Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 23, No. 3, 2026

Simultaneous Determination of Protein, Fat and Amylopectin in Rice Used for Baijiu Brewing Using Spectral Reconstruction

Haonan Yi[1], Liangliang Xie[1, 2, 3, *], Kaiyang Yuan[1], Anying Cai[1]

[1]School of Mechanical Engineering, Sichuan University of Science and Engineering, Yibin 644000, China
[2]Solid-state Brewing Technology Innovation Center of Sichuan, Luzhou 646000, China
[3]Brewing Science and Technology Key Laboratory of Sichuan Province, Yibin 644000, China
[*]Corresponding authors: xiehnpyll@163.com

Abstract: Rapid measurement of protein, fat and amylopectin is needed for quality control of rice used in Baijiu brewing. Hyperspectral imaging provides information for this task, but its equipment cost and scanning time constrain deployment. We combined RGB-based spectral reconstruction with multitask regression to determine the three components simultaneously. A multistage spectral-wise transformer (MST++) reconstructed 31 bands from 400 to 700 nm at 10 nm intervals. A Kolmogorov–Arnold network with customized gate control (KAN-CGC) then combined shared and task-specific experts, KAN prediction towers and uncertainty-weighted losses. Reconstruction models were developed using 400 samples; component models used 80 separate physical samples from eight rice varieties, with 12 region-of-interest spectra averaged per sample. MST++ achieved a test mean relative absolute error of 0.0240 and a peak signal-to-noise ratio of 39.2855 dB. After multiplicative scatter correction, KAN-CGC gave prediction coefficients of determination of 0.9824, 0.9612 and 0.9803 for protein, fat and amylopectin, respectively, with ratios of performance to deviation of 7.78, 5.24 and 7.36. The corresponding coefficients for directly acquired full spectra were 0.9887, 0.9701 and 0.9869. The processing workflow required 0.909 s per sample, compared with 18.728 s for hyperspectral acquisition and preprocessing, a 95.15% reduction. These results support RGB imaging and spectral reconstruction as a rapid approach to simultaneous component determination under the tested laboratory conditions.

Keywords: Brewing rice; spectral reconstruction; MST++; customized gate control; multitask learning; nondestructive analysis.

1. Introduction

Rice is a major substrate for the solid-state fermentation of Chinese Baijiu. Its protein, fat and amylopectin contents influence fermentation performance, liquor yield and flavor development [1, 2]. Starch supplies fermentable carbon, protein degradation provides amino acid precursors, and lipids contribute to organic acid and ester formation [3, 4]. Measuring these components together therefore offers a more complete assessment of raw materials than a single indicator. Conventional Kjeldahl, Soxhlet extraction and dual-wavelength spectrophotometric assays are destructive and time-consuming, limiting their use for high-throughput quality control [5].

Hyperspectral imaging (HSI) combines spatial and spectral information for nondestructive grain analysis. It has been used for multitask prediction of rice components [6], protein and fat determination in sorghum [7], and amylose and amylopectin measurement in sorghum [8]. Its practical deployment nevertheless requires costly instrumentation, relatively slow push-broom scanning and controlled illumination [9]. Spectral reconstruction addresses these constraints by estimating hyperspectral data from inexpensive RGB images [10]. Applications to agricultural products include tomato and other crop quality assessments [11, 12], and our previous work applied this strategy to online moisture determination in Baijiu Daqu [9].

Moving from one component to simultaneous determination introduces a further challenge: the three targets differ greatly in concentration and share overlapping spectral information. Shared representations can be useful, but unequal task contributions can impair multitask learning [13]. We therefore paired MST++, which models interband dependencies through spectral-wise attention [14], with a KAN-CGC regressor that combines shared and task-specific pathways. We first compared MST++, HSCNN-D and HRNet for reconstruction, then assessed spectral preprocessing and four regression models using reconstructed and directly acquired spectra. Accuracy, internal model responses and processing time were examined across eight rice varieties.

2. Materials and Methods

2.1. Sample preparation

Eight rice varieties were studied: Longyang 16, Nanjing 46, Nanjing 5055, Inner Mongolia fragrant rice, Wuchang fragrant rice, Xinglong 1, Xiaozhan rice and Yinzhen. Rice was spread evenly in 8 cm Petri dishes, with 60 independently prepared samples per variety (480 in total). Of these, 400 were assigned to spectral reconstruction, comprising 320 training and 80 test samples. The remaining 80 physical samples were reserved for component modeling and were excluded from reconstruction training. Twelve region-of-interest (ROI) spectra were averaged within each sample to obtain one modeling spectrum. Component models therefore used 80 sample-level spectra, divided into 64 calibration and 16 prediction samples. The reference protein, fat and amylopectin contents were 6.72 ± 0.83, 0.16 ± 0.05 and 81.00 ± 5.80 g/100 g, respectively. Fat was the lowest-concentration target.

2.2. Reference component measurements

Protein was measured by the Kjeldahl method according to GB 5009.5-2025. Ground, sieved rice was digested with a catalyst and concentrated sulfuric acid; distilled ammonia was absorbed in boric acid and titrated with standard acid, and nitrogen was converted to protein content. Fat was measured by Soxhlet extraction according to GB 5009.6-2025. Samples in preweighed filter-paper thimbles were extracted with anhydrous diethyl ether at 60–70 °C for 6–8 h. After solvent recovery, the residue was dried at 101–105 °C to constant mass, and fat content was calculated gravimetrically. Amylopectin was measured by dual-wavelength spectrophotometry: defatted, gelatinized samples were reacted with iodine, absorbance was measured at 620 and 510 nm, and concentration was obtained from calibration curves. Each assay was performed in triplicate and averaged for each sample.

2.3. Image acquisition and dataset construction

The HSI system comprised a Specim FX10E push-broom camera, two 150 W halogen lamps, a motorized translation stage and an acquisition computer (Fig. 1a). The image size was 630 × 1024 pixels, with a spectral resolution of 5 nm and 448 channels spanning 400–1000 nm. After a 15 min warm-up, images were acquired at 50 Hz with an 8.8 ms exposure and a stage speed of 10.84 mm/s. White and dark references were used to correct illumination and sensor effects:

$$R_c = \frac{R_0 - DN_d}{DN_w - DN_d} \quad (1)$$

Here, R_c is corrected reflectance, R_0 is the raw signal, and DN_w and DN_d are the white and dark references. Corrected data were restricted to 400–700 nm and resampled at 10 nm intervals to produce 31-band reference cubes. Pseudo-RGB images synthesized from the HSI data provided the spatial reference for registration and the paired RGB input for reconstruction training.

Industrial RGB images were acquired with a Daheng MER2-503-36U3C camera and an LED light source (Fig. 1b). The illumination setting was 1.1, exposure was 50 000 μs, gain was 1.00 dB, and continuous automatic white balance was enabled. Images were saved as PNG files and initially center-cropped to 1680 × 1680 pixels. Each cropped image was resized proportionally to 512 × 512 pixels without a second crop. The reconstructed 512 × 512 × 31 cube was then bilinearly interpolated to 1680 × 1680 pixels to align with the ROI mask on the original coordinate grid.

