

Lab Seminar

Intelligent Information Processing Lab

2023. 07. 12

이 새 봄

Contents

01 Previous Work

02 Work In Progress

03 Work In Future

01

Previous Work

- Citrus/Kiwi Disease Classification Service System
- Efficient Data Augmentation Method for Crop Disease

Citrus/Kiwi Disease Classification Service System

- 고품질 과수작물 통합 데이터를 바탕으로 Pre-trained 5개의 딥러닝 모델을 통해 병해충 분류 연구를 수행
- 고품질 과수작물 통합 데이터
 - ① 감귤, 키위 병해충을 인식하여 방제 계획을 수립할 수 있는 학습용 데이터 및 모델 구축
 - ② Pre-trained model : VGGNet, ResNet, DenseNet, EfficientNet, ViT
 - ③ AI-hub에서 데이터를 다운받을 수 있으며, 실제 연구 모델은 Docker 제공 + TTA 검증 완

Dataset (1/3)

- 고품질 과수작물 데이터 : 제주지역을 대상으로 감귤과 키위에 대해 고품질 과수작물 AI 학습용 데이터를 구축
- 감귤 : 20,000장, 6개의 클래스로 구성,



Citrus Fruit Normal



Citrus Fruit XCC



Citrus Leaf Red Mite



Citrus Leaf Normal



Citrus Leaf XCC



Citrus Leaf Aphid

Dataset (2/3)

- 고품질 과수작물 데이터 : 제주지역을 대상으로 감귤과 키위에 대해 고품질 과수작물 AI 학습용 데이터를 구축
- 키위 : 20,000장, 5개의 클래스로 구성,



Kiwi fruit bacterial soft rot



Kiwi fruit normal



Kiwi leaf thysanoptera



Kiwi leaf normal



Kiwi leaf spot

Dataset (3/3)

1. Kiwi / Citrus : AI Hub, 국립원예특작과학원 → Dataset : Training : Test = 7: 3 / Training : Validation : Test = 7 : 2: 1

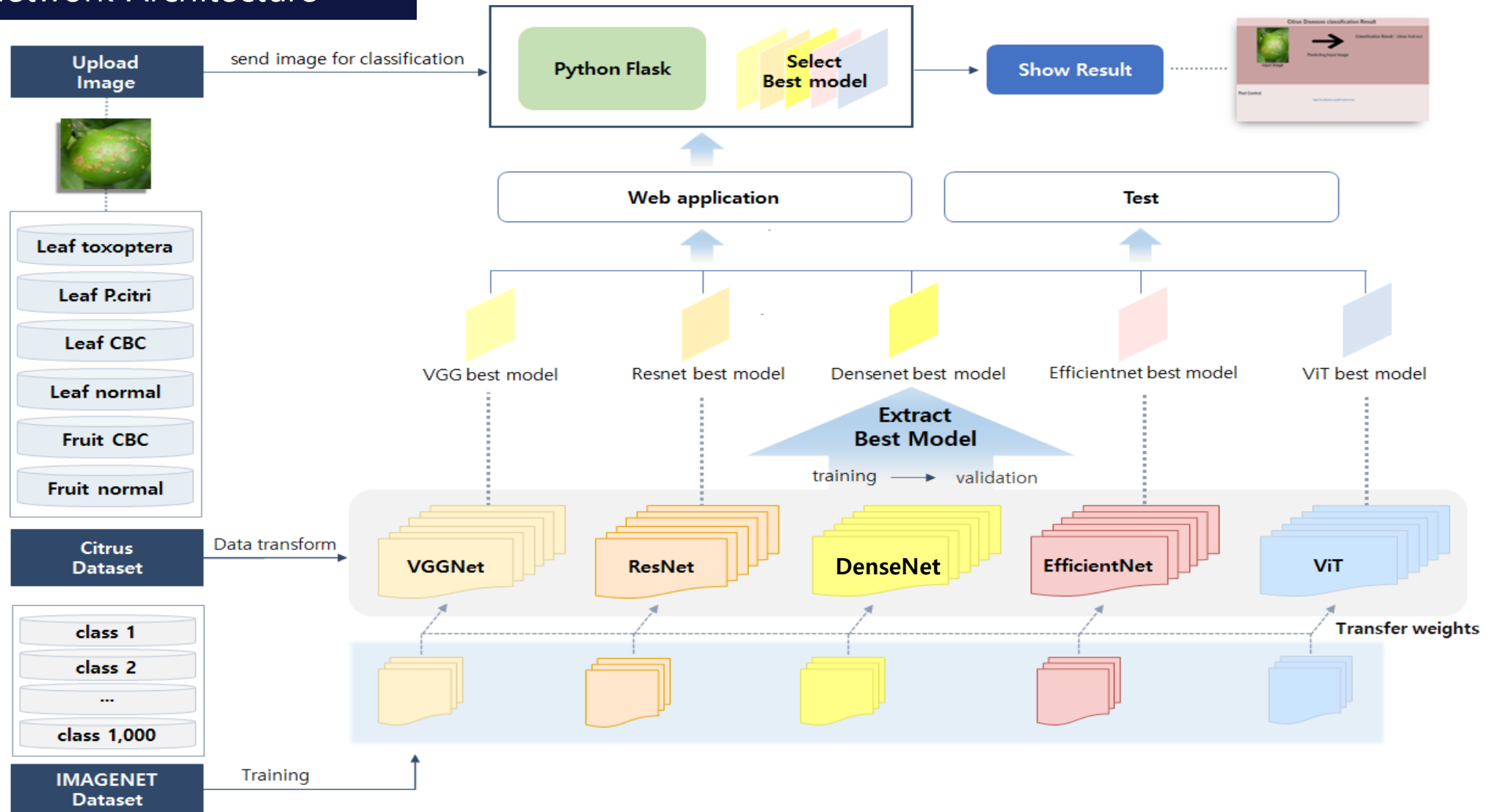
Kiwi disease dataset

| | Disease Type | Total |
|-------|-------------------------------|--------|
| 1 | Kiwi fruit healthy | 2,124 |
| 2 | Kiwi fruit bacterial soft rot | 1,737 |
| 3 | Kiwi leaf healthy | 2,876 |
| 4 | Kiwi leaf thysanoptera | 5,585 |
| 5 | Kiwi leaf spot | 7,678 |
| Total | | 20,000 |

Citrus disease dataset

| | Disease Type | Total |
|-------|----------------------------------|--------|
| 1 | Citrus fruit healthy | 2,545 |
| 2 | Citrus fruit CBC | 1,716 |
| 3 | Citrus leaf healthy | 2,455 |
| 4 | Citrus leaf CBC | 9,552 |
| 5 | Citrus leaf panonychus. citir | 1,814 |
| 6 | Citrus leaf toxoptera. citricida | 1,918 |
| Total | | 20,000 |

Network Architecture



Experiment Result

- 딥러닝 기반 과수작물 분류 모델 Validation 실험 결과 (Stratified K-fold cross-validation / k=5)

| Disease | Citrus | | Kiwi | |
|--------------|----------|----------|--------------------|--------------------|
| Model | F1 score | Accuracy | F1 score | Accuracy |
| VGGNet16 | 96.7 | 97.7 | 97.1 ± 0.6442 | 97.7 ± 0.1581 |
| ResNet50 | 97.6 | 98.3 | 96.46 ± 0.497 | 97.56 ± 0.2074 |
| DenseNet161 | 97.7 | 98.4 | 97.18 ± 0.5263 | 97.72 ± 0.3271 |
| EfficientNet | 98.2 | 98.8 | 90.46 ± 0.7162 | 91.88 ± 0.8556 |
| ViT | 98.2 | 98.8 | 97.34 ± 0.2881 | 98.18 ± 0.1924 |

Experiment Result

- 딥러닝 기반 과수작물 분류 모델 Test 실험 결과

| Disease | Citrus | | Kiwi | |
|--------------|----------|----------|-------------------|--------------------|
| Model | F1 score | Accuracy | F1 score | Accuracy |
| VGGNet16 | 97 | 97.9 | 97.8 ± 0.2074 | 99.06 ± 0.1161 |
| ResNet50 | 98 | 98.6 | 98.6 ± 0.1517 | 98.28 ± 0.1215 |
| DenseNet161 | 98.4 | 98.8 | 98.9 ± 0.0837 | 99.39 ± 0.0673 |
| EfficientNet | 98.6 | 99 | 98.7 ± 0.1789 | 99.16 ± 0.0811 |
| ViT | 97.5 | 98.1 | 98.7 ± 0.0837 | 99.17 ± 0.0778 |

Conclusion



Article

Automatic Classification Service System for Citrus Pest Recognition Based on Deep Learning

Saebom Lee ¹, Gyuho Choi ¹, Hyun-Cheol Park ² and Chang Choi ^{1,*}

¹ Department of Computer Engineering, Gachon University, Sujeong-gu, Seongnam-si 461-701, Gyeonggi-do, Republic of Korea

² Department of AI Software, Gachon University, Sujeong-gu, Seongnam-si 461-701, Gyeonggi-do, Republic of Korea

* Correspondence: changchoi@gachon.ac.kr

Abstract: Plant diseases are a major cause of reduction in agricultural output, which leads to severe economic losses and unstable food supply. The citrus plant is an economically important fruit crop grown and produced worldwide. However, citrus plants are easily affected by various factors, such as climate change, pests, and diseases, resulting in reduced yield and quality. Advances in computer vision in recent years have been widely used for plant disease detection and classification, providing opportunities for early disease detection, and resulting in improvements in agriculture. Particularly, the early and accurate detection of citrus diseases, which are vulnerable to pests, is very important to prevent the spread of pests and reduce crop damage. Research on citrus pest disease is ongoing, but it is difficult to apply research results to cultivation owing to a lack of datasets for research and limited types of pests. In this study, we built a dataset by self-collecting a total of 20,000 citrus pest images, including fruits and leaves, from actual cultivation sites. The constructed dataset was trained, verified, and tested using a model that had undergone five transfer learning steps. All models used in the experiment had an average accuracy of 97% or more and an average f1 score of 96% or more. We built a web application server using the EfficientNet-b0 model, which exhibited the best performance among the five learning models. The built web application tested citrus pest disease using image samples collected from websites other than the self-collected image samples and prepared data, and both samples correctly classified the disease. The citrus pest automatic diagnosis web system using the model proposed in this study plays a useful auxiliary role in recognizing and classifying citrus diseases. This can, in turn, help improve the overall quality of citrus fruits.

Keywords: agriculture; citrus disease classification; deep learning; web application



Citation: Lee, S.; Choi, G.; Park, H.-C.; Choi, C. Automatic Classification Service System for Citrus Pest Recognition Based on Deep Learning. *Sensors* **2022**, *22*, 8911. <https://doi.org/10.3390/s22228911>

Academic Editor: Yongwha Chung

- Journal : MDPI Sensors
- IF : 3.9
- Average JIF : Top 32.1%
- Status : Published 2022-11-18
- Kiwi : 임유리와 함께 작업 후 공동 1저자로 넣을 예정

01

Previous Work

- Citrus/Kiwi Disease Classification Service System
- Efficient Data Augmentation Method for Crop Disease

Efficient Data Augmentation Method for Crop Disease

- 컴퓨터 시스템의 발전으로 딥러닝을 적용한 다양한 농작물 병해충 인식 연구가 이루어지고 있음
- 현재 수행된 농작물 병해충 인식 연구는 다음과 같은 한계가 존재함
 - ① 데이터셋의 크기가 작고, 병해충의 감염된 농작물의 종류가 한정적임
 - ② 데이터 전처리 단계에서 grayscale과 수평 · 수직 뒤집기 등의 간단한 기하학적 데이터 증강 방법을 사용하여, 다양한 패턴을 갖고 있는 병해의 특징을 효율적으로 추출하지 못함

다양한 종류의 병해충 패턴들



① 갈변



② 점무늬



③ 가는 실을 곤 형태



④ 해충

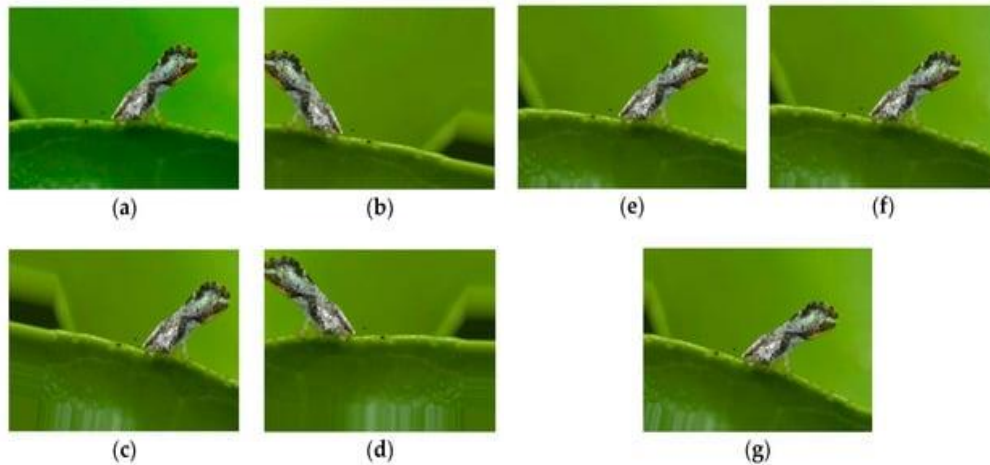
연구 목적

농작물 병해충을 분류를 위해 병해충에 발생한 패턴을 효과적으로 추출할 수 있는 데이터 증강 방법이 필요함

Related Work (1)

1. (2019) Citrus pests and diseases recognition model using weakly dense connected convolution network

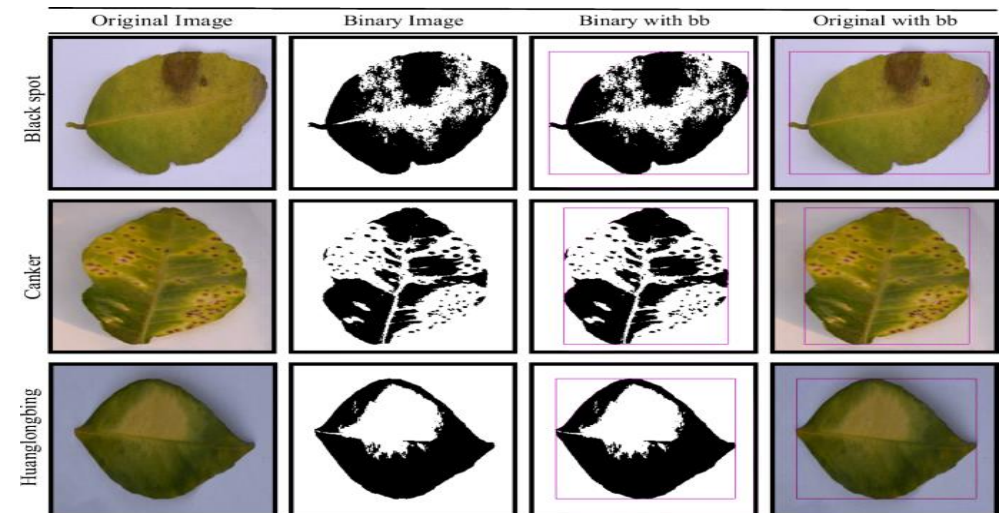
→ Random rotation, reflection, shift, and flip



