Journal of Digital Policy
Vol. 3. No. 3, pp. 9-14, 2024
ISSN 2951-245X
https://doi.org/10.23149/JDP.2024.3.3.009

적응형 가중치 잔차 블록을 적용한 다중 블록 구조 기반의 단일 영상 초해상도 기법

한현호
울산대학교 교양학부 조교수

Single Image Super Resolution using Multi Grouped Block with Adaptive Weighted Residual Blocks

Hyun Ho Han
Assistance Professor, College of General Education, University of Ulsan, Korea

요 약 본 논문은 단일 영상 기반의 초해상도에서 결과의 품질을 개선하기 위해 적응형 가중치를 적용한 잔차 블록으로 구성된 다중 블록 구조를 이용하는 방법을 제안하였다. 딥러닝을 이용한 초해상도를 생성하는 과정에서 품질 향상을 위한 가장 중요한 요소는 특징 추출 및 적용이다. 해상도가 낮아 이미 손실된 세부사항을 복원하기 위해 다양한 특징을 추출하는 것이 최우선이지만 네트워크의 구조가 깊어지거나 복잡해지는 등의 문제가 발생하기 때문에 실제 적용에서 제한사항이 있다. 따라서 특징 추출 과정은 효율적으로 구성하고 적용 과정을 개선하여 품질을 개선하였다. 이를 위해 최초 특징 추출 이후 다중 블록 구조를 구성하였고 블록 내부에는 중첩된 잔차 블록을 구성한 뒤 적응형 가중치를 적용하였다. 또한 최종 고해상도 복원을 위해 다중 커널을 이용한 영상 재구성 과정을 적용함으로써 결과물의 품질을 향상시켰다. 평가를 위해 원본 영상 대비 PSNR과 SSIM 값을 구하였고 기존 알고리즘과 비교하여 제안하는 방법의 성능 향상을 확인하였다.

주제어 : 초해상도, 적응형 가중치, 다중 블록, 특징, 딥러닝

Abstract In this paper, proposes a method using a multi block structure composed of residual blocks with adaptive weights to improve the quality of results in single image super resolution. In the process of generating super resolution images using deep learning, the most critical factor for enhancing quality is feature extraction and application. While extracting various features is essential for restoring fine details that have been lost due to low resolution, issues such as increased network depth and complexity pose challenges in practical implementation. Therefore, the feature extraction process was structured efficiently, and the application process was improved to enhance quality. To achieve this, a multi block structure was designed after the initial feature extraction, with nested residual blocks inside each block, where adaptive weights were applied. Additionally, for final high resolution reconstruction, a multi kernel image reconstruction process was employed, further improving the quality of the results. The performance of the proposed method was evaluated by calculating PSNR and SSIM values compared to the original image, and its superiority was demonstrated through comparisons with existing algorithms.

Key Words : Super Resolution, Adaptive Weight, Multi Grouped Block, Feature, Deep Learning

*Corresponding Author : Hyun Ho Han(hhhan@ulsan.ac.kr)
Received August 14, 2024
Revised September 7, 2024
Accepted September 20, 2024
Published September 28, 2024

1. 서론

단일 영상을 이용한 초해상도는 하나의 저해상도 영상을 이용해 고해상도 영상으로 복원하는 방법이다. 한 장면을 이용해 고해상도 복원을 수행하기에 다양한 컴퓨터 비전 분야에서 활용될 수 있다. 특정 대상을 감지하거나 얼굴의 인식, 의료 영상에서의 식별 보조 등의 과정에서 전처리 단계로써 사용되기도 하고, 고화질의 영상 콘텐츠를 제작, 복원하는데 활용되는 등이 대표적인 예 이다[1,2]. 활용 예에서도 확인할 수 있듯 초해상도의 결과의 품질이 매우 중요하다. 그러나 단일 영상을 이용한 초해상도는 이미 손실되어 존재하지 않는 화소값을 복원해야 하기 때문에 값을 추정하는 과정에서 명확한 정답이 없는 문제를 해결해야 한다. 이 과정에서 입력 영상 한 장 내에 있는 정보를 근거로 화질 복원을 수행하기에 결과물의 품질 수준이 떨어지게 된다[3,4].

