Energy Proceedings
Vol 60, 2025

ISSN 2004-2965

Dual-Weight Adversarial Learning for Partial Domain Adaptation in Cross-Domain HVAC Fault Inference#

Yutian Lei[3], Cheng Fan[1,2,3*], Yonghang Xie[3], Jiena Cai[3]

1 State Key Laboratory of Subtropical Building and Urban Science (Shenzhen University), Shenzhen, China

2 Key Laboratory for Resilient Infrastructures of Coastal Cities, Ministry of Education, Shenzhen University, Shenzhen, China

3 Sino-Australia Joint Research Center in BIM and Smart Construction, Shenzhen University, Shenzhen, China

(Corresponding Author: fancheng@szu.edu.cn)

ABSTRACT

Data-driven fault inference in building HVAC systems faces a critical challenge due to the absence of labeled operational data in the target working condition. To address this, transfer learning-based methods are proposed to leverage inference knowledge from labeled datasets acquired in distinct operational contexts, defined as the source domain. However, these methods typically assume that fault categories between source and target domains are the same. This assumption is invalid in real applications where the target domain often contains fewer fault types than the source. Such scenarios may introduce negative transfer caused by irrelevant source domain categories. To solve this problem, a novel dual-weight partial domain adaptation (DW-PDA) method is proposed, which implements selective knowledge transfer by the class-wise adversarial learning for category-aware feature alignment, and a dual-weight mechanism for dynamic sample selection. Experimental results on both HVAC air-side and water-side datasets demonstrate the effectiveness of the DW-PDA in partial domain adaptation scenarios. Compared to the global domain adaptation methods, the average accuracy improvement of the DW-PDA is 30.87% and 25.26% in AHU and chiller tasks, respectively.

Keywords: partial domain adaptation, transfer learning, fault inference, HVAC, building energy system

NONMENCLATURE

Abbreviations	
HVAC	Heating, ventilation and air-conditioning
AHU	Air handling units
DW-PDA	Dual-weighted partial domain adaptation
UDA	Unsupervised domain adaptation
PDA	Partial domain adaptation

1. INTRODUCTION

Effective fault inference is crucial for ensuring safety and energy efficiency in building energy systems[1]. Due to the scarcity of labeled building operational data, transfer learning has become a key solution[2]. These methods transfer inference knowledge from source domains (e.g., HVAC systems with abundant labeled operational data) to the target domains (e.g., HVAC systems without related labeled data). However, differences in data distributions between source and target domains—caused by variations in building equipment, seasons, or operation patterns—hinder reliable knowledge transfer[3]. To address this, methods based on unsupervised domain adaptation (UDA) align the data distributions using techniques like adversarial training or statistical matching[4, 5]. For instance, Lei et al. improved cross-building fault inference by minimizing the Earth Mover's Distance (EMD) between air handling unit (AHU) datasets from different buildings [3].

However, existing UDA methods typically assume source and target domains share identical fault categories[6]. In reality, target domains often contain only a subset of source labels. For example, short-term monitoring data in online fault inference may exclude rare fault types documented in source domains. When source domains include irrelevant fault categories (e.g., faults absent in target buildings), negative effects may be generated to impact knowledge transfer tasks.

Thus, the primary objective of partial domain adaptation (PDA) is to enhance positive transfer for shared fault classes while mitigating negative transfer

This is a paper for the 11th Applied Energy Symposium: Low Carbon Cities & Urban Energy Systems (CUE2025), July 18-22, 2024, Kitakyushu, Japan.

from outlier classes. Weighting source domain samples during training has emerged as an effective PDA strategy. In image classification and mechanical bearing diagnostics, existing weighting mechanisms can be divided into three types, including class-level, sample-level, and hybrid approaches[7, 8]. However, such methods remain underexplored in building field. Liang et al. proposed a class-level PDA method that determines source class relevance based on target sample prediction probabilities[4]. By embedding these weights into the loss function, their approach achieved a 14.5% overall accuracy improvement in chiller fault inference under variable operating conditions compared to unweighted baselines.

However, single-strategy weighting exhibits limited robustness during domain adaptation, particularly under significant cross-domain distribution shifts [9]. To enhance the applicability of HVAC fault inference models across varying working conditions, a novel dual-weighted partial domain adaptation (DW-PDA) method is proposed. The main contributions of this study are as follows.

