

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

아덱 폭주족

2023 연구데이터 분석활용 경진대회

Contents

- 1. 연구 개요
 - 1.1 문제 정의
 - 1.2 연구 목표
- 2. 연구 방법
 - 2.1 데이터 소개
 - 2.2 데이터 증강
 - 2.3 모델
- 3. 연구 결과
 - 3.1 연구 결과
- 4. 결론 및 기대 효과
 - 4.1 연구 결론
 - 4.2 기대 효과
- 5. 시연
 - 5.1 시연

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

1. 연구 개요 Introduction

1.1 연구 배경 및 목표

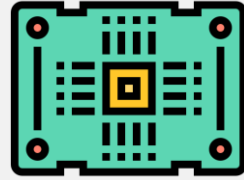
1.1 연구 배경

연구 개요

연구 방법

연구 결과

결론 및 기대효과



PCB(인쇄 회로 기판)

거의 대부분의 전자 장치에 포함되어 있으며, 전기 배선을 효율적으로 설계할 수 있도록 하여 전자기기 크기를 줄이고 성능을 높이는 역할을 함

01

제품의 불량률을 신속하고 정확하게

공정의 효율성을 높이고 경쟁력을 확보하기 위한 중요한 과제

02

적절하고 일관된 품질을 유지

품질 관리(Quality Control) 과정이 잘 수행되지 않는다면 신뢰성과 내구성에 영향

03

전문가에만 의존하는 것은 경제적 측면에서 비효율적

1달에 3억 개로 대량 생산되는 PCB 특성상 지나친 인력과 시간을 소모



공정의 자동화를 위해 **정확한 불량률**을 검출하고 **설명**할 수 있는 AI 시스템 구축

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

2. 연구 방법 Research Methods

- 2.1 데이터 소개
- 2.2 데이터 증강
- 2.3 모델

2.1 데이터 소개

2.1.1 클래스별 데이터 불균형

연구 개요

연구 방법

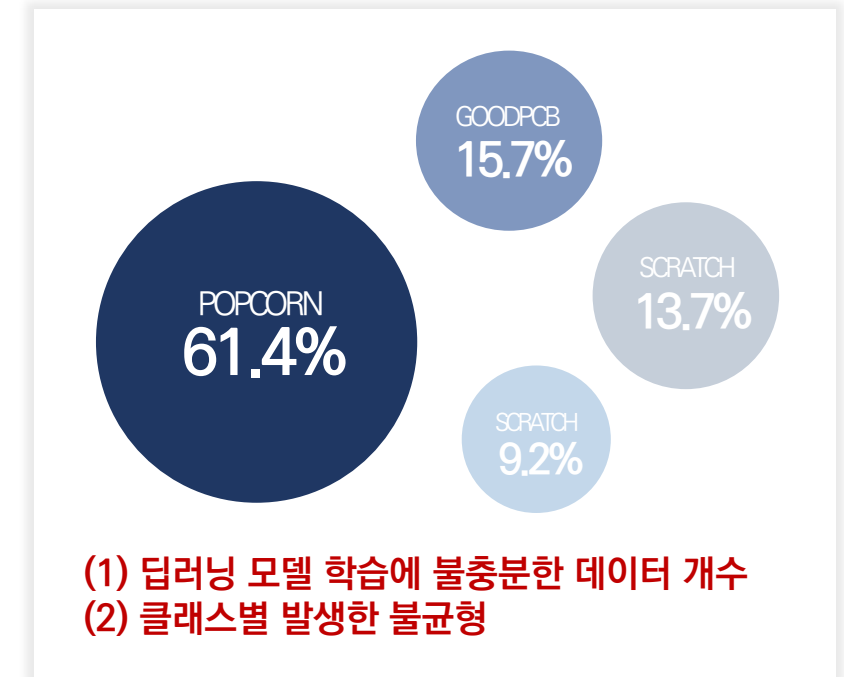
연구 결과

결론 및 기대효과

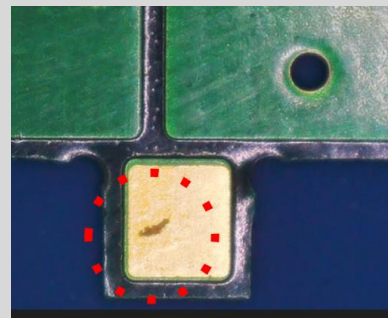
시연

1. 데이터 개수 및 비율

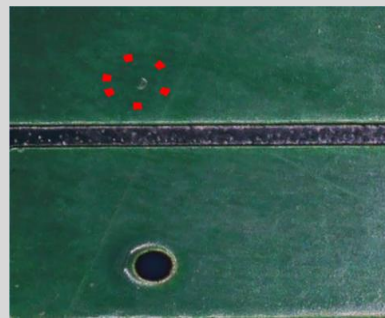
Class	데이터 개수
SCRATCH	79
DELAMINATION	53
POPCORN	353
GOODPCB	90
총합	575



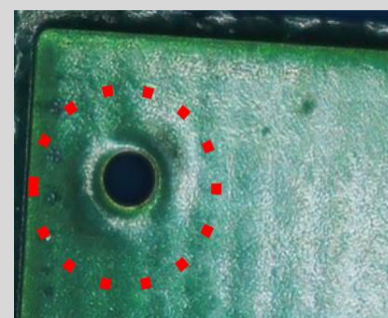
2. 데이터 구성



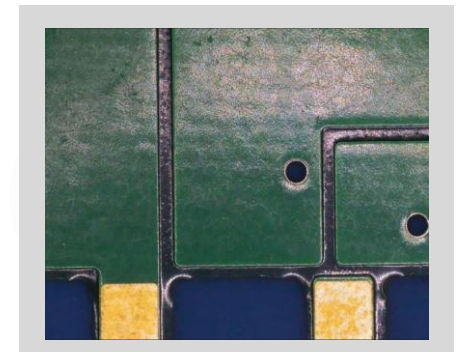
[SCRATCH]



[DELAMINATION]



[POPCORN]

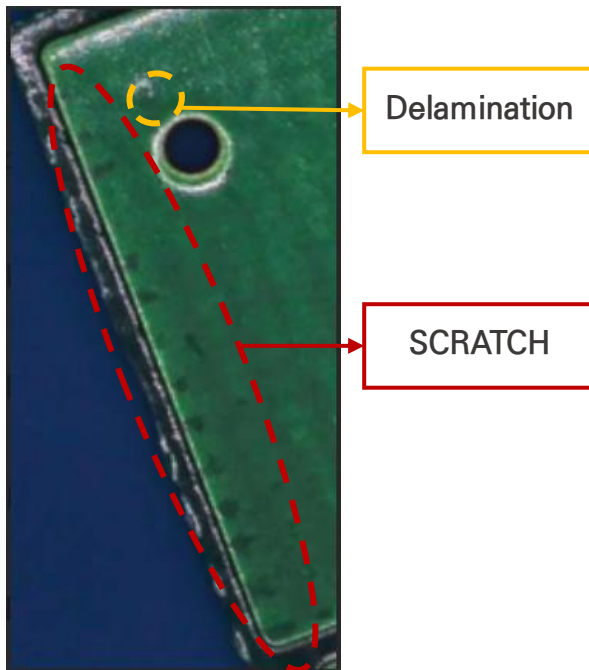


[GOODPCB]

