Wheat Detection and Counting Solutions in Global Wheat Head Detection Dataset with Performance-Oriented Strategies

Lujia Zhong[1] Xinliang Liu[2] Ying Wang[1] Lihua Zheng[1] Hong Liu[3] Minjuan Wang[1,3*]
[1]Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education
China Agricultural University, Beijing 100083, China
zljdanceholic@gmail.com leowyztl@163.com {zhenglh, minjuan}@cau.edu.cn

[2]National Engineering Laboratory for Agri-product Quality Traceability
Beijing Technology and Business University, Beijing 100048, China
liuxinl@btbu.edu.cn

[3]State Key Laboratory of Virtual Reality Technology and Systems
Beihang University, Beijing 100191, China
lh64@buaa.edu.cn

1. Introduction

As one of the three major grain crops, wheat is widely planted all over the world. Its planting and production have a direct relationship with people's food security and health safety. However, after increasing rapidly for decades, the rate of increment in wheat yields has slowed down since the early 1990s [3, 5]. According to the Food and Agriculture Organization of United Nations, the whole world's demand for wheat is expected to reach 850 million tons by 2050 [1], which means the supply may fall short of demand in the future. Wheat production has become ever more challenging worldwide.

Recently, precision agriculture is one of the many strategies designed to improve crop management and maximize crop yields. Precision agriculture relies on monitoring and measuring the growth of crops in real-time [2], which means a huge amount of crop data collected to explore the growing status needs well organization and analysis. However, analyzing such a sheer amount of crop data is overly time-consuming and labor-intensive. Yield estimation is one of the most important tasks in precision agriculture. However, traditional wheat yield estimation requires agricultural experts to manually count the heads of wheat, which is extremely challenging, error-prone, and obviously not cost-effective at all.

2. Proposed Method

To handle this challenging task, making it efficient and reducing miscellaneous labor and time cost, we apply several deep learning models as a strong baseline and propose FiveHeadYOLO and GhostYOLO as alternative solutions. Compared with 3 state-of-the-art algorithms and YOLOv5 variants, extensive experiments on the GWHD dataset demonstrates the effectiveness of the proposed FiveHeadYOLO and GhostYOLO for wheat detection and counting tasks. To exploit better model performance on wheat counting, we explore 4 strategies: Hyperparameter Evolution, Optimal Anchors, Optimal Input Size, and Model Ensembling [4], and conclude the better training or detection method. In addition, we develop an Android app to deploy the wheat detection model, make it possible to implement detection and counting tasks on an edge device in real-time. This system has the potential to be improved for agricultural applications in the future.

Table 1. The performance comparison of the competing models in wheat head counting on 3373 images. The FPS represents the inference speed of the model on the NVIDIA Tesla T4 GPU.

Model	Precision	Recall	F1	mAP@.5	MAE	RMSE	FPS
YOLOv5s	0.964	0.933	0.948	0.966	2.9532	3.8728	84
YOLOv5m	0.963	0.949	0.956	0.972	2.2179	3.1360	42
YOLOv5l	0.965	0.936	0.951	0.969	2.7848	3.7107	23
FCOS	0.967	0.918	0.942	0.920	3.5864	4.6086	2
Faster R-CNN	0.962	0.950	0.956	0.921	2.2069	3.2039	8
GhostYOLO	0.972	0.897	0.933	0.948	4.5197	5.6572	90
FiveHeadYOLO	0.957	0.937	0.947	0.955	2.7821	3.7122	58

3. Results and Discussion

With all the proposed models and all competing models are well-trained on the training dataset, we evaluated their performance on the GWHD dataset, which includes a total

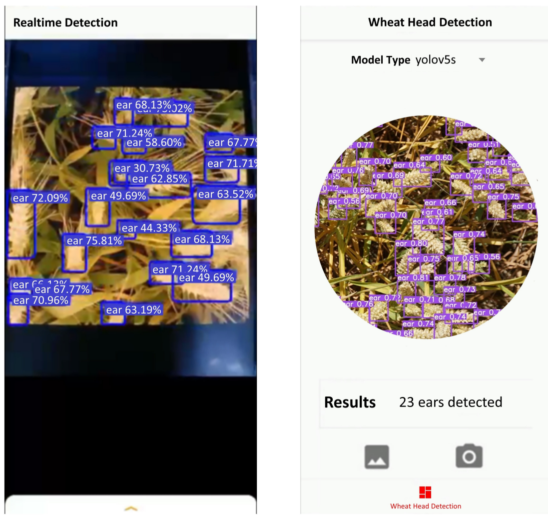

Figure 1. Software Main Interfaces. From left to right: image source selection and image detection and counting interface, wheat ear detection in real-time interface.

