

# 멀티 라벨링 기법을 통한 구조물 결함 분류 모델 개발

PIAI 인공지능연구원

Team: CRACK

홍창호, 박완욱

Thanks to : 김경영, 이수민

# 목차

## 1. 서론

- 주제 선정 배경 및 정의

## 2. 본론

- 모델 선정 및 개발, 결과분석

## 3. 결론

- 결론 및 한계점 제시

01.

## 주제 선정 배경

■ 증가하는 노후건축물에 따른 건물 점검 수요 증가

국내에서는 "시설물의 안전 및 유지관리에 관한 특별법"과 "제3종 시설물 안전 등급 평가 메뉴얼"을 통한 구조물 안전 점검에 대한 체계화된 정책 추진

향후 5년 이내에  
30년 이상 노후 건축물  
49.5%→ 65.8%



건물 안전 점검 수요 증가

향후 10년 이내에  
30년 이상 교량비율  
15%→ 45% 3배증가



구조물 노후화로 인한 유지관리 시스템의 필요



인공지능 기반의  
안전 점검 필요성 대두

건물 안전 점검은 점검자가 직접 육안 검사로 확인하는 관계로 늘어나는 수요를 뒷받침할 재정 및 인력이 필요.

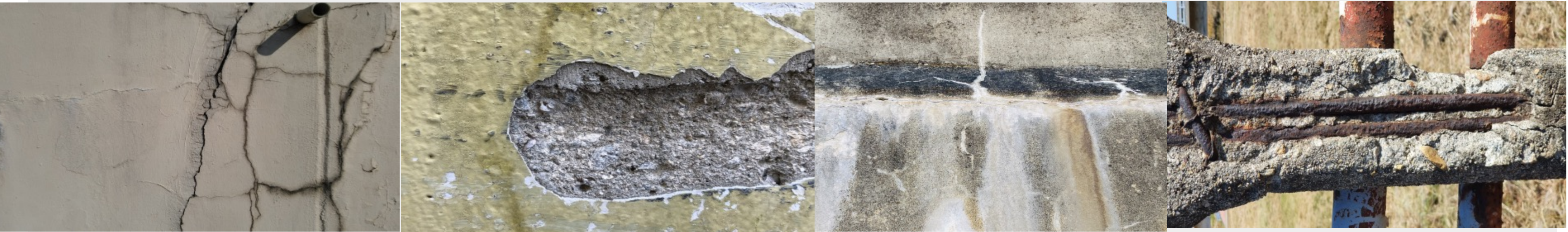
수요에 따른 인력의 공급을 보완하고자 인공지능 기반의 딥러닝 모델을 개발하고자함.



# 구조물 결함의 정의

## 구조물 결함이란

- 안전점검 또는 정밀 안전 진단 결과 발견된 구조상 주요 부분에 시공 상의 잘못으로 인한 중대한 하자.
- 기초, 교량, 터널, 항만, 댐 등의 내력상실 등 시설물의 내구성에 심각한 영향을 미치는 결함.
- 육안검사를 통한 구조물의 안전 점검 항목에는 균열, 박리, 박락, 백태 누수, 철근 노출등이 존재.

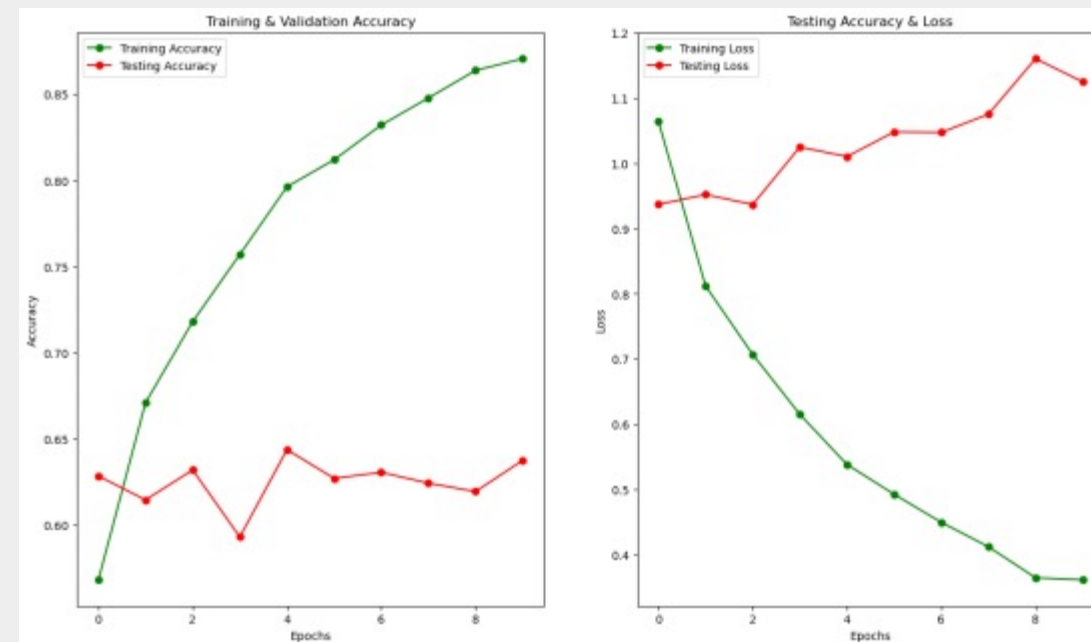


03.