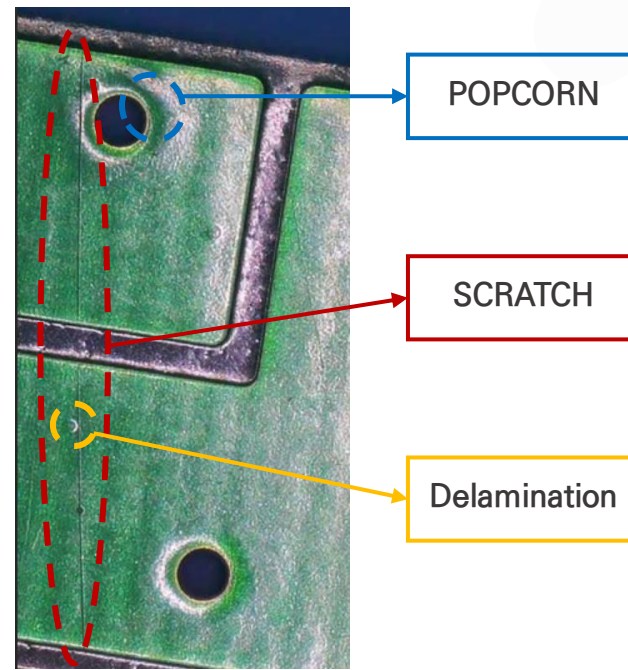
2.1 데이터 소개 2.1.2 중복 데이터

1. 중복 불량 데이터

- 제공 받은 단일 불량 Class에서 **중복된 불량 발생 가능성 존재 -> (3) Multi-Label Problem**
- 주어진 조건에 따라 중복 불량 의심 데이터 분류



[SCRATCH] NO.39



[POPCORN] NO.37

Class	데이터 개수
SCRATCH	79
DELAMINATION	53
POPCORN	353

Class	중복 의심 데이터
SCRATCH	17
DELAMINATION	9
POPCORN	142

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

2.1 데이터 소개 2.1.2 중복 데이터

연구 개요

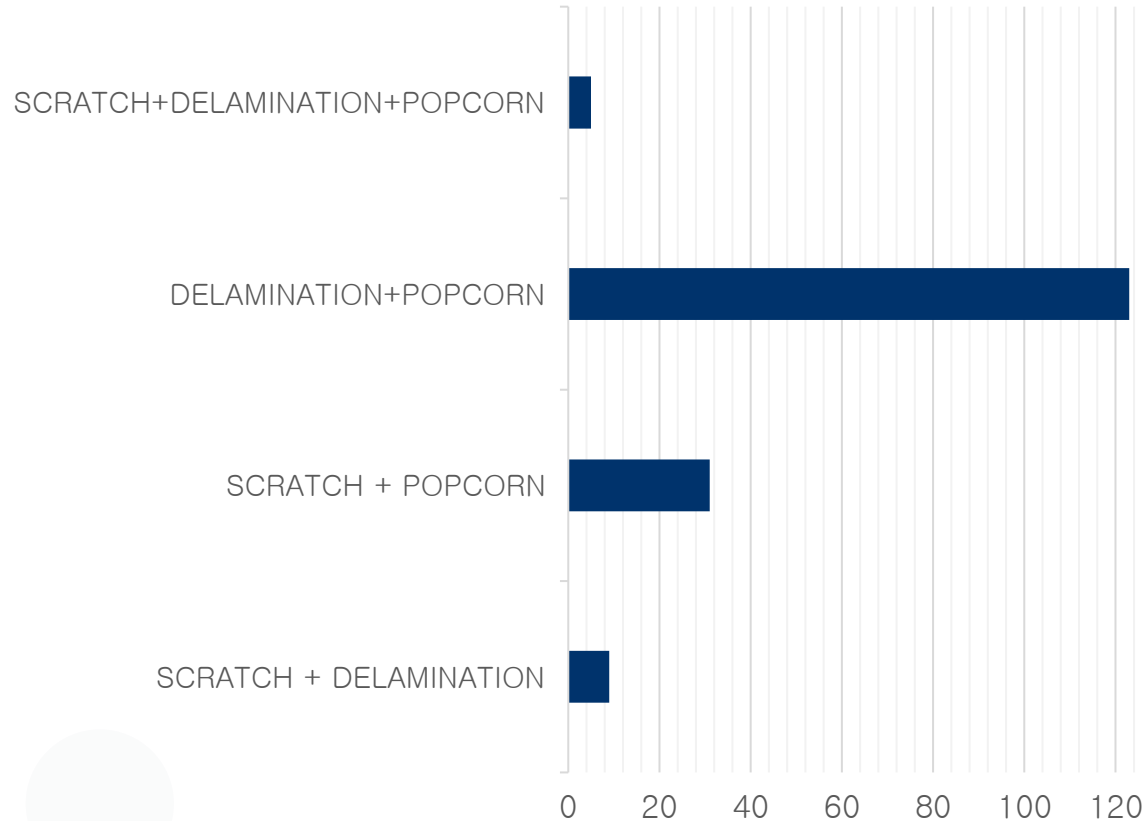
연구 방법

연구 결과

결론 및 기대효과

시연

1. 중복 불량 의심 데이터



클래스	SCRATCH DELAMINATION	SCRATCH POPCORN	DELAMINATION POPCORN	SCRATCH DELAMINATION POPCORN
개수	9	31	123	5

중복 불량 데이터의 문제점

- 비전문가 연구자가 라벨링 후 분류 모델 학습
- 명확한 정답 레이블이 주어지지 않아 올바른 학습 어려움
- 불량 유형 중복으로 판단되는 데이터의 양 역시 충분하지 않음
 - 중복 유형별로 불균형적으로 분포되어 있음

