

경량 트랜스포머에 기반한 흑백 항공영상의 초해상화

정태준*, 송용현**, 엄대용**

*국립한국교통대학교 대학원 토목공학과

**국립한국교통대학교 사회기반공학전공

e-mail:dyum@ut.ac.kr

Super-Resolution for Grayscale Aerial Imagery based on Lightweight Transformer

Tae-Jun Jeong*, Yong-Hyun Song**, Dae-Yong Um**

*Dept. of Civil Engineering, Graduate School, Korea National University of
Transportation

**Dept. of Civil Engineering, Korea National University of Transportation

요 약

흑백 항공사진은 장기간 지형 변화와 지적 현황 분석을 위한 핵심 자료이나, 1966-1974년 한-네덜란드 협정 촬영분은 약 50cm의 낮은 해상도로 정밀 분석에 한계가 있다. 본 연구는 경량 트랜스포머 기반 초해상화 모델인 Omni-SR을 적용하여 저해상도 흑백 항공영상의 시각적 품질을 향상하는 데 목적이 있다. 평가 결과, 제안 모델은 PSNR 28.9108 dB, SSIM 0.6732로 Bicubic(26.5205 dB, 0.6144) 대비 유의미한 개선효과를 확인할 수 있었다. 이는 적은 컴퓨팅 자원 환경에서도 기존 보간법보다 뛰어난 경량 학습 기반 접근의 실효성을 보여준다.

1. 서론

항공촬영 영상은 지도 제작, 지리정보시스템(GIS), 국방 감시 등 다양한 분야에서 핵심적인 기초 자료로 활용되고 있다. 이 가운데 1966년부터 1974년까지 한-네덜란드 협정에 따라 촬영된 과거의 흑백 항공사진은 장기간의 지형 변화 및 지적 현황 등을 분석할 수 있는 유일한 역사적 영상 자료이다. 그러나 이들 영상은 약 50~60cm 수준의 낮은 해상도를 지니고 있어, 현재 제공되는 25cm 해상도의 디지털 항공사진과 비교할 때 정밀한 분석 수행에 한계가 존재한다.

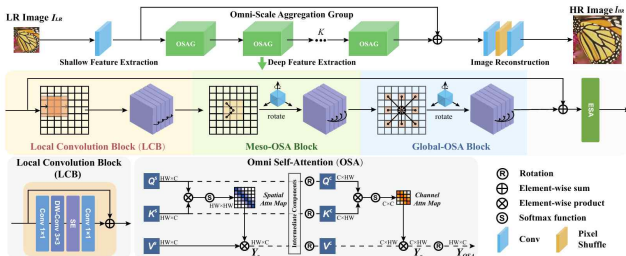
기존에는 영상의 해상도 향상을 위하여 Bilinear, Bicubic 등과 같은 보간 기법이 주로 활용되었다. 그러나 이러한 기법들은 픽셀 값을 평균화하는 과정에서 세밀한 질감 및 경계선 등 영상 고유의 고주파 성분을 손실시켜 결과적으로 이미지의 선명도를 저해하는 근본적인 제약을 내포하고 있다. 이러한 한계를 극복하고자, 2015년 최초의 딥러닝 기반 초해상화 모델인 SRCNN이 등장한 이후 원격탐사 영상의 해상도 및 시각적 품질을 비약적으로 개선하고 있다. 본 연구는 딥러닝 초해상화 모델 중 경량 트랜스포머 모델인 Omni-SR을 활용하여 성능을 검토하는 것을 목적으로 한다.

