

## WHEAT KERNEL SEGMENTATION AND COUNTING METHOD BASED ON IMPROVED HRNET

### 基于改进 HRNet 的小麦穗粒分割与计数方法

Aoqun HUANG<sup>1)</sup>, Junke ZHU<sup>\*1,2,3)</sup>, Zhicheng TANG<sup>1)</sup>,  
Shenke LI<sup>1)</sup>, Susu HUANG<sup>1)</sup>, Hongjian ZHAO<sup>1)</sup>, Yuxin ZHU<sup>3)</sup>

<sup>1)</sup>Shandong University of Technology, School of Agricultural Engineering and Food Science, Zibo / China;

<sup>2)</sup>Qilu Normal University, School of Life Sciences, Jinan / China;

<sup>3)</sup>Shandong Agricultural University, College of Agriculture, Tai'an / China;

Corresponding author: Junke ZHU; Tel: +8618905338833; Fax: +8605332786388; E-mail: zhujunke@126.com

DOI: <https://doi.org/10.35633/inmateh-78-78>

**Keywords:** wheat kernel counting; improved HRNet; semantic segmentation; crop phenotyping; precision agriculture

#### ABSTRACT

Aiming at the low accuracy of wheat kernel segmentation and counting caused by severe adhesion, blurred boundaries and complex field environment in wheat breeding tests, this study proposed an improved High-Resolution Network (HRNet) model optimized by integrating coordinate attention (CA), pyramid pooling module (PPM) and Lovasz-Softmax loss function. A pixel-level labeled wheat kernel segmentation dataset (3600 images, 5 national standard spike types) was constructed based on field samples from the Huang-Huai-Hai winter wheat region. The proposed model achieved a mean Intersection over Union (mIoU) of 88.3% on the laboratory test set, and a counting determination coefficient ( $R^2$ ) of 0.9135. For agricultural engineering application, the trained model was imported into a breeding phenotype analysis workstation, forming a complete application process of "field sampling → standardized image acquisition → batch model inference → automatic counting → phenotype data output". Field verification showed that the model realized high-throughput counting with an efficiency of 500 spikes per hour (38 times higher than manual counting), and maintained stable accuracy in complex field environments. This method can provide efficient technical support for high-throughput wheat phenotyping detection and precision agriculture.

#### 摘要

针对小麦育种试验中穗粒粘连严重、边界模糊、田间环境复杂导致的穗粒分割计数精度不足的问题，本研究融合坐标注意力机制（CA）、金字塔池化模块（PPM）与 Lovasz-Softmax 损失函数，提出一种改进的高分辨率网络（HRNet）模型。基于黄淮海冬麦区田间样本，构建了包含 3600 张图像、覆盖 5 种国标穗型的像素级标注小麦穗粒分割数据集。所提模型在实验室测试集上平均交并比（mIoU）达 88.3%，穗粒计数决定系数（ $R^2$ ）达 0.9135。面向农业工程应用，将训练完成的模型导入育种表型分析工作站，形成了“田间采样→标准化图像采集→批量模型推理→自动计数→表型数据输出”的完整应用流程。田间验证表明，该模型可实现高通量穗粒计数，作业效率达 500 穗 / 小时（较人工计数提升 38 倍），在复杂田间环境下精度保持稳定，可为小麦高通量表型检测与精准农业生产提供高效技术支撑。

#### INTRODUCTION

Wheat is the world's largest cereal crop and the second most important staple food crop in China, whose safe production is directly related to national food security, agricultural trade stability and rural industrial revitalization (Chen *et al.*, 2021). Wheat yield is mainly determined by three core components: spike number per unit area, kernel number per spike, and 1000-kernel weight (Du *et al.*, 2018). Among them, kernel number per spike is a key phenotypic index for wheat breeding, germplasm evaluation and precision cultivation management (Sun *et al.*, 2021). Therefore, high-throughput, high-precision and non-destructive automatic acquisition of kernel number per spike is the core premise to improve wheat breeding efficiency and support precision agricultural decision-making.

Automatic identification of wheat kernel number per spike is essentially a fine-grained instance segmentation and counting task, with core technical difficulties including small kernel size, severe adhesion between kernels, and blurred boundaries between kernels and glumes. Traditional kernel counting methods are mainly divided into manual counting and traditional machine vision-assisted counting (Teng *et al.*, 2025).

Manual counting has controllable accuracy but is inefficient and labor-intensive, while traditional image processing methods based on threshold segmentation and watershed algorithm are highly sensitive to kernel adhesion and illumination changes, with poor robustness (Li *et al.*, 2025). In recent years, image segmentation technology based on convolutional neural network has become the core research direction in crop phenomics, providing a new technical path for automatic detection of wheat spike traits (Li *et al.*, 2025; Teng *et al.*, 2025; Zhang, 2025).

