Tian *et al. Intell. Robot.* **2025**, *5*(1), 181-201
DOI: 10.20517/ir.2025.10

Intelligence & Robotics

Research Article

Open Access

Check for updates

Non-destructive prediction of apple SSC/TAC and firmness based on multilayer autoencoder and multilayer perceptron

Xu Tian[1], Lyuwen Huang[1,2], Mengqun Zhai[3], Mengyi Zhang[1], Pengju Hu[1], Mingjun Li[4], Liehong Ren[1]

[1]College of Information Engineering, Northwest A&F University, Xianyang 712100, Shaanxi, China.
[2]Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Xianyang 712100, Shaanxi, China.
[3]College of Mechanical and Electronic Engineering, Northwest A&F University, Xianyang 712100, Shaanxi, China.
[4]College of Horticulture, Northwest A&F University, Xianyang 712100, Shaanxi, China.

Correspondence to: Prof. Mengqun Zhai, College of Mechanical and Electronic Engineering, Northwest A&F University, No. 3, Taicheng Road, Yangling District, Xianyang 712100, Shaanxi, China. E-mail: zhaidreams@nwsuaf.edu.cn

How to cite this article: Tian, X.; Huang, L.; Zhai, M.; Zhang, M.; Hu, P.; Li, M.; Ren, L. Non-destructive prediction of apple SSC/TAC and firmness based on multilayer autoencoder and multilayer perceptron. *Intell. Robot.* **2025**, *5*(1), 181-201. https://dx. doi.org/10.20517/ir.2025.10

Received: 24 Oct 2024 **First Decision:** 11 Dec 2024 **Revised:** 8 Feb 2025 **Accepted:** 14 Feb 2025 **Published:** 27 Feb 2025

Academic Editors: Simon Yang, Mingzhe Liu **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

The physical and biochemical indices of apple fruit serve as crucial phenotypic parameters in genomic cultivation. Among them, the soluble solids content (SSC), titratable acid content (TAC), and firmness are the three most paramount parameters that directly reflect the inner quality of apples. To achieve a more accurate prediction of the internal physicochemical indicators, a novel non-destructive detection approach fused with nonlinear and multi-features using a multilayer autoencoder (MAE) was proposed. For non-destructive detection of internal physicochemical indicators, a dielectric spectrum device was employed to gather the electrical parameters of 300 Fuji genomic sample apples. These measurements were taken at nine distinct frequencies, spanning from 0.158 to 3,980 kHz. For the normal control group for validation, to precisely detect its physical and biochemical parameters, special physicochemical analysis apparatuses were utilized to collect data on firmness, SSCs, and TACs. To predict key genomic parameters such as firmness and SSC/TAC, three classical regression models were implemented and subject to comprehensive analysis. The experimental results reveal that the nonlinear feature variable selection based on MAE and multilayer perceptron (MLP) achieved the best prediction performance. Specifically, the correlation coefficients (R^2) for predicting firmness and SSC/TAC reached up to 0.88 and 0.82, respectively, with root mean square errors (RMSEs) of 0.66 and 2.08. Regarding state-of-the-art dimensionality reduction, MAE can be validated as a nonlinear feature extraction methodology for complex electrical parameters. It demonstrates robust applicability in predicting a diverse array of other genomic parameters.

Keywords: Apple, electrical parameters, multilayer autoencoder, nonlinear feature variable, firmness, SSC/TAC

1. INTRODUCTION

Firmness, along with the ratio of soluble solids content (SSC) to titratable acid content (TAC), denoted as SSC/TAC, constitutes direct and determinable biophysical and biochemical indicators reflective of the internal quality of apple fruit[1,2]. Nevertheless, traditional approaches of gauging firmness and SSC/TAC, which typically involve the employment of a penetration tester and a special sugar-acidity meter, are inherently destructive. Consequently, the fruit subjected to such tests cannot be made available for consumption. In contrast, the dielectric spectroscopy measurement, serving as an effective non-destructive detection[3-6], has been widely applied in the internal quality analysis of fruits, such as firmness[7,8], SSC[9-12] and maturity detection[13-16], and so on. With this dielectric spectroscopy, the electrical parameters could be effectively measured and comprehensively analyzed at different frequency points[17,18]. It should be noted that all the electrical parameters exhibit nonlinear and irregular characteristics. To accurately predict the firmness and SSC/TAC, it was a prevalent and essential practice to construct computational models by in-depth analysis of nonlinear characteristic parameters reflecting the internal qualities of the fruit. Therefore, devising a calculation model for non-destructive detection with higher accuracy has emerged as the focus of research efforts and has received increasing attention.

The variation in apple quality has always been correlated with variations in electrical parameters, and there has been a correlation established between dielectric properties and biophysical and biochemical indices. The dielectric property data can be captured non-destructively for analysis of biological genomic characteristics, as the biological characteristics are related to resonant frequencies of the fruit. Żywica *et al.* found an application of calculating electrical model parameters for evaluating the dilution of apple puree of the Champion variety[19]. Bian *et al.* demonstrated that there was a correlation between the electrical properties and internal qualities of Fuji apples, which reflected changes in the degree of damage and physiological state[20]. At nine specific frequency points ranging from 0.158 to 3,980 kHz, Cai *et al.* constructed the correlation between electrical characteristics and apple quality (taste brittleness, acidity, sweetness, palatability, pulp, aroma), and detected freshness and moldy core disease of Fuji apple[21]. Nevertheless, there exists a deficiency in the availability of electrical parameters for evaluating the firmness and the ratio of SSC/TAC for apple fruit.

Machine learning and deep learning have been applied in the field of agriculture[22]. All methods, characterized by their data-driven nature, are capable of extracting features, patterns, and trends from real-world data. This facilitates tasks such as dimensionality reduction[23], cluster analysis[24], classification prediction[25], and regression prediction[26]. Additionally, these methods fuse electrical spectroscopy parameters in order to determine the physicochemical contents of fruit. Liu *et al.* constructed a prediction model for SSC of persimmon fruit using partial least squares regression (PLSR) and least squares support vector regression (SVR)[27]. Using a generalized regression neural network, SVR and extreme learning machine (ELM) modeling, Guo *et al.* established a prediction model on SSC of three varieties of fresh apples, Fuji, Pink Lady, and Red Rome[28]. The same predictive models were also used to effectively evaluate the SSC and TAC for pears[29]. Reyes *et al.* used multiple linear regression (MLR) to predict SSC and pH of red delicious apples[30]. Bian *et al.* used linear regression to predict quality parameters such as sugar content and TAC of bruised apples[31]. Liu *et al.* used PLSR and ELM to predict the sugar content of melons[12]. Wang *et al.* applied linear regression to predict the TAC of kiwifruit[32]. Mohammed *et al.* utilized artificial neural networks (ANNs) and MLR to predict the sugar content, pH, total soluble solids (TSS), water activity, and moisture content (MC) of date fruit (Tamar)[33]. The electrical properties of fruit measured by dielectric spectra have been processed with ANN because of its diverse and multiple data analysis qualities and performance in data analysis, including fault tolerance, nonlinearity, self-learning, and generalization.

To reduce model complexity and simplify model input, most researchers use feature selection and feature extraction methods to reduce redundant variables[34]. The methodology has also been employed during the processing of electrical parameters. In feature selection, correlation analysis[35] and sparse principal component analysis (PCA)[36] are commonly used to select parameters that are strongly correlated with physicochemical indicators as input for prediction models. PCA was popular for feature extraction. Based on this method, Bian *et al.* selected two principal components to represent the dielectric parameters of bruised apples and established a prediction model of physicochemical indicators[37]. Albelda Aparisi *et al.* used PCA for feature extraction and achieved accurate classification of orange frost damage based on the extracted features[38]. Guo *et al.* used the method to realize the feature extraction of the apple dielectric spectrum[39]. The results show that the model performance based on PCA was the best, and the prediction correlation coefficient (R^2) was generally high. Liu *et al.* processed data and combined different prediction models to forecast the physical and chemical index content of pears[12]. Shang *et al.* employed PCA to extract features from the original variables for identifying apple varieties, and their experimental results demonstrate the potential of the method[40]. Wang *et al.* also used this method to simplify the input of the discriminant model and reduce the complexity of the model to construct the detection model for apple water core and moldy core disease[41]. Although the two methods mentioned above simplify the model, they only select a few variables or the linear characteristics between variables as the input for the prediction model, ignoring the nonlinear data characteristics between the sample variables.

Hinton *et al.* and Vincent *et al.* used neural networks and autoencoder (AE) for nonlinear feature extraction, achieving better results by representing high-dimensional data with low-dimensional data[42,43]. The AE was widely used in different machine learning tasks to extract key features from variables[44-47]. Regarding effective data encoding and decoding performance, AE was regarded as one of the important tools for nonlinear feature extraction in deep models, such as dynamic convolutional neural network (DCNN)[48] and deep auto-encoder models[49]. To achieve high performance in nonlinear feature extraction with deep AE, it was often used to reduce the dimensionality of sample variables and improve regression models.

Multilayer autoencoder (MAE) can learn deep nonlinear features of data by stacking multiple hidden layers. Each layer captures different levels of abstract representations of the data, thus better handling complex nonlinear relationships[50]. Compared to other new feature extraction algorithms, MAE is more suitable for extracting features from high-dimensional, nonlinear electrical parameters. Multilayer perceptron (MLP) uses nonlinear activation functions to introduce nonlinearity. This enables MLP to learn and represent complex nonlinear relationships, making them excel in handling nonlinear data[51]. Compared to some complex black-box models, the internal structure of MLP is relatively transparent, making them easier to understand and interpret[52]. Compared to other regression models, MLP is more suitable for predicting high-dimensional, small-sample, and nonlinear data.

