Article

An Augmented Model of Rutting Data Based on Radial Basis Neural Network

Zhuoxuan Li [1,2,†], Meng Tao [1,2,†], Jinde Cao [1,2,*] , Xinli Shi [3] , Tao Ma [4] and Wei Huang [5]

1   School of Mathematics, Southeast University, Nanjing 210096, China
2   Jiangsu Provincial Key Laboratory of Networked Collective Intelligence, Southeast University, Nanjing 210096, China
3   School of Cyber Science & Engineering, Southeast University, Nanjing 210096, China
4   School of Transportation, Southeast University, Nanjing 210096, China
5   Intelligent Transportation System Research Center, Southeast University, Nanjing 210096, China
*   Correspondence: jdcao@seu.edu.cn
†   These authors contributed equally to this work.

Abstract: The rutting depth is an important index to evaluate the damage degree of the pavement. Therefore, establishing an accurate rutting depth prediction model can guide pavement design and provide the necessary basis for pavement maintenance. However, the sample size of pavement rutting depth data is small, and the sampling is not standardized, which makes it hard to establish a prediction model with high accuracy. Based on the data of RIOHTrack's asphalt pavement structure, this study builds a reliable data-augmented model. In this paper, different asphalt rutting data augmented models based on Gaussian radial basis neural networks are constructed with the temperature and loading of asphalt pavements as the main features. Experimental results show that the method outperforms classical machine learning methods in data augmentation, with an average root mean square error of 3.95 and an average R-square of 0.957. Finally, the augmented data of rutting depth is constructed for training, and multiple neural network models are used for prediction. Compared with unaugmented data, the prediction accuracy is increased by 50%.

1. Introduction

In recent years, with the promotion of road transportation by economic development, the types and quantities of vehicles have been diversified, resulting in the rapid development of expressways. However, due to the increase in traffic flow and traffic axle load, the pavement will be damaged in varying degrees, such as cracks, permanent deformation, looseness, potholes, and slippage. As the main damage types, ruts and lateral cracks reduce the performance of the road surface, seriously affect driving comfort, threaten driving safety, and even cause damage to the road structure.

One of the primary pavement fatigue modes is the rutting in the asphalt and the underlying unbonded layer [1,2]. Rutting maintenance accounts for about 80% of asphalt pavement maintenance in China [3]. As one of the indicators for inspecting the degree of road damage, rutting has always played an indispensable and important role in road construction, inspection, and maintenance. An accurate rutting prediction model can guide pavement design and provide a foundation for pavement maintenance and repair.

By accurately predicting rutting depth, engineers can select the appropriate treatment and optimize the sequence and funds of the maintenance at the network level. Furthermore, good prediction allows for preventive maintenance throughout the pavement's life cycle [4]. In order to collect pavement indicators, such as the rutting depth of different pavement structures, a test road is usually constructed, and an actual vehicle loading test is carried

out. Since test roads are subjected to real traffic and environmental conditions, they are the most practical way to assess the impact of various parameters. It can be seen that a large number of models apply AASHO, NCAT, and other relevant data of large-scale test roads in the field loaded by actual vehicles. Although more accurate predictive models are available, these models cannot effectively reflect the situation of Chinese roads.

In November 2015, RIOHTrack (Research Institute of Highway MOT Track) was completed in China, aiming to verify the life-cycle service performance of long-life asphalt pavements [5]. Since then, China has fully grasped the evolution of asphalt pavements' whole-life multi-service performance. The conditions have been created to make up for the lack of original scientific data that has long constrained the innovation and development of the pavement discipline in China [6]. In the ROIHTrack, 19 main test pavement structures with different combinations of structural rigidity were erected to study and compare long-term performance and evolution. Relying on the automatic monitoring data of RIOHTrack, many prediction models and evaluation models for highway pavement have emerged [7–10].

However, RIOHTrack experiments are costly, and data collection also requires high costs. So far, an effective data enhancement model has yet to be proposed to augment the RIOHTrack experimental data credibly. In this case, how to efficiently use existing data and conduct credible data enhancement on existing data is very necessary. To generate the experimental data for RIOHTrack reliably, we propose a rutting data augmentation model based on Gaussian Radial Basis Neural Networks. These data provide sufficient data on which to build the rutting depth prediction model, improve the prediction accuracy of different neural network models, and validate the effectiveness of the augmented data model.

The main contributions of this paper are as follows:

(1) For the first time, the asphalt rutting depth data augmentation model based on the Gaussian Radial Basis Neural Network algorithm is proposed, which has higher accuracy than other machine learning algorithms and can reliably augment the RIOHtrack test data.

(2) Aiming at the effect of data enhancement, the experimental analysis is carried out on different asphalt structure pavements. For different neural network algorithms, using augmented data for rutting depth prediction improves the prediction accuracy by 30% compared with the original data. The effectiveness of the asphalt rutting data augmentation model based on the Gaussian Radial Basis Neural Network is verified.

The rest of this paper is organized as follows: Section 2 introduces the existing neural network-based rutting depth prediction work and the motivation of this paper. Section 3 introduces some basic concepts and principles of the algorithm. Section 4 builds a rutting data expansion model based on feature engineering and radial basis function neural networks; then, it compares various data expansion methods through numerical experiments and analyzes their advantages and disadvantages. Finally, Section 5 evaluates the data expansion model and looks forward to future work.

2. Related Word

The prediction and evaluation of rutting disease on asphalt pavement is an important index to study the degree of pavement damage and has been paid more and more attention by researchers. With the continuous development of computational science and data-driven methods [11–14], several researchers have studied the mechanical properties of asphalt mixtures from the perspective of theoretical analysis, aiming to accurately establish the intrinsic structure relationships of asphalt mixtures using viscoelastic, flexible layered systems and finite element analysis methods [6,10,15–18]. Many scholars have studied pavement metrics such as rutting depth using neural networks or statistical methods and proposed a data-driven approach to predict asphalt pavement metrics [7,19,20].

As a new tool in the field of modeling and prediction, artificial neural network (ANN) has become very popular in many fields, such as medicine, finance, transportation, etc. [21].

Neural networks (NNs) have been successfully utilized recently to solve many pavement engineering tasks, including pavement performance evaluation, pavement roughness estimation, fatigue cracking prediction, and maintenance decision optimization, due to their effectiveness in prediction and classification [22–25]. There are two main categories of NNs applications for rutting data mining. Different asphalt mixtures can be evaluated for the performance of rutting depth and then used as a guide to design and maintain pavements. The other is to build a rutting depth prediction model through NN, which provides an essential basis for the maintenance and repair of the road surface.