Domestic and foreign scholars have carried out extensive research on crop fine-grained phenotype segmentation, providing technical references for wheat kernel segmentation. Chen and Shi (2026) optimized DeepLabv3+ to improve the segmentation accuracy of small targets with blurred edges. Duan *et al.* (2018) constructed PanicleNet for field rice panicle segmentation, achieving good generalization performance across different years. Yi *et al.* (2024) improved the segmentation accuracy of adhesive grape clusters by embedding coordinate attention and pyramid pooling module into U-Net. For wheat kernel counting, Geng *et al.* (2025) constructed WSG-P2PNet for high-precision wheat kernel positioning and counting in the field, and Zhang (2023) realized automatic kernel number counting with a mean absolute error of 4.82.

High-Resolution Network (HRNet), a high-resolution semantic segmentation network proposed in 2019, maintains high-resolution feature transmission through a parallel multi-resolution branch structure, with inherent advantages in crop fine-grained phenotype segmentation (Sun *et al.*, 2019). Related studies have confirmed that attention mechanism and multi-scale feature fusion can effectively improve HRNet performance in crop segmentation (Algarni *et al.*, 2025; Geng, 2024; Kim *et al.*, 2025; Yang *et al.*, 2025), which provides a theoretical basis for this study.

However, existing technologies still have three core bottlenecks restricting the engineering application of wheat kernel counting: first, few studies have optimized HRNet targeting wheat kernel segmentation, with insufficient accuracy for adhesive and weak-boundary kernels; second, existing improvement strategies are mostly migration of general computer vision modules, lacking targeted design for actual wheat breeding scenarios; third, most studies remain at laboratory algorithm verification, lacking low-threshold application tools for grassroots breeding bases.

To address these bottlenecks, this study takes main wheat varieties in the Huang-Huai-Hai winter wheat region as the research object, and carries out the following research: (1) constructing a pixel-level labeled wheat kernel segmentation dataset covering 5 national standard spike types; (2) proposing an improved HRNet-CPL model for wheat kernel fine-grained segmentation; (3) verifying model performance through systematic comparative experiments; (4) verifying the engineering application effect of the model in actual wheat breeding scenarios.

## MATERIAL AND METHODS

### Data Collection

Samples were collected from the breeding test field of Zibo Hefeng Seed Industry Technology Co., Ltd. in Linzi District, Zibo City, Shandong Province (36°57'25"N, 118°13'20"E). To avoid the interference of field edge effect (Ou *et al.*, 2019), a total of 300 wheat samples of 5 spike types (60 plants per type) were collected from non-edge row areas of 5 independent test fields. All wheat spikes were individually packaged in numbered protective bags, and the front and back sides of the spikes were photographed with an iQOO 11S smartphone (original resolution of 1279×1706). After shooting, the corresponding wheat spikes were threshed, and the kernel number per spike was recorded one by one. The data collection method is shown in Fig. 1.

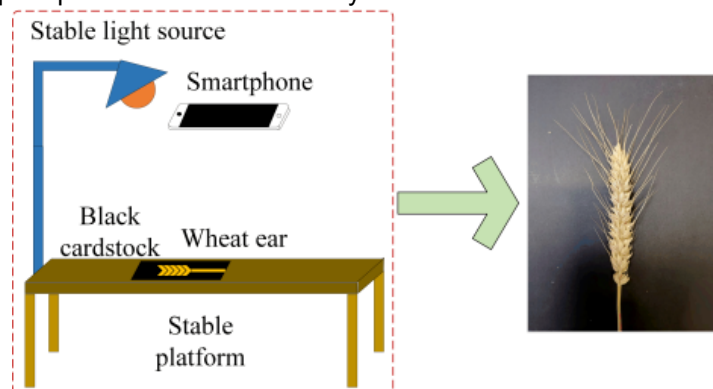


Fig. 1 - Schematic Diagram of Data Collection Methods

### Dataset Construction

The dataset construction process includes four core links: labeling, image normalization, data augmentation and dataset division.

(1) Data labeling: The open-source tool LabelMe was used to label the overall contour of wheat spikes and the fine contour of single kernels with irregular polygons, generating corresponding mask labels as the ground truth.

(2) Image normalization: The original 1279×1706 images were uniformly converted to 1280×1728 pixels through equal-proportion scaling + edge zero-padding, to provide standardized input for model training.

(3) Data augmentation: Brightness adjustment, horizontal and vertical flipping, Gaussian noise addition, and random rotation were performed on the original images, expanding the total dataset size to 3600 images.

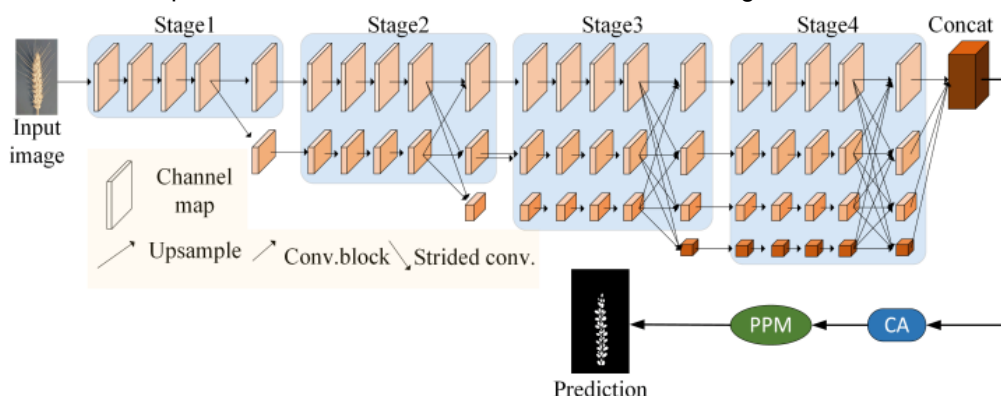
(4) Dataset division: The dataset was randomly divided into training set (2880 images) and test set (720 images) in a ratio of 8:2 for model training and performance verification.

