POLISH MARITIME RESEARCH 4 (128) 2025 Vol. 32; pp. 97-106
10.2478/pomr-2025-0053

AUV Fault Diagnosis Based on Multidimensional Temporal Classification Transformer

Weifeng Chen [1]
Xiaoyan Xu [1]
Mingzhi Chen [2]
Tomasz Tarasiuk [3,*]
[1] Shanghai Maritime University, Shanghai, China
[2] University of Shanghai for Science and Technology, Shanghai, China
[3] Gdynia Maritime University, Gdynia, Poland

* Corresponding author: *t.tarasiuk@we.umg.edu.pl (T. Tarasiuk)*

ABSTRACT

Autonomous underwater vehicles (AUVs) play a critical role in marine resource exploration, maritime patrol, and rescue operations, requiring reliable fault diagnosis for robust operation. This paper proposes a novel AUV fault diagnosis method based on the multidimensional temporal classification transformer (MTC-Transformer) deep learning algorithm. The MTC-Transformer enhances traditional transformer architectures by optimising them for multidimensional time-series data processing and fault classification. It features adaptive automatic feature extraction, superior multi-modal data handling, and improved scalability. Key innovations include gated fusion for combining features from outlier-processed data and kernel principal component analysis-reduced time-series data, a dedicated fault feature amplification mechanism, and a multi-layer perceptron head for classification. Trained and validated on the 'Haizhe' small quadrotor AUV fault dataset using double-layer randomised sampling to ensure generalisation, the model achieves exceptional accuracies of 99.51% (training) and 99.44% (validation). Comparative experiments demonstrate its superiority over WDCNN, LSTM-1DCNN, and other benchmarks in handling variable-length sequences and capturing long-range dependencies, confirming its high accuracy, robustness, and practical efficacy for AUV fault diagnosis.

Keywords: AUV, transformer, fault diagnosis and classification, time-series processing

INTRODUCTION

Marine resources possess significant potential for exploration and exploitation, yet complex and dynamic underwater conditions often lead to autonomous underwater vehicle (AUV) failures. Effective fault diagnosis methodologies enhance the fault tolerance and operational reliability of AUVs operating in prolonged hazardous conditions. Although conventional AUV fault diagnosis approaches demonstrate reasonable efficacy, most exhibit inherent limitations: Model-based methods (e.g. the t-test and decoupling matrix techniques for inertial measurement unit [IMU] fault detection [1]) require profound physical modelling expertise despite scenario-specific advantages; data-driven methods (e.g. stochastic resonance with wavelet reconstruction for weak propeller feature enhancement [2]) offer flexibility and low costs but impose stringent demands on data quality and feature selection algorithms; and knowledge-based systems (e.g. ontology-driven fault diagnosis frameworks [3]) enable dynamic rule-based updates yet remain constrained by the comprehensiveness of expert knowledge.

In contrast to the aforementioned approaches, machine learning-based AUV fault diagnosis methods significantly reduce the demand for specialised model expertise. By leveraging large-scale datasets, these techniques autonomously extract latent dependencies within complex data while exhibiting enhanced adaptability to novel scenarios. Recent advances primarily focus on parameter optimisation and training data refinement. Wang et al. [4] employed support vector domain description with optimised kernel mapping to construct minimal hyperspheres for anomaly detection. Pei et al. [5] integrated ensemble empirical mode decomposition-derived sample entropy features into particle swarm optimised support vector machines for effective fault classification. Hybrid frameworks demonstrate noTable improvements – Subha et al. [6] combined phase-space reconstruction with extreme learning machines for rapid anomaly diagnosis, while Li et al. [7] developed reinforcement learning-based integral extended observers with PD controllers for thruster fault tolerance. Das et al. [8] introduced genetic algorithm-supported ensemble learning (GASEL) to enhance diagnostic efficiency through feature reduction and model fusion.

These machine learning algorithms may exhibit limitations in large-scale data parallel processing, capturing long-range dependencies, and adapting to dynamic environmental conditions. Conversely, the transformer architecture demonstrates superior capabilities in data processing and adaptability. Building on the traditional transformer framework, this paper proposes a multidimensional temporal classification transformer (MTC-Transformer) model for fault diagnosis using multivariate time-series inputs. The principal innovations and contributions are summarised as follows:

(1) Optimised Architecture for Multivariate Time-Series Fault Diagnosis:

The MTC-Transformer integrates a gated fusion mechanism to dynamically combine features from statistical processing (boxplot with Spearman correlation) and nonlinear temporal reduction (kernel principal component analysis [KPCA]). This approach enhances feature representation while adaptively reducing dimensionality.

(2) Adaptive Handling of Variable-Length Sequences and Global Dependencies:

The MTC-Transformer employs KPCA for adaptive nonlinear dimensionality reduction of temporal data. This mechanism dynamically compresses variable-length sequences while preserving critical temporal dependencies, enabling efficient global feature extraction through subsequent multi-head self-attention layers.

(3) Automated Feature Engineering with Fault Amplification:

The model eliminates manual rule-setting through end-to-end parallel processing. Crucially, it introduces a fault feature amplification mechanism that magnifies deviations from normal operational patterns, increasing the sensitivity to anomalies while maintaining stability for healthy samples.

Extensive validation confirms the model's superiority over benchmarks (e.g. LSTM-1DCNN and WDCNN) in terms of accuracy and generalisation ('Results and Discussion' section). The remainder of this paper is structured as follows: The 'Selection and Processing of Dataset' section details dataset pre-processing and feature engineering, the 'MTC-Transformer Algorithm Model' section elaborates on the MTC-Transformer architecture, the 'Results and Discussion' section presents the experimental results, and the 'Conclusion' section concludes with future research directions.

SELECTION AND PROCESSING OF DATASET

The data sources selected in this paper are the datasets of the small AUV 'Haizhe' [9] collected by Zhejiang University under normal conditions and five other conditions, including 'abnormal load', 'abnormal depth sensor', 'serious propeller damage', and 'slight propeller damage' [10]. Fig. 1 shows the 'Haizhe' prototype.

