Engineering 56 (2026) 267–283

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Geotechnical Engineering—Article

Lightweight and Robust Cross-Domain Microseismic Signal Classification Framework with Bi-Classifier Adversarial Learning

Dingran Song [a], Feng Dai [a,b,*], Yi Liu [a], Hao Tan [a], Mingdong Wei [a]

[a] State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource and Hydropower, Sichuan University, Chengdu 610065, China
[b] College of Architecture and Environment, Sichuan University, Chengdu 610065, China

ARTICLE INFO

Article history:
Received 8 September 2024
Revised 23 July 2025
Accepted 20 October 2025
Available online 4 November 2025

Keywords:
Deep learning
Microseismic classification
Lightweight design
Noise robustness
Unsupervised domain adaptation

ABSTRACT

Automatic identification of microseismic (MS) signals is crucial for early disaster warning in deep underground engineering. However, three major challenges remain for practical deployment, namely limited resources, severe noise interference, and data scarcity. To address these issues, this study proposes the lightweight and robust entropy-regularized unsupervised domain adaptation framework (LRE-UDAF) for cross-domain MS signal classification. The framework comprises a lightweight and robust feature extractor and an unsupervised domain adaptation (UDA) module utilizing a bi-classifier disparity metric and entropy regularization. The feature extractor derives high-level representations from the preprocessed signals, which are subsequently fed into two classifiers to predict class probability. Through three-stage adversarial learning, the feature extractor and classifiers progressively align the distributions of the source and target domains, facilitating knowledge transfer from the labeled source to the unlabeled target domain. Source-domain experiments reveal that the feature extractor achieves high effectiveness, with a classification accuracy of up to 97.7%. Moreover, LRE-UDAF outperforms prevalent industry networks in terms of its lightweight design and robustness. Cross-domain experiments indicate that the proposed UDA method effectively mitigates domain shift with minimal unlabeled signals. Ablation and comparative experiments further validate the design effectiveness of the feature extractor and UDA modules. This framework presents an efficient solution for resource-constrained, noise-prone, and data-scarce environments in deep underground engineering, offering significant promise for practical implementations in early disaster warning.

© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Deep underground engineering projects frequently experience instability-induced disasters during construction owing to high geo-stress and strong dynamic disturbances, resulting in significant casualties and financial losses [1,2]. Microseismic (MS) monitoring, a real-time three-dimensional technique, is increasingly used for early warnings of deep underground disasters [3–5]. Manual processing of MS data remains common; however, it depends on experienced engineers and suffers from high latency, highlighting the growing need for automated, timely processing [6,7], as depicted in Fig. 1. As a prerequisite for MS data processing, automatic and accurate identification of MS signals is crucial for effective early warning of rock engineering disasters [8–10].

Conventional machine learning-based classification methods require extensive feature engineering, heavily relying on the domain expertise of researchers and advanced signal processing techniques [11–13]. Moreover, handcrafted features often inadequately capture the intrinsic characteristics of raw signals. In contrast, deep learning methods for MS signal classification enable end-to-end processing, from feature extraction to signal classification; the use of convolutional neural networks (CNNs) has been particularly prominent. For instance, Tang et al. [14] developed ResSCA, a CNN architecture utilizing residual skip connections and a dual-attention mechanism for MS signal classification. He et al. [15] proposed the universal automatic classification network, classifying MS signals of varying waveform sizes using convolutional layers, gate structures, and adaptive pooling layers. Ma et al. [16] proposed a bimodal CNN classification model that employs an attention mechanism to learn bimodal features simultaneously from the time and frequency domains. Additionally, Bi

* Corresponding author.
E-mail address: fengdai@scu.edu.cn (F. Dai).

https://doi.org/10.1016/j.eng.2025.10.023
2095-8099/© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

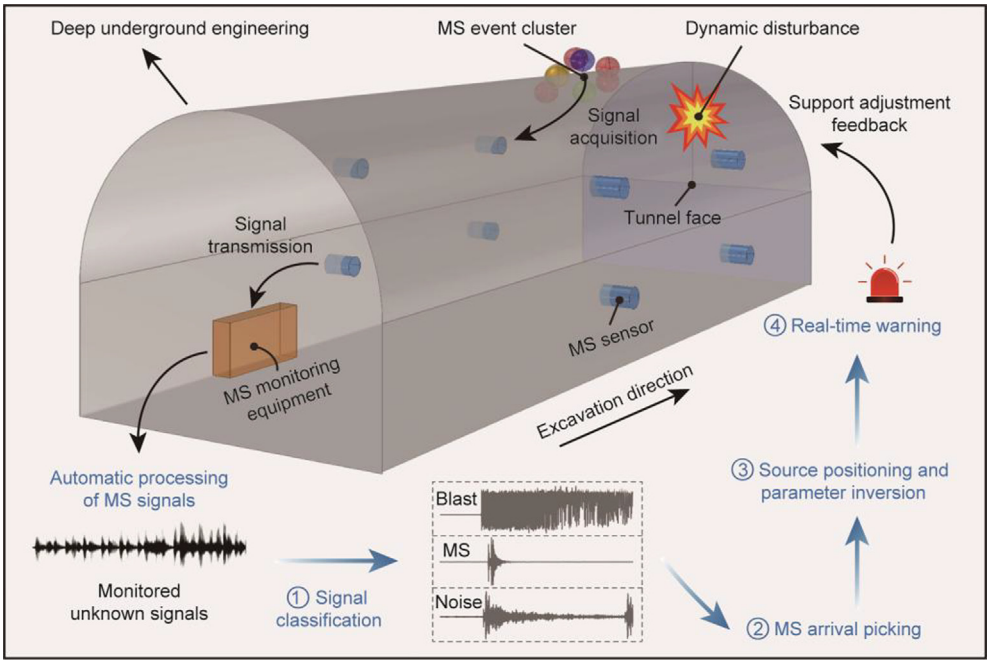

Fig. 1. Workflow for automatic processing of microseismic (MS) signals.

et al. [17] developed the explainable time–frequency CNN, which learns high-level fused features from both domains while generating fine-grained feature maps for classification explanations. Despite these advancements, MS classification algorithms still face practical challenges, such as lightweight models tailored to harsh deployment environments and robustness to complex on-site noise. Consequently, there is an urgent need for specialized neural network architectures to address these issues effectively.

The classification performance of deep learning methods heavily relies on a substantial number of well-labeled training samples [18,19]. The same model can show significant performance variations across different datasets. At the early stages of MS monitoring, limited MS data often impede the development of an optimal classification model. Furthermore, variations in geological conditions during excavation can impact the predictive performance of the model. Therefore, it is crucial to leverage available monitoring data from diverse distributions for the rapid development of effective deep learning models tailored to current scenarios.

Domain adaptation (DA), a prominent branch of transfer learning, effectively tackles challenges arising from discrepancies in data distribution [20,21]. When target tasks align, DA allows models trained on the source domain to adapt to related yet distinct target domains, often using only a limited number of samples from the target domain. This principle is illustrated in Fig. 2. When target domain data is unlabeled, it is referred to as unsupervised DA

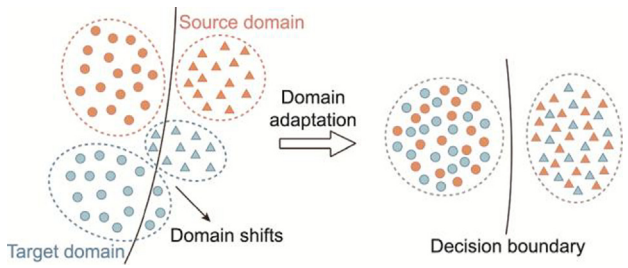

Fig. 2. Schematic of DA.

(UDA), which is essential across various industries [22,23]. Zhu et al. [24] proposed the multi-representation adaptation network, capturing diverse information to achieve cross-domain image classification through multi-representation alignment. Feng et al. [25] interpreted variations in electrocardiogram signals between individuals as domain shifts and proposed an unsupervised semantic-aware adaptive feature fusion network for intersubject arrhythmia detection. He et al. [26] developed a multilevel UDA framework with a multilevel DA optimization mechanism to address individual differences in electrocardiogram signal classification. Sun et al. [27] presented domain-invariant features evaluation and knowledge distillation, a UDA method for cross-domain bearing fault diagnosis based on domain-invariant feature evaluation and knowledge distillation. In routine MS monitoring, large volumes of unlabeled signals are detected daily, making UDA one of the most time- and resource-efficient transfer learning methods for such applications.

In summary, practical engineering deployment of MS signal classification algorithms faces three core challenges. First, the harsh deployment environments in deep underground engineering, such as high temperatures and weak network coverage, necessitate the development of lightweight models that enable real-time processing of large MS datasets. Second, complex on-site noise severely contaminates pure MS signals, significantly reducing the signal-to-noise ratio (SNR) and requiring models with robust noise resilience. Finally, in the initial phases of monitoring projects, a lack of labeled data complicates traditional supervised learning methods, leading to reduced model performance due to domain shifts. These challenges represent significant bottlenecks in implementing MS-classification algorithms, highlighting the urgent need for a systematic solution that addresses resource constraints, enhances noise robustness, and integrates UDA capabilities. In this study, we devised a novel UDA framework called the lightweight and robust entropy-regularized UDA framework (LRE-UDAF) for end-to-end MS signal classification. The main contributions of this study are as follows:

(1) A lightweight and robust feature extractor was designed, consisting primarily of an improved ShuffleNet unit (ISNU) module

with high computational efficiency and a dual attention adaptive residual shrinkage block (DAARSB) module with adaptive threshold denoising. Comparative evaluations demonstrate that the proposed feature extractor outperforms other state-of-the-art (SOTA) networks in terms of classification accuracy, model complexity, and noise robustness.

(2) A novel UDA framework for MS signal classification is proposed, leveraging a bi-classifier adversarial paradigm with entropy regularization to effectively mitigate distribution discrepancies. This framework addresses the issue of insufficient data in the early stages of MS monitoring and enhances the rapid adaptation capability of the model to geological changes.

(3) The effectiveness of the proposed framework is validated across two practical projects with distinct data distributions. The classification results and visual analyses pre- and post-DA demonstrate the exceptional transfer learning capabilities of the LRE-UDAF method.

The remainder of this study is organized as follows. Section 2 details the proposed method. Sections 3 and 4 present the experimental results for source and cross-domain classifications, respectively. Section 5 discusses the effectiveness of each component design through ablation studies and comparative analysis. Finally, we summarize the study in Section 6.

2. Methodology

To enhance clarity and ensure consistent terminology throughout this section, we provide a comprehensive glossary of all abbreviations used in the proposed UDA framework (Table 1).

In this study, we employed the UDA framework for cross-domain MS signal classification challenges. The labeled source domain is denoted as $\{X_s, Y_s\} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, where x_i^s represents the ith sample from the source domain, y_i^s is its corresponding label, and N_s is the total number of samples. The unlabeled target domain is denoted as $\{X_t\} = \{x_j^t\}_{j=1}^{N_t}$, where x_j^t denotes the jth unlabeled target domain sample and N_t is the number of samples. The source label Y_s and target label Y_t share the same label space, that is, $Y_s = Y_t = Y$. Moreover, the source data X_s and target data X_t have related but distinct distributions. The objective of the UDA is to train a model on labeled source data to predict unlabeled target data. The overall architecture of the proposed LRE-UDAF is shown in Fig. 3; it consists of a feature extractor F and two classifiers C_1 and C_2. Preprocessed signals sampled from both domains are fed to F to learn high-level features. C_1 and C_2 receive the flattened high-level representation and output the corresponding prediction probabilities. During DA, F engages in adversarial learning with the two classifiers, utilizing three main loss functions: supervised classification loss ($\mathscr{L}_{cls}(X_s, Y_s)$), entropy regularization loss ($\mathscr{L}_{ent}(X_t)$), and discrepancy loss ($\mathscr{L}_{cdd}(X_t)$). The detailed network structure and optimization process are as follows.

