Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 2893263, 15 pages
https://doi.org/10.1155/2020/2893263

Research Article

A Data Fusion Fault Diagnosis Method Based on LSTM and DWT for Satellite Reaction Flywheel

Dizhi Long,[1] **Xin Wen**,[1] **Junhong Wang**,[1] **and Bingyi Wei**[2]

[1]*College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*
[2]*Beijing Aerospace Automatic Control Institute, China Academy of Launch Vehicle Technology, Beijing, 100854, China*

Correspondence should be addressed to Dizhi Long; longdizhi@nuaa.edu.cn

Received 13 May 2020; Revised 17 July 2020; Accepted 4 August 2020; Published 11 September 2020

Academic Editor: Sajad Azizi

This paper presents a novel fault diagnosis method based on data fusion for a reaction flywheel of the satellite attitude system. Different from most traditional fault diagnosis techniques, the proposed solution simultaneously accomplishes fault detection and identification within parallel fusion blocks. The core of this method is independent fusion block, which uses a generalized ordered weighted average (GOWA) operator to complement the characteristics of output data from long short-term memory (LSTM) neural network and discrete wavelet transform (DWT) so as to enhance the reliability and rapidity of decision-making. Moreover, minibatch normalization is selected to address the problem of covariate shift, realize the adaptive processing of the dynamic information in the original data, and improve the convergence speed of the network. With the high-fidelity model of the reaction flywheel, three common faults are, respectively, injected to collect experimental data. Extensive experiment results show the efficacy of the proposed method and the excellent performance achieved by LSTM and DWT.

1. Introduction

In recent years, the aerospace industry has begun to be interested in preventive maintenance system. A technological shift is ongoing in system monitoring from traditional fixed interval maintenance (FIM) to condition-based maintenance (CBM) system [1, 2]. Compared with the former, CBM eliminates unnecessary preventive scheduled maintenance, reduces maintenance cost, and improves safety and reliability of the system. CBM includes a variety of technologies, such as performance monitoring and fault diagnosis, which ultimately enables the system to detect the faulty components and take appropriate measures [3].

Fault diagnosis has been widely concerned in CBM whose main task is to detect the fault occurrence time and identify the type of fault. Then, the diagnosis information is used to upgrade the maintenance operations from FIM to CBM [4, 5]. Nowadays, the intelligent technology represented by neural network further improves CBM, and thus many intelligent fault diagnosis methods have been studied by experts and scholars. Generally, the neural network-based diagnosis methods can be divided into two categories. One is to design an observer and estimator or fit the system function by the nonlinear fitting ability. Xin et al. [6] used the single hidden layer feedforward wavelet neural network to design an adaptive observer to estimate the state value of the system, whose residual is the detection signal of fault. Inspired by the concept of fault parameters, some researchers proposed a single fault detection method based on neural network fault parameter estimator [7, 8]. On this basis, Sobhani-Tehrani et al. [9] designed a new fault diagnosis structure, which solved the problem of detection and isolation of multiple faults in the system by paralleling multiple parameter estimators. Witczak et al. [10] proposed a method of $H\infty$ observer based on the state-space recurrent neural network (RNN). The approach guaranties a predefined disturbance attenuation level and convergence of the observer, as well as unknown input decoupling and state and actuator fault estimation. The other category adopts the feature extraction ability of neural network to detect and classify faults. The diagnosis procedure generally includes three steps, i.e., data collection, artificial feature extraction,

and health-state recognition [11, 12]. A novel health monitoring system for a variable air volume unit is developed in [13]. After generating fault features for various fault types via fuzzy logic, the artificial neural network classification technique is applied to fault classification. Sobie et al. [14] used training data gained from high-resolution simulations of roller bearing dynamics to train machine learning algorithms. Although simulated data cannot replace experimental data, they provide a stronger starting point for novel applications to further improve fault classifier performance. Barakat et al. [15] introduced the growing neural network to construct a diagnosis model for motor bearings, which obtained the higher diagnosis accuracy for a large number of data when compared with the conventional RBFN and the probabilistic neural network. A new approach named hybrid gradient boosting is proposed in [16] for fault detection in robotic arms. The method is based on the logistic regression model, and random forests, neural networks with machine learning algorithm, and xgboost models are used to boost the base line model.

At present, there is an end-to-end trend in the field of deep learning, which combines feature extraction and classifier design into a neural network. The scheme overcomes the shortcoming of traditional intelligent diagnosis method that needs manual intervention. The motivation of this idea is that neural network can automatically learn both features of original data and classifier so as to better adapt to fault diagnosis and improve performance [17–19]. In order to deal with the dynamic information of time series, recurrent neural network, such as long short-term memory (LSTM) neural network, and its variants have shown excellent performance, especially in the complex time series problems including signal processing, fault classification, and price prediction [20, 21]. By adjusting the inputs and outputs of the LSTM cell with the nonlinear gate units, LSTM can learn the dynamic information of the time series adaptively. However, it should be noted that the accuracy of this kind of method depends on the data's types and prior data of the system.

With the development of sensor technology, the types of system data collected are more abundant, which promote the development of data fusion technology and have been widely applied in different fields [22–24]. It can use the complementarity between data sets to improve the accuracy of decision process. In addition, the cooperative use of overlapping data reduces the uncertainty in the system and leads to more reliable results [25, 26]. Ordered weighted average (OWA) and its variants, as a robust tool to aggregate data from various sources, have become a representative data fusion method [27]. Now, the technology is also gradually applied in the field of fault diagnosis and prognosis to improve the ability of feature extraction and fault decision [28]. Rezamand et al. [29] proposed a novel real-time failure prognosis method for wind turbine bearings in which the OWA operator is applied to combine information obtained from various single features to provide relatively accurate predictions. A new type of fuzzy Petri nets (FPNs) was introduced in [30], which overcomes the deficiencies of traditional FPN by using intuitionistic fuzzy sets (IFSs) and intuitionistic fuzzy ordered weighted average (IFOWA) operators. This makes the model include a wide range of particular cases so that it can effectively handle the various uncertainties in knowledge acquisition and representation. Sánchez-Fernández et al. [31] achieved fault identification based on a scored ranking at two time points: early fault and steady fault. In each case, the OWA linguistic operator based on the regular increasing monotone (RIM) function can find the variables that are responsible for the fault. However, the specific application value of OWA still needs further study.

Reaction flywheel (RW) as the actuator of satellite attitude control system (ACS) is an important guarantee for the successful implementation of space mission. The RW consists of several highly nonlinear internal and external circuits. Generally, only feedback control signal and attitude sensor can be used for monitoring. Therefore, the fault diagnosis and health monitoring of RW is very challenging.

In this paper, the parallel fusion blocks are used to detect and identify the three common faults of RW: motor current reduction, bus voltage insufficient, and work temperature over-high. In each block, a data fusion method based on LSTM and DWT is designed. The main motivation to integrate LSTM and DWT methods is to improve the performance of fault diagnosis.

The main contributions of this paper are as follows:

(1) The parallel structure of fusion blocks reduces the calculation complexity, which allows the network to achieve higher accuracy and speed in identifying the initial fault.

(2) The LSTM can adaptively learn the dynamic information of sequential data. The minibatch normalization method is applied to solve the problem of covariate shift and improve the convergence of LSTM.

(3) The output information of LSTM and DWT is complemented by the GOWA operator, which overcomes the shortage of data types and improves the real-time detection ability and identification accuracy of fault diagnosis system.

2. Minibatch Normalization-Based Vanilla LSTM

Vanilla LSTM, improved form of LSTM, is most commonly used in sequential data prediction and classification [32]. The internal structure of the vanilla LSTM is illustrated in Figure 1.

In Figure 1,c_{t-1} and y_{t-1} represent the prior cell state and output, while c_t and y_t represent the current cell state and output. z_t, z_t^i, z_t^f, and z_t^o are the cell input, input gate, forget gate, and output gate, respectively. $\sigma$and$h$ denote the sigmoid function and tanh function. It can be seen from the figure that vanilla LSTM has similar architecture to RNN, but the former is based on a set of connected cells. Different from the RNN cell which overwrites the information directly, each cell of the vanilla LSTM contains cell input, three gates, and cell output. The cell output is recurrently connected back to the cell input and all the gates. This special design allows the vanilla LSTM to robustly remove or add information during long period of time.

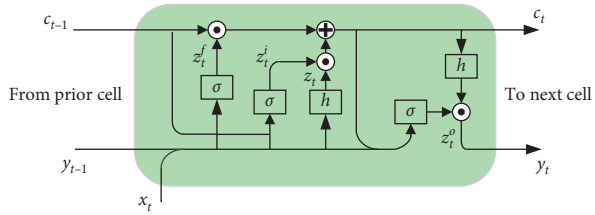

Figure 1: The architecture of the vanilla LSTM.

2.1. Minibatch Normalization. In deep neural network, the change of parameters in one layer usually has a serious impact on the distribution of subsequent layers. We refer to this phenomenon as covariate shift which can be addressed by normalizing layer input [33].

