Progressively Select and Reject Pseudolabeled Samples for Open-Set Domain Adaptation

Qian Wang ©, Fanlin Meng ©, Senior Member, IEEE, and Toby P. Breckon ©, Senior Member, IEEE

Abstract—Domain adaptation solves image classification problems in the target domain by taking advantage of the labeled source data and unlabeled target data. Usually, the source and target domains share the same set of classes. As a special case, open-set domain adaptation (OSDA) assumes there exist additional classes in the target domain but are not present in the source domain. To solve such a domain adaptation problem, our proposed method learns discriminative common subspaces for the source and target domains using a novel open-set locality preserving projection (OSLPP) algorithm. The source and target domain data are aligned in the learned common spaces classwise. To handle the open-set classification problem, our method progressively selects target samples to be pseudolabeled as known classes, rejects the outliers if they are detected as unknown classes, and leaves the remaining target samples as uncertain. The common subspace learning algorithm OSLPP simultaneously aligns the labeled source data and pseudolabeled target data from known classes and pushes the rejected target data away from the known classes. The common subspace learning and the pseudolabeled sample selection/rejection facilitate each other in an iterative learning framework and achieve state-of-the-art performance on four benchmark datasets Office-31, Office-Home, VisDA17, and Syn2Real-O with the average harmonic mean of open-set recognition accuracy (HOS) of 87.6%, 67.0%, 76.1%, and 65.6%, respectively.

Impact Statement—Traditional supervised machine learning methods require a large amount of labeled data for model training. In many real-world applications, however, collecting and labeling sufficient data is laborious and even impossible. One way to solve this problem is domain adaptation, which solves problems (e.g., image classification) in the target domain by taking advantage of the labeled data in a different but related domain (i.e., source domain). This article aims to address the OSDA problem in image classification. Our proposed approach relies on deep features and a feature transformation method so that it can not only achieve state-of-the-art performance on benchmark datasets but also is more practically useful due to its low training cost. Although we demonstrate the effectiveness of our proposed approach in the image classification tasks, we have found it also applicable in many other domains such as drug discovery and computational biology. We believe this work will benefit the broad community of artificial intelligence (AI) in a variety of applications.

Index Terms—Locality preserving projection, open-set domain adaptation (OSDA), pseudolabeling.

I. INTRODUCTION

ONE key to modern learning systems is the access to large-scale high-quality training data. Collecting and annotating a large amount of data for model training can be difficult and costly in some particular domains [1], [2], [3]. Given a target domain where the labeled training data are limited, one may promote the learning by exploiting annotated data from a source domain where annotated data are easier to access. For instance, object recognition from photos taken at night as a target domain task can be better addressed by exploiting more easily accessible photos taken in the daytime (i.e., source domain). However, the data distribution shift between the source and the target domains degrades the traditional transfer learning performance. To address the domain shift problem, domain adaptation techniques have been extensively studied in recent years. In particular, effective approaches have been proposed to solve the closed-set domain adaptation problems in which the source and target domains share the same set of classes [4], [5], [6], [7]. In many real-world applications, however, we may only have interest in a subset of the classes in the target domain and ignore the rest of them. In other cases, the annotated data in the source domain may not necessarily cover all classes in the target domain. Arising from such realistic scenarios, the open-set domain adaptation (OSDA) problem aiming to recognize the target-domain samples as one of the known classes (i.e., shared classes between the source and target domains), or the unified unknown class, has attracted much attention [8], [9], [10], [11], [12], [13], [14], [15].

Typical closed-set domain adaptation methods suffer from the negative transfer issue [16] when directly applied to the OSDA problems. Specifically, most closed-set domain adaptation methods take advantage of the prior knowledge that the source and target domains share the same set of classes and conditional distributions can be well aligned. With the existence of unknown classes in the target domain for OSDA, data belonging to these unknown classes will be mistakenly aligned with some known classes in the source domain. Many existing OSDA methods attempt to solve this problem by treating the unknown classes as a unified one and learning a classifier for $C + 1$ classes [9], [10], [12], [15]. Such a learning objective forces different
unknown classes to behave similarly (e.g., clustered compactly in the hidden representation space), increasing the difficulty of learning a performant classifier. We believe that such a learning object is not necessary for an OSDA problem, which aims to discriminate each known class from the rest. In contrast, we solve the OSDA problem by learning a common subspace where known classes are separated from each other as well as all the unknown classes, while the unknown classes are not necessarily compactly clustered.

To learn such a favorable common subspace from two domains, our proposed framework has two essential algorithms: 1) the algorithm of progressively selecting and rejecting pseudolabeled samples for unknown classes aware discriminative learning; and 2) the domain adaptation algorithm open-set locality preserving projection (OSLPP), our modified version of locality preserving projection (LPP) [17] for domain aligned common subspace learning. These two algorithms facilitate each other during favorable subspace learning for optimal OSDA.

The first algorithm is adapted from a typical UDA approach [16] by adding an unknown class sample rejection process in each iteration. Specifically, in [16], the target domain data are pseudolabeled as the known classes and progressively selected as supervision for the domain adaptation process. In our proposed method, to enable unknown classes aware discriminative learning, we additionally reject some pseudolabeled target samples as unknown classes if these samples are far away from all the known classes so that the unknown classes can be considered during common subspace learning using OSLPP. The remaining target samples are treated as uncertain and hence will not contribute to the domain alignment. It is important to select and reject the pseudolabeled samples progressively, so that the negative effect caused by incorrect pseudolabeling can be mitigated. To make the most of the pseudolabeled target samples, the numbers of selected and rejected target samples are monotonically increasing until no samples are remaining uncertain.

The second algorithm aims at aligning the source and target domains in a learned common subspace in which the source and target data from the known classes are aligned classwise, while the target domain data pseudolabeled as unknown classes are pushed away from the known classes. Inherited from the original LPP algorithm, the OSLPP has the capability of local structure preserving, which is important for model generalization from the source domain to the target domain. Different from the original LPP, OSLPP is a supervised learning algorithm taking advantage of the labeled source data and the pseudolabeled target data. OSLPP also differs from the supervised LPP used in [16] by taking into consideration the estimated unknown classes when learning the projection. In addition, the relatively loose constraints of the OSLPP algorithm allow the unknown classes distributed in a spacious region while separated from known classes. The OSLPP-based subspace learning and pseudolabeling are conducted alternately and repeated for a fixed number of iterations. These two algorithms facilitate each other iteratively, and the two domains are well aligned in the learned subspace where the recognition performance of target domain data is enhanced.

The contributions of this work can be summarized as follows.