Table 1
Glossary of abbreviations in Section 2.

Abbreviation	Full name
Conv	Convolutional layer
MaxPool	Max pooling layer
BN	Batch normalization
ReLU	Rectified linear unit
AvgPool	Average pooling layer
Concat	Concatenation operation
SoftMax	Softmax activation function
DW	Depthwise
MLP	Multilayer perceptron
ASPP	Atrous spatial pyramid pooling

2.1. Feature extractor module

The feature extractor module was designed to be a lightweight and robust architecture, as depicted in Fig. 4. This is achieved by integrating two key components, ISNU and DAARSB, each uniquely enhancing performance and adaptability.

2.1.1. ISNU module

Underground construction networks are often weak, and numerous MS events occur shortly after blasting, necessitating high calculation speed and accuracy of deep learning models. While deeper CNNs are commonly employed to improve the accuracy of the target task [28–30], deeper layers and more channels inevitably increase floating-point operations (FLOPs), adversely affecting computational efficiency. Consequently, more efficient model architectures are required to enhance prediction accuracy without additional FLOPs.

Recently, lightweight model designs have garnered considerable attention in deep learning, particularly under resource-constrained scenarios. To enhance computational efficiency, researchers have proposed various strategies, with DW separable and group convolutions widely employed to reduce model computational complexity while preserving representational capacity. However, these methods have limitations. For example, group convolution may restrict information flow between channels, potentially diminishing the representational power of the model. To address these issues, Ma et al. [31] proposed four practical guidelines for an efficient network architecture design and developed ShuffleNet-V2, which introduces the channel shuffle operation. This operation allows group convolutions to access the input information from different groups, facilitating communication between channel groups. Building on this, the ISNU integrates the "squeeze-and-excitation" (SE) block [32] to enhance prediction accuracy with minimal computational overhead. The SE block dynamically adjusts channel weights through "squeeze" and "excitation" operations, enhancing model robustness. The first linear transformation reduces the number of channels to one-fourth of the original, and the second applies the hard-sigmoid activation function to generate a channel-wise attention weight vector, which dynamically rescales the importance of each channel before the final linear transformation.

Fig. 5 illustrates the ISNU structure with a stride of 1 (Fig. 5(a)), enabling precise feature extraction without altering signal length and a stride of 2 (Fig. 5(b)), where input signals are downsampled to extract more abstract features. As shown in Fig. 4, the feature extractor module is built around the ISNU and is structured in three stages to progressively extract deeper-level features. Each stage begins with downsampling to reduce feature map size and extract abstract features, followed by refinement while maintaining the same dimensions.

2.1.2. DAARSB module

The noisy construction environment introduces various noise types that contaminate MS signals, significantly impairing the feature extraction capability of CNNs. To address this challenge, we propose the DAARSB module, inspired by deep residual shrinkage networks (DRSNs) [33]. This module combines threshold functions with deep learning to eliminate noise interference. Through iterative training, the deep learning model transforms feature representations into a space where essential information exhibits large absolute values, while redundant information approaches zero. Subsequently, the threshold function eliminates redundant information by setting it to zero, preserving only the most discriminative features. Compared to classical wavelet thresholding methods, the DAARSB module demonstrated superior noise

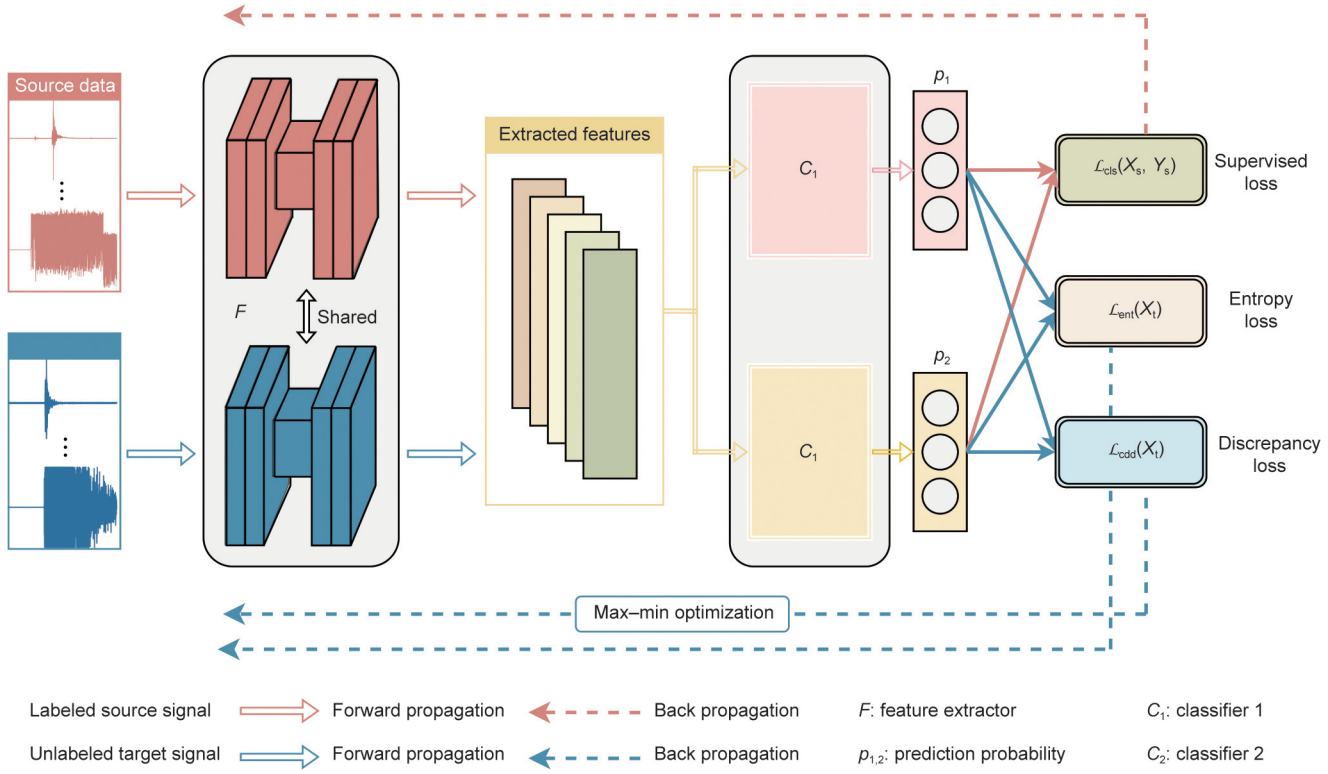

Fig. 3. Overview of the proposed LRE-UDAF. Max–min: maximum–minimum.

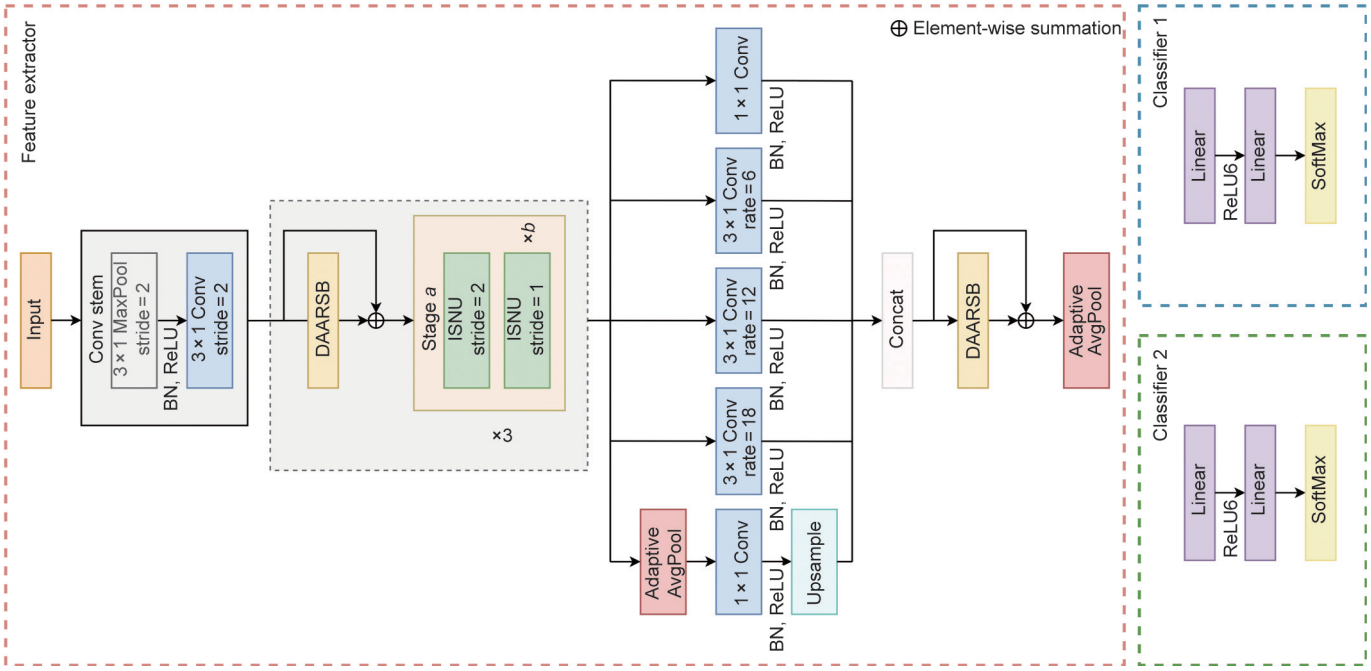

Fig. 4. Detailed architecture of LRE-UDAF. The notation "×*b*" indicates that the ISNU block (with stride = 1) is repeated *b* times within each stage. The term "Stage *a*" (where *a* = 1, 2, 3) denotes the repetition of the entire stage module three times in sequence.

suppression and feature robustness through its dynamic threshold learning and residual shrinkage architecture.

Key to the DAARSB module is the selection of the threshold function and adaptive determination of thresholds. Compared to DRSN, the DAARSB module achieved further optimization and enhancement. First, it incorporates an improved threshold function that addresses the limitations of the traditional hard and soft threshold functions, as shown in Eqs. (1) and (2), respectively.

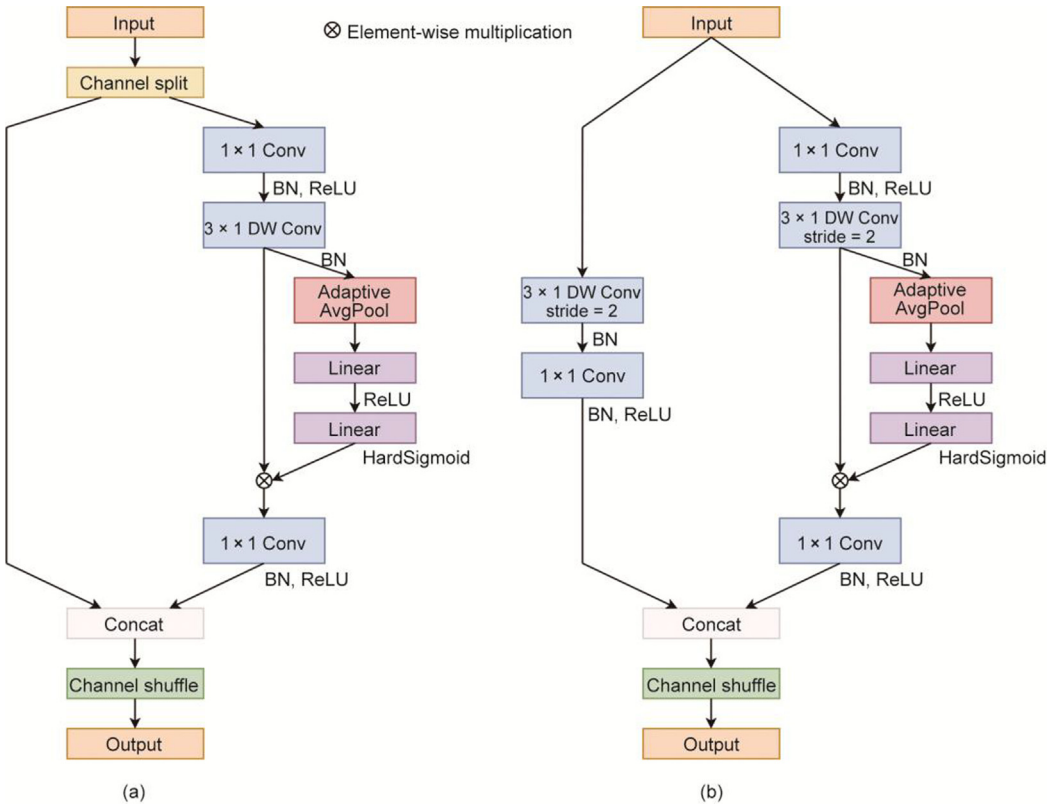

Fig. 5. Architecture of ISNU module: (a) stride = 1 and (b) stride = 2.

