

SRPAttention-GAN 기반 고해상도 철근 이미지 생성을 통한 철근 끝점 검출 성능 향상

박종찬*, 허지욱**, 김건우***

Enhancing Rebar Endpoint Detection Performance through High-Resolution Rebar Image Generation with SRPAttention-GAN

Jong-Chan Park*, Jee-Uk Heu**, and Gun-Woo Kim***

본 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1G1A1006381)

요약

현재 철근공장의 가공 공정에서 발생하는 불량품 생산을 최소화하기 위한 무인 자동화 연구가 진행되고 있다. 선행 연구에서는 YOLOv3 기반의 비전 카메라를 이용한 철근 이미지의 끝점을 검출하고 철근의 꼬임 유무를 파악하려 하였으나, 철근 가공 현장에서의 노이즈로 인해 이미지 품질이 저하되어 딥러닝 모델의 검출 성능이 감소하였다. 이에 본 논문에서는 이런 문제를 해결하기 위해 새로운 모델 SRPAttention-GAN을 제안한다. SRPAttention-GAN 모델은 저해상도의 철근 이미지를 고해상도 이미지로 변환하여 YOLOv5 모델에 학습시켜 검출 성능을 높인다. 실험 결과 SRPAttention-GAN 모델은 기존의 SRResNet-GAN 모델에 비해 PSNR이 약 0.3dB 더 높아졌음을 확인하였다. 또한, 원본 이미지와 제안 모델에서 생성한 이미지를 YOLOv5에 학습시켜 F1과 정밀도에서 더욱 개선된 결과를 얻을 수 있었으며, 이를 통해 철근의 끝점 검출 정확도를 향상시켰다.

Abstract

Many studies are currently underway on unmanned automation to reduce the production of defective products during the machining process in rebar factories. In previous studies, a vision camera using YOLOv3 was employed to detect the endpoint of rebar images and determine whether the rebar was twisted. However, the quality of the images deteriorated due to noise occurring at the rebar processing site, resulting in decreased detection performance by the deep learning model. In this paper, we propose a novel model called SRPAttention-GAN. This model transforms low-quality rebar images into high-resolution images and trains the YOLOv5 model to enhance detection performance. Experimental results confirmed that the PSNR of the SRPAttention-GAN model was approximately 0.3db higher than the existing SRResNet-GAN model. Furthermore, by training YOLOv5 on both the original images and the images generated by the proposed model, we were able to obtain improved results on the F1 and Precision, thereby enhancing the accuracy of rebar endpoint detection.

Keywords

rebar endpoint, srgan, cbam, high-resolution image generation, yolov5

* 경상국립대학교 AI융합공학과 박사과정
- ORCID: <https://orcid.org/0000-0002-1886-534X>
** 강남대학교 참인재대학 교양교수부 조교수
- ORCID <https://orcid.org/0000-0002-1666-1626>
*** 경상국립대학교 컴퓨터과학과 조교수(교신저자)
- ORCID: <https://orcid.org/0000-0001-5643-4797>

· Received: Aug. 09, 2023, Revised: Sep. 11, 2023, Accepted: Sep. 14, 2023
· Corresponding Author: Gun-Woo Kim
Dept. of Computer Science, College of Natural Science, Gyeongsang
National University, Jinju, Korea
Tel.: +82-55-772-3323, Email: gunwoo.kim@gnu.ac.kr

1. 서 론

인공지능과 로봇 기술의 발전에 따라 제조업의 자동화가 가속화되고 있다. 특히, 건설 현장에서는 건물의 토대를 콘크리트로 다질 때 널리 사용되는 철근을 다양한 형태로 가공하고, 이를 인공지능 기술과 로봇을 활용해 실시간으로 자동화하여 생산하는 연구의 필요성이 높아지고 있다. 기존 철근 가공 시스템은 그림 1과 같이 특정 모양의 가공 철근을 절단 전용 설비에서 생성한 후, 사람이 직접 수동으로 포장하고 적재하여 현장에 납품하였다. 이러한 일련의 수동 과정을 통해 철근의 불량률 줄이기 위한 교정 작업의 시간과 정확도는 작업자의 숙련도에 크게 의존하고 있다[1].

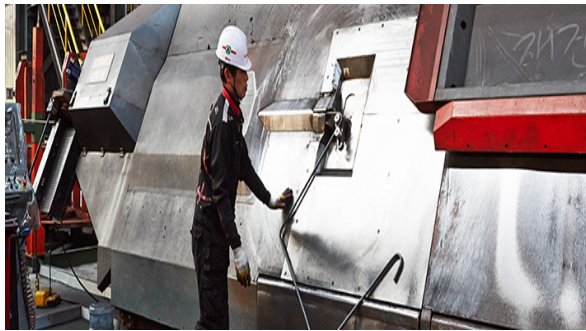


그림 1. 철근 가공 현장
Fig. 1. Rebar processing site

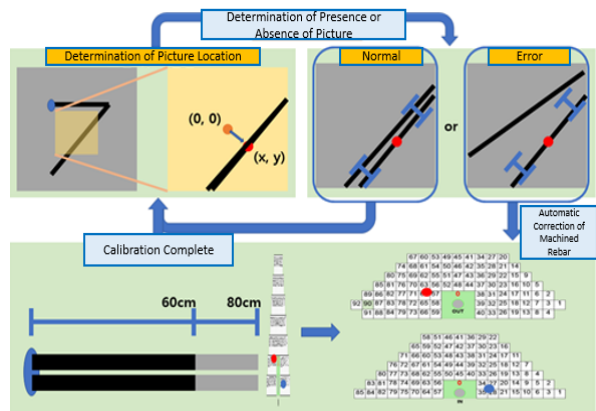
그러나 그림 1과 같이 현재의 수동적인 철근 가공 생산 공정은 품질과 안전 문제를 수반하고 있다. 이에 대한 해결책으로, 이전 연구에서는 가공 철근 공장의 핵심 설비인 절단, 절곡 및 적재 자동화 설비에 대한 실시간 영상분석을 통한 교정 자동화와 로봇 그립 최적화를 통한 가공 손실 감소 및 생산성 향상 등, 손실률을 최소화하는 자동화된 스마트 생산 시스템을 도입하였다[2].

기존 연구에서는, 그림 2와 같이 철근이 사출되는 지점에 머신비전 카메라를 설치하여 철근의 끝점을 검출하였다. 이 끝점의 위치 좌표 값을 비선형 회귀 모델에 학습시키고, 이를 통해 철근의 끝점 지점의 롤러 교정 값을 예측하여 사출 기계에 전송하였다. 이를 통해 실시간으로 철근을 자동 교정하였다. 또한 철근의 형상을 인식하여 그립 가능성 여부를 판단하였고, 이를 기반으로 그립 가능한 위치의

좌표 값을 도출하는 연구를 수행하였다[2][3].



(a) 머신비전 카메라 설치 환경
(a) Machine vision camera installation environment



(b) 실시간 교정 자동화 및 로봇 그립 최적화
(b) Real-time calibration automation and robot grip optimization

그림 2. 철근 끝점 검출을 위한 머신비전 카메라 환경 및 자동 교정을 위한 그립 최적화

Fig. 2. Machine vision camera environment for rebar end point detection and grip optimization for automatic calibration

실제 가공 공정 환경에서는 영상 품질과 조명 조건이 영상 분석의 정확성에 큰 영향을 미친다. 특히 조명이 부족하거나 불규칙한 조명 조건에서는 영상의 선명도와 명암 대비가 저하될 수 있으며, 이로 인해 객체 검출 및 특징 추출에 어려움을 겪을 수 있다. 실제로 철근 사출 동영상 및 이미지를 수집하는 과정에서 머신비전 카메라의 초점 거리, 렌즈, 명암 밝기 등을 조절한다고 하더라도 실시간 영상 분석이나 수집 중에 카메라가 흔들리거나 외부 환경에 의해 조명이 변경되면 노이즈가 발생한다. 이러한 노이즈가 낀 저해상도 이미지를 데이터 증강을 활용하더라도 분석에 큰 영향을 준다[4].

