

Research Article

An Intelligent Deep Learning Model for CO₂ Adsorption Prediction

Hanan Ahmed Hosni Mahmoud ¹, Nada Ali Hakami ² and Alaaeldin M. Hafez³

¹Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

²Jazan University, Computer Science Department, College of Computer Science and Information Technology, Saudi Arabia

³Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Correspondence should be addressed to Nada Ali Hakami; nmhakami@jazanu.edu.sa

Received 12 January 2022; Revised 9 April 2022; Accepted 13 April 2022; Published 10 May 2022

Academic Editor: Muhammad Raziq Rahimi Kooh

Copyright © 2022 Hanan Ahmed Hosni Mahmoud et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we propose a supervised deep learning neural network (D-CNN) approach to predict CO₂ adsorption from the textural and compositional features of biomass porous carbon waste and adsorption features. Both the textural and compositional features of biomass porous carbon waste are utilized as inputs for the D-CNN architecture. A deep learning neural network (D-CNN) is proposed to predict the adsorption rate of CO₂ on zeolites. The adsorbed amount will be classified and predicted by the D-CNN. Three tree machine learning models, namely, gradient decision model (GDM), scalable boosting tree model (SBT), and gradient variant decision tree model (GVD), were fused. A feature importance metric was proposed using feature permutation, and the effect of each feature on the target output variable was investigated. The important extracted features from the three employed model were fused and used as the fusion feature set in our proposed model: fusion matrix deep learning model (FMDL). A dataset of 1400 data items, on adsorbent type and various adsorption pressure, is used as inputs for the D-CNN model. Comparison of the proposed model is done against the three tree models, which utilizes a single training layer. The error measure of the D-CNN and the tree model architectures utilize the mean square error confirming the efficiency of 0.00003 for our model, 0.00062 for the SBT, 0.00091 for the GDM, and 0.00098 for the GVD, after 150 epochs. The produced weight matrix was able to predict the CO₂ adsorption under diverse process settings with high accuracy of 96.4%.

1. Introduction

The introduction of machine and deep learning models started in 1980s [1–5], and machine learning models have been at the front position for intelligent computer models for prediction and classification. Deep learning is a type of machine learning with more depth for feature extraction. Deep learning, with its dependence on feature extraction instead of clear training to achieve optimal solutions, accomplishes high performance in stochastic settings [6] and is thus associated in a way to simulate the inference thinking found in human brains, using supervised learning process [7]. Many scientists have studied machine learning and their mathematical and stochastic nature for numerous applica-

tions. This paper utilizes the mathematical nature of machine learning model in an organic application of CO₂ adsorption prediction. It first studies existing models of deep learning in such discipline.

As an advance of machine learning models, deep learning is distinguished from machine learning because it extracts deep features through unsupervised learning iteratively. Such feature extraction guarantees model independence from human control (human feature engineering). Also, deep learning model accuracy is powered by the existence of adequate data for training and assembles significant information and data correlations [8, 9].

Carbon capture is employed as an essential tool for dropping CO₂ emission rate [10–14], as CO₂ concentration

endures a steady increase. In [10], the authors concluded that process of CO₂ capture is considered an expensive operation, accounting for more than 60% of the total CO₂ capture and storing process cost. In [11], postcombustion CO₂ is considered as the main method for CO₂ capturing from manufacturing emission sources and also considered a cost-effective technique, but due to, its low CO₂ concentration with less than 20%, the key challenge of this operation is to develop of a cost-effective CO₂ capture process. Adsorption operation named regenerative amine solvent operation, for CO₂ capture, is not cost-effective and has high degree of corrosion, degradation loss, and toxicity [12]. Development of nonexpensive membranes with carbon permeability for CO₂ capture from gas is under consideration in the research community. In [13], authors proposed a solid porous carbon adsorption for second-generation carbon dioxide capture process. It is characterized with low cost, regulated pore arrangement, with low energy requirement. Also, biomass waste has encouraging, nonexpensive, and plentiful sources for manufacturing porous carbon adsorbents. Biowaste porous carbons are compounds that are extensively employed in ecological discard control and carbon emission. In [14], Biowaste porous carbon process (BWPC) for CO₂ capture can alleviate the ecological pollution triggered by biowaste management and can attain decarbonization and for climate change diminution.

The CO₂ adsorption of BWPC at various temperatures has been studied for explicating the thermodynamic features of the CO₂ adsorption operations and managing the CO₂ adsorption operation optimization. Thermodynamic features such as entropy and isostatic heat emission specify that CO₂ adsorption on solid adsorbents is controlled by physisorption [11, 12]. CO₂ adsorption improvement can be achieved using carbon selectivity and heteroatom [13, 14].

Deep learning model (DLM) is an intelligent model that undergoes training phase to accomplish classification from input data such as images with high precision. DLM is trained by utilizing supervised learning from labeled dataset [15]. Machine learning and deep learning in many adsorption fields, such as waste-to-energy transfiguration [16–20], compound sorption [21], and biowaste treatment [22, 23] has received intensive interest. Machine learning models comprise linear regression, *k*-nearest neighbors, support vector machines, neural networks, and deep learning models. A deep learning model has more layers than convolutional neural network (CNN) with deeper layers for higher classification accuracy [9]. Tree-based classification models are a type of supervised machine learning model, that utilize recursive data splitting in a binary manner by minimizing the mean sum of squares. Some of the standard models comprise decision trees, gradient decision models, random forest, and gradient boost models. Classification tree-based models are characterized by the capacity to work with small-size datasets, with less overfitting, and noisy counterpart's resistant features [24–27].

In this paper, we propose a data-driven model of mapping the CO₂ adsorption by BWPC based on textural and compositional properties and adsorption parameters. Mapping utilizes the adsorption pressure and temperature con-

duct of the CO₂ adsorption process. The core idea of the proposed research is to expose that machine learning techniques for projecting models and utilized to observe appreciated insights into the CO₂ capture. With this purpose, three tree machine learning methods, namely, gradient decision model (GDM), scalable boosting tree model (SBT), and gradient variant decision tree model (GVD), were employed and validated for the prediction accuracy for CO₂ capture. Using tree models, a feature metric was proposed using feature permutation, and the effect of each feature on the target output variable was investigated. The important extracted features from the three employed model were fused and used as the fusion feature set in our proposed model: fusion matrix deep learning model (FMDL).

Table 1 depicts different gas adsorption simulation deep learning prediction models.

In this paper, CO₂ adsorption data are used as input for the training layers and the adsorption rate is predicted by the D-CNN. Using the D-CNN model, adsorption of CO₂ on zeolite is predicted. This research employs the deep learning CNN to predict the amount of CO₂ adsorption.

This paper is structured as follows: Section 2 proposes the new methodology. Section 3 presents the data collection and statistics and. In Section 4, experimental results are demonstrated. Section 5 depicts a comparative study and result discussion. Section 6 depicts conclusions of the proposed work.

2. Model Description

Feature fusion process enhances neural network prediction performances after getting rid of redundant properties in the datasets. In our proposed model, we initiate feature fusion process and then we enter the learning process followed by the prediction process. A D-CNN with 12 convolutional layers and a dropping layer incorporated with feature fusion procedure. Accuracy of the D-CNN with different weights is computed and ordered so that the D-CNN can pick the attribute of the highest and the least imperative for each run of the D-CNN learning phase. The algorithm recaps itself to eliminate multiple input attributes. When the D-CNN will not reach a sufficient accuracy as defined in the model, the process will halt, and no more features will be removed. Dropout layer is employed to lessen the overfitting of the input training data. Frame work of the proposed model in Figure 1.

2.1. Model Work Frame. As seen in Figure 2, the work frame for the D-CNN model consists of many phase. The first phase is mainly to identify and gather data items based on the following parameters such as textural composition and adsorption parameters. The next phase is to input these data into the D-CNN network training phase. The input variables range between the values -1 and 1 and normalized according to the following equation:

$$D_{\text{norm}} = \frac{2(D - D_{\text{min}})}{(D_{\text{max}} - D_{\text{min}})} - 1. \quad (1)$$

TABLE 1: Summary of gas adsorption simulation deep learning prediction models for different gases using density, molecular, and vacancy volume.

