applied sciences

MDPI

Article

Low-Resource Spectral Feature Alignment for Cross-Condition Rolling Bearing Fault Diagnosis

Xiong Xiong , Linlin Liu * and Yu Cao

School of Information and Control Engineering, Liaoning Petrochemical University, Fushun 113001, China
* Correspondence: dl198182@163.com

Featured Application

A reproducible, label-blind-during-training target-alignment protocol for evaluating low-resource cross-condition bearing diagnosis.

Abstract

Cross-condition rolling-bearing diagnosis is difficult because operating-condition shifts alter both the amplitude distribution and spectral composition of vibration signals. This study evaluates whether a fixed spectral representation changes low-resource, label-blind unsupervised domain-adaptation performance when the learning backbone and optimization budget are controlled. Each 1024-point vibration window was standardized, transformed through $\log(1 + |\mathrm{rFFT}|)$, L2-normalized, and processed by a one-channel convolutional neural network. Source-only, Deep CORAL, maximum mean discrepancy (MMD), and domain-adversarial neural-network objectives were compared under a manifest-defined protocol that permits only 10% of a target-domain alignment pool to be used without labels. Target-query labels were unavailable to training and checkpoint selection; checkpoints were selected exclusively using source-domain validation Macro-F1. Five fixed training seeds were used. On CWRU hp0-to-hp3, FFT-MMD achieved 98.92 ± 0.34% Macro-F1, compared with 77.21 ± 5.16% for time-domain MMD; DANN achieved 99.23 ± 0.10%. On the reverse transfer, FFT-MMD achieved 96.84 ± 0.68% and DANN 98.35 ± 1.00%. On JNU-S600, FFT-MMD reached 90.48 ± 1.58%, whereas time-domain MMD reached 67.05 ± 3.75%. Time-domain MMD, however, outperformed FFT-MMD on the JNU-S800 and JNU-S1000 boundary tasks. The results therefore indicate a conditional representation effect, not a universal advantage of frequency-domain input.

Keywords: rolling bearing fault diagnosis; unsupervised domain adaptation; spectral representation; maximum mean discrepancy; domain-adversarial neural network; low-resource learning

Check for updates

Academic Editor: Emiliano Mucchi

Received: 30 August 2026
Revised: 31 October 2026
Accepted: 5 November 2026
Published: 8 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Rolling-element bearings are critical components of rotating machinery, and vibration-based diagnosis is widely used to support condition monitoring and maintenance decisions [1–3]. Deep neural networks can learn diagnostic features directly from vibration segments [4]; nevertheless, a model trained under one load or speed may degrade when evaluated under another operating condition. The change is not merely a classification problem: load and speed affect signal energy, characteristic-frequency placement, transmission paths, and the relative prominence of impulsive components. A high within-condition score is therefore not sufficient evidence that a method will transfer across operating conditions.

Domain adaptation offers one response to this problem. In the usual unsupervised formulation, source samples carry labels and target samples are available without labels during training, while target labels remain unavailable to the adaptation procedure [5]. Distributional objectives such as maximum mean discrepancy (MMD) encourage source and target embeddings to have similar mean embeddings in a reproducing-kernel Hilbert space [6,7]. Correlation alignment (CORAL) instead aligns second-order feature statistics [8]. Domain-adversarial neural networks (DANNs) learn an embedding that is useful for source classification while making the domain difficult to discriminate through gradient reversal [9]. These methods have been applied to cross-condition bearing diagnosis [10–14], but their conclusions depend strongly on signal representation, target-data access, data splitting, and the treatment of repeated runs.

Within machinery diagnosis, related work spans neural-network transfer under changing conditions, CNN-based load-robust learning, deep domain adaptation, and few-shot transfer [15–18]. Reviews emphasize that transferability, negative transfer, target-label scarcity, and the realism of the evaluation protocol remain central concerns [19–23]. More elaborate UDA objectives, including adversarial discriminative adaptation, classifier discrepancy, and joint-distribution alignment, are also available [24–26]. They are not added here as further competing methods because the purpose of this study is a controlled representation comparison using a small, fixed baseline set rather than a claim of state-of-the-art superiority.

Frequency-domain features are frequently used in vibration analysis because rotating-machine faults may manifest as structured spectral energy or modulation patterns [27]. Yet a frequency transform alone does not guarantee improved cross-condition transfer. In particular, a representation can remove nuisance variation in one task while discarding time-local information needed in another. This paper therefore does not propose a new Fourier transform or a new neural architecture. It studies a narrower and testable question: under a fixed one-channel CNN and a small, strictly label-blind target-alignment budget, does a normalized log-amplitude real FFT representation change the effectiveness of standard alignment objectives relative to time-domain input?

The comparison isolates input representation under a matched architecture rather than comparing alternative architectures. The input transform is fixed, the CNN is shared, the target-access budget is fixed, and all adaptation alternatives use the same checkpoint selection rule. This makes a negative result informative: if a time-domain condition is stronger on a particular transfer task, the result cannot be dismissed as an effect of a different backbone or a different amount of target data. Conversely, a positive spectral result is interpreted as evidence under this protocol only, rather than as a universal property of Fourier analysis. This distinction is particularly important for public bearing benchmarks, where high scores can otherwise hide decisions about window construction, data-role access, and model selection.

The study is designed to avoid several recurring threats to validity in fault-diagnosis experiments. Source training, source validation, target alignment, and target query roles are materialized in manifests before model fitting. Alignment files expose only signal paths to the training code; their label columns are not read. Target-query labels are opened only after source-validation checkpoint selection. Role separation is audited for duplicate output paths and raw-sample interval overlap. Five fixed seeds, rather than target windows, are the statistical replications. Results are reported as mean ± standard deviation, and no window-level p-values are used.

The contributions are threefold:

1. A manifest-defined, role-separated low-resource UDA protocol for bearing diagnosis that separates source validation, a label-free target-alignment pool, and a target query set.
2. A controlled comparison of time-domain and fixed log-rFFT representations using the same CNN backbone, optimization budget, and model-selection rule. This is a representation-and-protocol contribution, not a new adaptation algorithm or a claim of state-of-the-art performance.
3. A bidirectional CWRU analysis and multi-speed JNU analysis showing that spectral alignment can be beneficial on some condition shifts but is not uniformly superior to time-domain learning or domain-adversarial alignment. The negative boundary-task results are retained to delimit the scope of that conclusion.

2. Materials and Methods

2.1. Study Design and Data-Role Isolation

For each task, source-domain windows were divided into source training and source validation sets. A fixed 10% subset of the target-alignment pool was available without labels to CORAL, MMD, and DANN; source-only training did not consume target inputs. The target query set was not used for hyperparameter selection, early stopping, checkpoint selection, or training. A selected checkpoint was the epoch with the largest source-validation Macro-F1. Only then were target-query labels loaded for a single final evaluation.

The manifests retain source file, class, sample coordinate, role, and derived-window information. Before training, an audit found no missing saved windows, cross-role duplicate paths, or raw-sample interval overlaps in the five tasks. This establishes window-level separation only: roles can still use non-overlapping regions of the same recording. The waveform and spectrum symbols in Figure 1 identify the two fixed representations; they are not measured example signals.

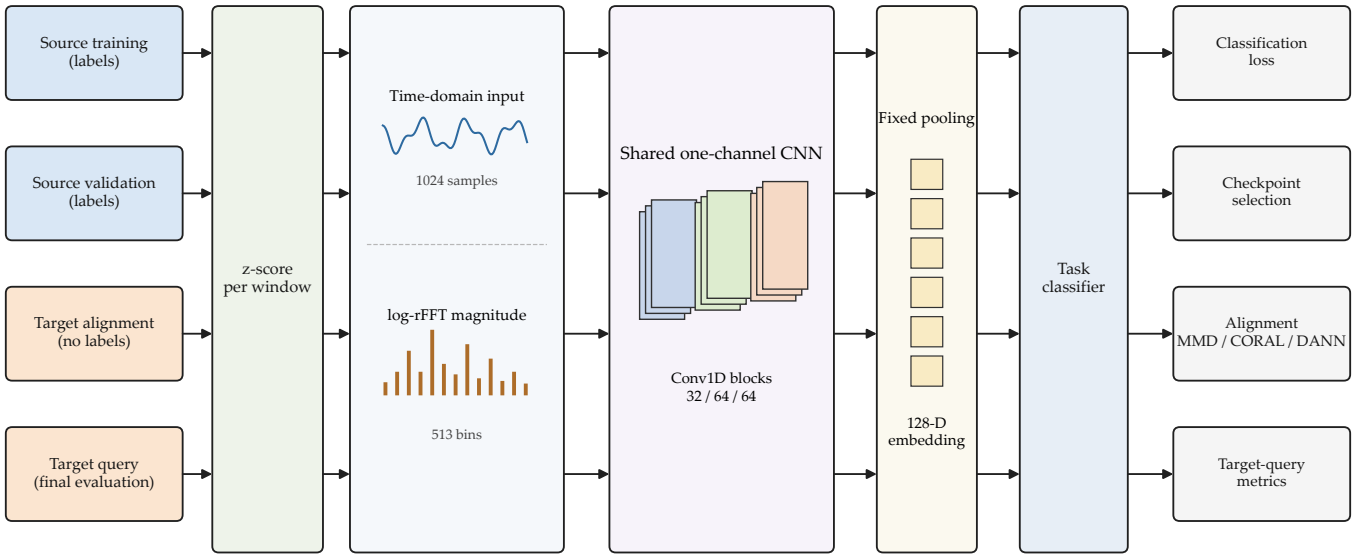


Figure 1. Manifest-defined data roles and permitted information flow for each transfer task.

