

RESEARCH ARTICLE

# Similarity Evaluation of Patent Drawings using ResNet and SIFT

Hyuna Kim\*, Seongju Kim, Cheonho Park, Kyungsoon Kim

Ph.D Candidate, IT Policy and Management, Soongsil University, Seoul, Republic of Korea

\*Corresponding Author: Hyuna Kim (hannah\_kim@naver.com)

## ABSTRACT

The study of the similarity evaluation and retrieval of patent documents is critical not only for the efficient management of patent literature, but also for the rapid and effective collection of information in industrial and technological fields. Patent drawings visually represent the outcomes of technological advancements and innovations, but have not been given as much importance as texts in the past.

This study evaluated the similarity of patent drawings for effective retrieval using the representative deep-learning model, ResNet-50, and the traditional computer vision algorithm, scale-invariant feature transform (SIFT). First, a classification experiment using 10,827 patent drawings was conducted to evaluate the similarity of the visual types, achieving a classification performance with an accuracy exceeding 95%. Second, a retrieval experiment using 5,000 technical drawings was conducted to compare the features of ResNet and SIFT based on their similarity. Finally, the retrieval and matching performances of ResNet and SIFT were evaluated using 50 original data samples and 4,800 augmented data samples created by various forms of editing. ResNet demonstrated an average matching performance of 72.54%, whereas SIFT achieved an average matching performance of 86.71%.

The findings reveal that, unlike ResNet-50, which compares similarity using the entire image information, SIFT evaluates similarity based on attribute information, such as key points within the image. Consequently, ResNet is advantageous for identifying visually similar images, whereas SIFT excels in identifying identical images.

## KEYWORDS

Patent drawings, binary images, classification and retrieval, ResNet-50, SIFT, similarity Comparison



## Open Access

**Citation:** Kim H, et al. 2024. Similarity Evaluation of Patent Drawings using ResNet and SIFT. The Journal of Intellectual Property 19(3), 131-153.

**DOI:** <https://doi.org/10.34122/jip.2024.19.3.7>

**Received:** May 4, 2024

**Revised:** June 11, 2024

**Accepted:** September 3, 2024

**Published:** September 30, 2024

**Copyright:** © 2024 Korea Institute of Intellectual Property

**Funding:** The author received manuscript fees for this article from Korea Institute of Intellectual Property.

**Conflict of interest:** No potential conflict of interest relevant to this article was reported.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

원저

# ResNet과 SIFT를 이용한 특허 도면의 유사도 평가 및 검색 연구\*

김현아\*, 김성주, 박천호, 김경순

송실대학교 일반대학원 IT정책경영학과 박사과정수료

\*교신저자: 김현아 (hannah\_kim@naver.com)

## 차례

1. 서론
2. 관련 연구
  - 2.1. 특허 도면의 개요
  - 2.2. 특허 도면 분류 및 검색에 관한 선행 연구
  - 2.3. ResNet-50
  - 2.4. SIFT
3. 데이터 및 방법론
  - 3.1. 데이터 수집 및 전처리 과정
  - 3.2. 특허 도면의 유형 분류 및 검색 방법
  - 3.3. 증강데이터를 이용한 특허 이미지 검색 방법
4. 실험 결과
  - 4.1. 특허 도면의 유형 분류 실험 결과
  - 4.2. 특허 도면의 유사도 평가 및 검색 실험 결과
  - 4.3. 증강데이터를 이용한 검색 실험 결과
5. 결론

## 국문초록

특허 문헌의 유사성 평가 및 검색 연구는 특허 문헌의 효율적인 관리뿐 아니라 산업 및 기술 분야에서 효율적이고 빠른 정보 수집을 위해 중요한 주제로 다뤄지고 있다. 특히 특허 도면은 산업 기술의 발전과 혁신의 결과물을 시각적으로 표현해왔으나 지금까지 텍스트에 비해 중요하게 다뤄지지 못한 측면이 있다.

본 연구는 효과적인 특허 도면의 검색을 위해 딥러닝의 대표 모델 ResNet-50과 전통적인 컴퓨터비전 알고리즘 SIFT를 이용하여 유사성을 평가하는 연구이다. 먼저 시각적 유형의 유사성을 평가하기 위해 총 10,827개의 특허도면을 이용한 유형 분류 실험을 진행했으며 분류성능은 95%가 넘는 Accuracy를 나타냈다. 두 번째로 기술도면 5,000개를 사용하여 ResNet과 SIFT를 이용한 검색 실험을 진행하여 각 모델의 유사성을 평가하는 특징을 살펴보았다. 마지막으로 원본 데이터 50개와 원본 데이터를 다양한 형태로 증강한 데이터 4,800개를 이용하여 편집 유형별로 검색 및 매칭한 결과, ResNet은 72.54%, SIFT는 86.71%의 평균적인 매칭 결과를 나타냈다.

연구 수행 결과, 이미지 전체의 정보를 이용하여 유사도를 비교하는 ResNet-50과 달리 SIFT는 이미지 내 특징점 등 속성 정보를 이용하여 유사도를 판단하므로 시각적으로 유사한 이미지를 찾는 일에는 ResNet이, 같은 이미지를 찾는 일에는 SIFT가 더 강점이 있는 것으로 평가할 수 있다.

## 주제어

특허 도면, 이진 이미지, 분류 및 검색, ResNet-50, SIFT, 유사도 비교

## 1. 서론

특허 문헌은 특허의 배타적 권리와 관련된 정보를 기록한 문서로 새로운 기술이나 발명에 대한 정보를 공개함으로써 기술 발전을 촉진하고, 기술적 지식의 교류를 도모하는 역할을 한다. 텍스트마이닝과 딥러닝 기술의 발전에 따라 특허 빅데이터는 기술 분류, 기술 및 예측, 동향 분석, 전략 수립 등에 활용하는 중요한 수단이 되었다. 또한 기술의 유사성을 평가하기 위해 주요 요약문이나 대표청구항을 직접 비교하는 연구도 증가하였다. 특허의 유사성 분석은 주로 기술의 중복이나 신규성 등을 판단하기 위한 목적으로 특허심사에서 가장 중요한 위치를 차지하고 있으나 대부분은 텍스트를 기반으로 한 연구에 중점을 두고 있으며 이에 비해 특허 도면에 관한 관심은 높지 않은 편이다.

특허 문헌에는 텍스트 외에도 기술을 설명하기 위한 다양한 유형의 이미지 즉 발명품이나 프로세스의 설명, 그림 도면 등을 포함하고 있다. 텍스트는 출원 언어나 용어의 다양성으로 단순한 텍스트 검색만으로는 특허 침해나 유사한 발명품을 탐지하기 쉽지 않다. 고광국(2012)은 특허 이미지 검색이 중요한 이유로 이미지가 특허 검색에서 결정적 역할을 할 수 있으며 이미지는 성격상 출원인의 언어나 용어의 변화로부터 독립적인 점을 꼽았다<sup>1)</sup>. 또한 특허 이미지는 ‘중요한 정보를 전달하는 물체 또는 개념에 대한 시각적 설명’으로 기술에 대해 더 빠른 이해를 얻을 수 있다<sup>2)</sup>는 점에서 텍스트가 제공할 수 어려운 직관성을 가지고 있다.

특허 이미지분석에 관한 연구는 2000년대부터 시작되었으며 주로 전통적인 비전시스템에 기반을 두고 사전에 정의된 특징 추출 알고리즘을 사용하여 이미지에서 유용한 정보를 추출하고 분석하는 데 중점을 두고 있다. 최근 10년간 딥러닝 및 신경망 기술의 발전은 특허 이미지분석 분야에도 많은 변화를 가져왔으며 이미지 분류, 유사성 검색, 텍스트 분석 등과 같은 다양한 연구에 활용되고 있다. 특히 EU 연구비 지원으로 만들어진 CLEP-IP 2011 데이터셋은 유럽 연구자들에게 많은 영감을 주었으며 미국 역시 특허청이 제공하는 특허 이미지 데이터셋을 활용하여 여러 연구가 진행되었다. 그러나 특허를 포함한 지식재산권 이미지 자료에 대한 국내 학술적 연구사례는 매우 미미한 상황이다. 또한 많은 선행 연구사례에서 딥러닝이 이미지분석 분야에서 비전시스템을 능가하는 기술로 자리 잡고 있으나 복잡한 시각적 유형을 가진 특허 도면 분석에는 여전히 한계를 보이고 있다.

본 연구에서는 특허 도면의 시각적 유사성을 평가하고 유사도가 높은 검색 방법을 찾기 위해 세 가지 실험을 진행하였다. 먼저 특허 도면의 다양한 시각적 유형의 유사성을 평가하기 위해 유형 분류 실험을 진행하였으며 총 10,827개의 데이터를 이용하였다. 두 번째로 기술 도면(Technical Drawing) 5,000개를 이용하여 CNN의 대표 모델, ResNet과 전통적인 비전시스템의 하나인 SIFT를 이용하여 특징 추출 방법을 연구하고 유사도 비교 평가를 수행하였다. 세 번째는 일본 데이터 50개와 증강데이터를 이용한 검색실험으로 2개 모델의 성능을 정량적으로 검증하고 비교하였다.

실험 결과, 특허 도면의 유형별 분류에는 95% 이상의 Accuracy를 나타냈으며 5개의 유형에서 그래프가 가장 높은 정밀도를 보였다. ResNet 기반의 유사도 평가에서는 질의 이미지와 같은 이미지에 대해 유사도 값이 1.0 또는 1.0에 가까웠으나 같은 디자인의 특허 도면이라 할지라도 이미지의 해상도나 방향이 다른 경우 유사도 값이 낮아지는 결과를 보였다. SIFT 기반의 유사도 평가에서는 이미지의 방향과 상관없이 같은 모양의 이미지와 유사도가 높았고 해상도가

1) 고광국, “개념기반 특허이미지 검색”, 한국과학기술정보연구원, 2012, pp. 1-6.