This paper proposes a novel prediction method on non-destructive detection of physiochemical parameters of apple, firmness and SSC/TAC, featuring the fusion of MAEs and regression models. The proposal is illustrated in Figure 1. The main contributions include correlation analysis of nonlinear genomics input features and optimal regression prediction for reducing data dimensionality. To begin with, three regression methods were applied to develop firmness and SSC/TAC prediction models based on full frequency spectrums (FFSs), PCA and MAE. In addition, the prediction performance of apple SSC/TAC and firmness after the fusion of three feature extraction methods and regression models was comparatively evaluated. Last but not least, the applicability of nonlinear feature extraction to the processing of electrical parameters was verified. The remainder of this paper is organized as follows: Section "MATERIALS AND METHODS"

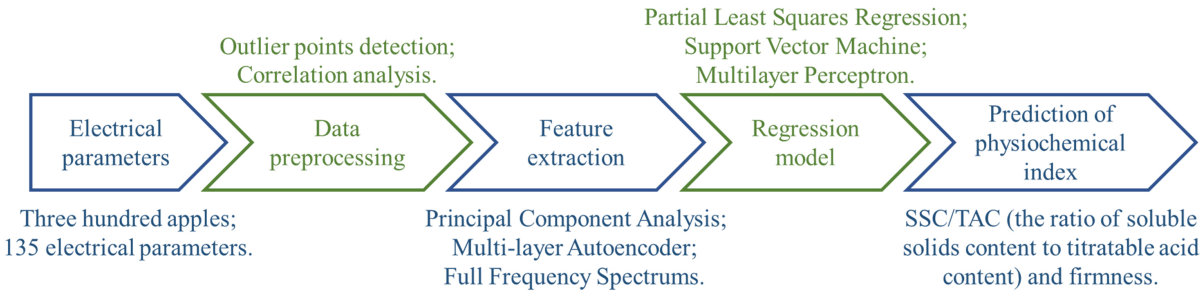

Figure 1. The complete data processing flowchart of proposed approach.

outlines the materials and methods for acquisition and establishment of genomic datasets, preprocessing and redefinition of the deep prediction model. Section "RESULTS" describes the results and provides a discussion of prediction accuracy. The final section summarizes our findings.

2. MATERIALS AND METHODS

2.1. Sample collection

All "Fuji" apple fruits were produced from Luochuan County in Shaanxi Province, which is regarded as a renowned prime area for apple-producing and cultivar breeding region on the Loess Plateau of China. Apple samples were collected from six different cultivations, as detailed in Table 1. All these samples were bagging-free apples. To preserve good discrepancies and stable genomics properties among various apple cultivars, 300 samples were randomly gathered from modern orchards. These samples were carefully selected to be free from any surface defects and damage. The samples were stored in an environment where the temperature was regulated at 4 ℃ and relative humidity was maintained at 90% until genomic analysis was conducted. This storage protocol was implemented to ensure the integrity and stability of the samples' genomic material, ensuring the reliability of subsequent research findings.

2.2. Electrical parameters measurement

Before starting the analysis of the genomic properties of the apples, it is necessary to ensure that all samples are placed outside of storage at room temperature for more than twelve hours but within twenty-four hours. This will help to maintain the temperature balance between the apples and the room, preventing other biochemical reactions from occurring. The electrical characteristics of each apple were then measured in a non-destructive manner using a GW Instek LCR-8205 high-frequency LCR meter [inductance (L), capacitance (C), and resistance (R); GW Instek Corp., Taiwan, China]. A total of 135 electrical parameters of each apple are presented in Table 2, including 15 electrical parameters, such as impedance $\vec{Z}$, phase angle θ, quality factor Q, and others, measured at nine frequencies spanning the range of 0.158 to 3,980 kHz.

As illustrated in Figure 2, the measurement process was conducted as follows. The apple designated for measurement was positioned with its maximum cross-section placed between two parallel electrode plates in a manual shielding box. This configuration ensured attainment of the maximum effective contact area between the apple and the two parallel electrode plates, allowing the apple to function as a dielectric medium for capacitance generation between the electrode plates. The LCR-8205 LCR meter was utilized to measure each electrical parameter at a specific frequency for every individual apple, with each measurement repeated three times. Concurrently, all the measured data were transmitted in real-time to a computer that was connected to the LCR-8205 meter via USB cables. For each apple, the final measurement value of every electrical characteristic parameter was derived by calculating the mean of three electrical parameters corresponding to a single frequency. Before conducting each dielectric measurement, the LCR-8205 meter

Table 1. Apple specimens under six different cultivations

Subgroup index	Cultivations	Samples
1	Bagging-free apples without spraying ecological film agent	93
2	Bagged apples with spraying ecological agent	57
3	Bagging-free apples with spraying ecological film agent	48
4	Bagging-free apples with spraying oligosaccharins	22
5	Apples planting with potash fertilizer	60
6	Apples with inter-row gramineous grasses in orchard	20
In total		300

Table 2. Description of electrical parameters

Electrical parameter	Meaning	Unit
$\vec{Z}$	Impedance	Ohm (Ω)
DEG	Phase angle	Degree (°)
Ls	Series inductance	Henry (*H*)
Lp	Parallel inductance	
Cs	Series capacitance	Farad (*F*)
Cp	Parallel capacitance	
Rs	Series resistance	Ohm (Ω)
Rp	Parallel resistance	
R	Resistance	
X	Reactance	
Y	Admittance	Siemens (*S*)
G	Conductance	
B	Susceptance	
Q	Quality factor	-
D	Dissipation factor	-

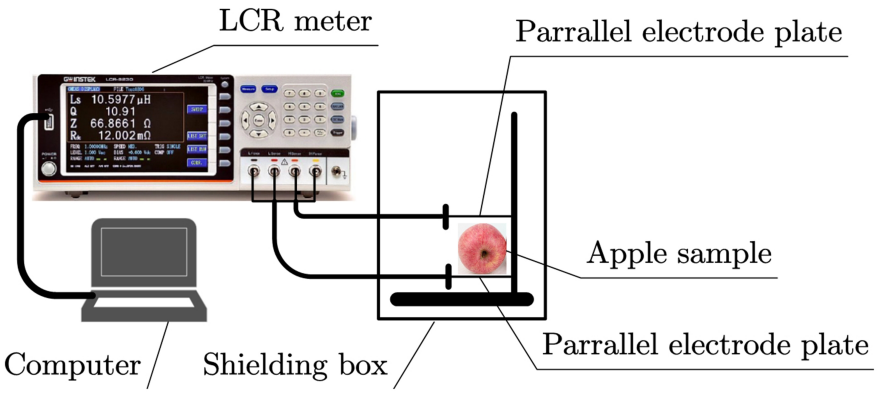

Figure 2. Electrical parameter measurement principle.

must undergo calibration following the procedures outlined in the user guide for opening and shorting circuit operations to ensure accuracy. However, given that the diameter of the samples ranged from 65 to 70 millimeters and their shapes and sizes were almost uniform, it can be postulated that neither the shape nor the size of the samples influence the outcomes of the experimental electrode potential during the measurement of the electrical parameters.

2.3. SSC/TAC and firmness measurement

During the process of collecting electrical characteristics of apples, it is important to promptly detect and analyze crucial physical and biochemical parameters such as SSC, TAC, the ratio of SSC/TAC and firmness. For the analysis of SSC and TAC of apples, a special sugar and acidity meter, ATAGO PAL-BX/ACID, was employed. Meanwhile, apple firmness was measured using food testing texture analysis instruments, specifically the FTC TMS-Pro (manufactured by Food Technology Corp. Ltd., U.S.A.) equipped with a five-millimeter-diameter probe. As with the previous procedures, each apple was analyzed three times, and the final value was determined by averaging the results. It is essential that the non-destructive detection of electrical characteristics be conducted before the analysis of SSC/TAC and firmness, because the physical and biochemical parameter analysis requires that almost the entire apple be destroyed and unusable.

2.4. Prediction model for physicochemical properties

2.4.1. Feature extraction of electrical parameters

Multicollinearity frequently emerges as an issue when using the collected electrical parameters to formulate a prediction model. Implementing effective feature extraction and selection strategies can mitigate the complexity of multiple analyses, enhance model robustness, and improve overall performance. Here, PCA[36,53] and MAE[50] were selected to extract effective feature information from electrical parameters, with the aim of predicting the SSC/TAC ratio and firmness of apples. PCA creates linear combinations of the original variables and can be utilized for both dimensionality reduction and data visualization. By decreasing the number of variables, PCA can effectively compress the data while preserving crucial information. It transforms a dataset into a new set of variables such that these components are chosen to capture the maximum variance in the data. Specifically, the first principal component accounts for the greatest possible variance, and each subsequent component accounts for the largest remaining variance.

MAE is a nonlinear feature extraction approach that consists of an encoder and a decoder. It compresses the input data into a low-dimensional representation (encoding) and subsequently reconstructs it back into the original data (decoding). This structure enables it to proficiently capture the intrinsic structure and nonlinear relationships within the data. MAE can automatically extract meaningful feature combinations and hierarchies from high-dimensional data, reducing dimensionality complexity while retaining essential information. Consequently, MAE proves to be especially well-suited for processing high-dimensional and nonlinear data. The schematic structure of MAE is illustrated in Figure 3.

