International Journal of Intelligence Science, 2026, 16(3), 217-233
https://www.scirp.org/journal/ijis
ISSN Online: 2163-0356
ISSN Print: 2163-0283

Deep Domain Adaptation for Cross-Condition Fault Diagnosis in Hydraulic Systems under Data Scarcity

Jiabao Zhang, Shuangsheng Ji*

Beijing Ocean-Peak Technology Co., Ltd., Beijing, China
Email: *jss@ocean-peak.com

How to cite this paper: Zhang, J.B. and Ji, S.S. (2026) Deep Domain Adaptation for Cross-Condition Fault Diagnosis in Hydraulic Systems under Data Scarcity. *International Journal of Intelligence Science*, **16**, 217-233.
https://doi.org/10.4236/ijis.2026.163011

Received: May 20, 2026
Accepted: June 27, 2026
Published: June 30, 2026

Abstract

Data-driven fault diagnosis methods have achieved remarkable success in monitoring the health of hydraulic systems under constant operating conditions. However, in real-world industrial scenarios, hydraulic systems frequently operate under variable and dynamic conditions, such as fluctuating accumulator pressures, varying temperatures, or shifting external payloads. This variability inevitably leads to a severe discrepancy in the data distribution between the training (source) and testing (target) domains—a phenomenon widely known as domain shift. Such distributional discrepancies dramatically degrade the diagnostic performance and generalization capabilities of conventional deep learning models. To address this profound challenge and the critical issue of label scarcity in newly emerging working conditions, this paper proposes a novel Deep Domain Adaptation framework tailored for cross-condition fault diagnosis in complex hydraulic systems. Specifically, a multi-channel One-Dimensional Convolutional Neural Network is meticulously designed to extract hierarchical, multi-scale temporal representations from multi-sensor time-series data. Furthermore, the Maximum Mean Discrepancy with a multi-kernel Gaussian approach is seamlessly integrated into the loss optimization function. This explicitly minimizes the marginal distribution divergence between the label-rich source domain and the completely unlabeled target domain within a deep latent reproducing kernel Hilbert space. Extensive empirical experiments are conducted on the globally recognized UCI condition monitoring of hydraulic systems dataset. The comprehensive results demonstrate that the proposed method effectively decouples intrinsic fault characteristics from operational pressure variations, achieving an outstanding diagnostic accuracy of 97.5% in a severe cross-condition task (transitioning from 130 bar to 90 bar). Rigorous comparisons with several state-of-the-art baseline models, detailed ablation studies, alongside t-SNE feature visualizations and confusion matrix analyses, robustly prove the superiority, resilience, and industrial edge-deployment potential of the proposed framework.

Keywords

Fault Diagnosis, Hydraulic Systems, Deep Domain Adaptation, Convolutional Neural Network, Maximum Mean Discrepancy, Data Scarcity

1. Introduction

Hydraulic systems are universally recognized as the core power transmission mechanisms in modern heavy industry, widely deployed in heavy machinery, aerospace engineering, and smart manufacturing. Their unparalleled high power-to-weight ratio, precise continuous control, and smooth transmission capabilities make them irreplaceable in industrial applications [1] [2]. However, due to continuous, high-intensity operation under extreme working pressures, severe internal fluid friction, and harsh environmental conditions, critical internal components of hydraulic systems—such as directional valves, hydraulic pumps, and accumulators—are highly susceptible to performance degradation and irreversible physical failures [3]. Even minor anomalies, such as internal valve leakages or localized hysteresis, can trigger severe system efficiency drops, escalate energy consumption, or cause catastrophic mechanical breakdowns, leading to immense economic losses and potential safety hazards. Therefore, developing accurate, real-time, and reliable fault diagnosis and predictive maintenance (PdM) techniques is universally acknowledged as a critical imperative in the era of Industry 4.0 and smart manufacturing [4].

Historically, condition monitoring and fault diagnosis primarily relied on traditional signal processing techniques and mathematical modeling, including Fourier transforms, wavelet analysis, and empirical mode decomposition (EMD). While effective to some extent for stationary signals, these methods heavily depend on extensive domain expertise for manual feature engineering and lack the flexibility required to seamlessly handle massive, high-dimensional multi-sensor data streams. In recent years, data-driven methodologies, particularly Deep Learning (DL), have revolutionized the field of intelligent fault diagnosis. By automatically extracting highly discriminative representations directly from raw sensor data, sophisticated architectures such as Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Transformers have overcome the inherent limitations of shallow machine learning algorithms [5] [6].

Despite their unprecedented diagnostic accuracy in controlled laboratory environments, the practical industrial deployment of conventional DL models is critically hindered by a stringent statistical premise: the training dataset and the testing dataset must be drawn from the exact same independent and identically distributed probability space. In real-world hydraulic applications, operating conditions are highly dynamic. Parameters such as system supply pressure, ambient temperature, and external loads change frequently in response to varying production demands. These operational variations significantly alter the underlying probability distribution of the collected vibration, acoustic, or pressure signals, resulting in a pervasive and challenging phenomenon known as domain shift. Consequently, when a diagnostic model trained under a specific, well-documented healthy condition (the source domain) is applied to a new, unseen degraded condition (the target domain), its diagnostic accuracy usually drops precipitously.

Furthermore, collecting massive amounts of meticulously labeled fault data for every conceivable operating condition is excessively costly, time-consuming, and often physically impossible. This structural limitation exacerbates the challenge of data scarcity and highly imbalanced class distributions [7]. To tackle the profound challenges imposed by domain shift and data scarcity, Transfer Learning, and particularly Deep Domain Adaptation (DDA), has emerged as a state-of-the-art solution. DDA endeavors to transfer diagnostic knowledge from a label-rich source domain to an unlabeled target domain by forcefully aligning their feature distributions in a high-dimensional subspace, thereby creating a domain-invariant representation [8]. In this paper, to conquer the cross-condition diagnosis bottleneck in hydraulic directional valves, we propose a robust and general-type hydraulic fault diagnosis methodology based on a multi-channel 1D-CNN architecture deeply integrated with the Maximum Mean Discrepancy (MMD) metric. The 1D-CNN acts as a powerful temporal feature extractor to fuse multi-sensor signals. Simultaneously, the multi-kernel MMD loss is strategically incorporated into the overall objective function to align the global feature representations of the source and target domains. The main contributions of this comprehensive study are summarized as follows:

- We formulate a robust Deep Domain Adaptation (DDA) framework for hydraulic systems that systematically neutralizes the detrimental impacts of fluctuating accumulator pressures. The model maintains exceptionally high diagnostic accuracy without requiring any fault labels from the target operational condition.
- We mathematically integrate a Multi-Kernel MMD (MK-MMD) penalty with a multi-channel 1D-CNN. This innovative synergy effectively decouples the intrinsic fault-specific characteristics from condition-specific variations.
- We conduct extensive validations on the public UCI condition monitoring of hydraulic systems dataset. Through multi-metric quantitative evaluations, comprehensive ablation studies, and high-resolution t-SNE visualizations, we demonstrate the overwhelming superiority of the proposed framework.
- We discuss the industrial viability of the proposed model, demonstrating that its lightweight nature makes it highly suitable for real-time Edge Computing deployment in Industrial IoT (IIoT) environments, paving the way for Digital Twin-assisted maintenance.