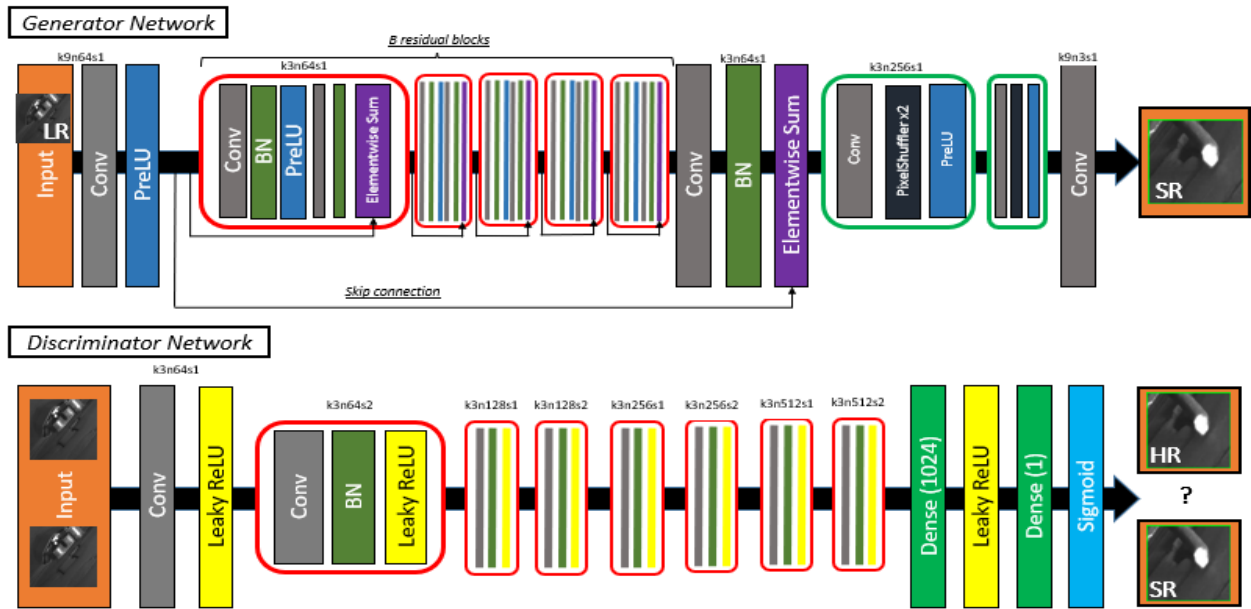


그림 3. SR-GAN 구조[5]
Fig. 3. Structure of SR-GAN[5]

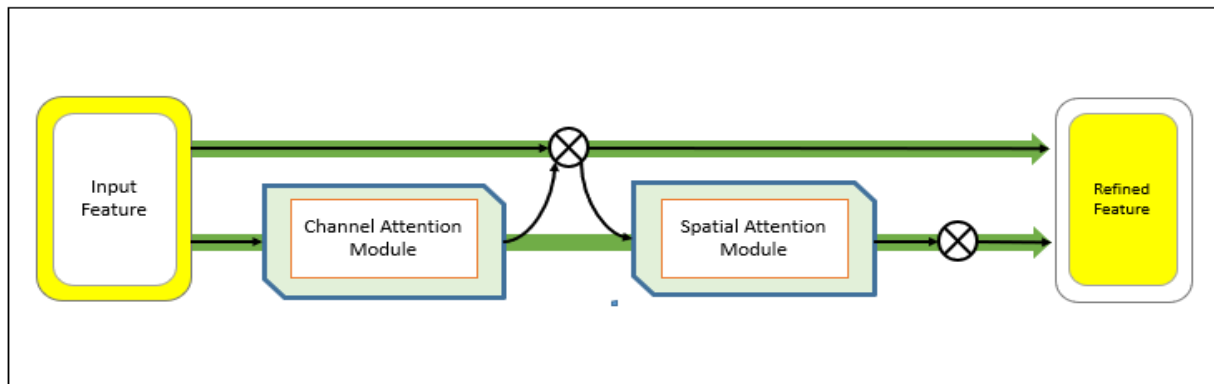


그림 4. CBAM의 attention 모듈 구조[6]
Fig. 4. Structure of CBAM's attention module[6]

일반적으로 저해상도 이미지를 고해상도 이미지로 생성하거나 복원하기 위해 SR-GAN (Super-Resolution Generative Adversarial Network)[5]가 널리 사용되고 있다. 이 모델의 경우 의료 이미지, 위성 및 항공 이미지, 보안 카메라, 재난 관리, 엔터테인먼트 산업 등에서 주로 활용되고 있다.

그러나, SR-GAN을 통해 저해상도 이미지에서 고해상도 이미지로의 변환 과정에서는, 모든 정보가 딥러닝 검출 모델이나 분류 모델에 학습 시 큰 영향을 주지 않는다는 한계점이 있다. 따라서, SR-GAN에 이미지의 특정 영역에 대한 더 세밀한 처리를 하는 기능을 추가하여, 해당 영역의 구조나

질감 등을 더 잘 보존하거나 강조한다면 이러한 세부적인 특징들로 인해 검출 모델의 성능 향상에 기여할 수 있을 것이다.

본 연구에서는 그림 3과 같이 SR-GAN에 CBAM (Convolutional Block Attention Module)[6]을 추가하는 방식을 제안한다. CBAM은 그림 4와 같은 구조를 가지며, 이미지의 특징을 더 잘 보존하거나 강조하는 역할을 한다. 이러한 CBAM 모듈을 SR-GAN의 Generator에 추가함으로써, 이미지 품질을 더욱 개선하고, 이를 통해 YOLO(You Only Look Once) v5[7]가 철근 검출을 더 잘 수행할 수 있게 하는 것이 이 연구의 주요 목표다.

II. 관련 연구

2.1 Channel Attention 메커니즘과 SR-GAN 네트워크

B. Liu 연구[8]에서는 SR-GAN을 기반으로 세 가지 개선 사항을 도입하여 연구하였다. 첫째로, 고주파수 특징(이미지나 신호에서 주파수가 높은 변화를 나타내는 부분)을 더 잘 표현하기 위해 attention 모듈을 활용하여, channel attention 모듈을 SR-GAN 네트워크에 추가하였다. 둘째로, 네트워크 성능을 향상시키기 위해 원래의 배치 정규화 레이어를 제외하였다. 마지막으로, 이미지의 노이즈 영향을 줄이기 위해 손실 함수를 수정하였다. 실험 결과는 제안한 방법이 다른 접근법들보다 우수함을 입증하며, 고주파수 정보를 효과적으로 복원하고 PSNR 향상을 달성하였다.

2.2 SAGAN: 이상 탐지를 위한 Skip-Attention GAN

G. Liu 연구[9]에서는 이미지 기반 이상 탐지를 위해 생성적 적대 신경망을 활용하여 이상치를 탐지하는 방법을 소개하였다. 기존의 접근 방식은 입력 이미지와 생성된 이미지 사이의 전역적 차이를 분석하여 이상치를 식별하는 것이었다. 그러나 실제 이상치는 종종 이미지의 지역적인 부분에서 발생하며, 이러한 지역 정보를 고려하지 않으면 신뢰할 수 없는 이상 탐지 결과를 얻을 수 있다. 본 연구에서는 SAGAN(Skip-Attention GAN)이라는 효율적인 이상 탐지 네트워크를 제안하였다. 이 네트워크는 로컬 정보를 더 잘 포착하기 위해 attention 모듈을 도입하여 이미지의 잠재적인 표현 정확도를 향상시켰다. 또한, 모델의 파라미터 수를 줄이기 위해 depthwise separable convolution을 적용하였다. 제안한 방법은 CIFAR-10 데이터셋과 자체 생성한 LBOT 데이터셋에서 평가하였으며, AUC(Area Under Curve) 지표를 사용하여 성능을 평가하였다. 실험 결과로는 두 데이터셋 모두에서 제안한 방법이 기존의 EGBAD, GANomaly 방법과 비교하여 평균적으로 10% 이상의 성능 향상을 보였다.

