PAPER • OPEN ACCESS

Multisource electrohydraulic servo valve fault status diagnostic algorithm based on a message propagation mechanism

To cite this article: Gao Wei *et al* 2023 *Meas. Sci. Technol.* **34** 055302

View the article online for updates and enhancements.

Multisource electrohydraulic servo valve fault status diagnostic algorithm based on a message propagation mechanism

Gao Wei[1]**, Sun Pengfei**[2] **, Ai Chao**[1,]*****, Wang Lei**[2]**, Chen Lijuan**[2]**, Chen Wenting**[1]**, Zheng Shuwei**[3] **and Yang Dong**[1]

[1] School of Mechanical Engineering of Yanshan University, Qinhuangdao 066004, People's Republic of China
[2] Nanjing Institute of Technology, Nanjing 211167, People's Republic of China
[3] China Hangfa Changchun Control Technology Co., Ltd, Chang Chun 130022, People's Republic of China

E-mail: aichao@ysu.edu.cn

CrossMark

Abstract

Fault identification of electrohydraulic servo valves is crucial to maintain the reliability and safety of high-precision electrohydraulic servo systems. Because the nonlinear characteristics and fault characteristics of electrohydraulic servo systems under noise conditions are implicit, it is difficult to obtain a large number of fault data of electrohydraulic servo valves. Therefore, an electrohydraulic servo valve fault diagnosis model based on characteristic distillation is proposed in this paper. First, the original fault data model is obtained based on an electrohydraulic servo valve fault test platform, the data are standardized, and the data of more than one cycle are extracted using a combination of down sampling and a sliding window for data enhancement. Second, a neural network fault diagnosis algorithm based on stack graph convolution is proposed, which is suitable for detecting different types of states (normal state, wear state, stuck state and coil short-circuit state) of electrohydraulic servo valves. The accuracy of the test set fluctuates between 0.7 and 1.0. Then, because there is a certain relationship between the characteristic smoothing phenomenon of a stack graph convolution model and the number of layers, a multilayer stack graph convolution model is bound to have problems such as model degradation. Therefore, a residual model is introduced into the stack model to improve the convergence speed of the model during the optimization process. The results show that the average accuracy of this method is 100%.

Keywords: electrohydraulic servo valve, fault diagnosis, characteristic distillation, stack plot convolution model, residual model

(Some figures may appear in colour only in the online journal)

***** Author to whom any correspondence should be addressed.

1. Introduction

Electrohydraulic control valves are key hydraulic components in industrial and aerospace applications to control electrohydraulic system movement. Electrohydraulic control valves have become increasingly digital, integrated and intelligent to meet the requirements of Industry 4.0 and the development of automation, digital technology and communication technology [1]. Because hydraulic valves usually work in harsh conditions and are severely disturbed by various paths, it is difficult to detect internal faults in hydraulic valves using traditional hydraulic testing techniques, such as pressure or flow sensors [2]. Therefore, research on intelligent fault diagnosis methods for servo valves is of great significance for improving the service quality, reducing the operation and maintenance costs of hydraulic systems and realizing hydraulic intelligence.

Therefore, much work has been done to improve the fault condition monitoring and diagnosis of hydraulic valves. Shi *et al* proposed an internal model based on an eigenmode function intrinsic mode function (IMF) and weighted densely connected convolutional networks for the wear and leakage of electrohydraulic directional valves due to oil pollution and frequent commutation. The leakage fault diagnosis method identifies the internal leakage type and wear location in an electrohydraulic directional valve [3]. Tang Shengnan *et al* used Bayesian optimization (BO) algorithm to automatically select parameters, and built an adaptive model named convolutional neural network (CNN)-BO based on Gaussian process BO, which can accurately complete the intelligent fault diagnosis of hydraulic pumps [4]. Shi *et al* proposed a two-stage multisensor information fusion method to diagnose the internal fault of hydraulic changeover valves [2] by using a vibration signal analysis method instead of the traditional hydraulic test method to solve problems such as the difficulty in obtaining the fault status of hydraulic valves and the low accuracy of the fault diagnosis of hydraulic valves. Guo, FY and others proposed a fault diagnosis method for reciprocating compressor valves based on a transfer learning convolutional neural network, focusing on the problem of the valve fault status for reciprocating compressors, and the fault recognition rate reached 98.32% [5]. Huang *et al* evaluated the working state of diesel engine valves using vibration response signal analysis and developed a fault diagnosis. They also proposed a method to effectively extract valve seat vibration characteristic parameters from the nonstationary vibration signals of diesel engine surfaces by using local wave decomposition and reconstruction technology, continuous wavelet transform time spectrum images, wavelet transforms and wavelet transform technology. Quantitative monitoring of the valve clearance in diesel engines and the fault diagnosis of abnormal valve clearance are realized [6]. Andrade *et al* developed a new methodology that uses fault emulation to obtain parameters similar to the development and application of methods for diagnosis of actuators in industrial control systems benchmark model for fault diagnosis of pneumatic control valves [7]. Ykla *et al* proposed a shallow-deep integrated fault diagnosis model based on DN50 electric valve signal extraction [8]. Yong *et al*

used the PSO-LeNet model in the CNN model to identify the five common fault states of the hydraulic piston pump using the acoustic signal. This model has the best stability and the highest recognition accuracy [9]. Liu *et al* used the limit learning machine method to diagnose faults such as the wear, jamming, zero deviation and limit of electrohydraulic servo valves, with an accuracy of more than 99% [10]. Jia Chunyu and others adopted a method based on a CNN and long short term memory (LSTM) to realize the fault prediction of electrohydraulic servo valves, and the accuracy of fault identification, such as valve core wear and throttling blockage, reached more than 95% [11]. Wang Wenqing used a method based on a CNN and gate recurrent unit (GRU) to predict the remaining life of electrohydraulic servo valves, and its accuracy reached more than 99% [12]. Chai and Jin used a deep generalized regression network and an improved fruit fly optimization algorithm-η. The method realized the fault diagnosis of electrohydraulic servo valves [13]. Chen *et al* adopted a stacked self-encoding algorithm and realized the fault diagnosis of electrohydraulic servo valves through layer-by-layer greedy training [14]. Based on the Dempster–Shafer theory, Ji *et al* solved the problem of information source conflicts and used a CNN, LSTM, random forest (RF) and other methods to realize the fault diagnosis of electrohydraulic servo valves [15]. Chao *et al* proposed a multi-sensor fusion method using a CNN. The experimental results show that the method improves the fault diagnosis performance of the plunger pump [16]. Chao *et al* adopted short-time Fourier transformation to convert the raw vibration data into spectrograms that act as input images of a modified LeNet-5 convolutional neural network, and proposed a method to convert 3D spectrograms based on spectral characteristics. Noise method. The results show that the method improves the diagnostic performance of the CNN model in noisy environments [17].

The above methods provide a certain direction for the research of hydraulic valve fault diagnosis methods. However, because the fault data of electrohydraulic servo valves are difficult to obtain and the amount of data is relatively large, the diagnostic accuracy is low. At the same time, most of the above algorithm models and methods optimized using laboratory high-sample-force equipment may not be applicable to fault diagnosis equipment. Therefore, to solve the above problems, this paper takes a MOOG G761-3004 electrohydraulic servo valve as the research object, obtains the state data of the electrohydraulic servo valve based on the electrohydraulic servo valve fault online detection platform, normalizes the fault data with different data standardization methods, and solves the problem of multisource data of the electrohydraulic servo valve. A data enhancement method is used to solve the problem of limited fault data. Aiming at the diversity and closeness of the fault states of electrohydraulic servo valves, taking the normal state, wear state, stuck state and coil short-circuit state of electrohydraulic servo valves as the research states, a neural network fault diagnosis algorithm based on a message mechanism is proposed. The obtained system state vectors are used as training data, and the adjacency matrix is used to describe the relationship between vectors in each state.

A residual model is introduced into the stack diagram classification model to improve the convergence speed and accuracy of the fault diagnosis model of electrohydraulic servo valves.

This paper mainly studies the fault diagnosis algorithm of electrohydraulic servo valves. The first part expounds the fault mechanism of electrohydraulic servo valves, and the second part obtains the fault data of electrohydraulic servo valves. The third part is the fault diagnosis algorithm and model verification, and the last part is the conclusion.

2. Model introduction

2.1. Working principle of electrohydraulic servo valves

As shown in figure 1, an electrohydraulic servo valve can be basically divided into the following structures: a force/torque motor, hydraulic amplifier (pilot stage and power stage), and feedback/balance mechanism.