Reference	Gas type	Model	Prediction results
[10]	Xe	Density parameter function	The adsorption ground truth is closer to the experimental results
[28]	CO ₂	Dunning's correlation sets	Moderate accuracy
[29]	CO ₂	Molecular simulations	Moderate accuracy
[22]	CO ₂	Scalable boosting tree model (SBT)	Satisfactory
[17]	CO ₂	The data were linearly correlated by Toth and sips equations	The sip model showed the least deviation
[30]	CO ₂ , He, and Ar	Gradient variant decision tree model (GVD)	Accurate for the adsorbed phase
[31]	CO ₂	The vendor and Langmuir metric	Vendor depicts less deviation than the Langmuir metric from the ground truth
[32]	Kr and N ₂	Vacancy solution method	Yeilds parameter optimization
Our proposed work: Fusion matrix deep learning model (FMDL)	CO ₂	Adsorption modeling using the D-CNN approach	High accuracy

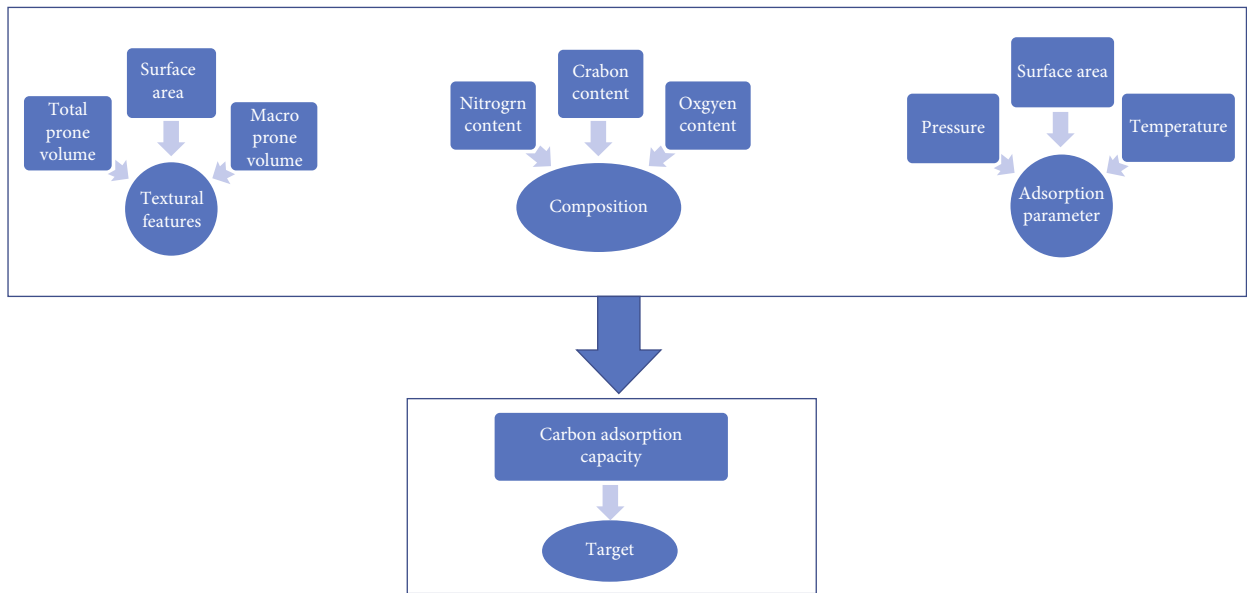


FIGURE 1: The proposed workflow defining model parameters, namely, textural, composition, and adsorption parameters.

The D_{norm} indicates normalized data item, D indicates the raw data, and D_{max} and D_{min} denote the maximum and minimum of D .

The proposed model will tune the D-CNN parameters and weights to enhance prediction performance. In the subsequent phase, the validation data is tuned using a dataset partition, and the verification is then executed using the validated inputs. The ratio of the correlation (R) and the mean square error are employed to evaluate the D-CNN prediction accuracy. The D-CNN model with the highest accuracy is selected and constructed. In various machine learning and regression models, the number of inner layers and the total neurons are accordingly updated. The training model halts when the optimal D-CNN reaches the best assessment for the parameters. The mean square error and R are employed

as valuation parameters, and the outputs are matched to the data set to decide on the best network.

The model simulation was implemented in the 32-bit Windows-10, using Python 3.6. The experiment was implemented using Keras on the Tensor platform. The model training was also performed on the tensor platform with Keras. The system was programmed using Intel® Core™ i7-7300 CPU @5.60GHz with 16M Cache, 32 GB RAM, and GTX 1070 video card. The model was implemented in Python 3.6, employing the Sickie-learn function library [28].

3. Data Collection and Statistics

3.1. Data Collection. Data collection was conducted by collecting data from literature review on BWPC for carbon

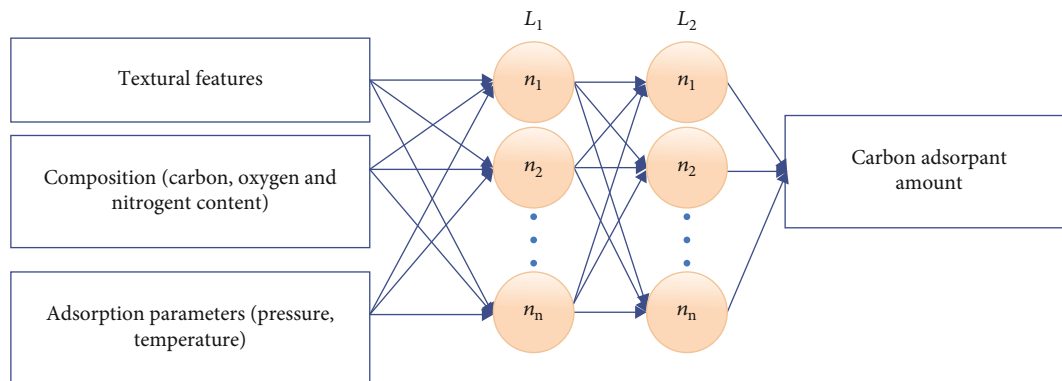


FIGURE 2: The deep learning model consists of many phases. The first phase is mainly to identify and gather data items based on the following parameters such as textural composition and adsorption parameters. The next phase is to input these data into the network training phase.

capture utilizing keywords such as biomass, porous carbon, waste, CO₂ adsorption, and CO₂ capture from indexed databases. 1400 data items were collected and utilized in our research [7, 8]. The common features for the dataset of CO₂ adsorbents BWPC incorporate adsorption capacity, cost-effectiveness, and adsorbent selectivity adsorption kinetics. In our research, we focus on the CO₂ adsorption capability achieved at various temperatures and air pressures versus the textural features and composition.

The data collection presumes the following assumptions:

- (1) Screened data were accepted, without bias towards data validity
- (2) The data were acquired from experiments performed by scientists. Data items that are not listed directly were extracted from listed figures utilizing Plot-Digitizer web crawler to extract the required data and cleaned to circumvent duplicates
- (3) The selected input features extracted and categorized into classified into three classes: (1) texture features, (2) BWPC compositions, and (3) adsorption properties such as temperature at which CO₂ adsorption were done
- (4) The primary texture features of the BWPC incorporate the surface area and the pore volume
- (5) The secondary texture features include macropore area and volume and the weight content of CO₂
- (6) The CO₂ uptake rate at various adsorption properties was utilized as the target optimization variable

3.2. Data Preprocessing. The collected data will be transferred into predefined units. At the data cleaning phase, missed data were found to be mostly total pore and macro pore volume. This missing data is due to the variance in the data published and the selection of textural features stated. Most papers described the surface area and total pore and macropore volumes. Therefore, many data items are either area or pore volume, while macropore volume was

not listed, and in alternate cases, the area and macropore volume were listed, but total pore volume was not listed. Data cleaning is very crucial for missing data reproach of total pore volume and macropore volume using machine learning. Data cleaning techniques are performed to evade removal of tuples with missing attributes.

