

2025년 추계학술발표대회 : 일반부문

지각된 어포던스 기반 건축 인지 AI - 생성형 모델을 넘어 인지 공간 해석으로 -

Perceived Affordance-Based Cognitive AI for Architecture - Beyond Generative Models Toward Spatial Interpretation -

박 준 민* 이 명 식**
Park, Jun-Min Lee, Myungsik-Sik

Abstract

This study addresses a key limitation of image-generative AI: it can make images but cannot read the spatial logic of architectural sketches. We present a cognition-based system, grounded in perceived affordance, that detects walls, doors, floors, and rooms; performs pixel-level semantic segmentation; and classifies elements by function for structural understanding. The pipeline uses CLIP (affordance matching), SAM (masking), Hough Transform (wall extraction), and Shapely/ezdxf (vectorization), trained and evaluated with FCN-ResNet50. It reframes AI as an intelligent tool for architectural analysis, robotics, and AR/VR.

키워드 : 인지형 AI, 지각된 어포던스, 건축 스케치 인식, 시맨틱 세그멘테이션, CLIP, SAM, 허프 변환, 벡터화, 공간 인지

Keywords : Cognitive AI, Perceived Affordance, Architectural Sketch Recognition, Semantic Segmentation, CLIP, Segment Anything Model (SAM), Hough Transform, Vectorization, Spatial Cognition

1. 서론

1.1 연구의 목적

인공지능(AI) 기술이 건축 설계에 혁신을 가져오고 있지만, 기존 이미지 생성 AI는 공간의 구조나 기능적 의미를 이해하지 못하고 시각적 패턴만 모방하는 명백한 한계를 가진다. 이러한 AI는 건축적 맥락에 대한 이해 없이 결과물을 생성하기에, 디자이너에게 시각적 영감을 줄 수 있으나 실제 설계 데이터로 활용되기에는 어려움이 따른다¹⁾. 본 연구는 이를 극복하기 위해 ‘지각된 어포던스(perceived affordance)’²⁾ 개념에 기반하여 ‘인지형 AI’ 개발을 목표로 한다. 제안하는 AI는 건축 스케치의 벽, 문과 같은 요소들을 단순한 시각적 정보가 아닌, ‘공간을 분할’하거나 ‘이동을 가능하게 하는’ 등 기능적 의미를 가진 구조로 인지하고 해석한다. 이를 통해 AI는 스케치를 해석하여 즉시 활용 가능한 DXF 파일로 변환하고, 초기 설계의 디지털 변환을 자동화함으로써 보다 창의적인 디자인 탐색에 집중할 수 있는 환경을 제공한다.

본 연구는 수작업에 의존하는 초기 설계 변환 과정을 자동화하여 설계 효율성을 극대화하고, 향수 빌딩 정보 모델링(BIM) 통합의 핵심 기반을 마련하는데 기여할 것이다.

2. 관련 연구와 이론적 배경

2.1 지각된 어포던스

어포던스는 미국의 생태 심리학자인 제임스 깁슨(James J. Gibson)에 의해서 처음 언급되었으며, 그의 저서 ‘The Ecological Approach to Visual Perception’에서 개념을 확장시켰다³⁾.

이는 환경이 제공하는 객관적 행위 가능성을 넘어, 사용자 경험과 지식을 통해 인지하는 기능적 속성을 말한다. 본 연구의 AI는 이 원리를 적용하여, 건축 스케치 속 추상적 선과 기호를 단순 픽셀이 아닌 ‘공간을 구획하는 벽’, ‘통과를 허용하는 문’과 같이 기능적 의미(어포던스)를 가진 요소로 ‘지각’하고 해석하도록 설계되었다. 이는 AI가 인간의 인지 과정처럼 도면의 의도를 파악하게 하는 핵심 기제이다.

2.2 이미지 분할 및 분류 기술

이미지 분할은 컴퓨터 비전에서 공간적 맥락을 이해하기 위한 핵심 기술이다. 그중 의미론적 분할(semantic segmentation)은 이미지 내의 모든 픽셀을 특정 클래스에 할당하여 동일 범주의 픽셀들을 그룹화한다. 이는 경계 상자(bounding box) 단위로 객체 위치만을 알려주는 객체 탐

*동국대 대학원 석사과정

**동국대 건축학과 교수, 건축공학과 공학 박사

(Corresponding author : Department of Architectural , Dongguk University, mslee@dongguk.edu)

¹⁾ Ciancio, F., &Kalisperis, L. N. (2022). The New Unreal: The Role of AI in Architectural Representation, Proceedings of the 40th eCAADe Conference, 2, 495-502.

