actuators

MDPI

Article

A Class-Specific Prototype and Multivariate Coupling-Aware Method for EHA Fault Time-Series Diagnosis

Guozhu Zhi, Kelin Zhong, Zhen Jia *, Zhihao Gao, Weijun Yan and Zhenbao Liu

School of Civil Aviation, Northwestern Polytechnical University, Xi'an 710072, China; 13519122683@mail.nwpu.edu.cn (G.Z.); bbwwkk@mail.nwpu.edu.cn (K.Z.); gaozhihao750@mail.nwpu.edu.cn (Z.G.); yanweijun@mail.nwpu.edu.cn (W.Y.); liuzhenbao@nwpu.edu.cn (Z.L.)
* Correspondence: jiazhen@nwpu.edu.cn

Abstract

In the multivariate time-series fault diagnosis task for aviation electro-hydrostatic actuators (EHA), the overall signal morphologies of different fault categories are relatively similar, while the key discriminative differences are hidden in local segments and variations in variable coupling. Therefore, existing Transformer-based methods usually have difficulty characterizing local specificity. To address this issue, this paper proposes a Local Prototype-Global Generic Dual-branch Transformer (LPG-Former). First, to obtain local information capable of characterizing class differences, a class-specific discriminative prototype (CDP) is constructed. The CDP selects discriminative time points from the time-series samples of each class to capture key local morphological variations, and constructs local prototypes carrying class-related local differential features. To further improve the ability of the CDP to capture multivariate fault coupling relationships, a multivariate coupling-aware prototype matching strategy (MCPM) is designed. The MCPM extends univariate prototypes into multivariate local prototype blocks and jointly measures local dissimilarity, variable correlation, and trend consistency, thereby enabling prototype learning with awareness of multivariate coupling relationships. Finally, to fuse local discriminative information and global temporal information, a dual-branch Transformer is constructed. LPG-Former encodes the differential features between the CDP and the best-fit subsequence (BFS) of the input sample through a Local Prototype Transformer, and complements global generic information through a Global Generic Transformer, thereby achieving collaborative dual-branch representation. Experimental results on an eight-class EHA operating-state dataset show that LPG-Former achieves an accuracy of 98.74% and an F1-score of 98.78%, significantly outperforming classical methods such as InceptionTime and TapNet.

Keywords: electro-hydrostatic actuator; fault diagnosis; multivariate time-series classification; prototype learning

Check for updates

Academic Editors: Roberto Strada and Paolo Righettini

Received: 3 June 2026
Revised: 2 July 2026
Accepted: 11 July 2026
Published: 13 July 2026

1. Introduction

With the in-depth development of more-electric aircraft technology, electro-hydrostatic actuator systems (EHAs) have been widely used to drive key flight-control surfaces such as ailerons, elevators, and rudders [1,2]. With prominent advantages including high power density, convenient installation and maintenance, and no requirement for a centralized hydraulic source, this system has established an irreplaceable core position in high-reliability engineering fields such as modern aerospace [3]. However, the highly integrated structural characteristics of EHA systems also lead to complex electromechanical–hydraulic multidomain strong coupling inside the system [4]. During actual service, EHA systems are exposed for long periods to severe environments such as high mechanical loads, intense vibration, and rapid switching of operating conditions. The long-term combined effects of coupled internal physical domains and harsh external alternating conditions can easily induce progressive performance degradation and sudden faults in key internal components [5]. As the terminal actuator in the flight-control chain, functional failure of an EHA system may cause loss of flight attitude control, resulting in severe casualties and major property losses [6]. Therefore, in-depth research on fault diagnosis and health-monitoring methods for electro-hydrostatic actuator systems is crucial for ensuring the reliable operation of aerospace vehicles [7].

Current mainstream research on EHA fault diagnosis mainly follows two technical routes: model-driven and data-driven approaches [8,9]. Traditional model-driven methods focus on constructing refined analytical mathematical models or state observers based on internal physical mechanisms to map and solve fault features. Xu et al. [10] proposed a model-driven fault-detection method for electro-hydraulic variable-displacement pumps. By introducing an unknown-input Kalman filter to estimate swashplate moment uncertainty and combining it with cumulative-sum residual evaluation, rapid detection of typical faults was achieved. Saeedzadeh et al. [11] constructed a real-time model-driven fault-detection and diagnosis strategy for electro-hydraulic actuators, achieving online identification of typical faults such as leakage and friction through the coordination of an updated interactive multiple-model structure and a smooth variable-structure filter. Nguyen et al. [12] designed a high-accuracy motion-control method based on an extended sliding-mode observer, improving control accuracy and system stability under complex nonlinear operating conditions by robustly estimating unknown states and disturbances in uncertain electro-hydraulic systems. However, such model-driven strategies rely heavily on accurate prior physical parameters and complete dynamic assumptions, and it is difficult to guarantee absolute physical modeling accuracy when facing complex strongly nonlinear dynamic disturbances and time-varying loads in actual operation [13].

With the rapid rise of intelligent computing technologies, data-driven methods such as neural networks and support vector machines have freed fault diagnosis from explicit physical differential equations and have demonstrated significant advantages in adaptive feature extraction and pattern recognition [14,15]. Yan et al. [16] designed a multisensor fusion diagnostic method based on a hypergraph structure, which significantly improved the stability of fault identification under complex operating conditions by characterizing high-order correlations among multisource monitoring signals. Shen et al. [17] proposed a multisource-domain fault-diagnosis method with an adaptive dual-domain obfuscation weighting strategy, effectively mitigating the influence of data-distribution differences across source domains on model generalization. Wang et al. [18] constructed an oversampling method based on Gaussian mixture models and boundary optimization, effectively improving the shortage of bearing-fault samples and the classification-boundary shift under class-imbalance conditions. Although the above data-driven approaches have made significant progress, diagnostic frameworks using a single learner are limited by fixed network structures and insufficient representation capacity when dealing with the variable alternating operating conditions of EHA systems and strongly coupled multisource heterogeneous monitoring signals, making them prone to local optima and overfitting [19,20]. These inherent limitations lead to poor generalization, limited recognition accuracy for early weak compound faults, and unstable classification boundaries, making existing intelligent decision models based on a single learner unable to fully satisfy the stringent requirements for diagnostic credibility and reliability in safety-critical fields such as aerospace [21,22].

To overcome the limitations of a single learner in handling complex diagnostic tasks, ensemble learning provides an effective improvement path for enhancing the accuracy and robustness of EHA fault diagnosis by collaboratively fusing multiple differentiated base models [23]. Hu et al. [24] proposed a dynamic domain-adaptive ensemble diagnostic framework, which significantly enhanced the fault-identification capability of mechanical equipment under complex operating conditions by fusing multiple differentiated diagnostic models and dynamically adjusting the domain-adaptation process. Liu et al. [25] proposed an evidential ensemble preference-guided real-time multimodal fault-diagnosis method, effectively improving the online diagnostic capability of complex dynamic systems under multiple operating modes by fusing fault-information preferences under different operating modes. Tong et al. [26] constructed an ensemble-learning-based multisensor information fusion diagnostic method, significantly improving the accuracy and stability of rolling-bearing fault identification by collaboratively fusing complementary fault information from different sensor signals. However, highly complex ensemble architectures still exhibit serious black-box characteristics in essence and cannot reveal a transparent logical mapping between input features and fault states [27]. This lack of interpretability makes it difficult for diagnostic systems to pass stringent aviation airworthiness safety evaluations and prevents them from providing maintenance personnel with confidence evidence that carries physical meaning [28].

To address the trust challenges caused by the black-box nature of ensemble models, explainable artificial intelligence provides intuitive and transparent decision logic, bringing important physical insights to intelligent health monitoring of aircraft actuators [29,30]. By quantifying the contribution weights of multisource monitoring features to diagnostic conclusions, this technology can assist operation and maintenance personnel in accurately identifying key physical representations that induce system-state deviations, thereby providing more persuasive decision support for EHA system maintenance [31,32].

This paper proposes an ensemble-learning EHA fault-diagnosis framework that considers both high-accuracy recognition performance and decision transparency. By deeply integrating interpretability strategies such as feature-importance attribution analysis and local interpretable models, the proposed framework mitigates the black-box barrier formed by the stacking of multiple base models at the algorithmic level, and verifies its effectiveness on a dataset covering typical faults such as leakage, friction, gas lock, and sensor drift. The main contributions of this paper are as follows:

(1). To address the problem that the overall inter-class morphology is similar and key differences are localized in EHA multivariate fault time series, this paper proposes a class-specific discriminative prototype construction method. By extracting discriminative time points, generating candidate local segments, and performing information-gain screening, this method constructs representative and discriminative local prototypes from different fault classes, enabling the model to characterize fine-grained class differences at the level of local fault patterns.

(2). To address the difficulty of univariate prototypes in representing the coupled fault-propagation characteristics of multiple EHA components, this paper designs a multivariate coupling-aware prototype matching strategy. This strategy extends class-specific local prototypes into multivariate prototype blocks and measures the matching relationship between prototypes and input samples from three aspects: local morphological differences, variable-correlation differences, and trend-consistency differences, thereby enhancing the model's ability to perceive coupling variations among multisource monitoring variables.

(3). To address the tendency of traditional Transformers to emphasize global generic features while ignoring class-specific local differences, this paper constructs a Local Prototype-Global Generic Dual-branch Transformer diagnostic framework. The framework uses a local prototype branch to encode CDP–BFS differential features and a global generic branch to supplement the overall dynamic information of the original time series, thereby achieving collaborative modeling of local discriminative information and global temporal information. Experimental results on the eight-class EHA operating-state dataset show that the proposed method outperforms multiple typical time-series classification methods in classification accuracy, missed-detection control, and feature separability.

The remainder of this paper is organized as follows. Section 2 introduces the operating principle of the electro-hydrostatic actuator system and provides a detailed analysis of typical fault modes. Section 3 details the architecture and implementation procedure of the proposed interpretable ensemble diagnostic method. Section 4 describes the setup of the flight-control semi-physical experimental platform and the overview of the dataset. Section 5 conducts an in-depth discussion of the diagnostic results and comparative interpretability analysis. Section 6 presents the conclusions of the study and discusses future research directions.

2. Related Theoretical and Engineering Basis

2.1. EHA System and Typical Fault Modes

An electro-hydrostatic actuator is a self-contained actuation system that integrates an electric drive, hydraulic power transmission, actuator motion output, and feedback control. In a typical aircraft aileron EHA, the servo motor drives the hydraulic pump to generate flow and pressure, and the actuator cylinder converts hydraulic energy into linear displacement to drive the control surface through the transmission mechanism. During this process, the motor-drive response, hydraulic pressure and flow response, actuator displacement response, and feedback signals are strongly coupled.

Typical EHA faults can be generally divided into hydraulic-transmission faults, mechanical-actuator faults, and feedback-related faults. Hydraulic leakage may reduce pressure-holding capability and change the flow-compensation demand. Cylinder deformation or aging may change friction, sealing, or motion constraints, thereby affecting displacement tracking and pressure fluctuation. Feedback-related stuck, gain, or bias faults may distort the measured state and further disturb the closed-loop control response. Therefore, EHA fault diagnosis should consider not only individual signal changes but also coupling variations among electrical, hydraulic, and mechanical responses.

2.2. Multivariate Time-Series Classification

Multivariate time-series classification (MTSC) [33] is an important task in time-series analysis. Its objective is to determine the class to which a sample belongs according to the dynamic variation patterns of multiple variables along the continuous time dimension. Compared with ordinary classification tasks, MTSC has higher modeling complexity [34]. This is because multivariate time-series samples simultaneously contain temporal dependencies and inter-variable coupling relationships. The dynamic trend of the same variable and the synchronous or lagged relationships among different variables may both affect the final classification result [35].

In this study, the terms "local" and "global" are defined according to the temporal scale of EHA multivariate time-series samples. "Local" refers to short subsequences or time windows within one sample, such as the segments bounded by discriminative points or the search windows used for CDP–BFS matching. These local segments mainly describe transient fault signatures, short-term fluctuations, and local variations in multivariate coupling. By contrast, "global" refers to the complete temporal evolution of one sample over the whole maneuver duration, which reflects the overall dynamic trend and long-range temporal dependency of the EHA system. Therefore, local features emphasize fine-grained fault-sensitive segments, whereas global features describe the overall response pattern of the complete EHA time-series sample.

In recent years, typical MTSC methods have mainly included deep-learning models such as 1D-CNN-LSTM [36], InceptionTime [37], TapNet [38], and Transformer. Among them, 1D-CNN-LSTM focuses on local feature extraction and temporal-dependency modeling; InceptionTime enhances local-pattern representation through multiscale convolution; TapNet improves class discriminability using attentional prototype learning; and Transformer relies on the self-attention mechanism to capture long-range dependencies and global contextual information. However, existing Transformer-based methods usually pay more attention to generic temporal features and insufficiently characterize local discriminative segments, making their recognition performance vulnerable in class-imbalance scenarios or when inter-class differences are subtle. Meanwhile, existing methods usually lack a dual-branch structure for collaborative modeling of local discriminative features and global generic features, making it difficult to fully exploit their complementary advantages.

2.3. Transformer-Based Time-Series Classification Methods

In recent years, Transformer has been widely applied to time-series classification tasks by virtue of its self-attention mechanism. Around sequential modeling, local morphological expression, and multiscale feature extraction for time series, Transformer-based studies have made improvements in multiple aspects. For example, ConvTran enhances the ability of Transformer to perceive temporal position information by introducing temporal absolute positional encoding and efficient relative positional encoding [39]. SVP-T starts from shape-level subsequences and uses variable-position self-attention to model associations among different local morphologies [40]. Mgformer extracts feature representations at different temporal granularities and inter-variable correlations through multiple Transformer groups and channel-attention masks [41].

Although the above methods improve the feature-modeling capability of Transformer in MTSC tasks, they mainly rely on generic temporal features and are relatively insufficient in characterizing class-specific local discriminative segments. Consequently, classification performance is easily affected in scenarios with class imbalance or similar overall inter-class patterns.

2.4. Time-Series Prototype Extraction Methods

Time-series prototypes aim to extract typical patterns capable of representing class characteristics from original sequences or latent-space representations for metric matching and pattern recognition. Unlike whole-sequence features, prototypes usually correspond to class-representative samples, local discriminative subsequences, or learnable latent-space centers, and can highlight key differences among different classes. In recent years, prototype learning has been used for time-series classification and interpretable modeling. By learning class-related prototypes, it enhances the ability of models to represent key temporal patterns [42]. In MTSC tasks, time-series prototypes can be regarded as typical local prototypes. By extracting subsequences that are highly related to a certain class and clearly different from other classes, they provide a more interpretable local basis for class determination [43].

Existing studies mainly improve time-series classification performance from three aspects: prototype selection, prototype optimization, and prototype matching. Overall, time-series prototype extraction can transform complex temporal data into a matching problem among class-representative patterns, providing an effective approach for improving classification accuracy, robustness, and interpretability.

3. Proposed Method

Before introducing the proposed method, the concept of a "sample" is first clarified. In this study, one sample refers to a complete multivariate time-series segment collected under one EHA operating state during one maneuver, rather than a single scalar value at a certain time instant.

Each EHA multivariate time-series sample contains monitoring variables associated with the motor-drive, hydraulic-transmission, and actuator-motion parts of the system. These signals are used to characterize the electrical response, hydraulic response, and mechanical motion response of the EHA during fault propagation.

For the fault-diagnosis task of aviation electro-hydrostatic actuators (EHAs), multivariate monitoring signals under different operating states often show high overall similarity, while fault differences are mainly reflected in subtle changes in local amplitude, local trends, and multivariate coupling relationships. Meanwhile, because EHA fault samples are costly to obtain, practical diagnosis often faces problems such as small samples and class imbalance. However, traditional Transformers focus on global generic feature modeling and tend to ignore local discriminative segments closely related to fault categories, resulting in limited discriminability and robustness in fine-grained fault-identification scenarios.

To address the above problems, this paper proposes a Local Prototype-Global Generic Dual-branch Transformer (LPG-Former). First, the method constructs a class-specific discriminative prototype (CDP) to extract local fault prototypes with class representativeness and discriminability from different operating states. Subsequently, a multivariate coupling-aware prototype matching strategy (MCPM) is designed. During CDP matching, local dissimilarity, correlation, and trend consistency with the input sample across multiple variables are simultaneously measured, thereby enhancing the ability of the CDP to perceive multivariate coupling relationships. Finally, a Local Prototype Transformer (LPT) and a Global Generic Transformer (GGT) are constructed to encode, respectively, the local discriminative information reflected by the CDP and the global generic information in the original time series. Ultimately, the feature representations output by the LPT and the GGT are concatenated and fused, then fed into a classification layer to generate the final fault-diagnosis result, while the loss function is updated to optimize the feature mapping, Transformer encoders, and classification-layer parameters.

