sustainability

MDPI

Article

A Novel Intelligent Prediction Model for Higher Heating Value of Sustainable Solid Biomass Fuel Based on Bayesian Optimized Deep Neural Network

Yaoxun Feng [1,2], Qing Xu [1,2,*] and Changqing Li [1,2]

[1] College of Ocean Engineering and Energy, Guangdong Ocean University, Zhanjiang 524088, China
[2] Guangdong Provincial Key Laboratory of Intelligent Equipment for South China Sea Marine Ranching, Guangdong Ocean University, Zhanjiang 524088, China
* Correspondence: xuqing@gdou.edu.cn

Abstract

Biomass energy is recognized as a clean and sustainable energy source and is leveraged as a key enabler for driving the low-carbon transition of the energy system and achieving sustainable development. The higher heating value of solid biomass fuels (HHV-SBF) is a key parameter in its catalytic conversion process, and HHV-SBF is of great significance for catalyst design and matching, as well as the selection of reaction process parameters. To address the limitations in accuracy and generalization capability of traditional prediction methods for estimating the HHV-SBF, a dataset is constructed in this study that correlates chemical elements, proximate analysis parameters, and biochemical components with the HHV-SBF. Key hyperparameters of the deep neural network (DNN) are optimized using the Bayesian optimization algorithm. A Bayesian optimization-based deep neural network (BO-DNN) model is developed for the intelligent prediction of the HHV-SBF. Results show that the coefficient of determination (R^2) of the BO-DNN model reaches 92.6%. Compared to multiple mainstream deep learning algorithms, its performance is improved by approximately 11.61%, and the mean square error is significantly reduced. The BO-DNN model demonstrates excellent generalization capability and stability. The findings of this study provide a theoretical basis for the rapid and accurate prediction of the HHV-SBF.

Keywords: sustainable solid biomass fuel; higher heating value (HHV); catalysts; Bayesian; deep neural network (DNN)

Check for updates

Academic Editor: Jan Baeyens

Received: 11 January 2026
Revised: 9 February 2026
Accepted: 9 February 2026
Published: 12 February 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The catalytic pyrolysis of solid biomass fuel (SBF) [1,2] is the core technology for realizing high-value utilization of biomass, which refers to the efficient conversion of complex biomass components into high-quality, low-oxygen-content renewable liquid fuels and chemicals through catalysts. The catalytic pyrolysis of SBF provides a feasible path for low-carbon transformation in fields such as transportation engineering and chemical engineering. The higher heating value (HHV) is regarded as a core indicator for evaluating the energy quality of SBF [3], and accurate prediction of HHV plays a key role in fuel quality assessment, optimization of thermochemical processes, and improvement of energy conversion efficiency. Traditional HHV determination mainly relies on technical means such as experimental measurements and empirical formula estimations [4]. Among these, the oxygen bomb calorimetry method is used as the standard measurement method. Although

accurate results are obtained, the process is considered cumbersome, time-consuming, and costly, making it difficult to meet the needs of rapid industrial detection. Laser detection technology is employed to infer the higher heating value by analyzing the content of elements such as carbon and hydrogen in the fuel. However, this technology is restricted by stringent requirements for equipment precision and is easily interfered with by factors such as ambient temperature and humidity during measurement, as well as uniformity of fuel particles [5], limiting its application scenarios. The empirical formula method based on chemical components [6] (such as the Friedl and Channiwala models), though simple and fast, is mostly built on specific types of samples and assumptions. When the fuel source or composition changes, the prediction accuracy is significantly reduced. Therefore, the development of HHV prediction methods that combine high accuracy and strong generalization capability has become an important research direction in the field of biomass energy.

With the rapid development of artificial intelligence technology, machine learning-based HHV prediction methods have been proven to have significant advantages [7,8]. Early studies typically employed machine learning models that did not require networks, such as support vector machines, random forests, and gradient boosting trees. By training on characteristics such as volatile compounds, fixed carbon, and ash content, preliminary prediction of HHV is achieved. To some extent, these methods overcome the limitations of traditional methods and provide new ideas for HHV prediction. However, when dealing with high-dimensional, heterogeneous, and complex nonlinear interactions in data (such as simultaneous fusion of multi-source features such as elemental analysis, industrial analysis, and biochemical components), machine learning models that do not require networks often rely on fine feature engineering to construct effective feature representations. There may be limitations in automatically mining deep and complex nonlinear correlations between multi-source features [9].

In recent years, deep learning theory [10–12] has provided new technical pathways for HHV prediction. Compared to machine learning models that do not require networks, deep neural networks (DNN) are able to automatically extract high-level features of data and more effectively characterize the complex mapping relationship between input variables and HHV, thanks to their deep nonlinear network structure. Research shows that DNN exhibits stronger adaptability and higher prediction accuracy when dealing with variable biomass data [13]. For example, Veza et al. [14] applied various neural network structures and significantly improved the prediction accuracy for biomass heating value. Kandpal et al. [15] combined eXtreme Gradient Boosting (XGBoost) with Adaptive Boosting (AdaBoost) algorithms, reducing data noise and multivariate interference, and enhancing model stability. Hu et al. [16] design an enhanced event-triggering mechanism that can address issues of incomplete and inaccurate modeling. The verification results indicate that the proposed detection unit and enhanced event-triggering mechanism are effective, and the correlation between alarm thresholds and key performance indicators is clarified. Hu and Ma [17] design a new switching function based on the Euler Lagrange framework and formulate compensation error dynamics using command filters. Subsequently, virtual control strategies and adaptive laws are proposed to estimate the upper bounds of uncertainty, enabling the synthesis of ETC schemes. In addition, strict finite time stability is established for different situations of switching functions. Finally, the effectiveness of the theoretical results is verified through numerical examples. However, the performance of DNN is highly dependent on hyperparameter configuration, including learning rate, batch size, and network depth. Improper hyperparameter combinations can lead to slow model convergence, entrapment in local optima, or overfitting. Traditional optimization methods are inefficient, becoming a key bottleneck restricting the performance of DNN models.

In terms of model input features, the HHV-SBF is comprehensively influenced by chemical elements, proximate analysis parameters, and biochemical components [18]. Studies indicate that carbon content is significantly positively correlated with HHV, while oxygen content shows a negative correlation, and the lignin-to-cellulose ratio also has an important impact on energy density [13]. Currently, most prediction models use only a single type of feature (such as only elemental analysis or only proximate analysis), lacking the integration of multi-source features, which limits the comprehensive characterization of fuel properties by the model. At the same time, the multicollinearity problem among different feature parameters increases the complexity of modeling. Therefore, how to scientifically select feature parameters and establish effective multi-feature fusion models has become an important link in improving HHV prediction accuracy.

In response to the above situation, this paper proposes a method for predicting the HHV-SBF based on a Bayesian optimized deep neural network (BO-DNN). This study first systematically integrates three types of features—chemical elements, proximate analysis parameters, and biochemical components—to construct a dataset that correlates comprehensive input features with output HHV. Then, the Bayesian optimization algorithm is introduced to achieve automatic global optimization of DNN hyperparameters. Based on the BO-DNN model, an intelligent prediction model for the HHV-SBF is established. Finally, BO-DNN is compared with various mainstream deep learning algorithms to evaluate the model's prediction performance. The research results of this paper provide a theoretical basis for the rapid and accurate prediction of HHV-SBF.

2. Analysis of SBF Characteristic Parameters

The HHV-SBF is influenced by various characteristic parameters, which often exhibit complex nonlinear relationships. The interaction of chemical elements in biomass fuel plays a significant role in determining HHV-SBF. Research shows that higher carbon (C) content typically leads to higher HHV, while hydrogen (H) content is critical in enhancing the energy density of the fuel. Additionally, studies have highlighted that HHV-SBF is also affected by the content of oxygen (O), nitrogen (N), and sulfur (S) [19]. Lower oxygen content generally corresponds to higher energy density, and although N and S are present in small quantities, their high combustion heat value contributes to the increase in HHV. Table 1 lists several empirical formulas used for HHV prediction. The biochemical components of SBF, including lignin, hemicellulose, and cellulose, also significantly impact the prediction of HHV-SBF. As the proportion of lignin increases, HHV-SBF rises significantly, especially when the lignin content exceeds 25%, at which point HHV-SBF can reach over 20 MJ/kg.

Table 1. Summary of established correlations used for predicting the HHV-SBF.