In this paper, a force feedback two-stage electrohydraulic servo valve is the research object. The hydraulic amplifier part of this type of valve consists of a double-nozzle bezel and a four-way spool valve, in which the double-nozzle bezel valve, which drives the armature iron rotation through a torque motor, controls the pressure of the two cavities to control the movement of the four-way spool, and the four-way spool is connected to the gear-plate assembly through the feedback rod.

When the input signal is zero, the armature iron is located in the middle of the upper and lower magnetic conductors, and the valve core is in the middle due to the restraint of its structure. When the input signal is not zero, the armature iron produces an electromagnetic torque. When the armature iron starts to turn counter clockwise, the bezel leaves the middle, and the spring tube bends. At this time, the throttle area of the right nozzle decreases, and the throttle area of the left nozzle increases, which causes the left-side pressure in the valve core pressure control cavity to decrease and the right-side pressure to increase. This causes the valve core to start to move to the left. At this time, as the spring tube continues to bend, when the spring tube feedback torque and electromagnetic moment offset each other, the armature iron bezel forces balance. The valve core will continue to move to the left, the feedback lever continues to bend, and the bezel starts time moving toward the mid-range. The valve core is subjected to an increase in pressure in the left cavity and the pressure in the right cavity is reduced. When the liquid pressure at both ends, the reaction force of the spring tube deformation, and the hydraulic power received by the valve core are balanced with each other, the valve core stops moving and is in a balanced position. The schematic is shown figure 2.

2.2. Mathematical model of an electrohydraulic servo valve

The basic equation for the torque motor circuit is:

$$2K_{\mathrm{u}}u_{\mathrm{g}} = (R_{\mathrm{c}} + r_{\mathrm{p}})\Delta i + 2K_{\mathrm{b}}\frac{\mathrm{d}\theta}{\mathrm{d}t} + 2L_{\mathrm{c}}\frac{\mathrm{d}\Delta i}{\mathrm{d}t} \qquad (1)$$

where, K_{u} is the unilateral gain of the amplifier, u_{g} is the amplifier input voltage signal (V), R_{c} is the resistance of each coil (Ω), r_{p} is the amplifier resistance in each coil loop (Ω), Δi is the current difference between two coils (A), θ is the armature iron angle (rad), K_{b} is the back electromotive force(EMF) constant of each coil (Wb), and L_{c} is the self-inductance coefficient of each coil (H).

Formula (1) is a Laplace transform, and the equation of moment motor motion is sorted as follows:

$$\Delta I = \frac{2K_{\mathrm{u}}U_{\mathrm{g}} - 2K_{\mathrm{b}}s\theta}{(R_{\mathrm{c}} + r_{\mathrm{p}})\left(1 + \frac{s}{\omega_{\mathrm{a}}}\right)} \qquad (2)$$

where, ω_{a} is the turning frequency of the control coil loop (Hz).

The motion equation for the armature baffle assembly is:

$$T_{\mathrm{d}} = J_{\mathrm{a}}\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + B_{\mathrm{a}}\frac{\mathrm{d}\theta}{\mathrm{d}t} + K_{\mathrm{a}}\theta + T_{\mathrm{L1}} + T_{\mathrm{L2}} \qquad (3)$$

where, T_{d} is the electromagnetic moment (N $\times$ mm), J_{a} is rotational inertia of the armature iron plate assembly (kg $\times$ mm^2), B_{a} is the viscous damping coefficient of the armature baffle assembly (N $\times$ s mm^{-1}), K_{a} is the spring tube stiffness (N mm^{-1}), T_{L1} is the load torque generated by the fluid flow of the nozzle to the bezel (N $\times$ mm), and T_{L2} is the load torque of the armature baffle assembly caused by feedback rod deformation (N $\times$ mm).

After the Laplace transformation of formula (3), the transfer function of the armature baffle assembly is obtained as follows:

$$\theta = \frac{\frac{1}{K_{\mathrm{mf}}}}{\frac{s^2}{\omega_{\mathrm{mf}}^2} + \frac{2\zeta_{\mathrm{mf}}}{\omega_{\mathrm{mf}}}s + 1}\left(K_{\mathrm{t}}\Delta I - K_{\mathrm{f}}(r+b)X_{\mathrm{v}} - rA_{\mathrm{N}}p_{\mathrm{Lp}}\right) \qquad (4)$$

where, ω_{mf} is the inherent frequency of the torque motor (Hz), ζ_{mf} is the mechanical damping ratio of the torque motor, K_{mf} is the combined stiffness of the torque motor (N mm^{-1}), K_{t} is electromagnetic force coefficient, $K_{\mathrm{t}} = B_{\mathrm{g}}\pi DN_{\mathrm{c}}$. B_{g} is magnetic induction intensity in working air gap, D is average diameter of coil, N_{c} is control coil turns. K_{f} is feedback rod stiffness, r is center distance of thin wall part of spring tube, b is the distance from the center of the feedback rod ball to the center of the nozzle, A_{N} is the nozzle hole area, P_{LP} is the load pressure difference between the two nozzle cavities.

The relationship between the bezel displacement and bezel transfer is:

$$X_{\mathrm{f}} = r\theta. \qquad (5)$$

The nozzle bezel valve control spool can be thought of as a valve-controlled hydraulic cylinder model with an inelastic load, so the transfer function of the spool is:

$$\frac{X_{\mathrm{v}}}{X_{\mathrm{f}}} = \frac{K_{\mathrm{qp}}/A_{\mathrm{v}}}{s\left(\frac{s^2}{\omega_{\mathrm{hp}}^2} + \frac{2\zeta_{\mathrm{hp}}}{\omega_{\mathrm{hp}}}s + 1\right)} \qquad (6)$$

where, K_{qp} is the inherent frequency of the torque motor (Hz), A_{v} is the spool core area (mm^2), ω_{hp} is the hydraulic inherent
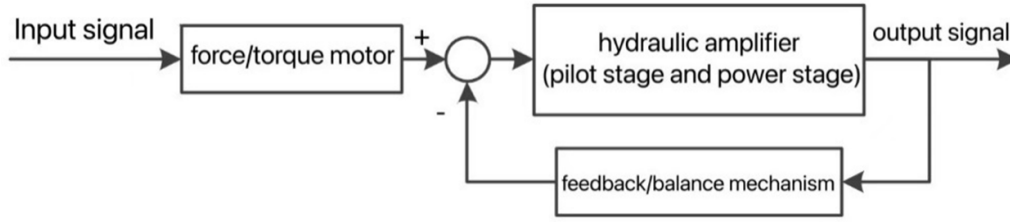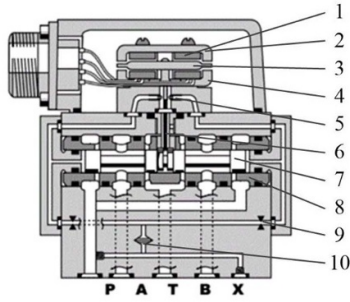
Figure 1. Working principle of an electrohydraulic servo valve.

1—Permanent magnet 2—Upper magnetic conductor 3—Armature 4—Lower magnetic conductor 5—Injector
6 — Spring tube7—Spool valve core 8—Slide valve sleeve 9—Fixed orifice 10—Butterfly filter

Figure 2. Force feedback two-stage electro-fluid servo valve diagram.

frequency of the spool (Hz), and ζ_{hp} is the hydraulic damping ratio of the spool.

Servo valves usually take current Δi as input parameter and no-load flow $q_0 = K_q x_v$ as output parameter. In this case, the transfer function of the servo valve can be expressed as:

$$\frac{Q_0}{\Delta I} = \frac{K_{sv}}{\left(\frac{s}{K_{vf}}+1\right)\left(\frac{s^2}{\omega_{mf}^2}+\frac{2\zeta'_{mf}}{\omega_{mf}}s+1\right)} \tag{7}$$

where, Q_0 and ΔI are Laplace transforms of q_0 and Δi, respectively. K_{sv} is flow gain of servo valve, $K_{sv} = \frac{K_t K_q}{(r+b)K_f}$, K_{vf} is open loop amplification factor of force feedback loop, $K_{vf} = \frac{r(r+b)K_f K_{qp}}{A_v[K_{an}+K_f(r+b)^2]}$, K_{an} is the net stiffness of the torque motor.

The four-sided spool flow equation is:

$$q_L = C_d W x_v \sqrt{\frac{1}{\rho}\left(p_s - \frac{x_v}{|x_v|}P_L\right)} \tag{8}$$

where, q_L is the spool valve load flow (l min^{-1}), C_d is the spool valve throttle port flow factor (l (min MPa)$^{-1}$), and W is the spool valve throttle port area gradient (mm).

2.3. Analysis of the failure mechanisms

2.3.1. Nozzle bezel valve fault. A nozzle bezel valve fault is simple, and the common fault modes are partial or total blockage of the nozzle or throttle hole and filter blockage.

