⊘ Source Models Leak What They Shouldn't ↛: Unlearning Zero-Shot Transfer in Domain Adaptation Through Adversarial Optimization

Arnav Devalapally[1,2,*,†], Poornima Jain[1,*], Kartik Srinivas[1,3,†], and Vineeth N. Balasubramanian[1,4]

[1]Indian Institute of Technology, Hyderabad [2]University of Michigan [3]Carnegie Mellon University [4]Microsoft Research

darnav@umich.edu, {ai24resch11002@, vineethnb@cse.}iith.ac.in, kartiksr@cs.cmu.edu, vineeth.nb@microsoft.com

Abstract

*The increasing adaptation of vision models across domains, such as satellite imagery and medical scans, has raised an emerging privacy risk: models may inadvertently retain and leak sensitive source-domain specific information in the target domain. This creates a compelling use case for machine unlearning to protect the privacy of sensitive source-domain data. Among adaptation techniques, source-free domain adaptation (SFDA) calls for an urgent need for machine unlearning (MU), where the source data itself is protected, yet the source model exposed during adaptation encodes its influence. Our experiments reveal that existing SFDA methods exhibit strong zero-shot performance on source-exclusive classes in the target domain, indicating they inadvertently leak knowledge of these classes into the target domain, even when they are not represented in the target data. We identify and address this risk by proposing an MU setting called SCADA-UL: **U**n**l**earning **S**ource-exclusive **C**l**A**sses in **D**omain **A**daptation. Existing MU methods do not address this setting as they are not designed to handle data distribution shifts. We propose a new unlearning method, where an adversarially generated forget class sample is unlearned by the model during the domain adaptation process using a novel rescaled labeling strategy and adversarial optimization. We also extend our study to two variants: a continual version of this problem setting and to one where the specific source classes to be forgotten may be unknown. Alongside theoretical interpretations, our comprehensive empirical results show that our method consistently outperforms baselines in the proposed setting while achieving retraining-level unlearning performance on benchmark datasets. Our code is available at https://github.com/D-Arnav/SCADA.*

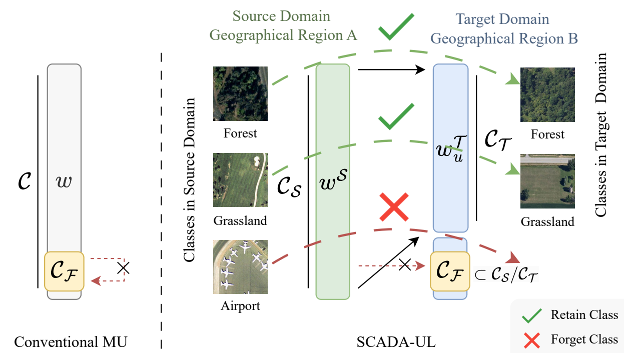

Figure 1. **Comparison of Conventional MU and Proposed SCADA-UL.** Conventional class-wise machine unlearning focuses on forgetting a subset of classes from a trained model. On the other hand, the proposed SCADA-UL aims to remove knowledge of source-exclusive classes (classes absent in target domain) while adapting a model to a new domain. For instance, if a land-use categorization model is adapted to a new geography, sensitive classes such as airports must not be transferred to the target domain.

1. Introduction

The widespread use of large-scale learning models trained on vast corpora of data has raised significant concerns in data ownership, copyrighting and privacy in recent times [26]. Recent legislation [6, 18] demands user data to be deleted on request, including any impact that the data may have on the output of a trained model. To address these concerns, Machine Unlearning (MU) has assumed increased importance, especially because trivial solutions such as retraining the model from scratch without the sensitive data may be expensive, slow or simply infeasible due to non-availability of the training data.

MU efforts over the last few years have broadly focused on methods to forget a subset of the training dataset [15, 21] or specific classes of data [4, 10, 44, 51, 62] in a given domain. These efforts have shown promising results and have validated the use of MU methods for protecting sensitive data as well as discarding unwanted data. However, finetuning and adaptation of models to different domains has played a key role in adoption of learning models for many years now, and unlearning in such non-stationary settings has seen little effort hitherto. We focus on addressing this need herein.

Why unlearning in domain adaptation? MU in domain adaptation (DA) settings has important real-world applica-

*Equal Contribution.

† Majority of work done at Indian Institute of Technology, Hyderabad

tions, especially in scenarios where the model is adapted from a source domain having classes with sensitive information. Consider a land-use categorization application (as in Fig 1) where a model is adapted to a new geography but sensitive regions such as government facilities, army regions, airports and other private land categories must not be transferred to the target domain. Similarly, in a fraud surveillance application, a model being adapted between two different environments (say, two countries) may need to forget certain classes due to legal restrictions. For another example, consider a disease diagnosis model initially trained on a dataset containing both mental and physical health conditions. If this model is later deployed in a hospital setting where patient privacy policies prohibit the use of mental health data, it becomes mandatory to forget the mental health-specific classes. Retaining such classes, even if their outputs are suppressed, poses a privacy risk. Internal representations might encode sensitive features associated with sensitive classes, inadvertently revealing protected information when processing new inputs. This can lead to unintended information leakage or even re-identification risks.

Image classifiers are particularly vulnerable to model inversion attacks [17, 36] that reconstruct inputs from confidence scores. [36] lists other common privacy leaks in classification models such as: (i) membership inference attack that reveals whether a person's record was in training; (ii) attribute inference that infers hidden/sensitive features from outputs; (iii) gradient leakage during training in federated/distributed settings that can reveal pixel-accurate images and labels [70]; and (iv) black-box model extraction (stealing a classifier via its prediction API). Therefore, merely masking or suppressing classifier outputs may not constitute true unlearning. Robust defenses must ensure minimal (ideally zero) residual traces of sensitive information in the adapted model parameters, an essential requirement for compliance with data protection regulations such as HIPAA and the GDPR.

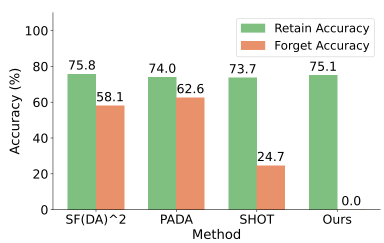

Figure 2. Existing SFDA/PDA methods leak source-exclusive classes in the target domain.

In this work, we propose a motivated methodology towards MU for source-free domain adaptation (SFDA) settings, where a model is adapted from a source domain to a target domain, and certain source-exclusive classes have to be unlearned, with no access to the source-domain data itself. This is a challenging, practical and new setting – which we call **Un**learning **S**ource-exclusive **Cl**A**sses in **D**omain **A**daptation (SCADA-UL) (Fig 1) – that has seen little concerted effort so far. Our empirical studies show that existing SFDA and Partial Domain Adaptation (PDA) methods [5, 28, 34] exhibit strong zero-shot performance on

Table 1. **Comparison of our setting with existing efforts**

Existing Efforts	Unlearning	Adaptation	Forget Data-free
SFDA/PDA [5, 28, 34]	✗	✓	N/A
MU [15, 21]	✓	✗	✗
Zero Glance MU [10, 51]	✓	✗	✓
SCADA-UL	✓	✓	✓

source-exclusive classes in the target domain, indicating they inadvertently leak knowledge of these classes into the target domain, even when they are not represented in the target data (Fig 2). We note that our setting is stricter than PDA: while PDA also mitigates the transfer of source-exclusive classes to the target domain, SCADA-UL explicitly requires *complete unlearning* of these classes. We compare against PDA methods in our experiments. Our method is also applicable in the Open Set Domain Adaptation setting, which we also explore in our empirical studies.

Existing MU methods typically require access to forget data for unlearning (Table 1), and therefore are not readily suitable for this task. Few MU methods do not require access to forget data [2, 10, 51]; however, these are not designed for domain-adapted models. Our experiments (see Appendix) reveal that applying such data-free MU methods in our setting results in poor performance, since they were not designed to handle a shift in the data distribution. We propose a new strategy to address SCADA-UL based on adversarial optimization, where an adversarially generated sample from a source-exclusive class to be unlearned (henceforth called a *forget class*) is thoroughly unlearned by the model during domain adaptation through a novel rescaled labeling strategy. Our analysis and empirical results show that this approach achieves effective unlearning in SFDA by progressively erasing representations of the forget classes. We also introduce two variants: Continual SCADA-UL (C-SCADA-UL), which addresses scenarios where classes must be forgotten across multiple unlearning requests, and Unknown Class SCADA-UL (UC-SCADA-UL), which deals with cases where source-exclusive classes are not known.

Our key contributions are summarized as follows: (i) We propose a practically relevant MU setting called SCADA-UL to address unlearning of source-exclusive classes during DA; (ii) We propose a new strategy to address SCADA-UL using adversarial optimization, where an adversarially generated forget class sample is unlearned by the model during DA through a novel rescaled labeling strategy; (iii) Our comprehensive suite of experiments across multiple datasets shows that the proposed method achieves retraining-level performance, outperforming all baselines in the proposed setting; (iv) We extend our work to two variants: C-SCADA-UL and UC-SCADA-UL, wherein our method shows promising performance; and (v) We also analyze the conceptual intuition behind our method and carry out ablation studies to study the impact of different design choices in our framework.

2. Related Work

Machine Unlearning. Machine unlearning (MU) is the process of forgetting samples or entire classes of data from a trained model. Existing MU methods for classifiers can be broadly categorized into exact or approximate unlearning [33, 43, 61]. Exact unlearning is achieved by efficiently retraining the model without using the forget data [63], while approximate unlearning methods aim to remove the influence of forget data without retraining; influence function-based methods estimate and remove the influence of the data to be removed (forget data) from the model weights [24, 37, 50, 58], gradient update-based methods perform gradient ascent on forget data [42, 59], model optimization-based methods finetune the model using different losses for forget and retain data [13, 21]. Other recent approaches to unlearning include bad-teacher and stochastic-teacher models, post-hoc dampening, and source-free unlearning methods [2, 7, 9, 16, 35, 56, 68]. MU has also been studied for Large Language Models [19, 66], federated learning [39, 69], generative models [40] and graph-based models [8]. Most methods however require access to the forget data, making them inapplicable in our setting where the source-exclusive class data is not available. Although some recent efforts [10, 20, 51] relax this requirement, they do not address unlearning in the context of domain adaptation.

Source-Free Domain Adaptation. Source-free domain adaptation (SFDA) [14], aims to adapt a model pre-trained on a source domain to a new target domain with access to only unlabeled target domain data. Finetuning-based SFDA methods such as SHOT [34] assign pseudo-labels to images based on the source model's prediction and refine the model in a self-supervised fashion in the target domain. [52] utilize target-sample selection to refine pseudo-labels for improved SFDA. [1, 38, 55] perform contrastive learning for SFDA. Few methods [12, 30, 31] use target data to update statistical information from the source model (for e.g., in batch norm layers), while minimizing the distance between source and target distributions. Clustering-based methods [32, 41] perform clustering in the target domain and update the source model through the cluster-assigned pseudo-labels. A recent method, SF(DA)^2 [28], uses spectral neighborhood clustering on the data augmentation graph formed by target data samples in the feature space of the source model, and represents one of the state-of-the-art for SFDA methods. Another recent method, UCon-SFDA [60], leverages uncertainty control in SFDA to further the SOTA in the field. Since source-free domain adaptation methods typically align the source and target domains globally, they transfer knowledge of all source classes to the target domain (as we show later). They do not address unlearning, thus necessitating this effort for simultaneous domain adaptation and unlearning.

Partial Domain Adaptation. Partial Domain Adaptation (PDA) addresses a related setting in domain adaptation where the target domain label space is a subset of the source domain label space. PDA methods use techniques such as suppressing source-exclusive classes through class- or instance-level reweighting [5, 23], attention-based entropy minimization [25] and alignment of source and target distributions [48, 64, 67]. Their main objective is to alleviate negative transfer of source-exclusive classes while improving target domain performance. These methods require access to source data (we include adaptations of such methods in our empirical studies for completeness).

3. SCADA Unlearning: Problem Setting

Preliminaries and Notations. Let x_i denote a training sample, and y_i be its corresponding label. In the domain adaptation context, let $\mathcal{D}^{\mathcal{S}} = \{(x_i^{\mathcal{S}}, y_i^{\mathcal{S}})\}_{i=1}^{n_s}$ denote the source domain dataset, and $\mathcal{D}^{\mathcal{T}} = \{(x_i^{\mathcal{T}})\}_{i=1}^{n_t}$ denote the target domain dataset. The set of source domain classes is denoted by $\mathcal{C}_{\mathcal{S}}$, and the set of target classes, assumed to be a subset of source classes in our setting, is denoted by $\mathcal{C}_{\mathcal{T}}$. The set of source-exclusive classes (or 'forget') classes is denoted by $\mathcal{C}_{\mathcal{F}} = \mathcal{C}_{\mathcal{S}} \setminus \mathcal{C}_{\mathcal{T}}$. A single forget class is denoted by $c_{\mathcal{F}} \in \mathcal{C}_{\mathcal{F}}$, and its complement, is denoted by by $c_{\mathcal{R}} = \mathcal{C}_{\mathcal{S}} \setminus c_{\mathcal{F}}$. The subset of source data where labels belong to $\mathcal{C}_{\mathcal{F}}$ is the source forget set $\mathcal{D}_f^{\mathcal{S}}$. Its complement, the source retain set is denoted by $\mathcal{D}_r^{\mathcal{S}} = \mathcal{D}^{\mathcal{S}} \setminus \mathcal{D}_f^{\mathcal{S}}$. To evaluate the performance of our method, we also introduce a target forget set $\mathcal{D}_f^{\mathcal{T}}$ on which our domain-adapted model is expected to perform poorly. The corresponding target retain set is given by $\mathcal{D}_r^{\mathcal{T}}$. We denote the source classifier training algorithm by $\mathcal{A}(\cdot)$, domain adaptation training algorithm by $\mathcal{B}(\cdot)$, and the unlearning algorithm by $\mathcal{U}(\cdot)$.

Proposed Problem Setting. As stated earlier, we introduce an unlearning setting within the source-free domain adaptation context called **Un**l**e**arning **S**ource-exclusive **C**l**A**sses in **D**omain **A**daptation, abbreviated to *SCADA-UL*. At a high level, this problem setting, similar to unlearning, seeks to produce a post-unlearning model that behaves similar to one never exposed to forget data. However, this setting presents some unique challenges: adaptation to a target domain, the absence of source data $\mathcal{D}^{\mathcal{S}}$, and the absence of target forget data $\mathcal{D}_f^{\mathcal{T}}$. While unavailability of source data is also an issue in SFDA, the non-availability of forget data in the target domain that is motivated by practical use case settings makes this problem challenging and non-trivial. Based on the definition of MU in [61], we formally define SCADA-UL as:

Definition 1. (SCADA-UL). Given a source model $w^{\mathcal{S}} = \mathcal{A}(\mathcal{D}^{\mathcal{S}})$, SCADA-UL is a process $\mathcal{U} : \{w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}}, \mathcal{C}_{\mathcal{F}}\} \rightarrow w_u^{\mathcal{T}}$ that produces an unlearned, adapted model $w_u^{\mathcal{T}}$.

In other words, $\mathcal{U}(w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}}, \mathcal{C}_{\mathcal{F}})$ maps the trained source model, target dataset, and forget classes to a model $w_u^{\mathcal{T}}$ that behaves as though it were adapted from a source model not

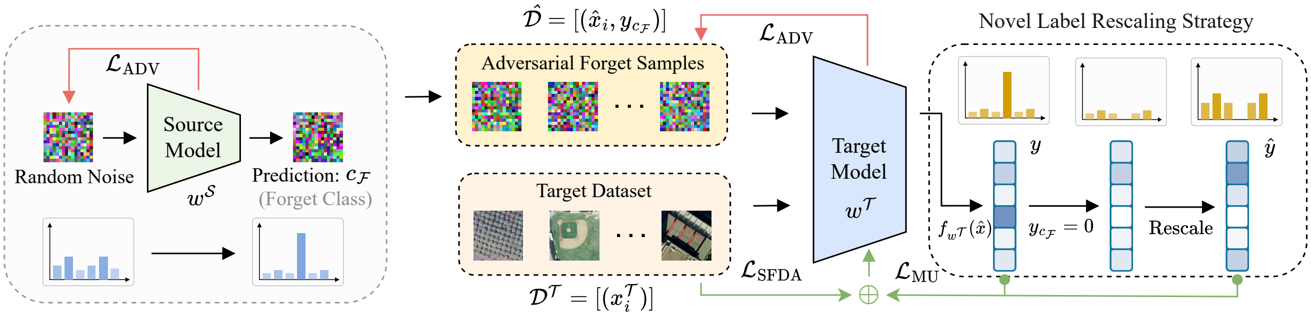

Figure 3. **Our overall method.** We adapt a source trained model $w^\mathcal{S}$ to a target domain $\mathcal{D}^\mathcal{T}$ while forgetting classes $c_\mathcal{F}$. We first create an initial adversarial sample $\hat{x}$ by *maximizing* its probability of belonging to the forget classes $c_\mathcal{F}$ by optimizing $\mathcal{L}_{\text{ADV}}(w^\mathcal{S}, \hat{x})$ (Eqn 3). In each subsequent iteration, the model $w^\mathcal{T}$ *minimizes*: (i) $\mathcal{L}_{\text{MU}}$ (5) using rescaled labels $\hat{y}$ to encourage unlearning the forget classes $c_\mathcal{F}$; and (ii) $\mathcal{L}_{\text{SFDA}}$ to simultaneously adapt to the target domain. Before the end of each iteration, the adversarial sample $\hat{x}$ is re-optimized to *maximize* its probability of belonging to the forget classes by maximizing -$\mathcal{L}_{\text{ADV}}$.

trained on $\mathcal{D}_f^\mathcal{S}$. A trivial solution here is to retrain the source model from scratch on $\mathcal{D}_r^\mathcal{S}$ and adapt it to the target domain using $\mathcal{D}_r^\mathcal{T}$, i.e. $\mathcal{B}(w_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$ where $w_r^\mathcal{S} = \mathcal{A}(\mathcal{D}_r^\mathcal{S})$. However, this is not possible due to unavailability of $\mathcal{D}^\mathcal{S}$ (we, however, use this as an upper bound in our experiments). On the other hand, naively domain adapting $\mathcal{B}(w^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$ would result in the model still containing information about the forget classes. This is used as an additional baseline.

We additionally define two variants of SCADA-UL under other constraints. In UC-SCADA-UL, we extend Defn 1 to scenarios where the forget classes $\mathcal{C}_\mathcal{F}$ are unknown, but their cardinality $|\mathcal{C}_\mathcal{F}|$ is known. Similarly, Continual SCADA-UL addresses Defn 1 in continual settings where disjoint forget class sets arrive in multiple steps *i*, denoted by ${\mathcal{C}_\mathcal{F}}^i$. Formal definitions are provided in the appendix.

4. Proposed Methodology

SCADA-UL requires approximating the behavior of a source model trained without forget data and adapted to the target domain. We model this task as obtaining a maximum a posteriori (MAP) estimate $p(w \mid w_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$, where $w_r^\mathcal{S}$ denotes the source model trained only on $\mathcal{D}_r^\mathcal{S}$. This posterior cannot be directly evaluated as it requires retraining the source model on the source retain dataset $\mathcal{D}_r^\mathcal{S}$, which is unavailable in our setting. To derive tractable objectives, we expand $p(w \mid \mathcal{D}_r^\mathcal{S}, {\mathcal{D}_f}^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$ using Bayes' rule: $p({\mathcal{D}_f}^\mathcal{S} \mid w, \mathcal{D}_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \cdot p(w \mid \mathcal{D}_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$. This can be written as: $p(w \mid \mathcal{D}_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \propto p(w \mid \mathcal{D}^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \,/\, p({\mathcal{D}_f}^\mathcal{S} \mid w)$ by rearranging and conditional independence (cf. Eq. 6 in [44]). Using model approximations gives us: $p(w \mid w_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \propto p(w \mid w^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \,/\, p({\mathcal{D}_f}^\mathcal{S} \mid w)$ (cf. Eq. 3 in [46]). We revisit our assumptions herein at the end of the section. Replacing the terms and maximizing the logarithm on both sides leads to our unlearning and adaptation objectives below.

$$\max_w \log p(w \mid w_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \equiv \max_w \big[\underbrace{\log p(w \mid w^\mathcal{S}, \mathcal{D}_r^\mathcal{T})}_{\text{Adaptation}} - \underbrace{\log p(\mathcal{D}_f^\mathcal{S} \mid w)}_{\text{Unlearning}}\big] \quad (1)$$

The left-hand side captures the ideal SCADA-UL objective in Expr 1; however, as discussed, this is infeasible to compute. The right-hand side represents an equivalent objective with similar source training and adaptation terms. As $w^\mathcal{S}$ is available in our setting, $\log p(w \mid w^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$ can be maximized by adapting $w^\mathcal{S}$ to the target domain. The other term on the RHS $-\log p(\mathcal{D}_f^\mathcal{S} \mid w)$ corresponds to unlearning. Implementing this term is non-trivial as both $\mathcal{D}_f^\mathcal{S}$ and $\mathcal{D}_f^\mathcal{T}$ are unavailable, and this motivates our method below.

4.1. Adversarial Optimization for SCADA-UL

Expr 1 showed that achieving SCADA-UL involves balancing both adaptation and unlearning objectives. One could achieve this objective by applying an existing data-free MU method on an adapted model or a source model before adaptation. However, existing data-free MU methods are designed for unlearning in a stationary single-domain setting and our empirical results reveal that they perform poorly when used both before and after adaptation (see Appendix). Thus, we propose carrying out both unlearning and adaptation processes simultaneously to achieve the proposed objective. The unlearning and adaptation processes, though, are misaligned goals - the unlearning process aims to avoid transfer of source-exclusive classes to the target domain, while the adaptation process tries to transfer source domain classes to the target domain. Hence, to achieve these simultaneously, the method must design objectives that are minimally conflicting (we elaborate this in our studies in Sec. 6). In particular, we use a novel rescaled labeling strategy (Eq. 4) in the unlearning process that helps reduce this conflict.

A key question in our SCADA-UL setting remains: *How to identify representative samples for a forget class $c_\mathcal{F}$ without access to its data*? We address this using adversarial sample generation. (We call a sample *adversarial*, since the model is expected to minimize its probability while maximizing its belongingness to the forget class it represents.) A second question that arises is: *Once representative samples for forget classes are identified, how can the model unlearn these samples while adapting simultaneously in a way that* *involves minimal conflict?* The opposing objectives of forget-class sample generation and unlearning lead to a minmax optimization formulation given by:

$$\min_{w^{\mathcal{T}}} \max_{\hat{x}} \left[\log p(\hat{\mathcal{D}} \mid w^{\mathcal{T}}) - \log p(w^{\mathcal{T}} \mid w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}})\right] \quad (2)$$

where $\hat{\mathcal{D}} = \{(\hat{x}_i, y_{c_{\mathcal{F}}})\}_{i=1}^{N_{\text{adv}}}$ and N_{adv} is the number of representative samples generated. This objective can be broken down into sample generation, unlearning, and adaptation steps which we detail below. For simplicity of explanation, without loss of generality, we focus our discussion on unlearning a single class $c_{\mathcal{F}}$, although this could be used to subsume a set of classes (or extended to multiple classes one after the other as described at the end of this section).

Generating Adversarial Samples. Generating representative samples in data-free settings has been studied in earlier work in other contexts [10]. In our case, we aim to generate inputs that elicit confident predictions from the model, thereby producing representative samples for the forget class. This is achieved by using the cross-entropy loss as our adversarial sample generation loss, $\mathcal{L}_{\text{ADV}}$, i.e.:

$$\mathcal{L}_{\text{ADV}}(w^{\mathcal{T}}, \hat{x}) = \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}, c_{\mathcal{F}}) \quad (3)$$

giving $\min_{\hat{x}} \mathcal{L}_{\text{ADV}}(w^{\mathcal{T}}, \hat{x}) \equiv \max_{\hat{x}} p(\hat{\mathcal{D}} \mid w^{\mathcal{T}})$. Our unlearning step (see Algorithm 1) starts by generating the adversarial samples $\hat{x}$ with a minimization of $\mathcal{L}_{\text{ADV}}$ with $w^{\mathcal{T}}$ initialized to the source model $w^{\mathcal{S}}$. We also minimize $\mathcal{L}_{\text{ADV}}$ *throughout the domain adaptation iterations*, allowing the generated samples to evolve alongside the adapting model.

Adapting and Unlearning with Minimal Conflict via Rescaled Labeling. Consider an adversarial sample $\hat{x}$ generated based on Expr. 3, i.e. $\hat{x}$ is representative of class $c_{\mathcal{F}}$. Let the output softmax distribution over all classes, $f_{w^{\mathcal{T}}}(\hat{x})$, for the sample be $y \in \mathbb{R}^d$ where d is the number of classes (retain and forget). Naturally, it would have the highest probability value for $c_{\mathcal{F}}$, as the sample is representative of the forget class. To promote unlearning, we propose exposing the model to false information $\mathcal{I}$ that the forget sample belongs to a different class $c_{\mathcal{R}}$ other than the forget class $c_{\mathcal{F}}$, i.e., there exists a class $c \in c_{\mathcal{R}}$, $c \neq c_{\mathcal{F}}$ with $y_c = 1$. From Bayes' Theorem, $p_{w^{\mathcal{T}}}(y = c \mid \mathcal{I}) \propto p(\mathcal{I} \mid y = c) \cdot p(y = c)$. Given $p(\mathcal{I} \mid y_{c_{\mathcal{F}}} = 1) = 0$ (since it contradicts the prior), and $p(\mathcal{I} \mid y_{c_{\mathcal{R}}}) = 1$, we obtain the final renormalized distribution as:

$$\hat{y} = p(y = c \mid \mathcal{I}) = \begin{cases} 0 & \text{if } i = c_{\mathcal{F}} \\ \frac{y_i}{\sum_{j \in c_{\mathcal{R}}} y_j} & \text{if } i \in c_{\mathcal{R}} \end{cases} \quad (4)$$

Equation (4) represents the proposed rescaled labeling we use for unlearning. This labeling strategy intuitively represents the ideal alternative label for the adversarial sample if it did not belong to the forget class. It redistributes the probabilities in proportion with the model's output for the input sample, hence causing minimal conflict with adaptation. Further, our analysis shows that this labeling strategy: (1) outperforms other strategies such as uniform or random labeling which either cause catastrophic forgetting or subpar unlearning and adaptation performance (Table 6); (2) aligns the post-unlearning model's output distributions more closely to that of a retrained model (see Appendix); and (3) results in larger gradient updates to the weights associated with retain classes compared to forget classes, thus achieving the unlearning objective (see Theorem 1 in Sec. 4.2). Using these labels, our MU loss term $\mathcal{L}_{\text{MU}}$ applies the cross entropy loss: $\mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y}) = \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}, \hat{y})$, giving us:

$$\min_{w^{\mathcal{T}}} \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y}) \equiv \min_{w^{\mathcal{T}}} p(\hat{\mathcal{D}} \mid w^{\mathcal{T}}) \quad (5)$$

For the domain adaptation process, we leverage foundational work in source-free domain adaptation (SFDA) [28, 34] to adapt the model to the target domain $\mathcal{D}_r^{\mathcal{T}}$ using a loss term that enforces latent alignment between source and target data distributions. We denote this loss as $\mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}})$:

$$\min_{w^{\mathcal{T}}} \mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}}) \equiv \max_{w^{\mathcal{T}}} \left[\log p(w^{\mathcal{T}} \mid w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}})\right] \quad (6)$$

Combining $\mathcal{L}_{\text{MU}}$ and $\mathcal{L}_{\text{SFDA}}$, our unlearn-and-adapt objective φ is given by:

$$\min_{w^{\mathcal{T}}} \left[\log p(\hat{\mathcal{D}} \mid w^{\mathcal{T}}) - \log p(w^{\mathcal{T}} \mid w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}})\right] \equiv \min_{w^{\mathcal{T}}} \underbrace{\mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}}) + \alpha \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}_{\varphi(w^{\mathcal{T}}, \hat{x}, x^{\mathcal{T}}, \hat{y})} \quad (7)$$

where $\alpha \geq 0$ is a hyperparameter.

Overall Optimization Objective. Combining the optimization objectives in Expr. 3 and 7 gives our final joint

Algorithm 1 : Adversarial Optimization for SCADA-UL

Inputs: Source model $w^{\mathcal{S}}$, target data $\mathcal{D}^{\mathcal{T}}$, forget classes $\mathcal{C}_{\mathcal{F}}$, SFDA loss $\mathcal{L}_{\text{SFDA}}$, trade-off α, learning rates η_1, η_2
Init: $w^{\mathcal{T}} \leftarrow w^{\mathcal{S}}$
for each $c_{\mathcal{F}} \in \mathcal{C}_{\mathcal{F}}$ **do**
  $\hat{x}_{c_{\mathcal{F}}} \leftarrow \arg\min_{\hat{x}} \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}, c_{\mathcal{F}})$ ▷ Eq. (3)
end for
for epoch $= 1$ to M **do**
  for each $c_{\mathcal{F}} \in \mathcal{C}_{\mathcal{F}}$ **do**
    for step $= 1$ to $N/|\mathcal{C}_{\mathcal{F}}|$ **do**
      $x^{\mathcal{T}} \sim \mathcal{D}^{\mathcal{T}}$; $y \leftarrow f_{w^{\mathcal{T}}}(\hat{x}_{c_{\mathcal{F}}})$
      $\hat{y}_i \leftarrow \begin{cases} 0 & \text{if } i = c_{\mathcal{F}} \\ \frac{y_i}{\sum_{j \neq c_{\mathcal{F}}} y_j} & \text{otherwise} \end{cases}$
      $\varphi = \mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}}) + \alpha \cdot \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}_{c_{\mathcal{F}}}, \hat{y})$
      $w^{\mathcal{T}} \leftarrow w^{\mathcal{T}} - \eta_1 \nabla_{w^{\mathcal{T}}} \varphi$ ▷ Eqs. (7, 8)
      $\hat{x}_{c_{\mathcal{F}}} \leftarrow \hat{x}_{c_{\mathcal{F}}} - \eta_2 \nabla_{\hat{x}_{c_{\mathcal{F}}}} \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}_{c_{\mathcal{F}}}, c_{\mathcal{F}})$ ▷ Eqs. (3, 8)
    end for
  end for
end for
return $w_u^{\mathcal{T}} \leftarrow w^{\mathcal{T}}$

Table 2. **Results for Multi-Class SCADA-UL on OfficeHome, Office31, DomainNet datasets.** We compare against 10 baselines: Original = model adapted to target domain without applying unlearning; Retrain = model adapted with source model retrained without forget data; Finetune = model finetuned on subset of target retain data; MU methods [2, 10, 15, 51, 53]; PDA/SFPDA methods [5, 34]. $A_{\mathcal{D}_f^\mathcal{T}}$ = accuracy on target forget set; $A_{\mathcal{D}_r^\mathcal{T}}$ = target retain accuracy; Score captures overall unlearning performance i.e. high $A_{\mathcal{D}_r^\mathcal{T}}$ + low $A_{\mathcal{D}_f^\mathcal{T}}$. *(Best result in bold, second-best underlined)*

Method	OfficeHome			Office31			DomainNet-126		
	$A_{\mathcal{D}_r^\mathcal{T}}\uparrow$	$A_{\mathcal{D}_f^\mathcal{T}}\downarrow$	Score $\uparrow$	$A_{\mathcal{D}_r^\mathcal{T}}\uparrow$	$A_{\mathcal{D}_f^\mathcal{T}}\downarrow$	Score $\uparrow$	$A_{\mathcal{D}_r^\mathcal{T}}\uparrow$	$A_{\mathcal{D}_f^\mathcal{T}}\downarrow$	Score $\uparrow$
Original (SF(DA)^2 [28])	$75.8_{\pm1.0}$	$58.1_{\pm2.2}$	$0.48_{\pm0.0}$	$76.8_{\pm1.2}$	$90.1_{\pm1.6}$	$0.40_{\pm0.0}$	$67.7_{\pm0.6}$	$38.7_{\pm2.9}$	$0.50_{\pm0.0}$
Retrain	$76.3_{\pm1.2}$	$0.0_{\pm0.0}$	$0.76_{\pm0.0}$	$77.4_{\pm1.4}$	$0.0_{\pm0.0}$	$0.77_{\pm0.0}$	$66.3_{\pm1.3}$	$0.0_{\pm0.0}$	$0.66_{\pm0.0}$
Finetune	$\mathbf{76.1}_{\pm0.9}$	$49.2_{\pm2.2}$	$0.51_{\pm0.0}$	$\underline{76.7}_{\pm1.4}$	$82.8_{\pm2.0}$	$0.42_{\pm0.0}$	$\underline{66.5}_{\pm0.9}$	$20.4_{\pm4.1}$	$\underline{0.56}_{\pm0.0}$
UNSIR [51]	$35.0_{\pm5.6}$	$\mathbf{0.0}_{\pm0.0}$	$0.35_{\pm0.1}$	$59.7_{\pm5.0}$	$\underline{16.3}_{\pm3.3}$	$0.53_{\pm0.1}$	$14.6_{\pm3.6}$	$\underline{0.4}_{\pm0.6}$	$0.14_{\pm0.0}$
ZSMU [10]	$71.1_{\pm6.1}$	$39.8_{\pm10.}$	$0.50_{\pm0.0}$	$77.0_{\pm1.1}$	$84.2_{\pm3.3}$	$0.42_{\pm0.0}$	$65.0_{\pm1.6}$	$34.5_{\pm6.6}$	$0.49_{\pm0.0}$
Lipschitz [15]	$58.6_{\pm11.}$	$25.4_{\pm16.}$	$0.46_{\pm0.1}$	$65.3_{\pm9.5}$	$28.4_{\pm16.}$	$0.52_{\pm0.1}$	$39.0_{\pm11.}$	$8.5_{\pm8.6}$	$0.35_{\pm0.1}$
Nabla Tau [53]	$63.2_{\pm2.9}$	$\underline{1.5}_{\pm2.0}$	$\underline{0.62}_{\pm0.0}$	$73.4_{\pm1.7}$	$28.0_{\pm7.0}$	$\underline{0.59}_{\pm0.0}$	$44.4_{\pm5.4}$	$1.3_{\pm1.9}$	$0.44_{\pm0.1}$
Unlearned(+) [2]	$\mathbf{76.1}_{\pm9.4}$	$38.6_{\pm7.3}$	$0.54_{\pm0.0}$	$\mathbf{77.9}_{\pm0.5}$	$85.3_{\pm2.5}$	$0.42_{\pm0.0}$	$\mathbf{66.9}_{\pm0.8}$	$26.6_{\pm4.8}$	$0.54_{\pm0.0}$
PADA [5]	$74.0_{\pm1.4}$	$62.6_{\pm1.3}$	$0.45_{\pm0.0}$	$76.2_{\pm1.6}$	$89.1_{\pm1.6}$	$0.40_{\pm0.0}$	$61.9_{\pm0.4}$	$46.6_{\pm1.1}$	$0.43_{\pm0.0}$
SHOT [34]	$73.7_{\pm1.3}$	$24.7_{\pm1.2}$	$0.59_{\pm0.0}$	$76.0_{\pm1.2}$	$68.0_{\pm1.7}$	$0.46_{\pm0.0}$	$67.8_{\pm0.9}$	$39.6_{\pm2.7}$	$0.50_{\pm0.0}$
Ours	$\underline{75.1}_{\pm1.3}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.75}_{\pm0.0}$	$76.6_{\pm1.2}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.77}_{\pm0.0}$	$65.6_{\pm0.9}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.66}_{\pm0.0}$

optimization objective: $\min_{w^\mathcal{T}} \min_{\hat{x}} \varphi(w^\mathcal{T}, \hat{x}, x^\mathcal{T}, \hat{y}) + \mathcal{L}_{\text{ADV}}(w^\mathcal{T}, \hat{x})$. We solve this joint optimization problem using an iterative approach. To ensure stable training and separate the effects of $\mathcal{L}_{\text{ADV}}$ and φ, we alternate between solving the following two problems:

$$\min_{w^\mathcal{T}} \varphi(w^\mathcal{T}, \hat{x}, x^\mathcal{T}, \hat{y}) \quad \& \quad \min_{\hat{x}} \mathcal{L}_{\text{ADV}}(w^\mathcal{T}, \hat{x}) \tag{8}$$

A single SGD step is performed on each objective iteratively until convergence. Algorithm 1 and Figure 3 summarize our overall learning procedure.

Connection to Objective (1). When the adversarial loss $\mathcal{L}_{\text{ADV}}$ (3) is sufficiently minimized on $\hat{x}$ ($\hat{x} \in \hat{\mathcal{D}}$) using the source model $w^\mathcal{S}$, the adversarial samples effectively approximate the source forget dataset: $p(\hat{\mathcal{D}} \mid w^\mathcal{T}) \approx p(\mathcal{D}_f^\mathcal{S} \mid w^\mathcal{T})$. From Expr. 2 and the above approximation, our method is expressed as:

$$\begin{aligned} &\min_{w^\mathcal{T}} \max_{\hat{x}} \left[\log p(\mathcal{D}_f^\mathcal{S} \mid w^\mathcal{T}) - \log p(w^\mathcal{T} \mid w^\mathcal{S}, \mathcal{D}_r^\mathcal{T})\right] \equiv \\ &\quad \max_{w^\mathcal{T}} \left[\log p(w^\mathcal{T} \mid w^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) - \log p(\mathcal{D}_f^\mathcal{S} \mid w^\mathcal{T})\right] \end{aligned} \tag{9}$$

The final step is obtained by assuming no further optimization on $\hat{x}$ after initialization. Under these conditions, our method (as in Expr. 9) becomes identical to Expr. 1. It may be noted that our final implementation remains to be Expr. 8, and Exprs. 1, 9 convey its intuition.

Unlearning Multiple Classes. We extend our approach to unlearn multiple classes by dividing each training epoch into $|\mathcal{C}_\mathcal{F}|$ steps and unlearning each class $c_\mathcal{F}$ during its respective step. This process is detailed in Algorithm 1.

Our methodology assumes that datasets with disjoint label spaces are assumed to be conditionally independent given the model weights, similar to prior work (cf. Eq. 6 in [44]). Studying this further, we compare the terms with and without this assumption, namely, $\log p(\mathcal{D}_f^\mathcal{S} \mid w)$ and $\log p(\mathcal{D}_f^\mathcal{S} \mid w, \mathcal{D}_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$. The first term minimizes the posterior on the forget data, while the second conditions on retain data while minimizing the posterior on the forget data. The second term corresponds to a stronger form of unlearning: forgetting a class requires removing its influence from all remaining classes. For illustration, consider a land-use classification scenario. If the model was trained to classify "urban" and "industrial" scenes, and "industrial" must be forgotten for security reasons, the second setting would require removing all related information from "urban" as well. While more stringent, such a requirement is not typically motivated in privacy-driven unlearning, where the goal is to *selectively forget only the specified class*. Thus, we adopt the first, more relaxed expression, which focuses solely on removing information related to the forget classes without conditioning on the retain data. Additionally, following earlier work in online and continual learning scenarios [3, 46], we use $p(w \mid \mathcal{D}_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T}) \approx p(w \mid w_r^\mathcal{S}, \mathcal{D}_r^\mathcal{T})$, which comes from considering the posterior $p(w \mid \mathcal{D}^\mathcal{S})$ as approximately Gaussian and centered at the MLE estimate $w^\mathcal{S}$ and Bayesian inference. More analysis of these conditions is included in the Appendix.

