An Overall Real-Time Mechanism for Classification and Quality Evaluation of Rice

Wanke Xia[1,2*], Ruoxin Peng[1*], Haoqi Chu[1], Xinlei Zhu[1], Zhiyu Yang[1], Yiting Zhao[2], Lili Yang[1†]

[1] China Agricultural University, [2] Tsinghua University
llyang@cau.edu.cn

Abstract

Rice is one of the most widely cultivated crops globally and has been developed into numerous varieties. The quality of rice during cultivation is primarily determined by its cultivar and characteristics. Traditionally, rice classification and quality assessment rely on manual visual inspection, a process that is both time-consuming and prone to errors. However, with advancements in machine vision technology, automating rice classification and quality evaluation based on its cultivar and characteristics has become increasingly feasible, enhancing both accuracy and efficiency. This study proposes a real-time evaluation mechanism for comprehensive rice grain assessment, integrating a one-stage object detection approach, a deep convolutional neural network, and traditional machine learning techniques. The proposed framework enables rice variety identification, grain completeness grading, and grain chalkiness evaluation. The rice grain dataset used in this study comprises approximately 20,000 images from six widely cultivated rice varieties in China. Experimental results demonstrate that the proposed mechanism achieves a mean Average Precision (mAP) of 99.14% in the object detection task and an accuracy of 97.89% in the classification task. Furthermore, the framework attains an average accuracy of 97.56% in grain completeness grading within the same rice variety, contributing to an effective quality evaluation system.

Datasets — https://huggingface.co/datasets/xwk25/RiceCC

1 Introduction

Rice is one of the most important crops world-wide and serves as a staple food for more than half of the global population (Fairhurst and Dobermann 2002). Its quality has a significant impact on dietary health. Additionally, the cultivar and characteristics of rice are key factors influencing its market value and selling price (Charoenthaikij et al. 2021). Therefore, the ability to assess rice quality rapidly and accurately is of critical importance. Rice quality is determined by a wide range of indicators. For instance, type quality is a fundamental factor in assessing the nutritional value of rice, while processing quality further influences its market competitiveness based on type quality (Verma, Srivastav, and Mohan 2018). Currently, manual sensory evaluation remains the most widely used method for rice quality assessment (Simonelli et al. 2017). However, this approach is susceptible to variations in lighting conditions, human eyesight, emotions, and other subjective factors, leading to slow identification speeds and an inability to meet the demands of rapid and objective evaluation.

With computer vision advancing rapidly, machine learning is increasingly applied in agriculture, enabling breakthroughs in crop detection (Zhang et al. 2024). It excels at efficient data analysis and processing, proving valuable for rice quality assessment (Robert Singh and Chaudhury 2020; Koklu, Cinar, and Taspinar 2021; Ahad et al. 2023). Several studies have demonstrated its effectiveness in this field. For example, (Moses et al. 2022) achieved single-class accuracies of 98.33%, 96.51%, 95.45%, 100%, 100%, 99.26%, and 98.72% for healthy, fully chalky, chalky discolored, semi-chalky, broken, discolored, and normal damage classes, respectively, attaining an overall classification accuracy of 98.37% using the EfficientNet-B0 architecture. (Din et al. 2024) developed the RiceNet model based on deep convolutional neural networks, using a dataset of over 4,700 images to classify five rice varieties, achieving an accuracy of 94%. (Lin et al. 2018) proposed a machine vision system based on deep convolutional neural networks for rice classification, obtaining an accuracy of 95.5%. (Zareiforoush et al. 2016) applied heuristic classification methods combined with computer vision to classify four rice types, achieving the highest classification accuracy of 98.72% using an artificial neural network topology. (Kurtulmuş and Ünal 2015) introduced a cost-effective method based on computer vision and machine learning, reporting an overall accuracy of 99.24% with the best predictive model.

In the field of rice quality evaluation, research utilizing computer vision in accordance with national or international standards remains limited. For instance, (Zareiforoush et al. 2016) provided only a qualitative grading of rice quality without employing quantitative criteria for comprehensive quality assessment and classification. Similarly, (Moses et al. 2022) categorized full and semi-chalky rice but did not measure chalkiness based on national standards.

To address prior study challenges, we propose an overall

*These authors contributed equally.
†Corresponding Author.

Accepted at the First International Workshop on AI in Agriculture (Agri AI), co-located with AAAI 2026.
Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Table 1: The basic information of each type of rice in our manually-collected dataset.