The fault mode and fault characteristics of the nozzle bezel valve are shown in table 1.

Table 1. Nozzle bezel valve fault mode and fault characteristics.

Fault mode	Fault characteristics
Nozzle or throttle hole is partially or completely blocked	Frequency is reduced, the resolution is reduced, and the system is seriously unstable
Filter element is blocked	Causes the frequency to drop, and the resolution decreases causes the system to seriously oscillate

Table 2. Power-stage spool fault modes and fault characteristics.

Fault mode	Fault characteristics
Edge wear	Leaks, fluid noise, large zero bias, system instability
Radial valve core sleeve wear	Leaks increase, zero bias increases, and gain decreases
Spool valve stuck	Waveform distortion, stuck

2.3.2. Power-stage spool valve fault. Faults of the power-stage spool valve mainly appear on the valve core and valve sleeve, which mainly include edge wear, radial filter element wear and slide valve sticking.

The fault modes and fault characteristics of the power-stage spool are shown in table 2.

3. Acquisition of the electrohydraulic servo valve fault data

Because the simulation data and real data are different, it is particularly important to obtain real fault data of electrohydraulic servo valves, and the fault data often contain considerable redundant information; therefore, it is necessary to solve a series of problems, such as the acquisition of the fault data of electrohydraulic servo valves and pre-processing.

3.1. Sample of electrohydraulic servo valve faults

3.1.1. Design of the pilot scheme. In this experiment, a common force feedback two-stage electrohydraulic servo valve is selected, and its model is Moog G761-3004. This model of servo valve is shown in figure 3. And the specific parameters of this type of electrohydraulic servo valve are shown in table 3.

Figure 3. Moog G761-3004 electrohydraulic servo valve.

Table 3. Mood-G761-3004 key technical parameters.

Technical parameter	Parameter range (value)
Work pressure	$\leqslant 31.5$ MPa
Rated flow error $\Delta p_n = 7$MPa	$\pm 10\%$
Symmetry	$< 10\%$
Resolution	$< 0.5\%$
Hysteresis loop	$< 3.0\%$

1-Tank 2-Hydraulic pump 3-Overflow valve 4-Cooler 5-Filter 6-One-way valve 7-Flow meter

8-Pressure sensor 9-Measured electrohydraulic servo valve 10-Variable throttle valve

Figure 4. Principles of the electrohydraulic servo valve test system.

3.1.2. Electrohydraulic servo valve test system. This paper mainly relies on an intelligent electrohydraulic servo valve performance test platform to obtain the fault data of the electrohydraulic servo valve. Its schematic diagram is shown in figure 4.

The upper test system is the test software. The test software realizes the static performance test and dynamic performance test, data storage and playback, report generation and so on.

The data acquisition module includes a data acquisition card and signal generator. The signal generator sends the command signal *i* sent by the data acquisition card or the command signal *i* generated by itself to the servo valve and feeds back the command signal *i* and valve core displacement feedback *s* to the data acquisition card. The data acquisition card collects the servo valve output flow *q*, valve inlet pressure p_p, valve outlet

pressure p_T, valve load port pressures p_A and p_B, servo valve command signal *i* and servo valve spool displacement feedback *s* during the hydraulic test system and transmits them to the upper computer.

The main components of the hydraulic test system are the oil source, which provides the test bench with a constant pressure, constant temperature and clean oil source to ensure that the servo valve can work properly, and the test bench. The test bench is equipped with flow sensors and four pressure sensors to measure system flow and pressure at the P, A, B and T valves of the electrohydraulic servo valve. The test bench also has adjustable flow valves to simulate servo valve loads, with different valve blocks compatible with different types of servo valves, to test the characteristics of the electrohydraulic servo valve. The test bench is shown in figure 5.

Four common states of the G761 electrohydraulic servo valve are selected: normal state, wear state, stuck state and coil short-circuit state.

During the test of the electrohydraulic servo valve, four pressure sensors and a flowmeter are used to measure the pressure and system flow at ports P, T, A and B. The command signal controls the servo valve through the servo amplifier and feeds back the command signal and valve core displacement to the upper data acquisition system. These data reflect the operation state of the electrohydraulic servo valve over a period of time, and some of them can be used as training samples. The internal leakage and flow characteristic curve of the electrohydraulic servo valve is shown in figure 6.

From figure 6 and formulas (7) and (8), the pressure, flow and other operating data in different states are obtained by testing the internal leakage characteristics and no-load flow characteristics of the servo valve.
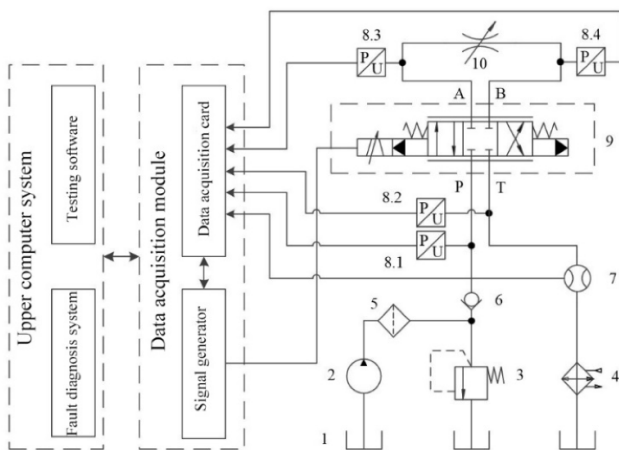
3.2. Data pre-processing

To prevent gradient explosion, the data need to be standardized after the test data are obtained, and the commonly used standardization methods are linear standardization using formula (9) and zero mean standardization using formula (10),

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{9}$$

$$x' = \frac{x - \mu}{\sigma}. \tag{10}$$

Both methods perform linear scaling of the data. Although the relative distribution state of the whole batch of data is not changed, if the data are divided into multiple batches for this operation, data with different characteristics may exhibit the same or similar characteristics after standardization. To preserve the data characteristics, different normalization methods are needed for each sensor's data.

The first method is to standardize the electrical signal of the sensor directly. For most pressure sensors, the feedback electrical signal is generally 0–10 V, or ± 10 V. For a gear flow meter, the feedback signal is often a pulsed signal, and

a) Electrohydraulic servo valve table

b) Servo valve signal generator

c) Operator′s station

d) Hydraulic oil source

Figure 5. Electrohydraulic servo valve test bench.

a. Internal leakage characteristics

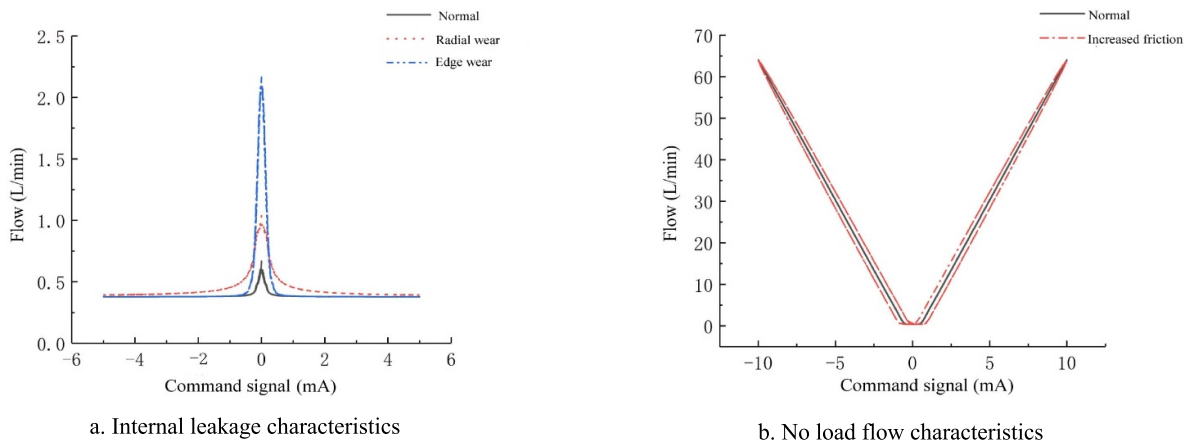b. No load flow characteristics

Figure 6. Test characteristic curve.

through an isolation conversion module, it can also be converted to 0–10 V, so the sensor signal can be linearly reduced.

This method is suitable for data generated by the same type of servo valve, but due to differences in performance of different types of servo valves, especially flow control valves such as the 3001, 3002, 3003, 3004, 3005 in the Moog-G761 series, the load flows rated at the rated pressure drop $\Delta p_N = 7$MPa are: 4 l min^{-1}, 10 l min^{-1}, 19 l min^{-1}, 38 l min^{-1}, 63 l min^{-1}, respectively. If sensor signal standardization is used, different scale characteristics will be produced when the flow characteristics of different types of servo valves are characterized.

