Open Set Domain Adaptation Methods in Deep Networks for Image Recognition

Lucas Fernando Alvarenga e Silva

Instituto de Computação
Universidade Estadual de Campinas – UNICAMP
13083-852, Campinas, SP – Brazil
Email: lucas.silva@ic.unicamp.br

Jurandy Almeida

Departamento de Computação
Universidade Federal de São Carlos – UFScar
18052-780, Sorocaba, SP – Brazil
Email: jurandy.almeida@ufscar.br

***Abstract*—Deep learning (DL) has revolutionized various fields through its remarkable capacity to learn from raw data. However, in uncontrolled environments like in the wild, the performance of these systems might degrade to some extent, especially with unlabeled datasets. Naive approaches train DL models on labeled datasets (source domains) that resemble the unlabeled test dataset (target domain), but nonetheless, this approach may not yield optimal results due to domain and category-shift problems. These issues have been the primary focus of Unsupervised Domain Adaptation (UDA) and Open Set Recognition research areas. To address the domain-shift problem, we introduced the Multi-Source Domain Alignment Layers (MS-DIAL), a structural solution for multi-source UDA. MS-DIAL aligns the source domains and the target domain at various levels of the feature space, individually achieving competitive results comparable to the state-of-the-art, and when combined with other UDA methods, it further enhances transferability by up to 30.64% in relative performance gains. Subsequently, we tackled the demanding setup of Open Set Domain Adaptation (OSDA), where both domain and category-shift issues coexist. Our proposed approach involves dealing with negatives, extracting a high-confidence set of unknown instances, and using them as a hard constraint to refine the classification boundaries of OSDA methods. We assessed our proposal in an extensive set of experiments, which achieved up to 5.8% of absolute performance gains.**

I. Introduction

In recent years, Deep Learning (DL) methods have yielded revolutionary outcomes in various computer vision research domains, mainly because of the Convolutional Neural Networks (CNNs). Nonetheless, most of these approaches have excelled in controlled environments, based on unrealistic assumptions of fully labeled data distributed into a Closed Set (CS) of categories [1]. For this reason, when such CNNs are posed to real-world problems, it is often needed to undertake expensive, time-consuming, or even impossible data labeling procedures [2], [3]. Another challenge arises during inference, where the model encounters unrestricted data instances, requiring the predictor to handle examples from unknown/unseen categories [1].

To tackle the reduced level of supervision, the common approach involves utilizing labeled datasets (source domains) similar to the unlabeled target dataset (target domain) to train a predictor. However, this strategy is not optimal. At first, the existence of *domain-shift* between the source and target domains might compromise the results, requiring the use of Unsupervised Domain Adaptation (UDA) methods to address this challenge [4]. Additionally, the use of unrestricted data during inference can introduce the problem of category-shift between the source and target domains, in which Open Set (OS) recognition methods are leveraged to discard unknown examples while accurately classifying examples that belong to known classes [1].

In this work, we address the practical and challenging scenario where both problems occur simultaneously. Specifically, we must deal with the recognition of unknown classes during inference (OS problem), while transfer knowledge from a label-rich source domain to an unlabeled target domain (UDA problem). This challenging problem was first observed by Busto and Gall [5], who termed it as Open Set Domain Adaptation (OSDA). Since then, this scenario has gained moment as a new research area with various contributions. For instance, researchers have proposed approaches based on Extreme Value Theory modeling techniques [6], self-supervised learning schemes [7], and adversarial learning [8], [9].

Historically in the literature on OSDA approaches [2], [7], it has been a common practice to divide the framework into separate parts that handle the domain-shift (*i.e.*, UDA) and category-shift (*i.e.*, OS) tasks individually. Following this convention, our investigation also comprises two main parts: the development of a UDA approach and an OSDA approach.

a) Part 1: Former UDA approaches attempted to adapt a single labeled source domain to a single unlabeled target domain by incorporating various regularization terms into the final loss function [10]. However, in cases where multiple correlated datasets are available for the target task (*e.g.*, digit recognition), each source domain can contribute with complementary information to enhance knowledge about the target domain [11], [12]. To address this, Multi-Source Unsupervised Domain Adaptation (MSDA) has been proposed, aiming to reduce domain-shift by adapting multiple source domains to a single target domain using loss terms that penalize feature discrepancies across domains [11]–[14].

