IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 April 2023, accepted 11 May 2023, date of publication 15 May 2023, date of current version 24 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3276450

RESEARCH ARTICLE

Impurity/Breakage Assessment of Vehicle-Mounted Dynamic Rice Grain Flow on Combine Harvester Based on Improved Deeplabv3+ and YOLOv4

QIAN ZHANG, JINPENG HU, LIZHANG XU, QIBING CAI, XUN YU, AND PENG LIU
School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China

Corresponding authors: Jinpeng Hu (hujinpeng2018@gmail.com) and Lizhang Xu (justxlz@ujs.edu.cn)

This work was supported in part by the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions under Grant PAPD-2018-87, in part by the Shandong Provincial Postdoctoral Innovation Project under Grant SDCX-ZG-202203051, and in part by the Key Research and Development Plan (Modern Agriculture) of Jiangsu Province under Grant BE2021339.

ABSTRACT It is difficult to extract small and dense objects with random state, such as grain and impurity, in image of vehicle-mounted dynamic rice grain flow on combine harvester. Therefore, this paper improves Deeplabv3+ by constructing MobileNetv2 in coding layer and adding ECA(Efficient Channel Attention) to Encoder and Decoder to improve extraction accuracy of high-dimensional features in images with a large number of objects with random state. In addition, the YOLOv4 is improved by using Mixup in preprocessing, constructing Mish in Neck and Head, adding ECA to Neck and Prediction of BackBone to improve training precision of small and dense objects and reducing effect of gradient disappearance. And the impurity/breakage rates are assessed based on relationship model between pixel area and quality, improved Deeplabv3+ and YOLOv4. The proposed method was verified by experiments with images acquired on intelligent combine harvester. Compared with existing Deeplabv3+, YOLOv4, U-NET, BP, the extraction accuracy by improved method increased by more than 4.01%. The average relative error and time of impurity/breakage assessment by proposed method were 7.69% and 1.56s. The proposed method can accurately and rapidly assess impurity/breakage rates for dynamic rice grain flow on combine harvester, and further realize closed-loop control of intelligent harvesting operation.

INDEX TERMS Breakage rate, grain flow, impurity rate, rice harvest, vehicle vision.

I. INTRODUCTION

China is a big country of rice production. It is vital to improve production and harvest level for ensuring national food security and enhance agricultural quality, efficiency and competitiveness [1]. With the development of rice harvesting mechanization and intelligent detection technology, the harvester in China are also gradually automated and accurate [2], [3]. The operation condition detecting which mainly contains impurity and breakage assessment is an important research direction of intelligent harvester [4], [5]. The impurity/breakage assessment is mainly located in the transfer and storage parts of the combine harvester after threshing and cleaning of crop. Different from the static detecting consistency on the tank for storing grains, the mode for transferring grains mainly includes scraper-type [6] and auger-type [7]. Compared with the scraper-type used mainly for large combine harvesters, the auger-type conveyer suitable for small and medium-sized combine harvesters can better meet the demand of agricultural machinery under moderate scale operation environment in our country. However, compared with the static sampling on the side wall of the scraper-type conveyer, the dynamic grain flow in the auger-type conveyer is faster and difficult for sampling. In addition, the objects

The associate editor coordinating the review of this manuscript and approving it for publication was Akansha Singh.

in the dynamic grain flow image are small, dense and random, which makes it more challenging to assess the impurity/breakage of vehicle-mounted dynamic rice grain flow on auger-type conveyer of combine harvester [8].

The main technologies available for grain impurity/breakage assessment include human visual, hyperspectral [9], terahertz [10], machine vision [11], etc. Compared with other methods, machine vision has the advantages of non-contact, fast speed and high accuracy, and is more suitable for real-time online vehicle-mounted dynamic grain flow impurity/breakage assessment. The image processing and neural network are the two main methods for assessing grain impurity/breakage in machine vision. The existing image processing includes k-means clustering, watershed, morphology, decision tree, Particle Swarm Optimization (PSO), Support Vector Machine (SVM) and so on. The k-means clustering [12], watershed [13] and morphology [14] are combined to segment impurities or broken grains in the image. It can obtain accurate results at a faster speed, but it is easy to be affected by light changes, dust interference, and image edge blurring. The decision tree can obtain feasible and good results for large data in a relatively short time, but it is difficult to improve the processing accuracy for features with strong correlation between broken and complete grains [15]. Both PSO and SVM are iterative classification algorithms. The PSO is based on iterative optimization similar to genetic algorithm [16], and the SVM is based on iterative training as a classifier. Considering that the initialization of PSO is a random solution, and the calculation amount of SVM is limited by the sample size [17], it is difficult to apply to the case of dense objects and high real-time requirements.

Compared with image processing, neural network is better in accuracy, adaptability, robustness and efficiency, and mainly includes shallow BP (Back Propagation) [18] and deep CNN (Convolutional Neural Network) [19] in impurity/breakage assessment. The BP has strong recognition and classification ability for grains and impurities, but it is easy to fall into local minimum, which makes it impossible to find the global optimal solution [20]. The CNN is a feedforward neural network with deep structure and convolution calculation, which is suitable for image processing. It improves the training efficiency of image data by local connection, weight sharing and pooling operation, and has better adaptability and robustness. At present, the CNNs used for impurity and grain extraction and classification are mainly RCNN [21], U-Net [22], YOLOv4 [23] and Deeplabv3 [24]. The one-step YOLOv4 can convert target extraction into end-to-end regression, and the Deeplabv3 based on DCNN and CRF can optimize object boundary, smooth noise and divide weights. Therefore, compared with two-step RCNN and full convolutional U-Net, the YOLOv4 and Deeplabv3 are more suitable for the impurity/breakage assessment of dynamic grain flow with dense objects and high real-time requirements.

However, for small and medium-sized combine harvesters with auger-type conveyer, the impurity/breakage detecting objects located in the grain conveyer are mobile. Compared with large grains, such as potato and corn, and regular shaped grains, such as rapeseed and soybean, the objects in the images of dynamic rice grain flow are small, dense and random. They all bring difficulties for rapid and accurate impurity/breakage assessment required by intelligent closed-loop control of online harvesting operations. Additionally, the existing Deeplabv3 with Xception without ability to change number of channels is easy to lose high-dimensional spatial features and has a unified attention mechanism for object detection [25]. The ReLU is easy to cause gradient disappearance, and Mosaic data enhancement for clipping and splicing images is difficult to improve the training effect of small and dense objects in existing YOLOv4 [26]. Therefore, it is difficult for existing Deeplabv3 and YOLOv4 to improve the accuracy and efficiency of impurity/breakage assessment of dynamic grain flow.

Based on the above research and problems, this paper improves the existing Deeplabv3+ and YOLOv4 based on coding layer, data enhancement optimization and attention mechanism, and proposes an impurity/breakage assessment method based on improved Deeplabv3+ and YOLOv4 for vehicle-mounted dynamic rice grain flow on combine harvester. It lays the foundation for on-line closed-loop control of intelligent operation of combine harvester based on machine vision [27].

II. MATERIALS AND METHODS

For the small and dense objects with random state, such as grain and impurity, in the image of vehicle-mounted dynamic rice grain flow on combine harvester, it is difficult to accurately extract by existing Deeplabv3+ and YOLOv4 with unified attention. Therefore, this paper improves existing Deeplabv3+ and YOLOv4 based oncoding layer optimization, data enhancement and attention mechanism, and proposes an impurity/breakage assessment method of vehicle-mounted dynamic rice grain flow on combine harvester.

A. IMPROVED DEEPLABV3+ AND YOLOv4 WITH ATTENTION MECHANISM OPTIMIZATION

1) EXISTING DEEPLABV3+ AND YOLOv4

a: EXISTING DEEPLABV3+

Based on Deeplabv3 structure, Deeplabv3+ network adds a decoding module for feature information refinement to improve the efficiency of image segmentation [28]. As shown in Fig. 1, the encoder adopts Xception model as the backbone network, and the depthwise separable convolution of different channels in the Xception model is adopted to extract multi-class feature information in image. After the 1×1 convolution processing in the spatial pyramid pooling module, three parallel 3×3 convolution processing with void rates of 6, 12 and 18, and global average pooling processing, a 1×1 convolution compression channel is used to extract high-level semantic features. In the decoder, the low-level semantic features extracted from the backbone network input

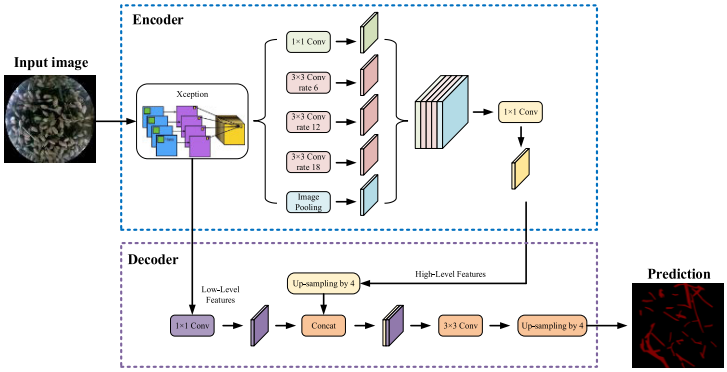


FIGURE 1. Existing Deeplabv3+.

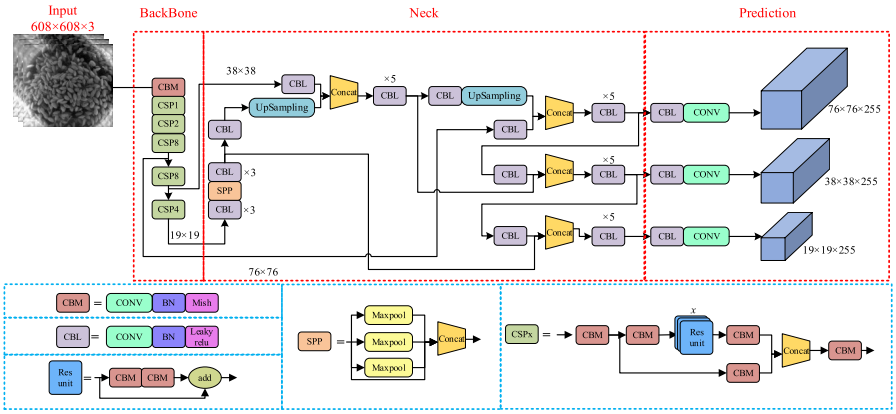


FIGURE 2. Existing YOLOv4.

layer after dimension reduction by 1 × 1 convolution are fused with the high-level semantic features obtained by up-sampling in encoder. The spatial information in the feature map is recovered based on 3 × 3 convolution, and the target boundary is refined by bilinear up-sampling to achieve semantic segmentation.