중복 의심 레이블을 가진 Image Data를
Multi-label Classification을 통해
분류를 시도하는 데 의의

추후 전문가에 의해 라벨링 완료된 중복 불량 데이터
학습 데이터셋에 추가하여 학습 정확도 향상 기대

2.1 데이터 소개 2.1.3 Flow Chart

1. 데이터 전처리 & 증강

[데이터 전처리]
Padding & Resize

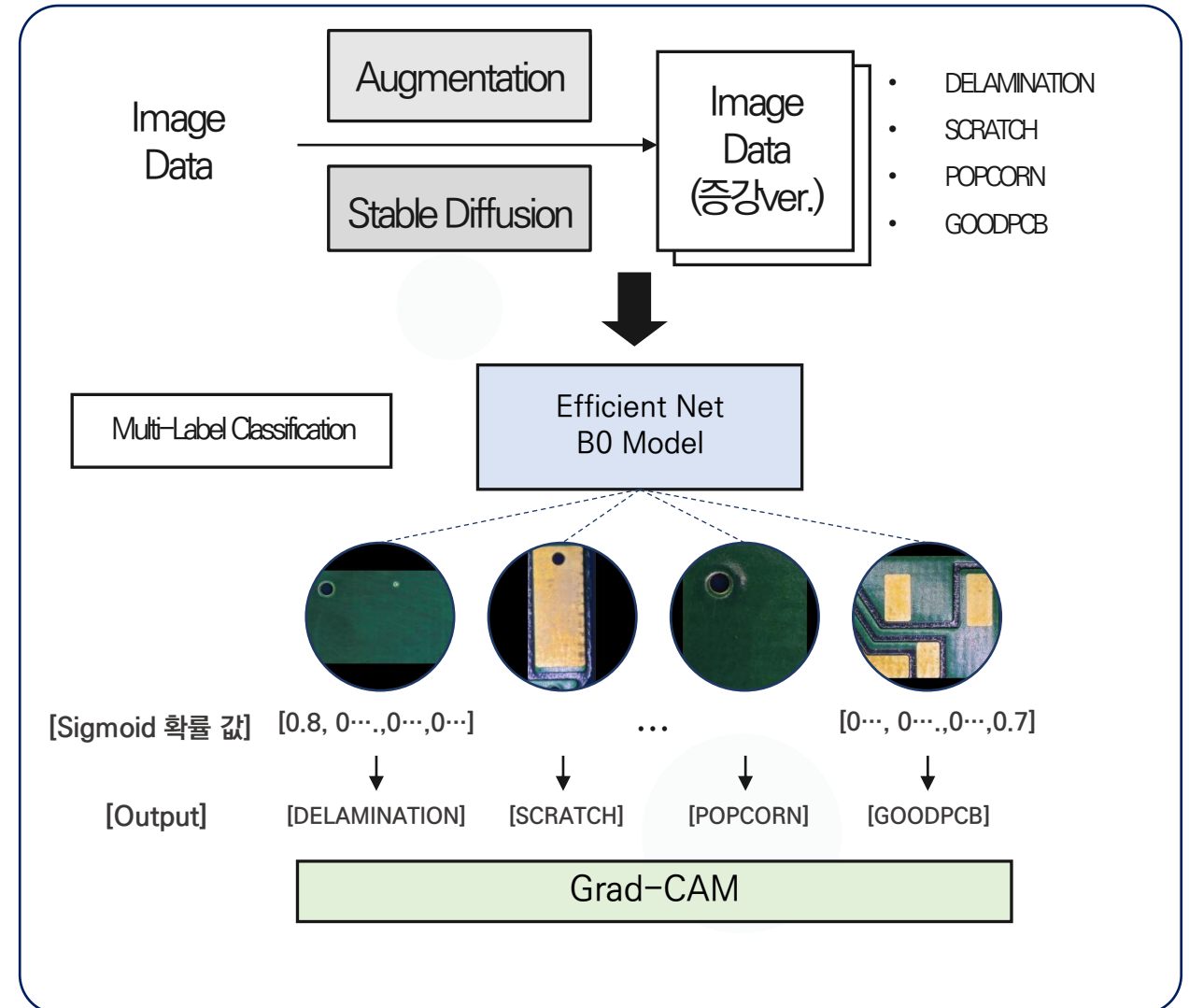
[Data Augmentation1]
회전&노이즈&반전

[Data Augmentation2]
Stable Diffusion

2. 모델 & 시각화

[Efficient Net]
중복 처리 알고리즘 추가

XAI : Grad-CAM 시각화



연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

2.2 데이터 증강 2.2.1 데이터 전처리

1. Padding & Resize



- Padding을 통해 가로x세로 비율을 유지하며, Resize로 인한 원본 이미지 비율 왜곡 방지
- 서로 다른 크기의 이미지를 (224x224)로 통일하는 Resize를 수행하여 모델
- 컨볼루션 신경망(Convolutional Neural Network, CNN)에서 입력 이미지의 불량 패턴 학습 원활

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

2.2 데이터 증강 2.2.2 데이터 Augmentation

연구 개요

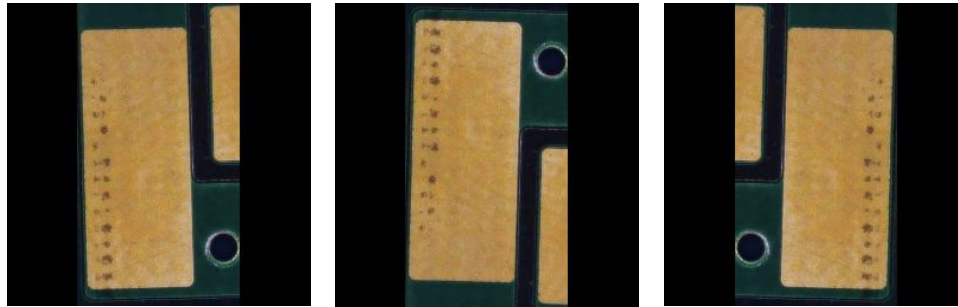
연구 방법

연구 결과

결론 및 기대효과

시연

1. 상하 반전 & 좌우 반전



SCRATCH NO.3

[Original]

SCRATCH NO.3

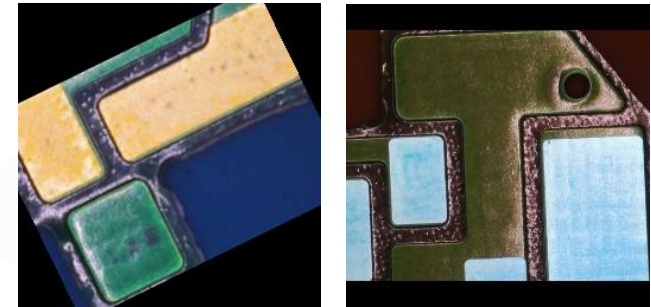
[상하반전]

SCRATCH NO.3

[좌우 반전]

Rotate
&
Noise

2. Rotate & Noise



SCRATCH

[Rotate]

POPCORN

[Noise]

- 원본 이미지에 대해 상하반전과 좌우 반전으로 원본 이미지 당 2장씩 새로운 이미지 생성
- 원본 이미지와 상하/좌우 반전된 이미지에서 무작위 추출하여 Rotate와 Noise 이미지 생성
- 이미지에 가우시안 노이즈(Gaussian Noise)를 추가 & $(-30^\circ \sim 30^\circ)$ 사이의 무작위 각도로 Rotate 이미지 생성
- 회전 시, 안티 앨리어싱(Anti-Aliasing) 방법으로 BICUBIC을 사용하여 계단 현상 완화 및 시각적 품질 향상
- 데이터 Augmentation으로 모델이 Train Data에 과적합 되는 것을 방지

