谷粒,不可避免地会触碰到下层谷粒,从而引入位置误差,数据集的精度将难以保证。

为解决上述 2 点问题,本文采用计算机生成的模拟数据集完成谷粒去遮挡任务。具体而言,使用空白谷粒形状遮挡单个谷粒的图像作为被遮挡谷粒的图像样本,即模拟经第 1 阶段处理后已去除上层谷粒后的图像,如图 3 所示。该数据集原始的图像尺寸为 64 像素×64 像素,由于尺寸太小会影响模型的计算,输入模型前将图像缩放至 128×128。原始数据集由 100 张图片构成,通过随机旋转、随机亮度等方式,数据集被增广至 8 000 张。

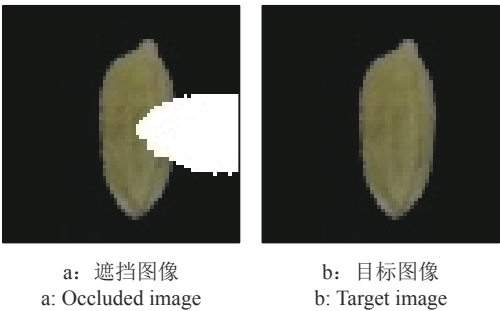


图 3 遮挡还原图像样本示例
Fig. 3 Samples of occlusion restoration images

2 解决 I2I 问题的轻量级网络

穗上谷粒原位计数的第 1 阶段“由谷粒稻穗图片得到谷粒分布概率图”及被遮挡谷粒还原任务的第 2 阶段“将残缺的下层谷粒进行补全”均为 I2I 问题,而深度学习中的语义分割网络可以解决 I2I 问题。此外,由于算法部署在考种机的树莓派开发板上,需要严格控制所需算力和模型体积,故选用带有 ASPP 模块的 MobileNet V3 轻量级网络。该网络保持了 MobileNet 系列深度可分离卷积的传统,大大压缩了计算量、提高了运算效率。除此之外,MobileNet V3 网络在前代的基础上引入了 Squeeze-and-Excite 结构,在前代 Swish 函数^[28]的基础上提出了 h-Swish 激活函数:

$$h\text{-Swish}(x) = x \frac{\text{ReLU6}(x + 3)}{6}, \tag{1}$$

式中: x 代表激活函数输入的自变量, ReLU6^[22] 是对 ReLU 激活函数的优化。

$$\text{ReLU6}(x) = \min[\max(x, 0), 6]. \tag{2}$$

MobileNet V3 使用神经架构搜索技术进行网络结构的搜索,从而进一步优化网络的性能。MobileNet V3 有 Large 和 Small 2 个版本,为了尽

可能地提升在边缘设备的计算性能, 本文选用文献 [24] 提出的 Small 版本的网络。在此网络基础上, 针对穗上谷粒计数和遮挡谷粒还原任务的不同特点, 将算法设计如下。

2.1 穗上谷粒计数网络设计

在该任务中, 网络需要预测每个像素点属于谷

粒标记 (前景) 还是背景, 所以是分类问题。完成这一任务的网络结构如图 4 所示。网络的输出是分别对应于正样本 (谷粒标记) 和负样本 (背景) 的 2 张概率图。由于最后一层前 logits 值被 Softmax 函数激活, 得到的 2 张概率图对应位置的概率相加等于 1。

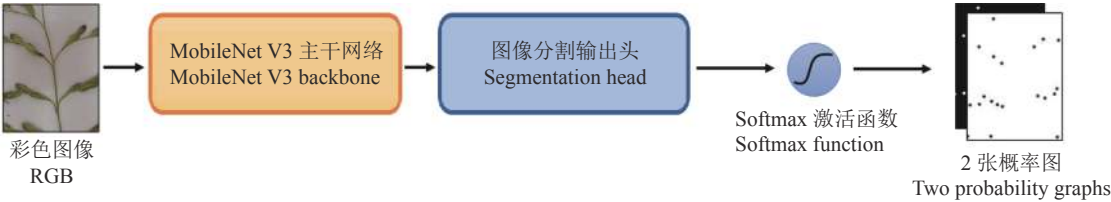


图 4 用于谷粒分割的 MobileNet V3 网络
Fig. 4 MobileNet V3 network used for grain segmentation