Since the full whitening of each layer's input is costly and not differentiable everywhere, we made two necessary simplifications in this paper. The first is that we normalize each scalar feature independently instead of whitening the features in layer inputs and outputs. For a layer with n-dimensional input $x = (x_1, \ldots, x_n)$, each dimension is normalized as

$$\widehat{x}_k = \frac{x_k - E[x_k]}{\sqrt{\mathrm{Var}[x_k]}}, \tag{1}$$

where the expectation and variance are calculated over the training data set. Note that simply normalizing each input to a layer may change what the layer can represent. To solve this problem, we introduce a set of parameters ρ_k and η_k for each activation x_k, which determine the mean and standard deviation of the normalized features:

$$y_k = \rho_k + \eta_k \widehat{x}_k. \tag{2}$$

When using stochastic optimization, normalizing activations by entire training set is impractical. Thus, the second simplification is that each minibatch produces estimates of the mean and variance of each activation. In this way, the statistical information used for normalization can fully participate in the gradient backpropagation.

Consider a minibatch ϑ of size m:

$$\vartheta = \{x_{1\ldots m}\}. \tag{3}$$

Let the normalized values be $\widehat{x}_{1\ldots m}$ and their linear transformations be $y_{1\ldots m}$, we refer to the transform

$$BN(\rho, \eta, x)\colon\ x_{1\ldots m} \longrightarrow y_{1\ldots m}. \tag{4}$$

The minibatch normalization transform is presented in Algorithm 1.

In Algorithm 1, ε is a regularization parameter added to the minibatch variance for numerical stability.

2.2. Forward Pass. Let x_t be the input at time t, and M and N are the number of inputs and LSTM cells, respectively. According to the architecture of the vanilla LSTM, the equations for layer forward pass can be written as

$$\begin{aligned}
\overline{z}_t &= W_z x_t + R_z y_{t-1} + b_z, z_t = g(\overline{z}_t),\\
\overline{z}_t^i &= W_i x_t + R_i y_{t-1} + p_i \odot c_{t-1} + b_i, z_t^i = \sigma(\overline{z}_t^i),\\
\overline{z}_t^f &= W_f x_t + R_f y_{t-1} + p_f \odot c_{t-1} + b_f, z_t^f = \sigma(\overline{z}_t^f),\\
c_t &= z_t \odot z_t^i + c_{t-1} \odot z_t^f,\\
\overline{z}_t^o &= W_o x_t + R_o y_{t-1} + p_o \odot c_t + b_o, z_t^o = \sigma(\overline{z}_t^o),\\
y_t &= h(c_t) \odot z_t^o,
\end{aligned} \tag{5}$$

where $W_{(\cdot)} \in \mathbb{R}^{N\times M}, R_{(\cdot)} \in \mathbb{R}^{N\times N}, P_{(\cdot)} \in \mathbb{R}^N$, and $b_{(\cdot)} \in \mathbb{R}^N$ represent the input weights, recurrent weights, peephole weights, and bias weights, respectively. σ and h represent the sigmoid function and tanh function; $\odot$ denotes the element-wise multiplication, and the initial states c_0 and y_0 are the parameters of the network.

Apply minibatch normalization to vanilla LSTM. In order to avoid unnecessary redundancy and over fitting, set $\rho = 0$ in minibatch normalization. According to Algorithm 1 and (5), the forward pass of minibatch normalization-based vanilla LSTM (hereinafter referred to as LSTM) is as follows:

$$\begin{aligned}
\begin{pmatrix} \overline{z}_t \\ \overline{z}_t^i \\ \overline{z}_t^f \\ \overline{z}_t^o \end{pmatrix} &= W_x x_t + BN(\eta_h, W_h \cdot y_{t-1}) + W_p \otimes c + b,\\
c_t &= z_t \odot z_t^i + c_{t-1} \odot z_t^f,\\
y_t &= h(BN(\eta_c, c_t)) \odot z_t^o,
\end{aligned} \tag{6}$$

where

$$\begin{aligned}
W_x &= \begin{bmatrix} W_z & W_i & W_f & W_o \end{bmatrix} \in \mathbb{R}^{4N\times M},\\
W_h &= \begin{bmatrix} R_z & R_i & R_f & R_o \end{bmatrix} \in \mathbb{R}^{4N\times N},\\
W_p \otimes c &= \begin{bmatrix} 0 & p_i \odot c_{t-1} & p_f \odot c_{t-1} & p_o \odot c_t \end{bmatrix}^T \in \mathbb{R}^{4N},\\
b &= \begin{bmatrix} b_z & b_i & b_f & b_o \end{bmatrix} \in \mathbb{R}^{4N}.
\end{aligned} \tag{7}$$

Because training data is standardized before training, there is no need to normalize $W_x x_t$.

2.3. Backpropagation through Time. The deltas inside the LSTM cell are calculated as

$$\begin{aligned}
\Delta y_t &= \Delta + R_z^T \Delta z_{t+1} + R_i^T \Delta z_{t+1}^i + R_f^T \Delta z_{t+1}^f + R_o^T \Delta z_{t+1}^o,\\
\Delta z_t^o &= \Delta y_t \odot h(c_t) \odot \sigma'(\overline{z}_t^o),\\
\Delta c_t &= \Delta y_t \odot z_t^o \odot h'(c_t) + p_o \odot \Delta z_t^o + p_i \odot \Delta z_{t+1}^i\\
&\quad + p_f \odot \Delta z_{t+1}^f + \Delta c_{t+1} \odot z_{t+1}^f,\\
\Delta z_t^f &= \Delta c_t \odot c_{t-1} \odot \sigma'(\overline{z}_t^f),\\
\Delta z_t^i &= \Delta c_t \odot z_t \odot \sigma'(\overline{z}_t^i),\\
\Delta z_t &= \Delta c_t \odot z_t^i \odot g(\overline{z}_t),
\end{aligned} \tag{8}$$

Input: $\vartheta = \{x_{1\ldots m}\}$; parameters to be learned ρ, η
Output: $y_i = BN(\rho, \eta, x_i)$
Minibatch variance: $\sigma_\vartheta^2 = (1/m)\sum_{i=1}^{m}(x_i - \mu_\vartheta)^2$
Normalize: $\widehat{x}_i = ((x_i - \mu_\vartheta)/\sqrt{\sigma_\vartheta^2 + \varepsilon})$
Scale and shift: $y_i = \rho\widehat{x}_i + \eta$

Algorithm 1: Minibatch normalizing transform.

where Δ denotes the delta passed down from the upper layer. Only when there is a layer to be trained below, the input deltas are needed, which can be calculated as follows:

$$\Delta x_t = \left(W_z^T \Delta z_t + W_i^T \Delta z_t^i + W_f^T \Delta z_t^f + W_o^T \Delta z_t^o\right). \tag{9}$$

Then, the gradients for weights are computed as

$$\begin{aligned}
\Delta W_{(\cdot)} &= \sum_{t=0}^{T} \langle \Delta(\cdot)_t, x_t \rangle, \\
\Delta p_i &= \sum_{t=0}^{T-1} c_t \odot \Delta z_{t+1}^i, \\
\Delta R_{(\cdot)} &= \sum_{t=0}^{T-1} \langle \Delta(\cdot)_{t+1}, y_t \rangle, \\
\Delta p_f &= \sum_{t=0}^{T-1} c_t \odot \Delta z_{t+1}^f, \\
\Delta b_{(\cdot)} &= \sum_{t=0}^{T} \Delta(\cdot)_t, \\
\Delta p_o &= \sum_{t=0}^{T} c_t \odot \Delta z_t^o,
\end{aligned} \tag{10}$$

where $(\cdot)$ denotes any of $\{z \;\; z^i \;\; z^f \;\; z^o\}$ and $\langle \cdot, \cdot \rangle$ denotes the outer product of two vectors.

3. The Proposed Method for Fault Diagnosis

In this section, a data fusion fault diagnosis method for the actuator of satellite attitude system is proposed. The main purpose is to design a fault diagnosis system to detect and identify the fault in RW. The timely fault diagnosis can not only give the satellite enough time to take appropriate measures before the further development of the fault in the system but also use the fault information to predict the remaining service life of the system. The following content defines the fault set of the actuator of satellite attitude system and introduces the detailed design process.

3.1. Fault Set of the Reaction Flywheel. The RW considered in this paper is ITHACO "type A" reaction flywheel that is manufactured by Goodrich Corporation. A high fidelity nonlinear model of the RW can be obtained from [34] and has been integrated into the ACS dynamics.

Three identical RWs are used in a three-axis stabilized satellite (regardless of the often redundant fourth RW), and each RW has a dedicated fault diagnosis block. Since the three RWs are identical, the results of the fault diagnosis block for only the pitch axis is studied. The training data set can be obtained from the closed-loop ACS simulation of the three-axis stabilized satellite. The moment of inertia of three axes is $I_x = (450\text{kg/m}^2), I_y = (200\text{kg/m}^2)$, and $I_z = (440\text{kg/m}^2)$. The initial angle of the satellite is 5, and the running time is 100 seconds.

In this paper, three types of common faults in RW are considered, including motor current reduction, bus voltage insufficient, and work temperature over-high. Generally, current and voltage faults are transient, while temperature fault accumulates slowly. Thus, fault in the motor current reduction is modeled and injected as variations in motor torque gain k_t. Fault in the bus voltage insufficient is modeled and injected as drops in the voltage of the power bus V_{bus}. Fault in the work temperature over-high is modeled and injected as the slope function in the standard temperature T. In other words, $k_t + \Delta I$, $V_{bus} + \Delta V$, and $T + \Delta T$ are used to replace k_t, V_{bus}, and T, where $\Delta I, \Delta V$, and ΔT are defined the fault parameters, collected in Table 1. Different fault parameters are injected into RW, respectively, to obtain the satellite pitch angle and feedback control signal. These measurements are very precise, and they are less affected by noise. Therefore, we add only 1% white noise to these measurements to mimic more realistic conditions.