# Multi-label

■ 단일 분류 결과

- Single-Label 모델의 Accuracy, Loss 그래프



- Confusion Matrix 분석

|            | Precision | Recall | F1-Score |
|------------|-----------|--------|----------|
| 박리         | 0.44      | 0.53   | 0.48     |
| 백태누수       | 0.64      | 0.73   | 0.68     |
| 철근노출       | 0.65      | 0.58   | 0.61     |
| 균열         | 0.76      | 0.61   | 0.67     |
|            |           |        |          |
| Weight Avg | 0.62      | 0.61   | 0.61     |
| Macro Avg  | 0.63      | 0.61   | 0.61     |
| Accuracy   | 0.61      |        |          |







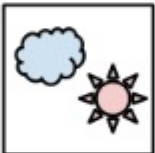

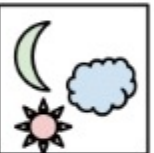
Single-Label 모델의 좋지 않는 성능으로 인해 피드백의 결과, 여러 결함이 동시에 일어나는 데이터 셋의 특성상 **Multi-Label** 을 활용하기로 함.



03.

# Multi-label

Multi-Class vs Multi-Label

|   | Multi-Class  | Multi-Label  |
|---|--|--|
| C = 3   |  |  |
| <br><br> | Samples<br>   | Samples<br>   |
|   | Labels (t)<br>[0 0 1] [1 0 0] [0 1 0]  | Labels (t)<br>[1 0 1] [0 1 0] [1 1 1]  |

산출물(output)에 따라 Single-label, Multi-label로 분류

**Single-label Classification :** 분류하고자 하는 이미지에 한 객체만 표현, 다른 객체가 표현 될 경우 그 중 대표하는 한 객체로 판별 및 학습 진행.  
분류 과정이 단순, 데이터들 공통적 특성을 집중적으로 학습하기에 적합

**Multi-label Classification :** 이미지에 포함된 모든 객체에 대한 라벨링이 진행, 해당 객체의 유무를 판별하는데 초점을 둠.  
위치 탐지가 없는 검출로 검출보다 쉬운 작업이지만 기본 분류에서 두드러진 최적화 과정이 없기에, 이를 해결하기 위한 다양한 접근들 존재.



03.

## Multi-label

■ Multi-label Classification의  
필요성



빨강 : 균열 / 노랑 : 박리 / 파랑 : 백태누수 / 초록 : 철근노출

해당 이미지에서 확인할 수 있듯이 구조물 결함들은 동시에 발생하는 경우가 존재. 따라 이미지내 존재하는 결함들을 명확히 분류하고자 Multi-label 기법을 사용하여 모델을 개발. 그에 따라 Loss function , Activation function 수정.

| Single-label Clasification                   | Multi-label Clasification            |
|--|--------------------------------------|
| Loss Function :<br>Categorical cross entropy | Loss function : binary cross entropy |
| Activation Function : Softmax                | Activation Function : Sigmoid        |

