

국내 최초 빅데이터 연합 동아리 BOAZ

BOAZ SR Project

: 경기는 계속되어야 한다

13기 시각화	정동욱
13기 분석	김영훈
13기 분석	장형주
14기 분석	홍승아

Contents

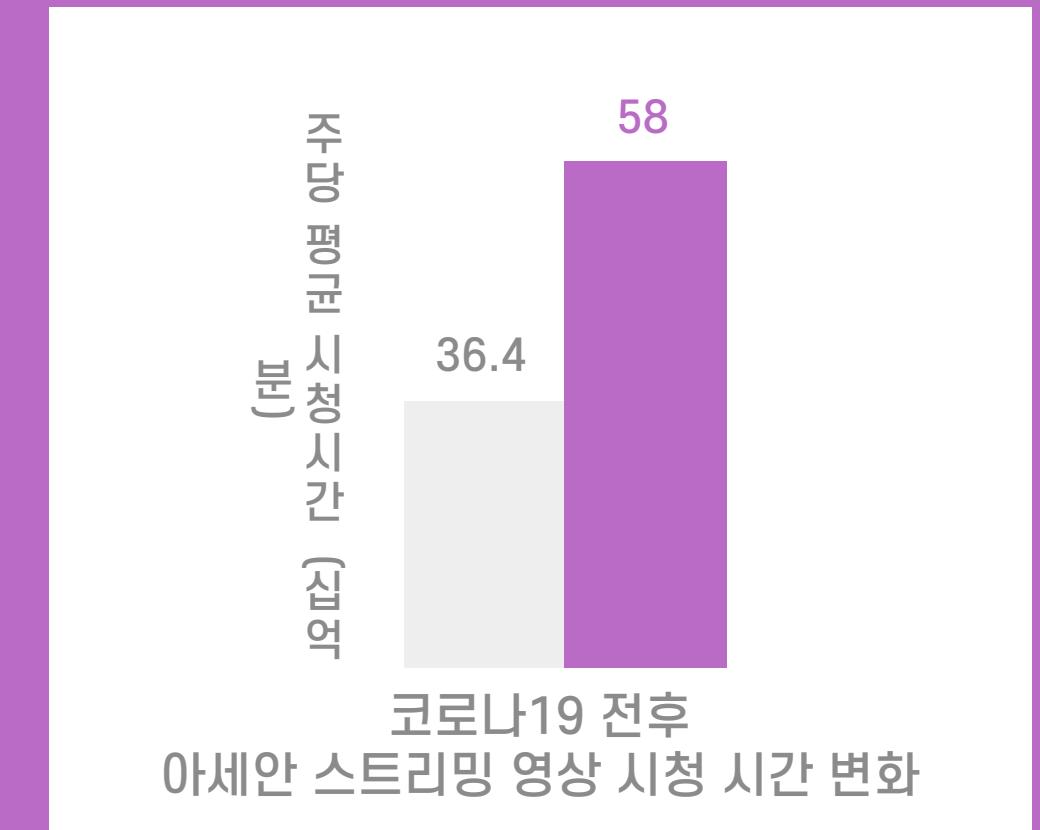
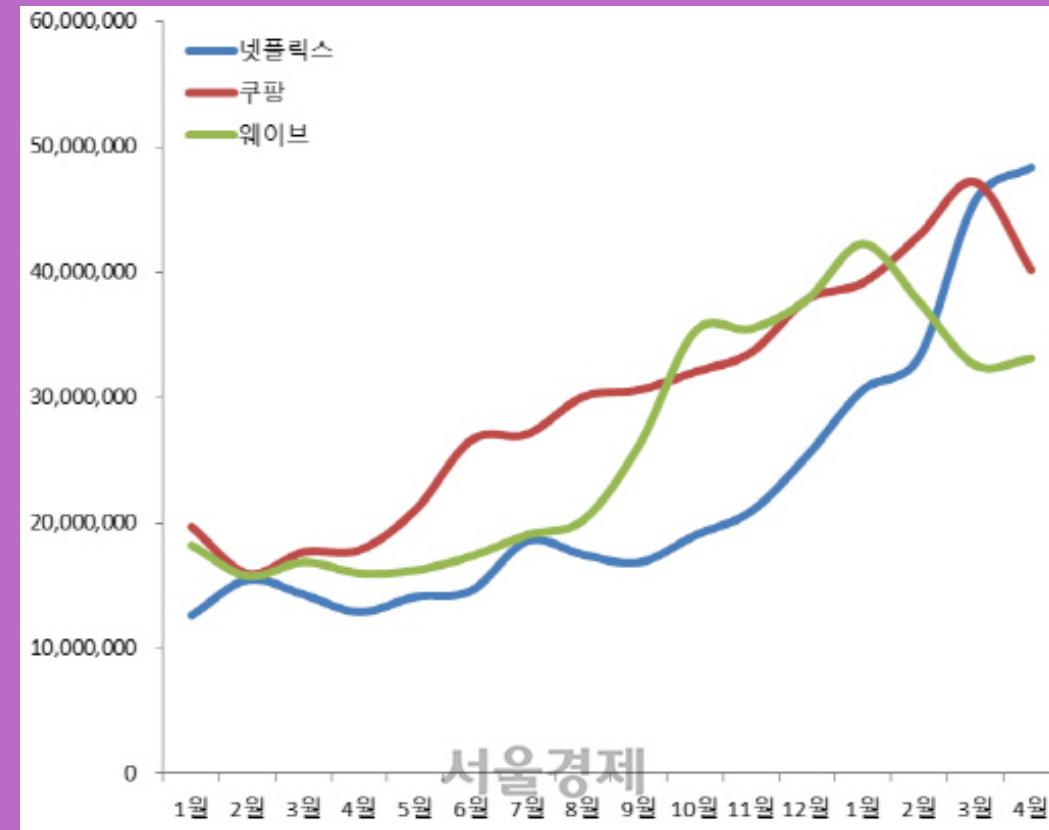
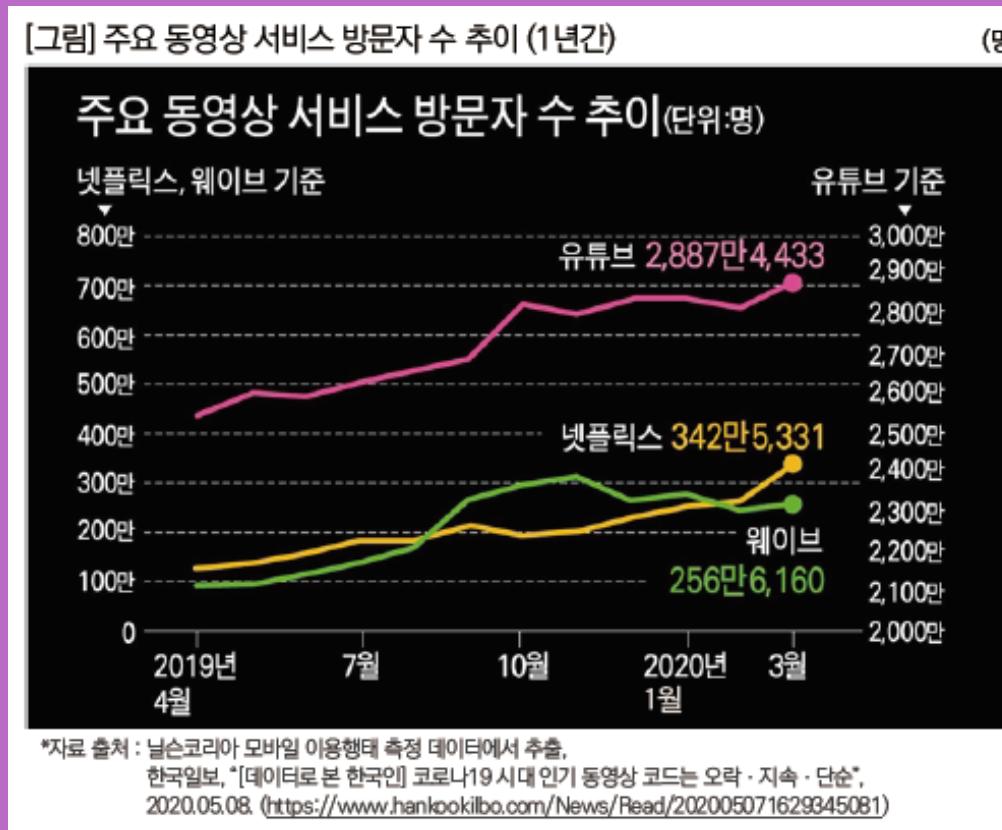
필요성	주제 및 목표	데이터셋	모델	결과	결론
스트리밍 서비스의 화질 개선 필요성	주제 선정 과정 실시간 축구 영상 화질 개선 목표	데이터셋 수집 - 실제 중계 영상 - 게임 플레이 영상	SR이란? Model 선정 전처리 Augmentation Training	이미지 결과 동영상 캡처본 추론 속도 전체 프로세스	의의 및 한계

01 ————— 필요성

스트리밍 서비스의
화질 개선 필요성

01 필요성

스트리밍 서비스 플랫폼 이용 증가



유튜브, 넷플릭스 이용자 수 증가

출처 : 널슨코리아

넷플릭스 이용 시간 전년 대비 186% 증가

출처 : 아이지에이웍스

일주일 평균 스트리밍 시간 60% 증가

출처 : ASEAN KOREA CENTRE (2020.06)

01 ━━━━━━ 필요성

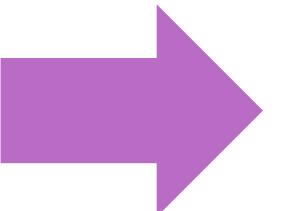
문제점 & 기존 해결 방법

네트워크 연결 불안정 → 영상 화질 저하

고화질로 설정 → 영상 끊김

저화질로 감상
낮은 화질로 인한 화면 깨짐 등 감상에 불편

캐시 삭제
매번 삭제해야 하는 번거로움



**디바이스 단위
스트리밍 동영상
화질 개선 필요!**

02 — 주제 및 목표

주제 선정 과정

실시간 축구 영상 화질 개선

목표

02 주제 주제 선정 과정

Neural Adaptive Content-Aware Internet Video Delivery

적응형 스트리밍 서비스에서의 화질 개선

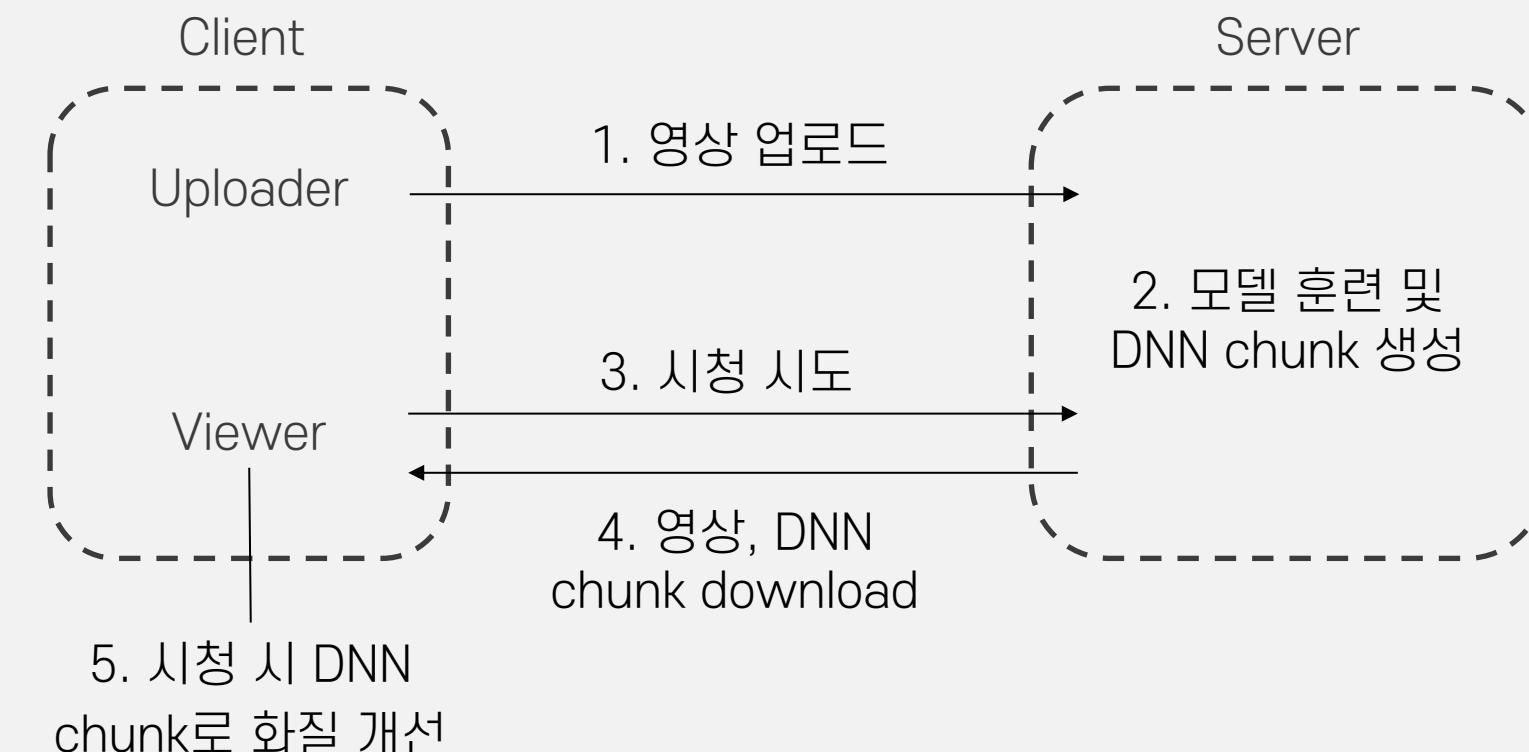
적응형 스트리밍 서비스?