3.2. The Data Fusion Fault Diagnosis Method. As Figure 2 illustrates, a parallel fault diagnosis framework is designed to detect and identify the fault.

It is indicated from Figure 2 that three parallel fusion blocks are, respectively, responsible for the diagnosis of motor current reduction, bus voltage insufficient, and work temperature over-high. In each block, the fusion method based on LSTM and DWT is used for fault detection and identification.

3.2.1. The Neural Network. Neural network is an effective method to estimate the complex nonlinear function, which is widely used in the field of fault diagnosis. It constructs a mapping between input and output of the system by the available data set [35]. In this paper, the satellite pitch angle $Y(t)$ and feedback control signal $U(t)$ are selected as inputs of neural network.

With the proposed LSTM, the diagnosis of fault data is straightforward. The offline modeling and online monitoring flowcharts are shown in Figure 3. The procedures of offline modeling and online monitoring are as follows:

Table 1: The RW fault types and health values.

Fault type	Fault parameter	Health value
Motor current reduction	ΔI	0.029 ± 0.006 (N·m/A)
Bus voltage insufficient	ΔV	3 ± 1.5 (V)
Work temperature over-high	ΔT	80 ± 20 (°C)

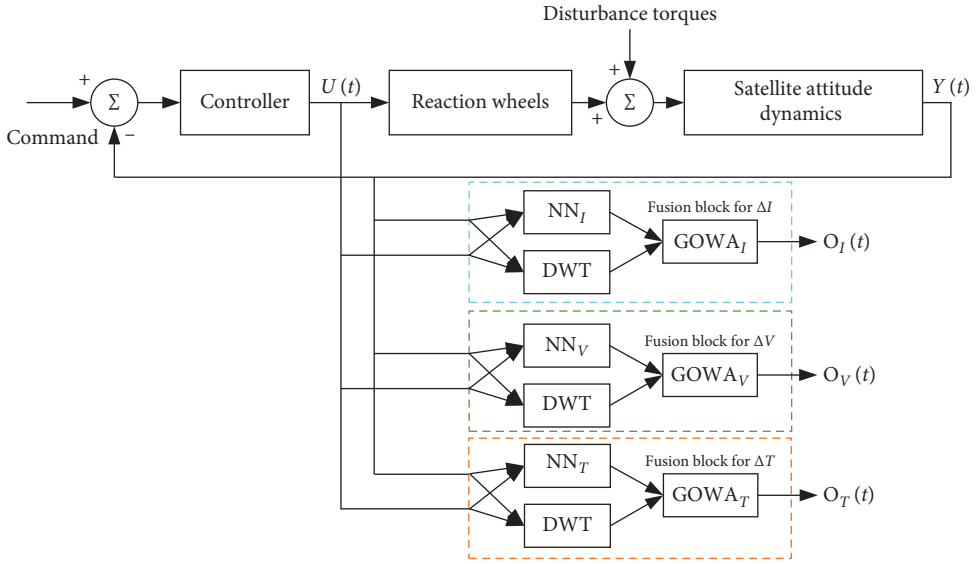

Figure 2: The structure of the proposed data fusion fault diagnosis method.

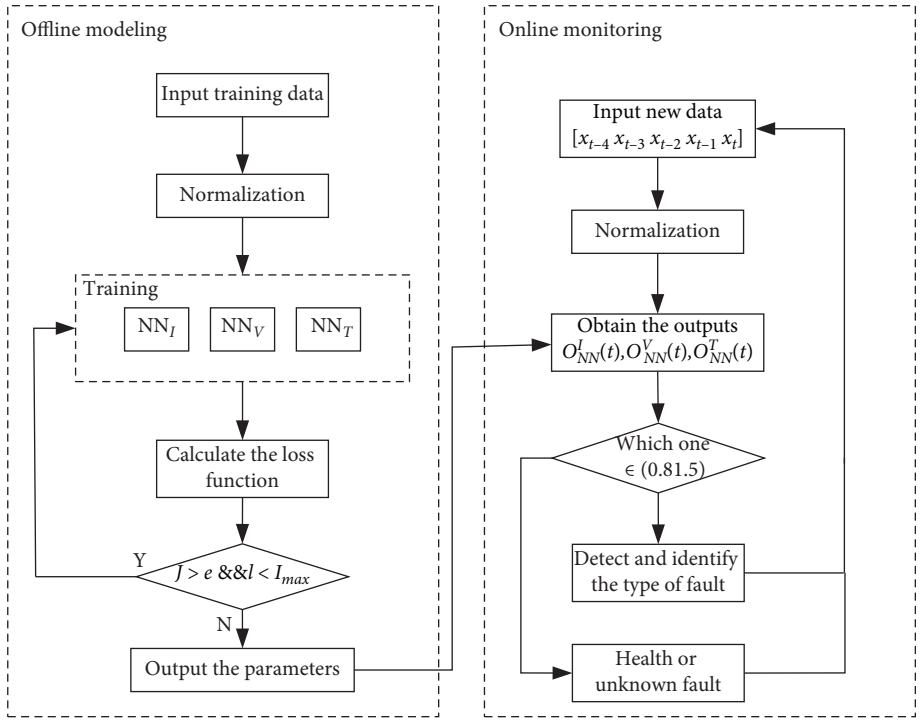

Figure 3: The steps of LSTM for fault diagnosis.

Offline modeling:

(1) Collect data of different fault types as training data.
(2) Normalize each feature of the training data.
(3) Train the three parallel LSTM with Adam.
(4) Calculate the loss function J. If $J > e$ and the number of iterations $l < I_{\max}$, go to (3), where e is a small positive number and $I_{\max}$ is the maximum number of iterations.
(5) Output the parameters of the neural network.

Online monitoring:

(1) Sample a new raw data as testing data:

$$\left[x_{t-4} \;\; x_{t-3} \;\; x_{t-2} \;\; x_{t-1} \;\; x_t\right], \quad t \geq 5. \tag{11}$$

(2) Normalize the testing data.
(3) Obtain the outputs of the NN for motor current reduction, bus voltage insufficient, and work temperature over-high $O^{\mathrm{I}}_{\mathrm{NN}}(\mathrm{t}), O^{\mathrm{V}}_{\mathrm{NN}}(\mathrm{t})$, and $O^{\mathrm{T}}_{\mathrm{NN}}(\mathrm{t})$.
(4) Detect and identify the type of fault according to the outputs.

3.2.2. The DWT Method. Wavelet transform can explore the local characteristics of signals and analyze signals with different time resolution and frequency resolution [36]. DWT is used to discretize the scale and translation of wavelet transform, and it can be used for adaptive time-frequency analysis of nonstationary signals. Moreover, DWT has the ability to capture frequency and location information of the signal. Therefore, DWT is an excellent tool for fault diagnosis.

Consider a signal $f(t) \in L_2(R)$, which can be constructed by the linear combinations of scaling functions and orthogonal wavelets as follows:

$$\begin{aligned} f(t) &= \sum_{i=-\infty}^{+\infty} a_{0,i}\theta_{0,i}(t) + \sum_{\mathrm{m}=-\infty}^{0}\sum_{i=-\infty}^{+\infty} d_{\mathrm{m,i}}\varphi_{\mathrm{m,i}}(t), \\ \theta_{\mathrm{m,i}}(t) &= 2^{(-\mathrm{m}/2)}\theta\left(2^{-m}t - k\right); m, k \in Z, \\ \varphi_{\mathrm{m,i}}(t) &= 2^{(-\mathrm{m}/2)}\varphi\left(2^{-m}t - k\right); m, k \in Z, \end{aligned} \tag{12}$$

where $\theta(t)$ and $\varphi(t)$ represent the scaling functions and orthogonal wavelets, respectively. And $m$and$i$ represent the dilation and translation factors, respectively. Approximation coefficients $a_{0,i}$ and detail coefficients $d_{m,i}$ can be computed by

$$\begin{aligned} a_{0,i} &= \langle f, \theta_{0,i} \rangle, \\ d_{m,i} &= \langle f, \varphi_{m,i} \rangle. \end{aligned} \tag{13}$$

As shown in Figures 4 and 5, the value of the detail coefficient will jump when the fault occurs in the system, and amplitudes are different for various faults. Therefore, the fault type can be identified by analyzing these coefficients.

