EXEMPLAR-FREE FEW-SHOT CLASS-INCREMENTAL LEARNING WITH FROZEN STATISTICS

Anonymous authors
Paper under double-blind review

ABSTRACT

Large-scale pretrained visual encoders have substantially raised the representation ceiling for few-shot class-incremental learning (FSCIL) on images and point clouds, allowing new classes to be learned from a few samples while old ones are retained. Yet many existing methods still replay raw exemplars, raising privacy concerns, or add prompts, auxiliary modules, or per-session gradient updates that increase complexity and overwrite previously learned knowledge. We show that exemplar-free FSCIL needs none of these: once the pretrained encoder is frozen, compact class statistics suffice. We propose ***FROzen STatistics*** (FROST), which replaces exemplar replay with two compact, complementary memories updated analytically in each session. A correlation memory accumulates sufficient statistics over features for a closed-form ridge solution, which is discriminative and needs no backpropagation but is biased toward data-rich base classes. A prototype memory keeps one normalised mean per class, so every class counts equally. Fusing the two combines discrimination with class balance and restores parity between base and novel classes. Across four point-cloud and three image benchmarks, FROST matches its exemplar-based variant and outperforms prior methods, most notably by **+14.7** points on CO3D and **+14.9** points on FGVC-Aircraft in terms of last-session accuracy. These results indicate that a dual-memory design over frozen features suffices for competitive exemplar-free FSCIL.

1 INTRODUCTION

Open-world perception in robotics, autonomous driving, and augmented reality continually encounters new object categories in both images and point clouds (Xiao et al., 2024; Yue et al., 2018; Thengane et al., 2025; Zhang et al., 2025), often under scarce supervision. Few-shot class-incremental learning (FSCIL) formalises this setting (Tao et al., 2020; Masana et al., 2022): after a data-rich base session, a model must absorb novel classes from only a handful of labelled examples per class, retain knowledge of all previously learned categories, and strictly never revisit raw past data. Large-scale pretrained encoders have raised the representation ceiling in 2D and 3D (Radford et al., 2021; Zhou et al., 2024b), yet turning a frozen feature map into a stable incremental classifier under few-shot novel supervision remains unresolved.

Despite strong pretrained backbones, existing FSCIL methods for images (Zhou et al., 2022; Belouadah & Popescu, 2020) and point clouds (Chowdhury et al., 2022; Tan & Xiang, 2024) still organise around weaker choices. Replay-based pipelines store raw exemplars to stabilise old classes, but this collides with privacy regulation and storage limits that often forbid revisiting past data. Exemplar-free alternatives instead introduce prompts, geometric auxiliary heads, or per-session gradient updates on adapters and classifiers (Cheraghian et al., 2024; Ahmadi et al., 2024; Li et al., 2025; Jiang et al., 2025); under few-shot increments these updates overfit the current batch and overwrite earlier-session knowledge. Neither route uses the structural opportunity of modern pretrained encoders: once representations are strong and stationary, new classes can be registered without raw replay or rewriting previous weights.

This opportunity is naturally realised by a closed-form ridge solution (Hoerl & Kennard, 1970; Rahimi & Recht, 2007; McDonnell et al., 2023): streaming feature correlations over a frozen embedding determine an exact linear head without optimisation. Under FSCIL, however, those unnormalised correlations inherit the base session's sample abundance, so the resulting logits sys-

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.

tematically favour data-rich base classes over scarce novel ones. We propose ***FROzen STatistics*** (FROST) to resolve this tension without replay, additional modules, or per-session optimisation. After base-session adaptation, the encoder is frozen and every new class is registered only through two complementary memories updated analytically during each session. A correlation memory accumulates feature correlations into a closed-form ridge head, preserving past-class discrimination by additive updates, yet remains biased toward data-rich base classes. A prototype memory corrects that class-mass bias with one normalised mean per class, balancing data-rich base and few-shot novel contributions. Fusing their logits couples ridge discrimination with prototype balance and restores base-novel parity.

This dual-memory design is validated across four point-cloud benchmarks (ModelNet40, ShapeNet, ScanObjectNN, CO3D) under in-domain and cross-domain protocols, and three image benchmarks (CUB-200, iNF200, FGVC-Aircraft). Exemplar-free FROST not only matches an exemplar-based counterpart, but also improves over prior methods on hard settings (including **+14.7** points on in-domain CO3D and **+14.9** points on FGVC-Aircraft in last-session accuracy) and leads all three cross-domain 3D transfer protocols. These results show that complementary class statistics on a frozen encoder suffice for strong exemplar-free FSCIL in both 2D and 3D.

Our contributions are threefold. First, we introduce FROST, an exemplar-free FSCIL framework that freezes an adapted pretrained encoder and registers new classes through complementary class statistics, without replay, prompt tuning, or incremental weight updates. Second, we pair a correlation memory for a closed-form ridge solution with a prototype memory of normalised class means; their fusion corrects base-class mass bias and restores base-novel parity. Third, across four point-cloud and three image benchmarks, including cross-domain 3D transfer, exemplar-free FROST matches an exemplar counterpart and substantially outperforms prior methods across both modalities.

2 Related Work

2.1 Pre-trained Models

Transferable visual representations for images and point clouds have progressed along a shared trajectory from specialised architectures to large-scale pretrained encoders. Image recognition matured through large-scale supervised pretraining (Deng et al., 2009) and vision transformers (Dosovitskiy et al., 2021). In parallel, 3D perception moved from multi-view images (Su et al., 2015), voxels (Maturana & Scherer, 2015), and meshes (Masci et al., 2015) to PointNet (Qi et al., 2017a) and PointNet++ (Qi et al., 2017b), then continuous convolutions (Li et al., 2018; Wu et al., 2019), dynamic graphs (Wang et al., 2019), and point transformers (Zhao et al., 2021; Wu et al., 2024). Large-scale foundation pretraining further raises the feature ceiling across modalities, including image-language models such as CLIP (Radford et al., 2021) and 3D multi-modal encoders such as Uni3D (Zhou et al., 2024b). These models transfer strongly under a fixed taxonomy, yet leave the open-world problem of label growth unresolved. Across both modalities, they remain closed-world by default: strong features alone provide no mechanism to absorb few-shot novel classes without rewriting established decision boundaries.

2.2 Few-Shot Class-Incremental Learning

FSCIL requires absorbing novel classes from few-shot supervision while retaining all previously observed labels (Tao et al., 2020; Masana et al., 2022). In 2D, early methods preserve prior knowledge through topology constraints (Tao et al., 2020), distillation (Li & Hoiem, 2017; Castro et al., 2018), classifier recalibration (Belouadah & Popescu, 2019; 2020), and subspace regularisation (Akyürek et al., 2022). Later designs expand few-shot plasticity with evolving classifiers or soft subnetworks (Zhang et al., 2021; Kang et al., 2023), prototype and margin refinements (Zhu et al., 2021; Zou et al., 2022), and forward-compatible or boosted feature allocation (Zhou et al., 2022; Wang et al., 2022a; Kim et al., 2023). In 3D, Micro. (Chowdhury et al., 2022) formalised benchmark protocols and synthetic-to-real transfers; Sem. (Cheraghian et al., 2021a) injects language-space class semantics via distillation and attention, following earlier zero-shot studies (Cheraghian et al., 2021b); Cross-Domain (Tan & Xiang, 2024) balances old and new classes under domain shift with soft- and hard-label replay. Across both modalities, a recurring assumption is that task-specific pa-

rameters must be updated from tiny labelled batches, which encourages overfitting and dependence on raw exemplar buffers or heavy regularisation.

2.3 Pre-trained Models and Analytical Solutions for FSCIL

Pretrained foundation models elevate the few-shot performance ceiling in both images and point clouds (Awais et al., 2023; Thengane et al., 2025), yet most FSCIL pipelines still couple those encoders with incremental adaptation. In 2D, prompt- and pool-based methods adapt frozen ViT or vision-language backbones with session-wise prompt updates or related optimisations (Jiang et al., 2025; Huang et al., 2024; Yoon et al., 2023; Park et al., 2024). In 3D, C3PR (Cheraghian et al., 2024) transfers CLIP (Radford et al., 2021) through multi-view depth rendering; FM (Ahmadi et al., 2024) adapts Uni3D (Zhou et al., 2024b) with prompts and dual caching; FLEG (Li et al., 2025) adds geometric branches to reduce language dependence. Across both modalities, these systems still rely on per-session prompts, auxiliary heads, or gradient updates that can overfit scarce novel batches and overwrite earlier knowledge. An alternative line freezes the backbone and updates only a closed-form analytical head. RanPAC (McDonnell et al., 2023) shows that random-projection ridge regression (Hoerl & Kennard, 1970; Rahimi & Recht, 2007) supports continual learning on frozen foundation embeddings in a modality-agnostic manner. Under FSCIL, however, unnormalised ridge statistics favour data-rich base classes, and few-shot retuning is unreliable. FROST corrects both with a dual-memory design that fuses ridge scores and normalised prototypes.

