sustainability

MDPI

Review

Review of Data-Driven Approaches for Wind Turbine Blade Icing Detection

Chang Cai [1], Jicai Guo [2,3], Xiaowen Song [2,3,*], Yanfeng Zhang [2,3], Jianxin Wu [2,3], Shufeng Tang [2,3], Yan Jia [4], Zhitai Xing [4] and Qing'an Li [1,*]

[1] Institute of Engineering Thermophysics, Chinese Academy of Sciences, Beijing 100190, China
[2] College of Mechanical Engineering, Inner Mongolia University of Technology, Hohhot 010051, China
[3] Inner Mongolia Key Laboratory of Special Service Intelligent Robotics, Hohhot 010051, China
[4] College of Energy and Power Engineering, Inner Mongolia University of Technology, Hohhot 010051, China
* Correspondence: sxw_istgut@imut.edu.cn (X.S.); liqingan@iet.cn (Q.L.)

Abstract: Onshore wind turbines are primarily installed in high-altitude areas with good wind energy resources. However, in winter, the blades are easy to ice, which will seriously impact their aerodynamic performance, as well as the power and service life of the wind turbine. Therefore, it is of great practical significance to predict wind turbine blade icing in advance and take measures to eliminate the adverse effects of icing. Along these lines, three approaches to supervisory control and data acquisition (SCADA) data feature selection were summarized in this work. The problems of imbalance between positive and negative sample datasets, the underutilization of SCADA data time series information, the scarcity of high-quality labeled data, and weak model generalization capabilities faced by data-driven approaches in wind turbine blade icing detection, were reviewed. Finally, some future trends in data-driven approaches were discussed. Our work provides guidance for the use of technical means in the actual detection of wind turbine blades. In addition, it also gives some insights to the further research of fault diagnosis technology.

Keywords: wind turbine; SCADA; data-driven approaches; blade icing detection

check for updates

Citation: Cai, C.; Guo, J.; Song, X.; Zhang, Y.; Wu, J.; Tang, S.; Jia, Y.; Xing, Z.; Li, Q. Review of Data-Driven Approaches for Wind Turbine Blade Icing Detection. *Sustainability* **2023**, *15*, 1617. https://doi.org/10.3390/su15021617

Academic Editors: Doug Arent, Adam Warren and Xiaolei Yang

Received: 15 November 2022
Revised: 5 January 2023
Accepted: 11 January 2023
Published: 13 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid growth of the economy and population, global energy consumption is expected to be doubled by 2050, and more severe energy challenges will arise in the future [1]. For the sustainable development of humanity and the protection of the ecological environment, China has already formulated the goal of carbon neutrality in 2060. Simultaneously, some of the world's largest countries are also beginning to take the concept of carbon neutrality seriously and practice it. To solve the above-mentioned problems, it is urgent to develop renewable energy sources. Under this perspective, wind energy is regarded as the most widely used renewable energy in the world. The total number and annual growth of global wind turbine assemblers since 2015 are depicted in Figure 1 [2]. As can be observed, from 2015 to 2021, the global total installed capacity continued to grow steadily, with a cumulative global installed capacity of 837 GW by 2021. From 2015 to 2019, the global annual growth of wind turbines was about 50 GW. After 2019, the annual growth of wind turbines reached 90 GW. Thus, it is foreseeable that wind energy will play a more critical role in the future.

To make full use of wind energy, wind farms are usually located in high-altitude areas, which are normally far away from human communities. Within 1000 m above sea level, the wind speed is increased by approximately 0.1 m/s for every 100 m of altitude gain, while the wind energy is proportional to the cube of the wind speed [3]. With the increase in altitude, the temperature will also decrease, and the density of the cold air is higher than that of the warm air, as a result, the wind energy potential in the alpine region is 10% higher than in other areas [4]. However, it is challenging to operate wind turbines in

areas over 1000 m, and some problems are often encountered, such as the failure of the measuring instruments [5]. In winter, wind turbines in alpine regions are easily exposed to low-hanging clouds in which blades are very prone to icing. A close-up of a wind turbine blade icing up is displayed in Figure 2 [6]. Nowadays, about one-third of the wind turbines around the world operate in cold environments [7], and most of them suffer from the blade icing problem. The most intuitive impact of blade icing is the reduction of the power generation of the wind turbines. Pérez et al. [8] analyzed 517 wind turbines in Spain and found that the power loss caused by the blade icing effect reached as much as 18 GWh in 29 months, which was estimated to be 550 GWh of the annual power loss due to blade icing over the country, that is equivalent to the electricity consumption of 200,000 households. In addition, Lehtomäki et al. [9] studied the wind turbine data at two sites in Canada and Sweden and concluded that blade icing increased the transverse fatigue load on the tower base and affected the service life of the wind turbines. Otherwise, the icing of the blades could increase noise pollution and affect the life of the surrounding staff. On top of that, ice shed and ice throw may destroy equipment, even causing people injury and death [10]. Therefore, early detection of blade icing is considered of great importance in ensuring the safe operation and power production of wind turbines.

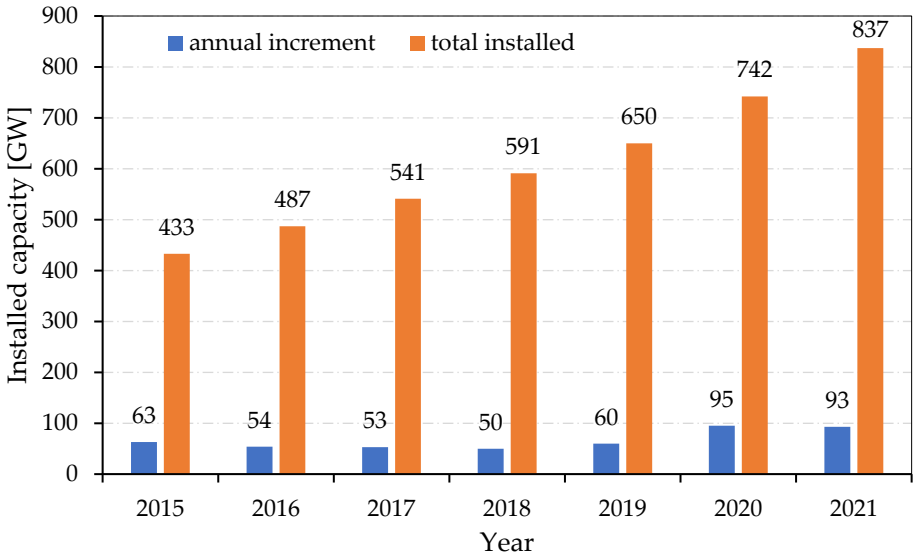

Figure 1. Global total installed wind turbines and annual growth from 2015 to 2021. [2].



Figure 2. Close-up of wind turbine blade icing (adapted from [6]).

At present, the icing environments of wind turbines are defined as when the temperature is below 0 °C and the relative humidity is over 95% [11]. However, Laakso et al. [12] found that in most cases, when the temperature was lower than the value of 0 °C, and the relative humidity was higher than 95%, the blades did not ice. In actual operation, the deviation between the actual power and the theoretical power of the wind turbine is used by the staff as the criterion for blade icing. When the deviation reached a specific value (i.e., 15% in usual), an alarm was triggered, and the wind turbine was then stopped. However, when the alarm was triggered, there was already a relatively large area of ice on the blade [13,14]. Although many wind turbines have been equipped with de-icing systems, the challenge in practice is to detect early the icing effect accurately and turn on the de-icing system in time. Later, the implementation of sensors was reported in the literature to detect changes in the physical properties of blades to determine whether the blades were iced. More specifically, Berbyuk et al. [15] used controlled acoustic waves propagating in the wind turbine blades to detect early icing. Interestingly, when the blade starts icing, the Fourier transform, amplitude attenuation, and the root mean square of the guided wave propagation changed significantly. Muñoz et al. [16] proposed a non-contact infrared sensor method to examine if blades were iced since the emitted thermal radiation signals by the blade surface are significantly different from those emitted during non-icing conditions. Kim et al. [17] detected the icing condition of the blade by arranging optical fiber optic sensors on the blade. The transition in Fresnel reflection due to the icing effect gives unusual intensity variations, which are used to identify the ice on the wind turbine blades. Although the above-mentioned methods are highly accurate in detecting blade ice, they are complex and highly dependent on expertise. Moreover, the majority of wind turbines in service do not have those specific sensing elements and often require additional installation costs. Therefore, these specialized ice identification methods are not used on many large-scale wind turbines in practice.

In recent years, SCADA systems have evolved the standard configuration of large-scale wind turbines. Moreover, the SCADA system contains many state monitoring data of the wind turbine's components or subsystems, such as environmental parameters, working conditions, and control parameters, which can comprehensively monitor the state of the wind turbines through an economical and convenient method [18]. Furthermore, improvements in data mining techniques have made it possible to detect blade icing based on the application of data-driven approaches.

Therefore, this paper reviewed and summarized the recent research on the data-driven approaches diagnosis of wind turbine blade icing. In the second section, the data-driven approaches and the feature selection were introduced and data imbalance issues were considered when modelling. In the third section, the time series information of SCADA data was explored. In the fourth section, the focus led on the semi-supervised learning methods, as well as the implementation of unlabeled data. Finally, in the fifth section, the application of the transfer learning method in blade icing detection was introduced.

2. Data-Driven Approaches in Blade Icing Detection

Data-driven fault diagnosis approaches is considered a method that can complete the fault diagnosis of the system by analyzing and processing the process operation data without knowing the accurate analytical model of the system [19]. Particularly, the data-driven turbine blade icing diagnosis approaches mainly focuses on the use of artificial intelligence algorithms to learn hidden information in the historical data to detect blade icing. According to the output variables, the data-driven icing diagnosis research can be divided into regression and classification tasks.