초해상도를 수행하기 위해 입력 영상 내의 정보를 활용해 복원하는 방법은 복원할 고해상도로 영상의 크기를 확장하면 값이 존재하는 화소 사이에 빈 화소가 존재하게 되는데 이를 화소 사이의 값으로 채우는 단순 보간법과 그 응용이 있다[5]. 이를 개선하기 위해 사전에 학습 데이터 셋을 이용해 저해상도-고해상도 패치 쌍 구조를 갖는 사전을 구성하고 입력 영상에서 저해상도 패치 단위의 분석을 거쳐 고해상도 패치를 매치하는 방법으로 초해상도를 수행하는 방법이 제안되었다[6]. 이후 CNN(Convolutional Neural Network)를 이용해 딥러닝을 기반으로 하는 초해상도 방법으로 발전되었다[7].

딥러닝 기반의 방법이 제안되면서 초해상도 결과 품질의 수준은 급성장하게 됐지만 딥러닝을 이용한 방법에서 품질을 개선하기 위해 네트워크의 매개변수가 많아지거나 깊이가 깊어져 연산하는데 소요되는 시간이 상당히 길어지는 문제가 있다[8]. 또한 모델의 규모가 커지더라도 세부 정보가 약화되어 오히려 결과물의 품질이 떨어질 수 있다. 제안하는 방법은 모델의 규모를 유지하면서 네트워크의 매개변수를 줄이고 깊이를 깊지 않게 구성할 수 있도록 효율적인 특징 추출 과정과 적용 과정에서의 활용을 개선하여 초해상도 결과의 품질을 향상시키는 방법을 제안하였다.

2. 관련 연구

최근의 단일 영상 기반의 초해상도는 딥러닝 기반의 방법이 주로 연구되고 있다. SRCNN(Super Resolution CNN)은 최초로 깊은 CNN 기반의 단일 영상 초해상도 모델로 제안되었다[9]. SRCNN은 기존의 보간법, 희소 표현 기반의 방법들과 비교했을 때 큰 성능 차이를 보여주었고 이는 초해상도의 연구 흐름이 딥러닝 기반으로 전환되는 기점이 됐다. 이후 SRCNN을 응용하여 네트워크의 깊이가 깊고 모델의 크기가 클수록 초해상도 결과가 개선될 수 있음을 제시한 VDSR(Very Deep Super Resolution)이 제안되었다[10]. 그러나 VDSR의 경우 모델의 규모로 인해 연산량의 문제가 발생할 수밖에 없었고 단순히 모델 규모를 키우는 것으로 해결하는 것이 아닌 구조를 개선하여 연산량 문제 및 품질 개선을 시도하려는 연구가 진행되었다. EDSR(Enhanced Deep Residual Networks SR)은 네트워크 내 절차를 간소화하면서도 품질을 개선하기 위해 배치 정규화 제거 및 계층 별 특징 출력 수 증가 등을 적용하였다[11]. 이후 제안된 RDN(Residual Dense Network)에서는 특징 추출 및 학습 절차의 강화를 위해 잔차 학습을 재귀적 구조로 구성하여 결과물의 품질을 향상시켰다[12]. IDN(Information Distillation Network)은 그룹 컨볼루션을 적용함과 동시에 짧은 경로의 특징과 긴 경로의 특징을 결합하여 모델 크기를 압축함으로써 효율적인 모델 구조를 제안하였다[13]. 딥러닝의 새로운 방향을 제시한 생성적 적대 신경망 GAN(Generative Adversarial Network)이 제안된 이후에는 정답이 없는 문제를 해결해야하는 초해상도 방법에서 학습된 데이터를 기반으로 손실 함수 기반 최적화를 통해 해결하려 했던 기존의 방법에서 생성자를 통해 초해상도 결과를 생성하고 그것을 판별자가 고해상도 영상과 구별되지 않는 수준까지 판별하여 최종 초해상도를 생성하도록 하는 방법들이 제안되었다[14,15]. GAN을 기반으로 하는 초해상도 방법에서는 더욱 현실적인 세부정보를 생성할 수 있지만 GAN의 구조적 특성 상 연산량이 많을 수밖에 없는 문제가 있다. 또한 생성자의 생성 결과와 판별자의 판단 기준에 따라 원 영상의 형태에 맞지 않는 인공물 또는 부자연스러운 질감이 생길 수 있는 문제가 있다.

