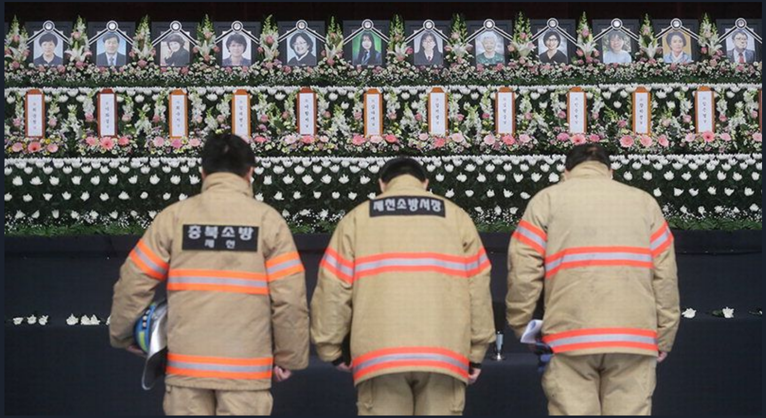
영정사진 화질 개선 딥러닝 모델 개발

이창수

목차

1. 프로젝트 배경
2. 파이프라인
3. 데이터 전처리
4. 모델링
5. 결과
6. 한계점 및 발전 방향

프로젝트 배경



끝까지 제대로 모셔야 하지 않을까요?...소방영웅들의 깨진 영정

소방청이 운영하는 온라인 순직소방관추모관에 올라온 고(故) 김범석 소방관의 사진. 낮은 화질의 사진을 올려 눈코입의 형태를 제대로 알아보기 힘들다. © 뉴스1

- 2017년 12월, 30여명의 사망자를 낸 ‘제천 휘트니스센터 화재’ 취재 때 본 ‘깨진 영정사진’들
- 순직 소방관 온라인 추모관, 쪽방촌 무연고 사망자 장례식 등등
- 이전부터 ‘딥러닝으로 해결할 수 있지 않을까?’ 막연히 생각하던 주제

파이프라인

kaggle → Google Drive → CO PRO → K

<CelebFaces Attributes (CelebA) Dataset>

train : 1700, val : 600, test : 30

https://www.kaggle.com/datasets/jessicali9530/celeba-dataset

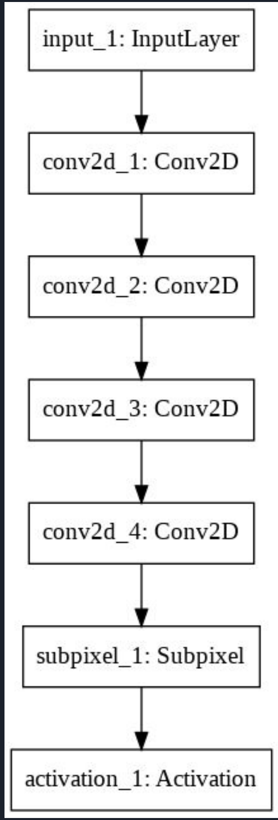
전처리


```
img_sample = cv2.imread(os.path.join(img_base_path, eval_list[123][0]))
h, w, _ = img_sample.shape
crop_sample = img_sample[int((h-w)/2):int(-(h-w)/2), :] ## 정사각형으로 바꿔주기
resized_sample = pyramid_reduce(crop_sample, downscale=4, multichannel=True) ## 4배 줄이기
```

모델링

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 44, 44, 3)	0
conv2d_1 (Conv2D)	(None, 44, 44, 64)	4864
conv2d_2 (Conv2D)	(None, 44, 44, 64)	36928
conv2d_3 (Conv2D)	(None, 44, 44, 32)	18464
conv2d_4 (Conv2D)	(None, 44, 44, 16)	4624
subpixel_1 (Subpixel)	(None, 176, 176, 3)	6960
activation_1 (Activation)	(None, 176, 176, 3)	0

Total params: 71,840
Trainable params: 71,840
Non-trainable params: 0



Is the deconvolution layer the same as a convolutional layer?

A note on Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network.

