Causal 추론이 가능한 골다공증 분류 모델 개발

박지후 김세한 윤영근

경희대학교 컴퓨터공학과

@khu.ac.kr, @khu.ac.kr, yuns49@khu.ac.kr

Development of an osteoporosis classification model capable of causal inference

, younggeun Yun

Department of Computer Science and Engineering, KyungHee University

Masked Autoencoder와 Proxy Tasks를 통한 Self-supervised Learning 방법을 활용하여 적은 데이터로도 효과적인 학습이 가능하도록합니다. 이를 위해 Masked Autoencoder를 사용해 중요한 특징을 자동으로 인코딩하고, Proxy Tasks를 통해 모델이 자기 자신의 출력을 예측하는 능력을 향상시며 다양한 방법론을 통해 어떤 모델이 효율적인지를 분석합니다. 특히, 어깨 X-ray 이미지 데이터를 활용하여 골다공증을 진단할 수 있는 데이터 효율적인 분류 모델을 개발하고, Causal 추론을 가능하게 하는 방법론을 개발하는 것이 목표입니다. 이를 위해 다음과 같은 주요 방법론들을 활용합니다

**1. 서 론**

골다공증으로 인한 골절은 환자의 생활 질을 크게 저하시킬 뿐만 아니라, 심각한 경우 사망에 이르게 할 수 있다. 이떄문에 골다공증을 정확하고 초기에 진단하는 것은 골절과 같은 심각한 합병증을 예방하며 치료에 수반되는 의료 비용을 줄일 수 있다. 기존 골다공증 진단의 표준 방법은 골밀도 측정이지만, 이는 비용이 많이 들고 모든 지역에서 쉽게 접근할 수 없어 비용 및 접근성 문제가 발생하게 된다. 또한 증상이 뚜렷하지 않아 많은 환자들이 질환을 인지하지 못하고 진단받지 못하는 경우가 많다.

최근 인공지능 기술, 특히 기계학습과 딥러닝의 발전은 의료 이미징 분야에서 정확도를 높이고 진단 속도를 개선할 수 있는 새로운 가능성을 열고 있다. 이를 통하여 X-ray와 같은 이미징 데이터를 활용해 골다공증을 더 정확하고 신속하게 진단할 수 있는 모델을 개발하는 것은 의료 접근성을 향상시키고 비용을 절감하는 효과적인 방법이 될 수 있음을 의미한다.

이러한 배경을 바탕으로, 흉부 X-ray 및 어깨 X-ray 이미지만을 사용하여 골다공증을 효과적으로 진단할 수 있는 인공지능 기반의 효율적인 분류 모델을 개발하는 것이 본 논문의 목표이며 이는 기존의 진단 방법을 보완하고, 더 많은 환자들이 초기에 정확한 진단을 받을 수 있게 함으로써 골다공증으로 인한 부담을 줄이는 데 기여할 것을 기대한다

본 논문은 2절에서 MAE와 Proxytask를 통한 모델 비교 방법에 대해 조사한다. 2절에서 조사한 내용을 바탕으로 3절에서는 다양한 모델간 비교 방식을 설계하고 이를 이용한 유의미한 causal 추론이 가능한 최선의 모델을 찾는다, 4절에서는 결론 및 향후 연구 방향을 제시한다.

**2. 연구배경 및 관련연구**

* 1. MAE(Masked Auto Incoder)

MAE는 입력 이미지에서 무작위로 패치를 마스킹하고, 인코더가 보이는 패치들만 처리하도록 하며, 디코더는 잠재적인 표현과 마스크 토큰을 통해 원본 이미지를 재구성한다. 높은 비율(예: 75%)의 이미지를 마스킹하는 것이 효과적인 자기지도 학습 과제를 생성하고, 이는 비전 태스크에서의 학습에 도움이 된다.

