

Received 27 October 2022, accepted 15 November 2022, date of publication 17 November 2022,
date of current version 28 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3223137

RESEARCH ARTICLE

Vision-Based Defect Classification and Weight Estimation of Rice Kernels

XIANG WANG¹, KAI WANG¹, XIAOHONG LI, AND SHIGUO LIAN, (Member, IEEE)

AI Innovation and Application Center, China Unicom, Beijing 100033, China

Corresponding authors: Shiguo Lian (liansg@chinaunicom.cn) and Kai Wang (wangk115@chinaunicom.cn)

ABSTRACT Rice is one of the main staple food in many areas of the world. The quality estimation of rice kernels are crucial in terms of both food safety and socio-economic impact. This was usually carried out by quality inspectors in the past, which may result in both objective and subjective inaccuracies. In this paper, we present an automatic visual quality estimation system of rice kernels, to classify the sampled rice kernels according to their types of flaws, and evaluate their quality via the weight ratios of the perspective kernel types. To compensate for the imbalance of different kernel numbers and classify kernels with multiple flaws accurately, we propose a multi-stage workflow which is able to locate the kernels in the captured image and classify their properties. We define a novel metric to measure the relative weight of each kernel in the image from its area, such that the relative weight of each type of kernels with regard to the all samples can be computed and used as the basis for rice quality estimation. Various experiments are carried out to show that our system is able to output precise results in a contactless way and replace tedious and error-prone manual works.

INDEX TERMS Rice defect classification, rice quality estimation, rice weight estimation, vision-based method.

I. INTRODUCTION

Rice is the staple food for people in many parts of the world, and its quality is related to food health as well as the economic interests of agricultural dealers, who make offer to the farmers according to the quality of the collected rice kernels. In the process of rice trading, the purchaser needs to evaluate the quality of the rice and then determine the purchase price. Historically, this job was largely relied on manual work. It requires professionals to manually pick and weigh different types of rice from randomly selected samples. Rice grades are assigned by calculating the weight ratio of defective rice in the overall samples. Manual rice quality estimation is quite tedious, as it requires experienced inspectors to identify and pick up the kernels with various defects one by one and weigh them carefully. The precision of the result is subject to the skill and conscientiousness of the inspectors [1].

With the development of computer vision technique in recent decades, lots of attempts have been made to classify rice types and defects automatically [2]. A wide range of

computer vision algorithms, from traditional geometric [3] to deep-learning based methods [4] have been utilized. However, some inherent problems in precise rice quality analysis still exist. There are many types of defects in rice. Some defective rice kernels are easy to identify, such as broken kernels and yellow-colored kernels. Some defective rice kernels are difficult to distinguish if the inspectors are not experienced professionals. These human-indiscernible defects are also difficult for vision-based methods. Most of the existing literature is aimed at the easily identifiable rice defects, and has not addressed the classification of rice kernel with dual properties. Meanwhile, weight estimation is a key step to automate the entire process. Whereas the use of visual method to estimate rice weight has been rarely studied.

To solve these problems which existing literature are not able to handle and replace traditional manual tasks with automated process, we propose a novel system for automatic quality estimation of rice kernels. We designed a hardware system to obtain high-quality images of rice kernels with little lighting interference and background noise. To detect and classify the rice kernels with various types of defects, we employed a multi-stage classification approach to perform

The associate editor coordinating the review of this manuscript and approving it for publication was Juan Wang¹.

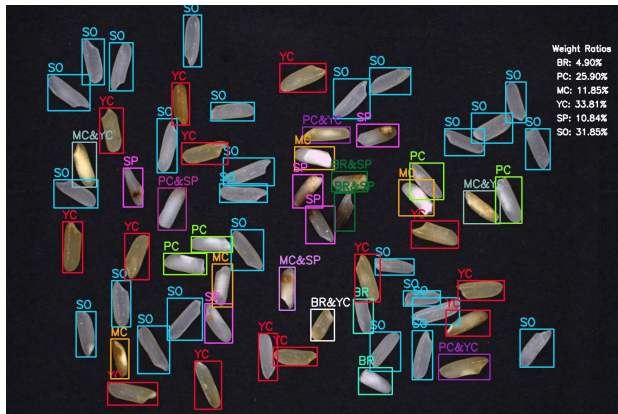


FIGURE 1. Rice quality estimation results with the proposed system. The type of each rice kernel is labeled on its bounding box, and the weight ratios are shown in the upper right area. PC, MC, YC, SP, BR and SO represent Partial-Chalky, Mass-Chalky, Yellow-Colored, Spotted, Broken and Sound kernels respectively.

multi-classification of rice flaws, such that a single kernel with dual defects can be identified and the classification accuracy for harder-to-distinguish defects is improved. Then we use a vision-based method to estimate the relative weights of kernels with each type of defect, by measuring its density via a novel weight-per-pixel metric and multiplying it with the segmented areas. An example of our result is shown in Fig. 1. With such workflow, the quality of the rice kernels can be automatically analyzed and presented to the inspectors, and tedious manual works can be replaced.

The contributions of this paper include:

- The introduction of a multi-stage method for detection and classification of rice kernels with various defects;
- The proposal of a weight-per-pixel metric for the estimation of weight ratios of each type of kernels;
- The design of a complete system which integrates these approaches to produce high-precision rice quality evaluation results automatically.

The rest of the paper is organized as follows: Section II reviews existing methods for rice classification and defect detection. Section III introduces the details of the proposed system and methods. Experimental evaluation is presented in Section IV, and Section V gives the conclusions.

II. RELATED WORK

During the last decades, various approaches for automatic rice classification and quality analysis have been proposed.

A. RICE CLASSIFICATION

Most work adopts the following route: extracting features with image processing algorithms and then classifying the rice based on these features. Kuo et al. [5] classified 30 varieties rice grains using image processing and sparse-representation-based classification (SRC). They captured the rice grain images by microscopy and identified the varieties of the grains based on the morphological, color, and

textural traits of the grain body, sterile lemmas, and brush. The average classification accuracy is 89.1%. For rice kernel images captured by ordinary optical cameras, Kambo and Yerpude [6] distinguished the variety of Basmati Rice Grain using K-NN and Principal Component Analysis (PCA), after preprocessing the images with smoothing and segmentation techniques. But the overall accuracy is only 79%. To improve the classification accuracy, Mousavirad et al. [7] presented a rice classification algorithm with optimal morphological features and back propagation neural network-based (BPNN). Here, 18 morphological features were extracted, and 6 features were selected. Silva and Sonnadara [8] combined neural network (NN) with PCA to classify the rice seed varieties. 34 features were extracted by some pre-processing operations before PCA was applied to perform dimensionality reduction, and one individual neural network was created for each feature set. They all need hand-engineered features and extract them with image processing algorithms.