$$v = \begin{cases} u & |u| \geq \tau \\ 0 & |u| < \tau \end{cases} \tag{1}$$

$$v = \begin{cases} u - \tau & u \geq \tau \\ 0 & |u| < \tau \\ u + \tau & u \leq -\tau \end{cases} \tag{2}$$

where u and v are the input and output features, respectively. Threshold τ is a positive parameter. The discontinuity of the hard threshold function at the threshold point can induce oscillations in reconstructed signals, while the soft threshold function can introduce distortion by reducing coefficients whose absolute values exceed the threshold. The improved threshold function in DAARSB mitigates these issues without introducing additional adjustment factors, enhancing robustness for noise reduction, and is incorporated as a nonlinear transformation layer in the deep learning model:

$$v = \begin{cases} \sqrt{u^2 - \tau^2} & |u| \geq \tau \\ 0 & |u| < \tau \end{cases} \tag{3}$$

Establishing a fixed threshold is challenging given the varying degrees of contamination in each signal. To address this issue, the DAARSB module employs a dual-attention framework that integrates channel and spatial attention mechanisms, enabling the model to adaptively set unique thresholds for different feature channels and spatial regions. Specifically, channel attention identifies and prioritizes feature channels that are critical for denoising and suppressing less important ones. Spatial attention focuses on identifying and preserving important spatial regions, enhancing the ability of the model to retain fine details.

In the DAARSB architecture in Fig. 6, an initial DW separable convolution condenses the input feature map, ensuring the learned thresholds are positive by taking absolute values. Subsequently, global maximum pooling (GMP) and global average pooling (GAP) extract finer-grained channel-level attention. Specifically, $A_c \in \mathbb{R}^{C\times 1}$ denotes the channel attention map, where C is the number of channels in the input feature map and $\mathbb{R}$ denotes the set of real numbers. A shared multi-layer perceptron with one hidden layer learns the scaling coefficients, which are scaled to (0, 1) using the sigmoid function. Finally, these coefficients are multiplied by the global maximum of each channel to obtain the channel-level thresholds. Similarly, GMP and GAP are applied to the feature map post-denoising to learn spatial-level attention. Specifically, $A_s \in \mathbb{R}^{1\times L}$ denotes the spatial attention map, where L represents the width of the feature map. The convolutional layers learn scaling coefficients, which are multiplied by the global average of each channel after being processed through a sigmoid function to derive the spatial-level threshold. By stacking multiple DAARSB modules, discernible features can be learned through nonlinear transformations and threshold shrinkage, effectively removing noise-related information.

2.2. UDA module

Traditional transfer learning methods typically rely on pre-trained weight loading and fine-tuning, which necessitate partially labeled target domain data for model adaptation. However, in routine MS monitoring, signal acquisition occurs continuously, and the collected data are entirely unlabeled, which significantly limits the applicability of traditional transfer learning techniques. In contrast, UDA methods do not require labeled target data, facilitating cross-

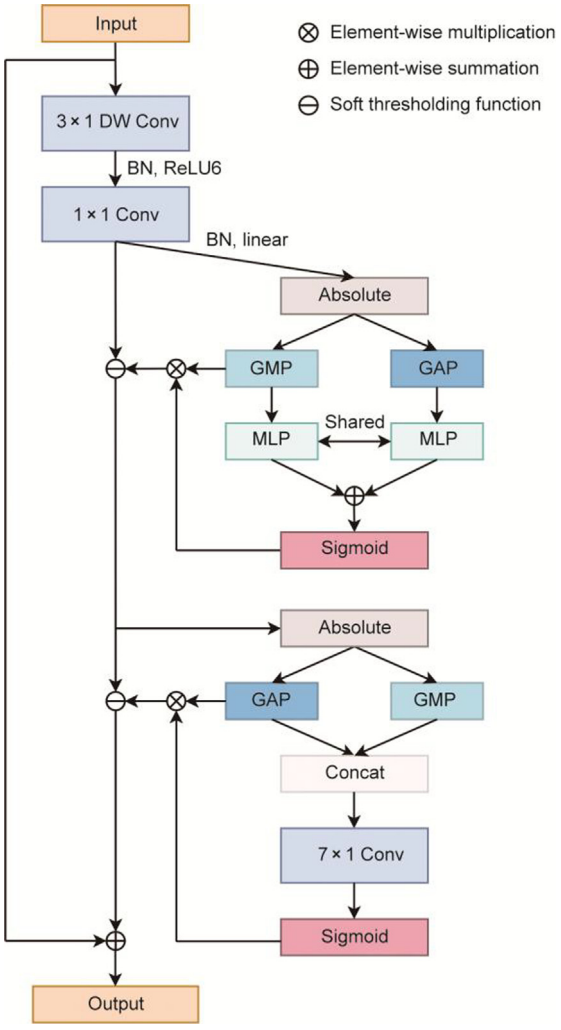

Fig. 6. Architecture of the DAARSB module.

domain knowledge transfer using only unlabeled signals. Existing UDA methods can be broadly categorized into discrepancy-based optimization and adversarial learning-based methods. Discrepancy-based optimization methods compute the distribution distance between the source and target domains using statistical moment-matching techniques, such as maximum mean discrepancy [34], and facilitate cross-DA by minimizing these discrepancy metrics. In contrast, adversarial learning-based UDA methods [35] introduce a domain discriminator to distinguish between source and target samples, while a generator is trained to produce feature representations that confuse the discriminator. However, these methods often overlook task-specific decision boundaries, focusing on the global alignment of feature distributions without adequately considering the unique characteristics of each domain. By contrast, the bi-classifier adversarial learning paradigm [36] adopted in this study enhances the target domain performance by integrating task-specific decision boundaries and intra-class alignment strategies.

As illustrated in Fig. 4, each classifier consists of two fully connected (FC) layers and a SoftMax layer. Let $p_1 \in R^{K\times 1}$ and $p_2 \in R^{K\times 1}$ be the SoftMax probability outputs of the two classifiers that satisfy

$$\sum_{k=1}^{K} p_l^k = 1,\ \text{s.t.}\ p_l^k \geq 0,\ \forall k = 1, \ldots, K;\ l = 1, 2 \tag{4}$$

where p_l^k denotes the kth element of the SoftMax probabilities of lth classifier (i.e., the classifier C_l classifies the sample into the kth class), and K is the total number of categories.

Adversarial learning with a bi-classifier paradigm formulates a two-player min–max game. In this framework, C_1 and C_2 maximize the discrepancy between their predictions on target samples, causing divergence that identifies target samples outside the support of the source domain. Conversely, the feature extractor F generates minimizes discrepancy, encouraging target samples to lie within the support of the source domain. Through iterative adversarial learning, our objective is to align the source and target feature distributions such that the support of the source domain encompasses that of the target domain is included by that of the source domain. The measurement of discrepancy loss between the two classifiers is crucial in this learning paradigm. Maximum classifier discrepancy (MCD) [36] uses the L_1 distance between predictions p_1 and p_2 (i.e., $|p_1 - p_2|_1$) to quantify disagreement. However, this metric only measures differences in predicted probabilities for corresponding class, ignoring correlations between classes. For instance, consider $p_1 = [0.33, 0.33, 0.34]$ and $p_2 = [0.33, 0.33, 0.34]$. Although the L_1 distance is zero, the predictions are still relatively uncertain across classes, exhibiting high entropy.

Therefore, a novel classifier discrepancy metric named classifier determinacy disparity (CDD) [37] was introduced to guide the UDA. The relevance matrix $\boldsymbol{M}$ of the bi-classifier prediction can be written as

$$\boldsymbol{M} = p_1 p_2^{\mathrm{T}} \tag{5}$$

where $M_{mn} = p_1^m p_2^n$ $(m, n \in \{1, 2, \ldots, K\})$ denotes the element in the mth row and nth column of a $K \times K$ square matrix $\boldsymbol{M}$. It represents the product of the probabilities that C_1 assigns the sample to the mth class and C_2 assigns it to the nth class. $\boldsymbol{M}$ captures the correlations of bi-classifier prediction across different categories. Maximizing the diagonal elements of matrix $\boldsymbol{M}$ implies minimizing the classifier discrepancy (i.e., making the two predictions consistent and correlated). Conversely, the non-diagonal elements of $\boldsymbol{M}$ matrix represent fine-grained confusion information between the two classifiers. Accordingly, the CDD can be defined as [37]:

$$\mathrm{CDD}(p_1, p_2) = \sum_{m,n=1}^{K} M_{mn} - \sum_{m=1}^{K} M_{mm} = \sum_{m\neq n}^{K} M_{mn} \tag{6}$$

The CDD encapsulates all probabilities where the predictions of C_1 and C_2 diverge. A minimum CDD value of zero is achieved only when the two probability distributions are identical and fully deterministic, such as $[0, 0, 1]$ and $[0, 0, 1]$.

2.3. Optimization

In this section, we present the three-step optimization process for LRE-UDAF.

Step A: For the UDA problem, it is essential that the model accurately classifies the source data. Therefore, the generator and classifiers are initially trained to minimize supervised loss on the source data. This step closely resembles training a conventional deep classification network, with the objective function defined as

$$\min_{\theta_f, \theta_{c1}, \theta_{c2}} \mathscr{L}_{\mathrm{cls}}(X_s, Y_s) = \frac{1}{2N_s} \sum_{i=1}^{N_s} \sum_{l=1}^{2} \mathscr{L}_{\mathrm{ce}}(p_{li}, y_i) \tag{7}$$

where θ_f, θ_{c1}, and θ_{c2} denote trainable parameters of F, C_1, and C_2, respectively, $\mathscr{L}_{\mathrm{ce}}$ is the cross-entropy loss, $p_{li} = C_l(F(x_i))$ represents the predicted probability vector output by the lth classifier for the ith source sample, and y_i is the true label.

Step B: Classifiers (C_1, C_2) are trained as discriminators for a fixed feature extractor F to maximize the discrepancy between their outputs on the target data. In addition, the

source-supervised loss is added to preserve source classification accuracy. The objective can be expressed as follows:

$$\min_{\theta_{c1},\theta_{c2}} \mathscr{L}_{\mathrm{cls}}(X_{\mathrm{s}}, Y_{\mathrm{s}}) - \mathscr{L}_{\mathrm{cdd}}(X_{\mathrm{t}}) \tag{8}$$

$$\mathscr{L}_{\mathrm{cdd}}(X_{\mathrm{t}}) = \frac{1}{N_{\mathrm{t}}} \sum_{i=1}^{N_{\mathrm{t}}} \mathrm{CDD}(p_{1i}, p_{2i}) \tag{9}$$

Step C: Finally, the feature extractor is trained to minimize the discrepancy between fixed classifiers. The objective is as follows:

$$\min_{\theta_{\mathrm{f}}} \mathscr{L}_{\mathrm{cdd}}(X_{\mathrm{t}}) \tag{10}$$

These three steps are performed iteratively until the maximum number of iterations is reached. The entire optimization procedure is illustrated in Fig. 7. In essence, the optimization centers around adversarial training, ensuring the generator and classifier accurately classify source samples.