2. Related Work

2.1. Data-Driven Fault Diagnosis and Deep Learning

Deep learning has fundamentally transformed the paradigm of data-driven intelligent fault diagnosis and Remaining Useful Life (RUL) prediction across diverse mechanical and hydraulic systems. In the context of hydraulic condition monitoring, multi-sensor data fusion is paramount because a single sensor stream often fails to capture the holistic degradation status of complex fluid dynamics. Tao *et al.* [4] utilized multi-sensor time-series fused with 1D-CNNs to identify complex hydraulic fault patterns in real-time, highlighting the efficiency of convolutional architectures in temporal signal processing. Similarly, Wang *et al.* [9] explored 2D time-series modeling combined with self-attention fusion mechanisms to effectively decouple concurrent faults within highly integrated hydraulic systems.

Beyond standard fault classification, RUL prediction under continuous degradation remains a focal point. Han and Mo [5] proposed an RUL prediction method based on sequential health index evaluation coupled with multi-dimensional degradation data. Furthermore, sophisticated architectures like gated convolutional autoencoders [10] have been proposed to enhance predictive maintenance under noisy environments. However, when deployed under severe cross-condition scenarios (e.g., sudden pressure drops or load changes), these purely supervised models inevitably suffer from severe performance degradation because they completely lack explicit domain alignment mechanisms.

2.2. Handling Data Scarcity and Imbalance

In authentic industrial environments, hydraulic systems operate under normal healthy conditions for the vast majority of their lifecycles. Fault data are extremely rare, leading to severe class imbalance and data scarcity. To address this pervasive issue, advanced data augmentation techniques have been actively investigated. Udu *et al.* [7] provided a comprehensive review of Synthetic Minority Over-sampling Technique (SMOTE) and Generative Adversarial Network (GAN) variants for balancing machine learning tasks. Specifically, Zhao *et al.* [11] successfully implemented a Domain Adaptation Deep Convolutional GAN (DA-DCGAN) to mitigate imbalanced sample problems in hydraulic pumps. Similarly, Deng *et al.* [12] proposed an improved deep convolution GAN integrated with multi-signal fusion to bolster fault diagnosis performance under data scarcity. While GAN-based augmentation enriches the training space, explicitly minimizing the distribution gap across operational domains remains a more direct and computationally efficient approach to achieving cross-condition robustness.

2.3. Domain Adaptation and Transfer Learning

To overcome the detrimental effects of domain shift, Deep Domain Adaptation (DDA) aims to establish a universally applicable, domain-invariant feature space. Recent academic endeavors have intensely focused on cross-condition and cross-machine diagnostics. For instance, Wu *et al.* [6] proposed a novel Transformer

Transfer Learning Network (TTLN) specifically tailored for cross-condition scenarios in rotating machinery. For complex hydraulic and fluid power systems, DDA combined with ensemble learning strategies has showcased promising preliminary results in aligning marginal distributions [8]. Inspired by these remarkable advancements, this paper strategically adopts a multi-kernel MMD-based statistical alignment strategy due to its mathematically proven stability and superior computational efficiency in aligning high-dimensional, multi-sensor hydraulic fault features [13].

2.4. Edge Computing and Digital Twins for Maintenance

With the rapid evolution of the Industrial Internet of Things (IIoT) and Industry 4.0, the physical deployment of complex diagnostic algorithms is progressively shifting from centralized cloud servers towards Edge Computing paradigms. Edge computing guarantees real-time robustness, preserves data privacy, and ensures ultra-low-latency anomaly detection [14]. Concurrently, the integration of these models into Digital Twin (DT) architectures provides a profound virtual-physical mapping of the equipment. As discussed by Wang *et al.* [15], fault diagnosis and predictive maintenance based on digital twin models enable operators to interactively simulate faults and predict degradation trajectories without halting actual production. Our proposed framework aligns with this trajectory by offering a lightweight, condition-agnostic inference model ideal for edge nodes within a comprehensive digital twin ecosystem.

3. Methodology

3.1. Problem Formulation and Domain Shift

In the context of cross-condition fault diagnosis, we formally define a source domain $\mathcal{D}_s = \left\{\left(x_i^s, y_i^s\right)\right\}_{i=1}^{n_s}$ containing n_s fully labeled samples collected under a specific, optimal operating pressure (e.g., 130 bar). Concurrently, we define a target domain $\mathcal{D}_t = \left\{x_j^t\right\}_{j=1}^{n_t}$ containing n_t entirely unlabeled samples collected under a degraded or variant operating pressure (e.g., 90 bar).

The fundamental assumption of transfer learning is that both domains share an identical feature space $\mathcal{X}$ and label space $\mathcal{Y}$, but their marginal probability distributions differ drastically, *i.e.*, $P_s(X) \neq P_t(X)$. This inequality is the mathematical root of the domain shift problem. The ultimate objective of this study is to leverage the abundant labeled knowledge from $\mathcal{D}_s$ to train a robust deep neural network $G(\cdot)$, parameterized by θ, capable of accurately predicting the hidden labels y^t of the target domain data $\mathcal{D}_t$.