No.	Equation	References
1	$HHV = 3.55\ C^2 - 232\ C - 1230\ H + 51.2 \times H + 131\ N + 2060\ O$	Friedl [20]
2	$HHV = 0.34191\ C + 1.1783\ H + 0.1005\ S - 0.1034\ O - 0.0151\ N$	Channiwala [21]
3	$HHV = -1.3675 + 0.3137\ C + 0.7009\ H + 0.0318\ O$	Sheng [22]
4	$HHV = -0.763 + 0.301\ C + 0.525\ H + 0.064\ O$	Jenkins [23]
5	$HHV = 0.341\ C + 1.323\ H + 0.0685 - 0.0153\ A - 0.1194(O + N)$	IGT [24]
6	$HHV = 0.352\ C + 0.944\ H + 0.105(S - O)$	Beckman [25]

Conversely, biomass fuels with higher hemicellulose content tend to exhibit lower HHV-SBF. Furthermore, a higher lignin-to-cellulose ratio correlates with an increase in the energy density of SBF[18]. In previous studies, proximate analysis parameters, such as volatile matter (VM), ash content (ASH), and fixed carbon (FC), have been commonly used for HHV-SBF prediction. The regression fitting results for these parameters in the present

dataset are shown in Figure 1. Although there is a specific relationship between these parameters and HHV, the mean squared error (MSE) and coefficient of determination (R^2) (Figure 1) indicate that these relationships are not typically linear. Consequently, simple linear regression models fail to capture the complex nonlinear relationships inherent in these data.

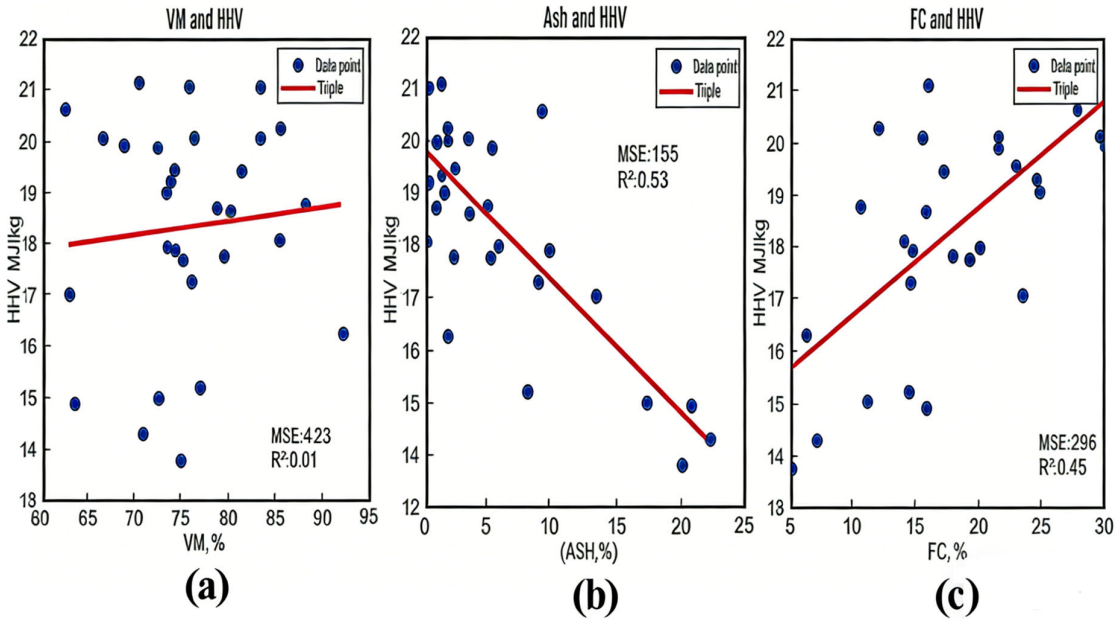

Figure 1. Relationship between volatile matter (**a**), ash content (**b**), fixed carbon (**c**), and higher heating value (HHV).

The correlation analysis between features and biomass high calorific value (HHV) in this study was quantitatively calculated using the Pearson correlation coefficient, which is a classic measure of the degree of linear correlation between continuous variables. The calculation formula is:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}} \quad (1)$$

where x_i represents the sample values of each input feature, y_i denotes the corresponding measured HHV values, $\overline{x}$ and $\overline{y}$ are the sample means of the features and HHV, respectively, and n is the total number of samples.

During the calculation process, all features were standardized and preprocessed to eliminate the influence of dimensional differences on correlation coefficients. At the same time, the statistical significance of the correlation was verified through bilateral significance tests ($p < 0.05$ was considered significant correlation). Finally, the correlation degree and significance between input features and between features and HHV were presented in the form of a correlation heatmap. This correlation analysis provides important theoretical and data support for the modeling selection and result interpretation of this study, which is reflected in two aspects: (1) It provides a basis for input feature screening and model architecture design. Through analysis, the core features significantly positively/negatively correlated with HHV were identified, and redundant features with weak correlation and severe multicollinearity with HHV were eliminated. The final 11 dimensional input features ensured the explanatory power of HHV and reduced the training complexity of the model. At the same time, they provided a data basis for the design of the input layer dimension and the optimization of the number of hidden layer neurons in the DNN model, avoiding overfitting or underfitting problems caused by unreasonable feature dimensions. And, (2) it provides support for the physical interpretation of the model prediction results; correlation

analysis clarifies the degree and direction of influence of each feature on HHV, which is mutually confirmed with the feature importance analysis results of the BO-DNN model. It can clearly explain the internal logic of the model prediction results, indicating that the model is not simply a black box prediction, but an accurate prediction based on the actual correlation law between features and HHV, improving the credibility and practical application value of the model prediction results.

This study conducted a correlation analysis between chemical elements, biochemical components, and proximate analysis parameters with a higher heating value (HHV), as shown in Figure 2. The heatmap illustrates the strength of correlations between different variables through color intensity and block size, with positive and negative values indicating the direction of the correlation. The results reveal a significant positive correlation between carbon content and HHV ($r > 0.9$), while oxygen content shows a significant negative correlation with HHV ($r < -0.9$). The nitrogen (N) and sulfur (S) content correlations with HHV are relatively low, exhibiting weak positive and weak negative correlations, respectively. Moreover, the impact of biochemical components such as cellulose, hemicellulose, and lignin on HHV is complex; most components show negative correlations, with hemicellulose being an exception. Therefore, when constructing the HHV-SBF prediction model, it is essential to include SBF characteristic parameters composed of chemical elements, biochemical components, and proximate analysis parameters as input features. By training the BO-DNN model, the complex nonlinear relationships between these input features and HHV can be learned, leading to more accurate predictions of the HHV-SBF.

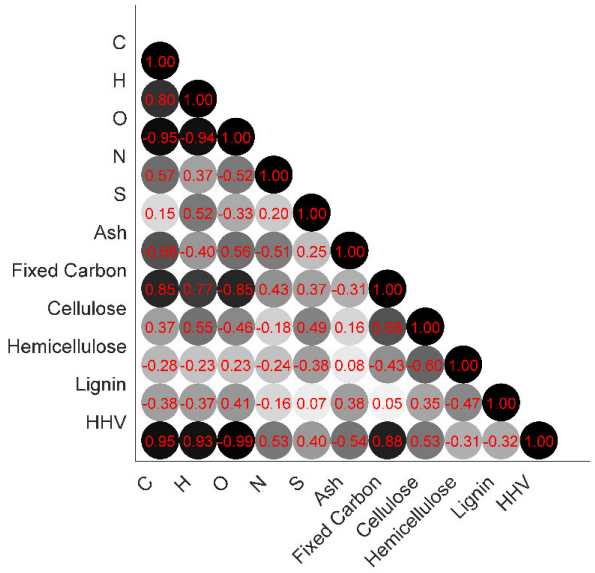

Figure 2. Pearson correlation coefficient matrix.

3. Methodology

3.1. DNN

A DNN is an extended deep learning architecture based on perceptrons that adjusts the network's weights and biases through the error backpropagation algorithm to minimize prediction errors [26]. Its main characteristic is that the working signals are propagated through the network layers, while the error signals are adjusted through backpropagation. DNNs typically consist of an input layer, multiple hidden layers, and an output layer, with the algorithm flow illustrated in Figure 3. In this study, the input layer comprises the characteristic parameters of SBF, which are utilized to train the model and enhance the prediction accuracy of the HHV-SBF.

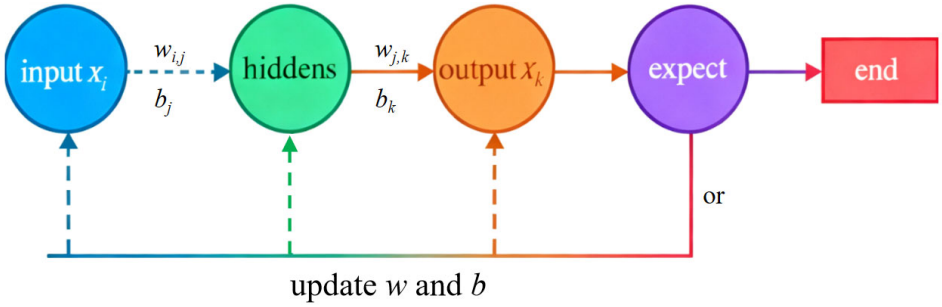

Figure 3. Algorithm flowchart of DNN.

Regarding the forward propagation of the DNN used to predict the HHV-SBF, its structure is illustrated in Figure 4. In the figure, the letter a represents the weights transmitted between the neurons of each layer, while b denotes the biases of the output neurons in each layer. The transmission process between the input layer and the hidden layers involves processing the input data through the weighted sum of weights and biases to generate the output for the hidden layers. This process provides a foundation for the subsequent nonlinear activation functions, allowing the results to be passed to the next layer.

