

Physical information embedded dual residual network for UAV actuators fault diagnosis

Fei Xue¹, Lining Tan^{1*} , Guodong Jin^{1*}, Chengxi Zhang², Hao Wang³, Jing Gao¹ and Ning Lv¹

¹ Rocket Force University of Engineering, Xi'an 710025, China

² School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China

³ School of Aeronautics, Harbin Institute of Technology, Harbin 150006, China

* Correspondence: tanlining021@hotmail.com (Tan L); jinguodong_army@163.com (Jin G)

Abstract

Accurate fault diagnosis for unmanned aerial vehicle (UAV) actuators is crucial for ensuring flight safety and completing various tasks. Deep learning models are one of the most popular methods in this field due to their powerful feature extraction capabilities. However, existing deep learning-based fault diagnosis methods lack the embedding of physical information and the ability to extract discriminative fault features from raw flight data, resulting in unsatisfactory performance in terms of noise robustness and interpretability. To improve both fault discriminability and robustness under noisy environments, a fault diagnosis method named PIEDRN is proposed. First, a dual residual channel fault feature enhancement network is proposed. By embedding physical information about UAV motion, this network accurately captures minor fault characteristics. To further improve classification accuracy for residual features of minor faults, a loss adaptation mechanism is designed and applied to enhance the Transformer network for fault diagnosis. A performance validation and an analysis of PIEDRN were conducted on the ALFA and FW-UAV datasets. The achieved diagnostic accuracy rate was consistently above 96% and reached 99.3% on several datasets. The experimental results demonstrate that this model provides significantly increased fault diagnosis accuracy under noisy conditions through the fusion of physical information. Compared with traditional deep learning approaches, the proposed method is more reliable and interpretable.

Keywords: UAV actuator, Fault diagnosis, Physical information embedding, Noise resistance, Deep learning

Citation: Xue F, Tan L, Jin G, Zhang C, Wang H, et al. 2026. Physical information embedded dual residual network for UAV actuators fault diagnosis. *International Journal of Micro Air Vehicles* 18: e017 <https://doi.org/10.48130/mav-0026-0018>

Introduction

UAVs serve as low-altitude economic carriers and have attracted extensive attention against the backdrop of rapid advances in engineering and artificial intelligence^[1]. Fixed-wing UAVs, which are characterized by high endurance and heavy payloads, have become key components in logistics, surveying, and mapping scenarios. SF Express has developed a fixed-wing UAV specifically for the logistics industry^[2]. The performance of such aircraft depends on the precise control of their actuators to maintain stable attitudes and trajectories; once actuator faults occur, the mission success rate decreases to nearly zero. Therefore, the accurate estimation of the fault states of fixed-wing UAV actuators is an urgent research issue^[3–5].

Two fault diagnosis approaches have been introduced: model-based and data-driven approaches. Model-based methods detect faults by establishing mathematical models of the target UAV system and comparing model-predicted responses with actual system measurements. These methods make use of prior physical knowledge and usually provide good interpretability and diagnostic reliability^[6]. Early studies mainly relied on linear models, such as Kalman-filter-based approaches for UAV sensor and actuator fault diagnosis^[7]. However, UAV systems are inherently nonlinear, and linear models are often insufficient for describing complex flight dynamics under varying operating conditions^[8]. To address this issue, nonlinear model-based methods, including nonlinear observers, equivalent-space methods, and adaptive robust fault estimation techniques, have been developed^[9]. Nevertheless, as UAV platforms and mission scenarios become increasingly diverse, establishing an accurate high-fidelity model for each specific UAV remains difficult and labor-intensive.

With the development of machine learning, data-driven fault diagnosis methods have gradually become an important research direction. These methods can automatically extract fault-related features from flight data and reduce the dependence on manual feature engineering^[10–15]. For spatial feature extraction, one-dimensional convolutional neural networks and depthwise separable convolution have been applied to fault diagnosis tasks^[16,17]. For temporal modeling, LSTM, BiLSTM, and attention-based recurrent models have shown advantages in capturing sequential fault features^[18,19]. In multisensor systems, graph convolutional networks and spatiotemporal graph neural networks can exploit sensor relationships to improve diagnostic robustness^[20,21]. More recently, Transformer-based models have demonstrated strong capability in long-sequence modeling, and hybrid architectures such as CNN-LSTM and GNN-Transformer models have become effective solutions for complex fault diagnosis tasks^[22].

Despite these advances, existing data-driven methods still face several limitations in UAV actuator fault diagnosis. Most deep learning models rely heavily on data-derived representations and fail to explicitly incorporate the physical operating rules of UAV systems. As a result, physical constraints, state-transition relationships, and fault-induced dynamic deviations are often ignored. For example, although Transformer-based and time-series models such as MTCL-UAV, TimesNet, and PatchTST can extract temporal features, they may confuse true fault characteristics with noise-induced fluctuations when physical constraints are absent^[23]. This issue becomes more serious under sensor noise, environmental disturbance, and measurement uncertainty, especially when fault features are weak or multiple fault modes coexist.

Embedding physical information into deep learning models provides a promising way to address these limitations. UAV motion equations, attitude dynamics constraints, and state-transition relationships describe the fundamental operating mechanisms of the aircraft. By introducing such physical information into feature extraction, the model can better highlight physically consistent fault-related deviations while suppressing irrelevant noise. This helps improve the discriminability, robustness, and interpretability of fault features, especially under complex and noisy flight conditions.

To address the above problems, a fault diagnosis network named PIEDRN is proposed. This approach combines normal and fault data to deeply embed physical information, thereby effectively mitigating the effect of sensor noise on fault diagnosis. Our contributions are listed below.

(1) A physical information-embedded dual residual network is proposed for UAV actuator fault diagnosis. By introducing UAV motion equations and state-transition residuals, the proposed model enhances fault-sensitive features and improves robustness against noise interference.

(2) A cross-attention fusion mechanism is designed to integrate raw temporal features with physics-guided residual features, enabling the model to capture complementary fault information from multichannel flight data.

(3) An adaptive loss Transformer classifier was designed to enhance the traditional Transformer network's capability to capture long-sequence fault features in data. This effectively improves the accuracy and robustness of fault diagnosis in noisy conditions.

(4) The experimental results demonstrate that PIEDRN comprehensively outperforms the existing methods on both the ALFA and FW-UAV datasets and maintains high accuracy and stability, especially under noisy conditions. Ablation and interpretability analyses further verify the effectiveness of physical information embedding.

The remainder of the article is organized as follows. In the section "Methodology", we discuss the methodology and architecture of PIEDRN. The section "Experimental design" presents a comparative experimental analysis and the results that were obtained across the two datasets. Finally, we conclude the paper.

Methodology

The proposed method unifies physical-equation residuals and data-driven state residuals into a physics-guided representation learning framework. During flight, UAVs can still fly normally even in the presence of minor faults. However, if these faults accumulate further, serious accidents may occur. In this study, a new network architecture for UAV fault diagnosis is proposed. The network consists primarily of a dual residual enhanced network and an adaptive

loss Transformer classifier module. The overall network architecture is shown in Fig. 1. For clarity, vectors are represented by bold lower-case letters, matrices by bold uppercase letters, and scalars by italic symbols throughout this paper.

Dual residual enhancement network

The actuator faults considered in this study mainly include control-surface effectiveness degradation and abnormal actuator responses. For a fixed-wing UAV, an actuator fault changes the actual control-surface deflection relative to the commanded deflection. This mismatch alters the aerodynamic forces and moments generated by the aircraft, which subsequently affects the attitude response, angular motion, acceleration, and velocity evolution. Under healthy conditions, the measured motion states are approximately consistent with the kinematic propagation and nominal state-transition relationship. In contrast, when actuator effectiveness degrades, the measured states deviate from the expected evolution because the actual aerodynamic response no longer matches the nominal control response. Therefore, such faults can be characterized by: (1) increased inconsistency between propagated and measured velocity states; and (2) increased deviation from the learned normal state-transition dynamics. These two characteristics motivate the physical residual branch and state-fitting residual branch of PIEDRN, respectively. The dual-residual design is thus intended to capture complementary fault-sensitive deviations from both explicit kinematic constraints and implicit nominal flight dynamics.

During data preprocessing, we typically apply physical models such as the Brownian motion model to perform operations such as noise reduction. These operations demonstrate the potential of physical information in data analysis. In this section, we introduce the proposed PIEDRN model. The network mainly consists of a dual residual enhancement network (DRN) and an adaptive loss Transformer classifier. The design is illustrated in detail in Fig. 1.

Physical equation residual line

The physical residual branch is designed based on the kinematic relationship between the measured UAV velocity and the velocity propagated from inertial measurements. Considering that actuator faults primarily affect attitude and acceleration, such abnormalities accumulate in velocity prediction and manifest as inconsistencies between predicted and measured states. Thus, residual signals from the physical motion equation provide fault-sensitive information while suppressing random sensor fluctuations. The model is derived under standard rigid body flight assumptions, i.e., the UAV is modeled as a rigid body, neglecting aerodynamic elastic effects, structural deformation, and high-frequency vibration dynamics.