2.3 Attention 기반 초해상도 GAN

B. Yang 연구[10]에서는 attention 모듈을 활용하여 개선된 attention 기반의 SR-GAN을 제안하였다. attention 모듈은 중요한 특징을 강조하고 중요하지 않은 특징을 억제하는 데 활용될 수 있어 네트워크 구조의 품질을 보장하고 SR-GAN의 생성자 네트워크를 최적화하는 데 도움이 된다. 실험 결과로는 Set5 데이터셋에서 attention 기반 SR-GAN 모델을 사용하면 더 적은 잔여 블록을 훈련하여 재구성 시간을 감소시킬 수 있음을 보여주었다. 이는 더 적은 매개변수와 계산 비용을 필요로 한다는 의미를 갖는다. 또한, Set14와 BSD100 데이터셋에서 선택한 이미지의 재구성도 개선되었으며, 재구성 시간도 단축되었음을 보여주었다.

III. SRPAttention-GAN 모델

3.1 철근 데이터셋

머신비전 카메라로부터 그림 5와 같이 동영상 9개를 수집하였다. 각각의 동영상으로부터 철근의 다양한 상황을 반영한 그림 6과 같은 416x416 고해상도 이미지 500장(HR) 그리고 104x104의 저해상도 이미지 250장(LR)을 추출하였다.

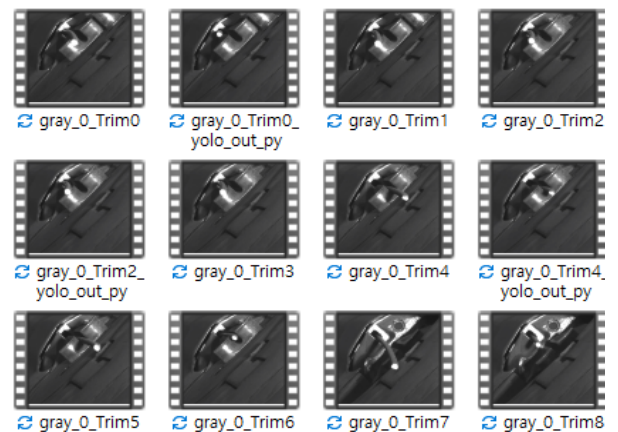


그림 5. 철근 사출 동영상 9개
Fig. 5. Rebar injection nine videos

3.2 철근 끝점 검출을 위한 구조

그림 7은 철근 끝점 검출 및 성능 향상 구조를 보여주고 있다. 머신비전 카메라로부터 철근 사출 동영상 9개를 수집한 후 해당 동영상에서 상대적으로 노이즈가 적은 416×416의 고해상도 이미지 500장과 104×104의 저해상도 이미지 250장을 추출한다.

이후 해당 두 가지 해상도의 이미지 데이터셋을 기존의 SRResNet-GAN과 제안하는 SRPAAttention-GAN에 각각 학습시켜, 이미지를 복원하고 생성하여 총 4가지 종류의 데이터를 생성한다. 철근 끝점의 검출 성능을 확인하기 위해, 고해상도 철근 이미지 500장을 학습시켜 생성된 두 가지 데이터셋을 YOLOv5에 학습시키는 구조를 사용한다.



(a) 416x416 고해상도 철근 이미지
(a) 416x416 high resolution rebar image



(b) 104x104 저해상도 철근 이미지
(b) 104x104 low resolution rebar image

그림 6. 철근 사출 동영상 기반 고해상도 및 저해상도 철근 이미지

Fig. 6. High-resolution and Low-resolution rebar images extracted from rebar injection videos

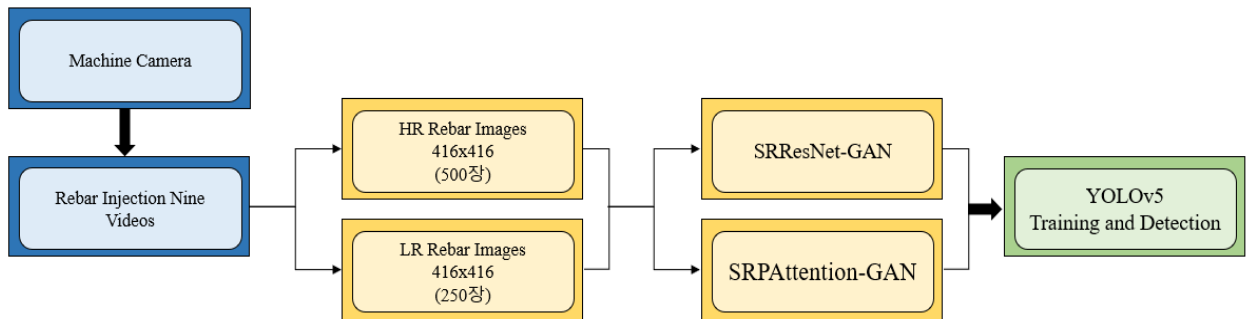


그림 7. 철근 끝점 검출을 위한 전체적인 프로세스

Fig. 7. Overall process for detecting rebar endpoints

3.3 제안하는 SRPAttention-GAN 구조

제안하는 구조는 그림 8과 같이 저해상도 이미지(LR, Low-Resolution)를 입력받아, CBAM의 attention 모듈과 함께 작동하는 SR generator를 이용해 고해상도 이미지(HR, High-Resolution)로 변환한다. CBAM의 attention 모듈은 그림 9와 같이 channel attention 모듈과 그림 10과 같이 spatial attention module로 구성되어 있다. 이 두 가지 모듈을 SR generator에 결합하는데 먼저 channel attention 모듈은 먼저 채널 간의 상호작용을 모델링 하며, 단일 채널(흑백 이미지)의 중요성을 계산하여 입력 이미지의 중요한 특징을 강조한다. 주로 이미지의 색조, 명암, 채도 등 픽셀의 속성 표현에 중요한 역할을 한다. 이로 인해 고해상도 이미지에서는 색상의 다

양성과 명암을 정확하게 재현할 수 있다. Spatial attention 모듈은 입력 이미지의 공간적 구조와 패턴을 고려하여 중요한 위치를 강조하고 generator가 해당 부분에 집중할 수 있도록 도와준다. 이를 통해 고해상도 이미지에서의 해상도, 선명도, 질감, 세부 구조 등을 잘 표현할 수 있다. 특히, 이미지 생성 시 공간 정보를 고려하지 않으면 이미지 생성 결과에 노이즈가 생길 수 있으므로, 고해상도 이미지의 품질과 신뢰성을 높이기 위해 spatial 정보와 channel 정보가 모두 중요하다. 마지막으로, SR discriminator는 생성된 고해상도 이미지를 이용해 진짜 HR 이미지와 가짜 HR 이미지를 구분하면서 경쟁 학습을 진행한다. 제안하는 SRPAttention-GAN 구조는 이미지의 중요한 특징을 강조하며, 고해상도 이미지의 생성에 도움을 주게 된다.

