

# 스터디\_8주차

## ▼ Automated Detection of COVID-19 in Chest Radiographs: Leveraging Machine Learning Approaches(2024)

### 1. 서론 (Introduction)

COVID-19 팬데믹을 유발한 코로나바이러스는 전 세계적으로 빠르게 확산되었으며, WHO는 2020년 3월 11일 이를 글로벌 보건 비상사태로 선포했다. 이러한 조치는 COVID-19의 급격한 확산 속도와 심각성을 강조하였으며, 코로나바이러스가 원인이 된 첫 번째 팬데믹으로 기록되었다. 전 세계 각국은 팬데믹의 영향을 줄이기 위해 사회적 거리 두기, 비행 제한, 국경 폐쇄, 공공 인식 캠페인 등의 다양한 정책을 시행했다.

WHO는 Cobas SARS-CoV-2 검사(Cobas 6800/8800 시스템과 호환 가능)와 Genesis Real-Time PCR 검사를 포함한 몇 가지 테스트를 권장했다. 하지만 이러한 테스트는 많은 시간과 비용이 소요된다. 이에 따라 CNN 기술을 활용하여 감염자 자동 감지 시스템을 개발함으로써 시간과 비용을 절약하고, 기존 검사 방법의 정확도를 보완하는 것이 가능하다.

CNN을 COVID-19 감지에 실질적으로 적용하는 데에는 아직 해결해야 할 과제가 많다. 특히, 의료 환경에서 CNN 도구를 원활하게 통합하기 위한 공통 프로토콜이 부족하며, AI 시스템과 인간 전문가 간 협력적 의사결정이 필요한 상황이다. 본 연구에서는 다중 분류(multiclassification) 모델을 개발하여 정상, COVID-19, 바이러스성 폐렴, 폐 혼탁의 네 가지 범주로 X-선 이미지를 분류하는 방법을 제안한다.

### 2. 선행연구 검토 (Literature Review)

COVID-19 환자의 의료 영상을 분석하는 연구는 활발히 진행되고 있으며, 흉부 X-선과 CT 영상을 활용하는 연구가 대표적이다.

- 한 연구에서는 DenseNet201, ResNet50V2, InceptionV3 모델을 사용하여 흉부 X-선 영상을 학습하고, 이를 기반으로 COVID-19 감염 여부를 예측하였다 【6】. 연구진은 538장의 COVID-19 X-선 이미지와 468장의 정상 X-선 이미지를 학습 및 검증에 사용하였다.
- 또 다른 연구에서는 COVID-19 환자의 194장과 정상인의 194장을 포함한 X-선 데이터셋을 활용하였다 【7】. 연구진은 전이 학습(Transfer Learning)을 활용하여 데이

터 부족 문제를 해결하였으며, SVM, KNN, 랜덤 포레스트(Random Forest), 다층 퍼셉트론(MLP)과 같은 다양한 기법을 적용하였다.

- MobileNet을 활용하여 COVID-19, 결핵, 바이러스성 폐렴, 세균성 폐렴, 정상 폐 상태를 구분하는 연구도 있었다 【8】. MobileNet 기반 모델은 98.80%의 높은 정확도를 기록하였다.
- Bassi와 Attux 【9】는 CNN 모델을 활용하여 150개의 X-선 이미지를 COVID-19, 폐렴, 정상 상태로 분류하였으며, 100%의 정확도를 기록하였다.
- Tuncer 등 【11】은 ResExLBP(Residual Exemplar Local Binary Pattern) 및 ReliefF 기반의 COVID-19 자동 감지 기법을 개발하였으며, 다양한 분류 모델(SVM, 의사결정 트리 등)을 평가하였다.

이와 같이, 기존 연구들은 대부분 이진 분류(binary classification)에 초점을 맞추었으나, 본 연구는 네 가지 범주로 분류하는 다중 분류 모델을 개발했다는 점에서 차별성이 있다.

---

### 3. 제안된 방법론 (Proposed Methodology)

본 연구의 목표는 흉부 X-선 이미지를 정상(Normal), COVID-19, 바이러스성 폐렴(Viral Pneumonia), 폐 혼탁(Lung Opacity) 네 가지 범주로 정확하게 분류하는 머신 러닝 모델을 개발하는 것이다. 이 장에서는 데이터 수집, 전처리 방법, 모델 선택 논리, 평가 지표 등에 대해 설명한다.

---

#### 3.1 데이터 전처리 (Data Preprocessing)

본 연구에서는 데이터 품질을 향상시키고 모델 학습을 최적화하기 위해 다음과 같은 전처리 과정을 수행하였다.

##### 1. 이미지 크기 표준화

- 모든 이미지를 **150×150 픽셀**로 변환하여 모델이 일관된 입력을 받을 수 있도록 함.

##### 2. 흑백 변환(Grayscale Conversion)

- 흉부 X-선 이미지는 컬러 정보가 불필요하므로 계산량을 줄이기 위해 흑백으로 변환.

