

IMAGE COLORIZATION

흑백 이미지 컬러화 프로젝트

최진영

이OO

오OO

이OO

목차

0. 팀원 소개 및 역할 분담
1. 프로젝트 목표
2. 프로젝트 진행 방식
3. 1차 모델링 - Pix2Pix
4. 2차 모델링 - CWGAN
5. 3차 모델링 - U-Net
6. 결론

1.

프로젝트 목표



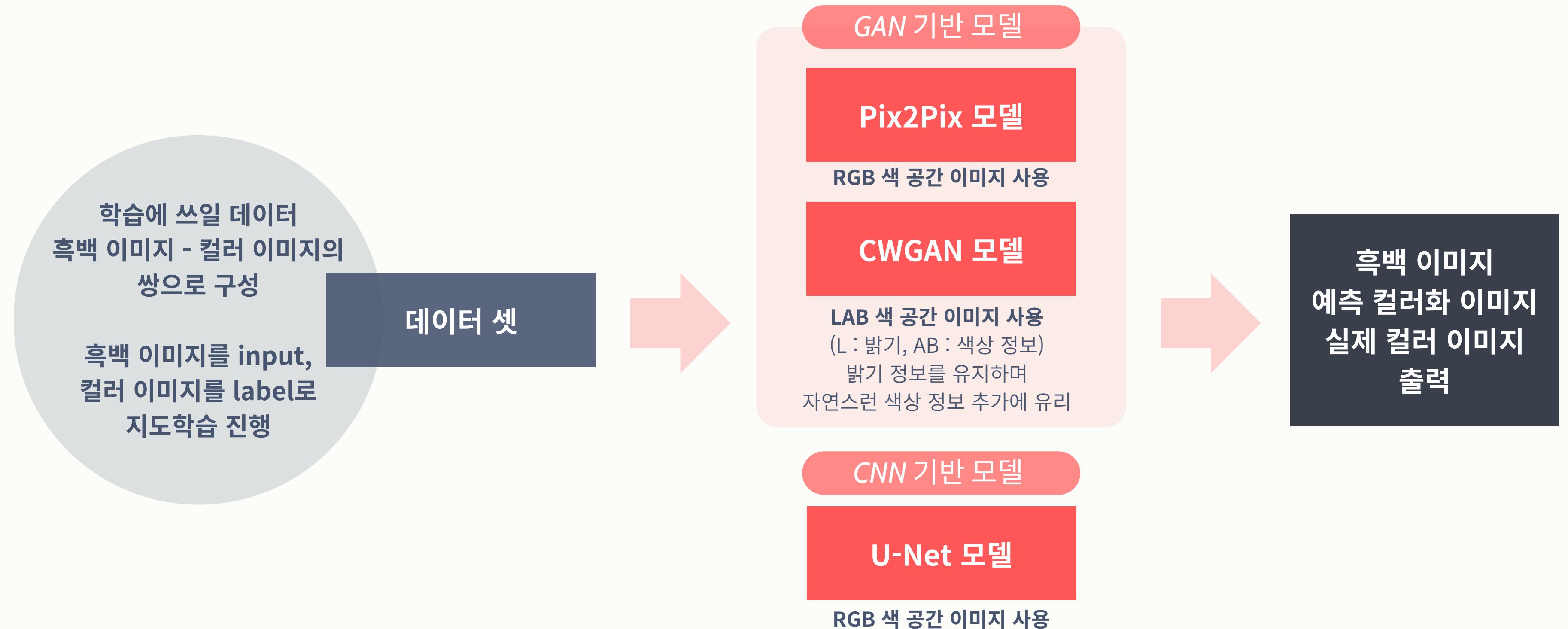
흑백 이미지를 컬러 이미지로 복원함으로써

개인적 측면에서는
추억을 되살릴 수 있고

사회적 측면에서는
역사 사료를 현대적으로 재조명함으로써
사회적 가치를 창출할 수 있다.

2.