### Baseline HRNet Model

HRNet is a high-resolution semantic segmentation network with a parallel multi-branch structure, which maintains high-resolution feature transmission throughout the network, and has inherent advantages in fine-grained crop phenotype segmentation (Sun *et al.*, 2019). This study selects HRNet-W32 as the baseline model to balance segmentation accuracy and engineering deployment efficiency.

### Improved HRNet-CPL Model

Aiming at the pain points of wheat kernel segmentation in complex field scenes, this study proposes an improved HRNet-CPL model, which is optimized from three aspects: feature enhancement, global context mining and loss function optimization. The network structure is shown in Fig. 2.



**Fig. 2 - Architecture of the Improved HRNet Model**

CA: Coordinate Attention module; PPM: Pyramid Pooling Module; Concat: feature concatenation; Upsample: upsampling operation; Conv.block: convolution block; Strided conv: strided convolution

1. Coordinate attention (CA) mechanism embedding: The CA mechanism is embedded after the multi-resolution branch feature fusion of the original HRNet. It encodes the position information of the channel dimension to accurately locate the spatial distribution of kernel targets, strengthens the weight of foreground effective features, and suppresses the interference of background irrelevant areas, helping the model focus on the effective features of kernel targets.

2. Pyramid pooling module (PPM) introduction: The PPM is introduced after the CA module to mine the global context information of features. It aggregates global features under different receptive fields through multi-scale pooling operations, making up for the lack of global context information after multi-resolution feature fusion, and improving the segmentation integrity of kernels with different spike shapes and adhesive overlapping areas, especially optimizing the depiction effect of kernel edge details.

3. Lovasz-Softmax loss function replacement: The original cross-entropy loss function is replaced with the Lovasz-Softmax loss function. It directly optimizes the Intersection over Union (IoU) index of the segmentation task, adjusts the gradient backpropagation weight of difficult samples, effectively alleviates the training deviation caused by category imbalance, and further improves the segmentation accuracy of kernel edges, overlapping areas and small targets.

**Coordinate Attention Mechanism**

Aiming at severe adhesion of adjacent wheat kernels and weakened boundary features between kernels and glumes in the yellow ripening period, the CA mechanism is embedded into the high-resolution feature branch of the improved HRNet. It completes joint modeling of channel and spatial features through two-way spatial coordinate encoding, retaining the spatial position information of the target and strengthening the feature expression of kernel boundaries and weak feature areas, with the detailed calculation process referring to *Hou et al. (2021)*.

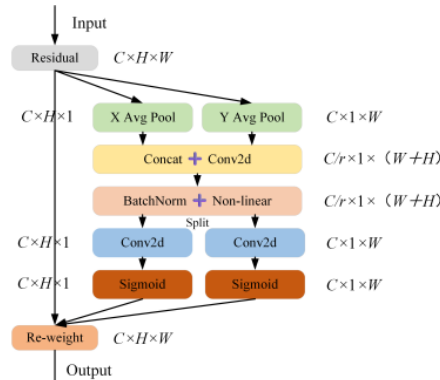


Fig. 3 - Structure of the CA Module

**Pyramid Pooling Module**

Aiming at the scale polymorphism of wheat kernels caused by variety differences and shooting distance changes, the PPM is connected after the CA attention enhanced features. It captures context information of different receptive fields through 4 parallel pooling branches, and aggregates local detail features and global layout features of kernels simultaneously, alleviating the segmentation accuracy fluctuation caused by kernel scale differences, with the specific calculation process referring to *Zhao et al. (2017)*.

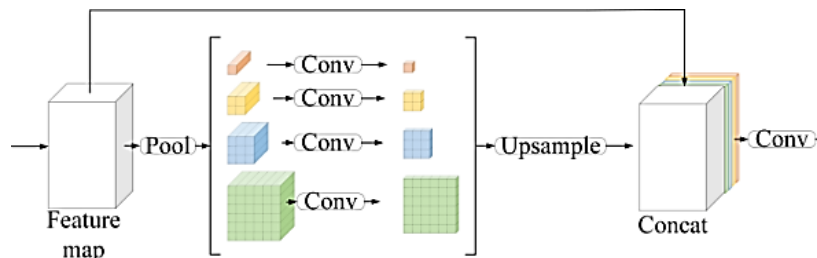


Fig. 4 - Schematic Diagram of the PPM

**Lovasz-Softmax Loss Function**

Aiming at the problem of unbalanced positive and negative samples and mismatching between the optimization objective of traditional cross-entropy loss and the core evaluation index IoU, the Lovasz-Softmax loss function is adopted as the main loss function of the model to replace the original cross-entropy loss. It directly optimizes the Intersection over Union (IoU) index of the segmentation task, adjusts the gradient backpropagation weight of difficult samples, effectively alleviates the training deviation caused by category imbalance, with the detailed principle referring to *Berman et al. (2018)*.