In this first step, we argue that solely relying on loss

This paper summarizes the main contributions of the M.Sc. project of Lucas Fernando Alvarenga e Silva function is insufficient to tackle domain-shift effectively, thus feature alignment at various levels of the network plays a crucial role in domain adaptation. To address this, we introduce the Multi-Source version of DomaIn Alignment Layers (MS-DIAL) [4]. These layers are integrated at different levels of any given DL model to enhance the network's transferability by redesigning its architectural components. Our experimental results demonstrate the significant effectiveness of this approach, showing relative gains of up to +30.64% in classification accuracies compared to state-of-the-art MSDA methods.

b) Part 2: The conventional strategy for OSDA methods involves drawing high-confidence known and unknown sets of samples from the target domain and attempting to align the known set with the source domain, while using the unknown set as negative supervision [2], [7], [9].

However, recent studies by Liu *et al.* [15] and Baktashmotlagh *et al.* [16] highlighted that examples from unknown categories from the target domain contains valuable information with complex semantics and potential correlations to known classes, are usually oversimplified by OSDA approaches. In light of this, we hypothesize that using an unknown set of examples of the target domain might refine the boundaries of the closed-set classifier and lead to improved classification performance. To address this hypothesis, we propose three different approaches:

1) **Original approach**: In which the high-confidence unknown set of target domain is simply used to refine the OSDA CS Classification boundaries.
2) **Augmentation approach**: Applying data augmentation techniques to randomly transform negatives before using them as the classification constraint.
3) **Generation approach**: Creating negative/adversarial examples through a Generative Adversarial Network (GAN) model trained with such negatives.

We compared all three approaches, observing performance gains in most OSDA tasks, with accuracy improvements of up to 5.8% in the Office-Home benchmark.

The remainder of this paper is organized as follows. Section II discusses related work. Section III describes the first step of our work, showing the MS-DIAL method and how it can be used for improving the transferability of any MSDA model. Section IV describes the second step of this work, regarding how we dealt with negative samples to improve OSDA methods. Finally, we offer our conclusions and directions for future work in Section V.

II. Related Works

This section presents relevant works from the literature regarding the first part of UDA, summarizing both single-source methods and multi-source approaches for domain adaptation, and the second part of OSDA, describing works directed to known samples and unknown samples from the target domain.

UDA methods for DL often involve adversarial training to transform a single source of samples towards target samples or overpower the feature extractor to fool the classifier by producing features closer to the target domain, as seen in Domain-Adversarial Neural Networks (DANN) [17] and Weighted Maximum Mean Discrepancy (WMMD) [18]. Alternatively, methods like DIAL [10] and AutoDIAL [19] aim to align feature distributions by using domain-specific normalization layers in a neural network. In Multi-Source setup, most MSDA methods adopt multi-flow models [20], [21] with a single feature extraction model shared by different domain-individual classifier heads [11]–[13], or multiple feature extraction and classifiers [14] to approximate the target distributions through joint optimization with proper loss-functions.

On the other hand, OSDA aims to tackle two main challenges: handle irrelevant categories in the target domain that do not appear in the source domain (the OS problem) and deal with potential data distribution misalignment between the source and target domains (the UDA problem). Despite abundant research on OS [22]–[24] and UDA [4], [25]–[27], OSDA has been less explored.

The concept of OSDA was introduced by Busto and Gall [5] and further developed by Saito *et al.* [8]. Existing techniques in the literature range from simple adversarial methods [9] to more complex approaches involving self-supervision [2] and combinations of contrastive learning, style transfer, and prototype hyperspherical learning [7]. Recently, novel studies have shed light on the complexity of unknown examples, which are often overlooked when simply added to the classifier as new logits. UOL [15] proposed a multi-unknown detector backed by gradient-graph annotations to design an unknown oriented feature space, while [16] suggested to generate source-like negative instances using a GAN and employing them as source closed/target supervision.

