

# SD-LoRA: SCALABLE DECOUPLED LOW-RANK ADAPTATION FOR CLASS INCREMENTAL LEARNING

**Yichen Wu<sup>1,2,\*</sup>, Hongming Piao<sup>1,\*</sup>, Long-Kai Huang<sup>4,†</sup>, Renzhen Wang<sup>3</sup>, Wanhua Li<sup>2</sup>,  
Hanspeter Pfister<sup>2</sup>, Deyu Meng<sup>3,6</sup>, Kede Ma<sup>1,†</sup>, Ying Wei<sup>5,†</sup>**

<sup>1</sup>City University of Hong Kong, <sup>2</sup>Harvard University, <sup>3</sup>Xi'an Jiaotong University,

<sup>4</sup>Tencent AI Lab, <sup>5</sup>Zhejiang University, <sup>6</sup>Pengcheng Laboratory

- Problem / objective
  - 기존 Continual Learning 방법들의 scalability 문제
- Contribution / Key idea
  - Scalable Decoupled LoRA (SD-LoRA)
    - LoRA에서 크기 및 방향을 분리하여 학습

- **SD-LoRA: CL method with foundation models**

1. Rehearsal-free
2. Inference-efficient
3. End-to-end optimized

- **two variants of SD-LoRA**

1. rank reduction
2. knowledge distillation

,based on the observations: "The importance of the incrementally learned LoRA directions diminishes as CL progresses"

- **How SD-LoRA mitigates catastrophic forgetting?**

aa