

A photograph of two individuals from the waist up, standing against a solid light blue background. On the left, a person with dark skin is wearing a bright green V-neck sweater over a pink collared shirt and an orange tie, paired with blue trousers. On the right, another person with dark skin is wearing a bright yellow V-neck sweater over a green t-shirt, paired with a magenta skirt. Both individuals have their arms crossed. The lighting is bright and even.

TEAM 모두연 포레버

IMAGE SUPER-RESOLUTION

AI 양재 허브 인공지능 오픈소스 경진대회



IMAGE SUPER-RESOLUTION

AI 양재 허브
인공지능 오픈 소스 경진대회

TEAM : 모두연 포레버
박수철. 장진우. 윤성국. 양성모

—

목차

- 대회 개요
- 데이터 특징
- Model Selection
- Methods
- Experiments



대회 개요

대회 개요



■ 주제

> 이미지 초 해상화(Image Super-Resolution)를 위한 AI 알고리즘 개발

■ 목적 및 배경

> 인공지능 오픈소스를 활용하여 이미지 초 해상화 문제를 해결함으로써 오픈 소스 생태계와 컴퓨터 비전 분야에 기여

■ 설명

> 품질이 저하된 저해상도 촬영 이미지(512X512)를 고품질의 고해상도 촬영 이미지(2048X2048)로 생성

■ 평가 산식

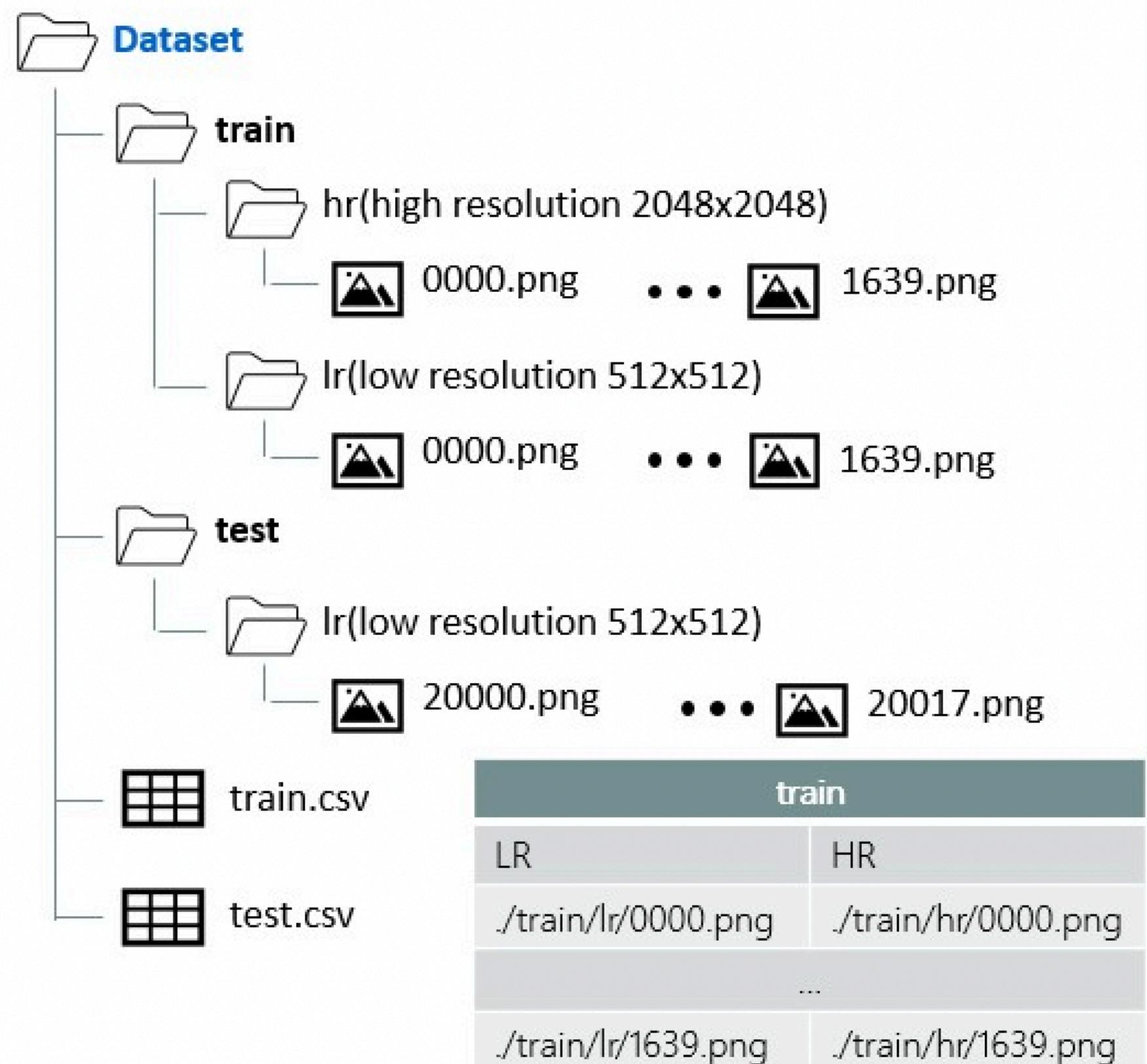
> PSNR(Peak Signal-to-Noise Ratio)