3.1. Class-Specific Discriminative Prototype (CDP)

To enhance the ability to characterize locally discriminative fault time series, this paper proposes a class-specific discriminative prototype construction method. The CDP is a local time-series prototype that satisfies two requirements: on the one hand, the prototype has high similarity to local segments in samples from the same operating state; on the other hand, it has large differences from local segments in samples from other operating states. Therefore, the CDP can serve as a local basis for discriminating EHA fault classes and provide input for subsequent Transformer modeling. In this study, a time-series prototype refers to a representative local subsequence that repeatedly appears in samples of a certain operating state and shows clear differences from samples of other operating states. Taking an actuator-displacement signal as an example, if a specific fault state repeatedly produces a short local deviation or tracking-lag pattern during the maneuver, the time points with the largest reconstruction errors are first selected as discriminative points. The signal segment bounded by adjacent discriminative points is then treated as a candidate prototype. This candidate segment is compared with the best-fit subsequences from samples of the same and different operating states. If it is consistently close to same-class samples but clearly different from other-class samples, it obtains a high information gain and is retained as a final CDP. In this way, the prototype-selection process transforms local signal differences into class-discriminative diagnostic evidence. The construction principle of the proposed CDP is shown in Figure 1.

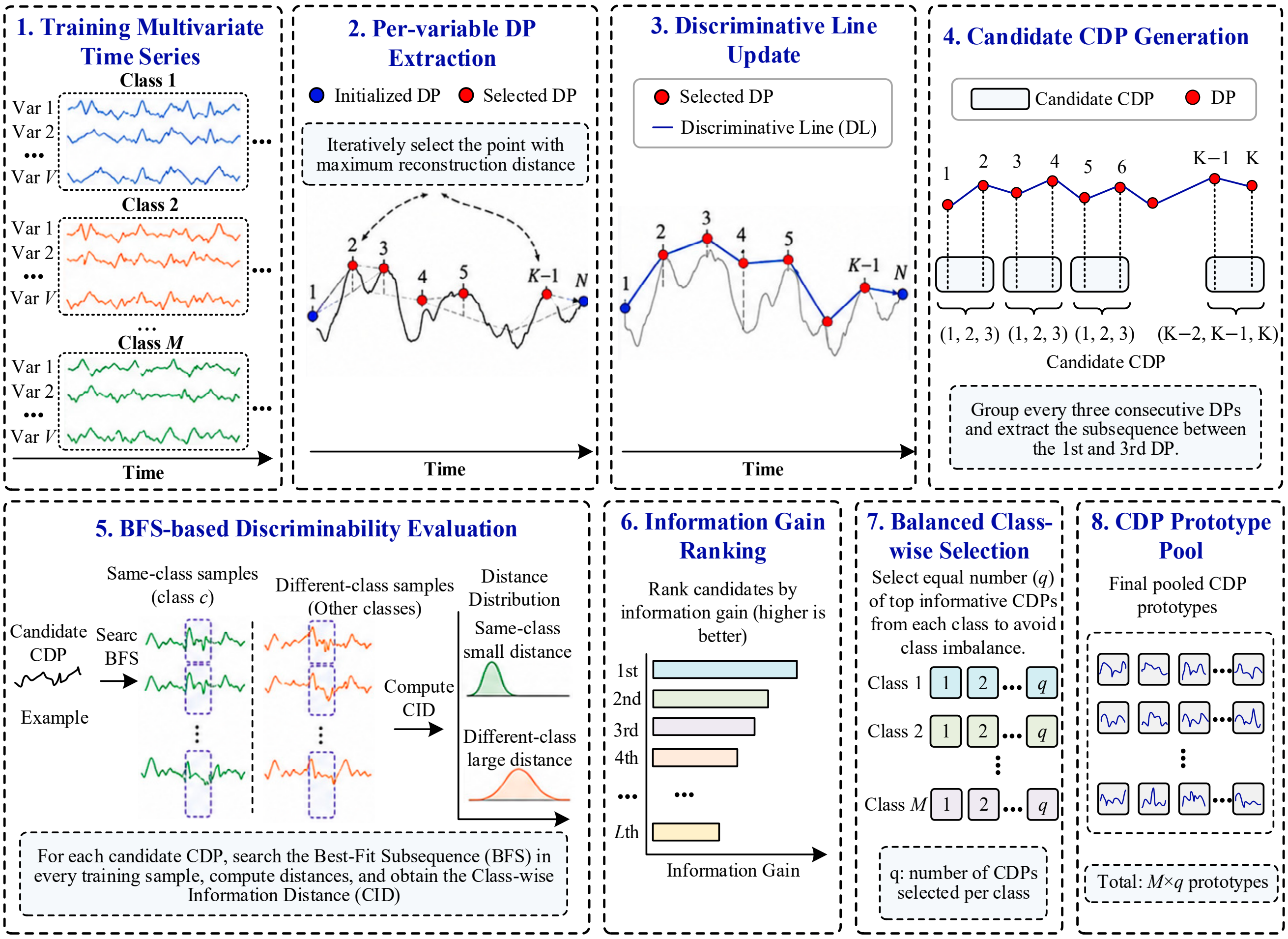

Figure 1. Construction process of class-specific discriminative prototypes based on discriminative time points and information-gain screening.

Specifically, discriminative points (DPs) are first extracted from each variable channel. The extracted DPs are then connected sequentially to form a discriminative line (DL). Next, the reconstruction distance from each remaining time point to the DL is calculated, and the time point with the largest distance is selected as the next DP to form a new DL. The above operation is repeated until a predetermined number of DPs has been selected. Finally, three consecutive DPs are taken as a group, and the original time-series segment between the first DP and the third DP is used as a candidate CDP.

In the initialization stage, the first and last time points of a time-series sample are used as two initial DPs, and the line connecting them is used as the initial DL. New DPs are then selected progressively according to the reconstruction distance. Let $RD(t)$ denote the reconstruction distance of a time point t in the sequence relative to the DL formed by the selected DPs. The newly added DP can be expressed as follows:

$$t^* = \text{argmax}_{1 \le t \le T} RD(t), \tag{1}$$

where $RD(t)$ denotes the reconstruction distance from time point t to the current reconstructed line segment, and the selected point denotes the newly added DP. This formula indicates that among all time points that have not yet been selected, the point with the largest reconstruction distance is selected as the new DP.

The v-th variable channel of a multivariate time-series sample X_n is denoted by $X_n^{(v)}$. After each new DP is added, candidate CDPs can be constructed using adjacent consecutive DPs. The starting and ending positions of a candidate CDP are denoted by l_r and r_r, respectively. Then, the candidate CDP can be expressed as follows:

$$G_r = X_n^{(v)}[l_r : r_r]. \tag{2}$$

After candidate CDPs are extracted, the best-fit subsequence (BFS) is further used to evaluate their class discriminability. Specifically, the role of the BFS is to search each training sample for the local segment most similar to the candidate CDP, thereby determining whether the sample contains a local fault pattern similar to the candidate prototype. If a sample and the candidate CDP belong to the same operating state, its BFS should generally have a small distance from the candidate CDP. If the sample belongs to another operating state, even the most similar BFS should still show a large distance from it. At this stage, CDP–BFS matching is conducted within the same signal channel. A candidate CDP extracted from one variable is compared only with local subsequences from the corresponding variable of each training sample. Same-class and different-class samples refer to samples with the same or different operating-state labels, rather than different signal channels. Cross-variable coupling is further introduced in the MCPM stage by extending selected CDPs into multivariate prototype blocks. Therefore, this paper calculates the matching distance between the candidate CDP and the BFS of each training sample, obtains its distance distributions on same-class and different-class samples, and further evaluates the discriminability of the candidate CDP for different operating states using information gain. A larger information gain indicates that the candidate CDP can better cluster same-class samples and separate different-class samples, and therefore has stronger class discriminability.

To evaluate the similarity between a candidate CDP and the BFS of a training sample, this paper introduces the complexity-invariant distance (CID). Compared with Euclidean distance, CID can consider both point-value differences and shape-complexity differences, and is therefore more suitable for measuring local fluctuations in EHA time series.

Specifically, given two time-series segments of the same length, X_a and X_b, representing the candidate CDP and the corresponding BFS, respectively, their complexity-invariant distance can be defined as follows:

$$\mathcal{D}_{cid}(X_a, X_b) = \|X_a - X_b\|_2 \cdot \frac{\max(\Gamma(X_a), \Gamma(X_b))}{\min(\Gamma(X_a), \Gamma(X_b)) + \eta}, \tag{3}$$

where the $\|\cdot\|_2$ denotes Euclidean distance, $\Gamma(\cdot)$ denotes the sequence-complexity estimation function, and η denotes a small constant used to avoid a zero denominator.

For a candidate CDP, the minimum matching distance between it and the BFS of sample X_n can be expressed as follows:

$$d_n(G_r) = \min_{\xi=1}^{T-L_r+1} \mathcal{D}_{cid}(X_n[\xi : \xi + L_r - 1], G_r), \tag{4}$$

where G_r is the candidate CDP, and L_r is the length of the candidate CDP. This formula indicates that the local segment most similar to the candidate CDP is searched in the input time-series sample, and the minimum distance is used as the matching distance between them.

After obtaining the minimum matching distance between the candidate CDP and all samples, information gain is further used to evaluate its class discriminability. Let the proportion of samples from the m-th operating state in the training set D be denoted by π_m; then the information entropy of the training set is

$$\mathcal{H}(\mathcal{D}) = -\sum_{m=1}^{M} \pi_m \log(\pi_m + \eta). \quad (5)$$

For a candidate G_r, the distances between it and all training samples are sorted in ascending order, and a candidate set of splitting thresholds can be constructed from the midpoints of adjacent distance values:

$$\Theta_r = \left\{ \vartheta_k = \frac{d_{(k)} + d_{(k+1)}}{2} \right\}_{k=1}^{N-1}, \quad (6)$$

The training set can be divided into two subsets with smaller and larger distances according to different thresholds:

$$\mathcal{D}^-(\vartheta) = \{(X_n, c_n) |\, d_n(G_r) \leq \vartheta\}, \quad (7)$$

$$\mathcal{D}^+(\vartheta) = \{(X_n, c_n) \mid d_n(G_r) > \vartheta\}. \quad (8)$$

The information gain of the candidate CDP under threshold ϑ can be expressed as

$$\mathcal{I}(G_r, \vartheta) = \mathcal{H}(\mathcal{D}) - \frac{|\mathcal{D}^-(\vartheta)|}{|\mathcal{D}|}\mathcal{H}(\mathcal{D}^-(\vartheta)) - \frac{|\mathcal{D}^+(\vartheta)|}{|\mathcal{D}|}\mathcal{H}(\mathcal{D}^+(\vartheta)). \quad (9)$$

The final information gain of the candidate CDP can be expressed as

$$\mathcal{I}(G_r) = \max_{\vartheta \in \Theta_r} \mathcal{I}(G_r, \vartheta). \quad (10)$$

A larger information gain indicates that the candidate CDP can reduce the class uncertainty after sample partitioning more effectively, and thus has stronger discriminative ability for different EHA operating states. Finally, to prevent the feature space from being dominated by majority-class samples under data-imbalance conditions, the same number of highly discriminative CDPs is selected for all classes.

Let the candidate CDP set of the m-th class be denoted as $\mathcal{G}_m$, and let the total number of retained CDPs be g. The g/M candidate prototypes with the highest information gains are selected from the m-th class as follows:

$$\Omega_m = \text{Top}_{g/M}\{G_r \in \mathcal{G}_m \mid \mathcal{I}(G_r)\}. \quad (11)$$

The final CDP prototype pool is obtained as

$$\Omega_{CDP} = \{\Omega_1, \Omega_2, \cdots \Omega_m\}. \quad (12)$$

Through the above process, a CDP prototype pool with local representativeness and class discriminability is extracted from EHA multivariate time series. This prototype pool can provide highly discriminative local feature information for different operating states.

Furthermore, the information-gain screening mechanism improves the stability of the extracted prototypes under practical perturbations. EHA signals often contain measurement noise, transient hydraulic fluctuations, and operating-condition variations caused by dynamic aerodynamic loads. As random noise and transient drift usually lack consistent morphology across samples from the same operating state, perturbation-dominated candidate segments cannot effectively partition the training set, and therefore obtain low information gain. These unstable segments are filtered out during prototype screening, whereas the retained class-specific discriminative prototypes correspond to statistically recurrent fault signatures. This mechanism enhances the robustness of prototype representation under complex dynamic interferences without requiring additional denoising procedures.

3.2. Multivariate Coupling-Aware Prototype Matching Strategy (MCPM)

During training, when highly discriminative CDPs are matched with BFSs in the input time series, the two need to be further mapped into learnable feature parameters and their feature differences need to be quantified before being input into the Transformer for encoding. However, because an EHA is composed of multiple coupled components such as a servo motor, piston pump, hydraulic actuator cylinder, and sensors, different fault states are often manifested through linkage relationships among variables such as current, displacement, velocity, and pressure. Therefore, simply comparing local morphological differences on a single variable usually cannot fully reflect fault propagation and state-change characteristics.

To this end, this paper proposes a multivariate coupling-aware prototype matching strategy (MCPM). This strategy consists of three steps. First, univariate CDPs are extended into multivariate local prototype blocks containing related variables. Second, the extraction-position information of CDPs is used for BFS search and matching, thereby minimizing search time. Finally, the differences between CDPs and BFSs are quantified as features and input into the Transformer for encoding. The workflow of the proposed multivariate coupling-aware prototype matching strategy is shown in Figure 2. The detailed principle is as follows.

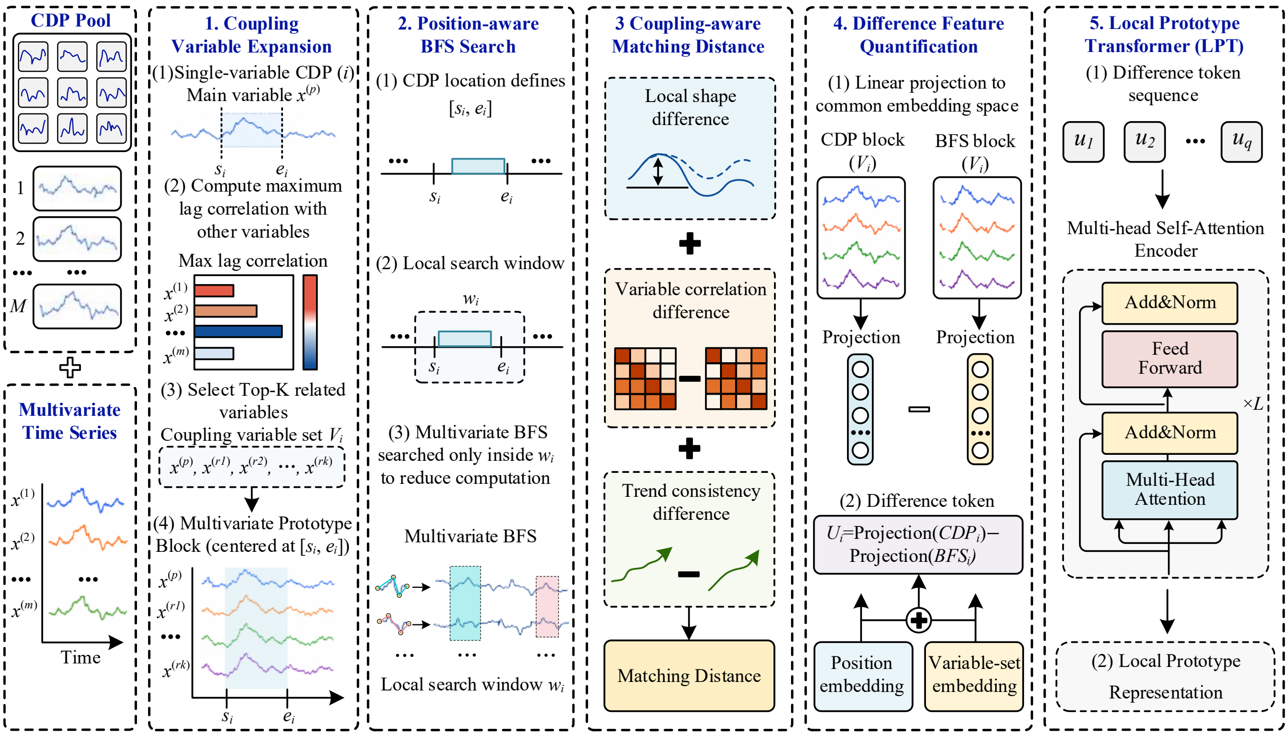

Figure 2. Multivariate coupling-aware CDP–BFS matching and local prototype feature encoding process.

