JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 26, No. 5, pp. 1161-1169, May. 2025

AI를 활용한 멀티모달 분석 기반 디자인 언어 이해: 가우디 건축 사례를 중심으로

박 라 미[1] · 유 진[1] · 최 유 리[1] · 오 효 정[2*]
[1]전북대학교 기록관리학과 석사과정
[2]전북대학교 문헌정보학과 교수, 문화융복합아카이빙연구소 공동연구원

AI-Based Multimodal Analysis to Understand Design Languages: A Case Study of the Gaudi Architecture

Lami Park[1] · Jin Yoo[1] · You-Ree Choi[1] · Hyo-Jung Oh[2*]

[1]Graduate School, Department of Archive & Record Management, Jeonbuk University, Jeonju 54896, Korea
[2]Professor, Department of Library & Information Science, Jeonbuk University, Jeonju 54896, Korea

[요 약]

멀티모달 학습 모델은 서로 다른 유형의 데이터를 결합하여 단일 데이터 유형만으로는 포착하기 어려운 맥락과 의미에 대한 이해를 가능하게 한다. 이러한 접근법은 디자인, 예술, 건축과 같은 창의적 분야에서도 활용될 수 있으며, 작품과 공간에 담긴 의미를 보다 체계적으로 해석하는 데 기여하고 있다. 특히, 창작자의 의도와 철학은 단순히 시각적 요소에만 국한되지 않으며, 텍스트 맥락과 결합될 때 더욱 명확하게 해석될 수 있다. 본 연구는 AI 기반 멀티모달 기술을 활용하여 디자인적 요소를 언어로 해석하는 방법론을 제안하며, 안토니 가우디의 건축 예술을 적용 사례로 선정하였다. 이를 위해 먼저 대중 및 전문가의 인식을 반영한 텍스트 데이터를 바탕으로 가우디 디자인 어휘사전을 구축하였다. 이미지 분석에는 다중분류체계와 CNN 기반 이진 분류기를 구축, 가우디 이미지 분류별 모델의 정확도는 각각 Interior 모델 0.89, Exterior 모델 0.87, External appearance 모델 0.95로 나타났다. 이후 텍스트와 이미지 분석 결과 간의 교집합을 도출하여 디자인 요소 설명문을 생성하였다. 이러한 과정을 통해 멀티모달 기반 I2T 생성 방법론이 가우디 건축물의 핵심 요소를 효과적으로 분석하고 해석할 수 있음을 확인하였으며 나아가 AI 기반 예술 교육 및 디지털 문화유산 보존 등 다양한 분야에서 활용할 수 있는 모형을 제시하였다.

[Abstract]

Multimodal learning models integrate diverse data types to understand contexts and meanings difficult to capture using single-mode data. This study proposes a methodology for interpreting design elements using AI-based multimodal technology, with the Antoni Gaudi architecture as a case study. A Gaudi design vocabulary was built using text data reflecting public and expert perceptions. For image analysis, multi-classification systems and CNN-based binary classifiers were developed, achieving accuracy rates of 0.89 (interior), 0.87 (exterior), and 0.95 (external appearance). Design descriptions were generated by identifying intersections between text and image results. The study confirmed that multimodal I2T generation methods effectively analyzed and interpreted key elements of the Gaudi architecture, presenting a model applicable to AI-based art education and digital heritage.

색인어 : 멀티모달 분석, I2T 기법, 가우디 건축, 디자인 언어, 다중분류체계
Keyword : Multimodal Analysis, Image-to-Text Method, Gaudi Architecture, Design Language, Multi-Classification System

http://dx.doi.org/10.9728/dcs.2025.26.5.1161

Received 21 March 2025; **Revised** 11 April 2025
Accepted 16 May 2025

***Corresponding Author; Hyo-Jung Oh**

Tel: +82-63-270-3208
E-mail: ohj@jbnu.ac.kr

Ⅰ. 서 론

현대 디지털 환경에서는 텍스트, 이미지, 음성 등 다양한 형태의 데이터가 공존하며, 인공지능(이하 AI) 기술의 발전으로 이를 효과적으로 처리하고 통합하는 방법이 활발히 연구되고 있다. 그중에서도 멀티모달(Multi-modal) 학습 모델은 서로 다른 유형의 데이터를 결합하여 통합적으로 처리함으로써 단일 유형의 정보만으로는 포착하기 어려운 맥락과 의미를 보다 깊이 이해할 수 있도록 돕는다. 이러한 접근법은 디자인, 예술, 건축과 같은 창의적 분야에서도 활용될 수 있으며, 작품과 공간에 담긴 의미를 보다 체계적으로 해석하는 데 기여하고 있다. 특히, 예술작품 창작자의 의도와 철학은 단순히 시각적 요소에만 국한되지 않으며, 텍스트적 맥락과 결합될 때 더욱 명확하게 해석될 수 있다. 이는 작품에 담긴 의미를 다층적으로 이해하는 데 도움을 주며, 창작 과정에서 사용된 조형 원리나 미학적 개념 또한 보다 정교하게 이해할 수 있도록 한다. 따라서 디자인 및 예술적 의미 체계를 보다 심층적으로 탐구하기 위해서는 기존의 텍스트 중심 또는 이미지 중심의 단일 모달리티(modality) 접근을 넘어서는 융합적 분석 방법론이 필요하다. 그러나 기존의 예술분야와 접목한 멀티모달 연구는 주로 객체 인식이나 이미지 생성과 같은 기술적 측면에 초점을 맞추었으며, 디자인이나 건축물에 내재된 철학적·미학적 요소를 언어적으로 해석하려는 시도는 상대적으로 부족한 실정이다.

이에 본 연구는 AI 기반 멀티모달 기술을 활용하여 디자인 언어를 해석하는 방법론을 제안하고자 한다. 특히, 독창적인 디자인 철학을 보여주는 대표적인 사례로 안토니 가우디(Antoni Gaudi)의 작품을 선정하였다. 스페인의 건축가인 가우디는 유기적인 디자인과 독창적인 건축 양식을 창조한 선구자로 평가받는다. 그의 건축물은 자연에서 영감을 얻은 곡선과 혁신적인 구조를 결합하여 독보적인 개성과 아름다움을 선보인다. 또한, 가우디의 작품들은 단순한 건축물을 넘어 예술과 철학이 녹아든 공간으로서, 건축계뿐만 아니라 문화, 예술, 경제적 측면에도 큰 영향을 미쳤다[1]. 가우디의 작품이 지닌 다차원적 가치와 영향력을 고려할 때, 이를 보다 체계적으로 분석하고 이해해야 연구할 필요성이 제기된다.

