

# 基于卷积神经网络的麦穗目标检测算法研究

王宇歌, 张涌, 黄林雄, 赵奉奎

(南京林业大学汽车与交通工程学院, 江苏 南京 210037)

✉515400100@qq.com; zyjs111@126.com; 1773361196@qq.com; zfk@njfu.edu.cn



**摘要:** 麦穗数量检测对于作物表型参数计算、产量预测和大田管理都具有重要的意义。为了解决人工计数工作量巨大且容易出错的问题, 提出了一种基于You Only Look Once (YOLO)的麦穗目标检测与计数方法。首先利用大量小麦图像对神经网络进行训练, 然后利用神经网络对小麦图像进行麦穗目标检测与计数, 最后对神经网络目标检测的准确率和召回率进行计算评估, 并通过分析检测结果验证其鲁棒性。分析结果显示, 所训练网络对麦穗检测的精确率为76.96%, 召回率为93.16%, 均值平均精度mean Average Precision (mAP)为89.52%。此外, 该模型可以检测不同生长时期的麦穗, 具有较高的鲁棒性。研究表明, 该方法对比其他麦穗计数方法准确高效, 可以实际应用小麦的产量估算上。

**关键词:** 目标检测; 产量预测; YOLO; 深度学习

**中图分类号:** TP391.4 **文献标识码:** A

## Research of Wheat Ear Target Detection based on Convolutional Neural Network

WANG Yuge, ZHANG Yong, HUANG Linxiong, ZHAO Fengkui

(College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China)

✉515400100@qq.com; zyjs111@126.com; 1773361196@qq.com; zfk@njfu.edu.cn

**Abstract:** Detecting the number of wheat ears is of great significance to the calculation of crop phenotypic parameters, yield prediction and field management. In order to solve the problem of heavy workload and error-prone manual counting, this paper proposes a wheat ear target detection and counting method based on You Only Look Once (YOLO). First, a large number of wheat images are used to train the deep neural network. Then, the neural network is used to detect and count wheat ears in the wheat images. Finally, the accuracy and recall rate of the neural network target detection are calculated and evaluated, and the robustness is verified by analyzing the detection results. The analysis results show that the trained network has an accuracy rate of 76.96% for wheat ear detection, a recall rate of 93.16%, and a mean Average Precision (mAP) of 89.52%. In addition, the model can detect wheat ears in different growth periods, and has high robustness. Studies have shown that this method is more accurate and efficient than other wheat ear counting methods, and can be applied to wheat yield estimation.

**Keywords:** target detection; output prediction; YOLO; deep learning

## 1 引言(Introduction)

小麦种植密度估算是小麦产量预测的重要手段, 也是小麦大田管理的重要依据。目前, 小麦产量预测方法有人工预测、年景预测<sup>[1]</sup>、基于遥感图像预测<sup>[2]</sup>和基于多元线性回归预

测<sup>[3]</sup>等。人工预测费时费力, 容易出错; 年景预测只适合地区大范围产量预测; 基于遥感图像预测准确率低; 基于多元线性回归预测受降水等变量影响较大, 准确率难以保证。相比之下, 视觉传感器可以获取丰富的纹理和颜色信息, 且成本

较低。近年来,机器视觉在麦穗检测研究中发挥着越来越大的作用。传统图像处理技术常使用移动窗法<sup>[4]</sup>或超像素分割法<sup>[5]</sup>采样子图像,从子图像中提取颜色或纹理特征,然后训练分类器,利用分类器识别麦穗,完成计数;或者通过图像处理方法突出麦穗,如将图像进行二值化处理,在去除粘连后识别麦穗<sup>[6]</sup>。

