

# 딥러닝 기반 위성영상 복원에 관한 연구

임지현, 이재열, 김민아, 정대원.

한국항공우주연구원

yimjh@kari.re.kr

## A Study on Satellite Image Restoration based on Deep Learning

Yim Ji Hyeon, Lee Jae yeol, Kim Min-A, Chung Dae Won.

Korea Aerospace Research Institute.

### 요약

영상에서 소실된 영역, 빈 영역이 존재할 경우 해당영역을 구조와 질감에 맞게 자연스럽게 채워 복원하는 영상 인페인팅(Image Inpainting)은 영상 처리 분야에서 특정 개체 제거, 소실 영역 복원 등 영상편집 목적으로 활용되고 있으며, 시각적으로 자연스럽게 얼마나 일관되게 재구성하였는지가 주요 평가 척도이다. 최근 딥러닝(Deep Learning) 기반의 영상 인페인팅 알고리즘이 초고해상도 영상 복원, 영상 복원의 처리 시간 단축, 정확도 개선에 좋은 성능을 보여주면서 대용량 영상 데이터인 위성영상 분야에서의 시도가 이루어지고 있다. 본 연구에서는 위성영상 Dataset을 인페인팅 알고리즘인 GLCIC(Globally and Locally Consistent Image Completion)모델에 적용하여 위성영상의 인페인팅 기술 적용 가능성과 향후 연구 방향성에 대해 검토하였다.

### I. 서론

영상처리 분야에서 왜곡된 영상을 처리하는 분야는 주요한 연구 중 하나로 과거부터 지속되어 왔다. 특히, 손상 및 공백이 발생한 영상을 복원하는 방법은 인페인팅(Inpainting)이라는 명칭으로 발전되어 왔다. 영상에서 특정 정보를 취득하거나 활용 시 해당 영역에 장애물(예, 워터마크 등)이 포함되어, 의도치 않은 손상이 발생할 경우 과거에는 수동적인 영상편집 방법으로 보정을 진행하였다. 그러나 영상의 활용도가 높아지고 고해상도 영상이 요구되는 현재에는 사람의 작업만으로는 한계점이 발생할 수밖에 없다.

이러한 단점을 해결하기 위해 초기에는 작고 단순한 구조의 영상의 복원에 적합한 텍스처 합성, 확산 기반 인페인팅 기술[1]을 사용하였고, 최근에는 딥러닝(Deep Learning)을 활용한 인페인팅 기술이 연구되고 있다[2]. 현재 사람의 수작업만큼 세부적인 디테일한 부분까지의 성능은 아니지만 충분히 활용이 가능한 수준까지 발전되어 왔다. 앞으로 딥러닝 기술이 고도화되고 대용량, 초고해상도 영상이 활용되는 분야(예, 위성영상 등)가 증대됨에 따라 영상 복원의 시도와 필요성이 증대될 것으로 예상된다[2].

위성영상의 경우 영상정보를 가공 및 분석하기 위해 영상 복원이 요구되는 경우가 발생하고 있다. 이러한 경우로는 먼저, 위성영상 생성 전 처리 과정에서 종종 발생하는 노이즈로 인해 일부 영역에 손실이 발생하는 경우이다. 노이즈로 인해 손실이 발생한 영상은 사용을 못하는 경우가 많지만 손실 영역이 복원된다면 위성영상의 사용이 가능해진다. 다음으로 광범위한 영역에 대해 지상의 많은 정보를 포함하고 있는 위성영상은 사용 목적에 따라 불필요한 정보들이 생기며 이를 제거해야 하는 경우가 발생하기도 한다. 영상에서 구름이 존재하는 경우가 이에 해당한다. 구름으로 덮은 지역은 정보 취득에 한계가 발생하며 이러한 경우 구름을 검출 및 제거 하고 해당지역의 복원이 필요하다[3].

본 논문에서는 대용량의 데이터이며, 영상의 복잡도가 높은 위성영상에 인페인팅 알고리즘을 적용시켜 위성영상을 자동으로 복원하는 방법에 대해 설명하고 실제 위성영상에 인페인팅 알고리즘인 GLCIC[4]을 활용한 결과를 제시함으로써 그 활용도에 대해 논하고자 한다.

### II. 관련 연구

영상 인페인팅(Image Inpainting) 기술은 영상에서 손상되거나 누락된 부분을 주변 영역을 활용하여 해당 영역의 픽셀을 재구성하는 기술을 말하며, 정보가 손실 및 누락된 부분을 주변과 일관성을 유지하면서 얼마나 자연스럽게 생성하는지가 중요하다. 일반적으로는 사진에 스크래치가 생긴 경우처럼 의도치 않게 손실되었거나 보이지 않는 부분의 픽셀을 생성하기 위해 사용되었으나 영상의 레이아웃 편집, 원치 않는 시각적 요소 제거, 얼굴 이미지 편집 등 다양한 분야에서 사용되고 있다.

