DIAL: Domain-Invariant Adversarial Learning for Sim-to-Real Robotic Manipulation

Ajay Kumar Tanwani, Daniel Zeng, Matthew Trepte, and Ken Goldberg
University of California, Berkeley

***Abstract*—Sample and time complexity of collecting/labelling real data can be very high for training vision-based deep learning models for robotic manipulation. Generating large-scale synthetic data in simulation is a feasible alternative, albeit the modelling inaccuracies do not generalize to the physical world. In this paper, we present a domain-invariant adversarial learning (DIAL) algorithm to adapt deep models to the physical environment with a small amount of real data. Existing approaches that only mitigate the covariate shift by aligning the marginal distributions across the domains and assume the conditional distributions to be domain-invariant can lead to ambiguous transfer in real scenarios. We propose to jointly align the marginal (input domains) and the conditional (output labels) distributions to mitigate the covariate and the conditional shift across the domains with regularized adversarial learning, and combine it with a soft metric learning triplet loss to make the conditional distributions disjoint in the shared feature space. Experiments on MNIST benchmark domains and vision based sim-to-real transfer of object recognition for mobile manipulation leads to state-of-the-art performance in adapting models across domains with small number of real examples. Videos, code and data are available at: https://sites.google.com/view/dial-sr**

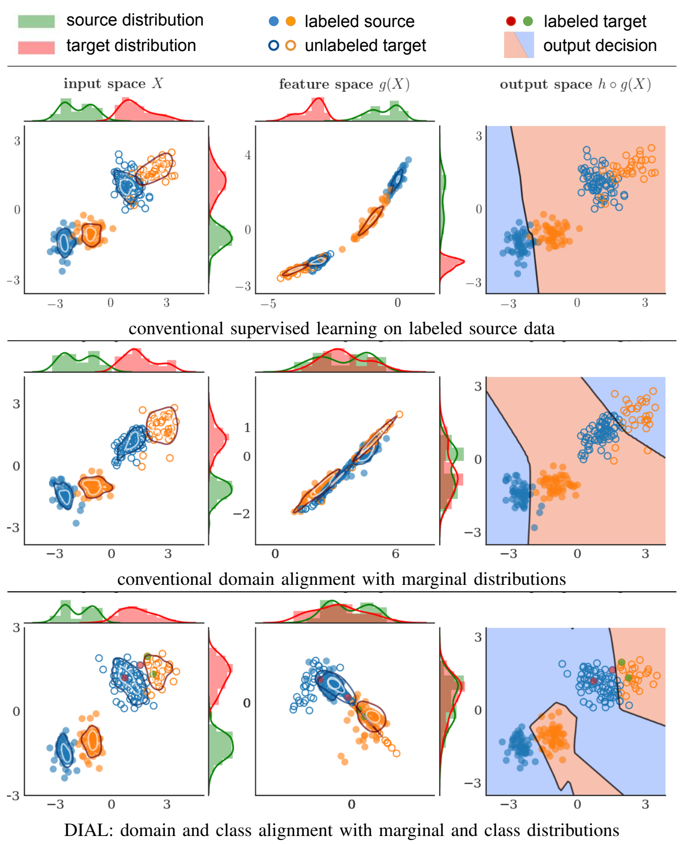

Fig. 1: Domain-invariant adversarial learning on 2D synthetic data with 2 classes: *(top)* conventional supervised learning on source domain does not generalize to the target domain drawn from a different distribution, *(middle)* unsupervised alignment of marginal distributions across domains leads to cross mapping of class categories with negative transfer, *(bottom)* DIAL leverages upon a few labeled target examples to semantically align both the marginal and the class distributions for semi-supervised domain adaptation.

I. Introduction

Deep learning models tend to perform well when plenty of labelled training data is available and the testing data is drawn from the same distribution as that of the training data. Collecting and labelling large scale training data for robotics applications, however, is time consuming and cumbersome [39, 24]. Additionally, the sample selection bias in data collection limits the model use to very specific environmental situations.

Training deep models in simulation for robot manipulation is becoming a popular alternative [55, 38, 18]. Despite the efforts to build good dynamic models and high fidelity simulators for scalable data collection, it is difficult to transfer the desired behaviour of robot in real environments. Model adaptation is a fundamental characteristic to make robots useful for a wide range of operating conditions in the real world. For example, we wish to adapt an object recognition or a grasping model trained from a large corpus of object meshes in the simulator to a different camera viewpoint of the robot, to a different robot, or to a different physical environment such as a household or a machine shop.

In this paper, we investigate the problem of domain adaptation to learn a deep model that can be transferred to a new domain [58]. We analyze the situation where we have a lot of labeled training data from the simulator or the *source* distribution, and we are interested in adapting the model with limited or no labelled training data drawn from the *target* distribution. Existing popular approaches to domain adaptation learn common feature transformations by aligning marginal distributions across domains [8, 57, 15, 46]. They implicitly assume that the class conditional distributions of the transformed features is also same across domains, i.e., if the model performs well on the source domain and the features overlap across domains, the model is going to perform well on the target domain. This, however, creates ambiguity in class alignment across source and target domains in the shared feature space that can result in a negative transfer to a new domain as shown in Fig. 1 (middle).

In this paper, we present a **domain-invariant adversarial learning** (DIAL) algorithm that: 1) resolves the ambiguity in domain adaptation by leveraging upon a few labeled examples of the target domain to align both the marginal and the conditional distributions in the shared feature space with adversarial learning, 2) increases the inter-class variance and reduces the intra-class variance of the shared feature space with a soft variant of metric learning triplet loss.

We apply the approach to vision-based robot decluttering, where a mobile robot grasps objects from a cluttered floor and sorts them into respective bins [54]. We seek to learn the deep models with the help of simulated images of object meshes and use a small number of labeled real world images to mitigate the domain shift between the simulator and the real images as well as the object distributions.

This paper makes the following contributions:

1) We propose a domain-invariant adversarial learning (DIAL) algorithm to mitigate the covariate and conditional shift by aligning the marginal and the conditional distributions across domains, and making them disjoint with a novel soft triplet distribution loss in a shared feature space.
2) Experiments on MNIST related benchmark domains and sim-to-real transfer of a single-shot object recognition model for mobile manipulation suggest performance improvement over state-of-the-art domain adaptation methods.

II. Related Work

Domain adaptation methods mitigate the discrepancy of the learner across different distributions by reducing the effect of covariate shift, prior distribution shift and/or sample selection bias [40, 58]. Deep invariant transferable representations can be learned across domains by matching feature and/or pixel distributions with similarity metrics like maximum mean discrepancy [29], shared or associative embedding [29, 13], reconstruction from latent features [4, 9] or adversarial learning [8, 3, 57, 15, 65].

Domain Adaptation with Adversarial Learning: Adversarial learning methods use the principle of Generative Adversarial Networks (GANs) [11] to train two networks in a zero-sum game: a discriminator tries to distinguish the target domain from the source domain, whereas a generator tries to fool the discriminator to make the source domain look like the target one. The source distribution is used as generator input instead of a normal distribution commonly used in GANs for sample efficiency. Ganin et al. proposed DANNs that learn a shared feature space by making marginal feature distributions similar with a gradient-reversal layer between feature extractor and domain classifier [8]. Tzeng et al. explicitly used inverted labels for adapting the target data only with an independent feature extractor instead of using negative gradient [57]. Associative and cyclic approaches learn a shared feature space such that the cycle of moving from a source sample to a target sample and back from a target sample to a source sample is consistent [50, 13, 15, 65]. These approaches imply that a provably low target error can be obtained by minimizing the marginal discrepancy between two classifiers in agreement with the theoretical analysis by Ben-David et al. [2], Mansour et al. [31] and Zhang et al. [63].