4.2. Method Interpretation

To conceptually understand why our method works, we study the rescaled labels and adversarial optimization strategy w.r.t. the gradients of the MU loss $\mathcal{L}_{\text{MU}}(w^\mathcal{T}, \hat{x}, \hat{y})$ in Expr. 5. The following analysis studies how the proposed strategy differentiates between the *retain* and *forget* classes in the given setting, during the training (domain adaptation) process.

Theorem 1. *Let τ denote the set of final layer weights of a neural network, and let τ_c represent the weights corresponding to the output neuron of class c. Consider two disjoint sets of classes: a forget set $c_\mathcal{F}$ and a retain set $c_\mathcal{R}$. If the proposed unlearning process (Algorithm 1) is applied with forget set $c_\mathcal{F}$, then the gradient magnitude of the MU loss* *with respect to $\tau_{c_{\mathcal{R}}}$ satisfies the following inequality:*

$$\left\| \frac{\partial \mathcal{L}_{MU}}{\partial \tau_{c_{\mathcal{R}}}} \right\| \geq \left(\frac{1}{\delta} - 1 \right) \left\| \frac{\partial \mathcal{L}_{MU}}{\partial \tau_{c_{\mathcal{F}}}} \right\|$$

for some constant $\delta \in (0, 1)$.

Proof is included in the Appendix Sec. A.1.

Observation 1. The above result shows that our method (adversarial sample generation with label rescaling) provably induces more gradient flow on the weights $\tau_{c_{\mathcal{R}}}$ as compared to $\tau_{c_{\mathcal{F}}}$. Hence, in the domain adaptation process, our model focuses on the weights $\tau_{c_{\mathcal{R}}}$, rather than $\tau_{c_{\mathcal{F}}}$, thus obtaining strong performance on the retain classes, while gradually (over the training iterations) unlearning the forget classes (due to weaker gradients in every iteration). Our method induces the model to find a representation ϕ that aligns the minimization of all the proposed objectives: $\mathcal{L}_{\text{ADV}}$, $\mathcal{L}_{\text{MU}}$ and $\mathcal{L}_{\text{SFDA}}$. Our above result shows that $\mathcal{L}_{\text{MU}}$ explicitly focuses on learning the retain classes, while $\mathcal{L}_{\text{ADV}}$ focuses on generating forget samples on which $\mathcal{L}_{\text{MU}}$ is applied. $\mathcal{L}_{\text{SFDA}}$ is responsible for domain adaptation, like other DA methods, and further accentuates the learning with higher gradients for the classes retained in the target domain, thus coherently achieving our overall objective of simultaneous unlearning and domain adaptation in the SCADA-UL setting.

4.3. Extensions to Additional Settings

C-SCADA-UL involves unlearning forget classes in sequential requests ${\mathcal{C}_{\mathcal{F}}}^i$ (see Appendix Defn 3). Our method naturally extends to address this setting: for the initial unlearning request, the approach remains unchanged, and for subsequent requests, we use a small subset of the target train dataset and reduce the number of training epochs (as the model has already been adapted to the target domain).

UC-SCADA-UL poses an additional constraint where the forget classes $\mathcal{C}_{\mathcal{F}}$ are unknown but its cardinality is known (see Appendix Defn 2). To address this setting, we draw inspiration from PADA [5], which identifies and downweighs source-exclusive classes for Partial Domain Adaptation. Building on this idea, we adapt the use of the γ term to our setting by selecting the bottom-ranked classes by γ as the forget classes. The implementation details are provided in the Appendix.

5. Experiments and Results

Datasets. We perform experiments on three widely used domain adaptation datasets: *OfficeHome:* [54] This comprises four domains: Art, Clipart, Product, and Real World. Each domain consists of 65 categories of common objects such as spoons, bicycles, and backpacks. *Office31:* [47] This contains three domains: Amazon, DSLR, Webcam. Each domain comprises 31 object categories encountered in office settings. *DomainNet:* [45] This is a large-scale dataset consisting of six domains and 345 categories. Following other works [28], we experiment on seven tasks in DomainNet-126, using four domains: Clipart, Painting, Real and Sketch. Going beyond existing work, we also study two real-world datasets: *CheXpert* [29] $\rightarrow$ *NIH Chest X-ray* [57] (medical dataset with chest radiology images of patients labeled with the detected abnormality [27]), and a land-use dataset [65, 71] consisting of 7 categories of scenes such as grasslands, residential areas, etc.

Baselines. As we are the first to introduce a setting with simultaneous adaptation and unlearning tasks, we compare our work with some crude baselines: *Original* (only SFDA) and adapted variants of existing MU methods: *Retrain*, *Finetune*, *UNSIR* [51], *ZSMU* [10], *Lipschitz unlearning* [15], *Nabla Tau* [53], *Unlearned(+)* [2], and PDA/SFPDA methods: *PADA* [5], *SHOT* [34]. Since these methods have not been designed for domain adaptation and unlearning, we carefully apply them to our setting; the implementations are provided in Appendix Sec. A.7. We follow SF(DA)^2 [28] for our $\mathcal{L}_{\text{SFDA}}$ loss. We include experiments with other SFDA loss terms such as SHOT [34] in Appendix Sec. A.5.6.

Performance Metrics. *Target Forget Accuracy ($A_{\mathcal{D}_f^{\mathcal{T}}}$)*: accuracy of the model on the target forget dataset (zero-shot). *Target Retain Accuracy ($A_{\mathcal{D}_r^{\mathcal{T}}}$)*: accuracy on the target retain dataset. *Unlearn Score* [11]: defined as $A_{\mathcal{D}_r^{\mathcal{T}}}/(100 + A_{\mathcal{D}_f^{\mathcal{T}}})$ *Time consumption*: the total time taken to unlearn forget classes and adapt to the target domain. *MIA Accuracy (MIA%)* quantifies the effectiveness of a membership inference attack (MIA) model in identifying forget classes. Conventional MU methods typically follow the MIA training procedure proposed in [22]; we however adapt this approach for class-level unlearning by modifying the discrimination task to differentiate between entropies of retain data and unseen class data. An ideal unlearning method would ensure forget classes are misclassified as unseen class data.

Results. We perform experiments for four different settings: single class SCADA-UL, multi class SCADA-UL, UC-SCADA-UL and C-SCADA-UL. Each experiment is run three times and the mean and standard deviation of the metrics are reported in all studies. For all tables, the best score for each metric is in bold, excluding original and retrain. Due to space constraints, we present results for SCADA-UL on a medical dataset, multi-class SCADA-UL on benchmark datasets, MIA accuracy on DomainNet herein, and the other results in the Appendix.

From Tab. 2, we see that in the multi class SCADA-UL setting with $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$, the original model as well as PDA methods (PADA, SHOT) exhibit strong zero-shot capabilities on $\mathcal{D}_f^{\mathcal{T}}$; however, a high accuracy is undesirable in the context of unlearning. Retraining on the other hand attains zero accuracy on $\mathcal{D}_f^{\mathcal{T}}$ for all tasks, while maintaining high accuracy on $\mathcal{D}_r^{\mathcal{T}}$. (This method serves only as a gold standard for comparison, since this data is otherwise unavailable). Finetuning fails to perform well, yielding results

Table 3. Results for SCADA-UL on medical DA benchmark: CheXpert → NIH Chest X-ray

Method	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	Score ↑
Original (SF(DA)^2 [28])	36.9$_{\pm 1.9}$	16.0$_{\pm 8.8}$	0.32$_{\pm 0.0}$
Retrain	39.6$_{\pm 2.0}$	0.0$_{\pm 0.0}$	0.40$_{\pm 0.0}$
Finetune	35.7$_{\pm 1.4}$	9.6$_{\pm 9.8}$	0.33$_{\pm 0.0}$
UNSIR [51]	27.5$_{\pm 2.1}$	**0.0**$_{\pm 0.0}$	0.27$_{\pm 0.0}$
ZSMU [10]	26.9$_{\pm 3.5}$	**0.0**$_{\pm 0.0}$	0.27$_{\pm 0.0}$
Lipschitz [15]	37.6$_{\pm 0.2}$	**0.0**$_{\pm 0.0}$	**0.38**$_{\pm 0.0}$
Nabla Tau [53]	28.5$_{\pm 1.9}$	<u>4.4</u>$_{\pm 4.1}$	0.27$_{\pm 0.0}$
Unlearned(+) [2]	**42.0**$_{\pm 1.44}$	15.1$_{\pm 1.9}$	<u>0.36</u>$_{\pm 0.0}$
PADA [5]	36.1$_{\pm 0.4}$	**0.0**$_{\pm 0.0}$	<u>0.36</u>$_{\pm 0.0}$
SHOT [34]	31.1$_{\pm 1.5}$	20.2$_{\pm 1.4}$	0.26$_{\pm 0.0}$
Ours	<u>38.1</u>$_{\pm 0.7}$	**0.0**$_{\pm 0.0}$	**0.38**$_{\pm 0.0}$

Table 4. Membership Inference Attack Accuracy (MIA%) results for each task in the DomainNet Dataset

Method	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Avg.
Original (SF(DA)^2 [28])	56.0$_{\pm 2.1}$	74.4$_{\pm 2.8}$	71.3$_{\pm 1.6}$	52.5$_{\pm 3.2}$	74.7$_{\pm 1.6}$	60.0$_{\pm 1.6}$	62.5$_{\pm 7.3}$	64.5$_{\pm 2.9}$
Retrain	40.9$_{\pm 9.1}$	41.2$_{\pm 6.3}$	66.8$_{\pm 2.6}$	41.7$_{\pm 9.2}$	58.6$_{\pm 9.5}$	58.2$_{\pm 4.5}$	64.1$_{\pm 5.7}$	53.1$_{\pm 6.7}$
Finetune	54.2$_{\pm 8.1}$	<u>29.5</u>$_{\pm 4.0}$	49.0$_{\pm 2.5}$	<u>27.9</u>$_{\pm 3.1}$	72.0$_{\pm 8.7}$	61.4$_{\pm 3.0}$	67.9$_{\pm 7.4}$	51.7$_{\pm 5.3}$
UNSIR [51]	65.6$_{\pm 18.}$	53.4$_{\pm 21.}$	45.9$_{\pm 2.4}$	56.9$_{\pm 6.6}$	66.3$_{\pm 6.9}$	54.2$_{\pm 0.9}$	66.8$_{\pm 1.5}$	58.5$_{\pm 8.2}$
ZSMU [10]	58.1$_{\pm 11.}$	66.0$_{\pm 8.2}$	68.5$_{\pm 7.2}$	49.8$_{\pm 7.3}$	41.6$_{\pm 7.9}$	59.2$_{\pm 6.0}$	56.6$_{\pm 5.4}$	57.1$_{\pm 7.6}$
Lipschitz [15]	**24.7**$_{\pm 18.}$	60.4$_{\pm 18.}$	**40.1**$_{\pm 9.6}$	**20.7**$_{\pm 13.}$	**36.7**$_{\pm 23.}$	<u>39.4</u>$_{\pm 10.}$	**48.2**$_{\pm 8.9}$	**38.6**$_{\pm 14.}$
Nabla Tau [53]	55.1$_{\pm 6.7}$	37.6$_{\pm 8.4}$	45.5$_{\pm 6.7}$	49.3$_{\pm 9.4}$	69.8$_{\pm 2.4}$	**38.8**$_{\pm 4.1}$	65.1$_{\pm 6.1}$	51.6$_{\pm 6.3}$
Unlearned(+) [2]	55.9$_{\pm 2.4}$	48.8$_{\pm 1.9}$	62.7$_{\pm 7.2}$	37.6$_{\pm 4.8}$	58.8$_{\pm 24.}$	59.1$_{\pm 5.1}$	58.4$_{\pm 3.9}$	54.2$_{\pm 7.7}$
PADA [5]	66.9$_{\pm 0.8}$	81.0$_{\pm 1.1}$	74.6$_{\pm 0.7}$	75.1$_{\pm 0.6}$	66.8$_{\pm 1.5}$	64.8$_{\pm 1.5}$	75.7$_{\pm 2.5}$	72.1$_{\pm 1.3}$
SHOT [34]	58.2$_{\pm 7.9}$	76.6$_{\pm 0.1}$	73.9$_{\pm 2.8}$	54.4$_{\pm 5.5}$	75.2$_{\pm 5.2}$	59.5$_{\pm 5.3}$	61.0$_{\pm 4.6}$	65.5$_{\pm 4.5}$
Ours	<u>52.7</u>$_{\pm 1.4}$	**22.0**$_{\pm 2.6}$	<u>40.3</u>$_{\pm 3.1}$	28.8$_{\pm 0.5}$	<u>44.3</u>$_{\pm 6.8}$	49.0$_{\pm 0.9}$	<u>58.0</u>$_{\pm 8.1}$	<u>42.2</u>$_{\pm 3.3}$

similar to the original model due to lack of unlearning on the forget class. Existing MU methods when adapted to our setting (as described in Appendix A.7.2) perform poorly too, either resulting in a significant drop in retain accuracy (UNSIR, Nabla Tau) or still maintaining a high forget accuracy (ZSMU, Lipschitz). We hypothesize methods like UNSIR, Nabla Tau and Unlearned (+) perform poorly in this setting since they were not designed to handle shift in the data distribution. In contrast, our method demonstrates strong performance, offering results comparable to retraining.

Table 3 shows results on a real-world medical domain adaptation benchmark for chest disease classification: CheXpert → NIH Chest X-ray. In medical applications like these, regulatory bodies may mandate that source-exclusive medical conditions are not transferred to the target domain. Additionally, our present Membership Inference Attack Accuracy (MIA%) in Tab. 4 shows that our method tends to either have the lowest or second lowest accuracy out of all baselines. Interestingly, we find that methods such as UNSIR, which achieve low forget accuracy, still have a high MIA%, while methods with higher forget accuracy such as Lipschitz have a low MIA%, corroborating the necessity of multiple metrics to evaluate such settings.

Full result tables, as well as results for UC-SCADA-UL, C-SCADA-UL, time consumption, another real-world dataset (scene classification), and other SFDA loss functions are in the Appendix Sec. A.5.

6. Analysis and Ablation Studies

We conduct comprehensive studies to evaluate key components of our method. All experiments measure multi-class SCADA-UL performance across 3 trials on the OfficeHome dataset, reporting retain accuracy (Acc $\mathcal{D}_r^{\mathcal{T}}$) and forget accuracy (Acc $\mathcal{D}_f^{\mathcal{T}}$) averaged across 12 tasks. Due to space constraints, we provide two ablations/studies here; the remaining studies (including on the α term and number of adversarial samples) are in Appendix.

Stage at which algorithm is applied. Tab. 5 shows our algorithm applied before, during, and after the domain adaptation process. In the after-adaptation scenario, we see a significant drop in accuracy due to *catastrophic forgetting*. While $\mathcal{L}_{\text{SFDA}}$ helps reverse this in the before-adaptation scenario, it has high forget accuracy which is undesirable. The proposed during -adaptation approach obtains good forget and retain performance via simultaneous optimization of unlearning and adaptation objectives.

Table 5. Study on stage of applying our algorithm

Stage	Acc $\mathcal{D}_r^{\mathcal{T}}$ ↑	Acc $\mathcal{D}_f^{\mathcal{T}}$ ↓
Before Adaptation	75.5$_{\pm 1.6}$	52.0$_{\pm 5.3}$
During Adaptation	75.1$_{\pm 1.3}$	0.0$_{\pm 0.0}$
After Adaptation	58.2$_{\pm 7.1}$	0.0$_{\pm 0.0}$

Labeling Strategy. Our proposed labeling strategy outperforms uniform and random labeling, as seen by the higher Acc $\mathcal{D}_r^{\mathcal{T}}$ and lower Acc $\mathcal{D}_f^{\mathcal{T}}$ in Tab. 6. Uniform labeling achieves moderate unlearning but underperforms our method, while random labeling fails in our setting, leading to catastrophic forgetting. This behavior is likely due to $\mathcal{L}_{\text{MU}}$ dominating over $\mathcal{L}_{\text{SFDA}}$, disrupting training stability. Furthermore, our labels align the post-unlearning model's output distributions more closely to that of the retrained model (See Appendix). This behavior likely explains the effectiveness of our labeling strategy – it provides a more natural objective for $\mathcal{L}_{\text{MU}}$, that minimizes conflict with retain classes.

Table 6. Choice of relabeling strategy

Strategy	Acc $\mathcal{D}_r^{\mathcal{T}}$ ↑	Acc $\mathcal{D}_f^{\mathcal{T}}$ ↓
Rescaled	75.1$_{\pm 1.3}$	0.0$_{\pm 0.0}$
Uniform	70.5$_{\pm 3.6}$	3.5$_{\pm 4.1}$
Random	1.5$_{\pm 0.4}$	0.0$_{\pm 0.0}$

7. Conclusions

In this work, we present a novel machine unlearning setting called **Un**learning **S**ource-exclusive **Cl**A**s**ses in **D**omain **A**daptation (SCADA-UL), along with two variants UC-SCADA-UL and C-SCADA-UL. This setting addresses the task of adapting a source model to a target domain while unlearning the specific source-exclusive classes, and has increased relevance as learning models get commonly adapted across domains today. To address these settings, we propose a new unlearning algorithm based on optimizing the model to iteratively forget the best estimate of forget classes throughout the domain adaptation process. Our thorough empirical and theoretical analysis highlights the effectiveness of our approach in the settings.

Acknowledgements

Arnav is supported by ACM India iKDD (Special Interest Group On Knowledge Discovery and Data Mining) and the University of Michigan Rackham Conference Travel Grant. Poornima is supported by the Google Research Grant and IIT-Hyderabad Conference Travel Support Grant. We are grateful for all the above institutions for their support to carry out this work. We are also grateful to our anonymous reviewers, program chairs and area chairs for their valuable feedback in improving the paper quality.

References

[1] Peshal Agarwal, Danda Pani Paudel, Jan-Nico Zaech, and Luc Van Gool. Unsupervised robust domain adaptation without source data. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2009–2018, 2022. 3

[2] Sk Miraj Ahmed, Umit Yigit Basaran, Dripta S. Raychaudhuri, Arindam Dutta, Rohit Kundu, Fahim Faisal Niloy, Basak Guler, and Amit K. Roy-Chowdhury. Towards source-free machine unlearning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4948–4957, 2025. 2, 3, 6, 7, 8, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36

[3] Abhishek Aich. Elastic weight consolidation (ewc): Nuts and bolts. *arXiv preprint arXiv:2105.04093*, 2021. 6, 19

[4] Thomas Baumhauer, Pascal Schöttle, and Matthias Zeppelzauer. Machine unlearning: Linear filtration for logit-based classifiers. *Machine Learning*, 111(9):3203–3226, 2022. 1

[5] Zhangjie Cao, Lijia Ma, Mingsheng Long, and Jianmin Wang. Partial adversarial domain adaptation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 135–150, 2018. 2, 3, 6, 7, 8, 13, 20, 21, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36

[6] CCPA. California consumer privacy act, 2018. 1

[7] Sungmin Cha, Sungjun Cho, Dasol Hwang, Honglak Lee, Taesup Moon, and Moontae Lee. Learning to unlearn: Instance-wise unlearning for pre-trained classifiers. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 11186–11194, 2024. 3

[8] Min Chen, Zhikun Zhang, Tianhao Wang, Michael Backes, Mathias Humbert, and Yang Zhang. Graph unlearning. In *Proceedings of the 2022 ACM SIGSAC conference on computer and communications security*, pages 499–513, 2022. 3

[9] Vikram S Chundawat, Ayush K Tarun, Murari Mandal, and Mohan Kankanhalli. Can bad teaching induce forgetting? unlearning in deep networks using an incompetent teacher. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7210–7217, 2023. 3

[10] Vikram S Chundawat, Ayush K Tarun, Murari Mandal, and Mohan Kankanhalli. Zero-shot machine unlearning. *IEEE Transactions on Information Forensics and Security*, 18:2345–2354, 2023. 1, 2, 3, 5, 6, 7, 8, 15, 16, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[11] Arnav Devalapally and Gowtham Valluri. A simple machine unlearning approach using elastic weight consolidation. In *International Conference on Recent Trends in AI Enabled Technologies*, pages 1–10. Springer, 2023. 7

[12] Cian Eastwood, Ian Mason, Chris Williams, and Bernhard Schölkopf. Source-free adaptation to measurement shift via bottom-up feature restoration. In *International Conference on Learning Representations*, 2022. 3

[13] Ali Ebrahimpour-Boroojeny, Hari Sundaram, and Varun Chandrasekaran. Not all wrong is bad: Using adversarial examples for unlearning. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, pages 14950–14971. Proceedings of Machine Learning Research (PMLR), 2025. 3

[14] Yuqi Fang, Pew-Thian Yap, Weili Lin, Hongtu Zhu, and Mingxia Liu. Source-free unsupervised domain adaptation: A survey. *Neural Networks*, 174:106230, 2024. 3

[15] Jack Foster, Kyle Fogarty, Stefan Schoepf, Cengiz Öztireli, and Alexandra Brintrup. Zero-shot machine unlearning at scale via lipschitz regularization. *arXiv preprint arXiv:2402.01401*, 2024. 1, 2, 6, 7, 8, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[16] Jack Foster, Stefan Schoepf, and Alexandra Brintrup. Fast machine unlearning without retraining through selective synaptic dampening. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 12043–12051, 2024. 3

[17] Matt Fredrikson, Somesh Jha, and Thomas Ristenpart. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, page 1322–1333, New York, NY, USA, 2015. Association for Computing Machinery. 2

[18] GDPR. General data protection regulation, 2016. 1

[19] Jiahui Geng, Qing Li, Herbert Woisetschlaeger, Zongxiong Chen, Fengyu Cai, Yuxia Wang, Preslav Nakov, Hans-Arno Jacobsen, and Fakhri Karray. A comprehensive survey of machine unlearning techniques for large language models, 2025. 3

[20] Ali Ghazal and Radwa El Shawi. Zero-shot machine unlearning using generative adversarial network. In *Advances in Knowledge Discovery and Data Mining: 29th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2025, Sydney, NSW, Australia, June 10–13, 2025, Proceedings, Part IV*, page 65–77, Berlin, Heidelberg, 2025. Springer-Verlag. 3

[21] Aditya Golatkar, Alessandro Achille, and Stefano Soatto. Eternal sunshine of the spotless net: Selective forgetting in deep networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9304–9312, 2020. 1, 2, 3

[22] Aditya Golatkar, Alessandro Achille, and Stefano Soatto. Forgetting outside the box: Scrubbing deep networks of information accessible from input-output observations. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIX 16*, pages 383–398. Springer, 2020. 7, 23

[23] Xiaojin Gu, Xing Yu, Yan Yang, Jian Sun, and Zongben Xu. Adversarial reweighting for partial domain adaptation. In

Advances in Neural Information Processing Systems, pages 14860–14872, 2021. 3

[24] Chuan Guo, Tom Goldstein, Awni Y. Hannun, and Laurens van der Maaten. Certified data removal from machine learning models. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, pages 3832–3842. PMLR, 2020. 3

[25] Peng Guo, Jun Zhu, and Yu Zhang. Selective partial domain adaptation. In *33rd British Machine Vision Conference 2022, BMVC 2022*, London, UK, 2022. 3

[26] Niv Haim, Gal Vardi, Gilad Yehudai, Ohad Shamir, and Michal Irani. Reconstructing training data from trained neural networks. *Advances in Neural Information Processing Systems*, 35:22911–22924, 2022. 1

[27] Bishi He, Yuanjiao Chen, Darong Zhu, and Zhe Xu. Domain adaptation via wasserstein distance and discrepancy metric for chest x-ray image classification. *Scientific Reports*, 14(1): 2690, 2024. 7, 24

[28] Uiwon Hwang, Jonghyun Lee, Juhyeon Shin, and Sungroh Yoon. SF(DA)^2: Source-free domain adaptation through the lens of data augmentation. In *The Twelfth International Conference on Learning Representations*, 2024. 2, 3, 5, 6, 7, 8, 16, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36

[29] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence*, pages 590–597, 2019. 7

[30] Masato Ishii and Masashi Sugiyama. Source-free domain adaptation via distributional alignment by matching batch normalization statistics, 2021. 3

[31] Marvin Klingner, Jan-Aike Termohlen, Jacob Ritterbach, and Tim Fingscheidt. Unsupervised BatchNorm Adaptation (UBNA): A Domain Adaptation Method for Semantic Segmentation Without Using Source Domain Representations . In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, pages 210–220, Los Alamitos, CA, USA, 2022. IEEE Computer Society. 3

[32] Jonghyun Lee, Dahuin Jung, Junho Yim, and Sung-Hoon Yoon. Confidence score for source-free unsupervised domain adaptation. In *International Conference on Machine Learning*, 2022. 3

[33] Na Li, Chunyi Zhou, Yansong Gao, Hui Chen, Zhi Zhang, Boyu Kuang, and Anmin Fu. Machine unlearning: Taxonomy, metrics, applications, challenges, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 2025. 3

[34] Jian Liang, Dapeng Hu, and Jiashi Feng. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In *International conference on machine learning*, pages 6028–6039. PMLR, 2020. 2, 3, 5, 6, 7, 8, 20, 21, 22, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[35] Bo Liu, Qiang Liu, and Peter Stone. Continual learning and private unlearning. In *Conference on Lifelong Learning Agents*, pages 243–254. PMLR, 2022. 3

[36] Hengzhu Liu, Ping Xiong, Tianqing Zhu, and Philip S. Yu. A survey on machine unlearning: Techniques and new emerged privacy risks. *J. Inf. Secur. Appl.*, 90(C), 2025. 2

[37] Jiawei Liu, Chenwang Wu, Defu Lian, and Enhong Chen. Efficient machine unlearning via influence approximation. *arXiv preprint arXiv:2507.23257*, 2025. 3

[38] Yuchen Liu, Yabo Chen, Wenrui Dai, Mengran Gou, Chun-Ting Huang, and Hongkai Xiong. Source-free domain adaptation with contrastive domain alignment and self-supervised exploration for face anti-spoofing. In *European Conference on Computer Vision*, pages 511–528. Springer, 2022. 3

[39] Yi Liu, Lei Xu, Xingliang Yuan, Cong Wang, and Bo Li. The right to be forgotten in federated learning: An efficient realization with rapid retraining. In *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, pages 1749–1758. IEEE, 2022. 3

[40] Zheyuan Liu, Guangyao Dou, Zhaoxuan Tan, Yijun Tian, and Meng Jiang. Machine unlearning in generative ai: A survey. *CoRR*, 2024. 3

[41] Shao-Yuan Lo, Poojan Oza, Sumanth Chennupati, Alejandro Galindo, and Vishal M Patel. Spatio-temporal pixel-level contrastive learning-based source-free domain adaptation for video semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10534–10543, 2023. 3

[42] Seth Neel, Aaron Roth, and Saeed Sharifi-Malvajerdi. Descent-to-delete: Gradient-based methods for machine unlearning. In *Algorithmic Learning Theory, 16-19 March 2021, Virtual Conference, Worldwide*, pages 931–962. PMLR, 2021. 3

[43] Thanh Tam Nguyen, Thanh Trung Huynh, Zhao Ren, Phi Le Nguyen, Alan Wee-Chung Liew, Hongzhi Yin, and Quoc Viet Hung Nguyen. A survey of machine unlearning. *ACM Trans. Intell. Syst. Technol.*, 16(5), 2025. 3

[44] Subhodip Panda, Shashwat Sourav, et al. Partially blinded unlearning: Class unlearning for deep networks from bayesian perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 6372–6380, 2025. 1, 4, 6

[45] Xingchao Peng, Qinxun Bai, Xide Xia, Zijun Huang, Kate Saenko, and Bo Wang. Moment matching for multi-source domain adaptation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1406–1415, 2019. 7

[46] Hippolyt Ritter, Aleksandar Botev, and David Barber. Online structured laplace approximations for overcoming catastrophic forgetting. *Advances in Neural Information Processing Systems*, 31, 2018. 4, 6, 19

[47] Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11*, pages 213–226. Springer, 2010. 7

[48] Kaichao Sheng, Ke Li, Xiaobo Zheng, Jian Liang, Weijie Dong, Fei Huang, Rongrong Ji, and Xiaoshuai Sun. On evolving attention towards domain adaptation, 2021. arXiv preprint arXiv:2106.04581. 3

[49] Shaoyue Song, Hongkai Yu, Zhenjiang Miao, Qiang Zhang, Yuewei Lin, and Song Wang. Domain adaptation for convolutional neural networks-based remote sensing scene classification. *IEEE Geoscience and Remote Sensing Letters*, 16(8): 1324–1328, 2019. 20, 23

[50] Ryutaro Tanno, Melanie F Pradier, Aditya Nori, and Yingzhen Li. Repairing neural networks by leaving the right past behind. *Advances in Neural Information Processing Systems*, 35:13132–13145, 2022. 3

[51] Ayush K Tarun, Vikram S Chundawat, Murari Mandal, and Mohan Kankanhalli. Fast yet effective machine unlearning. *IEEE Transactions on Neural Networks and Learning Systems*, 2023. 1, 2, 3, 6, 7, 8, 15, 16, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[52] Qing Tian and Lulu Kang. Source-free unsupervised domain adaptation through trust-guided partitioning and worst-case aligning. *Know.-Based Syst.*, 318(C), 2025. 3

[53] Daniel Trippa, Cesare Campagnano, Maria Sofia Bucarelli, Gabriele Tolomei, and Fabrizio Silvestri. $\nabla\tau$: Gradient-based and task-agnostic machine unlearning. *CoRR*, 2024. 6, 7, 8, 15, 16, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[54] Hemanth Venkateswara, Jose Eusebio, Shayok Chakraborty, and Sethuraman Panchanathan. Deep hashing network for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5018–5027, 2017. 7

[55] Rui Wang, Zuxuan Wu, Zejia Weng, Jingjing Chen, Guo-Jun Qi, and Yu-Gang Jiang. Cross-domain contrastive learning for unsupervised domain adaptation. *IEEE Transactions on Multimedia*, PP:1–1, 2022. 3

[56] Weiqi Wang, Chenhan Zhang, Zhiyi Tian, and Shui Yu. Machine unlearning via representation forgetting with parameter self-sharing. *IEEE Transactions on Information Forensics and Security*, 19:1099–1111, 2024. 3

[57] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017. 7

[58] Alexander Warnecke, Lukas Pirch, Christian Wressnegger, and Konrad Rieck. Machine unlearning of features and labels. In *Network and Distributed System Security Symposium (NDSS) 2023*, 2023. 3

[59] Yinjun Wu, Edgar Dobriban, and Susan Davidson. DeltaGrad: Rapid retraining of machine learning models. In *Proceedings of the 37th International Conference on Machine Learning*, pages 10355–10366. PMLR, 2020. 3

[60] Gezheng Xu, Hui Guo, Li Yi, Charles Ling, Boyu Wang, and Grace Yi. Revisiting source-free domain adaptation: a new perspective via uncertainty control. In *The Thirteenth International Conference on Learning Representations*, 2025. 3, 20, 21

[61] Heng Xu, Tianqing Zhu, Lefeng Zhang, Wanlei Zhou, and P. Yu. Machine unlearning: A survey. *ACM Computing Surveys*, 56:1 – 36, 2023. 3

[62] Tomoya Yamashita, Masanori Yamada, and Takashi Shibata. One-shot machine unlearning with mnemonic code. In *The 16th Asian Conference on Machine Learning (Conference Track)*, 2024. 1

[63] Haonan Yan, Xiaoguang Li, Ziyao Guo, Hui Li, Fenghua Li, and Xiaodong Lin. Arcane: An efficient architecture for exact machine unlearning. In *IJCAI*, page 19, 2022. 3

[64] Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, and Joost van de Weijer. Onering: A simple method for source-free open-partial domain adaptation. *arXiv preprint arXiv:2206.03600*, 2022. 3

[65] Yi Yang and Shawn Newsam. Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems*, pages 270–279, 2010. 7, 20

[66] Jin Yao, Eli Chien, Minxin Du, Xinyao Niu, Tianhao Wang, Zezhou Cheng, and Xiang Yue. Machine unlearning of pre-trained large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8403–8419, Bangkok, Thailand, 2024. Association for Computational Linguistics. 3

[67] Kai Yue, Wenhao Li, Jing Li, Zhen Chen, and Rui Huang. Source-free partial domain adaptation in rotating machinery fault diagnosis using deep hypothesis domain adaptation network. In *2022 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, pages 1–6. IEEE, 2022. 3

[68] Xulong Zhang, Jianzong Wang, Ning Cheng, Yifu Sun, Chuanyao Zhang, and Jing Xiao. Machine unlearning methodology based on stochastic teacher network. In *19th International Conference on Advanced Data Mining and Applications. (ADMA 2023)*, pages 250–261, 2023. 3

[69] Zhengyi Zhong, Weidong Bao, Ji Wang, Shuai Zhang, Jingxuan Zhou, Lingjuan Lyu, and Wei Yang Bryan Lim. Unlearning through knowledge overwriting: Reversible federated unlearning via selective sparse adapter, 2025. 3

[70] Ligeng Zhu, Zhijian Liu, and Song Han. Deep leakage from gradients. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019. 2

[71] Qin Zou, Lihao Ni, Tong Zhang, and Qian Wang. Deep learning based feature selection for remote sensing scene classification. *IEEE Geoscience and remote sensing letters*, 12(11):2321–2325, 2015. 7, 20

Appendix

Contents

A.1. Proof of Theorem 4.1

Theorem 1. *Let τ denote the set of final layer weights of a neural network, and let τ_c represent the weights corresponding to the output neuron of class c. Consider two disjoint sets of classes: a forget set $c_{\mathcal{F}}$ and a retain set $c_{\mathcal{R}}$. If the proposed unlearning process (Algorithm 1) is applied with forget set $c_{\mathcal{F}}$, then the gradient magnitude of the MU loss with respect to $\tau_{c_{\mathcal{R}}}$ satisfies the following inequality:*

$$\left\|\frac{\partial \mathcal{L}_{MU}}{\partial \tau_{c_{\mathcal{R}}}}\right\| \geq \left(\frac{1}{\delta} - 1\right) \left\|\frac{\partial \mathcal{L}_{MU}}{\partial \tau_{c_{\mathcal{F}}}}\right\|$$

for some constant $\delta \in (0, 1)$.

Proof. Let the pre-final layer output of an adversarial sample $\hat{x}$ generated by the model being adapted, $w^{\mathcal{T}}$, be denoted as $\phi(\hat{x}) = \hat{\phi}$.
Then, the MU loss in Expr. (5) is given by

$$\mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y}) = \sum_{i \in c_{\mathcal{R}}} -\hat{y}_i \log\left(\frac{\exp(\hat{\phi}^T \tau_i)}{\sum_{j \in C} \exp(\hat{\phi}^T \tau_j)}\right) \tag{A.10}$$

$$= \sum_{i \in c_R} -\hat{y}_i \hat{\phi}^T \tau_i + \log\left(\sum_{j \in C} \exp(\hat{\phi}^T \tau_j)\right) \tag{A.11}$$

The second equality holds due to our rescaled labeling strategy; $\sum_{i \in c_R} \hat{y}_i = 1$ (Expr. 4).

The gradients for any class c are given by

$$\frac{\partial \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}{\partial \tau_c} = -\hat{y}_c \hat{\phi} \mathbf{1}_{c \notin c_{\mathcal{F}}} + \underbrace{\frac{\exp(\hat{\phi}^T \tau_c)}{\sum_{j \in C} \exp(\hat{\phi}^T \tau_j)}}_{f_{w^{\mathcal{T}}}(\hat{x}) = y_c} \hat{\phi} \tag{A.12}$$

$$= \hat{\phi}(y_c - \mathbf{1}_{c \notin c_{\mathcal{F}}} \hat{y}_c) \tag{A.13}$$

Specifically, the norms of the gradients w.r.t. the weights in the case where $c \in c_{\mathcal{F}}$ or $c \in c_{\mathcal{R}}$, are given by

$$\frac{\partial \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}{\partial \tau_{c_{\mathcal{R}}}} = \|\hat{\phi}(y_c - \frac{y_c}{\sum_{j \in c_R} y_j})\| \tag{A.14}$$

$$\frac{\partial \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}{\partial \tau_{c_{\mathcal{F}}}} = \|\hat{\phi} y_c\| \tag{A.15}$$

We crucially observe that because of the nature of our adversarial sample generation in Expr. (3),

$$\sum_{j \in c_R} y_j \leq \delta$$

where $\delta \in (0, 1)$, because $P(\hat{x} \in c_R) = \sum_{j \in c_{\mathcal{R}}} y_j$ represents the softmax probability of the sample $\hat{x}$ belonging to any retain class, which is precisely what is minimized in $\mathcal{L}_{\text{ADV}}$. This leads to our main gradient flow inequality

$$\left\| \frac{\partial \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}{\partial \tau_{c_{\mathcal{R}}}} \right\| \geq \left(\frac{1}{\delta} - 1 \right) \left\| \frac{\partial \mathcal{L}_{\text{MU}}(w^{\mathcal{T}}, \hat{x}, \hat{y})}{\partial \tau_{c_{\mathcal{F}}}} \right\| \tag{A.16}$$

This inequality indicates that the gradient of the machine unlearning loss is more significant on the weights connected to the final-layer neurons of the classes that need to be retained, given that the adversarial sample is initially classified as part of a forget class $c_{\mathcal{F}}$ with a probability at least $1 - \delta$. □

A.2. Additional Algorithms

We provide an extended version of Algorithm 1 with full training details in Algorithm 2. Whenever this algorithm is invoked by another algorithm, we refer to it as *AO_SCADA_UL()* (**SCADA-UL** solved via **A**dversarial **O**ptimization) for convenience.
UC-SCADA-UL. Algorithm 3 describes the procedure for our algorithm adapted for the UC-SCADA-UL setting, where the identity of the forget classes is not known (Definition 2). It involves an additional forget class prediction step that estimates the classes that are most likely the forget classes based on the target dataset $\mathcal{D}^{\mathcal{T}}$. This estimation step utilizes a term $\gamma \in \mathbb{R}^d$ where d is the number of classes (similar to [5]). This term provides a relative measure of class relevance for a dataset. The bottom $R \cdot |\mathcal{C}_{\mathcal{F}}|$ classes associated with γ are selected as the predicted forget classes and the original algorithm (Algorithm 1) is applied using these classes.
C-SCADA-UL. Algorithm 4 shows the adaptation of our algorithm for the C-SCADA-UL setting (Definition 3), where the unlearning requests can be received over multiple time steps. For the first set of classes ${\mathcal{C}_{\mathcal{F}}}^1$, the process remains identical to the original algorithm (Algorithm 1). For subsequent classes, we use a subset of the target data $\mathcal{D}^{\mathcal{T}}_{\text{sub}}$ and apply the original algorithm to the previously adapted model $w^{\mathcal{T}}_u$.