이러한 설계를 통해 훈련 속도를 3배 이상 향상시키고 정확도를 개선할 수 있으며 큰 모델을 효율적으로 학습시킬 수 있다, 이러한 접근 방식은 이미지 인식에서 높은 용량의 모델이 잘 일반화되는 것을 가능하게 한다.[1]

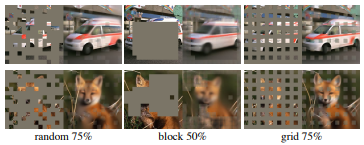


그림1

* 1. ViT(Vision Transform)

흉부 X-ray 이미지를 사용하여 ViT를 사전 훈련시키고, 이를 통해 ViT가 다중 라벨 흉부 질병 분류에서 상태 예술 CNN 모델과 비교할 때 유사하거나 더 나은 성능을 보인다. ViT에 대한 사전 훈련은 주로 마스크된 영역을 복원하는 작업을 통해 수행되며, 이미지의 90%를 마스킹하고 나머지 가시적인 패치로부터 복원을 시도한다. 이 방식은 흉부 X-ray 이미지의 특성상 높은 정보 중복성을 가지고 있기 때문에 효율적이다.

이를 통하여 ViT가 흉부 X-ray 이미지 분석에서 CNN을 대체할 가능성을 탐구하고, 의료 이미지 분야에서의 ViT 활용 방안을 제시합니다. [2]

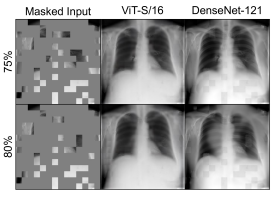


그림2

* 1. Grad-Cam

Grad-CAM(Gradient-weighted Class Activation Mapping)은 목표 개념에 대한 그라디언트를 활용하여 마지막 컨볼루션 레이어로의 그라디언트 흐름을 기반으로 중요 영역을 강조하는 지도를 생성한다. 다양한 CNN 모델에 적용 가능하며 고해상도 시각적 설명과 결합하여, 클래스 구별이 가능한 고해상도 시각화를 생성할 수 있다. 이를 통해 모델이 특정 결정을 내릴 때 중요하다고 판단한 이미지의 특정 영역을 상세하게 보여준다. 이미지 분류, 시각적 질문 응답(VQA) 모델에 적용되었다. [3]

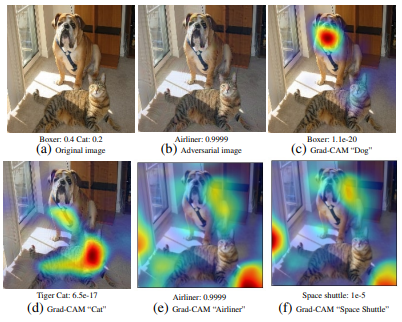


그림3

**3. MAE와 Proxytask를 활용한 골다공증 분류 모델 개발**

본 절에서는 제한된 X-ray 이미지 데이터를 이용하여 어깨 골다공증을 판별하는 모델 개발 방안을 제안한다. 상대적으로 적은 어깨 X-ray이미지데이터의 양을 보완하고 흉부 X-ray와 MAE, ViT등의 방법론을 통한 문제 해결을 위해 다음과 같은 5가지의 비교 방법을 제시한다.

1. ImageNet 사전학습 데이터 -> 어깨 X-ray (CNN, ViT)

이 접근법에서는 ImageNet 데이터로 사전 학습된 모델(CNN과 ViT)을 사용하여 어깨 X-ray 이미지를 분석하고 골다공증을 진단한다. 이미 효과적인 특징 추출 능력을 갖춘 모델을 사용하기 때문에, 데이터가 부족한 상황에서도 빠르게 초기 성능을 얻을 수 있으며 비교적 복잡한 모델 설계 없이, 널리 사용되는 네트워크를 적용할 수 있어 접근성이 높다.

이 방법은 시작 단계에서 빠른 결과를 원하거나, 특별히 희귀한 조건의 데이터에 접근하기 어려운 경우에 특히 유용할 수 있다. 그러나 의료 분야의 특성과 도메인 간 차이를 충분히 고려하고, 추가 실험을 통해 ImageNet 데이터로 학습된 모델이 의료 영상 분석에 효과적인지 평가해야 한다.

2. 흉부 X-ray MAE -> 어깨 X-ray (CNN, ViT)

흉부 X-ray 이미지를 사용하여 Masked Autoencoder (MAE)를 사전 훈련시킨 후, 어깨 X-ray 이미지 분석에 적용하여 골다공증을 판별하는 데 사용한다. 흉부 X-ray 이미지를 사용하여 MAE를 통한 사전 학습을 수행하는 과정은 비지도 학습 방식으로 진행되며, 이미지의 일부를 마스킹하고 남은 정보를 기반으로 원본 이미지를 재구성하여 모델이 중요한 특징을 학습하도록 한다..