2) Michal Kucer et al., “DeepPatent: Large scale patent drawing recognition and retrieval”, In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, IEEE, 2022, pp. 2309-2318.

다른 경우에는 차이를 보였다. 증강데이터를 이용한 성능검증 실험에서는 ‘shift’ 편집을 제외하고 평균적으로 SIFT의 성능이 ResNet보다 약 14% 높은 것으로 나타났다.

위와 같이 본 연구는 다각적인 접근 방법을 통해 특허 도면의 유사성 평가 방법과 기준을 찾고자 했다. 또한 그동안 한가지 방식으로 도면의 유사성을 평가해왔던 선행 연구와 달리 딥러닝과 전통적인 컴퓨터비전 알고리즘의 성능과 특성을 구체적으로 비교함으로써 두 가지 접근 방식의 장단점을 명확히 하였다. 기술정보를 시각화하고 이해도를 높일 수 있는 중요한 수단으로서 특허 이미지의 시각적 유사성 평가는 자동화된 특허 도면 검색과 다양한 응용 분야에서 활용될 수 있다.

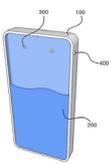
논문의 구성은 다음과 같다. 2장. 관련 연구에서는 특허 도면의 작성 규칙과 특징 등을 알아보고 특허 도면의 유사성 평가와 검색에 관한 선행 연구를 고찰하였다. 또한 본 실험에 활용된 ResNet과 SIFT 모델의 특징에 대해서도 살펴보았다. 3장. 데이터 및 방법론에서는 데이터수집 및 전처리 과정, 이미지 유형분류 및 검색실험 절차와 방법을 제시하였다. 4장. 실험 결과에서는 실험 결과 및 분석, 5장은 전체 연구의 요약과 평가를 기술하였다.

## 2. 관련 연구

### 2.1. 특허 도면의 개요

도면(또는 이미지)을 포함하고 있는 지식재산에는 특허와 실용신안, 상표, 디자인이 있다. 특허 발명은 물건의 발명과 방법의 발명, 물건을 생산하는 방법의 발명으로 구분할 수 있으며<sup>3)</sup> 실용신안은 ‘산업상 이용할 수 있는 물품의 형상·구조 또는 조합에 관한 고안’<sup>4)</sup>으로 물건의 발명에 국한되어 있는 특징이 있다. 통상적으로 특허 도면 검색에서 실용신안 도면도 함께 검색하여 확인하므로 포괄적으로 특허 도면에 포함할 수 있다. 디자인은 물품의 형상·모양·색채 또는 이들을 결합한 것<sup>5)</sup>이다. 다음의 사례와 같이 물건에 따라 특허(물건의 발명), 실용신안, 디자인이 유사한 이미지를 보유할 때가 있으나 디자인은 특허나 실용신안과 발명의 성격이 달라 통상적으로 특허 도면으로 분류하지 않는다.

<표1 특허/실용신안/디자인 도면의 유사 사례>

| 구분     | 특허 도면   | 실용신안 도면   | 디자인 도면  |
|--------|---|---|---|
| 발명의 명칭 | 온도변화 감지부가 구비된 휴대폰 케이스   | 휴대폰 케이스   | 휴대폰 케이스   |
| 도면     |  |  |  |
| 특징     | 도면에 부호 표시 있음.<br>때로 도면 내에 부호 설명을 포함하고 있음  | 도면에 부호 표시 있음  | 도면에 부호 표시 없음  |

3) 특허법 제2조 제3항.

4) 실용신안법 제4조 제1항.

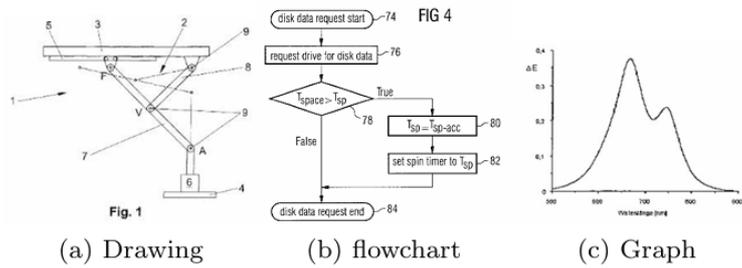
5) 디자인보호법 제2조 제1항.

특허 도면은 특허 출원의 중요한 구성 요소로 특허 명세서에 기술된 발명을 시각적으로 묘사하고 명확하게 하는 역할을 하므로 세계 특허청은 특허 도면의 작성 규칙을 정하고 있다.<sup>6)</sup> 한국은 특허법 및 시행규칙에서 평면도 또는 입면도를 흑백으로 선명하게 도시하도록 규정하고 있고 미국과 일본 역시 흑색 잉크 도면을 기본으로 하고 있다. 유럽은 영속성 있고 검고 충분히 짙고 두께가 균일하고 명확한, 색상 없는 선과 스트로크를 이용해야 하며 PCT는 내구성 있는 짙은 흑색, 균일한 두께와 선명한 선 및 획을 사용하도록 규정하고 있다. 즉, 발명의 내용을 표현하기 위해 불가피한 경우를 제외하고는 모두 흑백 도면을 사용하는 것이 세계적으로 특허 도면의 중요한 규칙이라 할 수 있다. 이러한 특허 도면의 작성 규칙은 특허 도면이 시각적으로도 명확하게 정의된 경계와 형상을 제공하며 특허 도면의 세부 사항을 뚜렷하게 표현할 수 있게 한다. 따라서 흑백 도면은 이미지분석 관점에서 분석 대상 자료의 대부분이 컬러 이미지가 아닌 이진 이미지(바이너리 이미지)라는 독특함을 제공하고 있다.

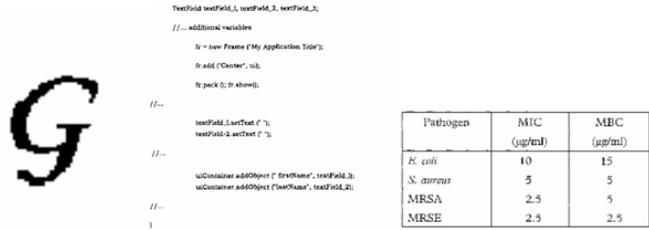
특허 도면의 또 다른 특징은 특허의 대상이 일정한 형상을 가진 물건/물체가 아닌, 화학발명이나 방법 발명같이 실질적으로 그 형상이 존재하지 않을 때 보통 화학 발명식이나 방법의 내용이나 순서를 표현한 순서도 등도 도면으로 활용하고 있다는 점이다. Alla G. Kravets et al.(2017)<sup>7)</sup>은 특허 도면의 유형을 기술 도면(Technical Drawing), 화학구조(Chemical structure), 프로그램 코드(Program code), 유전자 서열(Sequence of gen), 순서도 및 UML 다이어그램(Flowchart and UML diagram), 그래프(Diagram or Graph), 수학적식(Mathematical Formula), 테이블(Table) 등 8개로 구분하였으며 Jiang, S. et al.(2020)<sup>8)</sup>, Castorena, J. et al.(2020)<sup>9)</sup>은 기호(Symbol)를 추가하여 9가지로 구분하였다. 이렇듯 특허 도면의 유형은 도면을 시각적 관점으로 분류한 것으로 기술의 분류와는 무관하며, 일부 도면은 두 가지 이상의 유형이 하나의 도면에 표시되어 도면을 단순히 하나의 유형으로 자동 분류하기 쉽지 않은 사례도 많다.

- 
- 6) 특허청, “명세서/도면의 작성요령 및 분할출원 제도 개선에 관한 연구”, 특허청 특허심사 정책과, 2012, 4-70면.
- 7) Alla Kravets et al., “Patents images retrieval and convolutional neural network training dataset quality improvement”, In Proceedings of the IV International research conference “Information technologies in Science, Management, Social sphere and Medicine(ITSMSM 2017)”, Atlantis Press, 2017, pp. 287-293.
- 8) Shuo Jiang et al., “A convolutional neural network-based patent image retrieval method for design ideation”, In Proceedings of the ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol.9, V009T09A039, The American Society of Mechanical Engineers, 2020, pp. 1-11.
- 9) Juan Castorena et al., “Learning spatial relationships between samples of patent image shapes”, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2020, pp. 655-662.

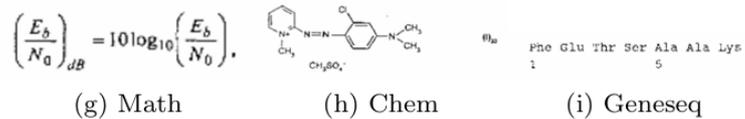
<그림1 CLEP-IP의 9가지 클래스 이미지>



(a) Drawing (b) flowchart (c) Graph



(d) Symbol (e) Program (f) Table



(g) Math (h) Chem (i) Geneseq

(출처 : Juan Castorena et al. 2020)

## 2.2. 특허 도면 분석 및 검색에 관한 선행 연구

이진 이미지는 픽셀값이 흰색 또는 검은색으로, 컬러 이미지와 같이 3개의 채널이 아닌 단일 채널을 갖는 이미지를 말한다. 이러한 특성으로 비전시스템은 이진 이미지를 대상으로 히스토그램 분석과 임계값을 이용한 이진화 변환, 침식과 팽창, 에지 검출과 허프 변환 등과 같은 기본적인 이미지 처리 알고리즘을 사용하여 객체의 윤곽을 추출하거나 노이즈를 제거하는 등의 방법을 사용해 왔다.<sup>10)</sup> 이후 이진 이미지분석뿐 아니라 특허 도면 분석에도 활용된, 컴퓨터비전의 대표적인 알고리즘은 SIFT와 Fisher Vector, AHDH, Local Primitives 등이 있다.