To make the data more standardized and uniform, the system pressure is usually adjusted to the rated pressure drop Δp_N of the servo valve during testing, and the load flow q_n under the rated pressure drop can be detected using the model. The standardization of the pressure and flow data can be expressed as formulas (11) and (12):

$$p' = \frac{p}{\Delta p_n} \tag{11}$$

$$q' = \frac{q}{q_n}. \tag{12}$$

If the maximum input electrical signal of the servo valve is rated as the test signal amplitude s_N, the standardization of the command signal is:

$$s' = \frac{s}{s_n}.$$ (13)

If the output maximum electrical signal of the servo valve displacement feedback sensor is the maximum forward displacement feedback d_{max} and the output minimum electrical signal is the maximum negative displacement feedback d_{min}, then the standardization of the displacement feedback is:

$$d' = \frac{d - d_{min}}{d_{max} - d_{min}}.$$ (14)

The standardized formula of the displacement feedback is slightly similar to the linear standardized formula, but there are differences. Linear standardization normalizes the data of a batch, where x_{max} and x_{min} are the maximum and minimum values in the batch, respectively, but the displacement feedback standardization marks the maximum positive displacement of the valve as 1 and the maximum negative displacement as 0, and the spool displacement will return to (0,1) after standardization. When there is no full-scale spool displacement in the whole batch, it will be standardized to (0,1), and the same data in different batches will be standardized to obtain different characteristics.

After data standardization, data nodes need to be built, and each data node represents the characteristics of the system at a certain moment because servo valves can be expressed as $N^{(i)} = \left(s^{(i)}, d^{(i)}, f^{(i)}, p_p^{(i)}, p_a^{(i)}, p_b^{(i)}, p_t^{(i)}\right)$. In this node vector, some characteristics of the current servo valve are already included, the test data can be constructed according to this method to fully connect them to the state vector collection that describes the characteristics of the servo valve. The collection element collapses and transforms to different feature dimensions to obtain different static characteristics of the servo valve. The data set can contain more characteristics than a single sensor data set, and through propagation and learning between nodes higher-order characteristics can be learned based on the guidance of the data set labels.

3.3. Data enhancement

Due to the limited test conditions, more electrohydraulic servo valve fault data cannot be obtained, so data enhancement is required. The results of the normalization of the electrohydraulic servo valve sample data are shown in figure 7.

For the electrohydraulic servo valve fault sample data, the data are enhanced by combining down sampling with a sliding window to extract data for more than one cycle.

According to the analysis of the data during the operation of the electrohydraulic servo valve, the clear characteristics of most of the data do not exist in a high-frequency band, so it can be down sampled. With the total length of the data unchanged, a set of data is broken down into sets of data by equidistant extraction to increase the sample size of the data.

After down sampling, a slide window extraction runs the data over a period of time, which generalizes the data and avoids the impact of phases of periodic data on the network model in post deployment applications. The enhanced results are shown in figure 8.

4. Neural network fault diagnosis algorithm based on a message propagation mechanism

In the traditional neural network algorithm, the input data of most network neurons is matrix operation in the form of one-dimensional vector, so the traditional neural network is often applied to the identification of single dimensional fault data, but too long data will cause too many training parameters and reduce the network performance.

The traditional neural network algorithm has a good effect in dealing with single dimension data, but there are some deficiencies in the face of multisource information system. Convolutional neural network can increase the number of data channels to extract multisource information, but the increase of the number of data channels ultimately leads to the increase of the amount of calculation in the process of data pre-processing and neural network calculation.

When the system state of a multisource information system is treated as a state node at every moment in a certain period of time, there must be some relationship between the nodes that can reflect the system characteristics.

In static test of electro-hydraulic servo valve, no-load flow characteristic test, as shown in figure 9, two state vectors S_1 and S_2 are selected when flow is equal. The angle between S_1 and S_2 is θ. If θ is not zero, then hysteresis exists in the current system. If more node characteristics in higher dimensions are selected, the fundamental failure that causes hysteresis can be determined by machine learning.

Similarly, for multisource information system, the system node state is taken as the input of neural network, and the neural network is trained through supervised learning to complete the fault diagnosis task of the system.

As shown in figure 10, to realize neural network fault diagnosis based on system state node data, it is necessary to find a neural network model based on learnable node characteristics, and the graphical neural network based on message propagation network (MPNN) can meet the above requirements.

The acquired system state vectors are used as training data, and the adjacency matrix is used to describe the relationship between the vectors in each state. For the multisource information system of the electro-hydraulic servo valve, the CNN [18] has advantages in processing multi-dimensional data. The advantage of graph convolutional networks (GCN) over CNN is that it can process more complex and irregular data structures. It can learn node feature information and structure information end-to-end at the same time. It is the best
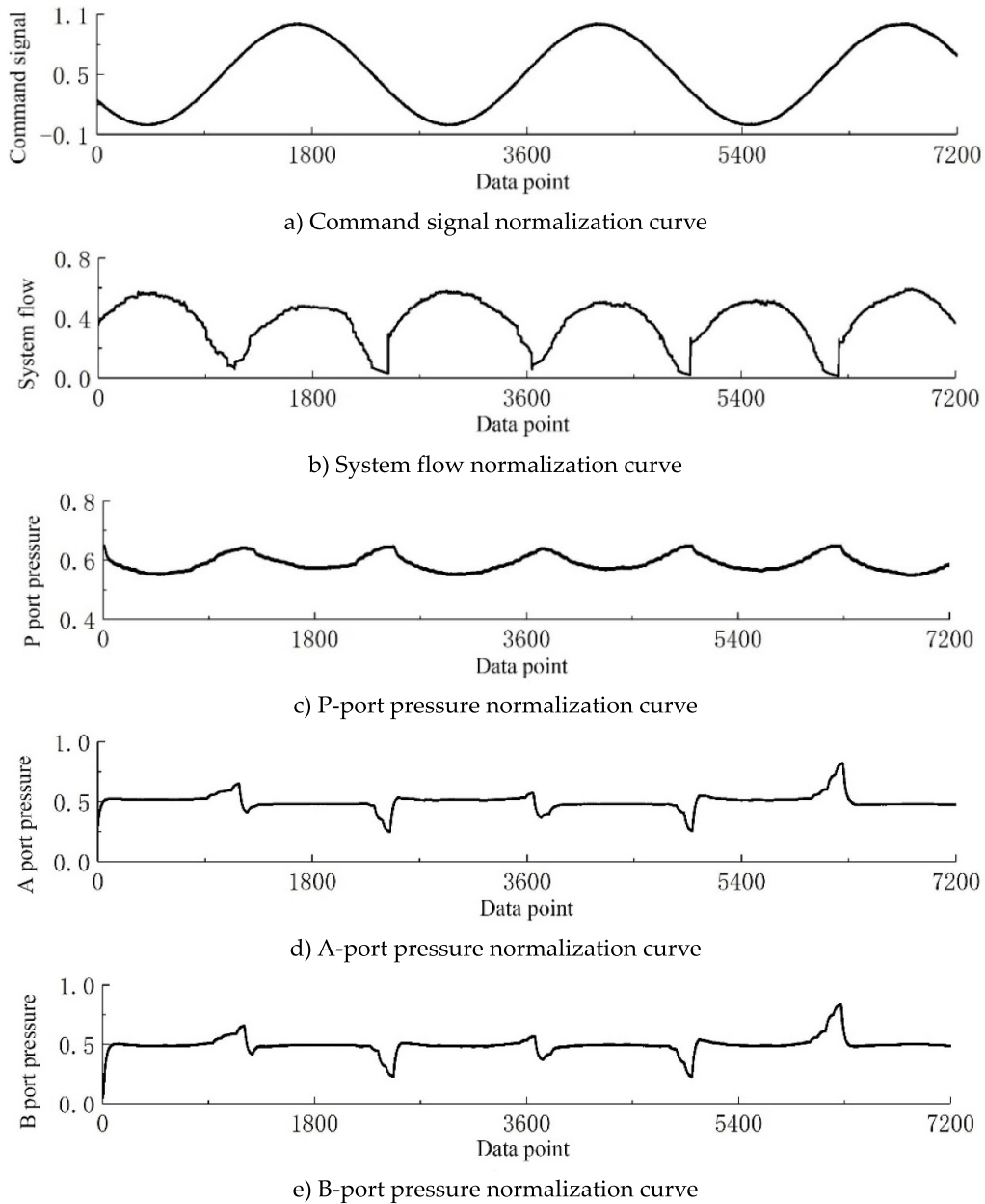
a) Command signal normalization curve

b) System flow normalization curve

c) P-port pressure normalization curve

d) A-port pressure normalization curve

e) B-port pressure normalization curve

Figure 7. Example of a normalized data sample.

choice for multi-dimensional data learning tasks at present. Moreover, graph convolution has extremely wide applicability and is applicable to nodes and graphs with arbitrary topology.