III. UDA: Domain Alignment Layers

Let $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_M\}$ be a finite set of labeled source domains sharing the same set $\mathcal{Y}$ of categories with an unlabeled target domain $\mathcal{T}$. Each source domain $\mathcal{S}_i = \{(\mathbf{x}_i^j, \mathbf{y}_i^j)\}_{j=1}^{N_i}$ refers to a set of tuples composed of N_i samples $\mathbf{x}_i^j$ and their respective labels $\mathbf{y}_i^j$. Since we do not know the labels of the target domain beforehand, the set $\mathcal{T} = \{\mathbf{x}_T^j\}_{j=1}^{N_T}$ comprises the target samples only. The final goal is to learn a function $f(\mathbf{x}_T; \theta)$ defined by a set of parameters θ, that diminish the domain-shift problem and better classify the target samples $\mathbf{x}_T \in \mathcal{T}$,

To address this problem for a single source domain ($M = 1$), Carlucci *et al.* [10] proposed a solution involving the alignment of feature distributions at various levels of a neural network using DomaIn Alignment Layers (DIAL). DIAL requires only one shared model across all domains, and each layer aims to bring all domain distributions to a common superposed distribution. Minor adjustments are then made through a jointly applied linear transformation on all distributions. In our study, we extend the applicability of DIAL to handle multiple source domains ($M > 1$), introducing an approach called MS-DIAL [25] that can be used in conjunction with any off-the-shelf MSDA methods.

MS-DIAL comprises a collection of domain-specific Batch Normalization (BN) layers [28] without individual affine transformations that aim to bring all domain distributions closer to a canonical distribution. Each domain's samples pass through a dedicated domain-wide BN. In sequence, the normalized features are grouped again and undergoes a linear transformation to further refine the alignment of data distributions following the optimized α and β linear scaling/stretching parameters.

However, MS-DIAL consists of a DL building block that expects feature vectors as inputs, thus requiring it to be embedded into off-the-shelf DL method before execution. For this, we follow the procedure detailed in Algorithm 1 to include MS-DIAL in state-of-the-art MSDA approaches[1].

Algorithm 1: Automatic MS-DIAL Insertion [4]

Input: DL model without MS-DIAL
Output: DL model with MS-DIAL

```
/* The loop below iterates through
   all layers to verify if they
   contain BN layers.                  */
if backbone model has BN layers then
    foreach layer l of the backbone model do
        if l is BN Layer then
            Replace it by MS-DIAL;
        end
    end
else
    foreach layer l of the backbone model do
        if l is a convolutional layer then
            Replace it by a building block formed by
              the same convolutional layer but now
              followed by MS-DIAL;
        end
        else if l is a fully-connected layer then
            Replace it by a building block formed by
              the same fully-connected layer but now
              followed by MS-DIAL;
        end
    end
end
/* The affine parameters of the
   original BN layers, if present,
   are copied to MS-DIAL.              */
```

a) Training: During the training process, MS-DIAL expects a joint mini-batch containing samples from all source domains and samples from the target domain. These mini-batches are passed through the deep learning model, and when reaching an MS-DIAL layer, the incoming mini-batch is split into domain sets, which are normalized by the domain-specific BN layers. Subsequently, the normalized features of all the domains jointly undergo a linear transformation using the α and β parameters and are forwarded to next DL layers. At the end, we calculate the loss function as a weighted combination (λ in Equation 3) of a classification term (Equation 1) and a distribution alignment term (Equation 2). The classification term involves the widely used cross-entropy loss function, while the distribution alignment refer to minimizing a Shannon Entropy of target samples in order to force the model to decide more confidently.

$$\mathcal{L}_{\mathcal{S}}(\theta) = -\sum_{i=1}^{M}\sum_{k=1}^{N_i} \mathbf{y}_i^k \log f_i(\mathbf{x}_i^k;\theta), \tag{1}$$

$$\mathcal{L}_{\mathcal{T}}(\theta) = -\sum_{k=1}^{N_T} f_T(\mathbf{x}_T^k;\theta)\log f_T(\mathbf{x}_T^k;\theta), \tag{2}$$

$$\mathcal{L}(\theta) = \mathcal{L}_{\mathcal{S}}(\theta) + \lambda\mathcal{L}_{\mathcal{T}}(\theta) \tag{3}$$

b) Inference:: During inference, our focus is solely on classifying samples from the target domain. Therefore, we disable the source domain-specific BN layers, and directly forward the target instances to their respective BN layer, effectively utilizing MS-DIAL as a simple BN layer.