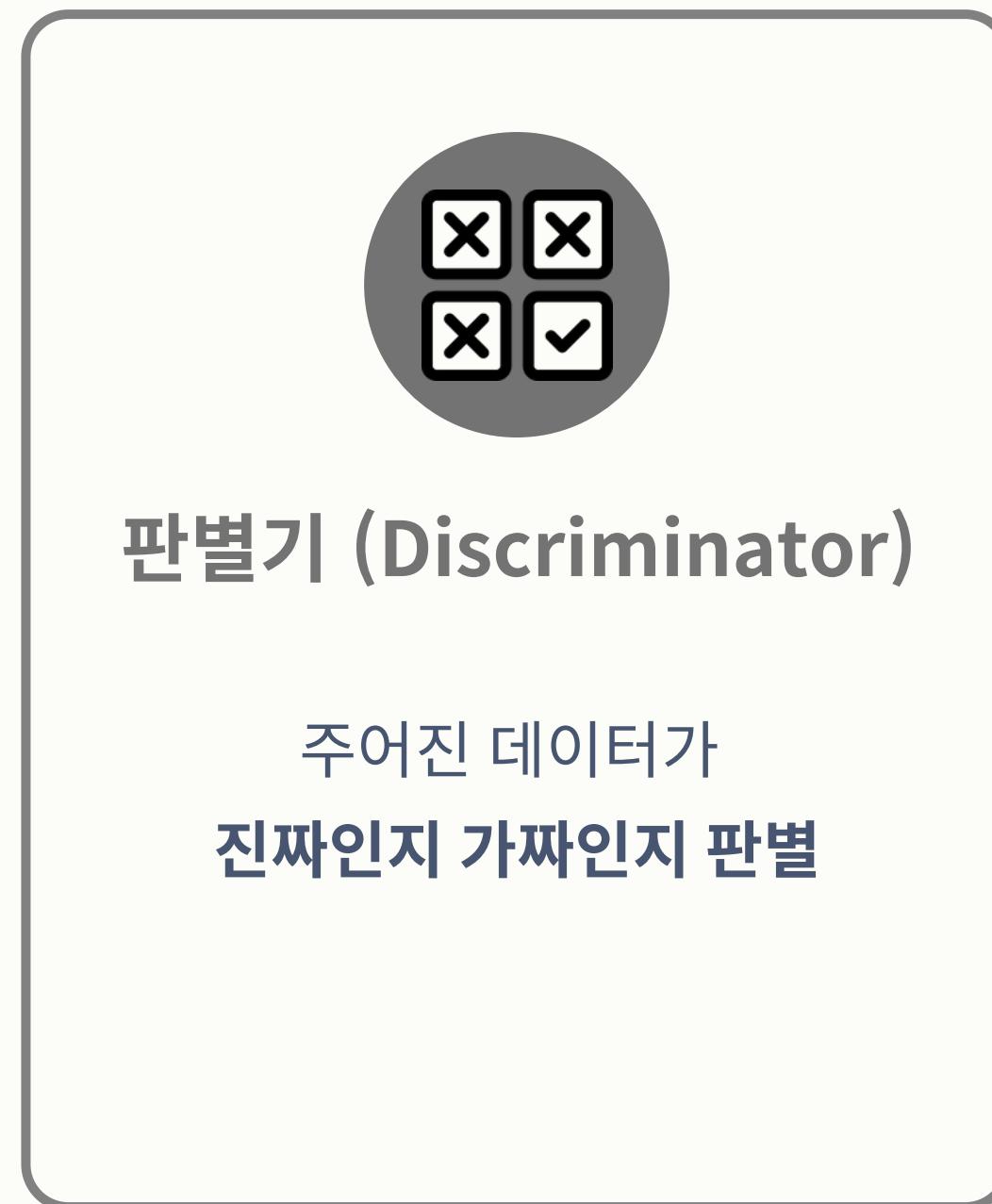
프로젝트 진행 방식



모델 소개 - GAN

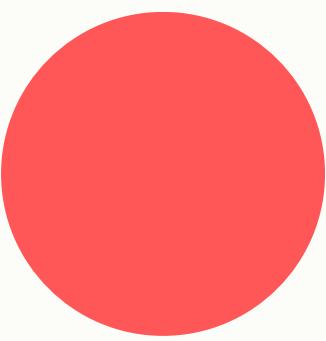


두 네트워크가
서로 경쟁하며 학습하기에
target에 가까운 데이터를
생성할 수 있다



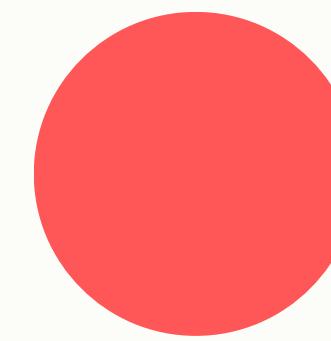
2. 프로젝트 진행 방식

모델 소개 - Pix2Pix & CWGAN



Pix2Pix

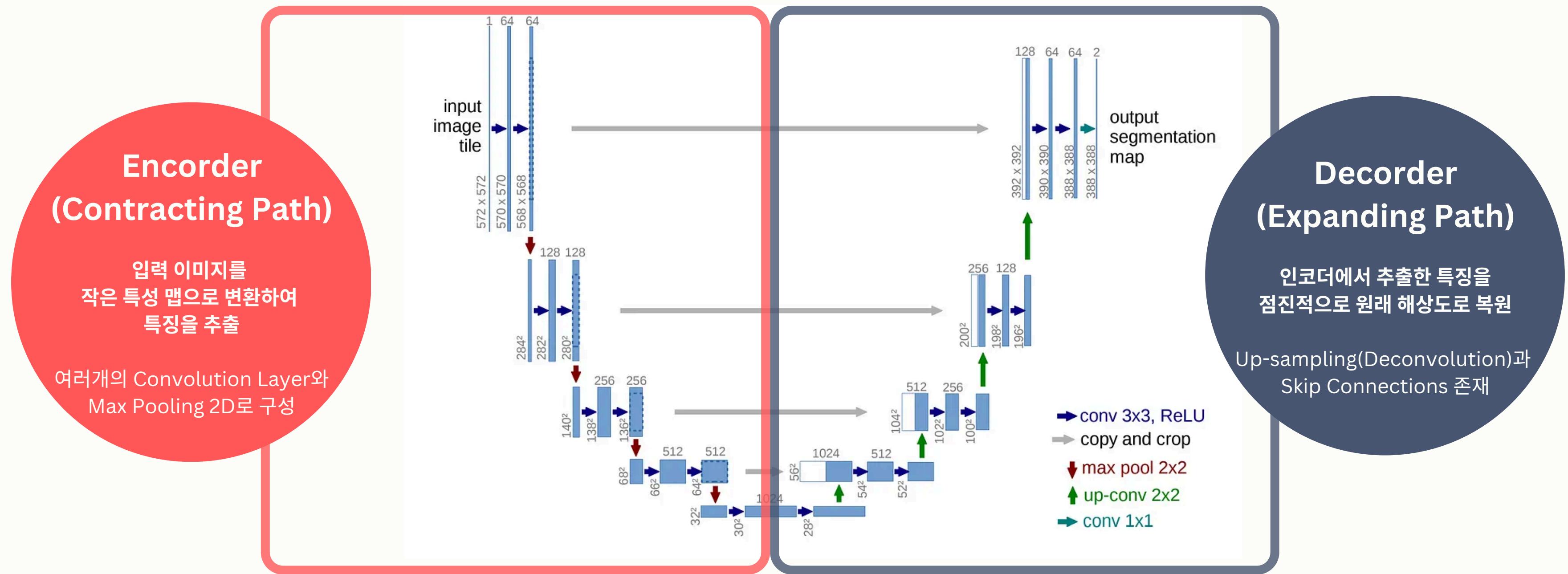
Condition GAN 모델 기반의
이미지-이미지 변환 모델
GAN 손실 (Binary Cross-Entropy
Loss), L1 손실을 주로 사용



CWGAN

Condition Wasserstein GAN
'Wasserstein 거리'
(두 확률 분포간의 거리)를 critic이
계산하여 생성자가 만든 데이터가
조건에 맞는지 구별한다.