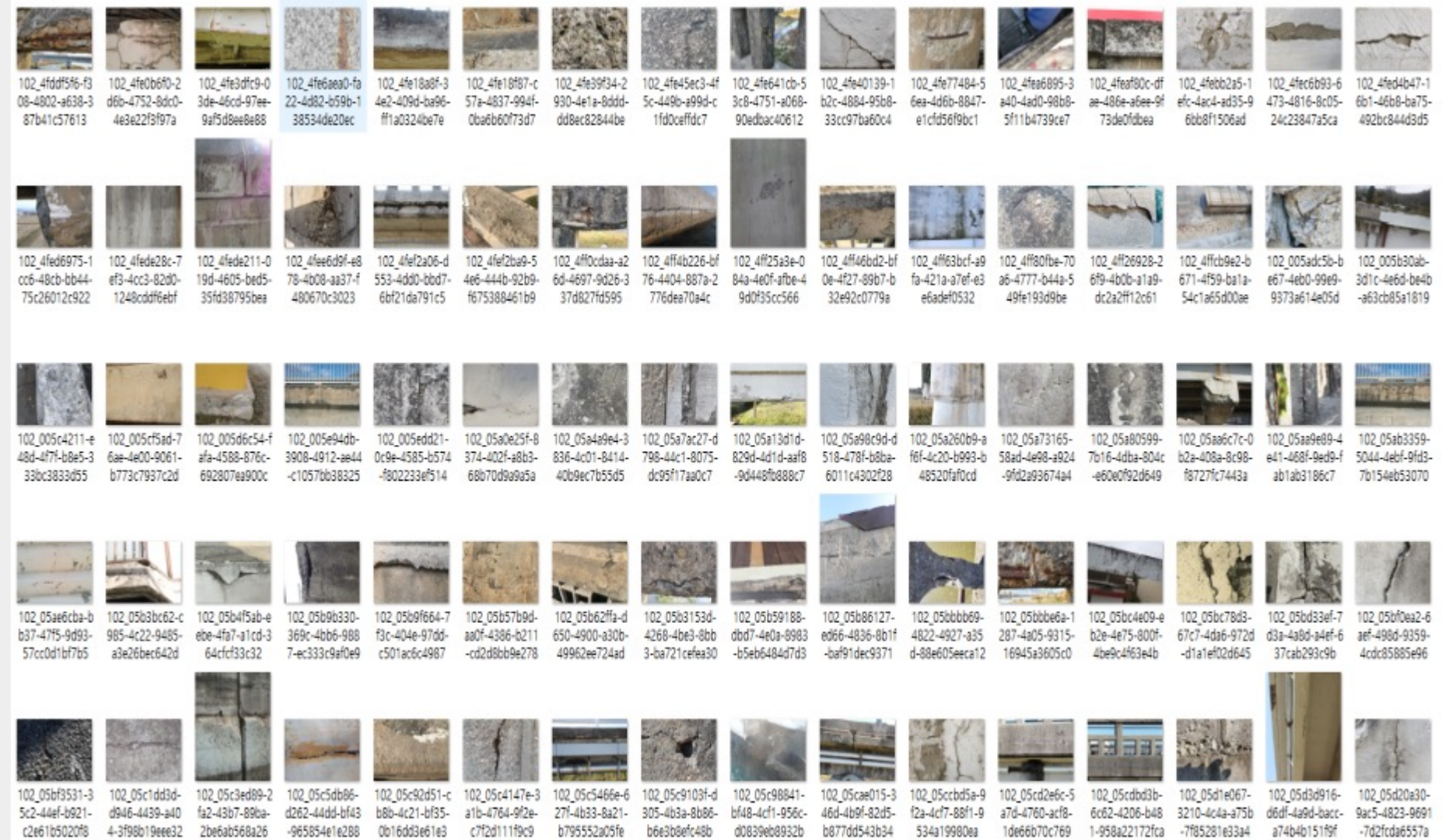


# 데이터 셋 구축

■ 데이터 셋 탐색

**AI-hub 데이터셋 : 고화질 카메라와 드론을 활용하여  
4가지 종류의 콘크리트 결함에 대해 35만건의 데이터로  
구성. 그 중 28000건의 데이터를 연구에 활용.**

**4가지 결함 클래스는 균열, 박리 박락, 백태 누수,  
철근 노출로 구성.**





# 데이터 셋 구축

## 데이터 셋 재라벨링

각 이미지마다 결함의 유무를 직접 확인하여 라벨링을 진행.

사람의 눈으로 확인하는 작업 특성상, 사진 환경에 따라 구분이 어려운 경우가 존재하여 구체적인 정제 및 라벨링 규칙 구축.

28,000건에서 1,048건을 정제하여 26,952건의 이미지를 활용할 예정.

| 결함       | 재라벨링 규칙   |
|----------|---|
| 균열       | 1. 좁은 폭으로 인한 비가시성 균열은 제외<br>2. 크기와 길이가 불충분한 균열은 제외                            |
| 박리<br>박락 | 1. 균열 부근이 아닌 박리 박락은 제외<br>2. 교량에서 흔적이 불충분한 박리는 제외<br>3. 돌과 자갈이 섞인 재질은 박리로 간주  |
| 백태<br>누수 | 1. 명확히 드러나지 않는 백태 누수는 제외<br>(진한 갈색, 흰색, 검은색)<br>2. 원인이 불분명한 흰색인 경우의 백태 누수는 제외 |
| 철근<br>노출 | 1. 철근 노출의 형태가 50% 이하인 경우는 제외<br>2. 철근으로 확정할 수 없는 경우 제외                        |

# 데이터 셋 구축

■ 구축된 데이터 셋 확인

26,952건의 이미지의 각 결함별 비율을 분석한 결과.

Single-label에서 클래스마다 비율이 25%씩 존재하였으나, Multi-label로 변경하면서 결함들이 2배에서 3.5배까지 증가한 점을 확인.

즉 단일 라벨이라고 명명된 데이터였으나, 다른 결함과 공존하는 경우가 많은 심각성 입증.

| 단일 결함<br>유형 | 라벨 수   | 비율(%) |
|-------------|--------|-------|
| 균열          | 12,294 | 45.6  |
| 박리<br>박락    | 24,196 | 89.7  |
| 백태<br>누수    | 12,514 | 46.4  |
| 철근<br>노출    | 9,741  | 36.1  |



05.