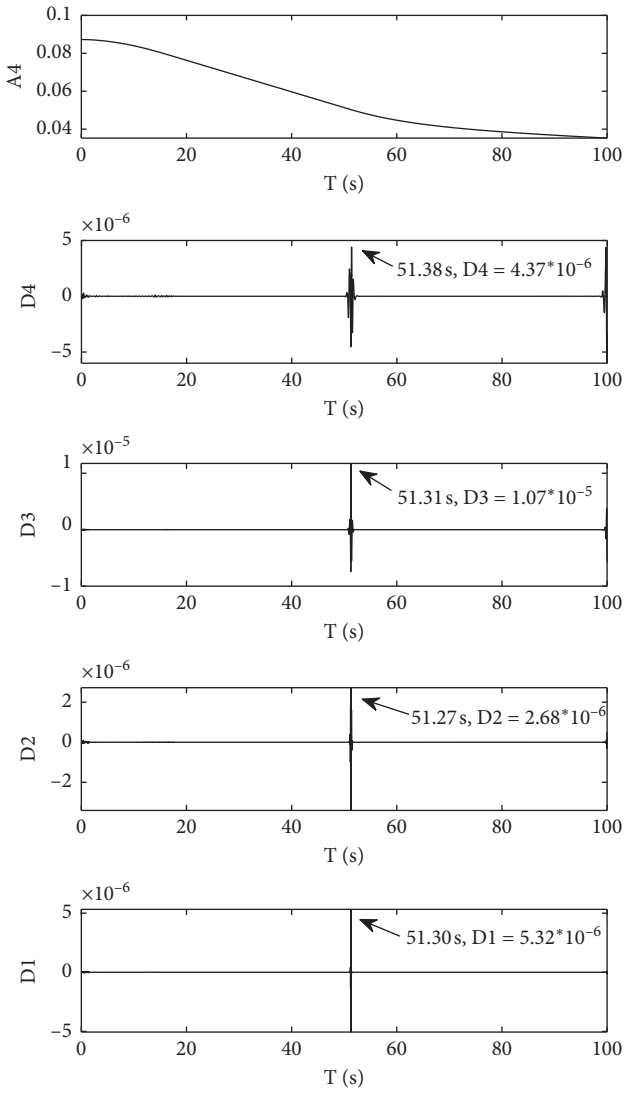

Figure 4: The approximation and detail coefficients of pitch angle in case of motor current fault.

3.2.3. The Fusion Decision Based on GOWA. The GOWA operator can aggregate data information more effectively and sensitively by adding an additional parameter to OWA, so it is considered as a tool for multi-information decision [37–39]. Due to the preprocessing of attitude data by LSTM and DWT, compared with other OWA diagnostic methods, the fusion method based on GOWA does not require to design a complex operator such as using intuitionistic fuzzy sets, linguistic operator, or induced continuous OWA. And only a simple decision process can achieve accurate diagnosis.

Figure 2 shows the internal diagram of the fault diagnosis system based on the data fusion method. Three parallel fusion blocks are used to identify different fault types. In each block, the GOWA operator is used to integrate the decision outputs of LSTM and DWT into a unique framework. The fusion outputs of every block are as follows:

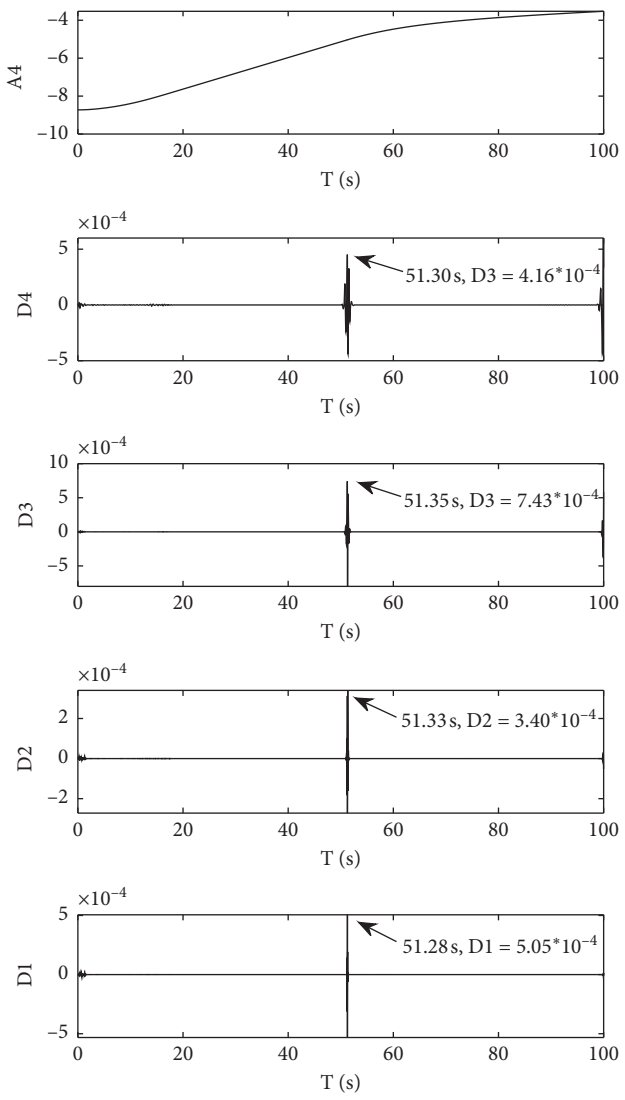

FIGURE 5: The approximation and detail coefficients of control signal in case of motor current fault.

$$O^{\mathrm{I}}_{\mathrm{GOWA}}(t) = \left(\omega_{\mathrm{NN,I}} O^{\mathrm{I}}_{\mathrm{NN}}(t)^{\lambda} + \omega_{\mathrm{DWT,I}} O_{\mathrm{DWT}}(t)^{\lambda}\right)^{(1/\lambda)},$$
$$O^{\mathrm{V}}_{\mathrm{GOWA}}(t) = \left(\omega_{\mathrm{NN,V}} O^{\mathrm{V}}_{\mathrm{NN}}(t)^{\lambda} + \omega_{\mathrm{DWT,V}} O_{\mathrm{DWT}}(t)^{\lambda}\right)^{(1/\lambda)},$$
$$O^{\mathrm{T}}_{\mathrm{GOWA}}(t) = \left(\omega_{\mathrm{NN,T}} O^{\mathrm{T}}_{\mathrm{NN}}(t)^{\lambda} + \omega_{\mathrm{DWT,T}} O_{\mathrm{DWT}}(t)^{\lambda}\right)^{(1/\lambda)},$$
$$\text{subject to :}$$
$$\omega_{\mathrm{NN,I}} + \omega_{\mathrm{DWT,I}} = 1,$$
$$\omega_{\mathrm{NN,V}} + \omega_{\mathrm{DWT,V}} = 1,$$
$$\omega_{\mathrm{NN,T}} + \omega_{\mathrm{DWT,T}} = 1, \tag{14}$$

where $O^{\mathrm{I}}_{\mathrm{GOWA}}(t), O^{\mathrm{V}}_{\mathrm{GOWA}}(t)$, and $O^{\mathrm{T}}_{\mathrm{GOWA}}(t)$ are the outputs of the GOWA operators for motor current reduction, bus voltage insufficient, and work temperature over-high, respectively. $O_{\mathrm{DWT}}(t)$ is the output of the DWT. $\omega_{\mathrm{NN},*}$ and $\omega_{\mathrm{DWT},*}$ ($* \in \{I, V, T\}$) denote the weight factors of the LSTM and DWT for different fault types. Generalized parameter λ is defined as 2.

4. Simulation Analysis

In this section, the effectiveness of the proposed fault diagnosis method for RW of the satellite attitude system is verified by simulation experiments.

4.1. Fault Scenarios. This paper considers three common faults of motor current reduction, bus voltage insufficient, and work temperature over-high in the ACS. Table 1 presents the health value range of these faults, so we can obtain fault data by injecting different faults into the simulation model.

4.1.1. Motor Current Fault. The motor has a torque constant k_{t}, which delivers a torque proportional to the current driver, i.e., $k_{\mathrm{t}} = (\tau_{\mathrm{m}}/I_{\mathrm{m}})$. Therefore, k_{t} can be used to reflect the changes in motor current. When k_{t} drops outside a certain range, the RW cannot provide sufficient control torque, which leads to fault of ACS.

4.1.2. Bus Voltage Fault. The bus voltage $\mathrm{V}_{\mathrm{bus}}$ needs to be set high enough to avoid insufficient voltage margin. When the bus voltage is insufficient, the EMF of the motor will rise so that the maximum torque that the motor can provide decreases. Eventually, it affects the stability of the satellite attitude.

4.1.3. Temperature Fault. Viscous friction τ_v is generated due to the bearing lubricant, and it has a strong sensitivity to the temperature T:

$$\tau_{\mathrm{v}} = (0.0049 - 0.00002(T + 30))\omega. \tag{15}$$

Note that it is the main friction in RW. Therefore, when the temperature exceeds the threshold, the generated friction will reduce the control torque and result in failure.

4.2. Experiment Setup

4.2.1. Implementation of LSTM. As shown in Figure 2, the pitch attitude $Y(t)$ and the feedback control signal $U(t)$ of the satellite are regarded as the inputs of the network with time steps 5. The LSTM contains 12 memory cells, and the output of memory is read through the full connection layer with activation function tanh to generate network output.

In training, the mini batch size and epochs are set to 30 and 50. The result in Figure 6 illustrates that the BN-based LSTM converges significantly faster to the baseline LSTM.

Under operating condition, any output value $\in\in(-0.5\ \ 0.8]$ is regarded as health, and any other value $\in\in(0.8\ \ 1.5)$ is regarded as fault. The value $\notin(-0.5\ \ 1.5)$ is considered as an unknown fault, and the output of the network is set to −2, indicating that an unknown fault is detected in the system.