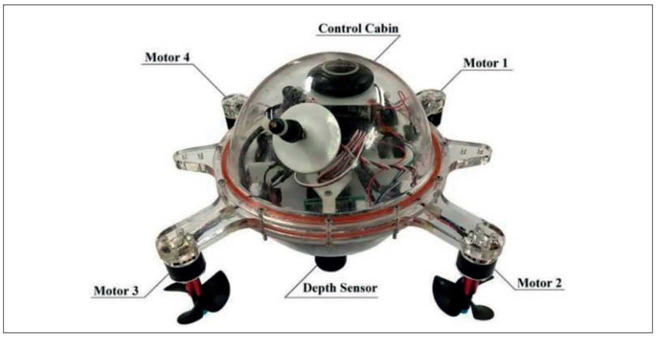

Fig. 1. The 'Haizhe' prototype

This dataset is a segmented continuous time-series sample and has been tested to confirm the reliability and validity of the data [11]. This dataset includes a total of 17 types of data, as shown in Table 1.

Tab. 1. Types of data collected by the 'Haizhe'

Type of data	Description of data
Time	The absolute point in time when the 'Haizhe' recorded the corresponding data
Pwm1~Pwm4	100 Hz Pwm wave high-level signal that controls four motors in the AUV (ms)
Depth	Depth feedback from the depth sensor in the AUV (m)
Press	Pressure feedback from the depth sensor in the AUV (Pa)
Voltage	The battery voltage used in the AUV (V)
Roll, Pitch, Yaw	Roll, pitch, and yaw angles of IMU feedback in the AUV (deg: °)
a_x, a_y, a_z	x-, y-, and z-axis accelerations of IMU feedback in the AUV (m/s^2)
w_{Roll}, w_{Pich}, w_{Yaw}	Angular velocity of IMU feedback around x, y, and z axes in the AUV (deg/s)

This study employs boxplot analysis [12] for AUV data integration and outlier detection, a standardised statistical method quantifying data dispersion through quartile ranges. The raw dataset comprises 192,598 observation groups, each containing 16-dimensional vectors (excluding temporal indices). Data anomalies exceeding 1.5× interquartile range thresholds were corrected via Winsorization at the 5th/95th percentiles. Fig. 2 visualises the processed distribution.

Given the inherent nonlinearity of underwater vehicle sensor data, Spearman's rank correlation coefficient [13] is preferentially adopted over Pearson's coefficient for robust correlation analysis, as it quantifies monotonic relationships rather than assuming linearity. This nonparametric method demonstrates superior resilience to outliers and enhanced efficacy for nonlinear dependencies, with its computational formulation expressed as follows:

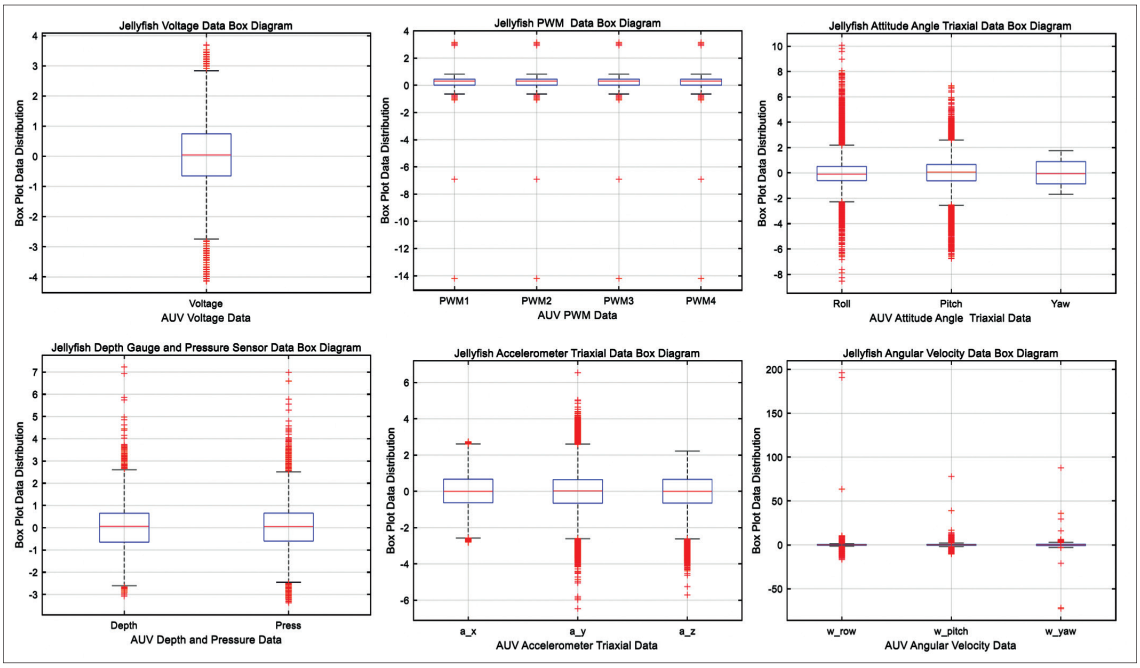

Fig. 2. Boxplot processing of 'Haizhe' AUV data

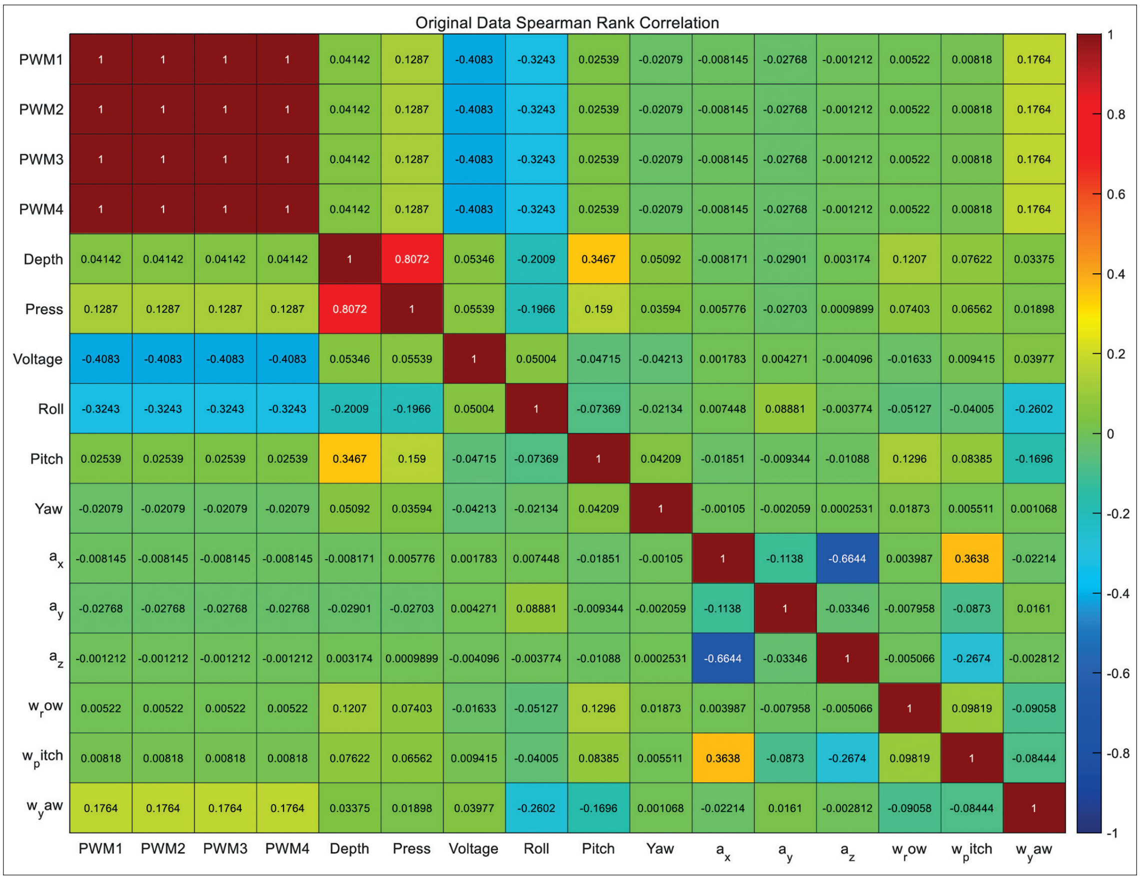

Fig. 3. Spearman rank correlation for original data

$$\rho_{ij} = 1 - \frac{6\sum_{k=1}^{n}(r_{x_i^n} - r_{x_j^n})^2}{n(n^2-1)} \quad (1)$$

In Eq. (1), ρ represents the Spearman correlation coefficient, n denotes the number of data samples used, and the sum of squared differences in the numerator magnifies the discrepancies between rank differences.

This study employs Spearman correlation analysis to conduct before-and-after comparisons during boxplot data processing. Fig. 3 and Fig. 4 show the corresponding coefficient analysis heatmaps.

A comparison of the heatmaps before and after boxplot data processing reveals that the method is effective for certain data features. For instance, the correlation coefficient between Depth

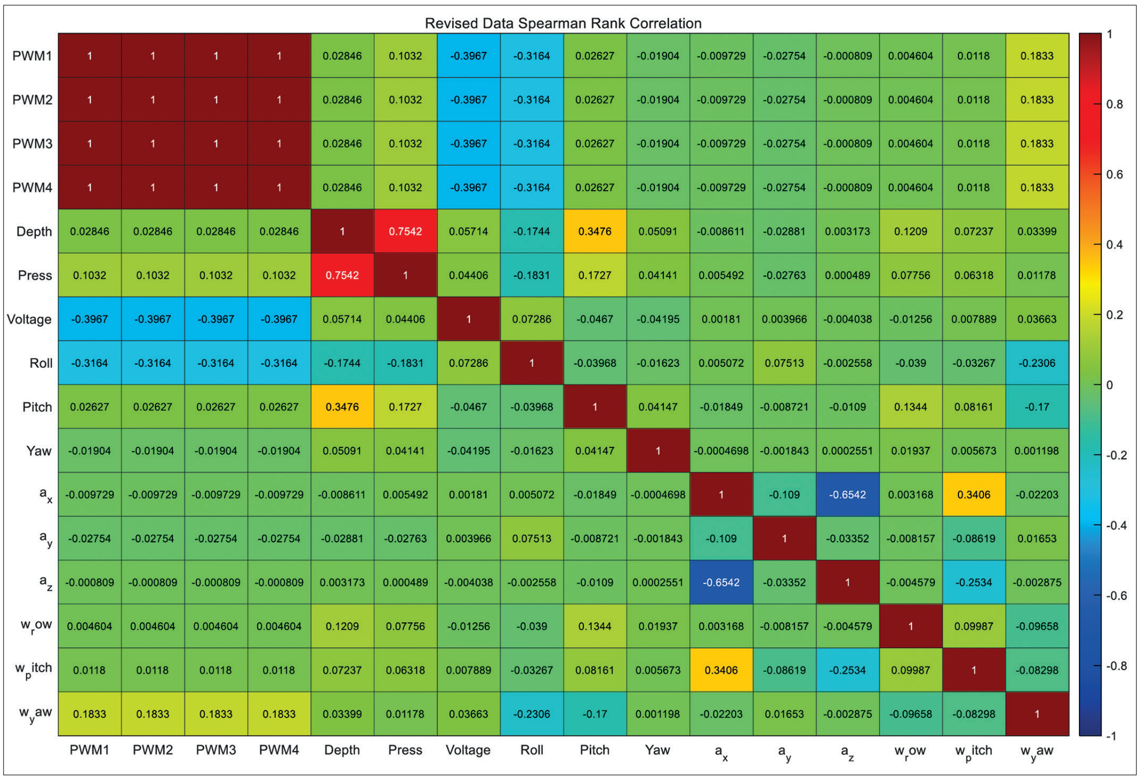

Fig. 4. Spearman rank correlation for revised data

and Press decreased from 0.8072 to 0.7542, while correlations among other data pairs also exhibited varying degrees of change. Based on the correlation analysis heatmap, dimensionality reduction can be implemented by setting correlation thresholds. For example, when setting a threshold of 0.8, the perfect correlation ($\rho = 1$) among the four PWM (Pulse Width Modulation) outputs suggests opportunities for weighted data fusion or data cleaning, achieving feature preservation and dimensionality reduction. Furthermore, applying a threshold of 0.6 identifies additional highly correlated pairs (e.g. depth-pressure and ax-az) as potential candidates for dimensionality reduction.

To enhance model convergence and generalisability, the dataset features undergo Z-score normalisation [14], followed by randomised node sampling and rolling-window segmentation. A dual-stage randomisation protocol was implemented: 5,000 training and 1,152 validation subsets were iteratively sampled from the processed data. Fig. 5 details the stratified selection procedure.