1) A novel partial transfer learning method is proposed for cross-domain building HVAC fault inference. It can selectively transfer task-relevant knowledge from labeled source domains to unlabeled target domains.

2) An innovative dual-level weighting mechanism is introduced, which dynamically adjusts the sample transferability through joint class-level and sample-level analysis. This dual-weight strategy can better identify the outlier class in the source domain.

3) Extensive experiments are conducted on two representative HVAC operation scenarios, including cross-load condition fault inference on chiller dataset and cross-seasonal condition fault inference on AHU dataset. The experiment results demonstrated the superiority of the proposed method and provided useful insights into practical application.

The remaining paper is organized as follows. Section 2 introduces research methodology. Section 3 details the data experiments setups. Results and conclusions are presented in Sections 4 and 5, respectively.

2. RESEARCH METHODOLOGY

2.1 Partial domain adaptation problem

Partial domain adaptation (PDA) is a challenging yet practical scenario in transfer learning, where the label space of the target domain $|C_T|$ is a subspace of the source domain $|C_S|$ (i.e., $C_T \in C_S$), and the detailed information about $|C_T|$ is unknown during adaptation[10]. To clearly describe the problem to be solved, some related notations are introduced as follows. Let $D_S = \{(X_i^S, Y_i^S)\}_{i=1}^{n_s}$ denote the labeled source domain with n_s samples, where X_i^S and Y_i^S represent the i_{th} sample and its label, respectively. The unlabeled target domain is defined as $D_T = \{(X_i^T)\}_{i=1}^{n_t}$, with n_t samples.

Unlike conventional UDA problem, which assumes that the distributions of the two domains are dissimilar (i.e., $P(X^S) \neq P(X^T)$), but their label spaces are the same. However, the PDA scenarios are more common in real-world applications. For PDA tasks, they must address two key challenges: 1) reducing the domain discrepancy between source and target domains ($P(X^S) \approx P(X^T)$), and 2) mitigating the negative effects caused by the outlier class samples of the source domain, which are not transferable.

2.2 The proposed DW-PDA method

To enhance positive transfer in PDA tasks for HVAC fault inference, a novel dual-weight partial domain adaptation (DW-PDA) method is proposed. Inspired by the sample selection mechanism of Selective Adversarial Networks [9], the DW-PDA implements a dual-weight strategy that simultaneously evaluates transferability weighting of each source sample at both class and sample levels. The overall framework of the proposed method is shown in Fig. 1. The feature extractor F learns domain-invariant representations through adversarial training with the discriminator$D$, effectively mitigating the cross-domain distribution discrepancy[11]. The classifier is designed to recognize the working states of the HVAC system. The hybrid weighting module mitigates negative transfer by assigning both class-level and sample-level weights to source samples, which subsequently reweight both the classification loss L_{cls} and adversarial loss L_{adv}. This

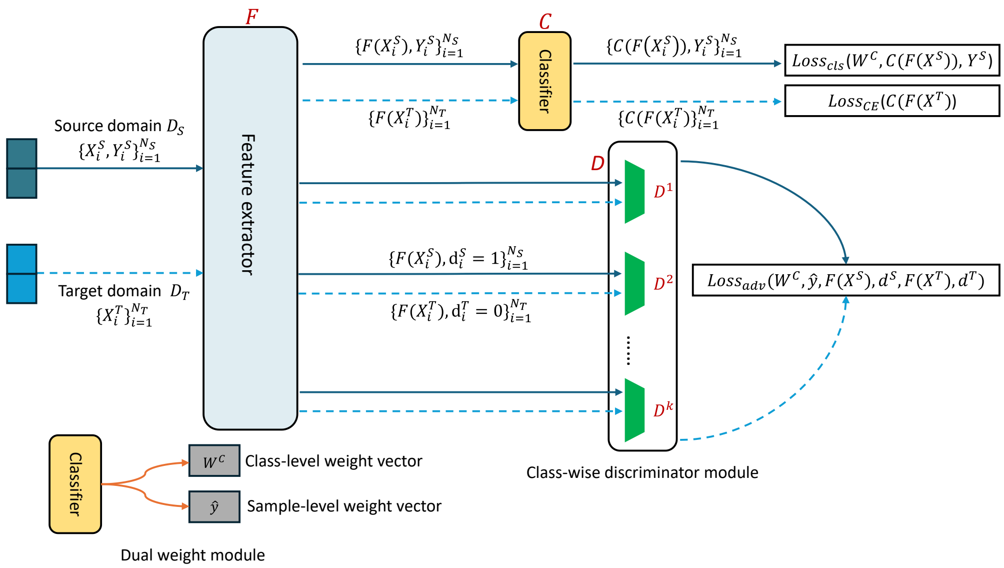

Fig. 1 The overall framework of the proposed DW-PDA

forces the model to prioritize learning from source samples exhibiting high transferability to the target task.