##### 3. 데이터 증강(Data Augmentation) 미적용

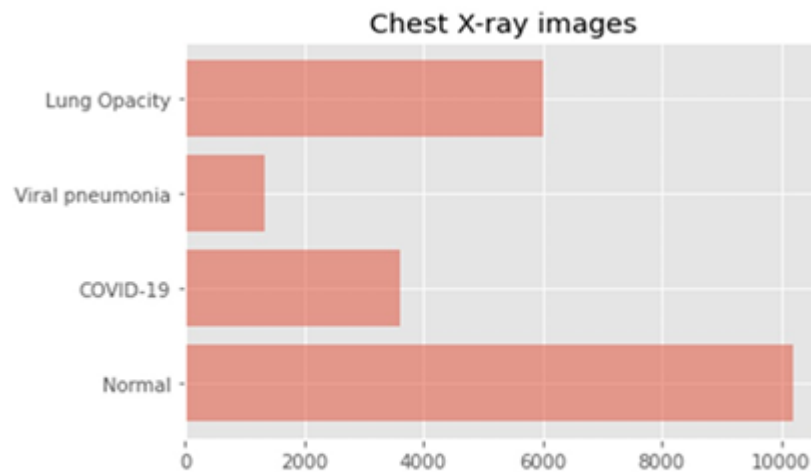
- 일반적으로 이미지 분류 작업에서는 회전, 뒤집기, 확대 등의 증강 기법을 적용하지만, 본 연구에서는 **의료 진단의 정확성을 유지하기 위해 원본 X-선 이미지를 그대로 사용.**

#### 4. 레이블 인코딩(Label Encoding)

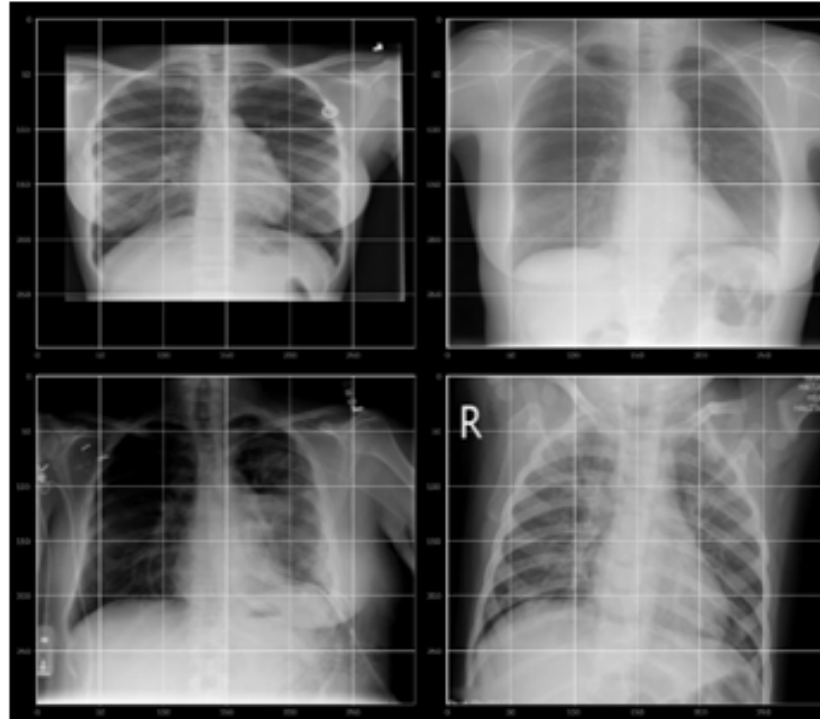
- 각 X-선 이미지를 해당하는 클래스(정상, COVID-19, 바이러스성 폐렴, 폐 혼탁)로 레이블링.
- 이진 벡터로 변환하는 **원-핫 인코딩(One-hot encoding)** 적용.

#### 5. 데이터셋 분할

- 전체 데이터셋의 **80%는 학습(training) 및 검증(validation) 용도로 사용하고,**
- **20%는 테스트(test) 용도로 분리하여** 모델의 일반화 성능을 평가.



**Fig. 1. Chest X-ray Image.**



**Fig. 2. X-ray Classification: ML Methodology.**

### 3.2 모델 구조 (Model Architecture)

본 연구에서는 VGG-16(Visual Geometry Group) 모델을 기반으로 하는 CNN(합성곱 신경망) 모델을 활용하였다.

#### 1. VGG-16 구조

- VGG-16은 옥스퍼드 대학교 Visual Geometry Group에서 개발한 신경망 모델로, **16개의 계층(13개 합성곱 층 + 3개 완전연결 층)**으로 구성되어 있음.
- 작은 **3×3 합성곱 필터**를 사용하여 **미세한 공간 정보를 유지하면서도 깊은 특징을 학습**할 수 있음.

#### 2. 특징

- 각 합성곱 블록은 **여러 개의 합성곱 층과 맥스 풀링(Max-Pooling) 층**으로 구성.
- **3×3 필터(스트라이드 1)**를 사용하여 이미지의 특징을 효과적으로 학습.
- **\*2×2 맥스 풀링 층(스트라이드 2)\*\***를 사용하여 차원을 축소하면서도 중요한 정보 유지.