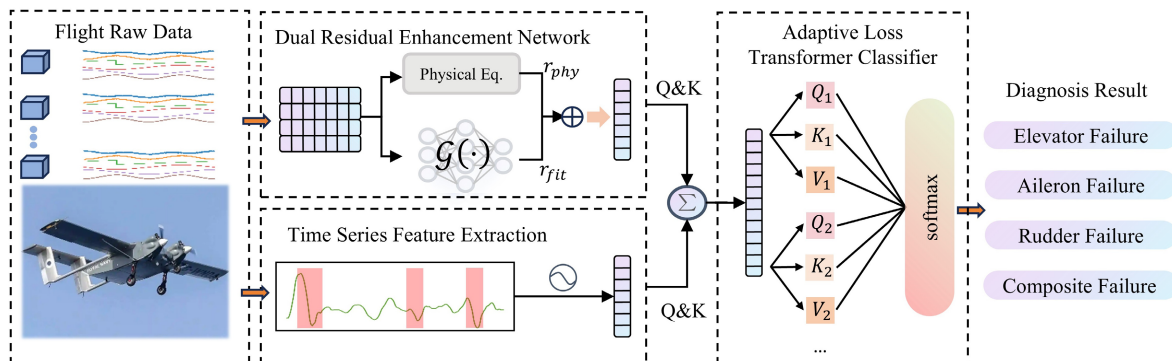


Fig. 1 Overall network architecture of the PIEDRN fault diagnosis model.

Since the full rotation matrix is employed for coordinate transformation, no small angle approximation is required. This is consistent with the datasets, which mainly consist of normal flight and moderate maneuvering conditions. Consequently, the kinematic propagation model is applicable to the operational conditions represented in the considered datasets.

For practical implementation, we assume that the sampling interval Δt is sufficiently small and that the inertial measurements are synchronized with the flight-state measurements. The body-frame acceleration is transformed into the navigation frame through the rotation matrix R_t , and gravity compensation is subsequently performed. Under normal flight conditions, the predicted velocity remains close to the measured velocity. In contrast, actuator faults introduce abnormal aerodynamic responses and control deviations, causing the residual magnitude to increase. This principle is based on the difference between the dynamic equation of an unmanned aerial vehicle and its actual measured motion state. The predicted velocity is defined as:

$$\hat{v}_t = v_{t-1} + (R_t a_t^b - g) \Delta t \quad (1)$$

where, $v_t \in \mathbb{R}^3$ denotes the measured velocity, $\hat{v}_t \in \mathbb{R}^3$ denotes the predicted velocity, $a_t^b \in \mathbb{R}^3$ denotes the body-frame acceleration, $R_t \in SO(3)$ denotes the rotation matrix, $g \in \mathbb{R}^3$ denotes the gravity vector. The physical residual is then defined as:

$$r_t^p = v_t - \hat{v}_t \quad (2)$$

Unlike conventional residual generators that require an accurate system model, the proposed residual is obtained from a simplified physical propagation equation and measured flight states. Consequently, it is computationally efficient and can be directly calculated from onboard sensor data. Since actuator faults alter the force and moment balance of the UAV, their effects are reflected in the mismatch between the predicted and actual motion states, making the physical residual a meaningful indicator of abnormal behavior. Thus, the residual vector is:

$$r_t^p = [r_{t,x}^p, r_{t,y}^p, r_{t,z}^p]^T \quad (3)$$

where, $r_{t,i}^p$ represents the residual along each axis.

State fitting residual line

Although the physical motion equation captures first-order kinematic inconsistencies, it cannot fully describe the nonlinear state-transition characteristics of UAV dynamics. To further exploit latent physical information, a state-fitting residual branch is introduced. This branch learns the nominal state evolution pattern from normal flight data and detects deviations caused by actuator faults through prediction errors. The strategy of fitting residuals for the state equation involves embedding the state equation of the target unmanned aerial vehicle into a multilayer perceptron (MLP), thereby deeply integrating physical laws and data-driven methods. The system dynamics of the UAV in its normal flight state can be expressed as follows:

$$x_{t+1} = f(x_t, u_t) + w_t \quad (4)$$

where, $x_t \in \mathbb{R}^n$ represents the state vector, $u_t \in \mathbb{R}^m$ represents the control input, and w_t represents process noise. A neural network $\hat{f}(\cdot)$ is used to approximate system dynamics:

$$\widehat{x}_{t+1} = \hat{f}(x_t, u_t) \quad (5)$$

The neural network is pretrained using only normal-flight data, where the target output is the measured next-step state x_{t+1} . The network parameters are optimized by minimizing the mean squared prediction error between x_{t+1} and $\widehat{x}_{t+1}$. In this way, the learned

model approximates the nominal state-transition function of the UAV. During inference, unseen fault conditions lead to larger prediction discrepancies because the observed state evolution no longer follows the learned normal-flight dynamics. The state residual is defined as:

$$r_t^s = x_{t+1} - \widehat{x}_{t+1} \quad (6)$$

During normal operation:

$$r_t^s \approx 0 \quad (7)$$

Under fault conditions, deviations occur:

$$r_t^s \neq 0 \quad (8)$$

where, $\hat{f}(\cdot)$ denotes the neural-network approximation of the physical state transition function $f(\cdot)$. The state residual can be interpreted as a data-driven approximation of the model mismatch. Under healthy operating conditions, the learned transition model accurately predicts future states and the residual remains small. When actuator effectiveness degrades, or abnormal control responses occur, the state-transition pattern deviates from the nominal dynamics, resulting in a noticeable increase in residual magnitude. Therefore, the state residual serves as a complementary fault-sensitive feature to the physical equation residual.

Cross-attention fusion module

When extracting physical information from the original training data, appropriate methods need to be used to achieve deep interactions between the extracted physical information and the original data. We construct a paired query-key approach to accomplish deep fusion between the original information and the physical residual information. We treat the hidden states that are obtained from the temporal features extracted from the original data as query vectors, whereas the physical residual information serves as the key vectors. By calculating the temporal similarity weights between the queries and keys ($Q \& K$, respectively), we achieve focused attention on the physical residual information. First, we construct the three projection matrices that are required for cross-attention:

$$Q_{cross} = H_x W_Q^{cross} \quad (9)$$

$$K_{cross} = H_r W_K^{cross} \quad (10)$$

$$V_{cross} = H_r W_V^{cross} \quad (11)$$

where, $Q_{cross} \in \mathbb{R}^{(T \times d_k)}$, $K_{cross} \in \mathbb{R}^{(T \times d_k)}$, and $V_{cross} \in \mathbb{R}^{(T \times d_k)}$ represent the query vector, key vector, and value vector, respectively. H_x and H_r denote the hidden state matrices of the original data and physical residual data after the temporal feature extraction process, respectively. W_Q^{cross} , W_K^{cross} , and W_V^{cross} represent the query projection weight matrix, key projection weight matrix, and value projection weight matrix, respectively. V^{cross} is the value vector. The distribution of the attention weights is as follows:

$$A = \text{softmax}\left(\frac{Q_{cross} K_{cross}^T}{\sqrt{d_k}}\right) \quad (12)$$

where, d_k represents the dimensionality of the query and key vectors. By adjusting d_k , numerical stability is maintained to ensure a reasonable attention weight distribution. The mixed feature representation is as follows:

$$Z_t = A \cdot V_{cross} \quad (13)$$

Adaptive loss Transformer classifier

On the basis of the acquired feature representations, we use an adaptive loss Transformer classifier to extract the key fault features

within the temporal dependencies, ultimately yielding probabilities for different fault modes. The fused feature Z_t is projected into a unified space:

$$\hat{Z}_t = Z_t W_{in} + b_{in} \quad (14)$$

The motivation of the adaptive loss function is to balance physical feature preservation and classification discrimination during different training stages. In the early stage, the residual-enhanced features generated by the dual residual network may not be fully aligned with the Transformer feature space. Therefore, a relatively larger feature-consistency weight encourages the classifier to preserve physically meaningful residual information and prevents the network from relying only on label-driven correlations. As training progresses, the optimization focus should gradually shift toward classification discrimination, so that inter-class separability can be further improved. This design follows a multi-objective optimization strategy, where the feature-consistency loss acts as a physical information regularization term, and the classification loss provides supervised fault-discrimination guidance. The adaptive coefficient $\alpha(\tau)$ dynamically adjusts the relative importance of these two objectives, enabling the model to achieve a better trade-off between physics-guided representation learning and fault classification accuracy. The specific formula is as follows:

$$\mathcal{L} = \alpha(\tau) \mathcal{L}_{feat} + (1 - \alpha(\tau)) \mathcal{L}_{cls} \quad (15)$$

The first term is the feature consistency loss:

$$\mathcal{L}_{feat} = \frac{1}{T} \sum_{t=1}^T |z_t - z_t^{ref}|^2 \quad (16)$$

where, z_t^{ref} denotes the reference feature obtained from the residual-enhanced encoder. The second loss is the classification loss, which increases the separability of the features:

$$\mathcal{L}_{cls} = - \sum_{c=1}^C y_c \log(\hat{y}_c) \quad (17)$$

where, y_c and $\hat{y}_c$ denote the ground-truth label and predicted probability of class c , respectively. The weighted calculation method is defined as follows:

$$\alpha(\tau) = \sigma(\eta * (\tau) / (\tau_{max}) + \beta) \quad (18)$$

where, τ , τ_{max} , η , and β denote the current epoch, the total number of epochs, and the hyperparameters, respectively. $\sigma(\cdot)$ is the Sigmoid. In this formulation, η controls the transition rate between feature-consistency learning and classification-oriented learning, while β determines the initial weighting bias. A larger η accelerates the shift toward classification optimization, whereas β adjusts the initial contribution of the physical feature-consistency constraint. The values of η and β were selected based on validation-set performance through Bayesian optimization to achieve a balanced trade-off between representation learning and fault discrimination. $\eta < 0$ controls the decay rate of the feature-consistency weight. The feedforward network dimension is set to four times d_{model} , with key parameters such as attention heads and network layers automatically optimized via the Optuna framework. After feature encoding by the Transformer, the output from the final time step is sent to a fully connected layer and mapped to the fault category space. The softmax function outputs prediction probabilities, thereby efficiently accomplishing both single- and multi-fault identification.