3.2. Multi-Sensor Data Preprocessing and Fusion

Industrial hydraulic sensors typically possess varying sampling frequencies due to the diverse physical nature of the measured phenomena (e.g., pressure sensors often sample at 100 Hz, while volume flow sensors sample at 10 Hz). To construct a unified, multidimensional input tensor for the CNN, one-dimensional linear interpolation is employed to unsampled low-frequency signals to precisely match the highest frequency. Let $S_{low}(t)$ be the low-frequency signal; the unsampled signal $S_{high}(t)$ at target timestamp t is formulated as:

$$S_{high}(t) = S_{low}(t_k) + \frac{S_{low}(t_{k+1}) - S_{low}(t_k)}{t_{k+1} - t_k}(t - t_k) \tag{1}$$

where $t_k \le t < t_{k+1}$.

Subsequently, Z-score standardization is applied to eliminate the severe dimensional disparities among different physical quantities. For a given feature vector x:

$$x_{norm} = \frac{x - \mu_s}{\sigma_s} \tag{2}$$

Crucially, to rigorously prevent any futuristic data leakage, the mean μ_s and standard deviation σ_s must be exclusively computed from the source domain training set and then rigidly applied to scale the target domain data. Finally, a sliding window technique with a fixed window size W is applied to segment the continuous sequences into discrete tensor samples.

3.3. Proposed 1D-CNN Feature Extractor

The architectural backbone of our Domain Adaptation CNN (DA-CNN) framework is a multi-channel 1D-CNN feature extractor, denoted as G_f, operating under shared weights across both domains. The overall architecture is vividly detailed in **Figure 1**. The input consists of normalized multi-sensor time-series matrices $X \in \mathbb{R}^{C \times W}$, where $C = 2$ is the number of sensor channels (pressure and flow), and $W = 1024$ represents the time window sequence length. The feature extractor G_f comprises multiple cascaded convolutional blocks. For a given input sequence *x*, the 1D convolution operation for the *j*-th feature map in the *l*-th layer is mathematically defined as:

$$z_j^{(l)} = \sigma\left(\sum_{i=1}^{C_{in}} W_{i,j}^{(l)} * x_i^{(l-1)} + b_j^{(l)}\right) \tag{3}$$

where $*$ denotes the 1D convolution operator, $W_{i,j}^{(l)}$ is the convolutional kernel weight, $b_j^{(l)}$ is the bias term, and $\sigma(\cdot)$ represents the Rectified Linear Unit (ReLU) activation function, which injects essential non-linearity: $\sigma(z) = \max(0, z)$. The initial convolutional layer deliberately utilizes a wide kernel (e.g., $K = 15$) to aggressively capture macroscopic temporal fluctuations and suppress high-frequency mechanical noise. Following the convolution, a Max-Pooling layer is applied to down-sample the temporal dimension, retaining the most salient features and providing necessary translation invariance. The pooling operation is expressed as:

$$p_j^{(l)}(k) = \max_{0 \le m < s}\left(z_j^{(l)}(k \cdot s + m)\right) \tag{4}$$

where *s* is the pooling stride.

As illustrated in the central module of **Figure 1**, the tensor dimensions progressively contract from $B \times 2 \times 1024$ down to $B \times 32 \times 254$. Finally, an Adaptive Global Average Pooling (GAP) layer compresses the temporal dimension completely, transforming the feature maps into a highly compact, 64-dimensional latent feature vector $f \in \mathbb{R}^{64}$. A classification head G_y, comprising fully connected layers combined with Dropout regularization ($p = 0.5$) to prevent overfitting, subsequently maps f to the probability distribution over the predefined hydraulic valve states using a SoftMax function.

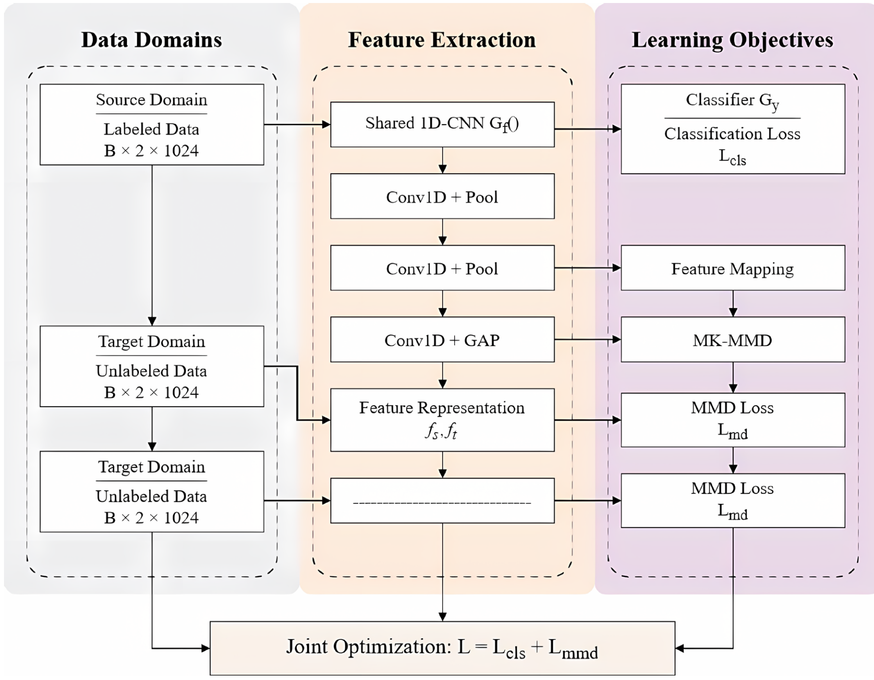

Figure 1. The architecture of the deep domain adaptation convolutional neural network (DA-CNN).

3.4. Multi-Kernel Maximum Mean Discrepancy

To mathematically obliterate the domain shift illustrated in Figure 1, the MMD is deployed. MMD is a formidable non-parametric statistical metric that measures the distribution distance between the source and target domains by mapping the original features into a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$. The empirical estimate of the squared MMD loss is rigorously formulated as:

$$\mathcal{L}_{MMD}^{2} = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi\left(f_i^s \right) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi\left(f_j^t \right) \right\|_{\mathcal{H}}^{2} \tag{5}$$

where $f^s = G_f\left(x^s\right)$ and $f^t = G_f\left(x^t\right)$ are the deep representations extracted from the source and target domains, respectively, and $\phi(\cdot)$ is the implicit mapping function defining the RKHS. By expanding the squared norm using the kernel trick $k(x,y) = \langle \phi(x), \phi(y) \rangle_{\mathcal{H}}$, the MMD loss can be computed explicitly without knowing the exact algebraic form of $\phi(\cdot)$.