미리 서버에 영상 제작자가 비디오 영상을 업로드 한 후 그 영상을 다시 시청할 수 있게끔 하는 서비스

한계

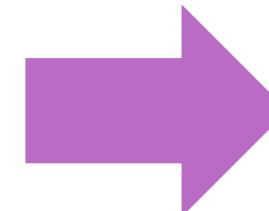
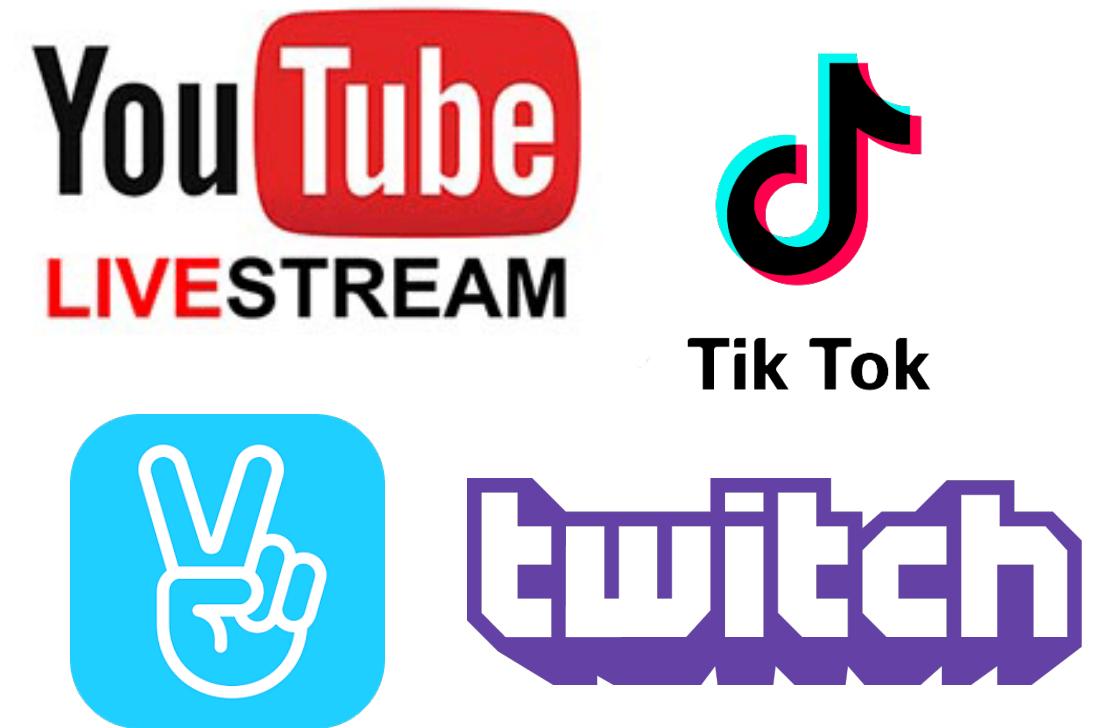
실시간 스트리밍 서비스에 적용 불가능

과정



02

주제 주제 선정 과정



실시간 스트리밍
화질 개선을 해보자 !

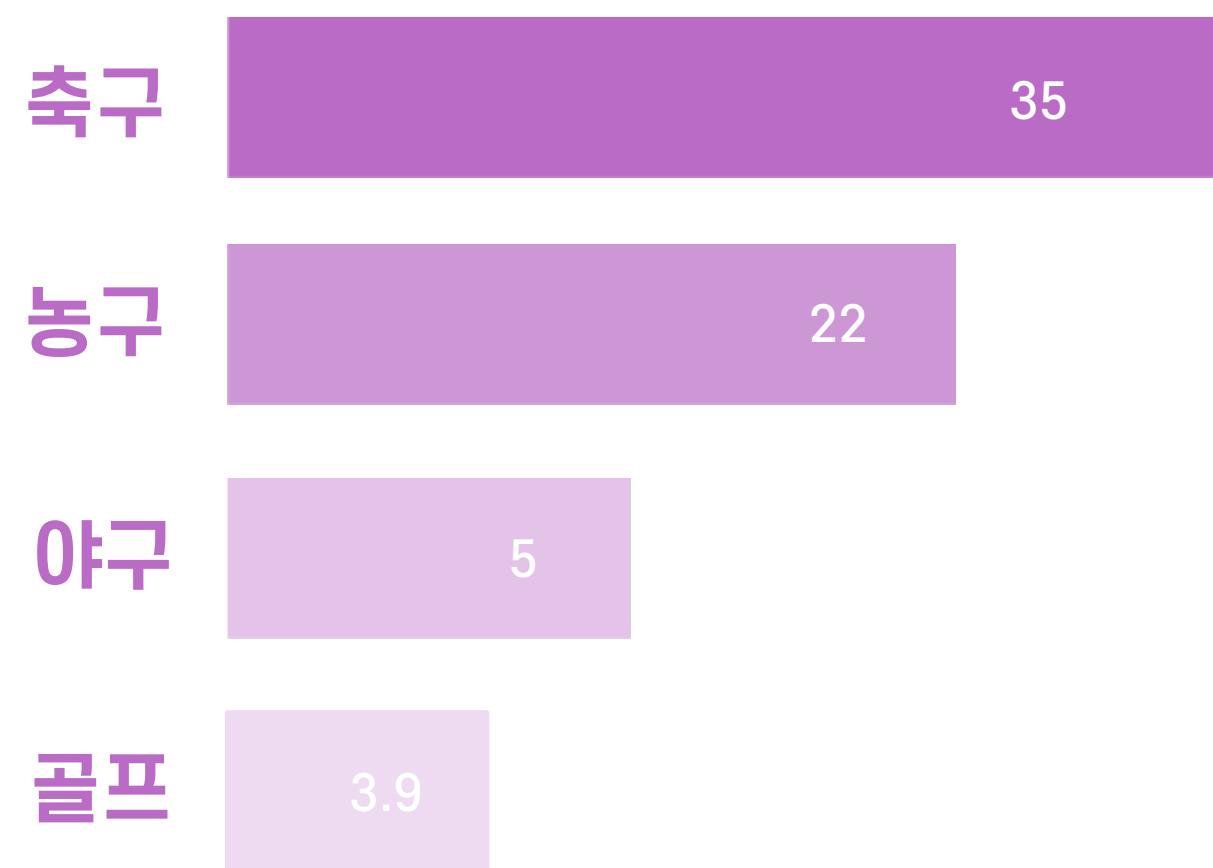
온라인 세미나, 온라인 콘서트, 실시간 게임 방송, 스포츠 경기 생방송 등
실시간 스트리밍 서비스 활용 증가

02

주제 실시간 축구 영상 화질 개선



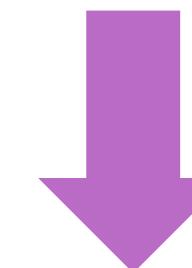
세계 스포츠 팬 수 1위, 축구



2020 기준 세계 스포츠 인기 순위

스포츠 스트리밍 시청 중 불편 경험

생중계 방송 BUT 불안정한 영상 화질



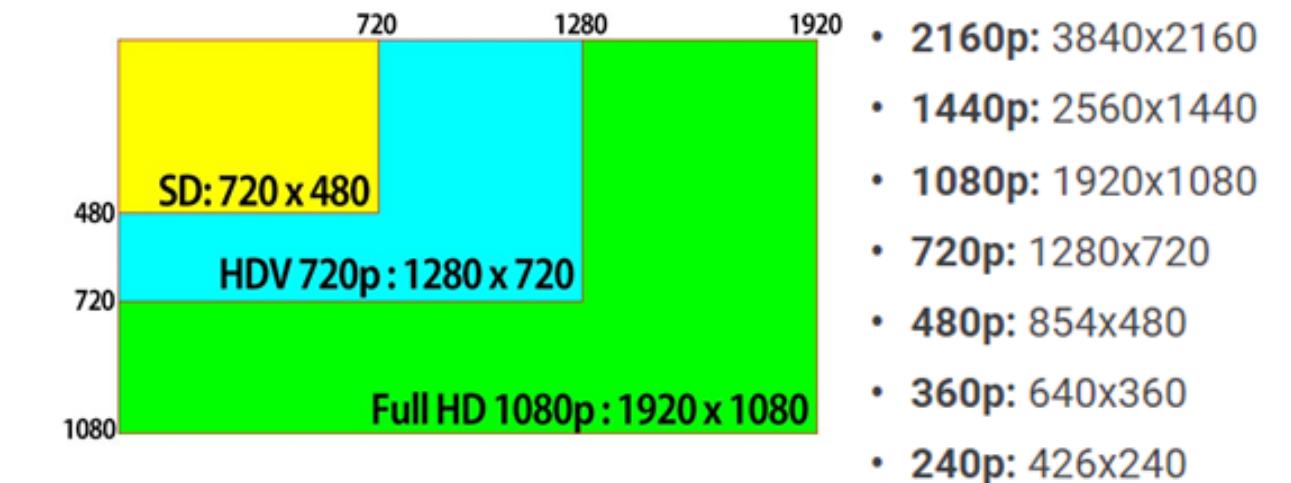
축구 생중계 영상
화질 개선 도전 !

02 _____ 주제 목표

✓ 저해상도 영상을 기준 화질 이상으로 올리는 SR 모델 구현

저해상도 : 144p or 240p / 기준 화질 : 720p

SR : Super Resolution



✓ 실시간 위한 모델 추론 속도 24fps 이상

24fps : 일반적인 프레임 재생 속도

동영상 해상도 기준

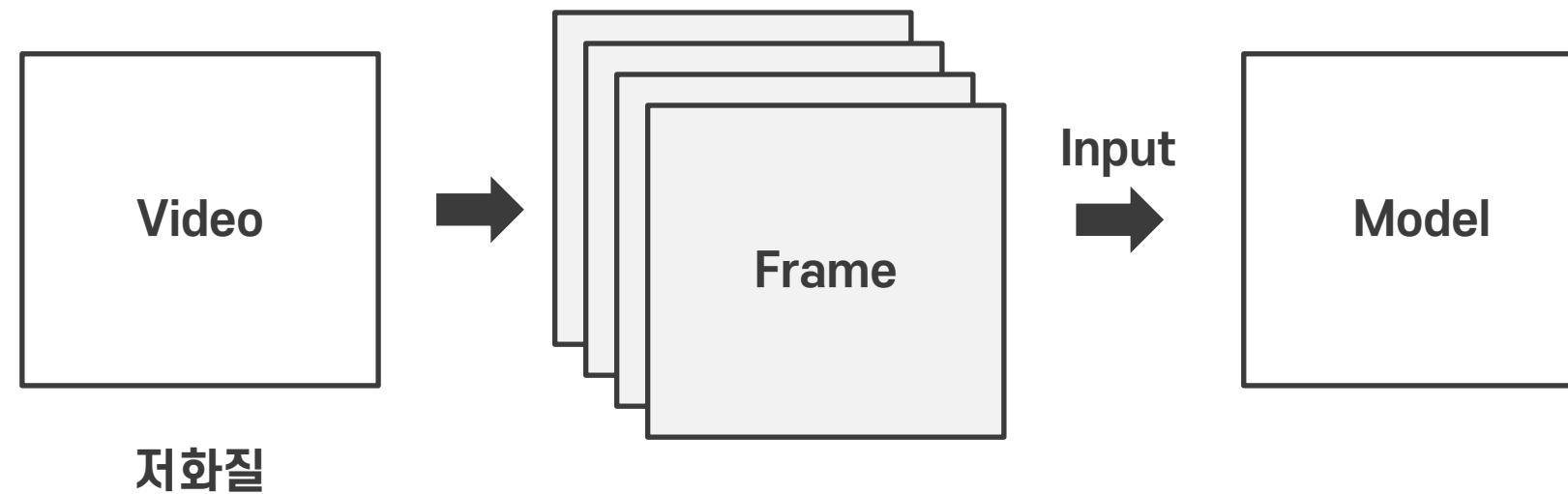
-144p, 240p, 360p, 480p… 세로 별의 픽셀 수 기준

-가로 : 세로 = 16 : 9 비율이 일반적

02 ━━━━ 주제 목표

✓ 동영상?

