

흉부 엑스레이 기반 폐렴 이진 분류 및 시각화

FastAI & PyTorch 기반 모델 비교와 Streamlit 시각화

발표자: 강충원

2025년 8월 8일

프로젝트 개요

프로젝트 목표

흉부 엑스레이 이미지를 활용한 폐렴 이진 분류 모델 개발 및 시각화 도구 구현을 통해 의료 진단 보조 시스템 구축

주요 기술 스택



FastAI & PyTorch

ResNet18/34, EfficientNetB0 모델 비교 실험



ResNet 아키텍처

딥러닝 모델 학습 및 전이학습 구현



Grad-CAM

모델의 의사결정 과정 시각화 및 해석



Streamlit

웹 기반 인터랙티브 시각화 애플리케이션 개발



정상 폐(좌)와 세균성 폐렴(중), 바이러스성 폐렴(우) 흉부 엑스레이 비교

데이터셋 설명

Kaggle 흉부 엑스레이 데이터셋

Kermay et al.에 의해 수집된 소아 환자 흉부 엑스레이 이미지 데이터셋

출처: Chest X-Ray Images (Pneumonia), Kaggle

데이터셋 구성

항목	내용
수집 기관	Guangzhou Women and Children's Medical Center (중국 광저우)
환자 대상	1세 ~ 5세의 소아 환자
전체 이미지 수	5,863장
클래스 수	2개 (NORMAL, PNEUMONIA)
train 폴더 구성	- NORMAL: 1,341장 - PNEUMONIA: 3,875장
test 폴더 구성	- NORMAL: 234장 - PNEUMONIA: 390장
라벨링 기준	전문가 2인의 판독 + 평가셋은 제3 전문가 확인

