



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

공학석사 학위논문

스타일 트랜스퍼를 활용한 평면도  
해석 프로세스의 범용성 강화

Increasing Generality of Floorplan Analysis  
Process using Style Transfer

2019 년 2 월

서울대학교 대학원

건설환경공학부

김 성 용

스타일 트랜스퍼를 활용한 평면도  
해석 프로세스의 범용성 강화

Increasing Generality of Floorplan Analysis  
Process using Style Transfer

지도 교수 김 용 일

이 논문을 공학석사 학위논문으로 제출함  
2018 년 12 월

서울대학교 대학원  
건설환경공학부  
김 성 용

김성용의 공학석사 학위논문을 인준함  
2018 년 12 월

위 원 장 \_\_\_\_\_ (인)

부위원장 \_\_\_\_\_ (인)

위 원 \_\_\_\_\_ (인)

## 초 록

최근 기술의 발달로 위치 추적, 내비게이션 등의 위치 기반 서비스가 실내까지 확장되어 관련 시장의 규모가 증가하는 추세이다. 실내 관련 서비스 및 연구에 대한 수요의 증대로 인해 그 기반이 되는 실내공간정보를 구축하는 것이 연구자들의 주요 관심사로 부상하고 있다. 그 중 도면을 활용한 접근은 간편하게 획득할 수 있고 접근성이 높다는 특징으로 인해 실내 정보의 효율적인 구축 방법으로 활용되고 있다. 이를 위한 기술적 측면의 쟁점은 수집된 다양한 포맷의 도면들에 적용할 수 있도록 하는 범용성의 확보이다. 도면 해석 분야는 이러한 흐름에 발맞춰 학습 알고리즘을 도입하여 새로운 포맷에도 기술의 적용이 가능한 방향으로 발전해왔다. 그러나 학습 기반의 도면 해석 프로세스는 학습 데이터를 구축하는 작업이 어렵고, 비교적 단순한 포맷의 도면들을 대상으로 수행되었기 때문에 실질적으로 기술을 활용하는데 한계가 있었다. 본 논문은 도면 해석 기술을 활용하기 위한 범용성의 확장을 목적으로 딥러닝 기반의 스타일 트랜스퍼를 적용하였다. 도면의 다양한 포맷에 대해서 실내 구조의 표현을 위한 다양한 스타일로 접근하여, 실내 구조를 잘 표현하도록 다양한 도면을 일관된 스타일로 변환시키며 필요한 정보를 출력하도록 딥러닝 네트워크를 설계하였다. 이러한 접근을 기반으로 학습 데이터셋을 구축하기 위한 방법과 이를 활용해 실내 구조를 구현하는 방법을 제시하여 도면 해석 프로세스를 개발하였다. 연구의 대상으로는 기존 공용 도면 데이터셋들에 비해 복잡하고 다양한 도면들로 구성된 세움터 도면 데이터셋을 새로 구축하여 사용하였으며, 개발한 해석 프로세스를 통해 구현한 벡터 포맷의 실내 정보를 기반으로 검증을 수행하였다. 기존 연구와의 비교를 통해 스타일 트랜스퍼의 적용이 성능을 향상시킴을 보였고, 방의 탐색률과 인지 정확도에서 각 87%와 85%의 결과를 얻어 단순한 포맷에 대한 이전의 연구들의 성능과 유사한 수준임을 확인하였다. 정리하면, 본 논문을 통해 복잡하고 다양한 포맷으로 구성된 도면 데이터셋에 대한 해석이 가능해졌고 이는 도면 해석 기술의 적용 범위를 확장시킨다는 의의가 있다.

**주요어** : 실내공간정보, 도면 해석, 벡터화, 스타일 트랜스퍼, 정수 계획법  
**학 번** : 2017 - 23581

# 목 차

제 1 장 서 론 .....	1
제 1 절 연구의 배경 .....	1
제 2 절 연구의 목표 .....	5
제 3 절 연구 흐름도 .....	8
제 2 장 관련 연구 .....	10
제 1 절 도면 데이터셋 .....	10
제 2 절 도면 해석의 연구 동향 .....	14
제 3 절 학습 알고리즘의 확장 .....	18
제 3 장 도면 데이터셋 구축 .....	20
제 1 절 세움터 도면 .....	20
제 2 절 실내 구조의 라벨링 .....	22
제 4 장 도면의 벡터화 프로세스 .....	26
제 1 절 스타일 트랜스퍼의 적용 .....	27
제 2 절 정수 계획법을 통한 보완 .....	32
제 5 장 결과 및 정확도 검증 .....	36
제 1 절 정확도 검증 방법론 .....	36
제 2 절 결과 및 검증 .....	39
제 3 절 도면 해석 프로세스의 딥러닝 활용 .....	43
제 6 결론 .....	45
참고문헌 .....	47

## 표 목차

[표 3-1] 세움터 도면 데이터셋 라벨링 예시.....	24
[표 4-1] 딥러닝 네트워크의 세부 구조.....	28
[표 5-1] 손실 함수 차이를 통한 비교군 설정 .....	39
[표 5-2] 모서리 정보 추출의 정확도 평가.....	39
[표 5-3] 도면 해석을 통한 실내 구조 구현 결과.....	41
[표 5-4] 방의 탐색률 및 인지 정확도 .....	41
[표 5-5] 학습 데이터 수에 따른 정밀도 및 재현율 .....	44

## 그림 목차

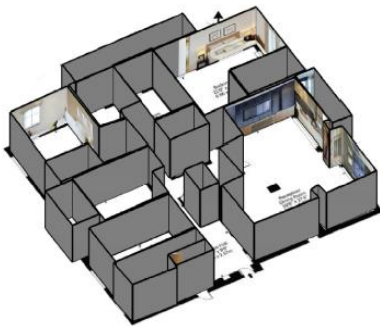
[그림 1-1] 실내공간정보 활용 예시.....	1
[그림 1-2] 사용자 참여형 데이터로서 도면의 활용 .....	2
[그림 1-3] 일반적인 도면 해석 프로세스.....	3
[그림 1-4] 본 논문의 도면 해석 프로세스.....	8
[그림 2-1] CVC 도면 데이터셋 예시 / 명칭(도면의 수) .....	11
[그림 2-2] Rakuten 도면 데이터셋 예시.....	13
[그림 2-3] 도면 해석 프로세스에서 학습 알고리즘의 확장 .....	18
[그림 3-1] 세움터 도면 데이터셋 예시.....	21
[그림 3-2] 실내 구조 표현을 위한 라벨링 .....	22
[그림 3-3] 세움터 도면 대상의 벽체 및 제안한 라벨링 예시.....	23
[그림 3-4] 모서리 클래스의 표현 (I, L, T, X, O).....	24
[그림 4-1] 도면의 벡터화 프로세스.....	26
[그림 4-2] 딥러닝 네트워크의 구조.....	27
[그림 4-3] 정수 계획법을 통한 딥러닝 출력물의 보완 .....	32
[그림 4-4] 후처리를 통한 벡터 포맷의 실내 구조 구현.....	35
[그림 5-1] CNN과 스타일 트랜스퍼의 정성적 비교.....	40
[그림 5-2] 학습 데이터 수에 따른 재현율 변화.....	43

# 제 1 장 서 론

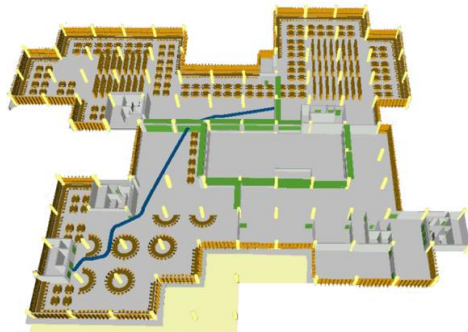
## 제 1 절 연구의 배경

최근 사물인터넷을 비롯한 기술의 발전으로, 실내에서 위치 추적이 가능해졌을 뿐만 아니라 네트워크, 내비게이션 등의 위치 기반 서비스 역시 실내로 확장되고 있다. 이에 따라 실내공간과 관련한 서비스 및 연구에 대한 수요가 급증하고, 그 기반이 되는 데이터베이스, 즉 실내공간정보를 구축하는 것이 연구자들의 주요한 관심사로 부상하고 있다. 실내공간정보는 IFC<sup>①</sup>, CityGML<sup>②</sup> 등의 표준 포맷을 통해 실내의 기하학적인 형상, 구성 요소들의 위상 관계, 의미론적 속성 등의 정보를 표현하는 데이터베이스를 일컫는다. 이러한 실내공간정보는 관련 연구나 서비스의 구체적인 목적에 따라 다양한 수준에서 구축되어왔다.

예컨대, Liu *et al.* (2015)은 표면 수준으로(surface-level) 구축된 실내공간에 텍스처 형태로 사진을 매핑했고, Chu *et al.* (2015)과 Wang *et al.* (2015)은 벽체 수준으로 구축된 실내공간에서 사진을 촬영한 위치를 특정하고 포인트 클라우드를 기반으로 한 내비게이션에 관한 연구를 수행했다. 또한, Lin *et al.* (2013)은 방과 출입구의 위상 수준으로 구축된 실내공간정보를 기반으로 실내 네트워크의 연구를 수행했다(그림 1-1). 이때, 기반이 되는 실내공간정보는 사진, 포인트 클라우드 등의 이중의 데이터와 연계되기 위해 주로 방의 구분하는



(a) From Liu *et al.* (2015)



(b) From Lin *et al.* (2013)

그림 1-1. 실내공간정보 활용 예시

① <http://www.buildingsmart-tech.org/>

② <https://www.citygml.org/about/>

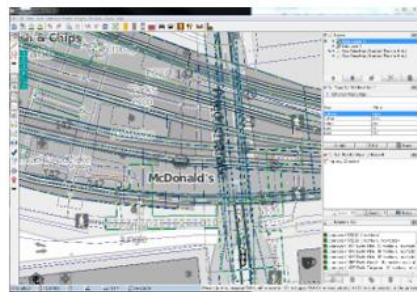
정도로 단순화되어 표현된다. 이를 통해 알 수 있듯이 실내공간정보는 실내 구조를 표현함으로써 다양한 연구의 기반이 되는 데이터로 활용되고 있다.

실내공간정보를 구축할 때 드론 영상, 3D 레이저 스캐닝, CAD 플랜, 도면 등의 다양한 재료를 사용할 수 있다. 이 때, 사용하는 재료의 특성에 따라 연구의 접근 방향과 결과물은 상이한 양상을 띤다(Gimenez *et al*, 2015). 본 연구는 도면에 초점을 맞추어, 우선 도면을 기반으로 실내공간정보를 구축하는 방식의 특징을 검토하고자 한다. 도면에 내재한 수많은 정보 중 어느 수준까지 재현할 것인지 택하는 것이 연구의 시작이며, 이는 연구의 목표와 도면의 포맷에 의존한다. 도면을 활용한 실내공간정보 구축의 쟁점은 형태가 유사한 CAD 플랜과 비교할 때 보다 분명하게 드러난다. CAD 플랜을 통한 접근은 벡터 형태의 정보를 기반으로 온전한 정보를 구축하는데 중점을 둔다. 반면, 도면을 통한 접근은 래스터화되어 의미가 사라진 정보를 복원하는 것을 핵심으로 한다. 도면은 다른 재료들에 비해 상대적으로 적은 비용으로 쉽게 구할 수 있고 개방되어 있다. Open Street Map, Google Map 등은 도면을 사용자 참여형 데이터로 활용하여, 도면을 직접 입력하고 수작업으로 실내 지도를 구축할 수 있도록 하는 플러그인을 제공하기도 한다(그림 1-2).

그러나 실내공간 구축 시 사용될 수 있는 도면은 그 포맷의 스펙트럼이 매우 넓어, 도면 해석의 어려움을 가중시킨다는 단점이 있다. 일반적인 건축 도면 외에도 인터넷 도면, 부동산 도면, 대피 안내도 등 실내 구조를 파악할 수 있는 래스터 형태의 도면 이미지들은 모두 도면 해석의 대상이 된다. 그런데 건축 도면은 통념과 달리 표준적인 포맷이



(a) Google Indoor Map



(b) Open Street Map

그림 1-2. 사용자 참여형 데이터로서 도면의 활용



고정되어 있지 않고 작성 기관이나 활용 목적 등에 따라 상이한 포맷으로 작성되고 있다. 특히 부동산 도면과 대피 안내도의 경우에는 색 정보가 활용되고 실내 구조의 표현 방법도 다르다.

요약하자면 도면에 기반해 실내공간정보를 구축하는 방식은 구축된 정보의 완결성보다는 구축 과정의 효율성을 중시하고 재료의 접근성이 높아 사용자 참여형으로 활용될 여지가 있는 한편, 활용되는 포맷이 제각각이라는 단점이 있다. 결과적으로 ‘도면 해석 프로세스를 서로 다른 수많은 포맷들에 적용할 수 있도록 한다’는 의미에서 기술의 범용성을 확보하는 것에 대한 사회적 수요가 존재한다.

도면 해석(floorplan analysis)이란 래스터 도면에서 사라진 기하 및 위상, 의미론적 정보들을 재현하여 실내 공간을 해석하는 연구 분야이다. 일반적인 프로세스는 그림1-3을 따른다(Gimenez *et al*, 2016). 래스터 도면에 대해 불필요한 정보를 제거하는 전처리를 수행한 후, 패턴 인식을 통해 기하학적인 특성을 기반으로 객체 후보군을 추출하고, 객체 후보군에 대한 구조 인식을 통해 객체 간의 위상 및 의미론적 정보를 구현하여 실내 구조를 완성한다.

기존의 도면 해석 연구들은 도면 포맷의 지역적인 특징을 활용하여 도면 해석 프로세스의 성능을 향상시키는 방식을 채택해왔다. 프로세스의 각 단계에서 특정한 도면 포맷에 의존적인 규칙을 결정하고 이에 기반하여 정보를 정제하는 것이다. 본 논문에서는 이러한 연구들을 ‘규칙 기반 프로세스’로 분류하고자 한다. 각 단계별로 널리 사용되어온 알고리즘들을 살펴보면 다음과 같다. 전처리 단계에서는 이진화(Otsu,

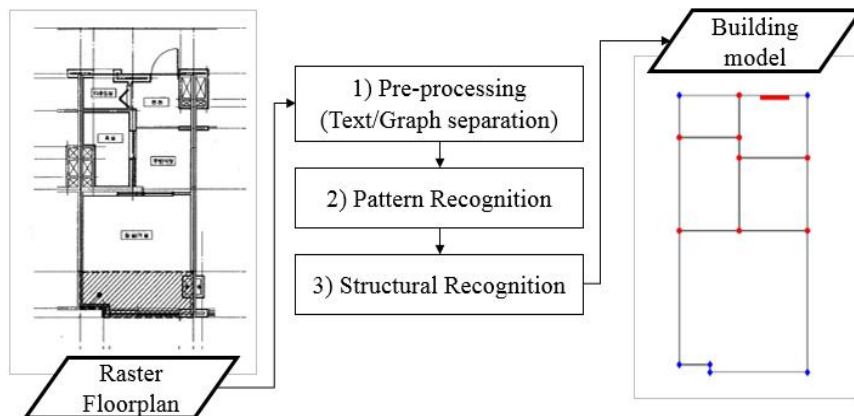


그림 1-3. 일반적인 도면 해석 프로세스

1979) 및 문자 요소의 분리(Tombre *et al*, 2002)가 수행되고, 패턴 인식 단계에서는 문자가 분리된 그래픽 요소에 대해 허프 변환(Hough transform) 기반의 알고리즘을 적용하여 객체 후보군들을 추출한다. 구조 인식 단계에서는 포맷의 특징을 활용한 규칙을 기반으로 벽체, 문 등의 의미론적인 속성 정보를 부여하여 건물 모델을 완성한다(Macé *et al*, 2010; Ahmed *et al*, 2011; Gimenez *et al*, 2016). 그러나 이러한 규칙 기반 프로세스는 도면 포맷의 특수한 성질을 기반으로 하기 때문에, 연구 대상인 특정 포맷에만 제한적으로 적용이 가능하고 포맷의 변화에 대응하지 못한다는 한계를 지닌다.