According to Simpson et al. [26], neural networks are used to model rutting based on 12 parameters, such as asphalt concrete (AC) thickness, base thickness, and AC voids. As compared to linear regression, NN's prediction performance is significantly better. Overfitting occurs when all data are trained, resulting in the duplication of features. Based on the long-term pavement performance (LTPP) database, Gong et al. [27] use NN training to predict rutting. In order to explain the NN algorithm, a random forest model is used to calculate the importance of pavement features. Based on artificial neural networks and genetic programming models, Mirabdolazimi et al. [28] predict the rutting depth. Experimental results are in good agreement with both artificial neural networks and genetic programming models. A neural network model has been used to investigate the complex relationships between asphalt properties and behavior, using a civil engineering application involving asphalt materials. Suo et al. [29] apply multiple regression techniques to develop a nonlinear model that can be used to study asphalt rutting behavior. This analysis can be used to design asphalt mixtures for new and repaired pavements. Xu et al. [30] develop a rutting prediction model using a random forest module on Python, suggesting that the proposed model can improve rutting prediction accuracy and identify the factors influencing the prediction. To establish a better performance rutting prediction model based on deep learning, Shang et al. [31] use a neural network (BPNN) and long short-term memory (LSTM) neural network combined with an attention model. It is found that the model has excellent prediction performance, with R^2 values of 0.821 for the training set and 0.796 for the testing set. Experiments have demonstrated that neural network algorithms perform well in modeling rutting prediction, but they require a lot of data to provide support.

When modeling and predicting road rutting using the NN, appropriate input data will directly affect the selection of parameters, model performance quality, and prediction results' accuracy. In addition, due to the uncertainty of field data and the short-term characteristics of indoor experimental data, if the amount of input data is insufficient, the NN cannot be accurately modeled. In the case of limited data, it is very necessary to enhance the original data.

Data augmentation was performed in the analysis of spectral data [32]. Data density is increased by adding Gaussian noise to the original data, thereby reducing prediction error. Data augmentation is becoming increasingly common in computer vision and pattern recognition, especially in image processing [33–35]. Take an image as an example. Data augmentation is simply deforming the image to augment the data set and deal with image deformation caused by different camera angles. The most commonly used methods are left and right flip, random crop, rotation, translation, noise perturbation, brightness and contrast transformation, and many other simples and efficient methods. In addition to enhancing image data, neural networks have been widely used in the enhancement tasks of time-series data. Guillermo et al. [36] develop a probabilistically composed electrocardio-gram generation algorithm, used synthetic data to enhance training convolutional neural networks, and achieved good classification results. Snow [37] uses multiple Generative Adversarial Network (GAN) stacks to simulate a variety of multivariate time-series data and apply it in the financial field. Li et al. [38] introduce the Transformer algorithm based on GAN to solve the problem that traditional algorithms cannot effectively model long sequence data points with irregular time relationships. However, few augmented models for asphalt rutting data are based on neural network algorithms. So far, the work has yet to

be conducted on the augmentation of rutting data, which shows that the augmentation of rutting data is meaningful.

In contrast to previous work, this paper makes use of Gaussian radial basis neural networks to develop an augmented model for rutting depth data with high accuracy. Moreover, it has good applicability to different types of asphalt pavements of RIOHTrack. It provides credible data support for the subsequent discovery of asphalt pavements' short-term and long-term service performance using rutting data.

3. Algorithm Introduction

3.1. Radial Basis Function Neural Network (RBFNN)

The use of neural networks has been successful in solving a variety of problems. A simple processing system consists of interconnected elements. A neural network aims to learn the nonlinear mapping between inputs and outputs based on sensor information. High-performance neural networks such as RBFNNs are among the most popular [39]. RBFNN is derived from function approximation and is a standard tool for function interpolation. RBFNN uses the radial basis function (RBF) as the 'base' of hidden units to form hidden layers. By mapping the low-dimensional input data into the high-dimensional space, the linearly inseparable problem in the low-dimensional space can be linearly separable in the high-dimensional space.

The RBF is the primary function in function space. They are the monotonic functions of the Euclidean distance between any point x in the space and a specific center c, which is denoted as $\phi(\|x - c\|)$. RBFs are commonly expressed in Table 1 with the Gaussian kernel being a popular choice [40,41].

Table 1. Kernel functions commonly used by RBFs.

Radial Basis Function	Mathematical Representation
Gaussian Function	$\phi(c) = exp(-\frac{c^2}{2\sigma^2})$
Multi-Quadric Function	$\phi(c) = (c^2 + \sigma^2)^{1/2}$
Thin Plate Spline Function	$\phi(c) = c^2 ln(c)$
Cubic Function	$\phi(c) = c^3$
Linear Function	$\phi(c) = c$

A RBFNN is a three-layer network, which includes input, hidden, and output layers, as shown in the following Figure 1.
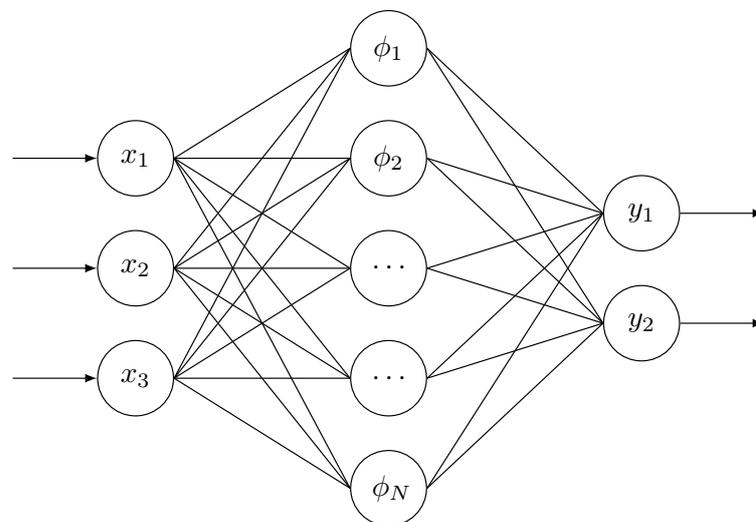

Figure 1. The structure of RBFNN.

For a set of n-dimensional input data X, the hidden layer contains N_T neurons and the radial basis function center of the i-th neuron is c_i. The c_i is also an n-dimensional vector. The mathematical expression of RBFNN can be expressed as [39]:

$$y = \sum_{i=1}^{N_T} \omega_i \phi(\|X - c_i\|_2) \tag{1}$$

where y is the output of the radial basis function; ω_i is the weight from the i-th neuron of the hidden layer to the output; ϕ is the RBF; and c_i is the center of the radial basis function ϕ_i. When the Gaussian function is selected as the kernel function, RBFNN can be expressed as

$$y = \sum_{i=1}^{N_T} \omega_i exp(-\frac{\|x - c_i\|^2}{2\sigma_i^2}), \tag{2}$$

where σ_i is the radius of ϕ_i. The key to RBFNN is to find N Gaussian radial basis functions ϕ_i and their corresponding weights ω_i.

Training a neural network involves adjusting its parameters to react to a stimulus to produce the desired response. The training of RBFNN algorithm consists of the following steps [42]:

Step 1. Preset the number N of neurons in the hidden layer;
Step 2. Determine the center c_i of the Gaussian radial basis function through a specific algorithm;
Step 3. Determine the radius σ_i of the Gaussian RBF through a specific algorithm;
Step 4. Calculate the weight matrix W from the hidden layer to the output layer.