Related Work (2)

2. (2022) Citrus disease detection and classification using end-to-end anchor-based deep learning model

→ 모든 이미지를 grayscale로 변경



Dataset

1. Kiwi / Citrus : AI Hub, 국립원예특작과학원

Kiwi disease dataset

| | Disease Type | Total |
|-------|-------------------------------|--------|
| 1 | Kiwi fruit healthy | 2,124 |
| 2 | Kiwi fruit bacterial soft rot | 1,737 |
| 3 | Kiwi leaf healthy | 2,876 |
| 4 | Kiwi leaf thysanoptera | 5,585 |
| 5 | Kiwi leaf spot | 7,678 |
| Total | | 20,000 |

Citrus disease dataset

| | Disease Type | Total |
|-------|----------------------------------|--------|
| 1 | Citrus fruit healthy | 2,545 |
| 2 | Citrus fruit CBC | 1,716 |
| 3 | Citrus leaf healthy | 2,455 |
| 4 | Citrus leaf CBC | 9,552 |
| 5 | Citrus leaf panonychus. citir | 1,814 |
| 6 | Citrus leaf toxoptera. citricida | 1,918 |
| Total | | 20,000 |

Dataset

2. Pepper bell / Potato / Tomato : PlantVillage

PlantVillage dataset

| | Disease Type | Total |
|---|--------------------------------|-------|
| 1 | Pepper bell leaf bacteria spot | 997 |
| 2 | Pepper bell leaf healthy | 1,478 |
| 3 | Potato leaf early blight | 1,000 |
| 4 | Potato leaf late blight | 1,000 |
| 5 | Potato leaf healthy | 152 |
| 6 | Tomato leaf target spot | 1,404 |
| 7 | Tomato leaf early blight | 1,000 |
| 8 | Tomato leaf late blight | 1,909 |

| | Disease Type | Total |
|-------|---------------------------|--------|
| 9 | Tomato leaf mold | 952 |
| 10 | Tomato leaf Septoria spot | 1,771 |
| 11 | Tomato leaf spider mites | 1,676 |
| 12 | Tomato leaf mosaic virus | 373 |
| 13 | Tomato leaf yellow virus | 3,209 |
| 14 | Tomato leaf healthy | 1,591 |
| 15 | Tomato leaf bacteria spot | 2,127 |
| Total | | 20,639 |

Imbalanced Data Problem

Dataset : Training : Test = 7: 3 (Stratified K-fold cross-validation / k=5)

Kiwi disease dataset used in the experiment

| | Disease Type | Training | Test | Total |
|---|-------------------------------|----------|-------|-------|
| 1 | Kiwi fruit healthy | 1,698 | 426 | 2,124 |
| 2 | Kiwi fruit bacterial soft rot | 1,389 | 348 | 1,737 |
| 3 | Kiwi leaf healthy | 2,300 | 576 | 2,876 |
| 4 | Kiwi leaf thysanoptera | 4,467 | 1,118 | 5,585 |
| 5 | Kiwi leaf thysanoptera | 6,142 | 1,536 | 5,585 |

Citrus disease dataset used in the experiment

| | Disease Type | Training | Test | Total |
|---|----------------------------------|----------|-------|-------|
| 1 | Citrus fruit healthy | 2,035 | 510 | 2,545 |
| 2 | Citrus fruit CBC | 1,372 | 344 | 1,716 |
| 3 | Citrus leaf healthy | 1,965 | 490 | 2,455 |
| 4 | Citrus leaf CBC | 7,642 | 1,910 | 9,552 |
| 5 | Citrus leaf panonychus. citir | 1,452 | 362 | 1,814 |
| 6 | Citrus leaf toxoptera. citricida | 1,534 | 384 | 1,918 |

Imbalanced Data Problem

Dataset : Training dataset : Test dataset = 7: 3



Class : 24 / Total images : 60,165

PlantVillage dataset in the experiment

| | Disease Type | Training | Test | Total |
|---|---------------------------|-----------|------|-------|
| 1 | Pepper bell bacteria spot | 779 → 800 | 200 | 1,000 |
| 2 | Pepper bell healthy | 1,182 | 296 | 1,478 |
| 3 | Potato early blight | 800 | 200 | 1,000 |
| 4 | Potato late blight | 800 | 200 | 1,000 |
| 5 | Potato healthy | X | | |
| 6 | Tomato target spot | 1,095 | 309 | 1,404 |
| 7 | Tomato early blight | 800 | 200 | 1,000 |
| 8 | Tomato late blight | 1,555 | 354 | 1,909 |

| | Disease Type | Training | Test | Total |
|-------|--------------------------|-----------|------|--------|
| 9 | Tomato leaf mold | 752 → 800 | 200 | 1,000 |
| 10 | Tomato Septoria spot | 1,432 | 339 | 1,771 |
| 11 | Tomato leaf spider mites | 1,319 | 357 | 1,676 |
| 12 | Tomato leaf mosaic virus | X | | |
| 13 | Tomato leaf yellow virus | 2,578 | 631 | 3,209 |
| 14 | Tomato leaf healthy | 1,269 | 322 | 1,591 |
| 15 | Tomato bacteria spot | 1,687 | 440 | 2,127 |
| Total | | | | 20,615 |

Noise Types

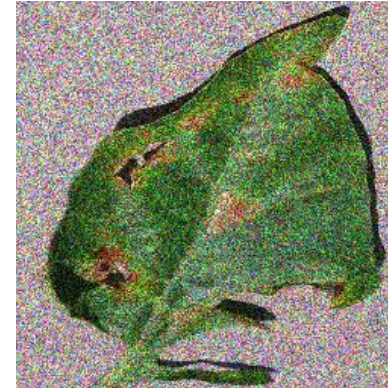
- Noise type : Impulse , Poisson , Gaussian , Uniform , Laplacian , **Multiplicative Gaussian**



(a) Original



(b) Impulse



(b) Poisson



(c) Gaussian



(d) Uniform



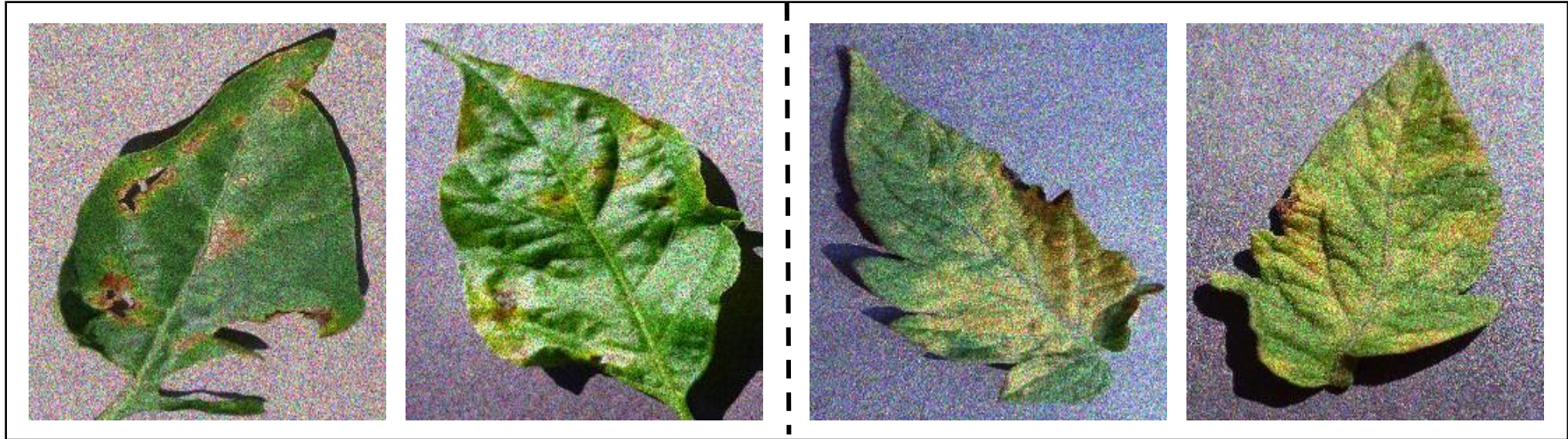
(e) Laplacian



(f) Multiplicative Gaussian

Image with added noise

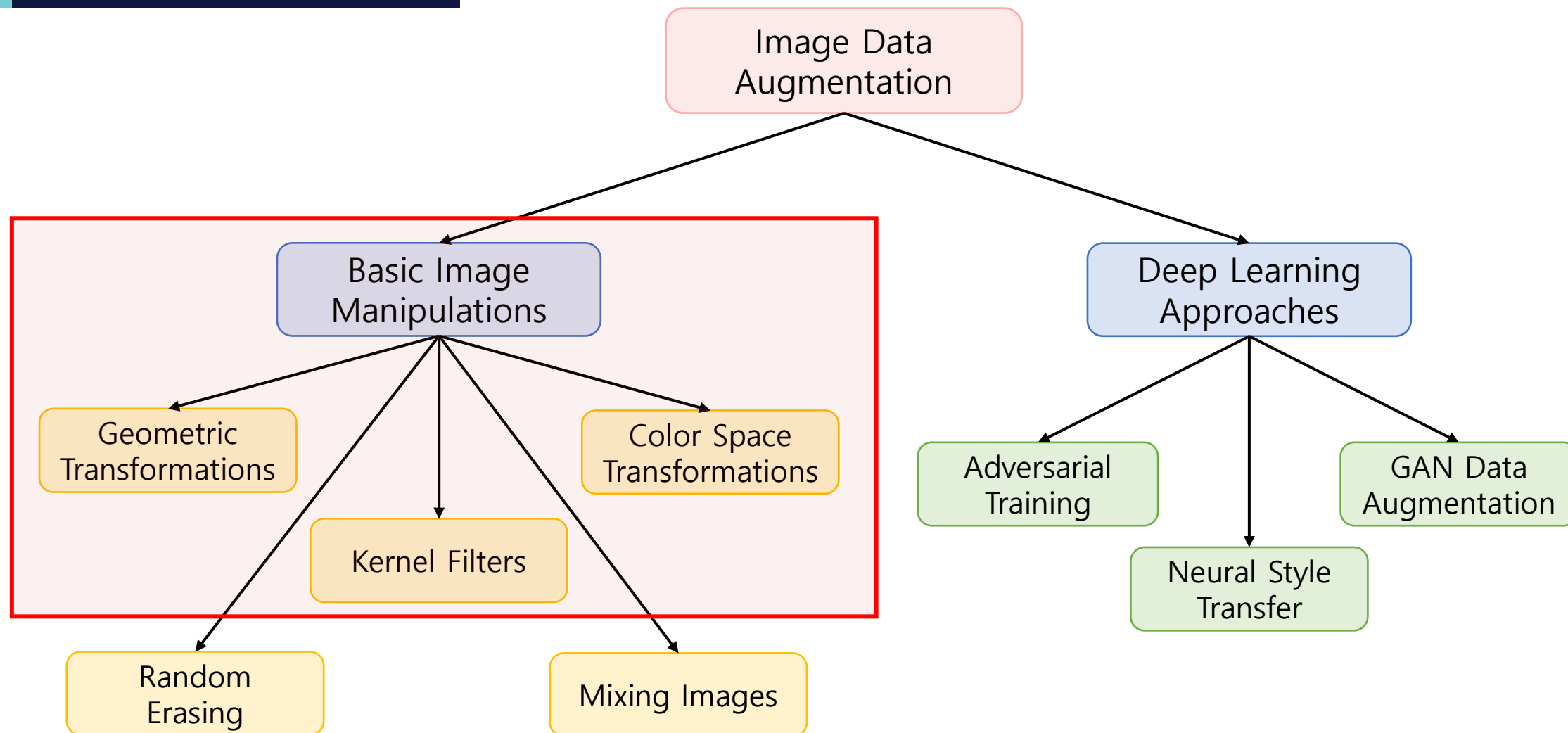
- Image with multiplicative Gaussian noise



(a) Pepper bell bacterial spot

(b) Tomato leaf mold

Data Augmentation Types



Sample image of crop disease (1)



(a) Kiwi fruit bacterial soft rot



(b) Kiwi fruit normal



(c) Kiwi leaf spot



(d) Pepper bell bacterial spot



(e) Pepper bell healthy



(f) Potato early blight



(g) Potato late blight



(h) Tomato target spot

Data augmentation methods applicable for experiments (1)



(a) Original



(b) Horizontal flip



(c) Vertical flip



(d) Rotation



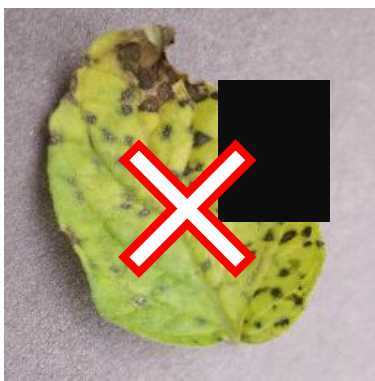
(e) Resize, crop



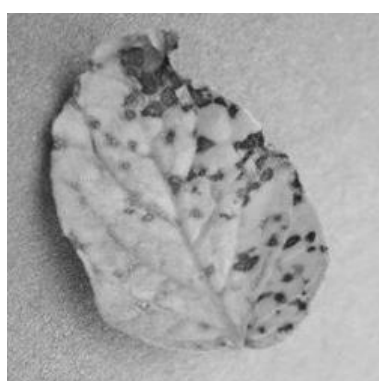
(f) Affine



(g) Perspective



(h) Erasing



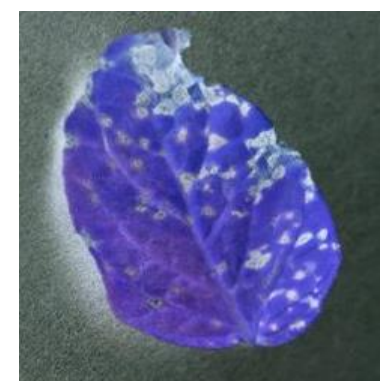
(i) Grayscale



(j) Gaussian blur



(k) Color jitter



(l) Invert

Sample image of crop disease (2)