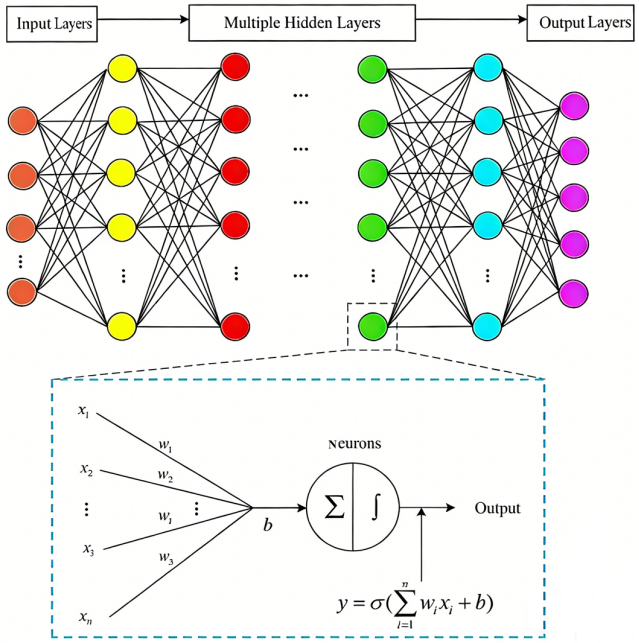

Figure 4. Schematic of DNN forward propagation.

In this study, i denotes the neuron indices of the input layer. We introduced the Sigmoid function as the activation function between the input and hidden layers to enable the neural network to process and represent complex nonlinear relationships effectively. The expression for the input value x_j of the hidden layer is as follows:

$$x_j = f_\sigma(S_j) \tag{2}$$

Here, j denotes the neuron indices of the hidden layer, S_j represents the transmitted values between the input layer and the hidden layers, and f_σ signifies the activation function Sigmoid. Therefore, the above expression can be modified as follows:

$$x_j = \frac{1}{1 + e^{-S_j}} \tag{3}$$

The transmitted values between the hidden layer and the output layer can be expressed as:

$$S_k = \sum_{j=1}^{n} w_{j,k} x_j + b_k \tag{4}$$

Here, k denotes the neuron indices of the output layer, S_k represents the transmitted values between the hidden layer and the output layer, and b_k signifies the biases of the output layer neurons. After processing through a linear activation function, the final predicted value of the output layer, $\hat{x}_k$, can be expressed as:

$$\hat{x}_k = S_k \tag{5}$$

During DNN backpropagation [27], the prediction of the HHV-SBF is treated as a regression problem. Consequently, we utilize the mean squared error as the loss function to evaluate the model's predictive performance.

$$E = \frac{1}{2} \sum_{k}^{n} (\hat{x}_k - x_k)^2 \tag{6}$$

Here, E denotes the loss value, and x_k represents the actual measured value of the HHV-SBF.

During the training process of the DNN, we employ the Adam algorithm to update all weights and biases in the network. The Adam algorithm calculates the gradients of the loss function concerning the network parameters, propagating the errors from the output layer back to the input layer to adjust the neurons' weights at each layer. The core idea of backpropagation is to compute the gradients layer by layer using the chain rule and to update the weights and biases with the Adam algorithm. After multiple rounds of training iterations, the model converges and is ultimately capable of effectively predicting the HHV-BSF.

We first calculate the gradients of each weight w and bias b with respect to the loss function E.

$$g_t^w = \frac{\partial E}{\partial w} \tag{7}$$

$$g_t^b = \frac{\partial E}{\partial b} \tag{8}$$

Here, g_t^w denotes the weight gradient at the t-th iteration, while g_t^b represents the bias gradient at the t-th iteration.

The Adam algorithm is favored in large-scale deep learning due to its ease of implementation, high computational efficiency, and rapid convergence. The parameter update rule is as follows:

The first-order moment estimate is the exponentially weighted average of the current gradient:

$$m_t^w = \beta_1 m_{t-1}^w + (1 - \beta_1) g_t^w \tag{9}$$

$$m_t^b = \beta_1 m_{t-1}^b + (1 - \beta_1) g_t^b \tag{10}$$

Here, m_t^w denotes the first-order moment of the weight, m_t^b represents the first-order moment of the bias, and β_1 is the first-order moment parameter.

The second-order moment estimate is the exponentially weighted average of the squared current gradient:

$$v_t^w = \beta_2 v_{t-1}^w + (1 - \beta_2)(g_t^w)^2 \tag{11}$$

$$v_t^b = \beta_2 v_{t-1}^b + (1 - \beta_2)(g_t^b)^2 \tag{12}$$

Here, v_t^w denotes the second-order moment of the weight, v_t^b represents the second-order moment of the bias, and β_2 is the second-order moment parameter.

Since m_t^w, m_t^b, v_t^w, and v_t^b initialized to zero, there is an inherent bias during the initial phase, particularly with small moment estimates in the early iterations. To correct this bias, Adam applies an unbiased estimate correction to the moment estimates:

$$\hat{m}_t^w = \frac{m_t^w}{1-\beta_1^t} \tag{13}$$

$$\hat{m}_t^b = \frac{m_t^b}{1-\beta_1^t} \tag{14}$$

$$\hat{v}_t^w = \frac{v_t^w}{1-\beta_2^t} \tag{15}$$

$$\hat{v}_t^b = \frac{v_t^b}{1-\beta_2^t} \tag{16}$$

Here, $\hat{m}_t^w$, $\hat{m}_t^b$, $\hat{v}_t^w$, and $\hat{v}_t^b$ represent the corrected first and second moment estimates for the weights and biases.

Ultimately, the weights and biases are updated using these corrected moment estimates:

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t^w}{\sqrt{\hat{v}_t^w} + \varepsilon} \tag{17}$$

$$b_t = b_{t-1} - \eta \frac{\hat{m}_t^b}{\sqrt{\hat{v}_t^b} + \varepsilon} \tag{18}$$

Here, η represents the learning rate, which determines the step size for each update, while ε is a numerical stability term designed to prevent zero values from occurring.

3.2. Bayesian Optimized DNN

In the process of training a DNN for predicting the characteristics of SBF, determining the optimal combination of model parameters is crucial [28]. To achieve this, Bayesian optimization is introduced to select the hyperparameter combinations for the DNN. Bayesian optimization is based on Bayes' theorem and uses a probabilistic model to fit the objective function, then analyzes the fit to select the points that are most likely to improve the model's performance for evaluation. This process iteratively updates the surrogate model, combining historical information to optimize the search strategy, thereby improving search efficiency while reducing the number of evaluations [29]. Bayesian optimization consists of two core components: the probabilistic surrogate model and the acquisition function. The probabilistic surrogate model provides predictions of the objective function values and their uncertainties, while the acquisition function determines the next set of hyperparameter values to explore, guiding the model towards the optimal solution. Through the iterative process of Bayesian optimization, the performance of the DNN model can be significantly improved, and the hyperparameter tuning process can be accelerated.

In the DNN model, the choice of hyperparameters plays a crucial role in model performance [30]. Based on the weight and bias update formulas of the backpropagation algorithm, it is evident that the initial learning rate (η) and mini-batch size (γ) directly affect the convergence performance of the model. Therefore, in this paper, we use these two hyperparameters (η and γ) to construct a new optimization combination to enhance the model's training effectiveness and convergence speed.

When selecting the surrogate function, this paper adopts the Gaussian process (GP), a non-parametric model. As a powerful Bayesian method, the prior distribution of the Gaussian process can be represented as:

$$f(y) = E(y) + \varepsilon \tag{19}$$

In Bayesian optimization, assume that $E(y)$ is the loss function predicted by the DNN and ε represents Gaussian noise. Given the observation data $D = (y_i, f_i)$, the posterior distribution of the Gaussian process can be updated using Bayes' theorem, which is expressed as:

$$p(E|D) = N(\hat{x}, \Sigma) \tag{20}$$

Here, $\hat{x}$ represents the predicted value of the objective function, and Σ represents the covariance matrix.

Bayesian optimization determines the next optimal hyperparameter combination through the acquisition function. In this study, expected improvement (EI) is employed as the acquisition function. EI offers advantages such as having fewer parameters, integrating the probability of improvement, and balancing the trade-off between exploration (depth) and exploitation (breadth). The acquisition function is defined as:

$$EI(y) = E[\max(E(y^+) - E(y), 0)] \tag{21}$$

Here, y^+ represents the current optimal DNN hyperparameter combination, and E denotes the expectation operation. When using Bayesian optimization for hyperparameter selection in the DNN model, the goal is to solve for the optimal hyperparameter combination using the following formula:

$$y^* = \text{argmin} f(y) \tag{22}$$

Here, y^* represents the hyperparameter combination, and argmin denotes the value at which the objective function reaches its minimum.

