Original Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.28922

Identification of milled rice varieties using machine vision

Wenmin FAN[1], Sen YANG[1*]

Abstract

Machine vision detection has the advantages of non-contact, non-destructive, and is suitable for large-scale and high-speed detection. However, existing studies only adopted one or one type image feature in the varieties identification of milled rice, and contained limited explanation of the influence of the image features on the identification accuracy. Thus, this paper developed an identification model for milled rice varieties based on multiple image features and explored the contributions of each image feature on the identification accuracy. The extracted image features were thirteen morphology features, eighteen color features, and four texture features. The training and testing data sets were one hundred and forty milled rice samples. The partial least squares algorithm was used to identify milled rice varieties. The multivariate data analysis was used to get the influence of the image features on the identification accuracy. Experiment results showed that, by selecting twenty features from the original thirty-five features, the identification accuracies were 100%, 70%, 100%, and 100% for milled rice varieties YG, WC, XS, and JN.

Keywords: milled rice; variety identification; machine vision; image feature.

Practical Application: Machine vision is used for rapid detection of milled rice varieties.

1 Introduction

Rice is the most important cereal crop worldwide in Asian countries because of its good taste and nutritional value (Murtaza et al., 2022; Munarko et al., 2022). Consumers pay more for premium products, so falsely marking geographical origins or adulteration can bring unfair economic advantage to fraudsters, which, on the other hand, undermines the credibility of the producers to consumers (Mongkontanawat et al., 2022; Lee et al., 2019). Thus, accurately deciding the origin of milled rice is desirable for consumers, retailers, and governmental authorities.

Several measurement methods had been developed to identify rice varieties, such as Raman spectroscopy (Hwang et al., 2012; Zhu et al., 2018; Sha et al., 2019), Multi-element fingerprinting (Gonzálvez et al., 2011; Cheajesadagul et al., 2013), and Near-infrared spectroscopy (Lin et al., 2012; Liu et al., 2020; Peijin et al., 2021). The methods can give exact results, but existing application limit because of the complex sample preparation, long testing process, and high instrument cost. The non-contact and non-destructive identification of mill rice varieties for large-scale and high-speed is still an urgent aim (Bagchi et al., 2016).

Machine vision technique does not have the above problems in the application. This technique is based on the image acquisition and image processing, which has the advantages of non-contact, non-destructive, and is suitable for large-scale and high-speed detection. Research efforts had focused on the potential of machine vision techniques for deciding the source or geographical origin of rice seed. Fayyazi evaluated the efficiency of morphological and textural features of rice seeds' images in varieties identification (Saeideh et al., 2017). Kantip et al. (2020) extracted image features of shape, color, and texture to classify rice seed varieties. Singh classified the bulk rice grain varieties using images features of color, texture, and wavelet (Singh & Chaudhury, 2016).

Besides, many researchers used the machine vision method to detect the rice quality features. Chen et al. (2019) developed a machine vision to inspect flawed rice kernels. Lin and Sun used image processing technique to detect milled rice chalkiness (Cheng et al., 2014; Lin et al., 2020). Singh et al. (2020) proposed a method that combines image processing and machine learning to measure the size and mass of rice kernels. Zareiforoush et al. (2015) qualitatively graded the milled rice for the degree of milling and percentage of broken using the image processing.

For the varieties classification of milled rice using machine vision, existing studies only adopted one image feature or one type image feature in the varieties classification of milled rice. Golpour identified white rice cultivars through developing an image processing algorithm based on the color features (Ranum et al., 2014). Yang (2021) combined the length-width ratio and the spectroscopy data to identify milled rice varieties. However, to our knowledge, no studies had previously published the influence of image features on the identification accuracy of milled rice varieties.

Thus, this paper developed an identification model for milled rice varieties based on multiple image features and explored the contributions of each image feature on the identification accuracy. Section 2 described the machine vision the milled rice samples, the multiple-image features, and the identification procedure. Section 3 presented the identification results based on different image features types, followed by discussing the influence of each image feature on the identifying milled rice varieties.