David G. Lowe(2004)<sup>11)</sup>는 이진 이미지에서 고유한 불변 특징을 추출하여 사물이나 장면의 다양한 관점(View)에서 객체 간 매칭 점을 찾는 방법을 제안하였다. 이 접근법은 이미지 데이터를 척도 불변 좌표로 변환하므로 SIFT(Scale Invariant Feature Transform)로 불리면서 이후 많은 연구에서 활용되었다. 박정범(2017)<sup>12)</sup>은 SIFT를 이용하여 특허 이미지의 특징점을 사용하여 문서 간의 유사성을 분석하는 연구를 진행한 바 있다.

SIFT가 이미지의 특징점을 찾아내고 그 특징을 설명하는 방법이라면 Fisher Vector는 이미지 전체에서 추출한 특징 벡터들을 통계적으로 모델링하고 그룹화하여 하나의 벡터로 표현하는 방법이다. 제록스 유럽 연구센터팀은 Clef-IP 2011 특허 이미지 분류 및 이미지 기반 특허

10) 최형일, 「원리를 설명하는 컴퓨터비전 최강의 수업」, 초판, 홍릉, 2023.

11) David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, Vol.60(2004), pp. 91-110.

12) 박정범, "Patent Document Similarity based on Image analysis by using the SIFT-Algorithm and OCR-Text", 배재대학교, 석사, 2017.

검색 연구에서 Fisher Vector를 활용하였다.<sup>13)</sup>

Sidiropoulos, P. et al.은 2010년, 2011년 연구에서 도면의 영역 분할과 히스토그램을 기반으로 하는 이진 이미지 검색 방법, ‘적응형 계층적 밀도 히스토그램’(Adaptive Hierarchical Density Histogram)을 제안하였다<sup>14)15)</sup>. AHDH는 이미지를 겹치지 않는 직사각형 영역으로 분해하고 각 영역의 밀도 히스토그램을 추출하는 방법이다. 즉 모든 영역의 기하학적 중심을 추정하여 영역을 분할하고 각 영역 내 픽셀값의 밀도를 계산한다. 밀도가 높은 영역은 세분화하여 다시 밀도를 계산하는 데 이런 과정을 반복하여 계층적인 구조를 형성하고 각 계층의 데이터 분포를 표현하는 방식이다. 이 연구팀의 일원이었던 Vrochidis et al.는 다른 연구에서 텍스트 기반/개념 기반/콘텐츠 기반 검색이 가능한 특허 이미지 검색 프레임워크를 제안하였으며 여기서 특허 도면의 특징 벡터를 추출하기 위해 AHDH를 이용하였다.<sup>16)17)18)</sup>

Local Primitives Histogram은 AHDH보다 단순한 방식으로 이미지를 격자 형태의 작은 영역으로 나눈 뒤 각 영역에서 지역적인 특징을 추출하여 히스토그램으로 표현한 것이다. N. Bhatti et al. (2018)<sup>19)</sup>은 그의 연구에서 Local Primitives Histogram에 기반한 측지선 컨텍스트 모델과 순환 컨텍스트 모델을 제안하였다.

딥러닝 기술의 발전으로 특허 도면의 분석과 검색 연구에도 딥러닝 모델이 적용되기 시작했다. Alla G. Kravets et al.(2017)은 훈련 데이터 세트의 품질을 개선하기 위해 이미지 크기별로 CNN 모델을 적용하여 유형별 분류 실험을 한 결과, 200X200(50%)과 400X400(51%) 이미지 간에 훈련 데이터 정확도는 큰 차이가 없는 것을 확인하였다.<sup>20)</sup> Castorena, J. et al. (2020)은 포인트 클라우드를 이용하여 특허 이미지 모양의 특징 벡터를 추출하였으며 포인트 클라우드의 밀도에 따라 분류 성능에 차이가 있는 것을 발견했다.<sup>21)</sup> 또한 여러 연구에서 디자인 관점의 특허 이미지 인식 및 검색 연구가 진행되었다. Jiang, S. et al.(2020)은 CLEP-IP 2011 데이터를 이용하여 시각 유형 예측과 IPC 예측을 통합한 Dual-VGG 모델을 선보였으며<sup>22)</sup> Kucer et

13) Gabriela Csurka et al., “XRCE’s Participation at patent image classification and image-based patent retrieval tasks of the Clef-IP 2011”, CEUR Workshop Proceedings, Vol.1177, CEUR-WS, 2011, pp. 1-7.

14) Panagiotis Sidiropoulos et al., “Adaptive hierarchical density histogram for complex binary image retrieval”, 2010 International Workshop on Content Based Multimedia Indexing (CBMI), IEEE, 2010, pp. 1-6.

15) Panagiotis Sidiropoulos et al., “Content-based binary image retrieval using the adaptive hierarchical density histogram. Pattern Recognition”, *Pattern Recognition*, Vol.44 No.4(2011), pp. 739-750.

16) Stefanos Vrochidis et al., “Towards content-based patent image retrieval : a framework perspective”, *World Patent Information*, Vol.32 No.2(2010), pp. 94-106.

17) Stefanos Vrochidis et al., “Concept-based patent image retrieval”, *World Patent Information*, Vol.34 No.4(2012), pp. 292-303.

18) Stefanos Vrochidis et al., “Enhancing patent search with content-based image retrieval”, Edit. by Paltoglou, G., Loizides, F., Hansen, P., Professional Search in the Modern World: COST Action IC1002 on Multilingual and Multifaceted Interactive Information Access, Springer, 2014, pp. 250-273.

19) Naeem Bhatti et al., “Contextual local primitives for binary patent image retrieval”, *Multimedia Tools and Applications*, Vol.77 No.7(2018), pp. 9111-9151.

20) Alla Kravets et al., “Patents images retrieval and convolutional neural network training dataset quality improvement”, In Proceedings of the IV International research conference “Information technologies in Science, Management, Social sphere and Medicine(ITSMSM 2017)”, Atlantis Press, 2017, pp. 287-293.

21) Juan Castorena et al., “Learning spatial relationships between samples of patent image shapes”, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2020, pp. 655-662.

22) Shuo Jiang et al., “A convolutional neural network-based patent image retrieval method for design ideation”, In Proceedings of the ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol.9, V009T09A039, The A

al.(2022)는 USPTO 데이터를 이용하여 대규모 데이터 세트인 DeepPatent를 구축하고 하이브리드 CNN-RNN 아키텍처를 사용한 디자인 특허 검색 모델, Patent-Net을 제안하였다.<sup>23)</sup> Ghauri, J.A. et al. (2023)도 특허 도면의 시각화 유형 및 관점 분류를 위해 CLIP 기반의 접근 방식과 ResNet-50 등 4개의 CNN 모델을 활용하여 분류 실험을 수행한 결과 고정 가중치가 있는 트랜스포머 모델보다 성능이 뛰어난 CNN 기반 모델의 성능이 더 우수하다는 결과를 얻었다.<sup>24)</sup> 한재석 et al. (2023)은 CNN 모델 5종을 이용하여 유사도 평가를 통해 검색시스템을 구현하는 연구에서 DenseNet-201과 EfficientNet-B1이 대규모 특허 도면 검색시스템에 적합함을 확인했다.<sup>25)</sup>

국내에서는 머신러닝을 이용한 상표 이미지 분류 연구(유대근, 2017)<sup>26)</sup>, CNN(윤재웅 et al., 2019)<sup>27)</sup>이나 어텐션 모형(윤재웅 et al., 2022)<sup>28)</sup>을 이용한 상표 검색 연구 등이 있으나 모두 상표를 대상으로 하였다. 드물게 SIFT와 OCR을 이용한 특허문서의 유사성 연구(박정범, 2017), 딥러닝 기반의 특허 도면 검색 연구(한재석 et al., 2023)<sup>29)</sup>가 진행되었으나 국내 특허 도면 연구는 여전히 초입 단계라고 할 수 있다.

### 2.3. ResNet-50

ResNet(Residual Network) 같은 CNN 모델의 이미지 특징 추출 방법은 컨볼루션(회선 연산)이라고 부르는 합성곱 연산을 통해 이루어진다. 합성곱 신경망은 입력 이미지를 받아 합성곱 층과 풀링층을 여러 번 반복 수행하여 필요한 특징을 추출한다. 마지막으로 추출된 값은 1차원 배열로 변환되는데 이때 1차원 배열에 담긴 값이 이미지의 feature를 한 줄로 이어 만든 값이다. 즉 커널 연산을 반복하여 이미지의 종합적인 특징을 포함한 값이라 할 수 있다.

ResNet은 심층 잔차 학습(Deep Residual Learning, DRL)을 위한 신경망 프레임워크의 하나로 네트워크 깊이를 증가시킬 때 발생하는 성능 저하 문제를 해결하기 위해 제안되었다.<sup>30)</sup> ResNet-50은 ResNet의 변형된 모델로 총 50개의 층으로 구성되어 있으며 ImageNet 데이터셋을 기반으로 사전 학습된 모델로 제공된다. ImageNet은 1,000개의 범주로 구성된 대규모 이미지 데이터셋으로 약 1,400만 개의 이미지와 그에 해당되는 레이블로 구성되어 있어 다양한 모델 생성에 활용되고 있다.<sup>31)</sup> 그러나 ImageNet은 다양한 범주의 사물, 동물, 식물 등을 포함하고 있어 ImageNet에서 사전 학습된 모델을 일반적인 자연 이미지와 다르게 형태가 뚜렷하고 추상적인 특허 도면에 직접 적용하는 것에 한계가 있다. Kucer et al.(2022)는 이러한 단점을

merican Society of Mechanical Engineers, 2020, pp. 1-11.