3 Methodology

3.1 Problem Formulation

Few-shot class-incremental learning (FSCIL) (Tao et al., 2020; Masana et al., 2022) considers a sequence of sessions $\mathcal{D} = \{\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_T\}$. Each session is a labelled set $\mathcal{D}_t = \{(\mathbf{x}_i^t, y_i^t)\}_{i=1}^{N_t}$, where the input $\mathbf{x}_i^t$ is a point cloud in the 3D setting or an image in the 2D setting, and $y_i^t \in \mathcal{C}^t$ is its class label. Session label sets are disjoint, $\mathcal{C}^i \cap \mathcal{C}^j = \emptyset$ for $i \neq j$, and the cumulative label space after session t is $\hat{\mathbf{C}}^t = \bigcup_{i=0}^{t} \mathcal{C}^i$. The base session ($t$=0) provides abundant supervision for every class in $\mathcal{C}^0$. Each subsequent novel session (t>0) introduces new classes under a strict K-shot budget: every class $c \in \mathcal{C}^t$ is observed through only K labelled examples, with $K \ll N_0/|\mathcal{C}^0|$. Training is strictly sequential and exemplar-free unless stated otherwise: when learning session t, the model may use only $\mathcal{D}_t$ and must not revisit raw samples from $\mathcal{D}_0, \ldots, \mathcal{D}_{t-1}$. After session t, evaluation is class-incremental over $\hat{\mathbf{C}}^t$ with no task identity at test time. The objective is to absorb novel classes under K-shot supervision while retaining all previously learned categories.

3.2 Method Overview

Fig. 1 illustrates the framework of FROST. The design principle is to keep the feature map stationary after a base-session alignment, then register every class through compact statistics rather than raw exemplars or incremental fine-tuning. Concretely, we align the pretrained model f_θ with the downstream data via a lightweight adapter over a base session; the adapted model f'_θ is then kept frozen throughout (Sec. 3.3). On this map, FROST maintains two complementary memories. CorrMem $(\mathbf{G}, \mathbf{Q})$ accumulates feature correlations into a closed-form ridge head that preserves past-class discrimination by additive updates, but inherits a base-class mass bias under FSCIL (Sec. 3.4). ProtoMem $\mathcal{M}_c = \{\mathbf{p}\}_c$ counters that bias with one normalised mean per class, so base and novel categories contribute symmetrically at scoring time (Sec. 3.5). Base data encoded by f'_θ initialise both memories; each novel session updates them analytically from K-shot data alone; and a fixed convex fusion of CorrMem and ProtoMem logits restores base-novel parity at inference (Sec. 3.6).

3.3 Alignment Adapter

We align the pretrained encoder f_θ to the downstream distribution by inserting a lightweight bottleneck adapter ϕ with inner width $d' \ll d$, without modifying the pretrained weights. Only ϕ is trained on the base session; the resulting adapted encoder and working representation are

$$f'_\theta = \phi \circ f_\theta, \qquad \mathbf{z} = f'_\theta(\mathbf{x}) \in \mathbb{R}^d, \tag{1}$$

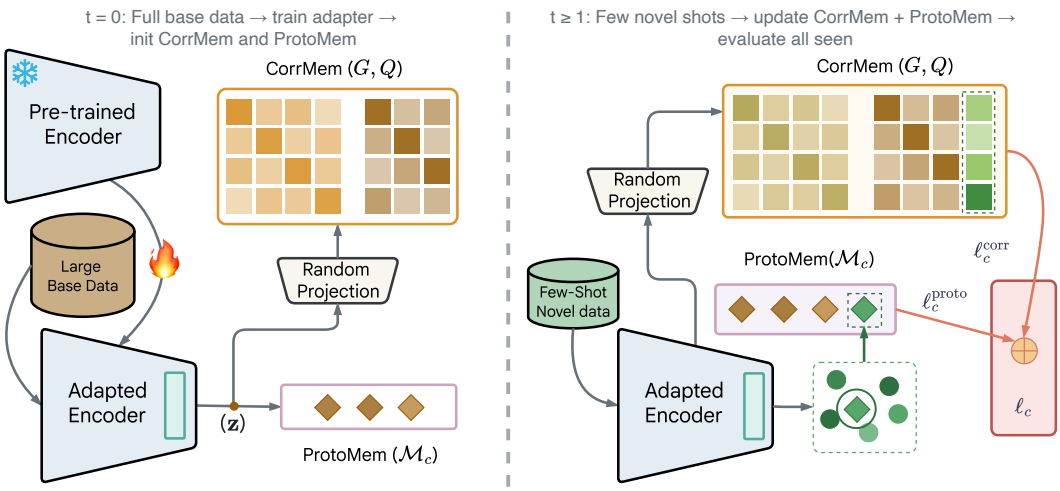

Figure 1: **Overview of FROST.** At session $t{=}0$, a base-only adapter is trained on the pretrained encoder and then frozen; the same base features initialise CorrMem $(\mathbf{G}, \mathbf{Q})$ and ProtoMem $\mathcal{M}_c$. For each novel session $t{\geq}1$, CorrMem accumulates K-shot statistics and expands $\mathbf{Q}$, while ProtoMem appends novel-class means. CorrMem and ProtoMem logits are fused to yield the final scores ℓ_c.

where f_θ maps inputs to $\mathbb{R}^d$ and $\phi : \mathbb{R}^d \to \mathbb{R}^d$. After the base session, f'_θ is frozen for the remainder of the sequence, so incremental sessions never rewrite encoder or adapter parameters and the feature map stays stationary for analytical memory updates. The bottleneck focuses capacity on task-relevant directions in $\mathbb{R}^{d'}$ while preserving the geometric (or visual) priors in f_θ, yielding a cheap base-session alignment that transfers to scarce novel classes under this fixed representation.

3.4 CorrMem: Global Correlation Memory

Closed-form ridge solution realises the structural opportunity of a frozen encoder: streaming feature correlations determine an exact linear head without optimisation (Rahimi & Recht, 2007; McDonnell et al., 2023). This avoids the failure mode of gradient-based heads, which rewrite shared classifier parameters and overwrite past decision boundaries under sequential updates (Masana et al., 2022). Because f'_θ already yields a stationary $\mathbf{z}$, streaming correlations can be accumulated without revisiting past inputs. To enlarge the linear capacity of the ridge head beyond embedding width d, we lift $\mathbf{z}$ into a fixed higher-dimensional space $\mathbb{R}^V$ with $V{>}d$ through a frozen random feature map:

$$\mathbf{h}(\mathbf{z}) = \sigma\big(W_{\text{rand}}^\top \mathbf{z}\big) \in \mathbb{R}^V, \tag{2}$$

where $W_{\text{rand}} \in \mathbb{R}^{d\times V}$ is drawn once from $\mathcal{N}(0,1)$ and kept frozen, and σ is an elementwise ReLU. With both f'_θ and W_{rand} frozen, the lifted map $\mathbf{h}$ remains identical across sessions and admits exact accumulation into additive sufficient statistics. Henceforth, for any input sample i, we write $\mathbf{h}_i \equiv \mathbf{h}(\mathbf{x}_i) \equiv \mathbf{h}(\mathbf{z}_i)$ interchangeably for brevity.

Statistics accumulation. To prevent past information loss during sequential learning, FROST summarises past experience using two streaming statistics, collectively termed *CorrMem* $(\mathbf{G}_t, \mathbf{Q}_t)$:

$$\mathbf{G}_t \in \mathbb{R}^{V\times V}, \qquad \mathbf{Q}_t \in \mathbb{R}^{V\times|\hat{\mathbf{C}}^t|}, \tag{3}$$

where $|\hat{\mathbf{C}}^t|$ denotes the total number of classes observed up to task t. The two matrices play complementary roles. First, $\mathbf{G}_t$ is *class-agnostic*: it accumulates the second-moment feature correlations and maintains a fixed $V \times V$ dimension across all tasks. Second, $\mathbf{Q}_t$ is *class-aware*: it is a feature-label accumulation matrix whose c-th column aggregates the mapped features $\mathbf{h}$ of class c. Consequently, $\mathbf{Q}_t$ expands by one column per novel class while past columns remain unchanged. Together, $(\mathbf{G}_t, \mathbf{Q}_t)$ fully determine the closed-form ridge solution, so past-class discrimination is preserved by additive updates rather than by storing raw exemplars or iterative tuning.

At $t{=}0$, CorrMem is initialised by encoding base data $\mathcal{D}_0$ with the frozen adapted encoder f'_θ and accumulating the lifted features into $(\mathbf{G}_0, \mathbf{Q}_0)$. At task $t{\geq}1$, given the current session $\mathcal{D}_t$, CorrMem updates by accumulation:

$$\mathbf{G}_t \leftarrow \mathbf{G}_{t-1} + \sum_{(\mathbf{x}_i, y_i)\in\mathcal{D}_t} \mathbf{h}_i\,\mathbf{h}_i^\top, \qquad \mathbf{Q}_t \leftarrow \mathbf{Q}_{t-1} + \sum_{(\mathbf{x}_i, y_i)\in\mathcal{D}_t} \mathbf{h}_i\,\mathbf{e}_{y_i}^\top, \tag{4}$$

where $\mathbf{e}_{y_i} \in \{0,1\}^{|\hat{\mathbf{C}}^t|}$ is the one-hot encoding of y_i, and $\mathbf{Q}_{t-1}$ is zero-padded for classes introduced at task t. The pair $(\mathbf{G}_{t-1}, \mathbf{Q}_{t-1})$ encodes the full contribution of all previous tasks in CorrMem, so no raw past input need be retained. For comparison with exemplar-based baselines we also report a configuration in which $\mathcal{D}_t$ is augmented with an exemplar set $\mathcal{E}_t$, i.e. $\mathcal{D}_t \leftarrow \mathcal{D}_t \cup \mathcal{E}_t$; unless otherwise stated, FROST does not store any past exemplars.