3.2.1. Multivariate Coupling-Aware BFS Search

Let the source sample of the i-th CDP be X_i, and let its start and end positions be s_i and e_i, respectively, with length $L_i = e_i - s_i + 1$. Considering that EHA fault differences are not only reflected in local morphological variations of a single variable, but may also appear as changes in synchronization, lag, or coupling relationships among multiple variables, a corresponding coupling-variable set is first constructed for each CDP.

Specifically, the maximum lagged correlation between the primary variable v_i and another variable q within the local time interval is calculated as

$$\kappa_{i,q} = \max_{\tau \in [-\tau_0, \tau_0]} \left| \mathrm{Corr}\left(X_i^{(v_i)}[s_i : e_i], X_i^{(q)}[s_i + \tau : e_i + \tau] \right) \right|. \tag{13}$$

According to the correlation magnitude, the K variables most strongly correlated with the primary variable are selected and combined with the primary variable to form a coupling-variable set:

$$\mathcal{M}_i = \{v_i\} \cup \mathrm{Top}K_{q \neq v_i}(\kappa_{i,q}), \tag{14}$$

where τ denotes the time lag, τ_0 denotes the maximum allowable lag range, and $\mathcal{M}_i$ denotes the coupling-variable set corresponding to the i-th CDP. Thus, the i-th CDP is no longer represented as a univariate local subsequence, but is extended into a multivariate local prototype block:

$$P_i = X_i^{\mathcal{M}_i}[s_i : e_i] \in \mathbb{R}^{|\mathcal{M}_i| \times L_i}. \tag{15}$$

Therefore, the proposed strategy can not only measure local morphological differences between CDP and BFS, but also further optimize the BFS search and matching process by measuring inter-variable correlation and trend consistency.

First, local dissimilarity can be defined as the weighted sum of the CIDs of all coupled variable channels:

$$\Delta_i^{loc}(j) = \sum_{q \in \mathcal{M}_i} \alpha_{i,q} \mathcal{D}_{cid}\left(X_i^{(q)}[j : j + L_i - 1], P_i^{(q)} \right), \tag{16}$$

where $\alpha_{i,q}$ is the weight of variable q, which can be obtained by normalizing its correlation with the primary variable.

Second, the inter-variable coupling relationship can be defined as the variable correlation between the prototype block and the candidate segment:

$$\Delta_i^{rel}(j) = \sum_{q<h} \left| \mathrm{Corr}\left(X_i^{(q)}[j : j + L_i - 1], X_i^{(h)}[j : j + L_i - 1] \right) - \mathrm{Corr}\left(P_i^{(q)}, P_i^{(h)} \right) \right|. \tag{17}$$

Meanwhile, to characterize the consistency of variation trends among different variables during changes in EHA operating states, a first-order difference trend is introduced:

$$\Delta_i^{slo}(j) = \sum_{q \in \mathcal{M}_i} \left[1 - \mathrm{Corr}\left(\nabla X^{(q)}[j : j + L_i - 1], \nabla P_i^{(q)} \right) \right]. \tag{18}$$

Before fusion, the three dissimilarity terms are first scale-aligned over the training set to avoid directly adding heterogeneous raw quantities:

$$\widetilde{\Delta}_i^k(j) = \frac{\Delta_i^k(j)}{\mu^k + \varepsilon}, \quad k \in \{loc, rel, slo\}, \tag{19}$$

where μ^k denotes the mean value of the corresponding dissimilarity term computed from the training CDP–BFS pairs. The parameter ε is a small constant. This normalization converts local morphology, variable correlation, and trend consistency into non-negative and dimensionless terms with comparable magnitudes. Finally, by integrating the three normalized components, the multivariate coupling-aware matching distance is defined as

$$\mathcal{D}_i^{mcp}(j) = \widetilde{\Delta}_i^{loc}(j) + \beta_1\,\widetilde{\Delta}_i^{rel}(j) + \beta_2\,\widetilde{\Delta}_i^{slo}(j). \tag{20}$$

The additive formulation represents a linear scalarization of a multiview matching criterion. Each term penalizes one specific aspect of the CDP–BFS mismatch. These aspects include local shape deviation, inter-variable coupling deviation, and trend inconsistency. This formulation provides a single searchable objective while preserving complementary information from different views. The local morphological term serves as the reference term with a fixed coefficient of 1. The coefficients β_1 and β_2 balance the contributions of the coupling and trend terms, respectively. After normalization, both coefficients are set to 1.0 based on the validation set. They are fixed during testing to avoid information leakage. The multivariate coupling-aware matching distance is then used to search for the multivariate BFS that best matches the CDP in the input sample. Compared with a matching method measuring only univariate morphological differences, this distance simultaneously characterizes changes in coupling relationships and trend-coordination differences among multiple variables. This multiview measurement significantly improves the diagnostic capability of CDP–BFS matching for EHA fault patterns.

To reduce the computational overhead caused by global search, the proposed strategy innovatively searches for BFS only near the CDP extraction position. The start and end positions of the position-aware search window w_i for the i-th CDP can be expressed as

$$w_{s_i} = \max(1, s_i - w); \tag{21}$$

$$w_{e_i} = \min(T - L_i + 1, s_i + w). \tag{22}$$

The primary variable and its coupled variables share the same search window to preserve their temporal correspondence within the multivariate prototype block. Allowing independent searches for different variables may align them with different oscillation cycles and thus distort the correlation and trend-consistency measurements. According to Equations (21) and (22), the search window contains at most $2w + 1$ candidate start positions and is adaptively truncated near the sequence boundaries.

The restricted search window is introduced not merely for computational acceleration, but also to preserve the temporal context of EHA fault signatures. In this study, the EHA responses are collected under a Dutch-roll maneuver, which imposes periodic sinusoidal excitation on the actuator system. Under this condition, morphologically similar segments may occur in adjacent cycles. Therefore, unrestricted global search or unconstrained time-warping may produce phase-mismatched matches by aligning a prototype with a physically irrelevant segment from another cycle. By limiting the BFS search to the vicinity of the original CDP extraction position, the proposed strategy enforces local phase consistency, reduces cross-cycle mismatching, and decreases search complexity without overlooking diagnostically relevant local features.

Within the search window, the start position that minimizes $\mathcal{D}_i^{mcp}(j)$ is used as the BFS start position:

$$j_i^{best} = \mathrm{argmin}_{j \in [w_{s_i}, w_{e_i}]} \mathcal{D}_i^{mcp}(j). \tag{23}$$

Starting from j_i^{best}, a subsequence of length L_i is extracted to obtain the multivariate BFS:

$$B_i = X_i[j_i^{best} : j_i^{best} + L_i - 1]. \tag{24}$$

Compared with univariate CDP, the above process extends CDP–BFS matching from a single local morphological segment to a multivariate local prototype block, and enables the BFS matching process to simultaneously perceive local morphological differences and changes in variable coupling relationships, making it more suitable for multivariate time-series fault-diagnosis tasks.

3.2.2. Quantification of CDP–BFS Differential Features

As shown in Figure 3, after obtaining the BFS corresponding to the i-th CDP, the CDP and BFS are further mapped into a unified embedding space, and the feature difference between them is calculated. Let $\mathcal{F}(B_i)$ and $\mathcal{F}(P_i)$ denote the linear projection functions of BFS and CDP, respectively; then, the differential feature corresponding to the i-th CDP–BFS pair is

$$h_i = \mathcal{F}(B_i) - \mathcal{F}(P_i), \tag{25}$$

where $h_i \in \mathbb{R}^{1\times d_l}$ denotes the embedding dimension of the differential feature. This differential feature not only reflects the degree of morphological deviation between BFS and CDP, but also contains matching differences caused by variations in variable coupling relationships. Therefore, for EHA fault diagnosis, h_i can be regarded as a representation of the deviation of the input sample from a typical local fault pattern.

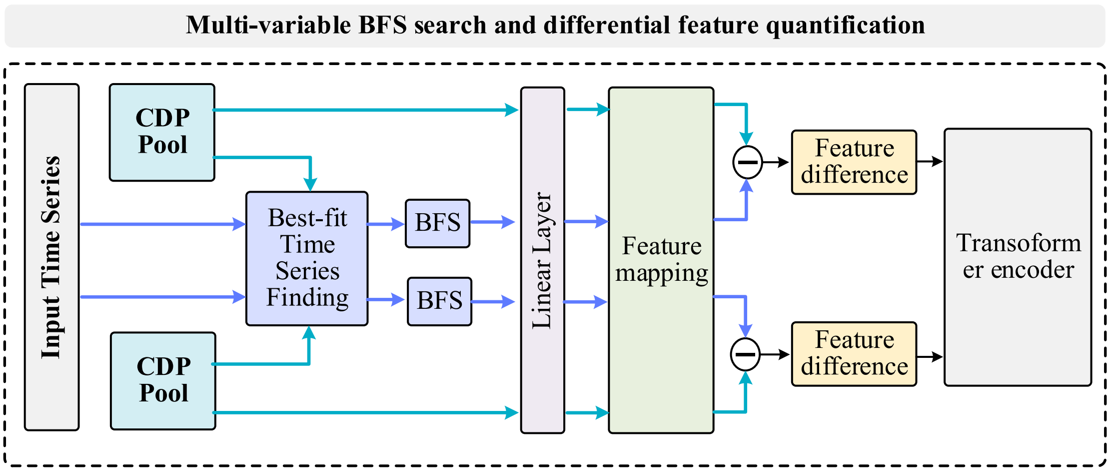

Figure 3. Quantification of feature differences based on CDP and BFS.

To preserve the temporal position and variable-source information of CDP, positional embedding is further introduced. The start position, end position, and coupling-variable-set embedding of CDP are added to the differential feature:

$$\hat{h}_i = h_i + e(s_i) + e(e_i) + e(\mathcal{M}_i), \tag{26}$$

where $e(\mathcal{M}_i)$ denotes the overall embedding of the coupling-variable set. Through the above processing, the differential feature simultaneously contains the local matching difference between CDP and BFS, temporal-position information, and variable-coupling information, and can be used as an input token for the subsequent Transformer.

3.2.3. Local Prototype Transformer

To further model dependencies among prototypes, this paper constructs a Local Prototype Transformer. Let the input differential-feature sequence be denoted as $\hat{H} =$

$\left\{\hat{h}_1, \hat{h}_2, \cdots \hat{h}_g\right\}$, where g is the number of CDP prototypes. For the *i*-th differential feature, its output after the multihead attention (MHA) mechanism is

$$r_i^{lpt} = \sum_{j=1}^{g} \omega_{i,j}\left(\hat{h}_j A_V\right), \quad (27)$$

where A_V denotes the value-mapping matrix, and $\omega_{i,j}$ denotes the attention weight from the *i*-th prototype differential feature to the *j*-th prototype differential feature:

$$\omega_{i,j} = \text{softmax}\left(\frac{(\hat{h}_i A_Q)(\hat{h}_j A_K)^T}{\sqrt{d_l}}\right), \quad (28)$$

where the corresponding matrices denote the query matrix and key matrix, respectively. Through this mechanism, LPT can learn correlations among the matching differences of different CDPs, allowing prototypes with similar fault-discrimination patterns to form stronger information interactions while suppressing interference from irrelevant or weakly discriminative prototypes.

3.3. Global Generic Transformer (GGT)

Although LPT can effectively characterize the matching differences between locally discriminative prototypes and input samples, relying only on local prototype features may ignore the overall dynamic trends and global information of time series. Therefore, to reduce the overfitting risk caused by the high sensitivity of local patterns, this paper further constructs a Global Generic Transformer to extract cross-class shared global generic features from multivariate time series.

Taking the original EHA multivariate time series as input, a convolutional mapping module is first used to extract generic temporal features. The convolutional mapping module consists of one-dimensional convolution, batch normalization, and a GELU activation function, and its calculation process is

$$\Phi_c(X) = \text{GELU}(\text{BN}(\text{Conv1}D(X))), \quad (29)$$

where the convolutional operation is used to extract local temporal patterns and basic inter-variable associations, batch normalization is used to stabilize feature distributions, and the GELU activation function is used to enhance nonlinear representation capability.

Subsequently, two convolutional mapping modules are used for consecutive feature extraction from the input sequence, yielding generic features:

$$Q = \Phi_c(\Phi_c(X)), \quad (30)$$

where $Q = \{q_1, q_2, \cdots q_n\}$ denotes the generic feature-token sequence extracted by the convolutional modules.

Because long-range dependencies still exist among the generic feature tokens, this paper further inputs them into a multihead attention module to model global correlations among different generic features. Considering that each token output by convolution lacks explicit positional information, learnable positional embedding is introduced, and the output of GGT is

$$r^{ggt} = \text{MHA}\left(Q + E_g\right). \quad (31)$$

Finally, because the input of the GGT branch is generic feature tokens for all classes, average pooling is used to obtain the global generic representation:

$$z_g^* = \text{AvgPool}\left(r^{ggt}\right), \tag{32}$$

where z_g^* denotes the final feature representation output by the Global Generic Transformer branch.

3.4. Local Prototype-Global Generic Dual-Branch Transformer (LPG-Former)

To fully fuse local prototype features and global generic features, this paper constructs a Local Prototype-Global Generic Dual-branch Transformer. The architecture of the proposed LPG-Former is shown in Figure 4. For the LPT branch, because the input tokens originate from the matching-difference features of different CDPs, they have explicit class-discriminative meaning. To avoid average pooling weakening the discriminative information carried by high-information-gain prototypes, this paper selects the token corresponding to the CDP with the highest information gain from the LPT output as the final representation of the local prototype branch, denoted as Z_*^{lpt}. It should be noted that this token has completed attention interaction with the differential features of other CDPs in LPT; therefore, it not only retains the dominant discriminative information of the prototype with the highest information gain, but also integrates contextual association features from other prototypes. For the GGT branch, the global generic representation Z_*^{ggt} has been obtained by average pooling as described above. Subsequently, the outputs of the two branches are concatenated:

$$Z = \text{concat}\left(Z_*^{lpt}, Z_*^{ggt}\right). \tag{33}$$

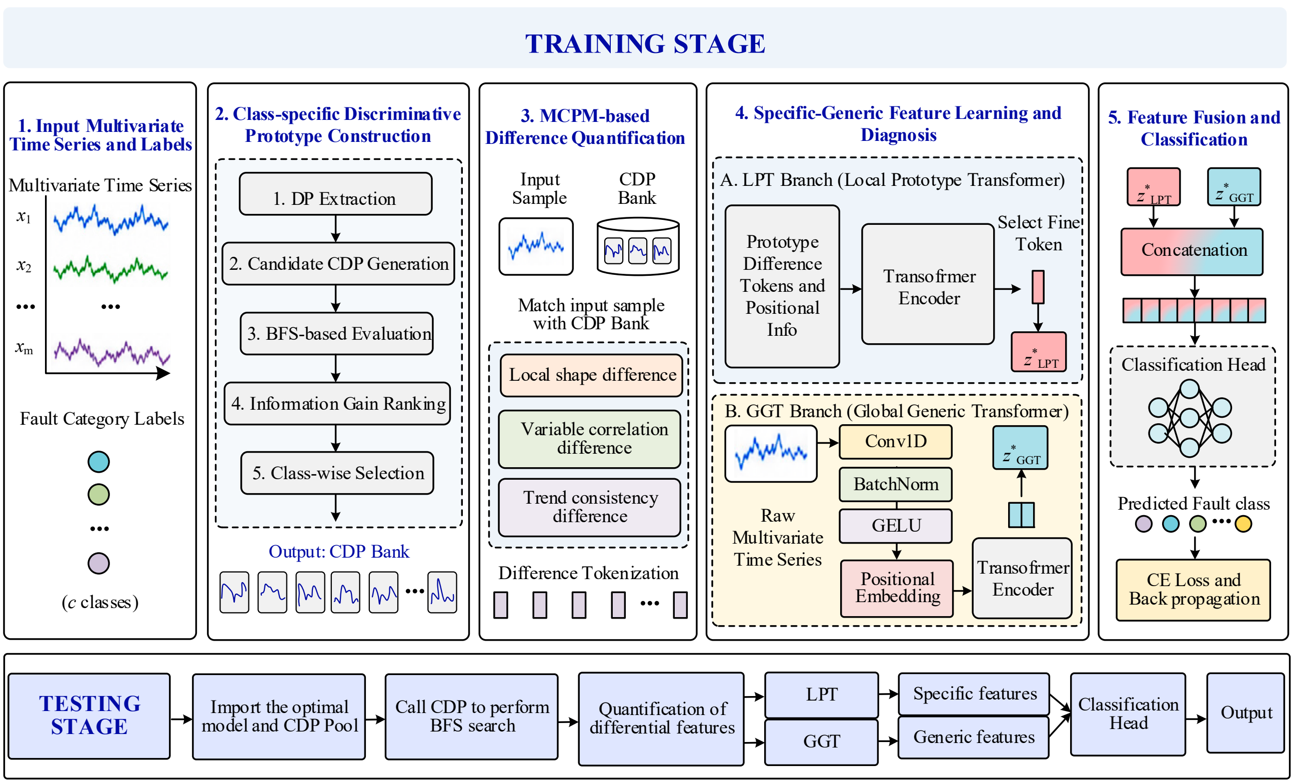

Figure 4. Overall architecture of the proposed LPG-Former.