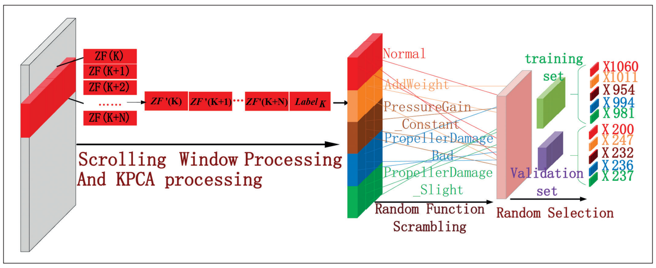

Fig. 5. Data selection based on double-layer random function

For the temporal sensor data, considering the nonlinear relationships and deeper feature dependencies within the data, this study employs KPCA with a Gaussian kernel for adaptive time-series dimensionality reduction. KPCA uses kernel functions (e.g. the Gaussian kernel) to implicitly map data into a higher-dimensional space, enabling more thorough capture of nonlinear patterns. The core formula of the Gaussian kernel function is expressed as follows:

$$k\,(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2\sigma^2}}\,. \qquad (2)$$

In Eq. (2), σ represents the kernel width parameter, governing the scale of feature space mapping. The primary computational steps involve kernel matrix computation, eigenvalue decomposition, and projection. This approach preserves more quantifiable information while reducing dimensionality, making it theoretically effective for processing coupled multi-sensor temporal data from underwater vehicles.

MTC-TRANSFORMER ALGORITHM MODEL

For the raw data, boxplot outlier processing and Spearman correlation-based dimensionality reduction are applied. For the temporal data, KPCA is used for dimensionality reduction. These two feature sets are effectively integrated by introducing a gated fusion mechanism to dynamically assign weights to each feature type:

$$Y_{fused} = G \odot X_{ofs} + (1-G) \odot X_{tfs}\,. \qquad (3)$$

In Eq. (3), G represents the generated learnable gating weight allocation matrix. The fused output Y_{fused} is then fed into the embedding layer of the MTC-Transformer to commence processing and training.

The temporal transformer is an important part of the MTC-Transformer algorithm module for processing time-series data. The first process, which consists of the time-series data entering the model, is roughly shown in Fig. 6.

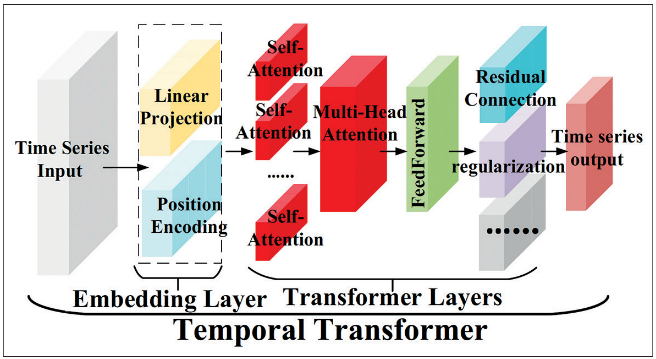

Fig. 6. Temporal transformer processes time-series data

When the model processes the incoming time-series data, the data are first transformed into a low-latitude feature vector by a linear transformation in the embedding layer, where the position encoding module helps the transformer model capture the dependence of different time steps in the data. The proposed model introduces a position encoding parameter based on a temporal convolution operation, which is a learnable matrix. The input time-series data are represented by $X \in R^{T \times C}$, where T is the length of the time series, C is the corresponding feature dimension, and the output of the model convolution operation is $Z \in R^{T' \times C'}$, which is expressed as follows:

$$Z_{i,j} = \sum_{n=1}^{K} X_{i+n-1,j} W_{k,j} + b_j + PE_{i,j} \,. \tag{4}$$

Here, K denotes the convolution kernel size, W represents the trainable kernel weight matrix, b is the bias vector, and PE indicates the position encoding tensor aligned with the temporal dimension of input sequence X. The position encoding parameters are deterministically generated via sinusoidal functions according to the following formulation for position pos and dimension index i:

$$PE_{pos,2i} = sin\left(\frac{pos}{10000^{2i/d_{model}}}\right),$$
$$PE_{pos,2i+1} = cos\left(\frac{pos}{10000^{2i/d_{model}}}\right). \tag{5}$$

The transformer architecture [15] comprises stacked layers of multi-head self-attention mechanisms and position-wise feedforward networks (FFNs), where each layer's input is derived from the preceding layer's output through successive nonlinear transformations. The self-attention module [16] computes context-aware attention weights that capture dependencies across sequence positions, enhancing the representational capacity. Fig. 7 schematises the weight calculation.

In Fig. 7, Q (query) represents the query vector part, which is used to specify the information attention part of the input sequence; K (key) represents the key vector, which is used to represent the feature pattern at each position of the input sequence and calculate the attention weight score of Q in V (value); and V represents the value vector. This contains specific information about each position in the sequence and is used as a weighted sum of the dot products of Q and K; $\sqrt{d_k}$ is used for scaling to prevent the gradient after softmax from being too small because the result of the dot product is too large.

Multi-head attention constitutes a core architectural element comprising parallelised self-attention mechanisms operating in distinct representation subspaces. This design enables the simultaneous learning of heterogeneous features through concurrent computation, generating enriched feature representations that enhance model expressiveness. The computational workflow follows:

$$Multi\text{–}Head(Q,K,V)=Concat(head_1,\ldots,\ head_h)W^o. \tag{6}$$

The position-wise FFN constitutes an essential module subsequent to self-attention, performing two-stage nonlinear transformations on attention outputs. This component enhances the representational capacity through feature abstraction, and it is formally expressed as follows:

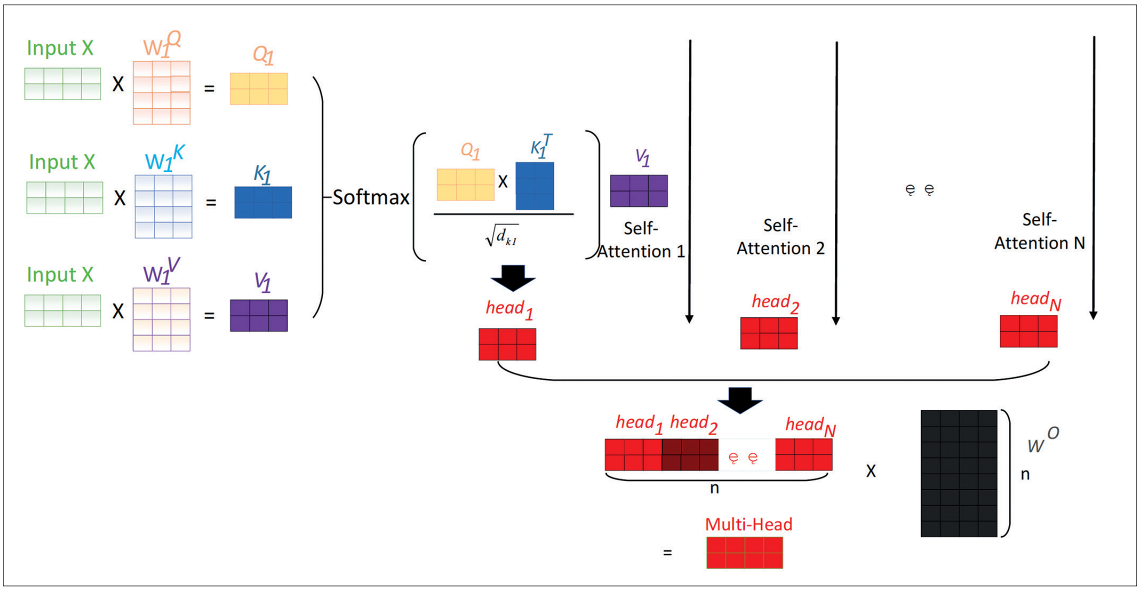

Fig. 7. Multi-head attention module

$$FFX(x) = max(0, xW_1+b_1)W_2+b_2. \tag{7}$$

The weight matrices W_1 and W_2 with the corresponding bias vectors b_1 and b_2 parameterise cascaded linear transformations, interleaved with nonlinear activations to form two fully connected layers. Crucially, while the encoder and decoder blocks maintain identical architectural configurations, their respective parameter sets remain independently optimised, without weight sharing across modules.

The residual connection mechanism mitigates gradient vanishing and exploding pathologies in deep architectures by preserving original input information through identity mapping, enhancing network stability during optimisation:

$$LayerNorm\big(X + SubLayer(X)\big). \tag{8}$$

Here, X denotes the input sequence tensor, SubLayer represents either the self-attention or feedforward submodule, and LayerNorm applies layer normalisation to the residual output, accelerating model convergence by stabilising gradient propagation during training.

A fault feature amplification module precedes the classification layer to heighten sensitivity to anomalous patterns. This component operates on the transformer's high-level output features h, applying a directional scaling transformation controlled by hyperparameter λ. The transformation projects feature deviations along the unit vector direction from the normal centroid, minimally perturbing nominal samples while selectively amplifying faults along their deviation axes. The scaling intensity λ is tuneable during model optimisation:

$$H_{mag}=h+\lambda \cdot ||h-\mu_{normal}||_2 \cdot \frac{h-\mu_{normal}}{||h-\mu_{normal}||_2 + \varepsilon} \tag{9}$$

The temporal transformer module processes raw time-series data, while the multi-layer perceptron (MLP) head maps the transformed features through a cascade of reshaping, layer normalisation, and linear projection layers. Specifically, the reshaping layer collapses temporal feature vectors into flattened 1D tensors of dimensionality [Batch_Size × (Seq_Len × Channels)] for batched linear processing.

Subsequent layer normalisation operates per-sample across feature dimensions, stabilising distributions independent of the batch size and promoting sample-wise feature homogeneity to accelerate convergence:

$$Y = \gamma \frac{X^l - \mu_L}{\sqrt{\sigma_L + \varepsilon}} + \beta. \tag{10}$$

The layer normalisation scheme incorporates trainable scaling (γ) and shifting (β) parameters beyond mean and variance statistics, enabling feature-wise affine transformations to enhance representational flexibility. A small positive constant ε ensures numerical stability by preventing division-by-zero singularities. Ultimately, the linear projection layer maps these normalised features to the diagnostic output space, completing the AUV fault classification pipeline.

The optimisation objective employs categorical cross-entropy loss [17], where model outputs undergo softmax normalisation to generate class probability distributions. These predictions are evaluated against ground-truth labels, and model parameters are iteratively refined via backpropagation to minimise the loss function through stochastic gradient descent:

$$L = -\frac{1}{N}\sum_{j}^{m}\sum_{i}^{n} T_{ji}\, log(Y_{ji}). \tag{11}$$

The loss function aggregates discrepancies across N samples, where T denotes the ground-truth label tensor and Y represents

