

Deep Learning

TermProject

22102304

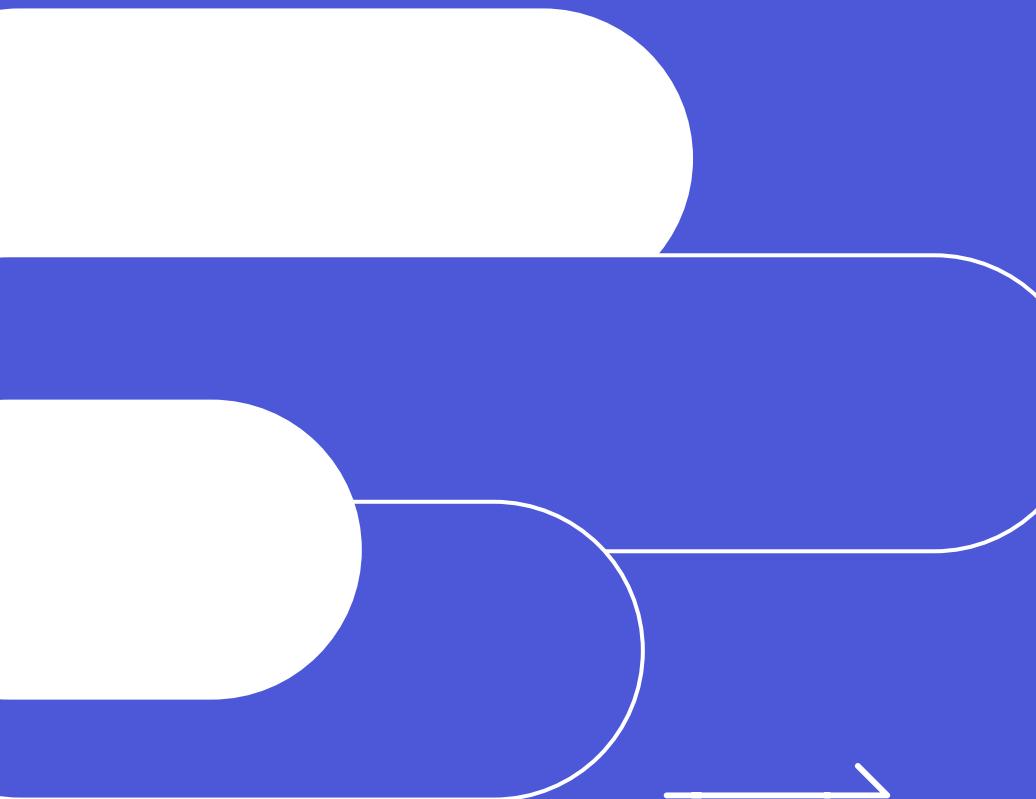
김나현

22102311

김승범

22102335

이수



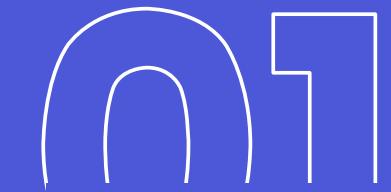
summary of one year's achievements

Project report.

DeepLearning

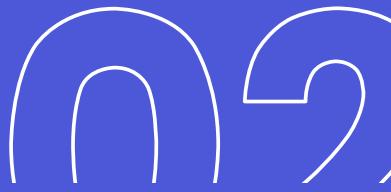
목차.

Project
report. →



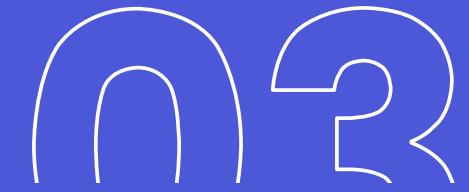
Introduction

인공지능 기반 나무 질병
연구의 중요성을 소개합니다



Data

이미지 전처리 단계에서 과
적합을 막기 위해 노력한 것
을 소개합니다



Model

Model의 Overview 및
근거에 대해 자세히 설명합
니다



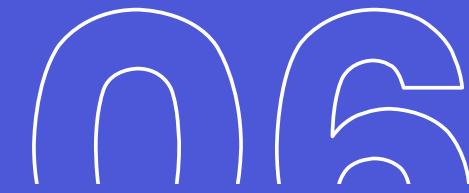
Results

Validation의 결과를 제시
합니다



Conclusion

프로젝트의 최종 결론을
제시합니다



Q&A

01

왜 나무 질병에 대해 연구해야하는가?

Why.

intro



실시간으로 수목관리 가능해진다

2023년부터 실행

엘젠이 아리울씨앤디와의 협약으로 인공지능(AI)을 활용한 조경 산업에 진출한다. 국내 조경 산업 시장의 규모가 7조 원에 이르며, 주택 구매자들의 주거환경에 대한 관심이 높아지고 있는 가운데, 두 기업은 음성인식과 영상인식 기술을 통한 AI키오스크를 개발하고 명품 조경을 선보일 계획이다. 또한, 엘젠은 자사 기술을 활용하여 분수를 제어하는 기술을 소개할 예정이다. [출처: 인공지능신문]

하나, 조기진단 및 모니터링

인공지능을 활용한 IoT 조경수목은 병해를 조기에 선별하고 예후를 예측하여 조기진단 및 모니터링이 가능합니다.

