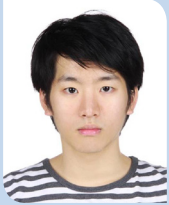


건축 분야에서 활용되는 인공지능 기반의 이미지 프로세싱 방법 소개



왕승현 한양대학교 건축공학부 겸임교수, wsh1019@hanyang.ac.kr

1. 서론

CCTV (Closed-Circuit Television), 디지털 카메라, 스마트폰, 레이저 스캐너 등 촬영 장비의 성능 향상과 함께, 보급화로 건설 분야에서 높은 품질의 방대한 양의 이미지 데이터 수집이 가능해졌다. 또한, 수집된 이미지를 분석해 안전관리, 공사관리 등 건설업의 전반적인 품질 관리 향상을 위한 다양한 작업이 수행되고 있다. 하지만, 많은 양의 데이터를 수작업으로 분석하는 것은 매우 오랜 시간이 걸리고, 많은 비용이 요구되기 때문에 자동화에 대한 필요성이 대두되고 있다.

이미지 프로세싱 (Image Processing)은 컴퓨팅 장치 (예: 데스크탑, 노트북, 스마트폰 등)를 이용해 사용자가 원하는 목적에 맞게 디지털 이미지를 변환하는 기술이다. 예를 들어, 이미지 복원 기술을 적용해 저화질의 이미지를 더 선명한 이미지로 가공하거나, 이미지 정합 기술을 이용해 여러 이미지를 하나의 파노라마 이미지로 생성할 수도 있다. 현재 이미지 프로세싱은 건축 분야에서 수집된 방대한 양의 이미지를 자동으로 분석하기 위한 방법으로 활발하게 사용되고 있다. 궁극적으로 이 자동화의 목표는 인간이 수행할 수 있는 수준까지 도달하는 것을 목표로 하고 있기 때문에, 특정 적용처에 있어서 최적의 방법을 개발하기 위해 다양한 이미지 프로세싱 기술들에 대한 적용 가능성에 대한 의논과 실험 등 여러 노력도 함께 수행되고 있다.

과거부터 현재까지 인공지능 기반의 다양한 방법론들이 이미지 프로세싱에 접목되어 건축 분야에 사용되어왔다. 과거

에는 주로 건축공학자, 인공지능 엔지니어 등 사람이 수작업으로 고안한 이미지의 특징을 사용하는 규칙 기반 또는 머신러닝 기반의 방법들(Rule-based or Machine Learning-based Approach) 이용해 문제를 해결해 왔다. 하지만, 특정 자동화 작업과 관련된 모든 이미지의 특징(예: 조도의 변화, 물체의 크기변화, 물체의 움직임의 변화 등)을 수작업으로 정확하게 판단하는 시간이 오랜 시간이 걸리며, 다양한 특징의 입력이 필요할 경우, 해당 분야 전문가라 할지라도 가능한 특징을 모두 고려하는 것은 매우 복잡한 작업이다. 따라서, 대체로 이 방법은 이미지의 특징이 매우 균일한 시나리오에서 제한되어 적용되었다. 이와는 반대로, 딥러닝 기반의 방법(Deep learning-based Approach)은 많은 양의 데이터를 이용해 기계가 직접 스스로 필요한 이미지의 특징을 학습하는 방법이다. 현재 이러한 딥러닝의 메커니즘을 이용한 자동화 방법은 기존의 이미지 프로세싱 방법들로 해결할 수 없었던 영역에 중점적으로 적용되고 있으며, 이미지 특징에 대한 다양한 변이가 동시에 발생하는 영역에서 높은 정확도를 보여주고 있다.

다양한 딥러닝 방법들이 적용되어 자동화의 수준을 높이고 있지만, 현재 3가지 작업: '이미지 분류(Image Classification)', '객체탐지(Object detection)', '시멘틱 세그멘테이션(Semantic Segmentation)'은 건축 분야에서 가장 활발히 사용되고 있는 기술들로 여겨지고 있다. 본 원고에서는 이 3가지 작업과 관련된 딥러닝 관련 기술들의 배경과 함께, 건축 분야에 적용된 다양한 연구 사례를 소개한다.



2. 이미지 분류

이미지 분류는 주어진 이미지 내에 어떤 종류의 사물이 존재하는지 분류해내는 작업이다. 이미지 분류는 앞에서 언급된 다른 작업과 비교했을 때, 연산량에 있어서 상대적으로 간단한 작업이지만, 건축 분야에서 여전히 활발하게 사용되고 있다. 분류에 사용되는 강력하고, 높은 정확도를 보이는 딥러닝 알고리즘은 AlexNet, VGG16, ResNet-50 (ResidualNet), ResNet-101 등이 있다.

