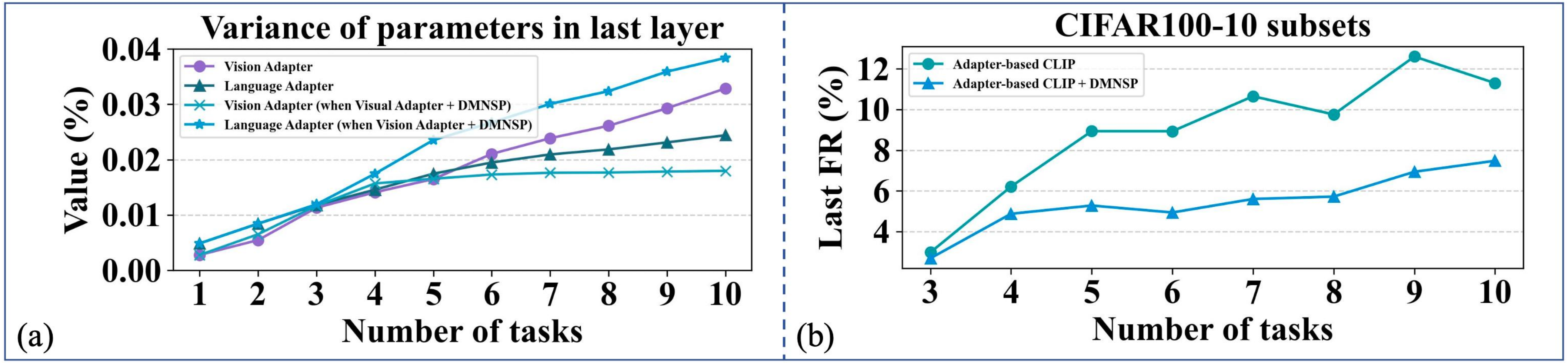
# Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning

# DMNSP

## Motivation

Overlooking modality-specific dynamics hinders effective forgetting mitigation in adapter-based VLM CL.

- **Suboptimal forgetting prevention**.
- **Inadequately explored modality parameter distribution shifts in VLM**.




**Visual modality**'s broader distribution increases forgetting risk.

### Our hypothesis

- Inhibiting shifts in **visual parameter distributions** can mitigate forgetting.



**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# DMNSP

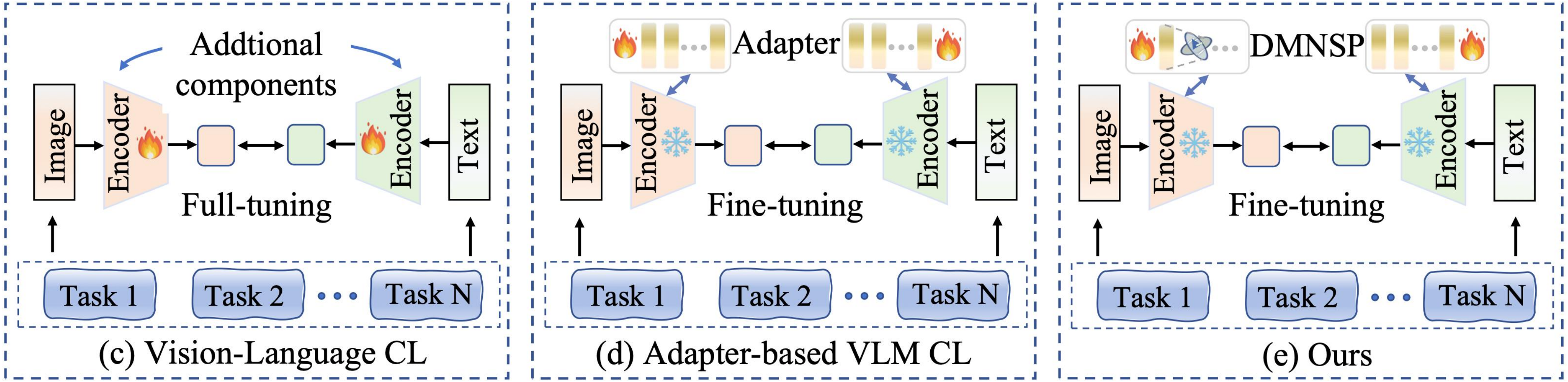
## Challenge

### Our hypothesis

- Inhibiting shifts in **visual parameter distributions** can mitigate forgetting.