The encoder was used to derive the hidden layer output h, as given by

$$h = f(WX + b_e) \tag{1}$$

where $f(\ )$ was referred to as the activation function. The rectified linear unit (ReLU) function is chosen; W represented the weight matrix of the encoder; b_e denoted the bias of the input; and X represents the input matrix comprising a total of 135 electrical parameters corresponding to each apple, which were adopted for the MAE analysis.

The decoded value is $\hat{X}$, with the decoding process as

$$\hat{X} = f(\hat{W}\hat{h} + b_d) \tag{2}$$

where $\hat{W}$ was the weight matrix of the decoder, set $\hat{W} = W^T$; b_d was the bias in the output layer; and $\hat{X}$ obtained herein was the reconstruction of the original data.

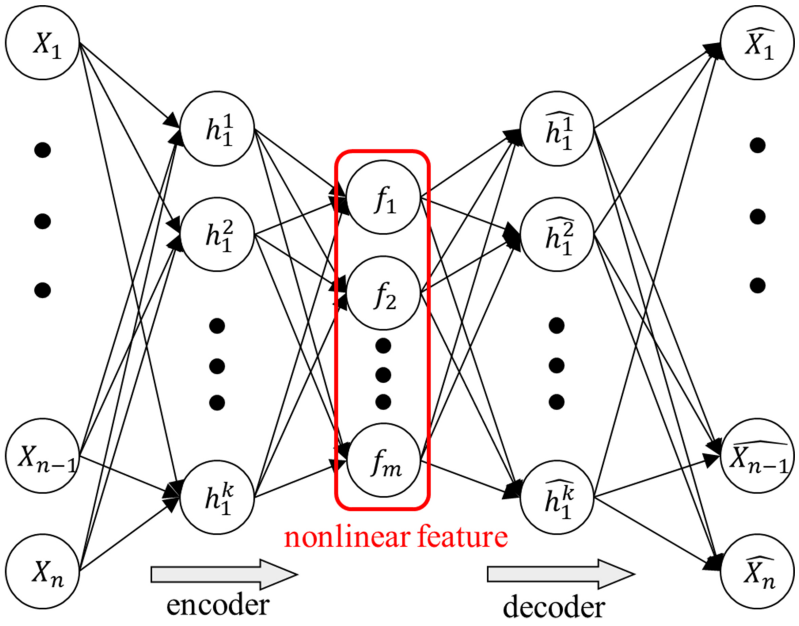

Figure 3. The structure of MAE. MAE: Multilayer autoencoder.

The reconstruction error, *Loss*, is expressed as

$$Loss = \sum_{i}^{s} \left(\widehat{X}_i - X_i\right)^2 \tag{3}$$

where s was the number of samples, X_i was the original input data, and $\widehat{X}_i$ was the reconstruction of the original data.

The determination of the number of hidden layers relied on calculating the exponent of the base of two, where the value fell between the number of input variables and the number of features to be extracted. The number of hidden layers was fine-tuned based on the number of features obtained. In this case, there are 135 input variables and 20 features to be extracted. The encoder was configured with three hidden layers (135 - 128 - 64 - 32 - 20), while the decoder mirrored the encoder's structure symmetrically (20 - 32 - 64 - 128 - 135). The exponents of the base of two, lying within the range defined by the number of input variables and the number of features, were 32, 64, or 128.

2.4.2. Regression model

MLP is a feedforward ANN that constructs intricate nonlinear models by combining multiple elementary artificial neurons. When applied to diverse data tasks, MLP exhibits outstanding performance and adaptability, particularly in scenarios involving small samples, nonlinearity, and high-dimensional data. The multilayered structure of the MLP, in conjunction with the implementation of activation functions, facilitates the network's capacity for profound abstraction and integration of input data. This ability to collect key information from data, enhance predictive accuracy and stability in scenarios with limited data, and learn and simulate complex data relationships leads to accurate predictions and analysis. Moreover, the training process of MLP can proficiently optimize weights via backpropagation algorithms so as to align with the characteristics of high-dimensional data. Algorithm 1 describes the detailed steps of the MAE-MLP model, including the training process of MAE for feature extraction and the training process of MLP for regression prediction.

Algorithm 1 MAE-MLP

Input:

(X_{cal},Y_{cal}): data of calibration set, where X_{cal} is electrical parameters, Y_{cal} is the true values of physicochemical properties, and the number of calibration set is cal;

lr: learning rate;

$fun_{activation}$: activation function for neurons [Rectified Linear Unit (ReLU)];

NUM_{epoch_MAE}: max iterations of MAE;

NUM_{epoch_MLP}: max iterations of MLP;

L_1: L1 regularization parameter;

L_2: L2 regularization parameter;

N_{MAE}: number of hidden layers in MAE;

N_{MLP}: number of hidden layers in MLP.

Output:

Trained MAE-MLP model with optimized weights and biases;

The predicted values of physicochemical properties in calibration set.

Step 1: Train MAE to extract features

1: Initialize weights W_l and biases b_l for each layer from 1 to N_{MAE} randomly.

2: **for** iter ← (1, NUM_{epoch_MAE}) **do**

3: **for** X_i in X_{cal} **do**

5: **for** l ← (1, N_{MAE}) **do**:

6: Calculate the weighted sum of inputs for layer l and apply the activation function to get activations for layer l: $H_l = fun_{activation}(W_l X_i + b_l)$

7: **end for**

8: Compute the reconstruction error: $Loss = \sum_i^{cal} \left(\widehat{X}_i - X_i\right)^2$;

9: Calculate the error gradient: $\nabla Loss$;

10: Update the weights W_l and biases b_l for layer l with L_1 and L_2 regularization:

11: $W_l \leftarrow W_l + lr\nabla Loss$, $b_l \leftarrow b_l + lr\nabla Loss$;

12: **end for**

13: **end for**

14: Save the optimal feature vector of calibration set: F_{cal};

Step 2: Train MLP to predict physicochemical properties

1: Initialize weights W_l and biases b_l for each layer from 1 to N_{MLP} randomly.

2: **for** iter ← (1, NUM_{epoch_MLP}) **do**

3: **for** F_i in F_{cal} **do**

4: **for** l ← (1, N_{MLP}) **do**:

5: Calculate the weighted sum of inputs for layer l and apply the activation function to get activations for layer l: $H_l = fun_{activation}(W_l F_i + b_l)$

6: **end for**

7: Compute the error at the output layer: $Loss = \sum_i^{cal} \left(\widehat{Y}_i - Y_i\right)^2$;

8: Calculate the error gradient: $\nabla Loss$;

9: Update the weights W_l and biases b_l: $W_l \leftarrow W_l + lr\nabla Loss$, $b_l \leftarrow b_l + lr\nabla Loss$;

10: **end for**

11: **end for**

12: save the predicted values of physicochemical properties in the calibration set;

13: save the optimized weights and biases.

Algorithm 1.

To validate the prediction accuracy, several performance evaluation indicators, such as the R^2 and the root mean square error (RMSE) were used. These indicators consisted of the RMSE of the calibration set (RMSEC), the R^2 of the calibration set (R_c^2), the RMSE of the prediction set (RMSEP) and the R^2 of the prediction set (R_p^2). A well-fitted model would typically exhibit an R^2 value approaching 1, while a smaller RMSE would signify stronger predictive capabilities. Additionally, it is of utmost importance to minimize the number of input variables to the greatest extent possible.

3. RESULTS

3.1. Experiment parameters configuration

To establish a prediction model elucidating the relationship between electrical parameters and SSC/TAC as well as firmness, three conventional regression methods were adopted: PLSR, SVR, and MLP. Each of these methods respectively utilized both FFS, which denotes the values of electrical parameters gauged at all frequency points, and the outcomes of feature extraction on electrical parameters as inputs. Concurrently, a comparative analysis was conducted on the prediction results.

The MAE model was trained using ReLU serving as the activation function. During the training process, a learning rate of 0.01 was set, and the model underwent 100 iterations, with a batch size of 5 for each batch. Adam was used as the optimizer to fine-tune this model. To prevent overfitting, L1 and L2 regularization operations were incorporated. To achieve the highest possible accuracy for the regression model, 10-fold cross-validation to fine-tune hyperparameters was employed. The ultimate objective was to identify the optimal hyperparameter combination required for the algorithm to achieve peak performance. In the process of predicting SSC/TAC, SVR was employed, utilizing the radial basis function (RBF) kernel to project the original data into a higher-dimensional space. The parameters were calibrated as follows: γ was set to 0.01 and the regularization parameter C was fixed to 5. Meanwhile, the hidden layer nodes in MLP model were configured as 60 and 30, respectively, employing the ReLU as activation function and mean squared error (MSE) as the loss function. In the Adam optimizer, an initial learning rate was set to 0.001 and a maximum iteration limit of 500. For the firmness prediction, the RBF kernel was again employed in SVR. Here, γ was adjusted to 0.001 and the regularization parameter C was held constant at 10. For the MLP model, its hidden layer nodes were configured as 40 and 29, respectively. All remaining parameters were consistent with the settings used in the previously described SSC/TAC prediction task.