Wenzhe Shi, Jose Caballero, Lucas Theis, Ferenc Huszar, Andrew Aitken, Alykhan Tejani, Johannes Totz, Christian Ledig, Zehan Wang
Twitter, Inc.[1]

In our CVPR 2016 paper [1], we proposed a novel network architecture to perform single image super-resolution (SR). Most existing convolutional neural network (CNN) based super-resolution methods [10,11] first upsample the image using a bicubic interpolation, then apply a convolutional network. We will refer to these types of networks as high-resolution (HR) networks because the images are upsampled first. Instead, we feed the low-resolution (LR) input directly to a sub-pixel CNN as shown in Fig.1:

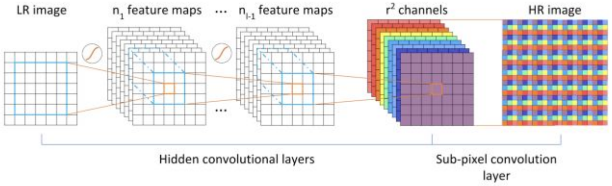


Figure 1: An illustration of the ESCPN framework where r denotes the upscaling ratio.

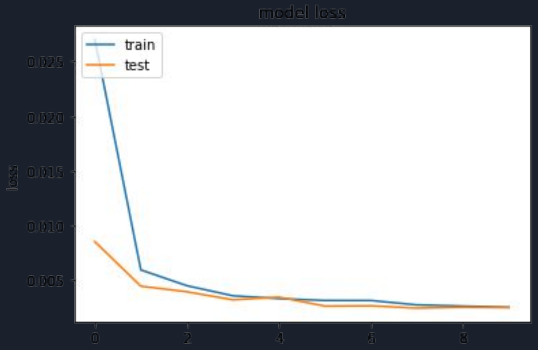
Let r denote the upscaling ratio - e.g if the input LR image is 1×1 then the output HR image will be $r \times r$. We then output r^2 number of channels instead of one high-resolution (HR) image and use periodic shuffling to recreate the HR image. The exact details about how our efficient sub-pixel convolutional layer works can be found in the paper. We will refer to our network as a LR network.

In this note, we want to focus on two aspects related to two questions most people asked us at CVPR when they saw this network. Firstly, how can r^2 channels magically become a HR image? And secondly, why are convolution in LR space a better choice? These are actually the key questions we tried to answer

https://github.com/atriumlts/subpixel

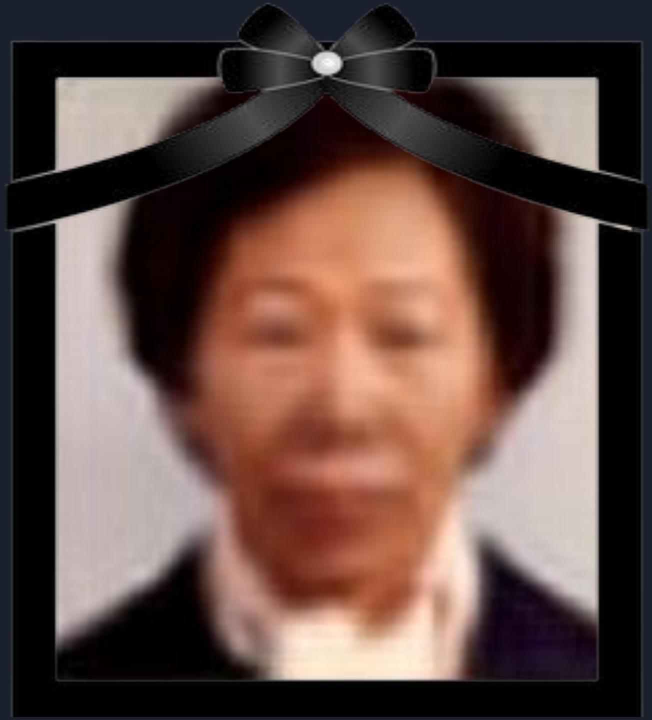
https://github.com/kairess/super_resolution

결과



train loss: 0.0025
val_loss: 0.0025

결과



한계점 및 추후 보완 방향

1. CPU 성능 등 물리적인 한계 존재
2. 비교적 작은 학습 데이터 -> 아쉬운 성능으로 이어짐
3. 학습 데이터 이미지 크기(4k~12k)에 따른 한계 존재

1. 모델의 학습 데이터셋 퀄리티&퀀티티 향상
2. SRGAN 모델링 시도