Shahid Bahonar University of Kerman
Biomechanism and Bioenergy Research
Online ISSN: 2821-1855
Homepage: https://bbr.uk.ac.ir
Iranian Society of Agricultural Machinery Engineering and Mechanization

A Hybrid Electronic Nose and Computer Vision System for Non Destructive Detection of Sulfur Dioxide (SO_2) Residues in Raisins

Mohammad Ghoushchian[1], Seyed Saeid Mohtasebi[1]✉, Shahin Rafiee[1]

[1] Department of Agricultural Machinery Engineering, Faculty of Agriculture, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran

✉ Corresponding author: mohtaseb@ut.ac.ir

ARTICLE INFO

Article type:

Research Article

Article history:

Received 11 May 2026

Received in revised form 29 June 2026

Accepted 20 August 2026

Available Online 30 September 2026

Keywords:

Electronic Nose, Gas Sensors, Computer Vision, Raisin Quality, Sulfur Dioxide

ABSTRACT

Sulfur dioxide (SO_2) is widely used as a preservative in raisins, but excessive residues harm health and restrict exports. Standard iodometric titration (ISIRI No. 545) is accurate yet destructive, slow, and requires a well-equipped laboratory. This study developed a low-cost, portable, non-destructive hybrid system combining an electronic nose (E-nose) and computer vision. Three raisin treatments (sun-dried, acid-dipped, and sulfur-fumigated-then-acid-dipped) were prepared, each with three replicates. The E-nose used eight MOS sensors (MQ9, MQ3, MQ5, TGS2620, MQ135, TGS822, TGS813, and MQ4). Sensor responses were recorded for 60 minutes. RAW images were captured in a dark box (2500 lux). After segmentation, 13 color components (RGB, HSV, L*a*b*, grayscale) were extracted, and six first-order statistical features per component gave 98 features per image. Gradient Boost Regression (GBR) achieved $R^2 = 0.9972$ and RMSE = 0.0031 on test data – far better than linear regression. Sensors MQ9, MQ3, MQ5, and TGS2620 were the most accurate (R^2 = 0.8668–0.9458). ANOVA showed significant effects of treatment and sensor ($p < 0.001$). An artificial neural network (two hidden layers, 6 and 8 neurons) reached R = 0.986 and RMSE = 0.0324. Fusing olfactory and visual data using PCA and SVM improved accuracy. The hybrid system is inexpensive (< $100), portable, battery-powered, and includes an Android app – a practical, real-time alternative to laboratory titration.

Cite this article: **Ghoushchian, M., Mohtasebi, S. S., & Rafiee, SH. (2026).** A Hybrid Electronic Nose and Computer Vision System for Non Destructive Detection of Sulfur Dioxide (SO_2) Residues in Raisins. *Biomechanism and Bioenergy Research*, *5*(3), 1-18. https://doi.org/10.22103/bbr.2026.27286.1154

Publisher: Shahid Bahonar University of Kerman
DOI: https://doi.org/10.22103/bbr.2026.27286.1154

INTRODUCTION

Raisins – dried grapes (*Vitis vinifera*) – are produced from different grape varieties using various drying methods. Iran is consistently among the world's top three raisin producers, alongside Turkey and the United States. A substantial share of Iranian raisins is exported to the European Union, Russia, and the Persian Gulf states, where product safety and compliance with international residue limits are strictly enforced.

In commercial raisin processing, sulfur dioxide (SO_2) has been used for decades. Its antioxidant and antimicrobial properties prevent microorganisms from growing, preserve the natural color and taste, and prevent browning during storage (Mischek & Krapfenbauer-Cermak, 2012). The most common application methods are direct fumigation with SO_2 gas produced by burning sulfur powder, or dipping the grapes in solutions of sulfite salts. Both leave residues in the final product. Sulfur is also used directly in vineyards as a pesticide (copper sulfate), so some residual sulfur can come from the field (Liu et al., 2018). Using sulfites during drying shortens the process (Stražanac et al., 2019; Türkyılmaz et al., 2013). SO_2 also protects fruits from insect attacks and helps keep the natural color by preventing enzymatic and non-enzymatic browning (Cantín et al., 2012; Guerrero & Cantos-Villar, 2015).

However, SO_2 is not harmless. Sulfites are allergens. They can cause severe allergic reactions in people with asthma and in those who have low sulfite oxidase activity (Soubra et al., 2007). There are also reports that these additives can cause skin reactions and even DNA damage (Meng et al., 2005; Vally et al., 2009). Consequently, food safety authorities have set maximum residue limits (MRLs). For example, the European Food Safety Authority (EFSA) has set an MRL of 2000 $mg \cdot kg^{-1}$ for dried apricots, peaches, and grapes, but only 500 $mg \cdot kg^{-1}$ for other dried fruits. The Codex Alimentarius recommends 1500 $mg \cdot kg^{-1}$ for raisins (Laganà et al., 2017).

Iran's national standard (ISIRI No. 569) is more detailed. It sets 700 $mg \cdot kg^{-1}$ for Sultana raisins and 1500 $mg \cdot kg^{-1}$ for golden (angora) raisins (ISIRI, 2014). For dried apricots the limit is 2000 $mg \cdot kg^{-1}$ (ISIRI, 2014; 2015), and for peach-based fruit snacks it is also 2000 $mg \cdot kg^{-1}$ (ISIRI, 2018). Therefore, measuring SO_2 residues accurately is not only a matter of consumer health but also a key requirement for maintaining export markets. The skin of raisins is less permeable to SO_2 compared to other dried fruits like apricots or pears, making high SO_2 concentrations more difficult to achieve and measure.

The conventional way to measure SO_2 residues in dried fruits is iodometric titration, described in Iranian National Standard No. 545 (ISIRI, 2014). The principle involves acidifying the sample, heating it to release SO_2, trapping the gas in a hydrogen peroxide solution (where it becomes sulfuric acid), and then titrating the acid with sodium hydroxide. Alternatively, the trapped SO_2 can be titrated directly with iodine in the presence of starch indicator.

These methods are accurate and reproducible. But they have several well-known disadvantages. They are destructive – once the sample is tested, it cannot be sold or reused. They are slow, taking several hours per analysis. They require a fully equipped laboratory, skilled chemists, fume hoods, distillation apparatus, and precise burettes, so they cannot be used on a factory floor or at a customs checkpoint. They are not real-time – by the time the result is obtained, the product has already left the production line. Other advanced techniques (ion chromatography, molecular absorption spectrometry, atomic absorption spectroscopy, UV fluorescence) are even more expensive and also lab-bound. Hence, there is a clear need for a non-destructive, fast, portable, and low-cost method for SO_2 detection in raisins.

Over the past twenty years, electronic nose (E-nose) technology has become a promising non-destructive tool for food quality evaluation. An E-nose uses an array of gas sensors to produce a unique "smell fingerprint" for each sample (Gardner & Bartlett, 1994). The response pattern can be analyzed by pattern recognition algorithms to classify samples or predict the concentration of specific compounds.

Among gas sensor technologies, metal-oxide-semiconductor (MOS) sensors are the most popular. They work by measuring changes in electrical resistance when gas molecules adsorb onto the heated metal oxide surface. Their main advantages are high sensitivity, fast response, long lifetime, moderate power consumption, and low cost (typically $10-20 per sensor). MOS sensors have been used in many food applications – detecting fungal contamination in grains, checking the freshness of meat and fish, monitoring fruit ripening, and identifying adulteration (Girmatsion et al., 2025).

But MOS sensors suffer from drift (baseline and sensitivity change over time, harming long-term reproducibility), cross-sensitivity (they respond to many different gases), non-linearity (resistance vs. concentration follows a power law or exponential curve), and environmental sensitivity (temperature and humidity strongly influence readings). Machine learning (ML) can help overcome these limitations. Unlike simple linear calibration, ML models capture non-linear relationships, use time or environmental variables to compensate for drift, and exploit cross-sensor correlations to improve selectivity (Girmatsion et al., 2025).

Among ML algorithms, ensemble methods have proven particularly effective for sensor data. Gradient Boost Regression (GBR) builds a strong predictive model from a collection of weak learners – typically shallow decision trees. The algorithm works sequentially: each new tree tries to predict the errors (residuals) of the previous trees. The final prediction is the sum of all tree predictions, each weighted by a learning rate (Friedman, 2001).

