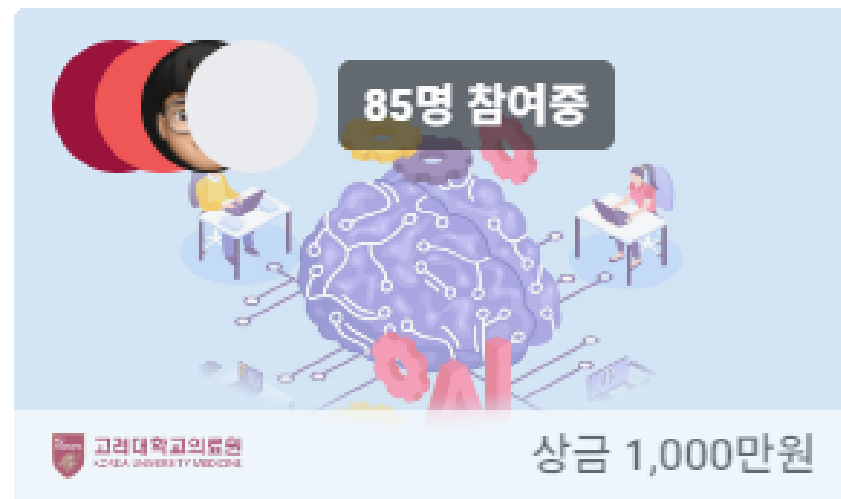


Group Project Presentation

최강 CV

발표자 송용호 | hello@reallygreatsite.com | +123-456-7890

공모전 주제



제1회 Medical AI (MAI) 경진대회

2024.10.02

D-28

경진대회 알고리즘 의료 유전자 비전

[배경]

인공지능(AI) 기술을 활용하여 사전에 제공된 의료 데이터를 기반으로 한 문제 해결을 목표로 하는 '제1회 MAI (Medical Artificial Intelligence) 경진대회'를 개최합니다.

이번 대회의 목표는 의료 분야에서 AI 활용의 저변을 확대하고, 참여자의 인공지능 역량을 강화하여 복잡한 의료 데이터를 효율적으로 분석하고 원하는 결과값을 산출할 수 있는 능력을 검증하는 것입니다.

이 대회는 의료 데이터의 활용도를 높이고, 이를 통해 의료 분야에서의 AI 기술 적용 가능성을 극대화하며, 인공지능 연구자들이 연구개발 의지를 더욱 고취할 수 있도록 기획되었습니다.

의료 데이터를 활용한 AI 기술이 실제 문제 해결에 어떻게 기여할 수 있는지 탐구하는 것이 목적입니다.

[주제]

H&E 염색된 조직 이미지로부터 유전자 발현 예측

[설명]

H&E 염색된 조직 이미지를 입력으로 받아, 해당 이미지에서 유전자 발현 데이터를 예측하는 AI 모델을 개발해야 합니다.

제공된 학습 데이터는 이미지와 유전자 발현 정보가 함께 제공되며, 이를 바탕으로 이미지와 유전자 발현 간의 관계를 모델이 학습해야 합니다.

평가 단계에서는 유전자 발현 정보가 제공되지 않는 새로운 이미지를 입력으로 받아, 이를 통해 유전자 발현 프로파일을 예측해야 합니다.

공모전 평가

평가 산식 : Pearson Correlation Coefficient (PCC)

$$\text{Score} = \max \left(\frac{\text{MeanCorr}_{\text{Cells}} + \text{MaxCorr}_{\text{Genes}} + \text{HEG}_{\text{Corr}} + \text{HVG}_{\text{Corr}}}{4}, 0 \right)$$

$\text{MeanCorr}_{\text{Cells}} = \frac{1}{n} \sum_{i=1}^n \text{Corr}(\mathbf{Pred}_i, \mathbf{True}_i)$: 샘플 i 에 대한 예측값과 실제값 사이의 상관관계수의 평균

$\text{MaxCorr}_{\text{Genes}} = \max(\text{Corr}(\mathbf{Pred}_j, \mathbf{True}_j))$: 유전자 j 에 대한 상관관계수 중 가장 큰 값

HEG_{Corr} : HEG 리스트에 속한 유전자들에 대한 상관관계수의 평균

HVG_{Corr} : HVG 리스트에 속한 유전자들에 대한 상관관계수의 평균

※ HEG, HVG 상관관계수 계산에 활용되는 유전자(Gene) 리스트는 공개하지 않습니다.

