Complex & Intelligent Systems (2025) 11:97
https://doi.org/10.1007/s40747-024-01747-y

ORIGINAL ARTICLE

Sample-prototype optimal transport-based universal domain adaptation for remote sensing image classification

Xiaosong Chen[1,2] · Yongbo Yang[3] · Dong Liu[1,4,5] · Shengsheng Wang[2]

Received: 25 May 2024 / Accepted: 15 December 2024 / Published online: 28 December 2024

Abstract

In recent years, there is a growing interest in domain adaptation for remote sensing image scene classification, particularly in universal domain adaptation, where both source and target domains possess their unique private categories. Existing methods often lack precision on remote sensing image datasets due to insufficient prior knowledge between the source and target domains. This study aims to effectively distinguish between common and private classes despite large intra-class sample discrepancies and small inter-class sample discrepancies in remote sensing images. To address these challenges, we propose Sample-Prototype Optimal Transport-Based Universal Domain Adaptation (SPOT). The proposed approach comprises two key components. Firstly, we utilize an unbalanced optimal transport algorithm along with a sample complement mechanism to identify common and private classes based on the optimal transport assignment matrix. Secondly, we leverage the optimal transport algorithm to enhance discriminability among different classes while promoting similarity within the same class. Experimental results demonstrate that SPOT significantly enhances classification accuracy and robustness in universal domain adaptation for remote sensing images, underscoring its efficacy in addressing the identified challenges.

Keywords universal domain adaptation · remote sensing · scene classification · optimal transport

Introduction

In recent years, there has been an increasing demand for high-precision land cover classification [1] across various sectors such as geographic information systems, environmental monitoring, agriculture, urban planning, and traffic analysis [2, 3]. This has led to research and application advancements in remote sensing image classification technology [4]. Various deep learning methods have been employed for scene classification in remote sensing images, including convolutional neural networks (CNNs [5–7]), wasserstein generative adversarial networks (WGANs [8]), soft computing techniques [9], transformers [10], gated recurrent unit network [11], hybrid neural network [12], cellular network [13], and others. However, advancements in remote sensing image sensor technology, coupled with factors such as geographic location, imaging conditions, weather changes, and climate conditions, have expanded the diversity of data resources available for remote sensing image classification. Hence, remote sensing images collected from different regions or at different times within the same region may exhibit disparities, leading to variations in the distribution of training and testing data used for remote sensing image classification,

✉ Dong Liu
liudong@xnu.edu.cn

Xiaosong Chen
chenxs22@mails.jlu.edu.cn

Yongbo Yang
15874316@qq.com

Shengsheng Wang
wss@jlu.edu.cn

1 Hunan Engineering Research Center of Advanced Embedded Computing and Intelligent Medical Systems, Xiangnan University, Chenzhou 423300, China

2 College of Computer Science and Technology, Jilin University, Changchun 130012, Jilin, China

3 Education Evaluation Center, Air Force Aviation University, Changchun 130000, Jilin, China

4 School of Computer and Artificial Intelligence, Xiangnan University, Chenzhou 423300, China

5 Key Laboratory of Medical Imaging and Artificial Intelligence of Hunan Province, Xiangnan University, Chenzhou 423300, China

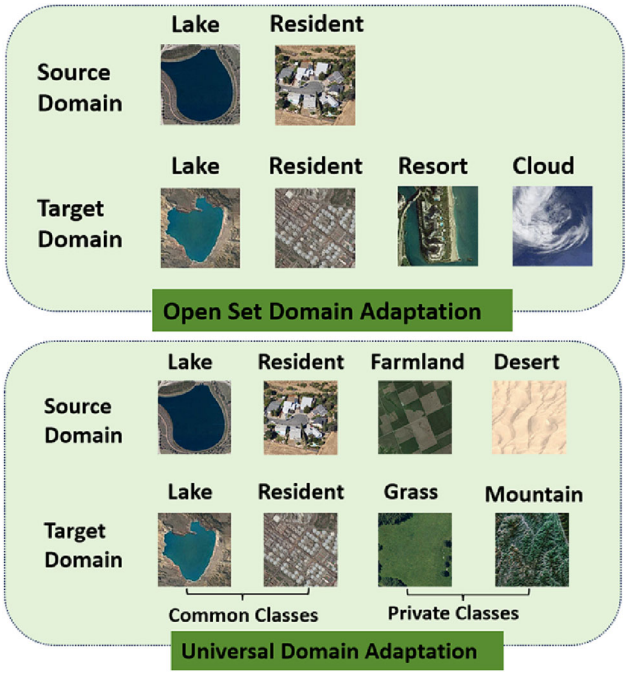

Fig. 1 Realistic Open-set and Universal DA scenarios.Open-set DA, the label set of source domain is a subset of the target domain labels. Universal DA, the source and target domains have a shared label set, as well as individual private label sets

resulting in a phenomenon known as domain gap [14]. Furthermore, dissimilar to traditional datasets such as Office, OfficeHome, and VisDA, remote sensing images require preprocessing before scene classification. This preprocessing includes operations such as cloud removal to enhance the clarity of the experimental images. Similarly, the ICA [15] algorithm involves analogous image deblurring operations. Remote sensing datasets exhibit significant variations in images of the same class within the same domain and across different domains. These variations arise from subtle similarities in sample features among images of the same class in remote sensing images and insufficient distinctiveness in sample features among images of different classes. To address the domain gap issue between the source domain dataset and the target domain dataset used for image classification, domain adaptation (DA [16]) methods have been proposed. Domain adaptation algorithms aim to train network models using source domain data to achieve high classification accuracy on diverse yet interconnected distributions of target domain data.

In the realm of remote sensing image scene classification, numerous contemporary domain adaptation methodologies concentrate on mitigating domain discrepancies through the acquisition of domain-invariant feature representations [17]. By understanding the relationship between the source label space and the target label space (i.e., category differences), domain adaptation can be categorized into closed-set domain adaptation [18], partial domain adaptation [19], and open-set domain adaptation [20]. Specifically, closed-set domain adaptation typically addresses the domain adaptation challenge by utilizing adversarial learning behaviors to align distributions in pixel [21], feature spaces [22], and output vectors [23], assuming a common label set between the source and target domains. However, in practical scenarios of remote sensing image scene classification, label sets for different scenes often exhibit private class distributions. To address this assumption, two alternative approaches have been proposed: partial domain adaptation, where the target domain label space is considered a subset of the source domain label space, and open-set domain adaptation [24], where the source domain label space is seen as a subset of the target domain label space. The lack of prior knowledge about the target domain label set presents a difficulty in choosing an appropriate domain adaptation method (closed-set domain adaptation, partial domain adaptation, or open-set domain adaptation). In such cases, a new domain adaptation approach called Universal Domain Adaptation (UniDA [25]) has been introduced.