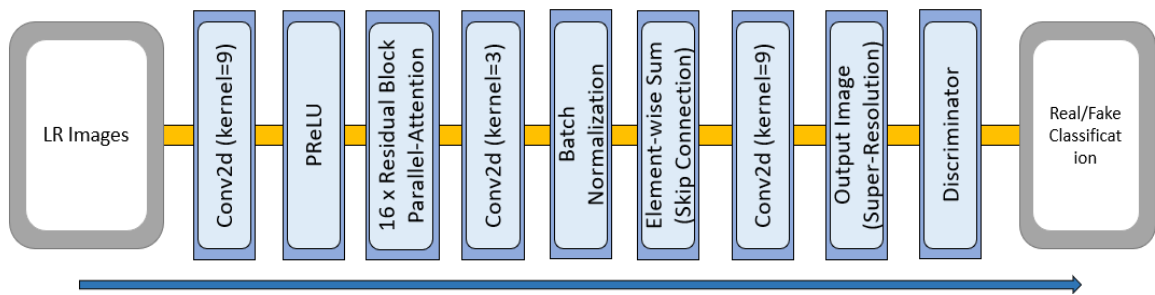


그림 8. 제안하는 SRPAttention-GAN 구조
Fig. 8. Proposed SRPAttention-GAN structure

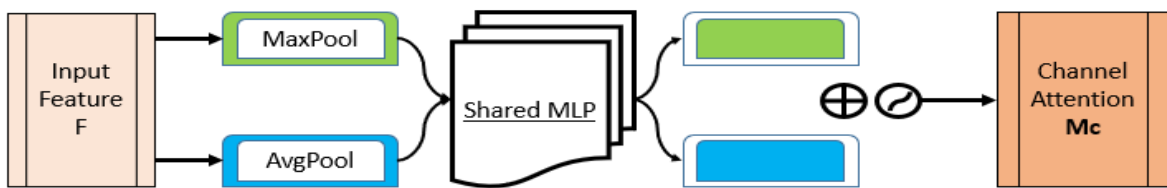


그림 9. CBAM의 channel attention module[6]
Fig. 9. Channel attention module of CBAM[6]

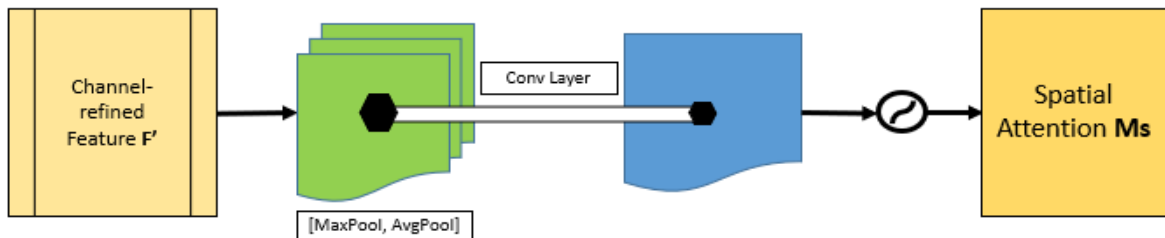


그림 10. CBAM의 spatial attention module[6]
Fig. 10. Spatial attention module of CBAM[6]

제안하는 모델의 핵심적인 요소인 residual block의 구조는 표 1과 같다. Residual block은 두 개의 Conv2d, 한 개의 BatchNorm2d, 그리고 한 개의 PReLU로 구성되어 있다. 이들은 모두 입력으로 64개의 채널을 가진 텐서를 받아 동일한 크기의 출력을 생성한다. 여기에는 CBAM의 attention 모듈이 추가되어 있으며, spatial와 channel은 병렬의 형태로 Parallel-Attention이 최종 addition이 되고 입력 텐서는 residual block의 출력에 더해져 skip connection을 형성한다. 이 과정은 입력과 출력 사이의 정보 흐름을 향상시킨다.

IV. 실험 방법 및 성능평가

본 연구의 실험 환경은 Intel(R) Core(TM) i9-13900 CPU, NVIDIA GeForce RTX 4080 16GB, RAM 32GB를 사용하여 진행하였다.

처음에, HR과 LR을 SRResNet-GAN과 제안하는 SRPAttention-GAN에 학습시킨 후 생성된 데이터셋은 총 4가지로, 이는 HR과 LR 각각에서 생성된 데이터셋 2가지씩을 포함한다. 또한 제안된 SRPAttention-GAN 모델의 성능 평가를 위해 PSNR (Peak Signal-to-Noise Ratio)와 SSIM(Structural Similarity Index Measure)를 사용하여 측정하였다. 이 이미지 품질 측정 결과는 표 2와 같다.

그림 11, 12는 표 2에 따른 이미지 품질 측정 결과를 히스토그램을 통해 보여주고 있다. 해당 결과를 통해 원본 HR 이미지를 사용하여 기존의 SRResNet-GAN과 제안하는 SRPAttention-GAN 모델을 이용하여 이미지를 생성해 비교하였을 때, SRPAttention-GAN 모델이 평균적으로 더 높은

PSNR 값을 달성한 것을 확인할 수 있다. 이는 SRPAttention-GAN 모델이 원본 이미지를 재구성하는데 SRResNet-GAN보다 더 효과적임을 나타내었다. 그러나 SSIM 점수의 경우, 두 모델 사이에 큰 차이를 보이지 않아, 두 모델이 이미지의 구조적 유사성을 유지하는 데 동등한 성능을 가진 것으로 보였다.

저해상도에서 고해상도로 이미지를 복원하거나 생성할 때, 초반 epoch 250-500구간에서 생성된 이미지는 SRResNet-GAN 모델이 더 높은 PSNR와 SSIM 값을 얻었다. 그러나 학습이 진행되면서 epoch 350-500구간에서 생성된 이미지의 품질을 측정한 결과, SRPAttention-GAN 모델이 더 높은 PSNR과 SSIM 값을 얻었다. 이 결과는 SRPAttention-GAN 모델의 성능이 학습 진행에 따라 향상됨을 보여주었다. 즉, 더 많은 epoch에서 학습을 진행하면, SRPAttention-GAN 모델로부터 더 높은 품질의 이미지를 생성할 수 있음을 보여준다.

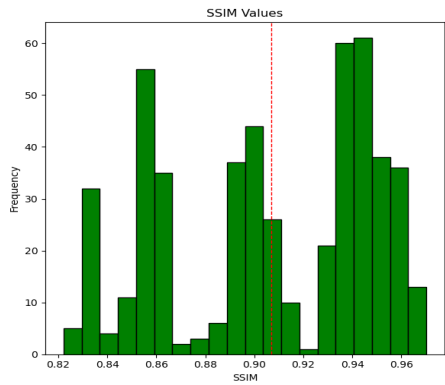
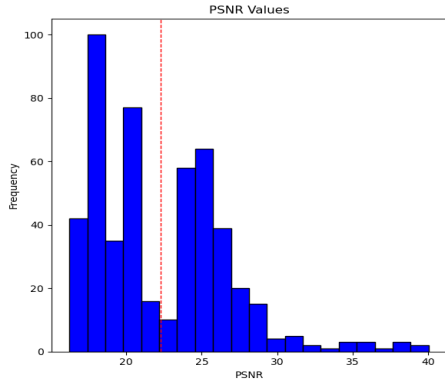
표 2. 이미지 품질 측정 결과 (PSNR = db, SSIM = 0-1)
Table 2. Image quality measurement results (PSNR = db, SSIM = 0-1)

모델(GAN)	Average PSNR	Average SSIM
SRResNet(HR)	22.204	0.907
SRPAttention(HR)	22.269	0.907
SRResNet(LR) (epoch 250-500)	23.033	0.913
SRPAttention(LR) (epoch 250-500)	22.589	0.911
SRResNet(LR) (epoch 350-500)	22.663	0.912
SRPAttention(LR) (epoch 350-500)	22.964	0.917

