

깊은 신경망 기반 영상 화질 개선 모델을 활용한 저품질 영상에서의 객체 검출 성능 향상

김 대 희*, 유 영 준*, 노 송**, 이 재 구°

Enhancing Object Detection in Low Quality Images Using Deep Neural Network Based Super-Resolution

Daehee Kim*, Youngjun Yoo*, Song Noh**, Jaekoo Lee°

요 약

객체 검출 기술은 입력 영상의 낮은 화질 및 해상도 등의 한계에 부딪혀서, 현실에서는 성능이 저하될 수 있다. 본 논문은 현실에서의 객체 검출 기술 성능을 실질적으로 향상시키고자, 깊은 신경망 기반 초해상도(Super-Resolution) 모델을 활용하는 방법을 제안한다. 우리는 실험으로 얼굴 검출과 일반적 객체 검출, 두 조건 모두에 대해 검증하였으며, 현실의 영상 환경을 시뮬레이션하기 위해 복합적인 영상 열화 방법을 사용하였다. 실험 과정으로, 먼저 전통적인 영상 크기 변환 방법인 쌍입방 보간법(Bicubic Interpolation)과 가우시안 노이즈, 가우시안 블러 등을 복합적으로 활용하여 영상을 열화 및 다운샘플링(Down-sampling)한 후, 쌍입방 보간법과 깊은 신경망 기반 최신 초해상도 모델인 EDSR(Enhanced Deep Residual Networks for Single Image Super-Resolution)을 사용하여 각각 영상을 복원한다. 이렇게 생성한 영상을 얼굴 검출용 깊은 신경망 모델로 RetinaFace, 일반적 객체 검출용 깊은 신경망 모델로는 EfficientDet-D7을 활용하여 평균 정밀도(Average Precision)를 비교 및 분석하였다. 이러한 과정을 Widerface와 COCO 데이터 집합을 기반으로 정량적 실험을 수행하였고, 여러 열화 방법 조건에서 객체 크기가 작은 경우, 얼굴 검출은 최대 54.8%, 일반적 객체 검출은 370.0%만큼 성능이 향상되는 것을 검증하였다. 따라서 깊은 신경망 기반 초해상도 모델을 활용하여, 현실에서 객체 검출 기술의 실질적 성능을 향상시킬 수 있을 것으로 판단된다.

Key Words : Super-Resolution, Deep Neural Network, Object Detection, Face Detection, Enhancing Performance

ABSTRACT

Obtaining potential gains from object detection in practical systems hinges on sufficient levels of image resolution. In this paper, a two-step guideline using deep neural network (DNN)-based super-resolution (SR) model for object detection is presented under the assumption that the DNN model takes multiple low-resolution images captured from the same scene. In the first step, enhanced deep residual networks for single image super-resolution (EDSR) is exploited to recover an intermediate high-resolution image. The second step is to

※ 본 연구는 2020년도 미래창조과학부의 재원으로 한국연구재단의 지원(No. NRF2018R1C1B5086441)과 과학기술정보통신부 및 정보통신산업진흥원의 '고성능 컴퓨팅 지원' 사업으로부터 지원을 받아 수행되었습니다.

• First Author : College of Computer Science, Kookmin University, frshield2@kookmin.ac.kr, 학생회원

° Corresponding Author : College of Computer Science, Kookmin University, jaekoo@kookmin.ac.kr, 정회원

* College of Computer Science, Kookmin University, junyoo96@kookmin.ac.kr, 학생회원

** Department of Information and Telecommunication Engineering, Incheon National University, songnoh@inu.ac.kr, 정회원

논문번호 : 202007-178-A-RE, Received July 31, 2020; Revised September 4, 2020; Accepted September 16, 2020

perform face and object detection based on RetinaFace and EfficientDet-D7. By capturing different training and test image formats, the resulting design employs a transfer learning method with pre-trained EDSR to further leverage detection performance. We adopt three degradation models for performance analysis and provide a practical guideline for DNN-based SR reconstruction. Numerical results show the effectiveness of the proposed method in improving small-sized object detection performance.

1. 서론

현재 깊은 신경망을 활용한 객체 검출 기술(Object Detection)은 비약적인 발전을 거듭하여 최근에는 YOLO(You Only Look Once)^[1]와 같이 뛰어난 실시간 성능을 보였다. 우리는 YOLO와 같은 최신 객체 검출 기술을 활용하여 영상에서 얼굴을 검출하고 비식별화 하는 방법을 기존 연구^[2]에서 제안했으나, 일반적인 현실에서 객체 검출 기술의 성능을 내기 힘든 문제를 인지하였다. 문제의 주요한 원인은 입력되는 영상의 화질이었으며, 그에 따라 영상 화질 개선의 필요성을 확인하였다.