23) Michal Kucer et al., "DeepPatent: Large scale patent drawing recognition and retrieval", In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, IEEE, 2022, pp. 2309-2318.

24) Junaid Ahmed Ghauri et al., "Classification of Visualization Types and Perspectives in Patents", In Proceedings of the International Conference on Theory and Practice of Digital Libraries, Springer Nature Switzerland, 2023, pp. 182-191.

25) 한재석 외 3인, "딥러닝 기반 특허 도면 검색 연구", 「한국정보과학회 학술발표논문집」, 2023년 6월(2023), 664-666면.

26) 유대근, "머신러닝을 활용한 특허상표이미지 분류", 경북대학교, 석사, 2017, 1면.

27) 윤재웅 외 5인, "합성곱신경망(Convolutional Neural Network)을 활용한 지능형 유사 상표검색 모형 개발", 「경영과 정보연구」, 제38권 제3호(2019), 55-80면.

28) 윤재웅 외 2인, "어텐션모형을 활용한 상표 이미지 내 도형 분류 모형 개발", 「경영과 정보연구」, 제41권 제4호(2022), 73-91면.

29) 한재석 외 3인, 앞의 논문, 664-666면.

30) Kaiming He et al., "Deep residual learning for image recognition", In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016, pp. 770-778.

31) Jia Deng et al., "Imagenet: A large-scale hierarchical image database", In Proceedings of the 2009 IEEE conference on computer vision and pattern recognition, IEEE, 2009, pp. 248-255.

극복하기 위해 대상의 시각적 특성이 있는 USPTO의 디자인 특허 도면 450,000개를 이용하여 ResNet-50을 재훈련하였으며 ImageNet만을 이용하였을 경우보다 검색 결과에서 더 높은 성능을 달성했다.<sup>32)</sup>Ghauri, J.A. et al. (2023)의 4개의 CNN모델(ResNet-50, RegNetY-16GF, EfficientNetV2- M 및 ResNeXt-101)과 2개의 트랜스포머 모델(CLIP+MLP)을 사용한 시각적 유형 분류 실험에서도 ResNetXt 모델이 트랜스포머 모델에 비해 우수한 것으로 나타났다.<sup>33)</sup>

이같이 ResNet은 모델의 개선, 특허 도면 데이터의 추가를 이용한 사전학습 모델 생성 등 특허 도면 이미지의 유형 분류 및 검색 분야에서도 가장 많이 활용되고 있다.

## 2.4. SIFT

SIFT(Scale Invariant Feature Transform)는 이미지에서 특징점을 추출하는 가장 유명한 알고리즘으로 일반적으로 객체 인식에 사용된다. SIFT는 이미지 크기 조정 및 회전에 대해 불변하고 조명 및 3D 카메라 시점의 변경에도 부분적으로 변하지 않는 특징을 가지고 있다.<sup>34)</sup> SIFT의 이미지 특징 추출은 스케일 공간 극값 탐지, 키포인트 현지화, 오리엔테이션 할당, 키포인트 설명자와 같은 4가지 작업을 거쳐 진행된다. 2개 이미지의 특징을 비교하기 위해서는 Feature Matching 단계를 거치는데 특징점 간 유클리드 거리(Distance)의 최소값을 매칭점으로 사용한다. 즉 첫 번째 최소거리와 두 번째 최소거리의 비율이 0.8을 넘으면 특징점을 매칭점으로 사용할 수 없는 계산 방법을 사용하여 좋은 매칭비율을 계산하며 이를 이용해 유사도를 측정할 수 있다.

## 3. 데이터 및 방법론

### 3.1. 데이터 수집 및 전처리 과정

특허 이미지를 이용한 다각적인 실험을 위해 국내 특허에서 다음과 같이 총 12,827개의 이미지를 활용하였다. 데이터는 기술 도면이 많은 B25J 섹션(매니플레이터; 매니플레이터 장치를 갖는 실)과 매니플레이터를 포함하고 있는 다른 섹션, 화학구조가 많은 C 섹션 등에서 무작위로 수집하였다. 도면의 유형에 따라 파일을 수동으로 분류하였으며 데이터가 적은 유형은 제외하고 총 5개의 유형을 사용했다.

32) Michal Kucer et al., "DeepPatent: Large scale patent drawing recognition and retrieval", In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, IEEE, 2022, pp. 2309-2318.

33) Junaid Ahmed Ghauri et al., "Classification of Visualization Types and Perspectives in Patents", In Proceedings of the International Conference on Theory and Practice of Digital Libraries, Springer Nature Switzerland, 2023, pp. 182-191.

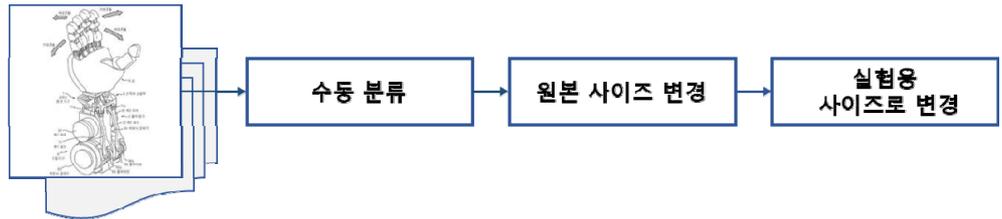
34) David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, Vol.60(2004), pp. 91-110.

&lt;표2 특허 도면 데이터 수집 정보&gt;

| 구분            | 개요                                     |
|---------------|--|
| 지역/기간         | 국내 출원 특허 (~2023년)                      |
| 특허 DB         | WIPS ON, 대표도면                          |
| 기술 분류         | B25J 및 Manipulator를 포함하고 있는 기타 섹션, C섹션 |
| 도면 파일 특징      | 한 변의 길이가 460픽셀                         |
| 이미지 포맷        | JPG                                    |
| (선별 후) 데이터 개수 | 12,827개                                |

수집한 도면 이미지는 한 변의 길이가 최대 460픽셀인 이미지로 우선 데이터의 사이즈를 통일하기 위해 모든 원본 이미지를 너비 460픽셀, 길이 460픽셀의 정사각형 사이즈로 변경하였다. 두 번째는 정사각형으로 변경된 데이터를 딥러닝 실험을 위해 다시 224\*224로 표준화하였는데 이는 ResNet 아키텍처가 학습 및 평가의 효율성을 위해 입력 이미지의 크기가 224x224에 최적화되어 있기 때문이다.

&lt;그림2 특허 도면 이미지의 수집 및 전처리 과정&gt;



유형 분류 실험에는 총 5개의 유형과 총 10,827개의 데이터를 사용했으며 훈련용 데이터와 테스트용 데이터는 약 7:3의 비율로 나누었다. 기술도면은 총 5,000개를 수집하였으나 과적합의 문제를 최소화하기 위해 3,000개만 유형 분류에 활용하였다. 특허 도면의 유형에는 Alla G. Kravets et al.(2017)<sup>35)</sup>의 연구를 인용하여 총 8개로 구분할 수 있으나 국내 특허 도면에서는 프로그램 코드, 수학적, 테이블 등 3가지 유형에 대한 데이터를 충분히 확보하기 어려워 이번 실험에서는 제외하였다.

&lt;표3 특허 도면 유형 분류 실험에 사용한 데이터 현황&gt;

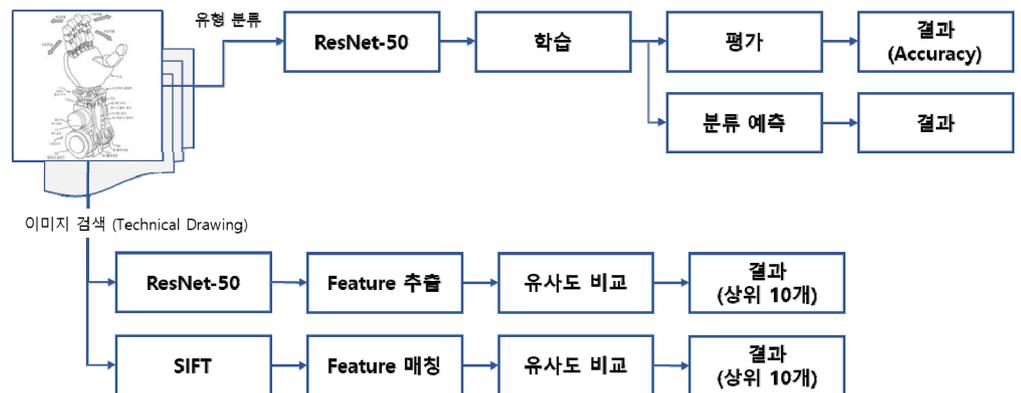
| 구분                                     | 훈련용    | 테스트용   |
|--|--------|--------|
| 기술도면 (Technical_drawing)               | 2,100개 | 900개   |
| 화학구조 (Chemical structure)              | 1,078개 | 455개   |
| 염기서열 (Sequence_of_gene)                | 370개   | 152개   |
| 순서도 및 다이어그램 (Flowchart & UML diagram ) | 2,014개 | 900개   |
| 그래프 (Graph)                            | 1,908개 | 950개   |
| 합계                                     | 7,470개 | 3,357개 |

또한 특허 이미지 검색 실험에서는 기술 도면이 많은 매니플레이터(manipulator)의 대표 도면 이미지 총 5,000개를 사용하였다. 매니플레이터는 인간의 팔과 유사한 동작을 제공하는 기계적인 장치로 자동화, 로봇 관련 기술에 속하여 주로 물건의 디자인 형상이나 설계도가 많은 분야이다. 따라서 매니플레이터가 속한 CPC 섹션, 주로 B25J 섹션과 매니플레이터를 포함하고 있는 다른 기술 분야로부터 도면을 수집하고 활용하였다.

