1. Complete Experimental Results

Figures 1 and 2 illustrate the comprehensive performance evolution trajectories across the entire incremental learning process on ImageNet-A and Cars-196 datasets, respectively. These figures compare our framework against 12 baseline methods across varying task granularities (5, 10, and 20 tasks). The results demonstrate that our method maintains a consistent and widening performance advantage as the number of incremental steps increases, validating its robustness in long-sequence incremental scenarios.

2. Ablation Studies

To empirically validate our hypothesis that catastrophic forgetting stems from the simultaneous collapse of the class-discriminative structure, we visualize the task-wise accuracy evolution on the Cars-196 dataset in Figure 3.

The "Base" heatmap (a) exposes the severity of structure collapse: without preservation mechanisms, the accuracy of the initial task precipitously drops from 94.8% to 28.7% by the final step. This degradation confirms that as the feature extractor evolves from $f_\theta^{(t-1)}$ to $f_\theta^{(t)}$, the unchecked deterioration of both intra-class integrity and inter-class separability renders the model incapable of distinguishing old classes.

The ablation results in (b) and (c) further isolate the contribution of each component. Employing only Adaptive Prototype Rectification (APR) (b) improves retention to 45.3% by dynamically correcting intra-class shape shifts, yet significant forgetting persists because the global geometric relationships remain unstable. Similarly, using only the Structural Consistency Constraint (SCC) (c) raises accuracy to 47.8% by stabilizing inter-class relationships, but this fails to address the intra-class structure degradation.

In contrast, our unified framework (d) demonstrates substantial synergistic efficacy. By simultaneously leveraging APR to preserve intra-class structure and SCC to maintain inter-class structure, we preserve the initial task accuracy at 63.0%—more than doubling the baseline retention rate. These results corroborate our core contribution: intra-class and inter-class structures are interdependent components, and their joint preservation is indispensable for preventing the collapse of the class-discriminative structure during continuous learning.

To quantify structural stability, Figure 4 tracks the mean L_2 distance of true class centers. The high deviation in "APR only", which is comparable to the Baseline , indicates that preserving intra-class structure alone—without inter-class constraints—cannot effectively preserve class-discriminative structure. In contrast, the lowest distance achieved by "Ours" demonstrates that the synergy of intra-class and inter-class preservation is indispensable to maximally mitigate catastrophic forgetting.

Figure 5 provides intuitive demonstration of feature space distortion and our method's effectiveness through t-SNE visualization.

In the ideal state after first task training, 10 categories form clear, compact, and mutually separated clustering structures. After incremental tasks, the baseline method leads to severe feature space degradation with blurred boundaries and overlapping categories. In contrast, our method effectively maintains inter-class separability and intra-class compactness, providing intuitive validation of our approach.

3. Implementation Details

All experiments were conducted on a server equipped with a NVIDIA RTX 4090 GPU. The software environment was based on Ubuntu 20.04, with key libraries including Python 3.9, PyTorch 2.4.1, and CUDA 12.4.

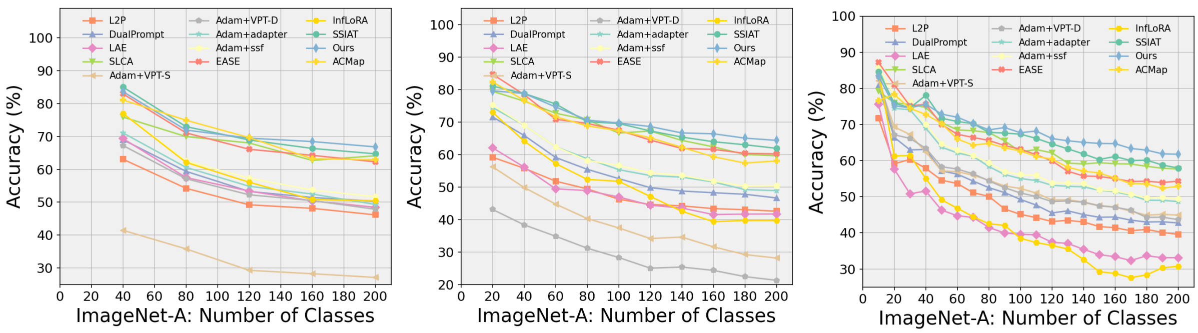


Figure 1. Performance evolution trends under different task settings on ImageNet-A dataset. Shows the changes in Last-Acc for our method compared with 12 baseline methods under 5-task (left), 10-task (middle), and 20-task (right) settings.