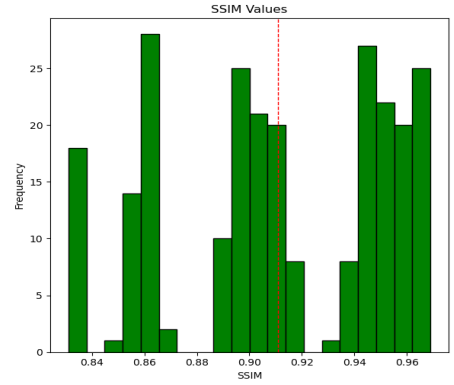
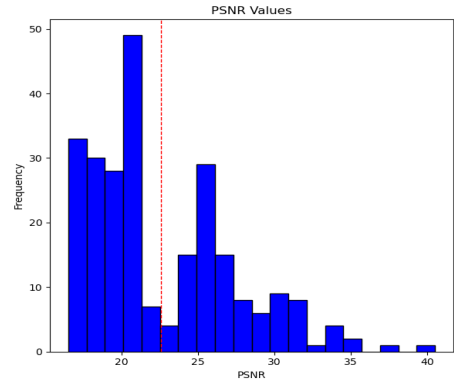
표 1. SRPAttention generator의 residual block 구조

Table 1. Structure of the residual block in the SRPAttention generator

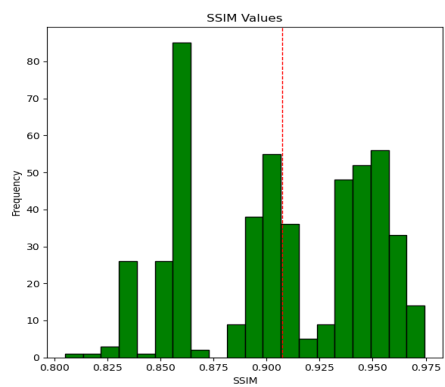
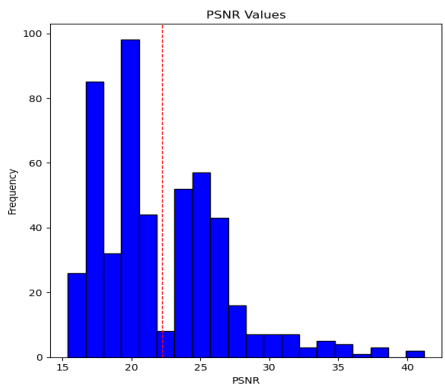
Layer	Output size	Kernel size	Stride	Padding	Activation
Conv2d	64	3	1	1	-
BatchNorm2d	64	-	-	-	-
PReLU	64	-	-	-	-
Conv2d	64	3	1	1	-
BatchNorm2d	64	-	-	-	-
Parallel-Attention	64	-	-	-	-



(a) SRResNet-GAN(HR)의 PSNR, SSIM 히스토그램
(a) SRResNet-GAN(HR)의 PSNR, SSIM histogram

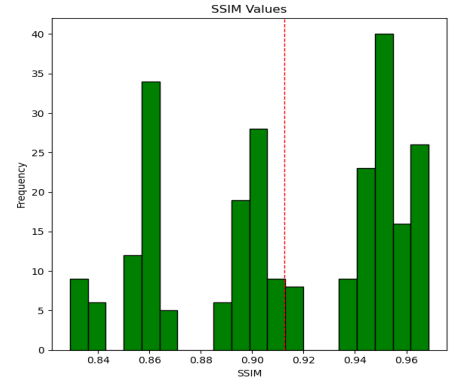
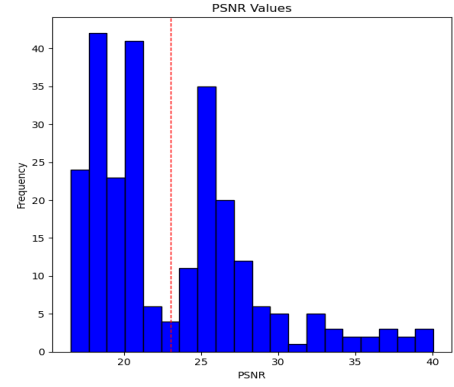


(a) SRResNet-GAN(LR)의 PSNR, SSIM 히스토그램
(a) SRResNet-GAN(LR)의 PSNR, SSIM histogram



(b) SRPAttention-GAN(HR)의 PSNR, SSIM 히스토그램
(b) SRPAttention-GAN(HR)의 PSNR, SSIM histogram

그림 11. HR-HR 이미지 품질 측정 결과
Fig. 11. HR-HR image quality measurement results



(b) SRPAttention-GAN(LR)의 PSNR, SSIM 히스토그램
(b) SRPAttention-GAN(LR)의 PSNR, SSIM histogram

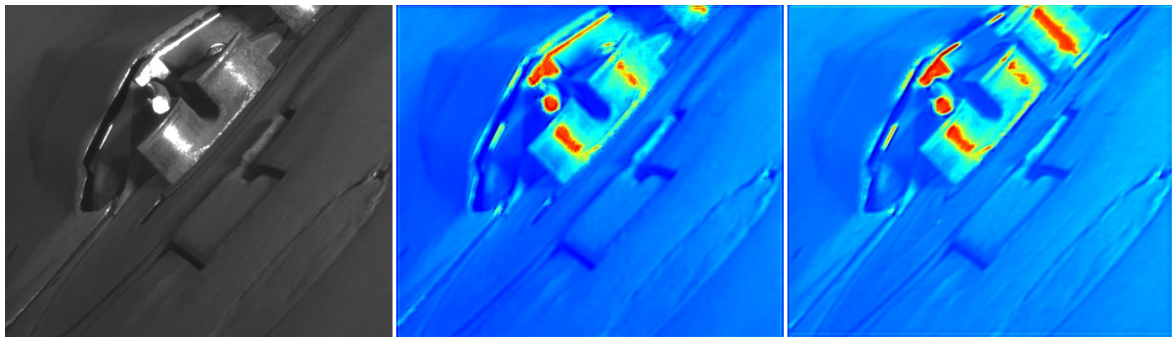
그림 12. LR-HR 이미지 품질 측정 결과 (epoch 250-500)
Fig. 12. LR-HR image quality measurement results (epoch 250-500)

그림 13의 (a)에서 (왼쪽)은 원본 HR 이미지고, (중앙)은 HR-HR로 바꾸는 SRResNet-GAN으로 생성된 이미지이다. (오른쪽)은 제안한 SRPAttention-GAN으로 생성된 이미지이다. (b)는 생성자에서 학습을 하면서 이미지를 생성할 때에 이미지 안에서 어떤 부분을

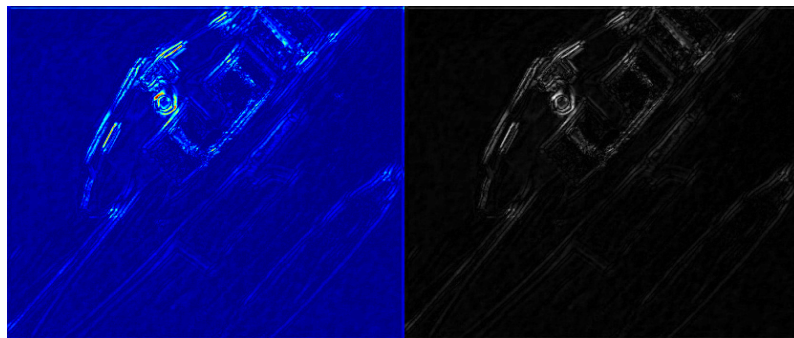
집중했는지 Grad-Cam 히트맵[11]으로 시각화한 결과이다. 마지막으로 (c)는 각각 SRResNet-GAN과 SRPAttention-GAN이 각각 이미지 영역에서 어떤 부분에 집중했는지 알기 위하여 시각화한 히트맵의 비교 결과이다.