이를 극복하기 위해 도면 해석 분야는 ‘학습 기반 프로세스’로 분류할 수 있는 새로운 프로세스를 개발하였다. 프로세스에 학습 알고리즘을 도입하여, 학습을 위한 정답 데이터가 마련되어 있는 한 ‘새로운 포맷에도 해당 프로세스를 적용할 수 있게 되었다’는 의미에서 프로세스의 범용성을 다소 확장시켰다. de las Heras *et al*. (2014)은 도면의 분할된 패치에 대한 영상 기술자 기반의 학습을 수행하였고, 서로 다른 4가지 포맷의 도면을 대상으로 각각의 학습을 통해 동일한 접근 방식이 작동하는 것을 보였다. 연구자들은 이미지 처리를 위한 학습 알고리즘 중 뛰어난 성능을 보이는 딥러닝을 도면 해석에 도입하기 시작했다. Dodge *et al*. (2017)은 딥러닝을 통해 벽체의 세그멘테이션을 수행했고, 건축 도면과 부동산 도면 각각을 대상으로 도면 해석 프로세스를 진행했다. Liu *et al*. (2017)은 부동산 도면을 대상으로 딥러닝을 통해 벡터 포맷의 실내 구조를 구현했다. 이러한 학습 기반 프로세스는 도면 해석 기술의 적용성을 일부 확장한 것으로 볼 수 있다.

이 절에서는 실내공간정보를 구축하는 도면 해석 기술에 대한 사회적 요구와 기술의 발전 방향을 살펴보았다. 실내 관련 서비스에 대한 수요가 증가하면서 그 기반이 되는 실내공간정보를 구축하는 것이 중요해지고 있으며, 특히 도면을 활용한 접근은 효율적으로 다양한 상황에 적용할 수 있는 범용적인 기술 개발을 핵심으로 한다. 이러한 범용성의 측면에서, 래스터 도면의 소실된 정보를 일정 수준으로 재현하기 위한 도면 해석 기술은 특정 도면 포맷에 의존적이라는 문제가 있다. 이를 극복하기 위해 도면 해석 기술은 규칙 기반 프로세스에서 학습 기반 프로세스로 전환되는 경향을 보이며, 이는 범용적인 실내공간정보 구축 기술에 대한 사회적인 요구에 상응한다고 할 수 있다.

## 제 2 절 연구의 목표

연구의 목표를 다루기에 앞서, ‘범용성’의 의미를 명확히 할 필요가 있다. 도면 해석 기술의 범용성은 두가지 측면에서 논의될 수 있다. 사회적 수요에 부합하는 범용성은 넓은 스펙트럼의 도면 포맷을 활용한 실내공간정보 구축의 의미를 가진다. 반면에 현재 기술적인 관점에서의 범용성은 새로운 포맷의 도면에도 유연하게 적용 가능한 도면 해석 프로세스의 확장을 의미한다. 이러한 두 가지 의미는 서로 관련을 맺고 있으며 기술이 발전해 나가면 중국에는 사회적 수요를 만족시킬 수 있게 된다.

일반적으로 규칙 기반의 프로세스는 새로운 포맷에 대응하지 못하고, 학습 기반의 프로세스의 경우에도 학습에 활용된 도면의 포맷에 대해서만 그 성능이 보장된다. de las Heras *et al.* (2015)은 영상 기술자 기반의 학습을 통한 도면 해석 프로세스를 다른 포맷에 적용할 경우에 통계적으로 유의미한 정도의 성능 감소가 나타남을 보였다. Dodge *et al.* (2017)은 건축 도면과 부동산 도면의 상이한 포맷을 대상으로 딥러닝 기반의 도면 해석을 진행하여, 학습에 활용되지 않은 포맷에 대해서는 성능이 크게 감소하는 것을 확인했다. 즉, 현재의 도면 해석 기술은 기본적으로 학습에 활용된 특정 포맷에 대해서만 기술의 적용이 가능하며, 이에 따라 본 연구에서는 기술적 측면에 초점을 맞춰 더 다양한 개별 포맷에도 적용할 수 있는 도면 해석 프로세스의 제안의 관점에서 범용성을 논의하고자 한다.

도면 해석 분야는 학습 알고리즘을 도입하여 새로운 포맷에 적용 가능한 방향으로 발전해 왔으나, 현재의 기술 수준에서는 여러 한계점들이 존재한다. 본 논문에서는 그 중에서도 범용성의 측면에서 다음의 두 가지 문제점에 초점을 맞추고자 한다. 첫 번째는 학습을 위한 도면 데이터셋의 구축이 어렵다는 점이고, 두 번째는 복잡한 도면에 대해서는 해석 성능이 검증되지 않았다는 점이다.

우선, 도면 해석 프로세스의 개발과 실제 기술의 활용에는 큰 갭이 있고, 이는 기술을 적용하려 하는 도면의 포맷에 대응되는 학습 데이터를 새로 구축해야 하기 때문에 발생한다. 최근 딥러닝이 도입되면서 필요한 학습 데이터의 양이 기하급수적으로 증가했고, 이는 도면 해석 기술을 활용하기 위해서는 수백 장의 도면을 수집한 뒤

라벨링을 수행하여 정답 데이터를 구축하는 작업이 선행되어야 함을 의미한다. 목표 도면과 유사한 포맷으로 구성된 도면들을 수집하는 것, 이를 바탕으로 정답 데이터를 구축하는 것 모두 기술 적용의 어려움으로 작용한다. 특히 도면의 정답 데이터를 구축하는 작업과 관련하여 많은 연구자들이 어려움을 토로한 바 있다. Macé *et al.* (2010)은 방을 구분하기 위해 벽체 정보 뿐만 아니라 색과 문자로 표현되는 방의 정보까지 활용해야 했고, 심지어 구분이 불가능한 경우도 존재하여 많은 경우 자의적인 판단에 의존했다고 하였다. de las Heras *et al.* (2014)에 따르면 한 장의 도면의 정답 데이터 생성을 위해 전문가가 30분 가량을 소모 하였고, Kim *et al.* (2018)은 같은 도면에 대해 다른 작업자가 정답 데이터를 생성할 경우 결과물이 달라지는 점을 지적했다. 이처럼 도면의 정답 데이터를 구축할 때 포맷에 대한 전문성이 요구되는 동시에 작업자의 주관을 배제하기 어렵기 때문에, 양질의 학습 데이터셋이 생성되기 어려운 실정이고 이는 기술 적용의 한계점으로 작용한다.

다음으로, 학습 기반 도면 해석 프로세스를 통해 새로운 포맷에 일부분 대응할 수 있게 된 것은 사실이나, 복잡한 도면에 대해서는 프로세스의 성능이 아직 검증되지 않았다. 이는 기존 연구들이 비교적 단순한 포맷의 도면들을 사용해왔기 때문이다. de las Heras *et al.* (2014)이 영상 기술자 기반의 벽체 추출이 특정 포맷의 건축 도면을 대상으로 잘 작동하는 것을 보인 한편, Dodge *et al.* (2017)은 부동산 도면 대상의 벽체 추출에서 영상 기술자 기반의 학습이 딥러닝 기반의 학습에 비해 성능이 낮은 것을 보였다. 요컨대 도면 해석 프로세스의 성능은 도면의 포맷, 학습 알고리즘의 종류와 접근 방법 등에 의존적이다. 기존 연구들은 비교적 단순한 포맷의 도면을 대상으로만 진행했기에, 복잡한 도면에 도면 해석 프로세스를 적용할 경우 그 성능이 보장되지 않는다.

실내공간정보 구축을 위해 현재의 도면 해석 기술을 적용하는 과정에서는 위와 같은 한계가 존재한다. 본 논문은 이러한 문제들을 개선하기 위한 목적으로 딥러닝 기반의 스타일 트랜스퍼를 도입한다. 도면의 포맷들을 실내 구조라는 콘텐츠를 표현하기 위한 다양한 스타일로 접근하여, 딥러닝 네트워크가 도면의 스타일을 변환시키는 과정에서 실내 구조의 정보를 출력하도록 한다. 구체적으로 실내 구조의 표현을 목적으로 다양한 도면들을 단순하고 일관된 포맷으로 변환하는

작업을 수행한다.

딥러닝을 분류에 활용한 기존 연구들 대비 스타일 트랜스퍼의 적용은 다음의 장점들이 있다. 먼저, 정답 데이터의 구축이 간단해진다. 접근 관점이 벽체의 추출이 아닌 스타일의 변환이므로, 기하학적인 형상보다 포맷의 일관성을 기준으로 하여 라벨링을 수행한다. 이러한 접근은 라벨링 기준의 엄밀성을 완화하여 작업 난이도를 낮춰주고, 특히 복잡한 도면들에 대해 정답 데이터의 구축을 간편하게 해준다. 이러한 관점에서 볼 때 실내 구조의 표현을 위한 라벨링 기준을 제시하는 것으로 이해할 수 있다. 두 번째로, 복잡한 도면에 대한 해석 성능을 향상시킨다. 스타일 트랜스퍼를 통한 도면의 해석은 복잡한 포맷의 도면을 간단한 포맷으로 변환시키는 관점으로 접근하여, 분류 기반의 방법들보다 복잡한 도면에 잘 대처할 수 있다. 이러한 장점들로 인해 스타일 트랜스퍼의 적용은 도면 해석 기술의 범용성 측면을 일부 개선시키는 것으로 볼 수 있다.

### 제 3 절 연구 흐름도

본 연구에서 개발하는 도면 해석의 범용적 프로세스의 개요는 다음과 같다(그림 1-4). 복잡하고 다양한 포맷의 도면들에 대해서 건물 요소들을 벡터 형태로 추출하여 실내공간정보를 구현한다. 이때 건물 요소들은 방의 외곽을 표현하는 벽체와 출입문으로, 이는 해당 객체들이 다양한 포맷의 도면들에서 공통적으로 추출 가능한 정보인 동시에 가장 핵심적으로 실내 구조를 표현하기 때문이다. 딥러닝 네트워크가 이미지를 핵심 콘텐츠와 스타일의 두 가지 측면에서 해석한다는 점에 착안하여(Gatys et al, 2016), 딥러닝 기반의 스타일 트랜스퍼를 활용하여 건물 요소들의 벡터화를 수행한다. 도면의 다양한 포맷은 실내 구조를 표현하기 위한 각기 다른 스타일로 정의할 수 있고, 딥러닝 네트워크는 이러한 도면의 스타일을 변환시키며 그 과정에서 실내 구조에 대한 정보를 출력한다.

구체적으로 스타일 트랜스퍼를 활용하여 다양한 도면들을 일관된 포맷으로 변환시키되, 이때 변환되는 포맷은 실내 구조를 잘 드러내는

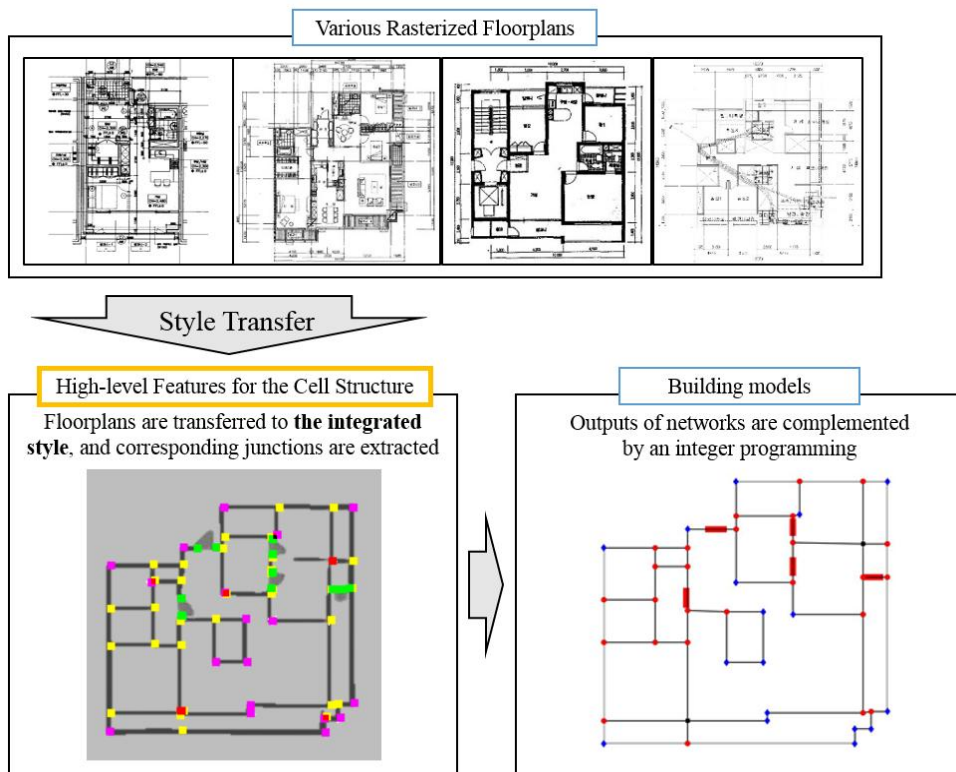


그림 1-4. 본 논문의 도면 해석 프로세스

것으로 선택 한다. 그와 동시에 네트워크는 실내 구조를 각 객체들의 모서리의 형태로 추출하고, 해당 모서리 정보를 기반으로 건물 요소들을 벡터화하여 최종적인 실내공간정보를 구현한다. 본 연구는 크게 학습 도면 데이터셋 구축과 도면 해석 프로세스의 두 단계로 구분되며, 구체적인 연구의 흐름은 아래와 같다.

먼저, 딥러닝 네트워크의 학습을 위한 정답 데이터셋을 구축한다. 기존 연구에 활용되던 도면 데이터셋은 단일 포맷으로 구성되거나 비교적 단순한 구조의 포맷으로 구성되어 있어 본 연구의 목적에 부합하지 않는다. 이에 건축행정시스템 세움터로부터 평면도, 시공 도면, 설비 도면 등의 원형의 도면들을 획득하여, 복잡성과 다양성이 강조된 새로운 도면 데이터셋을 구성하고 정답 데이터의 구축을 진행한다. 이때 라벨링은 실내 구조의 표현을 목적으로, 각 객체의 기하학적인 형상보다 정답 객체의 일관된 형태를 우선시하여 수행한다. 라벨링의 대상은 방의 외곽 벽체와 출입구이고, 각 객체에 대응되는 모서리 정보까지를 수집한다. 모서리 정보는 객체들을 벡터 형태로 구현하기 위한 목적으로 활용되며, 픽셀 위치와 함께 연결성 정보를 표현하는 5가지 클래스를 부여한 것이다. Liu *et al.* (2017)이 활용한 13개의 모서리 클래스와 비교하여 간편한 라벨링이 가능하고, 불완전한 정보는 정수 계획법을 통해 충분히 보완이 가능하다.

다음으로 구축한 도면 데이터셋을 바탕으로 도면의 벡터화 프로세스를 진행한다. 딥러닝 기반의 스타일 트랜스퍼를 통해 다양한 도면을 일관된 포맷으로 변환함과 동시에 모서리 정보를 히트맵의 형태로 출력한다. 각각의 작업을 수행하는 네트워크들은 학습되는 파라미터들을 공유하여 상호 보완적으로 학습되고, 최종적으로 스타일이 변환된 도면과 모서리 정보를 출력한다. 각 출력물은 벡터 형태로 실내 구조를 표현하기에 불완전한 형태이므로, 정수계획법을 활용하여 보완하는 작업이 필요하다. 딥러닝을 통해 출력된 모서리 후보군과 이를 활용하여 만든 벽체 후보군을 정수계획법의 변수로 정의하고, 적절한 목적 함수와 제약 조건을 설정한다. 목적 함수는 스타일 변환된 도면을 기반으로 정의하고, 제약 조건은 도면 특성을 활용하여 [연결성, 상호 배제, 폐합]의 조건으로 설정한다. 정수 계획법을 통해 도면의 특징이 반영되어 딥러닝의 출력물들이 보완되며, 모서리의 연결 정보를 표현하는 인접 행렬이 출력된다. 최종적으로 모서리의 좌표와 인접 행렬을 통해 벽체와 출입문으로 표현되는 벡터 포맷의 실내 구조를 구현한다.

## 제 2 장            관련 연구

### 제 1 절        도면 데이터셋

도면의 특성은 도면 해석 프로세스의 성능에 중요한 비중을 차지하고(de las Heras *et al.*, 2015), 이런 배경에서 연구자들은 공용 데이터셋을 구축하여 도면 해석 프로세스의 개발 및 성능 검증에 활용하고 있다. 도면 해석 프로세스들은 많은 경우 도면의 지역적인 특성을 사용하므로 기존에 활용되어온 도면 데이터셋의 분석은 필수적이다. 이번 절에서는 기존 도면 해석 분야에서 대표적으로 활용되던 ‘CVC 도면 데이터셋’과 ‘Rakuten 도면 데이터셋’에 대해서 다룬다.