따라서 본 연구의 궁극적인 목적은 가우디 건축물 이미지에 담긴 가우디만의 독창적인 요소를 언어로 표현, 즉 가우디만의 디자인적 요소를 어휘로 표현하는 것이다. 이를 위해 텍스트 및 이미지 데이터 분석 결과를 통합하여 가우디만의 디자인 언어를 해석하는 멀티모달 기반 I2T(Image-to-Text) 생성 기법을 개발하고자 한다. 본 연구에서 제안하는 멀티모달 기반 I2T 생성 기법은 향후 건축 교육, 디지털 문화유산 보존, 미술 교육 등 다양한 분야에서 활용될 수 있는 가능성을 지닌다. 이를 통해 디자인 사고와 언어적 이해를 위한 새로운 분석 틀을 제시하고 멀티모달에 기반한 예술 연구의 확장 가능성을 탐색할 수 있다.

Ⅱ. 이론적 배경 및 선행연구

2-1 멀티모달(Multi-Modal)

모달(Modal)이란 시각, 청각, 미각 등 감각을 의미하는 모달리티(Modality)의 줄임말로 이를 확장한 개념인 멀티모달(Multi-modal)은 여러 감각이 상호작용하는 현상을 의미한다. 특히, AI 분야에서는 텍스트, 오디오, 비디오 등 다양한 형태의 데이터를 통합적으로 처리하고 이해하며 생성할 수 있는 기술을 멀티모달이라 지칭한다[2]. 다양한 환경과 조건을 종합적으로 판단함으로써 고차원적인 문제를 해결하기 위해서는 여러 형태의 모달리티로부터 습득된 데이터를 학습하고 처리하는 멀티모달 접근 방식이 필요하다. 예를 들어, 이미지 분석만으로는 놓칠 수 있는 맥락적 정보를 텍스트 데이터를 통해 보완하거나, 텍스트로 표현하기 어려운 시각적 특징을 이미지 분석을 통해 포착할 수 있다[3]. 이러한 통합적 접근은 특히 건축물이나 예술작품과 같이 시각적 요소와 구조적 특성이 복합적으로 구성되어 있는 대상을 분석하는 데 효과적으로 활용될 수 있다.

본 연구에서는 텍스트의 맥락적 정보와 이미지에 담긴 시각적 요소를 언어로 표현함으로써, 디자인적 요소를 어휘로 변환하는 것을 목표로 한다. 이를 위해 다양한 형태의 데이터를 통합적으로 이해하고 분석할 수 있는 멀티모달 접근법을 연구의 방법론적 기반으로 설정하였다. 이러한 접근을 통해 디자인 요소를 보다 체계적으로 해석하고 분석할 수 있을 것으로 기대된다.

2-2 선행연구

선행연구로는 텍스트와 다른 유형의 데이터를 함께 활용하여 복합적인 정보를 분석하는 멀티모달 접근법을 적용한 연구를 중점적으로 살펴보았다. 황호현 외는 리뷰 데이터와 제품 정보를 이용한 멀티모달 감성분석 연구를 수행하였다. 이 연구에서는 텍스트 리뷰뿐만 아니라 제품의 카테고리 정보와 구매자가 입력하는 제품에 대한 메타정보를 활용하여 시퀀스 분류 모델 및 MLP(Multi Layer Perceptron)를 통한 멀티모달 감성분석 모델을 제안하였다[4]. 신미르와 신유현은 음성과 텍스트 데이터의 특징을 모두 결합한 멀티모달 학습으로 감정 분류를 진행하였다. 자연어 처리 분야에서 우수한 성능을 보인 BERT(Bidirectional Encoder Representations from Transformers) 및 파생 모델들의 성능을 비교하여 최적의 텍스트 추출 모델을 선정하고, 통합적 특성 학습을 통해 적은 데이터로도 높은 정확도를 달성할 수 있음을 입증하였다[3]. 김현아는 특허 문헌의 텍스트, 서지정보, 특허 도면을 통합 분석하여 유사성 평가의 정밀도를 향상시키는 멀티모달 기반 문헌 분석 모델을 제안하였다. 기존 단일 접근법의 한계를 극복하고 보다 정교한 분류 가능성을 제시한 점에서 의의

가 있다[5].

해외에서는 건축 및 도시 디자인 분야에서 멀티모달 분석을 적용한 연구를 살펴보았다. Veliz Reyes는 증강현실 기반의 건축 설계 교육 현장에서 물리적 모형, 디지털 자료, 인터뷰, 행위기반 상호작용 데이터를 통합하여 멀티모달 표현이 설계 이해에 미치는 영향을 분석하였다. 이를 통해 시각적·공간적 정보와 언어적 설명의 상호작용이 학습 효과를 높이는 데 기여함을 입증하였다[6]. Pistola 외는 도시 환경의 이미지성을 평가하기 위한 멀티모달 분석 시스템을 제안하였다. 이 연구는 시각적 장면 인식과 감정 분석, 색채 분석, 의미 분할 등의 시각적 정보와 함께 텍스트 분석을 결합하여 도시 이미지성 점수를 산출하였으며, 분석 결과를 텍스트 생성 알고리즘을 통해 도시 설계자가 활용할 수 있도록 구성하였다. 스페인의 도시 사례 연구를 통해 시스템의 활용 가능성도 입증하였다[7].