(a) (왼쪽) 원본 HR 이미지 (중앙) SRResNet-GAN을 통해 생성된 이미지 (오른쪽) SRPAttention-GAN에서 생성된 이미지
 (a) (Left) Original HR image (Center) image generated through SRResNet-GAN (Right) image generated through SRPAttention-GAN



(b) (왼쪽) 원본 HR 이미지 (중앙) SRResNet-GAN을 통해 생성된 이미지의 Grad-CAM 시각화 결과 (오른쪽) SRPAttention-GAN에서 생성한 이미지의 Grad-CAM 시각화 결과
 (b) (Left) Original HR image (Center) Grad-CAM visualization result of the image generated through SRResNet-GAN (Right) Grad-CAM visualization result of the image generated through SRPAttention-GAN



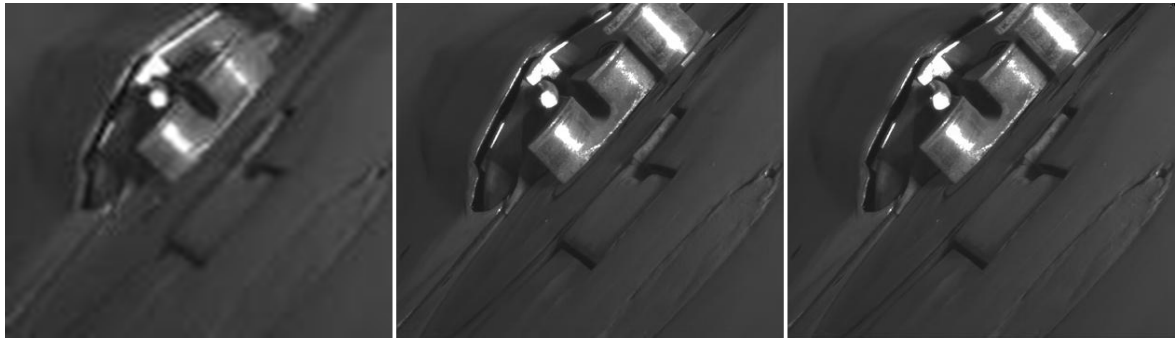
(c) 각각 SRResNet-GAN과 SRPAttention-GAN에서 생성한 이미지의 Grad-CAM시각화 결과의 비교 분석 결과 (밝은 부분일수록 SRPAttention-GAN에서 강조한 부분)
 (c) Comparison and analysis of Grad-CAM visualization results of images generated from SRResNet-GAN and SRPAttention-GAN (Bright areas indicate regions emphasized by SRPAttention-GAN)

그림 13. HR 철근 이미지를 통한 SRResNet-GAN과 SRPAttention-GAN에서 생성한 HR 철근 이미지들의 비교 분석
 Fig. 13. Comparative analysis of SRResNet-GAN through HR rebar images and HR rebar images created by SRPAttention-GAN

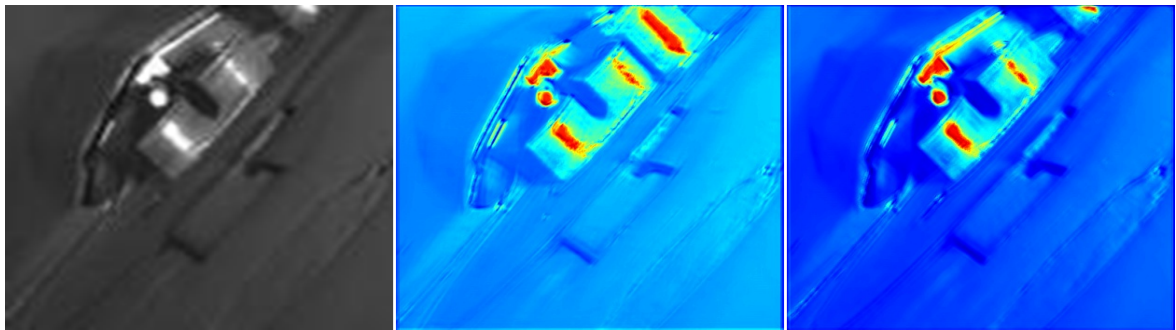
그림 14의 (a)는 (왼쪽)은 LR 이미지이고, (중앙)은 LR-HR로 변환하는 SRResNet-GAN을 통해 생성된 이미지이다. (오른쪽)은 LR-HR로 변환하는 SRPAttention-GAN에서 생성한 이미지이다. (b),(c)는 그림 14와 마찬가지로 시각화한 분석 결과이다.

Grad-CAM 히트맵을 통해서 두 모델이 어떤 이

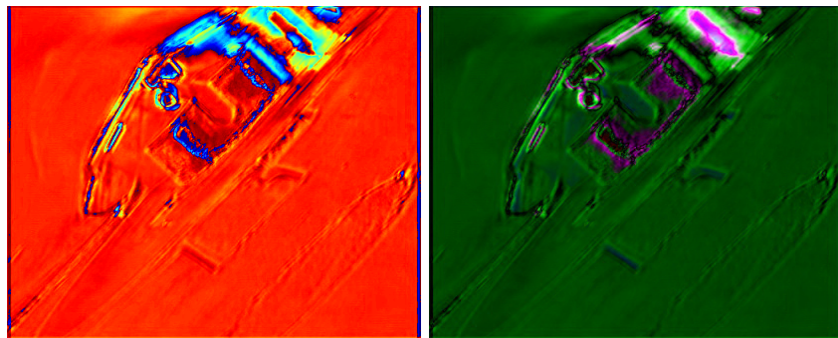
미지 영역에 주의를 기울였는지 확인할 수 있다. 그림 13, 14에서 (c)를 보면, 제안하는 SRPAttention-GAN에서 생성한 이미지는 밝은 부분일수록 SRPAttention-GAN에서 강조한 부분이며, 특히 이미지의 중앙 부분인 철근의 끝점에 주목하고 있는 것으로 보인다.



(a) (left) 원본 LR 이미지 (middle) SRResNet-GAN을 통해 생성된 이미지 (right) SRPAttention-GAN에서 생성된 이미지
(a) (Left) Original LR image (Middle) Image generated through SRResNet-GAN (Right) Image generated through SRPAttention-GAN



(b) (left) 원본 LR 이미지 (middle) SRResNet-GAN을 통해 생성된 이미지의 Grad-CAM 시각화 결과 (right) SRPAttention-GAN에서 생성된 이미지의 Grad-CAM 시각화 결과
(b) (Left) Original LR image (Middle) Grad-CAM visualization result of the image generated through SRResNet-GAN (Right) Grad-CAM visualization result of the image generated through SRPAttention-GAN



(c) 각각 SRResNet-GAN과 SRPAttention-GAN에서 생성한 이미지의 Grad-CAM시각화 결과의 비교 분석 결과(밝은 부분일수록 SRPAttention-GAN에서 강조한 부분)

(c) Comparison and analysis of Grad-CAM visualization results of images generated from SRResNet-GAN and SRPAttention-GAN (Bright areas indicating regions emphasized by SRPAttention-GAN)

그림 14. LR 철근 이미지를 통한 SRResNet-GAN과 SRPAttention-GAN에서 생성한 HR 철근 이미지들의 비교 분석
Fig. 14. Comparative analysis of SRResNet-GAN with LR rebar image and HR rebar images created by SRPAttention-GAN

반면 기존의 SRResNet-GAN에서는 어두운 부분 일수록 SRResNet-GAN에서 강조한 부분이지만, 이는 이미지 전체에 더 넓게 분산된 관심을 나타내고 있다. 즉, SRResNet-GAN은 철근의 다양한 부분에 주목하면서 이미지를 생성했다고 볼 수 있다. 이렇게 복원된 이미지의 특징이 약간 다르게 나타나긴 하지만, 특정 객체를 탐지하는 작업에서는 중요한 영향을 미칠 수 있다는 것을 알 수 있다. 이를 증명하기 위해 원본 철근 이미지 500장, HR-HR로 변환하고 복원한 SRPAttention-GAN에서 생성한 철근 이미지 500장, 그리고 SRResNet-GAN에서 생성한 철근 이미지 500장을 각각 YOLOv5 모델로 학습시켜 검출 성능을 실험했다. YOLOv5 학습을 위한 학습 하이퍼파라미터 환경은 표 3과 같으며, 데이터 증강 하이퍼파라미터는 표 4와 같다.