3. 제안하는 방법

제안하는 적응형 가중치 잔차 블록을 적용한 다중 블록 구조 기반의 단일 영상 초해상도 기법은 Fig. 1과 같은 흐름도를 갖는다.

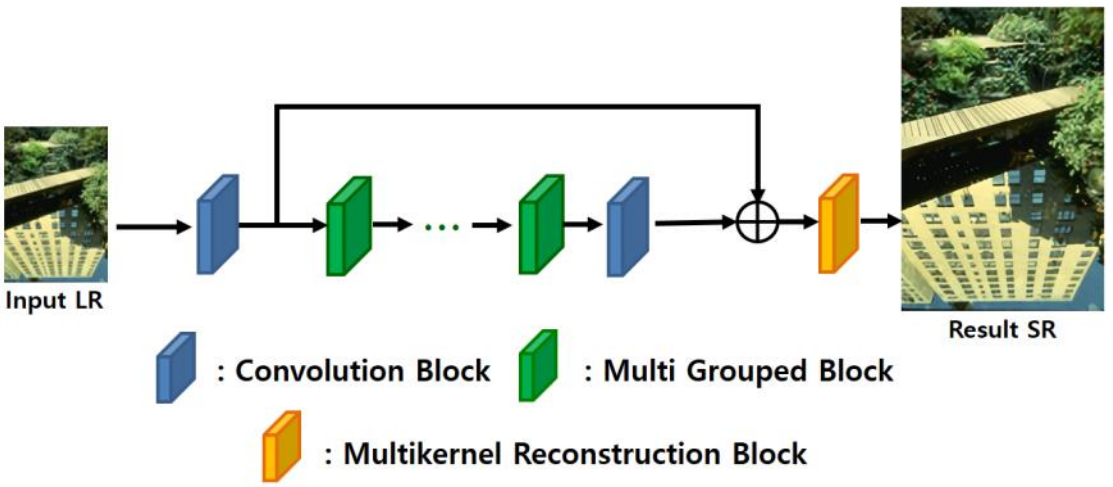

Fig. 1. Flowchart of proposed method

3.1 특징 추출

초해상도를 수행하기 위해 입력된 저해상도 영상 I_{LR} 으로부터 초기 저단계 특징을 추출하였다. 특징 추출을 위해 3x3 커널로 구성된 컨볼루션 계층$f_{ilf}()$을 사용하였다.

$$F_{init} = f_{ilf}(I_{LR}) \tag{1}$$

추출된 초기 저단계 특징 F_{init}으로부터 고단계 특징을 추출하기 위해 다중 블록 구조를 적용하였다. 다중 블록 구조의 내부는 n개의 잔여 블록과 1개의 Concat, 그리고 입력된 전단계 특징 F_{x-1}의 skip-connection이 있으며, 각각의 가중치를 적응적으로 적용할 수 있도록 $\alpha_{x_{rb}}$, $\alpha_{x_{sc}}$ 이 있고 이를 결합하는 구조로 구성되어 있다. 다중 블록 구조를 거칠 때 마다 점차 고단계 특징이 추출되고, 부가적인 특징 맵이 생성되어 세부사항 복원을 위한 용도로 활용된다. 수식 (2)에서 $Adap_f()$은 적응형 가중치 적용을 위한 함수이며 F_{RB_n}은 다중 블록 내 n번째 잔여 블록을 거쳐 skip-connection으로 획득된 결과를 의미한다.

$$F_x = Adap_f(Concat(F_{RB_1}, \cdots F_{RB_n}), \alpha_{x_{rb}}, F_{x-1}, \alpha_{x_{sc}}) \tag{2}$$

Fig. 2는 다중 블록 구조의 흐름도이다.