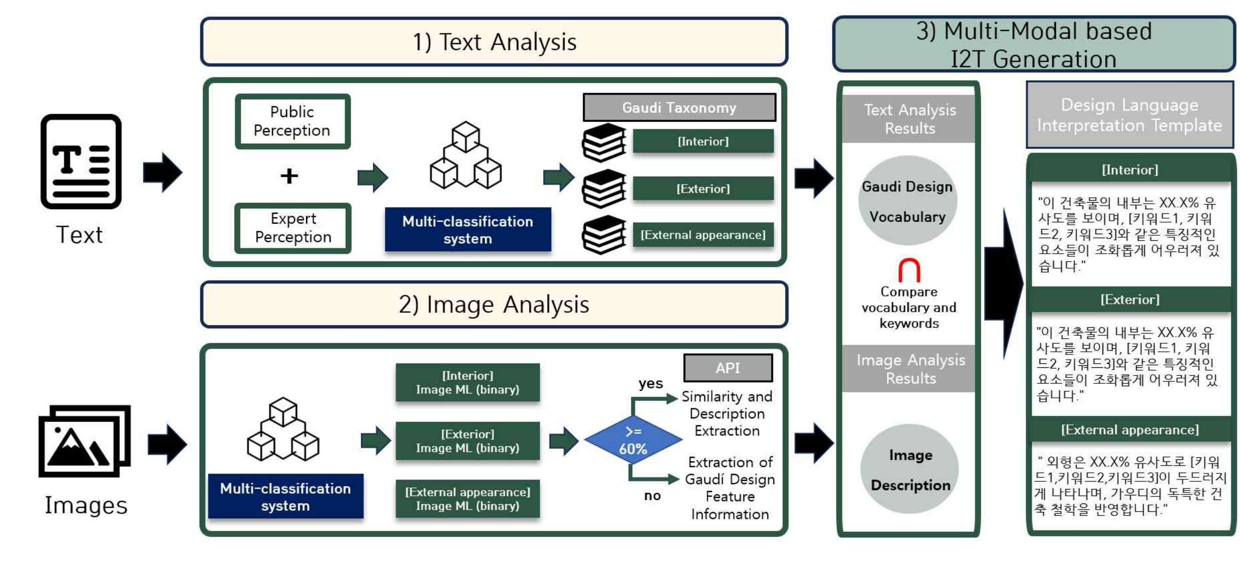
이상의 선행연구들과 같이 멀티모달 접근법은 단일 텍스트 기반 분석이 가진 한계를 극복하고, 다양한 유형의 데이터를 통합함으로써 정보 분석의 정확성과 효율성을 향상시키는 데 효과적임을 보여주고 있다. 그러나 대부분의 연구는 데이터를 병렬적으로 연결하거나, 분석 결과를 나열하는 데에 그치는 경우가 많다. 이와 달리 본 연구는 이미지에서 추출한 시각적 디자인 요소를 텍스트화하고, 이를 정형화된 언어 구조로 분석하여 의미를 재구성하는 과정을 포함한다는 점에서 차별성을 갖는다. 특히 시각적 특징과 텍스트 묘사를 유기적으로 연결함으로써, 기존 연구에서 다루지 못한 다층적 '디자인 언어'를 구성하고자 한다는 점에서 학문적 의의가 있다.

Ⅲ. 멀티모달 기반 디자인 언어 이해 모델

본 연구는 디자인 언어 이해에 있어 텍스트 분석과 이미지 분석의 멀티모달 접근법을 채택하였다. 그림 1은 본 연구의 분석방법을 도식화한 것으로, 먼저 텍스트 분석에서는 가우디 건축에 대한 대중의 인식(Public Perception)과 전문가의 인식(Expert Perception)을 구분하여 조사하였다. 대중 인식 분석에서는 블로그나 카페 등에서 일반인들이 가우디 건축에 대해 언급하며 사용한 어휘를 파악하였으며, 전문가 인식 분석에서는 관련 전문기관이나 학술 논문 등에서 출판, 보고된 문헌으로부터 가우디 건축에 대한 연구자들의 평론을 바탕으로 전문 용어를 추출하였다. 이후 건축물 유형에 따른 다중분류체계-내부(interior)와 외장(exterior), 외형(external appearance)-에 기반하여 각 문서 집합의 용어 가중치를 설정, 두 집합 간의 조합을 통해 가우디만의 디자인 어휘사전을 구축하였다. 또한, 수집된 문장들로부터 건축물에 대한 디자인적 요소를 설명하는 구문의 패턴을 정규화하여 템플릿으로 구축하였다.

이미지 분석에서는 앞서 언급한 다중분류체계를 기반으로 이진 분류 모델을 활용하여 가우디 건축 이미지를 탐지하는 방식으로 수행되었다. 주어진 이미지가 분류 모델에서 설정한 유사도 기준을 충족하는 경우와 그렇지 않은 경우에 따라 이미지 해석 API를 활용하여 해당 이미지의 설명문을 자동 생성하도록 설계하였다.

마지막으로, 다양한 데이터에 대한 특징을 동시에 결합할 수 있는 멀티모달 기반의 I2T Generation 기법을 적용, 이미



*It is a sentence generated based on a Korean taxonomy, and it is expressed in Korean because English translation is not possible.

그림 1. 분석방법 도식화

Fig. 1. Visualization of analysis method

지로부터 도출된 설명문과 텍스트 기반의 어휘사전을 비교하여 가우디만의 디자인 어휘를 교차 분석하였다. 분석 결과, 두 데이터에서 교집합이 존재하는 경우에는 구축된 디자인 요소 설명문의 템플릿을 활용하여 가우디만의 독창적인 디자인 어휘가 강조된 설명문을 생성하였다. 이러한 분석 과정의 세부 단계는 이어지는 하위 항목에서 상세히 설명한다.

3-1 텍스트 분석

1) 데이터 수집

본 연구에서는 먼저 텍스트 분석에 기반한 가우디만의 디자인 어휘사전을 구축하였다. 이를 위해 일반 대중의 인식과 전문가 집단의 어휘 표현을 각각 나누어 수집하였다. 기초 데이터 확보를 위한 웹 크롤링 검색어로 초기에는 “안토니 가우디”를 비롯해 “가우디”, “Gaudi” 등의 형태를 선정, 관련 웹 문서를 수집하였다. 그러나 수집된 문서 집합을 검토한 결과, 단순히 ‘가우디’라는 인명만으로는 가우디의 건축물에 대한 특징을 설명하는 글이 충분히 수집되지 못한다는 점을 보완, 검색어를 확대하였다. 확장된 검색어는 가우디의 대표 건축물인 까사밀라, 구엘 공원, 까사바뜨요, 사그리다 파밀리아 성당 등을 포함하였다. 가우디 작품에 대해 일반인들이 주로 언급하는 어휘를 수집하기 위한 대상으로는 네이버 블로그와 카페를 선정하였으며, 각각으로부터 상기 검색어를 포함하는 글 15,000건씩, 총 30,000건의 데이터를 확보하였다. 또한 디자인 전문가 집단의 어휘 표현을 포착하기 위해 가우디 월드 재단[9], 포털 가우디[10], 가우디 연구소[11] 등에서 출판, 게시한 연구 보고서와 간행물 150건을 수집하였다.