(a) Kiwi fruit bacterial soft rot



(b) Kiwi fruit normal



(c) Kiwi leaf spot



(d) Pepper bell bacterial spot



(e) Pepper bell healthy



(f) Potato early blight



(g) Potato late blight



(h) Tomato target spot

Data augmentation methods applicable for experiments (1)



(a) Original



(b) Horizontal flip



(c) Vertical flip



(d) Rotation



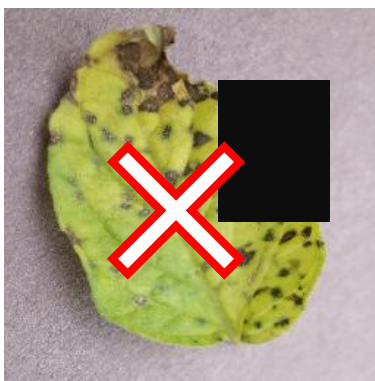
(e) Resize, crop



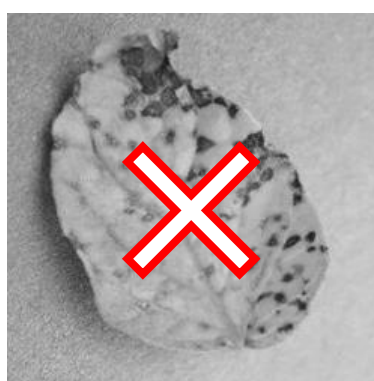
(f) Affine



(g) Perspective



(h) Erasing



(i) Grayscale



(j) Gaussian blur



(k) Color jitter



(l) Invert

Color Jitter

- 40% 이내에서 밝기, 대조, 채도 및 10% 이내에서 색조 적용 랜덤 적용



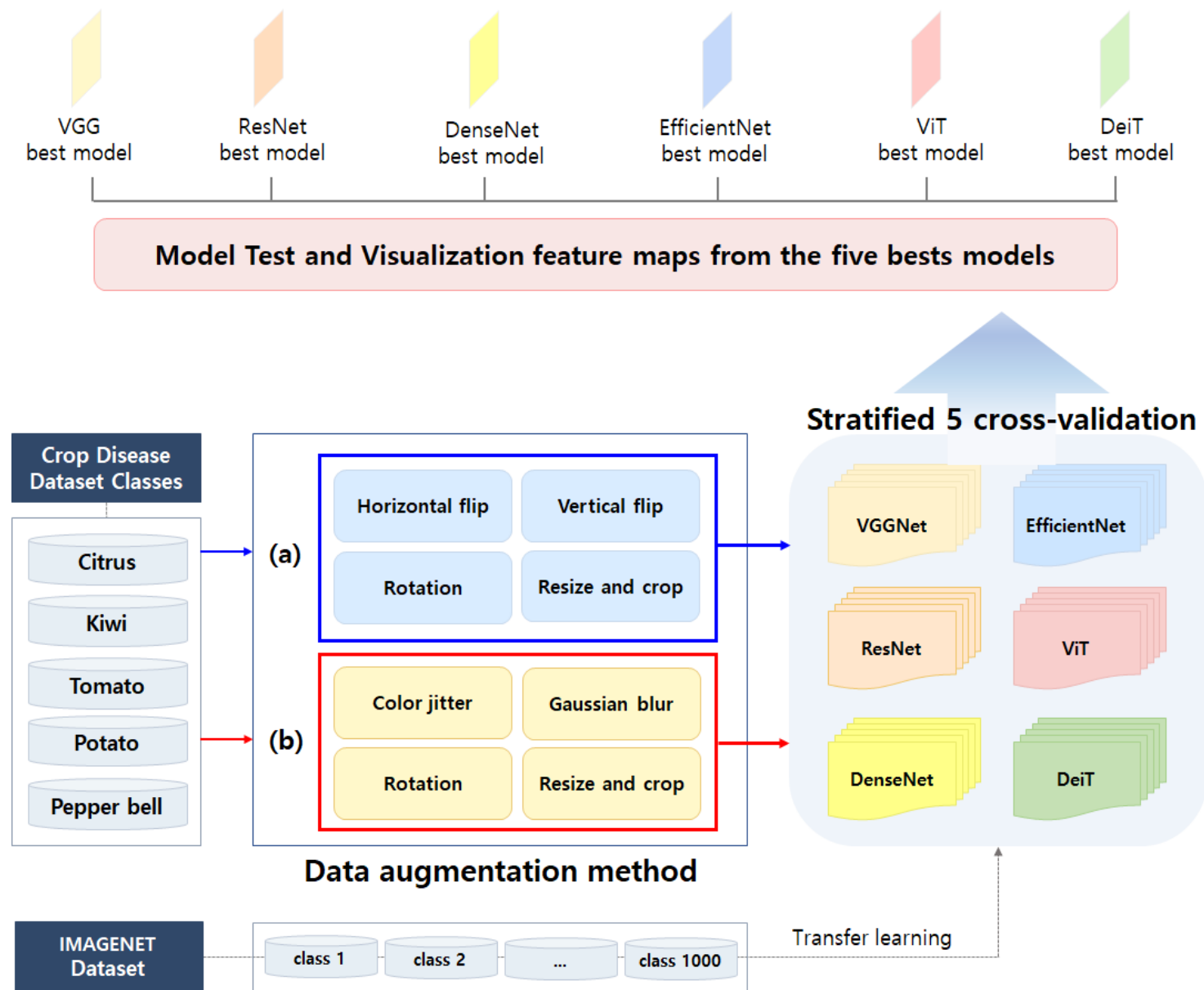
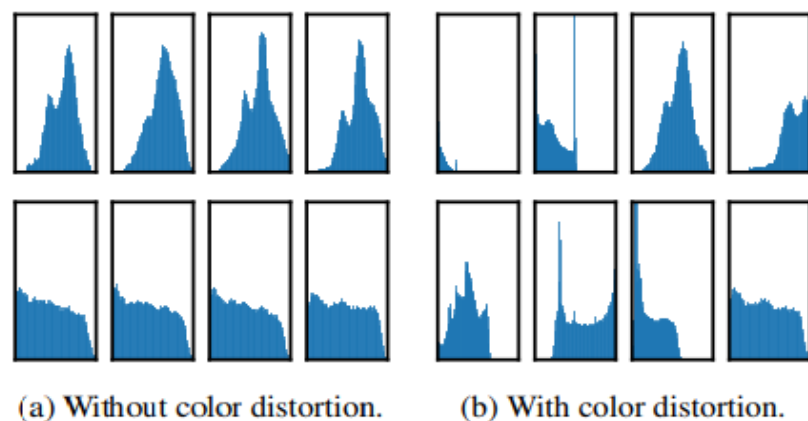
Original image



Color jitter 데이터 증강이 적용된 병해 이미지

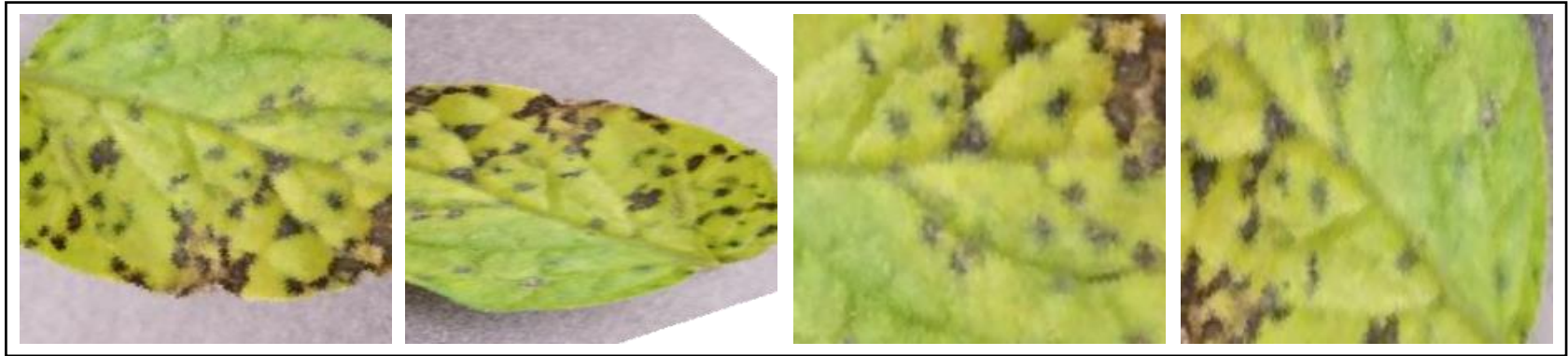
Network Architecture

- The overall workflow of the proposed network architecture



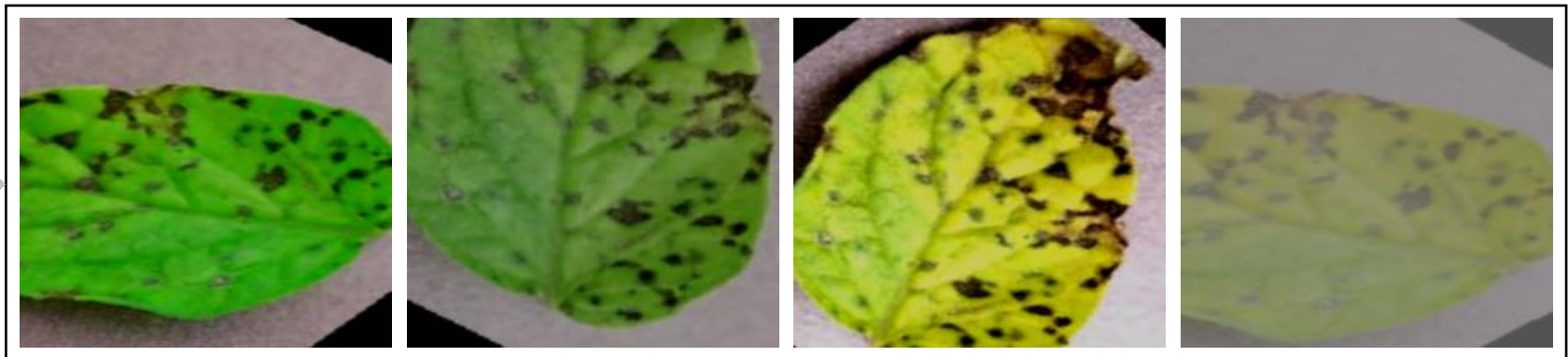
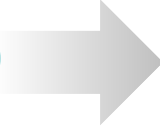
Sample image combined with data augmentation methods

a



horizontal flip + vertical flip + rotation + resize and crop

b



color jitter + vertical flip + rotation + resize and crop

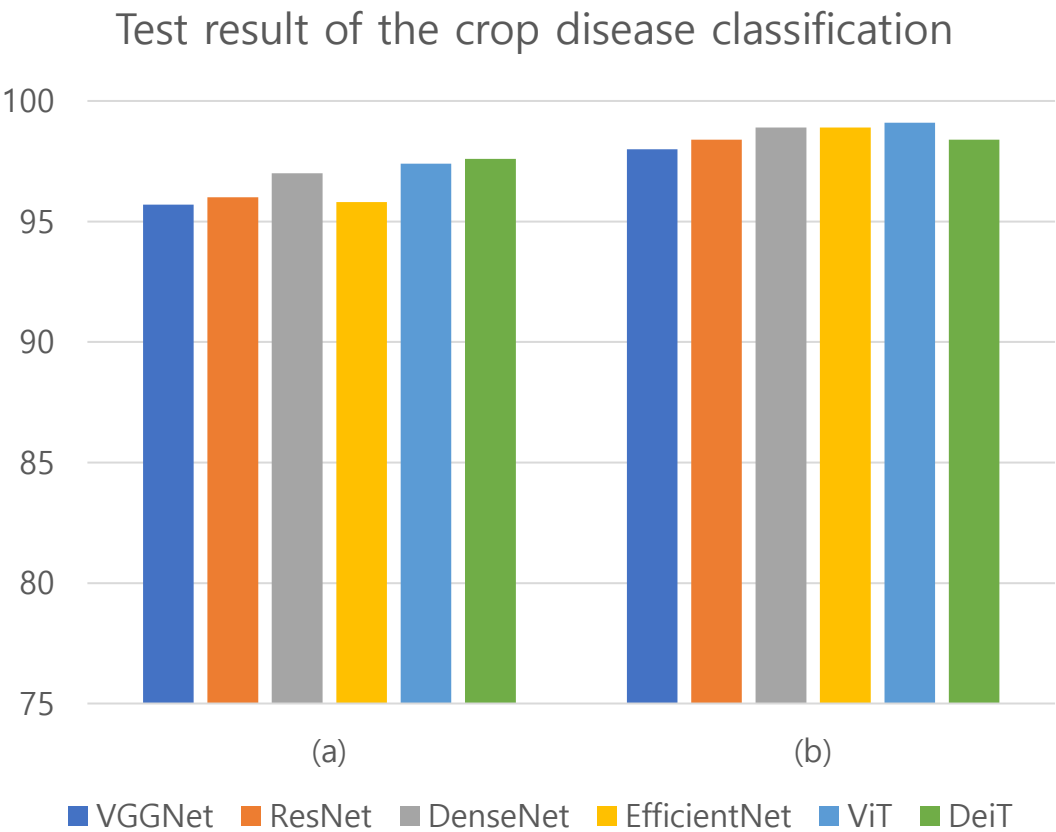
Experiment Result

- 딥러닝 기반 농작물 분류 모델 Validation 실험 결과 (Stratified K-fold cross-validation / k=5)

| Method | (a) | | | (b) | | |
|--------------|-------------------|--------------------|---------------|-------------------|--------------------|---------------|
| Model | F1 score | Accuracy | Training time | F1 score | Accuracy | Training time |
| VGGNet16 | 95.8 \pm 1.3024 | 97.47 \pm 0.0841 | 29 : 56 : 06 | 97.8 \pm 0.2074 | 99.06 \pm 0.1161 | 52 : 36 : 54 |
| ResNet50 | 96.2 \pm 0.3536 | 97.71 \pm 0.0727 | 23 : 05 : 57 | 98.6 \pm 0.1517 | 98.28 \pm 0.1215 | 57 : 14 : 01 |
| DenseNet161 | 97.5 \pm 0.2121 | 98.14 \pm 0.0857 | 50 : 28 : 07 | 98.9 \pm 0.0837 | 99.39 \pm 0.0673 | 59 : 11 : 22 |
| EfficientNet | 97.0 \pm 0.2345 | 97.4 \pm 0.1389 | 22 : 02 : 21 | 98.7 \pm 0.1789 | 99.16 \pm 0.0811 | 57 : 03 : 28 |
| ViT | 97.6 \pm 0.1095 | 98.16 \pm 0.0661 | 63 : 08 : 50 | 98.7 \pm 0.0837 | 99.17 \pm 0.0778 | 64 : 21 : 39 |
| DeiT | 95.5 \pm 0.0234 | 95.35 \pm 0.0287 | 75 : 38 : 36 | 95.9 \pm 0.024 | 95.64 \pm 0.029 | 75 : 47 : 28 |