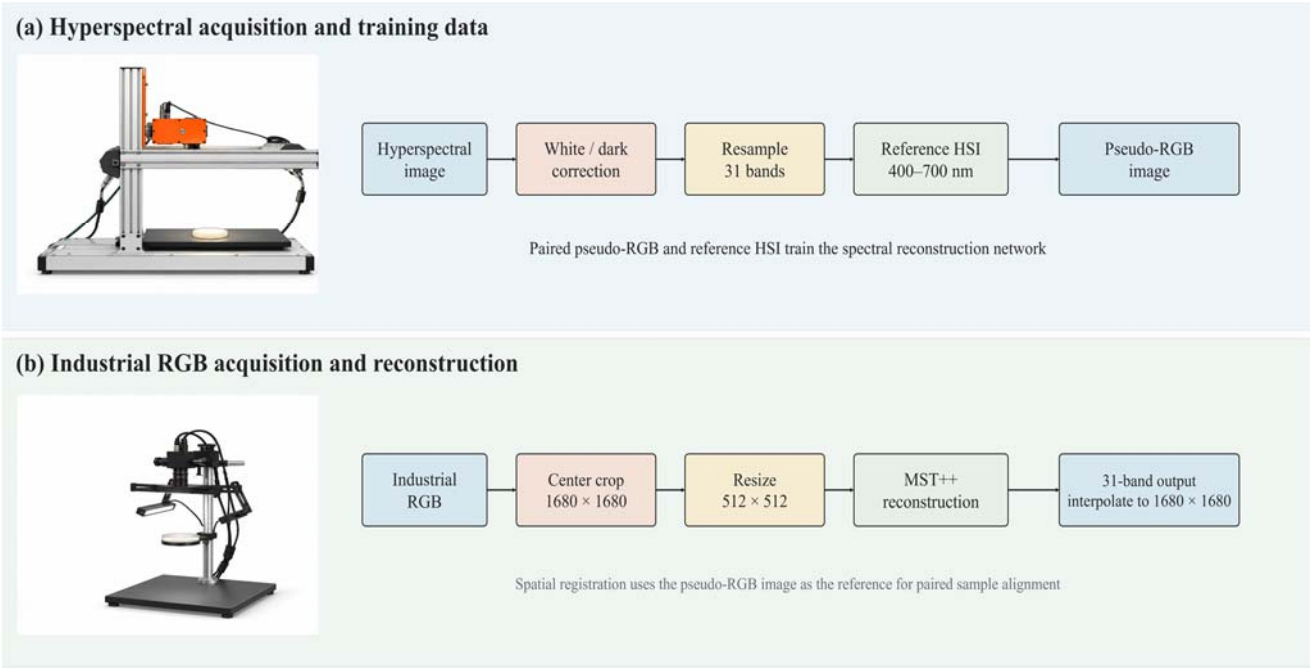


Figure 1. Image acquisition and spectral reconstruction workflow. (a) Hyperspectral acquisition, radiometric correction and construction of paired training data. (b) Industrial RGB acquisition, resizing, reconstruction and interpolation to the ROI coordinate grid.

2.4. ROI selection and spectral extraction

Grayscale conversion, structural feature extraction, thresholding and morphological filtering were used to obtain a binary mask excluding the Petri-dish background. Twelve ROIs were positioned within the mask over a circular sampling area centered on the sample with a radius of 500 pixels (Fig. 2). The spectrum of each ROI was calculated by averaging its pixels at each band:

$$R_i = \frac{1}{m}\sum_{j=1}^{m} R_{i,j} \quad (2)$$

R_i is mean reflectance at band i, m is the number of ROI pixels, and $R_{i,j}$ is the reflectance of pixel j at that band. The 80 physical samples yielded 960 ROI spectra. Averaging the 12 ROI spectra within each sample produced 80 modeling spectra, each linked to that sample's measured reference composition.

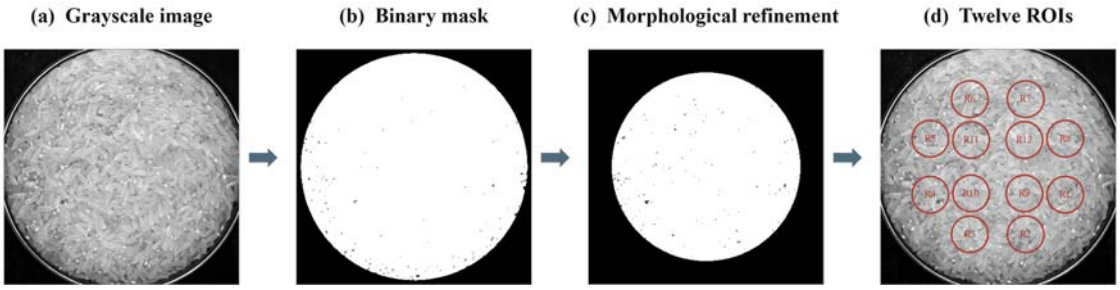


Figure 2. ROI extraction. (a) Grayscale image. (b) Binary mask. (c) Morphological refinement. (d) Twelve ROIs.

2.5. Spectral preprocessing

Four preprocessing options were compared: no preprocessing, multiplicative scatter correction (MSC), standard normal variate (SNV) transformation and Savitzky–Golay (SG) smoothing. MSC uses regression against a reference spectrum to reduce multiplicative and additive scatter effects [15]; SNV standardizes each spectrum [16]; and SG smoothing fits local polynomials to suppress noise while retaining spectral shape [17]. Partial least squares regression (PLSR) was used consistently to compare these options for all three components.

2.6. Spectral reconstruction models and evaluation

MST++ was compared with HSCNN-D and HRNet, representing dense convolutional reconstruction and multiresolution feature fusion, respectively [18, 19]. MST++ maps RGB images to spectral cubes through convolutional projections, cascaded single-stage spectral-wise transformers (SSTs) and a global residual connection [14] (Fig. 3a). Each SST uses a U-shaped encoder–decoder with 4 × 4 convolutional downsampling, 2 × 2 transposed-convolution upsampling and skip connections. Spectral-wise attention blocks (SABs) operate at successive scales and at the bottleneck (Fig. 3b). An SAB combines layer normalization, spectral-wise multihead self-attention (S-MSA) and a feed-forward network. S-MSA models dependencies along the spectral dimension, while skip connections preserve spatial information (Fig. 3c–e). HSCNN-D reuses convolutional features through dense connections; HRNet exchanges features across parallel resolution branches.