1) A novel framework is proposed for open set domain adaptation by learning a common subspace from both source and target domains using OSLPP, a novel algorithm aiming at aligning data from known classes and pushing away data from unknown classes.
2) An algorithm of progressively selecting and rejecting pseudolabeled target domain data is proposed to facilitate the domain adaptation algorithm.
3) Experiments are conducted on four commonly used datasets Office-31, Office-Home, VisDA17, and Syn2Real-O. The experimental results demonstrate our proposed method can achieve or outperform state-of-the-art performance.
4) We empirically demonstrate that hyperparameter values can tradeoff the accuracy of classifying known classes and the accuracy of detecting unknown classes in real-world OSDA problems.

II. RELATED WORK

In this section, we review existing related work in domain adaptation (including closed-set domain adaptation, partial domain adaptation, and universal domain adaptation), open-set recognition (OSR), and OSDA.

A. Domain Adaptation

Domain adaptation is a general technique being used to address various research problems. From the perspective of tasks, it can be applied to image classification [10], [18], object detection [19], image segmentation [20], person reidentification [21], [22], etc. From the perspective of supervision in the target domain, it can be categorized into unsupervised, semisupervised, or supervised domain adaptation although most existing works including ours focus on the unsupervised setting. From the perspective of label space overlapping between two domains, it has been formulated into closed-set domain adaptation where the label spaces are equivalent between two domains, partial...
domain adaptation (PDA) where the source domain contains extra classes [23], [24], [25]. OSDA which is our focus in this work, and universal domain adaptation where a versatile approach is expected to solve the problem without knowing it is a closed-set, open-set or partial domain adaptation problem [26], [27], [28], [29], [30], [31].

Focusing on the image classification tasks, the OSDA problem we attempt to address is unsupervised in the sense that there is no labeled target data. Our work is inspired by the unsupervised domain adaptation method proposed in [16]. Using a similar iterative learning framework, our method is dedicated to OSDA problems by detecting the unknown classes in the target domain and exploiting them to facilitate the common subspace learning for domain adaptation.

### B. OSDA

Existing OSDA methods borrow the successful ideas from unsupervised domain adaptation approaches and adapt them for OSDA by handling the unknown target samples in specific ways. These methods are distinct from each other in how the unknown target samples are detected and utilized. Saito et al. [8] learn a classifier to classify target samples into \( C + 1 \) classes \((C \text{ known classes and one unknown class})\) and use such pseudolabels to construct the loss function for classifier training. Pan et al. [32] try to explore the structure of the target data by self-ensembling with category-agnostic clusters (SE-CC) to improve the recognition of unknown classes in the target domain. Liu et al. [9] take one step further by progressive adaptation with selected pseudolabeled samples in the target domain. As this progressive learning algorithm has been proven effective for domain adaptation [16], our work adopts a similar progressive learning algorithm but utilizes novel strategies for data selection/rejection and OSLPP for domain alignment. Following the same direction, Fang et al. [12] use the samples classified as the unknown classes in the so-called open set difference loss term to enhance the ability to recognize the unknown classes of the learned classifier. However, this method introduces too many hyperparameters, making it difficult to use in practice.

Most aforementioned methods have coupled unknown class detection and domain adaptation modules; however, exceptions exist that address the OSDA problem in two stages [10], [33]. In the first stage of [10], a classifier for separating known and unknown classes is learned using rotation-based self-supervised learning. In the second stage, the source samples together with the detected known target samples are combined to train a classifier to classify target samples into either one of the known classes or the unknown class. The limitation of this method is that the two stages are trained sequentially; hence, the performance of the second stage relies on the accuracy of known–unknown class separability in the first stage and may lead to a suboptimal solution. Kundu et al. [33] aim at solving OSDA in a special scenario where source data are separated from the target data. Instead of training a model using the combination of source and target domain data, an inheritable model is first trained with the source data and subsequently adapted for the target data. The unknown classes are recognized by measuring the instance-level inheritability. In addition, the samples confidently pseudolabeled as unknown classes are used as target-domain supervision. Our work follows a similar idea of selecting the most confident pseudolabels as supervision information from the target domain, and such information is expected to better cluster the known classes and discriminate unknown classes far away from the known classes. The difference lies in that we use a novel manifold learning method OSLPP as opposed to the neural networks employed in [33] and achieve superior or comparable performance.

### III. Problem Formulation

Suppose we have a labeled dataset \( D^s = \{(x^s_i, y^s_i)\}, i = 1, 2, \ldots, n_s \) from the source domain \( S \), \( x^s_i \in \mathbb{R}^{d_0} \) represents the feature vector of \( i \)th labeled sample in the source domain, \( d_0 \) is the feature dimension and \( y^s_i \in \mathbb{Y}^s \) denotes the corresponding label. OSDA aims at classifying an unlabeled dataset \( D^t = \{x^t_i\}, i = 1, 2, \ldots, n_t \) from the target domain \( T \), where \( x^t_i \in \mathbb{R}^{d_0} \) represents the feature vector in the target domain. The target label space \( \mathbb{Y}^t \) is a union of the source label space \( \mathbb{Y}^s \) and an unknown label space \( \mathbb{Y}^\text{unk} \) (i.e., \( \mathbb{Y}^\text{unk} = \mathbb{Y}^t \setminus \mathbb{Y}^s \neq \emptyset \)). The classes shared by the source and target domains are known classes, while the rest of the target classes are unknown classes (\( \mathbb{Y}^\text{unk} \)). Samples from the target domain are expected to be classified as one of the known classes or the unified unknown class. It is assumed that both the labeled source domain data \( D^s \) and the unlabeled target domain data \( D^t \) are available during model training. To clarify, we will use the terms “select” and “reject” to represent the actions of confidently classifying/pseudolabeling a target sample into a known class or the unknown class, respectively. The rest target samples that are neither “selected” nor “rejected” are denoted as “uncertain.”

### IV. Method

We introduce our proposed approach to the OSDA problems in this section. The framework of the approach is first described and illustrated in Fig. 2. The original version of the LPP algorithm is briefly described to make the article self-contained. Subsequently, we give the details of two key algorithms in the framework: OSLPP and pseudolabel selection and rejection. Finally, the algorithm is summarized in Algorithm 1 and the notations are listed in Table I.

#### A. Overview

Our approach to OSDA is inspired by the existing approaches to unsupervised domain adaptation [5], [16], [34] and uses a
In the subspace learned by (c), class means are computed and used to selectively assign pseudolabels to the known classes; the target samples are rejected if their nearest neighbor is already rejected. Steps (c) and (d) are repeated for selectively assign pseudolabels to the target samples; on the other hand, target samples far away from all class means are rejected as unknown classes. Our proposed OSLPP is applied to the labeled source samples, selected pseudolabeled target samples, and rejected target samples so that the samples from the same class regardless of their domains are pulled close to each other while those rejected as unknown classes are pushed away from the known classes. In the subspace learned by (c), class means are computed and used to selectively assign pseudolabels to the known classes; the target samples are rejected if their nearest neighbor is already rejected. Steps (c) and (d) are repeated for times.