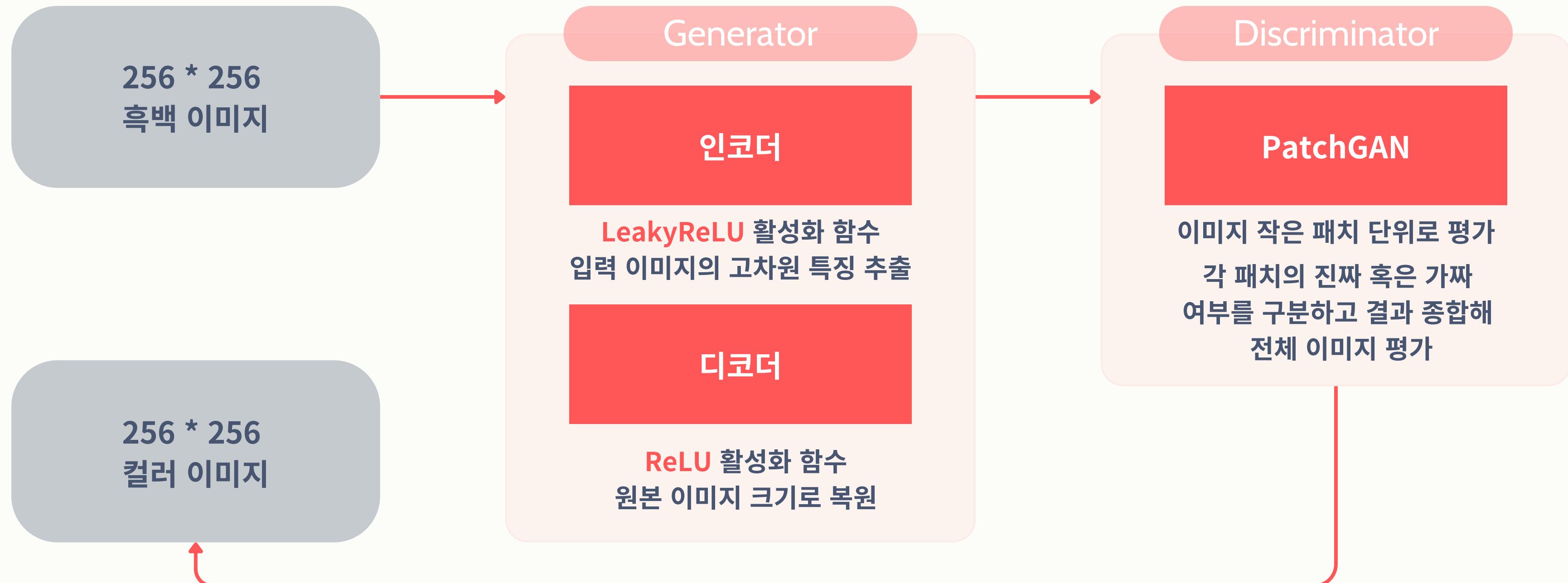
모델 소개 - U-Net



1차 데이터셋

- 데이터 형식: PNG 파일
- 해상도: 256 * 256 픽셀
- 데이터 수량: 10개 쌍

모델 구조



모델 평가

Real



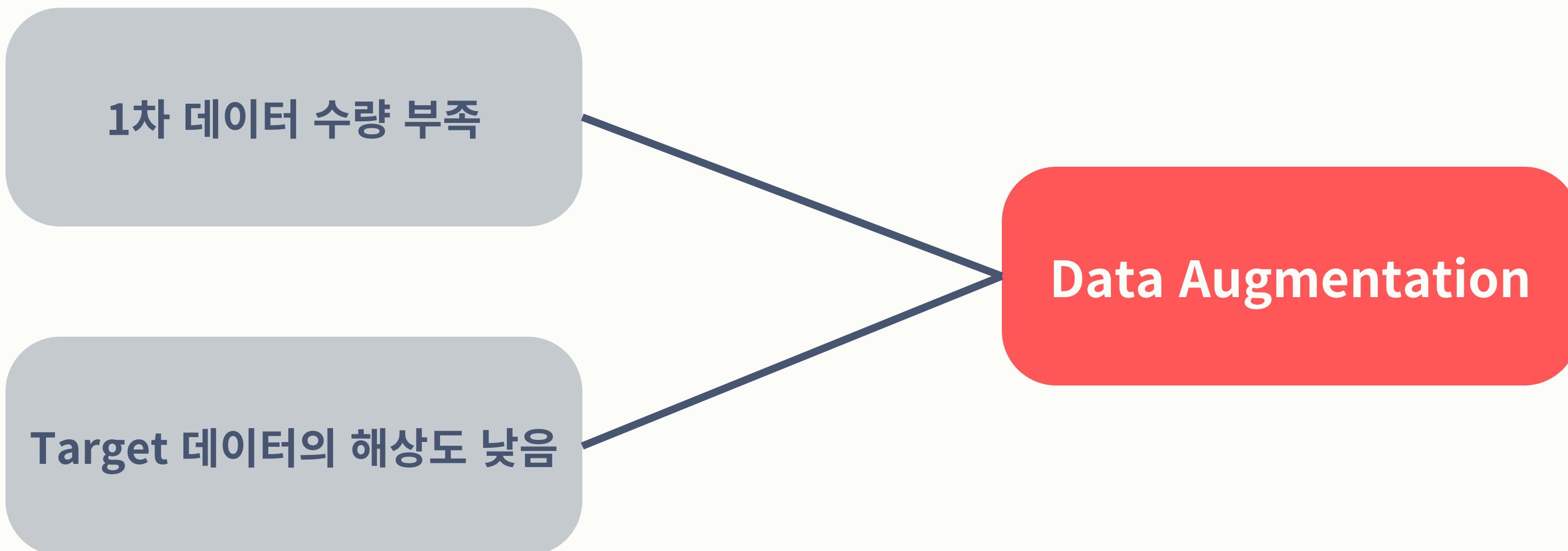
Generated



특정 부분 컬러화 잘 진행됨

낮은 해상도
파란색과 초록색 부분의 컬러화 미흡

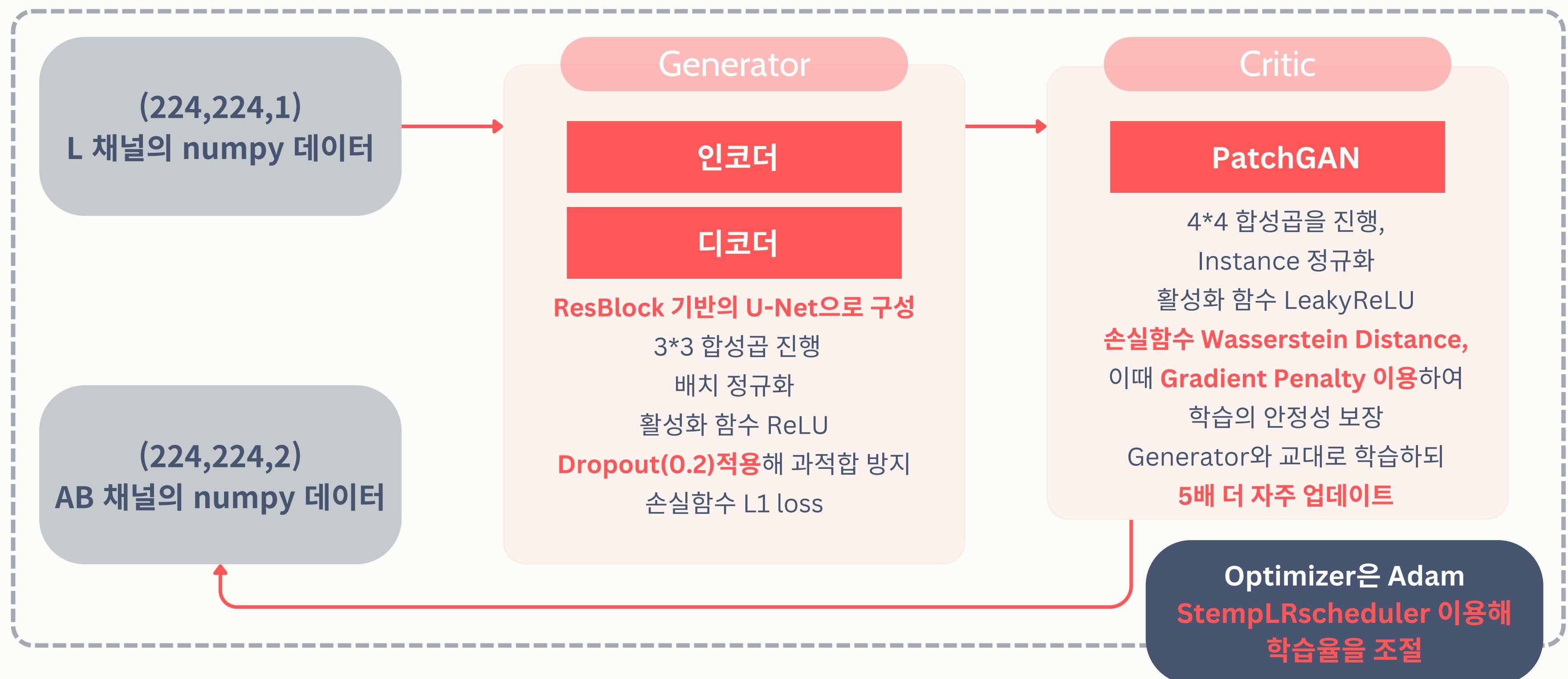
문제 분석



2차 데이터셋

- 데이터 형식: LAB 색 공간으로 정의된 numpy 데이터 셋
- 해상도: 224 * 224 픽셀
- 데이터 수량: 25000 쌍

모델 구조



이전 모델링과의 차이

1차

Generator가 일반적인
Conv2D와 TransposeConv
기반의 U-Net으로 구성

손실함수는 GAN loss와 L1 loss

Generator와 Discriminator
동일한 빈도 업데이트