Type Name	Abbr	Rice Variety	Avg Length (mm)	Avg Width (mm)	Phenotypes
Guangdong Simiao Rice	GD	Indica Rice	6.74	1.74	Low or semi opacity, slim figure
Northeastern Glutinous Rice	NM	Glutinous Rice	4.45	2.86	Wax white color, high opacity, plump figure
Wuchang Rice	WC	Indica Rice	6.63	2.44	High opacity, plump figure
Panjin Crab Field Rice	PJX	Japonica Rice	4.82	2.83	High opacity, plump figure
Wannian Gong Rice	WN	Indica Rice	6.81	2.20	Low or semi opacity, slim figure
Yanbian Rice	YB	Japonica Rice	4.59	2.62	Low or semi opacity, plump figure

real-time mechanism for classification and quality evaluation of rice. Our primary contributions are as follows:

- An improved one-stage object detection approach is employed to identify different rice types in mixed-species samples and analyze their distribution patterns.
- An improved CNN is utilized to assess the completeness of six rice varieties, achieving a 2.0% increase in accuracy compared to other classification models.
- The K-means clustering algorithm is applied to quantify the chalky area in rice grains, providing a precise measurement for quality evaluation.
- A dataset with 20,000 images spanning 6 major cultivated rice varieties in China is manually obtained, thereby furnishing a high-caliber data foundation.

2 Data Acquisition

We have primarily collected an image dataset comprising six widely cultivated rice varieties in China, with their basic characteristics detailed in Table 1. Phenotypic images of individual rice grains were captured using an industrial camera, and each grain was manually annotated based on its completeness and chalkiness. The hardware setups for data acquisition include a Sony CMOS sensor (providing real-time imaging at 1920×1080P resolution), an AOSVI T2-HD228S body microscope, a Philips monitor, and other peripheral devices. The image acquisition was conducted as follows. First, rice grains were evenly spread on a carrier table with a monochrome background. Next, optical parameters were adjusted to obtain clear, high-magnification images, after which the industrial camera was used to capture phenotypic images. Finally, the acquired images underwent manual screening and labeling to ensure data quality. Following this procedure, we collected approximately 20,000 images of Chinese rice grains against a solid-color background. These images were manually filtered and labeled to construct a dataset for rice object detection and completeness assessment. To enhance model's robustness, each rice variety dataset included both intact and broken grains in a ratio of approximately 10:1. Also, to investigate the effect of varying optical conditions on rice chalkiness detection, we selected 10 grains from each variety and captured their images under 5 different lighting environments, yielding a total of 300 images dedicated to the rice chalkiness experiment. Detailed image processing procedures applied in later experiments can be found at Appendix A.

3 Methods

3.1 The One-Stage Model for Object Detection

One-stage object detection models perform detection and classification in a single forward pass without prior candidate regions. The YOLO series, a leading framework, has advanced real-time object detection with high accuracy (Jocher et al. 2020). Our research enhances performance via attention to YOLO-v5. YOLO-v5 comprises three primary components, which are backbone, neck, and head. The backbone is responsible for extracting feature representations from the input image. The neck further processes and fuses these features. The head predicts bounding boxes, class probabilities, and objectness scores to assess object properties. To enhance accuracy and mAP, we replace the third C3 layer with SimAM (Yang et al. 2021) at the backbone, as shown in Figure 1. Unlike conventional attention mechanisms that typically introduce additional parameters or complex computations to emphasize or suppress features, SimAM adopts a self-induced approach that optimizes an energy function to assess the importance of individual neurons, as shown in Algorithm 1. This enables the derivation of an analytical solution for feature weighting without increasing the network's parameter count.

Algorithm 1: SimAM Implementation.

Input: X (input feature with shape $[N, C, H, W]$), λ (coefficient parameter)
Output: Attended features

$\mathbf{n} \leftarrow X.\text{shape}[2] \odot X.\text{shape}[3] - 1$ ▷ spatial size
$\mathbf{X}_{\text{mean}} \leftarrow X.\text{mean}(\text{dim} = [2, 3])$ ▷ mean of X over spatial dimensions
$\mathbf{d} \leftarrow (X - \mathbf{X}_{\text{mean}}).\text{pow}(2)$
$\mathbf{v} \leftarrow \mathbf{d}.\text{sum}(\text{dim} = [2, 3])/\mathbf{n}$ ▷ channel variance
$\mathbf{E_inv} \leftarrow \mathbf{d}/(4 \cdot (\mathbf{v} + \lambda)) + 0.5$
return $X \odot \text{sigmoid}(\mathbf{E_inv})$ ▷ return attended features

3.2 The CNN Model for Completeness Evaluation

We enhance ConvNeXt-Tiny (Liu et al. 2022) with attention specifically tailored for rice completeness recognition. The core building foundation of ConvNeXt-Tiny is the ConvNeXt Block, which consists of multiple group convolution operations and transformation layers. By stacking multiple ConvNeXt Blocks, ConvNeXt-Tiny can be constructed to meet the complexity of the task. In this architecture, the output of the ConvNeXt Blocks is connected to a global average pooling layer and fully connected layers, with classification ultimately performed using the softmax function.

