



레터논문 (Letter Paper)

방송공학회논문지 제29권 제2호, 2024년 3월 (JBE Vol.29, No.2, March 2024)

<https://doi.org/10.5909/JBE.2024.29.2.165>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## Super Resolution을 활용한 위성 영상의 표적 탐지 성능 분석

한승석<sup>a)</sup>, 이가은<sup>a)</sup>, 최해철<sup>a)†</sup>

### Analyzing Target Detection Performance in Satellite Images Using Super Resolution Techniques

Seungseok Han<sup>a)</sup>, Gaeun Lee<sup>a)</sup>, and Haechul Choi<sup>a)†</sup>

#### 요약

저용량 군집위성으로 획득한 영상은 중궤도 혹은 정지궤도 위성에 비해 화질 및 해상도가 낮기 때문에 감시 및 정찰의 효율성이 떨어진다. 본 논문은 저화질 위성영상에 초고해상화(Super-Resolution, SR) 기술을 적용하여 화질을 개선하며, 개선된 화질의 위성 영상으로 객체를 탐지할 때 그 성능을 분석한다. 실험에서는 위성영상에 대한 다양한 SR 신경망의 성능을 분석하고, SR 신경망과 탐지 신경망을 연동함으로써 SR을 사용하지 않은 경우와 비교하여 탐지 성능이 향상됨을 보인다. 제안 방법은 저화질의 위성영상에서 감시 및 정찰의 효율성을 극대화할 수 있을 것으로 기대한다.

#### Abstract

The images obtained from low-altitude cluster satellites suffers from lower quality and resolution compared to those acquired from satellites positioned in medium or geostationary orbits, leading to reduced efficiency in surveillance and reconnaissance. This paper proposes a method that enhances the quality of low-resolution satellite imagery by employing Super-Resolution (SR) technology and subsequently improves detection performance by identifying targets within the enhanced imagery. The experiments analyze the performance of various SR neural networks concerning satellite images and demonstrated that by integrating various SR neural networks with the prominent detection neural network, Rotated RetinaNet, it is possible to enhance detection performance. This combination of SR neural network and established detection framework leads to improvements in the overall accuracy and reliability of surveillance and reconnaissance.

Keyword : Super Resolution, Detection, Satellite Image, Neural Network

a) 국립한밭대학교 지능미디어공학과(Department of Intelligent Media Engineering, Hanbat National University)

† Corresponding Author : 최해철(Haechul Choi)

E-mail: [choihc@hanbat.ac.kr](mailto:choihc@hanbat.ac.kr)

Tel: +82-42-821-1149

ORCID: <https://orcid.org/0000-0002-7594-0828>

※ This work was supported by Korea Research Institute for defense Technology planning and advancement(KRIT) grant funded by the Korea government (DAPA(Defense Acquisition Program Administration)) (21-106-A00-007, Space-Layer Intelligent Communication Network Laboratory, 2022)

· Manuscript 10 January, 2024; Revised 8 February, 2024; Accepted 13 February, 2024.

## I. 서론

일반적으로 지구 관측 혹은 첩보위성은 지구 표면을 관찰하기 위해 고궤도에 위치한다. 하지만 고궤도의 위성은 고비용을 요구하므로 근래 저비용의 저궤도 저용량 군집위성에 대한 연구가 활발하다. 그러나 저용량 군집위성으로 획득한 영상은 중궤도 혹은 정지궤도 위성의 영상에 비해 화질 및 해상도가 낮고 국부적 영역의 부만 제공하기 때문에 높은 감시 및 정찰의 효율성을 갖기 어렵다. 이러한 문제를 보완하기 위해, 본 논문은 저궤도 군집위성에서 획득한 저화질 영상에 초고해상화(Super-Resolution, SR) 기술을 적용하여 화질을 개선시키고, 개선된 화질의 영상에서 표적을 탐지함으로써 탐지 성능을 향상하고자 한다. 실험에서는 최근 SR 분야에서 탁월한 성능을 보이는 다양한 신경망의 성능을 분석하고, 대표적 검출 신경망으로 각 SR 신경망을 적용한 초고해상화된 위성 영상에 대한 탐지 성능을 보임으로써, 제안하는 SR 신경망과 검출 신경망의 연동 프레임워크의 효율성을 제시한다.