近年来,基于卷积神经网络(Convolutional Neural Network, CNN)的机器学习方法在图像处理领域取得了突破性进展。在物体识别、目标检测与目标分割等应用上,基于CNN的图像处理在技术表现上远好于传统方法<sup>[7]</sup>。因此,越来越多的研究采用目标检测算法进行作物穗检测。目标检测方法可分为Two-stage和One-stage两类。Two-stage方法先采用一定的算法生成候选框,然后通过卷积神经网络对样本进行分类,典型方法包括R-CNN<sup>[8]</sup>及Fast R-CNN<sup>[9]</sup>目标检测模型,例如HE等<sup>[10]</sup>提出的SPP-net、REN等<sup>[11]</sup>提出的Faster R-CNN等。Two-stage目标检测方法准确率高,但是速度相对较慢。One-stage方法直接回归目标的类别概率和边界框,运算速度相比于Two-stage方法有较大提高,常用方法包括Single Shot MultiBox Detector<sup>[12]</sup>等。YOLO系列目标检测框架是一种非常典型的One-stage方法,包括王宇宁等改进的YOLO<sup>[13]</sup>,李鹏飞等改进的YOLOv2<sup>[14]</sup>,REDMON等提出的YOLOv3<sup>[15]</sup>,以及BOCHKOVSKIY等<sup>[16]</sup>提出的YOLOv4。One-stage和Two-stage方法均已应用在作物穗的检测上,如HASAN等<sup>[17]</sup>利用R-CNN进行麦穗检测,MADEC等<sup>[18]</sup>利用Faster R-CNN进行麦穗检测,还有研究开始使用One-stage目标检测算法,如利用YOLO等对作物穗进行目标检测<sup>[19]</sup>,从而进行作物穗计数。由于麦穗目标尺寸小,相互遮挡严重,而且还有叶片的遮挡,目标识别困难,因此,仍然有必要对麦穗目标检测开展进一步的研究。

本文研究基于YOLOv3的麦穗目标检测算法,对目标检测效果进行评价,并对比基于深度学习的麦穗计数与人工麦穗计数两种方式,验证该方法的可行性。

## 2 麦穗小目标检测机制(Detection mechanism of small wheat ear target)

### 2.1 YOLOv3目标检测原理

YOLO系列目标检测算法是一种典型的One-stage目标检测算法。该系列算法可以直接回归得出目标可能所处的位

置,显著提高了目标检测的速度,并在实现实时性的同时确保检测精度。YOLO系列算法在实践中被广泛使用。YOLOv3是YOLO系列目标检测算法的典型代表,是目前实际工程应用领域最优秀的目标检测算法之一。

YOLOv3网络结构如图1所示。由于ResNet<sup>[10]</sup>可以有效加深网络深度并提高预测精度,YOLOv3使用了残差网络Residual,引入了多个ResNet模块,并大量使用了残差的跳层连接,减轻了深度神经网络在加深深度时梯度消失的问题。YOLOv3最终设计了一个具有53个卷积层的神经网络,命名为Darknet-53,用以提取图像特征。

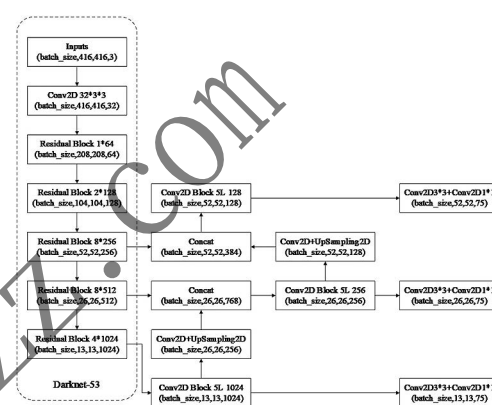


图1 YOLOv3网络结构

Fig.1 Network structure of YOLOv3

YOLOv3生成3个具有不同深度的特征图,来预测不同尺度的目标。3个特征层的shape分别为(52, 52, 256)、(26, 26, 512)、(13, 13, 1024),划分越精细的特征图所能检测的目标就越精细,有效提高了小目标的检测效果。YOLOv3通过将上采样后的特征图与较前层的特征图融合,例如将 $26 \times 26$ 的特征层上采样后与 $52 \times 52$ 的特征层融合,从而得到上采样的语义信息和前层特征图的细粒度信息,再进行卷积运算后对目标进行准确预测。