시간상 연속된 정지 사진들의 모음
각각의 정지 사진 하나가 ‘프레임’

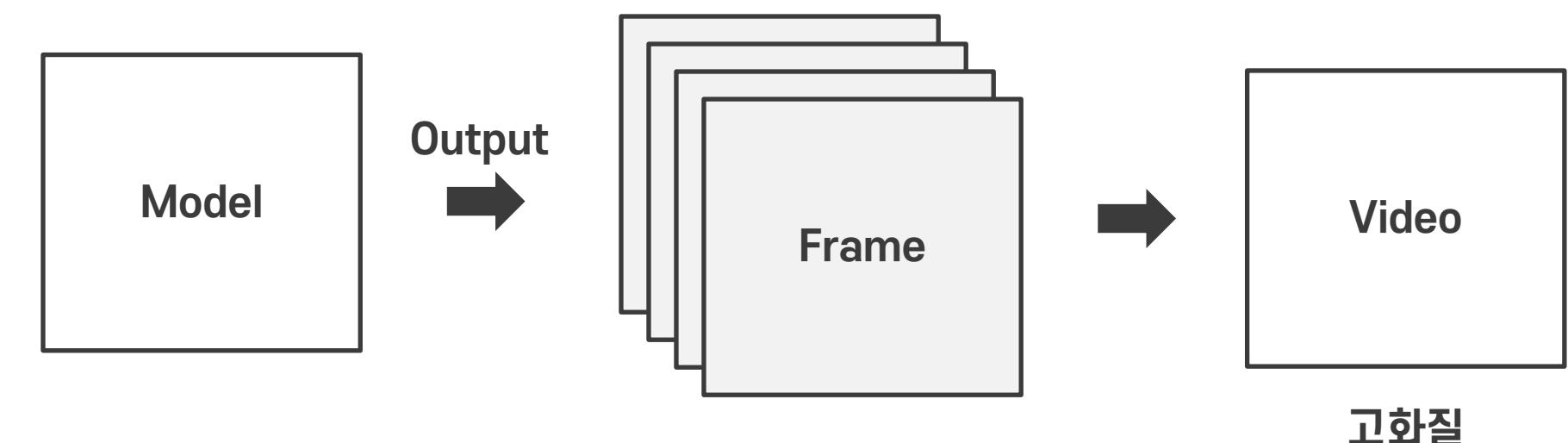


✓ FPS?

Frame Per Second. 초당 프레임

✓ 24fps 인 이유

일반적인 영상이 초당 24프레임으로 구성되므로
모델이 초당 24프레임 이상 화질 개선이 가능하다면
실시간 서비스가 가능할 것으로 예상



03—

데이터셋

실제 중계 영상

게임 플레이 영상

데이터셋 수집

03

데이터셋 실제 중계 영상

축구 중계 영상의 특징

빠른 화면 전환

Motion blur

신체의 움직임

한정적인 학습 요소

사람, 경기장, 공

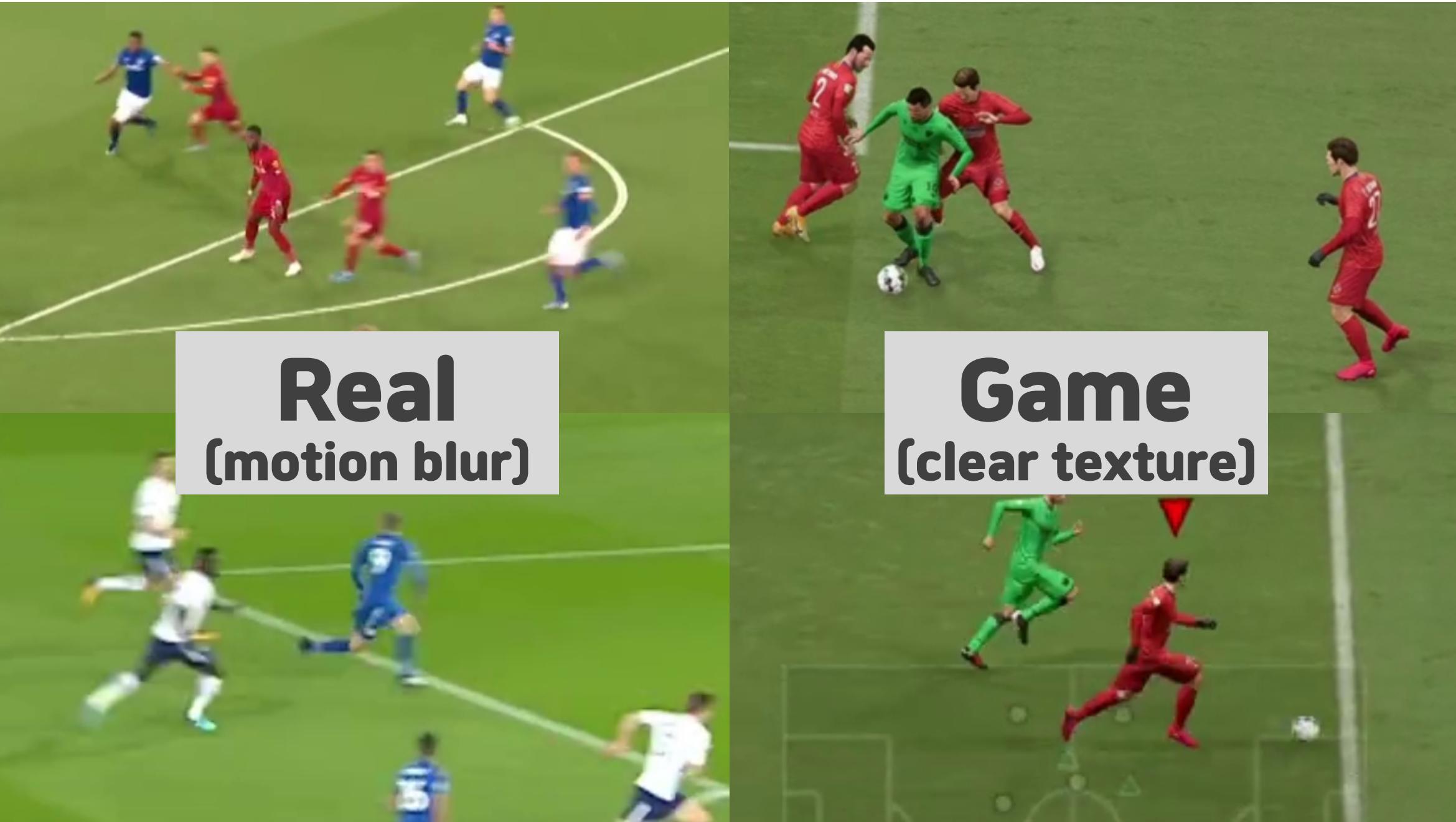
03

데이터셋 게임 영상과 중계 영상 비교



03

데이터셋 게임 영상과 중계 영상 비교



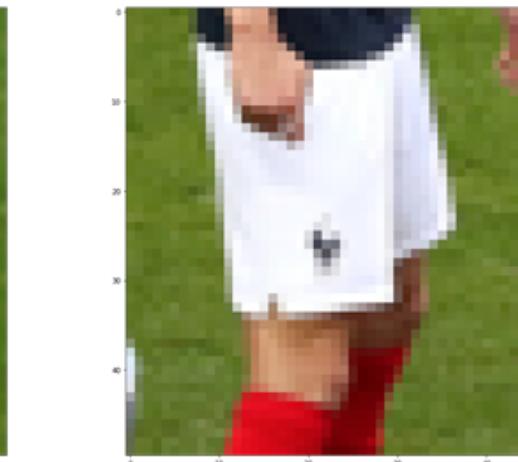
03

데이터셋 게임 플레이 영상

게임 데이터를 활용한 실험

train set : 게임 이미지 데이터 / test set : 축구 영상 이미지 데이터

모델 : SRGAN



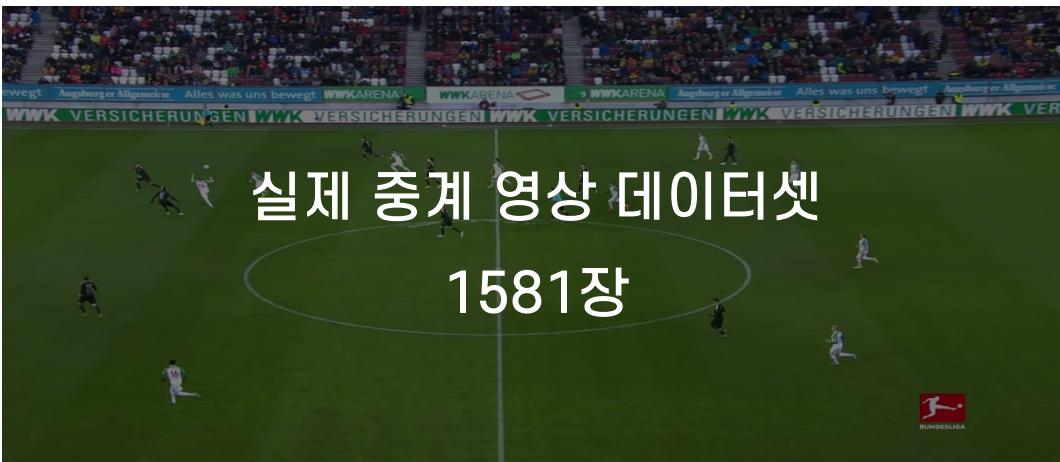
결과 : real world의 test set에 대해서도 효과를 보임 / 게임 속 데이터가 현실세계를 꽤 잘 반영

03

데이터셋 데이터셋 수집

실제 중계 영상과 게임 플레이 영상

FIFA, EPL, EA sports 공식유튜브 1080x1920 영상



총 2778장

04— 모델

SR01란?

Model 선정

전처리

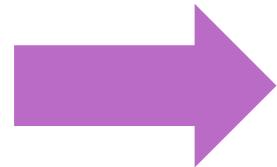
Augmentation

Training

04 _____ 모델 SR이란?

Super Resolution, 초해상화 : 화질 개선 알고리즘

저화질



고화질

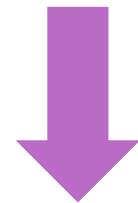


04 모델 SR이란?

단순 보간법(interpolation)

Nearest Neighbor, Bilinear, Bicubic

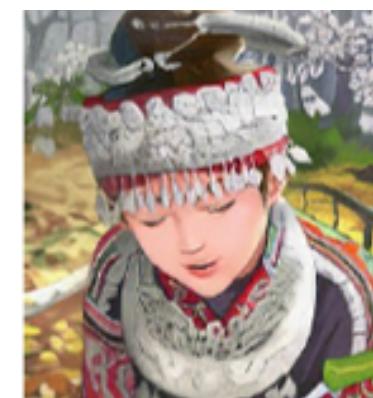
Bad ...



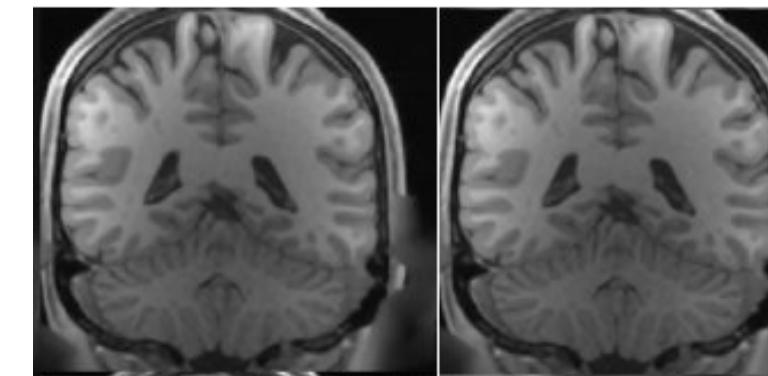
딥러닝 기반 Super-Resolution 모델

SRCNN, SRGAN, ESPCN, MDSR

Good !