A non-overlapping 1024-point window, rather than an independently acquired recording, is the split unit. A recording-level split and validation on independently acquired recordings were not performed. The reported scores therefore quantify transfer under this window-level protocol and cannot be interpreted as evidence of generalization to an independent recording, machine, or acquisition campaign.

The target-alignment loader receives inputs but does not read target labels, and the target-query split is excluded from training and checkpoint selection. The five seeds are stochastic training replications of this fixed protocol, not alternative test partitions; target-query windows are not treated as independent repetitions. Mean and standard deviation across seeds are thus descriptive summaries, and no threshold or window-level significance test is selected from target-query results.

2.2. Datasets and Transfer Tasks

The Case Western Reserve University (CWRU) Bearing Data Center provides seeded normal, inner-race, outer-race, and rolling-element fault recordings under multiple motor loads [28,29]. CWRU drive-end signals were used for the hp0-to-hp3 and hp3-to-hp0 load-transfer tasks. Ten classes were retained: normal and three defect sizes for each of the ball, inner-race, and outer-race locations. For each direction, the source train/validation sets contained 200/50 windows per class, the target-alignment pool contained 50 windows per class, and the target query set contained 200 windows per class. The CWRU apparatus is shown in Figure 2.

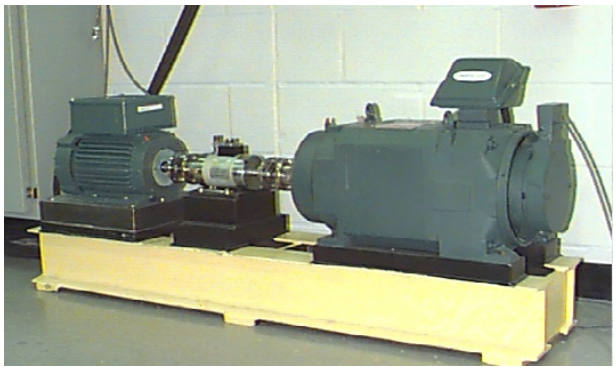


Figure 2. CWRU bearing-data test rig, reproduced from the CWRU Bearing Data Center [29].

The CWRU collection is widely used as a diagnostic benchmark, but its convenience does not remove the need to specify the splitting unit and data roles. Smith and Randall highlighted that evaluations based on the CWRU data require careful interpretation because task construction can materially influence conclusions [30]. In this work, a window is a derived analysis unit rather than an independent physical experiment; accordingly, seed-to-seed training replications, not individual windows, are the reporting units.

The Jiangnan University (JNU) benchmark contains normal, inner-race, outer-race, and rolling-element conditions collected at 600, 800, and 1000 rpm at a 50 kHz sampling frequency [10]. Signals were resampled from 50 kHz to 12 kHz by a rational 6/25 resampling ratio before windowing. For JNU-S600, the 800 and 1000 rpm recordings formed the source domain and the 600 rpm recordings formed the target domain. Analogously, JNU-S800 used 600 and 1000 rpm as source conditions, and JNU-S1000 used 600 and 800 rpm as source conditions. Each task used 160 source-training and 60 source-validation windows per class. The target alignment pool and target query set contained 30 and 60 windows per class, respectively, with a 10-window guard interval between time-ordered target regions. The corresponding JNU apparatus is shown in Figure 3.

All input windows had 1024 points. The CWRU and JNU tasks are reported separately; no cross-dataset transfer is claimed. Dataset properties and role counts are shown in Table 1.

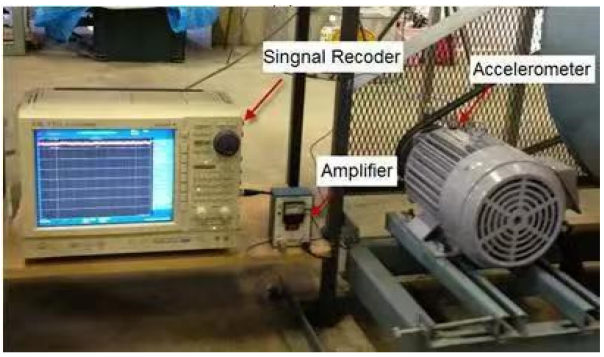


Figure 3. JNU rolling-bearing test rig, reproduced from Wang et al. [31] under CC BY 4.0.

Table 1. Dataset tasks and classwise window counts.

Task	Source	Target	Classes	Source: Train/Val	Target: Align/Query
CWRU hp0-to-hp3	0 hp	3 hp	10	200/50	50/200
CWRU hp3-to-hp0	3 hp	0 hp	10	200/50	50/200
JNU-S600	800, 1000 rpm	600 rpm	4	160/60	30/60
JNU-S800	600, 1000 rpm	800 rpm	4	160/60	30/60
JNU-S1000	600, 800 rpm	1000 rpm	4	160/60	30/60

In Table 2, "shared" denotes identical file identities assigned to separate roles with non-overlapping raw intervals.

Table 2. Record-level provenance and the remaining dependence boundary.

Task Family	Record Provenance, Role Allocation, and Remaining Dependence
CWRU hp0-to-hp3 and hp3-to-hp0	10/10 source train/validation and 10/10 target-alignment/query raw files have shared identities. One drive-end record represents each class–condition; alignment and query use distinct partitions separated by a 10% unused time gap. Thus, raw intervals are separated but the original records are shared.
JNU-S600, JNU-S800, and JNU-S1000	8/8 source train/validation and 4/4 target-alignment/query raw files have shared identities. For each target record, the first 30 windows form alignment, the next 10 a guard interval, and the final 60 query. Thus, raw intervals are separated but the original records are shared.

2.3. Class-Balanced Manifest Construction and Label Use

The equal per-class counts in Table 1 arise from an offline manifest and class-accounting step. Public benchmark class identifiers were retained to identify the original records, verify the counts, construct the balanced reporting tables, and score the final target-query set. They did not determine target-role membership: the same predefined partition rule was applied to each CWRU target record, and the same time-order rule was applied to each JNU target record. After the manifests were frozen, the target-alignment loader exposed only the signal path and sample coordinate. Consequently, "label-blind" in this article refers to adaptation and model selection, not to the absence of benchmark annotations during offline data accounting. Table 3 reports the permitted target-label access at each study stage.

Table 3. Target-label access by study stage.

Stage	Target Labels?	Permitted Use
Manifest construction and class accounting	Yes	Identify public benchmark records, verify category counts, and construct balanced reporting tables; this is an offline dataset-construction step.
Target-role allocation	No	Apply the same CWRU partition rule or JNU time-order rule to every target recording.
CORAL, MMD, and DANN training	No	Read target-alignment signal paths and sample coordinates only.
Checkpoint selection and hyperparameter choice	No	Select by source-validation Macro-F1 only.
Final target-query evaluation	Yes	Load labels once after checkpoint selection to calculate the reported metrics and confusion matrices.

2.4. Input Representations

For a window x, both representations begin with per-window z-score standardization. The time-domain representation retains the normalized 1024-point sequence. The spectral representation is

$$z = \frac{\log(1 + |\mathrm{rFFT}(x)|)}{\|\log(1 + |\mathrm{rFFT}(x)|)\|_2 + \epsilon}.$$

Here, rFFT retains the nonredundant 513 coefficients for a real-valued input, and ϵ avoids division by zero. Finally, feature-wise mean and standard deviation are estimated only from source-training windows and applied to the corresponding source-validation, target-alignment, and target-query representations. The spectral projection is fixed and contains no learned transform parameters.

The two branches share the layer configuration and optimization protocol but not the input length: each time-domain window contains 1024 samples, whereas the real FFT representation contains 513 nonredundant frequency bins. Consequently, using the same CNN architecture does not make the representations fully equivalent, because intermediate sequence lengths and the relative scale of convolutional receptive fields differ. Fixed 16-bin average pooling maps both final feature maps to 1024-dimensional vectors and controls the downstream feature dimensionality; it does not equalize input information or effective receptive fields. The results are therefore interpreted as a controlled architecture-and-protocol comparison of two fixed representations, rather than as a fully matched-input ablation.

2.5. Network and Adaptation Objectives

All conditions use the same single-input-channel CNN: Conv1D layers with 32, 64, and 64 channels, fixed 16-bin average pooling, a 128-dimensional embedding, and a task-specific classifier. The single-channel design avoids nominal multi-channel replication as an additional confound, but it does not remove the input-length difference noted above: the channel count, layer configuration, and downstream feature dimension are held fixed, so the comparison remains architectural and protocol-level rather than matched-input.