In the classification stage, the fused features are input into a linear classification layer, and class probability distributions are obtained through Softmax. The final predicted class is

$$\hat{y} = \text{argmax}(\text{softmax}(\text{Linear}(Z))). \tag{34}$$

During training, the cross-entropy loss function is used to optimize the overall network parameters. Its objective function is

$$\mathcal{L} = \mathcal{L}_{CE}(y, \hat{y}), \tag{35}$$

where y denotes the true class label, and the corresponding term denotes the model prediction.

4. EHA Fault Dataset

4.1. EHA Fault Experimental Platform

This paper uses an aviation electro-hydrostatic actuator (EHA) fault simulation experimental platform to construct a multivariate fault time-series dataset. The platform takes the aileron EHA system of a fixed-wing aircraft as the research object and uses the Dutch-roll maneuver as a typical flight condition. During the Dutch-roll maneuver, the aircraft heading angle and roll angle vary periodically in sinusoidal form, and the aileron deflection angle serves as the attitude-adjustment input that drives the EHA system. According to the aileron control-surface deflection demand, the hydraulic actuator-piston displacement command can be further obtained, and time-varying commands and composite loading conditions can be constructed by combining aerodynamic loads and axial structural vibration velocity, thereby simulating the dynamic response process of the EHA system under flight conditions. In the simulation platform, the Dutch-roll maneuver is applied to the EHA through two coupled inputs. First, the aileron deflection demand generated by the maneuver is converted into the actuator-piston displacement command, which serves as the motion input of the EHA closed-loop control system. Second, the corresponding time-varying aerodynamic load is introduced through the aileron-load module and acts on the actuator output side through the transmission mechanism. Therefore, the EHA is driven by both the displacement command and the external load variation during the maneuver, allowing the simulated response to reflect the coupled influence of flight-control motion and aerodynamic loading.

The established EHA system mainly consists of a servo motor, piston pump, hydraulic actuator cylinder, transmission mechanism, and aileron load, as shown in Figure 5. Specifically, the hydraulic power unit is represented by a servo-motor-driven piston pump, while the actuator is modeled as a two-chamber hydraulic cylinder. The servo motor is formulated in the d–q coordinate system to characterize the current–speed–torque relationship, and the pump-generated flow establishes the chamber-pressure difference that drives the actuator rod and subsequently deflects the aileron through the transmission mechanism. The servo motor drives the piston pump to output hydraulic flow. The hydraulic actuator cylinder generates displacement output under the pressure difference between the two chambers and drives the aileron control surface to deflect through the transmission mechanism. To describe the electrical–hydraulic–mechanical coupled dynamic characteristics of the EHA system, the platform establishes a servo-motor model, a hydraulic-subsystem model, and a load-coupling model. The servo-motor part describes the relationships among motor current, rotational speed, and output torque based on the d–q two-phase coordinate system. The hydraulic-subsystem part describes the dynamic coupling relationships among the flow, pressure, leakage, hydraulic oil compressibility, and external load of the working chamber and return chamber. The motor load torque is jointly determined by the hydraulic-system pressure difference and pump displacement. Through the above modeling process, the platform can relatively completely characterize the dynamic operating process of the EHA during flight maneuvering tasks.

In terms of fault injection, the platform simultaneously considers sensor faults, hydraulic-system faults, and mechanical-actuator faults. The green boxes in Figure 5 denote the fault-injection variables used to modify the corresponding model parameters or feedback signals. For sensor-related faults, these variables represent feedback signal bias, gain variation, or stuck value. For hydraulic and actuator faults, they represent changes in leakage, cylinder deformation or aging, and initial rod-displacement offset. For the sensor feedback loop, the control-feedback state variables mainly include actuator-cylinder displacement, servo-motor speed, and motor current, and bias faults, gain faults, and stuck faults are set for these feedback variables. For the hydraulic and actuator components, hydraulic oil leakage, actuator-cylinder deformation, actuator-cylinder aging, and initial rod-displacement deviation are further set. These faults cover typical failure forms in EHA systems, including sensor measurement anomalies, degradation of hydraulic transmission performance, and abnormalities in mechanical actuator components.

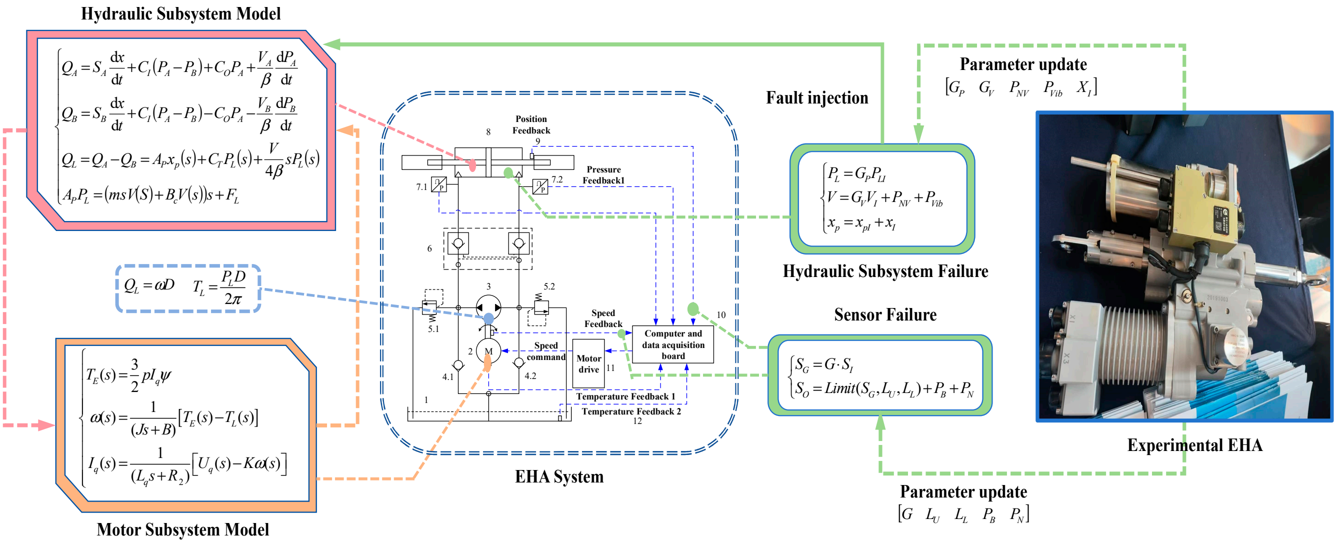

Figure 5. Aviation electro-hydrostatic actuator (EHA) fault simulation experimental platform.

The severities of the injected faults are quantified using the corresponding parameters in the lumped parameter dynamic model and the feedback signal model. For hydraulic oil leakage, a fixed leakage flow rate is not prescribed, because the pressure difference between the actuator chambers varies continuously during the Dutch-roll maneuver. Instead, leakage severity is represented by a leakage-related hydraulic coefficient that modifies the pressure and flow transmission behavior of the hydraulic subsystem. For actuator-cylinder deformation and aging, physical degradation such as abrasion, sealing deterioration, or diameter deviation is represented by equivalent dynamic parameters that affect mechanical resistance and actuator response. Initial rod-displacement deviation is quantified by the imposed initial displacement offset, while stuck, gain, and bias faults are quantified by the stuck limit, gain coefficient, and bias value in the feedback signal path, respectively. This parameter-based quantification is consistent with the governing equations of the simulation platform and preserves the time-varying coupling among current, pressure, flow, and displacement responses.

Based on this fault experimental platform, EHA dynamic response data under different operating states can be obtained by changing the fault-injection location, fault type, and fault parameters. Because the platform simultaneously considers flight-condition excitation, hydraulic-load variation, and multicomponent coupling relationships, the generated data can better reflect linkage variations among multivariate response signals during EHA fault propagation, providing a data basis for the multivariate time-series fault-diagnosis experiments in this paper.

4.2. Fault Data Acquisition

Based on the EHA fault simulation experimental platform, this paper collects multivariate time-series response signals under different operating states as fault-diagnosis data. It should be clarified that the dataset used in this study is generated from the EHA fault simulation experimental platform rather than directly collected from physical sensors on a hardware test bench. The monitored variables are obtained as simulated response outputs of the EHA model under the Dutch-roll maneuver and different fault-injection settings. Therefore, the term "data acquisition" in this study refers to the extraction and recording of multivariate simulated response signals from the simulation platform. Because the EHA system is composed of multiple coupled components such as a servo motor, piston pump, hydraulic actuator cylinder, sensors, and mechanical transmission mechanism, different fault states not only cause local amplitude or trend changes in a single variable, but also change the correlation and coordinated variation relationships among multiple monitoring variables. Therefore, this paper uses multivariate monitoring signals during EHA operation as model input to characterize dynamic response features under different fault states.

Specifically, the platform synchronously collects key state variables and response variables of the EHA under each operating state to form multivariate fault time-series samples. Compared with univariate signals, multivariate time series contain not only the temporal evolution pattern of each variable itself, but also variations in coupling relationships among different variables, and are therefore more suitable for characterizing fine-grained state differences during EHA fault propagation.

The input to LPG-Former comprises 11 synchronized signal channels: the difference between the motor output torque and the actuator-damping reaction torque DT, actual motor current I_R, actual motor speed N_R, actuator-cylinder pressure P, sensor-measured rod displacement S_D, sensor-measured motor speed S_N, sensor-measured motor current S_I motor voltage U, rod velocity V, actual rod displacement X_R, and rod-displacement command $(X_{R,\mathrm{Cmd}})$. The simulation-time variable T is used as a temporal index. Specifically, variations in motor current affect the motor torque and pump flow, while the resulting cylinder-pressure difference governs the actuator-rod motion. The rod-displacement feedback subsequently modifies the closed-loop current demand, forming a direct physical linkage among current, pressure, and displacement responses.

Because the dataset is generated by the EHA simulation platform, no physical hardware sensors are involved. The S_D, S_N, and S_I channels emulate rod-displacement, motor-speed, and motor-current sensors, respectively, whereas the remaining channels are extracted from the corresponding model states or control variables. All 11 channels are synchronously recorded over a 100 s Dutch-roll maneuver using a simulation time step of 0.01 s, corresponding to a sampling rate of 100 Hz. Each complete maneuver record therefore contains 10,001 synchronized time points per channel, including the initial and final instants.

To further illustrate the motion response of the aileron EHA under the Dutch-roll maneuver, Figure 6 presents the time history of the EHA actuator-rod displacement. The displayed maneuver spans 0–100 s with a simulation time interval of 0.01 s, and the actuator rod exhibits a periodic reciprocating motion corresponding to the oscillatory flight excitation.

https://doi.org/10.3390/act15070395

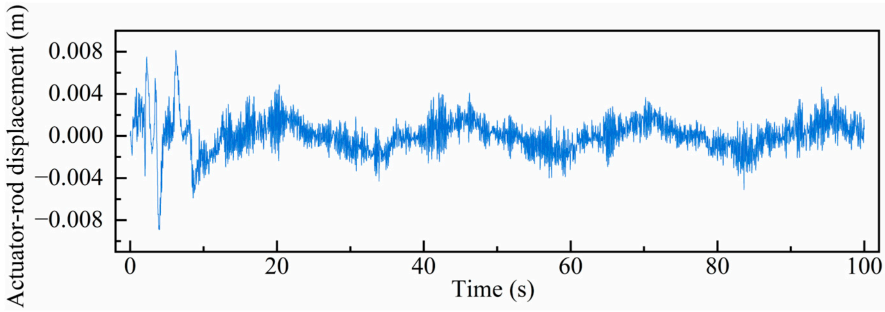

Figure 6. Time history of the aileron EHA actuator-rod displacement during the Dutch-roll maneuver.

To clarify the physical connection between the EHA dataset and the fault mechanisms, the operating states are further explained from the perspective of signal responses. The normal condition represents the baseline coordinated state of the motor, hydraulic transmission, and actuator motion. Initial rod-displacement deviation directly changes the initial actuator position and may induce abnormal tracking errors and transient current variations. Hydraulic oil leakage weakens pressure-holding and transmission efficiency, leading to pressure attenuation, increased flow compensation, and slower displacement response. Cylinder deformation and aging change the mechanical constraint, friction, or sealing characteristics of the actuator, which may cause abnormal pressure fluctuations, displacement trends, and current responses. Stuck, gain, and bias faults are associated with abnormal feedback or signal-transmission characteristics, and their effects can be propagated through the closed-loop control process. Therefore, EHA faults are reflected not only in local variations of individual signals, but also in coupling changes among current, pressure, flow, and displacement signals, which provides the physical basis for multivariate time-series diagnosis.

According to the fault-injection type, this paper constructs data for eight EHA operating states. Normal condition (NC) represents the normal operating state of the EHA; initial rod-displacement deviation (IRDD) simulates an actuator anomaly when the initial piston-rod position is offset; hydraulic oil leakage fault (HOLF) simulates abnormal pressure and flow transmission caused by hydraulic oil leakage; actuator-cylinder deformation fault (ACDF) and actuator-cylinder aging fault (ACAF) simulate changes in actuator dynamic characteristics caused by cylinder deformation and aging degradation, respectively; stuck fault (SF) simulates abnormal state variables caused by sensor or actuator sticking; and gain fault (GF) and bias fault (BF) simulate gain anomalies and offset anomalies in sensor measurements, respectively.

The selected fault types cover the main failure mechanisms of an aviation electro-hydrostatic actuator, including hydraulic transmission degradation, mechanical actuator abnormality, and feedback signal anomaly. These faults are selected because they directly disturb the coupled electrical, hydraulic, and mechanical responses and represent typical degradation modes in flight control actuation systems. Regarding fault severity, each fault category is generated using a representative severity setting that can produce observable dynamic deviations. As this study focuses on fault pattern classification rather than severity regression, the dataset is designed to distinguish different physical fault mechanisms from the resulting multivariate time-series responses.

The simulated responses are consistent with the expected physical behavior of the system. Hydraulic oil leakage occurs in the hydraulic transmission subsystem and weakens the pressure holding capability between the actuator chambers. This causes pressure attenuation and, under closed loop control, increases the flow compensation demand, slows the displacement response, and induces compensatory changes in motor current. Actuator-cylinder deformation and aging alter mechanical constraints, friction, or sealing characteristics, leading to abnormal pressure fluctuations, displacement tracking delay, and transient current variations. Initial rod-displacement deviation shifts the initial actuator position and may induce tracking error and transient motor current compensation. Stuck, gain, and bias faults occur in the feedback signal paths, where distorted feedback signals propagate through the controller and affect current, pressure, and displacement responses. Taken together, these signal responses indicate that an EHA fault may appear as a local deviation in an individual channel while simultaneously altering the response delay, inter-variable correlation, and trend consistency among current, pressure, and displacement. Therefore, jointly modeling multiple synchronized signals is necessary to distinguish fault states with similar single-channel waveforms. The state codes corresponding to different operating states in the specific experimental analysis are shown in Table 1.

Table 1. EHA operational status code.

Operational Status	Status Code
Normal Condition	NC
Initial Rod-Displacement Deviation	IRDD
Hydraulic Oil Leakage Fault	HOLF
Actuator-Cylinder Deformation Fault	ACDF
Actuator-Cylinder Aging Fault	ACAF
Stuck Fault	SF
Gain Fault	GF
Bias Fault	BF

The acquired EHA multivariate fault time series exhibit substantial fine-grained classification difficulty. On the one hand, the overall response morphologies under different operating states are relatively similar, and some classes differ only in amplitude, trend, or fluctuation degree over local time segments. On the other hand, the influence of faults on system states may be indirectly reflected through coupling relationships among multiple variables, and inter-class differences may not be obvious when only a single variable is observed. For example, stuck faults, gain faults, or bias faults may cause only weak amplitude changes within local time periods, whereas faults such as hydraulic oil leakage, actuator-cylinder deformation, and actuator-cylinder aging may simultaneously affect the coordinated relationships among motor-drive response, hydraulic-transmission state, and actuator-motion state. Therefore, this dataset has characteristics including subtle inter-class differences, hidden local discriminative information, and complex multivariate coupling relationships.