Algorithm 1 Open-Set LPP

Input: Labeled source data set \( \mathcal{D}^s = \{ (x_i^s, y_i^s) \}_i \), \( n_s \) and unlabeled target data set \( \mathcal{D}^t = \{ x_i^t \}_i \), \( n_t \), dimensionality of the subspace \( d \), number of iteration \( T \), number of initial rejected target samples \( n_r \).

Output: The projection matrix \( P \) and predicted labels \( \{ \hat{y}_i \} \) for target samples.

1: Initialize \( k = 0 \);
2: Learn the projection \( P_0 \) using only source data \( \mathcal{D}^s \);
3: Assign pseudo labels for all target data using Eq. (7-8);
4: Initialize \( n_r \) rejected samples of the lowest pseudo-labelling probabilities;
5: while \( k < T \) do
6: \( k \leftarrow k + 1 \);
7: Select a subset of pseudo-labelled target data \( S_k \subset \mathcal{D}^t \);
8: Reject samples using 1-Nearest-Neighbour and enrich the subset of rejected samples \( R_k \subset \mathcal{D}^t \);
9: Learn \( P_k \) using \( \mathcal{D}^s, S_k \) and \( R_k \);
10: Update pseudo labels for all target data using Eq.(7).
11: end while

similar iterative learning framework. The goal is to learn a common subspace based on the labeled source samples [c.f. Fig. 2(a)] and pseudolabeled target samples [c.f. Fig. 2(b)], so that the source and target domains can be well aligned in the learned subspace and the target samples can be recognized by a simple nearest neighbor method [c.f. Fig. 2(c)]. It has been proved that selective pseudolabeling performs better than the methods considering all the pseudolabeled target samples without the selection [16]. The subspace learning procedure and the selective pseudolabeling facilitate each other during the learning. Distinct from existing works, we propose a new subspace learning method OSLPP, an extension of LPP to the OSDA problems, and a novel selection–rejection algorithm to allow the OSLPP to learn a better subspace where known classes are aligned and unknown classes are pushed away [c.f. Fig. 2(d)].

B. Locality Preserving Projection

To make the article self-contained, we briefly describe the original LPP algorithm [17] based on which we propose our OSLPP in the next subsection. LPP aims at learning a favorable low-dimensional subspace where the local structures of data in the original feature space can be well preserved. To reduce the computational complexity, an unsupervised dimensionality reduction algorithm principal component analysis (PCA) [35] is first applied to the features. Suppose \( x_i \in \mathbb{R}^{d_0} \) and \( x_j \in \mathbb{R}^{d_0} \) are two data points in the PCA-reduced feature space, LPP aims at learning a projection matrix \( P \in \mathbb{R}^{d_0 \times d} \) \((d << d_0)\) so that data points close to each other in the original space will still be close in the projected subspace. The objective of LPP [16], [17] can be formulated as

\[
\min_P \sum_{i,j} ||P^T x_i - P^T x_j||^2 W_{ij}
\]

where \( W \) is the similarity matrix of the graph constructed by all the data points. According to [17], the edges of the graph can be created by either \( \epsilon \)-neighborhoods or \( k \)-nearest neighbors. The edge weights can be determined by the heat kernel \( W_{ij} = e^{-||x_i - x_j||^2/ \tau} \) or the simple binary assignment (i.e., all edges have the weights of 1). Note that LPP is an unsupervised learning method without the need for labeling information. In the following subsection, we will describe how to extend the LPP algorithm to solve the OSDA problems where there exist unknown classes in the target domain.

C. Open-Set LPP

Open set LPP aims at exploring the structural information underlying the labeled source data and the pseudolabeled target data including those pseudolabeled as known classes and unknown classes. We denote the combined labeled source and pseudolabeled target samples as \( X = \{ x_1^s, \ldots, x_{n_s}^s, x_1^t, \ldots, x_{n_t}^t \} \in \mathbb{R}^{d_0 \times (n_s + n_t)} \) where \( n_s \) is the number of pseudolabeled target samples. To simplify the notation, we omit the superscript and assign the subscripts from 1 to \( n_s + n_t \) for the
combined source and target samples, then the objective becomes the same as (1) but with a different way of constructing the similarity matrix $W \in \mathbb{R}^{(n_s+n_t) \times (n_s+n_t)}$

$$W_{ij} = \begin{cases} 1, & y_i = y_j \\ 0, & \text{otherwise} \end{cases}$$ (2)

where $y_i$ denotes either the ground-truth label of $x^s$ from the source domain, or the pseudolabel of $x^t$ from the target domain. It is noteworthy that the pseudolabels in (2) can be one of the known classes or the unified unknown class. Before the last iteration when all target samples are confidently recognized, only the confidently selected or rejected samples will be considered in (2) and the rest of the target samples will be treated as uncertain samples and have 0 similarity with all the other samples.

By optimizing the objective (1) with the similarity matrix defined as (2), the samples labeled or pseudolabeled as the same class will be projected to be close to each other regardless of which domain they are from. The samples from different classes will be implicitly separated in the learned subspace. Target samples rejected as unknown classes are treated as uncertain samples and have 0 similarity with all the known classes. Adding terms to the objective explicitly pushing the rejected samples far away from known classes implicitly. Adding terms to the objective explicitly pushing the rejected samples far away from known classes does not make a difference in our empirical study, e.g.,

$$\min_p \sum_{i,j} (||P^T x_i - P^T x_j||_2^2 W_{ij} - ||P^T x_i - P^T x_j||_2^2 H_{ij})$$ (3)

where the element $H_{ij} = 1$ if $x_i$ and $x_j$ are from known and unknown classes, respectively, and $H_{ij} = 0$ otherwise. This variation will not be further discussed in this article.

According to [17], optimizing the objective (1) is equivalent to solving the generalized eigenvalue problem

$$XDX^T p = \lambda (XLX^T + 1) p$$ (4)

where $L = D - W$ is the Laplacian matrix, $D$ is a diagonal matrix with $D_{ii} = \sum_j W_{ij}$, and the regularization term $tr(P^T P)$ is added for penalizing extreme values in the projection matrix $P$. Solving the generalized eigenvalue problem gives the optimal solution $P = [p_1, ..., p_d]$ where $p_1, ..., p_d$ are the eigenvectors corresponding to the largest $d$ eigenvalues.