#### 1.1 CVC 도면 데이터셋

CVC 도면 데이터셋<sup>③</sup>은 많은 연구자들이 오랜 기간 사용해온 데이터셋으로, 총 120여개의 고해상도의 건축 도면으로 구성되어 있다. Macé *et al.* (2010)의 연구에서 MAISON 건축사무소<sup>④</sup>의 도면을 수집하고, 벽체와 방의 라벨링을 수행하여 데이터셋을 처음 공개하였고, 이후에 de las Heras *et al.* (2014)의 연구에서 포맷의 다양성을 늘리기 위한 목적으로 도면을 추가로 수집하여 지금의 CVC 도면 데이터셋을 구축하였다. CVC 도면 데이터셋은 벽체의 표현 방식에 따라 4가지 유형으로 분류되어 있고, 그 유형과 분포는 그림 2-1과 같다: [BlackSet: 90개, TexturedSet: 10개, TexturedSet2: 18개, ParallelSet: 4개]. BlackSet은 벽체를 두꺼운 선으로 표현한 것이 특징으로, 해상도는 3,508 x 2,480 픽셀로 고정되어 있으며, 이는 MAISON 건축 사무소의 도면으로 CVC 데이터셋의 전신이다. TexturedSet은 벽체를 빗금 무늬로 표현하고, BlackSet에 비해 실내에 포함하는 정보가 자세하고, 표기법도 상이하다. 해상도는 1,000 x 2,000 픽셀 범위로 분포하고 있다. TexturedSet2는 TexturedSet과 같은 형태로 벽체를 표현하지만, 대상 건물의 규모 및 실내의 표기법 및 정보의 수준에 차이가 있고, 해상도는 7,383 x 5,671 픽셀 범위이다.

<sup>③</sup> <http://dag.cvc.uab.es/resources/floorplans/>

<sup>④</sup> <https://www.plans-maisons.com/>





CVC 도면 데이터셋에는 몇 가지 한계점이 있다. 먼저, 도면들을 4개의 포맷으로 분류하였으나 그 개수의 분포가 불균등하여 실질적으로 다양한 포맷의 데이터셋으로 활용되기 어렵고, 연구들의 검증의 비교 대상도 BlackSet에 국한되고 있다. 두 번째로, 도면의 수량이 학습 기반의 프로세스로 접근하기에 많이 부족하다. 이를 보완하기 위해, de las Heras *et al.* (2014)는 테스트 도면과 학습 도면들을 번갈아 가며 사용하는 방법(Leave-One-Out)을 사용했으나 이는 매 검증 시마다 다시 학습을 수행하는 임시방편의 수단이다. 또한, 100개 정도의 수량은 최근 도면 해석에 도입되고 있는 딥러닝을 적용하기에 턱 없이 부족하다. 이러한 점들로 인해 CVC 도면 데이터셋을 대상으로 수행한 도면 해석 프로세스들은 범용성 측면에서 한계가 있다.

## 1.2 Rakuten 도면 데이터셋

Rakuten 도면 데이터셋<sup>⑤</sup>은 Rakuten이라는 일본의 회사 부속 연구 기관에서 배포하는 500개의 부동산 도면 데이터셋이다. 해당 데이터셋은 LIFULL<sup>⑥</sup>의 부동산 회사에서 제공받은 도면들을 대상으로 벽체 라벨링이 수행되어 있는 형태이다. Dodge *et al.* (2017)은 데이터셋을 활용하여 딥러닝 기반으로 도면에서 벽체를 추출하는 연구를 수행했으며, Liu *et al.* (2017)은 Rakuten 데이터셋을 확장하여 벽체와 출입구의 모서리 및 방과 가구의 종류에 대한 추가적인 라벨링을 수행한 뒤 연구를 진행하였고 최종적으로 벽체와 출입구 및 가구의 정보를 재현하여 3D 실내공간을 구축했다.

Rakuten 도면 데이터셋은 CVC 데이터셋에 비해 상당히 많은 500개의 도면을 보유하고 있다. 색을 활용하여 정보를 표현하고, 해상도는 100 ~ 800 픽셀 수준으로 건축 도면에 비해 상대적으로 낮은 특징이 있다. 또한 부동산 도면의 특징상 실제 평면도나 시공 도면에 비해 실내 구조를 간략화하여 표현하며 벽체를 뚜렷하게 표시되어 있다. Rakuten 도면 데이터셋의 장점은 충분한 수의 데이터로 인해 딥러닝의 활용이 가능하다는 것이다. 하지만, 도면의 간단한 구조와 명확한 기준의 벽체는 딥러닝 적용의 필요성을 약화시키는 요소가 된다.

---

<sup>⑤</sup> [https://rit.rakuten.co.jp/data\\_release/](https://rit.rakuten.co.jp/data_release/)

<sup>⑥</sup> <https://lifull.com/>

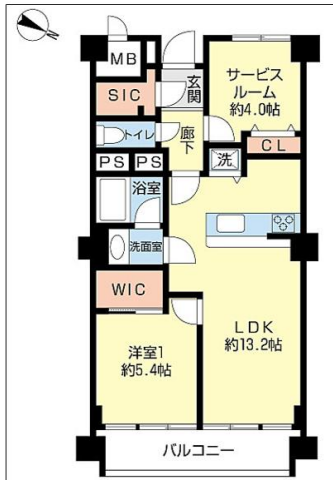


그림 2-2. Rakuten 도면 데이터셋 예시

## 제 2 절 도면 해석의 연구 동향

도면 해석이란 래스터 도면에서 목표로 하는 객체(문, 창문, 벽체 등)들을 추출한 뒤 위상 및 의미론적 정보를 파악하여 벡터 형태의 실내 구조를 구현하는 것을 의미한다. 도면 해석은 순차적인 프로세스 기반의 기술의 적용에 관한 분야로, 연구자들은 도면 해석 프로세스의 전체 틀 내에서 각자의 목적에 맞춰 알고리즘을 부분적으로 개선하는 방향으로 발전되어왔다. 예를 들면, Dosch and Masini (1999)는 그래프 이론을 기반으로 건물의 여러 층의 위상 관계에 집중하여 다층의 실내공간을 구현했고, Tombre *et al.* (2002)은 도면의 문자와 그래픽 정보의 분리에 집중하여 도면 해석 프로세스의 전처리 단계에서 큰 성과를 이뤘다. Macé *et al.* (2010)와 Ahmed *et al.* (2011)은 방의 구조 탐지 및 분할을 목적으로 연구를 진행했고, Gimenez *et al.* (2016)은 도면의 3D 구현을 목적으로 외벽, 내벽, 외곽 출구 등의 객체들의 세부적인 구분에 집중하여 연구를 수행했다. 이처럼 도면 해석 분야는 다양한 목적에 맞춰 프로세스를 개선시켜왔지만, 활용되는 알고리즘의 관점에서 규칙 기반의 프로세스에서 학습 기반의 프로세스로 전환되는 경향을 보이고 있다. 본 논문은 학습 알고리즘을 통한 도면 해석 기술의 범용성 확장에 집중하여 이러한 변화 흐름의 관점에서 기존 연구들을 분석한다.

### 2.1 규칙 기반의 프로세스

Macé *et al.* (2010)은 CVC 도면 데이터셋의 BlackSet 도면들을 대상으로 연구를 진행했고 도면 해석 프로세스는 다음과 같다. 도면 이미지를 선의 굵기를 기반으로 문자 요소와 그래픽 요소로 분리한 뒤, 허프 변환과 이미지 벡터화를 결합하여 도면에 존재하는 선형의 객체들을 탐지하였다. 포맷의 특징을 활용하여 벽체를 이미지 등고선(image contour) 기반으로 추출했고, 곡선을 탐지하여 출입문을 탐지했다. 추출한 벽체와 문을 기반으로 그래프의 볼록성(convexity) 활용하여 공간을 회귀적으로 분할하여 방의 구분을 수행했고 최종적으로는 방에 대한 탐색률과 인지 정확도에 대한 평가를 진행했다. BlackSet 포맷의 벽체 표현이 검은색으로 채워진 형태이고, 문자와 가구 등의 다른 객체들과 명확한 차이가 나는 특성을 활용한 기법으로 다른 포맷의 도면에는 적용할 수 없는 해석 프로세스이다.

Ahmed *et al.* (2011)도 BlackSet을 대상으로 연구를 수행하였다.

전처리 단계에서 Fletcher and Kasturi (1988)의 방법을 통해 도면 이미지를 텍스트와 그래픽 요소로 분리한 후, 그래픽 요소를 다시 두꺼운, 중간, 얇은 선으로 분류한다. 두꺼운 선과 중간 선을 벽체로 추출하고, 얇은 선을 대상으로 SURF(Bay *et al.*, 2006) 통해 출입구를 탐지한다. 포맷에 의존적인 임계값을 활용하여 추출한 벽체 사이의 틈을 메워 폐합된 방을 구성하고, 추출했던 문자 정보를 참조하여 추가적으로 방의 분할을 보완한다. 방의 탐색률 및 인지 정확도 측면에서 Macé의 연구보다 높은 성능을 보였으나, 그 과정에서 실증적인 임계값 기반으로 잘못 추출된 방과 벽체를 삭제하거나 그래픽 요소를 세 단계로 나누는 등의 BlackSet의 포맷의 표기법에 더욱 의존적인 방법들이 사용되었다.

Gimenez *et al.* (2016)은 도면을 활용한 건물의 3D 모델화를 목적으로 연구를 수행했다. 마찬가지로 CVC 도면 데이터셋의 BlackSet을 대상으로 연구를 수행하였고 기존 연구들의 최종결과물이 방의 구분이었던 것과 달리, 3D 모델의 구현을 목적으로 외벽, 내벽, 외곽 출입구, 내부 출입구 등의 요소들에 대한 분할된 추출을 수행하였다. 이때, 주목할 점은 기존의 연구들에 비해 벽체나 문의 추출에 대한 성능이 떨어졌지만, 3D 모델 구현 측면에서는 충분히 합리적인 결과임을 주장했다. Tombre *et al.* (2002)의 기법을 통해 도면의 문자 요소를 분리한 뒤, 그래픽 요소에서 허프 변환을 통해 직선과 곡선의 객체 후보군들을 추출한다. 추출한 객체 후보군을 포맷의 특징을 활용한 코드를 만들어 외벽, 내벽, 출입구 등으로 구분한다. 예를 들어, 기준치 근방의 평행한 두 선분이 벽체의 코드이고, 벽체에 붙어있는 호를 포함하는 평행한 선분들이 출입문의 코드이다. 도면 이미지의 요소들을 벡터화한 후 규칙을 통해 의미론적인 속성을 주입하는 방식으로 연구를 진행하였으나, BlackSet에 의존적인 규칙 기반의 코드를 적용한 만큼 다른 포맷의 도면에는 적용이 어렵다.

## 2.2 학습 기반의 프로세스

de las Heras *et al.* (2014)은 이전 연구들의 특정 포맷에만 적용 가능한 한계를 극복하기 위해 학습 기반의 알고리즘을 처음으로 도입하였다. Tombre의 기법을 통해 문자 요소를 분리한 그래픽 요소를 패치로 분할한 뒤 영상 기술자(Escalera *et al.*, 2009)를 추출하고, 각 벽체, 창문, 출입구, 배경에 대한 이미지 사전(Bag of Visual Word)을 구축하여 패치 수준의 세그멘테이션을 수행한다. 추출한 벽체를

정제하기 위해 세선화와 단순화를 알고리즘을 적용한 뒤, QGAR(Rendek *et al.*, 2004) 도구 기반의 벡터화를 통해 최종적인 벽체를 구현한다. 추출한 출입구에 대해서는 최단 경로를 찾는 A\* 알고리즘을 기반으로 벽체와의 위상을 고려한 출입구의 위치를 특정했다. 최종적으로 벽체의 추출 및 방의 탐색을 및 인지 정확도에 대한 평가를 수행하였으며, CVC 데이터셋의 4개의 유형의 포맷에 대해 각각 도면 해석 프로세스가 작동하는 것을 확인하였다. 학습 알고리즘의 도입을 통해 기술 적용의 범용성을 확장했으나, 학습 과정이 프로세스의 일부분에만 사용되었고 많은 전처리 및 후처리가 요구되는 불안정한 방법이다.

Dodge *et al.* (2017)는 Rakuten 도면 데이터셋을 대상으로 딥러닝을 적용하여 도면 해석에 접근하였다. 딥러닝 기반의 세그먼테이션을 통해 벽체를 추출하고, SURF를 활용하여 출입구를 추출했다. 최종적으로 추출한 벽체 픽셀에 높이 값을 주어, 3D 구현까지 수행하였다. 완전 연결 합성곱신경망(Long *et al.*, 2015) 기반의 딥러닝 네트워크를 활용하였으며, 커널의 보폭(stride)을 낮게 지정할수록 벽체 추출 성능이 향상되는 것을 확인하였다. 최종적으로 추출한 벽체를 대상으로 픽셀 기반의 검증을 수행했다. 주목할 점은 문자를 제거하는 전처리 과정을 생략했음에도 영상 기술자 기반의 학습보다 나은 성능을 보였다는 것이다. 제안된 프로세스를 추가적으로 CVC 도면 데이터셋에 대해서도 진행했으나 학습을 위한 도면 수량의 부족으로 인해 성능이 낮게 측정되었다.

Liu *et al.* (2017) 또한 Rakuten 데이터셋을 대상으로 도면 해석 프로세스를 진행했다. Dodge가 딥러닝을 객체의 세그먼테이션에 활용했다면, Liu는 딥러닝을 객체의 벡터화를 위한 중간 산물의 추출에 활용하였다. 딥러닝을 통해 각 객체의 모서리들의 위치와 방향성, 연결성 정보를 추출한 뒤, 이를 활용해서 선형의 벡터화된 객체들을 구현하였다. 이때 수행한 라벨링의 수준은 방향 및 연결성 정보를 세분화한 21개의 클래스의 모서리와 10개의 클래스의 방의 정보(부엌, 거실, 화장실 등), 8개의 클래스의 가구(발코니, 옷장, 싱크대 등)의 정보이다. 이러한 정답 데이터를 기반으로 He *et al.* (2016)의 딥러닝 네트워크를 활용하여 각 객체의 모서리 정보를 위한 히트맵의 추출과 방과 가구 정보에 대한 세그먼테이션을 수행한다. 딥러닝의 출력물을 활용하여 연결성, 상호 겹침, 폐합, 출입구 제약 조건 기반의 정수계획법을 수행하여 최종적인 벡터화된 건물 모델을 구현한다.

주목할 점은 이전 연구와는 달리 딥러닝을 높은 수준의 정보의 추출에 활용한 적용한 점이다. 세그먼테이션을 통한 객체 추출의 출력물은 여전히 래스터 형태이고, 최종적인 벡터 포맷을 구현하기 위해서는 추가적인 세션화, 벡터화 등의 작업이 필요하다. 제안된 논문에서는 실내 구현의 측면에서 더 높은 수준의 정보인 모서리를 출력하여 객체를 벡터 형태로 구현하였고, 이는 딥러닝 적용 측면에서 한단계 더 나아간 것이다. 하지만 학습데이터의 구축을 위해 요구되는 라벨링의 정보 구축 수준이 높고, 그 수량도 770장으로 기술의 활용을 어렵게 한다. 또한, 모서리 추출을 통한 벡터화의 접근 방법은 실내가 단순한 구조로 표현되는 Rakuten의 도면에서만 활용 가능하다. 구체적인 예시로 벽체의 기하학적인 형태가 세세하게 표현되어 있는 포맷의 경우에는 2개의 모서리로 벽체를 표현하기 어렵고, 벽체의 굽기가 일정하지 않을 경우에는 모서리의 위치가 모호한 문제가 발생한다. 이를 해결하기 위해서는 실내의 구조를 단순한 사각형의 조합으로 변환하는 전처리가 요구된다. 이러한 이유로 인해 제안된 방법을 복잡한 도면에 적용하기 위해서는 여러 한계가 존재한다.

### 제 3 절 학습 알고리즘의 확장

도면 해석의 분야는 범용성 확장을 목적으로 규칙 기반에서 학습 기반의 프로세스의 추세로 발전하고 있고, 이 과정에서 학습 알고리즘의 적용을 확장시켜 나가고 있다(그림 2-3). 학습 기반의 도면 해석 프로세스는 일반적으로 전처리, 패턴 인식, 구조 인식의 단계로 진행되고, 이때 패턴 인식 단계는 세그먼테이션과 벡터화로 세분화될 수 있다. 이러한 단계를 기준으로 도면 해석 프로세스들에서 활용된 학습 알고리즘의 역할을 비교하여 기존 연구들 및 본 논문의 차이점을 확인한다. 구체적인 방법은 적용한 학습 알고리즘들의 입력물과 출력물을 비교하는 것이다.