In order to illustrate the online work procedure of the blocks, it is assumed that a motor current fault occurred. In

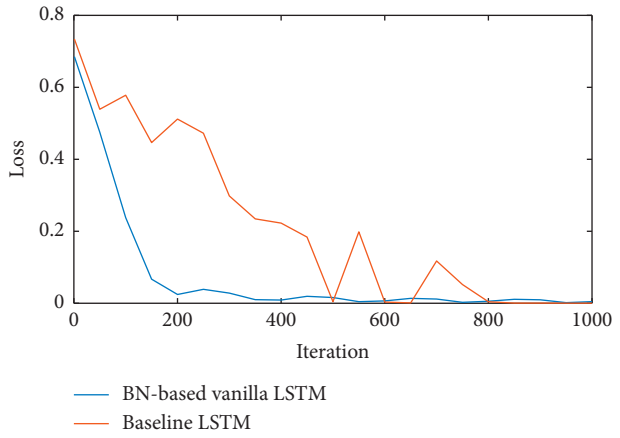

Figure 6: Comparison of convergence between baseline LSTM and BN-based vanilla LSTM for motor current reduction during training.

this case, $O_{NN}^{I}(t) \in (0.8\ \ 1.5)$, while $O_{NN}^{V}(t), O_{NN}^{T}(t) \notin (0.8\ \ 1.5)$. Furthermore, 100 tests are carried out to evaluate the accuracy of the LSTM for different faults. Table 2 illustrates the results, where label $I, V, T, H,$ and U denote motor current reduction, bus voltage insufficient, work temperature over-high, health, and unknown types, respectively. It is indicated from Table 2 that the error rate between voltage fault and temperature fault is higher than others. So, we can solve this problem by combining the outputs of LSTM and DWT.

4.2.2. Implementation of DWT. As mentioned above, the rapidly changing detail coefficients in the event of a fault can be used to detect the fault. For example, consider a motor current fault with the parameter $\Delta I = 0.008$ occurs in the system at $t = 50$s. Wavelet and level are selected to be db4 and 4, and Figures 4 and 5 show the behavior of the approximation and detail coefficients of the fault for satellite attitude and control signal, respectively.

Temperature fault will change very slowly in the process of temperature accumulation, and it is usually difficult to be observed at the beginning of fault. For motor current reduction and bus voltage insufficient, the sudden change of fault parameters will lead to higher detail coefficient. This feature is considered as the difference between them and work temperature over-high. So, the DWT identification mechanism is straightforward. The fault identification logic with DWT is as follows:

For pitch angle,

$$\begin{aligned} D3 &> 5 \times 10^{-6}, \\ D4 &> 4 \times 10^{-6}. \end{aligned} \tag{16}$$

For control signal,

$$\begin{aligned} D2 &> 3 \times 10^{-4}, \\ D3 &> 6 \times 10^{-4}, \\ D4 &> 4 \times 10^{-4}. \end{aligned} \tag{17}$$

Table 2: The RW fault types and health values.

Types	I	V	T	H
Identified as I	0.86	0.02	0.01	0.02
Identified as V	0.05	0.84	0.07	0.03
Identified as T	0.02	0.08	0.85	0.03
Identified as H	0.01	0.02	0.02	0.89
Identified as U	0.06	0.04	0.05	0.03

Table 3: Weighting factors of GOWA operators of different fault types.

Weighting factors	Value
$\omega_{NN,I}$	0.41
$\omega_{DWT,I}$	0.59
$\omega_{NN,V}$	0.37
$\omega_{DWT,V}$	0.63
$\omega_{NN,T}$	1
$\omega_{DWT,T}$	0

Table 4: Parameters of LSTM and DWT.

Parameters	Value
Number of sample	5
Learning rate	0.01
Minibatch size	30
Maximum epochs	50
Performance goal	0.0001
Activation function	Sigmoid and tanh
Number of memory cell	12
Type of wavelet	Daubechies 4

If (15) and (16) are both satisfied, DWT outputs "1," indicating that the system identifies the fault as motor current reduction or bus voltage insufficient. Otherwise, the output is "0". However, DWT cannot detect temperature fault. Thus, this paper combined the DWT method with LSTM as an aid fault diagnosis tool to improve the accuracy of fault diagnosis system.

4.2.3. Weights Determination of GOWA Operator. The function of GOWA operator is to fuse the outputs of LSTM and DWT to make fault decision. These weight factors can be determined by optimizing the following cost functions:

$$\min \sum_{i=1}^{M} \left(W\left(o_1^i, \ldots, o_n^i\right) - T^i \right)^2, \tag{18}$$

where W is a vector of weighting factors, o_n^i is the output of $n-$ th block at $i-th$ iteration, T^i denotes the ideal output of $i-$ th iteration, and "0" represents the health and "1" represents the fault.

Since DWT is unable to identify temperature fault, LSTM is only used to identify temperature fault. Table 3 presents the weighting factors of the fusion method. The design parameters of LSTM and DWT are summarized in Table 4.

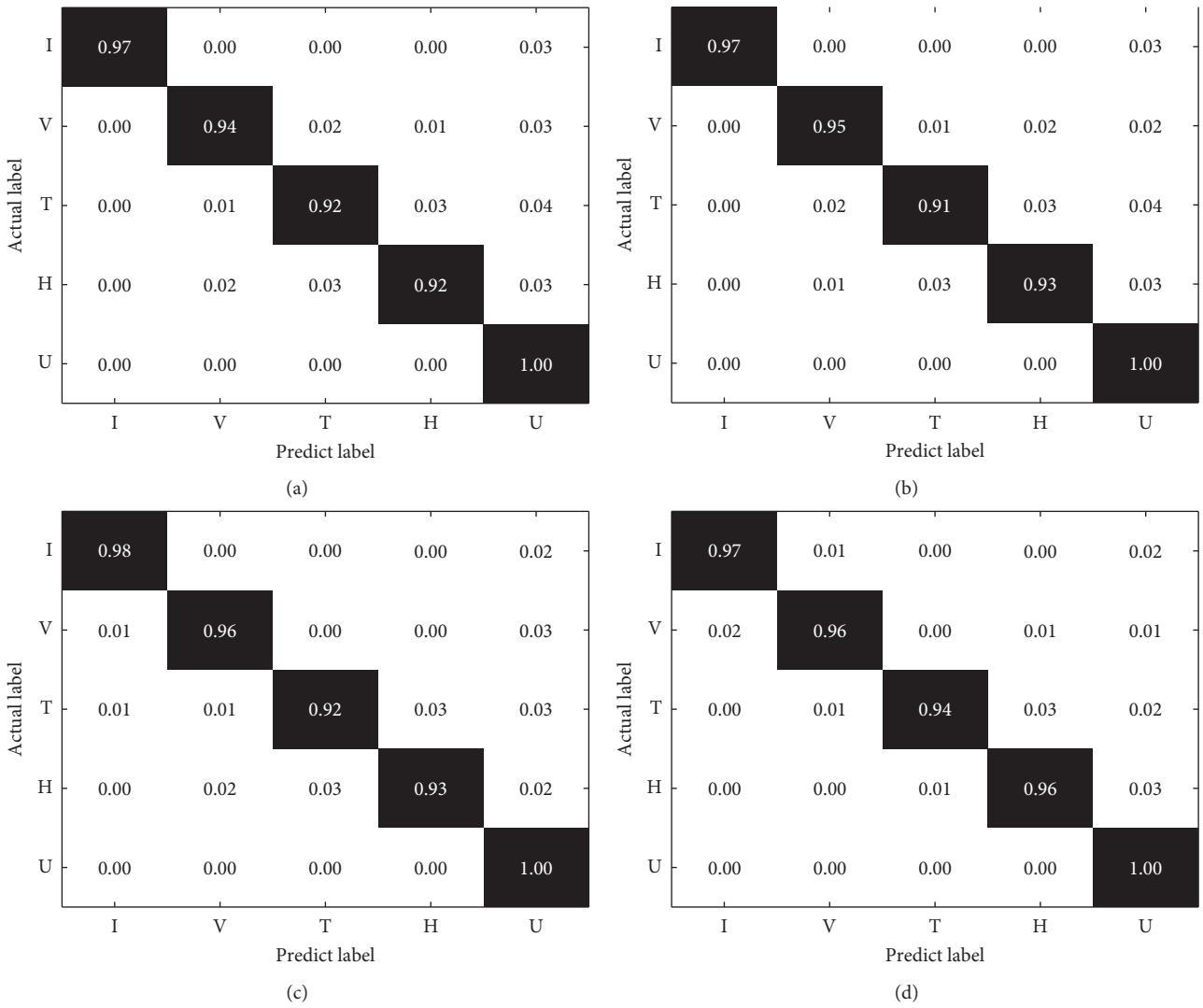

Figure 7: Confusion matrices of different methods. (a) DPCA + SVM. (b) DLDA + SVM. (c) DWT + MLP. (d) Proposed fusion method.

4.3. Performance Evaluation. In the experiment, the fault classification performance of DPCA + SVM, DLDA + SVM, DWT + MLP, and DWT + LSTM is compared. In DPCA and DLDA, 5 samples are concatenated to form the extended vectors. For SVM, we choose RBF as kernel with the parameter $\gamma = (1/\mathrm{d}f)$, where $\mathrm{d}f$ is the number of the extract features of DPCA or DLDA. The DWT + MLP is similar to the proposed fusion method, but LSTM is replaced with MLP. The results of the four fault identification methods are shown in Figure 7. In the confusion matrix, the rows represent the actual label and columns represent the predict label. The diagonal cells show where the actual and predict labels match. The nondiagonal cells show instances of test algorithm errors. It is noted that the unknown fault has no test data, and the default accuracy rate is 1.