Chow 외 6인 (2021) [1]의 연구에서는 건물 내부의 콘크리트의 상태를 자동으로 진단 하기 위한 방법을 개발하고자 했다. 더 구체적으로 설명하자면, 먼저, 건물 내부를 360도 카메라와 Lidar로 촬영해 이미지를 수집한 후, 딥러닝 기반의 이미지 분류 알고리즘을 적용해 콘크리트의 상태를 정상, 크랙, 박락으로 자동으로 분류하고자 했다<그림 1>. 딥러닝 모델은 ResNet-50 알고리즘을 사용했으며, 18,500개의 이미지를 학습시켰다. 테스트 결과 각 정확도는 86.8%, 90.6% and 96.2%였다.



그림 1. 콘크리트 상태의 분류 예시: 정상, 크랙, 박리

Xiong 외 2인은 (2020) [2] 드론으로 촬영한 이미지로부터 각 건물의 지진 피해 여부 (정상 또는 지진)를 판단하는 방법을 개발하고자 했다<그림 2>. 이 연구에서는 지오레퍼런싱 기술 (Georeferencing) 을 활용해 특정 건물의 상태를 촬영할 수 있는 위치에 드론을 보내어 이미지를 수집 한후, 딥러닝 알고리즘을 이용해 지진의 피해 여부만을 판단하고자 했다. 딥러닝 기반의 분류 알고리즘으로 VGG16이 사용되었으며, 총 2,000개 (정상건물: 1,000개, 지진피해 건물: 1,000개)를 이용해 학습시켰다. 그 결과, 지진으로 붕괴된 건물에 대한 판단 정확도는 89.39% 였다.

Gi1 외 2인 (2018) [3]의 연구에서는 각 현장의 현장 관리자가 촬영한 이미지를 총 27개의 공사의 종류 (예, 기초공사,



그림 2. 건물의 지진 분류 상태 예시

전기공사, 인테리어 공사 등) 로 분류하고자 했다<그림 3>. 이미지 분류 알고리즘으로 Google Inception v3을 1,208장의 이미지 데이터를 사용해 학습시켰다. 또한, 적은 수의 데이터로 효과적으로 학습하기 위해 이미 다른 분야에서 학습된 모델을 사용하는 전이 학습(Transfer Learning)을 사용했다. 그 결과, 27개의 공사의 종류에 대해 평균적으로 92.6%의 정확도를 보였다.



그림 3. 공사 종류 분류 예시

3. 객체 탐지

객체탐지는 관심영역 (Region of Interest)를 찾아내는 지역화(localization)와 영역 내의 항목을 분류를 함께 수행하는 작업이다. 즉, 이미지 분류와 지역화 알고리즘을 함께 결합한 것이다. 객체 탐지는 이러한 두 가지 작업을 동시에 수행해야 하기 때문에, 어려운 과제로 인식되었으나, 딥러닝 기반 기술의 적용은 현저히 낮았던 정확도를 비약적으로 향상시켰다.

딥러닝 기반의 객체탐지 분야의 알고리즘은 크게 1-stage방법과 2-stage방법의 계열들로 나눌 수 있다. 두 방법의 명확한 차이로서, 1-stage방법은 지역화와 이미지 분류 작업에 대한 연산을 한번 수행하는 반면, 2-stage방법은 지역화와 이미지 분류 작업을 추가적으로 더 수행한다. 그 결과,

1-stage방법은 연산속도는 빠르지만, 정확도가 더 낮고, 2-stage방법은 연산속도는 느리지만, 더 정확한 결과를 보여주고 있다. 각 계열의 대표 방법으로는 YOLO (You Only Look Once)와 Faster R-CNN (Faster Region Convolutional Neural Network)이 있다.

건축 분야에 객체탐지 방법을 적용한 다양한 연구사례들이 존재한다. 일반적으로 빠른 속도로 연산이 필요한 분야에서는 1-stage방법이 선호되고, 정확도를 더 선호하는 경우, 2-stage방법이 선택되어진다. Fang 외 7인은 (2018) [4] 건축 시공 현장에서 인부들의 안전모 착용 여부를 감지하고자 했다(그림 4). 이를 위한 방법으로 현장에서 CCTV를 설치해 데이터를 실시간으로 수집하고, 객체탐지 알고리즘 중 하나인 Faster R-CNN방법을 적용하는 방법을 제안하였다. Faster R-CNN의 학습을 위해 총 25개의 현장에서 수집한 81,000개의 데이터를 모델의 학습에 사용했다. 구현된 모델의 검증을 수행하기 위해 날씨, 조도, 스케일 등에 대한 변이에도 모델의 정확성이 우수한지 검증을 수행하였으며, 그 결과 평균적으로 95%의 정확도를 보였다.