4.1. Model of a troubleshooting algorithm based on stack diagram convolution

4.1.1. Model structure. The basic unit of the model is a graph convolution layer connected with a graph pooling layer. Graph convolution fuses the features of each node through a message propagation mechanism and then graph pooling is used to perform graph feature compression. When the basic unit completes the stack, a feature vector that can describe the whole graph is obtained through global pooling. The network structure is shown in figure 11.

MPNN [19] is a general summary of GNN model. It explains and explains some GNN models from the perspective of airspace. According to the paradigm of MPNN, its mathematical model is as follows:

Message function M is expressed as

$$M_{\text{TFC}}(h_i^{(k)}, h_j^{(k)}) = \sum_{j \in \text{N}(v_i)} \alpha_{ij} W_M^{(k)} h_j^{(k)} \tag{15}$$

where, k indicates the kth message propagation.

Where α_{ij} is expressed as:

$$\alpha_{ij} = \text{softmax}\left(\frac{\left(W_1^{(k)} h_i^{(k)}\right)^T \left(W_2^{(k)} h_j^{(k)}\right)}{\sqrt{d}} \right). \tag{16}$$

a) Command signal data enhancement sample 1

b) Command signal data enhancement sample 2

c) System traffic data enhancement sample 1

d) System traffic data enhancement sample 2

e) P-port pressure data enhancement sample 1

f) P-port pressure data enhancement sample 2

g) A-port pressure data enhancement sample 1

h) A-port pressure data enhancement sample 2

i) B-port pressure data enhancement sample 1

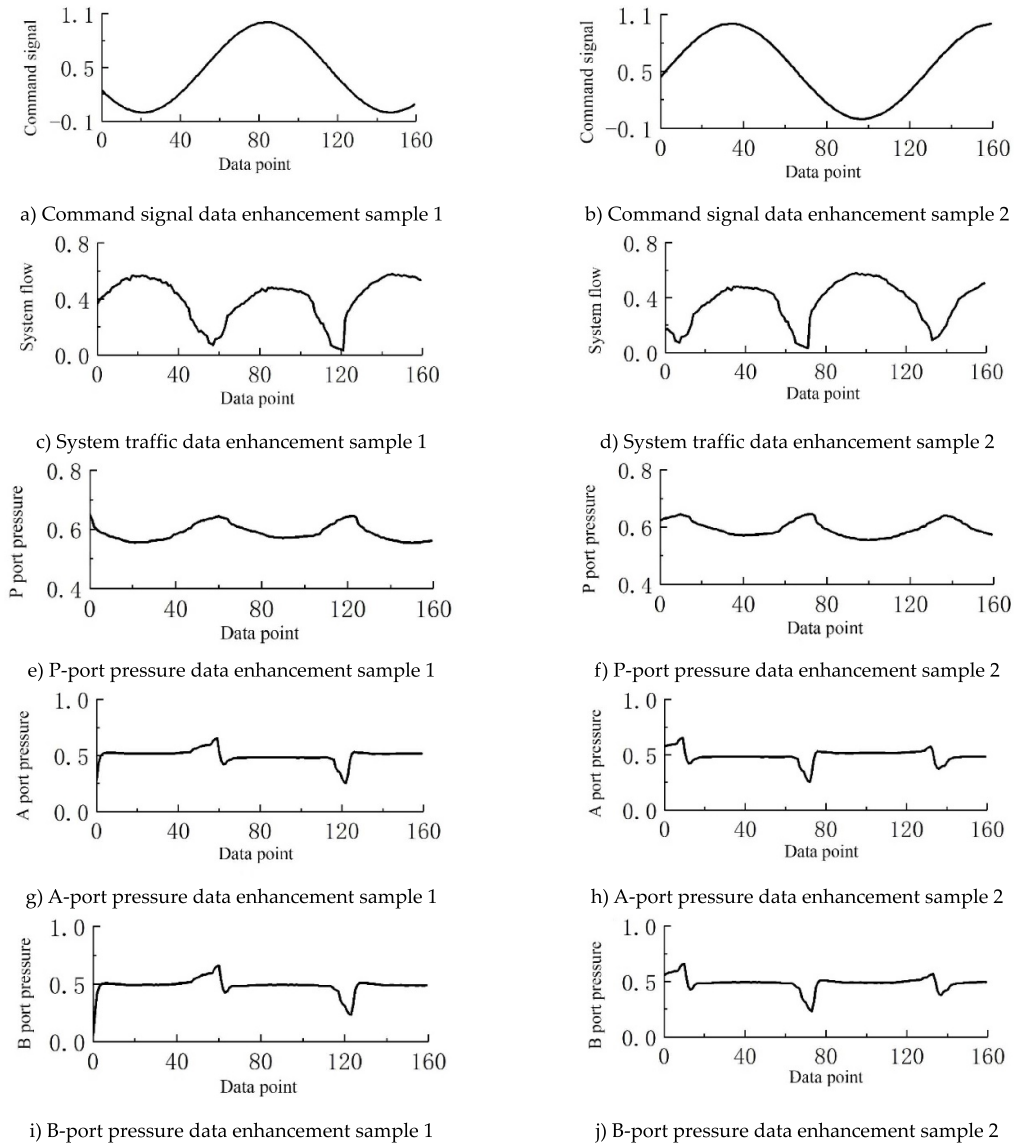j) B-port pressure data enhancement sample 2

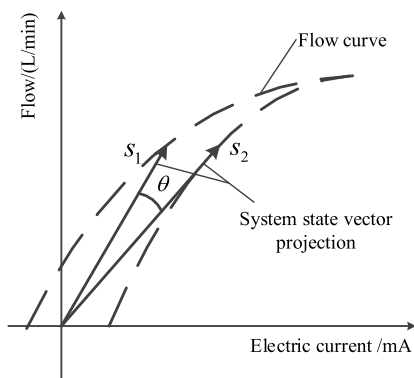**Figure 8.** Example of data enhancement samples.

Figure 9. System state diagram in static test of electro-hydraulic servo valve.

The update function U is expressed as:

$$U_{\text{GCN}}(m_i^{(k+1)}) = W_U^{(k)} h_i^{(k)} + m_i^{(k+1)}. \tag{17}$$
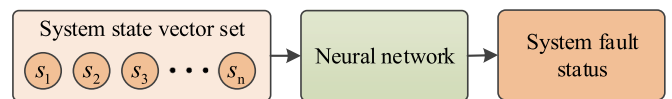

Figure 10. Principle of neural network fault diagnosis algorithm for multisource information system.

The pooling mechanism of TopK [20–22] is a process of continuously discarding nodes according to the characteristic data of different scales on the graph. It places the pooling scope on the full graph node. By setting the pooling rate k, $k \in (0, 1)$, then learning to obtain a value z that can characterize the node importance and sorting it. Finally, according to the node importance, down sample N nodes in the whole graph to KN nodes. The pooling mechanism principle of TopK is shown in formulas (18)–(20),

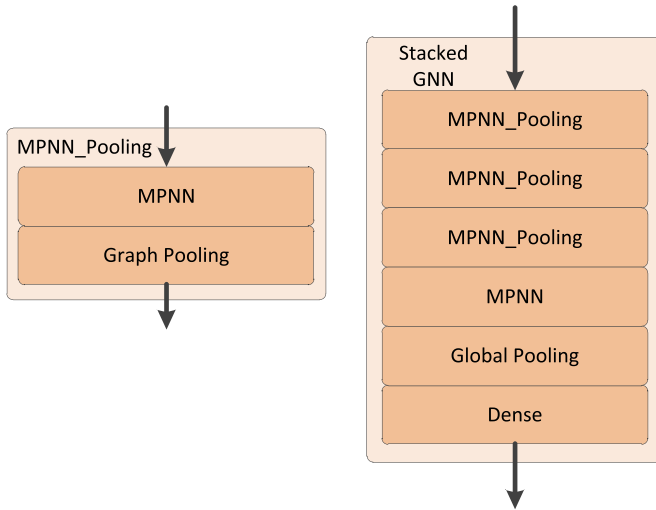$$i = \text{top} - \text{rank}(z, KN) \tag{18}$$

Figure 11. Stack basic unit structure and overall network structure.

$$X' = Xi \tag{19}$$

$$A' = A_{i,j} \tag{20}$$

where, X_i means to slice the characteristic matrix according to vector i, and A_{ij} means to slice the adjacent matrix according to vector i at the same time.

