



AI ARTWORK

AI 생성 그림 판별 프로젝트

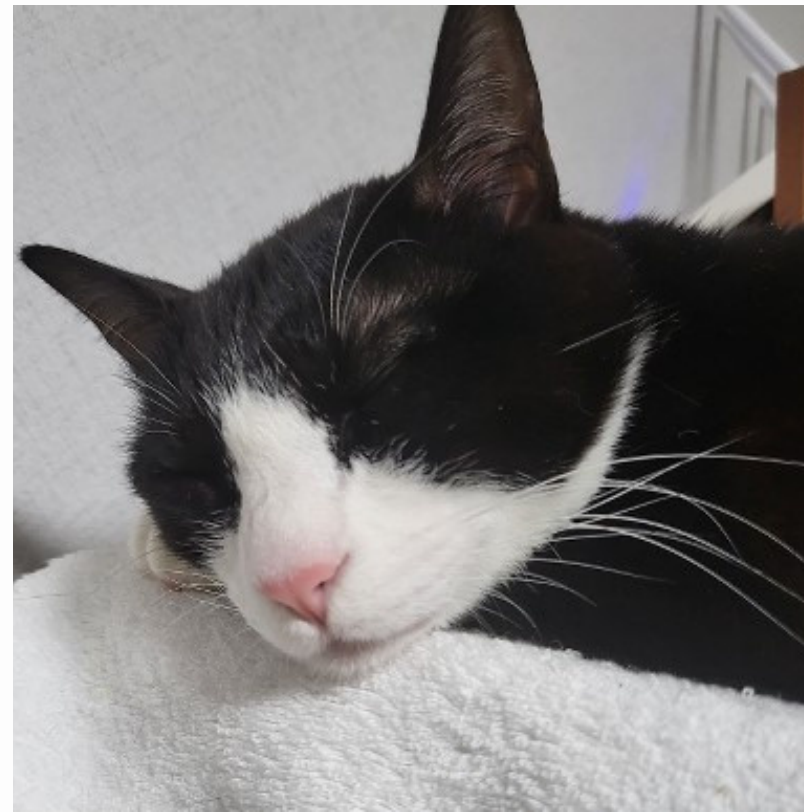
AIB18 장혜수

목 차

- 01 개요
- 02 데이터 준비
- 03 모델링
- 04 결론

01. 개요

문제 정의 및 목적



- 최근 AI 기술의 발전으로 인해 매우 현실적인 이미지를 생성할 수 있게 되었고, 이로 인해 AI가 생성한 이미지와 실제 이미지를 구분하는 것이 어려워졌다.
- 그림 대회 평가에 있어서 AI 출품작을 걸러내는 문제를 해결하려는 상황 가정

AI가 생성한 그림과 사람이 그린 그림을 구분할 수 있는 딥러닝 모델을 개발하고자 함

01. 개요




가설 설정






사람이 판별하기 힘든 가짜 그림을 딥러닝 모델로 구분할 수 있을까?