## 모델 개발

### ■ Backbone Network

특징 추출을 위한 Backbone Network : VGG, ResNet, DenseNet

| Transfer Learning을 위한 Pretrained Model 선정                                |   |  |
|--|---|--|
| VGG16  | ResNet50  | DenseNet201  |
| 합성곱 계층과 풀링 계층으로<br>구성되는 기본적인 CNN.<br>3x3의 작은 필터를 사용한<br>합성곱 계층을 연속으로 거침. | VGG 신경망을 기반으로<br>구성되어 있는 CNN.<br>Skip-connection을 도입한<br>Residual 블록을 활용해<br>깊은층에서의 성능 저하 방지. | ResNet과 유사한 CNN 구조로<br>feed forward 방식에서 각 레이어들이<br>모든 다른 레이어와 connect를<br>하는 것이 특징. |

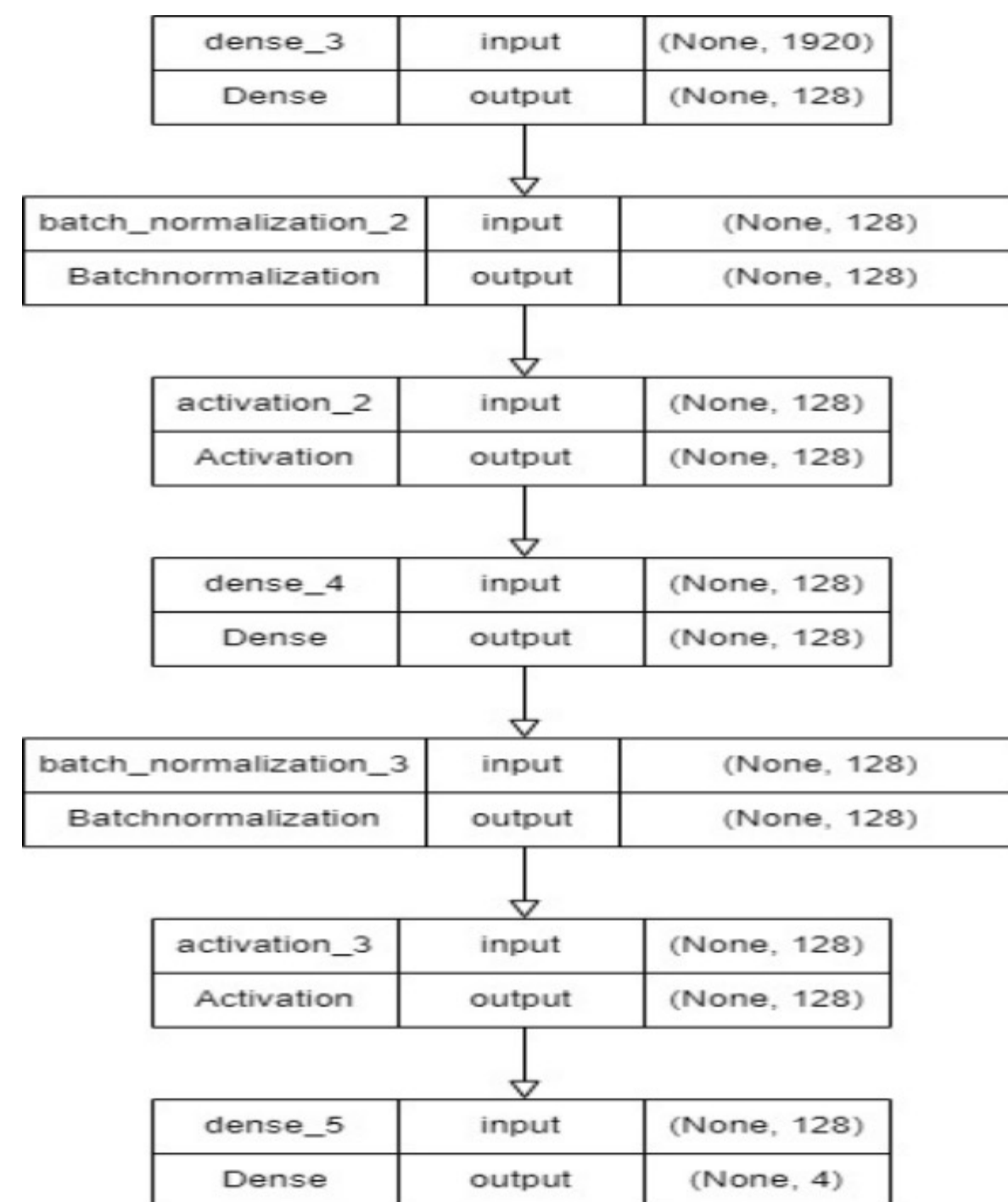
# 05.

