



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

실내 공간정보 자동 추출을 위한
핸드드로잉 건축도면의 머신러닝 기반
이진화 기법 개발

**Development of Machine Learning Based
Binarization Technique of Hand-drawn Floor Plans
for Automatic Extraction of Indoor Spatial
Information**

2022년 8월

서울대학교 대학원

건설환경공학부

서한유

실내 공간정보 자동 추출을 위한 핸드드로잉 건축도면의 머신러닝 기반 이진화 기법 개발

Development of Machine Learning Based Binarization Technique of Hand-drawn Floor Plans for Automatic Extraction of Indoor Spatial Information

지도교수 유 기 윤

이 논문을 공학석사 학위논문으로 제출함

2022년 4월

서울대학교 대학원

건설환경공학부

서 한 유

서한유의 공학석사 학위논문을 인준함

2022년 6월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

국문초록

최근 인공지능, 사물인터넷 등의 발전과 함께 사용자의 위치를 파악하여 실시간 정보를 제공하는 실내 위치기반 서비스에 대한 사회적 관심이 높다. 이러한 실내 위치기반 서비스의 활성화를 위해서는 실내 공간의 모습을 표현하는 실내 구조 형상화 및 모델링이 필수적이다. 이에 따라 레이저 스캐너, 건축도면 이미지, CAD 플랜 등 다양한 원천 데이터로부터 실내 공간을 재현하는 연구들이 진행되어 왔다. 특히 실내 공간정보를 자동 추출 기술은 수동 모델링 대비 경제적으로 매우 효율적이다. 이에 2차원 건축도면 이미지 데이터로부터 벽, 창문, 계단과 같은 실내 객체를 자동 추출하여 3D 모델링 데이터를 구축하는 도면 해석 연구가 활발히 진행 중에 있다. 기존의 2차원 사진 기반 도면 해석 연구들은 객체와 배경이 명확히 구분되며 객체가 일정한 색으로 표현된 전자 도면을 대상으로 연구를 수행하였다. 하지만, 펜과 잉크를 사용해 작성된 핸드드로잉 도면의 경우 기존 연구에 사용된 도면에 비해 노이즈가 많고 배경 패턴이 불규칙적이다. 또한 사용된 펜이나 잉크에 따라 객체의 색상값이 일정하지 않기 때문에 기존 실내 공간 객체 추출 알고리즘을 적용하는 데에 한계가 존재한다. 이에 본 연구는 노이즈가 심하고 불규칙적인 핸드드로잉 건축도면을 대상으로 실내 공간을 구성하는 객체와 배경을 구분하는 이진화를 수행하고자 한다. 본 연구는 전자 도면 대상의 기존 실내 공간정보 자동 추출 연구의 범위를 역사적 건축물이나 건축 연도가 오래되어 아날로그 방식으로 작성된 건축도면만 존재하는 건물을 대상으로 확장하는 것을 목표로 한다. 분석 데이터로서 1900년대 초반에 작성된 일제시기 건축도면을 활용하여 연구를 수행하였다. 본 연구에 사용된 일제시기 건축도면은 종이류 문화재 특성상 보관 및 디지털화 과정에서 다양한 형태의 노이즈가 존재하며 작성 시 사용된 필기류 종류에 따라 객체의 색상 값이 일정하지 못하다.

또한 핸드드로잉 건축도면 이미지마다 나타나는 노이즈의 픽셀값과 실내 객체의 선명도가 다르기 때문에 머신러닝 모델을 사용한 학습 기반 이진화 기법을 적용하였다. 이진화는 제거하고자 하는 노이즈의 형태에 따라 크게 두 가지 단계로 진행된다. 첫 번째 단계는 가우시안 혼합 모델을 사용하여 도면 이미지의 배경에 전체적으로 넓게 분포하는 노이즈를 감소시키는 단계이다. 두 번째 단계는 랜덤포레스트 모델을 기반으로 객체와 배경을 구분하는 특징을 추출하여 면적이 작고 다양한 형태의 노이즈를 학습 및 제거시키는 단계이다. 마지막으로 제안한 방법론에 대한 검증을 수행하기 위해 학습 과정에 사용되지 않은 테스트 셋에 대한 분류 모델 성능 평가와 최종 결과 이미지에 대한 이미지 품질 평가를 진행했다. 실험 결과, 분류 모델 성능 평가의 경우 랜덤포레스트 모델의 평균 정밀도 및 재현율은 각각 0.985 와 0.99 이고 최종 이진화 결과 이미지의 신호 대비 잡음 비 지표는 16.543 의 결과를 얻었다. 이진화 결과, 선행 연구 대비 다양한 두께로 구성된 벽, 창문, 가벽과 같은 실내 공간 객체와 배경을 성공적으로 분리하였다. 또한 모델의 일반화 성능 검증을 위해 베르사유 궁전 건축도면에 대해 본 연구의 이진화 알고리즘을 적용하였다. 적용 결과, 정밀도 및 재현율은 각각 0.998 와 0.969 이고 결과 이미지의 품질을 평가하는 지표 역시 테스트 셋과 유사하게 우수한 성능을 나타냈다. 본 연구는 기존 도면 해석 연구의 활용처를 핸드드로잉 건축도면으로 확장하는 기반을 마련했다는 점에서 의의가 있다.

주요어 : 도면 해석, 실내 공간정보, 머신러닝, 랜덤포레스트, 이진화

학 번 : 2019-23752

목 차

1. 서론.....	1
1.1 연구 배경 및 목적.....	1
1.2 이진화 연구 동향	4
1.2.1 규칙 기반 이진화 방법론.....	7
1.2.2 학습 기반 이진화 방법론.....	10
1.2.3 시사점 및 결론.....	12
1.3 연구 범위 및 방법.....	14
2. 연구 방법.....	17
2.1 데이터 수집 및 전처리.....	17
2.2 배경 예측 및 제거.....	19
2.2.1 픽셀값 빈도 분석.....	19
2.2.2 이상값 필터링.....	21
2.2.3 가우시안 혼합 모델.....	24
2.2.4 배경 제거 이미지 생성	26
2.3 머신러닝 기반 도면 이진화	27
2.3.1 특징 추출	27
2.3.1.1 통계적 특성.....	30
2.3.1.2 명암도 동시행렬의 통계적 특성	31
2.3.1.3 수직-수평 연속성 행렬.....	35
2.3.2 랜덤포레스트 모델.....	38
2.3.3 재귀적 특징 제거법	42
2.3.4 평가지표	44
2.4 후처리.....	47

3. 실험 적용 및 결과	49
3.1 데이터 수집 및 전처리 결과.....	49
3.2 배경 예측 및 제거 결과	52
3.3 특징 추출 결과.....	56
3.3.1 명암도 동시발생 행렬 특징 추출 결과	56
3.3.2 수직-수평 연속성 행렬 특징 추출 결과.....	58
3.4 머신러닝 기반 도면 이진화 평가 결과.....	59
3.4.1 특징 중요도 및 최적 특징 조합	59
3.4.2 분류 모델 성능 비교	63
3.4.3 이진화 결과 이미지의 품질 비교.....	65
3.5 머신러닝 기반 도면 이진화 적용 결과.....	68
3.5.1 소축척 도면에서의 이진화 적용 결과.....	70
3.5.2 대축척 도면에서의 이진화 적용 결과.....	72
3.6 다양한 핸드드로잉 건축도면의 이진화 평가 및 적용 결과.....	74
3.6.1 직선 객체로 구성된 베르사유 궁전 건축도면의 이진화.....	75
3.6.2 곡선 객체를 포함하는 베르사유 궁전 건축도면의 이진화 ...	76
4. 결론.....	79
참 고 문 헌.....	82
부 록.....	86
Abstract	112

그림 목 차

[그림 1-1] 기존 실내 공간정보 추출 연구에 사용된 도면 데이터 셋 .2	2
[그림 1-2] 본 연구에 사용된 핸드드로잉 건축도면..... 3	3
[그림 1-3] 연구 흐름도..... 14	14
[그림 2-1] 일제시기 건축도면 아카이브 홈페이지 화면..... 18	18
[그림 2-2] SupperAnnotate 라벨링 화면 예시..... 18	18
[그림 2-3] 도면 데이터 셋에 따른 픽셀값 히스토그램 예시..... 20	20
[그림 2-4] 4방향 변위 벡터 예시..... 31	31
[그림 2-5] 명암도 동시발생 행렬 구성 예시..... 32	32
[그림 2-6] DIBCO 데이터 셋과 핸드드로잉 건축도면 데이터 셋 비교 35	35
[그림 2-7] 연속성 특징 벡터 추출 (Vargas, 2018)..... 36	36
[그림 2-8] 수평 연속성 행렬 계산 예시 37	37
[그림 2-9] 수직 연속성 행렬 계산 예시 37	37
[그림 2-10] 랜덤포레스트 모델 학습 흐름도 38	38
[그림 2-11] 재귀적 특징 제거법 흐름도 43	43
[그림 2-12] 거리에 따른 오분류 픽셀 예시 (Lu, 2004) 45	45
[그림 2-13] 방향성에 따른 연결성 판단 종류..... 47	47
[그림 2-14] 연결 성분 분석에 따른 픽셀 라벨링 결과..... 48	48
[그림 3-1] 라벨링 결과 (도면 2번, 6번) 51	51
[그림 3-2] 양방향 필터 적용 전후 픽셀값 분포 예시 53	53
[그림 3-3] 가우시안 혼합 모형 적용 후 픽셀값 군집화 예시 54	54
[그림 3-4] 예측 배경 이미지와 배경 제거 이미지 예시..... 54	54
[그림 3-5] 배경 예측 및 제거 단계 이후 잔여 노이즈 예시 (1)..... 55	55

[그림 3-6] 배경 예측 및 제거 단계 이후 잔여 노이즈 예시 (2).....	55
[그림 3-7] 명암도 동시발생 행렬 특징 추출 예시	56
[그림 3-8] 특징 추출 위치에 따른 명암도 동시발생 행렬 평균과 표준 편차 간 관계성.....	57
[그림 3-9] 수직-수평 연속성 행렬 특징 추출 예시	58
[그림 3-10] 특징 개수에 따른 정확도 변화.....	59
[그림 3-11] 특징 별 상대적 중요도.....	60
[그림 3-12] 핵심 특징 추출 예시	61
[그림 3-13] 머신러닝 기반 이진화 적용 후 노이즈 제거 결과.....	68
[그림 3-14] 소축척 도면의 이진화 적용 결과.....	70
[그림 3-15] 소축척 도면의 선형객체 별 이진화 적용 결과	71
[그림 3-16] 대축척 도면의 이진화 적용 결과.....	72
[그림 3-17] 대축척 도면의 선형객체 별 이진화 적용 결과	73
[그림 3-18] 객체 형상에 따른 베르사유 궁전 건축도면 분류.....	74
[그림 3-19] 다양한 두께의 곡선 객체 이진화 결과.....	78
[그림 A-1] 대한의원의학교 평면도 (도면 14번) 이진화 결과.....	86
[그림 A-2] 경성세관출장소청사 평면도 (도면 41번) 이진화 결과...	87
[그림 A-3] 부산경찰서청사 설계도 (도면 43번) 이진화 결과.....	88
[그림 A-4] 갑오관사 설계도 (도면 48번) 이진화 결과.....	89
[그림 A-5] 경성의학전문학교 부속병원 외래진료소 평면도 (도면 38번) 이진화 결과.....	90
[그림 A-6] 해주자혜의원 본관 설계도 (도면 34번) 이진화 결과.....	91
[그림 B-1] 직선 객체로 구성된 베르사유 궁전 건축도면 1번 이진화 결과	92
[그림 B-2] 직선 객체로 구성된 베르사유 궁전 건축도면 2번 이진화 결과.....	93
[그림 B-3] 직선 객체로 구성된 베르사유 궁전 건축도면 3번 이진화	

결과.....	94
[그림 B-4] 직선 객체로 구성된 베르사유 궁전 건축도면 4번 이진화 결과.....	95
[그림 B-5] 직선 객체로 구성된 베르사유 궁전 건축도면 5번 이진화 결과.....	96
[그림 B-6] 직선 객체로 구성된 베르사유 궁전 건축도면 6번 이진화 결과.....	97
[그림 B-7] 직선 객체로 구성된 베르사유 궁전 건축도면 7번 이진화 결과.....	98
[그림 B-8] 직선 객체로 구성된 베르사유 궁전 건축도면 8번 이진화 결과.....	99
[그림 B-9] 직선 객체로 구성된 베르사유 궁전 건축도면 9번 이진화 결과.....	100
[그림 B-10] 직선 객체로 구성된 베르사유 궁전 건축도면 10번 이진화 결과.....	101
[그림 C-1] 곡선 객체를 포함하는 베르사유 궁전 건축도면 1번 이진화 결과	102
[그림 C-2] 곡선 객체를 포함하는 베르사유 궁전 건축도면 2번 이진화 결과.....	103
[그림 C-3] 곡선 객체를 포함하는 베르사유 궁전 건축도면 3번 이진화 결과.....	104
[그림 C-4] 곡선 객체를 포함하는 베르사유 궁전 건축도면 4번 이진화 결과.....	105
[그림 C-5] 곡선 객체를 포함하는 베르사유 궁전 건축도면 5번 이진화 결과.....	106
[그림 C-6] 곡선 객체를 포함하는 베르사유 궁전 건축도면 6번 이진화 결과.....	107
[그림 C-7] 곡선 객체를 포함하는 베르사유 궁전 건축도면 7번 이진화 결과.....	108
[그림 C-8] 곡선 객체를 포함하는 베르사유 궁전 건축도면 8번 이진화 결과.....	109

[그림 C-9] 곡선 객체를 포함하는 베르사유 궁전 건축도면 9번 이진화 결과.....	110
[그림 C-10] 곡선 객체를 포함하는 베르사유 궁전 건축도면 10번 이 진화 결과.....	111

표 목 차

[표 1-1] 문서 이진화 관련 연구 요약.....	6
[표 2-1] 선행 연구에서 사용된 필터 종류 예시.....	21
[표 2-2] 이진화를 위해 사용된 특징 분류.....	29
[표 2-3] 평가 항목에 따른 평가 지표 및 수식.....	46
[표 3-1] 데이터 셋 도면 목록.....	49
[표 3-2] 분류 모델 성능 비교표.....	64
[표 3-3] 이진화 결과 품질 비교표.....	67
[표 3-4] 직선 객체로 구성된 베르사유 궁전 건축도면의 이진화 평가 결과.....	75
[표 3-5] 곡선 객체를 포함하는 베르사유 궁전 건축도면의 이진화 평가 결과.....	77

용어정의

■ 선형객체

- : 건축도면은 벽, 창문, 문, 계단과 같이 실내공간을 구성하는 다양한 객체로 구성되는데, 본 연구에서는 이러한 객체들을 종합하여 선형객체라고 표현한다. 즉, 종이 위에 필기구를 사용해 작성된 모든 요소들은 선형객체이다.

■ 픽셀값(pixel value)

- : 디지털 이미지를 구성하는 각 픽셀의 색깔이나 밝기를 표시하는 값을 말한다. 본 연구에서 사용된 핸드드로잉 건축도면의 흑백 이미지의 경우 0에서 255사이의 값들 중 하나를 갖게 되는데, 픽셀값이 0인 경우는 가장 어두운 상태인 검정색을 나타내고 255는 가장 밝은 상태인 흰색을 나타낸다.

■ 이진화(binartization)

- : 이진화란 영상 내 모든 픽셀을 오로지 흑과 백으로만 표현하는 것을 말한다. 본 연구에서는 입력으로 사용되는 핸드드로잉 건축도면 이미지 내 모든 픽셀을 배경과 선형객체로 구분하여 배경에 해당하는 픽셀은 백으로, 선형객체에 해당하는 픽셀은 흑으로 표현하는 이진화를 수행하였다.

■ 노이즈(noise)

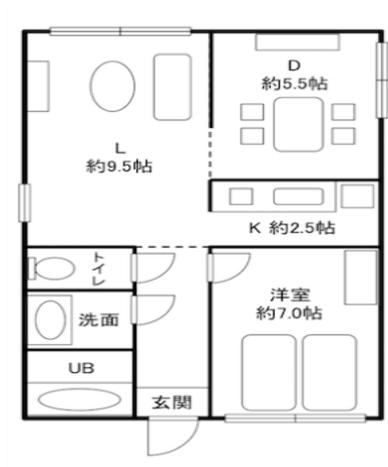
- : 핸드드로잉 건축도면에서 발생하는 노이즈란 종이의 노후화로 인해 발생하는 얼룩짐이나 보관 과정에서 생긴 접힘 흔적을 예로 들 수 있다. 이러한 노이즈는 주로 도면의 배경에 해당하는 영역에 나타나며, 일반적인 배경의 픽셀값보다 낮은 픽셀값을 갖는다. 본 연구에서는 이진화를 통해 노이즈에 해당하는 픽셀들의 픽셀값을 255로 변환해 배경화 함으로써 선형객체와 구별하였다.

1. 서론

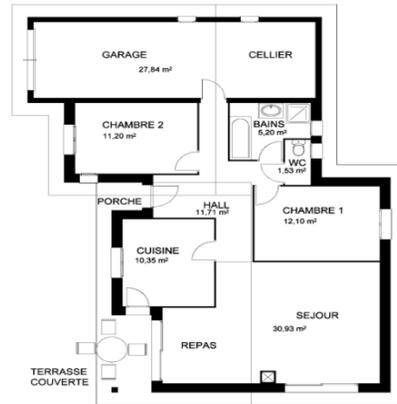
1.1 연구 배경 및 목적

최근 인공지능, 사물인터넷 등의 발전과 함께 사용자의 위치를 파악하여 실시간 정보를 제공하는 실내 위치기반 서비스에 대한 사회적 관심이 높다. 이러한 실내 위치기반 서비스의 활성화를 위해서는 실내 공간의 모습을 표현하는 실내 구조 형상화 및 모델링이 필수적이다. 이에 따라 레이저 스캐너, 건축도면 이미지, CAD플랜 등 다양한 원천 데이터로부터 실내 공간을 재현하는 연구들이 진행되어 왔다 (Kim *et al.*, 2021, Kim *et al.*, 2021). 이러한 기존 도면 해석 연구의 목표는 건축도면 이미지나 CAD파일 등과 같은 2차원 데이터로부터 기하학적 요소를 기반으로 하여 객체를 자동 추출해 실내 공간을 재현하는 것이다.

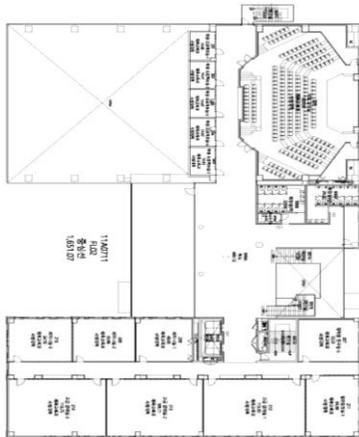
Ahmed *et al.* (2011)은 도면 이미지로부터 텍스트와 그래픽 성분을 분리한 후, 그래픽 성분에서 벽체의 패턴을 기반으로 방 구조를 인식하여 실내 공간을 재현하였고, Liu *et al.* (2017)은 CNN(Convolutional Neural Network)을 통해 도면 이미지로부터 다양한 종류를 지닌 모서리 정보를 추출한 뒤 정수 프로그래밍 기반의 벡터화 과정을 통해 3차원 실내 공간 모델을 재현하였다. 이와 같은 도면의 형태와 구성 요소로부터 특징을 추출하는 도면 해석 알고리즘들은 도면 데이터 고유 특성에 매우 의존적이다(김성용, 2018). 특히 머신러닝(Machine Learning)이나 딥러닝(Deep Learning)을 이용한 이미지 기반 도면 해석 연구의 경우, 이미지의 해상도와 노이즈 정도에 따라 추출되는 특징이 달라질 수 있기에 [그림 1-1]과 같이 도면과 배경 픽셀값이 명확히 구분되는 깨끗한 이미지의 사용이 필수적이다.



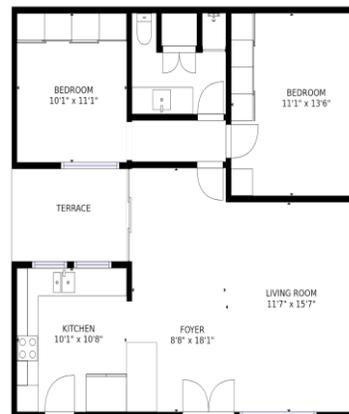
(a) Rakuten



(b) CVC-FP



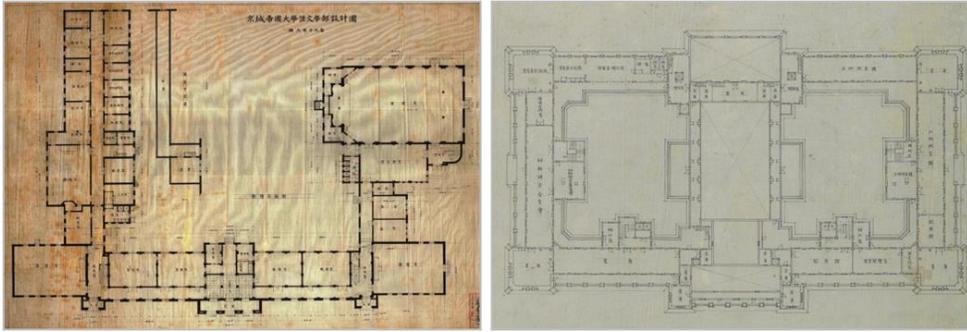
(c) 서울대학교



(d) CubiCasa-FP

[그림 1-1] 기존 실내 공간정보 추출 연구에 사용된 도면 데이터 셋

하지만 [그림 1-1]의 데이터 셋과 달리 핸드드로잉 도면의 경우, 배경 밝기가 불규칙적이며 사용된 펜이나 연필에 따라 도면 내 실내 객체의 색깔이 다르다([그림 1-2]).



[그림 1-2] 본 연구에 사용된 핸드드로잉 건축도면

특히 일제시기에 작성된 근대 건축도면과 같은 종이류 문화재의 경우 제작 연도가 100년 이상 되어 노이즈의 정도가 매우 심각한 수준이다. 이처럼 핸드드로잉 건축도면은 작성 도구에 따른 객체의 색깔 불일치, 종이의 노화로 인한 얼룩, 디지털화 과정에서 발생한 그림자와 같은 노이즈로 인해 배경과 실내 객체 간 명확한 픽셀값 구분이 어렵다. 또한 핸드드로잉 건축도면은 일반적으로 종이 도면을 스캔하여 디지털 파일로 보관되는데, 스캔 전 도면의 보관 상태에 따라 종이 구겨짐, 노후화에 따른 노이즈가 존재하며 스캔 파일 역시 동일한 노이즈가 남아있다. 이와 같은 이유로 핸드드로잉 건축도면은 기존 도면 해석 연구의 입력 데이터로 활용하기에 어려움이 존재했다. 만약 기존 도면 해석 연구를 핸드드로잉 건축도면으로 확장한다면 조선총독부 청사와 같은 현존하지 않는 역사적 건축물의 실내 공간을 시각적으로 재현할 수 있고, 청사진 형태로 보관되어 있는 건축물의 실내 공간 정보를 모델링 및 전산화할 수 있다. 따라서 본 연구는 기존 도면 해석 연구의 적용 범위를 핸드드로잉 건축도면으로 확장하기 위해 배경과 선형객체의 픽셀값을 구분하는 이진화 알고리즘을 개발하는 것을 목적으로 한다.

1.2 이진화 연구 동향

이미지 데이터를 활용한 연구에서는 노이즈를 축소하고 화질을 개선하는 이진화 단계가 매우 중요한 전처리 단계로 사용되어 왔다 (Mitianoudis *et al.*, 2015; Xiong *et al.*, 2021; Chen *et al.*, 2016). 이진화는 입력 이미지의 모든 픽셀을 배경(background)와 객체(object)라는 두 개의 하위 항목으로 분류하여 픽셀값을 0 또는 255로 만드는 연산이다. 이러한 이진화 연구는 특히 손으로 작성된 문서를 디지털화하기 위해 활발히 진행되었으며 특히 다양한 손 글씨로 작성된 과거 문헌을 바탕으로 구성된 DIBCO(International Competition on Document Image Binarization) 데이터 셋을 주로 활용하였다.

과거부터 가장 활발히 사용되어 오는 이진화 방법인 Otsu (1979) 알고리즘은 회색조로 구성된 영상 전체에 대하여 임의의 픽셀값 k 에 의해서 분할된 군집의 내부 분산(within-class variance)을 최소화하거나 혹은 군집 간 분산(between-class variance)을 최대화 시키는 전역 임계값(global thresholding value)을 찾는 기법이다. 하지만 Otsu의 이진화 방법은 입력 영상의 히스토그램이 쌍봉(bimodal)이라는 전제하에 효과적으로 작동한다. 핸드드로잉 건축도면 이미지나 과거 문헌 이미지와 같이 다양한 픽셀값이 분포하는 이미지에 적용하였을 때는 매우 낮은 효과를 보였다(Xiong *et al.*, 2021). 또한 이와 같은 전역 이진화 방법은 이미지 전체를 입력하여 결정된 하나의 임계값을 모든 픽셀에 적용하는 방법으로써, 이미지 내 명도가 일정하지 않은 대상에 적용하기 어려운 부분이 있다(Su *et al.*, 2010).

이후 이진화 알고리즘은 Otsu (1979)가 제안한 방법을 토대로 이미지를 국소 영역으로 분할하여 적용하는 이른바 지역 임계처리(local thresholding) 알고리즘이 제안되었다(Niblack, 1986; Sauvola *et al.*, 2000; Wolf *et al.*, 2003). 지역 임계값은 하나의 픽셀을 이진화하기 위해 각 픽셀에 인접한 이웃 픽셀을 참조하여 결정되기 때문에 윈도우

사이즈와 문자 객체의 두께에 민감하게 반응하는 한계가 존재했다(Su *et al.*, 2010).

과거 문헌 이미지를 활용한 이진화 연구들은 이러한 전통적인 이진화 알고리즘의 한계를 극복하고자 배경과 객체를 구분할 수 있는 특징(feature) 정보를 활용하였으며 특징 정보를 분류하는 과정에서 크게 규칙기반의 방법론과 학습기반의 방법론으로 나뉜다([표 1-1]). 규칙기반 방법론에서는 객체 픽셀 두께(stroke edge width), 국소 대조(local contrast) 등의 특징을 활용해 문자의 가장자리를 검출하여 이를 파라미터 삼아 규칙 기반으로 배경과 객체를 분리하는 방식으로 이진화를 진행하였다. 반면에 학습기반의 이진화 연구는 이미지로부터 배경과 객체를 구분하는 특징을 스스로 학습하는 방식을 따른다. 학습기반 방법론의 경우 연구자가 직접 특징을 추출하여 학습하는 머신러닝과 모델이 스스로 특징을 추출하여 학습하는 딥러닝으로 나눌 수 있다.

[표 1-1] 문서 이진화 관련 연구 요약

연구자	논문 제목	연구내용	
		규칙기반	학습기반
Lu (2010)	Document Image Binarization using Background Estimation and Stroke Edges	●	
Su (2010)	Binarization of Historical Document Images using the Local Maximum and Minimum	●	
Su (2013)	Robust Document Image Binarization Technique for Degraded Document Images	●	
Jia (2018)	Degraded Document Image Binarization using Structural Symmetry of Strokes	●	
Chen (2016)	Parallel Nonparametric Binarization for Degraded Document Images		●
Xiong (2018)	Degraded Historical Document Image Binarization using Local Features and Support Vector Machine		●
Bezmaternykh (2019)	U-net-bin: Hacking the Document Image Binarization Contest		●
Meng (2017)	Deep Networks for Degraded Document Image Binarization through Pyramid Reconstruction		●
Chris Tensmeyer, Tony Martinez (2017)	Document Image Binarization with Fully Convolutional Neural Networks		●

1.2.1 규칙 기반 이진화 방법론

Lu (2010)은 먼저 주어진 이미지(I)에 수직, 수평 방향으로 다항식 평활 필터(polynomial smoothing filter)를 반복 적용하여 배경 이미지(BG)를 추정하고, 추정된 배경 이미지를 활용해 정규화 이미지(\bar{I})를 생성하였다(식 1-1).

$$\bar{I} = \frac{C}{BG} \times I \quad (1-1)$$

다음 단계로 정규화된 이미지의 수직(V_v), 수평(V_h) 방향의 경사도(gradient)를 계산하여 식 (1-2)와 같이 문서 픽셀의 가장자리를 검출하였으며, 이후 인접 가장자리 픽셀까지의 평균 거리를 기반으로 지역 임계 처리를 위한 윈도우 사이즈를 결정하였다.

$$\begin{aligned} V_h(x, y) &= |\bar{I}(x, y + 1) - \bar{I}(x, y - 1)| \\ V_v(x, y) &= |\bar{I}(x + 1, y) - \bar{I}(x - 1, y)| \\ V(x, y) &= V_h(x, y) + V_v(x, y) \end{aligned} \quad (1-2)$$

그 결과, 이미지 내 문서 픽셀의 평균 두께를 고려하여 지역적인 임계처리를 통해 이진화를 수행했을 때, 전체 이미지를 한 번에 이진화를 수행한 것보다 결과가 우수했다. 하지만 배경과 문자의 픽셀값 차가 뚜렷하지 않은 이미지에 대해서는 문서 픽셀의 가장자리의 경사도 변화가 뚜렷하지 않아 좋지 못한 성능을 보였다.

Su (2010)은 3×3 크기의 슬라이딩 윈도우를 이동시키며 이미

지의 모든 픽셀을 대상으로 윈도우 내 최대, 최소값을 이용해 대조값 (local contrast)을 계산하여 문서 픽셀의 가장자리를 검출하였다. 윈도우 내 픽셀이 모두 배경에 해당한다면 대조값은 상대적으로 작고, 윈도우 내 픽셀에 배경과 문서에 해당하는 픽셀이 섞여있다면 대조값은 상대적으로 클 것이다. 이러한 배경에서 문서로 바뀌는 영역의 가장자리 픽셀의 대조값 변화를 탐지하여 가장자리를 검출하고, 검출된 가장자리 픽셀값과 유사한 인근 문서 픽셀을 추가적으로 추출하여 이진화를 진행하였다. Lu (2010)과 마찬가지로 Su (2010) 또한 배경과 문서의 픽셀값 차이가 유의미 하지 않은 영역에 대해서는 좋지 못한 성능을 보였다. 이는 두 방법론 모두 배경에서 문서로 바뀌는 영역의 픽셀값 변화가 두드러진다는 가정 하에 가장자리를 탐지하였기 때문이다.

Su (2013)은 이러한 한계를 가진 Su (2010)의 문서 가장자리 탐지 방법을 기반으로 배경과 문서 픽셀 간 대조값이 작은 영역에서도 적용 가능하도록 알고리즘을 고도화하였다. 구체적으로 식 (1-3)과 같이 Su (2010)의 대조값 계산 식은 픽셀값 변화가 큰 영역에서 분모가 커짐으로써 큰 값을 갖기 때문에 대조가 적은 영역—예컨대 배경이 문자만큼 어둡거나, 문자가 배경만큼 밝거나—의 탐지 성능이 낮다는 한계점이 발생하게 된다. 식 (1-3)에서 ϵ 는 분모가 0일 경우 추가되는 0에 가까운 상수이다.

$$C(x, y) = \frac{I_{\max}(x, y) - I_{\min}(x, y)}{I_{\max}(x, y) + I_{\min}(x, y) + \epsilon} \quad (1-3)$$

특히 식 (1-3)을 사용해 대조값을 구할 경우 문자 영역의 픽셀값이 배경만큼 밝을 때는 문자보다 분모가 훨씬 더 큰 값을 갖게 된다. 따라서 대조값이 과대 정규화(over normalization)되는 문제가 발생한

다. 이러한 문제를 극복하고자 Su (2013)은 식 (1-4)과 같이 대조값이 낮은 경우, 이미지의 표준 편차가 낮다는 점을 활용한 대조값 식을 제안하였다. 식 (1-4)에서는 식 (1-5)와 같은 표준 편차 기반 가중 계수 (α)을 사용하여 대조가 낮은 이미지에 대한 과소 표준화 경향을 보완하였다. 식 (1-5)에서 γ 는 대조값에 따라 변하는 상수이다.

$$C_{\alpha}(x, y) = \alpha C(x, y) + (1 - \alpha)(I_{\max}(x, y) - I_{\min}(x, y)) \quad (1-4)$$

$$\alpha = \left(\frac{STD}{128}\right)^{\gamma} \quad (1-5)$$

반면 Jia (2018)은 배경에서 문서로 바뀌는 영역의 픽셀의 경사도 값이 상대적으로 크다는 것과 해당 영역에서 경사도 방향이 바뀌는 특징을 활용하여 문서 가장자리를 검출하였다. 마지막으로 이미지 내 각 픽셀의 경사도 값과 방향을 대상으로 지역 임계값을 선정하여 이진화를 수행하였다.