²⁾ Gibson, J. J. (1979). The Ecological Approach to Visual Perception, Houghton Mifflin.

본 연구는 국토교통부/국토교통과학기술진흥원의 2025년도 지원으로 수행되었음(과제번호 RS-2021-KA163269).

³⁾ 박준민, 홍일태. (2022). 지각된 어포던스를 통한 공간 분석 - 레지오 에밀리아 공간 이론을 바탕으로 -. 한국실내디자인학회 논문집, 31(6), 67-75. 10.14774/JKIID.2022.31.6.067

지(object detection)와 달리, 픽셀 수준에서 보다 정밀한 윤곽 정보를 제공한다. 또한 이미지 분류(image classification)가 전체 이미지에 단일 레이블만을 부여하여 공간 정보를 상실하는 반면, 의미론적 분할은 공간적 구조를 보존할 수 있다. 인스턴트 분할(instance segmentation)은 동일 클래스 내에서도 개별 개체를 구분하지만, 의미론적 분할은 클래스 단위의 집합적 이해에 초점을 둔다.

전통적인 의미론적 분할 모델로는 FCN(Fully Convolutional Networks), U-Net, DeepLab 등이 대표적이며, 이들은 의학 영상, 자율주행, 건축 도면 인식 등 다양한 분야에서 활용되고 있다.

2.3 건축 스케치 - CAD/BIM 변환

건축 설계는 개념적 스케치에서 시작하여 개략 설계, 설계 개발, 건설 문서화로 이어지는 단계적 과정을 거친다⁴⁾. 이때 스케치를 CAD나 BIM과 같은 정밀하고 구조화된 데이터로 변화하는 과정은 필수적이지만, 여전히 수작업이나 트레이싱에 크게 의존하고 있다. 이러한 방식은 비효율적이며 오류 가능성이 크다. 문제의 본질은 기계가 인간 제도사처럼 스케치의 의도를 인지하고, 이를 편집 가능한 데이터로 변환할 수 있는지에 있다.

최근 AI 기반 도구들이 이러한 한계를 보완하며 설계 워크플로우에 통합되고 있다⁵⁾. 특히, 2D 평면도를 3D모델로 변화하는 툴이 등장했으나, 비정형적 스케치나 노이즈가 많은 이미지에는 여전히 취약하다. 그림자, 주석, 치수선, 불완전한 벽체 등은 AI의 인식을 방해하며, 방 해석 오류를 유발하기도 한다. 따라서 효과적인 변환을 위해서는 최소한 핵심 요소(벽, 문, 창문 등)만 포함된 구조화된 입력이 필요하다.

본 연구에서 제안하는 파이프라인은 이러한 조건을 만족하는 DXF파일을 자동 생성하여 스케치와 BIM사이의 간극을 매우는 중간 표현을 제공한다. 이를 통해 AI기반 설계 자동화, 디지털 트윈, 충돌 감지 등 BIM응용이 가능해지며, 초기 단계에서부터 AI가 건축가의 협력적 파트너로 작동할 수 있는 기반을 마련한다.

3. 지각된 어포던스 기반의 인지 AI 설계

3.1 전체 시스템 구성 및 단계별 프로세스

본 연구의 목표는 건축 스케치 이미지를 지각된 어포던스(perceived affordance) 개념에 기반하여 분할 분류 하고, 구조적으로 이해할 수 있는 인지형 AI를 개발하는 것이다. 연구는 Part1부터Part11까지 단계적으로 파이프라인 구축 과정을 거쳤으며, 각 모듈은 지각된 어포던스의 개념을 구현하는 데 기여한다.

4) Eastman, C., Teicholz, P., Sacks, R., &Liston, K. (2011). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. Wiley.

5) Liu, X., Xu, K., &Zheng, Y. (2021). "Automatic generation of 3D building models from architectural floor plans." *Automation in Construction*, 126, 103679.