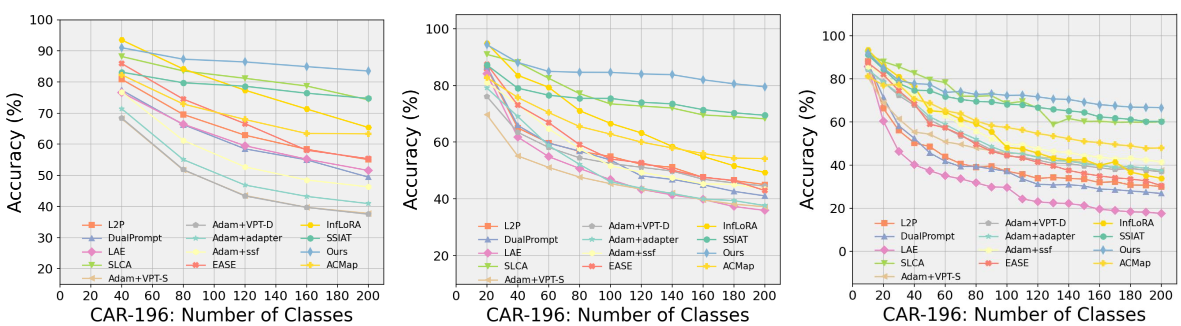


Figure 2. Performance evolution trends under different task settings on Cars-196 dataset. Shows the changes in Last-Acc for our method compared with 11 baseline methods under 5-task (left), 10-task (middle), and 20-task (right) settings.