3.2. Analysis of physiochemistry properties and datasets

The phenotypic characteristics of apple varieties, such as surface color, size, single fruit weight, shape index and maturity, exhibit substantial variability depending on the cultivation method. Table 3 shows the specific physicochemical properties of all samples in each genomics group, with noticeable disparities in the standard deviations of SSC/TAC. Noticeable fluctuations were observed in the minimum, maximum, mean values, standard deviation of SSC/TAC, and firmness. As shown in Table 3 for SSC/TAC of apples, the variation range between the minimum and maximum values was extensive. The standard deviations manifested a random, discrete, and irregular pattern, with mean values exceeding the minimum by a factor of two or more. For instance, Subgroup one had maximum and minimum values of 89.55 and 8.27 for SSC/TAC, respectively. Subgroup five exhibited a maximum standard deviation of 19.77, whereas Subgroup two had a minimum standard deviation of 5.87. Regarding the firmness of the apples, significant differences were also observed across the different subgroups. Subgroup one recorded a maximum value of 13.42 and a minimum value of 6.63. The standard deviation varied considerably as well, with Subgroup one exhibiting a maximum standard deviation of 1.22 and Subgroup four showing a minimum standard deviation of 0.73. These results indicate the necessity for a nonlinear analysis method to accommodate the extensive diversity in SSC/TAC and firmness across different cultivations, as the existing linear approaches might not adequately capture such complex variations.

Table 3. SSC in subgroup samples

Physiochemistry indicator	Subgroup index	Minimum	Maximum	Mean value	Standard deviation
SSC/TAC, %	1	8.27	89.55	26.02	11.96
	2	9.39	35.15	18.28	5.87
	3	17.11	76.22	33.47	12.59
	4	10.96	57.64	31.25	12.81
	5	13.56	82.60	33.50	19.77
	6	12.23	58.80	21.58	10.15
Firmness, N	1	6.63	13.42	10.08	1.22
	2	7.98	12.14	10.34	0.93
	3	8.74	12.93	10.09	0.86
	4	9.45	12.02	10.74	0.73
	5	6.83	11.49	8.33	0.79
	6	7.72	11.19	8.64	0.84

SSC: Soluble solids content; TAC: titratable acid content.

Electrical parameter measurement represents a non-destructive method for scrutinizing the effect of electrolytes on apples. Nevertheless, given the subtlety of this influence, it proves arduous to directly ascertain various physicochemical properties within the apple, such as SSC, TAC, and firmness. Additionally, electrical parameters gauged at a solitary frequency point fail to exhibit a direct and robust correlation when predicting the content of physicochemical properties in apples. The frequency range of electrical parameters was determined to be between 158 Hz and 3.98 MHz[54]. Table 2 showcases various electrical parameters, each endowed with a distinct meaning and corresponding dimensional unit. Despite the seemingly homogeneous initial distribution, there are significant dimensional variations among the parameters. In experiments, the magnitude of the data exerts a substantial influence on model training. Insufficient data volume might lead to certain data subsets being overlooked, thereby resulting in the omission of important information. Therefore, it is essential to subject all sampled data to a preprocessing method that ensures the retention of the original data information within the samples.

To accurately predict apple physicochemical properties (SSC/TAC and firmness) using electrical parameters, 300 apple samples were employed, which were randomly partitioned into two subsets: a calibration set comprising 255 samples and a prediction set consisting of 45 samples. The calibration set was dedicated to training the prediction model, whereas the prediction set was used to validate the accuracy of the predictions. Table 4 presents the random sample set and the range of values for the physicochemical indices. The calibration samples cover the prediction set well. The range of SSC/TAC values spanned from 8.27 to 89.55, and the range of firmness values extended from 6.63 to 13.42.

3.3. Preprocessing of electrical parameters and correlation analysis

To improve prediction performance regarding SSC/TAC and firmness, it is necessary to first remove outliers from each group of electrical parameters. Outliers, singular points, or redundant variables cannot fully represent the genuine variables among the electrical parameters. To ascertain the singular value of each electrical parameter at nine distinct frequencies, a statistical representation in the form of a box-whisker plot was used. This method calculates the first and third quartiles and identifies any outlier points, which are then replaced with the median. To eliminate the influence of different dimensions of these parameters on each apple sample, standardization procedures were applied.

Table 4. The results of dataset partitioning

Physiochemical indicator	Sample set	Sample sizes	Range	Mean ± Std (%)
SSC/TAC, %	Calibration set	255	8.27~89.55	27.74 ± 14.52
	Prediction set	45	12.23~68.77	24.95 ± 12.66
Firmness, N	Calibration set	255	6.63~13.42	9.95 ± 1.20
	Prediction set	45	7.10~11.49	8.51 ± 0.96

Std: Standard deviation; SSC: soluble solids content; TAC: titratable acid content.

To ensure the reliability of the electrical parameters of the samples, we employed box-whisker plots to remove any outliers, followed by normalization. After removing any abnormal values, the Kolmogorov-Smirnov test (K-S test)[55] was employed to determine the normality between the electrical parameters and physiochemistry indicators. If the *P*-value exceeded the significance level (*P*-value ≥ 0.05), it indicated that the electrical parameters generally followed a normal distribution. As shown in Figure 4, only 57 electrical parameters satisfied normal distribution. As a result, the data did not meet the prerequisites for Pearson's bivariate correlation analysis. Therefore, only Spearman correlation analysis[56] can be used to determine the correlation between electrical parameters and SSC/TAC as well as firmness. Figure 5 shows that there was no strong linear correlation between the electrical parameters and either SSC/TAC or firmness, with maximum correlation values reaching merely 0.34 and 0.33, respectively.

Through the comprehensive analysis of physicochemical properties and datasets, this paper pinpoints several inherent challenges, such as small sample size, high parameter dimensions, and pronounced nonlinearity. Therefore, given these circumstances, it becomes imperative to employ dimensionality reduction that is capable of preserving the original data information while also conducting nonlinear feature extraction. Additionally, the construction of nonlinear regression models is essential, as these models can yield favorable predictive outcomes even when dealing with small sample data. In light of this, MAE for feature extraction and MLP for regression prediction were adopted and analyzed.

3.4. Prediction of SSC/TAC and firmness

3.4.1. Feature of electrical parameters extracted by PCA and MAE

The optimally extracted features served as inputs for regression methods to fine-tune the parameters and optimize the prediction of physicochemical properties. A comparison was drawn between two feature extraction algorithms: the linear PCA and nonlinear MAE. Here, 40 main features were selected using PCA, attaining a cumulative contribution rate of 98.9%.

The RMSEC for the regression methods was calculated using MAE under diverse numbers of characteristic variables to determine the number of inputs for the prediction model. To avoid underfitting or overfitting problems in modeling, a full cross-validation strategy was applied to select the appropriate number of features. Figure 6 displayed the changes in RMSEC as the number of features increases in the prediction of SSC/TAC and firmness. MAE was combined with different regression algorithms for each prediction. Figure 6 showed that the number of features required for prediction varied, and that the RMSEC also varied across different fusion models. Notably, there was a significant difference between the predictions of the SSC/TAC.

Figure 6A and B presented the variations in prediction results using the MLP as the number of features increased. The RMSE decreased as the number of features increased, with a significant drop in prediction error for SSC/TAC and firmness when the number of features reached 40. However, if the number of features exceeded 40, the prediction error stabilized. Therefore, the final inputs for the prediction model

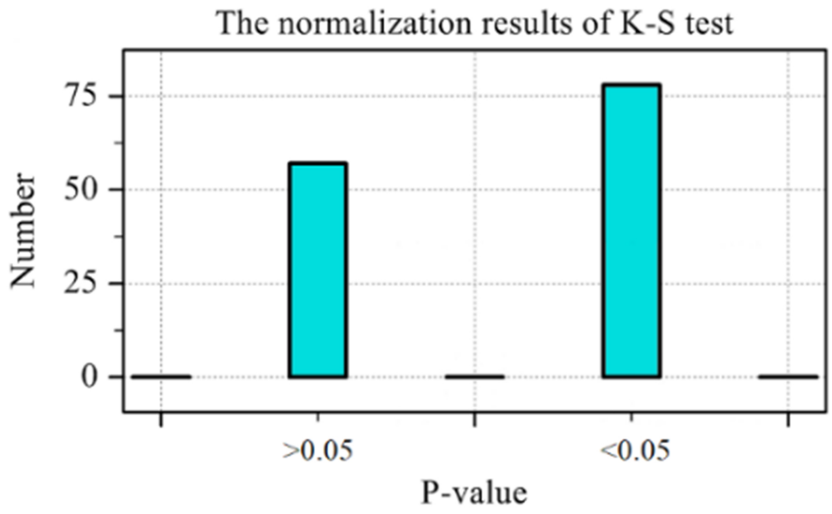

Figure 4. The normalization results of K-S test. K-S test: Kolmogorov-Smirnov test.