Training strategies and optimization

The training procedure comprises two stages. In the first stage, the dual residual network was trained exclusively on normal-flight

data to learn nominal physical and state-transition characteristics. After convergence, the residual-enhanced features were generated and combined with raw flight data through the cross-attention fusion module. In the second stage, the complete PIEDRN framework was fine-tuned end-to-end using both normal and fault samples, allowing the Transformer classifier to jointly optimize feature representation and fault discrimination. The UAV fault diagnosis procedure of the PIEDRN network is shown in Fig. 2. For all experiments, the Adam optimizer was employed for parameter optimization. The maximum number of training epochs was set to 200, and early stopping was triggered when the validation loss did not improve for 20 consecutive epochs. The batch size was fixed at 64. To improve reproducibility, all experiments were repeated using the same random seed (seed = 42). Model selection was performed according to the best validation accuracy.

Key hyperparameters (e.g., hidden layer dimensions, number of attention heads, dropout rate, and learning rate) are automatically optimized using Bayesian optimization algorithms within the Optuna framework. The Optuna framework was configured with 50 optimization trials. The search ranges included: learning rate $\in [1 \times 10^{-5}, 1 \times 10^{-2}]$, dropout rate $\in [0, 0.5]$, hidden dimension $\in [128, 512]$, number of attention heads $\in [2, 10]$, and number of Transformer layers $\in [1, 5]$. Additionally, gradient clipping, a learning rate scheduler, and early stopping mechanisms were applied to prevent overfitting and accelerate convergence. This design ensures stable performance under multiple fault scenarios and noisy environments.

The pseudocode for training and evaluating the PIEDRN is presented in Algorithm 1.

Experimental design

Dataset descriptions

The ALFA dataset was generated and disseminated by a research team from Carnegie Mellon University. It mainly provides real flight data for research on fault diagnosis and isolation as well as anomaly detection for UAVs^[24]. The dataset was collected using a modified Carbon-ZT-28 fixed-wing UAV and involved hardware such as a Holybro PX4 flight control system and an Nvidia Jetson TX2 computer. The dataset includes 47 autonomous flight records, which cover 8 different types of sudden failures. The total flight time includes 66 min of normal flight time and 13 min of post-failure flight data. Moreover, the dataset provides several hours of normal flight status data, including data from scenarios such as autonomous and manual flights.

The FW-UAV dataset was generated by a research team from the University of Toulouse, France. The dataset was collected on a UAV platform with a flying-wing layout using the Paparazzi autonomous driving system and a Raspberry Pi Zero control module^[25–27]. The dataset includes 11 flight records with a total time of more than 6 h, among which approximately 2 h are fault data. The fault types include a decrease in control surface efficiency (such as a 30% decrease in the aileron efficiency level), and the flight data include various flight scenarios (at different wind speeds).

For all datasets, we follow a consistent data partitioning strategy. Specifically, each dataset is randomly divided into training, validation, and test sets with a ratio of 70%/15%/15%. The validation set is strictly separated from the test set to avoid data leakage during hyperparameter optimization. For U1 and U2, the same model architecture is used, while the training strategy is adapted according to the fault complexity (binary vs multi-state faults). All experiments

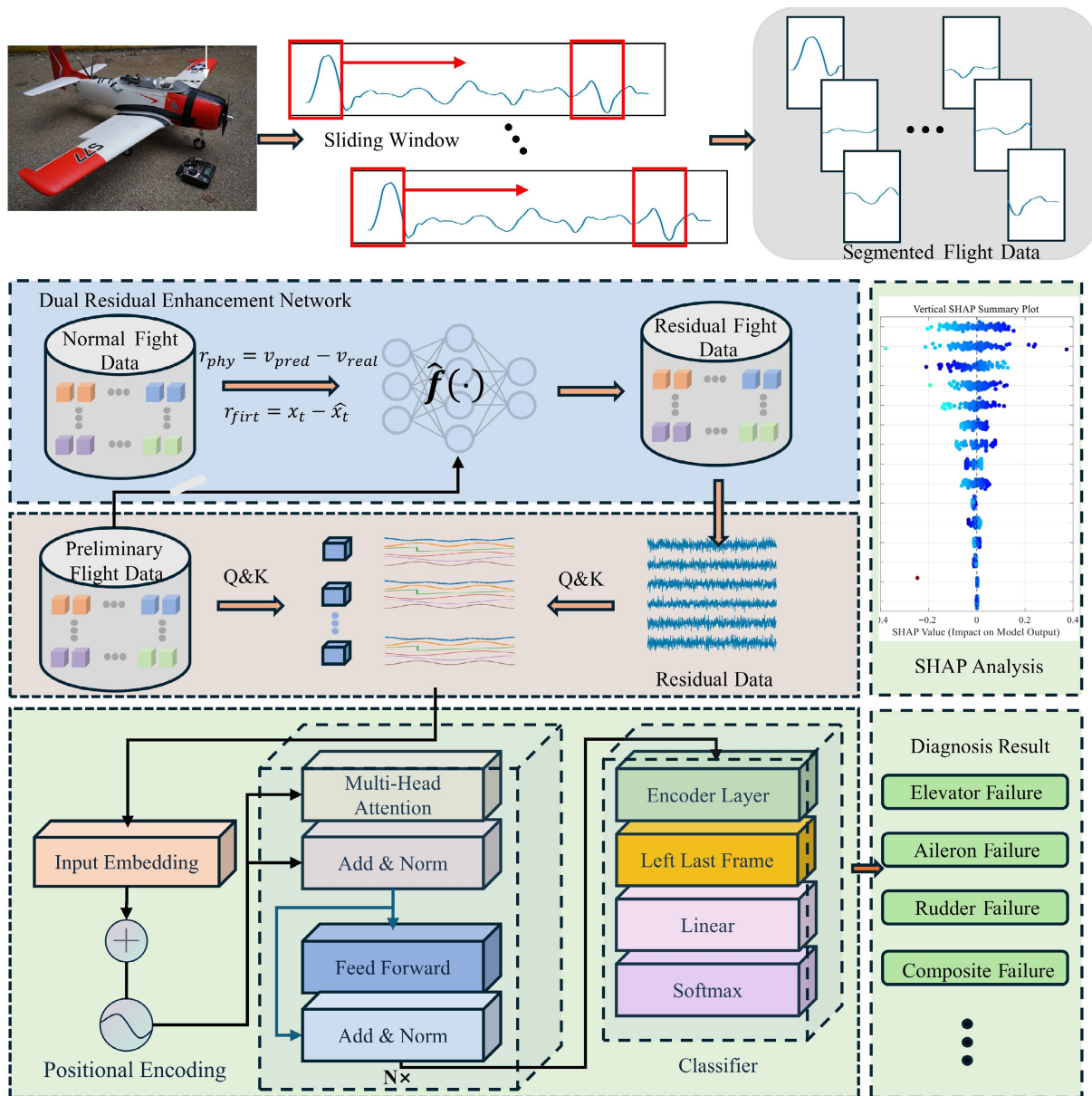


Fig. 2 UAV fault diagnosis procedure based on the proposed PIEDRN model.

are conducted under identical data preprocessing and evaluation protocols to ensure fair comparison. The dataset partitioning scheme is similar to that of Hu et al.^[23].

In this study, actuator faults are mainly categorized into control-surface effectiveness degradation and abnormal actuator response. Before training, all flight variables were synchronized according to their timestamps and normalized using z-score normalization. A sliding-window strategy was adopted to construct temporal samples, where each sample consisted of L consecutive time steps and inherited the fault label of the corresponding time interval. Missing values and corrupted records were removed during preprocessing. The same preprocessing procedure was applied to all compared methods to ensure a fair evaluation.