2. 연구 방법 및 분석 결과

2.1 연구 모델

Omni-SR (Omni Aggregation Networks for Lightweight Image Super-Resolution)은 Wang 등(2023)이 제안한 경량급 비전 트랜스포머(ViT) 기반의 초해상화 모델이다. 이 모델은 제한된 파라미터 내에서 성능을 극대화하기 위해 기존 경량 모델들이 가진 단방향적 어텐션(attention) 방식과 동차적(homogeneous) 정보 통합 구조의 한계를 해결하는 데 초점을 맞추었다. Omni-SR의 핵심은 Omni Self-Attention (OSA) 블록이다. OSA는 기존 어텐션 메커니즘이 공간(spatial) 또는 채널(channel) 차원 중 하나에만 집중했던 것과 달리, 두 차원의 픽셀 상호작용을 동시에 모델링하여 '모든 축(omni-axis)'에 걸친 잠재적 상관관계를 탐색한다. 이는 공간적 관계와 채널 간 정보 전송을 통합적으로 처리함으로써, 경량 모델에서도 더 풍부한 맥락 정보를 효율적으로 집계할 수 있게 한다. 두 번째 주요 혁신은 Omni-Scale Aggregation Group (OSAG) 이다. 기존 모델들이 단일 연산자를 반복적으로 쌓아 올려 유효 수용 영역(Effective Receptive Field) 확장에 비효율적이었던 문제를 해결하기 위해, OSAG는 세 가지 단계의 집계 방식을 통합했다. 구

체적으로, 지역적 패턴을 위한 로컬 컨볼루션 블록, 중간 규모의 상호작용을 위한 Meso-OSA 블록, 그리고 전역적 맥락 이해를 위한 Global-OSA 블록을 순차적으로 구성하여 다양한 규모의 텍스처 정보를 효과적으로 추출한다. 이와 관련한 프레임워크 구조는 그림 1과 같다.



[그림 1] Omni-SR 프레임워크의 구조

2.2 연구 과정

실험에 사용된 데이터는 2023년에 촬영된 고해상도 컬러(RGB) 항공사진을 활용하였다. 항공사진은 모델 학습에 앞서 다음과 같은 전처리 과정을 거쳤다. 먼저, 컬러 이미지를 흑백(gray-scale) 영상으로 변환하였으며 전처리된 영상은 GDAL(Geospatial Data Abstraction Library)을 이용하여 256×256 픽셀 크기의 패치(patch) 단위로 분할하여 데이터셋을 구축하였다.

원본 해상도(256×256)의 패치는 Ground Truth 영상으로 사용하였고, Low-Quality 입력 영상은 GT 영상에 2배 축소하여 128×128 의 패치로 제작한 후 Bicubic 보간 및 가우시안 노이즈(gaussian noise)를 추가하는 방식으로 생성하였다. 이는 의도적으로 화질을 저하시켜, 모델이 노이즈 제거 및 복원 능력을 학습하도록 하기 위함이다.

구축된 전체 데이터셋은 학습용 69,237장과 검증용 17,640장으로 분할하여 사용하였다. 또한, 모델의 일반화 성능을 높이기 위해 데이터 증강(data augmentation) 기법을 적용하였다.

모델 학습은 PyTorch 딥러닝 프레임워크를 기반으로 수행되었으며, 학습 가속화를 위해 NVIDIA GPU를 활용하였다. 손실 함수(loss function)는 생성된 영상과 GT 영상 간의 차이를 측정하기 위해 L1 손실을 사용하였다. 최적화 알고리즘으로는 AdamW를 채택하였다.

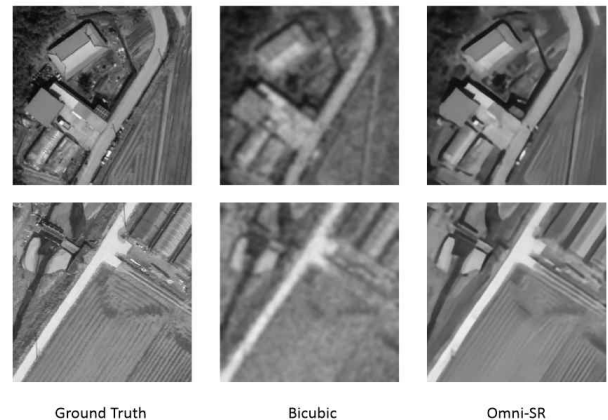
학습 과정에서 과적합(overfitting)을 방지하고 최적의 모델을 선택하기 위해 Early Stopping 기법을 적용하였다. 이 기법은 검증 세트(validation set)의 성능을 지속적으로 모니터링하여, 성능 개선이 더 이상 이루어지지 않거나 오히려 감소할 경우 학습을 조기에 중단하는 방식이다. 이를 통해 불필요한 추가 학습을 방지하고 모델의 일반화 성능을 극대화하였다.