A common trend among the existing methods is to align the marginal distributions across the domains. This, however, does not guarantee that that the samples from the same class across the domains are mapped nearby in the feature space (see Fig. 1 for a conceptual example). This lack of semantic alignment is a major limitation when the feature representation has conditional and/or label shift across the domains.

Label and Conditional Domain Adaptation: An under explored line of work addresses matching the conditional and/or the label distributions in the feature space. Related work in this direction make use of linear projections in the shared feature space [53, 28, 10], or estimate importance weights to address the label shift problem [26, 1]. Hoffman et al. [14] use domain discriminator for global alignment, and formulate multiple instance-based weak constraints for class specific adaptation [37]. The class specific adaptation makes a strong assumption that the proportion of pixels across each class category remains the same in the source and the target images. Chen et al. [6] use multiple domain discriminators for semantic segmentation of each class, in addition to a global discriminator for cross-domain alignment. Long et al. [30] take the outer product of the features and the class predictions as discriminator input to capture the cross-covariance between the classes and the features for conditional adversarial domain adaptation. Saito et al. [46] take domain-specific decision boundary into account with two classifiers as conditional discrepancy discriminators that adapt the target features to be within the support of the source domain. Shu et al. [51] combine domain adversarial training with conditional entropy regularizer to push the decision boundary far from the densely populated target samples. Wu et al. [59] relaxed the marginal distribution matching objective to bound the density ratio across domains with asymmetrical distribution alignment.

In comparison to these approaches, we propose a domain-invariant adversarial learning framework that provisions for both marginal and conditional adaptation of target data to be within the support of the source domain while encouraging the feature representation of each class to be disjoint with a soft variant of triplet loss.

Unsupervised vs Semi-Supervised Domain Adaptation: Learning discriminative class representations make use of target pseudo-labels to encourage a low-density separation between classes in the target domain [21]. However, this requires heuristically choosing a threshold for reliable prediction on unlabeled target data. Supervised [56] or semi-supervised methods [22, 61] provide a practical alternative where data is typically scarce (for example, when acquired via kinesthetic or teleoperation interfaces in robotics). Motiian et al. [33, 32] perform few-shot domain adaptation by sampling positive and negative pairs across labeled class categories of both domains to encode a four class domain-class discriminator. Saito et al. [47] present a semi-supervised adaptation approach by minmax optimization of conditional entropy. In this work, we leverage upon a few target examples to efficiently align conditional distributions across domains that stabilizes adaptation to challenging domains with large domain shift.

Simulation to Real Transfer: Perception and control policies learned in simulation often do not generalize to the real robots due to the modeling inaccuracies. Domain randomization methods [55, 44, 38, 41] treat the discrepancy between the domains as variability in the simulation parameters, assuming the real world distribution as one such randomized instance of the simulation environment. In contrast, domain adaptation methods learn an invariant mapping function for matching distributions between the simulator and the robot environment [60]. Saxena et al. [48] used synthetic objects to learn a vision based grasping model. Gupta et al. [12] learn invariant feature spaces to transfer skills across agents with reinforcement learning. Zhang et al. use adversarial discriminative learning for sim-to-real transfer of visuomotor policies for a reaching task. Saenko et al. [45] adapt the object recognition models to new domains. Nogues et al. [35] recognize objects by adversarial refinement of synthetic images. Chen et al. [7] and Zhu et al. [66] adapt the faster R-CNN segmentation model at image and instance level for cross-domain robustness of object recognition in an unsupervised manner. Hybrid methods combine domain randomization and domain adaptation to guide randomization of the simulator parameters to match the real distribution [5, 43, 17].

Online sim-to-real transfer learning approaches involving imitation or reinforcement are computationally very demanding, requiring an arm farm [24, 20] or hundreds of hours of robot training [39]. This paper, in contrast, focuses on a data-efficient approach to learn invariant features across domains while reducing the cost of labeling real data. We present a semi-supervised **domain invariant adversarial learning** approach that aligns the marginal and the conditional output distributions across the domains with regularized adversarial learning, and encourages discriminative representations between class categories with a soft variant of triplet loss. We show state-of-the-art performance on MNIST datasets and sim-to-real transfer of a single-shot object recognition for vision based decluttering with a mobile manipulator.

III. Problem Statement

We consider two environments: one belonging to the simulator or source domain $\langle D_S, \pi_S \rangle$ comprising of the dataset $\{(\boldsymbol{x}_i^S, \boldsymbol{y}_i^S)\}_{i=1}^{N_S}$ and the other belonging to the real or target domain $\langle D_T, \pi_T \rangle$ samples $\{(\boldsymbol{x}_i^T, \boldsymbol{y}_i^T)\}_{i=1}^{N_T}$ with a few labeled samples $N_S \gg N_T$. The samples are drawn from $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X}$ is the input space and $\mathcal{Y}$ is the output space and the superscripts S and T indicate the draw from two different distributions of source and target random variables $(X^S \times Y^S)$ and $(X^T \times Y^T)$ respectively. Each output labeling function is represented as $\pi : \mathcal{X} \rightarrow \mathbb{R}^{|\mathcal{Y}|}$, with π_S and π_T as the output policy of the source and the target domain. Let $\mathcal{Z}$ denote the intermediate representation space induced from $\mathcal{X}$ by a feature transformation $g : \mathcal{X} \rightarrow \mathcal{Z}$, which is mapped to the output conditional distribution by $f : \mathcal{Z} \rightarrow \mathcal{Y}$ under the transformation $X \xrightarrow{g} Z \xrightarrow{f} Y$. The composite function $f \circ g$ defines the policy π of a domain, and the loss of a candidate policy $\pi \in \mathcal{H}$ represents the disagreement with respect to the source policy π_S on domain D_S as $\epsilon_S(\pi) := \epsilon_S(\pi, \pi_S) = \mathbb{E}_{\boldsymbol{x} \sim D_S} \; [|\pi(\boldsymbol{x}) - \pi_S(\boldsymbol{x})|]$. For example, loss function for classification on target domain is represented by $\epsilon_T(\pi, \pi_T) = \mathbb{E}_{\boldsymbol{x} \sim D_T} \; [\mathbb{I}(\pi(\boldsymbol{x}) \neq \pi_T(\boldsymbol{x}))]$, while for regression as $\mathbb{E}_{\boldsymbol{x} \sim D_T} \; [\|\pi(\boldsymbol{x}) - \pi_T(\boldsymbol{x}))\|_2]$.