데이터 전처리

이미지 크기 조정: 224 × 224 픽셀

train 데이터를 기준으로 8:2로 나누어 validation 세트를 구성



Kaggle 흉부 엑스레이 데이터셋 샘플 이미지

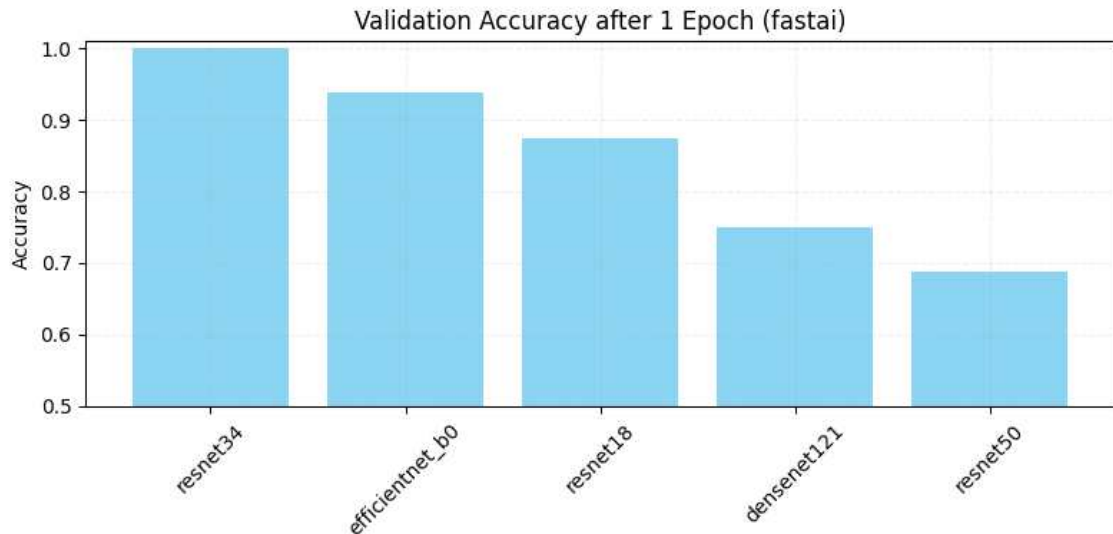
FastAI 기반 모델 비교

모델 비교 실험

흉부 엑스레이 이미지 분류를 위한 최적의 모델 선정을 위해 다양한 사전학습 모델 아키텍처 비교 실험 진행

ResNet34: 최종 모델 선택 이유

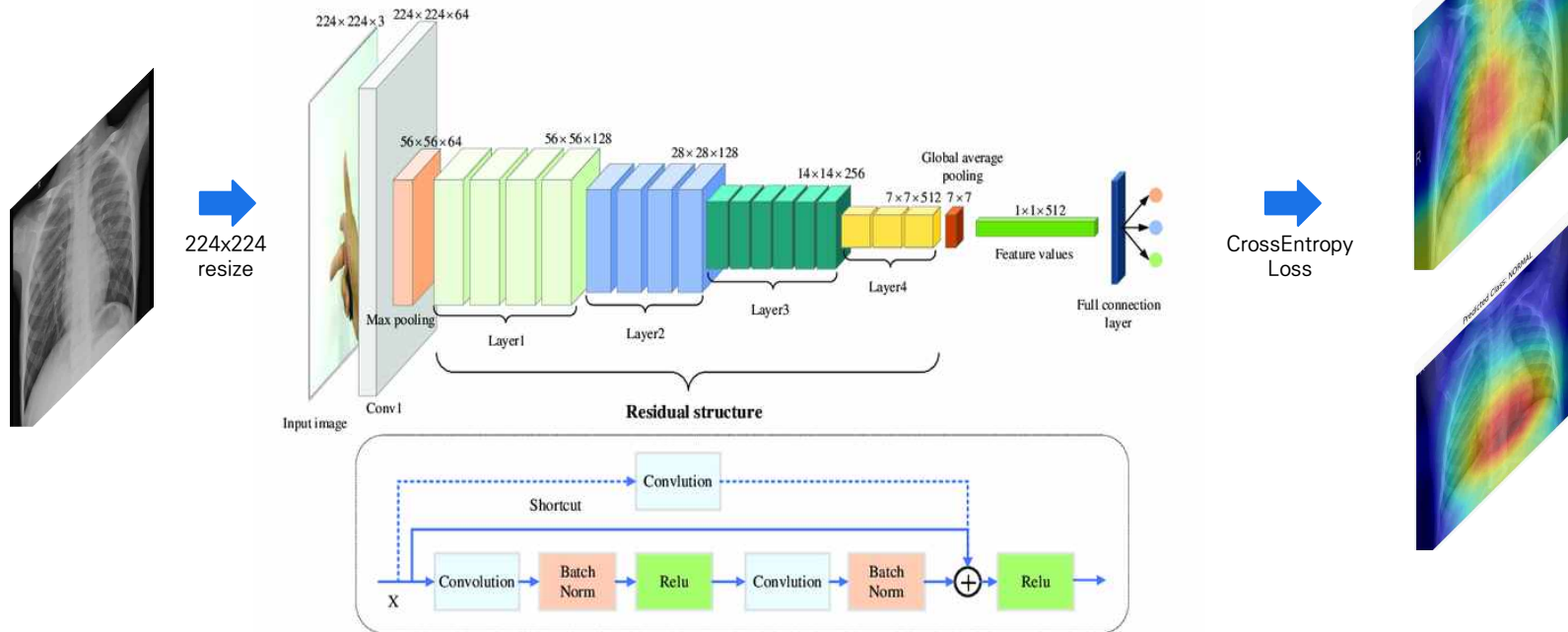
- 1 epoch 기준 가장 높은 Validation Accuracy
- 학습 안정성과 효율적인 수렴
- 과적합 위험은 낮고 표현력은 충분
- CNN 기반 Grad-CAM 해석에도 적합
- 다른 모델 대비 초기 성능 편차가 명확히 존재



최종 모델 구조 설명

ResNet34 아키텍처

34개 레이어로 구성된 잔차 신경망(Residual Neural Network) 구조



성능 평가

테스트 세트 성능 평가

ResNet34 모델의 테스트 세트에서의 성능 지표 및 ROC 곡선 분석

87.7%

정확도 (Accuracy)