The above data characteristics are highly consistent with the multivariate fault time-series diagnostic task studied in this paper. Traditional global feature-learning methods are easily affected by overall waveform similarity and have difficulty fully capturing local discriminative segments and variable-coupling variations. Therefore, this paper subsequently verifies the effectiveness of the proposed LPG-Former method on the EHA fine-grained fault-identification task using this dataset.

4.3. Division of the EHA Fault Dataset

This paper uses aviation electro-hydrostatic actuator multivariate fault time series as experimental data to verify the effectiveness of the proposed LPG-Former in fine-grained fault identification. The dataset contains eight operating states, namely, normal condition (NC), initial rod-displacement deviation (IRDD), hydraulic oil leakage fault (HOLF), actuator-cylinder deformation fault (ACDF), actuator-cylinder aging fault (ACAF), stuck fault (SF), gain fault (GF), and bias fault (BF).

To further evaluate the classification performance of the model under different sample-distribution conditions, this paper constructs a balanced dataset and a class-imbalanced dataset. The specific sample divisions of the two datasets are shown in Tables 2 and 3.

Table 2. Sample division of the eight EHA operating states under balanced data.

Number	Description	Training Set	Validation Set	Test Set
F0	Normal Condition	140	40	20
F1	Initial Rod-Displacement Deviation	140	40	20
F2	Hydraulic Oil Leakage Fault	140	40	20
F3	Actuator-Cylinder Deformation Fault	140	40	20
F4	Actuator-Cylinder Aging Fault	140	40	20
F5	Stuck Fault	140	40	20
F6	Gain Fault	140	40	20
F7	Bias Fault	140	40	20

Table 3. Sample division of the eight EHA operating states under class-imbalanced data.

Number	Description	Training Set	Validation Set	Test Set
F0	Normal Condition	85	40	20
F1	Initial Rod-Displacement Deviation	102	40	20
F2	Hydraulic Oil Leakage Fault	36	40	20
F3	Actuator-Cylinder Deformation Fault	58	40	20
F4	Actuator-Cylinder Aging Fault	91	40	20
F5	Stuck Fault	74	40	20
F6	Gain Fault	43	40	20
F7	Bias Fault	68	40	20

The balanced dataset maintains the same number of samples for each class in the training, validation, and test sets, and is used to evaluate the basic classification capability of the model under balanced class distribution. The class-imbalanced dataset adjusts only the number of samples in each class in the training set, while the validation and test sets still maintain the same number of samples for each class, ensuring fair comparison of different methods under the same evaluation distribution. This setting can simulate the shortage of samples for some classes in practical EHA fault diagnosis while avoiding interference in performance-evaluation results caused by changes in class proportions in the validation and test sets.

The class-imbalanced dataset is constructed as a controlled evaluation setting rather than a real fault-occurrence distribution, with IRDD set as the majority class to evaluate whether the model can still recognize minority hydraulic, mechanical, and feedback-related fault classes under biased training samples. The normal class is not set as the majority class because this imbalanced setting is designed to test the model under biased fault-sample availability rather than to reproduce the normal-dominant distribution commonly observed in practical long-term monitoring.

5. Experimental Results and Analysis

5.1. Evaluation Metrics

To comprehensively evaluate the recognition performance of the proposed LPG-Former in EHA multivariate fault time-series classification, this paper uses accuracy, precision, recall, F1-score, the confusion matrix, and the false negative rate (FNR) as evaluation metrics. For multiclass classification tasks, precision, recall, F1-score, and FNR are calculated for each class using a one-vs-rest strategy, and the overall evaluation values are obtained by macro averaging.

Accuracy represents the proportion of correctly classified samples among all samples and is used to measure the overall classification correctness of the model. It is defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{36}$$

where TP denotes the number of positive samples correctly classified as positive, FP denotes the number of negative samples incorrectly classified as positive, FN denotes the number of positive samples incorrectly classified as negative, and TN denotes the number of negative samples correctly classified as negative.

Precision represents the proportion of samples predicted as positive that are actually positive, and reflects the reliability of the model when predicting a sample as a certain class. It is defined as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{37}$$

Recall represents the proportion of true positive samples correctly identified by the model and measures the completeness of the model in recognizing samples of the true class. It is defined as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{38}$$

F1-score is the harmonic mean of precision and recall and is used to comprehensively evaluate the classification performance of the model. It is defined as follows:

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{39}$$

To further evaluate the missed-detection control capability of the model during fault identification, this paper introduces the false negative rate (FNR), which is defined as follows:

$$\text{FNR} = \frac{\text{FN}}{\text{TP} + \text{FN}}. \tag{40}$$

FNR represents the proportion of true positive samples incorrectly predicted as negative, namely, the probability that the target fault is not correctly identified. A lower FNR indicates a smaller missed-detection risk for the target fault and higher safety and reliability of fault identification.

The confusion matrix can intuitively display the classification and misclassification relationships among different fault classes. Its calculation method is shown in Table 4.

Table 4. Symbol definitions of the confusion matrix under the one-vs-rest evaluation strategy.

Confusion Matrix		**Actual Class**	
		Positive	**Negative**
Predicted Class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False negative (FN)	True negative (TN)

5.2. Experimental Environment and Parameter Settings of LPG-Former

All experiments in this paper were completed under a unified hardware and software environment. The hardware platform used an Intel(R) Core(TM) i7-14650HX CPU and an NVIDIA GeForce RTX 4060 GPU with 16 GB of video memory. The software environment was the Windows 11 operating system, the programming language was Python 3.11, the deep-learning framework was PyTorch 2.2.2, and the CUDA version was 12.1. To ensure reproducibility of the experimental results, all experiments were trained and tested under the same data division, input dimensions, and preprocessing methods. The main parameter settings of the proposed LPG-Former are shown in Table 5.

Table 5. Specific parameter settings of the proposed LPG-Former.

Module	Parameter	Value
CDP	Prototype length	50
	Number of candidate CDPs($\times T$)	0.2
	Number of prototypes retained per category	20
LPT	Local feature embedding dimension	128
	Local position embedding dimension	128
GGT	Global feature embedding dimension	128
	Global position embedding dimension	128
LPG-Former	Number of attention heads	4
	Transformer embedding dimension	128
	Feedforward network dimension	256
	Dropout	0.2
	Optimizer	RAdam
	Training epochs	200
	Initial learning rate	1×10^{-4}

The prototype-related and embedding-related parameters in Table 5 were determined according to the robustness analysis in Section 5.6, while the other training parameters follow the unified experimental settings used for all methods.

5.3. Introduction to Comparison Algorithms

To verify the effectiveness of the proposed LPG-Former, this paper selects four typical time-series classification methods as comparison algorithms: 1D-CNN-LSTM [36], InceptionTime [37], TapNet [38], and SiameseNet [44]. Among them, 1D-CNN-LSTM and InceptionTime are deep feature-learning methods, mainly used to extract temporal features and dynamic dependencies. TapNet and SiameseNet are prototype-learning and metric-learning methods, focusing on enhancing intra-class aggregation and inter-class separation in the embedding space. LPG-Former is a collaborative modeling method for local prototypes and global generic features, improving recognition of fine-grained fault differences through a dual-branch structure.

To ensure fairness of comparison, all methods use the same balanced dataset, preprocessing method, and training parameters. During training, the random seed is set to 42, the number of training epochs is 200, the batch size is 32, the optimizer is RAdam, the learning rate is 1×10^{-4}, the weight-decay coefficient is 5×10^{-4}, and dropout is set to 0.2.

5.4. Comparative Experiment (1D-CNN-LSTM, TapNet, InceptionTime, SiameseNet, and LPG-Former)

5.4.1. Classification Performance of Multivariate Fault Time Series

The classification performance of the five comparison methods on multivariate fault time series is shown in Figure 7 and Table 6. It should be noted that the results in the table are the means and standard deviations of five independent repeated experiments rather than single test results. Because the test set contains 160 samples, this paper uses the average results of multiple repeated experiments to more stably reflect the classification performance of different models. In addition, the precision, recall, and F1-score in the table are the mean macro-average results over five independent experiments, and the three metrics are counted independently. Therefore, the F1-score may differ slightly from the result directly calculated from the average precision and average recall in the table.

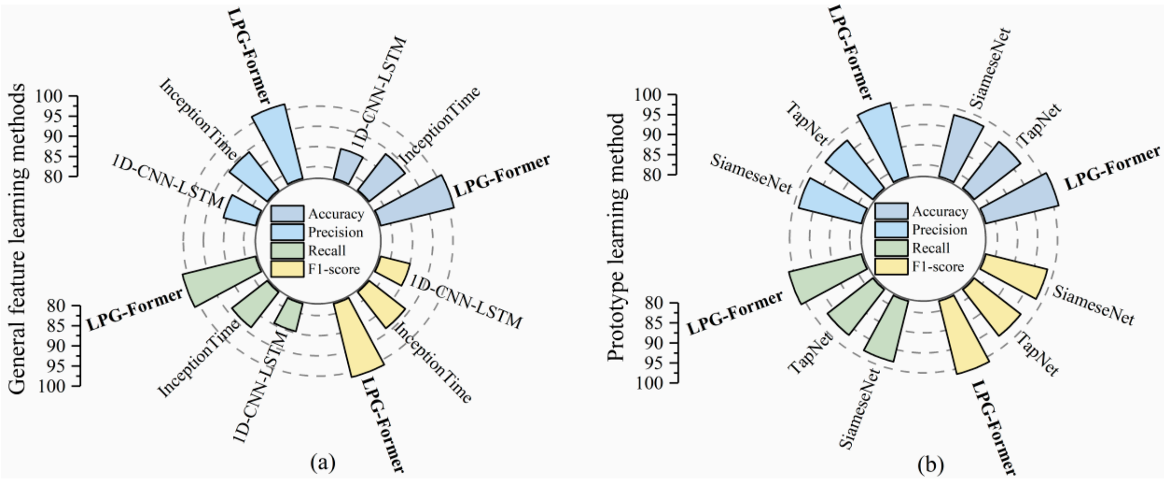

Figure 7. Classification performance of the comparison methods on multivariate fault time series: (**a**) general feature learning methods; (**b**) prototype-based learning methods.

Table 6. Specific classification results of the comparison methods on multivariate fault time series.

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
TapNet	94.67 ± 0.84	95.12 ± 0.78	94.43 ± 0.91	94.71 ± 0.86
InceptionTime	91.38 ± 1.17	91.92 ± 1.09	91.14 ± 1.26	91.41 ± 1.18
1D-CNN-LSTM	87.56 ± 1.64	88.21 ± 1.51	87.08 ± 1.73	87.49 ± 1.62
SiameseNet	91.83 ± 0.69	92.17 ± 0.62	92.61 ± 0.74	92.39 ± 0.67
LPG-Former	98.74 ± 0.31	98.96 ± 0.28	98.69 ± 0.35	98.78 ± 0.30

The specific data show that LPG-Former achieves the best results in all four metrics, namely accuracy, precision, recall, and F1-score, reaching 98.74%, 98.96%, 98.69%, and 98.78%, respectively. Compared with TapNet, LPG-Former improves accuracy, precision, recall, and F1-score by 4.07%, 3.84%, 4.26%, and 4.07%, respectively. Compared with 1D-CNN-LSTM, the improvements in the four metrics reach 11.18%, 10.75%, 11.61%, and 11.29%, respectively. Meanwhile, the standard deviations of all LPG-Former metrics are lower than those of the other comparison methods, indicating that it not only has higher classification accuracy but also better training stability.

Furthermore, the consistently low standard deviations across multiple independent runs indicate that the learned high-dimensional embedding space is stable under experimental variability. Although qualitative dimensionality-reduction results may vary with initialization and projection settings, the stable classification metrics show that LPG-Former reliably converges to a separable latent feature space. This consistency demonstrates the robust feature-representation capability of the proposed model and supports its diagnostic reliability under different training conditions.

From the perspective of performance differences, 1D-CNN-LSTM can combine convolutional and recurrent structures to extract local features and temporal dependencies, but its ability to characterize subtle inter-class differences and multivariate coupling relationships is limited; therefore, its overall performance is relatively low. InceptionTime enhances local morphological feature representation through multiscale convolution and performs better than 1D-CNN-LSTM, but it still mainly relies on generic convolutional features and lacks a local discriminative mechanism for specific fault classes. TapNet and SiameseNet introduce prototype-learning or metric-learning ideas and can enhance inter-class discriminability to some extent. However, these methods mainly measure sample similarity in the global embedding space and still insufficiently model fault-related local segments and variable-coupling variations explicitly. By contrast, LPG-Former extracts local fault prototypes through CDP and uses MCPM to quantify local morphological differences, variable-correlation differences, and trend-consistency differences between the CDP and BFS. At the same time, the LPT branch models correlations among local prototype matching differences, and the GGT branch supplements global generic features in the original time series, thereby achieving collaborative representation of local discriminative information and global temporal information.

Figure 8 shows the visualization results of the feature distributions learned by different methods. It can be seen that the features of some classes learned by the comparison methods overlap to varying degrees, indicating that they have difficulty forming clear inter-class boundaries when handling fine-grained EHA fault classes. Especially for classes with similar fault-response patterns, traditional generic feature-learning methods are easily affected by overall signal similarity, leading to insufficient inter-class separation in the feature space. By contrast, the feature distribution learned by LPG-Former is more compact, and the boundaries among different classes are clearer, indicating that the proposed method can effectively enhance intra-class compactness and inter-class separability.

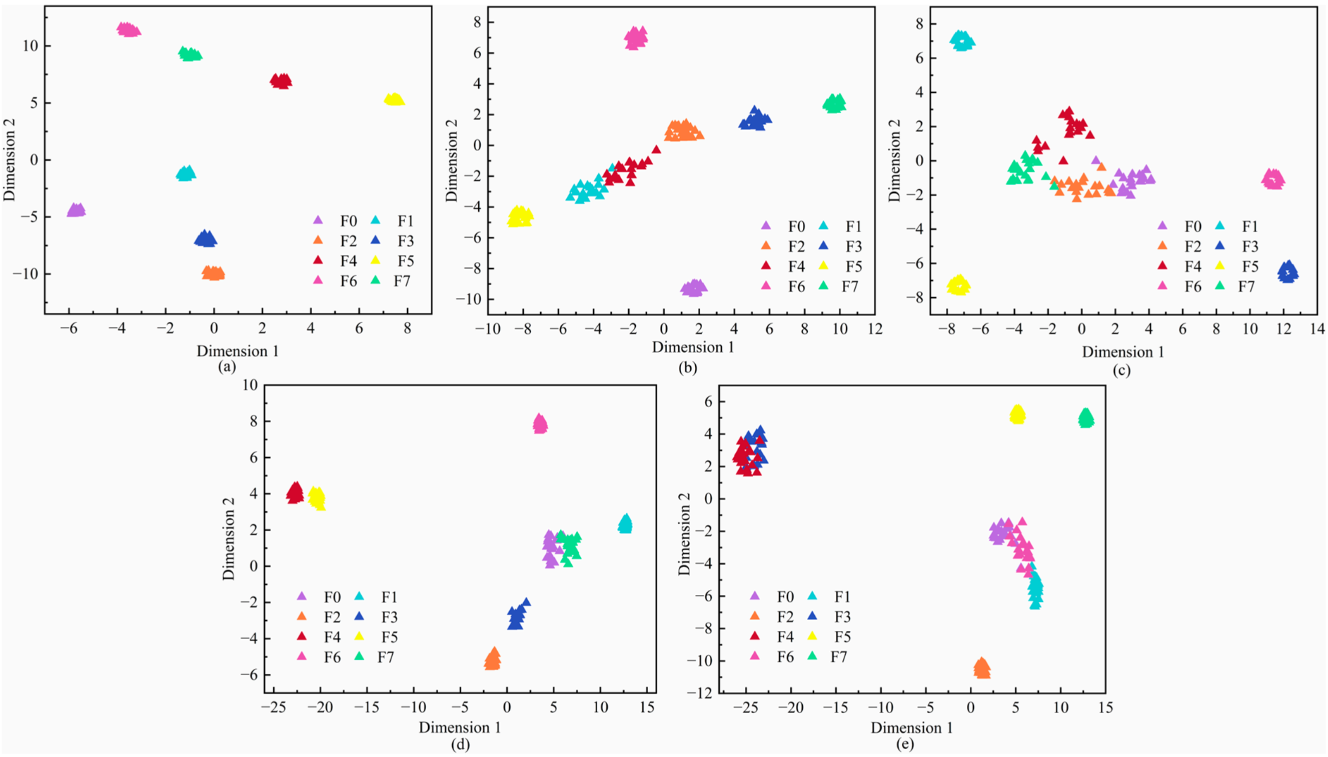

Figure 8. Visualization results of feature distributions for the comparison methods: (**a**) LPG-Former, (**b**) SiameseNet, (**c**) TapNet, (**d**) InceptionTime, (**e**) 1D-CNN-LSTM.