**EXPERIMENTAL SETUP**

**Environment Configuration**

The test operating system is Windows 11, with AMD EPYC 7713 CPU, NVIDIA GeForce RTX 3090 24GB GPU. The software environment includes CUDA 11.8, Cudnn 8.9.6, Python 3.9, and Pytorch 2.3.1 deep learning framework.

**Hyperparameter Setting**

The main training parameters are as follows: Batch size is 16, total training epochs is 300, stochastic gradient descent (SGD) optimizer is used with an initial learning rate of 0.01. The learning rate adopts a polynomial decay strategy with a power of 0.9, and the first 5 epochs are warm-up epochs. Weight decay is set to 0.0001, momentum is 0.9, and an early stopping strategy is adopted: training is terminated if the validation set mIoU does not improve for 30 consecutive epochs.

### **Evaluation Indicators**

In this study, two categories of evaluation indicators were set up to systematically verify the model performance: laboratory model performance indicators for algorithm effect verification, and field agricultural engineering application indicators for actual breeding scenario validation.

#### **Laboratory Model Performance Indicators**

This category of indicators is used to evaluate the segmentation and counting performance of the model in a controlled laboratory environment, including two dimensions: segmentation performance and kernel counting performance.

(1) Segmentation performance indicators: Precision (P, %), Recall (R, %), mean Intersection over Union (mIoU), and number of model parameters;

(2) Kernel counting performance indicators: Coefficient of Determination ( $R^2$ ), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Relative Error (MRE).

#### **Field Agricultural Engineering Application Indicators**

This category of indicators is used to evaluate the practical application effect of the trained model in actual wheat breeding scenarios, to verify the engineering application value of the model, including:

(1) Batch operation efficiency: The number of wheat spikes processed per hour by the model in batch mode, to characterize the high-throughput application performance of the model in actual breeding production;

(2) Single batch inference speed: The average processing time of 100 wheat spike images (front and back sides) by the trained model, to evaluate the processing efficiency of large-scale breeding samples;

(3) Field counting accuracy stability: The deviation of the model's counting results relative to the manual ground truth in the field test, to evaluate the robustness and stability of the model in complex actual breeding scenarios.

### **PRACTICAL APPLICATION VERIFICATION EXPERIMENT**

To verify the agricultural engineering application performance of the optimized HRNet-CPL model in actual wheat breeding scenarios, a systematic field application verification experiment was carried out in the breeding test field of Zibo Hefeng Seed Industry Technology Co., Ltd., which is consistent with the sample collection area of this study.

#### **Test Site and Materials**

The test site is located in the core wheat breeding area of the Huang-Huai-Hai winter wheat region, with flat terrain, uniform soil fertility, and consistent irrigation and fertilization management in the field. The test wheat varieties are 5 main planted varieties in the region, corresponding to the 5 national standard spike types (fusiform, club-shaped, rectangular, elliptical, conical) covered in the dataset. The test was carried out in the yellow ripening stage of wheat, which is the actual operation period of wheat breeding phenotype detection.

#### **Trained Model Preparation and Application Environment**

After the model training was completed and the optimal performance was verified through laboratory tests, the HRNet-CPL model weight file with the best mIoU on the test set was cured and exported, as the fixed model for agricultural engineering application. The application environment was built on a standard breeding phenotype analysis workstation, which is consistent with the model training hardware environment (Windows 11 system, NVIDIA GeForce RTX 3090 24GB GPU), and the trained model weight file was imported into the workstation to build a one-click batch processing pipeline for wheat kernel segmentation and counting, without secondary training or parameter adjustment for users.

#### **Field Application Test Design**

A randomized block test design was adopted, with 3 replicates set up. Each replicate contained 10 independent breeding plots, and 20 wheat spikes were randomly selected from the non-edge row area of each plot, with a total of 600 wheat spike samples for the test. For each sample, the following operations were carried out simultaneously to form a complete application process consistent with the actual breeding workflow:

1. Field sample collection: Wheat spikes were cut in the field, numbered and packaged individually, to simulate the actual sampling operation of breeding personnel in the wheat breeding season;

2. Standardized image acquisition: The iQOO 11S smartphone was used to capture front and back images of each wheat spike with a black cardstock as the uniform background, to ensure the consistency of image specifications and avoid extreme overexposure or underexposure, which is fully consistent with the image acquisition standard of the model training dataset;

3. Automatic counting by the trained model: The collected images were imported into the workstation in batches, and the trained HRNet-CPL model was called for automatic segmentation and counting, and the counting results of each sample were automatically exported to the phenotype data table;

4. Manual counting verification: The corresponding wheat spikes were individually threshed, and 2 experienced breeding personnel counted the kernels respectively, and the average value was taken as the ground truth of the field test, and the total manual counting time was recorded.