Compared with Deeplabv3, Deeplabv3 + replaces the original maximum pooling layers and dilated convolutions in Xception and decoder modules with the depthwise separable convolution layers. Compared with dilated convolutions, the depthwise convolution and pointwise convolution are combined for effectively reducing the computational complexity of model in the training process and constructing faster and stronger codec network.

b: EXISTING YOLOv4

Compared with YOLOv4-tiny with simplified network structure, the YOLOv4 is more suitable for small object detection. As shown in Fig. 2, the existing YOLOv4 mainly includes three parts: BackBone, Neck and Prediction [29]. Each part is mainly composed of five basic modules: CBM(Convolution, Batch Normalization and Mish), CBL(Convolution, Batch Normalization and Leaky-ReLU), Res unit(Residual units), CSPx(Center and scale prediction) and SPP(Spatial Pyramid Pooling). In the BackBone network, the CSPdarknet53 network structure including CSP and Darknet53 is used. The extracted feature maps are divided into X1 and X2 in CSPnet. X1 is directly input to the next layer and X2 is input to the next layer after Dense Block. In Neck structure, PANet and SPPNet are adopted. The rich spatial information of the bottom-up data flow and the rich semantic information of the top-down data flow are connected by PANet. The SPPNet applies maximum pooling with 5 × 5, 9 × 9, 13 × 13 sliding kernels for preserving space size. Then the feature maps from different kernel sizes are concatenated as output.

2) IMPROVED DEEPLABV3+ WITH CODING LAYER AND ATTENTION MECHANISM OPTIMIZATION

For existing Deeplabv3+ used for image segmentation, because the Depthwise Convolution of backbone feature extraction network Xception has no ability to change the number of channels, feature extraction is limited by the number of channels on the upper layer of Depthwise Convolution. It is easy to lose high-dimensional spatial features and affect segmentation accuracy. Additionally, for the small and dense objects with random state in the image of vehicle-mounted dynamic rice grain flow on combine harvester, it is more difficult to improve the accuracy and real-time of object segmentation by existing Deeplabv3+ with unified attention. Therefore, this paper improves the existing Deeplabv3+ based on coding layer and attention mechanism optimization, mainly as follows:

The backbone network is replaced. The Xception model of coding layer is replaced by the MobileNetv2 model.

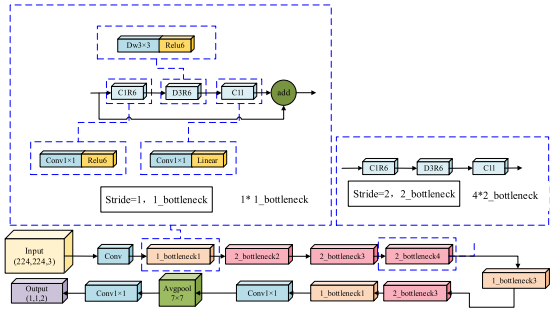


FIGURE 3. Network structure of MobileNetv2.

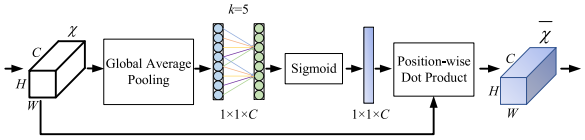


FIGURE 4. Efficient channel attention.

Compared with Xception, MobileNetv2 contains Inverted residual networks and Linear Bottleneck units [30], as shown in Fig. 3. When stride =1, the input image expands the number of channels based on 1 × 1 convolution (activation function is ReLU6), and based on 3 × 3 depthwise convolution (activation function is ReLU6). Then based on 1 × 1 pointwise convolution, the number of channels is compressed (activation function is linear). Finally, the two channels are added based on shortcut. When stride = 2, there is no shortcut because the feature map sizes of input and output are different. The number of channels is expanded by adding 1 × 1 convolution in depthwise convolution to reduce the limit of channel number for feature extraction, so as to increase the dimension of feature extraction and improve the segmentation accuracy.

ECA (Efficient Channel Attention) [31] is added in encoder and decoder of network, respectively. As shown in Fig. 4, the one-to-one weight update of the original channel is retained without reducing the channel dimension by designing an efficient channel attention mechanism for obtain a stability effect of one-to-one weight. In addition, local cross-channel interaction is used to reduce the increase of computational cost caused by non dimension reduction of channels. The coverage of cross-learning is determined by constructed mechanism of automatically adjusting for kernel size in ECA to improve the concentration of Deeplabv3+ model.

The improved Deeplabv3+ with coding layer and attention mechanism optimization is shown in Fig. 5. The MobileNetv2 is adopted as the backbone network in encoder part to extract two effective feature layers of high and low dimensions. After ECA and a large number of dilated convolutions in encoder, the high-dimensional feature layer with receptive field enlarged is sent to decoder. The low-dimensional feature layer after ECA module in decoder is fused with the high-dimensional feature layer obtained by encoder for obtaining semantic segmentation result.

3) IMPROVED YOLOv4 WITH DATA ENHANCEMENT AND ATTENTION MECHANISM OPTIMIZATION

The existing YOLOv4 uses Mosaic data enhancement which randomly clipping and splicing four images. For the vehicle-mounted dynamic rice grain flow image with small and dense objects on combine harvester, the clipping and splicing does not significantly improve the network training effect. In addition, the ReLU is easy to cause gradient disappearance and unified attention. All above make the accuracy and real-time performance of the existing YOLOv4 difficult to meet the requirements of impurity/breakage assessment for vehicle-mounted dynamic rice grain flow on combine harvester. Therefore, the existing YOLOv4 is improved based on data enhancement and attention mechanism optimization in this paper, as shown in Fig. 6, mainly as follows:

Mosaic is replaced by Mixup for data enhancement. For images with small and intensive objects with strong repeatability, Mixup [32] is used to preprocess the training data of YOLOv4 network in this paper, and new samples and labels are created by linear interpolation. Two images and labels are randomly selected for fusion according to a certain proportion, as shown in Fig. 7. Image enhancement is carried out based on (1), where, B_{x1} and B_{x2} represent batch samples respectively, B_{y1} and B_{y2} represent the label of B_{x1} and B_{x2}, M_x and M_y represent the batch samples and labels after mixing respectively, $\lambda = Beta(\alpha, \beta)$ is the mixing coefficient calculated by the Beta distribution of parameters α and β.

$$\begin{cases} M_x = \lambda B_{x1} + (1-\lambda)B_{x2} \\ M_y = \lambda B_{y1} + (1-\lambda)B_{y2} \end{cases} \tag{1}$$

For ReLU function, when the input is negative, the gradient is 0, which may cause the problem of gradient disappearing and weakening. Therefore, the ReLU activation functions in Neck middle layer and Head output layer are replaced by Mish activation function. As shown in Fig. 6, CBL module is replaced by CBM. The self-regularized non-monotonic neural activation function Mish is shown in (2). The smooth activation function allows better information to penetrate into the neural network, resulting in better accuracy and generalization.

$$f(x) = x \tanh(\ln(1 + e^x)) \tag{2}$$

In BackBone, The ECA modules are added in three-layer networks before output to Neck and Prediction. As shown in Fig. 4, the network aggregates features of different detection layers at different backbone layers to improve its concentration.

As shown in Fig. 6, improved YOLOv4 with data enhancement and attention mechanism optimization still includes BackBone, Neck and Prediction. In BackBone, which mainly extracts image features, the middle layer of Neck with attention mechanism not only fuses features of various scales, but also focuses on the main features and ignores the inessential features. In Prediction, a bottom-up feature pyramid with two PAN (Path Aggregation Network) structures is added behind

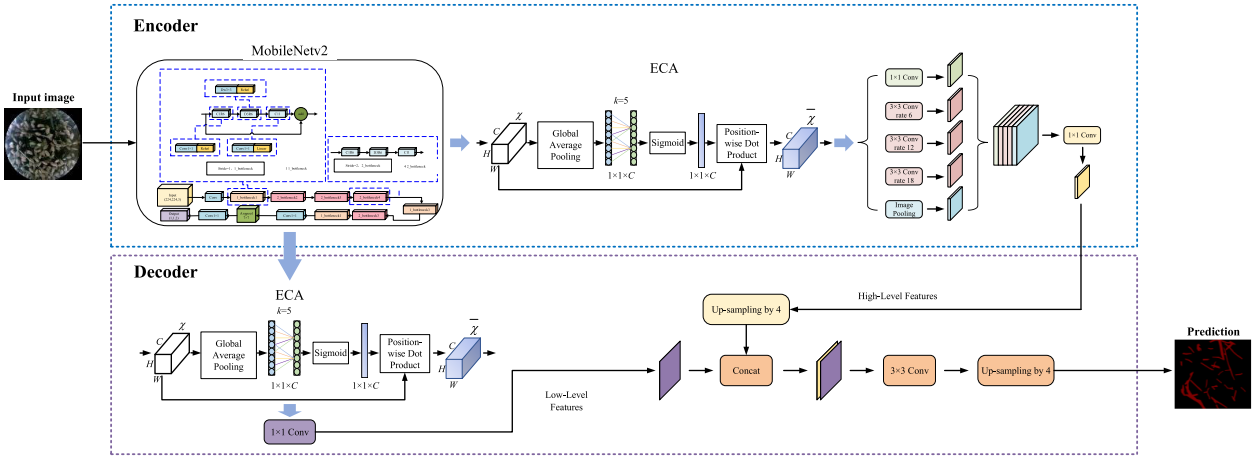


FIGURE 5. Improved Deeplabv3+ with coding layer and attention mechanism optimization.