0.93

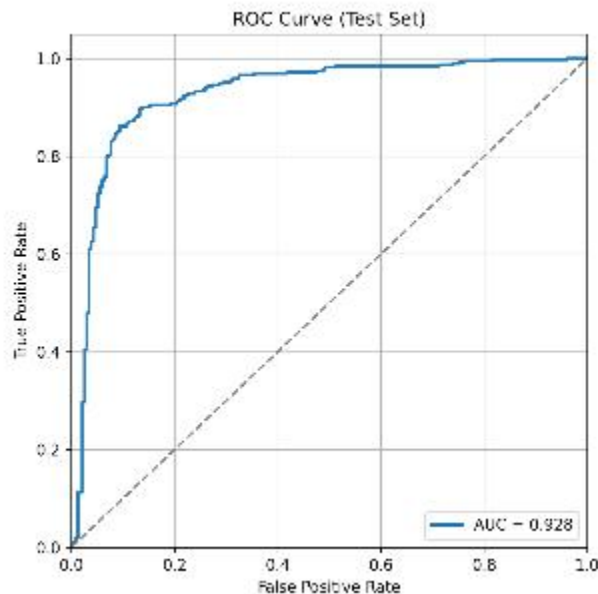
AUC

0.89

F1 점수

성능 해석

- 높은 AUC 값(0.93)은 모델이 정상과 폐렴 사례를 효과적으로 구분할 수 있음을 나타냅니다.
- 특히 의료 진단 보조 시스템에서는 높은 민감도(재현율)가 중요하며, 본 모델은 폐렴 사례의 87.7%를 올바르게 식별했습니다.



Confusion Matrix 분석

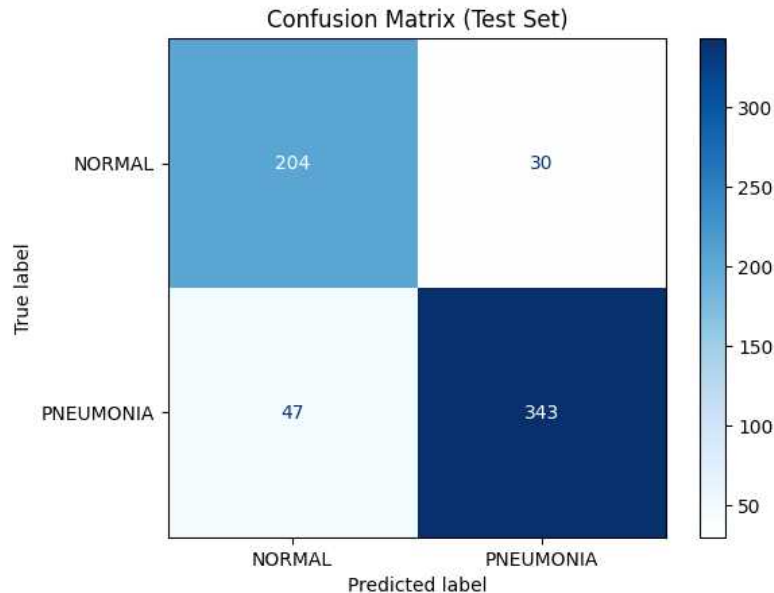
오분류 사례 분석

테스트 세트에서의 모델 예측 결과 분석 및 의학적 관점에서의 위험도 평가

Confusion Matrix 해석

- 본 모델은 폐렴 예측에서 높은 정확도를 보이며, 특히 폐렴 사례(TP)에 대해 강한 분류 성능을 나타냅니다.
- 하지만 실제 폐렴을 정상으로 오인하는 일부 사례(FN)가 존재하므로, 이는 향후 모델 보완 방향의 핵심 포인트입니다.