Figure 4 summarizes the common architecture. Both time and log-rFFT inputs enter as a single channel; the Conv1D blocks use 32, 64, and 64 channels with kernel sizes seven, seven, and five, respectively, and max pooling by two follows only the first two blocks. Fixed-bin average pooling maps either feature-map length to 16 bins per channel, giving $64 \times 16 = 1024$ features followed by a 128-dimensional embedding and a task classifier.

CORAL, MMD, and DANN operate on that embedding; DANN alone passes it through a gradient-reversal layer to the domain discriminator.

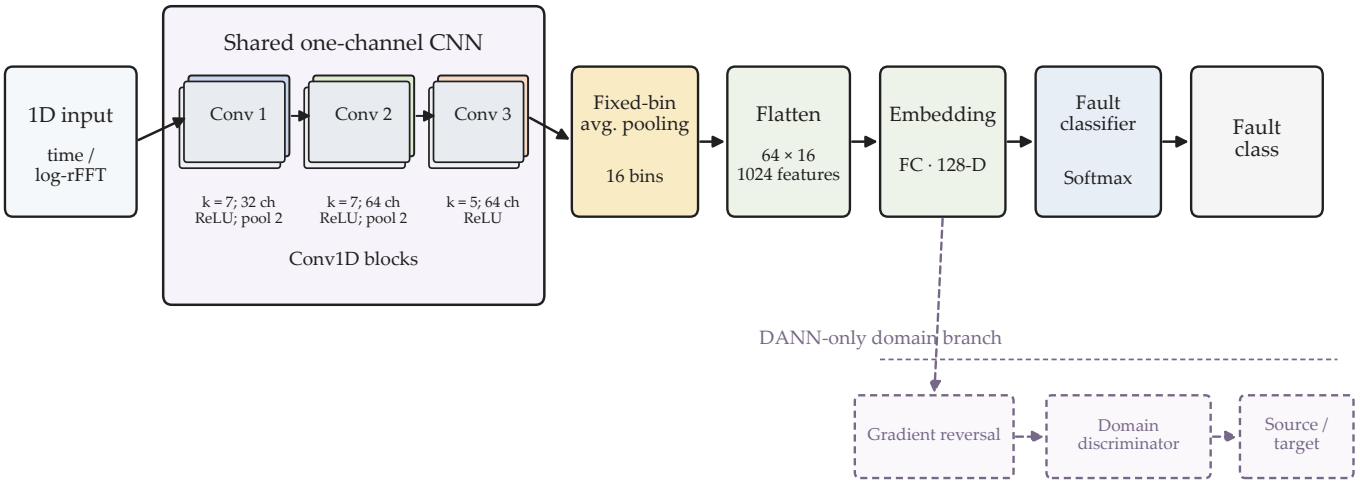


Figure 4. Common CNN configuration for 1024-sample time and 513-bin rFFT inputs. Fixed 16-bin pooling yields 1024 features; DANN adds a gradient-reversal discriminator to the 128-dimensional embedding.

Let f_θ denote the feature extractor and g_ϕ the classifier. For source mini-batch size n_s, x_i^s and y_i^s denote the ith source input and label. Source classification is optimized by the cross-entropy loss in Equation (1),

$$\mathcal{L}_{\text{cls}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \ell\left(g_\phi(f_\theta(x_i^s)), y_i^s\right). \tag{1}$$

For Deep CORAL, the loss aligns source and target embedding covariance matrices [8]. For MMD, n_t is the target mini-batch size, x_j^t is the jth target input, and the implementation uses an RBF-kernel discrepancy with mini-batch median bandwidth (Equation (2)),

$$\begin{aligned} \mathcal{L}_{\text{MMD}} = & \frac{1}{n_s^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} k(h_i^s, h_j^s) \\ & + \frac{1}{n_t^2} \sum_{i=1}^{n_t} \sum_{j=1}^{n_t} k(h_i^t, h_j^t) \\ & - \frac{2}{n_s n_t} \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} k(h_i^s, h_j^t). \end{aligned} \tag{2}$$

Here, $h_i^s = f_\theta(x_i^s)$ and $h_j^t = f_\theta(x_j^t)$ are embeddings, and k is the RBF kernel. The total CORAL or MMD loss is $\mathcal{L}_{\text{cls}} + \lambda \mathcal{L}_{\text{align}}$, where $\mathcal{L}_{\text{align}}$ is the active CORAL or MMD alignment term and $\lambda = 1$. DANN adds a domain discriminator with a gradient-reversal layer; its reversal coefficient follows the standard sigmoid schedule and reaches 1.0 by the end of training [9]. The source-only baseline minimizes only $\mathcal{L}_{\text{cls}}$.

For the MMD condition, the RBF bandwidth is the median pairwise squared distance within the joint source–target mini-batch, with a small positive lower bound for numerical stability. The MMD loss is therefore used as a training regularizer, not a separate target-domain validation criterion. For CORAL, centered source and target embedding covariance matrices are matched. For DANN, a small domain discriminator receives the same 128-dimensional embedding through a gradient-reversal layer. None of these objectives receives target fault labels. Keeping the classifier, optimizer, epochs, batch size, and

target-alignment fraction fixed across methods makes their differing objectives the primary intended source of variation, while recognizing that optimization dynamics can still differ.

2.6. Training, Reporting, and Implementation Controls

Each condition was trained for 50 epochs with Adam, batch size 32, and learning rate 10^{-4}. The five training seeds were 42, 2026, 0, 7, and 2027. The target-alignment subset seed was fixed at 20,261,006; for budget experiments, 10%, 25%, 50%, 75%, and 100% subsets were nested prefixes of the same fixed permutation. CUDA deterministic algorithms were enabled and TF32 was disabled. A synthetic forward/backward preflight was completed before formal runs.

The primary comparison uses CWRU hp0-to-hp3, CWRU hp3-to-hp0, and JNU-S600; each includes FFT source-only, FFT-CORAL, FFT-MMD, FFT-DANN, time-domain source-only, and time-domain MMD at the 10% alignment budget. To keep the controlled comparison tractable, the time-domain arm is limited to source-only and MMD, whereas CORAL and DANN are evaluated only with the FFT representation. JNU-S800 and JNU-S1000 serve as boundary-condition tasks under the same six conditions, fixed seeds, preprocessing, and source-validation-only checkpoint rule. FFT-MMD budget curves are reported only for the three primary tasks. Macro-F1 is the primary measure; accuracy and aggregate confusion matrices are secondary diagnostics. Means and standard deviations summarize five training replications, and paired seed differences are descriptive rather than independent target-window replications. The fixed training and reporting controls are summarized in Table 4; target-label permissions are detailed separately in Table 3.

Table 4. Fixed training and reporting controls.

Item	Setting
CNN backbone	One channel; Conv1D channels 32/64/64; 16-bin average pooling; 128-dimensional embedding
Optimization	Adam; learning rate 10^{-4}; batch size 32; 50 epochs
Adaptation weight	CORAL/MMD weight 1.0; DANN maximum reversal weight 1.0 with a sigmoid schedule
Training seeds	42, 2026, 0, 7, and 2027
Alignment-subset seed	20,261,006
Checkpoint selection	Maximum source-validation Macro-F1 only
Partition unit	Non-overlapping 1024-point window; not an independently acquired recording
Reporting unit	Independent training replication, not a target-query window
Determinism	Deterministic CUDA algorithms; TF32 disabled; preflight forward/backward check

2.7. Generative-AI Disclosure

Generative-AI assistance was limited to language editing, organizational suggestions, and figure-layout assistance. It was not used to generate or interpret experimental data, conduct analyses, or determine scientific conclusions. All output was reviewed and edited by the authors, who take full responsibility for this manuscript.

3. Results

3.1. Primary Cross-Condition Comparison

Table 5 reports all six defined method–representation conditions for each of the five transfer tasks. Time-domain CORAL and DANN were not included in this comparison; conclusions about those objectives are therefore restricted to the FFT representation. On CWRU hp0-to-hp3, FFT-MMD reached 98.92 ± 0.34% Macro-F1 versus 77.21 ± 5.16% for time-domain MMD, while FFT-DANN was highest at 99.23 ± 0.10%. This pattern indicates that the observed gain depends on the joint choice of representation and objective rather than on the alignment loss alone.

Table 5. Target-query Macro-F1 (%) at the 10% target-alignment budget (five-seed mean ± SD).