학습 기반의 알고리즘을 처음 적용한 de las Heras *et al.* (2014)는 영상 기술자를 활용했다. 학습 알고리즘의 입력물은 문자 요소를 분리하는 전처리를 수행한 도면을 그리드 분할한 형태이고, 출력물은 각 그리드 패치를 벽체와 창문, 출입구, 배경으로 분류한 것이다. 이는 거친 수준의 세그먼테이션을 수행한 것으로 래스터 형태의 객체 추출물이다. Dodge *et al.* (2017)는 딥러닝을 활용하여 벽체의 세그먼테이션을 수행했다. 입력물은 도면의 원본을 그대로 활용되고, 출력물은 픽셀 수준으로 벽체를 추출한 래스터 이미지이다. 이는 문자 요소의 제거 등의 전처리를 딥러닝 네트워크의 내부에서 자체적으로 수행한 것으로, 학습 알고리즘을 전처리 단계까지 확장하여 활용한 의의가 있다.

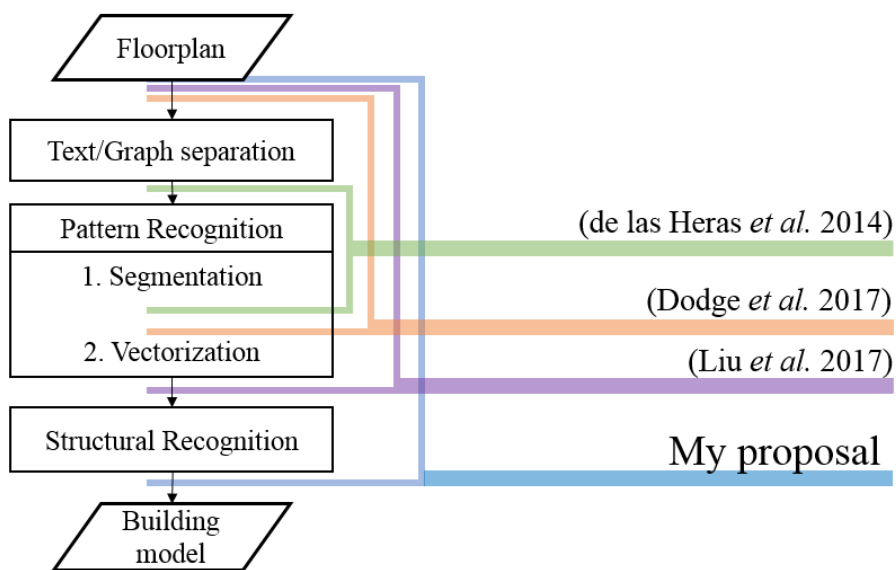


그림 2-3. 도면 해석 프로세스에서 학습 알고리즘의 확장



세부적으로는 딥러닝의 적용을 위해 도면의 크기 변환 등의 이미지 처리를 활용했지만 도면 해석 측면의 전처리는 수행되지 않았다. Liu *et al.* (2017)은 객체들의 벡터화를 위한 중간 산물을 출력하기 위해 딥러닝을 사용했다. 입력물은 마찬가지로 원본 도면이 사용되며, 출력물은 도면에 대한 세그멘테이션의 결과물과 히트맵으로 표현되는 모서리 정보들이다. 추출한 모서리 정보를 활용하여 벽체와 문의 벡터화를 진행하였다. 이는 딥러닝이 전처리를 넘어 패턴 인식의 세부 단계인 벡터화까지 확장되어 활용되는 형태이다. 본 논문의 도면 해석 프로세스는 학습 알고리즘의 적용을 한 단계 더 확장한다. 원본 도면의 입력에 대해서 기하학적 벽체가 아닌 방의 경계를 출력한다. 이는 학습 알고리즘이 내부적으로 벽체의 세션화 및 단순화를 추가적으로 수행한 것으로, 학습 알고리즘을 구조 인식 단계까지 확장하여 활용한 것이다. Liu *et al.* (2017)이 간단한 포맷을 대상으로 벽체를 구현했다면, 본 연구는 복잡한 포맷에 대한 방의 경계, 즉 실내의 구조를 구현한다. 이처럼 구현하고자 하는 목표를 벽체에서 실내의 구조로 변경하여, 간단한 포맷에만 적용 가능했던 기존의 도면 해석 프로세스들을 복잡한 도면에까지 확장하여 활용할 수 있게 하는 의의가 있다.

## 제 3 장 도면 데이터셋 구축

학습 기반의 도면 해석 프로세스를 활용하기 위해서는 학습이 선행되어야 하고, 이는 학습에 사용하기 위한 데이터셋을 구축해야 함을 의미한다. 학습 데이터셋의 구축은 목표로 하는 포맷으로 구성된 수백장의 도면을 수집하고 이에 대한 라벨링을 수행하여 이루어진다. 하지만 유사한 포맷의 도면들을 수집하고 일관된 기준으로 라벨링을 하는 작업, 즉 양질의 학습 데이터를 구축하기 위한 작업은 현실적으로 어렵기 때문에 도면 해석 기술을 활용하는데 장벽으로 작용한다.

본 연구는 다양하고 복잡한 포맷의 도면들로 구성된 데이터셋을 대상으로 도면 해석 프로세스를 진행했다. 양질의 학습 데이터셋을 구축하기 어려운 상황에 대한 학습 프로세스를 진행함으로써, 현실의 다양한 수준으로 수집되는 도면들에 대해서도 도면 해석 프로세스를 활용할 수 있도록 하였다. 이를 통해 도면 해석 기술의 범용성을 강화하였다. 이 장에서는 학습에 활용하기 위한 도면 데이터셋을 구축하는 과정을 다룬다. 먼저, 다양하고 복잡한 건축 도면들로 데이터셋을 구성했다. 그 뒤, 새로운 라벨링 방법을 제안하여 복잡하고 다양한 도면들에 대해서도 일관된 기준으로 라벨링을 수행하였다.

### 제 1 절 세움터 도면

복잡하고 다양한 도면의 수집은 국가의 건축행정 업무 전반을 관리하는 전산 시스템인 세움터(E-AIS)<sup>⑦</sup>를 활용했다. 세움터를 통해 스캔된 형태로 보관되는 400개의 도면 이미지를 획득하여 세움터 도면 데이터셋을 구성했다(그림 3-1). 세움터 도면 데이터셋은 건축 도면으로 분류되지만, 이를 구성하고 있는 개개별 도면의 포맷의 형태는 각양각색이다. 1) 대상 건물의 규모가 주택, 연립주택, 아파트 등으로 다양하고, 2) 벽체에 대한 표기 방법이 단순한 검은 색과 흰색의 평행선을 넘어 다양하고, 단일 도면 내에서도 다양한 형태로 표현하기도 한다. 또한, 벽체의 표현 범위도 내력 벽이나 기둥 및 단열재 등으로 다양하다. 3) 담고 있는 정보의 수준이 다양하여 단순한 가구 정보부터

<sup>⑦</sup> <https://www.eais.go.kr>

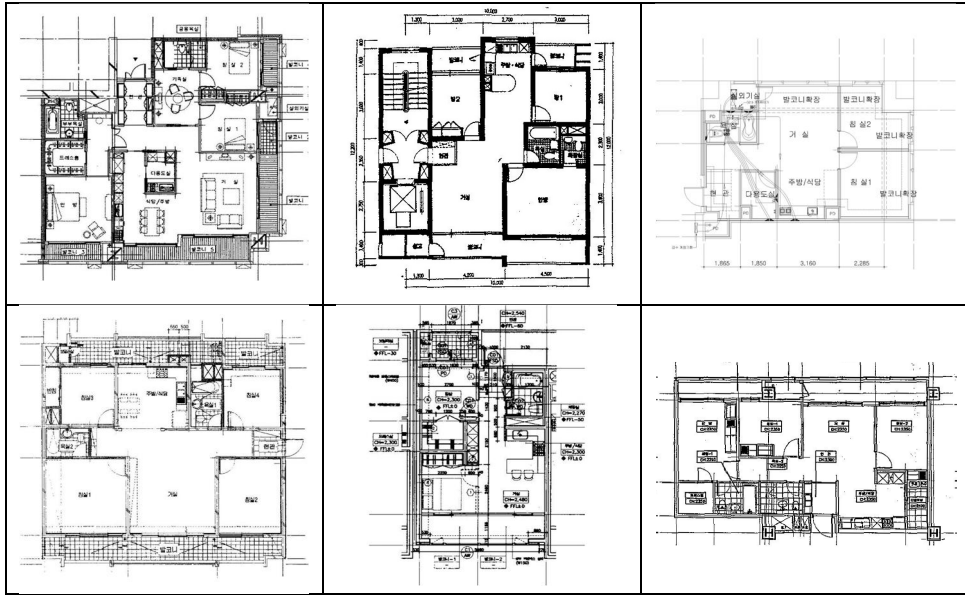


그림 3-1. 세움터 도면 데이터셋 예시

설비 및 시공 관련 정보까지 다양한 스펙트럼을 보이고 있고 이는 포맷의 다양성으로 나타난다.

세움터 도면 데이터셋은 기존의 공용 도면 데이터셋과 비교하여 포함하고 있는 도면의 포맷 종류가 다양하고, 복잡한 도면으로 구성되어 있는 특징이 있다. 세움터 도면은 스캔된 형태의 건축 도면으로 평면도, 시공 도면, 설비 도면 등의 원형의 포맷인 반면, CVC 도면은 이러한 원형의 포맷에서 내력벽, 기둥 및 설비 공간을 삭제하거나, 벽체를 일관된 형태로 바꾸는 등의 단순화 및 세션화 과정이 수행된 형태로 세움터 도면보다 간단한 포맷이다. 이러한 포맷의 복잡함으로 인해, 기존의 라벨링 방법을 그대로 활용할 경우 도면 해석 프로세스의 성능이 크게 감소한다. 상대적으로 단순한 형태의 CVC 도면에 대한 라벨링에 대해서도 연구자들은 많은 어려움을 토로해왔고, 실제로 세움터 도면의 경우에는 포맷이 복잡하여 기본적인 벽체의 기준도 모호한 도면들이 다수 존재한다.

## 제 2 절 실내 구조의 라벨링

일반적인 라벨링 작업은 대상 객체에 대한 판단 코드를 통해 수행된다. 예를 들면, 벽체를 표현하는 구조벽, 내력벽, 기둥 등의 다양한 요소들 중 어디까지를 벽체로 판단할 것인지를 기준 역할을 하는 코드를 결정하여, 벽체 대상의 일관된 형태의 라벨링을 수행한다. 하지만 세움터 데이터셋의 도면들은 포맷이 일관되지 않고 표현하는 정보의 양이 다양하다는 특징이 있어, 코드 기반의 라벨링을 위해서는 수 많은 사례들에 대한 코드의 결정이 필요하다. 이러한 접근은 효율적이지 못할뿐더러, 학습 데이터의 구축의 난이도를 높여 기술 적용의 범용성을 축소시킨다. 이를 극복하기 위해, 본 연구에서는 다양하고 복잡한 포맷들에 일관되게 적용할 수 있는 라벨링 방법을 제안한다. 이는 딥러닝 활용을 위해 적합한 형태와 다양한 포맷에서 일관되게 적용할 수 있는 방법의 두 가지 관점에서 수행된다. 기본적으로 본 연구는 딥러닝 기반의 스타일 트랜스퍼의 적용을 통해, 도면의 포맷과 실내의 구조를 스타일과 콘텐츠에 대응시킨 뒤 콘텐츠에 대한 정보를 출력한다. 이러한 접근에 맞춰 콘텐츠에 해당하는 실내 구조의 추출을 목적으로 도면의 라벨링을 진행한다.

즉, 제안하는 라벨링은 다양한 포맷에서 일관된 형태로 실내 구조를 표현할 수 있는 방법이다(그림 3-2). 제안의 핵심은 도면의 기하학적인 형상보다 실내 구조의 표현, 즉 방의 구분을 우선시하여 라벨링을 진행하는 것이다. 객체의 개별 사례를 다루는 판단 코드보다 상위의 개념을 제시하여 다양한 포맷의 도면들에서 비교적 일관된 형태로 라벨링을 수행할 수 있도록 한다. 구체적으로 벽체의 라벨링은

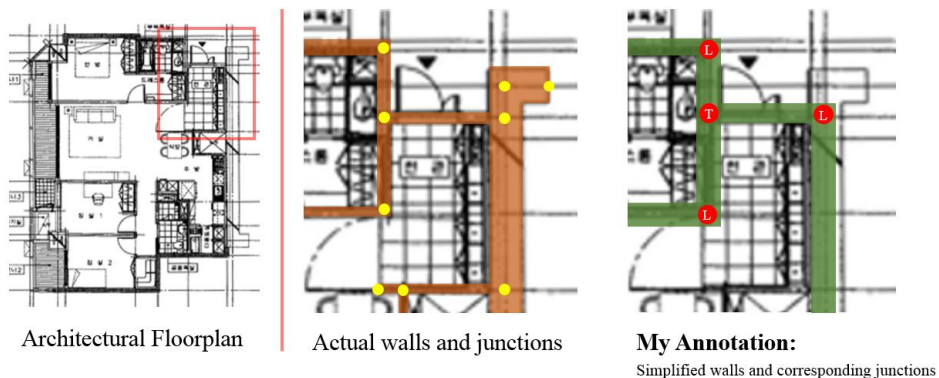


그림 3-2. 실내 구조 표현을 위한 라벨링

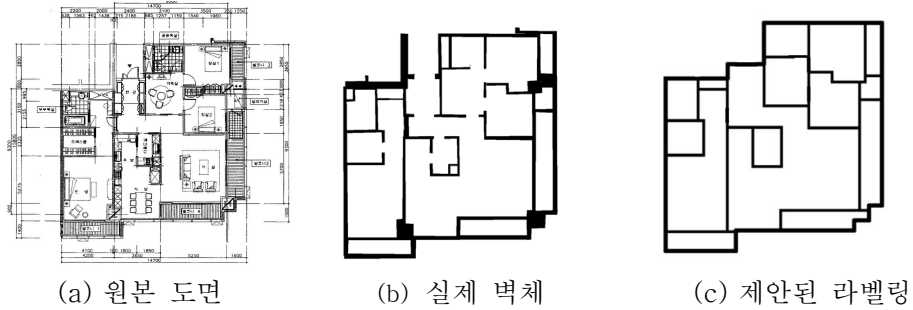


그림 3-3. 세움터 도면 대상의 벽체 및 제안한 라벨링 예시

방의 외곽 벽체를 대상으로 진행하며 형태는 다음과 같다(그림 3-3). 이때, 실제 벽체의 기하학적인 형상보다 방의 구조를 우선 기준으로 하여 방의 구분에 사용되지 않는 벽체를 생략하고, 방을 구분하지만 벽체로 표기되지 않는 부분을 표기하고, 벽체의 실제 두께보다 일관된 스타일의 형태로 표현하여 라벨링 작업을 수행한다.

실내 구조의 표현이라는 추상적인 개념이 개별 사례에 대한 판단을 직접 결정하는 코드보다 우선시되는 이유는 다음과 같다. 실제 도면의 라벨링 작업에서 기준이 모호한 문제가 발생하는 경우는 실내의 구조와 같은 도면의 핵심 정보를 표현하는 부분이 아닌, 외벽의 돌출된 부분을 벽체로 판단할 것인가 등의 지엽적인 부분들이다. 이러한 모호성은 정답 데이터의 일관성을 약하게 만들고 학습 알고리즘의 성능을 감소시키는 요인이 된다. 제안하는 라벨링 방법은 핵심 정보에만 집중하여 모호한 기준이 발생하는 상황을 회피하는 것을 목적으로 한다. 제안하는 방법의 큰 장점은 정답 데이터의 구축 작업을 용이하게 해주고, 복잡한 도면에 대해서도 비교적 일관된 형태의 라벨링이 가능하게 하는 것이다. 엄밀한 기준에서 판단 코드를 활용한 라벨링보다 정답 데이터의 일관성은 떨어지지만, 이는 딥러닝의 활용에 충분한 수준이고 특히 본 논문에서는 스타일 트랜스퍼를 적용하여 일관성이 떨어지는 부분을 보완하고 있다.

