LEARNING WHERE TO ADAPT: SPARSE DELTA LEARNING FOR ONLINE TEST-TIME ADAPTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern learning systems are often trained in controlled or simulated environments but deployed under shifting real-world conditions, creating a persistent training–deployment gap. Online Test-Time Adaptation (OTTA) addresses this gap by updating models from unlabeled target data, yet existing methods mainly focus on *how* to adapt while largely fixing *where* updates occur. We introduce **Sparse Delta Learning (SDL)**, which formulates OTTA as budgeted support learning over structured parameter groups. SDL separates support discovery from parameter adaptation. In the first stage, a dense functional update is coupled with a group-sparse structural variable, and Group Split-LBI traces an inverse-scale-space path to discover a data-dependent support under an explicit group budget. In the second stage, SDL initializes adaptation with the Stage-1 change inside the discovered support, applies the original TTA optimizer only within that support, and writes the refined update back to the persistent model. This formulation leaves the base adaptation objective unchanged and extends to fully connected, convolutional, and Transformer parameters through architecture-specific group definitions. On VisDA-C with DeiT-S/16 and SHOT, SDL improves final offline mean per-class accuracy by 14.29 percentage points over the strongest budget-matched sparse baseline while updating only 0.0213% of the model.

1 INTRODUCTION

Modern learning systems are increasingly deployed beyond the conditions represented in their training data. The mismatch between training and deployment is especially visible in embodied intelligence, particularly in robotics, where simulation is widely used to train and test robotic systems before real-world deployment. Differences between simulated and real environments in sensing, dynamics, actuation, and system modeling create a persistent sim-to-real gap (Aljalbout et al., 2026). More broadly, deployed models must cope with evolving, unlabeled target distributions without repeated retraining. Test-Time Adaptation (TTA) addresses this problem by updating a pre-trained model directly from unlabeled test data during inference (Liang et al., 2025).

We focus on online test-time adaptation (OTTA), where target batches arrive sequentially and each batch is observed once (Liang et al., 2025). After adapting to the current batch, the updated model is carried forward to the next one. Each adaptation step therefore affects both the current prediction and the starting point for all subsequent updates. Adaptation decisions in OTTA are thus cumulative rather than isolated.

Much of the TTA literature has focused on *how* the model should adapt. Existing methods design adaptation signals through entropy minimization (Wang et al., 2021), information maximization and pseudo-labeling (Liang et al., 2020), or more conservative entropy-based objectives (Zhang et al., 2025). For online streams, restoration and regularization are further used to stabilize accumulated updates (Wang et al., 2022; Niu et al., 2022). Recent work has also begun to control the update scope explicitly, for example through parameter-selective masks or dedicated adaptation modules (Tian & Lyu, 2024; Kim et al., 2026). Together, these works show that the update scope is itself an important design axis.

Existing selective approaches typically define the update scope through a specific module, masking criterion, or architectural design. In contrast, we formulate the adaptation support as a budget-constrained variable learned from each incoming target batch. In OTTA, both the incoming data and

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

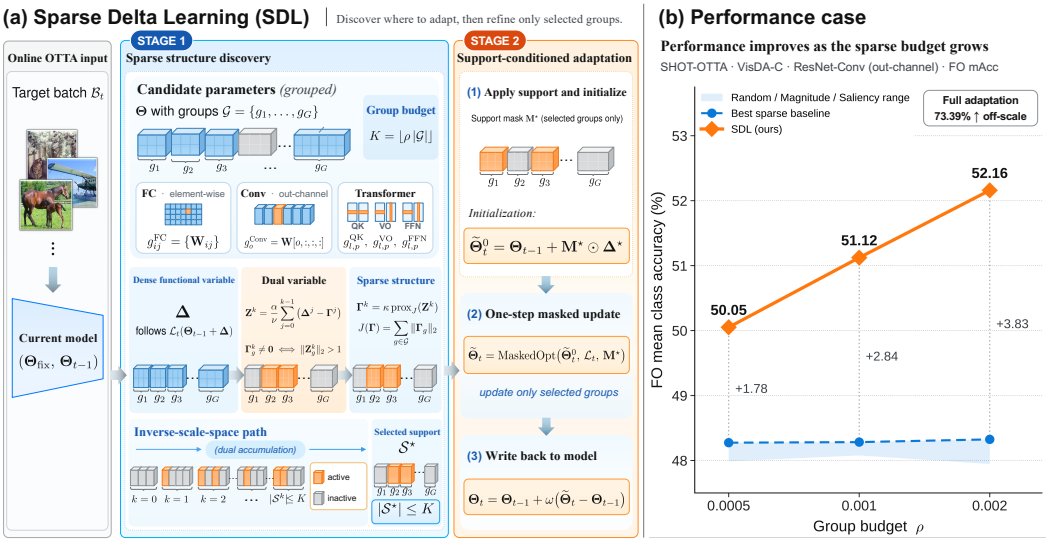

Figure 1: **Overview of Sparse Delta Learning (SDL). (a)** For each target batch, SDL follows an inverse-scale-space path to select a budgeted support of parameter groups, then adapts only these groups and writes the sparse update back to the online model. The grouping applies to fully connected, convolutional, and Transformer layers. **(b)** In a representative SHOT experiment on VisDA-C with the ResNet Conv track, SDL consistently outperforms the strongest sparse baseline, with the margin increasing from 1.78 to 3.83 percentage points as the group budget ρ grows.

the model state evolve along the stream, so the useful update directions may also change from one batch to the next. This motivates learning the adaptation support together with each test-time update.

The choice of adaptation support becomes especially important under a limited update budget. Even with the same adaptation objective and the same number of active parameter groups, performance can depend strongly on where that budget is placed. We therefore argue that OTTA should learn not only *how* to update a model, but also *where* those updates should occur. We formulate this as a *budgeted support-learning* problem over structured parameter groups.

We introduce **Sparse Delta Learning (SDL)** to solve this problem. As illustrated in Figure 1, SDL separates support discovery from parameter adaptation. In Stage 1, Group Split Linearized Bregman Iteration (Group Split-LBI) traces an inverse-scale-space path over candidate parameter groups and discovers a support under an explicit group budget (Burger et al., 2006; Osher et al., 2016; Fu et al., 2020). In Stage 2, SDL applies the original TTA optimizer only within the discovered support and writes the resulting update back to the persistent online model. The base adaptation objective remains unchanged; SDL determines where that objective is allowed to act. The same optimization framework applies across architectures, with scalar connections as groups for fully connected layers, output channels for convolutional layers, and QK, VO, and FFN coordinates for Transformers.

Our contributions are threefold:

- We formulate OTTA as a *budgeted support-learning* problem, where the adaptation support is inferred from the current target data rather than fixed in advance.
- We propose a two-stage SDL framework that combines Group Split-LBI support discovery with support-conditioned adaptation. A unified formulation applies to fully connected, convolutional, and Transformer layers through architecture-specific parameter groups.
- We evaluate SDL with SHOT and COME on Office-31 and VisDA-C using ResNet and DeiT backbones. Across architectures and budgets, SDL consistently outperforms budget-matched random, magnitude, and saliency selection under SHOT, while remaining competitive under COME. On VisDA-C with DeiT-S/16 and SHOT at $\rho = 0.001$, SDL improves final offline mean per-class accuracy by 14.29 percentage points over the strongest sparse baseline while updating only 0.0213% of the model.

2 Related Work

Test-Time Adaptation. TTA adapts a pre-trained model using unlabeled target data during inference. Representative methods use entropy minimization (Wang et al., 2021), information maximization and pseudo-labeling (Liang et al., 2020), or conservative entropy-based objectives (Zhang et al., 2025). For online adaptation, CoTTA uses stochastic restoration (Wang et al., 2022), EATA combines sample selection with Fisher-based regularization (Niu et al., 2022), and NCTTA combines entropy minimization with contrastive feature–classifier alignment (Chen et al., 2026). This line of work primarily studies the adaptation objective and its stability. SDL instead learns the parameter support on which the chosen objective acts.

Selective Adaptation. Recent work explicitly controls the test-time update scope. PSMT uses Fisher information to construct selective parameter masks (Tian & Lyu, 2024), while Buffer TTA introduces dedicated trainable buffer layers and keeps the backbone fixed (Kim et al., 2026). SDL instead treats the adaptation support as a per-batch, group-structured variable learned from target data under an explicit budget.

Linearized Bregman Iteration and Structured Sparsity. Linearized Bregman Iteration and inverse-scale-space methods trace sparse-to-dense solution paths (Osher et al., 2005; Burger et al., 2006; Osher et al., 2016). DessiLBI extends this principle to structured sparsity in deep networks through variable splitting and differential-inclusion dynamics (Fu et al., 2020). SPP further applies differential-inclusion paths to structured pruning of pre-trained Transformers (Ding et al., 2025). SDL repurposes this solution-path principle for per-batch adaptation support discovery rather than model pruning.

3 Methodology

Sparse Delta Learning (SDL) learns *where* a model should adapt at test time. For each target batch, SDL first traces a sparse structure-discovery path over candidate parameter groups, and then adapts the model only within the selected support. The optimization framework is shared across architectures; different models only require different definitions of the parameter groups.

3.1 Problem Setup

We consider online test-time adaptation (OTTA), where unlabeled target data arrive sequentially as $\{\mathcal{B}_t\}_{t=1}^{T}$. Each batch is observed once. The model adapts on $\mathcal{B}_t$, predicts with the updated parameters, and carries the resulting model to the next batch.

We divide the model parameters into frozen parameters $\mathbf{\Theta}_{\text{fix}}$ and candidate parameters $\mathbf{\Theta}$. Only the candidate parameters are adapted. Let $\mathbf{\Theta}_{t-1}$ denote their persistent value before processing $\mathcal{B}_t$. SDL directly uses the unsupervised loss of the underlying TTA method, $\mathcal{L}_t(\mathbf{\Theta}) = \mathcal{L}_{\text{adp}}(\mathcal{B}_t; \mathbf{\Theta}_{\text{fix}}, \mathbf{\Theta})$. SDL therefore changes the *adaptation support*, rather than replacing the base adaptation objective.

Throughout this section, bold uppercase symbols denote parameter tensors or collections of parameter tensors.

3.2 Sparse Delta Learning

Budgeted group-sparse adaptation. Let $\mathcal{G} = \{g_1, \ldots, g_G\}$ be a disjoint partition of the candidate parameters. A group is the basic unit of structure selection. For batch $\mathcal{B}_t$, SDL seeks an adaptation support $\mathcal{S}_t \subseteq \mathcal{G}$. Given a group budget ratio ρ, the maximum number of active groups is $K = \lfloor \rho |\mathcal{G}| \rfloor$. When group sizes differ, ρ is not the fraction of scalar parameters.

Let $\mathbf{\Delta}$ denote an additive change to the candidate parameters. For a support $\mathcal{S}$, $\Pi_{\mathcal{S}}(\mathbf{\Delta})$ keeps entries belonging to groups in $\mathcal{S}$ and sets all remaining entries to zero. The ideal budgeted adaptation problem is

$$\min_{\mathbf{\Delta},\, \mathcal{S} \subseteq \mathcal{G}} \mathcal{L}_t\left(\mathbf{\Theta}_{t-1} + \Pi_{\mathcal{S}}(\mathbf{\Delta})\right) \quad \text{s.t.} \quad |\mathcal{S}| \le K. \tag{1}$$

Directly searching over $\mathcal{S}$ is combinatorial. SDL instead generates a data-dependent sparse regularization path and uses the budget to select an early point on that path.

Variable splitting for structure discovery. A direct sparse penalty on $\mathbf{\Delta}$ would make the same variable responsible for both fitting the target batch and satisfying the structural constraint. SDL separates these two roles. We introduce a dense functional variable $\mathbf{\Delta}$ and a structural variable $\mathbf{\Gamma}$ with the same parameter layout, and couple them through

$$\overline{\mathcal{L}}_t\left(\mathbf{\Delta}, \mathbf{\Gamma}\right) = \mathcal{L}_t\left(\mathbf{\Theta}_{t-1} + \mathbf{\Delta}\right) + \frac{1}{2\nu}\left\|\mathbf{\Delta} - \mathbf{\Gamma}\right\|_F^2, \tag{2}$$

where $\nu > 0$ controls how tightly the functional and structural variables are coupled.

We impose the group penalty $J(\mathbf{\Gamma}) = \sum_{g\in\mathcal{G}} \|\mathbf{\Gamma}_g\|_2$ only on $\mathbf{\Gamma}$; for singleton groups it reduces to the ℓ_1 norm. Thus, $\mathbf{\Delta}$ remains free to follow the adaptation loss, while $\mathbf{\Gamma}$ records which groups have entered the sparse structure.

Inverse-scale-space structure path. Following inverse-scale-space dynamics (Burger et al., 2006; Osher et al., 2016; Huang et al., 2016; Fu et al., 2020), we introduce a dual variable $\mathbf{Z}$ and consider the following dynamics starting from $\mathbf{\Delta}(0) = \mathbf{\Gamma}(0) = \mathbf{Z}(0) = \mathbf{0}$,

$$\begin{aligned}
\frac{\dot{\mathbf{\Delta}}(s)}{\kappa} &= -\nabla_{\mathbf{\Delta}}\overline{\mathcal{L}}_t\left(\mathbf{\Delta}(s), \mathbf{\Gamma}(s)\right), \\
\dot{\mathbf{Z}}(s) &= -\nabla_{\mathbf{\Gamma}}\overline{\mathcal{L}}_t\left(\mathbf{\Delta}(s), \mathbf{\Gamma}(s)\right), \\
\mathbf{Z}(s) &\in \partial\left[J(\mathbf{\Gamma}(s)) + \frac{1}{2\kappa}\left\|\mathbf{\Gamma}(s)\right\|_F^2\right],
\end{aligned} \tag{3}$$

where s is algorithmic time and $\kappa > 0$ balances the ℓ_1 and ℓ_2 regularization in the elastic-net penalty. The dynamics couple the updates of the dense parameter $\mathbf{\Delta}$ and the sparse parameter $\mathbf{\Gamma}$. The dense parameter follows a gradient descent flow and fits the data better as the iteration proceeds. Through the ℓ_2 coupling term in Eq. (2), it drives the sparse parameter to select important groups. Specifically, starting from $\mathbf{Z}(0) = \mathbf{0}$, the dual variables of these groups surpass the threshold value 1, and the corresponding groups of $\mathbf{\Gamma}$ become nonzero. Groups selected earlier are more important. This is called the *inverse scale space* property, which has been established in Osher et al. (2005; 2016) and discussed in Fu et al. (2020) and Ding et al. (2025).

Group Split-LBI discretization. SDL solves Eq. (3) with Split LBI (Huang et al., 2016), a discrete approximation of these dynamics. With step size α, one step is

$$\begin{aligned}
\mathbf{\Delta}^{k+1} &= \mathbf{\Delta}^k - \alpha\kappa\,\nabla_{\mathbf{\Delta}}\overline{\mathcal{L}}_t(\mathbf{\Delta}^k, \mathbf{\Gamma}^k), \\
\mathbf{Z}^{k+1} &= \mathbf{Z}^k - \alpha\,\nabla_{\mathbf{\Gamma}}\overline{\mathcal{L}}_t(\mathbf{\Delta}^k, \mathbf{\Gamma}^k), \\
\mathbf{\Gamma}^{k+1} &= \kappa\,\mathrm{prox}_J(\mathbf{Z}^{k+1}).
\end{aligned} \tag{4}$$

This discretization has been used to explore structural sparsity when training deep networks (Fu et al., 2020) and to prune pre-trained Transformers (Ding et al., 2025). We use it in the online test-time setting, where the path restarts for every target batch and stops at the budget. The proximal operator acts on each group as $[\mathrm{prox}_J(\mathbf{Z})]_g = (1 - 1/\|\mathbf{Z}_g\|_2)_+\mathbf{Z}_g$, where $(x)_+ = \max(x, 0)$. Hence $\mathbf{\Gamma}_g^k \neq \mathbf{0}$ if and only if $\|\mathbf{Z}_g^k\|_2 > 1$. Because $\mathbf{Z}^0 = \mathbf{0}$, Eq. (4) gives $\mathbf{Z}_g^k = (\alpha/\nu)\sum_{j=0}^{k-1}(\mathbf{\Delta}_g^j - \mathbf{\Gamma}_g^j)$. Group activation therefore depends on the gap between $\mathbf{\Delta}$ and $\mathbf{\Gamma}$ accumulated along the Stage-1 trajectory. Gaps that keep the same direction reinforce one another and reach the threshold earlier, whereas gaps that change direction partially cancel. This accumulation distinguishes SDL from selecting groups by a single gradient or a static weight magnitude.