UniDA is a method designed to overcome the constraints imposed by differences in label sets between the source and target domains, encompassing a broad spectrrum of domain adaptation scenarios. UniDA involves the utilization of a shared label set while also maintaining private label sets for both the source and target domains, as shown in Fig. 1. Nevertheless, the UniDA approach confronts two principal challenges. To begin with, directly matching the source and target domains might give rise to the misalignment of samples belonging to different classes, resulting in a reduction in the accuracy of model classification. Secondly, target samples from the private label set of the target domain have to be labeled as "unknown" due to the absence of labeled training data from the source domain for these target categories. In the current realm of universal domain adaptation research, various transfer learning methods [26] have been proposed for conventional image classification datasets, distinguishing between common classes and private classes in the target domain by manually adjusting thresholds [27]. However, empirical validation often reveals that the application of these universal domain adaptation methods, which perform effectively on conventional datasets, to remote sensing image scene classification tasks often yields unsatisfactory results. A remote sensing image classification model trained solely on the source domain encounters difficulties in extracting comprehensive prior knowledge between the label sets of source and target domains. Furthermore, the diverse proportions of common class samples in the target domain dataset make the extension of classification models to more realistic universal domain adaptation tasks through manual threshold adjustments a challenging undertaking.

In view of the challenges faced by conventional universal domain adaptation methods in achieving high classification accuracy when applied to remote sensing image datasets, this research proposes the utilization of the Optimal Transport (OT) method. The OT approach has manifested significant classification accuracy in both closed-set [28] and open-set [29] domain adaptation scenarios. The optimal Transport algorithm serves as an effective optimization method that assess the similarity between distributions by determining the shortest distance calculation from one probability distribution to another [30]. In contrast to conventional methods, our proposed algorithm adeptly handles domain shift by leveraging the Optimal Transport framework to align distributions at both local and global levels. This dual alignment strategy not only enhances the classification robustness by accurately identifying and aligned common class samples but also maintains the integrity of class prototypes within the target domain. To facilitate the local mapping between target samples and source prototypes, a partial alignment method based on the unbalanced Optimal Transport algorithm [31] is introduced to identify common class samples in the target domain. The unbalanced Optimal Transport algorithm's ability to balance the sample distribution between common and private classes addresses the challenge of class imbalance, which is a prevalent in remote sensing datasets. By incorporating a sample completion mechanism, our proposed method ensures a comprehensive representation of all classes, thereby improving classification performance. The study presents a novel target sample clustering method within the framework of global mapping between target samples and target prototypes. This method utilizes the Optimal Transport [32] algorithm to establish mappings from target sample vectors to target prototype vectors. The clustering method is designed to enhance the discriminability of target categories while preserving local consistency, particularly advantageous for high-resolution remote sensing images. By mitigating the risk of overfitting to specific class counts, this method showcases versatility and efficacy across various datasets. Essentially, the Optimal Transport-based approach provides a robust, adaptable, and high-performing solution for remote sensing image classification in universal domain adaptation conditions. This innovative methodology ensures accurate and reliable classification across diverse and complex remote sensing datasets, establishing it as a superior alternative to existing approaches.

In summary, our contributions can be categorized into three main aspects:

1. The study suggests utilizing the unbalanced Optimal Transport algorithm with remote sensing datasets to detect common class samples in the target domain for domain alignment. Experimental findings demonstrate a significant improvement in aligning domain distributions, representing the first successful application of the Optimal Transport algorithm for universal domain adaptation in remote sensing image datasets.
2. A sample completion mechanism has been devised to automate the process of determining thresholds for selecting high-confidence common class samples, eliminating the need for manual intervention. This method utilizes the mean sample count as the threshold. Experimental evidence supports the effectiveness of this method in selecting high-confidence samples and its successful adaptation in real-world domain adaptation tasks, especially in situations where there are fluctuations in sample counts for common and private classes.
3. In addition to identifying common class samples, a global Optimal Transport algorithm is integrated to aid in mapping feature vectors between target samples and target prototypes. The clustering of target domain samples using this algorithm leads to enhanced feature distinctiveness among different classes and increased similarity within the same class. The experimental outcomes highlight substantial improvements in feature separability and classification accuracy.

Related works

Remote sensing image scene classification

Remote sensing image scene classification has garnered substantial attention in recent years due to its diverse applications in fields such as geographic information systems, environmental monitoring, agriculture, urban planning, and traffic analysis. Deep learning methods [33] have shown promising results in this domain. Common deep learning methods used for remote sensing image classification include autoencoders [34], convolutional neural networks (CNNs [5, 6]), wasserstein generative adversarial networks (WGANs [8]), category prototype-based memory networks [35], and transformers [36]. These methods are based on the assumption that the distributions of training and testing data are similar, although this may not always hold true due to variations in imaging conditions, geographic locations, and other factors. This discrepancy between domains, known as domain gap, poses a challenge to the effectiveness of classification tasks.

Universal domain adaptation

To address the domain gap issue, scholars have proposed domain adaptation (DA) algorithms [37]. Conventional DA methods aim to align the distributions of the source and target domains by obtaining domain-invariant feature representations. However, these methods often rely on assumptions about the relationship between the source and target domains,

which may not hold in real-world scenarios. Universal domain adaptation (UniDA) presents an innovative approach that aims to eliminate these assumptions and adapt to various unfamiliar domain adaptation settings. UniDA encompasses closed-set, partial, and open-set domain adaptation settings, making it applicable to a wide range of scenarios.

Optimal transport

The Optimal Transport (OT) algorithm has exhibited potential in addressing domain adaptation challenges [38]. OT entails an optimization problem that measures the similarity between probability distributions and seeks to minimize the distance computation from one distribution to another. While several domain adaptation methods based on OT have been proposed, most focus on global mapping from one sample to another under closed-set transfer learning conditions [39]. However, in universal domain adaptation, it is crucial to identify common class samples and perform cross-domain alignment. Therefore, there is a need for customized OT-based methods designed specifically for universal domain adaptation settings, especially in tasks such as remote sensing image scene classification.

Methodology

In this section, the UniDA problem is examined from two prespectives: distinguishing common samples and clustering target samples, as shown in Fig. 2.

Preliminary

In this section, an explanation of the background of the optimal transport algorithm is provided. Optimal transport refers to the process of identifying the most effective route to minimize the costs related to transferring one probability distribution to another. Following this, an overview of the original formulation of the optimal transport algorithm is provided.

Given two probability distribution vectors are $\alpha \in R^a$ and $\beta \in R^b$, where 1_d represents a d-dimensional vector of ones, the joint probability distribution matrix Q of the distributions α and β can be represented as follows:

$$U(\alpha, \beta) = \left\{ Q \in R_+^{a \times b} | Q1_b = \alpha, Q1_a = \beta \right\}. \tag{1}$$

The set $U(\alpha, \beta)$ can be interpreted as the collection of all possible joint probability distribution matrices Q for the probability distributions α and β, where the marginal distributions of the joint probability distribution matrix are α and β, respectively [26]. Given a cost matrix $M \in R_+^{a \times b}$ representing the similarity of distributions, the optimal transport algorithm minimizes the transportation cost by considering the product of the cost matrix M and all possible transportation matrices Q in $U(\alpha, \beta)$. It seeks the optimal coupling matrix Q^* that minimizes the transportation cost by optimizing the following minimization problem:

$$\min_{P \in (a,b)} \langle C, P \rangle \tag{2}$$

The optimal coupling matrix Q^* obtained from global optimal transport algorithms rigorously enforces the mass constraints between marginal distributions, rendering it inadequate for addressing partial alignment challenges in open-set or universal domain adaptation scenarios.