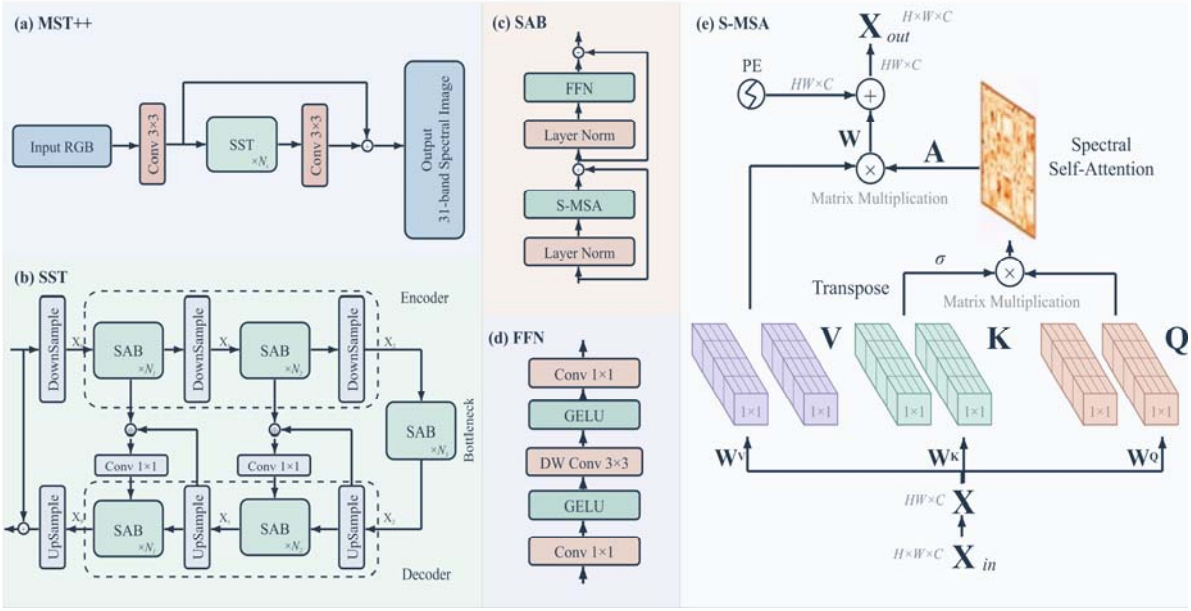


Figure 3. MST++ architecture. (a) Overall reconstruction network. (b) Single-stage spectral-wise transformer (SST). (c) Spectral-wise attention block (SAB). (d) Feed-forward network (FFN). (e) Spectral-wise multihead self-attention (S-MSA).

Reconstruction was evaluated using root mean square error (RMSE), mean relative absolute error (MRAE) and peak signal-to-noise ratio (PSNR):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(I_{G,j} - I_{R,j})^2} \quad (3)$$

$$\mathrm{MRAE} = \frac{1}{N}\sum_{j=1}^{N}\frac{|I_{G,j} - I_{R,j}|}{I_{G,j}} \quad (4)$$

$$\mathrm{PSNR} = 10\log_{10}\frac{\mathrm{MAX}^2}{\mathrm{MSE}} \quad (5)$$

H and W are image height and width, Λ is the number of bands (31), and I_G and I_R denote measured and reconstructed data. N = HWΛ is the total number of pixel–band values; the sums cover all of them. MAX was 1 for normalized reflectance. Lower RMSE and MRAE and higher PSNR indicate closer agreement with measured data.

2.7. Multicomponent regression models

KAN-CGC was designed to combine common spectral

information with task-specific features for protein, fat and amylopectin prediction (Fig. 4). A 31-dimensional spectrum first passed through two one-dimensional convolutional stages with kernel sizes of 7 and 3, each followed by batch normalization and GELU activation. Adaptive average pooling produced 64-dimensional base features. Two CGC layers were used, each containing two shared experts and two experts per task, giving eight experts per layer. Each task gate considered only the two shared experts and its own two experts; weights for the other experts were structural zeros. For an input feature vector e, the gate and mixture were g_t = softmax(W_t e) and $h_t = \sum_i g_{t,i}\ E_i(e)$, with normalization restricted to the candidate experts [20].

Each task mixture was passed to a KAN tower with dimensions 64 → 32 → 1. Learnable univariate B-spline functions model nonlinear relationships between learned features and component content [21]. The spline grid size was G = 5 and the order was k = 3. Training used homoscedastic uncertainty weighting, $L = \sum_t [L_t/(2\sigma_t^2) + \log \sigma_t]$, where L_t is the task-specific Huber loss and σ_t is a learned uncertainty parameter [22]. For model inspection, second-layer CGC gate weights were visualized as an 8 × 3 matrix. The first KAN layer of each tower was evaluated along the diagonal $z(x) = x \cdot 1_{64}$: all 64 input coordinates were set to the same normalized scalar x, and the 32 layer outputs were averaged. This yields a mean layer response along a defined feature-space path rather than an individual spline or a wavelength-specific response.

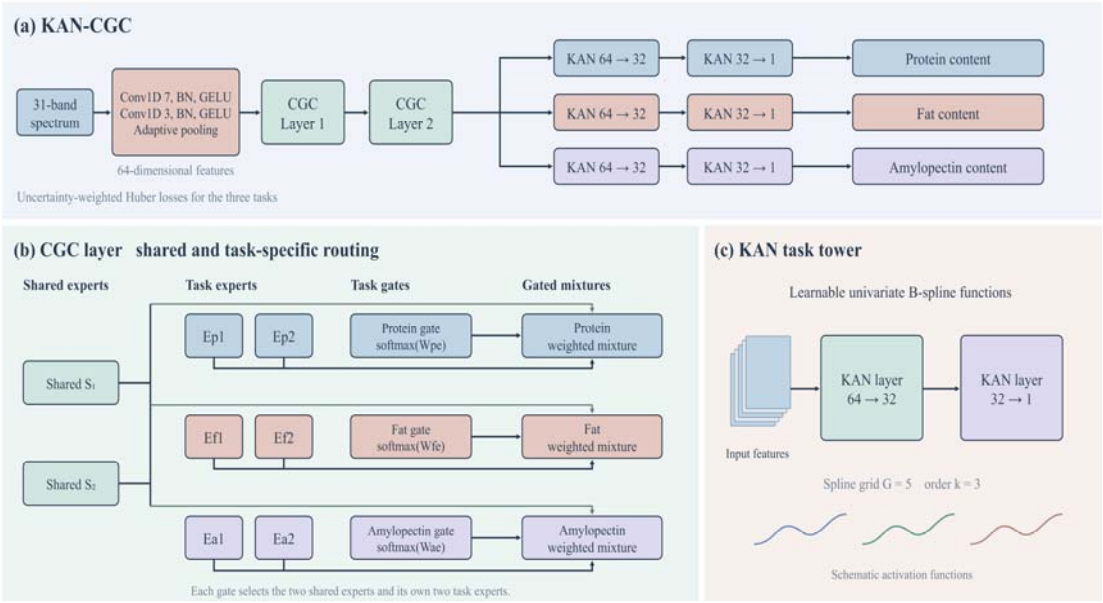


Figure 4. KAN-CGC architecture. (a) Spectral feature extraction, two CGC layers and three KAN task towers. (b) Shared and task-specific expert routing within a CGC layer. (c) KAN task tower.

Three comparison models were used. PLSR extracts latent variables that maximize covariance between predictors and responses; the number of latent variables was selected by cross-validation [23]. Support vector regression (SVR) used a radial basis function kernel, with the penalty and kernel parameters selected by grid search [24]. PLSR and SVR were fitted separately for each component. The cross-stitch network (CS-Net) retained task-specific feature branches and used learned linear combinations to exchange information between them [25]. This provided a soft-sharing multitask baseline for comparison with expert gating.