Path-based support discovery. Unlike a single-step selector based on one local statistic, SDL exposes a sequence of candidate supports as task-driven changes accumulate along the Group Split-LBI trajectory. Because the path is recomputed for each target batch from the current persistent model, the selected support can vary with the evolving stream. The group budget then selects a feasible point on this path, while the base TTA objective determines how parameters within the selected support are updated.

Budget-constrained early stopping. After iteration k, SDL extracts the current support as

$$\mathcal{S}^k = \left\{g \in \mathcal{G} : \|\mathbf{\Gamma}_g^k\|_2 > 0\right\}. \tag{5}$$

At each iteration, SDL accepts the current state whenever $|\mathcal{S}^k| \leq K$. If one step produces $|\mathcal{S}^k| > K$, that step is discarded and SDL returns the latest feasible state. The selected support therefore satisfies $|\mathcal{S}^\star| \leq K$. The final support may contain fewer than K groups if the path overshoots.

Algorithm 1 Sparse Delta Learning for Online Test-Time Adaptation

Input: target stream $\{\mathcal{B}_t\}_{t=1}^{T}$; group partition $\mathcal{G}$; budget ratio ρ; $\alpha, \kappa, \nu, \omega$; maximum Stage-1 steps $k_{\max}$
$K \leftarrow \lfloor \rho |\mathcal{G}| \rfloor$
for $t = 1$ **to** T **do**
  $\mathbf{\Delta}^0 \leftarrow \mathbf{0}$, $\mathbf{\Gamma}^0 \leftarrow \mathbf{0}$, $\mathbf{Z}^0 \leftarrow \mathbf{0}$, $\mathcal{S}^0 \leftarrow \emptyset$
  $(\mathbf{\Delta}^\star, \mathcal{S}^\star) \leftarrow (\mathbf{0}, \emptyset)$
  $k \leftarrow 0$
  while $k < k_{\max}$ **and** $|\mathcal{S}^k| < K$ **do**
    Update $\mathbf{\Delta}^{k+1}$, $\mathbf{Z}^{k+1}$, and $\mathbf{\Gamma}^{k+1}$ using Eq. (4)
    $\mathcal{S}^{k+1} \leftarrow \{g \in \mathcal{G} : \|\mathbf{\Gamma}_g^{k+1}\|_2 > 0\}$
    if $|\mathcal{S}^{k+1}| \leq K$ **then**
      $(\mathbf{\Delta}^\star, \mathcal{S}^\star) \leftarrow (\mathbf{\Delta}^{k+1}, \mathcal{S}^{k+1})$
    end if
    $k \leftarrow k + 1$
  end while
  $\mathbf{M}^\star \leftarrow \mathrm{Mask}(\mathcal{S}^\star)$
  $\widetilde{\mathbf{\Theta}}_t^0 \leftarrow \mathbf{\Theta}_{t-1} + \mathbf{M}^\star \odot \mathbf{\Delta}^\star$
  $\widetilde{\mathbf{\Theta}}_t \leftarrow \mathrm{MaskedOpt}(\widetilde{\mathbf{\Theta}}_t^0, \mathcal{L}_t, \mathbf{M}^\star)$
  $\mathbf{\Theta}_t \leftarrow \mathbf{\Theta}_{t-1} + \omega(\widetilde{\mathbf{\Theta}}_t - \mathbf{\Theta}_{t-1})$
  Predict $\mathcal{B}_t$ using $(\mathbf{\Theta}_{\mathrm{fix}}, \mathbf{\Theta}_t)$
end for

Support-conditioned adaptation. Let $\mathbf{M}^\star$ be the coordinate mask induced by $\mathcal{S}^\star$. Stage 2 starts from $\widetilde{\mathbf{\Theta}}_t^0 = \mathbf{\Theta}_{t-1} + \mathbf{M}^\star \odot \mathbf{\Delta}^\star$, where $\odot$ denotes element-wise multiplication. Only the Stage-1 change inside the discovered support is transferred.

SDL then takes one step of the base adaptation optimizer on $\mathcal{L}_t$. This step, denoted $\mathrm{MaskedOpt}$ in Algorithm 1, updates only the coordinates inside $\mathbf{M}^\star$ and leaves all others fixed. Stage 1 therefore selects the adaptation structure, while Stage 2 estimates the task update conditional on that structure.

The refined change is written back as $\mathbf{\Theta}_t = \mathbf{\Theta}_{t-1} + \omega \left(\widetilde{\mathbf{\Theta}}_t - \mathbf{\Theta}_{t-1} \right)$, with $\omega \in (0, 1]$. The local variables $\mathbf{\Delta}$, $\mathbf{\Gamma}$, and $\mathbf{Z}$ are then discarded, while $\mathbf{\Theta}_t$ is carried to the next target batch. Support discovery is thus local to each batch, whereas the adapted model evolves across the target stream.

Separating support discovery from adaptation. Stage 1 determines which parameter groups should participate in the current adaptation step, rather than replacing the optimizer of the underlying TTA method. Once the support is fixed, Stage 2 returns to the original adaptation objective and optimizer within that restricted parameter space. This separation isolates *where* adaptation occurs from *how* the selected parameters are updated, allowing SDL to be combined with different TTA objectives without changing their optimization rule.

Hyperparameters. The Stage-1 path is controlled by the step size α, the elastic-net balance κ, and the coupling strength ν. The budget ρ sets structural complexity, and ω sets how much of each sparse update the persistent model retains.

3.3 Architecture-Specific Instantiations

Grouping principle. Across architectures, we define groups around parameters that jointly implement one computational unit. This choice makes the support correspond to a structured adaptation decision rather than an arbitrary collection of individual weights. The granularity of that decision follows the architecture: scalar connections for fully connected layers, output feature channels for convolutional layers, and coupled attention or hidden coordinates for Transformers. The SDL optimization itself is unchanged across these cases, and only the partition $\mathcal{G}$ depends on the architecture.

Fully connected layers. Consider a fully connected layer $\mathbf{y} = \mathbf{W}\mathbf{x} + \mathbf{b}$. Each scalar weight and bias forms one group:

$$g_{ij}^{\mathrm{FC}} = \{\mathbf{W}_{ij}\}, \qquad g_i^b = \{\mathbf{b}_i\}. \tag{6}$$

The group penalty therefore reduces to element-wise ℓ_1 sparsity, and SDL selects individual FC connections for adaptation.

Convolutional layers. Consider a convolutional kernel $\mathbf{W} \in \mathbb{R}^{C_{\text{out}} \times C_{\text{in}} \times k_h \times k_w}$. We define one group for each output channel:

$$g_o^{\text{Conv}} = \mathbf{W}[o, :, :, :], \qquad o = 1, \ldots, C_{\text{out}}. \tag{7}$$

All kernel weights that produce one output feature map are therefore selected together, and the budget ρ is measured over output-channel groups.

Transformer layers. Consider Transformer block l with hidden width d. Let $\mathbf{W}_Q^{(l)}$, $\mathbf{W}_K^{(l)}$, $\mathbf{W}_V^{(l)}$, and $\mathbf{W}_O^{(l)}$ denote its attention projections. For each attention coordinate p, we define

$$g_{l,p}^{\text{QK}} = \left\{ \mathbf{W}_Q^{(l)}[p, :], \mathbf{W}_K^{(l)}[p, :] \right\}, \qquad g_{l,p}^{\text{VO}} = \left\{ \mathbf{W}_V^{(l)}[p, :], \mathbf{W}_O^{(l)}[:, p] \right\}. \tag{8}$$

The QK group collects query and key parameters associated with the same attention coordinate, while the VO group couples the parameters that write to and read from the same value coordinate.

For the FFN $\mathbf{h} = \mathbf{W}_2^{(l)} \phi(\mathbf{W}_1^{(l)} \mathbf{x})$, we define one group for each hidden coordinate p:

$$g_{l,p}^{\text{FFN}} = \left\{ \mathbf{W}_1^{(l)}[p, :], \mathbf{W}_2^{(l)}[:, p] \right\}. \tag{9}$$

This group contains the parameters that write to and read from one FFN hidden coordinate.

4 Experiments

4.1 Experimental Setup

Datasets and metrics. We evaluate SDL on Office-31 (Saenko et al., 2010) and VisDA-C (Peng et al., 2017). Office-31 contains 31 object categories from three domains: Amazon (A), DSLR (D), and Webcam (W). We evaluate all six directed domain transfers and report their mean classification accuracy. VisDA-C contains 12 categories and defines a synthetic-to-real transfer from the training split to the validation split. Because its class frequencies are uneven, we use mean per-class accuracy (mAcc) as the primary metric.

Online protocol. We use one-pass online test-time adaptation (OTTA). Target batches are processed sequentially in a fixed stream, each batch is used for one adaptation event, and the updated model is carried to the next batch. We report two complementary metrics. *Post-update accuracy* (PU) evaluates each batch immediately after its adaptation step, without any additional parameter update. *Final offline accuracy* (FO) evaluates the complete target set using the final model after the stream. PU measures performance along the online trajectory, while FO measures the model retained after adaptation.

Architectures and budgets. For ResNet-50 (He et al., 2016) on Office-31 and ResNet-101 on VisDA-C, we apply SDL either to the 256-dimensional bottleneck FC layer or to output-channel groups in the final residual stage. For DeiT-S/16 (Touvron et al., 2021), we apply SDL to the QK, VO, and FFN groups in the last three Transformer blocks. These settings cover three adaptation granularities: element-wise FC parameters, convolutional output-channel groups, and structured Transformer groups. All three use the same group-budget ratios $\rho \in \{0.0005, 0.001, 0.002\}$.

Compared methods. We instantiate SDL with two adaptation objectives, SHOT (Liang et al., 2020) and COME (Zhang et al., 2025). For each objective and architecture, we include three reference models: Source-only, Full-dense, and Candidate-dense. Full-dense updates all parameters except the classifier head, whereas Candidate-dense updates all parameters in the candidate space used by SDL. We further compare SDL with three sparse selectors under the same group budget. Random selects groups uniformly at random. Magnitude ranks groups by source-weight magnitude. Saliency ranks groups by the current weight–gradient saliency score. These comparisons isolate the effect of learning the adaptation support under a fixed budget. We also report Tent (Wang et al., 2021), EATA (Niu et al., 2022), and NCTTA (Chen et al., 2026) as standalone OTTA references.

Controlled comparison design. The dense and sparse controls separate several sources of improvement. Candidate-dense tests whether restricting adaptation to the SDL candidate space is sufficient without support selection. Random, Magnitude, and Saliency additionally hold both the

Table 1: Office-31 accuracy (%) for DeiT-S/16 at $\rho = 0.001$ ($K = 6$). Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	82.73	79.25	63.51	94.21	66.06	99.60	80.89	82.13	79.37	64.29	94.09	66.21	99.60	80.95
SHOT	Full-dense	84.34	83.40	63.65	94.84	67.94	99.80	82.33	85.74	85.66	63.51	95.60	68.23	99.80	83.09
SHOT	Candidate-dense	83.13	80.38	64.00	94.21	66.74	99.60	81.34	83.94	81.13	65.35	94.21	67.38	100.00	82.00
SHOT	Random	82.73	79.25	63.52	**94.21**	66.06	**99.60**	80.90	82.13	79.37	**64.30**	**94.09**	66.23	99.60	80.95
SHOT	Magnitude	82.73	79.25	63.51	**94.21**	66.06	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	99.60	80.95
SHOT	Saliency	82.73	79.25	63.54	**94.21**	66.10	**99.60**	80.91	82.13	79.37	64.29	**94.09**	66.24	99.60	80.95
SHOT	**SDL**	**83.33**	**79.75**	**63.79**	93.84	**66.28**	**99.60**	**81.10**	**83.94**	**80.13**	64.04	93.96	**66.84**	**99.80**	**81.45**
COME	Full-dense	82.73	79.50	63.54	94.21	65.99	99.60	80.93	82.53	79.50	64.50	94.09	66.42	99.60	81.11
COME	Candidate-dense	82.73	79.25	63.54	94.21	66.06	99.60	80.90	82.13	79.37	64.29	94.09	66.31	99.60	80.96
COME	Random	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	Magnitude	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	Saliency	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	**SDL**	**82.73**	**79.25**	**63.58**	**94.21**	66.03	**99.60**	**80.90**	**82.13**	**79.37**	64.25	**94.09**	**66.28**	**99.60**	**80.95**
Standard OTTA references															
Tent	DeiT full	83.13	81.26	57.88	94.59	65.55	99.80	80.37	83.73	82.26	51.08	94.72	63.40	99.80	79.17
Tent	Transformer	83.13	79.62	62.18	94.09	65.31	99.60	80.65	82.73	79.62	61.66	93.84	64.89	100.00	80.46
EATA	DeiT full	82.93	79.25	63.60	94.21	66.09	99.60	80.95	83.13	79.37	64.32	94.09	66.35	99.60	81.14
EATA	Transformer	82.73	79.25	63.53	94.21	66.12	99.60	80.91	82.13	79.37	64.29	94.09	66.24	99.60	80.95
NCTTA	DeiT full	83.53	79.37	63.64	94.09	65.09	99.60	80.89	83.33	80.88	62.62	94.47	64.22	100.00	80.92
NCTTA	Transformer	83.13	79.50	63.53	93.96	65.34	99.60	80.84	82.53	80.25	64.04	94.09	65.39	99.60	80.98

candidate space and group budget fixed, so their comparison with SDL directly tests how the support is chosen under the same structural budget. Full-dense provides a reference for adaptation without the candidate-space sparsity constraint.

Experimental configurations. For each objective, architecture, and budget, we use a fixed SDL configuration throughout evaluation. On Office-31, the same configuration is shared across all six directed transfers for a given objective, architecture, and budget. All methods use the same source model, target stream, and preprocessing. Candidate-dense and the sparse selectors use the same candidate space as SDL, while each baseline retains its prescribed optimizer and method-specific adaptation settings. Appendix A provides the complete experimental protocol and hyperparameter settings.

4.2 Main Results

We focus the main comparison on DeiT-S/16 at $\rho = 0.001$, corresponding to at most $K = 6$ active groups. The same architecture and support budget are used on Office-31 and VisDA-C. Complete results across architectures and budgets are reported in Appendices B and C.

Office-31: Learned support improves persistent adaptation under SHOT. As shown in Table 1, the fixed sparse selectors under SHOT remain close to Source-only, whereas SDL consistently improves under the same six-group budget. The advantage is modest in PU but becomes larger in FO: SDL exceeds the strongest sparse selector by 0.19 and 0.50 percentage points, respectively. Since FO evaluates the final model after the complete target stream, the larger FO separation is consistent with the discovered support remaining useful after repeated online write-back rather than benefiting only the current batch.

This pattern is also visible across individual transfers. SDL improves four of the six domain pairs under both PU and FO, indicating that the average gain is not driven by one favorable transfer. However, Candidate-dense and Full-dense remain stronger on average. Office-31 therefore represents a relatively mild setting: learning where to adapt improves over fixed sparse selection, but a six-group support does not capture all useful SHOT adaptation available to the dense models.

COME provides a useful contrast in the same table. Candidate-dense and all sparse variants remain almost unchanged from Source-only, and changing the selector produces essentially no separation. This suggests that support selection is not the main bottleneck in this regime. When adaptation within the candidate space itself produces little effective change, identifying a different sparse subset cannot yield a substantial improvement.