2.2.1 The dual-weight strategy

Since the target label space $|C_T|$ is unknown, the output probability distribution of the classifier C can serve as an indicator for identifying source-shared classes for source sample selection. Specifically, classes with consistently high prediction probabilities across target samples are more likely to belong to $|C_T|$. Thus, the sample-level weight vector of each sample across k health conditions is computed as Eq. (1).

$$\hat{y}_i = Softmax\left(C\left(F(X_i)\right)\right) \tag{1}$$

Moreover, the class-level weight vector is then obtained by averaging target sample weights per class, as shown in Eq. (2). To prevent numerical instability caused by vanishingly small weights, the class weights are normalized by their maximum element.

$$W_i^C = \frac{1}{n_t}\sum_{i=1}^{n_t} \hat{y}_i^T \tag{2}$$

2.2.2 The class-wise domain discriminator module

To achieve fine-grained UDA, K class-specific domain discriminators $\{D_k\}_{k=1}^{K}$ are employed, where each D_k is dedicated to aligning the feature distributions between source and target domain samples for class k. The sample-level weight represents the probability of assigning sample X_i to domain discriminator D_k.Thus, the adversarial loss for each D_k is computed as a weighted sum of sample losses. Furthermore, to mitigate negative transfer from outlier-class samples, class-level weights are applied to reweight the adversarial loss of D_k. The final objective, integrating both weighting strategies, is formulated in Eq. (3). Here, d_i denotes the domain label (1 for source, 0 for target), and L_{CE}^krepresents the cross-entropy loss for the k_{th} discriminator's domain classification task.

$$L_{adv} = \frac{1}{(n_s+n_t)} \times \sum_{k=1}^{|C_S|}\left[\left(W_i^{C_k}\right) \times \sum_{X_i\in(D_S\cup D_T)} \hat{y}_i^k L_{CE}^k\left(D_k\left(F(X_i)\right), d_i\right)\right] \tag{3}$$

2.2.3 The classifier module

Since outlier classes may cause negative transfer, their contribution to the classification loss is downweighted. The final class-weighted loss L_{cls} is formulated in Eq. (4).

$$L_{cls} = \frac{1}{(n_s)} \times \sum_{X_i\in(D_S)} W_i^{C_k} L_{CE}^k\left(C\left(F\left(X_i^S\right)\right), Y_i^S\right) \tag{4}$$

Furthermore, to enhance fault-type separability, an entropy minimization loss L_E, based on the entropy minimization principle [12], is introduced to sharpen target predictions. The loss L_E is shown in Eq. (5), where $H(\cdot)$ is the conditional-entropy loss function.

$$L_E = \frac{1}{(n_t)} \times \sum_{X_i\in(D_T)} H\left(C\left(F\left(X_i^T\right)\right)\right) \tag{5}$$

The proposed DW-PDA framework optimizes a composite objective function consisting of three key components, as formulated in Eq. (6). The trade-off parameters λ and β control the relative importance of each term in the overall optimization. Notably, β follows a progressive scheduling strategy, where its value gradually increases from 0 to 1 according to the training epoch (ep) as defined in Eq. (7).

$$Loss_{total} = L_{cls} + \lambda L_E + \beta L_{adv} \tag{6}$$

$$\beta = \frac{2}{\left(1 + \exp\left(-\frac{ep}{10}\right)\right)} - 1 \quad (7)$$

3. DATA EXPERIMENTS SETUPS

3.1 Description of experimental datasets and partial domain adaptation tasks

To ensure the generalizability of research results, two experimental datasets describing operations of HVAC air-side and water-side systems are used for analysis, respectively.