With the widespread application of deep learning and their excellent performance in image tasks, more and more rice classification work starts to adopt deep neural networks. Lin et al. [9] proposed a model using convolutional neural networks (CNN) for rice kernel classification which reached a 99.52% accuracy. However, they found the accuracy of classification is closely related to the preprocessing effect of the image, which was done with image enhancement operations, such as re-scaling, mean subtraction, and feature standardization. Qiu et al. [10] compared the performance of CNN, KNN, and SVM in variety identification of rice seed with hyperspectral images. The results showed that CNN is capable of analyzing spectral data and performed better than KNN and SVM in rice variety classification. Similar to this work, Chatnuntawech et al. [11] provided a rice classification algorithm for identification with the synergy between hyperspectral imaging and deep CNN. Although the classification accuracy was greatly improved over SVM, they depended on hyperspectral imaging which provides not only spatial information, but also spectral information. Patel proposed two methods for rice types classification with images of rice grains captured by a scanner. One used CNN with segmented rice images as input. Another used a pretrained VGG-16 model and transfer learning to achieve a better result. However, this all requires the rice grains to be segmented from the image first, and cannot process localization and classification by a single network.

B. RICE QUALITY ANALYSIS

Different solutions have been applied on rice grain analysis. These approaches can broadly be classified into geometric, statistical, and machine learning.

Geometric approaches consider morphological features as key factors to analysis. Ajay et al. [3] proposed a morphological method for the quality evaluation of milled rice, which only used shape descriptors and geometric features to detect broken kernels among milled rice samples. Asif et al. [12] used morphological features to determine the

quality of rice grains and distinguish five different varieties. These morphological features include major and minor axis length, eccentricity, perimeter, area and size of the grains. Mahale and Korde [1] applied image processing techniques to grade and evaluate rice grains based on grain size and shape, such as length, width and their ratio. In contrast, Ali et al. [13] proposed a low cost solution for rice quality analysis based on more features. They computed the average length, average width, the area and number of small rice grains, medium rice grains and broken rice grains to analysis rice quality. Although different morphological features were applied, they all need a series of complex image processing an preprocessing.

Statistical approaches primarily focus on summarizing data and making inferences according to the population. Mahajan and Kaur [14] proposed a method of quality analysis for three types of Indian Basmati rice grains including normal grains, long grains and small grains. They applied morphological closing and opening operations and top-hat transformation to calculate the length of the major and minor axes of rice. Histogram for major axis length and minor axis length was used to compute threshold values for distinguishing between normal grains, small grains, and long grains. For rice quality analysis, they only considered the length of the rice kernels.

With the successful application of machine learning and its excellent performance, Ngampak and Piamsa-Nga [15] proposed a method for finely classifying broken rice into small broken, broken, big broken and head rice. They used Least-Square Support Vector Machine (LS-SVM) with Radius Basis Function (RBF) kernel as their classifier. Kaur et al. [16] divided rice kernels into four grades using Multi-Class SVM. They categorized rice into head rice, broken rice and brewers according to the kernel shape, length and chalkiness. Traditional machine learning methods are difficult to achieve high identification accuracy for complicated rice defective. Reference [17] applied neural network and image processing approach for rice grain identification and grading on three varieties of Indian rice. Rice images, obtained from a flatbed scanner, were processed with a smoothing filter, a binarizing operation, and the watershed method. The length, width, and perimeter of the rice grains were extracted and input neural network to do classification. The neural network they applied was simple, which only contained one hidden layer. It was able to classify rice into sound, cracked, chalky, broken and damaged kernels. Agustin and Oh [18] proposed an automatic quality evaluation framework for milled rice kernels, in which a probabilistic neural network (PNN) classifier is used to detect defective rice including discolored, chalky, broken, immature, red, and damaged kernels. Feature extraction was still an important step for their classification. They used 24 color features as the input of the classification network. At the same time, a linear regression model was developed for estimating kernel weight with a given blob area. The constant bias in their regression model can counteract offset, but it can also make the estimated weight

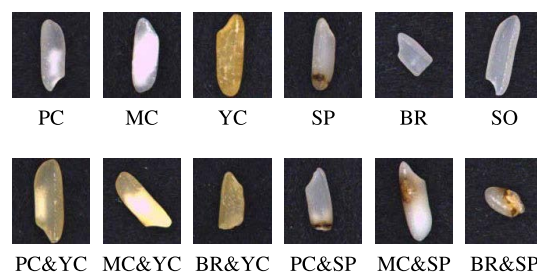


FIGURE 2. Example images of different rice kernel types. Top row: rice kernels with single property. Bottom row: rice kernels with dual properties.

negative, like the estimated weight of brewer kernels in their experiment.

While there are many different methods for detecting rice with different defects, none of them have classified rice with dual properties.

III. METHOD

In this section, the rice kernel quality evaluation problem will be formulated first. Then the system workflow, hardware setup for rice image acquisition and the methods for rice kernel classification and weight ratio estimation will be introduced in detail.

A. PROBLEM STATEMENT

In this paper, we focus on the detection of the main types of kernel flaws for new rice, including Partial-Chalky kernels (PC), Mass-Chalky kernels (MC), Yellow-Colored kernels (YC), Spotted kernels (SP), and Broken kernels (BR). A chalky kernel refers to a kernel with opaque white parts in endosperm. Depending on the percentage of this area with regard to the whole kernel, it is classified into PC ($\leq 50\%$) and MC ($> 50\%$). As its name suggests, a YC kernel refers to a kernel whose body is yellow. If the rice kernel has disease spot on the surface, it is regarded as a spotted kernel. Broken kernels are small pieces or particles of kernels whose length are less than $2/3$ of the average length of the whole kernels. If a kernel has none of the above defects, it is considered as a perfect kernel, also known as a Sound Kernel (SO).

It is also observed that some rice kernels could have dual properties: a partial-chalky kernel may also be yellow-colored, a spotted kernel may be incomplete (a broken kernel), etc. Common property combinations include: Partial-Chalky and Yellow-Colored (PC&YC), Mass-Chalky and Yellow-Colored (MC&YC), Broken and Yellow-Colored (BR&YC), Partial-Chalky and Spotted (PC&SP), Mass-Chalky and Spotted (MC&SP), Broken and Spotted (BR&SP). Kernels with other flaw combinations are rarely seen and thus can be ignored.

We need to classify the rice kernels with one of the above-mentioned single or dual properties, and the rice quality is estimated by computing the ratio of the weight of each type of kernels with regard to the sum of them. Note that only the weight ratio of kernels with each single types will

be calculated, and the weight of a kernel with dual properties will be summed up to the total weights of kernels with each of the two properties respectively.

