A Bidirectional Structure Constraint framework for Domain Generalization in Intelligent Fault Diagnosis

Wenjing Zhou[1], Liang Chen[2*], Hong Zhuang[3], Qitong Chen[4]

[1,2,3,4] *School of Mechanical and Electrical Engineering, Soochow University, Suzhou, 215000, China*
wjzhou1@stu.suda.edu.cn
ChenL@suda.edu.cn
hzhuang@suda.edu.cn
qtchen0730@163.com

ABSTRACT

Achieving robust generalization in intelligent fault diagnosis under diverse industrial conditions remains challenging. Most domain generalization (DG) methods focus on either feature compactness or category separation, seldom addressing both in a unified framework. To overcome this, we propose a Bidirectional Structure Constraint (BSC) framework comprising Momentum Feature Alignment (MFA) and Category Anchor Separation (CAS). MFA employs a momentum-driven strategy to capture domain-invariant features for each category, while CAS encourages learnable class anchors to repel each other in latent space, enhancing class separability. These objectives are jointly optimized in a multi-loss framework, enabling the model to learn representations that are both intra-class compact and inter-class distinct. Experiments on the Shandong University of Science and Technology (SDUST) rotating machinery fault diagnosis dataset show that BSC significantly improves cross-domain generalization.

1. INTRODUCTION

Intelligent fault diagnosis based on deep learning has been widely applied in industrial scenarios due to its ability to automatically extract representative features from raw sensor signals. However, most existing methods rely on the assumption that training and test data are drawn from the same distribution, which often does not hold in real-world applications(Zhao, Zio, & Shen, 2024). In practical scenarios, variations in operating conditions, sensor positions, or environmental noise lead to domain shifts, significantly degrading model performance when deployed in unseen domains (Zhou, Chu, Chen, Shen, & Chen, 2025).

To address this issue, various domain generalization (DG) approaches have been proposed, which aim to learn representations that generalize well across unseen domains. These methods can be roughly categorized into five key directions: data augmentation, domain alignment, feature disentanglement, meta-learning, and model interpretability (Xiao et al., 2025). Among these, domain alignment has been one of the most extensively studied strategies, which seeks to reduce the distribution discrepancy between source and target domains in a shared latent space (Chen, Li, Wu, Chen, & Shen, 2024). While effective in mitigating domain shift at a global level, many alignment-based methods perform only distribution-level matching, such as aligning marginal or conditional statistics, without explicitly preserving the class-level semantic structure. For example, Deep Correlation Alignment (Deep CORAL) (Sun & Saenko, 2016) aligns global feature distributions by minimizing the difference in covariance matrices between source and target domains. AND-Mask Gradient Aggregation (ANDMask) (Parascandolo, Neitz, Orvieto, Gresele, & Schölkopf, 2020) performs global feature alignment by enforcing consistent gradient directions across domains, promoting invariant representation learning without explicitly modeling class-level structure. Ma et al. (Ma et al., 2024) introduced a method that combines domain adversarial training with maximum mean discrepancy (MMD) to align global feature distributions across multiple domains. Pu et al. (Pu et al., 2024) proposed a Domain-Relevant Joint Distribution Alignment (DRJDA) strategy, which aligns joint and product distributions across domains via minimizing a relative chi-square divergence, effectively reducing global domain shift. These methods improve marginal alignment but do not ensure class-wise feature consistency. As a result, samples from different classes may be incorrectly aligned, leading to ambiguous decision boundaries and degraded classification performance.

Thus, a growing number of recent methods focus on promoting intra-class compactness by encouraging features from the

same class to cluster across domains. For instance, Wang et al. (B. Wang, Wen, Li, & Gao, 2023) proposed the Adaptive Class Center Generalization Network (ACCGN), which enforces class-wise feature compactness by aligning adaptive class centers across multiple source domains, while mitigating domain discrepancy through sparse domain regression. Shi et al.(Shi et al., 2023) proposed a Reliable Feature-Assisted Contrastive Generalization Network (RFACGN), which integrates contrastive learning with sample-adaptive weighting to enhance domain-invariant representation and improve feature discrimination under unseen machines and working conditions. Jia et al. (Jia, Chow, Wang, & Ma, 2024) proposed a Dynamic Balanced Dual Prototypical Network (DBDP-Net), which aligns both class and domain prototypes to promote intra-class compactness and domain-invariant representation learning. Qian et al. (Qian, Luo, & Qin, 2024) proposed an Adaptive Intermediate Class-wise Distribution Alignment (AICDA) model that simultaneously aligns global and class-conditional distributions via an adaptive intermediate representation, enhancing domain confusion and reducing negative transfer. While these methods effectively enhance intra-class consistency, they often lack mechanisms to enforce inter-class separability, which may lead to feature overlap and ambiguous decision boundaries. This imbalance between intra-class alignment and inter-class discrimination remains a fundamental limitation of existing domain alignment frameworks.