c) Experiments: We conducted a rigorous and extensive experimental evaluation to determine whether aligning domains in the feature space can enhance existing MSDA methods. For this, we followed the experimental protocol used by Wen *et al.* [13] and replicated all their results for DARN and all other reimplemented MSDA approaches, namely: DANN [17], MDAN [21], M3SDA [11], and MDMN [20]. Subsequently, we minimally modified the original code to incorporate MS-DIAL into each of the aforementioned methods and re-evaluated the results. Especially, this experimental evaluation was performed on a small-scale digit recognition task (MNIST, MNIST-M, SVHN, and Synth) and a large-scale object recognition task (Office-31 and Office-Home).

d) Results: Table I summarizes the relative gains between the *method w/o MS-DIAL* and the *method w/ MS-DIAL* evaluations for each benchmark. Roughly speaking, We could clearly notice that MS-DIAL is complementary to all the other MSDA methods, since for all the evaluated tasks its combination yielded superior results to those obtained by each method in isolation. Particularly, M3SDA presented the highest gains on the Digit recognition task. For object recognition with Office-31, in all cases the use of MS-DIAL improved the performance by a consistent margin, reaching on average, approximately, +3% of relative gains on the classification accuracy. For Office-Home, on the other hand, it showed up to 30.64% of relative gains when MS-DIAL was used alongside DANN and almost +5% when used alongside DARN, the state-of-the-art method.

IV. OSDA: Dealing with Negatives

Conventional OSDA techniques simply use thresholding strategies for rejecting unknown examples, while newer ones try aligning the target domain distribution with the source domain by incorporating target domain potentially known instances in the source-domain and unknown instances to supervise an extra logit associated to the unknown category [2], [7], [9]. Differently, our proposed approach concentrates solely on

[1] state-of-the-art methods of MSDA in 2021

TABLE I
RELATIVE GAINS (%) AMONG THE USE/DON'T USE OF MS-DIAL.

Methods	Digit Recognitions	Office-31	Office-Home
DANN [17]	+5.71%	+2.37%	+30.64%
M3SDA [11]	+7.07%	+4.54%	+25.81%
MDAN [21]	+3.04%	+2.39%	+7.49%
MDMN [20]	+4.14%	+2.69%	+7.20%
DARN [13]	+0.58%	+2.59%	+4.84%

the unknown categories. We extract highly confident unknown target instances and leverage them to enhance the classification boundaries of known categories for the OVANet approach.

OVANet, a recent Universal Domain Adaptation (UNDA) approach [27][2], has presented promising results in all closed-set, partially aligned for closed-set, partially aligned for open-set, and completely open-set UNDA contexts. This method utilizes a shared feature extractor G for either a closed-set classifier C and a set O of open-set binary classifiers to effectively learn precise class boundaries for known samples. Particularly, C is a linear model trained with cross-entropy loss, which enables it to accurately classify samples into one of the known categories L_s of the source domain based on softmax probabilities $p_c(y^k|\mathbf{x}_i)$, where $\mathbf{x}_i$ refers to the incoming instance and y^k to the model output for the k-th category. Conversely, the set $O = \{B_i\}_{i=1}^{|L_s|}$ consists of $|L_s|$ one-vs-all binary classifiers B, with each classifier B_i focused on determining the softmax probability of whether instances belong to the i-th category $(p_o(\hat{y}^k|\mathbf{x}_i))$ or not $(1-p_o(\hat{y}^k|\mathbf{x}_i))$. To achieve this, the method employs *Hard Negative Classifier Sampling* and *Open-set Entropy Minimization* loss functions, as described in Equation 2 of the paper [27].

a) Training: To enhance the known/unknown classification boundaries, our approach proposes a multi-step training process for OVANet. In the first half of the expected epochs, we train OVANet conventionally. In sequence, we employ a strengthened inference technique in order to extract a subset of highly probable negative samples $\overline{\mathbf{X}}$ from the target using an elevated threshold $(1 - p_o(\hat{y}^k|\mathbf{x}_i) > 0.9)$ and the actual knowledge of OVANet to label them as "unknown" (Figure 1). Sequentially, as depicted on Figure 2, these instances are then used to impose a novel negative supervision to fine-tune the binary classifiers in O for the remaining epochs in three different ways: *original*, *augmentation*, or *generation*. In the end, we enforce this desired effect by minimizing a new constraint that is added to the loss function (Equation 4).

$$\mathcal{L}_{ova}^{neg} = -\frac{1}{|L_s|}\sum_{k=1}^{|L_s|}\log(1-p_o(\hat{y}^k|\overline{\mathbf{x}}_i)) \quad (4)$$