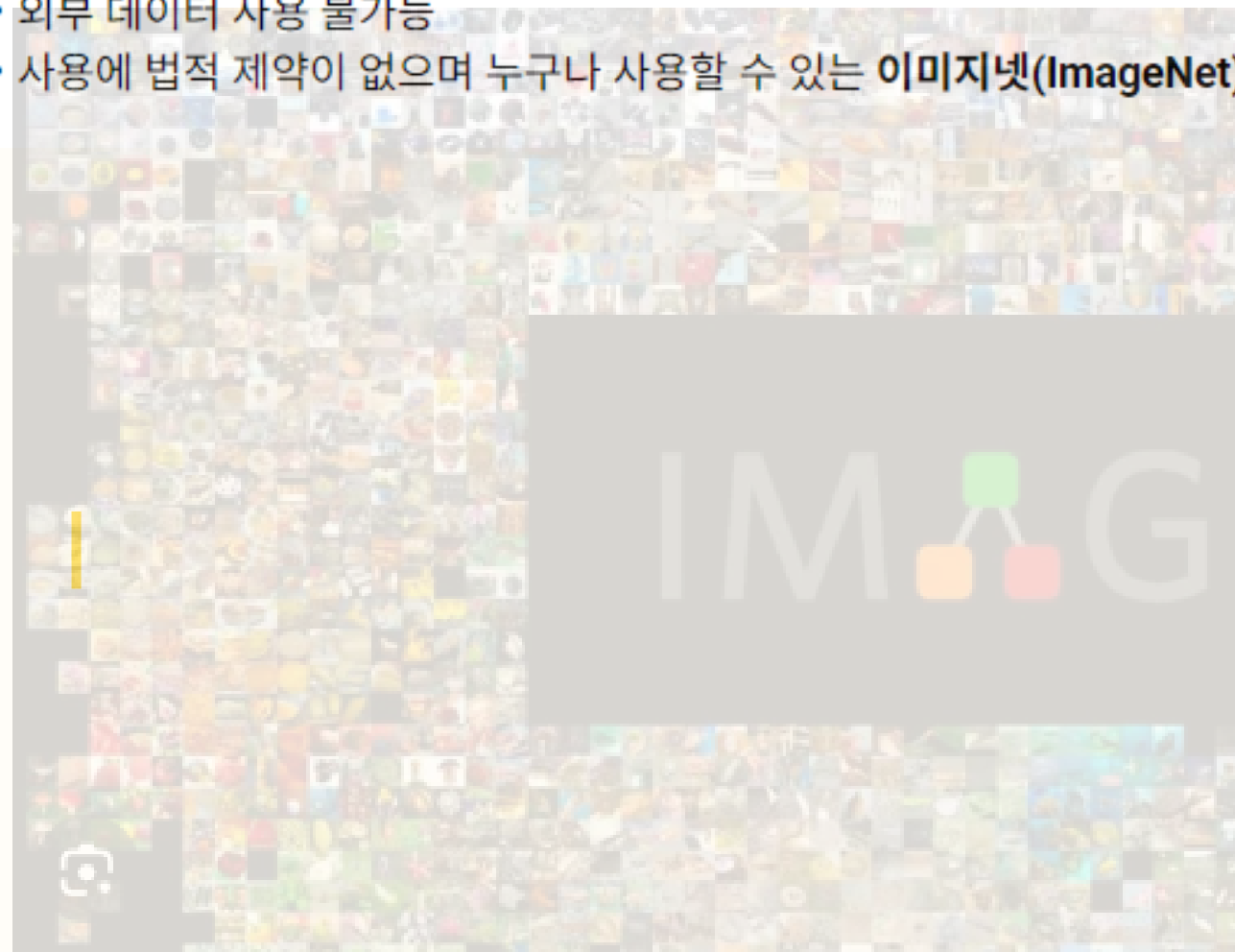
Public score : 전체 테스트 데이터 중 사전 샘플링된 50%

Private score : 전체 테스트 데이터 중 나머지 50%

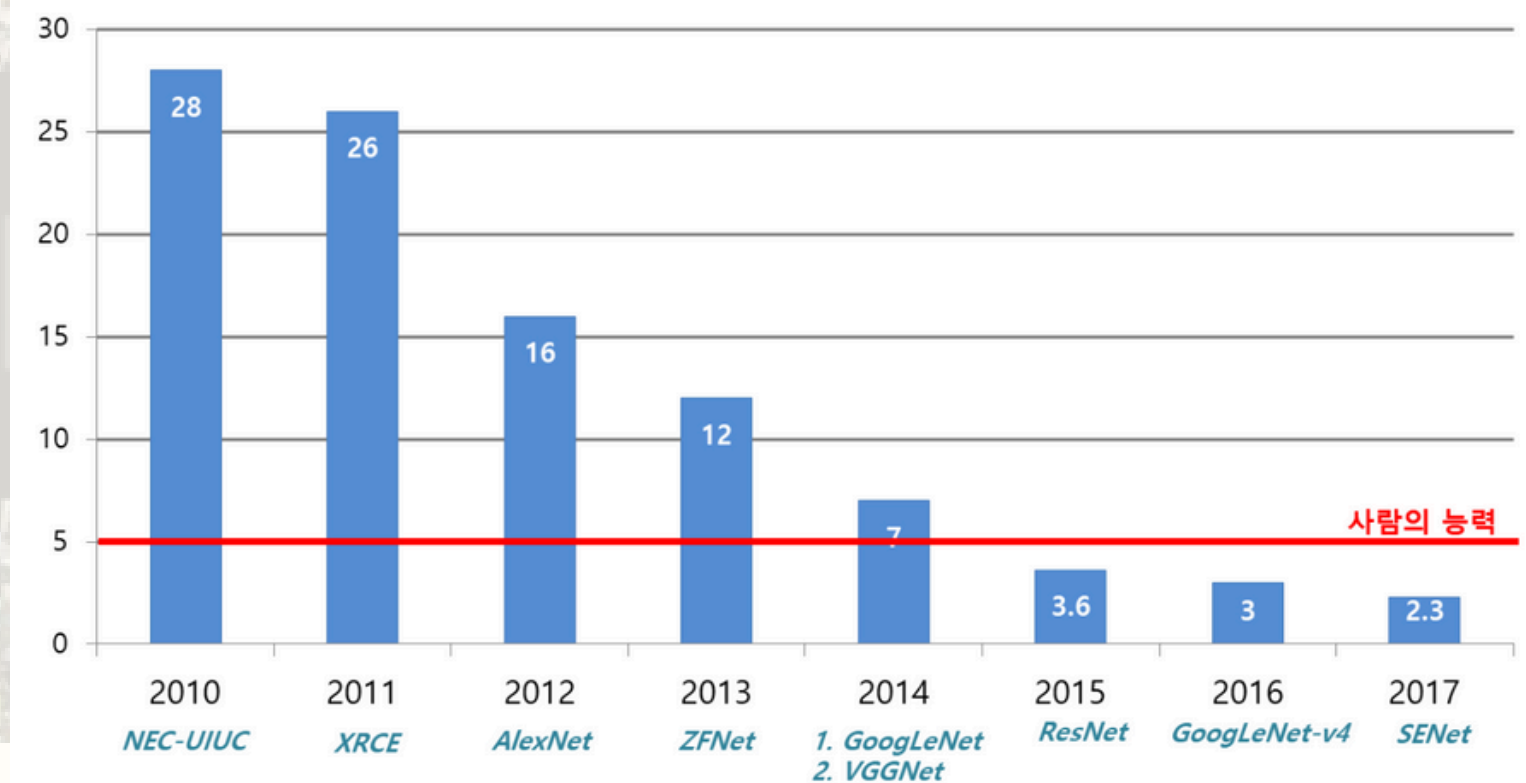
공모전 규칙

4. 외부 데이터 및 사전 학습 모델

- 외부 데이터 사용 불가능