由于该问题被建模为分类问题, 根据极大似然估计原理, 该模型的损失函数使用交叉熵 (Cross entropy loss, CE Loss), 即对于每个像素点 p 的损失 (l) 可计算如下:

$$l(y_p^*, y_p, \lambda) = -[\lambda y_p^* \log(y_p) + (1 - y_p^*) \log(1 - y_p)], \quad (3)$$

式中: y_p^* 为像素的真实标签, 取值为 0 (背景) 或 1 (前景, 即谷粒标记); y_p 为网络预测的概率, 是 (0, 1) 之间的浮点数; λ 是衡量正样本权重的超参数。在该任务中, 由于背景元素的像素数量远大于谷粒标签的像素数量, 即存在严重的样本不平衡问题, 因此需要调节该参数以加大正样本预测结果的权重。经过反复调参对比, 最终选择 $\lambda = 2$ 。

对于每个训练批次 (Mini-batch) 而言, 损失可计算如下:

$$CE\ Loss = \frac{1}{NH_1W} \sum_{i=1}^N \sum_{p \in Y_i} l(y_p^*, y_p, \lambda), \quad (4)$$

式中: N 表示 Mini-batch 中所含样本数量 (Batch size), H_1 表示图像高度, W 表示图像宽度, Y_i 表示第 i 张图像的像素点集合。

为在宏观尺度衡量语义分割的准确度, 本文使用交并比 (Intersection over union, IoU) 作为评估指标。以 0.5 为阈值对模型的输出概率图做二值化后, 按下式与真实标签图像求 IoU:

$$IoU = \frac{\text{sum}(Y^* \otimes Y_t)}{\text{sum}Y^* + \text{sum}Y_t - \text{sum}(Y^* \otimes Y_t) + \delta}, \quad (5)$$

式中: Y^* 为标签图像中取值为 1 的像素; Y_t 为二值化图像中取值为 1 的像素; 算子 $\otimes$ 表示哈达玛积; δ 是为防止出现除零错误的小量, 可取 $\delta = 10^{-5}$ 。

为了避免大量繁复的调参工作, 本文选用初始学习率为 0.001 的 Adam 优化器 [29], 对比了带重启

的余弦学习率衰减策略 [30] 以及指数衰减策略。这 2 种学习率衰减策略的情况如图 5a 所示。经过 50 轮的训练, 2 种学习率衰减策略在训练集上的损失函数如图 5b 所示。由此可见, 在该任务中指数衰减策略的表现优于余弦衰减策略。故使用指数衰减策略完成网络的训练。