02. 데이터 준비

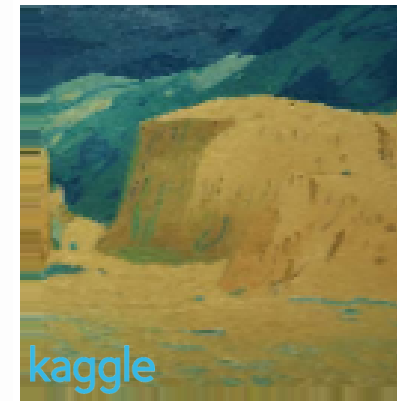
데이터 선정 이유 및 설명

 train <  FAKE - 1205개
 REAL - 1205개

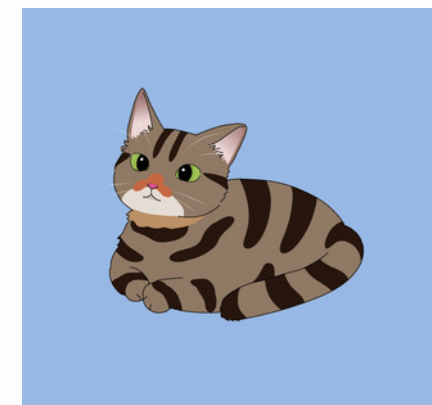
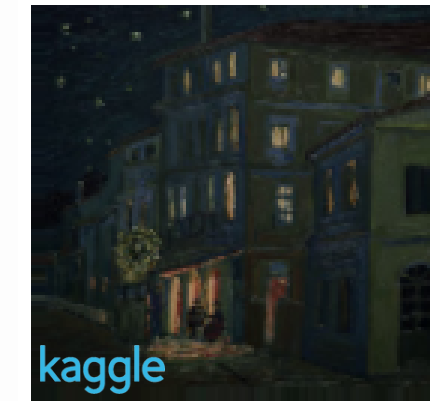