### 3.2. 특허도면의 유형분류 및 검색 방법

이미지의 유형분류는 이미지에서 유의미한 특징을 추출하는 과정을 포함하고 있다. 유형분류 모델은 유사한 이미지를 그룹화하거나 클러스터링하는 데 사용할 수 있으며 같은 유형으로 분류된 이미지는 시각적 관점에서 유사성이 높다고 할 수 있다. 특허 도면을 이용한 유형분류 및 검색실험의 연구 프레임워크는 다음과 같다.

<그림3 특허도면의 유형분류 및 검색 실험 프레임워크>



유형분류 실험은 수집된 5개 유형의 데이터셋(<표3> 참조)을 대상으로 ResNet-50을 이용한 학습(lr=0.001, batch\_size=64, epochs=5, optimizer=SGD) 및 평가를 진행했다.

두 번째로 이미지 검색실험은 ResNet-50과 SIFT를 이용한 유사도 비교 방법을 이용하였으며 두 가지 모델을 이용한 각 검색 방법의 특징을 살펴보면 다음과 같다. 먼저 ResNet-50을 이용한 방법을 보면 우선 검색 대상 데이터 5,000개의 Feature 값을 추출하여 파일로 저장하고(<그림4> 참조) 질의용 이미지의 Feature 값과 코사인 유사도 비교를 수행한다(<그림5> 참조).

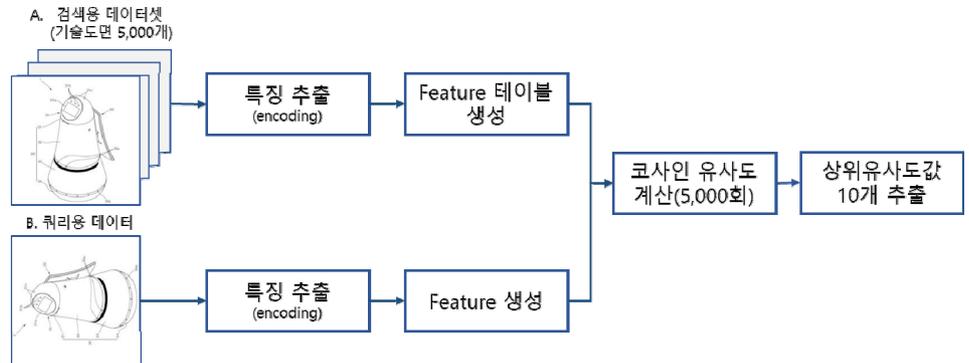
35) Alla Kravets et al., "Patents images retrieval and convolutional neural network training dataset quality improvement", In Proceedings of the IV International research conference "Information technologies in Science, Management, Social sphere and Medicine(ITSMSSM 2017)", Atlantis Press, 2017, pp. 287-293.

<그림4 ResNet-50을 이용하여 추출한 feature table>

| A  | B                               | C       | D       | E       | F       | G       | H       | I       | (K) | BZR     | BZS     | BZT     | BZU     | BZV     | BZW |
|----|---------------------------------|---------|---------|---------|---------|---------|---------|---------|-----|---------|---------|---------|---------|---------|-----|
| 1  | 0 2003021663713-pct00001-4.jpg  | 1.67436 | 0.33656 | 0.00956 | 0.23821 | 0.37728 | 0.14312 | 0.71545 |     | 0.06253 | 0.57367 | 0.02678 | 0.0013  | 0.23459 |     |
| 2  | 1 2004024081306-pct00001-4.jpg  | 0.0809  | 0.45962 | 0.00576 | 0.00538 | 0.29721 | 0.03749 | 0.25349 |     | 0.05326 | 0.05238 | 0.00553 | 0.00084 | 0.16512 |     |
| 3  | 2 2005042668599-pct00001-4.jpg  | 0.02215 | 0.15766 | 0.00076 | 0.09566 | 0.40756 | 0.13096 | 0.27261 |     | 0.00763 | 0.08196 | 0.00881 | 0.00354 | 0.08544 |     |
| 4  | 3 2005051013830-pct00001-4.jpg  | 0.38498 | 0.27892 | 0       | 0.0114  | 0.38498 | 0.03074 | 0.62    |     | 0.1966  | 0.07432 | 0.04325 | 0.00853 | 0.44085 |     |
| 5  | 4 2007005871417-pct00001-4.jpg  | 1.98055 | 0.15582 | 0.41401 | 2.03548 | 0.21316 | 1.47099 | 0.554   |     | 1.3577  | 0.49449 | 0.03713 | 0.00838 | 0.45956 |     |
| 6  | 5 2007056564850-pct00001-4.jpg  | 0.28741 | 0.29503 | 0.00063 | 0.02562 | 0.38784 | 0.1706  | 0.454   |     | 1.1852  | 0.2254  | 0.03502 | 0.00206 | 0.08159 |     |
| 7  | 6 2008005644719-pct00001-4.jpg  | 1.10202 | 0.0369  | 0.06483 | 0.94147 | 0.53502 | 0.34982 | 0.714   |     | 0.3209  | 0.52892 | 0.21211 | 0.35193 | 0.12399 |     |
| 8  | 7 2009026384581-pct00001-4.jpg  | 0.17834 | 0.28609 | 0.00176 | 0.37054 | 0.50586 | 0.05975 | 1.27935 |     | 0.05323 | 0.04157 | 0.00252 | 0.00452 | 0.5815  |     |
| 9  | 8 2009075981643-pct00001-4.jpg  | 1.38835 | 0.28458 | 0.07263 | 0.24586 | 0.34048 | 0.06548 | 0.35446 |     | 0.0754  | 0.7325  | 0.00486 | 0.00365 | 0.50621 |     |
| 10 | 9 2009079758425-pct00001-4.jpg  | 0.29981 | 0.12816 | 0.00249 | 0.19589 | 0.08672 | 0.14887 | 0.88146 |     | 0.20082 | 0.05675 | 0.1929  | 0.03487 | 0.80646 |     |
| 11 | 10 2012040942860-pct00074-4.jpg | 0.00837 | 0.02737 | 0.00154 | 0.1081  | 0.55318 | 0.13924 | 0.1691  |     | 0.11494 | 0.26249 | 0.00766 | 0       | 0.1293  |     |

index      파일명      2048개의 벡터값

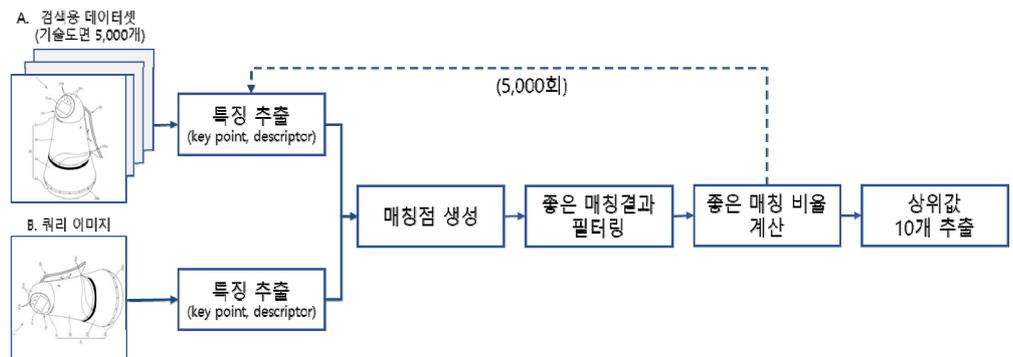
<그림5 ResNet을 이용한 유사도 비교 프로세스>



ResNet 같은 주요 CNN 아키텍처에는 평균 풀링레이어 다음에 FC(Fully Connected) 레이어가 추가되는데 ResNet 레이어의 출력 크기가 2,048개이다. 따라서 파일에는 5,000개의 이미지에 대해 인덱스와 파일경로와 파일명, 2,048개의 벡터값이 저장되어 있다. 다음으로 질의 이미지의 Feature 값과 검색 대상 Feature 값을 각각 코사인 유사도를 계산하여 가장 유사도가 높은 이미지를 10개 확보하여 그 결과를 보여준다. 코사인 유사도는 두 벡터가 얼마나 비슷한 방향을 향하고 있는지 측정하여 유사성을 결정한다. 따라서 해상도나 모양이 같은 이미지를 찾으면 유사도 값은 1이 되고 값이 1에 가까울수록 두 벡터가 비슷한 방향을 향하고 있음을 의미한다.

SIFT를 이용한 비교에서는 질의 이미지와 검색 대상 이미지의 특징을 매치 추출하고 매칭하는 과정으로 진행된다(<그림6> 참조).