Received 12 Mar, 2022
Accepted 02 May, 2022
[1] *College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin, China*
*Corresponding author: yangsen@nefu.edu.cn

2 Materials and methods

2.1 Machine vision and milled rice samples

The machine vision used a 1/2.5" CMOS (Complementary Metal Oxide Semiconductor) color digital camera (USB-500, TRWY) to record images. The Pixel and pixel size were 5 million and 2.2 μm × 2.2 μm. A zoom lens and a ring LED lamp were used to ensure the image quality. The rice sample was on a sample table 370 mm away from the zoom lens. The surrounding light was constant when capturing the image. Figure 1 shows the machine vision system used to record milled rice images.

This paper collected one hundred and forty milled rice samples from four rice-producing regions. Each variety has thirty-five samples, twenty-five for training and ten for testing. Table 1 shows the detailed information of milled rice samples.

Here, YG represents Yueguang rice (Liaoning, China), WC represents Wuchang rice (Heilongjiang, China), XS represents Xiangshui rice (Heilongjiang, China), and JN represents Jiangnan (Jiangxi, China). The milled rice samples were balanced by temperature and humidity before the image gaining. Figure 2 shows the original images of milled rice samples.

2.2 Image features

The image features used in this paper were three classes: morphology, texture, and color. The morphology features included thirteen subclasses, the texture features included four subclasses, and the color features included eighteen subclasses. Algorithms were developed using MATLAB to extract different image features. Table 2 shows the multiple image parameters used in this paper.

(1) The morphology features used in this paper were:

Region pixel number (S1): the number of pixels within its boundary.

Perimeter (C1): the length surrounds the image edge.

Length (L1): the smallest rectangle that can enclose the image.

Width (D1): the width of the smallest rectangle that can enclose the image.

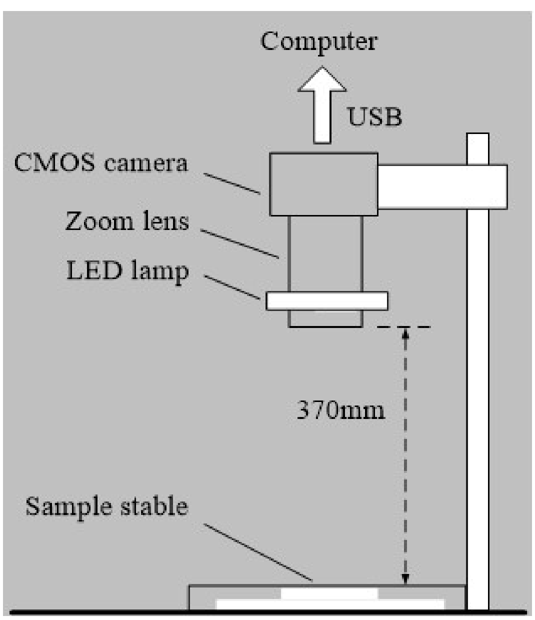

Figure 1. The machine vision system used to record milled rice images.

Table 1. The detailed information of milled rice samples.

Variety	Code Name	Serial Number	
		Training Set	Testing Set
Yueguang	YG	1-25	26-35
Wuchang	WC	36-60	61-70
Xiangshui	XS	71-95	96-105
Jiangnan	JN	106-130	131-140

Table 2. The multiple image parameters used in this paper.

Morphology	Texture	Color
S1; C1; L1; D1	CS	R-mean; G-mean; B-mean
K; R; CO	HT	R-std; G-std; B-std
SF1; SF2; SF3	CL	Y-mean; C_b-mean; C_r-mean
L; D; TA	EG	Y-std; C_b-std; C_r-std
		H-mean; S-mean; V-mean
		H-std; S-std; V-std

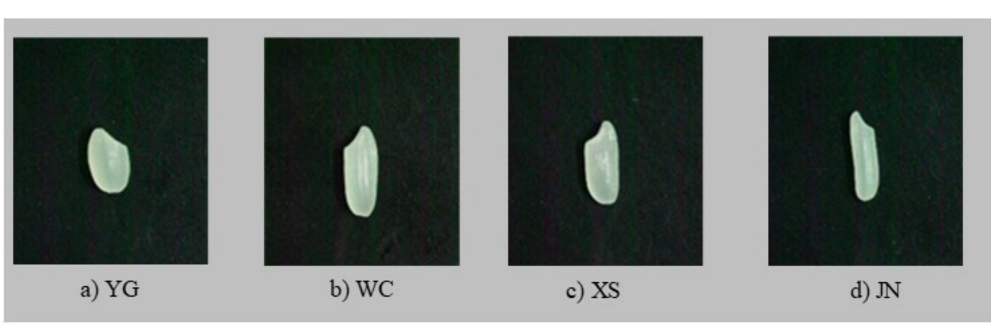

Figure 2. The original images of milled rice samples.

