

Error Estimation of Electric Energy Measurement Based on Wavelet Transform and Migration Neural Network

Meng-Shuang Liu*, Xu-Dong Wang

Electric Power Research Institute
State Grid Xinjiang Electric Power Co., Ltd., Urumqi 830000, P. R. China
xurongpeng_111@163.com, zhangtao_tao1@163.com

Chen Yang

Electric Power Research Institute
State Grid Xinjiang Electric Power Co., Ltd., Urumqi 830000, P. R. China

Chen Yang

College of Sciences
Adamson University, Manila 1008, Philippines
liwenyan_111@163.com

*Corresponding author: Meng-Shuang Liu

Received December 5, 2023, revised March 17, 2024, accepted May 29, 2024.

ABSTRACT. *Measuring device as the terminal equipment of the power grid is directly related to the fairness of the power transaction, so it is particularly important to detect the operation error of the measuring device. The current metering device inspection methods in electric power companies have the problems of high work intensity and serious manpower waste. Therefore, in order to improve the operation and maintenance level of the metering device, this work proposes a remote diagnosis method based on wavelet transform and migration neural network to complete the operation error estimation online. First, the metering error of the meter is characterised by forming an estimation population and establishing feature parameters, and then the remote estimation of the metering error is achieved by analysing the changes of the feature parameters. Then, based on the topological relationship, a remote estimation model of metering device operation error is constructed by using the massive power data from the collection system to locate the metering device whose operation error exceeds the limit. Finally, the model WT-Ada-LSTM, which combines deep learning, wavelet analysis and migration learning, is proposed. The WT-Ada-LSTM model is employed to monitor and analyse the feature covariates in real time to achieve online quantitative estimation of metering errors. The results show that the deviations between the results obtained by the proposed remote quantitative estimation method and the true measurement errors approximately satisfy the normal distribution, and the maximum values of the means are 0.0032% and 0.2871%, respectively.*

Keywords: Smart grid; Energy metering; Transfer learning; LSTM; Wavelet transform

1. **Introduction.** At present, the utilisation rate of power data in the smart grid is very low, and the metering data should not become dead data stored in the database [1, 2, 3], and it is necessary to process and mine the massive power data in depth with the help of rich big data processing methods to realise its value.

Meter metering device is the fundamental source of voltage data in the power system, and real-time mastery of its metering error status will help to maintain the fairness of cost settlement in the system [4, 5]. The widely used method of offline verification of metering error requires periodic checks to estimate the metering error by comparing it with the output of the standard. However, this method not only requires the equipment to be estimated to be disconnected from the power supply, but also requires bulky auxiliary equipment [6, 7], resulting in a large amount of workload for offline calibration, and at the same time, it is not possible to grasp the metering error status of the meter in real time.

Running error is an important parameter of meter operation, which is related to the accuracy of meter measurement and has important guiding significance for meter replacement and correction. At present, electric power companies mainly test the meter operation error by two ways: on-site testing [8] and laboratory testing [9]. If the error detection of the meter is needed, the power maintenance personnel should remove the meter from the site and send the meter to a professional organisation to test its metrological characteristics [10]. However, testing millions of smart meters is a time-consuming and expensive task. On the other hand, it is that the current testing methods cannot meet the requirement of fast and accurate replacement of defective and abnormal meters with a large lag. In order to realise the change of smart meter assessment methods from regular dismantling to real-time inspection, as well as to ensure that the metering devices are working in a good condition, it is a difficult task to improve the ability of smart meter abnormality detection.

Currently, online estimation methods of telegraph metering errors based on information-physics fusion [11] have gained widespread attention because they do not require the addition of extra sensors and can diagnose both long-term gradual and short-term mutation errors. However, due to the small number of abnormal meters in the substation, this type of method can only qualitatively determine whether the metering error state has changed according to whether the outliers appear or not, and cannot quantitatively analyse the specific amount of change in the metering error, so it is still difficult to accurately grasp the metering error of the meter. Aiming at the shortcomings of the existing online estimation, this work investigates the online quantitative estimation of metrological error methods.

1.1. Related Work. The current status of smart meter operation error research is mainly to explore the causes and influencing factors of the error in a deeper and deeper way, and to propose more accurate and reliable correction methods to ensure the accuracy and stability of the smart meter. With the development of technology and in-depth research, it is believed that the study of smart meter error will make more breakthrough progress.

Researchers have conducted an in-depth study on the influencing factors of smart meter errors, including the influence of environmental factors, voltage fluctuations, load changes and other factors on smart meter errors, in order to better understand the causes of errors and develop corresponding correction measures [12]. Researchers have explored the method of establishing an error monitoring system to monitor the errors of the smart meter in real time while studying the errors of smart meters and make corrections in time to ensure the accuracy of the meter [13]. Singh et al. [14] analysed the complexity of the causes of meter metering errors, and simulated by Monte Carlo method to obtain the effects of individual factors on metering errors such as signal frequency deviation, low current, high current, etc., and to obtain the combination error of each factor and the contribution of each factor to the error. Guo and Jin [15] Proposed a remote diagnosis

model of metering point anomaly based on tree topology, in-depth analysis of the physical significance of the meter operation error, and further introduced the calculation of weighted average, which improves the stability of the model for the error estimation, and the usability and reliability of the model are jointly demonstrated by laboratory and field validation. However, the method requires a large amount of historical data and is less accurate due to the non-negligible estimation error generated by the model. Considering that the data of the meter working under light load condition will affect the result of error estimation, McLoughlin et al. [16] proposed to filter the light load data by clustering method, and achieved online error estimation of large-scale smart meters by restricted recursive least squares method. The use of increasing the weight of recent electricity data to avoid the data saturation phenomenon helps to detect the abnormal operation status of the meter in time. Shi et al. [17] proposed to filter the abnormal data by decision tree and classify the data according to the estimated line loss rate to obtain the data sets with different energy consumption characteristics, and then construct the meter operation error analysis matrix to remotely estimate the operation error of the smart meter, and if the error exceeds the regulation will be classified as faulty meter.