표1. 파이프라인 구축 과정

구분	내용
PART1	CLIP 기반 affordance 매칭 (wall, door 등 단일 요소 분류 실험)
PART2	계획 생성기 (도면 + 요구 기능 → 작업 순서 API)
PART3	작업 순서 시각화 (Plotly 시퀀스 다이어그램)
PART4	대화형 인터페이스 (FastAPI + React 기반 질의응답 UI)
PART5	CLIP 성능 개선 (프롬프트 튜닝, 부정/긍정 샘플링)
PART6	윤곽선 기반 벡터화 (Shapely + ezdxf 활용, DXF 생성)
PART7	허프 변환 기반 직선 추출 → 벽체 구조화 및 벡터 변환
PART8	SAM 기반 마스크 분할, COCO 포맷 dataset.json 구축
PART9	FCN-ResNet50 학습 (image-mask 대응, 리사이즈/정합성 처리 포함)
PART10	테스트셋 구축
PART11	evaluate_test_simple.py로 검증

3.2 데이터 전처리 및 마스크 생성

첫 단계는 스케치에서 잠재적 공간 요소를 ‘지각’ 하는 과정이다. 이를 위해 Segment Anything Model(SAM)을 활용하여 자동 마스크를 생성하였다⁶⁾. SAM은 자연 이미지에서 강력한 제로샷 성능을 보이지만, 얇고 모호한 선이 많은 건축 스케치에는 그대로 적용하기 어렵다. 따라서 매개변수를 조정하여 정밀도보다 재현율을 우선시켰으며, 가능한 모든 잠재적 요소를 포괄하도록 설계하였다. 이후 생성된 마스크는 COCO 포맷으로 정리되어 학습 데이터셋으로 활용되었다.

데이터 부족 문제를 해결하기 위해 회전, 뒤집기, 색상 변화 등의 데이터 증강 기법을 적용하였다. 이는 다양한 드로잉 스타일을 반영하여 모델이 일반화할 수 있도록 돕는다. 또한 SAM 출력에서 발생하는 중복 모호 마스크는 combine_masks.py 스크립트를 통해 통합 정제하여, 후속 학습에서 안정적인 입력으로 사용되었다.

3.3 어포던스 기반 구조 인식 모델 설계

다음 단계는 분할된 요소를 실제 건축적 의미로 ‘해석’ 하는 과정이다.

표2. 지각된 어포던스 기반 해석

구분	내용
CLIP기 어포던스 매칭	CLIP을 활용하여 ‘wall, door, window, room’ 과 같은 텍스트 클래스와 분할된 이미지 마스크를 매칭하였다. classify_mask 함수는 각 마스크에 대한 텍스트-이미지 유사도를 산출하여 가장 높은 점수의 클래스를 할당한다. 점수가 낮을

6) Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... &Girshick, R. (2023). "Segment Anything." *arXiv preprint arXiv:2304.02643*.

	경우 'uncertain'으로 처리해 후속 보정 가능성을 열어두었다.
하프 변환 기반 직선 추출	벽체 구조의 인식을 위해 하프 변환을 적용하여 직선을 추출하고, 이를 벡터형태로 변환하였다. 이는 스케치의 벽을 기하학적으로 구조화하는데 기여하였다.
DXF 벡터화 및 보내기	mask_to_polygons 함수와 ezdxf 라이브러리를 활용하여 분류된 마스크를 CAD에서 활용 가능한 벡터 데이터로 변환하였다. 이 단계는 AI의 인식 결과를 실제 설계 워크플로우에 연결하는 핵심 과정이다.

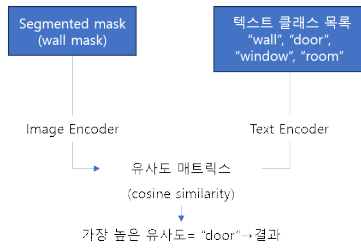


그림1. CLIP기 어포던스 매칭

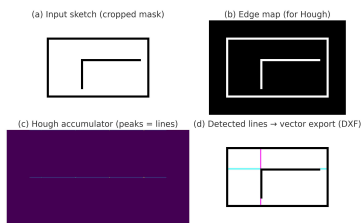


그림2. 하프 변환 기반 직선 추출

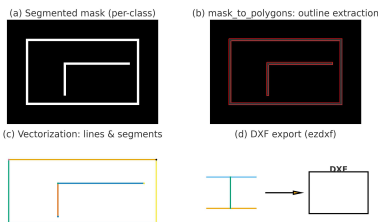


그림3. DXF 벡터화 및 보내기

3.4 세그멘테이션 학습 및 평가 파이프라인

SAM 과 CLIP의 제로샷 분류만으로 한계가 있으므로, FCN-ResNet50 기반의 지도 학습을 추가로 수행하였다. COCO 포맷 데이터셋을 이용해 ‘벽,문,창문,방,배경’을 학습시켰으며, 데이터 증강, 마스크 통합, 학습률 스케줄링(stepLR), Adam 옵티마이저 등 최적화 기법을 적용하였다.