데이터 특징

데이터 특징

■ 데이터 구조

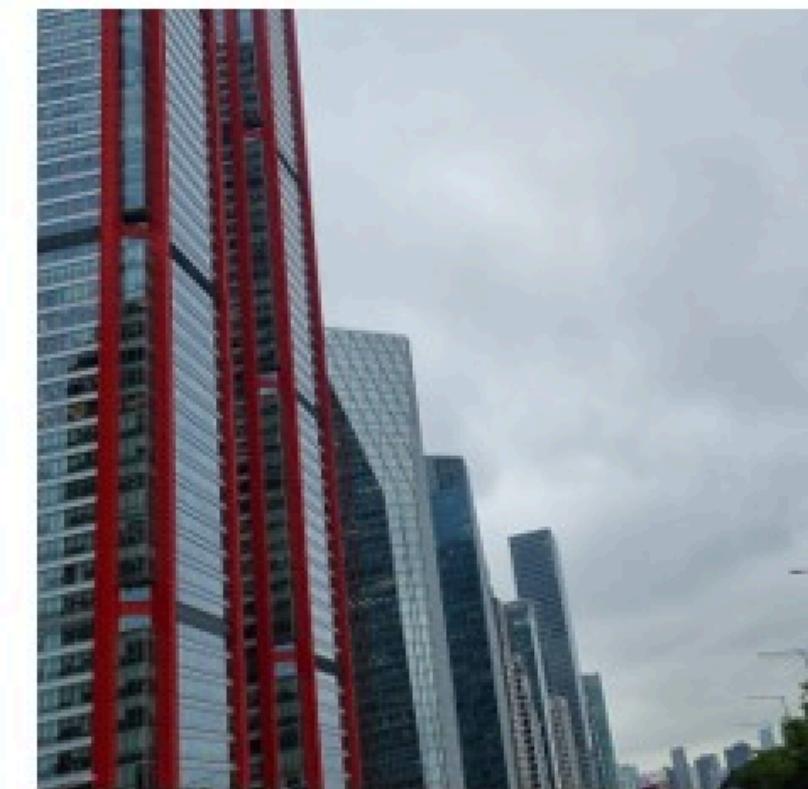
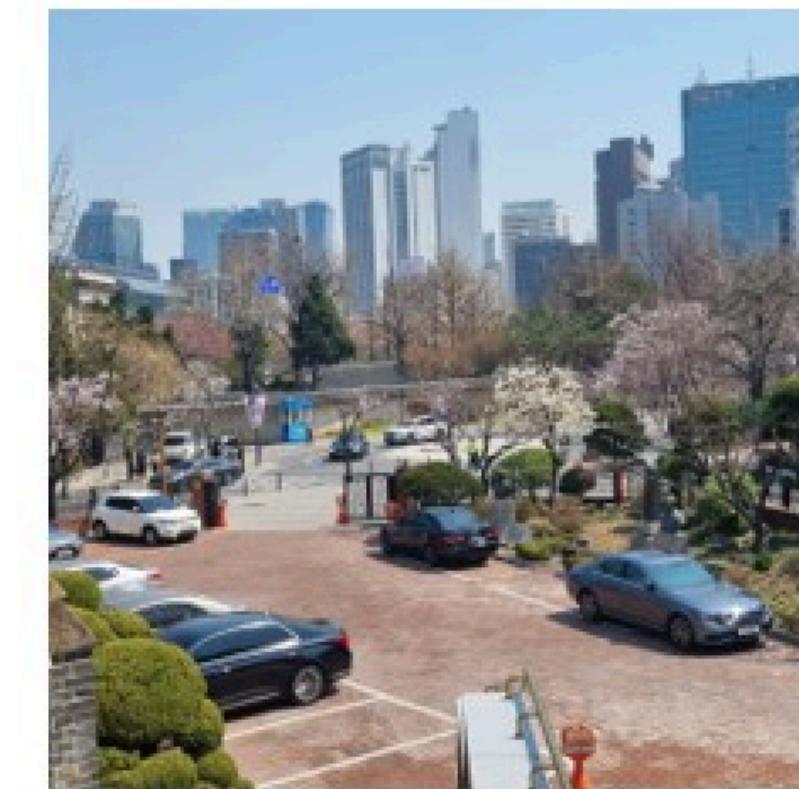
- > 고해상도(2048X2048)이미지와 저해상도(512X512) 이미지 총 1640쌍의 학습용 데이터
- > 저해상도 이미지 18개의 테스트 데이터
- > 테스트 데이터를 대상으로 초 해상화를 통해 생성한 고해상도 이미지를 대회 평가용으로 제출