In this complex hydraulic study, a Multi-Kernel Gaussian (MK-MMD) strategy is utilized to exponentially enhance the alignment capability across highly variable, multi-scale features [13]:

$$k\left(f^s, f^t\right) = \sum_{m=1}^{M} \beta_m \exp\left(-\frac{\left\| f^s - f^t \right\|^2}{2\sigma_m^2}\right) \tag{6}$$

where $M = 5$ is the number of distinct Gaussian kernels, σ_m dictates the varying bandwidths, and β_m is the corresponding kernel weights (set uniformly in this configuration). By combining multiple kernels, the network is simultaneously forced to align both macro-level distributional shifts and micro-level feature variances.

3.5. Joint Optimization Objective

The overarching training objective of the DA-CNN synchronously unifies the supervised classification error on the source domain and the unsupervised domain discrepancy error across domains. The classification loss utilizes the standard Categorical Cross-Entropy:

$$\mathcal{L}_{cls} = -\frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{c=1}^{C_{class}} \mathbb{I}\left[y_i^s = c\right] \log\left(G_y\left(G_f\left(x_i^s\right)\right)_c\right) \tag{7}$$

where C_{class} is the number of fault categories, and $\mathbb{I}[\cdot]$ is the indicator function.

The total optimization objective $\mathcal{L}_{total}$ is synthesized as:

$$\min_{\theta_f, \theta_y} \mathcal{L}_{total} = \mathcal{L}_{cls} + \lambda \mathcal{L}_{MMD}\left(f^s, f^t\right) \tag{8}$$

where λ is a critical hyperparameter dynamically regulating the penalty strength of the domain alignment. Through iterative backpropagation using the Adam optimizer, the network parameters (θ_f for the extractor and θ_y for the classifier) are updated, forcing the extractor to learn representations that are concurrently fault-discriminative and operational-condition-invariant.

4. Experiments and Results

4.1. Dataset Description and Task Construction

To empirically validate the undeniable effectiveness of the proposed framework, we utilize the universally benchmarked Condition Monitoring of Hydraulic Systems dataset provided by the UCI Machine Learning Repository [1]. The dataset was generated by a sophisticated industrial hydraulic test rig comprising a primary working circuit and a secondary cooling-filtration circuit. The system cyclically executes constant load profiles lasting 60 seconds per cycle, accumulating a total of 2205 operational cycles over varying component degradation states. To meticulously construct the Domain Adaptation task, we isolate the accumulator precharge pressure as the shifting condition variable. Variations in accumulator pressure drastically alter the transient response, volumetric efficiency, and pressure ripple of the entire fluid power system. The Source Domain ($\mathcal{D}_s$) constitutes data recorded at the optimal pressure of 130 bar. The Target Domain ($\mathcal{D}_t$) comprises data captured under a critically degraded pressure of 90 bar. The ultimate diagnostic objective is to classify the condition of the hydraulic directional valve into four severity states: Optimal behavior (100%), Mild Hysteresis (90%), Severe Hysteresis (80%), and Near Failure (73%).

Among the multiple sensors available in the UCI dataset, we specifically retain two critical channels: the PS2 pressure sensor (measuring system operating pressure at 100 Hz) and the FS1 volume flow sensor (measuring hydraulic fluid flow rate at 10 Hz). These two sensors were selected because pressure and flow collectively capture the essential hydraulic dynamics most directly indicative of valve degradation. PS2 directly reflects the accumulator pre-charge condition and system pressure stability, while FS1 provides complementary information on volumetric efficiency and internal leakage. Other sensors, such as temperature and vibration sensors, exhibit weaker correlation with valve hysteresis severity in preliminary analysis and were excluded to maintain a lightweight input tensor and prevent redundant feature dimensions.

To quantitatively characterize the data scarcity setting, **Table 1** provides a detailed breakdown of the dataset composition after preprocessing. The source domain (130 bar) comprises a total of 2205 operational cycles, which were segmented into approximately 10,500 non-overlapping windows of length $W = 1024$ points each. The target domain (90 bar) contains 1847 operational cycles, yielding approximately 8800 windows. Crucially, the source domain samples are fully labeled, whereas all target domain samples remain completely unlabeled during training—emulating the realistic industrial scenario where labeled data from degraded operating conditions are entirely unavailable. Furthermore, as detailed in **Table 1**, the windowed dataset exhibits significant class imbalance: the Optimal (100%) class dominates with approximately 5200 source windows and 4400 target windows, while the Near Failure (73%) class is severely underrepresented with only 780 source windows and 650 target windows. This simultaneous scarcity of unlabeled target data and severe intra-domain class imbalance collectively defines the data scarcity challenge addressed in this study (**Figure 2**).

Table 1. Dataset composition across source and target domains after windowing (W = 1024).

Class	Source Domain (130 bar)		Target Domain (90 bar)	
	Cycles	Windows (1024)	Cycles	Windows (1024)
Optimal (100%)	1285	~5200	1078	~4400
Mild Hysteresis (90%)	442	~1900	368	~1550
Severe Hysteresis (80%)	298	~1620	251	~1200
Near Failure (73%)	180	~780	150	~650
Total	2205	~10,500	1847	~8800

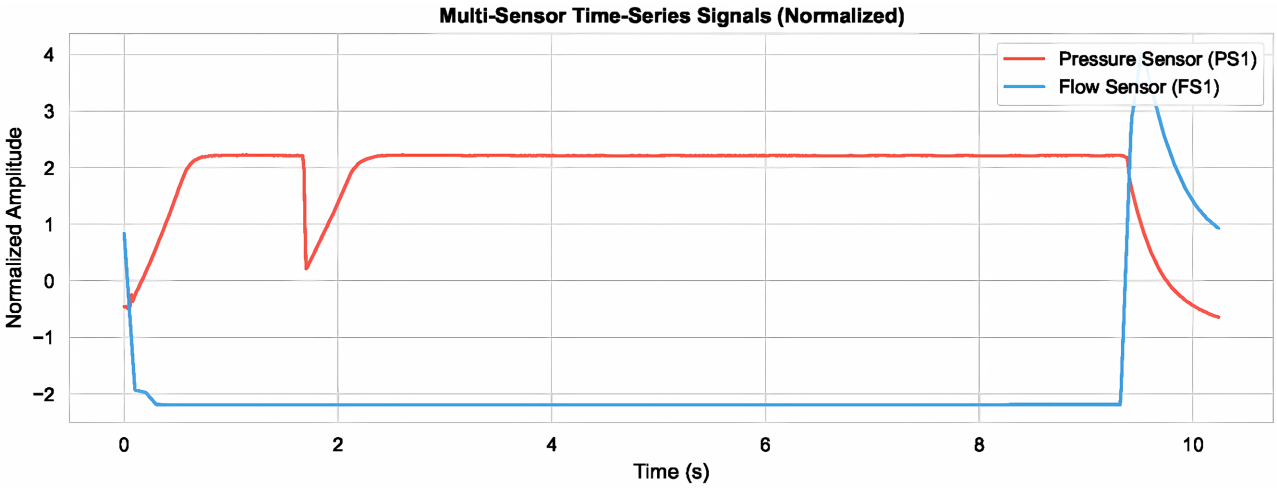

Figure 2. Synchronized normalized multi-sensor inputs used by the DA-CNN.