特征层预测结果shape分别为 $(N, 13, 13, 255)$ 、 $(N, 26, 26, 255)$ 、 $(N, 52, 52, 255)$ ,其中 $N$ 为图片数量。每个特征图网格单元上生成3个锚框,每个网格单元负责一个区域的检测,因此YOLOv3的边界框预测过程就是将特征图每个网格点加上它对应的坐标偏移量,相加的结果即为预测框的中心,然后再利用先验框和缩放尺度计算出预测框的长和宽,从而可以得到整个预测框的位置。

### 2.2 YOLOv3改进

神经网络训练之前,对YOLOv3的网络结构进行改进。

本文需要检测麦穗和背景,因此,只有一个类别,则输出特征层的最后一个维度为 $3 \times (5+1)=18$ ,其中“3”表示每个网格生成3个预测框;“5”则分别表示横纵坐标偏置、预测框长宽、置信度;“1”表示分类结果。因此,3个特征层的输出结果的尺寸改为 $(N, 13, 13, 18)$ 、 $(N, 26, 26, 18)$ 、 $(N, 52, 52, 18)$ ,其中 $N$ 为图片数量。

### 2.3 评价函数

目标检测通常使用准确率和召回率这两个指标来评估目标检测算法的性能。精确率(Precision)是指正确检测为正类的目标在所有判定为正类的目标中所占的比例;召回率(Recall)是指被正确检测为正类的目标在所有正类目标中所占的比例。精确率和召回率的计算公式分别如式(1)和式(2)所示。

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

式中, $TP$ 表示正确识别为正类的目标个数; $FP$ 表示错误识别为正类的目标个数; $FN$ 表示错误识别为负类的目标个数。

在实际的应用过程中,精确率和召回率两个指标无法同时满足要求,因此使用 $AP$ 值综合考虑两个指标,从而衡量网络的性能。 $AP$ 计算公式如式(3)所示。

$$AP = \sum_{k=1}^N Precision(k) \Delta Recall(k) \quad (3)$$

式中, $N$ 表示测试集中图片的数量; $Precision(k)$ 表示第 $k$ 张图片的 $Precision$ 值; $\Delta Recall(k)$ 表示从 $k-1$ 张图片到 $k$ 张图片 $Recall$ 值的变化情况。针对具有 $C$ 类目标的具体任务时,一般采用平均 $AP$ 值即 $mAP$ ,来衡量网络的性能,其计算如式(4)所示。

$$mAP = \frac{AP}{C} \quad (4)$$

## 3 实验分析(Analyses of experiment)

### 3.1 YOLOv3框架运行平台

麦穗目标检测程序运行环境为Ubuntu 18.04,采用keras开源框架,后端引擎为Tensorflow张量操作框架;硬件平台为Intel i9-9900K CPU,主频3.60 GHz,32 GB内存,显卡为NVIDIA RTX 2080Ti。

### 3.2 数据集制作

神经网络训练数据来源于文献[18]。然而该文献提供的数据中存在很多错误,标签位置与图像中麦穗目标不匹配,无法用于神经网络的训练。因此,本文校正了文献[18]提供的数据,查找标定错误的图像,并对图像进行翻转和旋转等操作,

使标签与麦穗目标匹配,最终获得有效图像222张。将其中80%的数据用于网络训练,其余20%的数据作为测试集用于测试网络的性能。在训练过程中,将前文所述80%数据中90%的数据作为训练集用于训练神经网络中的参数,其余10%的数据作为验证集用于验证模型性能的样本集合。由于采用了YOLOv3的预训练权重,设置batch\_size=4,经过150轮迭代后,损失函数趋于稳定。