 validation <  FAKE - 500개
 REAL - 500개

 test <  FAKE - 152개
 REAL - 152개

가짜



진짜

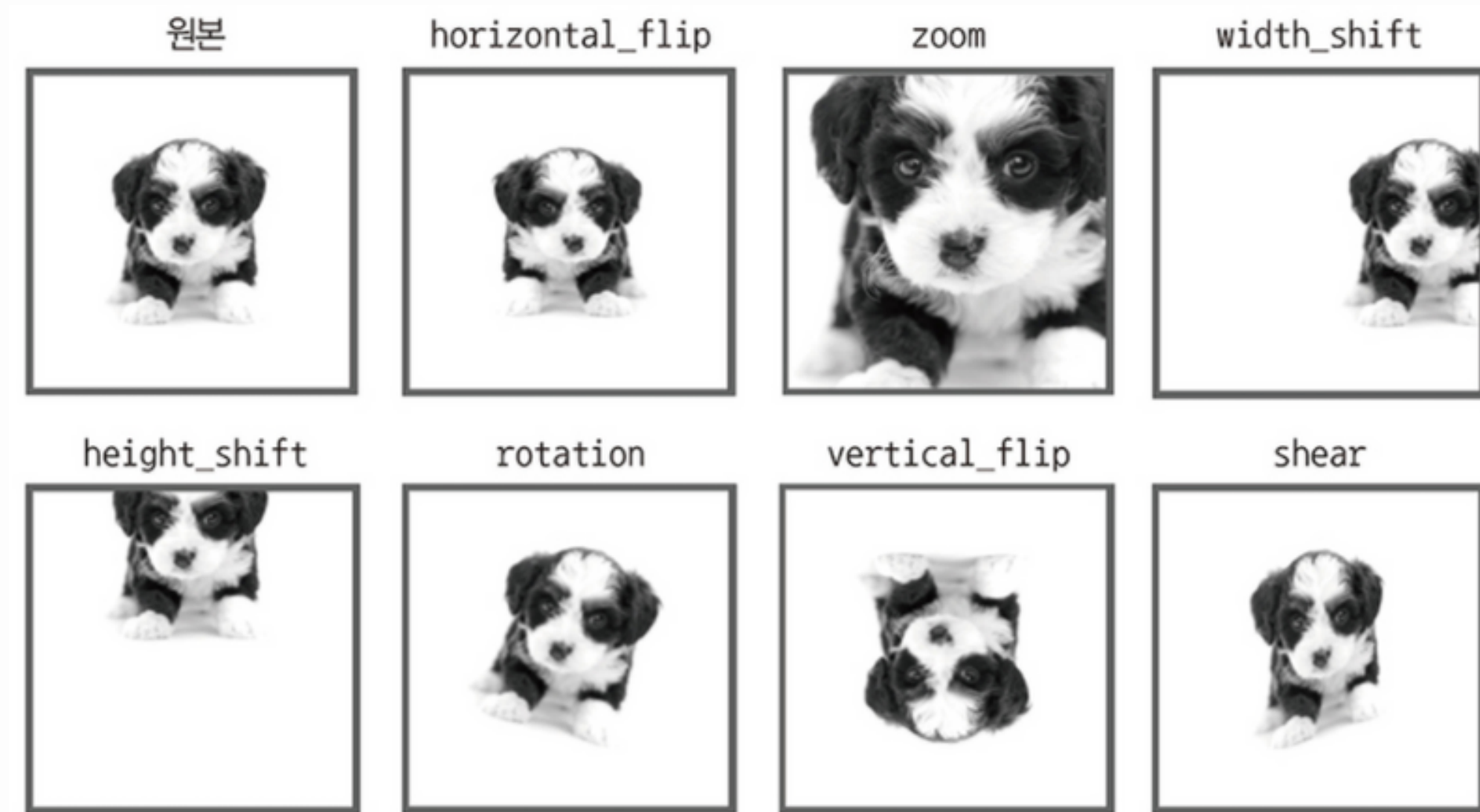
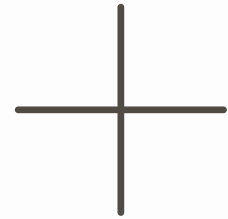