The goal is to learn the policy $\pi \in \mathcal{H}$ using the labeled source examples and a few or no labeled target examples such that the error on the target domain is low. We seek to minimize the joint discrepancy across the domains in a shared feature space and map the output policy on the embedded data to minimize the target error. Denoting the joint distribution of input domain and output labels across source and target domain as $\Pr(X^S, Y^S) = \Pr(Y^S | X^S) \cdot \Pr(X^S)$, and $\Pr(X^T, Y^T) = \Pr(Y^T | X^T) \cdot \Pr(X^T)$ respectively, the DIAL objective is to minimize,

$$
\begin{aligned}
\left| \log \Pr(X^S, Y^S) - \log \Pr(X^T, Y^T) \right| &= \left| \log \Pr(Y^S \mid X^S) + \log \Pr(X^S) \right| \\
&\quad - \left| \log \Pr(Y^T \mid X^T) + \log \Pr(X^T) \right|, \\
&= \left| \log \Pr(Y^S \mid X^S) - \log \Pr(Y^T \mid X^T) \right| + \left| \log \Pr(X^S) - \log \Pr(X^T) \right|, \\
&\approx \left| f \circ g(X^S) - f \circ g(X^T) \right| + \left| g(X^S) - g(X^T) \right|, \\
&= \underbrace{d_{\Pr(Y|X)} \left(D_S^Y, D_T^Y \right)}_{\text{conditional discrepancy}} + \underbrace{d_{\Pr(X)} \left(D_S^Z, D_T^Z \right)}_{\text{marginal discrepancy}}. \qquad (1)
\end{aligned}
$$

The joint discrepancy consequently depends on both the **conditional discrepancy** and **marginal discrepancy** between the source and the target domains, and is minimized by aligning the marginal and conditional distributions.

Definition 3.1: Marginal Distributions Alignment: *Given two domains $D_S = \{(\boldsymbol{x}_i^S, \boldsymbol{y}_i^S)\}_{i=1}^{N_S}$ and $D_T = \{(\boldsymbol{x}_i^T, \boldsymbol{y}_i^T)\}_{i=1}^{N_T}$ drawn from two different distributions $\Pr(X^S, Y^S) \neq \Pr(X^T, Y^T)$ with non-zero covariate shift $\mathrm{KL}\left(\Pr(X^S) \| \Pr(X^T)\right) > 0$, marginal alignment corresponds to finding the feature transformation $g : \mathcal{X} \rightarrow \mathcal{Z}$ such that the discrepancy between the transformed marginal distributions is minimized, i.e., $\Pr\left(g(X^S)\right) = \Pr\left(g(X^T)\right)$.*

Definition 3.2: Conditional Distributions Alignment: *Given two domains $D_S = \{(\boldsymbol{x}_i^S, \boldsymbol{y}_i^S)\}_{i=1}^{N_S}$ and $D_T = \{(\boldsymbol{x}_i^T, \boldsymbol{y}_i^T)\}_{i=1}^{N_T}$ drawn from random variables $(X^S \times Y^S)$ and $(X^T \times Y^T)$ with different output conditional probability distributions $\Pr\left(Y^S \mid X^S\right) \neq \Pr\left(Y^T \mid X^T\right)$, conditional alignment corresponds to finding the transformation $X \xrightarrow{g} Z \xrightarrow{f} Y$ such that the discrepancy between the transformed conditional distributions is minimized, i.e., $\Pr\left(Y^S \mid g(X^S)\right) = \Pr\left(Y^T \mid g(X^T)\right)$.*

Note that the methods aligning marginal distributions only assume the same output conditional distribution $\Pr\left(Y \mid X\right)$ across the domains for the adaptation to be effective. This assumption of different marginal distribution across domains, but similar conditional distribution is known as covariate shift [40]. Several methods attempt to find invariant transformation to mitigate covariate shift such that $\Pr\left(g(X)\right)$ is

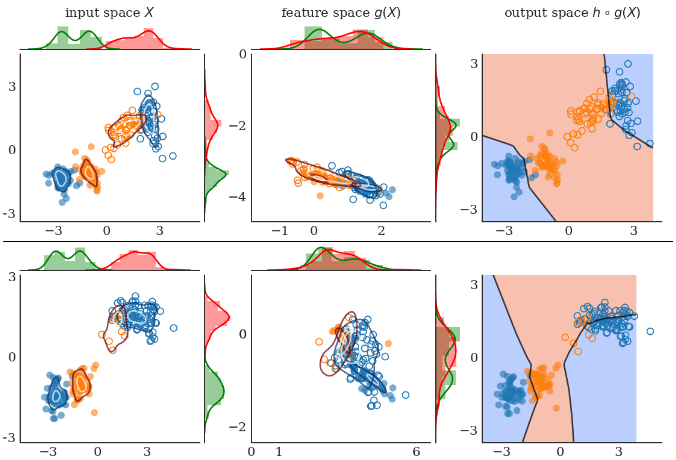

Fig. 2: Limitations of marginal distributions alignment only: *(top)* even though the source and target distributions are the same as in Fig. 1, swapping class labels of the target domain leads to cross label matching in unsupervised domain adaptation, *(bottom)* label shift with class ratio of 1.0 for the source and 0.1 for the target leads to uneven mixing of class samples.

similar across the domains by minimizing a domain discrepancy measure [8, 57, 46]. However, it is not clear if the assumption of $\Pr\left(Y \mid g(X)\right)$ also remains the same across the domains. As the target labels may not be available in unsupervised domain adaptation, the class conditional distributions are widely assumed to be true under the transformation $\Pr\left(Y^S \mid g(X^S)\right) = \Pr\left(Y^T \mid g(X^T)\right)$. In this paper, we consider the joint marginal and conditional distributions alignment problem across the domains. Note that it is a non-trivial problem since we have access to no or a few labels in the target domain. We consider the problem in a semi-supervised setting to resolve the ambiguity in conditional distributions alignment with a few labeled target examples.

IV. Limitations and Theoretical Insights

In this section, we present theoretical insights of a family of domain-invariant adversarial learning algorithms building upon the works of [2, 31, 64, 19, 62]. The analysis will lead us to highlight the limitations of existing approaches and formulate the DIAL algorithm in the subsequent section.

***Theorem** 4.1:* (Ben-Davide et al. [2]). Given two domains D_S and D_T, the error of a hypothesis $h \in \mathcal{H}$ in the target domain $\epsilon_T(h)$ is bounded by the sum of: 1) the error of the hypothesis in the source domain, 2) the discrepancy of the hypothesis class between the domains $d_{\mathcal{H}\Delta\mathcal{H}}(D_S^Z, D_T^Z) := 2\sup_{h,h' \in \mathcal{H}} |\epsilon_S(h,h') - \epsilon_T(h,h')|$, and 3) the best-in-class joint hypothesis error $\lambda_{\mathcal{H}} = \min_{h \in \mathcal{H}} \left[\epsilon_S(h) + \epsilon_T(h)\right]$.

$$\epsilon_T(h) \leq \epsilon_S(h) + \frac{1}{2} d_{\mathcal{H}\Delta\mathcal{H}}(D_S^Z, D_T^Z) + \lambda_{\mathcal{H}} \tag{2}$$

The $d_{\mathcal{H}\Delta\mathcal{H}}$ divergence can be empirically measured by training a classifier that discriminates between source and target instances, and subsequently minimized by aligning the marginal distributions between (unlabeled) source and target instances. The joint hypothesis error $\lambda_{\mathcal{H}}$ is widely assumed to be small, i.e., there exists a policy that performs well on the induced feature distributions of the source and the target examples after marginal alignment. More generally, the optimal joint error $\lambda_{\mathcal{H}}$ represents the cross-domain performance of the optimal policies $\min\{\epsilon_S(\pi_T), \epsilon_T(\pi_S)\}$. Hence, the upper bound on the target error is,

$$\begin{aligned}\epsilon_T(h) \leq \epsilon_S(h) + d_{\mathcal{H}\Delta\mathcal{H}}(D_S^Z, D_T^Z) + \\ \min\{\mathbb{E}_{D_S}|\pi_S - \pi_T|, \mathbb{E}_{D_T}|\pi_S - \pi_T|\}.\end{aligned} \tag{3}$$

In Fig. 1 (middle) showing conventional domain adaptation by aligning marginal distributions, the cross-overlapping area across class categories in the shared feature space (blue and orange) represents the optimal joint error given by $\epsilon_T(\pi_S, \pi_T)$ or $\epsilon_S(\pi_T, \pi_S)$. High joint error signifies that the conventional marginal alignment approaches fail in the presence of conditional or label shift between the domains. Intuitively, matching marginal distributions only across domains can result in mixing samples with different class labels, or in worst case swap labels of different classes resulting in negative transfer. Along the similar lines, Zhao et al. [64] presented an information theoretic lower bound implying that aligning the marginal distributions and minimizing the source error would lead to a large target error if the marginal label distributions are significantly different. Zhang et al. [62] presented a cross margin discrepancy measure and a more general upper bound for target error taking joint error into account.

