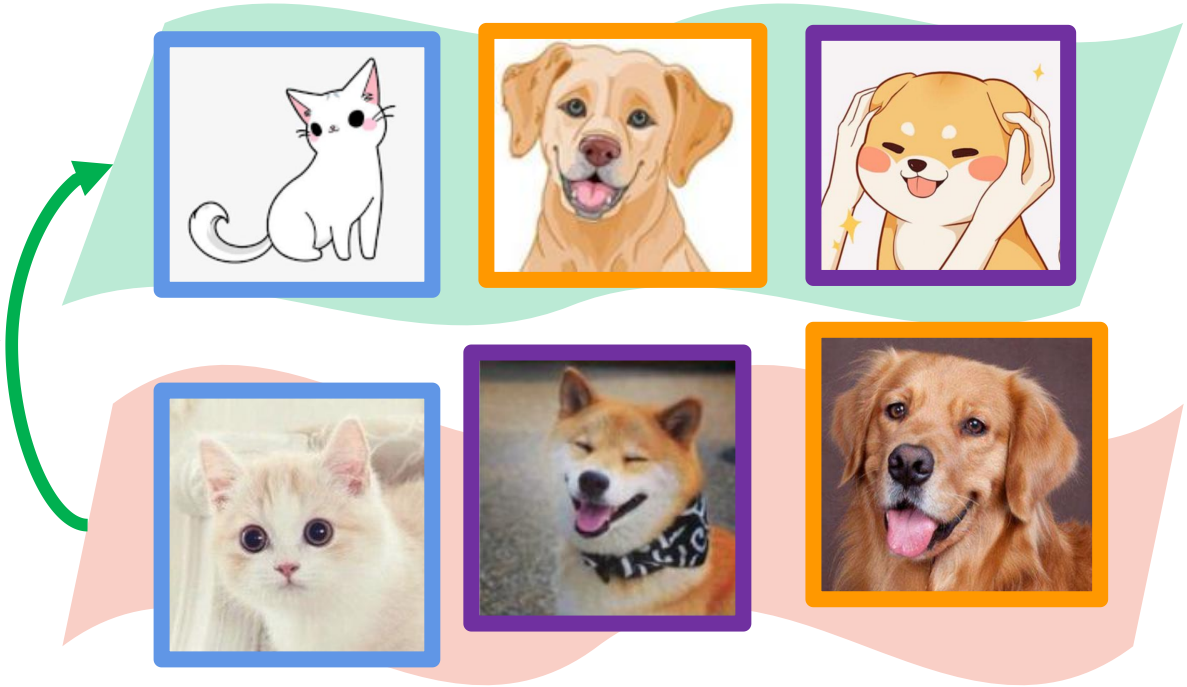
Unsupervised Domain Adaptation with Hierarchical Gradient Synchronization

Lanqing Hu, Meina Kan, Shiguang Shan, Xilin Chen
Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), China
lanqing.hu@vipl.ict.ac.cn, {kanmeina, sgshan, xlchen}@ict.ac.cn

