

이미지 고해상도 복원 서비스 설계 및 개발에 대한 연구

유원준, 오정석, 정지환, 조민정
명지대학교 ICT 융합대학 융합소프트웨어학부

요 약

본 논문에서는 이미지 고해상도 복원 플랫폼의 제안하고 프로토타입을 구현하였다. 프로토타입에서는 BSRGAN 모델 플랫폼을 기반으로 이미지를 고해상도로 복원할 수 있고, 애플리케이션 플랫폼을 구축하여 이미지를 영구적으로 저장하고 공유할 수 있다. 따라서 과거 아날로그 사진이 갖는 보관의 어려움을 극복하고, 기록물로서의 본래 가치를 재현할 수 있을 것이다.

Keywords: Generative model, Blind super resolution, Generative Adversarial Networks

1. 서론

사진의 사전적 의미는 '실물의 모양을 있는 그대로 그려 낸다'는 의미로, 기록물의 관점에서 사진은 지난 사실을 기록하고 그대로 재현할 수 있다. 이와 더불어, 사진은 기록한 순간의 감정과 경험을 전달하는 기능도 수행할 수 있다. 이를 통해 사진은 개인적인 경험 소유의 영역을 넘어, 그 자체로 사회적 가치를 지닐 수 있게 된다.

2. 문제 설정 및 목표

2.1 문제 정의

과거에는 아날로그 사진이 주로 사용되었으며, 이러한 사진들은 시간의 흐름에 따라 훼손되거나 품질이 저하된다. 또한, 초기 디지털카메라 기술의 한계 때문에 이때 찍은 사진들의 화질도 좋지 않았다. 이러한 저해상도 사진의 품질 저하 문제는 디지털 형식으로 보존하거나 온라인에서 공유하려는 경우 큰 제약 요인이 된다.

2.1.1 아날로그 사진 디지털화

사람들의 소통 공간이 오프라인 환경에서 온라인 환경으로 전환됨에 따라 사진 또한 디지털화되어야 한다.

2.1.2 아날로그 사진 복원

디지털화된 아날로그 사진은 해상도가 낮고 노이즈와 왜곡이 발생할 수 있다. 이러한 문제를 해결하기 위해 복원 기술이 필요하다.

2.1.3 아날로그 사진 디지털

초기 모바일 디바이스의 카메라로 촬영된 사진들은 해상도가 낮아 세부 정보가 흐려지거나 손실될 수 있다. 이러한 저해상도 사진을 고해상도로 복원할 기술이 필요하다.

2.2 연구 목표

본 연구의 구체적인 목표는 다음과 같다.

1. 과거 아날로그 사진을 디지털 형식으로 변환하고 저화질의 디지털 이미지를 고해상도 복원 기술을 활용하여 이미지의 해상도를 향상시킨다.
2. 아날로그 사진 및 저화질 디지털 이미지를 쉽게 복원하고 공유할 수 있도록 사용자 친화적인 인터페이스 구현한다.
3. 검증을 통해 개발한 앱의 성능을 평가하고, 기존의 기술 및 서비스와 비교하여 향상된 결과를 제시한다.