2.2 데이터 증강 2.2.3 DCGAN

연구 개요

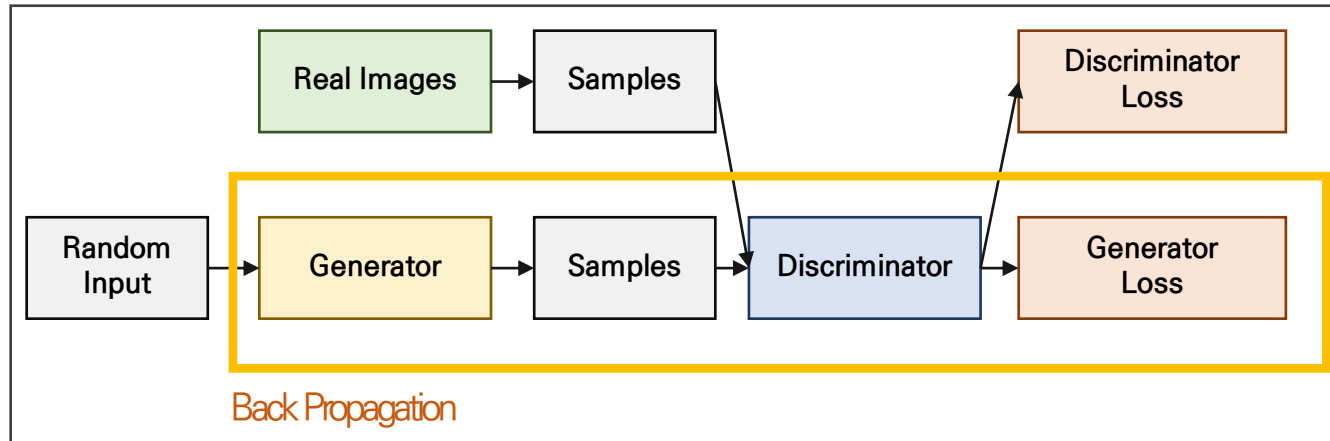
연구 방법

연구 결과

결론 및 기대효과

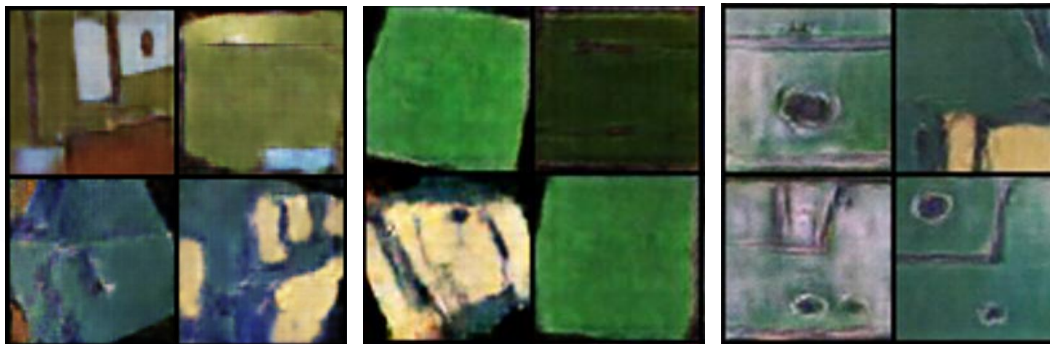
시연

1. DCGAN(Deep Convolutional Generative Adversarial Network)



- 생성적 적대 신경망(GAN)에 CNN 도입
- 실제 데이터를 모방하는 가짜 데이터를 생성
- 경쟁적인 학습 과정 속에서 generator와 discriminator 모두 점진적으로 개선

2. Image examples



[Delamination]

[Scratch]

[Popcorn]

- 고해상도의 이미지를 생성하기 위해서는 많은 이미지 데이터 필요
- 초기 데이터 부족 문제로 고품질의 이미지 생성 어려움
- 불량 패턴이 잘 보이지 않으며 왜곡 발생

2.2 데이터 증강 2.2.4 Stable Diffusion

연구 개요

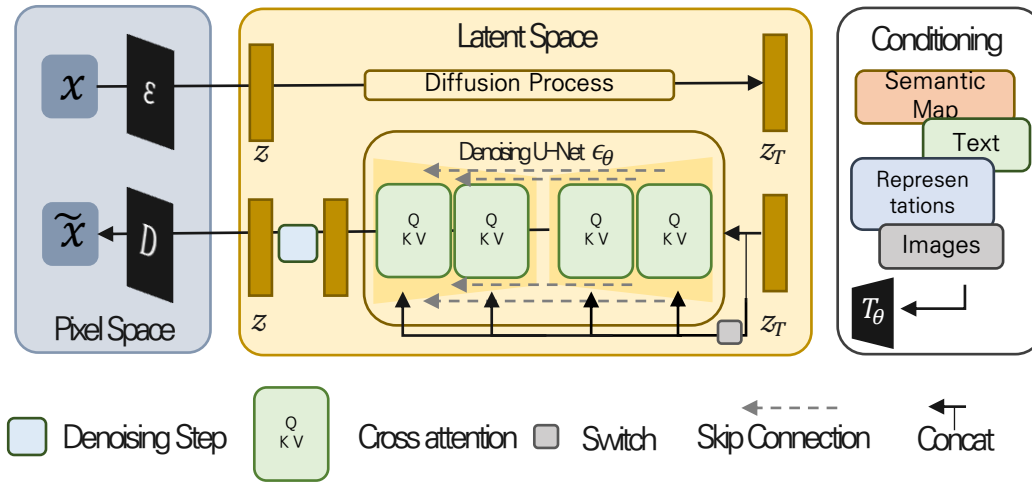
연구 방법

연구 결과

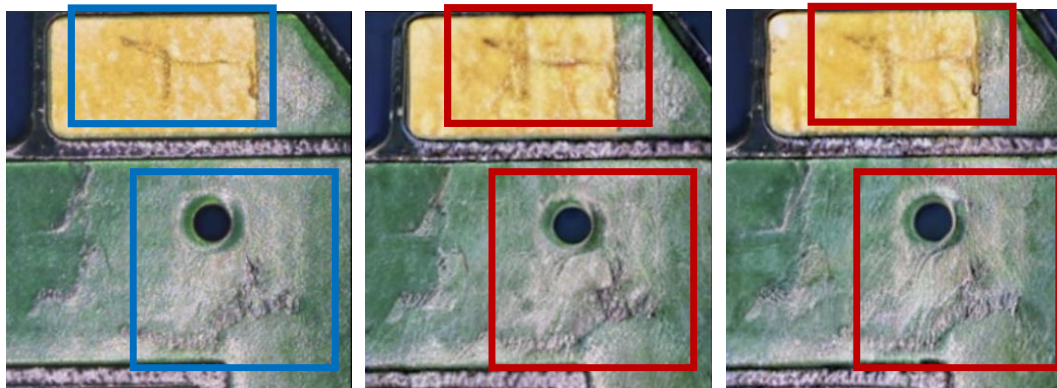
결론 및 기대효과

시연

1. Stable Diffusion



2. Image examples



[Original Image]

[Stable Diffusion Image]

[Stable Diffusion Image]

- 심층 생성 신경망의 일종인 잠재 확산 모델로 Diffusion process를 활용하는 딥러닝 모델
- 입력 이미지로부터 고품질의 이미지 생성
- 입력 이미지를 여러 단계로 거쳐 점진적으로 변환하면서 이미지를 개선

- Colab 기반의 Stable Diffusion Web UI를 활용
- 원본 이미지 당 2장씩 이미지 생성
- 불량 패턴이 미세하게 변화된 새로운 이미지 생성 가능

2.2 데이터 증강 2.2.5 데이터 개수 별 Loss & Accuracy

연구 개요

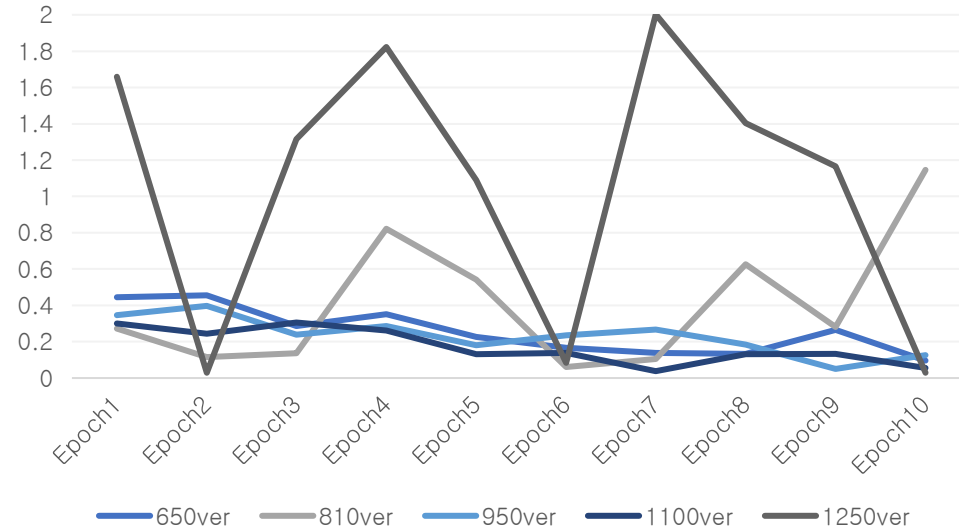
연구 방법

연구 결과

결론 및 기대효과

시연

1. Loss



2. Accuracy

표1. 데이터 개수 별 Class Accuracy

	Delamination	Popcorn	Scratch
650ver	95.77%	95.59%	92.65%
810ver	70.37%	89.50%	86.29%
950ver	95%	95.67%	96.96%
1100ver	88.40%	87.89%	89.66%
1250ver	97.78%	97.22%	93.88%