GBR has several attractive features: high accuracy by focusing on the most difficult samples, built-in regularization to prevent overfitting, easy handling of mixed data types, natural feature importance ranking, and robustness to outliers. Recent studies have applied GBR to gas sensor problems. Quan et al. designed a sensor array and obtained 100 % classification accuracy with a GBDT model (RMSE = 23.52 ppm) (Quan et al., 2024). Compared several ML models for noxious gas identification and reported that Gradient Boosting reached 100 % accuracy when three or more principal components were used. Yang et al. used GBR to model the microwave pyrolysis products of biomass and obtained $R^2 > 0.822$. Nevertheless, to the best of our knowledge, GBR has never been applied to an E-nose sensor array for SO_2 detection in raisins (Yang et al., 2022).

While an E-nose captures volatile compounds, computer vision records the external appearance – color, texture, shape, and size. For SO_2-treated raisins, color changes are particularly important because SO_2 bleaches pigments, making the raisins lighter and more uniform. Color can be measured objectively using digital imaging and color spaces such as RGB, HSV, and CIE L*a*b. *The L*a*b* space is especially useful because it separates lightness (L*) from two color axes: a** (green-red) and b* (blue-yellow). For dried apricots, SO_2 fumigation increases L* and decreases a* and b* (Guerrero & Cantos-Villar, 2015).

Computer vision was first used in agriculture and food processing for quality control and sorting (Pathmanaban et al., 2019). For raisins, researchers have tried different approaches. Khojastehnazhand and Ramezani used a simple video camera to separate dark and light raisins with about 92 % accuracy (Khojastehnazhand & Ramezani, 2020). Mehraban Sangatash et al. extracted color features (RGB, L, *a*, b*) and used a multilayer neural network to classify raisins into five SO_2 groups, achieving approximately 88 % accuracy. They also measured SO_2 by the national standard method. Another study by the same group (Mehraban Sangatash et al., 2015) analyzed 62 raisin samples from five cities in Khorasan province and reported average SO_2 residues ranging from 186 to 2484 $mg \cdot kg^{-1}$.

Despite these promising results, vision alone cannot detect SO_2 that is trapped inside the fruit or when color changes are very small at low concentrations. Similarly, an E-nose alone may miss color differences that are obvious to the eye.

This complementarity is why combining the two technologies makes sense.

Data fusion integrates information from two or more sources to obtain a more accurate or robust decision than any single source could provide. In food quality assessment, combining an E-nose with computer vision (and sometimes an electronic tongue) has been explored for several products. Kiani et al. fused E-nose and computer vision to detect saffron adulteration, achieving better performance than either method alone (Kiani et al., 2017). Khulal et al. combined an E-nose and a color sensor to monitor chicken meat spoilage (Khulal et al., 2016). Wang et al. used a real-time hybrid system to distinguish medicinal plants such as borage and fennel, achieving an average classification accuracy of approximately 93.5% (Wang et al., 2020). Guo et al. reported that combining E-nose, computer vision, and NIR spectroscopy improved tea quality grading to over 90 % accuracy (Guo et al., 2025). Liu et al. used hyperspectral imaging to detect sulfur-fumigated *Flos Lonicerae* (R = 0.911, RMSE = 0.335) (Liu et al., 2018). More recently, Chen et al. summarized progress in multi-modal flexible sensors and AI-driven fusion for agricultural quality detection, concluding that fusion substantially improves accuracy, reliability, and traceability (Chen et al., 2026). Wang and Wang achieved 99.06 % accuracy for peanut origin traceability using a hybrid neural network (Wang & Wang, 2026). Asadi et al. evaluated fusion of computer vision and E-nose for kiwifruit quality; PCA-SVM gave 93.33 % with single modalities, while fusion increased the rate to 100 % and improved regression R^2 from 90.2 % to 98.6 % (Asadi et al., 2024).

Nevertheless, despite this growing body of work on multi-modal fusion, no previous study has combined an E-nose and a computer vision system specifically for SO_2 detection in raisins. This is the gap that our research aims to fill.

Research gap and objectives

From the literature review, several important gaps remain:

- No study has built a hybrid E-nose and computer vision system for SO_2 detection in raisins.
- GBR has not been applied to model MOS sensor arrays for this application.
- The relative importance of individual sensors for detecting SO_2 in raisins has not been systematically evaluated using a robust ensemble method.
- Feature-level fusion of the two modalities has not been tested for this product and contaminant.

To address these gaps, the present study had the following objectives:

1. To design and construct a portable, low-cost hybrid system integrating an 8-sensor MOS-based E-nose with a computer vision unit.
2. To model E-nose sensor responses using GBR and compare with ordinary linear regression.
3. To extract features from images, correlate them with SO_2, and develop prediction models.
4. To fuse olfactory and visual features in order to improve classification accuracy.

MATERIALS AND METHODS

Sample Preparation

Two different seedless raisin samples from a vineyard in Qazvin, Iran, were used. From each sample, three treatments were prepared:

- **Sun-dried (control):** naturally dried without additives.
- **Acid-dipped:** immersed in a solution composed of water, 3–5% potassium carbonate (K_2CO_3), and approximately 1% edible oil (often referred to as 'Tizabi' in Persian) before sun-drying.
- **Sulfur-fumigated:** fumigated with SO_2 for 60 min (sulfur powder burned in a closed chamber), then sun-dried.

Each treatment had three replicates. To test the hybrid system's ability to distinguish between different SO_2 levels, two additional sun-dried varieties (Poloee from Qazvin and Fakhri from Maragheh) were fumigated at six time points (0, 30, 60, 90, 120, 150 min). A control (no fumigation) was also prepared. Figure 1 shows the fumigation stages.

Figure 1. Stages of fumigating raisins with sulfur gas (photograph of the fumigation chamber and process)

Measurement of Residual SO_2 (ISIRI No. 545)

Residual SO_2 was measured by iodometric titration following Iranian National Standard No. 545 (ISIRI, 2014). For each sample, 50.0 g of raisins were weighed into a 500 mL round-bottom flask. 200 mL deionised water and 10 mL of 4 M hydrochloric acid were added. The flask was connected to a steam distillation apparatus. The receiving flask contained 25 mL of 3 % hydrogen peroxide and a few drops of methyl red indicator. Distillation was carried out for 30 min, collecting about 150 mL of distillate. The distillate was titrated with 0.01 N iodine solution until a persistent blue-black color appeared (starch indicator near endpoint). A blank distillation (without sample) was run and its titer subtracted. All measurements were performed in triplicate. The SO_2 concentration ($mg \cdot kg^{-1}$) was calculated as $(V \times N \times 32.06 \times 1000) / m$, where V is the iodine volume (mL), N its normality (0.01), and m the sample mass (g).

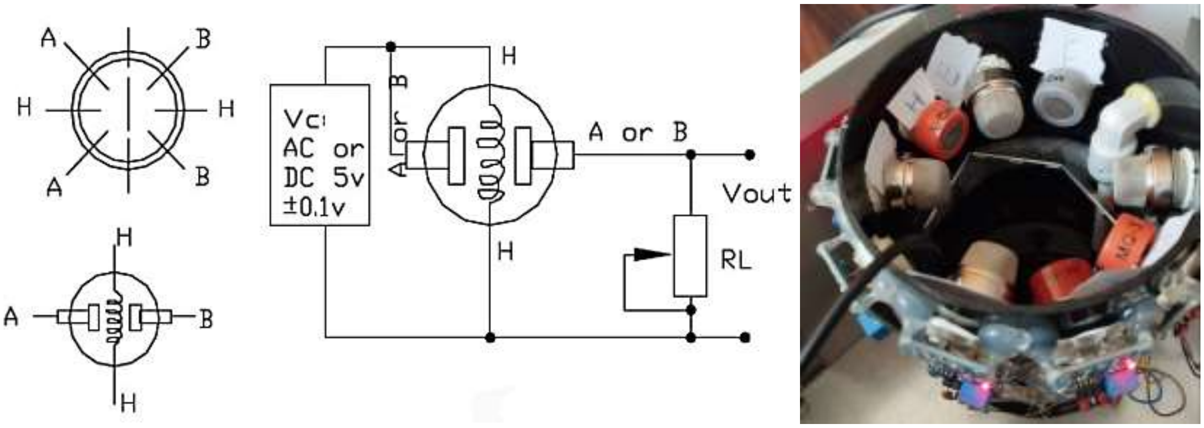
Development of the Hybrid System (Electronic Nose and Machine Vision)

Electronic nose system

The E-nose consisted of an odour chamber, an eight-sensor MOS array, a custom data acquisition board, pumps, and air-purification filters. Figure 2 shows the complete hybrid system.