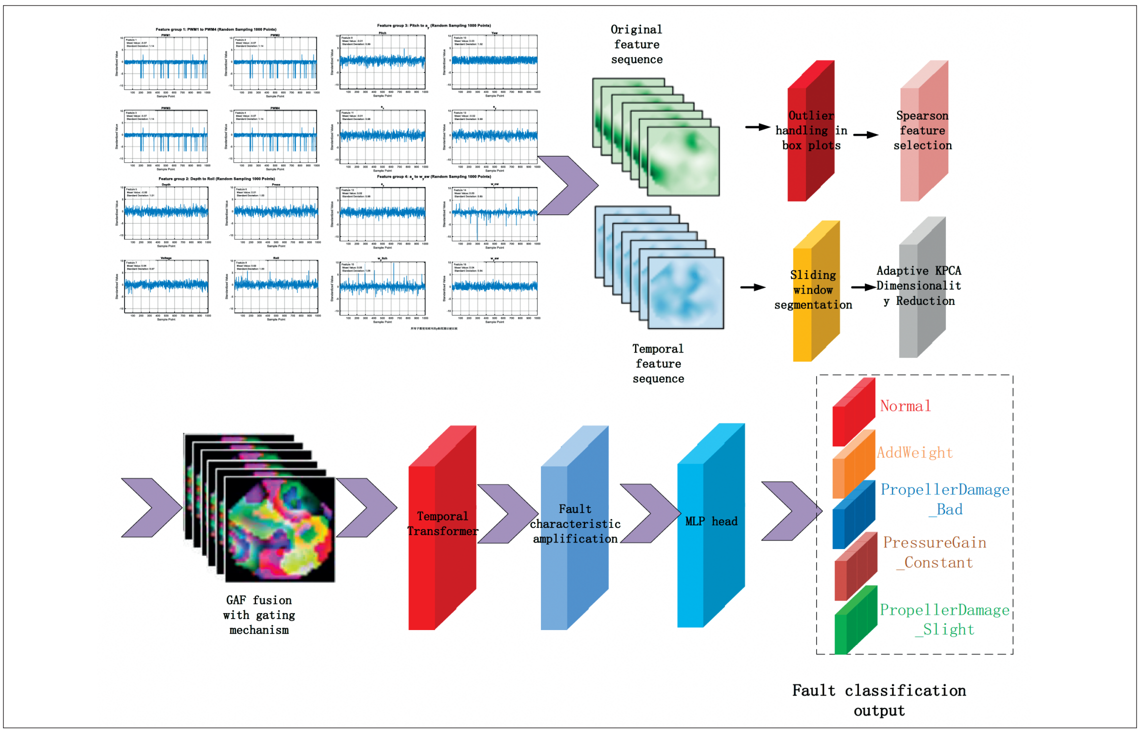

Fig. 8. Flowchart of the fault diagnosis model based on dual-channel data

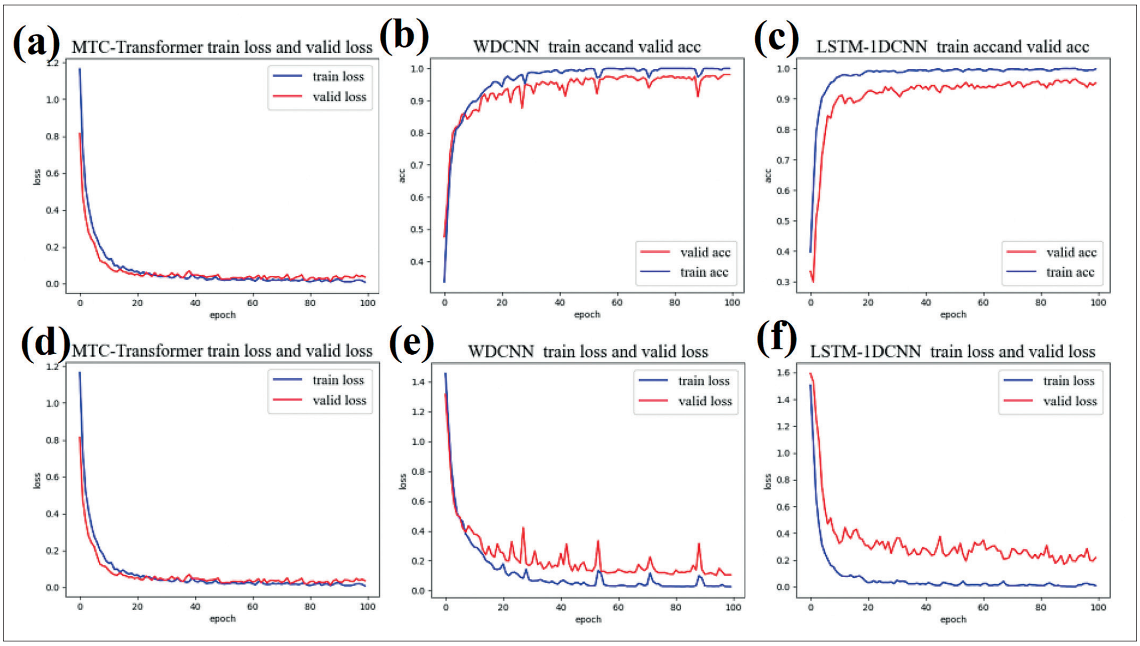

Fig. 9. Model training results

the model's probabilistic predictions. The minimisation of this metric correlates with enhanced predictive accuracy. On convergence, the MLP head generates diagnostic classifications for AUV faults [18], completing the inference pipeline.

The overall framework of the model primarily revolves around dimensionality reduction and feature enhancement. Fig. 8 depicts the complete implementation process.

RESULTS AND DISCUSSION

Comparative analysis against the WDCNN and LSTM benchmarks was conducted to validate the efficacy of the MTC-Transformer for multidimensional time-series AUV fault diagnosis. Fig. 9 illustrates the temporal trajectories of accuracy and loss across training iterations for all three architectures.

Among Fig. 9(a)–(f), panels a, b, and c are the model accuracy change graphs, and panels e, d, and f are the model loss function graphs. The figure indicates that the waveform change of the MTC-Transformer model is the smoothest and most stable, and the model has a faster convergence speed and higher initial accuracy. For a more intuitive comparison of the models, Table 2 records the training results of the three models.

Tab. 2. Comparison of model accuracy and loss

Algorithm model	Training accuracy	Validation accuracy	Training loss	Validation loss
MTC-Transformer	0.9951	0.9944	0.0158	0.0173
LSTM-1DCNN	0.9974	0.9539	0.0094	0.2121
WDCNN	0.9876	0.8856	0.0669	0.5200

Comparative analysis reveals that all three algorithms demonstrate diagnostic efficacy, although WDCNN exhibits limitations in global feature extraction and parallel processing for extended time-series sequences. While LSTM circumvents these issues, it remains susceptible to vanishing/exploding gradients in long-range dependencies. Conversely, the MTC-Transformer's inherent architectural advantages – massive parallelisability and superior long-term dependency modelling – yield an enhanced generalisation performance.

To better understand the diagnostic accuracy of the MTC-Transformer model for five types of faults, Fig. 10 also records the accuracy changes of various types of faults.

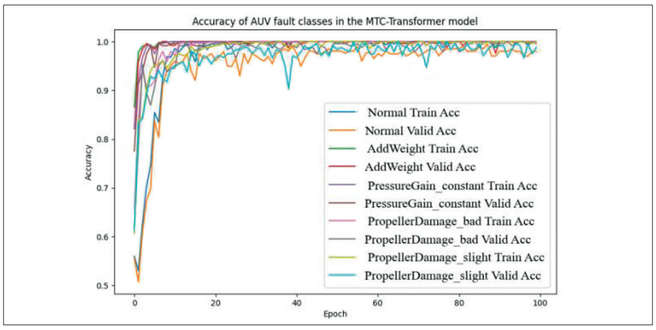

Fig. 10. MTC-Transformer fault accuracy of various AUV types

The visual analysis of convergence trajectories indicates a consistently superior diagnostic performance (>90% accuracy beyond 20–30 epochs) across diverse AUV fault types for the MTC-Transformer. To systematically benchmark model efficacy, Table 3 quantitatively compares the per-fault classification accuracy for all three architectures.