- Data Augmentation과 Stable Diffusion을 통해 650개, 810개, 950개, 1100개, 1250개의 데이터셋 구성
- 각 데이터 별로 모델을 학습한 후, Loss와 Label - Accuracy 측정
- 안정적인 Loss 추세와 높은 Label -Accuracy를 보이는 950개의 데이터셋을 최종 데이터셋으로 선정

2.2 데이터 증강 2.2.6 데이터 구성

1. Train/Test/Validation

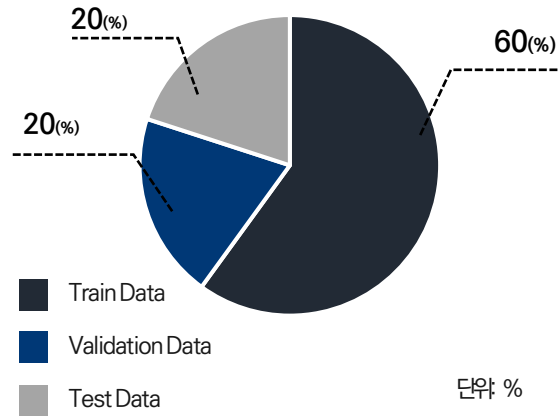


표1. 데이터 구성

	Train(전)	Test(전)	Train (후)	Test(후)	Validation	Total
SCRATCH	47	15	570	190	190	950
DELAMINATION	33	11	570	190	190	950
POPCORN	158	53	570	190	190	950
GOODPCB	67	23	570	190	190	950

2. Augmentation Ratio

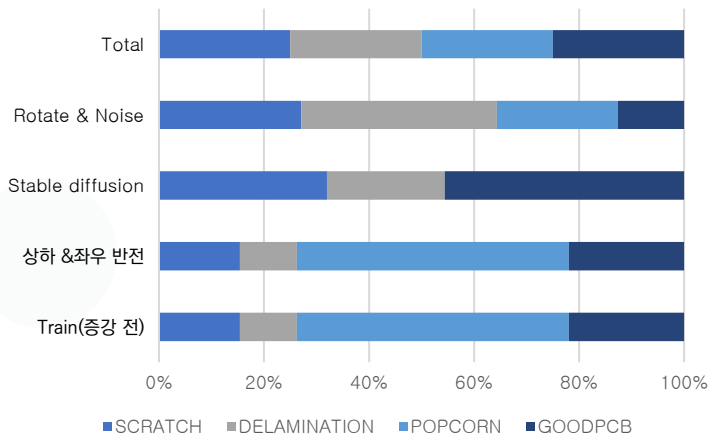


표2. Train Data Augmentation Ratio

	SCRATCH	DELAMINATION	POPCORN	GOODPCB
Train(전)	47	33	158	67
상하 &좌우 반전	94	66	316	134
Stable diffusion	282	198	0	402
Rotate & Noise	337	463	286	157
Total	760	760	760	760

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

2.2 데이터 증강

2.2.7 데이터 증강 전 후 성능 비교

연구 개요

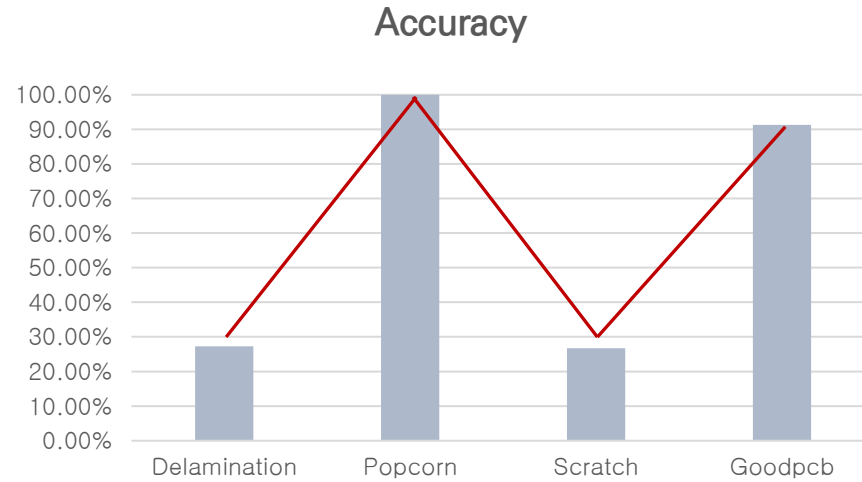
연구 방법

연구 결과

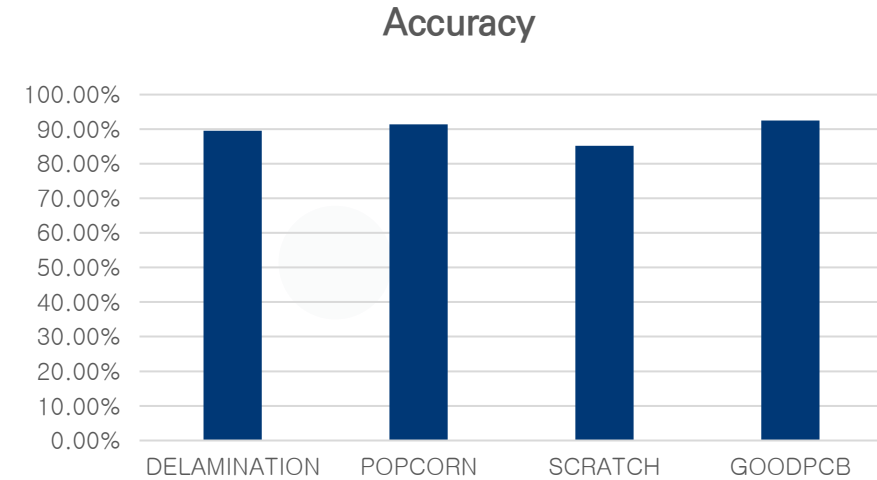
결론 및 기대효과

시연

1. 데이터 증강 전



2. 데이터 증강 후



Class	증강 전	증강 후
DELAMINATION	27.27%	89.47%
POPCORN	100.0%	91.32%
SCRATCH	26.67%	85.13%
GOODPCB	91.30%	92.50%

- 데이터 증강 전, 클래스 별 Accuracy 불균형 존재
- 데이터가 상대적으로 많은 POPCORN, GOODPCB는 높은 Accuracy를 보이지만 DELAMINATION, SCRATCH는 낮은 Accuracy를 보임
- 모델 학습을 위한 데이터 부족과 클래스별 불균형을 데이터 증강을 통해 해소하여 보다 높은 정확도를 가지는 모델 구축