Closed-form solution. Unlike the incremental setup of RanPAC (McDonnell et al., 2023), under FSCIL per-task retuning of penalty λ is unreliable: novel sessions provide only K shots per class, so validation is dominated by sampling noise. We instead select λ once on a 20% base hold-out set and keep it fixed throughout sessions. With λ fixed, the CorrMem weights W_t^{corr} at task t are

$$W_t^{\mathrm{corr}} = (\mathbf{G}_t + \lambda\mathbf{I})^{-1}\mathbf{Q}_t \in \mathbb{R}^{V\times|\hat{\mathbf{C}}^t|}, \tag{5}$$

where $\mathbf{I} \in \mathbb{R}^{V\times V}$ is the identity. For a query with embedding $\mathbf{z}$, CorrMem classifier logits are

$$\ell_c^{\mathrm{corr}}(\mathbf{z}) = \left[\mathbf{h}(\mathbf{z})^\top W_t^{\mathrm{corr}}\right]_c, \qquad c \in \hat{\mathbf{C}}^t. \tag{6}$$

This projection scales linearly with $|\hat{\mathbf{C}}^t|$ and requires no iterative optimisation at test time.

By constructing W_t^{corr} from $(\mathbf{G}_t, \mathbf{Q}_t)$ in closed form, CorrMem is discriminative and needs no backpropagation: the analytical head does not overwrite a shared parametric classifier. However, because $\mathbf{Q}_t$ accumulates unnormalised features, those correlations inherit the base session's sample abundance, so base classes with large n_c contribute far more mass than K-shot novel classes. This *class-mass imbalance* biases the analytical solver toward data-rich base classes at inference. A simple rescaling of each column of $\mathbf{Q}_t$ by $1/n_c$ is brittle under K-shot noise because novel-class columns become high-variance estimates while $\mathbf{G}_t$ remains base-dominated.

3.5 ProtoMem: Local Prototype Memory

We therefore introduce ProtoMem $\mathcal{M}_t$ to counter that bias outside the ridge solution: every observed class $c \in \hat{\mathbf{C}}^t$ is registered with a single ℓ_2-normalised mean $\mathbf{p}_c$, so base and novel categories contribute symmetrically at scoring time without raw exemplars and without rewriting W_t^{corr}.

Base initialisation. Similar to CorrMem, ProtoMem is seeded from base data encoded by the frozen adapted encoder f'_θ:

$$\mathcal{M}_0 = \{\mathbf{p}_c \mid c \in \mathcal{C}^0\}, \qquad \mathbf{p}_c = \mathrm{norm}(\bar{\mathbf{z}}_c), \tag{7}$$

where $\bar{\mathbf{z}}_c$ is the mean embedding of class c in $\mathcal{D}_0$. A single mean per class is preferred over multi-key caches: it is compact, and additional support features yield only marginal gains in our setting.

Incremental update. At each novel session $t{\geq}1$, task t introduces classes $\mathcal{C}^t \subset \hat{\mathbf{C}}^t$ with K support samples each from $\mathcal{D}_t$. Because f'_θ is frozen, past prototypes remain valid and need no re-encoding; ProtoMem only appends the new centroids:

$$\mathcal{M}_t \leftarrow \mathcal{M}_{t-1} \cup \left\{\mathbf{p}_c \mid c \in \mathcal{C}^t\right\}, \tag{8}$$

with $\mathbf{p}_c = \mathrm{norm}(\bar{\mathbf{z}}_c)$ computed from the K-shot support of class c. Past entries are left unchanged, so ProtoMem grows by $|\mathcal{C}^t|$ means per session while retaining all previously registered classes.

ProtoMem scoring. Inference matches a query embedding $\mathbf{z}$ to ProtoMem by cosine similarity, which realises the intended class-volume invariance:

$$\ell_c^{\mathrm{proto}}(\mathbf{z}) = \cos(\mathbf{z},\ \mathbf{p}_c), \qquad c \in \hat{\mathbf{C}}^t,\ \mathbf{p}_c \in \mathcal{M}_t. \tag{9}$$

Because every $\mathbf{p}_c$ is an ℓ_2-normalised centroid, these scores are invariant to sample volume n_c: a base class with abundant data and a novel class with K shots enter the score on the same footing.

3.6 Memory Fusion

Neither memory is sufficient alone: CorrMem underweights novel classes through class-mass imbalance, while ProtoMem balances classes but lacks the global discrimination of the analytical ridge head. We therefore fuse their logits with a fixed convex combination

$$\ell_c(\mathbf{z}) = \alpha\, \ell_c^{\mathrm{corr}}(\mathbf{z}) + (1-\alpha)\, \ell_c^{\mathrm{proto}}(\mathbf{z}), \tag{10}$$

where $\alpha \in [0,1]$ is a fixed fusion weight that controls how much the decision trusts CorrMem versus ProtoMem. The fused score couples CorrMem discrimination with ProtoMem volume invariance and restores base-novel parity at inference, without raw exemplar replay, prompt tuning, or incremental fine-tuning after the base session.

4 Experiments

4.1 Experimental Setup

Datasets and protocols. We evaluate FROST on point clouds and images under matched K=5-shot incremental schedules, covering in-domain retention, synthetic-to-real transfer, and cross-modal reuse of the same frozen dual-memory recipe. For point clouds we use ModelNet40 (Wu et al., 2015) and ShapeNet (Chang et al., 2015) (synthetic) with ScanObjNN (Uy et al., 2019) and CO3D (Reizenstein et al., 2021) (real scans), following the base-novel splits of Chowdhury et al. (2022). *In-domain* runs keep base and novel classes in one dataset (ModelNet40, ShapeNet, CO3D); *cross-domain* runs transfer a synthetic base session to real-scan novel classes (ModelNet→ScanObjNN, ShapeNet→ScanObjNN, ShapeNet→CO3D). For images we use CUB-200 (Wah et al., 2011), FGVC-Aircraft (Maji et al., 2013), and iNF200 (Van Horn et al., 2018) under the non-replay protocol of Jiang et al. (2025). Unless noted, results use the exemplar-free (EF) setting; an exemplar-based (EB) variant stores one sample from each previous class for comparison with prior methods. Full class counts and session schedules are listed in Appendix A.1.

Evaluation metrics. Following prior work (Chowdhury et al., 2022; Jiang et al., 2025), we report base-session accuracy $\mathcal{A}_b$ after task 0, last-session accuracy $\mathcal{A}_T$ on all classes observed by the final session, and average accuracy $\bar{\mathcal{A}} = \frac{1}{T+1}\sum_{t=0}^{T}\mathcal{A}_t$ over all sessions including the base session. Relative drop is $\Delta = 100 \times |\mathcal{A}_T - \mathcal{A}_b| \,/\, \mathcal{A}_b$. Because $\mathcal{A}_T$ and Δ can favour a large number of base classes, we also report the harmonic mean $\mathcal{H} = 2\,(\mathcal{A}_b\,\mathcal{A}_n) \,/\, (\mathcal{A}_b + \mathcal{A}_n)$ of final base- and novel-class accuracies. Unless stated otherwise, metrics are averages over five random seeds.

Implementation details. The dual-memory recipe is shared across modalities; only the frozen encoder and base-session adapter schedule change. Point-cloud runs use Uni3D-Base (Zhou et al., 2024b) with M=1024 FPS inputs; image runs use ViT-B/16 pretrained on ImageNet-21K with 224×224 inputs under the non-replay protocol of Jiang et al. (2025). In both cases a lightweight adapter ϕ (d'=64) is trained with SGD on the base session only and then frozen with the encoder; every novel session updates CorrMem and ProtoMem analytically, with no prompt tuning and no incremental gradient steps. We select the ridge penalty λ once on a 20% base hold-out, set the fusion weight α to 0.5, and set the CorrMem lift V to 2000. All runs use a single NVIDIA A100 (64 GB); further details on hyperparameters and preprocessing are provided in Appendix A.2.

Comparators. Reference baselines include Fine-Tuning (FT) without an explicit retention mechanism and an Oracle jointly trained on all base and novel classes. On point clouds we compare against classical 2D FSCIL methods adapted to 3D (LwF (Li & Hoiem, 2017), IL2M (Belouadah & Popescu, 2019), ScaIL (Belouadah & Popescu, 2020), EEIL (Castro et al., 2018), FACT (Zhou et al., 2022), Sem. (Cheraghian et al., 2021a)), dedicated 3D methods (Micro. (Chowdhury et al., 2022), C3PR (Cheraghian et al., 2024)), and foundation-model pipelines (FM (Ahmadi et al., 2024), FLEG (Li et al., 2025)). On images, under the non-replay protocol of Jiang et al. (2025), we report rehearsal and boosting methods (iCaRL (Rebuffi et al., 2017), FOSTER (Wang et al., 2022a)), recent FSCIL designs (FACT, Comp-FSCIL (Zou et al., 2024), NC-FSCIL (Yang et al., 2023), TEEN (Wang et al., 2023), RDI (Zhou et al., 2024a), Yourself (Tang et al., 2024), Approx. (Wang et al., 2024)), and prompt-based learners (L2P (Wang et al., 2022c), DualPrompt (Wang et al., 2022b), CODA-P (Smith et al., 2023), PriViLege (Park et al., 2024), LGSP-P (Jiang et al., 2025)).

Table 1: Comparison on in-domain protocols in terms of base-class accuracy $\mathcal{A}_b$, final-session accuracy $\mathcal{A}_T$, average accuracy $\bar{\mathcal{A}}$, and relative drop $\Delta \downarrow$. Best results are in **bold**.