The regression task predicts state variables of the wind turbines, such as the active power on the grid side, the rotor speed, etc. The operating parameters closely related to the prediction variables are used as the model's inputs, while the normal SCADA data are utilized to train the model. Then, the model's predicted values are compared with the actual measured values, and the judging criteria are set according to the statistical knowledge or

algorithm. If the residual error between the predicted and the actual measured values does not meet the judging criteria, the unit may be in the icing state. Guo et al. [20] analyzed the influence of the blade icing effect on the operation parameters of wind turbines. More specifically, the power, rotor speed, and ambient temperature were selected to detect blade icing. The normal behavior models of power and rotor speed were constructed by using the extreme gradient Boosting (XGBoost) method, and the sequential probability ratio test (SPRT) was introduced to analyze the model prediction residuals to identify the abnormal changes in the output power and rotor speed. Blade icing warnings were issued when the residual difference between the predicted and actual values of power and rotor speed were detected to be above the threshold at the same time and the ambient temperature was below zero. The authors found that the proposed blade icing detection method can warn wind turbines 5 h before the shutdown moment by icing. The principle of this technique is illustrated in Figure 3. Similarly, Fan et al. [21] established power and rotor speed models by using by adaptive boosting (AdaBoost) algorithm and introduced the exponentially weighted moving average method to analyze residuals.

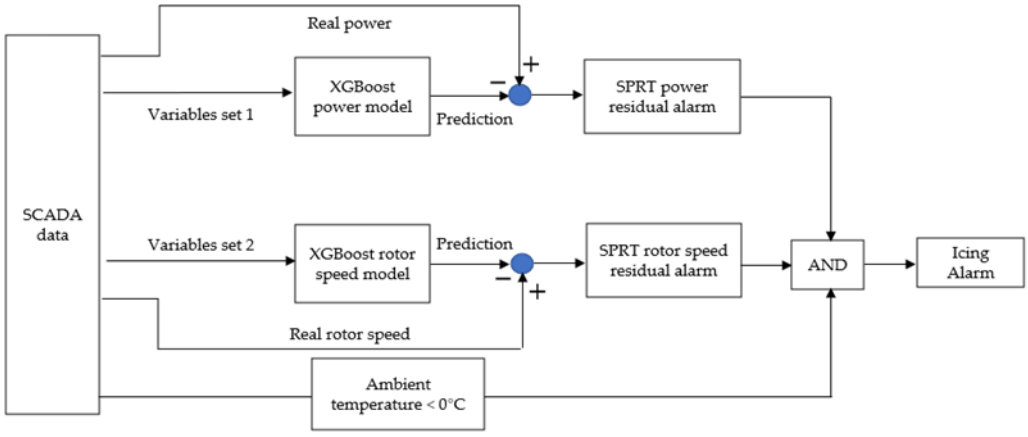

Figure 3. Principle of blade icing detection by regression task (adapted from [20]).

The operational parameters of the wind turbines predicted by the regression tasks are continuous variables. The classification task outputs are discrete variables that characterize if the SCADA data refers to icing. The typical evaluation indexes include accuracy, precision, recall, F1 score, etc. During the classification task, if the massive high-dimensional SCADA data are directly input into the learning model, too much redundant information will affect the model's performance. Moreover, most wind turbines work under normal conditions, resulting in a significant imbalance in the SCADA data. Therefore, feature selection and data imbalance must be considered in constructing the model.

2.1. Feature Selection Problem

In practice, the SCADA system records many state monitoring data of the wind turbine components or subsystems with hundreds of dimensions. Hence, the selection of the effective characteristics that can characterize the icing effect of wind turbines is a problem worthy of consideration. Although some variables have been selected in the literature according to relevant field knowledge, redundant information still prevents the model from achieving the desired effect. Because of the above-mentioned problems, the SCADA data feature selection has been examined in the literature. Li et al. [22] considered the actual situation of wind turbine icing by selecting two sensitive features, namely wind speed and power. The non-principal component direction value of the wind speed and power were the input of the classifier and by using logistic regression, the optimal model was found by setting different discriminant thresholds. Based on expert knowledge, Zhou et al. [23] selected the wind speed, power, environment temperature, and nacelle internal temperature from the SCADA data. The authors constructed four derived variables

including residuals of the normal fit speed and the generator speed, the residuals of the actual power and the theoretical power, the residual between the environment and the internal temperature, as well as the non-principal component direction values of the wind speed and power as the input of the support vector machine, whereas the particle swarm optimization was used to establish an icing detection model. Zhang et al. [24] selected wind speed and power as two essential characteristics and constructed six characteristics, such as: average wind speed (V); average power (P); the degree of deviation of P(C); the average rate of change of C (ΔC); the maximum value of C (max C); and the cumulative value of ΔC (sum ΔC), which were then used the random forest to detect blade iced. Additionally, Ma et al. [25] proposed a blade icing detection method based on a deep belief network (DBN) by selecting six different features, such as wind speed, power, and environment temperature, according to the relationship model between icing, and the wind turbine performance.

Nevertheless, the aforementioned method of manually extracting features is labor-intensive and strongly dependent on expert knowledge and skills [26]. The good or bad feature extraction affect also dramatically the performance of the model. Later, features were filtered based on algorithms. Ye et al. [27] proposed a feature screening method by combining the chi-square test and principal component analysis (PCA). The chi2 algorithm was used to select the top 10 feature parameters, and then the PCA algorithm was used to reduce the dimension of the feature parameters after the preliminary screening to reduce the redundancy between features. Besides, Peng et al. [28] extracted three features, yaw position, internal temperature, and generator speed, by using the highest correlation with wind turbine blade icing through dynamic PCA and Elman neural network. Then, they were input into the back propagation (BP) neural network for clustering judgment. From the experimental results, it was demonstrated that the prediction value of the Elman artificial neural network was better than that of support vector regression for the discrimination of the BP artificial neural network clustering. Li et al. [29] obtained the characteristic variables related to blade icing by using the extraction algorithm of the reward extremum and then a blade icing detection model was established based on a logistic regression algorithm from many historical monitoring data denoised by the wavelet transform method. Jia et al. [30] used a slow feature analysis (SFA) algorithm to extract the slow features of the original SCADA data as input from a support vector machine (SVM). The slow feature contained temporal information more in line with the wind turbine icing process. For a single wind turbine, when the number of slowly varying features was five, the model error rate in this paper's test set was 23.4%, while the minimum error rate of the SVM trained directly with the original SCADA data was 31.7%. It is also worth noting that if the data dimension was improperly selected, the model's performance worsened. When the feature dimension was reduced to three, the model's error rate reached 54.7%. Zhang et al. [31] proposed a feature extraction method based on MIV-tSNE. The mean impact value (MIV) algorithm was also used to eliminate variables with a MIV index of less than one. Then, the t-distributed stochastic neighbor embedding (tSNE) algorithm was employed to map the selected characteristic variables to specific dimensions as the input of the recurrent neural network. The model's accuracy on the test set can reach about 72% for a single wind turbine. Interestingly, compared to traditional recurrent neural networks (RNN), the prediction accuracy was improved by about 150%, and the model run-time was reduced by about 45%.

In addition, compelling features were extracted from the original SCADA data set based on domain knowledge and algorithms. Furthermore, Chen et al. [32] combined the mechanism model by using the decision tree for feature screening and selecting eight features, such as wind speed and power, and used a convolutional neural network (CNN) to construct an icing detection model for the wind turbine blades. Based on domain knowledge and the wrapper method, Di et al. [33] performed feature selection on the wind turbine SCADA data, and selected 10 features, such as power, generator speed, temperature difference environment and nacelle, and wind speed, while the XGBoost was applied to

construct a wind turbine blade iced detection model. Meng et al. [34] proposed a feature selection method based on the implementation of a recursive feature elimination -random forest (RF) method. According to the influencing factors of wind turbine icing and artificial experience, four different features of wind speed, power, internal nacelle temperature, and environment temperature were selected for the cross-feature construction. Then, the proposed method was used to screen the original and the newly constructed features, and the top 12 feature attributes were selected as the model's input. From the experimental results, it was suggested that the accuracy of the model test set increased after feature selection. Hai et al. [35] combined the Makkonen and the rotating object icing models. After performing PCA analysis, the wind speed (v) and the power, environment temperature, $v_{\omega}^3 + 1/2v_P^2$ (v_w refers to the directional component of the velocity of the water droplet on the surface of the accumulated ice in the rotating coordinate system normal to the surface, and v_P represents the rotating speed for wind blade), and the P_a/v^3 (P_a stands for the work absorbed by the blades from the wind energy) were selected as inputs to the support vector machine for the particle swarm optimization. Nie et al. [36] selected four essential features, such as wind speed and power, and the original feature set was expanded by feature crossover. The recursive feature elimination algorithm was also used to select 19 features with high weight scores, such as the wind speed and environment temperature, as the model's input. The data coupling was significantly reduced by PCA, and the improved sparrow search algorithm was introduced to establish the long short-term memory (LSTM) network icing detection model. Tang et al. [37] formulated 13 sensitive features, such as the wind speed power ratio and blade angle average, by performing a qualitative analysis of the icing mechanism and merging them with 26 features of the original dataset. Then, they were the recursive feature elimination cross validation algorithm was used to select 15 features, such as the generator speed, active power, and yaw position as the input of the fully connected neural network.

As far as feature engineering problems are concerned, the reviewed works in the literature can be classified into three types: mechanism-based feature selection, algorithm-based feature selection, and mechanism+algorithm-based feature selection. Table 1 displays a categorization of these works by their methodological method.

Table 1. Categorization of feature selection solutions by methodology.