A.3. Definitions and Motivation for UC-SCADA-UL and C-SCADA-UL

A.3.1. Formal Definitions

Definition 2. (UC-SCADA-UL). Unknown Class SCADA-UL is the process of learning a function $\mathcal{U} : \{w^{\mathcal{S}}, \mathcal{D}^{\mathcal{T}}_r, |\mathcal{C}_{\mathcal{F}}|\} \to w^{\mathcal{T}}_u$ that produces an unlearned, adapted model which behaves as though it were adapted from a source model not trained on $\mathcal{D}^{\mathcal{S}}_f$, all without requiring knowledge of $\mathcal{C}_{\mathcal{F}}$.

Definition 3. (C-SCADA-UL). Continual SCADA-UL is the process of learning a sequence of unlearning functions $\mathcal{U}^i(w^{\mathcal{T},i-1}, \mathcal{D}^{\mathcal{T}}_r, {\mathcal{C}_{\mathcal{F}}}^1 \cup \cdots \cup {\mathcal{C}_{\mathcal{F}}}^i)$ where $i > 0, w^{\mathcal{T},0} = w^{\mathcal{S}}$. Each function produces an unlearned, adapted model $w^{\mathcal{T},i}_u$ that behaves as though it were adapted from a source model that was not trained on the data associated with ${\mathcal{C}_{\mathcal{F}}}^1 \cup \cdots \cup {\mathcal{C}_{\mathcal{F}}}^i$.

We note that the term "continual" here specifically refers to sequential unlearning requests and not traditional continual learning where new classes are learned over tasks.

Algorithm 2 Adversarial Optimization for SCADA-UL (AO_SCADA_UL: Detailed Algorithm)

Inputs:
  Source model $w^{\mathcal{S}}$, target data $\mathcal{D}^{\mathcal{T}}$
Require:
  Forget classes $\mathcal{C}_{\mathcal{F}}$, SFDA loss $\mathcal{L}_{\text{SFDA}}$, loss trade-off α
  Learning rates η_1 (model), η_2 (adv. samples), η_{init} (init. step)
  Epochs M, total steps N, Initialization T_{init}
Init:
  $w^{\mathcal{T}} \leftarrow w^{\mathcal{S}}$
  Initialize optimizer and LR scheduler for $w^{\mathcal{T}}$

for each class $c_{\mathcal{F}} \in \mathcal{C}_{\mathcal{F}}$ **do**
  Initialize adversarial sample $\hat{x}_{c_{\mathcal{F}}}$ (random)
  for $t = 1$ to T_{init} **do**
    $\hat{x}_{c_{\mathcal{F}}} \leftarrow \hat{x}_{c_{\mathcal{F}}} - \eta_{\text{init}} \nabla_{\hat{x}_{c_{\mathcal{F}}}} \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}_{c_{\mathcal{F}}}, c_{\mathcal{F}})$ ▷ Adversarial sample "seed"
  end for
end for

for epoch $= 1$ to M **do**
  Iterate through target data
  for each forget class $c_{\mathcal{F}} \in \mathcal{C}_{\mathcal{F}}$ **do**
    for step $= 1$ to $N/|\mathcal{C}_{\mathcal{F}}|$ **do**
      Sample target image $x^{\mathcal{T}} \sim \mathcal{D}^{\mathcal{T}}$
      Compute SFDA loss $\mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}})$
      Obtain logits on adversarial sample:

$$y = f_{w^{\mathcal{T}}}(\hat{x}_{c_{\mathcal{F}}})$$

      Set $\hat{y}_{c_{\mathcal{F}}} = 0$; renormalize remaining logits:

$$\hat{y}_i = \frac{y_i}{\sum_{j \neq c_{\mathcal{F}}} y_j}$$

      Set $\varphi = \mathcal{L}_{\text{SFDA}}(w^{\mathcal{T}}, x^{\mathcal{T}}) + \alpha\, \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}_{c_{\mathcal{F}}}, \hat{y})$
      Update model parameters:

$$w^{\mathcal{T}} \leftarrow w^{\mathcal{T}} - \eta_1 \nabla_{w^{\mathcal{T}}} \varphi$$

      Apply LR scheduler
      Update adversarial sample:

$$\hat{x}_{c_{\mathcal{F}}} \leftarrow \hat{x}_{c_{\mathcal{F}}} - \eta_2 \nabla_{\hat{x}_{c_{\mathcal{F}}}} \mathcal{L}_{\text{CE}}(w^{\mathcal{T}}, \hat{x}_{c_{\mathcal{F}}}, c_{\mathcal{F}})$$

    end for
  end for
end for
return Final unlearned target model $w_u^{\mathcal{T}} = w^{\mathcal{T}}$

A.3.2. Real-World Motivations for Variants

In Sec. 1, we discuss real-world use cases that motivate this work. An increased use of finetuning and adapting models underlines the significance of the setting proposed in this work. To discuss this further, the SCADA unlearning setting (Definition 1) assumes prior knowledge of source-exclusive classes alongside the source model and unlabeled target dataset. However, this assumption may not hold in practice sometimes, and it may be difficult to infer the source-exclusive classes from the source model $w^{\mathcal{S}}$ and the given unlabeled target retain data $\mathcal{D}_r^{\mathcal{T}} = \{(x_{i,r}^{\mathcal{T}})\}_{i=1}^{n}$. (If the target dataset were labeled, identifying these classes would be straightforward as one could iterate through the dataset to compile the set of observed target

Algorithm 3 Adversarial Optimization for UC-SCADA Unlearning

Inputs: Source Model $w^{\mathcal{S}}$, target dataset $\mathcal{D}^{\mathcal{T}}$, number of forget classes $|\mathcal{C}_{\mathcal{F}}|$, SFDA loss $\mathcal{L}_{\text{SFDA}}$, trade-off α, learning rates η_1, η_2

Init: $w^{\mathcal{T}} = w^{\mathcal{S}}$; $\gamma = 0$

for each $x^{\mathcal{T}} \in \mathcal{D}^{\mathcal{T}}$ **do**
  $\gamma \leftarrow \gamma + f_{w^{\mathcal{T}}}(x^{\mathcal{T}})$
end for

$\mathcal{C}^*_{\mathcal{F}} \leftarrow$ Top_K$(-\gamma, R \cdot |\mathcal{C}_{\mathcal{F}}|)$ ▷ Lowest $R \cdot |\mathcal{C}_{\mathcal{F}}|$ classes according to γ
$w^{\mathcal{T}}_u \leftarrow$ AO_SCADA_UL$(w^{\mathcal{T}}, \mathcal{D}^{\mathcal{T}}, \mathcal{C}^*_{\mathcal{F}}, \mathcal{L}_{\text{SFDA}}, \alpha, \eta_1, \eta_2)$ ▷ Alg (1)

Return: $w^{\mathcal{T}}_u$

Algorithm 4 Adversarial Optimization for C-SCADA Unlearning

Inputs: Source Model $w^{\mathcal{S}}$, target dataset $\mathcal{D}^{\mathcal{T}}$, set of forget classes $\mathcal{C}_{\mathcal{F}}$, SFDA loss $\mathcal{L}_{\text{SFDA}}$, trade-off α, learning rates η_1, η_2

for each i, enumerating ${\mathcal{C}_{\mathcal{F}}}^i \in \mathcal{C}_{\mathcal{F}}$ **do**
  if $i = 1$ **then**
    $w^{\mathcal{T}}_u \leftarrow$ AO_SCADA_UL$(w^{\mathcal{S}}, \mathcal{D}^{\mathcal{T}}, {\mathcal{C}_{\mathcal{F}}}^i, \mathcal{L}_{\text{SFDA}}, \alpha, \eta_1, \eta_2)$ ▷ Alg (1)
  else
    $\mathcal{D}^{\mathcal{T}}_{\text{sub}} =$ Subset$(\mathcal{D}^{\mathcal{T}})$
    $w^{\mathcal{T}}_u \leftarrow$ AO_SCADA_UL$(w^{\mathcal{T}}_u, \mathcal{D}^{\mathcal{T}}_{\text{sub}}, {\mathcal{C}_{\mathcal{F}}}^i, \mathcal{L}_{\text{SFDA}}, \alpha, \eta_1, \eta_2)$
  end if
end for

Return: $w^{\mathcal{T}}_u$

labels $\mathcal{C}_{\mathcal{T}}$ and deduce the source-exclusive classes as $\mathcal{C}_{\mathcal{F}} = \mathcal{C}_{\mathcal{S}} \setminus \mathcal{C}_{\mathcal{T}}$.) Such a scenario occurs, for example, when one does not have access to the source domain label space (e.g. fraud categories of a given country in a fraud detection application). Our variants address such a constraint; we introduce the **UC-SCADA-UL** setting (Definition 2), where only the number of source-exclusive classes is assumed to be known. Practically, this quantity can be estimated using domain knowledge, such as the sizes of the source and target label spaces: if $|\mathcal{C}_{\mathcal{S}}|$ and $|\mathcal{C}_{\mathcal{T}}|$ are known, $|\mathcal{C}_{\mathcal{F}}|$ can be computed as $|\mathcal{C}_{\mathcal{S}}| - |\mathcal{C}_{\mathcal{T}}|$, even when specific class identities are unavailable. Additionally, in some scenarios only a subset of $\mathcal{C}_{\mathcal{F}}$ may be known before the adaptation process, with additional classes uncovered over time (eg., through user-initiated data removal requests). The **C-SCADA-UL** setting (Definition 3) formalizes this process by allowing sequential unlearning of disjoint subsets ${\mathcal{C}_{\mathcal{F}}}^i \subseteq \mathcal{C}_{\mathcal{F}}$ across multiple steps, requiring the model to dynamically discard knowledge of ${\mathcal{C}_{\mathcal{F}}}^i$.

A.4. Additional Analysis

A.4.1. Applying Existing MU Methods in SCADA-UL

We experiment with applying adapted versions of existing MU methods [10, 51, 53] in the single class SCADA-UL setting on DomainNet with $c_{\mathcal{F}} = 1$ (see Sec. A.7 for implementation details). We test three approaches: applying the method on the source model and subsequently adapting that model to the target domain ("Source → Unlearn → Adapt" in Table A.7), applying the method during the adaptation process by adding loss terms ("Source → (Unlearn + Adapt)" in Table A.7), and applying the method on the target adapted model ("Source → Adapt → Unlearn" in Table A.7). From the results in Tab. A.7,

Figure A.4. **Visual appearance of Adversarial Samples.** Top row shows examples of samples from the source forget class (Backpack) from OfficeHome Art domain, while the bottom row shows examples of adversarial samples of this class. Although these samples appear as random noise, they are confidently classified as a backpack by the model. Moreover, in additional studies (Fig. A.5), we show that these samples match the forget class even on a representational level)

Table A.7. **Existing Data-Free MU methods struggle in Domain Adaptation.** Existing methods [10, 51, 53] when applied either before (Source → Unlearn → Adapt), during (Source → (Unlearn + Adapt)), or after (Source → Adapt → Unlearn) the domain adaptation process perform poorly in our setting, motivating our method.

Method	**Source → Unlearn → Adapt**		**Source → (Unlearn + Adapt)**		**Source → Adapt → Unlearn**	
	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓
Original (SF(DA)^2 [28])	64.3$_{\pm 1.9}$	35.9$_{\pm 4.0}$	N/A	N/A	N/A	N/A
Retrain	65.2$_{\pm 1.9}$	0.0$_{\pm 0.0}$	N/A	N/A	N/A	N/A
UNSIR [51]	20.6$_{\pm 2.9}$	0.0$_{\pm 0.0}$	60.6$_{\pm 1.9}$	62.1$_{\pm 6.8}$	12.2$_{\pm 8.5}$	0.1$_{\pm 0.1}$
ZSMU [10]	62.5$_{\pm 1.6}$	29.6$_{\pm 7.8}$	62.7$_{\pm 0.7}$	53.7$_{\pm 3.2}$	58.1$_{\pm 13.}$	26.7$_{\pm 10.}$
Nabla Tau [53]	50.1$_{\pm 3.7}$	1.2$_{\pm 2.0}$	52.6$_{\pm 1.8}$	34.1$_{\pm 12.4}$	32.2$_{\pm 5.9}$	1.0$_{\pm 1.4}$

Table A.8. **Cosine similarity (S_C) between features of adversarial samples and real samples.** Adversarial sample features align more with those of the forget class both before and after adaptation. After adaptation, they align more with the target forget

Compared Features	S_C Before Adaptation	S_C After Adaptation
Adversarial - Source Forget	**0.360**	**0.371**
Adversarial - Source Retain	0.326	0.315
Adversarial - Target Forget	**0.359**	**0.366**
Adversarial - Target Retain	0.335	0.307

we see that all the approaches to applying existing methods perform poorly in the SCADA-UL setting. This limitation arises from their design, which does not account for varying data distributions. It highlights the need for a targeted MU solution for the SCADA-UL setting.

A.4.2. Study on Adversarial Samples

Adversarial samples are generated directly by the model through optimization in the input space. Visually, these samples differ significantly from real-samples (See Fig. A.4). However, the model confidently classifies these samples as a certain class. In Sec. 4, we mention that the model maximizes the probability of the adversarial samples belonging to the forget class. While the objective encourages adversarial samples to be classified as the forget class by maximizing its logit, no structural regularization is used to impose visual similarity of the adversarial samples to the forget class. In this section, we further analyze the features of adversarial samples to assess their similarity to the forget classes using cosine similarity. We experiment on a representative task from the DomainNet dataset: Clipart to Sketch.
Tab. A.8 shows a 36% cosine similarity of adversarial sample features with forget class features of both domains. This similarity increases by about 1% after the domain adaptation process, showing that over time the adversarial samples evolve to approximate the forget class better. On the other hand, their similarity with retain class features starts at 32-33% and decreases to 30-31% after adaptation. These results support our claim that adversarial samples are representative of the forget classes in Sec. 4.

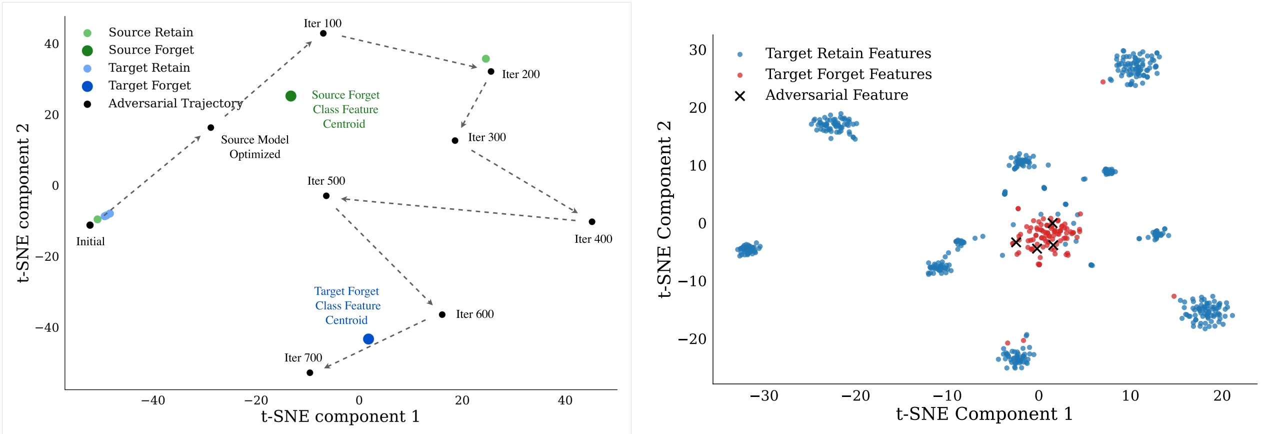

Figure A.5. **t-SNE plots of adversarial, retain, and forget class samples.** Left: Evolution of an adversarial sample over time. Initially, the sample lies far from the forget-class centroids (closer to retain-class centroids). While being optimized on the source model, it moves closer to the source forget-class centroid, and over iterations gradually converges toward the target forget-class centroid. This shows that the samples evolve alongside the model to best fit the class to forget in the target domain. Right: Final adversarial samples compared to t-SNE embeddings of 9 randomly selected retain classes $\{29, 3, 44, 58, 36, 2, 59, 47, 14\}$ and the forget class $c_{\mathcal{F}} = 1$. The final adversarial samples clearly align more closely with features from the forget class. Plotted for OfficeHome Art to Product.

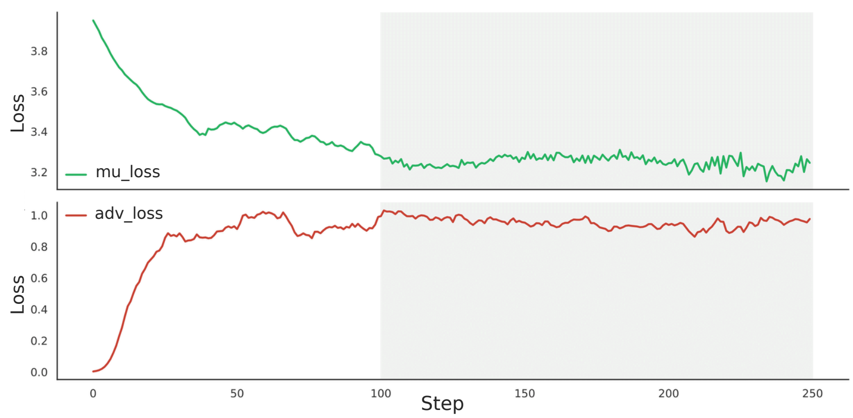

Figure A.6. **MU Loss ($\mathcal{L}_{\text{MU}}$) and Adversarial Loss ($\mathcal{L}_{\text{ADV}}$).** Initially, $\mathcal{L}_{\text{MU}}$ conflicts with $\mathcal{L}_{\text{ADV}}$, but equilibrium is reached after ~ 100 steps. This demonstrates a measured approach to unlearning where the model gradually finds a new representation ϕ that aligns with both objectives.

For qualitative analysis, we also visualize the evolution of adversarial sample features using t-SNE, as shown in Fig. A.5. We plot centroids for source retain and forget class features along with target retain and forget centroids, and overlay the adversarial trajectory across iterations. The visualization shows that adversarial samples begin far from the forget-class features, then gradually move toward the source forget class before shifting further toward the target forget class. This illustrates how the adversarial optimization drives samples to mimic the features of the class to forget as the model evolves from the source to target domain.

A.4.3. Visualizing Unlearning and Adversarial Loss Terms

We show a plot of the Unlearning Loss ($\mathcal{L}_{\text{MU}}$) and Adversarial Loss ($\mathcal{L}_{\text{ADV}}$) in Fig. A.6 for the OfficeHome Real-World-to-Art domain adaptation task. The figure reveals two key trends: firstly, the gradual decline in $\mathcal{L}_{\text{MU}}$ indicates a measured unlearning process, suggesting that the method does not aim to abruptly "erase" the forget classes (a process that still stores latent information about the forget classes). Instead, unlearning occurs adversarially, as evident from the mirrored relationship between $\mathcal{L}_{\text{MU}}$ and $\mathcal{L}_{\text{ADV}}$ where a decrease in one causes an increase in the other. These observations align with our discussion in Sec. 4.2, that is, a higher value of the gradients of $\mathcal{L}_{\text{MU}}$ for the retain classes gradually allows the unlearning of the forget classes during the training process. Together, these results demonstrate that our method performs thorough unlearning.

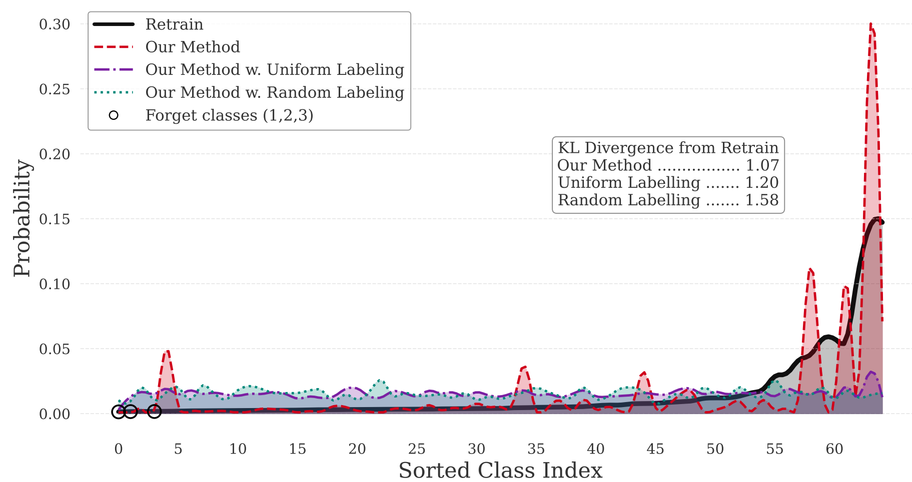

Figure A.7. **Softmax outputs of the target forget set for the retrained model compared with our method using Rescaled, Uniform, and Random Labeling.** After sorting classes by the retrained model's outputs, our method with Rescaled Labeling best matches the retrained distribution, while Uniform and Random Labeling deviate more significantly, as also reflected in their higher KL Divergence values. Note that we use interpolation to make the distributions more clear.

A.4.4. Visualizing Outputs Under Different Labeling Strategies

We visualize the average softmax outputs over the entire target forget dataset $\mathcal{D}_f^{\mathcal{T}}$ for four models: (1) **Retrain**: a model retrained on the source data and adapted to the target domain, (2) **Our Algorithm 1**: a model trained with our procedure using Rescaled Labeling, (3) **Uniform Labeling**: a variant where the rescaled labels $\hat{y}$ are replaced with a uniform distribution over the retain classes $c_{\mathcal{R}}$, and (4) **Random Labeling**: a variant where $\hat{y}$ is replaced with a randomly selected retain class.

Figure A.7 shows the sorted softmax outputs of all four models with respect to the retrained model's ordering. By sorting classes according to the retrained model's output distribution, the visualization highlights how closely each method matches the retrained behavior. Among the three variants of our method, the Rescaled Labeling strategy most closely follows the retrained curve, while Uniform and Random Labeling show more pronounced deviations. This further illustrates that Rescaled Labeling provides the best approximation to the output distribution of the gold-standard Retrain model.

A.4.5. Failure Case Analysis: UC-SCADA-UL

In Tab. A.25, specifically in the Sketch $\rightarrow$ Painting task, we observe that our method achieves a forget accuracy of 29.6%, which is an improvement over the original model's accuracy of 67.6%, but still far from the ideal 0% forget accuracy achieved by retraining.

We identify two key reasons for this gap.

1. **Low γ values for retain classes due to domain shift or inherent difficulty.** For example, classes like Compass, Cow, and Pencil exhibit 0% accuracy on the target domain and correspondingly low γ values (see Tab. A.9). This misleads our algorithm into detecting them as forget classes even though they are actually retain classes.
2. **Semantic or visual similarity between retain and forget classes.** Forget classes such as Great Wall of China receive higher γ values due to contributions from semantically similar classes like castle and streetlight (see Tab. A.10), making them harder to identify correctly as forget classes

A.4.6. Additional Ablation Studies

We performed additional ablation studies on the initialization of adversarial samples, loss trade-off parameter α and the number of adversarial samples, which are reported below. All experiments measure multi-class SCADA-UL performance across 3 trials on the OfficeHome dataset, reporting retain accuracy ($A_{\mathcal{D}_r^{\mathcal{T}}}$) and forget accuracy ($A_{\mathcal{D}_f^{\mathcal{T}}}$) averaged across 12 tasks.

Table A.11. Initializing adv samples

Initialize	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$
✓	$75.1_{\pm 1.3}$	$0.0_{\pm 0.0}$
✗	$75.0_{\pm 1.2}$	$15.2_{\pm 4.1}$

Table A.9. **Target accuracy of lowest γ value classes** (domain shift evident by low target accuracy)

Class Name	Target Acc	γ Value
Compass	0.0	0.010
Cow	0.0	0.012
Pencil	0.0	0.013
Peas	2.5	0.018
Cell Phone	0.0	0.022

Table A.10. **Top contributing classes for γ value of forget classes.** (semantic similarity is observed)

Forget Class	Candidate 1	Candidate 2	Candidate 3
Great Wall of China	castle	train	streetlight
Aircraft Carrier	submarine	cruise_ship	helicopter
Alarm Clock	compass	watermelon	dog

Table A.12. Effect of varying Loss Trade-off (α) hyperparameter study

α	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$
1.0	$75.0_{\pm 1.2}$	$15.2_{\pm 4.1}$
5.0	$75.2_{\pm 1.1}$	$2.6_{\pm 2.0}$
10.0	$75.1_{\pm 1.3}$	$0.0_{\pm 0.0}$
20.0	$75.2_{\pm 1.1}$	$0.0_{\pm 0.0}$

Table A.13. Study on # adversarial samples used during training

N_{adv}	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$
1	$70.9_{\pm 3.6}$	$0.8_{\pm 1.4}$
2	$72.9_{\pm 3.7}$	$0.0_{\pm 0.0}$
4	$75.1_{\pm 1.3}$	$0.0_{\pm 0.0}$
8	$75.2_{\pm 1.2}$	$0.0_{\pm 0.0}$
16	$75.3_{\pm 1.1}$	$0.0_{\pm 0.0}$

Table A.14. Study on number of train epochs

Epochs	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$
1	$75.6_{\pm 1.4}$	$18.2_{\pm 8.1}$
5	$75.1_{\pm 1.3}$	$0.0_{\pm 0.0}$
10	$74.9_{\pm 1.2}$	$0.4_{\pm 0.3}$

Initialization of Adversarial Samples. Tab. A.11 shows that initializing samples by a complete minimization of $\mathcal{L}_{\text{ADV}}$ on the source model is crucial for effective unlearning. This step enables the samples to be good representations of the forget classes, which are subsequently refined via gradient updates.
Loss Trade-off (α). As seen in Tab. A.12, balancing $\mathcal{L}_{\text{SFDA}}$ and $\mathcal{L}_{\text{MU}}$ requires $\alpha > 5.0$ to ensure thorough unlearning of forget classes, while values 10.0 and 20.0 yield comparable performance. This behavior can be explained by the distinct optimization objectives of the two losses, which stabilize after initial iterations (see Fig. A.6).
Below the threshold (5.0–10.0), $\mathcal{L}_{\text{SFDA}}$ marginally dominates $\mathcal{L}_{\text{MU}}$, leading to suboptimal unlearning performance.
Number of Adversarial Samples. Tab. A.13 shows that using very few adversarial samples, even 2 in this case, results in good unlearning performance. As the number of adversarial samples increases, retain accuracy improves, reaching close to its maximum value at 4 samples (which we use for our experiments in the main paper). Beyond this point, increasing the number of samples further does not lead to considerable improvement.
Number of Training Epochs. From Tab. A.14, we observe that using fewer epochs (e.g., 1) leads to incomplete forgetting, as seen by the higher forget accuracy. On the other hand, a larger number of epochs (e.g., 10) results in a slight drop in retain accuracy, likely due to continued application of $\mathcal{L}_{MU}$ negatively impacting the retained classes after the forget classes have already been unlearned. We use 5 epochs in our main experiments, as it provides a favorable trade-off, achieving complete forgetting while maintaining high retain accuracy.

A.4.7. Discussion of Assumptions in Our Method

Our method follows that (i) datasets with disjoint label spaces are conditionally independent given the model weights and (ii) source datasets can be approximated by their corresponding trained models. The use of (i) is explained and justified with an example at the end of Section 4.1. We further elaborate on (ii) here. For the approximation used in (ii), we follow such approximations used in Elastic Weight Consolidation [3] which employs a (diagonal) Laplace approximation to estimate posteriors across sequential tasks, and in online learning scenarios [46]. This approximation follows by assuming the posterior $p(w \mid \mathcal{D}^{\mathcal{S}})$ is approximately Gaussian and centered at the MLE estimate $w^{\mathcal{S}}$. i.e., $p(w \mid \mathcal{D}^{\mathcal{S}}) \approx \mathcal{N}(w^{\mathcal{S}}, \Sigma)$. By Bayesian inference, $p(w \mid \mathcal{D}^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}}) \propto p(\mathcal{D}_r^{\mathcal{T}} \mid w) \cdot p(w \mid \mathcal{D}^{\mathcal{S}})$. Substituting the Gaussian approximation yields $p(w \mid \mathcal{D}^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}}) \propto p(\mathcal{D}_r^{\mathcal{T}} \mid w) \cdot \mathcal{N}(w^{\mathcal{S}}, \Sigma) \approx p(w \mid w^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}})$. Similarly, we obtain $p(w \mid \mathcal{D}_r^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}}) \approx p(w \mid w_r^{\mathcal{S}}, \mathcal{D}_r^{\mathcal{T}})$.

A.5. Additional Experimental Results

A.5.1. Runtime Analysis of Methods

Table A.15 presents the runtime comparison of all methods for unlearning and adaptation to the target domain. While our method does not consistently achieve the lowest runtime across all datasets, its time cost remains significantly lower than that of retraining. We especially note that on the larger dataset setting (DomainNet), our method achieves the lowest training time, indicating its potential efficiency in large-scale settings. Additionally, the inference time is identical across all methods, as no extra computation is required during model inference.

Table A.15. Training Time for Each Method (in seconds)

Method	OfficeHome	Office31	DomainNet
Original (SF(DA)^2 [28])	307.9$_{\pm 0.6}$	239.9$_{\pm 2.0}$	623.6$_{\pm 5.3}$
Retrain	696.7$_{\pm 1.0}$	665.8$_{\pm 2.7}$	1109.3$_{\pm 1.9}$
Finetune	616.2$_{\pm 2.0}$	476.7$_{\pm 0.7}$	1239.0$_{\pm 2.3}$
UNSIR [51]	364.0$_{\pm 3.3}$	296.1$_{\pm 1.9}$	710.3$_{\pm 2.6}$
ZSMU [10]	403.9$_{\pm 0.2}$	334.5$_{\pm 1.3}$	738.2$_{\pm 3.2}$
Lipschitz [15]	484.8$_{\pm 2.8}$	419.4$_{\pm 0.4}$	789.5$_{\pm 0.5}$
Nabla Tau [53]	370.3$_{\pm 2.6}$	301.6$_{\pm 0.5}$	757.1$_{\pm 4.3}$
Unlearned(+) [2]	1073.8$_{\pm 514}$	396.1$_{\pm 59.}$	3017$_{\pm 1319}$
PADA [5]	**319.3**$_{\pm 1.1}$	**253.6**$_{\pm 2.1}$	691.2$_{\pm 1.0}$
SHOT [34]	560.1$_{\pm 0.3}$	538.8$_{\pm 1.3}$	730.6$_{\pm 1.6}$
Ours	382.2$_{\pm 5.4}$	319.6$_{\pm 2.0}$	**691.0**$_{\pm 2.7}$

Table A.16. SCADA-UL performance on Land-use Classification: UCMerced → RSSCN7

Method	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	Score ↑
Original (SF(DA)^2 [28])	76.2$_{\pm 1.3}$	24.3$_{\pm 28.}$	0.64$_{\pm 0.1}$
Retrain	76.2$_{\pm 3.8}$	0.0$_{\pm 0.0}$	0.76$_{\pm 0.0}$
Finetune	**76.2**$_{\pm 1.0}$	35.3$_{\pm 23.}$	0.57$_{\pm 0.1}$
UNSIR [51]	63.6$_{\pm 10.}$	**0.0**$_{\pm 0.0}$	0.64$_{\pm 0.1}$
ZSMU [10]	**76.2**$_{\pm 1.8}$	**0.0**$_{\pm 0.0}$	**0.76**$_{\pm 0.0}$
Lipschitz [15]	75.9$_{\pm 1.3}$	18.3$_{\pm 24.}$	0.66$_{\pm 0.1}$
Nabla Tau [53]	62.3$_{\pm 2.5}$	**0.0**$_{\pm 0.0}$	0.62$_{\pm 0.0}$
Unlearned(+) [2]	63.4$_{\pm 0.8}$	0.9$_{\pm 1.3}$	0.63$_{\pm 0.0}$
PADA [5]	64.4$_{\pm 1.4}$	2.7$_{\pm 3.8}$	0.63$_{\pm 0.0}$
SHOT [34]	65.2$_{\pm 1.9}$	14.5$_{\pm 0.9}$	0.57$_{\pm 0.0}$
Ours	75.9$_{\pm 1.0}$	**0.0**$_{\pm 0.0}$	**0.76**$_{\pm 0.0}$

A.5.2. Experiments on Land-use Classification Dataset

In Tab. A.16, we studied our approach on a land-use classification domain adaptation benchmark UC Merced [65] → RSSCN7 [71], similar to [49]. Privacy is a critical concern in such settings as certain categories of scenes (for e.g., government facilities) present in the source domain trained model must not be adapted to the target domain. Table A.16 indicates our method achieves good results on this real-world dataset.

A.5.3. SCADA Unlearning

Tables A.21 to A.23 present results for single-class SCADA unlearning on the OfficeHome, DomainNet and Office31 datasets respectively. Multi-class SCADA unlearning results are presented in Tabs. A.18 to A.20 tested on $\mathcal{C}_\mathcal{F} = \{1, 2, 3\}$. This is an extended version of Tab. 2 presented in the main paper, showing results for all the tasks within each dataset. Our method outperforms baselines w.r.t. all the metrics by achieving accuracies and unlearn score close to that of the retrained model in both single-class and multi-class unlearning settings.

A.5.4. UC-SCADA Unlearning

Tables A.24 to A.26 show the results in the UC-SCADA-UL setting (Def 2) with $\mathcal{C}_\mathcal{F} = \{1, 2, 3\}$, but these source-exclusive classes unknown to the model. Results show that our method outperforms the baselines w.r.t. the unlearn score metric on the OfficeHome and Office31 datasets. It maybe noted that finetuning is typically found to perform the best w.r.t. $A_{\mathcal{D}_r^\mathcal{T}}$ but its performance w.r.t. $A_{\mathcal{D}_f^\mathcal{T}}$ is poor. Similarly, w.r.t. $A_{\mathcal{D}_f^\mathcal{T}}$, UNSIR performs well in the UC-SCADA-UL setting, but it performs poorly w.r.t. $A_{\mathcal{D}_r^\mathcal{T}}$. On the DomainNet dataset, our method struggles with $A_{\mathcal{D}_f^\mathcal{T}}$, which is likely due to the inaccuracy in identifying the forget classes reliably using the γ parameter. We elaborate more on this in Sec. A.6.

A.5.5. C-SCADA Unlearning

Tables A.27 to A.29 show the results in the C-SCADA-UL setting (Def 3) where the source-exclusive classes are revealed over multiple time steps. For the experiments, we use $\mathcal{C_F}^1 = \{1, 2\}, \mathcal{C_F}^2 = \{3, 4\}, \mathcal{C_F}^3 = \{5, 6\}$. The accuracies presented are cumulative over all forget classes until the current time step (for e.g., $A_{\mathcal{D}_f^\mathcal{T}}$ in T2 is the forget accuracy over classes $\mathcal{C_F}^1 \cup \mathcal{C_F}^2 = \{1, 2, 3, 4\}$). This allows us to evaluate how effectively the model forgets the newly designated class sets while maintaining unlearning of the older forget classes. The results show that our method consistently achieves thorough unlearning of the new class sets at each step, outperforming the baselines w.r.t. the unlearn score metric.

A.5.6. Use of Other SFDA Loss Functions

Table A.17 shows results of experiments with two other SFDA methods: SHOT [34] and UCon-SFDA [60]. The target domain performance while using just the SFDA method *(Original)* and the target domain performance while using this method as $\mathcal{L}_{\text{SFDA}}$ in our proposed method are provided in terms of percentage improvement over the corresponding Unlearn Score metric of the previous method going from top to bottom. For example, Rows 5, 6 show our method achieves a 6.1% improvement in the Score while using UCon-SFDA as $\mathcal{L}_{\text{SFDA}}$ instead of (SF(DA)^2, when UCon-SFDA shows a 15.7% improvement in target domain performance over (SF(DA)^2. This indicates that while our method works well with different SFDA losses, the performance of our method improves with the SFDA method used. Table A.17 reports the metrics over all 7 source, target pairs in DomainNet-126 for a single forget class.

Table A.17. **Improved SFDA methods improve our method as well.** We observe that as the underlying SFDA method achieves better performance, our method also improves. This shows the adaptability of our approach to new SFDA loss terms and its potential to further benefit as future SFDA methods advance.

SFDA Loss	Method	Metric	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Average	% Improved
SHOT [34]	Original	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	70.1$_{\pm 1.1}$	57.4$_{\pm 1.9}$	61.1$_{\pm 1.0}$	83.7$_{\pm 0.4}$	60.6$_{\pm 3.4}$	69.1$_{\pm 1.0}$	75.5$_{\pm 0.6}$	68.2$_{\pm 1.3}$	-
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	88.7$_{\pm 0.9}$	47.0$_{\pm 20.6}$	57.8$_{\pm 2.5}$	86.7$_{\pm 0.1}$	27.7$_{\pm 4.8}$	41.8$_{\pm 0.6}$	43.1$_{\pm 2.3}$	56.1$_{\pm 4.5}$	
		Score $\uparrow$	0.37$_{\pm 0.0}$	0.39$_{\pm 0.0}$	0.39$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.53$_{\pm 0.0}$	0.44$_{\pm 0.0}$	
	Our Method	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	56.0$_{\pm 1.1}$	36.6$_{\pm 4.0}$	42.7$_{\pm 0.2}$	77.9$_{\pm 5.5}$	27.1$_{\pm 5.3}$	49.6$_{\pm 0.8}$	66.4$_{\pm 5.2}$	50.9$_{\pm 3.2}$	-
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	
		Score $\uparrow$	<u>0.56</u>$_{\pm 0.0}$	0.37$_{\pm 0.0}$	0.43$_{\pm 0.0}$	<u>0.78</u>$_{\pm 0.1}$	0.27$_{\pm 0.1}$	0.50$_{\pm 0.0}$	0.66$_{\pm 0.1}$	0.51$_{\pm 0.0}$	
SF(DA)^2 [28]	Original	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	71.3$_{\pm 0.3}$	66.6$_{\pm 0.4}$	61.7$_{\pm 0.8}$	78.3$_{\pm 0.2}$	55.9$_{\pm 1.5}$	65.1$_{\pm 0.9}$	75.0$_{\pm 0.4}$	67.7$_{\pm 0.6}$	+15.9%
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	77.6$_{\pm 5.7}$	36.9$_{\pm 6.9}$	35.6$_{\pm 1.3}$	74.1$_{\pm 12.1}$	3.2$_{\pm 0.4}$	17.5$_{\pm 1.0}$	6.1$_{\pm 0.3}$	35.9$_{\pm 4.0}$	
		Score $\uparrow$	0.40$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.46$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.54$_{\pm 0.0}$	0.55$_{\pm 0.0}$	0.71$_{\pm 0.0}$	0.51$_{\pm 0.0}$	
	Our Method	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	67.8$_{\pm 2.3}$	62.7$_{\pm 1.0}$	58.9$_{\pm 1.0}$	77.0$_{\pm 0.4}$	54.0$_{\pm 1.7}$	62.1$_{\pm 0.8}$	74.0$_{\pm 1.0}$	65.2$_{\pm 1.1}$	+27.4%
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	
		Score $\uparrow$	0.68$_{\pm 0.0}$	0.63$_{\pm 0.0}$	0.59$_{\pm 0.0}$	0.77$_{\pm 0.0}$	0.54$_{\pm 0.0}$	0.62$_{\pm 0.0}$	0.74$_{\pm 0.0}$	0.65$_{\pm 0.0}$	
UCon-SFDA [60]	Original	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	80.1$_{\pm 0.8}$	73.9$_{\pm 0.1}$	77.7$_{\pm 1.2}$	88.6$_{\pm 0.0}$	73.7$_{\pm 0.2}$	77.3$_{\pm 1.2}$	82.1$_{\pm 0.1}$	79.0$_{\pm 0.5}$	+15.7%
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	78.7$_{\pm 14.2}$	52.7$_{\pm 4.8}$	56.5$_{\pm 1.8}$	21.8$_{\pm 10.0}$	14.5$_{\pm 2.4}$	28.9$_{\pm 4.2}$	15.7$_{\pm 0.0}$	38.4$_{\pm 5.4}$	
		Score $\uparrow$	0.45$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.73$_{\pm 0.1}$	0.64$_{\pm 0.0}$	0.60$_{\pm 0.0}$	0.71$_{\pm 0.0}$	0.59$_{\pm 0.0}$	
	Our Method	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	71.7$_{\pm 2.4}$	67.8$_{\pm 0.2}$	60.6$_{\pm 0.6}$	83.6$_{\pm 0.4}$	56.1$_{\pm 0.5}$	64.9$_{\pm 0.7}$	75.4$_{\pm 0.1}$	68.6$_{\pm 0.7}$	+6.1%
		$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	
		Score $\uparrow$	0.72$_{\pm 0.0}$	0.68$_{\pm 0.0}$	0.61$_{\pm 0.0}$	0.84$_{\pm 0.0}$	0.56$_{\pm 0.0}$	0.65$_{\pm 0.0}$	0.75$_{\pm 0.0}$	0.69$_{\pm 0.0}$	

A.5.7. Extensions to Open-set Domain Adaptation

in Table A.30, we provide results on the open-set DA setting, replacing the SFDA loss term with the open-set DA version of SHOT [34] as described in the paper. This is evaluated on the OfficeHome dataset with source-only classes $\{1, 2, 3\}$ and target-only classes $\{34, 35, 36, 37, 38\}$. We report full target class accuracy (OS*), shared class accuracy (OS), forget accuracy, and unlearn score. The results show largely similar trends to the original SFDA setting, where existing methods perform poorly by dropping retain accuracy significantly (UNSIR, Lipschitz) or still maintaining high forget accuracy (ZSMU). In contrast, our method demonstrates strong unlearning performance while maintaining high retain accuracy, achieving the best overall unlearn score.