## II. 관련 연구

SR 분야의 기존 연구에는 Single-Image Super-Resolution (SISR)<sup>[1]</sup>, Reference-based Super-Resolution<sup>[2]</sup>, Multi-Image Super-Resolution<sup>[3]</sup>이 있다. SISR 모델에는 ESRGAN<sup>[4]</sup>, SwinIR<sup>[5]</sup>과 state-of-the-art (SOTA) 모델인 HAT-L<sup>[6]</sup> 등이 연구되었다. ESRGAN은 배치 정규화가 없는 Residual-in-Residual Dense Block을 가진 구조로 되어 있으며, relativistic GAN을 사용하여 판별자가 절대값이 아닌 상대적인 실제성을 예측하도록 하는 모델이다. SwinIR은 Swin Transformer를 기반으로 한 모델이다. 얇은 특성 추출, 깊은 특성 추출 및 고품질 이미지 재구성으로 구성되며, 깊은 특성 추출 모듈은 여러 개의 잔여 Swin Transformer 블록으로 구성되며, 각 블록에는 잔여 연결을 포함한 여러 개의 Swin Transformer 레이어가 존재한다. HAT-L은 채널 어텐션과 윈도우 기반 셀프 어텐션 스킵을 결합하여 전역 통계 활용과 강력한 로컬 맞춤 능력의 상호 보완적인 장점을 활용한

다. 또한 cross-window 정보를 더 잘 집계하고 이웃 윈도우 피처 간 상호 작용을 강화하기 위해 중첩된 cross-window 어텐션 모듈을 도입한 모델이다.

Detection 모델로는 mmrotate의 Rotated RetinaNet<sup>[7]</sup> 모델이 연구되었다. 이 모델은 위성영상을 기반으로 훈련이 되어있다. 위성 영상은 지상에서 촬영한 영상과 달리 객체가 다양한 각도로 회전되어 나타날 수 있으며, Rotated RetinaNet 모델은 이러한 객체의 회전이나 변형에 매우 강한 성능을 제공한다.

## III. 초고해상화를 통한 탐지 성능 향상

화질과 해상도가 낮은 저궤도 위성 영상에서의 표적 탐지는 고화질 고궤도 위성 영상에 비해 정확도에 한계가 있다. 이는 주로 세부 정보의 부족으로 인해 발생하는 문제이다. 고해상도의 위성 영상에서는 표적의 세부 사항과 윤곽이 뚜렷하게 보여 효과적으로 탐지할 수 있지만, 저해상도 영상에서는 이러한 세부 정보가 누락되거나 흐릿해져 객체를 정확하게 식별하기 어려워진다.

제안 방법은 이러한 문제의 해결책으로 그림 1과 같이 Super-Resolution (SR) 기술을 적용하여 저해상도 (Low Resolution, LR) 영상의 화질을 개선하여 객체의 세부 정보를 복원하는 전처리 과정을 갖는다. 이 전처리 과정을 통해 획득된 고화질 영상을 기존 탐지 신경망에 입력하여 표적 탐지 성능을 향상하고자 한다. 즉, 기존의 저해상도 영상에서 감지된 것보다 더 우수한 성능을 달성하기 위해 SR 방법과 탐지 방법을 연동하는 것이 제안 프레임워크이다.

제안 방법은 실험을 통해 기존의 다양한 SR 신경망에 대해 위성 영상에 대한 초고해상화 성능을 우선 비교하고, 각 SR 신경망으로 획득된 고해상도 위성 영상을 동일한 탐지 신경망에 입력했을 때 그 탐지 신경망의 성능 변화를 분석한다. SR 신경망으로는 높은 PSNR 성능을 보이는 대표적인 ESRGAN, SwinIR, HAT-L을 활용하였으며, 탐지 신경망은 객체의 방향 변화에 강한 Rotated RetinaNet을 기준으로 삼는다. 결론적으로 SR 신경망과 탐지 신경망을 연동 시킴으로써 위성 영상의 탐지 성능 향상 정도를 보인다.

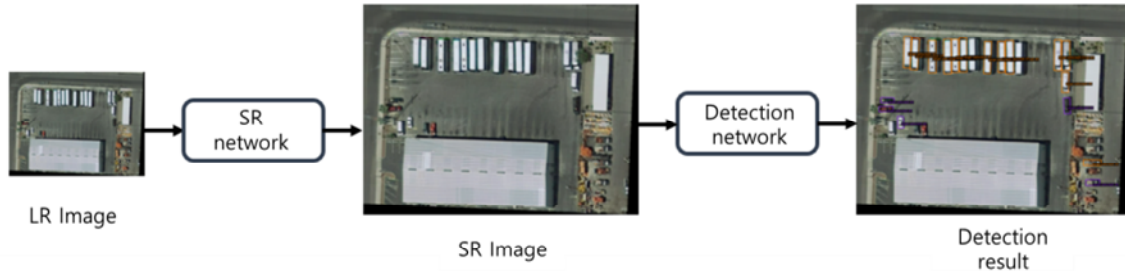


그림 1. 초고해상화를 전처리로 활용한 SR 및 표적 탐지 연동 프레임워크

Fig. 1. Integrated framework for Super-Resolution and Target Detection using preprocessing with super-resolution imaging