The self-attention pooling model (self-attention graph pooling, SAGPooling) [23], which uses a GNN to learn the importance of nodes. Compared with the design mode of global basis vector, this GNN-based method makes better use of graph structure information to learn the importance of nodes. Formula (21) is changed to obtain formula (22),

$$z = \frac{X_P}{\|P\|} \tag{21}$$

$$z = \text{GNN}(X, A) \tag{22}$$

where, p is the global basis vector.

In global average pooling, by obtaining the channel level average value in the node dimension, the batch graph level output is returned. Therefore, for a single graph G, the output is formula (23):

$$r = \frac{1}{N} \sum_{n=1}^{N} X_n. \tag{23}$$

In order to better integrate data, the convolution layer can adopt the GCN based on multihead attention, the pooling layer can adopt the attention graph pooling based on TopK, the global pooling adopts the global mean pooling, and the classification layer adopts the full connection layer.

Since the output of the network model is not normalized, it is necessary to add normalization to the loss function. First,

Table 4. Training parameters.

Training parameter	Parameter value
Rounds	400
Number of training set samples	300
Number of test set samples	130
Number of samples per batch of training set	20
Optimizer	Adam
Learning rate	0.01
Gradient attenuation coefficient	0.0005

adjust the network output to (0, 1) through the Softmax normalization layer. The Softmax function can be expressed as formula (23):

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}. \tag{24}$$

Secondly, the result of Softmax is logarithmically calculated, and then point multiplied with the one-hot code of the label to obtain the expression of the loss function (24):

$$\text{Loss} = -\sum_i y_i \ln a_i \tag{25}$$

where, y_i represents our real value and α_i represents the value calculated by softmax function.

Because of the filtering characteristics of the GCN itself, when the number of layers of the GCN is too large, the characteristics of all nodes will tend to be unified, and the phenomenon of feature smoothing will appear. When the number of layers of the GCN reaches five layers, its node characteristics will be approximately uniform. Therefore, when stacking GCN, it should be noted that the number of layers of a GCN should not exceed 5.

4.1.2. Parameter selection. The data set is randomly sampled to train the classification of the set on a test set, and because the sample size of the data set is small, it is divided into two data sets, an evaluation set is removed, and the accuracy on the test set is used instead of the evaluation set.

To prevent the uneven distribution of data in the test set from affecting the training effect of the neural network, the above steps can be repeated to separate several times, forming multiple independent training sets and test set samples. The average accuracy of the model on these independent test sets can be considered the true accuracy of the model on this batch of samples. This experiment uses a total of four data sets generated from the sample data for training. The specific training parameters are shown in table 4.

The Adam algorithm was chosen for this experiment. The Adam algorithm is an advanced version of the random gradient method, its learning rate can be adapted to the current gradient during optimization and the diagonal scaling of the gradient is invariable, making it ideal for solving problems with large-scale data or parameters. The algorithm also

Table 5. Structure and parameters of the graph classification model based on stack graph convolution.

Number of layers	Layer	Enter the dimensions	Output size
1	Transformer GCN	5	32
2	SAGPooling	32	32
3	Transformer GCN	32	64
4	SAGPooling	64	64
5	Transformer GCN	64	32
6	SAGPooling	32	32
7	Transformer GCN	32	16
8	Global mean pooling	16	16
9	Linear	16	4

has certain advantages in solving the problem of nonconvex optimization.

The network structure of the graph classification model based on stack chart convolution is shown in table 5. The head parameter in a Transformer GCN [24] model represents the number of cells it can count in parallel, and the model is set to 1. The dropout parameter represents the node drop rate, which improves the generalization of the network to prevent the phenomenon of overfitting during training, and the model is set to 0.2. The pooling rate of the SAGPooling layer is set to 0.5, representing the number of nodes that collapse in half of each pooling.

4.1.3. Training and analysis. The loss and training curves of the four data set batches that the stack classification model trains on are shown in figure 12.

The stack classification model has a double decline phenomenon on the four data set batches. This phenomenon describes the continuous decline and increase in the training loss and test loss in the neural network process with the increase in the number of rounds. It can be avoided by terminating training early.

The model has a slow convergence speed after the double decline phenomenon, and the accuracy fluctuation of the test set is clear. When the model is trending near convergence, the accuracy of the test set fluctuates between 0.7 and 1.0, which is closely related to its special pooling mechanism and network structure.

When using graph pooling, its essence is to discard the nodes on the whole graph based on some rules. At this time, some nodes will be directly discarded, and the discarded nodes will not participate in the error back propagation of the current operation. After performing graph pooling many times, some discarded nodes may be in progress, resulting in the need for the network to optimize the network through the nodes that appear here. This causes the error and accuracy to jitter. However, with optimization, this jitter phenomenon will gradually decrease. Its principle is shown in figure 13.

The stack classification model's training results on the four data sets batches are shown in table 6. It is clear that the

minimum error of the stack model converges above 10^{-2}, and the accuracy of the minimum error is distributed between 0.892 and 0.985, but the highest accuracy of the first three data set batches is 1.000, while the highest accuracy of the fourth data set batch is 0.938, combining the training curves of the four data set batches with the consideration of the training results. The training results of the model did not achieve the desired results.

To address some of the disadvantages of stack classification networks, a residual classification model is trained on the same data set.

4.2. Optimization of a fault diagnosis algorithm model

4.2.1. Model structure. Because the characteristic smoothing phenomenon of a stack graph convolution model has a certain relationship with the number of layers, a multilayer stack graph convolution model is bound to have problems such as model degradation. To prevent model degradation and improve the convergence speed of a model in the optimization process, a residual model can be introduced into the stack model.

A residual model first appeared in residual network [25], and its purpose is to solve the problems of gradient disappearance and gradient explosion of deep networks. In the process of deep network training, the backpropagation of the gradient propagates backward in the form of layer-by-layer multiplication, which will lead to the continuous multiplication of smaller gradients close to zero, thus affecting the gradient update speed of the previous layers of the network.

A schematic diagram of a residual model is shown in figure 14. The layer input and output are directly connected by identity mapping, and a residual model can be directly described by formula (26):

$$g(x) = f(x) + x. \tag{26}$$

A residual model changes the learning goal of the model. When the overall model reaches an optimal state at a certain depth, model degradation will occur if the depth continues to increase. In this case, it will be more troublesome to update the network weight again. After a residual model is introduced, to ensure that the next layer can continue to optimize after the model reaches an optimal state at one layer, it is only necessary to make the output of the next layer approach x, that is, the learning goal $f(x)$ will approach zero.

Introducing a residual model into a graph convolution network can effectively avoid the feature smoothing phenomenon of a graph convolution network during the forward propagation process. The network structure of a residual classification model is shown in figure 15.

Because a residual model itself is an identity map, that is, the dimensions of the input and output are the same, a residual model cannot be directly used after graph pooling. To solve the above problems, a unified number of hidden layer node
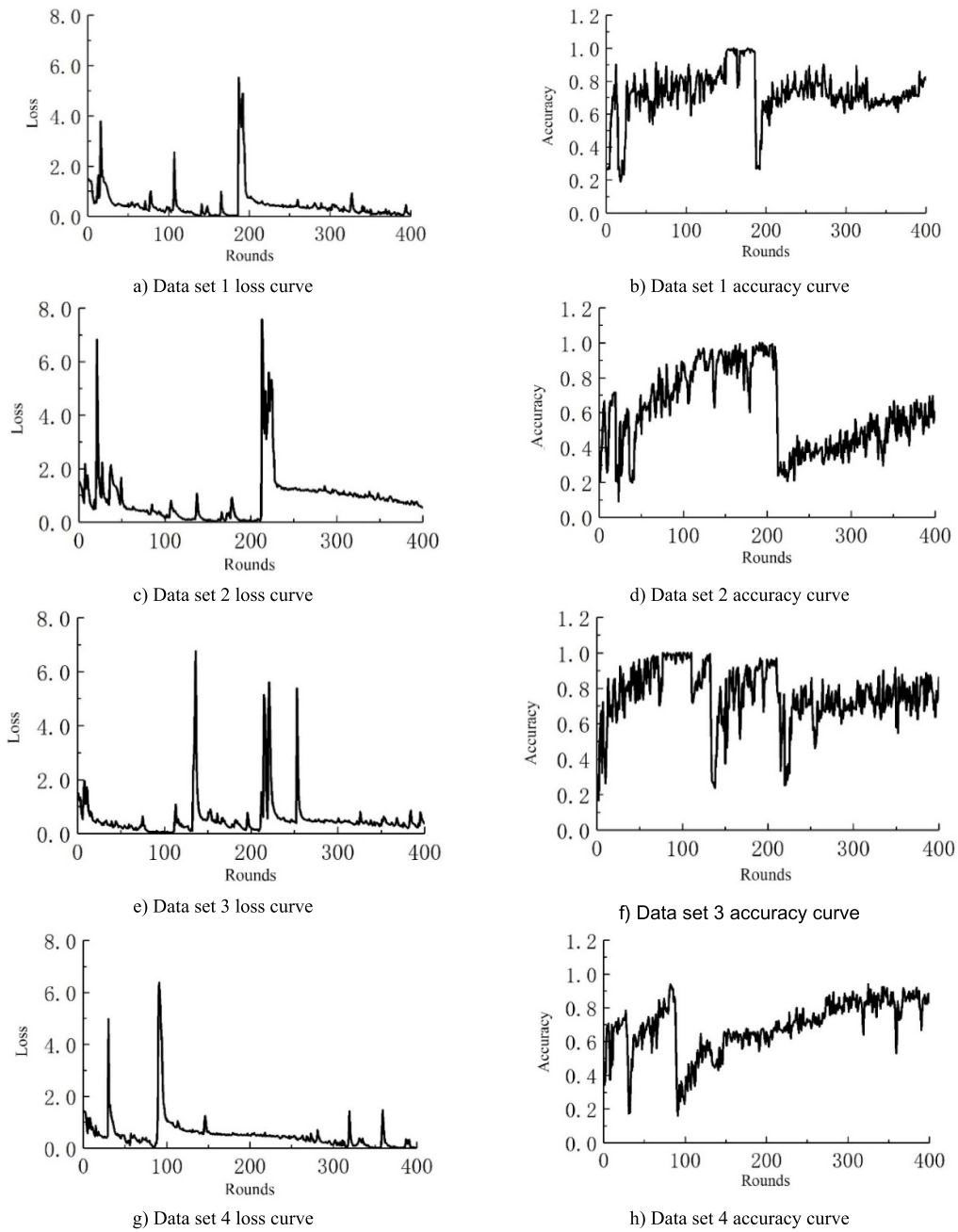
a) Data set 1 loss curve