In summary, the key of OSLPP is to treat the target samples differently when constructing the similarity matrix $W$ based on whether they are selected, rejected, or uncertain. In the following subsection, we will describe how to make such decisions for the target samples.

D. Pseudolabeled Sample Selection and Rejection

Once the projection matrix $P$ is learned, all the samples can be projected into the common subspace by

$$z^{s/t} = P^T x^{s/t}.$$ (5)

To enhance the separability of different classes, we apply zero-mean centering and L2 normalization to each projection $z$ so that it will be distributed on the surface of the unit hypersphere. The pseudo labeling can be done in the subspace by the nearest class mean (NCM) method [16]. Alternatively, one can train a parametric model (e.g., support vector machines and neural networks) for pseudolabelling. Our choice of NCM is inspired by its simplicity and effectiveness validated in prior works [16], [23], [36]. The class means are computed over all the labeled source samples and selected pseudolabeled target samples

$$\bar{z}_c = \frac{1}{n_c^e + \bar{n}_c} \left( \sum_{i \in x^{c}_c} z_i + \sum_{i \in \tilde{x}_c} z_i \right)$$ (6)

and the pseudolabel for a given target sample $x^t$ can be predicted as

$$\hat{y}^t = \arg\min_c p(\tilde{y}^t, c)$$ (7)

with the probability

$$p(\tilde{y}^t, c) = \frac{e^{-\text{dist}(\tilde{x}^t, \bar{z}_c)}}{\sum_{c'} e^{-\text{dist}(\tilde{x}^t, \bar{z}_{c'})}}$$ (8)

where $\text{dist}(a, b)$ is the Euclidean distance between $a$ and $b$.

In the $k$th ($k = 0, 1, ..., T$) iteration of learning, we select the top $\{k + 1\}/T$ of the target samples pseudolabeled as cth class for each class $c = 1, 2, ..., |C|$. These selected samples together with their corresponding pseudolabels will be used in the next iteration of subspace learning. To reject target samples as unknown classes, we use the 1-nearest-neighbour algorithm. Any target sample will be rejected if its nearest neighbor (excluding the samples neither selected nor rejected) in the subspace is a rejected sample. In the first iteration of learning, $n_s$ seed samples are rejected if their pseudolabeling probabilities are top $n_t$ lowest, where $n_t$ is a hyperparameter. The set of rejected samples will be enriched in each iteration without replacement. As a result, the value of $n_t$ will be hyperparameter trading off the accuracy of known classes and the unified unknown class.

Specifically, in the first iteration (i.e., $k = 0$), the projection matrix $P$ is learned with source data only. Subsequently, $1/T$ of the target samples are selected for the second iteration. In the second last iteration (i.e., $k = T - 1$), all target samples will be selected or rejected for the learning of $P$ in the final iteration. In the final iteration, we will get our final results without the need to select or reject. We summarize the proposed approach in Algorithm 1.

E. Computational Complexity

The complexity of PCA is $O(dn_t^2 + d_n^3)$. The complexity of OSLPP is $O(n_s^2dPCA + d_n^3)$, which is repeated for $T$ times and leads to approximately $O(T(n_s^2dPCA + d_n^3))$. In our experiments, we use a laptop with an Intel Core i5-7300HQ CPU and 32G memory RAM. Running 6 tasks of the Office31 dataset takes approximately 14 seconds and running 12 tasks of the Office-Home dataset takes approximately 10 min. For the VisDA17 and Syn2Real-O datasets, due to the large number of samples, it takes approximately 21 min to finish one task. This is much more efficient than the methods requiring the end-to-end training of a deep model, which usually takes hours on a graphics processing unit (GPU).

Regarding memory usage, when the number of samples $n = n_s + n_t$ is much greater than the dimensionality, the memory
complexity is $O(n^2)$. As a result, our method has the limitation of scaling up to extremely large datasets (e.g., $n > 100,000$) for which the neural networks-based approaches can be better choices.

V. EXPERIMENTS AND RESULTS

In this section, we demonstrate the experiments and results for validating the effectiveness of the proposed method. Specifically, we introduce the datasets, experimental settings, evaluation metrics, and experimental results.

A. Datasets

Four commonly used datasets for OSDA were employed in our experiments: Office-31 and Office-Home. Office31 [37] consists of three domains: Amazon (A), Webcam (W), and DSLR (D). There are 31 common classes for all three domains containing 4110 images in total. Following the open-set protocol employed in [8] and [10], we use the first ten classes in alphabetic order as the shared known classes in both source and target domains and the last 11 classes as the unknown classes in the target domain. Image features are extracted by the ResNet50 [38] model pretrained on ImageNet [39] without fine-tuning on the Office31 dataset. Office-Home [40] consists of four different domains: Artistic images (Ar), Clipart (Cl), Product images (Pr), and Real-World images (Rw). There are 65 object classes in each domain with a total number of 15,588 images. We follow [10] and use images from the first 25 classes in alphabetical order as the shared known classes in both domains and images from the remaining 40 classes as unknown classes in the target domain. Image features are extracted by the ResNet50 [38] model pretrained on ImageNet [39] without fine-tuning on the Office-Home dataset. VisDA17 [41] contains 12 categories in two domains. The Synthetic domain consists of 152,397 synthetic images from the train set, and the Real domain contains 55,388 real-world images from the validation set. We follow the setting in [8] using six categories as the known classes and the remaining six categories as the unknown classes. Syn2Real-O [42] (or VisDA18) is constructed from the VisDA17 and significantly increases the openness to 0.9 by adding additional unknown samples in the target domain. Following the official setting, the source domain contains 12 categories in two domains and images from the remaining 152,397 synthetic images from the train set, and the target domain consists of 152,397 synthetic images from the validation set.

B. Implementation Details

We implement the proposed method in MATLAB R2020b\textsuperscript{1}. We extract deep features from the penultimate layer of the pretrained models without fine-tuning them on the source data. The 2048-dimensional ResNet50 features or 4096-dimensional VGG16 features are first $l_2$ normalized [16] and the dimensionality is reduced by PCA [35] to 16, 512, 64, and 256 for Office31, Office-Home, VisDA17, and Syn2Real-O datasets, respectively. Applying PCA to the original high-dimensional deep features not only can reduce the computation cost of OSLPP but also benefit the performance of our experiments. The dimensionality $d$ of the learned subspace is set as 16, 128, 64, and 256, respectively. The number of iterations is set to ten for all datasets. The number of initially rejected samples $n_r$ is set to 140, 1200, 200, and 14,000, respectively. The sensitivity of our method to these hyperparameters will be discussed later in this section. To be noted, we only use 10\% of the source samples for the VisDA17 and Syn2Real-O datasets to reduce memory usage by our algorithm. The final results are the average over ten trials of experiments in which each trial uses 10\% evenly sampled source samples and all the target samples.