Consequently, we highlight two main limitations of aligning marginal distributions only with unsupervised domain adaptation shown in Fig. 2: 1) **cross label matching**: labels of different classes are swapped in the shared feature space, and 2) **label shift**: the distribution of class labels in the source and the target domains are different, leading to samples of one class being mixed with another class. In this work, we align both the marginal and the conditional discrepancies across domains and use a few labeled target samples to avoid ad-hoc mixing of class categories and negative transfer with cross label assignments.

V. DIAL Algorithm

The central idea of the domain-invariant adversarial learning method is to mitigate the domain shift by minimizing the joint distribution error in a shared feature space. The joint distribution of the source and the target domains is dependent upon both the marginal and the conditional distributions. To align the marginal distributions, we use a domain classifier to discriminate between source and target domains using the middle features that are adapted to deceive the domain classifier with regularized adversarial learning [8, 57, 51]. Aligning the conditional distributions, however, is non-trivial due to the lack of labeled target examples. As discussed before, ensuring marginal alignment in the presence of conditional discrepancy can cause more harm than good.

We align the conditional distributions by using class-wise domain discriminators for each class [14, 6]. The class-wise

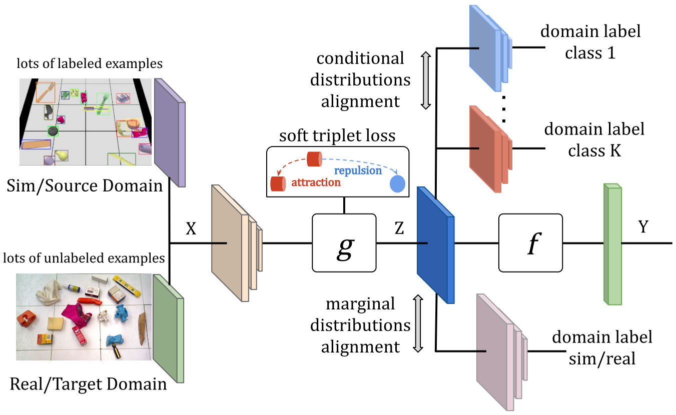

Fig. 3: DIAL aligns marginal and conditional distributions of source and target domains, and uses a soft metric learning triplet loss to make the feature distributions disjoint in a shared feature space.

discriminators suppress the conditional discrepancy between source and target domains across each class. Since we do not have labels for all the target examples, we leverage upon a few labeled target examples and extract pseudo-labels for the unlabeled examples by querying the output network. The output policy is trained on labeled source examples and unlabeled target data in a semi-supervised manner with entropy regularization. We further use a novel soft metric learning triplet loss to make the conditional distributions disjoint in the feature space. The loss reduces the divergence across similar class categories and vice versa, resulting in high inter-class variance and low intra-class variance of the shared feature distributions. The overall architecture of DIAL is shown in Fig. 3.

A. Marginal Distributions Alignment

Given the joint distribution $\Pr(X, Y) = \Pr(Y|X) \cdot \Pr(X)$, we align the marginal distributions of the source domain $\Pr(X^S)$ and the target domain $\Pr(X^T)$ with adversarial learning. The generator $g(X)$ encodes the data in a shared feature space, and the discriminator $D(X)$ predicts the binary domain label whether the data point is drawn from the source or the target distribution. The discriminator loss $\mathcal{L}_{ma}(g, D)$ is optimized in a supervised manner,

$$\min_{D} \mathcal{L}_{ma}(g, D) = -\mathbb{E}_{\boldsymbol{x}_s \sim X_s} \left[\log D\left(g(\boldsymbol{x}_s)\right)\right] - \mathbb{E}_{\boldsymbol{x}_t \sim X_t} \left[\log\left(1 - D\left(g(\boldsymbol{x}_t)\right)\right)\right]. \tag{4}$$

The generator subsequently adopts the target domain features to confuse the discriminator with inverted domain labels to avoid vanishing gradients [57]. Note that the gradient reversal layer can also be used [8]. The generator loss $\mathcal{L}_{ma}(g, D)$ adapts the feature extractor,

$$\min_{g} \quad \mathcal{L}_{ma}(g, D) = -\mathbb{E}_{\boldsymbol{x}_t \sim X_t} \left[\log D\left(g(\boldsymbol{x}_t)\right)\right]. \tag{5}$$

Without loss of generality, we adapt the feature extractor with respect to the target data only. The objective is optimized in a minmax fashion where the discriminator maximizes the empirical marginal discrepancy for a given features distribution, and the feature extractor minimizes the discrepancy by aligning the marginal distributions.

B. Conditional Distributions Alignment

Conditional distributions alignment can overcome the issues with marginal distributions alignment only, namely cross label matching and the shift in labeling distributions. Effective alignment of conditional distributions depends upon two factors: 1) target pseudo-labels extraction, and 2) balanced sampling of features per class across domains.

To this end, we leverage upon a few labeled target examples and train the output network with labeled source and target examples using the cross-entropy loss,

$$\mathcal{L}_{ca\_sc}(f, g) = \mathbb{E}_{\boldsymbol{x}_s, \boldsymbol{y}_s \sim (X_s, Y_s)} \left[-\boldsymbol{y}_s \log f\left(g(\boldsymbol{x}_s)\right)\right] + \mathbb{E}_{\boldsymbol{x}_t, \boldsymbol{y}_t \sim (X_t, Y_t)} \left[-\boldsymbol{y}_t \log f\left(g(\boldsymbol{x}_t)\right)\right]. \tag{6}$$

Additionally, we regularize the output network with entropy minimization on unlabeled target data that pushes the decision boundaries to be in well-separated low-density regions [36, 51],

$$\mathcal{L}_{ca\_te}(f, g) = \mathbb{E}_{\boldsymbol{x}_t \sim X_t} \left[-f\left(g(\boldsymbol{x}_t)\right) \log f\left(g(\boldsymbol{x}_t)\right)\right]. \tag{7}$$

We extract the pseudo-labels for the unlabeled target data, $\hat{\boldsymbol{y}}_t = \arg\max f\left(g(\boldsymbol{x}_t)\right)$, by querying the pre-trained source network and retain only top-n pseudo-labels of each class category based on their confidence. We sample with replacement to create a balanced mini-batch with half source and half labeled target examples, and augment the pool of target examples with the pseudo-labeled data after the pre-training stage only.