Training adversarial networks present challenges, including convergence difficulties and frequent training oscillations. To address these challenges, we incorporate entropy regularization, which guides the classifier to produce confident and low-entropy predictions for unlabeled target data. Specifically, it penalizes uncertain or ambiguous predictions (high-entropy) and encourages the model to make low-entropy, high-confidence decisions. Thus, this regularization term effectively stabilizes training by reducing oscillations and mitigating overfitting on low-confidence target data, enhancing generalization to unseen target domain data. The training process for LRE-UDAF is summarized in Algorithm 1. The optimal values of the classifier discrepancy coefficient α and the entropy regularization coefficient β were determined via cross-validation on the validation set, yielding $\alpha = 0.01$ and $\beta = 0.1$.

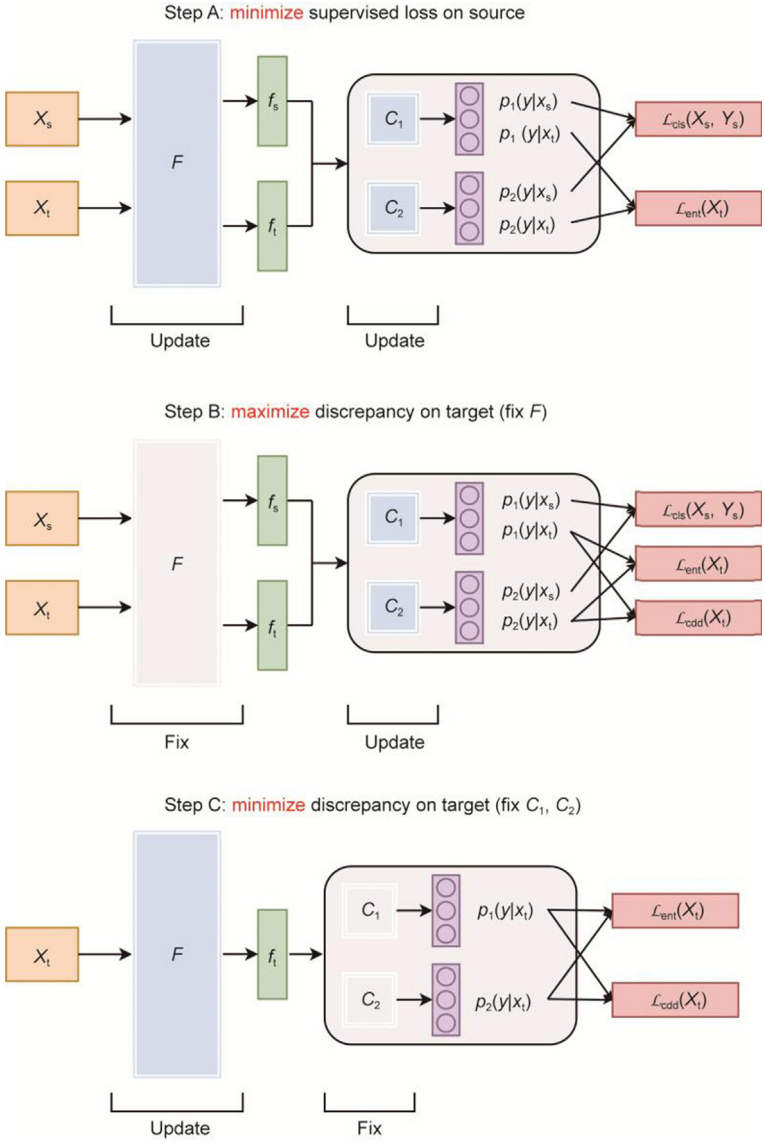

Fig. 7. LRE-UDAF optimization procedure. f_{s} and f_{t} denote the feature maps extracted by the shared feature extractor F from source domain inputs X_{s} and target domain inputs X_{t}, respectively. x_{s} and x_{t} denote the source domain sample and target domain sample, respectively; y denotes the corresponding label.

Algorithm 1. LRE-UDAF for cross-domain MS signal classification.

Input: Source signals $\{X_s, Y_s\} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$; target signals $\{X_t\} = \{x_j^t\}_{j=1}^{N_t}$; trade-off parameters α and β; batch size B; and maximum iteration

Output: Parameters θ_f, θ_{c1}, and θ_{c2} of feature extractor F and classifiers C_1, C_2

1. Initialize θ_f using pretrained parameters on source domain and randomly initialize θ_{c1}, θ_{c2}
2. **Repeat**
3. Sample mini-batch of B source signals and B target signals
4. Update θ_f, θ_{c1}, and θ_{c2} by minimizing supervised loss on source domain: $\min_{\theta_f,\theta_{c1},\theta_{c2}} \mathscr{L}_{cls}(X_s, Y_s) + \beta \mathscr{L}_{ent}(X_t)$
5. Update θ_{c1} and θ_{c2} by maximizing CDD on target domain: $\min_{\theta_{c1},\theta_{c2}} \mathscr{L}_{cls}(X_s, Y_s) - \alpha \mathscr{L}_{cdd}(X_t) + \beta \mathscr{L}_{ent}(X_t)$
6. Update θ_f by minimizing CDD on target domain: $\min_{\theta_f} \alpha \mathscr{L}_{cdd}(X_t) + \beta \mathscr{L}_{ent}(X_t)$
7. **Until** maximum iteration is reached

2.4. Proposed LRE-UDAF workflow

The workflow of LRE-UDAF is presented in Fig. 8 and comprises three concise steps:

Step one is data acquisition and preprocessing. The MS signals are collected from multiple projects, including one with substantial labeled signals (source domain) and others with limited unlabeled signals (target domain). All signals undergo maximum absolute scaling normalization to mitigate amplitude interference. The acquired signals are then divided into training, validation, and test sets for subsequent model training and evaluation.

Step two is model design and training. The LRE-UDAF consists of the lightweight ISNU module, noise-robust DAARSB module, and a dual-classifier DA module. Table 2 details the parameters for each module. The training occurs in two stages. Initially, a standard classification task is conducted on source-domain signals to establish foundational knowledge. Subsequently, the DA phase begins with the feature extractor, retaining the pretrained parameters while initializing the two classifiers randomly. It is crucial that the source data used in the second phase do not overlap with those used in the first phase to prevent performance degradation due to data leakage.

Step three involves performance testing and comparison. The target domain testing set is input into the trained model for MS signal classification, with results visualized from multiple perspectives.

3. Source classification experiments

A classification experiment was conducted in the source domain to comprehensively evaluate the performance of the feature extractor. Several industry-leading models were selected for comparative analysis, noted for their effectiveness in computer vision (CV), natural language processing, and mobile applications.

The primary objective of MS monitoring is to assess the stability of the surrounding rock and provide early warnings of potential deep underground disasters. Therefore, all signals are categorized into three types: blast signals (for tracking strong dynamic disturbances), MS signals (for assessing the conditions of surrounding rock fractures), and noise signals (other irrelevant signals). The source-domain classification task used 30 000 labeled single-channel acceleration waveforms collected by an MS monitoring system in a tunnel in southwest China. Each class contained 10 000 signals, each with 4000 sampling points. To address the limited number of blast signals, data enhancement methods, such as left–right panning and up–down flipping, were applied to balance the dataset. The training, validation, and test sets were split at an 8:1:1 ratio, with data labeling verified by experienced experts to reduce noise. The training hyperparameters for LRE-UDAF were determined through empirical knowledge and systematic cross-validation. The initial parameters were selected based on prior knowledge and relevant literature [28–31], followed by iterative fine-tuning according to validation set performance. Key training parameters are presented in Table 3. The model was optimized using the adaptive moment estimation (Adam) optimizer, a popular algorithm that computes adaptive learning rates for each parameter by estimating first- and second-order moments of the gradients. The "poly" strategy is utilized for learning rate adjustment [38]. If validation error did not decrease over 20 consecutive epochs, training was terminated. All the experiments were conducted on a computer equipped with an i5-12490F CPU (Intel Corporation, USA), an NVIDIA GeForce RTX 3070 Ti GPU (NVIDIA Corporation, USA), and 32 GB of random access memory (RAM).

3.1. Classification result analysis

To ensure the reliability and reproducibility of the experimental results, a random seed was fixed, and the experiment was repeated five times using the same training and validation sets. For each run, the final model obtained through early stopping was used to generate predictions on the test set. The test set performance was then averaged over five repetitions and reported as the final evaluation outcome. Multiple evaluation metrics, namely accuracy, precision, and recall, were used to evaluate the model performance [39]:

$$\text{Accuracy}_c = \frac{\text{TP}_c + \text{TN}_c}{\text{TP}_c + \text{FP}_c + \text{TN}_c + \text{FN}_c} \tag{11}$$

$$\text{Precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c} \tag{12}$$

$$\text{Recall}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c} \tag{13}$$

where c represents the cth class signal. TP (true positive), FP (false positive), TN (true negative), and FN (false negative) denote the correct identification of positive samples, misclassification of negative samples as positive, correct identification of negative samples, and misclassification of positive samples as negative, respectively. TP + FP is the number of samples predicted as positive, while TP + FN is the number of all positive samples in the reference label. Accuracy measures the overall classification correctness of the model (i.e., the ability to correctly classify all samples). Precision measures the ability of the model to correctly predict positive samples (i.e., the proportion of predicted positive samples that are actually positive). Recall measures the ability of the model to identify all positive samples (i.e., the proportion of actual positive samples correctly predicted by the model). These three indices complement one another and reflect the classification ability of network from different perspectives.

Fig. 9 illustrates the confusion matrix and corresponding evaluation indices for the source-domain classification results. The colorbar adjacent to the confusion matrix in step three indicates that darker shades correspond to a higher number of samples. The test set contains 1000 blast, 1000 MS, and 1000 noise signals, with precisions of 100%, 96.9%, and 96.2%, and recalls of 100%, 96.2%, and 96.9%, respectively. This indicates that all blast signals are correctly identified without misclassification, a critical factor

