

생성형 AI를 활용한 BIM 기반 건설현장 영상데이터 생성기술

BIM-based Image Dataset Generation Technology for Construction Site Using Generative AI

1. 서론

최근 딥러닝 기술의 급격한 발전과 함께, 건설 분야에 최신 딥러닝 기술 적용을 통한 건설현장 모니터링을 수행하고자 하는 연구가 많이 진행되고 있다. 이러한 기술의 도입은 건설 환경 디지털화를 위한 필수적인 요소임에도 디지털 데이터 및 신뢰도 부족 문제로 다른 산업 분야보다 기술 개발이 어려운 실정이다.

2022년부터 과학기술정보통신부와 한국지능정보사회진흥원(NIA)에서 인공지능 학습용 데이터 구축 지원사업[1]을 진행하며 건설현장의 데이터베이스를 구축하기 위한 지원을 하고 있지만 현재 수집된 건설현장 데이터를 실제 건설현장에 적용하기 위한 딥러닝 모델 학습에 활용하기에는 아래와 같은 여러 한계가 있다.

첫째는 건설현장 영상데이터의 제한적인 양이다. 컴퓨터 비전 분야는 ImageNet[2], MS COCO[3]와 같은 대규모 데이터셋 구축 및 공개를 통해 폭발적인 발전을 이룰 수 있었던 반면, 건설현장은 매우 제한적인 양의 데이터가 공개되어 있다 [4]. 건설현장 이미지는 사생활과 보안의 문제로 대중에 공개하는 것이 금지된 경우가 많으며[5], 특히 사람이 포함된 영상의 경우에 큰 제한이 있다. 이러한 문제가 최근 학습용 데이터 구축 사업을 통해 일부분 극복되기는 하였으나, 제한된 객체와 양식만을 대상으로 구축되어 확장성 및 다양성 측면에서 한계가 있다.

둘째는 건설현장 데이터베이스의 한정적 활용이다. 건설현장 모니터링은 개발된 딥러닝 모델에 따라 이미지 분류, 객체 탐지, 객체 분할, 깊이 추정과 같이 다양한 과제를 수행해야 하며, 과제에 따라 학습에 필요한 데이터의 양식이 다르다. 하지만, 기존에 구축되어 공개된 건설현장 영상데이터는 제한된 객체에 대해 객체 탐지와 객체 분할 과제만을 위한 데이터를 제공하고 있으며, 이러한 데이터 종류의 한계는 건설현장 모니터링을 위한 딥러닝 모델 개발의 확장성을 제한하고, 구축된 데이터베이스는 현재 한정적인 활용만이 가능한 실정이다.

셋째는 건설 환경의 다양성과 기존 데이터의 한계이다. 건설현장에서 획득되는 영상데이터는 객체 타입, 영상의 해상도, 조도, 카메라의 설치 위치, 건설 환경에 따라 다양한 특징을 가지게 된다. 이러한 다양성은 특정 소수의 건설현장에서 확보된 영상데이터를 기반으로 개발된 딥러닝 모델의 성능을 저하시키는데 큰 역할을 하며[6], 딥러닝 기반의 건설현장 모니터링 기술 현장적용의 한계로 남아있다. 컴퓨터 비전 분야에서는 이러한 문제를 대용량의 데이터베이스 구축을 통해 해결하였으나, 건설현장에서는 위와 같은 해결 방법을 적용하기에는 어려움이 있다.



김 형 수

한국과학기술원 응용과학연구소 연수연구원

건설 환경에는 BIM(Building Information Modeling), GIS(Geographic Information System), 디지털 트윈 모델과 같이 데이터 생성을 위해 활용할 수 있는 다양한 자원이 있으며, 위와 같은 데이터 부족의 문제를 극복하기 위해 다양한 방식으로 필요한 데이터를 확보하기 위한 연구가 진행되고 있다. 2장에서는 건설 환경에서 영상데이터 생성 연구를 진행한 사례를 소개하고자 한다.