둘, 생태계 보전

나무 질병의 조기 발견으로 인해 전염을 예방하고 퍼지는 것을 방지함으로써 산림 생태계를 보존하는데 도움이 됩니다.

셋, 임업사업 지원

임업 산업에 적용되어 나무 생산성을 향상시키고 임산물 품질을 개선할 수 있습니다. 예를 들어, 병해충이나 질병에 감염된 나무를 식별하여 신속하게 처리함으로써 수확량을 최적화 할 수 있습니다.



09

Data

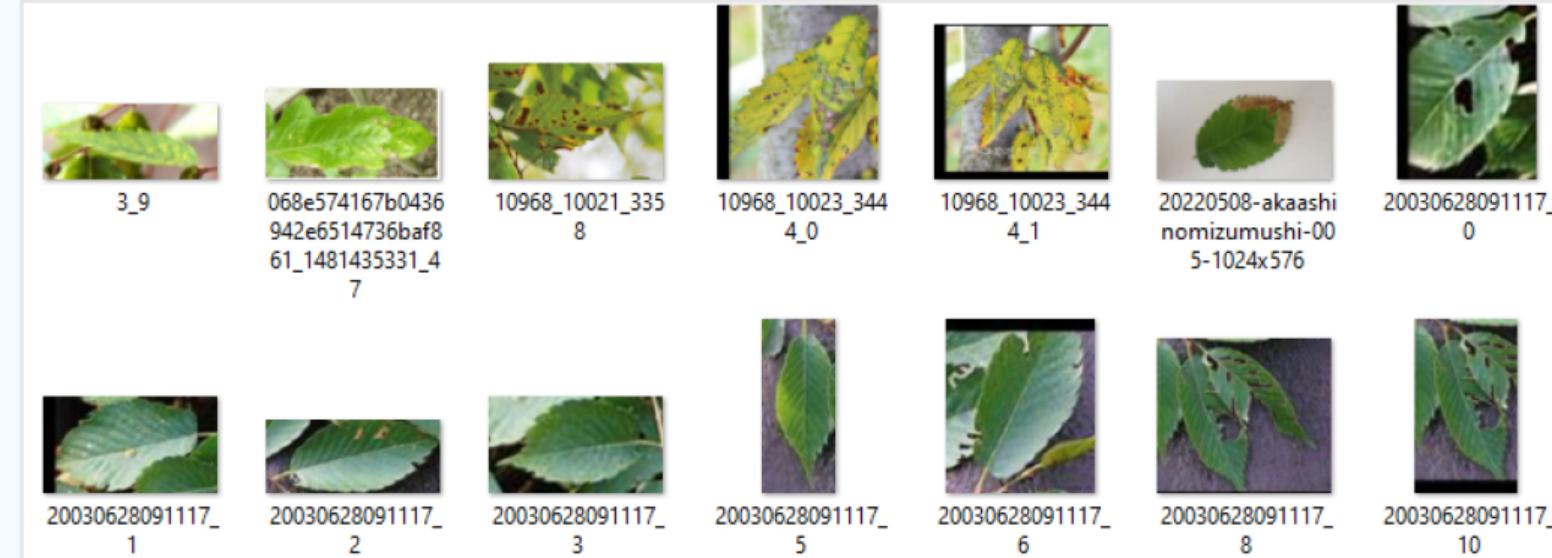
About.