### Contribution

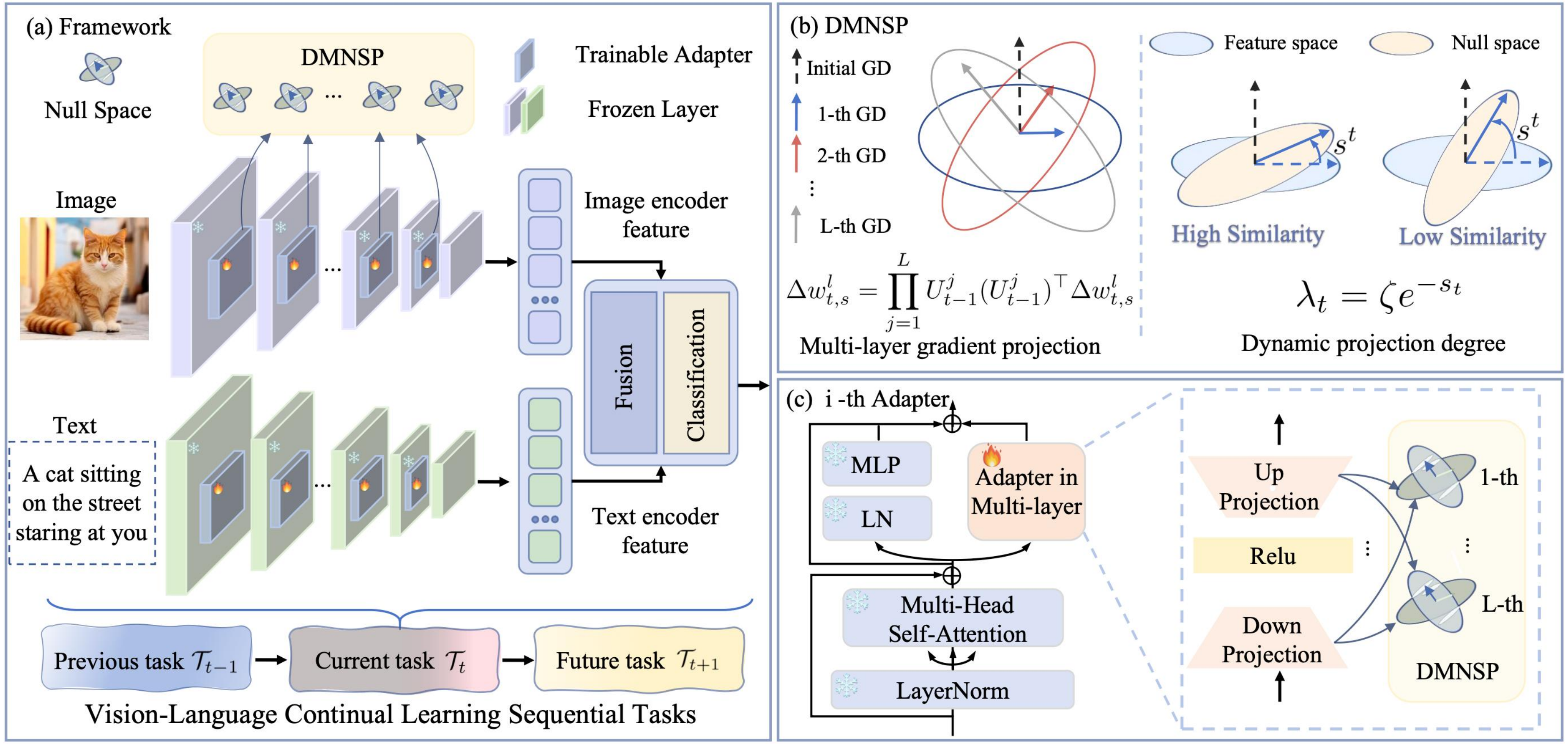
- We propose **asymmetric adapter training** to address modality-induced forgetting..
- We design **multi-layer** gradient projection with **dynamic** coefficients for CL.
- We achieve **state-of-the-art** in different CIL settings across different datasets.