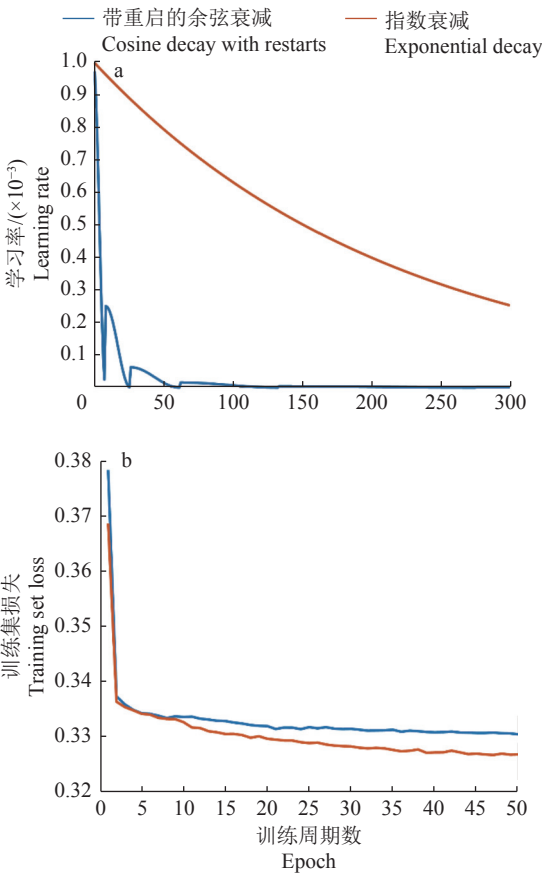


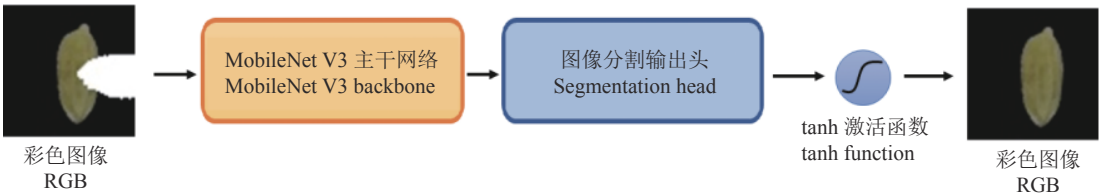
图 5 2 种学习率衰减策略
Fig. 5 Two strategies of learning rate decay

2.2 遮挡谷粒还原网络设计

该任务本质上是一种回归问题, 完成被遮挡谷

粒还原任务的网络结构如图 6 所示。模型的输入为被遮挡的谷粒图像, 模型的输出是完整的谷粒图

像。遮挡还原任务中, 网络不仅需要保持输入图像中未被遮挡的部分, 还需要还原已被遮挡的部分。



由于模型的输入及输出均为像素值在区间 $[-1,1]$ 的图像, 谷粒计数网络中的 Softmax 函数输出范围为 $(0,1)$, 因此不能沿用。如式 (6) 所示的双曲正切函数 $\tanh$ 可以完成从实数域 R 到区间 $(-1,1)$ 的映射, 因此可替代 Softmax 用作该任务中输出层的激活函数。

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{6}$$

式中: x 表示激活函数输入的自变量。

除此之外, 由于该模型完成回归任务, 故采用均方误差损失函数 (Mean square error loss, MSE Loss)。对于每个 Mini-batch, 其计算式如下:

$$\text{MSE Loss} = \frac{1}{NH_1W} \sum_{i=1}^N \sum_{p \in Y_i} (y_p - y_p^*)^2. \tag{7}$$

3 计数网络的训练、测试和部署

在建立训练集、测试集和确定网络结构后, 需对网络进行训练。为在多个训练模型中挑选最优的模型, 需在验证集中对网络的性能进行评估。在验证集中表现最优的模型将被选择为最终部署的模型, 通过 TensorFlow 导出并使用 TF Lite 解释器在考种机的开发板——树莓派端进行部署。最终, 为测试模型在树莓派上运行的性能, 将整个测试集在树莓派上进行推理, 与人工标注的结果进行比对, 得到模型的准确率和运行效率。

3.1 谷粒计数网络的训练

按照“1 问题建模与数据集制作”所述的方法分别建立容量为 900 的训练集和验证集, 每轮的训练阶段使用全部的训练集数据, 验证阶段在验证集中抽取 10 个 Mini-batch 完成验证。该模型训练过程中的各个超参数为初始学习率 0.001、衰减步数 2000、衰减率 0.96、批容量 4、总轮数 300。

将每轮训练在训练集和验证集上的平均损失

和 IoU 绘制成折线图 (图 7)。随着训练过程的不断推进, 模型在训练集和验证集的损失和 IoU 虽有小幅波动, 但最终趋于收敛。模型在验证集的损失降至 0.33 以下, IoU 上升至 0.98 以上。