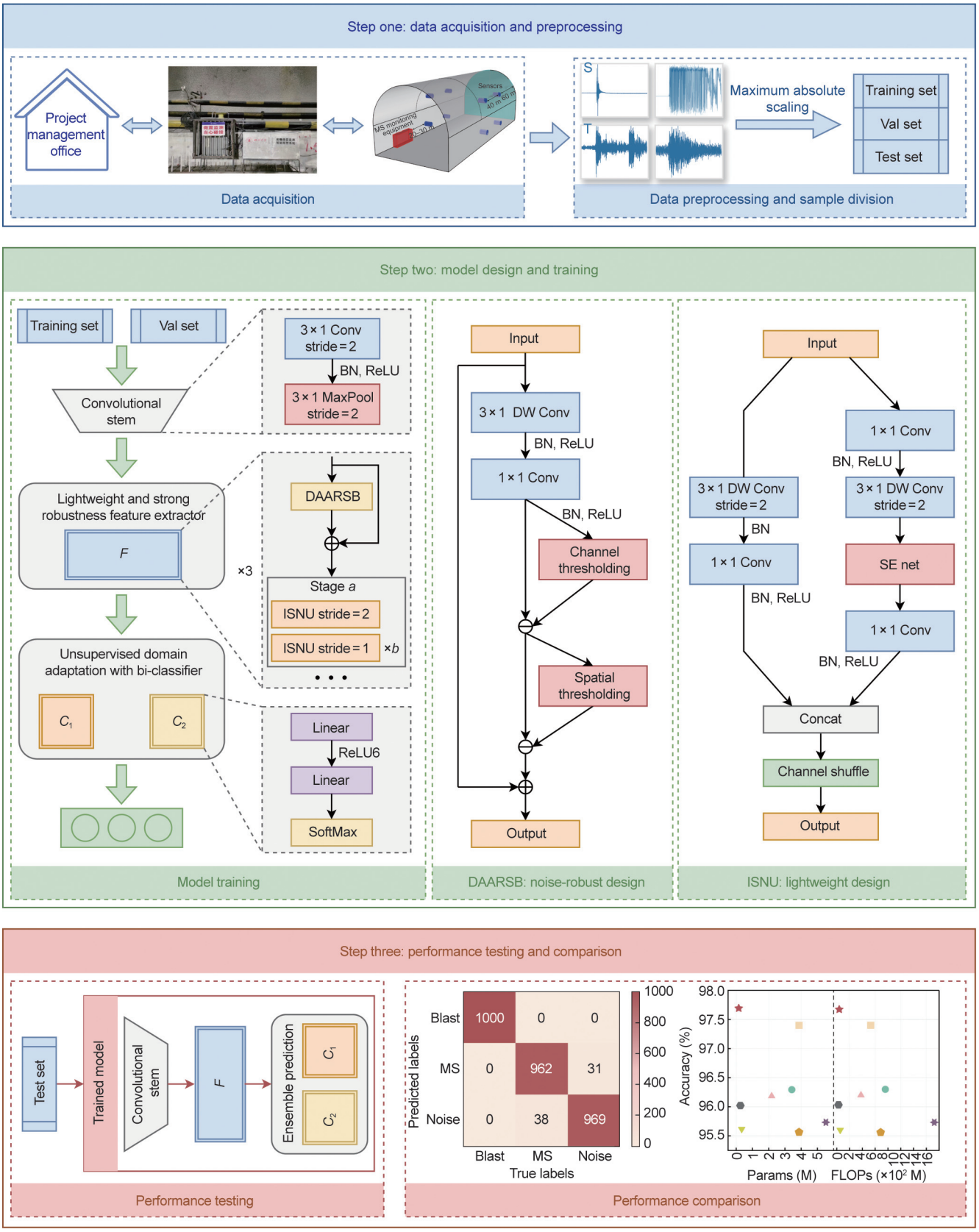

Fig. 8. Workflow of LRE-UDAF. "S" denotes the source domain; "T" denotes the target domain; "Val" denotes the validation set; "Params" represents the number of parameters; "M" denotes million. The darker shades of color indicate a higher number of samples in the confusion matrix of step three.

Table 2
Parameter configuration of LRE-UDAF.

Module	Sub-module	Layer	Output size	Stride	Repeat	Output channels
Feature extractor	Convolutional stem	Conv	2000 × 1	2	1	16
		MaxPool	1000 × 1	2	1	16
	DAARSB	—	1000 × 1	1	1	16
	Stage 1	ISNU	500 × 1	2	1	24
		ISNU	500 × 1	1	1	24
	DAARSB					
	—	500 × 1	1	1	24	
	Stage 2	ISNU	250 × 1	2	1	40
		ISNU	250 × 1	1	3	40
	DAARSB	—	250 × 1	1	1	40
	Stage 3	ISNU	125 × 1	2	1	48
		ISNU	125 × 1	1	1	48
	ASPP	—	125 × 1	1	1	48
	DAARSB	—	125 × 1	1	1	48
	Adaptive AvgPool	—	1 × 1	—	—	48
Classifier 1/2	—	FC1	—	—	—	80
	—	FC2	—	—	—	3

The input to the model is a one-dimensional signal of length 4000.

Table 3
Training parameter settings.

Parameter	Value
Batch size	16
Epochs	300
Learning rate	6×10^{-3}
Loss function	Cross-entropy (label smoothing = 0.1)
Optimizer	Adam

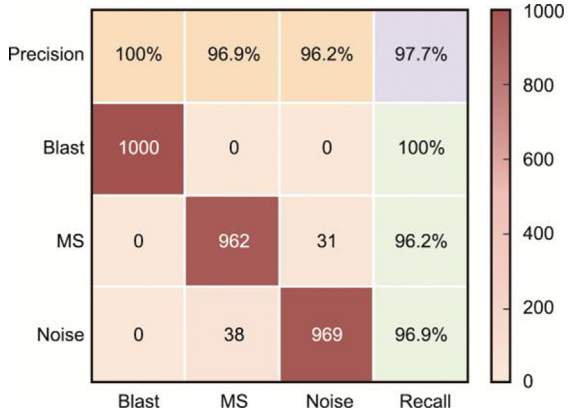

Fig. 9. Confusion matrix for the test set on source domain. The legend on the right illustrates the distribution of sample quantities.

since blast signals exhibit significantly higher energy and amplitude than MS and noise signals. Misclassifying these signals could severely compromise warning level assessments. A minor degree of confusion is observed between the MS and noise signals due to their waveform similarities. Even experienced experts may find it challenging to accurately distinguish them when relying solely on single-channel waveforms rather than analyzing the entire MS event. Moreover, the misclassified MS and noise signals share low amplitude and energy levels, making misclassification less impactful on warning level determination. The overall accuracy achieved on the test set is 97.7%, and when excluding labeling errors, the prediction accuracy significantly exceeds industry application standards. This finding underscores the robustness of the proposed feature extractor architecture in distinguishing among the three signal types.

3.2. Lightweight performance analysis

Seven leading models across diverse fields were selected for lightweight performance comparison. These include CNN architectures that are widely recognized and emerging in the CV field: GoogLeNet [30], RepViT [40], and StarNet [41]; lightweight models specifically designed for mobile applications: EfficientNet [42], MobileNet-V2 [43], and ShuffleNet-V2 [31]; along with Convformer-NSE, a hybrid CNN-transformer model for bearing fault diagnosis [44] was selected. Given that CV typically employs three-dimensional image inputs (channel, length, and width), whereas MS signals utilize two-dimensional inputs (channel and length), the appropriate CV model layers were modified to accommodate a two-dimensional input. Additionally, owing to the smaller size of the MS signal classification dataset compared to those in CV, the smallest versions of the CV model were used to ensure that the model capacity was appropriate for the dataset size. Five repeated experiments were performed for each model with parameters identical to those specified in our method to maintain consistent comparison conditions. The initial learning rate, which significantly affects the results, was determined using a grid search method based on validation set accuracy.

During training, validation loss and accuracy for each model were recorded over the first 100 epochs. As shown in Fig. 10(a), all models exhibit a rapid decline in validation loss during early training, stabilizing after approximately 40 epochs. Notably, LRE-UDAF consistently maintains the lowest validation loss throughout training. In Fig. 10(b), the validation accuracy trends align closely with those of validation loss. Among all models, EfficientNet demonstrates performance most comparable to the proposed method.

The lightweight performance comparison emphasizes three key aspects: model size, computational complexity, and prediction accuracy. The number of parameters (Params) directly determines the model size, representing the total number of trainable parameters (typically weights and biases) in the model. The FLOPs measure computational complexity, indicating the number of floating-point operations required for a single inference. Fig. 11 illustrates the relationship between Params, FLOPs, and accuracy for these models. The LRE-UDAF consistently appears in the top-left corner, signifying the fewest parameters, lowest computational complexity, and highest classification accuracy. Table 4 presents the detailed numerical values, showing that LRE-UDAF has the smallest Params and the highest accuracy. Although FLOPs for LRE-UDAF are slightly higher than those of Convformer-NSE, it maintains a comparative advantage overall. Although EfficientNet

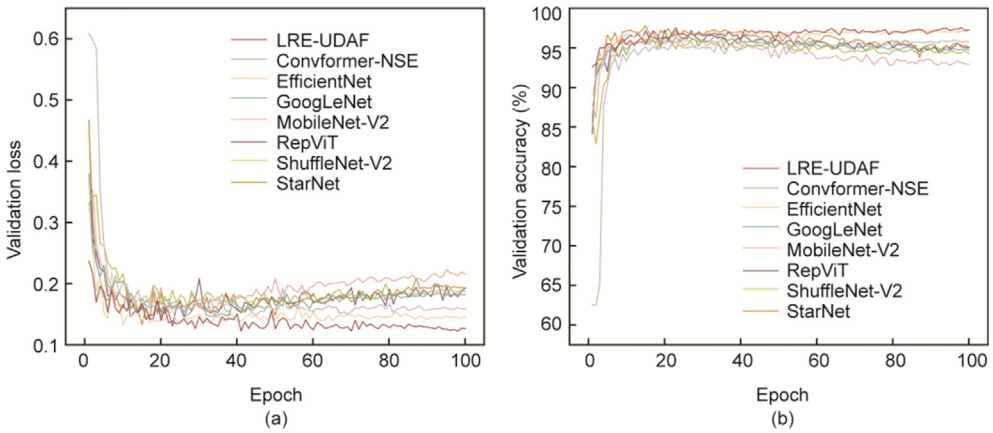

Fig. 10. (a) Validation loss and (b) accuracy for each model.

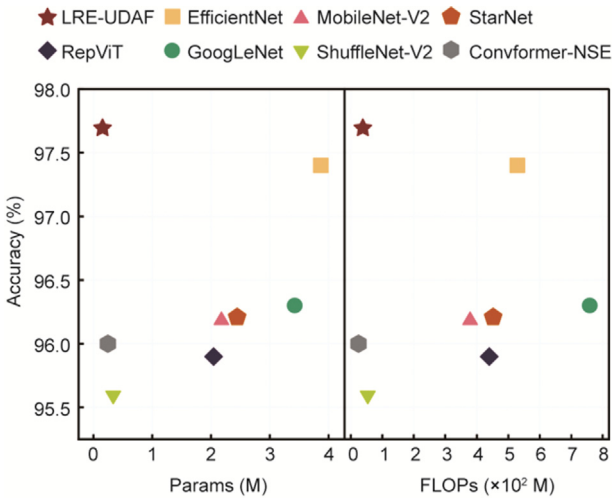

Fig. 11. Comparison of accuracy and model complexity. Params and FLOPs are reported in millions.

achieves an accuracy closest to that of our model, its Params and FLOPs are significantly higher. Fig. 10 shows that LRE-UDAF achieves relatively low loss and high accuracy in the first epoch, likely due to its relatively fewer trainable parameters, allowing for satisfactory performance after one training cycle. Although Convformer-NSE has a smaller number of parameters, its transformer component lacks the inherent inductive bias of CNNs.

3.3. Noise robustness analysis

To further investigate the sensitivity of the proposed method to noise, we synthesized MS signals with varying SNRs and compared their classification accuracy across different models. Noisy MS signals are generated by combining clean MS waveforms (with SNR above 40) acquired in the field with different types and amplitudes of noise waveforms. Two distinct noise types were selected for this analysis: Gaussian noise, simulating random disturbances, and field-acquired noise, representing common environmental interference in practical scenarios. By scaling their magnitudes to different degrees, we generated MS signals with various SNRs, as shown in Fig. 12. The SNR is calculated as follows:

$$\mathrm{SNR} = 20 \times \log_{10}(S_{\max}/N_{\max}) \tag{14}$$

where $S_{\max}$ and $N_{\max}$ denote the peak amplitudes of the MS and noise signals, respectively. We synthesized 100 noisy MS signals in six intervals: [10, 15), [15, 20), [20, 25), [25, 30), [30, 35), and [35, $+\infty$), recording the classification accuracy of models across these intervals.

Fig. 13 shows the classification accuracy of the different models across various SNR intervals, with LRE-UDAF achieving the highest accuracy throughout. When the SNR of the noisy MS signal exceeds 20, all the models achieve acceptable classification accuracy. However, within the SNR intervals of [10, 15) and [15, 20), our model significantly outperforms others. This observation suggests that while CNNs have inherent noise robustness, incorporating specialized noise-robust architectural designs can further enhance noise resistance. Table 4 lists the specific classification accuracy, further indicating that LRE-UDAF achieves 85% classification accuracy in the [10, 15) interval, compared to only 3% for StarNet. This trend continues in the [15, 20) interval. In the [20, $+\infty$) interval, most methods achieve nearly 100% accuracy. Analyses of field-acquired noisy MS signals reveal that most fall within the [20, $+\infty$) interval, with only a few heavily noise-corrupted signals in [10, 20) interval, and virtually none under 10. This indicates that

Table 4
Comparison of lightweight design and noise robustness for each model.