학습률 스케줄러, Dropout 없음

2차

ResBlock 기반의 U-Net 구성

Wasserstein Distance와
Gradient Penalty,
L1 loss 사용

Critic의 업데이트 빈도
Generator의 5배

학습률 스케줄러, Dropout 적용

모델 평가



Critic에서 쓰인 손실 함수의 값이
음수가 나오는 경우 발생

생성된 사진과 target 사진 차이 O
완벽한 컬러화 되지 않음

모델 개선



Loss 값이
음수

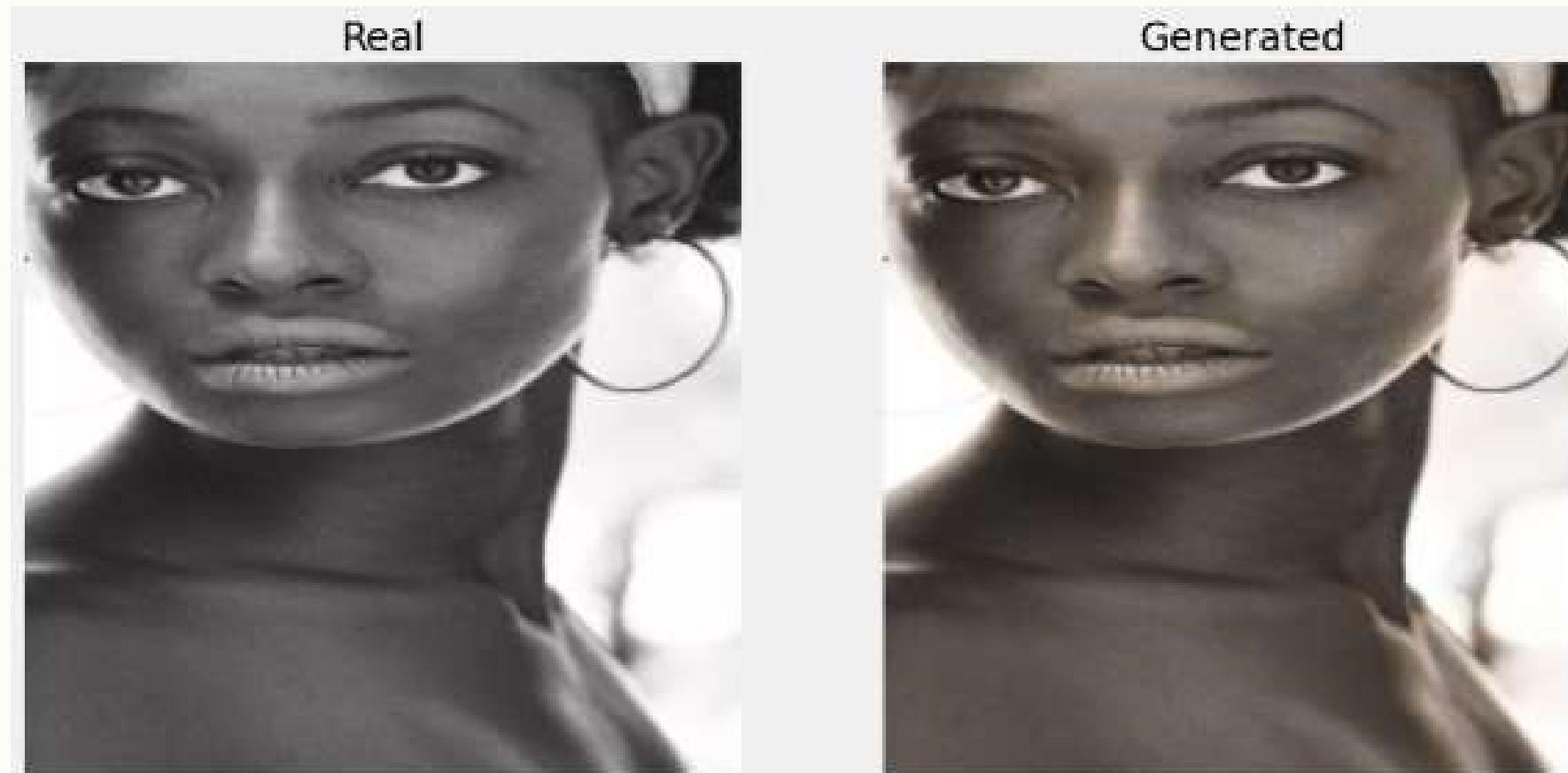
Wasserstein 거리 정의에
절댓값 적용

컬러화 미흡

VGG를 이용해
Perceptual loss 추가
L1 loss를 이용해
color loss 추가

가중치 변경

한계 및 개선할 점



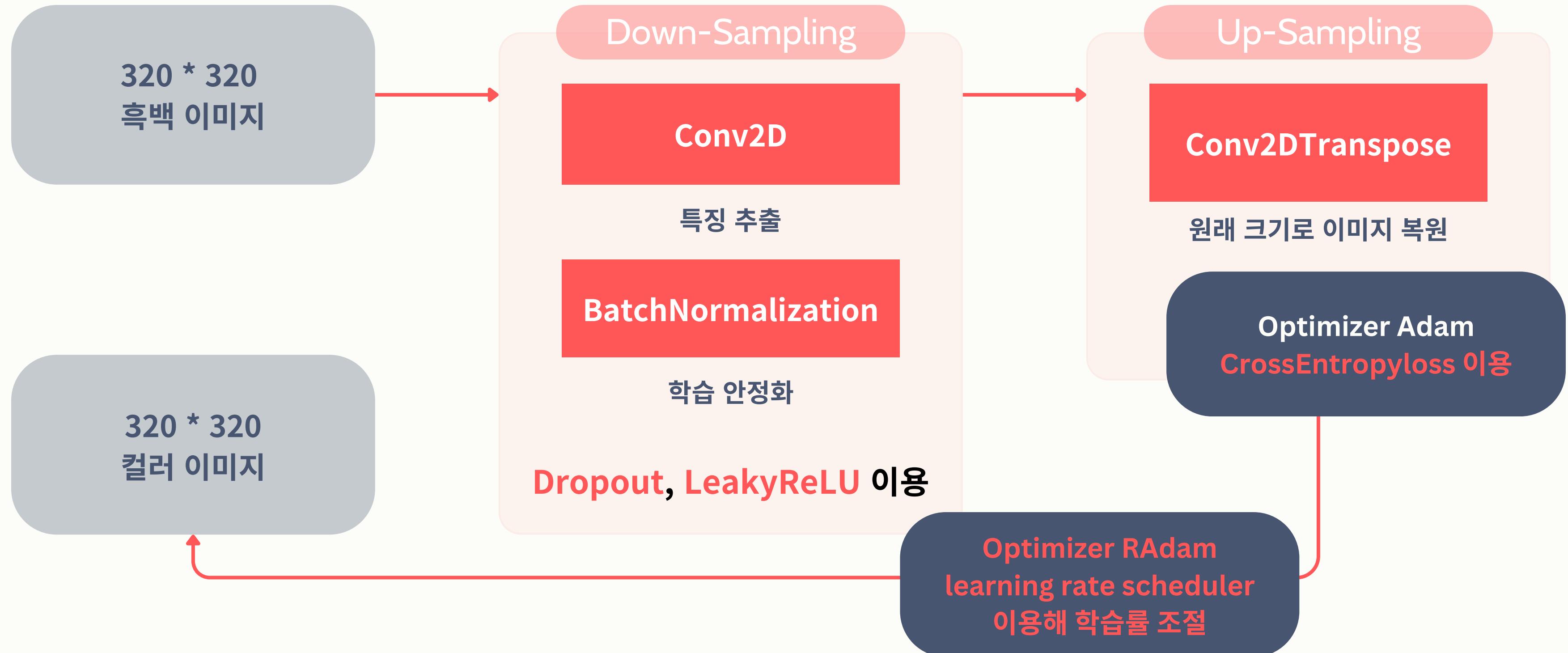
데이터 셋의 **target** 이미지 중
흑백 이미지가 섞여있었기에
학습에 방해가 되었을 수 있다.