In consideration of this, a new unbalanced optimal transport algorithm is proposed to address the limitation of preserving marginal distribution mass conservation, aiming to achieve partial alignment effects [40]. The formula for the unbalanced optimal transport algorithm is presented as follows:

$$UOT(M, \alpha, \beta) = argmax_{Q \in R_+} Tr\left(Q^\top M\right) \tag{3}$$

The optimization problem described above is addressed by employing the generalized Sinkhorn algorithm [30].

Distinguish samples from common classes

In the current domain adaptation research landscape, a prevalent methodology involves the identification of high-confidence samples belonging to common classes in the target domain to aid in domain alignment. This process typically involves labeling samples belonging to private classes in the target domain as an unknown class and manually adjusting thresholds to filter samples linked to common classes [41]. However, manual threshold adjustments are not suitable for practical domain adaptation scenarios where there are significant variations in the class distributions. Setting the threshold too high may lead to misclassifying some common class samples as unknown private classes, while setting it too low could result in misidentifying private class samples as common classes. Both circumstances have the potential to negatively affect the overall classification accuracy.

To more accurately distinguish samples belonging to common classes in the target domain, a method based on an unbalanced optimal transport algorithm is proposed to partially map target samples to the centroids of source domain classes. The partial alignment problem between target samples z_j^t and source prototypes c_i^t is addressed by our method from the perspective of universal domain adaptation. Given that universal domain adaptation allows for the existence of private classes in both the source and target domains, the

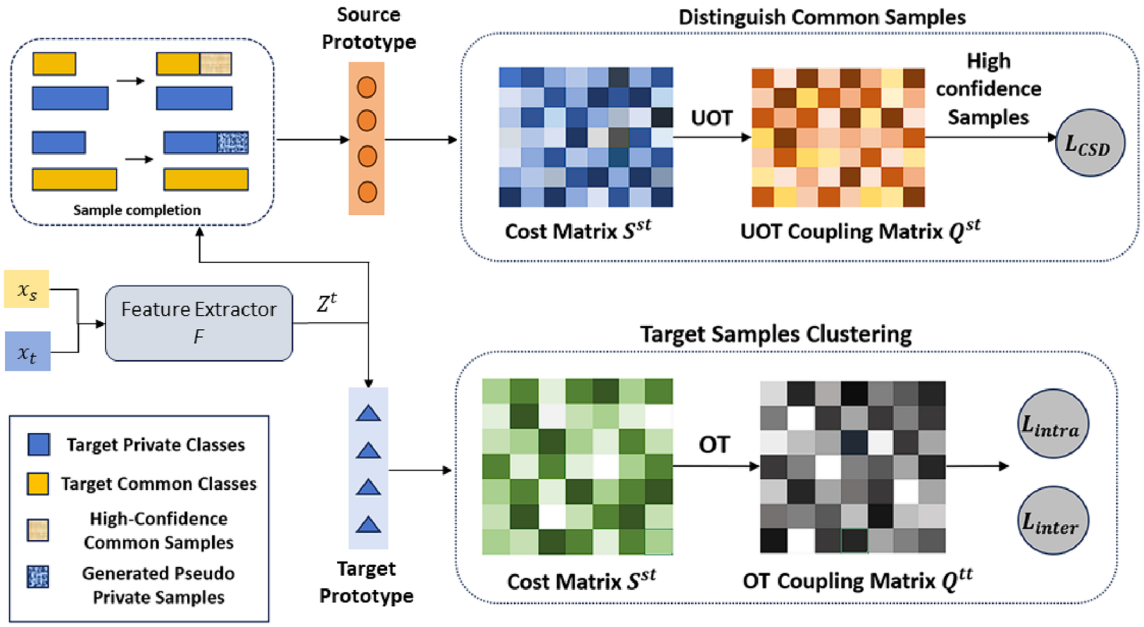

Fig. 2 Overview of the proposed SPOT. Our model consists of two parts: (1) Distinguish Common Samples: Initially, a sample completion mechanism is applied to the target samples, followed by partial alignment of target samples and source prototypes based on Unbalanced Optimal Transport (UOT). Subsequently, common class samples are filtered based on the mean of sample quantities. (2) Target Samples Clustering: A global mapping based on Optimal Transport (OT) is performed between the target samples and target prototypes, aiming to encourage consistency among samples of the same class while enhancing the global discriminability of samples from different classes

partial alignment method based on the unbalanced optimal transport algorithm can selectively extract common knowledge from both domains. The optimal assignment matrix Q^{st} obtained from the unbalanced optimal transport algorithm is represented as follows:

$$Q^{st} = UOT\left(S^{st}, \frac{1}{B}1_B, \frac{1}{|C_s|}1_{|C_s|}\right) \tag{4}$$

The cosine similarity matrix between target samples and source domain class prototypes serves as the cost matrix for the unbalanced optimal transport algorithm. In this algorithm, the initial marginal probability vector for source prototypes is set as $\beta^0 = \frac{1}{|C_s|}1_{|C_s|}$. However, since the source domain includes private classes, it is unreasonable to give equal weight to each source prototype in the marginal probability vector β. Instead, the marginal probability vector β needs to be updated based on a moving average using the allocation matrix Q^{st} from the previous iteration to better represent the true distribution of source prototypes. The vector $\beta^{'}$ used for updating the input marginal probability vector β in each iteration of the unbalanced optimal transport algorithm is the sum of each column in Q^{st}. The formulas for updating the vector β and the marginal probability vector for source prototypes in each iteration are as follows:

$$\beta^{(t+1)} = \mu\beta^{(t)} + (1-\mu)\beta', \beta'_i = \sum_{j=1}^{B} Q^{st}_{j,i} \tag{5}$$

Firstly, normalize the allocation matrix Q^{st}. Extract common knowledge from the normalized matrix Q^{st}, where the maximum value in each row of Q^{st} is set as the confidence w_i^t that the target sample belongs to the corresponding source domain class, and the corresponding source domain class is assigned as the pseudo-label for the target sample. It is established that the summation of each column in the allocation matrix Q^{st} represents the confidence, denoted as w_j^s, for each source domain category's affiliation with the common label set. The calculation formulas for target sample confidence w_i^t and source prototype confidence w_j^s are detailed below:

$$w_i^t = max(Q^{st}_{i,1}, Q^{st}_{i,2}, ..., Q^{st}_{i,|C_s|}), w_j^s = \sum_{i=1}^{B} Q^{st}_{i,j} \tag{6}$$

In the target and the source domains, higher confidence weights are attributed to common category samples. Thresholds are determined based on the sample quantity in the target domain and the number of categories in the source domain. Target samples with confidence levels above the

average number of target samples are recognized as common category samples in the target domain. Similarly, source prototypes with confidence levels exceeding the average number of categories in the source domain are classified as common category samples in the source domain. Unlike manual threshold setting, which requires subjective adjustments, the average number of target samples is derived from the inherent statistical attributes of the target domain. This enables the algorithm to dynamically adjust thresholds, catering to diverse dataset requirements without necessitating repetitive manual interventions for threshold calibration. Additionally, when confronted with non-uniform distributions of target domain samples, manual threshold specification may lead to over fitting. Conversely, employing the mean as the threshold can mitigate over fitting risks, thereby enhancing the algorithm's generalizability.