The linear correlation between the inputs is formulated using Pearson coefficient (ψ) as depicted in the following equation:

$$\psi = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{SD_x} \right) \left(\frac{y_i - \bar{y}}{SD_y} \right), \quad (2)$$

where ψ is the Pearson coefficient for the feature x and the target y , whereas \bar{x} and \bar{y} are the average values of the input x and output y , respectively. The value of ψ is in the range $\{-1, 1\}$.

3.3. The Collected Dataset (DS). The processed 1400 data items are exposed to several training phases by dividing the original dataset into random training set and another test set. 80% of the data items are labeled and utilized for the learning data, and the 20% partition are used as testing for the supervised models. Crossvalidation process is utilized to tune the parameters to concurrently enhance the model's prediction accuracy using k -fold [19]. In the k -fold validation process, the data will be partitioned into k -folds to be used for every training iteration, while one of them will be employed for validation. This aids in solving the overfitting challenge to solve the bias problem in machine learning model. We utilized different values of k (5, 7, 10, 12) and tuned the model to the best k value as depicted in Table 2. $K=7$ is found to be the best one in validation. The training dataset has 1120 data items (80% of 1400 total items). Therefore, using $k=7$ generates 8 partitions of 160 data items (Table 2).

3.4. Parameter Tuning. Three tree machine learning models were employed and fused to predict CO₂ adsorption on BWPCs. Recent research [34-39] have revealed the fitness

TABLE 2: Crossvalidation testing for $k = 5, 7, 10,$ and 12 -folds to determine the best performance k -validation value.

Testing method	Accuracy	Precision	Recall	F1-score
5-fold	93.60%	96.61%	93.16%	96.86%
7-fold	96.26%	95.31%	96.26%	96.01%
10-fold	91.35%	91.30%	92.62%	91.86%
12-fold	89.98%	89.96%	89.82%	89.95%

of tree machine learning models in small size datasets with less than 1200 data items when the number of selected features is between 5 and 20 features. Such data are usually collected from experimental research for midsize data.

The gradient decision model (GDM) is an ensemble learning model that syndicates several connected sequence decision trees [30]. Decision trees are considered weak learners. But with incorporation of gradient decision models, by totaling sequence decision trees in an ensemble will induce learning boosting. Each sequence decision subtree optimizes the solutions from the previous subtree, and the boosting algorithm induces high efficiency of the gradient decision model. The scalable boosting tree model (SBT) is a scalable decision tree with a trailing gradient decision algorithm, using many decision trees and biased quantile search to execute distributed computing. GVD is gradient variant decision tree model that utilizes fast feature bundling to enhance efficiency without compromising the accuracy. Then, as a summary, the three employed tree models are as follows: the gradient decision model (GDM), the scalable boosting tree model (SBT), and gradient variant decision tree model (GVD).

Parameter adjusting is the procedure of extracting parameters to attenuate the performance. Parameters are usually tuned in the training phase. Parameter tunings are performed using grid search and Bayesian inference [31]. In our research, multiple parameter tuning procedures are performed and tested, and then best parameters are selected with the highest accuracy. In the propose model, the grid search technique was performed for parameter tuning yielding a small set of input features (seven features).

3.4.1. Metrics. The performance metrics of the regression techniques are mostly use linear dependency metric (R) and the mean absolute error (MAE) [32, 33]. The greater the R and the lesser the MAE, the higher the accuracy, as depicted in the following equations:

$$R = -\frac{1/N \left(\sum_{i=0}^M |t_i - p_i| \right)}{\sum_{i=0}^M |p_i - m|}, \quad (3)$$

$$\text{MAE} = \frac{1}{N} \sum_i |t_i - p_i|,$$

where p_i and t_i are the predicted and the ground truth, respectively. M is the number of various classes. M is the average values of the predicted classes, and M is the number of the data items.

3.4.2. Feature Importance. Decision tree models face the challenge of computing the significance of an input feature and its effect on the output accuracy. Permutation mean decline accuracy (PMD) is a method utilized to identify the significance of each feature for tree prediction models by computing the changes in the model prediction performance when an input feature is used or not used. PMD is the average drop in the Gini score, which computes the contribution of each independent variable to the homogeneity of the tree nodes [34]. The greater the value of the PMD score, the significance of the input in the prediction. The model is fit for identifying the significance when the permutation count of the features is reasonable or the computation will be resource and time intensive. Our proposed model utilizes seven input features; the permutation model is utilized to compute the feature ensemble importance. Permutation significance accuracy can identify the respective importance of an input feature to the prediction model as a whole and the effect of the input on the target dependent variable. The PMD is a regression formula that lowers the effect of the inputs on the prediction of the deep learning model, with the exception of the single input of interest, thus denoting sensitivity analysis. Thus, by disregarding the impact of one independent input feature is attained. In our model, each single feature is used at an instance of time to measure its impact on the target output for each data item in the dataset. We should note that the GDM utilizes global and local estimates, and local sensitivity study from the PMD was utilized in this research. PMD allows the effect of the inputs on the output variables by computing their impact across individual data items.

3.4.3. Data Statistics. Analysis of the features and the target dependent variable is performed using raw data; for instance, the dataset consists of 1400 data items [7, 8] that were analyzed statistically, realizing the minimum and maximum values as well as the mean of the input values. The target dependent variable is also analyzed to gain insights. Figures 3–5 depict the representation of the data statistics for the inputs and the target output. The average values of CO_2 adsorbed on the porous biomass for the data composed was $3.38 \text{ mmol/g} \pm 1.58 \text{ mmol/g}$, with CO_2 adsorption of maximum 8.23 mmol/g at 0°C and 1 bar, and minimum of 0.24 mmol/g at 0.13 bar. The surface area is the texture feature that was described for the dataset. The reported surface area (SA) varies from 800 to $2436 \text{ m}^2/\text{g}$, with an average of $1442.70 \text{ m}^2/\text{g} \pm 755 \text{ m}^2/\text{g}$. The average values for pore volume (PV) are 0.77 with standard deviation 0.46 while macropore volume (MV) has an average of 0.52 with standard deviation of $0.49 \text{ cm}^3/\text{g}$. The results summaries are depicted in Figures 3–5. The SA and PV were considerably impacted by the carbonation treatments. The authors in [35], produced coconut shell porous for CO_2 adsorption using KOH galvanization. The outcomes depicted that the SA ranges from 880 to $2687 \text{ m}^2/\text{g}$, and the pore volume ranges from 0.378 to $1.329 \text{ cm}^3/\text{g}$. The greatest CO_2 uptake of 4.257 mmol/g at 26°C and 1 bar was produced with SA of $1478 \text{ m}^2/\text{g}$ and a pore volume of $0.67 \text{ cm}^3/\text{g}$. This results

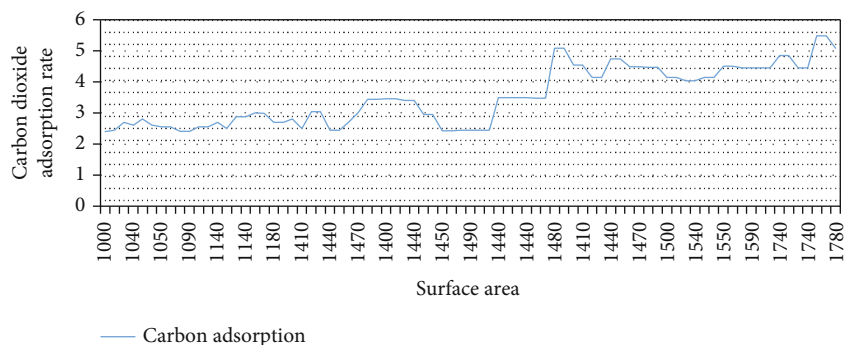


FIGURE 3: Representation of the data statistics for the input (CO_2 adsorption) versus the target output which is the surface area (data is collected from the dataset in [7, 8]).