A.6. Limitations

Our work addresses the task of adapting to new domains while unlearning a subset of classes present only in the source domain. Although our proposed method demonstrates strong performance in the SCADA-UL setting (including its extensions C-SCADA-UL and UC-SCADA-UL), our studies have been currently limited to image classification-based domain adaptation tasks. Domain adaptation is also popular for other tasks such as semantic segmentation or generative modeling. For example, adapting language models to company-specific documents to enhance accuracy and avoid irrelevant or incorrect responses, or adapting road sign segmentation models to new geographies containing only a subset of signs. Extending our method to such directions needs to be studied carefully, and would be interesting directions of future work.

Beyond the above, in our current implementation, we adapt the γ term from PADA [5] for identification of source-exclusive classes. However, we find that this is not always accurate, especially in datasets with large label-spaces such as DomainNet. Our solution for this is to select an excess number of classes and unlearn all of them to increase the likelihood for source-exclusive classes to be included. Better approaches to identify forget classes or unified methods (that don't require knowledge of forget classes) may help the UC-SCADA-UL setting in particular. Finally, although we present a theorem linking our method to gradients, establishing mathematical guarantees for unlearning itself is challenging, and would be a potential direction for future work.

A.7. Implementation Details

A.7.1. Compute Resources

Our experiments were executed on a Linux-based compute cluster each using a single Tesla V100-SXM3-32GB GPU limited to 20 CPU workers. The time taken for running each experiment ranges from 300-1500 seconds.

A.7.2. Adapting MU and PDA Methods to SCADA-UL

We applied the baseline methods during the domain adaptation process when feasible, similar to our approach. Otherwise, we evaluated both the "before DA" and "after DA" variants of each method, and adopted the variant that performed better with respect to the Unlearn Score metric. This turned out to be "before DA" for all these methods, since this variant preserved the majority of retain accuracy (and hence higher unlearn score). As none of the existing MU methods have been implemented on the OfficeHome, Office31 or DomainNet datasets, we found the best hyperparameters for all the baseline MU methods on these datasets. The hyperparameters were tuned to maximize the unlearn score of each method. **Retrain.** The source model is retrained without the forget data and this model is domain adapted. Since source data is inaccessible in our setting, this only serves as a gold standard for comparison. **Finetune.** The source model is finetuned on the retain data using the SFDA loss. **UNSIR (Unlearning by Selective Impair and Repair).** employs an impair step to reduce the model's performance on the forget class followed by a repair step to restore performance on the retain set. As this method was originally designed for unlearning where labeled retain data is available, we adapt it for application to our setting. While the noise generation step remains unchanged, both impair and repair steps use target data with pseudo-labels as $\mathcal{D}_{r_{sub}}$, as this represents the closest approximation to data in our setting. The algorithm is applied to the source model before the domain adaptation process. **ZSMU (Zero-shot Machine Unlearning).** It uses error minimizing-maximizing noise to achieve data-free unlearning. This method is already data-free and is therefore applied directly to the source model before the DA process. **Lipschitz Unlearning.** It achieves unlearning by enforcing local Lipschitz regularization, minimizing the change in model outputs with respect to perturbations of the forget samples. As this method utilizes a simple loss function, this loss term is directly swapped with $\mathcal{L}_{\text{MU}}$ in our method. Moreover, since it requires forget samples, they were replaced with adversarial samples generated from our method. **Nabla Tau.** It applies adaptive gradient ascent to forget data along with gradient descent on retain data. This method was also applied to the source model and adversarial forget and retain samples were used. **Unlearned(+).** It estimates the influence of forget data on model parameters and removes it. The influence function involves a gradient computed over the forget set and hessian estimated over the retain set. To apply the method in our setting, adversarial samples were generated to compute the forget data gradient over the source model, and pseudo-labeled retain data was used to estimate the hessian. The method was applied on the source model, and after a short SFDA warmup run, it was re-applied with gentler hyperparameters, and then SFDA was resumed until convergence. Taking the additional, small MU step helped to reduce the forget accuracy more, while maintaining retain accuracy. **PADA.** It introduces a γ term to downweight the contributions of source-exclusive classes during PDA. The method was modified to fit into our setting by applying the proposed class weight vector on all loss terms in SF(DA)^2, alleviating the need for access to source data. **SHOT.** It freezes the source classifier (hypothesis) and adapts the feature encoder for target domain learning. We used the SFPDA version of SHOT as described in the paper by setting the β term to 0.

A.7.3. Hyperparameters

Forget Classes. For single class unlearning in SCADA-UL, we set the forget class $c_{\mathcal{F}} = 1$, for multiple class forgetting in SCADA-UL, we set it to $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$, and for C-SCADA-UL, we set ${\mathcal{C}_{\mathcal{F}}}^1 = \{1, 2\}$, ${\mathcal{C}_{\mathcal{F}}}^2 = \{3, 4\}$, ${\mathcal{C}_{\mathcal{F}}}^3 = \{5, 6\}$. For scenes dataset, we used $c_{\mathcal{F}} = 1$ corresponding to class Grasslands. For medical dataset, we used $c_{\mathcal{F}} = 3$ corresponding to class Edema. For ablation studies, we used $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$ on the OfficeHome dataset. **Data split.** We used an 80-20 split for train and test data in all experiments. **Backbone.** vit-base-patch16-224 pretrained on ImageNet-1K. **Optimizer.** SGD optimizer with learning rate 1e-2, momentum 0.9, weight decay 1e-3 and nesterov set to True. **LR Scheduler.** Lambda scheduler with gamma set to 1e-3 and decay set to 0.9. **Epochs.** Source model is trained for 10 Epochs with 1000 steps per epoch using label smoothing with coefficient 0.1 (As it is standard practice in SFDA [28]). For SCADA-UL, we use 5 epochs and 1000 steps per epoch on OfficeHome, DomainNet, Medical, Scene datasets, and 10 epochs for Office31 on most methods including ours. **Loss Trade-off (α).** We used $\alpha = 10.0$ for most experiments with SF(DA)^2 [28] as $\mathcal{L}_{\text{SFDA}}$, for SHOT [34], we used $\alpha = 150.0$. **Method Hyperparameters.** As none of the baseline MU methods had implementations on any of our tested datasets, we found the best hyperparameters on these datasets. The hyperparameters were tuned to maximize retain-set accuracy and hence the unlearn score of each method. **Our Method.** We used 4 adversarial samples to compute $\mathcal{L}_{\text{MU}}$ in our method. **UNSIR.** We used 256 pseudo-labeled target samples with 32 noisy samples, trained over 5 epochs and 8 steps per epoch with learning rate

1e-1, alpha 2e-3, and mean vector $[1, 2, 3]$. We used SGD optimizer with learning rate 1e-1, momentum 0.9, weight decay 1e-3, and nesterov set to True for both impair and repair steps. Impair was done for only 8 batches of data as we found this led to minimum degradation in retain class performance. **ZSMU.** The error minimizing-maximizing noise consisted of 64 error minimizing samples and 32 error maximizing samples. We used SGD optimizer with learning rate 2e-1, momentum 0.9, weight decay 1e-3 and nesterov set to True and a Lambda scheduler with gamma set to 1e-3 and decay set to 0.75. The ZSMU process was applied for 4 epochs with 8 batches per epoch. **Lipschitz.** Optimizer and LR scheduler are identical to those in SCADA-UL. 4 adversarial samples were generated per forget class and every iteration, each of these samples were iterated through and perturbed five times and the loss was computed. For example, for 3 forget classes, a total of 3 x 4 = 12 adversarial samples were generated and each of these samples were perturbed 5 times (12 x 5 = 60 perturbations). The losses for each perturbation were added together giving us $\mathcal{L}_{\text{MU}}$ for that step. In the UC-SCADA-UL setting for domain net, we use 3 perturbations instead of 5 due to VRAM limitations. **Nabla Tau.** SGD optimizer with learning rate 5e-2, momentum 0.9, weight decay 1e-3 and nesterov set to True, lambda scheduler with gamma set to 1e-3 and decay set to 0.75. The alpha term was set to 0.02. This method was run for only 100 steps due to its poor performance for larger number of steps. **Unlearned(+).** We froze the backbone and linearized only the final ViT classifier layer, ran the Hessian solver for 10 inner iterations with a step size of 0.10 and an L2 curvature penalty of 3e-3; we scaled the one-hot forget targets to 0.7, clipped the tangent update at a global norm of 1e-3, and injected Gaussian noise of 1e-5 before writing the update back. Each forget class was represented by 12 adversarial samples that we re-optimized every iteration, and we drew 12 such minibatches per MU phase. We ran a single-epoch SFDA warmup and then applied a small MU phase configured identically except for a smaller step: weight decay 5e-3, step size 0.03, max update norm 8e-4; post this the SFDA run was continued. **PADA.** Identical optimizer and scheduler as SCADA-UL, the gamma term was computed over the entire target retain train dataset. **SHOT.** The method used a SGD optimizer with learning rate of backbone 1e-1, and fully connected and bottleneck and rotation classifier with 1.0. Lambda lr scheduler was used with gamma 1e-3 and decay 0.9. The adaptation process was run for 10 epochs with 1000 steps per epoch. In UC-SCADA-UL, given the inherent noise in the estimation of γ, we conservatively select a larger set of classes beyond the forget set. Empirically, we find that selecting the $3 \cdot |\mathcal{C}_{\mathcal{F}}|$ bottom-ranked classes yields the best results.

A.7.4. Metrics

MIA Accuracy (MIA%). As mentioned in Sec. 5, we extend the Membership Inference Attack Accuracy (MIA%) metric [22] to suit class-level unlearning. The model is trained to discriminate between the output entropies of retain class data and unseen or out-of-domain (OOD) data. This is implemented by selecting semantically non-overlapping class samples from other datasets (e.g., selecting OfficeHome samples of classes such as calendar, curtains, etc., when experimenting on DomainNet). An ideal method (such as retraining) would be unable to distinguish between the forget class and such OOD classes.

Forget Class False Negative Rate (Forget FNR): This gives the ratio of forget class samples classified as retain samples to the total forget class samples. Given a threshold k, the sample is classified as forget class if the predicted softmax score is less than k, else as retain class. Figure A.8 shows as we increase the threshold, the forget FNR decreases as fewer forget samples are classified as retain classes, as expected.

Retain Class False Positive Rate (Retain FPR): This gives the ratio of retain class samples classified as forget samples to the total retain class samples. Given a threshold k, the sample is classified as forget class if the predicted softmax score is less than k, else as retain class. Figure A.8 shows as we increase the threshold, the retain FPR increases as more retain samples are classified as forget classes, but it rises upto a maximum value of 0.2 in Clipart $\rightarrow$ Art and 0.12 in Art $\rightarrow$ Product, showing that the retain FPR of our method remains low across multiple thresholds. This indicates our method maintains retain class performance robustly across thresholds.

Figure A.8 also shows the forget FNR vs retain FPR plot. The plot indicates it is easy to choose a threshold which maximizes forget FNR with very low retain FPR, for instance 0.18 in Art $\rightarrow$ Product which maintains the forgetting to model utility trade-off in our method.

Calibration Error: Expected Calibration Error (ECE) quantifies the alignment between the model's predicted confidence and its actual accuracy (lower is better). We plot ECE on source model and the model after applying our method for both retain and forget splits: retain bars remain low after unlearning with our method, indicating the retain-class calibration is maintained, i.e., does not degrade reliability on retain classes, whereas forget ECE rises substantially, showing the model loses calibration on forget classes, thus becoming uncertain on forget class samples,

A.7.5. Real-world Dataset Implementations

Scenes Dataset. Similar to [49], we map 7 similar classes of UC Merced and RSSCN7 datasets (golf course $\rightarrow$ grassland, agricultural $\rightarrow$ farmland, storage tanks $\rightarrow$ industrial region, river $\rightarrow$ river and lake, forest $\rightarrow$ forest, dense residential $\rightarrow$

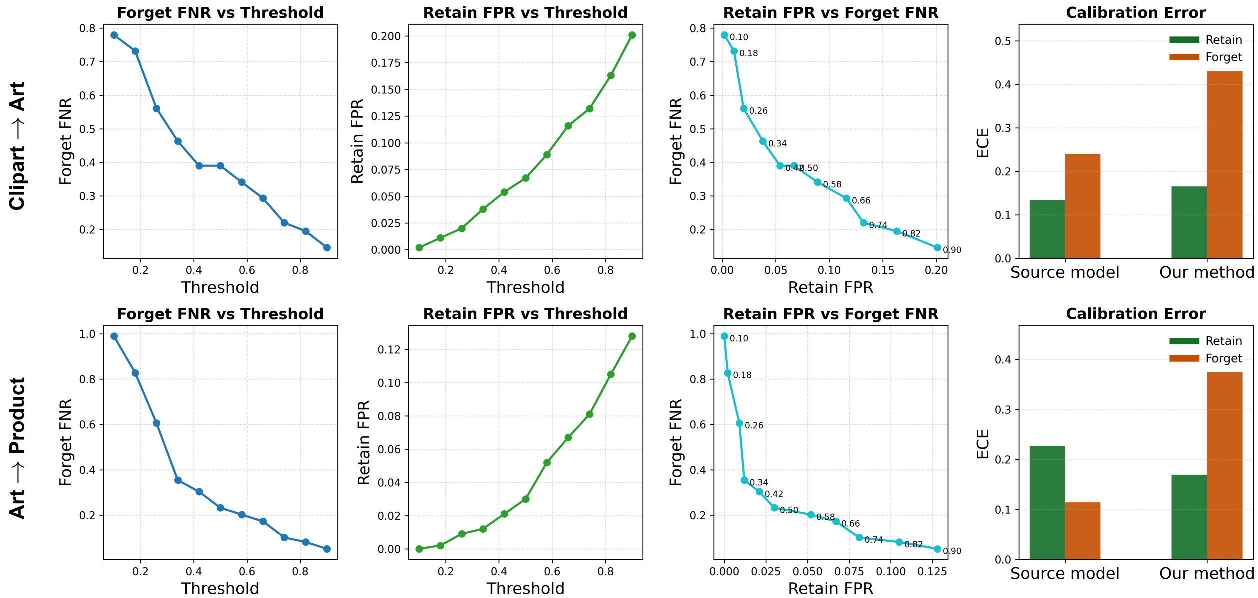

Figure A.8. **Forget/Retain Error and Calibration Analysis.** The figure shows the False Negative Rate (Forget FNR) of our method on forget-class samples across thresholds (first column); the False Positive Rate (Retain FPR) on retain-class samples across thresholds (second column); Forget FNR vs. Retain FPR (third column); and calibration error on retain and forget samples before and after unlearning (fourth column) for two OfficeHome settings. Our method achieves high Forget FNR (the model is confused on forget samples) and low Retain FPR (retain samples remain correctly classified) at appropriate thresholds. Calibration error is high for both forget and retain classes on the source model; after unlearning, retain-class calibration error stays low while forget-class calibration error increases, indicating greater model uncertainty on forget samples.

residential region, parking lot → parking lot). The forget class in our experiments is grassland/golf course. **Medical Dataset.** Similar to [27], we used 5 overlapping classes in both the datasets to conduct our experiments (Atelectasis, Cardiomegaly, Effusion, Consolidation and Edema). These datasets had some samples with multiple labels, we discarded these samples and retained samples with only single labels. Moreover, we used subsets of the datasets: ∼80,000 images for CheXpert , and ∼12,000 images for NIH Chest X-ray .

https://www.kaggle.com/datasets/ashery/chexpert
https://www.kaggle.com/datasets/khanfashee/nih-chest-x-ray-14-224x224-resized

Table A.18. **Results for Multi-Class SCADA Unlearning on OfficeHome.** Forget classes are $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	A → C	A → P	A → R	C → A	C → P	C → R	P → A	P → C	P → R	R → A	R → C	R → P	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	61.0$_{\pm2.0}$	82.2$_{\pm1.9}$	82.3$_{\pm0.4}$	76.9$_{\pm2.1}$	82.9$_{\pm1.2}$	82.5$_{\pm1.1}$	75.2$_{\pm1.7}$	58.2$_{\pm1.4}$	84.3$_{\pm1.6}$	78.5$_{\pm0.2}$	58.4$_{\pm1.8}$	87.0$_{\pm0.2}$	75.8$_{\pm1.3}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	28.7$_{\pm2.3}$	71.1$_{\pm6.2}$	86.1$_{\pm1.2}$	57.1$_{\pm1.4}$	69.0$_{\pm1.6}$	76.7$_{\pm2.8}$	47.8$_{\pm2.3}$	22.0$_{\pm0.5}$	72.8$_{\pm6.9}$	53.1$_{\pm4.2}$	35.5$_{\pm2.3}$	77.1$_{\pm2.0}$	58.1$_{\pm2.8}$
	Score ↑	0.47$_{\pm0.0}$	0.48$_{\pm0.0}$	0.44$_{\pm0.0}$	0.49$_{\pm0.0}$	0.49$_{\pm0.0}$	0.47$_{\pm0.0}$	0.51$_{\pm0.0}$	0.48$_{\pm0.0}$	0.49$_{\pm0.0}$	0.51$_{\pm0.0}$	0.43$_{\pm0.0}$	0.49$_{\pm0.0}$	0.48$_{\pm0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	60.8$_{\pm1.4}$	83.1$_{\pm0.8}$	82.2$_{\pm0.7}$	78.5$_{\pm2.5}$	82.6$_{\pm1.2}$	83.3$_{\pm1.8}$	76.4$_{\pm1.2}$	59.7$_{\pm3.1}$	84.4$_{\pm0.7}$	77.6$_{\pm1.6}$	59.9$_{\pm2.0}$	87.3$_{\pm0.5}$	76.3$_{\pm1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$	0.0$_{\pm0.0}$
	Score ↑	0.61$_{\pm0.0}$	0.83$_{\pm0.0}$	0.82$_{\pm0.0}$	0.79$_{\pm0.0}$	0.83$_{\pm0.0}$	0.83$_{\pm0.0}$	0.76$_{\pm0.0}$	0.60$_{\pm0.0}$	0.84$_{\pm0.0}$	0.78$_{\pm0.0}$	0.60$_{\pm0.0}$	0.87$_{\pm0.0}$	0.76$_{\pm0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**60.1**$_{\pm1.3}$	81.5$_{\pm1.3}$	**82.6**$_{\pm0.8}$	**77.8**$_{\pm2.1}$	$\underline{83.1}_{\pm0.3}$	**83.2**$_{\pm1.3}$	**76.5**$_{\pm0.7}$	58.3$_{\pm1.8}$	$\underline{84.4}_{\pm1.6}$	**79.3**$_{\pm0.5}$	58.6$_{\pm1.1}$	$\underline{87.3}_{\pm0.4}$	**76.1**$_{\pm1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	24.8$_{\pm3.8}$	66.5$_{\pm5.4}$	82.0$_{\pm2.7}$	44.5$_{\pm1.6}$	61.4$_{\pm3.0}$	75.1$_{\pm2.0}$	36.7$_{\pm1.9}$	$\underline{1.5}_{\pm0.5}$	59.6$_{\pm3.8}$	43.5$_{\pm4.0}$	21.3$_{\pm0.9}$	72.9$_{\pm2.8}$	49.2$_{\pm2.7}$
	Score ↑	0.48$_{\pm0.0}$	0.49$_{\pm0.0}$	0.45$_{\pm0.0}$	0.54$_{\pm0.0}$	0.51$_{\pm0.0}$	0.48$_{\pm0.0}$	0.56$_{\pm0.0}$	**0.57**$_{\pm0.0}$	0.53$_{\pm0.0}$	0.55$_{\pm0.0}$	0.48$_{\pm0.0}$	0.50$_{\pm0.0}$	0.51$_{\pm0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	25.4$_{\pm3.5}$	58.0$_{\pm2.5}$	48.1$_{\pm4.2}$	16.8$_{\pm11.}$	55.4$_{\pm13.}$	48.4$_{\pm2.7}$	11.6$_{\pm7.2}$	15.1$_{\pm11.}$	37.7$_{\pm12.}$	11.3$_{\pm7.2}$	28.6$_{\pm3.0}$	63.7$_{\pm3.1}$	35.0$_{\pm6.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	$\underline{0.2}_{\pm0.3}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$
	Score ↑	0.25$_{\pm0.0}$	0.58$_{\pm0.0}$	0.48$_{\pm0.0}$	0.17$_{\pm0.1}$	0.55$_{\pm0.1}$	0.48$_{\pm0.0}$	0.12$_{\pm0.1}$	0.15$_{\pm0.1}$	0.38$_{\pm0.1}$	0.11$_{\pm0.1}$	0.29$_{\pm0.0}$	0.64$_{\pm0.0}$	0.35$_{\pm0.1}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	58.9$_{\pm0.5}$	56.3$_{\pm43.}$	74.6$_{\pm12.}$	72.2$_{\pm3.0}$	79.2$_{\pm3.3}$	79.5$_{\pm2.7}$	$\underline{75.5}_{\pm1.4}$	55.2$_{\pm3.1}$	83.6$_{\pm1.3}$	76.2$_{\pm1.6}$	55.4$_{\pm0.8}$	86.1$_{\pm0.6}$	71.1$_{\pm6.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	20.9$_{\pm1.4}$	40.0$_{\pm34.}$	39.4$_{\pm26.}$	31.8$_{\pm2.8}$	41.2$_{\pm7.4}$	52.5$_{\pm3.3}$	47.8$_{\pm2.8}$	10.9$_{\pm5.5}$	60.8$_{\pm5.8}$	43.2$_{\pm10.}$	21.3$_{\pm9.2}$	67.5$_{\pm10.}$	39.8$_{\pm10.}$
	Score ↑	0.49$_{\pm0.0}$	0.36$_{\pm0.3}$	0.54$_{\pm0.0}$	0.55$_{\pm0.0}$	0.56$_{\pm0.0}$	0.52$_{\pm0.0}$	0.51$_{\pm0.0}$	$\underline{0.50}_{\pm0.0}$	0.52$_{\pm0.0}$	0.53$_{\pm0.0}$	0.46$_{\pm0.0}$	0.52$_{\pm0.0}$	0.50$_{\pm0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	42.8$_{\pm21.8}$	77.0$_{\pm2.2}$	79.8$_{\pm2.5}$	33.3$_{\pm13.4}$	71.5$_{\pm12.}$	70.3$_{\pm18.}$	25.1$_{\pm18.0}$	37.9$_{\pm14.}$	76.3$_{\pm6.2}$	65.0$_{\pm8.4}$	51.5$_{\pm7.2}$	72.7$_{\pm10.3}$	58.6$_{\pm11.}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$\underline{15.3}_{\pm13.}$	27.3$_{\pm15.}$	63.8$_{\pm18.}$	8.6$_{\pm15.}$	20.6$_{\pm35.}$	37.8$_{\pm32.}$	$\underline{2.2}_{\pm3.8}$	3.2$_{\pm5.5}$	37.3$_{\pm17.}$	$\underline{13.9}_{\pm11.}$	19.7$_{\pm1.2}$	55.6$_{\pm28.}$	25.4$_{\pm16.}$
	Score ↑	0.36$_{\pm0.2}$	0.61$_{\pm0.1}$	0.49$_{\pm0.0}$	0.31$_{\pm0.1}$	0.61$_{\pm0.1}$	0.51$_{\pm0.0}$	0.24$_{\pm0.2}$	0.36$_{\pm0.1}$	0.56$_{\pm0.0}$	0.57$_{\pm0.0}$	0.43$_{\pm0.1}$	0.47$_{\pm0.0}$	0.46$_{\pm0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	52.0$_{\pm2.2}$	74.5$_{\pm4.3}$	68.6$_{\pm1.7}$	61.1$_{\pm4.9}$	73.6$_{\pm2.7}$	69.8$_{\pm0.8}$	60.1$_{\pm4.0}$	46.7$_{\pm3.3}$	66.0$_{\pm3.5}$	61.3$_{\pm1.6}$	47.0$_{\pm3.4}$	77.0$_{\pm2.1}$	63.2$_{\pm2.9}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm0.0}$	$\underline{1.0}_{\pm1.7}$	$\underline{0.1}_{\pm0.2}$	$\underline{3.4}_{\pm4.3}$	0.3$_{\pm0.6}$	**0.0**$_{\pm0.0}$	5.9$_{\pm3.3}$	**0.0**$_{\pm0.0}$	$\underline{4.7}_{\pm8.2}$	**0.0**$_{\pm0.0}$	$\underline{1.2}_{\pm2.1}$	$\underline{1.8}_{\pm3.1}$	$\underline{1.5}_{\pm2.0}$
	Score ↑	$\underline{0.52}_{\pm0.0}$	$\underline{0.74}_{\pm0.0}$	$\underline{0.69}_{\pm0.0}$	$\underline{0.59}_{\pm0.1}$	$\underline{0.73}_{\pm0.0}$	0.70$_{\pm0.0}$	$\underline{0.57}_{\pm0.0}$	0.47$_{\pm0.0}$	0.63$_{\pm0.1}$	$\underline{0.61}_{\pm0.0}$	0.46$_{\pm0.0}$	$\underline{0.76}_{\pm0.0}$	$\underline{0.62}_{\pm0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	58.8$_{\pm4.5}$	**83.6**$_{\pm1.7}$	$\underline{81.5}_{\pm1.0}$	76.0$_{\pm0.8}$	82.6$_{\pm2.0}$	82.4$_{\pm0.6}$	75.4$_{\pm1.8}$	**60.5**$_{\pm1.6}$	**84.9**$_{\pm1.6}$	77.3$_{\pm0.8}$	$\underline{61.3}_{\pm0.9}$	**88.2**$_{\pm1.0}$	**76.1**$_{\pm1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	9.9$_{\pm2.3}$	59.3$_{\pm13.9}$	54.1$_{\pm12.1}$	27.3$_{\pm5.9}$	51.5$_{\pm6.2}$	57.3$_{\pm1.7}$	42.2$_{\pm5.9}$	21.1$_{\pm6.5}$	41.9$_{\pm8.3}$	33.8$_{\pm8.5}$	14.7$_{\pm9.8}$	50.1$_{\pm4.8}$	38.6$_{\pm7.3}$
	Score ↑	0.54$_{\pm0.05}$	0.53$_{\pm0.04}$	0.53$_{\pm0.05}$	0.60$_{\pm0.03}$	0.55$_{\pm0.01}$	0.52$_{\pm0.00}$	0.53$_{\pm0.01}$	0.50$_{\pm0.04}$	0.53$_{\pm0.02}$	0.58$_{\pm0.04}$	0.54$_{\pm0.04}$	0.52$_{\pm0.01}$	0.54$_{\pm0.03}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	58.4$_{\pm0.9}$	80.2$_{\pm2.4}$	80.8$_{\pm0.6}$	76.3$_{\pm1.7}$	80.7$_{\pm1.2}$	80.8$_{\pm0.7}$	73.5$_{\pm2.4}$	56.5$_{\pm1.7}$	82.7$_{\pm2.4}$	77.3$_{\pm1.4}$	54.5$_{\pm0.6}$	86.0$_{\pm1.1}$	74.0$_{\pm1.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	33.3$_{\pm0.3}$	77.4$_{\pm2.6}$	82.5$_{\pm0.8}$	64.8$_{\pm2.5}$	70.9$_{\pm2.5}$	73.3$_{\pm1.2}$	56.5$_{\pm0.9}$	31.0$_{\pm0.6}$	79.4$_{\pm0.8}$	60.5$_{\pm0.5}$	38.7$_{\pm2.1}$	83.0$_{\pm0.5}$	62.6$_{\pm1.3}$
	Score ↑	0.44$_{\pm0.0}$	0.45$_{\pm0.0}$	0.44$_{\pm0.0}$	0.46$_{\pm0.0}$	0.47$_{\pm0.0}$	0.47$_{\pm0.0}$	0.47$_{\pm0.0}$	0.43$_{\pm0.0}$	0.46$_{\pm0.0}$	0.48$_{\pm0.0}$	0.39$_{\pm0.0}$	0.47$_{\pm0.0}$	0.45$_{\pm0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	56.0$_{\pm2.1}$	76.8$_{\pm2.3}$	79.2$_{\pm1.1}$	74.4$_{\pm2.0}$	80.9$_{\pm1.2}$	80.0$_{\pm1.4}$	71.6$_{\pm1.0}$	$\underline{58.4}_{\pm1.3}$	81.4$_{\pm2.1}$	76.8$_{\pm0.5}$	**63.0**$_{\pm0.6}$	85.8$_{\pm0.6}$	73.7$_{\pm1.3}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	23.9$_{\pm1.0}$	27.6$_{\pm2.5}$	28.2$_{\pm1.7}$	31.2$_{\pm1.0}$	19.3$_{\pm0.8}$	$\underline{13.0}_{\pm0.8}$	29.6$_{\pm1.9}$	23.9$_{\pm2.3}$	15.0$_{\pm0.4}$	36.1$_{\pm0.9}$	28.8$_{\pm0.7}$	19.4$_{\pm0.3}$	24.7$_{\pm1.2}$
	Score ↑	0.45$_{\pm0.0}$	0.60$_{\pm0.0}$	0.62$_{\pm0.0}$	0.57$_{\pm0.0}$	0.68$_{\pm0.0}$	$\underline{0.71}_{\pm0.0}$	0.55$_{\pm0.0}$	0.47$_{\pm0.0}$	$\underline{0.71}_{\pm0.0}$	0.56$_{\pm0.0}$	$\underline{0.49}_{\pm0.0}$	0.72$_{\pm0.0}$	0.59$_{\pm0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$\underline{59.9}_{\pm1.2}$	$\underline{82.0}_{\pm2.1}$	81.1$_{\pm1.6}$	$\underline{76.6}_{\pm3.0}$	**83.3**$_{\pm1.4}$	$\underline{82.6}_{\pm1.9}$	74.0$_{\pm1.4}$	57.1$_{\pm2.0}$	83.4$_{\pm1.7}$	$\underline{77.8}_{\pm1.1}$	57.0$_{\pm0.8}$	86.7$_{\pm0.7}$	$\underline{75.1}_{\pm1.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$	**0.0**$_{\pm0.0}$
	Score ↑	**0.60**$_{\pm0.0}$	**0.82**$_{\pm0.0}$	**0.81**$_{\pm0.0}$	**0.77**$_{\pm0.0}$	**0.83**$_{\pm0.0}$	**0.83**$_{\pm0.0}$	**0.74**$_{\pm0.0}$	**0.57**$_{\pm0.0}$	**0.83**$_{\pm0.0}$	**0.78**$_{\pm0.0}$	**0.57**$_{\pm0.0}$	**0.87**$_{\pm0.0}$	**0.75**$_{\pm0.0}$

Table A.19. **Results for Multi-Class SCADA Unlearning on DomainNet.** Forget classes are $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$71.3_{\pm 0.4}$	$66.6_{\pm 0.5}$	$61.7_{\pm 1.0}$	$78.3_{\pm 0.2}$	$55.9_{\pm 1.9}$	$65.1_{\pm 1.1}$	$75.0_{\pm 0.5}$	$67.7_{\pm 0.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$67.6_{\pm 7.0}$	$55.9_{\pm 5.9}$	$42.9_{\pm 2.7}$	$60.6_{\pm 2.7}$	$6.5_{\pm 1.3}$	$14.6_{\pm 1.1}$	$22.8_{\pm 3.8}$	$38.7_{\pm 3.5}$
	Score ↑	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.49_{\pm 0.0}$	$0.53_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.61_{\pm 0.0}$	$0.50_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$71.0_{\pm 0.4}$	$65.3_{\pm 2.8}$	$58.6_{\pm 2.9}$	$78.2_{\pm 1.1}$	$54.2_{\pm 1.3}$	$62.9_{\pm 1.9}$	$74.1_{\pm 1.3}$	$66.3_{\pm 1.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$
	Score ↑	$0.71_{\pm 0.0}$	$0.65_{\pm 0.0}$	$0.59_{\pm 0.0}$	$0.78_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.63_{\pm 0.0}$	$0.74_{\pm 0.0}$	$0.66_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$68.9_{\pm 0.1}$	$\underline{64.9}_{\pm 1.3}$	$62.2_{\pm 1.4}$	$77.3_{\pm 0.2}$	$52.1_{\pm 2.3}$	$\underline{65.5}_{\pm 1.6}$	$74.2_{\pm 1.0}$	$\underline{66.5}_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$37.4_{\pm 7.1}$	$32.3_{\pm 3.2}$	$20.8_{\pm 7.7}$	$36.6_{\pm 3.8}$	$\underline{0.6}_{\pm 0.4}$	$5.3_{\pm 2.6}$	$10.1_{\pm 10.}$	$20.4_{\pm 5.0}$
	Score ↑	$\underline{0.50}_{\pm 0.0}$	$\underline{0.49}_{\pm 0.0}$	$\underline{0.52}_{\pm 0.0}$	$0.57_{\pm 0.0}$	$\underline{0.52}_{\pm 0.0}$	$\underline{0.62}_{\pm 0.0}$	$\underline{0.68}_{\pm 0.1}$	$\underline{0.56}_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$11.6_{\pm 3.6}$	$2.7_{\pm 1.0}$	$22.4_{\pm 3.9}$	$37.5_{\pm 4.9}$	$5.2_{\pm 4.5}$	$7.9_{\pm 5.6}$	$14.8_{\pm 7.6}$	$14.6_{\pm 4.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{3.0}_{\pm 5.1}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{0.4}_{\pm 0.7}$
	Score ↑	$0.12_{\pm 0.0}$	$0.03_{\pm 0.0}$	$0.22_{\pm 0.0}$	$0.37_{\pm 0.1}$	$0.05_{\pm 0.0}$	$0.08_{\pm 0.1}$	$0.15_{\pm 0.1}$	$0.14_{\pm 0.0}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$\underline{70.7}_{\pm 1.1}$	$63.1_{\pm 0.3}$	$61.1_{\pm 1.8}$	$\mathbf{78.6}_{\pm 1.3}$	$50.8_{\pm 0.7}$	$61.6_{\pm 1.5}$	$69.2_{\pm 4.2}$	$65.0_{\pm 1.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$61.4_{\pm 7.5}$	$47.1_{\pm 6.3}$	$31.5_{\pm 11.9}$	$50.5_{\pm 15.1}$	$19.5_{\pm 3.6}$	$13.4_{\pm 1.1}$	$18.2_{\pm 1.1}$	$34.5_{\pm 6.6}$
	Score ↑	$0.44_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.52_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.59_{\pm 0.0}$	$0.49_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$5.4_{\pm 4.2}$	$33.9_{\pm 27.}$	$39.1_{\pm 14.}$	$64.5_{\pm 3.4}$	$42.0_{\pm 6.3}$	$32.4_{\pm 11.}$	$55.7_{\pm 10.}$	$39.0_{\pm 11.}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$2.9_{\pm 5.1}$	$5.2_{\pm 8.9}$	$\underline{11.6}_{\pm 10.}$	$27.0_{\pm 18.}$	$1.6_{\pm 1.7}$	$\underline{0.3}_{\pm 0.5}$	$10.7_{\pm 14.}$	$8.5_{\pm 8.6}$
	Score ↑	$0.05_{\pm 0.0}$	$0.31_{\pm 0.2}$	$0.34_{\pm 0.1}$	$0.51_{\pm 0.1}$	$0.41_{\pm 0.1}$	$0.32_{\pm 0.1}$	$0.50_{\pm 0.0}$	$0.35_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$44.7_{\pm 3.2}$	$38.1_{\pm 5.4}$	$43.4_{\pm 3.3}$	$63.0_{\pm 3.3}$	$23.8_{\pm 18.}$	$47.3_{\pm 0.4}$	$50.6_{\pm 3.8}$	$44.4_{\pm 5.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\underline{0.8}_{\pm 1.4}$	$\underline{0.2}_{\pm 0.3}$	$\mathbf{0.0}_{\pm 0.0}$	$7.9_{\pm 11.}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{0.1}_{\pm 0.2}$	$1.3_{\pm 1.9}$
	Score ↑	$0.44_{\pm 0.0}$	$0.38_{\pm 0.1}$	$0.43_{\pm 0.0}$	$\underline{0.59}_{\pm 0.1}$	$0.24_{\pm 0.2}$	$0.47_{\pm 0.0}$	$0.51_{\pm 0.0}$	$0.44_{\pm 0.1}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$69.2_{\pm 0.4}$	$62.1_{\pm 3.6}$	$\mathbf{63.4}_{\pm 0.3}$	$76.2_{\pm 0.8}$	$\mathbf{56.1}_{\pm 0.2}$	$\mathbf{67.1}_{\pm 0.2}$	$74.3_{\pm 0.4}$	$66.9_{\pm 0.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$45.3_{\pm 8.3}$	$41.7_{\pm 1.1}$	$31.4_{\pm 7.8}$	$38.3_{\pm 2.4}$	$10.3_{\pm 12.9}$	$7.6_{\pm 0.0}$	$11.6_{\pm 0.9}$	$26.6_{\pm 4.8}$
	Score ↑	$0.48_{\pm 0.02}$	$0.44_{\pm 0.03}$	$0.48_{\pm 0.03}$	$0.55_{\pm 0.00}$	$0.51_{\pm 0.06}$	$0.62_{\pm 0.00}$	$0.67_{\pm 0.01}$	$0.54_{\pm 0.02}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$61.1_{\pm 0.3}$	$59.7_{\pm 0.7}$	$55.1_{\pm 0.3}$	$76.1_{\pm 0.5}$	$49.0_{\pm 0.3}$	$59.9_{\pm 0.4}$	$72.2_{\pm 0.3}$	$61.9_{\pm 0.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$45.5_{\pm 1.3}$	$70.1_{\pm 0.4}$	$57.1_{\pm 1.1}$	$71.6_{\pm 1.1}$	$22.3_{\pm 0.8}$	$26.8_{\pm 0.9}$	$32.7_{\pm 2.4}$	$46.6_{\pm 1.1}$
	Score ↑	$0.42_{\pm 0.0}$	$0.35_{\pm 0.0}$	$0.35_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.43_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$\mathbf{71.3}_{\pm 0.2}$	$\mathbf{66.4}_{\pm 1.6}$	$\underline{62.4}_{\pm 0.6}$	$\underline{78.5}_{\pm 0.4}$	$\underline{55.7}_{\pm 2.4}$	$65.1_{\pm 0.4}$	$\mathbf{74.8}_{\pm 0.9}$	$\mathbf{67.8}_{\pm 0.9}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$67.7_{\pm 4.5}$	$56.3_{\pm 5.1}$	$48.7_{\pm 3.0}$	$62.4_{\pm 3.3}$	$6.9_{\pm 2.1}$	$13.6_{\pm 0.5}$	$21.2_{\pm 0.5}$	$39.6_{\pm 2.7}$
	Score ↑	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.48_{\pm 0.0}$	$\underline{0.52}_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.62_{\pm 0.0}$	$0.50_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$65.3_{\pm 1.3}$	$63.7_{\pm 1.7}$	$60.1_{\pm 1.0}$	$77.0_{\pm 0.5}$	$55.0_{\pm 1.1}$	$63.5_{\pm 1.5}$	$\underline{74.7}_{\pm 0.4}$	$65.6_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$
	Score ↑	$\mathbf{0.65}_{\pm 0.0}$	$\mathbf{0.64}_{\pm 0.0}$	$\mathbf{0.60}_{\pm 0.0}$	$\mathbf{0.77}_{\pm 0.0}$	$\mathbf{0.55}_{\pm 0.0}$	$\mathbf{0.64}_{\pm 0.0}$	$\mathbf{0.75}_{\pm 0.0}$	$\mathbf{0.66}_{\pm 0.0}$