Confusion Matrix

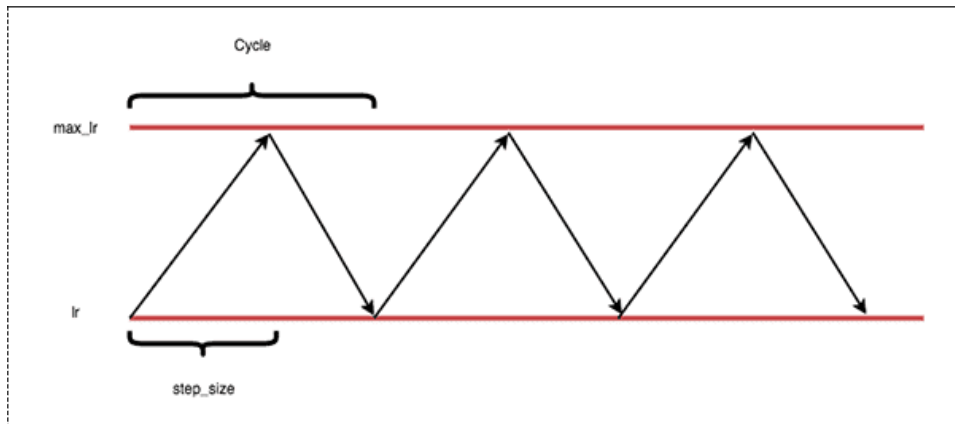


CyclicLR이란?

CyclicLR

LearningRate를 주기적으로 증가-감소시키면서 학습 효과를 극대화시키는 기법

```
scheduler = CyclicLR(
    optimizer,
    base_lr=1e-5,      # 최저 학습률
    max_lr=1e-3,       # 최고 학습률
    step_size_up=len(train_loader) * 2, # 얼마나 빨리 증가시킬지
    mode="triangular",  # 선형 증가-감소 (삼각형 파형)
    cycle_momentum=False # Adam은 모멘텀 기반이 아니라 반드시 False
)
```



- 기본 개념:
러닝레이트를 $\text{base_lr} \rightarrow \text{max_lr}$ 까지 증가했다가 다시 base_lr 로 감소하는 사이클을 반복
 - 목적:
단조롭게 감소시키는 기존 방식과 달리,
일정 주기로 증가/감소하며 local minima에서 벗어날 기회 제공
- 일반적으로 일정 수준의 일반화 성능 향상을 기대할 수 있음

Grad-CAM 시각화 예시

Gradient-weighted Class Activation Mapping

CNN 모델의 의사결정 과정을 시각화하여 모델이 이미지의 어떤 부분에 주목하는지 확인하는 기법

Grad-CAM 작동 원리

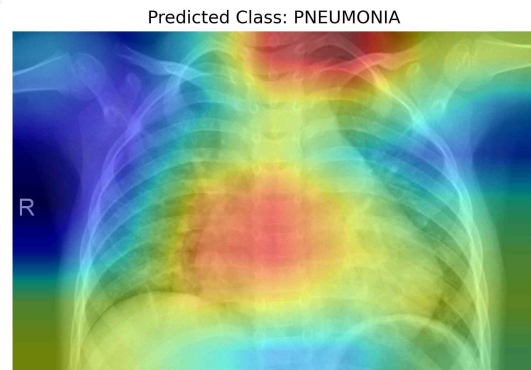
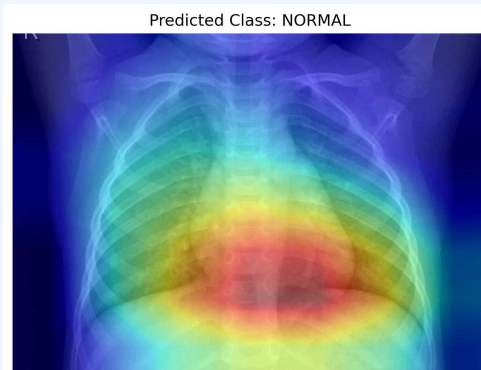
타겟 클래스(정상 or 폐렴)에 대한 그래디언트 계산