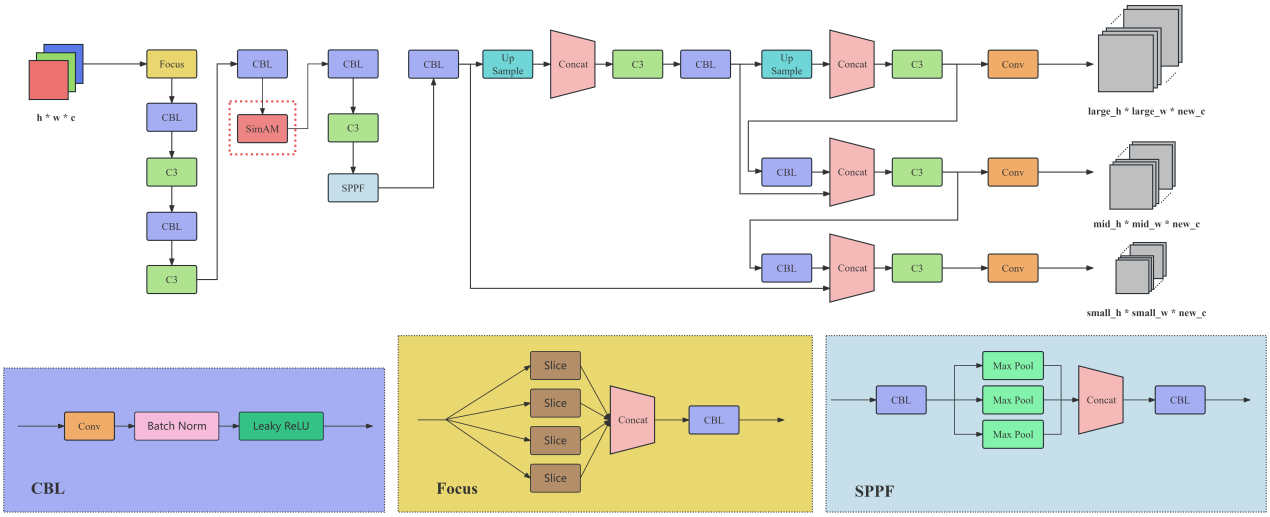

Figure 1: Framework of our improved YOLO-v5.

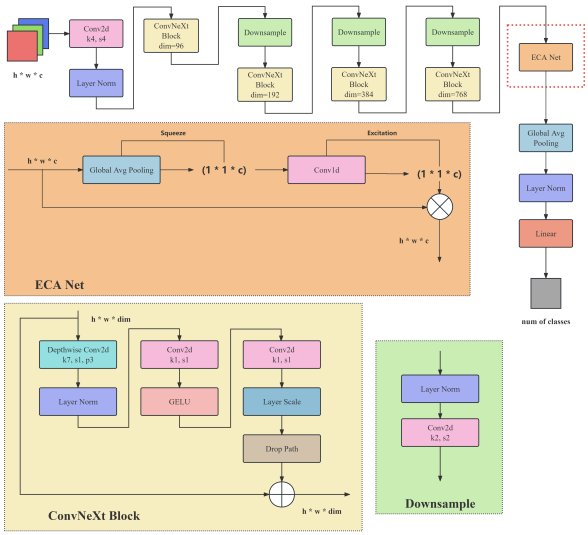

Figure 2: Framework of our improved ConvNeXt-Tiny.

To improve feature extraction and refine channel-wise attention, we modify ConvNeXt-Tiny by incorporating the Efficient Channel Attention (ECA) (Wang et al. 2020) between the final ConvNeXt Block and the global average pooling layer, as shown in Figure 2. ECA operates by computing attention weights independently for each feature map channel without considering spatial relationships between pixels or regions, allowing it to efficiently capture critical channel-wise information while avoiding computationally expensive pairwise interactions, thus well-suited for enhancing rice completeness recognition. Additionally, to reduce training costs and develop a high-accuracy yet practical model, we initialized the network using pre-trained ConvNeXt-Tiny weights from ImageNet-1K.

3.3 The K-means Algorithm for Chalk Recognition

Algorithm 2: K-means Algorithm for Chalk Recognition.

Input: Image pixels $\{p_1, ..., p_N\}$, clusters K, convergence threshold ϵ, category threshold T
Output: Segmented image with chalky (black) and non-chalky (white) regions

Initialize K cluster centers $\{\mu_1, \mu_2, ..., \mu_K\}$ randomly from image pixels
Repeat
  Assignment step: Assign each pixel p_i to the nearest cluster c
    where $c = \arg\min_j \|p_i - \mu_j\|^2$
  Update step: Recalculate each cluster center as the mean of assigned pixels
    $\mu_j = \frac{1}{|C_j|} \sum_{p_i \in C_j} p_i$ where C_j is the set of pixels in cluster j
Until Change in cluster centers is less than ϵ (convergence)
Segmentation:
for each pixel p_i in the image **do**
  Determine its cluster j from the assignment step
  if Cluster j meets chalky criteria based on T **then**
    Mark pixel as black (chalky region)
  else
    Mark pixel as white (non-chalky region)
  end if
end for
return Segmented image

The recognition of rice chalk involves segmenting chalky and non-chalky pixel areas within rice grains. Chalk refers

to the white, opaque portion in the rice endosperm, resulting from reduced translucency caused by gaps between endosperm starch granules. In chalk recognition, the K-means algorithm (as shown in Algorithm 2) is employed to assign image pixels to K cluster centers for image segmentation, and the determination of K is strictly based on the intrinsic spatial distribution characteristics of rice chalkiness.