그림 4. 안전모 착용 여부 결과

Cha 외 3인은 (2021) [5] UAV (Unmanned Aerial Vehicles)를 이용해 수집된 콘크리트와 철판에 대한 이미지로부터 총 6개의 결함(철판 부식 보통, 철판 부식 심함, 철판 바리, 콘크리트 크랙, 볼트 부식)으로 발생 위치와 함께 자동으로 분석하는 방법을 제안했다(그림 5). 사용된 객체탐지 알고리즘은 Faster R-CNN이며, 2,366개의 이미지를 이용해 학습시켰다. 그 결과, 평균적으로 90%의 정확도로 분류할 수 있었다.

Li 외 2인은 (2021) [6] 자재창고 등 보관 중인 다수의 철근의 개수를 자동으로 세기 위한 방법을 제안했다. 이를 위한 방법으로 YOLO기반의 객체탐지 알고리즘을 사용해 바운

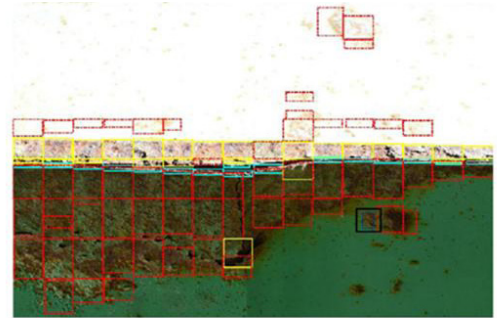


그림 5. 콘크리트와 철판 상태 감지 결과

딩 박스를 생성한 후, 바운딩 박스의 개수를 셈으로서, 철근의 수를 추정하고자 했다. YOLO알고리즘을 학습하기 위해 가능한 철근들이 식별될 수 있는 위치에서 일반 RGB 카메라를 사용해 589개의 철근 이미지를 수집했다(그림 6). 실제 테스트 결과 YOLO모델은 94.7%의 정확도를 보였다. 또한, 이 연구에서는 추가적으로 Faster R-CNN 알고리즘을 학습한 결과와 비교 실험을 수행하였으며, YOLO모델보다 더 높은 97.6%의 정확도를 보였다.

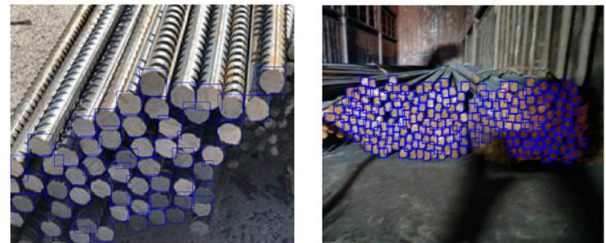
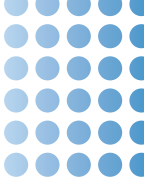


그림 6. 철근 개수 검출 결과

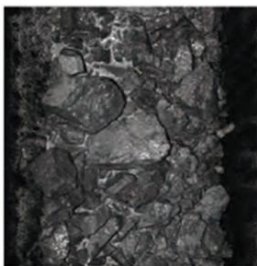
4. 시멘틱 세그멘테이션

시멘틱 세그멘테이션은 이미지 내에 관심 객체를 픽셀 단위로 영역을 정밀하게 분류하는 작업이다. 같은 관심 객체들은 동일한 픽셀로 수행된다. 시멘틱 세그멘테이션의 수행에 있어 많은 분야에서 현재 높은 정확도를 보여주며, 널리 사용되는 유명한 알고리즘으로는 U-Net, DeepLab 등이 있다. 건축분야에서 적용된 시멘틱세그멘테이션의 연구 사례는 다음과 같다.

Zhou 외 3인은 (2021) [7] 터널의 철길에 존재하는 돌의 사이즈를 자동으로 분석하기 위한 방법을 제안했다(그림 7). 그들은 TBM (Tunnel Boring Machines)을 이용해 철길을



일정한 높이로 촬영한 이미지를 사용했다. 또한, 사용된 딥러닝 알고리즘으로는 MSD-UNet (Dual UNet with Multi-Scale inputs and side-output) 기반의 방법을 사용했으며, 총 2,500개의 데이터를 이용해 모델을 학습 시켰다. 실험 결과, 각 픽셀의 사이즈는 0.387 mm/pixel임을 알 수 있었으며, 학습된 딥러닝 모델의 정확도는 80% ~ 100% 사이였으며, 평균 정확도는 약 87%였다.



