

Medical Image Encoders



CNN (Convolutional Neural Networks)

전통적이지만 강력한 영상 특징 추출 방법

- 공간적 계층 구조 학습
- 로컬 패턴 인식에 강점
- ResNet, DenseNet, EfficientNet
- 의료 영상의 텍스트, 경계 검출



Vision Transformer (ViT)

패치 기반 어텐션으로 글로벌 관계 학습

- Self-attention 메커니즘
- 긴 범위 의존성 포착
- 대규모 데이터셋에서 우수
- 병변 간 관계, 전체 맥락 이해



의료 특화 인코더

의료 도메인에 사전 학습된 모델

- MedCLIP, BiomedCLIP
- 대규모 의료 영상-텍스트 쌍 학습



Hybrid Architectures

CNN과 Transformer의 결합

- CNN: 로컬 특징 + ViT: 글로벌 관계
- CoAtNet, MaxViT

- 도메인 지식 내재화
 - 적은 데이터로 높은 성능
- 계산 효율성과 성능 균형
 - 다양한 스케일의 병변 검출

의료 영상 전처리 파이프라인



특징 추출

고차원 영상을 저차원 벡터 표현으로 변환하여 의미 있는
임상 특징 포착

다운스트림 태스크

추출된 특징을 분류, 검출, 세분화, 리포트 생성 등에 활용