Chalky areas typically aggregate in the middle region of rice grains, presenting as a single continuous opaque zone rather than discrete multiple regions. This spatial regularity indicates that the pixel set of chalky areas in rice grain images belongs to a single homogeneous cluster, while non-chalky pixels with consistent translucency and color form another dominant cluster. However, since the core goal of this study is to extract and calculate the area of the chalky region, we only need to focus on the cluster corresponding to chalky pixels. Thus, K is set to 1, where the K-means algorithm will iteratively converge to the cluster center that best represents chalky pixels. After clustering, we extract this single cluster, calculate its total pixel count, and convert it to the actual area of the rice chalky region based on the image resolution calibration parameters.

4 Experiments and Results

All experiments were conducted on a high-performance computing system equipped with an AMD EPYC 7543 32-core processor, an Nvidia RTX 3090, and Ubuntu 20.04 LTS. We primarily adhere to the Chinese National Standard *GB/T 1354-2018 Rice* as the assessment benchmark (ZHANG et al. 2021). Detailed descriptions about *GB/T 1354-2018 Rice* can be found at Appendix B.

4.1 Object Detection Experiment

We conducted a comparative analysis between our improved one-stage model and other two-stage models. The test accuracy for each model is presented in Table 2. The results indicate that two-stage models, when applied to both single-grain rice datasets and mixed-grain scenarios, fail to achieve high accuracy on the test set.

Table 2: The test accuracy of our improved YOLO-v5 and other two-stage methods in object detection experiment.

Model	Test Accuracy
Faster-RCNN (two-stage) (Ren et al. 2016)	87.33%
Tridentnet (two-stage) (Li et al. 2019)	93.69%
YOLO-v5 (one-stage)	95.05%
Ours (YOLO-v5 with SimAM, one-stage)	**97.89%**

Table 3: The validation accuracy, validation mAP, and test accuracy of the original and our improved YOLO-v5.

Model	Validation Accuracy	Validation mAP	Test Accuracy
YOLO-v5	97.54%	98.76%	95.05%
Ours(with SimAM)	**97.94%**	**99.14%**	**97.89%**

The confusion matrix of our improved YOLO-v5 on the validation set is presented in Figure 3. The model achieves a precision rate of 1.0 in GD, NM, and YB. Also, performance for WC, WN, and PJX is slightly lower but remained consistently above 0.95, indicating favorable overall outcomes.

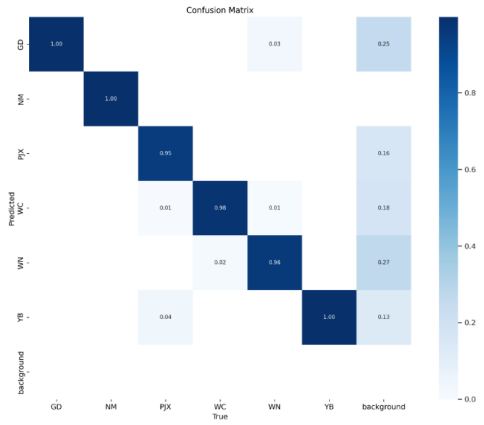

Figure 3: The confusion matrix of our improved YOLO-v5 on the validation set.

Example visual representation of our improved YOLO-v5 on the test set is illustrated in Figure 4. A cluster of NM instances is intricately interwoven with a sparse presence of GD. Results clearly demonstrate that the anchor boxes accurately and distinctly identify both NM and the interspersed GD, achieving commendable recognition performance.

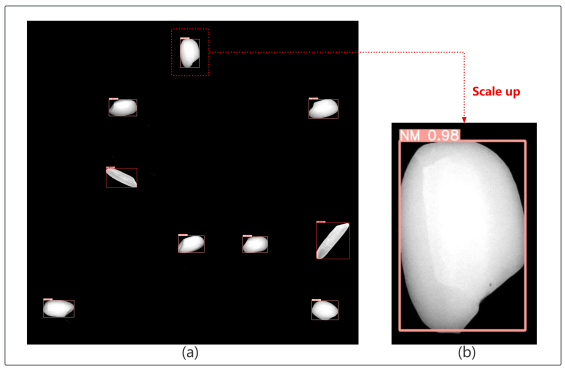

Figure 4: The example visual representation of our improved YOLO-v5 on the test set, wherein pink boxes denote NM instances, and red boxes signify GD instances. The upper-left corner of each box indicates the associated category and confidence level.