Super Resolution 활용 분야



04 모델 Model 선정

MSE

Pixel-wise image 의
차이 기반으로 정의

High texture detail
잡기 어려움

GAN

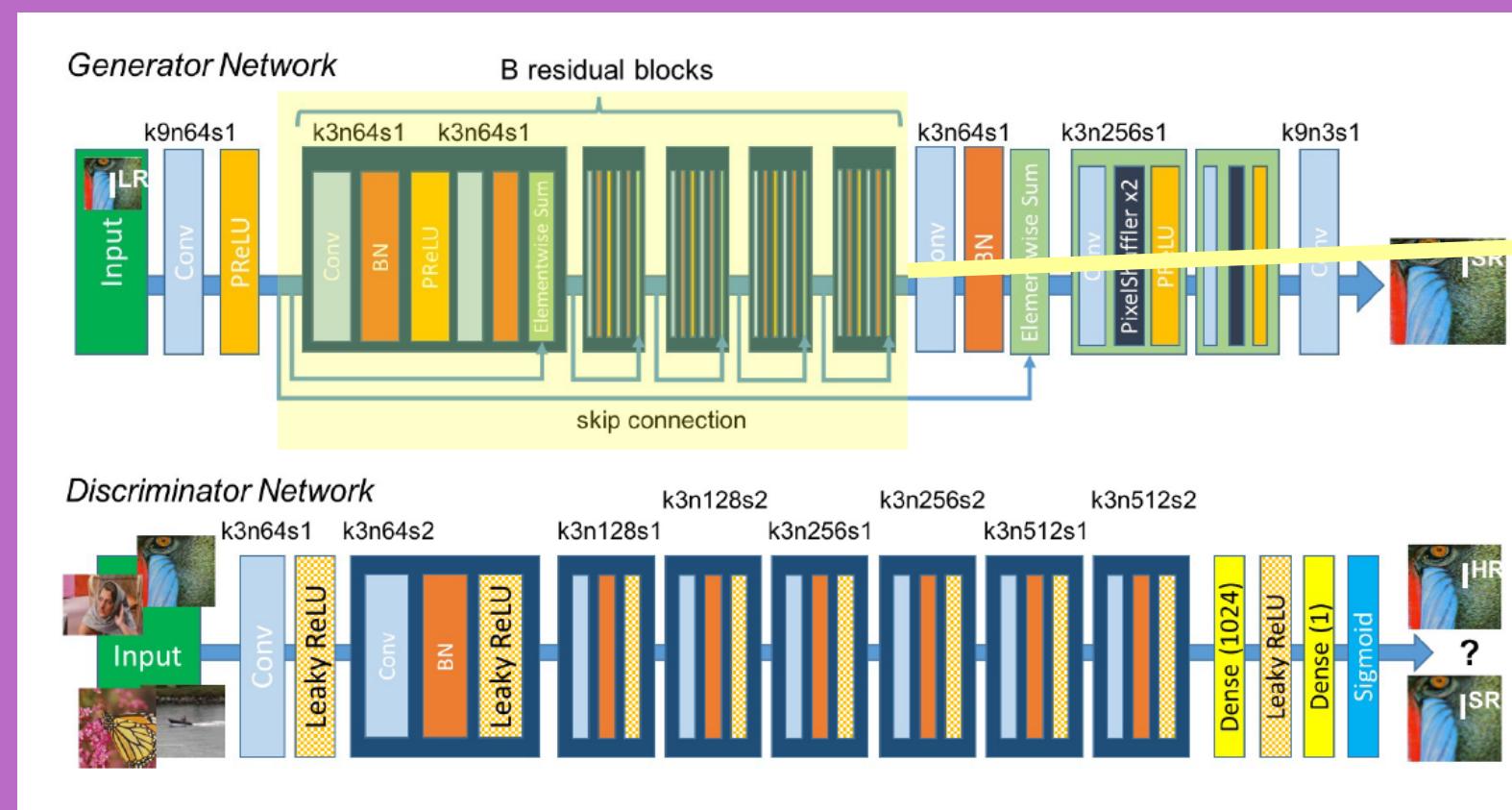
Perceptual similarity
기반하여 loss 구성

High texture detail
잘 잡아냄

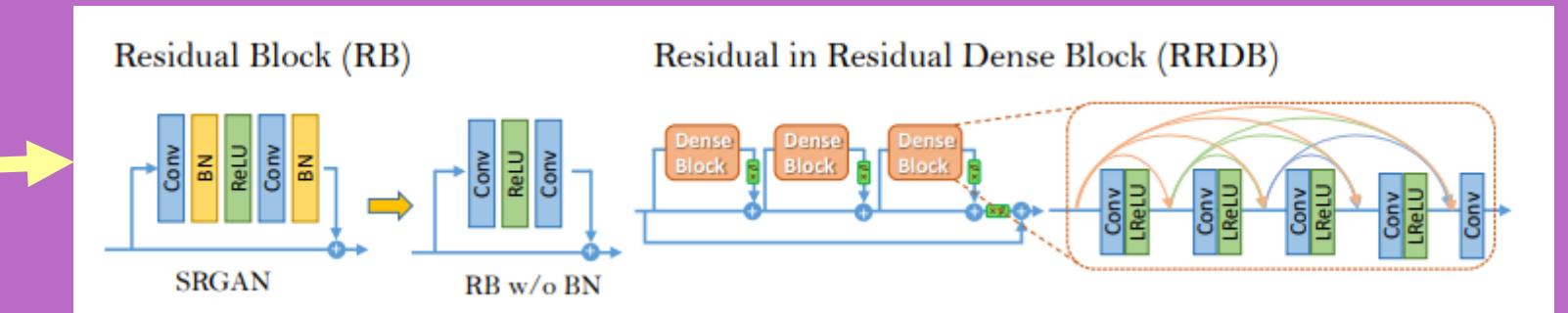
VS

SRGAN

SRGAN Network Architecture



ESRGAN



1. 전체 Batch Normalization layer의 삭제
2. 기존의 Residual Block을 RRDB로 대체
3. Discriminator 어느 쪽이 더 현실적인가?

04 모델 Model 선정

ESRGAN을 선택한 이유

1

1장짜리 모델

2

훈련 방식, 데이터 수집 용이

3

데이터 용량 최소화

4

단순한 모델 -> 빠른 추론 속도

5

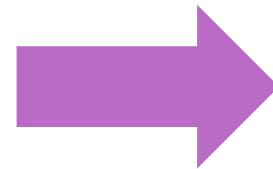
GAN 방식 -> 질감 Good

6

실험한 모델 중 가장 좋은 성능

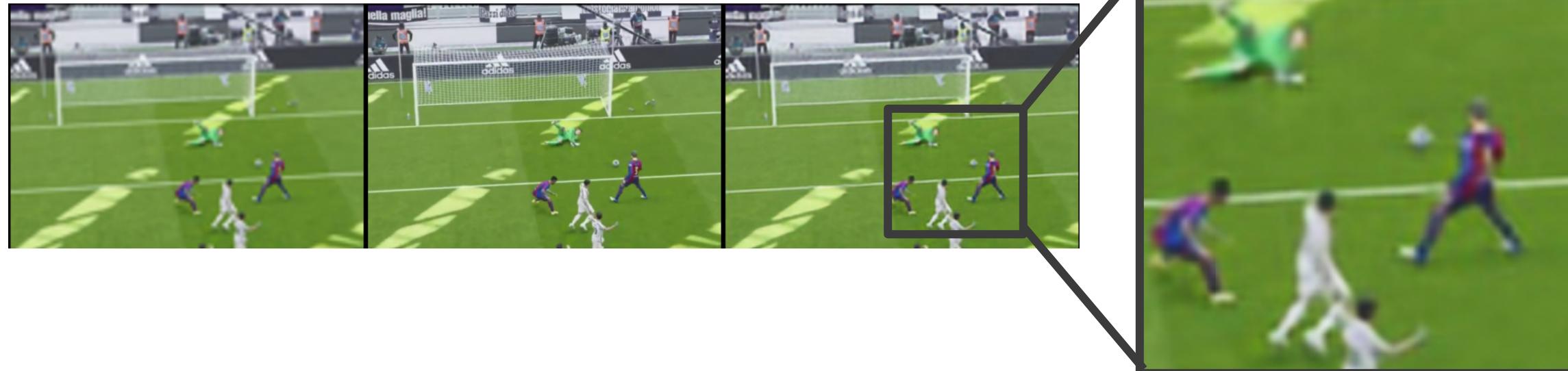
04 모델 전처리

데이터셋 학습 시
질감이 뚜렷하지 않음



RGB의 평균 분산값이 Threshold 이하일 경우 필터링

표준편차, 15

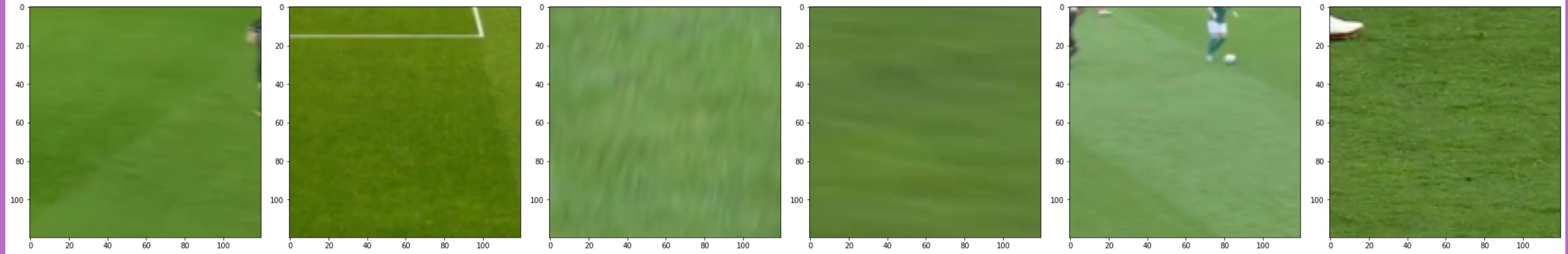


04 모델 전처리

Batch에 올라가는 사진 확인

사람의 동작, 관절 학습이 중요 !