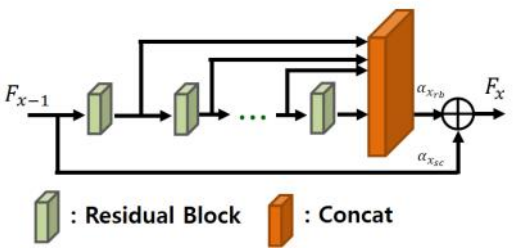

Fig. 2. Flowchart of Multi Grouped Block

이후 총 m개로 구성된 다중 블록을 거쳐 추출된 최종 특징 F_{MGB}를 3x3 컨볼루션을 이용해 서브픽셀 단위로 업스케일을 수행한 뒤 초기 저단계 특징 F_{init}의 skip-connection 결과와 결합하여 최종 특징 F_F를 생성한다.

$$F_F = conv3(F_{MGB}) \oplus F_{init} \tag{3}$$

3.2 다중커널을 이용한 영상 재구성

영상 재구성을 거쳐 초해상도를 생성할 때 기존 방법의 경우 고정된 하나의 커널 크기로 업스케일을 수행한 후 영상 재구성을 수행한다. 제안하는 방법은 생성된 최종 특징을 이용해 업스케일을 수행하기 전 다중 커널의 크기로 업스케일 변환 $REC_{K_n}()$을 수행한 뒤 각 결과물에 적응적 가중치를 반영하여 영상 재구성을 수행하

였다. 제안하는 방법에서는 3개의 커널 사이즈 3,5,7을 이용하였고 각각의 변환 결과에 적응적 가중치 계수 β_n을 적용하였다. 여기서 적응적 가중치 계수 β_n은 각각의 커널 크기에 따른 변환 결과를 비교하여 변화량으로 파악되는 화소 수를 분석하여 가중치를 적용하였다.

$$I_{SR} = Adap_{rec}(REC_{K_3}(F_F), \beta_3, REC_{K_5}(F_F), \beta_5, REC_{K_7}(F_F), \beta_7) \quad (4)$$

4. 실험 및 결과 고찰

본 논문의 실험 및 결과 고찰을 위해 초해상도 비교 평가 데이터셋으로 다수 활용되는 BSD100, Set5, Set14 데이터셋을 사용하여 실험을 수행하였다. 기존의 알고리즘인 VDSR, EDSR, RDN, IDN과 제안하는 방법을 시각적 및 정량적 비교 평가를 위해 각각의 평가 데이터셋으로 초해상도를 수행하였다. 초해상도 수행 방법은 원본 영상을 저해상도 크기로 축소하고 다시 원본 영상의 크기로 복원하는 방법을 이용하였다.

제안하는 방법의 비교 평가를 위해 정량적 비교를 수행하였다. 실험 방법은 기존 초해상도 방법과 제안하는 초해상도 방법을 각각 2, 3, 4배 초해상도를 수행하고, 초해상도 평가 지표로 활용되는 PSNR(Peak Signal to Noise Ratio), SSIM(Structural Similarity Index Measure)을 사용하였다.

Table 1. Quantative comparison (PSNR)

PSNR Comparison				
Model	Scale	BSD100	Set5	Set14
VDSR	2	31.90	37.53	33.03
EDSR		32.32	38.11	33.85
RDN		32.34	38.24	34.01
IDN		32.08	37.83	33.30
Proposed		32.42	38.32	34.10
VDSR	3	28.82	33.66	29.77
EDSR		29.25	34.65	30.44
RDN		29.26	34.71	30.57
IDN		28.95	34.11	29.99
Proposed		29.33	34.83	30.70
VDSR	4	27.29	31.35	28.01
EDSR		27.71	32.46	28.72
RDN		27.72	32.47	28.81
IDN		27.41	31.82	28.25
Proposed		27.80	32.63	28.92

Table 1은 PSNR 비교 결과이다. PSNR은 원본 영상 대비 화질 손실 정도를 값으로 표현한 지표로써 높은 값을 보일수록 원본에 가까운 것으로 보는 평가 지표이다. 제안하는 방법의 PSNR 결과가 모든 데이터셋에서 모든 배율의 초해상도를 수행했을 때 기존 방법 대비 우수함을 확인할 수 있었다.