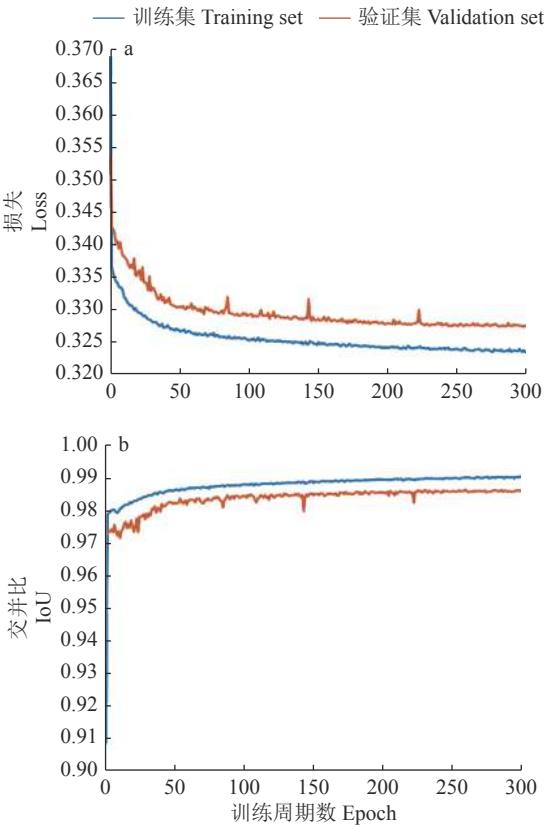


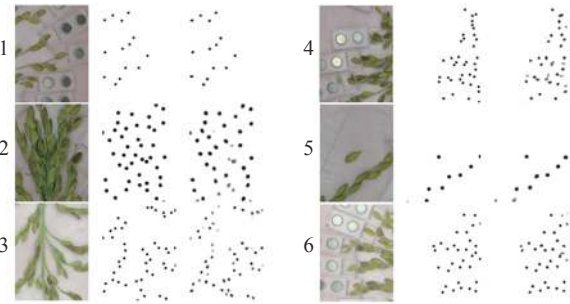
图 7 损失 (a) 和交并比 (b) 的可视化

Fig. 7 Visualization of loss (a) and IoU (b)

为更加直观地考察模型的性能, 将最优的模型参数保存, 部分预测结果可视化如图 8 所示。该模型对于谷粒相对稀疏的位置能够比较精确地预测, 而对于谷粒较为密集位置预测的置信度相对欠佳, 但准确率尚可。这是由于谷粒密集处谷粒间相互遮挡和变形严重, 识别难度将大大增加。

3.2 部署、测试与评估

3.2.1 模型部署 该模型在树莓派 4B 开发板的部署使用 TF Lite 解释器完成。经优化后的 MobileNet



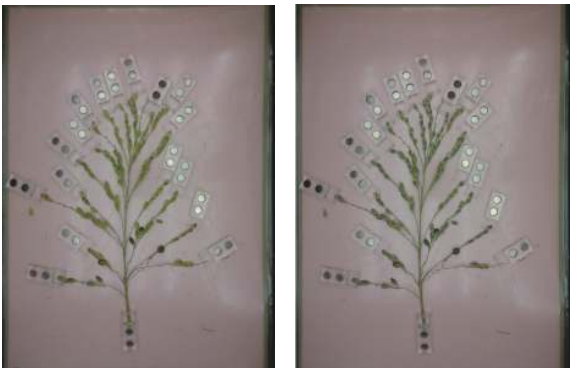
每组图片中, 左: 输入图片, 中: 标签图片, 右: 预测图片
In each group of pictures, the left is the input picture, the middle is the labeled picture, the right is the predicted picture

图 8 6 组遮挡还原图像样本

Fig. 8 Samples of six sets of occlusion restoration image

V3 模型文件仅占 1.7 MB, 在 PC 端前向推理 1 次需 7 s 左右, 在树莓派 4B 开发板运行 1 次前向推理耗时约 19 s。本文提出的算法模型的体积 (1.7 MB) 低于文献 [31] 提出的算法模型的体积 ($H=48$ 、 $K=10$: 8.8 MB、 $H=80$ 、 $K=16$: 22.5 MB, $H=112$ 、 $K=24$: 50.5 MB, $H=160$ 、 $K=32$: 89.8 MB), 本文提出的算法具有明显的轻量级特点。

为了使本文算法与多种分辨率的移动设备相兼容, 在进行前向推理时, 完整的图片将被按照 490×650 的大小切分成 40 块, 对于尺寸不足该大小的块, 使用双线性插值算法上采样到该尺寸。若图像被切分的块数不足 40, 将补 0 至 40 块。为充分发挥并行计算的优势, 该 40 个图像块将被打包成 1 个 Batch size 为 40 的 Mini-batch 输入 TF Lite 解释器, 进行前向推理得到 40 个结果图像。该 40 张结果图像将按照原来的尺寸和位置被拼成 1 张完整图片。将模型输出的 2 张概率图使用大津法进行二值化, 二值图中为 1 的部分将被标记为谷粒, 且使用蓝色被标注在原图像上, 得到的结果如图 9 所示。