Each of the three proposed approaches is performed as follows:

[2]https://github.com/VisionLearningGroup/OVANet

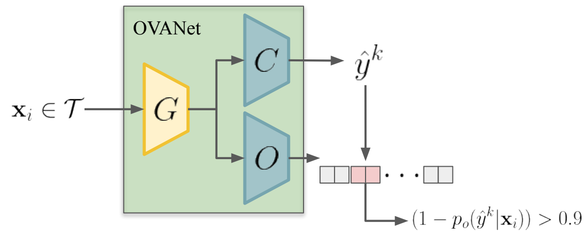

Fig. 1. The depicted procedure leverages a strengthened threshold of 0.9 to identify high-confidence unknown instances from the target domain. These instances are further used to enhance the classification boundaries of OVANet.

1) **Original** (Figure 2b): We simply use the negative samples in $\overline{\mathbf{X}}$ to adjust the binary classifiers in O.
2) **Augmentation** (Figure 2c): Before using the samples to fine tune the binary classifiers, we firstly apply data augmentation techniques to create new negative samples in $\overline{\mathbf{X}}$.
3) **Generation** We propose to use a GAN in order to generate synthetic samples resembling the negative samples of $\overline{\mathbf{X}}$. First, we suspend the training of OVANet after half of the epochs. Next, we train a GAN for a specific number of epochs, using $\overline{\mathbf{X}}$ as the training data. Once the GAN training is complete, we resume training OVANet, but this time, using the synthetically generated samples to fine-tune the binary classifiers in O. During this setup, we drive the fake samples to deceive both the discriminator and OVANet's current state, resulting in examples that align with some category in L_s according to the classifier C and are identified as "known" by the binary classifiers in O.

b) Inference: We adhere to OVANet's original inference procedure. At first, we assign an incoming instance to a pseudo-label $\hat{y}_i^t$ based on the maximum probability obtained from the classifier C. Subsequently, we determine whether the pseudo-label corresponds to "known" or "unknown" by utilizing the binary classifiers in O. If $p_o(\hat{y}_i^t|\mathbf{x}_i^t) \geq 0.5$, then $\hat{y}_i^t$ is considered the correct label; otherwise, the instance is deemed "unknown" and is rejected.

c) Experiments: We initially analyzed the probability distribution of OVANet's binary classifiers (O) for unknown and known instances from the training set of the target domain, finding that a threshold of 0.9 ensures that $\overline{\mathbf{X}}$ exclusively contains high-confidence unknown instances. Subsequently, we evaluated the *Original*, *Augmentation*, and *Generation* approaches on the Office-31 [29] and Office-Home [30] datasets, following the same experimental protocol as Saito and Saenko [27]. Specifically, the *original* approach simply forward the instances; *augmentation* employed random affine transformations (as proposed in [31]) and Gaussian blur with $\sigma = 0.1$; and the *generation* approach implemented the DCGAN proposed by Chen *et al.* [23]. In the end, we reported the Accuracy and H-Score [2] performance measures for each domain-shift scenario.

d) Results: Overally, although the results are highly dependent on the task, the use of unknown exploitation showed promising results. We highlight some results in Tables II and III from challenging domain-shifts like DSLR and Webcam for the Office-31 dataset and Clipart for the Office-Home dataset (refer to Section IV of [32] for a more in-depth discussion). Notably, in the *original* approach, we observed a 1% increase in both Accuracy and H-score for the task Amazon→DSLR on Office-31, and a 1.6% absolute gain in Accuracy for the Real-world→Clipart task on Office-Home. The *augmentation* approach showed the most favorable outcome for the Webcam→DSLR task, exhibiting absolute gains of 1.3% for both Accuracy and H-Score. and the Clipart→Product task on Office-Home showed a 1% improvement over the baseline. Lastly, the *generation* approach demonstrated enhancements of up to 1% in both Accuracy and H-score compared to the baseline on Office-31, while for the Office-Home it achieved notable progresses on Real-World→Product task, with 5.8% increase in Accuracy and 4.7% in H-score.

TABLE II
CLASSIFICATION ACCURACY (%) ACHIEVED BY OUR UNKNOWN EXPLOITATION STRATEGIES.