## RESULTS AND ANALYSIS

### Mainstream Model Comparison

To verify the segmentation performance of the improved model, it was compared with mainstream semantic segmentation models under the same training conditions (Table 1). The results showed that HRNet-CPL achieved the best segmentation performance, with a mIoU of 88.3%, which was 3.2 and 2.1 percentage points higher than the baseline HRNet and the optimal baseline model SegFormer, respectively. Meanwhile, the model only had a tiny parameter increment, with excellent lightweight characteristics for engineering deployment.

Table 1

Model	P/%	R/%	mIoU/%	Parameters(M)
U-Net	84.6	83.2	80.7	30.87
DeepLabV3	86.1	85.3	83.2	43.21
SegFormer	88.3	87.5	86.2	27.38
HRNet	87.4	86.9	85.1	29.56
HRNet-CPL	90.2	91.5	88.3	30.12

### Ablation Experiment

8 groups of ablation experiments were designed to verify the effect of each improved module (Table 2). The results showed that the three improved modules had a good synergistic complementary effect, and each module could significantly improve the segmentation accuracy of the model without excessive parameter increase.

Table 2

CA	PPM	Lovasz-Softmax	P/%	R/%	mIoU/%	Parameters/M
-	-	-	87.4	86.9	85.1	29.56
√	-	-	88.7	87.8	86.2	29.72
√	√	-	89.5	89.7	87.4	30.12
√	√	√	90.2	91.5	88.3	30.12
-	√	√	88.8	90.2	87.0	29.85
√	-	√	89.2	89.9	87.2	29.72
-	√	-	88.2	88.3	86.1	29.85
-	-	√	87.9	88.5	86.0	29.56

"√" indicates the use of this improvement, "-" indicates that the improvement is not used.

### Model Visualization Analysis

Wheat spike images of 5 different spike types were selected for visualization verification (Fig. 5). The results showed that compared with other mainstream models, HRNet-CPL could accurately restore the kernel contour, clearly distinguish the boundary of adhesive kernels, completely detect tiny kernels at both ends of the spike, and avoid background false detection.