2.3 모델 별 정확도 비교

2.3.1 MobileNet vs ResNet

연구 개요

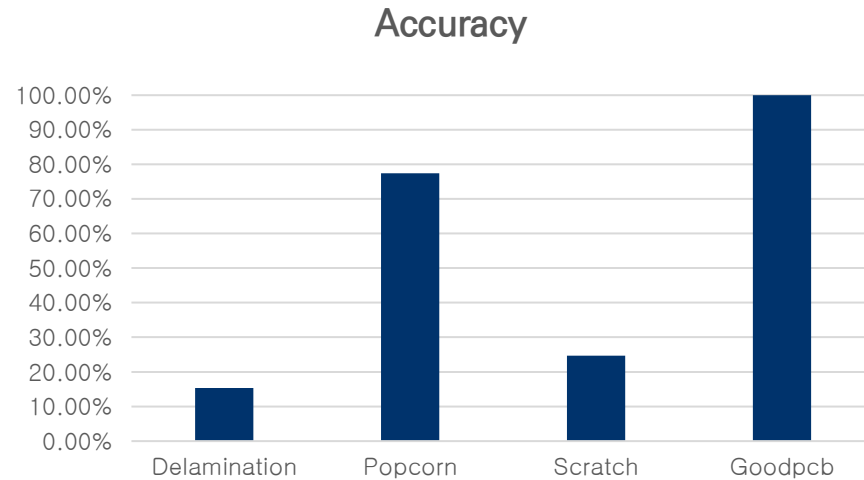
연구 방법

연구 결과

결론 및 기대효과

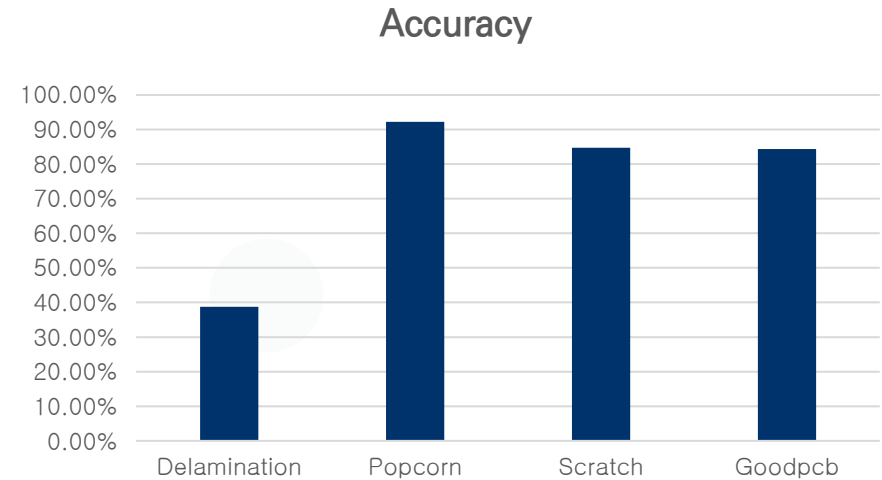
시연

1. MobileNet



Class	MobileNet	ResNet
DELAMINATION	15.30%	38.72%
POPCORN	77.40%	92.19%
SCRATCH	24.70%	84.72%
GOODPCB	100%	84.37%

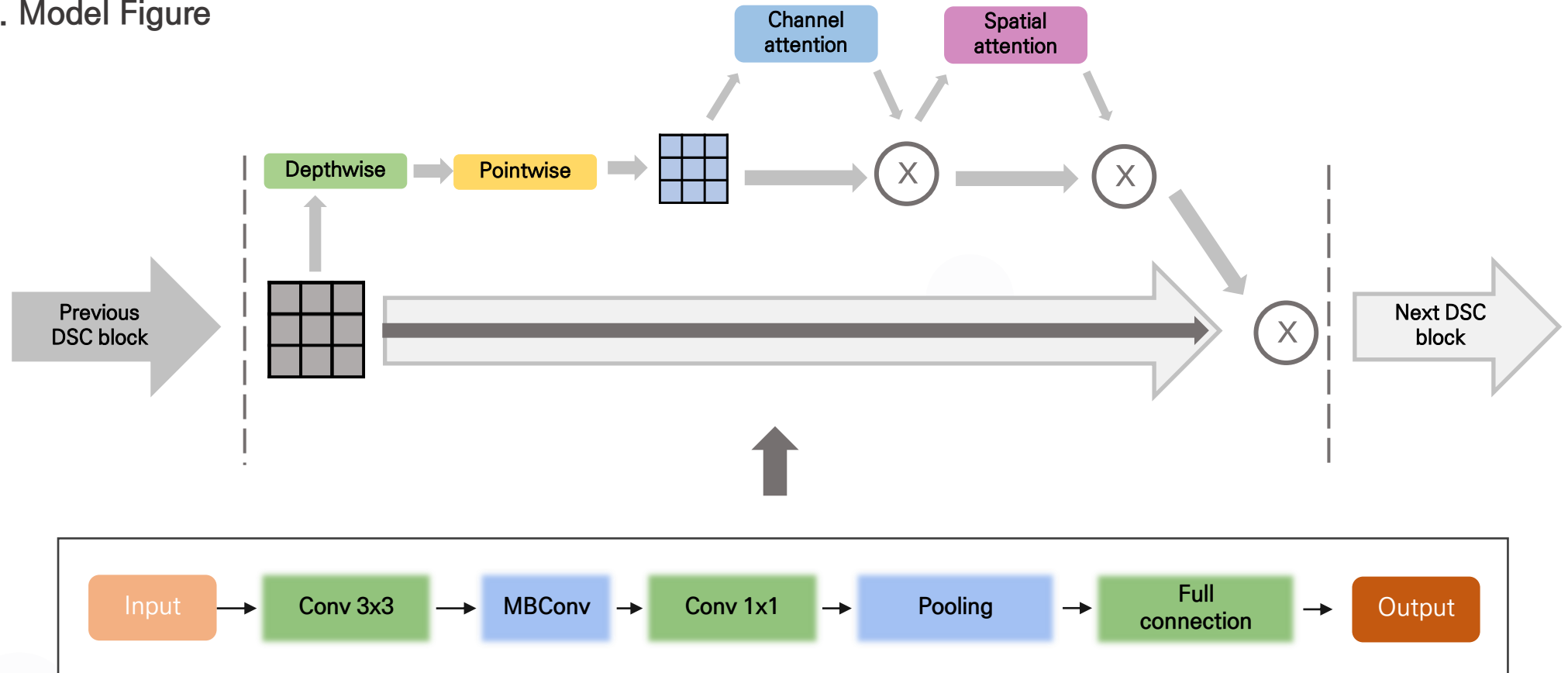
2. ResNet



Class	EfficientNet
DELAMINATION	89.47%
POPCORN	91.32%
SCRATCH	85.13%
GOODPCB	92.50%

2.3 모델 2.3.2 EfficientNetB0

1. Model Figure



- EfficientNetB0는 기존 ConvNet을 사용하여 scaling을 평가하는 방법에 더해 정확도와 FLOPS를 모두 최적화한 Classification 모델
- depth, width, image size 사이의 관계를 수식으로 표현, 세가지 요소의 균형을 맞추어 효율적으로 조절할 수 있는 compound scaling
- 적은 파라미터로 높은 정확도를 확보하여 더 가볍고 빠른 효율적인 모델

연구 개요

연구 방법

연구 결과

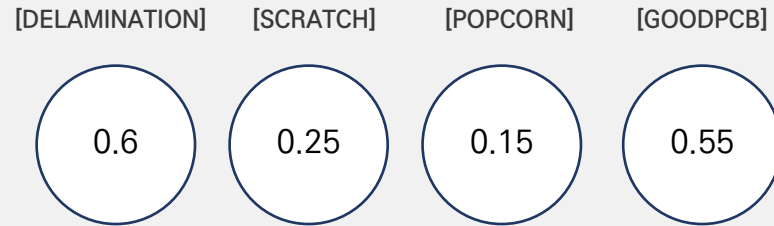
결론 및 기대효과

시연

2.3 모델 2.3.2 EfficientNetB0

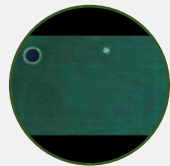
2. 중복 처리 알고리즘

01



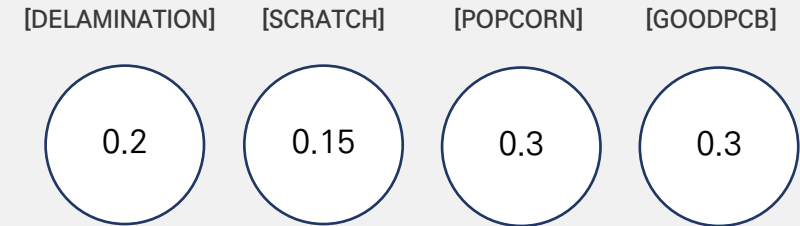
Threshold : 0.5

불량의 확률 값중 하나라도 임계 값을 초과



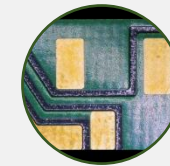
[DELAMINATION]