The sample set was partitioned by the sample set partitioning based on joint x–y distances (SPXY) method at the physical-sample level, with a 4:1 calibration-to-prediction ratio. Models within each comparison used the same partition. Hyperparameters were selected within training folds; the prediction set was excluded from model selection. Component prediction was evaluated using calibration and prediction coefficients of determination (R_c^2 and R_p^2), calibration and prediction RMSE (RMSEC and RMSEP), and the ratio of performance to deviation (RPD):

$$R_c^2 = 1 - \frac{\sum_{i=1}^{n_c} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_c} (y_i - \bar{y}_c)^2} \quad (6)$$

$$RMSEC = \sqrt{\frac{1}{n_c}\sum_{i=1}^{n_c} (y_i - \hat{y}_i)^2} \quad (7)$$

$$R_p^2 = 1 - \frac{\sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_p} (y_i - \bar{y}_p)^2} \quad (8)$$

$$RPD = \frac{SD_p}{RMSEP} \quad (9)$$

n_c and n_p are the calibration and prediction sample counts; y_i and $\hat{y}_i$ are measured and predicted contents; and $\bar{y}_c$ and $\bar{y}_p$ are the respective set means. RMSEP is calculated as in Eq. (7), using the prediction samples. SD_p is the standard deviation of prediction reference values. Higher R^2 and RPD and lower RMSE indicate better predictive performance.

3. Results and Discussion

3.1. Establishment of the spectral reconstruction mapping

Paired pseudo-RGB and hyperspectral images were divided into 128 × 128 pixel patches at a stride of 8 and augmented by random rotations and flips. RGB values were scaled to [0, 1]. The 512 × 512 input shown in the workflow

denotes whole-image inference after resizing, rather than training patch size. Models were trained in PyTorch on an Intel Core i7-13700 processor and NVIDIA RTX 4070 Ti GPU. Adam optimization used $\beta_1 = 0.9$, $\beta_2 = 0.999$, an initial learning rate of 4×10^{-4}, a batch size of 4 and 200 epochs. The training objective was MRAE with a smoothing term.

Training losses stabilized after approximately 150 epochs (Fig. 5). MST++ maintained the highest test PSNR during late training and achieved the best values among the three models in Table 1: MRAE = 0.0240, RMSE = 0.0109 and PSNR = 39.2855 dB. Its PSNR exceeded HSCNN-D and HRNet by 2.5762 and 1.0448 dB, respectively. MST++ was therefore selected for downstream component prediction.

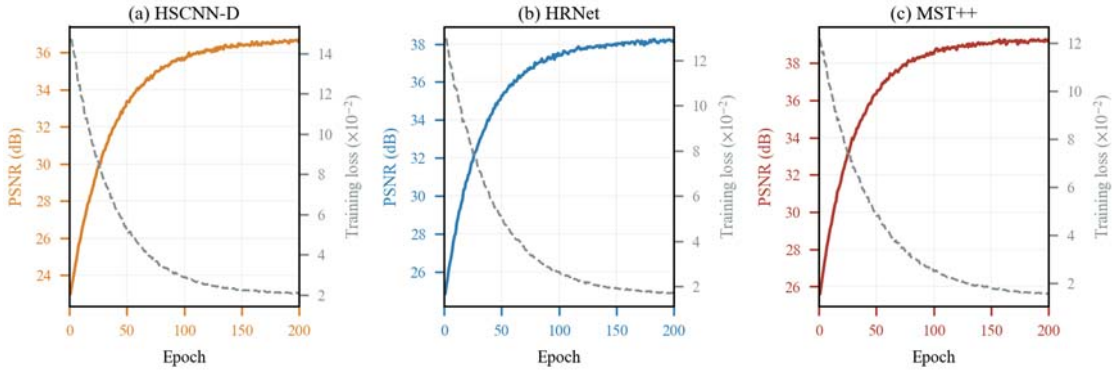


Figure 5. Training histories of the reconstruction models. (a) HSCNN-D. (b) HRNet. (c) MST++.

Table 1. Test performance of spectral reconstruction models.

Model	Test MRAE	Test RMSE	Test PSNR (dB)
HSCNN-D	0.0342	0.0147	36.7093
HRNet	0.0268	0.0123	38.2407
MST++	0.0240	0.0109	39.2855

All metrics were evaluated on the reconstruction test set.

3.2. Evaluation of reconstructed spectral curves and error maps

Measured and reconstructed spectra were extracted at the same spatial position from one sample of each variety (Fig. 6). All three models recovered the broad reflectance trend over 400–700 nm, but MST++ most closely followed the measured curves, HRNet ranked second, and HSCNN-D showed larger fluctuations. Deviations were more apparent near 400–430 and 690–700 nm than in the middle of the range. Reduced sensor response and noise near the spectral limits may contribute to these differences.

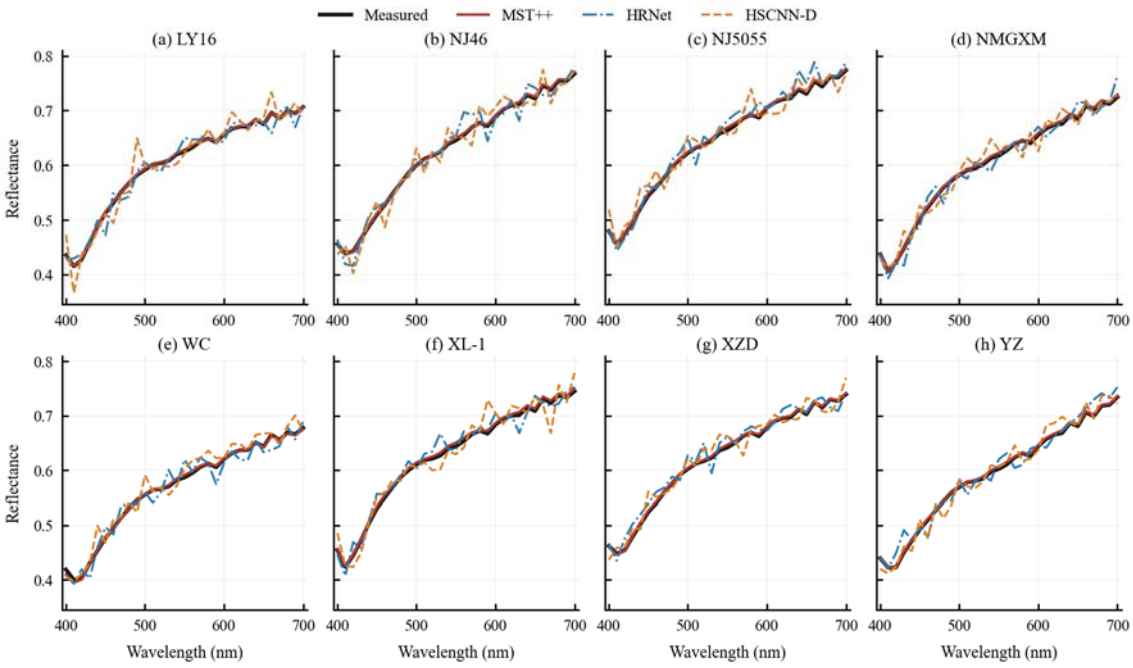


Figure 6. Measured and reconstructed spectra of eight rice varieties. (a) Longyang 16. (b) Nanjing 46. (c) Nanjing 5055. (d) Inner Mongolia fragrant rice. (e) Wuchang fragrant rice. (f) Xinglong 1. (g) Xiaozhan rice. (h) Yinzhen.