마지막 합성곱 레이어(layer4)의 activation map에 그래디언트를 가중 평균하여 클래스 활성화 맵 생성

ReLU 함수를 통해 양의 영향만 강조

히트맵으로 시각화하여, 모델이 예측 시 주목한 폐 영역을 확인

정상 vs 폐렴 Grad-CAM 비교



Grad-CAM의 한계 및 해석 주의점

Grad-CAM 해석의 한계점

시각화 기법의 한계를 이해하고 의료 진단에 책임 있게 활용하기 위한 주의사항

주요 한계점

비의학적 영역에 주목

→ 'R' 마커, 갈비뼈 외곽 등 병변과 무관한 위치에 집중

병변 중심 시각화 부족

→ 폐렴 이미지에서도 폐 내부 병변이 아닌 외곽 주목 현상 발생

Grad-CAM 해석의 모델 의존성

→ 같은 이미지라도 학습된 모델 상태에 따라 시각화 결과가 달라짐

개선 방향

이미지 전처리 강화

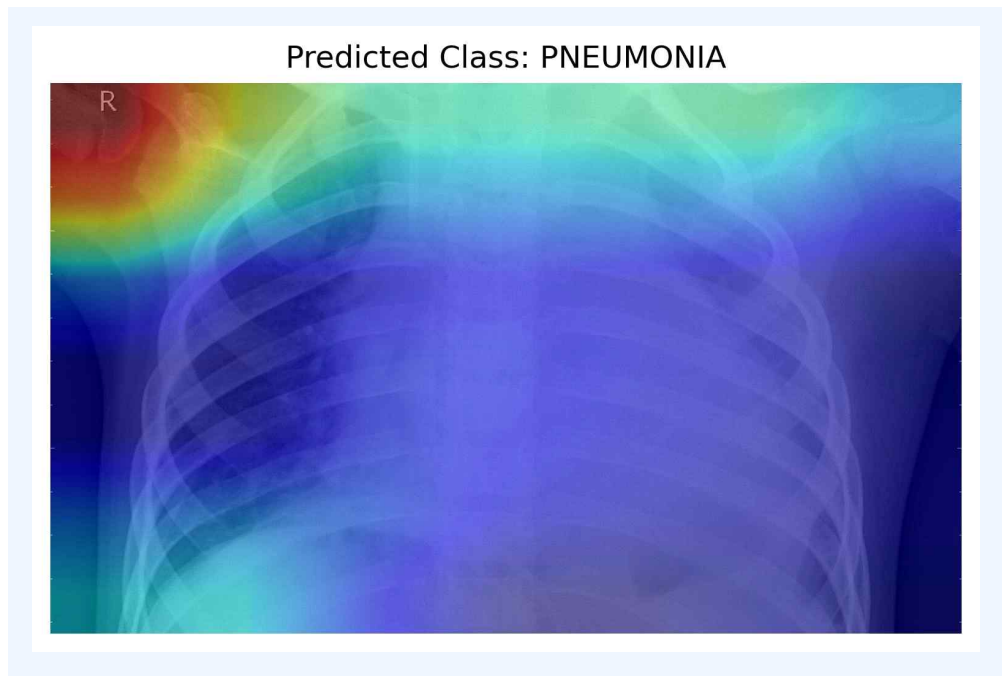
→ 'R', 'L' 마커 등 불필요한 요소 제거로 시각적 bias 최소화