Table 2: VisDA-C class accuracy (%) for DeiT-S/16 at $\rho = 0.001$ ($K = 6$). PU and FO are reported side by side. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	75.86	8.81	29.25	74.90	60.29	24.00	66.20	16.27	73.95	40.90	82.60	10.85	46.99	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.12
SHOT	Full-dense	79.98	43.02	62.17	40.24	66.42	51.13	66.99	23.15	80.37	55.55	78.21	35.04	56.86	80.94	71.77	71.28	36.33	75.14	75.86	59.54	43.62	76.52	46.34	76.94	36.73	62.58
SHOT	Candidate-dense	81.65	9.24	49.47	60.34	64.93	20.24	69.69	15.93	77.14	48.01	85.25	22.39	50.35	84.48	10.39	62.62	50.51	70.58	18.51	70.86	17.43	79.07	56.73	85.27	28.88	52.94
SHOT	Random	75.84	8.81	29.27	74.86	60.29	23.98	66.23	**16.28**	73.96	40.89	82.61	10.87	46.99	75.67	9.13	29.44	74.81	60.01	26.31	66.35	**16.05**	73.83	40.64	82.69	10.81	47.15
SHOT	Magnitude	75.84	8.81	29.25	74.91	60.29	24.00	66.22	16.27	73.93	40.86	82.60	10.85	46.98	75.64	9.12	29.42	74.87	59.97	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.13
SHOT	Saliency	76.03	8.78	29.36	**75.02**	60.24	24.05	66.44	15.97	73.97	40.95	82.29	10.87	47.00	76.08	9.06	29.89	**75.08**	59.94	26.31	66.86	15.75	73.80	40.46	82.32	11.01	47.21
SHOT	**SDL**	**83.60**	**34.47**	**61.83**	45.07	**86.57**	**29.30**	**69.88**	1.77	**81.29**	**54.23**	**84.56**	**29.60**	**55.18**	**79.46**	**69.38**	**68.91**	40.30	**89.58**	**79.23**	**69.63**	0.23	**81.47**	**42.83**	**85.65**	**31.31**	**61.50**
COME	Full-dense	77.56	8.29	23.60	80.49	59.22	21.98	65.94	14.60	75.62	38.76	78.28	8.35	46.06	78.61	7.57	17.53	85.87	58.17	21.16	65.77	12.65	77.18	35.73	72.99	6.40	44.97
COME	Candidate-dense	76.17	8.84	28.29	75.78	60.31	23.95	66.01	16.12	74.24	40.55	82.22	10.31	46.90	76.00	9.09	28.04	76.48	60.16	26.31	65.96	15.93	74.32	39.98	82.08	9.68	47.00
COME	Random	75.85	**8.81**	29.21	74.91	**60.29**	23.98	66.22	**16.30**	73.95	40.89	**82.60**	10.83	46.99	**75.67**	**9.13**	**29.42**	74.86	**60.02**	**26.31**	66.33	**16.03**	**73.84**	**40.64**	82.67	10.67	**47.13**
COME	Magnitude	75.84	**8.81**	**29.23**	74.92	**60.29**	**24.05**	66.20	**16.30**	**73.97**	**40.90**	**82.60**	**10.85**	**47.00**	**75.67**	9.12	**29.42**	74.87	59.99	**26.31**	66.30	16.00	**73.84**	**40.64**	**82.70**	**10.71**	**47.13**
COME	Saliency	**75.86**	**8.81**	29.19	74.92	**60.29**	**24.05**	66.23	16.27	73.95	**40.90**	**82.60**	**10.85**	46.99	**75.67**	9.12	29.38	74.90	59.97	**26.31**	66.32	16.02	73.82	40.60	82.65	10.69	47.12
COME	**SDL**	**75.86**	**8.81**	29.19	**74.94**	**60.29**	**24.05**	**66.25**	16.28	73.95	**40.90**	**82.60**	10.83	**47.00**	**75.67**	9.12	29.36	**74.94**	59.97	**26.31**	**66.39**	**16.03**	73.80	40.60	82.63	10.67	47.12
Standard OTTA references																											
Tent	DeiT full	28.39	1.29	2.64	98.26	10.34	2.31	10.66	1.07	12.84	6.09	9.82	0.78	15.38	0.55	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	8.38
Tent	Transformer	55.32	4.72	8.68	95.26	29.70	9.20	39.49	3.12	45.42	14.86	45.87	2.81	29.54	22.46	0.83	0.51	99.97	4.22	1.69	16.70	0.00	7.80	0.88	9.66	0.32	13.75
EATA	DeiT full	75.86	8.81	29.23	74.90	60.29	24.05	66.24	16.28	73.99	40.86	82.60	10.87	47.00	75.67	9.12	29.42	74.86	60.01	26.31	66.32	16.02	73.82	40.64	82.65	10.72	47.13
EATA	Transformer	75.86	8.81	29.25	74.90	60.29	24.05	66.22	16.28	73.95	40.90	82.60	10.85	47.00	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.84	40.64	82.65	10.71	47.12
NCTTA	DeiT full	49.34	2.96	7.16	95.99	24.13	7.18	25.97	2.93	28.75	14.16	23.44	1.98	23.67	20.57	0.03	0.09	99.98	1.45	0.34	1.26	0.00	0.75	0.66	0.02	0.00	10.43
NCTTA	Transformer	74.03	8.00	21.41	85.34	53.51	19.95	59.02	10.38	70.70	34.11	74.88	8.35	43.31	70.38	6.79	12.43	94.89	42.38	14.41	50.50	3.65	62.72	25.16	64.97	5.12	37.78

VisDA-C: Support selection becomes critical under SHOT. Table 2 shows a much stronger separation on VisDA-C. The ordering of the controlled variants is particularly informative: the fixed sparse selectors remain near Source-only, Candidate-dense provides a clear improvement, and SDL closes most of the remaining gap to Full-dense. Thus, sparsity alone is not sufficient, but simply updating the entire candidate space is not sufficient either. The key difference is which structured groups are selected for adaptation.

Under SHOT, SDL exceeds the strongest budget-matched sparse selector by 14.29 points in FO mAcc and Candidate-dense by 8.56 points, while remaining only 1.08 points below Full-dense. The gap between SDL and the fixed sparse selectors is also substantially larger in FO than in PU. As on Office-31, this indicates that the benefit of the learned support is not limited to immediate post-update predictions and becomes more pronounced in the model retained after the online stream.

The class-wise results in Table 2 show that the aggregate improvement is broad rather than driven by a single category. SDL improves 10 of the 12 classes over the strongest fixed sparse selector under both PU and FO, with large gains on several initially difficult categories. The exceptions are Car and Person, where performance drops sharply. The learned support therefore changes how the limited adaptation capacity is distributed across classes: substantial gains on most categories coexist with clear trade-offs on a smaller subset.

COME again shows little separation among candidate-space variants. Candidate-dense remains close to Source-only, and SDL is essentially tied with the fixed sparse selectors. Taken together with Office-31, the two objectives reveal a consistent interaction: support discovery is most useful when the base adaptation objective provides informative update directions within the candidate space. SDL can then concentrate those directions into a small support, but support selection alone cannot compensate for an ineffective update signal.

At this operating point, SDL activates at most 4,608 scalars, corresponding to only 0.0213% of DeiT-S/16. Complete parameter footprints and runtime measurements are reported in Appendix A.3.

4.3 Hyperparameter Sensitivity

We study hyperparameter sensitivity using SHOT on Office-31 with FC parameters at $\rho = 0.001$. Each sweep varies one hyperparameter while keeping the remaining settings fixed. Table 3 reports FO accuracy together with the maximum Stage-1 discovery length across the six transfers.

Discovery dynamics. Table 3 reveals a clear relationship between how the sparse path is traversed and the quality of the resulting support. Increasing α from 0.10 to 0.20 roughly halves the maximum discovery length, but reduces FO by more than two points. A larger step therefore accelerates traversal at the cost of a coarser path, which can miss useful intermediate supports.

The effects of κ and ν show that longer discovery is not necessarily more effective. Increasing κ from 1.0 to 1.5 leaves FO almost unchanged while requiring more iterations, whereas $\kappa = 2.0$ reaches the

Table 3: Controlled SDL sensitivity on Office-31 with SHOT, FC parameters, and $\rho = 0.001$. Each block varies one quantity and fixes the values shown in the third column. S1 max is the maximum number of Stage-1 iterations. Bold marks the best FO value within each block.

Varied	Value	Fixed quantities	FO	S1 max
α	.10	$\kappa = 1, \nu = .50, \omega = .20$	**79.559**	2268
	.15		78.518	1379
	.20		77.208	1121
κ	1.0	$\alpha = .10, \nu = .50, \omega = .20$	**79.559**	2268
	1.5		79.537	2501
	2.0		77.924	3000
ν	.25	$\alpha = .10, \kappa = 1, \omega = .20$	**80.099**	1623
	.50		79.559	2268
	1.0		78.968	3000
ω	.05	$\alpha = .10, \kappa = 1, \nu = .25$	78.361	1383
	.20		80.099	1623
	.30		**80.194**	1962

iteration limit and performs worse. For ν, the tighter-coupling setting achieves both the highest FO and a shorter discovery path, while $\nu = 1.0$ reaches the iteration limit. These results favor a regime in which structural support responds sufficiently quickly to the functional update without making the path excessively coarse or difficult to resolve.

Persistent write-back. The write-back coefficient exhibits a different trade-off. Increasing ω from 0.05 to 0.30 improves FO from 78.36% to 80.19%, showing that retaining more of each sparse update can improve the final adapted model. At the same time, the maximum discovery length increases from 1,383 to 1,962 iterations. A stronger write-back changes the starting model seen by subsequent batches more substantially, which can make later support discovery more demanding.

Overall, the sensitivity results suggest that SDL does not rely on making the discovery dynamics as aggressive as possible. Good performance is obtained when the sparse path is resolved at a moderate scale, while ω separately controls how strongly each discovered update is accumulated over the target stream.

5 Conclusions

This work reframes online test-time adaptation as a problem of learning not only *how* to update a model, but also *where* those updates should occur. Sparse Delta Learning (SDL) realizes this view by discovering a structured support for each target batch under an explicit group budget, then applying the original TTA optimizer only within that support. The same formulation extends across fully connected, convolutional, and Transformer parameters through architecture-specific group definitions. Across Office-31 and VisDA-C, our results show that sparsity alone is not sufficient: when the base adaptation objective provides useful update directions, learning where to place a limited update budget can substantially outperform fixed sparse selection and approach dense adaptation with a very small parameter footprint.

The controlled comparisons further reveal an interaction between the adaptation objective and the learned support. When the base objective provides informative directions, support learning can concentrate a limited update budget on a much more effective parameter subspace than fixed sparse selection. When the objective produces little effective change in the candidate space, as observed for COME in the Transformer setting, changing the support alone provides little benefit. These results suggest that the adaptation signal and its parameter support should be treated as complementary design axes in OTTA: the former determines *what* change is useful, while the latter determines *where* that change is allowed to act.

A remaining systems challenge is the cost of support discovery, since the current inverse-scale-space search dominates online latency. Developing faster approximations to the discovery path while preserving the quality of the selected support is therefore an important direction toward practical deployment-time support learning.

AI USE STATEMENT

We used generative AI tools only for language polishing and editing during the preparation of early manuscript drafts. All research ideas, methodologies, implementations, experiments, analyses, and conclusions were developed and verified by the authors. The authors reviewed all AI-assisted edits and take full responsibility for the final manuscript and its claims.

REPRODUCIBILITY STATEMENT

We will release the code used in this work to support reproducibility of the reported experiments and results. Detailed experimental settings and implementation details are provided in the main paper and appendix.

REFERENCES

Elie Aljalbout, Jiaxu Xing, Angel Romero, Iretiayo Akinola, Caelan Reed Garrett, Eric Heiden, Abhishek Gupta, Tucker Hermans, Yashraj Narang, Dieter Fox, Davide Scaramuzza, and Fabio Ramos. The reality gap in robotics: Challenges, solutions, and best practices. *Annual Review of Control, Robotics, and Autonomous Systems*, 9:403–432, 2026.

Martin Burger, Guy Gilboa, Stanley Osher, and Jinjun Xu. Nonlinear inverse scale space methods. *Communications in Mathematical Sciences*, 4(1):179–212, 2006.

Xiao Chen, Zhongjing Du, Jiazhen Huang, Xu Jiang, Li Lu, Jingyan Jiang, and Zhi Wang. Neural collapse in test-time adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 10567–10576, 2026.

Yizhuo Ding, Ke Fan, Yikai Wang, Xinwei Sun, and Yanwei Fu. Adaptive pruning of pretrained transformer via differential inclusions. In *International Conference on Learning Representations*, 2025.

Yanwei Fu, Chen Liu, Donghao Li, Xinwei Sun, Jinshan Zeng, and Yuan Yao. DessiLBI: Exploring structural sparsity of deep networks via differential inclusion paths. In *International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 3315–3326, 2020.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

Chendi Huang, Xinwei Sun, Jiechao Xiong, and Yuan Yao. Split LBI: An iterative regularization path with structural sparsity. In *Advances in Neural Information Processing Systems*, volume 29, pp. 3369–3377, 2016.

Hyeongyu Kim, GeonHui Han, and Dosik Hwang. Buffer layers for test-time adaptation. In *Advances in Neural Information Processing Systems*, volume 38, pp. 141127–141149, 2026.

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*, pp. 6028–6039. PMLR, 2020.

Jian Liang, Ran He, and Tieniu Tan. A comprehensive survey on test-time adaptation under distribution shifts. *International Journal of Computer Vision*, 133(1):31–64, 2025.

Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofo Chen, Shijian Zheng, Peilin Zhao, and Mingkui Tan. Efficient test-time model adaptation without forgetting. In *International Conference on Machine Learning*, pp. 16888–16905. PMLR, 2022.

Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005.

Stanley Osher, Feng Ruan, Jiechao Xiong, Yuan Yao, and Wotao Yin. Sparse recovery via differential inclusions. *Applied and Computational Harmonic Analysis*, 41(2):436–469, 2016.

Xingchao Peng, Ben Usman, Neela Kaushik, Judy Hoffman, Dequan Wang, and Kate Saenko. VisDA: The visual domain adaptation challenge. *arXiv preprint arXiv:1710.06924*, 2017.

Kate Saenko, Brian Kulis, Mario Fritz, and Trevor Darrell. Adapting visual category models to new domains. In *European Conference on Computer Vision*, pp. 213–226. Springer, 2010.

Jiaxu Tian and Fan Lyu. Parameter-selective continual test-time adaptation. In *Asian Conference on Computer Vision*, pp. 315–331. Springer, 2024.

Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Hervé Jégou. Training data-efficient image transformers & distillation through attention. In *International Conference on Machine Learning*, pp. 10347–10357. PMLR, 2021.

Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully test-time adaptation by entropy minimization. In *International Conference on Learning Representations*, 2021.

Qin Wang, Olga Fink, Luc Van Gool, and Dengxin Dai. Continual test-time domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7201–7211, 2022.

Qingyang Zhang, Yatao Bian, Xinke Kong, Peilin Zhao, and Changqing Zhang. COME: Test-time adaption by conservatively minimizing entropy. In *International Conference on Learning Representations*, 2025.

A Experimental Details

A.1 Datasets and Online Evaluation

Office-31 contains 2,817 Amazon images, 498 DSLR images, and 795 Webcam images from 31 classes. We evaluate A→D, A→W, D→A, D→W, W→A, and W→D. Accuracy is computed over samples within each transfer, and the Office-31 summary is the arithmetic mean of the six transfer accuracies. VisDA-C uses the synthetic `train` split as source and the 55,388-image real `validation` split as target. Its primary metric is the arithmetic mean of the accuracies of all 12 classes; we additionally retain sample-level overall accuracy and all class-wise accuracies.

The target stream is a fixed random permutation generated with seed 2026. Each stream is processed once without replay or dropped samples. Office-31 uses a batch size of 64, and VisDA-C uses 256; both use four data workers. Online images are bilinearly resized to 256×256, randomly cropped to 224×224, randomly flipped horizontally, and normalized with ImageNet statistics. PU is accumulated from a separate read-only forward pass after each update. After the stream, the final model is frozen and evaluated on the full target set using a 224×224 center crop to obtain FO. The adaptation objectives use no target labels.

A.2 Backbones, Candidate Spaces, and Budgets

Table 4: Candidate spaces used by SDL and the budget-matched sparse baselines. The integer budget is $K = \lfloor \rho |\mathcal{G}| \rfloor$ for $\rho \in \{0.0005, 0.001, 0.002\}$.

Track	Candidate parameters	Scalars	Groups	K
FC	ResNet bottleneck weight and bias	524,544	524,544	262 / 524 / 1,049
Conv	Nine convolution weights in ResNet `layer4`	12,845,056	9,216	4 / 9 / 18
Transformer	QKV, projection, and FFN weights in DeiT blocks 9–11	5,308,416	6,912	3 / 6 / 13

The FC track treats every scalar connection as one group and freezes its associated bottleneck BatchNorm layer and classifier. The Conv track groups all kernel weights that produce one output channel and freezes BatchNorm and classifier parameters. The Transformer track contains the 12 weight tensors `attn.qkv`, `attn.proj`, `mlp.fc1`, and `mlp.fc2` in blocks 9–11. It forms 6,912 QK, VO, and FFN groups, each containing 768 scalars; all biases, LayerNorm parameters, embeddings, earlier blocks, the final normalization layer, and the classification head are frozen.