라벨링의 대상은 세움터의 다양한 포맷에 공통적으로 존재하는 객체인 벽체와 출입구로 한다. 각 객체는 선형의 구조들의 조합으로 표현할 수 있으므로, 이를 벡터 포맷으로 구현하기 위해서 모서리 정보를 추출하는 방식으로 접근한다. 모서리 정보의 라벨링은 위치와 클래스를 대상으로 한다(그림 3-4). 모서리가 위치는 픽셀의 좌표이고 클래스는 5개로 나뉘는 되는 모서리의 연결성 정보이다. 벽체의

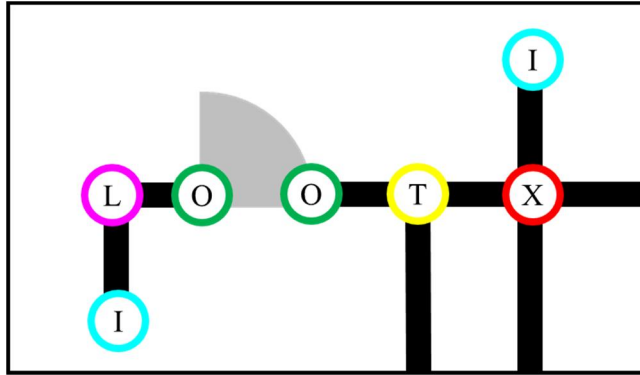


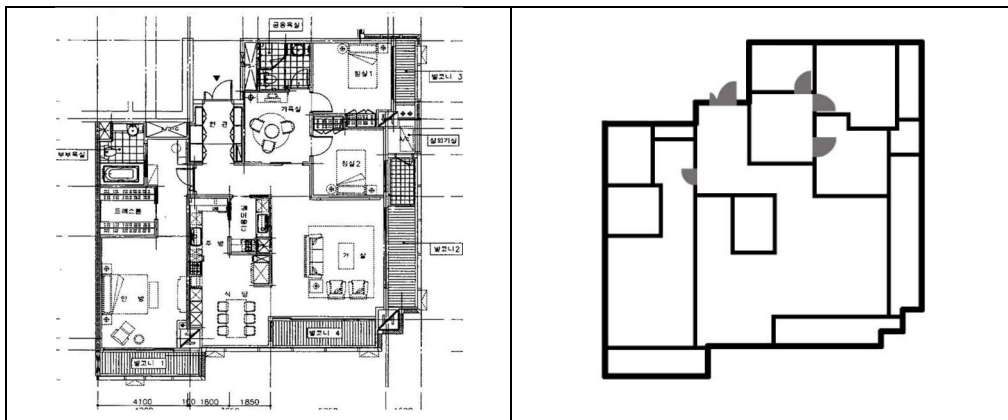
그림 3-4. 모서리 클래스의 표현

(I: 청록색, L: 자주색, T: 노란색, X: 빨간색, O: 초록색)

모서리는 I, L, T, X 4가지 클래스를 가지며, 이는 각 1, 2, 3, 4개의 벽체가 만나는 지점을 의미한다. 출입구의 모서리는 O(opening) 클래스를 가지며 벽체에 연결되는 출입문의 양 끝점을 통해 표현한다. 이와 같은 5개의 클래스로 표현되는 모서리 정보를 활용하여 실내 공간의 벡터화를 진행한다.

세움터 도면들을 대상으로 구축한 정답 데이터셋은 표 3-1에서 확인할 수 있다. 왼쪽의 원본 도면에 대해서 우측의 형태로 벽체와 출입구를 표현하였고, 대응되는 모서리 정보까지의 라벨링을 수행하였다.

표 3-1. 세움터 도면 데이터셋 라벨링 예시





## 제 4 장 도면의 벡터화 프로세스

도면의 실내 구조를 구현하기 위한 벡터화 프로세스의 과정은 다음과 같다(그림 4-1). 먼저, 딥러닝 네트워크를 활용하여 도면의 스타일 변환과 모서리 정보의 출력을 수행한다. 두 개의 네트워크는 학습되는 파라미터들을 공유하여, 마치 도면에서 동일한 피쳐 맵(feature map)을 추출하고 이를 기반으로 각 결과를 출력하는 형태로 작동한다. 딥러닝의 출력물인 스타일 변환된 도면과 모서리 히트맵(heat map)은 그 개별로는 벡터 포맷의 실내 구조를 표현하기에 불충분한 정보이므로 정수 계획법을 통해 상호 보완하는 과정을 거친다. 정수 계획법의 최종 출력물은 모서리의 연결성 정보를 표현하는 인접 행렬로, 이때 여러 제약 조건 및 목적 함수를 통해 도면의 특징을 반영하도록 한다. 출력된 인접 행렬을 통해 벡터 포맷의 실내 구조를 구현하여 도면 해석 프로세스를 완성한다.

딥러닝은 Python 기반의 Tensorflow<sup>⑧</sup>를 사용했고 정수계획법의 최적화는 Gurobi<sup>⑨</sup>를 활용하였다. 학습을 위해서는 GTX 1080 Ti를 통해 8시간이 걸렸고, 최종 결과의 출력에 있어서는 도면 하나당 2초 이내의 시간이 걸렸다.

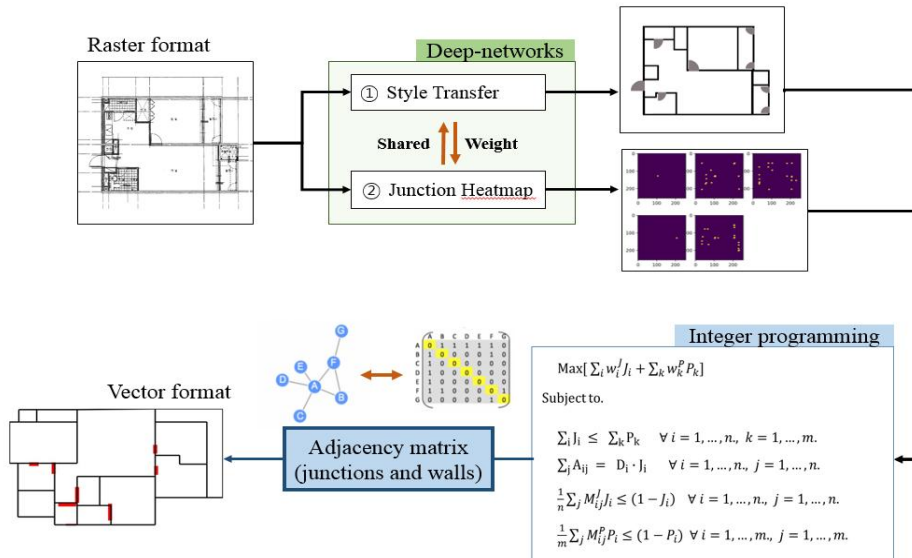


그림 4-1. 도면의 벡터화 프로세스

⑧ <https://www.tensorflow.org/>

⑨ <http://www.gurobi.com/>



## 제 1 절 스타일 트랜스퍼의 적용

딥러닝은 스타일 트랜스퍼와 모서리 정보의 추출의 두 가지 작업을 위해 활용되었다. 딥러닝 기반의 이미지의 처리는 다음의 관점으로 해석할 수 있다. 딥러닝 네트워크의 컨볼루션 레이어(convolution layer) 부분은 이미지의 피쳐 맵을 추출하는 역할로, 완전 연결 레이어(fully connected layer) 부분은 점수를 매겨 결과를 출력하는 결정 레이어의 역할로 대응시킬 수 있다. 연구에 사용하는 두 네트워크의 컨볼루션 레이어에 대응되는 부분의 파라미터를 공유하였고, 동일한 피쳐 맵을 기반으로 스타일 변환과 모서리 정보의 출력이 수행되도록 하였다.

총 400장의 세움터 도면 데이터셋 중에 350장을 랜덤 추출하여 학습에 사용했고 남은 50장은 검증에 사용하였다. 딥러닝을 위한 입력 이미지의 크기는 256x256 픽셀으로 고정했고, 스타일 변환과 모서리 히트맵의 출력물도 동일한 크기로 설정하였다. 도면은 가로와 세로의 비율이 중요한 정보를 포함하기 때문에 이 비율을 유지하며 이미지 크기를 변환시켜야 된다. 따라서 이미지를 잘라서 크기를 맞추는 크롭핑(cropping)이 아닌, 빈 공간을 덧대서 크기를 조절하는 패딩(padding)을 활용하였다. 또한, 데이터 증강을 위해 90도 단위의 회전을 및 이미지의 잘린 부분을 입력하는 랜덤 크롭핑, 색 정보를 변형하는 지터링(jittering)의 작업을 수행하였다.

### 1.1 네트워크의 구조

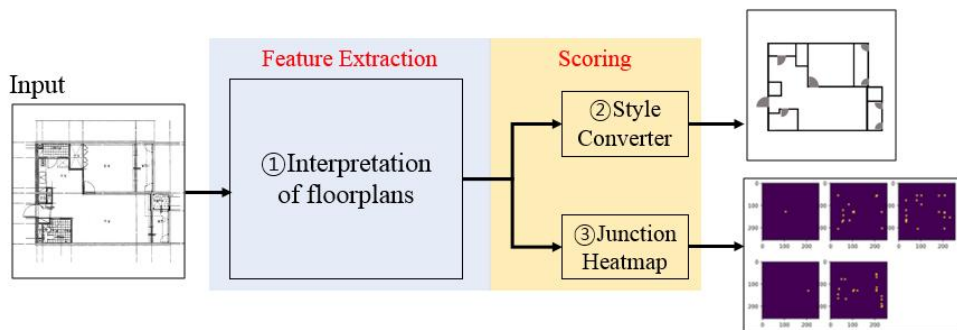


그림 4-2. 딥러닝 네트워크의 구조

네트워크의 구조(architecture of networks)는 다음과 같다(그림 4-2). ①의 부분은 도면의 피쳐 맵을 추출하는 역할, ②와 ③은 각각 추출된 피쳐 맵을 통해 점수를 매겨 정답을 출력하는 역할을 수행하였다.

②의 출력물은 변환된 스타일로 생성된 도면 이미지이고, ③의 출력물은 각 모서리 클래스에 대응되는 히트맵이다. 즉, [①, ②]가 스타일 트랜스퍼의 역할을 수행했고, [①, ③]은 히트맵을 통해 모서리 정보를 출력하였다.

네트워크는 다양한 연구를 참고하여, He *et al.* (2016)의 ResNet152를 수정하여 사용하였다. 먼저, ResNet152의 bottleneck 모듈의 보폭을 2픽셀에서 1픽셀로 줄였다. 이는 딥러닝의 출력물들이 1차원의 벡터가 아닌, 입력 이미지와 같은 크기의 2차원 행렬로 표현되기에 해상도를 높이기 위한 목적으로 수행됐다. 다음으로 후미의 완전 연결 신경망 부분을 제거하고, 스타일 변환과 모서리 히트맵을 위한 두개의 네트워크를 평행하게 연결하였다. 연결되는 네트워크는 디컨볼루션 레이어 (deconvolution layer)를 기반으로 줄어든 크기를 복구시키는 역할을 수행하고, 각 정답 출력물의 형태에 적합한 구조를 선택했다. 스타일 변환을 위한 부분은 (Zhu *et al.*, 2017)의 연구를 기반으로 적용했고, 모서리 히트맵을 위한 부분은 (Bulat and Tzimiropoulos, 2016)의 연구를 참고하였다. 히트맵 부분을 살펴보면, 64x64x2048의 피쳐 맵을 한 번에 5개의 클래스에 대응하는 64x64x5의 형태로 차원을 감소시킨 뒤, 다음 레이어에서 이미지의 크기를 디컨볼루션 레이어를 통해 4배 증가시켰다. 사용된 네트워크의 자세한 사항은 표4-1에서 확인 할 수 있다. 각 ①, ②, ③ 네트워크에 대해, (채널 수, 커널 크기, 보폭) / (커널 크기, 보폭)의 기호로 구조를 표기하였다.

표 4-1. 딥러닝 네트워크의 세부 구조

① Interpretation of floorplans				
1	2	3	4	5
1x conv layer (64, 7x7, 2x2) 1x pooling (3x3, 2x2)	3x bottleneck modules [(64, 1x1), (64, 3x3), (256, 1x1)]	8x bottleneck modules [(128, 1x1), (128, 3x3), (512, 1x1)]	38x bottleneck modules [(256, 1x1), (256, 3x3), (1024, 1x1)]	3x bottleneck modules [(512, 1x1), (512, 3x3), (2048, 1x1)]

② Style converter			③ Junction heatmap	
1	2	3	1	2
1x conv layer (64, 1x1, 1x1)	1x deconv layer (64, 4x4, 4x4)	1x conv layer (3, 7x7, 1x1)	1x conv layer (5, 1x1, 1x1)	1x deconv layer (5, 4x4, 4x4)

## 1.2 손실 함수

### 1.2.1 스타일 트랜스퍼

본 연구는 딥러닝 기반의 스타일 트랜스퍼를 위해 Conditional GAN (Isola et al, 2017)을 활용하였다. Conditional GAN의 전신이 되는 GAN(Goodfellow et al, 2014)의 기본 구조는 생성 네트워크와 식별 네트워크를 통해 이미지의 생성을 최대최소문제로 접근하는 것이다. 두 개의 딥러닝 네트워크가 '진품과 구분 불가능한 모조품을 생성하는 것'과 '진품과 모조품을 구분하는 것'의 역할을 수행하며 서로 경쟁적으로 성능을 향상 시키는 것이 목적이다. 이상적으로 학습이 진행된 경우 중국엔 진품과 구분이 안되는 모조품을 생성하게 된다. Conditional GAN은 진품과 모조품의 관계를 지정한 이미지 쌍 간의 관계로 대체시킨다. 다른 스타일의 이미지 쌍들을 활용하여 학습을 진행하고 최종적으로 딥러닝 네트워크는 입력한 이미지에 대한 스타일 변환을 수행한다.

입력 이미지를  $x$ , 그에 대응되는 다른 스타일의 이미지를  $y$ , 랜덤 노이즈 벡터를  $z$ 로 표기한다. 생성 네트워크는  $G$ , 식별 네트워크는  $D$ 이다. Conditional GAN의 목적은 생성 네트워크가  $x$ 와  $z$ 를  $y$ 에 매핑하는 것이다:  $G: \{x, z\} \rightarrow y$ . 이 때, 이상적인 식별 네트워크는 입력 받은 두 인수가 진짜 쌍의 관계이면 1, 생성된 가짜 쌍이면 0을 출력한다. 각 스타일 변환과  $L_1$  에 대한 손실 함수는 아래와 같다.

$$\mathcal{L}_{CGAN}(G, D) = E_{x, y \sim p_{date}(x, y)}[\log D(x, y)] \\ + E_{x \sim p_{date}(x), z \sim p_z(z)}[\log(1 - D(x, G(x, z)))]$$

$$\mathcal{L}_{L_1}(G) = E_{x, y \sim p_{date}(x, y), z \sim p_z(z)}[\|y - G(x, z)\|_1]$$

Conditional GAN의 최종적인 목적은 아래와 같이, 결합된 두 손실 함수를 통해 최대최소문제를 푸는 것이다. 학습된 생성 네트워크는 입력 이미지에 대응하여 스타일 변환된 이미지를 생성한다.

$$G^* = \operatorname{argmin}_G \max_D [\mathcal{L}_{CGAN}(G, D) + \lambda \cdot \mathcal{L}_{L_1}(G)]$$

스타일 트랜스퍼는 생성 네트워크와 식별 네트워크에 대한 최대최소문제를 풀기 위해서 순차적인 학습의 방식으로 접근한다.

구체적으로 식별 네트워크에 대한 최대 문제와 생성 네트워크에 대한 최소 문제의 반복을 통해 네트워크를 학습시킨다. 즉, 생성 네트워크의 학습은 스타일 손실 함수,  $\mathcal{L}_{Style}(= \mathcal{L}_{CGAN} + \lambda \cdot \mathcal{L}_{L_1})$ 를 최소화하는 문제로 접근할 수 있으므로, 본 논문에서는 최소문제라는 공통점을 활용하여 스타일 변환의 손실과 모서리 히트맵의 손실을 결합해 학습을 진행하였다.