3.3. Indicators of HHV-SBF

Using the BO-DNN model to predict the HHV-SBF is essentially a regression problem. For each SBF sample to be predicted, the network model outputs a scalar value as the result, representing the predicted heating value for the sample. By optimizing the learning rate and mini-batch size, the BO-DNN model significantly improves prediction accuracy, providing a reliable basis for HHV-SBF prediction.

Assuming a DNN model is used for HHV prediction, the last layer of the model outputs a scalar value representing the predicted HHV-SBF. To convert this scalar into a meaningful physical quantity, the model undergoes a Bayesian optimization process during training, continuously adjusting model parameters to minimize the loss function and improve prediction accuracy. In the final prediction phase, by fine-tuning the learning rate and mini-batch size of the DNN model, the output prediction value is ensured to be as close as possible to the actual HHV of the sample. This process can be expressed as:

$$\hat{x} = f_{DNN}(x, \mathrm{y}) \tag{23}$$

Here, $\hat{x}$ represents the predicted HHV-SBF, f_{DNN} denotes the prediction function modeled by the BO-DNN, x is the input fuel feature set, and y is the optimal combination of learning rate η and mini-batch size γ. The output layer of the BO-DNN model directly provides the prediction result, reflecting the higher heating value of the sample (as

shown in Figure 5). Thus, by optimizing the model parameters, the predicted HHV-SBF is accurately obtained.

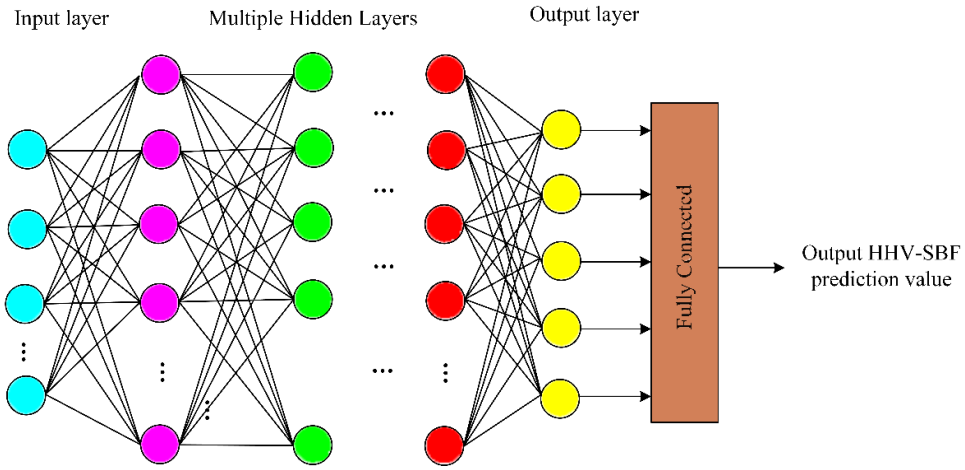

Figure 5. Architecture of BO-DNN higher heating value prediction model for SBF.

3.4. BO-DNN Model Architecture

(1) Design Basis for DNN Model Architecture

Network layers and number of neurons: The model adopts a concise structure of "input layer −1 hidden Layer—output Layer". The number of neurons in the input layer is 11, which is consistent with the input feature dimensions (chemical elements, industrial analysis parameters, biochemical component indicators). The number of hidden layer neurons is set to 7, which is determined through pre-experiments to ensure the model's fitting ability while avoiding overfitting and achieving efficient mapping of input features.

Activation function selection: Sigmoid function is selected as the activation function between the input layer and the hidden layer, which has good nonlinear mapping ability and can effectively handle the nonlinear relationship between biomass characteristics and HHV. The reason for choosing a linear activation function between the hidden layer and the output layer is that the HHV prediction in this study belongs to the continuous value regression task, and the linear activation function can directly output the predicted calorific value that conforms to the physical meaning.

Anti-overfitting measures: Dropout technology is introduced in the hidden layer and the dropout rate set to 0.3. This parameter is determined through pre-experiments and can randomly block the connections of some neurons, effectively avoiding overfitting of the model to the training data and improving the model's generalization ability.

(2) The selection criteria for Bayesian optimization hyperparameters

This study selected learning rate and small batch sample size as the core hyperparameters for Bayesian optimization, without optimizing parameters such as network layers and number of neurons. The main basis is as follows:

Targeted optimization objective: The core objective of this study is to improve the convergence speed and prediction accuracy of DNN models in HHV prediction tasks. The preliminary experimental results indicate that the learning rate directly determines the step size of model parameter updates. If it is too small, it will lead to slow convergence, while if it is too large, it is prone to training oscillations. The small batch sample size affects the stability and computational efficiency of gradient descent, which are the core sensitive hyperparameters that affect the performance of regression task models. In contrast, the architecture with 1 network layer and 7 hidden layer neurons showed stable fitting effects in the pre-experiment, but the adjustment space is limited.

Consideration of optimization efficiency: The core advantage of Bayesian optimization is to efficiently search for the optimal solution in the hyperparameter space through probabilistic surrogate models and collection functions. If too many hyperparameters are included, it will significantly increase the dimensionality and computational cost of optimization and reduce search efficiency. This study focuses on core sensitive hyperparameters, which significantly improves the training efficiency of the model while ensuring optimization effectiveness.

3.5. Prediction of HHV-SBF by BO-DNN Model

In the prediction of the HHV-SBF, the diversity and complexity of the data necessitate a thorough consideration of the interrelationships among various components. By employing the BO-DNN optimization method, multiple data sources, including chemical elements, proximate analysis parameters, and biochemical compositions, can be effectively integrated to achieve accurate HHV-SBF predictions. The BO-DNN model leverages the architecture of multi-layer DNN to uncover deep relationships among components, optimizing parameter settings and improving prediction accuracy. Its deep learning capability demonstrates significant advantages in handling complex biomass characteristics, thereby providing enhanced accuracy and reliability in HHV-SBF prediction. The specific steps of applying the BO-DNN model for HHV-SBF prediction are shown in Figure 6.

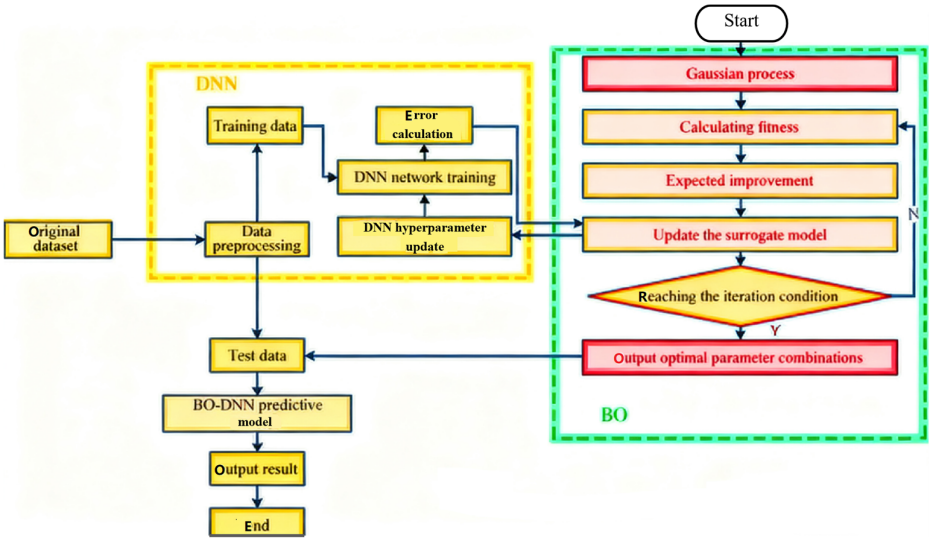

Figure 6. HHV prediction of SBF for BO-DNN models.

The network structure model consists of four main components: input layer, DNN layer, BO layer, and output layer. The detailed process is as follows:

(1) Data Preprocessing

Prepare the compositional data of SBF, divide it into training, testing, and validation sets in a certain ratio, and normalize the data;

(2) Input Layer

Convert the preprocessed compositional data into a format suitable for the deep learning model and input it in sequential form;

(3) DNN Layer

Build a DNN model, define the network structure, determine the required hyperparameters, and set the ranges of hyperparameters to be optimized (learning rate and mini-batch size);

(4) Bayesian Optimization Layer

Perform Gaussian process and expected improvement, using the output prediction error of the DNN model as the optimization objective function to optimize hyperparameters;

(5) Output Layer

Provide the optimized model performance, the optimal set of hyperparameters, and the corresponding BO-DNN model;

(6) Prediction and Validation

Use the BO-DNN model to predict the HHV-SBF on the test set, and output the prediction results, fitness, and accuracy to validate the model's performance.

4. Results and Analysis

4.1. Data Preparation

(1) Dataset information

The dataset of this study is integrated from public literature and online databases, with the core data source being the biomass characteristic database constructed by Nhuchen et al., supplemented by the biomass combustion calorific value dataset provided by Parikh et al. for external validation. The dataset covers various typical solid biomass fuels such as agricultural residues, woody plants, and seeds. Each sample contains 11 input features: chemical element features (C, H, O, N, S content), industrial analysis features (volatile matter (VM), ash content (ASH), fixed carbon proportion (FC)), biochemical component features (lignin, hemicellulose, cellulose content), and the output target variable is high calorific value (HHV). The final valid sample size in this paper is 334, which is divided by stratified sampling into: a training set of 234 samples (70%), a validation set of 50 samples (15%), and a test set of 50 samples (15%). The proportions of these components are shown in Figure 7.