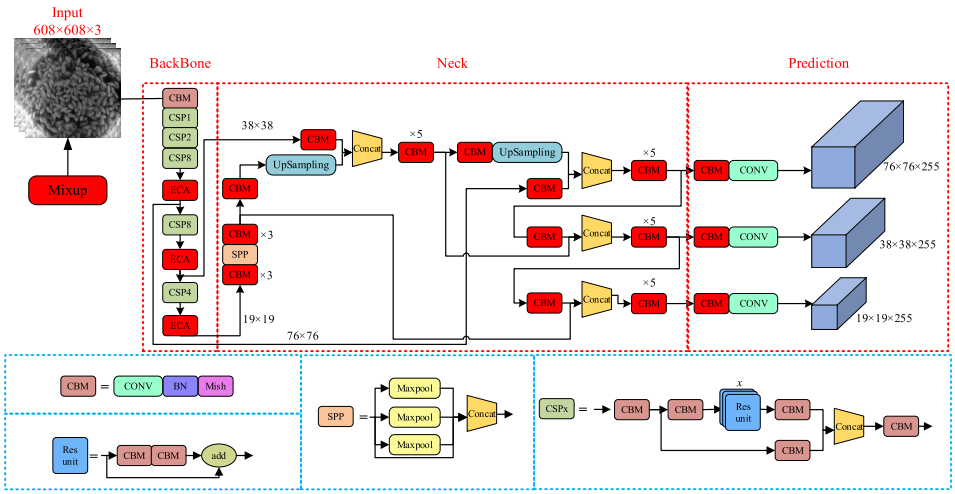


FIGURE 6. Improved YOLOv4 with data enhancement and attention mechanism optimization.

the FPN (Feature Pyramid Network) layer. The FPN layer transfers strong semantic features from top to bottom, and the feature pyramid transfers strong localization features from bottom to top to fuse features of different detection layers in different backbone layers.

B. DISCRETE ELEMENT ANALYSIS OF RICE GRAIN MOVEMENT AND DETECTING POINT LOCATING FOR IMPURITY/BREAKAGE ASSESSMENT ON CONVEYING AUGER OF COMBINE HARVESTER

The impurity/breakage assessment of rice grain for operation status monitoring of combine harvester is located in the transport and storage parts after threshing and cleaning. Compared with stacked and shielded grain tank, the grain conveyer with relatively separated grains at each time is more suitable for locating the detecting point of impurity/breakage. Therefore, the modeling and discrete element analysis of rice grain movement on grain conveyer of harvester are carried out to locate optimal detecting point of impurity/breakage, which lays a foundation for the impurity/breakage assessment of vehicle-mounted dynamic rice grain flow on combine harvester.

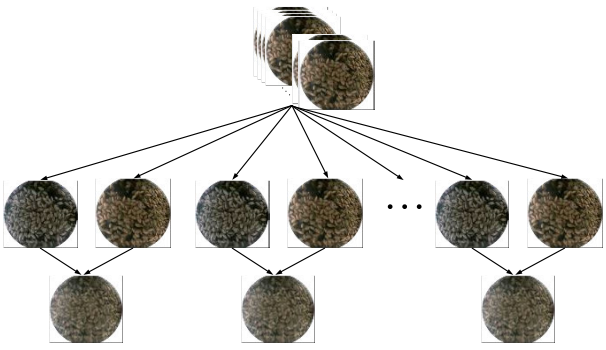


FIGURE 7. Mixup for data enhancement.

1) MODELING AND DISCRETE ELEMENT ANALYSIS OF RICE GRAIN MOVEMENT ON GRAIN CONVEYER OF COMBINE HARVESTER

Vertical grain conveyer is mainly composed of auger covering, blade of conveying auger, shaft, grain outlet, etc., as shown in Fig. 8. During the field operation, the blade drove the grains below the auger to the grain outlet. With the high speed rotation of blade, the grains quickly left the grain conveyer and are thrown into the grain tank with fan

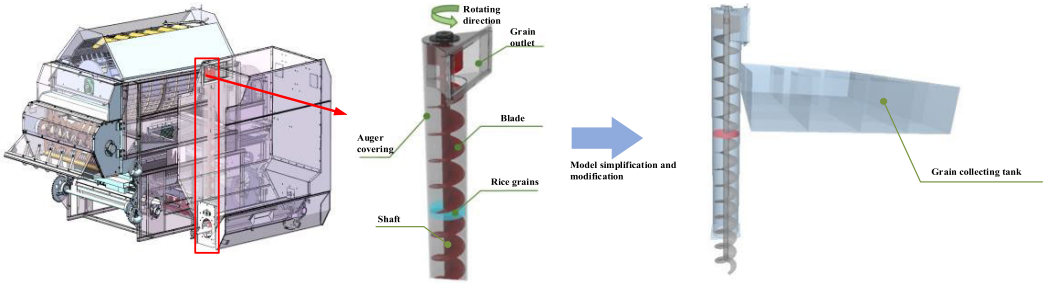


FIGURE 8. Grain conveyer of combine harvester.

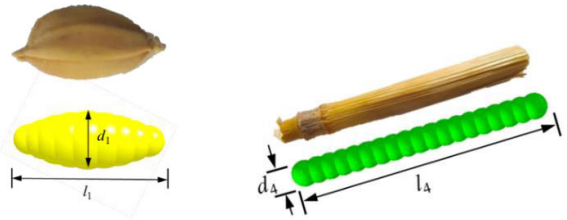


FIGURE 9. Rice grain and impurity particle models. (a) Rice grain, (b) Impurity.

TABLE 1. Mechanical parameters and contact coefficient values of grains, impurities and auger covering.

Material	Rice grain	Impurity	Auger covering
Poisson Ratio/MPa	0.25	0.4	0.3
Shear Modulus/Mpa	181	48	8×10^4
Density/kg·m-3	1320	227	7.8×10^3
Collision Recovery Coefficient	0.17	0.12	0.35
Static Friction Coefficient	0.56	0.72	0.70
Dynamic Friction Coefficient	0.10	0.01	0.01

shape. The grain movement speed during transportation can reach 1.2m/s [33]. It can be seen from Fig. 8 that the grain conveyer is a cylinder, and only has a plane at the side of grain outlet. In addition, as the grains are driven by blade in auger covering, if the grain status is detected at the side of auger covering, it will be easily interfered by blade. Therefore, the impurity/breakage detecting point is located on the side of grain outlet in this paper. The motion of rice particles in grain conveyer is simulated and the motion law of rice grains and impurities at the grain outlet is analyzed by EDEM [34] for locating optimal detecting point of impurity/breakage.

In order to reduce redundancy and improve simulation efficiency, the auger covering model is simplified by Merge function in Geometry. The transmission parts and power devices at the bottom of auger covering are removed, and the grain collecting tank is added, which is used to collect grains and impurities from the grain outlet. Based on the previous research results of our research group on rice discrete element simulation [35], [36], mechanical parameters and contact coefficient values of grains, impurities and auger covering are set, as shown in Table 1. The non-sliding contact model is set between grains, grains and impurities, grains and auger, impurities and auger. The rice particles with 7.62 mm length l1, 2.98 mm height h1 and 2.98 mm maximum diameter d1 are added. The cylindrical impurity particles with 32 mm length l4 and 3 mm diameter d4, as shown in Fig. 9.

A plane in the middle of auger covering is added and its property is set to Virtual. A particle factory on the plane is created to produce grain and impurity particles. The total yield and unit time yield of grain particles are 1×10^5 and 2×10^4, respectively. The total yield and unit time yield of impurity particles are 2×10^4 and 4×10^3, respectively. A circular rotation with speed of 974 r/min is added to shaft of grain conveyer. The time step is set as 1×10^{-6} s, the solution time is 5 s, the grid size is 6 mm, and the number of grids is 29,028,996. Based on the above model and optimized parameters, the discrete element analysis of rice grain movement on grain conveyer of harvester is carried out.

2) DETECTING POINT LOCATING FOR IMPURITY/BREAKAGE ASSESSMENT

The EDEM motion simulation based on rice grains shows that the position A outlined in the Fig. 10 is always in the state of grain and impurity flow and full in the process of grain transport. Considering the requirements of sampling coverage and single sampling amount of grains in grain conveyer for impurity/breakage assessment, the position A is used as the optimal detecting point for impurity/breakage assessment. In addition, considering the space limitation and operation smoothness of grain conveyer, the impurity/breakage assessment module is constructed on the outside of grain outlet corresponding to position A. The detection surface corresponding to position A is modified to be a high-transparence glass for realizing high-frequency sampling and assessment of grain in the grain conveyer by external detection module.

C. IMPURITY/BREAKAGE ASSESSMENT OF VEHICLE-MOUNTED DYNAMIC RICE GRAIN FLOW ON COMBINE HARVESTER

At the outlet of grain conveyer, the flowing grains are dense, and its state is random. The image of grain impacting glass of detecting point has many small objects and high-frequency spatial features. It is difficult to achieve high-precision

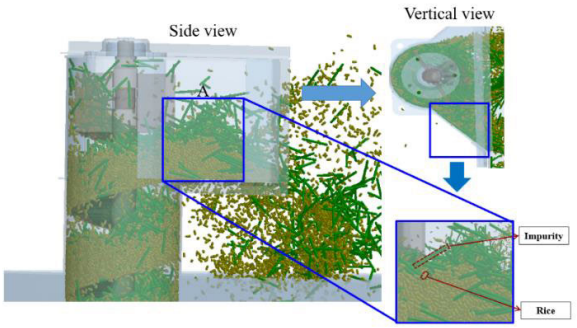


FIGURE 10. Optimal detecting point for impurity/breakage assessment.

