American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-11, Issue-09, pp-26-43
www.ajer.org
Research Paper Open Access

Wind Turbine Blade Icing Detection Via Deep Neural Network and Transfer Learning

Jiaohui Xu[1], Wen Tan[1],Tingshun Li[1],Zesan Liu[2], Zhijian Du[3]
[1]School of Control and Computer Engineering, North China Electric Power University, Beijing, China
[2]R & D, State Grid Information & Telecommunication Group Co., Beijing, China
[3]Beijing Xinwendao Education Technology Group Co. Ltd., Beijing, China
Corresponding Author:Wen Tan

ABSTRACT : *The number of wind turbines will be doubled in five years with carbon-peak and carbon-neutral commitment. It's a big problem that power generation and safety performance can be decreased because the blades of wind turbines in high latitudes often cover ice in winter. This paper proposes a hybrid model based on deep neural net and transfer learning which can detect whether blades of wind turbines in remote areas are been covered ice by supervisory control and data acquisition system (SCADA). In this study, both features and detecting model of blade icing are transferred between different wind turbines via transfer learning. In order to get more accuracy, some features of blade icing are transferred, which are testified more effective in previous studies. It's impossible that the models of dozens or even hundreds wind turbines in park are all trained separately, so it's necessary to build a general and effective model to detect icing of blades, which is trained from some typical wind turbine and then carry out to others. The initial model is obtained from 15# wind turbine by deep neural network (DNN), and then transfer to 21# wind turbine. Via dozens of experiments, a better model is obtained, which is testified more effective, lower error warning, more practical.*
KEYWORDS*SCADA; DNN; transfer learning; wind power; blade icing*.

Date of Submission: 01-09-2022 Date of acceptance: 13-09-2022

I. INTRODUCTION

In 2021, the Chinese government clearly put forward the goal of "carbon peak and carbon neutrality". Therefore, the electric power industry will become the main battlefield of energy transformation, and the construction of new energy will be further accelerated [1]. Wind power is one of the most important components of new energy. By the end of May this year, China's installed capacity of wind power was 290 million kW, an increase of 34.4% year-on-year [2]. Most wind power is installed in the middle and high latitudes, where conditions are harsh. In these areas, affected by cold weather conditions in winter, ice covering phenomenon of wind turbines is inevitable, which causes errors in wind speed measurement and control, reduces wind turbine output, causes mechanical and electrical faults, threatens the safe and reliable operation of wind turbines, and even affects the development of wind power [3,4,5].

At present, there is no mature commercial product on how to protect the wind turbine from the impact of icing [3]. Therefore, scholars from various countries have conducted in-depth research and put forward many research methods. Literature [3] points out that active blade heating is a reliable method to prevent the effect of ice coating, which is often used together with passive hydrophobic coating to reduce anti-ice energy consumption. It is also pointed out that the combination of dual anemometer and relative humidity measurement is an economical and reliable method for ice covering detection in the stage of wind resource assessment. After being put into operation, it is recommended to use ice-covered sensor and power curve method. However, literature [5] has conducted an in-depth study on the performance attenuation of wind turbine blades using superhydrophobic coating to prevent icing. On the premise of not adding deicing device, Wei Zhenhai and other researchers applied the operation data of wind turbineSCADA system to judge the icing situation of blades through monitoring information, and proposed corresponding solutions [6]. With the powerful computing power of computers, many researchers have proposed various artificial intelligence methods to remotely monitor wind turbineicing. In literature [7], deep automatic coding network was used to extract characteristic variables related to ice cover, and a remote monitoring model of ice cover was established. Literature [8] firstly selects two kinds of data highly related to leaf icing, and then uses optimized support vector regression to predict them. Finally, BP self-clustering algorithm is used to predict blade icing fault by multi-source fusion. The above methods are simple and do not involve the problem of model transfer and application to other wind turbines. If the model is only derived from a certain wind turbine, without considering transfer, there is more or less the problem of over-bridging, and the universality needs to be improved and further discussed and studied.

In conclusion, in the large-scale wind farms have hundreds of wind turbines, can't be installed each wind turbine ice monitoring equipment (high cost), also can't be set up for each specific model (training model to observed data of ice per all observations are not feasible), so in this paper from the observation of a few typical wind turbine of ice in the data, Based on the deep neural network technology, a wind turbine icing sensing model is established. The proposed model not only has high accuracy, but also has good generalization, which can be applied to other wind turbines by transfer learning technology, and also has high prediction accuracy. Therefore, this paper focuses on how to establish a portable ice covering monitoring model, and studies the transfer method, transfercontent and transferevaluation index.

Our research group has previously studied the influencing factors of icing, including the causes and performance factors of icing, analyzed the correlation between each factor and icing, and published relevant literature [9]. This study transfers the influencing factors of ice covering in literature[9], and the experiment shows that it can greatly improve the accuracy of the model and has good mobility. In addition, considering the practical performance, this paper tries not to take the accuracy rate of prediction as the only goal, but to establish an index system for comprehensive consideration based on the prediction accuracy rate, rate of return and missing rate, so as to comprehensively evaluate the advantages and disadvantages of the established model.

After many experiments, this paper obtains a good icing model, which not only has a high icing accuracy, but also has a good accuracy when transfer to other wind turbines. What is more valuable is that the missing rate of icing alarm is better, which greatly improves the practicability of the model. The experimental results show that transfer learning is feasible in icing warning. This research on wind turbine icing prediction based on transfer learning technology can also provide reference for other related transfer learning research.

II. DATA PREPROCESSING

2.1 Data source

The State Grid Corporation has built a large wind farm in the north of China, with hundreds of wind turbines and complete data. After careful selection, two wind turbines have relatively complete operation data and icing record, which can be used as the data training object of the wind turbine icing model. That is, the 15# wind turbine data is used to train the icing sensing model, and then the 21# wind turbine data is substituted into the model for testing, so as to test the transfer ability and prediction accuracy of the model. 21# wind turbine is just a randomly selected wind turbine in the wind field, and it is used to verify the characteristics of icing and the transfer ability of icing model, so it has certain representativeness and credibility. Of course, in order to actually test the universality of the icing model, it is necessary to verify the vast majority of types of wind turbines in the wind field one by one.

The wind turbine SCADA (Supervisory Control and Data Acquisition) system includes 26 monitoring points, which can be regarded as the 26 characteristics of the wind turbine, as shown in Table 1 below.

Table 1 Wind turbine detection variables

no	scada variables	no	scada variables
1	wind_speed	14	pitch1_moto_tmp
2	generator_speed	15	pitch2_moto_tmp
3	Power	16	pitch3_moto_tmp
4	wind_direction	17	acc_x
5	wind_direction_mean	18	acc_y
6	yaw_position	19	environment _tmp
7	yaw_speed	20	int_tmp
8	pitch1_angle	21	pitch1_ng5_tmp
9	pitch2_angle	22	pitch2_ng5_tmp
10	pitch3_angle	23	pitch3_ng5_tmp
11	pitch1_speed	24	pitch1_ng5_DC
12	pitch2_speed	25	pitch2_ng5_DC
13	pitch3_speed	26	pitch3_ng5_DC

The composition of the trained SCADA data is shown in Table 2 below.

Table 2 Data distribution of wind turbine samples

wind turbine data		number	percent
15#wind turbine	Normal data	170000	88%
	Icing data	23000	12%
	Combined data	193000	100%
21#wind turbine	Normal data	160000	94%
	Icing data	10000	6%
	Combined data	170000	100%

2.2 Data set partitioning

15# wind turbine contains more data, so the data of 15# wind turbine can be used as the training set. The 21# wind turbine is used as the transfer object to test the transferability of the model.

When 15# wind turbine is used as the training model dataset, in order to improve the accuracy of deep learning training, the dataset needs to be divided into training set, validation set and test set to train a better model and conduct cross-validation. The training set is used to train the model; Validation set is used to correct the model during training. Test set to test the training model. In view of the relatively large data sets, a ratio of 8:1:1 among the three data sets is better [10]. The subsequent experimental results show that a more ideal model can be trained after this partition.