2. 건설 환경 영상데이터 생성기술 동향

2.1 가상 환경을 활용한 건설 환경 학습용 데이터 생성기술

최근, 가상 환경을 구축할 수 있는 다양한 소프트웨어가 제공되면서 가상의 건설 환경을 구축하여 학습용 데이터를 구축하고자 하는 노력도 많이 진행되고 있다. 이러한 방식은 실제 건설현장의 영상이 없어도 딥러닝 모델을 학습시킬 수 있다는 것과 다양한 상황을 시뮬레이션하여 원하는 영상데이터를 얻을 수 있다는 장점이 있다. 이러한 장점은 건설현장이 가지고 있는 복잡한 환경을 구현하고 다양한 촬영 환경을 구현함으로써 데이터셋의 다양성을 확보하는데 큰 도움이 된다. 하지만, 가상 환경은 아직 실제 환경과의 괴리감으로 인해 발생하는 모델 성능 저하 문제를 해결하고 있지 못해 실제 산업현장에서 적용하기에는 추가적인 연구가 필요한 상황이다.

Soltani 등(2016)[7]은 건설현장에서 영상분석을 통해 굴착기를 탐지하는 모델을 학습하고자 하였고, 데이터셋 구축을 위하여 CAD 3D 모델을 활용하였다. 가상 굴착기 모델을 활용함으로써 원하는 굴착기 색상과 모델을 선택할 수 있었으며 이미지 촬영 조도를 설정하고 다양한 각도에서 촬영된 굴착기 이미지를 생성함으로써 많고 다양한 굴착기 데이터셋을 확보할 수 있었다. 해당 연구를 통해 가상 이미지를 통해 생성된 데이터셋으로 학습한 모델이 실제 현장에서 유의미한 탐지 성능을 확보할 수 있음이 확인되었다.

Neuhausen 등(2020)[8]은 Blender 2.80 software를 활용하여 건설 환경을 구축하고 Cycle 렌더링 엔진을 활용하여 실제 건설현장과 유사한 이미지를 생성하였다. 해당 연구에서는 건설현장에 등장하는 근로자를 top-view에서 탐지하는 모델을 학습하는 것을 목표로 다양한 배경에서의 영상데이터를 생성하였다. 실제 건설현장을 기반으로 가상 환경을 구축하여 실제 이미지와 상당히 유사한 영상데이터

를 확보할 수 있었고 근로자에 대한 3D 모델을 적용하였기 때문에 원하는 장면을 시뮬레이션하여 데이터셋을 구축하였다.

Assadzadeh 등(2022)[9]은 건설현장에서 굴착기의 포즈 탐지를 수행하기 위해서 다양한 포즈와 종류의 굴착기 이미지를 생성하는 기술을 제안하였다. 해당 기술에서는 굴착기의 3D CAD 모델을 기반으로 다양한 포즈와 텍스처를 가진 굴착기 모델을 생성하고 배경을 삽입하기 위해 도메인 랜더마이즈(Domain Randomization)를 수행하였다. 이를 통해 생성된 가상 이미지를 통해 단순 객체 탐지를 넘어서 복잡한 영상분석 모델을 학습하는 것이 가능함을 시사하였다.

이렇듯, 가상 환경을 구축하여 학습 데이터를 구축하는 기술은 다양한 객체를 포함하고 복잡한 배경을 가지고 있는 건설 환경에서 매우 유용하게 활용될 수 있음이 확인되었다. 가상 환경 구축을 통해 실제 건설현장에서 발생하기 힘든 상황을 시뮬레이션하고 해당 데이터를 확보하는 등 다양한 방법으로 활용될 수 있으며 객체 탐지뿐만 아니라 다양한 과제를 수행하는 영상기반 딥러닝 모델 학습을 위한 데이터 생성기술은 추후 건설 환경 디지털화 및 기술 첨단화 과정에서 필수적으로 요구되는 기술이라 할 수 있다.

2.2 BIM 기반 건설 환경 학습용 데이터 생성기술

BIM은 건축물의 3D 모델에 정보를 표현하는 객체지향 모델링 기술로 건축물의 설계부터 시공, 과거의 유지관리 정보를 디지털화하여 유지관리에서 해당 정보를 활용할 수 있는 특징이 있으며, 건축 환경의 디지털화 과정에서 필수적인 요소로 고려되고 있다. 또한, 우리나라는 현재 500억 이상의 도로사업에 BIM을 의무화하는 등 BIM의 활용을 적극적으로 장려하고 있으나, 설계나 시공단계의 3D 모델에 국한되어 활용되고 있다.