It can be seen from Figure 7 that the fusion method can correctly separate 97% of the motor current reduction, 96% of the bus voltage insufficient, 94% of the work temperature over-high, and 96% of the health condition. Compared with traditional "feature extraction + classifier" mode such as DPCA + SVM and DLDA + SVM, the proposed method has higher diagnostic accuracy for different fault types. In addition, with the adaptive processing ability of dynamic information in raw data, it also has a better performance than the fusion method of DWT with conventional single hidden layer feedforward network such as MLP which takes each data independently for training and ignores the correlation information between different data. At the same time, compared with the result of LSTM in Table 2, the proposed fusion method has a significantly lower error rate of fault diagnosis due to fusing the diagnostic information of DWT through GOWA. In summary, the discriminant power in the fusion method is larger than that in DPCA + SVM, DLDA + SVM, and DWT + MLP. Thus, the proposed fusion method should be more suitable and effective for fault identification.

Furthermore, in order to evaluate the real-time performance of fault diagnosis system, various faults consisting of the motor current reduction, bus voltage insufficient, and

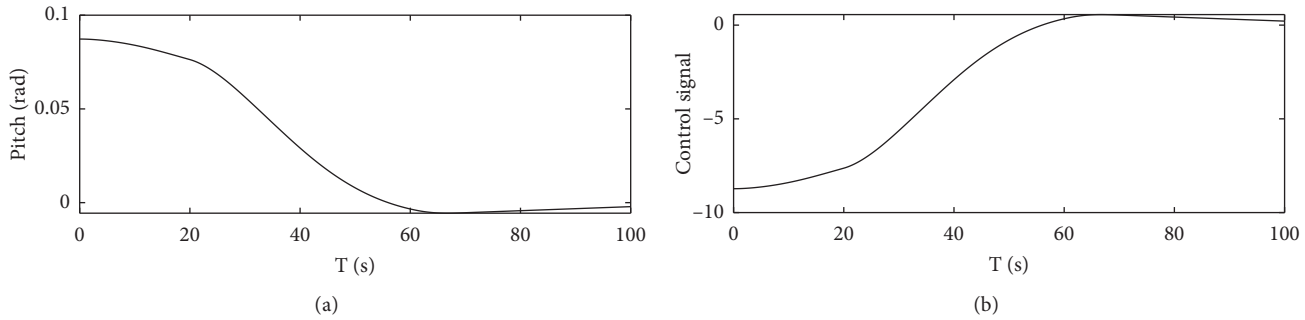

Figure 8: Measurements of pitch and control signal in case of motor current reduction.

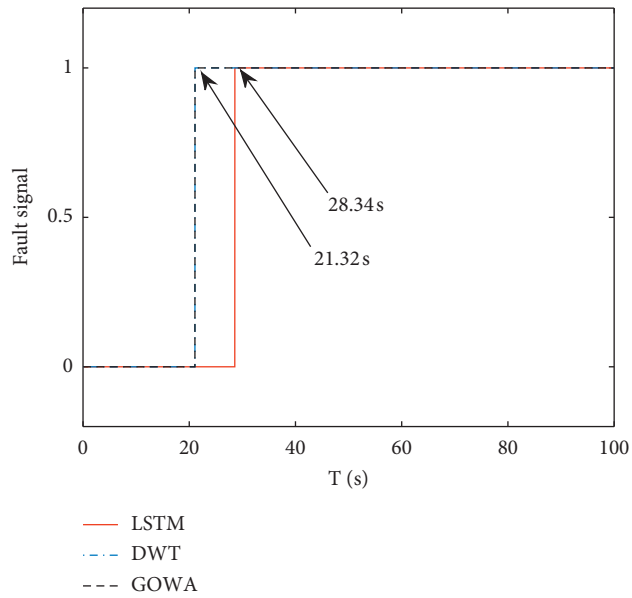

Figure 9: The fault diagnosis signals in the LSTM, DWT, and GOWA for motor current faults.

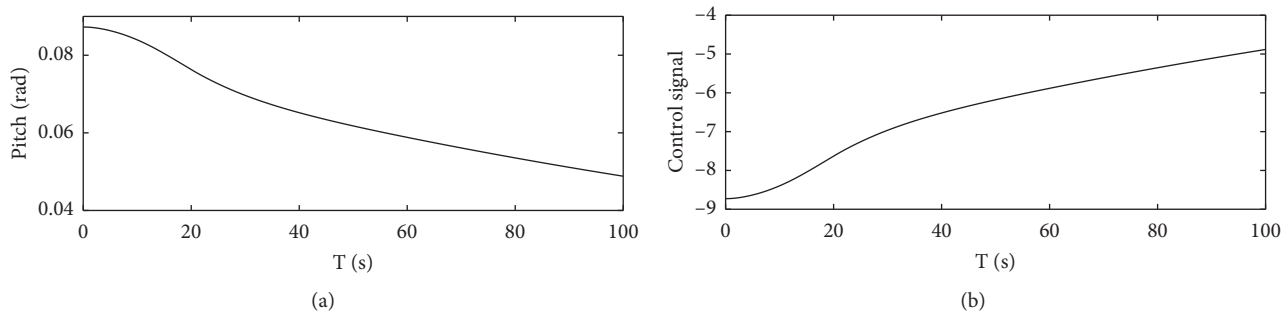

Figure 10: Measurements of pitch and control signals in case of bus voltage insufficient.

work temperature over-high are performed on the RW. Then, the test results obtained by the proposed fault diagnosis system using the control feedback signal and the pitch angle are discussed.

4.3.1. Motor Current Reduction. Consider a motor current fault with $\Delta I = -0.01$ at $t = 20$s. Figure 8 illustrates the change of satellite pitch attitude and feedback control signal in case of motor current reduction. Although the fault

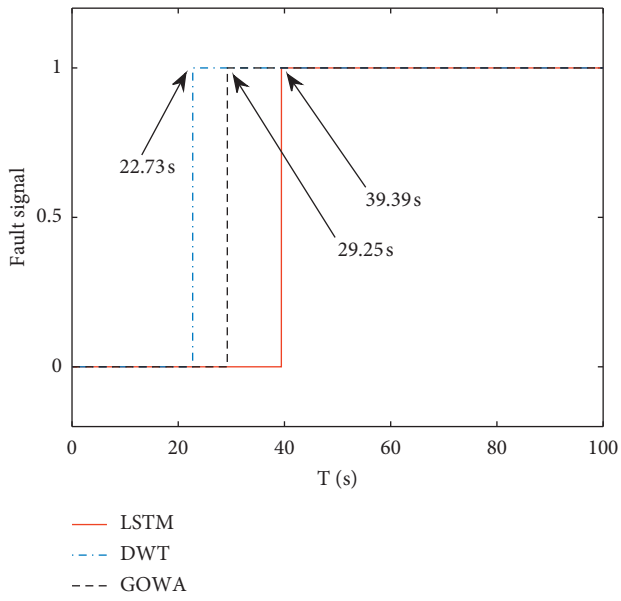

Figure 11: The fault diagnosis signals in the LSTM, DWT, and GOWA for bus voltage insufficient.

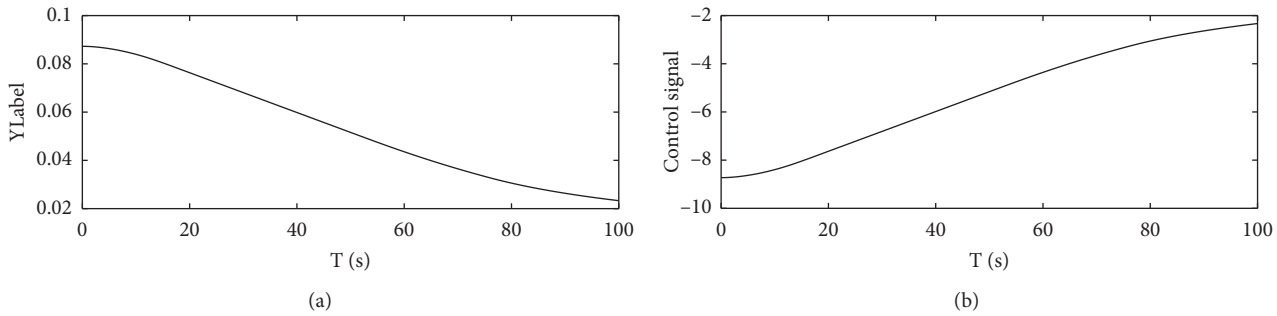

Figure 12: Measurements of pitch and control signal in case of bus voltage insufficient.

parameter is a small value, the fault has a sever effect in the system. Figure 9 presents the fault diagnosis signals in the LSTM, DWT, and GOWA for motor current faults.

Due to the severity of the fault, all fault diagnosis blocks can detect the fault quickly. However, compared with the LSTM method, DWT and GOWA have faster response speed.

4.3.2. Bus Voltage Insufficient. Consider a bus voltage fault with $\Delta V = -2$ at $t = 20$s. Figure 10 indicates the satellite pitch attitude output and feedback control signal. Figure 11 presents the LSTM, DWT, and GOWA output signals under bus voltage fault. Similarly, the detection speed of GOWA is faster than that of the LSTM method.