### 3.3 实验结果分析

为了验证YOLOv3在麦穗识别中的性能,基于上文所述的测试集,通过计算目标检测算法典型评价指标来对YOLOv3算法进行评价。三种评价指标计算结果如表1所示。由于本文所识别的为单一目标,因此,本文中 $mAP$ 值即为麦穗这一类别的 $AP$ 值。从表1中的数据可以得出,YOLOv3在麦穗检测中表现出较优的性能,其精确率为76.96%,召回率为93.16%。由于精确率和召回率取决于阈值设定,为了更准确地评估模型目标检测性能,绘制精确率-召回率变化曲线如图2所示,曲线包围面积即 $mAP$ 为89.52%。

表1评价指标计算结果

Tab.1 Calculation results of assessment criteria

| 精确率    | 召回率    | $mAP$ 值 |
|--------|--------|---------|
| 76.96% | 93.16% | 89.52%  |

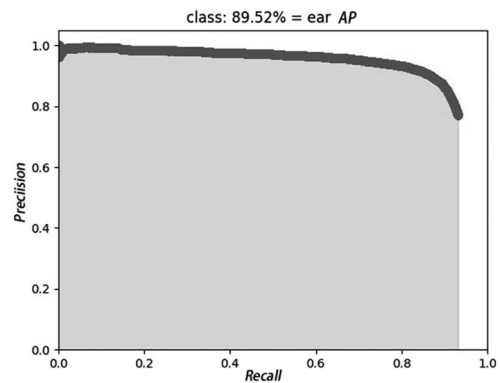


图2 P-R 曲线

Fig.2 P-R curve

为了验证YOLOv3在麦穗识别中具有较高的鲁棒性,本文选取了小麦在不同生长时期所拍摄的照片进行测试分析。测试结果如图3、图4和图5所示,测试图片按生长周期进行排序。各实例图片检测结果如表2所示,其中,检测成功的目标使用方框进行标注。在图3中,成功检测出138个目标,有8个目标未成功检测,准确率为94.52%;在图4中,成功检测

出100 个目标，有6 个目标未成功检测，准确率为94.34%；在图5中，成功检测出129 个目标，有1 个目标检测失败，准确率为99.23%。通过以上数据可知，该算法漏检少量麦穗，主要原因是麦穗目标比较小，麦穗与麦穗之间、叶片与麦穗之间相互遮挡。从上述图片的总体检测结果中可以看出，该算法在麦穗生长的各个时期均表现出较高的识别准确率，并且预测框的大小与位置也与麦穗非常匹配，可以认为该算法表现出较高的麦穗识别鲁棒性。