MAE는 대규모 비지도 데이터에서 강력한 특징 추출 능력을 제공하므로 흉부 X-ray 데이터의 복잡성을 학습하여 어깨 X-ray 이미지 분석에 적용할 때 더 나은 성능을 발휘할 것을 기대한다.

3. 흉부 X-ray MAE -> 어깨 X-ray Proxy Task -> 어깨 X-ray (CNN, ViT)

이 접근법은 흉부 X-ray 이미지를 이용해 MAE로 사전 학습을 실시한 후, 이를 어깨 X-ray 이미지에 적용하기 전에 Proxy Task를 한번더 거치는 방식입니다. 최종적으로, CNN과 Vision Transformer (ViT)를 사용하여 어깨 X-ray 이미지를 분석하고 골다공증을 진단합니다. 이 과정은 복잡하지만, 높은 정밀도와 특화된 진단 능력을 목표로 하며 최종적으로 미세조정된 모델을 이용하여 어깨 X-ray 이미지를 분석하고 골다공증을 진단합니다.

4. 흉부 X-ray MAE -> 어깨 X-ray MAE(center/random) -> 어깨 X-ray (CNN, ViT)

흉부 X-ray로 MAE를 먼저 학습하고, 이후 어깨 X-ray에 대해 추가로 MAE 학습을 수행합니다. 두 단계의 MAE 학습을 통해 더욱 정교한 특징 학습이 가능하며, 어깨 X-ray 이미지의 미묘한 변화를 더 잘 포착할 수 있습니다. 다만 학습 과정이 복잡하고, 두 데이터셋의 특성을 모두 잘 이해해야만 효과적인 학습이 가능합니다.

5. 흉부 X-ray MAE -> 흉부 X-ray proxy -> 어깨 X-ray(CNN, ViT

흉부 X-ray 데이터를 사용하여 Masked Autoencoder(MAE)를 사전 훈련시키고, 이후 흉부 X-ray에 대한 Proxy Task로 미세조정한 다음, 이 모델을 어깨 X-ray 이미지에 적용하여 골다공증을 진단하는 방법을 고려한다.

사전 훈련된 MAE 모델을 흉부 X-ray 데이터에 더 적합하게 조정하여 특정 질병의 식별 능력을 개선합니다. 흉부에서 학습된 특징 인코딩과 질병 판단 기법을 어깨 X-ray 이미지 분석에 적용하여 골다공증의 유무를 판단한다. Proxy Task를 통해 특정 질병 식별에 필요한 세밀한 조정이 가능하며, 이는 모델의 진단 정확도를 높일 수 있기를 기대한다.

4. 결론 및 향후연구

본 논문에서는 CNN흉부 X-ray를 이용한 Masked Autoencoder (MAE)를 통한 사전 훈련과 어깨 X-ray 이미지에 대한 응용은 의료 영상 분석에서 매우 유망한 접근법을 제시합니다. 이 연구는 특히 골다공증 진단과 같은 특정 의료 진단에서 더 나은 모델 성능과 일반화 능력을 추구하며 모델간 성능분석과 비교를 통한 유의미한 어깨 골다공증 판별 방법론을 제시하고자 한다.

향후 연구로는X-ray 이미지 뿐만 아니라 환자의 임상 데이터를 포함하는 다중 모델 접근 방식을 개발하여 진단 정확성을 추가로 향상시킬 수 있을 것으로 기대한다.

참고문헌

[1] Masked Autoencoders Are Scalable Vision Learners Kaiming He∗,† Xinlei Chen∗ Saining Xie Yanghao Li Piotr Dollar Ross Girshick, 2021

[2] Delving into Masked Autoencoders for Multi-Label Thorax Disease Classification Junfei Xiao Yutong Bai Alan Yuille Zongwei Zhou\*, 2022

[3] Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization Ramprasaath R. Selvaraju · Michael Cogswell · Abhishek Das · Ramakrishna Vedantam · Devi Parikh · Dhruv Batra