일반적으로 규칙 기반 방법론은 과거 문헌 이미지에서 문서의 가장자리에 해당하는 픽셀을 검출한 후, 해당 픽셀을 기준으로 가장자리 안쪽의 문서 픽셀을 추가적으로 탐지하고 이진화를 수행하였다. 하지만 과거 문헌 이미지에 따라 노이즈 정도와 종류가 다르기 때문에 이러한 규칙 기반 이진화는 일반화 성능이 떨어지며 이미지의 노이즈 정도에 따라 성능 편차가 심하다는 한계가 있다. 이에 따라 규칙 기반 이진화의 한계를 극복하고자 최근에는 데이터로부터 추출한 특징을 학습하여 이진화를 수행하는 학습 기반 방법론도 활발히 연구되고 있다.

1.2.2 학습 기반 이진화 방법론

배경과 문자의 픽셀값 차이라는 단일 특징을 기반으로 두 클래스의 경계를 탐지해 이진화를 수행했던 규칙 기반 방법론과 달리, 학습 기반 이진화 방법론은 다양한 특징을 추출해 배경과 객체 간 차이를 학습하는 방향으로 이진화를 수행한다. 학습 기반 방법론에서 머신러닝과 딥러닝의 가장 큰 차이점은 특징 추출 방법이다. 머신러닝은 연구자가 직접 특징들을 추출하여 모델에 입력하고, 딥러닝은 모델이 스스로 데이터에서 유의미한 특징을 추출하여 학습한다. 이미지 해석, 문장 분류, 얼굴 인식 등의 분야에 주로 사용되는 CNN은 합성곱층(convolution layer)과 풀링층(pooling layer)을 복합적으로 구성하여 특징을 추출하고 패턴을 파악하는 방식으로 학습이 진행된다. 특히 합성곱층은 각 픽셀의 인접 픽셀들을 조사해 특징을 파악하고 이를 하나의 특징맵(feature map)으로 도출하는 연산을 수행한다.

Chen (2016)은 규칙을 기반으로 적절한 임계값을 찾는 것 대신에 Sauvola 임계 알고리즘을 적용하여 다수의 이진화 이미지를 생성하고, 생성된 이진화 이미지로부터 픽셀값, 수직 수평 경사도 변화량, 라플라시안 값을 추출하여 선형 서포트 벡터 머신(SVM; Support Vector Machine)을 학습하였다. 여기서 서포트 벡터 머신은 입력된 4개 특징을 사용하여 노이즈에 대해 학습하고, 입력된 픽셀이 배경에 속할지 문자에 속할지 분류하여 이진화를 수행하였다.

Xiong (2018)은 이미지를 일괄 사이즈의 블록으로 분할하여 블록 내 대조값을 기준으로 세 개의 하위 목록으로 정리하였다. 첫 번째는 배경 픽셀만 포함하는 블록, 두 번째는 배경 픽셀과 문자 픽셀의 비가 1:1 인 블록, 마지막으로 문자 픽셀보다 배경 픽셀을 더 많이 포함하는 블록이다. 이때 첫 번째 블록은 배경 픽셀만 포함하고 있기에 임계처리가 필요 없다. 두 번째, 세 번째 블록에 속한 픽셀로부터 10가지 특징을 추출하였다. 이후 블록 별 특징을 서포트 벡터 머신에

입력하여 각 블록 별 최적 임계값을 찾아서 이진화 이미지를 생성하였다.

Bezmaternykh (2019)은 기존 의학 영상 처리 분야에 사용되었던 딥러닝 모델인 U-Net을 사용하여 이진화를 수행하였다. 과거 문헌 이미지를 128×128 사이즈의 작은 패치로 분할하여 모델을 학습하고, 정확도가 낮은 패치를 대상으로 이동(shifting), 가우시안 블러, 점 노이즈 추가 등의 데이터 증강(augmentation)을 통해 재학습(re-train)하는 방법을 통해 견고한 모델을 구축하였다.

Meng (2017)은 분해 단계와 재건축 단계로 이루어진 피라미드 계층 구조의 딥러닝 모델을 제안하였다. 모델은 분해 단계에서 배경과 문서 픽셀을 구분하는 특징을 다양한 해상도(scale)로 압축한 뒤, 압축된 특징맵으로부터 이진화 이미지를 재건축하였다.

1.2.3 시사점 및 결론

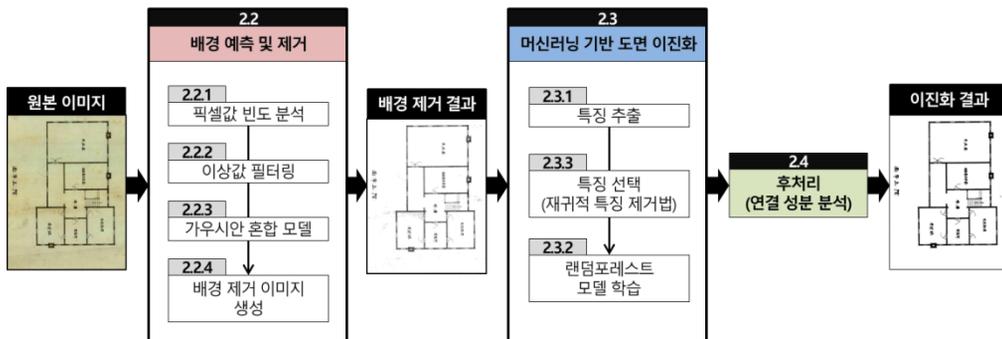
전역 임계값 이진화나 지역 임계값 이진화와 같은 전통적인 이진화 방법론들은 픽셀값 자체에 의존해 이진화를 진행하였다. 전역 임계값 이진화는 이미지 전체에 대하여 한 번에 이진화 연산을 수행했기 때문에 사이즈가 크거나 픽셀값이 다양한 이미지에서 성능이 좋지 못하였고, 지역 임계값은 국소 영역으로 분할하여 이진화를 수행하였지만 윈도우 사이즈에 지나치게 의존적이라는 한계점이 존재했다. 이에 Lu (2010), Su (2010), Su (2013)은 문서 데이터 셋의 특성을 바탕으로 문자의 두께를 파라미터로 활용하여 정밀한 임계 처리를 수행하였다. 1.2.2절에서 소개한 선행 연구들은 다음과 같이 요약할 수 있다. 첫째, 기존 이진화 연구들은 픽셀값의 통계적 특징을 바탕으로 이진화를 수행하였다. 둘째, 손 글씨 문서 데이터 셋을 대상으로 이진화를 수행한 연구들은 픽셀값의 통계적 특징 추출 시 결정해야 하는 윈도우 크기를 최적화하기 위해 문자 픽셀의 두께 정보를 활용하였다.

하지만 기존 선행 연구의 방법론을 건축도면에 적용할 때는 다음과 같은 한계점이 존재한다. 첫째, 문자 두께나 글자 경계와 같이 문서 데이터 셋을 대상으로 추출한 특징은 다양한 두께와 길이의 선으로 구성된 건축도면에 적용할 수 없다. 예를 들어 문서 데이터 셋은 각각의 문자 픽셀 두께가 일정하지만, 건축도면에서는 두꺼운 벽과 얇은 벽이 존재하기 때문에 각 객체마다 픽셀 두께가 상이하다. 둘째, 건축도면 내 실내 객체 요소의 연결성을 고려한 특징 추출이 불가능하다. 도면에서 각각의 객체는 서로 연결되어 있으며 이는 문서 내 각각의 문자의 픽셀 두께가 일정한 것과 같이 건축도면만의 고유한 특성이다. 일반적으로 문서 데이터 셋에서는 각각의 문자가 서로 연결되지 않고 독립적인 객체로 존재하지만, 건축도면의 경우 벽, 문, 창문 등의 실내 요소가 서로 연결되어 하나의 실내 공간을 형성한다. 따라서 객체 간 연결성은

두께가 얇은 선형객체나 배경과 유사한 픽셀값을 가진 선형객체와 배경을 구분할 수 있는 특징이다. 결론적으로 본 연구는 필터링이나 특징 추출을 위한 윈도우 크기를 최적화 하기 위해 객체의 픽셀 두께를 사용한 선행 연구와 달리 윈도우 크기를 3×3 으로 고정하여 파라미터 튜닝을 최소화하였다. 또한 기존 선행 연구에서 이진화를 위해 사용한 픽셀값의 통계적 특징과 더불어 도면 내 기하 요소의 수직 및 수평 연결성 정보를 활용하여 건축도면 데이터 셋의 특성을 고려한 이진화 알고리즘을 제안하였다.

1.3 연구 범위 및 방법

본 연구에서는 기존 도면 해석 연구에 사용된 전자 도면 대비 다양한 형태의 노이즈가 존재하는 건축도면을 대상으로 이진화 연구를 수행한다. 이를 위해 국가기록원에서 디지털 아카이브를 통해 제공하는 일제시기 건축도면을 사용하였다. 본 연구에서는 머신러닝 기반 분류 모델인 단일의사결정나무, 랜덤포레스트, 완전연결신경망을 사용하여 이진화를 수행하여 분류 성능이 가장 우수한 랜덤포레스트를 최종 모델로 선정하였다. 우수한 분류 성능과 더불어 랜덤포레스트는 학습 과정에 특징 중요도를 추출한다는 장점이 있다. 특징 중요도는 이후 실험 결과 분석 시 분류에 결정적인 역할을 하는 주요 변수 판단에 활용하였다. 본 연구에서 제안하는 이진화 기법은 제거하고자 하는 노이즈의 형태에 따라 크게 두 가지 단계로 진행된다. 첫 번째 단계는 가우시안 혼합 모델을 사용하여 도면 이미지의 배경에 전체적으로 넓게 분포하는 노이즈를 감소하는 단계이다. 두 번째 단계는 랜덤포레스트 모델을 기반으로 선형객체와 배경을 구분하는 특징을 추출하여 면적이 작고 다양한 형태의 노이즈를 학습하여 제거하는 단계이다([그림 1-3]).



[그림 1-3] 연구 흐름도

첫 번째 단계인 배경 예측 및 제거 단계에서는 RGB 값 3 개 채널로 이루어진 입력 도면 이미지를 회색조로 변환하여 1 개 채널로 축소하고, 픽셀값 분포가 쌍봉 분포(bimodal distribution)에 근사하도록 도면 이미지 전체에 3×3 양방향 필터(bilateral filter)를 적용한다. 이렇게 쌍봉분포에 근사한 픽셀값 밀도 함수를 가우시안 혼합 모델(GMM; Gaussian Mixture Model)에 입력하여 배경과 선형객체 군집으로 분류한 뒤, 선형객체 군집에 해당하는 픽셀값을 배경 픽셀값으로 대체 해 배경 이미지를 생성한다. 이렇게 생성된 배경 이미지는 기존 입력 도면 이미지와 배경에 해당하는 픽셀값은 유사하고, 도면에 해당하는 픽셀값은 차이가 생기게 된다. 따라서, 처음 입력한 도면 이미지에서 추정된 배경 이미지를 빼주어 배경 제거 이미지를 생성하면, 해당 이미지에서 배경에 해당하는 영역의 픽셀값은 0에 가깝고, 도면에 해당하는 영역은 유의미한 차이가 나는 픽셀값을 갖게 된다.

두 번째 단계인 학습 기반 이진화 알고리즘에서는 배경 예측 및 제거 단계에서 생성한 배경 제거 이미지 이미지를 입력자료로 활용한다. 그리고 특징 추출 단계에서는 이미지의 각 픽셀을 슬라이딩 윈도우 방식으로 탐색하여 총 24 개 특징들을 추출한다. 여기서 추출하는 특징은 크게 윈도우 내 픽셀들의 통계값, 명암도 동시발생 행렬(GLCM; Gray-Level Co-occurrence Matrix) 통계값, 중앙 및 이웃 픽셀값, 수직-수평 연속성 행렬 항목으로 이루어져 있다. 이러한 특징 값들을 바탕으로 노이즈 감소에 핵심적인 특징을 선택하기 위해 랜덤포레스트 모델을 통해 산출된 각 특징의 중요도로부터 최적의 변수를 선택하는 Guyon *et al.* (2002)이 제안한 재귀적 특징 제거법(RFE; Recursive Feature Elimination)을 활용하였다.

최종적으로는 최적 특징 조합으로 학습한 모델에 테스트 셋을 입력하여 모델의 성능을 검증하고, 연결 성분 분석(CCA; Connected-

component Analysis) 후처리를 통해 픽셀 단위의 작은 노이즈 픽셀을 제거하여 이진화 이미지를 생성하였다.

2. 연구 방법

2.1 데이터 수집 및 전처리

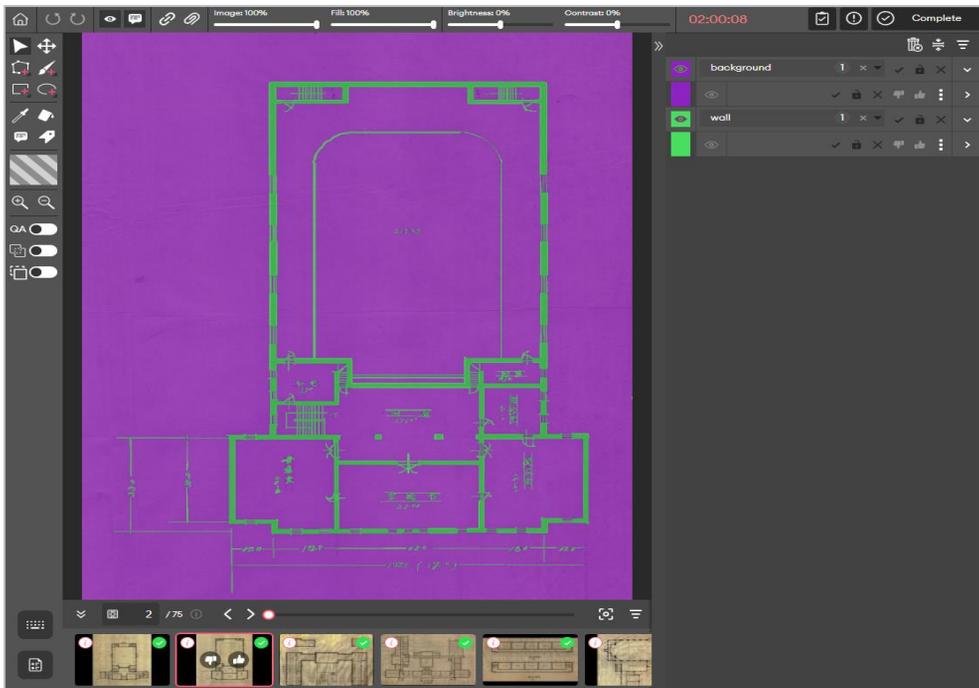
본 연구는 근대 건축도면 데이터를 수집하기 위해 국가기록원에서 제공하는 일제시기 건축도면 디지털 아카이브¹를 활용하였다([그림 2-1]). 해당 디지털 아카이브는 일제시기 조선총독부에서 진행된 공공건축 관련 자료들을 모아놓은 서비스로 2000년대 초반부터 2008년까지 정리된 도면을 온라인으로 제공한다. 본 연구에서는 평면도를 대상으로 총 50개의 도면을 수집하고 정답 데이터 셋을 구축하였다. 라벨링 도구로는 픽셀 단위 라벨링이 가능한 SuperAnnotate²를 사용하였다([그림 2-2]).

¹ <https://theme.archives.go.kr/next/plan/viewMain.do>

² <https://www.superannotate.com/>



[그림 2-1] 일제시기 건축도면 아카이브 홈페이지 화면



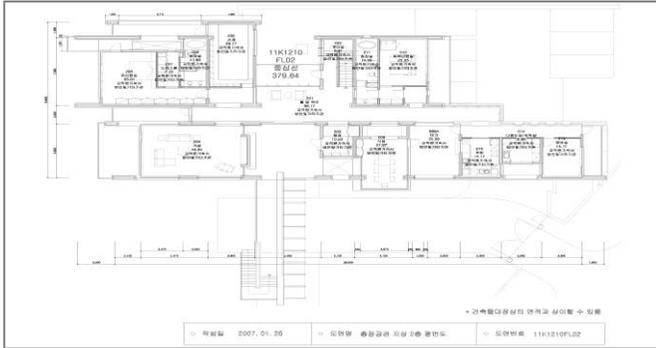
[그림 2-2] SupperAnnotate 라벨링 화면 예시

2.2 배경 예측 및 제거

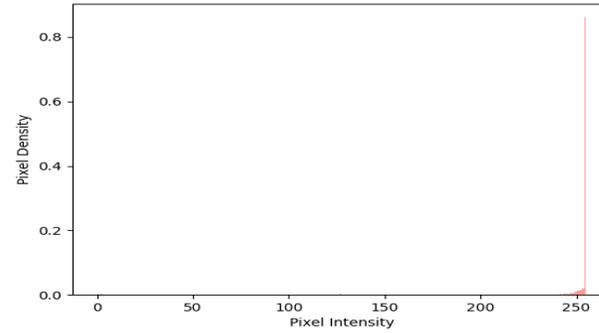
2.2.1 픽셀값 빈도 분석

일반적으로 이진화는 이미지의 픽셀값 빈도 히스토그램이 날카로운 두 개의 봉우리가 존재할 때 이상적인 결과를 도출한다. [그림 1-1]의 도면 데이터 셋처럼 배경과 도면의 픽셀값 차이가 명확한 도면의 경우 [그림 2-3]의 (b)과 같은 픽셀값 빈도를 보인다. [그림 2-3]의 (b)에서 X축은 픽셀값을 나타내며, Y는 각 픽셀값에 해당하는 픽셀의 빈도를 나타낸다. 선으로 실내 구조와 벽, 문, 창문 등을 표현하는 도면의 특성 상, 전체 면적 대비 도면에 해당하는 픽셀은 상대적으로 매우 작다. 그렇기에 픽셀값 빈도에서 0에 가까운 픽셀값을 가지는 픽셀의 개수가 255에 가까운 배경 픽셀의 개수 대비 매우 작다. 각 영역에 해당하는 픽셀의 비율은 큰 차이가 있지만, [그림 1-1]의 데이터 셋과 같이 도면과 배경의 픽셀값 차가 뚜렷한 도면의 픽셀값 빈도 히스토그램은 매우 날카로운 두 개의 봉우리를 보인다.

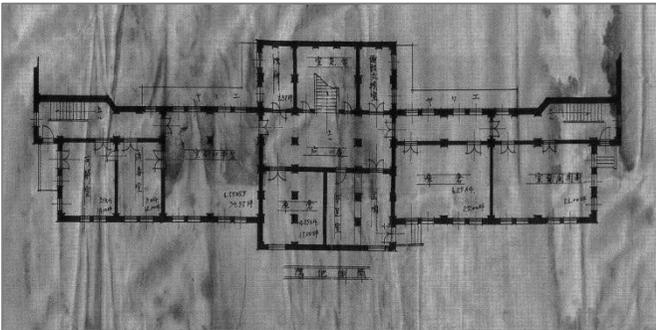
반면, 종이 비침, 불균일한 배경, 배경 잉크 얼룩, 불균일한 조명으로 인해 과거 문헌이나 도면 같은 문서는 다양한 픽셀값을 가지기 때문에 픽셀값 빈도 결과에서 배경 픽셀에 해당하는 군집을 특정하기에 어려움이 있다. 예를 들어, [그림 2-3]의 (c)와 같은 도면은 (d) 같은 픽셀값 빈도 히스토그램을 갖는데, (b)의 픽셀값 히스토그램 대비 배경과 도면의 군집 경계가 뚜렷하지 않다. 그래서 이러한 히스토그램은 쌍봉형태의 히스토그램으로 변환하는 과정과 이를 통해 배경과 도면을 분리하는 단계를 거쳐야 효과적인 이진화 작업을 수행할 수 있다 (Jana *et al.*, 2017)



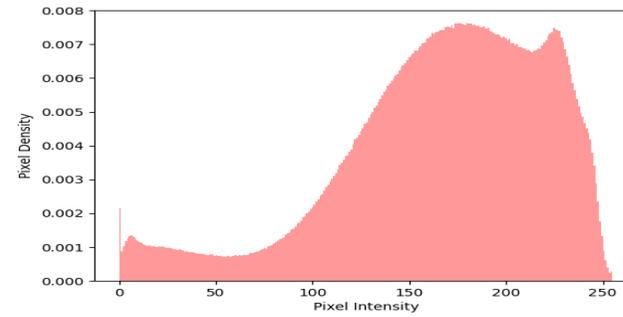
(a) 서울대학교 도면 데이터 셋



(b) 서울대학교 도면의 픽셀값 빈도 히스토그램



(c) 핸드드로잉 건축도면 데이터 셋



(d) 핸드드로잉 건축도면의 픽셀값 빈도 히스토그램

[그림 2-3] 도면 데이터 셋에 따른 픽셀값 히스토그램 예시

2.2.2 이상값 필터링

2.2.1절에서 설명한 바와 같이 배경과 선형 픽셀 군집을 분리하기 위해선 픽셀값 빈도 히스토그램이 쌍봉 형태여야 한다. 하지만 [그림 2-3]의 (d)를 보면 쌍봉 형태가 나타나지 않으며, 히스토그램 끝단에 이상치가 존재한다. 픽셀값 관점에서 0이나 255의 픽셀값은 이상치가 아니지만, 배경과 문자 군집이 섞여있다는 가정하에 두 개의 정규 분포를 따르는 픽셀 군집으로 분류하고자 하는 관점에서 이러한 양 끝단의 픽셀값 쏠림 현상은 이상치로 분류할 수 있다. 따라서, 선행 연구들에서는 이러한 픽셀값을 보정하기 위해 다양한 필터를 사용하였다([표 2-1]).

[표 2-1] 선행 연구에서 사용된 필터 종류 예시

연구자	사용 필터
Mitianoudis, N., & Papamarkos, N.(2015)	Median filter
Gatos <i>et al.</i> (2006)	Weiner filter
Lu <i>et al.</i> (2010)	Savitzky-Golay filter
Nina <i>et al.</i> (2011)	Bilateral filter
Xiong <i>et al.</i> (2018)	Morphological closing filter

이러한 필터 중 본 연구에서는 선명도를 증가시키고 노이즈를 감소시키는 비선형 필터인 양방향 필터(bilateral filter)를 사용하였다(Tomasi *et al.*, 1998). 양방향 필터는 가우시안 필터의 단점인 적용 시 경계 부분이 흐릿해지는 점을 극복하기 위해 범위 내 존재하는 이웃 픽셀들에 대하여 픽셀 단위로 유사 정도에 따라 가중치를 할당하고, 가중치 평균을 계산하여 픽셀값을 보정한다. 따라서, 가우시안 필터는 픽셀 간 공간적 거리차를 가중치로 사용하는 반면, 양방향 필터는 픽셀값 차이를 가중치로 사용하는 영역 필터($R_{i,j}$)와 가우시안 필터의 공간 필터

$(S_{i,j})$ 를 합성하여 필터링된다.

$$B'_i = \frac{1}{N_i} \sum_{j \in \Omega} R_{i,j} \cdot S_{i,j} \cdot B_j \quad (2-1)$$

양방향 필터의 형태는 식 (2-1)와 같으며, 여기서 i 는 필터링할 중앙 픽셀의 좌표를, j 는 필터링 범위로 설정된 윈도우 Ω 에 대응되는 픽셀값을, B_j 는 좌표 j 에 대한 픽셀값을, B'_i 는 좌표 i 에 대한 필터링 픽셀값을 의미한다. 또한, 양방향 필터의 $S_{i,j}$ 는 식 (2-2)에 나타난 바와 같이 유클리드 거리 계산법을 사용하여 좌표 i 와 좌표 j 의 공간 유사도를 계산한다.

$$S_{i,j} = \exp\left(\left(-\frac{\|i-j\|}{\sigma_s}\right)^2\right) \quad (2-2)$$

이때 σ_s 는 이웃 픽셀로 고려하는 거리를 조절하는 파라미터로, 해당 파라미터의 값이 커질수록 필터링이 적용되는 범위가 넓어지기 때문에 결과적으로는 더욱 흐릿한 영상을 생성한다.

$R_{i,j}$ 는 식 (2-3)와 같이 유클리드 거리 계산법을 바탕으로 i 와 j 의 픽셀값 유사도를 계산한다.

$$R_{i,j} = \exp\left(\left(-\frac{\|B_i - B_j\|}{\sigma_r}\right)^2\right) \quad (2-3)$$

이때 σ_r 은 픽셀값 범위를 정하는 파라미터로, 픽셀값 차가 설정한 파라미터를 초과하게 되면 다른 색으로 판단해 해당 영역에 객체의 가장자리가 있다고 가정하여 필터링 결과에 큰 영향을 미치지 못한다.

마지막으로 $N_{i,j}$ 는 필터링 결과를 정규화해주는 정규화 상수이며, 식 (2-4)와 같이 나타낼 수 있다.

$$N_{i,j} = \sum_{j \in \Omega} R_{i,j} \cdot S_{i,j} \quad (2-4)$$

위 식들을 바탕으로 양방향 필터에서는 중앙 픽셀값과 유사한 픽셀값을 가진 픽셀이 더 큰 가중치를 부여 받게 되고, 이 가중치가 곱해진 평균값이 계산된다. 따라서 밝은 픽셀들은 밝은 픽셀끼리, 어두운 픽셀들은 어두운 픽셀끼리 필터링 되어 경계 부분이 흐릿해지지 않는다. 즉, 양방향 필터는 중앙 픽셀과 밝기가 유사한 이웃 픽셀들에 대해서만 가우시안 필터를 적용하기 때문에 가우시안 필터 적용 시 경계가 흐릿해지는 한계점을 극복한 필터이다. 본 연구에서는 이미지 내 모든 픽셀에 대하여 3×3 슬라이딩 윈도우 방식으로 탐색하며 필터링을 진행하였다. 본 연구의 데이터 셋을 대상으로 파라미터 최적화를 위한 실험을 진행한 결과, 공간필터의 파라미터가 필터링 결과에 더욱 큰 영향을 주는 것으로 나타났다. 공간필터 파라미터가 커질수록 선형객체가 흐릿해지는 필터링 결과를 보였다. 따라서 필터링 시 고려하는 이웃의 범위를 최소 범위로 설정하여 공간필터의 파라미터인 σ_s 는 3으로 설정하였다. 영역필터의 파라미터인 σ_r 는 양방향 필터를 사용한 선행 연구인 Nina *et al.* (2011)을 참고하여 2로 설정하였다.

2.2.3 가우시안 혼합 모델

본 연구에서는 이상치를 제거한 픽셀값 밀도 함수를 각각 배경과 도면 군집으로 분류하기 위해 가우시안 혼합 모델을 적용하였다. 가우시안 혼합 모델은 확률 기반 군집화 알고리즘으로, 데이터가 여러 다른 모양의 가우시안 분포(Gaussian distribution)로 결합되어 있다는 가정 하에 개별 데이터를 동일한 가우시안 분포별로 묶어주는 비지도 학습 알고리즘이다(Bouveyron and Brunet-Saumard, 2014). 이러한 확률론적 군집화 방법의 핵심은 두 군집이 겹치는 구간을 고려할 수 있다는 것인데, 이러한 특징이 노이즈가 강한 데이터에 좋은 성능을 나타내는 것으로 연구된 바 있다(Kwon *et al.*, 2022). 따라서 본 연구에 사용된 핸드드로잉 건축도면과 같이 노이즈가 강한 데이터 셋에는 K-평균 군집화와 같은 계층적 군집 방법보다 가우시안 혼합 모델이 적합하다고 판단하였다.

가우시안 혼합 모델에서 이미지 내 각 픽셀의 가우시안 혼합 확률 밀도 함수(P)는 식 (2-5)와 같이 표현할 수 있다.

$$P(x_i) = \sum_{k=1}^K \pi_k N_k(x_i | \mu_k, \Sigma_k) \quad (2-5)$$

여기서 K 는 군집의 개수이며, π_k 는 0에서 1 사이의 값을 가지며 k 번째 가우시안 분포가 선택될 확률을 의미한다. 이때 $\sum_1^K \pi_k = 1$ 이다. N_k 는 픽셀 x_i 에 대하여 k 번째 평균(μ_k)과 k 번째 covariance matrix(Σ_k)를 갖는 k 번째 군집을 나타낸다.

k 번째 군집의 단일 다변량 가우시안 분포는 다음과 같이 정의된다:

$$N_k(x_i|\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x_i - \mu)^T \Sigma^{-1}(x_i - \mu)\right) \quad (2-6)$$

따라서 가우시안 분포를 추정하기 위한 모수는 $\theta = \{\pi, \mu, \Sigma\}$; $\pi \equiv \{\pi_1, \dots, \pi_k\}, \mu \equiv \{\mu_1, \dots, \mu_k\}, \Sigma \equiv \{\Sigma_1, \dots, \Sigma_k\}$ 로 정의할 수 있으며 각 모수의 로그 우도는 식 (2-7)과 같다.

$$L(\theta|x_i) = \sum_{i=1}^n \ln \left\{ \sum_{k=1}^K \pi_k N_k(x_i|\mu_k, \Sigma_k) \right\} \quad (2-7)$$

이러한 로그 우도를 바탕으로 가우시안 혼합모델의 각 모수는 기댓값-최대화(EM; Expectation-Maximization) 알고리즘을 사용하여 추정된다. 기댓값-최대화 알고리즘은 우선 모수를 초기화 하고, 초기화 된 모수에 대하여 로그 우도를 정의한다. 그 후 기댓값(expectation) 단계에서 각 데이터가 현재 모수로 이루어진 군집들 중 K 번째 군집으로부터 발생할 사후확률을 평가한다(식 2-8).

$$\gamma_k(x_i) = \frac{\pi_k N(x_i|\mu_k, \Sigma_k)}{\sum_{k=1}^K \pi_k N(x_i|\mu_k, \Sigma_k)} \quad (2-8)$$

최대화(maximization) 단계에서는 기댓값 단계에서 계산한 사후확률을 최대화하는 모수를 재추정한다. 최대화 단계에서 계산한 변수값은 다음

기댓값 단계에서 추정값으로 사용되며, 모수의 값이 수렴할 때까지 이와 같은 단계를 여러 번 반복한다. 이처럼 가우시안 혼합 모델을 사용해 픽셀값을 두 개의 군집으로 분리하고, 이후 이어지는 배경 밀도 함수 추출 단계에서는 선형객체 군집으로 분류된 픽셀들을 대상으로 픽셀값을 배경 군집과 유사하게 변환한다.

2.2.4 배경 제거 이미지 생성

본 연구에서 배경 예측이란, 2.2.3절에서 가우시안 혼합 모델을 사용해 선형 픽셀과 배경 픽셀 군집을 분리한 후, 선형 픽셀 군집에 해당하는 픽셀값을 배경 픽셀값으로 변경하는 과정을 의미한다. 이러한 과정을 통해 생성되는 배경 예측 이미지는 원본 이미지의 픽셀값과 각각 1:1로 비교하였을 때, 다음과 같은 차이를 보인다.

우선, 배경에 해당하는 영역은 배경 예측 이미지와 원본 이미지 모두 유사한 픽셀값을 갖는다. 반면 선형객체에 해당하는 영역은 배경 예측 이미지와 원본 이미지의 픽셀을 1:1로 대응하여 차이를 구하였을 때 큰 픽셀값 차이를 보인다. 즉, 두 이미지의 픽셀값 차이로 선형객체와 배경 영역을 구분할 수 있다는 것이다. 많은 선행 연구들에서 배경 영역을 예측하고, 예측된 배경 이미지를 원본 이미지에서 빼줌으로써 배경 노이즈를 효과적으로 제거할 수 있다고 언급한 바 있다(Jindal *et al.*, 2021). 본 연구에서도 이러한 픽셀값 차이를 바탕으로 배경 제거 이미지를 최종적으로 생성하였다.

2.3 머신러닝 기반 도면 이진화

본 절에서는 2.2절의 결과로 배경이 제거된 도면 이미지로부터 추출하는 특징과 이를 바탕으로 선형 픽셀의 특징 차이를 학습하기 위한 머신러닝 기반 알고리즘을 다룬다. 2.3.1절에서는 본 연구에서 배경과 선형객체를 구분하기 위해 추출하는 특징의 종류에 대해 자세히 설명한다. 2.3.2과 2.3.3절에서는 앞서 추출한 특징들을 바탕으로 랜덤포레스트 모델을 학습하는 과정과 학습 시 산출되는 특징 중요도에 대해 소개한다. 마지막 2.3.4절에서는 이후 성능 평가 시 사용되는 평가 지표에 대해 다룬다.