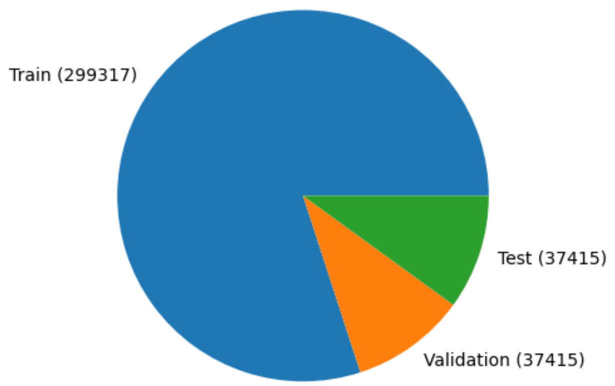


Fig. 1 Distribution diagram of train set, test set and validation set

2.3 Normalized treatment

In general, the value difference of each component in the sample is large, which will affect the prediction result because of the numerical difference. In order to reduce the impact of this numerical difference, it is necessary to normalize the sample data.

There are many Normalization methods, including Min-max Normalization and Z-score Normalization methods [10]. Considering that the model will be applied to real time prediction in the future, it is impossible to obtain the standard deviation in real time, so the Min-max standardization method is chosen.We use the min-max method to map the value of the characteristic variable to the range [0, 1]

$$x^* = \frac{x - \min}{\max - \min} \tag{1}$$

III. ICE COVERING DETECTION ALGORITHM

3.1Feature Transfer

Wind turbine blade icing is a complex and changeable process, which is not only related to environmental factors, such as temperature, wind speed and humidity, but also related to the running state of the wind turbine, such as speed and Angle. When using data analysis methods to build predictive models for wind turbine icing, the biggest challenge is to find out which features are associated with icing.

Relevant features can be transferred to this study by referring to previous research results of wind turbine icing and using transfer learning knowledge. Transfer learning is a new machine learning method that uses existing knowledge to solve problems in different but related fields. It relaxes two basic assumptions in traditional machine learning :(1) the training samples used for learning and the new test samples meet the conditions of independent and identical distribution; (2) sufficient training samples must be available to learn a good classification model [11,12]. In literature [13], the transfer learning method was used to transfer the vocalization features related to domestic cats previously studied to the current study, which achieved good results and fully proved the feasibility of feature transfer.

In this paper [9], feature vectors related to ice cover are deeply discussed, and a method to find feature vectors related to ice cover is proposed, which has high practicability. According to his research results, it can be known that the vector and correlation degree related to ice cover are shown in Table 3 below.

Table 3 Variables correlated with icing

no	the feature vectors	relevance
1	yaw_position	0.433590082
2	environment_tmp	-1.076909987
3	wind_Speed_Face	0.898209381
4	pitch_angle	-0.632307849
5	TD_inttmp	0.246596158
6	power	-2.162700181
7	TD_moto	0.348225667
8	wind_direction	-0.170664741
9	wind_Speed_Face_Mean	0.36449363
10	acc_x	-0.084550267
11	wind_speed	0.579723039
12	int_tmp	0.26705721

3.2Correlation analysis of icing characteristics

The characteristics of each monitoring point in the wind turbine SCADA system are some of the factors that constitute the wind turbine icing, such as the ambient temperature; Some factors are obvious differences compared with normal values after wind turbine icing, such as wind turbine output power, which can be called icing performance factors. Both genetic factors and performance factors can be used as the characteristics of ice cover monitoring model.

(1) Factor analysis of icing performance

The working process of wind turbine is that the wind turbine blade is driven by the outside wind, driving the generator in the wind turbine to work and generate electricity. When the wind turbine blade is covered with ice, the output power of the motor under the same wind force will be affected. Figure 2 below shows the distribution of wind speed wind_speed and wind turbine output power of sample 15# in the dimension of time.

As can be seen from the bottom of Figure 2, under the condition of non-icing, the wind speed is basically in direct proportion to the output power of the wind turbine within a certain range. The middle figure just reflects that there is no proportional relationship between the wind speed and the output power in the ice-covered state, and the wind speed is mainly concentrated in the low-power area. This indicates that refreezing does reduce the wind turbine power output.

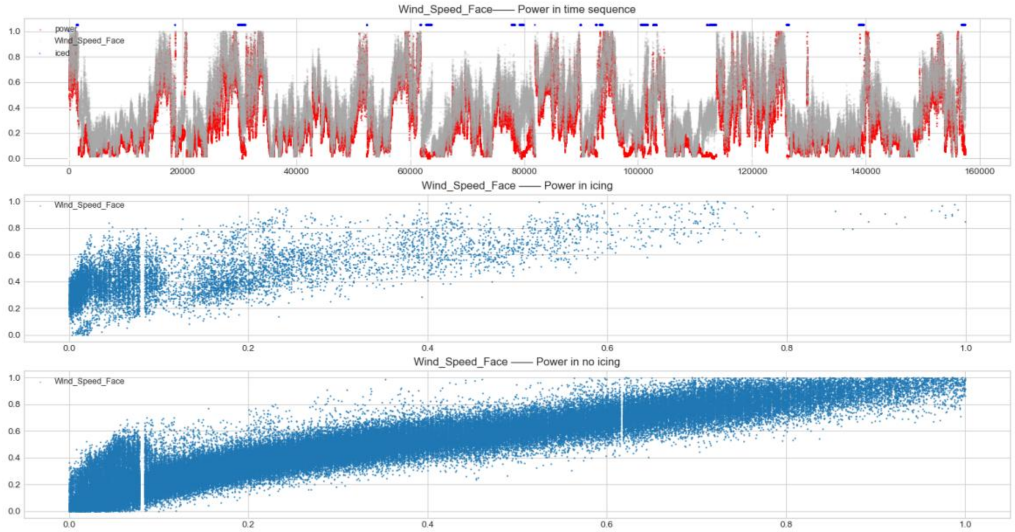


Fig. 2 Relation between wind speed and power of wind turbine
Fig. above: Distribution on time of wind speed(gray points) and power(red points)
(The upper horizontal lines in the figure are time periods of icing of wind turbine)
Fig. middle: Relation between wind speed and power in icing
Fig. Below: Relation between wind speed and power

This relationship is more intuitively reflected by enlarging part of Figure 2, as shown in Figure 3.

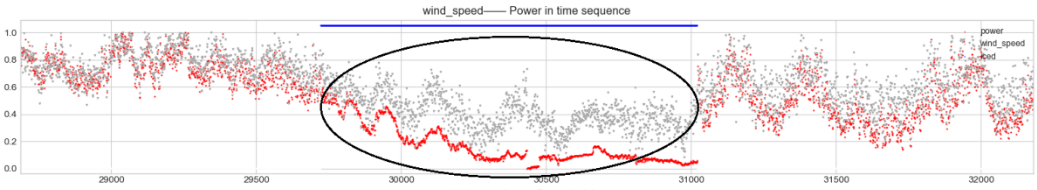


Fig. 3 Partial magnification of relation diagram between wind speed and power
(Circle part is icing)

It can be clearly seen from FIG. 3 that the wind speed and the output power of the wind turbine basically overlap under the condition of non-icing, while the wind speed is above the output power after the wind turbine is iced, indicating that the output power of the wind turbine is significantly reduced under the same wind speed after the wind turbine is iced.

The aerodynamic characteristics of the wind turbine will be affected obviously after icing. For example, pitch_angle, when the wind turbine is iced, the probability distribution of the large value is in the low value range around 0.2, as shown in Figure 4 below.