**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# DMNSP
## Framework




$$\Delta \mathbf{w}_{t,s}^{l} = \prod_{j=1}^{L} \mathbf{U}_{t-1}^{j}(\mathbf{U}_{t-1}^{j})^\top \mathbf{g}_{t,s}^{l}$$

$$\lambda_t = \zeta e^{-s_t}$$

$$\Delta \mathbf{w}_{t,s}^{l} = \prod_{j=1}^{L} \lambda_t^j \mathbf{U}_{t-1}^{j}(\mathbf{U}_{t-1}^{j})^\top \mathbf{g}_{t,s}^{l}$$



**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# DMNSP

## Experiments

| Method | Venue | 10 subset | | 20 subset | |
|---|---|---|---|---|---|
| | | Avg. ↑ | Last ↑ | Avg. ↑ | Last ↑ |
| L2P [42] | CVPR'22 | 80.83$_{\pm 1.39}$ | 74.60$_{\pm 0.90}$ | 78.39$_{\pm 0.94}$ | 72.09$_{\pm 1.12}$ |
| DualPrompt [41] | ECCV"22 | 81.39$_{\pm 1.25}$ | 74.87$_{\pm 0.85}$ | 79.12$_{\pm 1.27}$ | 71.69$_{\pm 1.06}$ |
| ESN [40] | AAAI'23 | 81.63$_{\pm 1.10}$ | 75.11$_{\pm 0.36}$ | 77.95$_{\pm 0.76}$ | 70.57$_{\pm 0.62}$ |
| CODAprompt [36] | CVPR'23 | 81.32$_{\pm 1.01}$ | 75.51$_{\pm 0.81}$ | 78.07$_{\pm 0.40}$ | 72.25$_{\pm 0.78}$ |
| LAE [12] | CVPR'23 | 76.71$_{\pm 0.10}$ | 71.70$_{\pm 0.39}$ | 73.72$_{\pm 0.05}$ | 66.98$_{\pm 0.35}$ |
| InfLoRA [25] | CVPR'24 | 80.82$_{\pm 0.24}$ | 75.65$_{\pm 0.14}$ | 77.28$_{\pm 0.45}$ | 71.01$_{\pm 0.45}$ |
| EASE [48] | CVPR'24 | 81.73 | 76.17 | - | - |
| CPrompt [13] | CVPR'24 | 82.92$_{\pm 0.70}$ | 77.14$_{\pm 0.11}$ | 81.46$_{\pm 0.93}$ | 74.79$_{\pm 0.28}$ |
| RAPF [18] | ECCV'24 | 85.85 | 79.62 | 86.28 | 79.62 |
| VPT - NSP$^2$ [29] | NeurIPS'24 | 84.84$_{\pm 0.12}$ | 78.88$_{\pm 0.50}$ | - | - |
| Adapter-based CLIP + GPM [34] | - | 86.18$_{\pm 0.08}$ | 80.11$_{\pm 0.12}$ | 86.73$_{\pm 0.19}$ | 80.65$_{\pm 0.14}$ |
| Adapter-based CLIP+ TRGP [26] | - | 86.24$_{\pm 0.10}$ | 81.29$_{\pm 0.09}$ | 86.07$_{\pm 0.17}$ | 81.21$_{\pm 0.13}$ |
| MoE4Adapters [45] + DMNSP | - | 86.45$_{\pm 0.11}$ | 81.87$_{\pm 0.16}$ | **87.06**$_{\pm 0.20}$ | 81.61$_{\pm 0.18}$ |
| Adapter-based CLIP + DMNSP | - | **87.49**$_{\pm 0.07}$ | **81.94**$_{\pm 0.15}$ | 86.81$_{\pm 0.16}$ | **82.72**$_{\pm 0.09}$ |

| Method | 5 subset | | 10 subset | | 20 subset | |
|---|---|---|---|---|---|---|
| | Avg. ↓ | Last ↓ | Avg. ↓ | Last ↓ | Avg. ↓ | Last ↓ |
| Adapter-based CLIP | 3.34 | 7.97 | 4.56 | 8.75 | 5.33 | 9.79 |
| Adapter-based CLIP + GPM | 3.30 | 7.98 | 4.50 | 8.69 | 5.35 | 9.27 |
| MoE4Adapters + DMNSP | 2.26 | **4.16** | 3.12 | 4.88 | **3.23** | **6.86** |
| Adapter-based CLIP + DMNSP | **2.15** | 4.84 | **2.65** | **4.84** | 3.65 | 7.11 |