2) 다중분류체계 설계

수집된 기초자료를 분석한 결과, 건축물에 설명문들이 각각 건축 외형의 형태에 대한 특징을 설명하는 문구와 건축 외장재나 외부 장식을 설명하는 문구, 건물 내부의 조명이나 기둥, 가구 등 인테리어적인 요소를 설명하는 문장으로 구별됨을 확인하였다. 이는 실제 본 연구에서 궁극적으로 해석하고자 하는 가우디 건축물 이미지 역시 위 같은 분류 기준으로 나뉠 수 있음을 시사한다. 따라서 본 연구에서는 이러한 건축물의 유형에 따라 다음과 같이 세 가지 분류로 나누어 파악하였다. 먼저 기둥이나 계단, 가구 등과 같이 건축물의 내부 요소에 대한 특징은 내부(Interior)로, 타일이나 외장재, 지붕 등과 같이 외부 장식에 대한 특징은 외장(Exterior)으로 나누었다. 마지막으로 건축물을 외부에서 바라본 전체 형상이나 모양 등에 대한 설명은 외형(External appearance)으로 구분하였다.

3) 어휘사전 구축

가우디의 디자인 어휘사전 구축을 위해 수집된 대중인식 문서 집합과 전문가 문서 집합을 대상으로 형태소 분석을 수행하여 주요 명사를 추출하고, 각각의 빈도수를 계산하였다. 이후, 빈도수를 기준으로 상위 30개 용어를 선정한 뒤, TF-IDF(Term Frequency – Inverse Document Frequency) 가중치를 적용하여 중요도를 평가하였다. 마지막으로 두 문서 집합에서 도출된 점수를 합산하여 1차 가우디 어휘사전을 구축하였다. 구축된 어휘사전은 Interior, Exterior, External appearance의 세 가지 다중분류와 디자인/구조, 자연, 물리적 요소, 테마의 4개 범주로 체계화하는 과정을 거쳐 표 1과 같은 최종 가우디 어휘사전을 구축하였다.

표 1. 가우디 어휘사전
Table 1. Gaudi taxonomy

Category	Interior	Exterior	External Appearance
Design /Structure	스테인드글라스, 아치, 포물선, 웅장, 체크	아치, 포물선, 궁전, 벽돌, 발코니, 직물, 완공, 조각	고딕, 첨탑, 조각, 궁전, 벽, 벽돌,직물,세라믹, 웅장
	곡선, 타일, 모자이크, 장식		
Nature	해변, 해, 숲, 나무, 식물, 꽃	자연, 물결, 파도, 산, 가지, 비늘, 토끼	자연, 물결, 파도, 해, 해변, 꽃, 식물, 나무
	바다		
Physical Elements	밤, 물결, 파도, 세라믹, 조각	해골	해골, 번개, 밤
	빛, 뼈, 돌		
Theme	블루	동화, 환상, 용	동화, 지중해

*It is a keyword extracted from Korean blogs and cafes, and it is expressed in Korean because English translation is not possible.

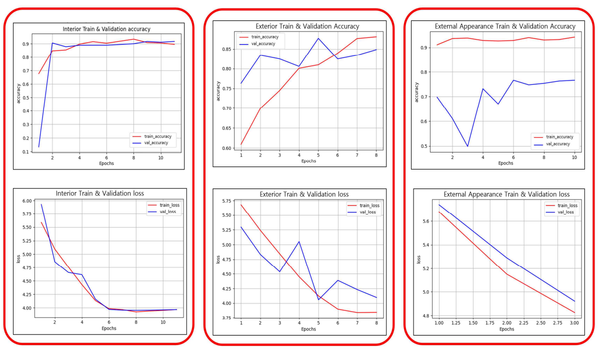
3-2 이미지 분석

가우디 디자인 요소 탐지를 위한 이미지 분석 역시 상기 텍스트 분석에서 적용한 다중분류체계에 기반한 분류 모델을 구축하였다. 활용한 알고리즘은 CNN(Convolutional Neural Network) 모델이며[12], 각 유형에 해당 여부를 판단하는 이진(binary) 분류 방식을 적용하였다. 학습 데이터는 유로피아나[13], 유네스코[14], 픽사베이[15] 등 학술 목적으로 저작권 사용이 허가된 이미지 전문 데이터베이스의 API를 통해 수집하였다.

표 2는 학습에 활용한 이미지 데이터의 통계정보로, 총 1,021건에 대한 분포도이다. 이미지 학습의 일관성을 위해 수집된 이미지에 대한 다음과 같은 전처리 과정을 수행하였다. 먼저 규격 및 형식을 통일하기 위해 데이터셋 크기를 224px로 조정 및 정규화(normalization)하였다. 이후 수집한 이미지 데이터 중 80%는 학습 데이터로, 나머지 20%는 테스트 데이터로 할당하였다. 또한, 모델 성능을 개선하기 위해 국가별 랜드마크 건축물 600건을 추가 수집하여 학습 데이터의 대조군으로 활용하였다. 데이터 증강은 Keras 패키지를 사용하여 수행하였으며, 모델이 에폭(Epoch) 단위로 실시간 변형된 이미지를 학습할 수 있도록 설정하였다[16].

표 2. 이미지 수집 데이터(건)
Table 2. Collected image data (cases)

	Interior	Exterior	External appearance	Total
Europeana	142	194	230	566
UNESCO	113	78	88	279
Pixabay	61	19	96	176
Total	316	291	414	1,021



(A) Interior　(B) Exterior　(C) External appearance

그림 2. 모델학습 손실 및 정확도
Fig. 2. Model training loss and accuracy

표 3. 최종 Epoch 손실 및 정확도
Table 3. Final Epoch loss and accuracy

	Train		Test	
	loss	accuracy	val_loss	accuracy
Interior	3.9754	0.8921	3.9644	0.9167
Exterior	3.8497	0.8711	3.0965	0.8483
External appearance	3.8518	0.9518	4.5702	0.6698

그림 2는 구축한 가우디 디자인 분류 모델의 결과를 시각화한 것이며, 표 3은 각 분류기의 최종 성능을 보여준다. 학습 결과, Interior 모델은 훈련 데이터와 검증 데이터에서모두 안정적인 성능을 보였다. 이는 실내 공간의 형태적 특성이 비교적 명확하고 일관된 패턴을 가지기 때문으로 해석된다. 반면, Exterior 모델의 경우 훈련 데이터에서 손실이 가장 낮았음에도 검증 데이터에서 상대적으로 낮은 정확도를 보였다. 이는 외부 구조의 다양한 형태적 특성이 충분히 반영되지 못했거나 곡선형 구조와 기하학적 패턴이 공존하는 경우가 많은 가우디 건축의 특성상 특정 패턴에 대한 학습이 편향되었을 가능성이 있음을 시사한다.