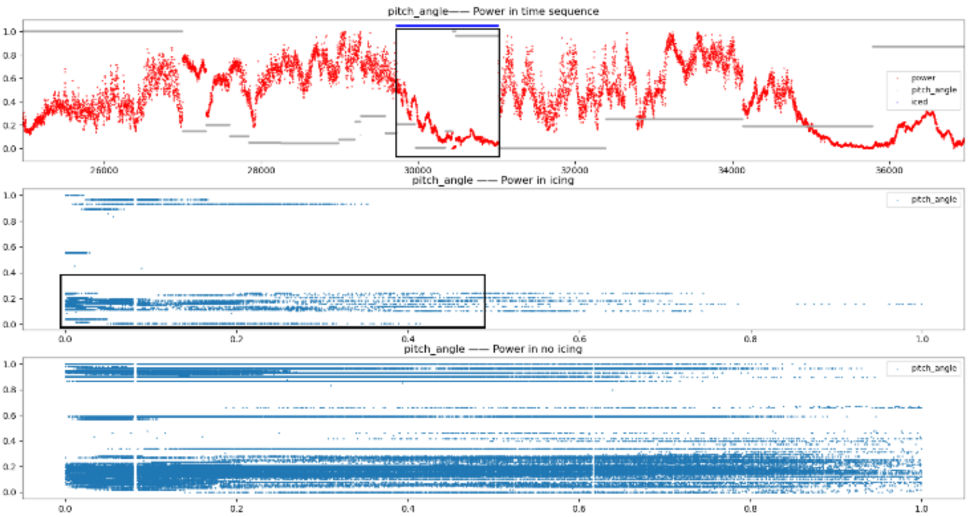


Fig. 4 Relation between pitch angle and power

(2) Analysis of cause factors of icing

According to common sense, the wind turbine blade icing needs to meet certain conditions: the ambient temperature is reduced to below freezing point; The air contains a certain amount of water vapor (humidity meets certain conditions); There is a certain temperature difference between inside and outside the wind turbine; Blade metal heat dissipation fast and so on.

Figure 5 shows the influence of ambient temperature on wind turbine icing.

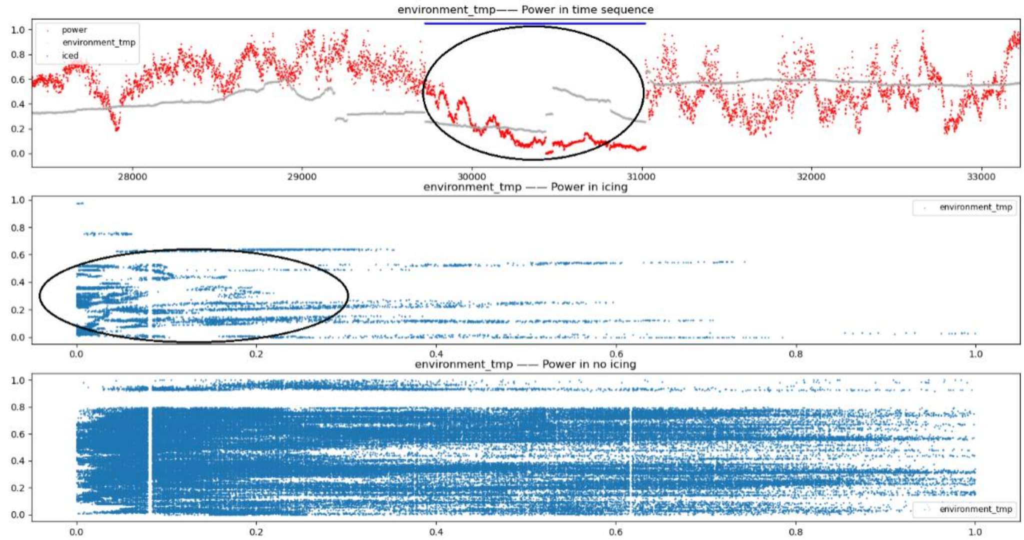


Fig. 5 Relation between ambient temperature and power

It can be seen from the figure that most of the icing is concentrated in the low-temperature area, but on the contrary, the low-temperature area is not necessarily covered with ice, and there are other factors that affect the icing.

A temperature difference feature (environment_TMP-int_tmp) is constructed by using environment_tmp and int_tmp in the wind turbine SCADA system. Figure 6 shows the influence distribution of the temperature difference feature of this structure on wind turbine icing.

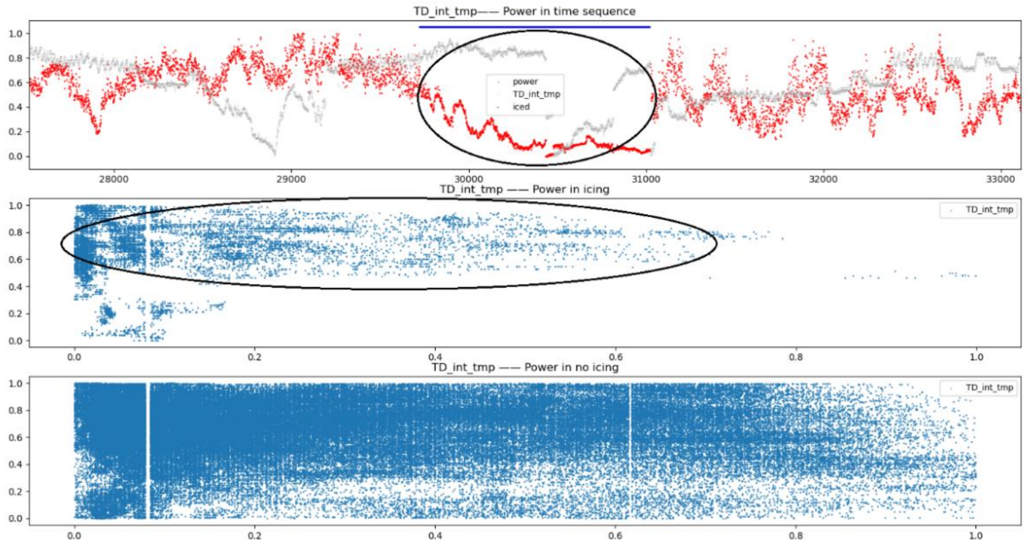


Fig. 6 Relation between temperature difference and power

As can be seen from the figure, the large probability distribution of temperature difference value in the ice-covered state is in the high value interval (circled part in the above figure and the middle figure).

In order to improve the prediction accuracy of the model, it is sometimes necessary to add some features that are not so obvious associated with icing. Here, no more details. In the experimental part below, the training results of the transfer feature model and the ordinary feature model will be compared.

3.3Deep neural network model

Obviously, there is a nonlinear relationship between wind turbine icing and each feature, and deep neural networks can theoretically bridge all kinds of nonlinear relationships [14,15,16]. To this end, a deep neural network as shown below can be constructed.

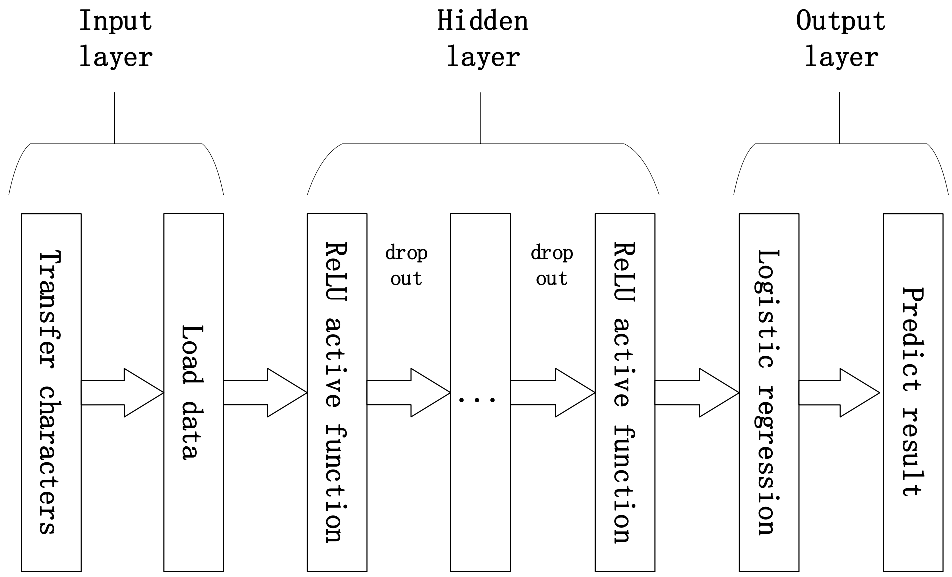


Fig. 7 DNN model

(1) **Hidden layer activation function**

The contractor shall provide at the reasonable satisfaction with ReLU (Rectified Linear Units) function for the hidden layer in the following form:

$$f\left(x\right)=\max(0,x) \tag{2}$$

Compared with traditional neural network activation functions, such as Logistic sigmoid and TANh, ReLU function has the following advantages:

- More efficient gradient descent and backpropagation, and avoid the problem of vanishing gradient;
- Reduces the amount of computation and simplifies the process.
- Increasing sparsity decouples the highly coupled variables and tolerates noise, making the model more robust.