To address the aforementioned limitations, we propose a novel Bidirectional Structure Constraint (BSC) framework that jointly enforces intra-class compactness and inter-class separability under domain shifts. Specifically, we design a Momentum Feature Alignment (MFA) mechanism to ensure that features from the same class maintain consistent activation patterns across domains by aligning their class-specific feature contributions with a momentum-updated reference. Meanwhile, a Category Anchor Separation (CAS) module explicitly encourages the semantic distance between class anchors in the latent space to increase, thereby improving class distinctiveness and preventing feature overlap. By combining these two objectives in a unified multi-loss framework, BSC preserves class semantics while mitigating domain-induced feature distortions, ultimately enhancing generalization to unseen target domains.The main contributions of this work are summarized as follows:

- We propose a BSC framework for DG in intelligent fault diagnosis, which simultaneously enforces intra-class compactness and inter-class separability through class-level structural modeling.
- We design two complementary modules, MFA and CAS, where MFA aligns class-specific feature contributions across domains, and CAS encourages separation between class anchors in the latent space.
- We validate the proposed method on the SDUST rotating machinery dataset, where it consistently outperforms state-of-the-art DG approaches in cross-domain diagnostic scenarios.

2. METHODOLOGY

2.1. Momentum Feature Alignment

To enforce intra-class consistency and promote domain-invariant representation learning, we introduce a Momentum Feature Alignment (MFA) mechanism, as shown in Figure 1. This mechanism aligns class-specific feature contributions in the latent space by leveraging a dynamically updated reference. Specifically, we interpret the classification layer weights as learned reference vectors for each class, referred to as anchors. For any input sample, its feature contribution under a particular anchor is defined as the element-wise multiplication between its feature vector and corresponding anchor. This operation reflects the extent to which each feature dimension is emphasized by the anchor during classification, essentially quantifying the anchor's reliance on specific latent factors.

To promote stability and alignment across domains, we maintain a class-wise momentum-updated contribution repository, denoted as $\mathbf{M}$. For each class, this repository stores the average feature contributions computed from historical samples. Formally, for class i, the feature contribution from anchor i for sample j is computed as:

$$\mathbf{m}_{i,j} = \mathbf{a}_i \odot \mathbf{z}_j, \tag{1}$$

where $\mathbf{a}_i$ denotes the anchor for class i, $\mathbf{z}_j$ is the extracted feature vector of sample j, and $\odot$ indicates element-wise multiplication.

The repository for class i is updated via an Exponentially Moving Average (EMA) using a momentum factor μ:

$$\mathbf{M}_i^{(t+1)} = (1-\mu)\cdot\mathbf{M}_i^{(t)} + \mu\cdot\operatorname{mean}\left(\{\mathbf{m}_{i,j} \mid y_j = i\}\right). \tag{2}$$

The MFA loss is computed by measuring the L2 norm between the current contribution of each sample and the corresponding class repository vector:

$$\mathcal{L}_{\text{MFA}}^{i} = \frac{1}{N_i}\sum_{j:y_j=i}\left\|\mathbf{m}_{i,j} - \mathbf{M}_i\right\|^2, \tag{3}$$

where N_i is the number of samples in class i within the current mini-batch. The MFA loss is then obtained by averaging over all present classes:

$$\mathcal{L}_{\text{MFA}} = \frac{1}{|\mathcal{Y}_b|}\sum_{i\in\mathcal{Y}_b}\mathcal{L}_{\text{MFA}}^{i}, \tag{4}$$

where $\mathcal{Y}_b$ is the set of class labels in the current mini-batch.

This mechanism ensures that, for samples of the same class,

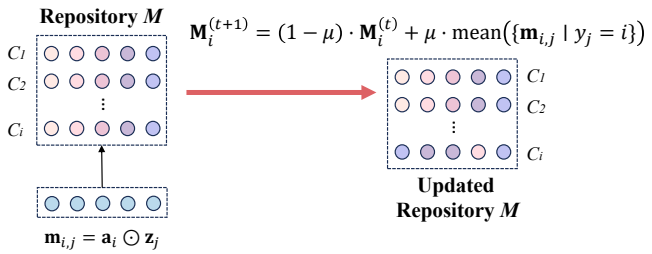

Figure 1. Operating principle of MFA method.