## 모델 개발

### Transfer learning & Fine tuning

pretrained 모델을 사용하여 다른 데이터셋에 적용한 Transfer Learning 방법과 기존의 분류기 (Classifier)를 제거하고 새로운 분류기를 추가한 모델로 학습하는 Fine-tuning 방법을 사용.

tf.keras.layers.Dense(128)과 BatchNormalization, relu 활성화 함수를 두번 반복하였고, 마지막에 타깃 결함의 개수에 맞는 Dense(4, activation='sigmoid')를 추가한 새로운 Classifier.





05.

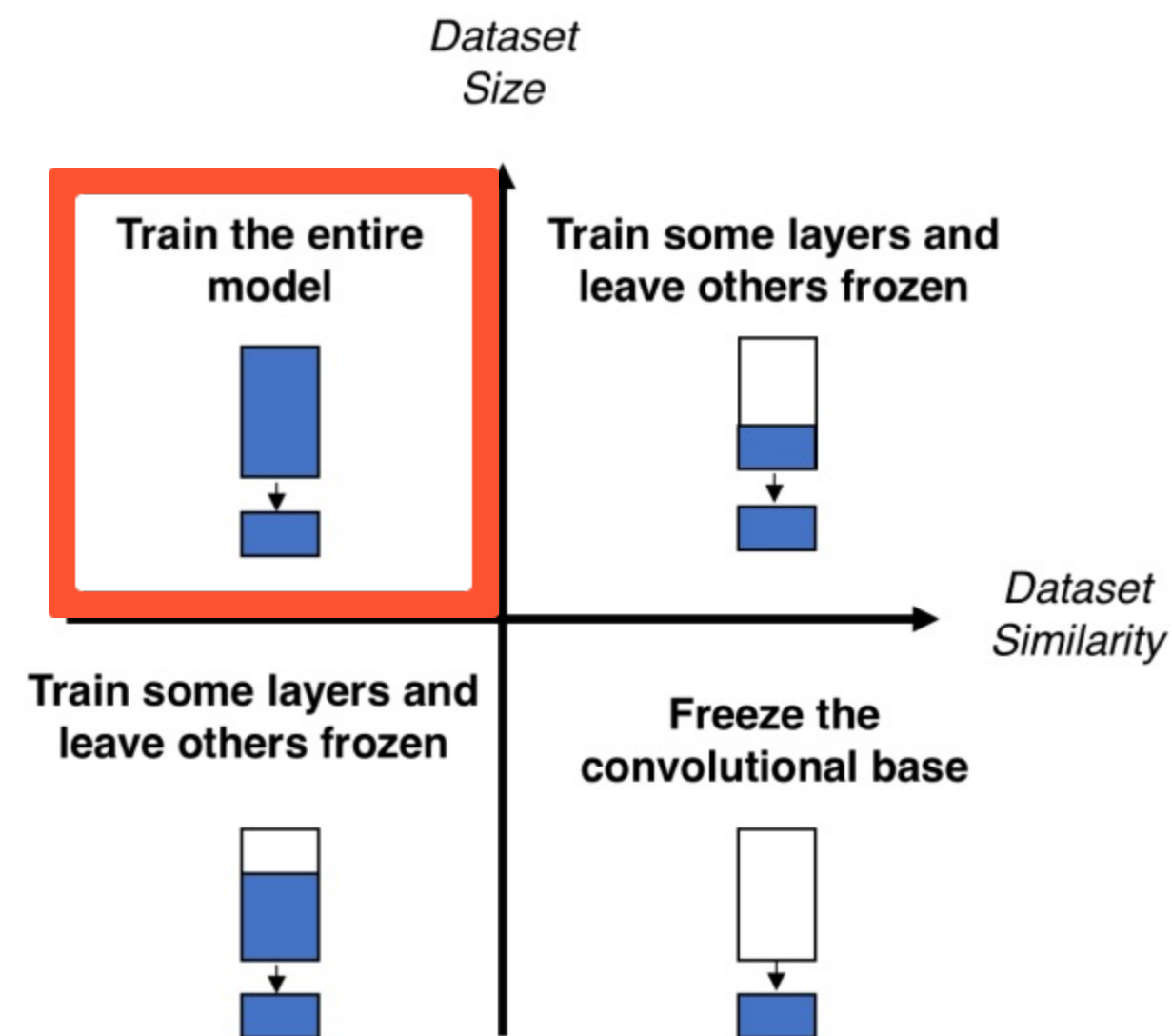
## 모델 개발

Transfer learning & Fine tuning

Fine-Tuning 방식에는 데이터의 유사성과 양에 따라서 학습시키는 방식이 4가지 존재.

본 모델에서 학습에 사용된 데이터는 기존의 'imagenet' 데이터셋과의 유사성이 없고 그 양이 약 27000장으로 많음.