(2) Number of hidden layers

Under normal circumstances, the number of hidden layers should not be too much. After several experiments, it is found that for the wind turbine icing samples, if the number of layers is small, the nonlinear bridging is not very good, the accuracy of the constructed model is not very good, and the convergence is slow. In the subsequent experiments, a deep neural network model with five hidden layers was selected, and the effect was good.

(3) Output layer design

Whether the wind turbine is iced or not is a dichotomous prediction problem, so the output layer selects the logistic regression model. Logistic regression is a classical linear classification model, and Sigmoid function is used to calculate the probability value of binary classification [17].

$$f(x)=\frac{1}{1+e^{x}} \tag{3}$$

(4) Cost function design

The cost function adopts the cross-entropy function [17], and the form is as follows:

$$\mathrm{H(X)}=-\frac{1}{N}\sum_{i=1}^{N}p(x_i)\log(q(x_i)) \tag{4}$$

Where $p(x_i)$ represents the true label value at sample$i$; $q(x_i)$ represents the forecast probability at sample.

Ice covering is a dichotomous classification problem, so the corresponding cross-entropy function can be expressed as:

$$H(X) = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(q(x_i)) + (1-y_i)\log(1-q(x_i))) \tag{5}$$

Where y_i is the true value at sample i, 1 or 0;

$q(x_i)$ represents the probability that the predicted value at sample i is 1, then 1- $q(x_i)$ is the probability that the predicted value is 0.

Clearly, in machine learning, cross-entropy can be expressed as the difference between the true probability distribution and the predicted probability distribution. The smaller the cross-entropy value is, the better the prediction effect of the model is.

(5) dropout parameters

In order to prevent the overfitting problem of the neural network trained model, dropout can be added to the hidden layer to discard neurons with a certain probability, so as to reduce the overfitting of the model to samples and enhance the transferability of the model [18,19].

The 5-layer deep neural network dropout constructed in this paper adopts different values in different layers. After many tests, the following values are finally adopted: 0, 0.4, 0.3, 0.2, 0.2.

(6) Generated NN model

According to the above method, an ice-covering sensing model can be obtained by using the training data of 15# wind turbine. Figure 8 is the DNN model obtained in a training and the schematic diagram after export.

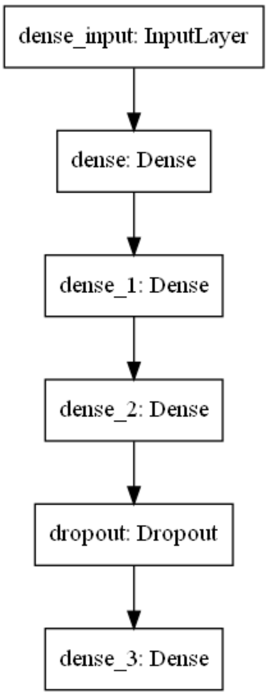


Fig.8 Sketch map of icing detecting DNN model

3.4 Ice cover model modification

Wind turbine icing is a continuous process in time, during a period of icing period, it is impossible to appear unicing, that is, icing has a certain "inertia". However, the prediction results given by the prediction model are judged according to the SCADA monitoring data at a certain moment, and the judgments before and after are isolated in the time dimension. This isolated judgment is not reasonable, and if the inertia of ice can be properly used, the accuracy of prediction can be improved to a certain extent. Therefore, it is necessary to modify the prediction results of the model.

(1) Latency Parameters

Literature [9] points out that the concept of "delay" in inertial system can be used for reference, and the prediction results before and after can be linked within the delay time. As a whole, the prediction results of the current model at a certain moment are no longer isolated, so as to improve the rationality and effectiveness of the prediction.

(2) Latency Modification Process

After adding the delay parameter, it is necessary to recalculate and revise all the predicted values of the prediction results within the delay. The modified model is shown in Figure 9 below.

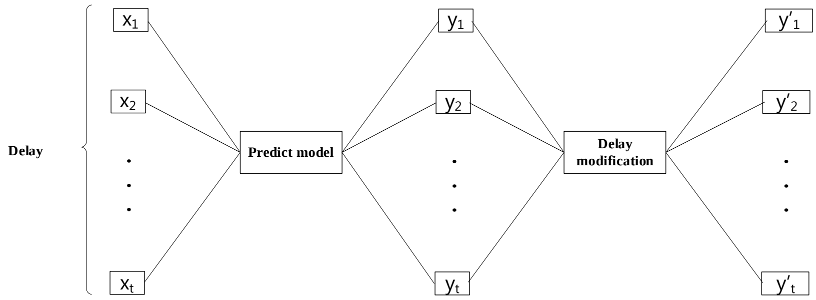


Fig. 9 Tuned model of time delay

(3) Effect of delay size

It can be seen from Figure 9 that the larger the value of delay is, the more prediction results it affects and the more possible correction values it has. However, more correction values do not necessarily improve accuracy. Because the direction of the correction is not necessarily the correction to the correct result, it is also possible to correct the correct result to the wrong value. In reference [9], the minimum icing time is used as the delay size, which is reasonable. In this paper, the method of automatic calculation by model training samples will be adopted to select the time delay within a certain range that can make the model predict the highest rate of return value, and then transferto the target samples together as model parameters.

In addition, when the icing model is applied to a practical project, the time delay represents the reaction time of the icing phenomenon predicted by the model. If the delay is too large, it means that the time required for the model to judge is about long, that is, the response of the prediction model is poor, and the real-time performance will be low.

IV. PERFORMANCE EVALUATION

4.1 Experimental Process

The 15# wind turbine was selected as the training sample data to train the icing model. After that, the model is transferred to the 21# wind turbine to test the model transfereffect (the evaluation of the transfereffect will be discussed in Section 4.3). The experimental process is shown in the figure below.