Table 2는 SSIM 비교 결과이다. SSIM은 영상 내 구조적 유사성을 평가하는 방식으로 PSNR 평가지표 대비 사람이 시각적으로 인식하는 형태를 평가하려는 지표로써 높은 값을 보일수록 원본에 가까운 것으로 보는 평가 지표이다. SSIM 평가에서도 제안하는 방법이 모든 데이터셋과 모든 배율의 초해상도에서도 기존 방법들 대비 향상된 결과를 보여주었다.

Table 2. Quantative comparison (SSIM)

SSIM Comparison				
Model	Scale	BSD100	Set5	Set14
VDSR	2	0.896	0.959	0.912
EDSR		0.901	0.960	0.920
RDN		0.902	0.961	0.921
IDN		0.898	0.960	0.915
Proposed		0.903	0.962	0.922
VDSR	3	0.798	0.921	0.831
EDSR		0.809	0.928	0.846
RDN		0.809	0.930	0.847
IDN		0.801	0.925	0.835
Proposed		0.811	0.931	0.849
VDSR	4	0.725	0.884	0.767
EDSR		0.742	0.897	0.788
RDN		0.742	0.899	0.787
IDN		0.729	0.890	0.773
Proposed		0.745	0.901	0.790

제안하는 방법의 시각적 개선을 비교하기 위해 기존의 방법들로 수행한 초해상도 영상과 제안하는 방법의 초해상도 결과의 일부분을 확대하여 비교하였다. 시각적 비교를 위해 고배율의 초해상도인 4배를 수행하였다.

Fig, 3은 영상 내에 반복적인 구조물의 형태가 존재하고 각 구조물의 크기가 크지 않아 고배율의 초해상도를 수행했을 때 명확한 형태 복원을 예상하기 힘든 유형이다. 기존의 초해상도 방법인 VDSR 에서는 창문 사이의 경계를 온전히 보존하지 못하고 흐려진 형태의 결과를 볼 수 있다. EDSR의 경우 창문의 간격과 형태를 일부 유지하였지만 흐려진 결과를 볼 수 있다. RDN은 창문의 형태와 간격을 유지하고 흐려짐이 적은 결과를 볼 수 있지만 각 형태의 구분 과정에서 평활화된 결과로 세부 질감이 복원되지 않은 결과를 볼 수 있다. IDN은 흐

려짐 결과를 최소화하려 했지만 형태가 불완전하게 복원됨을 확인하였다. 제안하는 방법은 창문 사이의 간격이 일정하고 건물 외벽의 질감을 일부 복원한 것을 확인할 수 있다.

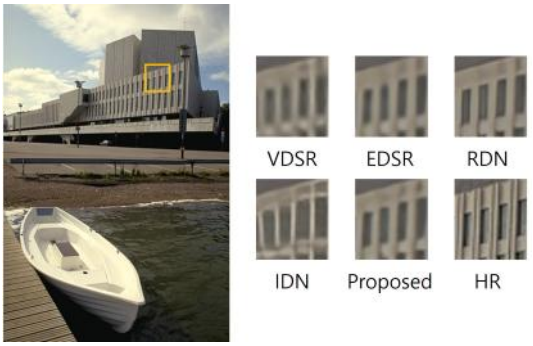

Fig. 3. Experimental result (78000 image in BSD100, scale 4)