3.2. Calculate the Parameters of the RBFNN

In the hidden layer, the number of neural units is determined by the needs of the described problem. When the number of neurons equals the number of input samples, a zero-error RBFNN is created with few training samples. Initially, a neural unit is established based on the largest error in the input data when the scale of training samples is large. The network is redesigned by gradually increasing the number of neurons to reduce error. The procedure ends when the error reaches the specified error performance or the number of neurons reaches the maximum. The amount of data n is usually much more than the number of neurons N.

3.2.1. Determine the Center by K-Means Clustering Algorithm

The K-means clustering algorithm is a typical data clustering algorithm widely used in various fields [43]. The key is to divide the data set into K clusters and minimize the distance sum of each data to the cluster center in each classification. For a given data set $D = \{x_i\}$, obtain clusters E_l from clustering, minimizing the sum of squares of errors denoted by SSE. The formula for calculating SSE is:

$$SSE = \sum_{l=1}^{K} \sum_{x \in E_l} dist(c_l, x)^2, \tag{3}$$

where E_l is the l-th cluster, and c_l is the center of cluster E_l. The primary process of the K-means clustering algorithm is as follows:

Step 1. As an initial cluster center, randomly select K objects from n sample data;
Step 2. The distance between each sample and each cluster center should be calculated separately, and the data should be assigned to the closest cluster center;
Step 3. The cluster centers of K should be recalculated once all data allocations are complete;
Step 4. Calculate the centers for each cluster based on the K clusters obtained in the previous calculation. If any cluster's center changes, proceed to step 2; otherwise, proceed to step 5;

Step 5. Terminate the algorithm and output the clustering results.

3.2.2. Determine Radius by k-Nearest Neighbor Algorithm

Next, determine the radius of the Gaussian radial basis function after finding its center. In the radius solution, distances between samples, distances between clusters, and distances between clusters are taken into account as measures of similarity. The k-nearest neighbor algorithm [44] is the most common method, also named as KNN, which is one of the simplest supervised machine learning algorithms used for classification. It is a classifier algorithm that classifies data by calculating the similarity of one data to another. In order to measure the similarity between samples, Euclidean distance, Manhattan distance, and Minkowski distance are used. Euclidean distance is the most commonly used, and its formula is as follows:

$$dist(x_s, x_t) = \sqrt{\sum_{j=1}^{m}(x_{tj} - x_{sj})^2}, \tag{4}$$

where x_s and x_t are different data sample. The primary process of the KNN algorithm is as follows:

Step 1. Determine the number k of neighbors;

Step 2. Calculate the distance of each node from the center by Equation (4), sort them in ascending order according to the size of the distance value, and the first k data samples after pre-sorting are selected;

Step 3. A root-mean-squared distance is calculated between the cluster and its k nearest neighbors. This is the value for σ_i, and its formula is expressed as:

$$\sigma_i = \sqrt{\frac{1}{k}\sum_{j=1}^{k}(x_{ij} - c_i)^2}, \tag{5}$$

where x_{ij} is the k neighbors of c_i.

In summary, the process of calculating the parameters of the Gaussian RBF is illustrated in Figure 2.
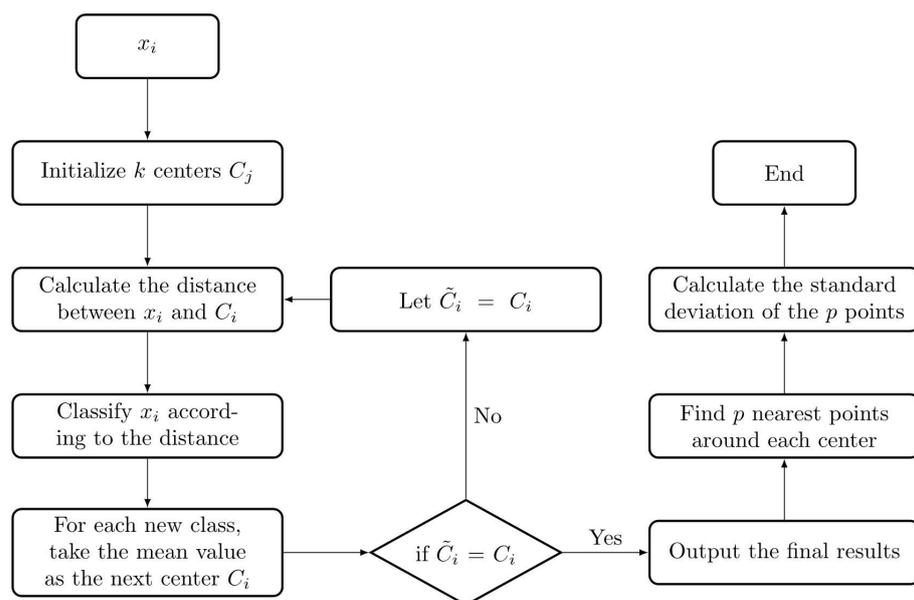

Figure 2. The process of calculating the parameters of the Gaussian RBF.

3.3. Calculate the Weights from the Hidden Layer to the Output Layer

After determining the number of neurons and the parameters of the Gaussian RBF, the next step is to determine the weights ω_i from the hidden layer to the output layer satisfing the following equations:

$$\begin{cases} w_1\phi_1(x_1) + \ldots + w_N\phi_N(x_1) = y_1 \\ w_1\phi_1(x_2) + \ldots + w_N\phi_N(x_2) = y_2 \\ \vdots \\ w_1\phi_1(x_n) + \ldots + w_N\phi_N(x_n) = y_n \end{cases}, \qquad (6)$$

The compact form of Equation (6) is

$$\phi W = y, \qquad (7)$$

where $\phi \in R^{n \times N}$, and $W = (\omega_1, \omega_2, \ldots, \omega_N)^T$. So, the weight matrix W is:

$$W = (\phi^T \phi)^{-1} \phi^T y. \qquad (8)$$

4. Experiment and Analysis

The measured rutting data come from the RIOHTrack. RIOHTrack includes 25 types of asphalt pavement structures in a runway-like layout totaling 2039 m. Over 1200 dynamic stress and strain sensors were embedded in 19 types of asphalt pavement structures in the main test section of the Ring Road in order to collect the mechanical response state of the asphalt pavement structure in real time under the coupling effect of load and environment. During the past five years, the collection work has accumulated relatively complete data of rutting depth about different pavement structures [45].