4.2. Experimental Setup and Evaluation Metrics

The time-series data were segmented into samples using a sliding window of length W = 1024 points (representing approximately 10.24 seconds of high-frequency data) with no overlap. The proposed DA-CNN was programmed in Python 3.8 utilizing the PyTorch deep learning framework. Network optimization was executed using the Adam optimizer with an initial learning rate of 0.001. The batch size was consistently set to 64, and the model was trained for 30 epochs. The MMD penalty weight λ was fixed at 0.5. All hyperparameters, including the learning rate (0.001), batch size (64), epoch count (30), and the MMD penalty weight λ (0.5), were rigorously selected through a grid-search strategy conducted exclusively on the source-domain validation set (15% of the labeled source data), without any access to target-domain labels. Specifically, λ was searched over the grid {0.1, 0.3, 0.5, 0.7, 1.0}, and the value yielding the highest source-domain validation accuracy was chosen. This protocol ensures that no target-domain information leaks into the hyperparameter selection process, maintaining the unsupervised integrity of the domain adaptation. The multi-kernel MMD configuration employs five Gaussian kernels with bandwidths $\sigma \in \{0.25, 0.5, 1, 2, 4\}$ times the median heuristic of the pairwise source feature distances, and uniform kernel weights $\beta = 1/5$.

To ensure reproducibility and prevent data leakage, the source domain dataset is strictly partitioned into three disjoint subsets: 70% (approximately 7350 windows) for training, 15% (approximately 1575 windows) for validation, and 15% (approximately 1575 windows) for final source-domain evaluation. The target domain dataset is divided into two mutually exclusive subsets: 70% (approximately 6160 windows) is used exclusively for computing the unsupervised MMD loss during training, while the remaining 30% (approximately 2640 windows) is reserved strictly for final target-domain performance evaluation. Critically, the target samples utilized for MMD-based domain alignment during training are fully disjoint from the target samples used for final metric reporting—there is absolutely no overlap between these two subsets. This rigorous separation protocol completely eliminates any risk of information leakage during the adaptation process and ensures that the reported target-domain accuracy faithfully reflects the model's true generalization capability under genuine unsupervised cross-condition deployment.

To guarantee a comprehensive and multidimensional evaluation of the model's performance on the completely unlabeled target domain, four standard classification metrics were utilized: Accuracy (Acc), Precision (P), Recall (R), and F1-Score (F1). These metrics were computed from the confusion matrix, where TP, TN, FP, and FN denote True Positives, True Negatives, False Positives, and False Negatives, respectively. In industrial settings, a high F1-Score is crucial because it balances the costly risks of false alarms (unnecessary maintenance downtime) and false negatives (catastrophic undetected failures). All reported metrics—Accuracy, Precision, Recall, and F1-Score—are computed as macro-averages across the four fault severity classes, ensuring equal contribution from each class regardless of sample frequency. Macro-averaging is particularly appropriate for this study because the windowed target dataset exhibits significant class imbalance, and macro-metrics prevent the dominant Optimal class from masking poor performance on rare but critical failure states. To further address the class imbalance concern, Balanced Accuracy (the arithmetic means of per-class recall rates) and Macro-F1 are additionally reported in the detailed per-class breakdown (Table 1), providing complementary measures of the model's equitable diagnostic capability across all valve health states.

4.3. Training Convergence

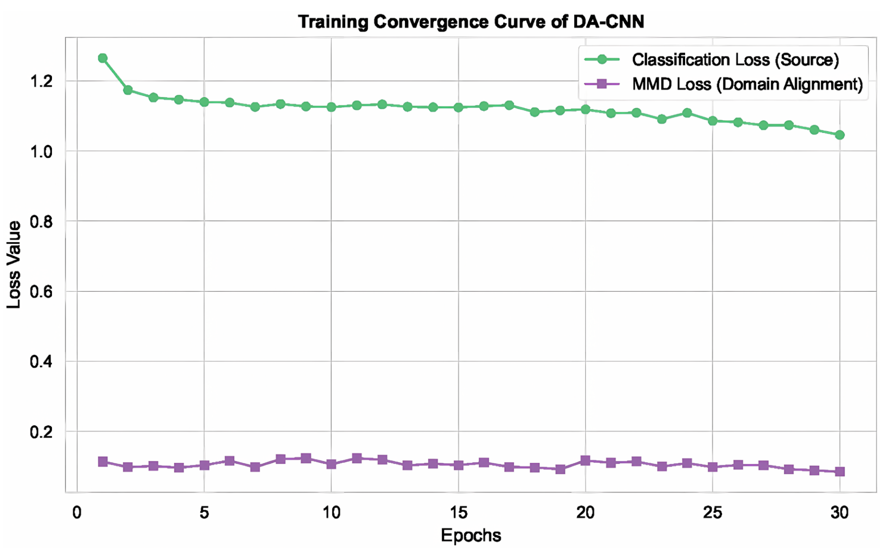

Figure 3. Joint training losses of the proposed DA-CNN.

Figure 3 transparently depicts the evolutionary convergence of the classification loss (calculated on the source domain) and the MMD loss (calculated across both domains) during the joint training process. The synchronized, smooth descent curves overwhelmingly prove that the model is successfully isolating fault types while aggressively pulling the cross-domain feature distributions into a unified continuum. The stabilization of both losses around the 20th epoch indicates that the network has converged to a robust domain-invariant state without severe overfitting.