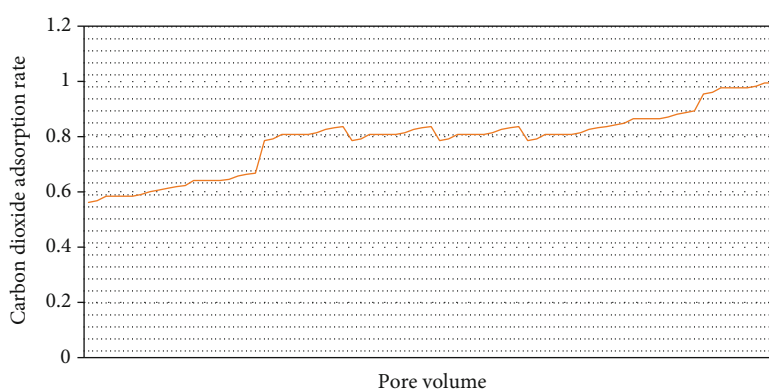


FIGURE 4: Representation of the data statistics for the input (CO_2 adsorption) versus the target output (pores volume) (data is collected from the dataset in [7, 8]).

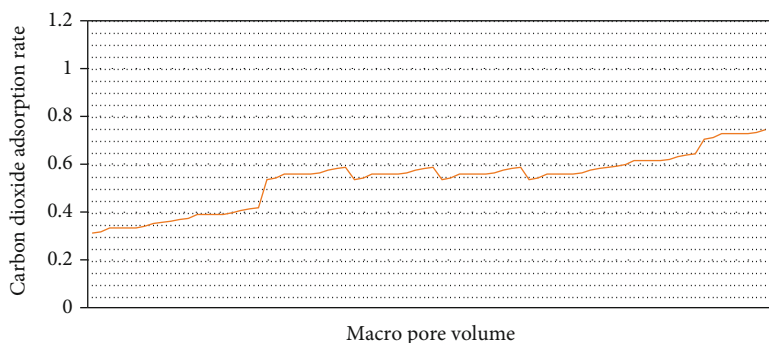


FIGURE 5: Representation of the data statistics for the input (CO_2 adsorption) versus the target output (macro pores volume) (data is collected from the dataset in [7, 8]).

infer that no simple path was available to produce the best porous carbons for CO_2 capture from biomass wastes.

4. Experiments

The three different feature extraction algorithms that were utilized to get best features are the gradient decision models, random forest, and gradient boost models, and the texture type, composition, and adsorption parameters are utilized as input data. The amount of adsorbed CO_2 is the target output data. The experimental dataset (1400 data items) was

distributed into training subset (85%) and validation subset (15%) at random. By measuring different parameters, the D-CNN activation functions are reached. The CPU utilized in the executing the experiment is an Intel i7-8200 CPU @ 3.30 GHz, and the memory is 16.00 GB. We used MATLAB R3012a. The experiments were executed 200 runs, and we took the average for the results.

4.1. Neuron Selection. A number of activation functions have been employed in this research. Sigmoid function is utilized for the hidden layers, and Purlin transfer is utilized for the

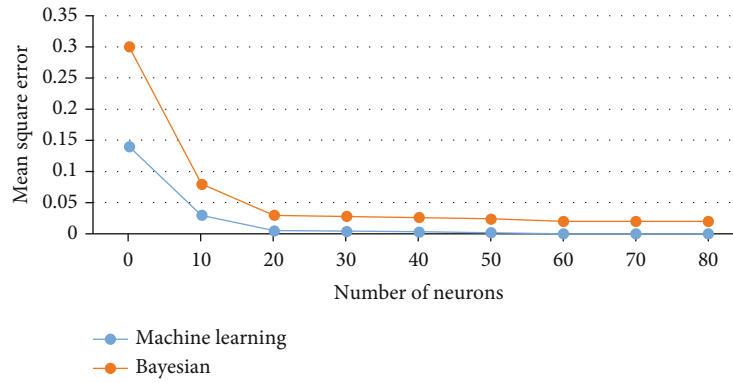
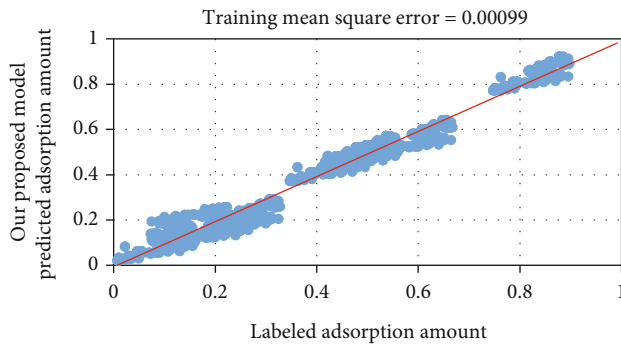


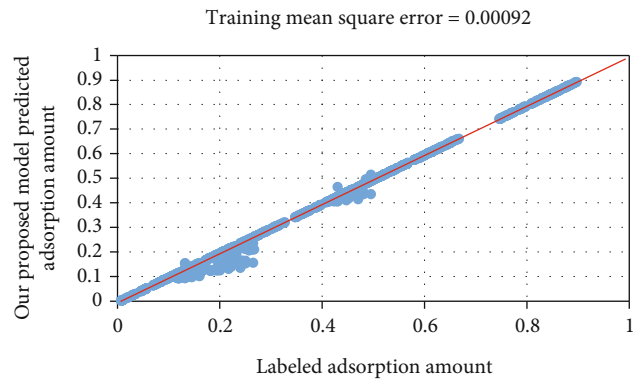
FIGURE 6: The optimum number of neurons is estimated using the least mean square error function value in the range from 1 to 80.

TABLE 3: Hidden layers' number against the mean square error and the correlation factors for different models, where the adsorption parameters (pressure and temperature) are selected using the Bayesian ordered function with the least mean square error values.

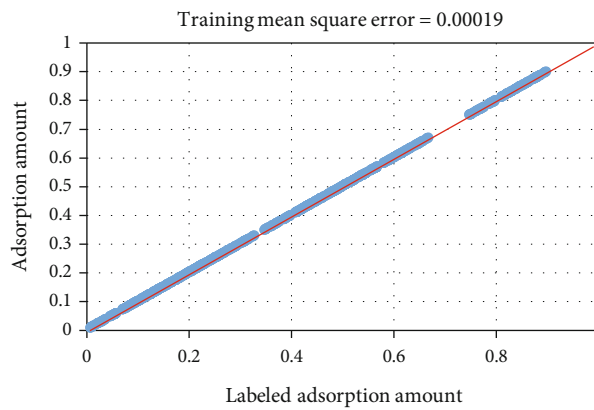
Model	Mean square error	Correlation	Epoch	Hidden layer number	Optimal number of neurons
Machine learning	0.000053	0.9898	62	[20]	25
Bayesian	0.000047	0.9798	100	[15]	65



(a) Correlation between the prediction of our model and the labeled adsorption data for mean square error of 0.00099



(b) Correlation between the prediction of our model and the labeled adsorption data for mean square error of 0.00092



(c) Correlation between the prediction of our model and the labeled adsorption data for mean square error of 0.00019

FIGURE 7: Correlation between the prediction of our model and the benchmark data for different mean square error versus the ground truth (labeled adsorption amount).