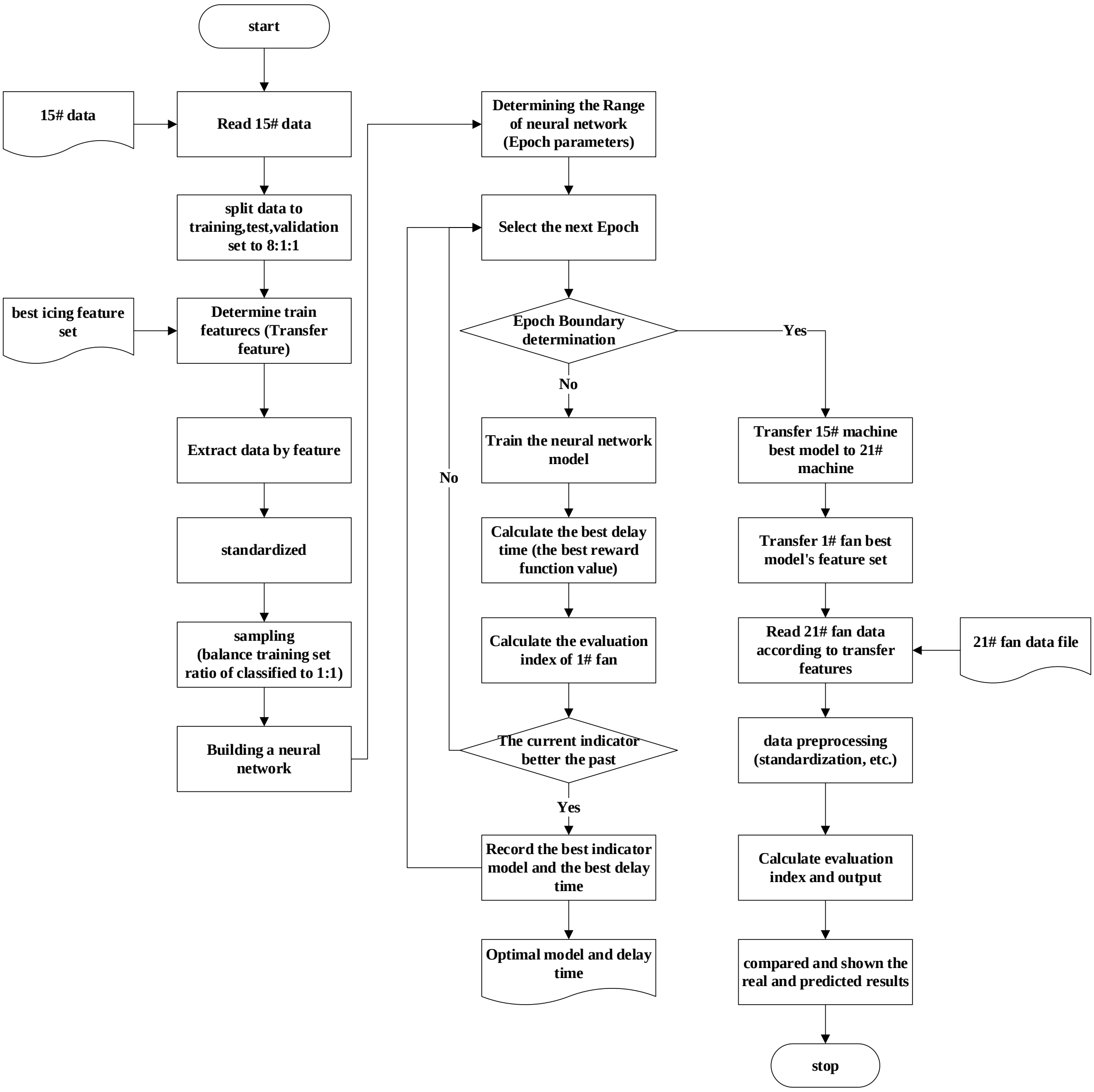
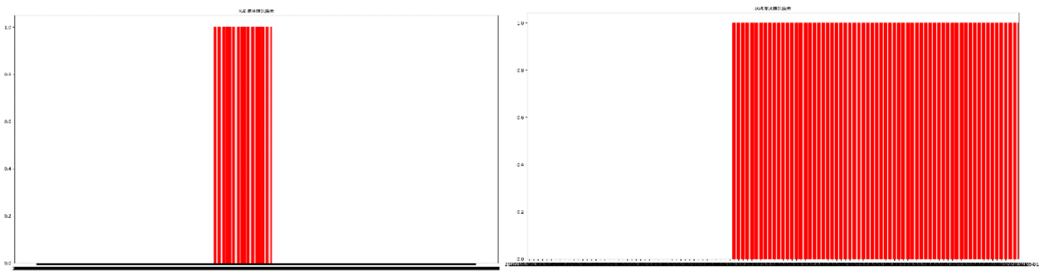


Fig. 10 Experiment process of model

4.2 Display method

Vision is the first sense of human beings, and graphic display is the most acceptable way. The wind turbine only has two states of icing and non-icing, so the time can be the horizontal axis, the icing is the value 1, and the non-icing is the value 0. Drawing a simple column diagram can intuitively show the icing situation of the wind turbine in a certain period of time.

Figure 11 below is a bar chart showing the actual ice covering of a wind turbine over a certain period of time (5000 monitoring points, approximately 11 hours). Figure 10-B is an enlarged view of Figure 10-A before and after freezing.



A B(Local amplification)

Fig. 11 Partial magnification of column chart of blade icing

This simple bar chart can visually show the icing situation of the wind turbine over a period of time. If the real icing situation of the wind turbine is displayed at the same time as the predicted situation, important indexes such as prediction accuracy can be visually compared.

Figure 12 below shows the comparison between the real icing situation of a wind turbine and the predicted situation of a model.

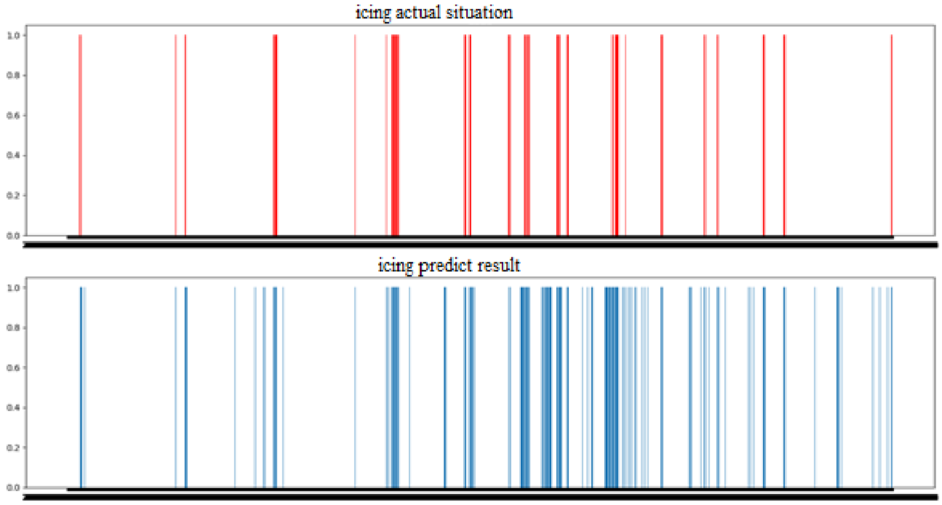


Fig. 12 Comparison chart between actual data and prediction ones

In order to compare model performance more intuitively, the above bar chart is undoubtedly a more appropriate way.Follow-up experimental results will be visually demonstrated and compared by this method.

3.3 Evaluating indicator

(1)accuracy

In general classification prediction problems, the accuracy of prediction is used to measure the advantages and disadvantages of the model.

$$P = \frac{N_{corr}}{N_{sum}} \times 100\% \quad (6)$$

P:accuracy；

N_{corr} :Predict the right amount;

N_{sum} :summer.

(2) Rate of return

Accuracy is only a general measure of prediction performance, which is too simple to reflect the overall quality of prediction. For example, Figure 4 is the comparison between the prediction result of a certain model and the actual result of the 21# wind turbine. The accuracy of this prediction reached 90%, but it can be clearly found from Figure 4 that the prediction result is not ideal. For large quantities of data, sample distribution has a crucial impact on accuracy. The 21# wind turbine sample shown in Figure 4 has 94% normal data and only 6% icing data. This means that in extreme cases, as long as the prediction results are all normal, the accuracy rate will be 94%. Therefore, accuracy alone cannot reflect the quality of prediction results. To this end, literature [9] designed a return function to measure the quality of prediction results.

$$p_{rew} = \frac{N_{pre\_0}}{N_{sum\_0}} \times 50\% + \frac{N_{pre\_1}}{N_{sum\_1}} \times 50\% \quad (7)$$

P_{rew} :rate value of return;

N_{pre_0} :The exact number of categorical values with 0 in the prediction result;

N_{sum_0} :The number of categories with 0 value in the sample;

N_{pre_1} :The exact number of classification value 1 in the prediction results;

N_{sum_1} : The number of categories with 1 value in the sample.

Formula (7) is to calculate the accuracy of different categories respectively, and then classify them into a reported result value according to the weight.

Obviously, the return function tries to avoid the influence of the sample classification quantity ratio on the prediction result, and can reflect the quality of the prediction result more than the simple accuracy. Therefore, in the subsequent experiments, the return function value will be used as an important indicator term.

(3) miss rate

Icing is an important type of wind turbine alarm. For alarm system, false alarm rate and false alarm rate are two of the most important indicators in practical work. In view of the serious impact of icing on the operation of the wind turbine, if the alarm is missed, it may cause serious consequences. Therefore, the prediction model should predict all icing alarms as far as possible.

$$P_{Miss} = 1 - \frac{N_{pre\_icing}}{N_{icing}} \times 100\% \quad (8)$$

P_{Miss} : miss rate;

N_{pre_icing} :Predicted ice cover times;

N_{icing} : Total icing times in the sample.