데이터 특징

■ 제공된 데이터 특징

- > 주로 수도권 도시 풍경을 촬영한 이미지
- > 저해상도 이미지는 원본 이미지를 다양한 방법으로 512X512 해상도로 감소시킨 것으로 보임
- > 빌딩, 창문 등 미세한 정보가 많아 이를 정확하게 복원하는 것이 PSNR 성능의 주요 요소로 판단



< 제공된 데이터 이미지 >

■ 유사 데이터 셋

- > Urban 100(Single Image Super-Resolution)
- > Cityscape(Semantic Segmentation)



< Ubran100(좌), Cityscape(우) >

MODEL SELECTION

초기 모델 선정

- 초기 모델 선정은 Papers with code의 Image Super-Resolution 관련 성능 Benchmarks를 참고 (특히, 유사 데이터 셋으로 판단한 Urban100-4x upscaling)
- 초기 선정 모델은 **HAT, SwinIR** 두 가지 모델로 선정

Rank	Model	PSNR ↑	SSIM	LPIPS	Perceptual Index	Paper	Code	Result	Year	Tags ⚡
1	<u>HAT-L</u>	28.60	0.8498			Activating More Pixels in Image Super-Resolution Transformer	🔗	📄	2022	
2	<u>HAT</u>	28.37	0.8447			Activating More Pixels in Image Super-Resolution Transformer	🔗	📄	2022	
3	<u>SwinFIR</u>	28.12	0.8393			SwinFIR: Revisiting the SwinIR with Fast Fourier Convolution and Improved Training for Image Super-Resolution	🔗	📄	2022	
4	<u>SwinIR</u>	27.45	0.8254			SwinIR: Image Restoration Using Swin Transformer	🔗	📄	2021	
5	<u>HBPN</u>	27.3	0.818			Hierarchical Back Projection Network for Image Super-Resolution	🔗	📄	2019	
6	<u>LTE</u>	27.24				Local Texture Estimator for Implicit Representation Function	🔗	📄	2021	
7	<u>SAN</u>	27.23	0.8169			Second-Order Attention Network for Single Image Super-Resolution	🔗	📄	2019	
8	<u>CSNLN</u>	27.22	0.8168			Image Super-Resolution with Cross-Scale Non-Local Attention and Exhaustive Self-Exemplars Mining	🔗	📄	2020	