Figure 9 presents the confusion matrix results of different methods. It can be seen that 1D-CNN-LSTM, InceptionTime, TapNet, and SiameseNet have certain misclassification results among some classes, indicating that when relying only on generic deep features, global embedding features, or a single metric-learning mechanism, the ability of the model to distinguish similar fault states is still limited. By contrast, the main diagonal of the LPG-Former confusion matrix is more prominent, and the off-diagonal misclassification ratios are significantly reduced, indicating that it has good recognition consistency across the eight EHA operating states. Combined with Table 6, LPG-Former can effectively reduce the risk that fault samples are missed or misclassified as other classes, significantly improving the reliability of multivariate fault time-series classification.

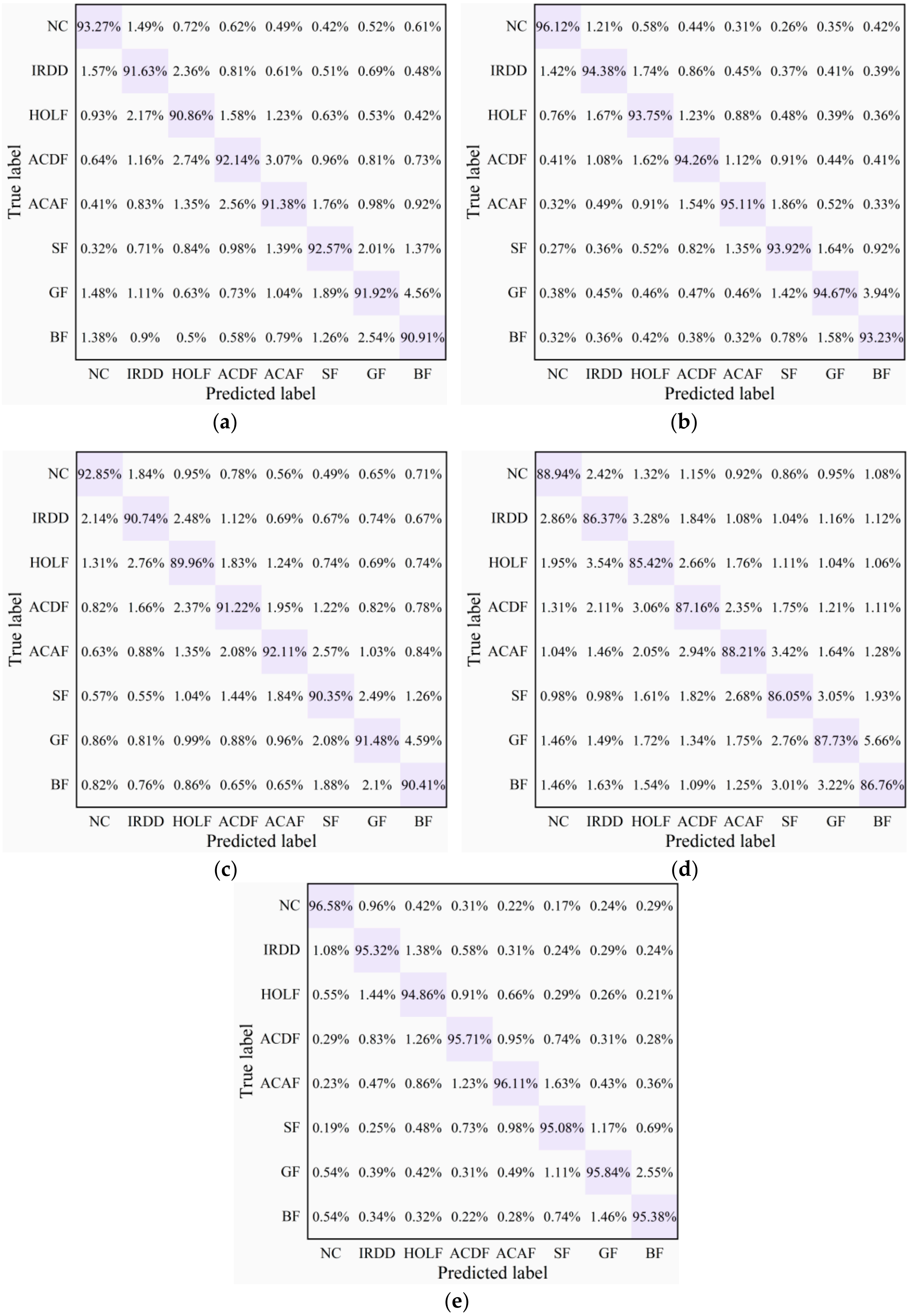

Figure 9. Confusion matrices of the comparison methods: (**a**) SiameseNet, (**b**) TapNet, (**c**) InceptionTime, (**d**) 1D-CNN-LSTM, (**e**) LPG-Former.

5.4.2. Computational-Complexity Analysis of Multivariate Fault Time-Series Classification

To evaluate the scalability of the proposed method for near-real-time condition monitoring, the computational complexity of the online inference phase is analyzed. The online diagnostic complexity is independent of the training dataset size, which mainly affects offline CDP extraction and model training. For the MCPM stage, let g, V, W, and L_r denote the number of prototypes, number of variables, restricted search-window size, and prototype length, respectively. The matching complexity is $\mathcal{O}(g \cdot V \cdot W \cdot L_r)$. As the position-aware search limits matching to a local window, the matching cost is decoupled from the total signal length. The Local Prototype Transformer processes g differential tokens with a self-attention complexity of $\mathcal{O}(g^2 \cdot d)$, where d is the embedding dimension. The Global Generic Transformer processes a downsampled generic sequence of length T$'$ with a complexity of $\mathcal{O}(T'^2 \cdot d)$. These bounds indicate that the online inference cost remains structurally manageable, and the corresponding empirical overhead is evaluated below.

To further evaluate the training and inference overheads of different methods in EHA multivariate fault time-series classification, this section tests each method from four aspects: classification accuracy, model size, training runtime memory, and single-sample inference time. The results are shown in Figure 10 and Table 7. Model size denotes the storage size occupied by the complete project file corresponding to the method and reflects the engineering storage overhead during model deployment. Runtime memory denotes the GPU memory occupied during model training. Single-sample inference time is the average time required for the model to infer one test sample after training is completed.

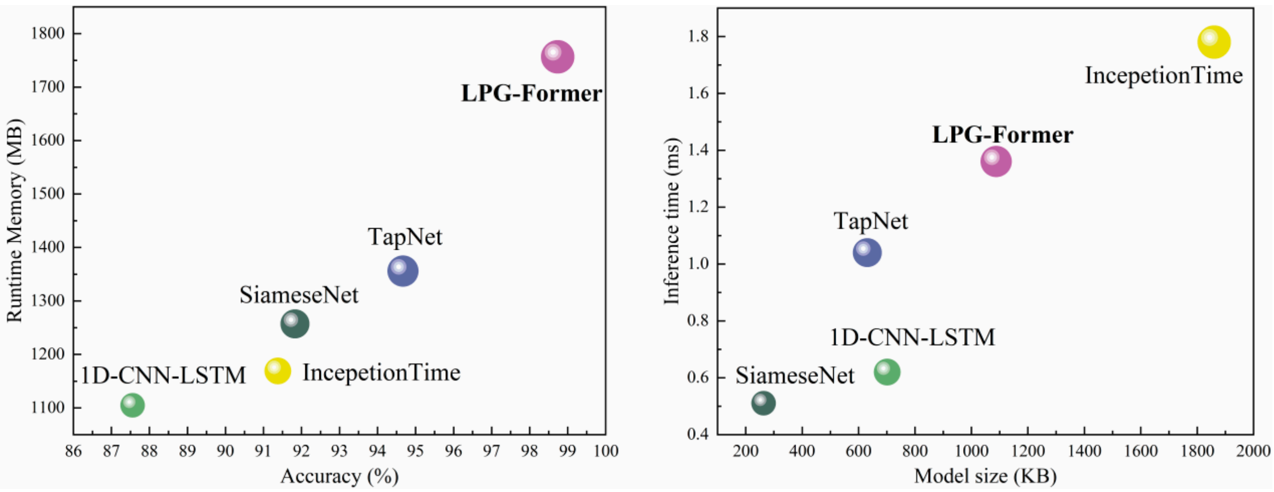

Figure 10. Accuracy, model size, and computational complexity of the comparison methods.

Table 7. Comparison of computational complexity among the comparison methods.

Method	Accuracy (%)	Model Size (KB)	Training Runtime Memory (MB)	Single Sample Inference Time (ms)
TapNet	94.67	630.98	1355.82	1.04
InceptionTime	91.38	1860.30	1168.97	1.78
1D-CNN-LSTM	87.56	701.33	1104.66	0.62
SiameseNet	91.83	263.27	1257.20	0.51
LPG-Former	98.74	1087.70	1756.58	1.36

As shown in Table 7, LPG-Former achieves the highest classification accuracy while maintaining relatively acceptable computational overhead. Specifically, the accuracy of LPG-Former reaches 98.74%, which is clearly higher than that of the other comparison methods. The model size of LPG-Former is 1087.70 KB, lower than the 1860.30 KB of InceptionTime and only higher than those of TapNet, 1D-CNN-LSTM, and SiameseNet. This indicates that although the proposed method introduces a dual-branch structure composed of a Local Prototype Transformer and a Global Generic Transformer, its engineering storage consumption remains within an acceptable range.

In terms of inference efficiency, the single-sample inference time of LPG-Former is 1.36 ms, slightly higher than those of TapNet, 1D-CNN-LSTM, and SiameseNet, but lower than the 1.78 ms of InceptionTime. The increase in inference time mainly comes from two aspects. First, the local prototype branch needs to encode CDP–BFS differential features to model dependencies among different discriminative prototypes. Second, the global generic branch needs to perform convolutional mapping and self-attention modeling on the original multivariate time series to supplement global information. Nevertheless, the inference time of LPG-Former remains at the millisecond level, indicating that it does not introduce excessive inference overhead while improving classification performance.

In terms of training runtime memory, the training runtime memory of LPG-Former is 1756.58 MB, higher than those of the other comparison methods. This is mainly because the dual-branch structure needs to store local prototype matching features and global generic features simultaneously, and retain intermediate representations required for attention calculation during Transformer encoding. By contrast, SiameseNet has the smallest model size and shortest inference time, but its accuracy is only 91.83%, indicating that simply pursuing lightweight design weakens the ability to represent complex multivariate fault features.

5.5. Ablation Experiments

5.5.1. Ablation Experiment on Discriminative Prototypes

To verify the effectiveness of CDP and its multivariate extension mechanism, this paper designs three groups of discriminative-prototype ablation experiments, as shown in Table 8. The three methods all use the same dual-branch Transformer architecture, with the difference lying in the input discriminative prototypes. Ablation method 1 uses the multivariate CDP proposed in this paper. Ablation method 2 uses univariate CDP to verify the contribution of multivariate coupling information to prototype discriminability. Ablation method 3 uses generic sequence prototypes to verify the necessity of class-specific local discriminative prototypes.

Table 8. Ablation experiment settings for the discriminative prototype module.

Ablation Method	Composition Mechanism	Purpose
Ablation method 1	Multivariable CDP + LPT + GGT	To verify the overall performance of the complete LPG-Former.
Ablation method 2	Single-variable CDP + LPT + GGT	To evaluate the influence of multivariate coupling information on prototype discriminability.
Ablation method 3	General prototype + LPT + GGT	To verify the necessity of class-specific local discriminative prototypes.

As shown in Figure 11 and Table 9, different prototype-construction methods have a significant influence on model classification performance. Ablation method 1 using multivariate CDP achieves the best results, with accuracy, precision, recall, and F1-score reaching 98.74%, 98.96%, 98.69%, and 98.78%, respectively. In comparison, ablation method 2 using only univariate CDP decreases to 93.42%, 93.73%, 93.21%, and 93.39% in the four metrics. The performance of ablation method 3 using generic sequence prototypes further decreases, with accuracy and F1-score of only 86.36% and 86.31%, respectively. This indicates that whether the prototype is class-specific and whether it contains multivariate coupling information are key factors affecting the fault-recognition performance of the model.

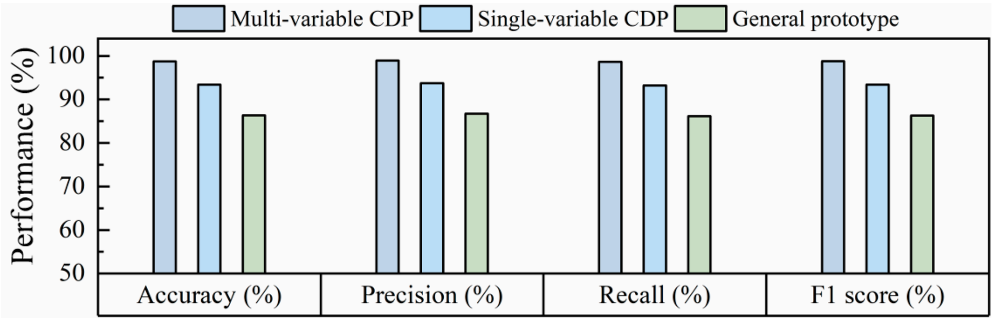

Figure 11. Influence of different prototype input strategies on the classification performance of LPG-Former.

Table 9. Ablation experimental results under different prototype inputs.

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Ablation 1	98.74 ± 0.31	98.96 ± 0.28	98.69 ± 0.35	98.78 ± 0.30
Ablation 2	93.42 ± 0.58	93.73 ± 0.62	93.21 ± 0.65	93.39 ± 0.60
Ablation 3	86.36 ± 0.84	86.75 ± 0.79	86.14 ± 0.86	86.31 ± 0.82

The comparison between ablation method 2 and ablation method 1 shows that after extending univariate CDP to multivariate CDP, accuracy increases by 5.32%. The main reason is that EHA fault states are not only reflected in local fluctuations of a single variable, but are also jointly manifested through coupling relationships and trend-consistency variations among variables such as current, displacement, and pressure. Although univariate CDP can capture local morphological differences, it easily ignores inter-variable linkage features during fault propagation. Further comparison with ablation method 3 shows that the performance of generic sequence prototypes is significantly lower than that of class-specific CDP. This indicates that relying only on global or class-irrelevant generic segments makes it difficult to effectively distinguish operating states in EHA that have similar overall morphology but different local fault responses. By contrast, multivariate CDP enables the prototype-matching process to more fully represent multivariate fault patterns through joint measurement of local morphological differences, variable correlation, and trend consistency, thus achieving the highest classification performance.

5.5.2. Ablation Experiment on the Dual-Branch Architecture

To verify the complementary roles of LPT and GGT in the proposed method, this section further conducts ablation experiments on the dual-branch architecture. The settings of the three ablation methods are shown in Table 10. Ablation method 4 retains only the LPT branch and uses the differential features between multivariate CDP and BFS as input to verify the effectiveness of locally discriminative prototype features. Ablation method 5 retains only the GGT branch and uses the original multivariate time series as input to verify the effect of global generic temporal features. Ablation method 6 uses both LPT and GGT, with local prototype differential features and global generic features as input, to verify the effectiveness of collaborative modeling of local discriminative information and global generic information.

Table 10. Design of ablation methods for the dual-branch Transformer architecture.

Ablation Method	Input Data	Architecture	Purpose
Ablation method 4	Differences between multivariate CDP and BFS.	Local Prototype Transformer	To verify the effectiveness of local discriminative prototype features.
Ablation method 5	Original multivariate time series.	Global Generic Transformer	To verify the effectiveness of global generic temporal features.
Ablation method 6	Differences between multivariate CDP and BFS + original multivariate time series.	LPG-Former	To verify the complementarity of local prototype features and global generic features.

As shown in Figure 12 and Table 11, when only the LPT branch is used, the model has an FNR of 7.45%, runtime memory of 1226.44 MB, and single-sample inference time of 0.94 ms. When only the GGT branch is used, the FNR increases to 13.59%, while runtime memory and inference time decrease to 918.73 MB and 0.72 ms, respectively. By contrast, the complete LPG-Former reduces the FNR to 1.31%, which is 6.14% and 12.28% lower than using only LPT and only GGT, respectively, indicating that the local prototype branch and the global generic branch have significant complementarity. In Figure 11, the circular marker corresponds to the left vertical axis, and the triangular marking corresponds to the right vertical axis.