Deep learning has a unique advantage in data mining, using deep learning to mine the temporal features of normal time series is effective whereas abnormal data is difficult to fit and predict, deep learning is an alternative approach to detecting the operational status of smart meters. The power information collected by the meter is a time series, and the time series is composed of data points according to a uniform time interval, and it contains the characteristics of the power user's power consumption behaviour. In recent years, methods based on deep learning have been widely used in time series analysis, and classification and prediction are the two outputs of deep learning networks, and their main idea is to extract the important features of the sequence and make predictions. Liu et al. [18] used a Long-Short Time Memory (LSTM) network to achieve smart meter operation error Abnormality detection and identification. Through the training of a large amount of smart meter data, a deep learning model was established to capture the anomalies in the operation of the meter, and the automatic identification and processing of the anomalies was realised. The method effectively improves the analysis accuracy and real-time performance of meter operation errors. Fekri et al. [19] used deep learning models such as Recurrent Neural Networks (RNN) for time series prediction of smart meter readings. By training and optimising the deep learning model with large-scale data, the accurate analysis of smart meter data was achieved, which improved the detection capability and accuracy of abnormal energy consumption and provided effective support for the operation and management of smart energy metering systems.

1.2. Motivation and contribution. Although deep neural networks have achieved a certain degree of improvement in prediction performance compared with traditional time series prediction models, there is still room for improvement in prediction performance. Through migration learning, the knowledge and features learnt from the source domain data can be used to help the target domain model to better generalise to new datasets, which improves the generalisation ability of the model. In addition, the time series of electricity data collected by smart meters is a non-smooth, non-linear series containing some noise. Therefore, in order to achieve the goal of accurate prediction, the problem of non-smoothness and non-linearity needs to be solved first.

The main innovations and contributions of this work include:

(1) The principle of remote estimation of meter measurement error is given. The metering error of the meter is characterised by forming an estimation population and establishing characteristic parameters, and then the remote estimation of the metering error is achieved by analysing the changes of the characteristic parameters.

(2) Based on the topological relationship, the massive power data from the collection system is used to construct a remote estimation model of the operation error of the metering device, and locate the metering device whose operation error exceeds the limit.

(3) In order to achieve real-time remote monitoring of the metering error of the meter, a model WT-Ada-LSTM that combines deep learning, Wavelet Transform (WT) and migration learning is proposed. The WT-Ada-LSTM model is used to monitor and analyse the feature covariates in real time in order to achieve online quantitative estimation of the metering error.

2. Principle of Remote Estimation of Meter Metering Error. Since the secondary output voltage signal is a randomly fluctuating signal affected by the primary side voltage, and its fluctuation range is much larger than the limit value specified by its accuracy level, it is difficult to judge whether the metering error is abnormal or not only based on the output signal of a single telegraph. For safety reasons, most of the high-voltage substations are equipped with multiple sets of meters, so it is possible to characterise the metering error of the meters by forming an estimation group and establishing characteristic parameters, and then realise remote estimation of the metering error by analysing the changes in the characteristic parameters.

In a high-voltage substation, the secondary output voltages of the meters in different groups at the same voltage level are made up of the bus voltage and the metering error. For the n th meter and the m th meter in the estimated group, the following relationship exists between the magnitude and phase of the secondary voltage and the primary voltage:

$$U_1 = \frac{k_{nr}U_n}{f_n + 1} = \frac{k_{mr}U_m}{f_m + 1} \tag{1}$$

$$\varphi_1 = \varphi_n - \delta_n = \varphi_m - \delta_m \tag{2}$$

where U_1 and φ_1 are the primary voltage amplitude and phase measured by the estimation group, U_n is the secondary output voltage amplitude of the n th meter, k_{nr} is the secondary rated ratio of the n th meter, f_n is the secondary ratio difference of the n th meter, φ_n is the secondary phase of the n th meter and δ_n is the secondary phase difference of the n th meter. Similarly, U_m is the secondary output voltage amplitude of the m th meter.

For the estimation population constructed from n_1 same-phase meters, the characteristic parameter ΔU_1 characterising the variation of the ratio difference of the estimation population and the characteristic parameter $\Delta \varphi_1$ characterising the variation of the phase difference can be established respectively.

$$\begin{aligned} \Delta U_1 &= \begin{bmatrix} \frac{U_1-U_2}{U_1+U_2} & \frac{U_1-U_3}{U_1+U_3} & \cdots & \frac{U_{(n_1-1)}-U_{n_1}}{U_{(n_1-1)}+U_{n_1}} \end{bmatrix} \\ &= \begin{bmatrix} \frac{f_1-f_2}{2+(f_1+f_2)} & \frac{f_1-f_3}{2+(f_1+f_3)} & \cdots & \frac{f_{(n_1-1)}-f_{n_1}}{2+(f_{(n_1-1)}+f_{n_1})} \end{bmatrix} \end{aligned} \tag{3}$$

$$\begin{aligned} \Delta \varphi_1 &= [\varphi_1 - \varphi_2 \quad \varphi_1 - \varphi_3 \quad \cdots \quad \varphi_{(n_1-1)} - \varphi_{n_1}] \\ &= [\delta_1 - \delta_2 \quad \delta_1 - \delta_3 \quad \cdots \quad \delta_{(n_1-1)} - \delta_{n_1}] \end{aligned} \tag{4}$$

The above characteristic coefficients exclude the influence of primary voltage fluctuations and relate only to the value of the metrological error within the estimation population, so they can well characterise the variation of the metrological error of the population. When

the measurement error within the estimated group is in normal operation, the above characteristic parameters will be in a stable state. However, when the measurement error within the estimated group changes due to factors such as capacitance breakdown or abnormal dielectric loss, the above characteristic parameter and its data distribution characteristics will also change, so the online quantitative estimation of measurement error can be achieved by real-time monitoring of the changes in the above characteristic parameter and the degree of change.

In order to validate the error estimation method, it is necessary to add additional measurement errors and corresponding attribute labels to simulate the operation data under different measurement errors, and to construct the set of eigenparameters containing information about different measurement errors. Then, a supervised learning approach is later used to monitor and estimate the eigenparameters of the population in real time, and thus obtain the metrological errors.

The way in which additional measurement errors were introduced to the offline modelling dataset is shown below:

$$U_{xi} = (1 + \Delta f_i)U_x \quad (5)$$

$$\varphi_{xi} = \Delta\varphi_i + \varphi_x \quad (6)$$

where U_{xi} is the magnitude of the x -th meter in the estimated population after the introduction of the additional measurement error, φ_{xi} is the phase of the x th meter in the estimated population after the introduction of the additional measurement error, Δf_i is the i -th additional ratio difference introduced, and $\Delta\varphi_i$ is the i -th additional phase difference introduced.

According to Equation (3) and Equation (4), the feature parameters ΔU_1 and $\Delta\varphi_1$ containing different measurement error information are obtained, and the corresponding measurement error attribute labels are introduced to facilitate the subsequent construction of supervised learning models.