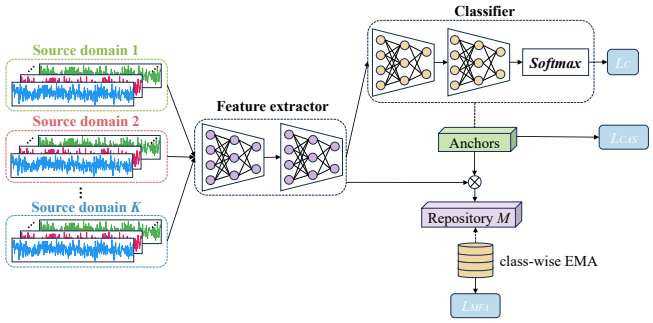

Figure 2. Schematic diagram of the overall structure of the BSC model.

their features contribute to classification in a consistent way under the same class anchor. This reduces the variation among samples of the same class across domains and promotes a more consistent latent representation. Consequently, MFA narrows the distributional gap between domains and improves training stability under domain shifts.

2.2. Category Anchor Separation

To enhance inter-class separability and prevent category overlap in the latent space, we introduce a Category Anchor Separation (CAS) mechanism as a complementary counterpart to MFA. While MFA promotes compactness by aligning feature contributions within each class, CAS explicitly encourages separation between different class anchors in the latent space. The core idea is that increasing the distance between anchors helps the model establish clearer decision boundaries, reducing ambiguity especially under domain shifts.

In our framework, each class is associated with a learnable anchor vector, which also serves as the weight of the classifier. These anchors act as class-level representatives that guide feature alignment. To enforce separation, we compute the pairwise distance between all anchors and penalize those that are too close. Specifically, given two class anchors $\mathbf{a}_i$ and $\mathbf{a}_j$, we define their similarity score as the negative distance between them, scaled by a temperature factor τ:

$$\mathbf{S}_{i,j} = -\frac{d(\mathbf{a}_i, \mathbf{a}_j)}{\tau}, \tag{5}$$

where $d(\cdot,\cdot)$ denotes the L2 distance, and τ is the temperature hyperparameter to modulate the sharpness of the similarity distribution.

To ensure that each anchor is only compared with other classes, self-comparisons are removed via masking. The CAS loss is then defined as:

$$\mathcal{L}_{\mathrm{CAS}} = \log\left(1 + \sum_{i\neq j}\exp(\mathbf{S}_{i,j})\right). \tag{6}$$

By minimizing this loss, the model learns to spread class anchors across the latent space, which encourages each class to occupy a distinct and non-overlapping region in the latent space, thereby reinforcing inter-class discrimination.

2.3. Bidirectional Structure Constraint Framework

The BSC framework jointly optimizes three complementary objectives: classification accuracy, intra-class consistency, and inter-class separability, as shown in Figure 2. These are enforced through a combination of standard cross-entropy loss, MFA loss and CAS loss, respectively.

The total loss function of BSC can be expressed as:

$$\mathcal{L}_{\mathrm{BSC}} = \mathcal{L}_{\mathrm{C}} + \alpha\mathcal{L}_{\mathrm{MFA}} + \beta\mathcal{L}_{\mathrm{CAS}}, \tag{7}$$

where $\mathcal{L}_{\mathrm{C}}$ is the standard cross-entropy loss between the predicted logits and ground truth labels. The coefficients α and β are hyperparameters controlling the relative importance of different loss.

3. EXPERIMENT

3.1. Datasets

3.1.1. SDUST Rotating Machinery Dataset

We conduct our experiments on the SDUST (J. Wang et al., 2024) rotating machinery fault diagnosis dataset, which is a widely used benchmark for DG in intelligent fault diagnosis. The SDUST test platform is shown in Figure 3. This dataset contains vibration signals collected from rotating machinery under different working conditions, simulating various load levels and fault types across multiple operational domains. To evaluate the cross-domain generalization performance of our method, we select four representative working conditions from the original dataset, covering diverse operating states and fault types, as shown in Table 1.

Based on these selected conditions, we construct four DG tasks by altering the source-target domain combinations. Each task involves training the model on data from multiple source domains and testing it on an unseen target domain

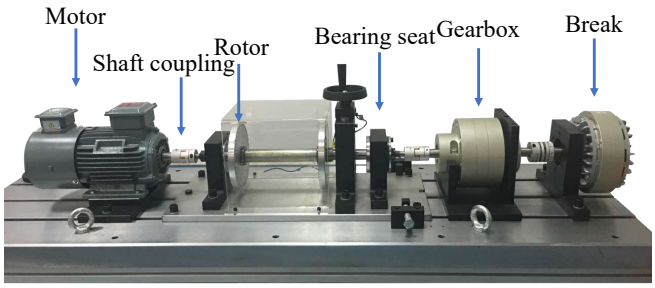

Figure 3. The SDUST bearing platform used for experiments.

Table 1. Four operating conditions.