02. 데이터 준비

ImageDataGenerator

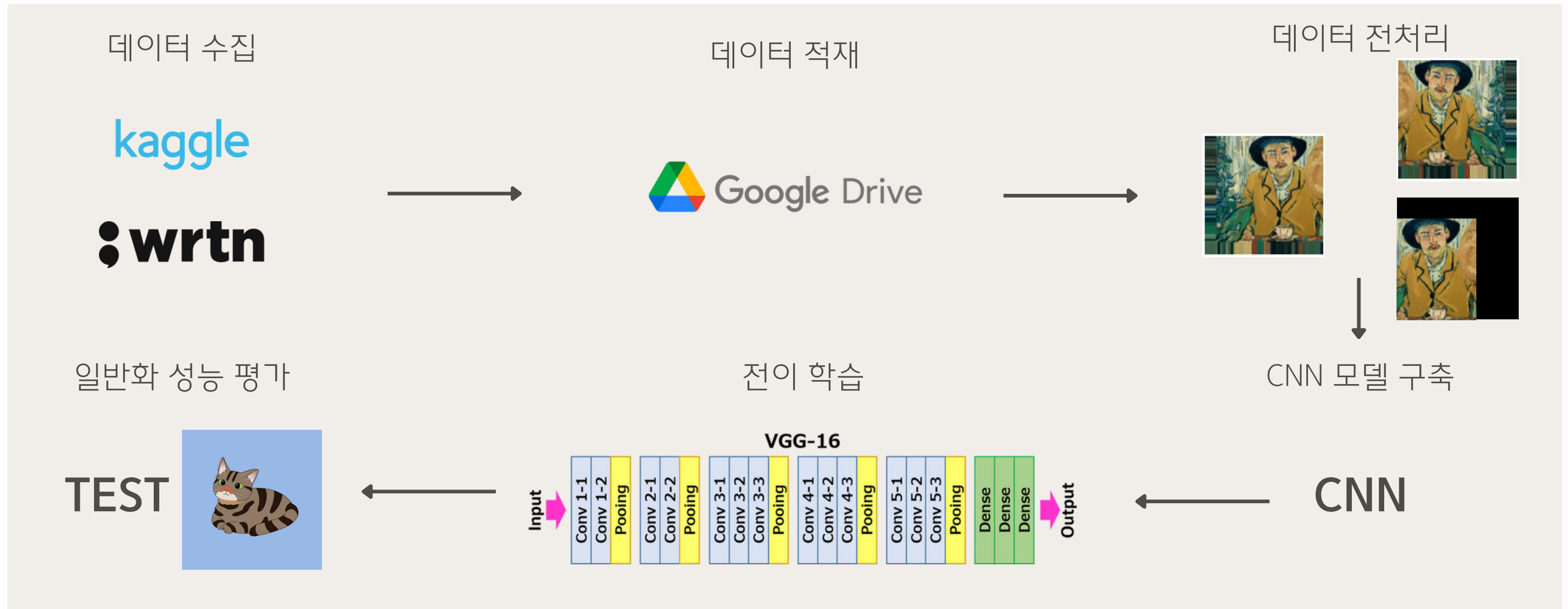
- 다양한 상황의 이미지를 잘 인식하도록 이미지 추가 생성

정규화



03. 모델링

딥러닝 파이프라인



03. 모델링

CNN, VGG16 모델 선택 이유

- CNN(Convolutional Neural Networks)

이미지 처리 작업에 대표적인 알고리즘