3. Topological modelling and data pre-processing.

3.1. Topological relationship model based on the law of conservation of energy.

Running error is an important parameter of meter operation, which is related to the accuracy of meter measurement and has important guiding significance for meter replacement and correction. At present, electric power companies mainly test the meter running error by two ways: on-site testing and laboratory testing.

To detect errors in electricity meters, power maintenance personnel should remove the meters from the site and send them to a professional organization to test their metrological characteristics. However, testing millions of smart meters is a time-consuming and expensive task. Therefore, this work constructs a remote estimation model for the operation error of metering devices based on topological relationships and uses the massive power data from the collection system to locate the metering devices whose operation error exceeds the limit.

Electricity, water, and gas can be analogized as a flow rate, and the meter, as a metering device for electricity, reads the magnitude of the flow rate. The connection of meters in the low-voltage area is simplified to form a cluster which is usually a tree topology, and the connection relationship between the master meter and the sub-meters constitutes the topology shown in Figure 1, where M_0 is the master meter of the area, M_1, M_2, \dots, M_p is the measurement sub-table of the region.

It is assumed that all nodes in the topology satisfy the law of energy conservation [20, 21]. Under the constraint of energy conservation, the energy flowing into a node is equal to the energy flowing out of the node. Then, the supply of electricity from

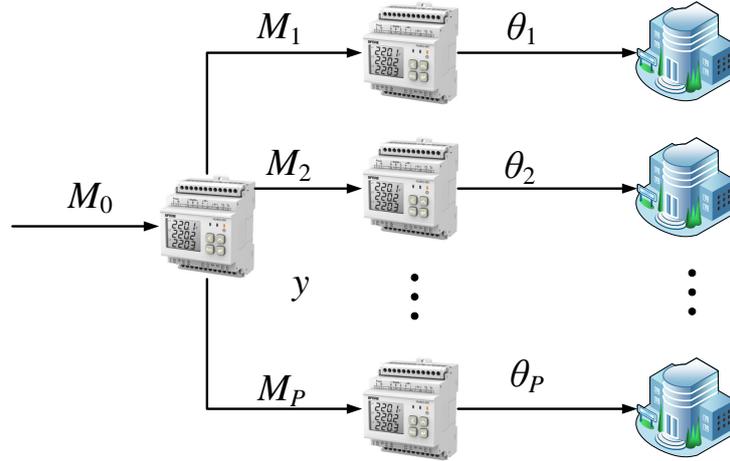


Figure 1. Tree topology for meter clusters

the total meter is equal to all the electricity sold during the same time period, which is approximately equal to the sum of the electricity consumption of each meter. Define $x_j(i)$ as the actual electricity consumption of meter j in the i -th cycle and $y(i)$ as the electricity supply in the i -th cycle. If there exist P users, the relationship between actual power supply and power consumption can be expressed as:

$$y(i) = \sum_{i=1}^P x(i) \quad (7)$$

The relative error of metering point j over the metering period i is shown below:

$$\epsilon_j = \frac{\theta_j(i) - x_j}{x_j} \quad (8)$$

where $\theta_j(i)$ is the measured value of meter j for the i -th cycle.

If there exist P users, the relationship between power supply and power consumption in the whole topological region [22] can be expressed as follows:

$$y = \sum_{i=1}^P \theta_j (1 - \epsilon_j) \quad (9)$$

Simplify the loss model for topological regions by defining leakage losses and line resistance losses as variable losses.

3.2. Data pre-processing. Not all electricity data collected by smart meters are valuable: some of them have null values due to collection or transmission problems, and thus have no value for direct use. In order to mitigate the negative impact of noise and incomplete data on the estimation effect of the model, this paper uses Local Outlier Factor (LOF) [23, 24] to eliminate erroneous values and interpolation to recover incomplete or worthless values.

LOF is the elimination of outliers in terms of the data density of each point. Firstly, by calculating the k -th distance of the point, the local k -th density of the point is calculated on the basis of the k -th distance, and finally the outlier factor of the point is obtained by bringing in Equation (10). The outlier factor indicates the degree of outlierness of a point, and a larger value indicates a higher degree of outlierness of the point; conversely,

a lower degree of outlieriness. Points with a high degree of outliers will be considered as collection outliers and set as null values. For example, in Figure 2, the red point is an outlier, and the size of the red point area indicates the outlier degree of the point.

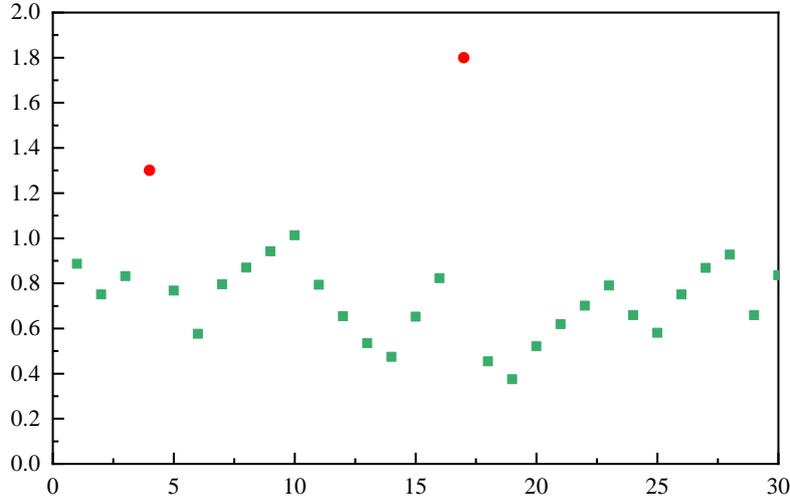


Figure 2. local outlier factor

The expression for the local outlier factor is shown below:

$$LOF_k(x_i) = \frac{\sum_{o \in N_k(x_i)} \frac{lrd_k(o)}{lrd_k(x_i)}}{k} \tag{10}$$

where $N_k(x_i)$ is the k -th distance neighbourhood of the point x_i , $lrd_k(x)$ represents the local k -th density of the point x_i and o is the neighbourhood point.