C. Evaluation Metrics

There exist several evaluation metrics for evaluating OSDA approaches [10]. OS is the mean per-class accuracy over all target domain images from known classes and unknown classes as one unified class. To measure the capabilities of recognizing the known and unknown classes, we use the metrics OS* which is the mean per-class accuracy over the shared known classes and UNK which is the accuracy of images from the unknown classes (as one unified class)

$$
\text{OS} = \frac{1}{|C_s| + 1} \sum_{i=1}^{|C_s|+1} \frac{|x : x \in D_t^i \land \hat{y}(x) = i|}{|x : x \in D_t^i|} \tag{9}
$$

$$
\text{OS}^* = \frac{1}{|C_s|} \sum_{i=1}^{|C_s|} \frac{|x : x \in D_t^i \land \hat{y}(x) = i|}{|x : x \in D_t^i|} \tag{10}
$$

$$
\text{UNK} = \frac{|x : x \in D_t^{unk} \land \hat{y}(x) = unk|}{|x : x \in D_t^{unk}|} \tag{11}
$$

where $|C_s|$ is the number of source domain classes (i.e., the number of shared known classes) in the class space $C_s$, $D_t^i$ denotes the dataset of $i$th class in the target domain, and $\hat{y}(x)$ is the predicted label of data sample $x$.

OS is a combination of OS* and UNK as $\text{OS} = (|C_s|/|C_s| + 1) \times \text{OS}^* + (1/|C_s| + 1) \times \text{UNK}$; however, as pointed out in [10], OS can be dominated by the accuracy of known classes since the unknown classes are treated as one unified class. One effective evaluation metric properly balancing the recognition performance for known and unknown classes is the harmonic mean of OS* and UNK

$$
\text{HOS} = \frac{2 \times \text{OS}^* \times \text{UNK}}{\text{OS}^* + \text{UNK}}. \tag{12}
$$

In most of our experiments, we report OS*, UNK, and harmonic mean of open-set recognition accuracy (HOS) for individual domain adaptation tasks as well as the average performance over all possible tasks for a dataset. In the ablation study, we report one extra evaluation metric $\text{ALL}$, which is the instancewise accuracy of all target samples

$$
\text{ALL} = \frac{\sum_{i=1}^{|C_s|+1} |x : x \in D_t^i \land \hat{y}(x) = i|}{\sum_{i=1}^{|C_s|+1} |x : x \in D_t^i|}. \tag{13}
$$

\textsuperscript{1}Code is available from https://github.com/hellowangqian/oslpp
TABLE II

OPEN-SET CLASSIFICATION RESULTS (%) ON THE VISDA17 DATASET USING EITHER RESNET50 FEATURES OR RESNET50-BASED DEEP MODELS

| Method          | OS* | AU* | UNK* | OS* | AU* | UNK* | OS* | AU* | UNK* | OS* | AU* | UNK* |
|-----------------|-----|-----|------|-----|-----|------|-----|-----|------|-----|-----|------|
| SynAdapt [9]    | 91.0| 63.9| 73.0 | 86.7| 61.6| 73.9 | 81.1| 65.9| 71.2 | 94.1| 55.5| 67.8 |
| OSBP [8]        | 90.5| 75.5| 82.4 | 86.8| 82.1| 75.4 | 92.4| 96.7| 97.2 | 73.0| 74.4| 73.7 |
| UAN [26]        | 95.6| 24.6| 35.7 | 95.5| 31.0| 45.5 | 93.5| 53.4| 68.0 | 99.8| 52.5| 68.5 |
| SE-CCI [32]     | 84.0| 46.6| 59.9 | 84.2| 64.4| 73.0 | 90.3| 12.2| 21.5 | 96.6| 55.9| 70.8 |
| ROS [10]        | 87.5| 77.8| 82.4 | 88.4| 76.7| 82.1 | 74.8| 81.2| 79.0 | 99.3| 93.0| 96.6 |
| OSLPP (ours)    | 92.6| 90.4| 91.3 | 89.2| 88.8| 89.0 | 82.2| 77.1| 79.3 | 96.9| 84.4| 92.7 |

Note: Bold values represent the best results and underlined values represent the second-best results.

TABLE III

OPEN-SET CLASSIFICATION RESULTS (%) ON HOME-DATASET USING EITHER RESNET50 FEATURES OR RESNET50-BASED DEEP MODELS

| Method          | OS* | AU* | UNK* | OS* | AU* | UNK* | OS* | AU* | UNK* |
|-----------------|-----|-----|------|-----|-----|------|-----|-----|------|
| SynAdapt [9]    | 93.3| 55.1| 68.7 | 93.4| 55.1| 68.7 | 93.4| 55.1| 68.7 |
| OSBP [8]        | 92.1| 66.4| 71.1 | 92.1| 66.4| 71.1 | 92.1| 66.4| 71.1 |
| UAN [26]        | 78.8| 78.8| 78.8 | 78.8| 78.8| 78.8 | 78.8| 78.8| 78.8 |
| SE-CCI [32]     | 80.3| 80.3| 80.3 | 80.3| 80.3| 80.3 | 80.3| 80.3| 80.3 |
| ROS [10]        | 81.0| 81.0| 81.0 | 81.0| 81.0| 81.0 | 81.0| 81.0| 81.0 |
| OSLPP (ours)    | 83.5| 83.5| 83.5 | 83.5| 83.5| 83.5 | 83.5| 83.5| 83.5 |

Note: Bold values represent the best results and underlined values represent the second-best results.

TABLE IV

OPEN-SET CLASSIFICATION RESULTS (%) ON THE VISDA17 DATASET USING VGG16 AND THE SYN2REAL-O-DATION USING RESNET50

| Method          | OS* | AU* | UNK* | OS* | AU* | UNK* | OS* | AU* | UNK* |
|-----------------|-----|-----|------|-----|-----|------|-----|-----|------|
| SynAdapt [9]    | 92.1| 66.4| 71.1 | 92.1| 66.4| 71.1 | 92.1| 66.4| 71.1 |
| OSLPP (ours)    | 93.5| 55.1| 68.7 | 93.5| 55.1| 68.7 | 93.5| 55.1| 68.7 |

Note: Bold values represent the best results and underlined values represent the second-best results.