The first dataset comes from the ASHRAE RP-1312 (denoted as the RP-1312), which is experimental data collected from AHU systems[13]. To simulate cross-seasonal fault inference under varying weather conditions, the summer operational data is selected as the source domain for training, while the winter data serves as the target domain for testing. In source domain, the dataset includes one normal class (Nor) and 7 major faults (i.e., exhaust air damper stuck at fully open (EA Damper100%), exhaust air damper stuck at fully close (EA Damper0%), outdoor air damper stuck at fully close (OA Damper0%), cooling coil valve stuck at fully open (CC Valve100%), outdoor air damper stuck at 45% open(OA Damper45%), return fan complete failure(RF-Fail), and heating coil valve leaking at stage 3 (HC Valve-leak). In this study, 4 PDA fault inference tasks are investigated on AHU dataset, detailed information is shown in Table 1. For example, in task A1, the target domain includes only normal and one fault type (i.e., EA Damper100%) samples, while the source domain contains normal and all 7 fault types, forming a partial domain adaptation scenario.

The second dataset is obtained from ASHRAE Research Project RP-1043(denoted as the RP-1043), featuring comprehensive experimental data that was collected from a 90-ton (316 kW) centrifugal chiller[14]. The chiller was evaluated under both healthy conditions and 7 representative fault scenarios, with each condition assessed at 4 distinct severity levels. The fault scenarios include excess oil (EO), non-condensable gas (NC), refrigerant overcharge (RO), refrigerant leakage (RL), reduced evaporator water flow (FWE), reduced condenser water flow (FWC) and condenser fouling (CF). To simulate cross-load conditions, the dataset is divided based on cooling load magnitude. Samples are sorted by descending cooling load, with the top 30% classified as the target domain representing high-load operation, and the remaining 70% designated as the source domain for normal-load conditions. In this study, 7 PDA fault inference tasks are investigated on this chiller dataset, detailed information is shown in Table 1.

Table. 1 The settings of partial domain adaptation task for HVAC fault inference

Dataset	Task Name	Domain information	Source classes	Target classes
RP-1312 (AHU)	A1	Summer operational data (Source) \ Winter operational data (Target)	Nor, 7 fault classes	Nor, EA Damper100%
	A2			Nor, EA Damper0%
	A3			Nor, OA Damper0%
	A4			Nor, CC Valve100%
RP-1043 (Chiller)	C1	High-cooling load operational data (Source) \ Normal-cooling load operational data (Target)	Nor, 7 fault classes	Nor, EO
	C2			Nor, NC
	C3			Nor, RO
	C4			Nor, RL
	C5			Nor, FWE
	C6			Nor, FWC
	C7			Nor, CF

3.2 Experimental details

In each partial domain adaptation task, 225 labeled samples per class were randomly selected from the source dataset. For the target domain, 150 unlabeled samples per class were extracted using the same sampling protocol. The label information of target samples is intentionally masked during training to simulate unsupervised scenarios. To ensure statistical reliability, this study conducted five independent random sampling trials and reported average classification accuracy.

To systematically validate the effectiveness of the proposed DW-PDA method, four baseline models were implemented for comparative experiments, comprising two supervised learning models and two transfer learning approaches. All neural network models used the same architecture for fair comparison. The feature extractor has four fully-connected layers. Each layer contains 128 neurons. Batch normalization and dropout (rate=0.3) follow every layer. The discriminator and classifier use similar fully-connected designs. Model optimization was uniformly performed using stochastic gradient descent (SGD) with a momentum coefficient of 0.9, coupled with L2 regularization (weight decay =1e-4).

Four baseline models were implemented for comprehensive comparison, with detailed configurations as follows:

1) FCNN (Fully Connected Neural Network). As a conventional supervised learning model, FCNN is trained exclusively on labeled source domain data. The learning rates were differentially set at 0.001 for AHU tasks and 0.005 for Chiller tasks.

2) Lightgbm (Light Gradient Boosting Machine). This non-deep learning method is implemented based on the gradient boosting decision tree framework[15]. Default hyperparameters were used throughout the experiments.

3) DANN (Domain-Adversarial Neural Network). This global UDA method utilizes adversarial training with gradient reversal layers. All source samples were uniformly weighted, and a learning rate of 0.001 was applied for both AHU and chiller tasks.

4) PADA (Partial Adversarial Domain Adaptation). As a classical PDA method, PADA incorporates class-level weighting mechanisms to adjust source domain sample contributions[4]. The learning rate was fixed at 0.005 for both task types. The class-wise sample weights in PADA were updated every 5 epochs, whereas DW-PDA implemented more frequent updates every 2 epochs.