식별 네트워크는 Conditional GAN과 마찬가지로 PatchGAN을 활용한다. PatchGAN은 식별 네트워크가 스타일 변화에 대한 참, 거짓으로 단순히 0, 1을 출력하는 것이 아닌, 이미지를 그리드 패치로 나누어 각 패치에 대해 판단을 수행해 2차원 행렬 형태로 결과를 출력하는 것이다. 이미지를 패치 단위로 해석해 지역적인 특성을 반영하며 스타일 변화를 수행하는 것이 목적으로, 이는 일부분을 통해서도 해석이 가능한 도면의 특성과도 잘 어울린다. 본 연구에서는 식별네트워크의 패치 크기가 각 70과 16인 경우에 대해 실험을 수행하여 결과를 비교하였다.

### 1.2.2 모서리 히트맵

모서리의 좌표와 클래스를 동시에 추출하기 위해 히트맵의 형태로 결과를 출력하였다. 라벨링된 모서리 정보는 픽셀 좌표 값과 클래스를 가지고 있고 학습을 위해 이 정보를 히트맵의 형태로 바꾸어야 한다. 학습에 사용하는 256x256 픽셀 크기를 기준으로, 5개의 모서리 클래스를 표현하는 256x256x5 형태로 히트맵을 구성하였다. 각 클래스 별로 모서리의 픽셀 좌표로부터 L1 거리가 4픽셀 이내인 부분을 정답 모서리로 표현하였다.

$$J^* = \operatorname{argmin}_j \mathcal{L}_{Junc},$$

$$\text{where } \mathcal{L}_{Junc}(J) = E_{x, j \sim p_{date}(x, j)} [\text{sigmoid\_cross\_entropy}(J(x), j)]$$

모서리 히트맵의 네트워크(J)는 손실 함수를 최소화하는 방향으로 학습했다. 이때, 모서리의 손실 함수( $\mathcal{L}_{Junc}$ )는 픽셀 기반의 크로스 엔트로피(pixel-wise sigmoid cross entropy)를 사용하였다. 모서리 라벨링 히트맵 데이터 j에 대해 목적 함수는 위치럼 표현된다. 크로스 엔트로피 손실을 픽셀 단위로 수행한 것으로, 구체적으로  $\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^W \sum_{j=1}^H [p_{ij}^n \log p_{ij}^{\hat{n}} + (1 - p_{ij}^n) \log(1 - p_{ij}^{\hat{n}})]$  의 연산을 통해

수행된다( $N$ = 모서리 클래스의 수,  $W$ =이미지 넓이,  $H$ =이미지 높이,  $p^n$  =정답,  $p^{\hat{n}}$ =sigmoid 출력 값).

### 1.2.3 손실 함수의 결합

그림4-2에 표현된 네트워크 구조를 스타일 트랜스퍼와 모서리 히트맵에 대응시키면 다음과 같다. 스타일 트랜스퍼의 생성 네트워크( $G$ )는 [①, ②]로 표현되고, 모서리 히트맵의 네트워크( $J$ )는 [①, ③]으로 표현된다. 이는 학습을 위해 각각의 손실 함수에 연결된 파라미터들을 보여준다. 예를 들면, 스타일 트랜스퍼의 손실 함수( $\mathcal{L}_{style}$ )를 통한 학습은 ③의 파라미터는 고정된 상태에서 손실 함수를 최소화하는 방향으로 [①, ②]의 파라미터에 대한 학습이 진행된다.

$$\text{minimize}_{(G, J)} [\mathcal{L}_{style} + \lambda \cdot \mathcal{L}_{junc}]$$

네트워크의 학습은 스타일 변환과 모서리 히트맵의 손실 함수를 결합하여 진행되었다. 학습은 전체 손실 함수( $\mathcal{L}_{style} + \lambda \cdot \mathcal{L}_j$ )의 최소화와 식별 네트워크의 손실 함수( $\mathcal{L}_{CGAN}$ )의 최대화의 반복을 통해 수행된다. 이때, 손실 함수의 비율을 조정하는  $\lambda$ 는 일반적으로 학습 데이터셋의 특성에 맞추어 실증적으로 결정된다. 하지만 본 논문에서는 실험적으로 결정되는 변수를 줄이기 위해  $\mathcal{L}_{style}$ 와  $\mathcal{L}_j$ 의 반복 학습을 통해  $\lambda$ 의 역할을 없앴다. 즉, 실질적으로  $G^* = \text{argmin}_G \mathcal{L}_{style}$ ,  $J^* = \text{argmin}_J \mathcal{L}_{junc}$ ,  $D^* = \text{argmin}_D \mathcal{L}_{CGAN}$ 의 세 개의 학습을 단계별로 반복하여 진행하였다.

## 제 2 절 정수 계획법을 통한 보완

딥러닝을 통해 스타일이 변환된 도면 이미지와 모서리 정보가 포함된 히트맵이 출력된다. 출력물은 벽체와 출입문에 대한 정보를 가지고 있지만, 각각은 벡터 형태의 실내 구조로 구현하기에 불완전한 정보들이다. 스타일 변환된 도면에 대해서는 세션화 및 벡터화 등의 과정이 추가적으로 요구되고, 모서리 히트맵에 대해서는 모서리의 연결성 정보를 구현하기 5개의 클래스로는 부족하다는 한계가 있다. 또한 같은 위치에 대한 여러 클래스의 모서리가 추출되는 등의 딥러닝의 성능과 관련한 문제도 있다. 그러므로 두 출력물을 상호 보완하며 벡터 포맷의 실내 정보를 구현하기 위한 과정이 필요하고, 이를 위해 정수 계획법을 사용하였다.

정수 계획법을 통해 실내의 구조를 벡터 포맷을 구현하기 위해서 오차의 임계값을 설정해야한다. 이는 출력된 모서리들이 같은 선 위에 존재한다고 판단하기 위해서 오차를 어느 정도까지 허용할 지, 모서리가 같은 장소에서 겹친다고 판단하기 위한 기준은 어느 정도인지 등에 대한 내용을 결정하기 위해 요구된다. 또한, 모서리 히트맵에서 범위로 표현되는 모서리를 하나의 점으로 결정하는 과정에서도 오차의 임계값을

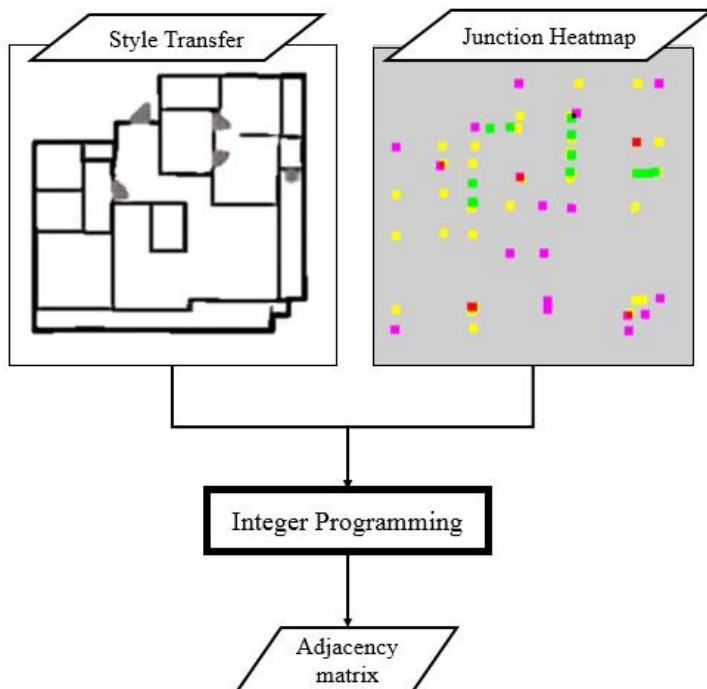


그림 4-3. 정수 계획법을 통한 딥러닝 출력물의 보완

사용한다. 본 연구에서 사용한 256x256 픽셀의 도면에 대해서 오차의 임계값은 4픽셀로 설정하였다. 이는 사람이 모서리의 위치를 라벨링하며 생기는 오차, 세움터 도면의 벽체 구조를 표현하기 위해 최소로 허용해야되는 오차 등을 고려해 실증적으로 결정되었다.

## 2.1 변수

$$J_i : i^{th} \text{ junction}$$

$$P_k : k^{th} \text{ wall primitive}$$

변수는 출력된 모서리( $J_i$ )와 벽체 후보군( $P_k$ )이고, 각각은 0과 1의 값을 통해 객체의 유무를 결정한다. 예를 들어,  $J_i$ 의 값이 1이면  $i$ 번째 모서리는 정답으로서 실내 구조 표현에 사용되는 것이고, 0이면 사용되지 않는다. 모서리와 벽체의 후보군(primitive)은 다음의 과정을 통해 추출되었다. 모서리는 모서리 히트맵에서 시그모이드의 값이 0.4 이상인 픽셀을 대상으로 오차의 임계값을 활용한 Non-Maximum Suppression을 수행하여 추출되었다. 딥러닝의 출력물에서 모서리가 존재할 확률이 40%인 곳을 대상으로 삼은 것으로 재현율을 높이기 위한 접근이다. 이 과정으로 각 클래스별 모서리들은 점으로 추출되었고, 모서리 후보군으로 활용된다. 벽체 후보군은 추출된 모서리를 통해 생성되며, 그 조건은 대각선의 벽체는 없는 것으로 가정하여 모서리 쌍이 수직이나 수평으로 마주보는 것으로 하였다.

## 2.2 목적 함수

$$\text{Max} \left[ \sum_i w_i^J J_i + \sum_k w_k^P P_k \right],$$

where  $w_i^J$  is  $\lambda \times$  class of  $i^{th}$  junction,  
and  $w_k^P$  is similarity score of  $k^{th}$  wall

목적 함수는 모서리와 벽체의 변수에 가중치가 곱해진 선형 결합의 형태로 최대화 문제의 형태이다. 이때, 가중치는 많은 벽체를 사용할 수록, 스타일 변환된 도면과 유사한 구조를 가질 수록 높은 가중치를 가지도록 설정하였다. 모서리의 가중치는 모서리 클래스를 기반으로 많은 숫자의 벽체가 만나는 모서리를 우대하여 더 많은 벽체가 사용되도록 하였다. 벽체 후보군의 가중치는 스타일 변환된 도면에서

구현된 벽체일 경우 높은 점수를 가지도록 설정하였고, 구체적으로 벽체후보군에서 오차의 임계값 이내의 공간에 존재하는 (벽체 픽셀의 수 - 배경 픽셀의 수)의 값이다. 이를 통해 스타일 변환된 도면에서 추출된 벽체를 최대한 사용하면서, 동시에 추출되지 않은 벽체를 배제하는 방향으로 최적화가 진행된다.

## 2.3 제약 조건

본 연구에서 추출한 모서리와 벽체 후보군의 특징을 고려하여 다음 4가지 제약 조건을 계획하여 사용하였다. 각각 폐합, 연결성, 모서리 상호 배제, 벽체 상호 배제의 조건이다.

폐합 조건. 모서리 후보군이 벽체 후보군보다 적게 추출되도록 하여 건물 단위의 폐합을 제약한다. 이를 통해 건물의 외곽 벽체가 열려있는 형태로 출력되지 않도록 한다.

$$\sum_i J_i \leq \sum_k P_k$$

연결성 조건. 각 모서리에서 만나는 벽체 후보군의 숫자가 모서리 클래스와 같도록 제약한다.

$$\sum_j A_{ij} = D_i \cdot J_i \text{ for } \forall i,$$

where  $D_i$  is class of  $i^{\text{th}}$  junction and

$A_{ij} = 1$  if  $i^{\text{th}}$  and  $j^{\text{th}}$  junctions form primitive, 0 otherwise

모서리와 벽체의 상호 배제 조건. 모서리와 벽체 후보군을 대상으로 겹치는 객체들이 존재하지 않도록 제약한다. 이때, 겹침의 정의는 오차의 임계값을 기반으로 한다.

$$\frac{1}{n} \sum_j M_{ij}^J J_i \leq (1 - J_i) \text{ for } \forall i$$

$$\frac{1}{m} \sum_j M_{kj}^P P_k \leq (1 - P_k) \text{ for } \forall k,$$

where  $n$  is num of junctions,  $m$  is num of walls,

$M_{ij}^J = 1$  if  $i^{\text{th}}$  and  $j^{\text{th}}$  junctions overlap, 0 otherwise,

$M_{kj}^P = 1$  if  $k^{\text{th}}$  and  $j^{\text{th}}$  walls overlap, 0 otherwise



## 2.4 최적화 및 후처리

위와 같은 목적 함수와 제약 조건을 통해 정수 계획법을 설계하였고, 그 연산은 최적화 솔루션 Gurobi를 통해 수행하였다. 정수 계획법을 통해 출력되는 결과는 모서리의 연결성 정보를 표현하는 인접 행렬이다. 모서리들의 픽셀 위치와 인접 행렬을 활용하여 최종적인 벡터 포맷의 실내 구조를 구현하였다(그림 4-4). 이때, 모서리 위치를 보정하기 위해서 벽체로 연결된 모서리들 간의 픽셀 좌표를 평균 내어 사용하였다.

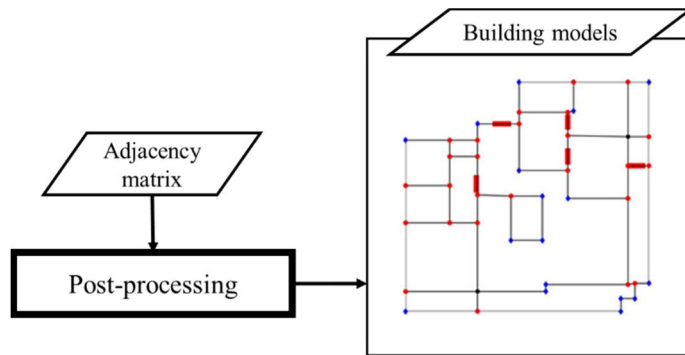


그림 4-4. 후처리를 통한 벡터 포맷의 실내 구조 구현

## 제 5 장            결과 및 정확도 검증

### 제 1 절            정확도 검증 방법론

도면 해석 프로세스들은 연구에 사용한 도면과 다른 유형의 포맷에 적용하기 힘든 특징이 있기 때문에 연구들 간의 직접적인 비교가 수행되기 어렵다. 이에 도면 해석 프로세스의 검증에는 고정적으로 널리 사용되는 방법이 존재하지 않고, 연구의 목적 및 접근 방법 등에 맞추어 다양한 방식을 활용한다. 기본적으로 래스터 형태의 추출물에 대해서는 JI(Jaccard Index)라는 단일 객체에 대한 IoU(Intersection over Union)를 기반으로 검증을 수행하고, 벡터 형태의 추출물은 매칭 테이블 기반의 점수 산출을 통해 검증한다. 이를 기반으로 하되 상황에 맞춰 다양한 방법을 활용한다. 예를 들면, Gimenez *et al.* (2016)는 CVC 도면 데이터셋의 BlackSet을 대상으로 벽체의 추출에 관한 연구를 진행했다. 같은 도면을 대상으로 수행된 이전의 연구들에 비해 낮은 JI 값을 획득했지만, 이는 연구 목적인 실내 공간의 3D 구현의 성능을 반영하지 못하므로 다른 방식의 검증을 새로 수행했다. 구체적으로 설명하면 래스터 추출물의 검증은 벡터 형태의 객체에 대한 추출 성능을 표현하지 못하므로, 벡터 형태로 구현된 벽체를 대상으로 길이 기반의 검증을 수행했다. 이처럼 도면 해석 연구들의 검증은 동일한 조건에서 수행되기 어렵기 때문에, 기존의 연구들과 간접적인 비교를 수행함과 동시에 연구의 목적이 충분히 달성됨을 보이는 방향으로 진행된다.