Method	Venue	ModelNet				ShapeNet				CO3D			
		$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$
Oracle	–	89.8	80.5	85.8	10.4	87.0	81.3	83.6	6.6	76.7	59.8	65.7	22.0
FT	–	89.8	3.0	22.0	96.7	87.0	0.4	17.5	99.5	76.7	0.8	16.2	99.0
LwF	TPAMI'17	89.8	3.1	28.3	96.5	87.0	1.8	29.4	97.9	76.7	1.0	17.2	98.7
IL2M	ICCV'19	89.8	53.6	63.9	40.3	87.0	49.3	54.4	43.3	76.7	21.9	33.8	71.4
ScaIL	WACV'20	89.8	56.5	67.3	37.1	87.0	45.4	54.5	47.8	76.7	27.5	38.7	64.1
EEIL	ECCV'18	89.8	55.6	69.6	38.1	87.0	65.8	72.2	24.4	76.7	32.8	50.9	57.2
FACT	CVPR'22	90.4	65.0	77.5	28.1	87.5	62.5	71.4	28.6	77.9	46.7	59.4	40.0
Semantic	CVPR'21	91.3	64.7	76.5	29.1	87.2	63.8	71.1	26.8	78.6	42.9	58.4	45.4
Microshape	ECCV'22	93.6	67.1	79.6	28.3	87.6	72.6	79.2	17.1	78.5	47.2	60.1	39.9
C3PR	ECCV'24	91.6	70.9	78.6	22.5	88.0	74.7	78.8	15.1	81.5	53.8	64.7	34.0
FLEG	ICML'25	**98.3**	87.3	93.2	11.2	**94.3**	83.0	88.9	12.0	**82.1**	57.3	68.2	30.2
FROST*	Ours (EB)	95.8	**91.5**	**93.5**	**4.5**	92.2	**89.3**	**90.2**	**3.1**	81.6	71.9	75.3	11.9
FROST	Ours (EF)	95.8	91.3	93.4	4.7	92.2	89.2	90.1	3.3	81.6	**72.0**	**75.3**	**11.8**

4.2 Experimental Comparisons

We compare on in-domain and cross-domain point-cloud protocols, then transfer the same frozen dual-memory recipe to image FSCIL under the non-replay setting. Unless noted, results are exemplar-free (EF), denoted FROST; FROST* stores one exemplar per old class (EB) for comparison with prior methods.

In-domain comparison. Tab. 1 reports in-domain results on ModelNet, ShapeNet, and CO3D against classical continual learners, prior 3D FSCIL methods, and a joint-training Oracle. Both EB and EF FROST lead all published competitors on $\mathcal{A}_T$, $\bar{\mathcal{A}}$, and Δ across the three datasets, and also surpass the Oracle on final accuracy and relative drop. On ModelNet, EB / EF reach 91.5% / 91.3% final accuracy versus FLEG's previous best of 87.3% (**+4.2 / +4.0**), reducing Δ from 11.2% to 4.5% / 4.7%. On ShapeNet the same ordering holds: 89.3% / 89.2% against FLEG's 83.0%, with Δ falling from 12.0% to 3.1% / 3.3% (**8.9 / 8.7** points lower). On CO3D, the hardest in-domain setting, the margin over the previous best widens: 71.9% / 72.0% versus FLEG's 57.3% (**+14.6 / +14.7**), with Δ reduced from 30.2% to 11.9% / 11.8%. FLEG records the highest base-session accuracy $\mathcal{A}_b$ on every dataset, but that base-class over-specialisation coincides with larger subsequent drops. Dual-memory fusion over frozen features instead accepts a modest base-session trade-off for stronger retention and novel-class plasticity; the near-identical EF and EB numbers indicate that CorrMem and ProtoMem already recover the gain usually attributed to raw exemplar buffers.

Cross-domain comparison. Tab. 2 reports cross-domain results on ShapeNet→CO3D, ModelNet→ScanObjNN, and ShapeNet→ScanObjNN against classical continual learners, prior 3D FSCIL methods, and a joint-training Oracle. Mirroring the in-domain findings, both EB and EF FROST improve final and average accuracy while reducing relative drop versus all published competitors. On ShapeNet→CO3D, EB reaches 82.9% final accuracy with a 10.2% relative drop, improving over FLEG (76.8%) by **+6.1** points and over Oracle (67.3%) by **+15.6** points; EF remains close at 82.5% with a 10.6% drop. On ModelNet→ScanObjNN, both variants reach 87.9% final accuracy with a 5.3% relative drop, improving over the previous best FLEG (86.8%, 7.5% drop) and over FM (79.2%, 9.7% drop). On ShapeNet→ScanObjNN, EB / EF lead the full comparator suite across all four metrics, reaching 93.4% base accuracy and 92.0% / 91.9% average accuracy (**+3.6 / +3.5** over FLEG; **+4.7 / +4.6** over FM), with relative drops of 2.8% / 2.9% versus FLEG's 5.1% and FM's 5.7%. Aggressive base-class tuning is brittle under domain shift; dual-memory updates on frozen features preserve discriminability without over-adapting to the source domain. EF-EB parity again indicates that closed-form correlation tracking and prototype matching generalise without stored exemplars.

Table 2: Comparison on cross-domain protocols in terms of base-class accuracy $\mathcal{A}_b$, final-session accuracy $\mathcal{A}_T$, average accuracy $\bar{\mathcal{A}}$, and relative drop $\Delta \downarrow$. Best results are in **bold**.

Method	Venue	ShapeNet→CO3D				ModelNet→ScanObjNN				ShapeNet→ScanObjNN			
		$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$
Oracle	–	81.0	67.3	74.0	16.9	88.4	71.2	78.3	19.5	81.4	78.7	80.6	3.3
FT	–	81.0	1.6	10.2	98.0	88.4	1.9	25.7	97.9	81.4	0.9	31.3	98.9
LwF	TPAMI'17	81.0	1.9	15.3	97.7	88.4	2.5	33.1	97.2	81.4	5.9	37.3	92.8
IL2M	ICCV'19	81.0	30.0	38.4	63.0	88.4	52.0	62.9	41.2	81.4	45.8	56.1	43.7
ScaIL	WACV'20	81.0	35.2	42.6	56.5	88.4	52.9	63.4	40.2	81.4	40.0	54.3	50.9
EEIL	ECCV'18	81.0	47.6	59.9	41.2	88.4	56.8	69.1	35.7	81.4	63.4	72.3	22.1
FACT	CVPR'22	81.4	55.9	65.4	31.3	89.1	63.5	73.4	28.7	82.3	66.8	73.4	18.8
Semantic	CVPR'21	80.6	55.2	63.2	31.5	88.5	64.2	73.6	27.5	81.3	62.9	70.0	22.6
Microshape	ECCV'22	82.6	57.1	67.8	30.9	89.3	65.1	74.0	27.1	82.5	67.1	73.9	18.7
C3PR	ECCV'24	83.6	63.2	72.0	24.4	88.3	67.8	75.6	23.2	84.5	71.9	77.4	14.9
FM	ACCV'24	87.3	72.6	79.9	16.8	87.7	79.2	83.3	9.7	90.8	85.6	87.3	5.7
FLEG	ICML'25	91.7	76.8	84.0	16.2	**93.8**	86.8	90.0	7.5	90.9	86.3	88.4	5.1
FROST*	Ours (EB)	**92.3**	**82.9**	**87.8**	**10.2**	92.8	87.9	90.2	5.3	**93.4**	**90.8**	**92.0**	**2.8**
FROST	Ours (EF)	**92.3**	82.5	87.6	10.6	92.8	**87.9**	**90.2**	**5.3**	**93.4**	90.7	91.9	2.9

Table 3: Comparison on image FSCIL benchmarks in terms of base-class accuracy $\mathcal{A}_b$, final-session accuracy $\mathcal{A}_T$, average accuracy $\bar{\mathcal{A}}$, and relative drop Δ. Best results are in **bold**.

(a) CUB-200

Method	Venue	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$
iCaRL	CVPR'17	82.43	55.92	63.83	32.16
FOSTER	ECCV'22	85.02	60.42	69.78	28.93
RDI	IJCAI'24	80.13	60.20	67.65	24.87
FACT	CVPR'22	84.32	70.91	75.18	15.90
Comp-FSCIL	ICML'24	83.67	72.80	76.72	12.99
NC-FSCIL	ICLR'23	83.52	73.80	77.32	11.64
L2P	CVPR'22	82.47	73.67	76.21	10.67
TEEN	NIPS'23	84.03	75.61	78.11	10.02
Yourself	ECCV'24	82.31	74.38	76.86	9.63
PriViLege	CVPR'24	82.21	75.08	77.50	8.67
DualPrompt	ECCV'22	83.51	76.53	78.53	8.36
CODA-P	CVPR'23	79.61	72.97	74.63	8.34
LGSP-P	ICCV'25	85.72	79.72	81.39	7.00
FROST	Ours	**90.10**	**84.96**	**85.80**	**5.71**

(b) iNF200

Method	Venue	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$
Yourself	ECCV'24	61.9	16.1	36.97	73.99
Approx.	ECCV'24	66.3	36.75	48.90	44.57
PriViLege	CVPR'24	63.1	48.1	55.53	**23.77**
LGSP-P	ICCV'25	67.0	48.6	57.62	27.46
FROST	Ours	**80.26**	**58.77**	**67.85**	26.78