4. DISCUSSION

4.1 The performance of the DW-PDA in partial domain adaptation

As shown in Tables 2 and 3, the fault inference performance of DW-PDA is evaluated across four PDA tasks for the AHU system and seven tasks for the chiller system. Figures 2 and 3 demonstrate the accuracy improvements achieved by DW-PDA over the four baseline models. The experimental results confirm the superior performance of DW-PDA in partial domain adaptation scenarios.

Specifically, DW-PDA achieves average accuracy improvements of 10.49% over LightGBM and 25.68% over FCNN in AHU tasks, with similar gains (20.71% and 7.83%) in chiller tasks. This substantial performance gap underscores the limitations of conventional supervised approaches when facing significant domain shifts, while highlighting DW-PDA's effectiveness in cross-domain feature distribution alignment through adversarial learning.

When compared to the global adaptation method (i.e., DANN), DW-PDA outperforms by an average margin of 30.88% and 25.26% in AHU and chiller tasks, respectively. This improvement demonstrates the capability of the DW-PDA to mitigate negative transfer effects from irrelevant source domain categories through its selective sample transfer mechanism.

Compared to the baseline method PADA (a classical partial domain adaptation approach), DW-PDA maintains average improvements of 13.44% and 0.71% in AHU and chiller tasks, respectively. The performance gain stems from two key innovations: 1) a dual-weight mechanism integrating class-level coarse filtering (global suppression of irrelevant categories) and instance-level fine calibration (adaptive sample weighting within shared classes), and 2) a category-aware adversarial module enabling fine-grained, class-conditional domain alignment. Furthermore, the advantage is more significant in challenging cross-seasonal conditions. For example, in task A3, DW-PDA achieves 91.80% accuracy. In contrast, PADA and ANN reach only 72.80% and 74.13% respectively. However, PADA exhibits marginally better performance than DW-PDA in certain chiller system scenarios. For example, results for task C6 reveal a 3.60% accuracy advantage for PADA (98.47%) over DW-PDA (94.87%). Nevertheless, this narrow performance gap primarily occurs in near-identical domain scenarios, whereas DW-PDA demonstrates superior robustness in practical applications.

Table. 2 Fault inference accuracy (%) of domain adaptation tasks on the AHU system

Task	Lightgbm	FCNN	DANN	PADA	DW-PDA
A1	**60.47**	28.53	43.20	38.73	58.60
A2	60.20	54.33	34.47	83.80	**91.80**
A3	90.13	74.13	76.80	72.80	**91.80**
A4	67.13	60.13	41.93	70.80	**77.67**

Table. 3 Fault inference accuracy (%) of domain adaptation tasks on the chiller system

Task	Lightgbm	FCNN	DANN	PADA	DW-PDA
C1	57.20	54.33	60.00	**88.80**	85.40
C2	71.80	86.67	69.33	86.27	**92.73**
C3	67.67	80.60	61.73	86.60	**87.67**
C4	44.53	60.40	41.93	**60.80**	54.53
C5	72.53	85.87	62.00	89.13	**94.07**
C6	71.40	87.13	62.93	**98.47**	94.87
C7	66.67	86.80	62.07	81.73	**87.53**

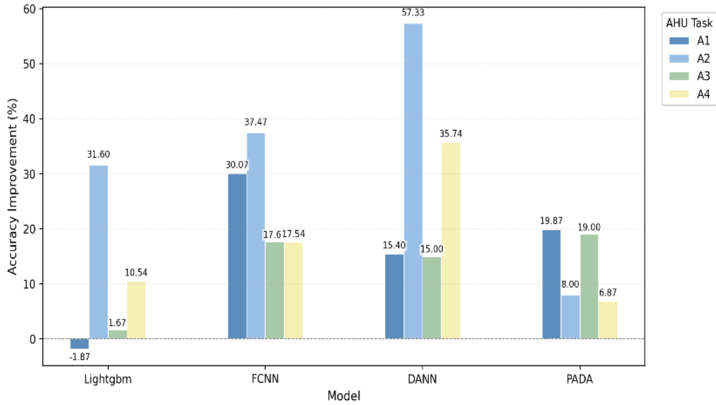

Fig.2 Performance improvement of DW-PDA over four baseline models on the AHU system