Equipment details and evaluation metrics

The network training platform uses an Intel Core i7-9700K processor and an NVIDIA RTX 2080Ti graphics card, which is equipped with 11 GB of memory. All the algorithms are implemented under the

PyTorch 2.0.0 framework. Owing to the adoption of the Bayesian optimization method, the hyperparameters that are used for different datasets are not exactly the same. These include the hidden layer dimensionality (HD), the dropout rate (DP), and the learning rate (LR) of the DRN, as well as the number of attention heads (AH), the hidden layer dimensionality (HD), the number of encoder and decoder layers (LY), the dropout rate, and the learning rate of PIEDRN. The specific parameters are shown in Table 1.

Four sets of core indicators are adopted to comprehensively evaluate the fault diagnosis performance of the tested methods. The accuracy rate (ACC) reflects the overall classification accuracy rate, and the precision (PRE) reflects the reliability of fault alarms. The recall rate (REC) characterizes the ability of the tested model to detect faults. The F1 score comprehensively reflects the average precision and recall rate and is applicable to imbalanced datasets. The complete network configuration, optimized hyperparameters, and evaluation protocol are reported in this paper to facilitate independent reproduction of the experimental results. All methods

Algorithm 1. PIEDRN core algorithm framework.

Require: Dataset D , loss weight α , epochs E_1, E_2 , batch size B
 Ensure: Best DRN model M_{best} , best Transformer model T_{best} , metrics Met
 1: Step 1: Data Preparation
 2: $D_{tr}, D_v, D_{te} \leftarrow \text{Split}(D)$
 3: $\text{Loader}_{tr} \leftarrow \text{DataLoader}(D_{tr}, B)$
 4: Step 2: DRN Pre-training (Residual Generation)
 5: $M_{\text{best}} \leftarrow \text{PreTrain_DRN}(D_{tr}, D_v, \alpha, E_1)$
 6: Step 3: Joint Transformer Training
 7: $T \leftarrow \text{Init_Transformer}(15)$
 8: $\text{opt}_T \leftarrow \text{Optimizer}(T)$
 9: $\text{max_acc} \leftarrow 0$
 10: for $e \leftarrow 1$ to E_2 do
 11: for each batch in Loader_{tr} do
 12: $r \leftarrow M_{\text{best}}[x_{\text{DRN}}] - [y_{\text{DRN}}]$
 13: $x_{\text{fus}} \leftarrow \text{Concat}([x_{\text{raw}}], r)$
 14: $\text{loss} \leftarrow \text{Loss_CE}(T(x_{\text{fus}}), [\text{label}])$
 15: $\text{Clip_Grad}(T)$
 16: end for
 17: $T_{\text{best}} \leftarrow T$
 18: end for
 19: Step 4: Evaluation
 20: $Met \leftarrow \text{Calc_Metrics}(T_{\text{best}}, D_{te})$
 21: return $M_{\text{best}}, T_{\text{best}}, Met$

were evaluated using the same dataset partition, preprocessing procedure, training protocol, and fixed random seed to ensure a controlled comparison. Although the proposed method was consistently validated across multiple datasets, noise levels, and ablation settings, repeated experiments with multiple random seeds and formal statistical significance tests will be conducted in future work to further quantify performance variability.

Baselines

To verify the performance of the proposed model, some deep learning methods were collected as baselines. The following is an overview of the baselines.

(1) MTCL-UAV^[23] is a multiscale Transformer-based model designed for UAV anomaly detection. It adopts a sparse mixture-of-experts architecture, where the Router adaptively selects temporal scales, and a contrastive-learning-enhanced dual attention mechanism captures intrapatch, interpatch, and neighboring patch relationships.

(2) DCdetector^[28] is a time-series anomaly detection model based on contrastive learning. It improves feature discrimination between normal and abnormal patterns, making it suitable for detecting subtle anomalies in complex temporal data.

(3) Anomaly Transformer^[29] is a Transformer-based model specifically developed for time-series anomaly detection. It models temporal dependencies through attention mechanisms and identifies anomalies by measuring abnormal association discrepancies.

(4) TimesNet^[30] is a deep learning model for general time-series analysis. It captures intraperiod and interperiod variations through multiscale temporal modeling and CNN-based structures, but it is relatively weaker than Transformer-based models in capturing global dependencies.

(5) PatchTST^[31] is a patch-based Transformer model for time-series analysis. It divides continuous sequences into patches to enhance local semantic representation and reduce modeling complexity, but it lacks adaptive multiscale modeling for complex UAV flight patterns.

Table 1. Model parameter configuration.

Dataset	DRN			PIEDRN				
	HD	DP	LR	AH	HD	LY	DP	LR
A1	462	0.2555	0.0007	8	512	2	0.3879	0.0002
A2	482	0.1265	0.0003	2	512	2	0.1843	0.0001
A3	272	0.3852	0.0008	8	512	3	0.3124	0.00001
A4	272	0.3852	0.0008	8	512	3	0.3124	0.00001
U1	465	/	0.0001	2	465	2	0.2241	0.0016
U2	220	/	0.0001	10	492	2	0.1081	0.0034

(6) CGAD^[32] is an anomaly detection model that combines causal convolutional and graph convolutional networks. It can capture dynamic patterns in flight data and often performs well in recall, although its complex modeling structure may limit generalization.

Performance comparison

All comparison methods are implemented under the same experimental setting and trained using their recommended configurations. For fairness, we re-implement all baseline models and tune their hyperparameters on the validation set when possible. All methods use identical training/testing splits.

The proposed PIEDRN model demonstrates outstanding performance advantages in the task of detecting faults in UAV actuators, and outperforms the existing methods in terms of all four indicators, as shown in Table 2. The training process of PIEDRN is illustrated in Fig. 3. A comparison of the A1 dataset before and after clustering is shown in Fig. 4. This comparison clearly demonstrates the performance of PIEDRN in feature clustering. The confusion matrices for various subdatasets are shown in Fig. 5. Owing to the limited amount of fault data in some cases, the sampling process cannot guarantee that seven fault types are consistently included. Across various subdatasets, PIEDRN consistently achieves strong classification performance.

The data in Table 2 indicate that the PIEDRN model ranks among the top performers across multiple key metrics, thus demonstrating its superior UAV fault diagnosis capabilities. PIEDRN achieves an average F1 score of 0.9749 and significantly outperforms the second-place MTCL-UAV model. This advantage is intrinsically linked to its partial feature extraction mechanism. The DRN effectively distinguishes noise from actual fault signals and amplifies the subtle fluctuations within the data that are indicative of faults. It then feeds the obtained fault states back into the output through an adaptive loss Transformer classifier. We segment the 12-dimensional data that are contained in dataset A1 into 12 channels (Ch), each of which comprises two dimensions. The diagnostic performance of the model across different channels is shown in Fig. 6. Channel 1 demonstrates superior feature representations and is easily diagnosed by the model, whereas on the other channels the model performs poorly. Crucially, the full-channel approach outperforms the single-channel methods across all the metrics, which confirms the ability of our model to integrate multisensor data and extract hidden fault information from multichannel data while mitigating interference.

Across datasets A1–A4, the adaptive loss Transformer classifier also demonstrates outstanding performance. The MTCL-UAV, DCdetector, and anomaly transformer models all incorporate this mechanism. Among them, the MTCL-UAV model uses multiple transformer architectures to jointly learn features and exhibits stronger feature extraction capabilities than a single transformer model does. However, owing to its limited depth in terms of

Table 2. Diagnosis and detection results on the ALFA dataset.

Dataset	Metrics	PIEDRN	MTCL-UAV	DCdetector	Anomaly transformer	TimesNet	PatchTST	CGAD
D1	ACC	0.9931	0.9384	0.9343	0.9470	0.9070	0.9151	0.8862
	Precision	0.9933	0.9244	0.8717	0.8788	0.8915	0.9057	0.7753
	Recall	0.9931	0.8417	0.8872	0.8410	0.7460	0.7648	0.8225
	F1-score	0.9930	0.8811	0.8794	0.8595	0.8123	0.8293	0.7982
D2	ACC	0.9724	0.9470	0.9135	0.9409	0.9266	0.9512	0.9078
	Precision	0.9727	0.7481	0.6799	0.7189	0.7385	0.7430	0.6310
	Recall	0.9724	0.9152	0.6950	0.9017	0.7177	0.8521	0.8653
	F1-score	0.9714	0.8233	0.6874	0.8000	0.7279	0.7938	0.7298
D3	ACC	0.9724	0.9299	0.9148	0.9263	0.9072	0.9129	0.8247
	Precision	0.9741	0.8047	0.7310	0.7358	0.7482	0.7752	0.4870
	Recall	0.9724	0.7136	0.7370	0.8619	0.6296	0.6386	1.000
	F1-score	0.9728	0.7564	0.7340	0.7939	0.6838	0.7003	0.6550
D4	ACC	0.9690	0.9256	0.9261	0.8981	0.8991	0.9076	0.8755
	Precision	0.9747	0.8404	0.8213	0.7882	0.8216	0.8260	0.7214
	Recall	0.9690	0.7965	0.7938	0.7137	0.6732	0.7181	0.7004
	F1-score	0.9698	0.8178	0.8073	0.7491	0.7400	0.7683	0.7107
Avg F1		0.9768	0.8197	0.7770	0.8006	0.7410	0.7729	0.7234