Length-width ratio (K): length/width.

Roundness (R): $R= (4\cdot\pi\cdot S1)/(C1)^2$.

Compactness (CO): $CO= [(4\cdot\pi\cdot S1)^{1/2}]/L1$.

Morphology feature 1 (SF1): $SF1=L1/S1$.

Morphology feature 2 (SF2): $SF2=S1/(L1^2)$.

Morphology feature 3 (SF3): $SF3=S1/[(0.5\cdot L1)\cdot(0.5\cdot D1)\cdot\pi]$.

Long axis length (L): The longest line that through the image.

Short axis length (D): The longest line that through the image and vertical to the long axis.

Convex area (TA): The number of pixels in the smallest convex polygon that can contain the area.

(2) The texture features used in this paper were:

Contrast (CS): Reflects the sharpness of the image and the degree of furrow depth of the texture.

Homogeneity (HT): Reflects the homogeneity of the texture by measuring the local variation.

Correlation (CL): Reflects the local gray correlation of the image.

Energy (EG): Reflects the uniformity of gray distribution and texture thickness.

(3) The color features used in this paper were:

This study used three color spaces: RGB, HSV, and YC_bC_r. The RGB Space contains color components Red (R), Green (G), and Blue (B). The HSV Space contains color components Hue (H), Saturation (S), and Value (V). The YC_bC_r Space contains color components Luminance (Y), Blue Chrominance (C_b), and Red Chrominance (C_r).

Here, we calculate two features for each color component: first moment and second moment. The first moment (mean) reflects the average intensity of each color component. The second moment (std) reflects the color variance of the measured region.

Thus, the color features used in this paper were the R-mean, G-mean, B-mean, H-mean, S-mean, V-mean, Y-mean, C_b-mean, C_r-mean, R-std, G-std, B-std, H-std, S-std, V-std, Y-std, C_b-std, C_r-std.

2.3 Identification procedure

This research used the Partial Least Squares (PLS) algorithm to find relations between image features and varieties. The inputs were different image features, the output was varieties number. The identification procedure of milled rice variety using the machine vision is shown in Figure 3:

(1) Gain milled rice image and assign variety number;

(2) Preprocess the milled rice image and calculate multiple image features using the MATLAB program;

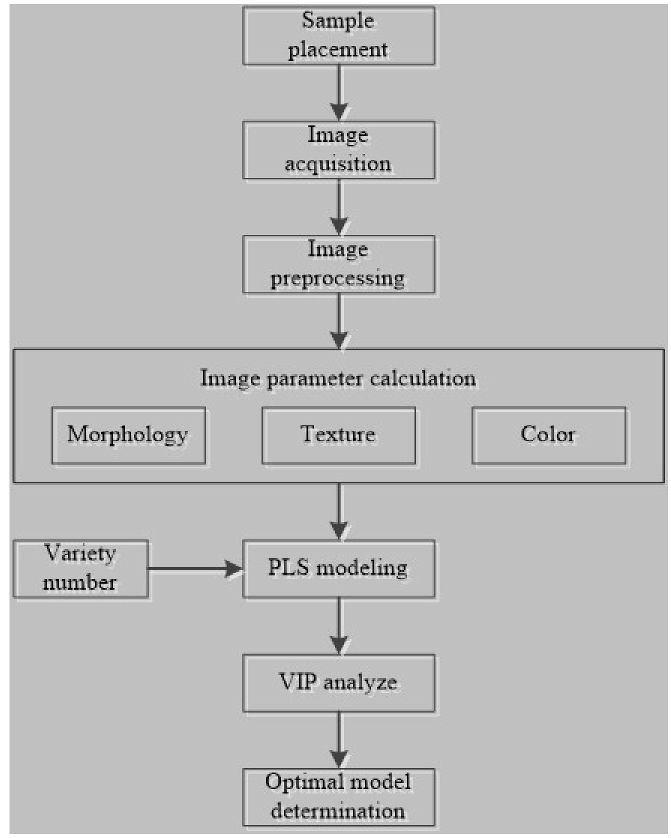

Figure 3. The identification procedure of milled rice variety using the machine vision.