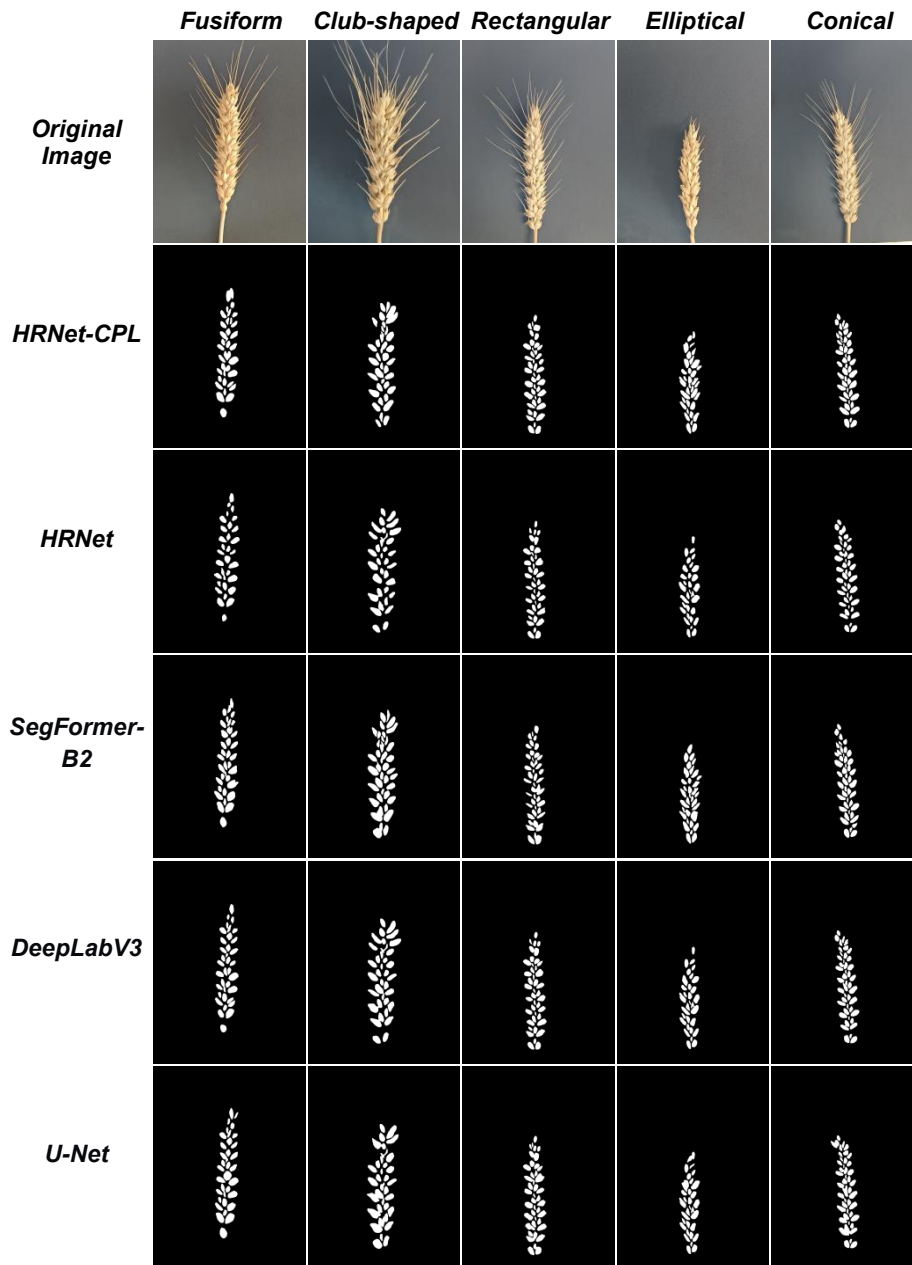


Fig. 5 - Segmentation Results of Wheat Grains by Different Models

**Kernel Counting Method**

For the segmentation prediction map, morphological erosion was first used to separate the adhesive target areas and suppress noise points. Then, the distance transformation algorithm was used to extract the central skeleton structure of the kernel area, followed by normalization and binarization processing. The opening operation was used to further filter out tiny noise points while retaining the integrity of the target area. Finally, the outer boundary of the target area was extracted based on the contour detection algorithm, and the effective contours were screened by setting the area threshold for counting, which was the final statistical value of the number of wheat kernels on one side of the spike. The counting process is shown in Fig. 6.

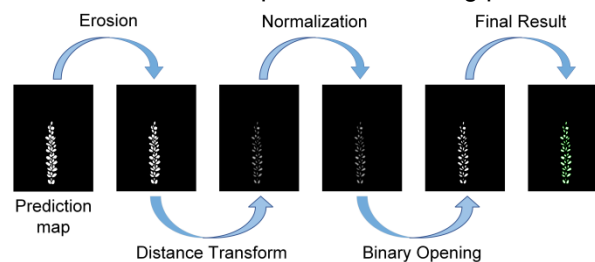


Fig. 6 - Schematic Diagram of the Wheat Grain Counting Method

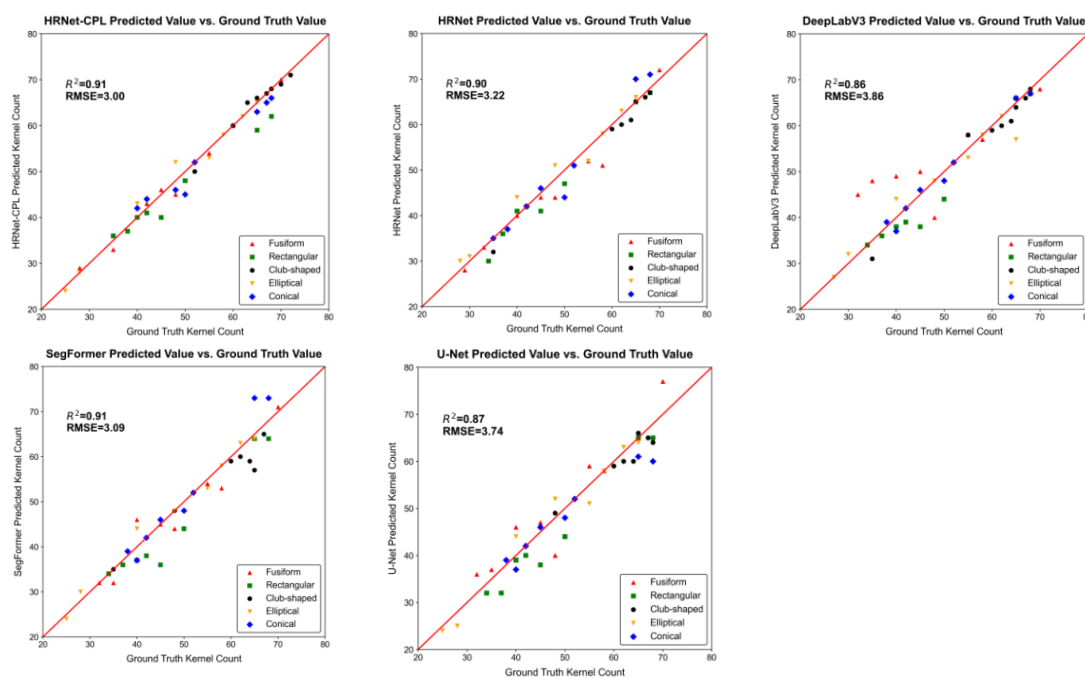
**Laboratory Counting Results and Analysis**

40 wheat spike samples not involved in model training were selected for counting performance verification (Table 3). The results showed that HRNet-CPL had the best overall counting performance, with a  $R^2$  of 0.9135 and the lowest RMSE of 2.9958, with a significantly better fitting degree with the ground truth than other models (Fig. 7).