(c) FGVC-Aircraft

Method	Venue	$\mathcal{A}_b$	$\mathcal{A}_T$	$\bar{\mathcal{A}}$	$\Delta \downarrow$
Yourself	ECCV'24	26.29	8.08	13.66	69.27
Approx.	ECCV'24	21.67	12.97	16.62	40.15
PriViLege	CVPR'24	22.75	20.04	21.16	**11.91**
LGSP-P	ICCV'25	27.67	23.87	25.82	13.73
FROST	Ours	**49.86**	**38.74**	**42.34**	22.29

Image benchmarks. Tab. 3 reports 2D results under the non-replay protocol of Jiang et al. (2025), against image FSCIL methods that rely on prompts, rehearsal, or session-wise adaptation. On CUB-200, exemplar-free FROST reaches 84.96% final accuracy with a 5.71% relative drop, improving over the previous best LGSP-P (79.72%, 7.00% drop) by **+5.24** points in $\mathcal{A}_T$ while raising base accuracy from 85.72% to 90.10%. On iNF200, it attains 58.77% final and 67.85% average accuracy, surpassing LGSP-P (48.6%) and PriViLege (48.1%) by **+10.17** and **+10.67** points in $\mathcal{A}_T$. On FGVC-Aircraft, the hardest of the three, FROST reaches 38.74% final accuracy, improving over LGSP-P (23.87%) by **+14.87** points, and lifts base accuracy from 27.67% to 49.86%. PriViLege reports a smaller relative drop Δ on Aircraft and iNF200, but at lower base accuracy. Under absolute final and average accuracy, which better reflect usable performance, FROST remains ahead; on CUB-200, where base accuracy is comparable across methods, it achieves the best Δ. As on CO3D, the largest gains appear where scarce novel supervision and fine-grained categories expose base bias, evidence that dual-memory fusion over frozen features supports exemplar-free FSCIL across modalities.

Harmonic-mean analysis. Fig. 2 tracks harmonic mean $\mathcal{H}$ over the three incremental tasks on ModelNet→ScanObjNN (left) and ShapeNet→ScanObjNN (right), using the same comparator suite as Tab. 2. On both transfers, FROST and FROST* remain near the top of the ranking and stay

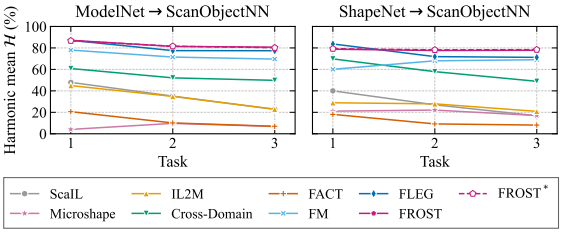

Figure 2: Harmonic mean $\mathcal{H}$ across incremental tasks under the cross-domain setting.

Table 4: Ablation on ModelNet (exemplar-free). Highlighted row is the default dual-memory configuration.

Setting	$\mathcal{A}_b$	$\mathcal{A}_n$	$\mathcal{A}_T$	$\Delta \downarrow$
Linear	93.62	21.21	4.39	95.31
Adapter + Linear	93.49	24.24	5.01	94.64
CorrMem	95.71	71.69	90.36	5.59
ProtoMem	94.00	**80.00**	88.79	5.55
Corr + ProtoMem	**95.82**	76.43	**91.27**	**4.75**

nearly flat as novel sessions arrive; the solid and dashed curves almost coincide, reinforcing the EF-EB parity observed in the tables. FLEG opens competitively, then loses base-novel balance as the label space expands, so its curve slopes downward on both panels. FM is transfer-dependent: comparatively flat under ModelNet→ScanObjNN, non-monotonic under ShapeNet→ScanObjNN, and below FROST on both panels. Cross-Domain sits in the mid-tier and declines with each task; Microshape stays low; classical 2D adaptations (ScaIL, IL2M, FACT) start weak and deteriorate further. Taken together, the trajectories show that CorrMem-ProtoMem fusion not only reaches the strongest $\mathcal{H}$, but also preserves that balance under synthetic-to-real shift more consistently than foundation-model and classical FSCIL baselines.

4.3 Ablation Studies

Tab. 4 isolates each component of the dual-memory design on ModelNet under the exemplar-free protocol. A strong frozen representation is not enough on its own: a linear head on frozen features, with or without the base-only adapter, preserves high base-session accuracy yet almost loses the novel stream, so last-session accuracy collapses below 5.1% (relative drop above 94%). *CorrMem* alone recovers closed-form ridge discrimination (90.36% final, 95.71% base) but under-serves few-shot classes at $\mathcal{A}_n$=71.69%, consistent with the predicted base-class mass bias. *ProtoMem* alone corrects that bias with one normalised mean per class, lifting novel accuracy to 80.00%, yet softens global discrimination to 88.79% final. Fusing the two couples ridge discrimination with prototype balance: 91.27% final accuracy, the lowest relative drop (4.75%), and the highest base score (95.82%), while novel accuracy (76.43%) sits between the single-memory extremes. Thus CorrMem supplies additive discrimination without replay, ProtoMem restores class-mass balance, and only their fusion recovers base-novel parity without exemplars or incremental updates.

Comprehensive implementation details, training schedules, full-sequence experiments, extended hyperparameter ablations, and further results are provided in the Appendix.

5 Conclusion

We presented FROST, an exemplar-free FSCIL framework that freezes a base-adapted encoder and registers new classes through one dual-memory recipe shared across images and point clouds. Once the feature map is stationary, raw exemplar replay, prompt tuning, auxiliary heads, and session-wise gradient updates are unnecessary: compact class statistics suffice. *CorrMem* accumulates feature correlations into a closed-form ridge head, preserving past-class discrimination by additive updates, yet inherits a base-class mass bias under few-shot novel supervision. *ProtoMem* corrects that bias with one normalised mean per class, placing data-rich base and scarce novel categories on equal footing. Fusing their logits couples ridge discrimination with prototype balance and restores base-novel parity; neither memory alone achieves this operating point. Across in-domain and cross-domain 3D protocols and non-replay image FSCIL benchmarks, exemplar-free FROST matches an exemplar counterpart and outperforms methods that rely on rehearsal, prompts, or incremental optimisation. The near match between exemplar-free and exemplar-based variants indicates that CorrMem and ProtoMem already recover the retention usually attributed to raw buffers. These results show that complementary class statistics on frozen pretrained features are sufficient for strong exemplar-free FSCIL across modalities.

AI USE STATEMENT

In this work, we used generative AI tools to help implement experimental and analysis code, prepare hyperparameter sweeps for the configuration tables, convert CSV results into LaTeX tables, and polish the wording of short passages of text. We did not use generative AI tools to generate synthetic datasets, develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proofs or assist in writing them, propose or refine scientific hypotheses as primary research contributions, design the core research methodology or experimental protocol, perform translation, clean or reformat datasets, or support qualitative or thematic data analysis. The disclosure items concerning proof writing and synthetic data generation are not applicable to this empirical study. In addition, we used generative AI tools for the following recommended-disclosure tasks: drafting and rewriting manuscript prose, editing for clarity and readability, suggesting paper structure and wording, brainstorming presentation choices, creating or editing plotting and figure-generation scripts, and formatting bibliographic entries. All AI-assisted text, code, and figures were reviewed, edited, and verified by the authors. Manuscript claims and citations were checked against primary sources and experimental logs; implementation and analysis code used for the reported results was tested by the authors; and all numerical results, including ablation metrics, were obtained from author-run experiments rather than from model-generated values. The authors take full responsibility for the final content of this work, including all text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work does not raise concerns under the ICLR Code of Ethics. Our study does not involve human subjects, user studies, or the collection of new data, and therefore did not require IRB approval. All experiments are conducted on publicly available benchmark datasets used in accordance with their original licenses and terms of use; we do not release any new datasets. The proposed method is a general-purpose approach to few-shot class-incremental learning on point clouds and images, and we do not foresee direct harmful applications, privacy or security risks, or discrimination and fairness concerns beyond those already inherent to standard 3D object recognition and image classification benchmarks. The authors declare no conflicts of interest or sponsorship that could have influenced this work.

REPRODUCIBILITY STATEMENT

We have made every effort to provide the full recipe needed to reproduce our results. The proposed method is described in Section 3, and the complete experimental setup, including benchmark datasets, class splits, incremental session protocol, backbone architecture, training schedule, and evaluation metrics, is detailed in Section 4. All hyperparameters and per-dataset configurations, together with data preprocessing and augmentation steps, are listed in Appendix A.2. To further support reproducibility, we provide an anonymous code archive as supplementary material, containing source code, configuration files, training scripts, scripts for generating the reported tables and figures, and checkpoint hashes; the code will be publicly released upon acceptance.

REFERENCES

Sahar Ahmadi, Ali Cheraghian, Morteza Saberi, Md Towsif Abir, Hamidreza Dastmalchi, Farookh Hussain, and Shafin Rahman. Foundation model-powered 3d few-shot class incremental learning via training-free adaptor. In *ACCV*, pp. 2282–2299, 2024.

Afra Feyza Akyürek, Ekin Akyürek, Derry Tanti Wijaya, and Jacob Andreas. Subspace regularizers for few-shot class incremental learning. *arXiv preprint arXiv:2110.07059*, 2022.

Muhammad Awais, Muzammal Naseer, Salman Khan, Rao Muhammad Anwer, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, and Fahad Shahbaz Khan. Foundational models defining a new era in vision: A survey and outlook. *arXiv:2307.13721*, 2023.