2.3.1 특징 추출

본 연구에서는 이미지를 3×3 크기의 윈도우로 탐색하며 총 24개의 특징을 추출하였는데, 이는 각각 윈도우 내 픽셀들의 통계값 특징 5개, 명암도 동시 발생 행렬 특징 9개, 수직-수평 연속성 행렬 특징 1개, 그리고 중앙 픽셀을 포함한 이웃 픽셀값 정보 9개로 구성된다([표 2-2]). 세부 절에서는 각 특징 항목에 대해 자세히 기술하였다.

통계값 특징, 명암도 동시 발생 행렬 특징, 픽셀값 정보들은 관심 영역 내 픽셀값과 직접적으로 관련된 영상 특징들이다. 특히 통계값이나 명암도 동시 발생 행렬들과 같은 특징들은 이미지에서 픽셀값에 따라 질감(texture)의 변화를 표현한다. 평균, 표준편차, 최대값, 최소값, 지역 대비값과 같은 통계적 특성은 단순히 이미지의 픽셀값 빈도 히스토그램을 사용한 1차 통계량(first order statistic)이며 명암도 동시 발생 행렬 특징은 거리와 방향 변화에 따른 픽셀값 빈도 히스토그램을 사용한 2차 통계량(second order statistic)이다.

앞서 언급한 영상 특징만 사용하였을 시, 두꺼운 벽의 경우 영상

내 특징이 매우 뚜렷하여 성공적으로 이진화를 수행할 수 있으나 상대적으로 얇게 표현된 가벽이나 문의 경우 두꺼운 벽 대비 특징이 뚜렷하지 못해 배경의 노이즈로 인식될 가능성이 존재한다. 이에 따라 본 연구는 이러한 영상 특징 이외에도 건축도면의 실내 요소들의 기하학적 특성을 고려한 수직-수평 연속성 행렬 특징을 추가로 추출하였다. 건축도면에 서는 실내 요소 객체가 개별적으로 존재하는 것이 아니라 선형객체를 구성하는 픽셀들이 연속적으로 상호 연결되어 있다. 본 연구는 이처럼 선형객체를 구성하는 픽셀이 연속적으로 상호 연결되어 있는 특징을 활용하고자 하였다.

[표 2-2] 이진화를 위해 사용된 특징 분류

대분류	특징명	연구자
통계적 특성	평균	Ahmadi <i>et al.</i> (2015), Xiong <i>et al.</i> (2018), Jana <i>et al.</i> (2017)
	표준편차	Ahmadi <i>et al.</i> (2015), Xiong <i>et al.</i> (2018), Jana <i>et al.</i> (2017)
	최대값	Ahmadi <i>et al.</i> (2015)
	최소값	Ahmadi <i>et al.</i> (2015)
	지역대비값	Su <i>et al.</i> (2012)
명암도 동시행렬의 통계적 특성	평균	Sehad <i>et al.</i> (2013), Bharathi <i>et al.</i> (2013)
	표준편차	Bharathi <i>et al.</i> (2013)
	Contrast (대조)	Westphal <i>et al.</i> (2018), Sehad <i>et al.</i> (2013), Bharathi <i>et al.</i> (2013)
	Dissimilarity (이질성)	Jain (2013), Bharathi <i>et al.</i> (2013)
	Homogeneity (유사성)	Westphal <i>et al.</i> (2018)
	ASM (균등성)	Jain (2013)
	Energy (에너지)	Westphal <i>et al.</i> (2018), Sehad <i>et al.</i> (2013)
	Peak (최고치)	Jain (2013)
	Entropy (명암도 변화)	Jain (2013), Bharathi <i>et al.</i> (2013)
수직-수평 연속성 행렬	수직-수평 연속성 행렬	Vargas (2018)
중앙 및 이웃 픽셀값	P1~P8, 중앙픽셀	Kasmin <i>et al.</i> (2019), Jana <i>et al.</i> (2017)

2.3.1.1 통계적 특성

본 연구에서 사용한 첫 번째 특징 군은 슬라이딩 윈도우 내 픽셀들의 통계적 특성이다. 3×3 크기의 윈도우로 각 픽셀을 탐색하면 중앙 픽셀을 포함해 9개의 이웃 픽셀의 픽셀값 정보를 얻을 수 있다. 이때 총 9개 픽셀값의 평균, 표준편차, 최댓값, 최소값, 지역 대비값을 추출하여 통계적 특성으로 사용하였다. 지역 대비값(LC)의 경우, 식 (2-9)를 사용하여 산출하였다.

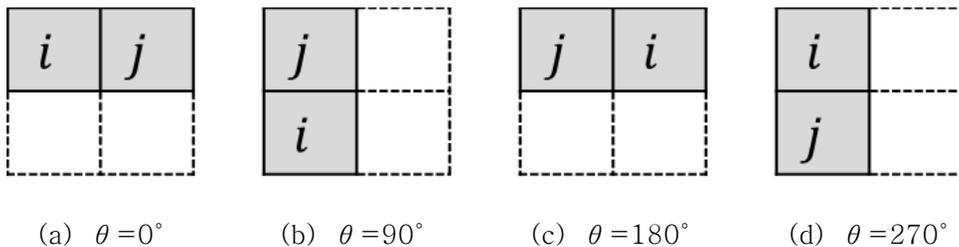
$$LC = \frac{\max - \min}{\max + \min} \quad (2-9)$$

이러한 통계적 특성들은 중앙 픽셀의 위치를 추정할 수 있게 한다. 예를 들어, 9개 픽셀의 평균이 255에 가까운 큰 값이라면, 중앙 픽셀이 매우 밝은 픽셀값을 가지는 배경 영역일 가능성이 높아진다. 이때 표준편차는 9개 픽셀이 모두 255에 가깝기 때문에 작은 값을 가지며, 지역 대비값도 0에 가까운 매우 작은 값을 가진다. 반면, 픽셀값 하나라도 0에 나머지 픽셀값과 크게 차이 나는 값을 가진다면, 표준편차와 지역 대비값이 크게 변하게 된다.

2.3.1.2 명암도 동시행렬의 통계적 특성

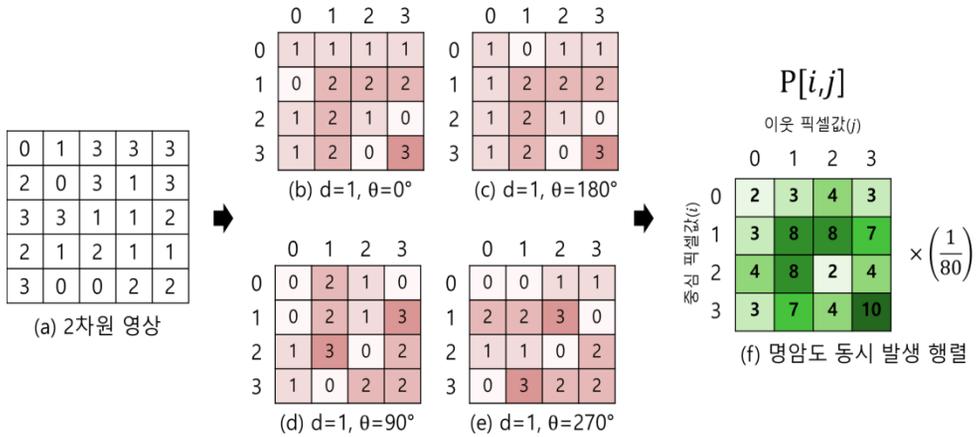
명암도 동시발생 행렬(GLCM; Grey Level Co-occurrence Matrix)은 영상 처리분야에서 텍스처 분석을 위해 도입된 개념으로, 2차원 영상에서 특정 거리를 두고 인접한 픽셀간에 얼마만큼의 픽셀값 차이가 나는지를 통계적으로 표현한 행렬이다(Haralick *et al.*, 1973). 픽셀값 자체를 참조하는 평균이나 표준 편차와 같은 통계적 특징과 달리, 명암도 동시발생 행렬은 두 픽셀 값 사이의 관계를 이용한 2차 통계 척도 기반 분석 방법이며, 이를 이용해 다양한 2차 통계량을 구할 수 있다. 명암도 동시발생 행렬은 픽셀과 이웃 픽셀의 관계를 나타내는 방법으로 배경과 선형객체의 질감 특성을 통계적으로 분석하는 데에 유용하다.

명암도 동시발생 행렬을 생성할 때 사용하는 두 가지 파라미터는 이웃의 범위를 결정하는 방향(θ)과 거리(d)이다. 이미지와 같은 2차원 영상에서 명암도 동시발생 행렬을 만드는 경우 일반적으로 상하좌우의 4방향이나, 대각 방향을 추가한 8방향 모델을 사용한다. [그림 2-4]는 이러한 4방향의 변위 벡터를 나타낸 것이다. 여기서 거리 d 는 1이다.



[그림 2-4] 4방향 변위 벡터 예시

[그림 2-5]는 0에서 3사이의 그레이 레벨을 가지는 5×5 크기의 2차원 영상(a)로부터 변위 벡터를 이용하여 명암도 동시발생 행렬을 구성하는 예시를 보인 것이다.



[그림 2-5] 명암도 동시발생 행렬 구성 예시

[그림 2-5]처럼 명암도 동시발생 행렬의 행과 열의 크기는 픽셀값의 단계 수이므로 4×4 크기의 행렬을 얻을 수 있다. 질감 분석에서 방향을 고려하지 않는 경우 일반적으로 이 행렬들을 합산한 하나의 행렬을 명암도 동시발생 행렬로 정의한다. 여기서 명암도 동시발생 행렬의 (0, 1) 값이 3이라는 것은 픽셀값이 0인 픽셀을 기준으로 그 이웃에 픽셀값 1이 오는 경우가 총 3번 출현함을 의미한다. 이를 바탕으로 [그림 2-5]의 (f)와 같이 명암도 동시발생 행렬을 구한 뒤 빈도수의 총합으로 나누어 정규화한 후, 이를 사용해 다음과 같은 2차 통계량을 계산할 수 있다.

$$Mean = \frac{\sum_{i=0}^{N-1} i(P_{i,j}) + \sum_{j=0}^{N-1} j(P_{i,j})}{2} \quad (2-10)$$

$$Std = \frac{\sum_{j=0}^{N-1} P_{i,j}(i - \mu_i)^2 + \sum_{i=0}^{N-1} P_{i,j}(j - \mu_j)^2}{2} \quad (2-11)$$

$$Dissimilarity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} (i - j)^2 \quad (2-12)$$

$$Contrast = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} |i - j| \quad (2-13)$$

$$Homogeneity = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P_{i,j}}{1 + |i - j|} \quad (2-14)$$

$$Energy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j}^2 \quad (2-15)$$

$$ASM = \sqrt{Energy} \quad (2-16)$$

$$Peak = Max(P_{i,j}) \quad (2-17)$$

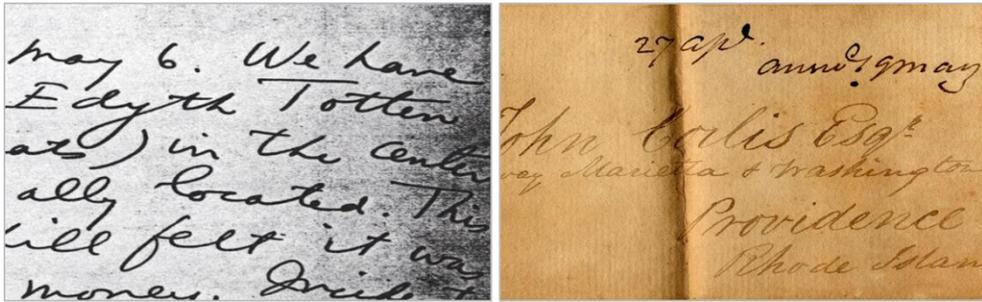
$$Entropy = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} (-\ln P_{i,j}) \quad (2-18)$$

여기서, $P_{i,j}$ 내 값은 원 영상에서 구해진 명암도 동시발생 행렬의 각 원소이다. 대조(contrast)는 명암도 차이를 측정하는 척도로, 명암도 동시발생 행렬의 대각선으로부터 멀리 떨어져 있는 $P_{i,j}$ 요소에 대하여 높은 가중치가 적용되며, 명암도가 크게 차이 나는 픽셀이 많을수록 더 큰 값을 가진다. 이질성(dissimilarity)은 대조와 유사하게 명암도 차이를 측정하기 위한 척도로 대조의 제곱값을 가지기 때문에 명암도 차이를 극대화한다. 반면, 유사성(homogeneity)은 $P_{i,j}$ 가 명암도 동시발생 행렬의 대각선에 모여있을 때 가장 큰 값을 가지게 된다. 그리고 균등성(ASM; Angular Second Moment)과 에너지(energy)는 명암도의 균일 정도를 측정하는 척도이며, 윈도우 내 픽셀들의 밝기 변화가 없을 때 큰 값을 갖는다. 즉, 모든 픽셀값이 유사할 때 최대값을 가지게 된다. 마치

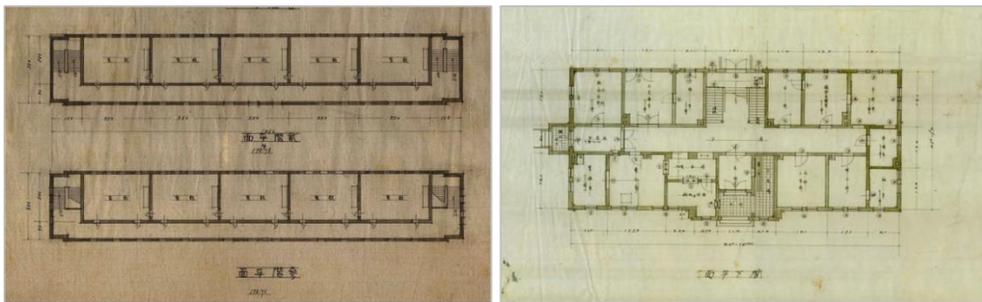
막으로 엔트로피(entropy)는 명암도 분포의 임의성(randomness)를 측정하는 척도로 윈도우 내 픽셀값의 변화가 많아서 $P_{i,j}$ 가 랜덤하게 분포하게 되면 큰 값을 갖는다.

2.3.1.3 수직-수평 연속성 행렬

명암도 동시 행렬 이외에도 본 연구는 도면의 특성을 고려하기 위해 Vargas (2018)이 제안한 Slice Transform 기반 수직-수평 연속성 행렬을 특징으로 활용하였다. 일반적으로 도면은 직선이 수직 혹은 수평으로 직교하여 공간을 분리하고, 여러 개의 선을 평행하게 배치하여 창문이나 계단실과 같은 실내 요소를 구성한다. 이와 다르게 [그림 2-6]과 같이 이진화 선행 연구에서 널리 사용된 DIBCO 데이터 셋은 손 글씨 데이터 셋이기 때문에 손 글씨 특성 상 곡선이 존재하며, 각각의 글자가 연결되는 부분 없이 독립적으로 존재한다. 반면 본 연구에서 사용한 도면은 다양한 길이의 직선이 수직-수평 방향으로 연결되어 있어 큰 차이점이 존재한다.



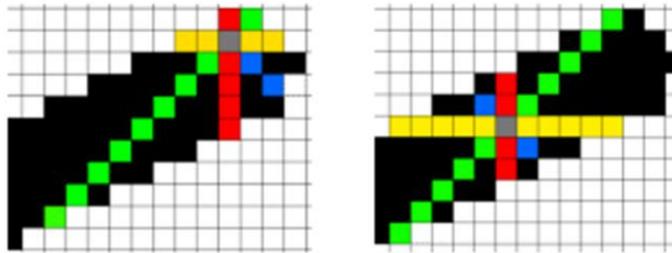
(a) DIBCO 데이터 셋



(b) 핸드드로잉 건축도면 데이터 셋

[그림 2-6] DIBCO 데이터 셋과 핸드드로잉 건축도면 데이터 셋 비교

Vargas (2018)이 제안한 Slice Transform 은 픽셀값이 255 가 아닌 모든 픽셀에 대하여 다음 그림과 같이 4 방향으로 연속성을 나타내는 특징 벡터를 추출하고 이를 합산한다. [그림 2-7]은 회색으로 칠해진 두 개의 픽셀에 대하여 각각 4 방향 연속성 특징 벡터를 추출하는 과정의 예시이며 각 방향은 다음과 같이 표현되었다: 빨간색(수직 방향 연속성), 노란색(수평 방향 연속성), 파란색(135 도 대각방향 연속성), 초록색(45 도 대각방향 연속성).



[그림 2-7] 연속성 특징 벡터 추출 (Vargas, 2018)

여기서 본 연구는 수직-수평 방향 연속성 행렬만 추출하여 사용하였으며, 가로 길이가 r , 세로 길이가 c 인 이미지 A 에 대한 수직-수평 방향 연속성 행렬 $M(A)$ 은 다음과 같이 표현될 수 있다.

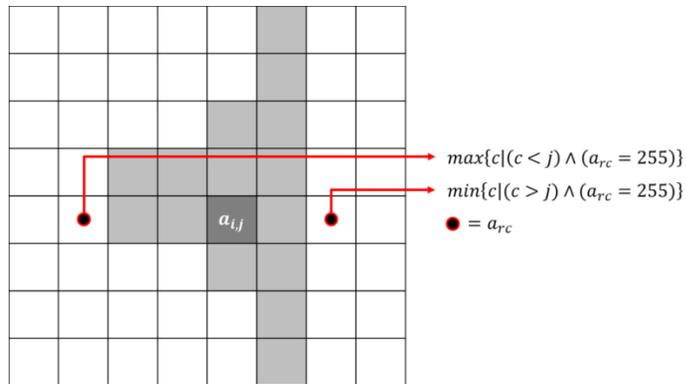
$$M(A) = \begin{cases} p_h + p_v, & \text{if } a_{ij} \neq 255 \\ 0, & \text{otherwise} \end{cases}$$

where:

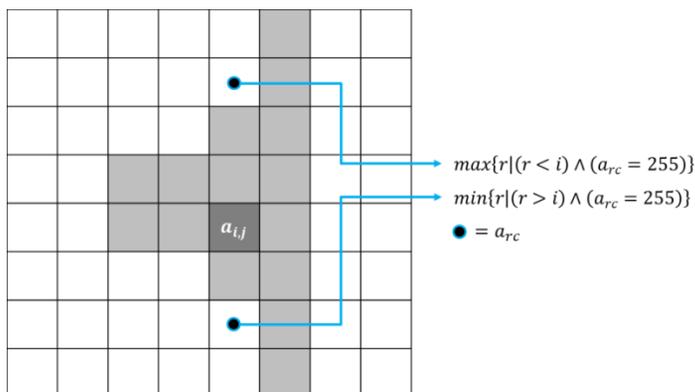
$$p_h = \min\{c | (c > j) \wedge (a_{rc} = 255)\} - \max\{c | (c < j) \wedge (a_{rc} = 255)\}, r = i \quad (2-19)$$

$$p_v = \min\{r | (r > i) \wedge (a_{rc} = 255)\} - \max\{r | (r < i) \wedge (a_{rc} = 255)\}, c = j$$

따라서, 픽셀값이 255가 아닌 픽셀 a_{ij} 은 수직 방향 연속성(p_v) 값과 수평 방향 연속성(p_h)값의 합으로 계산된다. 이때 수평 방향 연속성은 픽셀 a_{ij} 을 기준으로 우측에 있는 빈 픽셀의 열 인덱스 중 최소값과 좌측에 있는 빈 픽셀의 열 인덱스 중 최대값의 차이로 계산된다([그림 2-8]). 마찬가지로 수직 방향 연속성은 [그림 2-9]와 같이 픽셀 a_{ij} 을 기준으로 위쪽에 있는 빈 픽셀의 행 인덱스 중 최대값과 아래에 있는 빈 픽셀의 행 인덱스 중 최소값의 차이로 계산된다.



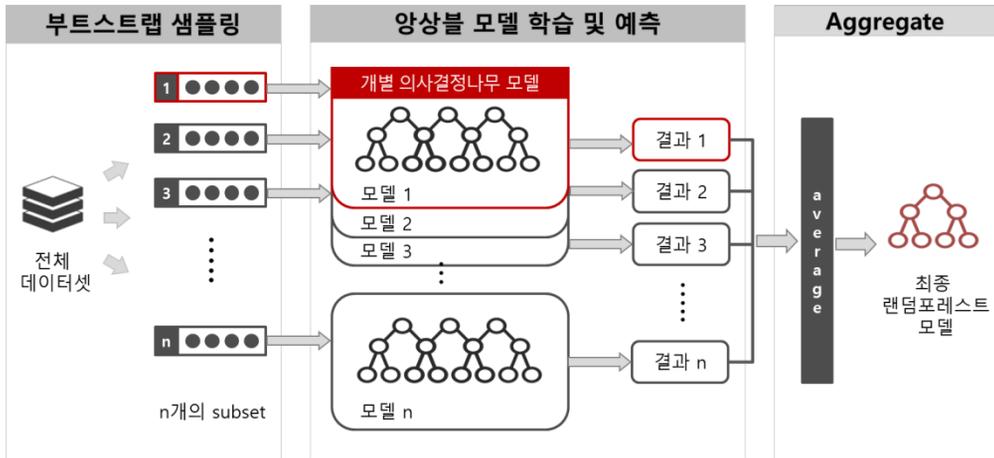
[그림 2-8] 수평 연속성 행렬 계산 예시



[그림 2-9] 수직 연속성 행렬 계산 예시

2.3.2 랜덤포레스트 모델

본 연구에서는 이진화를 위한 분류 모델로 랜덤포레스트 모델을 사용하였다. 해당 모델은 Breiman (2001)의 의사결정나무 모델을 고도화한 분류 모델이다. 랜덤포레스트는 [그림 2-10]과 같이 데이터 셋을 부트스트랩(Bootstrap) 샘플링 방법을 사용해 나누고, 이를 사용해 다수의 의사결정나무를 만든 후 이를 종합하여 결과를 도출한다. 따라서, 랜덤포레스트의 설명에 앞서 의사결정나무에 대한 설명이 필요하다.



[그림 2-10] 랜덤포레스트 모델 학습 흐름도

의사결정나무는 가장 기본적인 지도 학습 알고리즘으로 의사결정 규칙(나무) 생성, 가지치기, 타당성 평가, 해석의 절차로 진행된다 (James *et al.*, 2013). 먼저 의사결정규칙 생성 단계에서는 분석의 목적과 자료 구조에 따라 적절한 분리 기준과 정지규칙을 지정하여 생성한다. 이후 분류 오류를 높이는 가지를 제거한 후, 교차 타당성 평가를 통해 종속 변수와 독립 변수를 가장 잘 연결하는 나무를 구축한다.

이와 같은 의사결정나무 방법은 나무구조로 표현되어 결과 해석

이 용이하며 신경망 분석과 같은 예측 방법론에 비해 비교적 분석이 간단하고 속도가 빠르다는 장점이 있다. 하지만 교차 검증 방법을 거치는 데도 불구하고 입력 데이터에만 의존하여 모델을 생성하기 때문에 새로운 자료의 예측에서는 불안정하고 입력 데이터에 과적합할 위험이 있다 (Dietterich, 1995).

이러한 의사결정나무 모델의 단점을 극복하기 위해 배깅 (Bagging)과 앙상블 방법을 이용한 랜덤포레스트 모델이 제안되어 다양한 분야에서 활발히 활용되고 있다(Breiman, 1996; Breiman, 2001; Kwon *et al.*, 2022). 랜덤포레스트 모델은 다수의 의사결정나무의 결과를 통합하여 모델을 구성하기 때문에 과적합 문제를 해결할 수 있고, 각 의사결정나무끼리 독립적이기 때문에 학습 데이터의 노이즈에 강건하다는 장점이 있다.

랜덤포레스트 모델은 다수의 의사결정나무 모델에 의한 예측을 종합하는 앙상블 방법을 사용한다. 이때 각 의사결정나무 모델은 Breiman (1996)이 제안한 부트스트랩 샘플링 방식을 사용하여 생성한 학습 데이터 셋마다 개별적으로 구축된다. 즉, 각각의 의사결정나무 모델은 복원 추출을 통해 생성된 서로 다른 학습 데이터 셋을 사용하기 때문에 모델간 분산과 상관성이 적은 독립적인 모델을 구축할 수 있게 된다.

랜덤포레스트와 의사결정나무는 노드를 분할하는 방법에서도 차이를 갖는다. 의사결정나무는 분할을 위한 특징을 선택할 때 모든 특징을 고려하여 최적의 특징을 선택하는 반면, 랜덤포레스트는 전체 특징 중 일부를 무작위로 추출하여 최적의 분할을 만들어 나가는 임의의 노드 최적화 절차를 갖는다. 이때, 본 연구와 같은 분류 문제의 경우, 지니 지수(Gini index)를 최소화 하는 방향으로 노드를 분할하여 최적화 한다 (식 2-20). 여기서 c 는 사건의 수인데, 본 연구와 같은 이진 분류 문제의 경우 c 는 2가 된다.

$$\text{Gini Index}(s) = 1 - \sum_{i=1}^c P_i^2 \quad (2-20)$$

부트스트랩 샘플링 방식과 임의 노드 최적화와 같은 임의성에 의해 각 의사결정나무 모델이 서로 조금씩 다른 특성을 갖게 되며, 결과적으로는 일반화 성능이 향상된다. 또한, 임의 노드 최적화 과정에서 여러 특징을 무작위로 조합하기 때문에 본 연구와 같이 다수의 특징을 입력하는 경우에도 안정적인 결과를 도출할 수 있다(Caruana, Karampatziakis, & Yessenalina, 2008; Dudoit, Fridlyand & Speed, 2002). James (2013)은 임의 노드 최적화 절차를 통해 각각의 의사결정나무 모델이 무작위로 특징을 추출하여 학습하기 때문에 예측 성능이 강력한 특징이 있더라도 다양한 특징의 영향을 분석할 수 있는 장점이 있다고 서술한 바 있다.

또한, 랜덤포레스트는 부트스트랩 과정에서 표본으로 한 번도 뽑히지 않은 데이터를 바탕으로 Out-of-bag(이하 OOB) 자료를 테스트 셋으로 활용할 수 있다. 하지만 본 연구에서는 학습 데이터 셋과 테스트 셋을 따로 구성하여 모델 평가 자료로 사용하였다.

본 연구에서 딥러닝 기반 분류 모델이 아닌 머신러닝 기반 분류 모델을 선택한 이유는 크게 세 가지 이다. 첫 번째로 데이터 셋의 규모로 인한 한계이다. 딥러닝 모델은 모델이 스스로 데이터의 특성을 파악하기 위해 충분히 많은 양의 데이터가 필요하다. 본 연구에 사용된 데이터 셋의 개수는 총 50개이기 때문에 회전, 크기 변환, 밀림, 반사 등 다양한 데이터 증강 기법을 거치더라도 충분한 규모의 데이터 셋을 확보하는 데에 어려움이 존재했다. 반면 머신러닝 모델의 경우 데이터의 단위가 이미지 개수가 아니라 픽셀의 개수이기 때문에 가령 3000 ×

3000 크기의 이미지를 사용한다면 총 9,000,000개의 데이터를 확보하는 것과 같은 효과가 있다. 두 번째로, 본 연구에서 최종적으로 모델이 풀어야 하는 것은 이진 분류 문제이기 때문이다. 이진화는 입력으로 들어오는 픽셀의 특징 값을 바탕으로 배경과 선행객체 클래스 중 하나를 선택하는 이진 분류 문제인데, 이는 얼굴 인식이나 자율 주행과 같이 복잡한 추론을 통해 풀어야 하는 문제 대비 간단하다. 마지막으로 머신러닝 모델은 학습 과정에서 특징 중요도를 산출할 수 있으며 이를 통해 분류 과정에서 핵심적인 역할을 하는 특징을 확인할 수 있다. 또한 이후 알고리즘 확장 시 피쳐 엔지니어링을 위한 검증 정보로도 활용될 수 있다는 장점이 있다.

본 연구에서는 최종적으로 선택한 랜덤포레스트 모델과 더불어 동일한 특징 정보를 사용해 단일의사결정나무와 완전연결신경망에 대한 분류 학습도 진행하였다. 서포트 벡터 머신 모델의 학습은 고려하지 않았는데, 이는 서포트 벡터 머신이 학습 과정에서 많은 계산량을 필요하기 때문이다. 이러한 문제를 해결하고자 대량의 학습 데이터를 분할하고 분할된 학습 데이터 간 종속성이 없는 것처럼 병렬처리를 하는 연구들이 제안되었으나, 이러한 방법은 정확도 측면에서 차이가 있다는 문제가 발생한다(Alham, 2011). 본 연구의 학습 과정에서 사용된 이미지의 개수는 총 40개이며, 각 이미지당 대략 4,000,000 개 이상의 픽셀을 포함하고 있기에 학습 데이터의 양이 매우 크다. 또한 각 픽셀마다 24 개의 특징 정보를 포함하고 있기에 이러한 정보까지 포함하여 학습 시, 서포트 벡터 머신을 사용한 분석 방법은 컴퓨팅 자원이 매우 많이 요구되며 시간이 오래 걸린다는 단점이 존재한다. 이에 따라 본 연구에서는 서포트 벡터 머신을 이진화 분류 모델로 고려하지 않았다.

2.3.3 재귀적 특징 제거법

본 연구에서 추출한 영상 특징들과 같이 고차원의 자료를 머신러닝 모델에 학습시킬 시, 최적의 특징을 산출하는 특징 선택 과정은 모형의 과적합을 막고 정확도를 상승시킬 수 있다. 따라서, 본 연구에서는 랜덤포레스트 모델을 통해 산출된 각 특징 중요도(feature importance)로부터 최적의 변수를 선택하기 위해 Guyon *et al.* (2002)가 제안한 재귀적 특징 제거법(RFECV; Recursive Feature Elimination with Cross Validation)을 적용하였다.

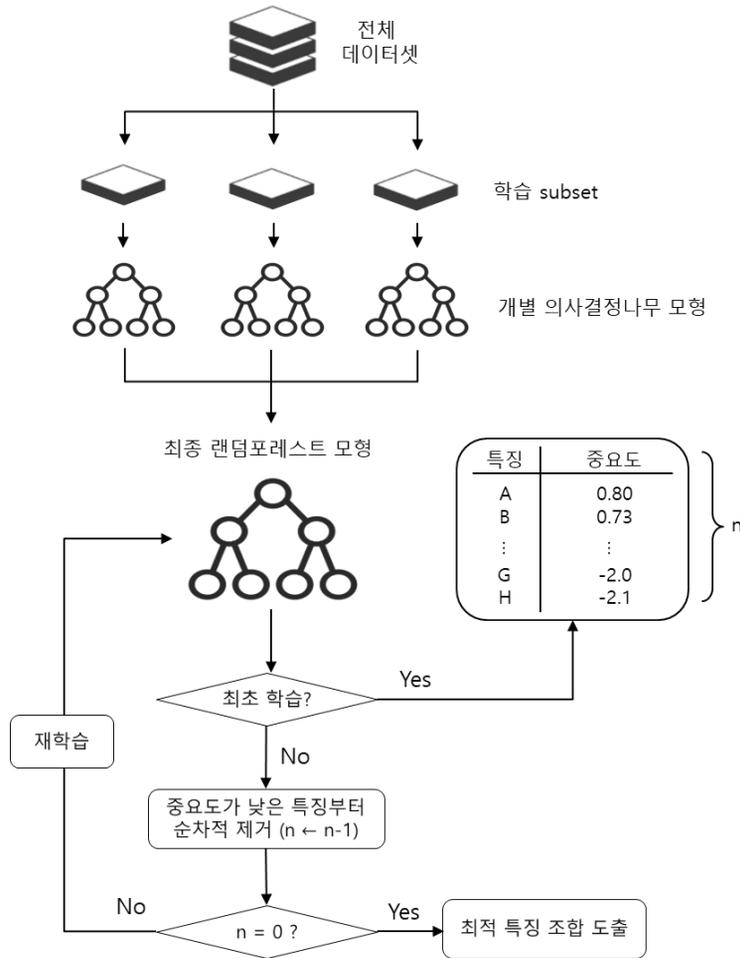
재귀적 특징 제거법의 설명에 앞서 특징 중요도를 산출하는 방법에 대해 설명해야 한다. 랜덤포레스트 모델에서 특징 중요도를 산출하는 식은 다음과 같다.

$$F(x) = \frac{1}{J} \sum_{j=1}^J (1 - Gini Index_j(x)) \quad (2-21)$$

식 (2-21)에서 x 는 입력 영상에서 추출한 특징이며, J 는 의사결정나무 모델의 개수이다. 즉, 특징 중요도는 랜덤포레스트 모델을 이루는 의사결정나무에서 각각의 특징이 불순도인 지니 지수를 얼마나 감소시키는 나타내며, 이는 각 특징 별 분류에 대한 기여도를 의미한다. 따라서, 중요한 특징일수록 해당 특징을 사용하여 노드를 분할하였을 때 불순도가 크게 감소한다는 것을 뜻한다.