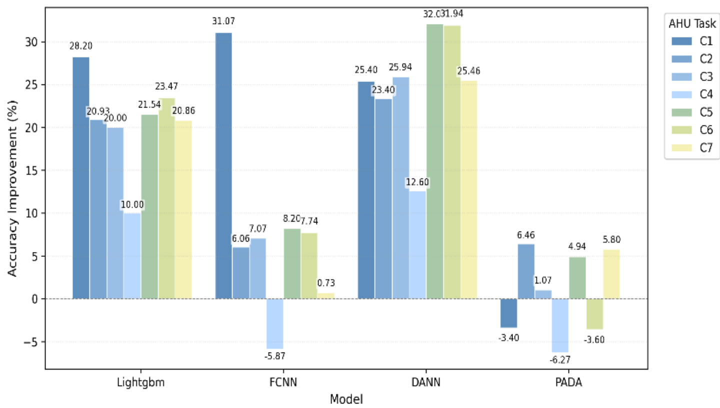

Fig.3 Performance improvement of DW-PDA over four baseline models on the chiller system

4.2 The importance of sample selection in transfer learning

Notably, conventional UDA methods (i.e., DANN) consistently underperform supervised models in PDA scenarios due to negative transfer effects. As evidenced in Task C5, DANN achieves only 62.0% fault inference accuracy compared to that of FCNN (85.87%). This performance degradation occurs because standard domain alignment may incorrectly match outlier source distributions with target features, biasing the classifier toward irrelevant source classes.

In contrast, the proposed DW-PDA demonstrates consistent performance improvements through its selective sample transfer mechanism. For clearer comparison, Fig. 4 shows the confusion matrices for Task C5. In this scenario, the source domain contains 8 health conditions, while the target domain comprises only 2 conditions (i.e., Nor and FWC). The visualization reveals that DW-PDA maintains superior target classification accuracy with minimal misclassifications, while DANN frequently associates target samples with source-specific classes. These results confirm the critical role of selective transfer in domain adaptation.

5. CONCLUSIONS

In this study, a novel dual-weight partial domain adaptation method, i.e., DW-PDA, is proposed for building system fault inference under partial domain adaptation scenarios. In PDA settings, the unlabeled target domain contains fewer fault categories than the source domain. To address this challenge, DW-PDA integrates sample-level and class-level weighting mechanisms to estimate source sample transferability, thereby mitigating negative transfer from irrelevant samples. Additionally, the method employs a class-wise discriminator module to achieve fine-grained cross-domain alignment of shared-class distributions. Extensive experiments on AHU and chiller systems validate DW-PDA's effectiveness. The main findings are as below.

1) DW-PDA demonstrates consistent performance advantages over both conventional supervised methods (i.e., FCNN and Lightgbm) and state-of-the-art UDA methods (i.e., DANN and PADA) across most scenarios.

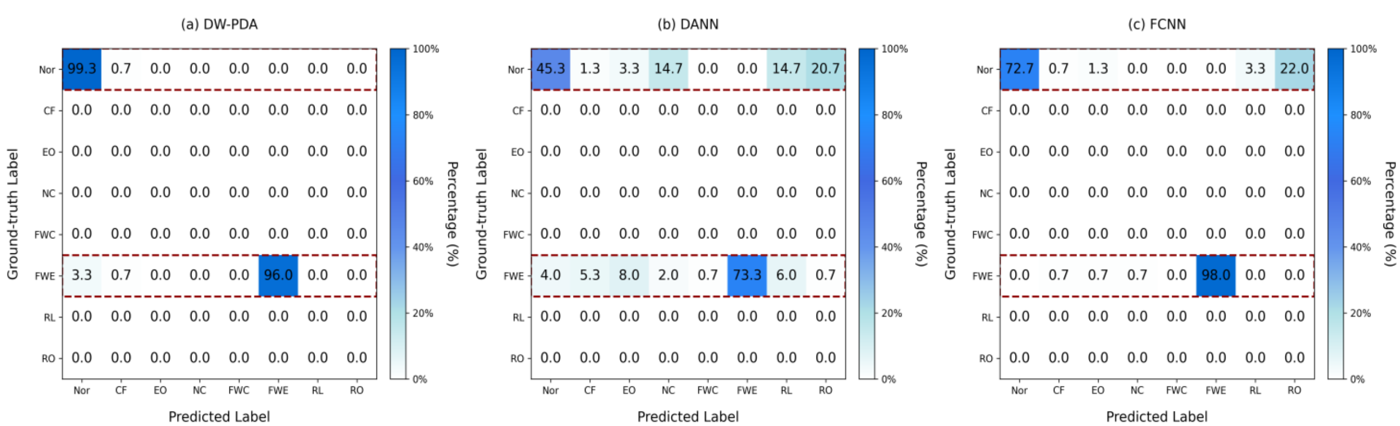

Fig.4 Performance comparison of DW-PDA, DANN and FCNN on Task C5 via confusion matrices

Specifically, it achieves 25.68% higher accuracy than FCNN in AHU systems while maintaining performance leads of 13.44% and 0.71% over PADA in AHU and chiller systems respectively. These results confirm the superior capability of the proposed method in handling complex operational variations across different HVAC system configurations.

