

CLIP Medical Variants: Contrastive Learning

CLIP (Contrastive Language-Image Pre-training): 영상과 텍스트를 공통 임베딩 공간에 정렬하여 **제로샷 학습** 가능



대조 학습 (Contrastive Learning)

- 매칭 영상-텍스트 쌍: 높은 유사도
- 비매칭 쌍: 낮은 유사도

영상-텍스트 정렬 효과

- 제로샷 분류 가능
- 텍스트 쿼리로 영상 검색

- InfoNCE loss로 구별 학습

- 멀티모달 표현 학습