The weight ratio can thus be calculated as follows. Given a set of rice kernels $\{K_i\}$, where $i \in \{PC, MC, YC, SP, BR, SO, PC\&YC, MC\&YC, BR\&YC, PC\&SP, MC\&SP, BR\&SP\}$, the total weight of all kernels is $\sum W_i$ (where W_i refers to the weight of the kernel type i), and the weight ratio R_t of each type of rice kernel, where $t \in \{PC, MC, YC, SP, BR, SO\}$, is formulated as:

$$\begin{aligned}
 R_{PC} &= \frac{W_{PC} + W_{PC\&YC} + W_{PC\&SP}}{\sum W_i}, \\
 R_{MC} &= \frac{W_{MC} + W_{MC\&YC} + W_{MC\&SP}}{\sum W_i}, \\
 R_{YC} &= \frac{W_{YC} + W_{PC\&YC} + W_{MC\&YC} + W_{BR\&YC}}{\sum W_i}, \\
 R_{SP} &= \frac{W_{SP} + W_{PC\&SP} + W_{MC\&SP} + W_{SP\&BR}}{\sum W_i}, \\
 R_{BR} &= \frac{W_{BR} + W_{BR\&YC} + W_{BR\&SP}}{\sum W_i}, \\
 R_{SO} &= \frac{W_{SO}}{\sum W_i}.
 \end{aligned} \tag{1}$$

As the weights of kernels with dual properties are calculated twice, the sum of all R_t will probably not be 1. However it can reflect the amount of the rice kernels with dual defects. Under the condition that the quantity of various types of rice kernels is constant, the more rice with dual properties, the greater the weight ratio of defective rice.

B. SYSTEM WORKFLOW

The workflow of our proposed system is illustrated in Fig. 3. First, a certain amount of samples are randomly selected from the rice kernels to be evaluated. Then RGB images containing all rice samples are captured. On the basis of the captured images, different types of rice kernels are located and identified, and the weight of these rice is estimated based on the projected area of the kernels in the RGB image. Finally, the weight ratio of various types of defective rice in the rice samples is calculated to judge the quality grade of the rice. Among them, the localization, classification and weight estimation are all based on vision methods. It can be done automatically without human involvement.

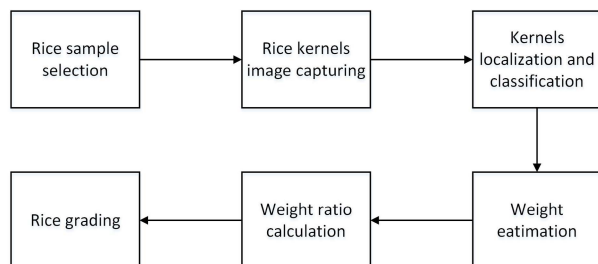


FIGURE 3. Workflow of the proposed rice quality estimation system.

C. HARDWARE SETUP

The hardware system for capturing the images of rice kernels is shown in Fig. 4. We used a white dome light as the light source to make the illumination evenly distributed and meanwhile avoid the influence of outside lightings. A black mat with rough material is used as the background to avoid light reflections. The kernels are placed on the mat manually by scattering through a griddle, such that they can be evenly distributed and have few overlaps. Overlapping and occlusion will affect the precise location and classification of rice kernels. A megapixel CCD camera and a lens with short focus length are used to capture the images of the kernels. The resolution of each image is 1280×960 . All these devices are fixed on a scaffold and are thus stable enough during the measurement process.



FIGURE 4. Outside (left) and inside (right) views of the hardware system.

The light source and camera are connected to a computing server via cables. After the kernels have been placed, the photograph process can be started by clicking a button, then the light will be turned on and the camera will capture the image and transmit it to the server for analysis.

D. MULTI-CLASSIFICATION OF RICE KERNEL FLAWS

Rice kernels can only be accurately distinguished by professionals in manual inspection work in the past. Traditional image classification methods relying on image features are neither effective for the recognition of rice flaws, as the boundaries between some types of rice kernels are not clear. Fig. 5 shows some rice kernels with different transparency. It is difficult to quantize the value of transparency and classify the rice kernels as chalky based on a given threshold. Different rice varieties have different appearance. Even for the same variety of rice from the same place of origin, the appearances of the kernels, such as the glossiness and transparency, will probably be different if stored in different environments. Fig. 6 shows an example which is difficult to effectively discriminate. According to China Standard GB/T 1354-2018 [19], a chalky kernel contains opaque white parts including white belly, white core, and white back. The rice kernels shown on the left side of Fig. 6 also have white opaque portions, but these white opaque portions are embryos, not chalky lumps. Whereas, the rice kernels shown on the right side of Fig. 6 are chalky kernels. This means it is almost impossible to determine whether the rice kernel is chalky based on solely the detected white opaque patches.



FIGURE 5. Rice kernels with different transparency.



FIGURE 6. Left: normal kernels and yellow kernel with white embryo; Right: chalky kernels.

To solve these problems, we employ deep neural networks to classify the rice flaws. It is straight forward to use the object detection algorithm–Yolo [20], to detect rice kernels in the whole image in one step. While the obtained kernel location is accurate, the classification precision of some types of rice is not high, especially for the partial-chalky and mass-chalky kernels. Therefore, we separately process PC and MC kernels with another Yolo and a classification network based on grayscale images inspired by the works [16], [17]. We designed a multi-stage workflow for rice kernel localization and flaw classification, which is shown in Fig. 7. The input image and the step-by-step outputs of the proposed method is demonstrated in Fig. 8.

This multi-stage workflow is divided into two branches (see Fig. 7). One is the detection of YC, SP, BR, and SO kernels. The other is the localization and classification of PC and MC kernels. The output of the two branches are fused together to form the final result. The whole process includes five stages: the detection of YC, SP, BR, and SO kernels, the detection of chalky kernels, segmentation, the classification of PC and MC kernels, and the fusion stage. The details of each stage will be described below.

1) DETECTION OF YC, SP, BR, AND SO KERNELS

The detection of YC, SP, BR, and SO kernels is implemented by Yolov5 [21], which shows advantages in both detection accuracy and speed over other methods.

As discussed in Section III-A, the output class of a rice kernel may either be one of the 6 classes with single properties, or those classes with dual properties, and all other possibilities such as a class with a property combination which are not mentioned before should be excluded. This is inherently a multi-classification problem. A Dual-property rice can be labeled as a new category which is independent of the single properties, or as one entity with two categories, or as two entities with different single class. As the number of the samples of dual-property rice is limited, the data imbalance will result in low detection accuracy of dual-property kernels if they are treated as independent categories. Both the other two options seem more suitable comparatively. Since Yolov5 supports multi-classification, annotating a kernel to two categories does not affect the calculation of the loss function, and it can be propagated backward during the network training.