4. 실험 및 결과 분석

4.1 실험 개요

본 연구에서 제안한 파이프라인은 건축스케치 이미지에서 공간 요소를 인식분류하고 구조적으로 해석하는 능력을 검증하기 위해 일련의 실험을 수행하였다. 데이터셋

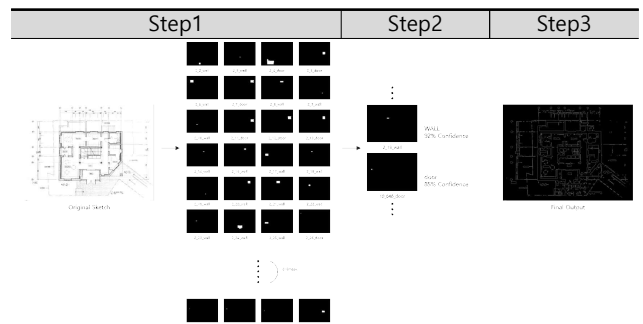
은 SAM을 통한 마스크 생성 및 통합 과정을 거쳐 구축되었으며, FCN-ResNet50 모델을 중심으로 학습과 평가가 진행되었다.

4.2 기존 이미지 생성형 AI와의 비교

제안된 모델 Stable Diffusion, DALL-E등과 같은 기존 이미지 생성형 AI와 목적과 접근 방식에서 본질적으로 차별된다. 생성형 AI는 텍스트 프롬프트를 기반으로 창의적 이미지를 “새롭게 만들어내는 것”에서 초점을 둔다. 반면 본 연구의 인지형 AI는 입력된 건축 스케치 내의 벽, 문, 창문, 방과 같은 공간 요소를 ‘지각된 어포던스(perceived affordance)’ 개념에 따라 인식하고, 기능적 의미를 해석하는 데 중점을 둔다.

따라서 제안된 AI는 단순한 이미지 생성 도구가 아니라, 공간적 맥락을 이해하고 구조를 인지하는 시스템으로서 기존 모델과의 차별성을 갖는다. 이는 건축 설계 및 분석 과정에서 AI를 협력적 파트너로 확장 할 수 있는 가능성을 제시한다.

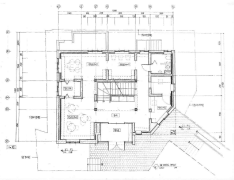
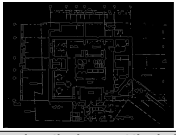
표3. 스케치 분석을 통한 분석 과정



상세설명	
구분	내용
Step1: 원본 스케치와 AI의 최초 인식	AI가 스케치에서 의미를 가질 수 있는 모든 잠재적 요소를 지각하는 단계이며, 노이즈가 많고 정제되지 않은 모습이다.
Step2: 지각된 어포던스 기반의 의미 부여(CLIP)	AI가 단순히 선을 인식하는 것을 넘어, ‘이것은 공간을 나누는 벽’, ‘저것은 통로를 가능하게 하는 문’이라는 기능적 의미(지각된 어포던스)를 부여하는 과정이다. 이는 인간 설계자가 도면을 이해하는 인지 과정과 유사하게 하려면 시도가 있다
Step3: 구조적 형태의 재구성(하프 변환 및 벡터화)	AI는 인식한 의미를 바탕으로 불규칙한 손그림 선을CAD에서 편집 가능한 정형화된 기하학적 재구성합니다. 이는 설계 초기 아이디어를 BIM과 같은 정밀 데이터로 연결하는 핵심적인 다리 역할을 한다.

〈표3〉원본 스케치(왼쪽)로부터 벽,문 등 건축 요소가 각각의 마스크로 분해(중앙)되며, 총 648개의 mask가 형성된다. 이후 마스크들은 재조합 과정을 거쳐 cad도면(오른쪽)으로 변환된다.