RIOHTrack covers all AC structure layers of China's high-grade highways as well as the flexible base thickness of thick asphalt pavement. The thickness of the asphalt concrete structure layer is 12, 18, 24, 28, 36, and 48 cm (or 52 cm). From the perspective of the type of base structure, it includes four typical structures: rigid base structure, semi-rigid base structure, flexible base structure, and full-thick asphalt pavement structure [5]. Figure 3 shows the layout of the pavement structure.
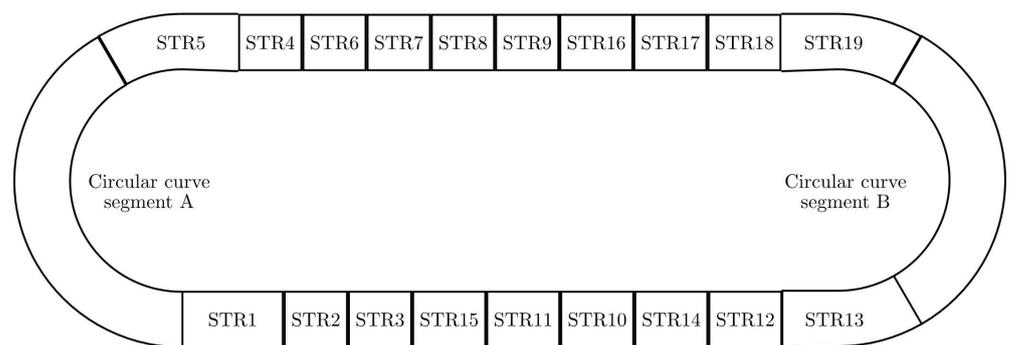

Figure 3. Structure of RIOHTrack.

This study uses measurement data of the full-scale pavement structure of 19 types of asphalt pavements as the data source. For each road surface, 103 efficient data samples are available, 52 of which are used for training, and the rest are used for testing. The features in the training set correlate highly with rutting depth. RMSE, MAE, MAPE, and R^2 indicators indicate the difference between the actual and predicted values. The smaller the RMSE, MAE, and MAPE values, the higher the model's accuracy. The closer R^2 is to 1, the higher the model's accuracy. The definitions of these accuracy metrics are shown in Table 2.

Table 2. Definitions of accuracy metrics.

Index	Formula
RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2}$
MAE	$\frac{1}{n}\sum_{i=1}^{n}\mid y_i - f(x_i)\mid$
MAPE	$\frac{100\%}{n}\sum_{i=1}^{n}\mid \frac{y_i - f(x_i)}{y_i}\mid$
R^2 score	$1 - \frac{\sum_{i=1}^{n}(y_i - f(x_i))^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

4.1. Model Building

The framework of the rutting data augmented model based on RBFNN is shown in Figure 4, which includes data preprocessing, feature engineering, and data augmentation.

Data preprocessing is a way of converting the raw data into a much-desired form so that valuable information can be derived from it, which is fed into the training model for successful medical decisions, diagnoses, and treatments [46]. This article deletes duplicate and missing data first; then, it converts the cumulative loading axis value into logarithmic form for consideration, since it is too large.
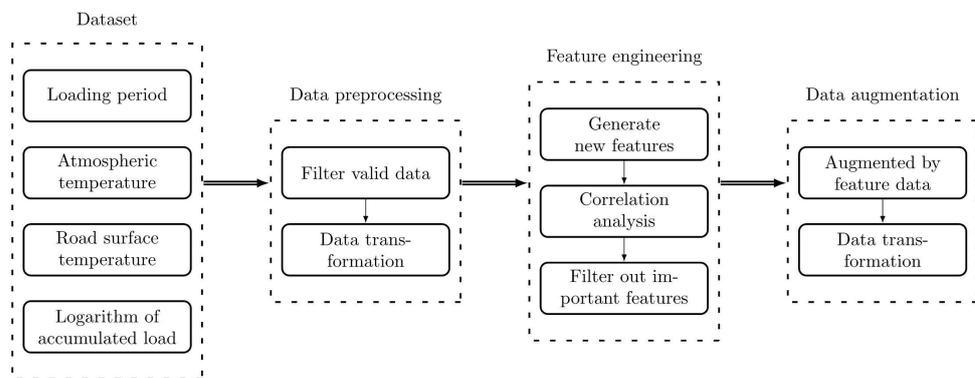

Figure 4. Framework of the rutting data expansion model.

Data features are created by transforming raw data into features that more accurately describe the problem and then are applied to the prediction model, improving its accuracy. The features in the data directly affect the predictive model and the obtained results. A recent rutting prediction model based on RIOHTrack automatic monitoring data fits the expression of the mechanical empirical constitutive equation (*M-E* model), which shows that the rutting depth is affected by temperature, load, deflection, and material parameters [47].

In this paper, the existing data are analyzed first, and the highly correlated features are removed according to the Pearson correlation coefficient to avoid the duplication of features in the modeling process. The existing data of the ring road are the loading period (LP), atmospheric temperature (AT), road surface temperature (RST), and logarithm of accumulated load (LAL). Taking the road data of STR2 as an example, six features, including single loading times (SLT) and difference in load change (DLC), are obtained by feature engineering. The data are shown in Table 3.

Table 3. Raw data of STR2.

LP	LAL	SLT	DLC	AT	RST	RD
N1	4.4533	0	0.56515204	3.8	13.52	15.67
N2	5.0184	0.56515204	−0.110689294	1.2	10.61	15.98
N3	5.4729	0.454462746	−0.24277208	0.9	11.08	14.81
N4	5.6846	0.211690666	−0.069787146	3.6	21.30	16.36
N5	5.8265	0.14190352	−0.036580855	11.4	30.75	8.53
. . .	. . .	. . .	. . .	. . .	. . .	. . .
N103	7.7407	0.006166064	−0.006166064	15.9	11.51	78.56

The heat map of the Pearson correlation coefficient matrix obtained from the correlation analysis is shown in Figure 5. If the correlation coefficient is above 0.8, it is considered that two features are highly correlated, and one feature is deleted to avoid feature duplication.

It can be seen that the features highly correlated with the rutting data are LP and LAL. However, there is also a high correlation between these two features, so only LAL is retained in this paper. In addition, this paper keeps AT due to the high correlation between AT and RST. Finally, the rutting data are extended by four essential features: AT, LAL, SLT, and DLC.
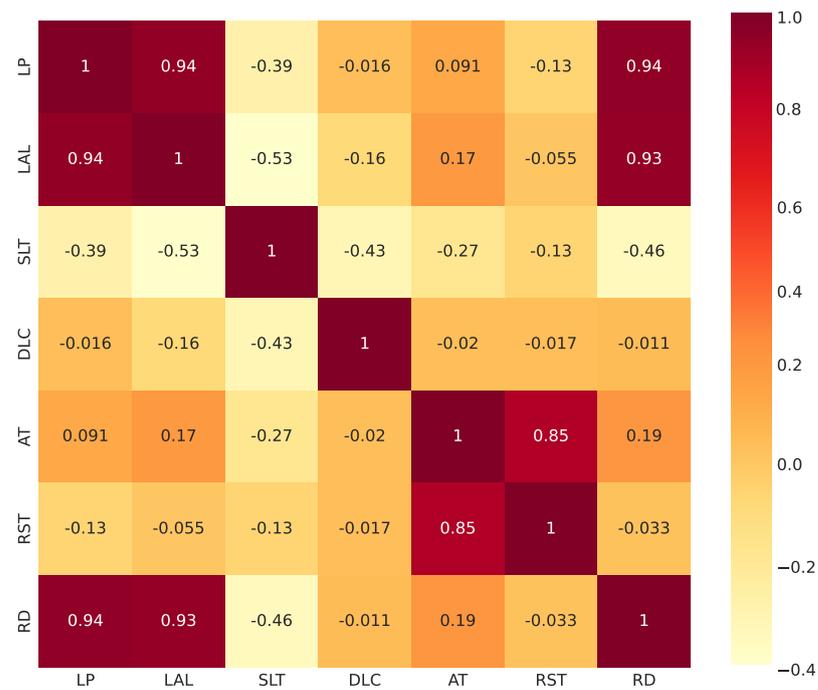

Figure 5. Correlation coefficient matrix heat-map.