<그림6 SIFT를 이용한 유사도 비교 프로세스>

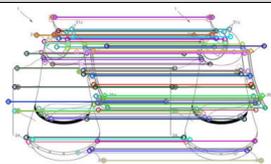


먼저 비교할 2개 이미지의 키-포인트(key point)와 설명자(descriptor)를 계산한다. 키-포인트는 공간 좌표값과 주변 영역의 크기, 반응 값, 방향각 등의 정보를 나타내며 설명자는 각 셀에서 구한 8개의 에지 방향 히스토그램값을 나타낸다. 설명자(descriptor)들의 거리를 계산하면 각 이미지의 매칭점을 생성하게 된다. 이후 매칭점의 순서 기준으로 두 개 매칭점의 거리(d1, d2)를 계산하여 앞의 값(d1)이 뒤의 값(d2)에 0.8을 곱한 값보다 작으면 좋은 매칭 결과라고 할 수 있다. 예를 들어, d1 = 144.298, d2=258.49라고 할 때  $d2 \times 0.8 = 206.79$ 가 된다. 이 경우  $d1 < (d2 \times 0.8)$  조건이 성립하여 좋은 매칭 결과가 되고 그렇지 못한 경우는 필터링되는 단계라고 할 수 있다. SIFT의 매칭 알고리즘은 각 특징점 벡터 간의 거리를 계산하여 이 거리 분포를 확률 밀도함수(Probability Density Function, PDF)로 나타낼 수 있는 데 0.8은 ‘올바른 일치(correct matches)’와 ‘잘못된 일치(incorrect matches)’를 나누는 임계값이라 할 수 있다. Lowe(2004)는 그의 연구에서 40,000개의 키-포인트 데이터베이스를 이용하여 올바른 일치에 대한 실험(Figure 11)<sup>36)</sup>을 진행하여 임계값을 확인했으며 이를 활용하는 개별 실험에서는 보통 0.8에서 0.7 사이의 값을 사용한다. 좋은 매칭 결과로 필터링된 매칭 점 쌍의 수가 확인되면 이를 두 개 이미지의 키-포인트 수의 총합으로 나누어 2를 곱해줌으로써 좋은 매칭 비율(good match ratio)을 얻을 수 있다. 질의 이미지와 비교 이미지의 좋은 매칭 비율이 클수록 유사도가 높은 것으로 볼 수 있다.

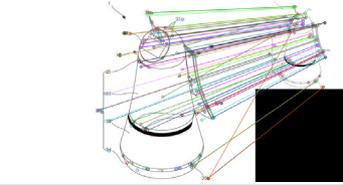
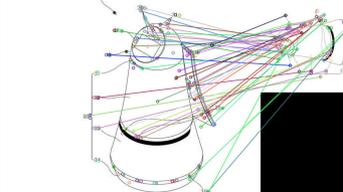
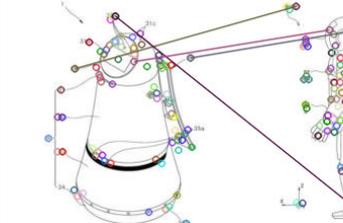
SIFT의 중요한 특징은 다음 5개의 비교사례(<표 4> 참고)를 통해 확인할 수 있다.

- (사례1) 해상도와 방향이 같은 경우, 즉 같은 모양의 이미지는 처음 매칭 결과(Matchers)와 필터링된 결과(Total good matches)의 값이 같고 평균 거리(Average Distance)는 0이며 좋은 매칭비율(Good match ratio)은 1을 나타냄
- (사례2) 해상도는 같으나 방향이 다른 경우 처음 매칭 결과는 98, 좋은 매칭 결과는 76, 평균 거리는 32.69, 좋은 매칭비율은 0.84로 높게 나타남
- (사례3) 해상도는 다르고 방향이 같은 경우 좋은 매칭비율은 0.37로 낮게 나타남
- (사례4) 해상도와 방향이 모두 다른 경우에는 사례3보다 평균 거리와 매칭비율이 낮게 나타남
- (사례5) 이미지가 전혀 다른 경우에는 평균 거리는 더 늘어나고 좋은 매칭비율은 0.03으로 매우 낮게 나타남

<표4 SIFT를 이용한 이미지 간 유사도 비교사례>

| 구분          | 매칭 결과 시각화 이미지  | 매칭 결과  |
|-------------|--|--|
| SIFT 비교 사례1 |  | Matchers : 98<br>Total good matches: 98<br>Average distance: 0.0<br>Good match ratio: 1.00   |
| SIFT 비교 사례2 |  | Matchers : 98<br>Total good matches: 76<br>Average distance: 32.69<br>Good match ratio: 0.84 |

36) David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, Vol.60(2004), pp. 91-110. See Figure 11.

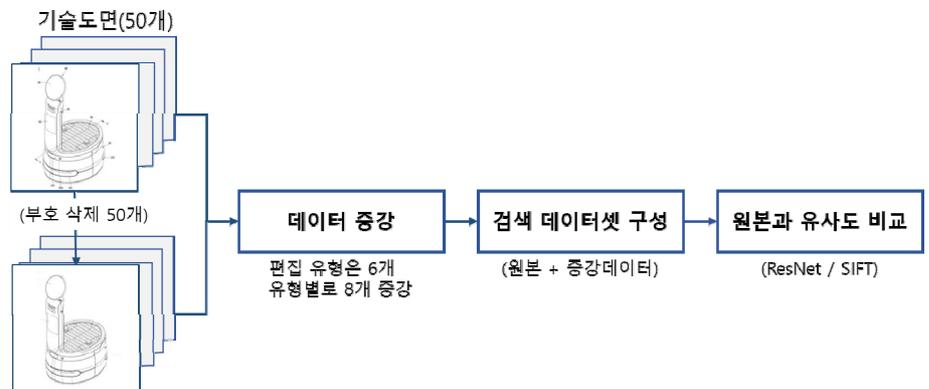
| 구분          | 매칭 결과 시각화 이미지  | 매칭 결과  |
|-------------|--|--|
| SIFT 비교 사례3 |  | Matchers : 228<br>Total good matches: 61<br>Average distance: 97.04<br>Good match ratio: 0.37  |
| SIFT 비교 사례4 |  | Matchers : 228<br>Total good matches: 53<br>Average distance: 107.76<br>Good match ratio: 0.34 |
| SIFT 비교 사례5 |  | Matchers : 98<br>Total good matches: 4<br>Average distance: 231.19<br>Good match ratio: 0.03   |

위 비교사례를 통해 확인할 수 있는 SIFT의 주요한 특징으로는 이미지 방향에 크게 좌우되지 않으나 해상도에는 영향을 받는다는 점이다. 또한 유사도가 떨어질수록 평균 거리와 좋은 매칭 비율도 낮아진다는 것을 확인할 수 있다.

### 3.3. 증강데이터를 이용한 특허 이미지 검색 방법

세 번째로 위 검색 실험 결과에서 나타난 각 모델의 정성적 특징을 정량적으로 평가하기 위한 실험을 진행하였다(<그림7> 참조). 실험방법은 검색 대상 이미지를 다양한 형태로 편집하여 증강하고 편집된 파일을 검색 데이터셋에 섞어 검색하였을 때 총 몇 개와 매칭되는지 평균을 내어 모델별 유사도 성능을 파악하는 방법이다. 증강데이터를 이용한 유사도 비교 프로세스는 다음과 같다.

<그림7 증강데이터를 이용한 유사도 비교 프로세스>



- 임의의 기술도면 50개를 선택한 후 이미지 유사성 비교에 영향을 줄 수 있는 부호를 삭제하여 부호가 없는 파일 50개를 추가로 생성하여 부호있음과 부호없음 파일로 구분함
- 이미지의 원활한 편집을 위해 도면의 배경 사이즈를 500x500픽셀로 확대함
- 편집의 유형은 부호 여부에 따라 각각 도면의 부분 삭제, (이미지 내 객체의) 크기 조정, 회전, (배경 이미지 내에서 상하좌우로) 이동, 노이즈 추가, 부분 삭제와 회전의 결합 등 6종류로 하고 편집 유형별로 8개의 파일을 생성함(예, 회전의 경우 원본 이미지를 45도씩 좌회전하여 총 8개의 이미지 생성). 이렇게 편집하여 증강된 데이터의 수는 총 4,800개(50\*2\*6\*8=4,800)

<그림8 도면 편집의 유형별 사례>



· 데이터 증강이 완료되면 기술도면 50개를 각각 편집 유형별 검색 데이터셋(원본 50개와 편집데이터 8개)과 유사도(ResNet은 코사인 유사도, SIFT는 좋은 매칭 비율)를 계산하여 상위 10개를 추출함

· 추출 결과에서 원본으로부터 증강된 데이터 개수를 확인하여 8로 나눈 뒤 이 값을 대상 파일을 찾은 비율로 사용함. 즉 증강데이터 8개로부터 해당 모델이 찾아낸 건수를 매칭 결과 성능으로 사용함(예, 찾은 개수가 7개이면  $7/8=0.875$ )

## 4. 실험 결과 및 분석

### 4.1. 특허도면의 유형 분류 실험 결과

특허 도면의 유형분류 실험은 95.4%의 Accuracy를 보였다. 혼동행렬(confusion matrix)은 분류 모델의 성능을 평가하는 데 사용되는 도구로 대각선 방향의 수치(867, 429, 135, 846, 926)<표5> 참조)는 각 유형의 TP(양성을 양성으로 예측)를 나타낸다, TP를 제외한 가로축의 합은 FN(양성을 음성으로 예측), 세로축의 합은 FP(음성을 양성으로 예측)를 나타낸다.

유형별 성능에서는 그래프(유형 4)가 가장 높은 정밀도(0.98)와 F1 점수(0.98)를 나타냈다.