External appearance 모델은 훈련 데이터에서 가장 높은 정확도(0.95)를 기록했으나, 검증 데이터에서는 0.67로 급감하며 과적합(overfitting) 가능성을 보였다. 이는 가우디 건축에서 내부 기둥, 천장 곡선 등이 외부 형태와 유사한 디자인 언어를 공유하는 경우가 많기 때문으로 분석된다. 즉, 건물 내부의 형태적 특징이 외장 디자인에도 영향을 미치며, 이로 인해 내부 공간과 외장 요소 간의 경계가 모호해지는 경향이 학습 과정에도 반영된 것으로 보인다. 특히 동일한 건물에서도 상반된 색감이나 재질이 존재하여 학습 데이터와 검증 데이터 간의 일관성이 낮아짐으로 인해 모델이 일반화되지 못했을 가능성이 크다. 이러한 과적합 문제를 해결하기 위해 L2 정규화와 Dropout 기법을 적용하는 방안을 고려할 수 있다. L2 정규화는 모델의 가중치에 패널티를 부여하여 복잡성을 줄이고[8], Dropout은 학습 과정에서 무작위로 뉴런을 비활성화함으로써 특정 특성에 대한 의존도를 낮추어 일반화 성능을 향상시킬 수 있다[17]. 또한 추가 데이터 증강을 통해 외부 구조의 다양한 변형을 생성하여 훈련 데이터의 다양성

Gaudi Taxonomy

Category	Interior	Exterior	External Appearance
Design /Structure	스테인드글라스, 아치, 포물선, 웅장, 체크	아치, 포물선, 궁전, 벽돌, 발코니, 직물, 완공, 조각	고딕, 첨탑, 조각, 궁전, 벽,벽돌, 직물, 세라믹, 웅장
	곡선, 타일, 모자이크, 장식		
Nature	해변, 해, 숲, 나무, 식물, 꽃	자연, 물결, 파도, 산, 가지, 비늘, 토끼	자연, 물결, 파도, 해, 해변, 꽃, 식물, 나무
	바다		
Physical Elements	밤, 물결, 파도, 세라믹, 조각	해골	해골, 번개, 밤
	빛, 뼈, 돌		
Theme	블루	환상, 용, 동화	지중해, 동화

∩

Design Elements in Images

Classification	Sentence
Intersection O	**곡선 형태**: 벽과 구조적 요소는 부드럽고 유려한 선을 보여주며, 자연에서 발견되는 유기적인 형태를 불러일으키고 경직된 건축적 형태와는 다릅니다. **질감 있는 표면**: 벽은 광택 있는 **타일**과 거친 질감의 혼합을 특징으로 하여 촉각적 탐험을 유도하고 시각적 경험에 깊이를 추가합니다. **자연광**: 크고 독특한 형태의 창문이 내부로 황금빛을 가득 들여보내어, 하루에 따라 변화하는 **빛**과 그림자의 역동적인 놀이를 만들어냅니다. **건축 세부 사항**: 손으로 제작된 난간과 창틀을 포함한 장식된 나무 요소는 복잡한 **조각**을 보여주어 전체 디자인에 장인의 터치를 추가합니다.
Intersection X	**색상 팔레트**: 부드러운 파스텔 색상과 생동감 있는 파란색의 조화로운 혼합이 기발한 분위기를 강조하며, 따뜻하고 초대하는 느낌을 제공합니다. **복잡한 패턴**: 기하학적 형태가 예술적으로 배열되어, 관람자의 시선을 공간 전역으로 이끄는 리듬을 생성합니다. **아르 누보의 영향**: 디자인은 아르 누보 운동에서 일반적으로 발견되는 흐르는 모티프와 장식적 플루리시를 통합하여 우아함과 예술적 장인 정신을 강조합니다. **유동적인 전환**: 공간은 다양한 영역의 매끄러운 통합을 특징으로 하여, 벽이 천장으로 녹아들어가면서 넓고 친밀한 느낌의 연속성과 개방감을 느끼게 합니다. **상징적인 모티프**: 물고기 비늘 패턴과 꽃 디자인과 같은 요소는 자연을 반영하며, 유기적인 세계와의 연결을 강화하고 조화로운 감각을 구현합니다. **역동적인 구성**: 전체 레이아웃은 움직임의 감각을 만들어내며, 시선을 위로 그리고 주위로 끌어들여 탐험을 초대하고 관람자를 시각적 여정에 참여시킵니다.

*It is a sentence generated based on a Korean taxonomy, and it is expressed in Korean because English translation is not possible.

그림 3. I2T Generation 분석 결과
Fig. 3. I2T Generation analysis results

을 확보하는 것도 모델의 일반화 능력을 높이는 데 효과적일 것으로 예상된다. 실제로 Dropout 기법과 L2 정규화를 External appearance 모델에 적용한 결과, 정확도는 기존 0.67에서 0.78로 상승하여 성능이 소폭 개선되었음을 확인할 수 있었다.

한편, 일반적으로 건물의 외부 형태가 외관 장식에도 영향을 미치기 때문에 Exterior와 External appearance 모델은 유사한 특성을 학습한 경향이 나타났다. 예를 들어, 건물이 물결치는 파도의 형태를 반영하여 설계되었을 경우 디자인 요소가 단순히 건물 외형에서만 나타나는 것이 아니라, 외장에서도 유사한 패턴이 적용되는 경향이 있었다. 이와 같은 현상은 어휘사전(Taxonomy)에도 반영되었으며 특정 디자인 요소가 내부와 외부, 외관에서 동일하게 나타나는 모습을 확인할 수 있었다. 이렇게 학습한 모델을 기반으로 이미지 간 유사도를 측정하였으며, 이후 단계에서는 이미지-텍스트 변환(I2T) 기법을 적용하여 가우디 건축 디자인의 언어적 해석을 진행한다. 이에 대한 구체적인 내용은 다음 절에서 서술한다.