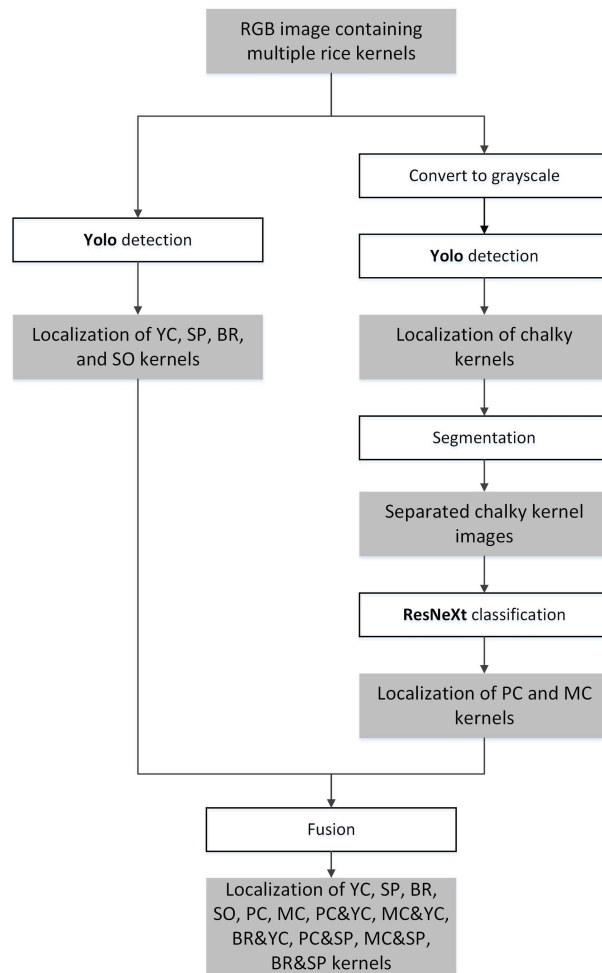


FIGURE 7. Multi-stage workflow for rice kernel detection and rice flaw classification. The applied neural networks are highlighted in bold.

On the other hand, Yolov5 predicts multiple bounding boxes per grid cell, so it can detect the same kernel repeatedly. Fig. 9 illustrates the differences between the annotations of the latter two schemes. On the left side of the figure, both kernels of dual-property have only one bounding box, but each bounding box is annotated as two categories. On the right side, the same two kernels of dual-property are both surrounded by two bounding boxes, each of which corresponds to one class. Here we chose to label dual-property rice kernels as two entities separately considering the convenience of annotation.

The trained Yolo model (denoted as Yolo-model-1) can locate and classify YC, SP, BR, SO, and other kernels, which is shown in Fig. 8-a. Rice kernels other than YC, SP, BR, and SO are classified as 'others'.

2) DETECTION OF CHALKY KERNELS

Fig. 10 shows the comparison of a color image and a gray image (in which the white chalky lumps are converted to black) of rice kernels. There are partial-chalky kernels, mass-chalky kernels, and also chalky kernels with dual properties, such as PC&YC and PC&SP kernels. It can be seen that the

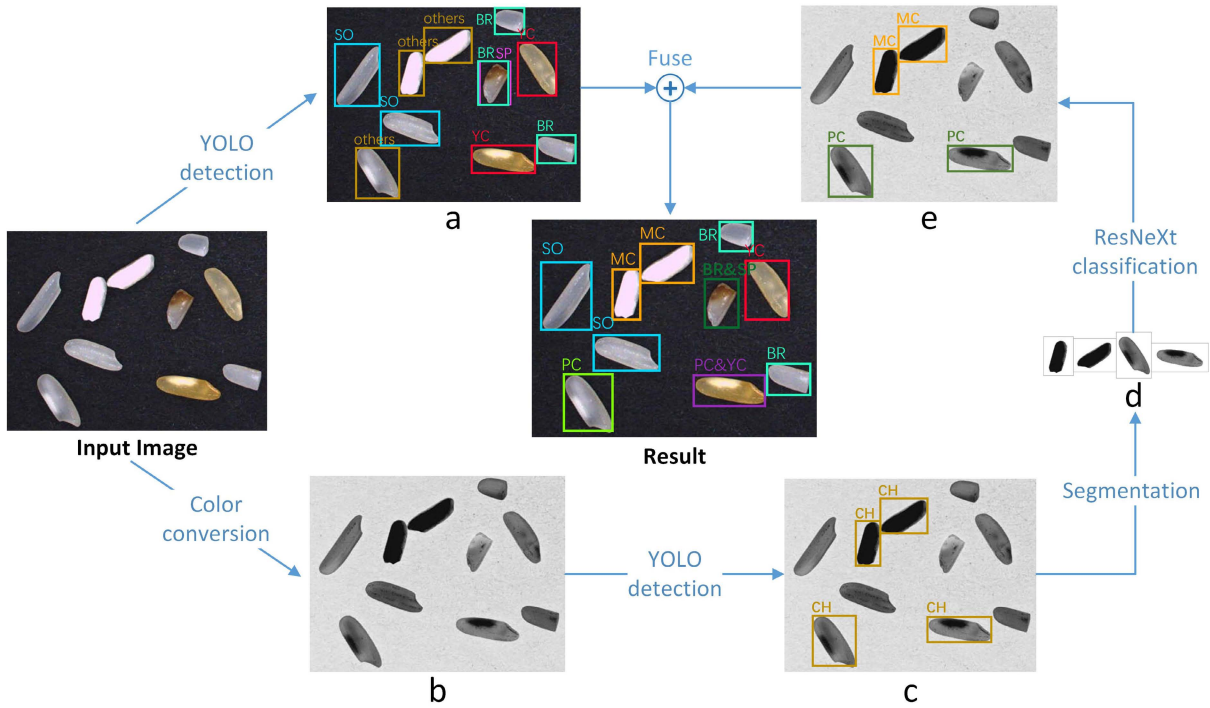


FIGURE 8. Input and the step-by-step outputs of the proposed method.

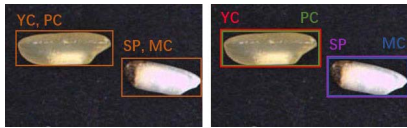


FIGURE 9. Left: each kernel has one bounding box and two classes; Right: each kernel has two bounding boxes, and each bounding box belongs to one category.

chalky parts are more obvious in the gray image which completely retains the chalky features and eliminate the occlusion of other features. The small white belly and white back in kernel 1 and 2 are not obvious in the RGB image, but are clearly shown in the grayscale image. Rice kernel 3 and 4 are yellow-colored with chalky lumps. Although the grayscale image loses their color information, the chalky patches are completely preserved and more prominent. Kernel 5 is a dual-property rice with disease spot and chalky lumps. In the grayscale image, it can be treated normally as other kernels to classify it as a PC kernel according to the area of its black parts.

The gray image in Fig. 10 is obtained using the following formula:

$$Y = 255 - 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B, \quad (2)$$

where Y is the grayscale value. R , G , and B are the three channel values of the color image respectively.

Taking the converted gray image as the input of Yolo detection network (Fig. 8-b), we trained another Yolo model

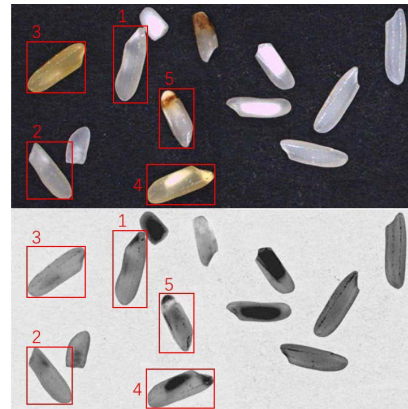


FIGURE 10. Comparison of the original RGB image and our converted gray image.

(denoted as Yolo-model-2), which is responsible for detection chalky kernels in gray images (Fig. 8-c).