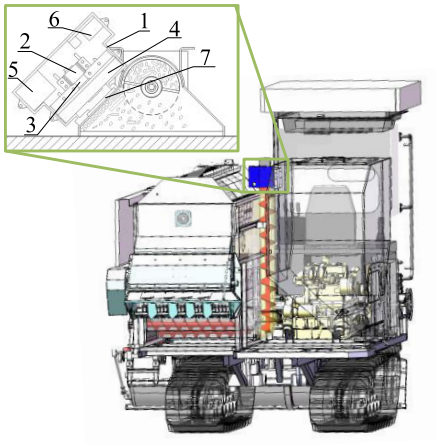


FIGURE 11. Image acquisition module of dynamic rice grain flow. 1 box, 2 camera, 3 lens, 4 light source, 5 light source controller, 6 embedded processor, 7 high-transparence glass.

real-time impurity/breakage assessment based on the existing Deeplabv3+ and YOLOv4. Therefore, an image acquisition module of vehicle-mounted dynamic rice grain flow is constructed in this paper. And the improved Deeplabv3+ and YOLOv4 with attention mechanism optimization is adopted to extract impurities, broken and complete grains in the dynamic grain flow. The impurity and breakage rates are assessed based on the relationship model between pixel area and quality.

1) IMAGE ACQUISITION AND PREPROCESSING OF VEHICLE-MOUNTED DYNAMIC RICE GRAIN FLOW

a: IMAGE ACQUISITION OF DYNAMIC RICE GRAIN FLOW

The image acquisition module of dynamic rice grain flow is installed at optimal detecting point of impurity/breakage, as shown in Fig. 11, which mainly includes box, camera, lens, light source, light source controller, embedded processor, and high-transparence glass. To prevent the interference of external light, the module is encapsulated in box, and the light is supplemented by an adjustable light source. A high-transparence glass is constructed between lens and grains to be detected for realizing high-frequency sampling and assessment of grain in the grain conveyer by external detection module. In order to reduce the influence of mirror imaging caused by glass reflection and brightness difference between two sides of glass on image quality, the low angle annular diffuse light source is adopted.

b: PREPROCESSING OF DYNAMIC RICE GRAIN FLOW IMAGES

For the factors of vibration, dust and reflect light, the acquired image may have problems such as shadow, edge blur, and noise, etc. Therefore, a nonlinear distortion model of camera imaging is constructed in this paper. A circular ROI (region of interest) is constructed according to the features of detecting region. Based on median filtering and high-frequency enhancement, the edge contrast of rice grains and impurities in the image is improved.

According to (3), ignoring high order component of distortion model, according to nonlinear distortions, such as radial distortion, centrifugal distortion and thin prism distortion, the mathematical model of total nonlinear distortion is established by taking the main distortion parameters.

$$\begin{cases} D_x = k_1 x_a r^2 + k_2 x_a r^4 + p_1(3x_a^2 + y_a^2) + 2p_2 x_a y_a + s_1 r^2 \\ D_y = k_1 y_a r^2 + k_2 y_a r^4 + 2p_1 x_a y_a + p_2(x_a^2 + 3y_a^2) + s_2 r^2 \end{cases} \tag{3}$$

$r^2 = x_a^2 + y_a^2$. the radial distortion is constrained by $x_a/y_a = x_r/y_r$. (x_a, y_a) and (x_r, y_r) are the image coordinates of the actual imaging and ideal imaging of camera, respectively. Transform (3) to obtain the camera distortion model:

$$\begin{bmatrix} x_a r^2 & x_a r^4 & 3x_a^2 + y_a^2 & 2x_a y_a & r^2 & 0 \\ y_a r^2 & y_a r^4 & 2x_a y_a & x_a^2 + 3y_a^2 & 0 & r^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \\ s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} x_r - x_a \\ y_r - y_a \end{bmatrix} \tag{4}$$

$D_k = [k_1\ k_2\ p_1\ p_2\ s_1\ _2]^T$ is the distortion parameter vector.

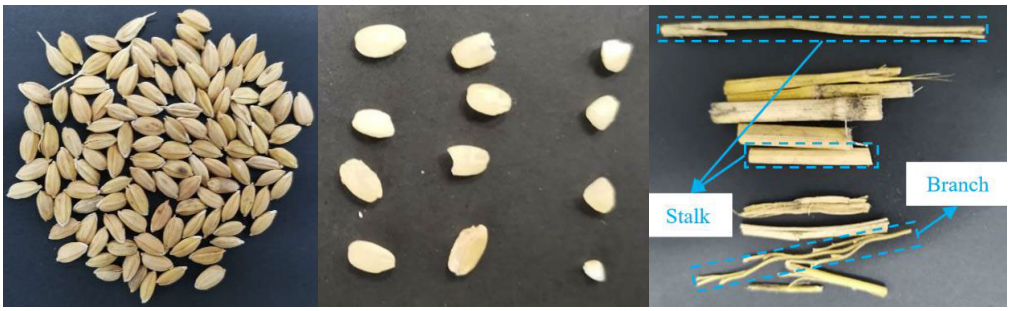
The circular illuminated region is constructed based on the size of light source, and is segmented based on gray histogram. The median filtering is used to smooth salt and pepper noise and pulse noise for maintaining edge features of image and reducing blur caused by noise reduction. According to (5), the filter is constructed to enhance image edges and reduce influence of shadow and edge blur on small object extraction. The preprocessing results are shown in Fig. 12.

$$P_1(x, y) = round((P(x, y) - \frac{1}{(2n+1)^2} \sum_{j=y-n}^{y+n} \sum_{i=x-n}^{x+n} P(i, j))^* a) + P(x, y) \tag{5}$$

$P(x, y)$ and $P_1(x, y)$ are the initial and enhanced pixel values at point (x, y), respectively, a is difference enhancement multiple, $n = (k-1)/2$, and k is the length and width of the filter, where k is odd and $k \geq 3$.

(a) (b) (c)

FIGURE 12. Preprocessed results. (a) Original image, (b) Corrected image, (c) Preprocessed image.



(a) (b) (c)

FIGURE 13. Objects in the image. (a) Complete grains, (b) Broken grains, (c) Impurities.

2) IMPURITY AND BROKEN GRAIN EXTRACTION BASED ON IMPROVED DEEPLABV3+ AND ELLIPSE ATTRIBUTES OF OBJECT

The objects in vehicle-mounted dynamic rice grain flow image mainly include complete and broken grains, impurities. The impurities mainly include branches and stalks, as shown in Fig. 13. The complete grain shape is oval and its shell is intact. The broken grains are incomplete elliptic and partially exposed. In impurities, the branches are curved and the stalks are straight. In order to improve the object extraction accuracy for a large amount of data, different methods are adopted to extract complete and broken grains, impurities in this paper, and the broken grains and impurities are extracted based on improved Deeplabv3+ and ellipse attributes.

a: EXTRACTION AND CLASSIFICATION OF IMPURITIES BASED ON IMPROVED DEEPLABV3+

From the rice images collected from field experiment, 400 images are selected as the original dataset without duplicate, consistent time and space, and omission. The original dataset is transformed and expanded by adding noise, data enhancement, scale change and affine transformation to obtain a dataset containing 6400 images. Labelme is used to label the transformed dataset, and the dataset is divided into training dataset and testing dataset according to 7:3. To accelerate the convergence speed of network for small samples dataset, transfer learning is used to train. Firstly, improve Deeplabv3+ is pre-trained based on large-scale Coco dataset to learn general features such as edges, lines and corners. Secondly, the training dataset of dynamic rice grain flow images is used to retrain improve Deeplabv3+ to learn object features. The network training adopts twice sampling freezing method. The first parameters are 5×10^{-4} learning rate, 4 batch size and 50 iterations. After freezing the weight of backbone network, the parameters are adjusted to 5×10^{-5} learning rate and 2 batch size. Finally, the network is verified and optimized on testing dataset for obtaining trained improved Deeplabv3+ network. It is used to extract impurities from dynamic rice grain flow images. And the impurities are subdivided based on features of edges and regions to obtain branch and stalk regions.

b: EXTRACTION OF BROKEN GRAINS BASED ON ELLIPSE ATTRIBUTES

The descriptors that can represent ellipse attributes of regions, such as anisometry E_d, bulkiness E_f, struct factor E_s of equivalent ellipse, circumcircle radius R_e, incircle radius R_i, and roundness R_n, are designed, as shown in (6)-(9) [37].

$$E_d = \frac{r_1}{r_2} \tag{6}$$

$$E_f = \frac{\pi r_1 r_2}{\alpha} \tag{7}$$

$$E_s = E_d \cdot E_f - 1 \tag{8}$$

$$R_n = 1 - \frac{d_p}{d_a} \tag{9}$$

r_1, r_2, α represent semi-major axis and semi-minor axis of equivalent ellipse, and the orientation of r_1 with regard to the x-axis of image coordinate, which can be calculated based on region geometric moments U_{pq} according to (10)-(12). d_a and d_p are the average distance and distance deviation from

the boundary to center of region, respectively.

$$r_1 = \sqrt{2(U_{20} + U_{02} + \sqrt{(U_{20} - U_{02})^2 + 4U_{11}^2})} \quad (10)$$

$$r_2 = \sqrt{2(U_{20} + U_{02} - \sqrt{(U_{20} - U_{02})^2 + 4U_{11}^2})} \quad (11)$$

$$\alpha = -\frac{1}{2}\arctan\frac{2U_{11}}{U_{02} - U_{20}} \quad (12)$$

Because of the randomness of broken grain region position, the position is independent of the region category. In order to avoid the influence of region position on geometric moment, the center of region is used to normalize the geometric moment according to (13)-(15).

$$N_{pq} = \frac{1}{a}\sum_{(x,y)\in R} x^p y^q \quad (13)$$

$$O_r = (N_{10}, N_{01}) = (\frac{1}{a}\sum_{(x,y)\in R} x^1 y^0, \frac{1}{a}\sum_{(x,y)\in R} x^0 y^1) \quad (14)$$

$$U_{pq} = \frac{1}{a}\sum_{(x,y)\in R} (x - N_{10})^p (y - N_{01})^q \quad (15)$$