D. Comparison With State-of-the-Art

The proposed method OSLPP is compared against recent state-of-the-art approaches to OSDA problems. The approaches we compare against include STA [9], OSBP [8], UAN [26], [32],[33], and [44], since they either use a different experimental protocol or different deep features in their experiments and make it difficult for a direct comparison with those considered in our work.

To make a fair and reliable comparison, we report the results of state-of-the-art methods reproduced by [10] for STA, OSBP, UAN, and ROS. For DAOD and PGL, the metrics UNK and OS* are calculated by (9)–(12) based on the reported OS and OS* values in their original papers.

As we can see in Table II, our proposed method achieves the best average HOS of 87.6% followed by 85.9% by [10] and 83.7% by [8]. Among the six adaptation tasks, our method performs the best in terms of HOS on four of them. The other methods usually perform well in recognizing the known classes (high OS*) but are bad at recognizing the unknown classes (low UNK). This is partially due to the bias caused by only considering the evaluation metrics of OS and OS*. A high OS* along with a low UNK means the method mistakenly classifies a large number of samples from unknown classes as one of the known classes. This caveat cannot be disclosed when using only OS* as the evaluation metrics.

On the more challenging Home-Dataset, as shown in Table III, our method achieves the best or the second best HOS on all 12 tasks and the best average HOS of 67.0% followed by 66.2% by [10] and 64.7% by [8]. The experimental results on both datasets are consistent in that methods achieving high OS* do not necessarily perform well in practice since they may make too many mistakes on the target samples from unknown classes as we can see from the results of UAN [26] and PGL [11].

On the VisDA17 and Syn2Real-O-datasets, our method also achieves the best performance in terms of HOS as shown in Table IV. PGL [11] performs the best in recognizing known classes but suffers a lower UNK. This drawback becomes obvious when the openness is high as in the Syn2Real-O-dataset for which the overall recognition accuracy of PGL is 51.5% [11] while our method achieves 69.4%.

E. Comparison With Universal Domain Adaptation Approaches

We compare our method with universal domain adaptation approaches on the OSDA setting. The universal domain adaptation approaches are usually more powerful with more complex modules designed and integrated. The results of HOS on three datasets are shown in Table V from which we can see our proposed OSLPP can achieve comparable or superior performance to most universal domain adaptation approaches. Although...
TABLE V
COMPARISON WITH SOTA UNIVERSAL DOMAIN
ADAPTATION APPROACHES

| Method          | Office31 | OfficeHome | VisDA17 | Average |
|----------------|----------|------------|---------|---------|
| DCC [29]       | 72.6     | 61.7       | 59.6    | 64.6    |
| DANCE [27]     | 79.8     | 63.0       | 67.5    | 70.1    |
| OvaNet [30]    | 91.7     | 64.0       | 66.1    | 73.9    |
| GATE [45]      | 89.5     | 69.1       | 70.8    | 76.5    |
| OSLPP (ours)   | 87.6     | 67.0       | 76.1    | 76.9    |

Note: Bold values represent the best results and underlined values represent the second-best results.

none of the investigated methods including ours performs the best on all three datasets, our proposed method achieves the best average performance with on HOS of 76.9%. In addition, our approach is more computationally efficient as it works on top of deep features in contrast to GATE [45] and OvaNet [30] both of which involve training a deep CNN model.

F. Effect of Hyperparameters

We investigate the effect of four hyperparameters in our method: the dimensionality of the PCA subspace $d_{PCA}$, the dimensionality of the OSLPP subspace $d$, the number of initially rejected samples $n_r$, and the number of iterations $T$. To these ends, we set the investigated hyperparameter to the values within a predefined set and the others fixed as the default in our experiments. The experiments are conducted on the Office31 and Office-Home datasets. The investigated values for four hyperparameters are $d_{PCA} \in \{1024, 512, 256, 128, 64, 32, 16, 8\}$, $d \in \{512, 256, 128, 64, 32, 16, 8\}$, $n_r \in \{40, 60, 80, 100, 120, 140, 160, 180, 200, 220\} \times 10$ for the Office-Home dataset since this dataset contains much more samples), and $T \in \{6, 8, 10, 12, 14, 16, 18, 20\}$, respectively.

The results are shown in Fig. 3. We report the average OS*, UNK, and HOS for each experiment. The optimal number of PCA dimensionality for Office31 is 16, and a smaller or greater value will lead to a slight performance drop for both OS* and UNK. This result means the ResNet50 features are discriminative enough to separate ten shared known classes in this simple dataset, so that the unsupervised PCA can extract the most useful information in the first 16 principal components while more dimensions hurt the performance slightly. Similarly, the optimal dimensionality of OSLPP subspace is 16 for Office31, and more dimensions lead to slightly worse performance. For the more challenging Office-Home dataset with 25 shared known classes and 40 unknown classes, the average HOS is less sensitive to the subspace dimensionality and optimal average HOS can be achieved within a large range of values for $d_{PCA}$ (i.e., $32$–$1024$) and $d$ (i.e., $32$–$512$). In addition, the values of $d_{PCA}$ and $d$ affect the tradeoff between the OS* and UNK although their harmonic mean HOS is marginally affected.

The effects of $n_r$ and $T$ are more understandable and consistent on the two datasets. On the one hand, the performance of our method in terms of HOS is not sensitive to these two hyperparameters given that the optimal HOS can be achieved with a large range of values for $n_r$ (i.e., $40$–$180$ for Office31 and $400$–$1800$ for Office-Home) and $T$ (i.e., $8$–$20$). On the other hand, both the $n_r$ and $T$ control the tradeoff between the recognition accuracy of known classes OS* and the accuracy of unknown classes UNK. Increasing the number of initially rejected samples $n_r$ will lead to more samples recognized as unknown classes hence an increased UNK. Increasing the number of iterations $T$ means selecting samples as known classes more slowly while the pace of rejecting samples as unknown classes is not affected. As a result, more samples will be rejected and recognized as unknown classes after more iterations hence a higher UNK can be achieved. Along with the improvement of UNK, our method suffers from the decrease of OS* although their harmonic mean HOS is stable.

Overall, our proposed method is less sensitive to hyperparameters. In addition, we can tradeoff the recognition accuracy of known classes and unknown classes by adjusting the hyperparameters of $n_r$ and $T$ according to the requirements in practice.