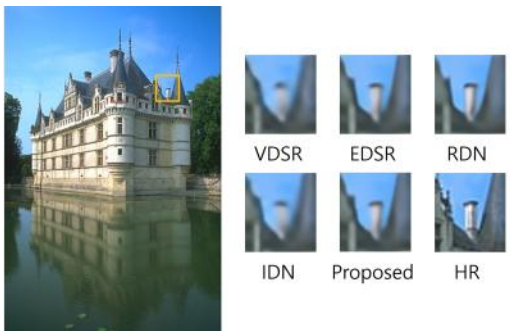

Fig. 4. Experimental result (102061 image in BSD100, scale 4)

Fig, 4는 작은 구조물에 대해 고배율의 초해상도를 수행했을 때 명확한 구분 및 복원을 수행하였는지 확인할 수 있는 유형이다. VDSR의 경우 전체적으로 흐려짐 상태로 복원되었으며 구조물의 형태가 원활히 복원되지 않음을 확인할 수 있다. EDSR은 일부 형태가 복원되었지만 흐려짐이 존재하는 것을 확인할 수 있고, RDN은 비교적 형태를 명확히 복원한 것으로 보이나 평활화된 처리 결과로 인해 각 구조물의 질감 구분이 명확하지 않은 형태로 복원되었다. IDN 또한 VDSR과 유사하게 명확한 형태 복원이 이루어지지 않고 흐려짐을 포함한 결과를 보였다. 제안하는 방법에서는 구조물의 복원 및 각 구조물의 질감 및 색상에 대해 비교적 구분이 된 것을 확인할 수 있다.

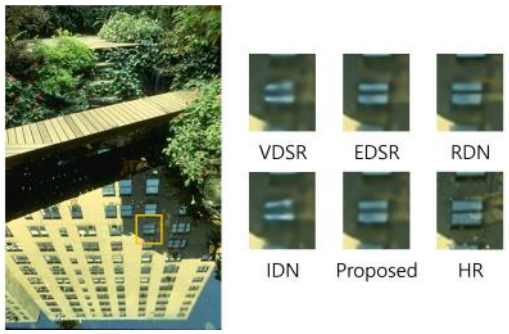

Fig. 5. Experimental result (148026 image in BSD100, scale 4)

Fig, 5는 구조물의 크기가 크지 않고 색 대비가 적으면서 좁은 영역이 존재할 때 고배율의 초해상도를 수행했을 때의 결과이다. VDSR과 IDN의 경우 창문의 형태와 창문 틀 경계를 원활히 복원하지 못하고 흐려짐 결과를 동반한 것을 확인할 수 있다. EDSR은 창문 형태를 복원한 것으로 보이나 흐려짐 형태와 창문 틀 경계가 확장된 결과를 보인다. RDN은 창문의 크기가 일부 변형되었으며, 세부사항이 명확하게 복원되지 않은 결과를 보인다. 제안하는 방법에서는 창문의 형태를 비교적 명확히 복원하였고, 창문의 틀 크기 또한 고배율의 초해상도임에도 원본과 유사한 형태를 보였으며, 질감과 같은 세부사항이 개선된 형태로 복원되었음을 확인할 수 있다.

5. 결론

본 논문에서는 적응형 가중치 잔차 블록을 적용한 다중 블록 구조를 이용해 초해상도 생성 과정에서의 효율적인 특징 추출과 적용 과정에서의 개선을 통해 품질을 향상시키는 방법을 제안하였다. 기존의 딥러닝 기반 초해상도에서 품질 향상을 위해 모델의 규모를 키우거나 네트워크 구조를 복잡하게 하는 등의 방법은 연산량이 다수 늘어나면서 세부정보가 약화되는 등의 개선 방향의 한계가 존재한다. 제안하는 방법은 잔차 블록의 구성을 변경하여 적응적 가중치를 적용하였고 이를 갖는 그룹된 블록 구조를 다중으로 배치함으로써 특징 추출을 효율적으로 하였다. 또한 목표 고해상도로 복원하는 과정에서 다중 커널 변환 및 가중 적용 방식을 활용함으로써 모델의 규모가 커지지 않으면서 품질이 향상된 결과

를 얻을 수 있었다.