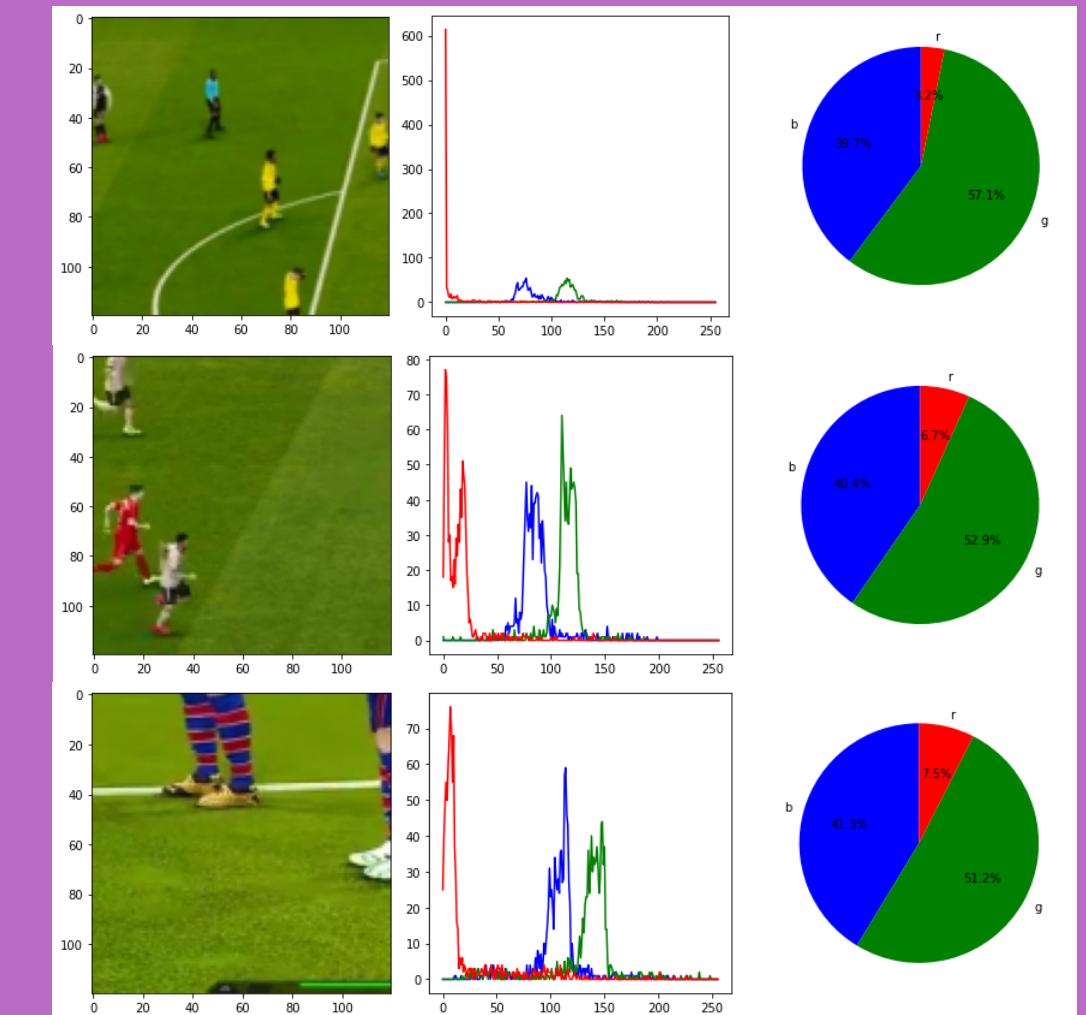
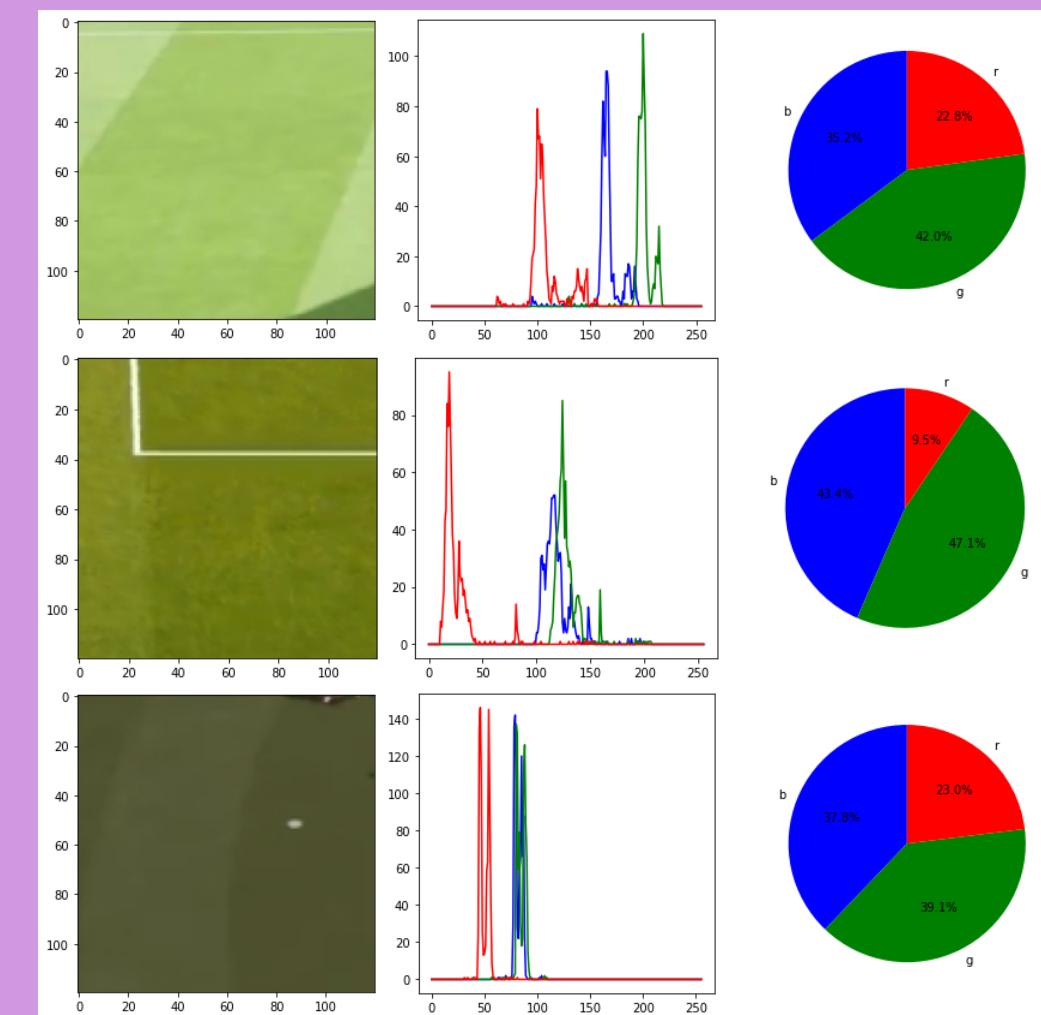
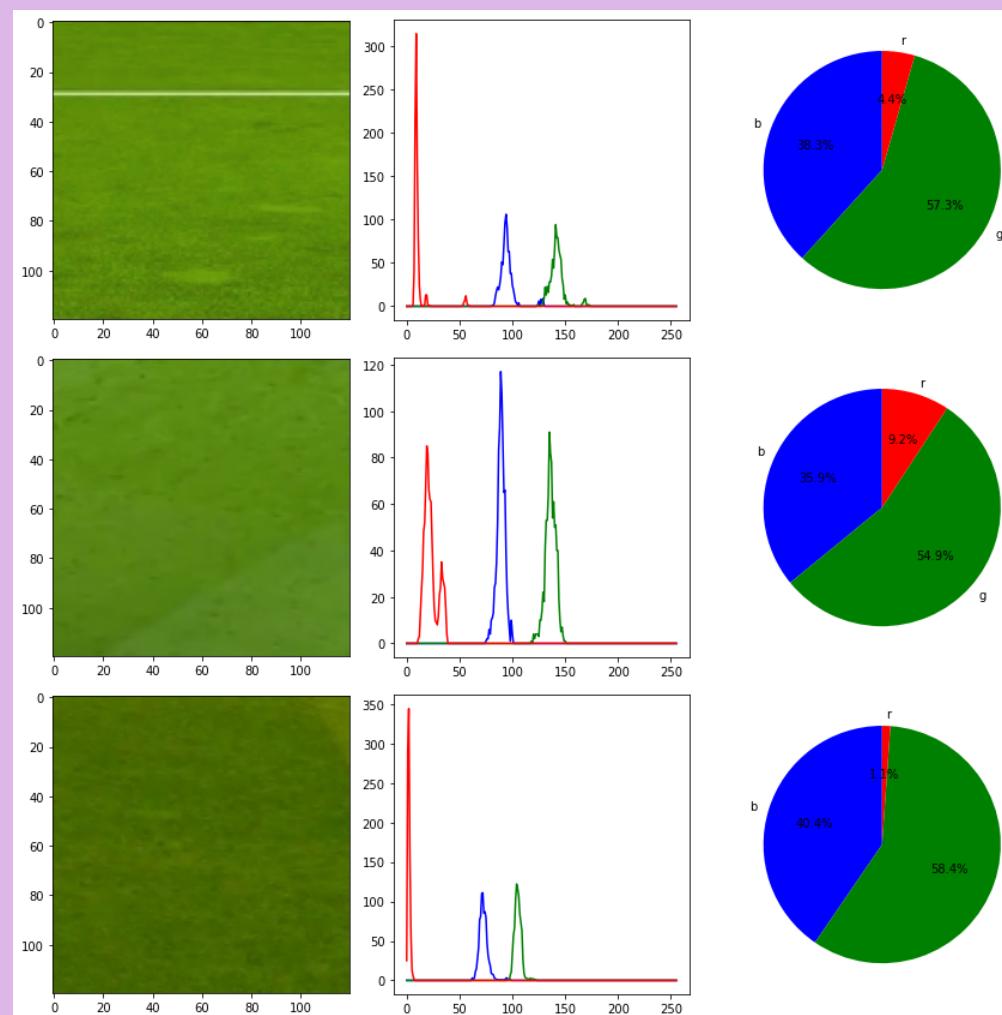
전체 사진에서 고정된 크기의 영역(ex. 120*120)을 crop해서 Batch에 올림 ➔ 모델 학습에 필요 없는 잔디만 crop되는 경우 다수 존재 (20~25%)



04 모델 전처리

RGB 이용해 필터링?

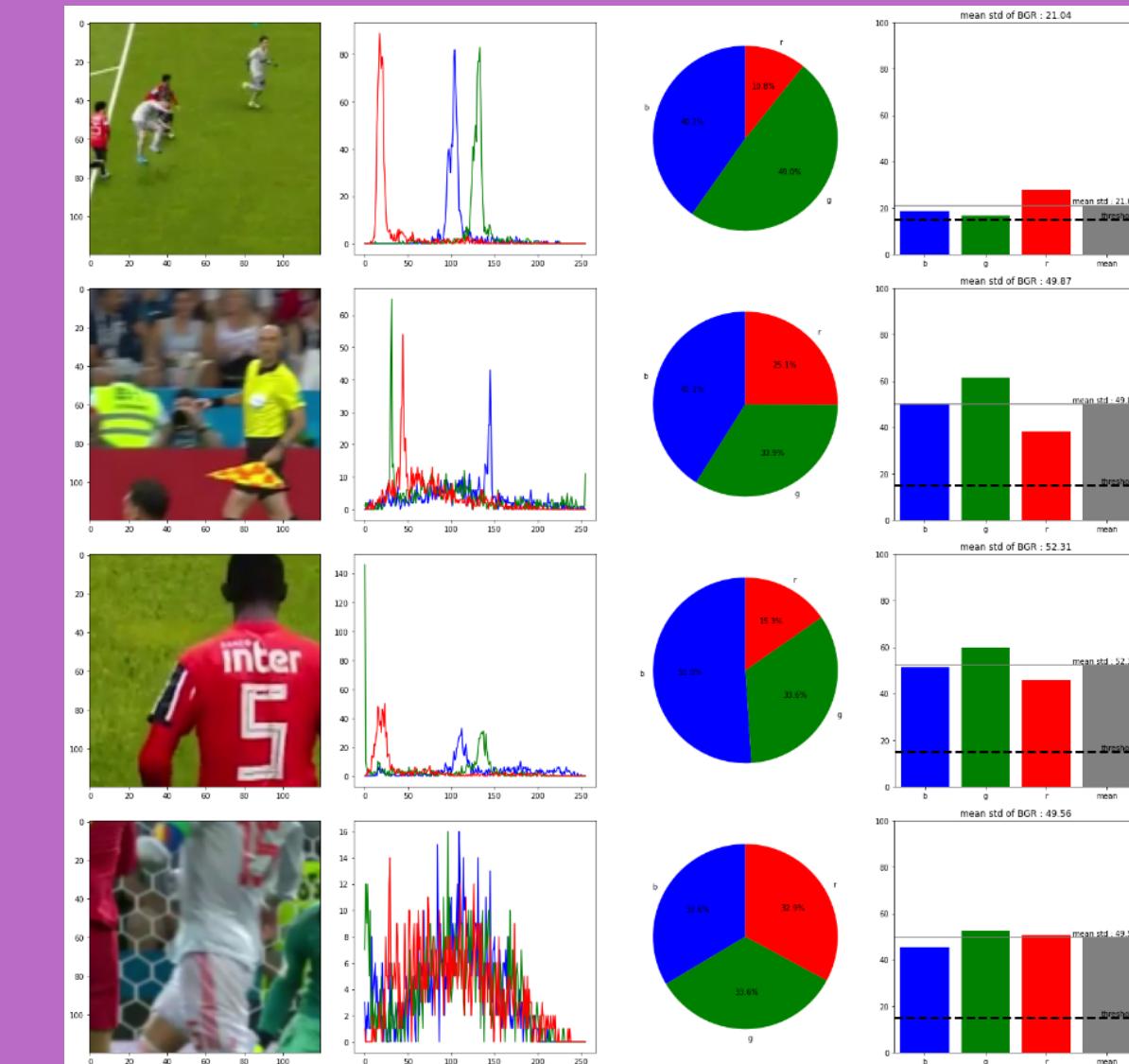
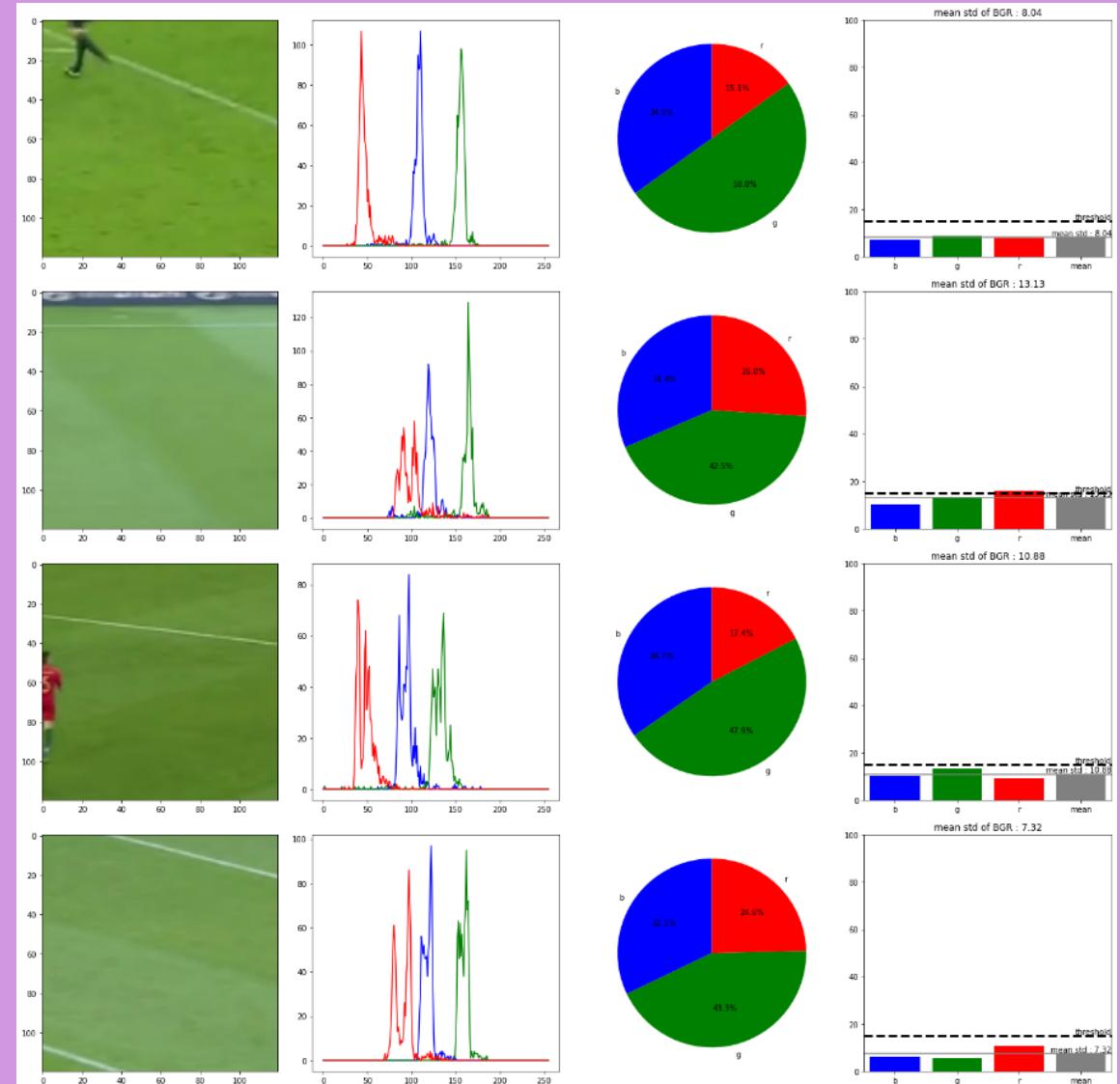
잔디가 녹색으로 보여도 Green 으로만 이루어진 게 아님 → 명확한 기준 잡을 수 없음



04 모델 전처리

RGB의 분산을 이용해 필터링!

