**Learning CNN on ViT: A Hybrid Model to Explicitly Class-specific Boundaries for Domain Adaptation**

**1. Introduction**

- ViT는 global representation을 포착 가능하고, 큰 dataset에 generalization 능력을 가짐

- 반면, CNN은 locality, spatial invariance의 inductive bias를 가정해, local feature를 잘 포착하며dataset이 작을 때 강점을 보임

-

**2. Related Works**

Co-training

-

**3. Methodology**

- ViT branch와 CNN branch로 우선 supervised training을 진행함

- 1) finding stage

- discrepancy loss는 두 probability output 간 차이의 class 평균으로 계산함