이 경우에도 근소하게 컬러화가 진행된 것에
의의가 있으나
데이터 셋을 변경할 필요가 있다.

3차 데이터셋

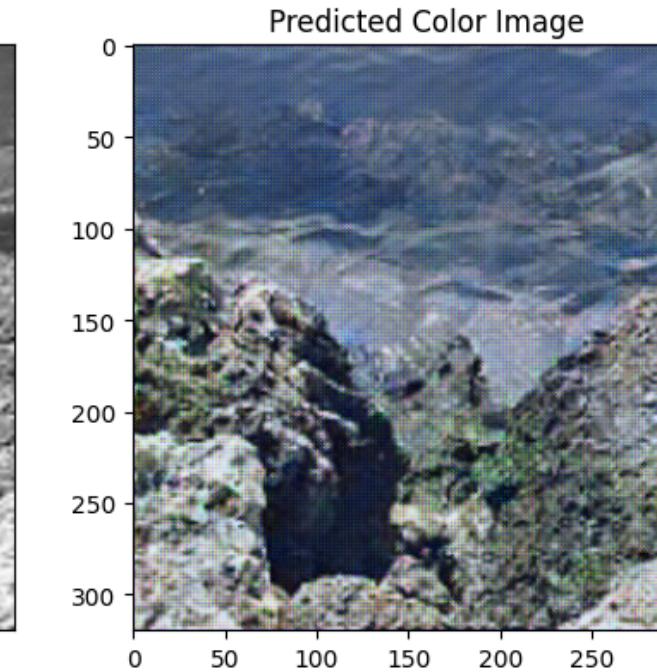
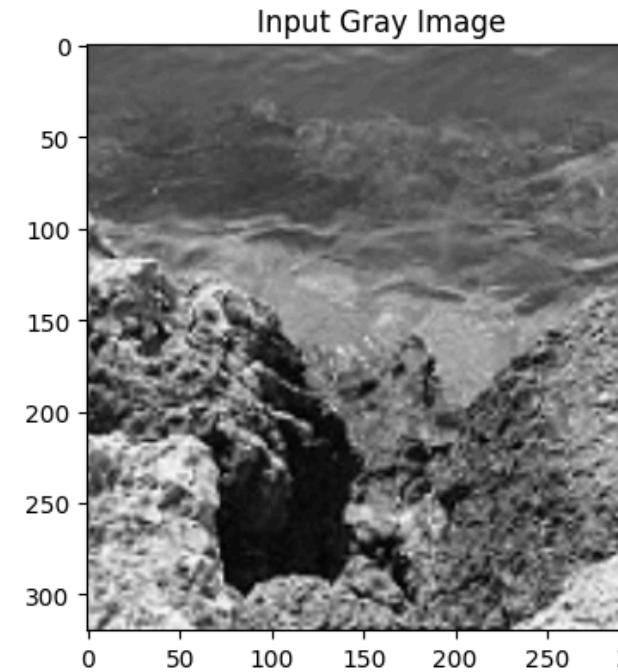
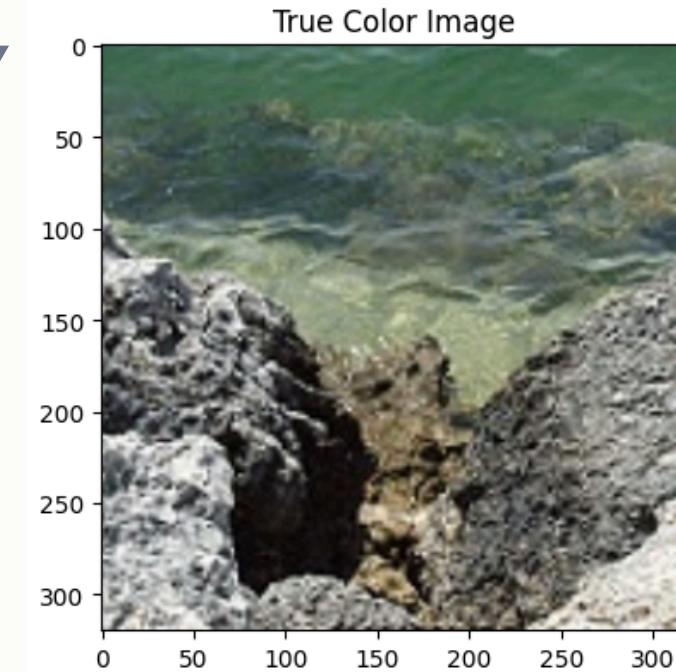
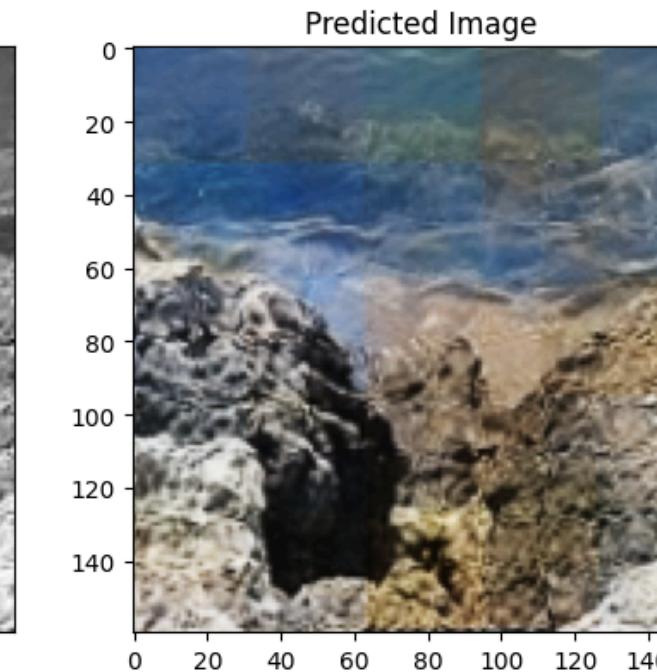
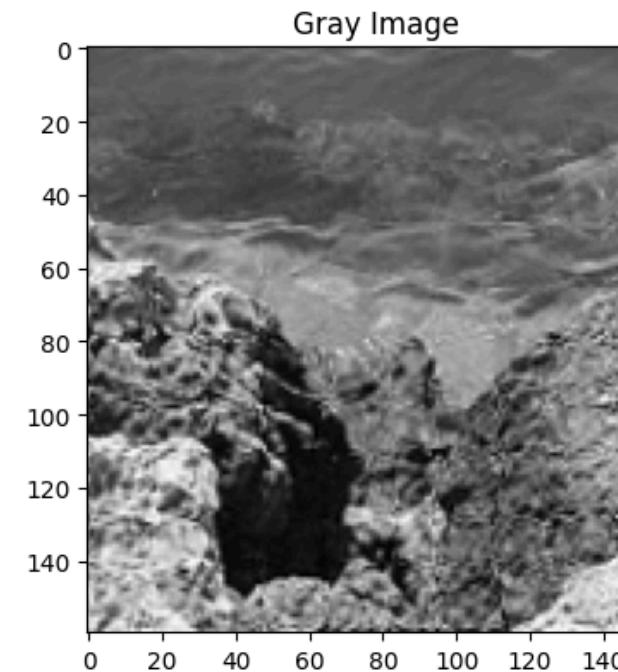
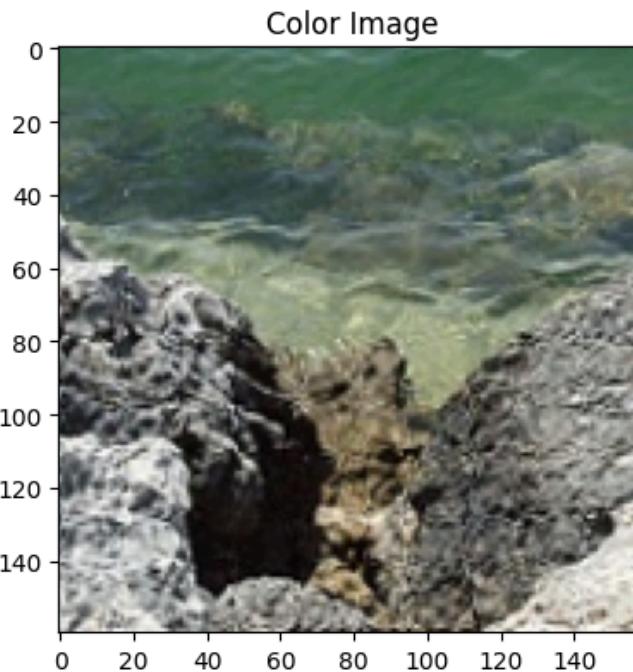
- 데이터 형식: JPG 파일
- 해상도: 320* 320 픽셀
- 데이터 수량: 6000개 쌍
- 풍경 데이터 이용