3. 제안하는 방법

3.1 시스템 아키텍처

그림 1. 본 연구에서 제공한 서비스의 시스템 아키텍처

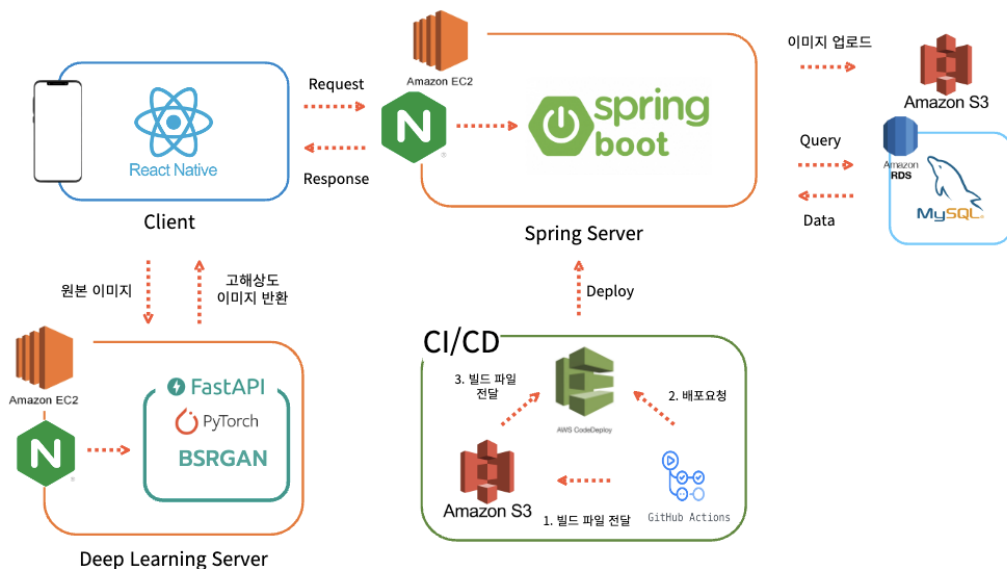


그림 1 은 본 논문에서 설계 및 개발한 서비스의 시스템 아키텍처의 그림이다.

프론트엔드에서는 React 를 사용하였으며, iOS 및 Android 기기 모두에서 실행할 수 있는 교차 플랫폼 앱을 구축하였다. 백엔드에서는 딥러닝 모델 추론을 위한 서버와 웹 빌드를 위한 서버로 각각 FastAPI, Spring boot 로 총 2 개의 서버 환경을 구축하였다. 이미지 데이터는 별도의 AWS S3 저장소에서 관리한다. 딥러닝 서버는 EC2 , FastAPI 를 통해 Pytorch 라이브러리를 사용하여 본 연구에서 사용하는 BSRGAN 딥러닝 모델을 구동시킨다. 또한 별도의 CI/CD 파이프라인을 구축하였고, 코드 품질 관리와 더불어 개발 후 자동 배포가 가능한 시스템을 갖추고 있다.

3.2 Single Image Super Resolution

SISR 방법이 low-resolution(LR) 이미지에서 high-resolution(HR) 이미지로의 사상이라면, degradation model 은 HR 이미지에서 LR 이미지로의 사상이다. 대표적인 degradation 모델은 traditional degradation 과 bicubic degradation 이 있다. 식 (1)은 traditional degradation 을 수학적으로 모델링한 수식이다.

$$y = (x \otimes k) \downarrow_s + n \quad (1)$$

$(x \otimes k)$ 는 Gaussian kernel k [14] degradation 이 존재하는 저해상도의 사진, $\downarrow_s$ 는 scale factor s 로 down-scaling 된 작업, n 은 표준편차를 σ 로 갖는 white Gaussian noise 를 의미한다.

Single image super resolution (SISR) 방법은 실제 사진에서 좋지 않은 성능을 보이지 못했다. 그 이유는 SISR 의 비약적인 가정 때문이다. 원본 데이터라고 생각하며 우리가 SISR 과정을 통해 얻어내고자 하는 고해상도의 사진과 실제 우리가 가지고 있는 데이터의 사이에 있어서 식(1)처럼 단순한 한가지의 degradation(열화)만을 가정하기 때문이다. 실제 사진에는 한가지의 degradation 만이 존재하기보다 다양한 종류의 degradation 이 여러 종류로 중첩되어 있을 수도 있고, 사람이 이해하기 힘든 복잡한 함수로 존재하는 degradation 일 수도 있다.

따라서 SISR 으로부터 blind SISR 로 연구가 계승되어 발전되었다. blind SISR 은 다양한 degradation, 복잡한 degradation 또는 degradation 에 대한 사전 지식을 모르는 상황을 가정한다. 이를 통해 사진에 대한 복원능력을 높인다. 본 논문에서 BSRGAN [1]을 기반으로 추가적인 모델 학습을 통해 설정한 문제 상황에 맞게 모바일환경에서 아날로그 사진을 촬영한 사진을 대상으로 blind SISR 모델로서 사용한다.

3.3 GAN 기반의 생성 모델

Generative Adversarial Networks(GAN)는 Generator 와 Discriminator 로 크게 두 부분으로 나눌 수 있다. Generator 는 입력 데이터의 분포를 모방하여 생성한 데이터와 실제 데이터를 discriminator 가 구분할 수 없게 하는 것이 학습의 목표가 되며, 반대로 discriminator 는 실제 데이터와 generator 가 생성한 데이터를 구별할 수 있는 것이 학습의 목표가 된다. 이러한 경쟁하는 학습과정을 Adversarial networks [2]라고 한다.