Method	Lightweight design			Strong noise robustness design					
	Params (M)	FLOPs (M)	Accuracy (%)	[10, 15)	[15, 20)	[20, 25)	[25, 30)	[30, 35)	[35, $+\infty$)
LRE-UDAF	**0.155**	38.237	**97.7**	**85%**	**94%**	**99%**	**100%**	**100%**	**100%**
Convformer-NSE	0.245	**24.059**	96	48%	91%	97%	100%	100%	100%
EfficientNet	3.868	529.144	97.4	62%	87%	98%	100%	100%	100%
GoogLeNet	3.424	758.340	96.3	31%	88%	97%	100%	100%	100%
MobileNet-V2	2.184	378.493	96.2	19%	91%	96%	99%	99%	100%
RepViT	2.040	439.422	95.9	8%	78%	97%	100%	99%	100%
ShuffleNet-V2	0.338	53.615	95.6	5%	51%	90%	100%	100%	100%
StarNet	2.401	451.817	96.2	3%	43%	96%	100%	99%	100%

Bold values indicate the optimal performance for each metric.

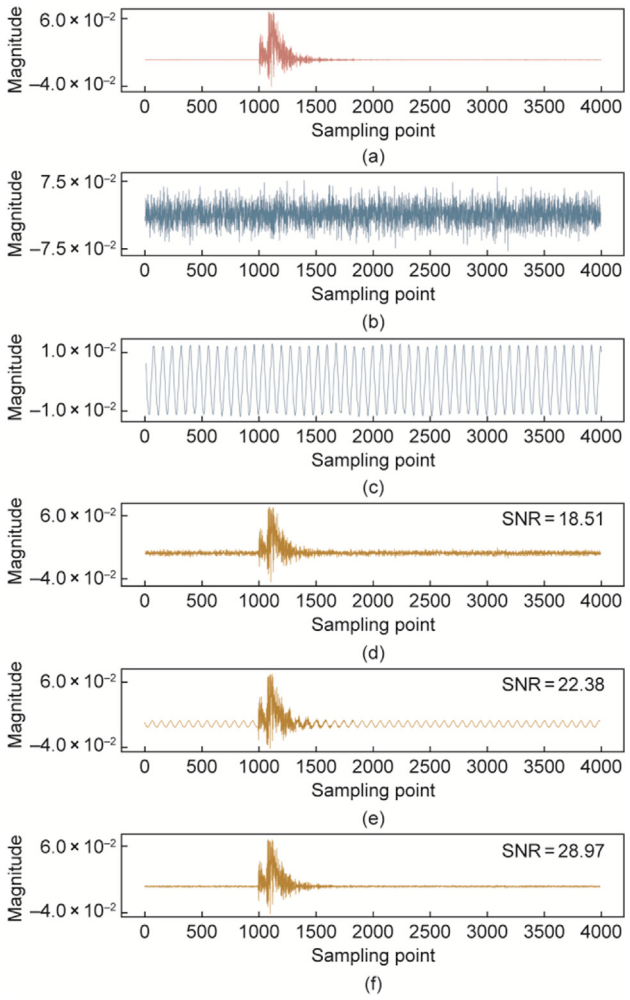

Fig. 12. Synthesis process of noisy MS waveforms. (a) Clean MS waveform, (b) Gaussian noise waveform, (c) field noise waveform, and (d–f) synthesized noisy MS waveforms with SNRs of 18.51, 22.38, and 28.97, respectively.

the proposed model is adequately robust for recognizing field-acquired noisy MS signals.

4. Cross-domain classification experiments

To evaluate the effectiveness of the UDA framework, we selected two target domains with distinct data distributions for cross-domain MS signal classification. The primary differences between the source and target domains are based on monitoring projects and equipment, as summarized in Table 5. The first target domain (referred to as DT2) originates from the same engineering project as the source domain (DT1), namely the DT project, a deep tunnel under construction in southwest China. The key distinction between DT1 and DT2 lies in the monitoring equipment used; DT1 uses the Chinese SINOSEISM MS monitoring system (Hubei Seaquake Technology Co., Ltd., China), whereas DT2 employs the Canadian Engineering Seismology Group (ESG) MS monitoring system. These systems differ in sampling frequency and signal length. The second target domain (deep hydropower, DHP) represents a different engineering context: the underground powerhouse of a large hydropower station in southwest China. Additionally, this domain uses the ESG system but uses a different sensor deployment method. While the DT project deploys sensors behind the tunnel face in a non-enclosed arrangement, the DHP project utilizes a layered, enclosed configuration, with the excavation face located within the three-dimensional space formed by the sensor array.

Given that the transfer tasks from DT1 to DT2 and from DT1 to DHP are the most representative owing to significant differences in equipment, signal characteristics, and engineering backgrounds, we provide a detailed analysis of these two transitions in Sections 4.1 and 4.2, respectively. These cases highlight the ability of the framework to handle intra- and interproject domain shifts, offering valuable insights into its performance under realistic engineering conditions. In contrast, the remaining four transfer directions (DT2→DT1, DT2→DHP, DHP→DT1, and DHP→DT2) are briefly summarized in Section 4.3 to validate the generalization capability of the framework. Furthermore, the UDA training setup slightly adjusts from the source-domain classification task, using an initial learning rate of 5×10^{-6} and the Adam optimizer.

4.1. Cross-domain transfer: DT1→DT2

Fig. 14(a) shows the equipment layout and typical signals collected at engineering site of DT2, while Fig. 14(b) presents the equipment layout and signal types from the DHP project. In both cases, the collected signals are downsampled to one-quarter of their original length, yielding 3750 sampling points per signal for DT2 and 3500 sampling points per signal for the DHP project. Table 6 presents evaluation indices before and after DA. The F_1-score, a harmonic mean of precision and recall, serves as a comprehensive evaluation metric. In MS signal classification, precision and recall are critical, requiring that positive samples be identified comprehensively while ensuring the correctness of these predictions. Therefore, a higher F_1-score indicates improved classification performance. Before DA, the classification results reveal imbalances, with significant discrepancies between precision and recall, resulting in a lower F_1-score and suboptimal classification performance. After applying DA, the classification results show a greater and significant F_1-score improvement. Although confusion between the MS and noise signals remains the primary classification error source, overall classification performance is markedly enhanced compared to pre-DA results. Furthermore, to visualize the positive effects of DA, we employed the *t*-distributed stochastic neighbor embedding (*t*-SNE) [45] method to visualize feature representations extracted by the feature extractor. Fig. 15 shows that DA clusters different signal types within the feature space, enhancing the discriminative power of previously mixed features. The overall classification accuracy improves from 77.7% to 94.3%, reflecting a significant 16.6% increase, thus satisfying on-site automatic classification requirements.

4.2. Cross-domain transfer: DT1→DHP

Table 7 outlines the classification results before and after DA for the DHP project. Post DA, the F_1-score for each signal type shows significant improvement, reducing confusion between MS and noise signals. Fig. 16 depicts the visual feature representation for signals in Case 2. The classification accuracy increases from 87.6% to 97.3%, a significant increase of 9.7%. Notably, in the feature space, blast signal representations are distinctly separated from the MS and noise signals, demonstrating effective discrimination. By contrast, the feature representations of the MS and noise signals exhibit some overlap, aligning with the classification results. Nevertheless, the significant improvement in the classification performance underscores the effectiveness of the proposed UDA framework in improving feature separability and facilitating robust cross-DA.

4.3. Other cross-domain transfer scenarios

To further validate the generalization capability of the proposed method across different transfer directions, additional experiments

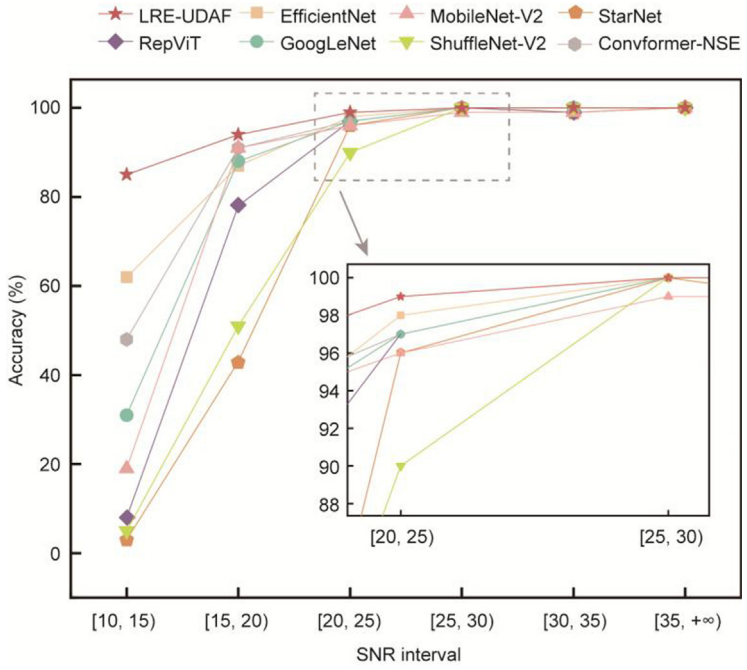

Fig. 13. Accuracy comparison of different models under varying SNRs.

Table 5
Comparison of profiles between different projects.

Project name	Project type	Equipment name	Layout	Sampling frequency (kHz)	Signal length
DT1	Tunnel	SINOSEISM	Non-contained	4	4 000
DT2	Tunnel	ESG	Non-contained	10	15 000
DHP	Underground cavern	ESG	Contained	10	14 000

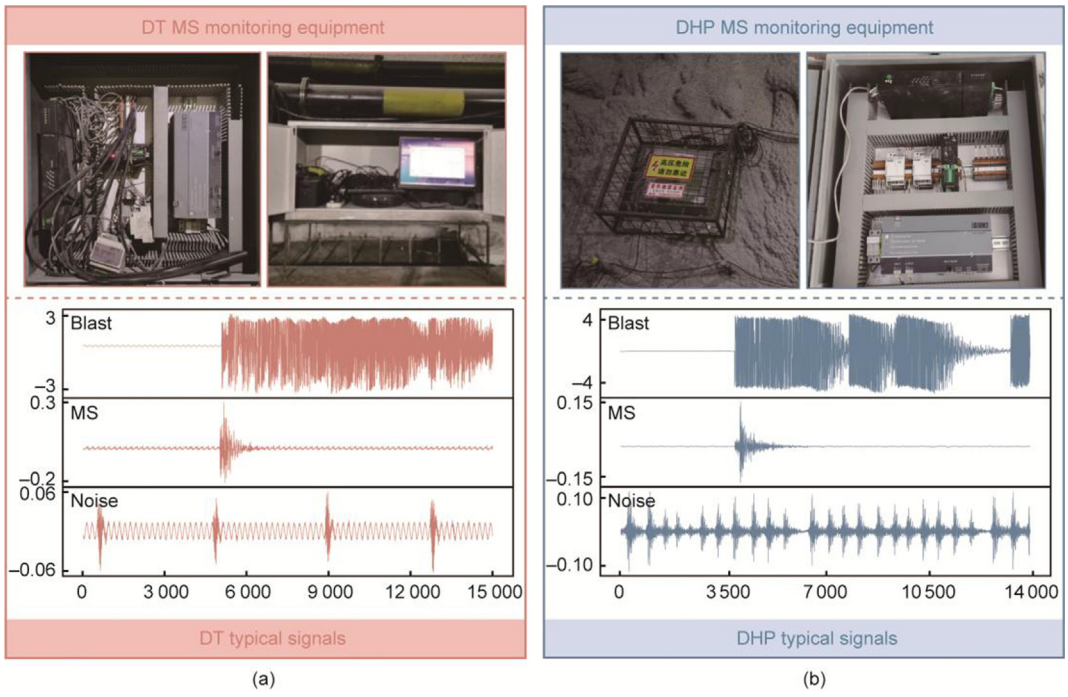

Fig. 14. Equipment and signal comparison between the (a) DT and (b) DHP projects.

Table 6
Evaluation indicators before and after DA in DT project.