Hierarchical Gradient Synchronization

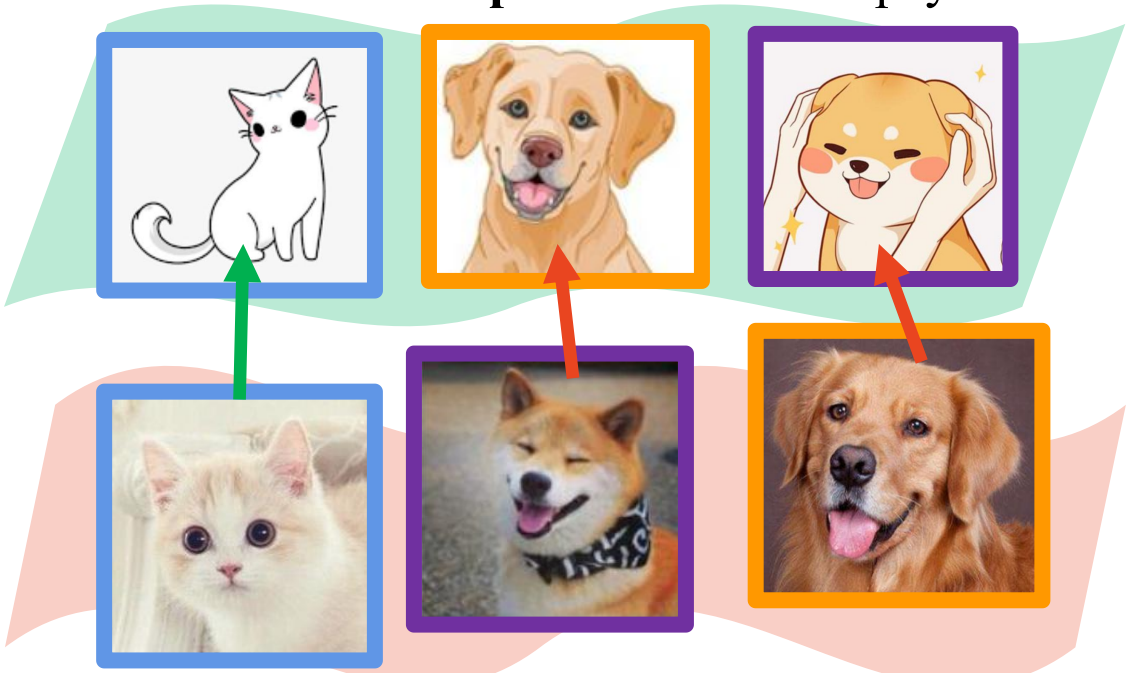
- **Problem**: Unsupervised Domain Adaptation (UDA)
  - Transfer the knowledge from labeled source to unlabeled target domain
    - Align the feature distribution
      - Marginal (global) and conditional (local) distribution

How to align distribution

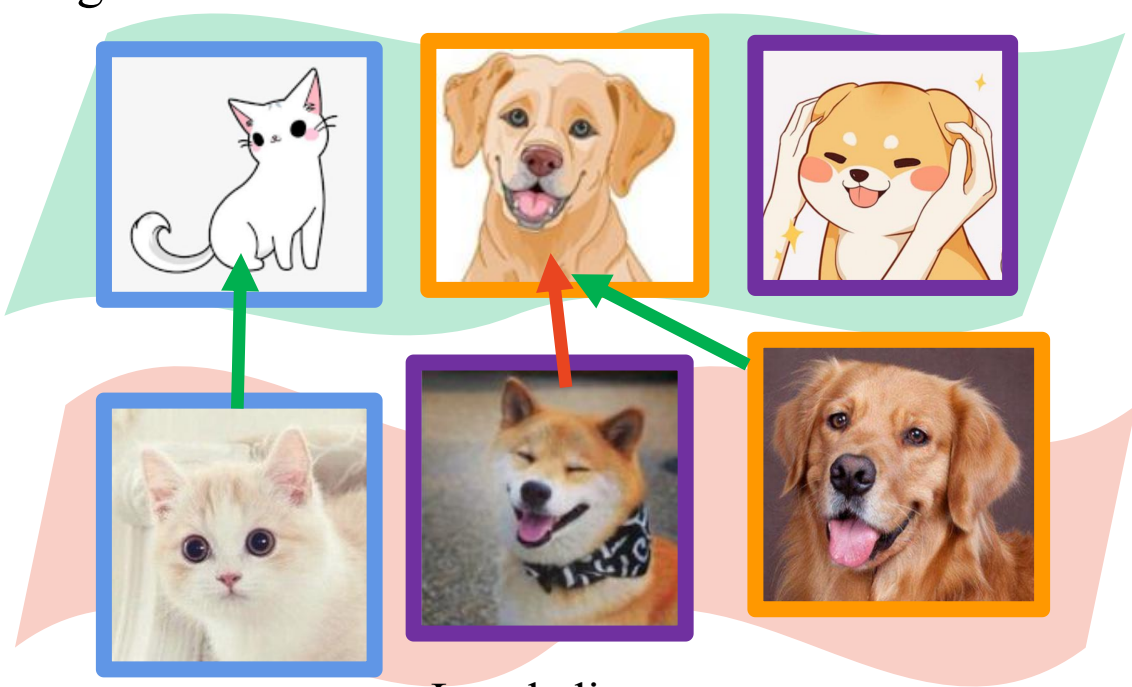


Limitation

- **Limitation** of existing methods
  - Independently aligning **global** and **local** distributions
    - Global alignment: Ignore fine category information->exist misaligned categories
    - Local alignment: Ignore global information->overfit to the same category
    - **Compromise** of simply combining them together



