

미니프로젝트 발표

Chest X-ray Image recognition & Pneumonia prediction model

2024. 06. 09(일)

김성학 남기범

목 차

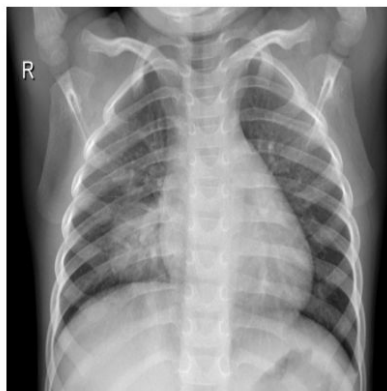
1. 프로젝트 개요
2. 예측모델 평가
3. 서비스 구성
4. 세부 구현 방법

1. 프로젝트 개요

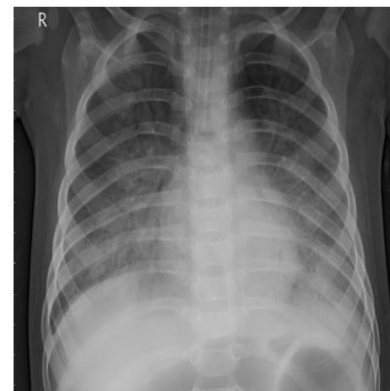
Pneumonia Detection



Normal



Bacterial Pneumonia



Viral Pneumonia

Model accuracy: 79.64%

1. 프로젝트 개요

분석 대상

- Kaggle에 공개된 흉부 X-ray 이미지 (~5,000개)
- Bacterial & viral pneumonia 통합, 정상 772개 Pneumonia 4281개

Pain Point

- 이미지 dataset 전처리 과정
- CNN parameter 수를 늘리는 것만으론 성능 개선 미약

분석 목표

- X-ray 이미지를 input으로 정상/Pneumonia 판별

프로세스

1. 데이터 전처리
2. CNN model 구축 및 학습
3. 모델 평가, 웹 서비스 서빙

2. 프로젝트 배경

- 흉부 X-ray 검사는 호흡기 및 폐질환 진단을 위한 중요 기법
- 일반적으로 내과 의사의 육안 평가로 질환 여부 판별
- 진단 정확도는 60-80% 내외 (CT 포함), 분석 시간 및 정확도 개선 요구됨

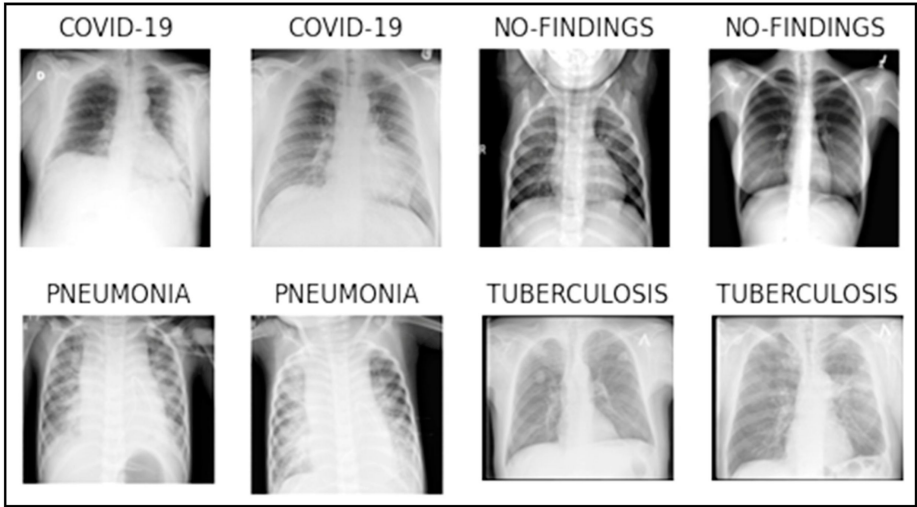
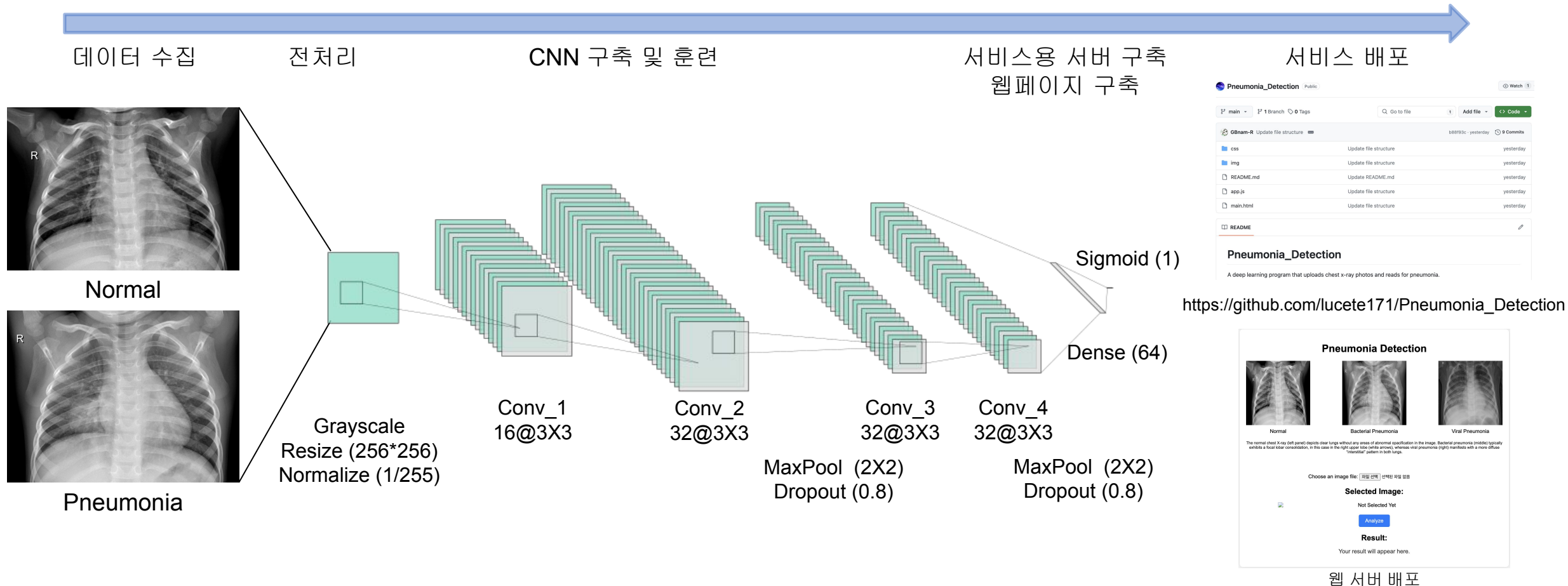