Table A.20. **Results for Multi-Class SCADA Unlearning on Office 31.** Forget classes are $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	A → D	A → W	D → A	D → W	W → A	W → D	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.1$_{\pm 1.0}$	86.3$_{\pm 1.9}$	67.0$_{\pm 1.2}$	79.9$_{\pm 2.2}$	72.0$_{\pm 0.6}$	78.5$_{\pm 1.8}$	76.8$_{\pm 1.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	84.8$_{\pm 2.8}$	99.5$_{\pm 0.9}$	93.6$_{\pm 1.0}$	88.5$_{\pm 2.8}$	97.7$_{\pm 2.0}$	76.6$_{\pm 2.3}$	90.1$_{\pm 2.0}$
	Score ↑	0.42$_{\pm 0.0}$	0.43$_{\pm 0.0}$	0.35$_{\pm 0.0}$	0.42$_{\pm 0.0}$	0.36$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.40$_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	76.0$_{\pm 5.0}$	87.0$_{\pm 0.7}$	68.2$_{\pm 1.2}$	83.3$_{\pm 1.0}$	72.0$_{\pm 1.3}$	78.0$_{\pm 1.5}$	77.4$_{\pm 1.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$
	Score ↑	0.76$_{\pm 0.1}$	0.87$_{\pm 0.0}$	0.68$_{\pm 0.0}$	0.83$_{\pm 0.0}$	0.72$_{\pm 0.0}$	0.78$_{\pm 0.0}$	0.77$_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	<u>77.2</u>$_{\pm 0.7}$	<u>86.5</u>$_{\pm 2.0}$	66.7$_{\pm 2.7}$	79.7$_{\pm 2.8}$	<u>72.0</u>$_{\pm 0.6}$	<u>78.4</u>$_{\pm 1.2}$	76.7$_{\pm 1.7}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	79.4$_{\pm 1.3}$	94.0$_{\pm 1.0}$	85.4$_{\pm 3.7}$	80.3$_{\pm 1.7}$	93.0$_{\pm 0.0}$	65.0$_{\pm 7.1}$	82.8$_{\pm 2.5}$
	Score ↑	0.43$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.36$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.37$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.42$_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	40.6$_{\pm 15.}$	73.1$_{\pm 5.2}$	61.7$_{\pm 4.3}$	63.5$_{\pm 4.5}$	71.6$_{\pm 1.1}$	47.8$_{\pm 6.1}$	59.7$_{\pm 6.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	<u>28.7</u>$_{\pm 20.}$	**0.0**$_{\pm 0.0}$	69.0$_{\pm 4.0}$	**0.0**$_{\pm 0.0}$	<u>16.3</u>$_{\pm 4.0}$
	Score ↑	0.41$_{\pm 0.2}$	<u>0.73</u>$_{\pm 0.1}$	<u>0.49</u>$_{\pm 0.1}$	0.63$_{\pm 0.0}$	0.42$_{\pm 0.0}$	0.48$_{\pm 0.1}$	0.53$_{\pm 0.1}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**78.2**$_{\pm 2.1}$	85.8$_{\pm 1.5}$	<u>67.0</u>$_{\pm 1.2}$	<u>80.1</u>$_{\pm 1.2}$	<u>72.0</u>$_{\pm 0.6}$	**78.6**$_{\pm 0.2}$	<u>77.0</u>$_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	74.9$_{\pm 11.}$	95.6$_{\pm 0.9}$	86.5$_{\pm 2.0}$	82.5$_{\pm 2.5}$	93.0$_{\pm 0.0}$	72.4$_{\pm 3.2}$	84.2$_{\pm 3.3}$
	Score ↑	0.45$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.36$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.37$_{\pm 0.0}$	0.46$_{\pm 0.0}$	0.42$_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	62.6$_{\pm 14.}$	63.7$_{\pm 15.}$	60.2$_{\pm 8.2}$	76.7$_{\pm 2.8}$	71.6$_{\pm 1.1}$	57.0$_{\pm 15.}$	65.3$_{\pm 9.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	<u>6.1</u>$_{\pm 5.5}$	<u>6.6</u>$_{\pm 11.}$	29.2$_{\pm 25.}$	60.1$_{\pm 25.}$	<u>60.8</u>$_{\pm 25.}$	<u>7.8</u>$_{\pm 6.4}$	28.4$_{\pm 16.}$
	Score ↑	0.59$_{\pm 0.1}$	0.59$_{\pm 0.1}$	0.47$_{\pm 0.1}$	0.49$_{\pm 0.1}$	<u>0.45</u>$_{\pm 0.1}$	0.53$_{\pm 0.1}$	0.52$_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.2$_{\pm 2.6}$	85.4$_{\pm 2.9}$	64.8$_{\pm 2.0}$	76.2$_{\pm 0.4}$	**72.0**$_{\pm 0.6}$	72.0$_{\pm 1.9}$	73.4$_{\pm 1.7}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	6.5$_{\pm 5.3}$	24.0$_{\pm 7.8}$	33.9$_{\pm 5.7}$	<u>7.1</u>$_{\pm 6.9}$	73.7$_{\pm 7.0}$	22.6$_{\pm 9.5}$	28.0$_{\pm 7.0}$
	Score ↑	<u>0.66</u>$_{\pm 0.1}$	0.69$_{\pm 0.0}$	0.48$_{\pm 0.0}$	<u>0.71</u>$_{\pm 0.0}$	0.41$_{\pm 0.0}$	<u>0.59</u>$_{\pm 0.1}$	<u>0.59</u>$_{\pm 0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.0$_{\pm 0.8}$	80.8$_{\pm 2.5}$	**80.6**$_{\pm 1.7}$	**86.0**$_{\pm 1.0}$	**78.9**$_{\pm 0.6}$	71.1$_{\pm 0.8}$	**77.9**$_{\pm 1.2}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	84.8$_{\pm 1.2}$	84.0$_{\pm 1.2}$	81.2$_{\pm 5.1}$	97.1$_{\pm 1.1}$	69.5$_{\pm 7.8}$	95.3$_{\pm 1.2}$	85.3$_{\pm 2.9}$
	Score ↑	0.38$_{\pm 0.01}$	0.44$_{\pm 0.02}$	0.45$_{\pm 0.00}$	0.44$_{\pm 0.01}$	0.47$_{\pm 0.02}$	0.37$_{\pm 0.00}$	0.42$_{\pm 0.01}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	76.8$_{\pm 0.4}$	<u>86.5</u>$_{\pm 1.5}$	66.3$_{\pm 2.4}$	78.5$_{\pm 2.9}$	<u>72.0</u>$_{\pm 0.6}$	77.4$_{\pm 2.2}$	76.2$_{\pm 1.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	76.8$_{\pm 3.1}$	100.$_{\pm 0.0}$	90.7$_{\pm 2.0}$	90.7$_{\pm 1.9}$	98.8$_{\pm 1.0}$	77.4$_{\pm 1.7}$	89.1$_{\pm 1.6}$
	Score ↑	0.43$_{\pm 0.0}$	0.43$_{\pm 0.0}$	0.35$_{\pm 0.0}$	0.41$_{\pm 0.0}$	0.36$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.40$_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	<u>77.2</u>$_{\pm 1.0}$	**87.0**$_{\pm 1.2}$	65.9$_{\pm 2.3}$	78.1$_{\pm 1.2}$	<u>72.0</u>$_{\pm 0.6}$	75.6$_{\pm 1.2}$	76.0$_{\pm 1.2}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	47.9$_{\pm 1.5}$	92.9$_{\pm 0.9}$	69.6$_{\pm 2.7}$	72.7$_{\pm 1.9}$	84.8$_{\pm 2.0}$	40.4$_{\pm 1.5}$	68.0$_{\pm 1.7}$
	Score ↑	0.52$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.39$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.39$_{\pm 0.0}$	0.54$_{\pm 0.0}$	0.46$_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.1$_{\pm 1.0}$	86.3$_{\pm 1.9}$	**67.0**$_{\pm 1.2}$	79.9$_{\pm 2.4}$	71.6$_{\pm 1.1}$	78.0$_{\pm 1.2}$	76.7$_{\pm 1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$
	Score ↑	**0.77**$_{\pm 0.0}$	**0.86**$_{\pm 0.0}$	**0.67**$_{\pm 0.0}$	**0.80**$_{\pm 0.0}$	**0.72**$_{\pm 0.0}$	**0.78**$_{\pm 0.0}$	**0.77**$_{\pm 0.0}$

Table A.21. **Results for Single-Class SCADA Unlearning on OfficeHome.** Forget class is $c_{\mathcal{F}} = 1$ *(Best result in bold, second-best underlined)*

Method	Metric	A → C	A → P	A → R	C → A	C → P	C → R	P → A	P → C	P → R	R → A	R → C	R → P	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$60.9_{\pm 0.5}$	$82.2_{\pm 1.0}$	$85.5_{\pm 1.8}$	$74.4_{\pm 2.4}$	$83.5_{\pm 1.0}$	$85.6_{\pm 1.2}$	$72.1_{\pm 3.3}$	$59.5_{\pm 0.8}$	$86.0_{\pm 0.7}$	$75.8_{\pm 1.4}$	$60.0_{\pm 2.0}$	$88.0_{\pm 1.4}$	$76.1_{\pm 1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$51.2_{\pm 9.8}$	$88.2_{\pm 1.5}$	$89.2_{\pm 1.5}$	$42.3_{\pm 7.5}$	$72.7_{\pm 2.7}$	$85.5_{\pm 1.2}$	$54.5_{\pm 8.6}$	$50.0_{\pm 6.4}$	$88.6_{\pm 2.9}$	$39.9_{\pm 2.8}$	$80.4_{\pm 3.6}$	$93.6_{\pm 1.5}$	$69.7_{\pm 4.2}$
	Score $\uparrow$	$0.40_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.52_{\pm 0.0}$	$0.48_{\pm 0.0}$	$0.46_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.46_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.33_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.45_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$62.1_{\pm 2.4}$	$83.4_{\pm 1.1}$	$86.0_{\pm 1.9}$	$74.7_{\pm 3.2}$	$82.6_{\pm 0.6}$	$86.2_{\pm 1.0}$	$72.4_{\pm 1.9}$	$60.0_{\pm 3.1}$	$86.5_{\pm 0.9}$	$75.0_{\pm 2.9}$	$61.1_{\pm 1.6}$	$88.3_{\pm 1.0}$	$76.5_{\pm 1.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$
	Score $\uparrow$	$0.62_{\pm 0.0}$	$0.83_{\pm 0.0}$	$0.86_{\pm 0.0}$	$0.75_{\pm 0.0}$	$0.83_{\pm 0.0}$	$0.86_{\pm 0.0}$	$0.72_{\pm 0.0}$	$0.60_{\pm 0.0}$	$0.87_{\pm 0.0}$	$0.75_{\pm 0.0}$	$0.61_{\pm 0.0}$	$0.88_{\pm 0.0}$	$0.77_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$\mathbf{60.7}_{\pm 0.5}$	$\mathbf{81.8}_{\pm 0.3}$	$\mathbf{85.3}_{\pm 1.8}$	$\mathbf{74.0}_{\pm 1.1}$	$82.9_{\pm 0.5}$	$\mathbf{85.8}_{\pm 0.7}$	$71.2_{\pm 1.2}$	$57.2_{\pm 0.8}$	$\mathbf{86.5}_{\pm 0.3}$	$\mathbf{75.2}_{\pm 0.9}$	$58.6_{\pm 1.6}$	$\underline{87.5}_{\pm 0.8}$	$\mathbf{75.5}_{\pm 0.9}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$42.9_{\pm 1.8}$	$74.4_{\pm 8.5}$	$82.8_{\pm 2.0}$	$35.8_{\pm 3.7}$	$69.4_{\pm 6.6}$	$82.5_{\pm 3.1}$	$52.8_{\pm 16.}$	$\mathbf{0.0}_{\pm 0.0}$	$81.5_{\pm 3.8}$	$50.4_{\pm 5.7}$	$74.4_{\pm 17.}$	$90.2_{\pm 6.5}$	$61.4_{\pm 6.3}$
	Score $\uparrow$	$0.42_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.49_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.47_{\pm 0.1}$	$\underline{0.57}_{\pm 0.0}$	$0.48_{\pm 0.0}$	$0.50_{\pm 0.0}$	$0.34_{\pm 0.0}$	$0.46_{\pm 0.0}$	$0.47_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$28.3_{\pm 1.0}$	$56.6_{\pm 1.2}$	$55.1_{\pm 5.0}$	$11.8_{\pm 6.4}$	$56.6_{\pm 2.4}$	$44.4_{\pm 4.3}$	$9.3_{\pm 5.6}$	$17.2_{\pm 6.1}$	$49.6_{\pm 6.8}$	$18.0_{\pm 13.}$	$24.4_{\pm 2.1}$	$61.7_{\pm 1.2}$	$36.1_{\pm 4.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$
	Score $\uparrow$	$0.28_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.55_{\pm 0.1}$	$0.12_{\pm 0.1}$	$0.57_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.09_{\pm 0.1}$	$0.17_{\pm 0.1}$	$0.50_{\pm 0.1}$	$0.18_{\pm 0.1}$	$0.24_{\pm 0.0}$	$0.62_{\pm 0.0}$	$0.36_{\pm 0.0}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$57.7_{\pm 1.3}$	$79.7_{\pm 1.4}$	$\underline{84.7}_{\pm 2.3}$	$68.6_{\pm 2.0}$	$81.5_{\pm 0.5}$	$83.1_{\pm 1.2}$	$\mathbf{71.6}_{\pm 2.1}$	$56.8_{\pm 3.2}$	$85.0_{\pm 3.0}$	$74.1_{\pm 1.5}$	$58.4_{\pm 1.6}$	$86.4_{\pm 0.2}$	$74.0_{\pm 1.7}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$25.0_{\pm 4.8}$	$78.1_{\pm 6.5}$	$87.9_{\pm 3.5}$	$18.7_{\pm 13.}$	$52.5_{\pm 5.3}$	$68.0_{\pm 7.2}$	$34.1_{\pm 8.8}$	$13.7_{\pm 12.}$	$76.4_{\pm 18.}$	$30.1_{\pm 7.4}$	$29.8_{\pm 10.}$	$79.8_{\pm 2.6}$	$49.5_{\pm 8.4}$
	Score $\uparrow$	$0.46_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.58_{\pm 0.1}$	$0.53_{\pm 0.0}$	$0.50_{\pm 0.0}$	$0.53_{\pm 0.0}$	$0.51_{\pm 0.1}$	$0.48_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.48_{\pm 0.0}$	$0.50_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$52.9_{\pm 4.9}$	$80.1_{\pm 1.6}$	$70.2_{\pm 18.}$	$65.3_{\pm 5.5}$	$72.4_{\pm 5.2}$	$81.4_{\pm 4.4}$	$64.6_{\pm 5.7}$	$44.5_{\pm 10.}$	$79.3_{\pm 6.5}$	$68.3_{\pm 7.2}$	$58.8_{\pm 1.7}$	$76.3_{\pm 9.1}$	$67.8_{\pm 6.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$14.9_{\pm 9.0}$	$51.9_{\pm 15.}$	$\underline{2.4}_{\pm 3.3}$	$\underline{5.7}_{\pm 9.9}$	$\underline{6.4}_{\pm 11.1}$	$49.2_{\pm 43.}$	$\underline{16.3}_{\pm 7.4}$	$\underline{2.4}_{\pm 4.1}$	$40.4_{\pm 24.}$	$16.3_{\pm 3.7}$	$43.4_{\pm 29.}$	$55.2_{\pm 47.}$	$25.4_{\pm 17.}$
	Score $\uparrow$	$0.46_{\pm 0.0}$	$0.53_{\pm 0.1}$	$\underline{0.68}_{\pm 0.2}$	$\underline{0.62}_{\pm 0.0}$	$0.68_{\pm 0.1}$	$0.58_{\pm 0.2}$	$\underline{0.56}_{\pm 0.1}$	$0.44_{\pm 0.1}$	$0.57_{\pm 0.1}$	$\underline{0.59}_{\pm 0.1}$	$0.42_{\pm 0.1}$	$0.52_{\pm 0.2}$	$0.55_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$50.0_{\pm 2.6}$	$71.7_{\pm 3.9}$	$67.5_{\pm 2.5}$	$53.7_{\pm 3.9}$	$74.9_{\pm 1.4}$	$71.0_{\pm 3.1}$	$55.5_{\pm 4.3}$	$45.4_{\pm 4.0}$	$65.3_{\pm 4.2}$	$56.0_{\pm 0.4}$	$44.9_{\pm 4.3}$	$74.1_{\pm 2.0}$	$60.8_{\pm 3.0}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\underline{3.6}_{\pm 6.2}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{1.6}_{\pm 2.8}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{0.4}_{\pm 0.8}$
	Score $\uparrow$	$\underline{0.48}_{\pm 0.0}$	$\underline{0.72}_{\pm 0.0}$	$0.67_{\pm 0.0}$	$0.54_{\pm 0.0}$	$\underline{0.75}_{\pm 0.0}$	$0.71_{\pm 0.0}$	$\underline{0.56}_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.65_{\pm 0.0}$	$0.55_{\pm 0.0}$	$0.45_{\pm 0.0}$	$\underline{0.74}_{\pm 0.0}$	$\underline{0.61}_{\pm 0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$54.5_{\pm 2.7}$	$76.8_{\pm 0.3}$	$78.5_{\pm 1.1}$	$\underline{73.3}_{\pm 1.6}$	$\underline{83.0}_{\pm 0.3}$	$83.6_{\pm 0.5}$	$\underline{71.4}_{\pm 3.1}$	$\mathbf{59.2}_{\pm 0.6}$	$85.2_{\pm 0.3}$	$73.6_{\pm 0.9}$	$\underline{61.9}_{\pm 1.8}$	$\mathbf{88.6}_{\pm 0.5}$	$74.1_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$23.2_{\pm 0.0}$	$78.3_{\pm 2.1}$	$75.8_{\pm 11.4}$	$34.2_{\pm 6.9}$	$67.2_{\pm 2.1}$	$83.4_{\pm 5.0}$	$34.1_{\pm 0.0}$	$47.3_{\pm 3.8}$	$81.8_{\pm 2.8}$	$31.7_{\pm 3.4}$	$84.8_{\pm 3.8}$	$87.9_{\pm 4.3}$	$60.8_{\pm 3.8}$
	Score $\uparrow$	$0.44_{\pm 0.02}$	$0.43_{\pm 0.01}$	$0.45_{\pm 0.02}$	$0.55_{\pm 0.02}$	$0.50_{\pm 0.01}$	$0.46_{\pm 0.01}$	$0.53_{\pm 0.02}$	$0.40_{\pm 0.02}$	$0.47_{\pm 0.01}$	$0.56_{\pm 0.02}$	$0.33_{\pm 0.00}$	$0.47_{\pm 0.01}$	$0.47_{\pm 0.01}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$58.9_{\pm 1.3}$	$80.4_{\pm 1.0}$	$84.0_{\pm 1.7}$	$72.9_{\pm 2.7}$	$81.4_{\pm 1.0}$	$83.6_{\pm 1.1}$	$70.6_{\pm 2.3}$	$56.7_{\pm 0.9}$	$\underline{85.8}_{\pm 1.7}$	$74.4_{\pm 1.6}$	$55.1_{\pm 1.1}$	$86.8_{\pm 1.1}$	$74.2_{\pm 1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$44.6_{\pm 1.8}$	$90.9_{\pm 1.0}$	$90.9_{\pm 1.0}$	$51.2_{\pm 2.5}$	$84.5_{\pm 2.1}$	$92.2_{\pm 1.5}$	$56.1_{\pm 4.9}$	$42.9_{\pm 1.8}$	$96.0_{\pm 1.1}$	$44.7_{\pm 2.8}$	$73.8_{\pm 6.8}$	$91.2_{\pm 1.5}$	$71.6_{\pm 2.4}$
	Score $\uparrow$	$0.41_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.48_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.51_{\pm 0.0}$	$0.32_{\pm 0.0}$	$0.45_{\pm 0.0}$	$0.43_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$56.9_{\pm 1.4}$	$78.7_{\pm 0.2}$	$82.4_{\pm 1.2}$	$71.8_{\pm 2.9}$	$81.9_{\pm 1.1}$	$82.9_{\pm 1.2}$	$68.7_{\pm 2.8}$	$\underline{59.1}_{\pm 0.6}$	$84.1_{\pm 0.8}$	$74.5_{\pm 2.8}$	$\mathbf{62.9}_{\pm 1.7}$	$86.8_{\pm 0.6}$	$74.2_{\pm 1.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$22.0_{\pm 4.5}$	$\underline{26.6}_{\pm 1.5}$	$25.9_{\pm 2.1}$	$34.9_{\pm 1.4}$	$20.5_{\pm 3.5}$	$\underline{13.4}_{\pm 1.2}$	$35.0_{\pm 2.8}$	$29.8_{\pm 2.7}$	$\underline{14.1}_{\pm 0.0}$	$34.1_{\pm 2.5}$	$\underline{26.8}_{\pm 1.8}$	$\underline{19.2}_{\pm 1.0}$	$25.2_{\pm 2.1}$
	Score $\uparrow$	$0.47_{\pm 0.0}$	$0.62_{\pm 0.0}$	$0.65_{\pm 0.0}$	$0.53_{\pm 0.0}$	$0.68_{\pm 0.0}$	$\underline{0.73}_{\pm 0.0}$	$0.51_{\pm 0.0}$	$0.46_{\pm 0.0}$	$\underline{0.74}_{\pm 0.0}$	$0.56_{\pm 0.0}$	$\underline{0.50}_{\pm 0.0}$	$0.73_{\pm 0.0}$	$0.60_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}} \uparrow$	$\underline{59.9}_{\pm 1.0}$	$\underline{81.5}_{\pm 1.2}$	$\mathbf{85.3}_{\pm 1.8}$	$\mathbf{74.0}_{\pm 2.1}$	$\mathbf{83.5}_{\pm 0.3}$	$\underline{85.3}_{\pm 1.1}$	$70.5_{\pm 3.3}$	$58.5_{\pm 2.2}$	$85.2_{\pm 0.9}$	$\underline{74.7}_{\pm 1.9}$	$59.4_{\pm 1.3}$	$87.4_{\pm 1.0}$	$\underline{75.4}_{\pm 1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}} \downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$
	Score $\uparrow$	$\mathbf{0.60}_{\pm 0.0}$	$\mathbf{0.81}_{\pm 0.0}$	$\mathbf{0.85}_{\pm 0.0}$	$\mathbf{0.74}_{\pm 0.0}$	$\mathbf{0.84}_{\pm 0.0}$	$\mathbf{0.85}_{\pm 0.0}$	$\mathbf{0.71}_{\pm 0.0}$	$\mathbf{0.58}_{\pm 0.0}$	$\mathbf{0.85}_{\pm 0.0}$	$\mathbf{0.75}_{\pm 0.0}$	$\mathbf{0.59}_{\pm 0.0}$	$\mathbf{0.87}_{\pm 0.0}$	$\mathbf{0.75}_{\pm 0.0}$

Table A.22. **Results for Single-Class SCADA Unlearning on DomainNet.** Forget class is $c_{\mathcal{F}} = 1$ *(Best result in bold, second-best underlined)*

Method	Metric	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$71.1_{\pm0.6}$	$65.5_{\pm0.4}$	$62.2_{\pm0.2}$	$78.2_{\pm0.3}$	$55.9_{\pm2.6}$	$65.1_{\pm0.8}$	$75.1_{\pm1.2}$	$67.6_{\pm0.9}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$77.6_{\pm5.7}$	$36.9_{\pm6.9}$	$35.6_{\pm1.3}$	$74.1_{\pm12.1}$	$3.2_{\pm0.4}$	$17.5_{\pm1.0}$	$6.1_{\pm0.3}$	$35.9_{\pm4.0}$
	Score ↑	$0.40_{\pm0.0}$	$0.48_{\pm0.0}$	$0.46_{\pm0.0}$	$0.45_{\pm0.0}$	$0.54_{\pm0.0}$	$0.55_{\pm0.0}$	$0.71_{\pm0.0}$	$0.51_{\pm0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$71.7_{\pm1.0}$	$66.6_{\pm1.1}$	$61.3_{\pm1.9}$	$78.3_{\pm0.3}$	$56.0_{\pm3.2}$	$65.4_{\pm1.1}$	$74.0_{\pm1.3}$	$67.6_{\pm1.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$
	Score ↑	$0.72_{\pm0.0}$	$0.67_{\pm0.0}$	$0.61_{\pm0.0}$	$0.78_{\pm0.0}$	$0.56_{\pm0.0}$	$0.65_{\pm0.0}$	$0.74_{\pm0.0}$	$0.68_{\pm0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$68.0_{\pm0.3}$	$\underline{63.3}_{\pm1.6}$	$\mathbf{62.8}_{\pm0.5}$	$77.1_{\pm0.3}$	$54.2_{\pm0.3}$	$\underline{65.8}_{\pm1.7}$	$73.5_{\pm0.5}$	$66.4_{\pm0.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$59.7_{\pm10.1}$	$\underline{0.5}_{\pm0.4}$	$13.2_{\pm4.8}$	$18.4_{\pm11.1}$	$\mathbf{0.0}_{\pm0.0}$	$5.7_{\pm4.0}$	$\underline{0.6}_{\pm0.7}$	$14.0_{\pm4.4}$
	Score ↑	$0.43_{\pm0.0}$	$\mathbf{0.63}_{\pm0.0}$	$\underline{0.56}_{\pm0.0}$	$\underline{0.65}_{\pm0.1}$	$\underline{0.54}_{\pm0.0}$	$\mathbf{0.62}_{\pm0.0}$	$\underline{0.73}_{\pm0.0}$	$\underline{0.59}_{\pm0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$8.9_{\pm7.3}$	$4.4_{\pm5.0}$	$19.7_{\pm10.7}$	$30.9_{\pm18.1}$	$2.0_{\pm1.4}$	$17.3_{\pm2.1}$	$18.3_{\pm8.7}$	$14.5_{\pm7.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
	Score ↑	$0.09_{\pm0.1}$	$0.04_{\pm0.0}$	$0.20_{\pm0.1}$	$0.31_{\pm0.2}$	$0.02_{\pm0.0}$	$0.17_{\pm0.0}$	$0.18_{\pm0.1}$	$0.15_{\pm0.1}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$\underline{69.9}_{\pm0.4}$	$56.8_{\pm2.0}$	$56.3_{\pm9.1}$	$\underline{78.0}_{\pm0.8}$	$51.4_{\pm1.6}$	$58.3_{\pm3.3}$	$67.2_{\pm9.2}$	$62.5_{\pm3.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$24.3_{\pm27.1}$	$2.2_{\pm0.4}$	$21.0_{\pm6.6}$	$29.4_{\pm19.7}$	$\underline{1.6}_{\pm2.2}$	$7.5_{\pm1.3}$	$1.5_{\pm1.6}$	$12.5_{\pm8.4}$
	Score ↑	$\underline{0.58}_{\pm0.1}$	$\underline{0.56}_{\pm0.0}$	$0.46_{\pm0.1}$	$0.61_{\pm0.1}$	$0.51_{\pm0.0}$	$0.54_{\pm0.0}$	$0.66_{\pm0.1}$	$0.56_{\pm0.1}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$40.6_{\pm34.4}$	$23.6_{\pm24.1}$	$53.1_{\pm4.5}$	$61.2_{\pm21.9}$	$38.2_{\pm5.9}$	$49.6_{\pm15.6}$	$53.0_{\pm14.9}$	$45.6_{\pm17.3}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$35.4_{\pm36.3}$	$\mathbf{0.0}_{\pm0.0}$	$14.1_{\pm9.7}$	$35.4_{\pm34.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{2.0}_{\pm3.5}$	$\mathbf{0.0}_{\pm0.0}$	$12.4_{\pm11.9}$
	Score ↑	$0.27_{\pm0.2}$	$0.24_{\pm0.2}$	$0.47_{\pm0.1}$	$0.44_{\pm0.1}$	$0.38_{\pm0.1}$	$0.48_{\pm0.1}$	$0.53_{\pm0.1}$	$0.40_{\pm0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$38.3_{\pm0.7}$	$37.6_{\pm1.8}$	$43.4_{\pm3.3}$	$62.4_{\pm1.0}$	$32.7_{\pm2.3}$	$40.5_{\pm6.6}$	$46.9_{\pm3.7}$	$43.1_{\pm2.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$18.4_{\pm10.1}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{0.3}_{\pm0.5}$	$\underline{13.9}_{\pm23.9}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{4.7}_{\pm4.9}$
	Score ↑	$0.33_{\pm0.0}$	$0.38_{\pm0.0}$	$0.43_{\pm0.0}$	$0.56_{\pm0.1}$	$0.33_{\pm0.0}$	$0.40_{\pm0.1}$	$0.47_{\pm0.0}$	$0.41_{\pm0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$69.0_{\pm0.9}$	$62.9_{\pm0.3}$	$\underline{62.2}_{\pm0.9}$	$77.1_{\pm0.5}$	$\mathbf{56.1}_{\pm1.9}$	$\mathbf{66.0}_{\pm1.7}$	$73.7_{\pm1.4}$	$\underline{66.7}_{\pm1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$44.0_{\pm37.8}$	$13.5_{\pm8.8}$	$29.3_{\pm3.8}$	$31.1_{\pm7.4}$	$0.9_{\pm0.8}$	$11.8_{\pm5.1}$	$3.1_{\pm1.9}$	$19.1_{\pm9.3}$
	Score ↑	$0.51_{\pm0.16}$	$0.56_{\pm0.04}$	$0.48_{\pm0.02}$	$0.59_{\pm0.03}$	$\mathbf{0.55}_{\pm0.02}$	$0.59_{\pm0.01}$	$0.71_{\pm0.01}$	$0.57_{\pm0.04}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$60.0_{\pm0.8}$	$59.3_{\pm0.9}$	$54.5_{\pm0.6}$	$75.8_{\pm0.1}$	$48.7_{\pm1.2}$	$59.8_{\pm0.6}$	$71.4_{\pm0.2}$	$61.3_{\pm0.6}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$\underline{9.0}_{\pm1.6}$	$60.6_{\pm4.0}$	$58.0_{\pm1.0}$	$62.0_{\pm1.2}$	$11.0_{\pm3.1}$	$30.4_{\pm1.0}$	$8.8_{\pm4.4}$	$34.3_{\pm2.3}$
	Score ↑	$0.55_{\pm0.0}$	$0.37_{\pm0.0}$	$0.35_{\pm0.0}$	$0.47_{\pm0.0}$	$0.44_{\pm0.0}$	$0.46_{\pm0.0}$	$0.66_{\pm0.0}$	$0.47_{\pm0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$\mathbf{71.0}_{\pm1.4}$	$\mathbf{66.1}_{\pm1.2}$	$\mathbf{62.8}_{\pm0.3}$	$\mathbf{78.4}_{\pm0.4}$	$\underline{55.4}_{\pm0.8}$	$65.2_{\pm1.0}$	$\mathbf{75.0}_{\pm0.9}$	$\mathbf{67.7}_{\pm0.8}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$74.9_{\pm5.5}$	$30.9_{\pm4.5}$	$37.9_{\pm4.3}$	$71.5_{\pm11.4}$	$3.6_{\pm0.4}$	$17.5_{\pm0.5}$	$5.0_{\pm0.7}$	$34.5_{\pm3.9}$
	Score ↑	$0.41_{\pm0.0}$	$0.51_{\pm0.0}$	$0.46_{\pm0.0}$	$0.46_{\pm0.0}$	$0.53_{\pm0.0}$	$\underline{0.55}_{\pm0.0}$	$0.71_{\pm0.0}$	$0.52_{\pm0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	$67.8_{\pm2.3}$	$62.7_{\pm1.0}$	$58.9_{\pm1.0}$	$77.0_{\pm0.4}$	$54.0_{\pm1.7}$	$62.1_{\pm0.8}$	$\underline{74.0}_{\pm1.0}$	$65.2_{\pm1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
	Score ↑	$\mathbf{0.68}_{\pm0.0}$	$\mathbf{0.63}_{\pm0.0}$	$\mathbf{0.59}_{\pm0.0}$	$\mathbf{0.77}_{\pm0.0}$	$\underline{0.54}_{\pm0.0}$	$\mathbf{0.62}_{\pm0.0}$	$\mathbf{0.74}_{\pm0.0}$	$\mathbf{0.65}_{\pm0.0}$

Table A.23. **Results for Single-Class SCADA Unlearning on Office 31.** Forget class is $c_{\mathcal{F}} = 1$ ***(Best result in bold, second-best underlined)***

Method	Metric	A → D	A → W	D → A	D → W	W → A	W → D	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	78.7$_{\pm 2.0}$	81.2$_{\pm 1.1}$	61.1$_{\pm 0.0}$	76.4$_{\pm 1.3}$	67.0$_{\pm 0.6}$	80.3$_{\pm 1.9}$	74.1$_{\pm 1.2}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	91.9$_{\pm 0.7}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	90.6$_{\pm 0.8}$	97.1$_{\pm 0.2}$
	Score ↑	0.41$_{\pm 0.0}$	0.41$_{\pm 0.0}$	0.31$_{\pm 0.0}$	0.38$_{\pm 0.0}$	0.34$_{\pm 0.0}$	0.42$_{\pm 0.0}$	0.38$_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	79.0$_{\pm 1.0}$	82.0$_{\pm 0.4}$	63.9$_{\pm 0.6}$	78.4$_{\pm 2.1}$	67.4$_{\pm 0.0}$	78.1$_{\pm 2.8}$	74.8$_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$
	Score ↑	0.79$_{\pm 0.0}$	0.82$_{\pm 0.0}$	0.64$_{\pm 0.0}$	0.78$_{\pm 0.0}$	0.67$_{\pm 0.0}$	0.78$_{\pm 0.0}$	0.75$_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**79.1**$_{\pm 1.8}$	81.2$_{\pm 1.3}$	59.6$_{\pm 0.6}$	76.0$_{\pm 1.7}$	67.0$_{\pm 0.6}$	**79.8**$_{\pm 1.7}$	73.8$_{\pm 1.3}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	88.2$_{\pm 1.4}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	89.0$_{\pm 0.0}$	96.2$_{\pm 0.2}$
	Score ↑	0.42$_{\pm 0.0}$	0.41$_{\pm 0.0}$	0.30$_{\pm 0.0}$	0.38$_{\pm 0.0}$	0.34$_{\pm 0.0}$	0.42$_{\pm 0.0}$	0.38$_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	40.5$_{\pm 26.}$	75.5$_{\pm 3.7}$	60.4$_{\pm 4.8}$	67.9$_{\pm 5.4}$	65.3$_{\pm 1.8}$	52.7$_{\pm 3.8}$	60.4$_{\pm 7.7}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	66.6$_{\pm 31.}$	**0.0**$_{\pm 0.0}$	100.0$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	27.8$_{\pm 5.2}$
	Score ↑	0.41$_{\pm 0.3}$	0.76$_{\pm 0.0}$	0.37$_{\pm 0.1}$	0.68$_{\pm 0.1}$	0.33$_{\pm 0.0}$	0.53$_{\pm 0.0}$	0.51$_{\pm 0.1}$
ZSMU [10]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.6$_{\pm 3.0}$	81.6$_{\pm 1.0}$	59.6$_{\pm 4.3}$	76.2$_{\pm 0.8}$	67.4$_{\pm 0.0}$	79.1$_{\pm 2.7}$	73.6$_{\pm 2.0}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	87.8$_{\pm 1.2}$	100.$_{\pm 0.0}$	100.0$_{\pm 0.0}$	100.$_{\pm 0.0}$	100.$_{\pm 0.0}$	88.2$_{\pm 1.8}$	96.0$_{\pm 0.5}$
	Score ↑	0.41$_{\pm 0.0}$	0.41$_{\pm 0.0}$	0.30$_{\pm 0.0}$	0.38$_{\pm 0.0}$	0.34$_{\pm 0.0}$	0.42$_{\pm 0.0}$	0.38$_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.4$_{\pm 2.0}$	79.0$_{\pm 2.9}$	53.7$_{\pm 13.}$	74.5$_{\pm 1.0}$	67.0$_{\pm 0.6}$	77.6$_{\pm 3.6}$	71.5$_{\pm 4.0}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	42.3$_{\pm 17.}$	15.9$_{\pm 23.}$	66.7$_{\pm 57.}$	33.3$_{\pm 57.}$	39.7$_{\pm 48.}$	7.7$_{\pm 13.}$	34.3$_{\pm 36.}$
	Score ↑	0.55$_{\pm 0.1}$	0.70$_{\pm 0.1}$	0.33$_{\pm 0.0}$	0.62$_{\pm 0.2}$	0.51$_{\pm 0.2}$	0.73$_{\pm 0.1}$	0.57$_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	69.2$_{\pm 3.4}$	**81.8**$_{\pm 0.7}$	58.9$_{\pm 2.8}$	75.3$_{\pm 1.3}$	66.3$_{\pm 1.8}$	71.9$_{\pm 3.5}$	70.6$_{\pm 2.3}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm 0.0}$	36.5$_{\pm 35.}$	74.6$_{\pm 9.9}$	**0.0**$_{\pm 0.0}$	100.$_{\pm 0.0}$	0.4$_{\pm 0.7}$	35.3$_{\pm 7.6}$
	Score ↑	0.69$_{\pm 0.0}$	0.62$_{\pm 0.1}$	0.34$_{\pm 0.0}$	0.75$_{\pm 0.0}$	0.33$_{\pm 0.0}$	0.72$_{\pm 0.0}$	0.58$_{\pm 0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	59.5$_{\pm 0.8}$	77.0$_{\pm 1.3}$	**79.0**$_{\pm 2.1}$	**81.9**$_{\pm 0.9}$	**80.5**$_{\pm 0.9}$	66.9$_{\pm 0.8}$	74.1$_{\pm 1.1}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	100.0$_{\pm 0.0}$	100.0$_{\pm 0.0}$	90.9$_{\pm 0.9}$	100.0$_{\pm 0.0}$	90.2$_{\pm 0.0}$	100.0$_{\pm 0.0}$	96.8$_{\pm 0.2}$
	Score ↑	0.30$_{\pm 0.00}$	0.39$_{\pm 0.01}$	0.41$_{\pm 0.01}$	0.41$_{\pm 0.01}$	0.42$_{\pm 0.00}$	0.33$_{\pm 0.00}$	0.38$_{\pm 0.01}$
PADA [5]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.2$_{\pm 1.9}$	81.4$_{\pm 1.0}$	61.1$_{\pm 1.8}$	75.8$_{\pm 1.0}$	67.0$_{\pm 0.6}$	77.3$_{\pm 0.6}$	73.3$_{\pm 1.2}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	92.3$_{\pm 0.7}$	100.0$_{\pm 0.0}$	100.0$_{\pm 0.0}$	100.0$_{\pm 0.0}$	100.0$_{\pm 0.0}$	92.3$_{\pm 0.7}$	97.4$_{\pm 0.2}$
	Score ↑	0.40$_{\pm 0.0}$	0.41$_{\pm 0.0}$	0.31$_{\pm 0.0}$	0.38$_{\pm 0.0}$	0.34$_{\pm 0.0}$	0.40$_{\pm 0.0}$	0.37$_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.0$_{\pm 2.4}$	**81.8**$_{\pm 1.3}$	60.0$_{\pm 1.8}$	75.1$_{\pm 1.6}$	67.0$_{\pm 0.6}$	75.1$_{\pm 1.2}$	72.7$_{\pm 1.5}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	54.9$_{\pm 0.0}$	6.4$_{\pm 2.7}$	1.6$_{\pm 2.8}$	1.6$_{\pm 2.8}$	3.2$_{\pm 2.8}$	51.2$_{\pm 1.2}$	19.8$_{\pm 2.0}$
	Score ↑	0.50$_{\pm 0.0}$	0.77$_{\pm 0.0}$	0.59$_{\pm 0.0}$	0.74$_{\pm 0.0}$	0.65$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.62$_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	77.9$_{\pm 2.0}$	81.2$_{\pm 1.1}$	62.8$_{\pm 1.6}$	76.2$_{\pm 1.0}$	67.0$_{\pm 0.6}$	79.7$_{\pm 2.3}$	**74.2**$_{\pm 1.4}$
	$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$
	Score ↑	**0.78**$_{\pm 0.0}$	**0.81**$_{\pm 0.0}$	**0.63**$_{\pm 0.0}$	**0.76**$_{\pm 0.0}$	**0.67**$_{\pm 0.0}$	**0.80**$_{\pm 0.0}$	**0.74**$_{\pm 0.0}$