As shown in Table 3, an ablation analysis is conducted between our improved YOLO-v5 and its predecessor. It is worth noting that our improved model performs better in terms of validation accuracy and validation mAP, with increases of 0.40% and 0.38%, respectively. Moreover, our improved model demonstrates an elevated accuracy of 97.89% on the test set, significantly outperforming the original model's accuracy of 95.05%. These results collectively affirm the superior precision and object detection capabilities of our improved YOLO-v5 in practical scenarios.

Table 4: The test accuracy of various classification methods and our improved ConvNeXt-Tiny.

Method	Average Accuracy	GD	NM	PJX	WC	WN	YB
Decision Tree	93.72%	85.47%	92.42%	97.56%	95.13%	95.13%	96.59%
Random Forest	94.27%	87.60%	93.68%	97.88%	95.03%	94.84%	96.60%
SVM	91.82%	87.75%	87.50%	97.72%	91.30%	91.30%	95.35%
Googlenet (Szegedy et al. 2015)	89.24%	82.60%	95.65%	95.00%	95.23%	86.95%	80.00%
Alexnet (Krizhevsky, Sutskever, and Hinton 2012)	94.90%	89.36%	95.74%	97.62%	**97.73%**	91.30%	97.62%
Mobilenet-v2 (Sandler et al. 2018)	90.45%	78.72%	91.49%	92.86%	95.45%	91.30%	92.86%
Resnet101 (He et al. 2016)	85.19%	80.85%	74.47%	90.00%	84.09%	91.30%	90.48%
Ours(ConvNeXt-Tiny with ECA)	**97.61%**	**98.80%**	**96.35%**	**98.80%**	96.10%	**95.92%**	**99.68%**

4.2 Completeness Experiment

We initially collected data for three levels of completeness, which are whole rice, large broken rice, and tiny broken rice. The dataset included measurements of their short axis, long axis, and area. The data quantity for each level in one rice type was varied, resulting in an approximate ratio of 2:1:1. Based on the definitions of broken rice outlined in *GB/T 1354-2018 Rice*, we utilized various classification methods to classify and grade the completeness level of rice grains. As shown in Table 4, our improved ConvNeXt-Tiny basically achieves the best performance in both average accuracy and accuracy in each rice type. Although Alexnet shows an accuracy of 97.73% in Wuchang rice(WC), our improved ConvNeXt-Tiny still ranks the second place with the accuracy of 96.10%.

We have also conducted ablation experiments on the ECA module to verify its effectiveness in the framework of completeness evaluation. As shown in Table 5, our improved ConvNeXt-Tiny outweighs the original ConvNeXt-Tiny in identifying the three rice completeness grades, with an average accuracy improvement of nearly 2%. Notably, significant increases in the test accuracy are observed for Guangdong Simiao rice(GD) and Wuchang rice(WC). The test accuracy of Guangdong Simiao rice improves from 88.65% to 98.80%. And the test accuracy of Wuchang rice increases from 95.44% to 96.10%.

Table 5: The test accuracy of the original and our improved ConvNeXt-Tiny.

Model	Average Accuracy	GD	NM	PJX	WC	WN	YB
ConvNeXt-Tiny	95.58%	88.65%	**96.35%**	98.28%	95.44%	95.37%	99.41%
Ours(with ECA)	**97.61%**	**98.80%**	**96.35%**	**98.80%**	**96.10%**	**95.92%**	**99.68%**

4.3 Chalk Experiment

According to *GB/T 1354-2018 Rice*, rice grains with a chalk content of greater than or equal to 2.0% are classified as first-grade premium rice, implying that lower chalkiness relates to higher rice quality. In chalk recognition, rice images are collected under 5 different luminance intensities using the same light source, then converted to grayscale for processing. We employed K-means to cluster the images, and by setting a category threshold, chalky regions are segmented, appearing black while non-chalky areas appear white. To accurately determine the chalky area, a geometric polygon fitting method was used to estimate the segmented area. The formula for area calculating is detailed in Section B.2.

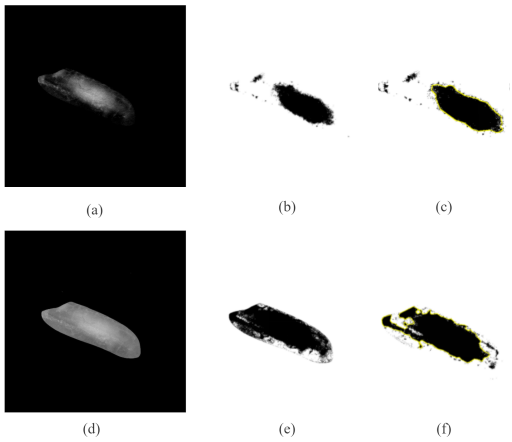

Figure 5: The processed images in chalk recognition of Guangdong Simiao Rice: (a) an original image at the brightness of level 2; (b) an image at the brightness of level 2 processed by K-means; (c) shows the theoretical area in chalk calculation surrounded by a yellow border; (d) an original image at the brightness of level 5; (e) an image at the brightness of level 5 processed by K-means; (f) shows the theoretical area in chalk calculation surrounded by a yellow border.