전체의 모델을 새로 학습하여, 사전학습 모델에서의 구조만 사용하는 방식을 채택.



05.

## 모델 개발

### Hyperparameter 선정

CosineAnnealingWarmRestarts를 활용하여 Learning rate를 epoch에 따라 지속적으로 감소시켜주면서, 한번씩 튀어오르게 하는 방식 채택.

Optimizer로 SGD momentum과 Adam, AdamW를 사용하였으나 Adam이 성능이 가장 좋았음.

Epochs가 높다고 하여 성능이 좋아지지 않는다고 판단하여 warmup을 2번정도 할 수 있는 60을 채택.

|               | Hyperparameter              |
|---------------|-----------------------------|
| Learning rate | 1e-4                        |
| Scheduler     | CosineAnnealingWarmRestarts |
| Optimizer     | Adam                        |
| Epochs        | 60                          |



# 05.

## 모델 개발

■ 평가 지표 및 모델 성능 비교

클래스별 데이터의 불균형한 분포에 영향이 적은 Weight Average를 적용한 Precision, Recall, F1-score값 추출.

3개의 모델의 비교한 결과 전반적으로 성능이 우수, 그 중 DenseNet201이 Precision 0.889, Recall 0.909, F1-Score 0.899로 가장 높은 성능을 보이는 것으로 확인.

|           | VGG16 | ResNet50 | DenseNet201 |
|-----------|-------|----------|-------------|
| Precision | 0.857 | 0.879    | 0.889       |
| Recall    | 0.910 | 0.895    | 0.909       |
| F1-Score  | 0.881 | 0.886    | 0.899       |