본 논문에서는 현실에서 객체 검출 기술의 얼굴 검출 성능을 향상시키고자, 입력 영상에 깊은 신경망 기반 초해상도(Super-Resolution) 모델을 활용하는 방법을 제안한다.

깊은 신경망 기반 초해상도 모델은 대표적으로 SRCNN(Image Super-Resolution Using Deep Convolutional Networks)^[3], SRGAN(Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network)^[4], EDSR(Enhanced Deep Residual Networks for Single Image Super-Resolution)^[5] 등이 있다. 이 중에서 EDSR은 국제 초해상도 모델 경쟁 대회인 Ntire 2017 Challenge^[11]에서 우승을 차지하였으며, 기존의 SRResNet^[3]에서 배치 정규화층을 제거하여 연산량을

감량하고, 감량한 연산량으로 모델의 층과 채널의 개수를 최대한 증가시켜 성능을 향상시킨 모델이다. EDSR은 최근에 발표된 DBPN^[13], RDN^[12] 등과 PSNR(Peak Signal-to-Noise Ratio)과 SSIM(Structural Similarity) 지표로 비교해보아도 성능을 비견할 수 있으므로, 최근의 초해상도 모델과 기존의 SRCNN, SRGAN과 같은 모델을 대표할 수 있다고 판단하여 EDSR을 실험에 사용하였다.

본 논문에서 실험은 얼굴 특화 검출과 일반적인 객체 검출에 대해 모두 진행되었다. 얼굴 검출 실험에서는 WIDER CHALLENGE 2019: Face Detection^[6]에서 평균 정밀도(Average Precision, AP) 기준으로 최고 수준의 성능을 보인 RetinaFace^[7] 모델을 사용하였고, 일반적인 객체 검출 실험에서는 널리 사용되는 데이터 집합인 COCO(Common Objects in Context)^[9]에 대해 최고 수준의 성능을 보인 EfficientDet-D7^[10]을 사용하였다. 또한, 현실의 저화질 영상 환경을 시뮬레이션하기 위해, 최근 초해상도 과업에 널리 사용되는 BI, BD, DN 영상 열화 방법^[12]을 통해 저화질 영상을 생성하여 실험하였다. 열화 방법은 본문 3절에서 상세히 설명하였다.

우리는 [그림 1]에 제시된 실험 개요에 따라 깊은 신경망을 활용하여 영상의 화질을 개선함으로써, 객체 검출의 정확도를 향상시킬 수 있는지 실험하였다. Widerface^[8]와 COCO 데이터 집합에 대한 실험 결과, RetinaFace와 EfficientDet-D7은 각각 작은 객체 검출에 대한 성능이 쌍입방 보간법 대비 최고 54.8%,

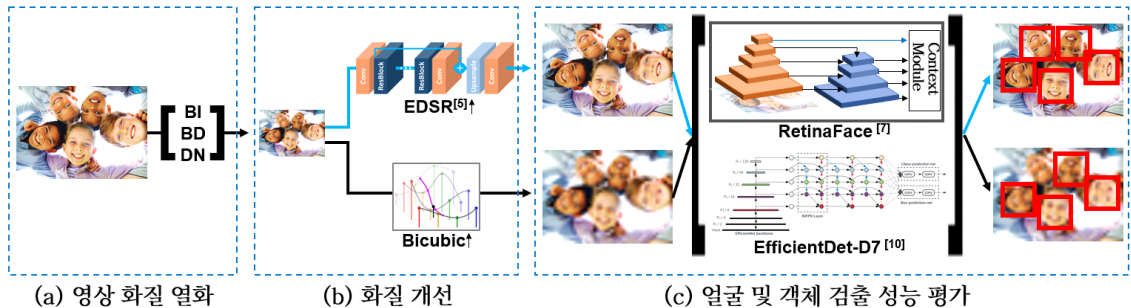


그림 1. 실험 개요
Fig. 1. Experiment overview

370.0%의 평균 정밀도(Average Precision)가 향상됨을 보였다. 따라서 깊은 신경망 기반 초해상도 모델을 활용하여, 현실에서 객체 검출 기술의 실질적 성능을 향상시킬 수 있을 것으로 판단된다.