이미지 전처리 단계에서 과적합을 막기 위해 노력
한 것을 소개합니다

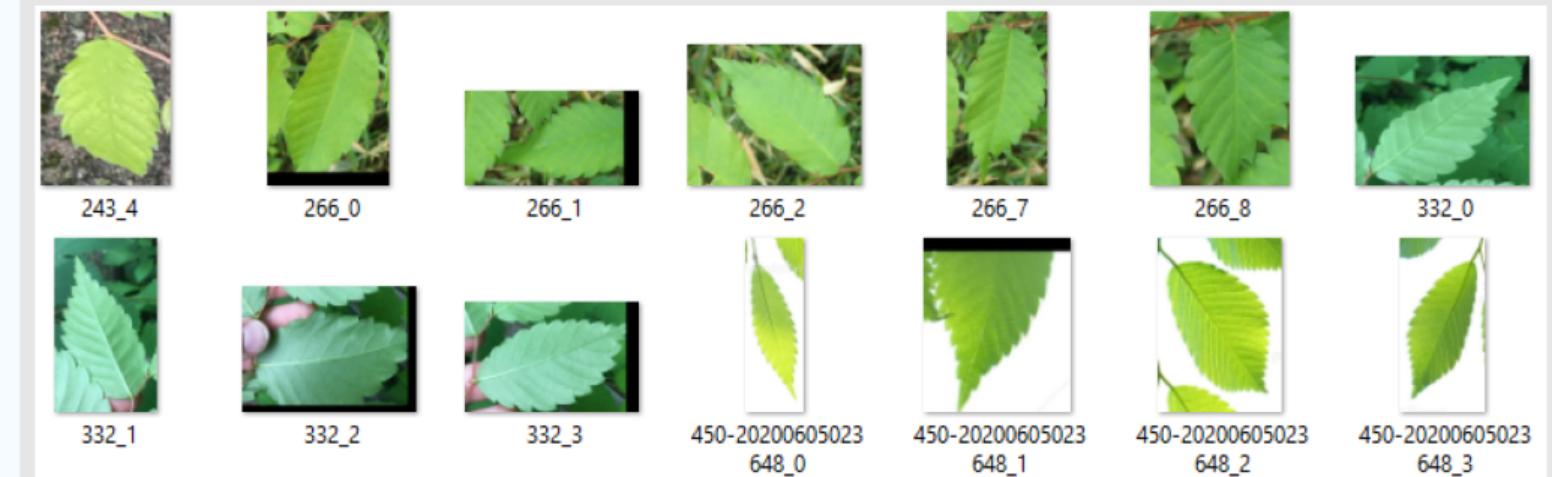


느티나무 질병 데이터셋

Disease



Healthy



데이터 개요

소개

- 느티나무 잎 개체 이미지, 정상과 비정상(병충해) 등이 매칭된 데이터

구축목적

- 지능형 조경수목 병해 조기선별 및 예후예측

09

Data

About.

이미지 전처리 단계에서 과적합을 막기 위해 노력한 것을 소개합니다



이미지 전처리 다양한 데이터 전처리 기법들

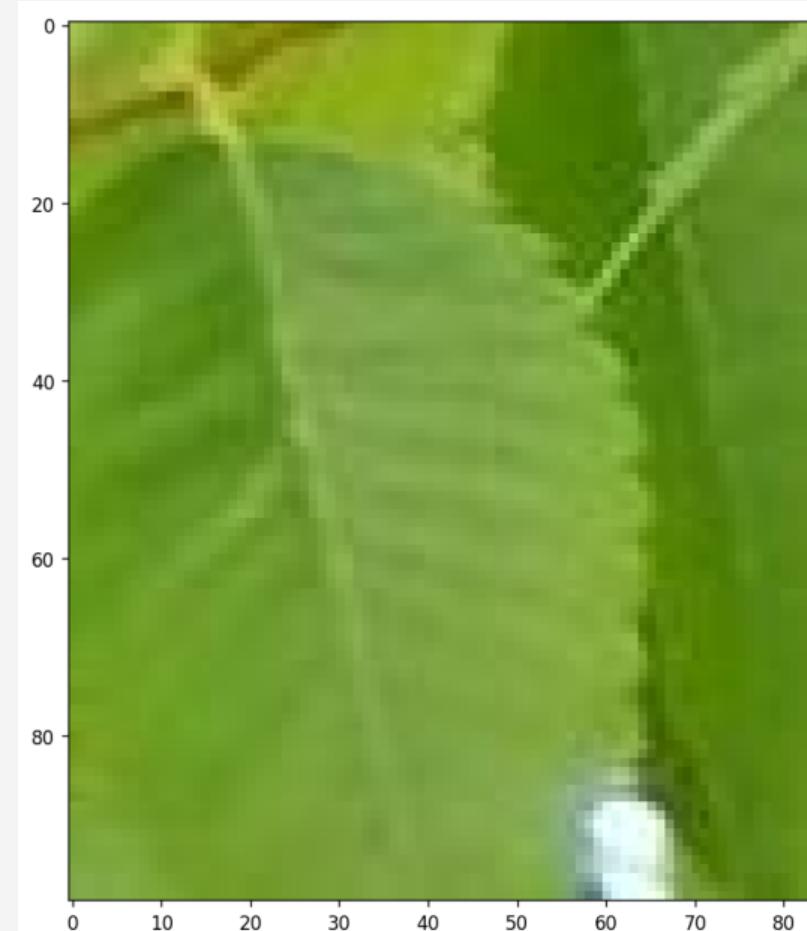
란초스보간법 및 제로패딩사용

정보손실 최소화

Data augmentation

Randomresized crop

$$S(x, y) = \sum_{i=\lfloor x \rfloor - a + 1}^{\lfloor x \rfloor + a} \sum_{j=\lfloor y \rfloor - a + 1}^{\lfloor y \rfloor + a} s_{ij} L(x - i)L(y - j).$$
$$L(x) = \begin{cases} 1 & \text{if } x = 0 \\ \frac{a \sin(\pi x) \sin(\pi x/a)}{\pi^2 x^2} & \text{if } 0 < |x| < a \\ 0 & \text{otherwise} \end{cases}$$



란초스보간법이란?