영상 인페인팅은 연구 시기별 복원영역의 픽셀 재구성 방법에 따라 크게 두 가지 방식으로 구분된다. 먼저 초창기 연구들은 대부분 주어진 영상 내에서 특징을 추출하여 활용하는 원본 영상에 대한 의존도가 높은 방식들이다. 가장 전통적인 확산 기반 영상 합성 기법[1]은 지역적인 특성 히스토그램 기반의 전역적 영상 통계를 활용하여 복원 영역 주변 픽셀 값으로부터 복원 영역으로 선형구조를 전이시키며 영상을 복원한다. 하지만 작고 좁은 영역에 대한 복원에만 좋은 성능을 보이는 한계점이 있다. 패치 기반 영상 합성 기법은 주변영역 텍스처의 샘플링 값으로 영역에 붙여 복원하는 방법이며, 확산 기법 보다 좀 더 자연스럽게 큰 영역을 복원할 수 있지만 주변 영역에는 포함하고 있지 않은 특이한 객체를 제거하는 경우에는 사용이 불가능하다.

이후에 패치 기반 영상 합성 기법에 기반한 여러 방법들이 연구되었으며 일관성 있게 영상을 채울 수 있는 부분에 집중한 Optimal Patch 기법[5], 실시간으로 높은 수준의 영상을 편집할 수 있는 PatchMatch 기법[6]이 등장하였다. 하지만 이러한 방법들은 전체적으로 복잡한 구조, 큰 영역의 영상을 복원하지 못하는 한계점을 가지고 있다.

두 번째로 이러한 한계점을 극복하기 위해 유사한 구조의 대용량 데이터베이스를 활용하는 방법들이다. 초창기 Hays and Efros[7]는 대용량 데이터베이스를 구축하고 입력영상과 가장 유사한 영상을 찾아 입력영상의 빈 영역을 채우는 방식을 제안했다. 최근에는 대용량 데이터베이스의 학습을 통

해 추출된 특징을 활용하여 영상 복원 예측 알고리즘을 개발하며 이 알고리즘을 통해 영상을 복원하는 딥러닝(Deep Learning) 기반 영상복원 방법에 대한 연구방향으로 흘러가고 있다.

딥러닝 기반 영상 복원 알고리즘은 CNN(Convolutional Neural Network), GAN(Generative adversarial networks)[8] 등의 학습 네트워크를 기반으로 개선 목적에 따라 이를 변형하고 융합 한다. 2014년 Kohler[9]는 MRI와 PET영상의 손실 데이터를 계산하기 위해 CNN 기반 영상 인페인팅 방법을 제안하였다. 이는 비교적 적은 용량의 데이터에만 적용되는 한계점이 있다. 2016년에 제안된 Context Encoder(CE)[10]는 큰 영역의 영상복원을 하기 위해 CNN 기반 인페인팅 방법을 확장한 형태로 GAN을 인페인팅 알고리즘 학습에 사용하였다. 최근에는 과거의 정형화 된 형태(예, 정사각형)의 Mask(빈 영역)에 대한 인페인팅 연구와는 다르게 다양한 형태의 Mask 인페인팅을 위한 Free-Form Inpainting[11], 인페인팅의 세부적인 디테일을 개선하기 위해 영상의 Edge 데이터를 활용하는 Edge Connet[12], 8K 정도의 Ultra High-Resolution 영상에 최적화된 Contextual Residual Aggregation[13] 등이 연구됨으로써 인페인팅 기술의 고도화가 활발히 진행 중이다.

### III. 위성영상 영상복원 실험

본 논문에서는 위성영상 복원 실험을 위해 GLCIC(Globally and Locally Image Completion)[4]알고리즘을 사용하였다. 해당 알고리즘은 다양한 영상(highly specific structures)에 적용가능하고 전역적, 지역적 영상 특징을 활용하여 일관된 인페인팅을 목표로 하기 때문에 영상의 복잡도가 높은 위성영상에 적합하다고 생각된다.