II. 현실 영상 환경 시뮬레이션

2.1 실험 설계

본 논문에서는 현실의 얼굴 검출과 일반적 객체 검출, 두 목적에 대해 영상 화질 개선의 영향을 실험하기 위하여, 각 목적에 널리 쓰이는 데이터 집합인 얼굴 검출용 Widerface와 일반적 객체 검출용 COCO를 사용하였다. Widerface는 다양한 장면에서 얼굴을 객체로 구성한 데이터 집합이며, COCO는 사람, 자전거, 차량, 비행기 등의 일상적인 장면에서의 객체로 구성된 데이터 집합이다. 또한 객체 검출의 성능을 최대화하여 실험하기 위해, 각 목적에 대해 현재 최고 수준의 깊은 신경망 기반 검출 알고리즘인 RetinaFace, EfficientDet-D7를 사용하였다.

[그림 1]은 전체적인 실험 과정을 묘사하였다. (a) 단계에서 전통적인 영상 크기 변환 방법인 쌍입방 보간법(Bicubic Interpolation)과 가우시안 블러(Gaussian Blur), 가우시안 노이즈(Gaussian Noise) 등을 복합적으로 이용하여 데이터 집합의 영상을 0.25배로 다운샘플링(Downsampling)한다. 이는 영상에 대해 저품질 영상에 초해상화를 적용하면서도 크기를 원본과 맞추기 위함이다. (b)단계에서는 (a)단계에서 다운샘플링한 영상을 입력으로 하여, EDSR과 쌍입방 보간법을 통해 원본과 같은 크기의 영상을 각각 생성한다. (c)단계에서는 생성한 영상 데이터 집합에 대해 RetinaFace와 EfficientDet-D7를 각각 사용하여 객체

를 검출하고, 평균 정밀도(Average Precision)를 계산한다.

2.1.1 전이 학습 (Transfer Learning)

사전 학습된 EDSR은 PNG(Portable Network Graphics) 파일 형식의 영상으로 구성된, 영상 복원 목적을 위한 고품질 영상 데이터 집합인 DIV2K^[11] 기반으로 학습되었다. 반면, 실험에 사용한 데이터 집합 모두 JPEG(Joint Photographic Experts Group) 파일 형식의 영상으로 구성되어 있다. JPEG 형식은 손실 압축 형식이므로 비손실 압축 형식인 PNG 형식과 영상의 분포가 다르다. 이를 신경망 관점에서 보면, 학습된 데이터와 테스트 데이터의 도메인이 다르다고 할 수 있다. [그림 4]에서 볼 수 있듯이, 이는 사전 학습된 EDSR의 성능을 현저히 떨어뜨리는 결과를 보였다. 우리는 이를 해결하기 위해 사전 학습된 EDSR의 매개변수를 기반으로, JPEG 형식의 영상 데이터를 전이 학습 하였다.

얼굴 검출 실험의 경우, Widerface 학습 집합에서 900개의 영상을 샘플링(Sampling) 하여 EDSR을 전이 학습 하였다. 일반적인 객체 검출 실험의 경우는, COCO 데이터 집합의 영상을 0.25배 다운샘플링과 4배 초해상화를 진행했을 때 해상도가 일치하지 않으므로, DIV2K 데이터 집합을 JPEG 형식으로 변환하여 전이 학습 하였다.

2.2 복합적 영상 열화 방법

최근 초해상도 과업에서는 쌍입방 보간법을 통한 축소와 가우시안 블러, 가우시안 노이즈를 복합적으로 활용하여 영상을 열화한다.^[12] 쌍입방 보간법을 통한 축소 영상으로 현실의 저화질 영상을 가정하기에는,



그림 2. Widerface[8]의 0_Parade_marchingband_1_355.jpg 영상을 DN 방법으로 열화하고, EDSR로 화질 개선한 예. 왼쪽부터 원본, DN 저화질 영상을 쌍입방 보간법으로 4배 확대, EDSR 출력
Fig. 2. 0_Parade_marchingband_1_355.jpg image in Widerface[8] that have been degraded by DN method and super-resolved with EDSR. From left to right, the original, DN image interpolated by bicubic upsampling, and EDSR output.

현실의 저화질 영상 분포가 훨씬 복잡하기 때문이다.