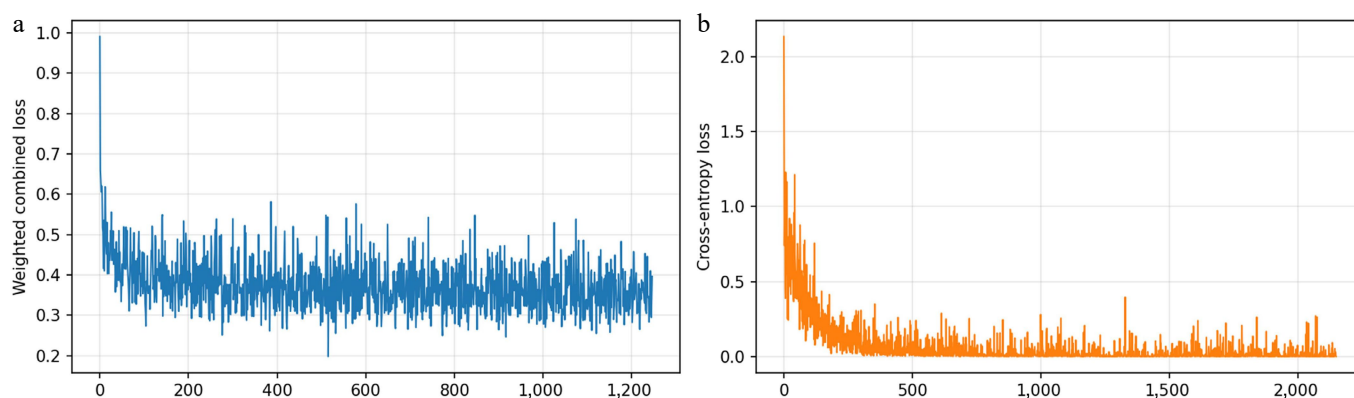


Fig. 3 (a) Training loss curve for the physical information fusion extractor, and (b) loss decay curve for the adaptive loss Transformer classifier.

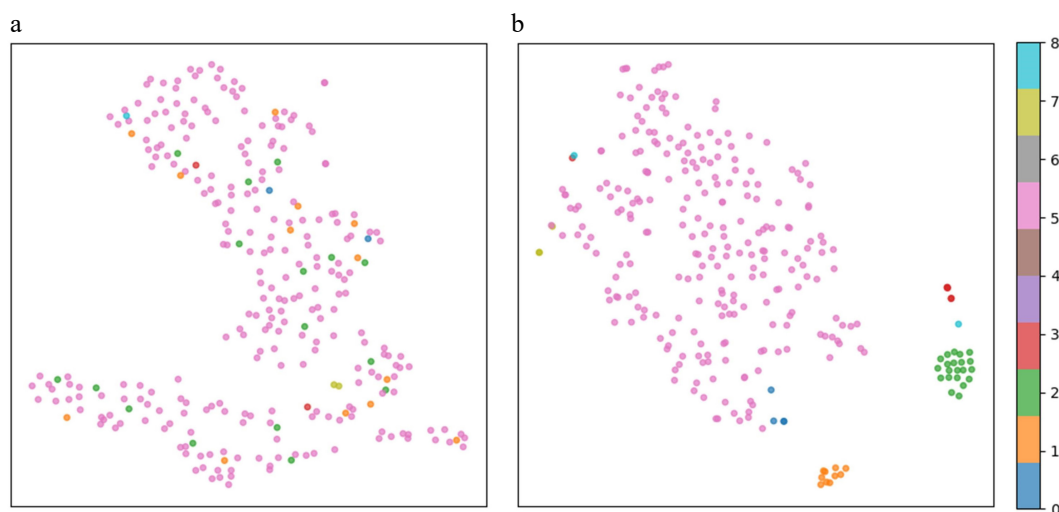


Fig. 4 (a) t-SNE plot without feature clustering, and (b) t-SNE plot after feature clustering by PIEDRN.

uncovering physical information and the coupling of fault features with noise, its discrimination ability falls short of that of the PIEDRN model. Although models such as TimesNet and PatchTST, which incorporate multihead attention, demonstrate superior time-frequency transformation capabilities and can extract long-term dependencies, their lack of physical constraints results in significantly lower REC and F1 scores on datasets A3 and A4. CGAD

attempts to incorporate causal graph structures, but its high modeling complexity limits its generalization ability. Although it achieves 1.0000 recall on A3, its precision decreases to 0.4870. These results demonstrate that in complex UAV systems, fault features are often subtle and easily mixed with noise. Without physical information, the performance of deep learning models significantly degrades.

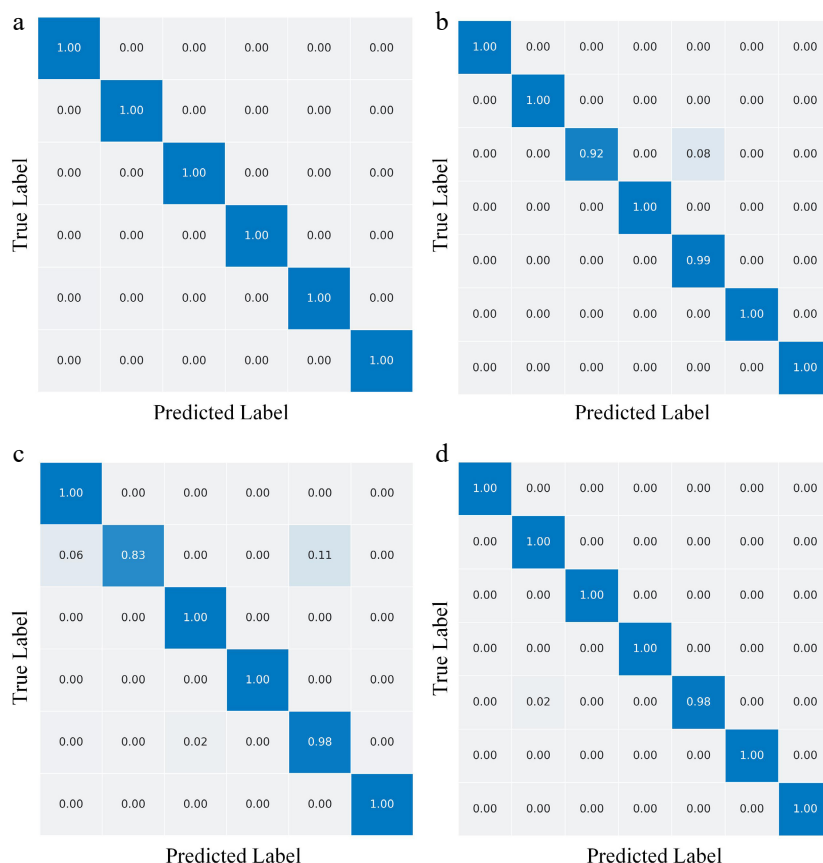


Fig. 5 Confusion matrix for datasets D1–D4.

The comparative results obtained on the FW-UAV dataset are shown in Table 3. The PIEDRN model achieves an average F1 score of 96.36% across datasets U1 and U2 and outperforms methods such as MTCL-UAV (95.68%) and DCDetector (94.58%). This outstanding performance demonstrates the high suitability of the DRN for fault detection in UAVs with flying-wing layouts. The model achieves 98.08% accuracy and recall scores on dataset U2, thus ensuring stable fault alarm triggering without missing hazardous states during multistate UAV failures. PIEDRN achieves superior overall performance across multiple fault modes on dataset U2 because of its focus on integrating physical information extraction with adaptive loss Transformer classifiers. However, the relatively

lower performance on U1 may be attributed to its single-fault structure, where physical residual enhancement provides limited additional information. The advantages of PIEDRN become more evident in multi-state and complex fault scenarios.

The core performance advantage of the model stems from its dual residual enhancement design. Unlike methods that rely on multiscale extraction or causal graphs, PIEDRN fundamentally enhances features by embedding physical equations (e.g., motion and state-fitting residuals), which reliably separates fault signals from noise. The subsequent multihead attention mechanism optimally fuses these physics-enhanced features with raw time-series data, capturing discriminative spatiotemporal patterns for multi-fault recognition.

The interpretable SHAP (Fig. 7) indicates that the physical information residual has the highest feature importance. For example, the mean absolute SHAP value for the pitch residual is 0.041, which is significantly higher than that of the original pitch data (0.027). This confirms that physical constraints amplify fault-induced deviations and improve the discriminability of fault-related features. Moreover, the hierarchical contribution of parameters and the strong synergy between control commands and their residuals demonstrate the ability of PIEDRN to capture sensor coupling and provide physically interpretable, high-fidelity diagnostics under noisy and complex fault conditions. The computational complexity of PIEDRN is mainly determined by the Transformer module, with time complexity approximately $O(T^2d)$. The DRN module introduces only linear overhead, making the method suitable for onboard UAV deployment under moderate sequence lengths.