이렇게 지니 지수를 바탕으로 랜덤포레스트에서 특징 중요도를 산출한 후, 재귀적 특징 제거법을 사용해 최적의 특징을 선택하였다. 재귀적 특징 제거법은 우선 앞서 도출된 특징 중요도를 기반으로 중요도가 낮은 특징부터 순차적으로 제외하여 랜덤포레스트 모델을 재학습하며 모

모델의 정확도가 최대인 지점에서의 특징 조합을 최적 특징으로 선택한다.
 이에 대한 구체적인 절차는 [그림 2-11]과 같다.



[그림 2-11] 재귀적 특징 제거법 흐름도

2.3.4 평가지표

본 연구에서 사용하는 평가지표는 평가 항목에 따라 모델 성능 평가와 이진화 결과 이미지 평가로 나누어 진다. 본 연구에서 사용한 머신러닝 모델의 분류 성능은 표와 같이 정확도, 정밀도, 재현율, F1-score로 평가하였다. 본 연구의 목적은 일제시기 건축도면처럼 노이즈가 심한 핸드드로잉 건축도면을 기존 개발된 도면 해석 연구에 사용될 수 있는 수준의 데이터로 변환하는 것이다. 따라서, 본 연구의 최종 출력 영상의 품질을 평가할 수 있는 지표를 추가로 사용하였다([표 2-3]).

모델 성능 평가 지표 중 정확도는 올바르게 예측된 픽셀의 개수를 전체 픽셀의 개수로 나눈 값이다. 본 연구에 사용된 데이터의 경우, 전체 이미지 대비 선형 픽셀이 차지하는 비율이 상대적으로 적기 때문에 정확도 지표만으로 모델의 성능을 평가하는 것은 부족하다. 따라서, 본 연구는 참으로 인식한 픽셀 중 실제로 참값인 픽셀의 개수를 나타내는 정밀도 지표도 추가적으로 사용하였다. 나아가 참값을 가지는 전체 픽셀 중 모델이 참이라고 예측한 픽셀의 개수의 비를 나타내는 재현율 지표도 함께 사용하였으며, 일반적으로 정밀도와 재현율은 서로 상충관계를 가지고 재현율은 연구자가 설정하는 파라미터 값에 따라 유동적으로 변한다. 따라서, 정밀도와 재현율을 동일한 중요성으로 두고 조화평균을 계산한 F1-score도 함께 고려하여 모델의 성능을 평가하였다.

본 연구에서 제안한 이진화 방법론을 거쳐 생성된 최종 출력 영상의 품질 평가는 PSNR(Peak Signal to Noise Ratio), NRM(Negative Rate Metric), DRD(Distance to Reciprocal Distortion), MPM(Misclassification Penalty Metric)지표를 사용하였다. PSNR 지표는 출력 이미지와 정답 이미지 간 유사도를 평가하는 지표이며, PSNR 값이 높을수록 두 이미지의 유사성이 높다고 판단한다. NRM은 모델이 맞추지 못한 픽셀에 대한 지표로 FN(False Negative)과 FP(False Positive) 지표의 평균으로 계산하기에 정답 이미지에 근사한 이진화 결

과 이미지일수록 NRM 값은 작아진다. MPM은 출력 이미지에서 잘못 분류된 픽셀에 대한 거리 기반 평가로 잘못 분류된 픽셀이 정답 이미지의 객체 경계로부터 떨어진 정도로 잘못 분류된 픽셀에 대한 평가를 진행한다. 만약 출력 이미지에서 도면 요소를 구성하는 픽셀의 경계에 있는 노이즈를 성공적으로 제거하였다면 MPM 값은 낮아지게 된다.

DRD는 Lu (2004)가 제안한 이진화 결과 이미지 평가 지표이며 Lu (2004)는 이진화 이미지로부터 글자나 기호와 같은 요소를 인식할 때 요소 중 뼈대가 되는 중앙 픽셀을 기준으로 해당 요소를 인식하기 때문에 중앙 픽셀로부터 오분류된 픽셀이 멀리 떨어질수록 글자 인식 수준이 낮아진다고 작성한 바 있다. 예를 들어 아래 [그림 2-12]의 경우 [그림 2-12 (a)]가 원본 이미지, [그림 2-12 (b)]가 각기 다른 형태로 일그러진(distorted) 이미지라면, [그림 2-12 (b)]의 좌측부터 우측으로 갈수록 오분류된 픽셀이 a 문자의 중앙 픽셀로부터 멀리 떨어진다.



(a) 정답 이미지 (b) 오분류 픽셀을 포함하는 결과 이미지 예시

[그림 2-12] 거리에 따른 오분류 픽셀 예시 (Lu, 2004)

[그림 2-12 (b)]의 결과 이미지는 모두 오분류된 픽셀을 포함하고 있는데, 가장 좌측 이미지의 경우는 오분류된 픽셀이 a 문자 인근에 위치해 있기 때문에 사용자가 문자를 인식하는 데에 큰 문제가 없다. 반면, [그림 2-12 (b)]의 가장 우측과 같이 오분류된 픽셀이 중앙 픽셀로부터 떨어져 있을수록 글자 인식이 어렵다. 따라서, DRD 지표는 세선화(skeletonization)을 통해 정답 이미지로부터 추출한 중앙 픽셀과 거리

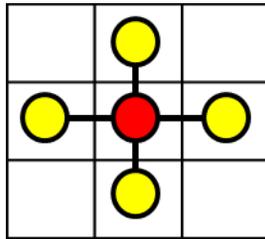
-대차 행렬을 사용하여 이진화 이미지의 일그러짐 정도를 측정할 때 오분류된 픽셀이 문자의 중심으로부터 멀리 떨어져 있을 수록 강한 패널티를 부여한다.

[표 2-3] 평가 항목에 따른 평가 지표 및 수식

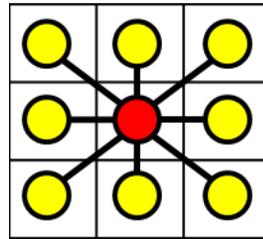
평가 항목	평가 지표	수식
분류 모델 성능 평가	정확도 (accuracy)	$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN}$
	정밀도 (precision)	$\text{Precision} = \frac{TP}{TP + FP}$
	재현율 (recall)	$\text{Recall} = \frac{TP}{TP + FN}$
	F1-score	$F1 = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right)$
이진화 결과 이미지 품질 평가	Peak Signal to Noise Ratio (PSNR)	$\text{PSNR} = 10 \times \log_{10} \frac{255^2}{MSE}$ where: $MSE = \frac{1}{X \times Y} \sum_{i=1}^X \sum_{j=1}^Y GT(i, j) - R(i, j) ^2$
	Negative Rate Metric (NRM)	$\text{NRM} = \frac{NR_{FN} + NR_{FP}}{2}$ where: $NR_{FN} = \frac{FN}{FN + FP}, NR_{FP} = \frac{FP}{FP + TN}$
	Distance Reciprocal Distortion (DRD)	$\text{DRD} = \frac{\sum_{k=1}^S \text{DRD}_k}{NUBN}$ where: $\text{DRD}_k = \sum_{i=-2}^2 \sum_{j=-2}^2 [GT_k(i, j) - R_k(x, y)] \times W_{N_m}(i, j)$
	Misclassifica tion Penalty Metric (MPM)	$\text{MPM} = \frac{1}{2D} \left(\sum_{i=1}^{FN} d_{FN}^i + \sum_{j=1}^{FP} d_{FP}^j \right)$ where D is the sum over all the pixel to contour distances of the ground truth object.

2.4 후처리

랜덤포레스트 모델을 통해 이진화가 완료된 이미지를 바탕으로 최종 후처리 단계로 이진화 선행 연구들(Jindal *et al.*, 2021; Jana *et al.*, 2017, Xiong *et al.*, 2021; Gatos *et al.*, 2006, Mitianoudis *et al.*, 2015)에서 일반적으로 사용되는 연결 성분 분석을 수행하여 면적이 작은 픽셀 단위 노이즈를 제거하였다. 이미지에서 픽셀들은 서로 연결된 형태로 나타나게 되는데 픽셀 간 연결성을 판단하는 방법에는 고려하는 방향에 따라 [그림 2-13]과 같이 두 가지 경우로 나타낼 수 있다. [그림 2-13 (a)]의 경우는 상하좌우로 인접한 4개의 픽셀을 고려하여 연결성을 판단하는 것이며 [그림 2-13 (b)]의 경우는 상하좌우와 더불어 대각선 방향으로 인접한 픽셀까지 고려하여 연결성을 판단하는 것이다.



(a) 상하좌우 방향



(b) 상하좌우 및 대각선 방향

[그림 2-13] 방향성에 따른 연결성 판단 종류

이러한 후처리 과정을 위해 본 연구에서는 OpenCV 라이브러리에서 제공하는 연결 성분 분석 함수를 사용하여 상하좌우 연결성을 판단하여 [그림 2-14]와 같이 선행 픽셀에 연결성 라벨을 부여하였다. 동일 라벨을 부여 받은 픽셀은 같은 색으로 표시되었으며, 이후 동일한 연결성 라벨을 부여 받은 픽셀의 개수가 30개 이하면, 해당

라벨을 부여 받은 픽셀 그룹이 노이즈라 판단하고 제거하였다. 그 결과 면적이 매우 작은 노이즈를 추가적으로 제거할 수 있었다.



[그림 2-14] 연결 성분 분석에 따른 픽셀 라벨링 결과

3. 실험 적용 및 결과

3.1 데이터 수집 및 전처리 결과

본 연구는 이진화를 위한 실험 데이터로 국가기록원에서 제공하는 일제시기 건축도면 디지털 아카이브에서 제공하는 도면 중 평면도 50개를 대상으로 [표 3-1]과 같이 데이터 셋을 구축하였다. 해당 데이터 셋은 병원, 대학교, 고등학교, 정부청사, 재판소, 미술관 등 다양한 목적의 건물로 이루어져 있다.

[표 3-1] 데이터 셋 도면 목록

번호	건물명	연도
1	경성제국대학 본부병대강당 평면도	1925
2	경성제국대학 본부병대강당 평면도	1925
3	경성제국대학 의학본부관 평면도	1926
4	경성제국대학 법문학부분관 평면도	1926
5	경성제국대학 법문학부분관 평면도	1926
6	경성제국대학 법문학부 설계도	1933
7	경성제국대학 도서관 설계도	1926
8	경성제국대학 도서관 설계도	1926
9	경성제국대학 도서관 설계도	1926
10	경성제국대학 도서관 설계도 1층평면	1933
11	경성제국대학 도서관 설계도 2층평면	1933
12	경성제국대학 예과 평면도	1923-1924
13	대한의원의학교 평면도	1909-1910
14	평양고등보통학교 평면도	1920
15	평양고등보통학교 평면도	1920
16	경성사범학교 부속소학교 평면도	1923
17	조선총독부상품진열관 이삼계평면도	1920
18	평리원급한성재판소평면도	1906-1907
19	평리원급한성재판소평면도	1906-1907
20	경성재판소수선모양채공사평면도	1909
21	경성재판소수선모양채공사평면도	1909
22	경성재판소수선모양채공사평면도	1909

23	공소원신축설계도(평면도)	1908
24	경성삼재판소합병청사신축평면도 - 지하실	1926-1927
25	경성삼재판소합병청사신축평면도 - 지하실	1926-1927
26	경성삼재판소합병청사신축평면도 - 2계	1926-1927
27	경성삼재판소합병청사신축평면도 - 3계	1926-1927
28	대한의원평면도 - 하계	1906
29	대한의원평면도 - 상계	1906
30	총독부의원외래본관계하평면도	1923
31	총독부의원외래본관계상평면도	1923
32	자혜의원계획설계도	1921
33	자혜의원계획설계도	1921
34	해주자혜의원본관신축설계도	1923
35	해주자혜의원본관신축설계도	1923
36	도립혜산진의원청사모양채평면도 1계	1927
37	도립혜산진의원청사모양채평면도 2계	1927
38	경성의학전문학교부속병원의외래진료소각계평면도	1927
39	인천세관청사분소실하물검사장설계지도	1911
40	진남포세관감시과청사신축설계도	1906-1924
41	경성세관출장소청사 청사재래평면도	1921
42	광화문총독부평사평면도 - 5계	1923-1924
43	부산경찰서청사신축공사설계도	1923
44	사세국개성출장소표본실건도급평면도	1912-1914
45	경성전매지국의주통공장신축평면도	1925
46	밀양영남루수선설계도	1918
47	시정오년조선물산공진회미술관신축설계도	1915
48	갑오관사신축공사설계도	1909
49	갑오관사신축공사설계도	1909
50	용산총독관저내외설비공사설계도	1920-1930

모델 학습을 위한 정답 데이터 셋은 픽셀 단위의 정답 데이터 라벨링이 가능한 SuperAnnotate 툴을 사용해 구축되었으며 정답 데이터 예시는 [그림 3-1]과 같다. 또한, 정답 데이터 구축 과정에서는 문자, 수치선, 도면 영역을 각각 구분하였지만 실제 학습 과정에서는 이들을 하나의 선형객체 클래스로 종합하여 사용하였다.



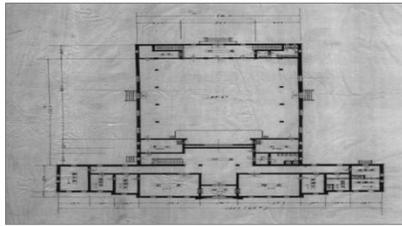
[그림 3-1] 라벨링 결과 (도면 2번, 6번)

3.2 배경 예측 및 제거 결과

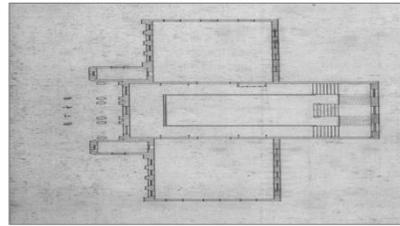
배경 예측 및 제거 방법은 1) 입력 이미지를 그레이 스케일로 변환한 뒤 양방향 필터로 이상치를 제거하는 단계; 2) 가우시안 혼합 모형을 사용해 이상치를 제거한 이미지의 픽셀 분포를 도면과 배경 군집으로 분리한 뒤 도면 군집에 해당하는 픽셀을 배경 군집 내 픽셀값으로 랜덤하게 대체하여 예측 배경 이미지를 생성하는 단계; 3) 그레이 스케일 입력 이미지에서 예측 배경 이미지를 빼줌으로써 배경 제거 이미지를 최종적으로 구하는 단계로 이루어진다.

[그림 3-2]에 이러한 각 단계의 결과를 시각화하였다. 해당 그림의 (a)~(c)는 원본 이미지, (d)~(f)는 해당 원본 이미지의 픽셀값 빈도, 그리고 (g)~(i)는 양방향 필터 적용 후 픽셀값 빈도이다. [그림 3-2 (a)]와 [그림 3-2 (b)]의 필터 적용 전 픽셀값 빈도의 경우, 양 끝 단인 0과 255에서 급격히 높은 빈도를 보였다. 이처럼 픽셀값에 이상치가 다수 존재하는 경우, 가우시안 혼합 모형을 기반으로 두 개의 단일 가우시안 분포 군집으로 분리하는 과정에 어려움이 존재한다. 따라서, 이러한 이상치 값을 인근 픽셀값으로 대체하여 최대한 쌍봉에 가까운 형태의 픽셀값 분포가 나오도록 양방향 필터를 적용하였다. 그리고 이러한 양방향 필터를 적용한 결과, [그림 3-2 (g)]와 [그림 3-2 (f)] 같이 기존의 이상치가 성공적으로 제거되고 픽셀값 빈도가 두 개의 정규 분포가 혼합된 형태로 변형되었다. 반면 [그림 3-2 (c)]의 경우, 원본 이미지의 픽셀값 분포가 필터 적용 전에도 쌍봉을 가지는 형태이다. 이러한 이미지는 양방향 필터 적용 없이도 가우시안 혼합 모형을 활용하여 뚜렷한 이진화가 가능하지만 본 연구에서 활용한 과거 도면들의 특성상 이러한 이상치가 발생하는 경우가 많기에 모든 도면에 대해서 일괄적으로 양방향 필터를 적용하였다.

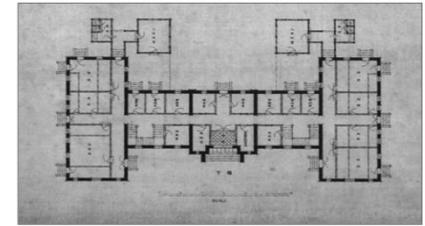
원본 이미지



(a)

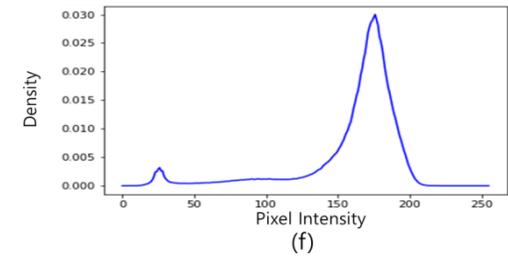
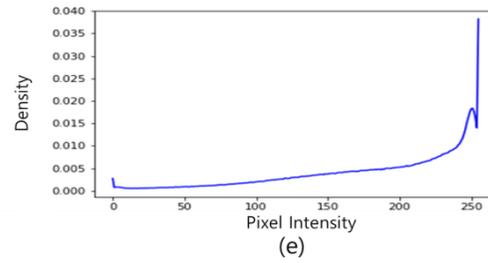
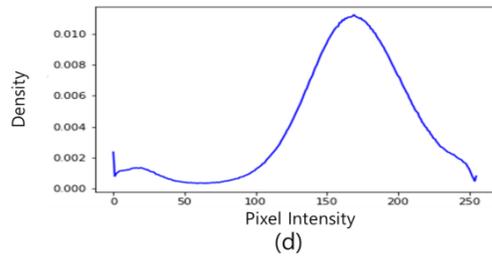


(b)

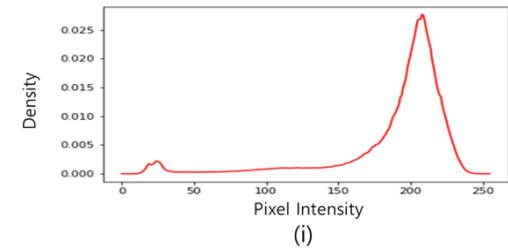
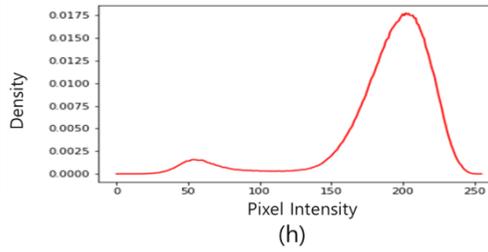
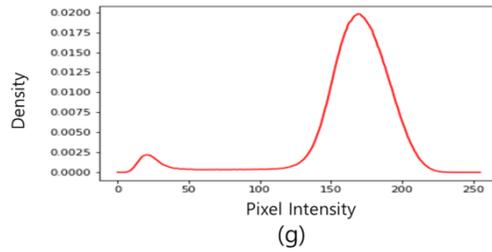


(c)

필터 적용 전 (前)

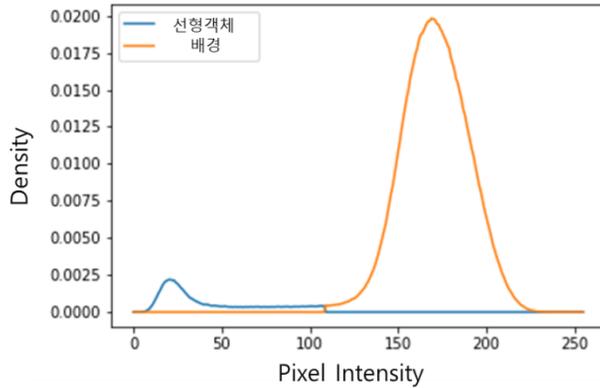


필터 적용 후 (後)



[그림 3-2] 양방향 필터 적용 전후 픽셀값 분포 예시

따라서, 양방향 필터를 적용한 도면들을 바탕으로 가우시안 혼합 모형을 통해 [그림 3-3]과 같이 선형객체 픽셀에 해당하는 정규 분포와 배경 픽셀에 해당하는 정규 분포로 군집화하였다.



[그림 3-3] 가우시안 혼합 모형 적용 후 픽셀값 군집화 예시

그 다음 단계로 배경 이미지를 추정하기 위해 선형 픽셀로 표시된 픽셀값을 배경 픽셀값으로 변환하여 예측 배경 이미지를 생성하였다. 이후 배경 제거를 위해 입력 이미지와 예측된 배경 이미지의 픽셀 간 픽셀값 차이를 통해 배경 제거 이미지를 계산하였다. 이러한 과정을 통해 배경에 전반적으로 넓게 존재하는 물결 무늬의 얼룩, 스캔 과정에서 발생한 반복적인 노이즈들은 대부분 지워지게 되었다 ([그림 3-4]).



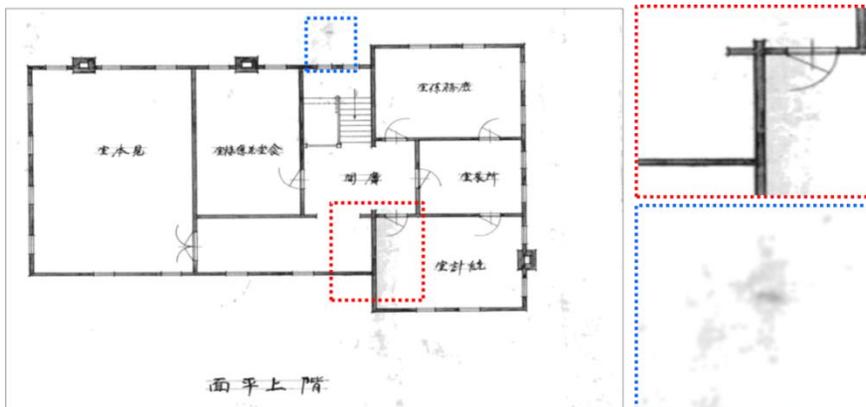
(a) 입력 이미지 (b) 예측 배경 이미지 (c) 배경 제거 이미지

[그림 3-4] 예측 배경 이미지와 배경 제거 이미지 예시

그러나, [그림 3-5]와 [그림 3-6]과 같이 도면과 유사한 픽셀 값을 가지면서 면적이 작은 얼룩과 도면 작성 과정에서 발생한 펜 흘림류의 노이즈는 제거하지 못하는 한계점이 존재하였다. 이러한 한계점은 이후 이어지는 머신러닝 기반 학습 과정을 통해 선형 픽셀과 노이즈 픽셀간 특징 차이를 학습하는 방향으로 극복하였다.



[그림 3-5] 배경 예측 및 제거 단계 이후 잔여 노이즈 예시 (1)

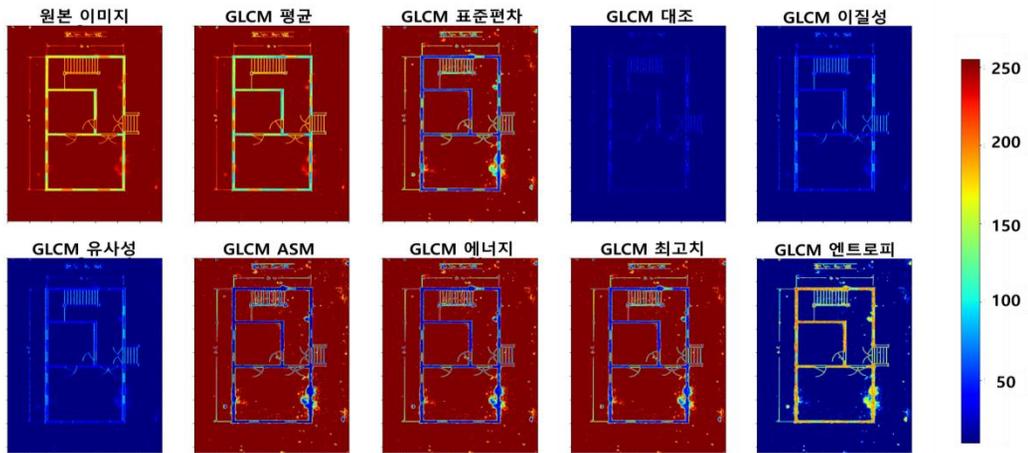


[그림 3-6] 배경 예측 및 제거 단계 이후 잔여 노이즈 예시 (2)

3.3 특징 추출 결과

3.3.1 명암도 동시발생 행렬 특징 추출 결과

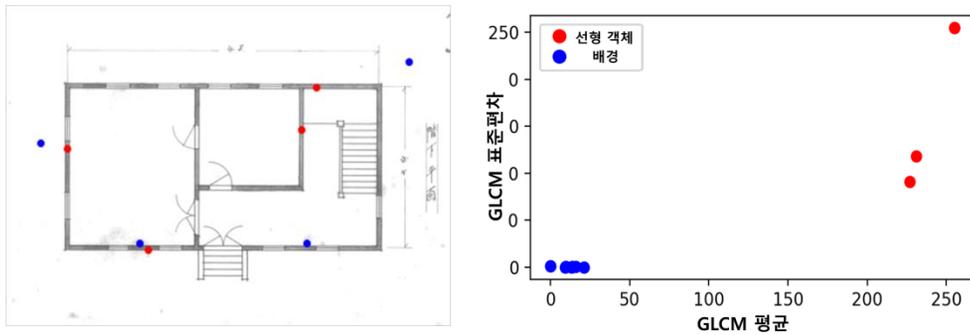
명암도 동시발생 행렬의 2차 통계는 3.2절에서 최종적으로 출력한 배경 제거 이미지로부터 추출한다. [그림 3-7]은 각 명암도 동시발생 행렬의 2차 통계량을 0에서 255 사이로 정규화하여 시각화한 결과이다.



[그림 3-7] 명암도 동시발생 행렬 특징 추출 예시

[그림 3-7]에서 나타난 바와 같이 명암도 동시 발생행렬의 평균과 표준편차는 도면과 배경을 잘 구분하는 특징이다. 반면, 이외의 대조, 이질성, 유사성은 배경과 도면 배경을 잘 구분하지 못하며, 나머지 특징들은 도면과 유사한 픽셀값을 가진 노이즈를 도면과 유사한 특징값으로 추출한다는 한계를 확인할 수 있었다. 명암도 동시 발생 행렬 구축 후 산출한 대조, 이질성, 유사성, 균등성, 에너지, 엔트로피와 같은 2차 통계량은 도면과 배경을 구분하는 유의미한 특징으로 활용되지 못하는 것을 확인할 수 있다. [그림 3-8]은 선형객체와 배경의 명암도 동시발생

행렬 평균과 표준편차 간 관계성을 나타내었다. 여기서 빨간색 점은 선형객체의 픽셀이며, 파란색 점은 배경 영역의 픽셀이고 아래의 그래프는 각 점에서 추출한 명암도 동시발생 행렬의 평균을 X축, 표준편차를 Y축에 표시한 것이다. [그림 3-8 (b)]에 나타난 바와 같이 선형객체 픽셀의 명암도 동시발생 행렬 평균 및 표준편차는 배경 픽셀 대비 상대적으로 큰 값이 추출되었다.



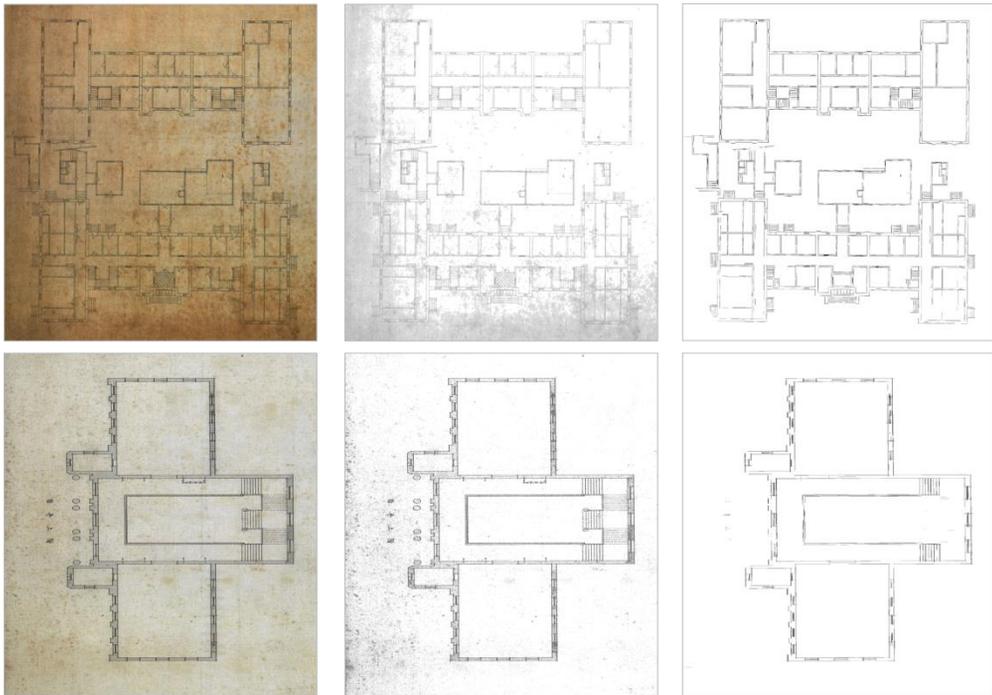
(a) 추출된 특징의 위치

(b) 위치에 따른 특징 값 산점도

[그림 3-8] 특징 추출 위치에 따른 명암도 동시발생 행렬 평균과 표준편차 간 관계성

3.3.2 수직-수평 연속성 행렬 특징 추출 결과

수직-수평 연속성 행렬은 특히 배경 예측 및 제거 단계를 거친 이후 여전히 배경 노이즈가 심한 도면에서 성능이 두드러진다. 아래 [그림 3-9]은 원본 이미지의 배경이 매우 어두워서 배경 제거 이후에도 노이즈가 여전히 남아있는 도면이다. 하지만 수직-수평 연속성 행렬은 이러한 이미지도 도면에 해당하는 픽셀의 특징값을 성공적으로 추출하였다. 이는 수직-수평 연속성 행렬이 동일한 픽셀값이 수직 및 수평 방향으로 연속적으로 나타나는 패턴을 탐지하는 방식으로 작동하기 때문이다. 그래서 배경 노이즈처럼 다양한 픽셀값이 넓은 영역에 걸쳐 나타나는 영역(배경)에서는 수직-수평 연속성 특징이 추출되지 않는다. 반면 선형 객체처럼 유사한 픽셀값이 연속적으로 나타나는 영역에서는 수직-수평 연속성 특징이 추출된다.



(a) 원본 이미지

(b) 배경 제거 이미지

(c) 수직-수평 연속성 행렬

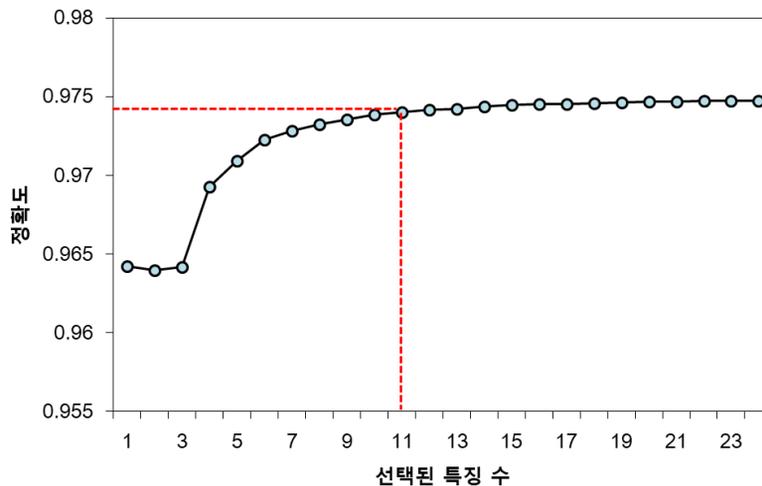
[그림 3-9] 수직-수평 연속성 행렬 특징 추출 예시

3.4 머신러닝 기반 도면 이진화 평가 결과

3.3절에서는 3.2절의 배경 제거 이미지를 바탕으로 총 24개의 특징을 추출한 결과를 살펴보았다. 본 절에서는 이렇게 추출된 특징을 사용해 학습한 랜덤포레스트 모델의 결과에 대해 살펴본다. 3.4.1절에서 각 특징 별 중요도와 재귀적 특징 제거법으로 도출한 최적 특징 조합에 대해 설명하고, 3.4.2절에서 선행 연구와 본 연구에서 제안하는 이진화 기법의 성능을 비교한다. 마지막으로 3.4.3절에서는 이진화 결과 이미지에 대한 품질 평가 결과를 살펴본다.