Input and Method	CWRU 0→3 hp	CWRU 3→0 hp	JNU-600	JNU-800	JNU-1000
FFT source-only	95.61 ± 0.40	95.00 ± 1.95	72.52 ± 1.81	96.50 ± 0.23	92.53 ± 0.82
FFT-CORAL	96.93 ± 1.26	97.51 ± 1.57	86.52 ± 1.04	96.41 ± 0.37	93.52 ± 0.76
FFT-MMD	98.92 ± 0.34	96.84 ± 0.68	90.48 ± 1.58	96.74 ± 0.46	95.68 ± 0.81
FFT-DANN	99.23 ± 0.10	98.35 ± 1.00	89.75 ± 2.02	96.74 ± 0.46	95.62 ± 1.08
Time source-only	62.36 ± 1.81	67.61 ± 8.94	44.53 ± 7.44	98.50 ± 0.62	98.92 ± 1.05
Time-MMD	77.21 ± 5.16	72.32 ± 4.26	67.05 ± 3.75	99.42 ± 0.63	99.42 ± 0.37

The reverse CWRU direction has a different ranking: FFT-DANN reached 98.35 ± 1.00%, whereas FFT-MMD reached 96.84 ± 0.68% and did not exceed FFT-CORAL. JNU-S600 provides a second favorable spectral-MMD case (90.48 ± 1.58% versus 67.05 ± 3.75% for time-domain MMD), but it does not establish superiority over every UDA objective. Reporting both CWRU directions and JNU-S600 prevents a one-task result from being presented as generic load robustness. Figure 5 visualizes the FFT-domain results at the common 10% target-alignment budget.

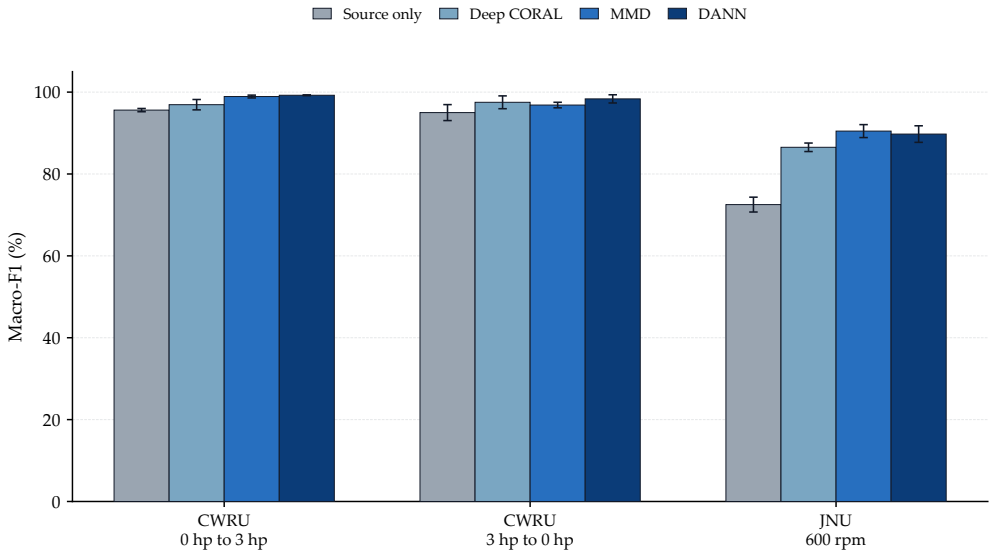


Figure 5. FFT-domain target-query Macro-F1 at the 10% alignment budget (five-seed mean ± SD).

3.2. Boundary Conditions

JNU-S800 and JNU-S1000 delimit the finding under the same complete six-condition coverage. Time-domain MMD exceeded FFT-MMD on both tasks (99.42 ± 0.63% versus 96.74 ± 0.46% at S800; 99.42 ± 0.37% versus 95.68 ± 0.81% at S1000). Thus, neither the fixed spectral representation nor one adaptation objective dominates every operating-condition shift.

3.3. Alignment-Budget and Class-Level Evidence

For the primary tasks, the FFT-MMD budget experiment varies the fraction of the label-free target-alignment pool. The five-seed means and standard deviations for the five fixed budgets are included in the experiment archive; the corresponding confusion matrices are provided in Appendix A. CWRU means stayed near ceiling (98.92–99.32% for hp0-to-hp3 and 96.84–99.27% for hp3-to-hp0), whereas JNU-S600 increased from 90.48% at 10% to 93.51% at 100%. These descriptive sequences do not establish a monotonic relationship;

the small CWRU differences should be interpreted with their seed-level variation. Figure 6 summarizes the corresponding FFT-MMD budget curves.

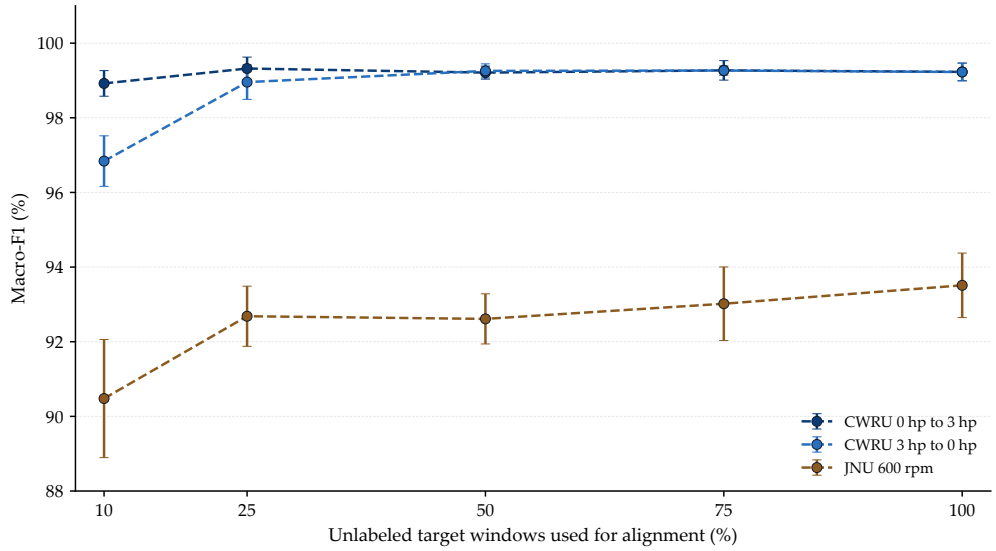
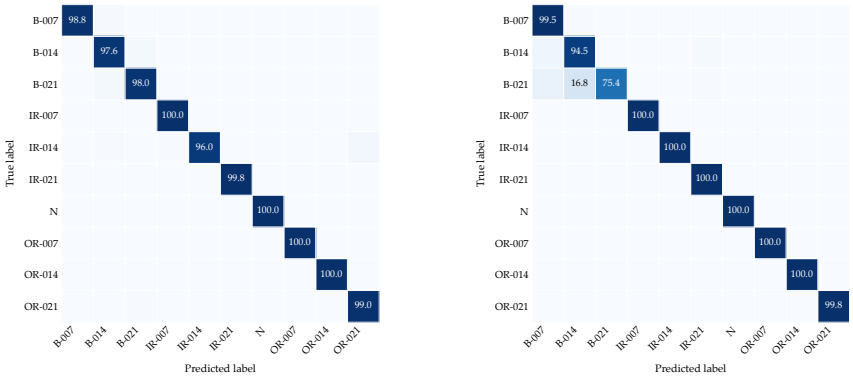


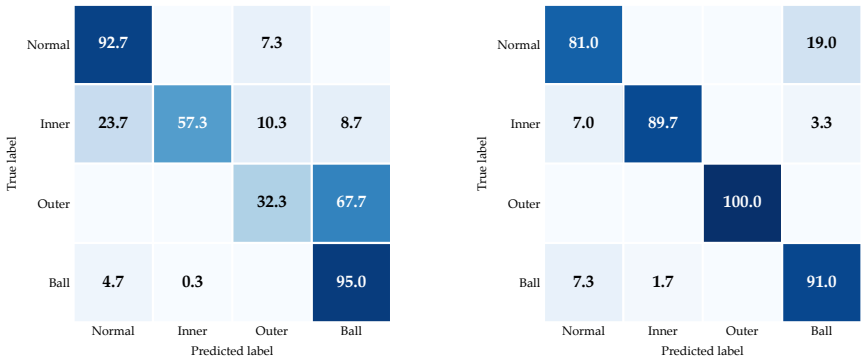
Figure 6. FFT-MMD Macro-F1 across target-alignment budgets (five-seed mean ± SD).

Aggregate confusion matrices provide class-level evidence without selecting a favorable seed. Each panel combines five seeds before row normalization on a fixed 0–100% scale (rows: true labels; columns: predicted labels). In the CWRU panels, B, IR, and OR denote ball, inner-race, and outer-race defects; the suffix is the defect size in thousandths of an inch. On hp3-to-hp0, the B021-to-B014 error fell from 31.60 ± 17.80% for FFT source-only to 16.80 ± 7.29% for FFT-MMD and 7.50 ± 9.63% for FFT-DANN. On JNU-S600, the Normal-to-Ball error fell from 71.67 ± 4.86% for FFT source-only to 19.00 ± 6.62% for FFT-MMD; DANN yielded 13.33 ± 9.86%. These descriptive class-boundary patterns do not identify a physical mechanism. Figure 7 shows representative panels; the complete method, budget, and boundary-condition atlas remains in Appendix A. Figure 8 shows the associated CWRU recall variation.