After the local outlier processing, the outliers of the measurements are set to the null value NaN. In this work, the interpolation method is used to supplement the original data and the null value after the outlier processing and the expression of the interpolation method is shown below:

$$f(x_i) = \begin{cases} \frac{(x_{i-1} + x_{i+1})}{2} & x_i \in \text{NaN}, x_{i-1}, x_{i+1} \notin \text{NaN} \\ 0 & x_i \in \text{NaN}, x_{i-1} \text{ or } x_{i+1} \notin \text{NaN} \\ x_i & x_i \notin \text{NaN} \end{cases} \tag{11}$$

4. Measurement error estimation based on WT and migration neural networks.

4.1. Multiscale analysis based on WT. The time series of electricity data collected by smart meters is a kind of non-smooth and non-linear series containing some noise. Therefore, in order to achieve the goal of accurate prediction, the problem of non-smoothness and non-linearity needs to be solved first. In order to make the original electricity data time series better to be predicted and thus achieve the idea of better prediction results, this work reduces the complexity of the series through multiscale analysis.

Multiscale analysis involves a number of different methods. Fourier spectral analysis requires prior calculation of the frequency parameters of the basis functions when dealing with time series data, and although this method can extract frequency domain features and achieve high resolution, it is difficult to effectively extract signal features in time domain analysis. The wavelet transform is a time-frequency analysis method that allows signals to be analysed in both the time and frequency domains. Translation and scaling are

the basic operations of the wavelet transform, which allows the signal to be transformed in time and frequency to extract the features of the signal. After obtaining the scale function $\varphi(x)$ and wavelet function $\psi(x)$, the functions that can be obtained are shown below:

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^jx - k) \tag{12}$$

If we let \mathcal{W}_j denote the space formed by the set of wavelet functions $\{\psi_{j_0,k} \mid k \in \mathbb{Z}\}$,

$$V_{j_0+1} = V_{j_0} \oplus \mathcal{W}_{j_0} \tag{13}$$

where \oplus denotes the straight sum of the resulting space.

In addition, the basis functions in \mathcal{W}_{j_0} and V_{j_0} are orthogonal.

$$\langle \varphi_{j_0,k}(x), \psi_{j_0,l}(x) \rangle = 0 \text{ for } k \neq l \tag{14}$$

In fact, the connection between the two, scale space and wavelet space, can be written in the following form:

$$\psi(x) = \sum_k h_\psi(k)\sqrt{2}\varphi(2x - k) \tag{15}$$

where $h_\psi = \{h_\psi(k) \mid k = 0, 1, 2, \dots, N\}$ denotes the wavelet function coefficients, which can be viewed as an ordered set.

Wavelet translation is actually the process of decomposition of different orthogonal bases.

$$h_\psi(k) = (-1)^k h_\varphi(1 - k) \tag{16}$$

The WT differs from the Fourier Transform in that it directly replaces the base of the Fourier Transform, turning what was once a trigonometric function into a wavelet base of fixed length. This gives you both the frequency and the moments.

Unlike the Fourier transform, the expression for the wavelet transform is shown below:

$$WT(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * \psi\left(\frac{t - \tau}{a}\right) dt \tag{17}$$

The wavelet transform can be determined in terms of two variables, the scale a and the translation τ . The expansion and contraction of the wavelet function depends on the scale a , and the corresponding translation of the wavelet function depends on τ . The frequency is then derived from the scale, and the moments are seen through the translation, as shown in Figure 3.

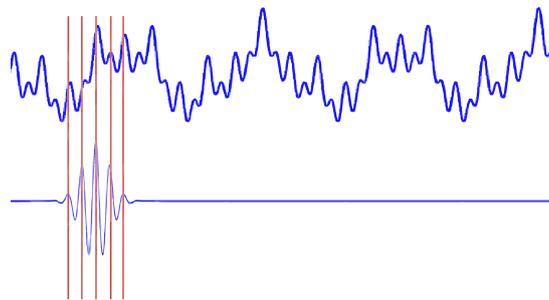


Figure 3. Schematic diagram of WT translation and expansion

As shown in Figure 3, when the sequence reaches such a tight state, it can be multiplied to get a larger value. At this point it is possible to know not only how much frequency is occupied, but also in which part of this tighter section is located. According to this idea, if all the scales are multiplied, we can know the frequency of each part of the whole sequence.

In addition, the mutation signal is also the place where WT is superior to Fourier transform, because in the process of Fourier's transform will inevitably appear the Gibbs effect, that is, transient strong transformation occurs, even because of the Gibson effect occurs in a very short paragraph, it is necessary to spend a lot of trigonometric functions to fit it. However, even with no wireless long trigonometric functions, the problem of not being able to fit accurately occurs, but the WT does not have this problem because of its characteristics.

4.2. LSTM-based estimation model. The LSTM-based estimation model consists of LSTM layer, Dropout layer, and Dense layer.

The inputs to the model are the feature parameters and containing information about the different measurement errors of the meter population, which are then trained by an LSTM neural network model. The model starts by placing an LSTM layer with 80 neural units that returns a three-dimensional tensor, after which a Dropout layer with a random neglect ratio of 0.2 is chosen to avoid overfitting, after which an LSTM layer with 100 neural units is placed and another Dropout layer with a ratio of 0.2 is added, and finally an output is defined as a one-dimensional by Dense Network.

Dropout network is used in the LSTM layer, the Dropout layer prevents overfitting during model training as it controls random disconnection of some nodes connected to the neural network.

$$g = h \cdot D(p) \quad (18)$$

where D stands for the Dropout operation, P is an adjustable hyperparameter, and h denotes the rate at which neural network units are disconnected in advance.

4.3. Migration learning model design. The Ada-LSTM model proposed in this paper is mainly composed of two core parts: temporal distribution description, and temporal distribution matching. Temporal distribution description is the period used to characterise the distribution information in the time series. Temporal distribution matching is to pair the above cycles for subsequent fitting of different temporal distribution cycles to build a time series prediction model.

Firstly, the training data are divided into cycles that can adequately characterise their distributional information through temporal distributional description, and then these cycles are distributionally matched using temporal distribution matching, after which the time series estimation model Ada-LSTM is built, and finally the trained model is used to predict the meter data series.

According to the principle of maximum entropy [27, 28], in the case of covariate transfer, in order to maximise the shared knowledge of the time series, this can be achieved by finding the periods that are least similar to each other. This process is shown in Figure 4 below.

(1) Time distribution description.

The time distribution description achieves this goal of segmenting the original time series by solving an optimisation problem that can be formulated as:

$$\max_{0 < K \leq K_0} \max_{n_1, n_2, \dots, n_K} \frac{1}{K} \sum_{1 \leq i \neq j \leq K} d(D_i, D_j) \quad s.t. \forall i, \Delta_1 < |D_i| < \Delta_2; \sum_i |D_i| = n \quad (19)$$

where d is the distance unit, Δ_1 and Δ_2 are both predefined parameters (to avoid very small values or very large values that may not capture distribution information).