3.4.1 특징 중요도 및 최적 특징 조합

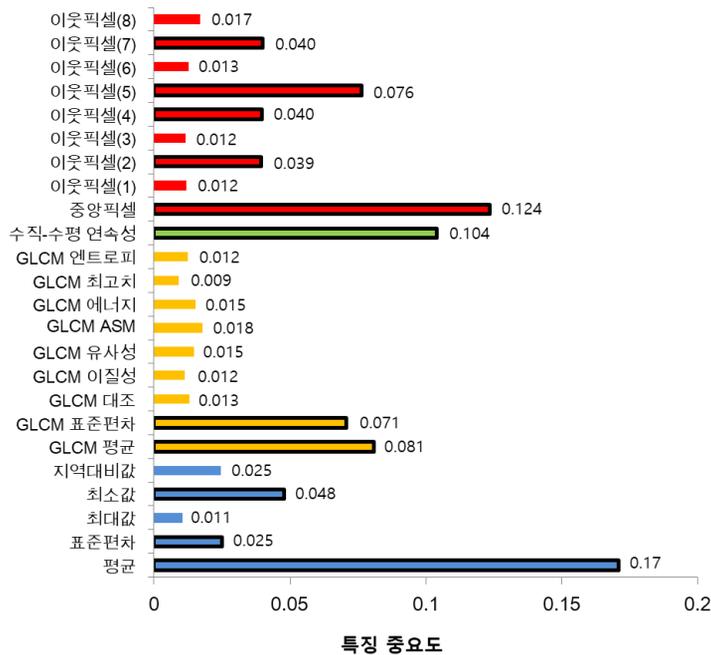
3.2절에서 최종적으로 배경이 제거된 이미지를 바탕으로 추출된 특징 24개 중 최적의 특징 조합을 선택하여 랜덤포레스트 모델을 구축하기 위해 우선적으로 재귀적 특징 제거법을 적용하였다.



[그림 3-10] 특징 개수에 따른 정확도 변화

[그림 3-10]은 재귀적 특징 제거법을 바탕으로 상대적 중요도가 낮은

특징부터 순차적으로 제거할 시 모델의 정확도 성능을 시각화 한 결과이다. [그림 3-10]에 나타난 바와 같이 24개의 특징을 모두 활용하였을 때 가장 높은 정확도인 0.975로 나타났다. 하지만, 특징을 11개 이상 사용할 시, 모델 성능이 크게 증가하지 못하고 0.975로 수렴하는 결과를 보였다. 그러므로 과적합을 방지하면서 동시에 높은 수준의 분류 성능을 나타내는 최적 특징 개수는 11개이다.

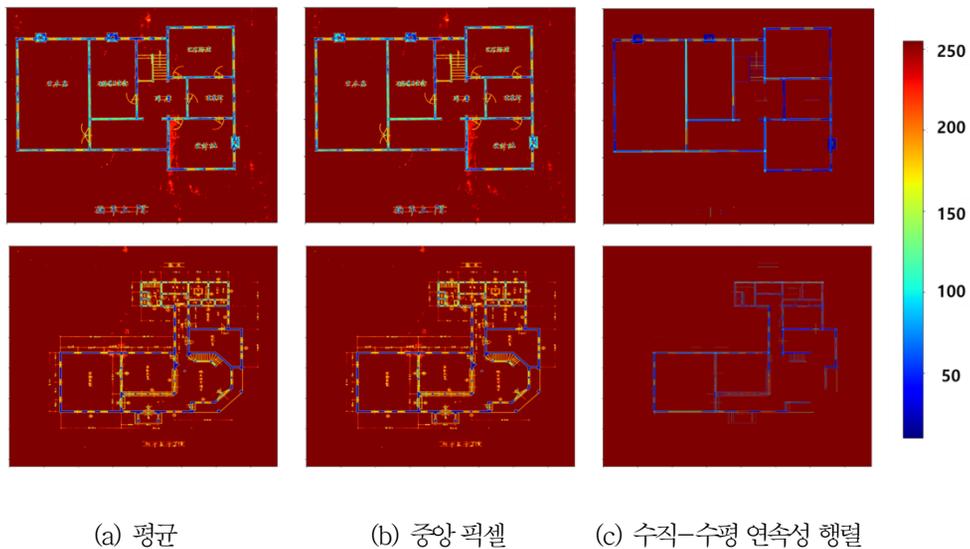


[그림 3-11] 특징 별 상대적 중요도

[그림 3-11]에서는 모델 학습에 사용한 각 특징들의 상대적 중요도를 나타내었으며 평균의 특징 중요도가 0.17로 가장 높음을 확인할 수 있다. 따라서 [그림 3-10]에서 선택된 특징 개수가 1개일 때는 평균이라는 특징 하나만 사용하여 분류를 진행했을 때의 정확도를 나타내는 것이다. [그림 3-10]에서 정확도 성능이 수렴하는 결과가 나오기 시작하였던 특징은 11개이며, 이때 선택된 특징은 [그림 3-11]에 굵은 테두리로 표시되어 있다. 선택된 특징을 나열하자면 통계적 특성에서 3개, 명암도

동시발생 행렬의 통계적 특성에서 2개, 수직-수평 연속성 행렬에서 1개, 중앙 및 이웃 픽셀값에서 5개이다. 특히 명암도 동시발생 행렬의 통계적 특성 9개 중 7개의 특징 중요도가 매우 낮았는데, 이는 본 연구에 사용된 건축도면 데이터 셋의 특성 상 질감 요소가 뚜렷하지 않기 때문으로 사료된다.

특징 별 상대적 중요도를 살펴본 결과, 평균, 중앙 픽셀, 수직-수평 연속성이 가장 중요한 특징으로 나타났으며 이러한 3개의 특징만 활용하여도 0.965이상의 정확도가 산출되었다. [그림 3-12]는 이러한 선형객체 추출을 위한 핵심 특징을 시각화한 것이다. [그림 3-12]에 나타난 바와 같이 평균과 중앙 픽셀의 경우 면적이 큰 노이즈와 선형 픽셀의 값이 확실히 구분되는 경향을 보였다.



[그림 3-12] 핵심 특징 추출 예시

수직-수평 연속성 행렬의 경우 도면과 선형객체의 특징 값 차이가 두드러졌다. 하지만 수직-수평 연속성 행렬은 문이나 계단과 같이 두께가

얇은 도면 요소의 픽셀 특징은 제대로 추출하지 못하는 경향을 보였다. 이처럼 특정 특징 하나만으로 선형객체와 배경을 완벽하게 구분하기 보다, 배경 노이즈 분류에 우수한 특징과 면적이 작고 진한 노이즈를 잘 분류하는 특징을 종합적으로 학습한 것으로 판단된다. 본 연구에서는 최종 성능 평가 시 최대 정확도를 나타내기 위해 24개 특징을 모두 사용하였지만 이후 추가 데이터 셋을 학습 시킬 때에는 핵심 특징 정보만 활용하여 보다 효율적이고 간편하게 이진화 모델을 구축할 수 있을 것으로 사료된다. 또한 본 연구에서는 동일한 24개 특징을 추출하여 단일의사결정나무 완전연결신경망 기반의 이진화 모델을 학습하였다. 그러나 랜덤포레스트 모델을 최종 모델로 선택하였는데, 이는 랜덤포레스트 모델이 타 머신러닝 모델 대비 크게 두 가지 장점이 있기 때문이다. 첫 번째로 랜덤포레스트 모델은 다수의 독립적인 결정트리 모형을 결합하여 해당 모형들에서 나온 결과들의 가중 평균을 통해 분류가 이루어지기 때문에 결정트리 모형의 빠른 학습 속도와 더불어 노이즈에 강하고 과적합 위험이 낮다. 이는 건축도면의 노이즈처럼 크기, 색, 밝기가 다양한 데이터에 크게 영향을 받지 않다고 해석할 수 있다. 두 번째로 랜덤포레스트 모델은 학습 과정에서 사용하는 지니 지수를 바탕으로 각 특징의 상대적 중요도를 산출할 수 있다. 이러한 상대적 중요도를 바탕으로 향후 핵심 특징만 사용하여 모델 고도화 및 학습 속도 개선이 가능하다. 또한 이진화 성능에 가장 중요한 특징에 대한 이해를 통해 추가적으로 유의미한 관련 특징을 추출할 수 있는 단서를 제공하기도 한다.

3.4.2 분류 모델 성능 비교

3.4.2절에서는 3.4.1절에서 제시한 랜덤포레스트 모델 기반 픽셀 분류기와 기존 선행 연구에서 제시한 이진화 방법론의 모델 성능을 비교하였다([표 3-2]). 성능 비교에는 전체 데이터 셋 50개 중 학습 과정에 사용한 도면 40개를 제외한 10개 도면을 사용하여 진행하였고, 각 테스트 셋의 성능을 종합하여 평균을 계산하였다.

그 결과, 머신러닝 기반 이진화를 제외한 기존 이진화 방법론들 모두 정밀도는 0.95이상으로 높은 편이었으나 재현율이 낮아 실제 도면인 픽셀들의 분류에 비교적 낮은 성능을 나타냈다. 반면, 머신러닝 기반 이진화는 모든 모델 성능 지표에서 0.96이상의 높은 성능을 보였다. 특히, 랜덤포레스트 모델은 재현율과 정밀도의 조화평균인 F1-score 지표가 0.987로, 가장 우수한 성능을 보여 기존 연구들과 타 머신러닝 모형의 성능 대비 상대적 우위를 보였다.

[표 3-2] 분류 모델 성능 비교표

분류	방법론	정밀도	재현율	정확도	F1-score	평균
전역이진화	Otsu (1979)	0.985	0.703	0.721	0.793	0.801
지역이진화	Niblack (1986)	0.954	0.56	0.575	0.708	0.700
	Sauvola (2000)	0.972	0.715	0.722	0.799	0.802
	Wolf (2002)	0.956	0.873	0.848	0.904	0.896
규칙 기반 이진화	Lu (2010)	0.981	0.911	0.902	0.944	0.935
	Su (2010)	0.965	0.797	0.792	0.853	0.852
머신러닝 기반 이진화	단일 의사결정나무	0.978	0.981	0.963	0.978	0.975
	랜덤포레스트	0.985	0.990	0.976	0.987	0.984
	완전연결신경망	0.981	0.992	0.975	0.986	0.983

3.4.3 이진화 결과 이미지의 품질 비교

본 절에서는 랜덤포레스트 모델 기반 픽셀 분류기와 기존 선행 연구에서 제시한 이진화 방법론의 모델들의 출력 이미지 품질을 측정하여 비교하였다([표 3-3]). 모델 성능 비교와 마찬가지로 학습 과정에 사용하지 않은 10개의 테스트 셋에 대하여 비교 검증을 수행하였다.

정답 이미지와 이진화 결과 이미지의 각 픽셀을 비교하여 노이즈 비율을 계산하는 PSNR 지표는 각 픽셀값 차이를 평균한 제곱근 오차 (Mean-Square Error; MSE)를 사용하여 계산한다. PSNR 지표가 높을 수록, 결과 이미지가 정답 이미지에 유사하다고 해석할 수 있다. 이러한 PSNR 지표는 랜덤포레스트 모델이 16.543으로 가장 우수한 성능을 보였으며, 타 연구들에 비해 전반적으로 머신러닝 기반 이진화 모형들의 성능이 우수하였다.

FN(False Negative)와 FP(False Positive)지표의 평균을 나타내는 NRM 지표는 그 값이 낮을수록 출력 이미지에서 잘못 분류된 픽셀이 없다고 해석할 수 있다. NRM 지표 역시 랜덤포레스트 모델이 0.093으로 가장 우수한 성능을 보였으며, 0.413로 가장 높은 값을 나타냈던 Nibalck (1986) 대비 약 4배 가량의 성능이 개선되었다.

MPM 지표 역시 오분류된 픽셀에 대한 평가이지만 오분류된 픽셀과 정답 객체의 경계 픽셀 간 거리를 고려할 수 있는 장점이 있다. MPM 지표 또한 그 값이 낮을수록 정답 이미지의 실내 객체의 경계와 결과 이미지의 실내 객체 경계가 유사한 것을 의미한다. 이러한 MPM은 랜덤포레스트 모델과 완전연결신경망이 모두 0.001로 가장 낮은 MPM 결과를 보였으며, 기존 연구들의 모형 대비 상당한 성능 개선을 확인할 수 있었다.

마지막으로, DRD 지표는 오분류된 픽셀이 객체의 중심에 인접한 지를 고려한 지표이다. DRD 지표는 오분류된 픽셀과 객체 중심 픽셀까

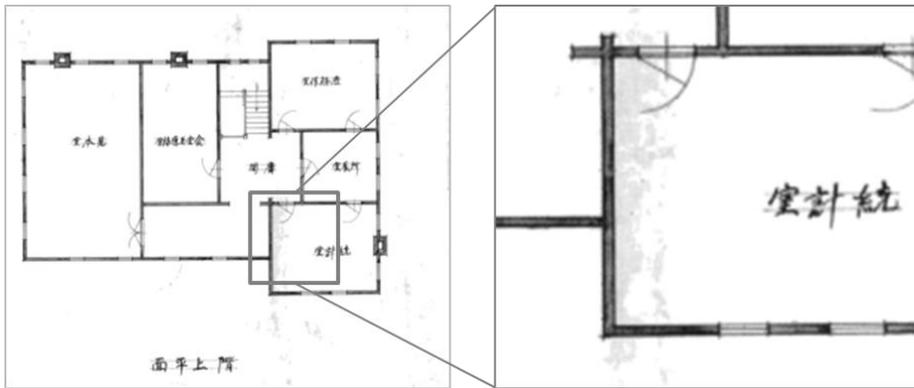
지의 거리를 가중치로 부여하여 멀리 떨어진 픽셀일수록 높은 가중치를 부여한다. 즉, 같은 오분류된 픽셀이더라도 정답 이미지에서 객체를 구성하는 뼈대 픽셀과 가까우면 영향이 낮고, 멀면 영향이 높은 오분류 픽셀로 판단한다. 해당 지표는 모든 모형들에서 비슷한 수준의 값이 산출되었으며, Wolf (2002)가 0.128로 가장 낮은 결과를 보였다. 하지만, 각 모형들에서 산출된 지표의 표준편차가 0.032 수준으로 모든 모형들이 낮은 결과를 산출하였다. 따라서, 성능 비교 결과에서 나타난 바와 같이 이진화 결과 이미지의 품질 역시 본 연구에서 제안한 랜덤포레스트 모델에서 가장 고품질의 이미지를 출력할 수 있었고 머신러닝 모형들이 기존 연구대비 더욱 도면 이진화에 적합하다는 결론을 얻을 수 있었다. 또한, 이러한 머신러닝 모형의 우수성은 본 연구에서 머신러닝 모형들의 입력 자료로 활용한 24가지의 특징이 도면 이진화에 매우 적합하기 때문으로 사료된다.

[표 3-3] 이진화 결과 품질 비교표

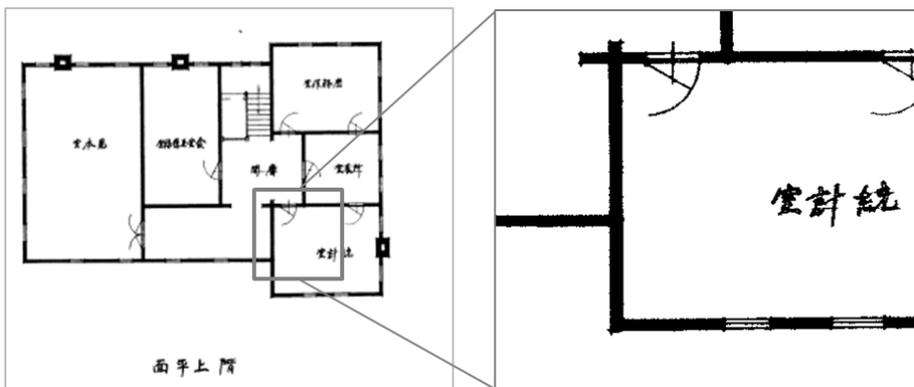
분류	방법론	PSNR	NRM	MPM	DRD
전역이진화	Otsu (1979)	8.346	0.202	0.136	0.218
지역이진화	Niblack (1986)	3.159	0.413	0.258	0.136
	Sauvola (2000)	8.069	0.249	0.138	0.203
	Wolf (2002)	10.009	0.294	0.060	0.128
규칙 기반 이진화	Lu (2010)	10.376	0.147	0.011	0.201
	Su (2010)	9.057	0.253	0.097	0.176
머신러닝 기반 이진화	단일 의사결정나무	14.570	0.132	0.002	0.195
	랜덤포레스트	16.543	0.093	0.001	0.212
	완전연결신경망	16.314	0.113	0.001	0.220

3.5 머신러닝 기반 도면 이진화 적용 결과

머신러닝 기반 도면 이진화는 배경 그림 [3-13 (a)]와 같이 예측 및 제거 단계 이후 여전히 남아있는 면적이 작고 다양한 형태의 노이즈를 처리하는 단계이다. 2.2절에서 수행되는 배경 제거 과정 이후에도 상대적으로 면적이 작고 선형객체와 유사한 색상값을 가지는 노이즈 남아 있는데, 본 절에서는 머신러닝 기반 분류 모델인 랜덤포레스트 모델을 사용해 잔여 노이즈를 감소한다. 앞서 추출한 24개의 특징 값을 바탕으로 입력 픽셀이 배경인지, 선형객체인지 판단하는 분류 모델을 통해 이진화를 수행하면 [그림 3-13 (b)]와 같이 잔여 노이즈마저 감소할 수 있다.



(a) 배경 제거 이미지(머신러닝 기반 이진화 적용 전)



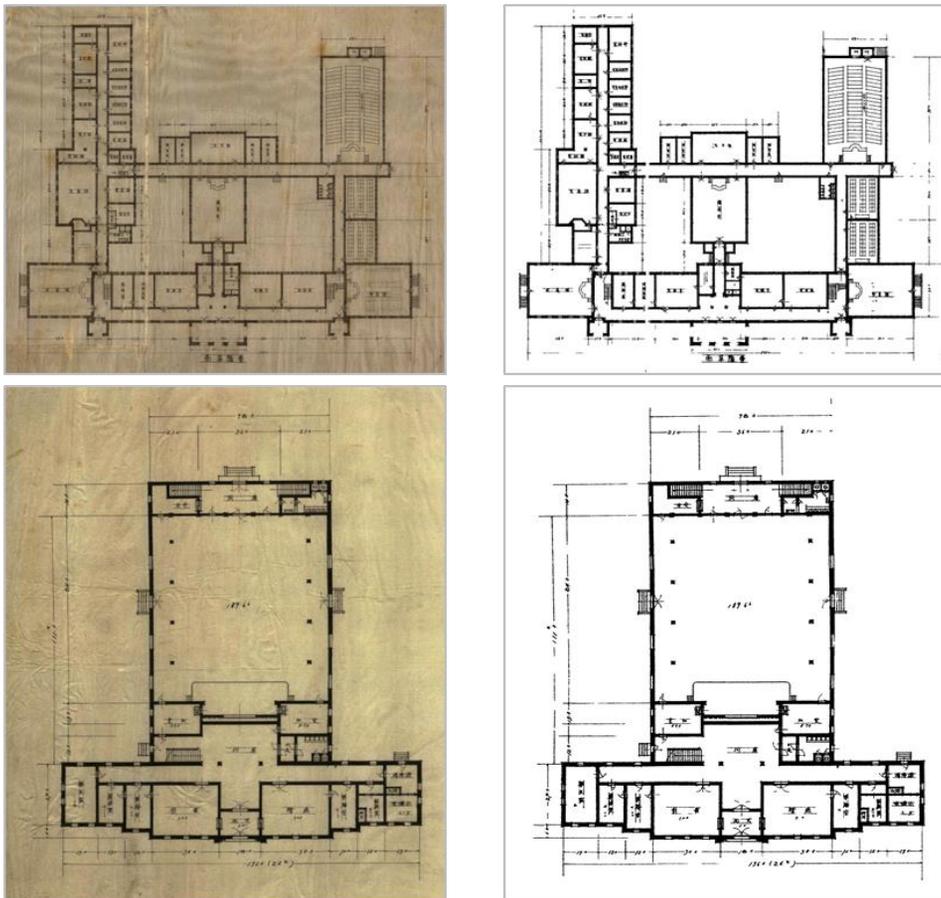
(b) 결과 이미지(머신러닝 기반 이진화 적용 후)

[그림 3-13] 머신러닝 기반 이진화 적용 후 노이즈 제거 결과

본 절에서는 테스트 셋 10개를 대상으로 랜덤포레스트 모델 적용 후 후 처리까지 수행하여 얻은 최종 이진화 이미지를 살펴본다. 결과 분석은 이미지의 해상도에 따라 구분하여 해석하였는데, 이는 건축도면의 축척에 따라 선형객체 크기가 상대적으로 크거나 작게 표현되기 때문이다. 따라서 도면 이진화 적용 결과, 다양한 축척의 건축도면에서 선형객체와 배경이 성공적으로 구분되었는지 분석하고자 한다. 3.5.1절은 축척이 작은 도면에서의 이진화 적용 결과를 다루며, 3.5.2절에서는 축척이 큰 도면에서의 이진화 적용 결과를 다룬다. 각 절에서는 선형객체를 세부적으로 벽, 문, 창문, 계단 기준으로 총 4개의 테스트 셋 결과를 다뤘으며 나머지 6개의 테스트 셋에 대한 결과 에 대한 결과 예시들은 <부록 A>에서 확인할 수 있다.

3.5.1 소축척 도면에서의 이진화 적용 결과

본 절에서는 소축척 도면의 이진화 적용 결과를 선형객체 별로 살펴본다. 일반적으로 소축척 도면에서는 전체 이미지 면적 대비 도면 요소가 작게 표현된다. [그림 3-14]은 테스트 셋 10개 중 소축척 도면 2개에 대한 이진화 적용 결과를 나타낸 것이다. 적용 결과, 원본 이미지의 배경에 존재하는 종이 구겨짐, 물결 무늬로 나타나는 불규칙적 배경 밝기와 같은 노이즈가 대부분 감소된 것을 확인할 수 있다.



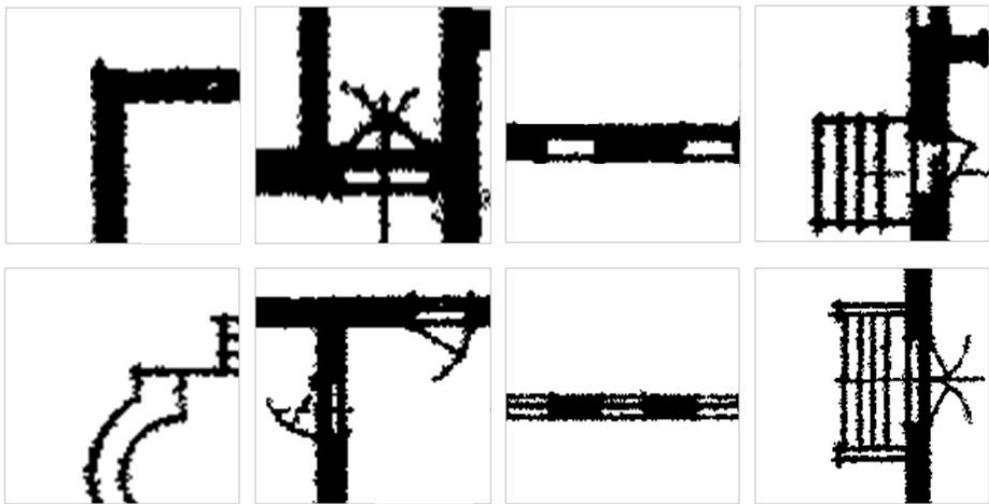
(a) 원본 이미지

(b) 이진화 결과 이미지

[그림 3-14] 소축척 도면의 이진화 적용 결과

다음으로 [그림 3-15]에서는 실내 공간정보 모델링 시 핵심적인 선형

객체의 이진화 적용 결과를 확대하였다. 이진화 결과, 다양한 두께의 벽체가 모두 끊기는 부분 없이 선형객체와 배경이 구분되었다. 도면 요소 중 크기가 작은 문과 창문 역시 유실되는 부분 없이 구분되었으며 두께가 얇은 다수의 선형객체로 구성되는 계단도 성공적으로 이진화 된 것을 확인할 수 있다.



(a) 벽

(b) 문

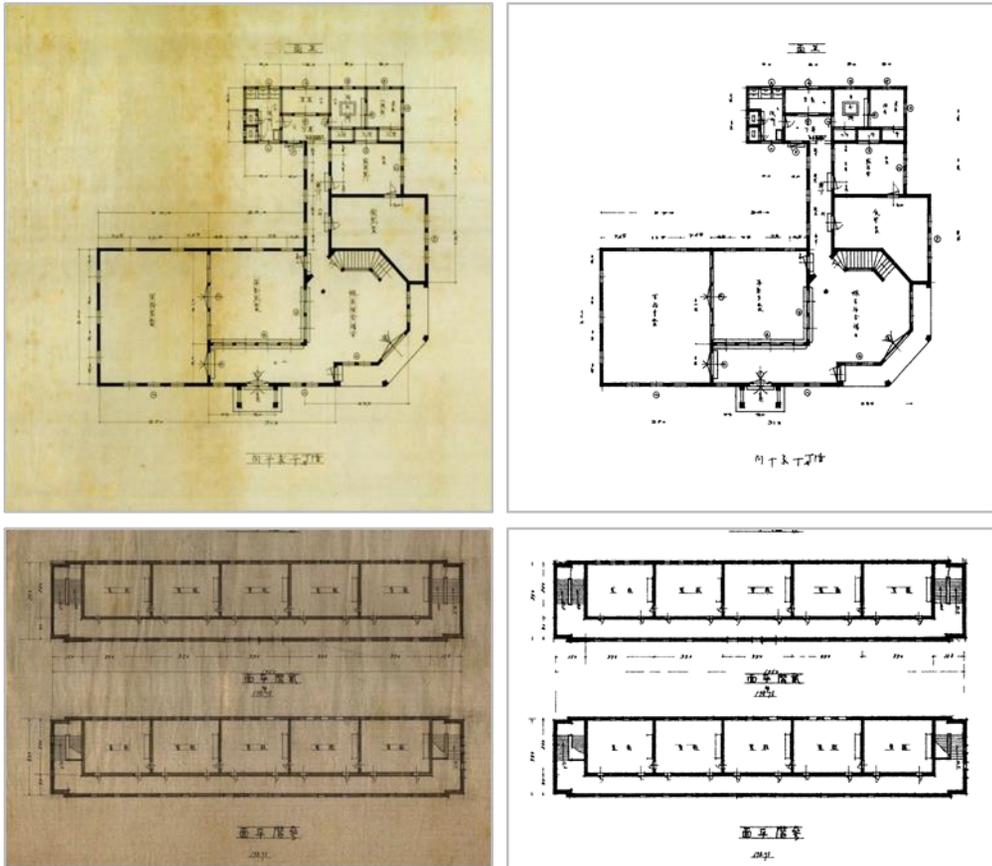
(c) 창문

(d) 계단

[그림 3-15] 소축척 도면의 선형객체 별 이진화 적용 결과

3.5.2 대축척 도면에서의 이진화 적용 결과

대축척 도면에서는 전체 이미지 면적 대비 도면 요소가 크게 표현된다. [그림 3-16]은 테스트 셋 10개 중 대축척 도면 2개에 대한 이진화 적용 결과를 나타낸 것이다. 소축척 도면의 이진화 적용 결과와 마찬가지로 다양한 형태의 배경 노이즈가 크게 감소하여 전반적으로 선형객체와 배경이 명확하게 구분된다.



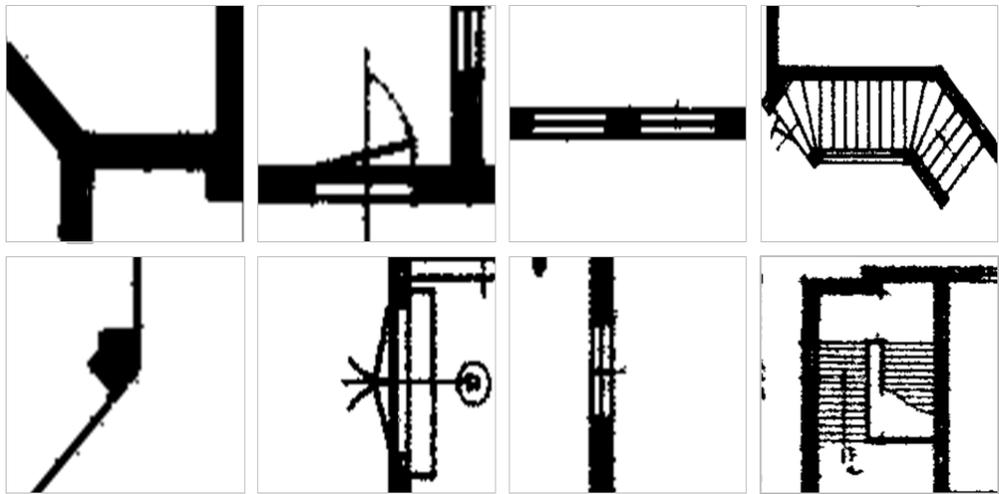
(a) 원본 이미지

(b) 이진화 결과 이미지

[그림 3-16] 대축척 도면의 이진화 적용 결과

[그림 3-17]은 소축척 도면의 선형객체 별 이진화 적용 결과를 보다 자세히 확대하였다. 이진화 적용 결과, 소축척과 유사하게 벽, 문,

창문, 계단 등 실내 객체가 배경과 명확히 구분되었다. 특히 창문이나 계단과 같이 두께가 얇은 다수의 선이 좁게 정렬된 형태의 객체의 경우는 소축척 보다 경계 인근 노이즈가 더욱 완벽히 감소한 것을 확인할 수 있었다.



(a) 벽

(b) 문

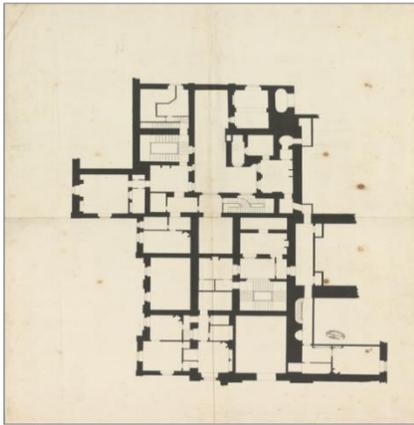
(c) 창문

(d) 계단

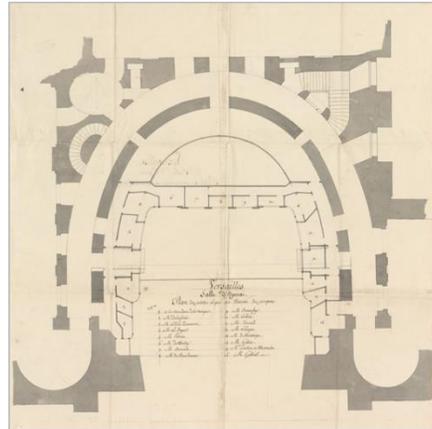
[그림 3-17] 대축척 도면의 선형객체 별 이진화 적용 결과

3.6 다양한 핸드드로잉 건축도면의 이진화 평가 및 적용 결과

본 연구에서 제안한 이진화 알고리즘의 범용적인 활용성을 검증하기 위해 다른 유형의 건축도면에 적용하여 성능 평가 및 적용 결과를 살펴보았다. 추가적으로 사용한 도면은 17세기와 18세기에 설계된 프랑스의 고전 궁전 중 하나인 베르사유 궁전의 건축도면이다. 본 절에서는 베르사유 궁전의 건축도면을 [그림 3-18]과 같이 직선 객체로만 구성된 도면과 곡선 객체를 포함하고 있는 도면으로 나누어 이진화 성능을 평가하였다.