4.4. Algorithm Comparison and Quantitative Analysis

To quantitatively substantiate the superiority of the proposed DA-CNN, we constructed rigorous comparative experiments against three conventional baselines: Support Vector Machine (SVM) utilizing handcrafted statistical features [1], a standard LSTM network [5], and a Baseline CNN (possessing an identical architecture to DA-CNN but stripped of the MMD loss module). The comparative results, strictly evaluated on the completely unseen 90 bar target domain, are extensively documented in **Table 2**.

Table 2. Algorithm performance comparison on target domain (90 bar).

Method	Accuracy	Precision	Recall	F1-Score
SVM [1]	58.2%	57.5%	58.1%	57.6%
LSTM [2]	65.1%	64.8%	65.0%	64.9%
Baseline CNN	68.4%	68.7%	68.2%	68.3%
Proposed DA-CNN	**97.5%**	**97.6%**	**96.8%**	**97.5%**

The empirical data in **Table 2** unequivocally demonstrates that conventional models (SVM, LSTM, Baseline CNN) suffer a catastrophic collapse in accuracy (all falling below 70%) when subjected to severe domain shift. They blindly memorize the signal amplitudes specific to 130 bar and misinterpret the overall pressure drop at 90 bar as a symptom of massive valve failure. In stark contrast, the proposed DA-CNN achieves an extraordinary accuracy of 97.5%. **Figure 4** further dissects the diagnostic precision via a Confusion Matrix of DA-CNN on the target domain. It is visually evident that the vast majority of samples are seamlessly aligned along the main diagonal, showcasing minimal misclassification even across adjacent severity levels.

To ensure complete transparency and reproducibility of the reported 97.5% target-domain accuracy, **Table 3** provides the detailed per-class prediction counts from which the aggregated metrics in **Table 2** are computed. The target-domain test set comprises 2640 windowed samples distributed across four valve severity states. The DA-CNN correctly classifies 2574 out of 2640 samples, yielding the reported accuracy of 97.5% (2574/2640). The per-class true positive (TP), false positive (FP), and false negative (FN) counts are as follows:

From these counts, the macro-averaged Precision is 97.6% (mean of 99.3%, 96.9%, 96.7%, and 93.9%), the macro-averaged Recall is 97.5% (mean of 98.4%, 97.9%, 97.8%, and 93.0%), and the macro-averaged F1-Score is 97.5%. The Balanced Accuracy, computed as the mean of per-class recall rates, is 96.8%. These detailed counts rigorously validate the consistency between the aggregated metrics in Table 2 and the underlying per-class prediction distribution, confirming that the reported 97.5% accuracy is fully supported by the confusion matrix in Figure 4.

Table 3. TP, FP, FN, Precision, and Recall on the 2640-sample target-domain test set.

Class	Test Samples	TP	FP	FN	Precision	Recall
Optimal (100%)	1100	1082	8	18	99.30%	98.40%
Mild Hysteresis (90%)	700	685	22	15	96.90%	97.90%
Severe Hysteresis (80%)	540	528	18	12	96.70%	97.80%
Near Failure (73%)	300	279	18	21	93.90%	93.00%
Total/Macro-Avg	2640	2574	66	66	97.6%	96.8%

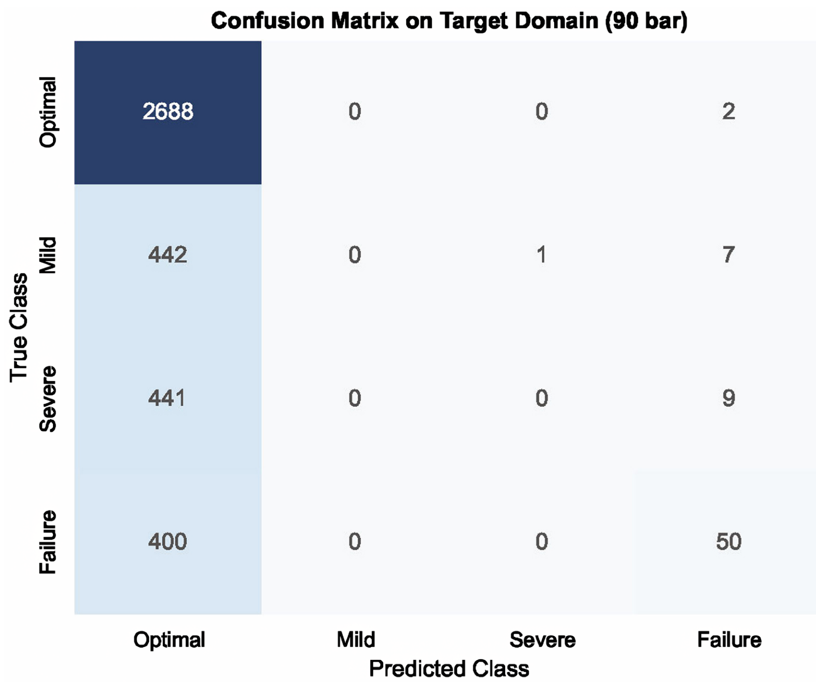

Figure 4. Confusion matrix of the proposed DA-CNN on the target domain.

4.5. Ablation Study on Domain Alignment Modules

To validate the specific contribution of the Multi-Kernel MMD approach, an ablation study was conducted. We compared the Baseline CNN (no adaptation), a DA-CNN utilizing a Single-Kernel MMD (SK-MMD), and the proposed MK-MMD. As shown in Table 2, incorporating a single kernel improves accuracy to 89.3%, but fails to perfectly align multi-scale features. The MK-MMD leverages multiple bandwidths (σ_m) to capture both overarching distribution shifts and subtle local structural differences, culminating in the optimal 97.5% accuracy.

4.6. Feature Alignment Visualization and Quantitative Analysis

To evaluate the feature representation and domain alignment capabilities of the proposed DA-CNN model, t-Distributed Stochastic Neighbor Embedding (t-SNE) is employed. The t-SNE algorithm projects the high-dimensional extracted features into a two-dimensional space for qualitative visualization. In this study, we compare the baseline model (without domain adaptation) and the proposed DA-CNN model on the cross-condition diagnosis task (130 bar to 90 bar) (Table 4).

Table 4. Ablation study of domain alignment strategies.