#### 1.1 참조 연구 선정

세움터 도면 데이터셋은 복잡성과 다양성이 강조된 형태로 기존의 도면 해석 프로세스를 적용하기 어렵다. 문자와 그래픽 요소의 분리가 포함된 해석 프로세스의 경우, 그 분리에 있어 도면 포맷에 의존하는 변수들이 요구되어 다양한 포맷에 대응하기 어려우며, 특히 복잡한 도면에 대해서는 성능이 크게 떨어진다. 주로 사용되는 Tombre의 문자 분리는 세움터 도면에서 갖게 보이는 문자와 그래픽이 겹치는 부분과 벽체와 유사한 격자형 패턴 무늬에 대응하지 못한다. 그래픽 요소의 분리가 불완전한 경우는 허프 변환 기반의 탐지 알고리즘들이 적용되기 어렵고, 또한 포맷의 지엽적인 특징을 활용한 연구들의 범용성이 떨어진다는 점과 맞물려 규칙 기반의 도면 해석 프로세스들은 세움터

도면 데이터셋에 대한 참조 대상으로 사용하기 부적절하다. 학습 기반의 도면 해석 프로세스들을 살펴보면, de las Heras *et al.* (2014)의 영상 기술자 기반의 알고리즘의 경우 딥러닝에 비해 성능을 기대하기 힘들고, Dodge *et al.* (2017)의 딥러닝 기반의 세그먼테이션을 통한 벽체의 추출과 JI를 통한 검증은, 본 연구의 목표가 벽체의 기하학적인 형상의 재현이 아니기 때문에 비교가 부적절하다. 이에 Rakuten 도면을 대상으로 본 연구와 유사한 모서리 추출을 통한 접근을 수행한 Liu *et al.* (2017)의 연구와 비교를 수행한다. 각 연구에서 사용한 데이터셋 및 그라벨링 수준 등의 세부 사항은 상이하지만, 프로세스의 핵심은 모서리 추출이라는 공통된 작업을 위해 각각 세그먼테이션과 스타일 트랜스퍼를 활용하는 것이다. 구체적으로 모서리 추출을 위해서 동일한 목적 함수를 사용하는 반면, 이와 병렬적으로 연결된 네트워크들이 각각 세그먼테이션과 스타일 트랜스퍼의 목적 함수를 사용한다. 이점에 착안하여 세움터 도면 데이터셋을 대상으로 각 연구의 방법을 적용 했을 때 모서리 정보의 추출 성능을 비교한다. 제안한 프로세스를 기존의 연구와 비교하여 복잡한 도면에 대한 스타일 트랜스퍼 적용의 의미를 확인한다.

## 1.2 정확도 검증 방법

본 연구에서는 제안한 프로세스에 대해 두가지 검증을 수행하였다. 복잡한 도면이 대상일 때 모서리 정보의 추출이 기존 연구보다 성능이 높음을 보였고, 매칭 테이블 기반의 방의 탐색률(detection rate) 및 인지 정확도(recognition accuracy)에 대한 검증을 통해 최종 벡터 출력물이 실내의 구조를 표현하는데 충분함을 보였다. 각각은 도면 해석 프로세스 내에서 딥러닝과 정수 계획법의 추출물에 대한 성능 평가에 대응된다. 400장으로 구성된 세움터 도면 데이터셋 중, 학습에 사용하고 남은 50개의 도면을 통해 검증을 수행하였다. 우선, 모서리 정보 출력물은 정밀도(precision)와 재현율(recall)으로 검증하였다. 모서리의 클래스와 위치를 모두 고려하여 정답을 판정했고, 이때 정답과의 픽셀 위치의 차이가 오차 임계값 이내일 경우 경우 일치하는 것으로 판단하였다. 다음으로 벡터 포맷의 실내의 구조에 대해 검증하기 위해 방의 탐색률과 인지 정확도를 사용하였다. 이는 Phillips and Chhabra (1999)이 제안한 매칭 점수표 기반의 평가이고, 다양한 도면 해석

연구에서 활용되는 지표이다(Macé *et al*, 2010; Ahmed *et al*, 2011; de las Heras *et al*, 2014). 이때 승인(acceptance)과 거부(rejection)의 임계값은 기존 연구를 참고하여 각 0.5, 0.1로 하였다(de las Heras *et al*, 2015).

## 제 2 절 결과 및 검증

### 2.1 스타일 트랜스퍼의 검증

본 연구에서는 복잡한 도면에 대한 해석 성능의 향상을 목적으로 스타일 트랜스퍼를 적용하였다. 딥러닝을 활용한 모서리 정보의 추출에 있어서 여러 비교군을 통해 성능을 검증했고, 이 과정에서 Liu *et al.* (2017)의 연구와 성능을 비교하였다.

표 5-1. 손실 함수의 차이를 통한 비교군 설정

	Loss function
①	$\mathcal{L}_J$
②	$\mathcal{L}_J + \lambda \cdot \mathcal{L}_{CNN}$
③	$\mathcal{L}_J + \lambda \cdot \mathcal{L}_{Style\_patch70}$
④	$\mathcal{L}_J + \lambda \cdot \mathcal{L}_{Style\_patch16}$

딥러닝 기반의 모서리 정보 추출에 대한 4가지 비교군을 통해 검증률 수행하였다. 각 방법은 본 연구의 네트워크의 구조를 공유하며, 다른 목적 함수를 통해 모서리 히트맵을 추출하였다(표 5-1). 이때,  $\mathcal{L}_J$ 은 모서리 히트맵을 위한 손실 함수이고,  $\mathcal{L}_{Style}$ 은 본 연구의 스타일 손실 함수이다.  $\mathcal{L}_{CNN}$ 은 CNN 기반의 세그멘테이션의 손실 함수로, 벽체와 출입구 및 배경 대상의 크로스 엔트로피(cross-entropy)를 통해 계산했다. patch 70과 patch 16은 스타일 트랜스퍼를 위한 식별자 네트워크의 패치 크기를 나타낸다. 정리하면, ①은 모서리 라벨링만 활용한 기본 네트워크이고, ②는 모서리 히트맵에 더해, 벽체와 출입구의 세그멘테이션을 활용한, Liu *et al.* (2017)의 알고리즘을 간접적으로 적용한 형태이다. [③과 ④]는 본 연구에서 제안하는 스타일 트랜스퍼 기반의 네트워크이며, 각각 다른 패치 크기의 식별자 네트워크 사용에 따른 결과 비교를 목적으로 한다. ① ~ ④의 손실 함수를 기반으로 추출한 모서리 정보의 평가를 통해, 기존의 연구 방법 대비 스타일 트랜스퍼 적용의 장점을 살펴보았다.

표 5-2. 모서리 정보 추출의 정확도 평가

	①	②	③	④
Precision	0.6487	<b>0.8427</b>	0.8394	0.8327
Recall	0.7281	0.9134	<b>0.9588</b>	0.9416

각 네트워크에 대한 모서리 정보 추출의 정확도 검증은 표 5-2와 같다. ①의 낮은 성능을 통해 네트워크의 파라미터의 공유가 역할을 수행하는 것을 확인할 수 있었다. 또한 CNN 기반의 네트워크보다(②) 스타일 트랜스퍼를 적용한 네트워크(③, ④)가 정밀도는 미세하게 낮고, 재현율은 큰 폭으로 높은 것을 확인하였다. 이는 CNN의 손실 함수가 GAN에 비해 상대적으로 안정적인 선택을 하는 것으로 해석할 수 있고, 정성적인 측면에서 벽체와 출입구에 대한 판단이 모호한 부분에서 스타일 트랜스퍼를 활용한 네트워크가 더 좋은 성능을 보이는 것으로 확인할 수 있었다(그림 5-1). 또한 ③과 ④의 경우, 패치 크기가 70인 네트워크가 16인 것에 비해 더 좋은 성능을 보였다. 작은 크기의 패치를 통한 지엽적인 해석보다 큰 패치를 통한 것이 실내의 구조 파악에 적합하다고 판단 할 수 있었다.

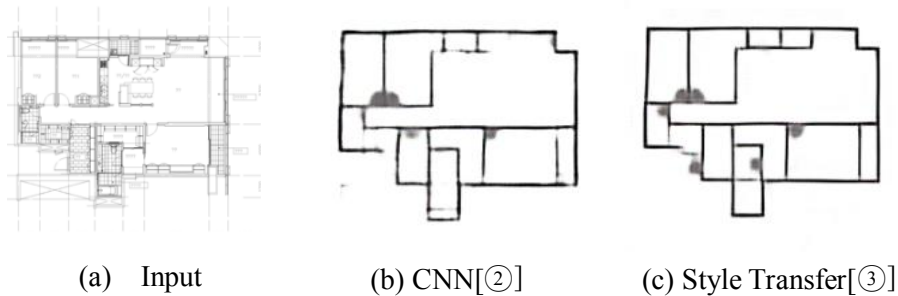


그림 5-1. CNN 통한 분류와 스타일 트랜스퍼의 정성적 비교

본 연구에서 제안하는 도면 해석 프로세스는 딥러닝을 통해 모서리를 추출하고, 이를 기반으로 실내의 구조를 구현하였다. 정수계획법을 통한 보완까지 고려한다면 모서리의 정밀도보다 재현율이 더 중요한 지표이다. 따라서 복잡한 도면의 경우 스타일 트랜스퍼의 적용을 통해 실내 구조 구현의 성능을 높일 수 있음을 간접적으로 확인할 수 있었다. 아래의 절에서 가장 높은 재현율을 보인 ③의 결과를 활용하여 실내의 구조를 구현하였다.

## 2.2 실내 구조 구현을 검증

매칭 테이블 기반으로 방의 탐색률 및 인지 정확도를 통해 실내 구조의 구현에 대한 평가를 수행하였다. 세움터 도면 라벨링은 벽체의 기하학적인 구조보다 실내의 구조, 즉 방의 구분을 우선 기준으로 삼아 구축되었다. 따라서, 세움터 데이터셋을 통한 도면 해석 프로세스는

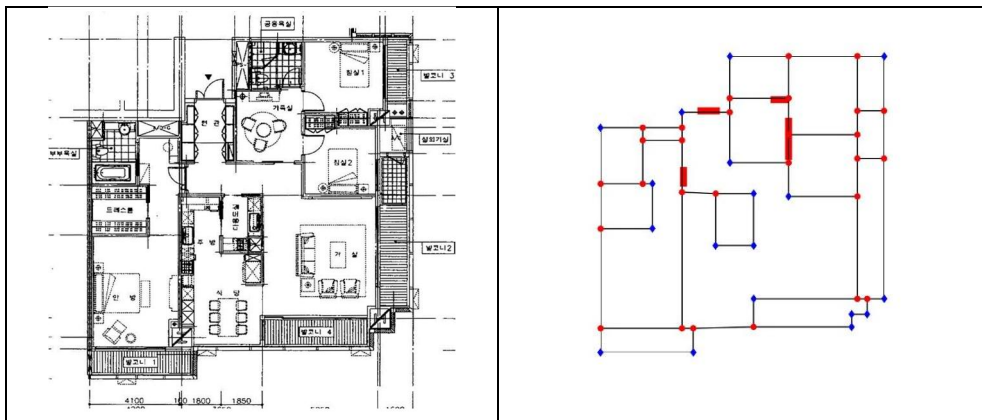
방의 구분을 기준으로 검증하는 것이 적합하다고 판단된다. 세움터 도면 데이터셋에 대해 표 5-3의 방의 탐색률과 인지 정확도를 얻었다. 각 87%와 85%의 결과이고 이는 단순한 포맷에 대한 기존의 연구들의 해석 성능과 유사한 수준이다. 다소 다대일 매칭(Many to one count)이 높게 나온 것은, 다양한 포맷의 해석이 목적인 만큼 문자 정보를 통한 방의 분할과 같은 프로세스를 수행하지 않았기 때문이다. 기존 연구들과 대상 도면 데이터셋이 다르기 때문에 직접적인 비교는 어렵지만, 복잡한 포맷으로 구성된 데이터셋을 사용했음에도 불구하고 충분히 실내의 구조를 구현 할 수 있음을 보였다.

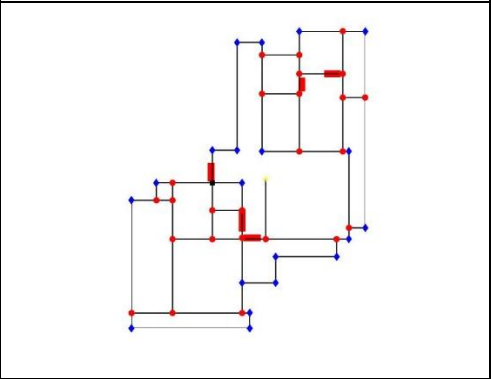
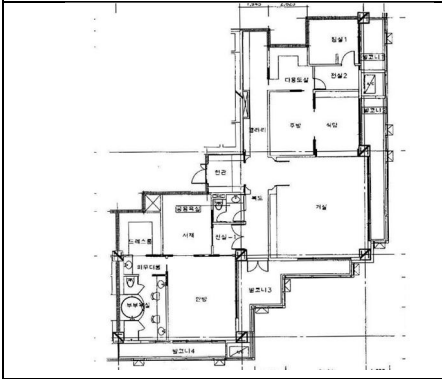
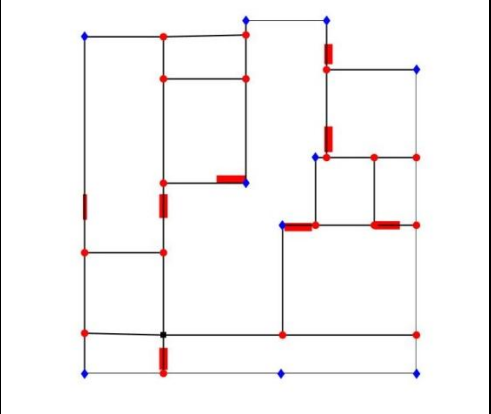
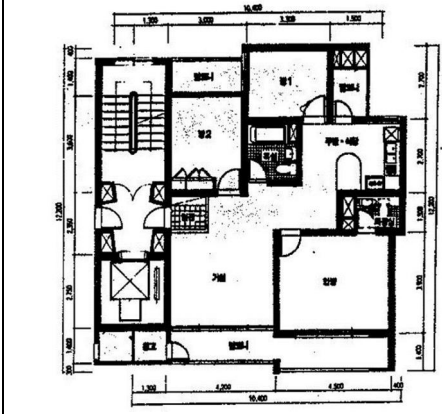
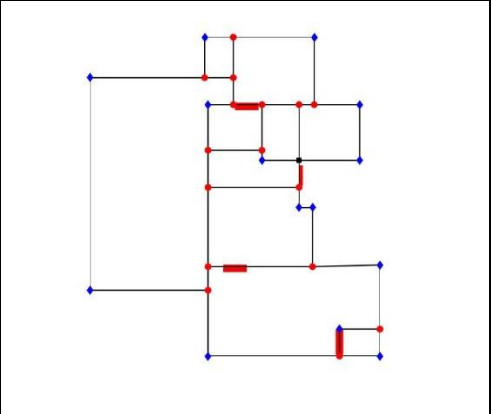
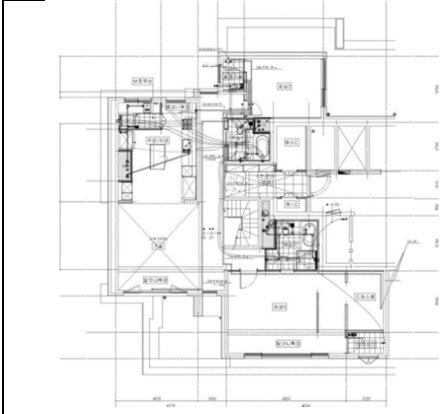
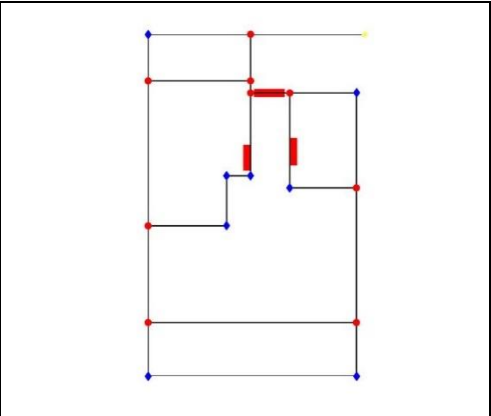
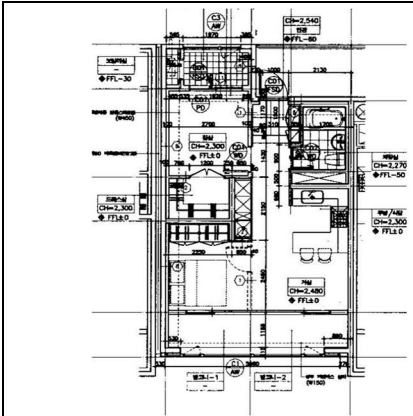
표 5-3. 방의 탐색률 및 인지 정확도

	Proposed
Detection rate (%)	87
Recognition acc. (%)	85
One to many count	1.2
Many to one count	2.4

표 5-4는 본 논문에서 제시하는 도면 해석 프로세스를 통해 구현한 실내 구조의 대표적인 예시들을 보여준다. 입력 도면에 대해서 벽체(검은선)와 출입구(붉은선)를 통해 실내의 구조를 표현하고, 이 때 모서리의 색은 클래스를 표현한다. 단일한 덩어닝 네트워크를 통해, 다양한 포맷의 도면들의 해석을 한 번에 수행하는 것을 확인할 수 있다.