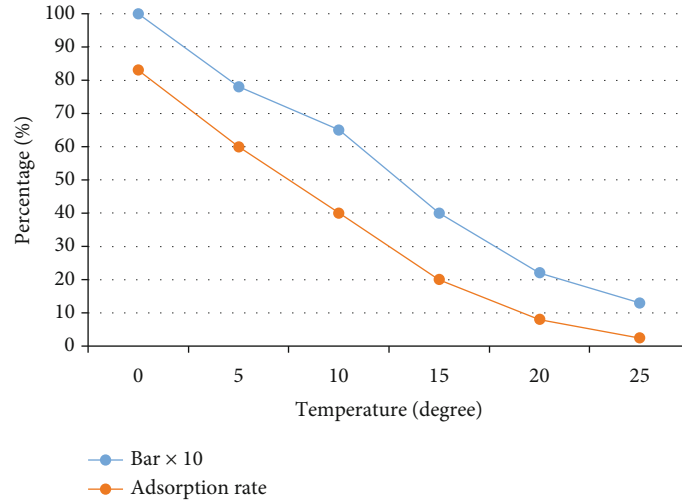


FIGURE 8: Adsorption rate and pressure versus temperature, with average surface area of $1442.70 \text{ m}^2/\text{g}$. The average values for pore volume (PV) are 0.77 , and macropore volume (MV) has an average of $0.52 \text{ cm}^3/\text{g}$.

TABLE 4: Pearson's coefficient matrix for all features included in the research. Insignificant correlation was found between the inputs, only textural features (surface area, pore volume, and macropore volume). The features with units are depicted as follows: surface area: (m^2/g), pore volume: (cm^3/g), macropore volume: (cm^3/g), H: hydrogen, N: nitrogen, O_2 : oxygen, and CO_2 adsorption (mmol/g).

	Surface area	Pore volume	Macropore volume	O_2	N	H	CO_2
Surface area	1	0.93	0.83	0.38	0.21	0.012	0.33
Pore volume	0.93	1	0.77	0.33	0.013	0.014	0.013
Macropore volume	0.83	0.77	1	0.29	0.31	0.22	0.31
O_2	0.38	0.34	0.31	1	0.12	0.31	0.12
N	0.21	0.19	0.17	0.15	1	0.25	0.014
H	0.012	0.013	0.011	0.011	0.014	1	0.22
CO_2	0.33	0.30	0.28	0.26	0.22	0.31	1

TABLE 5: Evaluation of the tree machine learning models using the benchmark data, the gradient decision model (GDM), the scalable boosting tree model (SBT), and gradient variant decision tree model (GVD) and our proposed model fusion matrix deep learning model (FMDL).

	Gradient decision model	Scalable boosting tree model	Gradient variant decision tree model	Fusion matrix deep learning model (FMDL)
Training correlation	0.97	0.91	0.94	0.97
7-fold correlation	0.78	0.77	0.78	0.82
Testing correlation	0.74	0.77	0.75	0.81
Mean absolute error (mmol/g)	0.67	0.77	0.64	0.52

output layers. The least mean square error function value and the maximum correlation (R) value are utilized to compute the optimum number of neurons that ranges from 1 to 80, as depicted in Figure 6. The adsorption parameters (pressure and temperature) are selected using the Bayesian ordered function with the least mean square error values as depicted in Table 3.

4.2. *D-CNN Architecture for Adsorption.* Both the machine learning and the Bayesian regression models have several hidden layers of 20 and 15 neurons and one layer of 65 neu-

rons. The D-CNN consists of input, hidden, and output layers. The input layer utilizes the input data for adsorption parameters and other inputs. The hidden layers are selected by the precision required. The hidden layers in this research are set to three layers which achieve suitable precision. The neurons in the hidden layers are set to 15 or 20. The Sigmoid function is employed in these layers. The output layer utilizes a linear transfer function, namely, Purlin.

The R parameter for the D-CNN model is depicted in Figure 7 and is nearly one ($R = 0.9989$). Figure 7 exhibits that the D-CNN model outputs and the labeled adsorption

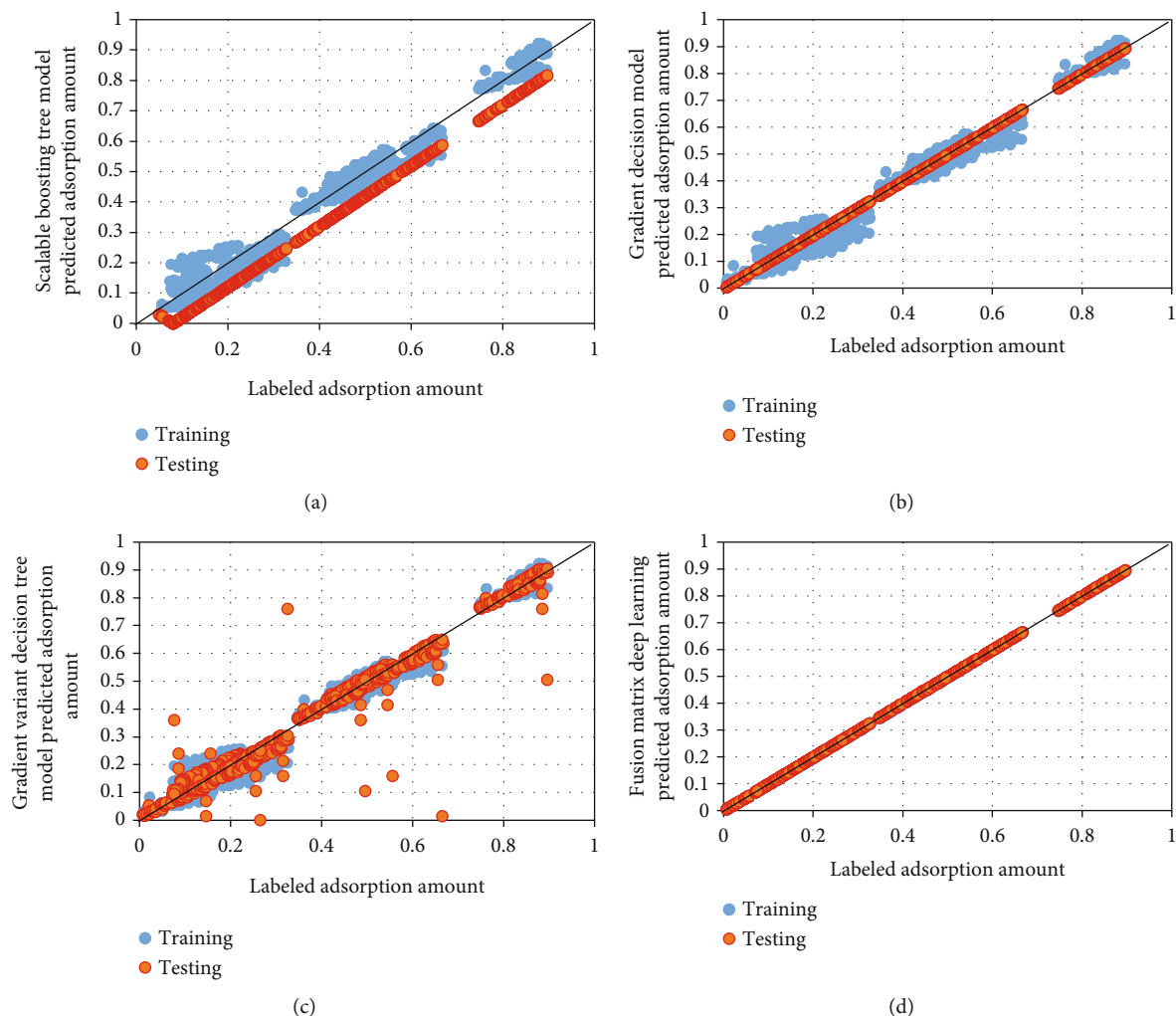


FIGURE 9: Prediction using scalable boosting tree model (a), gradient decision model (b), gradient variant decision tree model (c), and our proposed model Fusion matrix deep learning model (d). The red color represents 96% confidence of the regression for the test items. The black line is the equality line ($y = x$). The prediction accuracy of the models followed comparable trends but with diffused accuracy.