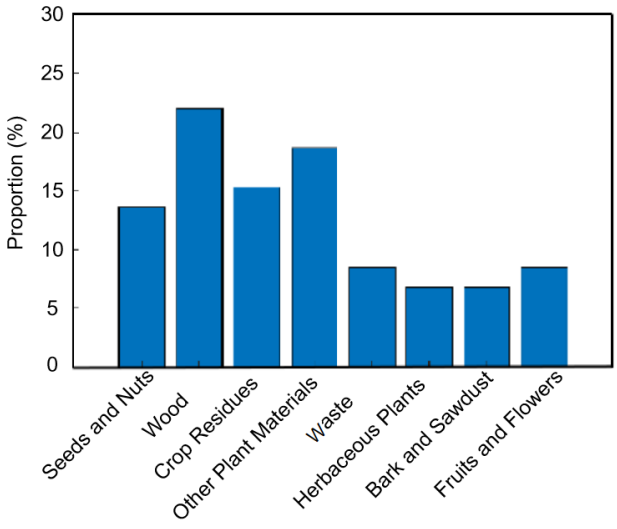

Figure 7. Proportions of solid biomass types.

(2) Data cleaning process

To ensure data quality, this study implemented strict data cleaning steps, as follows:

Outlier removal: Using the quartile method to identify and remove abnormal samples in each feature dimension.

Missing value filling: For samples with a small number of missing features, K-nearest neighbor algorithm is used for filling to avoid data loss caused by direct deletion of samples.

Data consistency verification: Due to the fact that the data sources come from different literature, there are differences in the indicator units of some samples. In this study, all component contents were uniformly converted to mass percentages (%), and HHV was uniformly converted to MJ/kg to ensure consistency in data dimensions.

(3) Dataset partitioning scheme

This study used stratified sampling to divide the dataset to ensure consistent distribution ratios of various types of biomass samples in the training, validation, and testing sets. The specific division ratio is:

Training set: accounting for 70%, used for model parameter fitting and training.

Validation set: accounting for 15%, used to monitor overfitting phenomena during model training and hyperparameter optimization of Bayesian optimization algorithm.

Test set: accounting for 15%, used to evaluate the final generalization performance of the model.

(4) Data preprocessing operation

To eliminate the impact of dimensional differences in various features on model training, this study standardized all input features and calculated all statistical measures for preprocessing operations based solely on the training set to avoid information leakage between the validation and test sets.

4.2. Performance Evaluation of the BO-DNN Model

Before applying the BO-DNN model to predict the HHV-SBF, it is essential to validate the algorithm's effectiveness. To achieve this, extensive testing was conducted on 23 classic benchmark functions [31]. These test functions cover a wide range of complexities and dimensions, aiming to evaluate the performance of the BO-DNN algorithm across various optimization challenges, including its convergence speed, global search capability, and ability to escape local optima.

Based on the test results shown in Figure 8, the BO-DNN model demonstrates a faster descent rate within the parameter space, allowing it to converge more quickly to the optimal solution. This indicates that the BO-DNN algorithm has significant advantages in handling high-dimensional, nonlinear, and unstable data. The improved BO-DNN algorithm, when applied to predict the HHV-SBF, efficiently identifies the optimal combination of DNN learning rate and batch size. This enhances the model's adaptability to the complexity of varying features and data distributions, significantly improving its generalization ability and prediction accuracy.

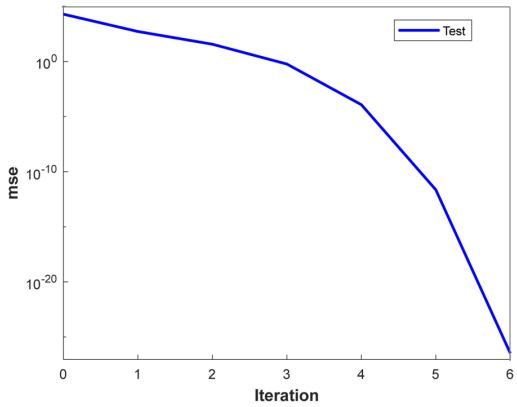

Figure 8. Results of parameter testing for the BO-DNN model.

4.3. Experimental Solution and Analysis

The BO-DNN network structure in this study consists of multiple DNN layers, including an input layer, multiple hidden layers, and an output layer. BO is used to optimize the learning rate and mini-batch size of the DNN model. The optimization range for these parameters is first defined, and BO is applied within this range to select the parameter combination with the highest prediction accuracy as the optimal result: batch size = 69 and learning rate = 10^{-6}. By leveraging BO to determine key parameters, the time cost associated with manual experimentation is significantly reduced, and both computational accuracy and efficiency are greatly improved.

To prevent overfitting, dropout technology is applied in the hidden layers, with a dropout rate set to 0.3. The training process combines the Adam optimization algorithm to accelerate the learning process and enhance convergence speed. Before the output layer, a Sigmoid activation function is employed to enhance the model's representational capability, and mean squared error (MSE) is used as the loss function to compute prediction errors. Figure 9 illustrates the loss function and prediction accuracy during the training process. Initially, the loss value decreases rapidly and eventually converges, achieving optimal performance in HHV prediction. The final loss function value is 0.14, and the coefficient of determination (R^2) reaches 0.926.

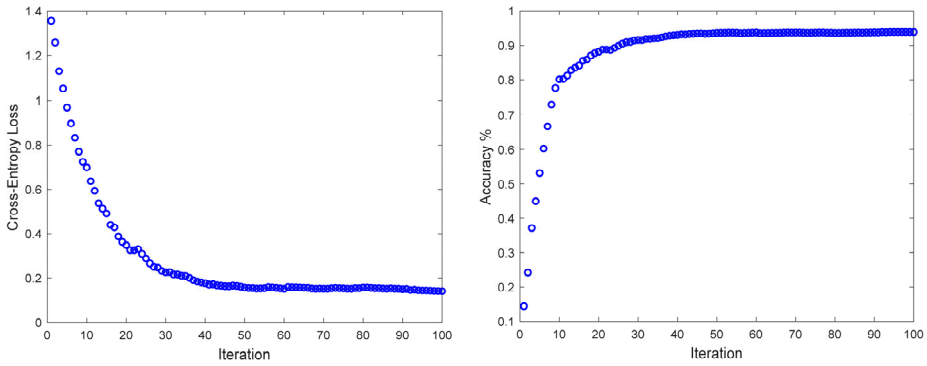

Figure 9. Changes in the loss function value and prediction accuracy with the number of iterations. (**a**) Loss function, (**b**) Prediction accuracy.

To evaluate the accuracy of the BO-DNN model, a comparative analysis was conducted between the BO-DNN and conventional DNN models. Validation dataset predictions of the higher heating value of SBF were used to compare the prediction accuracy of the two models. The specific results are shown in Figure 10.

By comparing the mean squared error (MSE) of the BO-DNN and traditional DNN models in predicting the HHV-SBF, the potential of the BO-DNN in this application is more intuitively demonstrated. Specifically, experimental results show that the BO-DNN model outperforms the traditional DNN model in predicting HHV for various SBF samples, especially in capturing the nonlinear relationships within the data. While the traditional DNN model is capable of capturing certain nonlinear features, its prediction error remains relatively high, particularly when dealing with complex data patterns, where its generalization ability is limited. On the other hand, BO-DNN, by incorporating Bayesian optimization, can automatically adjust the network structure and hyperparameters, significantly improving the prediction accuracy. This result highlights the distinct advantages of BO-DNN in addressing complex nonlinear issues and demonstrates its promising potential in HHV prediction of SBF. Figure 10 clearly shows the comparison between the two models based on the MSE metric, further validating the advantage of the BO-DNN model in enhancing prediction accuracy.