3) SEGMENTATION

After detecting the chalky kernels from the gray images, we need to segment it out from each detected bounding areas. Here we cutout the pixels of chalky kernels in gray images based on the kernels contours detected in color images, since the boundaries of rice kernels are easier to accurately extract in color images.

As has been introduced, we choose a black material as the photo background as it makes easier to distinguish the

kernels. However, this also resulted in the failure of simple contour extraction algorithms to extract the contours of spotted kernels which contain black disease spots. Therefore, we adopt the following processing for kernel contour extraction:

$$VS_img = \max(\text{erode}(S_img), V_img),$$

$$\text{Contours} = \text{findContours}(\text{Morph_Open}(VS_img)), \quad (3)$$

where S_img and V_img are the saturation and hue channels of the image which is converted from RGB color space to HSV (for hue, saturation, value) color space [22], and VS_img is the maximum image of eroded S_img and V_img . Contours can then be found by applying dilation, erosion and contour extraction operations provided in OpenCV [23]. Fig. 11 shows the comparison of the contours of two spotted kernels extracted from gray image and VS_img respectively. As can be seen, the lesion parts which are missing in the contours extracted with traditional method are contained in the results of our method.

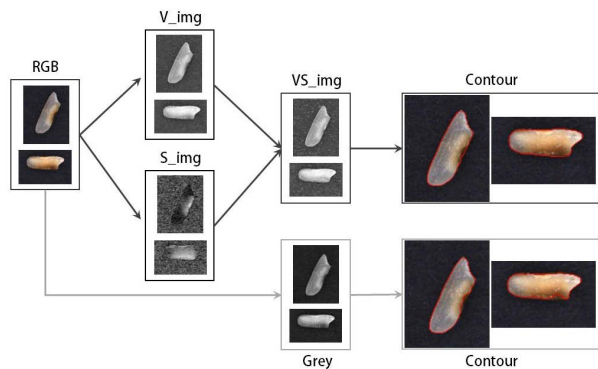


FIGURE 11. Comparison of contours extracted from gray image and our VS_img .

The segmentation is operated as follows: If the center point of the extracted kernel contour is close to a center point of the detected bounding boxes of chalky kernels, we label the pixels inside the contour belongs to a chalky kernel and cutout them to a blank template to construct individual image for each chalky kernel. An example is shown in Fig. 8-d.

4) CLASSIFICATION OF PC AND MC KERNELS

We employ a classification network, ResNeXt [24], to classify partial-chalky kernels and mass-chalky kernels. It does not require a long time training and the classification accuracy is satisfactory. Besides, its inference is fast. An example output of this stage is shown in Fig. 8-e.

5) FUSION

After the above steps, we have detected YC, SP, BR, and SO kernels from color images, and PC and MC kernels from gray images. As these detections are independent of each other, they need to be fused and undefined property combinations should be filtered out. As discussed in Section III-A, the

Algorithm 1 Fusion

Input: {Kernel K_i , Bounding box B_i , Class C_i , Confidence $Conf_i$ }, IOU threshold t

Output: {Kernel K_i , Class CC_i }

```

1: for each bounding box  $B_i$  do
2:   calculate its IOU with other bounding boxes
3:   if no IOU is larger than  $t$  then
4:      $K_i$  is classified into  $C_i$ ,  $CC_i = C_i$ 
5:   else if only one IOU with  $B_j$  is larger than  $t$  then
6:     if  $C_i \& C_j$  is within the predefined range then
7:        $K_i$  is classified into  $C_i \& C_j$ ,  $CC_i = C_i \& C_j$ 
8:     else
9:       if  $Conf_i$  is larger than  $Conf_j$  then
10:         $K_i$  is classified into  $C_i$ ,  $CC_i = C_i$ 
11:      else
12:         $K_i$  is classified into  $C_j$ ,  $CC_i = C_j$ 
13:   else
14:     find the largest two  $Conf_1$  and  $Conf_2$ 
15:     if  $C_1 \& C_2$  is within the predefined range then
16:        $K_i$  is classified into  $C_1 \& C_2$ ,  $CC_i = C_1 \& C_2$ 
17:     else
18:       if  $Conf_1$  is larger than  $Conf_2$  then
19:         $K_i$  is classified into  $C_1$ ,  $CC_i = C_1$ 
20:       else
21:         $K_i$  is classified into  $C_2$ ,  $CC_i = C_2$ 
22: return { $K_i$ ,  $CC_i$ }

```

output class of a rice kernel may either be one of the 6 classes with single properties, or those classes with dual properties, and all other possibilities such as a class with a property combination which are not mentioned before can be excluded.

The fusion operation is processed as follows: For each bounding box, we calculate its IOU (Intersection over Union) with other bounding boxes to get proposals of overlapping bounding boxes. If the bounding box has no overlaps with others, the rice kernel surrounded is classified into the class corresponding to this bounding box, i.e., one of the 6 classes with single properties. If the bounding box overlaps with another bounding box and the combination of the properties corresponding to the two bounding boxes is within the predefined range, the rice kernel surrounded by this bounding box is classified into a class with these dual properties. If the bounding box overlaps with more than one bounding box or the combination of the properties corresponding to the overlapped bounding boxes is not defined, we will rank the proposals with their confidence, filter out mutually exclusive terms with low confidence, and obtain the desired properties combination. The pseudo code of the fusion algorithm is shown in Algorithm 1. An example of the fusion result is shown in the center of Fig. 8.

E. WEIGHT RATIO ESTIMATION

After classifying each kernel, we need to estimate its weight and compute the ratio of weight for each kernel

Algorithm 2 Weight Ratio Calculation

Input: {Kernel K_i , Area A_i , Class CC_i }, weight-per-pixel metric ρ_t ($t \in \{PC, MC, YC, SP, BR, SO\}$)

Output: weight ratio R_t ($t \in \{PC, MC, YC, SP, BR, SO\}$)

```

1: initialize  $W_t$  ( $t \in \{PC, MC, YC, SP, BR, SO\}$ ) and  $W_{all}$ 
   to zero
2: for each  $\{K_i, A_i, CC_i\}$  do
3:   if  $CC_i$  is equal to  $C_t$  then
4:      $W_t + = \rho_t \cdot A_i$ 
5:      $W_{all} + = \rho_t \cdot A_i$ 
6:   else if  $CC_i$  is equal to  $C_m \& C_n$  then
7:      $W_m + = \rho_m \cdot A_i$ 
8:      $W_n + = \rho_n \cdot A_i$ 
9:      $W_{all} + = 0.5 \cdot \rho_m \cdot A_i + 0.5 \cdot \rho_n \cdot A_i$ 
10:  else
11:    pass
12: for each  $W_t$  do
13:    $R_t = W_t / W_{all}$ 
14: return  $\{R_t\}$ 

```

type. Traditionally, the inspectors weigh the kernels with a high-precision digital balance, which is quite tedious and time-consuming. Here we try to estimate the weight ratio using a vision-based approach. A formal way to do that is to get the volume of each kernel and multiply it with the averaged density of such kernel type. However, as the kernel size is quite small and most of its body is transparent, it is difficult to reconstruct it and compute its volume with vision-based methods. We noticed that the shape of the rice kernels is roughly the same except for the broken ones, so we imagined that the rice weight could be estimated in the 2D projection space. Moreover, what matters on the rice quality evaluation is the ratio of the kernel weight. That is, if the weight estimation for different types of rice have similar bias, the weight ratio will not be significantly affected.