As shown in Figure 5, the size of the chalky regions increases with higher luminance intensities. To ensure the reliability of the comparison benchmark, we invited several Chinese experts specializing in agriculture. They were asked to visually identified and determined the chalky areas of rice grains in the subset images independently, serving as the human visual assessment reference standard for subsequent analysis. Manual comparisons revealed that the chalky area's ratio to the total rice area in images captured under luminance level one closely aligns with human visual assessments. Thus, luminance level one was selected as the standard luminance for this experiment. Compared to the manual visual assessment method prescribed in *GB/T 1354-2018 Rice*, this automated approach is not only more precise but also significantly faster.

5 Conclusion

Our proposed overall mechanism, which consists of 3 primary steps, manifests the capability to accurately identify rice varieties and perform quality analyses. First, improved YOLO-v5 is implemented to enable real-time visualization of rice positions and classifications, achieving an accuracy of 97.89% and an mAP of 99.14%. Second, enhanced ConvNeXt-Tiny is employed to evaluate the completeness grade within the same variety, yielding an average accuracy exceeding 97%. Third, K-means is applied to extract the chalky regions of whole rice and calculate precise chalkiness measurements. All experiments are conducted on our manually collected dataset with 20,000 images spanning six major cultivated rice varieties in China. We hope that future research will focus more on establishing international standards for rice grain quality assessment and collecting multi-regional rice grain image datasets.

References

Ahad, M. T.; Li, Y.; Song, B.; and Bhuiyan, T. 2023. Comparison of CNN-based deep learning architectures for rice diseases classification. *Artificial Intelligence in Agriculture*, 9: 22–35.

Charoenthaikij, P.; Chaovanalikit, A.; Uan-On, T.; and Waimaleongora-ek, P. 2021. Quality of different rice cultivars and factors influencing consumer willingness-to-purchase rice. *International Journal of Food Science and Technology*, 56(5): 2452–2461.

Din, N. M. U.; Assad, A.; Dar, R. A.; Rasool, M.; Sabha, S. U.; Majeed, T.; Islam, Z. U.; Gulzar, W.; and Yaseen, A. 2024. RiceNet: A deep convolutional neural network approach for classification of rice varieties. *Expert Systems with Applications*, 235: 121214.

Fairhurst, T.; and Dobermann, A. 2002. Rice in the global food supply. *World*, 5(7,502): 454–349.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Jocher, G.; Stoken, A.; Borovec, J.; Changyu, L.; Hogan, A.; Diaconu, L.; Ingham, F.; Poznanski, J.; Fang, J.; Yu, L.; et al. 2020. ultralytics/yolov5: v3. 1-bug fixes and performance improvements. *Zenodo*.

Koklu, M.; Cinar, I.; and Taspinar, Y. S. 2021. Classification of rice varieties with deep learning methods. *Computers and electronics in agriculture*, 187: 106285.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

Kurtulmuş, F.; and Ünal, H. 2015. Discriminating rapeseed varieties using computer vision and machine learning. *Expert Systems with Applications*, 42(4): 1880–1891.

Li, Y.; Chen, Y.; Wang, N.; and Zhang, Z. 2019. Scale-aware trident networks for object detection. In *Proceedings of the IEEE/CVF international conference on computer vision*, 6054–6063.

Lin, P.; Li, X.; Chen, Y.; and He, Y. 2018. A deep convolutional neural network architecture for boosting image discrimination accuracy of rice species. *Food and Bioprocess Technology*, 11(4): 765–773.

Liu, Z.; Mao, H.; Wu, C.-Y.; Feichtenhofer, C.; Darrell, T.; and Xie, S. 2022. A convnet for the 2020s. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11976–11986.

Moses, K.; Miglani, A.; Kankar, P. K.; et al. 2022. Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset. *Computers and Electronics in Agriculture*, 195: 106811.

Ren, S.; He, K.; Girshick, R.; and Sun, J. 2016. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6): 1137–1149.

Robert Singh, K.; and Chaudhury, S. 2020. A cascade network for the classification of rice grain based on single rice kernel. *Complex & Intelligent Systems*, 6(2): 321–334.

Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; and Chen, L.-C. 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4510–4520.

Simonelli, C.; Galassi, L.; Cormegna, M.; and Bianchi, P. 2017. Chemical, physical, textural and sensory evaluation on Italian rice varieties. *Univ. J. Agric. Res*, 5(2): 104–112.

Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; and Rabinovich, A. 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1–9.