# 05.

## 모델 개발

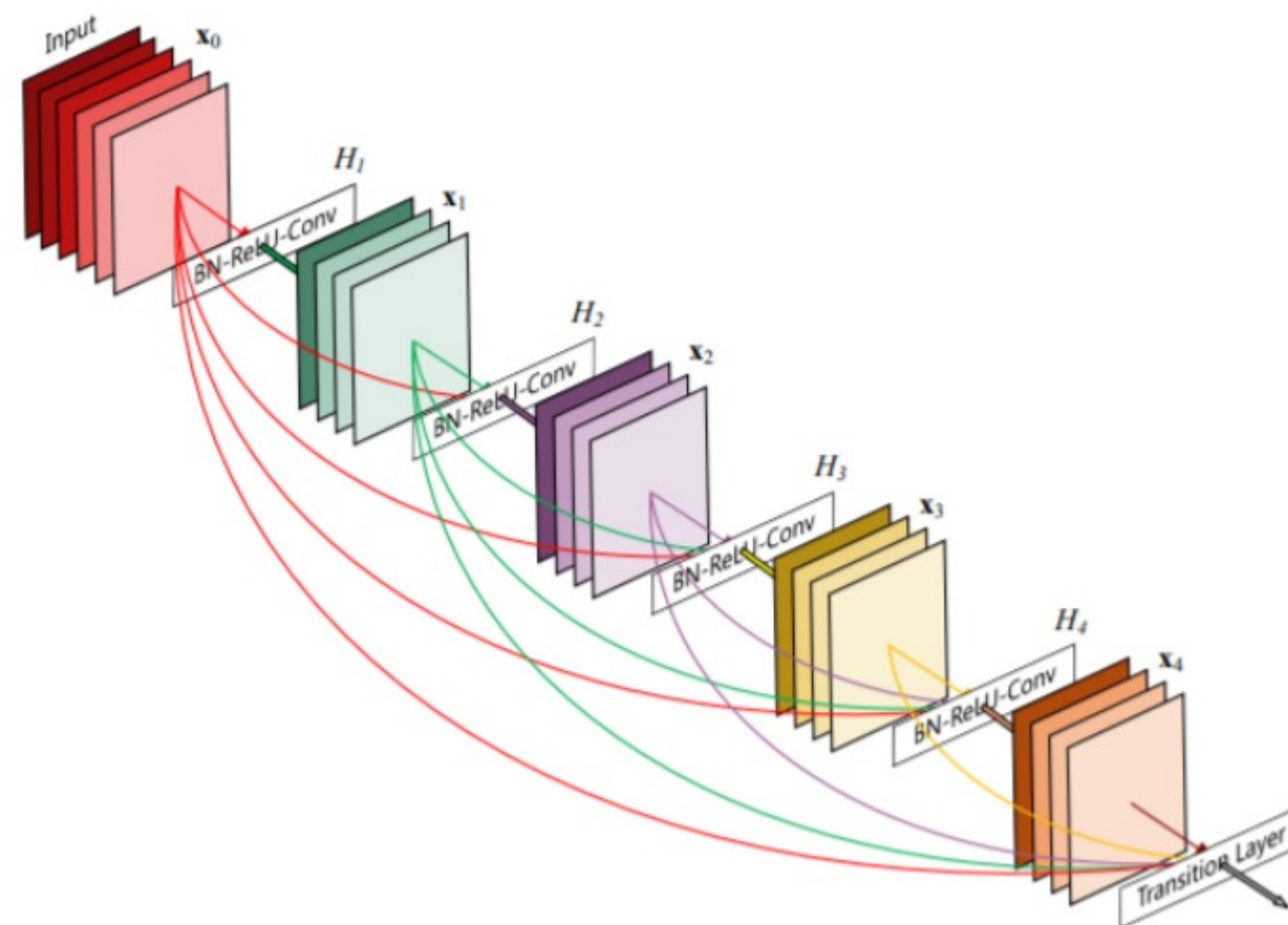
### DenseNet 모델 분석

입력값을 Summation(ResNet)이 아닌 Concatenation.

처음 레이어의 피쳐맵, 값들을 마지막 레이어까지 연결.

- Feature 재사용 가능.
- 정보가 소실되는 것을 방지.
- 오차역전파시 기울기 소실되는 것을 완화.
- 다양한 레이어 피쳐맵 학습에 따른 정규화 효과.

해당 장점을 통한 연산량, 속도, 성능 면에서 우수한 모습을 보여 본 연구모델에서 최종모델로 선정.



**Figure 1:** A 5-layer dense block with a growth rate of  $k = 4$ . Each layer takes all preceding feature-maps as input.

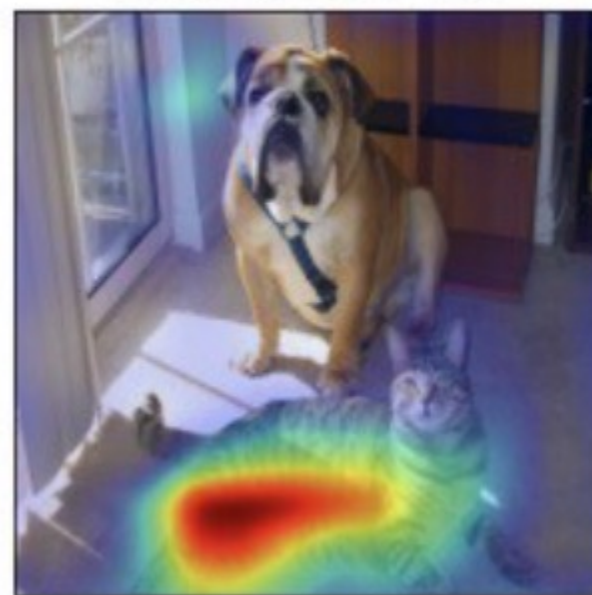


06.

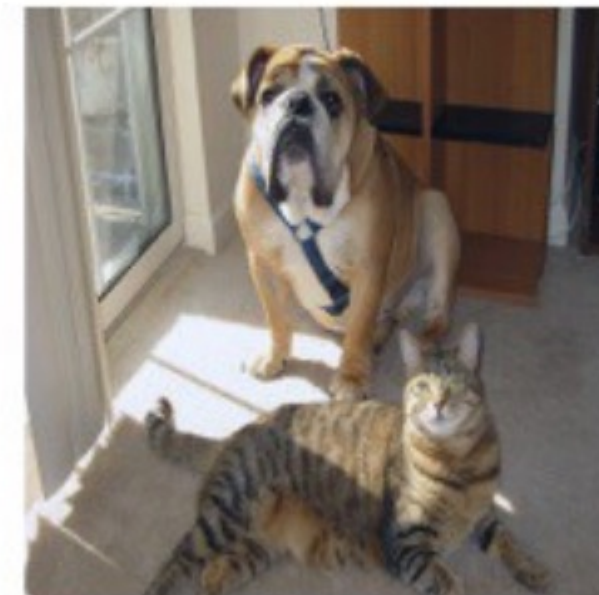
# Grad-CAM

■ Grad-CAM의 정의

Grad-CAM for "Cat"



Grad-CAM for "Dog"



- CNN 모델의 예측 결과에 대해 설명력과 해석력을 시각화하기 위한 기법으로 이미지의 중요한 영역을 강조하는 Localization Map을 만들어 시각화 함.
- Global Average Pooling 층이 꼭 필요한 CAM의 단점을 극복, 모델의 구조를 굳이 변경하지 않아도 됨.
- 적용할 수 있는 CNN 모델의 범위가 넓어짐.

06.