(**a**) CWRU 0→3 hp; FFT-MMD. (**b**) CWRU 3→0 hp; FFT-MMD.

Figure 7. *Cont.*



(**c**) JNU-S600; time-domain MMD. (**d**) JNU-S600; FFT-MMD.

Figure 7. Representative aggregate target-query confusion matrices at the 10% label-blind target-alignment budget. Cell colour encodes the row-normalised value on the common 0–100% scale.

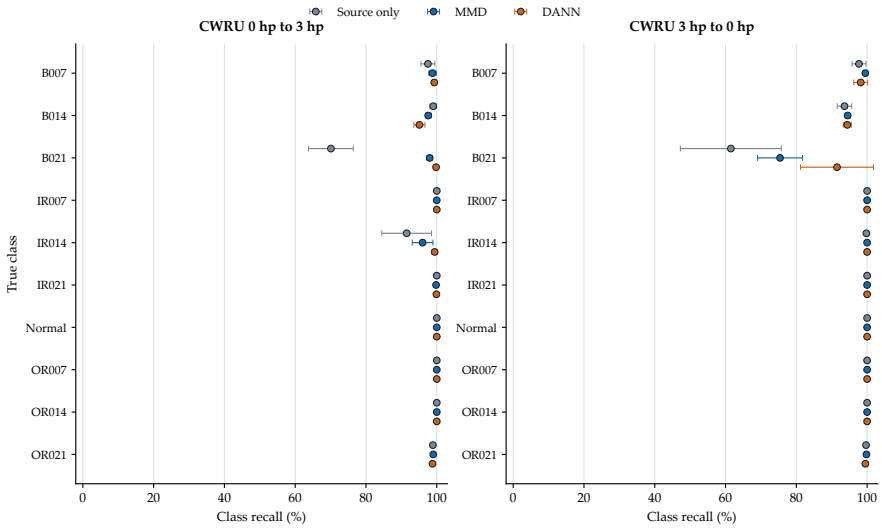


Figure 8. CWRU classwise target-query recall at the 10% alignment budget (five-seed mean ± SD).

4. Discussion

This controlled comparison tests a fixed log-amplitude spectral input rather than a new transform or architecture. FFT-MMD was favorable on the CWRU transfers and JNU-S600, DANN was best on the CWRU transfers, and time-domain MMD was stronger on JNU-S800 and JNU-S1000. The observed pattern therefore indicates a conditional representation effect, not evidence of a universal spectral advantage or a bearing-physics mechanism.

The deliberately simple baselines clarify that effect. MMD-based transfer, time–frequency fusion, and contrastive or adversarial variants have all been used for unlabeled bearing diagnosis [11–13]. In contrast, this study holds target-label access, backbone capacity, and checkpoint selection fixed while comparing three conventional objectives. The design complements those multi-component methods by showing that neither MMD nor the fixed spectral representation dominates every shift.

The safeguards do not remove all threats to validity. The manifests exclude duplicate output paths and overlapping raw intervals, but roles may use different regions of the same recording; the study therefore cannot claim recording-level independence. The JNU signals are resampled from 50 kHz to 12 kHz, and public seeded-fault benchmarks do not establish performance on evolving industrial damage, nonstationary noise, or new sensor

placements. Only source-only, CORAL, MMD, and DANN are evaluated; five seeds describe training variation, not independent machines or recordings. Finally, the budget study varies unlabeled alignment data, not labeled target data.

5. Conclusions

Under the manifest-defined, low-resource, label-blind UDA protocol, normalized log-rFFT input improved observed MMD performance on both CWRU transfers and JNU-S600, but not on every task. The contribution is empirical and protocol-oriented: with the same backbone, target-access budget, and source-only checkpoint rule, representation choice changes the observed effect of standard UDA objectives. The evidence is limited to class-balanced public benchmark recordings and does not demonstrate deployment performance or recording-level generalization. Validation on independent recordings, machines, and sensors is needed before deployment claims can be made.

Author Contributions: Conceptualization, X.X. and L.L.; methodology, X.X.; software, X.X. and Y.C.; validation, X.X. and Y.C.; formal analysis, L.L.; resources, X.X. and Y.C.; data curation, X.X. and Y.C.; writing—original draft preparation, X.X. and L.L.; writing—review and editing, L.L. and Y.C.; visualization, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of the data used in this study. The CWRU data were obtained from the Case Western Reserve University Bearing Data Center (Reference [28]); the JNU benchmark was obtained from the source described in Reference [10]. Reference [29] documents the CWRU test rig shown in Figure 2, and Reference [31] is the CC BY 4.0 source of the JNU rig shown in Figure 3. The training, evaluation, audit, and figure-export scripts, the frozen role manifests and protocol snapshots, the fixed seeds and seed-level results, the aggregate confusion matrices, and the data-materialization instructions are publicly available at Zenodo, Version 1.0.0, https://doi.org/10.5281/zenodo.23500472. The public archive includes a README and materialization-validation instructions for reconstructing the study windows and reproducing the reported workflow. The archive does not redistribute the third-party CWRU or JNU signal files, which must be obtained from their original providers. The complete aggregate confusion-matrix atlas is provided in Appendix A.

Acknowledgments: The authors acknowledge use of OpenAI GPT-5 (OpenAI, San Francisco, CA, USA) for language editing and figure-layout assistance.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

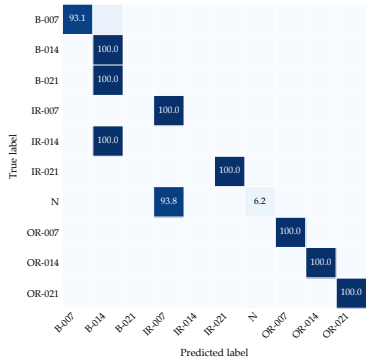
CNN	convolutional neural network
CORAL	correlation alignment
CWRU	Case Western Reserve University
DANN	domain-adversarial neural network
MMD	maximum mean discrepancy
UDA	unsupervised domain adaptation

Appendix A. Complete Aggregate Confusion-Matrix Atlas

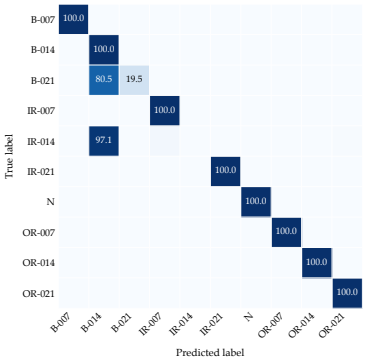
This appendix contains all 45 aggregate confusion matrices. Each panel combines the five fixed seeds before row normalization; rows are true labels and columns are predicted labels on a common 0–100% scale.

Appendix A.1. Primary Method Comparisons at the 10% Alignment Budget

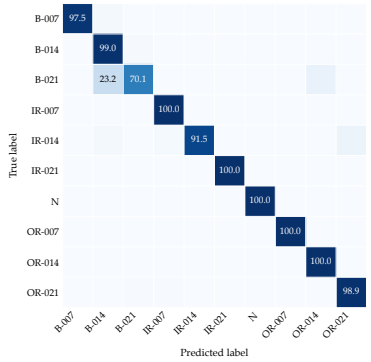
The six defined method–representation conditions are shown for the three primary tasks. In the CWRU panels, B, IR, and OR denote ball, inner-race, and outer-race defects; the suffix denotes the defect size in thousandths of an inch. JNU labels denote normal, inner-race, outer-race, and ball conditions. In all panels, cell colour encodes the row-normalised value on the same common 0–100% scale.



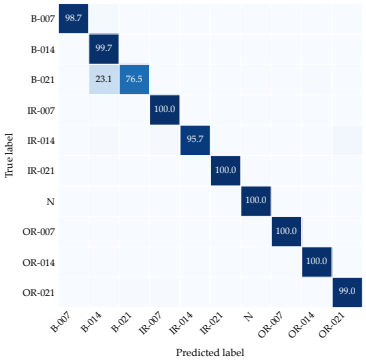
(**a**) Time source-only.



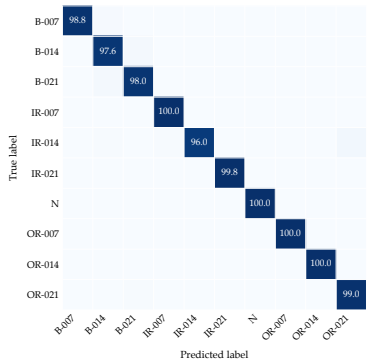
(**b**) Time MMD.



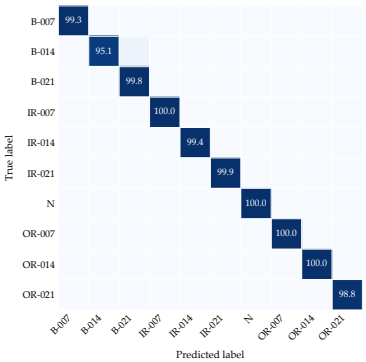
(**c**) FFT source-only.