MRAE maps of a randomly selected test sample showed the spatial distribution of errors at 400, 500, 600 and 700 nm

and their average across all bands (Fig. 7). Errors were highest at 400 nm, where vertical patterns and edge regions were visible for all models. Errors decreased at 500 nm. At 600 and 700 nm, HSCNN-D and HRNet retained brighter regions near the sample boundary, whereas MST++ was more uniform. The all-band maps followed the same ranking, supporting the quantitative comparison in Table 1.

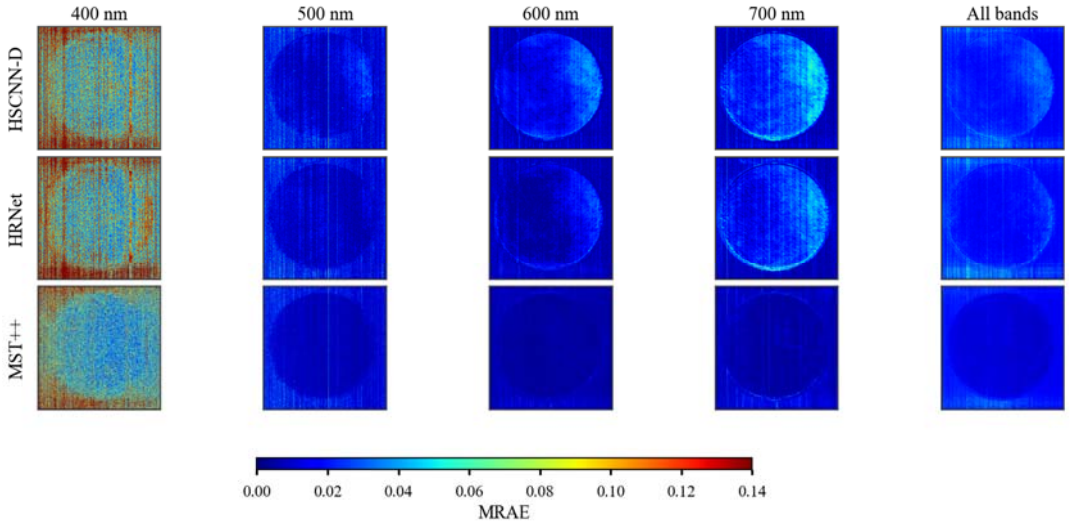


Figure 7. MRAE maps for three reconstruction methods.

3.3. Spatial quality of reconstructed band images

Actual reconstructed band slices from the same rice sample are shown at 460, 520, 580, 640 and 680 nm in Fig. 8. Grain boundaries in HSCNN-D images were blurred, with smoothed gaps and visible noise. HRNet preserved clearer contours but retained some diffusion at 680 nm. MST++ gave the clearest grain texture and intensity transitions relative to the original image, with limited noise at 460 nm. These visual differences agreed with the reconstruction metrics and supported the use of MST++ for spectral extraction.

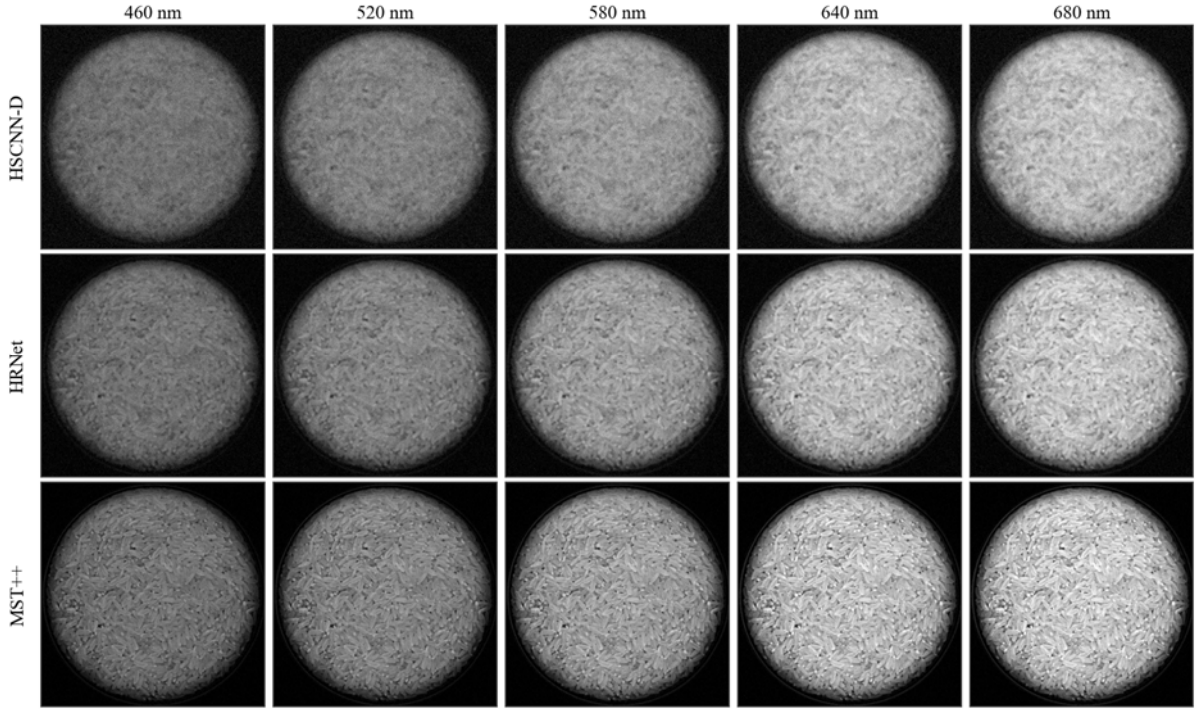


Figure 8. Actual spectral-band slices reconstructed by the three models for the same sample.

3.4. Selection of spectral preprocessing

MSC gave the highest prediction R^2 for all three components and all three reconstruction methods, whereas SG gave the lowest values (Table 2). For MST++ spectra, MSC produced R_p^2 values of 0.9017, 0.8493 and 0.8856, averaging 0.8789, an absolute increase of 0.0221 over no preprocessing. Directly acquired full spectra showed the same pattern: MSC increased the mean R_p^2 from 0.8674 to 0.8876 (Table 3). These results support scatter correction for both data sources. The lower accuracy after SG smoothing indicates that the chosen smoothing settings were less suitable for retaining predictive information. MSC was used for subsequent model comparisons.

Table 2. Prediction R² of PLSR models using reconstructed spectra with different preprocessing methods.

Reconstruction	Preprocessing	Protein R_p^2	Fat R_p^2	Amylopectin R_p^2	Mean R_p^2
HSCNN-D	None	0.8462	0.7924	0.8358	0.8248
HSCNN-D	SG	0.8305	0.7736	0.8189	0.8077
HSCNN-D	SNV	0.8691	0.8173	0.8524	0.8463
HSCNN-D	MSC	0.8783	0.8269	0.8602	0.8551
HRNet	None	0.8674	0.8102	0.8531	0.8436
HRNet	SG	0.8519	0.7948	0.8376	0.8281
HRNet	SNV	0.8862	0.8327	0.8709	0.8633
HRNet	MSC	0.8924	0.8401	0.8763	0.8696
MST++	None	0.8796	0.8261	0.8648	0.8568
MST++	SG	0.8634	0.8095	0.8497	0.8409
MST++	SNV	0.8941	0.8419	0.8782	0.8714
MST++	MSC	0.9017	0.8493	0.8856	0.8789

Mean R_p^2 is the arithmetic mean across the three components. SG, Savitzky–Golay smoothing; SNV, standard normal variate; MSC, multiplicative scatter correction.