A minimax game with adversarial learning aligns the conditional distribution of each class using a domain discriminator. First, the class discriminator $C(X)$ estimates the conditional discrepancy with respect to the source and target networks for a fixed feature extractor. Second, the generator adapts the feature extractor to minimize the conditional discrepancy for a fixed source and target network parameters, i.e., the adversarial loss for each class $k = 1 \ldots |\mathcal{Y}|$ is,

$$\min_{C} \mathcal{L}_{ca_k}(g, C) = -\mathbb{E}_{\boldsymbol{x}_s \sim X_s} \left[\log C\left(g(\boldsymbol{x}_s)\right)\right] - \mathbb{E}_{\boldsymbol{x}_t \sim X_t} \left[\log\left(1 - C\left(g(\boldsymbol{x}_t)\right)\right)\right], \tag{8}$$

$$\min_{g} \quad \mathcal{L}_{ca_k}(g, C) = -\mathbb{E}_{\boldsymbol{x}_t \sim X_t} \left[\log C\left(g(\boldsymbol{x}_t)\right)\right]. \tag{9}$$

Minimizing the conditional discrepancy penalizes the feature extractor to separate the cross domain overlap for each class to give low joint error $\epsilon_T(\pi_s, \pi_T)$ in Eq. (3) for provably effective domain adaptation.

C. Metric Learning with Triplet Distributions Loss

Aligning the marginal and conditional distributions brings source and target examples together in the shared feature space. To further increase the inter-class variance and reduce the intra-class variance, we propose a soft metric learning variant of triplet loss [49, 42] that operates on distributions of anchor, positive and negative examples (instead of tuples of individual examples). Given M labeled examples in a mini-batch, the loss posits that the Kullback-Leibler (KL)-divergence between anchor and positives distribution in the shared feature space is less than the KL-divergence between anchor and negatives distribution by some constant margin $\alpha_{tl} \in \mathbb{R}_+$. Mathematically,

$$\mathcal{L}_{tl} = \overbrace{\sum_{a=1}^{M}}^{\text{all anchors}} \Bigg[\overbrace{\frac{1}{M_p - 1} \sum_{\substack{p=1 \\ p \neq a}}^{M_p} \mathrm{KL}\left(\mathcal{N}\left(\bar{g}(\boldsymbol{x}_a), \sigma^2\right) \Big\| \mathcal{N}\left(\bar{g}(\boldsymbol{x}_p), \sigma^2\right)\right)}^{\text{all positives}}$$
$$- \overbrace{\frac{1}{M_n} \sum_{n=1}^{M_n} \mathrm{KL}\left(\mathcal{N}\left(\bar{g}(\boldsymbol{x}_a), \sigma^2\right) \Big\| \mathcal{N}\left(\bar{g}(\boldsymbol{x}_n), \sigma^2\right)\right)}^{\text{all negatives}} + \alpha_{tl} \Bigg]_+ , \tag{10}$$

where $\{.\}_+$ is the hinge loss, $\bar{g}(\boldsymbol{x})$ is normalized to extract scale-invariant features similar to [49], $\mathcal{N}\left(\bar{g}(\boldsymbol{x}_a), \sigma^2\right)$ is short for distribution of examples, $\left\{\mathcal{N}\left(\bar{g}(\boldsymbol{x}_i); \bar{g}(\boldsymbol{x}_a), \sigma^2\right) = \frac{\exp^{-\frac{1}{\sigma^2}\|\bar{g}(\boldsymbol{x}_i) - \bar{g}(\boldsymbol{x}_a)\|_2^2}}{\sum_{j=1}^{K} \exp^{-\frac{1}{\sigma^2}\|\bar{g}(\boldsymbol{x}_j) - \bar{g}(\boldsymbol{x}_a)\|_2^2}}\right\}_{i=1}^{K}$, with a Gaussian situated on normalized anchor example in the feature space $\bar{g}(\boldsymbol{x}_a)$, M_p and M_n are the number of positive and negative examples in a mini-batch, and $\sigma^2 \in \mathbb{R}$ is the hyper-parameter to control the variance of the distribution. The soft variant of triplet loss encourages the features distribution to be robust to outliers, while increasing the inter-class variance and reducing the intra-class variance across similar examples in the shared feature space.

VI. Experiments, Results and Discussions

In this section, we first benchmark the DIAL approach on digits domains, followed by its application to vision-based grasping with a mobile robot by sim2real transfer. We empirically investigate what representations transfer better with the source data only, and the effect of a few labeled target examples in transfer learning across domains.

A. Digits Datasets

We choose four widely used digits datasets (see Fig. 4): **MNIST** [23], **MNISTM** [8], **USPS** [16], and **SVHN** [34]. We select one dataset as the source and the other as the target for domain adaptation, namely: **MNIST→MNISTM**, **MNIST→USPS**, **SVHN→MNIST**, **USPS→SVHN**, **USPS→MNIST**, **MNIST→SVHN**.

Baselines: We compare the DIAL approach with state-of-the-art methods including DANN [8], associative domain adaptation (ADA) [13], reconstruction based domain adaptation (RDA) [9], MCD [46], and FADA [32].

Results: Table I lists the average test accuracy of the target domains in unsupervised and semi-supervised setting. Methods aligning marginal distributions only such as DANN often do not perform well due to lack of conditional alignment across source domains. Metric learning with triplet loss on the source domain increases the separation between class categories, which improves the performance of marginal alignment

Fig. 4: Sample image of label 5 from MNIST, MNISTM, SVHN and USPS digits datasets.

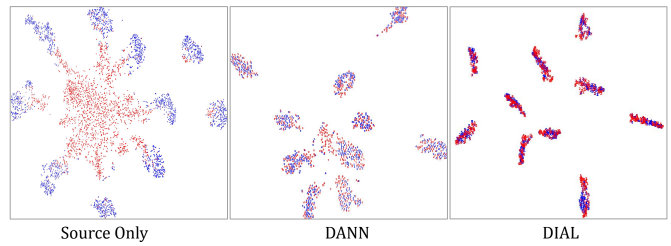

Fig. 5: T-SNE visualization of **MNIST→MNISTM** (source in blue, target in red). DIAL compactly clusters the class distributions across datasets for transfer learning in comparison to DANN and source only transfer.

methods such as DANN. Adding reconstruction loss on top of DANN to force the feature transformation to be invertible decreases the performance on the target domain [19]. This performance degradation is likely due to the additional constraints of having distinct features for each sample, making it difficult to align the marginal distributions. Associative domain adaptation imposes a cyclic loss to bring source and target examples close in the shared feature space, however, yields unreliable performance across datasets. MCD performs better across unsupervised baselines by minimizing the conditional discrepancy loss using two classifiers, however, gives unsatisfactory results with challenging adaptation situations such as **USPS→SVHN** and **MNIST→SVHN** like other unsupervised domain adaptation methods.

DIAL addresses the limitations of the existing approaches and outperforms other unsupervised and semi-supervised approaches with DANN, FADA in 1-shot, 5-shot and 10-shot scenario. DIAL uses a few target labels for effective conditional distributions alignment (see Fig. 5 for a qualitative comparison). We experimented with using pseudo-labels for target domains, but found its performance to be unstable, and/or only improve results when the source network performance on target data is already high. Consequently, the results presented here do not make use of the pseudo-labeled target data. Note that DIAL performs well across all datasets, especially on challenging problems of **USPS→SVHN** and **MNIST→SVHN** with large domain shift by aligning the marginal and the conditional distributions using only a few samples for the target class. Specifically, **MNIST → SVHN** increases 28.2% from 1-shot to 5-shot and 12.0% from 5-shot to 10-shots.

