

데이터 분석가

손소연

✉ 090484@naver.com

☎ 010-5703-9196

💻 <https://github.com/eejuais>



📁 경력사항

2022.11

● (주) 레드테이블 (인턴)

- 다국어 데이터 가공, 수집, 입력
- 데이터 전처리 및 데이터 가공 및 검수

2020 - 2022

● 달콤아이스크림 (자영업)

- 아이스크림 점포 운영
- 키오스크 매출 데이터 분석을 통한 상품 구성 관리

2021 - 2022

● (주) 에이치알엔

- 11번가 기프트콘 서비스 운영
- 기프트콘 브랜드별 실적관리
- 기프트콘 정산 데이터 관리

🎓 학력사항

2021 - 2024

● 서강대학교 AI·SW대학원

- 데이터사이언스·인공지능 석사 졸업
- 석사 학위논문 제출 “Vision Transformer를 활용한 반려동물 안구질환 이미지 분류” (2024년 7월)

2009 - 2014

● 동덕여자대학교

- 문헌정보학과 학사 졸업
- 일본어과 복수전공

🏆 수상내역

2022 공공데이터 아이디어 해커톤 우수상

- NIA 주관 해커톤 <‘같이’ 높여요, 공공데이터의 ‘가치’>
 - 기업 정보 데이터를 활용한 자립준비청년 취업매칭 앱 서비스 “동반JOB” 아이디어 제안

👤 ABOUT ME

- 일의 과정에서 재미를 찾는 분석가
- 커뮤니케이션에 강점을 가진 분석가
- 동료와 함께 발전해가는 분석가
- 목표 달성을 위한 도전에서 즐거움을 느끼는 분석가

📄 자격사항

- 빅데이터분석기사 (2023.07)
- SQLD(SQL개발자) (2023.04)
- 정보처리기사 (2013.12)

💡 SKILLS

- python
 - pandas / numpy / matplotlib / seaborn 등 다양한 라이브러리 활용하여 데이터 전처리, 탐색적 데이터 분석(EDA), 통계적 분석 및 시각화
 - PyTorch를 이용한 딥러닝 모델 구축 및 모델 학습, 평가
 - 다양한 컴퓨터 비전 및 자연어처리 프로젝트 수행
- SQL
 - DB에서 필요한 데이터를 추출하고 간단한 분석을 수행할 수 있으며 기본적인 쿼리 작성 및 데이터 관리
- 가설 검증에 필요한 통계적 지식



프로젝트 개요

- 기간 : 2024.03-2024.05
- 목적 : Vision Transformer 모델을 활용해 반려동물의 안구질환 이미지를 분류
- 결론 : 사전 학습된 ViT 모델을 로드하여 전이학습을 통해 반려동물의 4개 질환에 대한 분류 모델을 학습, 검증하여 평균 91.35%의 테스트 정확도를 도출
- 개발환경 : Google Colab, Python pytorch

프로세스

STEP 1. 문제정의

반려묘의 안검염과 결막염, 반려견의 귀양성 각막질환과 백내장의 각각 4개 질환에 대한 분류 모델을 학습, 검증

STEP 2. 데이터 수집

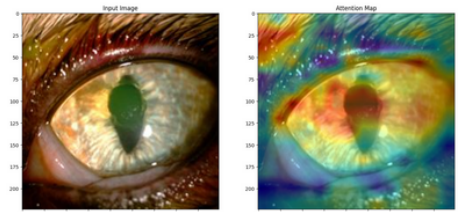
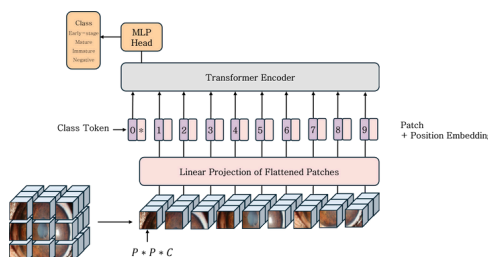
AI-Hub의 반려동물 안구질환 데이터 활용.
해당 이미지 데이터 57,328장을 사용

STEP 3. 데이터 전처리

반려묘의 데이터가 반려견에 비해 상대적으로 부족하여 Augmentation 기법 적용하여 이미지에 다양한 무작위 변형을 적용
Vision Transformer 모델에 맞게 입력 이미지 리사이징

STEP 4. 데이터 모델링

ViT-Base 모델 적용하여 전이학습 수행
드롭아웃 기법으로 과대적합 방지 / Optuna를 활용한 파라미터를 최적화
각 안구질환 이미지 샘플을 Attention Heat Map 시각화



STEP 5. 결론

평균 91.35%의 테스트 정확도, 질환 별로 편차가 크지 않은 고른 성과 도출.
전통적인 합성곱 신경망 계열 모델과의 비교 분석을 통해, 반려동물의 안구질환 이미지를 분류하는 데 ViT-Base 모델이 ResNet50과 EfficientNet-B5 모델에 비해 빠르고 안정적으로 학습을 진행할 수 있으며, 정확도 면에서 효율적이라는 결과 도출

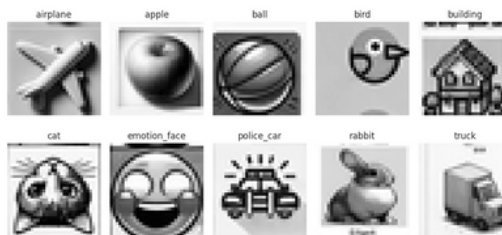
프로젝트 개요

- 기간 : 2025.04.15 ~ 2025.04.30
- 주제 : 아이콘 이미지 분류 AI 알고리즘 개발
- 목표 : 32*32 흑백 픽셀 데이터를 기반으로 아이콘의 종류 분류
- 개발환경 : Google Colab, Pytorch, timm

프로세스

STEP 1. 문제정의

- 32×32 Grayscale 이미지 데이터를 기반으로 10개 아이콘 종류 분류 모델 설계

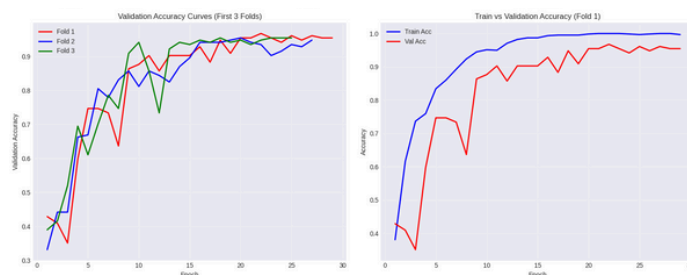


STEP 2. 데이터 전처리

- 흑백 이미지 → RGB 채널 변환 후 PIL 이미지 객체로 처리
- 약한 증강 기법 적용: RandomHorizontalFlip
- 정규화: Normalize(mean=[0.5], std=[0.5])

STEP 3. 모델 구조 및 학습 전략

- densenet121 모델 사용 (사전학습 미사용, timm 라이브러리 활용)
- 하이퍼파라미터 랜덤서치(총 10조합 실험)
- Stratified K-Fold 교차검증 (n=5)



STEP 4. 성과

- Public Score: 0.96
- Private Score: 0.97
- 리더보드 순위: 상위 10%
- 데이터 증강 및 랜덤서치를 통한 최적 모델 도출