Benchmark	Domain-Shift	Reproduced	Improvement
Office-31	Amazon→DSLR	86.8	**87.8** (Original)
	Webcam→DSLR	97.7	**99.0** (Augmentation)
	Amazon→Webcam	86.9	**87.8** (Generation)
Office-Home	Real World→Clipart	62.5	**64.1** (Original)
	Clipart→Product	66.7	**67.5** (Augmentation)
	Real World→Product	66.1	**71.9** (Generation)

TABLE III
CLASSIFICATION H-SCORE (%) ACHIEVED BY OUR UNKNOWN EXPLOITATION STRATEGIES.

Benchmark	Domain-Shift	Reproduced	Improvement
Office-31	Amazon→DSLR	88.4	**89.3** (Original)
	Webcam→DSLR	97.8	**99.0** (Augmentation)
	Amazon→Webcam	87.4	**88.4** (Generation)
Office-Home	Real World→Clipart	**58.6**	58.3 (Original)
	Clipart→Product	65.1	**65.7** (Augmentation)
	Real World→Product	66.4	**71.1** (Generation)

V. CONCLUSION

DL models have excelled in various tasks in recent years, but their performance often suffers in uncontrolled real-world environments due to the lack of supervision. To mitigate this, one approach is to train DL models on source domains similar to the unsupervised target domain. However, this alone may not be sufficient as it does not account for potential domain-shift and category-shift problems. In this dissertation, our focus was on addressing both these issues. We divided our development into two main stages: (1) Investigating and tackling the domain-shift in isolation using a proposed UDA approach, and (2) Handling the more challenging scenario where both domain-shift and category-shift problems occur simultaneously through an OSDA approach.

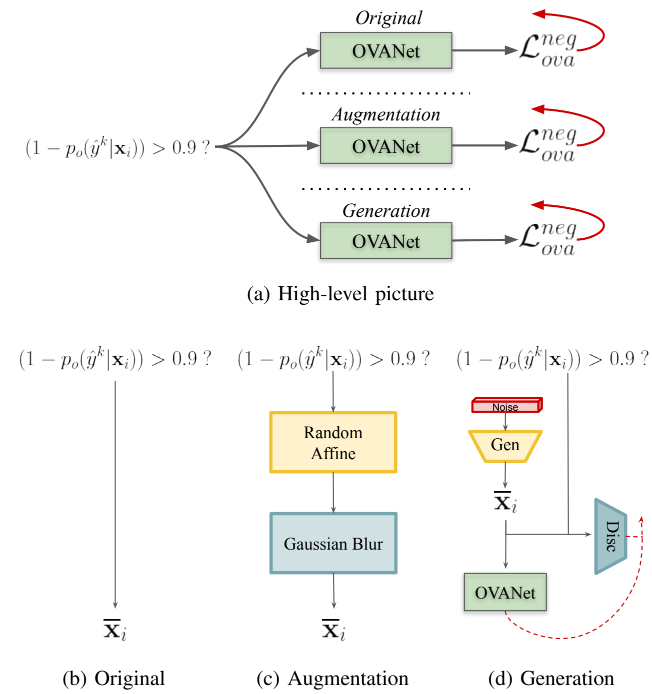

Fig. 2. Subfigure (a) shows the high-level strategy of our approach. Each of the three (b) (c) and (d) ways were evaluated in order to minimize the $\mathcal{L}_{ova}^{neg}$ novel loss function term. Particularly, (b) *original*, directly uses negative instances during the final training steps of OVANet. (c) *augmentation*, apply two affine transformations before continuing the OVANet training. (d) *generation* trains a DCGAN to replicate unknown instances while deceiving the actual OVANet. In this approach, the Gen module is optimized using errors calculated from both Disc and OVANet structures, as depicted by the red arrow in (d).

In the initial stage, we discover that aligning domains in different feature spaces yields better results compared to solely performing domain alignment on the output representations of deep models. To enhance the transferability of DL models, we propose embedding MS-DIAL [25] to align the source and target distributions in various levels of feature spaces. These layers can be seamlessly integrated into the network backbones of existing MSDA methods, leading to a substantial improvement in performance, with relative gains of up to +30.64% observed in their classification accuracies.