A.3 Efficiency and Sparsity

Table 5: SDL scalar footprint as a percentage of all model parameters. FC and Transformer entries are the maximum footprints implied by K. Conv entries are the mean realized footprints in the formal SHOT runs, since output-channel groups contain different numbers of scalars.

Dataset	Track	Model parameters	$\rho = 0.0005$	$\rho = 0.001$	$\rho = 0.002$
Office-31	FC	24,041,086	0.00109%	0.00218%	0.00436%
	Conv (realized mean)	24,041,086	0.03715%	0.08196%	0.17518%
	Transformer	21,677,599	0.01063%	0.02126%	0.04606%
VisDA-C	FC	43,028,312	0.00061%	0.00122%	0.00244%
	Conv (realized mean)	43,028,312	0.01841%	0.03986%	0.10282%
	Transformer	21,670,284	0.01063%	0.02126%	0.04607%

For FC and Transformer, the fixed group sizes map the three integer budgets to 262/524/1,049 and 2,304/4,608/9,984 scalars, respectively. For Conv, the formal SHOT means are 8,932/19,704/42,114 scalars on Office-31 and 7,920.7/17,153.2/44,242.0 on VisDA-C. These measured values preserve the distinction between group sparsity and scalar sparsity.

Representative parameter footprint. At $\rho = 0.001$, SDL selects at most six groups of 768 scalars, or 4,608 scalars in total. This is 0.0213% of DeiT-S/16 on both datasets, compared with 5,308,416 scalars in the Candidate-dense space. The two main tables therefore hold the architecture and scalar budget fixed while changing the dataset and adaptation objective. Table 5 reports the corresponding footprints for every architecture and budget.

Formal efficiency benchmark. We benchmark SHOT on all six Office-31 transfers with DeiT-S/16 at $\rho = 0.0005$ ($K = 3$). Candidate-dense and sparse methods share the same last-three-block candidate space; Full-dense updates the measured all-except-classifier scope. All rows use identical target streams, deterministic FP32 execution, batch size 64, and RTX 4090 hardware. Random is the aggregate of three fixed masks. The formal benchmark contains 48 runs, with one process per GPU. Runtime covers the complete online adaptation computation after a batch is available; in particular, SDL includes Stage-1 discovery, rollback, support construction, and the masked Stage-2 update. Data loading, read-only PU evaluation, checkpoint I/O, and final FO evaluation are excluded.

Table 6: Formal efficiency benchmark for SHOT on Office-31 with DeiT-S/16 at $\rho = 0.0005$ ($K = 3$). PU and FO are averaged equally over six transfers, mean latency pools online batches, and peak memory is the maximum allocated memory over transfers and Random masks. All columns are aggregated from the same formal executions.

Method	Updated scalars	Updated (%)	PU	FO	Time (s/batch)	Peak MB
Full-dense	21,665,664	99.94494	82.3282	83.0893	0.0789	4,119
Candidate-dense	5,308,416	24.48803	81.3440	82.0019	0.0400	1,214
Random	2,304	0.01063	80.8932	80.9505	0.0450	1,219
Magnitude	2,304	0.01063	80.8932	80.9466	0.0426	1,219
Saliency	2,304	0.01063	80.9050	80.9525	0.0468	1,219
SDL	2,304	0.01063	81.1247	81.3700	7.0549	1,547

Table 6 shows that SDL updates 2,304 scalars per batch, or 0.01063% of DeiT-S/16. This reduces the update footprint by 99.957% relative to Candidate-dense adaptation and by 99.989% relative to Full-dense adaptation. Its FO accuracy is 0.42 points higher than the strongest fixed sparse selector. This parameter efficiency does not translate into lower latency: SDL takes 7.05 seconds per batch, 150–166$\times$ longer than the three sparse selectors, because Stage 1 accounts for 99.22% of its online time. SDL uses 62.4% less peak allocated memory than Full-dense, but 27.0–27.5% more than the fixed sparse and Candidate-dense variants. Thus, the measured advantage is a smaller learned update subspace, while support discovery remains the main systems cost.

Table 7: Detailed timing and memory measurements for the formal efficiency benchmark. Amortized latency adds one-time mask or setup cost to the complete online adaptation time. Memory is reported in MiB.

Method	Mean (s)	Median (s)	P95 (s)	Amort. (s)	Peak alloc.	Peak reserv.
Full-dense	0.078899	0.074919	0.079421	0.078912	4118.510	6192
Candidate-dense	0.039967	0.035803	0.037841	0.039974	1213.505	1826
Random	0.045016	0.036608	0.038389	0.045553	1218.567	1832
Magnitude	0.042602	0.036811	0.055610	0.043701	1218.567	1832
Saliency	0.046827	0.040136	0.049620	0.046834	1218.594	1838
SDL	7.054882	6.289911	11.768097	7.054883	1547.435	2270

All 48 child runs pass the strict scientific checks, and all 1,040 expected online-batch records are present. The runs share the same GPU and software identity, target order, source checkpoint per source domain, and FP32 setting; no resumed run, off-mask update, budget violation, or classifier update enters the aggregate. Accuracy, updated parameters, latency, and memory in Tables 6 and 7 are aggregated from these same executions.

A.4 Controlled Dense and Sparse Baselines

Source-only performs inference without adaptation. **Full-dense** updates the complete feature extractor and bottleneck for ResNet, or every DeiT parameter except the classification head. **Candidate-**

dense updates every parameter in the candidate space of Table 4. All controlled sparse methods use the same candidate space and integer budget as SDL. Random uses three fixed masks with seeds 202600, 202601, and 202602, and we report their mean. Magnitude ranks groups once using their source-model ℓ_2 norms. Saliency recomputes $\|\boldsymbol{W}_g \odot \nabla_{\boldsymbol{W}_g} \mathcal{L}\|_2$ on each incoming batch and updates the global top-K groups.

Within each architecture and objective, Candidate-dense, Random, Magnitude, and Saliency use the same host optimizer. For ResNet, SHOT and COME use SGD with momentum 0.9, weight decay 10^{-3}, and Nesterov acceleration; the learning rate is 10^{-2} on Office-31 and 10^{-3} on VisDA-C. For DeiT-S/16, SHOT uses AdamW with learning rate 10^{-5}, while COME uses AdamW with learning rate 10^{-7}; both use $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and weight decay 10^{-2}. The SHOT loss is $0.3\mathcal{L}_{\text{pseudo}} + \mathcal{L}_{\text{ent}} + \mathcal{L}_{\text{div}}$. COME uses entropy of opinion with $p = 2$, temperature 1, and numerical constant 10^{-7}.

A.5 Standalone OTTA References

Table 8: Optimization settings for standalone OTTA references. Each method performs at most one update on an incoming batch.

Method	Objective and selection	Optimizer and fixed parameters
Tent (Wang et al., 2021)	Mean prediction entropy; no sample filtering	Adam, learning rate 10^{-3}, $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and zero weight decay
EATA (Niu et al., 2022)	Reliable and non-redundant sample filtering with reweighted entropy and source Fisher regularization	SGD, learning rate 2.5×10^{-4}, momentum 0.9, zero weight decay; entropy margin $0.4 \ln C$, diversity margin 0.05, and Fisher weight 2000
NCTTA (Chen et al., 2026)	Entropy with cosine InfoNCE alignment; entropy and distance weighting	SGD, learning rate 10^{-3}, momentum 0.9, Nesterov, zero weight decay; entropy threshold $0.4 \ln C$, top-$k = 10$, and scale 5

For ResNet, we evaluate normalization affine parameters in the full, FC, and Conv scopes. For DeiT-S/16, we evaluate all LayerNorm affine parameters and the LayerNorms in blocks 9–11 as the full and module scopes, respectively. The classifier and all ordinary Conv/Linear weights remain frozen in these standalone references.

Table 9: Trainable parameter footprint of the standalone OTTA references. Tent, EATA, and NCTTA share these scopes; percentages use the matching dataset-specific model, including its classification head, as denominator.

	Office-31		VisDA-C	
Trainable scope	**Scalars**	**Model ratio**	**Scalars**	**Model ratio**
ResNet full BatchNorm	53,632	0.22308%	105,856	0.24601%
ResNet FC BatchNorm	512	0.00213%	512	0.00119%
ResNet Conv BatchNorm	18,432	0.07667%	18,432	0.04284%
DeiT full LayerNorm	19,200	0.08857%	19,200	0.08860%
DeiT module LayerNorm	4,608	0.02126%	4,608	0.02126%

A.6 SDL Optimization and Hyperparameter Selection

Every online batch starts with fresh local $\boldsymbol{\Delta}$, $\boldsymbol{\Gamma}$, and $\mathbf{Z}$ states. In implementation, Stage 1 treats a group as active when $\|\boldsymbol{\Gamma}_g\|_2$ exceeds a numerical tolerance of 10^{-4}; for Transformer groups, the norm is first divided by $\sqrt{|g|}$. Stage 1 stops when it reaches the integer budget or the limit of 3,000 iterations. If an update crosses the budget, SDL restores the last feasible state; it does not trim or fill the support with a top-K operation. Stage 2 starts from the persistent model plus the masked Stage-1 delta and performs exactly one update while keeping off-support values and optimizer states fixed. The resulting delta is written back with coefficient ω.

We fix $\kappa = 1$ and select a complete (α, ν, ω) tuple for each objective, track, and budget. Office-31 uses the equal-weight FO mean over all six transfers, while VisDA-C uses FO on its synthetic-to-real stream. The chosen tuple is shared across the corresponding Office-31 transfers and is frozen before evaluation from the source checkpoint.

Coverage of selected configurations. Tables 10 and 11 contain one frozen tuple for every cell in the full Cartesian product of two datasets, two objectives, three tracks, and three budgets, for 36 configurations in total. Each Office-31 cell uses one tuple selected from the equal-weight mean over all six transfers and shares it across those transfers; each VisDA-C cell is selected for its single synthetic-to-real stream.

For the FC and Conv tracks, Stage 2 uses a fresh SGD optimizer with momentum 0.9, weight decay 10^{-3}, and Nesterov acceleration. For the Transformer track, it uses a fresh AdamW optimizer with $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and weight decay 10^{-2}. The single masked update uses a fixed learning rate for each reported setting and no learning-rate schedule. Because this is a routine optimizer setting rather than a structure-discovery parameter, it is stated here and omitted from Tables 10 and 11.

For $\rho = (0.0005, 0.001, 0.002)$, the Stage-2 learning rates for SHOT are $(.020, .020, .005)$, $(.010, .0025, .0025)$, and $(.005, .005, .00125)$ on Office-31 for the FC, Conv, and Transformer tracks, respectively; the corresponding VisDA-C values are $(.010, .0025, .005)$, $(.005, .010, .005)$, and $(.005, .010, .005)$. For COME, the Office-31 values are $(.0025, .0009375, .0025)$, $(.0053125, .00275, .015)$, and $(.020, .0020, .015)$, while the VisDA-C values are $(.005, .005, .005)$, $(.005, .005, .010)$, and $(10^{-5}, 10^{-5}, 10^{-5})$.

Table 10: Selected SDL configurations with the SHOT objective. For readability, ω is rounded to two significant digits.

Dataset	**Track**	ρ	K	α	κ	ν	ω
Office-31	FC	.0005	262	.100	1	.25	.30
	FC	.0010	524	.100	1	.25	.30
	FC	.0020	1049	.150	1	.50	.30
	Conv	.0005	4	.050	1	.50	.10
	Conv	.0010	9	.050	1	.50	.10
	Conv	.0020	18	.100	1	1.0	.025
	Transformer	.0005	3	.100	1	1.0	.10
	Transformer	.0010	6	.100	1	.50	.050
	Transformer	.0020	13	.200	1	1.0	.025
VisDA-C	FC	.0005	262	.125	1	.50	.0063
	FC	.0010	524	.100	1	.50	.0031
	FC	.0020	1049	.150	1	.50	.0016
	Conv	.0005	4	.025	1	.50	.013
	Conv	.0010	9	.025	1	.50	.025
	Conv	.0020	18	.050	1	1.0	.025
	Transformer	.0005	3	.200	1	.50	.20
	Transformer	.0010	6	.200	1	1.0	.20
	Transformer	.0020	13	.200	1	1.0	.20

Table 11: Selected SDL configurations with the COME objective. For readability, ω is rounded to two significant digits.

Dataset	**Track**	ρ	K	α	κ	ν	ω
Office-31	FC	.0005	262	.200	1	1.0	.0063
	FC	.0010	524	.150	1	1.0	.0070
	FC	.0020	1049	.150	1	1.0	.0094
	Conv	.0005	4	.150	1	.50	.0032
	Conv	.0010	9	.200	1	.50	.0033
	Conv	.0020	18	.025	1	1.0	.019
	Transformer	.0005	3	.150	1	.50	.010
	Transformer	.0010	6	.200	1	.25	.0018
	Transformer	.0020	13	.150	1	.25	.00050
VisDA-C	FC	.0005	262	.100	1	.50	.0031
	FC	.0010	524	.100	1	.50	.0016
	FC	.0020	1049	.100	1	.50	.0016
	Conv	.0005	4	.100	1	.25	.0016
	Conv	.0010	9	.100	1	.25	.00078
	Conv	.0020	18	.200	1	.50	.00078
	Transformer	.0005	3	.100	1	1.0	9.8×10^{-5}
	Transformer	.0010	6	.100	1	1.0	9.8×10^{-5}
	Transformer	.0020	13	.100	1	1.0	9.8×10^{-5}

A.7 Implementation and Reporting

All experiments are implemented in PyTorch and run in deterministic FP32 mode. The classifier is frozen in every controlled comparison. Random masks, target order, and augmentation randomness are generated from the recorded seeds. Accuracy runs store the resolved configuration, source-checkpoint identity, stream identity, per-batch metrics, final summary, and class-wise VisDA-C statistics. Efficiency runs use one experiment process per GPU and report adaptation time per batch, peak allocated GPU memory, Stage-1 steps, rollback frequency, and realized support.

B Complete Office-31 Results

The tables below report all six transfers and their equal-weight average for every budget. Each table keeps dense controls, budget-matched sparse selectors, SDL, and architecture-matched OTTA references together.

SDL exceeds the strongest sparse selector in all 12 objective–architecture PU/FO averages at $\rho = 0.0005$, in 9 of 12 at $\rho = 0.001$, and in 11 of 12 at $\rho = 0.002$. The largest gains occur for SHOT in the FC track and grow with the budget: PU/FO gains are 1.15/1.48 points at $\rho = 0.0005$, 2.17/2.53 at $\rho = 0.001$, and 2.78/3.45 at $\rho = 0.002$. The increasing margin shows that SHOT can use the additional scalar support selected by SDL, whereas the fixed selectors remain close to Source-only.

The Conv and Transformer tracks show smaller but stable improvements because each selected group moves many coupled parameters. For SHOT, SDL is the best sparse variant on both PU and FO at every budget in both structured tracks. COME produces narrower margins: its FC result ranges from a 0.33-point FO gain at $\rho = 0.0005$ to a 0.21-point PU loss at $\rho = 0.002$, and COME–Conv loses 0.14 FO points only at $\rho = 0.001$. Together with two ties at $\rho = 0.001$, these cases account for the four non-wins across 36 comparisons; SDL is strictly better in the other 32.