(3) Find out the relation between variety number and multiple image features using the PLS algorithm;

(4) Perform VIP analysis to decide the best image features and best model.

In this paper, the chosen image preprocessing methods contained gray-level transformation, median filter, threshold segmentation, opening operation, and image fusion. Figure 4 shows the relations of images, image preprocessing methods, and image features.

The preprocessed images of WC milled rice are shown in Figure 5.

3 Results and discussion

3.1 Varieties identification results based on different image feature types

Table 3 shows the identification accuracy of milled rice varieties using different image feature types. As shown in Table 3, the identification accuracy was the highest when using the morphology and texture features for modeling. The identification accuracy is the lowest when not using the morphological feature for modeling, proved the morphology features contributed the highest to the identification. Decrease of the identification accuracy happened when changing the texture features to

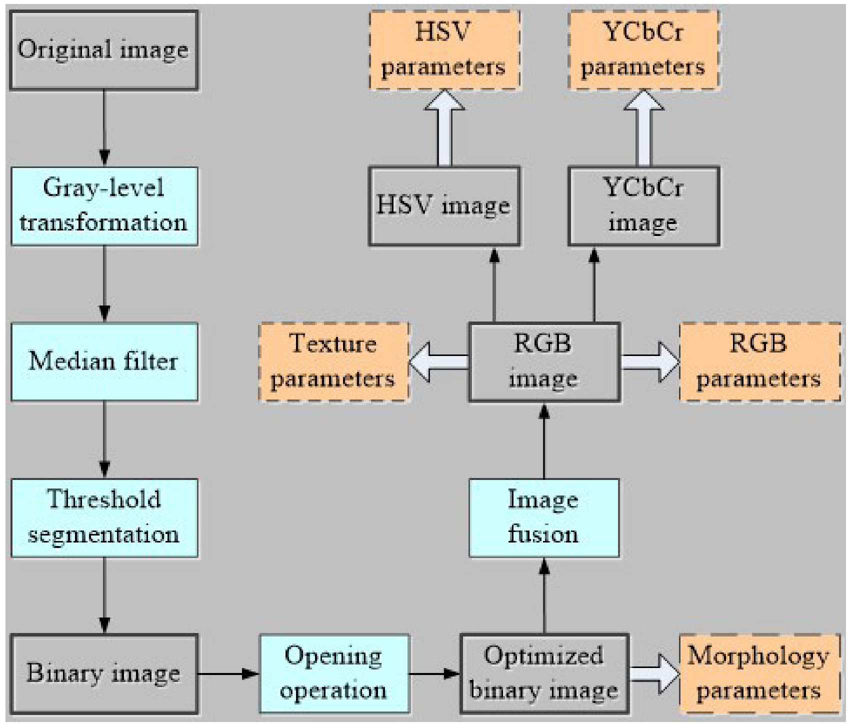

Figure 4. The relations of images, image preprocessing methods, and image features.