주어진 데이터 포인트들을 통해 주어진 범위에서 다항식으로 근사하는 보간법 중 하나입니다. 다수의 데이터 포인트가 주어진 경우, 데이터 포인트를 정확하게 지나는 하나의 다항식을 찾아내는 방법을 제공합니다. 이 다항식은 주어진 데이터를 완벽하게 표현하며, 추가적인 점에 대해서도 그 값을 예측할 수 있습니다.



Data

About.

이미지 전처리 단계에서 과적합을 막기 위해 노력
한 것을 소개합니다



이미지 전처리 다양한 데이터 전처리 기법들

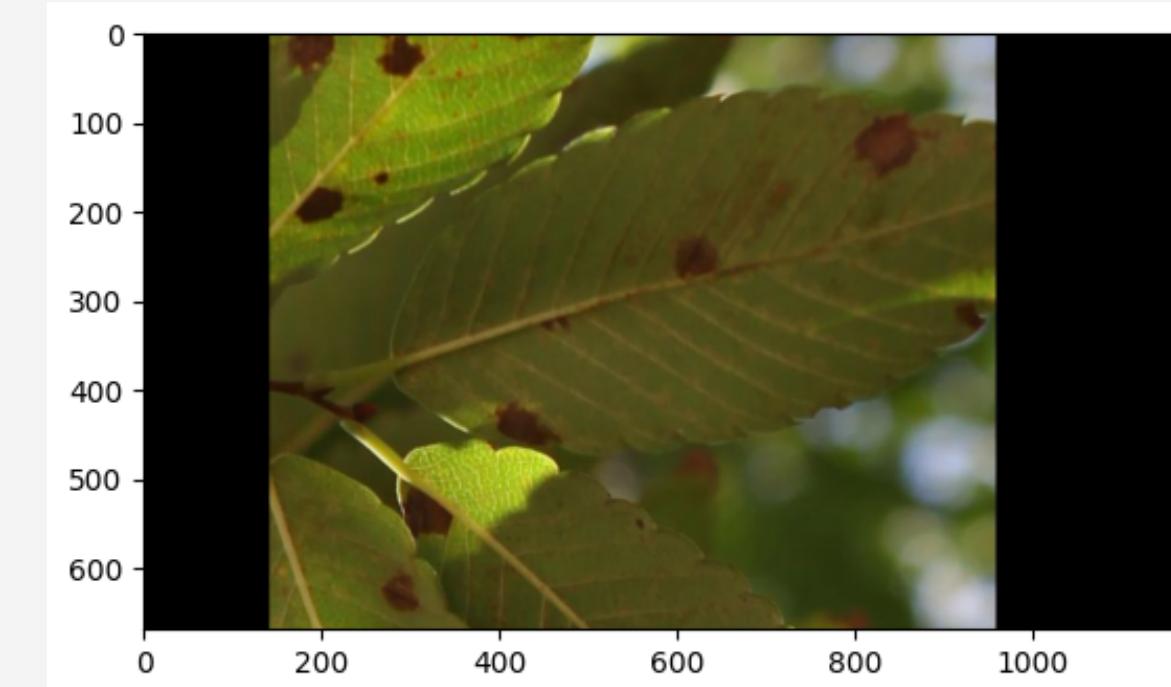
란초스보간법 및
제로패딩사용

정보손실
최소화

Data augmentation

Randomresized
crop

제로패딩: 입력 데이터 주위에 0으로 채워진 가상의 패딩을 추가하는 것을 의미합니다.
데이터의 입력과 출력의 공간적 크기를 일정하게 유지하거나 더 유용한 특성을 추출하기
위해 사용됩니다.



(width, height)의 잘라낼 크기를 설정하면, 그 크기만큼 랜덤으로 잘라낸 RandomCrop과 다르게 랜덤하게 자른 후에 주어진 크기만큼 Resize를 해주는 것을 의미합니다. 이러한 이유로 크기를 이미지의 크기보다 크게 넘겨도 오류가 나지 않습니다. 과적합을 막기 위해서 사용하였습니다.

03 Model

About.

Model의 Overview 및
근거에 대해 자세히 설명합니다

Model의 Overview

1) ResNet

residual block, skip connection을 사용한 모델입니다.

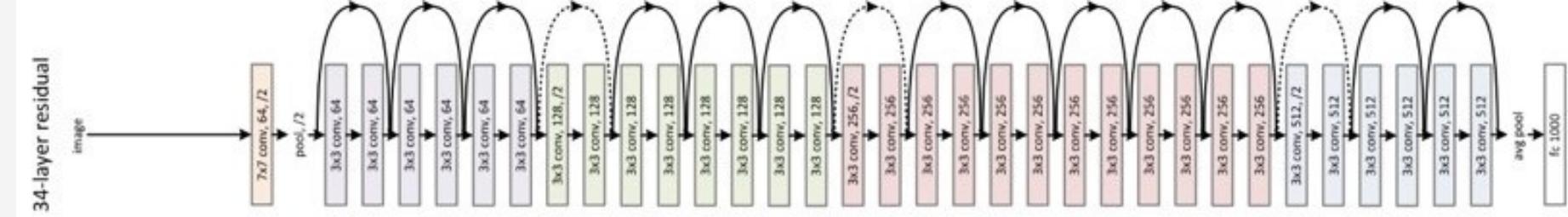
ResNet의 깊이가 조금만 깊어지더라도 데이터 수의 부족으로 인해

쉽게 과적합의 양상이 나타났고, 이 때문에 높은 정확도를 달성하기가 힘들었습니다.