Condition	Speed(RPM)	Load(N)
A	1500	0
B	2000	20
C	2500	40
D	3000	60

without using any target data during training. The details of the task settings are shown in Table 2. This experimental design allows us to comprehensively assess the domain robustness and transferability of the proposed BSC framework under realistic and challenging diagnostic scenarios.

3.1.2. SCARA Robot Dataset

We further evaluate the proposed BSC framework on the Selective Compliance Assembly Robot Arm (SCARA) dataset to assess its generalization performance across different mechanical systems. The SCARA fault diagnosis platform is developed at Soochow University and is designed to collect operational data from IR-C8 series high-speed SCARA robots, as shown in Figure 4. This dataset focuses on the ball screw system, a critical motion component in industrial robots.

During experiments, current signals are acquired from the ball screw drive at a sampling frequency of 8 kHz, and each sample contains 1024 data points. Four motor load conditions are considered: 0 kg, 3 kg, 6 kg, and 9 kg, representing distinct operational domains. Four health states are defined to characterize different fault types: normal (N), missing balls (L), helical nut stuck (S1) and spline nut stuck (S2).

For each category, 1600 samples are used for training and 800 samples for testing. Four domain generalization tasks are constructed by varying the load conditions of the source and target domains, as summarized in Table 3. For example,

Table 2. Details of 4 diagnosis tasks of SDUST.

Task	Source domain	Target domain
Task1	B,C,D	A
Task2	C,D	B
Task3	A,B	C
Task4	A,B,C	D

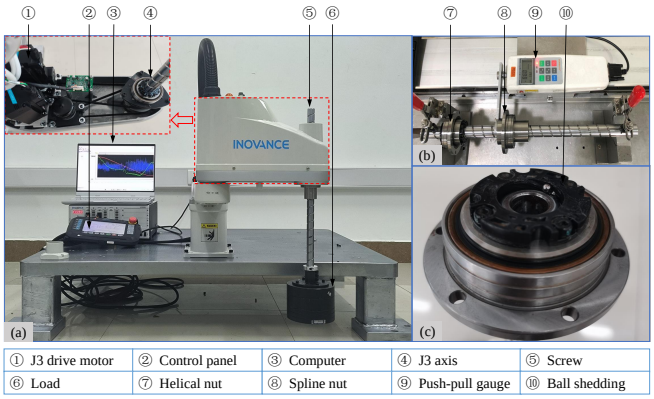

Figure 4. The SCARA bearing platform used for experiments.

Table 3. Details of 4 diagnosis tasks of SCARA.

Task	Source domain	Target domain
T0	3kg,6kg,9kg	0kg
T3	0kg,6kg,9kg	3kg
T6	0kg,3kg,9kg	6kg
T9	0kg,3kg,6kg	9kg

in Task T0, data collected under 3 kg, 6 kg, and 9 kg are used as source domains, while data under 0 kg serve as the target domain.

The SCARA dataset differs from SDUST in both signal modality (current vs. vibration) and mechanical structure (robotic ball screw vs. rotating machinery), thus providing a complementary benchmark to verify the robustness and cross-system adaptability of the proposed BSC framework.

3.2. Implementation Settings

In our implementation, we adopt a lightweight variant of ResNet (He, Zhang, Ren, & Sun, 2016) as the feature extractor, where only a single residual block is retained to reduce model complexity and facilitate efficient training. The overall architecture includes three key hyperparameters: μ in (2), α and β in (7).

While a lightweight ResNet was adopted to reflect edge-oriented constraints and isolate the contribution of BSC, the objective is backbone-agnostic and can scale to deeper CNNs or Transformer-style feature extractors, with expected accuracy gains at the expense of higher compute; a systematic backbone study across deployment budgets is left for future work.

To ensure fair and reproducible model selection, we follow the best practices recommended by the DomainBed (Gulrajani & Lopez-Paz, 2020) benchmark. Concretely, we adopt a domain-wise data partitioning scheme, in which 80% of the samples in each source domain are allocated for training and the remaining 20% are held out for validation. Hyper-

Table 4. Hyperparameter selecting.

Hyperparameter	Default	Range
momentum factor μ	10^{-2}	$[10^{-4},10^{-1}]$
loss weight α	1	$[10^{-3},1]$
loss weight β	1	$[10^{-3},1]$
step	2000	{1000,2000,...,5000}
learning rate	10^{-3}	$[10^{-4},10^{-2}]$
batch size	32	{32,64}

parameter optimization is performed via the Optuna framework. A predefined search space is specified for key parameters, from which 30 candidate configurations are sampled. Each configuration is independently evaluated over four runs, and the one with the best average validation accuracy is selected as the final setting.