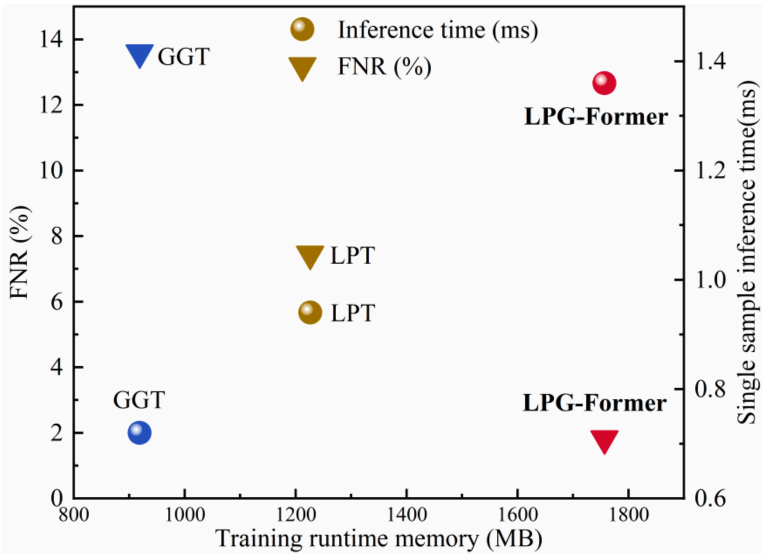

Figure 12. Comparison of inference overheads among LPT, GGT, and the LPG-Former dual-branch structure.

Table 11. Comparison of classification performance, training memory, and inference time among LPT, GGT, and LPG-Former.

Method	Architecture	FNR (%)	Training Runtime Memory (MB)	Single Sample Inference Time (ms)
Ablation 4	LPT	7.45	1226.44	0.94
Ablation 5	GGT	13.59	918.73	0.72
Ablation 6	LPT + GGT	1.31	1756.58	1.36

From the perspective of branch mechanisms, GGT focuses on extracting global temporal patterns and overall variable correlations from original multivariate sequences, but it insufficiently characterizes class-specific local fault segments and therefore easily leads to missed detections among similar fault classes. LPT uses CDP–BFS differential features to explicitly model the differences between input samples and typical fault local prototypes, enhancing the recognition capability for local morphological changes and variable-coupling anomalies. However, because LPT is highly specific, using it alone may ignore global information and increase the risk of overfitting. Therefore, the complete LPG-Former adopts a dual-branch architecture of LPT and GGT, improving fine-grained fault-identification capability while ensuring overall robustness, thereby achieving a lower missed-detection rate and better comprehensive performance.

In terms of computational overhead, although the dual-branch structure increases training runtime memory to 1756.58 MB and single-sample inference time to 1.36 ms, this overhead is not a simple superposition of the two single-branch models and remains within the millisecond-level inference range. Overall, LPG-Former achieves a significantly reduced missed-detection rate with moderate computational overhead, and the proposed dual-branch architecture significantly improves the accuracy and reliability of EHA multivariate fault time-series classification.

The significance of these architectural improvements is reflected not only in absolute accuracy but also in missed-detection reduction. The performance degradation observed in the ablation studies is larger than the random fluctuation range of the complete model reported in the baseline comparisons, indicating that the degradation is mainly caused by structural removal rather than stochastic training variation. More importantly, the false negative rate decreases from 13.59% with only the global generic branch and 7.45% with only the local prototype branch to 1.31% with the complete dual-branch architecture. This substantial reduction in missed-detection risk demonstrates that the fusion of local and global representations is critical for reliable EHA fault diagnosis in safety-critical applications.

5.6. Parameter Robustness Experiments

To analyze the influence of key parameters on the classification performance and computational overhead of LPG-Former, this paper conducts parameter robustness experiments around the candidate CDP extraction ratio, the final number of CDPs retained for each class, and the feature-embedding dimension. The experimental results are shown in Figure 13 and Tables 12–14. The candidate CDP extraction ratio controls the generation scale of candidate local prototypes and affects the coverage of key local fault segments. The final number of CDPs retained for each class controls the size of the CDP prototype pool and determines the amount of class-related local discriminative information available to the model. The feature-embedding dimension controls the representation capacity of prototype matching differences and global temporal features, affecting the feature-expression capability and computational overhead of the model.

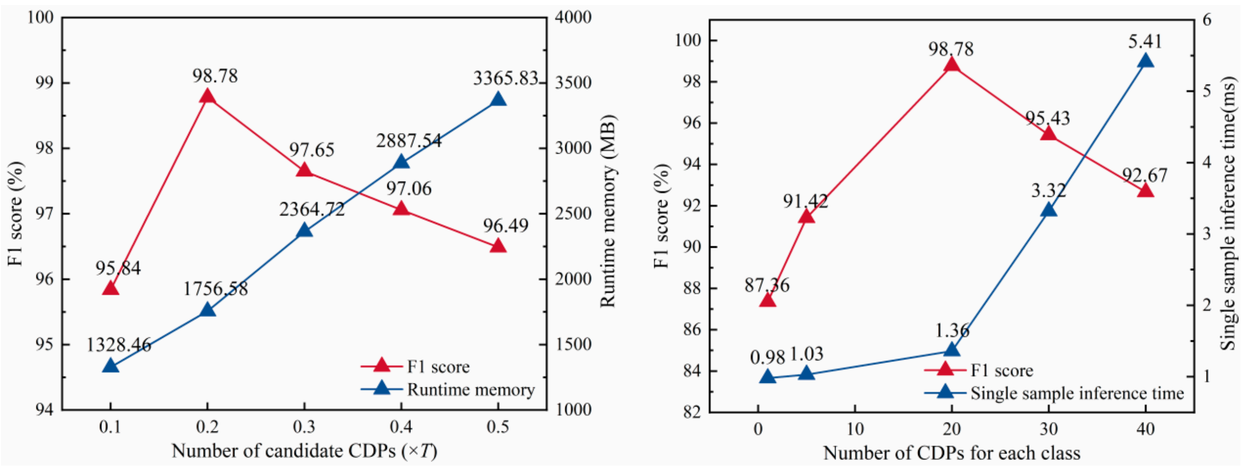

Figure 13. Influence of the number of CDPs on model performance and computational overhead.

Table 12. Influence of the candidate CDP extraction ratio on model performance and runtime memory.

Number of Candidate CDPs($\times T$)	F1-Score (%)	FNR (%)	Training Runtime Memory (MB)
0.1	95.84	4.78	1328.46
0.2	**98.78**	**1.31**	**1756.58**
0.3	97.65	2.06	2364.72
0.4	97.06	2.31	2887.54
0.5	96.49	3.12	3365.83

Table 13. Influence of the final number of retained CDPs for each class on model performance and inference time.

Number of CDPs for Each Class	F1-Score (%)	FNR (%)	Single Sample Inference Time (ms)
1	87.36	12.84	0.98
5	91.42	8.71	1.03
20	**98.78**	**1.31**	**1.36**
30	95.43	4.18	3.32
40	92.67	6.93	5.41

Table 14. Influence of the feature-embedding dimension on the classification performance of LPG-Former.

Local Embedding Dimension/ Global Embedding Dimension	Accuracy (%)			
	32	**64**	**128**	**256**
32	94.82	95.43	96.04	95.31
64	95.68	96.57	97.21	96.48
128	96.94	97.83	**98.74**	97.56
256	96.21	97.10	97.42	96.63

As shown in Figure 13 and Table 12, the candidate CDP extraction ratio has a direct influence on both model performance and training runtime memory. When the candidate CDP extraction ratio is 0.1 T (where T is the input sequence length), the F1-score is 95.84% and the FNR is 4.78%, indicating that too few candidate prototypes cannot sufficiently cover local fault morphologies in different classes. When the extraction ratio increases to 0.2 T, the F1-score reaches 98.78% and the FNR decreases to 1.31%, indicating that an appropriate increase in the number of candidate CDPs can provide a more sufficient candidate space for subsequent information-gain screening, thereby enhancing the discriminative capability of the final prototype pool. However, when the extraction ratio continues to increase to 0.5 T, the F1-score decreases to 96.49%, the FNR increases to 3.12%, and runtime memory increases from 1756.58 MB to 3365.83 MB. The reason is that excessive candidate CDPs introduce redundant segments and weakly discriminative segments, which not only significantly increases computational overhead but may also interfere with the model's learning of key local fault patterns.

Table 13 further shows the influence of the final number of CDPs retained for each class on model performance and inference time. It can be seen that when only one or five CDPs are retained for each class, the F1-scores are only 87.36% and 91.42%, while the FNRs reach 12.84% and 8.71%, respectively. This indicates that too few CDPs are insufficient to represent diverse local patterns within the same fault class. When 20 CDPs are retained for each class, the model obtains the optimal F1-score and the lowest FNR, and the single-sample inference time is only 1.36 ms, achieving a good balance between classification performance and inference efficiency. If the number is further increased to 30 or 40 CDPs, although the prototype coverage range expands, the F1-score instead decreases to 95.43% and 92.67%, and the inference time increases significantly to 3.32 ms and 5.41 ms. This indicates that excessive CDPs lead to prototype-space redundancy and overfitting and significantly increase the computational overhead of BFS search.

Table 14 presents the classification results of LPG-Former under different combinations of feature-embedding dimensions. Overall, as the embedding dimension increases from 32 to 128, model accuracy shows an upward trend, indicating that a higher-dimensional feature space helps improve the representation capability of CDP–BFS differential features and global temporal features. When both local and global feature-embedding dimensions are 128, the model achieves the highest accuracy. However, when the dimension continues to increase to 256, model performance does not improve further, and accuracy even decreases under some combinations. This is mainly because the EHA fault data scale is limited, and excessively high dimensions increase the number of model parameters and optimization difficulty, easily introducing redundant features.

The above results demonstrate that LPG-Former has good robustness to key parameters. Within a reasonable parameter range, the model consistently maintains high classification performance. This paper finally selects a candidate CDP extraction ratio of 0.2 T, 20 CDPs retained for each class, and a feature-embedding dimension of 128. This setting ensures sufficient expression of local discriminative information while effectively controlling runtime memory and single-sample inference time, thereby balancing classification accuracy, missed-detection control capability, and practical deployment efficiency.

5.7. Feature-Distribution Visualization Analysis Under Class-Imbalanced Scenarios

To verify the feature-representation capability of the proposed method under class-imbalanced scenarios, this paper further conducts experiments using the imbalanced dataset. In this setting, the validation and test sets maintain the same number of samples for each class, and only the number of samples in different classes in the training set is adjusted to simulate the shortage of samples for some classes in practical EHA fault diagnosis. The class with the largest number of training samples is F1, with 102 samples. The class with the smallest number of training samples is F2, with only 36 samples, showing an obvious imbalance in sample distribution.

Figure 14 presents the visualization results of the feature distributions of InceptionTime and LPG-Former under the class-imbalanced scenario. It can be seen that in the feature space learned by InceptionTime, some classes are still obviously close to or even overlapping with each other. Especially when the number of minority-class samples is insufficient, the model is more easily dominated by majority-class samples, resulting in unclear minority-class feature boundaries. This indicates that when relying only on multiscale convolution to extract generic temporal features, the model still insufficiently characterizes locally discriminative fault patterns under class-imbalance conditions and has difficulty stably forming feature representations with good inter-class separability.

By contrast, LPG-Former can form more compact and clearly separated feature clusters under imbalanced data conditions. Even for classes with fewer training samples, such as F2, F3, and F6, their feature distributions still maintain clear gaps from those of other classes, indicating that the proposed method has better robustness in minority-class recognition. The reason is that LPG-Former does not rely solely on the number of original samples to learn generic features. Instead, it extracts representative and discriminative local fault segments from each class through class-specific discriminative prototypes, and retains the same number of highly discriminative CDPs for different classes during prototype screening, thereby mitigating the dominance of majority-class samples over the feature space to some extent. In addition, LPG-Former performs joint modeling through CDP–BFS differential features, so even when the number of minority-class samples is small, it can form stable representations based on class-specific local patterns.

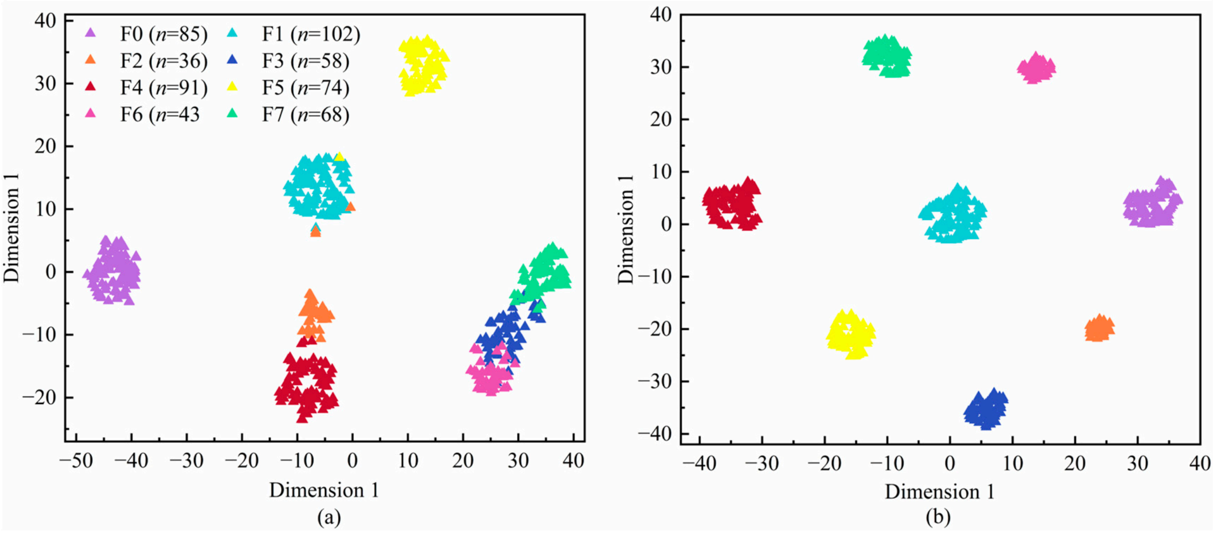

Figure 14. Feature-distribution comparison under class-imbalanced scenarios: (**a**) InceptionTime, (**b**) LPG-Former.

5.8. Verification Experiment on Discriminative Capability

To further verify the class-discriminative capability of class-specific discriminative prototypes, this paper designs a visualization experiment based on attention correlation. Specifically, one sample is randomly selected from the test set, and five CDPs with the highest information gains are selected from each of the eight operating states, resulting in 40 discriminative prototypes. To clarify the class source of each CDP, the CDPs are first initially numbered in class order: class 0 corresponds to I1-I5, class 1 corresponds to I6-I10, and so on, with class 7 corresponding to I36-I40. Subsequently, to avoid prior influence of the numbering order on visualization analysis, the 40 CDPs are randomly encoded as R1-R40 and input into the Local Prototype Transformer to extract the attention weight matrix for visualization analysis.

It should be noted that the class groups in this experiment are defined by operating-state labels rather than signal-channel types. Information gain is used only as a within-class screening criterion to select representative CDPs, whereas the attention weights characterize the relationships among their encoded representations. Therefore, the attention responses should not be interpreted as signal-level feature importance or as a global ranking of CDP information gains.

Figure 15 presents the visualization results of CDP attention correlation. Figure 15a shows the attention heat map under the randomly encoded R1-R40 state. It can be seen that even under random ordering, some CDPs show high attention responses, indicating that these prototypes have strong correlations in the local prototype feature space. Further back-tracking according to the mapping relationship between random codes and initial numbers reveals that most CDPs with high attention responses originate from the same fault class, indicating that same-class CDPs have high consistency in feature expression. Figure 15b further shows the attention heat map rearranged according to the initial numbering I1-I40. It can be observed that high-response regions are mainly concentrated within the class blocks composed of every five CDPs, whereas attention responses among different class blocks are generally weak.