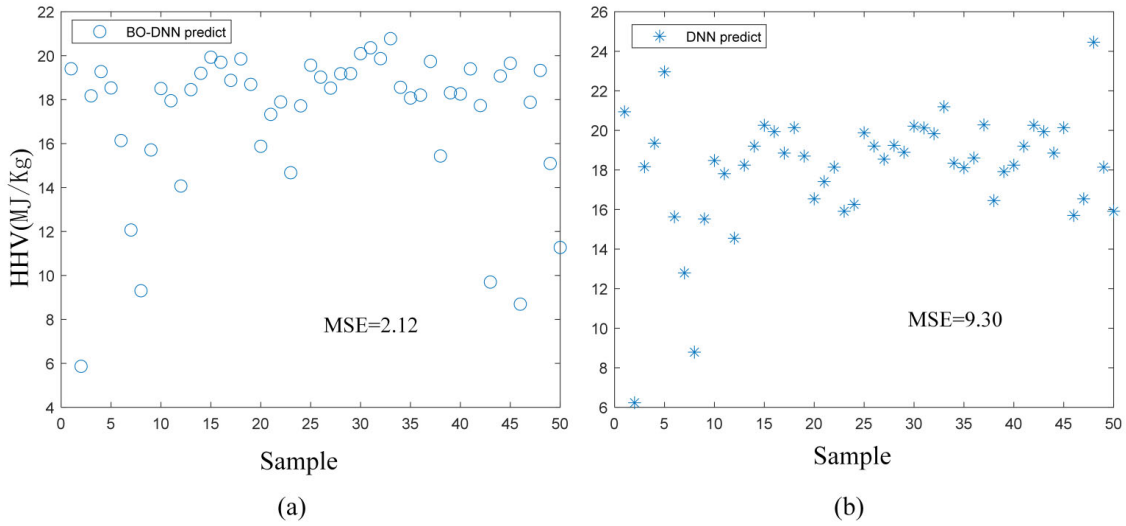

Figure 10. Prediction of HHV for SBF. (**a**) BO-DNN prediction results. (**b**) DNN prediction results.

4.4. Comparison with Other Intelligent Optimization Algorithms

To further validate the effectiveness of the BO-DNN algorithm, we compared its predictions with those from other intelligent optimization algorithm models. Figure 11 illustrates the comparison between the predicted and actual values of different models in the HHV-SBF prediction task. By analyzing and comparing the MSE and mean absolute percentage error (MAPE) of each model, we can more intuitively assess the strengths and weaknesses of the models in HHV-SBF prediction. The results show that, compared to other intelligent algorithm models, the BO-DNN model performs more effectively in addressing the HHV-SBF prediction problem. Specifically, the BO-DNN model not only provides higher prediction accuracy but also exhibits smaller error fluctuations, demonstrating superior stability. These results further confirm the significant advantages of the BO-DNN model in solving complex nonlinear problems, highlighting its broad application potential in HHV-SBF prediction tasks.

In Figure 11, the length of the error bars represents the error range or uncertainty of the data. The length of the error bars typically reflects the reliability of the measurement results, with the length being proportional to the size of the error. Longer error bars indicate larger errors for the data points, meaning the actual values may deviate more widely from the predicted or experimental values. Conversely, shorter error bars suggest smaller errors, with the actual values being more likely to cluster around the predicted values. By observing Figure 11d, we can clearly see that the error bars of the BO-DNN model are significantly shorter, indicating that the BO-DNN model exhibits smaller prediction errors compared to other methods when predicting HHV-SBF. This further validates the superiority of the BO-DNN model in improving prediction accuracy and stability, demonstrating its high accuracy and reliability in higher heating value prediction of SBF.

In order to fully consider the randomness of DNN training and the Bayesian optimization algorithm, this study conducted 30 independent repeated trainings on the BO-DNN model and comparison models (COA-DNN, HHO-DNN, MPA-DNN, traditional DNN). Each training session used different random seeds to initialize the network weights and hyperparameter search starting points, in order to eliminate the impact of randomness in a single experiment on the results. During the experiment, the key evaluation indicators of all models on the test set were recorded, including mean square error (MSE) and mean absolute percentage error (MAPE). Statistical analysis was conducted on the results of 30 repeated experiments to evaluate the stability of model performance. Figures 12 and 13 show the average MSE and MAPE values of each model, respectively. The results showed

that the MSE and MAPE mean of the BO-DNN model were significantly lower than those of COA-DNN, HHO-DNN, MPA-DNN, and traditional DNN, indicating that the performance advantage of the BO-DNN model is statistically significant and not caused by random factors.

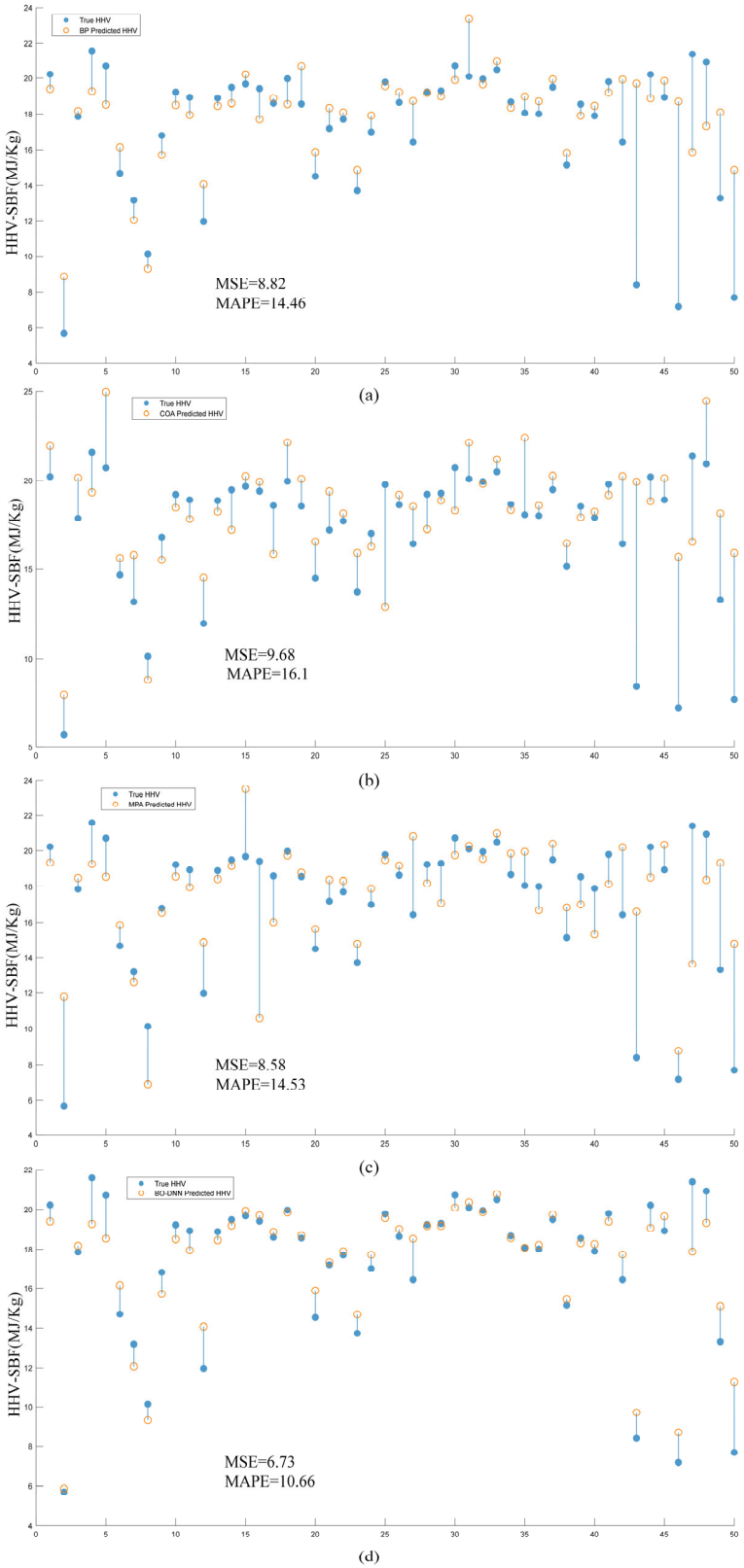

Figure 11. Prediction results of HHV-SBF using various intelligent algorithms. (**a**) COA-DNN prediction results, (**b**) HHO-DNN prediction results, (**c**) MPA-DNN prediction results, (**d**) BO-DNN prediction results.

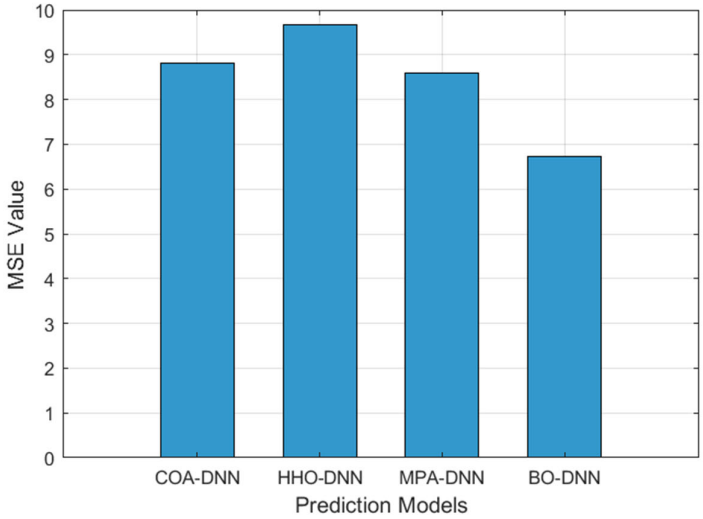

Figure 12. Comparison of MSE for predicting HHV using different models.