### 3. 추가 개선

- 기존 VGG-16 모델 외에도 **분리 가능 합성곱(Separable Convolution)** 및 **배치 정규화(Batch Normalization)** 기법을 적용하여 계산 효율성을 향상.
- **드롭아웃(Dropout)** 기법을 적용하여 과적합(Overfitting) 방지.
- **Adam 옵티마이저(Optimizer)**를 사용하여 학습률 조정.
- 다중 분류 문제 해결을 위해 **이진 크로스 엔트로피(Binary Cross-Entropy)** 손실 함수 사용.

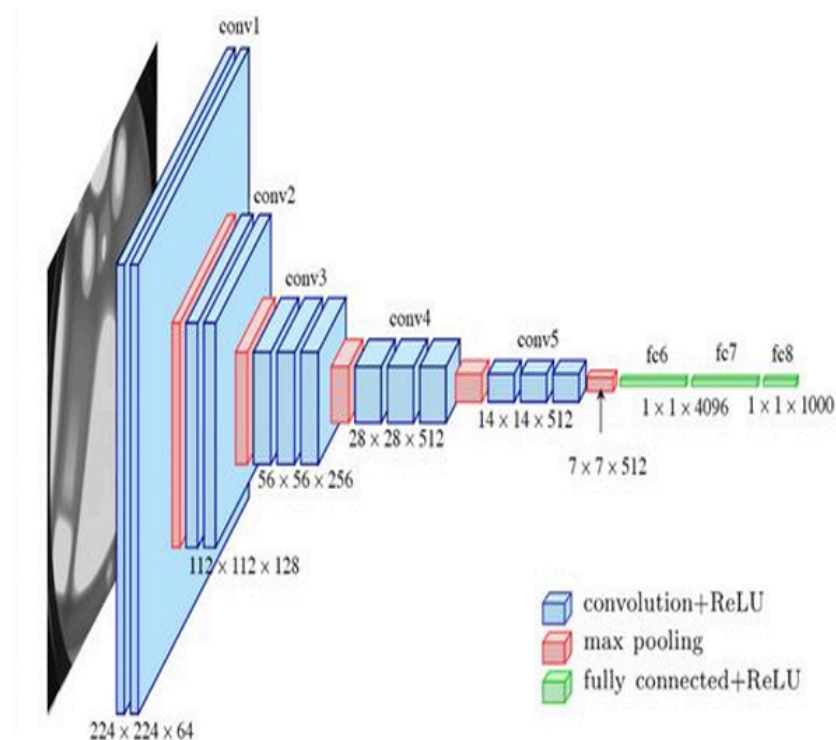


Fig. 3. Model Architecture.

## 4. 실험 결과 (Results)

본 연구에서는 제안된 모델의 성능을 다양한 평가 지표를 사용하여 분석하였다.

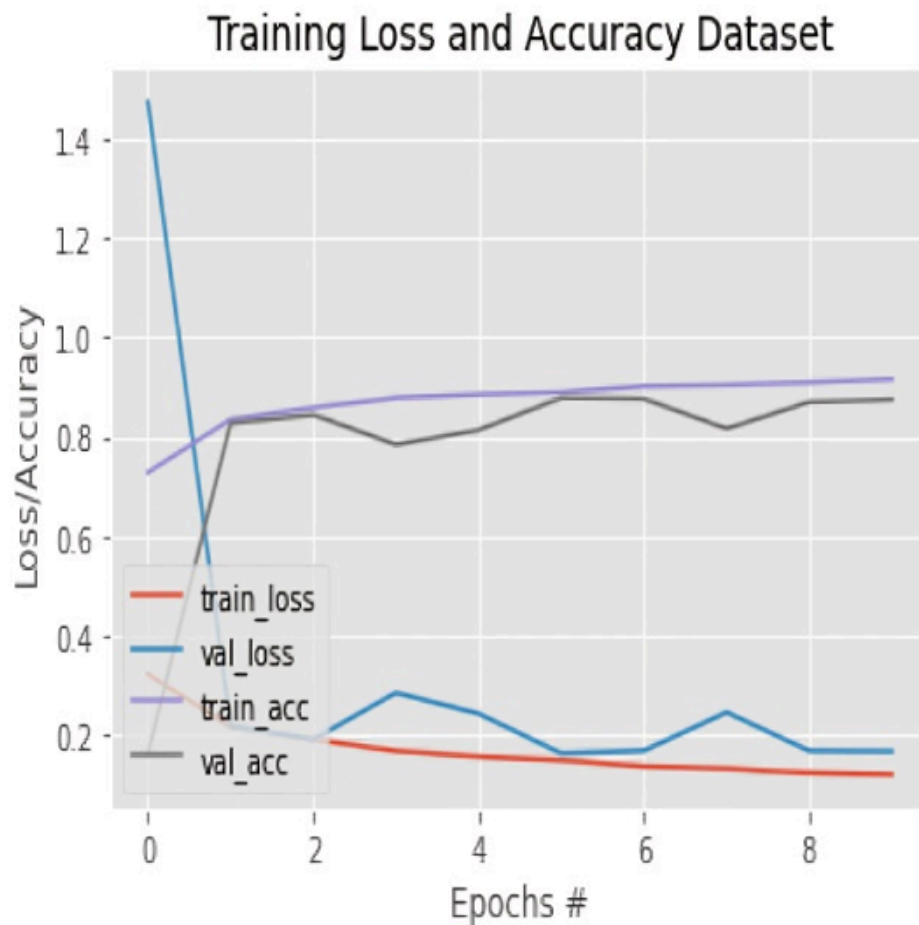
### 4.1 학습 과정 분석 (Epoch-wise Analysis)

- 모델을 총 10번의 에포크(Epoch) 동안 학습하고, 매 에포크마다 성능을 측정하였다.