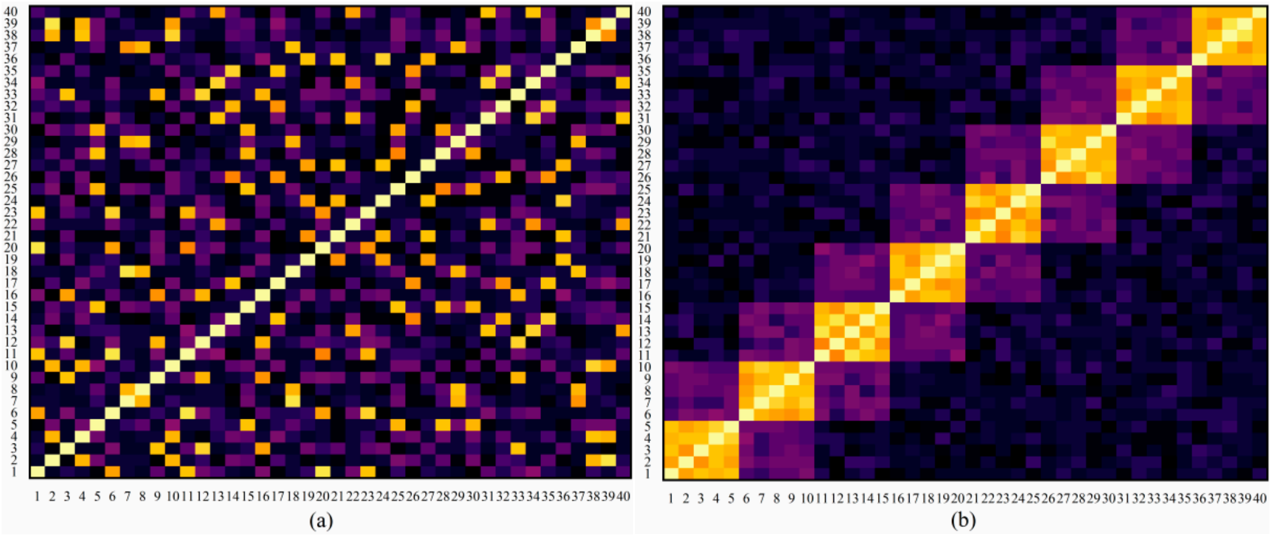

Figure 15. Attention-response heat maps of CDPs: (**a**) random numbering, (**b**) initial numbering.

The above results demonstrate from the perspective of attention response that the constructed CDPs are not random local segments, but local discriminative prototypes with strong class relevance. CDPs within the same class can form strong correlations in the attention space, indicating that they can jointly represent typical local patterns of that fault class. The weak attention responses among CDPs of different classes indicate that the prototypes of each class have clear differences in local morphology, variable-coupling relationships, and trend variations. Therefore, the proposed CDP can effectively enhance the consistency of same-class fault prototypes and the separability of different-class fault prototypes, providing discriminative input features for the Local Prototype Transformer.

To further relate the extracted prototypes to EHA fault physics, the multivariate CDPs can be interpreted as compact representations of electro-mechanical-hydraulic propagation patterns rather than purely statistical segments. For the hydraulic oil leakage fault, leakage causes pressure imbalance in the actuator chambers, and the closed-loop controller increases motor torque/current to compensate for the pressure loss. Thus, the prototype jointly captures pressure deviation and current compensation. For the actuator-cylinder deformation fault, cylinder deformation increases mechanical friction, leading to displacement tracking delay and a transient current rise. These mappings indicate that the high attention responses in Figure 15 correspond to physically meaningful coupled fault signatures.

6. Conclusions

To address the problems that the overall inter-class morphology in EHA multivariate fault time series is similar and locally discriminative fault segments are difficult to extract effectively, this paper proposes a Local Prototype-Global Generic Dual-branch Transformer. The method first constructs a class-specific discriminative prototype to extract local fault segments with class representativeness and discriminability. A multivariate coupling-aware prototype matching strategy is then designed to quantify the differences between CDPs and input samples from three aspects: local morphology, variable correlation, and trend consistency. Finally, collaborative modeling through the Local Prototype Transformer and Global Generic Transformer realizes the fusion of local discriminative information and global generic features, thereby achieving accurate diagnosis of fine-grained fault classes in EHA multivariate fault time series. The conclusions can be summarized as follows:

(1). LPG-Former achieves the best diagnostic performance in the eight-class EHA multivariate fault time-series recognition task, with accuracy, precision, recall, and F1-score reaching 98.74%, 98.96%, 98.69%, and 98.78%, respectively.
(2). The ablation experiments show that the proposed CDP and MCPM are key to improving the overall performance of the model. Compared with univariate CDP and generic sequence prototypes, multivariate CDP can more fully characterize local morphological variations, variable correlation, and trend consistency, thereby enhancing the model's discriminative capability for fine-grained fault features.
(3). The dual-branch structure further verifies the complementarity between local discriminative information and global generic information. The FNR of the complete LPG-Former decreases to 1.31%, which is clearly lower than the structures using LPT or GGT alone. Meanwhile, under reasonable parameter settings, the single-sample inference time of the model can be maintained at 1.36 ms.

Future work will further investigate incremental updating and lightweight deployment methods for the CDP prototype pool to improve the real-time performance and deployment feasibility of the model in online fault-monitoring scenarios.

Author Contributions: Conceptualization, G.Z. and Z.J.; methodology, G.Z. and Z.J.; software, K.Z.; validation, K.Z. and Z.G.; formal analysis, G.Z. and Z.J.; investigation, W.Y.; resources, Z.L.; data curation, Z.J.; writing—original draft preparation, G.Z. and Z.G.; writing—review and editing, G.Z. and K.Z.; visualization, Z.G.; supervision, Z.L.; project administration, Z.J. and Z.L.; funding acquisition, Z.J. and Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Key Program of the National Natural Science Foundation of China under Grant U2433208, in part by the Aeronautical Science Foundation of China under Grant ASFC-2018ZC53026, in part by the Postdoctoral Fellowship Program of CPSF under Grant GZC20252711, and in part by the Shaanxi Province Postdoctoral Research Funding Program under Grant 2025BSHSDZZ110.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data will be made available upon request.

Acknowledgments: The authors sincerely appreciate the contributions of all authors and the support of funding agencies.

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

References

1. Nie, S.L.; Gao, J.H.; Ma, Z.H.; Yin, F.L.; Ji, H. An online data-driven approach for performance prediction of electro-hydrostatic actuator with thermal-hydraulic modeling. *Reliab. Eng. Syst. Saf.* **2023**, *236*, 109289. [CrossRef]
2. Sun, M.W.; Fu, J.; Wang, C.; Wu, H.W.; Zhao, Z.R.; Wang, W.S.; Bai, L.L. System-level modeling and preliminary experiment study of a redundant EHA for more electric helicopter flight control. *Chin. J. Aeronaut.* **2025**, 103891. [CrossRef]
3. Zhang, J.H.; Li, X.G.; Lyu, F.; Schlegel, F.; Schmitz, K.; Liu, S.H.; Huang, W.D.; Xu, B. Critical operation range of electro-hydrostatic actuator pumps based on cylinder block tilting states. *Chin. J. Aeronaut.* **2025**, *38*, 103382. [CrossRef]
4. Xu, J.Q.; Jin, W.B.; Guo, H.; Yu, T.; Fan, W.H.; Jiao, Z.X. Modeling and analysis of oil frictional loss in wet-type permanent magnet synchronous motor for aerospace electro-hydrostatic actuator. *Chin. J. Aeronaut.* **2023**, *36*, 328–341. [CrossRef]
5. Fan, S.M.; Wang, S.P.; Wang, Q.Y.; Wang, X.J.; Liu, D.; Wu, X. Cumulative thermal coupling modeling and analysis of oil-immersed motor-pump assembly for electro-hydrostatic actuator. *Chin. J. Aeronaut.* **2025**, *38*, 103250. [CrossRef]
6. Ma, Z.H.; Liao, H.T.; Gao, J.H.; Nie, S.L.; Geng, Y.G. Physics-informed machine learning for degradation modeling of an electro-hydrostatic actuator system. *Reliab. Eng. Syst. Saf.* **2023**, *229*, 108898. [CrossRef]
7. Guo, Y.Q.; He, X. Active diagnosis of incipient actuator faults for stochastic systems. *IEEE Trans. Ind. Electron.* **2024**, *71*, 996–1005. [CrossRef]

8. Jia, F.L.; Cao, F.F.; He, X. Active fault-tolerant control against intermittent faults for state-constrained nonlinear systems. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 2389–2401. [CrossRef]
9. Shen, W.; Zhao, H.M. Fault tolerant control of nonlinear hydraulic systems with prescribed performance constraint. *ISA Trans.* **2022**, *131*, 1–14. [CrossRef] [PubMed]
10. Xu, B.; Huang, X.C.; Zhang, J.H.; Huang, W.D.; Lyu, F.; Xu, H.G. A fault detection method for a practical electro-hydraulic variable-displacement pump with unknown swashplate moment. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3513511. [CrossRef]
11. Saeedzadeh, A.; Habibi, S.; Alavi, M.; Setoodeh, P. A robust model-based strategy for real-time fault detection and diagnosis in an electro-hydraulic actuator using updated interactive multiple model smooth variable structure filter. *J. Dyn. Syst. Meas. Control* **2023**, *145*, 101006. [CrossRef]
12. Nguyen, M.H.; Dao, H.V.; Ahn, K.K. Extended sliding mode observer-based high-accuracy motion control for uncertain electro-hydraulic systems. *Int. J. Robust Nonlinear Control* **2023**, *33*, 1351–1370.
13. Gao, Y.C.; Shen, G.; Li, X.; Zhu, Z.C.; Wang, Q.G. An improved nonlinear extended disturbance observer for sliding mode control in electro-hydraulic servo system. *IEEE/ASME Trans. Mechatron.* **2025**, *30*, 2155–2164. [CrossRef]
14. Qian, Q.; Qin, Y.; Luo, J.; Wang, S.L. Partial transfer fault diagnosis by multiscale weight-selection adversarial network. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 4798–4806. [CrossRef]
15. Zhao, S.; Duan, Y.; Roy, N.; Zhang, B. A deep learning methodology based on adaptive multiscale CNN and enhanced highway LSTM for industrial process fault diagnosis. *Reliab. Eng. Syst. Saf.* **2024**, *249*, 110208. [CrossRef]
16. Yan, X.S.; Shi, Z.G.; Sun, Z.; Zhang, C.A. Multisensor fusion on hypergraph for fault diagnosis. *IEEE Trans. Ind. Inform.* **2024**, *20*, 10008–10018. [CrossRef]
17. Shen, C.Q.; Tian, J.; Zhu, J.; Shi, J.; Zhu, Z.K.; Wang, D. A new multisource domain bearing fault diagnosis method with adaptive dual-domain obfuscation weighting strategy. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3509311. [CrossRef]
18. Wang, Z.Y.; Liu, T.; Wu, X.; Liu, C. Application of an oversampling method based on GMM and boundary optimization in imbalance-bearing fault diagnosis. *IEEE Trans. Ind. Inform.* **2024**, *20*, 1931–1940. [CrossRef]
19. Zhao, Z.Q.; Jiao, Y.H. A fault diagnosis method for rotating machinery based on CNN with mixed information. *IEEE Trans. Ind. Inform.* **2023**, *19*, 9091–9101. [CrossRef]
20. Li, H.; Liu, T.; Wu, X.; Li, S.B. Correlated SVD and its application in bearing fault diagnosis. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 355–365. [CrossRef] [PubMed]
21. Li, T.F.; Zhou, Z.; Li, S.N.; Sun, C.; Yan, R.Q.; Chen, X.F. The emerging graph neural networks for intelligent fault diagnostics and prognostics: A guideline and a benchmark study. *Mech. Syst. Signal Process.* **2022**, *168*, 108653. [CrossRef]
22. Lv, H.X.; Chen, J.L.; Pan, T.Y.; Zhang, T.C.; Feng, Y.; Liu, S. Attention mechanism in intelligent fault diagnosis of machinery: A review of technique and application. *Measurement* **2022**, *199*, 111594. [CrossRef]
23. Zhang, T.C.; Chen, J.L.; Li, F.D.; Zhang, K.Y.; Lv, H.X.; He, S.L.; Xu, E.Y. Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions. *ISA Trans.* **2022**, *119*, 152–171. [CrossRef] [PubMed]
24. Hu, K.; He, Q.B.; Xu, H.; Cheng, C.M.; Peng, Z.K. Dynamic domain adaptive ensemble for intelligent fault diagnosis of machinery. *Knowl.-Based Syst.* **2025**, *314*, 113209. [CrossRef]
25. Liu, Z.Y.; Li, C.; He, X. Evidential ensemble preference-guided learning approach for real-time multimode fault diagnosis. *IEEE Trans. Ind. Inform.* **2024**, *20*, 5495–5504. [CrossRef]
26. Tong, J.Y.; Liu, C.; Bao, J.H.; Pan, H.Y.; Zheng, J.D. A novel ensemble learning-based multisensor information fusion method for rolling bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 9501712. [CrossRef]
27. Jang, K.; Pilario, K.E.S.; Lee, N.; Moon, I.; Na, J. Explainable artificial intelligence for fault diagnosis of industrial processes. *IEEE Trans. Ind. Inform.* **2025**, *21*, 4–11. [CrossRef]
28. Gwak, M.; Kim, M.S.; Yun, J.P.; Park, P.G. Robust and explainable fault diagnosis with power-perturbation-based decision boundary analysis of deep learning models. *IEEE Trans. Ind. Inform.* **2023**, *19*, 6982–6992. [CrossRef]
29. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. [CrossRef] [PubMed]
30. Barredo Arrieta, A.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
31. Kim, M.S.; Yun, J.P.; Park, P.G. Deep learning-based explainable fault diagnosis model with an individually grouped 1-D convolution for three-axis vibration signals. *IEEE Trans. Ind. Inform.* **2022**, *18*, 8807–8817. [CrossRef]
32. Wang, H.; Liu, Z.L.; Peng, D.D.; Qin, Y. Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5735–5745. [CrossRef]
33. Chu, C.; Shen, J.; Xiao, Q. Graph-Regularized Geometric Deep Learning for Motor Imagery EEG Decoding via Multi-Domain Fusion. *Pattern Recognit.* **2026**, *174*, 112976. [CrossRef]

34. Liu, W.; Pan, S.; Chang, S.; Huang, Q.; Jiang, N. TriFusNet: A Triple-Fusion Convolutional Network for Multivariate Time Series Classification. *IEEE Trans. Knowl. Data Eng.* **2026**, *38*, 4006–4019. [CrossRef]
35. Liu, X.; Fang, Z.; Lei, F.; Kang, H.; Han, H.; Shen, Y.; Dong, H.; Fu, X. A Novel Method to Study the Long Period Three-Dimensional Vibration Characteristics of Herringbone Gear with Asymmetry Pitch Deviation. *Mech. Syst. Signal Process.* **2025**, *224*, 112040. [CrossRef]
36. Alsumaidaee, Y.A.M.; Paw, J.K.S.; Yaw, C.T.; Tiong, S.K.; Chen, C.P.; Yusaf, T.; Benedict, F.; Kadirgama, K.; Hong, T.C.; Abdalla, A.N. Fault Detection for Medium Voltage Switchgear Using a Deep Learning Hybrid 1D-CNN-LSTM Model. *IEEE Access* **2023**, *11*, 97574–97589. [CrossRef]
37. Fawaz, H.I.; Lucas, B.; Forestier, G.; Pelletier, C.; Schmidt, D.F.; Weber, J.; Webb, G.I.; Idoumghar, L.; Muller, P.-A.; Petitjean, F. InceptionTime: Finding AlexNet for Time Series Classification. *Data Min. Knowl. Discov.* **2020**, *34*, 1936–1962. [CrossRef]
38. Zhang, X.; Gao, Y.; Lin, J.; Lu, C.-T. TapNet: Multivariate Time Series Classification with Attentional Prototypical Network. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 6845–6852. [CrossRef]
39. Foumani, N.M.; Tan, C.W.; Webb, G.I.; Salehi, M. Improving Position Encoding of Transformers for Multivariate Time Series Classification. *Data Min. Knowl. Discov.* **2024**, *38*, 22–48. [CrossRef]
40. Zuo, R.; Li, G.; Choi, B.; Bhowmick, S.S.; Mah, D.N.Y.; Wong, G.L. SVP-T: A Shape-Level Variable-Position Transformer for Multivariate Time Series Classification. *Proc. AAAI Conf. Artif. Intell.* **2023**, *37*, 11497–11505. [CrossRef]
41. Wen, J.; Zhang, N.; Lu, X.; Hu, Z.; Huang, H. Mgformer: Multi-Group Transformer for Multivariate Time Series Classification. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108633. [CrossRef]
42. Jie, J.; Wang, Z.; Yu, Z.; Ding, Y.; Guo, B. Prototype Learning for Medical Time Series Classification via Human–Machine Collaboration. *Sensors* **2024**, *24*, 2655. [CrossRef]
43. Le, X.-M.; Luo, L.; Aickelin, U.; Tran, M.-T. ShapeFormer: Shapelet Transformer for Multivariate Time Series Classification. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 1484–1494.
44. Xiong, P.; Li, Z.; Li, Y.; Huang, S.; Liu, C.; Gu, F. Fault Diagnosis of UAV Based on Adaptive Siamese Network with Limited Data. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3531711. [CrossRef]