The target common samples z_i^t and their associated pseudo-labels y_i^t selected in the preceding step are utilized to compute the inter-domain common class detection loss, denoted as L_{csd1}, employing a standard cross-entropy loss function. w_i^t represents the confidence of z_i^t, while $w^s_{y_i^t}$ indicates the confidence of the source domain category corresponding to the pseudo-label of z_i^t. The criteria for selecting target common samples z_i^t and the loss function L_{csd1} are expressed as follows:

$$\mathcal{L}_{csd1} = \mathcal{L}_{CE}(z_i^t, y_i^t), w_i^t > \frac{1}{B} \quad and \quad w^s_{y_i^t} > \frac{1}{||C_s||} \tag{7}$$

To further enhance the accuracy of common sample classification in our algorithm, we introduce an auxiliary loss function, denoted as L_{csd2}, within the common sample detection function. Similar to L_{csd1}, L_{csd2} constructs its loss function by filtering common samples using sample confidence weights. However, unlike L_{csd1}, the sample confidence weights $w_i^{'t}$ in L_{csd2} are obtained as the summation of weights in each row of the matrix Q^{st}. Evidently, $w_i^{'t}$ carries greater weights compared to w_i^t, resulting in a more lenient filtering criterion that selects a larger proportion of samples potentially belonging to the target common class. However this approach has a drawback as it may include non-common class samples in the training of the classification model, thereby reducing its classification accuracy. To leverage the advantages of both L_{csd1} and L_{csd2}, these two loss functions are simultaneously trained by introducing a hyperparameter μ, defined as the ratio between the current iteration count and the total number of iterations. Specifically, the hyperparameter of L_{csd1} is set to $1-\mu$, and that of L_{csd2} is set to μ. As the iteration count increases, the value of μ gradually approaches 1. The rationale behind this design is as follows: in the initial iterations when the model's classification accuracy is relatively low, a larger hyperparameter for L_{csd1} maximizes its influence, ensuring the selection of samples with higher potential to belong to common classes, thereby maintaining high sample quality for model training. Subsequently, as the iteration progresses and the model's classification accuracy enhances, the hyperparameter μ of L_{csd2} approaches 1. This adaptation results in a more lenient filtering criterion, facilitating the inclusion of a large pool of potential common class samples while maintaining sample quality. This synergistic interplay between the two loss functions ensures the selection of a appropriate number of high-quality common class samples for model training, thereby effectively preserving the model's classification accuracy. The computational formulas for $w_i^{'t}$ and $w_j^{'s}$ are detailed below:

$$w_i^{'t} = \sum_{j=1}^{|C_s|} Q^{st}_{i,j}, w_j^{'s} = \sum_{i=1}^{B} Q^{st}_{i,j} \tag{8}$$

Our approach to computing L_{csd2} is the same as for L_{csd1}, along with the criteria for selecting target common samples $z_i^{'t}$,represented as follows:

$$\mathcal{L}_{csd2} = \mathcal{L}_{CE}(z_i^{'t}, y_i^{'t}), w_i^{'t} > \frac{1}{B} \quad and \quad w^{'s}_{y_i^{'t}} > \frac{1}{||C_s||} \tag{9}$$

The overall loss function for common class distinguish is shown as follows.

$$L_{CSD} = (1-\mu)L_{csd1} + \mu L_{csd2} \tag{10}$$

Sample completion mechanism

In domain adaptation research domain, an imbalance in sample quantities across different categories, frequently witnessed in closed-set scenarios, can lead to a reduction in the effectiveness of classification models. These models often demonstrate a tendency to prioritize the prediction of classes with larger sample sizes, a phenomenon also remarked in studies on universal domain adaptation studies. In scenarios where a considerable disparity exists in sample quantities between common and private classes within the target domain, classification models may incorrectly identify certain common class samples as belonging to unknown categories. Moreover, as discussed in Sect. 3.2, establishing the average sample quantity as the threshold for distinguishing common class samples could worsen the prevalence of biased results in classification models, particularly when the number of common class samples is considerably smaller than that of private classes.

In order to improve the accurate differentiation between common and private class samples in classification models, a sample completion mechanism is proposed to address the issue of imbalanced sample quantities across classes. Our algorithm aims to identify the disparity in sample quantities

between common and private classes and augment samples for classes with fewer samples, thereby achieving a balance in sample quantities between the two type of classes. Initially, the cosine similarity matrix between target samples and source prototypes is computed. Subsequently, the maximum similarity of each sample is compared with a predefined threshold, denoted as γ, known as the rough boundary. Samples with similarity exceeding γ are classified as common class samples, while those below this threshold are categorized as private class samples. To address the deficiency in private class samples, a strategy is employed where the least similar source prototype to the target sample is utilized. This approach generates pseudo-private samples $\mathcal{Z}_i^t$ by averaging the feature vectors of the target sample and its least similar source prototype, as follows:

$$\mathcal{Z}_i^t = \frac{1}{2}(z_i^t + argmin_{c_k^s}(z_i^t c_k^s)) \tag{11}$$

To complete common class sample, a methodology similar to the one outlined in Sect. 3.2 is employed. The confidence level of each target sample belonging to source prototypes in the memory queue is determined using the unbalanced optimal transport algorithm. Samples meeting the criteria of exceeding the average sample quantity in confidence level are labeled as high-confidence source prototypes for enhancing common class. During this process, the marginal probability vector of source prototypes in the unbalanced optimal transport algorithm adopts the vector β as defined in Sect. 3.2. Consequently, updates are only needed when identifying common class samples in Section 3.2, with no updates necessary during the sample completion process.

Clustering of target samples

To enhance the compactness of feature representations within the same category and enhance the discriminative capability among different categories in the target domain, a domain clustering strategy is proposed based on the global optimal transportation algorithm. Firstly, K target sample features are randomly selected from the memory queue to serve as trainable target prototypes, and the most similar neighbor samples of the target samples for the current iteration are then identified and filtered out from the memory queue. A cosine similarity matrix, denoted as S^{tt}, between the target sample features Z^t and the most similar neighbor samples Z'^t with the target prototype C_t is used as the cost matrix for the global optimal transport algorithm. By resolving the optimal transport optimization problem, the target samples are aligned with their corresponding target prototypes.

$$Q^{tt} = OT\left(S^{tt}, \frac{1}{K}1_K, \frac{1}{2B}1_{2B}\right) \tag{12}$$

The allocation matrix Q^{tt} obtained from the solving algorithm generates individual row vectors q_i^{tt}, which represent the classification probability vector for the target sample z_i^t with respect to the target prototype. Each row vector q_i^{tt} of Q^{tt} is summed to 1/B; hence, Q^{tt} is multiplied by B to ensure that the sum of each row's classification vector is 1. To enhance the intra-class consistency, with the goal of minimizing the distances between samples of the same class, the cross-entropy loss function L_{intra} is constructed using the clustering prediction vector p_i^t for the target sample feature z_i^t and the optimal transport prediction vector q_i^{tt}. By minimizing L_{intra}, the feature extractor f is trained and the target prototype C_t is optimized.