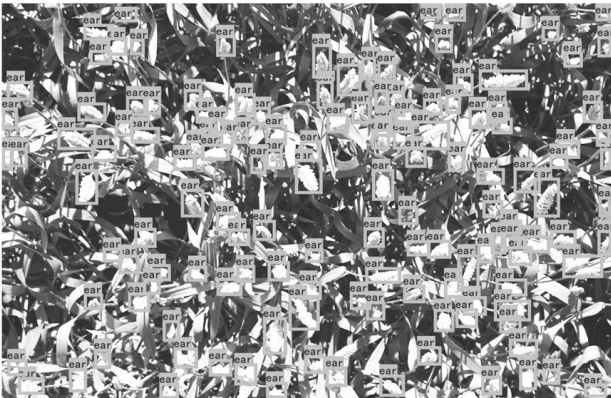


图3 麦穗检测结果1

Fig.3 Detection result of wheat ears 1

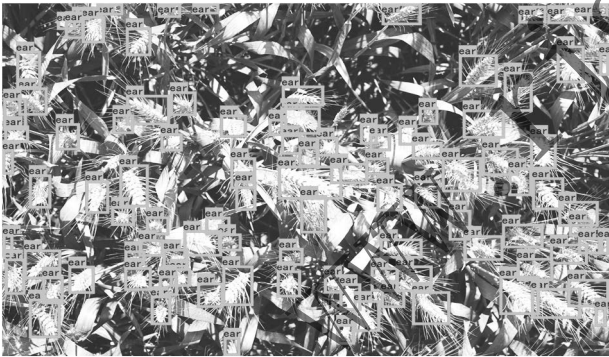


图4 麦穗检测结果2

Fig.4 Detection result of wheat ears 2



图5 麦穗检测结果3

Fig.5 Detection result of wheat ears 3

表2 各实例图片检测结果

| Tab.2 Detection results of all example pictures |      |          |
|-------------------------------------------------|------|----------|
| 检测出麦穗数量                                         | 真实数量 | 检测出麦穗百分比 |
| 138                                             | 146  | 94.52%   |
| 100                                             | 106  | 94.34%   |
| 129                                             | 130  | 99.23%   |

4 结论(Conclusion)

本文基于YOLOv3目标检测算法，对麦穗进行检测和计数。首先对开源的小麦麦穗数据集进行校正，并且适当划分训练集、验证集和测试集；其次，本文根据单类目标检测需求对YOLOv3进行改进，基于划分的数据集并结合YOLOv3的预训练权重对YOLOv3网络进行训练；最后，利用训练完成的神经网络对测试集数据进行测试。最终的实验结果证明，YOLOv3在麦穗检测这一任务中具有较高的准确率和*mAP*指标，且表现出较高的鲁棒性。同时也发现，由于麦穗目标较小，加之相互遮挡严重，仍存在部分麦穗漏检的情况，对麦穗的目标检测仍需要进一步的研究来提高准确率。

参考文献(References)

[1] 侯英雨,王建林,毛留喜,等.美国玉米和小麦产量动态预测遥感模型[J].生态学杂志,2009,28(10):2142-2146.

[2] 朱秀红,李秀珍,姚文军,等.基于SPSS的日照市小麦产量年景预测模型[J].中国农学通报,2010,026(02):295-297.

[3] 周永生,肖玉欢,黄润生.基于多元线性回归的广西粮食产量预测[J].南方农业学报,2011,42(09):1165-1167.

[4] 刘涛,孙成明,王力坚,等.基于图像处理技术的大田麦穗计数[J].农业机械学报,2014,45(2):282-290.

[5] 杜颖,蔡义承,谭昌伟,等.基于超像素分割的田间小麦穗数统计方法[J].中国农业科学,2019,52(01):21-33.

[6] 李毅念,杜世伟,姚敏,等.基于小麦群体图像的田间麦穗计数及产量预测方法[J].农业工程学报,2018,34(21):193-202.

[7] 翁杨,曾睿,吴陈铭,等.基于深度学习的农业植物表型研究综述[J].中国科学:生命科学,2019,49(06):698-716.

[8] 董静,耿达,郭迎港,等.室内环境下基于R-CNN的光照自适应物体检测[J].计算机工程与应用,2019,55(2):168-173.

[9] 莫宏伟,汪海波.基于Faster R-CNN的人体行为检测研究[J].智能系统学报,2018,13(6):967-973.

[10] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in Deep Convolutional Networks for visual recognition[J]. IEEE

- Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9):1904–1916.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137–1149.
- [12] 陈冰曲,邓涛.基于改进型SSD算法的目标车辆检测研究[J].重庆理工大学学报(自然科学),2019,33(1):58–63.
- [13] 王宇宁,庞智恒,袁德明.基于YOLO算法的车辆实时检测[J].武汉理工大学学报,2016,38(10):41–46.
- [14] 李鹏飞,刘瑶,李珣,等.YOLO9000模型的车辆多目标视频检测系统研究[J].计算机测量与控制,2019,27(8):21–24,29.
- [15] REDMON J, FARHADI A. YOLOv3: An incremental improvement[Z/OL]. [2021-05-26]. <https://arxiv.org/abs/1804.02767>.
- [16] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[Z/OL]. [2021-05-26]. <https://arxiv.org/abs/2004.10934>.
- [17] HASAN M M, CHOPIN J P, LAGA H, et al. Detection and analysis of wheat spikes using Convolutional Neural Networks[J]. Plant Methods, 2018, 14(1):100–113.
- [18] MADEC S, JIN X, LU H, et al. Ear density estimation from high resolution RGB imagery using deep learning technique[J]. Agricultural and Forest Meteorology, 2019(264):225–234.
- [19] 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.