Eden Belouadah and Adrian Popescu. Il2m: Class incremental learning with dual memory. In *ICCV*, pp. 583–592, 2019.

Eden Belouadah and Adrian Popescu. Scail: Classifier weights scaling for class incremental learning. In *WACV*, 2020.

Francisco M Castro, Manuel J Marín-Jiménez, Nicolás Guil, Cordelia Schmid, and Karteek Alahari. End-to-end incremental learning. In *ECCV*, 2018.

Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, et al. Shapenet: An information-rich 3d model repository. *arXiv:1512.03012*, 2015.

Ali Cheraghian, Shafin Rahman, Pengfei Fang, Soumava Kumar Roy, Lars Petersson, and Mehrtash Harandi. Semantic-aware knowledge distillation for few-shot class-incremental learning. In *CVPR*, 2021a.

Ali Cheraghian, Shafin Rahman, Sameera Ramasinghe, Pengfei Fang, Christian Simon, Lars Petersson, and Mehrtash Harandi. Synthesized feature based few-shot class-incremental learning on a mixture of subspaces. In *ICCV*, 2021b.

Ali Cheraghian, Zeeshan Hayder, Sameera Ramasinghe, Shafin Rahman, Javad Jafaryahya, Lars Petersson, and Mehrtash Harandi. Canonical shape projection is all you need for 3d few-shot class incremental learning. In *ECCV*, 2024.

Townim Chowdhury, Ali Cheraghian, Sameera Ramasinghe, Sahar Ahmadi, Morteza Saberi, and Shafin Rahman. Few-shot class-incremental learning for 3d point cloud objects. In *ECCV*, pp. 204–220. Springer, 2022.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, 2009.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. *ICLR*, 2021.

Arthur E. Hoerl and Robert W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

Zitong Huang, Ze Chen, Zhixing Chen, Erjin Zhou, Xinxing Xu, Rick Siow Mong Goh, Yong Liu, Wangmeng Zuo, and Chunmei Feng. Learning prompt with distribution-based feature replay for few-shot class-incremental learning. *arXiv preprint arXiv:2401.01598*, 2024.

Yongwei Jiang, Yixiong Zou, Yuhua Li, and Ruixuan Li. Revisiting pool-based prompt learning for few-shot class-incremental learning. In *ICCV*, pp. 1303–1313, 2025.

Haeyong Kang, Jaehong Yoon, Sultan Rizky Hikmawan Madjid, Sung Ju Hwang, and Chang D. Yoo. On the soft-subnetwork for few-shot class incremental learning. *arXiv preprint arXiv:2209.07529*, 2023.

Do-Yeon Kim, Dong-Jun Han, Jun Seo, and Jaekyun Moon. Warping the space: Weight space rotation for class-incremental few-shot learning. In *ICLR*, 2023.

Xiangqi Li, Libo Huang, Zhulin An, Weilun Feng, Chuanguang Yang, Boyu Diao, Fei Wang, and Yongjun Xu. Geometric feature embedding for effective 3D few-shot class incremental learning. In *ICML*, volume 267, pp. 34674–34687. PMLR, 2025.

Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on x-transformed points. In *NIPS*, 2018.

Zhizhong Li and Derek Hoiem. Learning without forgetting. *PAMI*, 40(12):2935–2947, 2017.

Subhransu Maji, Esa Rahtu, Juho Kannala, Matthew Blaschko, and Andrea Vedaldi. Fine-grained visual classification of aircraft. *arXiv preprint arXiv:1306.5151*, 2013.

Marc Masana, Xialei Liu, Bartłomiej Twardowski, Mikel Menta, Andrew D Bagdanov, and Joost Van De Weijer. Class-incremental learning: survey and performance evaluation on image classification. *TPAMI*, 45(5):5513–5533, 2022.

Jonathan Masci, Davide Boscaini, Michael Bronstein, and Pierre Vandergheynst. Geodesic convolutional neural networks on riemannian manifolds. In *ICCVW*, pp. 37–45, 2015.

Daniel Maturana and Sebastian Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. In *IROS*, 2015.

Mark D McDonnell, Dong Gong, Amin Parvaneh, Ehsan Abbasnejad, and Anton Van den Hengel. Ranpac: Random projections and pre-trained models for continual learning. *NIPS*, 36:12022–12053, 2023.

Keon-Hee Park, Kyungwoo Song, and Gyeong-Moon Park. Pre-trained vision and language transformers are few-shot incremental learners. *arXiv preprint arXiv:2404.02117*, 2024.

Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *CVPR*, 2017a.

Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *NIPS*, 2017b.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *ICML*, 2021.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *NIPS*, volume 20, 2007.

Sylvestre Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl, and Christoph H. Lampert. icarl: Incremental classifier and representation learning. In *CVPR*, 2017.

Jeremy Reizenstein, Roman Shapovalov, Philipp Henzler, Luca Sbordone, Patrick Labatut, and David Novotny. Common objects in 3d: Large-scale learning and evaluation of real-life 3d category reconstruction. In *ICCV*, 2021.

James Seale Smith, Leonid Karlinsky, Vyshnavi Gutta, Paola Cascante-Bonilla, Donghyun Kim, Assaf Arbelle, Rameswar Panda, Rogerio Feris, and Zsolt Kira. Coda-prompt: Continual decomposed attention-based prompting for rehearsal-free continual learning. *arXiv preprint arXiv:2211.13218*, 2023.

Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *ICCV*, pp. 945–953, 2015.

Yuwen Tan and Xiang Xiang. Cross-domain few-shot incremental learning for point-cloud recognition. In *WACV*, pp. 2307–2316, January 2024.

Yu-Ming Tang, Yi-Xing Peng, Jingke Meng, and Wei-Shi Zheng. Rethinking few-shot class-incremental learning: Learning from yourself. In *ECCV*, 2024.

Xiaoyu Tao, Xiaopeng Hong, Xinyuan Chang, Songlin Dong, Xing Wei, and Yihong Gong. Few-shot class-incremental learning. In *CVPR*, 2020.

Vishal Thengane, Xiatian Zhu, Salim Bouzerdoum, Son Lam Phung, and Yunpeng Li. Foundational models for 3d point clouds: A survey and outlook. *arXiv:2501.18594*, 2025.

Mikaela Angelina Uy, Quang-Hieu Pham, Binh-Son Hua, Thanh Nguyen, and Sai-Kit Yeung. Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data. In *ICCV*, pp. 1588–1597, 2019.

Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. The inaturalist species classification and detection dataset. In *CVPR*, pp. 8769–8778. IEEE, 2018.

Catherine Wah, Steve Branson, Peter Welinder, Pietro Perona, and Serge Belongie. The caltech-ucsd birds-200-2011 dataset. Technical report, California Institute of Technology, 2011.

Fu-Yun Wang, Da-Wei Zhou, Han-Jia Ye, and De-Chuan Zhan. Foster: Feature boosting and compression for class-incremental learning. *arXiv preprint arXiv:2204.04662*, 2022a.

Qi-Wei Wang, Da-Wei Zhou, Yi-Kai Zhang, De-Chuan Zhan, and Han-Jia Ye. Few-shot class-incremental learning via training-free prototype calibration. *arXiv preprint arXiv:2312.05229*, 2023.

Xuan Wang, Zhong Ji, Xiyao Liu, Yanwei Pang, and Jungong Han. On the approximation risk of few-shot class-incremental learning. In *ECCV*, 2024.

Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph cnn for learning on point clouds. *TOG*, 2019.

Zifeng Wang, Zizhao Zhang, Sayna Ebrahimi, Ruoxi Sun, Han Zhang, Chen-Yu Lee, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, and Tomas Pfister. Dualprompt: Complementary prompting for rehearsal-free continual learning. *arXiv preprint arXiv:2204.04799*, 2022b.

Zifeng Wang, Zizhao Zhang, Chen-Yu Lee, Han Zhang, Ruoxi Sun, Xiaoqi Ren, Guolong Su, Vincent Perot, Jennifer Dy, and Tomas Pfister. Learning to prompt for continual learning. *arXiv preprint arXiv:2112.08654*, 2022c.

Wenxuan Wu, Zhongang Qi, and Li Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *CVPR*, 2019.

Xiaoyang Wu, Li Jiang, Peng-Shuai Wang, Zhijian Liu, Xihui Liu, Yu Qiao, Wanli Ouyang, Tong He, and Hengshuang Zhao. Point transformer v3: Simpler faster stronger. In *CVPR*, pp. 4840–4851, 2024.

Zhirong Wu, Shuran Song, Aditya Khosla, Fisher Yu, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 3d shapenets: A deep representation for volumetric shapes. In *CVPR*, pp. 1912–1920, 2015.

Aoran Xiao, Xiaoqin Zhang, Ling Shao, and Shijian Lu. A survey of label-efficient deep learning for 3d point clouds. *TPAMI*, 46(12):9139–9160, 2024.

Yibo Yang, Haobo Yuan, Xiangtai Li, Zhouchen Lin, Philip Torr, and Dacheng Tao. Neural collapse inspired feature-classifier alignment for few-shot class-incremental learning. In *ICLR*, 2023.

In-Ug Yoon, Tae-Min Choi, Sun-Kyung Lee, Young-Min Kim, and Jong-Hwan Kim. Image-object-specific prompt learning for few-shot class-incremental learning. *arXiv preprint arXiv:2309.02833*, 2023.

Xiangyu Yue, Bichen Wu, Sanjit A Seshia, Kurt Keutzer, and Alberto L Sangiovanni-Vincentelli. A lidar point cloud generator: From a virtual world to autonomous driving. In *ICMR*, 2018.