Table 13: Office-31 accuracy (%) for the ResNet Conv track at $\rho = 0.0005$. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.12	72.58	64.70	94.97	65.45	97.99	79.30	82.73	75.85	66.92	95.35	66.77	97.79	80.90
SHOT	Random	77.51	**71.11**	61.66	93.21	63.08	**97.59**	77.36	79.52	72.24	60.94	94.34	62.30	**96.52**	77.64
SHOT	Magnitude	77.71	70.44	58.34	93.58	59.84	97.39	76.22	**79.92**	72.20	53.14	94.34	54.31	95.98	74.98
SHOT	Saliency	77.51	70.57	58.52	93.58	60.26	**97.59**	76.34	79.72	72.08	53.18	94.34	54.81	95.98	75.02
SHOT	**SDL**	**78.92**	71.07	**62.61**	**94.21**	**63.32**	96.59	**77.78**	79.52	**73.08**	**61.80**	**94.59**	**63.12**	95.18	**77.88**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	80.12	73.71	63.67	94.72	61.47	97.79	78.58	81.93	75.47	62.30	95.97	60.84	98.39	79.15
COME	Random	77.51	**71.07**	61.59	93.21	63.09	97.59	77.34	79.52	**72.20**	60.86	94.34	62.27	96.59	77.63
COME	Magnitude	77.51	**71.07**	**61.86**	93.21	62.96	97.59	77.37	79.52	71.95	**61.23**	94.34	62.30	96.59	77.66
COME	Saliency	77.51	**71.07**	61.79	93.21	**63.14**	97.59	77.38	79.32	71.95	61.06	**94.59**	**62.55**	96.59	77.67
COME	**SDL**	**77.71**	**71.07**	61.75	**93.33**	62.93	**97.79**	**77.43**	**79.72**	72.08	60.99	94.34	62.16	**96.79**	**77.68**
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	Conv	79.32	73.71	63.74	96.48	60.33	97.99	78.60	57.63	42.26	17.82	66.54	17.57	84.34	47.69
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	Conv	77.91	71.70	63.49	96.35	61.61	97.59	78.11	55.42	37.86	18.53	64.91	19.70	83.33	46.63
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	Conv	78.71	73.21	61.90	96.60	60.23	97.39	78.01	55.62	41.64	18.53	64.65	19.99	83.33	47.29

Table 12: Office-31 accuracy (%) for the ResNet FC track at $\rho = 0.0005$. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.72	73.08	58.17	94.84	59.94	98.19	77.49	83.13	75.85	59.28	95.85	61.87	97.99	79.00
SHOT	Random	77.51	71.07	61.62	93.21	63.12	97.59	77.35	79.52	72.20	60.86	94.34	62.25	96.59	77.63
SHOT	Magnitude	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	**60.99**	94.34	62.23	96.59	77.64
SHOT	Saliency	77.51	70.94	**61.65**	93.21	63.10	97.59	77.33	79.52	72.08	60.95	94.34	62.27	96.59	77.62
SHOT	**SDL**	**81.73**	**75.35**	56.71	**94.97**	**63.28**	**99.00**	**78.50**	**83.33**	**78.99**	55.45	**96.23**	**62.90**	**97.79**	**79.12**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	81.33	71.07	63.14	93.84	60.69	98.39	78.08	83.53	73.58	62.09	95.72	60.06	98.59	78.93
COME	Random	77.51	71.07	61.62	**93.21**	**63.12**	97.59	77.35	79.52	**72.20**	60.84	**94.34**	62.27	96.59	77.63
COME	Magnitude	77.51	71.07	61.58	**93.21**	63.10	97.59	77.34	79.52	**72.20**	60.81	**94.34**	62.27	96.59	77.62
COME	Saliency	77.71	71.07	62.25	**93.21**	62.75	**97.79**	77.46	**79.72**	71.95	61.41	**94.34**	62.73	96.39	77.75
COME	**SDL**	**78.11**	**71.32**	**62.75**	93.08	62.93	**97.79**	**77.66**	**79.72**	71.95	**61.77**	94.09	**63.19**	**97.79**	**78.08**
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	FC	80.72	72.58	62.36	95.60	62.96	97.59	78.63	64.26	50.06	27.01	72.45	28.15	88.55	55.08
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	FC	80.72	72.58	62.54	95.60	63.14	97.59	78.69	64.26	50.19	26.98	72.70	28.22	88.55	55.15
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	FC	80.72	72.58	62.54	95.60	63.10	97.59	78.69	64.26	50.19	27.01	72.70	28.26	88.55	55.16

C Complete VisDA-C Results

The main text reports the Transformer track at $\rho = 0.001$; this section provides the complete class-wise sweep over all three architectures and all three budgets. PU and FO are reported side by side in each table.

Across the full sweep, SDL improves mAcc over the strongest sparse selector in every SHOT and COME comparison in the FC and Conv tracks. For SHOT, the PU/FO gains at $\rho = 0.0005$ and $\rho = 0.001$ are 3.69/5.19 and 2.61/3.75 points in the FC track, and 1.25/1.78 and 1.64/2.84 points in the Conv track. At $\rho = 0.002$, the corresponding gains remain positive at 3.08/4.80 and 2.55/3.83

Table 14: Office-31 accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.0005$. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	82.73	79.25	63.51	94.21	66.06	99.60	80.89	82.13	79.37	64.29	94.09	66.21	99.60	80.95
SHOT	Full-dense	84.34	83.40	63.65	94.84	67.94	99.80	82.33	85.74	85.66	63.51	95.60	68.23	99.80	83.09
SHOT	Candidate-dense	83.13	80.38	64.00	94.21	66.74	99.60	81.34	83.94	81.13	65.35	94.21	67.38	100.00	82.00
SHOT	Random	82.73	79.25	63.51	**94.21**	66.06	**99.60**	80.89	82.13	79.37	**64.29**	**94.09**	66.23	99.60	80.95
SHOT	Magnitude	82.73	79.25	63.51	**94.21**	66.06	**99.60**	80.89	82.13	79.37	**64.29**	**94.09**	66.21	99.60	80.95
SHOT	Saliency	82.73	79.25	63.54	**94.21**	66.10	**99.60**	80.91	82.13	79.37	**64.29**	**94.09**	**66.24**	99.60	80.95
SHOT	**SDL**	**83.33**	**80.13**	**63.65**	93.84	**66.21**	**99.60**	**81.12**	**84.14**	**80.63**	64.04	93.84	65.78	**99.80**	**81.37**
COME	Full-dense	82.73	79.50	63.54	94.21	65.99	99.60	80.93	82.53	79.50	64.50	94.09	66.42	99.60	81.11
COME	Candidate-dense	82.73	79.25	63.54	94.21	66.06	99.60	80.90	82.13	79.37	64.29	94.09	66.31	99.60	80.96
COME	Random	**82.73**	79.25	63.51	**94.21**	**66.06**	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	**99.60**	80.95
COME	Magnitude	**82.73**	79.25	63.51	**94.21**	**66.06**	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	**99.60**	80.95
COME	Saliency	**82.73**	79.25	63.51	**94.21**	**66.06**	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	**99.60**	80.95
COME	**SDL**	**82.73**	**79.50**	**63.61**	94.09	**66.06**	**99.60**	**80.93**	**82.53**	**79.75**	**64.36**	93.96	**66.45**	**99.60**	**81.11**
Standard OTTA references															
Tent	DeiT full	83.13	81.26	57.88	94.59	65.55	99.80	80.37	83.73	82.26	51.08	94.72	63.40	99.80	79.17
Tent	Transformer	83.13	79.62	62.18	94.09	65.31	99.60	80.65	82.73	79.62	61.66	93.84	64.89	100.00	80.46
EATA	DeiT full	82.93	79.25	63.60	94.21	66.09	99.60	80.95	83.13	79.37	64.32	94.09	66.35	99.60	81.14
EATA	Transformer	82.73	79.25	63.53	94.21	66.12	99.60	80.91	82.13	79.37	64.29	94.09	66.24	99.60	80.95
NCTTA	DeiT full	83.53	79.37	63.64	94.09	65.09	99.60	80.89	83.33	80.88	62.62	94.47	64.22	100.00	80.92
NCTTA	Transformer	83.13	79.50	63.53	93.96	65.34	99.60	80.84	82.53	80.25	64.04	94.09	65.39	99.60	80.98

Table 15: Office-31 accuracy (%) for the ResNet FC track at $\rho = 0.001$. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.72	73.08	58.17	94.84	59.94	98.19	77.49	83.13	75.85	59.28	95.85	61.87	97.99	79.00
SHOT	Random	77.51	71.07	61.64	93.21	**63.10**	97.59	77.35	79.45	72.20	60.88	94.34	62.30	96.59	77.63
SHOT	Magnitude	77.51	71.07	**61.68**	93.21	**63.10**	97.59	77.36	79.52	72.20	**60.99**	94.34	62.30	96.59	77.66
SHOT	Saliency	77.51	71.07	61.58	93.21	62.96	97.59	77.32	79.52	72.08	60.70	94.34	62.16	96.59	77.56
SHOT	**SDL**	**83.13**	**77.86**	59.84	**94.72**	62.82	**98.80**	**79.53**	**84.14**	**81.64**	58.93	**95.47**	**62.80**	**98.19**	**80.19**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	81.33	71.07	63.14	93.84	60.69	98.39	78.08	83.53	73.58	62.09	95.72	60.06	98.59	78.93
COME	Random	77.51	71.07	61.64	93.21	63.08	97.59	77.35	79.45	72.20	60.89	**94.34**	62.36	96.45	77.62
COME	Magnitude	77.51	71.07	61.68	93.21	**63.10**	97.59	77.36	79.52	72.20	60.92	**94.34**	62.30	96.59	77.64
COME	Saliency	77.71	**71.32**	62.39	**93.33**	62.68	**97.79**	**77.54**	**79.72**	**72.33**	**61.77**	**94.34**	**62.87**	96.39	77.90
COME	**SDL**	**77.91**	**71.32**	**62.61**	92.96	62.68	**97.79**	**77.54**	79.52	72.20	61.70	93.84	62.73	**97.79**	**77.96**
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	FC	80.72	72.58	62.36	95.60	62.96	97.59	78.63	64.26	50.06	27.01	72.45	28.15	88.55	55.08
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	FC	80.72	72.58	62.54	95.60	63.14	97.59	78.69	64.26	50.19	26.98	72.70	28.22	88.55	55.15
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	FC	80.72	72.58	62.54	95.60	63.10	97.59	78.69	64.26	50.19	27.01	72.70	28.26	88.55	55.16

points. COME is less sensitive to the budget, but SDL retains a positive aggregate margin throughout both ResNet tracks.

The Transformer sweep separates the two objectives more strongly. Under SHOT, SDL's PU/FO margins increase from 4.91/8.74 points at $\rho = 0.0005$ to 8.18/14.29 at the main $\rho = 0.001$ setting, and remain large at 11.08/12.94 points for $\rho = 0.002$. FO is not monotonic across the last two budgets, but every budget exposes a much stronger subspace than the static selectors. Under COME, every sparse variant remains within 0.03 mAcc across the sweep. Thus, additional support expands what SDL can recover from the SHOT signal, while it has little effect when COME produces an almost inactive Transformer update.

Table 16: Office-31 accuracy (%) for the ResNet Conv track at $\rho = 0.001$. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.12	72.58	64.70	94.97	65.45	97.99	79.30	82.73	75.85	66.92	95.35	66.77	97.79	80.90
SHOT	Random	77.51	71.11	61.72	93.21	63.02	**97.59**	77.36	79.58	72.24	60.94	94.34	62.27	**96.45**	77.64
SHOT	Magnitude	77.51	70.57	57.88	93.71	59.55	97.39	76.10	79.72	71.82	52.86	94.34	54.10	95.98	74.80
SHOT	Saliency	78.11	70.82	57.67	**94.34**	59.45	**97.59**	76.33	79.72	71.57	53.57	93.96	54.45	95.98	74.88
SHOT	**SDL**	**79.12**	**71.82**	**62.61**	**94.34**	**63.32**	96.19	**77.90**	**80.72**	**73.58**	**62.02**	**94.72**	**62.98**	95.18	**78.20**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	80.12	73.71	63.67	94.72	61.47	97.79	78.58	81.93	75.47	62.30	95.97	60.84	98.39	79.15
COME	Random	77.51	71.07	61.62	93.21	63.08	97.59	77.35	79.52	**72.20**	60.86	94.34	62.29	96.59	77.63
COME	Magnitude	77.51	70.94	**61.86**	93.21	63.32	97.59	77.41	**79.72**	71.95	**61.27**	94.34	62.76	96.59	77.77
COME	Saliency	77.51	71.07	61.79	93.21	**63.64**	97.59	77.47	79.52	72.08	61.20	**94.59**	**62.94**	96.59	**77.82**
COME	**SDL**	**77.71**	**71.32**	61.83	**93.33**	62.89	**97.79**	**77.48**	**79.72**	72.08	60.74	94.34	62.44	**96.79**	77.68
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	Conv	79.32	73.71	63.74	96.48	60.33	97.99	78.60	57.63	42.26	17.82	66.54	17.57	84.34	47.69
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	Conv	77.91	71.70	63.49	96.35	61.61	97.59	78.11	55.42	37.86	18.53	64.91	19.70	83.33	46.63
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	Conv	78.71	73.21	61.90	96.60	60.23	97.39	78.01	55.62	41.64	18.53	64.65	19.99	83.33	47.29

Table 17: Office-31 accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.001$. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	82.73	79.25	63.51	94.21	66.06	99.60	80.89	82.13	79.37	64.29	94.09	66.21	99.60	80.95
SHOT	Full-dense	84.34	83.40	63.65	94.84	67.94	99.80	82.33	85.74	85.66	63.51	95.60	68.23	99.80	83.09
SHOT	Candidate-dense	83.13	80.38	64.00	94.21	66.74	99.60	81.34	83.94	81.13	65.35	94.21	67.38	100.00	82.00
SHOT	Random	82.73	79.25	63.52	**94.21**	66.06	**99.60**	80.90	82.13	79.37	**64.30**	**94.09**	66.23	99.60	80.95
SHOT	Magnitude	82.73	79.25	63.51	**94.21**	66.06	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	99.60	80.95
SHOT	Saliency	82.73	79.25	63.54	**94.21**	66.10	**99.60**	80.91	82.13	79.37	64.29	**94.09**	66.24	99.60	80.95
SHOT	**SDL**	**83.33**	**79.75**	**63.79**	93.84	**66.28**	**99.60**	**81.10**	**83.94**	**80.13**	64.04	93.96	**66.84**	**99.80**	**81.45**
COME	Full-dense	82.73	79.50	63.54	94.21	65.99	99.60	80.93	82.53	79.50	64.50	94.09	66.42	99.60	81.11
COME	Candidate-dense	82.73	79.25	63.54	94.21	66.06	99.60	80.90	82.13	79.37	64.29	94.09	66.31	99.60	80.96
COME	Random	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	Magnitude	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	Saliency	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	**80.95**
COME	**SDL**	**82.73**	**79.25**	**63.58**	**94.21**	66.03	**99.60**	**80.90**	**82.13**	**79.37**	64.25	**94.09**	**66.28**	**99.60**	**80.95**
Standard OTTA references															
Tent	DeiT full	83.13	81.26	57.88	94.59	65.55	99.80	80.37	83.73	82.26	51.08	94.72	63.40	99.80	79.17
Tent	Transformer	83.13	79.62	62.18	94.09	65.31	99.60	80.65	82.73	79.62	61.66	93.84	64.89	100.00	80.46
EATA	DeiT full	82.93	79.25	63.60	94.21	66.09	99.60	80.95	83.13	79.37	64.32	94.09	66.35	99.60	81.14
EATA	Transformer	82.73	79.25	63.53	94.21	66.12	99.60	80.91	82.13	79.37	64.29	94.09	66.24	99.60	80.95
NCTTA	DeiT full	83.53	79.37	63.64	94.09	65.09	99.60	80.89	83.33	80.88	62.62	94.47	64.22	100.00	80.92
NCTTA	Transformer	83.13	79.50	63.53	93.96	65.34	99.60	80.84	82.53	80.25	64.04	94.09	65.39	99.60	80.98

Table 18: Office-31 accuracy (%) for the ResNet FC track at $\rho = 0.002$. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.72	73.08	58.17	94.84	59.94	98.19	77.49	83.13	75.85	59.28	95.85	61.87	97.99	79.00
SHOT	Random	77.51	70.98	61.68	93.21	63.13	97.59	77.35	79.32	72.20	60.99	94.34	62.29	96.59	77.62
SHOT	Magnitude	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	61.06	94.34	62.30	96.59	77.67
SHOT	Saliency	77.51	71.32	61.61	93.21	63.00	97.79	77.41	79.72	71.95	60.60	94.47	62.30	96.59	77.60
SHOT	**SDL**	**82.13**	**78.49**	**61.90**	**95.09**	**64.74**	**98.80**	**80.19**	**85.34**	**80.63**	**61.98**	**95.35**	**65.25**	**98.19**	**81.12**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	81.33	71.07	63.14	93.84	60.69	98.39	78.08	83.53	73.58	62.09	95.72	60.06	98.59	78.93
COME	Random	77.51	71.07	61.65	93.21	**63.12**	97.59	77.36	79.32	**72.20**	61.05	**94.34**	62.32	96.45	77.61
COME	Magnitude	77.51	71.07	61.65	93.21	63.10	97.59	77.36	79.32	72.08	61.02	**94.34**	62.41	96.39	77.59
COME	Saliency	**77.71**	**71.45**	**62.61**	**93.46**	62.57	97.79	**77.60**	**79.72**	**72.20**	61.95	94.21	**62.90**	96.59	77.93
COME	**SDL**	77.51	**71.45**	62.36	92.83	62.22	**97.99**	77.39	79.52	**72.20**	**61.98**	93.96	62.51	**97.79**	**77.99**
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	FC	80.72	72.58	62.36	95.60	62.96	97.59	78.63	64.26	50.06	27.01	72.45	28.15	88.55	55.08
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	FC	80.72	72.58	62.54	95.60	63.14	97.59	78.69	64.26	50.19	26.98	72.70	28.22	88.55	55.15
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	FC	80.72	72.58	62.54	95.60	63.10	97.59	78.69	64.26	50.19	27.01	72.70	28.26	88.55	55.16