Table A.24. **Results for UC-SCADA Unlearning on OfficeHome.** Forget classes are $\mathcal{C}_{\mathcal{F}} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	A → C	A → P	A → R	C → A	C → P	C → R	P → A	P → C	P → R	R → A	R → C	R → P	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	61.0$_{\pm 1.6}$	82.2$_{\pm 1.5}$	82.3$_{\pm 0.3}$	76.9$_{\pm 1.7}$	82.9$_{\pm 1.0}$	82.5$_{\pm 0.9}$	75.2$_{\pm 1.4}$	58.2$_{\pm 1.2}$	84.3$_{\pm 1.3}$	78.5$_{\pm 0.2}$	58.6$_{\pm 1.1}$	87.0$_{\pm 0.1}$	75.8$_{\pm 1.0}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	28.7$_{\pm 1.9}$	71.1$_{\pm 5.0}$	86.1$_{\pm 1.0}$	57.1$_{\pm 1.1}$	69.0$_{\pm 1.3}$	76.7$_{\pm 2.3}$	47.8$_{\pm 1.9}$	22.0$_{\pm 0.4}$	72.8$_{\pm 5.7}$	53.1$_{\pm 3.4}$	35.8$_{\pm 1.4}$	77.1$_{\pm 1.7}$	58.1$_{\pm 2.2}$
	Score ↑	0.47$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.51$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.51$_{\pm 0.0}$	0.43$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.48$_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	60.8$_{\pm 1.2}$	83.1$_{\pm 0.6}$	82.2$_{\pm 0.6}$	78.5$_{\pm 2.1}$	82.6$_{\pm 0.9}$	83.3$_{\pm 1.5}$	76.4$_{\pm 0.9}$	59.7$_{\pm 2.5}$	84.4$_{\pm 0.6}$	77.6$_{\pm 1.3}$	59.9$_{\pm 1.6}$	87.3$_{\pm 0.4}$	76.3$_{\pm 1.2}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$	0.0$_{\pm 0.0}$
	Score ↑	0.61$_{\pm 0.0}$	0.83$_{\pm 0.0}$	0.82$_{\pm 0.0}$	0.79$_{\pm 0.0}$	0.83$_{\pm 0.0}$	0.83$_{\pm 0.0}$	0.76$_{\pm 0.0}$	0.60$_{\pm 0.0}$	0.84$_{\pm 0.0}$	0.78$_{\pm 0.0}$	0.60$_{\pm 0.0}$	0.87$_{\pm 0.0}$	0.76$_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	**60.1**$_{\pm 1.1}$	**81.5**$_{\pm 1.0}$	**82.6**$_{\pm 0.6}$	**77.8**$_{\pm 1.7}$	83.1$_{\pm 0.3}$	83.2$_{\pm 1.1}$	**76.5**$_{\pm 0.5}$	58.3$_{\pm 1.5}$	84.4$_{\pm 1.3}$	**79.3**$_{\pm 0.4}$	58.6$_{\pm 0.9}$	87.3$_{\pm 0.4}$	**76.1**$_{\pm 0.9}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	24.8$_{\pm 3.1}$	66.5$_{\pm 4.4}$	82.0$_{\pm 2.2}$	44.5$_{\pm 1.3}$	61.4$_{\pm 2.4}$	75.1$_{\pm 1.6}$	36.7$_{\pm 1.6}$	1.5$_{\pm 0.4}$	59.6$_{\pm 3.1}$	43.5$_{\pm 3.3}$	21.3$_{\pm 0.8}$	72.9$_{\pm 2.3}$	49.2$_{\pm 2.2}$
	Score ↑	0.48$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.54$_{\pm 0.0}$	0.51$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.56$_{\pm 0.0}$	**0.57**$_{\pm 0.0}$	0.53$_{\pm 0.0}$	0.55$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.51$_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	17.0$_{\pm 11.3}$	56.6$_{\pm 4.2}$	51.9$_{\pm 3.8}$	23.6$_{\pm 8.0}$	39.4$_{\pm 17.7}$	49.2$_{\pm 3.8}$	12.6$_{\pm 4.3}$	25.1$_{\pm 14.2}$	52.1$_{\pm 2.1}$	8.4$_{\pm 2.4}$	25.6$_{\pm 2.8}$	59.0$_{\pm 6.0}$	35.0$_{\pm 6.7}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$
	Score ↑	0.17$_{\pm 0.1}$	0.57$_{\pm 0.0}$	0.52$_{\pm 0.0}$	0.24$_{\pm 0.1}$	0.39$_{\pm 0.2}$	0.49$_{\pm 0.0}$	0.13$_{\pm 0.0}$	0.25$_{\pm 0.1}$	0.52$_{\pm 0.0}$	0.08$_{\pm 0.0}$	0.26$_{\pm 0.0}$	0.59$_{\pm 0.1}$	0.35$_{\pm 0.1}$
ZSMU [10]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	43.5$_{\pm 23.7}$	78.9$_{\pm 2.5}$	80.5$_{\pm 1.5}$	74.3$_{\pm 1.6}$	79.0$_{\pm 1.1}$	78.5$_{\pm 2.2}$	72.9$_{\pm 0.3}$	55.8$_{\pm 2.9}$	83.9$_{\pm 0.8}$	76.0$_{\pm 1.3}$	54.8$_{\pm 4.1}$	85.6$_{\pm 1.2}$	72.0$_{\pm 3.6}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	11.5$_{\pm 9.9}$	62.1$_{\pm 4.4}$	73.3$_{\pm 8.7}$	52.5$_{\pm 11.5}$	58.0$_{\pm 11.3}$	56.4$_{\pm 0.8}$	45.4$_{\pm 7.0}$	17.9$_{\pm 2.4}$	68.1$_{\pm 4.1}$	41.7$_{\pm 3.3}$	24.2$_{\pm 4.6}$	70.3$_{\pm 11.1}$	48.4$_{\pm 6.6}$
	Score ↑	0.38$_{\pm 0.2}$	0.49$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.54$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.50$_{\pm 0.0}$	0.48$_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	56.8$_{\pm 3.2}$	64.5$_{\pm 22.6}$	63.0$_{\pm 30.6}$	42.1$_{\pm 24.5}$	75.0$_{\pm 2.7}$	66.6$_{\pm 18.3}$	67.0$_{\pm 7.8}$	45.1$_{\pm 7.5}$	59.4$_{\pm 16.8}$	63.6$_{\pm 16.4}$	44.1$_{\pm 11.1}$	79.6$_{\pm 4.0}$	60.6$_{\pm 13.8}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	10.4$_{\pm 6.4}$	27.5$_{\pm 23.8}$	39.2$_{\pm 34.8}$	13.3$_{\pm 18.3}$	18.2$_{\pm 5.2}$	32.0$_{\pm 17.1}$	15.8$_{\pm 7.4}$	4.6$_{\pm 2.4}$	9.1$_{\pm 15.7}$	24.4$_{\pm 21.7}$	15.0$_{\pm 12.0}$	18.8$_{\pm 5.7}$	19.0$_{\pm 14.2}$
	Score ↑	0.51$_{\pm 0.0}$	0.49$_{\pm 0.1}$	0.43$_{\pm 0.1}$	0.36$_{\pm 0.2}$	0.64$_{\pm 0.0}$	0.50$_{\pm 0.1}$	0.58$_{\pm 0.1}$	0.43$_{\pm 0.1}$	0.54$_{\pm 0.1}$	0.51$_{\pm 0.1}$	0.38$_{\pm 0.1}$	0.67$_{\pm 0.0}$	0.50$_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	47.1$_{\pm 3.5}$	70.0$_{\pm 6.7}$	69.6$_{\pm 1.6}$	62.9$_{\pm 2.2}$	73.6$_{\pm 1.2}$	62.8$_{\pm 8.5}$	59.8$_{\pm 3.0}$	44.8$_{\pm 2.9}$	66.7$_{\pm 2.6}$	58.9$_{\pm 3.3}$	44.9$_{\pm 1.3}$	77.1$_{\pm 2.9}$	61.5$_{\pm 3.3}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	0.5$_{\pm 0.9}$	0.3$_{\pm 0.5}$	**0.0**$_{\pm 0.0}$	0.8$_{\pm 1.4}$	1.3$_{\pm 1.1}$	**0.0**$_{\pm 0.0}$	0.9$_{\pm 1.6}$	1.2$_{\pm 2.1}$	3.1$_{\pm 2.2}$	5.1$_{\pm 8.8}$	1.1$_{\pm 1.6}$
	Score ↑	0.47$_{\pm 0.0}$	0.70$_{\pm 0.1}$	0.69$_{\pm 0.0}$	0.63$_{\pm 0.0}$	0.74$_{\pm 0.0}$	0.62$_{\pm 0.1}$	0.59$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.66$_{\pm 0.0}$	0.58$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.74$_{\pm 0.1}$	0.61$_{\pm 0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	51.1$_{\pm 1.1}$	75.6$_{\pm 0.6}$	77.7$_{\pm 0.1}$	72.0$_{\pm 0.7}$	**83.6**$_{\pm 0.1}$	**83.6**$_{\pm 0.3}$	69.5$_{\pm 0.4}$	**59.7**$_{\pm 1.2}$	**86.1**$_{\pm 0.5}$	73.7$_{\pm 0.4}$	60.5$_{\pm 0.9}$	**88.4**$_{\pm 0.0}$	73.1$_{\pm 0.5}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	19.7$_{\pm 2.5}$	77.3$_{\pm 2.1}$	77.8$_{\pm 4.3}$	35.4$_{\pm 5.2}$	58.1$_{\pm 0.7}$	81.3$_{\pm 0.7}$	34.2$_{\pm 3.5}$	40.2$_{\pm 8.8}$	87.4$_{\pm 2.1}$	35.4$_{\pm 1.8}$	41.1$_{\pm 2.6}$	87.4$_{\pm 0.7}$	52.1$_{\pm 2.5}$
	Score ↑	0.43$_{\pm 0.02}$	0.43$_{\pm 0.01}$	0.44$_{\pm 0.01}$	0.53$_{\pm 0.02}$	0.53$_{\pm 0.00}$	0.46$_{\pm 0.00}$	0.52$_{\pm 0.01}$	0.43$_{\pm 0.04}$	0.46$_{\pm 0.00}$	0.55$_{\pm 0.00}$	0.43$_{\pm 0.01}$	0.47$_{\pm 0.00}$	0.48$_{\pm 0.01}$
PADA [5]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	58.4$_{\pm 0.9}$	80.2$_{\pm 2.4}$	80.8$_{\pm 0.6}$	76.3$_{\pm 1.7}$	80.7$_{\pm 1.2}$	80.8$_{\pm 0.7}$	73.5$_{\pm 2.4}$	56.5$_{\pm 1.7}$	82.7$_{\pm 2.4}$	77.3$_{\pm 1.4}$	54.5$_{\pm 0.6}$	86.0$_{\pm 1.1}$	74.0$_{\pm 1.4}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	33.3$_{\pm 0.3}$	77.4$_{\pm 2.6}$	82.5$_{\pm 0.8}$	64.8$_{\pm 2.5}$	70.9$_{\pm 2.5}$	73.3$_{\pm 1.2}$	56.5$_{\pm 0.9}$	31.0$_{\pm 0.6}$	79.4$_{\pm 0.8}$	60.5$_{\pm 0.5}$	38.7$_{\pm 2.1}$	83.0$_{\pm 0.5}$	62.6$_{\pm 1.3}$
	Score ↑	0.44$_{\pm 0.0}$	0.45$_{\pm 0.0}$	0.44$_{\pm 0.0}$	0.46$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.43$_{\pm 0.0}$	0.46$_{\pm 0.0}$	0.48$_{\pm 0.0}$	0.39$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.45$_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	56.0$_{\pm 2.1}$	76.8$_{\pm 2.3}$	79.2$_{\pm 1.1}$	74.4$_{\pm 2.0}$	80.9$_{\pm 1.2}$	80.0$_{\pm 1.4}$	71.6$_{\pm 1.0}$	58.4$_{\pm 1.3}$	81.4$_{\pm 2.1}$	76.8$_{\pm 0.5}$	**63.0**$_{\pm 0.6}$	85.8$_{\pm 0.6}$	73.7$_{\pm 1.3}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	23.9$_{\pm 1.0}$	27.6$_{\pm 2.5}$	28.2$_{\pm 1.7}$	31.2$_{\pm 1.0}$	19.3$_{\pm 0.8}$	13.0$_{\pm 0.8}$	29.6$_{\pm 1.9}$	23.9$_{\pm 2.3}$	15.0$_{\pm 0.4}$	36.1$_{\pm 0.9}$	28.8$_{\pm 0.7}$	19.4$_{\pm 0.3}$	24.7$_{\pm 1.2}$
	Score ↑	0.45$_{\pm 0.0}$	0.60$_{\pm 0.0}$	0.62$_{\pm 0.0}$	0.57$_{\pm 0.0}$	0.68$_{\pm 0.0}$	0.71$_{\pm 0.0}$	0.55$_{\pm 0.0}$	0.47$_{\pm 0.0}$	0.71$_{\pm 0.0}$	0.56$_{\pm 0.0}$	0.49$_{\pm 0.0}$	0.72$_{\pm 0.0}$	0.59$_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	59.5$_{\pm 1.9}$	79.5$_{\pm 1.0}$	78.7$_{\pm 1.2}$	74.8$_{\pm 0.7}$	80.6$_{\pm 1.3}$	78.8$_{\pm 1.2}$	73.7$_{\pm 0.4}$	55.7$_{\pm 0.6}$	80.7$_{\pm 0.5}$	74.2$_{\pm 0.8}$	56.7$_{\pm 0.5}$	82.8$_{\pm 0.8}$	73.0$_{\pm 0.9}$
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	0.1$_{\pm 0.2}$	**0.0**$_{\pm 0.0}$	4.1$_{\pm 5.8}$	**0.0**$_{\pm 0.0}$	**0.0**$_{\pm 0.0}$	12.8$_{\pm 9.1}$	**0.0**$_{\pm 0.0}$	1.4$_{\pm 1.3}$
	Score ↑	**0.60**$_{\pm 0.0}$	**0.80**$_{\pm 0.0}$	**0.79**$_{\pm 0.0}$	**0.75**$_{\pm 0.0}$	**0.81**$_{\pm 0.0}$	**0.79**$_{\pm 0.0}$	**0.74**$_{\pm 0.0}$	0.54$_{\pm 0.0}$	**0.81**$_{\pm 0.0}$	**0.74**$_{\pm 0.0}$	**0.51**$_{\pm 0.0}$	**0.83**$_{\pm 0.0}$	**0.72**$_{\pm 0.0}$

Table A.25. **Results for UC-SCADA Unlearning on DomainNet.** Forget classes are $\mathcal{C}_\mathcal{F} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Average
Original (SF(DA)[2] [28])	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$71.3_{\pm 0.3}$	$66.6_{\pm 0.4}$	$61.7_{\pm 0.8}$	$78.3_{\pm 0.2}$	$55.9_{\pm 1.5}$	$65.1_{\pm 0.9}$	$75.0_{\pm 0.4}$	$67.7_{\pm 0.6}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$67.6_{\pm 5.7}$	$55.9_{\pm 4.8}$	$42.9_{\pm 2.2}$	$60.6_{\pm 2.2}$	$6.5_{\pm 1.1}$	$14.6_{\pm 0.9}$	$22.8_{\pm 3.1}$	$38.7_{\pm 2.9}$
	Score ↑	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.49_{\pm 0.0}$	$0.53_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.61_{\pm 0.0}$	$0.50_{\pm 0.0}$
Retrain	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$71.0_{\pm 0.3}$	$65.3_{\pm 2.3}$	$58.6_{\pm 2.4}$	$78.2_{\pm 0.9}$	$54.2_{\pm 1.0}$	$62.9_{\pm 1.5}$	$74.1_{\pm 1.0}$	$66.3_{\pm 1.3}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$
	Score ↑	$0.71_{\pm 0.0}$	$0.65_{\pm 0.0}$	$0.59_{\pm 0.0}$	$0.78_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.63_{\pm 0.0}$	$0.74_{\pm 0.0}$	$0.66_{\pm 0.0}$
Finetune	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$68.9_{\pm 0.1}$	$\underline{64.9}_{\pm 1.1}$	$62.2_{\pm 1.2}$	$77.3_{\pm 0.2}$	$52.1_{\pm 1.9}$	$\underline{65.5}_{\pm 1.3}$	$\underline{74.2}_{\pm 0.8}$	$\underline{66.5}_{\pm 0.9}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$37.4_{\pm 5.8}$	$32.3_{\pm 2.6}$	$20.8_{\pm 6.3}$	$36.6_{\pm 3.1}$	$\underline{0.6}_{\pm 0.3}$	$5.3_{\pm 2.1}$	$10.1_{\pm 8.5}$	$20.4_{\pm 4.1}$
	Score ↑	$\underline{0.50}_{\pm 0.0}$	$\mathbf{0.49}_{\pm 0.0}$	$\mathbf{0.52}_{\pm 0.0}$	$\underline{0.57}_{\pm 0.0}$	$\mathbf{0.52}_{\pm 0.0}$	$\underline{0.62}_{\pm 0.0}$	$\underline{0.68}_{\pm 0.1}$	$\mathbf{0.56}_{\pm 0.0}$
UNSIR [51]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$14.1_{\pm 6.8}$	$1.5_{\pm 0.1}$	$25.2_{\pm 2.2}$	$40.4_{\pm 2.0}$	$1.5_{\pm 0.3}$	$14.4_{\pm 9.2}$	$23.1_{\pm 2.4}$	$17.2_{\pm 3.3}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{4.4}_{\pm 7.6}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.6}_{\pm 1.1}$
	Score ↑	$0.14_{\pm 0.1}$	$0.01_{\pm 0.0}$	$0.25_{\pm 0.0}$	$0.39_{\pm 0.0}$	$0.02_{\pm 0.0}$	$0.14_{\pm 0.1}$	$0.23_{\pm 0.0}$	$0.17_{\pm 0.0}$
ZSMU [10]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$\underline{70.0}_{\pm 0.9}$	$58.1_{\pm 3.1}$	$60.2_{\pm 2.2}$	$76.8_{\pm 1.5}$	$49.1_{\pm 1.0}$	$58.0_{\pm 4.1}$	$70.3_{\pm 1.7}$	$63.2_{\pm 2.1}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$67.0_{\pm 20.7}$	$58.6_{\pm 10.0}$	$39.5_{\pm 12.3}$	$49.6_{\pm 16.3}$	$23.8_{\pm 5.0}$	$13.8_{\pm 2.8}$	$25.2_{\pm 7.3}$	$39.7_{\pm 10.6}$
	Score ↑	$0.42_{\pm 0.1}$	$0.37_{\pm 0.0}$	$0.43_{\pm 0.1}$	$0.52_{\pm 0.1}$	$0.40_{\pm 0.0}$	$0.51_{\pm 0.0}$	$0.56_{\pm 0.0}$	$0.46_{\pm 0.0}$
Lipschitz [15]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$49.5_{\pm 14.9}$	$43.2_{\pm 17.5}$	$47.3_{\pm 11.9}$	$60.8_{\pm 13.4}$	$43.2_{\pm 5.0}$	$29.4_{\pm 21.0}$	$67.8_{\pm 2.2}$	$48.7_{\pm 12.3}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$20.2_{\pm 29.7}$	$\underline{20.2}_{\pm 19.5}$	$23.6_{\pm 20.5}$	$27.5_{\pm 22.0}$	$1.9_{\pm 2.5}$	$1.1_{\pm 2.0}$	$7.5_{\pm 9.2}$	$14.6_{\pm 15.1}$
	Score ↑	$0.41_{\pm 0.0}$	$0.35_{\pm 0.1}$	$0.38_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.29_{\pm 0.2}$	$0.63_{\pm 0.1}$	$0.42_{\pm 0.1}$
Nabla Tau [53]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$43.5_{\pm 2.0}$	$42.2_{\pm 1.7}$	$45.7_{\pm 2.3}$	$59.3_{\pm 4.9}$	$23.3_{\pm 19.1}$	$45.6_{\pm 3.6}$	$47.6_{\pm 1.3}$	$43.9_{\pm 5.0}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$\underline{12.3}_{\pm 3.8}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{0.3}_{\pm 0.5}$	$\mathbf{3.5}_{\pm 3.1}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{2.3}_{\pm 1.0}$
	Score ↑	$0.39_{\pm 0.0}$	$0.42_{\pm 0.0}$	$\underline{0.46}_{\pm 0.0}$	$\underline{0.57}_{\pm 0.1}$	$0.23_{\pm 0.2}$	$0.46_{\pm 0.0}$	$0.48_{\pm 0.0}$	$0.43_{\pm 0.0}$
Unlearned(+) [2]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$65.8_{\pm 1.2}$	$61.6_{\pm 1.2}$	$\mathbf{63.2}_{\pm 0.8}$	$77.2_{\pm 0.1}$	$\underline{52.8}_{\pm 1.7}$	$\mathbf{65.7}_{\pm 1.3}$	$74.1_{\pm 0.2}$	$65.8_{\pm 0.8}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$38.4_{\pm 5.2}$	$6.8_{\pm 3.8}$	$29.7_{\pm 1.8}$	$39.9_{\pm 8.3}$	$2.1_{\pm 0.9}$	$16.0_{\pm 5.4}$	$1.6_{\pm 0.4}$	$19.2_{\pm 3.0}$
	Score ↑	$0.48_{\pm 0.00}$	$0.58_{\pm 0.03}$	$0.49_{\pm 0.01}$	$0.55_{\pm 0.03}$	$0.52_{\pm 0.02}$	$0.57_{\pm 0.04}$	$0.73_{\pm 0.00}$	$0.56_{\pm 0.02}$
PADA [5]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$61.1_{\pm 0.3}$	$59.7_{\pm 0.7}$	$55.1_{\pm 0.3}$	$76.1_{\pm 0.5}$	$49.0_{\pm 0.3}$	$59.9_{\pm 0.4}$	$72.2_{\pm 0.3}$	$61.9_{\pm 0.4}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$45.5_{\pm 1.3}$	$70.1_{\pm 0.4}$	$57.1_{\pm 1.1}$	$71.6_{\pm 1.1}$	$22.3_{\pm 0.8}$	$26.8_{\pm 0.9}$	$32.7_{\pm 2.4}$	$46.6_{\pm 1.1}$
	Score ↑	$0.42_{\pm 0.0}$	$0.35_{\pm 0.0}$	$0.35_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.47_{\pm 0.0}$	$0.54_{\pm 0.0}$	$0.43_{\pm 0.0}$
SHOT [34]	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$\mathbf{71.3}_{\pm 0.2}$	$\mathbf{66.4}_{\pm 1.6}$	$\underline{62.4}_{\pm 0.6}$	$\mathbf{78.5}_{\pm 0.4}$	$\mathbf{55.7}_{\pm 2.4}$	$65.1_{\pm 0.4}$	$\mathbf{74.8}_{\pm 0.9}$	$\mathbf{67.8}_{\pm 0.9}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$67.7_{\pm 4.5}$	$56.3_{\pm 5.1}$	$48.7_{\pm 3.0}$	$62.4_{\pm 3.3}$	$6.9_{\pm 2.1}$	$13.6_{\pm 0.5}$	$21.2_{\pm 0.5}$	$39.6_{\pm 2.7}$
	Score ↑	$0.43_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.48_{\pm 0.0}$	$\mathbf{0.52}_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.62_{\pm 0.0}$	$0.50_{\pm 0.0}$
Ours	$A_{\mathcal{D}_r^\mathcal{T}} \uparrow$	$67.3_{\pm 1.4}$	$63.3_{\pm 0.7}$	$60.8_{\pm 1.8}$	$\underline{77.6}_{\pm 0.2}$	$52.4_{\pm 2.3}$	$63.3_{\pm 0.5}$	$72.7_{\pm 0.8}$	$65.3_{\pm 1.1}$
	$A_{\mathcal{D}_f^\mathcal{T}} \downarrow$	$29.6_{\pm 9.8}$	$36.8_{\pm 3.1}$	$36.3_{\pm 7.1}$	$28.7_{\pm 5.9}$	$4.5_{\pm 5.4}$	$\underline{0.7}_{\pm 1.0}$	$\underline{3.6}_{\pm 5.1}$	$20.1_{\pm 5.3}$
	Score ↑	$\mathbf{0.52}_{\pm 0.0}$	$\underline{0.46}_{\pm 0.0}$	$0.45_{\pm 0.0}$	$\mathbf{0.60}_{\pm 0.0}$	$\underline{0.50}_{\pm 0.0}$	$\mathbf{0.63}_{\pm 0.0}$	$\mathbf{0.70}_{\pm 0.0}$	$\underline{0.55}_{\pm 0.0}$

Table A.26. **Results for UC-SCADA Unlearning on Office 31.** Forget classes are $\mathcal{C}_\mathcal{F} = \{1, 2, 3\}$ *(Best result in bold, second-best underlined)*

Method	Metric	A → D	A → W	D → A	D → W	W → A	W → D	Average
Original (SF(DA)^2 [28])	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	67.0±0.9	79.9±1.8	77.1±0.9	86.3±1.5	78.5±1.5	72.0±0.5	76.8±1.2
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	93.6±0.8	88.5±2.3	84.8±2.3	99.5±0.8	76.6±1.9	97.7±1.6	90.1±1.6
	Score ↑	0.35±0.0	0.42±0.0	0.42±0.0	0.43±0.0	0.44±0.0	0.36±0.0	0.40±0.0
Retrain	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	68.2±0.9	83.3±0.8	76.0±4.1	87.0±0.6	78.0±1.2	72.0±1.0	77.4±1.4
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
	Score ↑	0.68±0.0	0.83±0.0	0.76±0.0	0.87±0.0	0.78±0.0	0.72±0.0	0.77±0.0
Finetune	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	66.7±2.2	79.7±2.3	77.2±0.6	**86.5**±1.6	78.4±1.0	72.0±0.5	**76.7**±1.4
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	85.4±3.0	80.3±1.3	79.4±1.1	94.0±0.8	65.0±5.8	93.0±0.0	82.8±2.0
	Score ↑	0.36±0.0	0.44±0.0	0.43±0.0	0.45±0.0	0.48±0.0	0.37±0.0	0.42±0.0
UNSIR [51]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	51.0±10.8	77.4±5.9	58.0±1.2	65.7±3.6	68.6±3.3	48.8±6.9	61.6±5.3
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	1.2±1.0	**0.0**±0.0	64.3±3.6	**0.0**±0.0	10.9±0.8
	Score ↑	0.51±0.1	**0.77**±0.1	0.57±0.0	0.66±0.0	0.42±0.0	0.49±0.1	0.57±0.1
ZSMU [10]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	76.7±0.7	85.9±2.0	63.3±3.7	78.1±3.0	71.2±0.6	74.7±0.6	75.0±1.8
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	68.6±6.4	97.3±1.0	92.4±5.3	85.3±4.3	92.4±2.0	78.0±4.3	85.6±3.9
	Score ↑	0.46±0.0	0.44±0.0	0.33±0.0	0.42±0.0	0.37±0.0	0.42±0.0	0.41±0.0
Lipschitz [15]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	72.3±2.5	76.2±16.3	58.3±6.6	75.8±5.5	70.9±0.6	66.4±15.6	70.0±7.9
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	47.2±22.6	33.9±32.1	37.4±6.7	60.1±17.1	47.4±33.4	9.4±14.2	39.2±21.0
	Score ↑	0.50±0.1	0.57±0.0	0.42±0.0	0.47±0.0	0.50±0.1	0.61±0.1	0.51±0.1
Nabla Tau [53]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	71.6±1.4	85.4±1.1	65.5±4.6	75.1±1.6	71.6±0.0	70.7±3.7	73.3±2.0
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	3.1±3.7	16.4±14.6	36.9±13.7	1.6±2.8	74.2±5.3	12.7±11.0	24.2±8.5
	Score ↑	**0.70**±0.0	0.74±0.1	0.48±0.1	0.74±0.0	0.41±0.0	0.63±0.0	0.62±0.0
Unlearned(+) [2]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	61.1±0.9	74.7±2.3	**78.6**±1.6	82.5±1.1	**79.7**±0.7	66.3±0.9	73.8±1.1
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	100.0±1.3	100.0±1.1	89.0±4.8	100.0±1.2	87.8±7.5	100.0±1.3	96.1±2.8
	Score ↑	0.30±0.01	0.37±0.02	0.42±0.01	0.41±0.01	0.42±0.02	0.33±0.01	0.38±0.01
PADA [5]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	76.8±0.4	86.5±1.5	66.3±2.4	78.5±2.9	72.0±0.6	**77.4**±2.2	76.2±1.6
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	76.8±3.1	100.±0.0	90.7±2.0	90.7±1.9	98.8±1.0	77.4±1.7	89.1±1.6
	Score ↑	0.43±0.0	0.43±0.0	0.35±0.0	0.41±0.0	0.36±0.0	0.44±0.0	0.40±0.0
SHOT [34]	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	**77.2**±1.0	**87.0**±1.2	65.9±2.3	78.1±1.2	72.0±0.6	75.6±1.2	76.0±1.2
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	47.9±1.5	92.9±0.9	69.6±2.7	72.7±1.9	84.8±2.0	40.4±1.5	68.0±1.7
	Score ↑	0.52±0.0	0.45±0.0	0.39±0.0	0.45±0.0	0.39±0.0	0.54±0.0	0.46±0.0
Ours	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	60.6±0.6	72.8±2.0	72.1±1.7	77.4±1.9	71.9±1.4	67.4±2.9	70.4±1.7
	$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
	Score ↑	0.61±0.0	0.73±0.0	**0.72**±0.0	**0.77**±0.0	**0.72**±0.0	**0.67**±0.0	**0.70**±0.0

Table A.27. **C-SCADA Unlearning performance on the OfficeHome dataset.** Forget classes are $\mathcal{C_F}^1 = \{1,2\}, \mathcal{C_F}^2 = \{3,4\}, \mathcal{C_F}^3 = \{5,6\}$