Ian Goodfellow 에 의해 처음 2014 년에 제안된 Vanilla GAN [2]을 시작으로 SRGAN [3], CycleGAN [4] 그리고 BSRGAN [1] 등 여러 GAN 기반의 생성 모델들이 발전되어 왔다.

3.4 BSRGAN

학습 과정에 사용하는 저해상도 사진을 만드는 Degradation model 의 아키텍처가 BSRGAN [4]의 핵심 아이디어이다. Degradation 을 가할 때 blur, downsampling, noise 그리고 random shuffle 4 가지 방법을 사용해 저해상도 사진을 만든다.

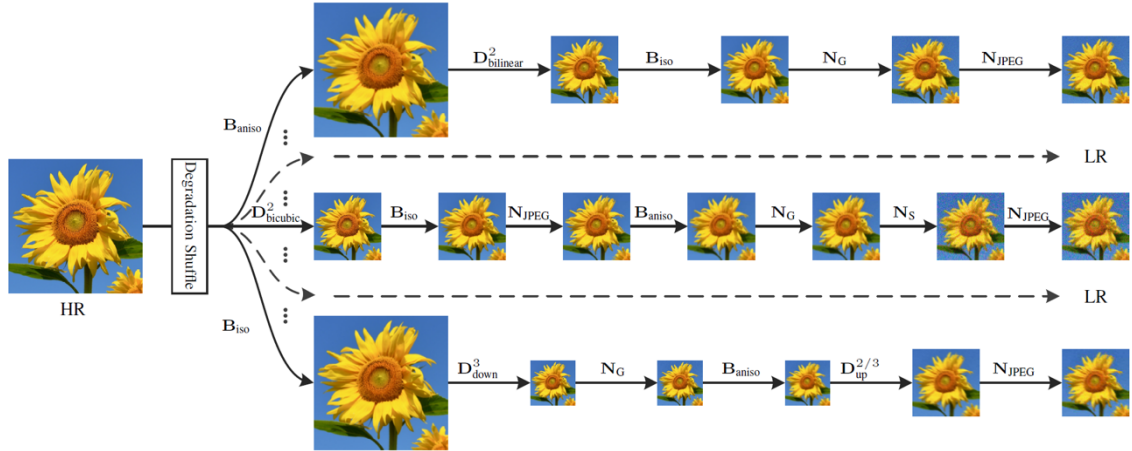
Blur 에서는 isotropic Gaussian blur B_{iso} , anisotropic Gaussian blur B_{aniso} 두 가지 방법을 사용하는데 Gaussian blur 란 2 차원 행렬로 표현할 수 있는 Gaussian function 을 사용해서 이미지에 대해 convolution 연산을 수행하는 것이다.

Downsampling 에서는 scale factor s 를 사용하여 nearest neighbor downsampling, bilinear downsampling, bicubic downsampling 그리고 bilinear 와 bicubic downsampling 중 하나를 무작위로 수행하고 다시 upsampling 하는 down-up-sampling [1] $D^s = \{D_{nearest}^s, D_{bilinear}^s, D_{bicubic}^s, D_{down-up}^s\}$ 4 가지 방법이 있다. 4 가지 방법 중에서 한 가지를 무작위로 선택하여 downsampling 을 진행한다.

Noise 에서는 Gaussian noise N_G , JPEG compression noise N_{JPEG} , Processed camera sensor noise N_S 3 가지의 noise 종류가 있다. Gaussian noise 는 아무런 정보가 없는 노이즈로서 가장 보수적인 노이즈이다 [11]. JPEG 는 가장 잘 알려진 사진 압축방식으로, 사진 데이터의 저장 및 전송에 주로 쓰이는 압축방식이다. JPEG compression noise 는 이 때 발생하는 손실 노이즈이다. Processed camera sensor noise 는 디지털카메라에서 발생하는 노이즈이다. 디지털카메라는 Image signal processing(ISP) 과정을 통해 카메라 센서를 거쳐 디지털 이미지를 얻는데 이 이미지는 non-Gaussian noise [12]의 한 종류인 센서 노이즈가 포함되어 있다.