모델 구조



5.3차 모델링 - U-NET

화질 개선



- 이미지를 down-sampling하고 up-sampling 하는 과정에서 기존보다 세분화된 픽셀 단위의 학습이 가능
- 160 * 160에서 320 * 320로 확장
- Perception loss를 Mae가 아닌 **VGG19 네트워크 사용**, 세부적인 구조를 더 잘 학습하도록 함 (픽셀 값의 손실을 줄이고 자연스럽게 색 복원)
- down-sampling에서 크기를 (3,3)에서 (5,5)로 확장하여 더 넓은 영역에서 정보 추출
- batch size 16에서 8로 조정하여 세부 패턴을 효과적으로 학습

이전 모델링과의 차이

1차, 2차

U-Net 구조를 가진
Generator 클래스

2차

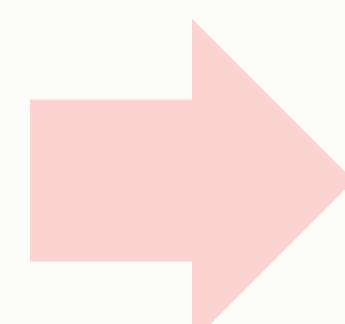
VGG16 모델

3차

U-Net 구조 기반 CNN 모델

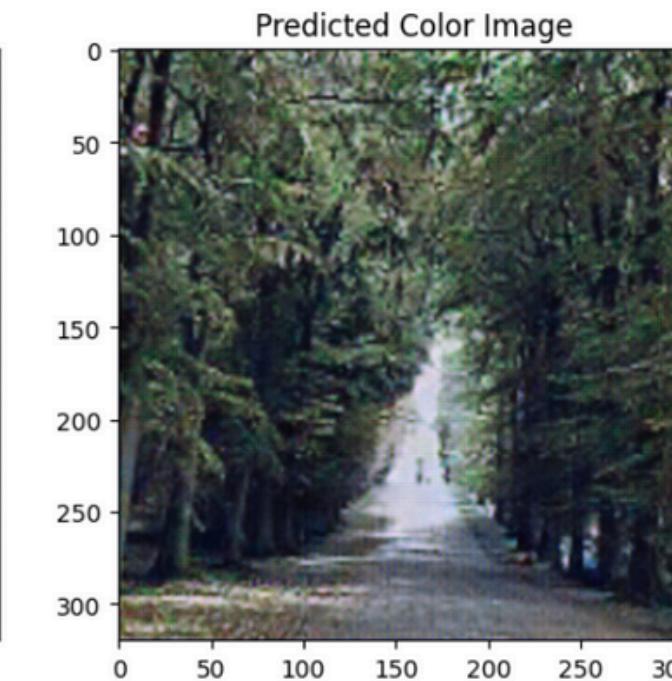
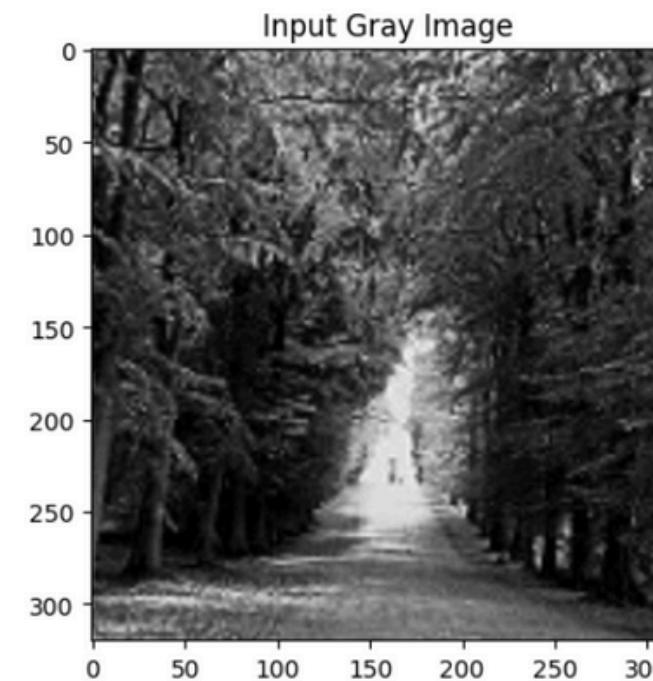
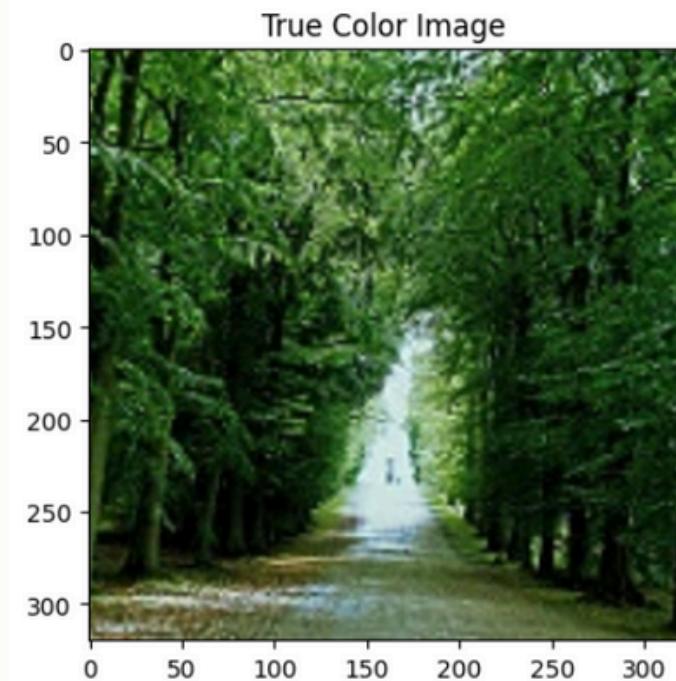
3차