최종 모델 선정

- 초기 검토 모델들은 Transformer 계열의 모델
- Transformer 계열의 모델은 다양하고 방대한 데이터에 대해서는 효과적이지만 데이터가 **한정적인** 경우에는 좋지 않을 것으로 판단
- 따라서, CNN 계열의 RRDBNet을 Baseline으로 사용
(BasicSR repository : https://github.com/XPixelGroup/BasicSR/blob/master/basicsr/archs/rrdbnet_arch.py)

METHODS

학습방법

- **Dataset : 대회에서 주어진 데이터만 사용**

> 저해상도를 만드는 방법이 공유되지 않았기 때문에 추가 데이터는 불필요할 것으로 판단

- **Augmentation : Rotation, Flip**

> 특정 이미지 패턴에 모델이 편향되지 않도록 train 데이터에 rotation, flip 적용

- **다양한 Patch Size에 대해 검토**

> 빌딩 등에서 자주 발생하는 Moire pattern 등의 다양한 패턴들의 영향을 고려하여 다양한 Patch Size에 대해서 성능 검토

- **Loss function : MSELoss**

> 성능 지표로 PSNR만 고려하기 때문에, MSELoss를 Loss function

- **MultiStepLR 적용**

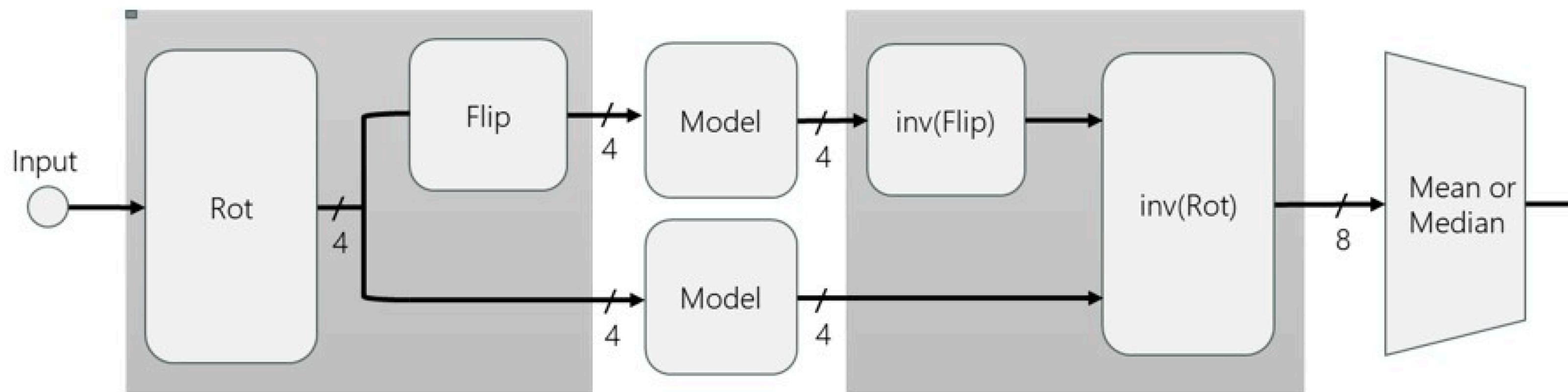
> 200,000 iter마다 gamma를 0.5 줄어들도록 설정

- **Post Processing : Geometric Self-Ensemble**

> 후 처리 과정으로 Geometric Self-Ensemble을 이용하여 성능 향상

GEOMETRIC SELF-ENSEMBLE

- EDSR Paper에서 언급된 ‘NTIRE 2017 SISR challenge’를 우승할 때 사용했던 방법인 Geometric Self-Ensemble 방법 채용. [1]