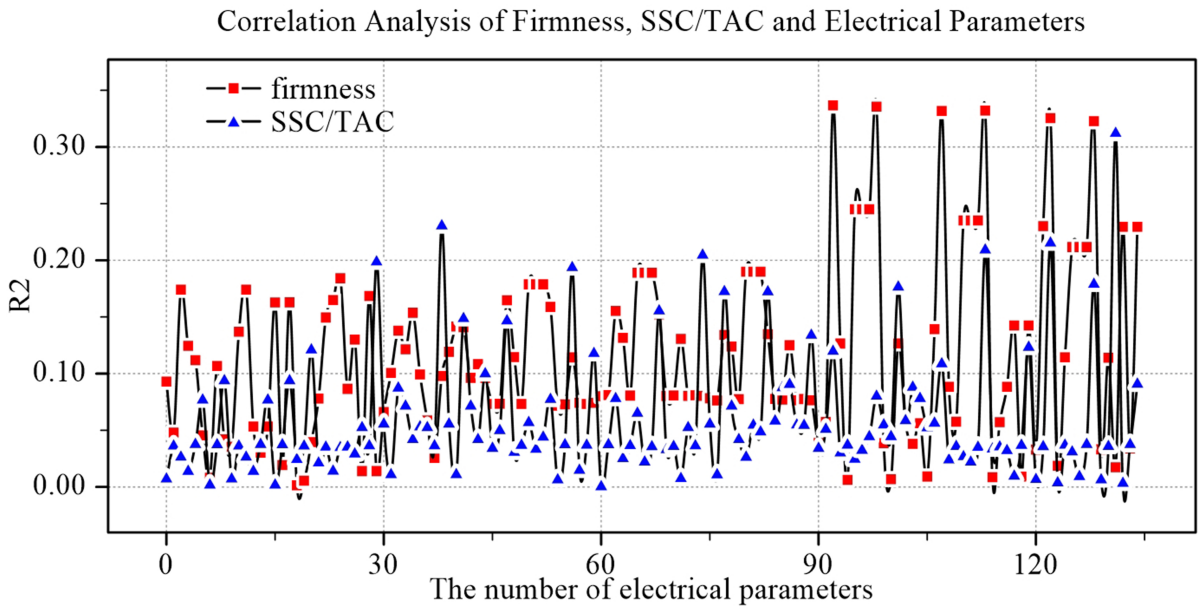

Figure 5. The results of correlation analysis.

were selected with fewer features and lower RMSE. Finally, the number of features for SSC/TAC prediction was determined to be 46, and for firmness prediction, it was determined to be 67.

Figure 6C and D illustrated the correlation between prediction error and the number of features using SVR. The prediction error of SSC/TAC initially decreased and then increased, reaching its minimum when the number of features ranged from 50 to 60. If the number was greater than 65, the prediction error of firmness was at a lower level. The minimum error recorded was approximately 2.46. The SSC/TAC and firmness prediction models had 57 and 68 input features, respectively.

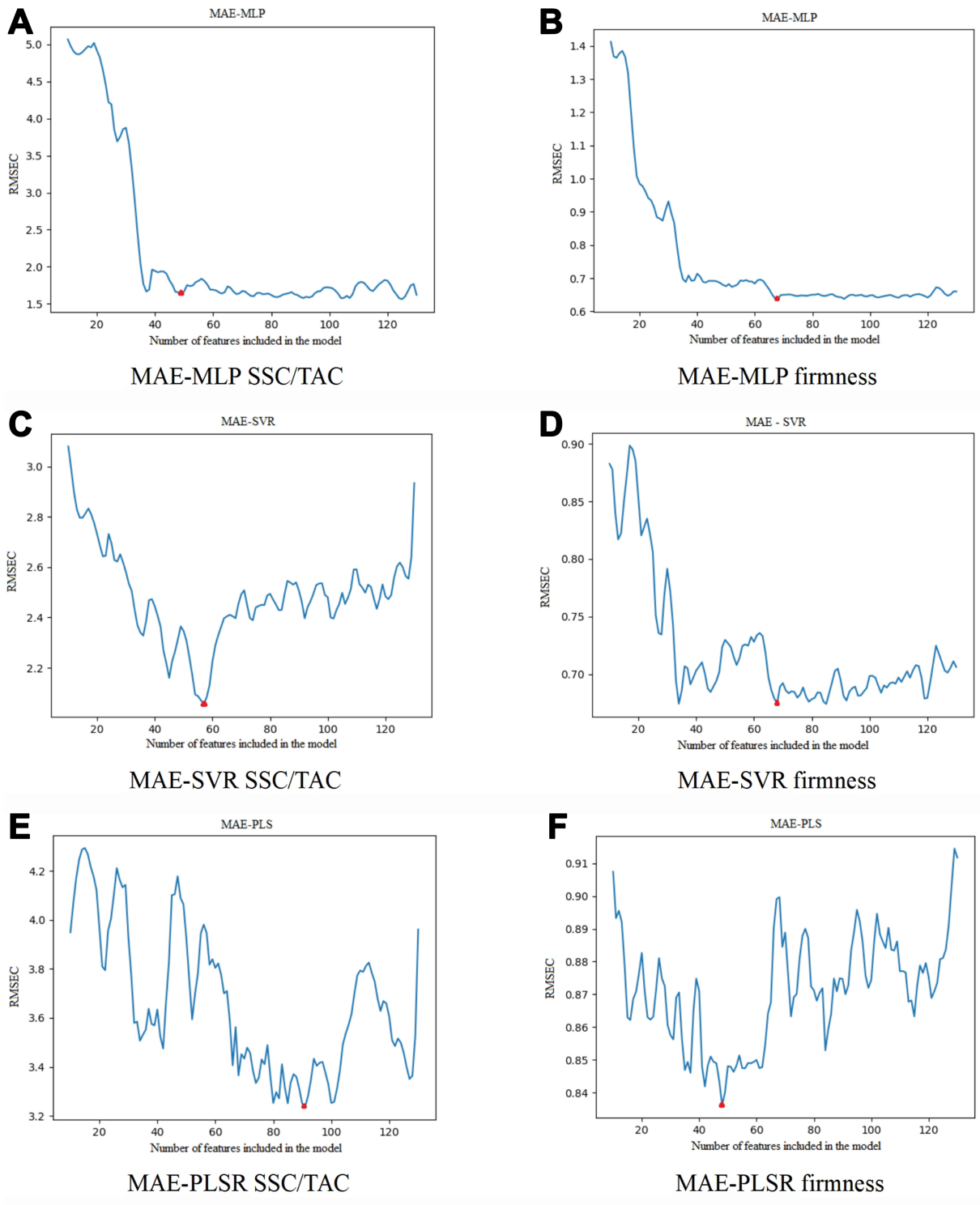

Figure 6. RMSE variation curve of MAE combined with different regression models. The extracted numbers of features were indicated by the red triangle in the figure. RMSE: Root mean square error; MAE: multilayer autoencoder.

Figure 6E and F showed the influence of increasing the number of features on the prediction error using the PLSR model. The change tendency presented as V-shaped, with the RMSE of both SSC/TAC and firmness

changing more dramatically than the other two models. The range of RMSE variation was relatively small. The study found that the prediction error of SSC/TAC was the smallest; the prediction error for SSC/TAC reached its minimum when the number of features ranged from 80 to 100. Similarly, for firmness, the prediction error was the lowest when the feature ranged between 40 and 60. Therefore, the model necessitated 91 input features for optimal prediction of SSC/TAC and 49 input features for firmness.

The final number of input features for different regression models was shown in Table 5. The number of features varied depending on the regression model used alongside the MAE. Nevertheless, the MAE enables the selection of more features compared to PCA. In the case of predicting the SSC/TAC of apples, the MLP based on MAE (MAE-MLP) selected the fewest features but still had six more eigenvalues than the PCA. MAE-PLSR selected the most features, with a total of 45 extractions, surpassing the number chosen by MAE-MLP. MAE-SVR selected eleven features, which was more than MAE-MLP. Meanwhile, regarding the firmness of the apple, the feature numbers selected by MAE-MLP and MAE-SVR were relatively close, with a difference of merely one. Significantly, both models selected significantly more features than PCA. However, the MAE-PLSR model selected the fewest features, with 18 and 19 fewer than MAE-MLP and MAE-SVR, respectively.

3.4.2. Prediction results

3.4.2.1. PLSR model

Based on dialectical FFS, PCA, and MAE, the PLSR prediction models of SSC/TAC and firmness were constructed. The results are shown in Table 6. Generally, in terms of predicting SSC/TAC and firmness, the feature selection of PLSR models based on PCA and MAE was comparable to that of the models based on FFS. Among them, the PLSR model based on MAE performed the best, followed by PCA, and the FFS-based was the worst. For the prediction of SSC/TAC, R^2 values were 0.57, 0.60, and 0.69 in sequence, and the RMSE of predictions was 5.33, 4.63, and 4.05, respectively. For firmness prediction, the PCA-PLSR and MAE-PLSR models had similar prediction performance with no significant difference between them. More specifically, during the feature extraction process, MAE managed to extract 49 data features, whereas PCA only extracted 40. Thus, the PCA-PLSR model is simpler and more effective than the MAE-PLSR model. For the prediction of SSC/TAC, the prediction performance of FFS-PLSR was the poorest (R_p^2 = 0.57, RMSEP = 5.33). In contrast, the performance of PCA-PLSR and MAE-PLSR was essentially the same, with a discrepancy of only 0.07 and 0.09 in R_c^2 and R_p^2, respectively. The employment of MAE resulted in a significantly higher number of features compared to PCA-PLSR, and the RMSE of MAE-PLSR was reduced by 0.69 and 0.58 on the calibration and prediction sets, respectively. Although both MAE-PLSR and PCA-PLSR models were used to predict apple's SSC/TAC and firmness, the former showed slightly better performance than the latter. Thus, the outcomes imply that the application of MAE did not significantly improve the prediction performance when employing the PLSR.

3.4.2.2. SVR model

The prediction results of the SVR model were shown in Table 7. Its predictive performance was similar to that of the PLSR prediction model. The SVR model based on MAE demonstrated the best performance, followed by the PCA-SVR model, while the FFS-SVR model performed the worst. The data from Table 7 indicated that the prediction performance for firmness was better than that for SSC/TAC. Regarding firmness, the prediction performance of PCA-SVR and MAE-SVR did not change much. Compared to FFS-SVR in the calibration set, the R^2 of PCA-SVR and MAE-SVR was increased by 0.03 and 0.04, respectively, and RMSEC was decreased by 0.19 and 0.32. In the prediction set, the R^2 increased by 0.07 and 0.10, respectively, and RMSEP was decreased by 0.25 and 0.28. For the prediction of SSC/TAC, the prediction performance of the MAE-SVR model was significantly improved compared to the FFS-SVR model. The

Table 5. The results of feature quantity selection

Model	The number of features	
	SSC/TAC	Firmness
MAE-MLP	46	67
MAE-SVR	57	68
MAE-PLSR	91	49

SSC: Soluble solids content; TAC: titratable acid content; MAE: multilayer autoencoder; MLP: multilayer perceptron; SVR: support vector regression; PLSR: partial least squares regression.