Figure 2. A complete view of the constructed hybrid system (electronic nose + machine vision) as displayed at the research exhibition.

All sensors were of MOS type, including six MQ and two TGS sensors (S1=MQ-9, S2=MQ-3, S3=MQ-5, S4=TGS-2620, S5=MQ-135, S6=TGS-822, S7=TGS-813, S8=MQ-4). They were mounted on a PCB and wired according to the manufacturer's datasheets. All sensors were heated for one week before measurements (TGS sensors require seven days). Figure 3 shows the sensor array.



(a) (b)

Figure 3. Sensor array used in the electronic nose system (a) Circuit diagram of one of the sensors used in the electronic nose system and the pin configuration, (b) photograph of the custom PCB with eight MOS sensors.

Control circuit: A four-channel relay module (5 V trigger) controlled three electronic valves and the pump.

Power supply: A 12 V supply provided power (converted to 5 V for control circuit and sensors; 12 V for the pump).

Purified gas: Two activated-carbon capsules and two 12 V diaphragm pumps supplied clean air for baseline and cleaning.

Computer vision system

A dark chamber with ring-light illumination from above at 45° was integrated with the odour chamber. Illumination was approximately 2500 lux, and the camera-to-sample distance was 150 mm. The light source was a 12 V LED strip with 144 SMD chips. A Xiaomi Redmi Note 12 phone (50 MP sensor) was used for imaging. RAW images were taken using the ON1 PHOTO Android app to avoid internal color processing, ensuring consistent and reproducible color components. Figure 4 shows the imaging chamber and the lighting setup, as well as the ON1 PHOTO software environment used for capturing RAW images.

(a) (b)

Figure 4. (a) ON1 PHOTO software interface for RAW image acquisition, (b) Imaging chamber and its illumination system.

Image Processing

To prepare the captured images, a custom-written MATLAB (R2023a) code was used to detect edges, separate the raisins from the background, and remove the stems. Through this process, individual raisin grains were segmented so that only the raisins remained in the image. Grayscale images were then extracted from the resulting cropped color images. Image intensity histograms were employed for image processing. In addition to the grayscale images, various color components – including L*a*b*, RGB, and HSV – were extracted from the color images, giving a total of 13 components for analysis. Subsequently, the following statistical features were extracted for each of the obtained components.

Statistical Features from Images

For each of the 13 components, the intensity histogram was normalized, and the following first-order statistical features were extracted: mean, standard deviation, smoothness, skewness, uniformity (energy), and entropy. This yielded $6 \times 13 = 98$ features per image. Sensitivity analysis and principal component analysis (PCA) were used to select the best features for predicting SO_2. Artificial neural networks (ANN) were then used to model the relationship between selected features and SO_2.

Data Acquisition System

A custom board was designed using an ATmega2560 microcontroller, an HC-05 Bluetooth module, a computer interface, a temperature/humidity sensor (DHT22), and a display. Coding was done in the Arduino IDE. Figure 5 shows the data acquisition board.

Figure 5. Data logging process from the olfactory sensors using the designed Arduino board.

Data Acquisition

Each experiment consisted of three stages: (1) baseline correction (5 min), (2) sample odor injection (60 min), and (3) sensor chamber cleaning (90 min). All experiments were carried out at approximately constant temperature and humidity (25±2 °C, 40±5 %), which were recorded. The custom board connects to a computer via cable and to a smartphone via Bluetooth; both interfaces can be used.

Sensor Response Modeling

To evaluate prediction quality, R^2, RMSE, MAE, and RRMSE were used. After normalizing the sensor data, correlation matrices of sensor responses (with each other and with time) were extracted for different treatments and replicates. Graphs comparing predicted vs. actual values were plotted for ordinary regression and GBR. GBR hyperparameters were: learning rate = 0.01, max depth = 4, number of estimators = 500, subsample = 0.5. To avoid overfitting, we performed 5-fold cross-validation on the training set, yielding consistent results (R^2 = 0.995 ± 0.002). ANOVA was performed for the GBR model for each prediction quality component, separately for treatments, sensors, and their combinations.

Fusion of Visual and Olfactory Data

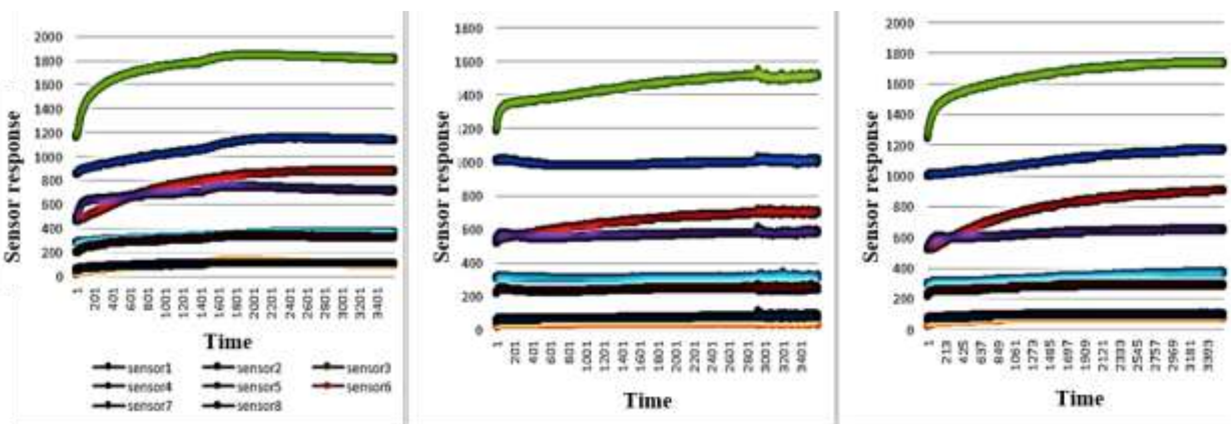
Feature-level and decision-level fusion strategies were used. KNN, SVM, and multinomial logistic regression (MLR) were used for classification modeling. PCA reduced the E-nose data to 5 components and the image features to 15 components; these were concatenated (20 features) and input to an SVM classifier.

RESULTS AND DISCUSSION

Sensor Response Curves

Three experiments were conducted: (1) sensor response recording under the three treatments, (2) calibration using six fumigation times, and (3) image feature extraction and fusion.

Figure 6 shows the average responses of the eight MOS sensors (three replicates) for the three treatments. Each treatment produced a distinct response pattern. Sensors S1 (MQ-9), S2 (MQ-3), and S3 (MQ-5) reacted fastest and strongest. S6-S8 (TGS822, TGS813, MQ-4) gave weak signals and showed larger baseline drift.



(a) (b) (c)

Figure 6. Data logging graph of the eight olfactory sensors showing real-time responses (a) Acid dipped (b) Sulfur fumigated (c) Sun dried (control)

Correlation Matrices

Tables 1, 2, and 3 present the correlation matrices among the eight different sensors for the three treatments (acid-dipped, sulfur-fumigated, and sun-dried). These matrices were generated using correlation analysis on data collected over time. Each cell in these tables shows the correlation coefficient between two sensors. The correlation indicates a direct or inverse relationship between two variables. A zero correlation between two sensors means there is no relationship between them. A correlation of 1 indicates a perfect direct or inverse relationship. A positive correlation between two sensors indicates a direct relationship. For example, the correlation between Sensor 1 and Sensor 2 is 0.976, showing that the two sensors are related and exhibit similar changes.