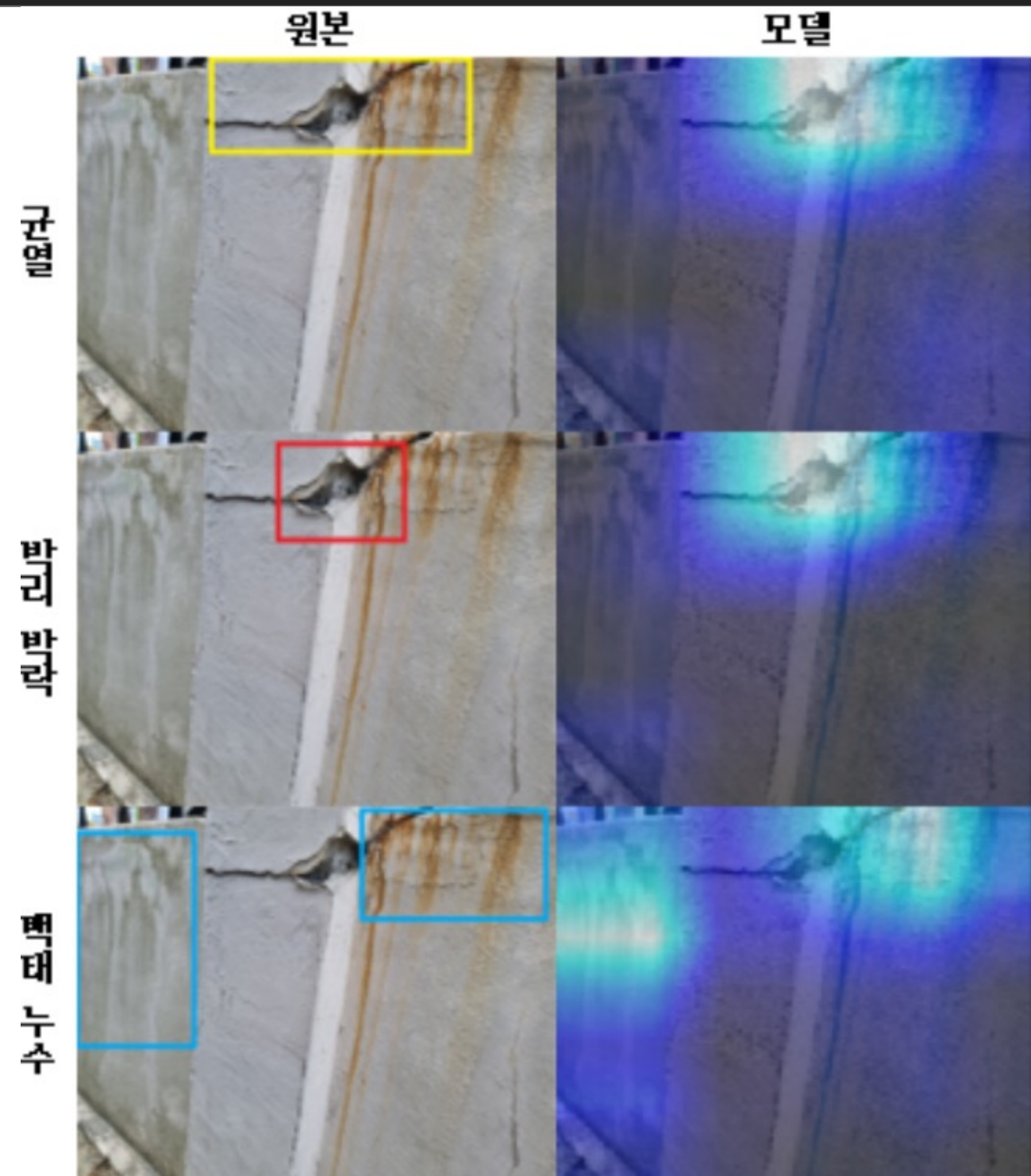
# Grad-CAM

■ Grad-CAM 활용한 모델 결과 분석

해당 이미지는 균열, 박리박락, 백태누수가 포함된 이미지로, 색이 밝을 수록 높은 가중치로 두고 모델이 판단한것.

균열과 박리 박락 부분은 경계박스와 Grad\_CAM 시각화 부분이 거의 일치하며, 백태누수 부분은 박리박락 하단 부분보다 왼쪽 부분에 가중치를 두고 예측.

해당 결과를 보아 DenseNet201이 다중 결함을 분류하는데 강건함을 보이는 것으로 판단.



07.

## 결론

■ 본 연구의 의의

01

기존의 연구보다 많은 데이터 셋을 구축하여  
데이터 기반의 학습력을 높인 점.



02

균열에만 국한된 연구를 결함으로 확장시켜,  
결함의 특성을 반영한 다중 라벨 분류를  
적용한 점.



03

Grad-CAM을 통해 모델의 해석력, 설명력을  
확인한 점.



# 07.

## 결론

■ 본 연구의 한계점

01

Transformer 기반의 모델 구축에  
성공했으나 높은 성능을 위한 학습과정에  
어려움을 겪음



02

시간상의 제약으로 하이퍼파라미터 튜닝을  
통한 성능 고도화에 실패함.



03

텐서플로의 언어적 제약으로 더 많은  
하이퍼파라미터들을 활용하지 못하였음.