3-3 멀티모달 기반 I2T 생성

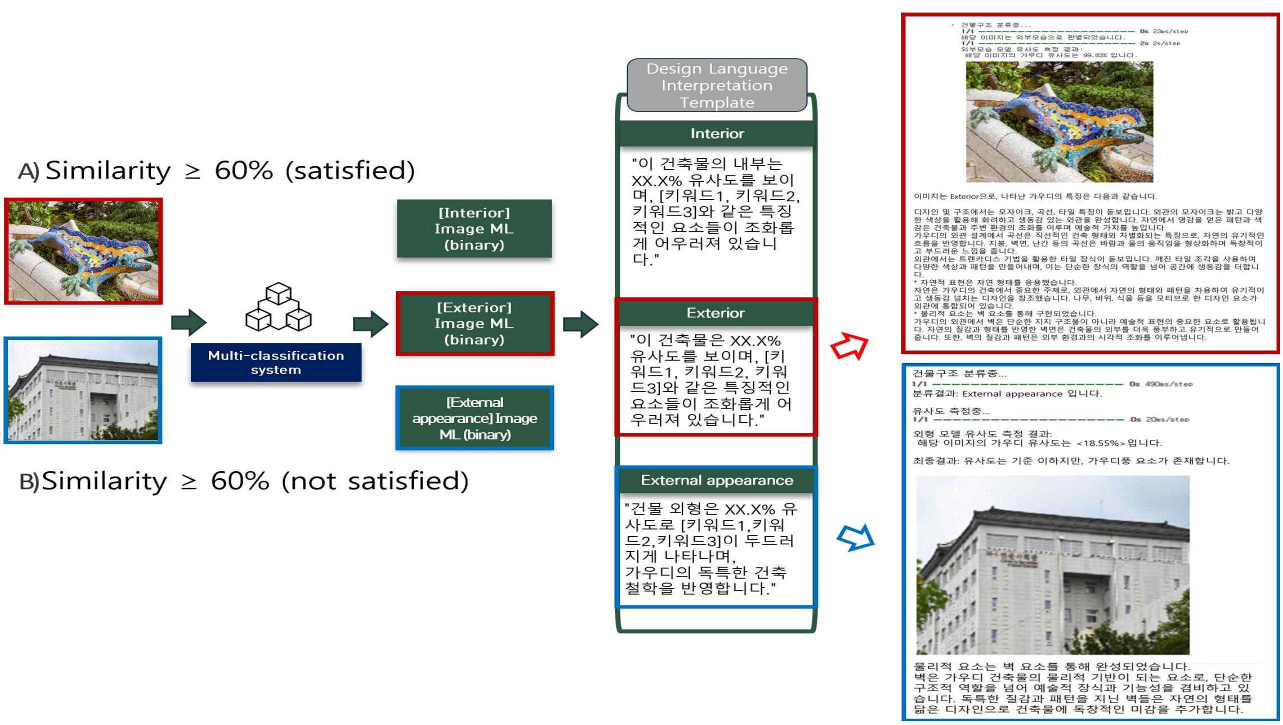
본 연구에서는 상기한 텍스트 분석 결과와 이미지 분석 결과를 결합해 가우디 디자인 언어를 설명해주는 멀티모달 기반의 I2T(Image-to-Text) 생성 기법을 고안하였다. 먼저, 이미지 해석 API(Google Gemini, Chat GPT)를 이용하여 주어진 이미지의 시각적 특징을 추출하고 이를 바탕으로 주요 키워드와 설명문을 자동 생성하였다. 이후 이미지 내 디자인 요소와 가우디 디자인 어휘사전 간의 교집합을 도출하는 방식을 적용하여 의미적으로 연관된 텍스트를 생성하도록 하였다.

그림 3은 입력 이미지 내 디자인 요소와 가우디 어휘사전의 키워드를 비교한 것이다. 동일한 키워드가 Interior, Exterior, External appearance 모델에서 중복 출현하는 경우, 문맥에 맞게 해석될 수 있도록 템플릿 기반의 연결 구조를 설정하였다. Interior는 내부 공간의 특징이 강조되도록, Exterior는 지붕, 벽면, 난간 등의 요소가 부각되도록, External appearance는 건축물 전체의 형태와 흐름이 강조되도록 구성하였다. 더불어 가우디 어휘사전을 기반으로 '디자인/구조', '자연', '물리적 요소', '테마'의 네 가지 범주로 템플릿을 세분화하였다. 이를 통해 입력 이미지에서 탐지된 가우디 디자인 언어 키워드가 어떤 건물 구조에 속하며, 어떤 관점에서 분석되었는지 명확히 파악할 수 있도록 하였다.

Ⅳ. 디자인 언어 이해 활용모형(Use Case)

4-1 디자인 언어 이해 해석 결과

그림 4는 입력 이미지에 따른 가우디 디자인 언어 이해 결과에 따른 해석문을 예시한 것이다. 그림 4의 A 이미지는 스페인 바르셀로나 구엘공원에 위치한 도마뱀 조각상 이미지이



*It is a sentence generated based on a Korean taxonomy, and it is expressed in Korean because English translation is not possible.

그림 4. 이미지 유사도 출력 결과 예시

Fig. 4. Example of image similarity output results

고, B 이미지는 J대학교 건물 이미지로, 가우디풍 이미지 유사도는 각각 99.83%와 18.55%로 판명되었다. 한편, 가우디풍 디자인 건축물 이미지와 일반 건축 이미지 간 구분의 명확성을 확보하기 위한 사전실험 결과, 그림 5와 같이 50% 기준은 과도한 오분류를, 70% 기준은 지나치게 엄격한 필터링을 유발하였기에 이 둘 사이에서 가장 균형 잡힌 결과를 보인 60%를 기준치로 채택하였다.

Image	Similarity Score	Ground truth	Threshold		
			50%	60%	70%
	72%	Gaudi	O	O	O
	54%	Not Gaudi	X	O	O
	63%	Gaudi	O	O	X

* O: True, X: False

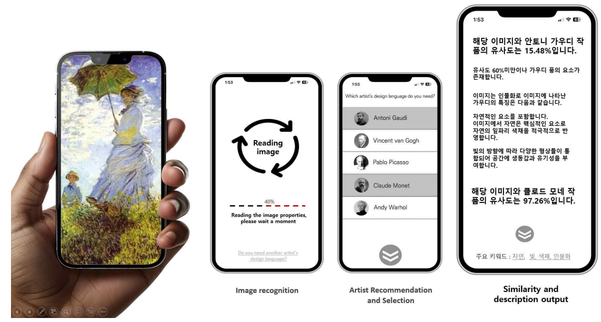
그림 5. 유사도 임계값에 따른 이미지 분류 결과 비교
Fig. 5. Comparison of classification results by similarity thresholds

먼저, 입력 이미지(A)는 다중분류체계를 통해 Exterior로 식별되었으며, Exterior 분류 모델에서 측정된 가우디 이미지 유사도는 99.83%로 기준치(60%)를 충족하였다. 이후 이미지 해석 API를 통해 설명문이 출력되고 가우디 어휘사전과의 I2T 생성 과정을 거쳐 Exterior 템플릿 형식에 맞춘 정보가 구성되었다. 최종적으로 '외관', '지붕', '벽면', '난간'과 같은 Exterior 템플릿에 기반해 '모자이크', '곡선', '타일', '트렌카디스 기법' 등의 가우디만의 독창적인 키워드로 해석된 설명문이 도출됨을 알 수 있다.