표 3. YOLOv5 학습을 위한 하이퍼파라미터
Table 3. Hyperparameters for learning YOLOv5

Hyperparameter	Value
learning rate	0.01
momentum	0.937
weight_decay	0.0005
warmup_epochs	3.0
warmup_momentum	0.8
warmup_bias_lr	0.1
box	0.05
cls	0.5
clw_pw	1.0
obj	1.0
obj_pw	1.0
iou_t	0.2
anchor_t	4.0

표 4. YOLOv5 학습을 위한 데이터 증강 하이퍼파라미터
Table 4. Data augmentation Hyperparameters for YOLOv5 learning

Hyperparameter	Value
hsv_h	0.015
hsv_s	0.7
hsv_v	0.4
degrees	0.0
translate	0.1
scale	0.5
shear	0.0
perspective	0.0
flipud	0.0
fliplr	0.5
mosaic	1.0
mixup	0.0

YOLOv5에 학습 및 검출 결과는 그림 15, 16, 17에 각각 표시되어 있다. 학습 결과, 모든 데이터셋은 평균 mAP 0.955를 기록했다. 그러나 더 세부적으로 살펴보면, 원본 HR 철근 이미지 500장을 학습시켰을 때의 F1와 정밀도는 각각 0.66과 0.771로 나타났다. SRResNet-GAN에서 생성된 HR 철근 이미지 500장을 학습시켰을 경우 0.7373와 0.788의 결과를 보였다. 반면, 제안한 SRPAttention-GAN 모델에서 생성된 HR 철근 이미지 500장을 YOLOv5에 학습시켰을 경우 F1와 정밀도가 각각 0.743과 0.793으로, 가장 높은 성능을 보였다. 이는 YOLOv5에서 F1와 정밀도 성능을 통해 모델이 객체를 더 잘 인식하고, 예측이 더 정확하다는 것을 의미한다. 즉, 모델이 철근 이미지에서 객체의 위치와 클래스를 더 정확하게 예측함을 알 수 있다.

V. 결론 및 향후 과제

본 논문에서는 철근 가공 현장에서 발생하는 노이즈로 인한 이미지 품질 저하와 딥러닝 모델의 검출 성능 향상을 위한 새로운 모델을 제안하였다. 본 논문에서는 SR-GAN과 Parallel Attention Mechanism을 결합한 SRPAttention-GAN 모델을 사용하여, 저해상도의 철근 이미지를 고해상도 이미지로 변환하였다. 이렇게 변환된 이미지는 YOLOv5 모델을 통해 더 높은 품질의 철근 끝점 검출을 수행하였다.

실험 결과를 통해 제안하는 SRPAttention-GAN 모델은 기존의 SRResNet-GAN 모델 대비 PSNR이 약 0.3dB 더 높음을 확인하였다. 또한, 변환된 이미지를 YOLOv5에 학습시킨 결과, F1 점수와 정밀도에서 더욱 개선된 결과를 얻었다. 이는 SRPAttention-GAN 모델이 고해상도 이미지를 생성하여 철근 끝점을 더 정확하게 검출하는 성능을 보여주었다. 성능 향상을 위하여 철근 끝점 검출의 정확성을 향상시키는 새로운 방법론을 제시하였으며, 본 논문에서 제안한 방법을 통해 철근 가공 공정에서의 불량품 생성을 최소화하는 데 기여할 것으로 기대된다. 또한, 다른 이미지 품질이 감소하는 환경에서의 객체 검출 문제에도 적용할 수 있을 것으로 예상된다.

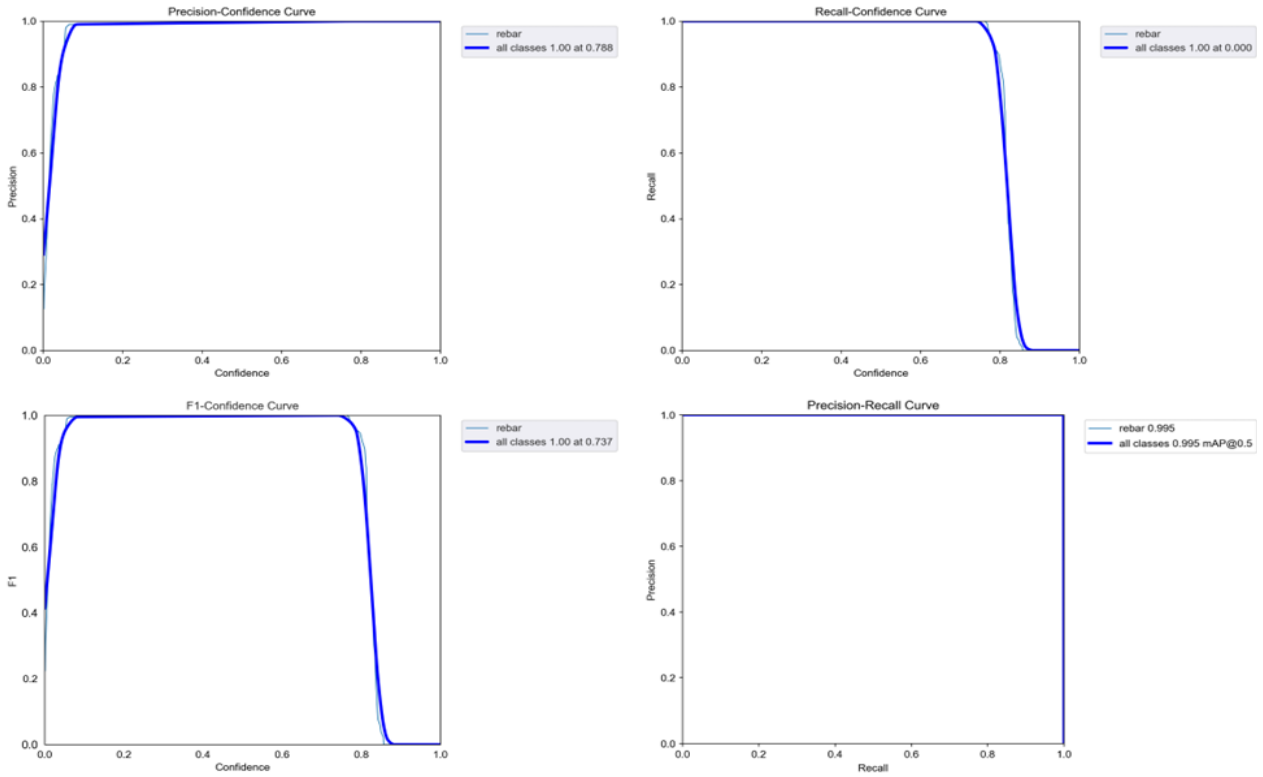


그림 16. SRResNet-GAN에서 생성된 HR 철근 이미지 500장을 YOLOv5에 학습하여 얻은 성능 결과
 Fig. 16. Performance results obtained by learning 500 HR rebar images from SRResNet-GAN to YOLOv5