## 作者简介:

王宇歌(1999–),女,本科生.研究领域:汽车电子.

张涌(1971–),男,博士,教授.研究领域:智能车底盘线控.

黄林雄(1996–),男,硕士生.研究领域:汽车电子.

赵奉奎(1986–),男,博士,讲师.研究领域:智能车环境感知技术.本文通讯作者.

(上接第25页)

占据海运网络的中心位置。但是亚洲各港口间竞争也比较激烈,新加坡港是世界海运网络最大的中转中心,使其综合中心性的位置排在第二。因此,随着海运网络的不断发展,我国还需要不断提高港口竞争力以加强在世界海运网络中的地位。

## 6 结论(Conclusion)

以港口为节点,以港口间航次连接作为边、航次数量作为权重构建海运复杂网络。对海运复杂网络进行分析,分别计算网络的度中心性、中间中心性、接近中心性,将三种中心性作为决策属性,港口作为决策方案。熵权-折中妥协法是一种主观和客观相结合的多属性决策方法,使用此方法对港口的中心地位进行排序,能够合理地反映出海运网络中各港口的中心地位,为港口的发展和政策的制定提供科学的依据。

根据排序结果可以看出各个港口的差异较大,排名前四的深圳港、新加坡港、香港港和上海港的折中值差距很小,竞争非常激烈,且都位于海运复杂网络的重要位置,海运复杂网络稳定性依赖于这些港口的稳定发展。使用将加权的复杂网络与EW-VIKOR多属性决策相结合的方法研究港口的中心地位,此方法为全球贸易发展中港口的建设提供了更多的参考。

## 参考文献(References)

- [1] 关晓光,李振福,陈雪,等.基于航次加权的世界集装箱港口复杂网络结构及港口中心性研究[J/OL].大连海事大学学

报,2020,46(03):11–21.

- [2] 刘婵娟,胡志华.“21世纪海上丝绸之路”海运网络空间格局及其复杂性研究[J].世界地理研究,2018,27(03):11–18.

- [3] 陈芙蓉,张建同,罗梅丰.“21世纪海上丝绸之路”沿线港口地位综合评价[J].运筹与管理,2020,29(01):157–164.

- [4] 伍静,张笛,万程鹏.基于复杂网络的海上丝绸之路沿线港口地位评价[J].交通信息与安全,2019,37(03):101–108.

- [5] JIANG L P, JIA Y, ZHANG C, et al. Analysis of topology and routing strategy of container shipping network on "Maritime Silk Road" [J]. Informatics and Systems, 2019(21):72–79.

- [6] ASADABADI A, MILLER-HOOKS E. Maritime port network resiliency and reliability through co-opetition[J]. Logistics and Transportation Review, 2020(137):1–17.

- [7] 杨忍,车乃夏,彭澎,等.“海上丝绸之路”沿线重要港口竞争力评价[J].地球信息科学学报,2018,20(05):623–631.

- [8] 陈芙蓉,张建同.“一带一路”倡议下我国沿海港口竞争力评价与比较研究[J/OL].工业工程与管理,2021,26(03):1–7.

- [9] 段雪妍,余思勤,王效俐,等.基于UTAGMS方法的内河港口竞争力排序[J].上海海事大学学报,2015,36(01):6–11.

## 作者简介:

刘苗苗(1993–),女,硕士生.研究领域:决策分析.

蒋艳(1974–),女,博士,副教授.研究领域:多目标决策.