Table 6. The results of PLSR model prediction

Model		R_c^2	RMSEC	R_p^2	RMSEP
Firmness	FFS	0.64	1.14	0.59	1.29
	PCA	0.67	1.12	0.61	1.20
	MAE	0.71	1.04	0.63	1.13
SSC/TAC	FFS	0.61	4.62	0.57	5.33
	PCA	0.65	3.89	0.60	4.63
	MAE	0.72	3.20	0.69	4.05

PLSR: Partial least squares regression; R_c^2: R^2 of the calibration set; RMSEC: RMSE of the calibration set; R_p^2: R^2 of the prediction set; RMSEP: RMSE of the prediction set; FFS: full frequency spectrum; PCA: principal component analysis; SSC: soluble solids content; TAC: titratable acid content; MAE: multilayer autoencoder.

Table 7. The results of SVR model prediction

Model		R_c^2	RMSEC	R_p^2	RMSEP
Firmness	FFS	0.74	0.99	0.70	1.06
	PCA	0.79	0.78	0.77	0.81
	MAE	0.86	0.67	0.80	0.78
SSC/TAC	FFS	0.72	3.18	0.70	3.53
	PCA	0.76	2.82	0.72	3.19
	MAE	0.83	2.10	0.81	2.23

SVR: Support vector regression; R_c^2: R^2 of the calibration set; RMSEC: RMSE of the calibration set; R_p^2: R^2 of the prediction set; RMSEP: RMSE of the prediction set; FFS: full frequency spectrum; PCA: principal component analysis; MAE: multilayer autoencoder; SSC: soluble solids content; TAC: titratable acid content.

model's prediction error was reduced by 1.08 and the R^2 was increased by 0.11 on the calibration set. Similarly, on the prediction set, the RMSEP of the model was reduced by 1.30 and the R^2 was increased by 0.11. Additionally, compared to PCA-SVR, RMSEP of the model was reduced by 0.96 and the R_p^2 was increased by 0.09. The MAE-SVR model exhibited only a marginally superior calibration performance compared to PCA-SVR when predicting firmness or SSC/TAC. Its prediction performance achieved significant improvement. This indicates that, when employing the SVR model to predict SSC/TAC, the MAE feature extraction method was more efficacious than PCA and FFS.

3.4.2.3. MLP model

Table 8 showed the results of using MLP models to predict SSC/TAC and firmness, which follow a similar trend to the PLSR and SVR models. The prediction results for firmness of apples did not change significantly for the FFS-MLP and the MLP based on PCA (PCA-MLP) models, with the prediction error of PCA-MLP being similar to that of FFS-MLP. In particular, the prediction errors of MAE-MLP on the calibration and prediction sets were diminished significantly. On the calibration set, MAE-MLP had a

Table 8. The results of MLP model prediction

Model		R_c^2	RMSEC	R_p^2	RMSEP
Firmness	FFS	0.80	0.78	0.74	0.89
	PCA	0.83	0.70	0.76	0.83
	MAE	**0.92**	**0.61**	**0.88**	**0.66**
SSC/TAC	FFS	0.78	2.79	0.71	3.23
	PCA	0.82	2.09	0.78	2.77
	MAE	**0.85**	**1.62**	**0.82**	**2.08**

MLP: Multilayer perceptron; R_c^2: R^2 of the calibration set; RMSEC: RMSE of the calibration set; R_p^2: R^2 of the prediction set; RMSEP: RMSE of the prediction set; FFS: full frequency spectrum; PCA: principal component analysis; MAE: multilayer autoencoder; SSC: soluble solids content; TAC: titratable acid content.

higher R_c^2 value of 0.92 and a lower RMSEC value of 0.61 compared to PCA-MLP, which had an R_c^2 value of 0.83 and a RMSEC value of 0.70. Similarly, on the prediction set, MAE-MLP had a higher R_p^2 value of 0.88 and lower RMSEP value of 0.66 than PCA-MLP, which had an R_p^2 value of 0.76 and an RMSEP value of 0.83. When compared with FFS-MLP, the R^2 of MAE-MLP on the prediction set increased by 18.9 %, and the RMSE of prediction decreased by 25.8%. In the case of predicting SSC/TAC of apple, MAE-MLP outperformed PCA-MLP in terms of the calibration performance (R_c^2 = 0.85, RMSEC = 1.62) and the prediction performance (R_p^2 = 0.82, RMSEP = 2.08) compared with MAE. The R_c^2 and R_p^2 values increased by 0.03 and 0.04, respectively, while the RMSE of the calibration and the prediction sets decreased by 0.47 and 0.69. Among three models, the FFS-MLP model had the poorest prediction performance for apple's SSC/TAC.

The final results indicated that the fusion of the nonlinear feature extraction with MAE-based and MLP had better prediction performance in the prediction of apple's SSC/TAC and firmness. Among the aforementioned models, MAE-MLP achieved the best prediction results when predicting both SSC/TAC and firmness, as illustrated in Figures 7 and 8.

4. DISCUSSIONS

In firmness prediction, a comparative analysis of model performance revealed that the FFS prediction model, specifically FFS-MLP, outperformed the FFS-SVR and FFS-PLSR models. Among these models, FFS-PLSR exhibited the poorest prediction performance. Among the models based on PCA, PCA-MLP demonstrated the best performance when compared with PCA-PLSR and PCA-SVR. The MAE was used to extract the characteristics of the sample data and subsequently construct the prediction model. It outperformed other prediction models in terms of performance. MAE-MLP exhibited the highest R^2 in both the calibration and prediction sets, along with lower RMSEC and RMSEP. The predicted R^2 reached up to 0.88, and the predicted RMSE was 0.66.

In predicting SSC/TAC for apples, the FFS-MLP model demonstrated the best performance when utilizing FFS data as input to the regression model. It was followed by the FFS-SVR model, while the FFS-PLSR model delivered the poorest. Among the models based on PCA, the PCA-MLP model outperformed the PCA-SVR and PCA-PLSR models. Additionally, the models based on MAE had similar prediction results.

Upon analyzing the overall prediction results, it was evident that MLP regression method had a pronounced advantage. The primary features obtained through the PCA method showed negligible impact on the models' prediction performance. In contrast, employing the MAE-based prediction model resulted in a substantial reduction in prediction errors. To illustrate this point, consider the MLP model. In the

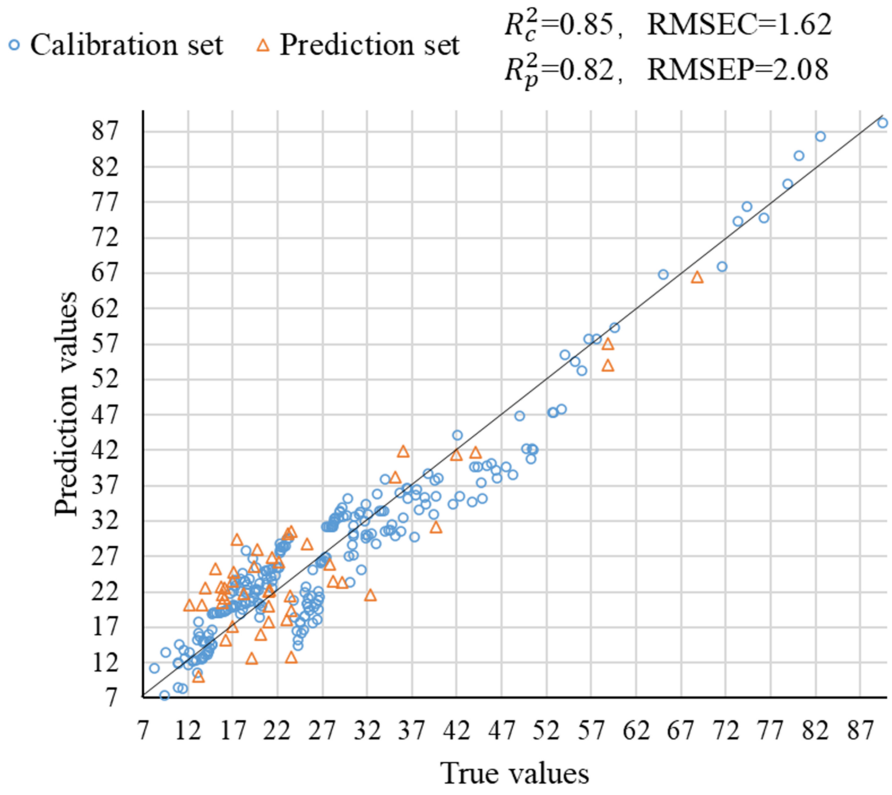

Figure 7. Optimal prediction results of SSC/TAC. SSC: Soluble solids content; TAC: titratable acid content.

calibration set, compared to the PCA-MLP and FFS-MLP models, the MAE-MLP model achieved a reduction in prediction error of 0.47. In the prediction set, the reduction in prediction error was 0.69 and 1.15. The main rationale for this is the fact that the FFS-based models retain all the original information of the electrical spectral data, which increases model complexity and introduces a large amount of redundant information. Meanwhile, the predictive results also reflect that the MAE method can effectively extract the original information of electrical parameters, improve prediction performance, and simplify the model complexity to some extent.