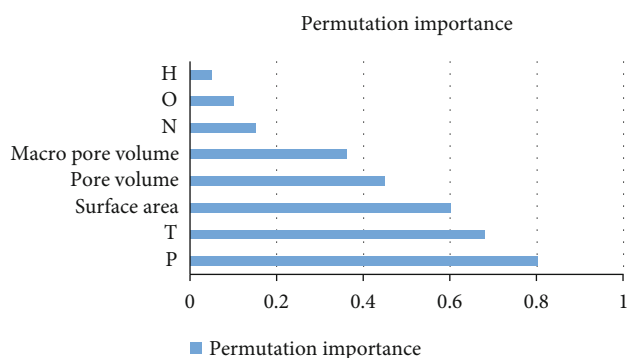


FIGURE 10: Feature importance analysis where the impact of each factor on the target output. The experiment illustrates the permutation importance on the output.

amount from the benchmark dataset have exactly value of correlation.

The machine learning model with the Bayesian regression technique for the adsorption procedure is fitted to normalize the impact. These results indicate that the predicted D-CNN model fit the labeled adsorption dataset concisely. The regression correlation coefficient (R) of the proposed model is 0.99989 and 0.99784, respectively. We can denote that the developed D-CNN offers a prediction accuracy that is consistent with the benchmark labeled dataset. Based on the experimental results, the proposed model avoided being caught in local optima by altering the radial basis range function. The results have also shown that the RBF model can perform the same functions as the MLP model in most datasets.

To study the relationship of the adsorption parameters (i. e., pressure, temperature, and adsorbents amount) and to identify each variable impact on the adsorption amount, charts of a response surface for the proposed model predicted

TABLE 6: Recall, precision, and F -measure metric comparison.

Models	Results			Enhancement of our proposed regarding other adsorption prediction models		
	Recall	Precision	F -measure	Recall	Precision	F -measure
Model 1 scikit machine learning model [36]	0.91	0.92	0.91	+5.59%	+6.52%	+6.59%
Model 2 interpretable machine learning model [37]	0.92	0.91	0.92	+4.35%	+7.67%	+6.59%
Model 3 light gradient boosting machine [31]	0.92	0.915	0.93	+4.35%	+7.10%	+4.30%
Our proposed model	0.96	0.98	0.97	—	—	—

versus the labeled benchmark data are depicted in Figure 8. The average values of CO_2 adsorbed on the porous biomass for the data composed were $3.38 \text{ mmol/g} \pm 1.58 \text{ mmol/g}$, with CO_2 adsorption of maximum 8.23 mmol/g at 0°C and 1 bar, and minimum of 0.24 mmol/g at 25°C and 0.13 bar. The surface area is the texture feature that was described for the dataset. The reported surface area (SA) varies from 800 to $2436 \text{ m}^2/\text{g}$, with an average of $1442.70 \text{ m}^2/\text{g} \pm 755 \text{ m}^2/\text{g}$. The average values for pore volume (PV) are 0.77 with standard deviation 0.46 while macropore volume (MV) has an average of 0.52 with standard deviation of $0.49 \text{ cm}^3/\text{g}$. The result summaries are depicted in Figure 8. The SA and PV were considerably impacted by the carbonation treatments. The authors in [38-41] produced coconut shell porous for CO_2 adsorption using KOH galvanization. The outcomes depicted that the SA ranges from 880 to $2687 \text{ m}^2/\text{g}$ and the pores volume ranges from 0.378 to $1.329 \text{ cm}^3/\text{g}$. The greatest CO_2 uptake of 4.257 mmol/g at 26°C and 1 bar was produced with SA of $1478 \text{ m}^2/\text{g}$ and a pore volume of $0.67 \text{ cm}^3/\text{g}$. This results infer that no simple path was available to produce the best porous carbons for CO_2 capture from biomass wastes.

The proposed model underwent experiments and the results are summarized using the optimized D-CNN weights for adsorption rate prediction. The results are summarized in Table 4.

In the linear dependency correlation between the input independent variables, a high positive correlation was perceived in the textural features including the surface area, prone, and macropore volume. The Pearson coefficient for those variables is greater than 0.736 , indicating a high correlation. Only textural properties have high correlation but no substantial correlation was perceived for the other input variables with Pearson coefficient between -0.5 and 0.5 . Table 3 represents the Pearson coefficient matrix. The shortage of a correlation between the inputs aided in maintaining all of them for constructing the prediction method, as every single feature contributes independently to the prediction. There is a high correlation between the textural features, and this set of inputs contained a large portion of the missing information in the raw data.

The overfitting is countered as the 7-fold validation is reached in terms of R and thus enhancing model generalization as depicted in Table 5.

Figure 9 displays the joint plots of the actual cases versus predicted cases of CO_2 adsorption, as computed by

the three tree models. Although the GDM and SBT presented analogous performances in the training and cross-validation, GDM outperformed SBT within higher test R at 0.85 and 0.78 and lesser MAE at 0.63 and 0.70 . These experimental results prove that all the models have similar performances. In general, the GDM outperformed the SBT and GVD with less overfitting yielding the minimum R , proving the generalization competence. Our fusion matrix deep learning model outperforms the three models when they act separately.

4.3. Feature Analysis. The permutation mean decline accuracy (PMD) is utilized to define the impact of the inputs, which contained the compositions and textural features versus the adsorption parameters on the output target (CO_2 adsorption rate). This study was performed for the GDM model, which is indicated as the best model in performance in our research. Figure 10 depicts the impact of each input on the target output. The experiment illustrates the permutation importance for each independent variable on the output. A high permutation importance value for the input indicates a reduction in the model accuracy when the factor is not employed. Thus, a factor with a high permutation importance specifies a weighty impact on the accuracy.

5. Comparison Study and Discussion

5.1. Comparison Study. We conducted a comparative study of our model versus similar machine and deep learning published models. We completed the comparison on the collected dataset (DS). The comparison is portrayed in Table 6 of the recall, precision, and F -measure metrics. The results designate that our adsorption prediction model outperforms other models in adsorption precision. Our system demonstrated enhancement of 6.59% , 6.59% , and 4.3% , in the F -measure w.r.t. the other state of the art adsorption prediction model. It is found that our model's performance is greater than peer models.

Table 7 demonstrates the statistical metrics for the compared adsorption prediction models. Table 8 portrays the confusion matrix of the accuracy, specificity, and sensitivity for the compared adsorption prediction models for three adsorption states (high, moderate, and low) for temperature 25°C and pressure of 1 bar.

5.2. Discussion. In the experimental study, we applied accuracy, specificity, and sensitivity metrics. The results

TABLE 7: Statistic metrics for the compared model with and without feature fusion.

	Model 1 [36] without the feature fusion	Model 1 [36] with feature fusion	Model 2 [37] without feature fusion	Model 2 [37] with feature fusion	Model 3 [31] without feature fusion	Model 3 [31] with feature fusion	Our proposed model with future fusion
Correctly classified	0.712	0.92	0.646	0.91	0.747	0.915	0.98
Incorrectly classified	0.288	0.08	0.364	0.09	0.374	0.085	0.02
Kappa coefficient (interqualitative reliability)	0.201	0.315	0.166	0.309	0.177	0.411	0.514
Mean square error	0.521	0.441	0.583	0.412	0.583	0.413	0.211

TABLE 8: Confusion matrix for the accuracy, specificity, and sensitivity for the compared adsorption prediction models for three adsorption states (high, moderate, and low) for temperature 25°C and pressure of 1 bar.