(**d**) FFT-CORAL.

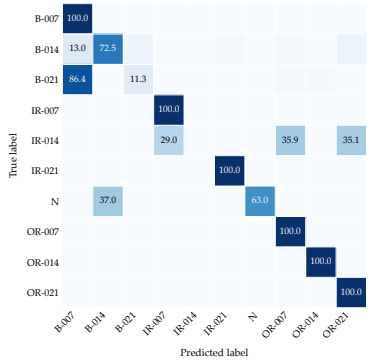


(**e**) FFT-MMD.

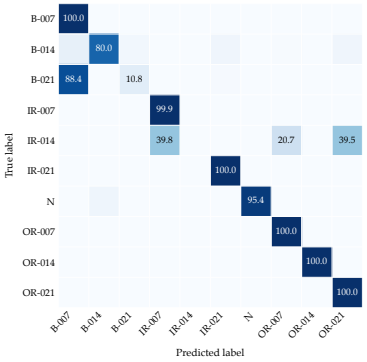


(**f**) FFT-DANN.

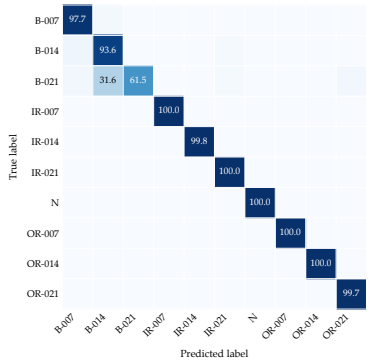
Figure A1. CWRU hp0-to-hp3, six methods at the 10% alignment budget.



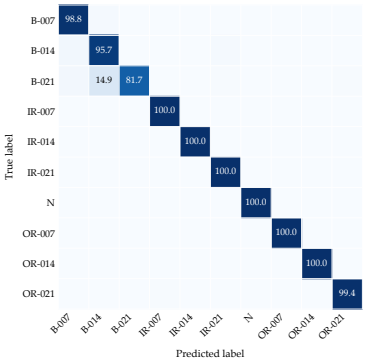
(**a**) Time source-only.



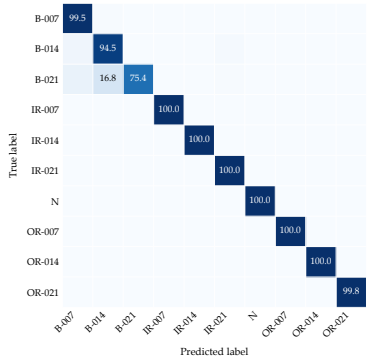
(**b**) Time MMD.



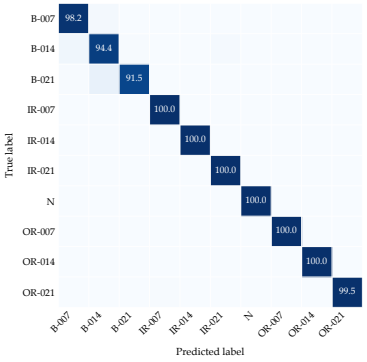
(**c**) FFT source-only.



(**d**) FFT-CORAL.



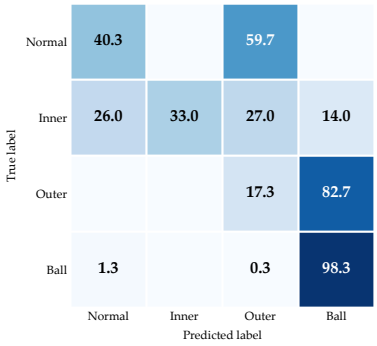
(**e**) FFT-MMD.



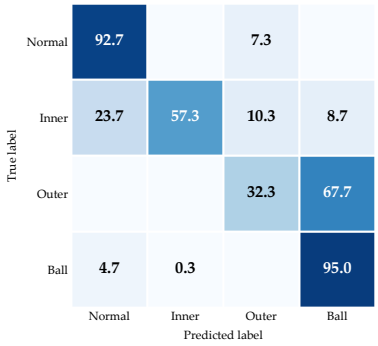
(**f**) FFT-DANN.

Figure A2. CWRU hp3-to-hp0, six methods at the 10% alignment budget.

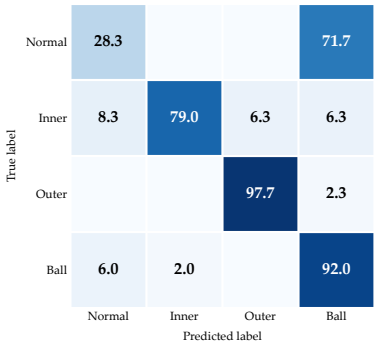
https://doi.org/10.3390/app162111053



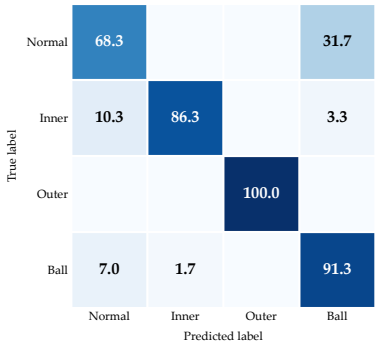
(**a**) Time source-only.



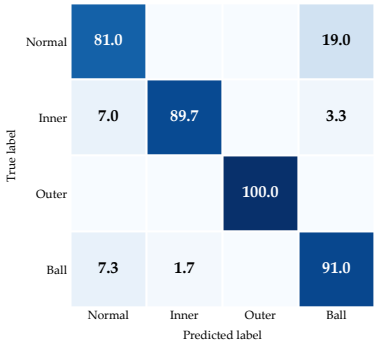
(**b**) Time MMD.



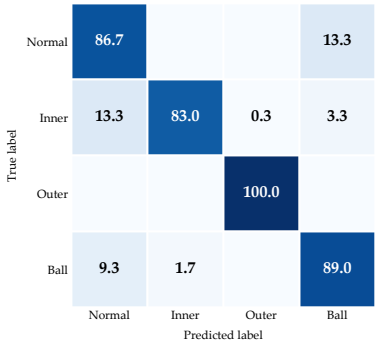
(**c**) FFT source-only.



(**d**) FFT-CORAL.



(**e**) FFT-MMD.

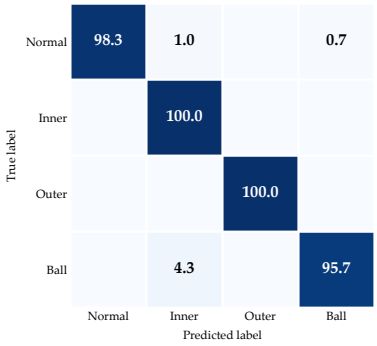


(**f**) FFT-DANN.

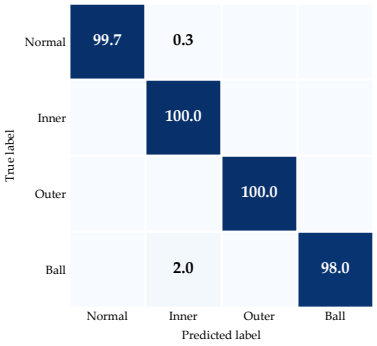
Figure A3. JNU-S600, six methods at the 10% alignment budget.

Appendix A.2. Boundary-Condition Tasks at the 10% Alignment Budget

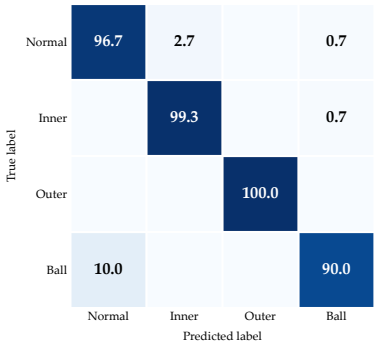
All six defined method–representation conditions were completed for both boundary tasks. Each panel aggregates the five fixed seeds before row normalization; the common scale facilitates visual comparison but does not constitute a window-level test.



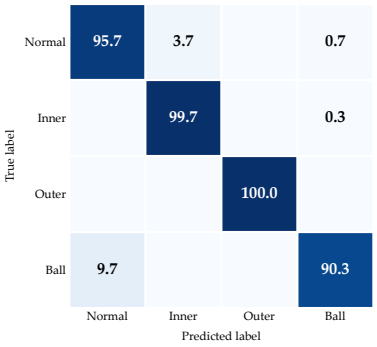
(**a**) Time source-only.



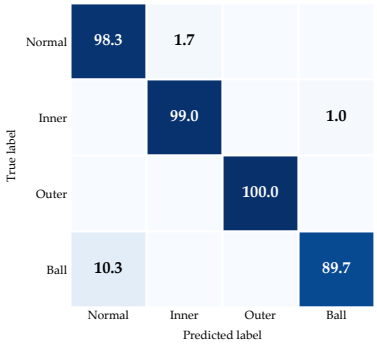
(**b**) Time MMD.