Tab. 3. The accuracy for each type of AUV fault

Algorithm model	MTC-Transformer	LSTM-1DCNN	WDCNN
Normal	0.9900	0.9183	0.9519
AddWeight	1.0000	0.9961	1.0000
PressureGain_constant	0.9957	0.9750	0.9875
PropellerDamage_bad	0.9958	0.9553	0.9756
PropellerDamage_slight	1.0000	0.9518	0.9799

Tab. 4. Model accuracy based on different time-series lengths

Algorithm model	MTC-Transformer		LSTM-1DCNN		WDCNN	
Sequence length	Training accuracy	Validation accuracy	Training accuracy	Validation accuracy	Training accuracy	Validation accuracy
4 × 16	0.9895	0.9888	0.9950	0.9133	0.9874	0.9043
8 × 16	0.9909	0.9840	0.9953	0.9445	0.9810	0.8955
16 × 16	0.9951	0.9944	0.9974	0.9539	0.9876	0.8856
24 × 16	0.9989	0.9990	0.9903	0.8912	0.9956	0.9538
30 × 16	0.9975	0.9995	0.9942	0.8848	0.9971	0.9792
36 × 16	0.9985	0.9996	0.9905	0.8879	0.9982	0.9865

Quantitative analysis confirms the efficacy of temporal feature processing for model training, with the MTC-Transformer demonstrating accelerated convergence and enhanced generalisation compared to its LSTM and WDCNN counterparts. To evaluate the robustness to variable-length sequences, we trained all architectures on 16-dimensional features segmented into progressively extended temporal windows, yielding the comparative results shown in Table 4.

For a visual comparative analysis of model performance differentials, Fig. 11 plots temporal accuracy trajectories across training epochs derived from experimental results.

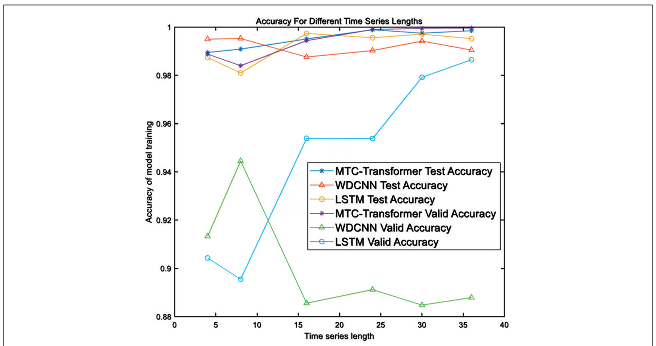

Fig. 11. Model accuracy based on different time-series lengths

Comparative validation reveals that all architectures achieve a high training accuracy across variable-length inputs, yet the MTC-Transformer maintains superior stability on the validation set. Building on prior 'Haizhe' AUV diagnostic studies [9, 19, 20], Fig. 12 benchmarks our method against state-of-the-art approaches via multidimensional radar chart visualisation of accuracy metrics.

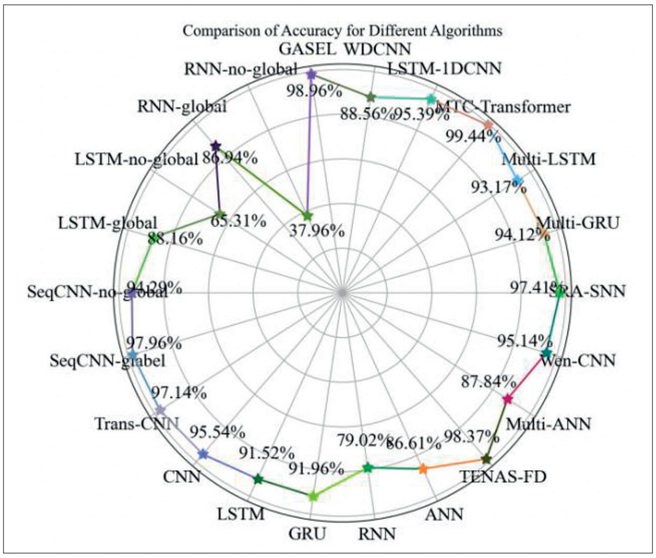

Fig. 12. Radar chart of algorithmic accuracy

Table 5 also records the training results of various algorithms to analyse the training results more intuitively.

Benchmarking analysis demonstrates that most algorithms achieve competitive performance on the 'Haizhe' dataset, with the proposed MTC-Transformer attaining state-of-the-art accuracy. This empirically validates the architecture's superior diagnostic precision and generalisation capacity relative to existing methodologies.

Tab. 5. Some algorithm training results based on 'Haizhe'

Authors	Algorithm model	Accuracy (%)	Authors	Algorithm model	Accuracy (%)
Ours	MTC-Transformer	99.44	Wang et al. [19]	CNN	95.54
	LSTM-1DCNN	95.39		LSTM	91.52
	WDCNN	88.56		GRU	91.96
Das et al. [5]	GASEL	98.96		RNN	79.02
Ji et al. [9]	RNN-no-global	37.96		ANN	86.61
	RNN-global	86.94	Pei et al. [20]	TENAS-FD	98.37
	LSTM-no-global	65.31		Multi-ANN	87.74
	LSTM-global	88.16		Wen-CNN	95.14
	SeqCNN-no-global	94.29		SRA-SNN	97.41
	SeqCNN-glabel	97.96		Multi-GRU	94.12
Wang et al. [19]	Trans-CNN	97.14		Multi-LSTM	93.17

CONCLUSION

The proposed MTC-Transformer framework for multidimensional time-series fault diagnosis demonstrates an effective performance on the 'Haizhe' dataset, with experimental results confirming the following:

(1) Gated Feature Fusion Mechanism: Integrating outlier-filtered static features (via boxplot/Spearman processing) and temporal features (via KPCA dimensionality reduction) through learnable weight allocation enhanced cross-modal feature representation.

(2) Fault Feature Amplification: Deliberately magnifying the deviation vectors of faulty samples using the operation $h + \lambda \cdot ||h - \mu_{normal}||_2 \cdot \hat{u}$ significantly boosted sensitivity to fault patterns while preserving normal samples.

(3) Generalisation Capability: To assess the cross-dataset generalisation performance of the proposed method for time-series data, we conducted preliminary evaluations on underwater acoustic signals using the SonAIr Dataset [21]. After convergence training, the model consistently achieved a classification accuracy exceeding 95%, demonstrating robust transferability to heterogeneous temporal domains without domain-specific adaptation.