$$\mathcal{L}_{intra} = \frac{1}{B}\sum_{i=1}^{B} l(q_i^{tt}, p_i^t) \tag{13}$$

The cross-entropy loss function $l(q_i^{tt}, p_i^t)$ can be represented as:

$$l(q_i^{tt}, p_i^t) = -\sum_{k=1}^{K} q_{i,k}^{tt} log p_{i,k}^t \tag{14}$$

Moreover, to enhance the global discriminability of features across different classes, the optimal transport prediction vectors are swapped between target sample features and their nearest neighbor sample features to construct the cross-entropy loss function L_{inter}.

$$\mathcal{L}_{inter} = \frac{1}{2B}\sum_{i=1}^{B}[l(q_{B+i}^{tt}, p_i^t) + l(q_i^{tt}, p_{B+i}^t)] \tag{15}$$

Overall train objective

Integrating loss function $\mathcal{L}_{cls}$ from the source dataset, along with the common class distinguish loss $\mathcal{L}_{CSD}$, the target cluster loss $\mathcal{L}_{intra}$, and the inter-cluster loss $\mathcal{L}_{inter}$, a comprehensive objective can be expressed as

$$\mathcal{L}_{overall} = \mathcal{L}_{cls} + \lambda(\mathcal{L}_{CSD} + \mathcal{L}_{intra} + \mathcal{L}_{inter}) \tag{16}$$

Experiments

Experimental setup

Datasets

To demonstrate the effectiveness of our proposed algorithm, the research utilized the RSSCN7, AID, UC Merced(UCM), and NWPU-RESISC45 datasets as cross-domain scene classification datasets for remote sensing images. These datasets

Fig. 3 Some images from five common classes from four datasets

were obtained from Google Earth. The RSSCN7 dataset [42] consists of 7 remote sensing scene categories, with 400 images per category, resulting in a total of 2800 RGB remote sensing scene images, each sized at 400 × 400 pixels. The AID [43] dataset is a high-resolution satellite image dataset containing approximately 10,000 images distributed among 30 classes, with each image sized at 600 × 600 pixels. The UC Merced dataset [44] is a conventional remote sensing image scene classification dataset comprising 2100 remote sensing images belonging to 21 scene categories, with each scene category containing 100 images sized at 256 × 256 pixels. The NWPU-RESISC45 dataset [45] consists of 45 scene categories, with each category containing 700 remote sensing images sized at 256 × 256 pixels, totaling 31,500 images. The spatial resolutions of the remote sensing images in each dataset range from 0.2 ms to 30 ms.

The selection of the RSSCN7 dataset as the source domain dataset was based on its relatively small number of categories and its higher Jaccard index in comparison to the other three datasets. The remaining three datasets were as target domain datasets for the development of three cross-domain scene classification tasks for remote sensing images. For the RSSCN7 → AID and RSSCN7 → NWPU tasks, the shared label set consisted 6 categories, namely farmland, forest, dense residential area, river, parking lot, and industrial area. In contrast, the RSSCN7 → UCM task had one less category in the shared label set, specifically excluding the industrial area. The complexity of the domain adaptation tasks was heightened by the high Jaccard index between the AID and NWPU-RESISC45 datasets, along with a larger number of shared label categories. Consequently, a fourth remote sensing image classification task, AID → NWPU, was created. Sample images of these four datasets are shown in Fig. 3. Additional information regarding the four cross-domain universal domain adaptation remote sensing image classification tasks are presented in Table 1.

Evaluation criteria

During the testing phase, all private category samples in the target domain are uniformly classified into an unknown class. Consequently, the number of categories in the target domain comprises the number of common categories plus one unknown category. The classification accuracy for each category is computed. The unbalanced optimal transport algorithm is employed to determine the maximum classification probability for each target sample. Subsequently, the threshold is established as the number of target samples, and if the predicted probability exceeds the mean of the sample numbers, the sample label is assigned to the source domain category corresponding to the maximum probability; otherwise, the sample is classified as an unknown class. The classification accuracy for each category, in addition to the average accuracy for common categories and the unknown category, serves as the representation of our final testing results.

Experimental details

The study employs an optimal transport solver based on the unbalanced optimal transport algorithm. The network model consists of a feature extractor denoted as F and a classifier denoted as C, initially trained on the source domain. The feature extractor adopts the pre-trained ResNet-50 [46] network as the initial backbone, comprising a backbone layer and a projection layer. The classifier network consists of three fully connected layers, with ReLU layers between the first two layers and BatchNorm layers between the last two layers. Pre-training of the network model with data from the source domain data was conducted to establish the cross-entropy loss function. The aggregator used for clustering target sample features is a single-layer fully connected network, preceded by data normalization. During the training phase, the batch size of 24 samples was set for both the source and target domain data. Stochastic gradient descent (SGD) was employed to train the network model. In all cross-domain scene classification experiments, the initial learning rate for the pre-trained ResNet network was set to 1×10^{-3}, while for the classifier and aggregator, it was set to 1×10^{-2}. The momentum for the SGD method across all networks was set to 0.9. For the cross-domain classification tasks RSSCN7→UCM, RSSCN7→AID, and RSSCN7→NWPU, the total training steps were set to 40,000 epochs, and capacity of the memory queue for target samples was set to 2000. Due to the larger data scale of the AID→NWPU task, the total training steps were set to 80,000

Table 1 Four Universal DA Tasks for Remote Sensing Scene Classification

Task	Shared	Source Private	Target Private	ξ
RSSCN2NWPU	6	1	39	0.12
RSSCN2UCM	5	2	16	0.18
RSSCN2AID	6	1	24	0.12
AID2NWPU	19	11	26	0.27

epochs, and the memory queue's capacity for target samples was set to 4000.

All experiments were conducted on a 1080Ti GPU with 11GB of memory. The experiments RSSCN7→UCM, RSSCN7→AID, and RSSCN→NWPU each required approximately 4.5 h to complete, while the experiment AID→NWPU took around 9 h.

Experimental results

Compared methods

The SPOT algorithm we propose is assessed in comparison with various state-of-the-art universal domain adaptation algorithms within the realm of remote sensing image classification:

1. Source-only: The network model is trained solely on the source domain data, with data from the target domain input directly into the trained model for testing purposes.
2. UDA [25]: UDA paper is regarded as a seminal contribution in the realm of computer vision, specifically addressing universal domain adaptation research. It introduces the concept of sample transferable similarity weights, which are computed based on domain similarity and information entropy. This methodology aids in distinguishing between common class samples and private class samples in the target domain.
3. DANCE [47]: DANCE integrates neighborhood clustering with self-supervised learning and entropy-based feature alignment and rejection. Rather than relying solely on source categories, neighborhood clustering minimizes the entropy of the similarity distribution between target samples and source prototypes. This approach promotes the separation of known and unknown target samples, thereby preventing misalignment.
4. I-UAN [26]: I-UAN is an improvement upon the UDA method. In I-UAN, the sample transferable similarity weights are obtained by subtracting the second-highest prediction probability vector from the maximum prediction probability vector of the target samples. These weights are established according to the confidence levels of the target samples in distinguishing between common class and private class samples within the target domain.