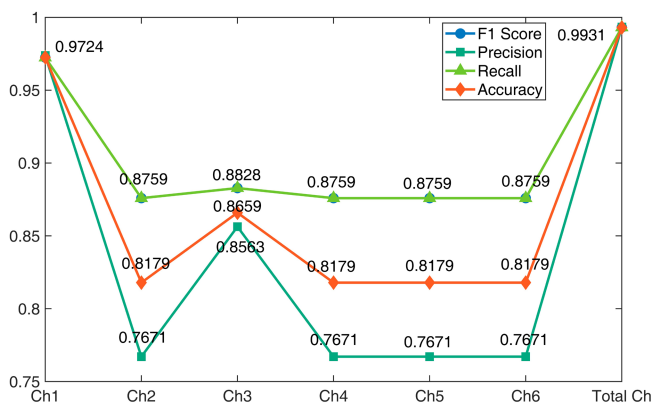


Fig. 6 Comparison of the classification performances of PIEDRN across different channels.

Table 3. The fault diagnosis results obtained on the FW-UAV dataset.

Metrics	Dataset U1				Dataset U2				Average F1
	ACC	PRE	REC	F1 score	ACC	PRE	REC	F1 score	
OCSVM	0.9754	0.9795	0.9945	0.9715	0.8879	0.7270	1.0000	0.8419	0.9067
IForest	0.9823	0.9630	0.9931	0.9778	0.8934	0.7368	1.0000	0.8485	0.9132
XGBoost	0.9190	0.8509	1.0000	0.9194	0.9761	0.9254	0.9253	0.9206	0.9200
AutoEncoder	0.9875	0.9737	0.9831	0.9784	0.9399	0.8323	1.0000	0.9085	0.9435
CGAD	0.9986	0.9733	0.9826	0.9779	0.9540	0.9124	0.9358	0.9240	0.9419
Autoformer	0.9709	0.9804	0.9562	0.9682	0.9429	0.8928	0.9192	0.9058	0.937
FEDformer	0.9793	0.9823	0.9728	0.9775	0.9432	0.8956	0.9168	0.9061	0.9418
PatchTST	0.9699	0.9643	0.9708	0.9675	0.9474	0.8668	0.9735	0.9171	0.9423
TimesNet	0.9555	0.9253	0.9833	0.9534	0.9296	0.8233	0.9733	0.8920	0.9227
Anomaly	0.9714	0.9418	1.0000	0.9700	0.9739	0.9196	0.9463	0.9328	0.9514
DCdetector	0.9662	0.9437	0.9858	0.9643	0.9546	0.8906	0.9669	0.9272	0.9458
MTCL-UAV	0.9807	0.9641	0.9954	0.9795	0.9583	0.8859	0.9878	0.9341	0.9568
PIEDRN	0.9474	0.9462	0.9474	0.9474	0.9808	0.9808	0.9808	0.9807	0.9636

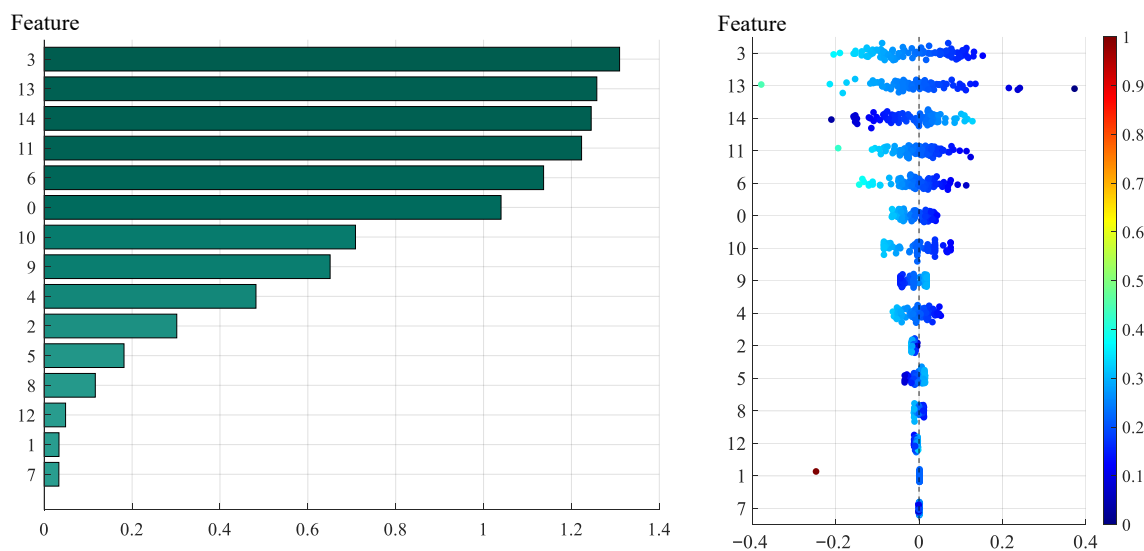


Fig. 7 SHAP visualization graph. (a) Mean values of various feature weights, and (b) a scatter plot produced with various feature weights.

Sensitivity and robustness analysis

The hyperparameters of the model are varied on the ALFA dataset, and noise is introduced to test the sensitivity and robustness of the model.

Sensitivity analysis

We also study the influences of the three key parameters of the PIEDRN model on the actuator fault detection results obtained for fixed-wing UAVs. The changes in the parameters and the model detection results are shown in Fig. 8. With respect to the number of PIEDRN layers, the accuracy of the model exhibits a clear inverted U-shaped trend. When the number of layers increases from 1 to 2, the F1 score increases from 97.24% to 99.31%, indicating that a network with more layers has better performance. However, when the number of layers increases to 5, the performance decreases to 97.99%, suggesting that the use of too many layers may lead to model overfitting.

When d_{model} is 3, the optimal performance is 99.31%. This finding indicates that when d_{model} is 3, the model can fully utilize the feature information acquired by the DRN and effectively filter the influence of sensor noise. Moreover, when d_{model} is 5, the performance decreases to 94.61%. This suggests that owing to their excessively high dimensionality, the coupling effects between aerodynamic parameters may affect the prediction of the model. Therefore,

in terms of model embeddability, appropriate model dimensions can both ensure the performance of the model and reduce the imposed equipment load.

A sensitivity analysis of the hidden layer dimensionality of the DRN reveals multiple peak characteristics. Within the range from 350 to 650, the model maintains a stable performance level of more than 99.31%, and it reaches its optimal performance at 500 dimensions. This phenomenon indicates that PIEDRN is strongly robust to the capacity of the DRN, but several performance decline points also exist. When the dimensionality increases from 200 to 500, the F1 score increases by 1.25%, whereas it decreases by 0.71% when the dimensionality is between 750 and 800. This finding indicates that the DRN generates the most accurate residuals at 500 dimensions, and the recall rate data also confirm this finding. Therefore, in engineering applications, the optimal configuration of PIEDRN includes 2 stacked layers plus d_{model} 3 and 500 residual hidden layers. This combination ensures a detection accuracy of more than 99.6% while having lower dimensionality requirements, demonstrating the lightweight advantage of the DRN in engineering practice.

When we analyze the loss weights, we find that different loss weights have different optimal values for indicators such as accuracy. When $\lambda = 0.7$, the recall rate reaches a peak of 99.31%, which reflects that the model is almost not prone to missed detections. When λ is 0.4, the accuracy is the highest, which is suitable for the