Chi Zhang, Nan Song, Guosheng Lin, Yun Zheng, Pan Pan, and Yinghui Xu. Few-shot incremental learning with continually evolved classifiers. In *CVPR*, 2021.

Tongxu Zhang, Zhiming Liang, and Bei Wang. A survey of medical point cloud shape learning: Registration, reconstruction and variation. In *BIBM*, pp. 5777–5782. IEEE, 2025.

Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip HS Torr, and Vladlen Koltun. Point transformer. In *CVPR*, pp. 16259–16268, 2021.

Da-Wei Zhou, Fu-Yun Wang, Han-Jia Ye, Liang Ma, Shiliang Pu, and De-Chuan Zhan. Forward compatible few-shot class-incremental learning. In *CVPR*, 2022.

Haichen Zhou, Yixiong Zou, Ruixuan Li, Yuhua Li, and Kui Xiao. Delve into base-novel confusion: Redundancy exploration for few-shot class-incremental learning. *arXiv preprint arXiv:2405.04918*, 2024a.

Junsheng Zhou, Jinsheng Wang, Baorui Ma, Yu-Shen Liu, Tiejun Huang, and Xinlong Wang. Uni3d: Exploring unified 3d representation at scale. In *ICLR*, 2024b.

Kai Zhu, Yang Cao, Wei Zhai, Jie Cheng, and Zheng-Jun Zha. Self-promoted prototype refinement for few-shot class-incremental learning. *arXiv preprint arXiv:2107.08918*, 2021.

Yixiong Zou, Shanghang Zhang, Yuhua Li, and Ruixuan Li. Margin-based few-shot class-incremental learning with class-level overfitting mitigation. *NIPS*, 35:27267–27279, 2022.

Yixiong Zou, Shanghang Zhang, Haichen Zhou, Yuhua Li, and Ruixuan Li. Compositional few-shot class-incremental learning. *arXiv preprint arXiv:2405.17022*, 2024.

Table 5: Dataset splits used in Sec. 4. Point-cloud base/novel class counts follow Chowdhury et al. (2022). Image schedules follow the non-replay protocol of Jiang et al. (2025), where Nw-Ks denotes N-way K-shot and novel classes arrive over 10 incremental sessions.

Modality	Protocol	Dataset	Base	Novel
Point cloud	In-domain	ModelNet40	20 classes	20 classes
Point cloud	In-domain	ShapeNet	20 classes	30 classes
Point cloud	In-domain	CO3D	25 classes	25 classes
Point cloud	Cross-domain	ModelNet→ScanObjNN	26 classes	11 classes
Point cloud	Cross-domain	ShapeNet→ScanObjNN	44 classes	15 classes
Point cloud	Cross-domain	ShapeNet→CO3D	15 classes	50 classes
Image	Fine-grained	CUB-200 (birds; 11,788 imgs)	100w-30s	10×(10w-5s)
Image	Fine-grained	iNF200 (fungi; 10,000 imgs)	100w-50s	10×(10w-5s)
Image	Fine-grained	FGVC-Aircraft (10,000 imgs)	50w-30s	10×(5w-5s)

Table 6: Implementation details for point-cloud (3D) and image (2D) experiments.

Setting	Point clouds (3D)	Images (2D)
Encoder	Uni3D-Base (Zhou et al., 2024b)	ViT-B/16 (ImageNet-21K)
Protocol	Splits of Chowdhury et al. (2022)	Non-replay of Jiang et al. (2025)
Input	FPS, M=1024 points	224×224
Shot budget	K=5	K=5
Adapter ϕ	Base session only, then frozen	Base session only, then frozen
Bottleneck width d'	64	64
Adapter epochs	10	100 (CUB/iNF), 120 (Aircraft)
Batch size	8	48
Optimiser	SGD	SGD
Learning rate	1×10^{-3}	1×10^{-2}
Weight decay	1×10^{-4}	5×10^{-4}
CorrMem lift V	2000	2000
Ridge penalty λ	Once on 20% base hold-out	Once on 20% base hold-out
Fusion weight α	0.5	0.5

A Additional Experimental Details

In addition to Sec. 4.1, this appendix records the full dataset splits and the default implementation recipe used for all reported runs.

A.1 Dataset and Split

Tab. 5 lists every protocol used in Sec. 4. Point-cloud in-domain runs draw both sessions from one dataset; cross-domain runs use a synthetic base session and real-scan novel classes (Chowdhury et al., 2022). Image experiments follow Jiang et al. (2025): CUB-200 uses a 100-way 30-shot base session then ten 10-way 5-shot novel sessions; iNF200 uses a 100-way 50-shot base then the same novel schedule; FGVC-Aircraft uses a 50-way 30-shot base then ten 5-way 5-shot novel sessions. These three fine-grained sets avoid ImageNet class overlap that affects CIFAR and miniImageNet under ImageNet-21K pretraining. Unless noted, reported numbers use the exemplar-free (EF) setting; the exemplar-based (EB) variant stores one sample from each previous class.

A.2 Implementation Details

Tab. 6 summarises the default 3D and 2D recipes side by side. Both modalities share the same dual-memory design: a base-only adapter of width d'=64, then a frozen f'_θ; CorrMem with lift V=2000 and a single λ chosen on a 20% base hold-out; ProtoMem class means; and fixed fusion α=0.5. Novel sessions never store exemplars in the default EF setting, never tune prompts, and never run incremental SGD; they only accumulate CorrMem statistics and append ProtoMem centroids. What differs is modality plumbing: Uni3D-Base with FPS M=1024 under the splits of Chowdhury

Table 7: Session-wise accuracy on in-domain protocols (ModelNet, ShapeNet, CO3D). $\mathcal{A}_b$ is base-session accuracy; $\mathcal{A}_t$ is accuracy after novel session t; $\Delta\downarrow$ is the relative drop. Best results are in **bold**.

Method	ModelNet						ShapeNet								CO3D						
	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\Delta\downarrow$	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\mathcal{A}_5$	$\mathcal{A}_6$	$\Delta\downarrow$	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\mathcal{A}_5$	$\Delta\downarrow$
Oracle	89.8	88.2	87.0	83.5	80.5	10.4	87.0	85.2	84.3	83.0	82.5	82.2	81.3	6.6	76.7	69.4	64.8	62.7	60.7	59.8	22.0
FT	89.8	9.7	4.3	3.3	3.0	96.7	87.0	25.7	6.8	1.3	0.9	0.6	0.4	99.5	76.7	11.2	3.6	3.2	1.8	0.8	99.0
LwF	89.8	36.0	9.1	3.6	3.1	96.5	87.0	60.8	33.5	15.9	3.8	3.1	1.8	97.9	76.7	14.7	4.7	3.5	2.3	1.0	98.7
IL2M	89.8	65.5	58.4	52.3	53.6	40.3	87.0	58.6	45.7	40.7	50.1	49.4	49.3	43.3	76.7	31.5	27.7	18.1	27.1	21.9	71.4
ScaIL	89.8	66.8	64.5	58.7	56.5	37.1	87.0	56.6	51.8	44.3	50.3	46.3	45.4	47.8	76.7	39.5	34.1	24.1	30.1	27.5	64.1
EEIL	89.8	75.4	67.2	60.1	55.6	38.1	87.0	77.7	73.2	69.3	66.4	65.9	65.8	24.4	76.7	61.4	52.4	42.8	39.5	32.8	57.2
FACT	90.4	81.3	77.1	73.5	65.0	28.1	87.5	75.3	71.4	69.9	67.5	65.7	62.5	28.6	77.9	67.1	59.7	54.8	50.2	46.7	40.0
Semantic	91.3	82.2	74.3	70.0	64.7	29.1	87.2	74.9	68.1	69.0	68.1	66.9	63.8	26.8	78.6	66.9	59.2	53.6	49.1	42.9	45.4
Microshape	93.6	83.1	78.2	75.8	67.1	28.3	87.6	83.2	81.5	79.0	76.8	73.5	72.6	17.1	78.5	67.3	60.1	56.1	51.4	47.2	39.9
C3PR	91.6	82.3	75.8	72.2	70.9	22.5	88.0	81.6	77.8	76.7	76.9	76.2	74.7	15.1	81.5	69.4	66.5	63.0	54.2	53.8	34.0
FLEG	**98.3**	**95.5**	**93.5**	91.4	87.3	11.2	**94.3**	**91.4**	89.7	89.2	88.4	86.2	83.0	12.0	**82.1**	75.4	69.7	66.7	57.7	57.3	30.2
FROST*	95.8	93.5	**93.5**	**93.0**	**91.5**	**4.5**	92.2	91.0	**90.2**	**89.9**	**89.6**	**88.9**	**89.3**	**3.1**	81.6	77.2	**74.8**	73.7	**72.5**	71.9	11.9
FROST	95.8	93.6	93.3	92.8	91.3	4.7	92.2	91.0	90.1	89.8	89.4	88.8	89.2	3.3	81.6	**77.3**	**74.8**	**73.8**	72.3	**72.0**	**11.8**

et al. (2022) versus ViT-B/16 with 224×224 inputs under the non-replay protocol of Jiang et al. (2025), together with a shorter, smaller-batch base SGD schedule for point clouds than for images. Relative to prior FSCIL systems, the recipe therefore keeps one frozen dual-memory pipeline across modalities instead of session-wise prompts, auxiliary heads, or raw exemplar replay.