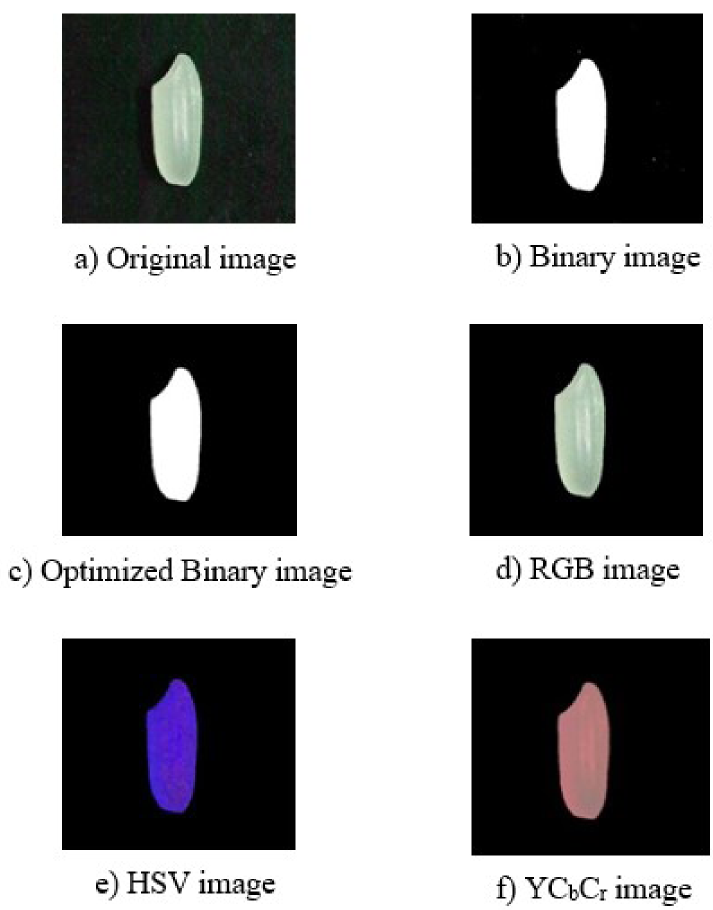

Figure 5. The preprocessed images of WC milled rice.

Table 3. The identification accuracy of milled rice varieties using different image feature types.

Image feature types	Identification accuracy			
	YG	WC	XS	JN
Morphology, texture, and color	100%	70%	90%	100%
Morphology and texture	100%	70%	100%	100%
Morphology and color	100%	60%	90%	100%
Texture and color	90%	50%	100%	80%

Table 4. The identification accuracies when selecting different VIP thresholds for the variable selection.

VIP threshold	Identification accuracy			
	YG	WC	XS	JN
0.90	100%	70%	100%	100%
0.93	100%	70%	100%	100%
0.94	100%	60%	100%	100%
0.95	100%	60%	90%	100%
1.00	100%	50%	80%	100%

color features, which proved the texture features owed a higher contribution than color features. The identification accuracy of YG and JN was 100% since they were externally different from the others, as they had stretched or plump shapes. We supposed the likeness of XS and WC resulted in the 70% identification accuracy for WC. Besides, the identification accuracy was not the highest when using all image features for modeling, which proved that selecting too many image features will lessen the identification accuracy.

3.2 Analysis results of variable importance in projection

Some image features are less important and the identification performance will decline when using too many redundant features. A good insight in the variable importance and model accuracy can be achieved by using multivariate data analysis. A measure that summarizes the importance is the Variable Importance for the Projection (VIP), which can give the influence of the input variables on the model. This paper performed VIP analysis in a Windows environment using SIMCA software. VIP values of 1.0 and above are most relevant for explaining Y, while VIP values below 0.5 shows unimportant (Jonsson, 2011). Figure 6 shows the calculated VIP values for each image feature.

Table 4 shows the identification accuracies when selecting different VIP thresholds for the variable selection. As shown in Table 4, the identification accuracies were different with various VIP thresholds. When the VIP threshold was bigger than 0.93, the identification accuracy decreased, which proved the loss of valid features. When the VIP threshold was smaller than 0.93, the identification accuracy remained unchanged. Thus, the selected VIP threshold was 0.93 in this paper.

Table 5 shows the twenty image features used for the best model. For these twenty features, the morphology feature K was the most important in given model (VIP=1.302), while the color feature Y-std was the least important feature (VIP=0.938). The developed best model showed an accuracy of 100%, 70%, 100%, and 100% for the milled rice varieties YG, WC, XS, and JN.

Table 5. The twenty image parameters used for the optimal model.

Class	Subclass
Morphology	S1, C1, L1, D1, K, R, CO, SF1, SF2, SF3, L, D, TA
Color	R-std, G-std, H-mean, H-std, V-std, Y-std
Texture	CL

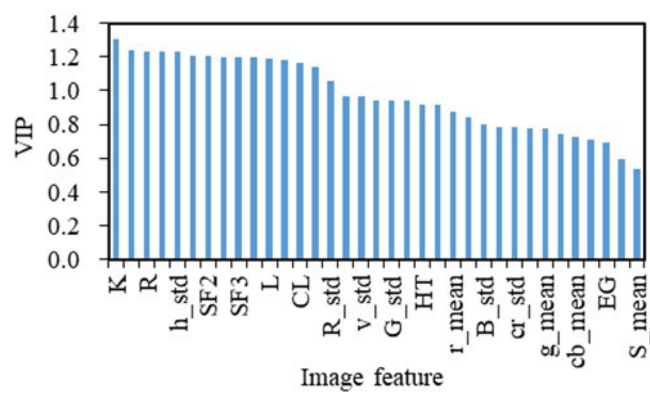

Figure 6. The calculated VIP values for each image feature.