BI(Bicubic Downsampling)는 쌍입방 보간법(Bicubic Interpolation)을 통해 영상을 다운샘플링하는 방법이며, **BD**(Blur and Downsampling)는 표준편차(Standard Deviation)이 1.6인 7×7 크기의 가우시안 커널을 원본 영상에 사용하여 영상을 블러링(Blurring) 한 후, 쌍입방 보간법을 통해 다운샘플링하는 방법이다. **DN**(Downsampling and Noise)은 먼저 쌍입방 보간법으로 원본 영상을 다운샘플링한 후, 노이즈 레벨 30의 가우시안 노이즈를 더하는 방법이며, [그림 2]에서 생성된 영상의 형태를 확인할 수 있다.^[12] 본 논문에서는 현실 환경과 유사하도록 BI, BD, DN 열화 방법을 통해 영상을 0.25배로 축소하여 실험을 진행하였다.

III. 실험

3.1 얼굴 검출 성능 향상 실험

얼굴 검출 실험 결과는 [그림 3]에 각각 BI, BD,

DN 열화 방법에 대한 각 초해상도 모델의 성능 향상 결과를 그래프로 나타냈다. 우리는 EDSR을 전이 학습한 결과를 [그림 3] 상단의 범주에 RetinaFace - EDSRX4(Widerface)로 표기하였다. 또한 원본 데이터 집합을 사용한 결과를 RetinaFace로, 사전 학습된 EDSR로 생성한 데이터 집합을 사용한 결과를 RetinaFace-EDSRX4로, 쌍입방 보간법으로 생성한 데이터 집합을 사용한 결과를 RetinaFace-BicubicX4로 표기하였다. 사전 학습된 EDSR은 BI 열화 방법으로 학습되었으므로, BD, DN은 사전학습된 EDSR이 존재하지 않는다. [그림 3]에서 (a)는 300픽셀 이상, (b)는 50~300픽셀, (c)는 10~50픽셀 크기에 해당하는 얼굴이 포함된 영상 데이터 집합이다.

[그림 3]의 (a)와 (b)에서는 쌍입방 보간법에 비해 사전 학습된 EDSR 모델을 사용하였을 때, RetinaFace의 성능이 BI의 경우 0.4%, 1.3%로 미미하게 향상된다. 그러나 (c)에서는 8.6%까지 증가하며, 전이 학습한 EDSR을 사용하면 10.3%까지 성능이 향상된다. BD의 경우는 EDSR이 (a)와 (b)에서 1.3%,