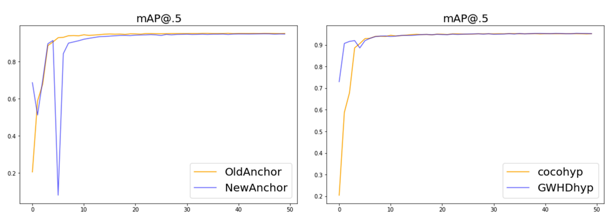

Figure 2. Training process visualization. From left to right: trained with anchors from COCO and GWHD, trained with hyperparameters from COCO and GWHD.

Table 2. The performance comparison of the YOLOv5 model with different input image sizes.

imgSize	precision	recall	mAP@.5	mAP@.5:.95	infer/ms	NMS/ms
256	0.7152	0.5202	0.5532	0.204	5.140	3.234
384	0.9043	0.8132	0.8842	0.4471	9.445	3.622
512	0.9401	0.8879	0.9457	0.5345	16.95	2.965
640	0.9454	0.9108	0.9594	0.562	32.67	3.146
768	0.9462	0.9193	0.9641	0.5723	44.73	3.589
896	0.9487	0.9217	0.966	0.5767	54.77	3.673
1024	0.9466	0.9231	0.9662	0.5749	78.89	4.138
1152	0.9417	0.9279	0.9657	0.5748	91.68	4.338
1280	0.9414	0.9257	0.9646	0.5704	106.6	4.870
1408	0.9382	0.9223	0.9628	0.5660	121.1	5.112
1536	0.9351	0.9192	0.9604	0.5592	160.0	5.595

number of 147793 ears. Table 1 shows the detection and counting performances of all models. As shown in Table 1, our proposed models compete with other methods by varying extents.

We perform several ablations to better understand the contributions of 4 training and inference strategies. The training process with strategies of hyperparameter evolution and optimal anchors is visualized as shown in Figure 2. It can be seen that both the hyperparameters combination obtained through hyperparameter evolution and suitable anchors from GWHD dataset significantly accelerate the convergence process of the model.

Table 3. The performance comparison Among single YOLOv5 models and different model combinations with NMS and mean ensemble.

Model	Precision	Recall	mAP@.5	mAP@.5:.95	Infer/ms	NMS/ms
mean_s_m	0.957	0.933	0.974	0.595	35.4	4.7
mean_s_l	0.950	0.932	0.972	0.588	52.4	4.6
mean_m_l	0.957	0.937	0.976	0.598	67.0	4.2
YOLOv5s	0.947	0.923	0.966	0.575	11.9	5.6
YOLOv5m	0.955	0.939	0.972	0.593	23.8	3.7
YOLOv5l	0.947	0.934	0.969	0.578	41.8	3.7
nms_s_m	0.945	0.939	0.972	0.59	35.1	6.0
nms_s_l	0.942	0.933	0.970	0.581	55.0	10.5
nms_m_l	0.947	0.942	0.973	0.592	65.0	6.5

After inference on the YOLOv5 model with different input image sizes, we compare the detection performance among input sizes. The results are shown in Table 2. Based on the results, we recommend that image input size should be 896 if optimal detection performance is required and input size of 768 is recommendable if less inference time is more important and keeping a good detection performance is still needed.

In our work, the combination strategies include NMS, max, and mean ensemble. After full training, we compare the ensemble results between single models and ensemble models. The mean strategy improves the model performance more than the max ensemble. The comparison between the mean and NMS ensemble with multiple models are shown in Table 3.

Besides, the developed application provides the function of real-time and image-based detection and counting by using the local camera, which are shown in Figure 1.

4. Conclusion

In this work, we propose GhostYOLO and FiveHeadYOLO and compare their performance with the sort-of-the-art object detection models, suggesting 2 practical solutions based on different scenarios. Notably, we explore 4 strategies for the training and inference and conclude the optimal method to improve performance on wheat ears detection and counting task. To bring the effective models to practical application, we quantize the model for faster inference with little performance loss and develop an Android app to implement real-time and image-based detection and counting. We hope the proposed models and the strategies can facilitate the research and development in wheat ears detection and counting in the agricultural field.

References

[1] Nikos Alexandratos and Jelle Bruinsma. World agriculture towards 2030/2050: the 2012 revision. 2012.

[2] Delfina Barabaschi, Alessandro Tondelli, Francesca Desiderio, Andrea Volante, Patrizia Vaccino, Giampiero Valè, and Luigi Cattivelli. Next generation breeding. *Plant Science*, 242:3–13, 2016.

[3] Nadine Brisson, Philippe Gate, David Gouache, Gilles Charmet, François-Xavier Oury, and Frédéric Huard. Why are wheat yields stagnating in europe? a comprehensive data analysis for france. *Field crops research*, 119(1):201–212, 2010.

[4] Glenn Jocher. Hyperparameter evolution. `https://github.com/ultralytics/yolov5/issues/607`.

[5] Bernhard Schauberger, Tamara Ben-Ari, David Makowski, Tomomichi Kato, Hiromi Kato, and Philippe Ciais. Yield trends, variability and stagnation analysis of major crops in france over more than a century. *Scientific reports*, 8(1):1–12, 2018.