원본 이미지



세그멘테이션 결과

그림 7. 터널 철길 이미지 세그멘테이션 결과

Wang 외 4인은 (2022) [8] 건축 시공 현장에서 발생하는 이미지를 시멘틱 세그멘테이션을 적용하는 연구를 수행했다(그림 8). 이 시멘틱 이미지의 잠재적인 활용처로는 이미지 분류, 객체탐지 등의 분석에 불필요한 클래스는 제거하고, 필요한 클래스 이미지만을 남긴 시멘틱 이미지를 사용함으로써 필요한 클래스의 작업에 집중할 수 있다. 그들은 인부, 타워크레인, 굴삭기, 트럭 등 총 12개의 클래스에 대한 859개의 이미지를 학습 데이터로 사용했으며, 시멘틱 세그멘테이션 알고리즘으로 DeeplabV3+을 이용했다. 학습된 모델을 새로운 테스트 데이터에 적용한 결과 평균 정확도는 92.62%였다.



원본 이미지



세그멘테이션 결과

그림 8. 공사현장 세그멘테이션 결과 (인부, 타워크레인)

5. 결론

이 원고에서는 건설 분야에서 자동화를 위해 딥러닝 기반의 이미지 프로세싱을 적용한 연구의 사례를 알아보았다. 소개한 연구들의 결과로부터 딥러닝의 적용은 많은 작업에서 높은 정확도를 보이고 있다는 것을 알 수 있었다. 또한, 이러한 연구결과는 딥러닝 방법론의 적용이 건축 분야의 다른 많은 작업에서 높은 정확도를 줄 수 있는 잠재성을 보여 주고 있다. 다시 말해서, 딥러닝을 이용한 이미지 프로세싱의 활용은 위에서 언급된 사례를 제외하고도, 자동화를 위한 도구로서 많은 범위로의 적용이 가능하다. 하지만, 신뢰할만한 정확도를 달성하기 위해 딥러닝기반의 방법은 많은 데이터를 요구하기 때문에, 구축 비용에 있어서 많은 자원이 투입이 요구된다. 따라서, 상대적으로 적은 데이터를 이용해 자동화 구현이 가능한 규칙 기반이나 머신러닝 방법과 비교했을 때, 딥러닝이 모든 상황에 있어서 대안이 될 수 없다는 것을 의미한다. 이런 이유로 각 자동화 프로세스에 있어서 딥러닝 적용이 적합한지에 대해서 해당 문제에서의 이미지의 변이성과 기존의 유사한 사례의 충분한 검토가 필요할 것으로 판단된다.

참고문헌

1. J.K. Chow, K. fu Liu, P.S. Tan, Z. Su, J. Wu, Z. Li, Y.H. Wang, Automated defect inspection of concrete structures, *Autom. Constr.* (2021). <https://doi.org/10.1016/j.autcon.2021.103959>.
2. C. Xiong, Q. Li, X. Lu, Automated regional seismic damage assessment of buildings using an unmanned aerial vehicle and a convolutional neural network, *Autom. Constr.* (2020). <https://doi.org/10.1016/j.autcon.2019.102994>.
3. D. Gil, G. Lee, K. Jeon, Classification of images from construction sites using a deep-learning algorithm, in: *ISARC 2018 - 35th Int. Symp. Autom. Robot. Constr. Int. AEC/FM Hackathon Futur. Build. Things*, 2018. <https://doi.org/10.22260/isarc2018/0024>.
4. Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T.M. Rose, W. An, Detecting non-hardhat-use by a deep learning method from far-field surveillance videos, *Autom. Constr.* (2018). <https://doi.org/10.1016/j.autcon.2017.09.018>.
5. Y. Li, Y. Lu, J. Chen, A deep learning approach for real-time rebar counting on the construction site based on YOLOv3 detector, *Autom. Constr.* (2021). <https://doi.org/10.1016/j>.

autcon.2021.103602.

6. Y.J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, O. Büyükoztürk, Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types, *Comput. Civ. Infrastruct. Eng.* (2018). <https://doi.org/10.1111/mice.12334>.
7. X. Zhou, Q. Gong, Y. Liu, L. Yin, Automatic segmentation of TBM muck images via a deep-learning approach to estimate the size and shape of rock chips, *Autom. Constr.* (2021). <https://doi.org/10.1016/j.autcon.2021.103685>.
8. Z. Wang, Y. Zhang, K.M. Mosalam, Y. Gao, S.L. Huang, Deep semantic segmentation for visual understanding on construction sites, *Comput. Civ. Infrastruct. Eng.* (2022). <https://doi.org/10.1111/mice.12701>.