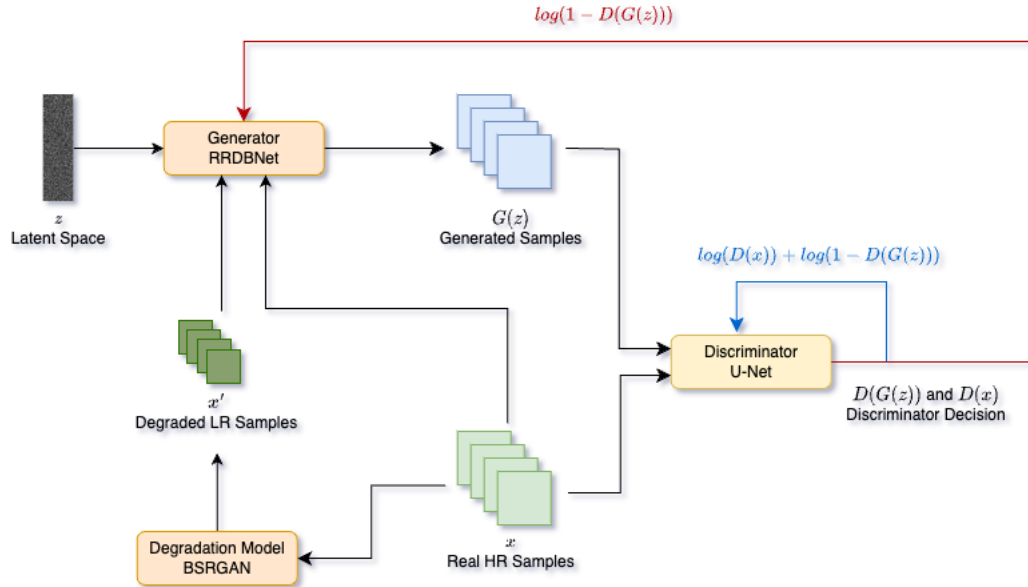
Random shuffle 에서 위에 언급한 blur, downsampling, noise 방법들을 무작위로 섞은 순차적이고 연속적인 degradation $\{B_{iso}, B_{aniso}, D^s, N_G, N_{JPEG}, N_S\}$ 을 수행하여 low-resolution (LR) 이미지를 만든다. 그림 2 은 원래 HR 이미지 데이터에 대해 degradation 과정을 거쳐 LR 이미지 데이터를 만드는 과정에 대한 그림이다.

그림 2. BSRGAN 이 Degradation 수행 과정 [1]



본 논문에서는 Degradation model로 BSRGAN, Generator로 RRDBNet [17], Discriminator로 U-Net [18]을 사용한다. 그림 3은 이미지 고해상도 복원 서비스에 설계하고 구현한 딥러닝 모델의 알고리즘 파이프라인 아키텍처이다.

그림 3. 전체 알고리즘 파이프라인 아키텍처



4. 데이터셋

4.1 Pre-training, fine-tuning 데이터셋

본 논문의 문제상황이자 연구 목표인 아날로그 사진 및 저화질 디지털 이미지의 복원에 맞게 범용의 Blind 상황에 대한 이미지 고해상도 작업이고, image degradation prior[1]를 학습하기 위해 pre-training 데이터셋으로 학습한 모델 파라미터를 weight initialization에 사용하였다. Pre-training 데이터셋은 BSRGAN의 Generator인 RRDBNet [17]의 weight initialization에 사용한 데이터셋으로 DIV2K [5], Flick2K [6,

7], WED [8], FFHQ [9] 데이터셋이 포함되어 있다. 또한 DRealSR [10] 데이터셋으로 추가적인 fine-tuning 학습을 진행하였다.

4.2 Evaluation 데이터셋

모델 검증을 위해 Pre-training 데이터셋과 Fine-tuning 데이터셋이 아닌 별개의 데이터셋을 사용한다. DIV2K [5] 의 validation 데이터셋인 DIV2KRK [15]와 RealSRSet [1] 총 120 장의 이미지 데이터 샘플을 가진 검증 데이터셋을 모델 검증에 사용하였다. 표 1 은 모델 평가에 사용한 데이터셋의 예시 이미지 데이터 정리표다.