건설 환경에서 발생하는 데이터 부족 문제는 대용량의 영상데이터를 수집하는 것뿐만 아니라 대용량의 데이터를 수작업으로 레이블링하는 정제 작업이 큰 문제로 지적됐다. 이를 해결하기 위해 초기 BIM 기반의 데이터 생성 연구들은 수집된 영상데이터를 기반으로 레이블링 작업을 자동화하는 레이블링 자동화(Auto-labeling) 연구에 초점을 맞추어 진행되었다.



그림 1 4D BIM 및 포인트 클라우드를 활용한 레이블링 생성[10]

Braun and Borrmann(2019)[10]는 4D BIM에 포인트 클라우드를 함께 고려하여 이미지 레이블링 정보를 자동으로 생성하는 모델을 제안하였다. 4D BIM을 보유하고 있는 건설현장에 대해 다양한 각도에서 영상을 촬영하여 해당 현장의 포인트 클라우드 모델을 확보한 뒤 BIM에서 확보한 객체 정보와 포인트 클라우드 정보를 정합하여 레이블링 이미지를 생성한다. 해당 기술은 객체 인식, 객체 분할과 같은 다양한 딥러닝 분야에 활용할 수 있으며, 그림 1과 같이 복잡한 이미지에서도 효과적으로 구성 요소를 레이블링하는 모습을 확인할 수 있다.

Alawadhi와 Yan(2020)[11]은 BIM 모델과 초현실적인 렌더링 기술(Hyper realistic rendering)을 활용하여 구조물의 객체 분할 학습용 데이터를 생성하는 모델을 제안하였다. 해당 연구에서는 BIM 모델에 포토리얼리스틱 렌더링

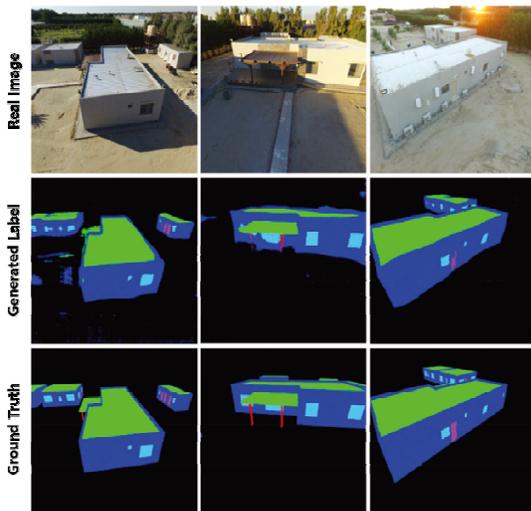
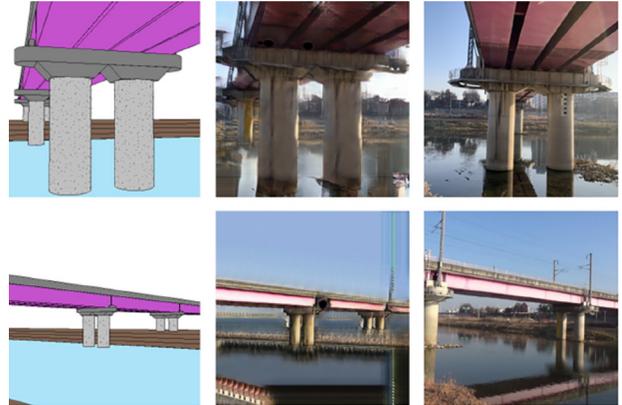


그림 2 BIM 기반 객체 분할 레이블링 생성 사례[11]



BIM 생성 이미지 실제 이미지
그림 3 사회기반시설 영상데이터 생성 사례[13]

을 적용하여 실제 촬영한 이미지와 유사한 이미지를 생성하고 BIM에서 추출된 객체 분할 레이블링 이미지와 함께 pix2pix[12] 모델을 학습하였다. 학습된 모델은 실제 촬영된 영상을 입력받으면 객체 분할 레이블링을 수행하는 레이블링 자동화 및 학습용 데이터 생성을 수행할 수 있다. 학습된 모델에서 생성된 객체분할 레이블링 이미지는 실측 자료와 0.642의 mIOU를 보이며 실제 딥러닝 모델 학습에 활용할 수 있는 정도의 정확도를 보여 BIM 기반의 데이터 생성기술이 실제 현장에 적용 가능성을 입증하였다.