G. Ablation Study

An ablation study is conducted on the VisDA17 dataset by removing different components of the proposed method. Specifically, we remove the component of PCA (denoted as “w/o PCA”), the component of OSLPP (denoted as “w/o OSLPP”), and the component of pseudolabel selection (denoted as “w/o selection”), respectively. We also replace the OSLPP algorithm with LPP (i.e., without considering the target samples rejected as unknown classes) and denote it as “w/o rejection.” The experimental results are shown in Table VI. Overall, our proposed method with the OSLPP algorithm performs the best in all four benchmark datasets. Replacing the OSLPP with the conventional LPP (i.e., w/o rejection) leads to a marginal performance drop due to the lower UNK. Such a difference in performance is accounted for by the fact that OSLPP considers the unknown classes as a unified one during the discriminative subspace learning while LPP does not. The worst HOS is obtained by the method “w/o selection” in which the pseudolabeled target samples are not progressively selected but overconfidently included for the subspace learning in the next iteration. As a result, the known classes are recognized with a better chance (i.e., high OS*) at the cost of recognition accuracy of unknown classes (i.e., extremely low UNK). When the OSLPP is removed from our framework (i.e., w/o OSLPP), the recognition is done in the PCA subspace and the performance is significantly worse than our proposed method with OSLPP. Similarly, the removal of PCA also causes a significant performance drop when compared with our proposed approach.

H. The Effect of the Number of Source Samples

We conduct experiments to investigate how the number of source samples affects the performance of our method. To this end, we evenly select 50%, 20%, 10%, 5%, 2%, or 1% of the source samples in the experiments and compare the results against the baseline when 100% (or 10% for the VisDA17 and Syn2Real-O datasets) of the source samples are used. The experimental results on four datasets are shown in Tables VII–IX.

For the Office31 dataset, a significant performance drop can be observed when the fraction of source data decreases to 10%.
The performance drop is mainly caused by the adaptation tasks where the domain D or the domain W servers as the source domain. Note that the total numbers of samples in these two domains (i.e., D and W) are 498 and 795, respectively. A fraction of 10% samples means there could exist some classes with zero samples in the source data. Similar phenomena can be seen in Table VIII. It can be concluded that our proposed method can achieve comparably good performance even if only a very small fraction (e.g., 10%) of source samples are used.

For the VisDA17 dataset, using less than 10% source-dominant data surprisingly gives better results as shown in Table IX due to the better recognition accuracy of unknown classes. On the other hand, using less than 10% source-dominant data gives consistently worse performance for the Syn2Real-O dataset due to the decreasing accuracy of recognizing known classes (i.e., OS*).

### I. Openness Analysis

We follow [10] to make an empirical openness analysis on the Office-Home dataset. Specifically, we set 40 (class IDs: 1–40, 16–55, 26–65), 25 (class IDs: 1–25, 26–50, 51–65), and 10 (class IDs: 1–10, 11–20, 21–30) samples as the source samples, respectively. The results are presented in Table X.
TABLE XI
OPENNESS ANALYSIS ON THE OFFICE-HOME DATASET

| Method     | 40 Known Classes | 25 Known Classes |
|------------|------------------|------------------|
|            | OS* UNK OS HOS   | OS* UNK OS HOS   |
| STAum [9]  |                  |                  |
| OSBP [8]   |                  |                  |
| ROS [10]   |                  |                  |
| OSLPP (ours) |                |                  |
|            |                  |                  |
| STAum [9]  |                  |                  |
| OSBP [8]   |                  |                  |
| ROS [10]   |                  |                  |
| OSLPP (ours) |                |                  |

Note: Bold values represent the best results and underlined values represent the second-best results.

TABLE XI
COMPARISON OF OPEN-SET RECOGNITION PERFORMANCE (HOS) WITH AND WITHOUT THE DOMAIN SHIFT

| Target Domain → | Pr | Rw |
|-----------------|----|----|
| Source Domain → | Ar Cl Rw Pr Ar Cl Pr Rw |

Note: Bold values represent the best results and underlined values represent the second-best results.

IDS: 1–10, 11–20, 21–30, and 5 (class IDs: 1–5, 6–10, 11–15) classes as known classes, respectively. The class IDs from 1 to 65 are defined by ranking all the classes in alphabetical order.

The results are shown in Table X, and the results of three comparative methods are from [10]. Our proposed method OSLPP outperforms others in settings where there are 40/25 known classes and is comparable to ROS [10] when there are 10/5 known classes. These results indicate our method is more advantageous when the openness is moderate.

J. Extend to Transductive Open-Set Recognition

To investigate how our proposed approach works for the transductive OSR problems without domain shift, we design and conduct a series of experiments. On the Office-Home dataset, we divide each domain into two disjoint subsets: one for training and the other for testing: 1) We simulate the OSR problem under the transductive setting. The training data and the test data are from the same domain but the test data contain additional classes that are not present in the training data. 2) We simulate the OSDA problem. The training data and the test data are from the source and target domains, respectively. For both (a) and (b), we apply our full method and the ablated one without the OSLPP algorithm for comparison. The results are shown in Table XI. Based on the experimental results, following conclusions can be drawn: 1) our proposed approach can be directly applied to the OSR problems (i.e., the tasks of Pr→Pr and Rw→Rw) under the transductive setting; 2) for the transductive OSR, the high performance of our approach is mainly contributed by the progressive sample selection and rejection strategy, given that removing the OSLPP algorithms only degrades the performance marginally; and 3) for OSDA, the OSLPP algorithm plays a more crucial role in bridging the domain gap, thereby significant performance drops are observed when it is not used.

In summary, the experimental results in Table XI demonstrate our proposed method also works for OSR under the transductive setting, and it is the OSLPP algorithm that enables our proposed approach to bridge the domain gap when it does exist.

K. Visual Inspection

We take the task Art→Clipart as an example to inspect how the class separability can be improved in the learned subspace by visualizing the features with the t-SNE technique. To avoid clutters, we select the first ten shared known classes and the first ten unknown classes for visualization. As shown in Fig. 4, the samples from the source domain (“o”) and the target domain (“×”) in the learned common subspace are better aligned classwise, while the unknown classes (“+”) are also better separated from the known classes.

Other than the original features (a) and the features learned by our proposed OSLPP (c), we also visualize the features learned by LDA (b) as a representative of existing methods which push unknown classes into a compacted region in the learned space as illustrated in Fig. 1 and demonstrated in Fig. 4. As a result, our proposed OSLPP algorithm outperforms LDA in the domain adaptation tasks (e.g., the HOS values are 67.0% and 65.0%, respectively, on the Office-Home dataset).

VI. DISCUSSION AND CONCLUSION

We address the OSDA problem in the image classification domain by proposing a novel OSLPP algorithm and a progressive pseudolabeled sample selection and rejection algorithm. The OSLPP adapts the original LPP algorithm to the OSDA scenario...
by considering the labeled source samples and pseudolabeled target samples, which have been either selected or rejected. Experimental results on two benchmark datasets demonstrate our proposed method can perform comparably or outperform state-of-the-art approaches to OSDA and is more efficient. The method is also less sensitive to the hyperparameters in terms of the harmonic mean but provides the flexibility of trading off the accuracy of known and unknown classes, which can be useful in real-world applications.