According to the correlation values, it can be concluded that some sensor pairs have strong positive relationships. For instance, in the acid-dipped treatment, Sensors 1 and 2 have a strong positive correlation. Also, Sensors 4 and 5 show a strong positive correlation (r = 0.966). On the other hand, some sensor pairs show negative correlations. These results can be used in designing control and prediction systems. For example, given the positive relationship between Sensor 1 and Sensor 2, one can predict that if the value of Sensor 1 increases, the value of Sensor 2 will also increase accordingly.

Table 1. Correlation matrix of gas sensors for the acid-dipped treatment.

	Time	**Sensor1**	**Sensor2**	**Sensor3**	**Sensor4**	**Sensor5**	**Sensor6**	**Sensor7**	**Sensor8**
Time	1	0/915876	0/891561	0/727372	0/709883	0/953525	0/605007	0/777876	0/79014
Sensor1	0/915876	1	0/976056	0/884079	0/907079	0/972177	0/843345	0/902621	0/941014
Sensor2	0/891561	0/976056	1	0/932071	0/907399	0/9416	0/878706	0/935074	0/956826
Sensor3	0/727372	0/884079	0/932071	1	0/950336	0/816418	0/929412	0/933298	0/964324
Sensor4	0/709883	0/907079	0/907399	0/950336	1	0/83248	0/945793	0/919864	0/965705
Sensor5	0/953525	0/972177	0/9416	0/816418	0/83248	1	0/751569	0/872977	0/891544
Sensor6	0/605007	0/843345	0/878706	0/929412	0/945793	0/751569	1	0/913676	0/941191
Sensor7	0/777876	0/902621	0/935074	0/933298	0/919864	0/872977	0/913676	1	0/957355
Sensor8	0/79014	0/941014	0/956826	0/964324	0/965705	0/891544	0/941191	0/957355	1

Table 2. Correlation matrix of gas sensors for the sulfur-fumigated treatment.

	Time	**Sensor1**	**Sensor2**	**Sensor3**	**Sensor4**	**Sensor5**	**Sensor6**	**Sensor7**	**Sensor8**
Time	1	0/437001	0/952921	0/936539	0/767123	0/39632	0/752082	0/690451	0/628973
Sensor1	0/437001	1	0/261158	0/25618	0/611412	0/603359	0/238775	0/308056	0/479587
Sensor2	0/952921	0/261158	1	0/980572	0/724904	0/295434	0/785909	0/696384	0/639157
Sensor3	0/936539	0/25618	0/980572	1	0/756758	0/277445	0/772703	0/689256	0/660113
Sensor4	0/767123	0/611412	0/724904	0/756758	1	0/520018	0/649446	0/622233	0/77358
Sensor5	0/39632	0/603359	0/295434	0/277445	0/520018	1	0/4067	0/511558	0/500869
Sensor6	0/752082	0/238775	0/785909	0/772703	0/649446	0/4067	1	0/702541	0/617868
Sensor7	0/690451	0/308056	0/696384	0/689256	0/622233	0/511558	0/702541	1	0/656943
Sensor8	0/628973	0/479587	0/639157	0/660113	0/77358	0/500869	0/617868	0/656943	1

Table 3. Correlation matrix of gas sensors for the sun-dried treatment.

	Time	**Sensor1**	**Sensor2**	**Sensor3**	**Sensor4**	**Sensor5**	**Sensor6**	**Sensor7**	**Sensor8**
Time	1	0/984689	0/936501	0/891912	0/90792	0/97289	0/793613	0/758693	0/892275
Sensor1	0/984689	1	0/965980	0/924816	0/949282	0/981974	0/843438	0/797339	0/933601
Sensor2	0/936501	0/965980	1	0/978778	0/9450636	0/941478	0/912593	0/826619	0/941955
Sensor3	0/891912	0/924816	0/978778	1	0/946352	0/901890	0/923081	0/818141	0/934693
Sensor4	0/90792	0/949282	0/945063	0/946352	1	0/940539	0/887764	0/837009	0/960576
Sensor5	0/97289	0/981974	0/941478	0/901890	0/940539	1	0/846992	0/822205	0/935099
Sensor6	0/793613	0/843438	0/912593	0/923081	0/887764	0/846992	1	0/909156	0/910115
Sensor7	0/758693	0/797339	0/826619	0/818141	0/837009	0/822205	0/909156	1	0/876470
Sensor8	0/892275	0/933601	0/941955	0/934693	0/960576	0/935099	0/910115	0/876470	1

Machine Learning Algorithms

Figure 7 compares ordinary regression and GBR. GBR provided much more accurate predictions. Table 4 shows GBR metrics: R^2 = 0.9972 (test), RMSE = 0.0031, MAE = 0.0026, RRMSE = 0.0209. Figure 8 shows that GBR outperformed ordinary regression for every sensor.

As can be seen from the figures, the Gradient Boost Regression model provided more accurate and better predictions for all sensors. Therefore, this model was further analyzed using model quality metrics.

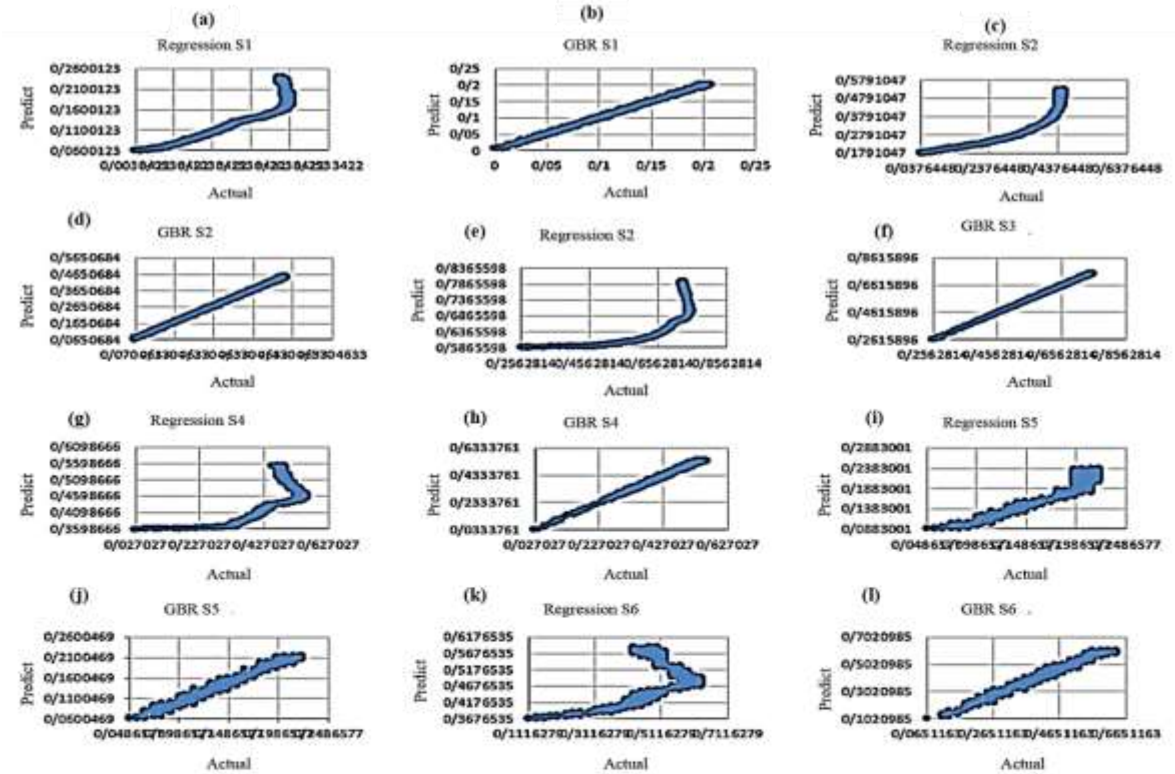


Figure 7. Comparison of actual vs. predicted sensor values for ordinary regression and gradient boost regression models (test data).