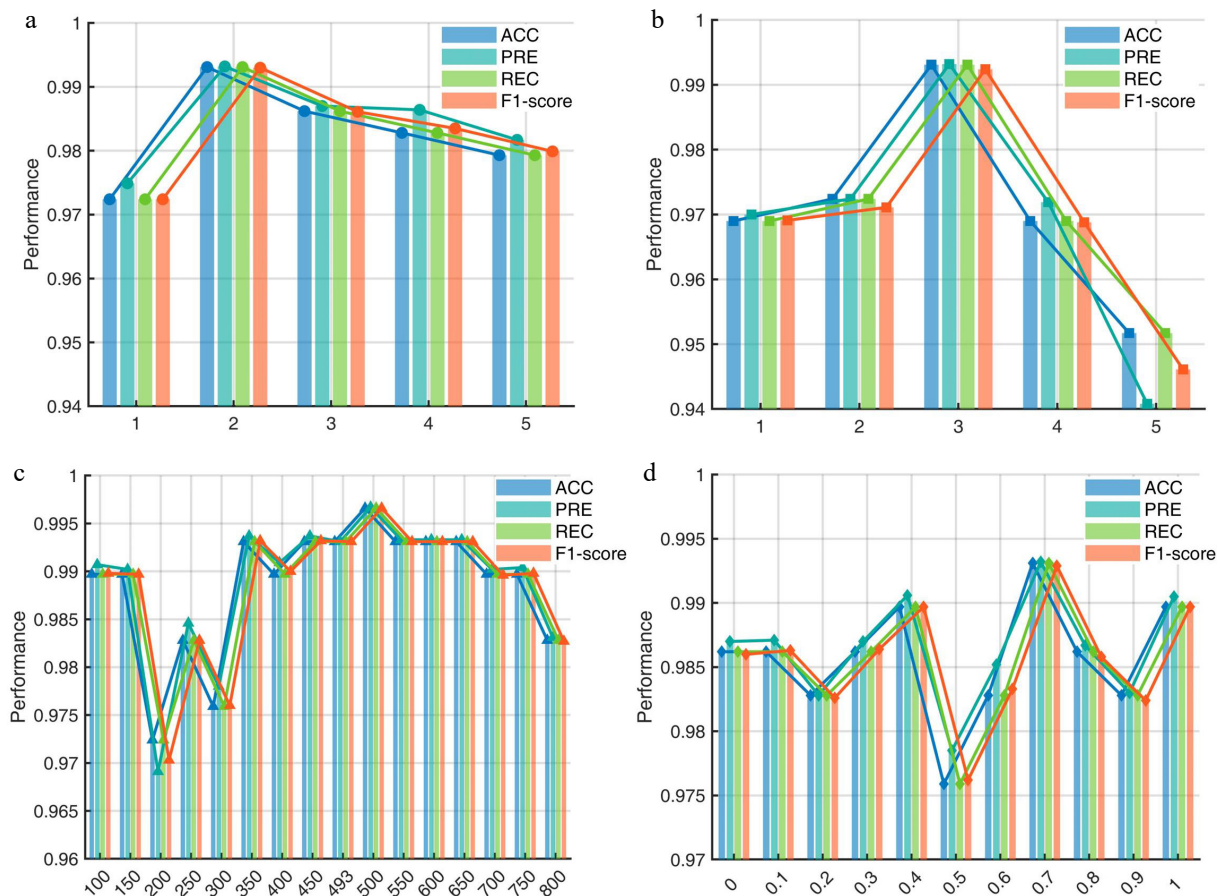


Fig. 8 Hyperparameter sensitivity change graph: (a) represents the number of stacked layers; (b) is the number of heads contained in PIEDRN; (c) is the number of hidden layers contained in the DRN; and (d) is the size of the weights.

low false-alarm rate requirement. Therefore, in engineering practice, the value of λ can be changed according to the needs of the given scenario to obtain the best effect.

Robustness analysis

Noise is independently added to each sensor channel following an additive white Gaussian noise (AWGN) model^[33]. All signals are normalized before noise injection to ensure consistent signal-to-noise ratio (SNR) definitions across datasets. This setup simulates realistic sensor disturbances in UAV flight environments. The noise types and parameter settings are not chosen arbitrarily; they are specifically designed based on typical noise phenomena observed during actual flight tests of our laboratory UAV. Two noise environments with 10% and 20% noise are designed. We also compare our model with several other models evaluated on the same dataset, as detailed in Table 4. The two noise environments are simulated by adding white Gaussian noise to the raw data, with the standard deviations of the noise set to 10% and 20% of the standard deviation of the signal, which results in SNRs of approximately 20 and 14 dB, respectively. The superior performance of our model under noise, as detailed in Table 4, is therefore not only a product of the physics-informed residual extraction process of the DRN and the attention mechanism of PIEDRN but also enabled by the redundancy and complementary information provided by the multichannel data fusion strategy.

The data in Table 4 indicate that the PIEDRN model has strong anti-noise performance. The high accuracy (99.31%) and strong robustness of PIEDRN are attributed to the introduction of the

Table 4. Comparison among the impacts of noise on the performance of different models.

Noise type	Method	Evaluation metrics			
		ACC	PRE	REC	F1 score
Normal data	SAE ^[34]	0.9455	0.9434	0.7116	0.7981
	LSTM-AE ^[35]	0.9496	0.8866	0.8021	0.8301
	LSTM-SAE ^[36]	0.9432	0.8483	0.7949	0.8286
	Stacked-LSTM ^[37]	0.8719	0.7091	0.3188	0.4119
	ltransformer ^[38]	0.9236	0.9660	0.5520	0.6917
	TMSxier ^[39]	0.9643	0.9860	0.7977	0.8734
	RSTAE ^[40]	0.9695	0.9592	0.8426	0.8934
	PIEDRN	0.9931	0.9933	0.9931	0.9930
20 dB SNR	SAE	0.9431	0.9870	0.6521	0.7754
	LSTM-AE	0.9479	0.9368	0.7203	0.8063
	LSTM-SAE	0.9373	0.8738	0.7314	0.7790
	Stacked-LSTM	0.8655	0.6707	0.2872	0.3770
	ltransformer	0.8733	0.3112	0.7406	0.4138
	TMSxier	0.9231	0.8743	0.5992	0.6886
	RSTAE	0.9521	0.9360	0.7599	0.8319
	PIEDRN	0.9897	0.9897	0.9897	0.9895
14 dB SNR	SAE	0.9365	0.9955	0.6101	0.7489
	LSTM-AE	0.9380	0.9031	0.7101	0.7794
	LSTM-SAE	0.9331	0.8536	0.7308	0.7691
	Stacked-LSTM	0.8597	0.6835	0.1285	0.2000
	ltransformer	0.8380	0.3562	0.1415	0.1859
	TMSxier	0.9104	0.8930	0.4672	0.6031
	RSTAE	0.9436	0.9498	0.4672	0.7939
	PIEDRN	0.9862	0.9888	0.9862	0.9865

DRN. The DRN filters out environmental noise through physical constraints, thereby enabling PIEDRN to focus on true fault features, which differs sharply from traditional data-driven models. Moreover, the hyperparameter tuning effect of Bayesian optimization reduces the sensitivity of the model to noise, from 99.33% in the noise-free environment to 98.88% in the 20% noise environment, thus avoiding the problem of a significant decrease in the accuracy of models such as the ltransformer in noisy environments.

In contrast, the F1 scores of traditional methods, such as the stacked LSTM model, are reduced to 20% in a 20% noise environment; thus, these methods no longer have practical engineering value. However, owing to the multiple prevention mechanisms of the PIEDRN model, the nonlinear model that is constructed by the DRN achieves noise filtering, and the attention mechanism of PIEDRN performs secondary filtering on the acquired features. In combination with the selection of the optimal hyperparameter combination via Bayesian optimization, the three mechanisms jointly ensure the robustness of the model.

Ablation experiments

To validate the effectiveness of each component module in PIEDRN, we conducted ablation experiments on the physical residual branch, the state residual branch, and the adaptive loss mechanism, respectively. All experiments were conducted under the same data split and training settings. The results are shown in Table 5.

As shown in Table 5, the full PIEDRN model achieved the best performance, with ACC and F1-score reaching 99.31% and 99.30%, respectively. When the physical residual branch was removed, the F1-score dropped to 95.16%, a decrease of 4.14% compared to the full model. This indicates that the physical residuals, constructed based on dynamics equations, can effectively enhance failure-related features and improve the distinguishability between failure and normal states. After removing the state residual branch, the overall F1-score dropped to 95.60%, a decrease of 3.70% compared to the full model. This result indicates that state residuals can effectively characterize the extent to which flight states deviate from normal dynamic evolution patterns, providing supplementary discriminative information for fault diagnosis. Furthermore, when fixed loss weights were used and the adaptive loss mechanism was removed, the model's performance declined further. The F1-score dropped to 94.63%. Dynamic weighting of the two loss terms strikes an improved balance between feature learning and classification discrimination, leading to enhanced diagnostic performance.

Overall, all three modules made positive contributions to the final performance. Among them, the physical residual branch and the state residual branch together constitute the core feature enhancement mechanism of PIEDRN. The adaptive loss strategy further enhances the model's ability to learn complex fault features. The results of the ablation experiments validate the rationality and effectiveness of the proposed network architecture design.

Conclusions

In this study, a fault diagnosis framework for UAV actuators is proposed. The network includes a dual residual enhancement

network and a Transformer classifier. Through a dual parallel residual enhancement mechanism, it physically enhances raw data by embedding them into physical constraints such as UAV motion equations and state fitting equations. This design generates residual features that describe typical physical characteristics and fully reveal hidden correlations and fault information within the data while preserving the dynamic characteristics of the original time series data. By applying cross-attention mechanisms, the network achieves deep fusion between physically enhanced residuals and original temporal features, thereby providing more recognizable inputs for subsequent diagnosis. Afterward, the mixed features are fed into the Transformer network, which applies its multihead attention mechanism to capture time-space relevance and multifault mode characteristics within long sequences, enabling precise diagnosis across multiple fault statuses. The model achieves over 98% accuracy on both the ALFA and FW-UAV datasets across all the compared metrics and maintains robust anti-interference capability. Future work should focus on incorporating the dynamic equations of the aircraft into the network to improve interpretability and generalizability, thereby promoting the application of physically informed neural networks in the health management of complex systems.

Author contributions

The authors confirm their contributions to the paper as follows: contributed to all aspects of the work, including conceptualization, methodology, investigation, formal analysis, writing – original draft, and writing – review and editing: Xue F; responsible for funding acquisition and supervision: Tan L, Jin G; contributed to the later-stage review and editing of the manuscript: Zhang C, Wang H, Gao J, Lv N. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

During the course of this research, we did not receive any form of financial support. All the necessary funds were raised by the researchers themselves or their affiliated institutions. The design, implementation, data analysis, and conclusion writing of this study were not influenced by any funding sponsors.