- 사용에 법적 제약이 없으며 누구나 사용할 수 있는 **이미지넷(ImageNet) 데이터셋**으로 사전학습된 오픈소스 모델 사용 가능



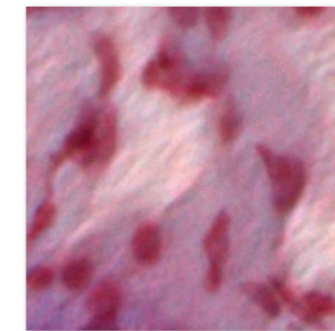
우승 알고리즘의 분류 에러율(%)



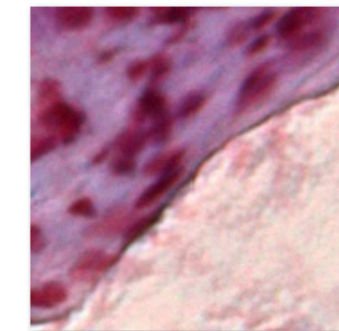
데이터 소개

DataSet Info

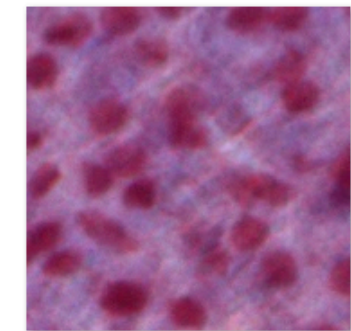
- 학습용 H&E 염색된 조직 이미지 샘플 6992개
- 평가용 H&E 염색된 조직 이미지 샘플 2277개
[train_csv]
 - ID : 샘플 ID
 - path : H&E 염색된 조직 이미지의 경로
- AL645608.7 ~ AL592183.1 : 각 유전자의 발현 정보 (유전자 총 3467개 존재)
[test.csv] :
 - ID : 샘플 ID
 - path : H&E 염색된 조직 이미지의 경로