가장 많은 데이터를 사용했음에도 기술도면(유형 0)의 정밀도(precision)가 0.92로 비교적 낮은 이유는 기술도면의 다양성과 복잡성에 기인한 것으로 판단된다. 일정한 시각적 특징을 보이는 다른 유형과 다르게 기술도면은 매우 다양한 형태(장치의 디자인, 장치의 설계도 등)로 구성되어 있으며 특히 기술적 분류에 따라 다양성과 복잡성은 더 크게 나타난다. 염기서열(유형 2)의 경우 정밀도는 0.96으로 예측은 정확하나 재현율(recall)은 전체 지표 중에서 가장 낮은 0.89가 나왔다. 낮은 재현율로 인해 염기서열의 F1 점수 역시 0.92로 다른 유형보다 상대적으로 낮은 편이다. 이는 다른 유형보다 염기서열의 데이터 갯수가 매우 적었고 해당 유형이 충분히 학습되지 않아 오분류가 많이 나타난 결과로 볼 수 있다. 이같이 F1 점수는 정밀도와 재현율의 조화평균으로 데이터의 불균형을 나타내는 단서이기도 하다.

종합하면 일부 지표에서 상대적으로 낮은 값들이 있었지만, 전체적으로 특허도면의 유형분류 결과는 우수하다고 평가할 수 있다.

<표5 특허도면의 유형분류 성능 및 혼동행렬>

| 구분                | precision | recall | F1 score | Confusion Matrix  |
|-------------------|-----------|--------|----------|---|
| 기술도면(유형 0)        | 0.92      | 0.96   | 0.94     | <pre>[[867 12  0 18  3]  [ 14 429 4  5  3]  [  4  3 135  0 10]  [ 43  6  0 846  5]  [ 15  4  2  3 926]]</pre> |
| 화학구조(유형 1)        | 0.94      | 0.94   | 0.94     |   |
| 염기서열(유형 2)        | 0.96      | 0.89   | 0.92     |   |
| 순서도 및 다이어그램(유형 3) | 0.97      | 0.94   | 0.95     |   |
| 그래프(유형 4)         | 0.98      | 0.97   | 0.98     |   |

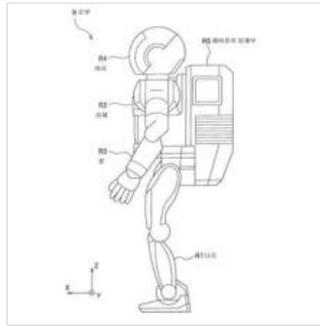
#### 4.2. 특허도면의 유사도 평가 및 검색 실험 결과

이미지 검색 실험은 다양한 경우를 가정하여 다음과 같이 진행했다.

- 검색 대상 데이터에 같은 이미지가 있는 경우, 같은 파일 이미지로 검색
- 검색 대상 데이터에 같은 이미지가 있는 경우, 해상도가 높은 질의 이미지를 사용하거나 이미지의 방향을 바꾼 질의 이미지를 사용하여 검색(90도로 회전, 상하 대칭, 좌우대칭으로 편집한 이미지 사용)
- 구글에서 내려받은 사진 이미지로 검색
- 검색 대상 데이터에 전혀 같은 모양의 이미지가 없는 경우의 검색

검색 대상 데이터에 포함된 이미지로 검색한 결과, ResNet-50을 이용한 검색 결과 5건(<표 6> 참조)을 보면 일치하는 이미지 2건을 검색하였으며 이미지가 같으면 유사도가 1.0(<표 6>의 검색 결과 1), 이미지의 미세한 차이가 있으면 0.974(<표 6>의 검색 결과2)로 나왔다. 검색 결과 3은 로봇의 형상이라는 측면에서 유사성을 나타내고 있으나 나머지 결과는 유사도는 높으나 질의 이미지와 많은 차이를 나타내고 있다.

<그림9 질의 이미지 사례>



<표6 ResNet-50을 이용한 검색 결과 (유사도 상위 5개)>

| 검색 결과1   | 검색 결과2  | 검색 결과3  | 검색 결과4  | 검색 결과5   |
|--|---|---|---|--|
|  |   |   |   |  |
| file:<br>kr1120050426685<br>99-pct00001-4.jpg<br>similarity: 1.000 | file:<br>kr00000754475b1<br>p-4.jpg<br>similarity: 0.9744 | file:<br>kr00000635593b1<br>p-4.jpg<br>similarity: 0.9640 | file:<br>kr00001274139b1<br>p-4.jpg<br>similarity: 0.9582 | file:<br>kr00001398880b1<br>p-4.jpg<br> <br>similarity: 0.9562 |

SIFT에서도 ResNet-50과 마찬가지로 같은 이미지 1건(<표 7>의 검색 결과1), 미세한 차이가 있는 이미지 1건(<표 7>의 검색 결과2)을 1, 2위로 찾았으며 검색 결과 3 이후에는 시각적 유사성을 찾아보기 어려운 결과를 확인했다. <표 6>의 검색 결과2는 시각적으로 미세한 차이가 있으나 good match ratio는 0.2480으로 매우 낮은 점수를 보였다.

<표7 SIFT를 이용한 검색 결과 (좋은 매칭 비율 기준 상위 5개)>

| 검색 결과1  | 검색 결과2   | 검색 결과3   | 검색 결과4   | 검색 결과5   |
|---|--|--|--|--|
|   |  |  |  |  |
| file:<br>kr1120050426685<br>99-pct00001-4.jpg<br>good match ratio:<br>1.0 | file:<br>kr00000754475b1<br>p-4.jpg<br>good match ratio:<br>0.2480 | file:<br>kr00001579036b1<br>p-4.jpg<br>good match ratio:<br>0.1650 | file:<br>kr00002247398b1<br>p-4.jpg<br>good match ratio:<br>0.1372 | file:<br>kr00002561338b1<br>p-4.jpg<br>good match ratio:<br>0.1266 |

다양한 경우를 가정한 이미지 검색 실험 결과, 검색 대상 데이터베이스에 같은 이미지가 있을 때 상위 10건 내에서 원하는 결과를 얻었으나 ResNet-50은 이미지 회전, SIFT는 이미지 해상도에 약한 것으로 나타났다.

<표8 ResNet-50과 SIFT를 이용한 검색실험 결과>

| 검색 실험 구분  | ResNet-50                                     | SIFT  |
|---|---|---|
| 검색 대상 데이터에 같은 이미지가 있는 경우                          | 같은 이미지나 미세한 차이가 있는 경우는 상위 검색 결과로 나타남. 즉 검색 성공 | 같거나 유사한 이미지는 상위 검색 결과로 나타나 검색 성공. 다만 질의 이미지의 해상도가 높을 때 ratio는 낮게 나타남          |
| 검색 대상 데이터에 같은 이미지가 있고 이미지의 방향을 바꾼 질의 이미지를 사용하여 검색 | 90도 회전과 상하 대칭의 경우 검색이 되지 않음                   | 이미지의 방향과 상관없이 모두 10위 내에 검색 성공. 다만, 해상도가 크고 상하 대칭인 이미지의 경우 매칭 결과를 찾지 못한 경우도 있음 |
| 사진 이미지(구글 검색)로 질의한 경우                             | 시각적으로 유사한 사진 이미지를 주로 검색함.                     | 사진 이미지를 주로 검색하였으나 시각적 유사성 없음  |
| 검색 대상 데이터에 유사한 이미지가 없는 경우                         | 검색 실패. 유사도 값에 매우 차이가 있는 이미지 검색                | 검색 실패. 시각적 유사성 없음   |

이미지 검색 실험의 두 번째 평가요인은 각각 상위 검색 결과를 만들어 낸 기준값에 관한 것이다. ResNet 인코딩 데이터의 유사성을 판단하는 코사인 유사도 값은 두 벡터의 방향이 다를 경우 최대 -1이 나올 수 있어 상위값으로 검색 결과를 판단할 때는 사용할 수 없다. 또한 SIFT의 good match ratio는 ResNet의 유사도와 달리 두 이미지의 속성을 비교한 결과로 두 이미지가 얼마나 잘 맞는지에 관한 값이다. 이미지의 회전에 대해서는 ratio 값이 큰 차이가 없는 반면 해상도에 따른 차이는 매우 크게 나타난다. 따라서 코사인 유사도와 좋은 매칭 비율만으로도 다 검색의 성공/실패를 나타내는 정량적인 기준으로 사용할 수 없음을 알 수 있다.

### 4.3. 증강데이터를 이용한 검색 실험 결과

증강데이터를 이용하여 유사도를 평가하고 매칭(검색) 결과를 이용하여 정량적 성능을 검증한 실험 결과는 다음과 같다.