B. Vision-Based Decluttering by Sim2Real Transfer

Robot decluttering is a promising application of service robots in a broad variety of unstructured environments such as home, office and machine shops. We consider a mobile robot that observes the state of the floor as a RGB image

TABLE I: Normalized average test accuracy on target domains of Digits datasets with unsupervised and semi-supervised domain adaptation. DIAL performs well across all target domains in comparison to other baselines. *results reported from [32].

Methods		MNIST→ MNISTM	MNIST→ USPS	SVHN→ MNIST	USPS→ SVHN	USPS→ MNIST	MNIST→ SVHN
unsupervised							
RDA		0.534	0.823	0.627	0.155	0.715	0.195
DANN		0.798	0.873	0.674	0.153	0.751	0.194
Triplet		0.704	0.886	0.805	0.180	0.832	0.212
ADA		**0.895**	0.212	0.960	0.256	0.570	**0.359**
MCD		0.770	**0.941**	**0.978**	**0.288**	**0.932**	0.330
semi-supervised							
DANN	1	0.758	0.872	0.764	0.151	0.777	0.228
	5	0.763	0.920	0.793	0.252	0.850	0.353
	10	0.796	0.923	0.806	0.354	0.914	0.405
FADA*	1	–	0.891	0.728	0.275	0.811	0.377
	5		0.934	0.861	0.379	0.911	0.461
	7		0.944	0.872	0.429	0.915	0.470
DIAL	1	0.786	0.894	0.773	0.400	0.949	0.495
	5	0.941	0.946	0.864	0.610	0.945	0.683
	10	**0.948**	**0.951**	**0.903**	**0.802**	**0.962**	**0.837**

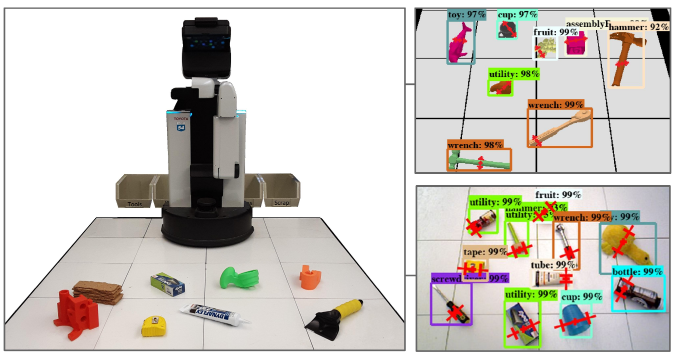

Fig. 6: *(left)* Experimental setup for decluttering objects into bins with HSR, *(right)* Object recognition and grasp planning model output on a simulated image on *(top)* and real image on *(bottom)*.

$\boldsymbol{I}_t^c \in \mathbb{R}^{640\times480\times3}$ and a depth image $\boldsymbol{I}_t^d \in \mathbb{R}^{640\times480}$. The robot recognizes the objects $\{o_i\}_{i=1}^N$ as belonging to the object categories $o_i \in \{1\ldots C\}$, and subsequently plan a grasp action $\boldsymbol{y_t} \in \mathbb{R}^4$ corresponding to the 3D object position and the planar orientation of the most likely recognized object. After grasping an object, the robot places the object into appropriate bins (see Fig. 6 for an overview).

In this work, we show the application of DIAL to learn a domain-invariant object recognition model by sim2real transfer in a semi-supervised manner. The cropped depth image from the output bounding box of the object recognition model is fed as input to the DexNet grasp planning model [52]. Depth images are readily invariant to the simulator and the real environment. The grasp planning model samples antipodal pairs on the cropped depth image of the object and outputs the top ranked grasp for the robot to pick and place the object into corresponding bin.

Simulation and Real Dataset: We simulate the decluttering environment in a Pybullet simulator. The simulated dataset comprises of $20,000$ synthetic RGB and depth images of cluttered object meshes on floor. Each image randomly contains between $5-25$ objects that are split across 12 categories, namely screwdriver, wrench, fruit, cup, bottle, assembly part, hammer, scissors, tape, toy, tube and utility. We use domain randomization to vary the background texture and color of object meshes in each image, and store the ground-truth bounding box locations, object categories, segmentation masks and grasp candidates locations.

The physical dataset comprises of 212 real RGB and depth images collected with the Toyota HSR looking at 1.2 sq. meter white tiled floor as shown in Fig. 6. We hand-label the bounding boxes and object categories similar to the synthetic classes above.

Sim2Real Transfer: We use the MobileNet-Single Shot MultiBox Detector (SSD) [27, 25] algorithm with focal loss and feature pyramids as the base model for object recognition [54]. The input image is fed to a pre-trained VGG16 network, followed by feature resolution maps and a feature pyramid network, before being fed to the output class prediction and box prediction networks. We modify the base model by adding two convolutional layers before the class logits and box prediction network for creating a shared feature across domains. We add a domain classifier and duplicate the class and box prediction network for source and target domain. The choice of the feature space and size of convolution layers is empirically selected to give better performance with the unlabeled real data.

Results: We divide the simulated and the real datasets into 60% training with 127 labeled target examples and 40% evaluation sets. Results are summarized in Table II. We observe

TABLE II: Comparative experiments for learning deep object recognition for simulation to reality transfer. Metrics include mean Average Precision (mAP) on real images, classification accuracy on synthetic test images sim_eval, real test images real_eval and silhouette score SS. DIAL performs better than other compared approaches across both domains.

Methods	mAP	sim_eval	real_eval	SS
Sim Only	0.13	95.7	26.8	0.08
Real Only	0.62	24.6	85.9	0.42
Sim + Real	0.33	95.5	70.6	0.19
Triplet	–	94.5	76.2	0.35
DANN	0.61	93.2	84.4	0.30
MCD	0.65	94.1	89.6	0.48
DIAL	**0.69**	**94.2**	**91.0**	**0.69**

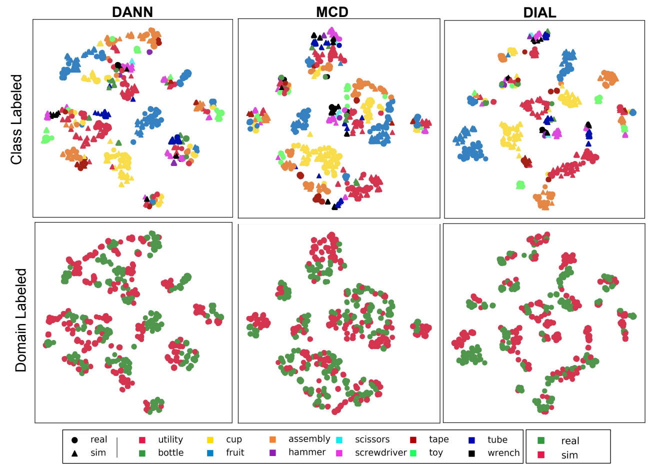

Fig. 7: T-SNE visualization of feature vectors with class labels and domain labels. DIAL effectively clusters the class categories for sim-to-real transfer with a silhouette score of 0.69 compared to 0.48 and 0.30 for MCD and DANN.

that the object recognition model trained on synthetic data only gives poor performance on real data with 26.8% accuracy, in comparison to 85.9% accuracy obtained with labeled training data. Naively combining the synthetic and real data in a mini-batch is also sub-optimal. Using triplet loss on labeled source examples preserves the structure of the features for transfer to real examples. DANN improves performance in both domains by aligning marginal distributions. MCD further improves the performance in domain adaptation with conditional alignment of distributions. DIAL outperforms the compared approaches by combining marginal and conditional distributions alignment with triplet distributions loss (see Fig. 7 for a qualitative comparison of learned features).