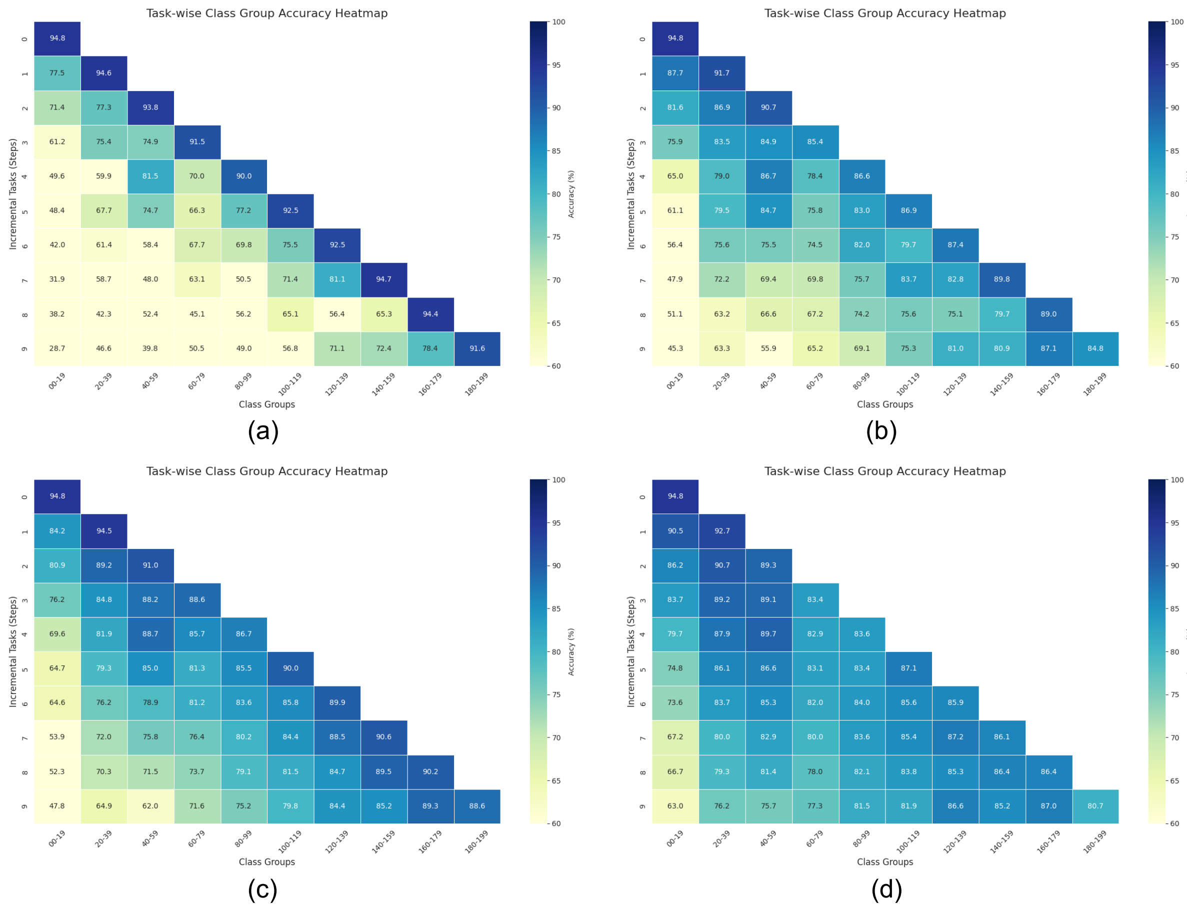


Figure 3. Ablation study of task-wise accuracy on the Cars-196 dataset (10 tasks). We compare the performance evolution of: (a) the Baseline, (b) Only APR (Intra-class preservation), (c) Only SCC (Inter-class preservation), and (d) Our unified framework. The horizontal axis represents the class groups corresponding to each task, while the vertical axis denotes the incremental learning steps. Each cell displays the test accuracy of a specific task group after the corresponding incremental step. The rapid "cooling" (lighter colors) in (a), (b), and (c) indicates severe forgetting, whereas the sustained high accuracy in the lower-left triangle of (d) validates that simultaneously preserving intra-class and inter-class structures significantly mitigates catastrophic forgetting.

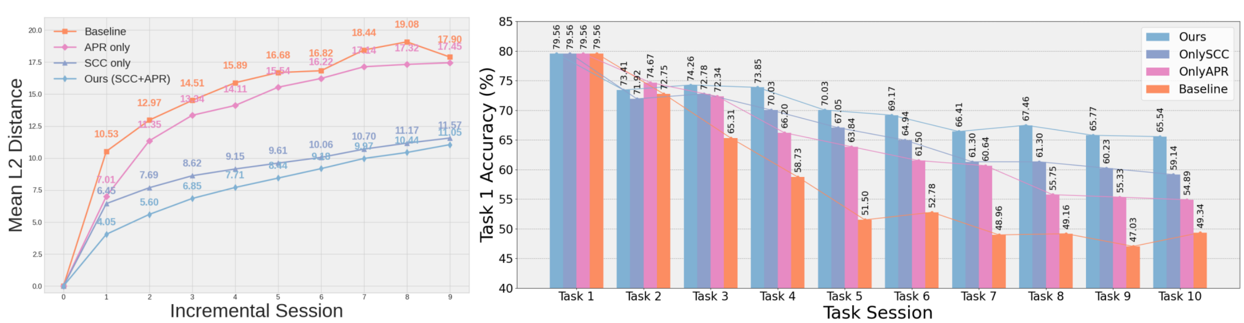
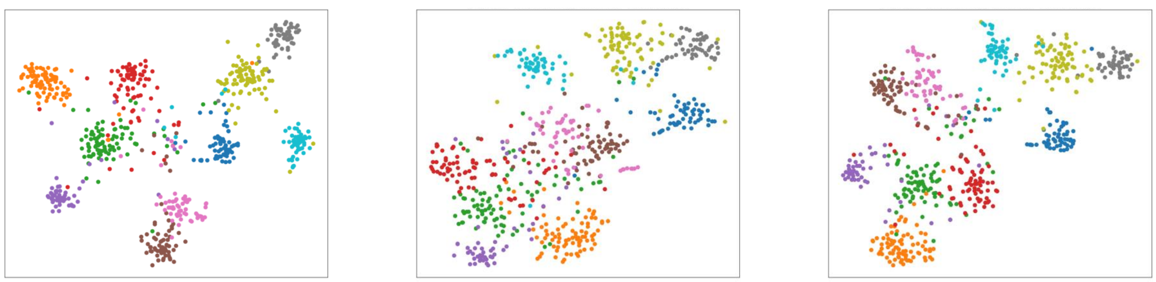


Figure 4. Prototype analysis under 10-task setting on ImageNet-A dataset. The left plot shows the changes in average L2 distance between the first task prototype centers and their true feature centers as tasks progress; the right plot shows the degradation of first task class accuracy during subsequent task learning processes.



Task1 Training Features

Baseline Features After Incremental Tasks

Our Method Features Under Same Conditions

Figure 5. t-SNE feature visualization comparison of the first task's 10 classes on ImageNet-A dataset. **Left**: Ideal feature distribution immediately after first task training; **Middle**: Baseline method after incremental tasks showing severe degradation; **Right**: Our method maintaining clear separability.