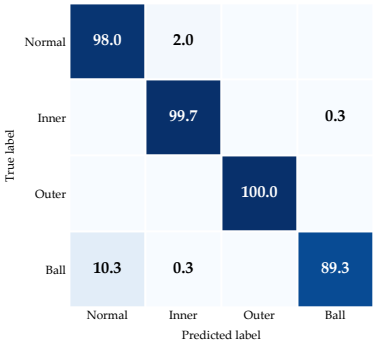
(**c**) FFT source-only.



(**d**) FFT-CORAL.

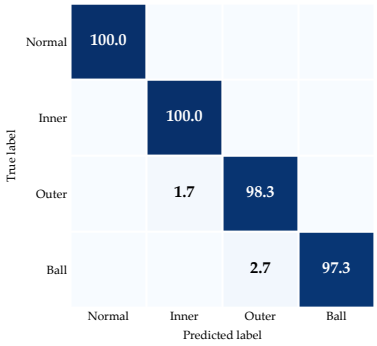


(**e**) FFT-MMD.

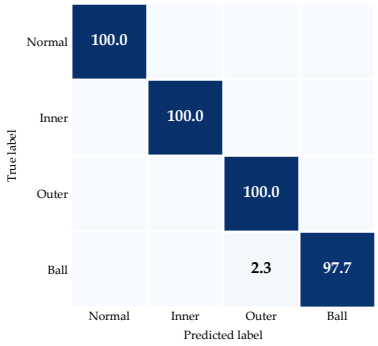


(**f**) FFT-DANN.

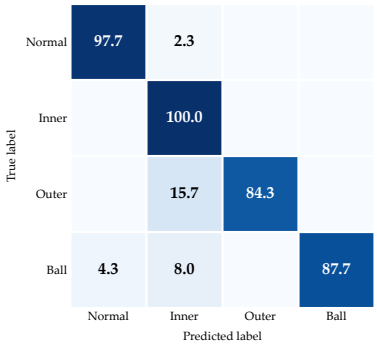
Figure A4. JNU-S800, six methods at the 10% alignment budget.



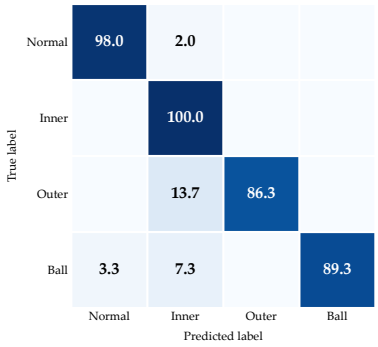
(**a**) Time source-only.



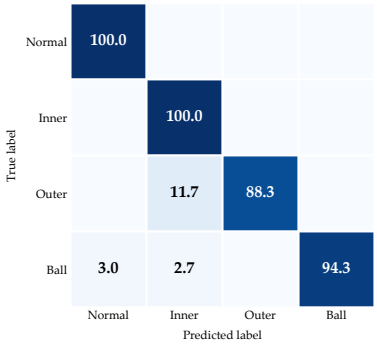
(**b**) Time MMD.



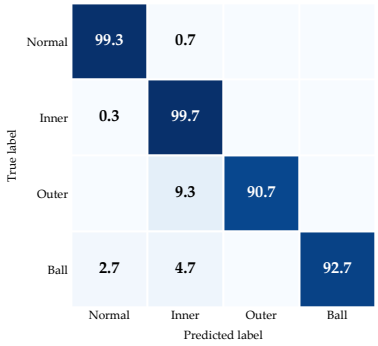
(**c**) FFT source-only.



(**d**) FFT-CORAL.



(**e**) FFT-MMD.

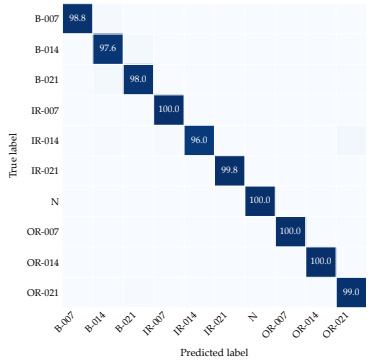


(**f**) FFT-DANN.

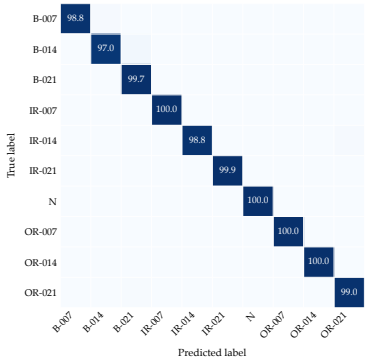
Figure A5. JNU-S1000, six methods at the 10% alignment budget.

Appendix A.3. FFT-MMD Target-Alignment-Budget Conditions

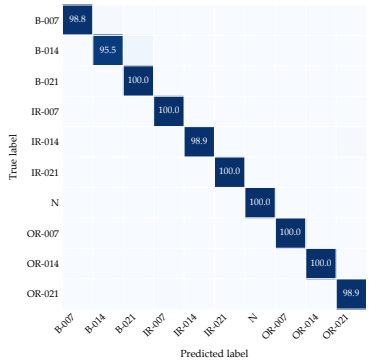
The panels below report the fixed 10%, 25%, 50%, 75%, and 100% fractions of the unlabeled target-alignment pool used by FFT-MMD. They are descriptive budget conditions, not a post hoc selection procedure.



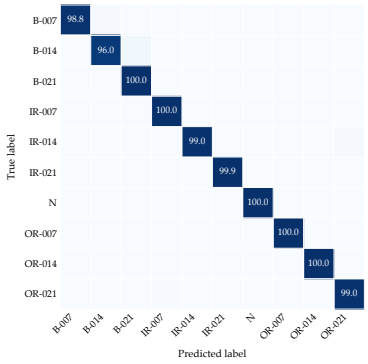
(**a**) 10%.



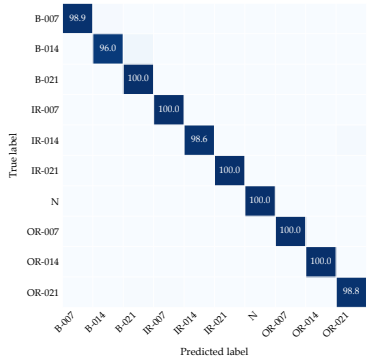
(**b**) 25%.



(**c**) 50%.

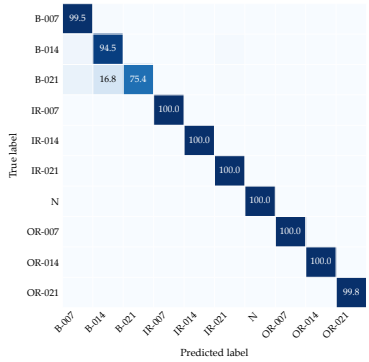


(**d**) 75%.

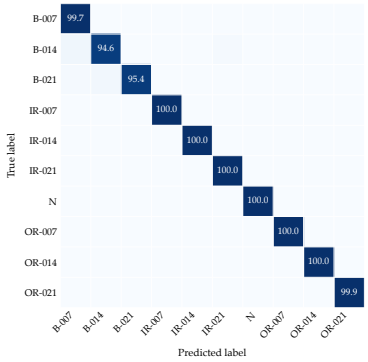


(**e**) 100%.

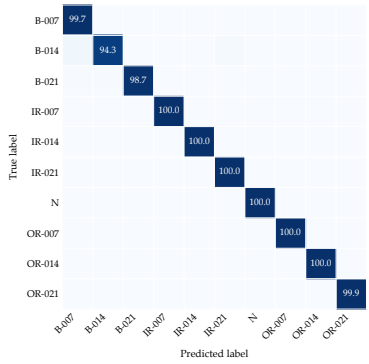
Figure A6. CWRU hp0-to-hp3, FFT-MMD across alignment budgets.



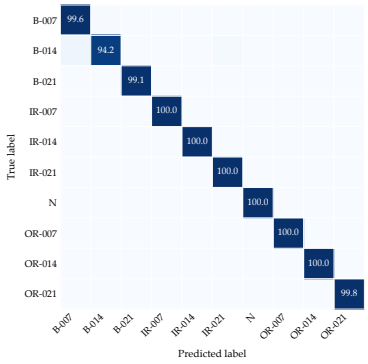
(**a**) 10%.



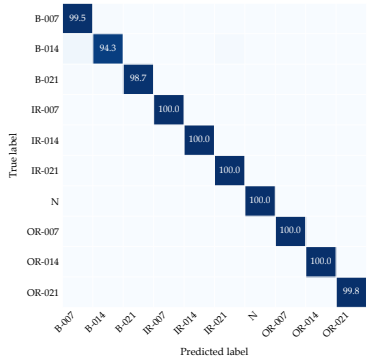
(**b**) 25%.



(**c**) 50%.