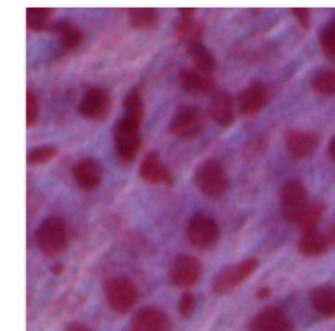
TRAIN_6990.png



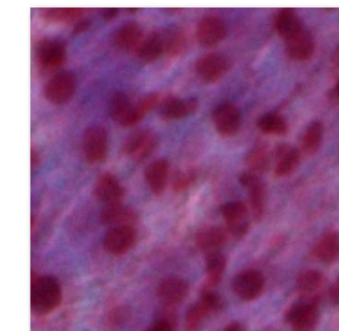
TRAIN_6991.png



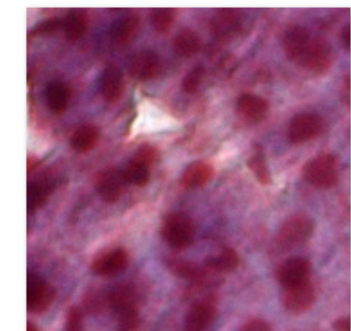
TRAIN_6965.png



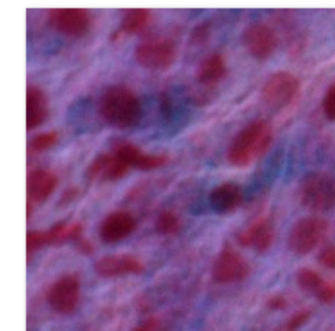
TRAIN_6971.png



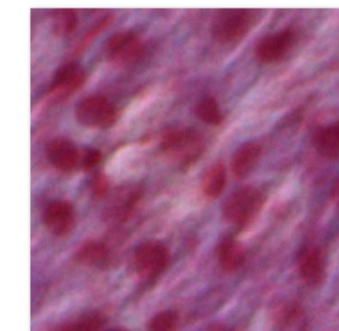
TRAIN_6972.png



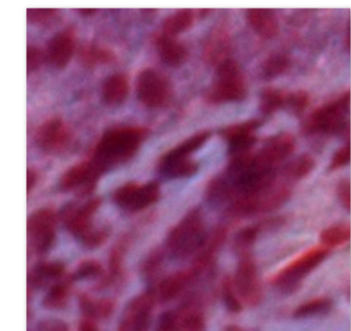
TRAIN_6973.png



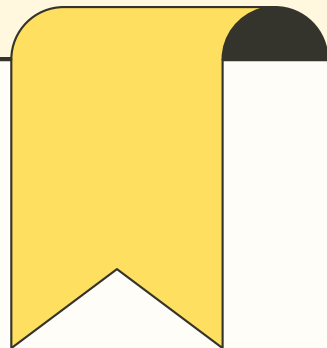
TRAIN_6979.png



TRAIN_6980.png



TRAIN_6981.png



대회 일정

Group Project Presentation

[세부일정]

- 대회 기간 : 2024년 10월 02일(수) 10:00 ~ 2024년 10월 31일(목) 10:00
- 팀 병합 마감 : 2024년 10월 24일(목) 23:59
- 대회 종료 : 2024년 10월 31일(목) 10:00
- 코드 및 PPT 제출 : 2024년 10월 31일(목) 12:00 ~ 2024년 11월 04일(월) 10:00
- 코드 검증 및 내부 평가 : 2024년 11월 04일(월) 12:00 ~ 2024년 11월 14일(목) 10:00
- 최종 수상자 발표 : 2024년 11월 14일(목) 10:00
- 오프라인 시상식 : 2024년 11월 15일(금) 고려대학교 의과대학 내 행사진행

프로젝트 주제

제목 | 너 어디야?

이미지 인식 후 해당 이미지가 어디에서 촬영되었는지 알아내는 프로젝트입니다.

학교 내 장소에 대해 잘 모르는 신입생, 복학생들에게 유용할 수 있습니다.

사진을 통해 현재 자신의 위치를 알아내거나 친구끼리 위치를 공유할 수 있습니다.

프로젝트 목표

프론티어관, 상상관, 미래관에서 찍힌 이미지를 보고 장소 분류

| 데이터 수집

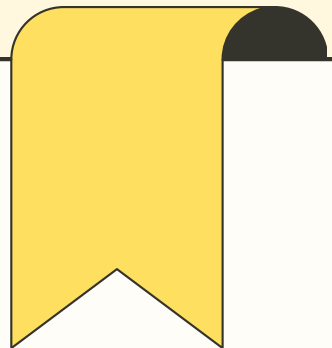
직접 이미지 촬영 후 수집

| 데이터 처리

이미지 크기 조정 및 전처리

| 모델 튜닝

모델 학습 및 파라미터 튜닝

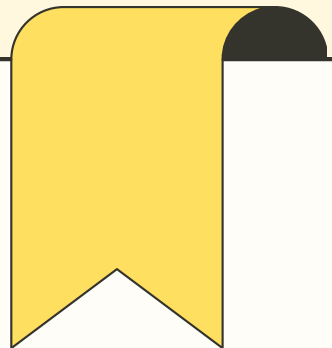


Group Project Presentation

Q & A

프로젝트에 대해 궁금한 점은 질문해주세요

발표자 이수진 | hello@reallygreatsite.com | +123-456-7890



Group Project Presentation

감사합니다

발표자 이수진 | hello@reallygreatsite.com | +123-456-7890