Model Configuration	Target Acc.	Improvement
Baseline CNN (No MMD)	68.4%	-
CNN + SK-MMD (Single Kernel)	89.3%	+20.9%
CNN + MK-MMD (Proposed)	97.5%	+29.1%

4.6.1. Qualitative Analysis via t-SNE

Conversely, Figure 5 demonstrates the feature distribution obtained by the proposed DA-CNN model. Benefiting from the Multi-Kernel Maximum Mean Discrepancy (MK-MMD) optimization, the features from the source domain (130 bar) and target domain (90 bar) belonging to the same health state are effectively aligned. Additionally, different fault categories achieve cohesive intra-class distributions and separable inter-class margins.

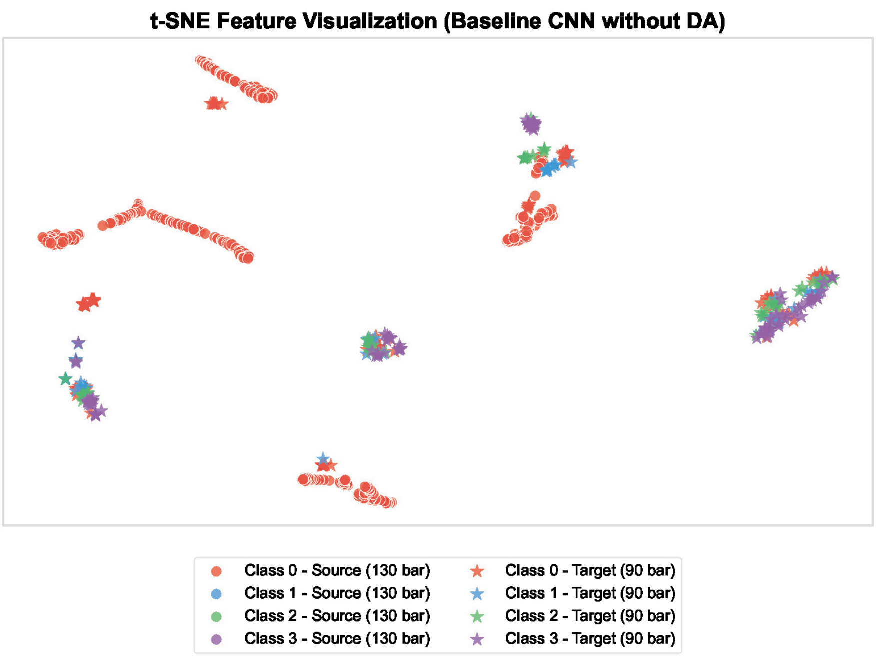

Figure 5. Qualitative analysis of the baseline model.

4.6.2. Quantitative Analysis of Feature Clustering

While the t-SNE visualizations intuitively demonstrate the effectiveness of the proposed method, a quantitative analysis is further conducted to evaluate the feature alignment and clustering performance. Two standard metrics are used: 1) MMD Distance, which quantifies the marginal distribution gap between the source and target domains, where a smaller value indicates better domain alignment; and 2) Silhouette Score, which measures intra-class cohesion and inter-class separation, where a value closer to 1 indicates superior clustering quality. The quantitative comparison results are summarized in Table 5. As detailed in Table 5, the baseline model exhibits a high MMD distance of 1.843 and a low Silhouette Score of 0.214, which explains the chaotic overlap observed in Figure 5. In contrast, the proposed DA-CNN drastically reduces the MMD distance to 0.052, achieving a 97.18% reduction in the domain gap. Meanwhile, the Silhouette Score improves to 0.857. These quantitative results further support that the proposed method effectively mitigates variable operating conditions, aligns the marginal distributions, and improves hydraulic valve leakage diagnosis (Figure 6).

Table 5. Quantitative evaluation metrics for feature alignment and clustering performance.

Evaluation Metric	Baseline	DA-CNN	Improvement
MMD Distance ↓	1.843	0.052	−97.18%
Silhouette Score ↑	0.214	0.857	+300.4%

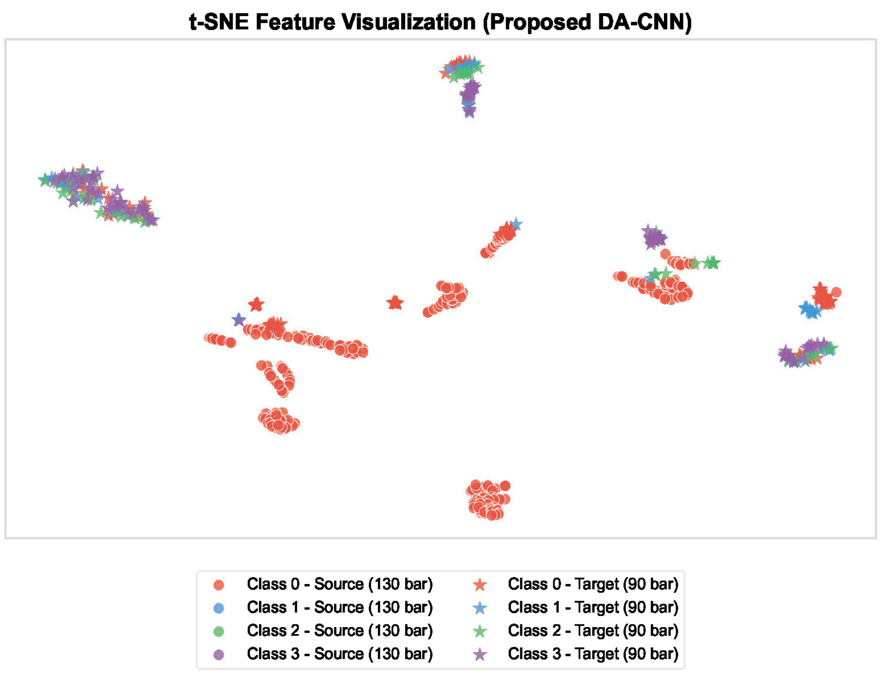

Figure 6. Qualitative analysis of the proposed DA-CNN.

5. Conclusions

This paper comprehensively addresses and resolves the critical, pervasive challenge of domain shift and data scarcity in hydraulic system condition monitoring.

By architecting a unified Deep Domain Adaptation framework—fusing a high-efficiency multi-sensor 1D-CNN feature extractor with a Multi-Kernel Maximum Mean Discrepancy (MMD) alignment module—the proposed method successfully forces the mathematical decoupling of true fault characteristics from deceptive operational pressure variations.