2) The experimental results underscore the critical importance of selective sample transfer in partial domain adaptation. For example, the severe performance degradation of DANN, which achieved only 62.0% accuracy compared to FCNN's 85.87% in Task C5, clearly illustrates the consequences of negative transfer. In contrast, the robust performance of DW-PDA and PADA validates their effectiveness of its sample selection strategy for successful domain adaptation and knowledge transfer.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of this research by the National Natural Science Foundation of China (No. 52278117), Guangdong Basic and Applied Basic Research Foundation (No. 2024A1515011549), Shenzhen Science and Technology Program (No. 20220531101800001, 20240813143330039 and 20220810160221001), and the Philosophical and Social Science Program of Shenzhen (SZ2024C006).

REFERENCE

[1] G. Li, J. Xiong, S. Sun, J. Chen, Validation of virtual sensor-assisted Bayesian inference-based in-situ sensor calibration strategy for building HVAC systems, Building Simulation 16(2) (2023) 185-203.
[2] C. Fan, Q. Wu, Y. Zhao, L. Mo, Integrating active learning and semi-supervised learning for improved data-driven HVAC fault diagnosis performance, Applied Energy 356 (2024) 122356.
[3] Y. Lei, C. Fan, H. He, Y. Xie, Enabling efficient cross-building HVAC fault inferences through novel unsupervised domain adaptation methods, Building and Environment (2025) 112678.
[4] X. Liang, P. Li, S. Chen, X. Jin, Z. Du, Partial domain adaption based prediction calibration methodology for fault detection and diagnosis of chillers under variable operational condition scenarios, Building and Environment 217 (2022) 109099.
[5] B. Feng, Q. Zhou, J. Xing, Q. Yang, Y. Chen, Z. Deng, Attention-empowered transfer learning method for HVAC sensor fault diagnosis in dynamic building environments, Building and Environment 250 (2024) 111148.
[6] X. Zhu, K. Chen, B. Anduv, X. Jin, Z. Du, Transfer learning based methodology for migration and application of fault detection and diagnosis between building chillers for improving energy efficiency, Building and Environment 200 (2021) 107957.
[7] L. Guo, Y. Yu, Y. Liu, H. Gao, T. Chen, Measurement, Reconstruction domain adaptation transfer network for partial transfer learning of machinery fault diagnostics, IEEE Transactions on Instrumentation 71 (2021) 1-10.
[8] Y. Zhang, J. Ji, Z. Ren, Q. Ni, F. Gu, K. Feng, K. Yu, J. Ge, Z. Lei, Z. Liu, Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing, Reliability Engineering System Safety 234 (2023) 109186.
[9] Z. Cao, M. Long, J. Wang, M.I. Jordan, Partial transfer learning with selective adversarial networks, Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 2724-2732.
[10] Z. Cao, L. Ma, M. Long, J. Wang, Partial adversarial domain adaptation, Proceedings of the European conference on computer vision (ECCV), 2018, pp. 135-150.
[11] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. March, V. Lempitsky, Domain-adversarial training of neural networks, Journal of machine learning research 17(59) (2016) 1-35.
[12] Z. Cao, K. You, M. Long, J. Wang, Q. Yang, Learning to transfer examples for partial domain adaptation, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 2985-2994.
[13] S. Li, Development and validation of a dynamic air handling unit model, Part I, ASHRAE Transactions 116 (2010) 45.
[14] M. Comstock, J. Braun, E. Groll, The sensitivity of chiller performance to common faults, ASHRAE Transactions 108(1) (2002) 467.
[15] M. Li, H. Tao, M. Liu, T. He, Study on enhanced fault diagnosis of chiller units in HVAC systems under the imbalanced data environment using GA-Optimized LightGBM, Energy and Buildings (2025) 115360.