- 첫 번째 에포크에서 **훈련 정확도(Training Accuracy)**는 **72.76%**, \*검증 정확도 (Validation Accuracy)는 16%로 낮았음.
- 그러나 점차 향상되어 **10번째 에포크에서 훈련 정확도는 91.54%, 검증 정확도는 87.48%**에 도달.

## 4.2 손실과 정확도 분석 (Loss and Accuracy)

- 손실 함수로 **이진 크로스 엔트로피(Binary Cross-Entropy)**를 사용.
- 첫 번째 모델에서 **최종 손실 값은 훈련 데이터에서 0.1184, 검증 데이터에서 0.1653.**
- 두 번째 모델에서는 손실이 **0.0354(훈련 데이터), 0.0457(검증 데이터)**로 더 낮아졌음.



**Fig. 4. Training Loss and Accuracy Dataset.**

---

### 4.3 평균 절대 오차(Mean Absolute Error, MAE)

- 훈련 데이터에서 **MAE는 0.0695**, 검증 데이터에서 **0.0761**을 기록.
- 이는 이상적인 값(0.08~0.09)보다 낮아 모델이 정확한 예측을 수행했음을 나타냄.

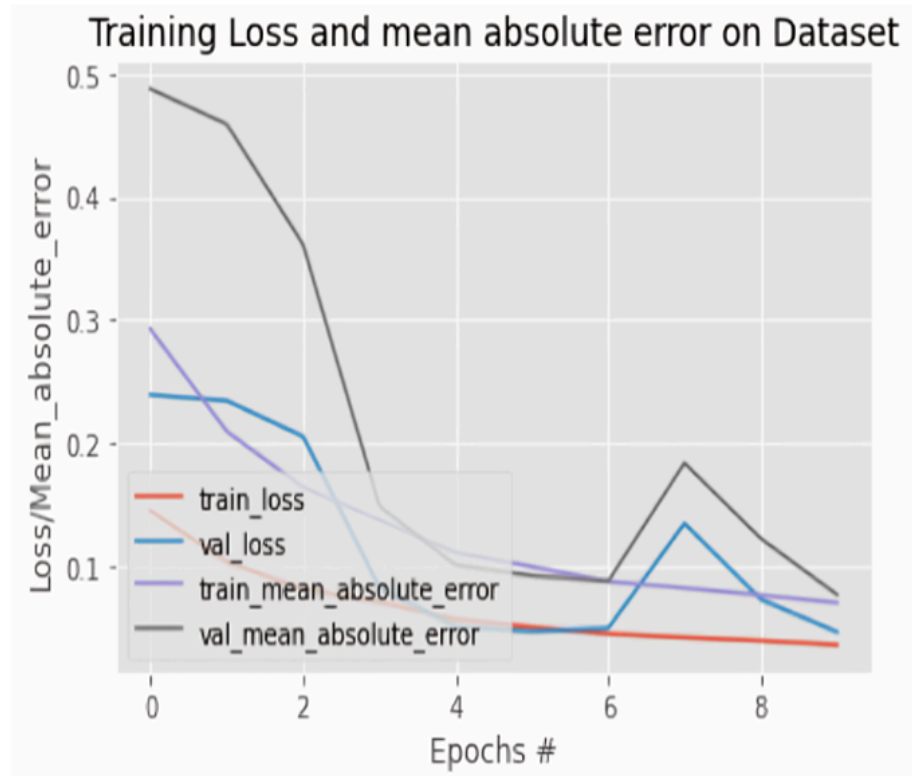


Fig. 5. Training Loss and mean absolute error on Dataset.

---

### 4.4 테스트 데이터 평가 (Test Accuracy and Loss)

- 테스트 데이터에서 **88.14%의 정확도**와 **손실 0.0449**를 기록.
- 검증 데이터와 일관된 결과를 보이며 모델이 새로운 데이터에 대해 일반화 능력을 가짐을 확인.

---

### 4.5 혼동 행렬(Confusion Matrix) 분석

- COVID-19 및 정상(Normal) 데이터에 대해 높은 성능을 보였으며, 전반적으로 높은 정밀도(Precision)와 재현율(Recall)을 달성.
- F1-score(정밀도와 재현율의 조화 평균) 또한 모든 클래스에서 높게 나타남.