표4. 연구 결과 분석

원본	
	
지각된 어포던스 AI 결과	기존 생성형 AI (Raster to Vector)
	
건축적 맥락을 '해석'하여 변환 단순히 선을 따라서 그리는 것이 아닌 '벽'은 공간을 분할하고 '문'은 이동을 가능하게 하는 등 요소의 기능(지각된 어포던스)을 먼저 인지한다.	모든 요소를 '무분별하게' 인식 스케치의 모든 픽셀을 의미 구분 없이 동일한 선으로 취급하여 기계적으로 추적한다.
정보의 구조화 및 정제 건축적 요소(벽 등)는 기하학적으로 반듯한 벡터선으로 구조화 한다. 불필요한 구조적 (텍스트), 치수선 등은 건축요소와 분리하여 처리한다.	정체되지 않는 데이터 벽체, 치수선, 주석, 노이즈까지 모두 뒤섞인 벡터선으로 변환된다. 뒤섞인 벡터선으로 변환된다. 정리되어 있지 않아 후속 작업이 필수적이다.
결과 설계자의 의도가 반영된, cad가 생성됨	결과 단순한 디지털 복사본

4.3 성능 평가

테스트셋 기반으로 모델 성능을 검증한 결과, 최종적으로 **픽셀 정확도 75.14%**와 **평균 교차합(mIoU) 40.68%**를 달성하였다. 이는 모델이 공간 요소의 존재 여부를 비교적 안정적으로 인식하고 있음을 보여주며, 동시에 경계 분할 정밀도에서는 개선 여지가 있음을 시사한다.

표5. 성능 평가 테스트 결과

지표	값
Pixel Accuracy	75.14
mIoU	40.68

특히 벽의 모서리, 문과 벽의 접합부, 가구와 벽체의 구분 등 복잡한 경계 영역에서 오류가 관찰되었다. 이는 건축 스케치가 지닌 얇고 불명확한 선, 다양한 스타일적 표현 방식으로 인한 도전 과제에서 기인한다.

5. 결론

본 연구는 기존 이미지 생성 AI가 갖는 공간 해석 능력의 한계를 극복하고자, ‘지각된 어포던스’ 개념에 기반한 인지형 AI를 개발했다. 제안된 AI는 SAM, CLI, 허프 변

환 등을 결합한 파이프라인을 통해 건축 스케치를 기능적, 구조적으로 해석하고, 이를 CAD에서 활용 가능한 DXF파일로 자동 변환한다.

실험 결과 픽셀 75.14%, mIoU 40.68%를 달성하여 건축 스케치의 다양한 스타일 속에서도 공간 요소를 안정적으로 인식할 수 있는 가능성을 확인했다. 본 연구의 가장 큰 의의는 AI를 단순한 이미지 생성 도구가 아닌, 설계자의 초기 아이디어를 이해하고 디지털 설계 과정으로 자동 연결해주는 지능적인 협력 파트너로 재정의했다는 점에 있다.

본 연구의 AI는 초기 단계에서 의미 있는 성능을 보였으나 다음과 같은 한계가 존재한다.

표5. AI의 한계에 대한 연구 보완사항

구분	내용
세그멘테이션 정밀도	미세한 경계 요소에서 성능저하
스타일 다양성	학습 데이터에 없는 드로잉 스타일에 대한 일반화 한계
어포던스 해석 모호성	단순 선 요소가 벽인지 판다 어려움
확장성 문제	다층 건물, 대규모 도면에서대한 적용 한계
연상 효율성	고해상도 이미지 처리 시 실시간 적용의 제약

향후에는 데이터셋 규모 확대, 다중 모달 학습(테스트 도면 병행), 경량화 네트워크 적용을 통해 이러한 한계점을 극복할 수 있을 것으로 기대된다.

참고문헌

- Ciancio, F., &Kalisperis, L. N. (2022). The New Unreal: The Role of AI in Architectural Representation, Proceedings of the 40th eCAADe Conference, 2, 495-502.
- Gibson, J. J. (1979). The Ecological Approach to Visual Perception, Houghton Mifflin.
- 박준민, 홍일태. (2022). 지각된 어포던스를 통한 공간 분석 - 레지오 에밀리아 공간 이론을 바탕으로 -. 한국실내디자인학회 논문집, 31(6), 67-75. 10.14774/JKIID.2022.31.6.067
- Eastman, C., Teicholz, P., Sacks, R., &Liston, K. (2011). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. Wiley.
- Liu, X., Xu, K., &Zheng, Y. (2021). "Automatic generation of 3D building models from architectural floor plans." *Automation in Construction*, 126, 103679.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... &Girshick, R. (2023). "Segment Anything." *arXiv preprint arXiv:2304.02643*.