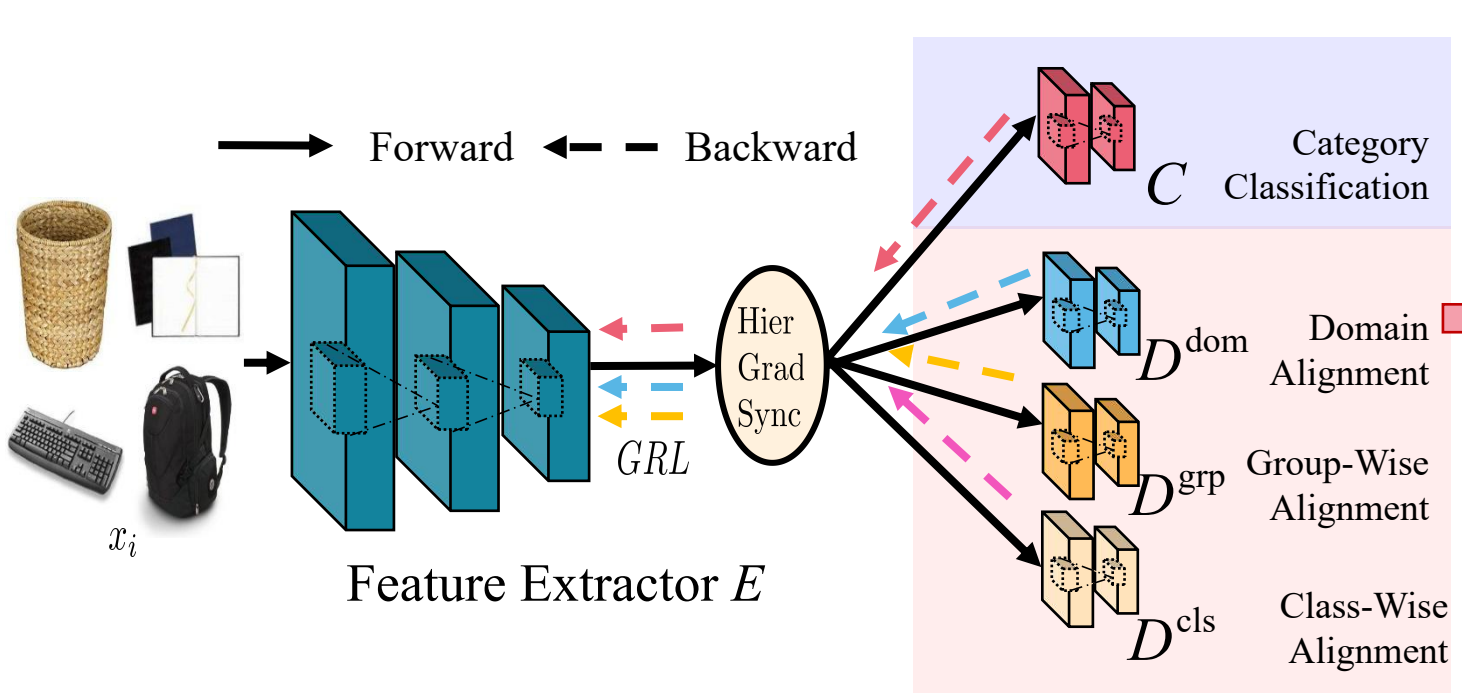
Global alignment

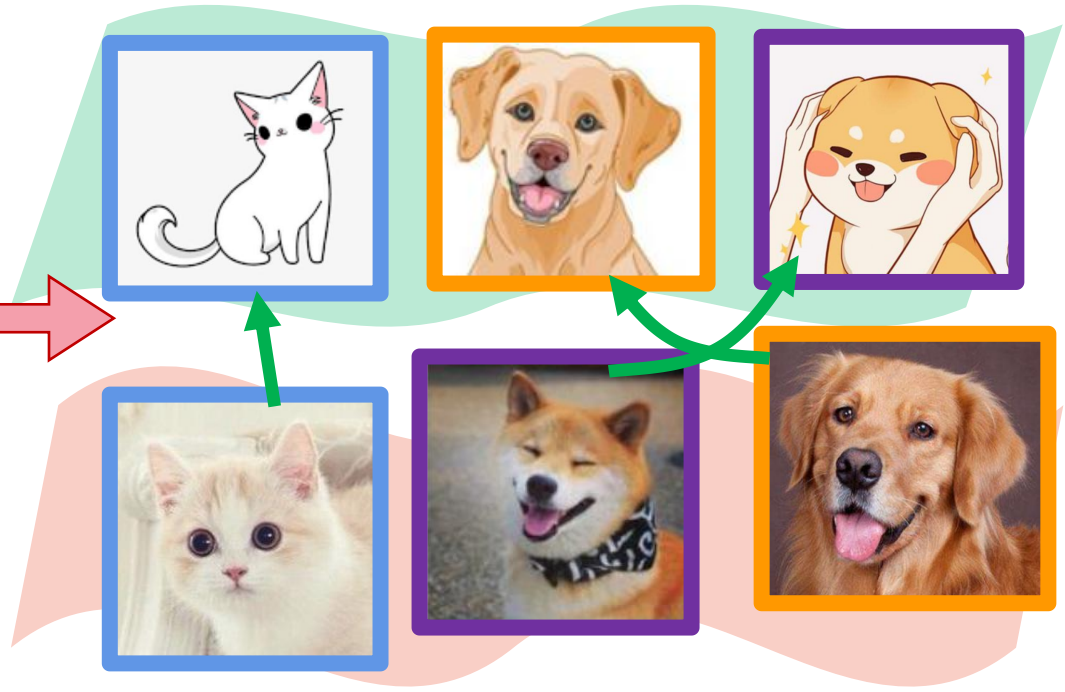


Local alignment

Method

- Make the global and local alignments consistent
  - **Gradient Synchronization:** make the global and local alignment gradients synchronization on direction and magnitude
  - Obtain better performance on commonly used UDA datasets