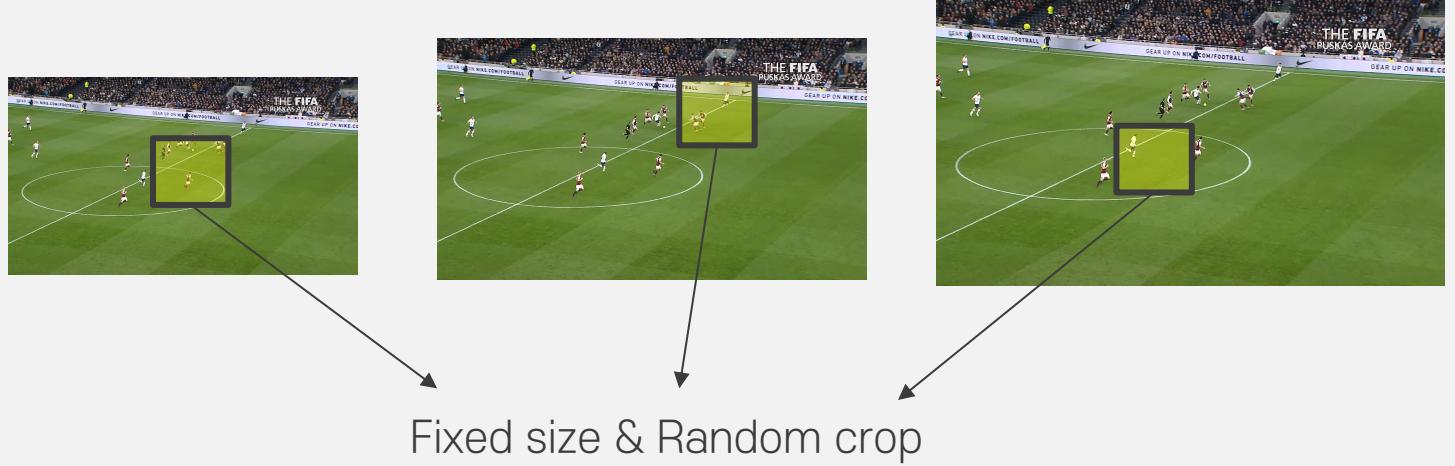
RGB 히스토그램을 보고, 잔디만 있는 경우 색상 분포의 분산이 작다는 것을 발견



04 모델 Augmentation

✓ Resize

Multi-scale에 대응하기 위해 다양한 해상도로 resize



✓ Horizontal Flip

실제 중계 앵글 모습에 집중해서 피팅

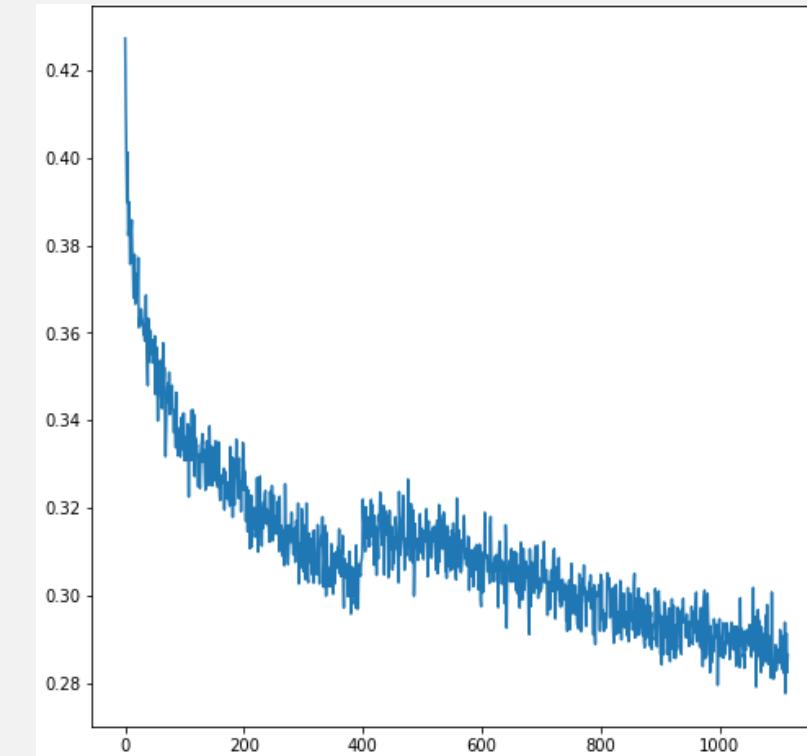
✓ 밝기, 채도, 명암 랜덤 미세 조정

강한 수치 조정 시 원본 데이터의 특징 상실 위험

04 모델 Training

✓ 환경	Google Colab Pro
✓ batch / crop size	8 / 256*256
✓ epoch	1150 epoch
✓ Training 시간	1epoch 당 약 2분

✓ 훈련그래프



05— 결과

이미지 결과

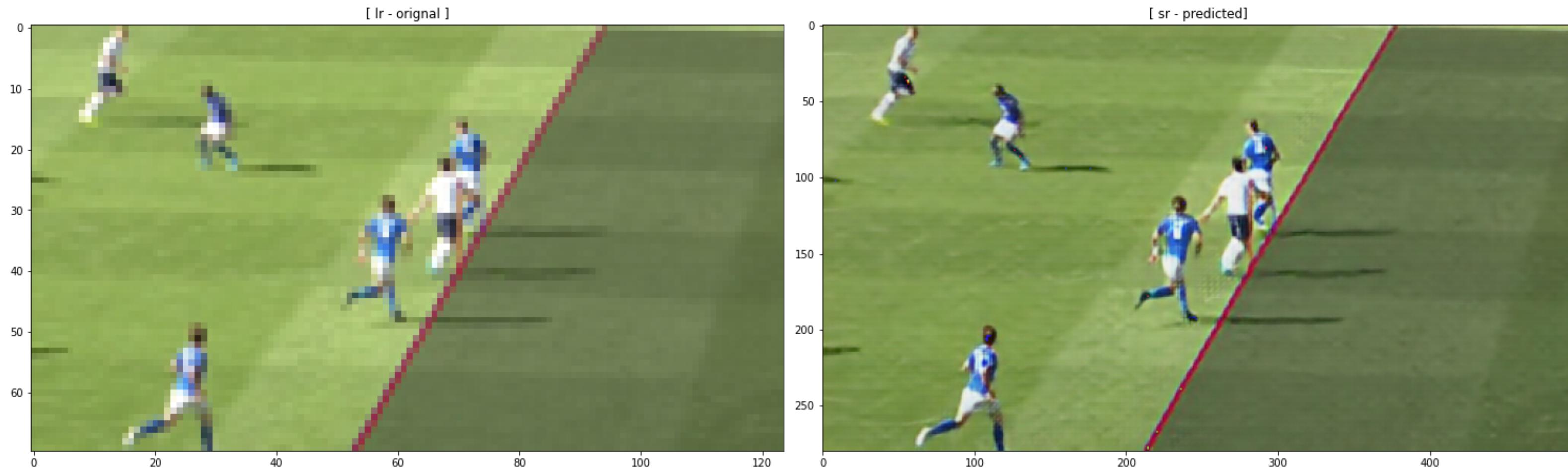
동영상 캡처본

추론 속도

전체 프로세스

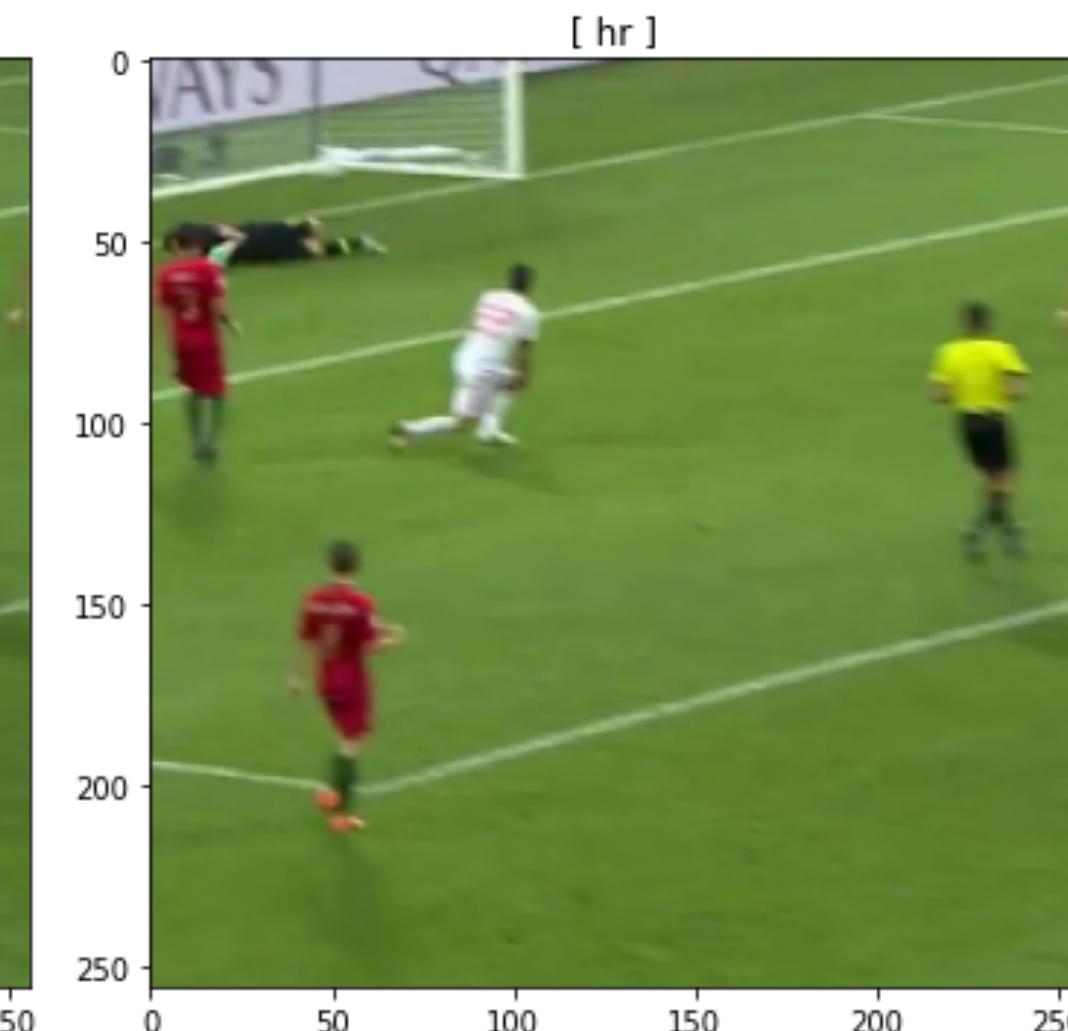
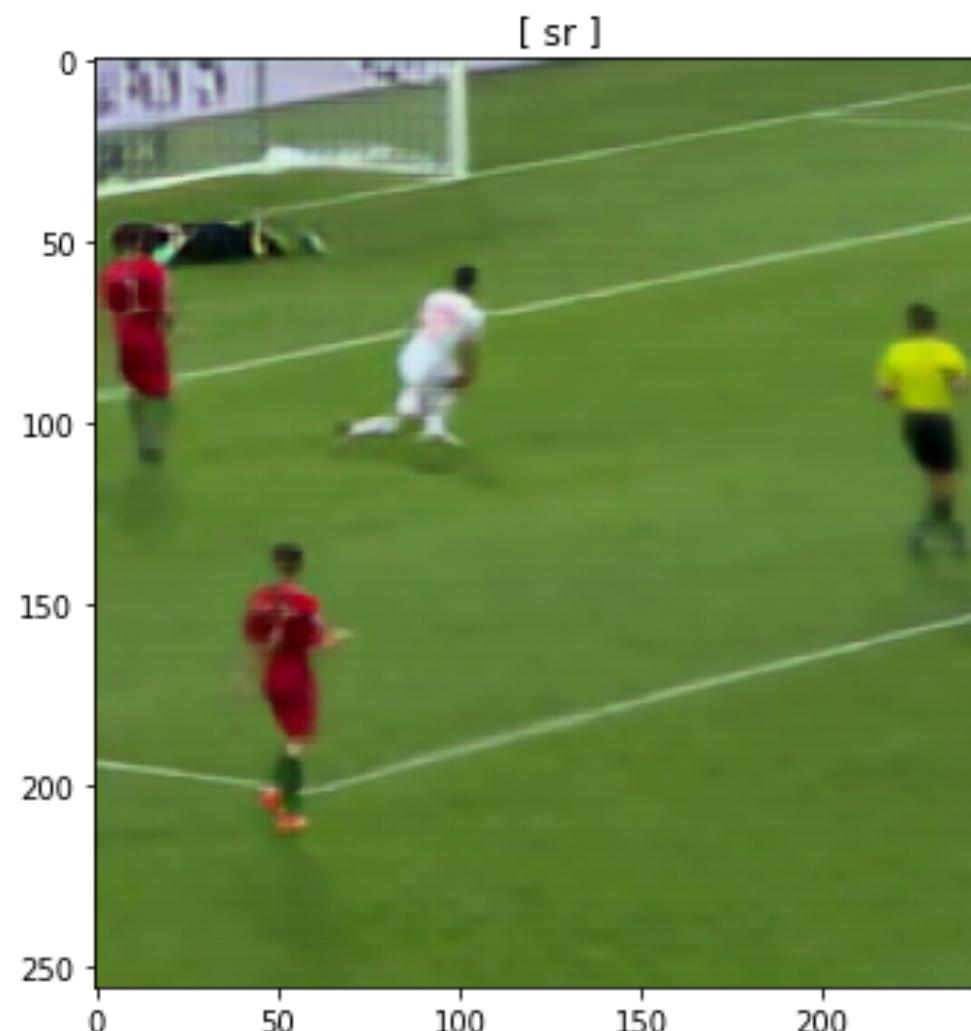
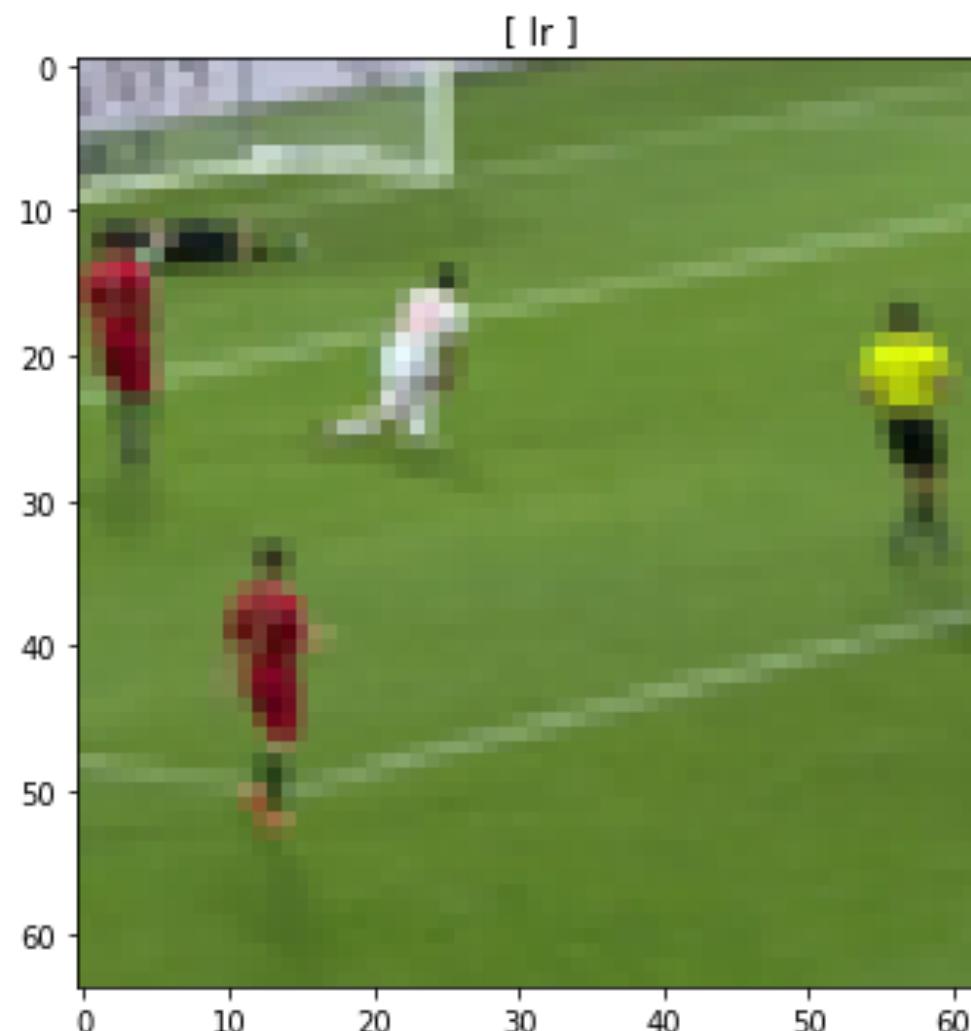
05 _____ 결과 이미지 결과