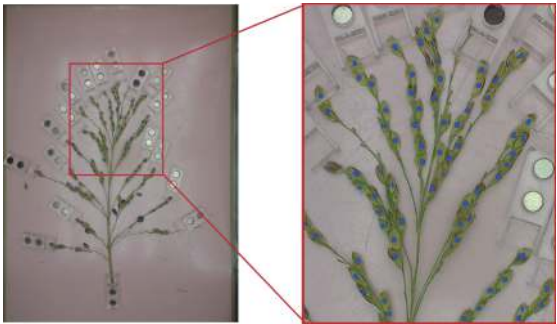


a: 输入图像
a: Input image
b: 可视化图像
b: Visualized image

图 9 穗上谷粒计数结果的可视化

Fig. 9 Visualization of on-panicle grain counting results

3.2.2 常规测试集下的评估 采集 45 张与训练集背景、光照条件完全一致的图像作为测试集, 如图 9a 所示。由图 9b 结果可见, 该算法对于穗上谷粒的计数总体上较为精确。考虑到训练时模型对于谷粒相对密集区域的预测难度较大, 故将该测试图像谷粒最密集的区域放大, 如图 10 所示。



a: 谷粒计数结果
a: Grain counting result
b: 局部放大
b: Local magnification

图 10 谷粒计数结果及其局部放大图像

Fig. 10 Grain counting result and its local magnification image

由图 10 可以看出, 该算法在谷粒相对密集的区域也有不俗的标记效果, 对于饱满粒的识别较为精确, 对于瘪粒和遮挡粒仅有少量遗漏。在 45 张测试样本上测试该模型的性能, 得到模型的平均准确率为 96.81%, 超过了文献 [32] 中小波方法的准确率 (94%) 及文献 [33] 中 P-TRAP 算法的准确率 (86%)。

另外, 为了评估模型在不同切片大小情况下的性能, 选取 6 种不同尺寸但长宽比相同的切片, 测试结果如表 1 所示。可以看到, 所选用的 650×490 是使得模型性能最优的切片尺寸, 其余尺寸均会一定程度上损伤性能, 且过大的切片会严重影响模型性能。

表 1 不同切片尺寸下算法的准确率对比

Table 1 Accuracy comparison of algorithms for different slicing sizes

长度/像素 Length	宽度/像素 Width	平均准确率/% Mean accuracy
780	588	91.56
748	564	92.28
715	539	94.53
650	490	96.81
682	514	95.66
620	467	96.73
566	426	94.25

3.2.3 域偏移测试集下的评估 常规测试集中测得的准确率能够体现该模型的基本性能, 但却无法衡量模型的泛化性能。模型的泛化性能是模型投入实际应用所需要的一项重要指标, 它体现了算法对环境的鲁棒性。在存在域偏移的测试集下测试模型性能能够很好地体现模型的泛化性。

在统计学中, 传统测试集的数据分布应当与训练集数据分布共同服从于总体分布, 而存在域偏移的数据集中, 测试集与训练集的数据分布服从各自的总体, 具有明显的区别。由于其违背训练集、测试集共同服从于总体分布的统计假设, 深度学习模型在域偏移数据集下的准确率往往会大幅下降^[34]。

为采集具有域偏移的图像样本, 使用来自五常市古农水稻种植专业合作社的稻花香‘五优稻 4 号’水稻, 在新一代考种机平台上以不同的背景和光照条件采集 50 张如图 11 所示的图像。新版本的考种机将原有的顶光灯拍照更改为背光灯拍照, 且将图像的背景由白色塑料板更改为蓝色背景板。除此之外, 由于图像采集的季节并非水稻收获季节, 用于采集图像的稻穗样本是枯黄的陈年样本。以上诸多因素导致新采集的图像已经与原图像相差甚远, 为算法的泛化能力和鲁棒性带来了很大挑战。