Inspired by these observations, we propose to compute the weight ratio of rice kernels via their 2D projected area instead of volume. Specifically, we propose a new weight-per-pixel metric ρ_t ($t \in \{PC, MC, YC, SP, BR, SO\}$), which represents the averaged density of rice kernels of type t per pixel, and the relative weight of rice kernels of type t can be obtained by multiplying it with the projected area on the 2D image in pixels. After collecting the weight of all types of rice kernels in this way, the ratios are computed via Equation 1. Pseudo code for the calculation of the weight ratios are shown in Algorithm 2.

To get the area of a kernel, we use the contour detection method mentioned previously and calculate the area inside the kernel contour. It should also be pointed out that as the distance between the lens and the background mat is fixed, the projection matrix is constant, hence the scales of all the kernels are uniform.

To get the weight-per-pixel ρ_t , we first collected a large quantity of samples K_t for each type of kernels. These

TABLE 1. Rice density calculation.

	SO	PC	MC	YC	SP	BR
Kernel amount	1227	952	1088	810	1014	1425
Weight (g)	21.9459	17.4861	15.8578	9.2693	10.3913	12.6766
Area (pixels)	4.11E6	3.22E6	2.98E6	1.84E6	2.1E6	2.42E6
Density (g/pixel)	5.32E-6	5.42E-6	5.31E-6	5.03E-6	4.93E-6	5.23E-6

samples only consisted of rice kernels with single property. The total weight values W_t of kernels K_t for type t were obtained by using a professional electronic balance, whose accuracy is 0.1 mg. The projection area of each type of rice was measured multiple times. We used the average area A_t and the weight value of K_t to calculate the density ρ_t , which is shown in Equation 4.

$$\rho_t = \frac{W_t}{A_t}. \quad (4)$$

It is observed that for a certain category of rice, the density fluctuates little. Whereas the density needs to be re-calculated when estimating other rice categories. For dual-property kernels, their weights will be calculated twice, and the results will probably vary as the densities used in the two calculations are different. We take half of each as the weight of the dual-property rice kernel when calculating the total weight of the rice samples. Because dual-property kernels only account for a small portion of all the kernels, this weight error has little effect on the overall result and can therefore be ignored.

IV. EXPERIMENTAL EVALUATION**A. TRAINING SETUP****1) TRAINING DATA**

Using the hardware setup introduced earlier, we collected a total of 322 images as training data, validation data and test data. Each image contains a variable number of rice kernels, mostly between 40 and 80. The total number of rice kernels is over 20,000. All data are annotated by professionals with many years of experience. They are used for the training and testing of our rice detection and classification model. The dataset is not publicly available due to commercial interest, but portion of it can be requested from the corresponding author.

In order to calculate the density ρ_t for each type of kernels via equation 4, we first invited experts to manually classify some rice samples into different types and weighted them separately with a professional electronic balance, which has a measurement accuracy of 0.1 mg. We then took photos of these rice samples and calculated their projection area. After that, the density of different types of rice kernels can be calculated. Table 1 shows the number of kernels we collected, their weighed weight, projection area and the rice density calculated.

2) TRAINING DETAILS

Our detection and classification networks are implemented with Pytorch [25]. Both the detection model Yolo-model-1 and Yolo-model-2 are trained for 300 epochs from the pre-trained yolov5x model using SGD optimizer [26], with learning rate of 0.01 and batch size of 16. Then the trained models are fine-tuned for 300 epochs with a small learning rate 0.0032. The size of the input image is 640×640 . The classification network ResNeXt is trained for 100 epochs from a pretrained model using Adam optimizer [27], with learning rate of 0.001 and batch size of 32. The size of the input image is 224×224 .

B. EVALUATION

1) CLASSIFICATION RESULTS

A 5-fold cross validation is applied to test our multi-stage workflow for the multi-classification of rice kernels. The total 322 rice images were shuffled randomly and split into training data (292 images), validation data (15 images), and test data (15 images) for 5 times. We trained and tested our model 5 times with these datasets. Table 2 shows the test result, where the ground truth is obtained by letting professional inspectors to classify the kernels one by one. We compare our model with the native Yolov5 [21], Yolov5 plus postprocessing (Yolov5+p), EfficientDet-D2 [28] (EDet), EfficientDet-D2 plus postprocessing (EDet+p), DynamicHead+Resnet50+ATSS [29] (DyHead), and DynamicHead+Resnet50+ATSS plus postprocessing (DyHead+p). This postprocessing refers to the operation of filtering out impossible property combinations from the detection result, as described in Section III-D.

TABLE 2. 5-fold classification results of Yolo, Yolo plus postprocessing, EfficientDet, EfficientDet plus postprocessing, DynamicHead, DynamicHead plus postprocessing, and our model.

Precision(%)	SO	BR	PC	MC	YC	SP
Yolov5	88.20	93.77	75.76	91.11	93.22	88.35
Yolov5+p	92.28	96.34	81.63	95.89	94.86	90.49
EDet	89.61	96.27	78.00	82.95	97.30	94.07
EDet+p	92.24	96.95	81.43	86.22	98.11	95.54
DyHead	91.36	96.48	80.00	87.29	96.58	94.13
DyHead+p	92.67	97.03	81.93	88.62	97.69	94.59
Our model	94.25	97.83	86.58	86.99	97.61	93.07
Recall(%)	SO	BR	PC	MC	YC	SP
Yolov5	97.43	96.65	85.75	82.22	94.87	86.72
Yolov5+p	95.27	95.88	81.27	78.07	94.67	85.98
EDet	94.52	89.08	76.69	91.04	92.50	67.34
EDet+p	92.86	88.41	72.91	88.64	92.31	67.16
DyHead	93.94	91.95	75.64	83.56	83.63	68.08
DyHead+p	92.36	90.80	73.79	82.22	83.43	67.71
Our model	95.35	95.21	83.38	96.52	96.84	86.72
F1 Score	SO	BR	PC	MC	YC	SP
Yolov5	92.58	95.19	80.45	86.44	94.04	87.52
Yolov5+p	93.75	96.11	81.45	86.07	94.77	88.17
EDet	92.00	92.54	77.34	86.81	94.84	78.49
EDet+p	92.55	92.48	76.94	87.41	95.12	78.87
DyHead	92.63	94.16	77.76	85.38	89.64	79.01
DyHead+p	92.51	93.81	77.65	85.30	90.00	78.92
Our model	94.80	96.50	84.95	91.51	97.23	89.78

We calculate the classification precision, recall and F1 score of the six types of rice kernels. The dual-property kernels are counted into each single type. As can be seen from Table 2, the average precision of our model for the six types are all larger than 85%. The precision for sound kernels, broken kernels, yellow-colored kernels, and spotted kernels are over 90%. Compared with Yolov5, EfficientDet-D2, and DynamicHead+Resnet50+ATSS, our model performs the best in all types of rice classification. Although the classification precision of mass-chalky kernel, yellow-colored kernel, and spotted kernel is lower than other models, the recall rate is improved, and the F1 score is still the highest.