b) Data set 1 accuracy curve

c) Data set 2 loss curve

d) Data set 2 accuracy curve

e) Data set 3 loss curve

f) Data set 3 accuracy curve

g) Data set 4 loss curve

h) Data set 4 accuracy curve

Figure 12. Loss and accuracy curves of the stack classification model on the four data set batches.
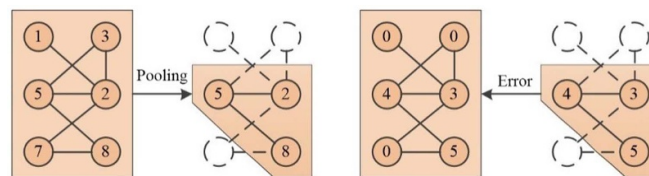
Figure 13. Forward propagation and error inverse transmission during the pooling process.

features can be determined according to the number of node features and classification, and then a readout layer is added after graph pooling; that is, global pooling is used to realize identity mapping.

4.2.2. Parameter selection. The network structure of the graph classification model based on residual graph convolution is shown in table 7. The model is connected by a residual jump, and the feature number of hidden layer nodes

Table 6. Stack classification model training results.

	Data set 1	Data set 2	Data set 3	Data set 4
Minimum loss round	180	203	107	355
Minimum loss	2.501×10^{-2}	3.741×10^{-2}	3.258×10^{-2}	1.227×10^{-2}
Minimum loss round accuracy	0.973	0.985	0.977	0.892
Maximum accuracy	1.000	1.000	1.000	0.938

$$g(x) = f(x) + x$$

Figure 14. Residual model principle.

Figure 15. Residual classification model.

Table 7. Structure and parameters of the graph classification model based on residual graph convolution.

Number of layers	Layer	Enter size	Output size
1	Transformer GCN	6	64
2	SAGPooling	64	64
3	Global mean pooling	64	64
4	Global max pooling	64	64
5	Transformer GCN	64	64
6	SAGPooling	64	64
7	Global mean pooling	64	64
8	Global max pooling	64	64
9	Transformer GCN	64	64
10	SAGPooling	64	64
11	Global mean pooling	64	64
12	Global max pooling	64	64
13	Dense	128	64
14	Dense	64	32
15	Dense	32	4

must be equal. Here, the feature number of hidden layer nodes is 64. The full connection layer has three layers. The first two layers reduce the dimension of features, and the output size of the last layer is equal to the number of classifications.

4.2.3. Training and analysis. The loss and training curves obtained using training the residual classification model on the four data set batches are shown in figure 16.

The residual classification model also exhibited double drops on the four data set batches. Compared with the stack classification model, the residual classification model has much smaller training errors each time a double drop occurs, and its convergence speed is also faster.

As shown in figure 16, the residual classification model overcomes the disadvantage of the large fluctuations in the accuracy of the stack classification model during training.

When it approaches a convergence state, the accuracy of the test set is stable above 0.9 and maintained at 1.0 for a long time.

The training results of the residual classification model on the four data set batches are shown in table 8. It is clear that the minimum error of the residual model converges to approximately 10^{-3}, the accuracy at the minimum error is 1.000, and the highest accuracy of the four data set batches is 1.000. The training results of the model achieves the expected results.

a) Data set 1 loss curve

b) Data set 1 accuracy curve

c) Data set 2 loss curve

d) Data set 2 accuracy curve

e) Data set 3 loss curve

f) Data set 3 accuracy curve

g) Data set 4 loss curve

h) Data set 4 accuracy curve

Figure 16. Loss and accuracy curves of the residual classification model on the four data set batches.

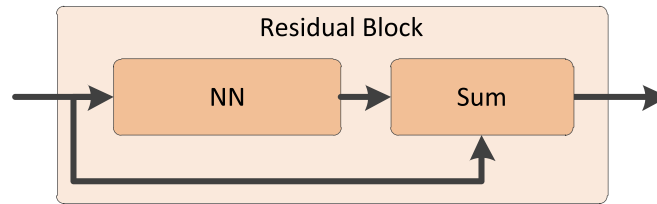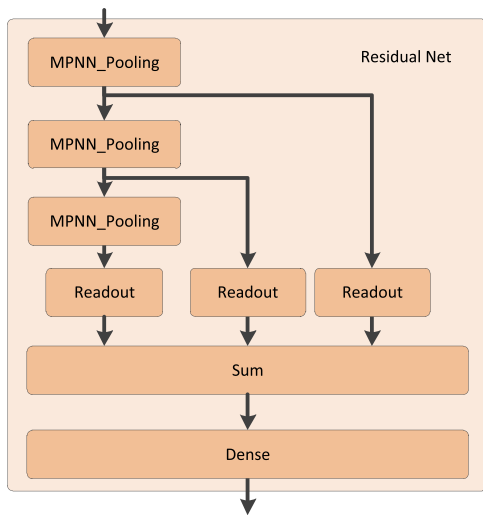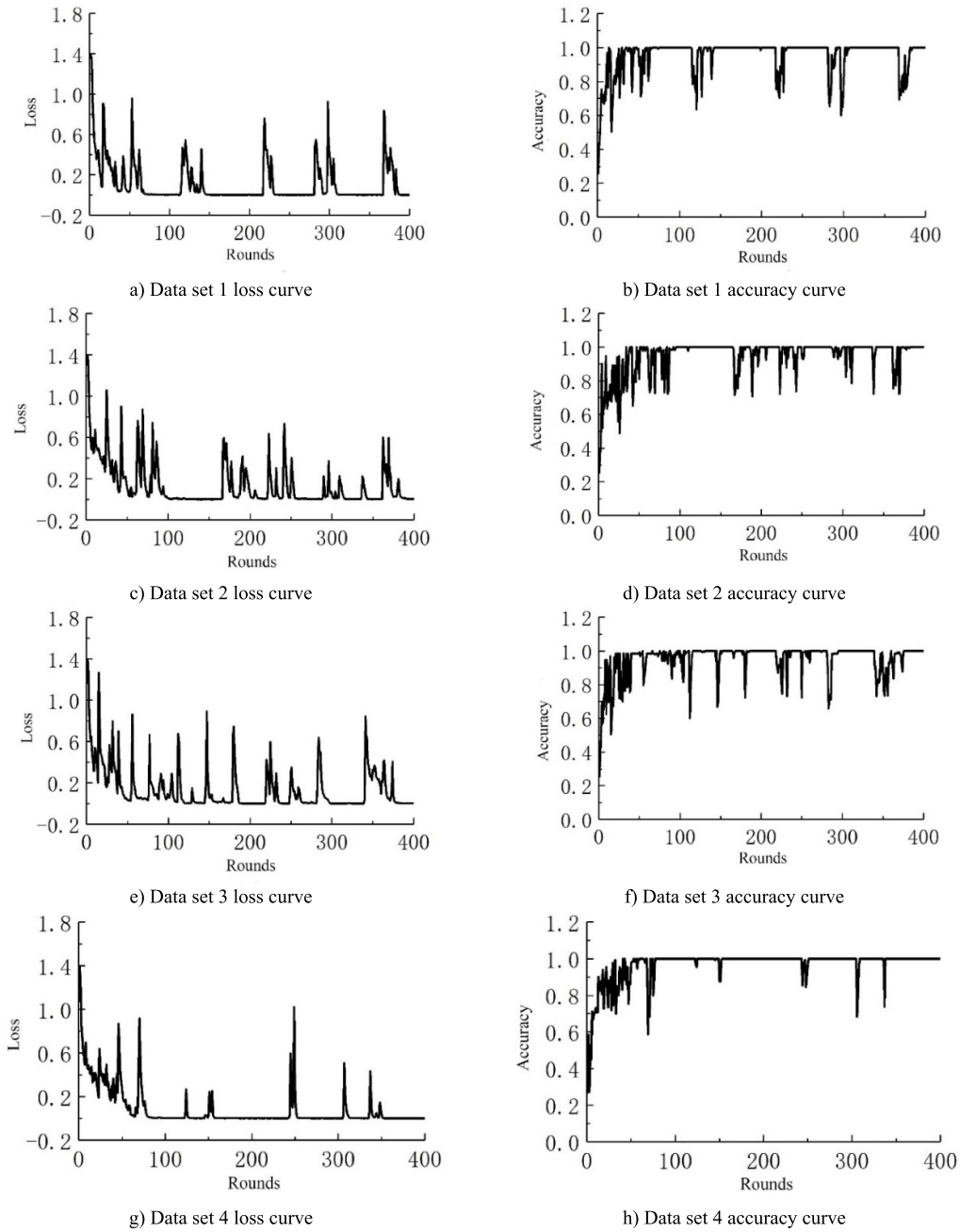**Table 8.** Residual classification model training results.