VGG19 모델



5.3차 모델링 - U-NET

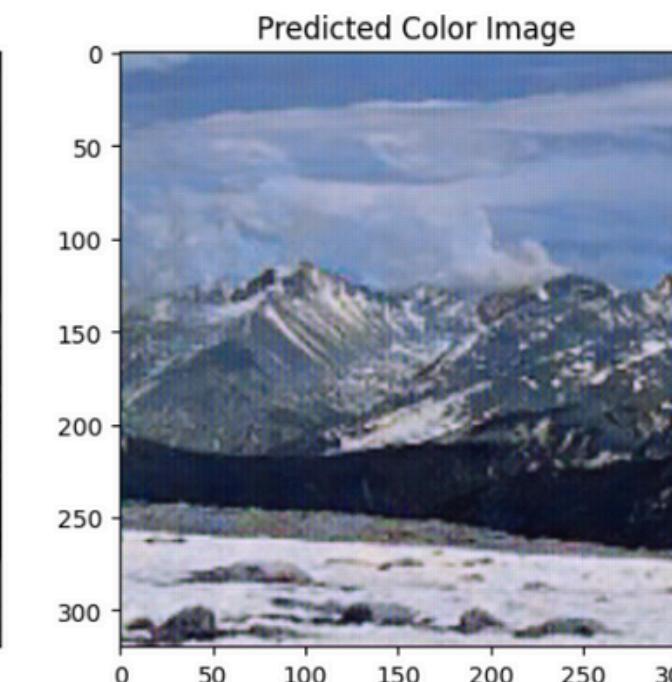
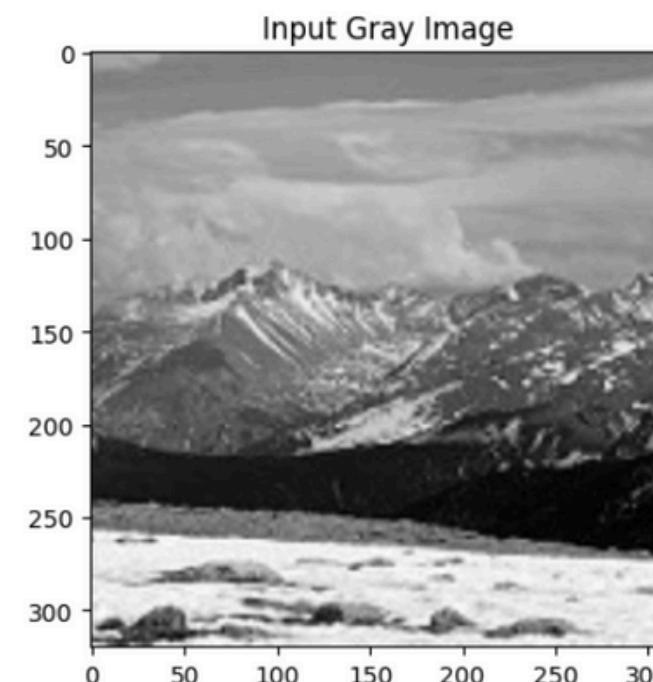
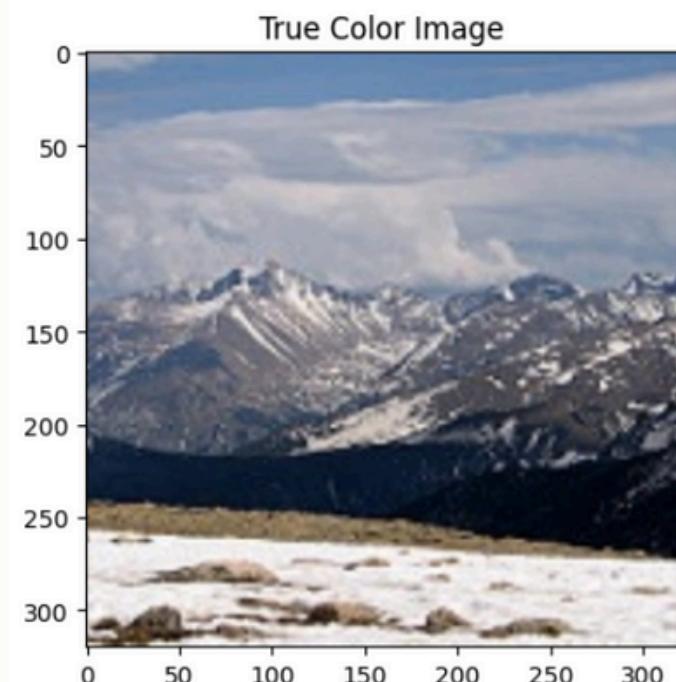
모델 평가