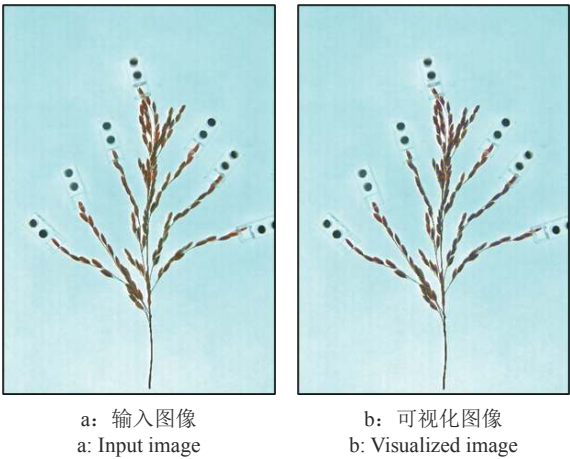


图 11 新采集样本的计数结果
Fig. 11 Counting results on newly collected samples

从图 11 可以看出, 该算法对于新采集的图像样本总体上也有较好的识别效果。为进一步考量算法的性能, 将图像中谷粒较为密集的区域放大, 如图 12 所示。

如图 12 所示, 在该区域内算法能够标记出绝大部分谷粒, 个别谷粒的标记较小但仍有标记。按照连通域计数算法的要求, 图像中所有出现标记的连续区域无论大小都会被标记为谷粒, 因此这些谷粒也会被计算在其中。总体而言, 在该区域仅有小

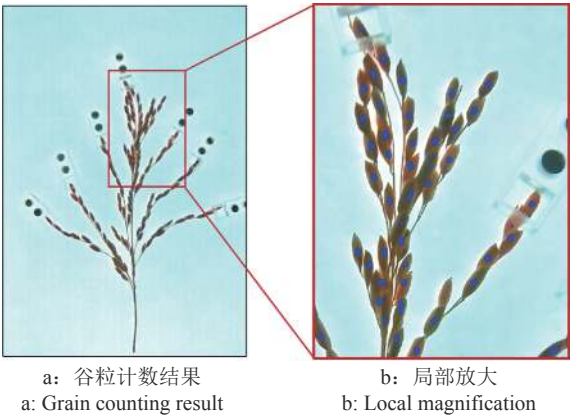


图 12 域偏移数据集下的谷粒计数结果及其局部放大图像
Fig. 12 Grain counting result and its local magnification image under domain shift data set

部分被遮挡严重的谷粒未能被标记, 该算法在新采集的图像中表现良好。

为不失一般性, 在新采集的 50 张图像中对谷粒计数算法进行准确率计算, 结果为 76.17%, 相比常规测试集下降 20.64%。如文献 [34] 所述, 大多深度学习模型在存在域偏移的数据集下性能将下跌 30% 以上, 相比较而言, 本文的算法在新采集的样本集上也有一定的泛化性能, 且未来通过域适应的方法, 性能有进一步提升的空间。

4 遮挡还原网络的训练和测试

4.1 遮挡还原网络的训练

由于训练的过程与谷粒计数的网络别无二致, 此处不再赘述。训练使用的超参数为初始学习率 0.01, 学习率衰减策略为带重启的余弦衰减策略, 批容量为 8, 总轮数为 300。

其中关于余弦衰减的参数为学习率衰减步数 1 000, 学习率衰减率 0.96, t_mul 参数 2.0, m_mul 参数 0.25, alpha 参数 0.000 1。

4.2 测试与评估

使用部分新图像测试模型的性能, 可视化结果如图 13 所示。从图 13 可以看出, 该遮挡还原网络能够比较好地还原谷粒缺失区域的大致形状和颜色, 但是网络恢复的区域相对比较模糊, 纹理等细

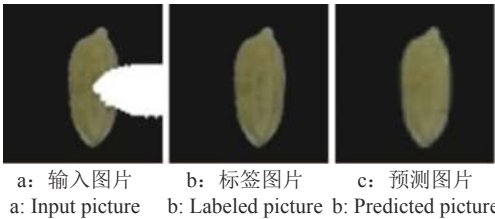


图 13 遮挡还原结果示例
Fig. 13 Example of occlusion restoration result

节也随之丢失, 细节方面相较于计算机视觉图像恢复领域中主流的 Conditional GAN^[35] 虽有一定的距离, 但对于几何特征能够比较好地还原, 已经可以满足考种机对集合形状测量的要求。

为准确量化模型的效果, 使用 100 张图像组成的测试集对该遮挡还原模型的性能进行评估。使用面积、周长、长度、宽度、颜色分数 5 个指标进行评估。其中颜色分数指标定义为下式:

$$\text{Color score} = \frac{G}{R+G+B}, \tag{8}$$

式中: R 、 G 、 B 分别表示红色、绿色、蓝色的像素点数。

本文提出的遮挡还原算法与文献 [31] 提出的遮挡还原算法的平均准确率对比如表 2 所示。从统计结果可以看出, 本文提出的算法在大部分指标中均超过文献 [31] 所提出的算法。对于考种机的使用需求而言, 该模型能够十分准确地恢复谷粒图像的几何特征, 且保持了模型的轻量级特点。