2) VisionTransformer(ViT)

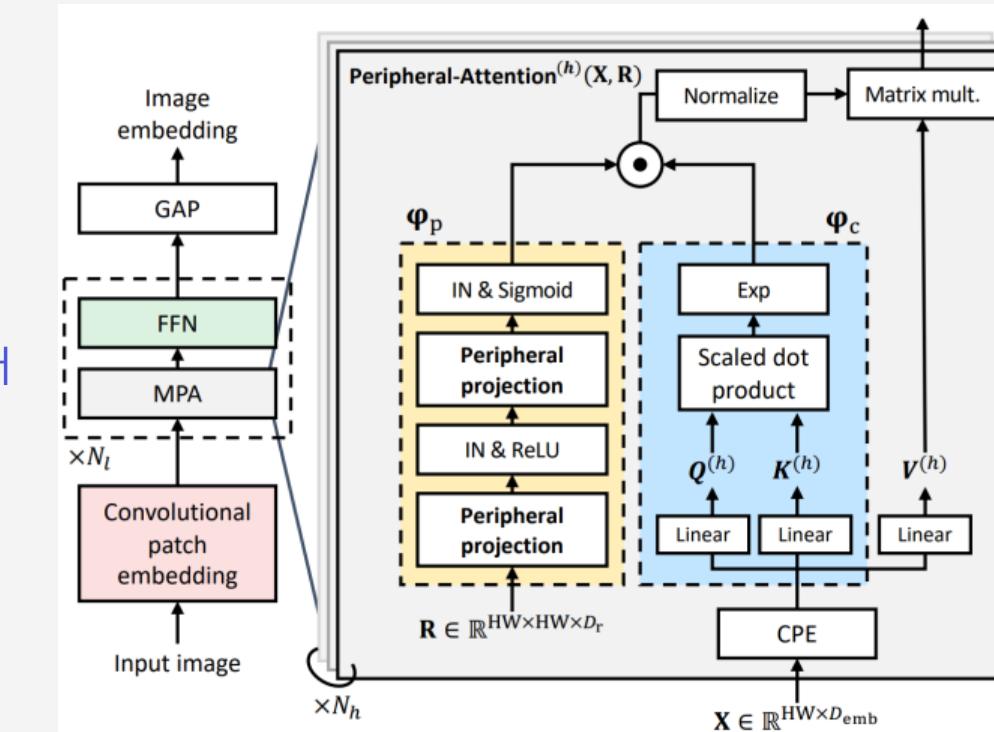
이미지를 격자로 분할하고 각 격자의 정보를 펼쳐서 일렬로 나열한 후, 이를 트랜스포머 모델에 입력으로 하는 VisionTransformer 모델입니다.

ImageNet-21k pretrained 모델을 기반으로 finetuning함으로써 높은 성능을 보였지만, 데이터가 부족하다는 근본적인 문제가 남아있었습니다.



3) VisionTransformer(ViT) + PlantVillage Dataset

Imagenet21k로 pretrain된 ViT모델을 기반으로 PlantVillage 데이터에 대해 finetuning후,
해당 모델로 느티나무 질병 데이터에 대해 Kfold로 finetuning한뒤 soft voting방식으로 양상불한 결과
더 나은 정확도를 관찰할 수 있었습니다.



K fold crossvalidation

- 더 일반적인 모델
- 더 신뢰성 있는 평가를 진행

03

Model

About.

Model의 Overview 및
근거에 대해 자세히 설명합니다



Model의 차별점

VisionTransformer(ViT) finetuning PlantVillage Dataset

imagenet-21k pretrained VisionTransformer 모델을
PlantVillage 데이터셋에 대해 finetuning 한 뒤,
해당 모델을 기반으로 느티나무 질병 데이터에 대해 K fold finetuning을 진행하여 ($k=5$)
학습된 K개의 모델을 Soft Voting 방식으로 ensemble하여 최종 모델을 구성하였습니다.
기존의 finetuning에 비해 중간에 PlantVillage로의 1차적인 finetuning 과정을 통해
식물질병 탐지에 관한 지식을 학습함으로써 더 나은 정확도를 보일 수 있었습니다.



PlantVillage Dataset

54303개의 healthy & unhealthy leaf images로 이루어진
데이터셋으로, 식물 종류와 질병 여부에 따라 총 15개의 클래스로
분류되어있는 대규모 식물 질병 데이터셋입니다.