Methods	References	Base Features	Advantages	Disadvantages
Mechanism-based method	Li et al. [22]	Wind speed, power, etc.	The extracted features are highly accurate and relatively simple to model.	Manual extraction of features is labor-intensive and overly dependent on domain knowledge.
	Zhou et al. [23]	The residuals of actual power and the theoretical power, etc.		
	Zhang et al. [24]	Degree of deviation of P(C), the maximum value of C (max C), etc.		
	Ma et al. [25]	Wind speed, power, environment temperature etc.		
Algorithm-based method	Ye et al. [27]	Generator speed, pitch angle, motor temperature, etc.	Selecting features from the algorithm level reduces labor and decreases model complexity and run time.	May be losing useful information.
	Peng et al. [28]	Yaw position, internal temperature, generator speed, etc.		
	Li et al. [29]	Yaw position, Wind Speed Face (Upwind speed), pitch angle, etc.		
	Jia et al. [30]	Dimensionality reduction of raw 26-dimensional SCADA data using the SFA algorithm works best at dimension 5.		
	Zhang et al. [31]	Pitch1_ng5_tmp, Generator speed, Pitch2_moto_tmp, etc.		
Mechanism+algorithm-based method	Chen et al. [32]	Angle_avg (the average of three pitch angle), etc.	Avoid the impact of manual feature selection and the loss of some information, more comprehensive consideration, required.	More labor-intensive, potentially increasing the complexity of the model.
	Di et al. [33]	Power, generator speed, etc.		
	Meng et al. [34]	Wind speed, power, internal nacelle temperature, environment temperature, etc.		
	Hai et al. [35]	Wind speed, power, environment temperature, etc.		
	Nie et al. [36]	Wind speed, power, etc.		
	Tang et al. [37]	The wind speed power ratio, blade angle average, etc.		

2.2. Data Imbalance Problem

In the supervised learning mode, the classification tasks need to learn the hidden information of the normal and iced data in the training stage. However, most of the working time of the unit is in the normal state, and only a few of the time is in an icing state. Therefore, the data recorded by SCADA show a great imbalance. The model trained with unbalanced data gives a biased prediction [38]. The sample imbalance of SCADA data has been also explored in the literature, and the solutions can be roughly divided into two categories. One is data-based, such as large sample under-sampling or small sample over-sampling. In the previously reported work, Chen et al. [32] eliminated the influence of sample imbalance by using 28 s average under-sampling normal samples. Ye et al. [27] used the synthetic minority over-sampling technique (SMOTE) algorithm to oversample a few samples for the problem of sample imbalance, whereas the Tomek links algorithm was employed to clean the noise samples and boundary samples in the data set. Finally, the authors used the random forest algorithm to establish the icing classification model. On top of that, Xu et al. [39] first removed some samples that were obviously not iced according to the actual situation of the blade icing effect, and then an under-sampling algorithm was proposed based on the similarity function to remove the redundant normal samples. The SMOTE algorithm was also to generate iced samples to ensure data balance. Similarly, Zhang et al. [24] utilized the SMOTE algorithm to increase the number of iced samples.

Since the SMOTE algorithm can synthesize a few classes of samples indiscriminately, it leads to some redundancy in the synthesized samples, which can induce the boundary confusion problem. For this reason, some novel oversampling algorithms have been introduced for solving the data imbalance problem. Ge et al. [13] proposed an improved SMOTE algorithm to conquer the data imbalance. More specifically, by introducing the concept of sample density and dividing the concentration region to balance the sample distribution, the mini-batch k-means algorithm (MBK-SMOTE) can avoid the redundancy of the synthesized samples by using the SMOTE algorithm. The author used the random forest to construct a model to detect when the wind turbine blades are iced. It was found that the prediction effect of MBK-SMOTE was better than that of the other SMOTE methods. Peng et al. [40] also proposed an over-sampling balancing algorithm, namely the safe circle synthetic minority oversampling technology (SC-SMOTE) to optimize the data set, and an improved k-nearest neighbor (KNN) method was used to construct a wind turbine blade icing fault diagnosis model. By carrying out extensive experiments, the authors proved that compared with the SMOTE algorithm, the SC-SMOTE algorithm had a better fault diagnosis accuracy, which was 10.3% better than the SMOTE algorithm. Tang et al. [37] utilized SVM-SMOTE to oversample a small number of iced samples and randomly under-sampling most normal samples to balance the data set of the iced data with normal data. Jiang et al. [41] used the adaptive synthetic method to oversample the wind turbine blade icing (minority class) samples and improve the unbalanced data set of wind turbines. The other was algorithm-based, such as using cost-sensitive or ensemble learning to reduce the impact of data imbalance. In addition, Chen et al. [42] proposed a deep neural network based on triplet loss. This method mapped SCADA data to a hypersphere, and after triplet loss learning, the distance of similar data was near but further for different data in the reconstruction space. By using such processing methods, the impact of data imbalance can be significantly reduced. The local and global features of the data were also considered to introduce skip connections in the deep neural networks (DNN) model, and the SCADA data of two other wind turbines were verified. The accuracy of this model can reach more than 80%, the recall rate reached about 60%, and the overall performance of the model was better than the traditional artificial neural network. The model framework is depicted in Figure 4.

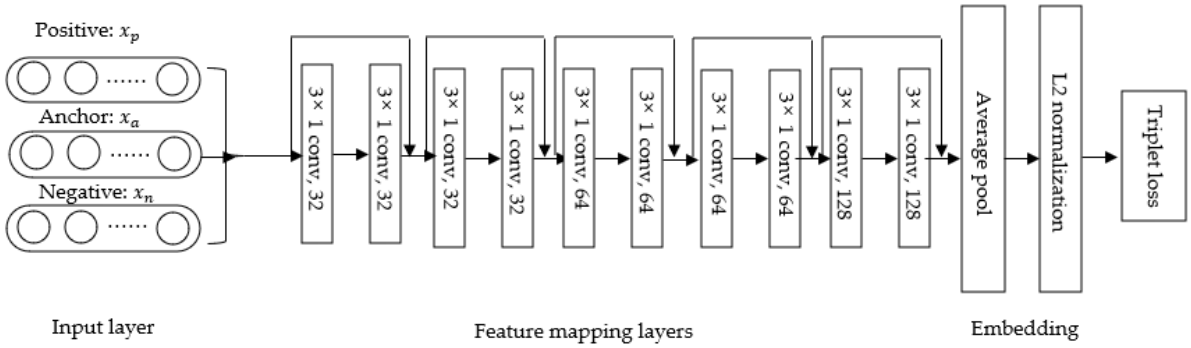

Figure 4. Deep neural network based on triple loss (adapted from [42]).

In addition, Peng et al. [43] introduced a focal loss function to replace the traditional cross-entropy loss function. As a result, the impact of the data imbalance was weakened by reducing the weight of the simple sample loss during the training process since focusing on the training of samples with less data that are difficult to predict. Tong et al. [44] proposed a novel adaptive weighting strategy to eliminate the influence of data imbalance from the perspective of sample distribution information by using the support vector data description method. The authors combined it with the traditional fixed weighting strategy to develop an adaptive weighted kernel extreme learning machine (AWKELM) algorithm. Moreover, the superiority of the AWKELM algorithm was verified on two wind turbines, with a 23% improvement in the F1 score performance over WKELM. Ding et al. [45] combined the ideas of under-sampling and ensemble learning to under-sample normal data to obtain several subsets and assign icing data to each subset so that each subset was class balanced. Then, several classification models were obtained by training each subset using temporal convolutional networks (TCN). Finally, the prediction results of all models were averaged as the final prediction results of the samples. The model framework is displayed in Figure 5.

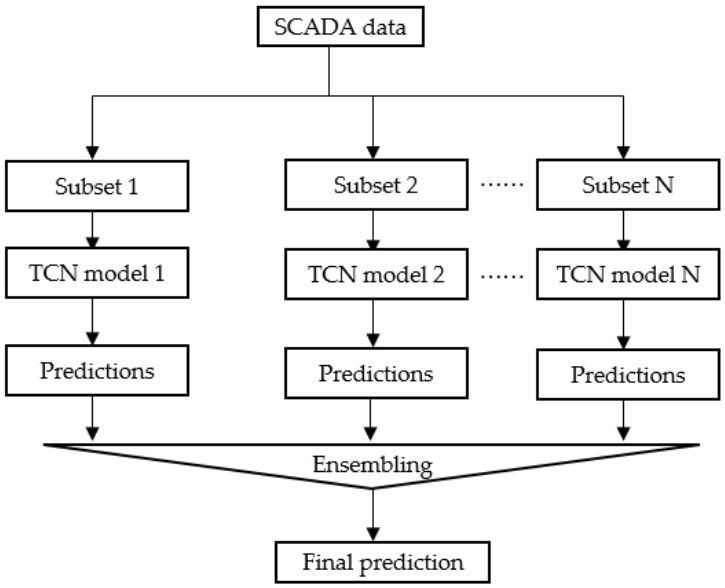

Figure 5. Blade icing method based on under-sampling and ensemble learning (adapted from [45]).

In summary, to solve the problem of data imbalance, data-based and algorithm-based methods are proposed, as presented in Table 2. The flow of the wind turbine blade icing classification model is roughly shown in Figure 6. First, the historical data recorded by the SCADA system was selected features and processed for data imbalance. Then, the model was modelled by using artificial intelligence algorithms. Finally, the model was evaluated and optimized by using the test set.

Table 2. Categorization of data imbalance solutions by methodology.

Methods	References	Characteristics
Data-based	Ge et al. [13] Zhang et al. [24] Ye et al. [27] Chen et al. [32] Tang et al. [37] Xu et al. [39] Peng et al. [40] Jiang et al. [41]	This method eliminates the effects of data imbalance, but incorrect down-sampling can lose helpful information, and incorrect oversampling can lead to overfitting problems.
Algorithm-based	Chen et al. [42] Peng et al. [43] Tong et al. [44] Ding et al. [45]	This method eliminates the effects of data imbalance but may suffer from hyperparameters or thresholds that are difficult to determine.