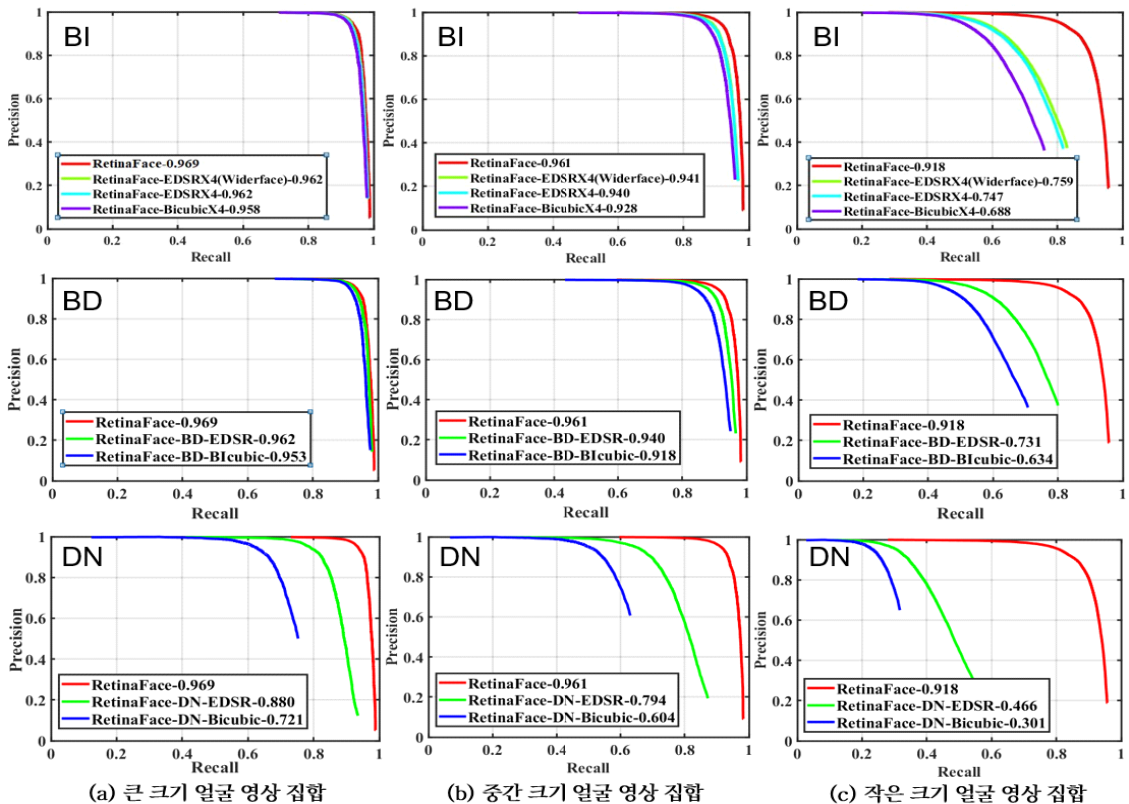


그림 3. BI, BD, DN 열화 방법의 각 영상 화질 개선에 따른, Widerface[8] 데이터 집합에 대한 RetinaFace[7]의 얼굴 검출 평균 정밀도(AP) 변화.

Fig. 3. Face detection AP variation of RetinaFace[7] for the Widerface[8], on BI, BD, DN degradation methods

표 1. Widerface[8] 각 얼굴 크기에 대한 EDSR[5]-JPEG의 객체 검출 성능 향상률 (%)
Table 1. Face detection improvement rate of EDSR-JPEG on each face scale in Widerface (%)

Face Scale	Small	Medium	Large
BI	8.6	1.3	0.4
BD	15.3	2.4	1.3
DN	54.8	31.5	22.1
Average	26.2	11.7	7.9

2.4%로 BI보다 약간 더 높은 수치를 보였고, (c)에서는 15.3%로 더욱 상승하였다. DN은 (a)와 (b)에서 22.1%, 31.5%, (c)에서는 54.8%로 전체적으로 매우 높은 성능 향상률을 보였다.

[표 1]은 [그림 3]의 각 열화 방법에서, 각 얼굴 크기에 대한 EDSR의 쌍입방 보간법에 비한 성능 향상률을 표시하였다. [표 1]을 보면, BI, BD 열화 방법에서 크거나 중간 크기의 얼굴 영상은 쌍입방 보간법으로 확대하더라도 RetinaFace 모델이 얼굴을 검출할 수 있도록 원본의 얼굴 특징을 충분히 복원하는 것을 확인할 수 있다. 하지만 (c)의 영상은 얼굴 크기가 작아 확대된 영상이 과하게 부드러워져 RetinaFace가 얼굴을 검출할 정도로 충분한 특징을 복원하지 못하는 것으로 볼 수 있다. 또한, [표 1]을 확인하면 DN, BD, BI 순서로 열화 방법이 복잡할수록 더 성능 향상률이 높은 결과를 보였다. DN 조건에서는 [그림 2]에서 확인할 수 있듯이 EDSR이 노이즈를 제거하는 효과도 갖고 있으므로, 성능 향상률이 매우 큰 영향을

준 것으로 판단된다. 더불어 최근 다수의 초해상도 모델들이 DN 조건에서 학습할 경우 잡음 제거의 효과를 가지므로, 비슷한 양상의 성능 향상 결과가 도출될 것으로 예상된다.

3.2 일반적 객체 검출 성능 향상 실험

[표 2]는 EfficientDet-D7을 사용하여 COCO 데이터 집합에 대해 각 초해상도 방법의 평균 정밀도와 평균 재현율(Average Recall)을 측정된 결과이다. [표 2]의 BI 열화 방법에서는 사전학습된 EDSR(EDSR-Pretrain)을 사용한 실험 결과가 쌍입방 보간법을 사용한 결과보다도 성능이 낮게 측정되었다. 이는 [그림 4]에서 확인할 수 있듯이, 데이터 도메인의 차이로 인해 영상의 질감을 제대로 복원하지 못한 결과가 준 영향으로 볼 수 있다. 그렇다고 하더라도, 쌍입방 보간법에 비해 사전 학습된 EDSR이 오히려 성능 저하된 결과를 보인 것은 RetinaFace의 결과와 다른 양상인데, RetinaFace는 얼굴의 위치를 파악하는 알고리즘에 자기 지도 학습(Self-supervised Learning) 등을 활용하여 얼굴만을 더 잘 검출할 수 있도록 특화하였다.^[7] 따라서 일반적인 객체 검출을 목적으로 하는 EfficientDet-D7은, 데이터의 도메인이 달라 EDSR이 제대로 된 결과를 내지 못한 것에 영향을 더 크게 받았고, 그에 따라 성능이 저하된 결과를 보인 것으로 예상된다.