반면 이미지(B)는 입력 후 다중분류체계를 통해 External appearance로 식별되었으나, 유사도가 18.55%로 기준치(60%)보다 낮게 측정되었다. 다만 일부 건물 특징을 설명함에 있어 가우디 디자인 분류 모델 결과에서 가우디 디자인 요소가 포함된 것으로 판별되었으며, 이에 따라 "유사도는 기준 이하이나 가우디풍 요소가 존재합니다"라는 보조 정보가 제시되었다. 이후 가우디 어휘사전과의 교집합(I2T Generation) 과정을 거쳐 템플릿에 맞게 정보가 구성되었으며, 가우디 디자인의 물리적 요소 유사점인 '벽'의 설명문이 최종 템플릿으로 출력되었다.

이처럼 본 연구는 단순히 이미지 유사도에 의존하지 않고, 유사도가 낮더라도 세부 디자인 요소와의 연결성을 분석하여 설명문을 생성하는 유연한 분석체계를 제시한다는 점에서 기존 분석 방식과의 차별성을 보여준다. 이는 향후 건축 또는 예술작품의 자동 해석 시스템 개발에 있어 보다 정밀하고 직관적인 데이터 기반 지원 도구로 활용될 수 있는 가능성을 제시한다.

4-2 이용자 활용모형



*It is a sentence generated based on a Korean taxonomy, and it is expressed in Korean because English translation is not possible.
그림 6. 디자인 언어 이해 앱 목업
Fig. 6. Application mock-up

본 연구에서 제안한 디자인 언어 이해 모델은 예술작품의 해석 및 전시와 같은 다양한 분야에서 활용될 수 있다. 그림 6은 본 연구의 결과를 실시간 이미지 분석 서비스와 연계하여 모바일 애플리케이션으로 확장한 예시이다. 해당 애플리케이션은 이미지 인식, 예술가 추천 및 선택, 유사도 분석 및 설명문 출력의 단계를 거쳐 실행된다.

이 시스템은 사전에 구축한 예술작품 데이터셋을 기반으로 학습된 다중분류체계를 통해, 특정 작가의 스타일을 식별하고 분석하는 방식으로 작동한다. 이를 위해 색상, 구도, 질감, 선의 활용, 대상 등의 다양한 시각적 요소를 평가하여 사용자가 업로드한 이미지가 어느 작가의 스타일과 유사한지 판단한다. 먼저, 이미지 인식 단계에서 사용자가 사진을 촬영하거나 특정 이미지를 애플리케이션에 업로드하면, 학습된 다중분류체계를 통해 자동으로 분류가 이루어진다. 이후 예술가 추천 및 선택 단계에서는 분류된 이미지와 스타일이 유사한 예술가 목록을 제공하며, 사용자는 이를 참고하여 특정 작가를 선택할 수 있다. 마지막으로 유사도 분석 및 설명문 출력 단계에서는 이진 분류 모델을 활용하여 업로드된 이미지와 선택된 작가의 스타일 유사도를 평가한다. 동시에 이미지 해석 API를 활용해 생성된 설명문과 주요 키워드를 출력한 후, 이를 사전에 구축한 작가별 어휘사전과 비교·분석하여 최종적으로 공통된 문장과 키워드를 도출한다. 또한 이미지에서 추출된 주요 키워드를 기반으로 동일한 키워드를 공유하는 작품을 추가로 추천하도록 설계하였다. 이를 통해 사용자는 단순한 스타일 매칭을 넘어 특정 작가의 특징이 어떤 작품에 어떻게 반영되었는지를 보다 심층적으로 탐색할 수 있다.

본 연구에서 제안한 멀티모달 기반 이미지 분석 시스템은 예술 교육 분야에서도 적용될 수 있다. 특히 미술 교육에서 기존의 개념 중심 학습을 보완하고, 시대별 미술사조의 특징을 이해하는데 있어 본 시스템이 보조적인 학습 도구의 역할을 할 수 있다. 나아가 본 시스템은 교육적 활용을 넘어 미술 감상과 디지털 아카이브 구축 등 다양한 분야에서도 활용될

수 있다. 이를 통해 사용자는 특정 작품과 스타일이 유사한 예술가나 작품을 탐색할 수 있으며, 기술의 발전에 따라 미술 작품의 자동 분류 및 메타데이터 생성과 같은 응용이 가능할 것으로 예상된다.

Ⅴ. 결 론

본 연구는 안토니 가우디의 건축물에 담긴 독창적인 디자인 요소를 분석하고, 이를 디자인 언어로 받아들이고 해석하였다. 가우디 건축의 시각적 특징을 효과적으로 분석하기 위해 멀티모달 접근 방식을 적용하여 텍스트 데이터와 이미지 데이터를 결합하는 방법론을 도출하였다. 특히 I2T(Image-to-Text) 생성 기법을 고안하여 이미지에서 의미 있는 정보를 추출하고 이를 정교한 어휘사전과 비교하는 과정을 설계함으로써 설명문의 정확도를 높였다.

본 연구는 가우디라는 특정 건축가의 작품을 중심으로 진행되었기 때문에, 다양한 건축 양식이나 예술가들에게 일반화하기에는 한계가 존재한다. 또한 텍스트와 이미지 데이터에 기반한 제한적인 멀티모달 접근으로 인해 음향, 공간감 등 건축물의 다차원적 경험을 충분히 반영하지 못한 한계가 있다. 이를 보완하고자 가우디 건축물의 시각적 디자인 요소를 체계화한 어휘사전을 구축하였으나, 해당 어휘사전 역시 건축물이 내포한 문화적·역사적 의미까지는 반영하지 못하였다. 이로 인해 가우디 건축에 대한 총체적 특징을 완전히 파악하는 데 한계가 있으며, 후속 연구에서는 건축물이 형성된 문화적·시대적 맥락을 반영한 심층 분석을 통해 보다 종합적인 이해를 도모할 필요가 있다.

향후 연구에서는 본 연구에서 제안한 AI 활용 분석 도구를 미술 교육에 접목하는 방안을 모색하고자 한다. 특히, 중학교 미술 교과과정에서 학습 도구로서의 이용 가능성을 탐색하고, 이를 통해 AI 활용 분석 도구가 학습자의 예술적 이해도를 향상시키는 데 실질적으로 기여할 수 있는지 검증할 예정이다. 또한 연구 결과를 바탕으로 보다 효과적인 학습 도구를 개발하고, AI를 활용한 예술 교육이 실제 교육 환경에서 어떤 역할을 할 수 있을지 검토하고자 한다.