Our method suffers from the common issue of how to set proper hyperparameters to adjust the recognition accuracy of known classes and unknown classes. In a real-world scenario, we do not have prior knowledge of how many unlabeled target samples are from unknown classes if there are any. In our method, this issue corresponds to the question of how to set the values of \( n_p \) and \( T \) given a domain adaptation task to achieve the best performance. This is also the key problem in universal domain adaptation problems [26] and will be left to future work.

REFERENCES

[1] Q. Wang and T. P. Breckon, “Generalized zero-shot domain adaptation via coupled conditional variational autoencoders,” *Neural Netw.*, vol. 163, pp. 40–52, Jun. 2023.

[2] Z. Yi, W. Shang, T. Xu, and X. Wu, “Neighborhood preserving and weighted subspace learning method for drift compensation in gas sensor,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 6, pp. 3530–3541, Jun. 2022.

[3] J. Tian, J. Zhang, W. Li, and D. Xu, “VDM-DA: Virtual domain modeling for source data-free domain adaptation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 6, pp. 3749–3760, Jun. 2022.

[4] X. Xu, H. He, H. Zhang, Y. Xu, and S. He, “Unsupervised domain adaptation via importance sampling,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4688–4699, Dec. 2020.

[5] C. Chen et al., “Progressive feature alignment for unsupervised domain adaptation,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 627–636.

[6] C. Chen, Z. Chen, B. Jiang, and X. Jin, “Joint domain alignment and discriminative feature learning for unsupervised deep domain adaptation,” in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 3296–3303.

[7] M. Meng, Z. Wu, T. Liang, J. Yu, and J. Wu, “Exploring fine-grained cluster structure knowledge for unsupervised domain adaptation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 8, pp. 5481–5494, Aug. 2022.

[8] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, “Open set domain adaptation by backpropagation,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 153–168.

[9] H. Liu, Z. Cao, M. Long, J. Wang, and Q. Yang, “Separate to adapt: Open set domain adaptation via progressive separation,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2927–2936.

[10] S. Bucci, M. R. Loghmani, and T. Tommasi, “On the effectiveness of image rotation for open set domain adaptation,” in *Proc. Eur. Conf. Comput. Vis.*, New York, NY, USA: Springer-Verlag, 2020, pp. 422–438.

[11] Y. Luo, Z. Wang, Z. Huang, and M. Baktashmotlagh, “Progressive graph learning for open-set domain adaptation,” in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2020, pp. 6468–6478.

[12] Z. Fang, J. Lu, F. Liu, J. Xuan, and G. Zhang, “Open set domain adaptation: Theoretical bound and algorithm,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 10, pp. 4309–4322, Oct. 2021.

[13] J. Liu, M. Jing, J. Li, K. Lu, and H. T. Shen, “Open set domain adaptation via joint alignment and category separation,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 9, pp. 6168–6199, Sep. 2023.

[14] M. Jing, J. Li, L. Zhu, Z. Ding, K. Lu, and Y. Yang, “Balanced open set domain adaptation via centroid alignment,” in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 9, 2021, pp. 8013–8020.

[15] Z. Feng, C. Xu, and D. Tao, “Open-set hypothesis transfer with semantic consistency,” *IEEE Trans. Image Process.*, vol. 30, pp. 6473–6484, 2021.
[41] D. Bashkirova et al., “VisDA-2021 competition universal domain adaptation to improve performance on out-of-distribution data,” 2021.

[42] X. Peng, B. Usman, K. Saito, N. Kaushik, J. Hoffman, and K. Saenko, “Syn2real: A new benchmark for synthetic-to-real visual domain adaptation,” 2018, arXiv:1806.09755.

[43] Y. Ganin et al., “Domain-adversarial training of neural networks,” J. Mach. Learn. Res., vol. 17, no. 1, pp. 2096–2030, 2016.

[44] P. Panareda Busto and J. Gall, “Open set domain adaptation,” in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 754–763.

[45] L. Chen, Y. Lou, J. He, T. Bai, and M. Deng, “Geometric anchor correspondence mining with uncertainty modeling for universal domain adaptation,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 16134–16143.

Qian Wang received the B.Sc. degree in electronic engineering and the master’s degree in biomedical engineering from the University of Science and Technology of China, Hefei, China, in 2010 and 2013, respectively, and the Ph.D. degree in machine learning from The University of Manchester, Manchester, U.K., in 2017. He is a Visiting Scholar with the Department of Computer Science, Durham University, Durham, U.K. He was a Deep Learning Researcher with MediaTek Research Cambridge, Cambridge, U.K., and a Research Associate with the Department of Computer Science, Durham University. His research interests include deep learning and computer vision.

Fanlin Meng (Senior Member, IEEE) received the B.Sc. degree in automation from China University of Mining and Technology, Xuzhou, China, in 2008, the M.Sc. degree in systems engineering from Xiamen University, Xiamen, China, in 2011, and the Ph.D. degree in computer science from the University of Manchester, Manchester, U.K., in 2015. He is currently a Lecturer in data science with the University of Manchester. His primary research interests include energy market, smart energy and mobility, operations research, machine learning, and game theory and optimization.

Dr. Meng is a fellow of British Computer Society and a Member of EPSRC Peer Review College.

Toby P. Breckon (Senior Member, IEEE) received the Ph.D. degree in informatics (computer vision) from the University of Edinburgh, Edinburgh, U.K., in 2006. He is currently a Professor with the Departments of Engineering and Computer Science, Durham University, Durham, U.K. His key research interests include the domain of computer vision and image processing, and he leads a range of research activity in this area. He has been a Visiting Member of Faculty with the Ecole Supérieure des Technologies Industrielles Avancées, Bidart, France; Northwestern Polytechnical University, Xi’an, Shaanxi, China; Shanghai Jiao Tong University, Shanghai, China; and Waseda University, Tokyo, Japan.

Dr. Breckon is a Chartered Engineer, a Chartered Scientist, and a Fellow of the British Computer Society. In addition, he is an Accredited Senior Imaging Scientist and a Fellow of the Royal Photographic Society. He led the development of image-based automatic threat detection for the 2008 UK MoD Grand Challenge winners [R.J. Mitchell Trophy, (2008), IET Innovation Award (2009)]. His work is recognized as a recipient of the Royal Photographic Society Selwyn Award for early career contribution to imaging science (2011). http://www.durham.ac.uk/toby.breckon/.