	Data set 1	Data set 2	Data set 3	Data set 4
Minimum loss round	204	155	256	234
Minimum loss	0.874×10^{-3}	2.099×10^{-3}	0.946×10^{-3}	1.241×10^{-3}
Minimum loss round accuracy	1.000	1.000	1.000	1.000
Maximum accuracy	1.000	1.000	1.000	1.000

5. Summary

First, a mathematical model of a force feedback two-stage electrohydraulic servo valve is established, and the fault characteristics of the electrohydraulic servo valve under typical fault modes are explored. Then, the fault data of the electrohydraulic servo valve are obtained and normalized. Finally, a fault diagnosis algorithm based on a MPNN is proposed.

According to the experimental results, the stack classification model has a slower convergence due to its

special pooling mechanism and network structure, and the accuracy fluctuates more severely during the convergence process. To address this, a residual model is added to the stack classification model, which overcomes the loss of node information caused by the stack classification model due to the pooling mechanism. After many experiments, the residual classification model converges faster than the stack classification model. The characteristics of a small convergence error and high accuracy after convergence prove that a graph neural network has certain advantages in the application to fault diagnosis of electrohydraulic servo valves. This can help realize the healthy operation and maintenance of an electrohydraulic servo system.

However, there are some shortcomings in this article. Because the common faults of the test object are relatively single, all failure modes cannot be reproduced. The experimental object is single, and the data set cannot contain fault samples of multiple models of electrohydraulic servo valves. Therefore, in follow-up research, the focus should also be placed on the construction of the data set. A data-driven fault diagnosis algorithm has higher requirements for the data, so it is necessary to construct complete electrohydraulic servo valve fault data set. Due to the limitation of the data, this paper did not test data of different lengths. Subsequent researchers can study the application of a graph neural network model based on data of different lengths to the fault diagnosis of electrohydraulic servo valves.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID iD

Sun Pengfei ⓘ https://orcid.org/0000-0003-4147-3341

References

[1] Bing X, Shen J, Liu S, Su Q and Zhang J 2020 Research and development of electro-hydraulic control valves oriented to industry 4.0: a review *Chin. J. Mech. Eng.* **33** 13–32

[2] Shi J, Yi J, Ren Y, Li Y, Zhong Q, Tang H and Chen L 2021 Fault diagnosis in a hydraulic directional valve using a two-stage multi-sensor information fusion *Measurement* **179** 109460

[3] Shi C, Tang H and Mupfukirei L R 2021 A fault diagnosis method for an electro-hydraulic directional valve based on intrinsic mode functions and weighted densely connected convolutional networks *Meas. Sci. Technol.* **32** 08401

[4] Tang S, Zhu Y and Yuan S 2022 Intelligent fault diagnosis of hydraulic piston pump based on deep learning and Bayesian optimization *ISA Trans.* **129** 555–63

[5] Guo F Y, Zhang Y C, Wang Y, Ren PJ and Wang P 2021 Fault diagnosis of reciprocating compressor valve based on transfer learning convolutional neural network *Math. Probl. Eng.* **2021** 8891424

[6] Huang C M, Li J, Wang X, Liao J, Yu H, Chen C-C and Wang K-C 2021 Monitoring of valve gap in diesel engine based on vibration response feature extraction *Sens. Mater.* **33** 2365–83

[7] Andrade A, Lopes K, Lima B and Maitelli A 2021 Development of a methodology using artificial neural network in the detection and diagnosis of faults for pneumatic control valves *Sensors* **21** 853

[8] Ykla B, Wen Z B, Aa B, Zhou XQ, Peng MJ and Chao N 2020 A multi-layer approach to DN 50 electric valve fault diagnosis using shallow-deep intelligent models *Nucl. Eng. Technol.* **53** 148–63

[9] Zhu Y, Li G, Tang S, Wang R, Su H and Wang C 2022 Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN *Appl. Acoust.* **192** 108718

[10] Liu C, Wang Y, Pan T and Zheng G 2020 Fault diagnosis of electro-hydraulic servo valve using extreme learning machine *Int. Trans. Electr. Energy Syst.* **30** e12419

[11] Chunyu J, Kaixuan K, Wei G, Dong YA, Li-juan CH and Chao AI 2020 Fault prediction of electro—hydraulic servo valve based on CNN + LSTM neural network *Hydraul. Pneum.* **12** 173–81

[12] Wenqing W 2019 Research on fault prognostic of electro-hydraulic servo valve based on CNN and GRU (Shanghai: Shanghai Jiaotong University) pp 38–61

[13] Chai J and Jin X 2017 Research on fault diagnosis of servo valve based on deep learning *2017 Prognostics and System Health Management Conf. (Phm-harbin)* (IEEE) pp 1–7

[14] Chen G, Yang D, Gao W, Chen L, Hua Z and Ai C 2019 Research on fault method of electro-hydraulic servo valve based on deep neural network *2019 IEEE 8th Int. Conf. on Fluid Power and Mechatronics (FPM)* (IEEE) pp 393–7

[15] Ji X, Ren Y, Tang H, Shi C and Xiang J 2020 An intelligent fault diagnosis approach based on Dempster-Shafer theory for hydraulic valves *Measurement* **165** 108129

[16] Chao Q, Gao H H, Tao* J F, Wang Y, Zhou J and Liu C 2022 Adaptive decision-level fusion strategy for the fault diagnosis of axial piston pumps using multiple channels of vibration signals *Sci. China Technol. Sci.* **65** 470–80

[17] Chao Q, Wei X, Lei J, Tao J and Liu C 2022 Improving accuracy of cavitation severity recognition in axial piston pumps by denoising time–frequency images *Meas. Sci. Technol.* **33** 055116

[18] Shengnan T, Yong* Z and Shouqi Y 2022 Intelligent fault identification of hydraulic pump using deep adaptive normalized CNN and synchrosqueezed wavelet transform *Reliab. Eng. Syst. Saf.* **224** 108560

[19] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl GE 2017 Neural message passing for quantum chemistry *Int. Conf. on Machine Learning* PMLR pp 1263–72

[20] Cangea C, Veličković P, Jovanović N, Kipf T and Liò P 2018 Towards sparse hierarchical graph classifiers (arXiv:1811.01287)

[21] Knyazev B, Taylor G W and Amer M 2019 Understanding attention and generalization in graph neural networks *Advances in Neural Information Processing Systems* vol 32 (arXiv:1905.02850)

[22] Lee J, Lee I and Kang J 2019 Self-attention graph pooling *Int. Conf. on Machine Learning* PMLR pp 3734–43

[23] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 770–8

[24] Shi Y, Huang Z, Feng S, Zhong H, Wang W and Sun Y 2020 Masked label prediction: unified massage passing model for semi-supervised classification (arXiv:2009.03509)

[25] Hinton G, Vinyals O and Dean J 2015 Distilling the knowledge in a neural network (arXiv:1503.02531)