2) WEIGHT ESTIMATION RESULTS

Five sets of rice samples are used to test the effectiveness of the weight estimation method proposed in this paper. Each set contains a different amount of kernels in different types. First, the rice samples were classified using our trained multi-stage model. Then we estimated the weight of each type of rice kernels with their projected area and density. At last, the weight ratios were calculated according to equation 1. The rice samples have been sorted and precisely weighed by professionals, and we will compare our calculation result with these expert data.

Table 3 shows the evaluation result. It can be seen that all mean errors are less than 2%, and the maximum errors are not greater than 3%. In comparison, the error of the mass-chalky kernel is the largest, and the error of yellow-colored kernel is the smallest.

TABLE 3. Test results of weight ratio of defective rice.

	BR	PC	MC	YC	SP
mean error(%)	1.8926	1.4620	1.9038	0.4108	1.7482
max error(%)	2.4514	2.7315	2.9977	1.4765	2.4498

C. DISCUSSION

1) SINGLE DENSITY VS. VARIOUS DENSITY

According to our intuition, broken kernels and sound kernels have the same physical properties but different sizes, so their densities should be the same. But on the 2D projection surface, broken kernel of the same weight should have more edge parts than whole kernel, so the projected density of broken kernels should be less than that of the whole kernels. Chalky, yellow-colored and spotted kernels all have character changes, so their densities should be different from sound rice, and their 2D projection densities should also be different. But how much do the projected densities of different types of rice kernels differ? Can the same density be used to estimate the weight ratios for these types of rice kernels? To answer these questions, we conducted a test.

Some rice samples classified by experts are used for this test. We randomly divided each of the five types of defective rices into three groups. The remaining sound rice was divided into two groups. They occupied the majority of all

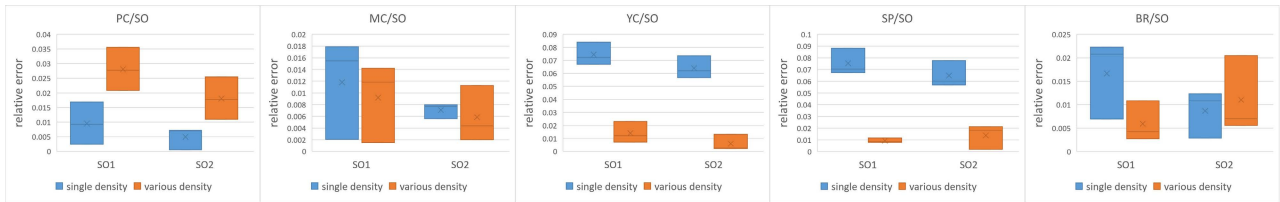


FIGURE 12. Box diagrams of the relative error of estimated weight percentage calculated with a single density and various densities. PC, MC, YC, SP, BR and SO represent Partial-Chalky, Mass-Chalky, Yellow-Colored, Spotted, Broken and Sound kernels respectively.

TABLE 4. The rice samples used for density test.

	kernel amount	accurate weight(g)	area (pixels)	estimated weight(g)	weight error(mg)
SO 1	148	2.6927	494365	2.6346	58.1
SO 2	338	5.9928	1111058	5.9211	71.7
PC 1	97	1.7933	330068.5	1.7913	2.0
PC 2	109	1.9313	360577.5	1.9569	-25.6
PC 3	127	2.3933	443462.5	2.4067	-13.4
MC 1	111	1.8700	348635.5	1.8513	18.7
MC 2	115	1.9726	362904.5	1.9271	45.5
MC 3	133	2.2357	417796.5	2.2186	17.1
YC 1	134	1.6182	322061.5	1.6199	-1.7
YC 2	149	1.8479	363798.5	1.8298	18.1
YC 3	144	1.8580	364007	1.8308	27.2
SP 1	109	1.3254	260486.5	1.2857	39.7
SP 2	139	1.6427	328200.5	1.6199	22.8
SP 3	144	1.7985	352404.5	1.7394	59.1
BR 1	98	0.9846	182024	0.9529	31.7
BR 2	129	1.2161	227903	1.1931	23.0
BR 3	137	1.3338	250334	1.3105	23.3

rice samples. Then the accurate weight values, projected area values and estimated weight were obtained respectively. Table 4 shows the detail.

If a single density value is used, the weight ratio is equal to the area ratio. We calculate the accurate weight percentage, area percentage and estimated weight percentage of the five types of defective kernels relative to the sound kernels, where the area percentage is equal to the estimated weight percentage when different types of rice kernels are supposed to have the same density. In order to analyze the effect of rice density on rice quality estimation, we compare the relative error of area percentage relative to accurate weight percentage and the relative error of estimated weight percentage relative to accurate weight percentage.

Figure 12 shows the box diagrams of the relative error calculated for the five types of defective rice. As can be seen from the chart, the error median line using the same density is higher than the error median line using respective densities for mass-chalky kernels, yellow-colored kernels, spotted kernels and broken kernels. Only the partial-chalky kernels have different result from our expectation, which means they are more suitable to apply the same density as sound kernels. We infer the reason is that the density of PC is very close to that of SO, since the chalky part in PC is small. Therefore, the error of the area ratio is smaller, and using an inaccurate projection density will introduce greater deviation.

V. CONCLUSION

In this paper, we propose an integrated system for rice quality estimation. A multi-stage workflow is employed to locate and classify the rice kernels with possibly overlapped types of flaws. A density per pixel metric is proposed to measure the weight ratios of kernels with different types, such that the quality estimation can be performed with a fully vision-based approach. We performed various experiments to show the advantage and usefulness of our system and proved that rice quality estimation can be carried out automatically and tedious manual works can be replaced. Further investigation of how to reduce the requirements on the quantity and quality of the training data will be a meaningful direction. Because inviting professionals to label data is often not available, and data labelled by non-professionals often has errors. Therefore, it is also important to improve the robustness of the network model to labeled data.