The learning objective of the optimisation problem is to maximise the difference between the average distributions of the cycles while determining the value of K and on the basis of the corresponding cycles so that the distributions of each cycle are as diverse as possible,

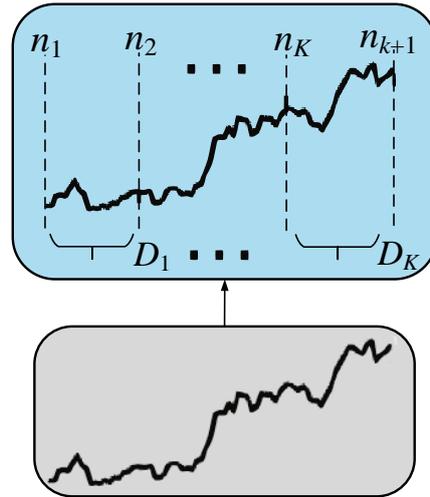


Figure 4. Schematic representation of the time distribution

at which point the predictive model has a better ability to generalise. The principle of segmentation can be explained in detail using the principle of maximum entropy. Firstly, it is necessary to find the most different periods rather than the most similar ones, and in the absence of prior assumptions about the segmentation of time series data, the entropy of the total distribution can only be maximised by diversifying the distribution of each period as much as possible [29], so that a model that can flexibly cope with future data can be built. At the same time, since there is no prior information about the test set data, the model is invisible during the training process, so it is more reasonable to train the model in the worst case scenario, where different cycles can be simulated and the distributions of these cycles can be learnt.

(2) Temporal distribution matching.

The time series distribution matching module was designed to match distributions from different periods by putting them together [30]. Learning common knowledge about different periods and using this knowledge to match their distributions for unknown sequences. As a result, the model trained by this step is able to generalise well on never-before-seen test sets compared to methods relying only on local or statistical information, the structure of which is shown in Figure 5.

The loss function for time-distribution matching prediction can be formulated as follows:

$$L_{\text{pred}}(\theta) = \frac{1}{K} \sum_{j=1}^K \frac{1}{|D_j|} \sum_{i=1}^{|D_j|} l(y_i, M(x_i; \theta)) \quad (20)$$

where (x_i^j, y_i^j) denotes the i -th labelled segment from the time period D_j , $l(\cdot)$ is a loss function (e.g., the MSE loss function), and θ is a model parameter with learning capability.

While matching two LSTM cell distributions, it needs to be able to capture the temporal dependencies. Therefore, this work introduces an importance vector α to learn the relative importance of hidden states within the LSTM, where all hidden states are weighted by a normalised α . Firstly, the network parameters θ are pre-trained using all cycle-to-cycle data, and then this paper uses a boosting-based importance estimation algorithm to learn α as follows:

$$\alpha_{i,j}^{t,(n+1)} = \begin{cases} \alpha_{i,j}^{t,(n)} \times G\left(d_{i,j}^{t,(n)}, d_{i,j}^{t,(n-1)}\right) & \text{if } d_{i,j}^{t,(n)} \neq d_{i,j}^{t,(n-1)} \\ \alpha_{i,j}^{t,(n)} & \text{otherwise} \end{cases} \quad (21)$$

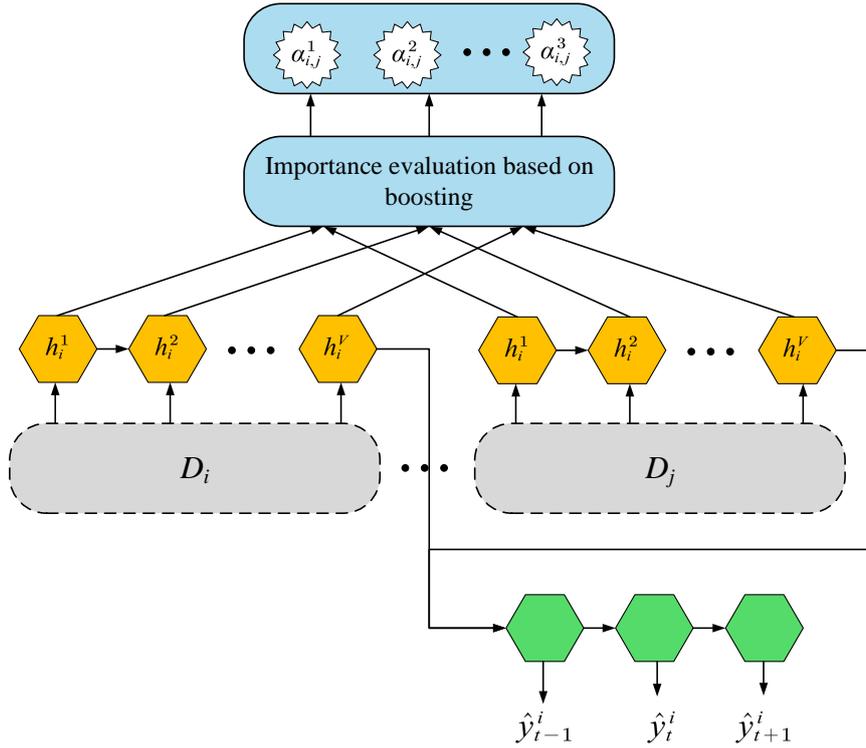


Figure 5. Schematic diagram of time distribution matching

$$G \left(d_{i,j}^{t,(n)}, d_{i,j}^{t,(n-1)} \right) = \left(1 + \sigma \left(d_{i,j}^{t,(n)} - d_{i,j}^{t,(n-1)} \right) \right) \tag{22}$$

where $G \left(d_{i,j}^{t,(n)}, d_{i,j}^{t,(n-1)} \right)$ denotes the update function computed at different learning stages, $d_{i,j}^{t,(n)}$ denotes the difference in distribution at the t -th time step in the n -th iteration, and σ denotes the sigmoid function.

4.4. WT-Ada-LSTM based metering error estimation model. In order to achieve real-time remote monitoring of the metering error of the meter, the WT-Ada-LSTM model is used to monitor and analyse the characteristic covariates in real-time in order to achieve online quantitative estimation of the metering error.

Step 1: The essence of estimating the measurement error using Ada-LSTM model is to perform supervised learning on the feature parameters and containing different measurement error information of the population respectively, which in turn form more representative high-dimensional features.

Step 2: Complete the quantitative estimation of meter metering error based on the good mapping capability of high dimensional features.