Table 19: Office-31 accuracy (%) for the ResNet Conv track at $\rho = 0.002$. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	77.51	71.07	61.61	93.21	63.10	97.59	77.35	79.52	72.20	60.81	94.34	62.23	96.59	77.61
SHOT	Full-dense	77.11	76.73	66.58	95.97	61.72	99.00	79.52	81.12	84.53	69.26	96.86	63.44	99.20	82.40
SHOT	Candidate-dense	80.12	72.58	64.70	94.97	65.45	97.99	79.30	82.73	75.85	66.92	95.35	66.77	97.79	80.90
SHOT	Random	77.51	71.11	61.70	93.25	**63.13**	**97.59**	77.38	79.58	72.24	61.05	94.34	**62.34**	**96.45**	77.67
SHOT	Magnitude	77.91	70.31	56.85	93.71	58.70	97.39	75.81	79.52	71.95	52.43	93.96	54.42	95.78	74.68
SHOT	Saliency	77.91	70.06	57.07	**94.21**	58.59	97.39	75.87	80.12	71.57	53.64	93.96	54.70	95.98	75.00
SHOT	**SDL**	**78.31**	**71.32**	**62.85**	93.96	62.82	**97.59**	**77.81**	**80.72**	**72.83**	**61.45**	**94.72**	62.16	96.19	**78.01**
COME	Full-dense	76.51	75.60	63.35	95.47	60.05	99.00	78.33	79.32	82.01	65.57	96.73	61.31	99.20	80.69
COME	Candidate-dense	80.12	73.71	63.67	94.72	61.47	97.79	78.58	81.93	75.47	62.30	95.97	60.84	98.39	79.15
COME	Random	77.51	71.07	61.65	93.21	63.09	97.59	77.35	79.52	72.20	60.86	94.34	62.28	96.59	77.63
COME	Magnitude	77.51	**71.19**	**62.00**	93.21	63.32	97.59	77.47	79.32	72.08	**61.52**	94.34	62.90	96.59	77.79
COME	Saliency	77.51	71.07	61.79	**93.33**	**63.64**	**97.79**	77.52	**79.72**	71.95	61.16	94.59	**63.15**	96.59	77.86
COME	**SDL**	**77.91**	71.07	61.83	**93.33**	63.60	**97.79**	**77.59**	**79.72**	**72.70**	60.92	**94.84**	63.01	**96.79**	**78.00**
Standard OTTA references															
Tent	ResNet full	79.12	78.74	67.22	96.48	61.08	98.19	80.14	55.82	42.64	17.71	62.77	16.79	82.73	46.41
Tent	Conv	79.32	73.71	63.74	96.48	60.33	97.99	78.60	57.63	42.26	17.82	66.54	17.57	84.34	47.69
EATA	ResNet full	76.10	73.84	59.45	95.60	57.71	97.79	76.75	47.99	34.72	14.13	58.87	14.84	80.12	41.78
EATA	Conv	77.91	71.70	63.49	96.35	61.61	97.59	78.11	55.42	37.86	18.53	64.91	19.70	83.33	46.63
NCTTA	ResNet full	76.91	75.60	58.84	95.72	58.63	98.19	77.32	50.80	40.25	15.34	61.89	15.69	81.12	44.18
NCTTA	Conv	78.71	73.21	61.90	96.60	60.23	97.39	78.01	55.62	41.64	18.53	64.65	19.99	83.33	47.29

Table 20: Office-31 accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.002$. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU							FO						
		A→D	A→W	D→A	D→W	W→A	W→D	Avg.	A→D	A→W	D→A	D→W	W→A	W→D	Avg.
Controlled adaptation															
–	Source-only	82.73	79.25	63.51	94.21	66.06	99.60	80.89	82.13	79.37	64.29	94.09	66.21	99.60	80.95
SHOT	Full-dense	84.34	83.40	63.65	94.84	67.94	99.80	82.33	85.74	85.66	63.51	95.60	68.23	99.80	83.09
SHOT	Candidate-dense	83.13	80.38	64.00	94.21	66.74	99.60	81.34	83.94	81.13	65.35	94.21	67.38	100.00	82.00
SHOT	Random	82.73	79.25	63.52	**94.21**	66.06	**99.60**	80.90	82.13	79.37	**64.31**	**94.09**	66.24	**99.60**	80.96
SHOT	Magnitude	82.73	79.25	63.51	**94.21**	66.06	**99.60**	80.89	82.13	79.37	64.29	**94.09**	66.21	**99.60**	80.95
SHOT	Saliency	82.73	79.25	63.58	**94.21**	66.10	**99.60**	80.91	82.13	79.37	64.25	**94.09**	66.31	**99.60**	80.96
SHOT	**SDL**	**82.93**	**79.62**	**63.76**	93.84	**66.31**	**99.60**	**81.01**	**83.13**	**79.50**	63.93	93.96	**66.92**	**99.60**	**81.17**
COME	Full-dense	82.73	79.50	63.54	94.21	65.99	99.60	80.93	82.53	79.50	64.50	94.09	66.42	99.60	81.11
COME	Candidate-dense	82.73	79.25	63.54	94.21	66.06	99.60	80.90	82.13	79.37	64.29	94.09	66.31	99.60	80.96
COME	Random	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	80.95
COME	Magnitude	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	80.95
COME	Saliency	**82.73**	**79.25**	63.51	**94.21**	**66.06**	**99.60**	80.89	**82.13**	**79.37**	**64.29**	**94.09**	66.21	**99.60**	80.95
COME	**SDL**	**82.73**	**79.25**	**63.58**	**94.21**	66.03	**99.60**	**80.90**	**82.13**	**79.37**	**64.29**	**94.09**	**66.31**	**99.60**	**80.96**
Standard OTTA references															
Tent	DeiT full	83.13	81.26	57.88	94.59	65.55	99.80	80.37	83.73	82.26	51.08	94.72	63.40	99.80	79.17
Tent	Transformer	83.13	79.62	62.18	94.09	65.31	99.60	80.65	82.73	79.62	61.66	93.84	64.89	100.00	80.46
EATA	DeiT full	82.93	79.25	63.60	94.21	66.09	99.60	80.95	83.13	79.37	64.32	94.09	66.35	99.60	81.14
EATA	Transformer	82.73	79.25	63.53	94.21	66.12	99.60	80.91	82.13	79.37	64.29	94.09	66.24	99.60	80.95
NCTTA	DeiT full	83.53	79.37	63.64	94.09	65.09	99.60	80.89	83.33	80.88	62.62	94.47	64.22	100.00	80.92
NCTTA	Transformer	83.13	79.50	63.53	93.96	65.34	99.60	80.84	82.53	80.25	64.04	94.09	65.39	99.60	80.98

Table 21: VisDA-C class accuracy (%) for the ResNet FC track at $\rho = 0.0005$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU												FO										
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	57.32	11.19	79.21	65.55	71.41	0.77	86.99	15.30	39.26	68.52	69.14	0.67	47.11	60.89	12.78	81.19	65.10	74.14	0.10	86.30	26.00	37.94	72.91	70.56	0.32	49.02
SHOT	Random	54.20	29.10	49.52	74.76	65.69	4.43	76.51	21.52	67.66	46.95	**77.57**	6.47	47.87	53.72	28.71	50.68	74.99	66.43	4.67	76.38	21.91	**67.85**	48.50	**78.27**	6.30	48.20
SHOT	Magnitude	54.22	29.12	49.55	74.76	65.76	4.43	76.50	21.50	67.64	47.00	**77.57**	6.47	47.88	53.73	28.72	**50.73**	74.95	66.51	4.67	76.38	21.80	67.82	**48.53**	78.23	6.33	48.20
SHOT	Saliency	53.81	28.40	**49.68**	**75.82**	65.17	4.43	**77.05**	21.10	66.85	46.16	77.27	6.07	47.65	53.04	27.71	50.66	**76.31**	65.76	4.67	**77.28**	21.07	66.26	46.91	77.88	5.70	47.77
SHOT	**SDL**	**54.69**	**44.29**	48.27	70.12	**75.70**	**18.41**	74.31	**22.95**	**73.27**	**51.47**	76.89	**8.51**	**51.57**	**55.02**	**57.64**	47.97	64.37	**79.66**	**41.40**	70.62	**22.12**	66.87	48.36	76.82	**9.86**	**53.39**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	59.57	39.40	69.49	49.99	68.09	34.12	73.72	29.57	62.19	44.06	71.20	24.19	52.13	59.98	44.75	70.28	47.27	70.67	32.77	68.13	33.45	66.87	46.12	68.08	28.23	53.05
COME	Random	54.21	29.03	49.43	74.75	65.48	4.48	76.52	**21.50**	67.67	46.87	**77.66**	6.47	47.84	53.73	28.59	50.53	74.97	66.23	4.67	76.38	**21.77**	67.86	48.40	**78.34**	6.36	48.15
COME	Magnitude	**54.28**	**29.12**	49.42	74.78	65.66	4.43	76.48	**21.50**	67.77	46.87	77.57	6.47	47.86	**53.76**	**28.75**	50.58	**74.98**	66.40	4.67	76.36	21.73	67.88	48.49	78.28	6.34	48.19
COME	Saliency	53.98	26.96	51.34	**74.79**	65.68	5.78	79.06	20.55	65.42	44.41	77.41	6.56	47.66	53.29	25.18	52.86	74.70	66.36	6.55	80.26	20.07	64.17	43.71	77.88	6.62	47.64
COME	**SDL**	53.81	27.42	**63.37**	65.05	**67.09**	**9.59**	**80.26**	21.35	**69.88**	**50.33**	72.14	**13.14**	**49.45**	51.15	25.50	**73.99**	54.01	**66.94**	**10.22**	**82.31**	19.38	**68.89**	**53.35**	68.72	**21.74**	**49.68**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	FC	12.75	7.71	96.06	8.27	11.70	22.60	24.93	4.95	8.27	8.90	26.16	4.06	19.70	0.00	0.49	95.03	0.78	0.55	0.24	1.22	0.30	0.44	2.32	0.59	0.40	8.53
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	FC	48.57	33.78	80.30	41.45	46.13	62.02	73.78	19.73	35.55	25.91	73.32	18.66	46.60	0.74	12.78	28.12	19.32	10.13	18.46	22.93	10.65	11.10	7.58	30.52	12.78	15.43
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	FC	45.80	34.13	81.30	40.61	41.97	53.93	72.83	17.88	29.44	42.35	71.91	18.69	45.90	0.41	13.35	24.86	17.74	6.22	14.60	18.89	9.55	7.19	28.23	28.02	11.41	15.04

Table 22: VisDA-C class accuracy (%) for the ResNet Conv track at $\rho = 0.0005$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU												FO										
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	60.83	36.37	77.48	64.70	76.15	1.83	83.16	40.62	58.58	64.27	70.52	13.54	54.01	63.16	58.70	76.78	57.58	77.38	1.30	80.25	54.30	60.83	63.57	74.29	29.65	58.15
SHOT	Random	54.43	**29.03**	50.02	74.70	65.78	**4.43**	76.50	21.52	**67.88**	47.01	77.40	**6.44**	47.93	53.99	**28.77**	51.23	74.84	66.60	**4.71**	76.39	21.77	**68.21**	48.33	78.11	**6.34**	48.27
SHOT	Magnitude	54.53	26.04	**50.30**	74.76	69.50	3.95	78.69	19.85	67.22	47.22	**77.45**	5.35	47.91	54.14	22.91	**51.24**	74.90	72.35	3.57	80.26	18.57	67.09	48.22	**78.21**	4.40	47.99
SHOT	Saliency	54.74	26.48	49.62	75.35	69.64	3.90	77.90	19.55	**67.88**	47.83	77.36	5.30	47.96	54.72	23.80	50.30	75.84	72.29	3.47	78.92	18.40	68.17	48.84	77.95	4.45	48.09
SHOT	**SDL**	**70.46**	23.34	43.35	**78.11**	**73.18**	2.60	**79.40**	**21.95**	66.32	**52.56**	73.70	5.59	**49.21**	**83.21**	18.36	38.23	**80.46**	**77.13**	2.41	**81.42**	**23.32**	64.98	**55.37**	70.87	4.85	**50.05**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	46.57	12.72	30.49	84.02	57.22	19.61	87.04	8.88	48.03	21.66	65.53	3.19	40.41	34.15	5.53	2.79	93.29	45.64	12.24	85.33	2.88	30.91	10.26	47.69	0.65	30.95
COME	Random	54.25	29.05	49.57	74.75	65.76	4.43	76.37	**21.49**	67.78	46.91	77.55	6.43	47.86	53.79	28.69	50.68	74.95	66.62	4.67	76.27	21.71	68.10	48.28	78.22	**6.29**	48.19
COME	Magnitude	**54.36**	29.07	49.59	74.77	65.81	4.43	76.50	21.48	67.75	**47.04**	77.53	**6.47**	47.90	53.95	28.69	50.81	75.06	**66.79**	4.63	76.43	**21.80**	68.10	**48.66**	78.31	**6.29**	48.29
COME	Saliency	54.06	28.66	48.93	75.04	65.40	4.43	**76.71**	21.45	67.53	46.95	**77.69**	6.42	47.77	53.48	28.23	49.13	75.39	66.36	4.67	**76.78**	21.38	67.51	48.01	**78.56**	6.22	47.98
COME	**SDL**	54.17	**33.93**	**55.76**	**75.20**	**66.49**	**9.06**	73.83	19.65	**73.25**	43.93	73.65	5.19	**48.67**	**54.06**	**36.69**	**58.42**	**76.79**	66.64	**13.59**	72.58	17.90	**76.19**	41.74	69.97	3.84	**49.03**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	Conv	0.00	0.09	0.00	0.42	0.00	9.16	0.43	0.00	0.02	0.00	97.69	0.00	8.98	0.00	0.00	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	98.21	0.00	8.22
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	Conv	0.66	18.50	2.94	5.85	1.17	86.60	31.99	0.73	5.78	0.53	77.48	0.79	19.42	0.00	11.31	8.83	26.15	5.84	35.86	26.98	2.10	15.81	0.00	71.79	3.05	17.31
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	Conv	0.14	11.97	2.05	49.88	0.15	49.35	14.91	0.25	2.53	0.09	40.16	0.11	14.30	0.00	0.00	0.04	99.73	0.04	0.00	0.00	0.00	0.00	0.00	0.02	0.00	8.32