表 2 遮挡还原模型的平均准确率对比

Table 2 Comparison of average accuracy of occlusion restoring model %

指标 Index	文献[31]算法 Algorithm in reference [31]	本文算法 Algorithm in this paper
面积 Area	91.66	99.76
周长 Perimeter	96.64	97.90
长度 Length	99.92	99.55
宽度 Width	96.70	98.11
颜色分数 Color score	93.17	97.86

5 结论

本文针对自动化考种机设计了一种基于深度学习的轻量级 I2I 水稻穗上谷粒表型分析算法框架。针对水稻考种机的穗上谷粒原位计数任务和被遮挡谷粒还原任务, 本文将复杂的、难以实现的任务整体拆分为若干个简单的、容易实现的算法单元, 并将上述 2 种任务的核心阶段统一建模为 I2I 的映射问题。针对这 2 种 I2I 问题, 本文根据任务特点分别设计了样本集的制作方法, 改造了轻量级卷积神经网络 MobileNet V3, 使其适用于 I2I 问题, 选择了合理的优化器和超参数对其进行训练, 并使用 TF Lite 将其部署在考种机的树莓派 4B 开发板上。

为了验证模型的效果, 分别为原位计数和遮挡还原任务建立相应的测试集。经测试, 该算法在穗上谷粒原位计数任务中平均准确率为 96.81%, 明显

优于现有方法; 该算法模型体积仅为 1.7 MB, 能够在树莓派 4B 开发板上以每次前向推理耗时约 19 s 的速度运行, 具有轻量级和快速性的特点; 该模型于存在域偏移的样本集上准确率为 76.17%, 相比原测试集下降 20.64%, 低于大部分模型的下降比例, 说明该模型具有一定的泛化能力, 且未来通过域适应的方法准确率还有进一步提升的空间。该算法在遮挡谷粒的还原任务中, 谷粒恢复图像在面积、周长、长度、宽度和颜色分数 5 个指标的准确率均高于 97%, 明显优于现有方法, 说明本文提出的方法能够良好地还原被遮挡谷粒的几何形状。

综上所述, 本文提出的基于深度学习的 I2I 方法能够较好地解决穗上谷粒原位计数与被遮挡谷粒还原任务, 且具有轻量级的优点。

参考文献:

[1] ZHANG Q. Strategies for developing green super rice[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(42): 16402-16409.

[2] CHENG S H, CAO L Y, ZHUANG J Y, et al. Super hybrid rice breeding in China: Achievements and prospects[J]. *Journal of Integrative Plant Biology*, 2007, 49(6): 805-810.

[3] TANABATA T, SHIBAYA T, HORI K, et al. Smart-Grain: High-throughput phenotyping software for measuring seed shape through image analysis[J]. *Plant physiology*, 2012, 160(4): 1871-1880.

[4] KNECHT A C, CAMPBELL M T, CAPREZ A, et al. Image Harvest: An open-source platform for high-throughput plant image processing and analysis[J]. *Journal of Experimental Botany*, 2016, 67(11): 3587-3599.

[5] CHEN Y T, LIU X, YANG M H. Multi-instance object segmentation with occlusion handling[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015.

[6] DOLLÁR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: A benchmark[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009.

[7] 李秀智, 李家豪, 张祥银, 等. 基于深度学习的机器人最优抓取姿态检测方法[J]. *仪器仪表学报*, 2020, 41(5): 108-117.

[8] 甘海明, 岳学军, 洪添胜, 等. 基于深度学习的龙眼叶片叶绿素含量预测的高光谱反演模型[J]. *华南农业大学学报*, 2018, 39(3): 102-110.

[9] 史红栩, 李修华, 李民赞, 等. 基于深度学习的香蕉病害远程诊断系统[J]. *华南农业大学学报*, 2020, 41(6): 92-99.