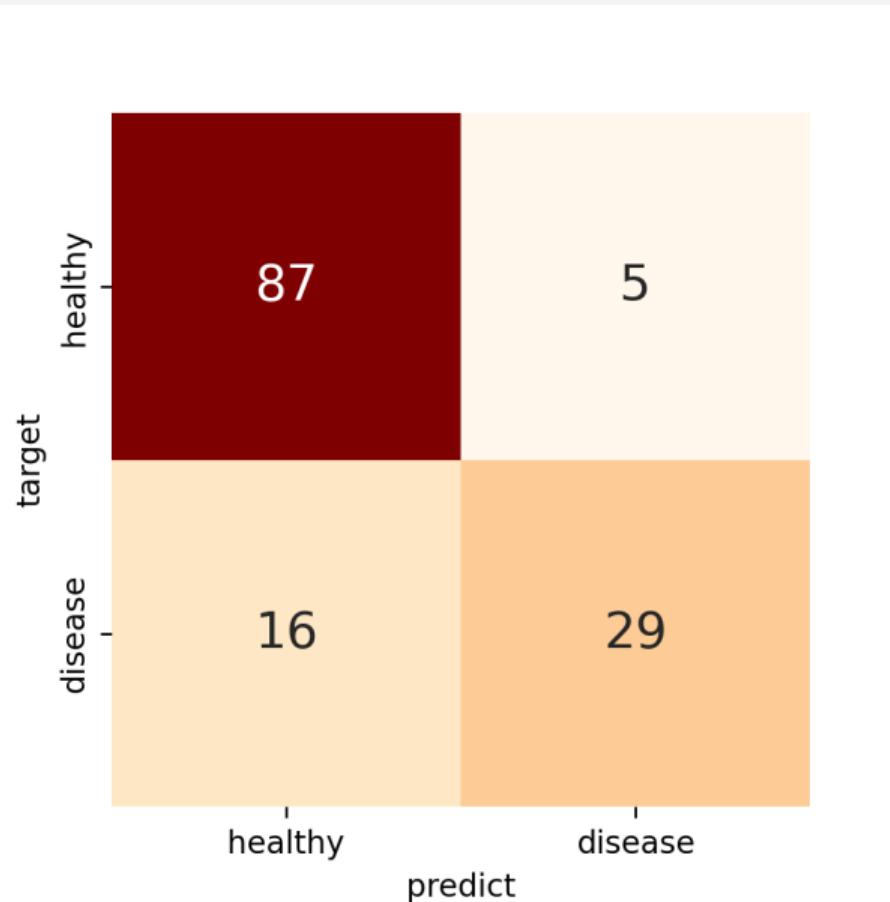
0/4

Results

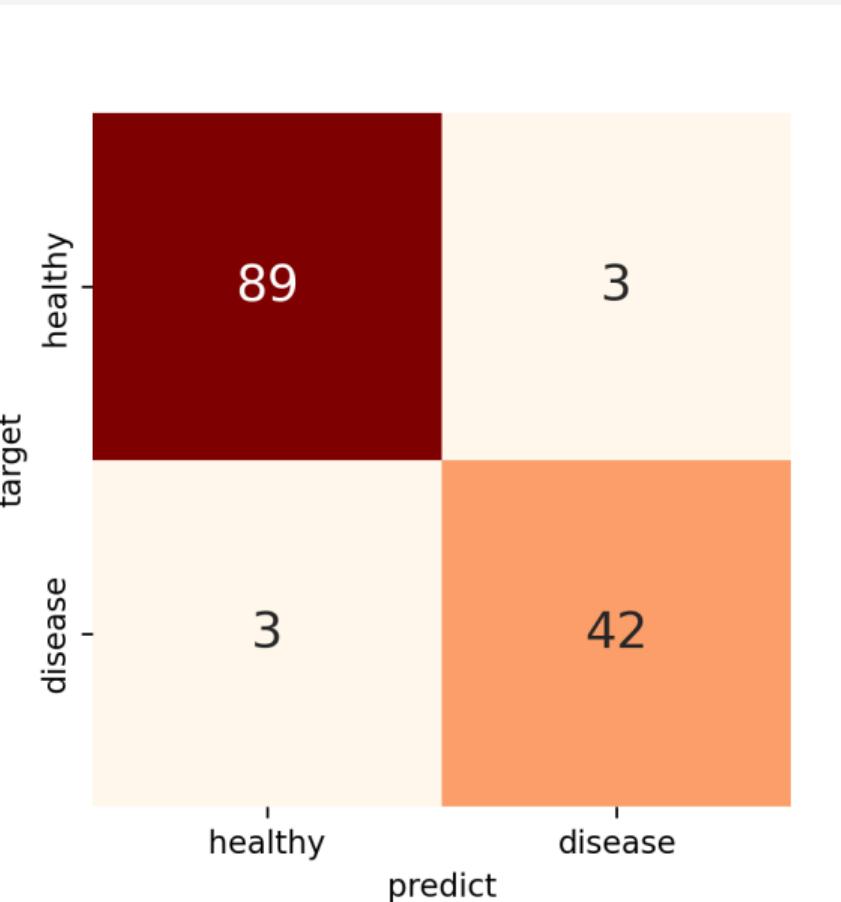
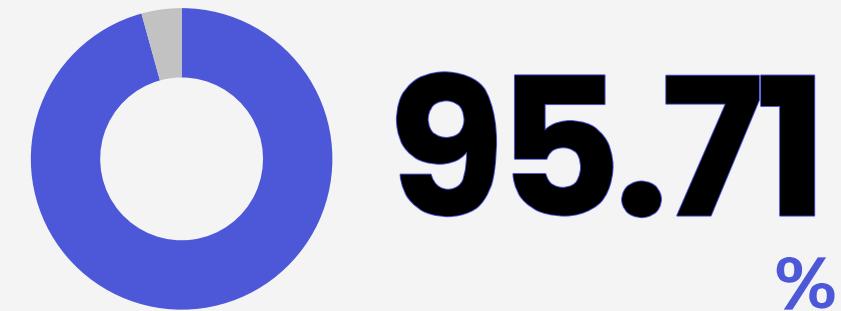
About.