**Confusion Metrix**

Covid Detection with VGG-16 Confusion Matrix

1760	46	225	18
16	672	40	0
113	20	1063	0
6	12	6	236

**Fig. 6. Confusion Matrix.**

## 4.6 학습률 조정 (Learning Rate Adjustments)

- 초기에 학습률 0.0001을 사용하였으나, 학습 진행에 따라 점진적으로 감소시켜 0.000106까지 조정.
- 적응형 학습률 덕분에 학습이 안정적으로 진행됨.



Accuracy Score : 0.8814079848806993

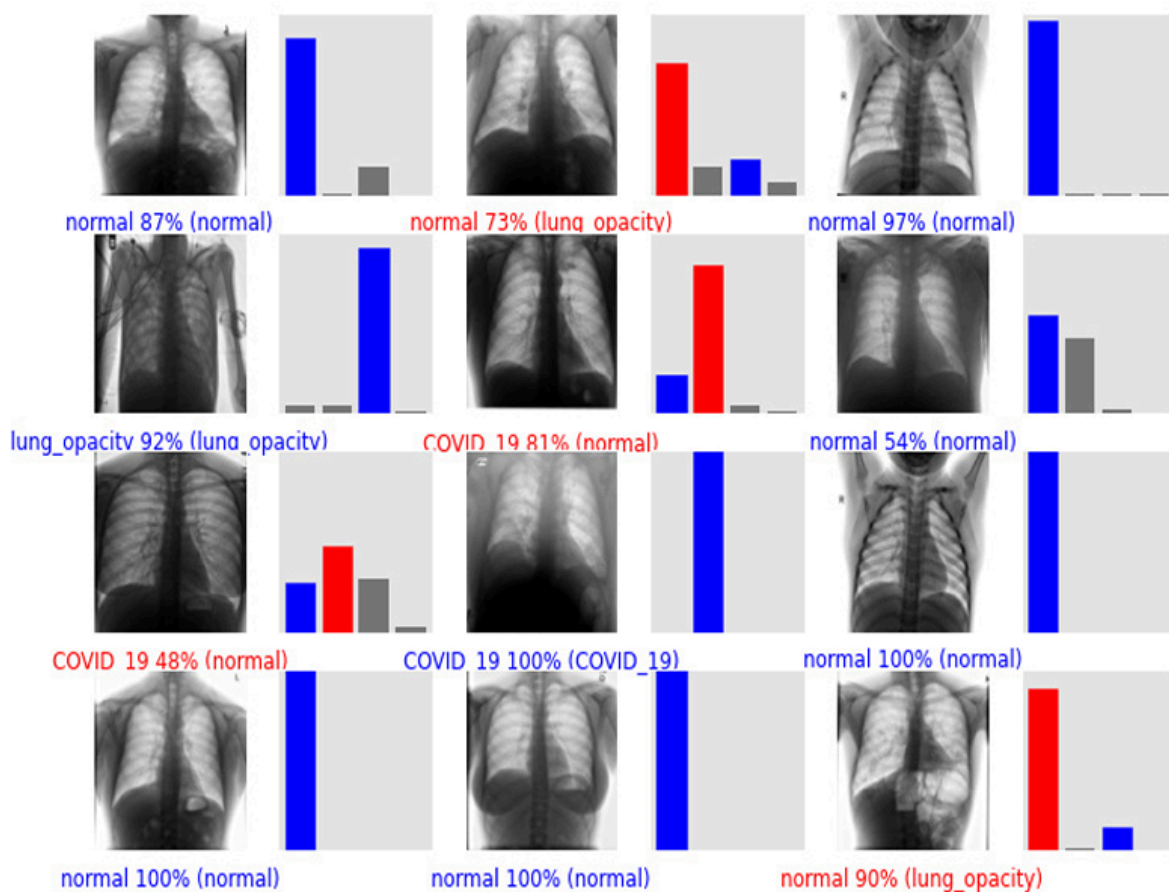
Report :

	percision	recall	f1-score	support
0	0.93	0.86	0.89	2049
1	0.90	0.92	0.91	728
2	0.80	0.89	0.84	1196
3	0.93	0.91	0.92	260
accuracy			0.88	4233
macro svg	0.89	0.89	0.89	4233
weighted avg	0.89	0.88	0.88	4233

---

## 4.7 모델 예측 결과 시각화 (Visual Interpretation of Predictions)

- 모델의 예측 정확성을 직관적으로 확인하기 위해 **색상별 바 그래프** 및 **시각적 이미지 분석**을 수행.
- 파란색 막대는 모델이 정확하게 예측한 경우를, 빨간색 막대는 잘못된 예측을 의미.



**Fig. 8. Model Predictions Visualization.**

## 5. 연구의 의의 및 시사점 (Significance and Implications)

본 연구에서 개발한 모델은 **COVID-19** 및 기타 폐 질환을 신속하고 정확하게 감지할 수 있는 도구로 활용될 수 있다.

특히, 높은 정확도를 유지하면서도 다중 분류 문제를 해결할 수 있는 점에서 실질적인 의료 응용 가능성이 높다.

그러나, 본 모델이 다양한 인구 집단과 의료 환경에서 동일한 성능을 유지하는지에 대한 추가 연구가 필요하다.