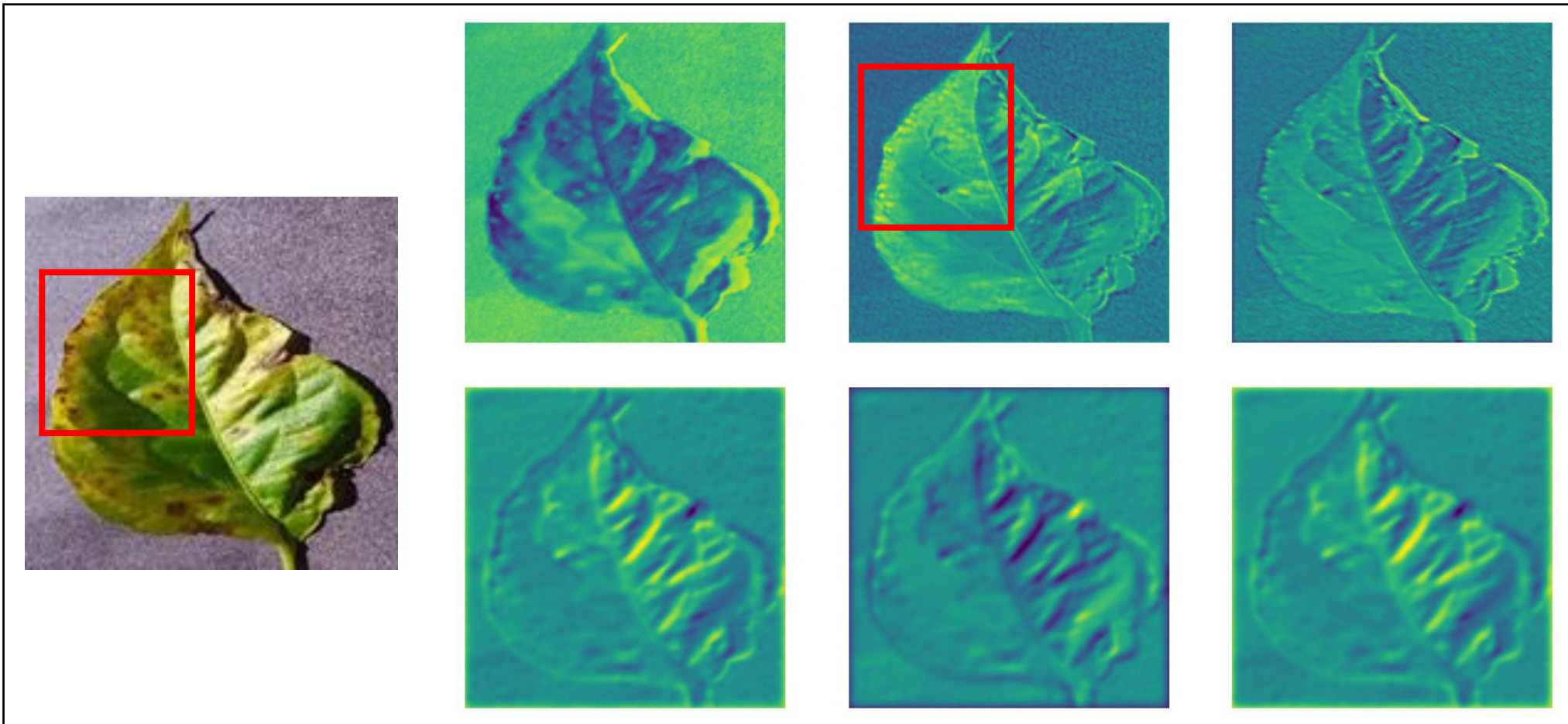
Experiment Result

- 딥러닝 기반 농작물 분류 모델 Test 실험 결과

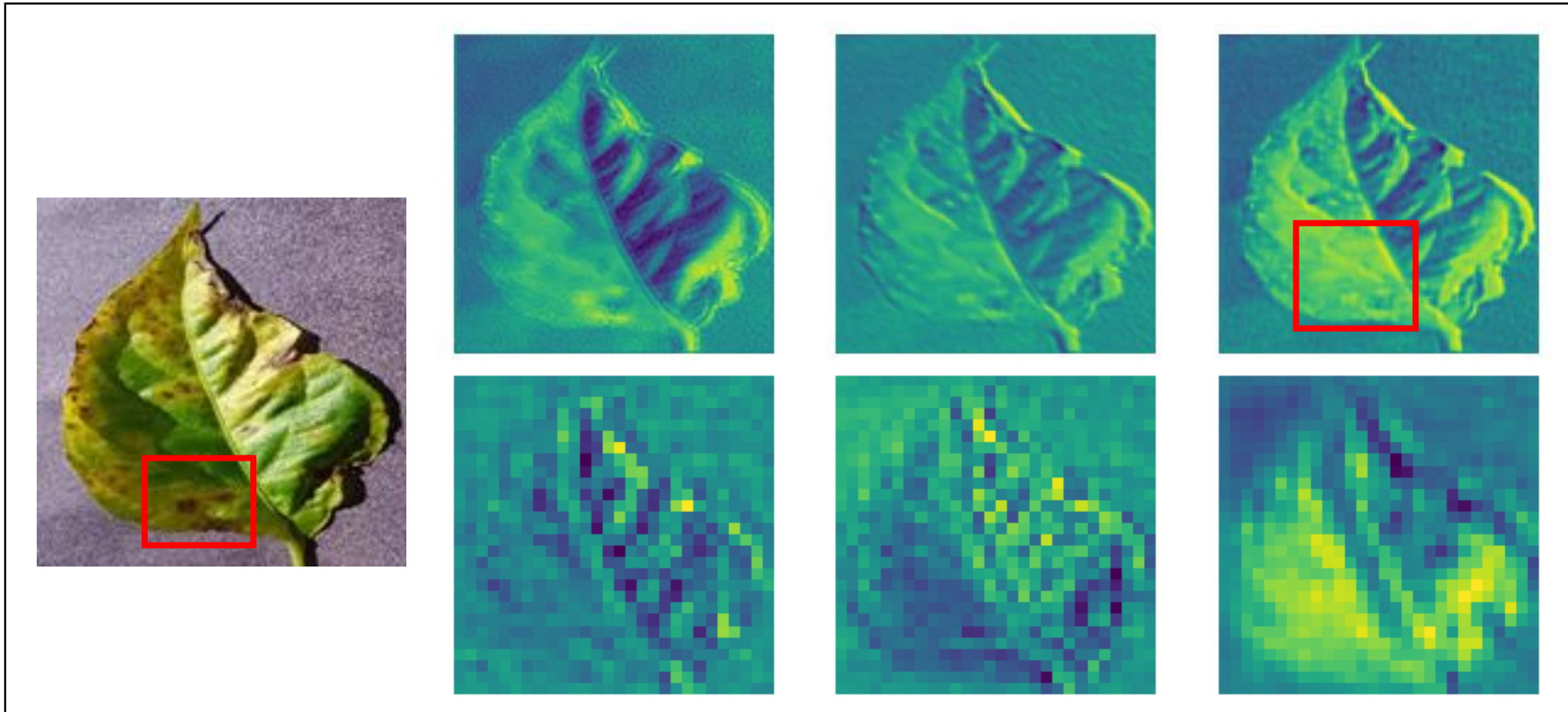


| Model | F1 score | Accuracy | Recall | Precision |
|--------------|----------|----------|--------|-----------|
| VGGNet16 | 95.7 | 97.3 | 96.59 | 95.86 |
| | 98 | 97.9 | 98.06 | 97.87 |
| ResNet50 | 96 | 96.9 | 96.9 | 96.07 |
| | 98.4 | 98.3 | 98.4 | 98.46 |
| DenseNet161 | 97 | 98 | 97.85 | 97.16 |
| | 98.9 | 98.8 | 97.85 | 97.16 |
| EfficientNet | 95.8 | 96.9 | 95.7 | 96.96 |
| | 98.9 | 98.7 | 98.8 | 98.92 |
| ViT | 97.4 | 98.2 | 97.42 | 98.45 |
| | 99.1 | 98.9 | 98.96 | 99.13 |
| DeiT | 97.6 | 98.37 | 97.61 | 98.52 |
| | 98.4 | 98.19 | 98.31 | 98.38 |

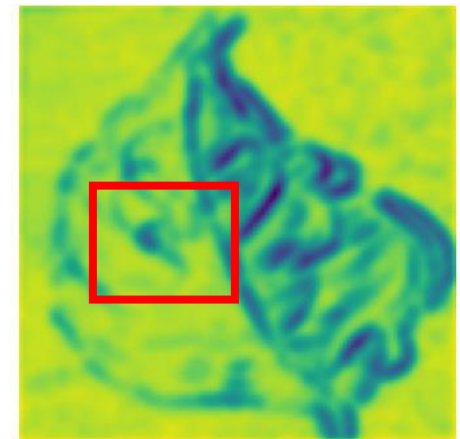
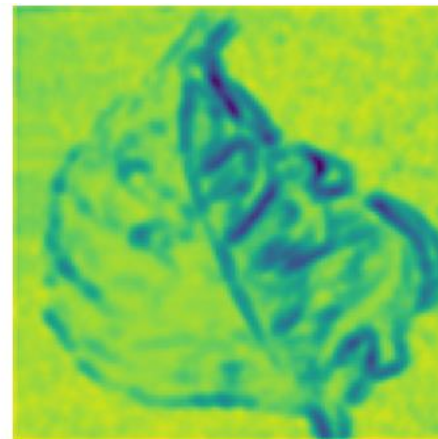
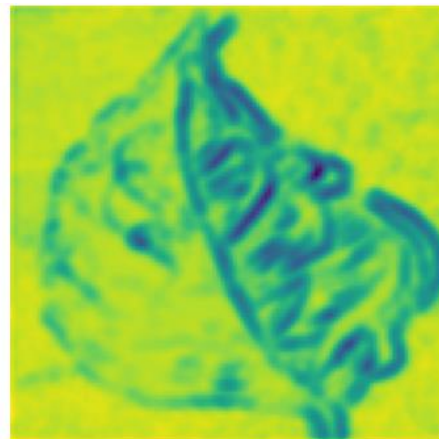
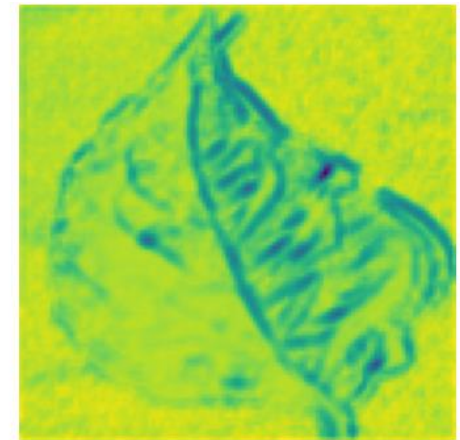
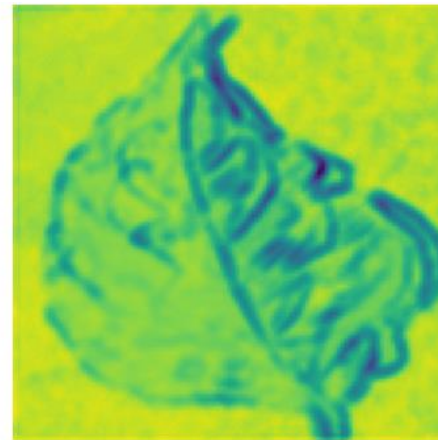
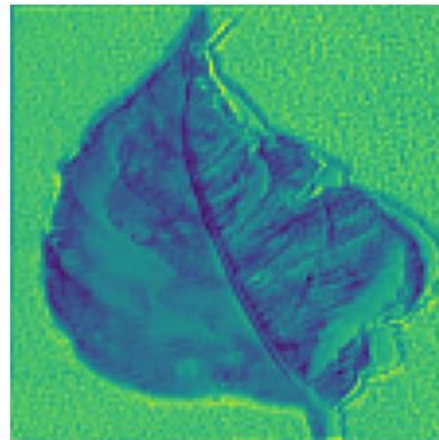
Feature map (VGGNet)



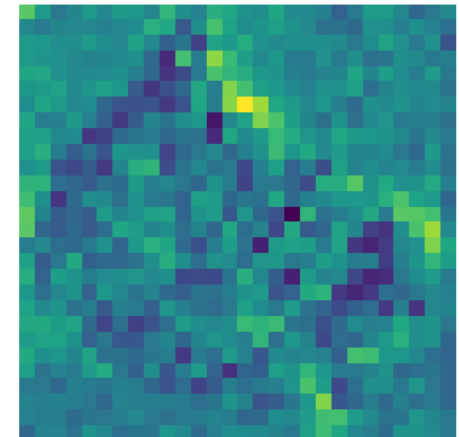
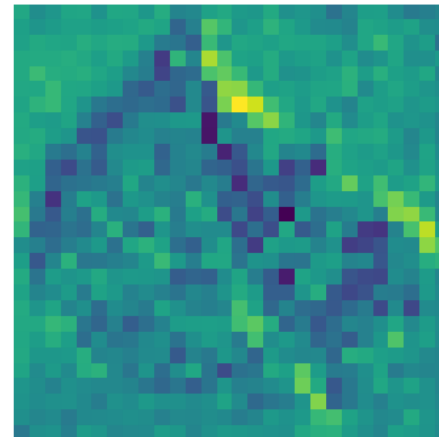
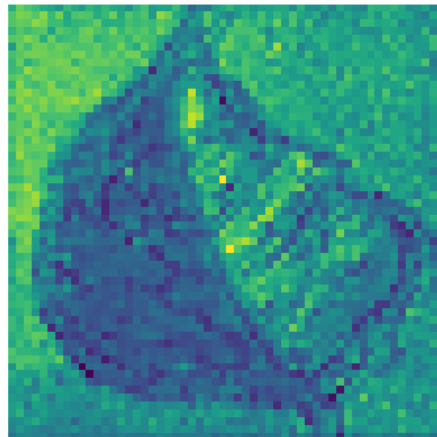
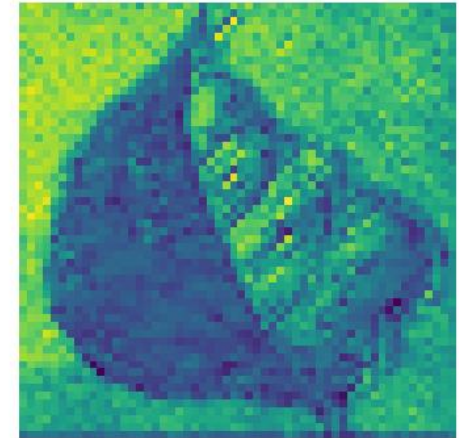
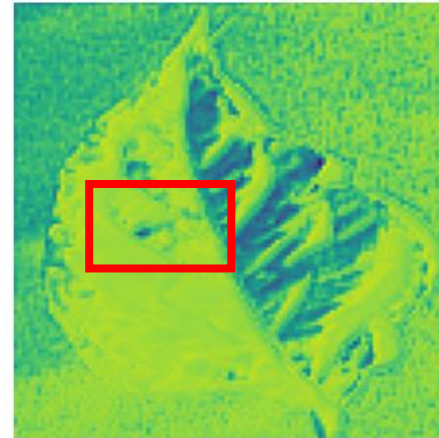
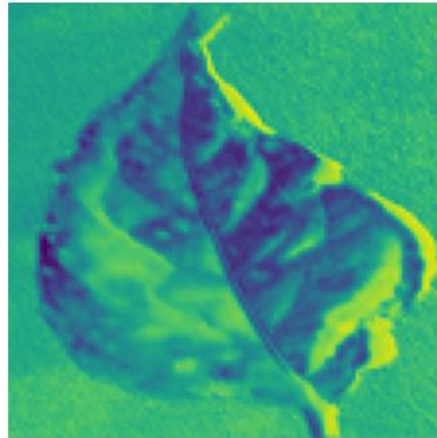
Feature map (ResNet)