REFERENCES

- [1] B. Mahale and S. Korde, "Rice quality analysis using image processing techniques," in *Proc. Int. Conf. Conver. Technol.*, Apr. 2014, pp. 1–5.
- [2] S. B. Ahmed, S. F. Ali, and A. Z. Khan, "On the frontiers of Rice grain analysis, classification and quality grading: A review," *IEEE Access*, vol. 9, pp. 160779–160796, 2021.
- [3] G. Ajay, M. Suneel, K. K. Kumar, and P. S. Prasad, "Quality evaluation of Rice grains using morphological methods," *Int. J. Soft Comput. Eng.*, vol. 2, no. 6, pp. 35–37, 2013.
- [4] K. Kiratiranapruk, P. Temniranrat, A. Kitvimonrat, W. Sinthupinyo, and S. Patarapuwadol, "Using deep learning techniques to detect Rice diseases from images of Rice fields," in *Proc. Int. Conf. Ind., Eng. Other Appl. Appl. Intell. Syst.* Springer, 2020, pp. 225–237.
- [5] T.-Y. Kuo, C.-L. Chung, S.-Y. Chen, H.-A. Lin, and Y.-F. Kuo, "Identifying Rice grains using image analysis and sparse-representation-based classification," *Comput. Electron. Agricult.*, vol. 127, pp. 716–725, Sep. 2016.
- [6] R. Kambo and A. Yerpude, "Classification of basmati Rice grain variety using image processing and principal component analysis," 2014, *arXiv:1405.7626*.
- [7] S. J. Mousavirad, F. A. Tab, and K. Mollazade, "Design of an expert system for Rice kernel identification using optimal morphological features and back propagation neural network," *Int. J. Appl. Inf. Syst.*, vol. 3, pp. 33–37, Jul. 2012.
- [8] C. S. Silva and D. Sonnadara, "Classification of Rice grains using neural networks," *Tech. Rep.*, 2013.
- [9] P. Lin, Y. Chen, J. He, and X. Fu, "Determination of the varieties of Rice kernels based on machine vision and deep learning technology," in *Proc. 10th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2017, pp. 169–172.
- [10] Z. Qiu, J. Chen, Y. Zhao, S. Zhu, Y. He, and C. Zhang, "Variety identification of single Rice seed using hyperspectral imaging combined with convolutional neural network," *Appl. Sci.*, vol. 8, no. 2, p. 212, 2018.
- [11] I. Chatnuntawech, K. Tantisantisom, P. Khanchaitit, T. Boonkoom, B. Bilgic, and E. Chuangsuwanich, "Rice classification using spatio-spectral deep convolutional neural network," 2018, *arXiv:1805.11491*.

- [12] M. J. Asif, T. Shahbaz, S. Tahir Hussain Rizvi, and S. Iqbal, "Rice grain identification and quality analysis using image processing based on principal component analysis," in *Proc. Int. Symp. Recent Adv. Electr. Eng. (RAEE)*, Oct. 2018, pp. 1–6.
- [13] S. F. Ali, H. Jamil, R. Jamil, I. Torij, and S. Naz, "Low cost solution for Rice quality analysis using morphological parameters and its comparison with standard measurements," in *Proc. Int. Multi-Topic Conf. (INMIC)*, Nov. 2017, pp. 1–6.
- [14] S. Mahajan and S. Kaur, "Quality analysis of Indian basmati Rice grains using top-hat transformation," *Int. J. Comput. Appl.*, vol. 94, no. 15, pp. 42–48, May 2014.
- [15] D. Ngampak and P. Piamsa-nga, "Image analysis of broken Rice grains of khao daw Mali rice," in *Proc. 7th Int. Conf. Knowl. Smart Technol. (KST)*, Jan. 2015, pp. 115–120.
- [16] H. Kaur and B. Singh, "Classification and grading Rice using multi-class SVM," *Int. J. Sci. Res. Publications*, vol. 3, no. 4, pp. 1–5, 2013.
- [17] B. Verma, "Image processing techniques for grading & classification of Rice," in *Proc. Int. Conf. Comput. Commun. Technol. (ICCCT)*, Sep. 2010, pp. 220–223.
- [18] O. C. Agustin and B.-J. Oh, "Automatic milled Rice quality analysis," in *Proc. 2nd Int. Conf. Future Gener. Commun. Netw.*, Dec. 2008, pp. 112–115.
- [19] R. Tang, *National Standard for Rice*, China's State Administration for Market Regulation, Standard GB/T 1354-2018, 2019.
- [20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.
- [21] G. Jocher, "Ultralytics/YOLOv5: V6.0–YOLOv5n 'Nano' models, Roboflow integration, tensorflow export, OpenCV DNN support," Tech. Rep., 2021, doi: [10.5281/zenodo.5563715](https://doi.org/10.5281/zenodo.5563715).
- [22] D. Li, "Hsv color space," in *Encyclopedia of Microfluidics and Nanofluidics*. 2008, p. 793.
- [23] G. Bradski, "The openCV library," *Dr. Dobb's J., Softw. Tools Prof. Programmer*, vol. 25, no. 11, pp. 120–123, 2000.
- [24] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," 2016, *arXiv:1611.05431*.
- [25] A. Paszke, S. Gross, F. Massa, and A. Lerer, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2019, pp. 8026–8037.
- [26] L. Bottou and O. Bousquet, "The tradeoffs of large scale learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 20, 2007, p. 1–8.
- [27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [28] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and efficient object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10781–10790.
- [29] X. Dai, Y. Chen, B. Xiao, D. Chen, M. Liu, L. Yuan, and L. Zhang, "Dynamic head: Unifying object detection heads with attentions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 7373–7382.



XIANG WANG received the B.S. degree in automation from Xidian University, Xi'an, China, in 2009, and the Ph.D. degree from BUAA, Beijing, China, in 2016.

Since 2019, she has been a Researcher with the AI Innovation and Application Center, China Unicom. Before that, she worked at CloudMinds Technologies Inc., and the Information Science Research Institute of China Electronics Technology Group Corporation. Her research interests include computer vision, augmented reality, and human–computer interaction.



KAI WANG received the Ph.D. degree from Nanyang Technological University, Singapore, in 2013. He has been the AI Director with the AI Innovation and Application Center, China Unicom, since 2019. Before that, he worked at CloudMinds Technologies Inc., and the Central Research Institute of Huawei Technologies. He has published over 20 refereed papers on international journals and conferences and granted more than 40 patents. His research interests include computer vision, computer graphics, and human–computer interaction.



XIAOHONG LI received the M.S. degree in biomedical engineering from Northeastern University, Shenyang, China, in 2020. From 2018 to 2019, she was a Guest Intern at the Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences. Since 2020, she has been employed with the AI Innovation and Application Center, China Unicom. Her research interests include 3D reconstruction, depth estimation, digital human, vision-based detection and classification, and other related directions.



SHIGUO LIAN (Member, IEEE) received the Ph.D. degree in control science and engineering from the Nanjing University of Science and Technology, China, in 2005.

During the Ph.D. study, he has been a Research Assistant at the City University of Hong Kong, in 2004. From 2005 to 2019, he has been at the France Telecom Research and Development Center, Huawei Central Research Institute, and the CloudMinds Technologies for research work. He is currently the Chief AI Scientist and the Head of the AI Innovation and Application Center, China Unicom. He is the author of more than 100 refereed international articles, co-edited more than ten books, and held more than 100 patents. His research interests include artificial intelligence, human–machine interaction, multimedia communication, and their applications in industry, city, and so on. He is a member of several IEEE Technical Committees and on the editor board of several refereed international journals.