(a) 직선 객체로만 구성된 도면



(b) 곡선 객체를 포함하는 도면

[그림 3-18] 객체 형상에 따른 베르사유 궁전 건축도면 분류

검증 과정을 위해 선정한 총 20개의 도면들은 앞 절과 마찬가지로 소축척과 대축척 도면을 포함하고 있으며 다양한 형태의 배경 노이즈가 존재한다. 3.6.1절에서는 직선 객체로만 구성된 베르사유 궁전의 건축도면을 대상으로 분류 모델 성능 평가와 이진화 결과 품질 평가에 대해 설명한다. 이어지는 3.6.2절에서는 곡선 객체를 포함하는 베르사유 궁전의 건축도면의 이진화 결과를 살펴본다.

3.6.1 직선 객체로 구성된 베르사유 궁전 건축도면의 이진화

직선 객체로 구성된 베르사유 궁전 건축도면을 대상으로 수행한 분류 모델 성능 평가와 이진화 결과 품질 평가 결과는 [표 3-4]와 같다. 분류 모델 성능은 테스트 셋과 비슷하게 우수한 것으로 나타났다. 테스트 셋의 정밀도와 재현율의 조화평균인 F1-score 지표가 0.984였으며 직선 객체로 구성된 베르사유 궁전 건축도면의 F1-score 지표는 0.983이다. 따라서 본 모델에서 제안한 이진화 알고리즘의 분류 성능은 두 개의 데이터 셋에서 모두 안정적인 분류 성능을 달성한 것으로 보인다. 분류 모델 성능뿐만 아니라 이진화 결과 이미지의 품질 또한 테스트 셋만큼 우수한 것으로 나타났지만 DRD 지표는 테스트 셋의 결과가 다소 우수했다. 이는 테스트 셋 대비 직선 객체로 구성된 베르사유 궁전의 건축도면 이진화 결과에서 선형객체 인근 잔여 노이즈의 양이 많기 때문이다. 직선 객체로 구성된 베르사유 궁전 건축도면은 테스트 셋 대비 두께가 다양한 벽들이 직교하여 복잡한 형태의 실내 공간을 구성한다. 특히 두께가 얇은 내벽의 경우, 수직-수평 연속성 행렬과 같은 핵심 특징의 특징 값이 두드러지게 추출되지 않기 때문에 노이즈 감소에 한계가 있는 것으로 추정된다.

[표 3-4] 직선 객체로 구성된 베르사유 궁전 건축도면의 이진화 평가 결과

데이터	분류 모델 성능 평가				이진화 결과 품질 평가			
	정밀도	재현율	정확도	F1-score	PSNR	NRM	MPM	DRD
직선 객체로 구성된 베르사유 궁전 건축도면	0.998	0.969	0.969	0.983	15.72	0.020	0.002	1.678
테스트 셋	0.985	0.990	0.976	0.987	16.54	0.093	0.001	0.212

일반화 성능 검증을 목적으로 사용한 직선 객체로 구성된 베르사유 궁전의 건축도면 10개에 대한 이진화 적용 결과는 <부록 B>에서 확인할 수 있다. 직선 객체로 구성된 베르사유 궁전의 건축도면의 경우 테스트 셋 대비 선형객체의 두께 편차가 매우 크다. 뿐만 아니라 매우 두껍게 표현된 벽과 벽 사이에 매우 얇게 표현된 창문이 존재하는 형태로 건물이 표현되어 있어 두께 편차가 큰 실내 객체(벽, 창문)들이 상호 연결되어 있는 공간 구조를 갖는다. 따라서 이진화 과정에서 창문을 표현하는 얇은 선형객체가 끊어질 경우 이후 이어질 실내 공간정보 추출에 어려움이 존재하기 때문에 이진화 과정에서 이러한 얇은 선형객체의 정보 유실을 최소화 하는 것이 중요하다. 한 번도 학습하지 않은 직선 객체로 구성된 베르사유 궁전 건축도면을 대상으로 수행한 이진화, 다양한 두께의 선형 객체 정보 유실 없이 성공적으로 분류하는 것을 확인할 수 있었다.

3.6.2 곡선 객체를 포함하는 베르사유 궁전 건축도면의 이진화

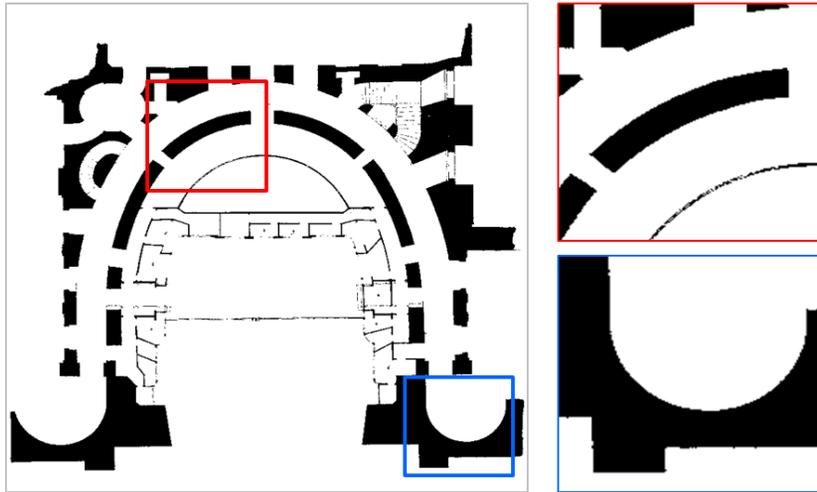
곡선 객체를 포함하는 베르사유 궁전 건축도면은 3.6.1절에서 사용된 도면과 다르게 둥근 호 모양의 벽을 포함하고 있다. 본 연구에서는 핵심 특징으로 나타난 수직-수평 연속성 행렬의 한계를 검증하기 위해 곡선 객체를 포함하는 베르사유 궁전 건축도면을 사용해 추가 검증을 수행하였다. 앞서 3.4.1절에서 본 연구에 사용된 24개 특징에 대한 특징 중요도 산출 결과에 대해 설명하였다. 산출 결과, 평균, 중앙 픽셀, 수직-수평 연속성 행렬이 핵심 특징으로 나타났다. 이 중 수직-수평 연속성 행렬은 유사한 픽셀값이 수직 혹은 수평으로 연속해서 나타나는 경우에만 추출되기 때문에 곡선으로 이루어진 객체 영역에서는 추출되지 않는다는 한계점이 존재한다. 하지만 해당 특징을 제외한 나머지 두 개의 핵심 특징(평균, 중앙픽셀)은 곡선으로 이루어진 객체에서도 추출할 수 있다. 이러한 세 개의 핵심 특징들 간 상호보완적인 관계로 인해 곡선 객체를 포함하는 도면에서도 성공적으로 이진화를 수행할 수 있다. 이에

본 절에서는 곡선 객체를 포함하는 도면을 대상으로 이진화를 수행한 결과를 살펴본다. 베르사유 궁전의 건축도면 데이터 셋으로부터 총 10개의 도면을 선별하였으며, 자세한 이진화 적용 결과는 <부록 C>에서 확인할 수 있다. 선별된 도면들은 원형 및 호와 같은 곡선으로 구성된 벽체가 다수 포함하고 있다. 3.6.1절과 마찬가지로 해당 도면들은 한 번도 학습한 적 없는 도면들이며 이진화 평가 결과는 [표 3-5]와 같다.

[표 3-5] 곡선 객체를 포함하는 베르사유 궁전 건축도면의 이진화 평가 결과

데이터	분류 모델 성능 평가				이진화 결과 품질 평가			
	정밀도	재현율	정확도	F1-score	PSNR	NRM	MPM	DRD
곡선 객체를 포함하는 베르사유 궁전 건축도면	0.999	0.969	0.971	0.983	15.69	0.020	0.002	0.242
테스트 셋	0.985	0.990	0.976	0.987	16.54	0.093	0.001	0.212

[표 3-5]에 나타난 바와 같이, 곡선 객체를 포함하는 건축도면을 대상으로 한 이진화 결과의 F1-score는 0.983으로 매우 우수한 성능을 보였다. 이진화 결과 이미지의 품질 평가 지표 중 신호 대비 잡음 비를 나타내는 PSNR 지표는 테스트 셋보다 우수한 것으로 나타났다.



[그림 3-19] 다양한 두께의 곡선 객체 이진화 결과

[그림 3-19]는 다양한 두께의 곡선 객체를 포함하는 도면의 이진화 결과이다. 두꺼운 벽체의 경우 곡선 형상이 끊기는 부분 없이 연결되는 결과를 보였다. 얇은 벽체의 경우 전체적인 곡선 형상은 유지하였으나, 일부 픽셀이 유실되는 한계를 보였다.

4. 결론

건축도면 이미지로부터 실내 공간정보를 자동 추출하는 기존 도면 해석 연구들은 배경과 실내 객체가 명확히 구분되는 깨끗한 전자도면을 사용하였다. 하지만 역사적 건축물이나 건축 연도가 오래된 건물들의 도면은 다양한 필기구를 사용하여 손으로 작성되었으며 종이의 구겨짐 및 노후화로 심각한 노이즈가 존재한다. 이에 본 연구는 기존 도면 해석 연구를 핸드드로잉 건축도면으로 확장하기 위해 노이즈가 심한 건축도면 이미지를 대상으로 실내 객체와 배경을 구분하는 학습 기반 이진화 연구를 진행하였다.

이진화 관련 선행 연구들은 손 글씨로 작성된 문서 데이터 셋을 대상으로 주로 진행되어 왔으며 픽셀값의 통계적 특징이나 문자 픽셀의 두께 등의 특징을 활용하였다. 하지만 기존 선행 연구의 방법론을 건축도면에 적용할 때는 다음과 같은 한계점이 존재한다. 첫째, 문자 두께나 글자 경계와 같이 문서 데이터 셋을 대상으로 추출한 특징은 다양한 두께와 길이의 선으로 구성된 건축도면에 적용할 수 없다. 둘째, 건축도면 내 실내 객체 요소의 연결성을 고려한 특징 추출이 불가능하다. 이러한 한계점을 극복하고자 본 연구는 기존 이진화 선행 연구에서 활용한 픽셀값의 통계적 특징과 더불어 건축도면의 고유 특성인 실내 공간 객체 간 연결성을 학습을 위한 특징으로 사용하였다. 모델 학습에 사용된 데이터 셋은 일제시기 건축도면이며, 총 50 개의 이미지 중 40 개를 학습 데이터로 사용하였고 나머지 10 개는 테스트 셋으로 사용하였다.

본 연구의 이진화 단계는 두 단계에 거쳐 수행되었다. 첫 번째 단계는 가우시안 혼합 모델을 사용해 노이즈와 선형객체 픽셀을 두 개의 군집으로 분리하고, 노이즈 군집에 해당하는 픽셀들의 픽셀값을 변경하여 배경 추정 이미지를 생성하는 단계이다. 이렇게 생성된 배경

추정 이미지와 원본 이미지 간 일대일 픽셀값 차이를 기반으로 전체적인 배경 노이즈를 제거하였다. 가우시안 혼합 모델을 통해 선형객체와 분명한 픽셀값 차이가 존재하는 노이즈는 대부분 감소하는 것을 확인할 수 있었다. 하지만 면적이 작고 선형객체와 유사한 픽셀값을 갖는 노이즈는 감소하지 못하는 한계가 존재하여, 이러한 노이즈는 두 번째 단계인 랜덤포레스트 모델로 감소하였다.

랜덤포레스트 모델을 활용한 이진화 단계에서는 총 24 개의 특징을 추출하여 선형객체와 배경을 구분하고자 하였다. 제안한 방법으로 특징 24 개를 모두 사용하였을 때 이진화 성능의 정확도는 0.976였으며, 상대적 중요도가 높은 핵심 특징 3개만 사용하여도 정확도 0.97이 넘는 우수한 성능을 나타냈다. 다양한 축척의 도면으로 구성된 테스트 셋과 베르사유 궁전 건축도면에 제안한 방법을 적용한 결과, 픽셀값 관련 통계와 연결성 정보를 함께 활용함으로써 면적이 작고 두께가 얇은 실내 객체 요소가 소실 되지 않는 것을 확인할 수 있었다. 분류 성능에 핵심적인 특징 3개 중 수직-수평 연속성 행렬 특성 상 곡선을 포함하는 도면의 이진화 성능은 낮을 것으로 예상했으나, 곡선 객체를 포함하는 베르사유 궁전 건축도면을 대상으로 이진화 성능을 검증한 결과, 이진화 결과만큼 우수한 결과를 나타냈다. 이는 단일 특징을 기반으로 픽셀을 분류하기 보다 핵심 특징 3개(평균, 중앙 픽셀, 수직-수평 연속성 행렬)를 종합적으로 사용하여 이진화를 수행하기 때문으로 사료된다.

핸드드로잉 건축도면을 대상으로 이진화를 수행할 때에 가장 중요한 것은 실내 공간을 구성하는 선형객체 정보의 유실을 최소화 하는 것이다. 선형객체 정보의 유실화는 핸드드로잉 건축도면의 이진화 이후 후속될 실내공간 정보 추출 성능에 큰 영향을 미치기 때문이다. 본 연구의 이진화 결과 이미지를 살펴본 결과, 정보 유실 가능성이 높은 객체는 쪽문, 얇은 내벽, 창문과 같이 대체로 두께가 얇고 면적이 작은

선형객체들이었다. 이와 같은 선형객체들은 두께가 얇기 때문에 픽셀값만 사용하여 이진화를 수행하기 보다, 인근 객체들로부터 추출할 수 있는 추가 정보를 함께 활용해야 한다. 본 연구의 경우 수직-수평 연속성 행렬이라는 특징을 사용하여 유사한 픽셀값을 가진 인근 객체들과의 연결성 정보를 추가적으로 활용하였다.

본 연구에서 제안한 랜덤포레스트 기반 이진화 모델은 전체적으로 우수한 이진화 성능을 보였으나, 노이즈의 픽셀값이 선형객체와 매우 유사한 경우 이진화 후에도 여전히 남아있는 한계가 있다. 노이즈의 면적 자체는 작지만 선형객체의 경계에 존재하므로 실내 공간정보 자동 추출 시 객체 인식에 영향을 야기할 수 있다. 따라서 선형객체에 인접한 작은 노이즈를 추가적으로 제거할 수 있는 방법에 대한 연구가 필요하다.

본 연구는 문서 데이터 셋을 대상으로 주로 진행되어 오던 이진화 연구를 도면 데이터 셋으로 확장하고, 도면 데이터 셋에 나타나는 고유한 특징을 활용하여 이진화를 수행하였다는 점에서 의의가 있다. 또한, 본 연구에서 제안한 이진화 기법을 통해 실내 공간정보 자동 추출 연구의 활용처를 핸드드로잉 건축도면으로 확장하는 기반을 마련했다는 것에 의미가 있다.

참 고 문 헌

- 김성용. (2019). 스타일 트랜스퍼를 활용한 평면도 해석 프로세스의 범용성 강화 (Doctoral dissertation, 서울대학교 대학원).
- 김지엽. (2020). 규모가 크고 복잡한 건물의 지하 정보 추출 및 벡터화를 통한 도면 해석 연구 (Doctoral dissertation, 서울대학교 대학원).
- Alham, N. K. (2011). Parallelizing support vector machines for scalable image annotation (Doctoral dissertation, Brunel University School of Engineering and Design PhD Theses).
- Ahmadi, E., Azimifar, Z., Shams, M., Famouri, M., & Shafiee, M. J. (2015). Document image binarization using a discriminative structural classifier. *Pattern recognition letters*, 63, 36–42.
- Ahmed, S., Liwicki, M., Weber, M., & Dengel, A. (2011, September). Improved automatic analysis of architectural floor plans. In 2011 International Conference on Document Analysis and Recognition (pp. 864–869). IEEE.
- Bharathi, V. C., & Geetha, M. K. (2013). Segregated handwritten character recognition using GLCM features. *International Journal of Computer Applications*, 84(2).
- Bouveyron, C., & Brunet-Saumard, C. (2014). Model-based clustering of high-dimensional data: A review. *Computational Statistics & Data Analysis*, 71, 52–78.
- Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123–140.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- C. Wolf, J.-M. Jolion, Extraction and recognition of artificial text in multimedia documents. *Pattern Anal. Appl.* 6(4), 309–326 (2003). <https://doi.org/10.1007/s10044-003-0197-7>
- Cabrera Vargas, D. A. (2018). Wall extraction and room detection for multi-unit architectural floor plans (Doctoral dissertation).
- Caruana, R., Karampatziakis, N., & Yessenalina, A. (2008). An empirical

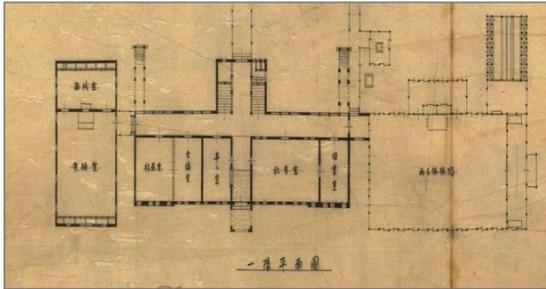
- evaluation of supervised learning in high dimensions. Proceedings of the International Conference on Machine Learning (ICML 2008), 96–103.
- Chen, X., Lin, L., & Gao, Y. (2016). Parallel nonparametric binarization for degraded document images. *Neurocomputing*, 189, 43–52.
- Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM Computing Surveys (CSUR)*, 27(3), 326–327.
- Dudoit, S., Fridlyand, J., & Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association*, 97, 77–87.
- Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. *Pattern recognition*, 39(3), 317–327.
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). “Gene selection for cancer classification using support vector machines.” *Machine Learning*, Vol. 46, pp. 389–422.
- Lu, H., Kot, A. C., & Shi, Y. Q. (2004). Distance–reciprocal distortion measure for binary document images. *IEEE Signal Processing Letters*, 11(2), 228–231.
- Jain, S. (2013). Brain cancer classification using GLCM based feature extraction in artificial neural network. *International Journal of Computer Science & Engineering Technology*, 4(7), 966–970.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. New York, NY: Springer.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. New York, NY: Springer.
- Jana, P., Ghosh, S., Bera, S. K., & Sarkar, R. (2017, December). Handwritten document image binarization: An adaptive K-means based approach. In *2017 IEEE Calcutta Conference (CALCON)* (pp. 226–230). IEEE.
- Jia, F., Shi, C., He, K., Wang, C., & Xiao, B. (2018). Degraded document image binarization using structural symmetry of strokes. *Pattern Recognition*, 74, 225–240.
- Jindal, H., Kumar, M., Tomar, A., & Malik, A. (2021, April). Degraded Document Image Binarization using Novel Background Estimation

- Technique. In 2021 6th International Conference for Convergence in Technology (I2CT) (pp. 1–8). IEEE.
- Kasmin, F., Othman, Z., Sakinah, S., & Ahmad, S. (2019). Pixel-wise classification using support vector machine for binarization of degraded historical document image. *vol, 15*, 1329–1336.
- Kim, H., Kim, S., & Yu, K. (2021). Automatic Extraction of Indoor Spatial Information from Floor Plan Image: A Patch-Based Deep Learning Methodology Application on Large-Scale Complex Buildings. *ISPRS International Journal of Geo-Information*, 10(12), 828.
- Kim, S., Park, S., Kim, H., & Yu, K. (2021). Deep Floor Plan Analysis for Complicated Drawings Based on Style Transfer. *Journal of Computing in Civil Engineering*, 35(2), 04020066.
- Kwon, S., Seo, I. W., Noh, H., & Kim, B. (2022). Hyperspectral retrievals of suspended sediment using cluster-based machine learning regression in shallow waters. *Science of The Total Environment*, 833, 155168.
- Liu, C., Wu, J., Kohli, P., & Furukawa, Y. (2017). Raster-to-vector: Revisiting floorplan transformation. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2195–2203).
- Lu, S., Su, B., & Tan, C. L. (2010). Document image binarization using background estimation and stroke edges. *International Journal on Document Analysis and Recognition (IJ DAR)*, 13(4), 303–314.
- Mitianoudis, N., & Papamarkos, N. (2015). Document image binarization using local features and Gaussian mixture modeling. *Image and Vision Computing*, 38, 33–51.
- Niblack, W. (1985). *An introduction to digital image processing*. Strandberg Publishing Company.
- Nina, O., Morse, B., & Barrett, W. (2011, January). A recursive Otsu thresholding method for scanned document binarization. In *2011 IEEE Workshop on Applications of Computer Vision (WACV)* (pp. 307–314). IEEE.
- Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621.
- Sauvola, J., & Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern recognition*, 33(2), 225–236.

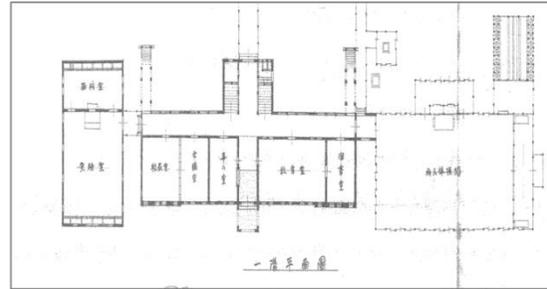
- Sehad, A., Chibani, Y., Cheriet, M., & Yaddaden, Y. (2013, September). Ancient degraded document image binarization based on texture features. In 2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA) (pp. 189–193). IEEE.
- Su, B., Lu, S., & Tan, C. L. (2010, June). Binarization of historical document images using the local maximum and minimum. In Proceedings of the 9th IAPR International Workshop on Document Analysis Systems (pp. 159–166).
- Su, B., Lu, S., & Tan, C. L. (2013). Robust document image binarization technique for degraded document images. *IEEE transactions on image processing*, 22(4), 1408–1417.
- Tomasi, C., & Manduchi, R. (1998, January). Bilateral filtering for gray and color images. In Sixth international conference on computer vision (IEEE Cat. No. 98CH36271) (pp. 839–846). IEEE.
- Xiong, W., Jia, X., Xu, J., Xiong, Z., Liu, M., & Wang, J. (2018, August). Historical document image binarization using background estimation and energy minimization. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 3716–3721). IEEE.
- Westphal, F., Grahn, H., & Lavesson, N. (2018). Efficient document image binarization using heterogeneous computing and parameter tuning. *International Journal on Document Analysis and Recognition (IJ DAR)*, 21(1), 41–58.
- Xiong, W., Zhou, L., Yue, L., Li, L., & Wang, S. (2021). An enhanced binarization framework for degraded historical document images. *EURASIP Journal on Image and Video Processing*, 2021(1), 1–24.