Signal	Conditions	Precision (%)	Recall (%)	F_1
Blast	Before	**96.3**	86.2	0.910
	After	95.1	**100.0**	**0.975**
MS	Before	**99.6**	50.4	0.669
	After	98.2	**89.6**	**0.937**
Noise	Before	60.4	**96.5**	0.743
	After	**90.0**	93.2	**0.916**

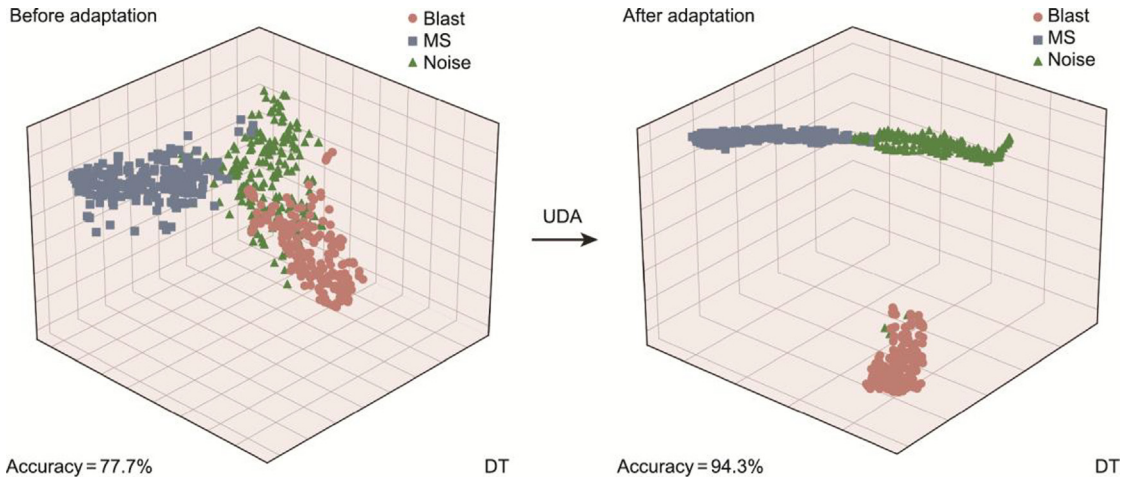

Fig. 15. Feature visualization before and after DA in the DT project.

Table 7
Evaluation indicators before and after DA in the DHP project.

Signal	Conditions	Precision (%)	Recall (%)	F_1
Blast	Before	**99.9**	85.2	0.920
	After	99.0	**100.0**	**0.995**
MS	Before	**97.4**	80.0	0.878
	After	96.5	**96.3**	**0.964**
Noise	Before	73.8	**97.8**	0.841
	After	**96.3**	95.5	**0.959**

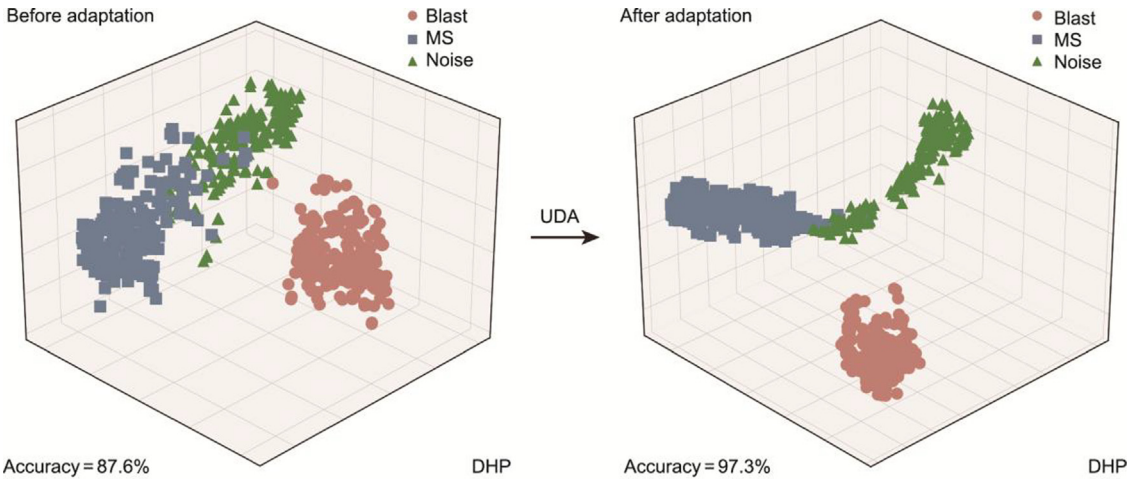

Fig. 16. Feature visualization before and after DA in the DHP project.

were conducted in four transfer scenarios: DT2→DT1, DT2→DHP, DHP→DT1, and DHP→DT2. The experimental results are presented in Table 8. Before DA, the initial classification accuracies across the tasks vary substantially, indicating significant data distribution shifts between the source and target domains. After applying the proposed UDA method, the classification performance improves substantially across all transfer directions. The accuracy for the DT2→DT1 transfer increases from 75.7% to 92.7%, showing the

Table 8
Comparison of classification accuracy before and after DA across other transfer directions.

Transfer scenarios	Before DA	After DA
DT2→DT1	75.7%	92.7%
DT2→DHP	93.1%	98.2%
DHP→DT1	84.0%	95.8%
DHP→DT2	86.4%	96.4%

largest improvement. This demonstrates the adaptability of the model across cross-domain tasks involving different monitoring devices and significant signal length variations. Moreover, as a signal source from a complex underground engineering environment, DHP demonstrates high adaptation stability when transferred to the DT domains, with accuracies exceeding 9%. These results confirm the robustness and effectiveness of LRE-UDAF across multiple transfer directions.

5. Discussion

5.1. Ablation study

To validate the design effectiveness of the feature extractor in the LRE-UDAF framework, we conducted a systematic ablation study by removing or replacing its core components: the ISNU and DAARSB modules. The results summarized in Table 9 highlight their indispensable role in achieving optimal performance.

The removal of the ISNU module reduces classification accuracy, from 97.7% to 93.5%, and affecting noise robustness, particularly in low-SNR scenarios (e.g., the accuracy drops by 63% in the [10, 15) SNR interval). Although the reduction in model size (Params and FLOPs) appears advantageous, this trade-off highlights the critical role of ISNU in balancing computational efficiency with high representational ability. Similarly, removing the DAARSB module results in a 2.3% reduction in accuracy and a substantial decline in noise robustness, emphasizing its importance in adaptive noise suppression. Replacing the ISNU module with ShuffleNet-V2 blocks (lacking SE blocks) slightly reduces the model size and decreases accuracy and denoising performance, illustrating the potential of the SE module to enhance performance. Furthermore, replacing the DAARSB module with the DRSN module increases the model size while reducing classification accuracy and noise robustness, demonstrating the effectiveness of the DAARSB module.

Collectively, these results confirmed that the ISNU and DAARSB modules are uniquely tailored to address challenges in MS signal classification. Integrating these modules within the LRE-UDAF framework yields the highest classification accuracy and superior noise robustness, demonstrating their synergistic effect on balancing computational efficiency with overall performance.

5.2. UDA technique comparisons

To validate the superiority of the proposed LRE-UDAF framework, we compared it systematically with other popular UDA methods (namely margin disparity discrepancy (MDD) [46] and minimum class confusion (MCC) [47]) and ablation variants of LRE-UDAF.

As quantitatively summarized in Table 10, LRE-UDAF achieves the highest cross-domain classification accuracy in DT1→DT2 (Case 1, 94.3%) and DT1→DHP (Case 2, 97.3%), outperforming all benchmarks. Ablation experiments reveal that removing entropy regularization causes accuracy reductions of 2.5% (Case 1) and 1.6% (Case 2), emphasizing its role in stabilizing the training dynamics and reducing the prediction uncertainty for unlabeled target data. Furthermore, substituting the CDD metric with MCD reduces classification accuracy from 94.3% to 89.8% in Case 1 and from 97.3% to 94.9% in Case 2. This decline underscores the superiority of CDD in capturing fine-grained interclass confusion patterns compared to MCD, which relies on simplistic L_1 distance metrics between classifier outputs. Other SOTA UDA methods also exhibit varying degrees of performance degradation. Notably, MDD, which introduces margin disparity discrepancy, a theoretically grounded measure of domain shift derived from margin-based generalization bounds, achieves 85.0% (Case 1) and 94.7% (Case 2). The MCC, designed to minimize class confusion by encouraging well-calibrated and confident predictions in the target domain, performs better, yielding 89.5% (Case 1) and 96.7% (Case 2). These results highlight the limitations of MDD and MCC in fully addressing complex task-specific class-level discrepancies. By contrast, the LRE-UDAF framework synergistically integrates bi-classifier adversarial learning (via CDD) and entropy regularization to achieve precise domain alignment and robust classification.

5.3. Discussion on future directions

Although the proposed LRE-UDAF framework demonstrates significant advantages in cross-domain MS signal classification tasks, it exhibits several limitations that warrant further discussion and future research.

(1) Hyperparameter sensitivity. The UDA module exhibits high sensitivity to hyperparameters, particularly the classifier discrepancy coefficient α and entropy regularization β. These parameters require extensive manual tuning through grid search and cross-validation, which is time-consuming and may lead to suboptimal solutions that lack adaptability under varying domain conditions. Future research will focus on integrating meta-learning algorithms [48] or adaptive weighting mechanisms [49] to adjust these hyperparameters dynamically during training. Such approaches can eliminate manual tuning and ensure that the model adapts more effectively to diverse domain shifts.

(2) Single-channel dependency. The current design of the LRE-UDAF framework is optimized for single-channel waveforms, fulfilling lightweight and robust design requirements. However, this focus limits its ability to capture multidimensional features in complex scenarios, potentially impairing performance in tasks requiring rich information from multichannel signals. Future work will explore extending the framework to support multichannel inputs, enabling the model to leverage richer spatial and temporal

Table 9
Ablation study on the feature extractor module.

Method	Accuracy (%)	Params (M)	FLOPs (M)	Accuracy of noise robustness design (%)	
				[10, 15)	[15, 20)
Feature extractor of LRE-UDAF	**97.7**	**0.155**	**38.237**	**85**	**94**
Without ISNU	93.5 (−4.2)	0.127	23.955	22 (−63)	66 (−28)
Without DAARSB	95.4 (−2.3)	0.105	30.529	31 (−54)	61 (−33)
Replace ISNU	97.3 (−0.4)	0.150	38.227	81 (−4)	93 (−1)
Replace DAARSB	97.5 (−0.2)	0.201	62.364	83 (−2)	89 (−5)

Table 10
Performance comparison of different UDA techniques.

Method	DT1→DT2	DT1→DHP
LRE-UDAF	**94.3%**	**97.3%**
Without entropy regularization	91.8% (−2.5%)	95.7% (−1.6%)
Replace CDD with MCD metric	89.8% (−4.5%)	94.9% (−2.4%)
MDD	85.0% (−9.3%)	94.7% (−2.6%)
MCC	89.5% (−4.8%)	96.7% (−0.6%)

correlations. A key challenge in processing multichannel signals is the potential presence of faulty sensors within an MS event (e.g., an eight-channel event may include "bad channels" dominated by noise due to sensor malfunction). To ensure a robust overall judgment of MS events, attention mechanisms [50] can be incorporated to effectively weigh the contributions of different channels, focusing on reliable signals while mitigating the impact of noisy or faulty channels.

(3) Interpretability of domain alignment. Although *t*-SNE visualization qualitatively demonstrates effective feature alignment between source and target domains, the interpretability of domain-invariant features remains insufficiently explored. This limits the transparency and acceptability of the model in practical engineering applications. Future research will explore techniques such as saliency maps [51], attention visualization, or explainable artificial intelligence (XAI) tools [52] that offer deeper insights into the decision-making process, enhancing engineers' comprehension of how the model identifies critical patterns and aligns features across domains, thereby improving its credibility and usability in real-world scenarios.