Method	Task	Acc.	A → C	A → P	A → R	C → A	C → P	C → R	P → A	P → C	P → R	R → A	R → C	R → P	Average
Original (SF(DA)[2] [28])	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	63.6±0.4	80.3±0.3	86.7±0.8	78.6±0.4	81.9±0.9	87.4±0.4	76.5±1.1	62.2±0.3	89.5±0.4	81.3±1.4	63.4±1.5	87.1±0.2	78.2±0.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	43.8±1.1	61.4±2.4	75.4±1.1	59.8±3.7	55.7±1.3	66.9±3.2	46.0±2.6	26.2±0.8	62.1±1.2	58.4±4.8	40.3±0.6	63.4±0.6	55.0±2.0
		Score ↑	0.44±0.0	0.50±0.0	0.49±0.0	0.49±0.0	0.53±0.0	0.52±0.0	0.52±0.0	0.49±0.0	0.55±0.0	0.51±0.0	0.45±0.0	0.53±0.0	0.50±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	63.6±0.4	80.3±0.3	86.7±0.8	78.6±0.4	81.9±0.9	87.4±0.4	76.5±1.1	62.2±0.3	89.5±0.4	81.3±1.4	63.4±1.5	87.1±0.2	78.2±0.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	44.2±1.1	62.2±2.4	76.1±1.0	60.6±3.7	56.3±1.3	67.5±3.2	46.7±2.6	26.5±0.9	62.6±1.2	59.2±4.9	40.6±0.6	64.2±0.6	55.6±2.0
		Score ↑	0.44±0.0	0.50±0.0	0.49±0.0	0.49±0.0	0.52±0.0	0.52±0.0	0.52±0.0	0.49±0.0	0.55±0.0	0.51±0.0	0.45±0.0	0.53±0.0	0.50±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	63.6±0.4	80.3±0.3	86.7±0.8	78.6±0.4	81.9±0.9	87.4±0.4	76.5±1.1	62.2±0.3	89.5±0.4	81.3±1.4	63.4±1.5	87.1±0.2	78.2±0.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	44.3±1.1	62.4±2.4	76.3±1.1	60.9±3.8	56.6±1.3	67.7±3.2	46.9±2.6	26.5±0.8	62.8±1.2	59.5±5.0	40.8±0.6	64.5±0.6	55.8±2.0
		Score ↑	0.44±0.0	0.49±0.0	0.49±0.0	0.49±0.0	0.52±0.0	0.52±0.0	0.52±0.0	0.49±0.0	0.55±0.0	0.51±0.0	0.45±0.0	0.53±0.0	0.50±0.0
Retrain	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	65.0±1.7	80.3±1.2	86.6±1.4	78.2±1.6	83.3±0.4	87.8±0.8	78.1±1.4	63.3±1.3	88.1±0.7	77.0±2.6	65.6±0.2	85.2±0.6	78.2±1.2
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.65±0.0	0.80±0.0	0.87±0.0	0.78±0.0	0.83±0.0	0.88±0.0	0.78±0.0	0.63±0.0	0.88±0.0	0.77±0.0	0.66±0.0	0.85±0.0	0.78±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	65.0±1.7	80.3±1.2	86.6±1.4	78.2±1.6	83.3±0.4	87.8±0.8	78.1±1.4	63.3±1.3	88.1±0.7	77.0±2.6	65.6±0.2	85.2±0.6	78.2±1.2
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.65±0.0	0.80±0.0	0.87±0.0	0.78±0.0	0.83±0.0	0.88±0.0	0.78±0.0	0.63±0.0	0.88±0.0	0.77±0.0	0.66±0.0	0.85±0.0	0.78±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	65.0±1.7	80.3±1.2	86.6±1.4	78.2±1.6	83.3±0.4	87.8±0.8	78.1±1.4	63.3±1.3	88.1±0.7	77.0±2.6	65.6±0.2	85.2±0.6	78.2±1.2
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.65±0.0	0.80±0.0	0.87±0.0	0.78±0.0	0.83±0.0	0.88±0.0	0.78±0.0	0.63±0.0	0.88±0.0	0.77±0.0	0.66±0.0	0.85±0.0	0.78±0.0
Finetune	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	62.8±0.4	**79.4**±0.8	85.8±0.9	77.2±1.2	**82.0**±0.5	**86.7**±0.7	75.7±1.7	61.8±0.4	**89.5**±0.4	**80.9**±2.0	62.0±0.8	85.8±1.2	**77.5**±0.9
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	38.5±1.6	42.1±1.9	66.7±1.6	37.5±2.7	47.0±1.4	60.7±3.0	24.6±0.5	15.6±3.7	53.4±2.5	32.0±1.1	37.0±3.0	55.7±2.8	42.6±2.2
		Score ↑	0.45±0.0	0.56±0.0	0.51±0.0	0.56±0.0	0.56±0.0	0.54±0.0	0.61±0.0	0.54±0.0	0.58±0.0	0.61±0.0	0.45±0.0	0.55±0.0	0.54±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	62.8±0.4	**79.4**±0.8	85.8±0.9	77.2±1.2	82.0±0.5	**86.7**±0.7	75.7±1.7	**61.8**±0.4	**89.5**±0.4	**80.9**±2.0	62.0±0.8	85.8±1.2	77.5±0.9
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	38.8±1.6	42.6±1.9	67.3±1.6	38.0±2.7	47.6±1.5	61.2±3.1	24.9±0.5	15.8±3.7	53.9±2.5	32.5±1.1	37.3±3.0	56.3±2.8	43.0±2.2
		Score ↑	0.45±0.0	0.56±0.0	0.51±0.0	0.56±0.0	0.56±0.0	0.54±0.0	0.61±0.0	0.53±0.0	0.58±0.0	0.61±0.0	0.45±0.0	0.55±0.0	0.54±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	**62.8**±0.4	**79.4**±0.8	85.8±0.9	77.2±1.2	**82.0**±0.5	**86.7**±0.7	75.7±1.7	**61.8**±0.4	**89.5**±0.4	**80.9**±2.0	62.0±0.8	85.8±1.2	77.5±0.9
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	39.0±1.6	42.8±1.9	67.5±1.6	38.2±2.7	47.8±1.5	61.4±3.1	25.0±0.5	15.8±3.7	54.0±2.5	32.6±1.1	37.4±3.0	56.6±2.9	43.2±2.2
		Score ↑	0.45±0.0	0.56±0.0	0.51±0.0	0.56±0.0	0.56±0.0	0.54±0.0	0.61±0.0	0.53±0.0	0.58±0.0	0.61±0.0	0.45±0.0	0.55±0.0	0.54±0.0
UNSIR [51]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	30.9±2.7	54.3±3.8	52.5±3.3	16.0±5.8	60.2±4.7	48.1±2.2	17.8±2.5	15.5±9.3	52.0±2.0	16.0±10.0	24.5±9.1	59.1±2.0	37.2±4.8
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
		Score ↑	0.31±0.0	0.54±0.0	0.53±0.0	0.16±0.1	0.60±0.0	0.48±0.0	0.18±0.0	0.15±0.1	0.52±0.0	0.16±0.1	0.24±0.1	0.59±0.0	0.37±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	23.3±2.1	48.5±5.5	44.3±4.5	12.2±3.6	54.6±4.1	35.3±3.6	12.2±0.8	12.5±7.2	38.8±2.1	10.7±6.5	19.2±7.2	49.5±1.8	30.1±4.1
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
		Score ↑	0.23±0.0	0.48±0.1	0.44±0.0	0.12±0.0	0.55±0.0	0.35±0.0	0.12±0.0	0.12±0.1	0.39±0.0	0.11±0.1	0.19±0.1	0.50±0.0	0.30±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	19.6±2.4	43.8±3.9	39.9±3.5	10.9±2.1	49.4±5.0	31.3±4.6	10.2±0.9	11.7±6.9	27.1±5.7	8.7±4.7	15.7±5.4	47.0±3.3	26.3±4.0
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
		Score ↑	0.20±0.0	0.44±0.0	0.40±0.0	0.11±0.0	0.49±0.0	0.31±0.0	0.10±0.0	0.12±0.1	0.27±0.1	0.09±0.0	0.16±0.1	0.47±0.0	0.26±0.0
ZSMU [10]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	40.5±33.1	78.5±1.1	85.2±2.0	75.8±1.6	77.7±1.5	78.7±9.2	75.9±2.9	61.3±1.0	87.5±1.8	78.8±0.7	56.7±3.4	84.1±0.3	73.4±4.9
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	14.0±15.7	57.0±3.9	71.5±4.6	32.9±4.6	42.8±2.7	40.3±29.1	40.7±9.6	15.0±4.5	69.2±7.5	35.1±3.4	20.2±12.3	62.4±1.8	41.8±8.3
		Score ↑	0.34±0.3	0.50±0.0	0.50±0.0	0.57±0.0	0.54±0.0	0.57±0.1	0.54±0.0	0.53±0.0	0.52±0.0	0.58±0.0	0.47±0.0	0.52±0.0	0.52±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	40.4±33.7	36.2±39.3	84.7±2.1	49.4±37.0	76.9±0.8	78.0±9.9	72.6±5.4	39.7±32.8	86.7±2.2	51.9±43.9	55.1±3.0	84.3±1.1	63.0±17.6
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	15.0±13.3	19.2±33.3	82.5±3.7	29.5±25.7	54.6±10.7	64.3±14.5	30.4±4.7	9.4±8.4	67.5±3.0	34.0±29.5	19.7±7.9	61.6±4.7	40.7±13.3
		Score ↑	0.33±0.3	0.26±0.2	0.46±0.0	0.35±0.2	0.50±0.0	0.47±0.0	0.56±0.0	0.35±0.3	0.52±0.0	0.35±0.3	0.46±0.0	0.52±0.0	0.43±0.1
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	21.8±33.3	37.5±37.1	58.3±48.5	49.8±36.6	77.3±1.2	78.4±9.1	72.4±4.7	40.2±32.6	86.9±1.9	51.5±43.8	55.0±3.4	69.7±27.0	58.2±23.3
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	11.6±17.5	19.0±30.4	46.4±40.2	20.4±18.0	42.9±9.2	55.0±11.7	36.2±10.7	10.8±9.6	55.4±4.8	28.4±25.3	26.6±4.6	32.9±24.9	32.1±17.2
		Score ↑	0.17±0.2	0.28±0.2	0.35±0.3	0.39±0.3	0.54±0.0	0.50±0.0	0.54±0.1	0.35±0.3	0.56±0.0	0.36±0.3	0.43±0.0	0.51±0.1	0.42±0.2
Lipschitz [15]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	38.1±15.2	70.9±10.4	67.7±22.1	61.7±13.0	65.4±20.1	56.5±23.7	55.5±4.8	54.4±1.8	79.0±1.6	67.2±4.0	51.4±8.4	85.1±1.0	62.7±10.5
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	7.0±12.1	25.5±20.2	35.7±35.8	**0.0**±0.0	19.0±31.4	14.7±12.8	**0.0**±0.0	**0.0**±0.0	19.8±22.5	3.0±3.3	16.8±16.8	76.1±4.4	18.1±13.3
		Score ↑	0.35±0.1	0.58±0.2	0.50±0.1	0.62±0.1	0.56±0.2	0.48±0.2	0.55±0.0	0.54±0.0	0.67±0.1	0.65±0.0	0.45±0.1	0.48±0.0	0.54±0.1
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	46.5±3.7	68.4±10.3	68.1±13.4	57.7±28.6	72.5±12.0	55.4±15.7	49.1±26.9	33.0±12.2	83.3±3.8	76.0±2.1	50.8±20.2	79.5±4.2	61.7±12.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	9.5±9.4	8.7±15.1	15.4±24.7	27.4±23.8	19.7±19.1	3.4±5.9	1.0±1.8	**0.0**±0.0	26.7±9.5	4.0±6.9	6.8±5.4	42.7±22.1	13.8±12.0
		Score ↑	0.43±0.1	0.63±0.1	0.59±0.0	0.43±0.2	0.61±0.0	0.54±0.2	0.48±0.3	0.33±0.1	0.66±0.0	0.73±0.1	0.47±0.2	0.57±0.1	0.54±0.1
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	46.1±6.6	75.2±2.7	56.9±24.9	45.6±19.4	68.2±4.2	64.8±17.6	50.2±9.1	48.5±5.1	60.4±38.5	69.4±13.5	58.8±6.3	79.8±4.2	60.3±12.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	9.8±8.0	12.9±22.4	5.3±4.8	1.0±1.7	3.5±6.0	6.0±5.2	3.8±4.5	5.0±2.0	10.2±10.9	6.3±8.1	17.9±4.7	31.0±6.8	9.4±7.1
		Score ↑	0.42±0.1	0.68±0.1	0.54±0.2	0.45±0.2	0.66±0.0	0.61±0.1	0.48±0.1	0.46±0.0	0.53±0.3	0.65±0.1	0.50±0.1	0.61±0.0	0.55±0.1
Nabla Tau [53]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	52.8±2.7	71.5±1.6	72.2±0.8	55.7±10.3	70.3±3.5	76.2±1.5	61.1±1.9	48.8±2.4	75.0±2.2	63.4±2.0	48.9±3.0	72.6±4.9	64.0±3.1
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	1.3±2.3	**0.0**±0.0	**0.0**±0.0	1.3±2.3	0.6±1.0	3.1±4.0	6.5±8.1	**0.0**±0.0	**0.0**±0.0	0.9±1.5	**0.0**±0.0	**0.0**±0.0	1.1±1.6
		Score ↑	0.52±0.0	0.71±0.0	0.72±0.0	0.55±0.1	0.70±0.0	0.74±0.0	0.58±0.0	0.49±0.0	0.75±0.0	0.63±0.0	0.49±0.0	0.73±0.0	0.63±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	50.8±3.1	69.2±2.4	69.0±0.5	52.5±7.5	69.0±2.9	70.7±2.3	57.9±0.8	47.2±1.8	71.4±2.6	60.2±0.3	46.5±2.5	72.0±3.1	61.4±2.5
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	2.1±3.7	**0.0**±0.0	**0.0**±0.0	3.6±3.2	0.1±0.2	1.8±3.1	0.3±0.6	**0.0**±0.0	**0.0**±0.0	0.2±0.3	**0.0**±0.0	**0.0**±0.0	0.7±0.9
		Score ↑	0.50±0.0	0.69±0.0	0.69±0.0	0.51±0.1	0.69±0.0	0.69±0.0	0.58±0.0	0.47±0.0	0.71±0.0	0.60±0.0	0.47±0.0	0.72±0.0	0.61±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	49.1±1.5	68.9±2.7	67.2±0.8	52.4±7.9	69.2±2.7	68.9±3.0	57.5±0.6	45.8±0.9	71.2±3.5	60.8±1.9	46.2±2.7	72.1±4.2	60.8±2.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	5.4±1.8	**0.0**±0.0	**0.0**±0.0	1.8±1.6	0.5±0.8	0.7±1.2	2.0±1.2	4.9±2.4	**0.0**±0.0	0.2±0.3	5.7±4.6	**0.0**±0.0	1.8±1.2
		Score ↑	0.47±0.0	0.69±0.0	0.67±0.0	0.52±0.1	0.69±0.0	0.68±0.0	0.56±0.0	0.44±0.0	0.71±0.0	0.61±0.0	0.44±0.0	0.72±0.0	0.60±0.0
Unlearned(+) [2]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	54.5±1.1	76.2±1.5	78.2±1.2	**78.5**±0.9	81.9±1.3	86.1±1.6	78.0±1.4	60.6±1.7	88.3±1.0	82.4±1.3	65.4±2.1	85.6±1.4	76.3±1.3
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	10.1±2.2	51.8±3.1	65.1±1.9	31.9±8.3	62.7±5.4	71.0±2.3	22.8±3.2	10.9±2.5	50.6±4.1	44.2±1.6	30.7±1.8	75.6±1.5	43.9±2.4
		Score ↑	0.50±0.02	0.50±0.03	0.47±0.02	0.60±0.03	0.50±0.02	0.50±0.02	0.64±0.02	0.55±0.01	0.59±0.02	0.57±0.02	0.50±0.02	0.49±0.02	0.53±0.02
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	55.6±1.2	75.9±1.0	77.7±1.1	**77.9**±0.8	81.0±1.4	87.3±0.9	77.3±1.3	60.4±1.1	88.6±0.9	82.4±1.0	65.7±1.5	85.7±1.1	76.3±1.1
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	15.0±2.1	38.5±1.9	57.9±2.2	40.4±1.8	48.6±1.6	56.0±2.3	11.2±1.8	2.2±0.9	51.1±2.4	37.8±2.6	23.0±0.4	55.5±2.5	36.4±1.9
		Score ↑	0.48±0.02	0.55±0.02	0.49±0.02	0.55±0.02	0.55±0.02	0.56±0.02	0.69±0.02	0.59±0.01	0.59±0.02	0.60±0.02	0.53±0.02	0.55±0.02	0.56±0.02
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	55.1±0.9	75.9±1.1	77.7±0.8	78.0±0.9	81.1±1.0	87.2±0.7	77.5±1.0	60.0±1.2	88.6±1.1	82.2±1.3	65.6±0.8	85.6±0.7	76.2±0.9
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	25.9±1.3	27.0±1.1	38.3±1.5	19.6±1.0	31.7±1.2	37.9±1.4	14.7±1.3	6.7±1.1	33.4±1.7	41.7±1.4	31.2±1.3	38.6±1.5	28.9±1.2
		Score ↑	0.44±0.02	0.60±0.02	0.56±0.01	0.65±0.02	0.62±0.02	0.63±0.01	0.68±0.01	0.56±0.02	0.66±0.01	0.58±0.02	0.50±0.01	0.62±0.01	0.59±0.02
PADA [5]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	61.8±0.8	77.8±1.4	85.1±1.1	76.5±1.0	79.7±1.7	85.4±0.5	75.9±1.7	59.1±0.4	88.9±1.1	80.2±0.8	57.7±0.8	85.6±0.8	76.1±1.0
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	37.0±2.5	75.9±0.6	80.6±2.4	54.1±2.0	74.3±0.9	80.8±0.6	48.9±1.9	34.4±0.9	79.7±2.4	48.9±1.9	41.1±2.8	83.0±2.6	61.6±1.8
		Score ↑	0.45±0.0	0.44±0.0	0.47±0.0	0.50±0.0	0.46±0.0	0.47±0.0	0.51±0.0	0.44±0.0	0.50±0.0	0.54±0.0	0.41±0.0	0.47±0.0	0.47±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	60.1±0.2	76.3±1.3	84.7±2.0	75.7±0.9	79.0±1.7	84.9±1.0	75.2±2.5	58.4±0.7	88.6±1.0	79.8±1.1	55.8±1.3	84.6±0.8	75.3±1.2
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	34.1±1.7	75.0±3.2	81.0±2.2	61.8±2.0	69.2±1.4	72.7±0.3	48.6±1.3	24.4±1.0	74.5±2.2	63.7±0.6	35.7±4.7	76.9±2.4	59.8±1.9
		Score ↑	0.45±0.0	0.44±0.0	0.47±0.0	0.47±0.0	0.47±0.0	0.49±0.0	0.51±0.0	0.47±0.0	0.51±0.0	0.49±0.0	0.41±0.0	0.48±0.0	0.47±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	59.0±0.4	76.1±1.0	84.1±2.4	74.9±1.9	78.6±1.6	84.4±1.5	74.7±2.4	57.7±0.9	88.5±1.1	80.0±1.3	53.0±2.1	84.4±0.5	74.6±1.4
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	29.8±1.4	59.6±2.9	63.4±3.7	40.7±1.5	51.2±1.8	53.4±2.5	33.9±0.2	13.5±2.5	56.5±2.1	51.3±2.0	27.0±2.9	57.2±2.4	44.8±2.2
		Score ↑	0.45±0.0	0.48±0.0	0.51±0.0	0.53±0.0	0.52±0.0	0.55±0.0	0.56±0.0	0.51±0.0	0.57±0.0	0.53±0.0	0.42±0.0	0.54±0.0	0.51±0.0
SHOT [34]	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	58.5±1.4	74.8±1.5	83.6±1.4	76.4±1.9	79.8±0.9	83.7±1.0	72.2±1.3	**63.1**±1.1	86.5±1.0	79.8±2.3	**66.6**±1.2	84.2±1.2	75.8±1.3
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	30.2±1.4	38.3±0.6	35.6±0.4	31.7±0.8	29.8±1.0	22.2±0.8	29.9±0.5	26.0±0.7	25.5±0.3	35.8±0.8	30.7±1.2	32.2±0.3	30.7±0.7
		Score ↑	0.45±0.0	0.54±0.0	0.62±0.0	0.58±0.0	0.61±0.0	0.69±0.0	0.56±0.0	0.50±0.0	0.69±0.0	0.59±0.0	0.51±0.0	0.64±0.0	0.58±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	58.3±1.1	74.0±1.1	82.4±0.9	76.5±2.5	79.7±0.8	83.5±1.4	71.1±1.2	**61.8**±0.7	86.1±1.3	79.0±2.1	**66.3**±1.0	83.7±1.0	75.2±1.2
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	30.1±2.0	37.5±1.2	35.0±2.0	30.6±2.0	30.6±2.1	22.6±1.2	28.6±0.5	24.2±0.5	25.0±0.2	33.4±0.5	29.7±2.2	31.7±1.1	29.9±1.3
		Score ↑	0.45±0.0	0.54±0.0	0.61±0.0	0.59±0.0	0.61±0.0	0.68±0.0	0.55±0.0	0.50±0.0	0.69±0.0	0.59±0.0	0.51±0.0	0.64±0.0	0.58±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	57.6±1.3	72.9±1.1	81.8±0.8	75.7±2.7	79.8±0.8	82.9±1.2	70.8±1.7	61.2±0.8	86.2±1.3	78.7±1.6	**65.7**±1.3	83.5±1.0	74.7±1.3
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	28.0±2.0	36.9±0.3	35.0±1.1	30.4±1.7	31.3±2.2	22.4±0.4	28.3±1.3	23.6±0.4	24.8±0.5	32.8±0.7	29.0±2.0	30.7±0.6	29.4±1.1
		Score ↑	0.45±0.0	0.53±0.0	0.61±0.0	0.58±0.0	0.61±0.0	0.68±0.0	0.55±0.0	0.49±0.0	0.69±0.0	0.59±0.0	0.51±0.0	0.64±0.0	0.58±0.0
Ours	T1	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	**63.1**±0.8	79.1±0.3	**86.2**±0.8	**77.4**±1.1	80.5±0.6	86.6±0.5	**76.1**±1.0	61.0±0.7	89.2±0.3	79.9±1.3	62.1±0.7	**86.6**±0.5	77.3±0.7
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
		Score ↑	**0.63**±0.0	**0.79**±0.0	**0.86**±0.0	**0.77**±0.0	**0.81**±0.0	**0.87**±0.0	**0.76**±0.0	**0.61**±0.0	**0.89**±0.0	**0.80**±0.0	**0.62**±0.0	**0.87**±0.0	**0.77**±0.0
	T2	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	**62.9**±0.7	79.0±1.2	**86.6**±0.6	**77.8**±1.4	**82.1**±0.8	86.6±0.3	**76.8**±1.5	60.5±1.1	88.8±0.4	80.7±2.5	63.2±1.4	**86.7**±0.6	**77.6**±1.0
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	2.4±0.5	**0.0**±0.0	8.9±3.2	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	11.4±8.1	1.9±1.0
		Score ↑	**0.63**±0.0	**0.79**±0.0	**0.87**±0.0	**0.76**±0.0	**0.82**±0.0	**0.80**±0.0	**0.77**±0.0	**0.61**±0.0	**0.89**±0.0	**0.81**±0.0	**0.63**±0.0	**0.78**±0.1	**0.76**±0.0
	T3	$A_{\mathcal{D}_r^\mathcal{T}}$ ↑	62.7±0.4	79.1±1.1	**86.2**±0.5	**78.2**±1.3	81.8±1.3	86.5±0.1	**76.8**±1.6	60.2±0.6	89.0±0.4	80.0±2.1	63.7±1.6	**87.1**±0.6	**77.6**±1.0
		$A_{\mathcal{D}_f^\mathcal{T}}$ ↓	4.4±6.3	**0.0**±0.0	**0.0**±0.0	3.8±1.0	**0.0**±0.0	5.0±3.0	0.1±0.1	0.6±0.5	2.0±2.8	**0.0**±0.0	3.3±1.7	7.3±5.2	2.2±1.7
		Score ↑	**0.60**±0.0	**0.79**±0.0	**0.86**±0.0	**0.75**±0.0	**0.82**±0.0	**0.82**±0.0	**0.77**±0.0	**0.60**±0.0	**0.87**±0.0	**0.80**±0.0	**0.62**±0.0	**0.81**±0.0	**0.76**±0.0

Table A.28. **C-SCADA Unlearning performance on the DomainNet dataset.** Forget classes are ${\mathcal{C}_{\mathcal{F}}}^1 = \{1, 2\}, {\mathcal{C}_{\mathcal{F}}}^2 = \{3, 4\}, {\mathcal{C}_{\mathcal{F}}}^3 = \{5, 6\}$

Method	Task	Acc.	s → p	c → s	p → c	p → r	r → s	r → c	r → p	Average
Original (SF(DA)[2] [28])	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$71.8_{\pm0.5}$	$64.9_{\pm0.0}$	$62.1_{\pm0.6}$	$78.2_{\pm0.1}$	$55.6_{\pm0.8}$	$65.8_{\pm1.6}$	$75.4_{\pm0.7}$	$67.7_{\pm0.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$37.2_{\pm1.6}$	$37.2_{\pm2.9}$	$37.3_{\pm1.6}$	$63.2_{\pm3.6}$	$17.3_{\pm2.1}$	$10.0_{\pm0.6}$	$25.5_{\pm3.5}$	$32.5_{\pm2.3}$
		Score ↑	$0.52_{\pm0.0}$	$0.47_{\pm0.0}$	$0.45_{\pm0.0}$	$0.48_{\pm0.0}$	$0.47_{\pm0.0}$	$0.60_{\pm0.0}$	$0.60_{\pm0.0}$	$0.51_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$71.8_{\pm0.5}$	$64.9_{\pm0.0}$	$62.1_{\pm0.6}$	$78.2_{\pm0.1}$	$55.6_{\pm0.8}$	$65.8_{\pm1.6}$	$75.4_{\pm0.7}$	$67.7_{\pm0.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$37.3_{\pm1.6}$	$37.4_{\pm3.0}$	$37.7_{\pm1.6}$	$63.3_{\pm3.6}$	$17.4_{\pm2.1}$	$10.1_{\pm0.6}$	$25.7_{\pm3.5}$	$32.7_{\pm2.3}$
		Score ↑	$0.52_{\pm0.0}$	$0.47_{\pm0.0}$	$0.45_{\pm0.0}$	$0.48_{\pm0.0}$	$0.47_{\pm0.0}$	$0.60_{\pm0.0}$	$0.60_{\pm0.0}$	$0.51_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$71.8_{\pm0.5}$	$64.9_{\pm0.0}$	$62.1_{\pm0.6}$	$78.2_{\pm0.1}$	$55.6_{\pm0.8}$	$65.8_{\pm1.6}$	$75.4_{\pm0.7}$	$67.7_{\pm0.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$37.4_{\pm1.6}$	$37.5_{\pm3.0}$	$37.8_{\pm1.6}$	$63.3_{\pm3.6}$	$17.4_{\pm2.1}$	$10.1_{\pm0.6}$	$25.7_{\pm3.5}$	$32.7_{\pm2.3}$
		Score ↑	$0.52_{\pm0.0}$	$0.47_{\pm0.0}$	$0.45_{\pm0.0}$	$0.48_{\pm0.0}$	$0.47_{\pm0.0}$	$0.60_{\pm0.0}$	$0.60_{\pm0.0}$	$0.51_{\pm0.0}$
Retrain	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$72.6_{\pm0.6}$	$64.3_{\pm0.8}$	$60.1_{\pm2.0}$	$79.9_{\pm0.4}$	$57.0_{\pm0.1}$	$63.6_{\pm0.4}$	$75.0_{\pm0.6}$	$67.5_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$
		Score ↑	$0.73_{\pm0.0}$	$0.64_{\pm0.0}$	$0.60_{\pm0.0}$	$0.80_{\pm0.0}$	$0.57_{\pm0.0}$	$0.64_{\pm0.0}$	$0.75_{\pm0.0}$	$0.68_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$72.6_{\pm0.6}$	$64.3_{\pm0.8}$	$60.1_{\pm2.0}$	$79.9_{\pm0.4}$	$57.0_{\pm0.1}$	$63.6_{\pm0.4}$	$75.0_{\pm0.6}$	$67.5_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$
		Score ↑	$0.73_{\pm0.0}$	$0.64_{\pm0.0}$	$0.60_{\pm0.0}$	$0.80_{\pm0.0}$	$0.57_{\pm0.0}$	$0.64_{\pm0.0}$	$0.75_{\pm0.0}$	$0.68_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$72.6_{\pm0.6}$	$64.3_{\pm0.8}$	$60.1_{\pm2.0}$	$79.9_{\pm0.4}$	$57.0_{\pm0.1}$	$63.6_{\pm0.4}$	$75.0_{\pm0.6}$	$67.5_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$	$0.0_{\pm0.0}$
		Score ↑	$0.73_{\pm0.0}$	$0.64_{\pm0.0}$	$0.60_{\pm0.0}$	$0.80_{\pm0.0}$	$0.57_{\pm0.0}$	$0.64_{\pm0.0}$	$0.75_{\pm0.0}$	$0.68_{\pm0.0}$
Finetune	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\mathbf{70.4}_{\pm1.2}$	$\mathbf{64.8}_{\pm0.9}$	$\mathbf{62.9}_{\pm0.3}$	$\mathbf{78.0}_{\pm0.5}$	$\mathbf{56.4}_{\pm0.4}$	$\mathbf{67.0}_{\pm1.0}$	$\underline{75.1}_{\pm0.8}$	$\mathbf{67.8}_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$13.1_{\pm1.2}$	$16.8_{\pm2.6}$	$28.2_{\pm4.8}$	$38.5_{\pm9.6}$	$9.9_{\pm0.2}$	$4.9_{\pm0.7}$	$14.7_{\pm1.6}$	$18.0_{\pm3.0}$
		Score ↑	$\underline{0.62}_{\pm0.0}$	$\underline{0.55}_{\pm0.0}$	$\underline{0.49}_{\pm0.0}$	$0.57_{\pm0.0}$	$\underline{0.51}_{\pm0.0}$	$\mathbf{0.64}_{\pm0.0}$	$\underline{0.66}_{\pm0.0}$	$\underline{0.58}_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\mathbf{70.4}_{\pm1.2}$	$\mathbf{64.8}_{\pm0.9}$	$\mathbf{62.9}_{\pm0.3}$	$\underline{78.0}_{\pm0.5}$	$\mathbf{56.4}_{\pm0.4}$	$\mathbf{67.0}_{\pm1.0}$	$\mathbf{75.1}_{\pm0.8}$	$\mathbf{67.8}_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$13.2_{\pm1.2}$	$\underline{16.9}_{\pm2.6}$	$28.4_{\pm4.8}$	$38.6_{\pm9.6}$	$10.0_{\pm0.2}$	$5.0_{\pm0.7}$	$14.7_{\pm1.6}$	$18.1_{\pm3.0}$
		Score ↑	$\underline{0.62}_{\pm0.0}$	$\underline{0.55}_{\pm0.0}$	$\underline{0.49}_{\pm0.0}$	$0.57_{\pm0.0}$	$\underline{0.51}_{\pm0.0}$	$\mathbf{0.64}_{\pm0.0}$	$0.66_{\pm0.0}$	$\underline{0.58}_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\mathbf{70.4}_{\pm1.2}$	$\mathbf{64.8}_{\pm0.9}$	$\mathbf{62.9}_{\pm0.3}$	$\mathbf{78.0}_{\pm0.5}$	$\mathbf{56.4}_{\pm0.4}$	$\mathbf{67.0}_{\pm1.0}$	$\mathbf{75.1}_{\pm0.8}$	$\mathbf{67.8}_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$13.2_{\pm1.2}$	$16.9_{\pm2.6}$	$28.5_{\pm4.8}$	$38.6_{\pm9.6}$	$10.0_{\pm0.2}$	$5.0_{\pm0.7}$	$14.7_{\pm1.6}$	$18.1_{\pm3.0}$
		Score ↑	$\underline{0.62}_{\pm0.0}$	$\underline{0.55}_{\pm0.0}$	$\underline{0.49}_{\pm0.0}$	$0.57_{\pm0.0}$	$\underline{0.51}_{\pm0.0}$	$\underline{0.64}_{\pm0.0}$	$\underline{0.65}_{\pm0.0}$	$\underline{0.58}_{\pm0.0}$
UNSIR [51]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$15.8_{\pm6.0}$	$7.3_{\pm5.9}$	$16.9_{\pm2.3}$	$38.9_{\pm1.3}$	$1.2_{\pm0.3}$	$24.8_{\pm3.4}$	$22.4_{\pm4.2}$	$18.2_{\pm3.3}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
		Score ↑	$0.16_{\pm0.1}$	$0.07_{\pm0.1}$	$0.17_{\pm0.0}$	$0.39_{\pm0.0}$	$0.01_{\pm0.0}$	$0.25_{\pm0.0}$	$0.22_{\pm0.0}$	$0.18_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$9.6_{\pm3.8}$	$4.5_{\pm2.3}$	$11.2_{\pm1.4}$	$25.8_{\pm2.3}$	$1.3_{\pm0.3}$	$17.6_{\pm2.7}$	$13.7_{\pm2.5}$	$11.9_{\pm2.2}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
		Score ↑	$0.10_{\pm0.0}$	$0.05_{\pm0.0}$	$0.11_{\pm0.0}$	$0.26_{\pm0.0}$	$0.01_{\pm0.0}$	$0.18_{\pm0.0}$	$0.14_{\pm0.0}$	$0.12_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$8.0_{\pm2.9}$	$3.6_{\pm1.3}$	$9.4_{\pm1.1}$	$21.1_{\pm1.2}$	$1.2_{\pm0.3}$	$15.6_{\pm2.7}$	$11.6_{\pm2.8}$	$10.1_{\pm1.8}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
		Score ↑	$0.08_{\pm0.0}$	$0.04_{\pm0.0}$	$0.09_{\pm0.0}$	$0.21_{\pm0.0}$	$0.01_{\pm0.0}$	$0.16_{\pm0.0}$	$0.12_{\pm0.0}$	$0.10_{\pm0.0}$
ZSMU [10]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\underline{69.4}_{\pm1.4}$	$61.2_{\pm1.8}$	$58.7_{\pm1.4}$	$\underline{77.9}_{\pm0.1}$	$46.3_{\pm10.1}$	$60.8_{\pm0.3}$	$71.7_{\pm2.2}$	$63.7_{\pm2.5}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$42.9_{\pm17.5}$	$16.4_{\pm16.2}$	$25.7_{\pm2.7}$	$49.2_{\pm6.0}$	$6.2_{\pm4.7}$	$14.5_{\pm4.1}$	$18.0_{\pm8.0}$	$24.7_{\pm8.5}$
		Score ↑	$0.49_{\pm0.1}$	$0.53_{\pm0.1}$	$0.47_{\pm0.0}$	$0.52_{\pm0.0}$	$0.43_{\pm0.1}$	$0.53_{\pm0.0}$	$0.61_{\pm0.1}$	$0.51_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\underline{69.3}_{\pm1.2}$	$58.1_{\pm1.7}$	$58.9_{\pm2.9}$	$\mathbf{78.5}_{\pm1.4}$	$46.0_{\pm5.9}$	$60.3_{\pm1.0}$	$70.4_{\pm3.3}$	$63.1_{\pm2.5}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$25.7_{\pm9.2}$	$26.4_{\pm5.8}$	$29.8_{\pm10.9}$	$51.3_{\pm4.7}$	$8.0_{\pm4.5}$	$11.7_{\pm5.4}$	$9.8_{\pm3.9}$	$23.2_{\pm6.3}$
		Score ↑	$0.55_{\pm0.0}$	$0.46_{\pm0.0}$	$0.46_{\pm0.1}$	$0.52_{\pm0.0}$	$0.43_{\pm0.0}$	$\underline{0.54}_{\pm0.0}$	$0.64_{\pm0.0}$	$0.51_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$68.4_{\pm1.2}$	$44.3_{\pm11.7}$	$38.7_{\pm29.7}$	$75.0_{\pm7.7}$	$45.2_{\pm5.7}$	$60.1_{\pm0.2}$	$69.4_{\pm4.4}$	$57.3_{\pm8.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$34.3_{\pm3.1}$	$\underline{13.4}_{\pm5.3}$	$15.9_{\pm14.8}$	$38.5_{\pm22.4}$	$10.8_{\pm0.6}$	$12.1_{\pm5.2}$	$14.0_{\pm4.5}$	$19.8_{\pm8.0}$
		Score ↑	$0.51_{\pm0.0}$	$0.39_{\pm0.1}$	$0.32_{\pm0.2}$	$0.55_{\pm0.1}$	$0.41_{\pm0.1}$	$0.54_{\pm0.0}$	$0.61_{\pm0.0}$	$0.47_{\pm0.1}$
Lipschitz [15]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$20.6_{\pm26.0}$	$26.1_{\pm16.3}$	$46.5_{\pm10.3}$	$74.7_{\pm1.4}$	$40.1_{\pm14.9}$	$40.0_{\pm4.3}$	$61.9_{\pm13.1}$	$44.3_{\pm12.3}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\underline{3.7}_{\pm6.4}$	$2.3_{\pm3.2}$	$16.0_{\pm12.7}$	$19.4_{\pm5.9}$	$\underline{0.3}_{\pm0.5}$	$\underline{0.2}_{\pm0.4}$	$\underline{2.7}_{\pm4.6}$	$6.4_{\pm4.8}$
		Score ↑	$0.19_{\pm0.2}$	$0.25_{\pm0.2}$	$0.40_{\pm0.1}$	$0.63_{\pm0.0}$	$0.40_{\pm0.1}$	$0.40_{\pm0.0}$	$0.60_{\pm0.1}$	$0.41_{\pm0.1}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$20.9_{\pm24.9}$	$15.1_{\pm7.4}$	$43.5_{\pm21.2}$	$73.4_{\pm2.1}$	$45.2_{\pm3.6}$	$46.1_{\pm12.8}$	$68.0_{\pm2.6}$	$44.6_{\pm10.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$1.5_{\pm2.5}$	$\mathbf{0.0}_{\pm0.0}$	$9.1_{\pm15.8}$	$9.7_{\pm16.4}$	$\underline{0.1}_{\pm0.2}$	$\underline{0.5}_{\pm0.6}$	$\underline{0.6}_{\pm1.0}$	$3.1_{\pm5.2}$
		Score ↑	$0.20_{\pm0.2}$	$0.15_{\pm0.1}$	$0.39_{\pm0.2}$	$\underline{0.68}_{\pm0.1}$	$0.45_{\pm0.0}$	$0.46_{\pm0.1}$	$\underline{0.68}_{\pm0.0}$	$0.43_{\pm0.1}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$35.9_{\pm28.0}$	$17.1_{\pm15.3}$	$48.2_{\pm19.1}$	$50.1_{\pm22.6}$	$45.3_{\pm9.0}$	$43.0_{\pm6.8}$	$63.6_{\pm5.7}$	$43.3_{\pm15.2}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$2.3_{\pm2.0}$	$\mathbf{0.0}_{\pm0.0}$	$6.4_{\pm8.6}$	$\underline{4.5}_{\pm7.8}$	$\underline{2.2}_{\pm3.6}$	$\underline{0.5}_{\pm0.6}$	$\underline{0.4}_{\pm0.7}$	$2.3_{\pm3.3}$
		Score ↑	$0.35_{\pm0.3}$	$0.17_{\pm0.2}$	$0.45_{\pm0.2}$	$0.47_{\pm0.2}$	$0.44_{\pm0.1}$	$0.43_{\pm0.1}$	$0.63_{\pm0.1}$	$0.42_{\pm0.1}$
Nabla Tau [53]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$42.8_{\pm4.3}$	$42.6_{\pm2.3}$	$43.4_{\pm1.3}$	$62.0_{\pm3.1}$	$2.6_{\pm1.3}$	$44.7_{\pm2.9}$	$47.7_{\pm1.8}$	$40.8_{\pm2.4}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$4.1_{\pm3.6}$	$\underline{0.2}_{\pm0.3}$	$\underline{0.2}_{\pm0.4}$	$\underline{3.3}_{\pm4.6}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{1.1}_{\pm1.3}$
		Score ↑	$0.41_{\pm0.1}$	$0.43_{\pm0.0}$	$0.43_{\pm0.0}$	$0.60_{\pm0.0}$	$0.03_{\pm0.0}$	$0.45_{\pm0.0}$	$0.48_{\pm0.0}$	$0.40_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$35.8_{\pm5.1}$	$35.4_{\pm2.9}$	$39.9_{\pm2.0}$	$55.9_{\pm2.8}$	$2.4_{\pm1.1}$	$40.5_{\pm2.8}$	$43.8_{\pm2.8}$	$36.3_{\pm2.8}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\underline{0.7}_{\pm1.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{5.3}_{\pm9.2}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$0.9_{\pm1.5}$
		Score ↑	$0.36_{\pm0.0}$	$0.35_{\pm0.0}$	$0.40_{\pm0.0}$	$0.53_{\pm0.1}$	$0.02_{\pm0.0}$	$0.40_{\pm0.0}$	$0.44_{\pm0.0}$	$0.36_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$34.5_{\pm4.8}$	$32.7_{\pm2.5}$	$37.8_{\pm2.2}$	$53.4_{\pm6.7}$	$2.3_{\pm1.0}$	$38.9_{\pm2.2}$	$41.1_{\pm2.0}$	$34.4_{\pm3.0}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\underline{0.9}_{\pm1.4}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{1.8}_{\pm1.6}$	$7.4_{\pm12.9}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{1.5}_{\pm2.3}$
		Score ↑	$0.34_{\pm0.0}$	$0.33_{\pm0.0}$	$0.37_{\pm0.0}$	$0.51_{\pm0.1}$	$0.02_{\pm0.0}$	$0.39_{\pm0.0}$	$0.41_{\pm0.0}$	$0.34_{\pm0.0}$
Unlearned(+) [2]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$63.8_{\pm1.0}$	$61.5_{\pm0.8}$	$77.4_{\pm1.2}$	$64.0_{\pm0.9}$	$72.6_{\pm3.0}$	$54.9_{\pm0.7}$	$69.5_{\pm0.6}$	$66.2_{\pm0.5}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$21.8_{\pm1.5}$	$23.7_{\pm2.1}$	$51.6_{\pm3.0}$	$8.6_{\pm1.1}$	$6.8_{\pm1.0}$	$5.0_{\pm1.2}$	$41.9_{\pm2.3}$	$22.8_{\pm1.8}$
		Score ↑	$0.52_{\pm0.0}$	$0.50_{\pm0.0}$	$0.51_{\pm0.0}$	$0.59_{\pm0.0}$	$0.68_{\pm0.0}$	$0.52_{\pm0.0}$	$0.49_{\pm0.0}$	$0.54_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$62.3_{\pm0.9}$	$62.1_{\pm0.7}$	$76.3_{\pm0.8}$	$64.5_{\pm1.0}$	$71.0_{\pm1.4}$	$53.5_{\pm0.9}$	$67.0_{\pm0.8}$	$65.2_{\pm0.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$19.9_{\pm1.3}$	$7.1_{\pm1.2}$	$43.4_{\pm2.0}$	$2.5_{\pm0.8}$	$2.4_{\pm0.9}$	$6.0_{\pm1.1}$	$18.1_{\pm1.0}$	$14.2_{\pm0.9}$
		Score ↑	$0.52_{\pm0.0}$	$0.58_{\pm0.0}$	$0.53_{\pm0.0}$	$0.63_{\pm0.0}$	$0.69_{\pm0.0}$	$0.50_{\pm0.0}$	$0.57_{\pm0.0}$	$0.58_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$61.5_{\pm0.8}$	$61.6_{\pm0.9}$	$76.5_{\pm1.0}$	$64.2_{\pm0.9}$	$71.3_{\pm1.7}$	$54.7_{\pm0.8}$	$65.4_{\pm0.7}$	$65.0_{\pm0.5}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$6.1_{\pm0.8}$	$2.2_{\pm0.7}$	$23.3_{\pm1.9}$	$1.8_{\pm0.7}$	$5.4_{\pm1.4}$	$5.8_{\pm0.9}$	$6.1_{\pm1.0}$	$7.2_{\pm1.1}$
		Score ↑	$0.58_{\pm0.0}$	$0.60_{\pm0.0}$	$0.62_{\pm0.0}$	$0.63_{\pm0.0}$	$0.68_{\pm0.0}$	$0.52_{\pm0.0}$	$0.62_{\pm0.0}$	$0.61_{\pm0.0}$
PADA [5]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$62.4_{\pm0.8}$	$59.9_{\pm0.6}$	$54.8_{\pm0.7}$	$76.2_{\pm0.1}$	$49.6_{\pm0.4}$	$60.1_{\pm1.7}$	$72.8_{\pm0.5}$	$62.3_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$43.5_{\pm0.4}$	$51.2_{\pm0.5}$	$52.4_{\pm1.0}$	$76.7_{\pm1.3}$	$12.3_{\pm2.0}$	$31.6_{\pm1.8}$	$24.9_{\pm4.6}$	$41.8_{\pm1.6}$
		Score ↑	$0.43_{\pm0.0}$	$0.40_{\pm0.0}$	$0.36_{\pm0.0}$	$0.43_{\pm0.0}$	$0.44_{\pm0.0}$	$0.46_{\pm0.0}$	$0.58_{\pm0.0}$	$0.44_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$40.3_{\pm6.8}$	$52.5_{\pm1.8}$	$50.7_{\pm0.7}$	$74.6_{\pm0.1}$	$45.3_{\pm1.4}$	$56.7_{\pm1.6}$	$71.4_{\pm0.3}$	$55.9_{\pm1.8}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$9.2_{\pm3.0}$	$57.0_{\pm2.4}$	$43.8_{\pm0.5}$	$69.3_{\pm2.9}$	$19.4_{\pm1.0}$	$19.4_{\pm1.1}$	$18.9_{\pm2.7}$	$33.9_{\pm1.9}$
		Score ↑	$0.37_{\pm0.1}$	$0.33_{\pm0.0}$	$0.35_{\pm0.0}$	$0.44_{\pm0.0}$	$0.38_{\pm0.0}$	$0.48_{\pm0.0}$	$0.60_{\pm0.0}$	$0.42_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$9.8_{\pm8.0}$	$37.4_{\pm3.1}$	$44.0_{\pm1.9}$	$73.6_{\pm0.1}$	$40.6_{\pm3.4}$	$54.4_{\pm1.3}$	$70.3_{\pm0.3}$	$47.2_{\pm2.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$5.6_{\pm9.6}$	$32.8_{\pm4.0}$	$31.3_{\pm3.0}$	$47.4_{\pm1.6}$	$18.1_{\pm2.4}$	$16.6_{\pm3.4}$	$25.3_{\pm0.5}$	$25.3_{\pm3.5}$
		Score ↑	$0.09_{\pm0.1}$	$0.28_{\pm0.0}$	$0.34_{\pm0.0}$	$0.50_{\pm0.0}$	$0.34_{\pm0.0}$	$0.47_{\pm0.0}$	$0.56_{\pm0.0}$	$0.37_{\pm0.0}$
SHOT [34]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$65.8_{\pm1.0}$	$47.7_{\pm1.8}$	$53.1_{\pm1.0}$	$77.4_{\pm0.2}$	$51.6_{\pm6.6}$	$57.4_{\pm0.9}$	$\mathbf{75.2}_{\pm0.4}$	$61.2_{\pm1.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$20.1_{\pm0.8}$	$20.9_{\pm0.5}$	$23.6_{\pm0.8}$	$19.9_{\pm0.3}$	$18.4_{\pm0.9}$	$19.8_{\pm0.9}$	$21.1_{\pm0.5}$	$20.5_{\pm0.7}$
		Score ↑	$0.55_{\pm0.0}$	$0.39_{\pm0.0}$	$0.43_{\pm0.0}$	$\underline{0.65}_{\pm0.0}$	$0.44_{\pm0.1}$	$0.48_{\pm0.0}$	$0.62_{\pm0.0}$	$0.51_{\pm0.0}$
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$63.8_{\pm0.8}$	$45.0_{\pm1.6}$	$49.7_{\pm3.1}$	$76.0_{\pm0.4}$	$50.4_{\pm4.1}$	$56.7_{\pm1.2}$	$73.1_{\pm0.8}$	$59.2_{\pm1.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$18.8_{\pm0.8}$	$20.1_{\pm0.4}$	$20.8_{\pm1.6}$	$19.8_{\pm0.0}$	$17.3_{\pm0.3}$	$17.9_{\pm0.4}$	$20.2_{\pm0.2}$	$19.3_{\pm0.5}$
		Score ↑	$0.54_{\pm0.0}$	$0.37_{\pm0.0}$	$0.41_{\pm0.0}$	$0.63_{\pm0.0}$	$0.43_{\pm0.0}$	$0.48_{\pm0.0}$	$0.61_{\pm0.0}$	$0.50_{\pm0.0}$
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$62.7_{\pm1.1}$	$44.1_{\pm2.5}$	$49.0_{\pm3.2}$	$74.6_{\pm0.5}$	$49.4_{\pm3.9}$	$56.4_{\pm1.4}$	$72.0_{\pm0.6}$	$58.3_{\pm1.9}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$17.9_{\pm0.6}$	$19.5_{\pm0.5}$	$19.8_{\pm1.5}$	$19.9_{\pm0.3}$	$17.0_{\pm0.5}$	$17.0_{\pm1.2}$	$19.5_{\pm0.7}$	$18.7_{\pm0.8}$
		Score ↑	$0.53_{\pm0.0}$	$0.37_{\pm0.0}$	$0.41_{\pm0.0}$	$\underline{0.62}_{\pm0.0}$	$0.42_{\pm0.0}$	$0.48_{\pm0.0}$	$0.60_{\pm0.0}$	$0.49_{\pm0.0}$
Ours		$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$68.0_{\pm0.5}$	$\underline{63.3}_{\pm1.2}$	$\underline{59.3}_{\pm0.7}$	$77.0_{\pm0.1}$	$\underline{53.3}_{\pm0.9}$	$\underline{63.4}_{\pm0.7}$	$74.6_{\pm0.1}$	$\underline{65.5}_{\pm0.6}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$
		Score ↑	$\mathbf{0.68}_{\pm0.0}$	$\mathbf{0.63}_{\pm0.0}$	$\mathbf{0.59}_{\pm0.0}$	$\mathbf{0.77}_{\pm0.0}$	$\mathbf{0.53}_{\pm0.0}$	$\underline{0.63}_{\pm0.0}$	$\mathbf{0.75}_{\pm0.0}$	$\mathbf{0.66}_{\pm0.0}$
		$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$67.3_{\pm1.3}$	$\underline{61.1}_{\pm1.7}$	$\underline{61.6}_{\pm1.4}$	$77.2_{\pm0.4}$	$\underline{54.1}_{\pm0.9}$	$\underline{63.9}_{\pm0.2}$	$\underline{73.5}_{\pm0.6}$	$\underline{65.5}_{\pm0.9}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{2.6}_{\pm3.4}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\underline{0.4}_{\pm0.5}$
		Score ↑	$\mathbf{0.67}_{\pm0.0}$	$\mathbf{0.61}_{\pm0.0}$	$\mathbf{0.60}_{\pm0.0}$	$\mathbf{0.77}_{\pm0.0}$	$\mathbf{0.54}_{\pm0.0}$	$\mathbf{0.64}_{\pm0.0}$	$\mathbf{0.73}_{\pm0.0}$	$\mathbf{0.65}_{\pm0.0}$
		$A_{\mathcal{D}_r^{\mathcal{T}}}\uparrow$	$\underline{69.2}_{\pm0.3}$	$\underline{62.7}_{\pm1.4}$	$\underline{61.0}_{\pm1.6}$	$\underline{77.2}_{\pm0.1}$	$\underline{53.8}_{\pm0.8}$	$\underline{65.6}_{\pm0.6}$	$\underline{73.9}_{\pm0.4}$	$\underline{66.2}_{\pm0.7}$
		$A_{\mathcal{D}_f^{\mathcal{T}}}\downarrow$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$4.9_{\pm4.0}$	$12.3_{\pm0.5}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$\mathbf{0.0}_{\pm0.0}$	$2.5_{\pm0.6}$
		Score ↑	$\mathbf{0.69}_{\pm0.0}$	$\mathbf{0.63}_{\pm0.0}$	$\mathbf{0.58}_{\pm0.0}$	$\mathbf{0.69}_{\pm0.0}$	$\mathbf{0.54}_{\pm0.0}$	$\mathbf{0.66}_{\pm0.0}$	$\mathbf{0.74}_{\pm0.0}$	$\mathbf{0.65}_{\pm0.0}$