R is the region, (x, y) is the image coordinates of pixels in the region, a is the area of region, O_r is the center of region, (N_{10}, N_{01}) is the image coordinate of O_r, N_{pq} is the geometric moment normalized by region area, U_{pq} is the geometric moment normalized by area and position. d_a and d_p are also calculated based on O_r, according to (16)-(17). The normalized geometric moment is adopted to define the ellipse equivalent to the region. The center of equivalent ellipse is the same as the center of region. Based on the constructed feature descriptors, a multi-dimensional feature vector with ellipse attributes is constructed for describing and classifying exposed regions of broken grains for extracting broken grains from images.

$$d_a = \frac{1}{a}\sum_{(x,y)\in R} \sqrt{(x - N_{10})^2 + (y - N_{01})^2} \quad (16)$$

$$d_p = \sqrt{\frac{1}{a}\sum_{(x,y)\in R} (\sqrt{(x - N_{10})^2 + (y - N_{01})^2} - d_a)^2} \quad (17)$$

3) COMPLETE GRAIN EXTRACTION BASED ON IMPROVED YOLOv4

Compared with impurities, the number of complete grains in an image can reach up to 300. Therefore, compared with dataset for improved Deeplabv3+, the dataset for improved YOLOv4 is structured according to the rules of non-duplicate, differentiation in time and space, non-omission, and complete distribution for coverages of grains on images. The 400 selected original images are transformed and expanded to obtain a dataset containing 6400 images. The transformed dataset is labelled by Labelimg based on rectangular boxes, and is divided into training dataset and testing dataset according to 6:4. Transfer learning is still used to train, and the improved YOLOv4 is pre-trained based on PASCAL VOC2012 dataset to learn general features and retrain based on training dataset to learn object features. For network training by twice sampling freezing method, the first parameters are 1×10^{-3} learning rate, 1 batch size and 40 iterations. After freezing the weight of backbone network, the parameters are adjusted to 1×10^{-4} learning rate and 1 batch size. The network is verified and optimized on testing dataset for obtaining trained improved YOLOv4. It is used to extract complete grains from dynamic rice grain flow images and count the number of complete grains in an image.

4) IMPURITY/BREAKAGE ASSESSMENT BASED ON THE RELATIONSHIP MODEL BETWEEN PIXEL AREA AND QUALITY

The impurity/breakage assessment is related to the quality of object, but the extracted regions of complete and broken grains, impurities in images only has pixel and quantity information. It is necessary to construct relationship model between pixel area and quality to accurately and rapidly assess the impurity and breakage rates.

a: CONSTRUCTION OF RELATIONSHIP MODEL BETWEEN PIXEL AREA AND QUALITY

For complete grains, the relationship between grain quantity and total quality is calculated based on 1000 grain weight obtained by precision electronics balance with 0.001g accuracy and manual counting. For broken grains and impurities, the pixel area and quality are obtained by static test stand for relationship fitting. The constructed image acquisition module of dynamic rice grain flow is installed at static test stand for simulating actual detecting environment, as shown in Fig. 14. The samples of broken grains, branches and stalks are placed on the detection plane by category for acquiring image with single category sample. The objects in images are extracted by proposed method in this paper to obtain a pixel set in which each region corresponds to the sample one by one, as shown in Fig. 15. The quality set in which each data corresponds to the sample one by one is obtained by electronics balance.

From Fig. 16 we can see that there is a linear relationship between pixel and quality, so the least square method is used to fit. The relationship model between pixel area and quality for broken grains, branches and stalks obtained are as follows:

$$M_r = 1.0543 \times 10^{-5} \times S_r \quad (18)$$

$$M_s = 3.1732 \times 10^{-6} \times S_s - 0.02859 \quad (19)$$

$$M_b = 1.6348 \times 10^{-6} \times S_b + 0.016 \quad (20)$$

The M_r, M_s and M_b are qualities of broken grains, branches and stalks, respectively and its unit is g. The Sr, Ss and Sb are pixel areas of broken grains, branches and stalks, respectively and its unit is pixel.

b: ASSESSMENT OF IMPURITY AND BREAKAGE RATES

Based on the relationship model between pixel area and quality, an assessment model of impurity and breakage rates

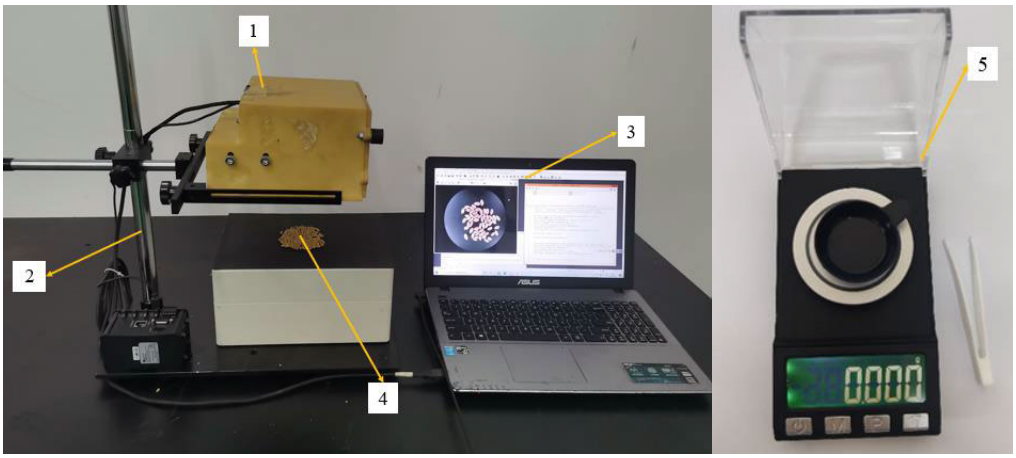
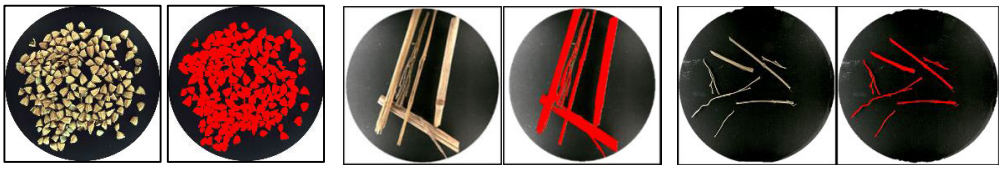
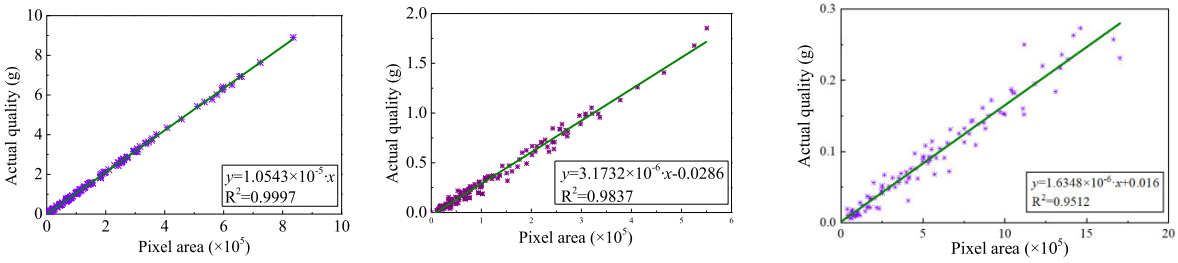


FIGURE 14. Static test stand. 1 Image acquisition module, 2 brace, 3 Processor, 4 Test objects, 5 electronics balance.



(a) (b) (c)

FIGURE 15. Sample pixel extraction. (a) Broken grains, (b) Branches, (c) Stalks.



(a) (b) (c)

FIGURE 16. Fitting for pixel and quality. (a) Broken grains, (b) Branches, (c) Stalks.

is constructed. If m is 1000 grain weight of complete grains, n_t is quantity of complete grains extracted from an image at time t during harvest operation, S_{rt}, S_{st} and S_{bt} are pixel areas of broken grains, branches and stalks at time t, respectively, then the impurity rate at time t is as (21), shown at the bottom of the page.

The M_{rt}, M_{st}, M_{bt} and M_{ct} are the total quantities of broken grains, branches, stalks and complete grains in an image at time t, respectively. The breakage rate at time t is:

$$Z_{dt} = \frac{M_{rt}}{M_{rt} + M_{ct}} = \frac{1.04 \times 10^{-5} \times S_{rt}}{1.04 \times 10^{-5} \times S_{rt} + m \times n_t/1000} \tag{22}$$

III. RESULTS

A. EXPERIMENTAL PLATFORM

To verify the effectiveness and advantages of proposed method, the image acquisition module of vehicle-mounted dynamic rice grain flow was installed on the 4LZ-6A multi-functional intelligent crawler-type combine harvester developed by our team, as shown in Fig. 17. The processor was Jetson Xavier NX with Ubuntu18.04 system, Python3.8 com-

$$Z_{it} = \frac{M_{st} + M_{bt}}{M_{rt} + M_{st} + M_{bt} + M_{ct}} = \frac{\left(3.1732 \times 10^{-6} \times S_{st} - 0.02859\right) + \left(1.6348 \times 10^{-6} \times S_{bt} + 0.016\right)}{1.04 \times 10^{-5} \times S_{rt} + \left(3.1732 \times 10^{-6} \times S_{st} - 0.02859\right) + \left(1.6348 \times 10^{-6} \times S_{bt} + 0.016\right) + m \times n_t/1000} \tag{21}$$

FIGURE 17. Experimental platform.

pilation environment, Opencv, PaddlePaddle, Pytorch deep learning framework and CUDA architecture. The experiments were conducted in rice fields in Picheng Town, Zhenjiang City and Wujiang National Modern Agriculture Demonstration Zone, Suzhou.

B. EXPERIMENTAL METHODS

Based on the images collected in the field, the comparative experiments were conducted on the improved Deeplabv3+, improved YOLOv4 and the existing methods. According to trained network models, the effectiveness and advantages of the assessment method for impurity and breakage rates were verified in the field experiments.