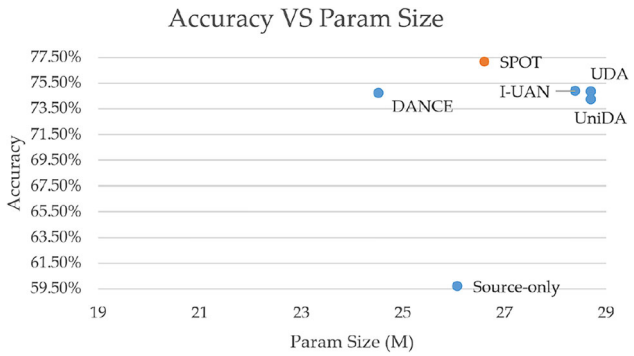

Fig. 4 Accuracy vs. Parameter Size for SPOT and Baseline Methods

5. Universal DA [48]: This method utilizes sample transferable similarity weights, which are calculated based on domain similarity and the highest probability of predicted class probabilities. These weights distinguish between common class and private class samples in the target domain based on a predefined threshold.

Results for the task RSSCN7→NWPU

The experimental results are presented in Table 2. The data clearly demonstrates that our proposed SPOT method achieves the highest classification accuracy for both common class samples and all target samples in comparison to other universal domain adaptation methods. More specifically, the classification accuracy for common class samples outperforms the second-highest method, I-UAN, by 1.31 percentage points, while the classification accuracy for all target samples is 2.27 percentage points higher than that of the I-UAN method. These results demonstrate the effectiveness of our proposed approach, which utilizes the optimal transport algorithm within the framework of universal domain adaptation.

Results for the task RSSCN7→UCM

The experimental results are presented in Table 3. It can be observed that our method outperforms all other methods in the study in terms of the classification accuracy for unknown class samples. Specifically, our approach surpasses the accuracy of the second-highest method, UDA, by 3.94 percentage points.

Table 2 Classification Accuracy of Different Methods on RSSCN7→NWPU

Methods	Grass	Field	River	Forest	Resident	Parking	Unknown	Avg-Share	Avg
Source-only	74.00	70.14	34.14	97.71	67.71	74.00	0.23	69.62	59.71
UDA [25]	74.29	85.85	**80.42**	92.71	86.00	87.42	**17.23**	84.45	74.85
DANCE [47]	**86.42**	70.85	77.28	**98.85**	94.71	95.00	0.02	87.18	74.73
I-UAN [26]	78.42	75.71	80.00	98.00	**97.42**	92.00	2.64	86.93	74.89
UniDA [48]	84.57	83.00	69.85	97.71	93.85	89.57	1.13	86.43	74.24
SPOT	80.86	**87.00**	80.14	92.29	90.26	**98.86**	10.71	**88.24**	**77.16**

Bold values indicate the best results

Table 3 Classification Accuracy of Different Methods on RSSCN7→UCM

Methods	Field	River	Forest	Resident	Parking	Unknown	Avg-Share	Avg
Source-only	93.00	60.00	98.00	42.00	67.00	1.63	72.00	60.27
UDA [25]	92.00	77.00	98.00	70.00	76.00	8.62	82.60	70.27
DANCE [47]	94.00	69.00	98.00	68.00	**100.00**	0.03	85.80	71.50
I-UAN [26]	**94.00**	**92.00**	98.00	**88.00**	81.00	0.06	**90.60**	**75.51**
UniDA [48]	92.00	89.00	98.00	79.00	89.00	0.12	87.20	74.52
SPOT	85.00	80.00	**98.00**	67.00	90.00	**12.56**	84.00	72.09

Bold values indicate the best results

Table 4 Classification Accuracy of Different Methods on RSSCN7→AID

Methods	Grass	Field	River	Forest	Resident	Parking	Unknown	Avg-Share	Avg
Source-only	68.46	79.19	63.41	100	90.73	80.00	0.12	80.3	68.84
UDA [25]	64.28	94.32	65.85	100.0	91.21	88.20	18.71	83.98	74.66
DANCE [47]	65.00	92.43	52.68	98.80	94.63	99.23	0.71	83.79	71.92
I-UAN [26]	61.78	**95.40**	59.26	98.80	98.29	**99.23**	1.22	**85.46**	73.43
UniDA [48]	**72.85**	91.62	45.36	99.20	**98.29**	96.41	1.68	83.96	72.20
SPOT	62.86	94.32	**68.05**	**100.0**	90.24	90.51	**19.89**	84.33	**75.13**

Bold values indicate the best results

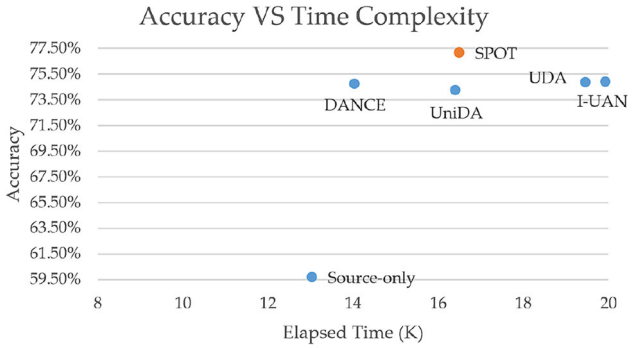

Fig. 5 Accuracy vs. Time Complexity for SPOT and Baseline Methods

Results for the task RSSCN7→AID

The experimental results are presented in Table 4. The data clearly indicate that our proposed SPOT method achieves the highest average classification accuracy across all target classes, including the unknown class, surpassing the second-highest UniDA method by 0.4 percentage points. Additionally, the classification accuracy for common class samples is second only to the I-UAN method, with a value of 0.8433.

Results for the task AID→NWPU

The experimental results are presented in Table 5. It can be observed that our proposed SPOT method achieves an average sample classification accuracy across all classes that exceeds the source-only method by 21.8 percentage points, ranking second only to the I-UAN method. Furthermore, it achieves the highest classification accuracy for unknown classes. These results effectively demonstrate the effectiveness of the UniDA-OT method in tackling more complex universal domain adaptation tasks.

Comparative analysis of accuracy and parameter size

Figure 4 depicts the comparison among the proposed method, SPOT, and several baseline methods in terms of accuracy versus parameter size. The x-axis represents the parameter size in millions (M), while the y-axis shows the accuracy percentage.

Table 5 Classification Accuracy of Different Methods on AID→NWPU

Methods	AI	BA	BE	BR	CH	CO	DR	DE	FA	FO	ME	MR	MO	PA	RS	RI	SR	ST	ST	Unknown	Avg-Share	Avg
Source-only	62.43	0.00	0.00	0.00	78.43	54.57	50.71	89.86	89	86	75.43	66.71	93.43	78.14	0.00	66.00	78.86	0.00	93.57	0.01	53.99	51.42
UDA[25]	47.43	71.42	**98.00**	90.86	75.28	57.86	65.43	74.57	88.43	86.57	88.28	75.57	46.43	81.86	77.57	71.43	83.57	68.42	90.43	16.14	75.76	72.78
DANCE[47]	44.14	69.71	84.85	81.14	68.13	58.85	77.14	73.57	88.57	93.00	77.14	76.71	78.57	74.00	74.71	72.14	76.14	67.71	92.71	1.86	75.21	71.54
I-UAN[26]	**52.57**	**81.43**	89.57	**91.43**	**83.71**	**61.00**	83.71	79.14	**96.28**	**94.86**	**91.00**	**82.29**	**84.57**	**86.28**	**80.71**	**79.00**	**92.71**	**85.00**	**94.43**	7.53	**83.67**	**79.86**
UniDA[48]	40.00	75.57	88.86	74.00	61.14	57.43	**83.86**	78.29	93.14	89.28	82.43	63.86	79.43	81.71	75.28	74.71	75.43	63.57	90.71	0.61	75.20	71.47
SPOT	50.00	76.14	97.86	90.42	76.14	52.71	62.28	**79.29**	88.71	75.14	87.43	73.71	63.43	81.43	80.00	70.28	83.43	65.57	90.71	**19.86**	76.04	73.23