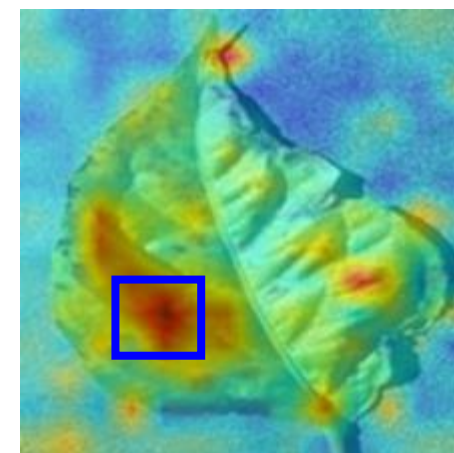
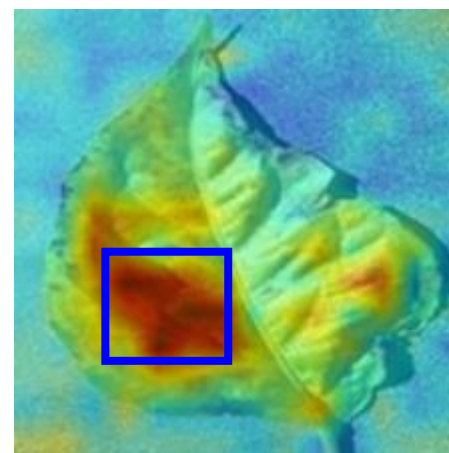
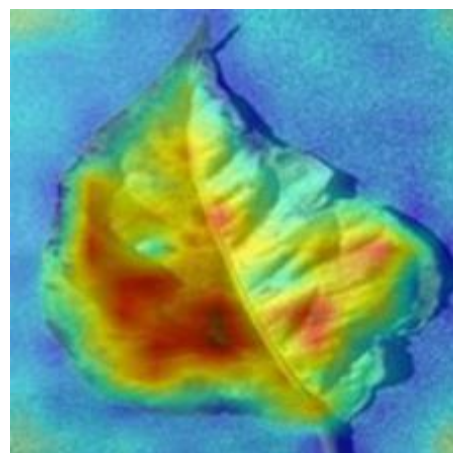
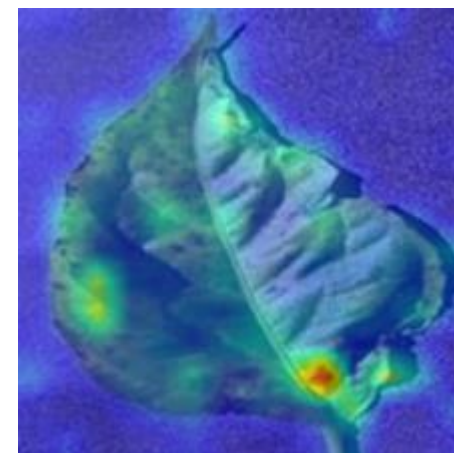
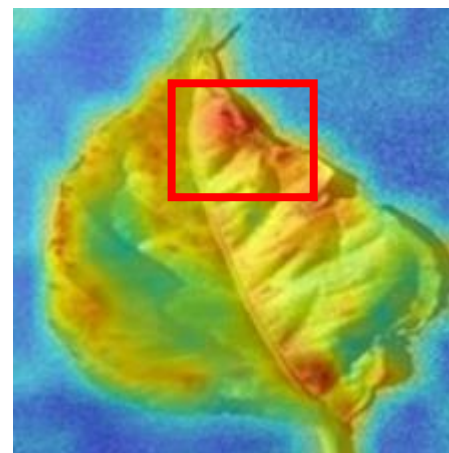
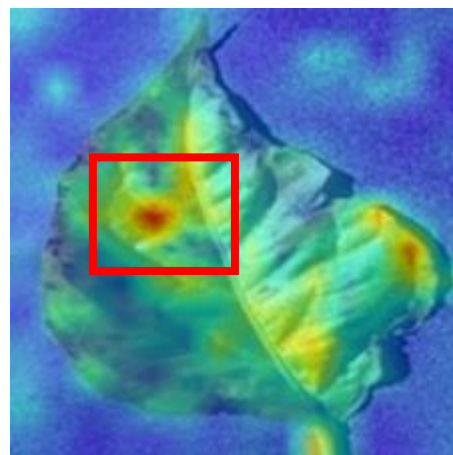
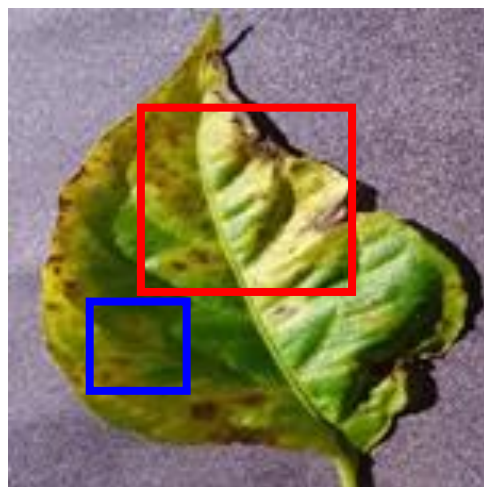
Feature map (DesnetNet)



Feature map (EfficientNet)



Feature map (ViT, DeiT)

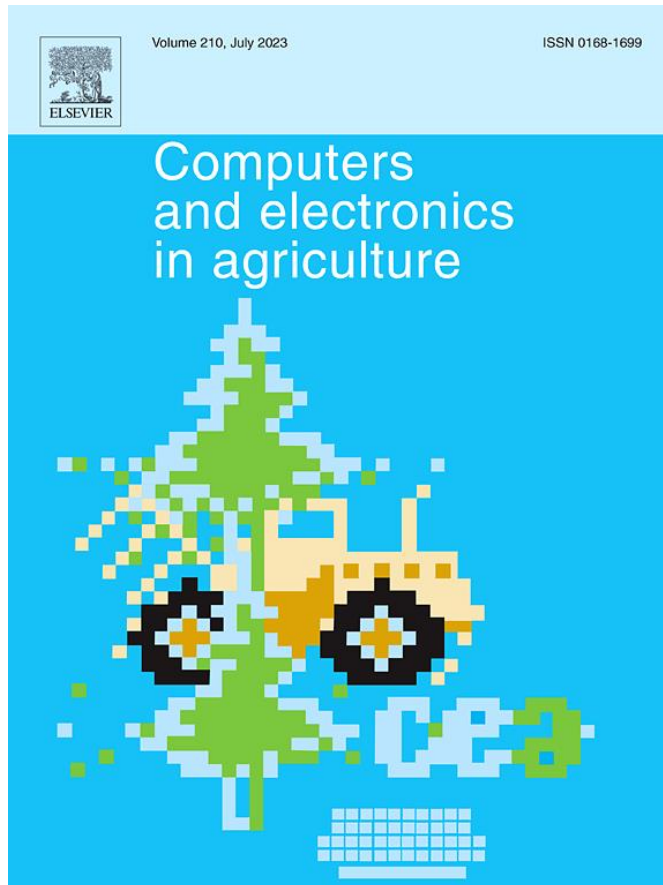


min

mean

max

Conclusion



- Computers and Electronics in Agriculture 투고 예정
: 농경학, 원예, 임업, 양식업을 포함하여 농업 문제를 해결하기 위한
컴퓨터 하드웨어, 소프트웨어, 전자 계측, 시스템의 개발 및 응용 분야
- Cite Score : 13.6 (Q1)
- IF : 8.3
- Average JIF : Top 7.5

02

Work In Progress

- Adversarial Attack for ECG user authentication

Adversarial Attack for ECG user Authentication

- 주제 : 적대적 공격을 통한 심전도 사용자인증 보안기술 취약점 분석
- 2023년 「여대학원생 공학연구팀제 지원사업」
- 연구책임자 : 이새봄
- 참여연구원 : 이명희(3학년), 나경민(3학년), 장예정(4학년)
- 연구기간 : 2023. 04. 01 ~ 2023. 10. 31 (8,000,000원)

Introduction

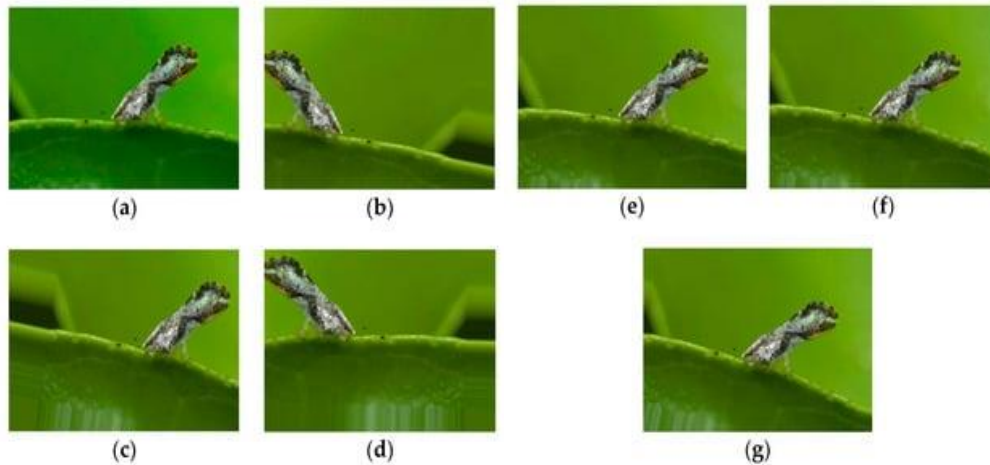
- 목표 : 적대적 공격을 통한 딥러닝 기반 심전도 사용자인증 기술의 안전성 분석 및 보안 취약성 검증
- Dataset : 김예진 연구원으로부터 심전도 전처리 데이터셋을 제공받음 → Training : Validation : Test = 7 : 2 : 1

| | | | |
|--------|------------------------------------|----------------------|---------------|
| 측정기간 | 2016.08.23 ~ 2016.12.27 | 측정담당자 | 최규호 |
| 측정 인원 | 100명 : 조선대학교 IT융합 대학 대학원생 및 학부생 | 피험자 상태 및 조건 | 의자에 앉은 편안한 상태 |
| 측정 시간 | 1회 측정 시간 : 10초 총 60회 측정 | 데이터 sampling rate | 50만 Hz |
| 심전도 유형 | 심전도 Lead-I | 전극 유형 | 습식 전극 |

Related Work (1)

1. (2019) Citrus pests and diseases recognition model using weakly dense connected convolution network

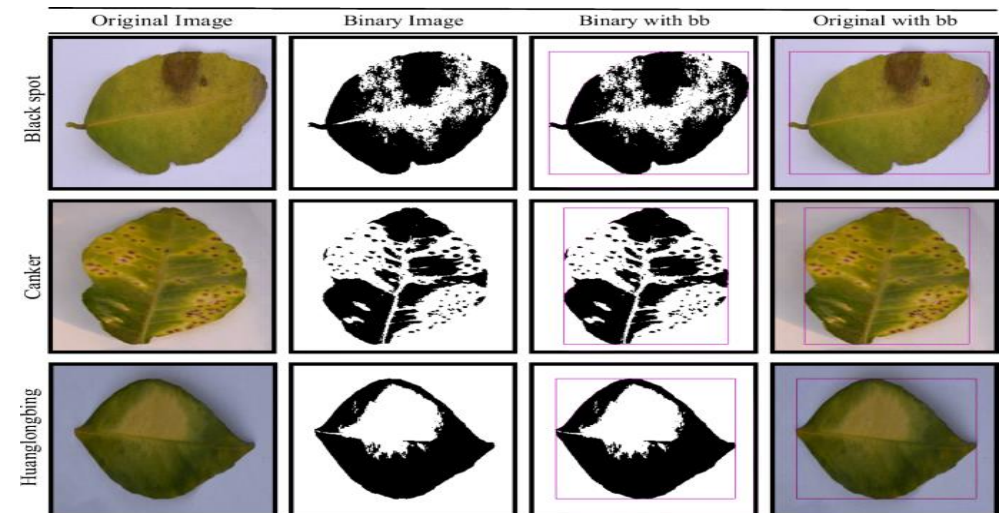
→ Random rotation, reflection, shift, and flip



Related Work (2)

2. (2022) Citrus disease detection and classification using end-to-end anchor-based deep learning model

→ 모든 이미지를 grayscale로 변경



Related Work (1)

1. (2023) Adversarial examples: attacks and defences on medical deep learning systems