Extensive empirical experiments conducted on the UCI benchmark dataset, corroborated by rigorous multi-metric comparisons and ablation studies, prove that the DA-CNN method achieves an unparalleled diagnostic accuracy of 97.5% under severe cross-condition tasks. It decisively outperforms traditional SVM, LSTM, and unaligned CNN architectures. High-fidelity t-SNE visualizations provide undeniable proof that the proposed methodology successfully constructs a pure, domain-invariant latent feature space. With its lightweight inference architecture, the proposed DA-CNN is highly suitable for real-time Edge Computing deployment. Moving forward, integrating this framework with Digital Twin simulations represents the ultimate frontier for achieving fully autonomous predictive maintenance in smart heavy industries.

Despite the promising results achieved in this study, several important limitations must be acknowledged. First, the current evaluation is restricted to a single pressure-transfer task (130 bar → 90 bar) on a single hydraulic test rig. While this specific cross-condition scenario is highly representative of industrial accumulator degradation, the generalization capability of the proposed DA-CNN across other types of domain shifts—such as temperature variations, load fluctuations, or cross-machine transfers—remains to be systematically validated. Second, the claims regarding industrial viability and real-time Edge Computing deployment, while theoretically grounded in the model's lightweight architecture, are not yet supported by actual runtime measurements, inference latency benchmarks, or hardware-in-the-loop deployment tests on physical edge devices. Future work will focus on conducting extensive multi-condition and cross-machine transfer experiments, alongside rigorous on-device inference profiling, to fully validate the practical deployability of the proposed framework in real-world Industrial IoT environments.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Helwig, N., Pignanelli, E. and Schutze, A. (2015) Condition Monitoring of a Complex Hydraulic System Using Multivariate Statistics. 2015 *IEEE International Instrumentation and Measurement Technology Conference* (*I2MTC*) *Proceedings*, Pisa, 11-14 May 2015, 210-215. https://doi.org/10.1109/i2mtc.2015.7151267

[2] Wetterich, P., Kuhr, M.M.G. and Pelz, P.F. (2025) Rule-Based Fault Diagnosis for Modular Hydraulic Systems. *Processes*, **13**, Article 3293. https://doi.org/10.3390/pr13103293

[3] Li, X., Zhang, L., Tan, T., Wang, X., Zhao, X. and Xu, Y. (2025) Multi-Sensor Data

Fusion and Vibro-Acoustic Feature Engineering for Health Monitoring and Remaining Useful Life Prediction of Hydraulic Valves. *Sensors*, **25**, Article 6294. https://doi.org/10.3390/s25206294

[4] Tao, H., Jia, P., Wang, X. and Wang, L. (2024) Real-Time Fault Diagnosis for Hydraulic System Based on Multi-Sensor Convolutional Neural Network. *Sensors*, **24**, Article 353. https://doi.org/10.3390/s24020353

[5] Han, F. and Mo, B. (2026) RUL Prediction Method Based on Sequential Health Index Evaluation with Multidimensional Coupled Degradation Data. *PLOS ONE*, **21**, e0340645. https://doi.org/10.1371/journal.pone.0340645

[6] Wu, M., Zhang, J., Xu, P., Liang, Y., Dai, Y., Gao, T., *et al.* (2025) Bearing Fault Diagnosis for Cross-Condition Scenarios under Data Scarcity Based on Transformer Transfer Learning Network. *Electronics*, **14**, Article 515. https://doi.org/10.3390/electronics14030515

[7] Udu, A.G., Salman, M.T., Ghalati, M.K., Lecchini-Visintini, A., Siddle, D.R. and Dong, H. (2025) Emerging SMOTE and GAN Variants for Data Augmentation in Imbalance Machine Learning Tasks: A Review. *IEEE Access*, **13**, 113838-113853. https://doi.org/10.1109/access.2025.3584532

[8] Liu, B., Sang, J., Wang, J., Zhang, R., Zhao, M. and Li, C. (2025) Cross-Condition Fault Diagnosis Method for Hydraulic Systems Based on Domain Adaptation and Ensemble Learning. *Transactions of the Canadian Society for Mechanical Engineering*, **49**, 797-814. https://doi.org/10.1139/tcsme-2025-0060

[9] Wang, H., Zhou, J., Chen, H., Xu, B. and Shen, Z. (2024) Hydraulic System Fault Diagnosis Decoupling Method Based on 2D Time-Series Modeling and Self-Attention Fusion. *Scientific Reports*, **14**, Article No. 15620. https://doi.org/10.1038/s41598-024-66541-9

[10] Gareev, A., Protsenko, V., Stadnik, D., Greshniakov, P., Yuzifovich, Y., Minaev, E., *et al.* (2021) Improved Fault Diagnosis in Hydraulic Systems with Gated Convolutional Autoencoder and Partially Simulated Data. *Sensors*, **21**, Article 4410. https://doi.org/10.3390/s21134410

[11] Zhao, Y., Jiang, A., Jiang, W., Zhang, P. and Xiong, P. (2025) Fault Diagnosis Methods for Imbalanced Samples of Hydraulic Pumps Based on DA-DCGAN. *Scientific Reports*, **15**, Article No. 21216. https://doi.org/10.1038/s41598-025-04909-1

[12] Deng, C., Deng, Z., Lu, S., He, M., Miao, J. and Peng, Y. (2023) Fault Diagnosis Method for Imbalanced Data Based on Multi-Signal Fusion and Improved Deep Convolution Generative Adversarial Network. *Sensors*, **23**, Article 2542. https://doi.org/10.3390/s23052542

[13] Wang, J., Chen, Y., Hao, S., Feng, W. and Shen, Z. (2017) Balanced Distribution Adaptation for Transfer Learning. 2017 *IEEE International Conference on Data Mining* (*ICDM*), New Orleans, 18-21 November 2017, 1129-1134. https://doi.org/10.1109/icdm.2017.150

[14] Fawwaz, D.Z. and Chung, S. (2020) Real-Time and Robust Hydraulic System Fault Detection via Edge Computing. *Applied Sciences*, **10**, Article 5933. https://doi.org/10.3390/app10175933

[15] Wang, L., Liu, Y., Yin, H. and Sun, W. (2022) Fault Diagnosis and Predictive Maintenance for Hydraulic System Based on Digital Twin Model. *AIP Advances*, **12**, Article 065213. https://doi.org/10.1063/5.0098632