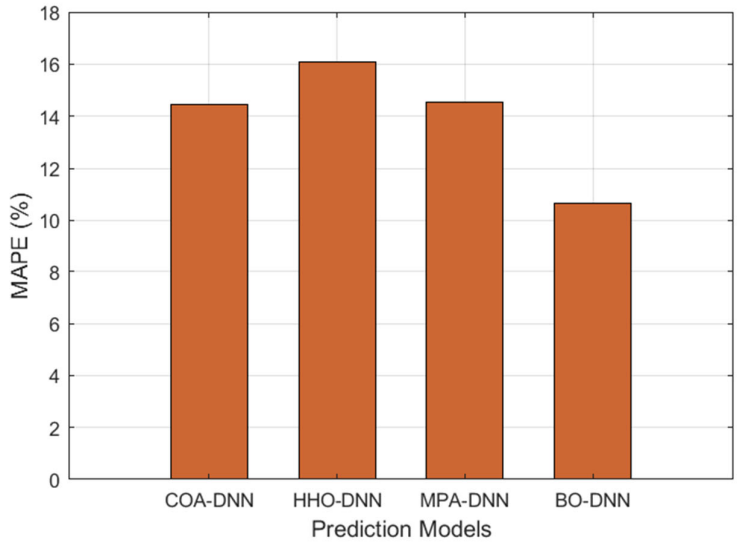

Figure 13. Comparison of MAPE for predicting HHV using different models.

4.5. Practical Application of the BO-DNN Model

Intelligent algorithm models often encounter the issue of overfitting in practical applications, where the model performs well on training data but poorly on new or validation data. To assess the practical applicability of the BO-DNN model, we applied it to predict the HHV of various SBF data samples to evaluate its generalization ability. For this purpose, we used the biomass combustion total heating value data sample provided by Parikh et al. [32] for cross-validation, with the prediction results shown in Figure 14.

From the prediction results in Figure 14 (Figure 14a), there is a high consistency between the predicted HHV values and the actual values, with most data points being close to the actual values, indicating that the BO-DNN model effectively captures the relationship between the characteristics of SBF and their HHV. Additionally, we analyzed the error by calculating the difference between the predicted and actual values, as shown in Figure 12. The majority of the prediction errors are concentrated between −0.6 and 0.6 MJ/Kg, further demonstrating the high accuracy and good generalization ability of the BO-DNN model for HHV-SBF prediction. In conclusion, the BO-DNN model exhibits excellent predictive performance and strong adaptability in practical applications, making it an effective tool for predicting the heating values of various SBF.

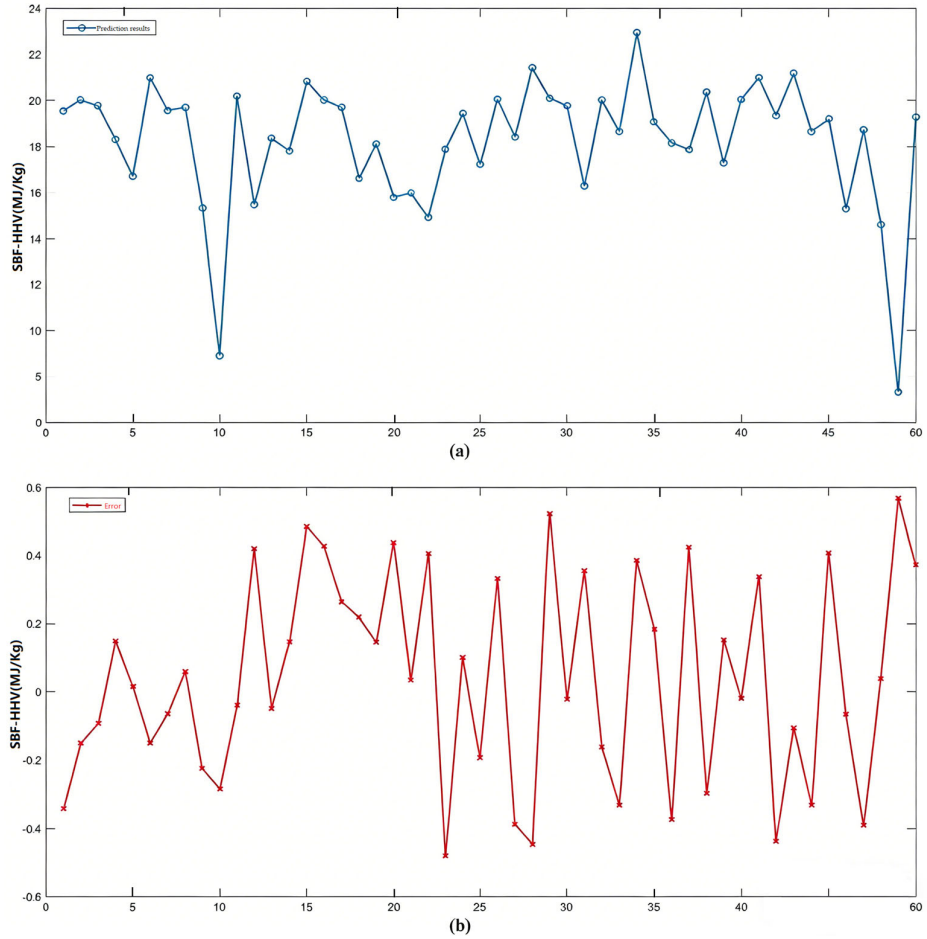

Figure 14. Sample prediction results and error fluctuation. (**a**) Prediction results, (**b**) error fluctuation.

The innovation of the BO-DNN model proposed in this paper is mainly reflected in the following three aspects:

(1) Construction of a dataset based on multi-source feature fusion

Most existing HHV prediction models are typically based on a single type of feature (such as elemental analysis only) for modeling, failing to fully utilize the multidimensional information that affects HHV. This article systematically integrates three types of features: chemical elements, industrial analysis parameters, and biochemical components, and constructs a comprehensive input feature set for the first time. This fusion of multi-source heterogeneous features more comprehensively reflects the complex properties of solid biomass fuels, providing richer discriminative information for the model, thereby improving the systematicity and accuracy of prediction.

(2) Bayesian optimization hyperparameter automatic tuning strategy for HHV prediction

Although the combination of Bayesian optimization and DNN is not entirely new in the field of machine learning, most existing research adopts the general strategy of "full parameter traversal optimization", which has the problems of high computational cost and weak targeting. This article focuses on the regression task characteristics of HHV prediction, accurately identifying the two core hyperparameters that have the greatest impact on model convergence and prediction accuracy: learning rate and small batch sample size. Using Gaussian process as a proxy model, combined with expected lift as the acquisition function, we efficiently searched for the optimal combination in the hyperparameter space. The experimental results showed that the optimized parameter combination (batch size = 69, learning rate = 10^{-6}) significantly improved the convergence speed and prediction stability

of the model. This targeted optimization strategy has not been reported in existing research on biomass HHV prediction.

(3) Significant performance improvement achieved on HHV prediction tasks

This article compares the BO-DNN model with various mainstream deep learning models through rigorous experimental design. The experimental results show that the BO-DNN model improves the coefficient of determination (R^2) of HHV to 92.6%, which is about 11.61% higher than the traditional DNN model. The mean square error is significantly reduced, and it exhibits good generalization ability and stability on the external validation set. These results demonstrate the superiority of the proposed method in solving complex nonlinear prediction problems in specific fields, rather than simply stacking algorithms.

5. Discussion

The core value of this study lies in the rapid and accurate prediction of high calorific value using conventional biomass characteristic indicators as input, replacing the problems of long detection cycle, high cost, and complex operation of traditional oxygen bomb calorimetry [33]. It can meet the rapid detection needs of bulk raw materials in large-scale scenarios such as biomass power generation and formed fuel production, providing efficient technical support for raw material grading and energy planning. At the same time, the concise model architecture reduces deployment costs and is more suitable for industrial practical applications. The model can be widely used for biomass raw material sorting and grading, energy project planning, and rapid on-site testing by grassroots institutions, greatly reducing the detection threshold and cost. In terms of generalization ability, the model has been validated by external independent datasets and has a certain degree of generalization ability for common biomass from different sources. The limitations of this article mainly lie in the fact that the dataset is mainly composed of common biomass, and there is insufficient coverage of extreme environmental biomass, new cultivated varieties, and industrial biomass waste.

The future research directions mainly include: (1) building an open-source and continuously updated database of biomass fuel characteristics, covering a wider range of raw material types and geographical sources; and (2) exploring transfer learning frameworks to enable models to quickly adapt to new types of biomass using existing knowledge.

6. Conclusions

Aiming at the current bottlenecks in the prediction of the HHV-SBF, this paper integrates deep learning with intelligent optimization algorithms to propose a BO-DNN intelligent prediction model which effectively solves the shortcomings of traditional methods in terms of HHV-SBF prediction accuracy and generalization ability. The research results of this project are of great significance for the design and matching of catalysts and the selection of reaction process parameters in the SBF catalytic pyrolysis process. The main conclusions are as follows:

(1) By integrating multi-source data, including chemical elements, proximate analysis parameters, and biochemical components, a composite feature set is constructed that systematically characterizes the influencing factors of the HHV-SBF.
(2) The Bayesian optimization algorithm is employed to automatically adjust hyperparameters of the DNN, such as the learning rate and batch size. This overcomes the inefficiency of traditional parameter tuning methods and significantly improves the model's convergence speed and stability.
(3) Results indicate that the coefficient of determination (R2) of the BO-DNN model reaches 92.6%, showing significant improvement compared to traditional DNN

and other intelligent algorithms. Furthermore, the model maintains small error fluctuations on the external validation set, demonstrating good precision and generalization capability.