[표 2]에서도 BI 열화 방법에서 전이 학습한 EDSR(EDSR-JPEG)로 생성한 COCO를 사용했을 때,

표 2. BI, BD, DN 열화 영상을 영상 화질 개선한 COCO[9] 데이터 집합과 원본 영상에 대한 EfficientDet-D7[10]의 객체 검출 평균 정밀도(AP)와 평균 재현율(Average Recall, AR) 변화
Table 2. Object detection AP, AR variation of RetinaFace for super-resolved images of degraded COCO validation set

BI Degradation	Super-Resolution Method			
	Bicubic	EDSR[5]-Pretrain	EDSR[5]-JPEG	HR
AP [IoU=0.50:0.95 area= all maxDets=100]	0.312	0.267	0.357 (14.4%)	0.526
AP [IoU=0.50 area= all maxDets=100]	0.464	0.403	0.526 (13.4%)	0.716
AP [IoU=0.75 area= all maxDets=100]	0.327	0.278	0.374 (14.4%)	0.569
AP [IoU=0.50:0.95 area= small maxDets=100]	0.104	0.070	0.125 (20.2%)	0.353
AP [IoU=0.50:0.95 area=medium maxDets=100]	0.328	0.271	0.387 (18.0%)	0.559
AP [IoU=0.50:0.95 area= large maxDets=100]	0.532	0.477	0.578 (9.6%)	0.650
AR [IoU=0.50:0.95 area= all maxDets= 1]	0.283	0.259	0.309 (9.2%)	0.393
AR [IoU=0.50:0.95 area= all maxDets= 10]	0.440	0.400	0.477 (8.4%)	0.639
AR [IoU=0.50:0.95 area= all maxDets=100]	0.466	0.424	0.505 (8.4%)	0.677
AR [IoU=0.50:0.95 area= small maxDets=100]	0.195	0.150	0.226 (15.9%)	0.510
AR [IoU=0.50:0.95 area=medium maxDets=100]	0.511	0.454	0.564 (10.4%)	0.711
AR [IoU=0.50:0.95 area= large maxDets=100]	0.717	0.691	0.753 (5.0%)	0.812

BD Degradation	Bicubic	EDSR-JPEG	HR
AP [IoU=0.50:0.95 area= all maxDets=100]	0.376	0.399 (6.1%)	0.526
AP [IoU=0.50 area= all maxDets=100]	0.547	0.577 (5.4%)	0.716
AP [IoU=0.75 area= all maxDets=100]	0.397	0.423 (6.5%)	0.569
AP [IoU=0.50:0.95 area= small maxDets=100]	0.144	0.164 (13.9%)	0.353
AP [IoU=0.50:0.95 area=medium maxDets=100]	0.413	0.439 (6.3%)	0.559
AP [IoU=0.50:0.95 area= large maxDets=100]	0.596	0.617 (3.5%)	0.650
AR [IoU=0.50:0.95 area= all maxDets= 1]	0.318	0.334 (5.0%)	0.393
AR [IoU=0.50:0.95 area= all maxDets= 10]	0.498	0.516 (3.6%)	0.639
AR [IoU=0.50:0.95 area= all maxDets=100]	0.525	0.546 (4.0%)	0.677
AR [IoU=0.50:0.95 area= small maxDets=100]	0.248	0.283 (14.1%)	0.510
AR [IoU=0.50:0.95 area=medium maxDets=100]	0.587	0.606 (3.2%)	0.711
AR [IoU=0.50:0.95 area= large maxDets=100]	0.767	0.777 (1.3%)	0.812
DN Degradation	Bicubic	EDSR-JPEG	HR
AP [IoU=0.50:0.95 area= all maxDets=100]	0.074	0.223 (201.4%)	0.526
AP [IoU=0.50 area= all maxDets=100]	0.119	0.347 (191.6%)	0.716
AP [IoU=0.75 area= all maxDets=100]	0.075	0.226 (201.3%)	0.569
AP [IoU=0.50:0.95 area= small maxDets=100]	0.010	0.047 (370.0%)	0.353
AP [IoU=0.50:0.95 area=medium maxDets=100]	0.075	0.214 (185.3%)	0.559
AP [IoU=0.50:0.95 area= large maxDets=100]	0.146	0.433 (196.6%)	0.650
AR [IoU=0.50:0.95 area= all maxDets= 1]	0.120	0.229 (90.8%)	0.393
AR [IoU=0.50:0.95 area= all maxDets= 10]	0.186	0.349 (87.6%)	0.639
AR [IoU=0.50:0.95 area= all maxDets=100]	0.194	0.369 (90.2%)	0.677
AR [IoU=0.50:0.95 area= small maxDets=100]	0.026	0.099 (280.8%)	0.510
AR [IoU=0.50:0.95 area=medium maxDets=100]	0.171	0.391 (128.7%)	0.711
AR [IoU=0.50:0.95 area= large maxDets=100]	0.388	0.665 (70.9%)	0.812