4.2. Performance of RBFNN-Based Rutting Augmentation Model

Designed pavement structures in RIOHTrack can be further classified into six categories based on their structural combination and asphalt layer thickness [47]. There are six types of pavement structures selected in this paper for expansion, with STR2 representing thin semi-rigid AC base structures, STR8 representing common semi-rigid base structures, STR5 representing rigid composite base structures, STR12 representing an inverted structure, STR11 representing thick AC base structures, and STR18 representing full depth AC structures. Figure 6 shows the differences between the augmented data and the actual data in the test set of road data. There are five nodes in the input layer, n nodes in the hidden layer, and one node in the output layer of the RBFNN structure. The hidden layer neural units can be obtained through training, and in the RBFNN of the six road structures, there are 28, 22, 17, 30, 22, and 28 units. The accuracy indicators of the six types of data augmentation are shown in Table 4.

Table 4. RBF performance of different asphalt pavements.

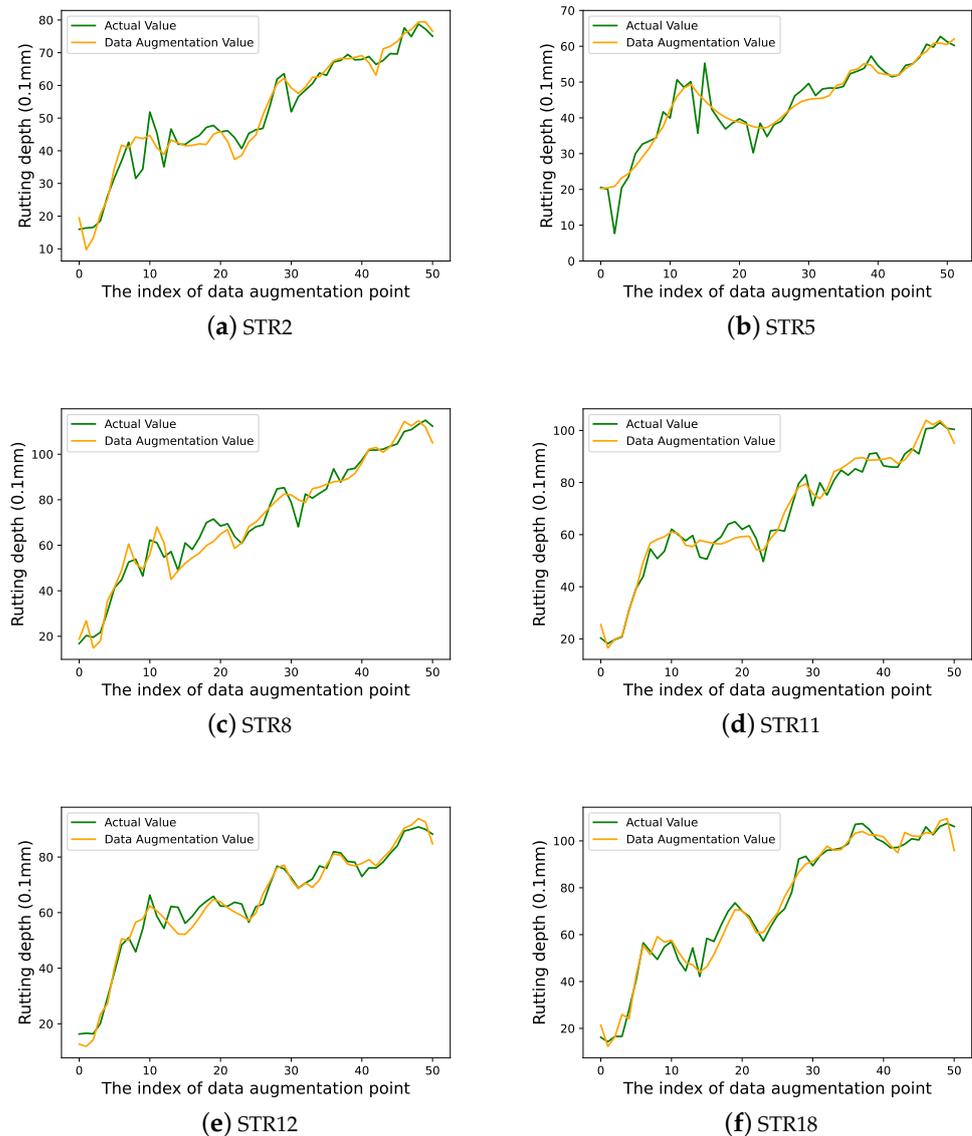| Type	RMSE	MAE	MAPE	R^2
STR2	3.757	2.874	0.074	0.9503
STR8	5.010	4.034	0.072	0.9620
STR5	3.524	2.229	0.081	0.9137
STR12	3.376	2.628	0.054	0.9680
STR11	3.862	3.186	0.053	0.9702
STR18	4.142	3.143	0.063	0.9776

Figure 6. Comparison of six types of pavement rutting depth augmented data and real data.

As shown in Table 4, the RMSE and MAE of the six types of structures are both single digits, and the values of MAPE are all less than 10%. Except that R^2 of STR5 is 0.9137, the R^2 values of other types are all above 0.95. The results show that the RBFNN-based data augmentation model proposed in this paper can make full use of feature information and achieve better augmentation accuracy. Hence, the model is suitable for all pavement structure data and has good generalization ability.

The residual distribution should be zero mean to ensure a good effect of the rutting depth interpolation augmentation model. Therefore, it is necessary to test the residuals. RBFNN augments the rutting data of six road surfaces, the residual histogram of the result is shown in Figure 7, and the probability density map of residual error is shown in Figure 8.