Step 3: In order to further improve the accuracy of the prediction results, the proposed algorithm firstly enhances the non-stationarity of the original exchange rate series by decomposing the original exchange rate series into a number of sub-sequences using the WT model.

Step 4: The Ada-LSTM model with migration learning is then used to predict each subsequence separately.

Step 5: Finally, the prediction results of each subsequence are combined and an error reduction operation (SVR) is added.

4.5. Supervised tuning of estimation models. Since the layer-by-layer unsupervised learning can only guarantee that the feature extraction capability of each Ada-LSTM layer is optimal, it cannot guarantee that the performance of the model consisting of multi-layer Ada-LSTMs can also be optimal. Therefore, the weights and biases are fine-tuned using the back-propagation method in the training phase of supervised tuning to optimize the feature extraction capability of the whole multi-layer Ada-LSTM model.

After completing the pre-training of each layer of Ada-LSTM, the output of this model is obtained by combining the classifiers and the loss function is constructed by combining the theoretical output as follows:

$$\text{loss} = \frac{1}{n_2} \sum_{i=1}^{n_2} (\hat{Y}_i - Y_i)^2 \quad (23)$$

where \hat{Y}_i is the true output of the i -th sample in the feature parameter and Y_i is the corresponding theoretical output.

And then the parameters of Ada-LSTM network are fine-tuned using gradient descent as follows:

$$(\omega^{(t+1)}, a^{(t+1)}, b^{(t+1)}) = (\omega^{(t)}, a^{(t)}, b^{(t)}) + \gamma_t \frac{\partial \text{loss}}{\partial (\omega^{(t)}, a^{(t)}, b^{(t)})} \quad (24)$$

where γ_t is the backpropagation learning rate during supervised tuning [31].

5. Experimental results and analyses.

5.1. Experimental design. The amount of electricity meter data is 200 days of data from 30 users. In order to simulate the real residential area operating error over the limit of the meter, the experiment adopts the form of artificial injection of error, that is, embedding a certain amount of offset in the meter measurement value. According to the technical specifications of the smart meter, the meter with over-limit operation error means that the measured value is too large compared with the actual value. The relative error of the meter is more than 2%, which is judged as the meter running over the limit.

After constructing the error estimation model, the 9th and 22nd meters are randomly selected to inject offsets respectively, i.e., their values are artificially altered to simulate meter measurement errors on the basis of the original measurement data, in order to test the effect of the error estimation model as well as to simulate the effect of the presence of meter operation error overruns on the error estimation of other meters in the topology. The error injection is shown in Table 1:

Table 1. Error injection situation

	Injection of offset meters	Injected offset
Example 1	No. 9	4 %
Example 2	No. 22	13 %

5.2. Running error remote estimation. Before and after fault injection, the topology data is used to construct a WT-Ada-LSTM based measurement error estimation model, and the results of the running error estimation are shown in Figure 6 and Figure 7.

Before the error injection, the running error of meter No. 9 is -1.15%, 4% of the error injection, the running error of meter No. 9 is 2.83%, which is consistent with the error injection situation, and verifies the validity of the running error remote estimation model (WT-Ada-LSTM) in this paper. The error injection has an effect on the error solution of other meters, but the effect is small. For example, before and after the error injection,

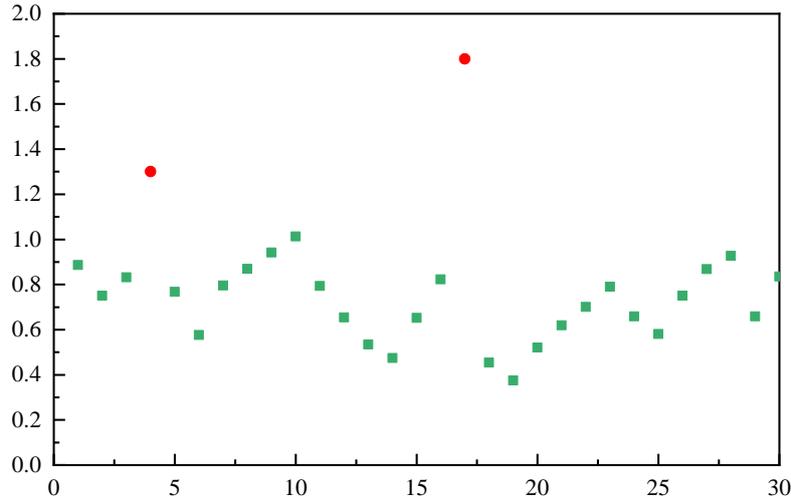


Figure 6. Error estimation result of No. 9

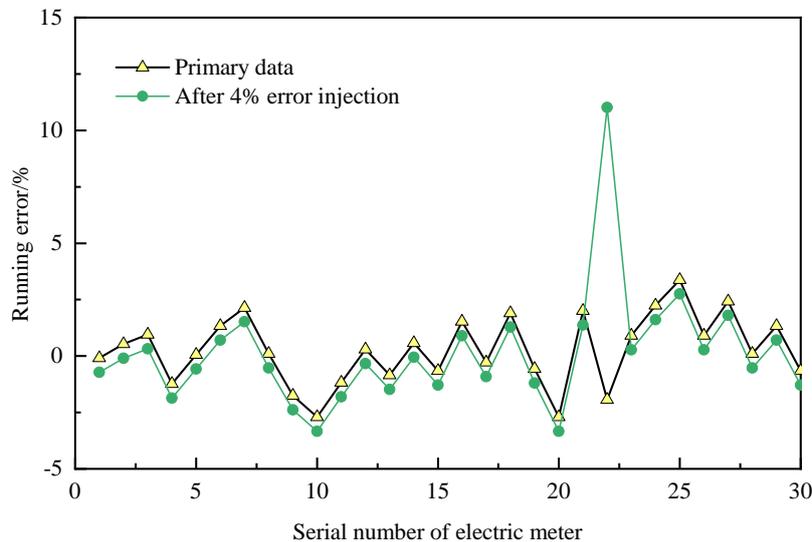


Figure 7. Error estimation result of No. 22

the operating errors of meter No. 17 are -4.7% and -5.3%, respectively. Therefore, the error injection amplifies the error of other meters, but the change is small.