## 6. 결론 (Conclusion)

COVID-19 확산이 계속됨에 따라 신속하고 정확한 감지가 필수적이다.

본 연구에서는 CNN 기반의 머신 러닝 모델을 개발하여 흉부 X-선 영상을 분석하였으며, 정상, COVID-19, 바이러스성 폐렴, 폐 혼탁 네 가지 범주로 분류하는 데 성공하였다.

---

## 7. 향후 연구 방향 (Future Directions)

- 흉부 X-선뿐만 아니라 CT 영상과 텍스트 데이터를 추가하여 모델 성능 개선.
- 전이 학습(Transfer Learning)을 활용하여 대규모 데이터셋에서 모델 학습.
- AI 모델의 설명 가능성(Explainability)을 높여 의료진이 신뢰할 수 있는 시스템 구축.
- COVID-19 외에도 폐 질환 전반을 예측할 수 있도록 모델 확장.

▼ CNN 기반의 이종 신경망 융합을 이용한 물체 인식 및 파지점 예측

## 1. 개요 및 연구 배경

본 논문은 로봇 시스템에서 CNN(Convolutional Neural Network)을 이용하여 물체 인식 및 파지점(grasping point) 예측을 수행하는 연구를 다루고 있다. 최근 CNN을 활용한 딥러닝 기법이 발전하면서, 로봇 비전 분야에서 물체 인식 및 특징점 검출 성능이 크게 향상되었다.

그러나, CNN 기반 알고리즘을 임베디드 시스템에 적용하는 데는 한계가 있다. 일반적으로 CNN은 연산량이 많아 고성능 GPU가 필요한데, 임베디드 환경에서는 컴퓨팅 파워가 제한되기 때문이다.

본 연구에서는 이종 신경망 융합(Fusion of Different Neural Networks) 방법을 적용하여 이러한 문제를 해결하고자 한다. Mask R-CNN에서 배경이 제거된 이미지를 Fully Convolutional Network(FCN)에 입력하여 물체를 인식하고, 보다 정확한 파지점을 예측하는 방법을 제안하였다.

이 시스템은 ROS(Robot Operating System) 환경에서 구현되었으며, 서버-클라이언트 방식으로 신경망을 분산 처리함으로써 연산 속도와 정확도를 동시에 향상시켰다.

---

## 2. 로봇 시스템 및 하드웨어 구성

### 2.1 하드웨어 구성

- Jetson Tx2 Development Kit: NVIDIA의 임베디드 보드로, CNN 연산을 담당

- **TurtleBot3 Waffle:** 이동형 로봇 플랫폼
- **LDS-01(LIDAR) Sensor:** SLAM 및 Navigation에 사용
- **OpenCR & DYNAMIXEL XM-430:** 로봇 팔 및 모터 제어
- **Xtion Pro Live Kinect Sensor & Logitech C920 Camera:** 물체 인식용 카메라
- **5-DOF(자유도) Robot Manipulator:** 파지점 예측 후 실제 물체를 잡기 위한 로봇 팔

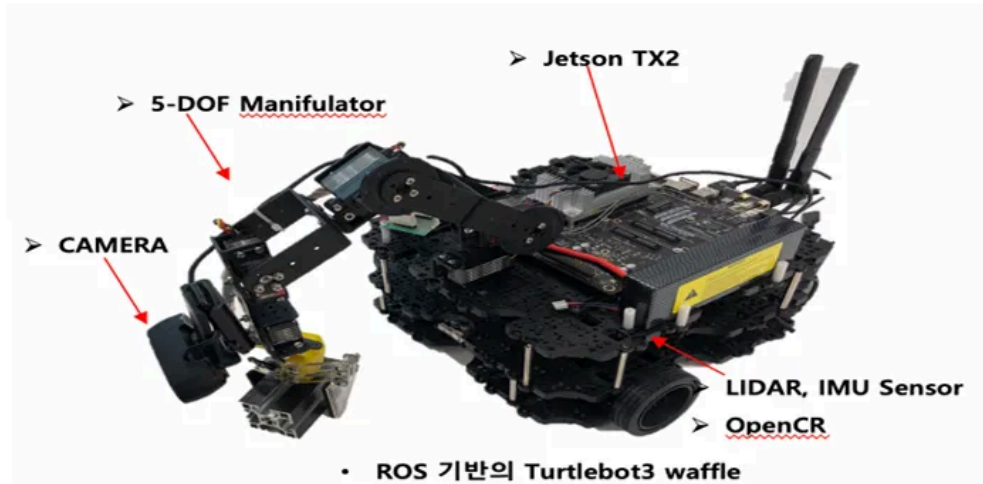


그림 1. ROS 기반의 모바일 로봇

## 2.2 소프트웨어 구성

- 운영체제: Ubuntu 16.04
- ROS 버전: ROS-Kinetic
- 딥러닝 프레임워크: TensorFlow 1.18.0, Keras
- 프로그래밍 언어: Python 2.7

## 3. 물체 파지점 검출 알고리즘

### 3.1 Mask R-CNN을 이용한 Object Detection

Mask R-CNN은 기존 R-CNN → Fast R-CNN → Faster R-CNN의 발전 과정에서 개발된 알고리즘으로, 픽셀 단위의 객체 분할(Segmentation)이 가능하다.