In the subsequent experiments, the missing rate will be counted。

V. THE EXPERIMENTAL RESULTS

In order to test the transferability of the model, this experiment not only compares the transferresults of neural network and deep neural network, but also compares the transferresults of the training model based on all features and the training model based on transferfeatures, as well as the transferresults after taking appropriate delay. In addition to the key indicators such as prediction accuracy, return value and missing rate, the bar chart is also used for intuitive comparison.

5.1 Neural networks and deep neural networks

Taking the 15# wind turbine data as the training data, a better model of neural network and deep neural network can be obtained. After that, the model is transfered to the 21# wind turbine respectively for prediction, calculation of main indicators, and comparison with the actual value.

The obtained good neural network model is used to predict 15# and 21# wind turbines, and the main indicators are shown in Table 4 below.

Table 5 Main indicators of prediction results by using NN model

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	87.61	86.88	87.59	87.24	0
21#	91.98	63.12	92.16	77.64	7.69

The bar chart is used to compare the actual results of the 21# wind turbine with the predicted results of the transfermodel, as shown in Figure 13 below.

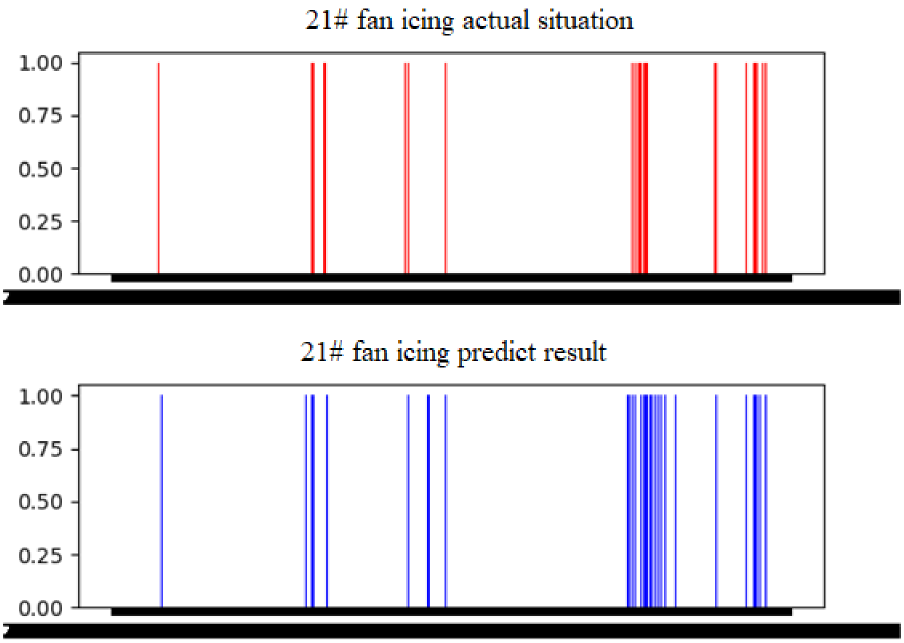


Fig. 13 Comparison chart between actual situation of 21# wind turbine and prediction results by using NN model

The obtained good deep neural network model is used to predict 15# and 21# wind turbines, and the main indicators are shown in Table 5 below.

Table 5 Main indicators of prediction results by using DNN

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	91.33	93.50	91.19	92.34	4
21#	90.84	65.79	92.42	79.10	0

The bar chart is used to compare the actual results of the 21# wind turbine with the predicted results of the transfermodel, as shown in Figure 14 below.

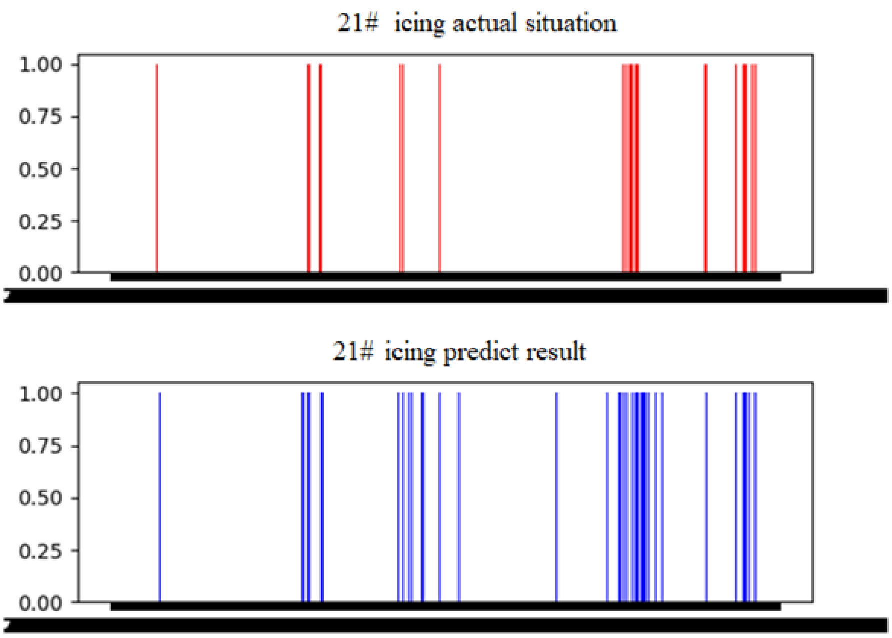


Fig. 14 Comparison chart between actual situation of 21# wind turbine and prediction results by using DNN model

5.2. Transfer characteristics

Literature [9] points out that the features in Table 1 have a great influence on icing, and relatively good results can be obtained when linear logistic regression model is used for prediction. Therefore, the above characteristics can be transferred to obtain better experimental results.

After the features in Table 1 are transfered, a new training model can be obtained by using deep neural network technology with good transferperformance. Then, the obtained model is used to predict 15# and 21# wind turbines respectively. The main indicators and comparison figures are shown in Table 6 and Figure 15 below, respectively.

Table 6 Main indicators of prediction results by using DNN model trained from transferred variables

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	89.02	95.05	88.93	91.99	0
21#	92.29	73.31	93.65	83.48	0

By comparing the results in Table 6 and Table 5, it is found that the prediction accuracy, rate of return and false report rate are all improved to different degrees after the transfer feature is adopted. This indicates that the features given in these references [9] are not only valid and the conclusion is correct, but also can be transfered from one wind turbine to another wind turbine. It shows that these transferable characteristics are the cause or performance factors of wind turbine icing.

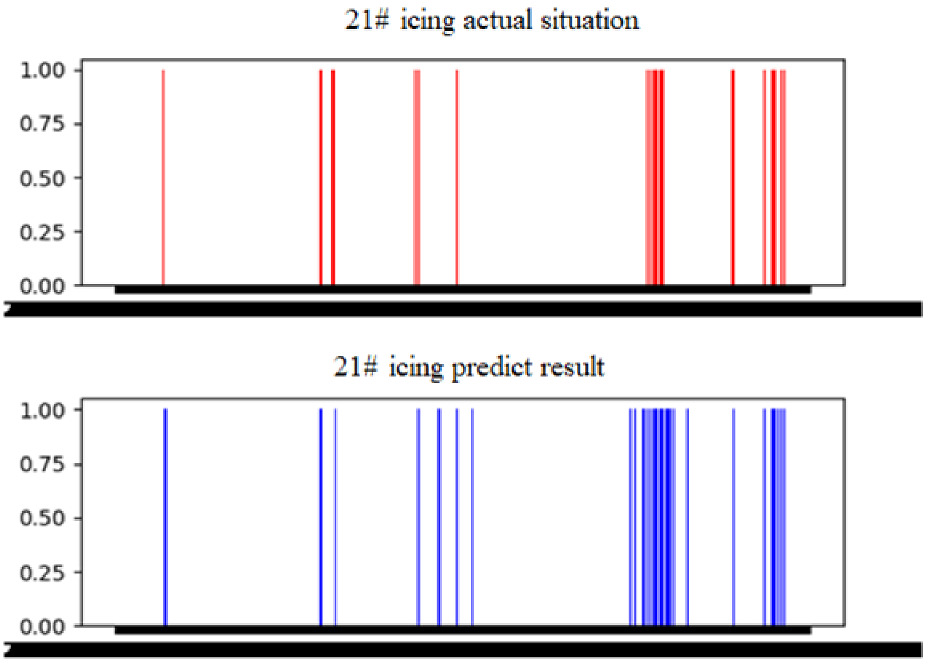


Fig. 15 Comparison chart between actualsituation of 21# wind turbine and prediction results by using DNN model trained from transferred variables

If the transferfeatures in Table 1 are combined with the monitoring points in SCADA, a full feature is constructed, a full feature model is trained again, and then transferred to the 21# wind turbine to test whether the experimental results will be better, which is worth discussing.

