Supplementary Document

This document provides information on preparing a supplemental information for inclusion with primary article submission to IEEE Transactions on Industrial Cyber-Physical Systems journal. This document, which include supplementary information such as the confusion matrix of testing accuracy for the partial transfer task 7 cross different approaches and ablation study for the proposed approach, will be published as a PDF linked to the primary article.

1. CONFUSION MATRIX

The diagnostic results considering the DT update strategy for all tasks are summarized in Table VIII of the primary article, using precision as the evaluation metric. The general comparison in Table VIII demonstrates the clear superiority of the framework developed in addressing partial transfer learning challenges, especially in cases with significant label space discrepancies between the source and target domains. To further assess the negative impact of outlier source data across different approaches, Fig. S1 depicts the confusion matrix to test accuracy for the partial transfer task 7. Here, it becomes evident that a significant portion of the inner fault and outer fault data are inaccurately diagnosed by feature-based transfer models (MK-MMD and WDANN), resulting in unexpected negative transfer. In contrast, the SAN and IWAN approaches significantly improve positive transfer for the partial task 7. However, both adversarial-based and feature-based transfer strategies tend to misclassify most target data originally labeled as IF and OF into the H condition, significantly impacting transferability and potentially inducing negative transfer.

This negative transfer phenomenon may stem from several factors. Firstly, there could be a bias towards classifying instances as healthy due to imbalanced data in the target domain (as indicated in Table. III of primary article, the ratio of health condition data to each faulty condition is 3 to 1). Secondly, the models may tend to learn common feature modes from labels IF, OF, and H to adapt to both source and target domains while failing to capture discriminative features between them. The quality of generated or real data, or the fault size, further affects the domain-shared feature learning process, leading to unexpected negative transfer. In contrast, the proposed model effectively suppresses this negative transfer effect by employing a dual-attention mechanism, thereby significantly enhancing diagnostic performance.

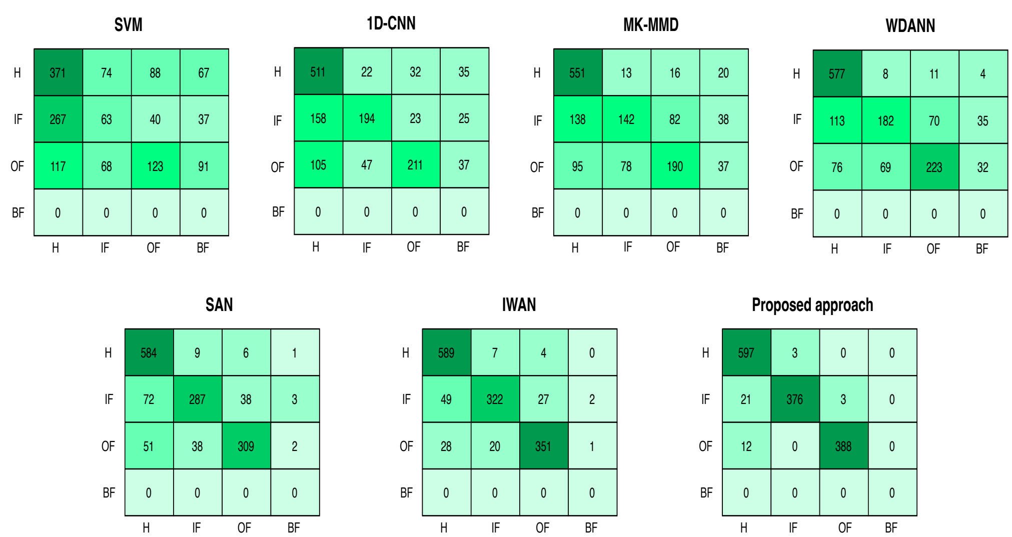

Fig. S1. Comparison analysis. Confusion matrices for task 7 with the DT updating strategy.

The diagnostic results for partial TL tasks without considering the DT update strategy are presented in Table IX of primary article . These results demonstrate that the proposed model consistently outperforms other approaches across all tasks, showcasing its transfer efficiency. Again the confusion matrix is exploited to explore the efficiency of the developed approach also in this case. Illustrative results for the task 7 are depicted in Fig .S2. It can be noticed that all transfer learning approaches reduce their transferability on all labels. However, our proposed approach still achieves better performance compared to the others.