Conflict of interest

The authors declare that they have no conflict of interest.

Dates

Received 16 May 2026; Revised 1 July 2026; Accepted 3 August 2026; Published online 8 October 2026

References

- [1] Xiong P, Li Z, Li Y, Huang S, Liu C, et al. 2023. Fault diagnosis of UAV based on adaptive Siamese network with limited data. *IEEE Transactions on Instrumentation and Measurement* 72:1–11
- [2] Qiao G, Liu G, Shi Z, Wang Y, Ma S, et al. 2018. A review of electromechanical actuators for More/All Electric aircraft systems. *Proceedings of*

Table 5. Results of ablation experiments conducted on the ALFA dataset.

State	ACC	PRE	REC	F1 score
PIEDRN	0.9931	0.9933	0.9931	0.9930
Without DRN	0.9621	0.9658	0.9621	0.9625
Without λ	0.9897	0.9863	0.9897	0.9877

- the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 232:4128–4151
- [3] Han T, Liu C, Wu R, Jiang D. 2021. Deep transfer learning with limited data for machinery fault diagnosis. *Applied Soft Computing* 103:107150
 - [4] Chen Y, Wang Y, Yu Y, Wang J, Gao J. 2023. A fault diagnosis method for the autonomous underwater vehicle via meta-self-attention multi-scale CNN. *Journal of Marine Science and Engineering* 11:1121
 - [5] Zhou D, Huang D, Tie M, Zhang X, Hao J, et al. 2025. A fault diagnosis method based on interpretable machine learning model and decision visualization for HVs. *Applied Intelligence* 55:270
 - [6] Raissi M, Perdikaris P, Karniadakis GE. 2019. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378:686–707
 - [7] Arias Chao M, Kulkarni C, Goebel K, Fink O. 2022. Fusing physics-based and deep learning models for prognostics. *Reliability Engineering & System Safety* 217:107961
 - [8] Chao Q, Hu Y, Liu C. 2025. Physics informed neural networks for detecting the wear of friction pairs in axial piston pumps. *Reliability Engineering & System Safety* 261:111144
 - [9] Peng J, Kimmig A, Wang D, Niu Z, Fan Z, et al. 2023. A systematic review of data-driven approaches to fault diagnosis and early warning. *Journal of Intelligent Manufacturing* 34:3277–3304
 - [10] Wu Y, Sicard B, Gadsden SA. 2024. Physics-informed machine learning: a comprehensive review on applications in anomaly detection and condition monitoring. *Expert Systems with Applications* 255:124678
 - [11] Liu D, Wang N, Guo K, Wang B. 2023. Ensemble transfer learning based cross-domain UAV actuator fault detection. *IEEE Sensors Journal* 23:16363–16372
 - [12] Wang B, Liu D, Peng Y, Peng X. 2020. Multivariate regression-based fault detection and recovery of UAV flight data. *IEEE Transactions on Instrumentation and Measurement* 69:3527–3537
 - [13] Sun X, Hu Y, Qin Y, Zhang Y. 2024. Risk assessment of unmanned aerial vehicle accidents based on data-driven Bayesian networks. *Reliability Engineering & System Safety* 248:110185
 - [14] Lusch B, Kutz JN, Brunton SL. 2018. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications* 9:4950
 - [15] Swain S, Khilar PM, Senapati BR, Swain RR. 2024. Comprehensive fault diagnosis in UAV-assisted sensor networks: a three-phase automated approach. *Computer Communications* 219:76–91
 - [16] Janssens O, Slavkovikj V, Vervisch B, Stockman K, Loccufier M, et al. 2016. Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration* 377:331–345
 - [17] Chen X, Zhang B, Gao D. 2021. Bearing fault diagnosis base on multi-scale CNN and LSTM model. *Journal of Intelligent Manufacturing* 32:971–987
 - [18] Xie Y, Zhao J, Qiang B, Mi L, Tang C, et al. 2021. Attention mechanism-based CNN-LSTM model for wind turbine fault prediction using SSN ontology annotation. *Wireless Communications and Mobile Computing* 2021:6627588
 - [19] Wang S, Liu Z, Jia Z, Qin X. 2025. DFF-Net: an intelligent recognition network with high precision for unbalanced flight regimes of unmanned aerial vehicles. *Expert Systems with Applications* 274:126929
 - [20] D'Amato E, Nardi VA, Notaro I, Scordamaglia V. 2021. A particle filtering approach for fault detection and isolation of UAV IMU sensors: design, implementation and sensitivity analysis. *Sensors* 21:3066
 - [21] Liu Z, Wang L, Song Y, Dang Q, Ma B. 2022. Fault diagnosis and accommodation for multi-actuator faults of a fixed-wing unmanned aerial vehicle. *Measurement Science and Technology* 33:075903
 - [22] Sun Y, Zhu D, Wang Y, Fu Y, Tian Z. 2025. GTC: GNN-Transformer co-contrastive learning for self-supervised heterogeneous graph representation. *Neural Networks* 181:106645
 - [23] Hu G, Zhou Z, Li Z, Dong Z, Fang J, et al. 2025. Multiscale transformers with contrastive learning for UAV anomaly detection. *IEEE Transactions on Instrumentation and Measurement* 74:2530715
 - [24] Keipour A, Mousaei M, Scherer S. 2021. ALFA: a dataset for UAV fault and anomaly detection. *The International Journal of Robotics Research* 40:515–520
 - [25] Bronz M, Baskaya E, Delahaye D, Puechmore S. 2020. Real-time fault detection on small fixed-wing UAVs using machine learning. *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC), San Antonio, TX, USA, 11–15 October, 2020. USA: IEEE. pp. 1–10 doi: 10.1109/dasc50938.2020.9256800*
 - [26] Zhang Y, Li S, Yang L, Zhu Y, Liao Z, et al. 2024. FW-UAV fault diagnosis based on multilevel task knowledge supplement network under small samples. *IEEE Transactions on Instrumentation and Measurement* 73:3528814
 - [27] Keipour A, Mousaei M, Scherer S. 2019. Automatic real-time anomaly detection for autonomous aerial vehicles. *2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May, 2019. USA: IEEE. pp. 5679–5685 doi: 10.1109/ICRA.2019.8794286*
 - [28] Yang Y, Zhang C, Zhou T, Wen Q, Sun L. 2023. DCdetector: dual attention contrastive representation learning for time series anomaly detection. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Long Beach, CA, USA, 6–10 August, 2023. USA: IEEE. pp. 3033–3045 doi: 10.1145/3580305.3599295*
 - [29] Xu J, Wu H, Wang J, Long M. 2021. Anomaly transformer: time series anomaly detection with association discrepancy. *arXiv Preprint:2110.02642*
 - [30] Wu H, Hu T, Liu Y, Zhou H, Wang J, et al. 2022. TimesNet: temporal 2d-variation modeling for general time series analysis. *arXiv Preprint:2210.02186*
 - [31] Nie Y. 2022. A time series is worth 64 words: long-term forecasting with transformers. *arXiv Preprint:2211.14730*
 - [32] Wan Y, Zhang D, Liu D, Xiao F. 2024. CGAD: a novel contrastive learning-based framework for anomaly detection in attributed networks. *Neurocomputing* 609:128379
 - [33] Zhang W, Li C, Peng G, Chen Y, Zhang Z. 2018. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing* 100:439–453
 - [34] Park KH, Park E, Kim HK. 2021. Unsupervised fault detection on unmanned aerial vehicles: encoding and thresholding approach. *Sensors* 21:2208
 - [35] Park JH, Shanbhag S, Chang DE. 2022. Model-free unsupervised anomaly detection of a general robotic system using a stacked LSTM and its application to a fixed-wing unmanned aerial vehicle. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Kyoto, Japan, 23–27 October, 2022. USA: IEEE. pp. 4287–4293 doi: 10.1109/IROS47612.2022.9981950*
 - [36] Tlili F, Ayed S, Chaari L, Ouni B. 2022. Artificial intelligence based approach for fault and anomaly detection within UAVs. In: *Advanced Information Networking and Applications*, eds. Barolli L, Hussain F, Enokido T. Cham: Springer International Publishing. pp. 297–308 doi: 10.1007/978-3-030-99584-3_26
 - [37] Bell V, Rengasamy D, Rothwell B, Figueredo GP. 2022. Anomaly detection for unmanned aerial vehicle sensor data using a stacked recurrent autoencoder method with dynamic thresholding. *arXiv Preprint:2203.04734*
 - [38] Liu Y, Hu T, Zhang H, Wu H, Wang S, et al. 2023. iTransformer: inverted transformers are effective for time series forecasting. *arXiv Preprint:2310.06625*
 - [39] Chen SA, Li CL, Yoder N, Arik SO, Pfister T. 2023. TSMixer: an all-MLP architecture for time series forecasting. *arXiv Preprint:2303.06053*
 - [40] Jiang G, Nan P, Zhang J, Li Y, Li X. 2024. Robust spatial-temporal autoencoder for unsupervised anomaly detection of unmanned aerial vehicle with flight data. *IEEE Transactions on Instrumentation and Measurement* 73:3526014



Copyright: © 2026 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.