**Table 3**

**Counting Performance of Different Models**

Model Name	$R^2$	RMSE	MAE	MRE
HRNet-CPL	0.9135	2.9958	2.425	0.0520
HRNet	0.8998	3.2249	2.500	0.0535
SegFormer	0.9080	3.0903	2.250	0.0500
DeepLabV3	0.8566	3.8568	3.225	0.0734
U-Net	0.8651	3.7417	3.000	0.0644



**Fig. 7 - Scatter Plot of Counting Accuracy for Different Models**

**Field Application Verification Results**

The field verification test based on 600 wheat spike samples showed that the imported trained HRNet-CPL model maintained high counting accuracy in actual breeding scenarios, with an overall counting  $R^2$  of 0.9012, RMSE of 3.1200, MAE of 2.5800 and MRE of 0.0562. The model's accuracy reduction was less than 2% compared with laboratory test results, showing strong robustness and generalization performance without secondary fine-tuning, and its counting MRE for 5 national standard spike types was controlled within 0.07 with good adaptability to different wheat varieties. The operation efficiency comparison (Table 4) showed that the model only took 12 min to process 100 wheat spikes, with an hourly processing capacity of 500 spikes (38 times higher than manual threshing and counting), which can significantly reduce breeders' labor intensity and shorten the phenotype detection cycle for large-scale germplasm processing in the wheat breeding season.

**Table 4**

**Comparison of field operation efficiency between the trained model counting and manual counting**

Counting method	Processing time per 100 spikes (min)	Hourly processing capacity (spikes)
Trained HRNet-CPL batch counting	12	500
Manual threshing and counting	462	13

## CONCLUSIONS

Aiming at the engineering pain points of low efficiency, high labor intensity and large error of manual wheat kernel counting in actual wheat breeding production, as well as the insufficient segmentation accuracy of existing models for adhesive and small kernels, this study carried out research on wheat kernel segmentation and counting method based on improved HRNet for agricultural engineering application.

First, a pixel-level labeled wheat kernel segmentation dataset (3600 images, 5 national standard spike types) was constructed based on field wheat samples from the Huang-Huai-Hai winter wheat region. An improved HRNet-CPL model was proposed based on HRNet-W32, with targeted optimizations including CA embedding, PPM introduction and Lovasz-Softmax loss replacement. The experimental results showed that the model significantly outperformed mainstream segmentation models such as U-Net, DeepLabV3 and SegFormer in segmentation accuracy, with only 0.56 M parameter increment.

For agricultural engineering application, the trained model can be directly imported into the breeding phenotype analysis workstation to realize batch automatic counting of wheat kernels. Field verification test showed that the model maintained high counting accuracy in actual breeding scenarios, with a batch operation efficiency 38 times higher than that of manual counting, and the counting error fully meets the requirements of wheat breeding phenotypic detection. In summary, the method proposed in this study can provide efficient and stable technical support for high-throughput wheat phenotyping detection in the Huang-Huai-Hai winter wheat region. Future work will focus on optimizing the model's adaptability to non-standard images and expanding its application to other grain crop phenotypic traits.

## ACKNOWLEDGEMENT

This research was supported by the Shandong Provincial Key R&D Program (Agricultural Improved Seed Project, 2022LZGCQY002). The authors would like to thank Zibo Hefeng Seed Industry Technology Co., Ltd. for providing the test field and wheat breeding samples for this study.

## REFERENCES

- [1] Algarni A., Naseer A., Alshehri M., AlQahtani Y., Alshahrani A., Park J. (2025). Hybrid HRNet-Swin Transformer: Multi-Scale Feature Fusion for Aerial Segmentation and Classification, *Computers, Materials & Continua*, vol.85, no.1, pp.1981-1998, Tech Science Press, Los Angeles/USA. DOI: <https://doi.org/10.32604/cmc.2025.038547>.
- [2] Berman M., Triki A. R., Blaschko M. B. (2018). The Lovasz-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.4413-4421, Salt Lake City, USA. DOI: <https://doi.org/10.1109/CVPR.2018.00464>.
- [3] Chen J. W., Li Q., Tan Q. X., Gui S. Q., Wang X., Yi F. J., Jiang D., Zhou J. (2021). Field wheat yield measurement combining lightweight wheat spike detection model and offline Android software development, *Transactions of the Chinese Society of Agricultural Engineering*, vol.37, no.19, pp.156-164, Chinese Society of Agricultural Engineering, Beijing, China. DOI: <https://doi.org/10.11975/j.issn.1002-6819.2021.19.018>. (in Chinese with English abstract)
- [4] Chen Y. N., Shi Z. (2026). Image semantic segmentation algorithm based on improved DeepLabv3+ model, *Journal of Heilongjiang University of Science and Technology*, vol.36, no.1, pp.163-168, Heilongjiang University of Science and Technology Press, Harbin, China. (in Chinese with English abstract)
- [5] Duan L. F., Xiong X., Liu Q., Yang W. N., Huang C. L. (2018). Field rice panicle segmentation based on deep fully convolutional network, *Transactions of the Chinese Society of Agricultural Engineering*, vol.34, no.12, pp.202-209, Chinese Society of Agricultural Engineering, Beijing, China. DOI: <https://doi.org/10.11975/j.issn.1002-6819.2018.12.025>. (in Chinese with English abstract)
- [6] Du S. W., Li Y. N., Yao M., Li L., Ding Q. S., He R. Y. (2018). Kernel counting method based on image segmentation of spikelets in wheat spikes, *Journal of Nanjing Agricultural University*, vol.41, no.4, pp.742-751, Nanjing Agricultural University Press, Nanjing, China. DOI: <https://doi.org/10.7685/jnau.201711037>. (in Chinese with English abstract)
- [7] Geng Q. (2024). Wheat kernel segmentation method based on improved HRNet with attention mechanism, *Agricultural Engineering International: CIGR Journal*, vol.26, no.1, pp.112-120, International Commission of Agricultural and Biosystems Engineering, Beijing/China.