02



Threshold : 0.5

모든 불량클래스의 확률 값 < 임계 값



[GOODPCB]

- 하나의 이미지에 대한 불량 클래스 예측 확률 중 임계 값 0.5를 넘는 확률을 해당 이미지의 불량 클래스로 판단
- 임계 값을 넘긴 불량 클래스가 여러 개 존재할 경우 중복 불량으로 판단
- 정상 클래스에 대한 예측 값이 임계 값을 넘지 못하더라도, 불량 클래스에 대한 예측 값이 모두 임계 값을 넘지 못했을 경우 정상으로 판단

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

2.3 모델 2.3.3 XAI: Grad-CAM

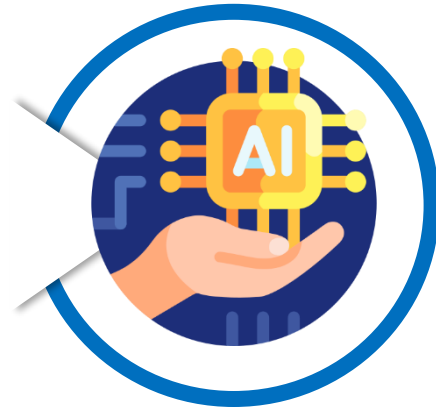
연구 개요

연구 방법

연구 결과

결론 및 기대효과

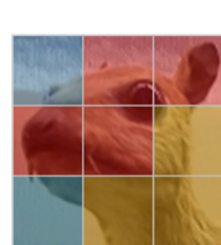
시연



XAI(eXplainable Artificial Intelligence)란?

설명 가능한 인공지능으로, 모델의 의사결정 분기점이 기하급수적으로 늘어남에 따라 큰 성능의 개선이 이루어지고 있는 반면, 왜 그렇게 판단했는지 그 결과에 대한 해석이 어려운 실정입니다. 모델의 블랙박스를 이해 가능한 수준으로 분해하여 연구자의 해석 가능성 및 신뢰도 제고를 도움

1. Grad-CAM



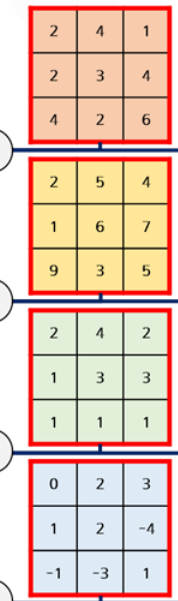
Pixel-wise sum



Heatmap



Feature map



AVG

3

AVG

5

AVG

2

AVG

1

 w_1^c w_2^c w_3^c w_4^c

Dog

 S^c

Camel

Cat

Cow

- 이미지를 분류하는 CNN 기반의 AI 모델에서 특정 클래스를 분류할 때 중요하게 고려한 영역을 표시해주는 것
- 마지막 전연결층이 존재하더라도 출력에 대한 특정 층의 Gradient를 계산하면 CAM을 얻을 수 있도록 기존 CAM의 단점을 개선

2.3 모델

2.3.3 XAI: Grad-CAM

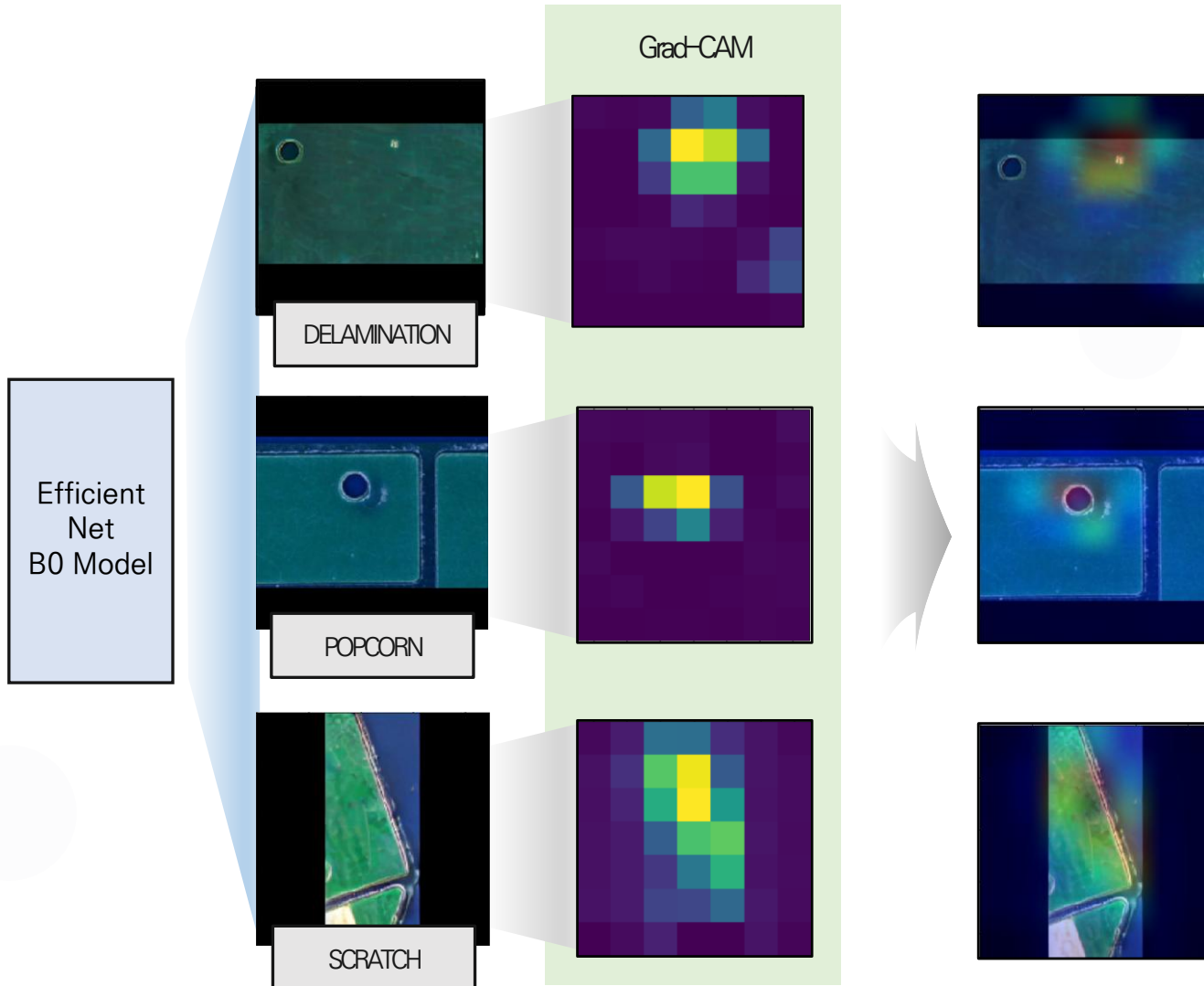
연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연