축구선수



05 _____ 결과 이미지 결과

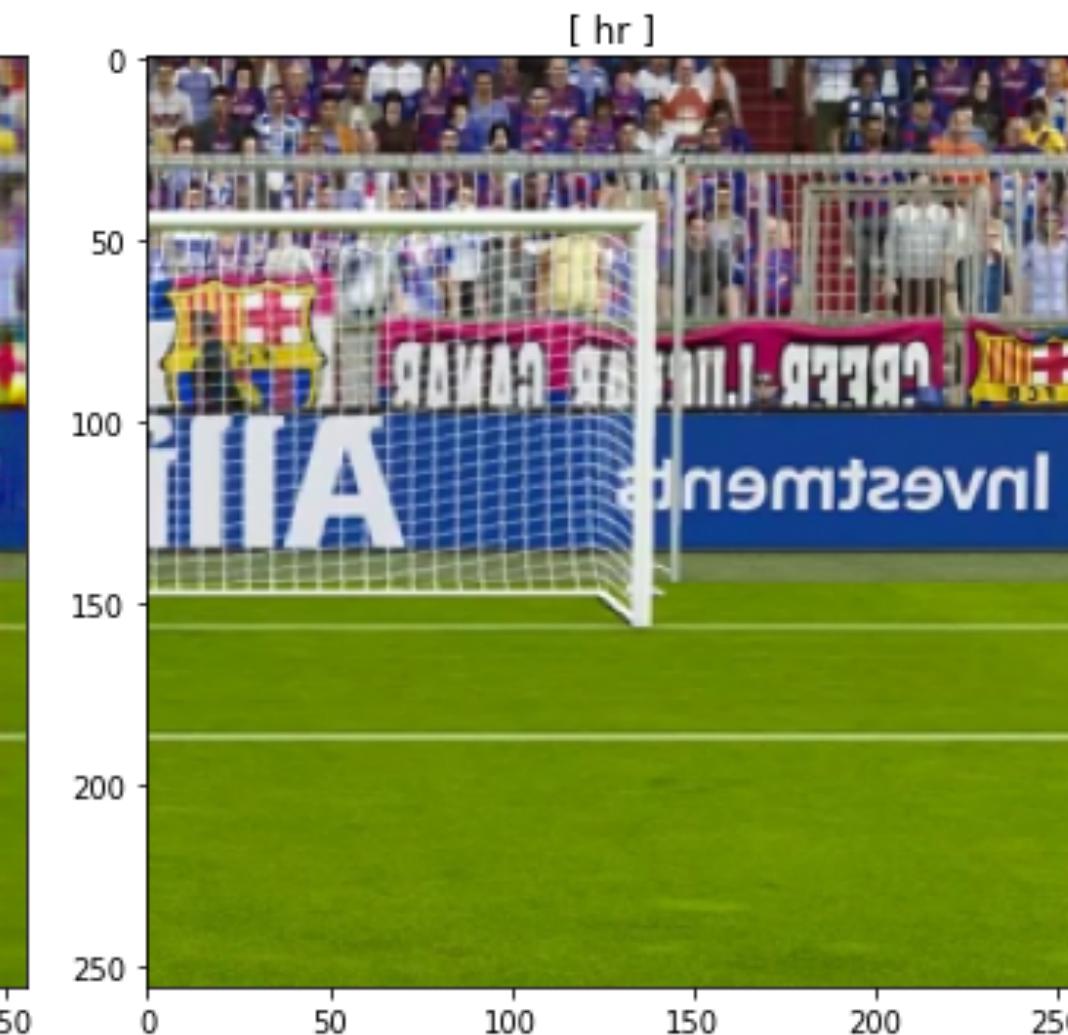
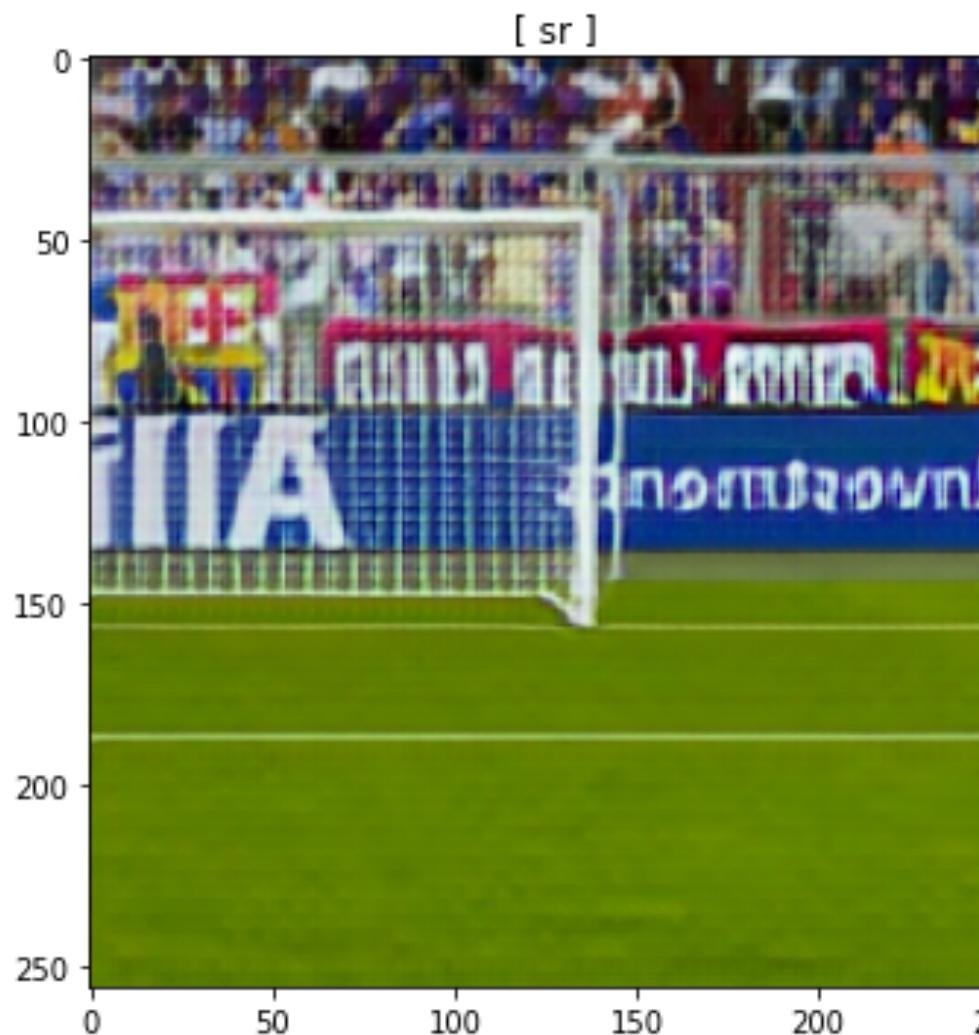
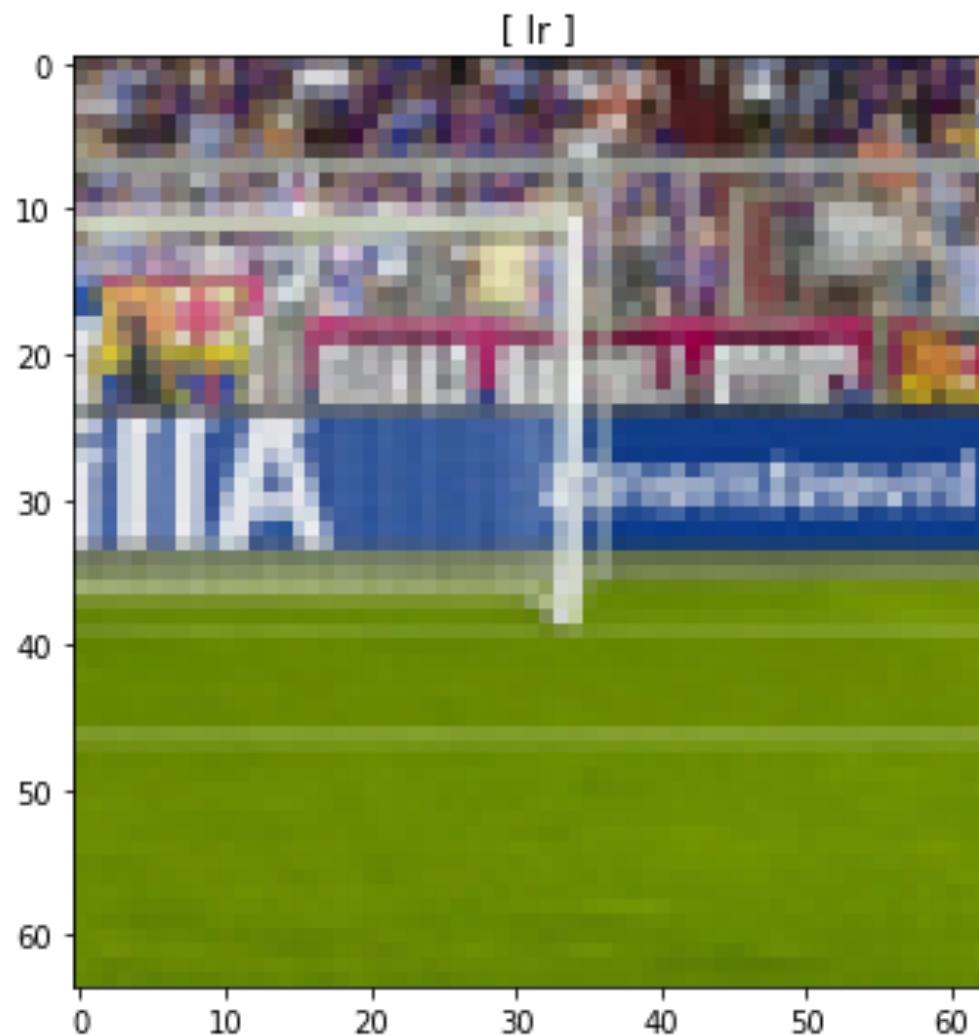
축구선수, 경기장 선