부 록

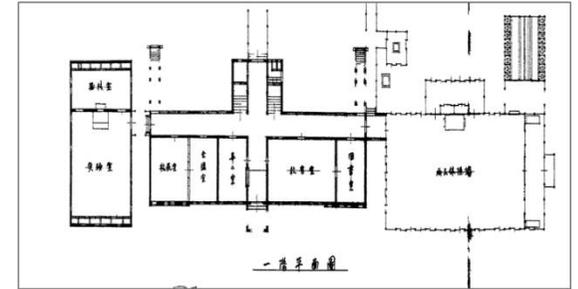
<부록 A> 테스트 셋 이진화 결과



(a) 원본 이미지

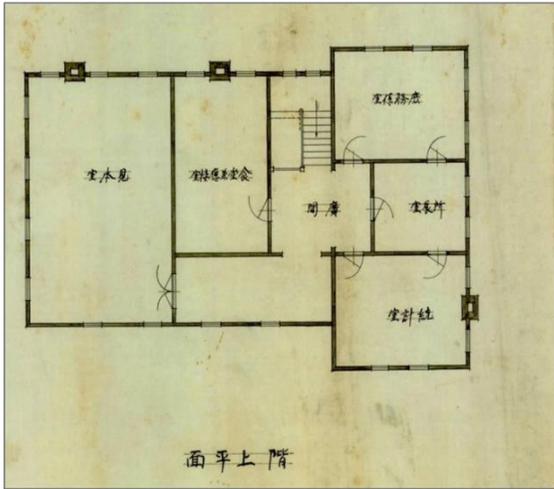


(b) 배경 제거 이미지

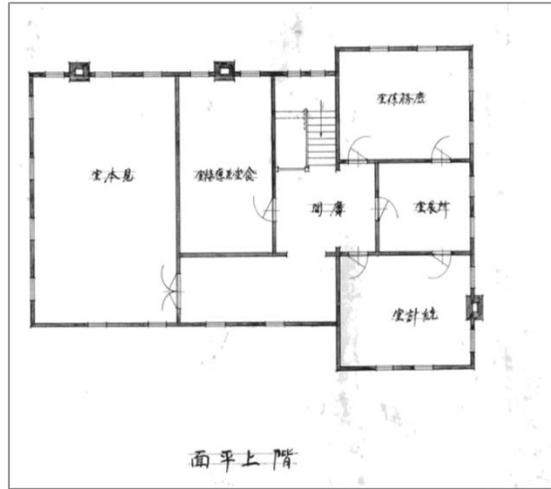


(c) 이진화 결과 이미지

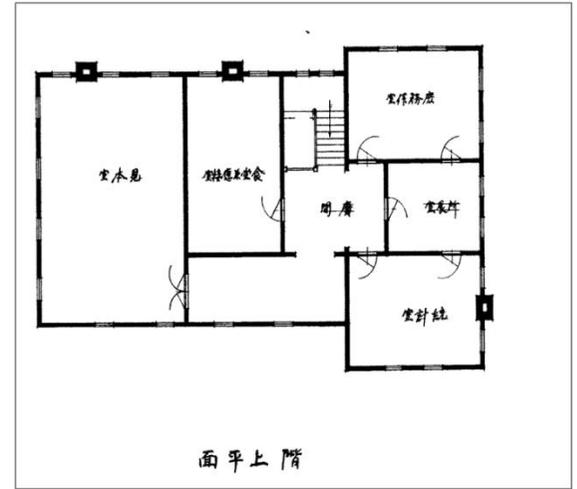
[그림 A-1] 대한의원학교 평면도 (도면 14번) 이진화 결과



(a) 원본 이미지

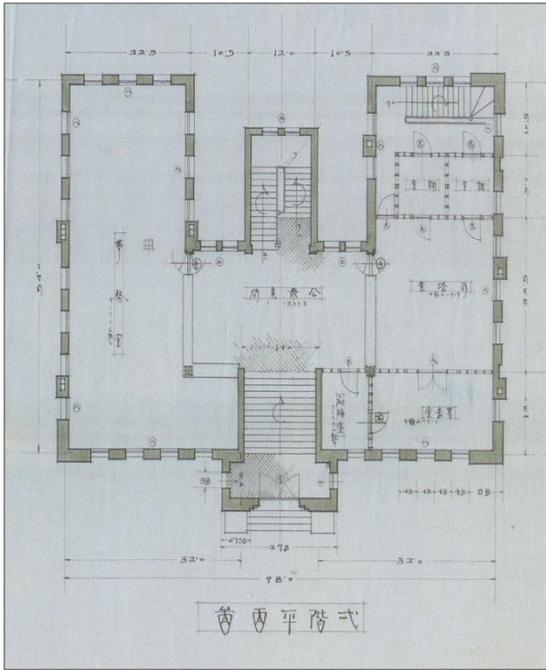


(b) 배경 제거 이미지

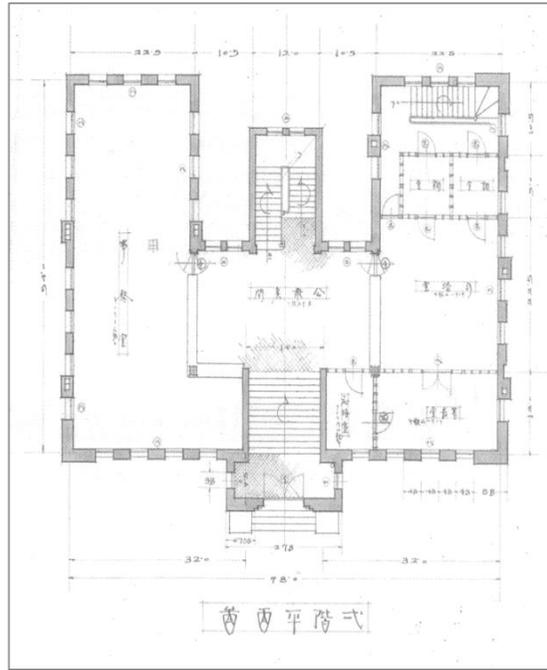


(c) 이진화 결과 이미지

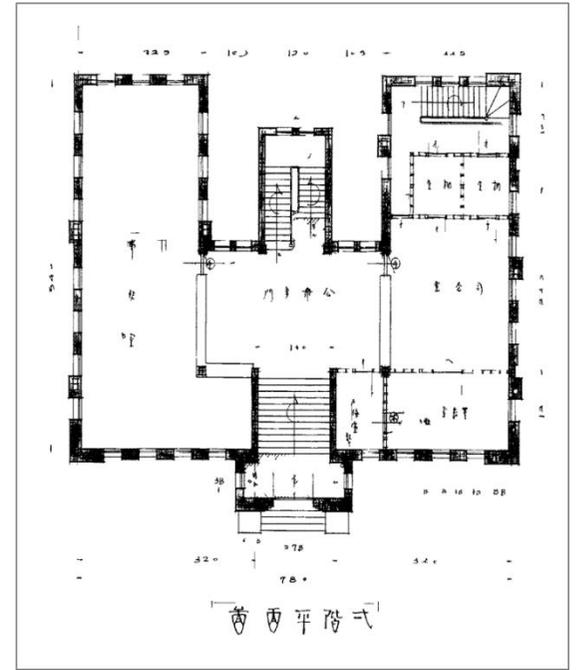
[그림 A-2] 경성세관출장소청사 평면도 (도면 41번) 이진화 결과



(a) 원본 이미지

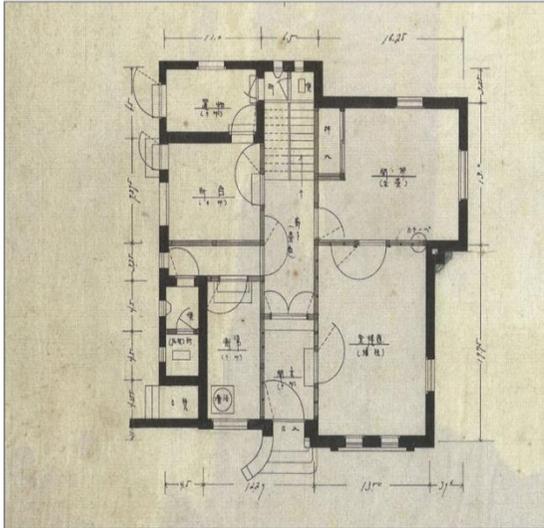


(b) 배경 제거 이미지

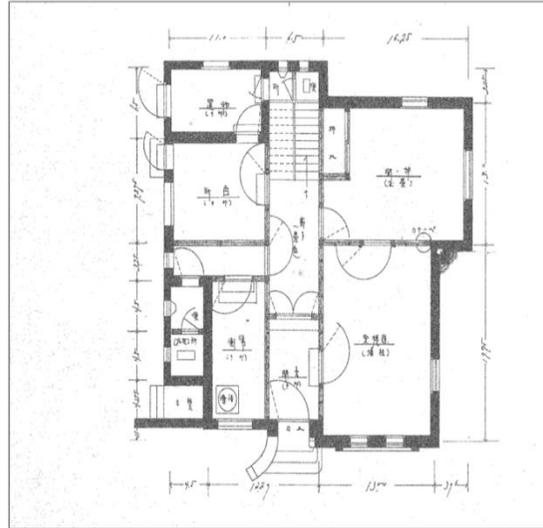


(c) 이진화 결과 이미지

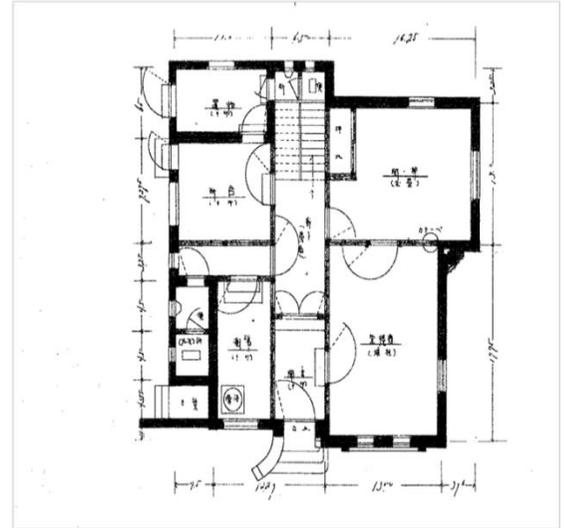
[그림 A-3] 부산경찰서청사 설계도 (도면 43번) 이진화 결과



(a) 원본 이미지

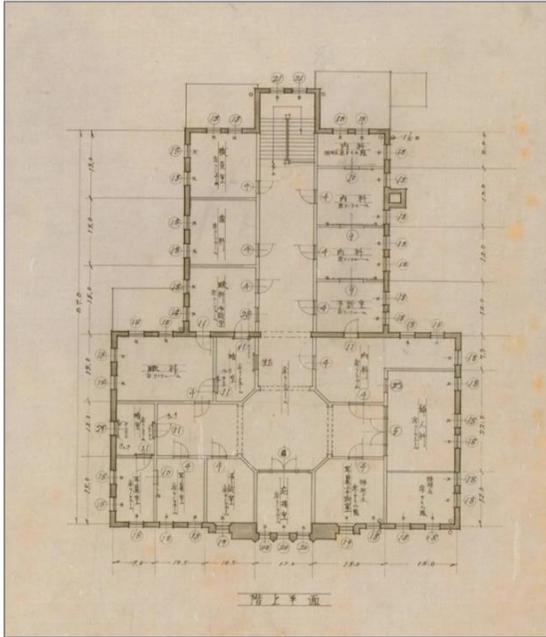


(b) 배경 제거 이미지

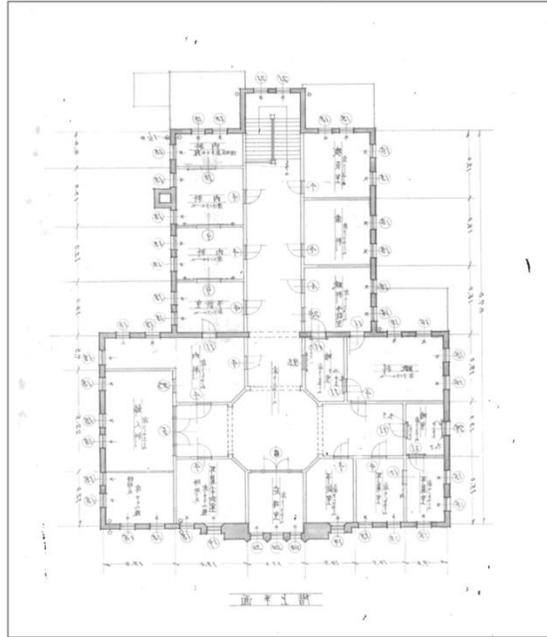


(c) 이진화 결과 이미지

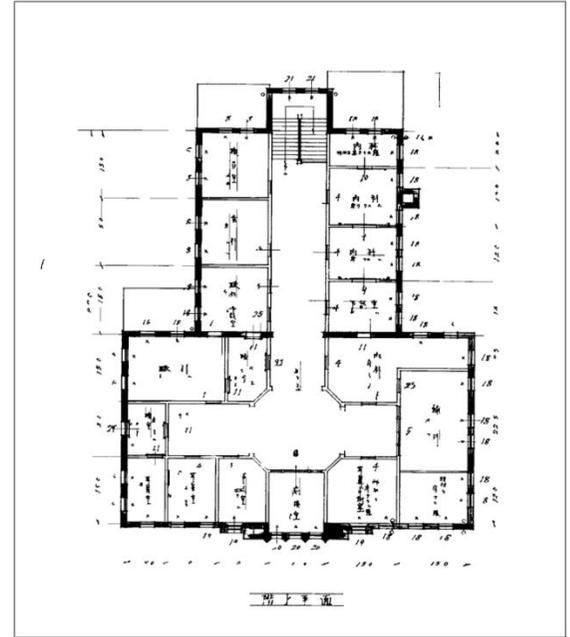
[그림 A-4] 갑오관사 설계도 (도면 48번) 이진화 결과



(a) 원본 이미지

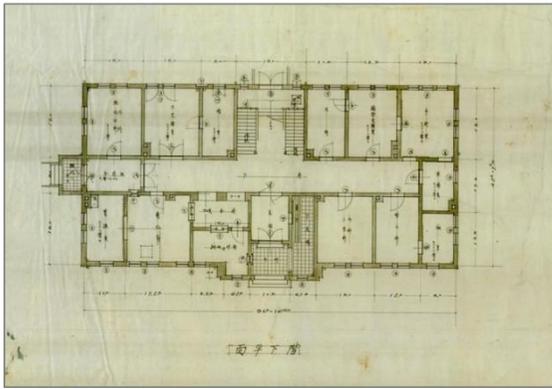


(b) 배경 제거 이미지

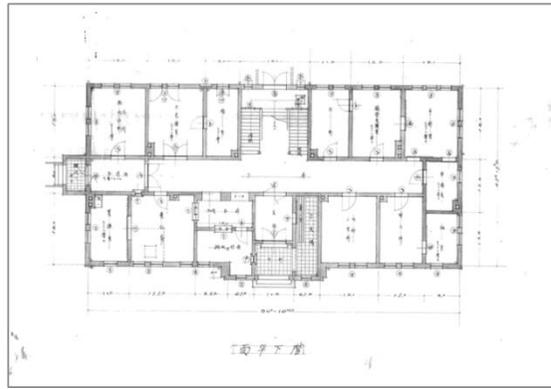


(c) 이진화 결과 이미지

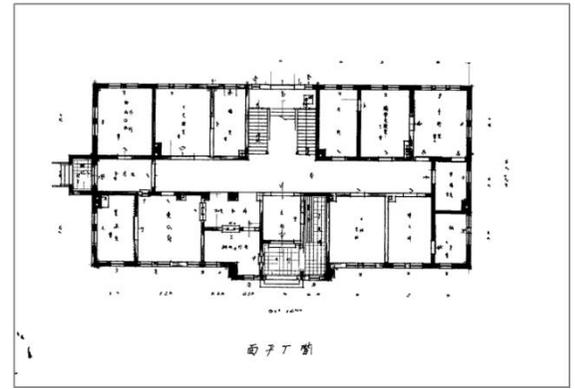
[그림 A-5] 경성의학전문학교 부속병원 외래진료소 평면도 (도면 38번) 이진화 결과



(a) 원본 이미지



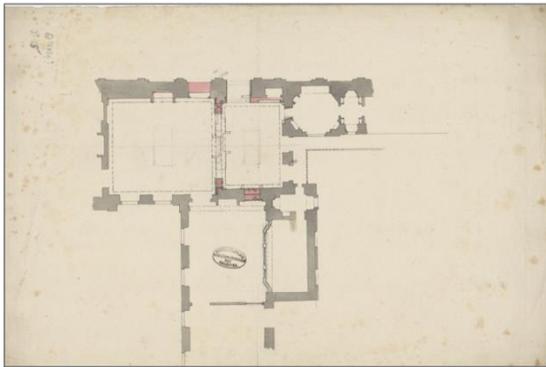
(b) 배경 제거 이미지



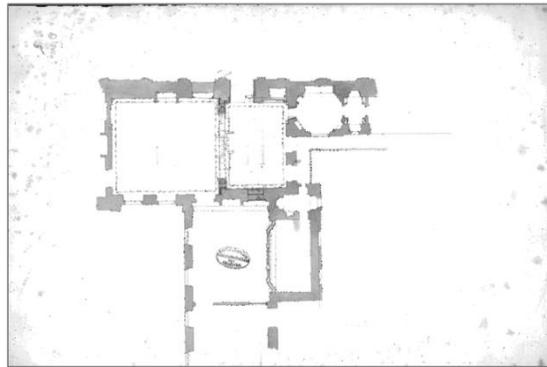
(c) 이진화 결과 이미지

[그림 A-6] 해주자혜의원 본관 설계도 (도면 34번) 이진화 결과

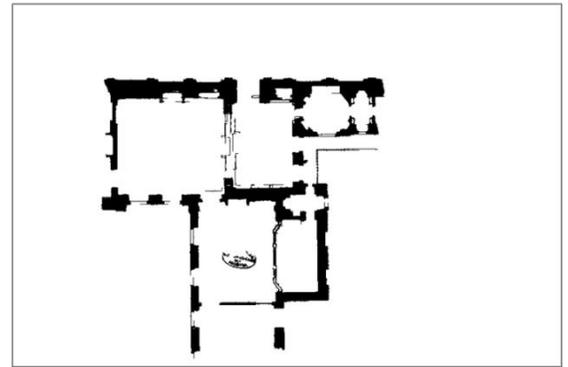
<부록 B> 직선 객체로 구성된 베르사유 궁전 건축도면의 이진화 결과



(a) 원본 이미지

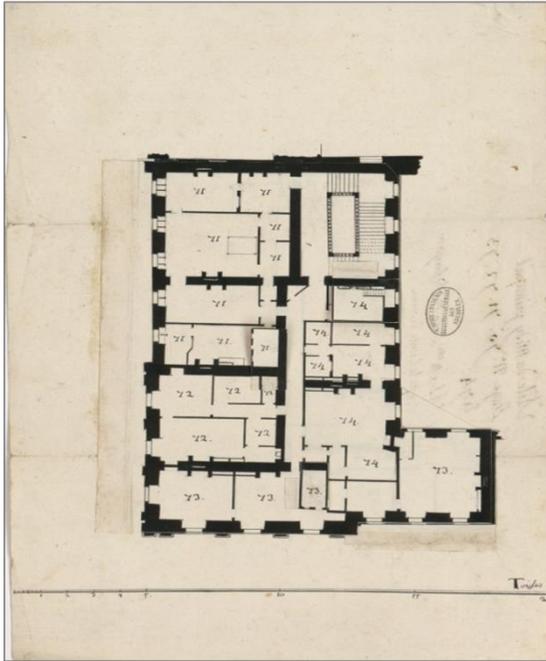


(b) 배경 제거 이미지

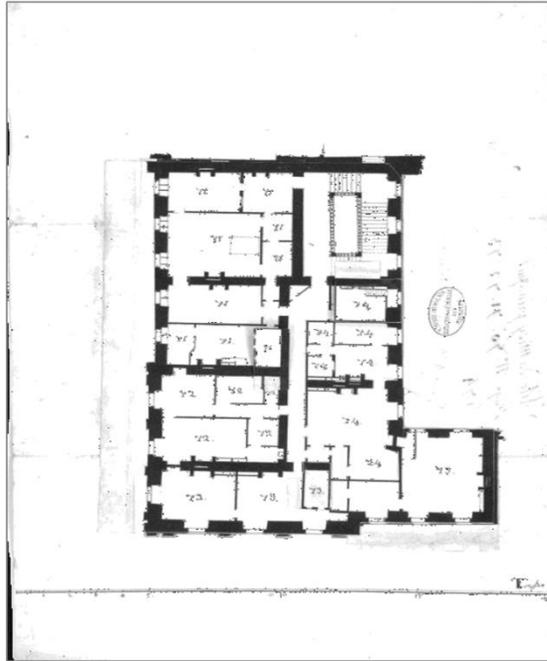


(c) 이진화 결과 이미지

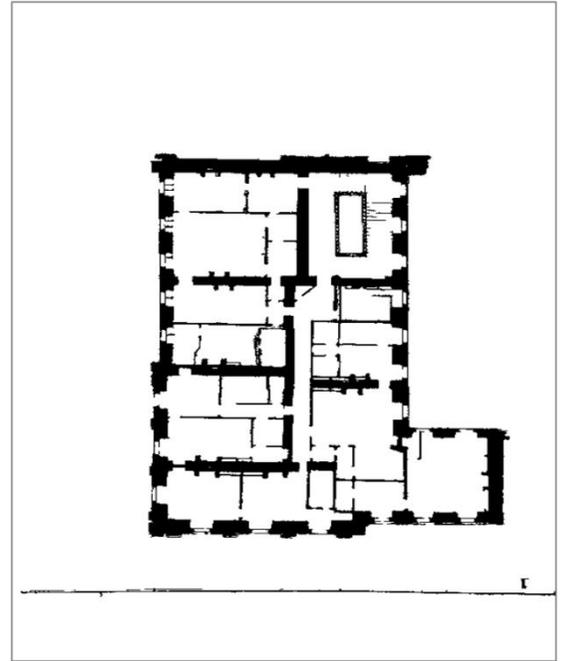
[그림 B-1] 직선 객체로 구성된 베르사유 궁전 건축도면 1번 이진화 결과



(a) 원본 이미지

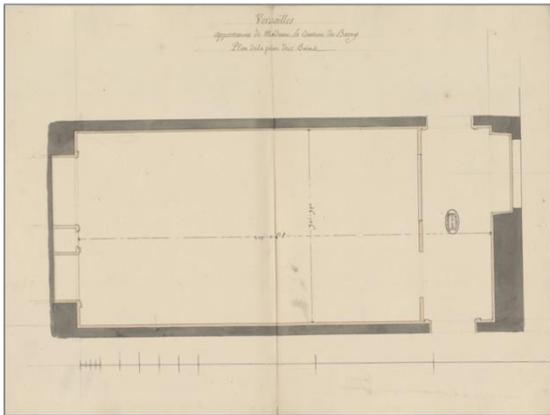


(b) 배경 제거 이미지

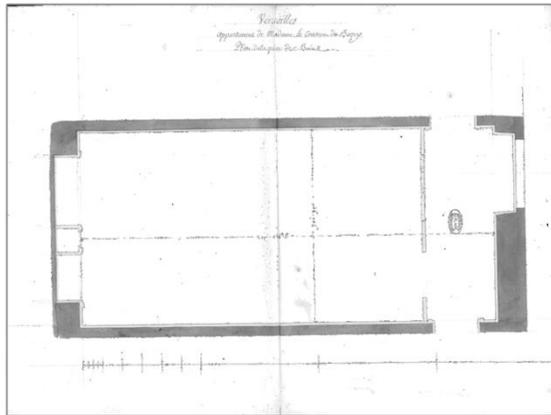


(c) 이진화 결과 이미지

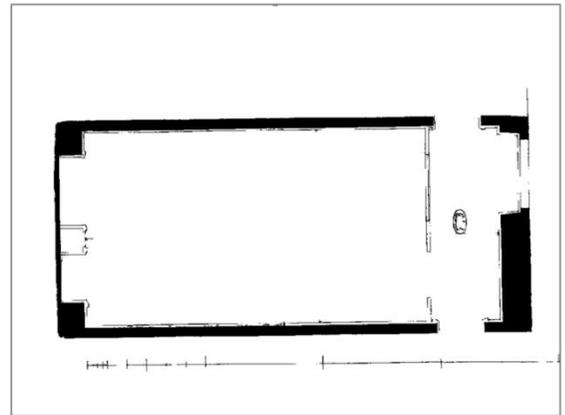
[그림 B-2] 직선 객체로 구성된 베르사유 궁전 건축도면 2번 이진화 결과



(a) 원본 이미지

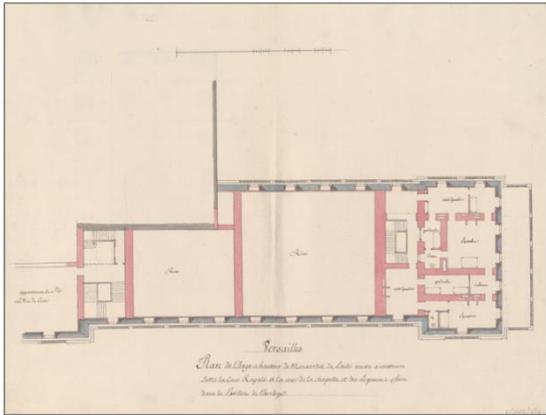


(b) 배경 제거 이미지

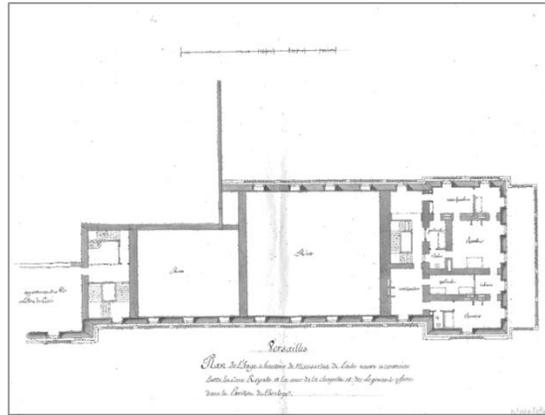


(c) 이진화 결과 이미지

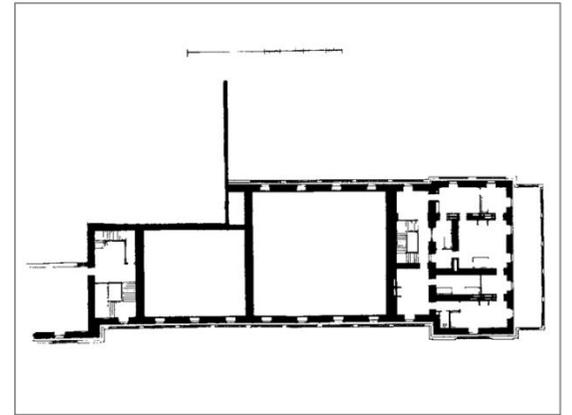
[그림 B-3] 직선 객체로 구성된 베르사유 궁전 건축도면 3번 이진화 결과



(a) 원본 이미지

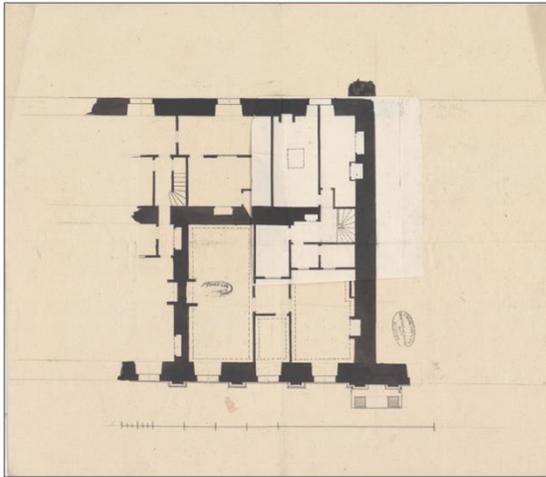


(b) 배경 제거 이미지

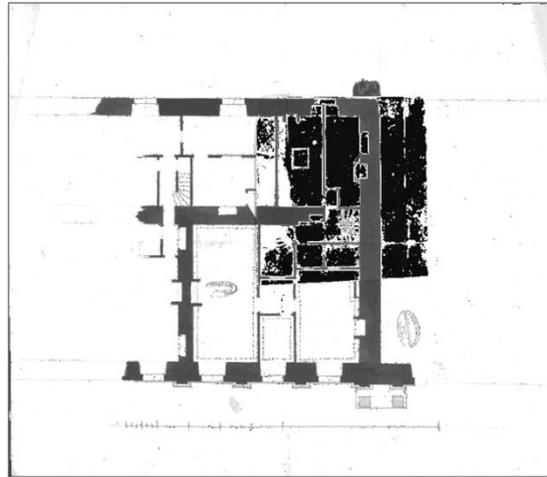


(c) 이진화 결과 이미지

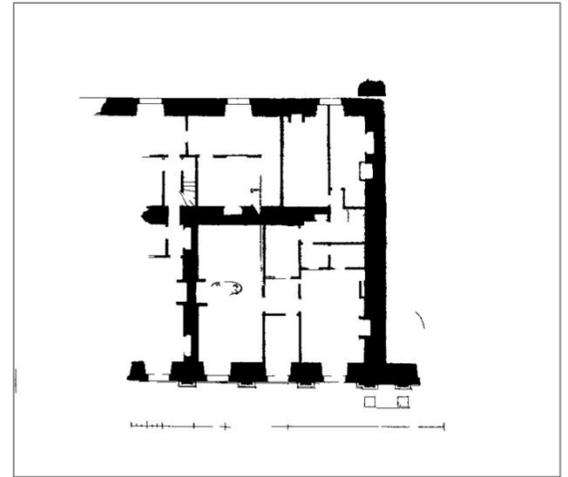
[그림 B-4] 직선 객체로 구성된 베르사유 궁전 건축도면 4번 이진화 결과



(a) 원본 이미지

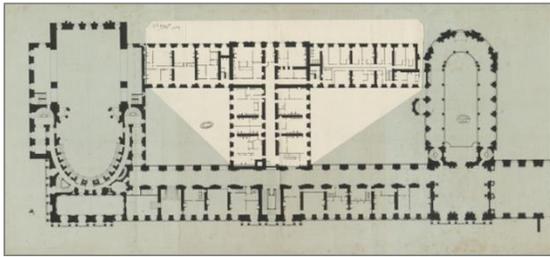


(b) 배경 제거 이미지

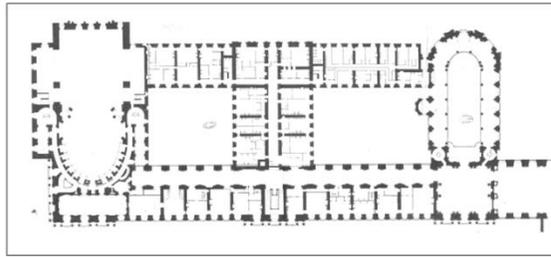


(c) 이진화 결과 이미지

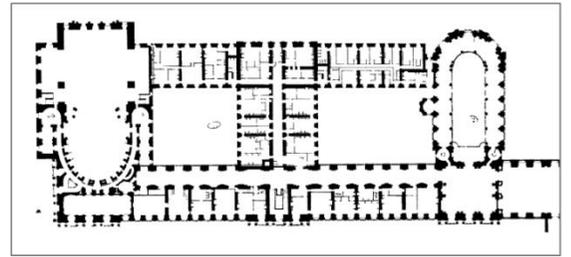
[그림 B-5] 직선 객체로 구성된 베르사유 궁전 건축도면 5번 이진화 결과



(a) 원본 이미지

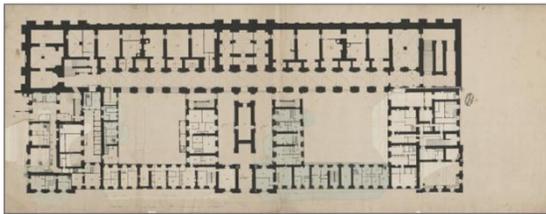


(b) 배경 제거 이미지

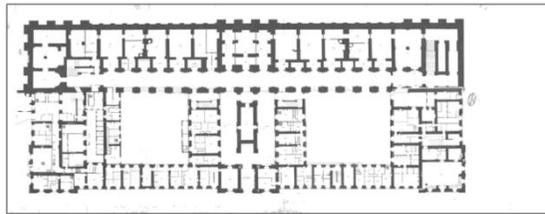


(c) 이진화 결과 이미지

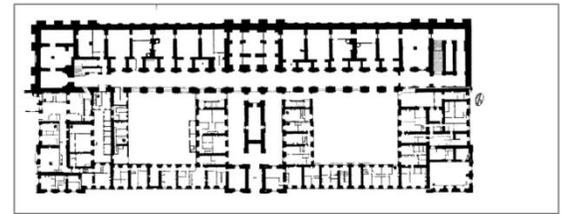
[그림 B-6] 직선 객체로 구성된 베르사유 궁전 건축도면 6번 이진화 결과



(a) 원본 이미지

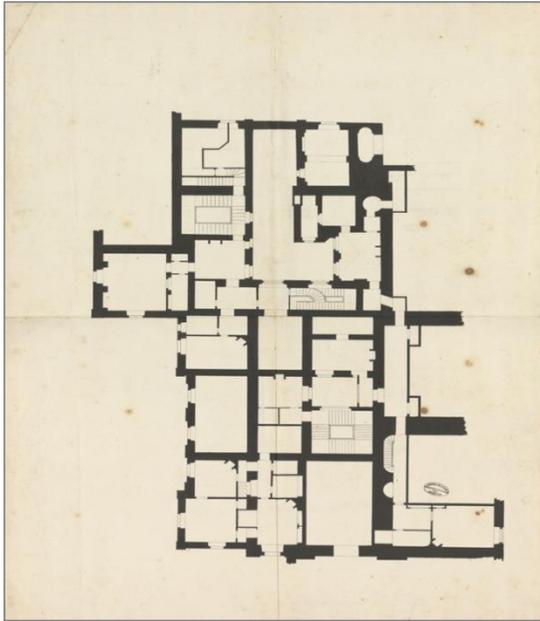


(b) 배경 제거 이미지

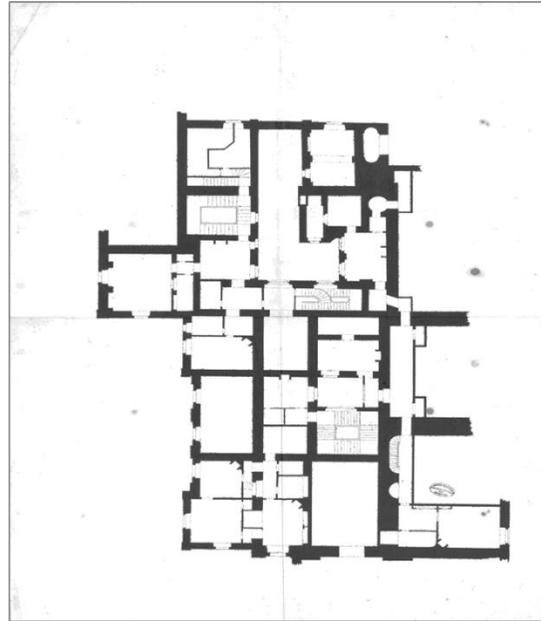


(c) 이진화 결과 이미지

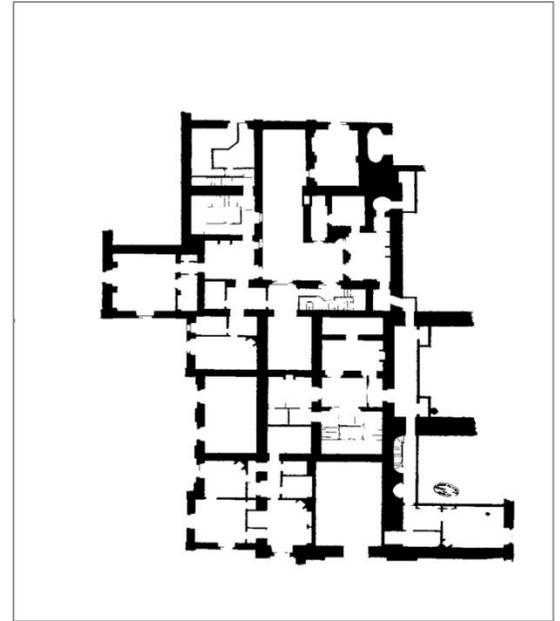
[그림 B-7] 직선 객체로 구성된 베르사유 궁전 건축도면 7번 이진화 결과



(a) 원본 이미지

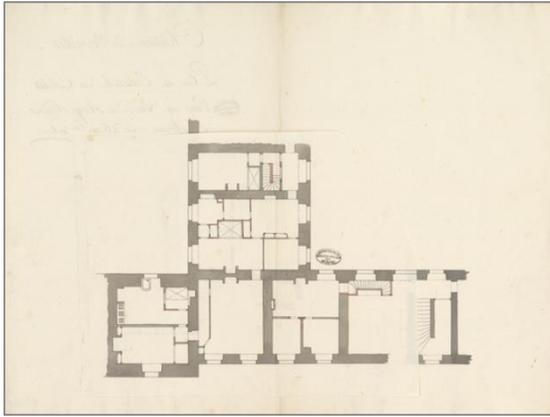


(b) 배경 제거 이미지

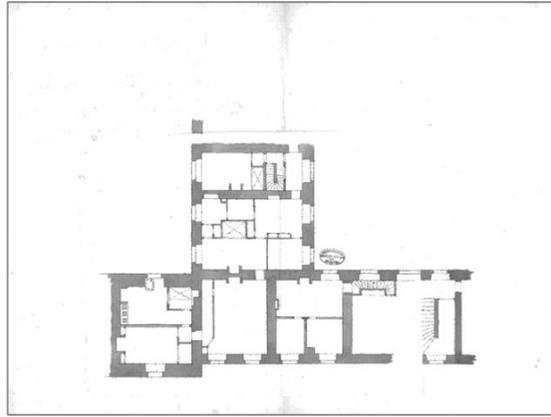


(c) 이진화 결과 이미지

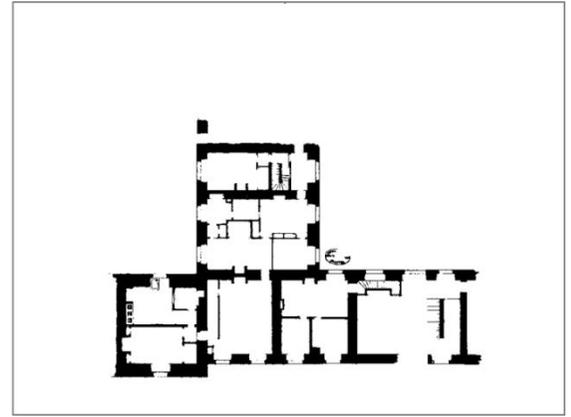
[그림 B-8] 직선 객체로 구성된 베르사유 궁전 건축도면 8번 이진화 결과



(a) 원본 이미지

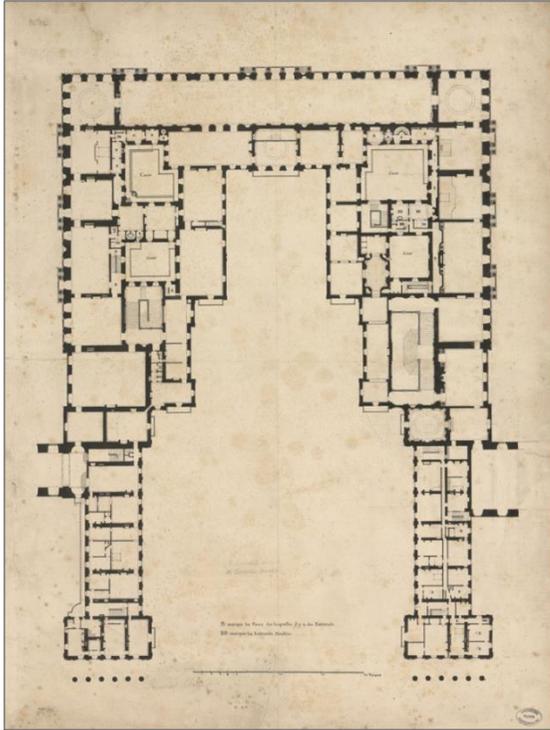


(b) 배경 제거 이미지

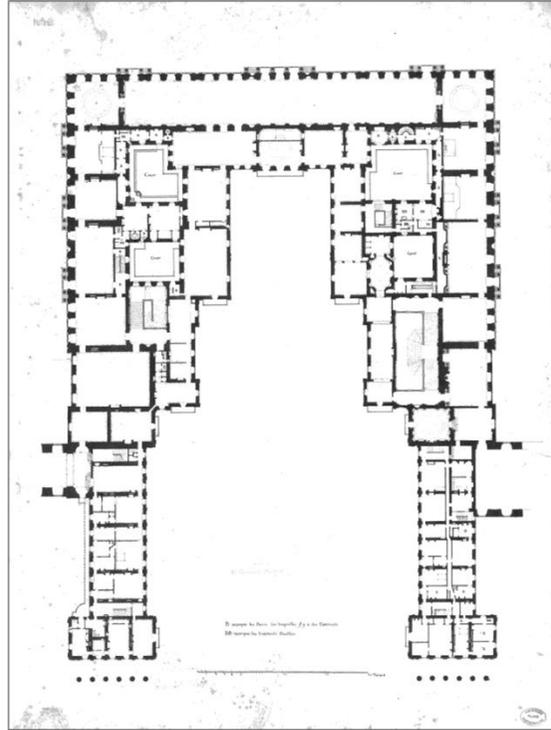


(c) 이진화 결과 이미지

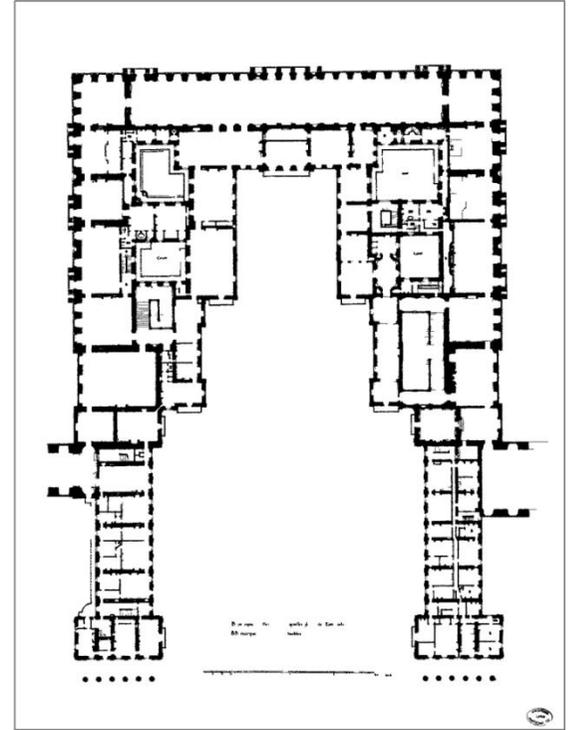
[그림 B-9] 직선 객체로 구성된 베르사유 궁전 건축도면 9번 이진화 결과