1) EXPERIMENT METHOD FOR COMPARISON OF IMPROVED DEEPLABV3+ WITH EXISTING NETWORK

The datasets of images from all experimental fields were used for impurity and broken grain extraction by trained existing and improved Deeplabv3+ model. The artificial visual markers were used as the real values. The experimental results were evaluated based on MIoU (mean intersection over union) and MPA (mean pixel accuracy) indicators. As shown in (23)-(24), MPA is the cumulative average of proportions of the correct pixel number for each class by detection in the total pixel number. MIoU is the cumulative average of the ratios of intersection to union for predicted results and real values of each class. The p_{ij} represents the pixel number of class i detected as class j. The extraction accuracy of impurities and broken grains is calculated based on pixel area ratio, according to (25). The A_p and A_r are the extracted and actual pixel areas of impurities and broken grains, respectively.

$$MPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}} \tag{23}$$

$$MIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ij}}{\sum_{j=0}^{k}p_{ij}+\sum_{j=0}^{k}p_{ji}-p_{ii}} \tag{24}$$

$$A_e = \frac{A_p}{A_r} \tag{25}$$

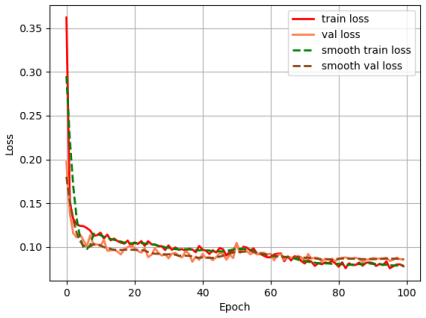


FIGURE 18. Training loss curves of improved DeepLabv3+.

2) EXPERIMENT METHOD FOR COMPARISON OF IMPROVED YOLOV4 WITH EXISTING NETWORK

The improved YOLOv4 was experimentally compared with the existing neural networks, such as YOLOv4, U-NET and BP, and traditional image processing algorithms, such as template matching and feature extraction. The complete grain extraction methods based on U-NET and BP referred to references [38] and [20]. The grain feature extraction method based on threshold segmentation and region classification and the grain template matching method based on edge detection and shape matching were designed in this paper. Based on the datasets of images from experimental fields, the experimental results were evaluated based on Precision, Recall and mAP indicators according to (26)- (28). TP is true positives in confusion matrix for network evaluation, FP is false positives, FN is false negatives, n is the number of classification category, and p(r) is precision-recall curve. According to (29), the accuracy of complete grain extraction was calculated for an image. The A_r and A_c are the actual and extracted number of complete grains, respectively.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{26}$$

$$Recall = \frac{TP}{TP+FN} \tag{27}$$

$$mAP = \frac{\int_0^1 p(r)dr}{n} \tag{28}$$

$$A_a = 1 - \frac{|A_r - A_c|}{A_r} \tag{29}$$

3) EXPERIMENTAL METHOD FOR IMPURITY/BREAKAGE ASSESSMENT

In the actual field operation, the experiments were carried out with 1 m/s operation speed, 200mm cutting height and 5s beat for impurity/breakage assessment. Five groups of experiments were conducted in five different fields. Each experiment was at a distance of 20 m, and four detection data were obtained in each group, with a total of 20 data results. During 5 s beat interval of each experiment in each group, the field operation was stopped and the samples in grain tank were sampled manually. Manual statistical average results of 5 samples were used as the actual value of impurity and breakage rate in this beat. Before the next beat, the grain

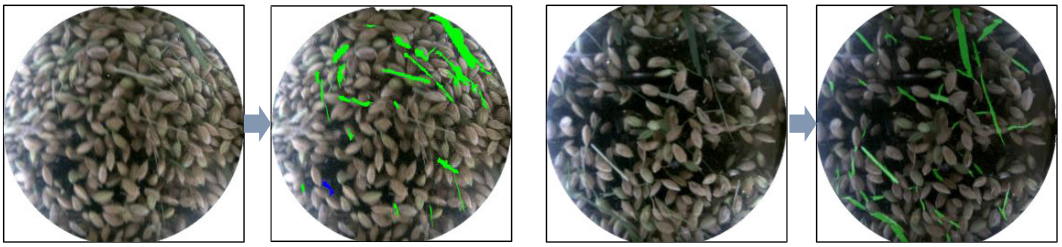


FIGURE 19. Extraction of impurities and broken grains.

TABLE 2. Comparison of impurity extraction.

Methods	MIoU/%	MPA/%	$\overline{t_d}$ /s
Deeplabv3+ Existing Deeplabv3+	70.08	75.24	0.14
Deeplabv3+ Improved Deeplabv3+	74.21	80.16	0.08

tank was emptied to ensure the accuracy of the actual value in each beat. According to (30)-(31), the assessment accuracy of impurity and breakage rates were calculated by relative error. The e_i, Z_{ir} and Z_{ic} are the relative error, actual and detected values of impurity rate, respectively. The e_d, Z_{dr} and Z_{dc} are the relative error, actual and detected values of breakage rate, respectively.

$$e_i = \frac{|Z_{ir} - Z_{ic}|}{Z_{ir}} \tag{30}$$

$$e_d = \frac{|Z_{dr} - Z_{dc}|}{Z_{dr}} \tag{31}$$

C. EXPERIMENTAL RESULTS AND DISCUSSION

1) COMPARISON OF EXISTING AND IMPROVED DEEPLABV3+ FOR IMPURITY AND BROKEN GRAIN EXTRACTION

The loss curve of improved Deeplabv3+ for impurity extraction on training dataset is shown in Fig. 18, and the training loss reached 0.078. The trained existing and improved Deeplabv3+ were used for impurity extraction, and the MIoU and MPA of the networks were calculated on testing dataset according to (23)-(24). The average time $\overline{t_d}$ of impurity extraction for single frame by the two networks was calculated. The indicators are shown in Table 2.

From Table 2, we can see that, compared with existing Deeplabv3+, the MioU and MPA of improved Deeplabv3+ for impurity extraction increased by 4.13% and 4.92%, respectively, and average time decreased by 0.06 s. The improved Deeplabv3+ and ellipse attributes were used to extract impurities and broken grains. According to (25), the extraction accuracy of all images in testing dataset were calculated and the average accuracy for impurity and broken grain extraction was 93.24% and 94.58%, respectively. The extraction results are shown in Fig. 19, in which green regions are impurities and blue regions are broken grains.

TABLE 3. Comparison of complete grain extraction.

Methods	YOLOv4 improved YOLOv4	YOLOv4 Existing YOLOv4	U-Net	BP
Precision/%	98.65	95.78	91.15	81.25
Recall/%	90.07	84.41	79.28	73.54
mAP/%	91.47	85.62	78.89	65.99
A_a /%	99.16	95.15	90.01	64.06
$\overline{t_y}$ /s	0.72	0.70	4.75	2.53

2) COMPARISON OF IMPROVED YOLOv4 AND EXISTING METHODS FOR COMPLETE GRAIN EXTRACTION

a: COMPARISON OF IMPROVED YOLOv4 AND EXISTING NEURAL NETWORKS

The loss curve of improved YOLOv4 for complete grain extraction on training dataset is shown in Fig. 20, and the training loss reached 2.080. The trained improved YOLOv4 was analyzed by comparing with existing YOLOv4, U-Net and BP on testing dataset. According to (26)-(29), the Precision, Recall, mAP, A_a and average computation time $\overline{t_y}$ of the networks were calculated, respectively, as shown in Table 3.

From Table 3, we can see that, compared with the existing YOLOv4, U-NET and BP, the precision of improved YOLOv4 in complete grain extraction increased by 2.87%, 7.5% and 17.4%, the Recall increased by 5.66%, 10.79% and 16.53%, mAP increased by 5.85%, 12.58% and 25.48%, and A_a increased by 4.01%, 9.15% and 35.1%, respectively. Compared with U-Net and BP, the computation time of improved YOLOv4 decreased by 1.81-4.03 s. Compared with existing YOLOv4, the improved YOLOv4 is more complex, but the increase in computation time is small, which is basically the same as before. The existing and improved YOLOv4 both meet the lowest beat requirement of 1s for impurity/breakage assessment.

b: COMPARISON OF IMPROVED YOLOv4 AND TRADITIONAL IMAGE PROCESSING ALGORITHMS

The trained improved YOLOv4 was analyzed by comparing with template matching and feature extraction algorithms on testing dataset. According to (29), the A_a and computation time t_y of the algorithms were calculated, respectively. The calculation results of 100 randomly selected images were shown in Fig. 21. The average A_a of the template matching

TABLE 4. Assessment results of impurity and breakage rates.

groups	Impurity rates			Breakage rates		
	assessing values	Actual values	assessing time	assessing values	Actual values	assessing time
1	3.70	3.80	1.54	0.67	0.59	1.06
2	3.07	3.16	1.62	0.89	0.79	1.18
3	2.55	2.60	1.59	0.66	0.71	1.12
4	1.96	2.14	1.52	0.53	0.57	1.03
5	2.07	2.13	1.52	1.30	1.43	1.05
6	1.77	1.87	1.51	1.20	1.29	1.02
7	0.75	0.83	1.6	0.77	0.85	1.09
8	0.24	0.27	1.58	0.75	0.84	1.12
9	0.95	1.02	1.6	0.53	0.59	1.23
10	2.53	2.48	1.51	0.74	0.82	1.08
11	2.28	2.43	1.5	0.63	0.71	1.01
12	2.11	2.18	1.54	0.77	0.85	1.11
13	1.36	1.49	1.52	0.58	0.61	1.06
14	1.35	1.43	1.56	0.42	0.44	1.09
15	1.20	1.27	1.62	0.83	0.91	1.16
16	1.58	1.74	1.53	1.53	1.71	1.07
17	3.42	3.74	1.51	0.86	0.94	1.06
18	2.35	2.53	1.58	0.98	1.08	1.13
19	2.65	2.88	1.56	1.23	1.38	1.08
20	1.68	1.84	1.52	0.76	0.83	1.06

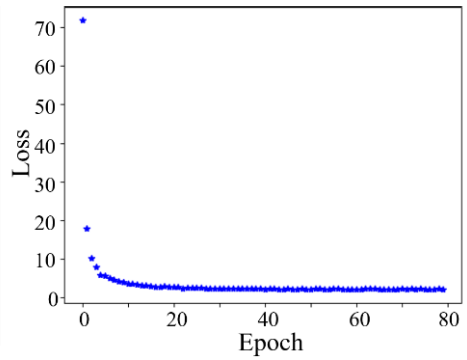


FIGURE 20. Training loss curves of improved YOLOv4.

and feature extraction algorithms were 94.21% and 77.78%, and the average t_y were 0.37s and 0.024s, respectively. The three algorithms all meet the lowest beat requirement for impurity/breakage assessment. Compared with the template matching and feature extraction algorithms, the average A_a of improved YOLOv4 increased by 4.95% and 21.38%, respectively.