TABLE 3: Scenarios Based on Assignment of Uncertain Chest Radiography and CT Impressions

Performance Value	Value (95% CI)				
	Scenario 1 ^a	Scenario 2 ^a	Scenario 3 ^a	Scenario 4 ^a	Scenario 5 ^b
Sensitivity (%)	51.9 (48.9–54.9)	59.6 (55.6–63.7)	13.4 (11.4–15.4)	19.6 (16.4–22.0)	32.7 (27.7–37.8)
Specificity (%)	71.3 (68.9–73.7)	67.1 (65–69.2)	97.7 (96.9–98.5)	96.4 (95.6–97.3)	96.8 (95.7–97.9)
PPV ^c					
At prevalence of 54.7%	68.4 (66.2–70.6)	68.6 (66.6–70.7)	87.6 (83.4–91.7)	86.8 (83.5–90.1)	92.5 (89.9–95.1)
At prevalence of 44.9%	59.4 (56.9–61.8)	59.6 (57.3–61.9)	82.6 (77.1–88.1)	81.6 (77.3–85.9)	89.2 (85.7–92.9)
At prevalence of 23.2%	35.1 (32.8–37.5)	35.4 (33.2–37.5)	63.8 (54.9–72.6)	62.2 (55.4–69.0)	75.5 (68.6–82.5)
NPV ^c					
At prevalence of 54.7%	55.0 (53.3–56.8)	57.9 (55.3–60.5)	48.3 (47.7–48.9)	49.8 (48.8–50.9)	54.4 (52.5–56.2)
At prevalence of 44.9%	64.5 (62.9–66.1)	67.1 (64.8–69.4)	58.1 (57.5–58.7)	59.5 (58.5–60.5)	63.8 (62.1–65.6)
At prevalence of 23.2%	83.0 (82.0–84.0)	84.6 (83.2–86.0)	78.9 (78.5–79.3)	79.9 (79.2–80.5)	82.6 (81.6–83.7)
Positive LR	1.81 (1.63–2.00)	1.81 (1.65–1.99)	5.74 (3.92–8.38)	5.51 (4.12–7.36)	10.3 (7.07–15.1)
Negative LR	0.675 (0.629–0.724)	0.601 (0.541–0.668)	0.887 (0.865–0.909)	0.833 (0.799–0.869)	0.695 (0.644–0.749)
OR	2.68 (2.27–3.17)	3.01 (2.48–3.66)	6.47 (4.35–9.62)	6.61 (4.79–9.12)	14.9 (9.73–22.7)
Accuracy (%)	62.6 (60.6–64.5)	65.4 (63.4–67.3)	59.8 (57.8–61.8)	78.5 (76.9–80.2)	80.4 (78.2–82.6)

3. 추진 전략

- Kaggle dataset에서 chest X-ray 이미지 확보
- Bacterial/viral pneumonia로 세분화 되어있는 기존 label을 Normal/Pneumonia로 단순화
- 전이 학습 없이 CNN 모델 구축 및 훈련



4. 개발 환경

- 운영 체제: Linux (Google Colab)
- 개발 언어: Python, node.js
- 개발 도구: Google Colab
- 라이브러리: tensorflow, numpy, matplotlib, math, os, cv2, sklearn

The screenshot shows the GitHub repository page for 'Pneumonia_Detection'. The repository is public and has 1 watch. The main branch is selected, with 1 branch and 0 tags. A search bar and 'Add file' button are visible. The commit history shows a recent commit by GBnam-R titled 'Update file structure' with 9 commits. The file list includes css, img, README.md, app.js, and main.html, all updated yesterday. The README section is partially visible, showing the title 'Pneumonia_Detection' and a description: 'A deep learning program that uploads chest x-ray photos and reads for pneumonia.'