Although the proposed model demonstrates promising training results on the 'Haizhe' dataset, several areas warrant further improvement:

(1) Drawing inspiration from the Trans-CNN algorithm [19], which employs convolutional neural networks (CNNs) to extract features for transformer training, our model could be enhanced by integrating multi-scale CNN feature extraction with KPCA to reduce feature redundancy and achieve feature augmentation.

(2) The TENAS-FD algorithm [20] introduces a scoring mechanism to directly select the optimal architecture from candidate networks without training. Our model could adopt similar strategies to improve computational efficiency.

(3) An AUV operating in underwater environments may encounter not only known faults but also novel fault patterns with broader spatial distributions. Future work should therefore focus on recognising unseen fault modes.

ACKNOWLEDGMENT

Project Funding: This research was supported by the National Natural Science Foundation of China (Grant No. 52371331)

REFERENCES

1. Liu X, Ju Y, Liu X, et al. An IMU fault diagnosis and information reconstruction method based on analytical redundancy for autonomous underwater vehicle. IEEE Sensors Journal 2022, 22(12): 12127–12138. https://doi.org/10.1109/JSEN.2022.3174340.

2. Liu WX, Wang YJ, Liu X, et al. Weak thruster fault detection for AUV based on stochastic resonance and wavelet reconstruction. Journal of Central South University 2016. https://doi.org/10.1007/s11771-016-3352-1.

3. Alexander T, Alexander Z, Eduard M, Valeria G, Alexander I. Description and fault diagnosis in autonomous underwater vehicles based on ontologies. 2022 International Conference on Ocean Studies (ICOS), Vladivostok, Russian Federation, 2022, pp. 39–44. https://doi.org/ 10.1109/ICOS55803.2022.10033380.

4. Wang Y, Zhang W, Di F, et al. An AUV thruster fault diagnosis method based on the improved SVDD. 2018 IEEE 8th International Conference on Underwater System Technology: Theory and Applications (USYS), IEEE, 2018, pp. 1–5. https://doi.org/10.1109/USYS.2018.8778887.

5. Pei X, Sun X, Zeng Q, et al. Research on propeller fault diagnosis of autonomous underwater vehicle. 2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), IEEE, 2023, pp. 1–6. https://doi.org/10.1109/SAFEPROCESS58597.2023.10295702.

6. Subha TD, Subash TD, Jane KSC, et al. Autonomous under water vehicle based on extreme learning machine for sensor fault diagnosistics. Materials Today: Proceedings 2020, 24: 2394–2402. https://doi.org/10.1016/j.matpr.2020.03.769.

7. Li Z, Wang M, Ma G, et al. Adaptive reinforcement learning fault-tolerant control for AUVs with thruster faults based on the integral extended state observer. Ocean Engineering 2023, 271: 113722. https://doi.org/10.1016/j.oceaneng.2023.113722.

8. Das DB, Birant D. GASEL: Genetic algorithm-supported ensemble learning for fault detection in autonomous underwater vehicles. Ocean Engineering, 2023, 272: 113844. https://doi.org/10.1016/j.oceaneng.2023.113844.

9. Ji D, Wang R, Zhai Y, et al. Dynamic modeling of quadrotor AUV using a novel CFD simulation. Ocean Engineering 2021, 237: 109651. https://doi.org/10.1016/j.oceaneng.2021.

10. Daxiong J, Xin Y, Shuo L, et al. Autonomous underwater vehicle fault diagnosis dataset. Data in Brief 2021, 39: 107477. https://doi.org/10.1016/j.dib.2021.107477.

11. Ji D, Yao X, Li S, et al. Model-free fault diagnosis for autonomous underwater vehicles using sequence convolutional neural network. Ocean Engineering 2021, 232: 108874. https://doi.org/10.1016/j.oceaneng.2021.108874.

12. Moeini B, Haack H, Fairley N, et al. Box plots: A simple graphical tool for visualizing overfitting in peak fitting as demonstrated with x-ray photoelectron spectroscopy data. Journal of Electron Spectroscopy and Related Phenomena 2021, 250: 147094. https://doi.org/10.1016/j.elspec.2021.147094.

13. Xia J, Shen Q. Correlation analysis of different vegeTable categories based on Spearman rank correlation coefficient. 2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2024, pp. 548–551. https://doi.org/10.1109/IMCEC59810.2024.10575437.

14. Rahadian H, Bandong S, Widyotriatmo A, et al. Image encoding selection based on Pearson correlation coefficient for time series anomaly detection. Alexandria Engineering Journal 2023, 82: 304–322. https://doi.org/10.1016/j.aej.2023.09.070.

15. Yuan G, Shohei M, Yuki M, et al. Spatio-temporal interpreTable neural network for solar irradiation prediction using transformer. Energy Buildings 2023, 297. https://doi.org/10.1016/j.enbuild.2023.113461.

16. Zhou X, Zhou S, Zhang Y, et al. GDALR: Global dual attention and local representations in transformer for surface defect detection. Measurement 2024: 114398. https://doi.org/10.1016/j.measurement.2024.114398.

17. Kurian NC, Meshram PS, Patil A, Patel S, Sethi A. Sample specific generalized cross entropy for robust histology image classification. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 2021, pp. 1934–1938. https://doi.org/10.1109/ISBI48211.2021.9434169.

18. Daly LM, Siamak L, Weng WL, et al. FlowTransformer: A transformer framework for flow-based network intrusion detection systems. Expert Systems with Applications 2024, 241. https://doi.org/10.1016/j.eswa.2023.122564.

19. Wang J, Zhao X, Pei S, Luo X. State monitoring and diagnosis of autonomous underwater vehicles based on transformer. 2023 5th International Conference on System Reliability and Safety Engineering (SRSE), Beijing, China, 2023, pp. 220–225. https://doi.org/10.1109/SRSE59585.2023.10336070.

20. Pei S, Wang H, Han T. Time-efficient neural architecture search for autonomous underwater vehicle fault diagnosis. IEEE Transactions on Instrumentation and Measurement 2023, 72: 1–11. https://doi.org/10.1109/TIM.2023.3327477.

21. Uvaydov D, Unal D, Enhos K, Demirors E, Melodia T. SonAIr: Real-time deep learning for underwater acoustic spectrum sensing and classification. 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), Pafos, Cyprus, 2023, pp. 9–16. https://doi.org/10.1109/DCOSS-IoT58021.2023.00011.