6. Conclusions

To address the engineering deployment challenges of MS signal classification, this study proposes a novel UDA framework called LRE-UDAF, which integrates two core components: a lightweight and robust feature extractor and a bi-classifier adversarial UDA module. The feature extractor achieves computational efficiency and adaptive denoising by integrating ISNU and DAARSB modules, effectively suppressing noise interference while preserving discriminative features and balancing model complexity with performance. The UDA module employs a bi-classifier adversarial learning paradigm enhanced by the CDD metric and entropy regularization. Through three-stage adversarial training, the source and target domain distributions are progressively aligned, ensuring robust cross-domain generalization.

The performance of LRE-UDAF was comprehensively evaluated across multiple dimensions. The experimental results in source-domain classification show that, compared to other SOTA models, the feature extractor yields superior classification accuracy, model size, and computational complexity. Moreover, in the noise signal recognition tests, LRE-UDAF achieved the highest accuracy across each SNR interval. Cross-domain classification experiments reveal that UDA significantly enhances feature representation, as well as the overall classification accuracy. A systematic ablation study and comparative analysis further validate the effectiveness of each component.

CRediT authorship contribution statement

Dingran Song: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation. **Feng Dai:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Yi Liu:** Writing – review & editing, Supervision, Funding acquisition. **Hao Tan:** Writing – review & editing, Visualization, Supervision. **Mingdong Wei:** Writing – review & editing, Visualization, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful for the financial support from the National Natural Science Foundation of China (52225904, 52039007, and 42377144) and the Natural Science Foundation of Sichuan Province (2023NSFSC0377). This work was supported by the New Cornerstone Science Foundation through the XPLORER PRIZE.

References

[1] Li A, Liu Y, Dai F, Liu K, Wang K. Deformation mechanisms of sidewall in layered rock strata dipping steeply against the inner space of large underground powerhouse cavern. Tunn Undergr Space Technol 2022;120:104305.
[2] Hu J, He M, Li H, Tao Z, Liu D, Cheng T, et al. Rockburst hazard control using the excavation compensation method (ECM): a case study in the Qinling Water Conveyance Tunnel. Engineering 2024;34(3):154–63.
[3] Song D, Liu Y, Dai F, Jiang R, Li A. Quantitative prediction of surrounding rock deformation via an energy-based damage model combining with microseismic monitoring. Tunn Undergr Space Technol 2024;147:105711.
[4] Jiang R, Dai F, Liu Y, Li A. Fast marching method for microseismic source location in cavern-containing rockmass: performance analysis and engineering application. Engineering 2021;7(7):1023–34.
[5] Zhao Y, Du Y, Yan Q. Challenges, progress, and prospects of ultra-long deep tunnels in the extremely complex environment of the Qinghai–Xizang Plateau. Engineering 2025;44(1):162–83.
[6] Feng X, Liu J, Chen B, Xiao Y, Feng G, Zhang F. Monitoring, warning, and control of rockburst in deep metal mines. Engineering 2017;3(4):538–45.
[7] Yin X, Liu Q, Huang X, Pan Y. Real-time prediction of rockburst intensity using an integrated CNN-Adam-Bo algorithm based on microseismic data and its engineering application. Tunn Undergr Space Technol 2021;117:104133.
[8] Li J, Tang S, Li K, Zhang S, Tang L, Cao L, et al. Automatic recognition and classification of microseismic waveforms based on computer vision. Tunn Undergr Space Technol 2022;121:104327.
[9] Shu H, Dawod A. Microseismic monitoring signal waveform recognition and classification: review of contemporary techniques. Appl Sci 2023;13(23):12739.
[10] He J, Li H, Tuo X, Wen X, Rong W, He X. Strong noise-tolerance deep learning network for automatic microseismic events classification. IEEE Trans Geosci 2022;60:5918109.
[11] Ding L, Chen Z, Pan Y, Song B. Mine microseismic time series data integrated classification based on improved wavelet decomposition and ELM. Cognit Comput 2022;14(4):1526–46.
[12] Jiang R, Dai F, Liu Y, Li A. A novel method for automatic identification of rock fracture signals in microseismic monitoring. Measurement 2021;175:109129.
[13] Jiang R, Dai F, Liu Y, Wei M. An automatic classification method for microseismic events and blasts during rock excavation of underground caverns. Tunn Undergr Space Technol 2020;101:103425.
[14] Tang S, Wang J, Tang C. Identification of microseismic events in rock engineering by a convolutional neural network combined with an attention mechanism. Rock Mech Rock Eng 2021;54(1):47–69.
[15] He Z, Jia M, Wang L. UACNet: a universal automatic classification network for microseismic signals regardless of waveform size and sampling rate. Eng Appl Artif Intel 2023;126:107088.
[16] Ma C, Zhang H, Lu X, Ji X, Li T, Fang Y, et al. A novel microseismic classification model based on bimodal neurons in an artificial neural network. Tunn Undergr Space Technol 2023;131:104791.
[17] Bi X, Zhang C, He Y, Zhao X, Sun Y, Ma Y. Explainable time–frequency convolutional neural network for microseismic waveform classification. Inf Sci 2021;546:883–96.
[18] Sun C, Shrivastava A, Singh S, Gupta A. Revisiting unreasonable effectiveness of data in deep learning era. In: Proceedings of the 16th IEEE International Conference on Computer Vision (ICCV); 2017 Oct 22–29; Venice, Italy. New York City: IEEE; 2017. p. 843–52.
[19] Zhou Y, Meng S, Lou Y, Kong Q. Physics-informed deep learning-based real-time structural response prediction method. Engineering 2024;35(4):140–57.
[20] Tzeng E, Hoffman J, Saenko K, Darrell T. Adversarial discriminative domain adaptation. In: Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2017 Jul 21–26; Honolulu, HI, USA. New York City: IEEE; 2017. p. 2962–71.

[21] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A comprehensive survey on transfer learning. Proc IEEE 2021;109(1):43–76.
[22] Ganin Y, Lempitsky V. Unsupervised domain adaptation by backpropagation. In: Proceedings of the 32nd International Conference on Machine Learning; 2015 Jul 7–9; Lille, France. Cambridge: JMLR; 2015. p. 1180–9.
[23] Zhao S, Yue X, Zhang S, Li B, Zhao H, Wu B, et al. A review of single-source deep unsupervised visual domain adaptation. IEEE Trans Neural Netw Learn Syst 2022;33(2):473–93.
[24] Zhu Y, Zhuang F, Wang J, Chen J, Shi Z, Wu W, et al. Multi-representation adaptation network for cross-domain image classification. Neural Netw 2019;119:214–21.
[25] Feng P, Fu J, Ge Z, Wang H, Zhou Y, Zhou B, et al. Unsupervised semantic-aware adaptive feature fusion network for arrhythmia detection. Inf Sci 2022;582:509–28.
[26] He Z, Chen Y, Yuan S, Zhao J, Yuan Z, Polat K, et al. A novel unsupervised domain adaptation framework based on graph convolutional network and multi-level feature alignment for inter-subject ECG classification. Expert Syst Appl 2023;221:119711.
[27] Sun K, Bo L, Ran H, Tang Z, Bi Y. Unsupervised domain adaptation method based on domain-invariant features evaluation and knowledge distillation for bearing fault diagnosis. IEEE Trans Instrum 2023;72:3531810.
[28] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27–30; Seattle, WA, USA. New York City: IEEE; 2016. p. 770–8.
[29] Krizhevsky A, Sutskever I, Hinton GE. ImageNET classification with deep convolutional neural networks. Commun ACM 2017;60(6):84–90.
[30] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015 Jun 7–12; Boston, MA, USA. New York City: IEEE; 2015. p. 1–9.
[31] Ma N, Zhang X, Zheng H, Sun J. ShuffleNet V2: practical guidelines for efficient CNN architecture design. In: Proceedings of the 15th European Conference on Computer Vision (ECCV); 2018 Sep 8–14; Munich, Germany. Berlin: Springer; 2018. p. 122–38.
[32] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018 Jun 18–23; Salt Lake City, UT, USA. New York City: IEEE; 2018. p. 7132–41.
[33] Zhao M, Zhong S, Fu X, Tang B, Pecht M. Deep residual shrinkage networks for fault diagnosis. IEEE Trans Industr Inform 2020;16(7):4681–90.
[34] Long M, Cao Y, Wang J, Jordan M. Learning transferable features with deep adaptation networks. In: Bach F, Blei D, editors. Proceedings of the 32nd International Conference on Machine Learning (ICML); 2015 Jul 6–11; Lille, France. Cambridge: JMLR; 2015. p. 97–105.
[35] Ganin Y, Ustinova E, Ajakan H, Germain P, Larochelle H, Laviolette F, et al. Domain-adversarial training of neural networks. J Mach Learn Res 2016;17(1):2096–130.
[36] Saito K, Watanabe K, Ushiku Y, Harada T. Maximum classifier discrepancy for unsupervised domain adaptation. In: Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018 Jun 18–23; Salt Lake City, UT, USA. New York City: IEEE; 2018. p. 3723–32.
[37] Li S, Lv F, Xie B, Liu C, Liang J, Qin C, et al. Bi-classifier determinacy maximization for unsupervised domain adaptation. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence/33rd Conference on Innovative Applications of Artificial Intelligence/11th Symposium on Educational Advances in Artificial Intelligence; 2021 Feb 2–9; online. Red Hook: Curran Associates, Inc.; 2021. p. 8455–64.
[38] Zheng S, Lu J, Zhao H, Zhu X, Luo Z, Wang Y, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2021 Jun 19–25; Nashville, TN, USA. New York City: IEEE; 2021. p. 6877–86.
[39] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag 2009;45(4):427–37.
[40] Wang A, Chen H, Lin Z, Han J, Di G. RepViT: revisiting mobile CNN from vit perspective. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2024 Jun 16–22; Seattle, WA, USA. New York City: IEEE; 2024. p. 15909–20.
[41] Ma X, Dai X, Bai Y, Wang Y, Fu Y. Rewrite the stars. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2024 Jun 16–22; Seattle, WA, USA. New York City: IEEE; 2024. p. 5694–703.
[42] Tan M, Le Q. EfficientNet: rethinking model scaling for convolutional neural networks. PMLR 2019;97:6105–14.
[43] Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L. MobileNetV2: inverted residuals and linear bottlenecks. In: Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018 Jun 18–23; Salt Lake City, UT, USA. New York City: IEEE; 2018. p. 4510–20.
[44] Han S, Shao H, Cheng J, Yang X, Cai B. Convformer-NSE: a novel end-to-end gearbox fault diagnosis framework under heavy noise using joint global and local information. IEEE/ASME Trans Mechatron 2023;28(1):340–9.
[45] Maaten V, Hinton G. Visualizing data using *t*-SNE. J Mach Learn Res 2008;92:579–605.
[46] Zhang Y, Liu T, Long M, Jordan MI. Bridging theory and algorithm for domain adaptation. PMLR 2019;97:7404–13.
[47] Jin Y, Wang X, Long M, Wang J. Minimum class confusion for versatile domain adaptation. In: Vedaldi A, Bischof H, Brox T, Frahm JM, editors. Proceedings of the 16th European Conference on Computer Vision (ECCV); 2020 Aug 23–28; Glasgow, UK. Berlin: Springer; 2020. p. 464–80.
[48] Hospedales T, Antoniou A, Micaelli P, Storkey A. Meta-learning in neural networks: a survey. IEEE Trans Pattern Anal Mach Intell 2022;44(9):5149–69.
[49] Xiang Z, Peng W, Liu X, Yao W. Self-adaptive loss balanced physics-informed neural networks. Neurocomputing 2022;496:11–34.
[50] Wang Q, Wu B, Zhu PF, Li P, Zuo W, Hu Q. ECA-Net: efficient channel attention for deep convolutional neural networks. 2013. arXiv: 1910.03151.
[51] Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps. 2013. arXiv: 1312.6034.
[52] Arrieta AB, Diaz-Rodriguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 2020;58:82–115.