[10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016

- [11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017.
- [12] LIU C, CHEN L C, SCHROFF F, et al. Auto-DeepLab: Hierarchical neural architecture search for semantic image segmentation[C]//2019 IEEE Conference on Computer Vision and Pattern Recognition. Long beach: IEEE, 2019.
- [13] ASHA C, NARASIMHADHAN A. Vehicle counting for traffic management system using YOLO and correlation filter[C]//2018 IEEE International Conference on Electronics, Computing and Communication Technologies. Bangalore: IEEE, 2018.
- [14] OLTEAN G, FLOREA C, ORGHIDAN R, et al. Towards real time vehicle counting using yolo-tiny and fast motion estimation[C]//2019 IEEE International Symposium for Design and Technology in Electronic Packaging. Cluj-Napoca: IEEE, 2019.
- [15] RAD R M, SAEEDI P, AU J, et al. Blastomere cell counting and centroid localization in microscopic images of human embryo[C]//2018 IEEE International Workshop on Multimedia Signal Processing. Vancouver: IEEE, 2018.
- [16] CHEN J, FAN Y, WANG T, et al. Automatic segmentation and counting of aphid nymphs on leaves using convolutional neural networks[J]. *Agronomy*, 2018, 8(8): 129. doi: 10.3390/agronomy8080129.
- [17] FALK T, MAI D, BENSCH R, et al. U-Net: Deep learning for cell counting, detection, and morphometry[J]. *Nature methods*, 2019, 16(1): 67-70.
- [18] CHAUDHURY S, ROY H. Can fully convolutional networks perform well for general image restoration problems?[C]//2017 IEEE International Conference on Machine Vision Applications. Nagoya: IEEE, 2017.
- [19] XIONG R, LIU G, QU Y, et al. Depth map inpainting using a fully convolutional network[C]//2019 IEEE International Conference on Robotics and Biomimetics. Dali: IEEE, 2019.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. arXiv: 1409.1556 (2014-09-04) [2022-02-12]. <https://doi.org/10.48550/arXiv.1409.1556>.
- [22] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. arXiv: 1704.04861 (2017-04-17) [2022-02-12]. <https://doi.org/10.48550/arXiv.1704.04861>.
- [23] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
- [24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE International Conference on Computer Vision. Seoul: IEEE, 2019.
- [25] YANG T J, HOWARD A, CHEN B, et al. Netadapt: Platform-aware neural network adaptation for mobile applications[EB/OL]. arXiv: 1804.03230 (2018-04-09) [2022-02-12]. <https://doi.org/10.48550/arXiv.1804.03230>.
- [26] VANSCHOREN J. Meta-learning: A survey[EB/OL]. arXiv: 1810.03548 (2018-10-08) [2022-02-12]. <https://doi.org/10.48550/arXiv.1810.03548>.
- [27] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. arXiv: 1706.05587 (2017-06-17) [2022-02-12]. <https://doi.org/10.48550/arXiv.1706.05587>.
- [28] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[EB/OL]. arXiv: 1710.05941 (2017-10-16) [2022-02-12]. <https://doi.org/10.48550/arXiv.1710.05941>.
- [29] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. arXiv: 1412.6980 (2014-12-22) [2022-02-12]. <https://doi.org/10.48550/arXiv.1412.6980>.
- [30] LOSHCHILOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts[EB/OL]. arXiv: 1608.03983 (2016-08-13) [2022-02-12]. <https://doi.org/10.48550/arXiv.1608.03983>.
- [31] 马志宏. 基于深度学习的水稻粒穗复杂性状图像分析方法[D]. 上海: 上海交通大学, 2018.
- [32] GONG L, LIN K, WANG T, et al. Image-based on-panicle rice [*Oryza sativa* L.] grain counting with a prior edge wavelet correction model[J]. *Agronomy*, 2018, 8(6): 91. doi: 10.3390/agronomy8060091.
- [33] AL-TAM F, ADAM H, DOS ANJOS A, et al. P-TRAP: A panicle trait phenotyping tool[J]. *BMC Plant Biology*, 2013, 13: 122. doi: 10.1186/1471-2229-13-122.
- [34] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C/OL]//Proceedings of the 32nd International Conference on Machine Learning. PMLR, 2015, 37: 1180-1189. [2022-02-08]. <http://proceedings.mlr.press/v37/ganin15.html>.
- [35] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks[EB/OL]. arXiv: 1611.07004 (2016-11-21) [2022-02-12]. <https://doi.org/10.48550/arXiv.1611.07004>.