표 5-4. 도면 해석을 통한 실내구조 구현 결과







### 제 3 절 도면 해석 프로세스의 딥러닝 활용

정확도 검증을 통해 제안하는 도면 해석 프로세스가 다양하고 복잡한 포맷의 도면들의 실내 구조를 충분히 구현 할 수 있음을 확인하였다. 도면 해석의 기술 적용의 범용성을 확장하는 측면에서 딥러닝은 큰 장점을 가지고 있지만, 많은 수의 정답 데이터의 준비가 어려워 그 적용에 한계가 있다. 본 연구에서는 350개의 도면을 학습에 활용했지만, Dodge *et al.* (2017)와 Liu *et al.* (2017)는 각 500개와 770개의 도면을 사용했다. 딥러닝 학습에 필요한 데이터의 수는 데이터셋의 특성, 목표 정확도, 네트워크의 구조 등의 상황에 맞추어 다양하게 요구되며, 학습 데이터가 많을수록 성능이 뛰어나다는 일반론적인 접근은 있지만 그 수량에 대해 구체적으로 다뤄지고 있지는 않다. 딥러닝 활용을 위해 데이터 수의 대략적인 판단 기준을 제시하는 것을 목표로, 학습에 사용된 도면의 수의 변화에 따른 모서리 추출의 정확도 및 재현율의 변동을 확인하였다(표 5-5).

그림 5-2는 학습 데이터 수에 따른 모서리의 재현율 변화를 가시화한 것이다. 이 정보를 통해 도면 해석 프로세스에 딥러닝을 적용할 시, 필요한 학습데이터의 규모를 간접적으로 확인할 수 있다. 예를 들어, 세움터 도면 데이터셋을 활용한 본 연구의 경우는 350개 수량에서 학습 데이터 수가 조금만 줄어도 성능이 큰 폭으로 감소하는 것을 통해 확인 할 수 있고, 동시에 학습 데이터 수의 증가에 따라 성능 증진이 가능함을 추정할 수 있다. 향후 딥러닝을 활용한 도면 해석

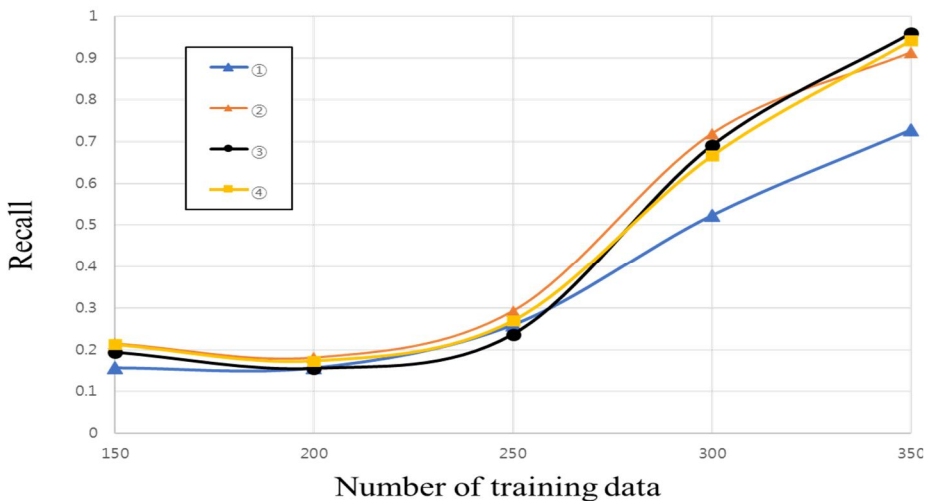


그림 5-2. 학습 데이터 수에 따른 재현율 변화

분야의 연구 데이터가 지속적으로 축적될 경우, 해당 정보를 참조하여 기술을 더 적절히 활용할 수 있을 것이다.

표 5-3. 학습 데이터 수에 따른 정밀도 및 재현율

# of data	Precision				Recall			
	①	②	③	④	①	②	③	④
150	0.3312	0.3228	0.3294	0.3227	0.1573	0.2144	0.1942	0.2126
200	0.3056	0.3814	0.3749	0.3762	0.1570	0.1808	0.1554	0.1734
250	0.3732	0.3857	0.3906	0.3958	0.2600	0.2930	0.2371	0.2696
300	0.5351	0.5187	0.5105	0.5254	0.5238	0.7194	0.6905	0.6661
350	0.6487	0.8427	0.8394	0.8327	0.7281	0.9134	0.9588	0.9416

## 제 6 장            결    론

본 논문에서는 도면 해석 기술의 범용성을 강화하기 위해 복잡하고 다양한 포맷들로 구성된 도면 데이터셋을 대상으로 활용 가능한 도면 해석 프로세스를 개발하였다. 개발한 프로세스는 딥러닝을 활용한 학습에 기반을 둔 것으로, 학습 데이터셋을 구축하기 위한 방법과 학습을 통해 실내 구조를 구현하는 방법의 두 단계로 구성되어 있다. 연구의 대상은 기존 공용 도면 데이터셋들에 비해 복잡하고 다양한 도면들로 구성된 세움터 도면 데이터셋을 새로 구축하여 사용하였다.

두 단계에 대해 보다 자세히 설명하자면 우선 첫 단계로, 실내의 구조를 대상으로 하는 라벨링을 통해 학습 데이터셋을 구축하였다. 이는 도면의 라벨링을 수행할 때 발생하는 모호하고 주관적인 판단 기준들로 야기되는 문제점들을 개선하기 위한 것으로, 객체들의 기하학적인 형태보다 실내 구조의 표현을 우선 기준으로 삼아 진행하였다. 구체적인 라벨링의 대상 객체는 방의 구분을 위한 벽체와 출입구이고 각 객체의 위치 및 모서리에 대한 정보를 표현하였다.

다음 두 번째 단계로, 구축한 정답 데이터셋을 활용하여 도면의 벡터화 프로세스를 진행하였다. 딥러닝 기반의 스타일 트랜스퍼를 활용하여 대상 객체들의 위치와 모서리 정보를 추출했고, 추출된 정보를 기반으로 정수 계획법을 활용하여 실내 구조를 벡터 형태로 구현하였다. 스타일 트랜스퍼는 Conditional GAN을 활용하여 입력된 도면을 실내 구조를 표현하기 위한 형태로 재생성하는 관점으로 접근했고, 그 출력물은 스타일 변환된 도면과 입력 도면에 대응되는 모서리 정보이다. 정수 계획법은 모서리의 연결성 정보를 표현하는 인접 행렬을 출력하기 위해 사용되었으며, 이때 모서리의 정보, 방의 폐합, 벽체와 모서리의 상호 배제 등의 도면의 특징을 반영하는 제약 조건을 활용하여 출력물의 완성도를 높였다. 이로부터 출력된 인접 행렬을 통해 각 객체를 벡터 포맷으로 구현하였다.

본 논문에서는 제안한 프로세스의 유효성을 알아보기 위해 최종적으로 구현한 벡터 포맷의 실내 정보를 기반으로 도면 해석 프로세스를 검증하였다. 개발한 프로세스는 방의 경계를 통해 표현되는 실내 구조의 구현을 목표로 하기 때문에, 방의 탐색률과 인지 정확도를 활용하여 검증을 수행하였다. 세움터 도면 데이터셋에 대해 각 87%와

85%의 결과를 얻었고, 이는 단순한 포맷에 대한 기존의 연구들의 해석 성능과 유사한 수준이다. 또한 세움터 도면 데이터셋에 대한 해석 성능을 최신 연구와 비교하였고, 본 논문이 최종 구현물에 핵심적인 역할을 하는 객체 및 모서리 정보의 출력에서 더 나은 성능을 보이는 것을 확인하였다. 정리하면, 본 논문을 통해 복잡하고 다양한 포맷으로 구성된 도면 데이터셋에 대한 해석이 가능해졌고 이는 도면 해석 기술의 적용 범위를 확장시킨 의의가 있다.

### ■ 연구의 한계 및 보완점

본 논문의 해석 프로세스는 건물의 외곽과 직교하지 않는 벽체, 즉 대각선으로 표현되는 벽체를 구현하지 못한다. 정수 계획법의 벽체 후보군을 생성하는 과정에서 대각선의 벽체를 의도적으로 배제했고 이는 세움터 도면 데이터셋에 대각선의 벽체가 존재하지 않기 때문이다. 간단히 벽체 후보군을 대각선 벽체까지 확장하여 이를 해결할 수 있다. 이를 통해 도면 데이터셋에 대각선의 벽체가 존재하는 경우에도 본 해석 프로세스의 활용이 가능하다.

본 논문에서는 딥러닝을 구조 인식 단계까지 확장하여 활용했고, 벽체의 단순화와 세선화를 비롯해 방의 분할까지를 알고리즘이 내재적으로 수행한다. 이는 실내의 구조를 구현하는 작업을 규칙이 아닌 데이터 기반으로 접근하여 다양하고 복잡한 포맷의 도면에 대해 접근할 수 있는 장점이 있으나, 접근 방법에 따라 정보를 잘못 구현하는 경우가 생긴다. 본 논문의 프로세스에서는 문의 형태로만 출입구를 다루기 때문에 개방된 형태로 구분된 방들에 대한 출입 정보가 손실된다. 예를 들면 거실과 주방 등의 열린 공간에 대해서 본 해석 프로세스는 각각을 닫힌 방으로 출력하여 출입구의 정보를 잃어버린다. 추후에 딥러닝의 출력물에 대한 세분화된 접근 혹은 추가적인 후처리 등을 통해 보완하는 과정이 필요하다.

## 참고 문헌

- AHMED, S., LIWICKI, M., WEBER, M., and DENGEL, A., 2011. Improved automatic analysis of architectural floor plans. In *Document Analysis and Recognition (ICDAR), 2011 International Conference on IEEE*, 864–869.
- BAY, H., TUYTELAARS, T., and VAN GOOL, L., 2006. Surf: Speeded up robust features. In *European conference on computer vision* Springer, 404–417.
- BULAT, A. and TZIMIROPOULOS, G., 2016. Human pose estimation via convolutional part heatmap regression. In *European Conference on Computer Vision* Springer, 717–732.
- CHU, H., KI KIM, D., and CHEN, T., 2015. You are here: Mimicking the human thinking process in reading floor-plans. In *Proceedings of the IEEE International Conference on Computer Vision*, 2210–2218.
- DE LAS HERAS, L.-P., AHMED, S., LIWICKI, M., VALVENY, E., and SANCHEZ, G., 2014. Statistical segmentation and structural recognition for floor plan interpretation. *International Journal on Document Analysis and Recognition (IJDAR)* 17, 3, 221–237.
- DE LAS HERAS, L.-P., TERRADES, O.R., ROBLES, S., and SANCHEZ, G., 2015. CVC-FP and SGT: a new database for structural floor plan analysis and its groundtruthing tool. *International Journal on Document Analysis and Recognition (IJDAR)* 18, 1, 15–30.
- DODGE, S., XU, J., and STENGER, B., 2017. Parsing floor plan images. In *Machine Vision Applications (MVA), 2017 Fifteenth IAPR International Conference on IEEE*, 358–361.
- DOSCH, P. and MASINI, G., 1999. Reconstruction of the 3d structure of a building from the 2d drawings of its floors. In *Document Analysis and Recognition, 1999. ICDAR'99. Proceedings of the Fifth International Conference on IEEE*, 487–490.
- ESCALERA, S., FORNÈS, A., PUJOL, O., ESCUDERO, A., and RADEVA, P., 2009. Circular blurred shape model for symbol spotting in documents. In *Image Processing (ICIP), 2009 16th IEEE International Conference on IEEE*, 2005–2008.
- FLETCHER, L.A. and KASTURI, R., 1988. A robust algorithm for text string separation from mixed text/graphics images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 6, 910–

- GATYS, L.A., ECKER, A.S., and BETHGE, M., 2016. Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* IEEE, 2414–2423.
- GIMENEZ, L., HIPPOLYTE, J.-L., ROBERT, S., SUARD, F., and ZREIK, K., 2015. reconstruction of 3D building information models from 2D scanned plans. *Journal of Building Engineering* 2, 24–35.
- GIMENEZ, L., ROBERT, S., SUARD, F., and ZREIK, K., 2016. Automatic reconstruction of 3D building models from scanned 2D floor plans. *Automation in Construction* 63, 48–56.
- GOODFELLOW, I., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A., and BENGIO, Y., 2014. Generative adversarial nets. In *Advances in neural information processing systems*, 2672–2680.
- HE, K., ZHANG, X., REN, S., and SUN, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.
- ISOLA, P., ZHU, J.-Y., ZHOU, T., and EFROS, A.A., 2017. Image-to-image translation with conditional adversarial networks. *arXiv preprint*.
- KIM, S., PARK, S., and YU, K., 2018. Application of Style Transfer in the Vectorization Process of Floorplans (Short Paper). In *LIPICs–Leibniz International Proceedings in Informatics Schloss Dagstuhl–Leibniz–Zentrum fuer Informatik*.
- LIN, Y.-H., LIU, Y.-S., GAO, G., HAN, X.-G., LAI, C.-Y., and GU, M., 2013. The IFC-based path planning for 3D indoor spaces. *Advanced Engineering Informatics* 27, 2, 189–205.
- [19] LIU, C., SCHWING, A.G., KUNDU, K., URTASUN, R., and FIDLER, S., 2015. Rent3d: Floor-plan priors for monocular layout estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3413–3421.
- LIU, C., WU, J., KOHLI, P., and FURUKAWA, Y., 2017. Raster-to-Vector: Revisiting Floorplan Transformation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2195–2203.
- LONG, J., SHELHAMER, E., and DARRELL, T., 2015. Fully convolutional

- models for semantic segmentation. In *CVPR*, 4.
- MACĚ, S., LOCTEAU, H., VALVENY, E., and TABBONE, S., 2010. A system to detect rooms in architectural floor plan images. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems* ACM, 167–174.
- OTSU, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* 9, 1, 62–66.
- PHILLIPS, I.T. and CHHABRA, A.K., 1999. Empirical performance evaluation of graphics recognition systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 9, 849–870.
- RENDEK, J., MASINI, G., DOSCH, P., and TOMBRE, K., 2004. The search for genericity in graphics recognition applications: Design issues of the qgar software system. In *International Workshop on Document Analysis Systems* Springer, 366–377.
- TOMBRE, K., TABBONE, S., PÉLISSIER, L., LAMIROY, B., and DOSCH, P., 2002. Text/graphics separation revisited. In *International Workshop on Document Analysis Systems* Springer, 200–211.
- WANG, S., FIDLER, S., and URTASUN, R., 2015. Lost shopping! monocular localization in large indoor spaces. In *Proceedings of the IEEE International Conference on Computer Vision*, 2695–2703.
- ZHU, J.-Y., PARK, T., ISOLA, P., and EFROS, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593*.

## Abstract

# Increasing Generality of Floorplan Analysis Process using Style Transfer

Kim, Seongyong

Department of Civil and Environmental Engineering

The Graduate School

Seoul National University

Owing to recent technological advances, location-based services such as tracking, navigation have expanded indoor and the size of the related markets has been increased. As the demand for indoor services and researches increases, constructing indoor information is emerging as a primary concern. Among them is the method using a floorplan, viewed as an effective way to construct information since a floorplan has a high accessibility. It is required to improve versatility of technology to apply to a wide range of floorplan formats. Floorplan analysis has been developed in this way by introducing learning algorithms. However there is a limitation on utilizing technology since previous studies were performed with comparatively simple floorplans and it is hard to construct training data set. The main aims of this research are improving performance of analyzing complex floorplans by applying style transfer, and developing floorplan analysis process which can be generally utilized by suggesting simple criteria to construct labelling data. More specifically, this research constructs new dataset which is more various and intricate than the previous one, and performs vectorization of building factors using deep learning based style transfer. Regarding a format of floorplan as a style, deep learning network outputs information on indoor structure in the process of transferring a style of floorplan. Analysis performance on complex floorplans has been improved by using style transfer, and in terms



of constructing indoor structure, it shows 87% of detection rate and 85% of recognition accuracy of a room, which is a similar level as the previous researches. To sum up, style transfer lets us do floorplan analysis on complex and diverse formats, easily constructing labelling data.

**Keywords : Indoor Information, Floorplan Analysis, Style Transfer, Integer Programming**

**Student Number : 2017-23581**