The experimental results of full features are shown in Table 7 and Figure 15 below.

Table 7 Main indicators of prediction results by using DNN model trained from all variables

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	85.66	96.63	84.91	90.77	0
21#	90.94	72.32	92.12	82.22	7.69

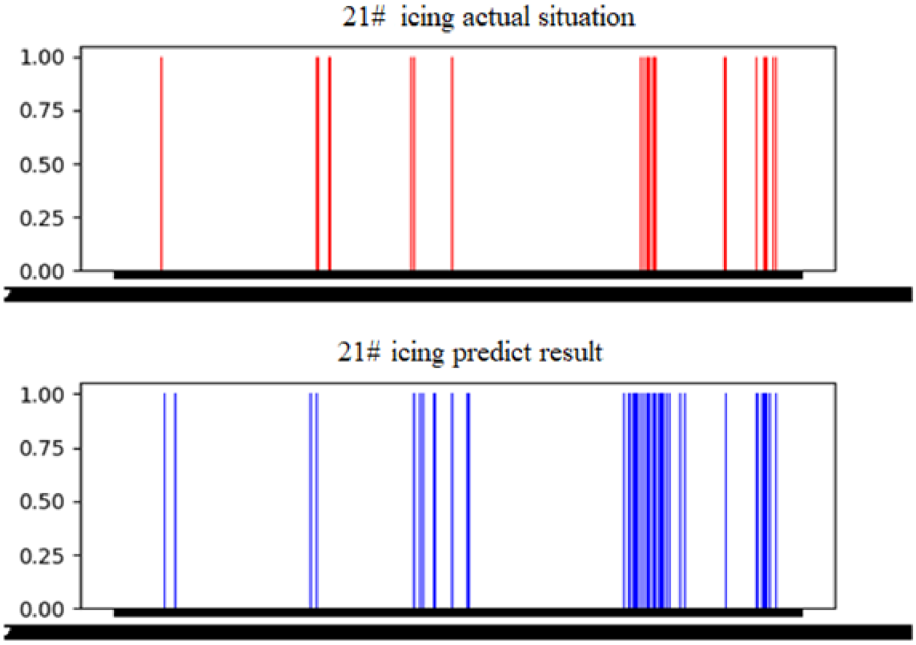


Fig. 16 Comparison chart between actual situation of 21# wind turbine and prediction results by using DNN model trained from all variables

By comparing the results of Table 6 and Table 7, it can be found that the rate of return does not rise but falls, and there is an under-reporting phenomenon after the transfer. This shows that the more features are not the better when using neural networks for training; When transferring, more features are not always better. If the appropriate characteristic parameters can be found, a more ideal model can be obtained. This aspect confirms the conclusion of literature [9].

5.3 Delay Function

Reference [9] points out that since icing is an inertial process, better prediction accuracy can be obtained by using time delay. However, the literature does not indicate how to obtain better delay. Using the algorithms in Sections 3.6 and 3.7, it is possible to find a more cost-effective delay size.

Through experiments, the time delay interval for 15# to obtain the best prediction result is 107 monitoring points, which is about 15 minutes.

After the delay parameter is added to adjust the prediction result, the main indicators are shown in Table 8 below. The comparison between the adjusted prediction result and the actual result is shown in Figure 11 below.

Table 8 Main indicators of prediction results by using time delay

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	93.25	99.17	92.85	96.01	4
21#	93.46	73.95	94.69	84.32	7.69

Compared with the data in Table 7, after adding the delay parameter, the accuracy rate and rate of return are greatly improved, but the false alarm rate deteriorates to a certain extent, and some icing alarms are eliminated by "delay".

5.4The adjusted icing model

The above model can be further tuned to further improve the parameters on the premise of keeping the model transferability.

According to the experience provided by field engineers, when the wind turbine is iced, the blade Angle of the wind turbine will inevitably change. In the feature variables of icing, the monitoring points related to leaf Angle (Pitch1_angle, PitCH2_angle, Pitch3_angle) are added to Table 1 as the new feature variable group, namely, the leaf Angle feature vector is added, and the method of Section 3.6 is again used to construct the deep neural network prediction model. The results obtained this time are relatively ideal, and the main parameters are shown in Table 9 below.

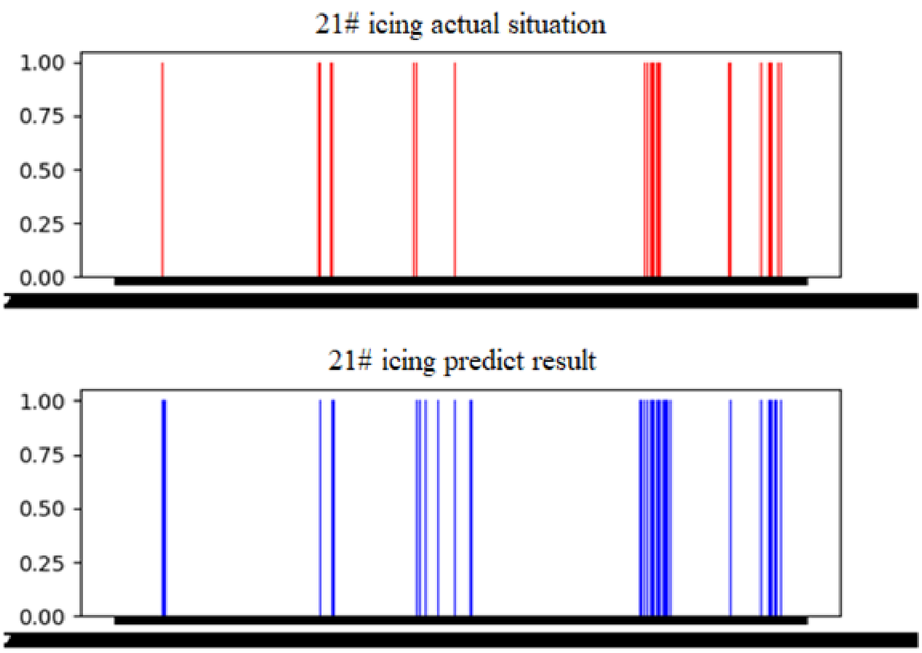


Fig. 17 Comparison chart between actual situation of 21# wind turbine and prediction results by using time delay

Table 9 Main indicators of prediction results after tuning

wind turbine	accuracy	Reward rate			miss rate
		Icing accuracy	Normal accuracy	Reward rate	
15#	91.63	98.64	91.20	94.92	0
21#	89.97	88.79	90.23	89.51	0

Relative to the previous tables, it is not difficult to find, in addition to accuracy, the rate of return has been greatly improved. This means that the prediction of the normal state and the ice-covered state has been improved, especially the prediction accuracy of the ice-covered state has jumped forward and reached more than 88% after the transfer. This improvement shows that when icing occurs, the ice on the blade has an obvious influence on blade Angle, which is consistent with the experience of field engineers.

