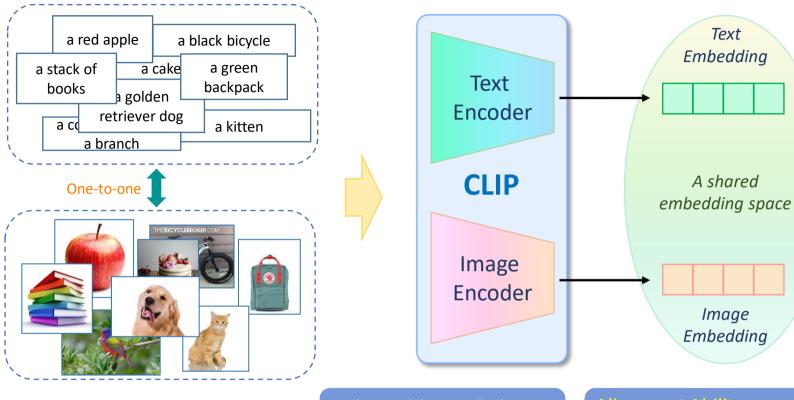
Why CLIP is Ideal for Domain Generalization (DG) and Domain Adaptation (DA)?



Pre-trained with 400,000,000 image-text pairs which span various domains and tasks

Rich World Knowledge

Encodes extensive semantic and contextual information

Alignment Ability

Aligns textual and visual modalities in a shared space

Zero-Shot Capacity

Perform unseen tasks without task-specific fine-tuning

Robustness

Maintains performance under domain shift

Generalization

Performs well on unseen data distributions

Adaptation

Adapts to new tasks with minimal supervision

Feature Distribution Alignment

Bridges the gap between source and target domains

Prior Knowledge Utilization

Leverages pre-trained knowledge for rapid domain adaptation