Figure 8 shows that the residuals of the model results in this paper are close to a normal distribution. To better test the distribution of model prediction residuals, the K-S test [48] is used to test whether the residuals of the model in this paper obey the normal distribution to verify the model's reliability. When the significance level is 0.05, the statistic calculated by the K-S test is 0.06294, and the *p* value is 0.16181 (>0.05), which does not violate the original hypothesis, indicating that the prediction residuals of the model obey the normal distribution.
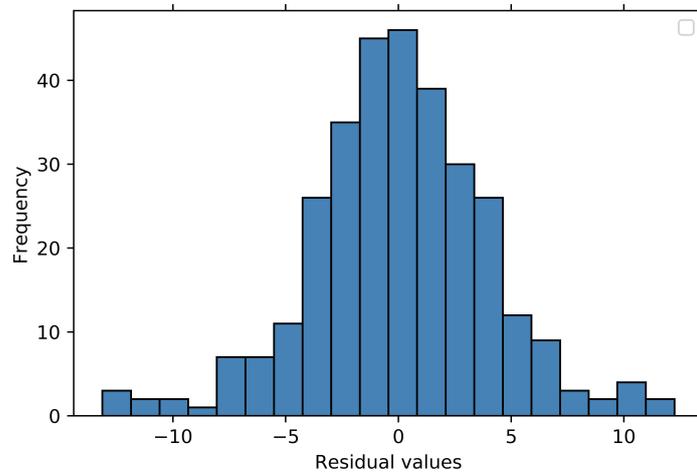
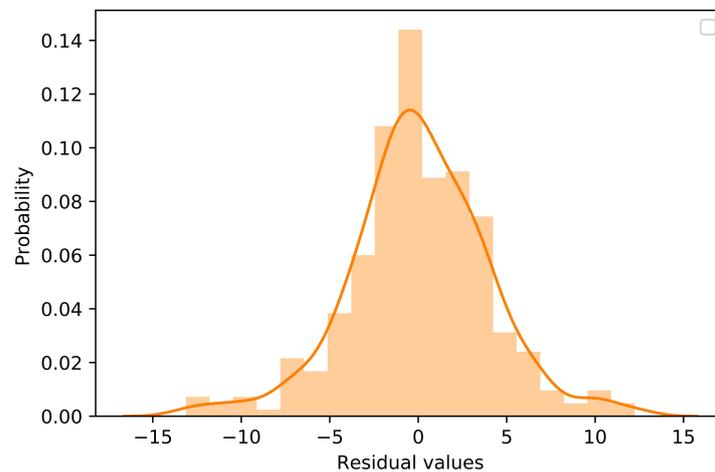**Figure 7.** Histogram of the predicted residuals of the model in this paper.

Figure 8. The predicted residual probability distribution of the model in this paper.

4.3. Model Comparison and Evaluation

This paper also compares the performance of RBFNN, KNN, DT, RF, XGBoost, MLP and SVM in terms of RMSE, MAE, MAPE and R^2. Using the Scikit learn [49] library for Python, the baseline models are implemented using recommended parameters.

RBFNN in this paper is almost the best algorithm used to augment the above six types of road surfaces. The average results of six types of pavement indicators are used for simplicity. The comparison results of all algorithms are shown in Table 5.

Table 5. Algorithmics performance of different asphalt pavements.

Algorithm	Mean RMSE	Mean MAE	Mean MAPE	Mean R^2
RBFNN	3.945	3.016	0.066	0.957
KNN	4.252	3.273	0.067	0.948
DT	5.403	3.961	0.088	0.906
RF	4.348	3.218	0.069	0.943
XGBoost	5.006	3.678	0.080	0.926
MLP	10.378	8.494	0.175	0.528
SVM	13.352	9.476	0.272	0.571

In Table 5, the RBFNN model is more accurate than the baseline model for data augmented and is the best performer under different metrics. It can be seen that the data augmen-

tation model proposed in this paper not only has a good data augmentation effect but also has a good generalization ability and can adapt to highly complex pavement structures.

4.4. Performance of the Augmented Dataset on Other Tasks on Rutting Depth

The regression estimation of rutting depth is tested to verify whether the augmented rutting data can be used in other tasks. This paper compares the training results before and after data augmentation for different NN algorithms, such as RBFNN, MLP, and ELM algorithms. For RBFNN, ELM, and MLP, 9 input nodes, 28 hidden nodes, and 1 output node are set up to perform regression estimation on the rutting depths of six pavement structures of STR2, STR5, STR8, STR11, STR12, and STR18, respectively. Boxplots of the performance of each algorithm on RMSE, MAE, and MAPE are shown in Figure 9.
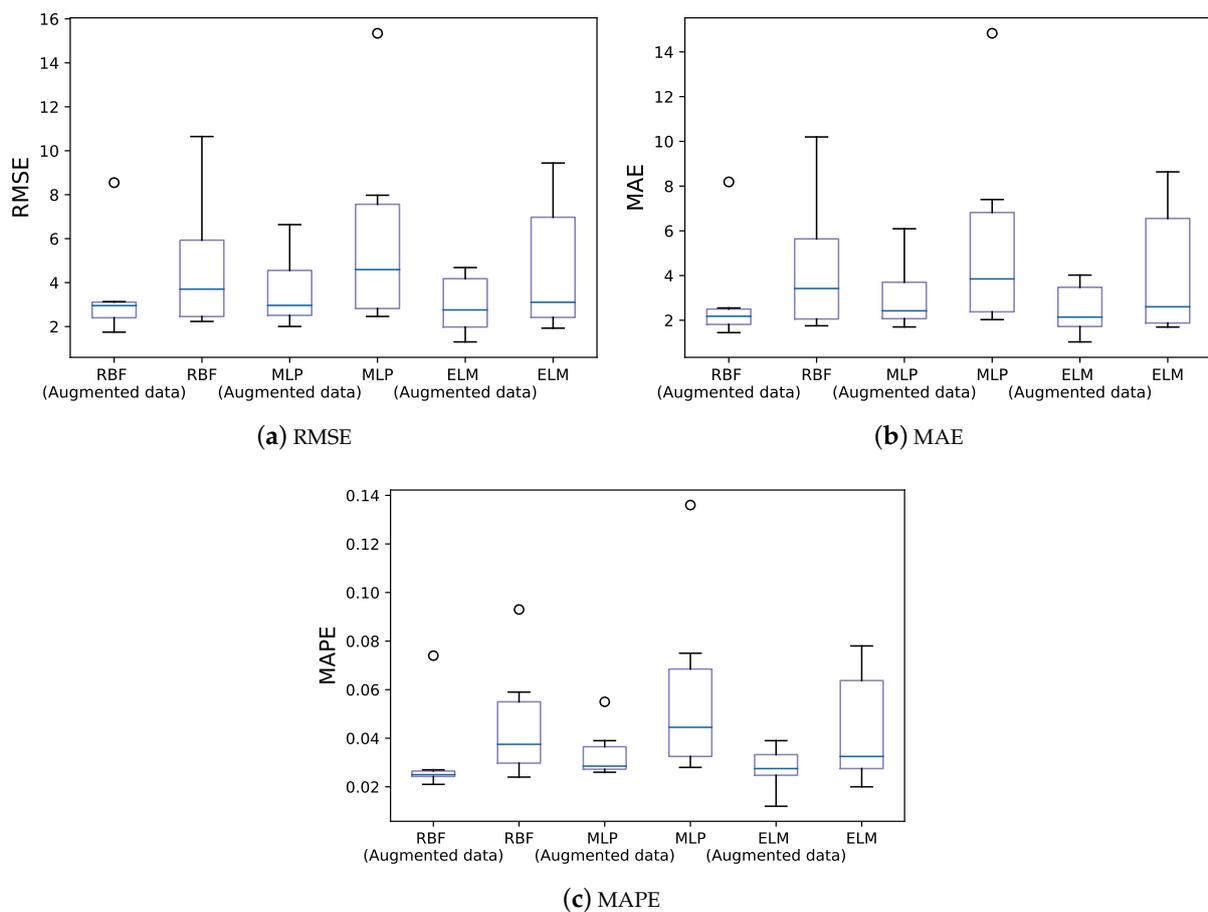

(**a**) RMSE

(**b**) MAE

(**c**) MAPE

Figure 9. Boxplots of accuracy index with different algorithms.