In the second step, we assessed three different approaches to leverage knowledge from high-confidence negative instances and enhance the boundaries of the closed-set classifier. These approaches include the *original*, *augmentation*, and *generation* negative instance techniques that were used alongside a novel disalignment loss function constraint. By employing these approaches, we observed performance improvements in most OSDA tasks, achieving absolute gains of up to 1.3% in both Accuracy and H-Score on Office-31, and 5.8% in Accuracy and 4.7% in H-Score on Office-Home.

For future work, in the first step, concerning the UDA approaches, we plan to investigate the significance of data balancing for enhancing MS-DIAL performance; the impact of using various reference distributions in MS-DIAL [10], and the assessment of transferability measures [33] before the fine-tuning procedure in DL models. In the second stage, we particularly envisage evaluating newer and more robust GANs [34] to generate higher-quality examples, visualize the classification feature space, and investigate the use of other OSDA approaches. Additionally, it would be relevant to evaluate both methods on new datasets, such as DomainNet [11].

Scientific Publications

Two scientific papers, both as first author, were published on renowned conferences. They are directly related to the contributions introduced by this dissertation, one presenting the MS-DIAL was published in [4] and another improving OSDA tasks through unknown exploitation was accepted for publication at SIBGRAPI 2023 [32]. The latter is the result of an international collaboration with the University of Trento, Italy, where a research internship was carried out. During the MSc, two other works we co-authored were developed by our research group and published in [35] and [36].

Acknowledgments

We would like to thank the IDeepS project founded by LNCC/MCTI via resources of the SDumont supercomputer; and the supervision of Prof. Dr. Nicu Sebe, and the valuable suggestions of Prof. Dr. Paolo Rota and Prof. Dr. Zhun Zhong during the author's research internship in the MHUG research group at the University of Trento, Italy. This research was supported by the São Paulo Research Foundation - FAPESP (grants 2017/25908-6, 2020/08770-3, 2021/13348-1) and the Brazilian National Council for Scientific and Technological Development - CNPq (grants 314868/2020-8).

References

[1] C. Geng, S. Huang, and S. Chen, "Recent advances in open set recognition: A survey," *TPAMI*, vol. 43, no. 10, pp. 3614–3631, 2021.

[2] S. Bucci, M. R. Loghmani, and T. Tommasi, "On the effectiveness of image rotation for open set domain adaptation," in *ECCV*, 2020, pp. 422–438.

[3] C. Saltori, P. Rota, N. Sebe, and J. Almeida, "Low-budget label query through domain alignment enforcement," *Computer Vision and Image Understanding*, vol. 222, p. 103485, 2022.

[4] L. F. A. Silva, D. C. G. Pedronette, F. A. Faria, J. P. Papa, and J. Almeida, "Improving transferability of domain adaptation networks through domain alignment layers," in *SIBGRAPI*, 2021, pp. 168—-175.

[5] P. P. Busto and J. Gall, "Open set domain adaptation," in *ICCV*, 2017, pp. 754–763.

[6] Y. Xu and D. Klabjan, "Open set domain adaptation by extreme value theory," *CoRR*, vol. abs/2101.02561, 2021.

[7] S. Bucci, F. C. Borlino, B. Caputo, and T. Tommasi, "Distance-based hyperspherical classification for multi-source open-set domain adaptation," in *WACV*, 2022, pp. 1030–1039.

[8] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "Open set domain adaptation by backpropagation," in *ECCV*, 2018, pp. 156–171.

[9] S. Rakshit, D. Tamboli, P. S. Meshram, B. Banerjee, G. Roig, and S. Chaudhuri, "Multi-source open-set deep adversarial domain adaptation," in *ECCV*, 2020, pp. 735–750.

[10] F. M. Carlucci, L. Porzi, B. Caputo, E. Ricci, and S. R. Bulò, "Just DIAL: domain alignment layers for unsupervised domain adaptation," in *ICIAP*, 2017, pp. 357–369.

[11] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, "Moment matching for multi-source domain adaptation," in *ICCV*, 2019, pp. 1406–1415.

[12] R. Xu, Z. Chen, W. Zuo, J. Yan, and L. Lin, "Deep cocktail network: Multi-source unsupervised domain adaptation with category shift," in *CVPR*, 2018, pp. 3964–3973.

[13] J. Wen, R. Greiner, and D. Schuurmans, "Domain aggregation networks for multi-source domain adaptation," in *ICML*, 2020, pp. 10 927–10 937.