Verma, D. K.; Srivastav, P. P.; and Mohan, M. 2018. Nutritional quality evaluation of different rice cultivars. In *Agronomic rice practices and postharvest processing*, 331–394. Apple Academic Press.

Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; and Hu, Q. 2020. ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11534–11542.

Yang, L.; Zhang, R.-Y.; Li, L.; and Xie, X. 2021. Simam: A simple, parameter-free attention module for convolutional neural networks. In *International conference on machine learning*, 11863–11874. PMLR.

Zareiforoush, H.; Minaei, S.; Alizadeh, M. R.; and Banakar, A. 2016. Qualitative classification of milled rice grains using computer vision and metaheuristic techniques. *Journal of food science and technology*, 53(1): 118–131.

ZHANG, J.; Zhenzhen, X.; Zhang, X.; Zheng, D.; and Youxiang, Z. 2021. Overview of quality and safety standards for rice in China. *China Rice*, 27(4): 77.

Zhang, Z.; Xiao, Y.; Chen, Z.; Chen, X.; and Li, X. 2024. Gait recognition for farm workers via multi-scale temporal feature perception. *Computers and Electronics in Agriculture*, 226: 109353.

A Image Processing

A.1 Image Processing for Object Detection

Some of the original rice images exhibited darker coloration, which obscured essential features such as translucency, chalkiness, contours, and other visual characteristics necessary for accurate observation. To enhance the visibility of these attributes, this study initially applied grayscale processing to the darker rice images. The maximum and minimum pixel values were computed, and each pixel value was remapped to a range of 0 to 255 using a lookup table, effectively replacing the original pixel values. This processing step enhanced the overall clarity, improving contour definition and sharpness, as illustrated in Figure 6. Subsequently, image brightness was adjusted to further emphasize fine details and enhance contrast. Taking Guangdong Simiao rice as an example, the final brightness-adjusted image is shown in Figure 7 (a), demonstrating a notable improvement in brightness and visibility compared to Figure 6 (a). This enhancement facilitates more precise human and computational analysis.

To simulate real-world conditions where multiple rice varieties may be intermixed, we employed a randomized mixing approach to construct a test dataset for rice identification. This dataset enables the evaluation of model accuracy in classifying rice varieties under practical scenarios.

In summary, the grayscale and brightness-adjusted rice images effectively highlight key visual features such as grain outline, internal texture, chalkiness, and translucency. These enhancements ensure that the images meet the requirements for classification and recognition tasks, making them suitable as input data for relevant deep learning models.

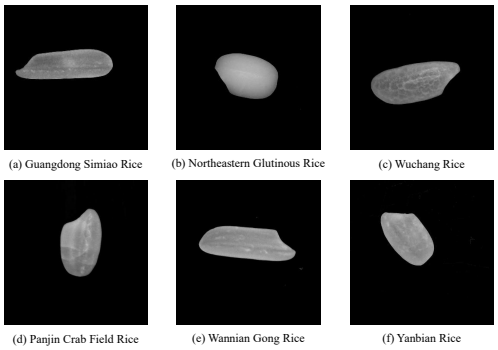

Figure 6: Example images of 6 types of Chinese rice sampled from our manually-collected dataset after grayscale processing: (a) Guangdong Simiao Rice, whose abbr is GD; (b) Northeastern Glutinous Rice, whose abbr is NM; (c) Wuchang Rice, whose abbr is WC; (d) Panjin Crab Field Rice, whose abbr is PJX; (e) Wannian Gong Rice, whose abbr is WN; (f) Yanbian Rice, whose abbr is YB.

A.2 Image Processing for Completeness Evaluation

For processing the broken rice image dataset, the same initial procedures were applied. In grading the completeness of rice grains, shape and contour features play a crucial role. Compared to grayscale images, binarized images more clearly delineate the shape and contours of rice grains, thereby improving the accuracy of subsequent classification algorithms. For instance, the preliminary binarized image of Guangdong Simiao rice, shown in Figure 7 (b), exhibits significantly more distinguishable geometric features than the original grayscale image in Figure 6 (a).

However, several small, unwanted white regions were present in the images. To simplify the dataset and retain only the significant and relevant areas, all white regions with a pixel area smaller than 40,000 were filled with black. Additionally, to enhance dataset quality, a median filtering algorithm was applied to remove edge noise, as illustrated in Figure 7 (c).

In summary, the rice images processed using the methods described in previous contexts, followed by binarization and noise reduction techniques, exhibit enhanced shape and contour clarity. These improvements contribute to more effective subsequent quality evaluation of rice grains.

A.3 Image Processing for Chalk Recognition

In studies that evaluate rice chalk using machine vision, accurate chalk discrimination requires determining a specific light intensity. To investigate the impact of different light intensities on the recognition algorithm, this study established five distinct brightness levels. Using Guangdong Simiao rice as an example, Figure 7 (d) and Figure 7 (e) illustrate the results obtained at brightness levels 2 and 5, respectively.