Mask R-CNN을 실시간으로 처리하기 위해 **MobileNet을 Backbone으로 사용**하였다. MobileNet은 VGG Net이나 ResNet과 비교해 연산 속도가 빠르며, 정확도 차이가 크지 않기 때문에 임베디드 시스템에서 효율적으로 사용될 수 있다.

- **학습 데이터:** 한 물체당 약 500장의 다양한 배경에서 촬영한 이미지 사용
- **사전 학습된 COCO Dataset을 이용하여 Fine-Tuning**
- **Epoch 수: 20** (짧은 학습으로도 우수한 성능 확보)
- **ROS 환경에서 실행 시 17fps의 속도로 처리** (기존 1~2fps 대비 약 10배 속도 향상)

Mask R-CNN을 활용하여 배경을 제거한 후, **다른 신경망(FCN)의 입력 데이터로 활용**하였다.

▼ backbone이란?

### Backbone이란?

- **\*Backbone(백본)\***은 딥러닝에서 기본적인 특징(feature)을 추출하는 주요 신경망 구조를 의미한다. 주로 **CNN 기반의 모델**에서 입력된 이미지를 처리하여 중요한 특징을 추출하는 역할을 한다.

### Backbone의 역할

- **입력 이미지에서 중요한 특징을 추출 (Feature Extraction)**
- **고차원 특징 맵(Feature Map) 생성**
- **추출된 특징을 바탕으로 후속 네트워크(Detection, Segmentation 등)에서 활용** 즉, **Backbone은 CNN의 "Feature Extractor" 역할을 한다.**

## 3.2 Fully Convolutional Network (FCN)

FCN은 기존 CNN에서 **Fully Connected Layer를 제거하고, Up-Sampling을 추가**하여 픽셀 단위의 예측이 가능하도록 만든 네트워크이다. 본 논문에서는 얇은 신경망의 CNN을 사용하여, 실시간성 측면에서 강점이 있는 U-net을 사용하였다.

- **FCN의 단점:** 경계 부분 예측 정확도가 낮음
- **해결 방법:** Mask R-CNN을 활용하여 배경 제거 후 FCN에 입력함으로써 성능 개선
- **Jetson TX2에서 처리** (컴퓨팅 파워가 제한적이므로 경량화된 FCN 모델 사용)

- U-Net의 가장 큰 약점은 경계 부분에 대한 예측 정확도가 떨어지는 점이다. 이 점을 해결하면서 실시간성을 맞추기 위해 U-Net 앞에 Mask R-CNN을 이용하여 이미지를 전처리를 진행하였다.

#### ▼ U-net이란?

- U커브 형태의 Convolutional Layer를 쌓은 구조로, 이미지 Segmentation을 하기 위한 네트워크 구조이다.
- U-Net은 크게 2가지로 나뉘어져있는데, 그림 2-5의 좌측부분에 해당하는 downsampling 과정과 우측부분에 해당하는 upsampling 과정이 있다. Downsampling과정을 Constracting Path 라고도 하는데, 이 path에서의 역할은 이미지의 전체적인 구성을 보는 것이다. 즉, 이미지에서 어떤 물체가 있는지에 대한 정보 등을 얻게된다. Upsampling 과정은 Expanding Path 라고도 하고, 이 path에서의 역할은 물체의 위치를 찾는 것이다.

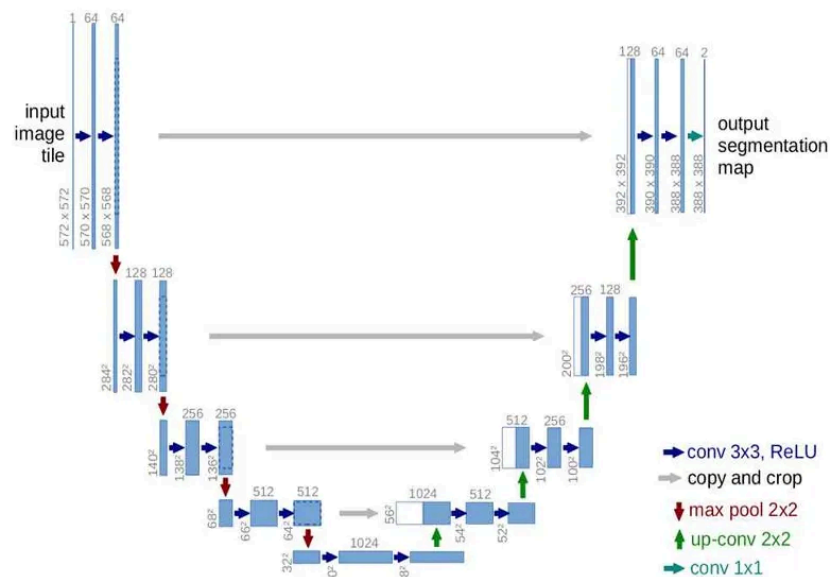




그림 2. 특징점 추출을 위한 Image, Annotation data

### 3.3 ROS를 이용한 시스템 구성

전체 시스템은 ROS에서 Node 단위로 구성되었으며, 주요 노드는 다음과 같다:

1. 이미지 처리 노드: Mask R-CNN 및 FCN을 실행하여 파지점 예측
2. SLAM & Navigation 노드: 로봇의 이동 경로 설정
3. 5-DOF 로봇 암 노드: 로봇 팔 제어
4. 모터 제어 노드: DYNAMIXEL을 이용한 모터 동작 수행

이러한 노드 구조를 통해, 연산 부담을 서버와 클라이언트로 분산하여 처리 속도를 개선하였다.