쌍입방 보간법 대비 객체 검출 성능이 14.4% 향상되었으며, 특히 작은 객체에 대해 20.2% 향상되어 중간 크기 이상 객체에 비해 성능 향상이 두드러졌다. BD의 경우는 모든 객체에 6.1%, 작은 객체에 13.9%가 향상되어 비교적 약간 낮은 향상률을 보였고, DN은 모든 객체에 201.4%, 작은 객체에 370.0%의 비교적 매우 높은 성능 향상률을 보였다. 이는 앞서 [그림 2]에서 보았던 것처럼 EDSR 초해상도 모델의 잡음 제거 효과가 미친 영향으로 보인다.

실험을 종합한 결과, 객체 크기가 작은 영상에서 깊은 신경망 기반 초해상도 모델 EDSR을 사용하면, 쌍입방 보간법을 사용했을 때보다, 얼굴 검출 성능은 최대 54.8%, 일반적 객체 검출 성능은 최대 370.0%만큼 향상되었다. 또한 [표 2]와 [그림 3]의 (c)에서 사전 학습된 EDSR과 전이 학습한 EDSR의 성능 수치 차이에서 확인할 수 있듯이, 데이터의 도메인이 적합하도록 모델을 전이 학습 하면 객체 검출의 성능을 더욱

향상시킬 수 있다.

IV. 결 론

본 논문에서는 영상의 화질이 충분히 좋지 않으면 객체 검출 기술의 성능이 저하되는 문제를 극복하기 위해, 깊은 신경망 기반 초해상도 모델을 활용하는 방법을 제안하였다. 그리고 복잡한 영상 열화 방법으로 현실의 저화질 영상 환경을 시뮬레이션하여, 얼굴 검출과 일반적 객체 검출 두 가지 경우에 대한 실험을 통해 타당성을 검증하였다. 실험 결과로, 객체 크기가 작은 영상에서 깊은 신경망 기반 초해상도 모델 EDSR을 사용한다면, 쌍입방 보간법을 사용하는 경우보다 RetinaFace의 성능은 최대 54.8%, EfficientDet-D7은 최대 370.0%만큼 향상됨을 정량적으로 산출하였다. 결과적으로 우리는 영상의 화질이 좋지 않을 때 깊은 신경망을 활용하여 화질을 개선함

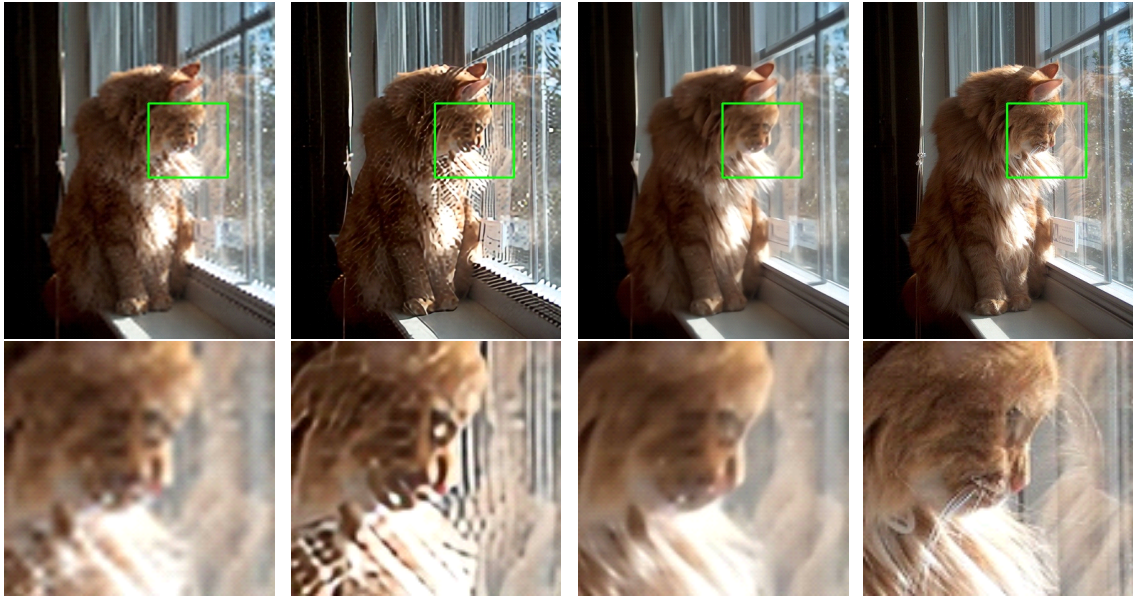


그림 4. COCO[9]데이터 집합의 000000574810.jpg 영상을 쌍입방 보간법으로 0.25배 다운샘플링 한 후 각 방법으로 초해상화한 결과. 왼쪽부터 쌍입방 보간법, 사전 학습된 EDSR[5], JPEG 형식으로 변환한 DIV2K[11]로 전이 학습된 EDSR, 원본.