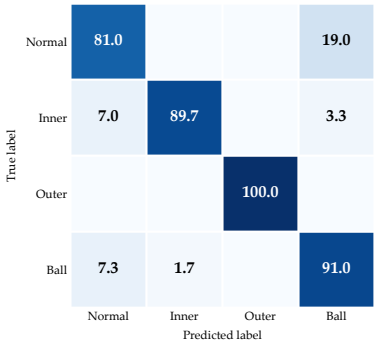


(**d**) 75%.

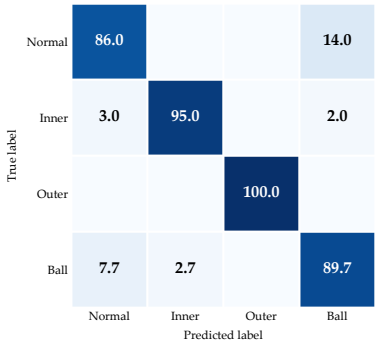


(**e**) 100%.

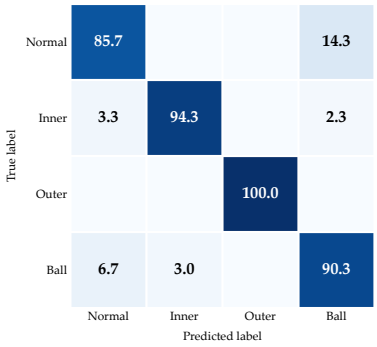
Figure A7. CWRU hp3-to-hp0, FFT-MMD across alignment budgets.



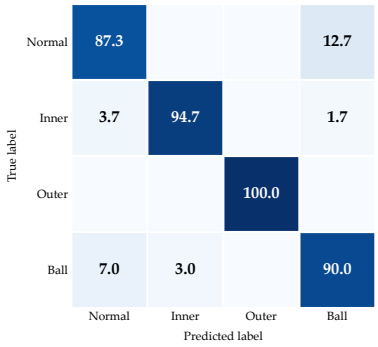
(**a**) 10%.



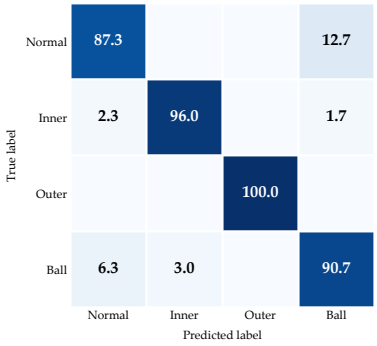
(**b**) 25%.



(**c**) 50%.



(**d**) 75%.



(**e**) 100%.

Figure A8. JNU-S600, FFT-MMD across alignment budgets.

References

1. Hoang, D.-T.; Kang, H.-J. A Survey on Deep Learning Based Bearing Fault Diagnosis. *Neurocomputing* **2019**, *335*, 327–335. [CrossRef]
2. Peng, B.; Bi, Y.; Xue, B.; Zhang, M.; Wan, S. A Survey on Fault Diagnosis of Rolling Bearings. *Algorithms* **2022**, *15*, 347. [CrossRef]
3. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of Machine Learning to Machine Fault Diagnosis: A Review and Roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [CrossRef]
4. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5990–5998. [CrossRef]
5. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]

6. Gretton, A.; Borgwardt, K.M.; Rasch, M.J.; Schölkopf, B.; Smola, A. A Kernel Two-Sample Test. *J. Mach. Learn. Res.* **2012**, *13*, 723–773.
7. Long, M.; Cao, Y.; Wang, J.; Jordan, M.I. Learning Transferable Features with Deep Adaptation Networks. In *Proceedings of the 32nd International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2015; Volume 37, pp. 97–105.
8. Sun, B.; Saenko, K. Deep CORAL: Correlation Alignment for Deep Domain Adaptation. In *Computer Vision–ECCV 2016 Workshops*; Springer International Publishing: Cham, Switzerland, 2016; pp. 443–450. [CrossRef]
9. Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; Lempitsky, V. Domain-Adversarial Training of Neural Networks. *J. Mach. Learn. Res.* **2016**, *17*, 1–35.
10. He, J.; Li, X.; Chen, Y.; Chen, D.; Guo, J.; Zhou, Y. Deep Transfer Learning Method Based on 1D-CNN for Bearing Fault Diagnosis. *Shock Vib.* **2021**, 6687331. [CrossRef]
11. Li, X.; Jiang, H.; Wang, R.; Niu, M. Rolling Bearing Fault Diagnosis Using Optimal Ensemble Deep Transfer Network. *Knowl.-Based Syst.* **2021**, *213*, 106695. [CrossRef]
12. Tao, H.; Qiu, J.; Chen, Y.; Stojanovic, V.; Cheng, L. Unsupervised Cross-Domain Rolling Bearing Fault Diagnosis Based on Time-Frequency Information Fusion. *J. Frankl. Inst.* **2023**, *360*, 1454–1477. [CrossRef]
13. An, Y.; Zhang, K.; Chai, Y.; Liu, Q.; Huang, X. Domain Adaptation Network Base on Contrastive Learning for Bearings Fault Diagnosis under Variable Working Conditions. *Expert Syst. Appl.* **2023**, *212*, 118802. [CrossRef]
14. Wu, Z.; Jiang, H.; Lu, T.; Zhao, K. A Deep Transfer Maximum Classifier Discrepancy Method for Rolling Bearing Fault Diagnosis under Few Labeled Data. *Knowl.-Based Syst.* **2020**, *208*, 105814. [CrossRef]
15. Zhang, R.; Tao, H.; Wu, L.; Guan, Y. Transfer Learning With Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions. *IEEE Access* **2017**, *5*, 14347–14357. [CrossRef]
16. Zhang, W.; Li, C.; Peng, G.; Chen, Y.; Zhang, Z. A Deep Convolutional Neural Network with New Training Methods for Bearing Fault Diagnosis Under Noisy Environment and Different Working Load. *Mech. Syst. Signal Process.* **2018**, *100*, 439–453. [CrossRef]
17. Lu, W.; Liang, B.; Cheng, Y.; Meng, D.; Yang, J.; Zhang, T. Deep Model Based Domain Adaptation for Fault Diagnosis. *IEEE Trans. Ind. Electron.* **2017**, *64*, 2296–2305. [CrossRef]
18. Wu, J.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Few-Shot Transfer Learning for Intelligent Fault Diagnosis of Machine. *Measurement* **2020**, *166*, 108202. [CrossRef]
19. Li, C.; Zhang, S.; Qin, Y.; Estupinan, E. A Systematic Review of Deep Transfer Learning for Machinery Fault Diagnosis. *Neurocomputing* **2020**, *407*, 121–135. [CrossRef]
20. Wilson, G.; Cook, D.J. A Survey of Unsupervised Deep Domain Adaptation. *ACM Trans. Intell. Syst. Technol.* **2020**, *11*, 51:1–51:46. [CrossRef] [PubMed]
21. Bhuiyan, M.R.; Uddin, J. Deep Transfer Learning Models for Industrial Fault Diagnosis Using Vibration and Acoustic Sensors Data: A Review. *Vibration* **2023**, *6*, 218–238. [CrossRef]
22. Guo, Y.; Zhang, J.; Sun, B.; Wang, Y. Adversarial Deep Transfer Learning in Fault Diagnosis: Progress, Challenges, and Future Prospects. *Sensors* **2023**, *23*, 7263. [CrossRef] [PubMed]
23. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A Survey of Transfer Learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
24. Tzeng, E.; Hoffman, J.; Saenko, K.; Darrell, T. Adversarial Discriminative Domain Adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2017; pp. 7167–7176. [CrossRef]
25. Saito, K.; Watanabe, K.; Ushiku, Y.; Harada, T. Maximum Classifier Discrepancy for Unsupervised Domain Adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: New York, NY, USA, 2018; pp. 3723–3732. [CrossRef]
26. Long, M.; Zhu, H.; Wang, J.; Jordan, M.I. Deep Transfer Learning with Joint Adaptation Networks. In *Proceedings of the 34th International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2017; Volume 70, pp. 2208–2217.
27. Feng, Z.; Liang, M.; Chu, F. Recent Advances in Time-Frequency Analysis Methods for Machinery Fault Diagnosis: A Review with Application Examples. *Mech. Syst. Signal Process.* **2013**, *38*, 165–205. [CrossRef]
28. Case Western Reserve University Bearing Data Center. Bearing Data Center. Available online: https://engineering.case.edu/bearingdatacenter/welcome (accessed on 7 October 2026).
29. Case Western Reserve University Bearing Data Center. Apparatus & Procedures. Available online: https://engineering.case.edu/bearingdatacenter/apparatus-and-procedures (accessed on 7 October 2026).
30. Smith, W.A.; Randall, R.B. Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study. *Mech. Syst. Signal Process.* **2015**, *64–65*, 100–131. [CrossRef]
31. Wang, J.; Wang, Y.; Yu, J.; Li, Q.; Wang, H.; Zhou, X. Research on Bearing Fault Diagnosis Method for Varying Operating Conditions Based on Spatiotemporal Feature Fusion. *Sensors* **2025**, *25*, 3789. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.