표 1. DIV2KRK[15]와 RealSRSet [1]의 예시 이미지 데이터 정리표



5. 실험

5.1 모델 학습

본 논문에서는 DRealSR 전체 데이터를 모두 학습에 사용하였다. 학습환경은 AMD Ryzen 7 5700X / 32GB RAM / Nvidia RTX 3080 12GB / Ubuntu 20.04 LTS 에서 진행하였다. 하이퍼파라미터(Hyperparameter)는 해당 논문에서 제시한 기본 설정으로 주어진 그래픽카드와 메모리 하드웨어 환경에 적합한 하이퍼파라미터를 맞추어 batch size 를 4 로 설정해 학습을 진행하였다. RRDBNet [17]과 U-net [18] 모두 Adam 을 optimizer 로 learning rate 는 $5e-5$ 로 사용하였다. 모델의 Iteration 은 65000 까지 진행하였고, overfitting 을 고려하여 모델 평가 메트릭인 Average PSNR, RRDBNet 의 loss, U-net 의 loss 를 확인하였다. 표 2 는 모델 학습에 사용한 하이퍼파라미터에 관한 표다. Patch size 는 Degradation 을 가할 때 이미지를 작은 크기로 나눌 때의 해상도이고, scale factor 는 저해상도 이미지와 고해상도 이미지의 해상도 크기 비율이다. 배치 당 데이터 수인 batch size 는 둘 다 4 이고, learning rate 는 $5e-5$ 으로 설정하였다. Optimizer 는 Adam 을 사용하였고 RRDB 가 3 개로 구성된 RRDBNet 을 사용하였다.

표 2. 모델 학습에 사용한 하이퍼파라미터

Model		Hyperparameter	
Degradation	BSRGAN	patch size	72*72
Generator G	RRDBNet	scale factor	4
		batch size	4
		learning rate	5e-5
		optimizer	Adam
		RRDB (Residual in residual dense block)	3
Discriminator D	U-Net	batch size	4
		learning rate	5e-5
		optimizer	Adam

5.2 실험 결과

정량적으로 Generator G가 생성한 고해상도 이미지와 원래 고해상도 이미지와 유사한지 평가하기 위해 이미지의 품질을 평가하는 데 사용하는 Average PSNR(Peak Signal-to-Noise Ratio)을 사용하였다. Average PSNR의 수식은 다음과 같다.

$$MSE = \frac{1}{2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

$$\text{Average PSNR} = \frac{1}{m} \sum_{i=1}^m 10 \log \frac{255^2}{\text{MSE}_i} \quad (3)$$

식 2에서 Y_i 는 원래 고해상도 이미지의 RGB 픽셀값이고 $\hat{Y}_i$ 는 G가 생성한 이미지의 RGB 픽셀값이다. 총 n 개의 픽셀이 존재하는 같은 해상도의 이미지 쌍에 대한 MSE를 계산한다. MSE만으로는 지각적으로 의미 있는 이미지 품질 척도를 알 수 없으므로 식 2를 사용해 식 3을 유도해 사용한다. 식 3에서 255는 8-bit 이미지의 최대 픽셀값을 의미하고 이를 검증 데이터셋 내의 모든 이미지에 대한 평균을 산출하였다. 표 3은 각각의 모델이 생성한 이미지 데이터에 대한 Average PSNR 계산 결과를 나타낸 표이고 표 3의 굵은 숫자는 가장 높은 Average PSNR을 의미한다.

표 3. 각 모델의 Average PSNR 계산 결과

Model	Fine-tuned BSRGAN (Ours)	KernelGAN [15]	DSAR [16]	BSRGAN [1]
Average PSNR (dB)	28.2	26.1	28.2	26.2

그림 5는 본 연구의 모델 학습 과정에서 iteration 200 단위마다 측정한 Average PSNR로 iteration 65000까지 학습을 진행하였다.