<표9 ResNet과 SIFT의 이미지 매칭 결과>

| 구분1   | 구분2                              | ResNet  | SIFT    | (구분1) 평균 |
|-------|----------------------------------|---------|---------|----------|
| 부호 없음 | 부분 삭제(erase part)                | 87.75%  | 99.75%  | /        |
|       | 노이즈 추가(noise)                    | 7.25%   | 47.75%  |          |
|       | 크기 조정(resize)                    | 87.50%  | 98.75%  |          |
|       | 회전(rotate)                       | 79.75%  | 99.00%  |          |
|       | 이동(shift)                        | 100.00% | 100.00% |          |
|       | 부분 삭제 및 회전 (erase part & rotate) | 65.25%  | 97.50%  |          |
|       | (부호 없음 평균)                       | 71.25%  | 90.46%  | 80.85%   |

| 구분1   | 구분2                              | ResNet  | SIFT    | (구분1) 평균 |
|-------|----------------------------------|---------|---------|----------|
| 부호 있음 | 부분 삭제(erase part)                | 88.50%  | 99.75%  |          |
|       | 노이즈 추가(noise)                    | 16.75%  | 9.50%   |          |
|       | 크기 조정(resize)                    | 95.00%  | 95.50%  |          |
|       | 회전(rotate)                       | 80.00%  | 97.50%  |          |
|       | 이동(shift)                        | 100.00% | 100.00% |          |
|       | 부분 삭제 및 회전 (erase part & rotate) | 62.75%  | 95.50%  |          |
|       | (부호 있음 평균)                       | 73.83%  | 82.96%  | 78.40%   |
|       | 전체 평균                            | 72.54%  | 86.71%  | 79.53%   |

위 실험 결과는 ResNet과 SIFT 모두 이미지 편집 조건에 따라 매칭 성능이 매우 다양하게 나타나는 것을 보여주고 있다. 그뿐 아니라 실험 환경에 따라 매칭 성능이 다소 변경되기도 했는데 본 연구에서는 낮은 값을 사용하였다. 위 결과에 따르면 ‘부호없음’의 매칭 평균은 80.85%, 부호가 있는 원본 도면은 78.40%로 도면의 부호가 유사성에 미치는 영향은 예상보다 큰 차이가 없는 것으로 확인되었다. 이동(shift)의 경우는 이미지의 편집 없이 위치만 변경된 것으로 매칭에는 전혀 문제가 없는 것을 볼 수 있으며 ‘부분 삭제와 회전’에서는 ResNet의 매칭 평균이 대폭 낮아지는 것을 볼 수 있다. 또한 편집 유형에서 ‘노이즈 추가’ 평균이 가장 낮게 나타났는데 ResNet은 ‘부호없음’에서 7.25%, SIFT는 ‘부호있음’에서 9.5%를 나타냈다. 노이즈는 이미지의 각 픽셀값에 일정한 패턴을 추가하는 작업으로 정도에 따라 차이가 있겠으나 같은 위치에 있는 픽셀이나 에지와 같은 주요 특징에 왜곡을 발생시키는 점에서 다른 편집 유형보다 유사성 측정이 더 어려운 것으로 판단된다. 결론적으로 이동(shift)의 경우를 제외하고 대부분의 편집 유형에서 ResNet보다 SIFT의 매칭 성능이 좀 더 나은 성능을 보여주었다.

## 5. 결론

본 연구는 특허 도면의 시각적 유사성을 평가하고 유사도가 높은 검색 방법을 찾기 위해 딥러닝의 대표적인 모델 ResNet-50과 고전적인 컴퓨터비전 알고리즘 SIFT를 이용하여 다각적인 측면으로 도면의 유사성을 평가하는 연구이다. 특허도면의 5개 유형분류 실험결과는 기술도면의 정밀도가 가장 낮게 나타나 기술도면의 시각적 다양성과 복잡성이 큰 것으로 평가되었다. 두 개 모델을 이용한 기술 도면 검색 실험에서 ResNet-50은 이미지 회전에, SIFT는 이미지 해상도에 약한 것을 알 수 있었다. 마지막으로 증강데이터를 이용한 검색 및 매칭 실험에서는 전반적인 매칭 성능은 SIFT가 조금 더 높게 나타났고 두 개 모델 모두 noise가 있는 이미지에는 낮은 매칭 성능을 보였다. 즉, ResNet은 사전 훈련된 모델을 이용하여 추상적이고 전역적인 정보를 생성하므로 전반적으로 유사한 이미지를 유사도가 높은 것으로 평가한다. SIFT는 이미지의 특징점을 이용하여 직접적인 비교를 수행하므로 같은 모양의 이미지를 찾는 일에 더 강한 것을 알 수 있다.

특허 문헌의 유사성 평가 연구는 특허 기술 분류체계를 개선하고 선행기술 검색, 출원 기술의 신규성 분석, 특허 침해 방지 등 기술전문가의 정성적 분석에 의존했던 업무를 일부분 자동화하고 개선하는 데 기여하고 있다. 그러나 그동안 텍스트를 중심으로 이루어져 온 특허 문헌 분석에서 본 연구는 도면 이미지를 중심으로 도면 검색에 필요한 유사성 평가 방법과 기준을 찾고자

했다. 특허 도면은 기술의 중요한 정보 전달 수단으로서 해당 정보의 이해도와 활용성에 영향을 줄 수 있다. 따라서 특허 도면의 유사성 평가 연구는 특허 이미지에 대한 이해를 높이고, 자동화된 특허도면 검색, 장치 설계나 디자인에 대한 기술정보 제공, 위조 제품 감지 등 실제 응용 분야에서의 적용 가능성을 높일 수 있다. 그러나 실험 대상이 국내 특허, 일부 섹션에 국한되어 있고 대표도면만 대상으로 한 점, 도면의 유형에서 염기서열 유형 데이터의 부족, 테이블과 같은 유형의 도면이 제외된 점, 데이터 증강을 이용한 실험에서 좀 더 다양한 사례를 추가하지 못한 점 등은 본 연구의 한계라고 할 수 있다.

본 연구와 관련된 후속 연구에서는 국내외 특허의 다양한 섹션과 부족했던 유형의 데이터를 추가하여 데이터를 확장할 필요가 있다. 또한 도면 이미지와 텍스트를 함께 활용할 수 있는 연구도 필요하다. 최근 발전하고 있는 이미지 캡션 생성 모델은 이미지를 보고 그에 맞는 설명을 자동으로 생성하며 비주얼-언어 모델은 이미지와 텍스트를 동시에 처리하여 두 도메인 간의 관계를 학습하는 방식이다. 이미 이미지 캡션 모델을 이용하여 디자인 특허 이미지를 분류하는 연구사례도 있어 기술도면을 시각적 관점에서 더 상세히 분류하고 설명자료를 연결하면 도면 데이터를 대상으로 텍스트 기반의 검색이 가능하며 특정 기술의 시계열적 변화를 살펴볼 때도 유용한 솔루션이 될 수 있다. 즉 후속 연구에서는 시각적 유사성 평가에서 개념적 유사성을 평가하고 통합하는, 멀티모달 접근 방법의 유사성 평가 연구로 발전되어야 할 것이다.

## 참고 문헌(References)

### 단행본(국내 및 동양)

최형일, 「원리를 설명하는 컴퓨터비전 최강의 수업」, 초판, 홍릉, 2023.

### 단행본(서양)

Stefanos Vrochidis et al., “Enhancing patent search with content-based image retrieval”, Edit. by Paltoglou, G., Loizides, F., Hansen, P., Professional Search in the Modern World: COST Action IC1002 on Multilingual and Multifaceted Interactive Information Access, Springer, 2014.

### 학술지(국내 및 동양)

윤재웅 외 5인, “합성곱신경망(Convolutional Neural Network)을 활용한 지능형 유사상표검색 모형 개발”, 「경영과 정보연구」, 제38권 제3호(2019).

윤재웅 외 2인, “어텐션모형을 활용한 상표 이미지 내 도형 분류 모형 개발”, 「경영과 정보연구」, 제41권 제4호(2022).

한재석 외 3인, “딥러닝 기반 특허 도면 검색 연구”, 「한국정보과학회 학술발표논문집」, 2023년 6월(2023).

### 학술지(서양)

David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, Vol.60(2004).

Panagiotis Sidiropoulos et al., “Content-based binary image retrieval using the adaptive hierarchical density histogram. Pattern Recognition”, *Pattern Recognition*, Vol.44 No.4(2011)

Stefanos Vrochidis et al., “Towards content-based patent image retrieval : a framework perspective”, *World Patent Information*, Vol.32 No.2(2010).

Stefanos Vrochidis et al., “Concept-based patent image retrieval”, *World Patent Information*, Vol.34 No.4(2012).

Naeem Bhatti et al., “Contextual local primitives for binary patent image retrieval”, *Multimedia Tools and Applications*, Vol.77 no.7(2018).

### 학위논문(국내 및 동양)

유대근, “머신러닝을 활용한 특허상표이미지 분류”, 경북대학교, 석사, 2017.

박정범, “Patent Document Similarity based on Image analysis by using the SIFT-Algorithm and OCR-Text”, 배재대학교, 석사, 2017.

### 연구보고서

고광국, “개념기반 특허이미지 검색”, 한국과학기술정보연구원, 2012.

특허청, “명세서/도면의 작성요령 및 분할출원 제도 개선에 관한 연구”, 특허청 특허심사 정책과, 2012.

### 기타자료

Alla Kravets et al., “Patents images retrieval and convolutional neural network training dataset quality improvement”, In Proceedings of the IV International research conference “Information technologies in Science, Management, Social sphere and Medicine (ITSMSSM

- 2017)", Atlantis Press, 2017.
- Gabriela Csurka et al., "XRCE's Participation at patent image classification and image-based patent retrieval tasks of the Clef-IP 2011", CEUR Workshop Proceedings, Vol.1177, CEUR-WS, 2011.
- Jia Deng et al., "Imagenet: A large-scale hierarchical image database", In Proceedings of the 2009 IEEE conference on computer vision and pattern recognition, IEEE, 2009.
- Junaid Ahmed Ghauri et al., "Classification of Visualization Types and Perspectives in Patents", In Proceedings of the International Conference on Theory and Practice of Digital Libraries, Springer Nature Switzerland, 2023
- Juan Castorena et al., "Learning spatial relationships between samples of patent image shapes", 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2020.
- Kaiming He et al., "Deep residual learning for image recognition", In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2016.
- Michal Kucer et al., "DeepPatent: Large scale patent drawing recognition and retrieval", In Proceedings of the 2022 IEEE/CVF Winter Conference on Applications of Computer Vision, IEEE, 2022.
- Panagiotis Sidiropoulos et al., "Adaptive hierarchical density histogram for complex binary image retrieval", 2010 International Workshop on Content Based Multimedia Indexing (CBMI), IEEE, 2010.
- Shuo Jiang et al., "A convolutional neural network-based patent image retrieval method for design ideation", In Proceedings of the ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol.9, V009T09A039, The American Society of Mechanical Engineers, 2020.