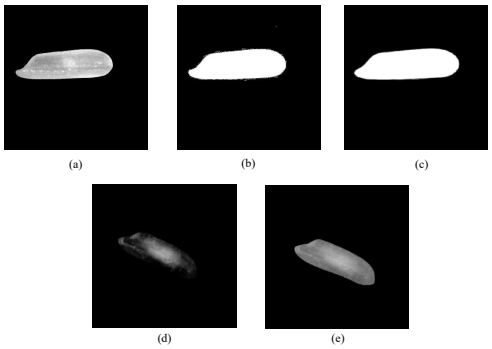

Figure 7: Example images in image processing: (a) an example image of Guangdong Simiao Rice after brightness enhancement; (b) an example image of Guangdong Simiao Rice after preliminary binarization processing; (c) an example image of Guangdong Simiao Rice after removing small domains and implementing median filtering; (d) an example image of Guangdong Simiao Rice after grayscale processing at the brightness of level 2, which is a low brightness; (e) an example image of Guangdong Simiao Rice after grayscale processing at the brightness of level 5, which is a high brightness.

Table 6: General rice quality standards of different rice varieties based on GB/T 1354 - 2018 Rice (ZHANG et al. 2021).

Rice Variety	Evaluated Level	Broken Rice Rate	Small Broken Rice Rate	Chalk Rate	Admixture Rate
Indica Rice	1	≤15.0%	≤1.0%	≤2.0%	≤5.0%
	2	≤20.0%	≤1.5%	≤5.0%	
	3	≤30.0%	≤2.0%	≤8.0%	
Japonica Rice	1	≤10.0%	≤1.0%	≤2.0%	≤5.0%
	2	≤15.0%	≤1.5%	≤4.0%	
	3	≤20.0%	≤2.0%	≤6.0%	
Glutinous Rice from Japonica branch	1	≤10.0%	≤1.5%	/	≤5.0%
	2	≤15.0%	≤2.0%		
Glutinous Rice from Indica branch	1	≤15.0%	≤2.0%	/	≤5.0%
	2	≤25.0%	≤2.5%		

B Detailed Descriptions about GB/T 1354 - 2018 Rice

B.1 The Standard of Rice Completeness

According to the classification criteria outlined in the Chinese National Standard GB/T 1354-2018 Rice, whole grains are defined as rice grains that remain intact in all parts except the embryo. Broken rice, on the other hand, refers to rice grains that are incomplete and have a length less than three-quarters of the average length of whole grains from the same batch, as retained on a 1.0 mm round-hole sieve. Within the category of broken rice, sizeable broken rice includes incomplete grains with a length less than three-quarters of the average length of whole grains from the same batch but retained on a 2.0 mm round-hole sieve. Tiny broken rice refers to fragments that pass through a 2.0 mm round-hole sieve but are retained on a 1.0 mm round-hole sieve.

To facilitate a more detailed analysis and assessment of broken rice in rice samples, we classified the rice into three categories based on the degree of completeness as defined by the standard: whole grains (with a length not less than three-quarters of the average length of whole grains from the same batch), sizeable broken rice, and tiny broken rice. Subsequently, the rice samples were labeled according to the identification criteria specified in the standard document.

For experimental validation, we used a sample of no less than 10 g of rice and conducted broken rice classification tests. The ratio of sizeable broken rice to tiny broken rice was calculated following the method prescribed in the standard, as described by Equation 1.

$$X_1 = \frac{m_1}{m} \times 100\% \tag{1}$$

where X_1 represents the broken rice rate, m_1 is the mass of broken rice in grams (g), and m is the mass of the sample in grams (g).

Referring to the calculation method for rice's broken rice rate in the standard document, we calculate it using Equation 2.

$$X_2 = \frac{m_2}{m} \times 100\% \tag{2}$$

where X_2 represents the broken rice rate, m_2 is the mass of broken rice, where broken rice includes the mass of sizeable and small broken rice, measured in grams (g), and m is the mass of the sample in grams (g).

B.2 The Standard of Rice Chalk

According to the detection method for rice chalk specified in the Chinese National Standard GB/T 1354-2018 Rice, the procedure begins by selecting 100 intact rice grains (n0) from the suspended rice sample. From this selection, the chalky rice grains (n1) are identified and separated. Subsequently, ten chalky rice grains are randomly selected from this group for further analysis. These selected grains are placed flat and observed from the front to visually assess the chalky projection area as a percentage of the total projection area of the intact rice grains. The average of these percentages is then calculated to quantify the degree of chalkiness. The chalk content is calculated using Equation 3, and the result is expressed as a percentage (%).

$$D = W_D \times \frac{n_1}{n_0} \tag{3}$$

where D represents chalkiness, expressed in percentage (%), W_D represents the chalky size, also in percentage (%), n_1 is the number of chalky rice grains in the sample, and n_0 is the total number of rice grains in the sample.