Table 3. Prediction R² of PLSR models using directly acquired full spectra.

Preprocessing	Protein R_p^2	Fat R_p^2	Amylopectin R_p^2	Mean R_p^2
None	0.8914	0.8382	0.8726	0.8674
SG	0.8772	0.8215	0.8583	0.8523
SNV	0.9028	0.8506	0.8841	0.8792
MSC	0.9106	0.8594	0.8928	0.8876

Mean R_p^2 is the arithmetic mean across the three components.

3.5. Component prediction from reconstructed spectra

With MSC-treated MST++ spectra, KAN-CGC achieved the highest prediction accuracy among the four regressors for every component (Table 4). Its R_p^2 values were 0.9824, 0.9612 and 0.9803 for protein, fat and amylopectin, with RMSEP values of 0.1048, 0.0081 and 0.8177 g/100 g, respectively. Fat prediction improved from R_p^2 = 0.8493 with PLSR and 0.8716 with SVR to 0.9256 with CS-Net and 0.9612 with KAN-CGC. RPD values of 7.78, 5.24 and 7.36 indicated useful quantitative performance within the evaluated sample range. Calibration and prediction points were distributed close to the identity line (Fig. 9).

Table 4. Component prediction using MSC-treated reconstructed spectra.

Component	Model	R_c^2	RMSEC	R_p^2	RMSEP	RPD
Protein	PLSR	0.9177	0.1996	0.9017	0.2476	3.29
	SVR	0.9452	0.1629	0.9312	0.2072	3.94
	CS-Net	0.9723	0.1158	0.9603	0.1574	5.18
	KAN-CGC	0.9924	0.0607	0.9824	0.1048	7.78
Fat	PLSR	0.8653	0.0121	0.8493	0.0159	2.66
	SVR	0.8856	0.0112	0.8716	0.0147	2.88
	CS-Net	0.9376	0.0082	0.9256	0.0112	3.79
	KAN-CGC	0.9712	0.0056	0.9612	0.0081	5.24
Amylopectin	PLSR	0.9016	1.1732	0.8856	1.9705	3.05
	SVR	0.9268	1.0119	0.9128	1.7203	3.50
	CS-Net	0.9641	0.7087	0.9521	1.2750	4.72
	KAN-CGC	0.9903	0.3684	0.9803	0.8177	7.36

The 80 physical samples were partitioned into 64 calibration and 16 prediction samples by SPXY. RMSEC and RMSEP are in g/100 g.

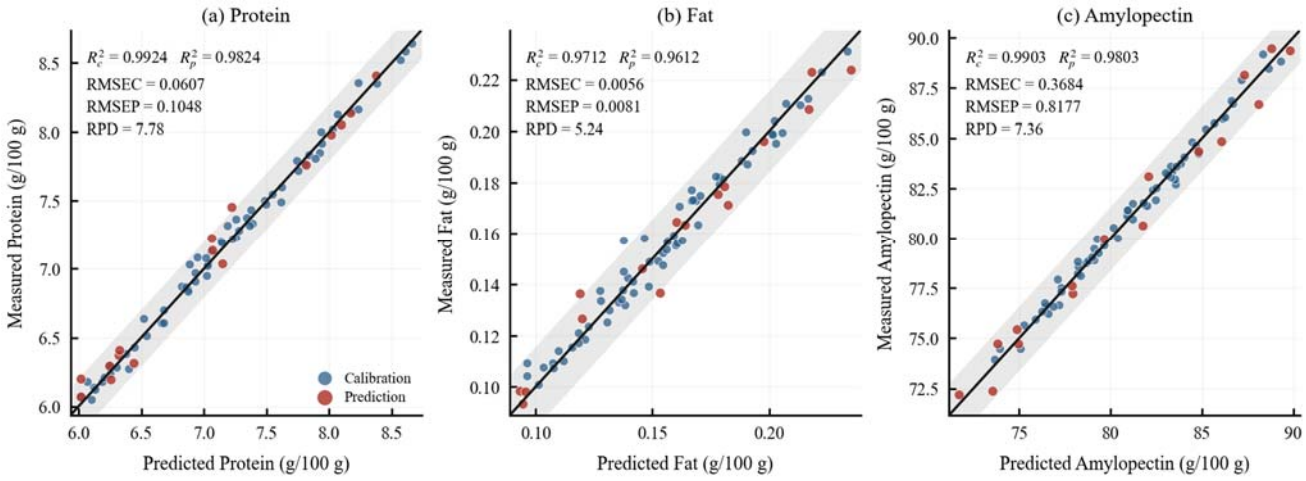


Figure 9. KAN-CGC prediction using reconstructed spectra. (a) Protein. (b) Fat. (c) Amylopectin.

Second-layer CGC weights showed that both shared experts contributed to all tasks, with weights ranging from 0.187 to 0.325 (Fig. 10a). Protein-specific experts received weights of 0.302 and 0.127, fat-specific experts 0.336 and 0.233, and amylopectin-specific experts 0.298 and 0.221. Noncandidate experts were excluded structurally. Along the diagonal input scan defined in Section 2.7, the protein tower had an increasing, S-shaped mean first-layer response; the fat response decreased before a slight terminal increase; and the amylopectin response increased, decreased and then increased again (Fig. 10b). These curves reveal different task-dependent feature mappings. Their input is the learned 64-dimensional feature space, so wavelength attribution would require a separate analysis.