이미지 분류 모델이
분류 시 중점을 둔 부분 확인
→ 분류 결과 해석에 도움

→ 시각적으로 이해하고 평가한 뒤,
AI 모델 성능 향상을 위해
지도 및 불량 원인 추적 가능

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

3. 연구 결과 Research Results

3.1 연구 결과

3.1 연구 결과

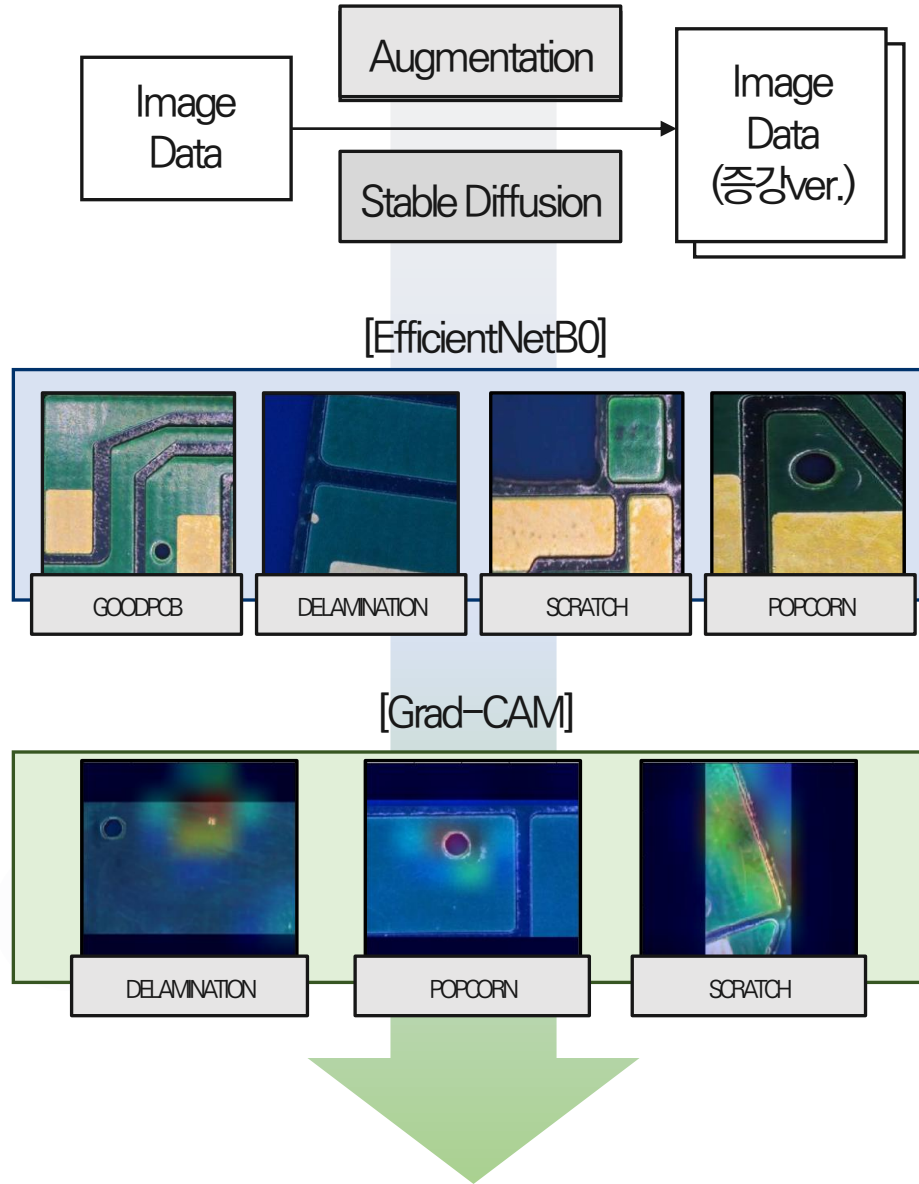
연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연



1. 평가지표 - Accuracy

Class	EfficientNet
DELAMINATION	92.76%
POPCORN	90.92%
SCRATCH	87.76%
GOODPCB	96.44%

4. 결론 및 기대효과

Conclusion & Expected Effects

4.1 결론

4.2 기대 효과

4.1 결론 및 기대효과

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

PCB 인쇄 회로 기판을
빠른 속도로
정상과 불량을 분류하는
자동 분류 AI 시스템

- ✓ 소수의 양/불 데이터를 이용해 라벨링되지 않은 양/불 데이터를 높은 정확도로 빠르게 분류
- ✓ 딥러닝 모델을 통해 이미지를 실시간으로 분류하여 불량 제품을 식별하고 자동으로 거부하는 시스템 구축

① 불량 원인 파악 및 추적

단순 양/불 데이터를 이진 분류하는 것에서 나아가 불량 패턴별로 분류하고, XAI기법을 이용하여 시각화한 구체적인 불량 부분을 파악함으로써 불량 발생 빈도가 높은 영역을 도출하여 그 부분에 대한 공정 에 문제가 있음을 파악하고 개선

② 자동화를 통한 생산성 향상

대량으로 생산되는 PCB 제품에 대해 업무를 자동화함으로써 다른 핵심적인 업무에 기업 역량을 집중하여 비효율적인 자원 운용을 방지하고 높은 신뢰성에 따른 수출 향상에 기여

③ 데이터 증강을 통한 정확도 확보

제조 공정 특성상 양품 이미지 대비 불량 이미지가 적을 수밖에 없으므로 불균형한 데이터셋을 가지는 문제를 본 연구에서 제안하는 방법을 통해 해결해 학습에 적합한 형태로 만들어 높은 정확도를 확보

④ 명확한 불량품 판단 기준 제시

기존에는 전문가의 경험에 의존하는 것이 필수적이기 때문에 기업마다 방식이 다르고 판정에 대한 정형화된 수치가 없어 일관적인 품질 관리에 어려움이 있는 점을 해결하여 딥러닝 기반의 불량 분류 모델로 보다 객관적인 분류 기준을 제안

4.2 향후 연구

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

01 신규 불량 패턴 탐지 모델 구현

스크래치, 박리, 들뜸 외에도 존재할 수 있는 신규 불량 패턴 역시 자동으로 분류할 수 있도록 하는 모델을 개발

02 모델의 범용성 확대

다른 제품군에 있어서 불량 판별에 적합하도록 모델의 파라미터, 구조 등의 수정을 통해 활용 영역 확대

03 제조 현장에 적용하도록 경량화

메모리 사용량과 연산 시간을 감소시켜 실시간으로 빠르게 분류가 이루어져야 하는 실제 공정에 적용될 수 있도록 경량화

04 불량 영역 시각화 정확도 개선

불량 패턴의 정확한 위치 정보를 파악하기 위해 Grad-CAM 뿐만 아니라 추가적인 후처리 과정 수행하여 방법론 강화

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

5. 시연 Test

5.1 시연

5.1 시연

연구 개요

연구 방법

연구 결과

결론 및 기대효과

시연

The screenshot shows a Jupyter Notebook titled "서버 코드_파이썬_제출" (Server Code_Python_Submission) with a "TRUSTED" badge and "Kernel o". The notebook contains the following code cells:

```

In [ ]: pip install tensorflow==2.13.0

In [ ]: pip install torch

In [ ]: pip install numpy==1.17

In [ ]: pip install cython

In [ ]: pip install opencv-python-headless

학습 모델 불러오기

In [ ]: from tensorflow.keras.models import load_model

        model_folder_path = '.data/'

        model = load_model(model_folder_path + 'efficientnet_model.h5')

Efficientnet 모델

In [ ]: import os
        import numpy as np
        from tensorflow.keras.preprocessing.image import load_img, img_to_array
        from tensorflow.keras.applications import EfficientNetB0
  
```

딥러닝을 이용한 XAI 기반 PCB 불량 탐지 및 분류

XAI-based PCB defect detection and classification using deep learning

감사합니다