Table 4 presents the prediction quality metrics of the best Gradient Boost Regression model for the sensor data, split into training and test sets. The "Training" column shows the model prediction quality metrics for the training data, and the "Test" column shows those for the test data. The coefficient of determination (R^2) for the training and test data was 0.9976 and 0.9972, respectively, indicating that the model fits the data very well. The root mean square error (RMSE) for the training and test data was 0.0028 and 0.0031, respectively. This metric shows that the model's average prediction error for the test data is slightly higher than for the training data. The mean absolute error (MAE) for the training and test data was 0.0024 and 0.0026, respectively. The relative root mean square error (RRMSE) for the training and test data was 0.0187 and 0.0209, respectively. Based on these metric values, the model appears to fit the data well and has excellent predictive quality for the target variable on both training and test data. Overall, considering the high values of the prediction quality metrics, it can be concluded that the designed Gradient Boost Regression model fits the dataset well and can effectively predict the target variable.

Table 4. Prediction quality metrics of the best Gradient Boost Regression model.

	Train	Test
rsquare	0.99759	0.997244
RMSE	0.002817	0.003098
RRMSE	0.018669	0.020985
MAE	0.002401	0.002647

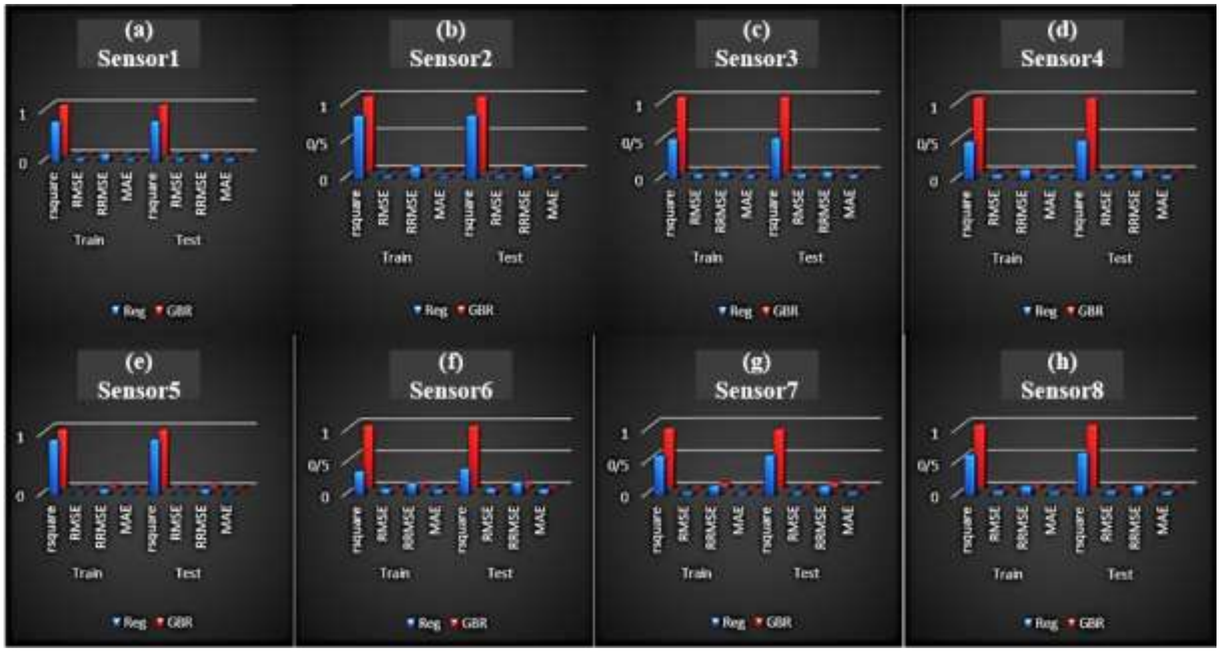


Figure 8. Comparison of prediction quality (R^2 and RMSE) for ordinary regression and gradient boost regression models across all eight sensors.

Table 5 shows ANOVA results. Treatment ($F = 8.88$, $p = 2.9\times10^{-5}$) and sensor ($F = 7.16$, $p = 6.6\times10^{-7}$) were significant; replicate was not ($p = 0.57$). Table 6 lists the GBR hyperparameters.

The values in Table 5 represent various ANOVA statistics that help evaluate model performance. In the presented ANOVA, the sum of squares for the intercept is 70.60344, indicating the amount of variation in the dependent variable that cannot be explained by the intercept. The F-statistic for each independent variable measures the ratio of between-group variance to within-group variance. An F-value greater than 1 indicates a significant difference. These parameters were determined experimentally; depending on the data and the problem at hand, different parameter choices might be better. For example, increasing the maximum depth may improve algorithm accuracy but could also cause overfitting. Likewise, the appropriate value for each parameter should be chosen based on the characteristics of the data and the problem.

Table 5. Analysis of variance (ANOVA) for the performance of the Gradient Boost Regression model.

	sum_sq	**df**	**F**	**PR(>F)**
Intercept	70.60344	1	3949.466	5.35E-81
C(Treatment, Sum)	0.476324	3	8.881653	2.9E-05
C(Sensore, Sum)	0.895736	7	7.158049	6.61E-07
C(Replication, Sum)	0.035953	3	0.670383	0.572222
Residual	1.751917	98		

Table 6. Specifications of the Gradient Boost Regression machine learning algorithm.

	Sensor 1	**Sensor 2**	**Sensor 3**	**Sensor 4**	**Sensor 5**	**Sensor 6**	**Sensor 7**	**Senso r8**
Learning_Rate	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Max_Depth	4	4	4	4	4	4	4	4
N_Estimators	500	500	500	500	500	500	500	500
Subsample	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

Significant Difference Tests

Figure 9 (panels a, b, c) shows that the acid-dipped treatment had the highest R^2, significantly different from the sulfur-fumigated and sun-dried treatments ($p < 0.05$). No significant differences were observed for RMSE, MAE, or RRMSE.

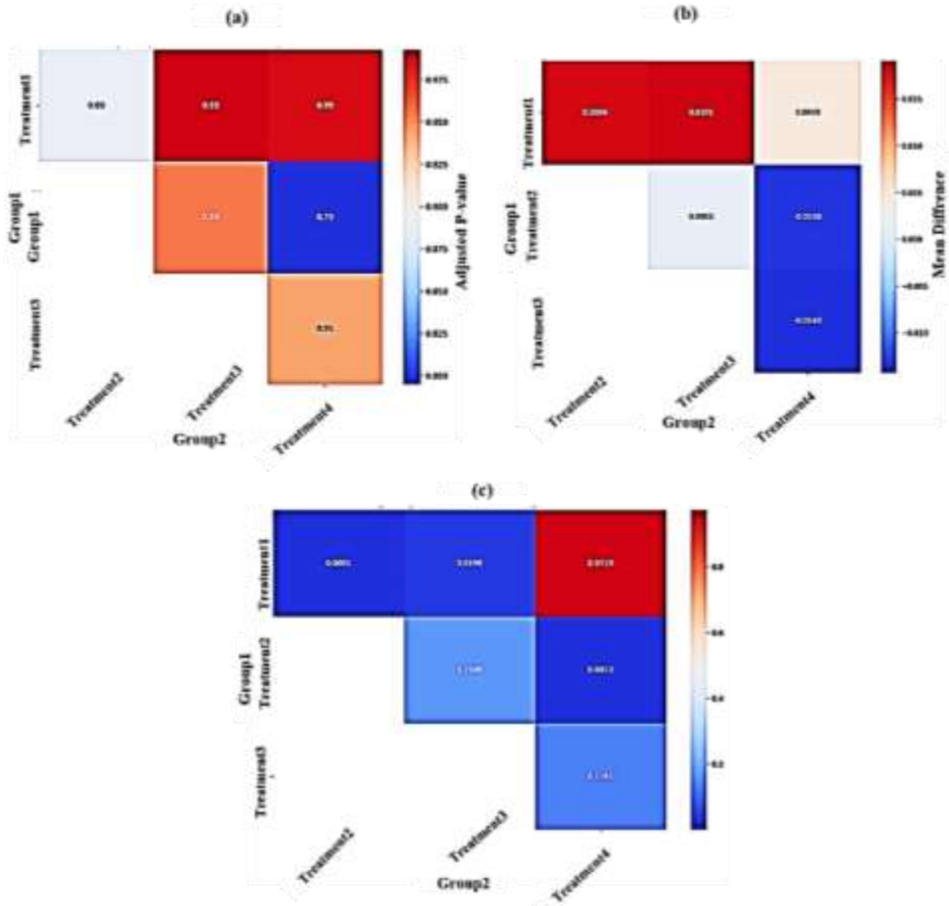


Figure 9. Results of significant difference tests between treatment means: (a) RMSE, (b) RRMSE, (c) R^2.