The optimal configuration obtained from the Optuna search corresponds to $\alpha = 0.3843$ and $\beta = 0.1009$, yielding the highest average diagnostic accuracy across all tasks. A further inspection reveals that excessively large α values cause the model to overemphasize intra-class compactness, which may suppress inter-class diversity and lead to feature redundancy. Conversely, an overly large β amplifies the inter-class repulsion effect, resulting in unstable convergence and degraded feature alignment within the same class. The chosen balance effectively harmonizes these two objectives, ensuring both consistent intra-class alignment and clear inter-class separability, while maintaining stable optimization and robust cross-domain generalization.

3.3. Comparative experimental analysis

To evaluate the effectiveness of the proposed BSC framework, we compare it against several representative domain generalization (DG) methods on the SDUST rotating machinery dataset and the SCARA robot dataset. The selected baselines include: Empirical Risk Minimization (ERM) (Vapnik, 2013), Domain-Adversarial Neural Network (DANN) (Ganin & Lempitsky, 2015), Domain-Adversarial Conditional Entropy (DANNCE) (Sicilia, Zhao, & Hwang, 2023), Invariant Risk Minimization (IRM) (Arjovsky, Bottou, Gulrajani, & Lopez-Paz, 2019), Mixup (Yan, Song, Li, Zou, & Ren, 2020), Maximum Mean Discrepancy (MMD) (Li, Pan, Wang, & Kot, 2018) and causal invariant representation learning based on maximum mean discrepancy (CauslRL_MMD) (Chevalley, Bunne, Krause, & Bauer, 2022).

Table 5 summarizes the diagnostic accuracy on four DG tasks. As shown, the proposed BSC framework achieves the highest average accuracy of 88.0%, surpassing all baseline methods. Notably, it also yields the best performance in two out of four tasks (Task1, Task3, and Task4), and ranks second in Task2 and Task3.

Compared to the ERM baseline, which does not incorporate any domain alignment strategy, BSC improves the average accuracy by 1.7 percentage points, demonstrating the advantage of structure-aware class-level modeling. While DANN and DANNCE introduce domain adversarial training, their performance fluctuates across tasks and remains lower than BSC, indicating the limitations of global distribution alignment without explicit class separation. IRM and Mixup exhibit the weakest performance on Task2, ranking at the bottom among all methods and revealing their limited generalization capability under complex domain shifts. This suggests that regularization-based or invariant risk approaches may not fully capture domain-specific semantics. CauslRL_MMD shows competitive performance on Task2 and achieves the second-best overall average (87.4%), highlighting the potential of causality-based representation learning. Nevertheless, BSC consistently outperforms it across most tasks, suggesting that jointly enforcing intra-class compactness and inter-class separability leads to better generalization under complex domain shifts.

Table 5. Diagnostic results of SDUST comparative experiments.

Methods	Task1	Task2	Task3	Task4	Average
ERM	71.6±1.9	88.2±3.1	89.4±1.1	95.8±1.2	86.3
DANN	73.1±2.0	84.9±2.3	87.4±1.5	94.4±1.6	85.0
DANNCE	72.3±2.4	83.8±3.3	88.2±1.5	95.1±1.8	84.9
IRM	72.6±2.9	82.7±2.8	84.8±2.1	86.7±4.4	81.7
Mixup	70.3±1.7	73.9±2.0	**94.1±1.9**	91.4±2.4	82.4
MMD	72.7±1.4	83.7±3.2	88.2±0.7	93.2±1.7	84.5
CauslRL_MMD	72.4±2.1	**93.2±2.1**	89.1±1.0	94.8±1.6	87.4
BSC	**76.4±1.9**	88.5±0.8	90.2±0.9	**96.7±0.4**	**88.0**

In terms of computational efficiency, the proposed BSC framework introduces negligible additional overhead compared with standard ERM-based training. The momentum update in MFA and the pairwise anchor computation in CAS increase the overall computation by approximately 4–6%. This moderate increase is attributed mainly to the EMA updates in MFA. The inference speed remains effectively unchanged, as both modules are only active during training.

To further assess the cross-system generalization of the BSC framework, additional experiments are performed on the SCARA robot dataset. The diagnostic results on four DG tasks are summarized in Table 6.

As observed, the proposed BSC framework again achieves the best overall performance, with average accuracy of 92.75% across the four tasks. The result is higher than that of all baseline methods. In particular, BSC shows remarkable stability under domain shifts with large load variations, where adversarial approaches such as DANN and DANNCE experience significant performance degradation. The results suggest that the bidirectional structural constraints effectively preserve class semantics and mitigate domain-induced distor-

Table 6. Diagnostic results of SCARA comparative experiments.