4 Conclusions

This work presented a new machine vision-based measurement method for varieties identification of milled rice. A model was developed based on the morphology features, texture features, and color features. The identification accuracy will decrease when using all image features for modeling. The morphological features have the highest contribution to identifying milled rice variety. The model is best when selecting 0.93 as the VIP threshold for varieties selection. At this point, the model inputs were twenty image features from the original thirty-five image features. At last, the identification accuracies were 100%, 70%, 100%, and 100% for milled rice varieties YG, WC, XS, and JN.

However, this paper got the experiment results under a laboratory setting, so it had some limits. A long-term project contains: (1) Add more milled rice varieties and use more image features; (2) Establish a variety identification model based on combining near-infrared spectroscopy and machine vision.

Conflict of interest

The author declare that they have no conflict of interest.

References

Bagchi, T. B., Sharma, S., & Chattopadhyay, K. (2016). Development of NIRS models to predict protein and amylose content of brown rice and proximate compositions of rice bran. *Food Chemistry*, 191, 21-27. http://dx.doi.org/10.1016/j.foodchem.2015.05.038. PMid:26258697.

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chemistry*, 141(4), 3504-3509. http://dx.doi.org/10.1016/j.foodchem.2013.06.060. PMid:23993513.

Chen, S., Xiong, J., Guo, W., Bu, R., Zheng, Z., Chen, Y., Yang, Z., & Lin, R. (2019). Colored rice quality inspection system using machine vision. *Journal of Cereal Science*, 88, 87-95. http://dx.doi.org/10.1016/j.jcs.2019.05.010.

Cheng, M. S., Tao, L., Cheng, X. J., Ting, T., Dou, D. G., Li, J. W., Ying, Y. C., & Xiu, M. L. (2014). Evaluation and analysis the chalkiness of connected rice kernels based on image processing technology and support vector machine. *Journal of Cereal Science*, 2(60), 426-432. http://dx.doi.org/10.1016/j.jcs.2014.04.009.

Gonzálvez, A., Armenta, S., & de la Guardia, M. (2011). Geographical traceability of "Arròs de Valencia" rice grain based on mineral element composition. *Food Chemistry*, 126(3), 1254-1260. http://dx.doi.org/10.1016/j.foodchem.2010.11.032.

Hwang, J., Kang, S., Lee, K., & Chung, H. (2012). Enhanced Raman spectroscopic discrimination of the geographical origins of rice samples via transmission spectral collection through packed grains. *Talanta*, 101, 488-494. http://dx.doi.org/10.1016/j.talanta.2012.10.001. PMid:23158353.

Jonsson, P. (2011). Classification of road conditions: from camera images and weather data. In *2011 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA) Proceedings* (pp. 1-6). New York: IEEE. http://dx.doi.org/10.1109/CIMSA.2011.6059917.

Kantip, K., Pitchayagan, T., Wasin, S., Panintorn, P., Kosom, C., Supanit, P., & Anchalee, P. (2020). Development of paddy rice seed classification process using machine learning techniques for automatic grading machine. *Journal of Sensors*, 2020, 1-14. http://dx.doi.org/10.1155/2020/7041310.

Lee, Y., Dias-Morse, P. N., & Meullenet, J.-F. (2019). Effect of rice variety and milling fraction on the starch gelatinization and rheological properties of rice milk. *Food Science and Technology*, 39(4), 1047-1051. http://dx.doi.org/10.1590/fst.17118.

Lin, P., Zhang, H., He, J. Q., Zou, Z., & Chen, Y. (2020). Research on chalky rice detection based on visible spectrogram and deep neural network technology. *Spectroscopy and Spectral Analysis*, 40, 233-238. http://dx.doi.org/10.3964/j.issn.1000-0593(2020)01-0233-06.