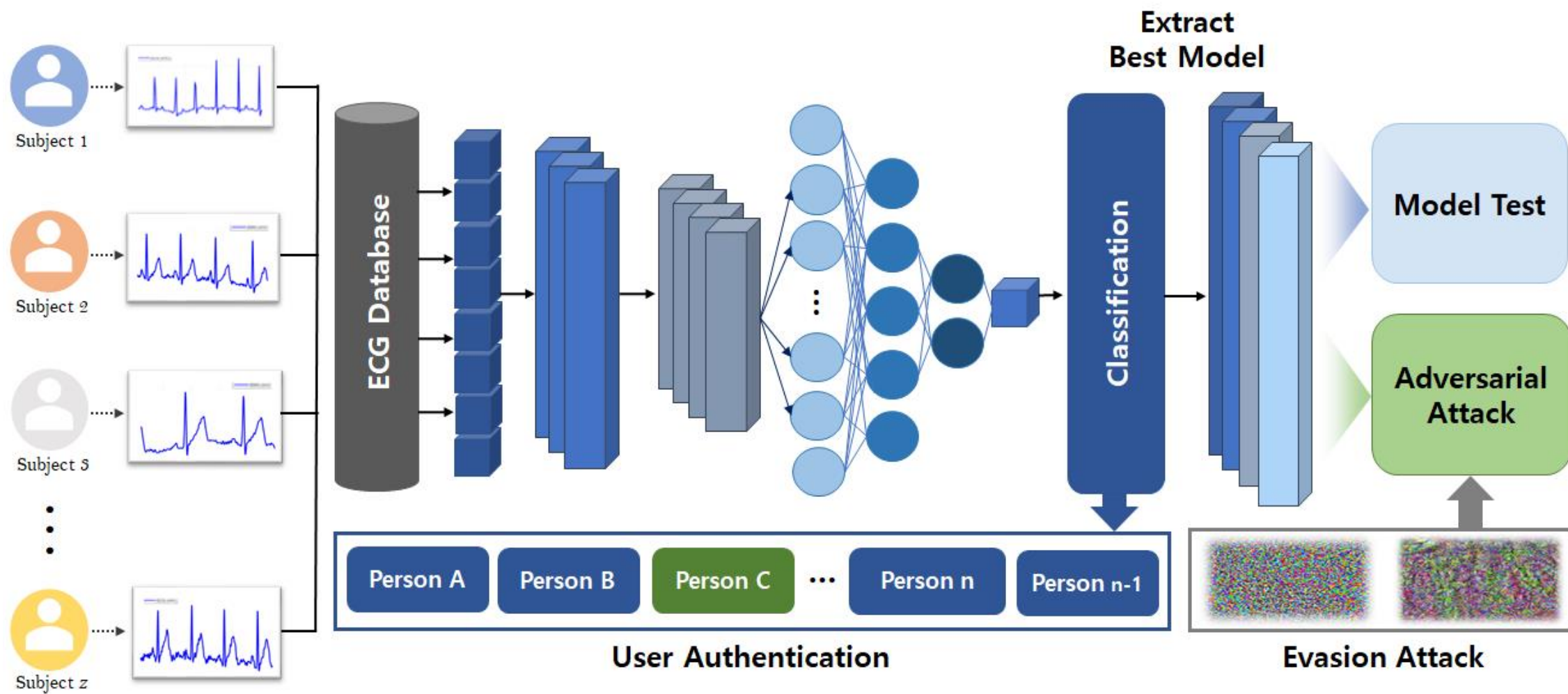
- Dataset : MNIST(Medical and Non-Medical)
 - CXR : 흉부 X선 이미지
 - BS : 혈액 도말 이미지
 - DR : 당뇨병성 망박명증
- Evasion Attack : **FGSM**
- 실험 모델 : VGGNet 19
- 공격 성공률 (epsilon = 0.0003)
 - CXR : 51.84%, BS : 48.1%, DR : 47.8%

Related Work (2)

2. (2020) Deep learning models for electrocardiograms are susceptible to adversarial attack

- Dataset : 2017 PhysioNet/CinC Challenge
 - Training set : Test set = 9 : 1
- Evasion Attack : **FGSM , PGD**
- 실험 모델 : 2017 PhysioNet/CinC Challenge 에서 우승한 13계층 convolution network
- 공격 성공률
 - 50% 이상

Method



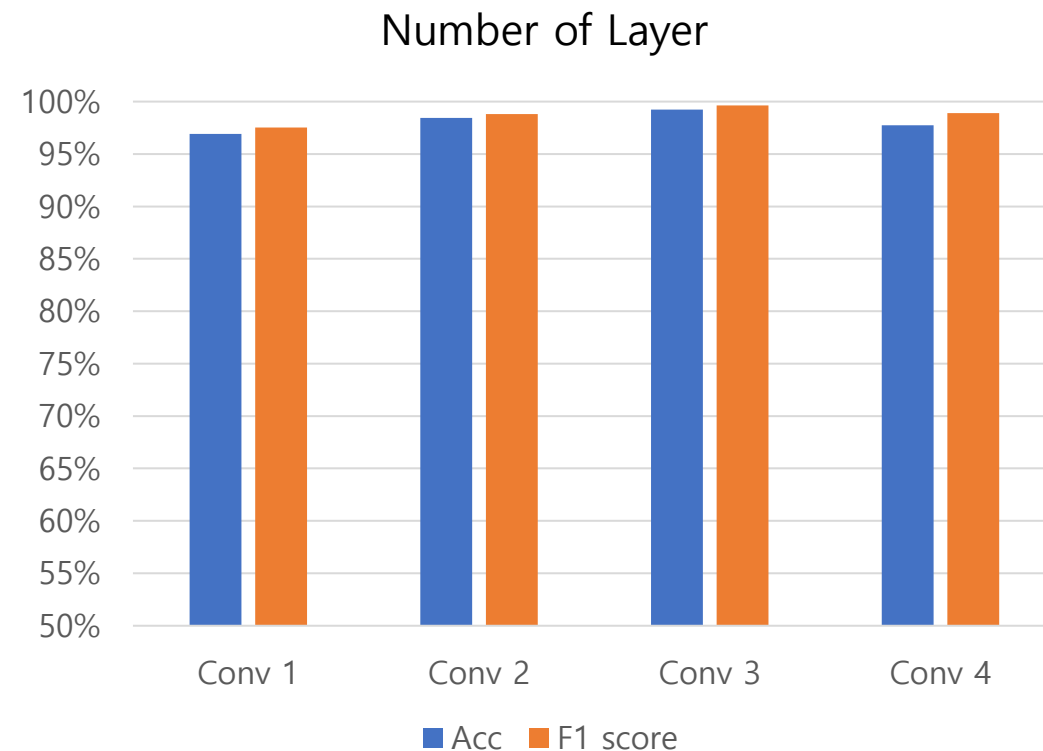
Experiment Result (1/5)

- 1D CNN Layer depth 에 따른 성능 비교 실험

| Number of Layer | Accuracy | F1 score |
|-----------------|----------|----------|
| Conv 1 | 96.92% | 97.53% |
| Conv 2 | 98.44% | 98.83% |
| Conv 3 | 99.25% | 99.65% |
| Conv 4 | 97.75% | 98.91% |

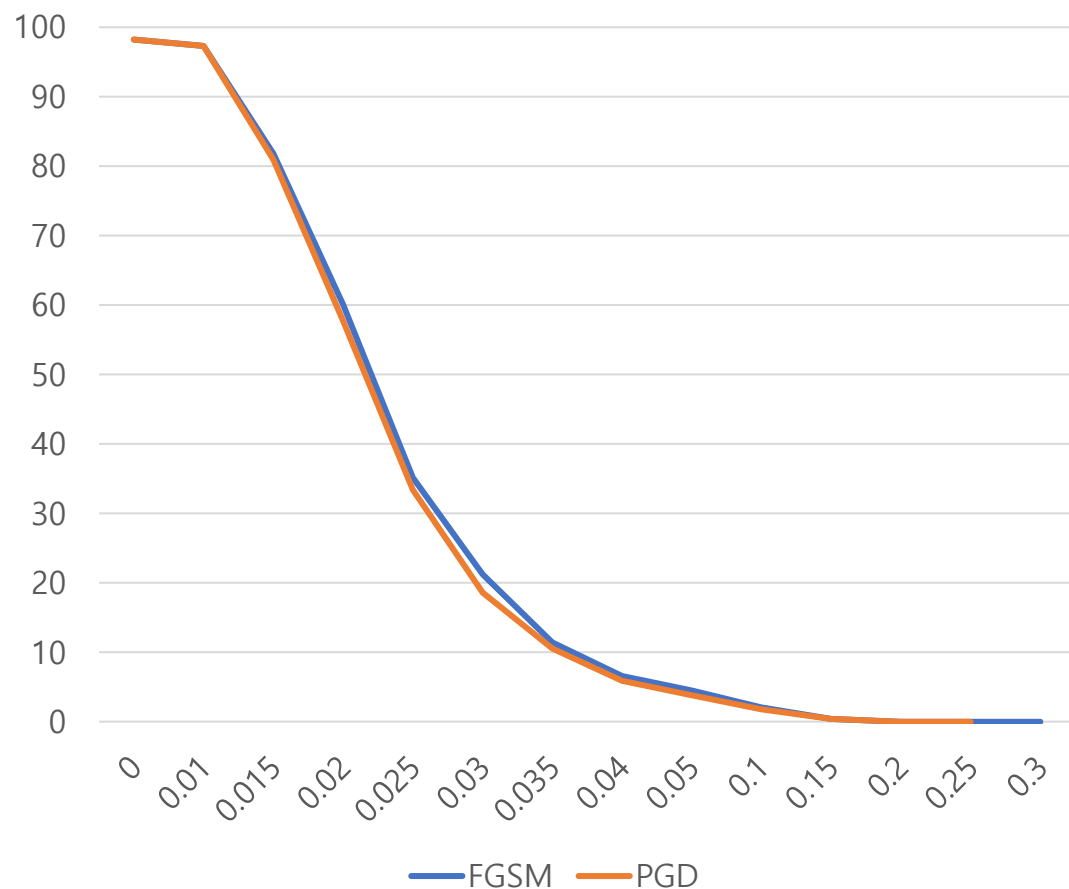
| Test Dataset | Accuracy | F1 score |
|--------------|----------|----------|
| Conv 3 | 98.22% | 98.21% |

Conv 3에서 가장 높은 성능을 도출한 가중치 선택

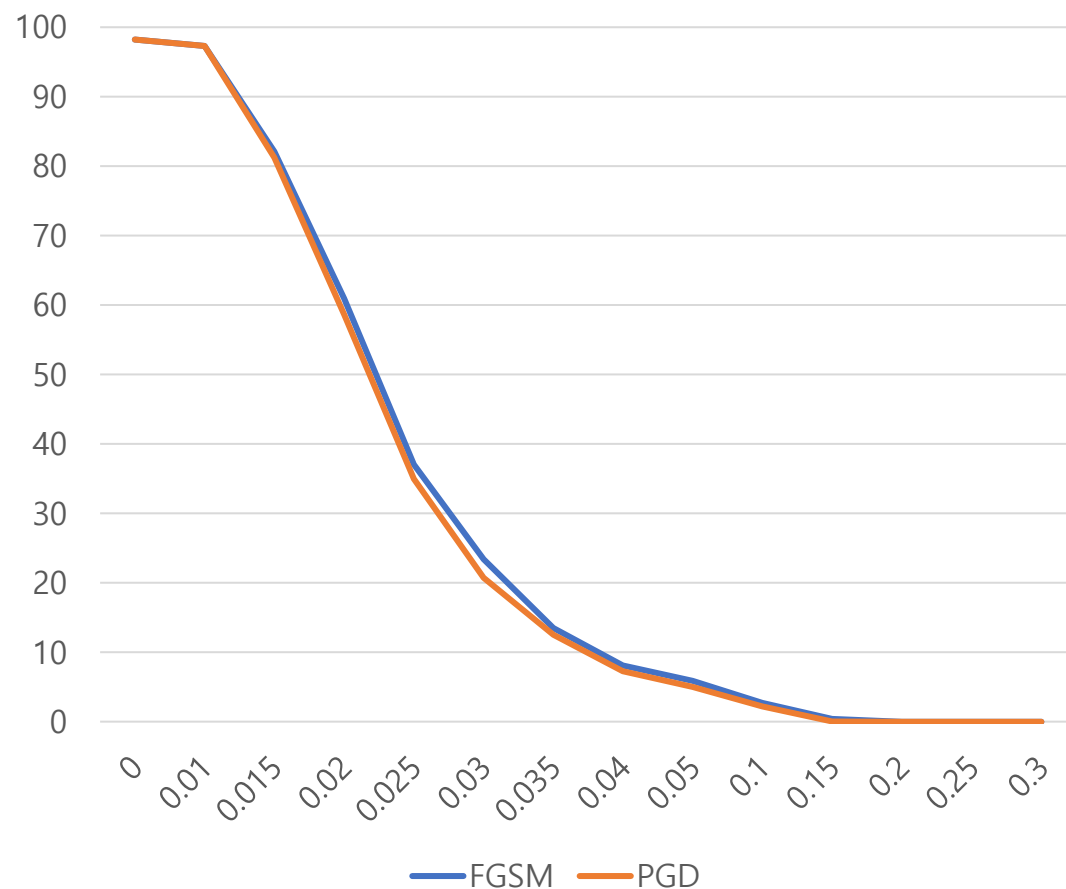


Experiment Result (2/5)