Table A.29. **C-SCADA unlearning performance on the Office31 dataset.** Forget classes are $\mathcal{C_F}^1 = \{1, 2\}, \mathcal{C_F}^2 = \{3, 4\}, \mathcal{C_F}^3 = \{5, 6\}$

Method	Task	Acc.	A→D	A→W	D→A	D→W	W→A	W→D	Average
Original (SF(DA)[2] [28])	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.6±2.2	89.9±1.3	73.5±1.2	97.4±1.0	75.3±0.3	80.6±0.6	81.5±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	74.1±1.2	75.6±0.8	75.3±1.5	83.2±1.1	75.6±0.5	90.3±0.9	79.0±1.0
		Score ↑	0.42±0.0	0.51±0.0	0.42±0.0	0.53±0.0	0.43±0.0	0.42±0.0	0.46±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.6±2.2	89.9±1.3	73.5±1.2	97.4±1.0	75.3±0.3	80.6±0.6	81.5±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	77.4±1.2	77.8±0.8	76.0±1.5	85.7±1.1	76.3±0.5	94.3±0.9	81.2±1.0
		Score ↑	0.41±0.0	0.51±0.0	0.42±0.0	0.52±0.0	0.43±0.0	0.41±0.0	0.45±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.6±2.2	89.9±1.3	73.5±1.2	97.4±1.0	75.3±0.3	80.6±0.6	81.5±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	78.5±1.2	78.6±0.8	76.3±1.5	86.5±1.1	76.6±0.5	95.7±0.9	82.0±1.0
		Score ↑	0.41±0.0	0.50±0.0	0.42±0.0	0.52±0.0	0.43±0.0	0.41±0.0	0.45±0.0
Retrain	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	75.5±2.4	94.8±1.0	67.8±9.2	97.4±1.0	72.1±1.1	81.0±0.0	81.5±2.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.76±0.0	0.95±0.0	0.68±0.1	0.97±0.0	0.72±0.0	0.81±0.0	0.81±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	75.5±2.4	94.8±1.0	67.8±9.2	97.4±1.0	72.1±1.1	81.0±0.0	81.5±2.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.76±0.0	0.95±0.0	0.68±0.1	0.97±0.0	0.72±0.0	0.81±0.0	0.81±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	75.5±2.4	94.8±1.0	67.8±9.2	97.4±1.0	72.1±1.1	81.0±0.0	81.5±2.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
		Score ↑	0.76±0.0	0.95±0.0	0.68±0.1	0.97±0.0	0.72±0.0	0.81±0.0	0.81±0.0
Finetune	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.5±3.2	88.9±1.0	71.7±1.7	97.1±1.0	75.5±1.2	**81.0**±0.0	80.8±1.3
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	70.1±1.6	67.1±1.5	61.1±4.8	78.3±0.5	59.7±2.0	86.6±1.2	70.5±1.9
		Score ↑	0.41±0.0	0.53±0.0	0.45±0.0	0.54±0.0	0.47±0.0	0.43±0.0	0.47±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.5±3.2	88.9±1.0	71.7±1.7	97.1±1.0	75.5±1.2	**81.0**±0.0	80.8±1.3
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	73.2±1.6	69.1±1.6	61.8±4.8	80.7±0.5	60.2±2.0	90.4±1.2	72.6±2.0
		Score ↑	0.41±0.0	0.53±0.0	0.44±0.0	0.54±0.0	0.47±0.0	0.43±0.0	0.47±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.5±3.2	88.9±1.0	71.7±1.7	97.1±1.0	75.5±1.2	**81.0**±0.0	80.8±1.3
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	74.3±1.6	69.8±1.6	62.0±4.8	81.5±0.5	60.4±2.0	91.8±1.2	73.3±2.0
		Score ↑	0.40±0.0	0.52±0.0	0.44±0.0	0.54±0.0	0.47±0.0	0.42±0.0	0.47±0.0
UNSIR [51]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	65.0±2.2	75.5±9.5	45.0±7.8	89.4±3.8	50.2±1.8	79.7±1.0	67.5±4.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	38.3±5.7	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	72.2±0.0	18.4±1.0
		Score ↑	0.47±0.0	0.75±0.1	0.45±0.1	0.89±0.0	0.50±0.0	0.46±0.0	0.59±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	62.5±2.2	73.1±7.0	36.2±8.3	65.1±23.0	34.8±4.5	72.1±8.2	57.3±8.8
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	11.0±7.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	50.0±18.0	10.2±4.2
		Score ↑	0.56±0.0	0.73±0.1	0.36±0.1	0.65±0.2	0.35±0.0	0.48±0.0	0.52±0.1
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	61.6±3.3	66.1±8.9	31.4±7.5	61.5±22.5	28.3±6.4	70.0±11.1	53.1±10.0
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	7.2±5.9	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	34.3±26.8	6.9±5.4
		Score ↑	0.58±0.0	0.66±0.1	0.31±0.1	0.61±0.2	0.28±0.1	0.53±0.0	0.50±0.1
ZSMU [10]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.0±2.8	97.1±1.2	72.1±4.5	90.7±2.0	80.2±1.4	72.1±2.7	80.7±2.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	81.8±2.6	84.5±0.0	79.6±3.7	81.6±2.7	81.5±0.0	77.1±6.1	81.0±2.5
		Score ↑	0.40±0.0	0.53±0.0	0.40±0.0	0.50±0.0	0.44±0.0	0.41±0.0	0.45±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.0±2.3	97.4±1.6	70.9±3.8	91.2±2.5	80.2±1.4	50.7±35.1	77.1±7.8
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	65.9±12.0	86.8±2.9	74.4±2.1	75.6±4.2	88.2±0.7	49.5±42.9	73.4±10.8
		Score ↑	0.44±0.0	0.52±0.0	0.41±0.0	0.52±0.0	0.43±0.0	0.31±0.2	0.44±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	71.3±1.8	97.4±1.6	70.5±4.5	85.5±9.2	**80.2**±1.4	51.5±34.5	76.1±8.9
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	73.2±6.4	83.4±1.7	73.5±2.2	71.1±18.6	88.5±2.4	52.4±45.2	73.7±12.7
		Score ↑	0.41±0.0	0.53±0.0	0.41±0.0	0.50±0.0	0.43±0.0	0.31±0.2	0.43±0.0
Lipschitz [15]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	45.4±26.5	96.6±0.5	67.1±4.4	84.0±5.6	78.9±2.7	67.6±6.7	73.3±7.7
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**±0.0	26.4±4.3	43.2±7.8	32.2±35.0	40.8±36.5	25.8±22.3	28.1±17.7
		Score ↑	0.45±0.3	0.76±0.0	0.47±0.0	0.66±0.1	0.59±0.2	0.54±0.0	0.58±0.1
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	58.6±9.1	95.9±3.8	62.0±14.2	89.1±2.8	**81.0**±0.0	67.4±6.6	75.7±6.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**±0.0	25.2±13.6	9.0±8.1	35.7±31.4	29.3±8.6	14.2±7.3	18.9±11.5
		Score ↑	0.59±0.1	0.77±0.1	0.56±0.1	0.68±0.2	0.63±0.0	0.59±0.0	0.64±0.1
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	63.8±5.9	97.4±1.2	**72.6**±1.4	82.7±13.9	70.0±19.0	65.7±13.7	75.4±9.2
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	24.2±13.3	12.8±11.5	34.3±16.5	18.9±16.4	45.8±29.4	24.7±26.3	26.8±18.9
		Score ↑	0.51±0.0	0.87±0.1	0.55±0.1	0.69±0.0	0.48±0.1	0.53±0.1	0.61±0.1
Nabla Tau [53]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	66.0±3.2	96.1±2.3	67.1±2.5	86.5±3.1	80.2±1.4	67.5±1.6	77.2±2.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	1.0±1.8	6.9±7.9	22.8±18.5	0.6±1.0	75.3±2.1	**0.0**±0.0	17.8±5.2
		Score ↑	0.65±0.0	0.90±0.1	0.56±0.1	0.86±0.0	0.46±0.0	0.67±0.0	0.68±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	63.6±3.4	95.9±1.2	66.7±1.9	86.3±3.2	80.2±1.4	64.9±1.6	76.3±2.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	3.1±2.7	1.9±3.3	9.7±10.4	**0.0**±0.0	59.8±11.7	5.7±5.0	13.4±5.5
		Score ↑	0.62±0.0	**0.94**±0.0	0.61±0.1	0.86±0.0	0.50±0.0	0.62±0.0	0.69±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	64.1±3.2	95.9±1.2	66.7±1.9	86.3±3.2	**80.2**±1.4	65.4±1.0	76.4±2.0
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	7.0±7.7	0.4±0.8	3.1±2.7	**0.0**±0.0	46.4±14.0	7.8±9.7	10.8±5.8
		Score ↑	0.60±0.0	**0.95**±0.0	0.65±0.0	0.86±0.0	0.55±0.0	0.61±0.1	0.70±0.0
Unlearned(+) [2]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.9±0.9	91.5±2.1	73.5±1.5	97.7±1.2	75.6±0.8	81.0±1.0	81.7±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	83.3±1.3	82.8±1.5	79.1±2.1	84.5±1.2	76.7±1.4	81.5±1.0	81.3±1.4
		Score ↑	0.39±0.0	0.50±0.0	0.41±0.0	0.53±0.0	0.43±0.0	0.45±0.0	0.45±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	75.9±0.8	92.2±2.0	73.0±1.6	97.7±1.0	76.0±1.2	81.0±1.3	82.6±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	59.8±1.4	69.8±1.7	41.3±1.9	81.4±1.1	40.9±2.0	79.3±1.4	62.1±1.6
		Score ↑	0.47±0.0	0.54±0.0	0.52±0.0	0.54±0.0	0.54±0.0	0.45±0.0	0.51±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	75.9±1.0	92.2±2.3	72.6±1.7	97.7±1.1	75.6±0.9	81.0±1.2	82.5±1.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	58.9±1.1	61.1±1.5	27.5±1.7	69.4±1.2	48.9±1.4	77.6±1.0	57.2±1.3
		Score ↑	0.48±0.0	0.57±0.0	0.57±0.0	0.58±0.0	0.51±0.0	0.46±0.0	0.53±0.0
PADA [5]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**72.4**±0.5	**97.4**±1.2	70.9±3.3	91.2±0.5	**80.6**±0.8	72.4±0.6	80.8±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	87.3±0.3	84.5±0.0	83.3±0.0	84.5±0.0	82.7±1.0	84.5±1.0	84.5±0.4
		Score ↑	0.39±0.0	0.53±0.0	0.39±0.0	0.49±0.0	0.44±0.0	0.39±0.0	0.44±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	71.9±2.3	97.2±0.9	69.2±2.7	91.7±0.4	79.7±1.3	72.0±0.7	80.3±1.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	56.4±5.3	88.4±1.2	72.0±2.5	77.1±2.9	88.2±0.7	55.9±5.7	73.0±3.0
		Score ↑	0.46±0.0	0.52±0.0	0.40±0.0	0.52±0.0	0.42±0.0	0.46±0.0	0.46±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.0±1.6	97.4±1.2	69.2±1.4	91.7±0.4	79.7±1.3	72.1±0.7	80.4±1.1
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	49.9±3.0	80.9±1.1	70.7±4.2	71.3±0.7	88.1±2.7	62.2±7.3	70.5±3.2
		Score ↑	0.48±0.0	0.54±0.0	0.41±0.0	0.54±0.0	0.42±0.0	0.45±0.0	0.47±0.0
SHOT [34]	T1	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	71.1±0.7	**97.4**±2.0	70.5±4.5	89.4±0.9	**80.6**±0.8	70.1±1.3	79.8±1.7
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	41.2±1.3	66.5±2.7	73.4±2.7	62.4±3.1	92.2±1.6	41.9±0.8	62.9±2.0
		Score ↑	0.50±0.0	0.59±0.0	0.41±0.0	0.55±0.0	0.42±0.0	0.49±0.0	0.49±0.0
	T2	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.6±0.9	**97.7**±1.6	**72.1**±5.8	89.1±1.3	80.2±1.4	69.3±0.5	79.8±1.9
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	39.5±2.9	66.4±2.2	73.1±1.6	62.2±1.2	90.8±1.8	40.8±0.7	62.1±1.7
		Score ↑	0.51±0.0	0.59±0.0	0.42±0.0	0.55±0.0	0.42±0.0	0.49±0.0	0.50±0.0
	T3	$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	70.1±0.9	**97.7**±1.6	72.1±5.8	89.1±1.3	**80.2**±1.4	69.2±0.3	79.7±1.9
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	37.1±2.4	65.3±1.0	71.8±2.4	61.9±1.6	89.1±1.5	40.4±0.6	60.9±1.6
		Score ↑	0.51±0.0	0.59±0.0	0.42±0.0	0.55±0.0	0.42±0.0	0.49±0.0	0.50±0.0
Ours		$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	72.2±2.1	91.2±1.3	**72.5**±1.1	**97.4**±1.0	74.4±0.6	80.2±1.2	**81.3**±1.2
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0	**0.0**±0.0
		Score ↑	**0.72**±0.0	**0.91**±0.0	**0.73**±0.0	**0.97**±0.0	**0.74**±0.0	**0.80**±0.0	**0.81**±0.0
		$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**72.6**±2.6	91.2±1.3	71.8±1.4	**97.4**±1.0	74.6±0.6	80.2±1.2	**81.3**±1.3
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**0.0**±0.0	**0.0**±0.0	3.1±4.4	**0.0**±0.0	2.0±2.8	**0.0**±0.0	**0.8**±1.2
		Score ↑	**0.73**±0.0	0.91±0.0	**0.70**±0.0	**0.97**±0.0	**0.73**±0.0	**0.80**±0.0	**0.81**±0.0
		$A_{\mathcal{D}_r^{\mathcal{T}}}$ ↑	**72.6**±2.6	90.9±1.5	71.8±1.7	**97.4**±1.0	74.6±0.7	80.2±1.2	**81.2**±1.4
		$A_{\mathcal{D}_f^{\mathcal{T}}}$ ↓	**1.9**±2.6	**0.0**±0.0	5.0±2.6	**0.0**±0.0	15.7±3.4	**0.0**±0.0	**3.8**±1.5
		Score ↑	**0.71**±0.0	0.91±0.0	**0.68**±0.0	**0.97**±0.0	**0.65**±0.0	**0.80**±0.0	**0.79**±0.0

Table A.30. **Results for Multi-Class SCADA Unlearning on Openset Domain Adaptation on OfficeHome.** Implemented with SHOT-ODA as the open-set domain adaptation loss

Method	Metric	A $\rightarrow$ C	A $\rightarrow$ P	A $\rightarrow$ R	C $\rightarrow$ A	C $\rightarrow$ P	C $\rightarrow$ R	P $\rightarrow$ A	P $\rightarrow$ C	P $\rightarrow$ R	R $\rightarrow$ A	R $\rightarrow$ C	R $\rightarrow$ P	Average
Original (SHOT [34])	OS Acc. $\uparrow$	$54.5_{\pm 1.0}$	$69.8_{\pm 1.6}$	$75.9_{\pm 0.9}$	$70.3_{\pm 1.3}$	$72.5_{\pm 0.5}$	$76.9_{\pm 0.5}$	$67.7_{\pm 2.6}$	$49.1_{\pm 1.8}$	$75.3_{\pm 0.8}$	$72.8_{\pm 1.5}$	$55.4_{\pm 0.8}$	$76.5_{\pm 1.3}$	$68.0_{\pm 1.2}$
	OS* Acc. $\uparrow$	$60.2_{\pm 1.2}$	$77.0_{\pm 1.8}$	$81.8_{\pm 0.9}$	$75.5_{\pm 1.3}$	$79.9_{\pm 0.6}$	$83.0_{\pm 0.5}$	$72.7_{\pm 2.8}$	$54.3_{\pm 1.9}$	$81.3_{\pm 0.9}$	$78.2_{\pm 1.7}$	$61.2_{\pm 0.8}$	$84.3_{\pm 1.4}$	$74.1_{\pm 1.3}$
	Forget Acc. $\downarrow$	$43.3_{\pm 0.5}$	$85.6_{\pm 1.2}$	$86.2_{\pm 0.7}$	$68.2_{\pm 2.7}$	$78.9_{\pm 1.3}$	$79.1_{\pm 1.5}$	$65.4_{\pm 0.5}$	$36.5_{\pm 4.0}$	$83.6_{\pm 1.5}$	$70.7_{\pm 1.1}$	$45.0_{\pm 3.3}$	$89.9_{\pm 1.6}$	$69.4_{\pm 1.6}$
	Score $\uparrow$	$0.38_{\pm 0.0}$	$0.38_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.36_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.38_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.40_{\pm 0.0}$
Retrain	OS Acc. $\uparrow$	$51.7_{\pm 5.5}$	$67.9_{\pm 1.2}$	$75.2_{\pm 1.2}$	$70.3_{\pm 0.9}$	$71.5_{\pm 1.2}$	$76.4_{\pm 0.8}$	$70.7_{\pm 2.8}$	$51.5_{\pm 3.7}$	$77.2_{\pm 0.4}$	$74.2_{\pm 0.2}$	$58.0_{\pm 1.5}$	$77.4_{\pm 2.4}$	$68.5_{\pm 1.8}$
	OS* Acc. $\uparrow$	$57.2_{\pm 6.1}$	$74.9_{\pm 1.3}$	$81.1_{\pm 1.2}$	$75.6_{\pm 1.0}$	$78.9_{\pm 1.3}$	$82.4_{\pm 0.9}$	$75.9_{\pm 3.1}$	$56.9_{\pm 4.1}$	$83.2_{\pm 0.5}$	$79.7_{\pm 0.2}$	$64.1_{\pm 1.7}$	$85.4_{\pm 2.7}$	$74.6_{\pm 2.0}$
	Forget Acc. $\downarrow$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$	$0.0_{\pm 0.0}$
	Score $\uparrow$	$0.52_{\pm 0.1}$	$0.68_{\pm 0.0}$	$0.75_{\pm 0.0}$	$0.70_{\pm 0.0}$	$0.72_{\pm 0.0}$	$0.76_{\pm 0.0}$	$0.71_{\pm 0.0}$	$0.52_{\pm 0.0}$	$0.77_{\pm 0.0}$	$0.74_{\pm 0.0}$	$0.58_{\pm 0.0}$	$0.77_{\pm 0.0}$	$0.69_{\pm 0.0}$
Finetune	OS Acc. $\uparrow$	$\mathbf{51.1}_{\pm 2.5}$	$\mathbf{68.9}_{\pm 1.1}$	$\mathbf{75.7}_{\pm 0.8}$	$\mathbf{69.6}_{\pm 1.8}$	$\mathbf{71.7}_{\pm 0.9}$	$\mathbf{76.3}_{\pm 0.5}$	$\mathbf{67.1}_{\pm 2.3}$	$\mathbf{49.5}_{\pm 1.4}$	$\mathbf{75.3}_{\pm 0.5}$	$\mathbf{72.7}_{\pm 1.1}$	$\mathbf{53.1}_{\pm 0.8}$	$\mathbf{75.9}_{\pm 1.3}$	$\mathbf{67.2}_{\pm 1.2}$
	OS* Acc. $\uparrow$	$56.5_{\pm 2.7}$	$76.0_{\pm 1.1}$	$81.7_{\pm 0.8}$	$74.8_{\pm 1.9}$	$79.1_{\pm 1.0}$	$82.3_{\pm 0.6}$	$72.1_{\pm 2.4}$	$54.7_{\pm 1.5}$	$81.2_{\pm 0.5}$	$78.0_{\pm 1.2}$	$58.7_{\pm 0.9}$	$83.7_{\pm 1.5}$	$73.2_{\pm 1.3}$
	Forget Acc. $\downarrow$	$38.5_{\pm 4.4}$	$85.1_{\pm 1.8}$	$83.6_{\pm 0.2}$	$63.3_{\pm 2.3}$	$77.4_{\pm 2.5}$	$73.6_{\pm 1.1}$	$64.5_{\pm 1.9}$	$35.3_{\pm 5.6}$	$82.1_{\pm 1.4}$	$67.0_{\pm 1.9}$	$39.8_{\pm 2.3}$	$87.3_{\pm 1.5}$	$66.5_{\pm 2.3}$
	Score $\uparrow$	$0.37_{\pm 0.0}$	$0.37_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.37_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.44_{\pm 0.0}$	$0.38_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.40_{\pm 0.0}$
UNSIR [51]	OS Acc. $\uparrow$	$26.2_{\pm 12.5}$	$57.3_{\pm 3.8}$	$57.6_{\pm 1.2}$	$12.2_{\pm 9.0}$	$57.0_{\pm 2.4}$	$59.0_{\pm 0.7}$	$16.1_{\pm 6.1}$	$26.0_{\pm 7.9}$	$57.4_{\pm 3.0}$	$26.8_{\pm 9.4}$	$21.6_{\pm 10.1}$	$62.1_{\pm 2.7}$	$39.9_{\pm 5.7}$
	OS* Acc. $\uparrow$	$29.0_{\pm 13.9}$	$63.2_{\pm 4.1}$	$62.1_{\pm 1.3}$	$13.1_{\pm 9.7}$	$62.9_{\pm 2.6}$	$63.6_{\pm 0.7}$	$17.2_{\pm 6.6}$	$28.7_{\pm 8.7}$	$61.9_{\pm 3.3}$	$28.8_{\pm 10.0}$	$23.9_{\pm 11.1}$	$68.5_{\pm 3.0}$	$43.6_{\pm 6.2}$
	Forget Acc. $\downarrow$	$\mathbf{0.3}_{\pm 0.5}$	$\mathbf{0.2}_{\pm 0.3}$	$\mathbf{0.5}_{\pm 0.9}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.6}_{\pm 1.1}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.1}_{\pm 0.2}$
	Score $\uparrow$	$0.26_{\pm 0.1}$	$0.57_{\pm 0.0}$	$0.57_{\pm 0.0}$	$0.12_{\pm 0.1}$	$0.57_{\pm 0.0}$	$0.59_{\pm 0.0}$	$0.16_{\pm 0.1}$	$0.26_{\pm 0.1}$	$0.57_{\pm 0.0}$	$0.27_{\pm 0.1}$	$0.22_{\pm 0.1}$	$0.62_{\pm 0.0}$	$0.40_{\pm 0.1}$
ZSMU [10]	OS Acc. $\uparrow$	$45.6_{\pm 3.1}$	$64.2_{\pm 2.3}$	$72.2_{\pm 3.5}$	$71.7_{\pm 1.5}$	$55.2_{\pm 25.6}$	$65.0_{\pm 15.1}$	$59.3_{\pm 3.2}$	$40.1_{\pm 2.6}$	$72.7_{\pm 2.4}$	$70.2_{\pm 3.3}$	$49.4_{\pm 3.8}$	$73.1_{\pm 2.3}$	$61.6_{\pm 5.7}$
	OS* Acc. $\uparrow$	$49.7_{\pm 3.8}$	$70.0_{\pm 1.8}$	$77.7_{\pm 3.5}$	$76.6_{\pm 1.2}$	$60.6_{\pm 28.0}$	$69.8_{\pm 16.1}$	$63.6_{\pm 3.4}$	$44.1_{\pm 3.0}$	$78.4_{\pm 2.5}$	$75.4_{\pm 3.5}$	$54.3_{\pm 3.9}$	$80.1_{\pm 1.8}$	$66.7_{\pm 6.0}$
	Forget Acc. $\downarrow$	$24.9_{\pm 7.1}$	$72.4_{\pm 5.7}$	$78.6_{\pm 3.3}$	$57.4_{\pm 6.4}$	$44.8_{\pm 40.1}$	$49.2_{\pm 40.9}$	$44.1_{\pm 4.2}$	$25.4_{\pm 5.8}$	$72.9_{\pm 3.8}$	$61.7_{\pm 11.4}$	$29.4_{\pm 4.7}$	$81.9_{\pm 3.8}$	$53.6_{\pm 11.4}$
	Score $\uparrow$	$0.37_{\pm 0.0}$	$0.37_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.46_{\pm 0.0}$	$0.37_{\pm 0.1}$	$0.44_{\pm 0.0}$	$0.41_{\pm 0.0}$	$0.32_{\pm 0.0}$	$0.42_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.38_{\pm 0.0}$	$0.40_{\pm 0.0}$	$0.40_{\pm 0.0}$
Lipschitz [15]	OS Acc. $\uparrow$	$1.4_{\pm 0.7}$	$23.3_{\pm 19.8}$	$15.3_{\pm 23.9}$	$8.6_{\pm 11.6}$	$9.6_{\pm 6.4}$	$25.7_{\pm 22.6}$	$1.5_{\pm 0.6}$	$2.2_{\pm 1.1}$	$18.4_{\pm 29.5}$	$2.2_{\pm 1.1}$	$5.2_{\pm 6.1}$	$52.2_{\pm 14.0}$	$13.8_{\pm 11.4}$
	OS* Acc. $\uparrow$	$1.6_{\pm 0.8}$	$25.7_{\pm 21.8}$	$16.5_{\pm 25.8}$	$9.3_{\pm 12.5}$	$10.6_{\pm 7.2}$	$27.7_{\pm 24.4}$	$1.6_{\pm 0.6}$	$2.5_{\pm 1.2}$	$19.8_{\pm 31.8}$	$2.4_{\pm 1.2}$	$5.7_{\pm 6.7}$	$57.6_{\pm 15.4}$	$15.1_{\pm 12.4}$
	Forget Acc. $\downarrow$	$\mathbf{0.0}_{\pm 0.0}$	$5.9_{\pm 9.0}$	$4.6_{\pm 8.0}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$19.5_{\pm 23.4}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$11.9_{\pm 20.7}$	$\mathbf{0.0}_{\pm 0.0}$	$\mathbf{0.0}_{\pm 0.0}$	$32.7_{\pm 12.2}$	$6.2_{\pm 6.1}$
	Score $\uparrow$	$0.01_{\pm 0.0}$	$0.21_{\pm 0.2}$	$0.14_{\pm 0.2}$	$0.09_{\pm 0.1}$	$0.10_{\pm 0.1}$	$0.20_{\pm 0.2}$	$0.02_{\pm 0.0}$	$0.02_{\pm 0.0}$	$0.14_{\pm 0.2}$	$0.02_{\pm 0.0}$	$0.05_{\pm 0.1}$	$0.39_{\pm 0.1}$	$0.12_{\pm 0.1}$
Nabla Tau [53]	OS Acc. $\uparrow$	$49.7_{\pm 1.6}$	$68.1_{\pm 0.3}$	$69.7_{\pm 2.4}$	$68.1_{\pm 1.6}$	$68.3_{\pm 2.3}$	$72.4_{\pm 0.8}$	$62.1_{\pm 3.4}$	$44.7_{\pm 0.9}$	$71.5_{\pm 2.6}$	$68.4_{\pm 3.5}$	$48.7_{\pm 2.0}$	$73.7_{\pm 1.2}$	$63.8_{\pm 1.9}$
	OS* Acc. $\uparrow$	$54.9_{\pm 1.8}$	$75.2_{\pm 0.3}$	$75.2_{\pm 2.6}$	$73.2_{\pm 1.6}$	$75.4_{\pm 2.6}$	$78.1_{\pm 0.9}$	$66.7_{\pm 3.6}$	$49.4_{\pm 0.9}$	$77.1_{\pm 2.8}$	$73.5_{\pm 3.8}$	$53.8_{\pm 2.2}$	$81.3_{\pm 1.3}$	$69.5_{\pm 2.0}$
	Forget Acc. $\downarrow$	$6.3_{\pm 3.0}$	$12.7_{\pm 8.7}$	$5.7_{\pm 5.9}$	$0.6_{\pm 1.1}$	$\mathbf{0.0}_{\pm 0.0}$	$0.3_{\pm 0.2}$	$5.2_{\pm 1.9}$	$5.2_{\pm 1.9}$	$\mathbf{0.0}_{\pm 0.0}$	$8.7_{\pm 7.0}$	$2.7_{\pm 1.2}$	$11.8_{\pm 4.7}$	$4.9_{\pm 3.0}$
	Score $\uparrow$	$0.47_{\pm 0.0}$	$0.61_{\pm 0.1}$	$0.66_{\pm 0.0}$	$0.68_{\pm 0.0}$	$0.68_{\pm 0.0}$	$0.72_{\pm 0.0}$	$0.59_{\pm 0.0}$	$0.43_{\pm 0.0}$	$0.71_{\pm 0.0}$	$0.63_{\pm 0.1}$	$0.47_{\pm 0.0}$	$0.66_{\pm 0.0}$	$0.61_{\pm 0.0}$
Ours	OS Acc. $\uparrow$	$46.2_{\pm 3.4}$	$67.3_{\pm 1.6}$	$73.3_{\pm 0.8}$	$68.8_{\pm 3.5}$	$68.3_{\pm 4.8}$	$75.8_{\pm 0.3}$	$64.9_{\pm 3.0}$	$27.3_{\pm 8.9}$	$72.7_{\pm 1.0}$	$71.7_{\pm 1.5}$	$45.4_{\pm 1.9}$	$74.2_{\pm 3.9}$	$63.0_{\pm 2.9}$
	OS* Acc. $\uparrow$	$\mathbf{51.1}_{\pm 3.7}$	$74.2_{\pm 1.7}$	$79.1_{\pm 0.9}$	$73.9_{\pm 3.7}$	$75.4_{\pm 5.2}$	$81.8_{\pm 0.3}$	$69.8_{\pm 3.3}$	$30.2_{\pm 9.8}$	$78.4_{\pm 1.1}$	$77.0_{\pm 1.6}$	$50.2_{\pm 2.1}$	$81.8_{\pm 4.4}$	$68.6_{\pm 3.1}$
	Forget Acc. $\downarrow$	$\underline{0.5}_{\pm 0.5}$	$\underline{2.3}_{\pm 3.2}$	$12.0_{\pm 12.9}$	$\mathbf{0.0}_{\pm 0.0}$	$\underline{3.1}_{\pm 2.8}$	$\underline{3.0}_{\pm 3.6}$	$\underline{1.6}_{\pm 1.4}$	$\mathbf{0.0}_{\pm 0.0}$	$15.6_{\pm 14.3}$	$\underline{0.9}_{\pm 1.6}$	$\underline{0.2}_{\pm 0.3}$	$\underline{7.6}_{\pm 10.0}$	$\underline{3.9}_{\pm 4.2}$
	Score $\uparrow$	$\underline{0.46}_{\pm 0.0}$	$\mathbf{0.66}_{\pm 0.0}$	$\mathbf{0.66}_{\pm 0.1}$	$\mathbf{0.69}_{\pm 0.0}$	$\underline{0.66}_{\pm 0.1}$	$\mathbf{0.74}_{\pm 0.0}$	$\mathbf{0.64}_{\pm 0.0}$	$0.27_{\pm 0.1}$	$\underline{0.63}_{\pm 0.1}$	$\mathbf{0.71}_{\pm 0.0}$	$\underline{0.45}_{\pm 0.0}$	$\mathbf{0.69}_{\pm 0.0}$	$\mathbf{0.61}_{\pm 0.0}$