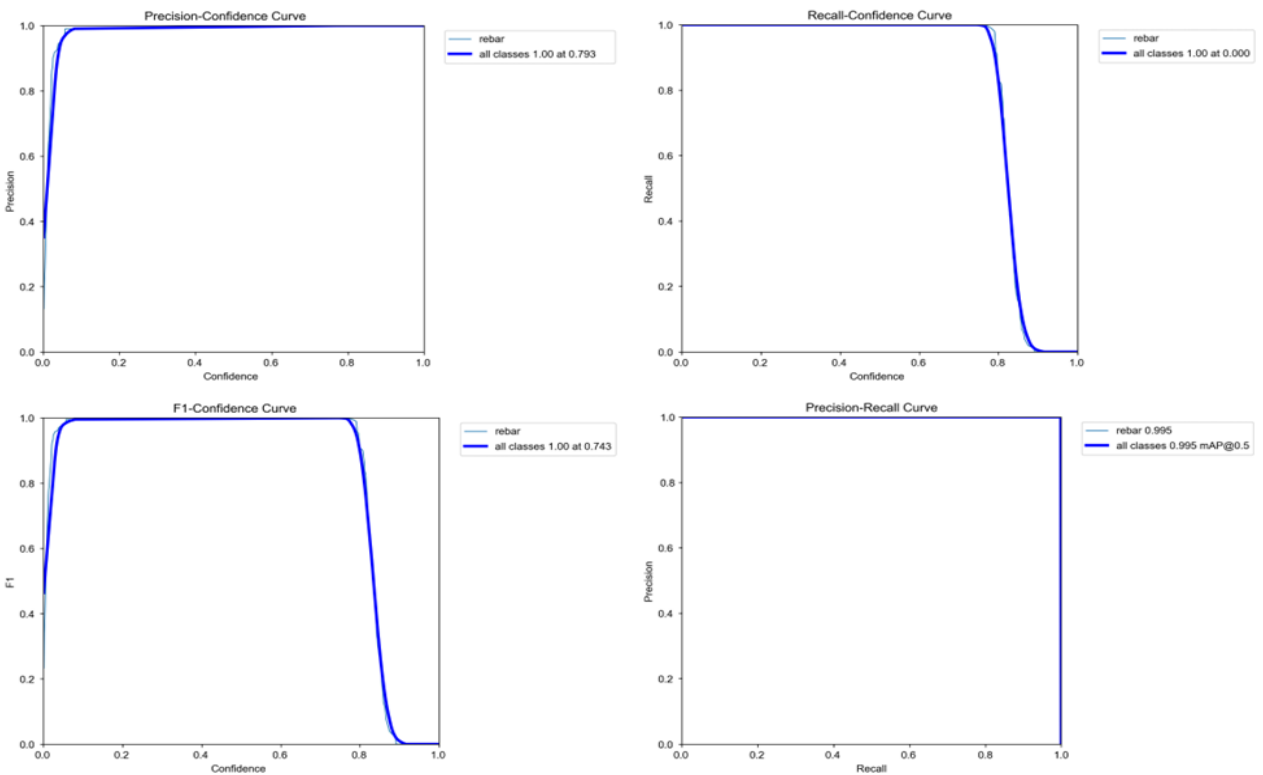


그림 17. SRPAttention-GAN에서 생성된 HR 철근 이미지 500장을 YOLOv5에 학습하여 얻은 성능 결과
 Fig. 17. Performance results obtained by learning 500 HR rebar images generated by SRPAttention-GAN in YOLOv5

향후 연구에서는 향상된 철근 끝점 검출을 통해 얻은 좌표를 선형 / 비선형 회귀 모델을 이용하여 학습시켜, 예측 지점을 표시하고 실제 철근의 위치와 예측 지점의 오차율을 줄이는 방안을 모색할 계획이다. 추가적으로, 철근 가공 사출 공정의 특성을 고려하여 기존의 회귀 모델에서 예측 함수 수식을 개선하여 오차율을 줄이는 연구를 진행할 예정이다.

References

- [1] Z. Fan, J. Lu, B. Qiu, T. Jiang, K. An, A. N. Josephraj, and Ch. Wei, "Automated steel bar counting and center localization with convolutional neural networks", arXiv:1906.00891, Jun. 2019. <https://doi.org/10.48550/arXiv.1906.00891>.
- [2] J. Han, H.-Y. Lim, and D.-S. Kang, "GPC: Grip Position Coordinate Detection Model for YOLO-based Rebar Shape", Proc. of KIIT Conference, Jeju, Korea, pp. 145-148, Nov, 2021.
- [3] J.-C. Park, H.-Y. Lim, and D.-S. Kang, "Predicting Rebar endpoints using sin exponential regression model", arXiv:2110.08955, Oct, 2021. <https://doi.org/10.48550/arXiv.2110.08955>.
- [4] J.-C. Park and G.-W. Kim, "Real-time Twist Rebar Detection System exploiting GAN-based Data Augmentation technique", CEUR Workshop Proceedings, Japan, Vol. 3362, Dec. 2022.
- [5] C. Ledig, et al., "Photo-realistic single image super-resolution using a generative adversarial network", Proc. of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, pp. 105-114, Jul. 2017. <https://doi.org/10.1109/CVPR.2017.19>.
- [6] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module", Proc. of the European conference on computer vision (ECCV), Munich, Germany, Vol. 11211, pp. 3-19, Sep. 2018. https://doi.org/10.1007/978-3-030-01234-2_1.
- [7] A. B. Wang, Y. Holoborodko, and M. Shvets, "YOLOv5: A Unified Approach to Real-Time Object Detection", GitHub repository, <https://github.com/ultralytics/yolov5> [accessed: Jul. 01, 2023].
- [8] B. Liu and J. Chen, "A super resolution algorithm based on attention mechanism and srgan network", IEEE Access, Vol. 9, pp. 139138-139145, Jul. 2021. <https://doi.org/10.1109/ACCESS.2021.3100069>.
- [9] G. Liu, S. Lan, T. Zhang, W. Huang, and W. Wang, "SAGAN: skip-attention GAN for anomaly detection", 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, pp. 2468-2742, Sep. 2021. <https://doi.org/10.1109/ICIP42928.2021.9506332>.
- [10] B. Yang, K. Xie, Z. Yang, and M. Yang, "Super-resolution generative adversarial networks based on attention model", 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, pp. 781-786, Dec. 2020. <https://doi.org/10.1109/ICCC51575.2020.9345044>.
- [11] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization", Proc. of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, pp. 618-626, Oct. 2017. <https://doi.org/10.1109/ICCV.2017.74>.

저자소개

박종찬 (Jong-Chan Park)



2022년 3월 ~ 현재 :
 경상국립대학교 AI융합공학과
 박사과정
 관심분야 : 영상처리, 인공지능,
 이미지 증강, 스마트 팩토리

허 지 욱 (Jee-Uk Heu)



2007년 2월: 한림대학교
컴퓨터공학과(공학사)
2009년 2월 : 한양대학교
컴퓨터공학과(공학석사)
2016년 2월 : 한양대학교
컴퓨터공학과(공학박사)
2021년 3월~ 현재 : 강남대학교

참인재대학 교양교수부 조교수

관심분야 : 인공지능, 시멘틱 헬스케어, 데이터마이닝

김 건 우 (Gun-Woo Kim)



2006년 12월 : 호주뉴캐슬대학교
컴퓨터공학과(공학사)
2007년 9월 : 호주뉴캐슬대학교
정보공학과(공학석사)
2017년 8월 : 한양대학교
컴퓨터공학과(공학박사)
2021년 9월 ~ 현재 :

경상국립대학교 컴퓨터과학부 조교수

관심분야 : 인공지능, 시멘틱 헬스케어, 데이터마이닝