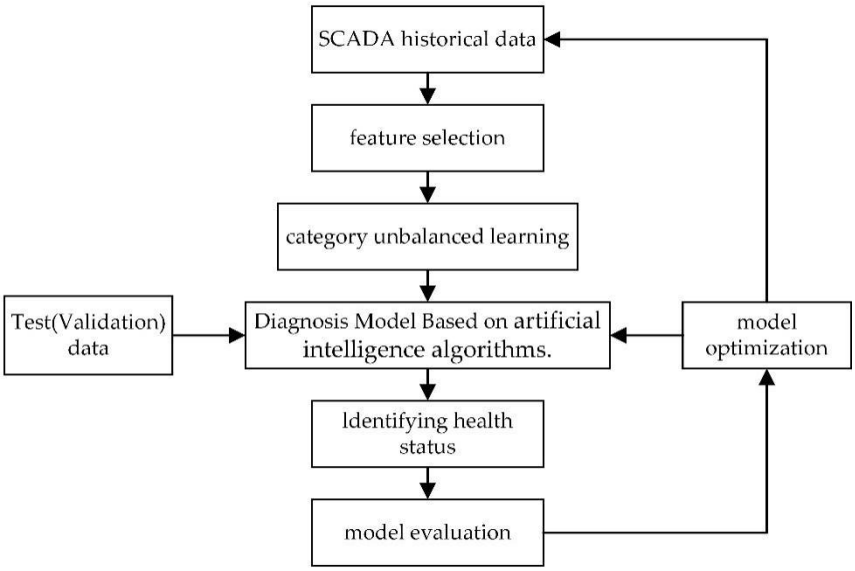

Figure 6. Flow chart of wind turbine blade icing classification model (adapted from [27]).

3. Research on the Time Dimension of SCADA Data

Wind turbine blade icing is considered a gradual process, and there is a correlation between before and after in the time dimension. Hence, the SCADA data are essentially time series data, which is mostly ignored by the previously-mentioned works. In addition, to solve the data imbalance problem, the SMOTE algorithm has been also used to generate icing samples that will destroy the characteristics of the SCADA data in the time dimension. Wang et al. [46] constructed time variation features based on the original data to solve this problem, while the DBN was used to construct the blade icing detection model. By considering the time dimensional information of the SCADA data, Li et al. [47] used a sliding window to intercept time series segments from the original data as the training samples. In this way, each training sample represents the state of the wind turbines over a period of time rather than a moment in time, making it thus easier for the algorithm to learn the hidden knowledge of the SCADA data. To better learn the information between the time sequence fragments, Yuan et al. [48] proposed a wavelet-based fully convolutional neural network (FCNN), namely Wavelet FCNN. The model framework is illustrated in Figure 7. The time-frequency domain information of SCADA data was learned by enforcing an orthogonal discrete wavelet transform. Based on this approach, the detection model was constructed by using the ensemble learning and CNN, and the global pooling layer was used instead of the fully connected layer to achieve the deep mining of SCADA time series samples. To further improve the robustness of this model, an anomaly monitoring

algorithm combining the sliding windows and majority voting was designed during the simulation deployment stage. After the model predicted the real-time data, the multi-voting mechanism was used to determine whether the current data were abnormal by considering the wind turbine's state over a while.

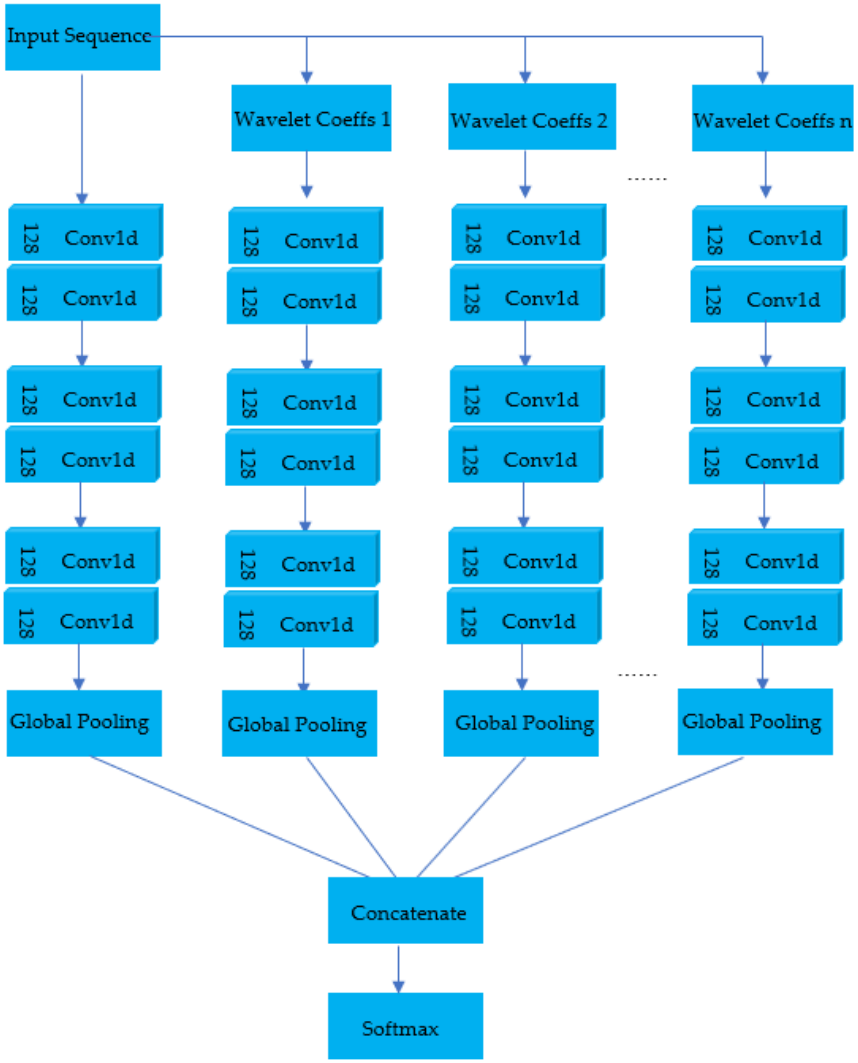

Figure 7. The architecture of Wavelet FCNN (adapted from [48]).

Compared with the traditional FCNN model, the accuracy, precision, and F1 score were remarkably improved after the wavelet transform method, and the recall was slightly decreased. However, there was only a 28.4% (precision) probability that the predicted iced samples were accurate, which was not enough for applying to the wind turbine blade icing detection. According to the literature [49,50], the FCNN architecture does not adequately learn the time dependency of the SCADA data. Therefore, Wang et al. [49] proposed a wavelet-based multi-scale long LSTM network, namely Wavelet LSTM, for learning the global and local features of multivariate SCADA data. The LSTM network is a kind of neural network that is used to learn sequential data in deep learning. It can detect the long-term dependence of SCADA data and avoid the gradient disappearance of the RNN. After using the LSTM structure and the wavelet transform, the accuracy and recall rate of the model reached more than 90%, and the accuracy reached 37.6%, which was nearly 32% higher than that of Wavelet FCNN. Tian et al. [51] also proposed a parallel network architecture composed of LSTM and CNN to improve the model's ability to extract SCADA data in both temporal and spatial dimensions, as can be observed from Figure 8.

At the same time, by considering that the time series samples still have the problem of sample imbalance, a data enhancement algorithm was proposed. During the sampling process, through the non-overlapping segmentation of the normal data, the overlapping segmentation of the icing data increased the number of icing samples to balance the data. The model scored more balanced precision and recall on the test set for individual wind turbines than Wavelet FCNN, reaching 63% and 61%, respectively.

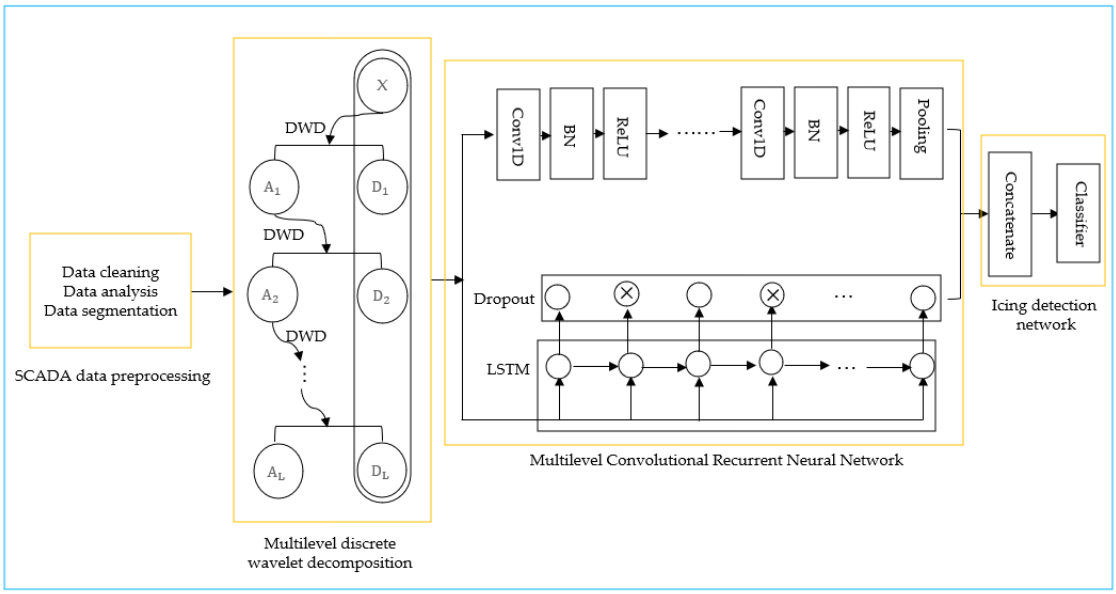

Figure 8. The parallel structure of LSTM and CNN (adapted from [50]).

By considering the efficiency and complexity of the model operation, Li et al. [51] introduced the gated recurrent unit (GRU) to construct the icing detection model. GRU is a simplified version of LSTM [52]. Compared with the LSTM network, this structure is simpler and more time-efficient. Moreover, the authors used BiGRU, where the model introduced information about future moments in the learning process to better capture the temporal dependencies and improve sensitivity to time.

The wind turbine blade icing is a gradual process that is correlated backward and forward in the time dimension, while time-series models may be more suitable for the problem. Hence, the application of 1D-CNN, LSTM, and GRU to wind turbine icing detection was examined. Wavelet FCNN performs a wavelet transform on Scada data and then uses the detail coefficients at different scales as the input features, which may not be learned by CNNs. CNNs are concerned with the connections between different sensors, whereas LSTMs are more of a metric for time steps. Compared to LSTM, GRU is somewhat simpler in structure, and BIGRU can introduce information from future moments for the model to learn, by capturing the temporal sensitivity better.