이미지의 공간적 구조를 이해하는 데 유리한 구조

이미지의 작은 영역에서 특징을 학습하고 이러한 지역 특징을 활용하여 더 큰 패턴을 학습하는 방식으로 작동

- VGG16(Visual Geometry Group, 16 layers)

2014 이미지 인식 대회(ILSVRC)에서 2번째로 높은 성능

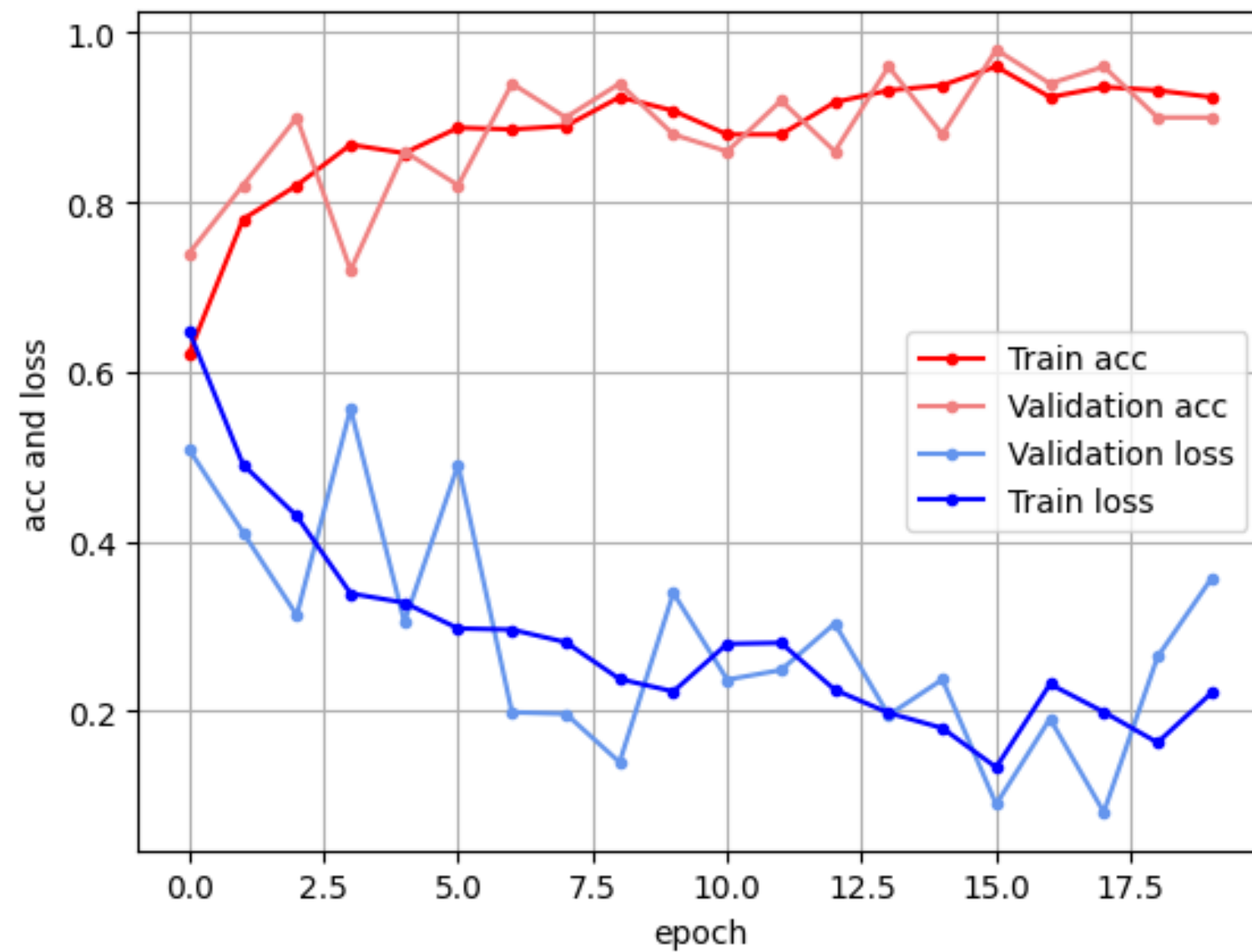
이미지 처리 작업에 대한 강력한 성능으로 잘 알려진 특정한 CNN 아키텍처