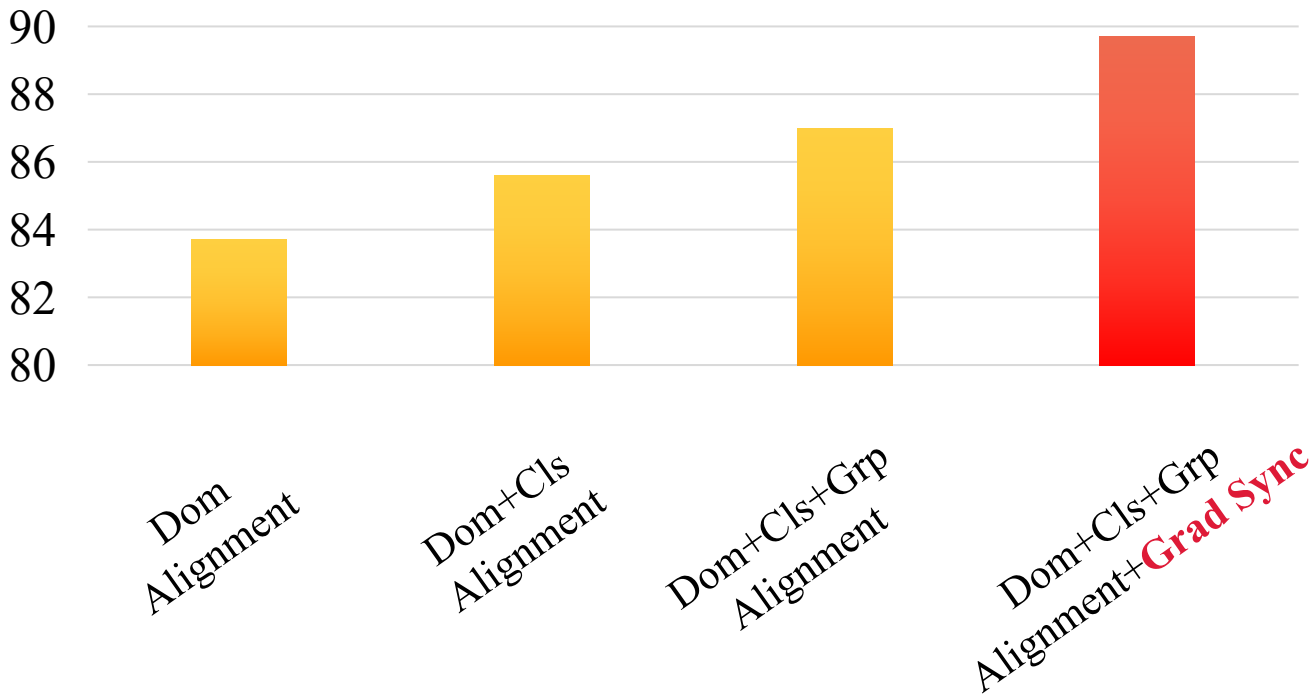




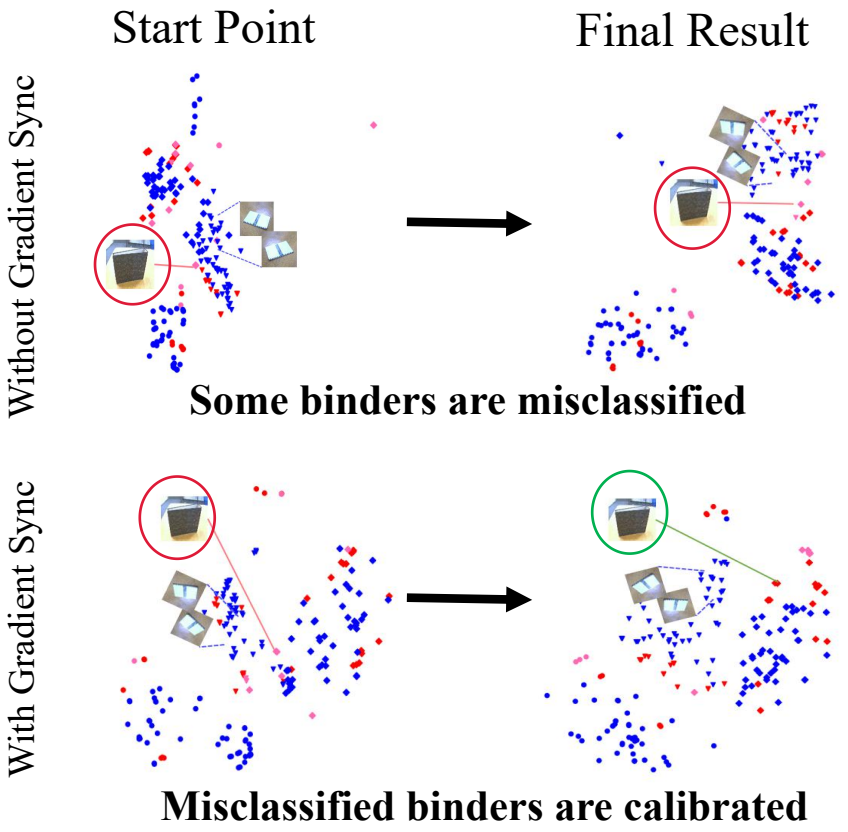
The three aligning path (**green**) should be **synchronized** to be in the same **direction** and **magnitude** in each step.

Main Result

Ablation Study on Classication Task on A->W of Office-31



Visualization of Optimization process



Discovering the **relation** between global and local alignments, e.g., **gradient synchronization**, shows its rationality for UDA problem and deserves more exploration.