Bold values indicate the best results

AI Airport, *BA* Baseballfield, *BE* Beach, *BR* Bridge, *CH* Church, *CO* Commercial, *DR* Dense Residential, *DE* Desert, *FA* Farmland, *FO* Forest, ME:Meadow, MR:Medium Residential, MO:Mountain, *PA* Parking, *RS* Railway Station, *RI* River, *SR* Sparse Residential, *ST* Stadium, *ST.* Storage Tanks

Table 6 Evaluation of the effectiveness of L_{CSD},L_{inter} and L_{intra}

L_{CSD}	L_{CSD}^{-}	L_{inter}	L_{intra}	RSSCN7→NWPU	RSSCN7→UCM	RSSCN7→AID	AID→NWPU
✓				77.34	76.00	76.45	65.87
✓			✓	82.93	79.00	82.42	74.92
✓		✓		83.75	81.00	81.78	75.74
		✓	✓	81.35	78.00	79.41	70.89
	✓	✓	✓	86.89	82.00	**85.06**	75.34
✓		✓	✓	**88.24**	**84.00**	84.33	**76.04**

Bold values indicate the best results

Our method, SPOT, attains the highest accuracy rate of 77.16% with a parameter size of 26.61M, indicating a significant improvement compared to other methods. The UDA method closely follows with the next best performance, achieving an accuracy of 74.85% with a larger parameter size of approximately 29 M. These results unequivocally demonstrate that SPOT offers a superior balance of accuracy and parameter efficiency, outperforming all other compared methods within the given experimental setup.

Comparative analysis of accuracy and time complexity

Figure 5 illustrates a performance evaluation that compares our proposed method, SPOT, with several baseline methods in terms of accuracy versus time complexity. The x-axis represents the elapsed time in thousands of seconds (K), while the y-axis shows the accuracy percentage.

Our method, SPOT, achieves the highest accuracy of 77.16% within a time frame of 16.5K seconds. This result signifies that SPOT not only offers superior accuracy but also maintains a reasonable computational cost compared to other methods. Subsequently, the I-UAN method achieves an accuracy of 74.89% within an elapsed time of approximately 20K seconds. The UDA method demonstrates comparable a similar time complexity, albeit with a slightly lower accuracy rate of 74.85%. These results clearly demonstrate that SPOT offers a superior balance between accuracy and time complexity, outperforming all other compared methods in the specified experimental setup.

Ablation study

To accurately evaluate the effectiveness of the three loss functions, L_{CSD}, L_{intra}, L_{inter}, as well as the effectiveness of the sample completion mechanism in L_{CSD}, the network model was trained using different combinations of these four components.

1. The Contribution of the Loss Functions are L_{inter} and L_{intra}. As shown in Table 6, the simultaneous presence of L_{inter} and L_{intra} significantly enhances the model's classification accuracy. This enhancement is particularly evident in the scene classification task RSSCN7→NWPU, where it improves by 10.9 percentage points compared to training the model with only L_{CSD}. Moreover, the contribution of L_{inter} to the model's classification accuracy surpasses that of L_{intra}, highlighting the importance of performing global discriminative clustering on target domain samples.
2. The Contribution of the Loss Function L_{CSD}. In the scene classification task RSSCN7→NWPU, the network model was trained using the loss function L_{CSD}, exhibiting a significant improvement in classification accuracy, surpassing the model trained solely on the source domain data by 17.65 percentage points. When compared to the model trained using the loss functions L_{inter} and L_{intra}, the classification accuracy improves by 6.89 percentage points. This indicates that selecting common class samples to construct the cross-entropy loss function can significantly enhance the model's classification accuracy.
3. The Contribution of the Sample Completion Mechanism. The loss function L_{CSD} without the sample completion mechanism is denoted as L_{CSD}^{-}. It can be observed that the contribution of the sample completion mechanism is most pronounced in the scene classification task RSSCN7→UCM, resulting in an improvement of 2 percentage points in the model's classification accuracy.

Conclusion

This study presents a novel approach to scene classification within remote-sensing image datasets by utilizing the Optimal Transport algorithm in the context of universal domain adaptation settings. Universal domain adaptation methods eliminate the need for prior knowledge regarding the relationship between source and target domains, making them applicable to all unknown domain adaptation scenarios. This characteristic is advantageous for practical applications in remote sensing image scene classification. We propose the implementation of the unbalanced Optimal Transport algorithm to enhance the representation of prevalent classes in

the target domain and to aid in the identification of common class samples for cross-domain alignment. Additionally, we introduce a target domain sample clustering method based on the global Optimal Transport algorithm, which enhances the similarity among samples within the same class while encouraging the discriminability of samples across different classes. The effectiveness of our proposed method is demonstrated through experimental results obtained from four remote sensing image classification tasks conducted under universal domain adaptation settings.

Acknowledgements This research was funded by Natural Science Foundation of Hunan Province (No. 2023JJ50392) and Scientific Research Fund of Hunan Provincial Education Department (No. 23A0588, 22C0564); Aid Program for Science and Technology Innovative Research Team in Higher Educational Institutions of Hunan Province.

Data availability The datasets supporting the findings of this study are publicly available and included within the article.

Declarations

Conflict of 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.