Model	Precision	Precision	Recall	<i>F</i> -measure	ROC	Category
Model 1 [36]	0.812	0.812	0.741	0.811	0.831	High adsorption
	0.732	0.732	0.763	0.733	0.733	Moderate adsorption
	0.712	0.712	0.741	0.711	0.791	Low adsorption
Model 2 [37]	0.812	0.812	0.741	0.811	0.831	High adsorption
	0.732	0.732	0.763	0.733	0.733	Moderate adsorption
	0.712	0.712	0.741	0.711	0.791	Low adsorption
Model 3 [31]	0.812	0.812	0.841	0.811	0.871	High adsorption
	0.732	0.732	0.763	0.733	0.733	Moderate adsorption
	0.612	0.612	0.641	0.611	0.671	Low adsorption
Our proposed model	0.912	0.912	0.941	0.911	0.971	High adsorption
	0.972	0.972	0.967	0.957	0.987	Moderate adsorption
	0.912	0.912	0.941	0.911	0.991	Low adsorption

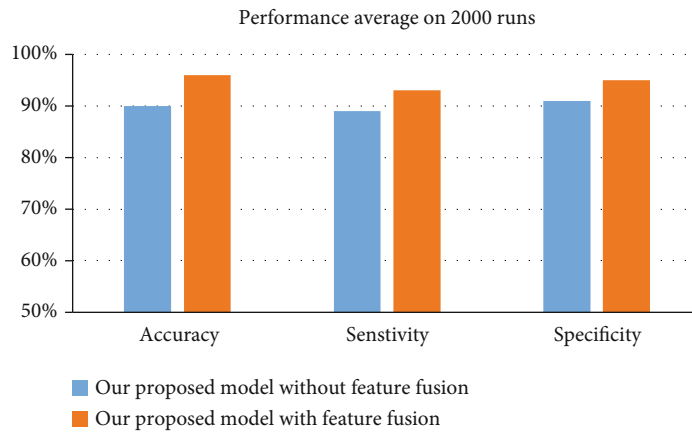


FIGURE 11: Accuracy results for proper feature fusion.

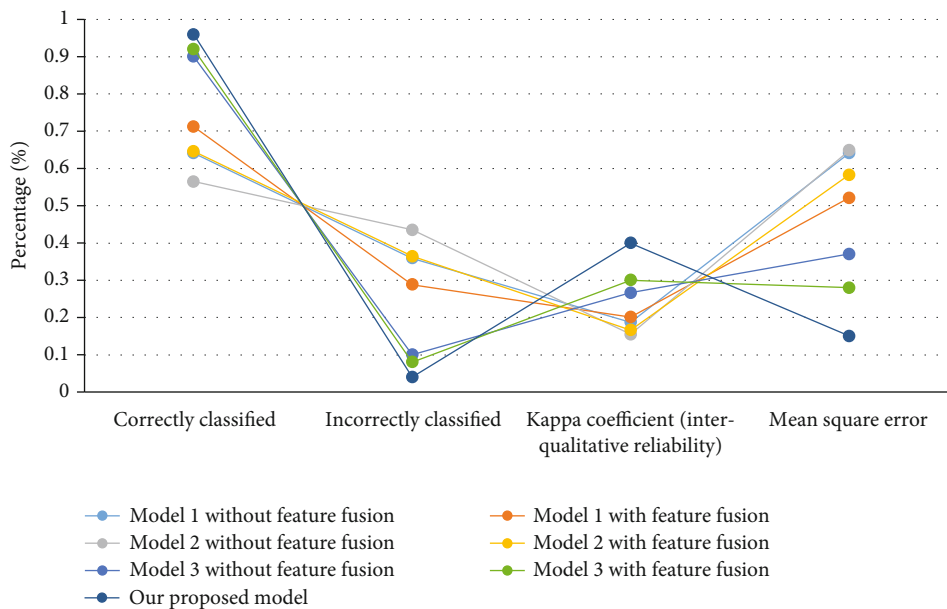


FIGURE 12: Correctly classified versus incorrectly classified case comparison for different models with and without feature fusion.

TABLE 9: Comparison of the prediction time in seconds with the same training dataset.

Method	Execution time (sec)
Our proposed model	4.42×10^{-3}
Model 1 without feature fusion	$12.05 \times 10^{-1} \pm 8.09 \times 10^{-3}$
Model 1 with feature fusion	$7.05 \times 10^{-1} \pm 4.09 \times 10^{-3}$
Model 2 without feature fusion	$17.31 \times 10^{+1} \pm 2.03 \times 10^{-1}$
Model 2 with feature fusion	$12.42 \times 10^{+1} \pm 7.19 \times 10^{-3}$
Model 3 without feature fusion	$117.42 \times 10^{+1} \pm 12.09 \times 10^{-1}$
Model 3 with feature fusion	$67.17 \times 10^{+1} \pm 7.09 \times 10^{-1}$

demonstrate that by employing feature fusion, all prediction models including our proposed model were more able in terms of accuracy with respect to the same classifiers without feature fusion. The best accuracy level for adsorption detection was attained by our proposed D-CNN classifier, which gained 98% accuracy outperforming other classifiers by about 6%.

The experiments with feature fusion designated that suitable feature space fusion can enhance the results by a realistic margin. Accuracy outcomes of these cases are represented in Figure 11.

Figure 12 exhibits correctly predicted cases and incorrectly predicted cases. The results display upgrading with feature fusion. The mean square error stayed less considerably when using feature fusion. The Kappa metric for the all compared prediction model also was better when feature fusion was incorporated. This suggests that feature fusion upsurges accuracy because it fuses all relevant features. Of all the compared models, our model attained the maximum improvement with feature fusion.

Table 9 depicts the comparison study of execution time of classifying CO₂ adsorption with the same training dataset. Our model with feature fusion has the least prediction time (in contrary to training time because of more features incorporated in training). Model 1 is the next in prediction time with feature fusion still slower than our model by an order 2.

6. Conclusions

In this paper, we employed a supervised deep learning model for CO₂ adsorption prediction from fused adsorption features. Both the textural and compositional features of biomass porous carbon waste are utilized as inputs for the D-CNN architecture. The deep learning neural network (D-CNN) predicts the adsorption rate of CO₂ on zeolites. 1400 data items of different adsorbent rates and adsorption pressure is built and used as inputs for the D-CNN model. The adsorbed rate will be classified and predicted by the D-CNN. The correlation (*R*) for the deep learning model and Bayesian model models was 0.9998 and 0.9978, respectively. The produced weight matrix was able to predict the CO₂ adsorption under diverse process settings with high accuracy of 96.4%.

The permutation importance of fused features yields the following observations: the pressure and temperature are the weightiest parameters impacting the model prediction accuracy. The textural features (SA, TV, MV) in the order of declining precedence and the last important are the compositional factors on the feature importance analysis, and the significance of the features was individually observed and categorized into three classes, with the adsorption being the utmost important one.

A comparison of our model versus deep learning published models (recall, precision and F-measure, execution time) was performed. The comparison results indicate that our adsorption prediction D-CNN model is better than other models in adsorption precision. Our system demonstrated enhancement of 6.59%, 6.59%, and 4.3%, in the *F*-measure versus the other models. Our model exhibits fast computational time with an average execution time of 4.2 seconds which is better than all other models by half the time at least.

These results indicate that adsorption parameters highly impacted the CO₂ adsorption rate. For example, we will find a decrease in the CO₂ adsorption with the more temperature and less pressure.

Data Availability

The data presented in this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflict of interest to report regarding the present study.

Acknowledgments

This research was funded by Princess Nourah Bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R113), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

- [1] M. Jahandar Lashaki, S. Khiavi, and Sayari, "Leveraging machine learning to optimize CO₂ adsorption," *Chemical Society Review*, vol. 48, p. 8820, 2019.
- [2] S. P. Soniya and L. Singh, "A review on advances in deep learning," in *2015 IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions (WCI)*, pp. 1–6, Paris, France, 2015.
- [3] J. Li, L. Pan, M. Suvarna, Y. W. Tong, and X. Wang, "Fuel properties of hydrochar and pyrochar: prediction and exploration with machine learning," *Applied Energy*, vol. 269, article 115166, 2020.
- [4] J. Li, X. Zhu, Y. Li, Y. W. Tong, Y. S. Ok, and X. Wang, "Multi-task prediction and optimization of hydrochar properties from high-moisture municipal solid waste: application of machine learning on waste-to-resource," *Journal of Cleaner Production*, vol. 278, article 128928, 2021.