| Method | 10 subset | | 20 subset | | 50 subset | |
|---|---|---|---|---|---|---|
| | Avg. ↓ | Last ↓ | Avg. ↓ | Last ↓ | Avg. ↓ | Last ↓ |
| Adapter-based CLIP | 7.14 | 11.39 | 7.43 | 12.61 | 9.34 | 14.52 |
| Adapter-based CLIP + GPM | 7.14 | 11.29 | 7.38 | 12.61 | 9.33 | 14.51 |
| MoE4Adapters + DMNSP | 4.95 | 9.28 | 5.95 | 10.67 | 7.47 | 12.44 |
| Adapter-based CLIP + DMNSP | **4.36** | **7.48** | **4.73** | **9.52** | **7.39** | **12.42** |

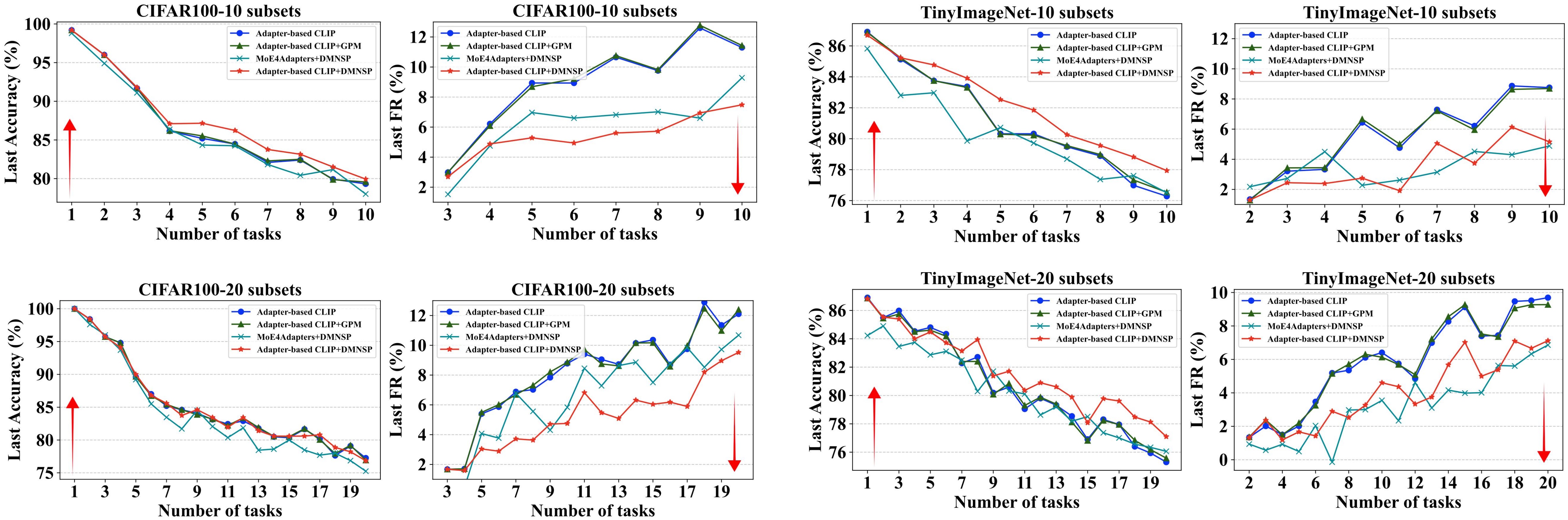
- Achieve competitive results across **5** datasets and over **10** settings.
- Achieve a **lower** forgetting rate.