References

1. Zhu Q, Guo X, Deng W, Shi S, Guan Q, Zhong Y, Zhang L, Li D (2022) Land-use/land-cover change detection based on a siamese global learning framework for high spatial resolution remote sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing 184:63–78
2. Hassan MA, Javed R, Granelli F, Gen X, Rizwan M, Ali SH, Junaid H, Ullah S et al (2023) Intelligent transportation systems in smart city: a systematic survey. In: 2023 International Conference on Robotics and Automation in Industry (ICRAI), pp. 1–9. IEEE
3. Hassan MA, Javed AR, Hassan T, Band SS, Sitharthan R, Rizwan M (2022) Reinforcing communication on the internet of aerial vehicles. IEEE Transactions on Green Communications and Networking 6(3):1288–1297
4. Tuia D, Persello C, Bruzzone L (2021) Recent advances in domain adaptation for the classification of remote sensing data. arXiv preprint arXiv:2104.07778
5. Liu H, Ghadimi N (2024) Hybrid convolutional neural network and flexible dwarf mongoose optimization algorithm for strong kidney stone diagnosis. Biomedical Signal Processing and Control 91:106024
6. Han M, Zhao S, Yin H, Hu G, Ghadimi N (2024) Timely detection of skin cancer: An ai-based approach on the basis of the integration of echo state network and adapted seasons optimization algorithm. Biomedical Signal Processing and Control 94:106324
7. Cai X, Li X, Razmjooy N, Ghadimi N (2021) Breast cancer diagnosis by convolutional neural network and advanced thermal exchange optimization algorithm. Computational and Mathematical Methods in Medicine 2021(1):5595180
8. Arjovsky M, Chintala S, Bottou L (2017) Wasserstein generative adversarial networks. PMLR
9. Xu Z, Sheykhahmad FR, Ghadimi N, Razmjooy N (2020) Computer-aided diagnosis of skin cancer based on soft computing techniques. Open Medicine 15(1):860–871
10. Yu HQ (2024) Attention enhanced siamese neural network for face validation. Artificial Intelligence and Applications 2:21–27
11. Zhang L, Zhang J, Gao W, Bai F, Li N, Ghadimi N (2024) A deep learning outline aimed at prompt skin cancer detection utilizing gated recurrent unit networks and improved orca predation algorithm. Biomedical Signal Processing and Control 90:105858
12. Razmjooy N, Sheykhahmad FR, Ghadimi N (2018) A hybrid neural network-world cup optimization algorithm for melanoma detection. Open Medicine 13(1):9–16
13. Ali S, Hassan MA, Granelli F, Wang W, Sampedro GA, Al Hejaili A, Bouazzi I (2024) Optimizing multi-tier cellular networks with deep learning for 6g consumer electronics communications. IEEE Transactions on Consumer Electronics
14. Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan JW (2010) A theory of learning from different domains. Machine learning 79:151–175
15. Razmjooy N, Ramezani M, Ghadimi N (2017) Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal. International Journal of Fuzzy Systems 19:1144–1156
16. Saenko K, Kulis B, Fritz M, Darrell T (2010) Adapting visual category models to new domains. In: Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11, pp. 213–226. Springer
17. Wittich D, Rottensteiner F (2021) Appearance based deep domain adaptation for the classification of aerial images. ISPRS Journal of Photogrammetry and Remote Sensing 180:82–102
18. Xu Q, Ouyang C, Jiang T, Yuan X, Fan X, Cheng D (2022) Mffenet and adanet: a robust deep transfer learning method and its application in high precision and fast cross-scene recognition of earthquake-induced landslides. Landslides 19(7):1617–1647
19. Cao Z, Ma L, Long M, Wang J (2018) Partial adversarial domain adaptation. In: Proceedings of the European conference on computer vision (ECCV), pp. 135–150
20. Saito K, Yamamoto S, Ushiku Y, Harada T (2018) Open set domain adaptation by backpropagation. In: Proceedings of the European conference on computer vision (ECCV), pp. 153–168
21. Guo J, Yang J, Yue H, Liu X, Li K (2021) Unsupervised domain-invariant feature learning for cloud detection of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 60:1–15
22. Tasar O, Tarabalka Y, Giros A, Alliez P, Clerc S (2020) Standardgan: Multi-source domain adaptation for semantic segmentation of very high resolution satellite images by data standardization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 192–193
23. Tasar O, Tarabalka Y, Giros A, Alliez P, Clerc S (2020) Standardgan: Multi-source domain adaptation for semantic segmentation

of very high resolution satellite images by data standardization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 192–193
24. Li Y, Wang S, Wang B (2023) Dual teacher-student based separation mechanism for open set domain adaptation. Knowledge-Based Systems 272:110600
25. You K, Long M, Cao, Z, Wang J, Jordan MI (2019) Universal domain adaptation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2720–2729
26. Yin Y, Yang Z, Wu X, Hu H (2021) Pseudo-margin-based universal domain adaptation. Knowledge-Based Systems 229:107315
27. Kundu JN, Venkat N, Babu RV, et al (2020) Universal source-free domain adaptation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4544–4553
28. Xu R, Liu P, Wang L, Chen C, Wang J (2020) Reliable weighted optimal transport for unsupervised domain adaptation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4394–4403
29. Xu R, Liu P, Zhang Y, Cai F, Wang J, Liang S, Ying H, Yin J (2020) Joint partial optimal transport for open set domain adaptation. In: IJCAI, pp. 2540–2546
30. Cuturi M (2013) Sinkhorn distances: Lightspeed computation of optimal transport. Advances in neural information processing systems **26**
31. Fatras K, Séjourné T, Flamary R, Courty N (2021) Unbalanced minibatch optimal transport; applications to domain adaptation. In: International Conference on Machine Learning, pp. 3186–3197. PMLR
32. Nguyen K, Nguyen D, Pham T, Ho N et al (2022) Improving minibatch optimal transport via partial transportation. In: International Conference on Machine Learning, pp. 16656–16690. PMLR
33. Remmelzwaal L (2023) Object detection and tracking for crate and bottle identification in a bottling plant using deep learning. In: Artificial Intelligence and Applications, vol. 1, pp. 191–195
34. Tschannen M, Bachem O, Lucic M (2018) Recent advances in autoencoder-based representation learning. arXiv preprint arXiv:1812.05069
35. Hua Y, Mou L, Lin J, Heidler K, Zhu XX (2021) Aerial scene understanding in the wild: Multi-scene recognition via prototype-based memory networks. ISPRS Journal of Photogrammetry and Remote Sensing 177:89–102
36. Long S, Zhao Z, Pi J, Wang S, Wang J (2023) Beyond attentive tokens: Incorporating token importance and diversity for efficient vision transformers. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10334–10343
37. Xia Y, Yun LJ, Yang C (2023) Transferable adversarial masked self-distillation for unsupervised domain adaptation. Complex & Intelligent Systems 9(6):6567–6580
38. Flamary R, Courty N, Tuia D, Rakotomamonjy A (2016) Optimal transport for domain adaptation. IEEE Trans. Pattern Anal. Mach. Intell 1(1–40):2
39. Courty N, Flamary R, Tuia D (2014) Domain adaptation with regularized optimal transport. In: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2014, Nancy, France, September 15-19, 2014. Proceedings, Part I 14, pp. 274–289. Springer
40. Chizat L, Peyré G, Schmitzer B, Vialard FX (2018) Scaling algorithms for unbalanced optimal transport problems. Mathematics of Computation 87(314):2563–2609
41. Saito K, Kim D, Sclaroff S, Saenko K (2020) Universal domain adaptation through self supervision. Advances in neural information processing systems 33:16282–16292
42. Zou Q, Ni L, Zhang T, Wang Q (2015) Deep learning based feature selection for remote sensing scene classification. IEEE Geoscience and remote sensing letters 12(11):2321–2325
43. Xia GS, Hu J, Hu F, Shi B, Bai X, Zhong Y, Zhang L, Lu X (2017) Aid: A benchmark data set for performance evaluation of aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing 55(7):3965–3981
44. Yang Y, Newsam S (2010) Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pp. 270–279
45. Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: Benchmark and state of the art. Proceedings of the IEEE 105(10):1865–1883
46. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M et al (2015) Imagenet large scale visual recognition challenge. International journal of computer vision 115:211–252
47. Saito K, Kim D, Sclaroff S, Saenko K (2020) Universal domain adaptation through self supervision. Advances in neural information processing systems 33:16282–16292
48. Xu Q, Shi Y, Yuan X, Zhu XX (2023) Universal domain adaptation for remote sensing image scene classification. IEEE Transactions on Geoscience and Remote Sensing 61:1–15