Methods	T0	T3	T6	T9	Average
ERM	89.3±1.5	92.3±1.7	92.7±1.2	77.1±2.1	87.8
DANN	91.0±1.8	94.3±2.2	88.7±1.5	56.6±2.8	82.6
DANNCE	89.0±2.9	92.2±2.4	86.8±1.9	75.4±3.1	85.8
IRM	56.5±2.8	93.7±2.6	78.3±2.1	46.6±3.6	68.8
Mixup	95.8±1.8	**95.2±2.0**	**96.5±1.7**	66.9±1.8	88.6
MMD	88.9±1.1	91.8±1.5	86.1±1.2	43.0±1.2	77.4
CauslRL_MMD	56.9±2.3	66.8±1.0	62.4±1.6	52.3±2.3	59.6
BSC	**97.4±0.8**	94.2±1.0	95.2±0.6	**84.3±1.1**	**92.8**

tions, even when applied to a completely different mechanical system.

Notably, the strong performance on the SCARA dataset demonstrates that the proposed BSC framework is not limited to vibration-based diagnostics but can be extended to current-signal-based and robotic systems, reflecting its scalability, robustness, and practical applicability in diversified industrial scenarios.

3.4. Visualization Analysis

To further illustrate the effectiveness of the proposed BSC framework in learning discriminative and domain-invariant representations, we employ t-SNE in Task 1 to visualize the feature distributions of different models on the test domains. Figure 5 presents the two-dimensional embeddings generated by t-SNE for eight different methods: ERM, DANN, DANNCE, IRM, Mixup, MMD, CauslRL_MMD, and our proposed BSC.

From the visualizations, it is evident that ERM, which does not account for domain shift, produces highly scattered and overlapping feature clusters, indicating poor generalization. DANN and DANNCE, which incorporate domain-adversarial training, exhibit slightly improved separation, yet still suffer from ambiguous class boundaries and inconsistent feature distributions. IRM and Mixup also fail to form compact and well-separated clusters, with significant overlap across categories, especially in complex domain shift conditions. MMD and CauslRL_MMD achieve relatively better alignment, with reduced inter-domain discrepancy. However, their class-wise boundaries remain fuzzy, and some categories still exhibit significant intra-class dispersion. In contrast, the proposed BSC framework demonstrates the most compact and clearly separated feature clusters. Each class forms a tight group, with minimal overlap across categories, highlighting the effectiveness of BSC in simultaneously promoting intra-class compactness and inter-class separability.

3.5. Ablation Study

To investigate the individual contributions of the MFA and CAS modules, we conduct ablation experiments based on the ERM baseline. Specifically, we construct three variants: ERM+MFA, ERM+CAS, and the full BSC model that integrates both modules. The diagnostic results on four cross-domain tasks from the SDUST dataset are reported in Table 7.

Table 7. Diagnostic results of ablation experiments.

Methods	Task1	Task2	Task3	Task4	Average
ERM	71.6±1.9	88.2±3.1	89.4±1.1	95.8±1.2	86.3
ERM+MFA	75.6±2.1	87.8±1.0	**90.4±1.1**	**96.8±0.8**	87.7
ERM+CAS	74.1±2.0	87.3±2.4	90.1±2.1	97.0±1.7	87.1
BSC	**76.4±1.9**	**88.5±0.8**	90.2±0.9	96.7±0.4	**88.0**

Compared to the vanilla ERM baseline, incorporating either MFA or CAS consistently improves the domain generalization performance. ERM+MFA achieves an average accuracy of 87.7%, indicating that aligning class-specific feature contributions helps maintain intra-class consistency across domains. ERM+CAS also brings performance gains, with an average accuracy of 87.1%, demonstrating the benefit of enhancing inter-class separability through anchor repulsion. The full BSC model, which jointly optimizes both objectives, further improves the average accuracy to 88.0%, outperforming all ablated variants. This confirms the complementary nature of MFA and CAS, and highlights the effectiveness of jointly enforcing intra-class compactness and inter-class distinctiveness in enhancing generalization to unseen domains.

4. CONCLUSION

In this paper, we proposed a novel Bidirectional Structure Constraint (BSC) framework for DG in intelligent fault diagnosis. Unlike existing approaches that focus solely on either intra-class compactness or inter-class separation, our method integrates both objectives through a unified design of class-level structural constraints. Specifically, the MFA module aligns class-specific feature contributions across domains to improve intra-class consistency, while the CAS module promotes semantic separation among class anchors to enhance inter-class discriminability. These two objectives are jointly optimized within a multi-loss learning paradigm, enabling the model to learn robust and transferable representations that generalize well to unseen domains. Extensive experiments conducted on the SDUST rotating machinery dataset demonstrate that our method consistently outperforms existing DG techniques, verifying its effectiveness and applicability under diverse industrial conditions.