Hong 등(2021)[13]은 BIM 모델을 활용하여 사회기반시설의 학습용 이미지를 생성하는 기술을 제안하였다. 해당 기술에서는 BIM 모델 실사화를 위해 CycleGAN[14] 모델을 활용하였으며 실내 구조물 객체 분할뿐만 아니라 야외 사회기반시설 객체 분할을 수행하는 모델을 효과적으로 학습하는 데이터셋을 생성하였다.

위의 그림 3과 같이 BIM 모델을 활용하여 생성된 이미지는 실제 현장의 모습을 효과적으로 모사하여 딥러닝 모델 성능 향상에 큰 역할을 하고 있다. 또한, 생성형 AI를 활용하여 BIM 모델을 학습용 데이터로 변환하는 연구가 실제 촬영한 영상과 상당히 유사한 이미지를 생성할 수 있다는 것이 입증되어 이러한 연구 방향은 앞으로 건설 환경 디지털화를 달성하는데 큰 역할을 할 수 있을 것으로 기대된다.

2.3 생성형 AI를 활용한 건설 환경 학습용 영상데이터 생성기술

생성형 AI는 이미지, 텍스트, 음악과 같은 유형의 데이터 형태로 새로운 데이터를 생성하는 인공지능 기술의 한 종

류로 실제 데이터와 유사한 데이터를 생성할 수 있다. 가상 환경을 통해 합성 데이터는 실제 현장 데이터와의 분포 차이로 딥러닝 모델 적용 단계에서 안 좋은 영향을 줄 수 있으며, 합성 데이터를 실제 데이터와 유사하게 변형하여 딥러닝 성능을 향상하는 것은 주요 과제로 남아있다.

기존에 건설 환경에서는 생성형 AI를 통해 데이터 증강을 수행하고자 하는 연구가 많이 진행되었다. 데이터 증강 기술은 부족한 데이터셋을 추가로 확보하기 위해 기존에 보유하고 있는 데이터를 기반으로 새로운 데이터를 생성하는 기술로 데이터 부족 문제를 효과적으로 해결하여 딥러닝 학습 과정에서 보편적으로 적용되고 있다. 과거의 데이터 증강기술은 단순히 데이터를 합치고 자르고 뒤집는 등 1차원적인 데이터 변형 수준이었지만 최근 생성형 AI 기술의 적용이 효과적으로 데이터 증강 효과를 보임이 입증되면서 다양한 생성형 AI를 건설 환경에 적용하고 있다.

Bang 등(2020)[15]이 제안한 기술은 기존에 확보된 데이터에서 객체 정보와 배경을 추출하여 무작위 조합을 시도하여 새로운 이미지 데이터를 생성한다. 객체와 배경 정보를 조합하는 과정에서 GAN 모델을 통해 객체 채워 넣기(Inpainting)을 수행하는데 GAN 모델은 객체를 효과적으로 합성시켜 자연스러운 이미지를 생성해내는 역할을 한다.

Back 등(2022)[16]은 conditional GAN[17]과 적대적 학습을 통해 데이터 생성을 수행하는 기술을 제안하였는데 생성된 이미지의 대상 객체는 하나의 이미지에서 생성되었지만 다양한 색상과 텍스처를 가지며 실제 이미지와 상당한 유사성을 가진다. 해당 기술로 구축된 데이터셋은 객체 분류를 수행하는 딥러닝 모델을 효과적으로 학습하였으며 기존에 데이터만으로 학습한 모델보다 높은 성능을 보여 생성형 AI의 데이터 증강 효과를 입증하였다.

3. 생성형 AI를 활용한 BIM 기반 건설현장 영상 데이터 생성기술

3.1 필요성

신속하고 효율적인 건설 환경 디지털화를 위해서는 건설 현장 데이터 부족 문제를 해결하고 학습용 데이터의 다양성과 확장성을 확보할 수 있는 기술 개발이 필요하다. BIM 모델과 생성형 AI는 이미 건설 환경에서 주목받고 있는 기

술들이며 건설 환경의 첨단화와 자동화를 달성하는데 필수적인 요소이다. 건설 환경 영상 데이터베이스 구축 기술은 건설현장에 적용되는 시각 장면 인식기술을 획기적으로 발전시킬 수 있으며 이를 통해 효과적인 모니터링 기술 개발을 달성할 수 있다.