05

결과 이미지 결과

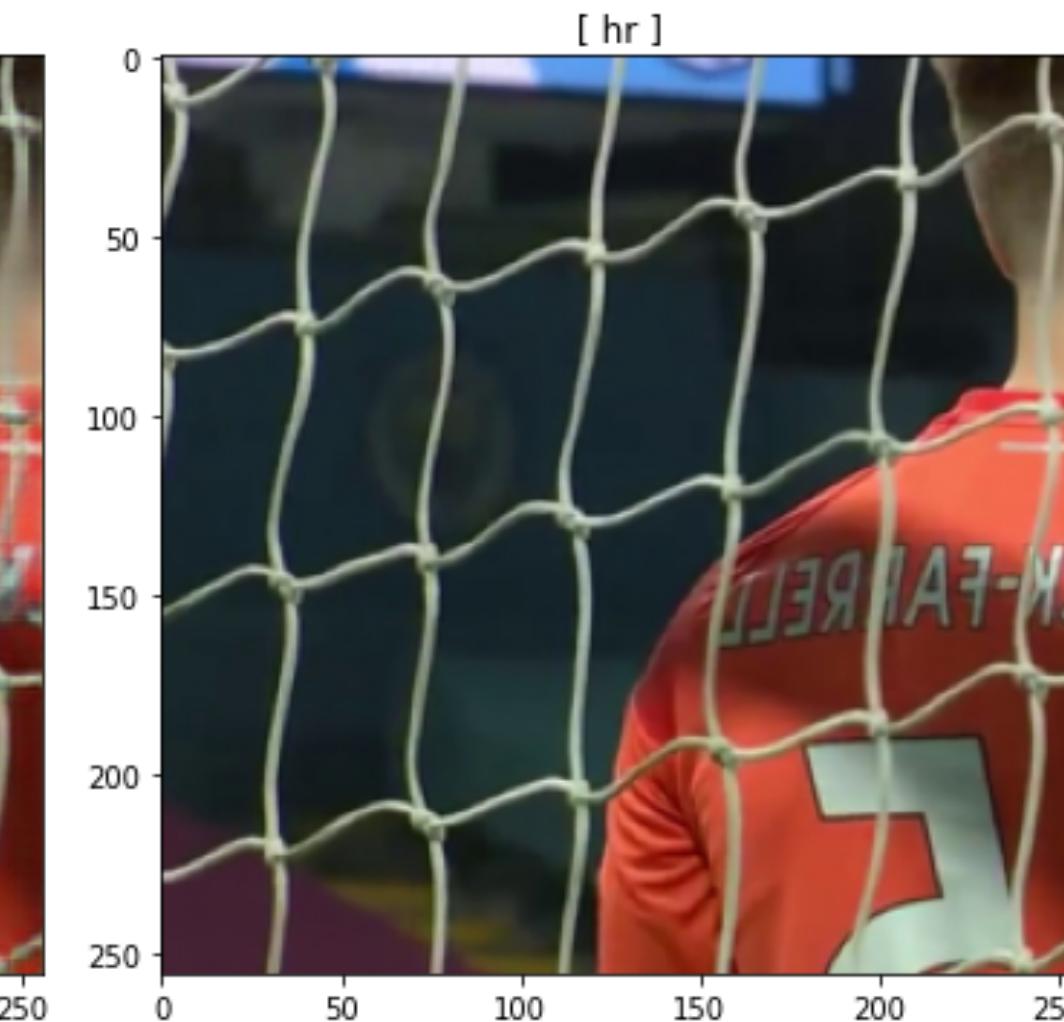
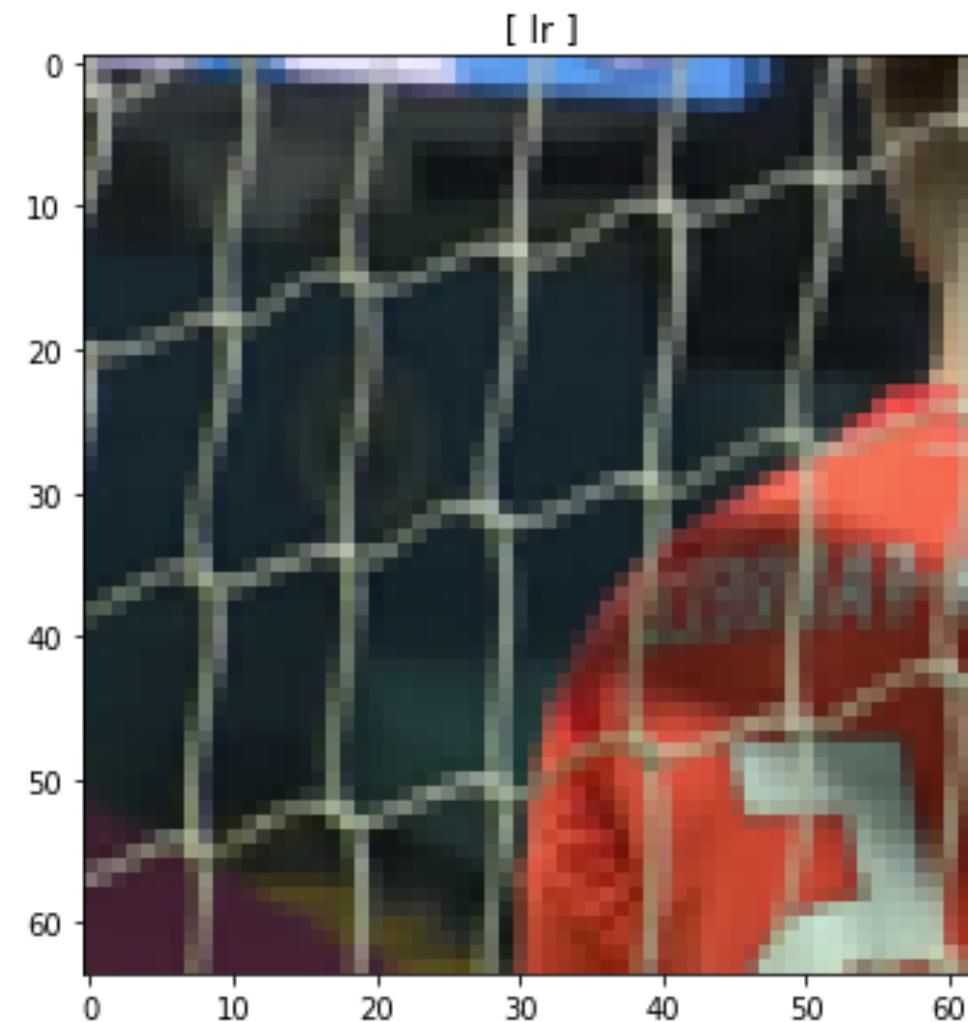
글자, 관중석



05

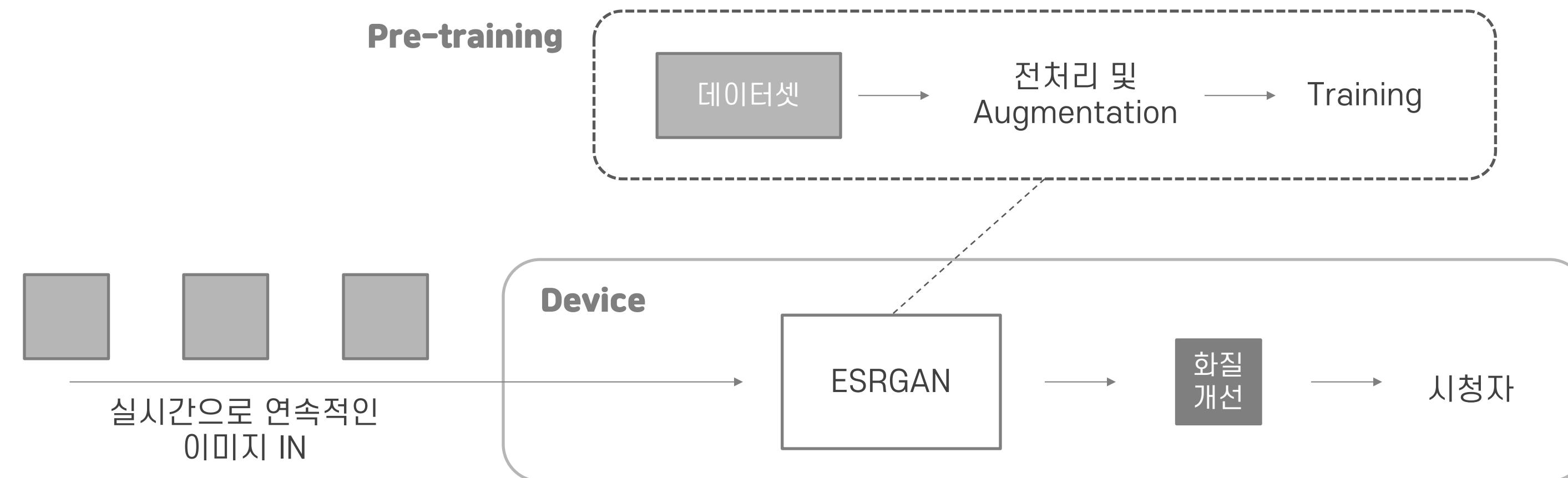
결과 이미지 결과

글자, 골대망



05

결과 전체 프로세스



05 _____ 결과 추론 속도

추론 속도(Inference Speed)

	input	sec/frame	frame/sec
0	120p	0.01581	63.22
1	144p	0.02010	49.73
2	180p	0.03110	32.13
3	240p	0.05680	17.61
4	270p	0.06950	14.39

- 추론 속도 : 학습된 모델에 저화질 영상을 input으로 하여 테스트할 때의 초당 프레임 추출 속도
- 120p, 144p, 180p 추론 속도 : 목표 속도 이상
- Input 영상 화질이 높아질수록 추론 속도 감소, 시간 증가

05 _____ 결과 추론 속도

추론 속도(Inference Speed)

input	sec/frame	frame/sec
0 120p	0.01581	63.22
1 144p	0.02010	49.73
2 180p	0.03110	32.13
3 240p	0.05680	17.61
4 270p	0.06950	14.39

- 추론 속도 : 학습된 모델에 저화질 영상을 input으로 하여 테스트할 때의 초당 프레임 추출 속도
- 120p, 144p, 180p 추론 속도 모드에 따른 영상의 화질 개선 실시간으로 구현
- Input 영상 화질이 높아질수록 추론 속도 감소, 시간 증가

결과

06 ————— 결론

의의 및 한계

06

결론 의의 및 한계

의의

- 이미지 뿐만 아니라 영상 화질 개선이 가능함
- 입력 영상이 저화질일 때 실시간 성능을 낼 수 있음
- 기존에 시도되지 않았던 축구라는 분야에서 한정된 데이터셋으로 품질향상 효과를 달성
- 하나의 모델로 다양한 해상도, 모든 축구경기에 적용 가능

한계

- 네트워크 포함한 전체적인 시스템 보다는 모델 중심 결과
 - 수치보다 실질 체감 화질이 떨어짐
-

06

결론 Reference

Code

ESRGAN-PyTorch : <https://github.com/Lornatang/ESRGAN-PyTorch>

NAS_public : <https://ina.kaist.ac.kr/~nas/>

Paper

NAS : Yeo, Hyunho, et al. "Neural adaptive content-aware internet video delivery." *13th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 18)*. 2018.

ESRGAN : Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.

SRCNN : Dong, Chao, et al. "Image super-resolution using deep convolutional networks." *IEEE transactions on pattern analysis and machine intelligence* 38.2 (2015): 295-307.

ESPCN : Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

SRGAN : Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

DRDVSR : Tao, Xin, et al. "Detail-revealing deep video super-resolution." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

RBPN : Haris, Muhammad, Gregory Shakhnarovich, and Norimichi Ukita. "Recurrent back-projection network for video super-resolution." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019.

iseebetter : Chadha, Aman, John Britto, and M. Mani Roja. "iSeeBetter: Spatio-temporal video super-resolution using recurrent generative back-projection networks." *Computational Visual Media* 6.3 (2020): 307-317.

Source Video

"박지성 가슴 트래핑에 이은 완발슛! 16강행 결정 짓습니다!" 2002년 한일 월드컵 포르투갈전 (다시 보는 골장면) / SBS / 2018 러시아 월드컵 https://www.youtube.com/watch?v=t7-wFHYs0_0

Q&A

감사합니다!