In future research, we will further explore the integration of the proposed Momentum Feature Alignment (MFA) mechanism with edge computing and real-time monitoring systems, enabling on-device inference and adaptive updating for low-latency fault diagnosis in industrial environments. Moreover, we plan to extend the scalability of the BSC framework to the fault diagnosis of industrial robots, where complex multi-joint dynamics and heterogeneous sensory data pose addi-

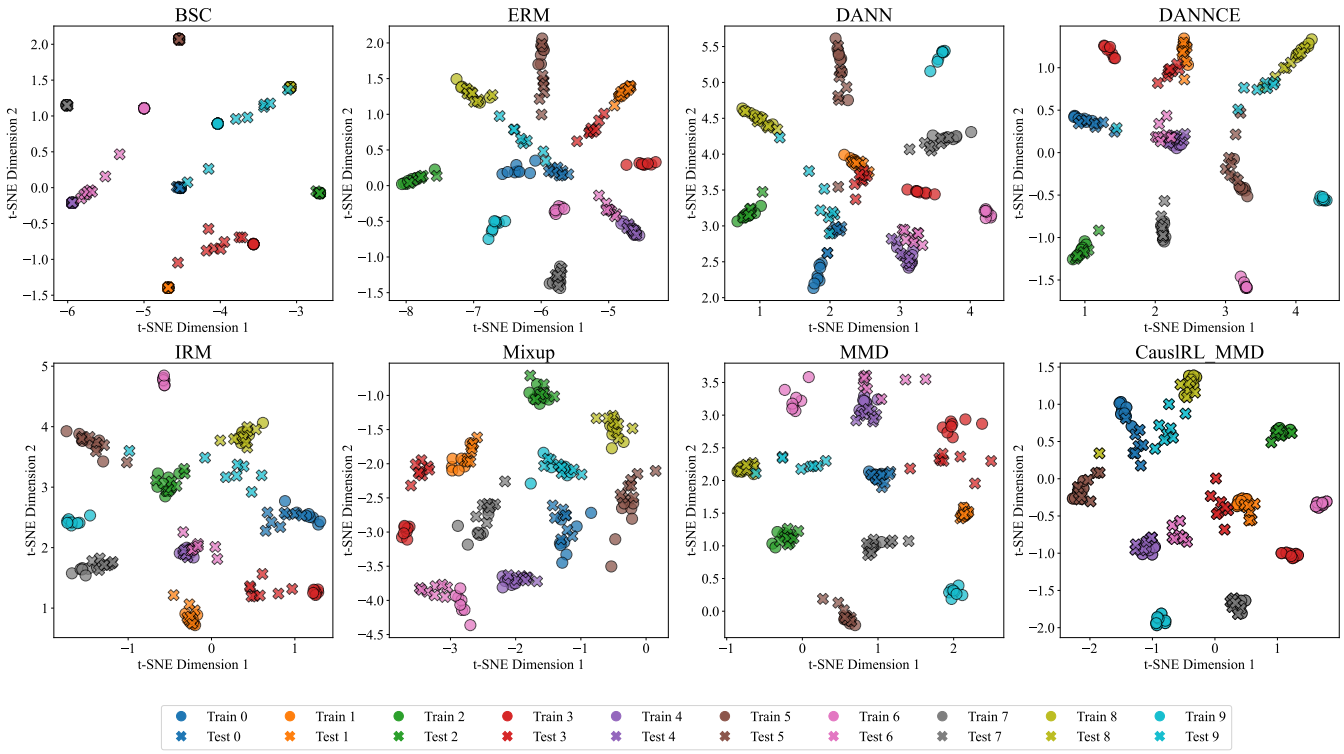

Figure 5. The t-SNE visualization of the feature representations learned by different algorithms on Task 1.

tional challenges for domain generalization. By combining MFA and CAS with edge intelligence and continual learning paradigms, our future work aims to establish a more robust, scalable, and adaptive DG framework for real-time intelligent health monitoring in industrial applications.

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China under Grants 52375114, 52272440, and 52175056, and by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under Grant KYCX25_3465.

REFERENCES

Arjovsky, M., Bottou, L., Gulrajani, I., & Lopez-Paz, D. (2019). Invariant risk minimization. *arXiv preprint arXiv:1907.02893*.

Chen, Q., Li, Q., Wu, S., Chen, L., & Shen, C. (2024). Fault diagnosis for ball screws in industrial robots under variable and inaccessible working conditions with non-vibration signals. *Advanced Engineering Informatics*, *62*, 102617.

Chevalley, M., Bunne, C., Krause, A., & Bauer, S. (2022). Invariant causal mechanisms through distribution matching. *arXiv preprint arXiv:2206.11646*.

Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. In *International conference on machine learning* (pp. 1180–1189).

Gulrajani, I., & Lopez-Paz, D. (2020). In search of lost domain generalization. *arXiv preprint arXiv:2007.01434*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 770–778).