- Evasion Attack Accuracy 하락률

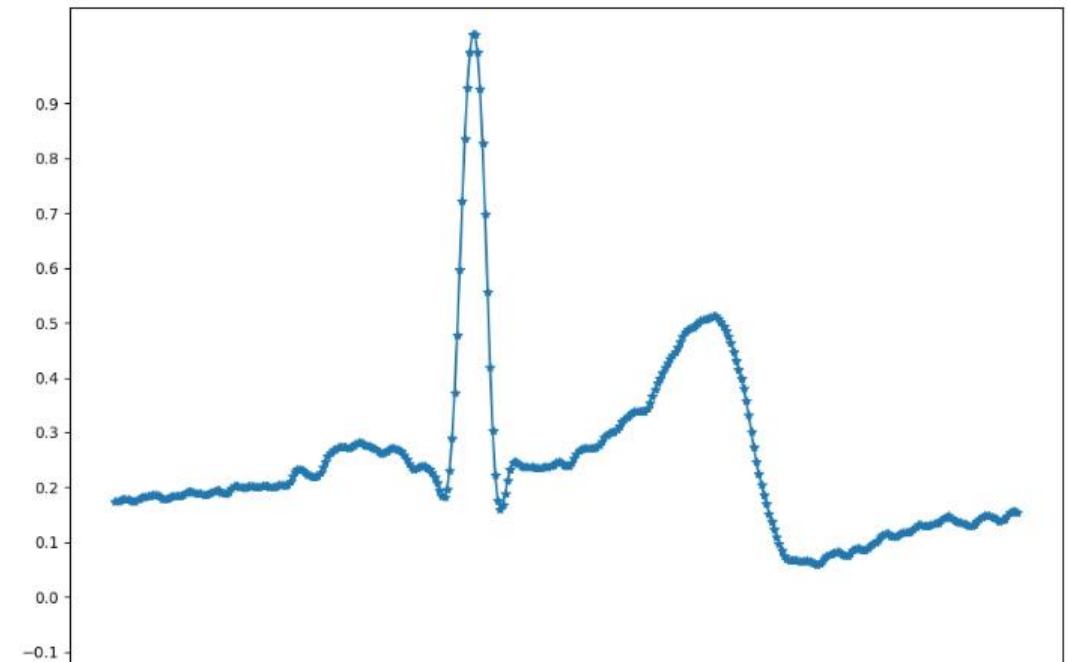
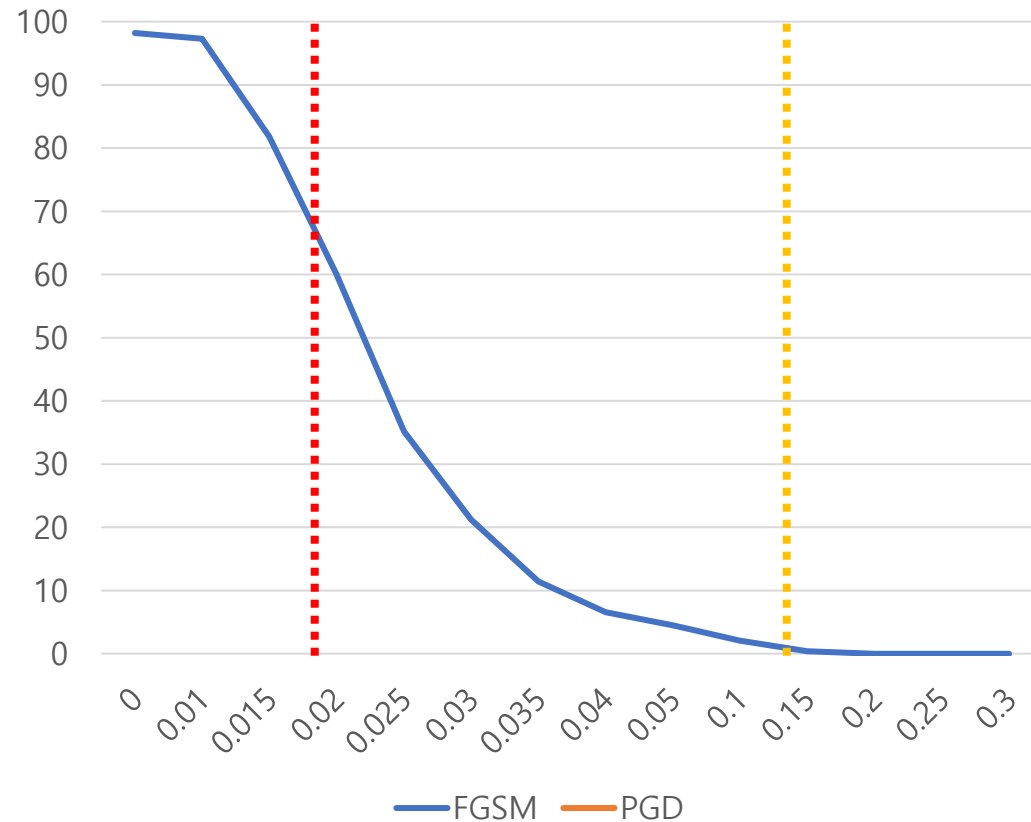


- Evasion Attack F1 score 하락률



Experiment Result (3/5)

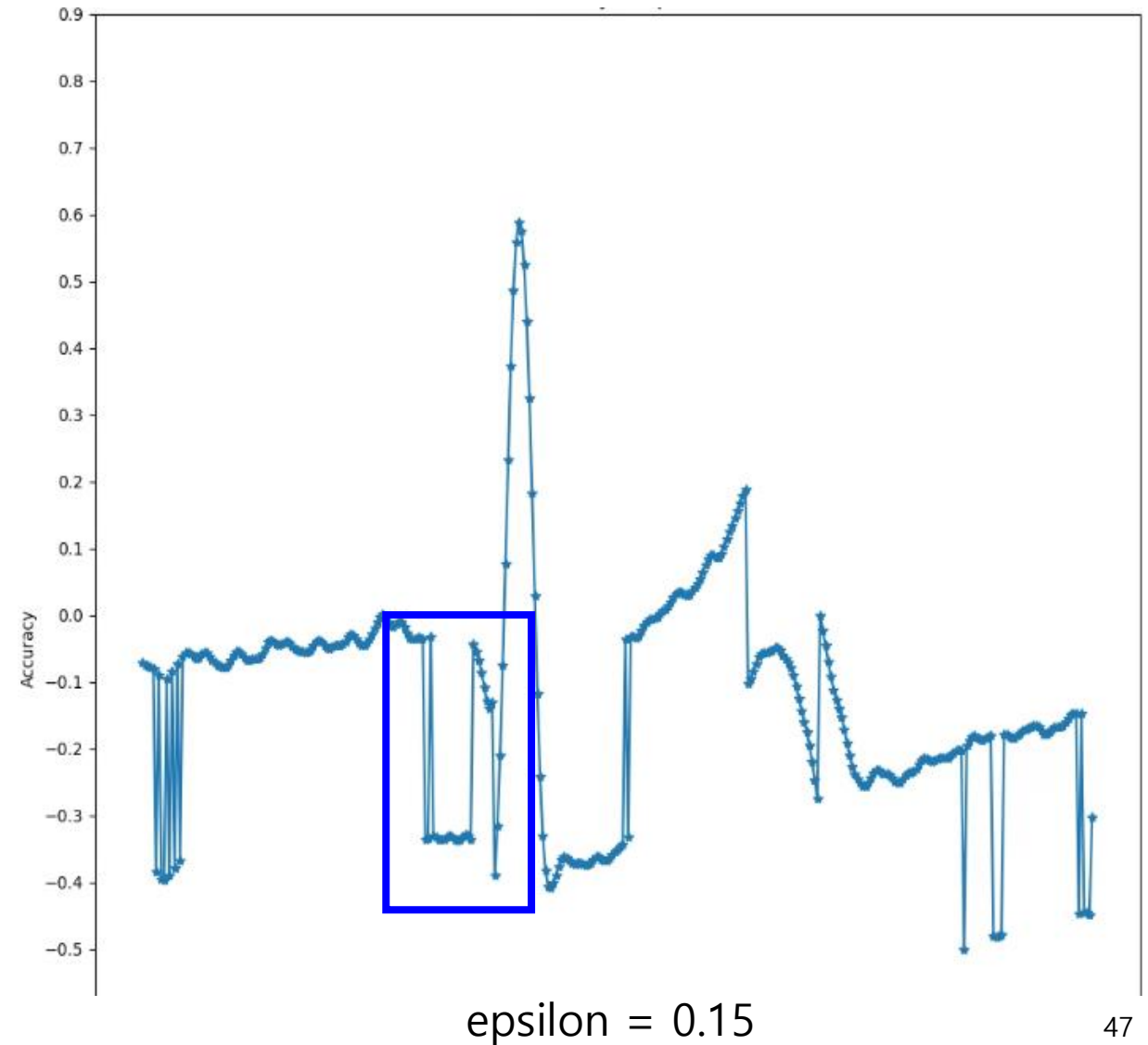
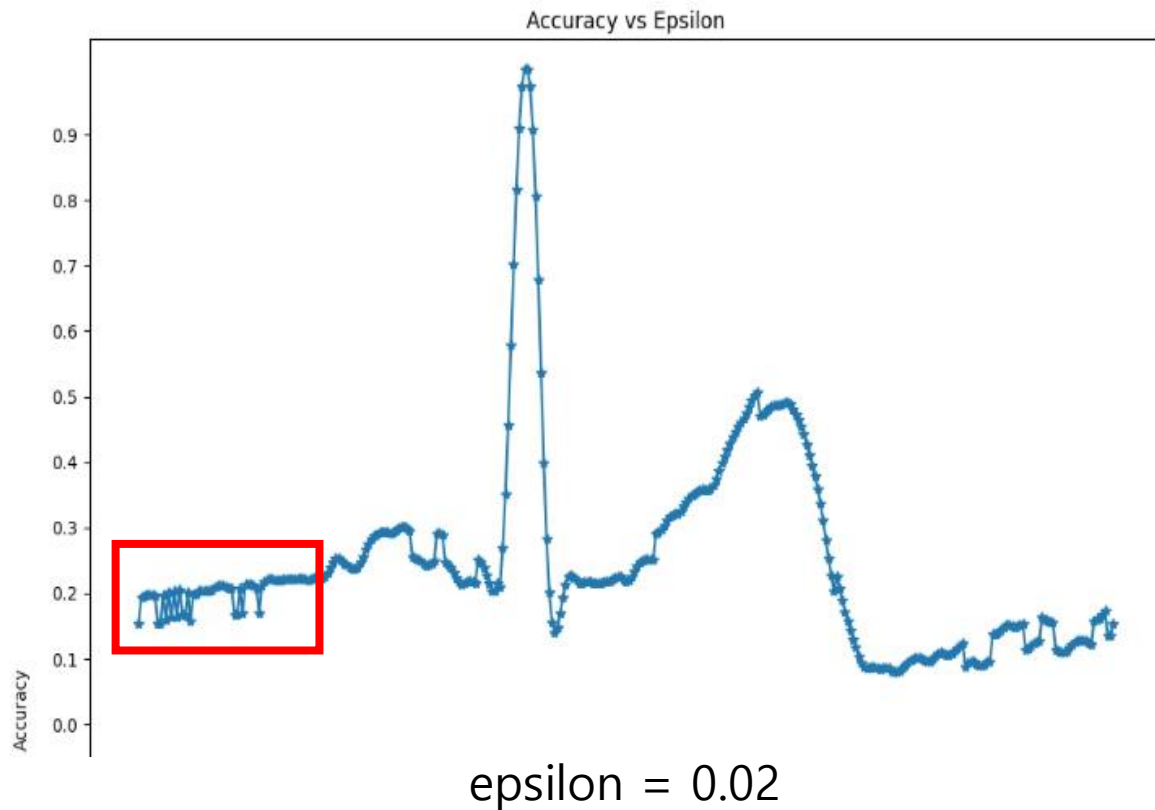
- FGSM (Fast Gradient Sign Method) vs PGD (Projected Gradient Descent) ECG Signal



원본 ECG

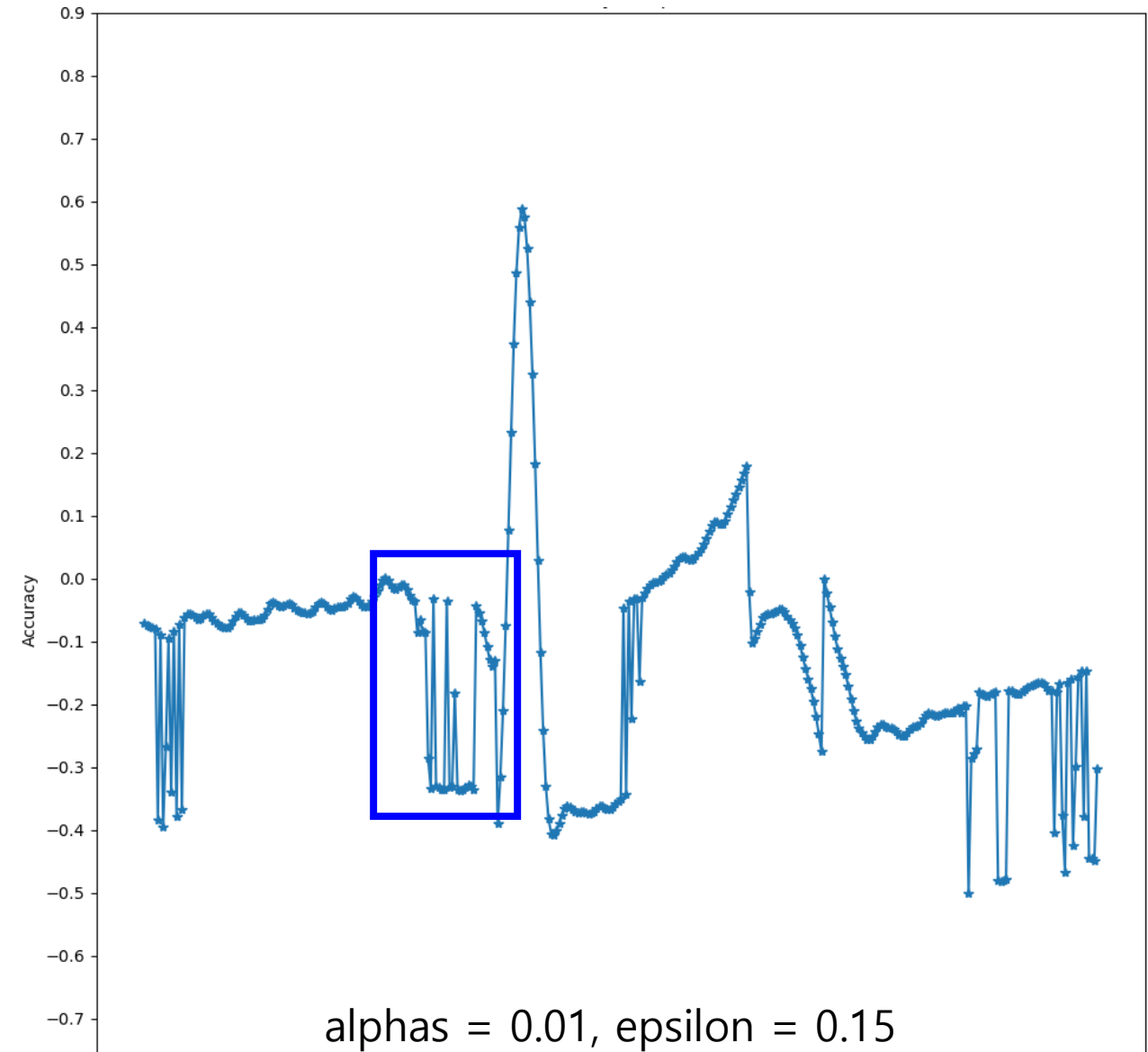
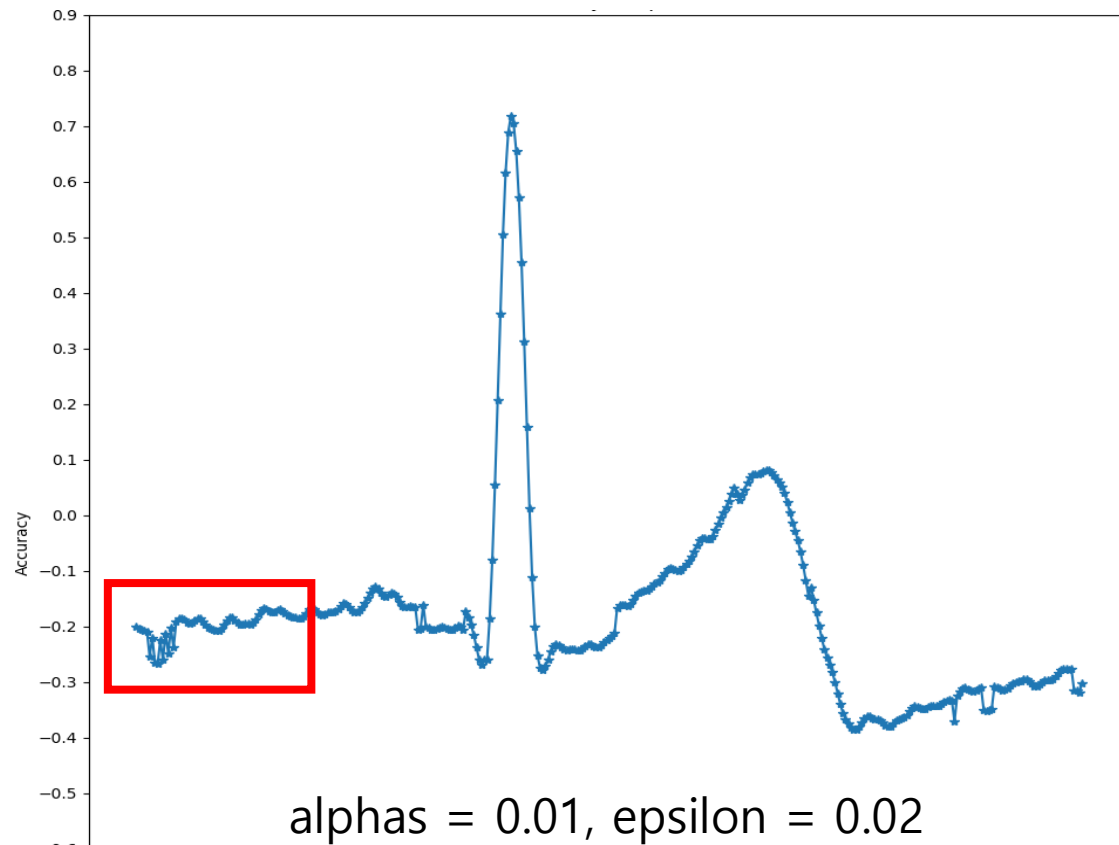
Experiment Result (4/5)