Figure 9 presents the results of the test for significant differences between treatment means. Treatments 1 to 4 correspond to acid-dipped, factory-processed, sulfur-fumigated, and sun-dried treatments, respectively. The results show that the modelling R^2 for the acid-dipped treatment was significantly different from (and better than) both the factory-processed and the sulfur-fumigated treatments. Also, the factory-processed treatment was significantly different from the sun-dried treatment, with the sun-dried treatment giving better results. This indicates that the acid-dipped treatment had the highest modelling capability and therefore the most predictable sensor responses, which is understandable given the more distinct odor produced by the acid. The factory-processed raisins showed the lowest modelling capability among the treatments, which is reasonable because their exact processing history was unknown compared to the other treatments. However, no significant differences were observed between treatment means for RMSE, MAE, or RRMSE.

Figure 10 (panels a, b, c) shows sensor comparisons. Sensors 1-3 had significantly lower errors than sensors 5-8. Sensors 1, 2, and 3 performed similarly well (no significant differences among them). Sensors 6,7,8 were the weakest.

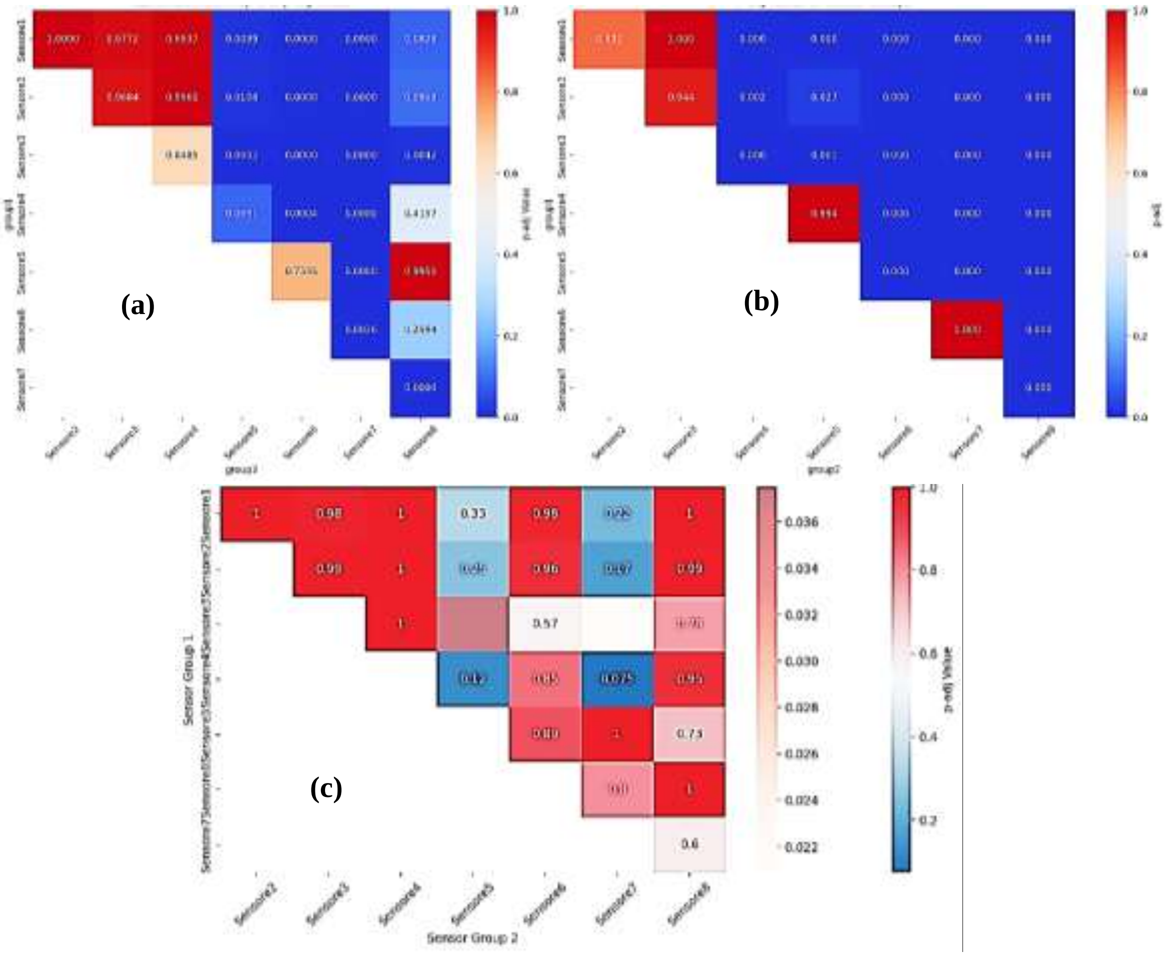


Figure 10. Results of significant difference tests between sensor means: (a) RMSE, (b) RRMSE, (c) R^2.

Figure 10 compares the mean responses of each sensor and examines whether there are significant differences between them. For example, the first comparison shows the mean difference between Sensor 1 and Sensor 2 is 0.0012, with a p-adj value of 0.0911, indicating no significant difference. The CI ranges from –0.0018 to 0.0043, meaning that the difference falls within this interval with 95 % confidence. The reject value indicates acceptance or rejection of the null hypothesis (no significant difference), shown in blue or red. If reject is True, the null hypothesis is rejected, meaning a significant difference exists. If reject is False, the null hypothesis is accepted, meaning no significant difference. In this figure, for comparisons 1-2, 1-3, 2-3, 2-5, 4-5, and 6-7, the null hypothesis was accepted (no significant difference). For the remaining comparisons, the null hypothesis was rejected (significant differences exist).

Figure 11 (panels a, b, c) confirms that sensors 1-3 consistently outperformed sensors 6-8 across all treatments.