### 3.3 ROS를 이용한 시스템 구성

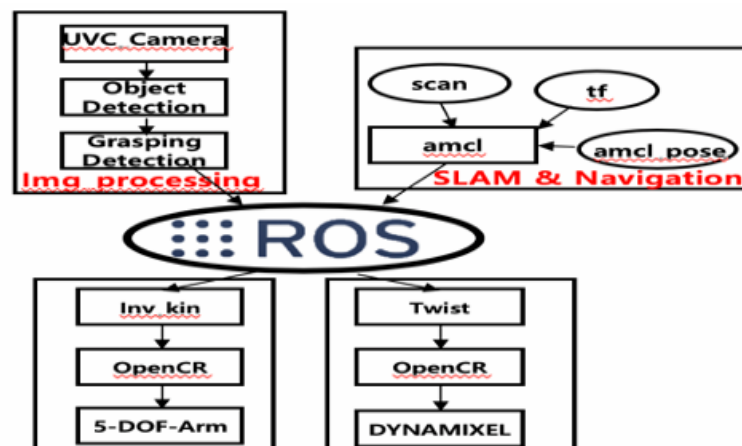


그림 3. 전체 제어 시스템 구성

## 4. 실험 및 성능 분석

### 4.1 실험 결과

- 기존 임베디드 환경에서 Mask R-CNN 단독 실행 시: 1~2fps (성능 부족)
- ROS 서버-클라이언트 방식 적용 후: 17fps (약 10배 속도 개선)
- 단일 신경망 대비 이중 신경망 융합 시 정확도 15% 향상

이러한 결과를 통해, 임베디드 시스템에서도 데스크탑 수준의 파지점 예측 성능을 확보할 수 있음을 확인하였다.

### 4.2 특징점 인식을 이용한 파지점 및 각도 예측

파지점을 단순한 좌표가 아니라 각도까지 포함하여 예측할 수 있도록 구현하였다.

- **Mask R-CNN**: 물체 인식 및 배경 제거
- **FCN**: 최종 파지점 및 각도 예측
- **ROS 기반 로봇 암 제어**: 예측된 파지점을 활용하여 물체를 안정적으로 파지

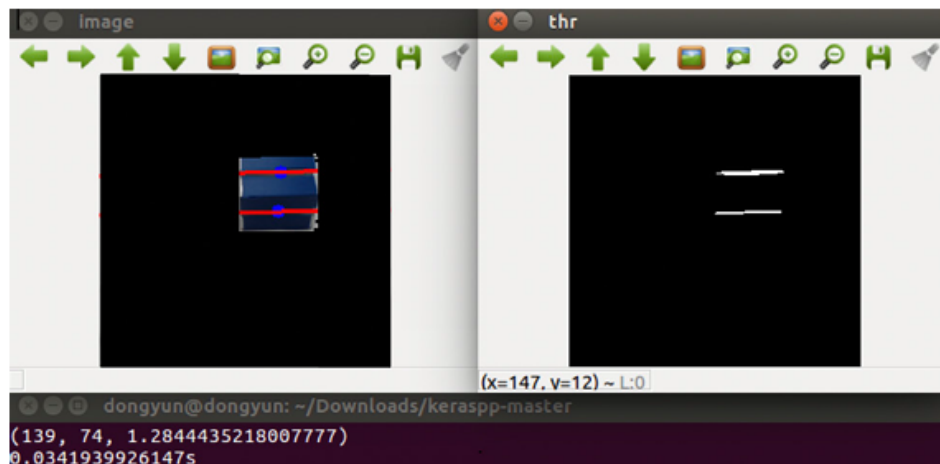


그림 4. 특징점 인식을 이용한 파지점 및 각도 예측

## 5. 결론 및 향후 연구 방향

### 5.1 연구의 주요 성과



- 이중 신경망(Mask R-CNN + FCN) 융합을 통해 정확도 및 속도 개선
- 임베디드 환경에서도 실시간 파지점 예측 가능
- ROS 기반 서버-클라이언트 분산 처리 기법을 적용하여 속도 10배 향상
- 정확도를 15% 향상시키면서도 실시간성 유지

## 5.2 향후 연구 방향

- 보다 다양한 환경(조명 변화, 배경 변화)에서도 성능 개선
- 경량화된 딥러닝 모델을 적용하여 임베디드 시스템 최적화
- 로봇 팔의 파지 동작 최적화를 위한 강화학습 적용