As shown in Figure 9, it is easy to find that the model accuracy has been significantly improved for the neural network algorithms under two different settings of the data set, using the augmented data as the training set. The comparison results of different algorithms are shown in Tables 6 and 7.

Table 6. Performance of different algorithms using raw data as training set on rutting depth regression estimation.

Algorithm	Average RMSE	Average MAE	Average MAPE
RBF	4.8327	4.4620	4.67%
MLP	6.2957	5.7092	5.97%
ELM	4.6720	4.1737	4.38%

Table 7. Performance of different algorithms using augmented data as training set on rutting depth regression estimation.

Algorithm	Average RMSE	Average MAE	Average MAPE
RBF	3.5965	3.0492	3.27%
MLP	3.6690	3.1193	3.40%
ELM	2.9782	2.4725	2.75%

Compared with the raw dataset, the Average MAPE of the model trained with the augmented dataset, the maximum boost, the minimum boost, and the average boost are 75.49%, 42.86%, and 59.25%, respectively. The Average RMSE has an enormous improvement. The maximum, minimum and average improvements are 71.59%, 34.37%, and 54.28%, respectively. The maximum increase in Average MAE, minimum, and average are 83.03%, 46.34%, and 66.05%, respectively. It is found from the experiment that using the RBFNN model to enhance the original data of the rutting data can effectively improve the prediction accuracy of the neural network algorithm for the rutting depth regression compared with the original data. It is verified that the augmented model is beneficial to the data mining tasks of rutting depth.

5. Conclusions

This paper proposes a rutting data augmentation model based on a radial basis function neural network, and the pavement rutting data of six different structures are augmented by screening out important, relevant features. The test experiments show that the expansion performance of the proposed model is better than other traditional models and neural network models in terms of accuracy indicators. Moreover, the model has better generalization performance. An experiment is tested for the rutting depth regression estimation, and it is found that using augmented data as the training set can effectively improve the regression accuracy of the neural network algorithm. This experiment verifies that the data augmentation method can effectively improve data quality.

In summary, the rutting data augmentation model proposed in this paper has the following advantages:

1. Avoid the uncertainty error caused by the use of field data and increase the amount of data, which can further improve the accuracy of the neural network-based prediction model;
2. Generate relatively regular data and try some new algorithms to explore the periodic characteristics of the data;
3. Effectively reduce the measurement cost. It can perform low-density measurements of rutting data and high-density measurements of easy-to-measure data such as temperature and axle order.

In further research, we hope to introduce the constitutive equation of rutting depth into the data augmentation model, aiming to transform the data-driven model to both a mechanistic and data-driven model.

Author Contributions: Conceptualization and methodology: Z.L. and J.C.; formal analysis and validation: Z.L., M.T. and X.S.; writing and visualization: T.M. and W.H. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

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

References

1. Domingos, M.D.I.; Faxina, A.L. Susceptibility of asphalt binders to rutting: Literature review. *J. Mater. Civ. Eng.* **2016**, *28*, 04015134. [CrossRef]
2. Norouzi, A.; Kim, D.; Richard Kim, Y. Numerical evaluation of pavement design parameters for the fatigue cracking and rutting performance of asphalt pavements. *Mater. Struct.* **2016**, *49*, 3619–3634. [CrossRef]
3. Sha, Q. *Premature Damage and Its Preservative Measures of Bituminous Pavement on Expressway*; China Communications Press: Beijing, China, 2001.
4. Fang, M.; Han, C.; Xiao, Y.; Han, Z.; Wu, S.; Cheng, M. Prediction modelling of rutting depth index for asphalt pavement using de-noising method. *Int. J. Pavement Eng.* **2020**, *21*, 895–907. [CrossRef]
5. Wang, X. Design of pavement structure and material for full-scale test track. *J. Highw. Transp. Res. Dev.* **2017**, *34*, 30–37.
6. Zhang, L.; Zhou, X.; Wang, X. Research progress of long-life asphalt pavement behavior based on the RIOHTrack full-scale accelerated loading test. *Chin. Sci. Bull.* **2020**, *65*, 3247–3258. [CrossRef]
7. Liu, H.; Cao, J.; Huang, W.; Shi, X.; Wang, X. Complex network approach for the evaluation of asphalt pavement design and construction: A longitudinal study. *Sci. China Inf. Sci.* **2022**, *65*, 172204. [CrossRef]
8. Liu, H.; Cao, J.; Huang, W.; Shi, X.; Zhou, X. A data-driven approach to the evaluation of asphalt pavement structures using falling weight deflectometer. *Discret. Contin. Dyn. Syst.-S* **2022**, *15*, 3223–3241. [CrossRef]
9. Li, Z.; Shi, X.; Cao, J.; Wang, X.; Huang, W. CPSO-XGBoost segmented regression model for asphalt pavement deflection basin area prediction. *Sci. China Technol. Sci.* **2022**, *65*, 1470–1481. [CrossRef]
10. Li, S.; Fan, M.; Xu, L.; Tian, W.; Yu, H.; Xu, K. Rutting performance of semi-rigid base pavement in RIOHTrack and laboratory evaluation. *Front. Mater.* **2021**, *7*, 590604. [CrossRef]
11. Nejad, R.M.; Sina, N.; Moghadam, D.G.; Branco, R.; Macek, W.; Berto, F. Artificial neural network based fatigue life assessment of friction stir welding AA2024-T351 aluminum alloy and multi-objective optimization of welding parameters. *Int. J. Fatigue* **2022**, *160*, 106840. [CrossRef]
12. Nejad, R.M.; Sina, N.; Ma, W.; Liu, Z.; Berto, F.; Gholami, A. Optimization of fatigue life of pearlitic Grade 900A steel based on the combination of genetic algorithm and artificial neural network. *Int. J. Fatigue* **2022**, *162*, 106975. [CrossRef]
13. Narayan, M.V.; Mohd, R.; Nadeem, R. Dunkl analogue of Sz á sz Schurer Beta bivariate operators. *Math. Found. Comput.* **2022**, *5*, 315–333. [CrossRef]
14. Mishra, L.N.; Raiz, M.; Mishra, V.N. Tauberian theorems for weighted means of double sequences in intuitionistic fuzzy normed spaces. *Yugosl. J. Oper. Res.* **2022**, *32*, 377–388. [CrossRef]
15. Xiao, Q.; Wang, X.D.; Zhou, X.Y.; Zhang, L.; Guan, W. Temperature Correction Method of Deflection Basin and Stress/Strain Response of Asphalt Pavement. In *Accelerated Pavement Testing to Transport Infrastructure Innovation*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 602–611.
16. Wu, J.; Wang, X.; Xiao, Q.; Zhang, L.; Guan, W.; Zhang, C. Temperature correction method of sensor measured texture depth index based on equivalent temperature of asphalt surface layer. *Int. J. Pavement Res. Technol.* **2021**, *14*, 450–458. [CrossRef]
17. Wu, Y.; Zhou, X.; Wang, X.; Shan, L. Long-Term Service Performance of Hard-Grade Asphalt Concrete Base Pavement Based on Accelerated Loading Test of Full-Scale Structure. *Sustainability* **2022**, *14*, 9712. [CrossRef]
18. Liu, X.; Wang, X.; Wu, Y.; Zhou, X. Modulus Back-Calculation of Four-Layer System Based on the Characteristic Parameters of Deflection Basin. *J. Test. Eval.* **2022**, *50*, 1887–1905. [CrossRef]
19. Uwanuakwa, I.D.; Ali, S.I.A.; Hasan, M.R.M.; Akpinar, P.; Sani, A.; Shariff, K.A. Artificial intelligence prediction of rutting and fatigue parameters in modified asphalt binders. *Appl. Sci.* **2020**, *10*, 7764. [CrossRef]
20. Huang, W.; Liang, S.; Wei, Y. Surface deflection-based reliability analysis of asphalt pavement design. *Sci. China Technol. Sci.* **2020**, *63*, 1824–1836. [CrossRef]
21. Liu, H.; Zhang, J.; Liu, Q.; Cao, J. Minimum spanning tree based graph neural network for emotion classification using EEG. *Neural Netw.* **2022**, *145*, 308–318. [CrossRef]
22. Guo, W.; Zhang, J.; Cao, D.; Yao, H. Cost-effective assessment of in-service asphalt pavement condition based on Random Forests and regression analysis. *Constr. Build. Mater.* **2022**, *330*, 127219. [CrossRef]
23. Zhang, X.; Ji, C. Asphalt pavement roughness prediction based on gray GM (1, 1|sin) model. *Int. J. Comput. Intell. Syst.* **2019**, *12*, 897–902. [CrossRef]
24. Ahmed, I.; Thom, N.; Zaidi, S.B.A.; Ahmed, N.; Carvajal-Munoz, J.S.; Rahman, T.; Ahmad, N. A mechanistic approach to evaluate the fatigue life of inverted pavements. *Constr. Build. Mater.* **2021**, *311*, 125288. [CrossRef]
25. Bao, S.; Han, K.; Zhang, L.; Luo, X.; Chen, S. Pavement Maintenance Decision Making Based on Optimization Models. *Appl. Sci.* **2021**, *11*, 9706. [CrossRef]
26. Simpson, A.L.; Daleiden, J.F.; Hadley, W.O. Rutting analysis from a different perspective. *Transp. Res. Rec.* **1995**, *1473*, 9–17.
27. Gong, H.; Sun, Y.; Mei, Z.; Huang, B. Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks. *Constr. Build. Mater.* **2018**, *190*, 710–718. [CrossRef]
28. Mirabdolazimi, S.; Shafabakhsh, G. Rutting depth prediction of hot mix asphalts modified with forta fiber using artificial neural networks and genetic programming technique. *Constr. Build. Mater.* **2017**, *148*, 666–674. [CrossRef]
29. Suo, Z.; Wong, W.G. Nonlinear properties analysis on rutting behaviour of bituminous materials with different air void contents. *Constr. Build. Mater.* **2009**, *23*, 3492–3498. [CrossRef]