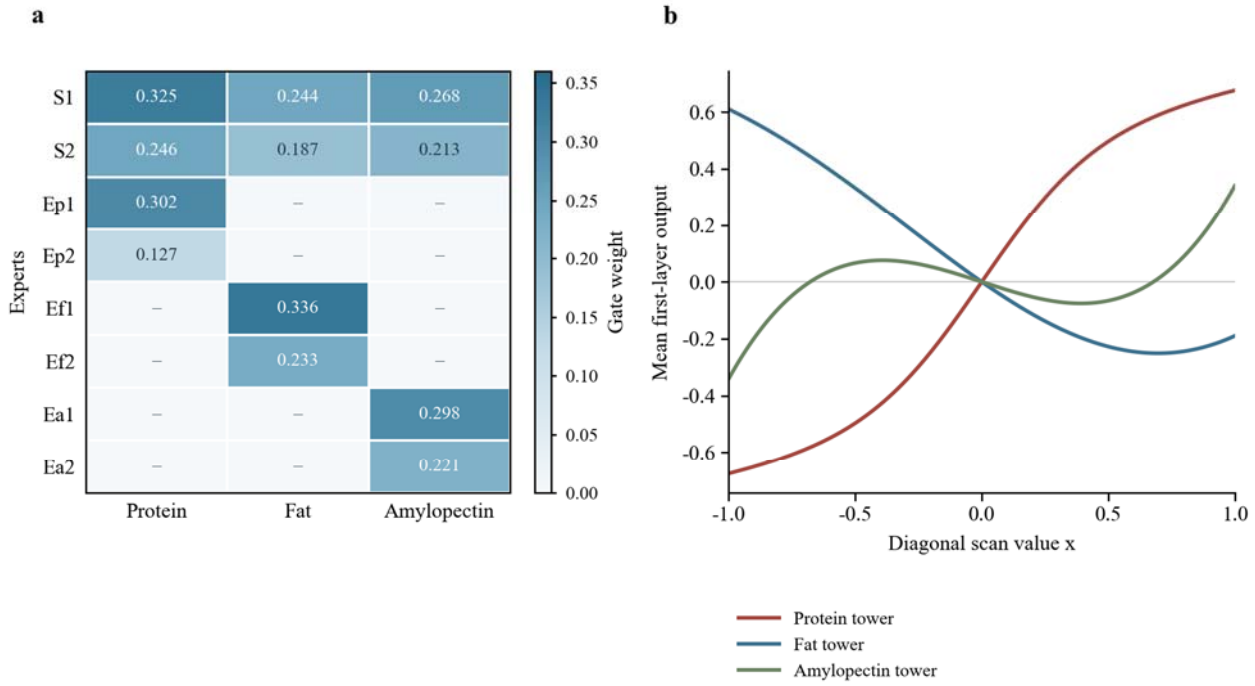


Figure 10. KAN-CGC gate weights and feature-space responses. (a) Second-layer CGC weights. (b) Mean outputs of each task.

3.6. Component prediction from directly acquired full spectra

Models developed from the 448 directly acquired bands (400–1000 nm) followed the same overall ranking, with KAN-CGC again giving the highest R_p^2 for all components (Table 5). Its protein, fat and amylopectin R_p^2 values were 0.9887, 0.9701 and 0.9869, respectively, with RPD values of 9.72, 5.97 and 9.02. The mean R_p^2 was 0.9819 for full spectra and 0.9746 for reconstructed spectra, a difference of approximately 0.0073, or 0.73 percentage points. Thus, reconstructed visible spectra retained substantial predictive information, while directly acquired spectra provided higher accuracy in these comparisons.

Table 5. Component prediction using MSC-treated directly acquired full spectra.

Component	Model	R_c^2	RMSEC	R_p^2	RMSEP	RPD
Protein	PLSR	0.9266	0.1473	0.9106	0.2244	3.45
	SVR	0.9477	0.1244	0.9337	0.1932	4.01
	CS-Net	0.9838	0.0692	0.9718	0.1260	6.15
	KAN-CGC	0.9980	0.0243	0.9887	0.0798	9.72
Fat	PLSR	0.8754	0.0114	0.8594	0.0169	2.75
	SVR	0.8926	0.0106	0.8786	0.0157	2.96
	CS-Net	0.9522	0.0071	0.9402	0.0110	4.22
	KAN-CGC	0.9801	0.0046	0.9701	0.0078	5.97
Amylopectin	PLSR	0.9088	1.4933	0.8928	1.8732	3.15
	SVR	0.9287	1.3204	0.9147	1.6709	3.54
	CS-Net	0.9773	0.7450	0.9653	1.0657	5.54
	KAN-CGC	0.9969	0.2753	0.9869	0.6548	9.02

Full spectra comprised 448 bands over 400–1000 nm. SPXY partitioning used 64 calibration and 16 prediction samples. RMSEC and RMSEP are in g/100 g.

3.7. Comparison of acquisition and processing time

HSI acquisition required stage positioning followed by push-broom scanning. The stage traveled 119.8 mm to the start position at 20 mm/s and scanned 136.43 mm at 10.84 mm/s. Acquisition took 18.575 s, and white/dark correction and band extraction from approximately 700 MB of data took another 0.153 s, totaling 18.728 s (Table 6). The RGB workflow required 0.098 s for image acquisition and conversion to a matrix, 0.782 s for MST++ reconstruction, 0.023 s for preprocessing and 0.006 s for KAN-CGC inference, totaling 0.909 s. This was 95.15% shorter than the HSI acquisition and preprocessing time. Timing began after warm-up and parameter setup; Table 6 identifies the operations included in each total.

Table 6. Per-sample acquisition and processing times.

Method	Acquisition (s)	Reconstruction (s)	Preprocessing (s)	Inference (s)	Total (s)
HSI	18.575	—	0.153	—	18.728
RGB reconstruction	0.098	0.782	0.023	0.006	0.909

Times are in seconds and exclude warm-up and parameter setup. A dash indicates that the operation was not included as a separate timing entry in the workflow.

3.8. Discussion

This study extends RGB-based spectral reconstruction from single-component measurement to simultaneous prediction of three rice components. The results connect reconstruction quality to downstream regression: MST++ gave the lowest reconstruction errors, MSC improved the PLSR baseline, and KAN-CGC achieved the best component predictions among the evaluated regressors. With 1.42 million parameters, MST++ improved PSNR by 2.58 dB over HSCNN-D. The resulting reconstructed spectra supported a mean R_p^2 of 0.9746 with KAN-CGC, compared with 0.8789 with PLSR.

The three targets span a wide concentration range, making fat particularly demanding. Candidate-restricted gates provide shared and task-specific feature routes, KAN towers allow flexible nonlinear mappings, and uncertainty weighting adjusts task losses. Together, these mechanisms provide a plausible basis for the observed multitask performance. Gate weights and diagonal-scan responses show how the trained model allocates experts and transforms features; isolating the contribution of each mechanism would require ablation experiments. The shared experts remain common optimization pathways, so expert routing should not be interpreted as complete gradient isolation.

Our previous Daqu study reported R_p^2 = 0.9984 and RMSEP = 0.4093 g/100 g for moisture [9]. The lower coefficients here concern different targets and sample matrices, including a fat range of only 0.09–0.24 g/100 g, and therefore do not constitute a direct accuracy comparison. Relative to directly acquired full rice spectra, reconstruction reduced the processing time to below 1 s while giving slightly lower component prediction accuracy. This trade-off supports further development of a lower-cost acquisition route for brewing raw materials.

The present validation was performed under static laboratory imaging conditions. Transfer to other varieties, origins and factories will require broader sampling and external validation. Spline grid size and order were set empirically and warrant systematic optimization. Conveyor vibration, changing illumination and industrial operating conditions also require testing before online deployment. These steps will establish how far the accuracy and timing observed here extend to production environments.

4. Conclusions

RGB-based spectral reconstruction combined with KAN-CGC enabled simultaneous prediction of protein, fat and amylopectin in brewing rice. MST++ reconstructed 31 bands over 400–700 nm with test MRAE = 0.0240, RMSE = 0.0109 and PSNR = 39.2855 dB. MSC gave the best preprocessing results. Using reconstructed spectra, KAN-CGC achieved R_p^2 values of 0.9824, 0.9612 and 0.9803 for the three components, compared with 0.9887, 0.9701 and 0.9869 using directly

acquired full spectra. The difference in mean R_p^2 was approximately 0.73 percentage points. The RGB workflow took 0.909 s per sample, 95.15% less time than HSI acquisition and preprocessing. This approach offers a promising route to rapid multicomponent quality assessment, with production-line validation needed to establish online performance.