평가

PSNR

최대: 18.37dB
평균: 15.63dB

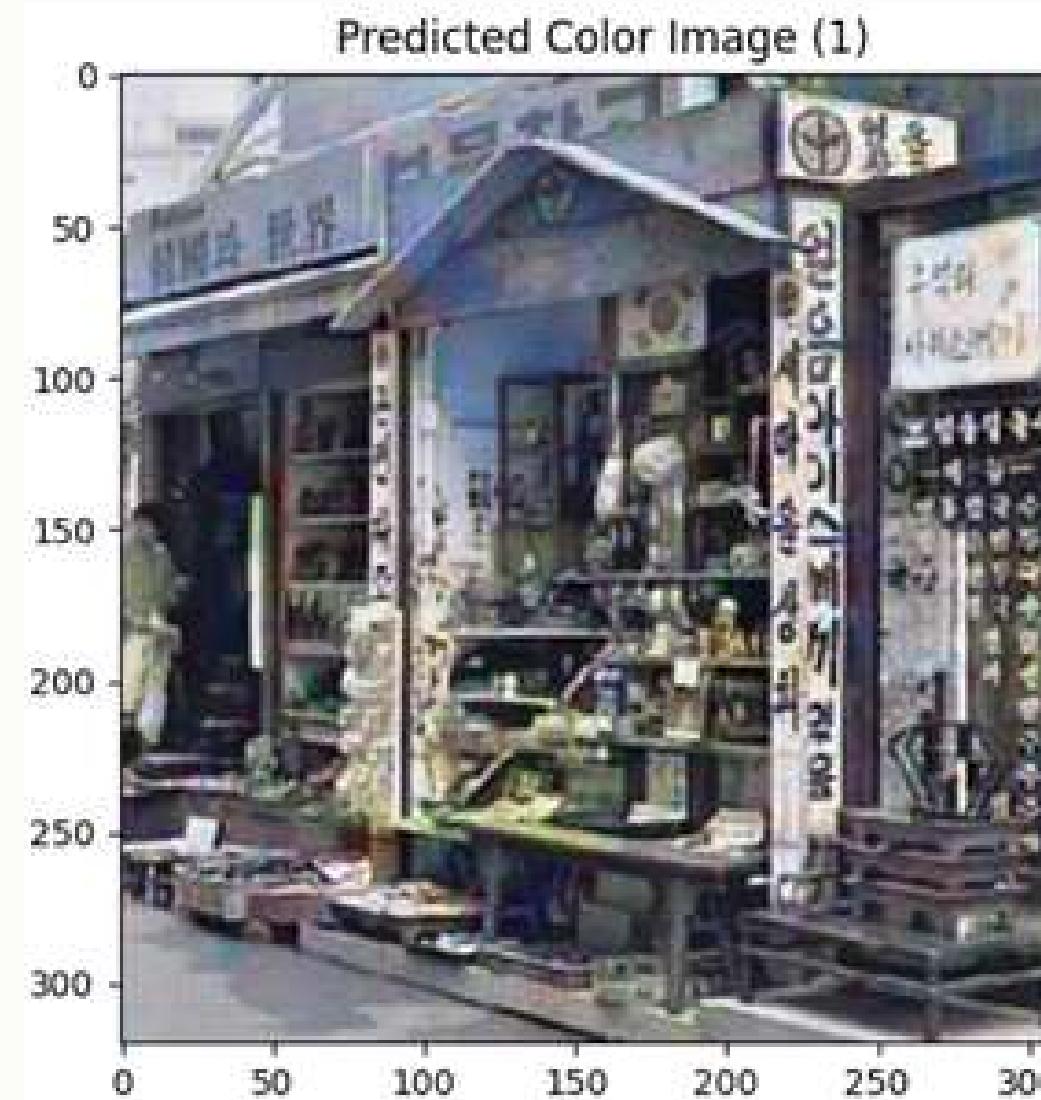
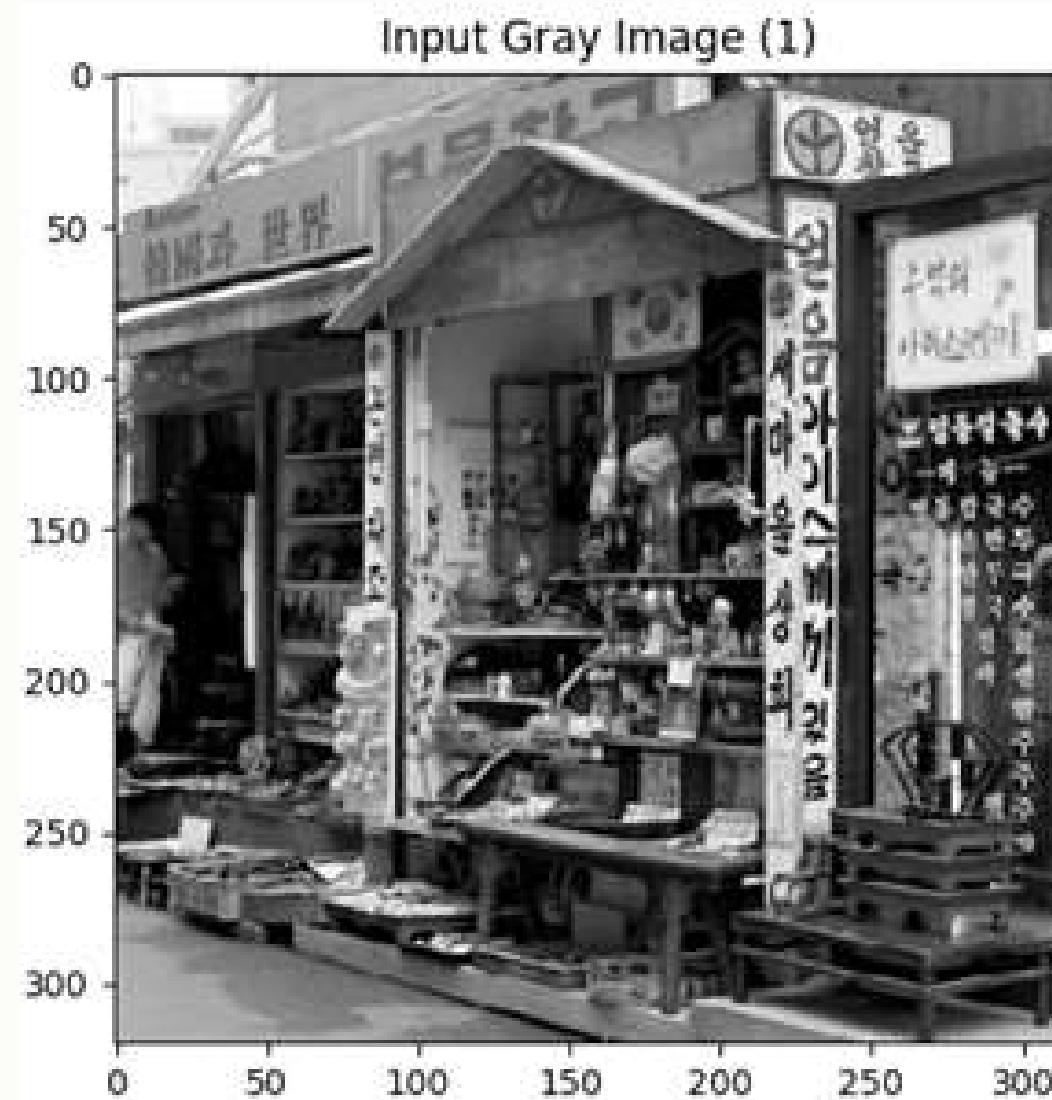


SSIM

최대: 0.4604
평균: 0.2181

6. 결론

테스트 및 의의



원래의 목적대로 과거의 사진을
컬러화 한 결과
흑백 사진을 상당히 높은 성능으로
컬러화 성공
훈련세트에 포함되지 않은 이미지도
효과적으로 컬러화 가능

한계 및 개선할 점

사람에 대한 컬러화 미흡

사람 위주의 데이터 이용해
추가 학습

추후 목표

흑백 동영상 컬러 동영상으로 변환 시키는 모델 구축