- [8] Geng Q., Xu X., Ma X. M., Li L., Xu F., Gao B. B., Ma Y. T., Huang J. X., Yang J. Y., Yao X. C. (2025). WSG-P2PNet: A deep learning framework for counting and locating wheat spike grains in the open field environment, *Computers and Electronics in Agriculture*, vol.235, pp.110314, Elsevier, Amsterdam, Netherlands. DOI: <https://doi.org/10.1016/j.compag.2025.110314>.
- [9] Hou Q. B., Zhou D. Q., Feng J. S. (2021). Coordinate Attention for Efficient Mobile Network Design, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.13713-13722, Nashville, USA. DOI: <https://doi.org/10.1109/CVPR46437.2021.01350>.
- [10] Kim J. S., Park S. W., Yang H. S., Sim C. B., Jung S. H. (2025). CSA-HRNet: Channel and Spatial Attention High-Resolution Networks for Semantic Segmentation, *Journal of Signal Processing Systems*, vol.98, no.1, pp.6-16, Springer, New York, USA. DOI: <https://doi.org/10.1007/s11265-024-02090-9>.
- [11] Li L., Hassan M. A., Wang D. X., Wan G. L., Beegum S., Rasheed A., Xia X. C., He Y., Zhang Y., He Z. H., Liu J. D., Xiao Y. G. (2025). RGB imaging and computer vision-based approaches for identifying spike number loci for wheat, *Plant Phenomics*, vol.7, pp.100051, *Nanjing Agricultural University Press*, Nanjing, China. DOI: <https://doi.org/10.34133/plantphenomics.2024.100051>.
- [12] Ou X. Q., Ren X. J., Li X. H., Ouyang J. (2019). Effects of edge advantage and inner row performance of wheat varieties on plot yield, *Crops*, no.1, pp.97-102, Crops Society of China, Beijing, China. DOI: <https://doi.org/10.16035/j.issn.1001-7283.2019.01.016>. (in Chinese with English abstract)
- [13] Sun K., Xiao B., Liu D., Wang J. D. (2019). Deep high-resolution representation learning for human pose estimation, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.5693-5703, Long Beach, USA. DOI: <https://doi.org/10.1109/CVPR.2019.00584>.
- [14] Sun J., Yang K. F., Luo Y. Q., Shen J. F., Wu X. H., Qian L. (2021). Multi-scale perception wheat spike counting method based on UAV images, *Transactions of the Chinese Society of Agricultural Engineering*, vol.37, no.23, pp.136-144, Chinese Society of Agricultural Engineering, Beijing, China. DOI: <https://doi.org/10.11975/j.issn.1002-6819.2021.23.016>. (in Chinese with English abstract)
- [15] Teng G. F., Yan M. Y., Yao J. F., Li J. M. (2025). Field wheat spike recognition and counting based on improved YOLOv7 model, *Journal of Chinese Agricultural Mechanization*, vol.46, no.12, pp.77-85, Chinese Society of Agricultural Mechanization, Beijing, China. DOI: <https://doi.org/10.13733/j.jcam.issn.2095-5553.2025.12.011>. (in Chinese with English abstract)
- [16] Yang X. H., Li H. Q., Zhu W., Zuo Y. (2025). RSHRNet: Improved HRNet-based semantic segmentation for UAV rice seedling images in mechanical transplanting quality assessment, *Computers and Electronics in Agriculture*, vol.234, pp.110273, Elsevier, Amsterdam, Netherlands. DOI: <https://doi.org/10.1016/j.compag.2025.110273>.
- [17] Yi X. M., Zhou Y., Wu P., Wang G. Y., Mo L. F., Chola M., Fu X. Y., Qian P. X. (2024). U-Net with Coordinate Attention and VGGNet: A Grape Image Segmentation Algorithm Based on Fusion Pyramid Pooling and the Dual-Attention Mechanism, *Agronomy*, vol.14, no.5, pp.925, MDPI, Basel, Switzerland. DOI: <https://doi.org/10.3390/agronomy14050925>.
- [18] Zhang L. F. (2023). Wheat spike phenotypic parameter extraction method based on improved HRNet, *Journal of Intelligent & Fuzzy Systems*, vol.45, no.3, pp.4567-4578, IOS Press, Amsterdam, Netherlands. DOI: <https://doi.org/10.3233/JIFS-223456>.
- [19] Zhang Y. C. (2025). Advances in deep learning-based wheat spike detection and counting technology, *Smart Agricultural Technology*, vol.5, pp.100234, Elsevier, Amsterdam, Netherlands. DOI: <https://doi.org/10.1016/j.atech.2024.100234>.
- [20] Zhao H. S., Shi J. P., Qi X. J., Wang X. G., Jia J. Y. (2017). Pyramid Scene Parsing Network, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2881-2890, Honolulu, USA. DOI: <https://doi.org/10.1109/CVPR.2017.660>.