**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# DMNSP

## Visualization



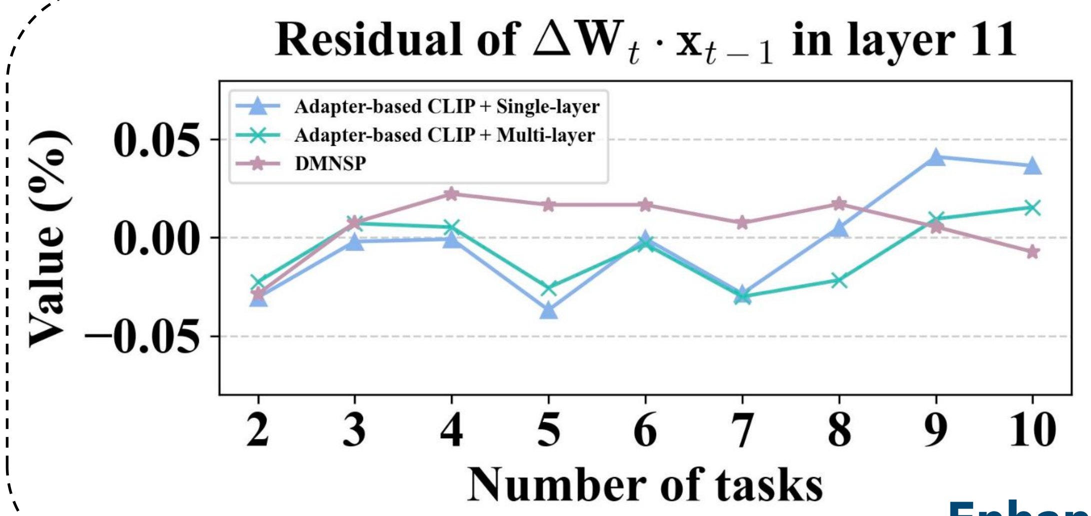

➢ Achieve a **higher** last accuracy curve and a **lower** forgetting curve.

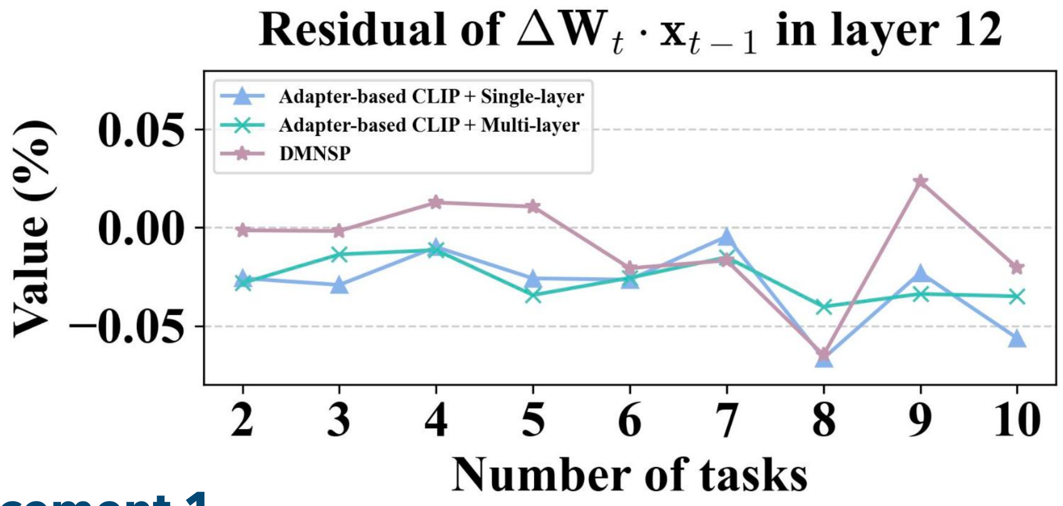


**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# DMNSP

## New Enhancements




**Enhancement 1**

| Method | Train Params ↓ | Memory Use ↓ | Times ↓ |
|---|---|---|---|
| LWF [24] | 149.6M | 32172MiB | 1.54s/it |
| LWF-VR [6] | 149.6M | 32236MiB | 1.51s/it |
| ZSCL [47] | 149.6M | 26290MiB | 3.94s/it |
| MoE-Adapters [45] | 59.8M | 22358MiB | 1.58s/it |
| Ours | **7.8M** | **21116MiB** | **0.23s/it** |

**Enhancement 2**



**[ICCV 2025]** Dynamic Multi-Layer Null Space Projection for Vision-Language Continual Learning.

# Thank you for listening.