4. Semi-Supervised Learning Method in Blade Icing Detection

Most of the aforementioned data-driven icing detection approaches are supervised learning and have the disadvantage of being overly dependent on labels. However, the accurate labeling of large amounts of data can be very challenging. Under this direction, the semi-supervised learning approach is a learning framework that addresses this problem. It allows predictive models to be built by using a small amount of labeled data and a large amount of unlabeled data [53,54]. In this regard, Yi et al. [55] proposed a graph-based semi-supervised learning system. It first used mutual information to extract features from data to eliminate redundant variables. Then, the graph data are constructed from processed data and a small amount of labeled data is used to represent the relationship between the samples. Finally, through the iterative operation of the label propagation algorithm of the

graph, the initial label information is propagated to the data for classification. Autoencoder (AE) is also an effective unsupervised learning method to learn the potential representations of the input data [56]. Autoencoder have been also introduced for icing detection in wind turbines. Yuan et al. [57] proposed a wavelet enhanced autoencoder model (Wavelet AE), whose specific architecture is depicted in Figure 9. The input SCADA data were first decomposed into multi-level wavelet detail coefficients. Then, at each scale layer, the SCADA data or wavelet detail coefficients were passed through the CNN encoder and the LSTM encoder, and the hidden states of the LSTM encoder were concatenated to form a global code. In the decoding stage, the fully concatenated layer first mapped the global code to the initial hidden state at each scale. Then the LSTM decoder and CNN decoder reconstructed the SCADA data and wavelet detail coefficients as the way to learn the low-dimensional representation of SCADA data, and then the encoder part was used as a feature extractor. A SoftMax layer was added as a classifier to supervise the model's training using labels. Moreover, the method's effectiveness was verified in real datasets, and the model accuracy and precision metrics reached more than 80%. Additionally, Liu et al. [28] applied deep auto-encoders to learn multilevel fault features from SCADA data adaptively and used the idea of ensemble learning to construct an integrated icing detection model by adding classifiers after each hidden layer of the network. This method has been also tested on another wind turbine, on which the model's accuracy was as high as 96%, and the MCC coefficient was 0.783, which was better than RF, SVM, shallow neural networks, and DNN.

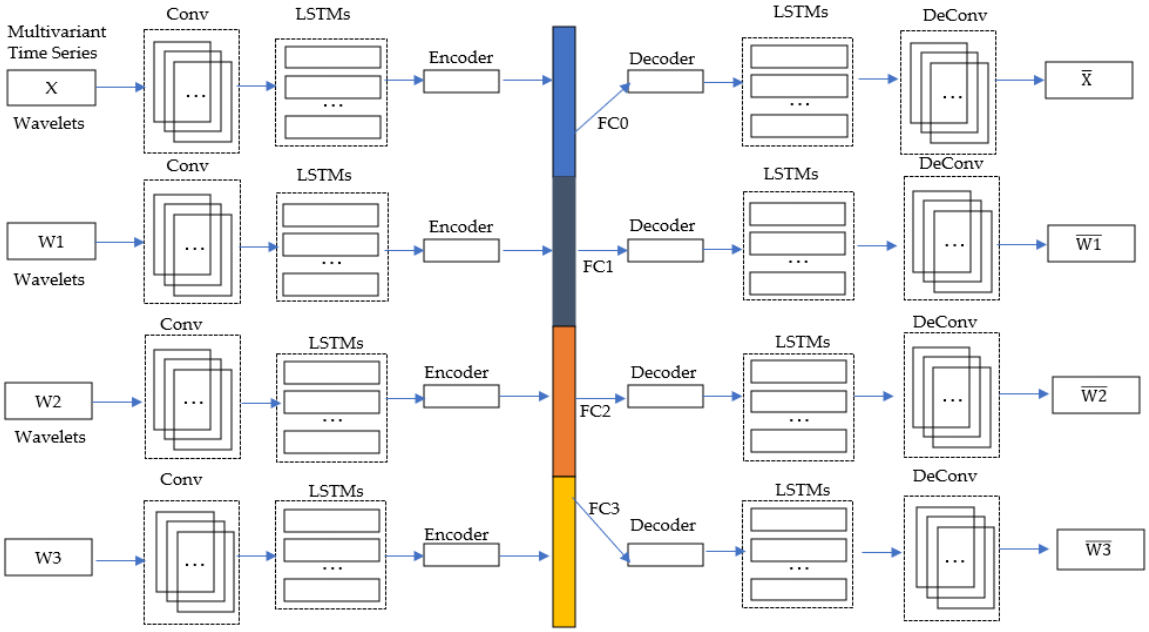

Figure 9. The architecture of Wavelet AE (adapted from [57]).

Similarly, Yi et al. [58] proposed a wind turbine icing fault detection method based on discriminative feature learning. The stacked autoencoder (SAE) model trained by the normal data was used to extract deep features of the sample (i.e., the original data, the features extracted by SAE, and the error vector), and then the extracted deep features were mapped to the low-dimensional space by sparse linear discriminant analysis for classifier training. The better classification was attained when considering the time series relationships. Cheng et al. [59] proposed a new depth class imbalance semi-supervised model for icing detection. The specific principle is depicted in Figure 10. The method first projected the imbalanced dataset with only partially labeled data into the feature space with the feature encoder and then utilized the prototype network to calculate the prototypes of each category (i.e., icing or non-icing) from labeled data. Finally, it updated the labeled data prototype by comparing the distance between the unlabeled data prototype (unknown

class) and the labeled data prototype (known class). This method used only unlabeled data for training the optimization model. In the testing phase, each test sample was mapped to the feature space by the feature encoder and then classified by the trained classifier.

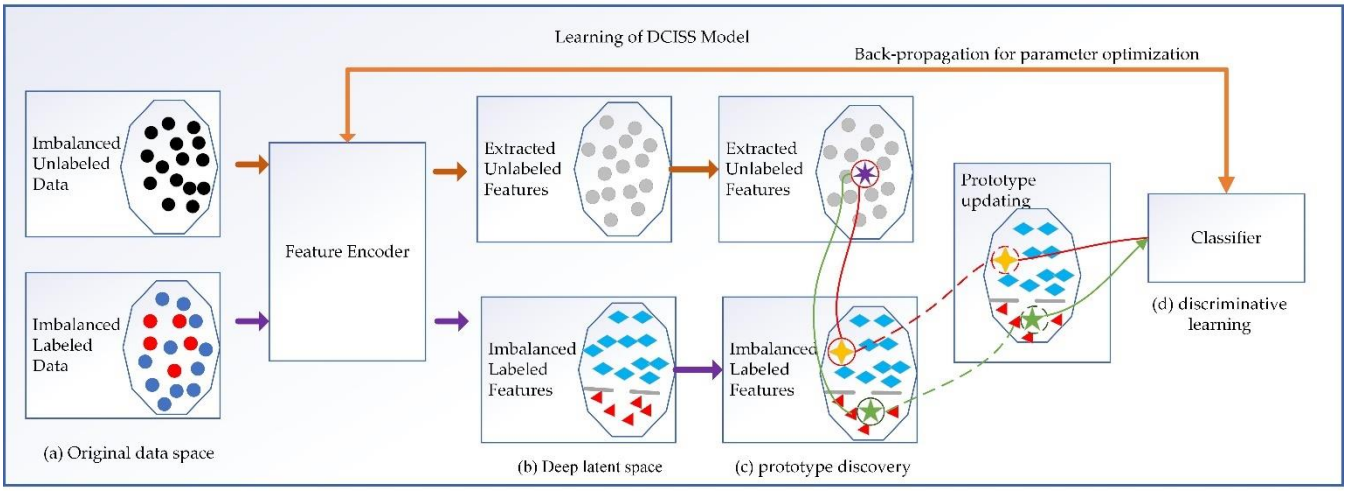

Figure 10. Imbalanced semi-supervised model (adapted from [59]).

Supervised learning requires high-quality labeled data, whereas in the industry there is a large amount of unlabeled data and a severe shortage of high-quality labeled data. As a result, in this section, the application of semi-supervision was reviewed for leaf icing detection. While for graph neural networks it is worth considering when constructing graph data, for AE the implementation of AE unsupervised learning to SCADA low-dimensional representations for distinguishing normal from icing data should be considered.

5. Transfer Learning Method in Blade Icing Detection

In some cases, obtaining a large amount of training data for the target domain is difficult and expensive. Therefore, it is of great significance to learn from the relevant source domain to transfer knowledge to the target domain and build a high-performance model for the target domain, which is the motivation for transfer learning [60]. For wind turbines, different geographical locations and environments lead to specific differences in the operation data of wind turbines. The model based on a single wind turbine has the problem of insufficient generalization ability when it is directly applied to other wind turbines. Meanwhile, the retraining model for the target wind turbine has the problem of inadequate data and label quality. Therefore, Zhang et al. [61] proposed a CNN detection framework based on the inductive transfer learning (ITL) method. As shown in Figure 11, the author used many source domain data to train FNN to construct a high-performance icing detection model for wind turbine B. Based on the icing detection model of wind turbine B, the author used a small amount of SCADA data (target domain data) of wind turbine A to quickly update the model to construct the icing detection model of wind turbine A. The authors also compared it with the direct application of FNN. Interestingly, the inductive transfer learning based FNN can detect about 79% of the ice data, which is 14% more accurate than the direct application of FNN.

Yun et al. [62] proposed an adaptive inductive transfer learning (adaptive ITL) method for detecting wind turbines. The underlying principle is displayed in Figure 12. First, the source domain data was used to train the base classifier and to predict the target domain data. Then, based on the prediction results, the predicted icing and non-icing data were used to train the precision auxiliary classifier and recall auxiliary classifier, respectively. The basic classifier first classified each test sample in the testing phase. Then, the samples were fed into the auxiliary classifier based on the prediction results, and the output of the auxiliary classifier was used as the final prediction result. The author used RF, AdaBoost,

FNN, and other classifiers, to verify that the classifier's performance was improved with adaptive ITL.