GLCIC는 CNN과 GAN을 결합한 이중구조로 총 3개의 Network이 사용된다. CNN 기반 Singe Completion Network은 영상 인페인팅 작업을 수행하고, GAN 기반 Global and Local Context Discriminator Network은 실제적인 완전한 영상으로 복원하기 위해 사용된다. Completion Network은 영상 인페인팅 작업을 속이고 두 개의 Discriminator Network은 인페인팅이 되었는지 여부를 결정하도록 지속적으로 훈련하며, 인페인팅 성능을 높여간다. 특히, GLCIC의 CNN에서는 Standard Convolution layers를 사용하는 대신 Convolutional kernel을 입력 맵에 분산시켜 가중치를 부여하는 Dilated Convolutional layers[14]를 사용하였는데, 이는 가능한 학습의 수를 늘리지 않고 입력으로 사용할 수 있는 영상의 영역을 늘리기 위해 사용되었다. loss function으로는 weighted Mean Squared Error(MSE) loss와 Generative Adversarial network(GAN) loss[8]가 혼합되어 사용되었으며, 이는 안정적인 훈련과 자연스러운 결과에 영향을 미친다.

본 논문에서 GLCIC 알고리즘에 위성영상의 복원 가능성을 평가하기 위해서 사용된 데이터는 NWPU(Northwestern Polytechnical University)에서 제작한 NWPU-RESISC45[15](Remote Sensing Image Scene Classification) 위성영상 데이터이다. 총 31,500장(해상도 256x256)의 데이터 세트로 위성에서 관측될 수 있는 전반적인 공간적 특징을 잘 반영한 45개의 클래스(commercial area, overpass, airplane, airport, bridge, mountain, industrial area 등)로 나뉘어 있으며 각 클래스에 700장의 영상이 포함되어 있다. 영상들은 클래스 내의 다양성을 높이기 위해 Translation, Spatial resolution, Veiwpoint, Object pose, illumination, Background, Occlusion 등의 조건이 반영되어 있다.

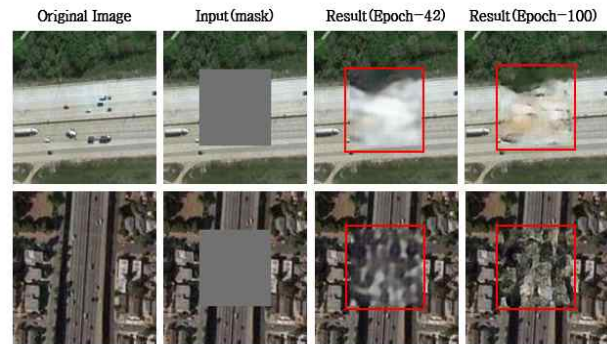
학습을 위해 총 31,500장의 영상 중 45개의 클래스에서 균등한 비율로 학습데이터로 5,461장 선별한 후 학습데이터의 특징을 반영하여 30장의 테스트 데이터를 추가적으로 선별하였다. 256x256 해상도인 데이터 셋을

128x128 해상도로 변환하였으며 128x128 크기 영상의 중앙에 64x64 mask를 생성하였다.

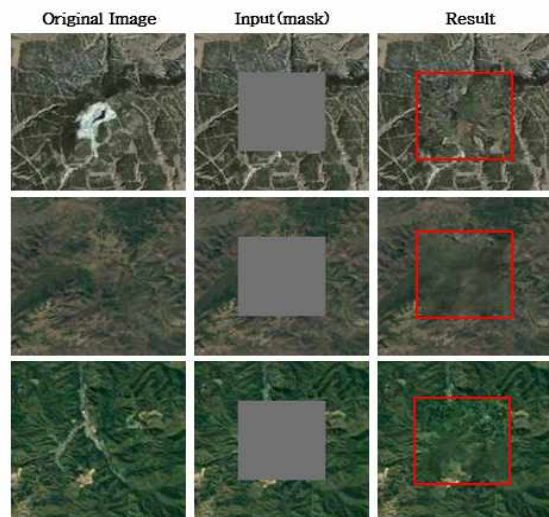
학습은 64x64 크기의 영역을 복원하는 방식으로 총 100 Epoch을 수행하였고, 대략 120시간이 소요되었다.

30장의 테스트 데이터로 확인한 학습 결과는 아래와 같으며, 총 100 Epoch이 수행되면서 변화하는 중간 Epoch의 visual performance결과도 측정하였다.