In summary, the electrical spectrum parameters were used to determine the SSC/TAC and firmness content of apple fruit. To develop predictions for physical-chemical indices, PLSR, SVR, and MLP models were employed using FFS, PCA, and MAE. Nonlinear models were found to surpass linear ones in performance. MAE emerged as an effective multivariate data analysis method, capable of extract relevant features between multiple variables and achieving good performance on various nonlinear regression models. It can be effectively used to extract and enhance nonlinear dynamic features of electrical parameters within apple fruit. The experimental results signified that the collection of electrical parameters can serve as an effective means to predict SSC/TAC and firmness. In the extraction of apple electrical parameter features, MAE showed optimal performance.

5. CONCLUSIONS

This paper has proposed a novel MAE and MLP model (MAE-MLP) to predict SSC, TAC and firmness within electrical parameters in a non-destructive way, which are the three most paramount parameters that directly reflect the inner quality of apples. The diverse prediction models of apple SSC/TAC and firmness

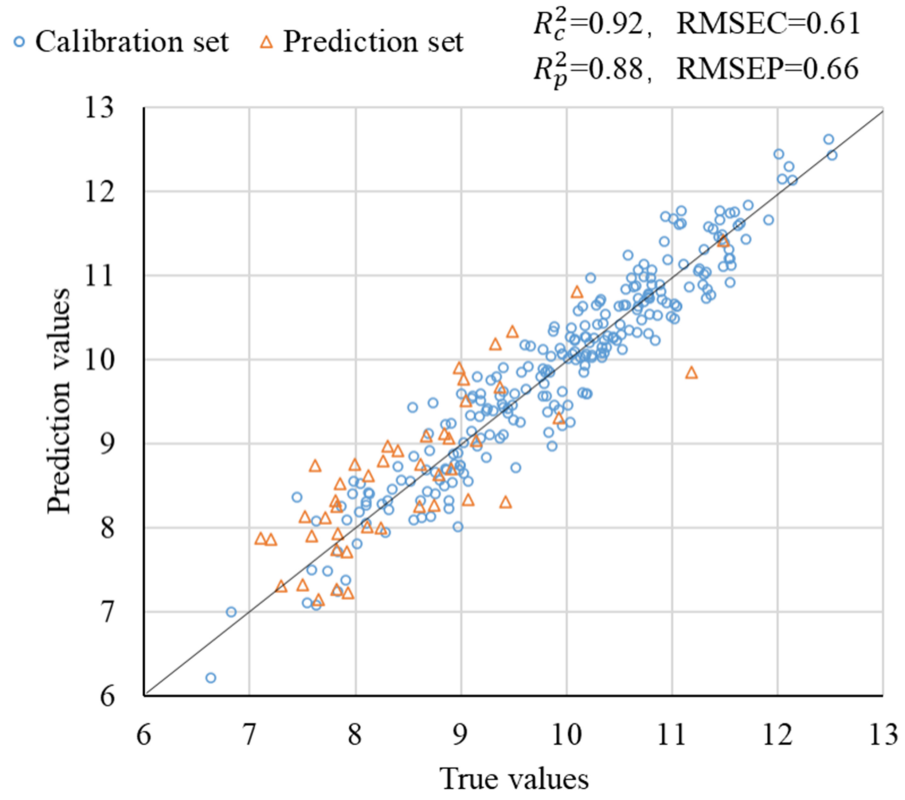

Figure 8. Optimal prediction results of firmness.

were put forward, which were based on the FFS, PCA and MAE methods. Three regression models were constructed and analyzed, and the prediction results were shown in Tables 6-8. The results revealed that all prediction models under the FFS, PCA, and MAE methods can be applied, with the MLP modeling method performing better overall than the other methods. Regardless of the regression model employed, the best prediction performance was based on MAE, followed by PCA, with FFS performing the poorest. The cumulative contribution rate of the first 40 principal components extracted from all variables by the PCA method was close to 98.9%, which reflected the original information of the electrical parameters relatively comprehensively. However, all models based on this method had higher prediction errors compared to MAE. Although MAE extracted different features for modeling from all electrical parameters and had a greater number of features than the PCA method, the model had a lower RMSE. Among them, MAE-MLP had the best performance in predicting firmness (R_c^2 = 0.92, RMSEC = 0.61, R_p^2 = 0.88, RMSEP = 0.66) and also had the best effect in predicting SSC/TAC (R_c^2 = 0.85, RMSEC = 1.62, R_p^2 = 0.82, RMSEP = 2.08). Therefore, MAE demonstrated effective feature extraction for nonlinear parameters and wide applicability in predicting the internal quality parameters of fruits and vegetables. Additionally, it emphasized the potential of electric spectrum parameters for further research on predicting physical and chemical criteria, providing new theories and methods for future studies on non-destructive detection of apples. The proposed approach exhibited promise in developing prediction models for determining SSC/TAC and firmness of apples. In future research, it may be worthwhile to consider, optimize, and implement various nonlinear feature extraction methods to predict SSC/TAC and firmness.

DECLARATIONS

Acknowledgments

We thank the Editor-in-Chief and all reviewers for their comments.

Authors' contributions

Conceived and proposed this study: Tian, X.; Huang, L.
Conducted key experiments and collected data: Tian, X.; Zhai, M.; Zhang, M.
Wrote and revised the manuscript.: Tian, X.; Hu, P.
Controlled the quality and checked the results of the manuscript: Huang, L.; Zhai, M.
Provided data support and experimental environment: Li, M.
Assisted in collecting data and revising the paper: Ren, L.

Availability of data and materials

The data and charts supporting the experimental results of this study are available within the article. The datasets used in the current study are available from the corresponding author upon reasonable request.

Financial support and sponsorship

This work was fully supported by SCI-TECH INNOVATION AND ACHIEVEMENTS PROMOTIONS in Northwest A&F University, grand number Z222021411, partially funded by SUBTASKS OF NATIONAL KEY R&D PROJECTS, grant number 2020YFD1100601.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Teixeira, G. C. M.; Junior, J. S. P.; Mattiuz, B.; et al. Spraying of calcium carbonate nanoparticles on pineapple fruit reduces sunburn damage. *S. Afr. J. Bot.* **2022**, *148*, 643-51. DOI
2. Zeng, S.; Cai, X.; Guo, W.; Zhang, Z.; Yang, S. Differences in optical properties and internal qualities of 'Fuji' apple produced in different areas of the Loess Plateau region. *Eur. J. Agron.* **2022**, *140*, 126608. DOI
3. Serrano-Finetti, E.; Castillo, E.; Alejos, S.; León, H. L. Toward noninvasive monitoring of plant leaf water content by electrical impedance spectroscopy. *Comput. Electron. Agric.* **2023**, *210*, 107907. DOI
4. Van Haeverbeke M, De Baets B, Stock M. Plant impedance spectroscopy: a review of modeling approaches and applications. *Front. Plant. Sci.* **2023**, *14*, 1187573. DOI PubMed PMC
5. Astashev, M. E.; Konchekov, E. M.; Kolik, L. V.; Gudkov, S. V. Electric impedance spectroscopy in trees condition analysis: theory and experiment. *Sensors* **2022**, *22*, 8310. DOI PubMed PMC
6. Cheng, J.; Yu, P.; Huang, Y.; Zhang, G.; Lu, C.; Jiang, X. Application status and prospect of impedance spectroscopy in agricultural product quality detection. *Agriculture* **2022**, *12*, 1525. DOI
7. Fathizadeh, Z.; Aboonajmi, M.; Hassan-Beygi, S. R. Nondestructive methods for determining the firmness of apple fruit flesh. *Inf. ProcessAgric.* **2021**, *8*, 515-27. DOI
8. Hiruta, T.; Sasaki, K.; Hosoya, N.; Maeda, S.; Kajiwara, I. Firmness evaluation of postharvest pear fruit during storage based on a vibration experiment technique using a dielectric elastomer actuator. *Postharvest. Biol. Technol.* **2021**, *182*, 111697. DOI
9. Feng, L.; Zhang, M.; Dong, Z.; Guo, J.; Zhang, H.; Liu, Z. Electrical impedance spectroscopy: potential non-destructive method for aflatoxin B1 in peanut. *Food. Meas.* **2024**, *18*, 9353-63. DOI