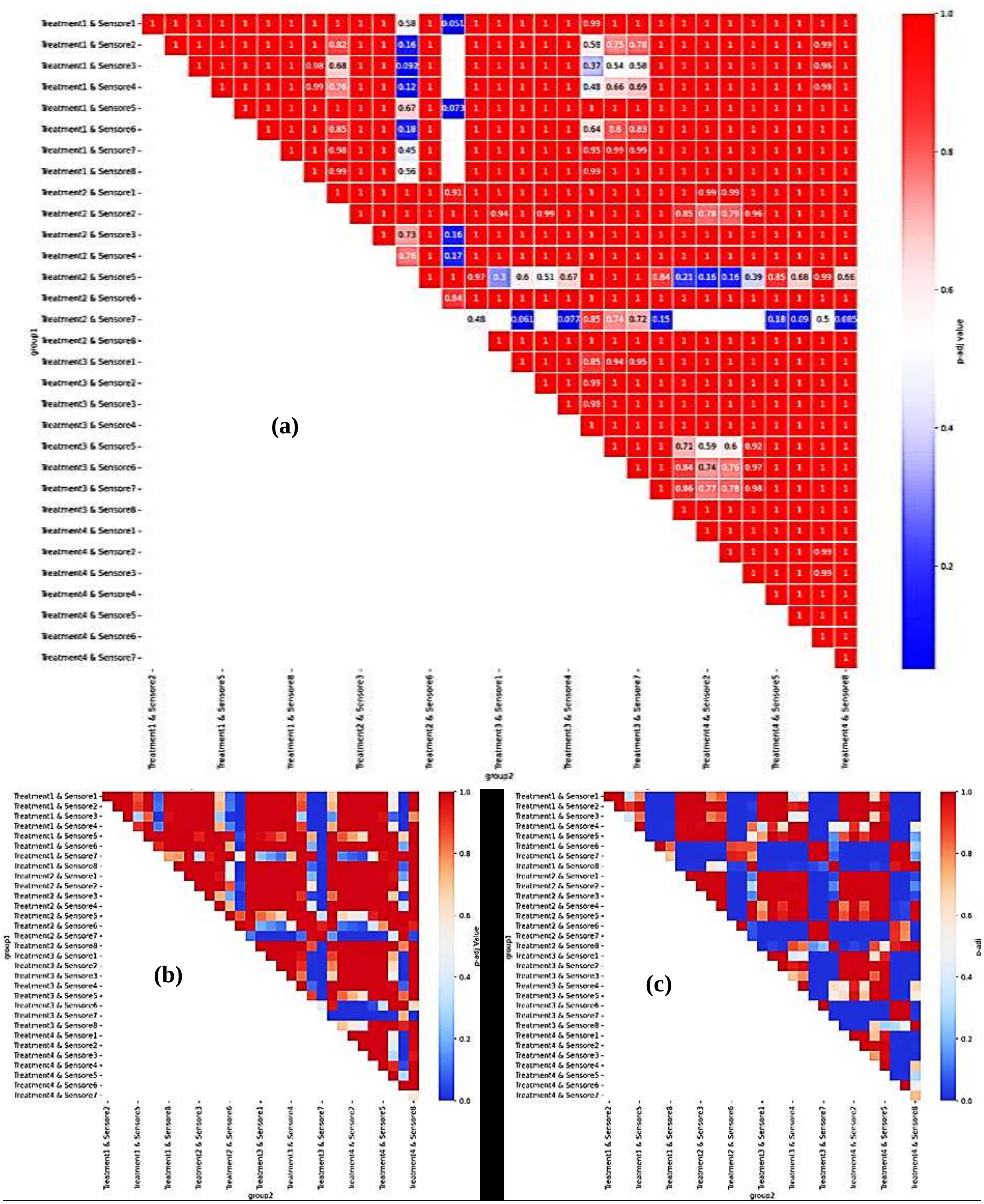
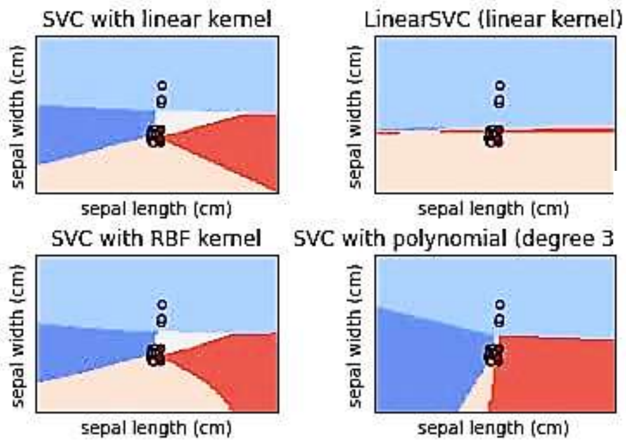


Fig ure 11. Results of significant difference tests between sensor means across different treatments: (a) RMSE, (b) RRMSE, (c) R^2.

PCA score and loading plots identified the most influential sensors. For the sun-dried control, S2 (MQ-3) performed best; for the sulfur-fumigated-then-acid-dipped sample, S2 then S3; for the acid-dipped sample, S2 and S1 then S5. Sensors S1, S2, and S3 (MQ-9, MQ-3, MQ-5) were the best overall.

Data Fusion: Combining E-Nose and Computer Vision

The images presented in Figure 12 show the results of applying the Support Vector Machine (SVM) algorithm with different kernels to the Iris dataset. This dataset is a classic machine learning dataset used for classification.



(a) (b) (c) (d)

Figure 12. Results of applying the SVM algorithm with different kernels to the fused data.

Each of the four plots shown in the figure represents an SVM model with a different kernel:

- SVC with linear kernel: SVM model with a linear kernel.
- LinearSVC (linear kernel): A different implementation of SVM with a linear kernel.
- SVC with RBF kernel: SVM model with a radial basis function (RBF) kernel.
- SVC with polynomial (degree 3) kernel: SVM model with a third-degree polynomial kernel.

In these plots, the horizontal and vertical axes represent the sepal length and sepal width, respectively. The boundary lines separating the data represent the decision boundaries of the SVM model. The region where each class lies is the area where the model predicts that new samples belong to that class.

Kernels: Different kernels determine how the data is mapped into a higher-dimensional space. The linear kernel is suitable for linearly separable data, while RBF and polynomial kernels are more appropriate for non-linear data.

Decision boundaries: The decision boundaries in each plot show how the model classifies new data. The more complex these boundaries, the more capable the model is at modelling complex relationships between features.

Overfitting and underfitting: If the decision boundaries are too complex, the model may overfit the training data and perform poorly on test data. Conversely, if the boundaries are too simple, the model may not capture the patterns in the data (underfitting).

Model performance: To evaluate the performance of different models, metrics such as accuracy, precision, recall, F1-score, and the confusion matrix can be used. These metrics indicate how well the model can classify new samples.

Comparison of models:

Linear kernel: Linear kernel models do not perform well on this dataset because the Iris data is not completely linearly separable.

RBF kernel: The RBF kernel model appears to perform better than the other models, as its decision boundaries are more complex and better fit the data.

Polynomial kernel: The polynomial kernel model also performs acceptably, but due to its higher complexity, it may overfit the training data.

CONCLUSIONS

Based on the presented images, it can be concluded that the SVM model with the RBF kernel performs better than the other models on this data.

Parameter tuning: The performance of SVM models strongly depends on the choice of parameters such as C (penalty parameter) and gamma (RBF kernel parameter).

Other features, such as petal length and petal width, could also be used to improve model performance.

Feature selection techniques can be used to reduce data dimensionality and improve model performance.

Therefore, as can be seen, data fusion has successfully enhanced the discrimination power between treatments. Combining both methods, with very good accuracy, can classify samples and distinguish acid-dipped and sulfur-fumigated samples from sun-dried (untreated) samples. Using the designed system, it is easy to identify harmful samples from healthy ones. It was also found that for the electronic nose system, using only three sensors – S1, S2, and S3 (MQ9, MQ3, MQ5) – which had the best performance for sample identification and classification, is sufficient. This conclusion leads to a reduction in the final system cost for the customer.

After examining similar trends and Figure 9 (results of significant difference tests between sensor means across different treatments on test data), it was also found that Sensor 7 in the acid-dipped treatment had a significantly higher R^2 than in the factory-processed treatment, which confirms the treatment comparison results.

The GBR model demonstrated excellent predictive capability ($R^2 = 0.9972$). This is attributed to its ability to capture non-linear interactions between sensors, such as ratios of sensor responses, which ordinary linear regression cannot. The high R^2 values for sensors S1-S4 indicate that these sensors have strong and consistent responses to SO_2-related VOCs, while S5-S8 are less reliable and could be omitted in a cost-reduced device.

The correlation matrices revealed that redundancy changes with treatment. In the acid-dipped treatment, S1 and S2 are highly correlated; but in the sulfur-fumigated treatment, their correlation drops to 0.261, meaning they provide complementary information. This suggests that keeping both may still be beneficial when the sample type is unknown.

The hybrid system is non-destructive, so samples can be sold after testing. The 60-min exposure gives excellent predictions, but we found that the first 30 min already yields $R^2 > 0.99$; thus a shortened protocol is feasible. Component cost is under \$100 (excluding the smartphone). The device fits in a 30×20×15 cm box, weighs about 2 kg, and can be powered by a 12 V battery. The Android app displays traffic-light results (green = acceptable, yellow = caution, red = exceed limit), making it usable by non-specialists.

Based on all analyses, the following conclusions were drawn:

1. The GBR model modelled sensor responses very well ($R^2 = 0.9972$, RMSE = 0.0209 on test data), far outperforming ordinary linear regression.
2. Correlation analysis revealed strong relationships between some sensor pairs, and these correlations changed across treatments, indicating complementary information.
3. Sensors 1-4 (MQ-9, MQ-3, MQ-5, TGS-2620) were the most accurate and predictable; sensors 5-8 (MQ-135, TGS-822, TGS-813, MQ-4) were weaker.
4. The acid-dipped treatment was the most predictable; the sun-dried control was less predictable but still well modelled.
5. Both olfactory and vision systems separated the three treatments with high accuracy. Feature-level fusion of the two modalities significantly improved classification accuracy to 97.2 % ($p < 0.01$), demonstrating that the hybrid system outperforms either modality alone.