CRediT authorship contribution statement

Haonan Yi: Writing – original draft, Writing – review & editing. Liangliang Xie: Supervision. Kaiyang Yuan: Resources. Anying Cai: Resources.

Acknowledgments

This research was supported by the Solid-state Brewing Technology Innovation Center of Sichuan (GNLMLX202509).

Declaration of competing interest

The authors declare that they do not have any conflict of financial interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] Liu, M.-K., Liu, C.-Y., Tang, Y.-M., Liu, Y., Su, Y., Tian, X.-H., Feng, J., & Ni, X.-L. (2026). Innovative use of unhulled rice in Baijiu brewing: Impact on flavor and microbial composition. Food Chemistry: X, 34, 103569. https://doi.org/10.1016/j.fochx.2026.103569

[2] Yin, X., Yoshizaki, Y., Ikenaga, M., Han, X.-L., Okutsu, K., Futagami, T., Tamaki, H., & Takamine, K. (2020). Manufactural impact of the solid-state saccharification process in rice-flavor baijiu production. Journal of Bioscience and Bioengineering, 129(3), 315–321. https://doi.org/10.1016/j.jbiosc.2019.09.017

[3] Tse, T. J., Wiens, D. J., & Reaney, M. J. T. (2021). Production of bioethanol—A review of factors affecting ethanol yield. Fermentation, 7(4), 268. https://doi.org/10.3390/fermentation7040268

[4] El Hosry, L., Elias, V., Chamoun, V., Halawi, M., Cayot, P., Nehme, A., & Bou-Maroun, E. (2025). Maillard reaction: Mechanism, influencing parameters, advantages, disadvantages, and food industrial applications: A review. Foods, 14(11), 1881. https://doi.org/10.3390/foods14111881

[5] Galanakis, C. M. (Ed.). (2021). Innovative food analysis. Academic Press. https://doi.org/10.1016/C2019-0-00106-9

[6] Xia, Y., Tian, J., Zhou, Y., Huang, D., Xie, L., Hu, X., Yang, H., & Shang, J. (2026). Rapid multi-task detection of rice components using convolutional bidirectional long short-term fusion network combined with hyperspectral imaging. Food Control, 181, 111701. https://doi.org/10.1016/j.foodcont.2025.111701

[7] Fei, X., Jiang, X., Lei, Y., Tian, J., Hu, X., Bu, Y., Huang, D., & Luo, H. (2023). The rapid non-destructive detection of the protein and fat contents of sorghum based on hyperspectral imaging. Food Analytical Methods, 16(11–12), 1690–1701. https://doi.org/10.1007/s12161-023-02529-x

[8] Huang, H., Hu, X., Tian, J., Jiang, X., Sun, T., Luo, H., & Huang, D. (2021). Rapid and nondestructive prediction of amylose and amylopectin contents in sorghum based on hyperspectral imaging. Food Chemistry, 359, 129954. https://doi.org/10.1016/j.foodchem.2021.129954

[9] Xie, L., Zhang, H., Tian, J., Yu, S., Zhang, X., Yang, H., Hu, X., Chen, M., Wang, R., Cai, A., Yang, J., & Liu, S. (2025). Study on the real-time online determination of the moisture content of Baijiu Daqu during fermentation: New application of spectral reconstruction technology. Journal of Food Composition and Analysis, 146, 107904. https://doi.org/10.1016/j.jfca.2025.107904

[10] Zhang, J., Su, R., Fu, Q., Ren, W., Heide, F., & Nie, Y. (2022). A survey on computational spectral reconstruction methods from RGB to hyperspectral imaging. Scientific Reports, 12(1), 11905. https://doi.org/10.1038/s41598-022-16223-1

[11] Zhao, J., Kechasov, D., Rewald, B., Bodner, G., Verheul, M., Clarke, N., & Clarke, J. L. (2020). Deep learning in hyperspectral image reconstruction from single RGB images—A case study on tomato quality parameters. Remote Sensing, 12(19), 3258. https://doi.org/10.3390/rs12193258

[12] Ahmed, M. T., Monjur, O., & Kamruzzaman, M. (2024). Deep learning-based hyperspectral image reconstruction for quality assessment of agro-product. Journal of Food Engineering, 382, 112223. https://doi.org/10.1016/j.jfoodeng.2024.112223

[13] Lee, H. B., Yang, E., & Hwang, S. J. (2018). Deep asymmetric multi-task feature learning. In Proceedings of the 35th International Conference on Machine Learning (Vol. 80, pp. 2956–2964). https://proceedings.mlr.press/v80/lee18d.html

[14] Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., Pfister, H., Timofte, R., & Van Gool, L. (2022). MST++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 744–754). https://doi.org/10.1109/CVPRW56347.2022.00090

[15] Maleki, M. R., Mouazen, A. M., Ramon, H., & De Baerdemaeker, J. (2007). Multiplicative scatter correction during on-line measurement with near infrared spectroscopy. Biosystems Engineering, 96(3), 427–433. https://doi.org/10.1016/j.biosystemseng.2006.11.014

[16] Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1993). Correction to the description of standard normal variate (SNV) and de-trend (DT) transformations. Journal of Near Infrared Spectroscopy, 1(3), 185–186. https://doi.org/10.1255/jnirs.21

[17] Zimmermann, B., & Kohler, A. (2013). Optimizing Savitzky–Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. Applied Spectroscopy, 67(8), 892–902. https://doi.org/10.1366/12-06723

[18] Shi, Z., Chen, C., Xiong, Z., Liu, D., & Wu, F. (2018). HSCNN+: Advanced CNN-based hyperspectral recovery from RGB images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 939–947). https://doi.org/10.1109/CVPRW.2018.00139

[19] Yu, C., Xiao, B., Gao, C., Yuan, L., Zhang, L., Sang, N., & Wang, J. (2021). Lite-HRNet: A lightweight high-resolution network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10440–10450). https://doi.org/10.1109/CVPR46437.2021.01030

[20] Tang, H., Liu, J., Zhao, M., & Gong, X. (2020). Progressive layered extraction (PLE): A novel multi-task learning (MTL) model for personalized recommendations. In Proceedings of the 14th ACM Conference on Recommender Systems (pp. 269–278). https://doi.org/10.1145/3383313.3412236

[21] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., & Tegmark, M. (2024). KAN:

Kolmogorov-Arnold networks. arXiv preprint. https://doi.org/10.48550/arXiv.2404.19756

[22] Kendall, A., Gal, Y., & Cipolla, R. (2018). Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7482–7491). https://doi.org/10.1109/CVPR.2018.00781

[23] Carrascal, L. M., Galván, I., & Gordo, O. (2009). Partial least squares regression as an alternative to current regression methods used in ecology. Oikos, 118(5), 681–690. https://doi.org/10.1111/j.1600-0706.2008.16881.x

[24] Qiu, J., Sheffler, W., Baker, D., & Noble, W. S. (2008). Ranking predicted protein structures with support vector regression. Proteins, 71(3), 1175–1182. https://doi.org/10.1002/prot.21809

[25] Misra, I., Shrivastava, A., Gupta, A., & Hebert, M. (2016). Cross-stitch networks for multi-task learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3994–4003). https://doi.org/10.1109/CVPR.2016.433