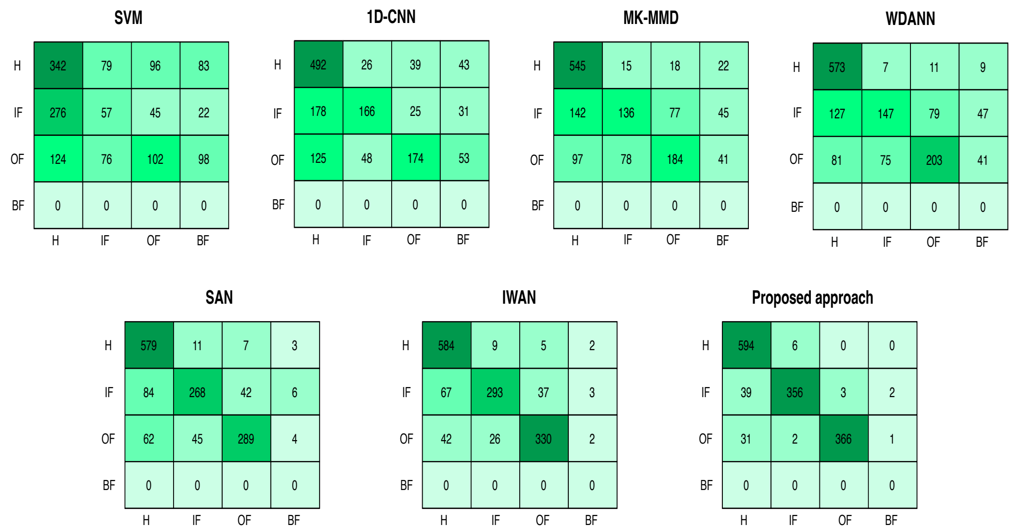

Fig. S2. Comparison analysis. Confusion matrices for task 7 without the DT updating strategy.

2. ABLATION EXPERIMENT

To verify the effectiveness of the dual-attention mechanism, ablation experiments were conducted, comparing the performance of the proposed approach both with and without this mechanism across various domain adaptation tasks. These experiments aimed to isolate the impact of the dual-attention mechanism, providing a clear understanding of its contribution to improving the overall model performance.

Table S1. Diagnostic accuracy (percentage) comparison of the proposed approach with and without the dual-attention mechanism

	Proposed PTL	
Task	with dual-attention mechanism	without dual-attention mechanism
Task 1	95.4	93.6
Task 2	97.6	85.8
Task 3	94.8	80.5
Task 4	92.5	74.4
Task 5	94.5	78.2
Task 6	94.1	91.8
Task 7	97.2	82.7
Task 8	91.6	70.9
Task 9	89.6	68.1
Task 10	93.9	72.3
Task 11	95.2	83.4
Task 12	94.4	77.6
Task 13	96.6	74.1
Average	94.26	79.49

The experimental results demonstrate that introducing the dual-attention mechanism significantly enhances the performance of the proposed approach, especially in partial transfer learning tasks. As shown in Table S1, the proposed approach with the dual-attention mechanism consistently achieves superior domain adaptation performance across all transfer tasks. On average, the accuracy of the proposed approach with the dual-attention mechanism is approximately 94.2%, while the version without this mechanism shows a marked decrease to 79.49%.

A closer examination of the results reveals interesting patterns across different types of transfer tasks. In closed-set transfer tasks, where both the source and target domains share the same label space (e.g., Task 1 and Task 6), both the proposed approach with and without the dual-attention mechanism exhibit comparable performance, indicating that the mechanism may not be as critical in scenarios where the label distribution is fully aligned. However, in partial transfer tasks, where the target domain contains fewer labels than the source domain, the performance gap becomes more pronounced. For instance, in Task 7, which represents a partial transfer task, the proposed approach without the dual-attention mechanism reaches an accuracy of 82.7%, which is 14.5% lower than the version equipped with the dual-attention mechanism. This suggests that the dual-attention mechanism plays a crucial role in handling the domain misalignment and reducing negative transfer in more complex transfer settings.

Overall, these results confirm that incorporating the dual-attention mechanism into the discriminator significantly enhances the proposed approach's domain adaptation capability, particularly in partial transfer tasks where domain shift is more challenging to overcome.