REFERENCES

[1] Y. Zhang, Y. Huang, K. Wang, G. Qi & J. Zhu. (2023). Single image super-resolution reconstruction with preservation of structure and texture details. *Mathematics, 11(1)*, 216.
DOI : 10.3390/math11010216

[2] H. Chen et al. (2022). Real-world single image super-resolution: A brief review. *Information Fusion, 79*, 124-145.
DOI : 10.1016/j.inffus.2021.09.005

[3] H. Al-Mekhlafi & S. Liu. (2024). Single image super-resolution: a comprehensive review and recent insight. *Frontiers of Computer Science, 18(1),* 181702.
DOI : 10.1007/s11704-023-2588-9

[4] K. P. Gunasekaran. (2023, May). *Ultra sharp: Study of single image super resolution using residual dense network.* In 2023 IEEE 3rd International Conference on Computer Communication and Artificial Intelligence (CCAI) (pp. 261-266). IEEE.
DOI : 10.1109/CCAI57533.2023.10201303

[5] J. B. Huang, A., Singh & N. Ahuja. (2015). *Single image super-resolution from transformed self-exemplars.* In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5197-5206).
DOI : 10.1109/CVPR.2015.7299156

[6] J. Yang, J. Wright, T. S. Huang & Y. Ma. (2010). Image super-resolution via sparse representation. *IEEE transactions on image processing, 19(11),* 2861-2873.
DOI : 10.1109/TIP.2010.2050625

[7] B. Liu & D. Ait-Boudaoud. (2020). Effective image super resolution via hierarchical convolutional neural network. *Neurocomputing, 374,* 109-116.
DOI : 10.1016/j.neucom.2019.09.035

[8] S. M. A. Bashir, Y. Wang, M. Khan & Y. Niu. (2021). A comprehensive review of deep learning-based single image super-resolution. *PeerJ Computer Science, 7,* e621.
DOI : 10.7717/peerj-cs.621/retraction

[9] C. Dong, C. C. Loy, K. He & X. Tang. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence, 38(2),* 295-307.
DOI : 10.1109/TPAMI.2015.2439281

[10] J. Kim, J. K. Lee & K. M. Lee. (2016). Accurate image super-resolution using very deep convolutional networks. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646-1654).
DOI : 10.1109/CVPR.2016.182

[11] B. Lim, S. Son, H. Kim, S. Nah & K. Mu Lee. (2017). *Enhanced deep residual networks for single image super-resolution.* In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 136-144).
DOI : 10.48550/arXiv.1707.02921

[12] Y. Zhang, Y. Tian, Y. Kong, B. Zhong & Y. Fu. (2018). *Residual dense network for image super-resolution.* In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2472-2481).
DOI : 10.48550/arXiv.1802.08797

[13] Z. Hui, X. Wang & X. Gao. (2018). *Fast and accurate single image super-resolution via information distillation network.* In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 723-731).
DOI : 10.48550/arXiv.1803.09454

[14] G. Wu, J. Jiang & X. Liu. (2023). *A practical contrastive learning framework for single-image super-resolution. I*EEE Transactions on Neural Networks and Learning Systems.
DOI : 10.1109/TNNLS.2023.3290038

[15] K. S. Reddy, V. P. Vijayan, A. D. Gupta, P. Singh, R. G. Vidhya & D. Kapila. (2022, April). *Implementation of super resolution in images based on generative Adversarial network.* In 2022 8th International Conference on Smart Structures and Systems (ICSSS) (pp. 01-07). IEEE.
DOI: 10.1109/ICSSS54381.2022.9782170

한 현 호(Hyun Ho Han) [종신회원]

· 2009년 2월 : 광운대학교 컴퓨터공학과 (공학사)
· 2011년 2월 : 광운대학교 정보콘텐츠 대학원 유비쿼터스컴퓨팅학과 (공학석사)
· 2019년 8월 : 광운대학교 정보디스플레이학과(공학박사)
· 2020년 3월 ~ 현재 : 울산대학교 교양학부 조교수
· 관심분야 : 무선 네트워크, 영상인식, 3D 영상처리, 기계 학습, 딥러닝
· E-Mail : hhhan@ulsan.ac.kr