Accuracy와 Confusion matrix를
이용해 Validation의 결과를 제시합니다

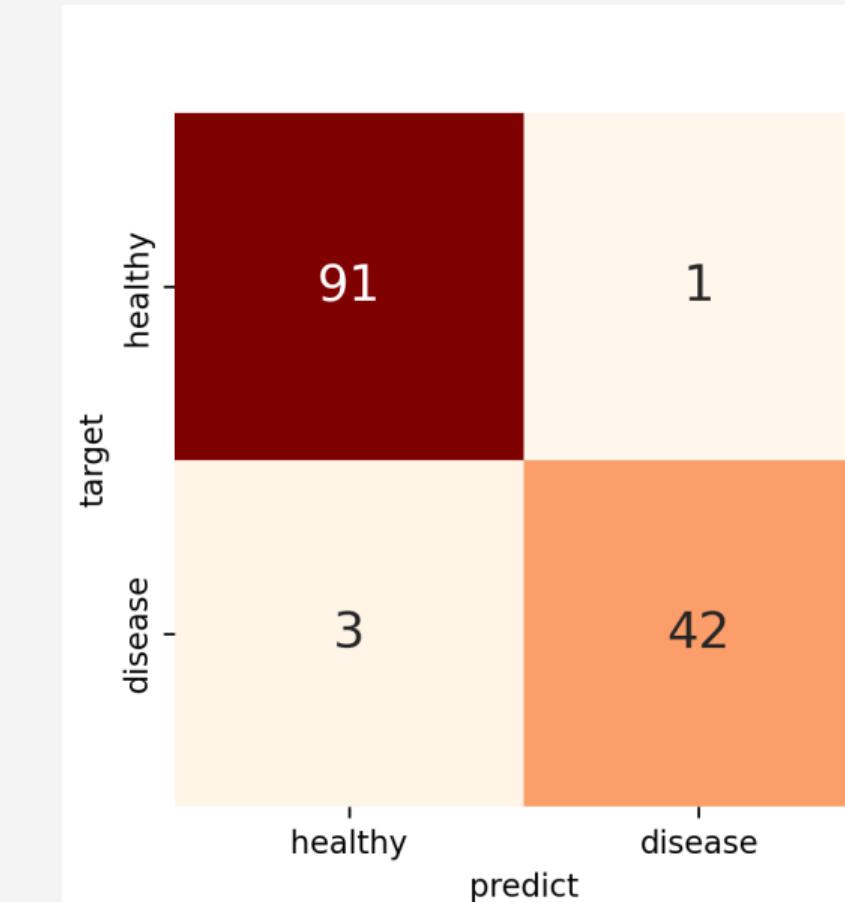
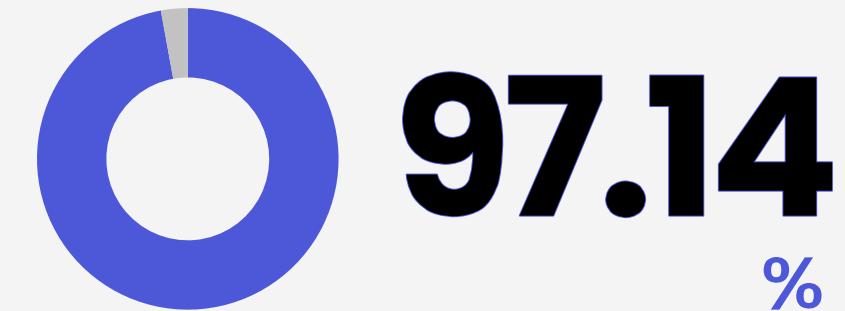
Resnet



ViT



ViT (PlantVillage)





Results

About.