- FGSM (Fast Gradient Sign Method) ECG Signal



Experiment Result (5/5)

- PGD (Projected Gradient Descent) ECG Signal



Conclusion

- 적대적 공격은 딥러닝 모델의 취약점을 공격하는 기술로, 심전도 사용자인증에 매우 취약함을 알 수 있음
- FGSM, PGD 모두 성공적인 성능 하락도를 확인할 수 있음
- 이를 통해, 적대적 공격을 방어할 수 있는 방법을 연구할 필요가 존재함
- 1D CNN외 다른 모델들(LSTM, BiLSTM)을 사용하여, FGSM과 PGD의 공격 성공률을 연구할 예정

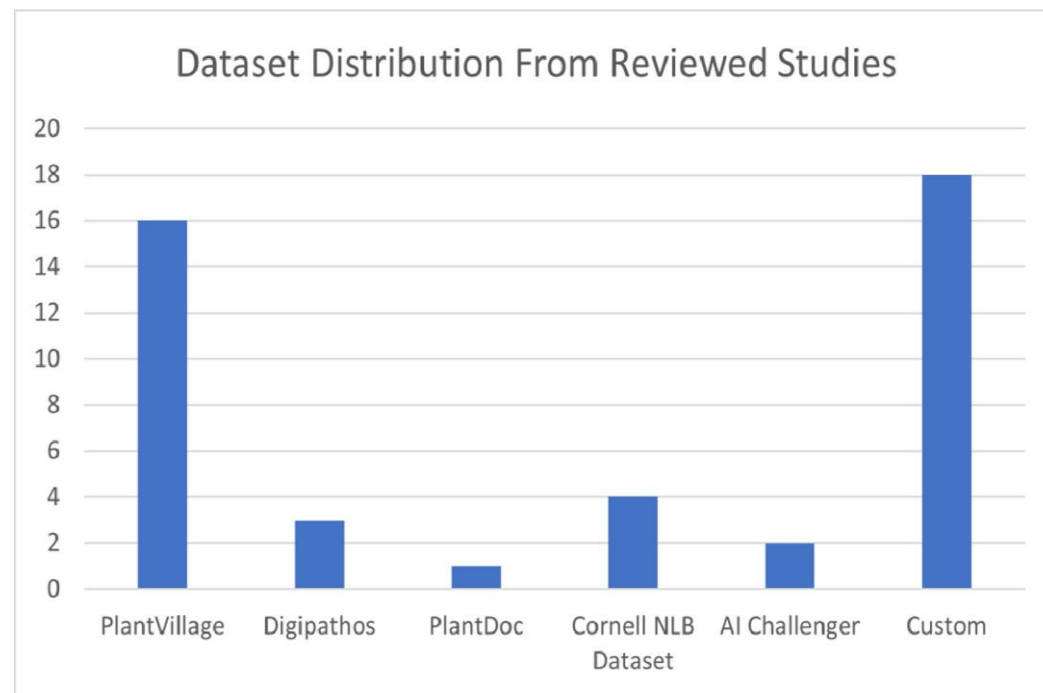
03

Work In Future

- Generating Kiwi Data through Data Diffusion Model
- Follow-up research : Adversarial Defense for ECG user authentication

Generating Kiwi Data through Data Diffusion Model

- 주제 : **Generating Kiwi Data through Data Diffusion Model**
 - 농업 데이터셋은 Open dataset이 매우 희귀함
 - 현재 농작물 질병 인식 연구는 PlantVillage Dataset을 사용한 연구와 Custom Dataset을 사용한 연구가 주를 이룸
- Diffusion Model을 통해 농작물 데이터 합성 연구가 필요



농작물 질병 연구에 수행된 Dataset (2022)

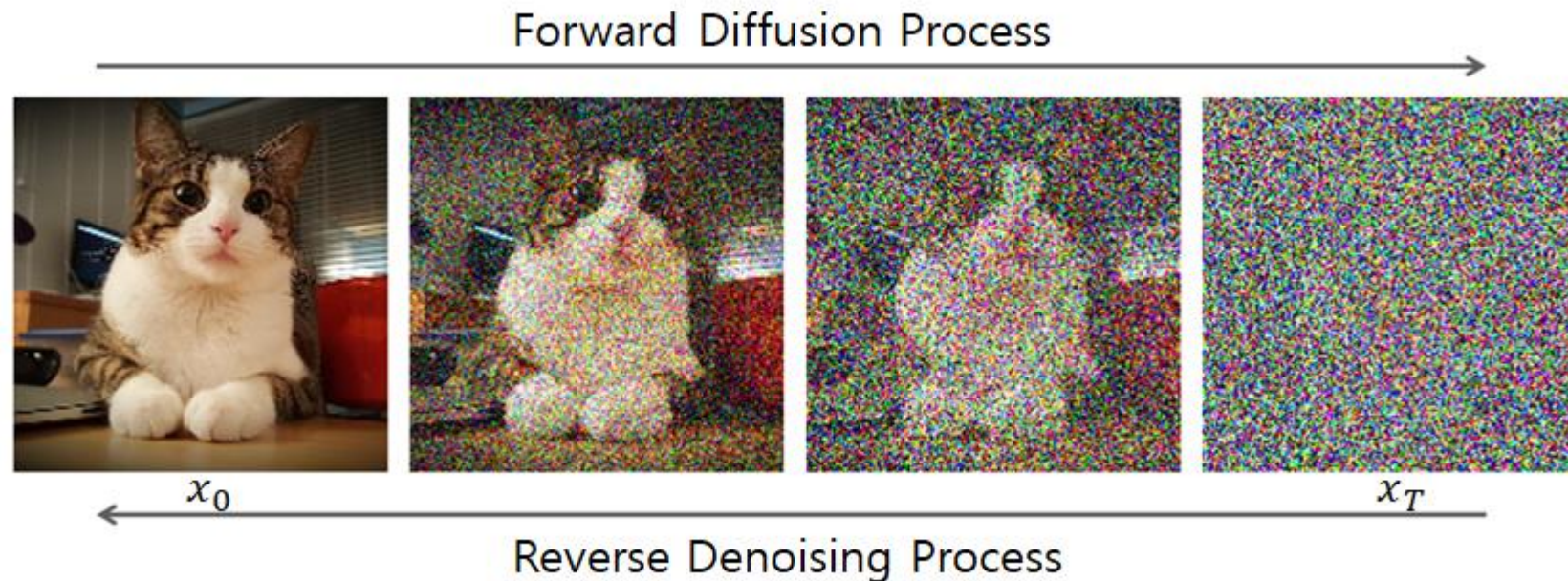
Generating Kiwi Data through Data Diffusion Model

- 목표 : Stable Diffusion model을 통해 kiwi dataset을 합성하는 연구를 수행하고자 함
- Generator Model / GAN(Generative Adversarial Networks) : 생성자(Generative)와 판별자(Discriminator) 2개의 신경망을 경쟁시키며 학습하며, 생성자가 실제 데이터와 유사한 가짜 데이터를 생성하도록 함
→ 새로운 데이터를 처음부터 만들어내는 생성
- Diffusion Model : 이미지나 데이터의 특징을 유지하면서 노이즈를 추가/ 세부 사항을 흐리게 만들어 데이터를 변형함
→ 기존 데이터를 가공하거나 조합하여 새로운 데이터를 생성함

Generating Kiwi Data through Data Diffusion Model

- Diffusion Model

: Input image에 Noise를 여러 단계에 걸쳐 추가하고, 정규 분포를 가진 노이즈를 제거함으로써, input image와 유사한 확률 분포를 가진 이미지를 생성하는 모델



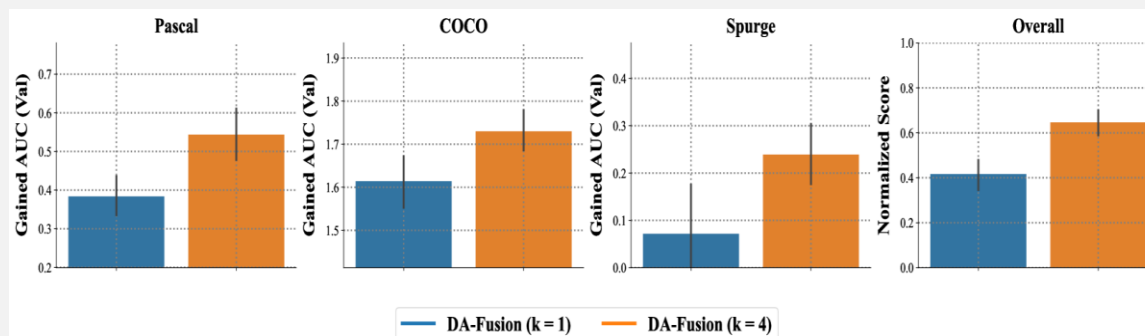
Generating Kiwi Data through Data Diffusion Model

- Diffusion Model을 통한 농작물 데이터 합성 연구가 매우 적음

관련 연구

(arXiv 2023) Effective Data Augmentation With Diffusion Models

- Dataset : Pascal, COCO, Leafy Spurge
- Model : DA-Fusion (Image-to - Image Diffusion Model에서 착안)
- Performance :



Generating Kiwi Data through Data Diffusion Model

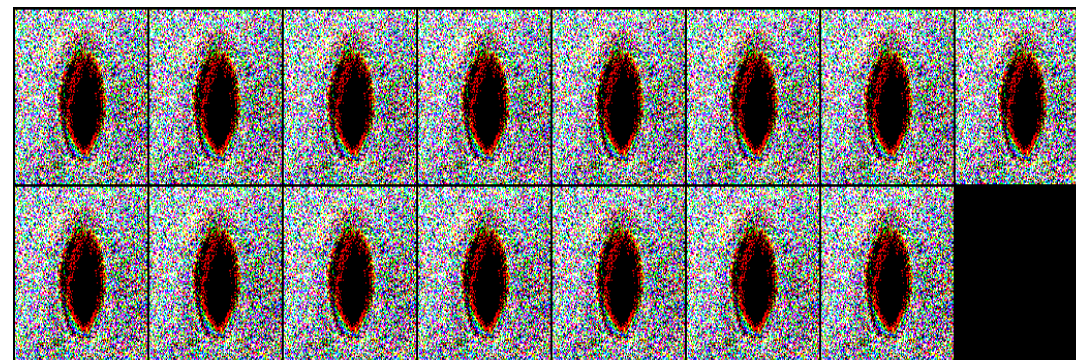
- 현재 GAN 기반 Kiwi Data Generating 연구를 수행 중, 아직 유의미한 성능을 도출하지는 못함

[Generator Model]

- ① 4개의 Convolution layer
- ② Input – Dense – Embedding – Leaky ReLU – Conv

[Discriminator model]

- ① 4개의 Convolution layer
- ② Input – Embedding – Dense – Leaky ReLU – Conv – Flatten – Dropout



→ Diffusion Model을 사용하여 데이터 합성을 연구 수행 (Stable Diffusion Model 사용 예정)

Generating Kiwi Data through Data Diffusion Model

- 해당 연구를 수행하여, 2024년 박사과정생 연구장려금 사업에 지원할 예정
- 2023년 기준 : 연구비 20,000,000원 / 연구기간 : 1~2년
- 최종 목표는 농업 데이터셋을 합성을 통해 데이터를 생성하고, 데이터 생성 모델과 데이터 합성 모델을 비교해 보고자 함

03

Work In Future

- Generating Kiwi Data through Data Diffusion Model
- Follow-up research : Adversarial Defense for ECG user authentication

Adversarial Defense for ECG user authentication

- 주제 : 적대적 공격에 강인한 심전도 사용자인증 보안기술 연구 (후속연구)
- 2024년 「여대학원생 공학연구팀제 지원사업」
- 연구책임자 : 이새봄
- 참여연구원 : 미정
- 연구기간 : 2024. 04. 01 ~ 2024. 10. 31 (8,000,000원)

감사합니다

발표 경청해 주셔서 감사합니다

SaeBom Lee 이새봄
Department of Computer Engineering, Gachon University | Researcher

Tel. +82-31-750-8822 Mobile. +82-10-6641-9390
E-mail. dltoqha@gachon.ac.kr