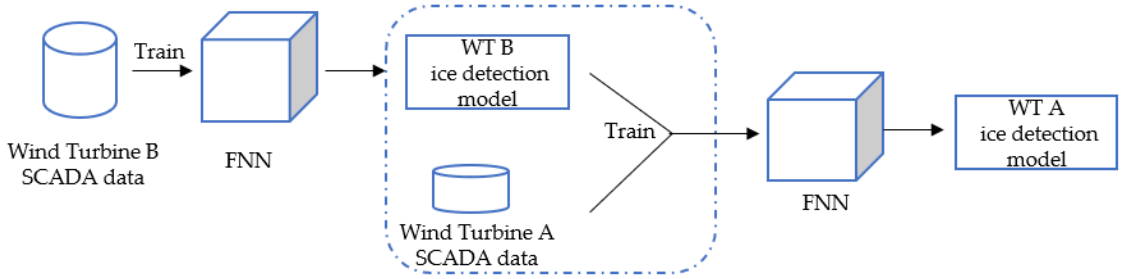

Figure 11. The framework of the inductive transfer learning based FNN (adapted from [61]).

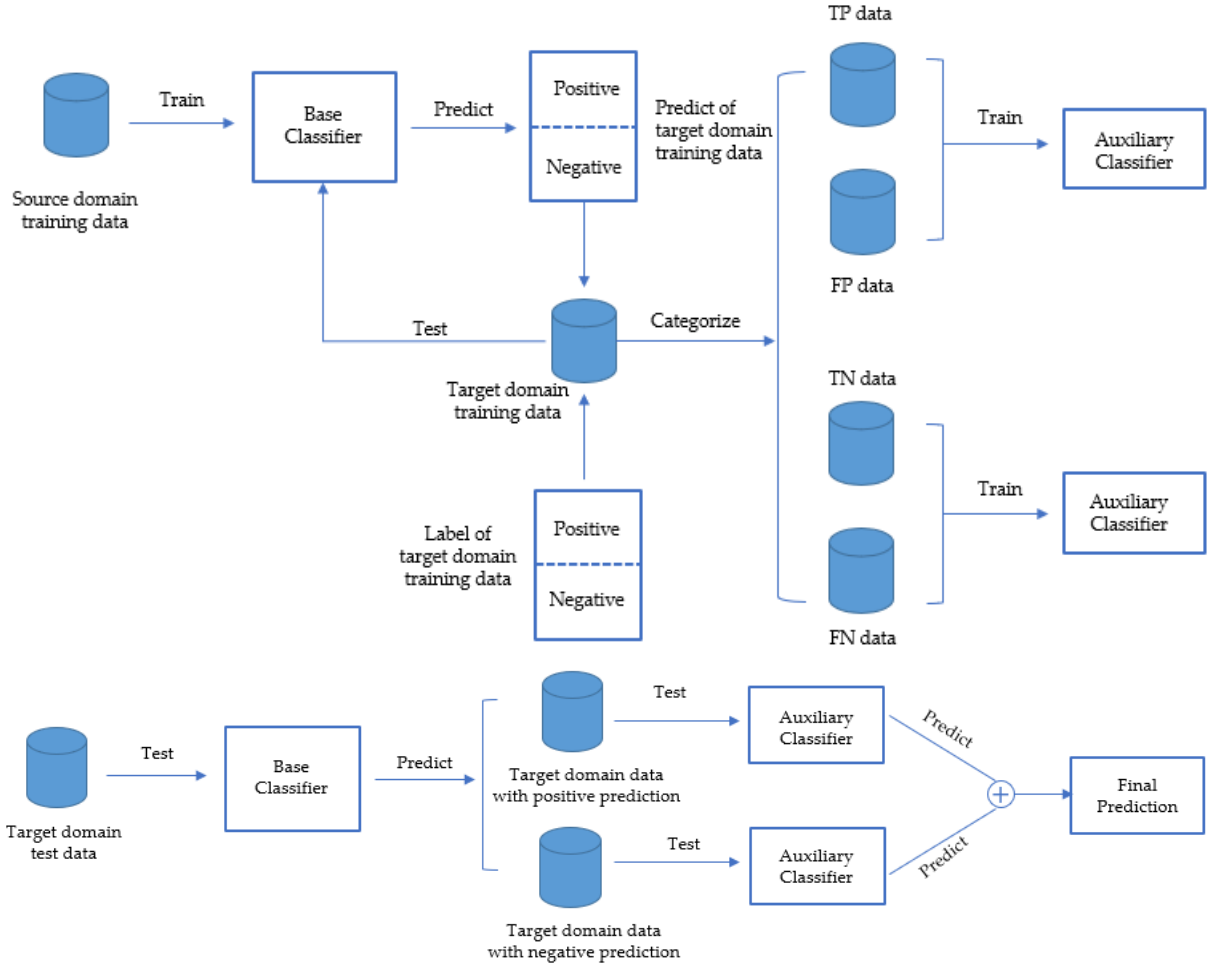

Figure 12. Adaptive inductive transfer learning (adaptive ITL) method (adapted from [62]).

Similarly, Chen et al. [63] applied the TrAdaBoost algorithm, which is another transfer learning algorithm, for constructing classifiers on a small amount of data in the target field to detect wind turbines. The authors claimed that the generalization ability of the TrAdaBoost algorithm in other wind turbines was better than that of the traditional machine learning KNN and random forest algorithms. The principle was to iteratively update their weights based on the contribution of each sample to the classification in the source and target domain dataset. The weight updating strategy for the target domain samples was the same as AdaBoost but the weights of sample models in the source domain were updated by using different techniques [64]. Liu et al. [65] applied the idea of domain

adaptation in transfer learning to the icing detection of wind turbine blades and used the temporal convolutional network to construct the model, namely domain adaptation temporal convolutional network (DATCN). Domain adaptation alternately trained feature extractors and domain discriminators to spontaneously learn a Nash equilibrium state through the idea of confrontation, completed the transfer learning from the source domain wind turbine to the target domain wind turbine, and reduced the influence of the different wind turbines due to different geographical locations, environmental factors, and operating conditions. Compared with the TCN model without migration, the DATCN model was more adaptive.

ITL and adaptive ITL are essentially a way to achieve high performance in a target task by transferring knowledge from the source task to the target task. TrAdaBoost is an Instance-based TL, where a small amount of newly labeled data combined with weighted old data are used to construct a high-quality classification model for the new data. Domain adaptation is a more tightly constrained form of TL, which aims to use the well-labeled data in the source domain to learn an accurate model that can be applied to a target domain with no or little labeling. By considering the time cost of collecting data from multiple wind turbines and the quality of the labels, it is still helpful to use transfer learning to achieve a multi-wind turbine diagnostic model.

6. Research Gaps and Challenges

There are several essential issues to be considered about the application of data-driven approaches in icing detection of wind turbines in the future:

The vast majority of the current models are offline predictions under conditions using test sets to simulate the online monitoring phase and are not applied in reality. In the future, it should be considered to deploy the model to wind turbines for online learning and to continuously learn and update based on the new data generated by wind turbine operation.

In terms of the feature redundancy and feature selection for SCADA data, although some of the features in the SCADA data have a low level of contribution to icing prediction, there is a high risk that some helpful information will be lost by artificially removing some features and that the features will not be the same depending on the filtering algorithm. Since humans do not know what the model needs, the model can be allowed to learn on its own. When building the model, the attention mechanism is embedded in the deep neural network. Through performing continuous iterative updates, the model learns to give more weight to the robust features and less weight to the weak features [66]. For example, Cheng et al. [67] proposed a temporal attention-based convolutional neural network, which introduced the temporal attention module into the convolutional neural network to learn the importance of sensors and time steps. Thus, discriminative features in the raw data can be automatically identified. Of course, the location of the attention mechanism embedding is also a question worth exploring.

The development of data-driven approaches based on a large amount of data is also required, especially to build models with sufficient generalization capabilities, implying that different wind farm data need to be accessed. The SCADA data contain vast information and also involve the issues of commercial confidentiality of wind farms, resulting in the sharing of data in the wind turbine domain not being well developed. In response, Cheng et al. [68] proposed a federation learning (FL) based blade icing detection model and also validated the model's reliability on the data from two wind farms. As a distributed machine learning paradigm, FL does not require network transmission of large amounts of data, such as centralized learning, and there is a threat of data leakage. FL enables participants to model jointly without sharing data, break data isolated islands technically, and achieve AI collaboration [69,70]. The issue of surpassing the existing bottlenecks and effectively protecting the privacy and ownership of wind turbine data is an important research direction in the future.

Artificial intelligence algorithms are generally regarded as "black boxes", which can make decisions alone. However, their poor interpretability is not well-understood. Studying the blade surface icing mechanism from the mechanism level and combining the blade icing mechanism model with the data-driven approaches to predict the manifestation of an icing event in advance, prevent icing, and even eliminate icing, has high research values.

To achieve the localization of blade icing, Hacıefendioğlu et al. [6,71] used pre-trained models of classical convolutional neural networks from deep learning for the accurate detection of blade icing by using image data and investigated the various visualization techniques for locating the location of blade icing. The study demonstrated that Gradient-weighted Class Activation Mapping for Resnet-50 outperformed the other combinations. In order to clearly show the icing boundary, the image semantic segmentation network U-NET was applied to detect wind turbine blades. SCADA data detected the status of wind turbines from a global perspective, while the image data intuitively determined whether the blades are iced from a local perspective. In the future, deep learning can be used to fuse information from multiple sources to improve the accuracy of wind turbine fault detection and to locate blade icing.

7. Conclusions

The fault detection of wind turbine blade icing is of great importance for wind farms to improve the power generation efficiency and the operational life of wind turbines. With the widespread use of wind turbine SCADA systems and the rapid development of artificial intelligence algorithms, data-driven approaches show great potential in the fault diagnosis of wind turbine blade icing.