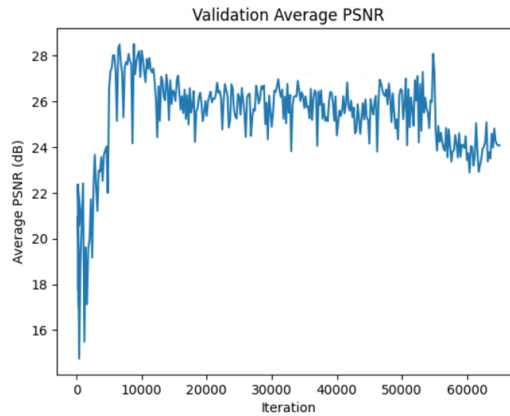


그림 5. 검증 데이터셋에 대한 Average PSNR 계산 결과

그림 6은 본 연구의 모델 학습 과정에서 iteration 200 단위마다 측정한 Generator RRDBNet의 loss와 Discriminator U-Net의 loss다. Loss 함수에는 Generator와 discriminator를 고려한 pixel-wise reconstruction loss, L1 loss와 정성적 평가를 위해 VGG perceptual loss [3]가 포함되어있다.

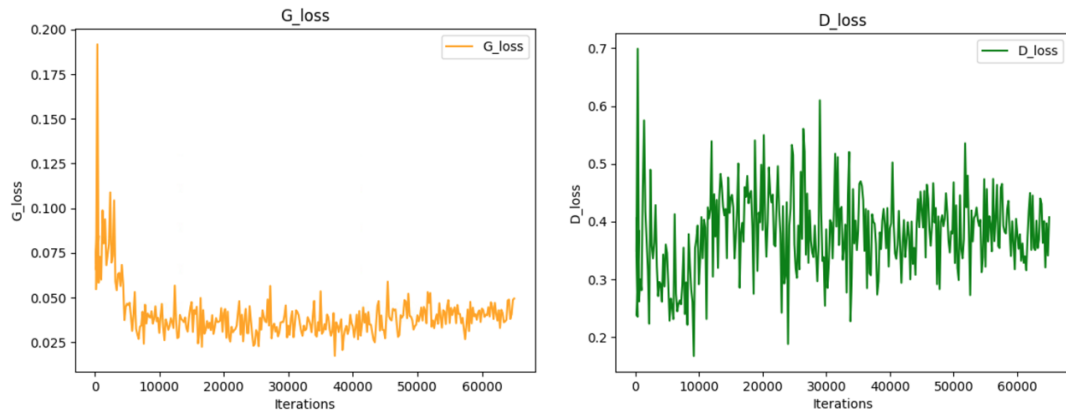


그림 6. 학습 데이터셋에 대한 RRDBNet 의 loss 와 U-Net 의 loss 계산 결과

표 4. 원본, BSRGAN(scale factor 2), BSRGAN(scale factor 4) 결과 이미지 데이터 비교 정리



표 4 는 원본 2151*1512 해상도의 이미지 데이터를 모델 학습 및 추론 환경에 맞게 해상도를 낮춘 후 각 scale factor 에 따라 고해상도로 복원한 정리표이다.

5.4 서비스 검증

서비스의 정량적인 검증을 위해 이미지 고해상도 복원 서비스에 대한 SNS 상의 불특정 다수에게 설문조사를 진행하였다. 그 결과 우리 서비스에 관심을 갖는 타겟층은 20 대 여성이었으며 사진에 관심이 많고 감성적인 사람이 다수를 이루었다. 기능 및 서비스에 대한 측정을 진행한 결과 복원된 사진의 화질과 정확도에 대한 평가로 평점 3.74 점, 앱의 전반적인 만족도에 대한 평가로 평점 3.82 점으로 평점 3.5 점 이상을 달성하였다.