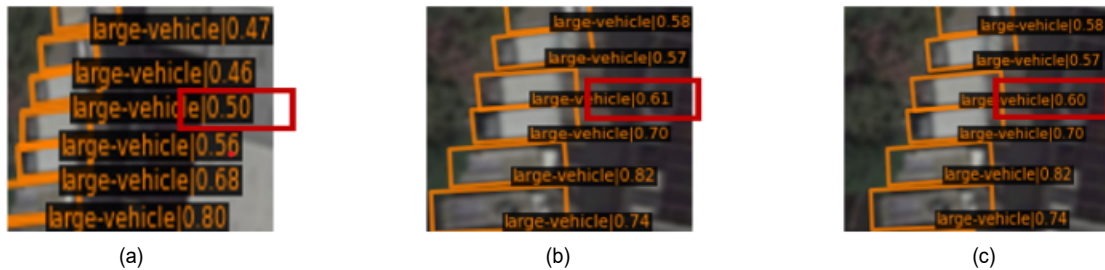


그림 2. 탐지 결과: (a) 저화질 영상, (b) SwinIR을 적용하여 2배 초고해상화한 영상, (c) HAT-L을 적용하여 2배 초고해상화한 영상  
 Fig. 2. Detection Results: (a) Low-resolution image, (b) 2x SR image using SwinIR, (c) 2x SR image using HAT-L

#### IV. 실험 결과

SR 모델 중 Reference-based Super-Resolution, Multi-Image Super-Resolution은 참조 영상과 동일 지역의 복수 영상에 대한 데이터셋을 추가적으로 요구하기에 일반적으로 많이 활용되는 SISR 모델을 기반으로 본 실험을 수행하였다. SISR을 위한 신경망 중에서 높은 SR 성능을 보이는 대표적인 ESRGAN, SwinIR과 SOTA 모델인 HAT-L을 사용하였다. 탐지 신경망은 객체의 방향 변화가 많은 위성 영상에 강인한 Rotated RetinaNet을 이용하였다. 실험에서는 제안하는 방법의 탐지 정확도를 평가하기 위해 널리 활용되는 mean Average Precision (mAP)과 SR 성능을 평가하기 위한 Peak Signal-to-noise ratio (PSNR)을 사용하였다. 데이터셋은 위성 영상으로 구성된 DOTA Dataset<sup>[8]</sup> 중  $1000 \times 1000$  이하의 56장을 이용하였다.

표 1에서는 저화질 위성영상, 공간해상도를 가로와 세로 모두 2배 보간할 때의 SwinIR, HAT-L의 성능을 보여주고 있다. PSNR 결과는 SwinIR과 HAT-L이 비슷하게 나타났다. 탐지 성능인 mAP 결과는 저화질의 위성영상을 Rotated

RetinaNet에 입력한 경우가 0.56으로 제일 낮았고, SwinIR이 0.58로 가장 높은 성능을 보였다. 그림 2를 참고하면 SwinIR의 결과가 저화질의 위성영상과 HAT-L의 결과에 비해 다소 높은 정확도를 보임을 알 수 있다. 표 2에서는 저화질 위성영상의 공간해상도를 4배 확장한 ESRGAN, SwinIR, HAT-L의 성능을 보여주고 있다. PSNR 결과는 SR 모델 중 HAT-L이 가장 높은 성능을 보였다. 그러나 mAP 결과는 ESRGAN이 0.65로 가장 높게 나타났다. 이때 저화질의 위성영상은 mAP가 0.5이기에 SR을 통해 0.15의 탐지 성능이 향상됨을 확인할 수 있다. 그림 3을 참고하면 탐지된 객체에 대해서 탐지 신뢰도는 HAT-L가 높은 경우가 많으나 이는 다른 객체로 오탐지하는 경우가 있었다. 종합적으로는 ESRGAN이 높은 mAP 정확도를 나타냈다. PSNR 측면에서는 ESRGAN보다 HAT-L이 가장 좋은 성능을 나타내는 것을 보아 이는 인지적 화질 측면에서의 성능과 탐지 과업에서의 성능이 절대적으로 비례하지는 않음을 확인할 수 있다. 결론적으로 본 실험은 특히 객체의 크기가 작은 위성 영상에서는 2배 혹은 4배 초고해상화한 영상으로 탐지하는 경우 탐지 성능이 향상될 수 있음을 증명한다.