30. Xu, Y.; Du, Y. Using random forest algorithm to build rutting prediction model of asphalt pavement. *J. Henan Urban Constr. Inst.* **2022**, *31*, 43–48.

31. Shang, Q.; Tian, B.; Li, S.; Quan, L. Research on Prediction Modle of LSTM-BPNN Feature fusion for Asphalt Pavement Rutting. *J. China Foreign Highw.* **2021**, *41*, 6.

32. Conlin, A.; Martin, E.; Morris, A. Data augmentation: An alternative approach to the analysis of spectroscopic data. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 161–173. [CrossRef]

33. Johnson, V.E.; Wong, W.H.; Hu, X.; Chen, C.T. Data augmentation schemes applied to image restoration. In *Medical Images: Formation, Handling and Evaluation*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 345–360.

34. Antoniou, A.; Storkey, A.; Edwards, H. Data augmentation generative adversarial networks. *arXiv* **2017**, arXiv:1711.04340.

35. Inoue, H. Data augmentation by pairing samples for images classification. *arXiv* **2018**, arXiv:1801.02929.

36. Jimenez-Perez, G.; Acosta, J.; Alcaine, A.; Camara, O. Generalizing electrocardiogram delineation: Training convolutional neural networks with synthetic data augmentation. *arXiv* **2021**, arXiv:2111.12996.

37. Snow, D. MTSS-GAN: Multivariate Time Series Simulation Generative Adversarial Networks. 2020. Available online: https://ssrn.com/abstract=3616557 (accessed on 26 June 2020).

38. Li, X.; Metsis, V.; Wang, H.; Ngu, A.H.H. TTS-GAN: A Transformer-based Time-Series Generative Adversarial Network. *arXiv* **2022**. arXiv:2202.02691.

39. Montazer, G.A.; Giveki, D.; Karami, M.; Rastegar, H. Radial basis function neural networks: A review. *Comput. Rev. J* **2018**, *1*, 52–74.

40. Franke, C.; Schaback, R. Convergence order estimates of meshless collocation methods using radial basis functions. *Adv. Comput. Math.* **1998**, *8*, 381–399. [CrossRef]

41. Baxter, B. The interpolation theory of radial basis functions. *arXiv* **2010**, arXiv:1006.2443.

42. Broomhead, D.S.; Lowe, D. *Radial Basis FUNCTIONS, Multi-Variable Functional Interpolation and Adaptive Networks*; Technical Report; Royal Signals and Radar Establishment Malvern: Malvern, UK, 1988.

43. MacQueen, J. Classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: San Francisco, CA, USA, 1967; pp. 281–297.

44. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]

45. Wang, X.; Zhou, X.; Guan, W.; Xiao, Q. Characteristics and analysis of the mechanical response inside the structure of asphalt pavement. *Chin. Sci. Bull.* **2020**, *65*, 3298–3307. [CrossRef]

46. Piuri, V.; Raj, S.; Genovese, A.; Srivastava, R. *Trends in Deep Learning Methodologies: Algorithms, Applications, and Systems*; Academic Press: Cambridge, MA, USA, 2020.

47. Liu, G.; Chen, L.; Qian, Z.; Zhang, Y.; Ren, H. Rutting prediction models for asphalt pavements with different base types based on RIOHTrack full-scale track. *Constr. Build. Mater.* **2021**, *305*, 124793. [CrossRef]

48. Young, I.T. Proof without prejudice: Use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources. *J. Histochem. Cytochem.* **1977**, *25*, 935–941. [CrossRef] [PubMed]

49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.