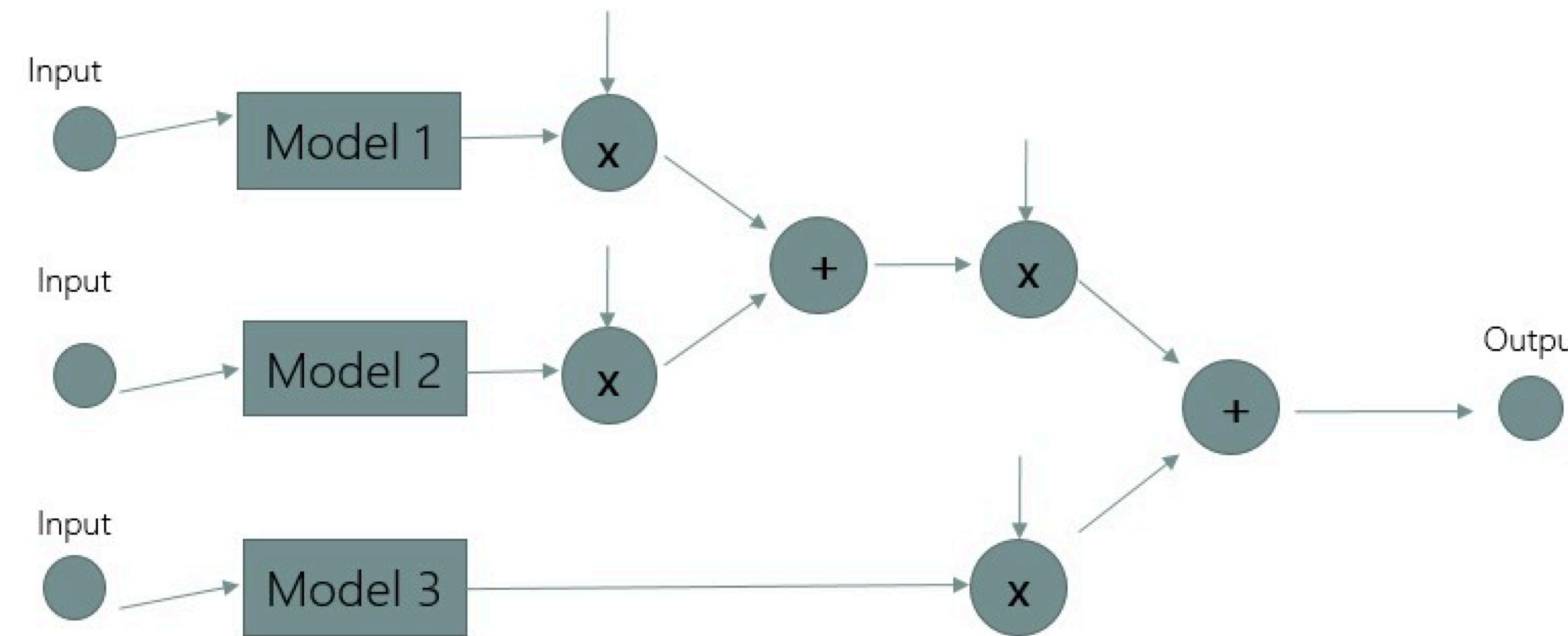
- Rotation, Flip을 이용하여 학습된 모델에 좌우 반전, 4가지 회전(0, 90, 180, 270)을 적용한 이미지 8개를 통과시켜 나온 결과들을 평균 혹은 중앙값을 취하여 성능을 개선시킴.
(최종 모델은 평균 값 이용)