REFERENCES

Asadi, M., Ghasemnezhad, M., Bakhshipour, A., Olfati, J.-A., & Mirjalili, M. H. (2024). Predicting the quality attributes related to geographical growing regions in red-fleshed kiwifruit by data fusion of electronic nose and computer vision systems. *BMC Plant Biology*, *24*(1), 13. https://doi.org/10.1186/s12870-023-04661-6

Cantín, C. M., Minas, I. S., Goulas, V., Jiménez, M., Manganaris, G. A., Michailides, T. J., & Crisosto, C. H. (2012). Sulfur dioxide fumigation alone or in combination with CO2-enriched atmosphere extends the market life of highbush blueberry fruit. *Postharvest Biology and Technology*, *67*, 84–91. https://doi.org/10.1016/j.postharvbio.2011.12.006

Chen, Z., Zhang, G., & Zhang, F. (2026). Multimodal AI for real-time food safety and quality: from sensors to foundation models, edge deployment, and regulation. *Food science & nutrition*, *14*(2), e71534. https://doi.org/10.1002/fsn3.71534

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gardner, J. W., & Bartlett, P. N. (1994). A brief history of electronic noses. *Sensors and Actuators B: Chemical*, *18*(1-3), 210–211. https://doi.org/10.1016/0925-4005(94)87085-3

Girmatsion, M., Tang, X., Zhang, Q., & Li, P. (2025). Progress in machine learning-supported electronic nose and hyperspectral imaging technologies for food safety assessment: A review. *Food Research International*, *209*, 116285. https://doi.org/10.1016/j.foodres.2025.116285

Guerrero, R. F., & Cantos-Villar, E. (2015). Demonstrating the efficiency of sulphur dioxide replacements in wine: A parameter review. *Trends in Food Science & Technology*, *42*(1), 27–43. https://doi.org/10.1016/j.tifs.2014.11.004

Guo, J., Zhang, K., Adade, S. Y. S. S., Lin, J., Lin, H., & Chen, Q. (2025). Tea grading, blending, and matching based on computer vision and deep learning. *Journal of the Science of Food and Agriculture*, *105*(6), 3239–3251. https://doi.org/10.1002/jsfa.14088

Institute of Standards and Industrial Research of Iran (ISIRI). (2014). *Gheisi (whole dried apricot) specifications and test methods* (National Standard No. 13, 4th rev.). Tehran, Iran.

Institute of Standards and Industrial Research of Iran (ISIRI). (2015). *Dried apricots: Specification and test methods* (National Standard No. 11, 5th rev.). Tehran, Iran.

Institute of Standards and Industrial Research of Iran (ISIRI). (2017). *Dried fruits: Determination of sulfur dioxide* (National Standard No. 569). Tehran, Iran.

Institute of Standards and Industrial Research of Iran (ISIRI). (2018). *Specification and methods of test for fruit snack (fruit paste)* (National Standard No. 3308, 2nd rev.). Tehran, Iran.

Khojastehnazhand, M., & Ramezani, H. (2020). Machine vision system for classification of bulk raisins using texture features. *Journal of Food Engineering*, *271*, 109864. https://doi.org/10.1016/j.jfoodeng.2019.109864

Khulal, U., Zhao, J., Hu, W., & Chen, Q. (2016). Nondestructive quantifying total volatile basic nitrogen (TVB-N) content in chicken using hyperspectral imaging (HSI) technique combined with different data dimension reduction algorithms. *Food chemistry*, *197*, 1191–1199. https://doi.org/0.1016/j.foodchem.2015.11.084

Kiani, S., Minaei, S., & Ghasemi-Varnamkhasti, M. (2017). Integration of computer vision and electronic nose as non-destructive systems for saffron adulteration detection. *Computers and Electronics in Agriculture*, *141*, 46–53. https://doi.org/10.1016/j.compag.2017.06.018

Laganà, P., Avventuroso, E., Romano, G., Gioffré, M. E., Patanè, P., Parisi, S., Moscato, U., & Delia, S. (2017). The Codex Alimentarius and the European legislation on food additives. In *Chemistry and hygiene of food additives* (pp. 23–32). Springer. https://doi.org/10.1007/978-3-319-57042-6_2

Liu, Y., Wang, Q., Xu, Q., Feng, J., Yu, H., & Yin, Y. (2018). Non-destructive detection of Flos Lonicerae treated by sulfur fumigation based on hyperspectral imaging. *Journal of Food Measurement and Characterization*, *12*(4), 2809–2818. https://doi.org/10.1007/s11694-018-9896-z

Mehraban Sangatash, M., Mahmoodi, Z., Poorazarang, H., Vosough, A. S., & Nozary Aval, Y. (2015). Determination of sulfur dioxide residue and some physicochemical properties of raisin varieties produced in Khorasan Razavi

Province. *Journal of Food Hygiene*, *5*(3), 49–58. (In Persian)

Meng, Z., Qin, G., & Zhang, B. (2005). DNA damage in mice treated with sulfur dioxide by inhalation. *Environmental and molecular mutagenesis*, *46*(3), 150–155. https://doi.org/10.1002/em.20142

Mischek, D., & Krapfenbauer-Cermak, C. (2012). Exposure assessment of food preservatives (sulphites, benzoic and sorbic acid) in Austria. *Food Additives & Contaminants: Part A*, *29*(3), 371–382. https://doi.org/10.1080/19440049.2011.643415

Pathmanaban, P., Gnanavel, B., & Anandan, S. S. (2019). Recent application of imaging techniques for fruit quality assessment. *Trends in Food Science & Technology*, *94*, 32–42. https://doi.org/10.1016/j.tifs.2019.10.004

Quan, H., Sun, J., Zhou, T., Chen, T., Niu, Z., & Ma, X. (2024). Research on gas recognition based on gas sensor array and feature analysis. *IEEE Sensors Journal*, *24*(15), 24958–24971. https://doi.org/10.1109/JSEN.2024.3409680

Soubra, L., Sarkis, D., Hilan, C., & Verger, P. (2007). Dietary exposure of children and teenagers to benzoates, sulphites, butylhydroxyanisol (BHA) and butylhydroxytoluen (BHT) in Beirut (Lebanon). *Regulatory toxicology and pharmacology*, *47*(1), 68–77. https://doi.org/10.1016/j.yrtph.2006.07.005

Stražanac, D., Gross-Bošković, A., Hengl, B., Bašić, S., & Sokolić, D. (2019). Use of sulphur dioxide and sulphites (E 220–E 228) in canning of crab meat. *Meso: prvi hrvatski časopis o mesu*, *21*(3.), 269–278. https://doi.org/10.31727/m.21.3.5

Türkyılmaz, M., Tağı, Ş., Dereli, U., & Özkan, M. (2013). Effects of various pressing programs and yields on the antioxidant activity, antimicrobial activity, phenolic content and colour of pomegranate juices. *Food chemistry*, *138*(2-3), 1810–1818. https://doi.org/10.1016/j.foodchem.2012.11.100

Vally, H., Misso, N. L., & Madan, V. (2009). Clinical effects of sulphite additives. *Clinical & Experimental Allergy*, *39*(11), 1643–1651. https://doi.org/10.1111/j.1365-2222.2009.03362.x

Wang, C. N., Li, Y. L., Gong, F. L., Zhang, Y. H., Fang, S. M., & Zhang, H. L. (2020). Advances in doped ZnO nanostructures for gas sensor. *The Chemical Record*, *20*(12), 1553–1567. https://doi.org/10.1002/tcr.202000088

Wang, Q., & Wang, T. (2026). Machine learning-driven sensor fusion for precision agriculture: current applications and future perspectives. *Sensor Review*, *46*(1), 1–26. https://doi.org/10.1108/SR-08-2025-0613

Yang, Y., Shahbeik, H., Shafizadeh, A., Masoudnia, N., Rafiee, S., Zhang, Y., Pan, J., Tabatabaei, M., & Aghbashlo, M. (2022). Biomass microwave pyrolysis characterization by machine learning for sustainable rural biorefineries. *Renewable Energy*, *201*, 70–86. https://doi.org/10.1016/j.renene.2022.11.028