DDSF: Robust Few-Shot Learning via Disentangled Subspaces with Determinantal Point Process

Xulun Ye[1*] Yifan Mei[1*] Kun Zhou[2†] Zelei Wu[1] Jieyu Zhao [1†]
Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo, China
[2]School of Architecture & Urban Planning, Shenzhen University, Shenzhen, China
{yexulun, 2411100090, zhao_jieyu}@nbu.edu.cn, zhoukun@szu.edu.cn, zeleiwu214@gmail.com

A. Extended Experimental Results

A.1. Robustness in Extreme Low-Shot Settings

We adopted the 10-shot protocol in the main paper to align with DETA++ [53]. Here, we evaluate robustness under 5-shot and extreme 1-shot settings (where affected classes are entirely noise-dominated). As shown in Tab. A.1, DDSF maintains remarkable stability with limited samples. Under the 5-shot setting with 30% noise, DDSF achieves 79.9% accuracy, dropping only 5.4% compared to the 10-shot setting (85.3%). In the 1-shot 30% noise setting, it still reaches 60.6% (+13.8% over baseline).

While we acknowledge the inherent bias in subspace estimation under extreme data scarcity (e.g., 1-shot), DDSF's consistent superiority demonstrates the robustness of the Filter-Repair-Expand paradigm. In future work, we plan to explore incorporating generative augmentation to further stabilize subspace estimation in such limited scenarios.

Table A.1. Performance under low-shot settings.

Method	1-shot (0%)	1-shot (30%)	5-shot (0%)	5-shot (30%)
Baseline	58.0	46.8	80.2	71.0
DDSF	**71.9**	**60.6**	**88.1**	**79.9**

A.2. Cross-Domain Generalization on Real-World Datasets

To verify practical generalization, we evaluate our model—trained exclusively on the Meta-Dataset—directly on three distinct cross-domain datasets: EuroSAT (remote sensing), CropDisease (agriculture), and Animal-10N (real-world natural noise). Notably, the Animal-10N dataset inherently contains approximately 8% natural label noise.

Results in Tab. A.2 demonstrate strong cross-domain generalization. DDSF achieves peak accuracies of 89.1% on EuroSAT and 92.5% on CropDisease, while also effectively handling the inherent noise in Animal-10N to maintain a robust 81.1% accuracy.

Table A.2. Zero-shot transfer performance.

Dataset	5-S (0%)	5-S (30%)	10-S (0%)	10-S (30%)
EuroSAT	87.7	76.0	89.1	81.7
CropDisease	90.1	81.3	92.5	85.7
Animal-10N	77.2	-	81.1	-

B. Algorithm Analysis and Visualizations

B.1. Computational Overhead Analysis

We analyze the computational efficiency of the proposed DDSF framework compared to the baseline. As shown in Tab. B.1, the training time marginally increases by only 0.5 s per task, and inference latency increases by a mere 4 ms per image. This high efficiency is achieved because: (1) the diffusion-based repair process is lightweight, requiring only $T = 15$ steps operated in a low-dimensional feature space; and (2) although the DPP matrix operations theoretically possess $O(N^3)$ complexity, the actual support set size N in few-shot tasks is extremely small (e.g., $N = 10$), rendering the computational cost essentially negligible.

Table B.1. Computational overhead comparison.

Method	Train. (s/task)	Infer. (ms/img)	Diff Steps (T)
Baseline	1.53 s	16 ms	-
DDSF	2.06 s	20 ms	15

B.2. Visualization of Feature Disentanglement

To intuitively validate the effectiveness of our feature disentanglement module, we provide t-SNE visualizations of the feature space in Fig. B.1. In the original feature space

*Equal contribution.
†Corresponding author.

(Fig. B.1a), noisy samples blur the decision boundaries and distort the class representations. In contrast, after processing through our framework, the purified unique subspace (Fig. B.1b) exhibits clear and distinct cluster separation. By rectifying noise via the DPP-guided repair module and maximizing the subspace volume to ensure diversity, DDSF effectively constructs robust and discriminative class representations.

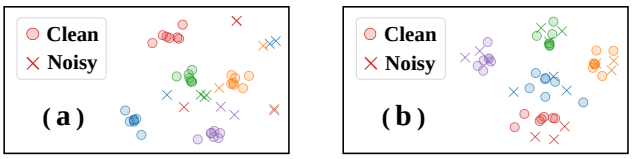

Figure B.1. t-SNE visualization. (a) Original noisy features exhibit blurred boundaries. (b) DDSF's purified subspace achieves distinct cluster separation.