- [5] X. Zhu, X. Wang, and Y. S. Ok, "The application of machine learning methods for prediction of metal sorption onto biochars," *Journal of Hazardous Materials*, vol. 378, article 120727, 2019.
- [6] K. Zhang, S. Zhong, and H. Zhang, "Predicting aqueous adsorption of organic compounds onto biochars, carbon nanotubes, granular activated carbons, and resins with machine learning," *Environmental Science & Technology*, vol. 54, no. 11, pp. 7008–7018, 2020.
- [7] H.-n. Guo, S.-b. Wu, Y.-j. Tian, J. Zhang, and H.-t. Liu, "Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: a review," *Bioresource Technology*, vol. 319, article 124114, 2021.
- [8] D. Cha, S. Park, M. S. Kim et al., "Prediction of oxidant exposures and Micropollutant abatement during ozonation using a machine learning method," *Environmental Science & Technology*, vol. 55, no. 1, pp. 709–718, 2021.
- [9] T. Chunfang, W. Lindan, C. Qian, and X. Zhu, "Study of Adsorption Characteristics of Activated Carbon to Chromium from Aquatic Solution," in *2011 International Conference on Computer Distributed Control and Intelligent Environmental Monitoring*, pp. 1323–1326, 2011.
- [10] J. Wang, L. Huang, R. Yang et al., "Recent advances in solid sorbents for CO₂ emission and new development trends," *Energy & Environmental Science*, vol. 7, pp. 8478–8518, 2014.
- [11] A. E. Creamer and B. Gao, "Carbon-based adsorbents for post-combustion CO₂ capture: a critical review," *Environmental Science & Technology*, vol. 50, no. 14, pp. 7276–7289, 2016.
- [12] G. Singh, K. S. Lakhi, S. Sil et al., "Biomass derived porous carbon for CO₂ capture," *Carbon*, vol. 148, pp. 164–186, 2019.
- [13] V. K. Singh and E. Anil Kumar, "Measurement and analysis of adsorption isotherms of CO₂ on activated carbon," *Applied Thermal Engineering*, vol. 97, pp. 77–86, 2016.
- [14] W.-J. Liu, H. Jiang, and H.-Q. Yu, "Emerging applications of biochar-based materials for energy storage and conversion," *Energy & Environmental Science*, vol. 12, no. 6, pp. 1751–1779, 2019.
- [15] X. Yuan, S. Li, S. Jeon, S. Deng, L. Zhao, and K. B. Lee, "Valorization of waste polyethylene terephthalate plastic into N-doped macroporous carbon for CO₂ emission through a one-pot synthesis," *Journal of Hazardous Materials*, vol. 899, article 128010, 2020.
- [16] G. Singh, J. Lee, A. Karakoti et al., "Emerging trends in porous materials for CO₂ capture and conversion," *Chemical Society Reviews*, vol. 49, pp. 4404–4860, 2020.
- [17] X. Ma, Y. Yang, Q. Wu et al., "Underlying mechanism of CO₂ uptake onto biomass-based porous carbons: do adsorbents capture CO₂ chiefly through narrow micropores?," *Fuel*, vol. 282, article 118727, 2020.
- [18] A. I. Osman, "Mass spectrometry study of lignocellulosic biomass combustion and pyrolysis with NO_x removal," *Renewable Energy*, vol. 146, pp. 484–496, 2020.
- [19] J. Wang, X. Yuan, S. Deng et al., "Waste polyethylene terephthalate (PET) plastics-derived activated carbon for CO₂ capture: a route to a closed carbon loop," *Green Chemistry*, vol. 22, p. 6886, 2020.
- [20] P. D. Dissanayake, S. W. Choi, A. D. Igalavithana et al., "Sustainable gasification biochar as a high efficiency adsorbent for CO₂ capture: a facile method to designer biochar fabrication," *Renewable and Sustainable Energy Reviews*, vol. 124, article 109785, 2020.
- [21] Global Monitoring Laboratory, "National Oceanic and Atmospheric Administration Trends in atmospheric carbon dioxide," <https://www.esrl.noaa.gov/gmd/ccgg/trends/>.
- [22] X. Yuan, J. G. Lee, H. Yun et al., "High capacitance of dipicolinic acid-intercalated MXene in neutral water-based electrolyte," *Chemical Engineering Journal*, vol. 399, article 125850, 2020.
- [23] A. D. Igalavithana, S. W. Choi, J. Shang et al., "Carbon dioxide capture in biochar produced from pine sawdust and paper mill sludge: effect of porous structure and surface chemistry," *Science of The Total Environment*, vol. 789, article 189845, 2020.
- [24] "Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: a review," *Technology*, vol. 319, article 124114, 2021.
- [25] H. Mashhadimoslem, M. Vafaenia, M. Safarzadeh, A. Ghaemi, F. Fathalian, and A. Maleki, "Development of predictive models for activated carbon synthesis from different biomass for CO₂ adsorption using artificial neural networks," *Industrial & Engineering Chemistry Research*, vol. 60, no. 38, pp. 13950–13966, 2021.
- [26] M. Belhachemi, M. M. de Castro, M. Casco, A. Sepúlveda-Escribano, and F. Rodríguez-Reinoso, "Adsorption of hydrogen on activated carbons prepared by thermal activation: Hydrogen storage," in *2019 10th International Renewable Energy Congress (IREC)*, pp. 1–5, 2019.
- [27] A. P. M. Michel, A. E. Morrison, V. L. Preston, C. T. Marx, B. C. Colson, and H. K. White, "Rapid identification of marine plastic debris via spectroscopic techniques and machine learning classifiers," *Environmental Science & Technology*, vol. 54, no. 17, pp. 10630–10637, 2020.
- [28] A. V. Konstantinov and L. V. Utkin, "Interpretable machine learning with an ensemble of gradient boosting machines," *Knowledge-Based Systems*, vol. 222, article 106998, 2021.
- [29] X. Zhu, D. C. W. Tsang, L. Wang et al., "Machine learning exploration of the critical factors for CO₂ adsorption capacity on porous carbon materials at different pressures," *Journal of Cleaner Production*, vol. 273, article 122915, 2020.
- [30] J. Fan, X. Ma, L. Wu, F. Zhang, X. Yu, and W. Zeng, "Light gradient boosting machine: an efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data," *Agricultural Water Management*, vol. 225, article 105758, 2019.
- [31] X. Yuan, S. I. Im, S. W. Choi, and K. B. Lee, "Removal of Cu(II) ions from aqueous solutions using petroleum coke-derived microporous carbon: investigation of adsorption equilibrium and kinetics," *Adsorption*, vol. 25, no. 6, pp. 1205–1218, 2019.
- [32] D. Liu, L. Li, A. Rostami-Hodjegan, F. Y. Bois, and M. Jamei, "Considerations and caveats when applying global sensitivity analysis methods to physiologically based pharmacokinetic models," *The AAPS Journal*, vol. 22, no. 5, p. 98, 2020.
- [33] "Scikit – learn, Machine Learning in Python," May 2021, <https://scikit-learn.org/stable/>.
- [34] H. W. Pennline, *Sorbent Research for the Capture of Carbon Dioxide*, U.S. Department of Energy, National Energy Technology Laboratory, 2016.
- [35] D. Chen, J. Mei, H. Li et al., "Combined pretreatment with torrefaction and washing using torrefaction liquid products to yield upgraded biomass and pyrolysis products," *Bioresource Technology*, vol. 228, pp. 62–68, 2017.

- [36] C. E. Golden, M. J. Rothrock, and A. Mishra Jr., "Comparison between random forest and gradient boosting machine methods for predicting *Listeria* spp. prevalence in the environment of pastured poultry farms," *Food Research International*, vol. 122, pp. 47–55, 2019.
- [37] Y. Chen, H. Yang, Q. Yang, H. Hao, B. Zhu, and H. Chen, "Torrefaction of agriculture straws and its application on biomass pyrolysis poly-generation," *Bioresource Technology*, vol. 156, pp. 70–77, 2014.