4.3.3. Work Temperature Over-High. Consider a temperature fault with a slope of 0.4 in the system as follows:

$$\Delta T = \begin{cases} 0, & t < 20, \\ 0, & t \geq 20. \end{cases} \tag{19}$$

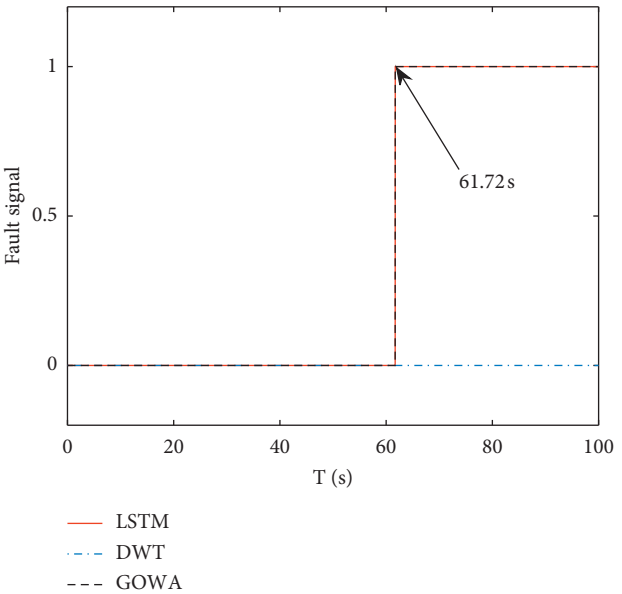

Figure 13: The fault diagnosis signals in the LSTM, DWT, and GOWA for motor current reduction work temperature over-high.

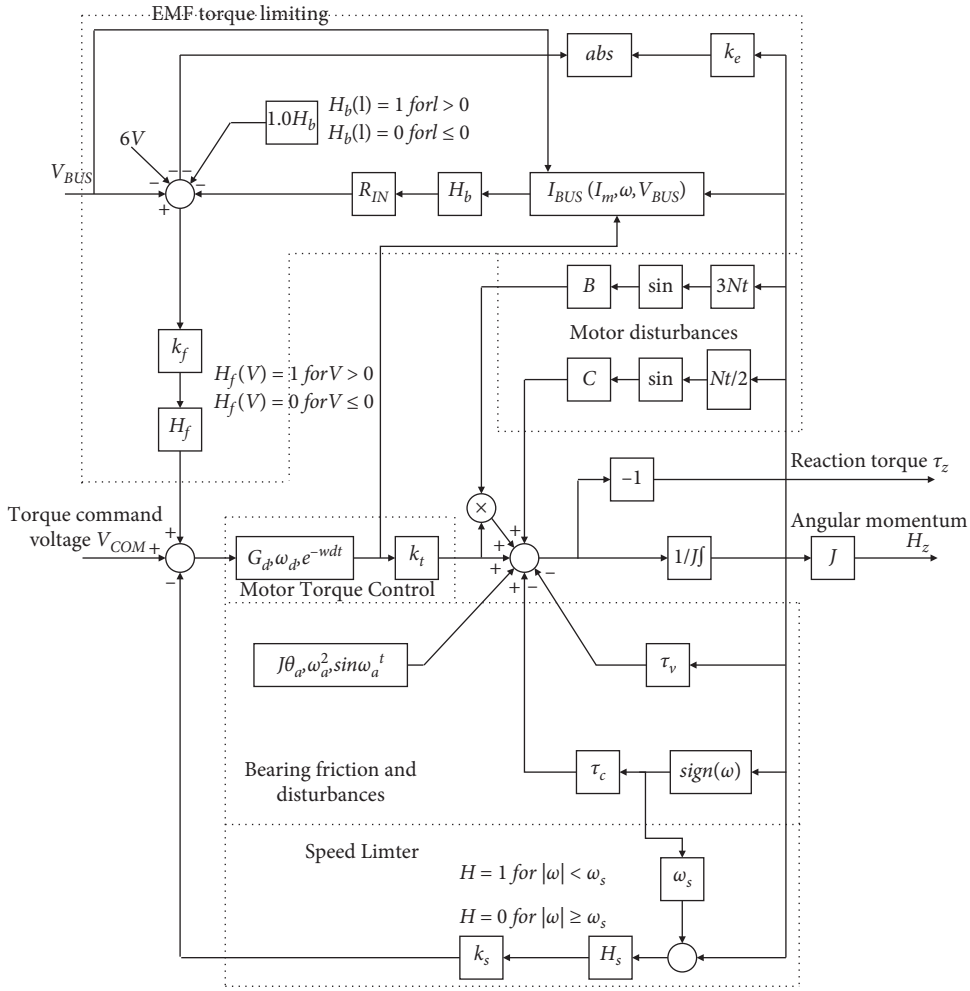

Figure 14: Detailed reaction wheel block diagram.

Table 5: Parameters of the reaction flywheel.

Parameters	Nomenclature	Unit	Value
V_{bus}	Bus voltage	V	—
J	Flywheel inertia	$N{\cdot}m{\cdot}s^2$	0.0077
ω_d	Driver bandwidth	rad/s	20
k_f	Voltage feedback gain	—	0.5
k_s	Overspeed circuit gain	V/(rad/s)	95
ω_s	Overspeed circuit threshold	rad/s	690
τ_c	Coulomb friction	N·m	0.002
k_t	Motor torque constant	N·m/A	0.029
G_d	Driver gain	A/V	0.19
T	Operating temperature	°C	—
ω_a	Torque noise high-pass filter frequency	rad/s	0.2
N	Number of motor poles	—	36
k_e	Motor back EMF	V/(rad/s)	0.029
R_{IN}	Input resistance	Ω	2
B	Cogging torque amplitude	N·m	0
C	Motor torque ripple coefficient	—	0.22
R_B	Bridge resistance	Ω	2
θ_a	Torque noise angle deviation	rad	0.05
P_q	Quiescent power	W	3

Figure 12 illustrates the change of satellite pitch attitude and feedback control signal in case of temperature fault. Figure 13 illustrates the fault signals in the LSTM, DWT, and GOWA as the operating temperature continues to rise.

It is worth noting that due to the slow change of system state caused by fault, DWT alone cannot detect the occurrence of temperature fault. Therefore, only the output of LSTM plays a decisive role in fault detection in this condition.

5. Conclusions

Due to the inherent nonlinearities of RW and satellite attitude dynamics, as well as the impact of disturbance on the satellite, it is a challenging problem to effectively and accurately diagnose the RW. In this paper, a fusion method based on LSTM and DWT is proposed to solve the problem of fault detection and identification of the actuator. Three common faults of RW, motor current reduction, bus voltage insufficient, and work temperature over-high, are researched. Then, three parallel fusion blocks are developed to detect and identify these faults.

In each block, fault information from LSTM and DWT is fused by the GOWA operator, so fault types can be synthetically determined. In addition, due to the use of LSTM, the dynamic information of process data can be used adaptively to improve the reliability of the system. Compared with DPCA + SVM, LDA + SVM, and DWT + MLP algorithms, the fusion method has better performance in fault identification accuracy. Moreover, compared with the single LSTM method, this method has better real-time fault detection ability, more sensitive, and faster response to the fault.

Finally, there will be many directions to be explored about fault diagnosis and prognosis in future. Some of those are as follows:

(1) In the process of fault detection, transfer learning can be used to learn the identified unknown fault to enhance network intelligence

(2) After the fault diagnosis, the prediction of remaining useful life of the system can be further studied by using the time series prediction ability of LSTM

Appendix

A block diagram representation of this RW model is shown in Figure 14.

The RW is mainly composed of three nonlinear parts: EMF torque limiting loop, bearing friction and disturbance, and speed limiter. Let the motor current I_m and angular velocity ω be the states of the flywheel, the voltage signal V_{Com} is output from the controller as the input, and the torque $\tau_{r\omega}$ generated by the RW as the output. The state equation of the RW can be written as

$$\dot{I}_m = G_d\omega_d\left(f_3(\omega, I_m) - f_5(\omega)\right) - \omega_d I_m + G_d\omega_d V_{\mathrm{Com}}, \tag{A.1}$$

$$\dot{\omega} = \left(\frac{1}{J}\right)\left[f_1(\omega) + k_t I_m\left(f_2(\omega) + 1\right) - \tau_v\omega - \tau_c f_4(\omega) + \tau_{\mathrm{noise}}\right], \tag{A.2}$$

$$\tau_{r\omega} = f_1(\omega) + k_t I_m\left(f_2(\omega) + 1\right) - \tau_v\omega - \tau_c f_4(\omega) + \tau_{\mathrm{noise}}, \tag{A.3}$$

where f_1 and f_2 are the functions due to motor disturbances, f_3 is derived from the EMF torque limiting loop, f_4 is the sigmoidal function replacing the discontinuous sign function in the Coulomb friction, and f_5 is the speed limiter. Their expressions are as follows:

$$f_1(\omega) = C\sin\left(\frac{Nt}{2}\right)\omega, \tag{A.4}$$

$$f_2(\omega) = B\sin 3Nt\omega, \tag{A.5}$$

$$f_3(\omega, I_{\mathrm{m}}) = \frac{\exp\left(-aV\left(\omega, I_{\mathrm{m}}, V_{\mathrm{bus}}\right)\right)}{1 + \exp\left(-aV\left(\omega, I_{\mathrm{m}}, V_{\mathrm{bus}}\right)\right)} V\left(\omega, I_m, V_{\mathrm{BUS}}\right), \tag{A.6}$$