In this work, the application of data-driven approaches to wind turbine blade icing detection was reviewed and the problem of feature selection and category imbalance in classification tasks with data-driven approaches was analyzed. In this regard, three approaches to feature selection were summarized including mechanism-based, algorithm-based, and mechanism+algorithm-based. The previously reported works in the literature on data and algorithms that are used to deal with the category imbalance problem were also reviewed. It should be also noted that SCADA data are essentially time series data, which should be considered when icing detection models are developed. The success of a wind turbine blade diagnostic model requires a relatively large number of labeled samples, which is difficult to be satisfied in real applications. As a result, to bridge this gap, semi-supervised learning has been introduced for icing fault diagnosis of wind turbines. Due to the different geographical locations and environmental factors of the different wind turbines, which lead to poor generalization of the model to other wind turbines, in response, knowledge learned from one wind turbine was transferred to another, and TL was introduced to the task of wind turbine icing fault diagnosis. Finally, the challenges of data-driven approaches based on data were discussed, and an outlook for future research was provided.

Author Contributions: Conceptualization, X.S. and Q.L.; writing of the manuscript, J.G. and X.S.; updating of the text, X.S. and C.C.; literature searches, Y.Z., Y.J. and Z.X.; critical reviewing of the manuscript, J.W., S.T. and C.C.; organization and editing of the manuscript, J.G., X.S. and Y.Z.; project administration: C.C.; supervision: Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This Research was funded by the National Natural Science Foundation of China (Grant No. 52176212) and the National Engineering Research Center for Offshore Windpower (Grant No. HSFD22001). The research was also supported by the Doctor Foundation of the Inner Mongolia University of Technology (BS2021031), We are also very grateful to the Program for Innovative Research Team in Universities of Inner Mongolia Autonomous Region (NMGIRT2213) and Inner Mongolia Science and Technology Program under Grant (2021GG0336).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

SCADA	supervisory control and data acquisition
XGBoost	extreme gradient Boosting
SPRT	sequential probability ratio test
AdaBoost	adaptive boosting
DBN	deep belief network
PCA	principal component analysis
BP	back propagation
SFA	slow feature analysis
SVM	support vector machine
MIV	mean impact value
tSNE	t-distributed stochastic neighbor embedding
RNN	Recurrent Neural Networks
CNN	convolutional neural network
RF	random forest
LSTM	long short-term memory
SMOTE	synthetic minority over-sampling technique
MBK-SMOTE	mini-batch k-means smote
SC-SMOTE	safe circle smote
KNN	k-nearest neighbor
DNN	deep neural networks
AWKELM	adaptive weighted kernel extreme learning machine
TCN	temporal convolutional network
FCNN	fully convolutional neural network
GRU	gated recurrent unit
AE	autoencoder
SAE	stacked autoencoder
ITL	inductive transfer learning
DATCN	domain adaptation temporal convolutional network
FL	federation learning

References

1. USEIA. Available online: https://www.eia.gov/outlooks/ieo (accessed on 3 March 2022).
2. Council, G.W.E. Global Wind Report 2022. Available online: https://gwec.net/global-wind-report-2022/ (accessed on 4 April 2022).
3. Parent, O.; Ilinca, A. Anti-icing and de-icing techniques for wind turbines: Critical review. *Cold Reg. Sci. Technol.* **2011**, *65*, 88–96. [CrossRef]
4. Fortin, G.; Perron, J.; Ilinca, A. Behaviour and Modeling of Cup Anemometers under Icing Conditions. In *IWAIS XI*; Montreal, QC, Canada, June 2005. Available online: https://www.researchgate.net/publication/259758149_Behaviour_and_Modeling_of_Cup_Anemometers_under_Icing_Conditions (accessed on 5 January 2023).
5. Laakso, T.; Holttinen, H.; Ronsten, G.; Tallhaug, L.; Horbaty, R.; Baring-Gould, I.; Lacroix, A.; Peltola, E.; Tammelin, B. State-of-the-art of wind energy in cold climates. *IEA Annex XIX* **2003**, *24*, 53.
6. Hacıefendioğlu, K.; Başağa, H.B.; Yavuz, Z.; Karimi, M.T. Intelligent ice detection on wind turbine blades using semantic segmentation and class activation map approaches based on deep learning method. *Renew. Energy* **2022**, *182*, 1–16. [CrossRef]
7. Ronsten, G.; Wallenius, T.; Hulkkonen, M.; Baring-Gould, I.; Cattin, R.; Durstewitz, M.; Krenn, A.; Laakso, T.; Lacroix, A.; Tallhaug, L. State-of-the-art of Wind Energy in Cold Climates. *IEA Wind Task*. 2012. Available online: https://www.tuulivoimayhdistys.fi/media/task19_sota_weincc_2012_approved.pdf (accessed on 5 January 2023).
8. Pérez, J.M.P.; Márquez, F.P.G.; Hernández, D.R. Economic viability analysis for icing blades detection in wind turbines. *J. Clean. Prod.* **2016**, *135*, 1150–1160. [CrossRef]
9. Lehtomäki, V.; Rissanen, S.; Wadham-Gagnon, M.; Sandel, K.; Moser, W.; Jacob, D. Fatigue loads of iced turbines: Two case studies. *J. Wind. Eng. Ind. Aerodyn.* **2016**, *158*, 37–50. [CrossRef]

10. Shu, L.; Qiu, G.; Hu, Q.; Jiang, X.; Wang, P.; Xia, H. Numerical Model and Field Experimental Investigation of Threshold Heat Flux of Wind Turbine De-icing. *Proc. CSEE* **2018**, *38*, 3997–4003. [CrossRef]
11. Fikke, S.M.; Ronsten, G.; Heimo, A.; Kunz, S.; Ostrozlik, M.; Persson, P.; Sabata, J.; Wareing, B.; Wichura, B.; Chum, J. *COST 727: Atmospheric Icing on Structures: Measurements and Data Collection on Icing: State of the Art*; Meteo Schweiz: Geneva, Switzerland, 2006.
12. Laakso, T.I.; Peltola, E.; Tammelin, B. *Wind Turbines in Icing Environment: Improvement of Tools for Siting, Certification and Operation: New Icetools*; Ilmatieteen laitos: Helsinki, Finland, 2005; p. 127, ISBN 951-697-618-2.
13. Ge, Y.; Yue, D.; Chen, L. Prediction of wind turbine blades icing based on MBK-SMOTE and random forest in imbalanced data set. In Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 26–28 November 2017; pp. 1–6.
14. Tao, T.; Liu, Y.; Qiao, Y.; Gao, L.; Lu, J.; Zhang, C.; Wang, Y. Wind turbine blade icing diagnosis using hybrid features and Stacked-XGBoost algorithm. *Renew. Energy* **2021**, *180*, 1004–1013. [CrossRef]
15. Berbyuk, V.; Peterson, B.; Möller, J. Towards early ice detection on wind turbine blades using acoustic waves. In Proceedings of the Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security 2014, San Diego, CA, USA, 10 March 2014; pp. 96–106.
16. Muñoz, C.Q.G.; Márquez, F.P.G.; Tomás, J.M.S. Ice detection using thermal infrared radiometry on wind turbine blades. *Measurement* **2016**, *93*, 157–163. [CrossRef]
17. Kim, D.-g.; Sampath, U.; Kim, H.; Song, M. A fiber-optic ice detection system for large-scale wind turbine blades. In Proceedings of the Optical Modeling and Performance Predictions IX, San Diego, CA, USA, 7–8 August 2017; pp. 52–57.
18. Maldonado-Correa, J.; Martín-Martínez, S.; Artigao, E.; Gómez-Lázaro, E. Using SCADA data for wind turbine condition monitoring: A systematic literature review. *Energies* **2020**, *13*, 3132. [CrossRef]
19. Zhou, D.; Hu, Y. Fault Diagnosis Techniques for Dynamic Systems. *Acta Autom. Sin.* **2009**, *35*, 748–758. [CrossRef]
20. Guo, P.; Infield, D. Wind turbine blade icing detection with multi-model collaborative monitoring method. *Renew. Energy* **2021**, *179*, 1098–1105. [CrossRef]
21. Fan, D.; Liu, B.; Guo, P. Wind turbine blades icing detection with multi-parameter models based on AdaBoost algorithm. *Huadian Technol.* **2021**, *43*, 20–26. [CrossRef]
22. Li, N.; Yan, T.; Li, N.; Kong, D.; Liu, Q.; Lei, Y. Ice Detection Method by Using SCADA Data on Wind Turbine Blades. *Power Gener. Technol.* **2018**, *39*, 58–62. [CrossRef]
23. Zhou, G.; Tan, W.; Zhang, D. Ice detection for wind turbine blades based on PSO-SVM method. *J. Phys. Conf. Ser.* **2018**, *1087*, 022036.
24. Zhang, L.; Liu, K.; Wang, Y.; Omariba, Z.B. Ice detection model of wind turbine blades based on random forest classifier. *Energies* **2018**, *11*, 2548. [CrossRef]
25. Ma, J.; Ma, L.; Tian, X. Wind Turbine Blade Icing Prediction Based on Deep Belief Network. In Proceedings of the 2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Hohhot, China, 24–26 October 2019; pp. 26–263.
26. Liu, Y.; Cheng, H.; Kong, X.; Wang, Q.; Cui, H. Intelligent wind turbine blade icing detection using supervisory control and data acquisition data and ensemble deep learning. *Energy Sci. Eng.* **2019**, *7*, 2633–2645. [CrossRef]
27. Ye, C.; Qiu, Y.; Feng, Y. Faults Diagnosis of Wind Turbine Blade Icing based on Data Mining. *Noise Vib. Control.* **2018**, *28*, 643–647. [CrossRef]
28. Peng, C.; Tang, Z.; Chen, Q.; Wang, S.; Zhou, X.; Chi, H. Short-Term Prediction of Generator Blade Ice Fault Based on Multi-AN. In Proceedings of the 2019 Seventh International Conference on Advanced Cloud and Big Data (CBD), Suzhou, China, 21–22 September 2019; pp. 291–296.
29. Li, T.; Tan, W.; Liu, Z. A Hybrid Model Based on Logistic Regression Algorithm and Extraction Algorithm Using Reward Extremum to Real-Time Detect Blade Icing Alarm. *Int. J. Pattern Recognit. Artif. Intell.* **2019**, *33*, 1955016. [CrossRef]
30. Jia, P.; Chen, G. Wind Power Icing Fault Diagnosis Based on Slow Feature Analysis and Support Vector Machines. In Proceedings of the 2020 10th International Conference on Power and Energy Systems (ICPES), Chengdu, China, 25–27 December 2020; pp. 398–403.
31. Zhang, Z.; Fan, B.; Liu, Y.; Zhang, P.; Wang, J.; Du, W. Rapid warning of wind turbine blade icing based on MIV-tSNE-RNN. *J. Mech. Sci. Technol.* **2021**, *35*, 5453–5459. [CrossRef]
32. Chen, X.; Lei, D.; Xu, G. Prediction of icing fault of wind turbine blades based on deep learning. In Proceedings of the 2019 IEEE 2nd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 22–24 November 2019; pp. 295–299.
33. Di, R.; Chen, Z.; Lv, J. Application of XGBoost based on SMOTE in prediction of wind turbine blades icing. *Inf. Technol.* **2019**, *43*, 81–85. [CrossRef]
34. Meng, H.; Huang, X.; Liu, J.; Han, Z. Forecast of fan blade icing combing with random forest and SVM. *Electr. Meas. Instrum.* **2020**, *57*, 66–71. [CrossRef]
35. Hai, T.; Fan, H.; Wang, K.; Liu, Z.; Chen, Y. Icing Fault Diagnosis of Wind Turbines Based on PSO-SVM Algorithm. *Smart Power* **2021**, *49*, 1–6+74.