Decluttering with Toyota HSR: We test the performance of the trained models on the mobile Toyota HSR robot for grasping objects from floor and putting them in target bins [54]. The overall accuracy of the domain invariant object recognition and the grasping model on the robot is 90.14% and 86.85%, respectively, for a total of decluttering 185 objects across 213 grasp attempts. We observe that the robot performs well in grasping compliant objects and objects with well-defined geometry such as cylinders, screwdrivers, tape, cups, bottles, utilities, and assembly parts.

VII. Conclusion

In this paper, we have presented a domain-invariant adversarial learning approach for sim-to-real robotic manipulation. We have identified the limitations of existing unsupervised domain adaptation methods that lead to ambiguous transfer with aligning marginal distributions only. The proposed DIAL approach leverages upon a few target examples of the real environment to align the marginal and the conditional distributions of source and target domains, and uses a soft metric learning triplet loss to make the feature distributions disjoint in a shared feature space. We have presented a vision-based decluttering application, that uses sim-to-real transfer of a single-shot object recognition model to successfully declutter a wide variety of objects.

References

[1] Kamyar Azizzadenesheli, Anqi Liu, Fanny Yang, and Animashree Anandkumar. Regularized learning for domain adaptation under label shifts. *CoRR*, abs/1903.09734, 2019. URL http://arxiv.org/abs/1903.09734.

[2] Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine Learning*, 79(1): 151–175, May 2010. doi: 10.1007/s10994-009-5152-4.

[3] Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. Unsupervised pixel-level domain adaptation with generative adversarial networks. *CoRR*, abs/1612.05424, 2016.

[4] Konstantinos Bousmalis, George Trigeorgis, Nathan Silberman, Dilip Krishnan, and Dumitru Erhan. Domain separation networks. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 343–351. Curran Associates, Inc., 2016. URL http://papers.nips.cc/paper/6254-domain-separation-networks.pdf.

[5] Yevgen Chebotar, Ankur Handa, Viktor Makoviychuk, Miles Macklin, Jan Issac, Nathan D. Ratliff, and Dieter Fox. Closing the sim-to-real loop: Adapting simulation randomization with real world experience. *CoRR*, abs/1810.05687, 2018.

[6] Yi-Hsin Chen, Wei-Yu Chen, Yu-Ting Chen, Bo-Cheng Tsai, Yu-Chiang Frank Wang, and Min Sun. No more discrimination: Cross city adaptation of road scene segmenters. *CoRR*, abs/1704.08509, 2017.

[7] Yuhua Chen, Wen Li, Christos Sakaridis, Dengxin Dai, and Luc Van Gool. Domain adaptive faster R-CNN for object detection in the wild. *CoRR*, abs/1803.03243, 2018. URL http://arxiv.org/abs/1803.03243.

[8] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.*, 17(1):2096–2030, January 2016. ISSN 1532-4435.

[9] Muhammad Ghifary, W. Bastiaan Kleijn, Mengjie Zhang, David Balduzzi, and Wen Li. Deep reconstruction-classification networks for unsupervised domain adaptation. *CoRR*, abs/1607.03516, 2016.

[10] Mingming Gong, Kun Zhang, Tongliang Liu, Dacheng Tao, Clark Glymour, and Bernhard Schölkopf. Domain adaptation with conditional transferable components. In Maria Florina

Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2839–2848, New York, New York, USA, 20–22 Jun 2016. PMLR.

[11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. 2014.

[12] Abhishek Gupta, Coline Devin, Yuxuan Liu, Pieter Abbeel, and Sergey Levine. Learning invariant feature spaces to transfer skills with reinforcement learning. *CoRR*, abs/1703.02949, 2017.

[13] Philip Häusser, Thomas Frerix, Alexander Mordvintsev, and Daniel Cremers. Associative domain adaptation. *CoRR*, abs/1708.00938, 2017.

[14] Judy Hoffman, Dequan Wang, Fisher Yu, and Trevor Darrell. Fcns in the wild: Pixel-level adversarial and constraint-based adaptation. *CoRR*, abs/1612.02649, 2016.

[15] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei A. Efros, and Trevor Darrell. Cycada: Cycle-consistent adversarial domain adaptation. *CoRR*, abs/1711.03213, 2017. URL http://arxiv.org/abs/1711.03213.

[16] J. J. Hull. A database for handwritten text recognition research. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(5):550–554, 1994. doi: 10.1109/34.291440.

[17] Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks. *CoRR*, abs/1812.07252, 2018.

[18] Rae Jeong, Yusuf Aytar, David Khosid, Yuxiang Zhou, Jackie Kay, Thomas Lampe, Konstantinos Bousmalis, and Francesco Nori. Self-supervised sim-to-real adaptation for visual robotic manipulation, 2019.

[19] Fredrik D. Johansson, David Sontag, and Rajesh Ranganath. Support and invertibility in domain-invariant representations. In *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS 2019, 16-18 April 2019, Naha, Okinawa, Japan*, pages 527–536, 2019.

[20] Dmitry Kalashnikov, Alex Irpan, Peter Pastor, Julian Ibarz, Alexander Herzog, Eric Jang, Deirdre Quillen, Ethan Holly, Mrinal Kalakrishnan, Vincent Vanhoucke, and Sergey Levine. Qt-opt: Scalable deep reinforcement learning for vision-based robotic manipulation. *CoRR*, abs/1806.10293, 2018. URL http://arxiv.org/abs/1806.10293.

[21] Guoliang Kang, Lu Jiang, Yi Yang, and Alexander G. Hauptmann. Contrastive adaptation network for unsupervised domain adaptation. *CoRR*, abs/1901.00976, 2019. URL http://arxiv.org/abs/1901.00976.

[22] Durk P Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling. Semi-supervised learning with deep generative models. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 3581–3589. Curran Associates, Inc., 2014. URL http://papers.nips.cc/paper/5352-semi-supervised-learning-with-deep-generative-models.pdf.

[23] Yann LeCun and Corinna Cortes. MNIST handwritten digit database. 2010.

[24] Sergey Levine, Peter Pastor, Alex Krizhevsky, Julian Ibarz, and Deirdre Quillen. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research*, 37(4-5):421–436, 2018. doi: 10.1177/0278364917710318.

[25] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017. doi: 10.1109/iccv.2017.324. URL http://dx.doi.org/10.1109/ICCV.2017.324.

[26] Zachary C. Lipton, Yu-Xiang Wang, and Alexander J. Smola. Detecting and correcting for label shift with black box predictors. *CoRR*, abs/1802.03916, 2018.

[27] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu, and Alexander C. Berg. SSD: single shot multibox detector. *CoRR*, abs/1512.02325, 2015. URL http://arxiv.org/abs/1512.02325.

[28] M. Long, J. Wang, G. Ding, J. Sun, and P. S. Yu. Transfer feature learning with joint distribution adaptation. In *2013 IEEE International Conference on Computer Vision*, pages 2200–2207, 2013. doi: 10.1109/ICCV.2013.274.