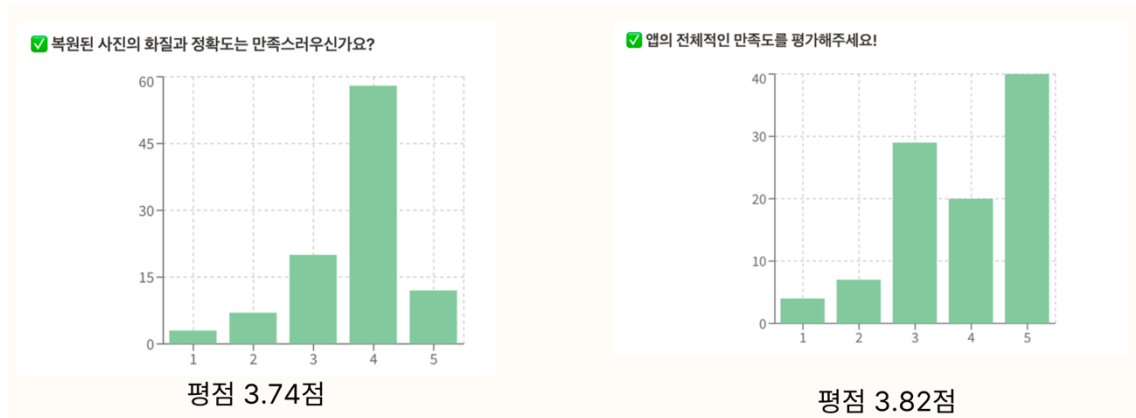


그림 7. 기능 및 서비스 측정 설문조사 결과

6. 결론

본 논문에서는 이미지 고해상도 복원 플랫폼의 제안하고 프로토타입을 구현하였다. 프로토타입에서는 BSRGAN 모델 플랫폼을 기반으로 degradation 을 모르는 이미지를 고해상도로 복원할 수 있고, 애플리케이션 플랫폼을 구축하여 이미지를 영구적으로 저장하고 공유할 수 있다. 따라서 과거 아날로그 사진이 갖는 보관의 어려움을 극복하고, 기록물로서의 본래 가치를 재현할 수 있을 것이다. 본 논문의 연구는 아날로그 사진을 촬영한 디지털 이미지라는 degradation 이 사전에 정의되지 않은 상황에서의 이미지 복원을 하고 이를 모바일 어플리케이션에서 서비스를 제공했다는 점에서 의의가 있으나 영상이 아닌 이미지만을 그 대상으로 한 것, 모델 학습 환경의 부족한 것을 극복하지 못하고 이미지의 해상도를 낮추어 추론한 것, 정성적 평가에 대해, 모델 배포 서버의 비용 한계에 맞추어 모델 양자화를 하지 못한 것이 한계점이다.

참고 문헌

- [1] Zhang, Kai, et al. "Designing a practical degradation model for deep blind image super-resolution." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.
- [2] Goodfellow, Ian, et al. "Generative adversarial networks." Communications of the ACM 63.11 (2020): 139–144.
- [3] 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.
- [4] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.
- [5] Agustsson, Eirikur, and Radu Timofte. "Ntire 2017 challenge on single image super-resolution: Dataset and study." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.

- [6] Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.
- [7] Timofte, Radu, et al. "Ntire 2017 challenge on single image super-resolution: Methods and results." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017.
- [8] Ma, Kede, et al. "Waterloo exploration database: New challenges for image quality assessment models." IEEE Transactions on Image Processing 26.2 (2016): 1004–1016.
- [9] Karras, Tero, Samuli Laine, and Timo Aila. "A style-based generator architecture for generative adversarial networks." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.
- [10] Wei, Pengxu, et al. "Component divide-and-conquer for real-world image super-resolution." Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII 16. Springer International Publishing, 2020.
- [11] Sangwoo Park, Erchin Serpedin, and Khalid Qaraqe. "Gaussian assumption: The least favorable but the most useful [lecture notes]." IEEE Signal Processing Magazine 30.3 (2013): 183–186.
- [12] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In CVPR, pages 1586–1595, 2017.
- [13] Vint, David, Matthew Anderson, Yuhao Yang, Christos Ilioudis, Gaetano Di Caterina, and Carmine Clemente. "Automatic Target Recognition for Low Resolution Foliage Penetrating SAR Images Using CNNs and GANs" Remote Sensing 13, no. 4 (2021): 596.
- [14] N. Efrat, D. Glasner, A. Apartsin, B. Nadler, and A. Levin, "Accurate Blur Models vs. Image Priors in Single Image Super-resolution," 2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2013, pp. 2832–2839,
- [15] Bell-Kligler, Sefi, Assaf Shocher, and Michal Irani. "Blind super-resolution kernel estimation using an internal-gan." Advances in Neural Information Processing Systems 32 (2019).
- [16] Wang, Longguang, et al. "Unsupervised degradation representation learning for blind super-resolution. " Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.
- [17] Wang, Xintao, et al. "Esrgan: Enhanced super-resolution generative adversarial networks." Proceedings of the European conference on computer vision (ECCV) workshops. 2018.
- [18] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015.