The extraction results of complete grains by the three algorithms are shown in Fig. 22, in which the red boxes/ regions are the detection results of complete grains. From Fig. 22, we can see that, the feature extraction algorithm was easily affected by uneven illumination and grayscale, which made the segmentation error larger. Compared with feature extraction algorithm, the extraction accuracy of template matching algorithm was improved, but still was easily affected by the changes of object scale, rotation and occlusion. Compared

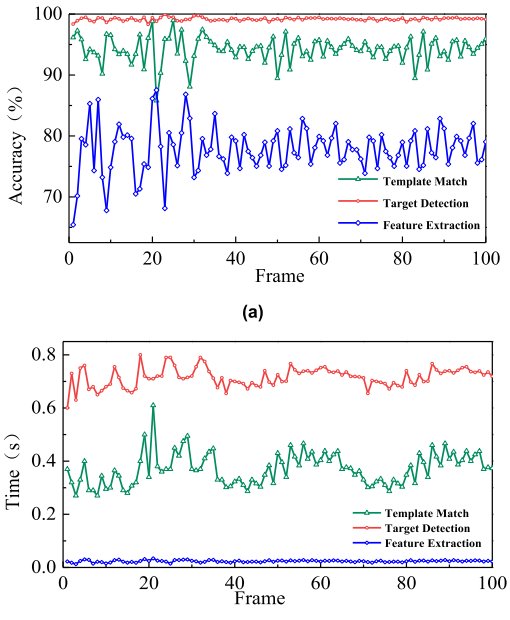
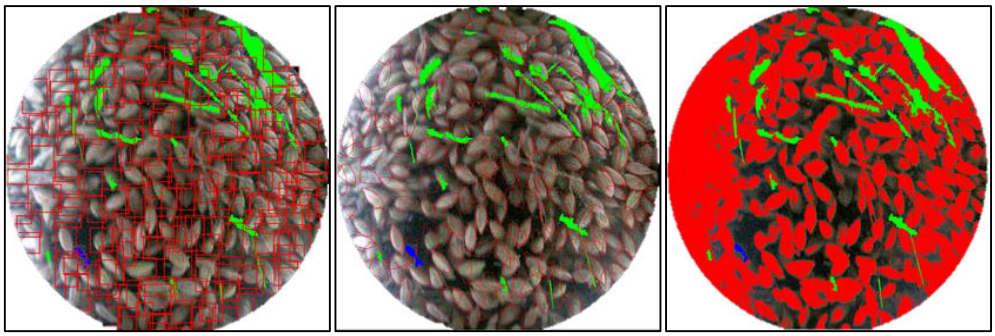


FIGURE 21. Comparison of complete grain extraction. (a) A_a, (b) t_y.

with other algorithms, the improved YOLOv4 had higher detection accuracy and better robustness.

3) FIELD EXPERIMENTS FOR IMPURITY/BREAKAGE ASSESSMENT

According to (21)-(22), the number of complete grains, the pixel areas of broken grains and impurities in image obtained



(a) (b) (c)

FIGURE 22. Extraction results of complete grains. (a) Improved YOLOv4, (b) Template matching, (c) Feature extraction.

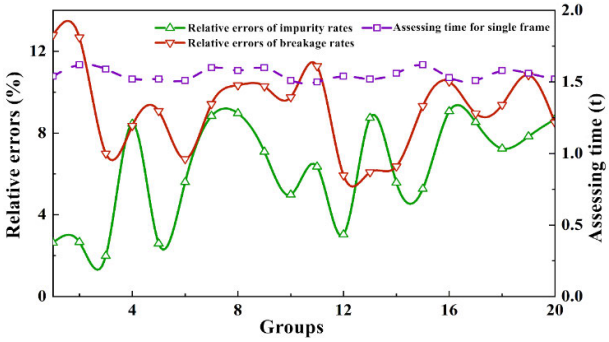


FIGURE 23. Assessing time and relative errors of impurity and breakage rates.

by improved Deeplabv3+ and YOLOv4 were used to assess impurity and breakage rates. The 20 data results are shown in Table 4. The calculation of impurity rate was based on the qualities of complete and broken grains, branches and stalks. The calculation of breakage rate was based on the qualities of complete and broken grains. The assessing time included the time of image processing, quality fitting and calculation of impurity and breakage rates, so the assessing time of breakage rate was less than that of impurity rate.

According to (30)-(31), the relative errors of assessment were calculated based on Table 4. As shown in Fig. 23, the relative errors of impurity rates stabilized at 2.6%-9.1%, and the average relative error was 6.20%. The relative errors of breakage rates stabilized at 6.0%-12.8%, and the average relative error was 9.18%. The assessing time of impurity and breakage rates stabilized at 1.5 s-1.6 s, and the average assessing time was 1.56s. The experimental results verified the effectiveness and advantages of the proposed method for impurity/breakage assessment in this paper.

IV. DISCUSSION AND CONCLUSION

For the small and dense objects with random state, such as grain and impurity, in the image of vehicle-mounted dynamic rice grain flow on combine harvester, it is difficult to accurately extract by existing Deeplabv3+ and YOLOv4 with unified attention. Therefore, this paper improved existing Deeplabv3+ and YOLOv4 based on coding layer optimization, data enhancement and attention mechanism, and proposed an impurity/breakage assessment method of vehicle-mounted dynamic rice grain flow on combine harvester. The following conclusions can be drawn:

1) An improved Deeplabv3+ model with coding layer and attention mechanism optimization was proposed. The Xception model of coding layer in constructed network model was replaced by the MobileNetv2 model. The ECA was added in encoder and decoder of network, respectively. It solved the problems that the existing Deeplabv3 with Xception without ability to change number of channels is easy to lose high-dimensional spatial features, has a unified attention mechanism for object detection, and are difficult to apply to extract small and dense objects with random state in dynamic rice grain flow image.

2) An improved YOLOv4 with data enhancement and attention mechanism optimization was proposed. Mosaic was replaced by Mixup for data enhancement in constructed network. The ReLU activation functions in Neck middle layer and Head output layer were replaced by Mish activation function. The ECA modules were added in three-layer networks before output to Neck and Prediction in BackBone. It solved the problems that the clipping and splicing does not significantly improve the network training effect for images with small and dense objects, and the ReLU is easy to cause gradient disappearance.

3) The modeling and discrete element analysis of rice grain movement on grain conveyer of harvester were carried out to locate optimal detecting point of impurity/breakage. The relationship model between complete grains, broken grains, impurities and pixel area, object quantity was constructed. An impurity/breakage assessment method of vehicle-mounted dynamic rice grain flow on combine harvester based on improved Deeplabv3+ and YOLOv4 was proposed.

4) The proposed impurity/breakage assessment method of vehicle-mounted dynamic rice grain flow on combine harvester was adopted to the 4LZ-6A multi-functional intelligent crawler-type combine harvester developed by our team for comparative experiments. The improved Deeplabv3+ and ellipse attributes were adopted to impurity and broken

grain extraction. Compared with existing Deeplabv3+, the MioU and MPA of improved Deeplabv3+ increased by 4.13% and 4.92%, and average time decreased by 0.06s. The improved YOLOv4 was adopted to complete grain extraction. Compared with the existing YOLOv4, U-NET and BP, the accuracy increased by 4.01%, 9.15% and 35.1%, respectively. Compared with the template matching and feature extraction algorithms, the accuracy increased by 4.95% and 21.38%, respectively. Based on proposed impurity/breakage assessment method, the average relative error of impurity rates was 6.20%, the average relative error of breakage rates was 9.18%, and the average assessing time was 1.56s. The improved Deeplabv3+ and YOLOv4 can be used for real-time impurity/breakage assessment for vehicle-mounted dynamic rice grain flow on combine harvester in real time, and improving the assessment accuracy to further realize closed-loop control of intelligent harvesting operation.

The improved Deeplabv3+ and YOLOv4 with attention mechanism optimization can be used to improve the accuracy and efficiency for impurity/breakage assessment of vehicle-mounted dynamic rice grain flow on combine harvester. The research in this paper focuses on impurity/breakage assessment on conveying auger of combine harvester. In the future, this method can be further extended to other nodes in the harvester operation process to lay the foundation for high-precision closed-loop control of intelligent operation of harvesters.

REFERENCES

[1] F. Wang, F. Wang, J. Hu, L. Xie, and X. Yao, "Rice yield estimation based on an NPP model with a changing harvest index," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2953–2959, 2020.

[2] Y. Luo, L. Wei, L. Xu, Q. Zhang, J. Liu, Q. Cai, and W. Zhang, "Stereo-vision-based multi-crop harvesting edge detection for precise automatic steering of combine harvester," *Biosyst. Eng.*, vol. 215, pp. 115–128, Mar. 2022.

[3] Q. Zhang, Z. H. Zhao, and G. Q. Gao, "Fuzzy comprehensive evaluation for grasping prioritization of stacked fruit clusters based on relative hierarchy factor set," *Agronomy-Basel*, vol. 12, no. 3, p. 22, Mar. 2022.

[4] J. Chen, Y. Lian, R. Zou, S. Zhang, X. Ning, and M. Han, "Real-time grain breakage sensing for Rice combine harvesters using machine vision technology," *Int. J. Agricult. Biol. Eng.*, vol. 13, no. 3, pp. 194–199, 2020.