13개의 컨볼루션 레이어와 풀링 레이어를 통해 이미지에서 저수준 특징에서 고수준 특징까지 추출하고, 이를 바탕으로 이미지를 분류

03. 모델링

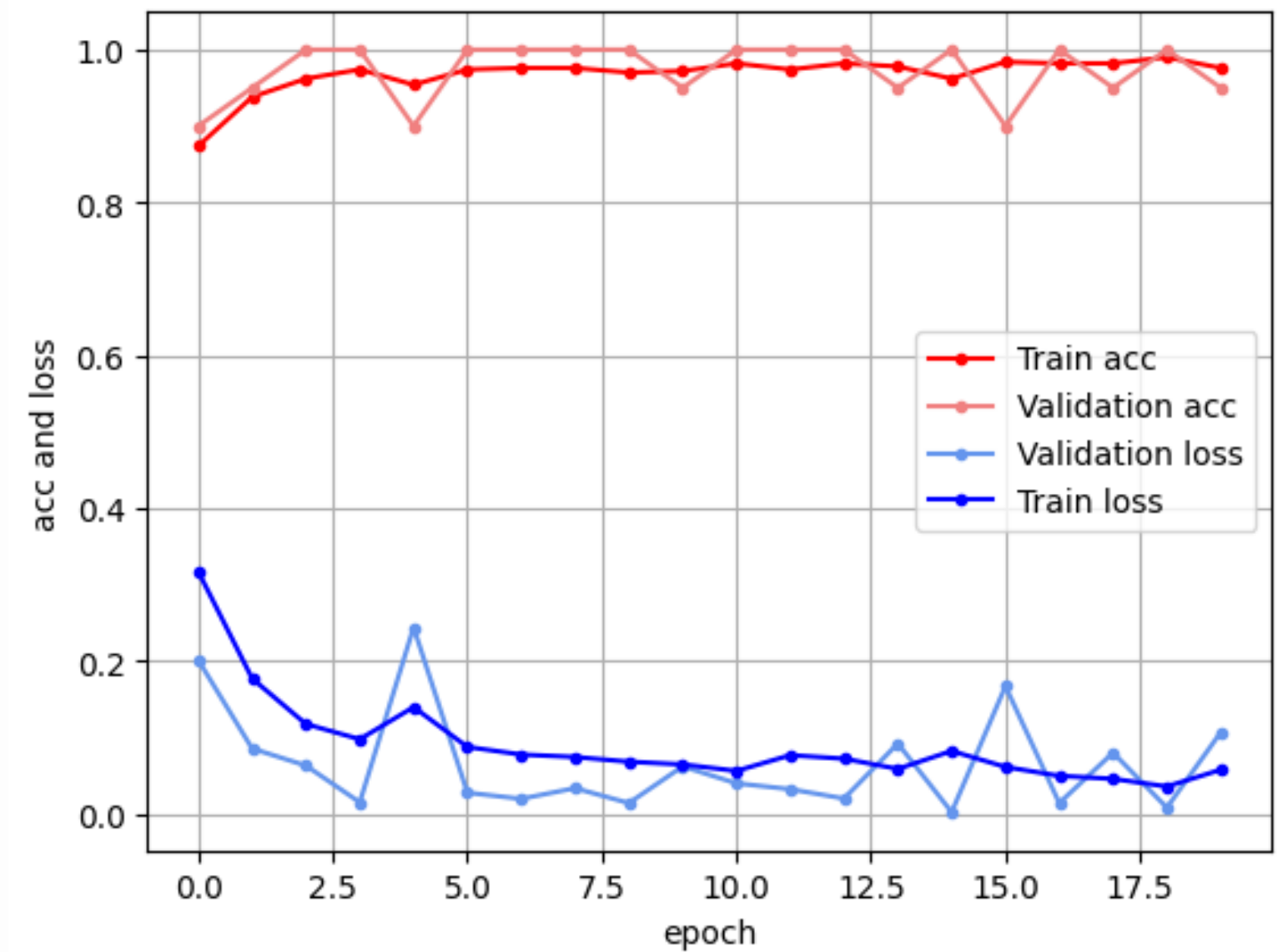
CNN, VGG16 모델 학습

CNN



val_loss: 0.3564 - val_accuracy: 0.9000

VGG16



val_loss: 0.1061 - val_accuracy: 0.9500

03. 모델링

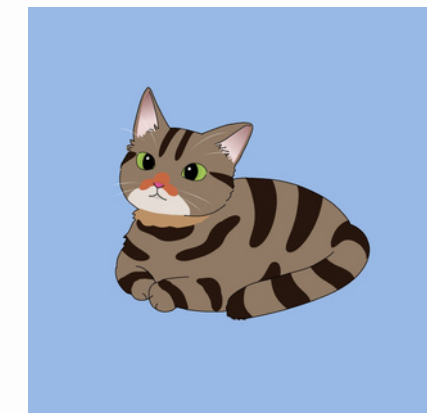
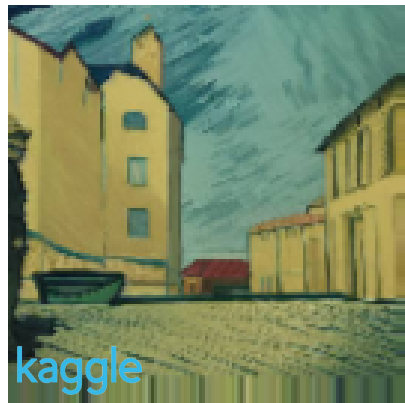
전 이 학습 모델 평가

```
# 테스트 데이터셋에 대한 성능 평가  
test_loss, test_accuracy = finetune_model.evaluate(test_generator)
```

```
61/61 [=====] - 105s 2s/step - loss: 3.7629 - accuracy: 0.5263
```