Table 23: VisDA-C class accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.0005$. PU and FO are reported side by side. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	75.86	8.81	29.25	74.90	60.29	24.00	66.20	16.27	73.95	40.90	82.60	10.85	46.99	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.12
SHOT	Full-dense	79.98	43.02	62.17	40.24	66.42	51.13	66.99	23.15	80.37	55.55	78.21	35.04	56.86	80.94	71.77	71.28	36.33	75.14	75.86	59.54	43.62	76.52	46.34	76.94	36.73	62.58
SHOT	Candidate-dense	81.65	9.24	49.47	60.34	64.93	20.24	69.69	15.93	77.14	48.01	85.25	22.39	50.35	84.48	10.39	62.62	50.51	70.58	18.51	70.86	17.43	79.07	56.73	85.27	28.88	52.94
SHOT	Random	75.85	**8.81**	29.27	74.87	60.29	23.95	66.24	**16.29**	73.95	40.89	**82.61**	10.86	46.99	75.65	**9.13**	29.42	74.81	60.02	26.31	66.32	**16.02**	73.82	40.64	**82.70**	10.80	47.14
SHOT	Magnitude	75.86	**8.81**	29.25	74.89	60.29	24.00	66.24	16.27	73.95	40.86	82.60	10.85	46.99	75.64	9.12	29.42	74.87	59.97	26.31	66.29	**16.02**	73.82	40.64	82.67	10.71	47.12
SHOT	Saliency	75.92	**8.81**	29.34	**75.10**	60.22	**24.10**	66.27	16.15	**73.99**	41.03	82.39	10.83	47.01	**75.75**	9.09	29.68	**75.10**	60.01	26.36	66.46	15.95	**73.99**	40.60	82.39	10.78	47.18
SHOT	**SDL**	**80.44**	6.85	**68.25**	51.95	**81.97**	16.53	**76.85**	2.50	72.79	**62.03**	80.22	**22.69**	**51.92**	75.59	6.56	**79.25**	48.70	**87.19**	**51.13**	**75.53**	0.38	73.49	**67.78**	79.49	**25.96**	**55.92**
COME	Full-dense	77.56	8.29	23.60	80.49	59.22	21.98	65.94	14.60	75.62	38.76	78.28	8.35	46.06	78.61	7.57	17.53	85.87	58.17	21.16	65.77	12.65	77.18	35.73	72.99	6.40	44.97
COME	Candidate-dense	76.17	8.84	28.29	75.78	60.31	23.95	66.01	16.12	74.24	40.55	82.22	10.31	46.90	76.00	9.09	28.04	76.48	60.16	26.31	65.96	15.93	74.32	39.98	82.08	9.68	47.00
COME	Random	**75.86**	**8.81**	29.22	74.91	**60.29**	24.02	66.22	**16.29**	73.94	**40.90**	**82.60**	10.84	46.99	75.65	**9.13**	**29.42**	74.86	**60.01**	**26.31**	66.31	16.02	**73.84**	**40.66**	82.67	10.70	**47.13**
COME	Magnitude	75.84	**8.81**	**29.25**	74.89	**60.29**	**24.05**	66.23	16.27	**73.97**	**40.90**	**82.60**	**10.85**	**47.00**	75.64	9.12	**29.42**	74.84	59.99	**26.31**	66.32	16.02	**73.84**	40.64	**82.70**	**10.71**	**47.13**
COME	Saliency	**75.86**	**8.81**	29.19	74.92	**60.29**	**24.05**	66.23	16.27	73.95	**40.90**	**82.60**	**10.85**	46.99	**75.67**	9.12	29.40	74.87	59.99	**26.31**	66.30	16.02	73.82	40.60	82.65	**10.71**	47.12
COME	**SDL**	**75.86**	**8.81**	29.19	**74.93**	**60.29**	**24.05**	**66.25**	16.28	73.95	**40.90**	**82.60**	**10.85**	**47.00**	75.65	9.12	29.40	**74.90**	59.99	**26.31**	**66.37**	**16.03**	73.82	40.60	82.65	10.69	**47.13**
Standard OTTA references																											
Tent	DeiT full	28.39	1.29	2.64	98.26	10.34	2.31	10.66	1.07	12.84	6.09	9.82	0.78	15.38	0.55	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	8.38
Tent	Transformer	55.32	4.72	8.68	95.26	29.70	9.20	39.49	3.12	45.42	14.86	45.87	2.81	29.54	22.46	0.83	0.51	99.97	4.22	1.69	16.70	0.00	7.80	0.88	9.66	0.32	13.75
EATA	DeiT full	75.86	8.81	29.23	74.90	60.29	24.05	66.24	16.28	73.99	40.86	82.60	10.87	47.00	75.67	9.12	29.42	74.86	60.01	26.31	66.32	16.02	73.82	40.64	82.65	10.72	47.13
EATA	Transformer	75.86	8.81	29.25	74.90	60.29	24.05	66.22	16.28	73.95	40.90	82.60	10.85	47.00	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.84	40.64	82.65	10.71	47.12
NCTTA	DeiT full	49.34	2.96	7.16	95.99	24.13	7.18	25.97	2.93	28.75	14.16	23.44	1.98	23.67	20.57	0.03	0.09	99.98	1.45	0.34	1.26	0.00	0.75	0.66	0.02	0.00	10.43
NCTTA	Transformer	74.03	8.00	21.41	85.34	53.51	19.95	59.02	10.38	70.70	34.11	74.88	8.35	43.31	70.38	6.79	12.43	94.89	42.38	14.41	50.50	3.65	62.72	25.16	64.97	5.12	37.78

Table 24: VisDA-C class accuracy (%) for the ResNet FC track at $\rho = 0.001$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	57.32	11.19	79.21	65.55	71.41	0.77	86.99	15.30	39.26	68.52	69.14	0.67	47.11	60.89	12.78	81.19	65.10	74.14	0.10	86.30	26.00	37.94	72.91	70.56	0.32	49.02
SHOT	Random	54.18	29.02	49.68	74.76	65.65	4.43	76.57	**21.50**	67.60	46.92	77.56	6.46	47.86	53.68	28.60	50.93	74.99	66.45	4.67	76.42	**21.82**	67.74	48.44	78.25	6.27	48.19
SHOT	Magnitude	54.25	29.01	49.68	74.73	65.70	4.43	76.64	21.48	67.53	**47.04**	77.57	**6.47**	47.88	53.70	28.60	50.96	74.95	66.60	4.67	76.57	21.70	67.66	**48.62**	78.23	**6.33**	48.22
SHOT	Saliency	53.62	27.89	50.30	**76.05**	65.10	4.43	77.59	20.77	66.34	45.73	77.01	5.79	47.55	52.69	26.96	**51.51**	**76.77**	65.40	4.63	**77.98**	20.45	65.18	45.90	77.53	5.30	47.52
SHOT	**SDL**	**54.55**	**38.45**	**50.53**	73.46	**72.42**	**14.94**	**78.11**	20.57	**74.76**	43.75	**78.05**	6.25	**50.49**	**54.80**	**46.33**	50.49	72.93	**75.27**	**32.77**	77.30	16.77	**76.30**	36.56	**78.68**	5.39	**51.97**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	59.57	39.40	69.49	49.99	68.09	34.12	73.72	29.57	62.19	44.06	71.20	24.19	52.13	59.98	44.75	70.28	47.27	70.67	32.77	68.13	33.45	66.87	46.12	68.08	28.23	53.05
COME	Random	54.22	28.82	49.50	74.74	65.48	4.48	76.58	21.44	67.63	46.75	**77.63**	6.47	47.81	53.72	28.53	50.58	**75.03**	66.28	4.79	76.48	21.63	67.69	48.17	**78.34**	6.37	48.13
COME	Magnitude	54.22	**29.07**	49.59	**74.75**	65.68	4.48	76.59	21.45	67.60	46.78	77.60	6.47	47.86	53.73	**28.60**	50.85	74.97	66.45	4.77	76.42	21.55	67.80	48.22	78.28	6.36	48.17
COME	Saliency	53.95	26.27	52.39	74.70	65.38	6.84	**80.12**	19.70	64.32	42.74	77.27	6.67	47.53	53.29	23.83	54.46	74.45	65.76	8.48	**81.85**	18.85	62.21	42.04	77.69	6.78	47.48
COME	**SDL**	**54.42**	28.46	**58.78**	68.00	**67.36**	**9.35**	78.90	**22.30**	**68.67**	**50.28**	75.35	**10.92**	**49.40**	**54.09**	28.20	**66.67**	60.26	**68.11**	**11.04**	80.37	**22.18**	**67.95**	**55.37**	74.55	**16.69**	**50.46**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	FC	12.75	7.71	96.06	8.27	11.70	22.60	24.93	4.95	8.27	8.90	26.16	4.06	19.70	0.00	0.49	95.03	0.78	0.55	0.24	1.22	0.30	0.44	2.32	0.59	0.40	8.53
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	FC	48.57	33.78	80.30	41.45	46.13	62.02	73.78	19.73	35.55	25.91	73.32	18.66	46.60	0.74	12.78	28.12	19.32	10.13	18.46	22.93	10.65	11.10	7.58	30.52	12.78	15.43
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	FC	45.80	34.13	81.30	40.61	41.97	53.93	72.83	17.88	29.44	42.35	71.91	18.69	45.90	0.41	13.35	24.86	17.74	6.22	14.60	18.89	9.55	7.19	28.23	28.02	11.41	15.04

Table 25: VisDA-C class accuracy (%) for the ResNet Conv track at $\rho = 0.001$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	60.83	36.37	77.48	64.70	76.15	1.83	83.16	40.62	58.58	64.27	70.52	13.54	54.01	63.16	58.70	76.78	57.58	77.38	1.30	80.25	54.30	60.83	63.57	74.29	29.65	58.15
SHOT	Random	54.42	**29.02**	**50.03**	74.70	65.78	**4.43**	76.54	21.52	67.83	47.01	**77.38**	**6.45**	47.93	54.00	28.75	**51.25**	74.82	66.58	**4.69**	76.42	21.79	68.19	48.43	**78.10**	**6.37**	48.28
SHOT	Magnitude	54.42	25.81	48.66	75.40	70.47	3.71	78.04	19.73	69.27	**49.10**	76.37	5.05	48.00	54.11	22.45	48.63	75.70	73.50	3.13	79.09	18.48	70.32	**51.34**	76.46	3.71	48.08
SHOT	Saliency	55.02	26.27	49.00	76.02	70.33	3.57	78.17	19.38	**69.40**	48.93	76.02	5.05	48.09	54.58	23.80	49.40	76.78	73.10	3.42	79.12	17.73	70.04	50.28	75.97	3.93	48.18
SHOT	**SDL**	**77.34**	25.52	38.98	**79.82**	**77.77**	2.60	**85.68**	**26.23**	67.55	45.42	66.53	3.30	**49.73**	**84.59**	**30.91**	42.05	**78.33**	**80.15**	2.27	**90.48**	**30.38**	**73.20**	35.69	63.43	2.02	**51.12**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	46.57	12.72	30.49	84.02	57.22	19.61	87.04	8.88	48.03	21.66	65.53	3.19	40.41	34.15	5.53	2.79	93.29	45.64	12.24	85.33	2.88	30.91	10.26	47.69	0.65	30.95
COME	Random	**54.25**	29.05	49.57	74.75	65.79	4.43	76.37	**21.50**	67.80	46.88	77.54	**6.43**	47.86	53.78	28.69	50.64	74.94	66.60	4.67	76.28	21.72	68.18	48.18	78.21	**6.30**	48.18
COME	Magnitude	**54.25**	29.07	49.62	74.78	65.68	4.43	76.59	21.48	67.86	**47.09**	77.41	6.42	47.89	53.81	28.66	50.79	75.02	66.70	4.63	76.54	**21.90**	68.21	**48.66**	78.09	6.22	48.27
COME	Saliency	53.78	28.32	48.12	**75.40**	65.51	4.48	**77.02**	21.00	67.27	46.65	**77.83**	6.33	47.64	53.02	27.57	48.19	**75.87**	66.30	4.82	**77.21**	20.93	67.18	47.48	**78.66**	6.04	47.77
COME	**SDL**	54.14	**32.29**	**54.33**	74.80	**66.51**	**7.28**	74.95	20.55	**71.16**	45.07	75.66	5.73	**48.54**	**53.95**	**35.22**	**58.83**	74.43	**67.23**	**12.00**	73.93	18.80	**75.09**	43.97	73.51	4.49	**49.29**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	Conv	0.00	0.09	0.00	0.42	0.00	9.16	0.43	0.00	0.02	0.00	97.69	0.00	8.98	0.00	0.00	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	98.21	0.00	8.22
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	Conv	0.66	18.50	2.94	5.85	1.17	86.60	31.99	0.73	5.78	0.53	77.48	0.79	19.42	0.00	11.31	8.83	26.15	5.84	35.86	26.98	2.10	15.81	0.00	71.79	3.05	17.31
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	Conv	0.14	11.97	2.05	49.88	0.15	49.35	14.91	0.25	2.53	0.09	40.16	0.11	14.30	0.00	0.00	0.04	99.73	0.04	0.00	0.00	0.00	0.00	0.00	0.02	0.00	8.32

Budget $\rho = 0.002$.

Table 26: VisDA-C class accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.001$. PU and FO are reported side by side. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	75.86	8.81	29.25	74.90	60.29	24.00	66.20	16.27	73.95	40.90	82.60	10.85	46.99	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.12
SHOT	Full-dense	79.98	43.02	62.17	40.24	66.42	51.13	66.99	23.15	80.37	55.55	78.21	35.04	56.86	80.94	71.77	71.28	36.33	75.14	75.86	59.54	43.62	76.52	46.34	76.94	36.73	62.58
SHOT	Candidate-dense	81.65	9.24	49.47	60.34	64.93	20.24	69.69	15.93	77.14	48.01	85.25	22.39	50.35	84.48	10.39	62.62	50.51	70.58	18.51	70.86	17.43	79.07	56.73	85.27	28.88	52.94
SHOT	Random	75.84	8.81	29.27	74.86	60.29	23.98	66.23	**16.28**	73.96	40.89	82.61	10.87	46.99	75.67	9.13	29.44	74.81	60.01	26.31	66.35	**16.05**	73.83	40.64	82.69	10.81	47.15
SHOT	Magnitude	75.84	8.81	29.25	74.91	60.29	24.00	66.22	16.27	73.93	40.86	82.60	10.85	46.98	75.64	9.12	29.42	74.87	59.97	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.13
SHOT	Saliency	76.03	8.78	29.36	**75.02**	60.24	24.05	66.44	15.97	73.97	40.95	82.29	10.87	47.00	76.08	9.06	29.89	**75.08**	59.94	26.31	66.86	15.75	73.80	40.46	82.32	11.01	47.21
SHOT	**SDL**	**83.60**	**34.47**	**61.83**	45.07	**86.57**	**29.30**	**69.88**	1.77	**81.29**	**54.23**	**84.56**	**29.60**	**55.18**	**79.46**	**69.38**	**68.91**	40.30	**89.58**	**79.23**	**69.63**	0.23	**81.47**	**42.83**	**85.65**	**31.31**	**61.50**
COME	Full-dense	77.56	8.29	23.60	80.49	59.22	21.98	65.94	14.60	75.62	38.76	78.28	8.35	46.06	78.61	7.57	17.53	85.87	58.17	21.16	65.77	12.65	77.18	35.73	72.99	6.40	44.97
COME	Candidate-dense	76.17	8.84	28.29	75.78	60.31	23.95	66.01	16.12	74.24	40.55	82.22	10.31	46.90	76.00	9.09	28.04	76.48	60.16	26.31	65.96	15.93	74.32	39.98	82.08	9.68	47.00
COME	Random	75.85	**8.81**	29.21	74.91	**60.29**	23.98	66.22	**16.30**	73.95	40.89	**82.60**	10.83	46.99	**75.67**	**9.13**	**29.42**	74.86	**60.02**	**26.31**	66.33	**16.03**	**73.84**	**40.64**	82.67	10.67	**47.13**
COME	Magnitude	75.84	**8.81**	**29.23**	74.92	**60.29**	**24.05**	66.20	**16.30**	**73.97**	**40.90**	**82.60**	**10.85**	**47.00**	**75.67**	9.12	**29.42**	74.87	59.99	**26.31**	66.30	16.00	**73.84**	**40.64**	**82.70**	**10.71**	**47.13**
COME	Saliency	**75.86**	**8.81**	29.19	74.92	**60.29**	**24.05**	66.23	16.27	73.95	**40.90**	**82.60**	**10.85**	46.99	**75.67**	9.12	29.38	74.90	59.97	**26.31**	66.32	16.02	73.82	40.60	82.65	10.69	47.12
COME	**SDL**	**75.86**	**8.81**	29.19	**74.94**	**60.29**	**24.05**	**66.25**	16.28	73.95	**40.90**	**82.60**	10.83	**47.00**	**75.67**	9.12	29.36	**74.94**	59.97	**26.31**	**66.39**	**16.03**	73.80	40.60	82.63	10.67	47.12
Standard OTTA references																											
Tent	DeiT full	28.39	1.29	2.64	98.26	10.34	2.31	10.66	1.07	12.84	6.09	9.82	0.78	15.38	0.55	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	8.38
Tent	Transformer	55.32	4.72	8.68	95.26	29.70	9.20	39.49	3.12	45.42	14.86	45.87	2.81	29.54	22.46	0.83	0.51	99.97	4.22	1.69	16.70	0.00	7.80	0.88	9.66	0.32	13.75
EATA	DeiT full	75.86	8.81	29.23	74.90	60.29	24.05	66.24	16.28	73.99	40.86	82.60	10.87	47.00	75.67	9.12	29.42	74.86	60.01	26.31	66.32	16.02	73.82	40.64	82.65	10.72	47.13
EATA	Transformer	75.86	8.81	29.25	74.90	60.29	24.05	66.22	16.28	73.95	40.90	82.60	10.85	47.00	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.84	40.64	82.65	10.71	47.12
NCTTA	DeiT full	49.34	2.96	7.16	95.99	24.13	7.18	25.97	2.93	28.75	14.16	23.44	1.98	23.67	20.57	0.03	0.09	99.98	1.45	0.34	1.26	0.00	0.75	0.66	0.02	0.00	10.43
NCTTA	Transformer	74.03	8.00	21.41	85.34	53.51	19.95	59.02	10.38	70.70	34.11	74.88	8.35	43.31	70.38	6.79	12.43	94.89	42.38	14.41	50.50	3.65	62.72	25.16	64.97	5.12	37.78