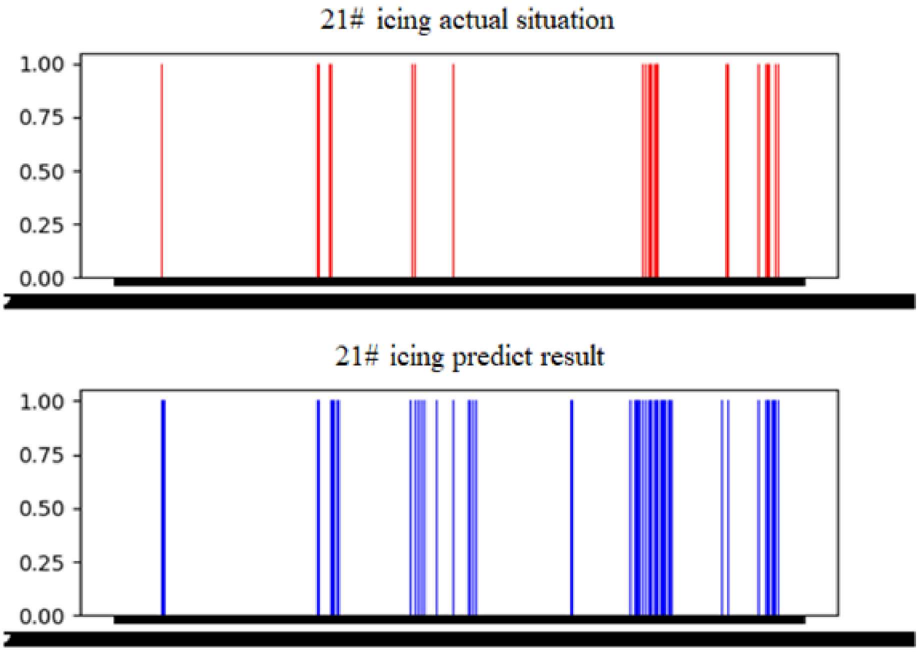


Fig. 18 Comparison chart between actual situationof 21# wind turbine and prediction results by using tuned model

5.5About the false alarm rate

False alarm rate is an important index of alarm system. If false alarms occur frequently in the system, the system gives a warning when no alarm actually occurs, which will reduce the practical performance of the alarm system. However, if blindly accurate, it is likely to under-report. This is, in a way, contradictory. By contrast, underreporting is intolerable.

From the comparison of the above images, it can be seen that while trying to solve the problem of missing alarms, false alarms always exist, and the false alarm rate is not low (it is not difficult to find this problem by comparing the upper and lower figures).

In practice, in order to deal with the problem of false positives in the system, the "video linkage" measure can be generally adopted. When the system determines that a wind turbine has icing, the engineer in the central control room needs to confirm again with the help of the camera attached to the wind turbine to eliminate false positives.

VI. CONCLUSION

Compared with the feature vectors formed by each monitoring point in SCADA, the prediction of wind turbine icing is highly nonlinear. Due to the many uncertainties involved, it has been difficult to obtain high precision prediction. Many researchers at home and abroad basically focus on the method of model acquisition, and have made some progress. However, the prediction models obtained by various machine learning methods often have high performance for the wind turbines from the training samples, with good accuracy, rate of return and false report rate. However, for other wind turbines, the prediction performance is rarely mentioned or unclear.

Aiming at the problem that the training model is difficult to be extended to other wind turbines, this paper not only describes how to train the wind turbine icing model based on deep neural network, but also focuses on how to build a transferable model using transfer learning technology to improve the transfer performance of the model. Aiming at the problem that the prediction of wind turbine icing is difficult, the problem of what to transferand how to transferis solved. The icing transferalgorithm based on DNN is designed creatively. In addition, to establish the evaluation index of the ice forecast model, points out that only by the traditional forecasting accuracy is not enough, can't reaction model performance, returns indicate the classification forecast also need to improve the quality of model prediction, need non-response rates as guidance, to show that forecasting model is feasible, whether has the engineering application value. Based on the calculation results, it can be concluded that the proposed method provides an effective method for the comprehensive prediction of wind turbine icing in wind farms due to the comprehensive consideration of the accuracy, quality and practical applicability of the prediction.

Next steps could include: I) considering other data, such as video images, as input to further improve accuracy and utility; Ii) How to reduce the false alarm rate of the model; Iii) Apply the model to the actual wind field to test the practicability and transfereffect of the model.

REFERENCES

[1]. KeSun.The power industry is the main battlefield of energy transformation under the goal of carbon peak and carbon neutralization.[EB/OL]. (2021-08-24) [2021-09-25].http://www.chinasmartgrid.com.cn/news/20210824/639568.shtml

[2]. Dongjun Zhang, Haiyan Qin. wind power will get broad prospects with helping "double carbon" green development[EB/OL]. (2021-07-08) [2021-09-25]. https://news.bjx.com.cn/html/20210708/1162787.shtml.

[3]. LIU Taowen, MENG Wenchuan, DAI Chengwei, et al. An Overview of Wind Turbine Generator Anti-Freezing and Deicing[J]. Hubei Electric Power, 2019, 43(01):14-21.

[4]. LI Cheng. Discussion on the Influence and Countermeasures of Wind Turbine Icing[J]. Science and Technology Innovation Herald, 2019, 15:93-94.

[5]. JIANG Xingliang, ZHOU Hongyu, HE Kai, et al. Anti-icing Performance Degradation for Wind Blades with Superhydrophobic Coatings[J]. High Voltage Engineering, 2019, 15:93-94.

[6]. XIONG Changquan, HE Zeqi, ZHANG Yuning, et al. Short-term Icing Status Prediction Model of Wind Turbine Blades Based on Bi-LSTM and SVM Models[J]. Sichuan Electric Power Technology, 021, 44(03): 88-94.

[7]. Liu Y, H Cheng, Kong X, et al. Intelligent wind turbine blade icing detection using supervisory control and data acquisition data and ensemble deep learning[J]. Energy Science & Engineering, 2019, 7(6).

[8]. Chi H, Peng C, He J, et al. Short Term Ice Prediction by Using SCADA Data on Wind Turbine Blades[J]. China Computer & Communication, 2019.

[9]. Li Tingshun, Tan Wen, Liu Zesan. A Hybrid Model Based on Logistic Regression Algorithm and Extraction Algorithm Using Reward Extremum to Real-Time Detect Blade Icing Alarm[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2019(36).

[10]. LI Hang. Statistical Learning Methods[M]. Tsinghua University Press.

[11]. LI Nianhua. Research on Key Problems of Transfer Learning in Deep Neural Network[D]. University of Electronic Science and Technology of China, 2018.

[12]. QIU Ningjia, WANG Xiaoxia, WANG Peng, et al. Research on Convolutional Neural Network Algorithm Combined with Transfer Learning Model[J]. Computer Engineering and Applications, 2020(5):43-48.

[13]. Pandeya Y R , Lee J . Domestic Cat Sound Classification Using Transfer Learning[J]. International Journal of Fuzzy Logic & Intelligent Systems, 2018, 18.

[14]. WEN Kunzhe, WEI Yuke, DONG Xiaohua. Survey of Application of Deep Convolution Neural Network in Image Aesthetic Evaluation[J]. Computer Engineering and Applications, 2019, 055(015):13-23,58.

[15]. Zhao Chongwen. A Survey on Artificial Neural Network[J]. Shanxi Electronic Technology, 2020, 210, 03:96-98.

[16]. LI Tingshun, Wang Wei, Liu Zesan. GELM-WNN Forecasting Method of Power Load Considering Uncertain Interval[J]. Computer Engineering, 2019, 45(7):315-320.

[17]. CHEN Yuting, Zhang Fengyu. Research on CNN Traffic Sign Recognition Method Based on Cross E ntropy[J]. Software Guide, 2018, 17(12):28-32.

[18]. DANG Jianwu, CONG Xiaoqing. Research on Hybrid Stock Index Forecasting Model Based on CNN and GRU[J]. Computer Engineering and Applications, 2021,57(16):167-174.

[19]. Srivastava N , Hinton G , Krizhevsky A , et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting[J]. Journal of Machine Learning Research, 2014, 15(1):1929-1958.

Wen Tan, et. al. "Wind Turbine Blade Icing Detection Via Deep Neural Network and Transfer Learning."*American Journal of Engineering Research (AJER)*, vol. 11(09), 2022, pp. 26-43.