[5] Z. Huang, J. Xue, B. Ming, K. Wang, R. Xie, P. Hou, and S. Li, "Analysis of factors affecting the impurity rate of mechanically-harvested maize grain in China," *Int. J. Agricult. Biol. Eng.*, vol. 13, no. 5, pp. 17–22, 2020.

[6] H. Zeng, J. Lei, J. Tao, and C. Liu, "Yield monitoring for grain combine harvester based on monocular vision," *Trans. Chin. Soc. Agricult. Machinery*, vol. 52, no. 12, pp. 281–289, 2021.

[7] D. Geng, D. Tan, G. Su, Z. Wang, Z. Wang, and X. Ji, "Optimization and experimental verification of grain yield monitoring system based on pressure sensors," *Trans. Chin. Soc. Agricult. Eng.*, vol. 37, no. 9, pp. 245–252, 2021.

[8] M. Chen, C. Jin, Y. Ni, T. Yang, and G. Zhang, "Online field performance evaluation system of a grain combine harvester," *Comput. Electron. Agricult.*, vol. 198, Jul. 2022, Art. no. 107047.

[9] M. Chen, Y. Ni, C. Jin, Z. Liu, and J. Xu, "Spectral inversion model of the crushing rate of soybean under mechanized harvesting," *Food Sci. Technol.*, vol. 42, p. 9, May 2022.

[10] Y. Shen, Y. Yin, B. Li, C. Zhao, and G. Li, "Detection of impurities in wheat using terahertz spectral imaging and convolutional neural networks," *Comput. Electron. Agricult.*, vol. 181, Feb. 2021, Art. no. 105931.

[11] M. Chen, Y. Ni, C. Jin, J. Xu, and G. Zhang, "Online monitoring method of mechanized soybean harvest quality based on machine vision," *Trans. Chin. Soc. Agricult. Machinery*, vol. 52, no. 1, pp. 91–98, 2021.

[12] X. Wei and M. Zhang, "Extraction of crop height and cut-edge information based on binocular vision," *Trans. Chin. Soc. Agric. Mach.*, vol. 53, no. 3, pp. 225–233, Jan. 2022.

[13] M. A. Momin, K. Yamamoto, M. Miyamoto, N. Kondo, and T. Grift, "Machine vision based soybean quality evaluation," *Comput. Electron. Agricult.*, vol. 140, pp. 452–460, Aug. 2017.

[14] J. Chen, Y. Gu, Y. Lian, and M. Han, "Online recognition method of impurities and broken paddy grains based on machine vision," *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agricult. Eng.*, vol. 34, no. 13, pp. 187–194, 2018.

[15] J. Chen, Y. Lian, and Y. Li, "Real-time grain impurity sensing for Rice combine harvesters using image processing and decision-tree algorithm," *Comput. Electron. Agricult.*, vol. 175, Aug. 2020, Art. no. 105591.

[16] S. Agaazizi, M. Rasekh, Y. Abbaspour-Gilandeh, and M. H. Kianmehr, "Identification of impurity in wheat mass based on video processing using artificial neural network and PSO algorithm," *J. Food Process. Preservation*, vol. 45, no. 1, p. 13, Jan. 2021.

[17] D. Ngampak and P. Piamsa-Nga, "Image analysis of broken Rice grains of Khao Dawk Mali rice," in *Proc. 7th Int. Conf. Knowl. Smart Technol. (KST)*, Jan. 2015, pp. 115–120.

[18] D. Liu, X. Ning, Z. Li, D. Yang, H. Li, and L. Gao, "Discriminating and elimination of damaged soybean seeds based on image characteristics," *J. Stored Products Res.*, vol. 60, pp. 67–74, Jan. 2015.

[19] S. Zhu, J. Zhuo, H. Huang, and G. Li, "Wheat grain integrity image detection system based on CNN," *Trans. Chin. Soc. Agricult. Machinery*, vol. 51, no. 5, pp. 36–42, 2020.

[20] K. Wu, M. Zhang, G. Wang, X. Chen, and J. Wu, "A continuous single-layer discrete tiling system for online detection of corn impurities and breakage rates," *Agricult.-Basel*, vol. 12, no. 7, p. 18, Jul. 2022.

[21] D. Xu, L. Wang, F. Li, Y. Guo, and K. Xing, "Application of improved faster RCNN in food insect object detection," *J. Chin. Cereals Oils Assoc.*, vol. 37, no. 4, pp. 178–186, 2022.

[22] K. Zou, Q. Liao, F. Zhang, X. Che, and C. Zhang, "A segmentation network for smart weed management in wheat fields," *Comput. Electron. Agricult.*, vol. 202, Nov. 2022, Art. no. 107303.

[23] A. M. Roy, R. Bose, and J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," *Neural Comput. Appl.*, vol. 34, no. 5, pp. 3895–3921, Mar. 2022.

[24] S. K. Skovsen, M. S. Laursen, R. K. Kristensen, J. Rasmussen, M. Dyrmann, J. Eriksen, R. Gislum, R. N. Jorgensen, and H. Karstoft, "Robust species distribution mapping of crop mixtures using color images and convolutional neural networks," *Sensors*, vol. 21, no. 1, p. 28, Jan. 2021.

[25] D. Zhang, Y. Ding, P. Chen, X. Zhang, Z. Pan, and D. Liang, "Automatic extraction of wheat lodging area based on transfer learning method and deeplabv3+ network," *Comput. Electron. Agricult.*, vol. 179, Dec. 2020, Art. no. 105845.

[26] L. Wang, Y. Zhao, S. Liu, Y. Li, S. Chen, and Y. Lan, "Precision detection of dense plums in orchards using the improved YOLOv4 model," *Frontiers Plant Sci.*, vol. 13, p. 14, Mar. 2022.

[27] Z. Na and P. Younghwan, "A research on the classification of intelligence level of unmanned grain harvester," *J. Korea Converg. Soc.*, vol. 11, no. 5, pp. 165–173, 2020.

[28] L. C. E. Chen, Y. K. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder–decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 833–851.

[29] C. X. Zhang, F. Kang, and Y. X. Wang, "An improved apple object detection method based on lightweight YOLOv4 in complex backgrounds," *Remote Sens.*, vol. 14, no. 17, p. 23, Sep. 2022.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[31] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.

[32] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "*mixup*: Beyond empirical risk minimization," in *Proc. 6th Int. Conf. Learn. Represent.*, 2018, pp. 1–13.

[33] Z. Liang, Y. Li, J. D. Baerdemaeker, L. Xu, and W. Saeys, "Development and testing of a multi-duct cleaning device for tangential-longitudinal flow Rice combine harvesters," *Biosyst. Eng.*, vol. 182, pp. 95–106, Jun. 2019.

[34] J. Yuan, C. Wu, H. Li, X. Qi, X. Xiao, and X. Shi, "Movement rules and screening characteristics of rice-threshed mixture separation through a cylinder sieve," *Comput. Electron. Agricult.*, vol. 154, pp. 320–329, Nov. 2018.
[35] Z. Ma, M. Han, Y. Li, H. Gao, E. Lu, F. A. Chandio, and K. Ma, "Motion of cereal particles on variable-amplitude sieve as determined by high-speed image analysis," *Comput. Electron. Agricult.*, vol. 174, Jul. 2020, Art. no. 105465.
[36] L. Xu, C. Wei, Z. Liang, X. Chai, Y. Li, and Q. Liu, "Development of rapeseed cleaning loss monitoring system and experiments in a combine harvester," *Biosyst. Eng.*, vol. 178, pp. 118–130, Feb. 2019.
[37] Q. Zhang and G. Gao, "Grasping point detection of randomly placed fruit cluster using adaptive morphology segmentation and principal component classification of multiple features," *IEEE Access*, vol. 7, pp. 158035–158050, 2019.
[38] J. Chen, M. Han, Y. Lian, and S. Zhang, "Segmentation of impurity Rice grain images based on U-Net model," *Trans. Chin. Soc. Agricult. Eng.*, vol. 36, no. 10, pp. 174–180, 2020.

QIAN ZHANG was born in 1991. She received the B.Eng., M.S., and Ph.D. degrees from Jiangsu University, in 2013, 2016, and 2020, respectively. Since 2020, she has been a Research Assistant with the College of Agricultural Engineering, Jiangsu University. Her research interests include machine vision in agriculture, operation status of harvester, and environmental information detection.

JINPENG HU was born in 1996. He received the B.Eng. and M.S. degrees from the Henan University of Science and Technology, in 2013 and 2017, respectively. He is currently pursuing the Ph.D. degree with the College of Agricultural Engineering, Jiangsu University. His research interests include intelligent detection methods and control technology for combine harvesting equipment.

LIZHANG XU was born in 1979. He received the B.Eng., M.S., and Ph.D. degrees from Jiangsu University, in 2000, 2003, and 2009, respectively.

From 2005 to 2011, he was a Research Assistant with the College of Agricultural Engineering, Jiangsu University. From 2011 to 2015, he was an Associate Researcher with the College of Agricultural Engineering. Since 2015, he has been a Researcher with the College of Agricultural Engineering. His research interests include environmental information perception and precision operation of harvesting machinery, unmanned harvesting machinery and artificial intelligence systems, design theory, and structural innovation of multi crop harvesting machinery. He was a recipient of First Prize for Science and Technology Progress of the Ministry of Education, Second Prize of National Technological Invention Award, and 16th China Youth Science and Technology Award.

QIBING CAI was born in 1996. He received the B.Eng. degree in agricultural mechanization and automation from Shenyang Agricultural University, in 2019, and the M.S. degree in agricultural mechanization from Jiangsu University, in 2022. His research interest includes harvester operation state perception.

XUN YU was born in 1998. He received the B.Eng. degree from the Xuzhou Institute of Technology, in 2021. He is currently pursuing the M.S. degree with the College of Agricultural Engineering, Jiangsu University. His research interest includes harvester operation state perception.

PENG LIU was born in 1996. He received the B.Eng. degree in automotive engineering from Jiangsu University, in 2019, where he is currently pursuing the Ph.D. degree with the College of Agricultural Engineering. His research interests include intelligent combine harvester working environment information perception and precision operation.