(a) 원본 이미지



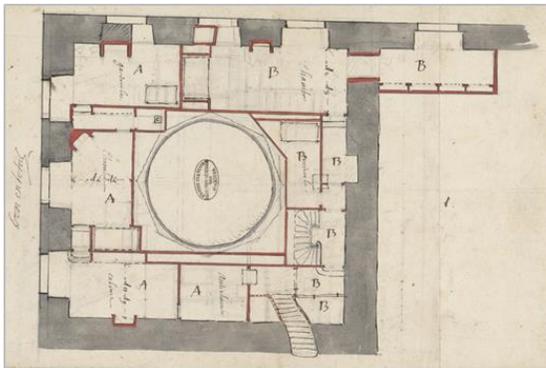
(b) 배경 제거 이미지



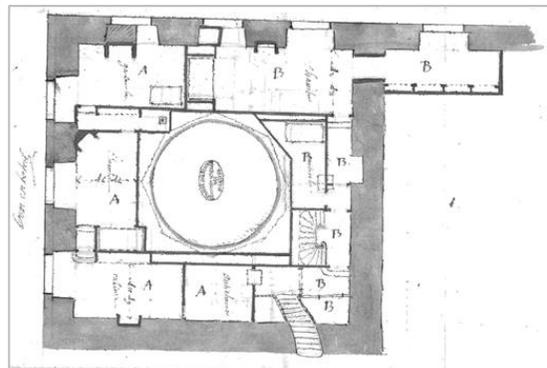
(c) 이진화 결과 이미지

[그림 B-10] 직선 객체로 구성된 베르사유 궁전 건축도면 10번 이진화 결과

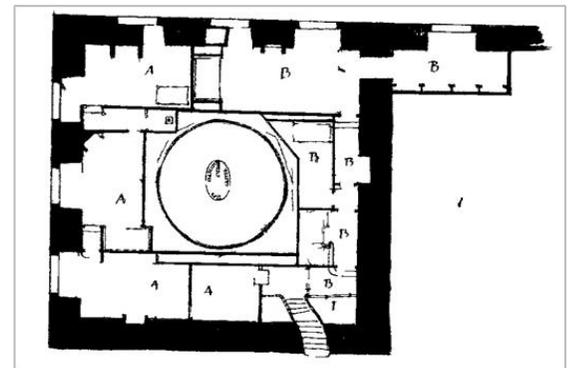
<부록 C> 곡선 객체를 포함하는 베르사유 궁전 건축도면의 이진화 결과



(a) 원본 이미지

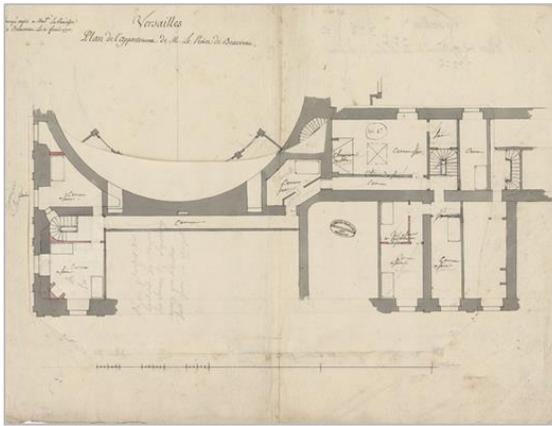


(b) 배경 제거 이미지

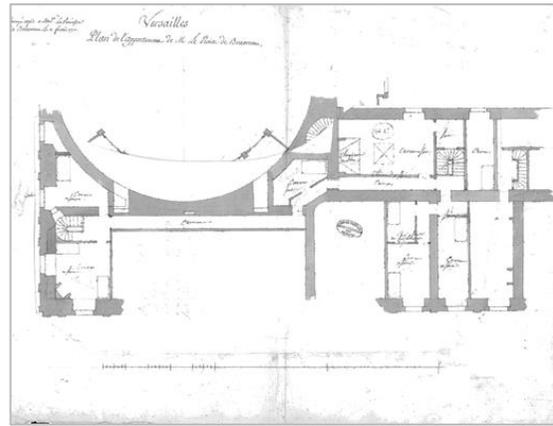


(c) 이진화 결과 이미지

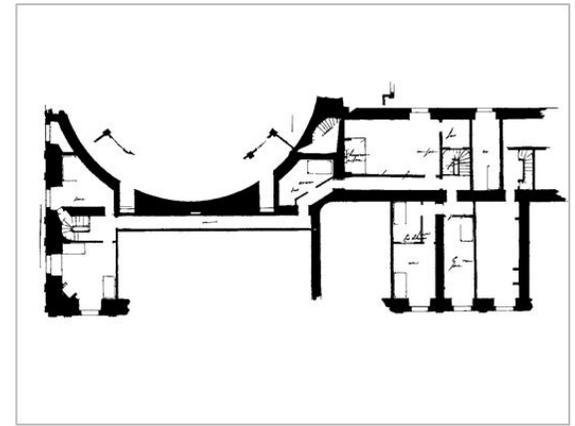
[그림 C-1] 곡선 객체를 포함하는 베르사유 궁전 건축도면 1번 이진화 결과



(a) 원본 이미지

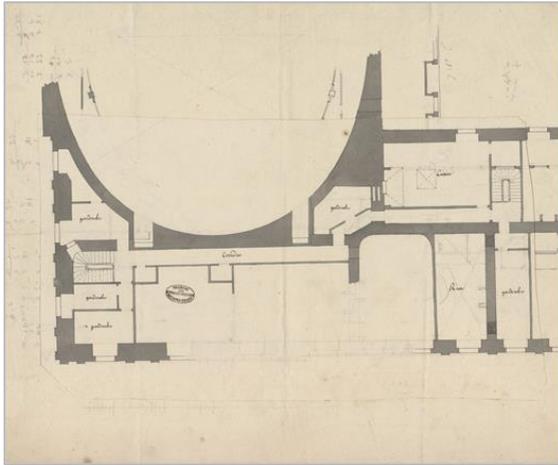


(b) 배경 제거 이미지

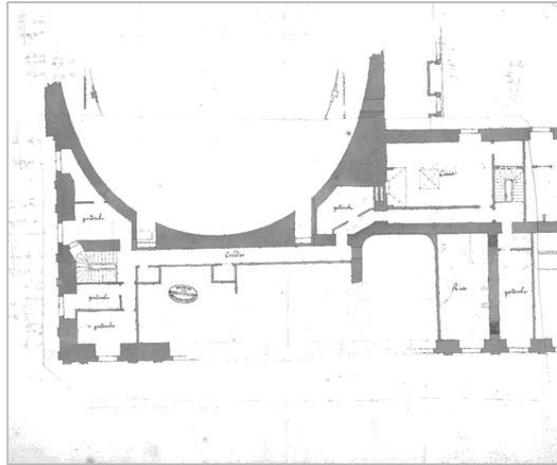


(c) 이진화 결과 이미지

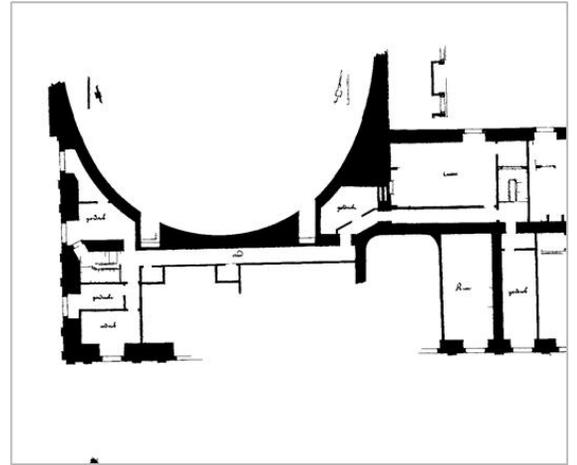
[그림 C-2] 곡선 객체를 포함하는 베르사유 궁전 건축도면 2번 이진화 결과



(a) 원본 이미지

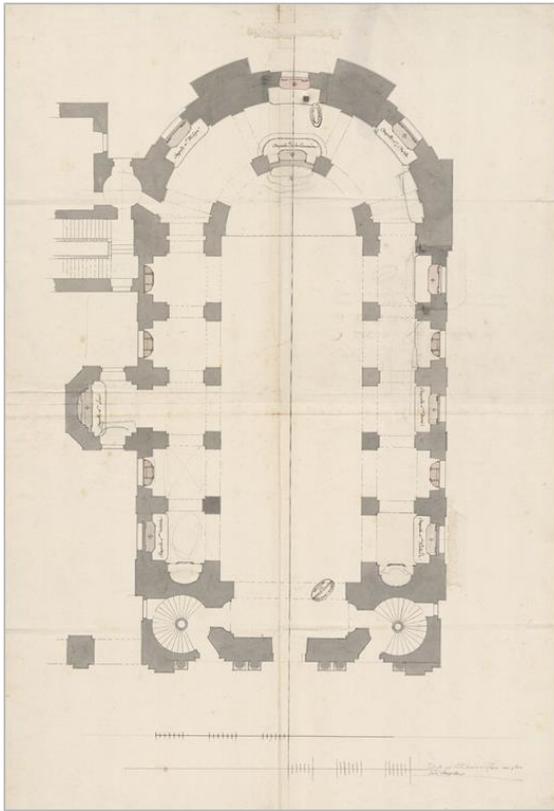


(b) 배경 제거 이미지

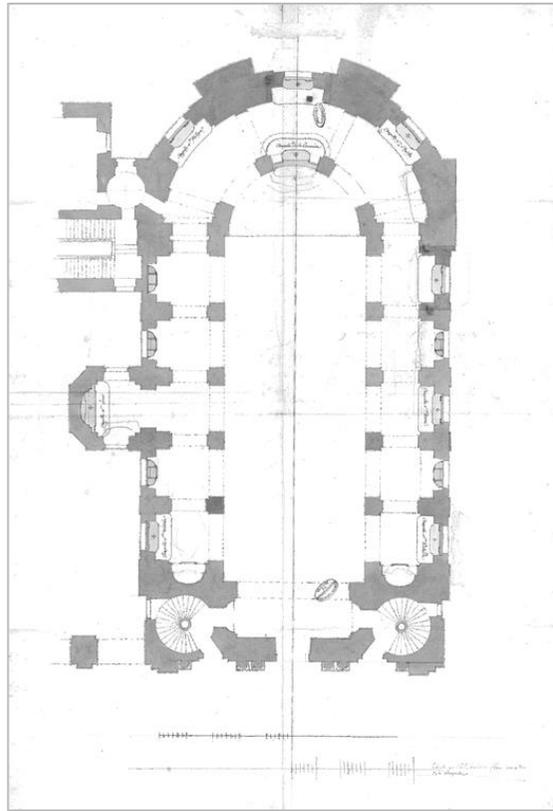


(c) 이진화 결과 이미지

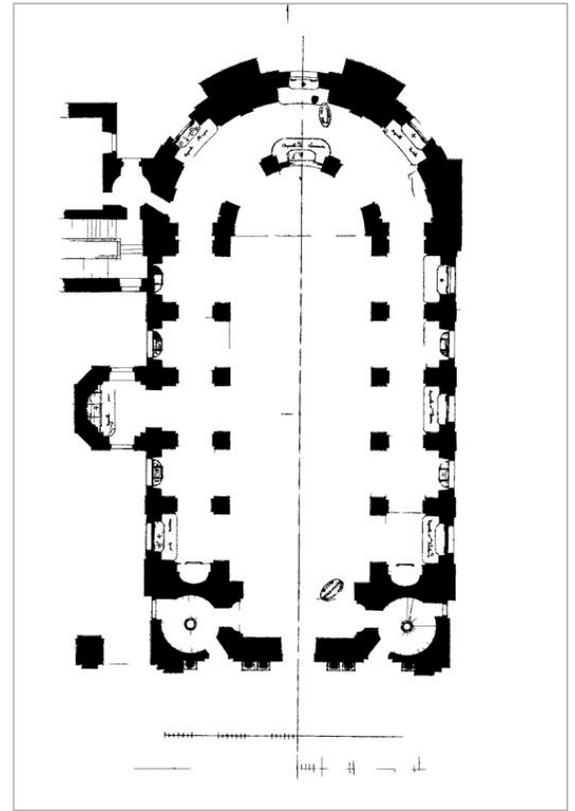
[그림 C-3] 곡선 객체를 포함하는 베르사유 궁전 건축도면 3번 이진화 결과



(a) 원본 이미지

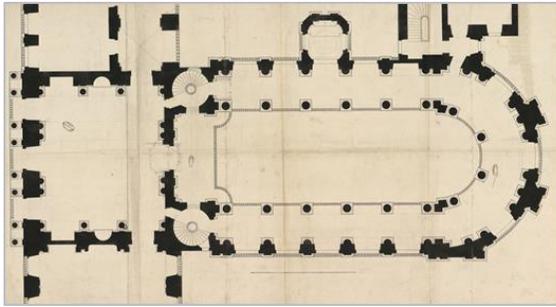


(b) 배경 제거 이미지

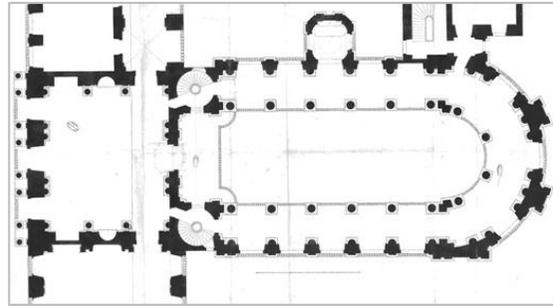


(c) 이진화 결과 이미지

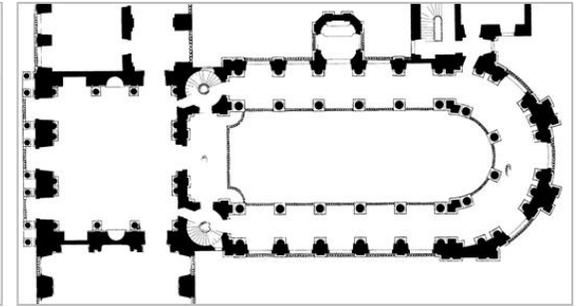
[그림 C-4] 곡선 객체를 포함하는 베르사유 궁전 건축도면 4번 이진화 결과



(a) 원본 이미지

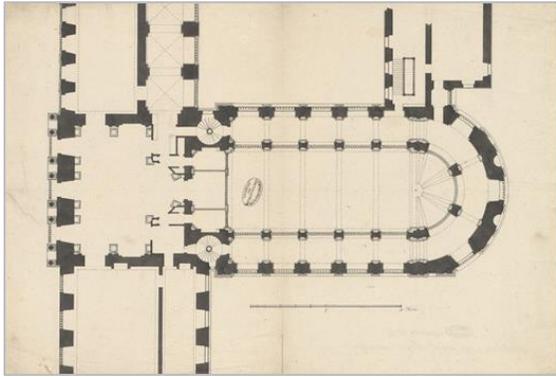


(b) 배경 제거 이미지

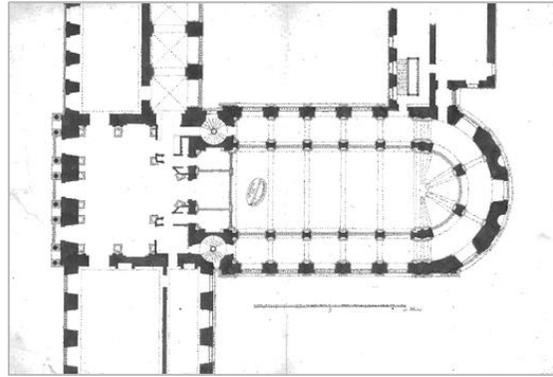


(c) 이진화 결과 이미지

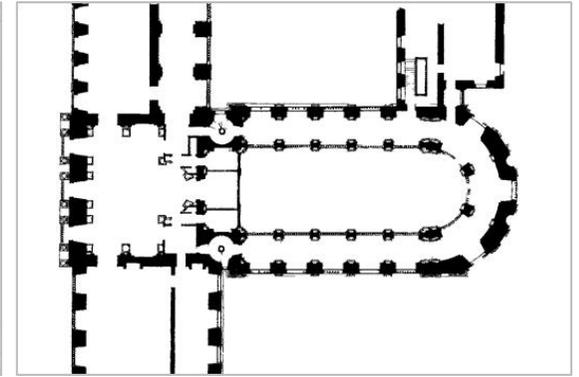
[그림 C-5] 곡선 객체를 포함하는 베르사유 궁전 건축도면 5번 이진화 결과



(a) 원본 이미지

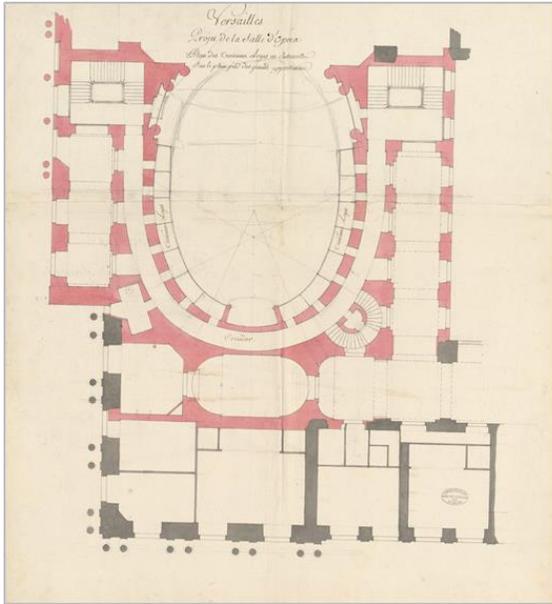


(b) 배경 제거 이미지

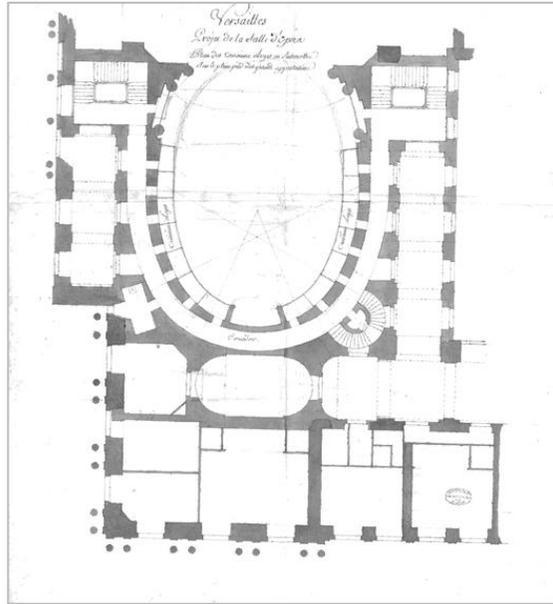


(c) 이진화 결과 이미지

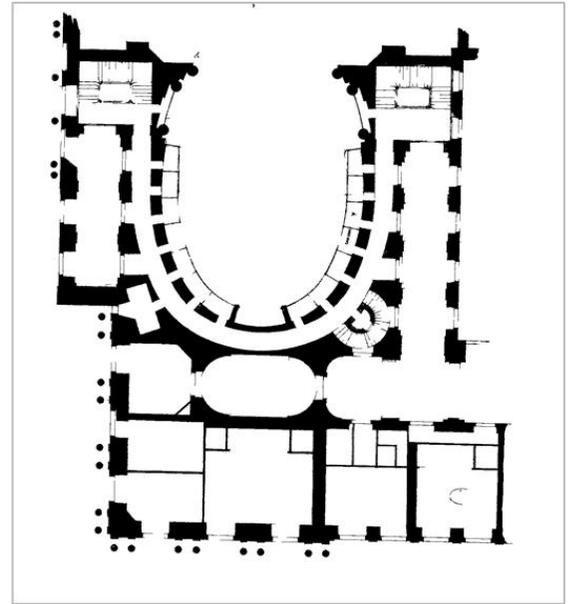
[그림 C-6] 곡선 객체를 포함하는 베르사유 궁전 건축도면 6번 이진화 결과



(a) 원본 이미지

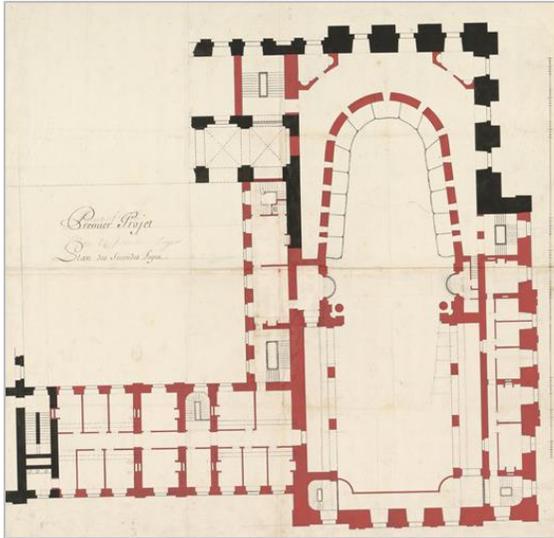


(b) 배경 제거 이미지

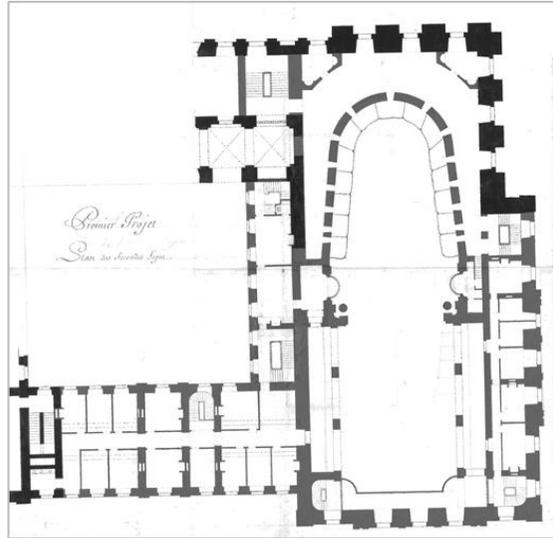


(c) 이진화 결과 이미지

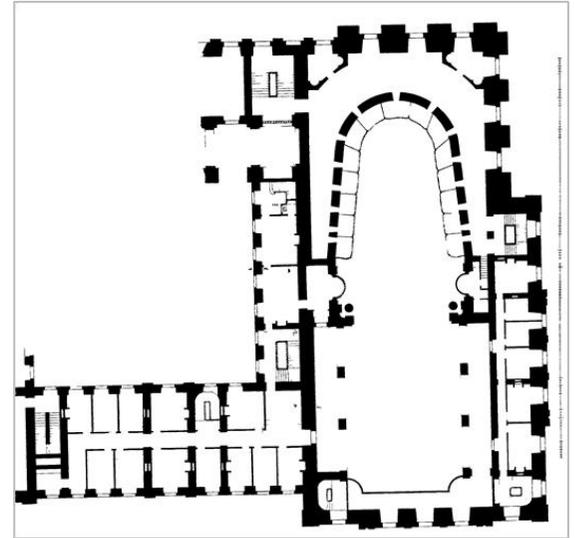
[그림 C-7] 곡선 객체를 포함하는 베르사유 궁전 건축도면 7번 이진화 결과



(a) 원본 이미지

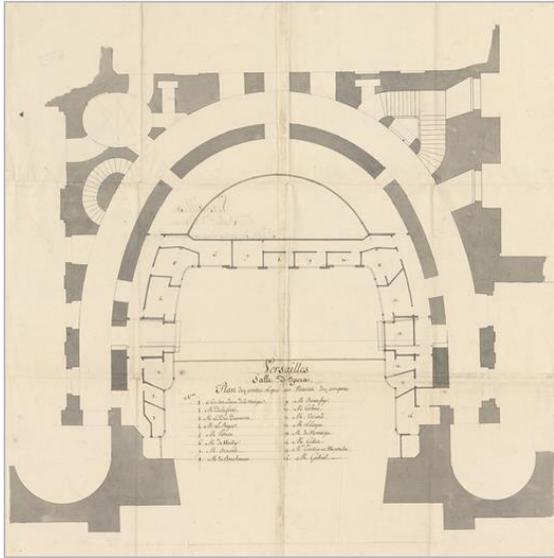


(b) 배경 제거 이미지

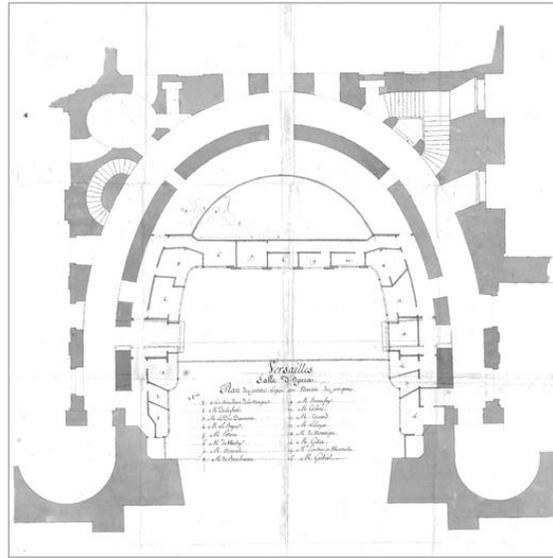


(c) 이진화 결과 이미지

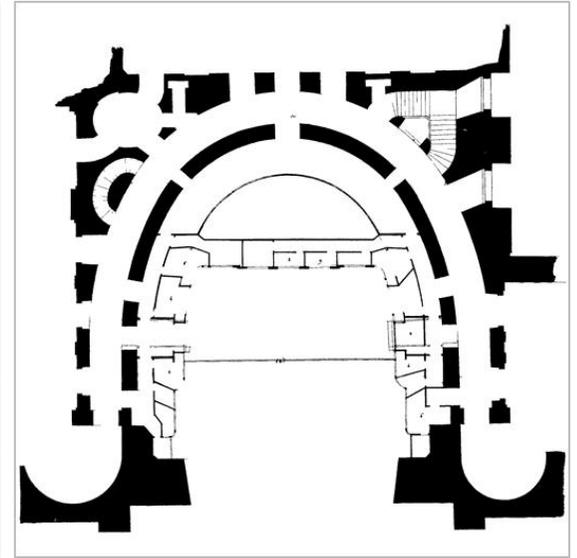
[그림 C-8] 곡선 객체를 포함하는 베르사유 궁전 건축도면 8번 이진화 결과



(a) 원본 이미지

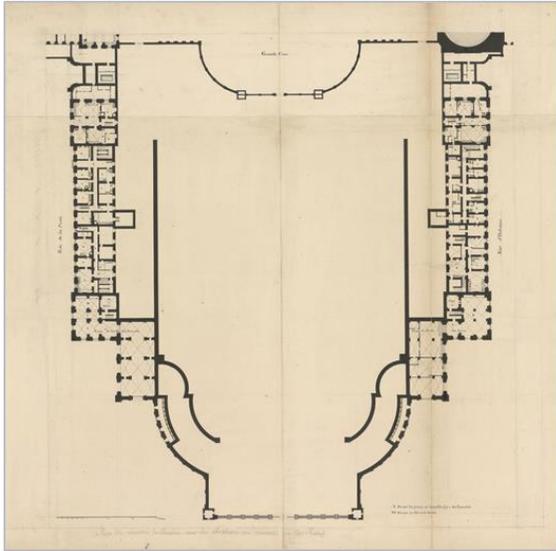


(b) 배경 제거 이미지

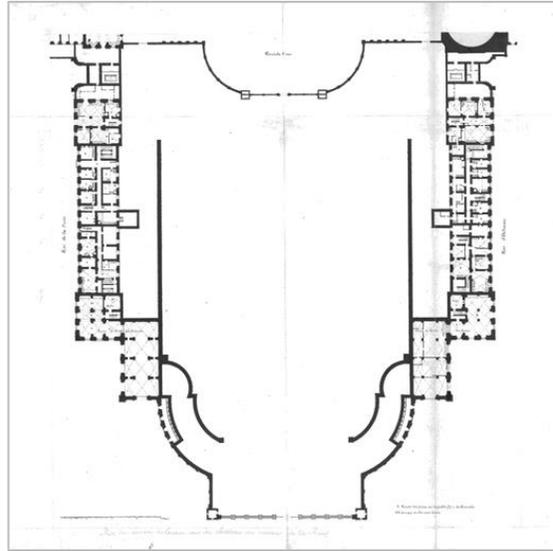


(c) 이진화 결과 이미지

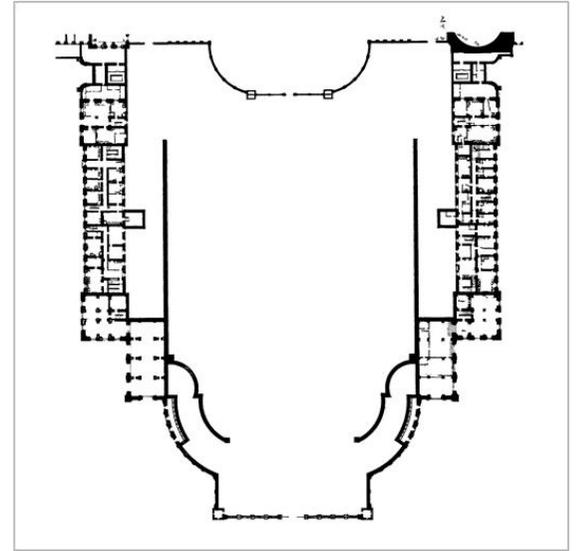
[그림 C-9] 곡선 객체를 포함하는 베르사유 궁전 건축도면 9번 이진화 결과



(a) 원본 이미지



(b) 배경 제거 이미지



(c) 이진화 결과 이미지

[그림 C-10] 곡선 객체를 포함하는 베르사유 궁전 건축도면 10번 이진화 결과

Abstract

Development of Machine Learning Based Binarization Technique of Hand-drawn Floor Plans for Automatic Extraction of Indoor Spatial Information

Suh, Hanew

Department of Civil and Environmental Engineering
College of Engineering
The Graduate School
Seoul National University

Along with the recent development of artificial intelligence and the Internet of Things, social interest in indoor location-based services providing real-time information from user location is getting high. For location-based service development, indoor spatial modelling is essential to represent indoor topology.

Therefore, many studies have been conducted to extract indoor structure information from various types of data such as laser scanners, architectural drawing images, and CAD plans. In particular, the automatic extraction technology of indoor space

information is economically efficient compared to manual modeling, so algorithms for automatic extraction of floor plan entities like walls, windows, and stairs from 2D floor plan image are actively developed. Previous studies mostly used “clean” floor images that floor plan entities and background are clearly distinguished. However, in the case of hand-drawing architectural floor plans created using various types of pens and ink, there are large numbers of noise in background. In addition, since the pixel intensities of every floor plan entities are not constant depending on the pen or ink used, there is a limit to applying the previous algorithms. Therefore, this study aims to perform binarization to distinguish floor plan entities from background with noise and irregular patterns. The purpose of this study is to expand the scope of previous floor plan analysis studies to historical and old buildings.

For dataset, we use architectural drawings of the Japanese colonial period written in the early 1900s. The Japanese architectural drawings used in this study have various types of noise made during the process of storage and digitization. Also, floor plan entities consist of all different colors depending on the type of materials used. We apply learning-based binarization algorithm and our algorithm can be divided into two main steps. The first step is to reduce the noise that is widely distributed across the background of the drawing image using a Gaussian mixture model. The second step is to extract features that distinguish objects and backgrounds based on the random forest model, and to learn various forms of small noise. For evaluation, we perform the classification

performance of suggested algorithm on test set. Our binarization algorithm results in 98.5% precision and 99.0% F1–score rate. This study has two main contributions. First, our algorithm successfully distinguishes various types of floor plan entities with different thickness. Second, study scope of automatic extraction of spatial information from floor plan image can be expanded from electronic floor plan image to hand–drawing architectural floor plans.

Keywords : Floor Plan Analysis, Indoor Spatial Information, Machine Learning, Binarization, Random Forest

Student Number : 2019–23752