36. Nie, F.; Li, Q.; Huang, Q.; Huang, L. Wing icing detection based on dimension fusion and SSA-LSTM. *Transducer Microsyst. Technol.* **2022**, *41*, 118–121. [CrossRef]
37. Tang, Z.; Xiao, Y.; Li, Y. Wind turbine blade icing diagnosis method based on deep learning. *Mod. Electron. Tech.* **2022**, *45*, 77–82. [CrossRef]
38. Madasamy, K.; Ramaswami, M. Data imbalance and classifiers: Impact and solutions from a big data perspective. *International J. Comput. Intell. Res.* **2017**, *13*, 2267–2281.
39. Xu, J.; Tan, W.; Li, T. Predicting fan blade icing by using particle swarm optimization and support vector machine algorithm. *Comput. Electr. Eng.* **2020**, *87*, 106751. [CrossRef]
40. Peng, C.; Chen, Q.; Zhang, L.; Wan, L.; Yuan, X. Research on fault diagnosis of wind power generator blade based on SC-SMOTE and kNN. *J. Inf. Process. Syst.* **2020**, *16*, 870–881.
41. Jiang, N.; Yan, M.; Li, N. Icing detection of wind turbine blade based on the time-dimensional upsampling convolutional neural network. *Control. Decis.* **2022**, *37*, 2017–2025. [CrossRef]
42. Chen, L.; Xu, G.; Zhang, Q.; Zhang, X. Learning deep representation of imbalanced SCADA data for fault detection of wind turbines. *Measurement* **2019**, *139*, 370–379. [CrossRef]
43. Peng, D.; Liu, C.; Desmet, W.; Gryllias, K. An improved 2DCNN with focal loss function for blade icing detection of wind turbines under imbalanced SCADA data. In Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering, Virtual, Online, 16–17 February 2021; American Society of Mechanical Engineers: Cancún, Mexico, 2021; p. V001T01A018.
44. Tong, R.; Li, P.; Lang, X.; Liang, J.; Cao, M. A novel adaptive weighted kernel extreme learning machine algorithm and its application in wind turbine blade icing fault detection. *Measurement* **2021**, *185*, 110009. [CrossRef]
45. Ding, S.; Wang, Z.; Zhang, J.; Han, F.; Gu, X.; Song, G. A PCC-Ensemble-TCN model for wind turbine icing detection using class-imbalanced and label-missing SCADA data. *Int. J. Distrib. Sens. Netw.* **2021**, *17*, 15501477211057737. [CrossRef]
46. Wang, J.; Tang, Z.; Lu, P. Ice detection and voice alarm of wind turbine blades based on belief network. *Int. J. Speech Technol.* **2021**, 1–10. [CrossRef]
47. Li, M.; He, X. An accurate detection method for turbine icing issues using LSTM network. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Moscow, Russia, 27 May–6 June 2019; p. 032109.
48. Yuan, B.; Wang, C.; Jiang, F.; Long, M.; Yu, P.S.; Liu, Y. WaveletFCNN: A Deep Time Series Classification Model for Wind Turbine Blade Icing Detection. *arXiv* **2019**. [CrossRef]
49. Wang, X.; Zheng, Z.; Jiang, G.; He, Q.; Xie, P. Detecting Wind Turbine Blade Icing with a Multiscale Long Short-Term Memory Network. *Energies* **2022**, *15*, 2864. [CrossRef]
50. Tian, W.; Cheng, X.; Li, G.; Shi, F.; Chen, S.; Zhang, H. A multilevel convolutional recurrent neural network for blade icing detection of wind turbine. *IEEE Sens. J.* **2021**, *21*, 20311–20323. [CrossRef]
51. Li, Y.; Hou, L.; Tang, M.; Sun, Q.; Chen, J.; Song, W.; Yao, W.; Cao, L. Prediction of wind turbine blades icing based on feature Selection and 1D-CNN-SBiGRU. *Multimed. Tools Appl.* **2022**, *81*, 4365–4385. [CrossRef]
52. Greff, K.; Srivastava, R.K.; Koutník, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 2222–2232. [CrossRef]
53. Blum, A.; Mitchell, T. Combining labeled and unlabeled data with co-training. In Proceedings of the Eleventh Annual Conference on Computational Learning Theory, Madison, WI, USA, 24–26 July 1998; pp. 92–100.
54. Zhu, X.J. Semi-Supervised Learning Literature Survey. 2005. Available online: http://digital.library.wisc.edu/1793/60444 (accessed on 3 June 2022).
55. Yi, H.; Jiang, Q. Graph-based semisupervised learning for icing fault detection of wind turbine blade. *Meas. Sci. Technol.* **2020**, *32*, 035117. [CrossRef]
56. Ng, A. Sparse autoencoder. *CS294A Lect. Notes* **2011**, *72*, 1–19.
57. Yuan, B.; Wang, C.; Luo, C.; Jiang, F.; Long, M.; Yu, P.S.; Liu, Y. WaveletAE: A wavelet-enhanced autoencoder for wind turbine blade icing detection. *arXiv* **2019**, *preprint*. arXiv:1902.05625.
58. Yi, H.; Jiang, Q. Discriminative feature learning for blade icing fault detection of wind turbine. *Meas. Sci. Technol.* **2020**, *31*, 115102. [CrossRef]
59. Cheng, X.; Shi, F.; Liu, X.; Zhao, M.; Chen, S. A novel deep class-imbalanced semisupervised model for wind turbine blade icing detection. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 2558–2570. [CrossRef] [PubMed]
60. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
61. Zhang, C.; Bin, J.; Liu, Z. Wind turbine ice assessment through inductive transfer learning. In Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, I2MTC 2018, Houston, TX, USA, 14–17 May 2018; pp. 1–6.
62. Yun, H.; Zhang, C.; Hou, C.; Liu, Z. An adaptive approach for ice detection in wind turbine with inductive transfer learning. *IEEE Access* **2019**, *7*, 122205–122213. [CrossRef]
63. Chen, W.; Qiu, Y.; Feng, Y.; Li, Y.; Kusiak, A. Diagnosis of wind turbine faults with transfer learning algorithms. *Renew. Energy* **2021**, *163*, 2053–2067. [CrossRef]
64. Dai, W.Y.; Yang, Q.X. Boosting for transfer learning. In Proceedings of the Twenty-Fourth International Conference (ICML 2007), Corvallis, OR, USA, 20–24 June 2007; Volume 193, p. 200.
65. Liu, X. Diagnosis of Wind Turbine Blade Icing Based on SCADA Data. Master's Thesis, Kunming University of Science and Technology, Kunming, China, September 2021.

66. Guo, M.-H.; Xu, T.-X.; Liu, J.-J.; Liu, Z.-N.; Jiang, P.-T.; Mu, T.-J.; Zhang, S.-H.; Martin, R.R.; Cheng, M.-M.; Hu, S.-M. Attention mechanisms in computer vision: A survey. *Comput. Vis. Media* **2022**, *14*, 1–38. [CrossRef]
67. Cheng, X.; Shi, F.; Zhao, M.; Li, G.; Zhang, H.; Chen, S. Temporal attention convolutional neural network for estimation of icing probability on wind turbine blades. *IEEE Trans. Ind. Electron.* **2021**, *69*, 6371–6380. [CrossRef]
68. Cheng, X.; Shi, F.; Liu, Y.; Liu, X.; Huang, L. Wind turbine blade icing detection: A federated learning approach. *Energy* **2022**, *254*, 124441. [CrossRef]
69. Zhang, C.; Xie, Y.; Bai, H.; Yu, B.; Li, W.; Gao, Y. A survey on federated learning. *Knowl.-Based Syst.* **2021**, *216*, 106775. [CrossRef]
70. AbdulRahman, S.; Tout, H.; Ould-Slimane, H.; Mourad, A.; Talhi, C.; Guizani, M. A survey on federated learning: The journey from centralized to distributed on-site learning and beyond. *IEEE Internet Things J.* **2020**, *8*, 5476–5497. [CrossRef]
71. Hacıefendioğlu, K.; Başağa, H.B.; Ayas, S.; Karimi, M. Automatic detection of icing wind turbine using deep learning method. *Wind. Struct.* **2022**, *34*, 511–523. [CrossRef]

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