[1] <https://arxiv.org/pdf/1707.02921.pdf>

(ADDITIONAL) SOFT VOTING

- 팀원들의 각각의 모델의 pnsr값에 근거하여 weight를 설정하여 soft voting 수행

그림 업데이트 필요



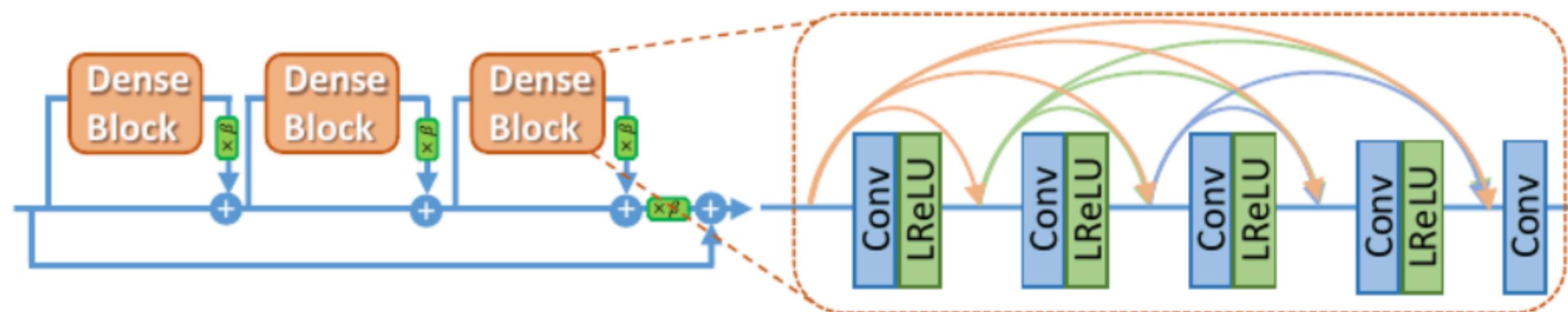
- Voting은 각 모델 별 성능 편차가 존재하여 큰 개선 효과가 없었음

NETWORK ARCHITECTURE

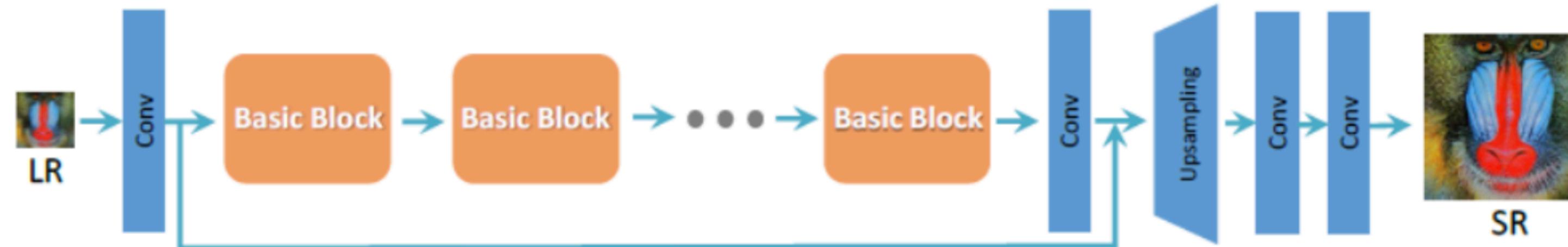
■ Baseline으로 RRDBNet을 이용

> RRDB : 5개의 Conv로 이루어진 Dense Block을 3번 통과 후 통과 전 이미지와 weighted sum

Residual in Residual Dense Block (RRDB)



> RRDBNet Architecture (Basic Block : RRDB)



EXPERIMENTS

EXPERIMENTS

- RRDBNet 기본 architecture를 baseline으로 아래와 같은 실험들을 진행
 1. Model hyper-parameters 조정 (num_feat, num_grow_ch)
 2. Training 시 Model 자체에 rotation ensemble을 출력단에 이용
 3. Body(RRDB)의 위치 변경
 4. CLIP Embedding 활용
 5. Gated Convolution 사용
 6. Pretrained Model 사용