해당 기술 개발 및 건설현장의 원활한 적용을 위해서는 아래와 같은 항목의 연구가 수행되어야 한다.

1) BIM 기반 합성 이미지 데이터 생성 알고리즘 개발

BIM 기반의 학습용 데이터 생성을 위해서는 건설현장에 적합한 BIM 라이브러리를 선정해야 하며, BIM 3D 모델상에서 목표 객체를 추가하고 2D 이미지를 추출할 수 있는 알고리즘을 개발하여야 한다. 해당 과정에서 BIM의 객체 정보를 활용하여 레이블링 이미지를 자동으로 생성하고 이미지를 정합할 수 있어야 한다.

2) 합성 이미지 데이터 실사화를 위한 생성형 AI 개발 및 학습

다양한 이미지 생성형 AI 모델을 후보군으로 건설현장 이미지 실사화에 가장 적합한 생성형 AI를 선정하고 개발하여야 한다. 이 과정에서 생성형 AI를 통한 실사화 후에도 레이블링 이미지와의 정합성을 유지하는 것이 가장 중요하며 생성형 AI의 실사화 능력을 제대로 평가할 수 있는 객관적인 성능 지표를 제안할 필요가 있다.

3.2 건설현장 기술 적용 방향 및 이점

생성형 AI를 활용한 BIM 기반 학습용 데이터 생성기술은 건설현장의 데이터를 신속하게 생성하며 높은 신뢰도를 가지는 딥러닝 모델의 학습에 활용될 수 있다. 또한, 이미지 분류, 객체 탐지, 객체 분할, 깊이 추정과 같은 원하는

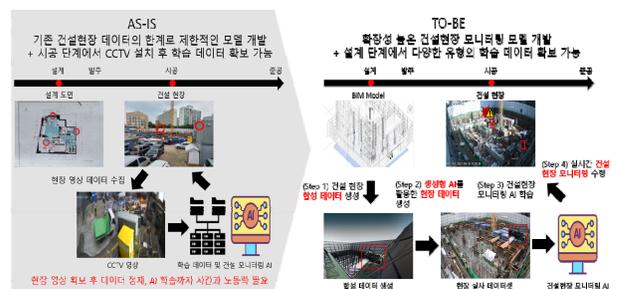


그림 4 기술 적용 시나리오 및 이점

타입의 학습용 데이터를 생성할 수 있으며, 목표 객체를 자유롭게 설정하여 영상데이터 수집 및 정제에 걸리는 시간을 단축할 수 있다. 그리고 건설공사의 설계단계에서 학습용 데이터를 확보하여 딥러닝 모델을 학습할 수 있으므로, 건설공사의 시공단계 시작부터 별도의 데이터 수집 과정없이 건설현장 모니터링을 수행할 수 있다. 마지막으로 BIM 기반으로 생성된 데이터는 건설현장 모니터링 대상 현장에 최적화된 데이터를 생성하여 높은 성능의 딥러닝 모델을 학습할 수 있다.

그리고 생성형 AI를 활용함으로써 얻을 수 있는 다양한 이점도 존재하는데, BIM을 통해 생성된 건설현장 영상데이터를 딥러닝 모델 학습에 활용하여 실제 건설현장에서 촬영한 영상데이터를 활용한 것과 비슷한 성능의 더 다양하고 확장성있는 모델 학습이 가능하다. 추가로, 생성형 AI 기능을 이미지 실사화에 국한하지 않고, 단일 영상으로부터 다양한 영상데이터 생성을 통해 학습 데이터의 양과 다양성 측면에서의 한계를 해결할 수 있을 것이다.

3.3 활용방안

시각 장면 이해(Visual scene understanding) 기술 개발은 건설 환경 디지털화의 필수 요소로 디지털 트윈, 안전 모니터링, 공정 관리와 같은 다양한 과업에 활용이 가능한 기술이다. 건설현장 학습용 데이터 생성기술은 건설현장 영상데이터 부족 문제를 해결하고 다양한 딥러닝 모델 과제에 맞는 학습용 데이터 생성을 통해 영상기반 건설현장 모니터링 기술개발의 원천 기술로써 활용할 수 있을 것이다.