[14] S. Zhao, G. Wang, S. Zhang, Y. Gu, Y. Li, Z. Song, P. Xu, R. Hu, H. Chai, and K. Keutzer, "Multi-source distilling domain adaptation," in *AAAI*, 2020, pp. 12 975–12 983.

[15] J. Liu, X. Guo, and Y. Yuan, "Unknown-oriented learning for open set domain adaptation," in *ECCV*, 2022, pp. 334–350.

[16] M. Baktashmotlagh, T. Chen, and M. Salzmann, "Learning to generate the unknowns as a remedy to the open-set domain shift," in *WACV*, 2022, pp. 3737–3746.

[17] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. S. Lempitsky, "Domain-adversarial training of neural networks," *JMLR*, vol. 17, pp. 59:1–59:35, 2016.

[18] H. Yan, Y. Ding, P. Li, Q. Wang, Y. Xu, and W. Zuo, "Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation," in *CVPR*, 2017, pp. 945–954.

[19] F. M. Carlucci, L. Porzi, B. Caputo, E. Ricci, and S. R. Bulò, "Autodial: Automatic domain alignment layers," in *ICCV*, 2017, pp. 5077–5085.

[20] Y. Li, M. Murias, G. Dawson, and D. E. Carlson, "Extracting relationships by multi-domain matching," in *NeurIPS*, 2018, pp. 6799–6810.

[21] H. Zhao, S. Zhang, G. Wu, J. M. F. Moura, J. P. Costeira, and G. J. Gordon, "Adversarial multiple source domain adaptation," in *NeurIPS*, 2018, pp. 8568–8579.

[22] L. Neal, M. Olson, X. Fern, W.-K. Wong, and F. Li, "Open set learning with counterfactual images," in *ECCV*, 2018, pp. 620–635.

[23] G. Chen, P. Peng, X. Wang, and Y. Tian, "Adversarial reciprocal points learning for open set recognition," *TPAMI*, vol. 44, no. 11, pp. 8065–8081, 2022.

[24] S. Vaze, K. Han, A. Vedaldi, and A. Zisserman, "Open-set recognition: A good closed-set classifier is all you need," in *ICLR*, 2022.

[25] L. F. A. Silva and J. Almeida, "Ms-dial: Multi-source domain alignment layers for unsupervised domain adaptation," in *WVC*, 2020, pp. 111–116.

[26] K. Saito, D. Kim, S. Sclaroff, and K. Saenko, "Universal domain adaptation through self-supervision," in *NeurIPS*, 2020.

[27] K. Saito and K. Saenko, "Ovanet: One-vs-all network for universal domain adaptation," in *ICCV*, 2021, pp. 9000–9009.

[28] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *ICML*, 2015, pp. 448–456.

[29] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in *ECCV*, 2010, pp. 213–226.

[30] H. Venkateswara, J. Eusebio, S. Chakraborty, and S. Panchanathan, "Deep hashing network for unsupervised domain adaptation," in *CVPR*, 2017, pp. 5385–5394.

[31] G. French, M. Mackiewicz, and M. Fisher, "Self-ensembling for visual domain adaptation," in *ICLR*, 2018.

[32] L. F. A. Silva, N. Sebe, and J. Almeida, "Tightening classification boundaries in open set domain adaptation through unknown exploitation," in *SIBGRAPI*, 2023, pp. 1–6.

[33] A. Agostinelli, J. R. R. Uijlings, T. Mensink, and V. Ferrari, "Transferability metrics for selecting source model ensembles," in *CVPR*, 2022, pp. 7926–7936.

[34] Z. Wang, Q. She, and T. E. Ward, "Generative adversarial networks in computer vision: A survey and taxonomy," *ACS*, vol. 54, no. 2, pp. 37:1–37:38, 2022.

[35] G. C. Rocha, H. M. Paiva, D. G. Sanches, D. Fiks, R. M. Castro, and L. F. A. Silva, "Information system for epidemic control: a computational solution addressing successful experiences and main challenges," *Library Hi Tech*, vol. 39, no. 3, pp. 834–854, 2021.

[36] M. Miranda, L. F. A. e Silva, S. Santos, V. S. Jr, T. Körting, and J. Almeida, "A high-spatial resolution dataset and few-shot deep learning benchmark for image classification," in *SIBGRAPI*, 2022, pp. 19–24.