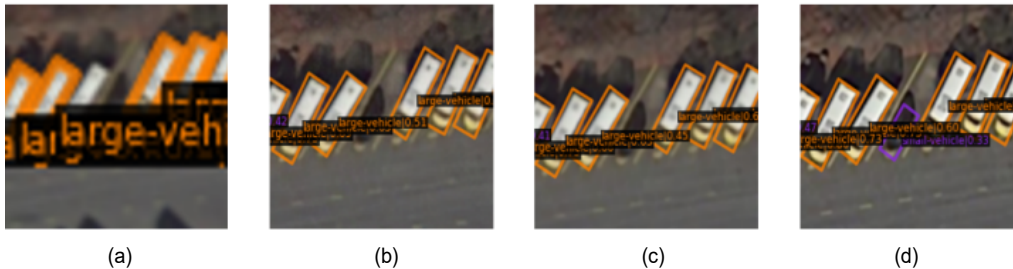


그림 3. 탐지 결과: (a) 저화질 영상, (b) ESRGAN을 적용하여 4배 초고해상화한 영상, (c) SwinIR을 적용하여 4배 초고해상화한 영상, (d) HAT-L을 적용하여 4배 초고해상화한 영상  
 Fig. 3. Detection Results: (a) Low-resolution image, (b) 4x SR image using ESRGAN, (c) 4x SR image using SwinIR, (d) 4x SR image using HAT-L

표 1. 저화질 영상과 이의 2배 초고해상화한 영상에 대한 초고해상화 성능 (PSNR) 및 탐지 성능(mAP)

Table 1. SR Performance (PSNR) and Detection Performance (mAP) for Low-Resolution Images and its 2x Super-Resolution Images

x2	PSNR	mAP
LR image	-	0.56
SwinIR	25.98	0.58
HAT-L	25.98	0.57

표 2. 저화질 영상과 이의 4배 초고해상화한 영상에 대한 초고해상화 성능 (PSNR) 및 탐지 성능(mAP)

Table 2. SR Performance (PSNR) and Detection Performance (mAP) for Low-Resolution Images and its 4x Super-Resolution Images

x4	PSNR	mAP
LR image	-	0.50
ESRGAN	26.71	0.65
SwinIR	27.05	0.58
HAT-L	29.56	0.61

## V. 결론

본 논문에서는 저화질 위성영상에 대한 탐지 성능을 향상 시키는 것을 목적으로 한다. 이를 위해 저화질 위성영상에 SR 기술을 적용하여 화질을 개선시키고, 개선된 화질의 영상으로 표적을 탐지하는 방법을 제안하였다. 실험 결과, 저화질의 위성영상에서 표적을 탐지하는 것보다 SR 기술을 적용한 영상의 탐지 성능이 더욱 높은 것을 확인하였다. 또한 실험에서는 인지적 화질 측면에서의 SR 성능과 탐지 과업에서의 SR 성능이 비례하지 않음도 확인할 수 있었다. 향후 연구로써 탐지에 적합한 SR에 대한 연구를 추진할 계획이다.

## 참고 문헌 (References)

- [1] W. Yang, X. Zhang, Y. Tian, W. Wang, J. -H. Xue and Q.Liao, "Deep Learning for Single Image Super-Resolution: A Brief Review," in IEEE Transactions on Multimedia, vol.21, No.12, pp.3106-3121, Dec 2019. doi: <https://doi.org/10.1109/TMM.2019.2919431>
- [2] Cao, Jiezhong, et al. "Reference-based image super-resolution with deformable attention transformer," European conference on computer vision, Switzerland, pp.325-345, 2022. doi: [https://doi.org/10.1007/978-3-031-19797-0\\_19](https://doi.org/10.1007/978-3-031-19797-0_19)
- [3] M. Kawulok, P. Benecki, S. Piechaczek, K. Hrynczenko, D. Kostrzewa and J. Nalepa, "Deep Learning for Multiple-Image Super-Resolution," in IEEE Geoscience and Remote Sensing Letters, vol.17, no.6, pp.1062-1066, June 2020. doi: <https://doi.org/10.1109/LGRS.2019.2940483>
- [4] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao and C. -C. Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," Proceedings of the European conference on computer vision workshops, 2018. doi: <https://doi.org/10.48550/arXiv.1809.00219>
- [5] j. Liang, j. Cao, G. Sun, K. Zhang, L. -V. Gool and R. Timofte, "SwinirL Image restoration using swin transformer," Proceedings of the IEEE/CVF international conference on computer vision, pp.1833-1844, 2021. doi: <https://doi.org/10.48550/arXiv.2108.10257>
- [6] X. Chen, X. Wang, J. Zhou, Y. Qiao and C. Dong, "Activating more pixels in image super-resolution transformer," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp.22367-22377, 2023. doi: <https://doi.org/10.48550/arXiv.2205.04437>
- [7] T. -Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal loss for dense object detection," Proceedings of the IEEE international conference on computer vision, pp.2980-2988, 2017. doi: <https://doi.org/10.48550/arXiv.1708.02002>
- [8] J. Ding, N. Xue, G. -S. Xia, X. Bai, W. Yang, M. -Y. Yang, S. Belongie, J. Luo, M. Datcu, M. Pelillo and L. Zhang, "Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.44, No.11, pp.7778-7796, Nov 2022. doi: <https://doi.org/10.1109/TPAMI.2021.3117983>