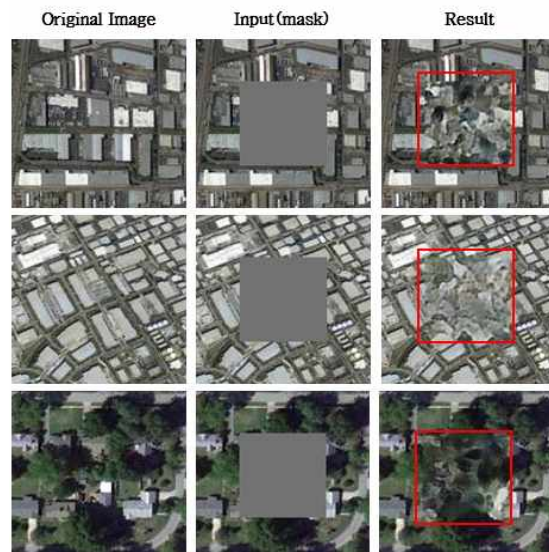
[그림1] 위성영상의 Epoch에 따른 영상복원 결과



[그림2] 단조로운 위성영상에서의 영상복원 결과



[그림3] 복잡도높은 위성영상에서의 영상복원 결과



본 논문에서는 위성영상을 인페인팅 알고리즘을 활용하여 정형화된 64x64 크기의 mask를 복원하였다. Epoch이 증가 할수록 Original Image의 구조를 잡아가는 것을 볼 수 있었고 [그림2]와 [그림3]에서 알 수 있듯이 단조로운 위성영상에서는 시각적으로 꽤 자연스러운 복원력을 보였으며 건물밀집지역, 도로와 같은 복잡도가 높은 위성영상에서는 복원의 완성도가 떨어지는 걸 볼 수 있었다.

기존의 인페인팅 알고리즘들이 대부분 몇 십만 장 단위의 학습데이터를 사용하는 것 대비 본 실험은 약 5000장의 학습데이터를 사용한 점을 감안했을 때 향후 학습데이터를 증량하고 학습네트워크의 보완만으로도 위성영상에서의 영상복원의 성능을 기대해 볼 수 있을 것 같다.

#### IV. 결론 및 향후 연구

본 논문에서 구현된 위성영상 복원 알고리즘은 비교적 단조로운 텍스처의 영상에서는 자연스러운 결과물을 생성했지만, 위성영상에서 자주 관측되는 도시, 마을 등의 복잡한 텍스처의 영상에서는 아직 미비한 인페인팅 결과를 보였다. 향후 위성영상 편집에 최적화된 인페인팅 알고리즘 연구하기 위해 대용량의 위성영상 데이터 구축이 필요해 보이며, 복잡한 텍스처를 복원하기 위한 추가적인 기술융합이 필요해 보인다.

#### ACKNOWLEDGMENT

본 연구는 한국항공우주연구원의 위성정보활용사업의 지원을 받아 수행되었습니다.

#### 참 고 문 헌

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image Inpainting," *Proceedings of SIGGRAPH*, pp.417-424, 2000.
- [2] 안지영, 김태영, and 배성호. "Image Inpainting 모델을 이용한 위성사진의 손실 복구." *한국정보과학회 학술발표논문집* (2018): 2125-2127.
- [3] 이미희, 천은지, and 어양담. "재난 모니터링을 위한 Landsat 8 호 영상의 구름 탐지 및 복원 연구." *대한원격탐사학회지* 35.5 (2019): 861-871.
- [4] Iizuka, Satoshi, Edgar Simo-Serra, and Hiroshi Ishikawa. "Globally and locally consistent image completion." *ACM Transactions on Graphics (ToG)* 36.4 (2017): 1-14.
- [5] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. "Simultaneous Structure and Texture Image Inpainting," *IEEE Trans. Image Processing*, Vol.12, No.8, pp.882-889, 2003.
- [6] Barnes, Connelly, et al. "The generalized patchmatch correspondence algorithm." *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2010.
- [7] Hays, James, and Alexei A. Efros. "Scene completion using millions of photographs." *ACM Transactions on Graphics (TOG)* 26.3 (2007): 4-es.
- [8] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.
- [9] Köhler, Rolf, et al. "Mask-specific inpainting with deep neural networks." *German Conference on Pattern Recognition*. Springer, Cham, 2014.
- [10] Pathak, Deepak, et al. "Context encoders: Feature learning by inpainting." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [11] Yu, Jiahui, et al. "Free-form image inpainting with gated convolution." *Proceedings of the IEEE International Conference on Computer Vision*. 2019.
- [12] Nazeri, Kamyar, et al. "Edgeconnect: Generative image inpainting with adversarial edge learning." *arXiv preprint arXiv:1901.00212* (2019).
- [13] Yi, Zili, et al. "Contextual Residual Aggregation for Ultra High-Resolution Image Inpainting." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.
- [14] Yu, Fisher, and Vladlen Koltun. "Multi-scale context aggregation by dilated convolutions." *arXiv preprint arXiv:1511.07122* (2015).
- [15] Cheng, Gong, Junwei Han, and Xiaoqiang Lu. "Remote sensing image scene classification: Benchmark and state of the art." *Proceedings of the IEEE* 105.10 (2017): 1865-1883.