B Additional Experimental Results

B.1 Full Session-Wise Results for 3D

Tabs. 7 and 8 expand the condensed in-domain and cross-domain summaries of Sec. 4 into accuracy after every novel session. The session curves make the failure mode of classical continual learners explicit: FT and LwF collapse within one or two novel sessions on every protocol, while replay-style adaptations (IL2M, ScaIL, EEIL) slow the drop but still lose tens of points by the final session. Prior 3D FSCIL methods occupy a middle band; among them FLEG is the strongest early-session competitor, yet its trajectory declines more steeply than FROST once novel classes accumulate.

Tab. 7 reports the in-domain trajectories. On ModelNet, FLEG starts highest ($\mathcal{A}_b$=98.3%) but falls from 95.5% after the first novel session to 87.3% at the end (Δ=11.2%), whereas EB / EF FROST start lower at the base session (95.8%) and finish at 91.5% / 91.3% with Δ=4.5% / 4.7%, overtaking FLEG from $\mathcal{A}_3$ onward and also beating the joint-training Oracle (80.5%, Δ=10.4%). ShapeNet shows the same pattern over six novel sessions: FLEG declines from 94.3% to 83.0%, while FROST stays at or above 88.8% throughout and ends at 89.3% / 89.2%. CO3D remains the hardest in-domain protocol; every baseline records Δ>20%, yet FROST retains 71.9% / 72.0% at the end against FLEG's 57.3% and Oracle's 59.8%. Across all three datasets the EB and EF curves stay within a few tenths of a point after every session, so the closed-form CorrMem and ProtoMem updates already recover the retention that exemplar buffers are intended to provide.

Tab. 8 reports the cross-domain trajectories under domain shift from synthetic base classes to real-scan novel classes. ShapeNet→CO3D is the longest schedule (ten novel sessions): FLEG drops from 91.7% to 76.8% (Δ=16.2%), while EB / EF FROST descend more gradually from 92.3% to 82.9% / 82.5% (Δ=10.2% / 10.6%) and remain ahead of every published competitor after each session. On ModelNet→ScanObjNN, FLEG again records the strongest base accuracy (93.8%), but FROST matches or exceeds it from $\mathcal{A}_2$ onward and finishes at 87.9% with Δ=5.3%, versus 86.8% and 7.5% for FLEG. On ShapeNet→ScanObjNN the gap opens immediately: FROST leads after every session, ends at 90.8% / 90.7%, and reduces Δ to 2.8% / 2.9% against FLEG's 5.1% and FM's 5.7%. As in the in-domain setting, EB and EF remain nearly coincident along each curve, confirming that exemplar-free dual-memory updates retain their advantage when the novel distribution departs from the base domain.

Table 8: Session-wise accuracy on cross-domain protocols (ShapeNet→CO3D, ModelNet→ScanObjNN, ShapeNet→ScanObjNN). $\mathcal{A}_b$ is base-session accuracy; $\mathcal{A}_t$ is accuracy after novel session t; $\Delta\downarrow$ is the relative drop. Best results are in **bold**.

Method	ShapeNet→CO3D												ModelNet→ScanObjNN					ShapeNet→ScanObjNN				
	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$	$\mathcal{A}_5$	$\mathcal{A}_6$	$\mathcal{A}_7$	$\mathcal{A}_8$	$\mathcal{A}_9$	$\mathcal{A}_{10}$	$\Delta\downarrow$	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\Delta\downarrow$	$\mathcal{A}_b$	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\Delta\downarrow$
Oracle	81.0	79.5	78.3	75.2	75.1	74.8	72.3	71.3	70.0	68.8	67.3	16.9	88.4	79.7	74.0	71.2	19.5	81.4	82.5	79.8	78.7	3.3
FT	81.0	20.2	2.3	1.7	0.8	1.0	1.0	1.3	0.9	0.5	1.6	98.0	88.4	6.4	6.0	1.9	97.9	81.4	38.7	4.0	0.9	98.9
LwF	81.0	57.4	19.3	2.3	1.0	0.9	0.8	1.3	1.1	0.8	1.9	97.7	88.4	35.8	5.8	2.5	97.2	81.4	47.9	14.0	5.9	92.8
IL2M	81.0	45.6	36.8	35.1	31.8	33.3	34.0	31.5	30.6	32.3	30.0	63.0	88.4	58.2	52.9	52.0	41.2	81.4	53.2	43.9	45.8	43.7
ScaIL	81.0	50.1	45.7	39.1	39.0	37.9	38.0	36.0	33.7	33.0	35.2	56.5	88.4	56.5	55.9	52.9	40.2	81.4	49.0	46.7	40.0	50.9
EEIL	81.0	75.2	69.3	63.2	60.5	57.9	53.0	51.9	51.3	47.8	47.6	41.2	88.4	70.2	61.0	56.8	35.7	81.4	74.5	69.8	63.4	22.1
FACT	81.4	76.0	70.3	68.1	65.8	63.5	63.0	60.1	58.2	57.5	55.9	31.3	89.1	72.5	68.3	63.5	28.7	82.3	74.6	69.9	66.8	18.8
Semantic	80.6	69.5	66.5	62.9	63.2	63.0	61.2	58.3	58.1	57.2	55.2	31.5	88.5	73.9	67.7	64.2	27.5	81.3	70.6	65.2	62.9	22.6
Microshape	82.6	77.9	73.9	72.7	67.7	66.2	65.4	63.4	60.6	58.1	57.1	30.9	89.3	73.2	68.4	65.1	27.1	82.5	74.8	71.2	67.1	18.7
C3PR	83.6	80.0	77.8	75.4	72.8	72.3	70.3	67.9	64.9	64.1	63.2	24.4	88.3	75.7	70.6	67.8	23.2	84.5	77.8	75.5	71.9	14.9
FM	87.3	86.2	84.4	82.2	80.7	79.6	78.2	76.8	76.1	74.5	72.6	16.8	87.7	84.7	81.5	79.2	9.7	90.8	86.5	86.4	85.6	5.7
FLEG	91.7	90.5	89.0	86.8	84.9	83.3	82.3	81.1	79.8	77.8	76.8	16.2	**93.8**	**91.9**	87.5	86.8	7.5	90.9	89.1	87.1	86.3	5.1
FROST*	**92.3**	**91.7**	**90.9**	**89.7**	**88.7**	**87.8**	**86.8**	**85.7**	**84.9**	**84.2**	**82.9**	**10.2**	92.8	91.2	**88.9**	**87.9**	**5.3**	**93.4**	**92.3**	**91.4**	**90.8**	**2.8**
FROST	**92.3**	91.6	90.8	89.5	88.6	87.5	86.6	85.5	84.6	83.9	82.5	10.6	92.8	91.2	88.8	**87.9**	**5.3**	**93.4**	**92.3**	91.2	90.7	2.9

Table 9: Hyperparameter sensitivity on ModelNet (exemplar-free). Highlighted rows are the defaults. (a) Fusion weight α. (b) Adapter width d'. (c) CorrMem lift V.

(a) Fusion α

α	$\mathcal{A}_b$	$\mathcal{A}_n$	$\mathcal{A}_T$	$\Delta\downarrow$
0.1	95.10	**79.14**	90.22	5.13
0.3	95.57	77.73	91.10	**4.67**
0.5	**95.82**	76.43	**91.27**	4.75
0.7	95.81	74.63	91.03	4.99
0.9	95.72	72.90	90.68	5.27

(b) Adapter d'

d'	$\mathcal{A}_b$	$\mathcal{A}_n$	$\mathcal{A}_T$	$\Delta\downarrow$
32	**95.82**	76.00	91.22	4.80
64	**95.82**	76.43	91.27	4.75
128	95.75	76.75	91.30	4.65
256	95.73	**77.41**	**91.40**	**4.52**

(c) CorrMem lift V

V	$\mathcal{A}_b$	$\mathcal{A}_n$	$\mathcal{A}_T$	$\Delta\downarrow$
500	95.33	76.71	90.73	4.83
1000	95.35	**77.26**	91.01	**4.55**
2000	**95.82**	76.43	**91.27**	4.75
4000	95.64	75.29	91.00	4.85

B.2 Hyperparameter Ablations

Tab. 9 reports ModelNet exemplar-free sweeps of the fusion weight α, adapter bottleneck d', and CorrMem lift V, with defaults highlighted. In panel (a), following Eq. equation 10, decreasing α (e.g. 0.1) places more weight on ProtoMem, which treats every class through one normalised mean and therefore lifts novel-class accuracy to its peak ($\mathcal{A}_n$=79.14%), but weakens last-session accuracy and increases relative drop. Increasing α (e.g. 0.9) places more weight on CorrMem, recovering base-class discrimination while novel accuracy falls and Δ rises again. The default α=0.5 balances the two memories and attains the strongest $\mathcal{A}_T$. Panel (b) sweeps the adapter bottleneck d'. Widening d' from 32 to 256 steadily raises novel and last-session accuracy and lowers Δ, while base accuracy stays nearly flat, so extra adapter capacity mainly helps few-shot plasticity after the encoder is frozen. Gains beyond d'=64 are modest (91.27% to 91.40% in $\mathcal{A}_T$), so the default keeps a compact bottleneck without sacrificing the dual-memory operating point. Panel (c) varies the random-feature width of CorrMem. V=500, 1000, and 2000 stay close on last-session accuracy (90.73% to 91.27%), indicating that CorrMem is robust to the lift dimension once the ridge head has enough capacity; V=2000 is strongest on $\mathcal{A}_T$, V=1000 edges it on Δ, and V=4000 does not help further.