Table 27: VisDA-C class accuracy (%) for the ResNet FC track at $\rho = 0.002$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and FC scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	57.32	11.19	79.21	65.55	71.41	0.77	86.99	15.30	39.26	68.52	69.14	0.67	47.11	60.89	12.78	81.19	65.10	74.14	0.10	86.30	26.00	37.94	72.91	70.56	0.32	49.02
SHOT	Random	54.18	28.84	49.99	74.74	65.61	4.43	76.75	21.45	67.44	46.87	77.54	6.45	47.86	53.69	28.45	51.31	75.00	66.42	4.67	76.71	21.63	67.50	48.25	78.23	6.18	48.17
SHOT	Magnitude	54.20	28.86	49.96	74.71	65.61	4.43	76.81	21.45	67.42	47.04	**77.55**	6.45	47.88	53.70	28.43	51.24	74.97	66.45	4.67	76.83	21.50	67.55	48.31	**78.26**	6.24	48.18
SHOT	Saliency	53.35	27.34	**51.41**	**76.35**	65.02	4.48	**78.42**	20.23	65.47	44.85	76.82	5.53	47.44	52.19	25.12	**53.75**	**76.82**	65.19	4.58	**79.35**	19.52	63.97	44.72	77.12	4.85	47.27
SHOT	**SDL**	**56.17**	**40.60**	49.27	72.73	**73.86**	**10.22**	76.02	**24.73**	**73.73**	**49.76**	77.27	**7.14**	**50.96**	**58.56**	**50.91**	49.34	70.22	**78.43**	**18.27**	74.50	**27.20**	**74.13**	**49.10**	78.07	**7.10**	**52.98**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	59.57	39.40	69.49	49.99	68.09	34.12	73.72	29.57	62.19	44.06	71.20	24.19	52.13	59.98	44.75	70.28	47.27	70.67	32.77	68.13	33.45	66.87	46.12	68.08	28.23	53.05
COME	Random	54.22	28.52	49.61	74.76	65.49	4.63	76.75	21.35	67.44	46.44	**77.75**	6.49	47.79	53.74	28.16	50.75	75.06	66.18	5.00	76.71	21.45	67.33	47.74	**78.43**	6.39	48.08
COME	Magnitude	54.22	**28.84**	49.45	**74.87**	65.49	4.53	76.57	21.30	67.55	46.56	77.67	6.47	47.79	53.73	28.49	50.51	**75.15**	66.21	4.87	76.54	21.30	67.53	47.92	78.38	6.34	48.08
COME	Saliency	53.92	25.47	53.45	74.23	65.34	8.39	**81.18**	18.90	62.59	41.21	77.19	6.90	47.40	53.46	22.48	56.16	73.59	65.44	11.47	**83.51**	17.93	59.44	39.81	77.69	7.12	47.34
COME	**SDL**	**54.94**	28.72	**58.02**	68.27	**67.30**	**9.73**	78.42	**22.77**	**68.63**	**50.15**	75.94	**10.40**	**49.44**	**54.96**	**29.07**	**64.99**	60.69	**68.15**	**11.76**	79.38	**23.20**	**68.41**	**55.02**	75.52	**15.92**	**50.59**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	FC	12.75	7.71	96.06	8.27	11.70	22.60	24.93	4.95	8.27	8.90	26.16	4.06	19.70	0.00	0.49	95.03	0.78	0.55	0.24	1.22	0.30	0.44	2.32	0.59	0.40	8.53
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	FC	48.57	33.78	80.30	41.45	46.13	62.02	73.78	19.73	35.55	25.91	73.32	18.66	46.60	0.74	12.78	28.12	19.32	10.13	18.46	22.93	10.65	11.10	7.58	30.52	12.78	15.43
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	FC	45.80	34.13	81.30	40.61	41.97	53.93	72.83	17.88	29.44	42.35	71.91	18.69	45.90	0.41	13.35	24.86	17.74	6.22	14.60	18.89	9.55	7.19	28.23	28.02	11.41	15.04

Table 28: VisDA-C class accuracy (%) for the ResNet Conv track at $\rho = 0.002$. PU and FO are reported side by side. Standard OTTA references use the ResNet full and Conv scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	54.22	29.18	49.40	74.76	65.70	4.43	76.38	21.57	67.77	47.04	77.60	6.47	47.88	53.73	28.84	50.43	75.00	66.40	4.67	76.31	21.88	68.02	48.58	78.33	6.36	48.21
SHOT	Full-dense	94.02	60.72	83.86	58.19	88.45	68.24	84.02	62.27	75.56	61.20	82.08	26.60	70.44	94.84	64.23	84.80	61.51	90.39	77.06	86.40	67.90	79.53	64.62	83.03	26.35	73.39
SHOT	Candidate-dense	60.83	36.37	77.48	64.70	76.15	1.83	83.16	40.62	58.58	64.27	70.52	13.54	54.01	63.16	58.70	76.78	57.58	77.38	1.30	80.25	54.30	60.83	63.57	74.29	29.65	58.15
SHOT	Random	54.48	**29.01**	**50.26**	74.71	65.83	4.40	76.50	**21.53**	67.82	47.01	**77.35**	**6.40**	47.94	54.05	**28.77**	**51.66**	74.89	66.87	4.69	76.38	**21.78**	68.13	48.38	**78.02**	**6.29**	48.33
SHOT	Magnitude	54.55	25.41	47.57	75.41	71.35	3.71	78.74	20.98	68.87	**49.72**	76.56	5.73	48.22	54.42	21.58	46.31	75.72	74.67	3.18	79.97	20.90	69.40	**52.04**	76.44	5.17	48.32
SHOT	Saliency	54.80	25.32	48.51	76.59	71.16	3.33	79.28	19.07	68.70	49.01	75.40	4.90	48.01	54.28	21.93	48.57	**77.63**	74.12	2.79	**80.47**	17.07	69.29	50.77	74.83	3.60	47.95
SHOT	**SDL**	**78.14**	19.19	39.66	**76.75**	**80.71**	**37.78**	**80.07**	15.90	**71.22**	42.70	64.12	2.94	**50.77**	**79.02**	9.01	50.58	70.68	**86.21**	**88.67**	79.06	6.25	**83.47**	16.88	53.59	2.50	**52.16**
COME	Full-dense	92.59	58.33	82.41	54.80	86.19	55.76	83.08	51.92	69.69	59.89	82.01	28.48	67.09	94.27	60.23	83.67	58.20	88.83	55.33	85.42	52.98	73.97	61.77	83.57	26.57	68.73
COME	Candidate-dense	46.57	12.72	30.49	84.02	57.22	19.61	87.04	8.88	48.03	21.66	65.53	3.19	40.41	34.15	5.53	2.79	93.29	45.64	12.24	85.33	2.88	30.91	10.26	47.69	0.65	30.95
COME	Random	54.27	29.09	49.72	74.76	65.81	4.43	76.38	21.49	67.85	46.73	77.50	**6.43**	47.87	53.80	28.79	50.90	74.98	66.61	4.71	76.24	**21.70**	68.14	48.09	78.17	**6.25**	48.20
COME	Magnitude	**54.44**	29.01	49.38	74.91	65.76	4.43	76.55	**21.52**	67.95	**47.09**	77.38	6.40	47.90	**53.92**	28.66	50.45	75.27	66.72	4.63	76.54	21.68	68.28	**48.71**	78.07	6.22	48.26
COME	Saliency	53.59	28.00	47.78	**75.82**	65.64	4.77	**77.45**	20.65	67.25	45.90	**77.88**	6.27	47.58	52.52	27.05	47.61	**76.35**	66.66	5.01	**77.88**	19.98	67.07	46.30	**78.61**	6.06	47.59
COME	**SDL**	53.95	**31.89**	**54.71**	74.78	**66.40**	**7.42**	75.43	20.48	**70.61**	44.72	75.59	5.66	**48.47**	53.90	**34.76**	**59.51**	74.31	**67.15**	**12.29**	74.45	18.82	**74.24**	43.66	73.63	4.54	**49.27**
Standard OTTA references																											
Tent	ResNet full	78.30	21.81	74.29	74.33	79.30	94.22	85.73	64.70	66.70	8.33	75.97	1.95	60.47	9.74	0.03	7.14	40.46	16.82	40.29	27.19	14.30	4.51	0.00	14.09	0.00	14.55
Tent	Conv	0.00	0.09	0.00	0.42	0.00	9.16	0.43	0.00	0.02	0.00	97.69	0.00	8.98	0.00	0.00	0.00	0.47	0.00	0.00	0.00	0.00	0.00	0.00	98.21	0.00	8.22
EATA	ResNet full	88.65	56.58	82.32	50.66	80.64	54.27	80.99	49.85	62.45	58.18	81.04	30.16	64.65	11.77	9.12	22.35	14.84	14.90	15.28	15.55	11.47	10.35	12.71	20.25	10.72	14.11
EATA	Conv	0.66	18.50	2.94	5.85	1.17	86.60	31.99	0.73	5.78	0.53	77.48	0.79	19.42	0.00	11.31	8.83	26.15	5.84	35.86	26.98	2.10	15.81	0.00	71.79	3.05	17.31
NCTTA	ResNet full	91.69	31.77	80.90	74.81	89.60	39.66	93.17	36.05	66.61	21.74	88.31	5.57	59.99	18.49	0.06	22.86	33.44	21.27	0.14	33.13	0.35	8.90	0.04	24.72	0.02	13.62
NCTTA	Conv	0.14	11.97	2.05	49.88	0.15	49.35	14.91	0.25	2.53	0.09	40.16	0.11	14.30	0.00	0.00	0.04	99.73	0.04	0.00	0.00	0.00	0.00	0.00	0.02	0.00	8.32

Table 29: VisDA-C class accuracy (%) for the DeiT-S/16 Transformer track at $\rho = 0.002$. PU and FO are reported side by side. Standard OTTA references use the DeiT full and Transformer scopes. Bold values are best among Random, Magnitude, Saliency, and SDL within each objective; dark cells are best over the complete column. SDL rows are shaded.

Objective	Variant	PU													FO											
		Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc	Aero.	Bike	Bus	Car	Horse	Knife	Moto.	Person	Plant	Skate	Train	Truck	mAcc
Controlled adaptation																											
–	Source-only	75.86	8.81	29.25	74.90	60.29	24.00	66.20	16.27	73.95	40.90	82.60	10.85	46.99	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.82	40.64	82.67	10.71	47.12
SHOT	Full-dense	79.98	43.02	62.17	40.24	66.42	51.13	66.99	23.15	80.37	55.55	78.21	35.04	56.86	80.94	71.77	71.28	36.33	75.14	75.86	59.54	43.62	76.52	46.34	76.94	36.73	62.58
SHOT	Candidate-dense	81.65	9.24	49.47	60.34	64.93	20.24	69.69	15.93	77.14	48.01	85.25	22.39	50.35	84.48	10.39	62.62	50.51	70.58	18.51	70.86	17.43	79.07	56.73	85.27	28.88	52.94
SHOT	Random	75.85	8.81	29.27	74.86	60.29	24.02	66.24	**16.32**	73.96	40.95	**82.62**	10.89	47.00	75.72	9.13	29.47	74.79	60.03	26.35	66.33	**16.07**	73.86	**40.68**	**82.70**	10.83	47.16
SHOT	Magnitude	75.84	8.81	29.23	74.90	60.29	24.05	66.22	16.25	73.95	40.86	82.60	10.85	46.99	75.64	9.12	29.42	**74.87**	59.94	26.31	66.32	16.02	73.84	40.64	82.65	10.72	47.13
SHOT	Saliency	76.36	8.66	29.38	**74.91**	60.43	23.81	67.34	15.80	74.06	40.90	82.22	10.90	47.06	76.74	8.81	30.11	74.84	60.58	25.73	**68.93**	15.35	74.10	39.98	82.15	11.18	47.38
SHOT	**SDL**	**86.42**	**39.17**	**73.50**	48.35	**83.16**	**61.69**	**70.46**	1.48	**84.19**	**41.34**	80.71	**27.16**	**58.14**	**90.04**	**71.25**	**73.60**	42.29	**85.78**	**85.83**	54.16	0.05	**85.14**	28.50	80.10	**27.05**	**60.32**
COME	Full-dense	77.56	8.29	23.60	80.49	59.22	21.98	65.94	14.60	75.62	38.76	78.28	8.35	46.06	78.61	7.57	17.53	85.87	58.17	21.16	65.77	12.65	77.18	35.73	72.99	6.40	44.97
COME	Candidate-dense	76.17	8.84	28.29	75.78	60.31	23.95	66.01	16.12	74.24	40.55	82.22	10.31	46.90	76.00	9.09	28.04	76.48	60.16	26.31	65.96	15.93	74.32	39.98	82.08	9.68	47.00
COME	Random	**75.89**	**8.81**	29.21	74.95	**60.29**	**24.06**	66.21	16.27	**73.99**	**40.95**	**82.60**	10.81	**47.01**	**75.71**	**9.13**	29.38	74.92	**60.04**	**26.34**	66.29	16.02	**73.88**	**40.67**	82.64	10.62	**47.14**
COME	Magnitude	75.84	**8.81**	**29.25**	74.90	**60.29**	24.05	66.23	16.27	73.95	40.90	**82.60**	**10.85**	46.99	75.64	9.12	**29.42**	74.84	59.99	26.31	66.30	16.02	73.84	40.64	**82.70**	**10.71**	47.13
COME	Saliency	75.86	**8.81**	29.19	74.92	**60.29**	24.05	66.23	16.27	73.97	40.90	**82.60**	10.81	46.99	75.67	9.12	29.36	74.94	59.97	26.31	66.34	16.00	73.80	40.60	82.62	10.65	47.12
COME	**SDL**	**75.89**	**8.81**	29.13	**74.97**	**60.29**	23.95	**66.25**	**16.28**	73.93	40.90	**82.60**	10.82	46.98	75.67	9.12	29.30	**74.98**	59.97	26.31	**66.37**	**16.03**	73.80	40.55	82.63	10.65	47.12
Standard OTTA references																											
Tent	DeiT full	28.39	1.29	2.64	98.26	10.34	2.31	10.66	1.07	12.84	6.09	9.82	0.78	15.38	0.55	0.00	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	8.38
Tent	Transformer	55.32	4.72	8.68	95.26	29.70	9.20	39.49	3.12	45.42	14.86	45.87	2.81	29.54	22.46	0.83	0.51	99.97	4.22	1.69	16.70	0.00	7.80	0.88	9.66	0.32	13.75
EATA	DeiT full	75.86	8.81	29.23	74.90	60.29	24.05	66.24	16.28	73.99	40.86	82.60	10.87	47.00	75.67	9.12	29.42	74.86	60.01	26.31	66.32	16.02	73.82	40.64	82.65	10.72	47.13
EATA	Transformer	75.86	8.81	29.25	74.90	60.29	24.05	66.22	16.28	73.95	40.90	82.60	10.85	47.00	75.64	9.12	29.42	74.84	59.99	26.31	66.30	16.02	73.84	40.64	82.65	10.71	47.12
NCTTA	DeiT full	49.34	2.96	7.16	95.99	24.13	7.18	25.97	2.93	28.75	14.16	23.44	1.98	23.67	20.57	0.03	0.09	99.98	1.45	0.34	1.26	0.00	0.75	0.66	0.02	0.00	10.43
NCTTA	Transformer	74.03	8.00	21.41	85.34	53.51	19.95	59.02	10.38	70.70	34.11	74.88	8.35	43.31	70.38	6.79	12.43	94.89	42.38	14.41	50.50	3.65	62.72	25.16	64.97	5.12	37.78