그림 5 BIM 기반 건설현장 영상데이터 생성기술 활용 시나리오

4. 결론

현재의 건설현장 모니터링 기술은 데이터 부족과 같은 제약에 직면하고 있다. 따라서, 건설 환경 디지털화를 위해서는 건설현장 데이터 부족 문제를 극복하고 학습용 데이터의 다양성 및 확장성을 보장하는 기술적 발전이 필수적이다. 이미 주목받고 있는 BIM 모델과 생성형 AI는 이러한 기술적 발전을 이끌 수 있는 주요 도구로 인식되고 있으며 특히, 건설 환경 영상 데이터베이스 구축 기술은 시각 장면 인식 기술의 혁신을 이룰 것으로 기대된다.

딥러닝 및 인공지능 기술의 발전에 기여한 ImageNet, MS COCO와 같은 방대하고 정제된 데이터셋은 건설 분야에서 혁신을 이끌 수 있다. 건설현장 영상데이터 생성기술이 다양한 딥러닝 과제를 수행할 수 있는 정제된 데이터를 생성할 수 있다면, 이는 건설현장 모니터링 분야에서 인공지능 기술과의 융합을 통한 혁신을 기대할 수 있는 중요한 도구가 될 것이다.

또한, 건설현장의 영상데이터 수집과 저장, 정제에 따른 노동력과 비용의 부담은 건설 환경 디지털화의 지연 요인 중 하나이다. 학습용 데이터 생성기술의 발전은 이러한 경제적 비용을 효과적으로 저감할 수 있으며, 노동 집약적인 산업 구조를 개편하고 기술 첨단화를 통한 효율화로 건설산업의 생산성과 안정성을 향상시킬 수 있다. 이러한 변화는 국민 인식의 변화를 불러 일으키며, 건설산업의 발전과 함께 국가 경제에 긍정적인 기여를 할 것으로 기대된다.

참고문헌

1. 과기정통부, 16일 ‘디지털 뉴딜 사업설명회’ 개최, 2020; <https://www.irobotnews.com/news/articleView.html?idxno=21083>
2. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.
3. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13 (pp. 740-755).

- Springer International Publishing.
4. Duan, R., Deng, H., Tian, M., Deng, Y., & Lin, J. (2022). SODA: A large-scale open site object detection dataset for deep learning in construction. *Automation in Construction*, 142, 104499.
 5. Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., & Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. *Automation in Construction*, 122, 103482.
 6. Kim, J., & Chi, S. (2021). A few-shot learning approach for database-free vision-based monitoring on construction sites. *Automation in Construction*, 124, 103566.
 7. Soltani, M. M., Zhu, Z., & Hammad, A. (2016). Automated annotation for visual recognition of construction resources using synthetic images. *Automation in Construction*, 62, 14-23.
 8. Neuhausen, M., Herbers, P., & König, M. (2020). Using synthetic data to improve and evaluate the tracking performance of construction workers on site. *Applied Sciences*, 10(14), 4948.
 9. Assadzadeh, A., Arashpour, M., Brilakis, I., Ngo, T., & Konstantinou, E. (2022). Vision-based excavator pose estimation using synthetically generated datasets with domain randomization. *Automation in Construction*, 134, 104089.
 10. Braun, A., & Borrmann, A. (2019). Combining inverse photogrammetry and BIM for automated labeling of construction site images for machine learning. *Automation in Construction*, 106, 102879.
 11. Alawadhi, M., & Yan, W. (2021). BIM hyperreality: Data synthesis using BIM and hyperrealistic rendering for deep learning. *arXiv preprint arXiv:2105.04103*.
 12. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).
 13. Hong, Y., Park, S., Kim, H., & Kim, H. (2021). Synthetic data generation using building information models. *Automation in Construction*, 130, 103871.
 14. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).
 15. Bang, S., Baek, F., Park, S., Kim, W., & Kim, H. (2020). Image augmentation to improve construction resource detection using generative adversarial networks, cut-and-paste, and image transformation techniques. *Automation in Construction*, 115, 103198.
 16. Baek, F., Kim, D., Park, S., Kim, H., & Lee, S. (2022). Conditional generative adversarial networks with adversarial attack and defense for generative data augmentation. *Journal of Computing in Civil Engineering*, 36(3), 04022001.
 17. Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*. 