Jia, L., Chow, T. W., Wang, Y., & Ma, J. (2024). Dynamic balanced dual prototypical domain generalization for cross-machine fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *73*, 1–10.

Li, H., Pan, S. J., Wang, S., & Kot, A. C. (2018). Domain generalization with adversarial feature learning. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 5400–5409).

Ma, Z., Fu, L., Dun, G., Tan, D., Xu, F., & Zhang, L. (2024). A robust domain distribution alignment discriminative network driven by physical samples for rotor-bearing fault diagnosis. *Knowledge-Based Systems*, *300*, 112216.

Parascandolo, G., Neitz, A., Orvieto, A., Gresele, L., & Schölkopf, B. (2020). Learning explanations that are hard to vary. *arXiv preprint arXiv:2009.00329*.

Pu, H., Teng, S., Xiao, D., Xu, L., Luo, J., & Qin, Y. (2024).

Domain generalization for machine compound fault diagnosis by domain-relevant joint distribution alignment. *Advanced Engineering Informatics*, *62*, 102771.

Qian, Q., Luo, J., & Qin, Y. (2024). Adaptive intermediate class-wise distribution alignment: A universal domain adaptation and generalization method for machine fault diagnosis. *IEEE transactions on neural networks and learning systems*, *36*(3), 4296–4310.

Shi, Z., Chen, J., Zhang, X., Zi, Y., Li, C., & Chen, J. (2023). A reliable feature-assisted contrastive generalization net for intelligent fault diagnosis under unseen machines and working conditions. *Mechanical Systems and Signal Processing*, *188*, 110011.

Sicilia, A., Zhao, X., & Hwang, S. J. (2023). Domain adversarial neural networks for domain generalization: When it works and how to improve. *Machine Learning*, *112*(7), 2685–2721.

Sun, B., & Saenko, K. (2016). Deep coral: Correlation alignment for deep domain adaptation. In *European conference on computer vision* (pp. 443–450).

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Wang, B., Wen, L., Li, X., & Gao, L. (2023). Adaptive class center generalization network: A sparse domain-regressive framework for bearing fault diagnosis under unknown working conditions. *IEEE Transactions on Instrumentation and Measurement*, *72*, 1–11.

Wang, J., Zhang, X., Zhang, Z., Han, B., Jiang, X., Bao, H., & Jiang, X. (2024). Attention guided multi-wavelet adversarial network for cross domain fault diagnosis. *Knowledge-based systems*, *284*, 111285.

Xiao, Y., Shao, H., Yan, S., Wang, J., Peng, Y., & Liu, B. (2025). Domain generalization for rotating machinery fault diagnosis: A survey. *Advanced Engineering Informatics*, *64*, 103063.

Yan, S., Song, H., Li, N., Zou, L., & Ren, L. (2020). Improve unsupervised domain adaptation with mixup training. *arXiv preprint arXiv:2001.00677*.

Zhao, C., Zio, E., & Shen, W. (2024). Domain generalization for cross-domain fault diagnosis: An application-oriented perspective and a benchmark study. *Reliability Engineering & System Safety*, *245*, 109964.

Zhou, W., Chu, L., Chen, Q., Shen, C., & Chen, L. (2025). Adversarial weighted multi-source domain adaptation for compound fault diagnosis. *Measurement Science and Technology*.

BIOGRAPHIES

Wenjing Zhou received the B.Eng. degree in light chemical engineering from Soochow University, Suzhou, China, in 2023. She is currently pursuing the M.S. degree in control science and engineering at Soochow University, Suzhou, China. Her research interests include intelligent fault diagnosis of rotating machinery, domain adaptation, domain generalization, and transfer learning.

Liang Chen received the Ph.D. degree in control engineering from a joint Ph.D. program with Zhejiang University, Hangzhou, China, and TU Berlin, Berlin, Germany, in 2009. He is currently a professor with the Department of Automation Engineering, the School of Mechanical and Electric Engineering, Soochow University, Suzhou, China. His research interests include intelligent control and deep learning-based intelligent sensoring and fault diagnosis.

Hong Zhuang received the B.Eng. degree in Electrical Engineering and Automation from Soochow University Wen Zheng College, Suzhou, China. She is currently pursuing further studies in Engineering Management at Nanjing University, Nanjing, China. Presently, she serves as a research assistant at the School of Mechanical and Electric Engineering, Soochow University. Her research interests include transfer learning and domain generalization.

Qitong Chen received the B.Eng. degree in electrical engineering from Heilongjiang Bayi Agricultural University, Daqing, China, in 2021. He is currently pursuing the Ph.D. degree in Computer Science and Technology at Soochow University, Suzhou, China. His research interests include fault diagnosis of industrial robots, fault diagnosis of rotating machinery, domain generalization, adversarial learning, few-shot learning, and transfer learning.