Author Contributions: Conceptualization, Q.X. and Y.F.; methodology, Q.X. and Y.F.; software, Q.X. and Y.F.; validation, Q.X. and Y.F.; formal analysis, Q.X. and Y.F.; writing—original draft preparation, Q.X. and Y.F.; writing—review and editing, Q.X., Y.F. and C.L.; funding acquisition, Q.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Nos. 52476190 and 52376171).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in this study are openly available in the Zenodo repository under the following link: https://doi.org/10.5281/zenodo.17541227, accesed on 6 November 2025.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li, H.; Yang, S.; Zhao, W.; Xu, Z.; Zhao, S.; Liu, X. Prediction of the physicochemical properties of woody biomass using linear prediction and artificial neural networks. *Energy Sources Part A Recovery Util. Environ. Eff.* **2016**, *38*, 1569–1573. [CrossRef]
2. Bartocci, P.; Zampilli, M.; Bidini, G.; Fantozzi, F. Hydrogen-rich gas production through steam gasification of charcoal pellet. *Appl. Therm. Eng.* **2018**, *132*, 817–823. [CrossRef]
3. Faramarzi, A.; Heidarinejad, M.; Mirjalili, S.; Gandomi, A.H. Marine Predators Algorithm: A nature-inspired metaheuristic. *Expert Syst. Appl.* **2020**, *152*, 113377. [CrossRef]
4. Shi, Z.; Ferrari, G.; Ai, P.; Marinello, F.; Pezzuolo, A. Artificial Intelligence for Biomass Detection, Production and Energy Usage in Rural Areas: A review of Technologies and Applications. *Sustain. Energy Technol. Assess.* **2023**, *60*, 103548. [CrossRef]
5. Kim, S.U.; Yun, J.S.; Ri, J.S.; Hong, K.I.; Chon, O.S.; Ri, J.H.; Ri, J.M. Improvement of HHV prediction model of biomass based on the ultimate analysis. *J. Renew. Sustain. Energy* **2021**, *13*, 053101. [CrossRef]
6. Maksimuk, Y.; Antonava, Z.; Krouk, V.; Korsakova, A.; Kursevich, V. Prediction of higher heating value (HHV) based on the structural composition for biomass. *Fuel* **2021**, *299*, 120860. [CrossRef]
7. Ascher, S.; Watson, I.; You, S. Machine learning methods for modelling the gasification and pyrolysis of biomass and waste. *Renew. Sustain. Energy Rev.* **2022**, *155*, 111902. [CrossRef]
8. Abdollahi, S.A.; Ranjbar, S.F.; Razeghi Jahromi, D. Applying feature selection and machine learning techniques to estimate the biomass higher heating value. *Sci. Rep.* **2023**, *13*, 16093. [CrossRef]
9. Onsree, T.; Tippayawong, N.; Phithakkitnukoon, S.; Lauterbach, J. Interpretable machine-learning model with a collaborative game approach to predict yields and higher heating value of torrefied biomass. *Energy* **2022**, *249*, 123676. [CrossRef]
10. Aghel, B.; Yahya, S.I.; Rezaei, A.; Alobaid, F. A dynamic recurrent neural network for predicting higher heating value of biomass. *Int. J. Mol. Sci.* **2023**, *24*, 5780. [CrossRef]
11. Güleç, F.; Pekaslan, D.; Williams, O.; Lester, E. Predictability of higher heating value of biomass feedstocks via proximate and ultimate analyses–A comprehensive study of artificial neural network applications. *Fuel* **2022**, *320*, 123944. [CrossRef]
12. Jakšić, O.; Jakšić, Z.; Guha, K.; Silva, A.G.; Laskar, N.M. Comparing artificial neural network algorithms for prediction of higher heating value for different types of biomass. *Soft Comput. Fusion Found. Methodol. Appl.* **2023**, *27*, 5933–5950. [CrossRef]
13. Ghugare, S.B.; Tiwary, S.; Elangovan, V.; Tambe, S.S. Prediction of higher heating value of solid biomass fuels using artificial intelligence formalisms. *Bioenergy Res.* **2014**, *7*, 681–692. [CrossRef]
14. Veza, I.; Panchal, H.; Paristiawan, P.A.; Idris, M.; Fattah, I.R.; Putra, N.R.; Silambarasan, R. Improved prediction accuracy of biomass heating value using proximate analysis with various ANN training algorithms. *Results Eng.* **2022**, *16*, 100688. [CrossRef]
15. Kandpal, S.; Tagade, A.; Sawarkar, A.N. Critical insights into ensemble learning with decision trees for the prediction of biochar yield and higher heating value from pyrolysis of biomass. *Bioresour. Technol.* **2024**, *411*, 131321. [CrossRef]
16. Hu, Z.J.; Su, R.; Ling, K.V.; Guo, Y.; Ma, R.J. Resilient event-triggered MPC for load frequency regulation with wind turbines under false data injection attacks. *IEEE Trans. Autom. Sci. Eng.* **2024**, *21*, 7073–7083. [CrossRef]
17. Hu, Z.J.; Ma, R.J. Adaptive event-triggered tracking control via switching functions. *Automatica* **2026**, *185*, 112813. [CrossRef]

18. Yin, C.Y. Prediction of higher heating values of biomass from proximate and ultimate analyses. *Fuel* **2011**, *90*, 1128–1132. [CrossRef]
19. Esteves, B.; Sen, U.; Pereira, H. Influence of chemical composition on heating value of biomass: A review and bibliometric analysis. *Energies* **2023**, *16*, 4226. [CrossRef]
20. Friedl, A.; Padouvas, E.; Rotter, H.; Varmuza, K. Prediction of heating values of biomass fuel from elemental composition. *Anal. Chim. Acta* **2005**, *544*, 191–198. [CrossRef]
21. Channiwala, S.A.; Parikh, P.P. A unified correlation for estimating HHV of solid, liquid and gaseous fuels. *Fuel* **2002**, *81*, 1051–1063. [CrossRef]
22. Sheng, C.; Azevedo, J.L.T. Estimating the higher heating value of biomass fuels from basic analysis data. *Biomass Bioenergy* **2005**, *28*, 499–507. [CrossRef]
23. Jenkins, B.M.; Ebeling, J.M. Correlation of Physical and Chemical Properties of Terrestrial Biomass with Conversion. In Proceedings of the IGT symposium: Energy from Biomass and Waste IX, Lake Buena Vista, FL, USA, 28 January–1 February 1985.
24. Institute of Gas Technology; United States Division of Coal Conversion. *Coal Conversion Systems: Technical Data Book*; US Government Printing Office: Washington, DC, USA, 1978.
25. Beckman, D.; Elliott, D.; Gevert, B.; Hörnell, C.; Kjellström, B.; Östman, A.; Solantausta, Y.; Tulenheimo, V. *Techno-Economic Assessment of Selected Biomass Liquefaction Processes*; Technical Research Centre of Finland: Espoo, Finland, 1990.
26. Wang, R.; Zha, B. A research on the optimal design of BP neural network based on improved GEP. *Int. J. Pattern Recognit. Artif. Intell.* **2019**, *33*, 1959007. [CrossRef]
27. Chen, J.; Dong, C.; He, G.; Zhang, X. A method for indoor Wi-Fi location based on improved back propagation neuralNetwork. *Turk. J. Electr. Eng. Comput. Sci.* **2019**, *27*, 2511–2525. [CrossRef]
28. Ding, S.; Su, C.; Yu, J. An optimizing BP neural network algorithm based on genetic algorithm. *Artif. Intell. Rev.* **2011**, *36*, 153–162. [CrossRef]
29. Greenhill, S.; Rana, S.; Gupta, S.; Vellanki, P.; Venkatesh, S. Bayesian optimization for adaptive experimental design: A review. *IEEE Access* **2020**, *8*, 13937–13948. [CrossRef]
30. Springenberg, J.T.; Klein, A.; Falkner, S.; Hutter, F. Bayesian optimization with robust Bayesian neural networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*.
31. Maraş, A.; Erol, Ç. Emerging Trends in Classification with Imbalanced Datasets: A Bibliometric Analysis of Progression. *Bilişim Teknol. Derg.* **2022**, *15*, 275–288. [CrossRef]
32. Parikh, J.; Channiwala, S.A.; Ghosal, G.K. A correlation for calculating HHV from proximate analysis of solid fuels. *Fuel* **2005**, *84*, 487–494. [CrossRef]
33. Zhao, C.; Lu, X.; Zhang, Y. Prediction of the higher heating value of biomass based on multiple classification methods. *Biomass Convers. Biorefin.* **2024**, *15*, 5183–5191. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.