$$f_4(\omega) = \frac{1 - \exp(-a\omega)}{1 + \exp(-a\omega)}, \tag{A.7}$$

$$f_5(\omega) = \left(\frac{1}{2}\right)\left(\frac{k_s\left(\omega - \omega_s f_4(\omega)\right)}{1 + \exp\left(-a\left(\omega - \omega_s\right)\right)} + \frac{k_s\left(\omega - \omega_s f_4(\omega)\right)}{1 + \exp\left(a\left(\omega + \omega_s\right)\right)}\right), \tag{A.8}$$

$$V\left(\omega, I_{\mathrm{m}}, V_{\mathrm{bus}}\right) = k_f\left[V_{\mathrm{bus}} - 6 - \frac{1 + R_{\mathrm{in}} I_{\mathrm{bus}}}{1 + \exp\left(-aI_{\mathrm{bus}}\right)} - \frac{1 - \exp\left(-ak_e\omega\right)}{1 + \exp\left(-ak_e\omega\right)} k_e\omega\right], \tag{A.9}$$

where the typical constants are defined in Table 5.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This research was partially supported by the Foundation of Graduate Innovation Center in Nanjing University of

Aeronautics and Astronautics under Grant no. kfjj20181507.

References

[1] S. J. Wu, N. Gebraeel, M. A. Lawley, and Y. Yih, "A neural network integrated decision support system for condition-based optimal predictive maintenance policy," *IEEE Transactions on Systems Man & Cybernetics Part A Systems & Humans*, vol. 37, pp. 226–236, 2007.

[2] G. Vachtsevanos, F. Lewis, M. Roemer, A. Hess, and B. Wu, *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, Wiley, Hoboken, NJ, USA, 2007.

[3] W. Wang, "A two-stage prognosis model in condition based maintenance," *European Journal of Operational Research*, vol. 182, no. 3, pp. 1177–1187, 2007.

[4] S. Choi, E. Pazouki, J. Baek, and H. R. Bahrami, "Iterative condition monitoring and fault diagnosis scheme of electric motor for harsh industrial application," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1760–1769, 2015.

[5] M. A. Siddiqui, S. I. Butt, A. A. Baqai, J. Lu, and F. Zhang, "A novel idea for optimizing condition-based maintenance using genetic algorithms and continuous event simulation techniques," *Mathematical Problems in Engineering*, vol. 2017, pp. 1–10, Article ID 6061234, 2017.

[6] W. Xin, W. Jiayi, and L. Xin, "A feed-forward wavelet neural network adaptive observer-based fault detection technique for spacecraft attitude control systems," *Chinese Journal of Electronics*, vol. 27, no. 1, pp. 102–108, 2018.

[7] A. Barua and K. Khorasani, "Hierarchical fault diagnosis and fuzzy rule-based reasoning for satellites formation flight," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 47, no. 4, pp. 2435–2456, 2011.

[8] M. Mrugalski, "An unscented kalman filter in designing dynamic GMDH neural networks for robust fault detection," *International Journal of Applied Mathematics & Computer Science*, vol. 23, no. 1, pp. 157–169, 2013.

[9] E. Sobhani-Tehrani, H. A. Talebi, and K. Khorasani, "Hybrid fault diagnosis of nonlinear systems using neural parameter estimators," *Neural Networks*, vol. 50, pp. 12–32, 2014.

[10] P. Witczak, K. Patan, M. Witczak, and M. Mrugalski, "A neural network approach to simultaneous state and actuator fault estimation under unknown input decoupling," *Neurocomputing*, vol. 250, pp. 65–75, 2017.

[11] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106587, 2020.

[12] Y. Lei, J. Lin, M. J. Zuo, and Z. He, "Condition monitoring and fault diagnosis of planetary gearboxes: A review," *Measurement*, vol. 48, pp. 292–305, 2014.

[13] W. H. Allen, A. Rubaai, and R. Chawla, "Fuzzy neural network-based health monitoring for HVAC system variable-air-volume unit," *IEEE Transactions on Industry Applications*, vol. 52, no. 3, pp. 2513–2524, 2016.

[14] C. Sobie, C. Freitas, and M. Nicolai, "Simulation-driven machine learning: Bearing fault classification," *Mechanical Systems and Signal Processing*, vol. 99, pp. 403–419, 2018.

[15] M. Barakat, M. E. Badaoui, and F. Guillet, "Hard competitive growing neural network for the diagnosis of small bearing faults," *Mechanical Systems & Signal Processing*, vol. 37, no. 1-2, pp. 276–292, 2013.

[16] M. A. Costa, B. Wullt, M. Norrlöf, and S. Gunnarsson, "Failure detection in robotic arms using statistical modeling, machine learning and hybrid gradient boosting," *Measurement*, vol. 146, pp. 425–436, 2019.

[17] S. J. Qin, "Survey on data-driven industrial process monitoring and diagnosis," *Annual Reviews in Control*, vol. 36, no. 2, pp. 220–234, 2012.

[18] Z. Ge, Z. Song, S. X. Ding, and B. Huang, "Data mining and analytics in the process industry: The role of machine learning," *IEEE Access*, vol. 5, pp. 20590–20616, 2017.

[19] Q. Jiang, X. Yan, and B. Huang, "Review and perspectives of data-driven distributed monitoring for industrial plant-wide processes," *Industrial & Engineering Chemistry Research*, vol. 58, no. 29, pp. 12899–12912, 2019.

[20] Z. Sun, H. Sun, and J. Zhang, "Multistep wind speed and wind power prediction based on a predictive deep belief network and an optimized random forest," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–15, Article ID 6231745, 2018.

[21] C. Meng and L. Jia, "Maxout neurons for deep convolutional and LSTM neural networks in speech recognition," *Speech Communication*, vol. 77, pp. 53–64, 2015.

[22] W. Jiang, B. Wei, X. Qin, J. Zhan, and Y. Tang, "Sensor data fusion based on a new conflict measure," *Mathematical Problems in Engineering*, vol. 2016, no. 8, pp. 1–11, Article ID 5769061, 2016.

[23] D. Novak, M. Mihelj, and M. Munih, "A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing," *Interacting with Computers*, vol. 24, no. 3, pp. 154–172, 2012.

[24] M. Kordestani, M. F. Samadi, M. Saif, and K. Khorasani, "A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4990–5001, 2018.

[25] H. Shi, L. P. Yan, B. Liu, and J. Zhu, "A sequential asynchronous multirate multisensor data fusion algorithm for state estimation," *Chinese Journal of Electronics*, vol. 17, no. 4, pp. 630–632, 2008.

[26] A. Abbasi and R. Eslamloueyan, "Determination of binary diffusion coefficients of hydrocarbon mixtures using MLP and ANFIS networks based on QSPR method," *Chemometrics & Intelligent Laboratory Systems*, vol. 132, pp. 39–51, 2014.

[27] R. R. Yager, J. Kacprzyk, and G. Beliakov, *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*, Springer, Berlin, Germany, 2011.

[28] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani, "Failure prognosis and applications–A survey of recent literature," *IEEE Transactions on Reliability*, pp. 1–21, 2019.

[29] M. Rezamand, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "An integrated feature-based failure prognosis method for wind turbine bearings," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 3, pp. 1468–1478, 2020.

[30] H.-C. Liu, J.-X. You, X.-Y. You, and Q. Su, "Fuzzy Petri nets using intuitionistic fuzzy sets and ordered weighted averaging operators," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1839–1850, 2016.

[31] A. Sánchez-Fernández, G. I. Sainz-Palmero, J. M. Benítez, and M. J. Fuente, "Linguistic OWA and two time-windows based fault identification in wide plants," *Computers & Chemical Engineering*, vol. 115, pp. 412–430, 2018.

[32] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2017.

[33] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, http://arxiv.org/abs/1502.03167.
[34] B. Bialke, "High fidelity mathematical modeling of reaction wheel performance," in *Proceedings of the 21st Annual American Astronautical Society Guidance and Control Conference*, Breckenridge, CO, USA, February 1998.
[35] M. Zakaria and M. Al-Shebany, "Artificial neural network &58; A brief overview," *International Journal of Engineering Research & Applications*, vol. 4, pp. 7–12, 2014.
[36] Y. Zhao, D. Baleanu, C. Cattani, D.-F. Cheng, and X.-J. Yang, "Local fractional discrete wavelet transform for solving signals on cantor sets," *Mathematical Problems in Engineering*, vol. 2013, pp. 1–14, Article ID 560932, 2013.
[37] D.-F. Li, "The GOWA operator based approach to multi-attribute decision making using intuitionistic fuzzy sets," *Mathematical and Computer Modelling*, vol. 53, no. 5-6, pp. 1182–1196, 2011.
[38] L.-G. Zhou and H.-y. Chen, "Generalized ordered weighted logarithm aggregation operators and their applications to group decision making," *International Journal of Intelligent Systems*, vol. 25, no. 7, 2010.
[39] R. R. Yager, "Generalized OWA aggregation operators," *Fuzzy Optimization and Decision Making*, vol. 3, no. 1, pp. 93–107, 2004.