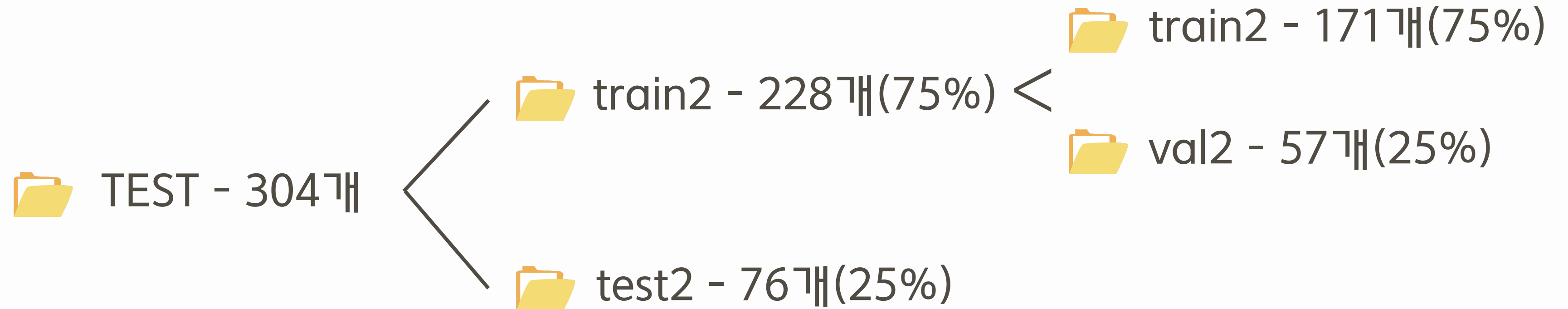
- Chance Level인 0.5와 비교하여 성능이 좋지 않음

→ 학습한 그림들과 평가할 그림이 너무 달라서 그런걸까?



03. 모델링

새로 학습 및 평가



epoch 1

```
loss: 0.1773 - accuracy: 0.9200 - val_loss: 0.0034 - val_accuracy: 1.0000
```

테스트 데이터셋에 대한 성능 평가

```
test_loss, test_accuracy = finetune_model2.evaluate(test_generator2)
```

```
8/8 [=====] - 9s 1s/step - loss: 0.0036 - accuracy: 1.0000
```

03. 모델링

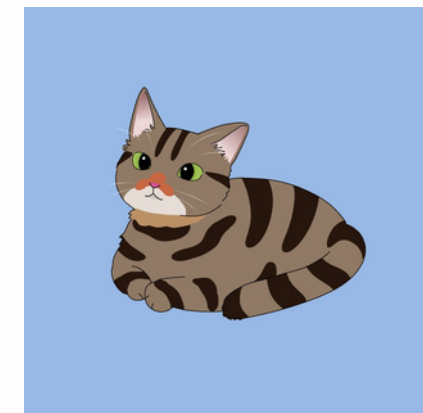
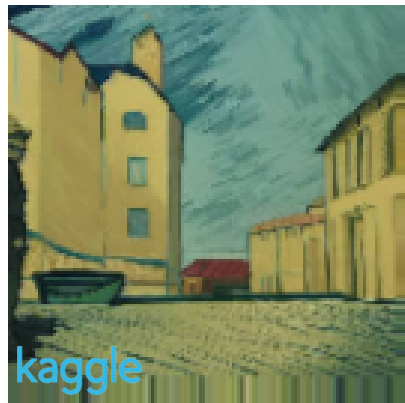
첫 전이학습 모델 평가

```
# 테스트 데이터셋에 대한 성능 평가
test_loss, test_accuracy = finetune_model.evaluate(test_generator)
```

```
61/61 [=====] - 105s 2s/step - loss: 3.7629 - accuracy: 0.5263
```

- Chance Level인 0.5와 비교하여 성능이 좋지 않음

→ kaggle 데이터에 과적합된 것으로 판단



04. 결론

문제 해결 및 가설 해소

문제: AI가 생성한 그림을 사람이 판별하기 매우 어렵다.

- 딥러닝 모델을 통해 검증 정확도 95% 정도로 AI 생성 그림과 사람이 그린 그림을 분류할 수 있었다.
- 온라인에서 공유되는 이미지의 진위를 판별하는데 도움이 될 것이다.

04. 결론

한계점 및 향후 계획

- 데이터셋의 편향으로 인해 다른 스타일의 이미지를 구분하는 데에 어려움
→ 다양한 스타일의 이미지를 포함하는 풍부한 데이터셋을 수집
- 모델 성능을 개선하기 위해 추가적인 데이터셋을 수집하고, 최신 AI 이미지 생성 기술을 계속 모니터링하여 새로운 기술이 도입된 이미지를 구분할 수 있는 능력을 모델에 반영할 수 있도록 함



감사합니다