통계적인 관점을 이용해 Validation의 결과를 분석합니다



검정(McNema 방법)

- McNema 방법: 통계적으로 두 가지 상호 배타적인 범주를 가진 반복 측정 자료를 분석하는 데 사용되는 검정 방법 중 하나입니다. 이 방법은 두 가지 처리나 조건 간의 차이를 확인하기 위해 사용됩니다. 주로 2x2 교차표로 표시된 데이터를 분석하는 데 활용됩니다. 따라서 Confusion Matrix를 분석하기 위해 이 방법으로 검정하였습니다.
- H0(귀무가설): my model is oracle v.s. H1(대립가설): my model is not oracle [oracle: 이상적인 모델]
- If My model is oracle :Real Value는 A와 D에 몰립니다.

A	B
C	D

A: TP, B: FP, C:FN, D:TN

$$\chi^2 = \frac{(b - c)^2}{b + c} \sim (1)$$



Results

About.

통계적인 관점을 이용해 Validation의
결과를 분석합니다



검정(McNema 방법)

- FP = FN 일 경우에는 McNema 방법을 할 수 없습니다. ($\because b-c = 0$)
- 따라서 이 경우에는 신뢰구간을 이용하여 검정을 합니다. (ViTs의 경우)

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{Total Samples (TP} + \text{FP} + \text{FN} + \text{TN})}$$

$$\text{SE} = \sqrt{\frac{\text{Accuracy} \times (1 - \text{Accuracy})}{\text{Total Samples (n)}}}$$



Results

About.



검정 결과 $[X(1) = 5.02 \text{ (95% 유의수준)}]$

Resnet

- 검정 통계량 : 5.761 ($5.761 > 5.02$)
- 귀무가설 기각
- 이상적인 모델 X

ViTs

- P-value : 0.0122 ($0.0122 < 0.05$)
- 귀무가설 기각
- 이상적인 모델 X

ViTs(plantvillage)

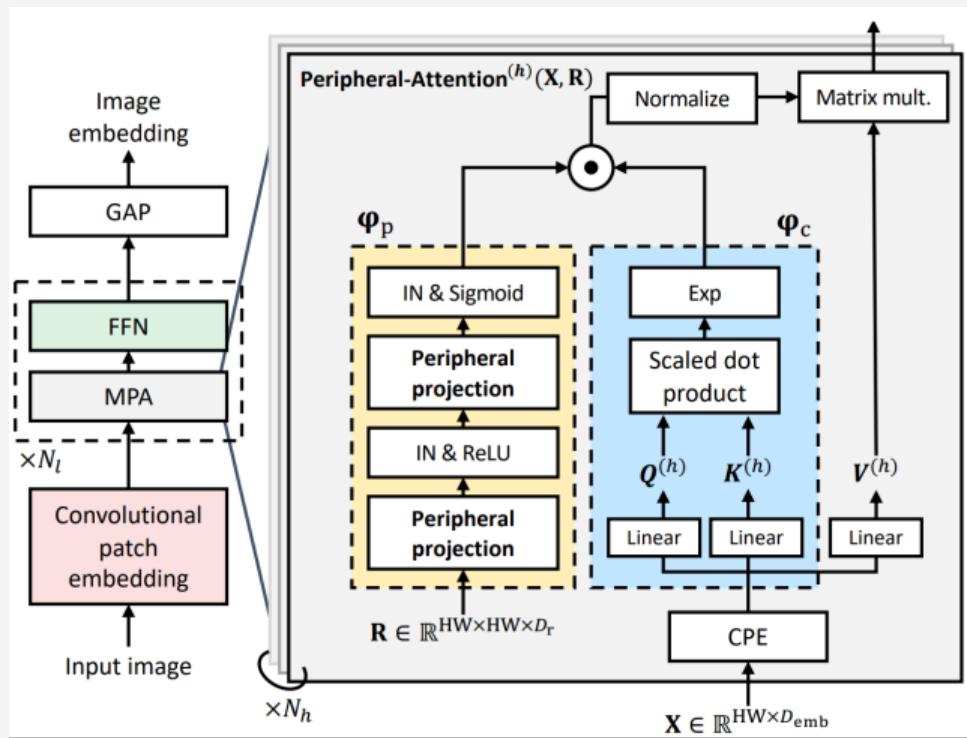
- 검정 통계량 : 1 ($1 < 5.02$)
- 귀무가설 채택
- 이상적인 모델 O

05

Conclusion

About.

프로젝트의 최종 결론을
제시합니다



최종 모델

VisionTransformer
WithPlantVillage

최고의 정확도

97.14%

검정 결과 귀무가설 채택, 이상적인 모델

06

Q&A

About.

궁금한 점에 대해 질문하는 시간입니다



Thank You

Reference

Dosovitskiy et al., (2020) "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE", Google Research, Brain Team.