2.3 분석 결과

표 1의 참조 기반 평가 결과, 본 연구에서 제안하는 경량 트랜스포머 모델은 PSNR 28.9108 dB, SSIM 0.6732를 달성하여 기존의 Bicubic 보간법 대비 뚜렷한 성능 향상을 보였다. 높은 PSNR 수치는 원본 대비 재구성된 영상의 평균제곱오차(RMSE)가 유의미하게 감소했음을 나타내며, SSIM 값의 향상은 영상의 경계와 질감 같은 주요 구조적 정보가 효과적으로 보존되었음을 시사한다. 그림 2를 보면 이러한 성능적 우위는 트랜스포머 모델이 영상 내의 전역적(global) 및 국소적(local) 상관관계를 동시에 학습하여 고주파 성분을 정밀하게 복원하고, 블러링(blurring)이나 링잉(ringing) 현상과 같은 시각적 왜곡을 효과적으로 억제하기 때문으로 분석된다.

[표 1] 초해상화 품질 평가

Division	PSNR	SSIM
Omni-SR	28.9108	0.6732
Bicubic	26.5205	0.6144



[그림 2] 초해상화 이미지 비교

3. 결론

본 연구는 경량 트랜스포머 기반 초해상화 모델인 Omni-SR을 흑백 항공영상에 적용하고, 전통적 보간법인 Bicubic과 정량적으로 비교하였다. 그 결과, Omni-SR은 Bicubic을 일관되게 상회하는 성능을 보였으며, 이는 제한된 연산 자원 환경에서도 기존 보간법을 대체할 수 있는 경량 학습 기반 접근의 실효성을 보여준다. 다만 본 연구는 $\times 2$ 배율 실험에 국한되었고 다양한 경량 초해상화 모델과의 직접 비교가 이루어지지 않았다는 한계를 지닌다. 향후 연구에서는 $\times 4$ 이상 고배율 확대에 대한 일반화 성능과 안정성 검증, 최신 경량 모델들과의 체계적 비교, 그리고 추론 시간-메모리 사용량을 포함한 성능분석이 필요하다.

감사의 글

본 논문은 2025년 국립한국교통대학교의 지원을 받아 연구되었음.

참고문헌

- [1] H. Wang, X. Chen, B. Ni, Y. Liu and J. Liu, “Omni Aggregation Networks for Lightweight Image Super-Resolution”, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Vancouver, Canada, pp. 22378-22387, June 2023.
- [2] H. D. Seo, H. Y. So and E. M. Kim, “Quantitative Evaluation of Super-resolution Drone Images Generated Using Deep Learning”, Journal of Cadastre & Land InformatiX, Vol.53, No.2, pp. 5-18, 2023.
- [3] J. T. An and H. C. Yun, “Analysis of the Impact of Downsampling Methods and Deep Learning Models for Single Image Super-resolution”, Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, Vol.43, No.3, pp. 397-408, 2025
- [4] C. Dong, C. C. Loy, K. He and X. Tang, “Image Super-Resolution Using Deep Convolutional Networks”, European Conference on Computer Vision (ECCV), pp.184-199, 2014.
- [5] S. W. Kim and H. H. Park, “Super-Resolution Using NLSA Mechanism”, Journal of Korea Institute of Convergence Signal Processing, Vol.23, No.1, pp. 8-14, 2022.
- [6] J. W. Lee, S. W. Yoon and K. C. Lee, “Transformer-Based Deep Learning Models” for ERS SAR Image Super-Resolution“, Korean Journal of Remote Sensing, Vol.41, No.1, pp.143-152, 2025.
- [7] J. Kim, J. K. Lee and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks”, In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, USA, June 26-July 1, pp.1646-1654, 2016.