[29] Mingsheng Long, Han Zhu, Jianmin Wang, and Michael I. Jordan. Unsupervised domain adaptation with residual transfer networks. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, pages 136–144, USA, 2016. Curran Associates Inc. ISBN 978-1-5108-3881-9. URL http://dl.acm.org/citation.cfm?id=3157096.3157112.

[30] Mingsheng Long, ZHANGJIE CAO, Jianmin Wang, and Michael I Jordan. Conditional adversarial domain adaptation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 1640–1650. 2018.

[31] Yishay Mansour, Mehryar Mohri, and Afshin Rostamizadeh. Domain adaptation: Learning bounds and algorithms. *CoRR*, abs/0902.3430, 2009.

[32] Saeid Motiian, Quinn Jones, Seyed Mehdi Iranmanesh, and Gianfranco Doretto. Few-shot adversarial domain adaptation. *CoRR*, abs/1711.02536, 2017.

[33] Saeid Motiian, Marco Piccirilli, Donald A. Adjeroh, and Gianfranco Doretto. Unified deep supervised domain adaptation and generalization. *CoRR*, abs/1709.10190, 2017.

[34] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Ng. Reading digits in natural images with unsupervised feature learning. *NIPS*, 01 2011.

[35] Fernando Camaro Nogue, Andrew Huie, and Sakyasingha Dasgupta. Object detection using domain randomization and generative adversarial refinement of synthetic images. *CoRR*, abs/1805.11778, 2018.

[36] Avital Oliver, Augustus Odena, Colin Raffel, Ekin D. Cubuk, and Ian J. Goodfellow. Realistic evaluation of deep semi-supervised learning algorithms. *CoRR*, abs/1804.09170, 2018. URL http://arxiv.org/abs/1804.09170.

[37] Deepak Pathak, Philipp Krähenbühl, and Trevor Darrell. Constrained convolutional neural networks for weakly supervised segmentation. *CoRR*, abs/1506.03648, 2015.

[38] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. *CoRR*, abs/1710.06537, 2017.

[39] L. Pinto and A. Gupta. Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3406–3413, 2016. doi: 10.1109/ICRA.2016.7487517.

[40] Joaquin Quionero-Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D. Lawrence. *Dataset Shift in Machine Learning*. The MIT Press, 2009. ISBN 0262170051, 9780262170055.

[41] Xinyi Ren, Jianlan Luo, Eugen Solowjow, Juan Aparicio Ojea, Abhishek Gupta, Aviv Tamar, and Pieter Abbeel. Domain randomization for active pose estimation. *CoRR*, abs/1903.03953, 2019.

[42] Oren Rippel, Manohar Paluri, Piotr Dollar, and Lubomir Bourdev. Metric learning with adaptive density discrimination. *arXiv*

preprint arXiv:1511.05939, 2015.
[43] Nataniel Ruiz, Samuel Schulter, and Manmohan Chandraker. Learning to simulate. *CoRR*, abs/1810.02513, 2018.
[44] Fereshteh Sadeghi and Sergey Levine. (cad)$ˆ2$rl: Real single-image flight without a single real image. *CoRR*, abs/1611.04201, 2016.
[45] Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In *Proceedings of the 11th European Conference on Computer Vision*, ECCV, pages 213–226, 2010.
[46] Kuniaki Saito, Kohei Watanabe, Yoshitaka Ushiku, and Tatsuya Harada. Maximum classifier discrepancy for unsupervised domain adaptation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
[47] Kuniaki Saito, Donghyun Kim, Stan Sclaroff, Trevor Darrell, and Kate Saenko. Semi-supervised domain adaptation via minimax entropy. *CoRR*, abs/1904.06487, 2019. URL http://arxiv.org/abs/1904.06487.
[48] Ashutosh Saxena, Justin Driemeyer, and Andrew Y. Ng. Robotic grasping of novel objects using vision. *Int. J. Rob. Res.*, 27(2): 157–173, 2008. doi: 10.1177/0278364907087172.
[49] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. *CoRR*, abs/1503.03832, 2015. URL http://arxiv.org/abs/1503.03832.
[50] Ozan Sener, Hyun Oh Song, Ashutosh Saxena, and Silvio Savarese. Learning transferrable representations for unsupervised domain adaptation. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems 29*, pages 2110–2118. 2016.
[51] Rui Shu, Hung H. Bui, Hirokazu Narui, and Stefano Ermon. A dirt-t approach to unsupervised domain adaptation, 2018.
[52] Benno Staub, Ajay Kumar Tanwani, Jeffrey Mahler, Michel Breyer, Michael Laskey, Yutaka Takaoka, Max Bajracharya, Roland Siegwart, and Ken Goldberg. Dex-Net MM: Deep Grasping for Surface Decluttering with a Low-Precision Mobile Manipulator. In *IEEE International Conference on Automation Science and Engineering (CASE)*, pages 1–7, 2019.
[53] Qian Sun, Rita Chattopadhyay, Sethuraman Panchanathan, and Jieping Ye. A two-stage weighting framework for multi-source domain adaptation. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 505–513. 2011.
[54] A. K. Tanwani, N. Mor, J. Kubiatowicz, J. E. Gonzalez, and K. Goldberg. A Fog Robotics Approach to Deep Robot Learning: Application to Object Recognition and Grasp Planning in Surface Decluttering. In *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, 2019.
[55] Joshua Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. *CoRR*, abs/1703.06907, 2017.
[56] Eric Tzeng, Judy Hoffman, Trevor Darrell, and Kate Saenko. Simultaneous deep transfer across domains and tasks. *CoRR*, abs/1510.02192, 2015.
[57] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. *CoRR*, abs/1702.05464, 2017. URL http://arxiv.org/abs/1702.05464.
[58] Mei Wang and Weihong Deng. Deep visual domain adaptation: A survey. *CoRR*, abs/1802.03601, 2018. URL http://arxiv.org/abs/1802.03601.
[59] Yifan Wu, Ezra Winston, Divyansh Kaushik, and Zachary C. Lipton. Domain adaptation with asymmetrically-relaxed distribution alignment. *CoRR*, abs/1903.01689, 2019. URL http://arxiv.org/abs/1903.01689.
[60] Markus Wulfmeier, Ingmar Posner, and Pieter Abbeel. Mutual alignment transfer learning. *CoRR*, abs/1707.07907, 2017. URL http://arxiv.org/abs/1707.07907.
[61] T. Yao, Yingwei Pan, C. Ngo, Houqiang Li, and Tao Mei. Semi-supervised domain adaptation with subspace learning for visual recognition. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2142–2150, 2015. doi: 10.1109/CVPR.2015.7298826.
[62] Dexuan Zhang and Tatsuya Harada. A general upper bound for unsupervised domain adaptation, 2019.
[63] Yuchen Zhang, Tianle Liu, Mingsheng Long, and Michael I. Jordan. Bridging theory and algorithm for domain adaptation. *CoRR*, abs/1904.05801, 2019.
[64] Han Zhao, Remi Tachet des Combes, Kun Zhang, and Geoffrey J. Gordon. On learning invariant representation for domain adaptation. *CoRR*, abs/1901.09453, 2019.
[65] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *CoRR*, abs/1703.10593, 2017. URL http://arxiv.org/abs/1703.10593.
[66] Xinge Zhu, Jiangmiao Pang, Ceyuan Yang, Jianping Shi, and Dahua Lin. Adapting object detectors via selective cross-domain alignment. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.