Lin, W.-S., Yang, C.-M., & Kuo, B.-J. (2012). Classifying cultivars of rice (*Oryza sativa* L.) based on corrected canopy reflectance spectra data using the orthogonal projections to latent structures (O-PLS) method. *Chemometrics and Intelligent Laboratory Systems*, 115, 25-36. http://dx.doi.org/10.1016/j.chemolab.2012.04.005.

Liu, Y. C., Li, Y. Y., Peng, Y. K., Yan, S., & Han, D. (2020). Application of two-dimensional correlation spectra in the identification of adulterated rice. *Spectroscopy and Spectral Analysis*, 40, 1559-1564. http://dx.doi.org/10.3964/j.issn.1000-0593(2020)05-1559-06.

Mongkontanawat, N., Ueda, Y., & Yasuda, S. (2022). Increased total polyphenol content, antioxidant capacity and γ-aminobutyric acid content of roasted germinated native Thai black rice and its microstructure. *Food Science and Technology*, 42, e34521. http://dx.doi.org/10.1590/fst.34521.

Munarko, H., Sitanggang, A. B., Kusnandar, F., & Budijanto, S. (2022). Germination of five Indonesian brown rice: evaluation of antioxidant, bioactive compounds, fatty acids and pasting properties. *Food Science and Technology*, 42, e19721. http://dx.doi.org/10.1590/fst.19721.

Murtaza, G., Huma, N., Sharif, M. K., & Zia, M. A. (2022). Probing a best suited brown rice cultivar for the development of extrudates with special reference to Physico-chemical, microstructure and sensory evaluation. *Food Science and Technology*, 42, e103521. http://dx.doi.org/10.1590/fst.103521.

Peijin, T., Kevin, L. J., Tingting, W., Elejalde, U., Hongchao, Z., Yuanrong, J., & Wenming, C. (2021). Rapid identification of the variety and geographical origin of Wuyou No.4 rice by fourier transform near-infrared spectroscopy coupled with chemometrics. *Journal of Cereal Science*, 102, 103322. http://dx.doi.org/10.1016/j.jcs.2021.103322.

Ranum, P., Peña-Rosas, J. P., & Garcia-Casal, M. N. (2014). Global maize production, utilization, and consumption. *Czech Journal of Food Sciences*, 32, 280-287. http://dx.doi.org/10.1111/nyas.12396.

Saeideh, F., Mohammad, H. A. F., Abbas, R., Amirhassan, M. S., & Hassan, S. (2017). Identification and classification of three iranian rice varieties in mixed bulks using image processing and MLP neural network. *Journal of Food Engineering*, 13, 20160121. http://dx.doi.org/10.1515/ijfe-2016-0121.

Sha, M., Gui, D., Zhang, Z., Ji, X., Shi, X., Liu, J., & Zhang, D. (2019). Evaluation of sample pretreatment method for geographic authentication of rice using Raman spectroscopy. *Journal of Food Measurement and Characterization*, 13(3), 1705-1712. http://dx.doi.org/10.1007/s11694-019-00087-7.

Singh, K. S., & Chaudhury, S. (2016). Efficient technique for rice grain classification using back-propagation neural network and wavelet decomposition. *IET Computer Vision*, 10(8), 780-787. http://dx.doi.org/10.1049/iet-cvi.2015.0486.

Singh, S. K., Vidyarthi, S. K., & Tiwari, R. (2020). Machine learnt image processing to predict weight and size of rice kernels. *Journal of Food Engineering*, 274, 109828. http://dx.doi.org/10.1016/j.jfoodeng.2019.109828.

Yang, S. (2021). Development of an integrated variety and appearance quality measurement system for milled rice. *Journal of Food Measurement and Characterization*, 15(5), 4679-4685. http://dx.doi.org/10.1007/s11694-021-01041-2.

Zareiforoush, H., Minaei, S., Alizadeh, M. R., & Banakar, A. (2015). A hybrid intelligent approach based on computer vision and fuzzy logic for quality measurement of milled rice. *Measurement.*, 66, 26-34. http://dx.doi.org/10.1016/j.measurement.2015.01.022.

Zhu, L., Sun, J., Wu, G., Wang, Y., Zhang, H., Wang, L., Qian, H., & Qi, X. G. (2018). Identification of rice varieties and determination of their geographical origin in China using Raman spectroscopy. *Journal of Cereal Science*, 82, 175-182. http://dx.doi.org/10.1016/j.jcs.2018.06.010.