본 연구는 가우디 건축의 디자인 요소를 분석하는 새로운 방법론을 제시함과 동시에 건축 및 예술 분야에서 멀티모달 분석방법의 기초를 마련했다는 점에서 의의를 가진다. 특히, 기존 연구들이 건축 양식을 주로 정성적 평가 또는 시각 자료의 분류 수준에서 머물렀던 반면, 본 연구는 AI 기반 멀티모달 기술을 통해 구체적인 시각 정보와 언어 데이터를 통합적으로 분석함으로써 데이터 기반의 정량적 비교가 가능해졌다는 점에서 실질적인 분석 도구로서의 활용 가능성을 입증하였다. 더불어 기존의 건축 분석 연구들이 주로 텍스트 중심의 서술적 방식이나 단순 이미지 분류에 의존했던 반면, 본 연구는 AI 기반의 이미지 인식과 텍스트 생성을 유기적으로 결합하여 건축 양식 분석에 새로운 패러다임을 제시하였다는 점에서 차별성을 갖는다. 이러한 접근 방식은 특정 예술가의 스타일을 보다 체계적으로 분석할 수 있도록 지원하며, 향후 다양한 예술작품 해석 및 시각 데이터 분석 연구에서도 전이 학습(Transfer Learning), 사전훈련(Pre-training) 등 AI 모델 개발에 활용될 수 있는 확장 가능성을 지닌다. 특히 이미지 기반 예술 연구에 텍스트 해석을 결합하는 방식이 가지는 학술적·실용적 가치를 확인했다는 점에서 연구의 의미가 있다. 향후 연구에서는 본 연구의 한계를 보완하고, 분석 방식을 보다 정교화하여 데이터 해석의 새로운 가능성을 탐색하고 더욱 풍부한 학술적 통찰을 제공할 수 있기를 기대한다.

감사의 글

본 논문은 2025년도 전북대학교 연구기반 조성비 지원에 의하여 연구되었음.

참고문헌

[1] R. W. Kim, “About Antoni Gaudi,” *Journal of Sookmyung Design Research*, Vol. 35, pp. 11-16, 2023.
[2] D. H. Kwon, “Analysis of Prompt Elements and Use Cases in Image-Generating AI: Focusing on Midjourney, Stable Diffusion, Firefly, DALL·E,” *Journal of Digital Contents Society*, Vol. 25, No. 2, pp. 341-354, 2024. https://doi.org/10.9728/dcs.2024.25.2.341
[3] M. Shin and Y. Shin, “Efficient Emotion Classification Method Based on Multimodal Approach Using Limited Speech and Text Data,” *The Transactions of the Korea Information Processing Society*, Vol. 13, No. 4, pp. 174-180, 2024.
[4] H. Hwang, K. Lee, J. Yu, and Y. Lee, “Multimodal Sentiment Analysis Using Review Data and Product Information,” *The Journal of Society for e-Business Studies*, Vol. 27, No. 1, pp. 15-28, 2022. https://doi.org/10.7838/jsebs.2022.27.1.015
[5] H. Kim, A Study on Patent Document Similarity Evaluation and Classification Precision Enhancement Using a Multimodal Approach, Ph.D. Dissertation, Graduate School of Soongsil University, Seoul, 2024.
[6] A. Veliz Reyes, “A Multimodal Study of Augmented Reality in the Architectural Design Studio,” *International Journal of Technology and Design Education*, Vol. 35, No. 1, pp. 283-303, April 2024. https://doi.org/10.1007/s10798-024-09895-5
[7] T. Pistola, N. Georgakopoulou, A. Shvets, K. Chatzistavros,

V.-R. Xefteris, A. Táboas García, ... and L. Kompatsiaris, "Imageability-Based Multi-modal Analysis of Urban Environments for Architects and Artists," in *Proceedings of ICIAP 2022 Workshops*, Lecce: Italy, pp. 198-209, May 2022. https://doi.org/10.1007/978-3-031-13321-3_18
[8] P.-H. Jeon, "A Study on Polynomial Neural Networks for Stabilized Deep Networks Structure," *The Transactions of The Korean Institute of Electrical Engineers*, Vol. 66, No. 12, pp. 1772-1781, 2017. https://doi.org/10.5370/KIEE.2017.66.12.1772
[9] Gaudí Foundation. Gaudí Foundation Official Website [Internet]. Available: www.gaudifoundation.org.
[10] Portal Gaudí. Portal Gaudí [Internet]. Available: www.portalgaudi.org/en/.
[11] Colònia Güell. Descubre la Cripta Gaudí de la Colonia Güell, el tesoro mejor guardado de Gaudí [Internet]. Available: www.gaudicoloniaguell.org/.
[12] Keras. Models API [Internet]. Available: https://keras.io/api/models/.
[13] Europeana. Discover Europe's Digital Cultural Heritage [Internet]. Available: www.europeana.eu.
[14] UNESCO World Heritage Convention. Works of Antoni Gaudí [Internet]. Available: whc.unesco.org/en/list/320/
[15] Pixabay. Free Images & Videos [Internet]. Available: www.pixabay.com/ko/.
[16] TensorFlow. ImageDataGenerator [Internet]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator
[17] Keras. Dropout Layer [Internet]. Available: https://keras.io/api/layers/regularization_layers/dropout/

박라미(Lami Park)

2024년 : 전북대학교 문헌정보학과 졸업

2024년～현　재: 전북대학교 기록관리학과 석사과정
※관심분야 : 데이터 분석, 장기보존

유진(Jin Yoo)

2024년 : 전북대학교 문헌정보학과 졸업

2024년～현　재: 전북대학교 기록관리학과 석사과정
※관심분야: 데이터 분석, 텍스트 마이닝

최유리(You-Ree Choi)

2018년 : 전북대학교 경영학과 졸업

2024년～현　재: 전북대학교 기록관리학과 석사과정
※관심분야 : 데이터 분석, 텍스트 마이닝

오효정(Hyo-Jung Oh)

2008년 : 한국과학기술원 컴퓨터공학과 공학박사

2000년～2015년: 한국전자통신연구원 지능정보연구본부 책임연구원
2015년～현　재: 전북대학교 문헌정보학과 교수, 문화융복합아카이빙연구소 공동연구원
※관심분야 : 정보검색, 텍스트마이닝, 빅데이터정보처리 등