Before the error injection, the operating error of meter No. 22 is -1.93%. 13% of the error injection results in an operating error of 11.02% for meter No. 22. When the error injection is too large, it can be found that the error estimation result of the fault injection point is more accurate, but the deviation of the running error estimation of other meters without fault injection is amplified, which is caused by the perturbation of the importance vector α . In practice, each time the running error is estimated online, a few meters with excessive error values are located first. After replacing the few meters with excessive error values, re-run the remote error estimation again for all the meters in the topological relationship to avoid reducing the effect of error injection. The two sets of experiments verify that the WT-Ada-LSTM model can accurately estimate the operating errors of all meters under the condition of sufficient data.

5.3. Results of quantitative error estimation. The deviations between the estimation results and the true measurement errors for the population of meters in the statistical

topology with different measurement errors are used to analyse the effectiveness of the present quantitative estimation method. The statistical distribution of the estimated deviations for meter No. 22, for example, is shown in Figure 8 and Figure 9.

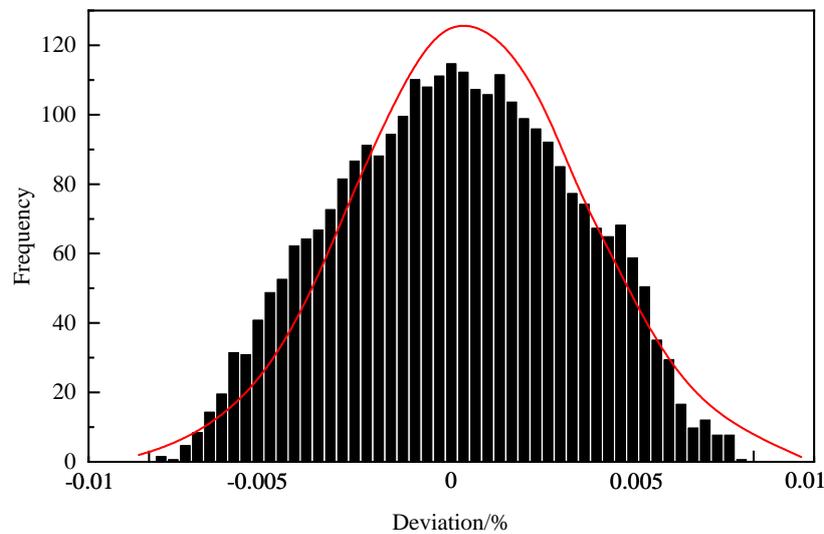


Figure 8. Distribution of deviations between estimated and true values (ratio difference)

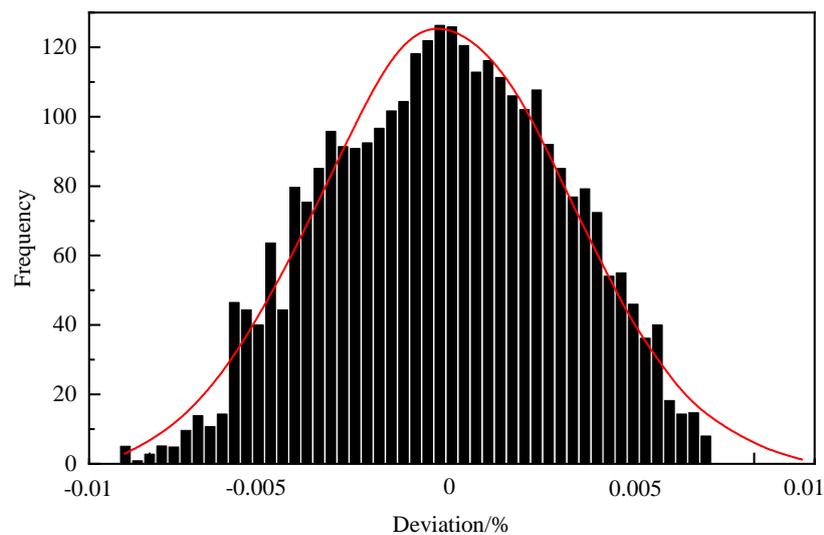


Figure 9. Distribution of deviations between estimated and true values (phase difference)

The mean value of deviation, standard deviation of deviation and maximum value of deviation between the estimation results and the true measurement error for 30 meters under different measurement errors are shown in Table 2.

It can be seen that for the ratio difference deviation and phase difference deviation in the estimation results, the deviation between the results obtained by the proposed remote quantitative estimation method and the true measurement error approximately satisfies the normal distribution, and the maximum values of the mean values are 0.0032% and 0.2871%, respectively, the maximum values of the standard deviation of the deviation are 0.0036% and 0.3514%, respectively, and the maximum values of the deviation are 0.0132% and 0.9853%, so quantitative estimation of measurement error can be achieved.

Table 2. Estimated bias information

	Deviation of ratio difference/%			Deviation of phase difference/%		
	Average value	Standard deviation	Maximum values	Average value	Standard deviation	Maximum values
No.1	0.0026	0.0031	0.0086	0.2971	0.3424	0.9853
No.2	0.0032	0.0036	0.0132	0.262	0.3104	0.7579
No.3	0.0030	0.0034	0.0117	0.2627	0.3089	0.8418
...
No.20	0.0027	0.0032	0.0099	0.2854	0.3414	0.9848

6. Conclusion. A model WT-Ada-LSTM combining deep learning, wavelet analysis and migration learning is proposed in this work. This model not only has the ability to highly fit nonlinear features, but also prevents gradient vanishing and overfitting, which can improve the learning ability of deep learning from the root cause to achieve higher accuracy of error estimation. The experimental results show that although the time series prediction model with the addition of migration learning is more complex than before, the model can learn deeper features in the time series due to the increase of the trained parameters, and thus has better prediction performance. The WT-Ada-LSTM model has a lower estimation error performance and is more adaptable to the nonlinear and nonsmooth time series of meter data. In order to reduce the complexity of the model, subsequent studies will try to ensemble the empirical modal decomposition method to replace the wavelet transform in order to further improve the practicality.

Acknowledgement. This work supported by the Science and technology project of the headquarters of State Grid Corporation of China (No. 5600-2019168a-0-0-00).