File	Commit Message	Commit Date
css	Update file structure	yesterday
img	Update file structure	yesterday
README.md	Update README.md	yesterday
app.js	Update file structure	yesterday
main.html	Update file structure	yesterday

Pneumonia_Detection

A deep learning program that uploads chest x-ray photos and reads for pneumonia.

2. 예측모델 평가

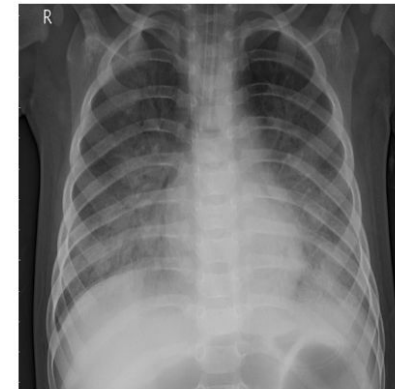
Pneumonia Detection



Normal



Bacterial Pneumonia



Viral Pneumonia

Model accuracy: 79.64%

1. 모델 평가

- 최적 모델 accuracy: 79.64%
- 이미지 장당 소요 시간: ~300 ms (parameter 수: 7,644,033 개)
- CNN parameter 수를 늘리는 것만으론 정확도 개선 나타나지 않음
- 이미지 sharpening 처리 후에도 개선되지 않음



```
1 test_loss, test_accuracy = loadedModel.evaluate(test_images, test_labels)
2 print("테스트 데이터 정확도:", test_accuracy)
```



```
20/20 [=====] - 1s 20ms/step - loss: 0.7128 - accuracy: 0.7965
테스트 데이터 정확도: 0.7964743375778198
```

3. 서비스 구성

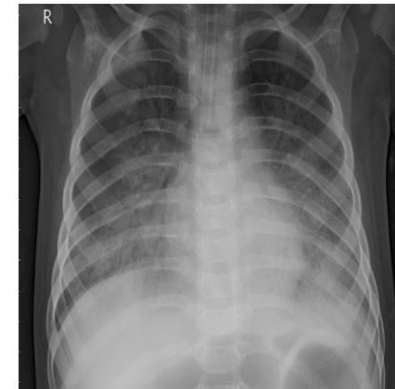
Pneumonia Detection



Normal



Bacterial Pneumonia

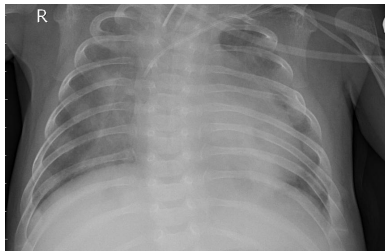


Viral Pneumonia

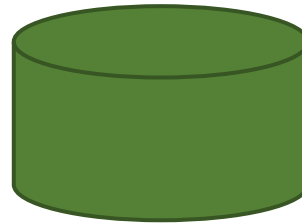
Model accuracy: 79.64%

1. 서비스 구성

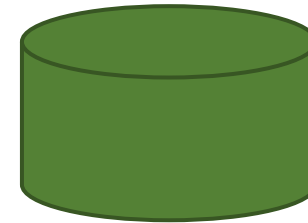
- CNN 예측모델은 Google Colab로 연결 (ngrok)
- 이미지 파일을 **string format**으로 분석서버로 연결
- 예측치 > 0.7 이상인 경우 **pneumonia**로 판정



New X-ray scan



웹서버



분석서버

4. 의의 및 기대효과

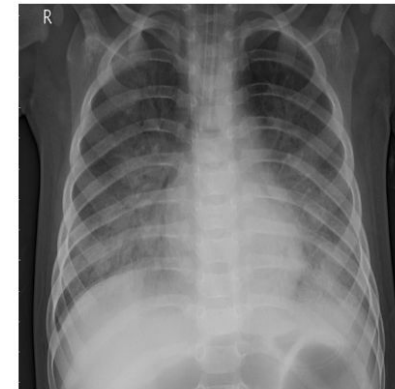
Pneumonia Detection



Normal



Bacterial Pneumonia



Viral Pneumonia

Model accuracy: 79.64%

1. 의의 및 기대효과

1. Dataset 내에서 임상 수준의 정확도 확보
2. 이미지 1장당 연산 시간 < 0.5초, 신속한 진단 보조 가능
3. AI 예측모델 구축 및 서비스 배포 구현
4. 추가 개선 방향성: Data augmentation, filter 도입으로 dataset 개선

감사합니다

End of Document