의료 전문가 피드백 기반 검증

→ 실제 폐 병변 위치와 비교하여 Grad-CAM 결과 검토

시각화 품질 평가 지표 도입 고려

→ CAM 결과를 bounding box 또는 세그멘테이션 기반 병변 위치와 비교 분석

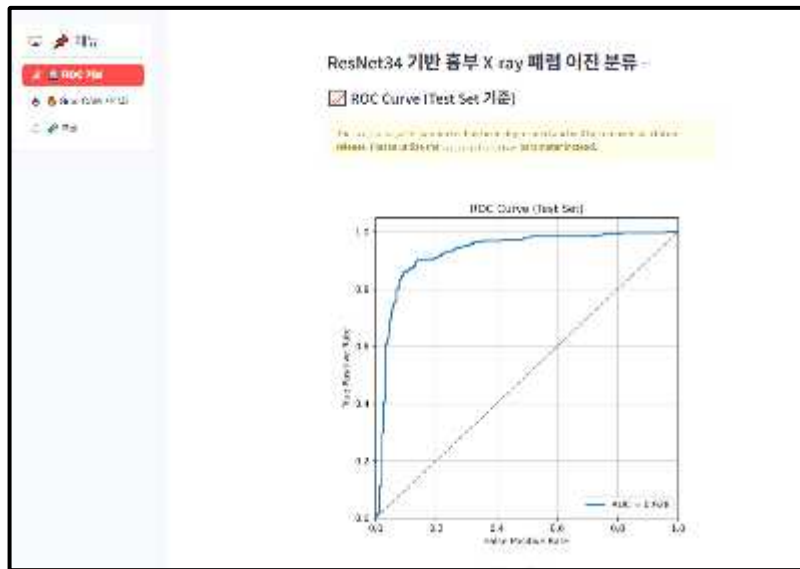


Streamlit 앱 구성 흐름

웹 기반 폐렴 진단 보조 애플리케이션

Streamlit 프레임워크를 활용한 사용자 친화적 인터페이스와 실시간 예측 시각화 구현

ROC 커브 시각화



GradCAM 시각화



Streamlit 앱 구성 흐름

웹 기반 폐렴 진단 보조 애플리케이션

Streamlit 프레임워크를 활용한 사용자 친화적 인터페이스와 실시간 예측 시각화 구현

테스트 추론 페이지



파일 업로드 및 추론



결론 및 느낀 점

프로젝트 결론

전이학습과 해석 가능성을 통한 의료 AI 시스템의 신뢰성 및 투명성 향상

느낀 점

본 프로젝트를 통해 FastAI 기반의 여러 사전학습 모델을 비교 실험하고, 복잡한 모델이 항상 최선은 아님을 체감할 수 있었습니다.

단순한 구조의 ResNet34가 가장 높은 검증 정확도를 기록하며 효율적인 선택임을 입증했고, CyclicLR 스케줄러를 직접 적용해 학습률 변화가 성능에 미치는 영향을 실험적으로 경험할 수 있었습니다.

또한 Grad-CAM의 작동 원리를 이해하고, 이를 통해 모델의 의사결정 과정을 시각적으로 확인할 수 있었습니다.

아쉬운 점

다만 아쉬운 점으로는, JPEG가 아닌 의료용 포맷(DICOM)을 직접 처리해보는 경험이나, 하이퍼컬럼, TTA(Test-Time Augmentation) 등의 기법을 데이터셋 특성상 적용하지 못했던 점이 남습니다.

참고 자료 및 코드 링크

학술 참고 자료

프로젝트에 활용된 주요 논문 및 학술 자료

Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning

Kermary, D. S., Goldbaum, M., Cai, W., et al.

Cell, 172(5), 1122-1131.e9, 2018

Deep Residual Learning for Image Recognition

He, K., Zhang, X., Ren, S., & Sun, J.

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D.

International Journal of Computer Vision, 128(2), 336-359, 2020

Test-Time Augmentation with Uncertainty Estimation for Deep Learning-Based Medical Image Segmentation

Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., & Vercauteren, T.

IEEE Transactions on Medical Imaging, 38(9), 2165-2175, 2019

데이터셋



Chest X-Ray Images (Pneumonia)

Kaggle Dataset



NIH Chest X-ray Dataset

National Institutes of Health

코드 및 리소스



프로젝트 GitHub 저장소

github.com/merware4969/chest_project.git



FastAI 문서

docs.fast.ai



PyTorch 튜토리얼

pytorch.org/tutorials



Streamlit 데모 앱

streamlit_app.py