REFERENCES

- [1] J. Wu, K. Ota, M. Dong, J. Li, and H. Wang, "Big data analysis-based security situational awareness for smart grid," *IEEE Transactions on Big Data*, vol. 4, no. 3, pp. 408–417, 2016.
- [2] C. Tu, X. He, Z. Shuai, and F. Jiang, "Big data issues in smart grid—A review," *Renewable and Sustainable Energy Reviews*, vol. 79, pp. 1099–1107, 2017.
- [3] K. Wang, Y. Wang, X. Hu, Y. Sun, D.-J. Deng, A. Vinel, and Y. Zhang, "Wireless big data computing in smart grid," *IEEE Wireless Communications*, vol. 24, no. 2, pp. 58–64, 2017.
- [4] A. Tolba, and Z. Al-Makhadmeh, "Predictive data analysis approach for securing medical data in smart grid healthcare systems," *Future Generation Computer Systems*, vol. 117, pp. 87–96, 2021.
- [5] D. Syed, A. Zainab, A. Ghayeb, S. S. Refaat, H. Abu-Rub, and O. Bouhali, "Smart grid big data analytics: Survey of technologies, techniques, and applications," *IEEE Access*, vol. 9, pp. 59564–59585, 2020.
- [6] H. Daki, A. El Hannani, A. Aqqal, A. Haidine, and A. Dahbi, "Big Data management in smart grid: concepts, requirements and implementation," *Journal of Big Data*, vol. 4, no. 1, pp. 1–19, 2017.
- [7] Y. Zhang, T. Huang, and E. F. Bompard, "Big data analytics in smart grids: a review," *Energy Informatics*, vol. 1, no. 1, pp. 1–24, 2018.
- [8] R. Shyam, B. G. HB, S. Kumar, P. Poornachandran, and K. Soman, "Apache spark a big data analytics platform for smart grid," *Procedia Technology*, vol. 21, pp. 171–178, 2015.
- [9] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Communications Magazine*, vol. 50, no. 8, pp. 24–29, 2012.
- [10] Y. Liu, W. Guo, C.-I. Fan, L. Chang, and C. Cheng, "A practical privacy-preserving data aggregation (3PDA) scheme for smart grid," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1767–1774, 2018.
- [11] M. Panda, "Intelligent data analysis for sustainable smart grids using hybrid classification by genetic algorithm based discretization," *Intelligent Decision Technologies*, vol. 11, no. 2, pp. 137–151, 2017.

- [12] M. Ghorbanian, S. H. Dolatabadi, and P. Siano, "Big data issues in smart grids: A survey," *IEEE Systems Journal*, vol. 13, no. 4, pp. 4158–4168, 2019.
- [13] R. F. Arritt, and R. C. Dugan, "Distribution system analysis and the future smart grid," *IEEE Transactions on Industry Applications*, vol. 47, no. 6, pp. 2343–2350, 2011.
- [14] R. Singh, B. C. Pal, R. A. Jabr, and R. B. Vinter, "Meter placement for distribution system state estimation: An ordinal optimization approach," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2328–2335, 2011.
- [15] J. Guo, and Z. Jin, "Autonomous algorithm for relative error of generalized flow meters in tree topology," *Measurement*, vol. 44, no. 9, pp. 1592–1597, 2011.
- [16] F. McLoughlin, A. Duffy, and M. Conlon, "A clustering approach to domestic electricity load profile characterisation using smart metering data," *Applied Energy*, vol. 141, pp. 190–199, 2015.
- [17] S. Shi, Z. Xu, and Y. Xiao, "Load identification method of household smart meter based on decision tree algorithm," *International Journal of Global Energy Issues*, vol. 44, no. 5-6, pp. 440–453, 2022.
- [18] M. Liu, D. Liu, G. Sun, Y. Zhao, D. Wang, F. Liu, X. Fang, Q. He, and D. Xu, "Deep learning detection of inaccurate smart electricity meters: A case study," *IEEE Industrial Electronics Magazine*, vol. 14, no. 4, pp. 79–90, 2020.
- [19] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, "Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network," *Applied Energy*, vol. 282, pp. 116177, 2021.
- [20] C.-M. Chen, Q. Miao, F. Khan, G. Srivastava, and S. Kumari, "Sustainable Secure Communication in Consumer-Centric Electric Vehicle Charging in Industry 5.0 Environments," *IEEE Transactions on Consumer Electronics*, 2023. [Online]. Available: <https://doi.org/10.1109/TCE.2023.3338818>
- [21] S. P. Xu, K. Wang, M. R. Hassan, M. M. Hassan, and C.-M. Chen, "An Interpretive Perspective: Adversarial Trojaning Attack on Neural-Architecture-Search Enabled Edge AI Systems," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 503–510, 2023.
- [22] T.-Y. Wu, Y.-Q. Lee, C.-M. Chen, Y. Tian, and N. A. Al-Nabhan, "An enhanced pairing-based authentication scheme for smart grid communications," *Journal of Ambient Intelligence and Humanized Computing*, 2021. [Online]. Available: <https://doi.org/10.1007/s12652-020-02740-2>
- [23] S.-M. Zhang, X. Su, X.-H. Jiang, M.-L. Chen, and T.-Y. Wu, "A traffic prediction method of bicycle-sharing based on long and short term memory network," *Journal of Network Intelligence*, vol. 4, no. 2, pp. 17–29, 2019.
- [24] A. Boukerche, L. Zheng, and O. Alfandi, "Outlier detection: Methods, models, and classification," *ACM Computing Surveys*, vol. 53, no. 3, pp. 1–37, 2020.
- [25] G. Othman, and D. Q. Zeebaree, "The applications of discrete wavelet transform in image processing: A review," *Journal of Soft Computing and Data Mining*, vol. 1, no. 2, pp. 31–43, 2020.
- [26] M. Jalayer, C. Orsenigo, and C. Vercellis, "Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms," *Computers in Industry*, vol. 125, 103378, 2021.
- [27] C. Li, S. Zhang, Y. Qin, and E. Estupinan, "A systematic review of deep transfer learning for machinery fault diagnosis," *Neurocomputing*, vol. 407, pp. 121–135, 2020.
- [28] R. Zhang, and N. El-Gohary, "A deep neural network-based method for deep information extraction using transfer learning strategies to support automated compliance checking," *Automation in Construction*, vol. 132, 103834, 2021.
- [29] Q. Dai, X.-M. Wu, J. Xiao, X. Shen, and D. Wang, "Graph transfer learning via adversarial domain adaptation with graph convolution," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 5, pp. 4908–4922, 2022.
- [30] Y. Lockner, C. Hopmann, and W. Zhao, "Transfer learning with artificial neural networks between injection molding processes and different polymer materials," *Journal of Manufacturing Processes*, vol. 73, pp. 395–408, 2022.
- [31] M. Subramanian, K. Shanmugavadivel, and P. Nandhini, "On fine-tuning deep learning models using transfer learning and hyper-parameters optimization for disease identification in maize leaves," *Neural Computing and Applications*, vol. 34, no. 16, pp. 13951–13968, 2022.