10. Tang, Y.; Zhang, H.; Liang, Q.; Xia, Y.; Che, J.; Liu, Y. Non-destructive testing of the internal quality of korla fragrant pears based on dielectric properties. *Horticulturae* **2024**, *10*, 572. DOI
11. Yu, Y.; Yao, M. Is this pear sweeter than this apple? A universal SSC model for fruits with similar physicochemical properties. *Biosyst. Eng.* **2023**, *226*, 116-31. DOI
12. Liu, D.; Wang, E.; Wang, G.; Wang, P.; Wang, C.; Wang, Z. Non-destructive sugar content assessment of multiple cultivars of melons by dielectric properties. *J. Sci. Food. Agric.* **2021**, *101*, 4308-14. DOI
13. Feng, L.; Gao, J.; Sui, X.; Weng, T.; Kong, A. Effect of fruit ripeness on electrical impedance spectrum parameters. *LWT.* **2024**, *208*, 116751. DOI
14. Yang, Z.; Amin, A.; Zhang, Y.; Wang, X.; Chen, G.; Abdelhamid, M. A. Design of a tomato sorting device based on the multisine-FSR composite measurement. *Agronomy* **2023**, *13*, 1778. DOI
15. Velásquez, S.; Franco, A. P.; Peña, N.; Bohórquez, J. C.; Gutierrez, N. Effect of coffee cherry maturity on the performance of the drying process of the bean: sorption isotherms and dielectric spectroscopy. *Food. Control.* **2021**, *123*, 107692. DOI
16. Ibba, P.; Falco, A.; Abera, B. D.; Cantarella, G.; Petti, L.; Lugli, P. Bio-impedance and circuit parameters: an analysis for tracking fruit ripening. *Postharvest. Biol. Technol.* **2020**, *159*, 110978. DOI
17. An, J.; Luo, X.; Xiong, L.; Tang, X.; Lan, H. Discrimination of inner injury of Korla fragrant pear based on multi-electrical parameters. *Foods* **2023**, *12*, 1805. DOI PubMed PMC
18. Ibba, P.; Crepaldi, M.; Cantarella, G.; et al. Design and validation of a portable AD5933–based impedance analyzer for smart agriculture. *IEEE. Access.* **2021**, *9*, 63656-75. DOI
19. Żywica, R.; Pierzynowska-Korniak, G.; Wójcik, J. Application of food products electrical model parameters for evaluation of apple purée dilution. *J. Food. Eng.* **2005**, *67*, 413-8. DOI
20. Bian, H.; Tu, P. The simultaneous monitoring of physiological change of apple based on dielectric parameters in static pressure. *J. Chin. Inst. Food. Sci. Technol.* **2019**, *19*, 279-85. DOI
21. Cai, C.; Li, X.; Ma, H.; et al. Non-destructive detection of freshness grade for apple fruit based on bio-impedance properties. *Trans. Chin. Soc. Agric. Mach.* **2013**, *44*, 147-52. DOI
22. Zhou, Z.; Zhang, Y.; Gu, Z.; Yang, S. X. Deep learning approaches for object recognition in plant diseases: a review. *Intell. Robot.* **2023**, *3*, 514-37. DOI
23. Berahmand, K.; Bahadori, S.; Abadeh, M. N.; Li, Y.; Xu, Y. SDAC-DA: semi-supervised deep attributed clustering using dual autoencoder. *IEEE. Trans. Knowl. Data. Eng.* **2024**, *36*, 6989-7002. DOI
24. Berahmand, K.; Li, Y.; Xu, Y. DAC-HPP: deep attributed clustering with high-order proximity preserve. *Neural. Comput. Applic.* **2023**, *35*, 24493-511. DOI
25. Yongnian, Z.; Yinhe, C.; Yihua, B.; Xiaochan, W.; Jieyu, X. Tomato maturity detection based on bioelectrical impedance spectroscopy. *Comput. Electron. Agric.* **2024**, *227*, 109553. DOI
26. Badfar, M.; Yildirim, M.; Chinnam, R. State-of-charge estimation across battery chemistries: a novel regression-based method and insights from unsupervised domain adaptation. *J. Power. Sources.* **2025**, *628*, 235760. DOI
27. Liu, D.; Guo, W. Nondestructive determination of soluble solids content of persimmons by using dielectric spectroscopy. *Int. J. Food. Prop.* **2017**, *20*, S2596-611. DOI
28. Guo, W.; Fang, L.; Liu, D.; Wang, Z. Determination of soluble solids content and firmness of pears during ripening by using dielectric spectroscopy. *Comput. Electron. Agric.* **2015**, *117*, 226-33. DOI
29. Guo, W.; Shang, L.; Zhu, X.; Nelson, S. O. Nondestructive detection of soluble solids content of apples from dielectric spectra with ANN and chemometric methods. *Food. Bioprocess. Technol.* **2015**, *8*, 1126-38. DOI
30. Reyes, A.; Yarlequé, M.; Castro, W.; Chuquizuta, S. Determination of dielectric properties of the red delicious apple and its correlation with quality parameters. In *2017 Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL)*, Singapore, Nov 19-22, 2017; IEEE, 2017; pp. 19-22. DOI
31. Bian, H.; Shi, P.; Tu, P. Determination of physicochemical quality of bruised apple using dielectric properties. *Food. Meas.* **2020**, *14*, 2590-9. DOI
32. Wang, D.; Li, L.; Liang, J.; et al. Nondestructive detection of kiwifruit treated with *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea by electrical method. *J. Food. Process. Preserv.* **2020**, *44*, e14860. DOI
33. Mohammed, M.; Munir, M.; Aljabr, A. Prediction of date fruit quality attributes during cold storage based on their electrical properties using artificial neural networks models. *Foods* **2022**, *11*, 1666. DOI PubMed PMC
34. Wang, D.; Yang, S. X. Intelligent feature extraction, data fusion and detection of concrete bridge cracks: current development and challenges. *Intell. Robot.* **2022**, *2*, 391-406. DOI
35. Weng, Y. K.; Chen, J.; Cheng, C. W.; Chen, C. Use of modern regression analysis in the dielectric properties of foods. *Foods* **2020**, *9*, 1472. DOI PubMed PMC
36. Niu, Y.; Wang, D.; Ye, L.; et al. Nondestructive detection of kiwifruit infected with *Penicillium expansum* based on electrical properties. *Postharvest. Biol. Technol.* **2023**, *195*, 112150. DOI
37. Bian, H.; Tu, P.; Hua-li, X.; Shi, P. Quality predictions for bruised apples based on dielectric properties. *J. Food. Process. Preserv.* **2019**, *43*, e14006. DOI
38. Aparisi P, Fortes Sanchez E, Contat Rodrigo L, Masot Peris R, Laguarda-miro N. A rapid electrochemical impedance spectroscopy and sensor-based method for monitoring freeze-damage in tangerines. *IEEE. Sensors. J.* **2021**, *21*, 12009-18. DOI

39. Guo, W.; Shang, L.; Wang, M.; Zhu, X. Soluble solids content detection of postharvest apples based on frequency spectrum of dielectric parameters. *Trans. Chin. Soc. Agric. Mach.* **2013**, *44*, 132-7. DOI
40. Shang, L.; Guo, W.; Nelson, S. O. Apple variety identification based on dielectric spectra and chemometric methods. *Food. Anal. Methods.* **2015**, *8*, 1042-52. DOI
41. Wang, R.; Wang, D.; Ren, X.; Ma, H. Nondestructive detection of apple watercore disease based on electric features. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 129-36. DOI
42. Hinton, G. E.; Salakhutdinov, R. R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504-7. DOI PubMed
43. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P. A. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, 2008; pp. 1096-103. DOI
44. Berahmand, K.; Daneshfar, F.; Salehi, E. S.; Li, Y.; Xu, Y. Autoencoders and their applications in machine learning: a survey. *Artif. Intell. Rev.* **2024**, *57*, 10662. DOI
45. Cacciarelli, D.; Kulahci, M. Hidden dimensions of the data: PCA vs autoencoders. *Qual. Eng.* **2023**, *35*, 741-50. DOI
46. Tsakiridis, N. L.; Samarinas, N.; Kokkas, S.; Kalopesa, E.; Tziolas, N. V.; Zalidis, G. C. In situ grape ripeness estimation via hyperspectral imaging and deep autoencoders. *Comput. Electron. Agric.* **2023**, *212*, 108098. DOI
47. Plakias, S.; Boutalis, Y. S. A novel information processing method based on an ensemble of auto-encoders for unsupervised fault detection. *Comput. Ind.* **2022**, *142*, 103743. DOI
48. Yuan, X.; Qi, S.; Wang, Y.; Xia, H. A dynamic CNN for nonlinear dynamic feature learning in soft sensor modeling of industrial process data. *Control. Eng. Pract.* **2020**, *104*, 104614. DOI
49. Meng, L.; Ding, S.; Xue, Y. Research on denoising sparse autoencoder. *Int. J. Mach. Learn. Cyber.* **2017**, *8*, 1719-29. DOI
50. Li, P.; Pei, Y.; Li, J. A comprehensive survey on design and application of autoencoder in deep learning. *Appl. Soft. Comput.* **2023**, *138*, 110176. DOI
51. Ahmed, S. A software framework for predicting the maize yield using modified multi-layer perceptron. *Sustainability* **2023**, *15*, 3017. DOI
52. Shang, L.; Wang, J.; Schäfer, D.; et al. Surrogate modelling of a detailed farm-level model using deep learning. *J. Agric. Econ.* **2024**, *75*, 235-60. DOI
53. Zhang, H.; Liu, Y.; Tang, Y.; Lan, H.; Niu, H.; Zhang, H. Non-destructive detection of the fruit firmness of Korla fragrant pear based on electrical properties. *Int. J. Agric. Biol. Eng.* **2022**, *15*, 216-21. DOI
54. Cai, C.; Li, Y.; Ma, H.; Li, X. Nondestructive classification of internal quality of apple based on dielectric feature selection. *Trans. Chin. Soc. Agric. Eng.* **2013**, *29*, 279-87. https://www.cabidigitallibrary.org/doi/full/10.5555/20143067192. (accessed 2025-02-21)
55. Drezner, Z.; Turel, O.; Zerom, D. A modified Kolmogorov–Smirnov test for normality. *Commun. Stat. Simul. Comput.* **2010**, *39*, 693-704. DOI
56. Pearson, E. S.; Snow, B. A. S. Tests for rank correlation coefficients: III. Distribution of the transformed Kendall coefficient. *Biometrika* **1962**, *49*, 185-91. DOI