Fig. 4. 000000574810.jpg image in COCO $0.25\times$ bicubic downsampled and super-resolved with EDSR. From left to right, bicubic upsampling, EDSR pretrained on DIV2K, and EDSR-JPEG trained on DIV2K-JPEG, and the original

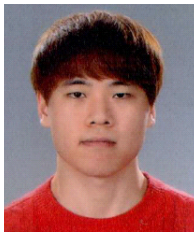
으로써, 객체 검출 기술의 실질적인 성능을 향상시킬 수 있음을 확인하였다.

References

- [1] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [2] S. Heo, D. Kim, Y. Kim, and J. Lee, "Real-time face de-identification in visual media using a deep neural network for object detection," in *Proc. KICS Summer Conf.*, vol. 69, no. 1, pp. 609-610, Jeju Island, Korea, Jun. 2019.
- [3] C. Dong, et al., "Learning a deep convolutional network for image super-resolution," in *Eur. Conf. Computer Vision*, Zurich, Switzerland, Sep. 2014.
- [4] C. Ledig, et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 4681-4690, Hawaii, USA, Jul. 2017.
- [5] B. Lim, et al., "Enhanced deep residual networks for single image super-resolution" in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Wkshps.*, pp. 136-144, Hawaii, USA, Jul. 2017.
- [6] Y. Liu, et al., *WIDER Face and Person Challenge 2019*(2019), Retrieved Jul. 28, 2020, <https://wider-challenge.org/2019.html>
- [7] J. Deng, et al., "RetinaFace: Single-shot multi-level face localisation in the wild," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 5203-5212, Jun. 2020.
- [8] S. Yang, et al., "Wider face: A face detection benchmark," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 5525-5533, Nevada, USA, Jun. 2016.
- [9] T.-Y. Lin, et al., "Microsoft coco: Common objects in context," in *Eur. Conf. Computer Vision*, pp. 740-755, Zurich, Switzerland, Sep. 2014.
- [10] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, pp. 10781-10790, Jun. 2020.

- [11] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Wkshps.*, pp. 126-135, Hawaii, USA, Jul. 2017.
- [12] Y. Zhang, et al., "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 2472-2481, Utah, USA, Jun. 2018.
- [13] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1664-1673, Utah, USA, Jun. 2018.

김 대 희 (Daehee Kim)*



2020년 2월 : 국민대학교 소프트웨어학부 졸업
 2020년 3월~현재 : 국민대학교 컴퓨터공학과 석사
 <관심분야> 인공지능, 초해상도(Super-Resolution)
 [ORCID:0000-0001-9676-9604]

유 영 준 (Youngjun Yoo)



2020년 2월 : 국민대학교 소프트웨어학부 졸업
 2020년 3월~현재 : 국민대학교 컴퓨터공학과 석사
 <관심분야> 인공지능, 초해상도(Super-Resolution)
 [ORCID:0000-0001-9836-8732]

노 송 (Song Noh)



2008년 2월 : 숭실대학교 정보통신전자공학부 (공학사)
 2010년 2월 : 한국과학기술원 전기 및 전자공학부 (공학석사)
 2015년 12월 : 미국 Purdue University, Electrical and Computer Engineering (공학박사)

2015년 12월~2018년 7월 : 미국 Intel Corp. 시스템 엔지니어
 2018년 9월~현재 : 인천대학교 정보통신공학과 조교수
 <관심분야> 신호처리, 스마트 시스템 설계
 [ORCID:0000-0002-0106-7106]

이 재 구 (Jaekoo Lee)*



2011년~2013년 : LG전자 CTO 부문 (주임연구원)
 2018년 : 서울대학교 전기컴퓨터공학부 박사
 2018년 : SK 텔레콤 ICT기술원 매니저
 2018년~현재 : 국민대학교 소프트웨어학부 조교수

<관심분야> 인공지능, 기계학습, Data Science
 [ORCID:0000-0002-5947-5487]