

딥러닝 기반 원전 설계 및 건설 문서의 화학조성 테이블 자동 검출

이상훈*, 문성빈**¹, 오영진**², 박상원***, 이해연****

Deep Learning-based Automated Detection of Chemical Composition Tables in Design and Construction Documents of Nuclear Power Plants

Sang-Hoon Lee*, Seong-Bin Mun**¹, Young-Jin Oh**², Sang-Won Park***, and Hae-Yeoun Lee****

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. RS-2022-KP002852)

요약

원자력 발전소의 배관감육 관리를 위해서는 정확한 화학조성 데이터 확보가 필수적이다. 그러나 이 정보는 수천 페이지에 달하는 품질증빙서류에 흩어져 있으며, 현재까지는 수작업에 크게 의존해 왔다. 본 연구에서는 이러한 비효율성을 해소하기 위해, 품질증빙서류에서 관심 페이지만을 분류하고 페이지 내 화학조성 정보가 담긴 테이블 영역을 정밀 검출하는 2단계 자동화 딥러닝 파이프라인을 제안한다. 제안한 파이프라인은 다양한 문서 양식과 복잡한 내포 테이블 구조에서도 강건하게 동작하도록 설계되었다. 실험 결과, 페이지 분류는 99% 이상의 정확도, 테이블 검출은 mAP50-95 기준 0.85 이상의 성능을 달성하였다. 이러한 성과는 배관감육 평가에 필요한 핵심 정보 추출 과정을 자동화하고, 향후 완전한 데이터베이스 구축을 위한 기반 기술을 제시한다는 점에서 의의가 있다.

Abstract

Accurate chemical composition data is essential for managing pipe wall thinning in nuclear power plants, but this information is dispersed across thousands of pages in Quality Verification Documents (QVDs) and largely relies on manual inspection. To address this inefficiency, we propose a two-stage automated deep learning pipeline that classifies relevant pages and detects table regions containing chemical composition information, designed to be robust across diverse document formats and complex nested table structures. Experimental results show page classification accuracy over 99% and table detection mAP50-95 above 0.85, highlighting the pipeline's effectiveness in automating key information extraction for pipe wall thinning assessment and providing a foundation for a fully automated database.

Keywords

deep learning, classification, object detection, pipe wall thinning, QVD, CMTR

* 국립금오공과대학교 디지털융합공학과 박사과정
한국전력기술(주) 디지털솔루션연구소

- ORCID: <https://orcid.org/0009-0000-9828-1063>

** 한국전력기술(주) 디지털솔루션연구소

- ORCID¹: <http://orcid.org/0009-0009-1048-2473>

- ORCID²: <http://orcid.org/0000-0001-7187-9997>

*** 국립금오공과대학교 컴퓨터공학전공 학사과정

- ORCID: <http://orcid.org/0009-0007-2039-7067>

**** 국립금오공과대학교 컴퓨터소프트웨어공학과 교수(교신저자)
- ORCID: <http://orcid.org/0000-0002-6081-1492>

• Received: Aug. 29, 2025, Revised: Sep. 16, 2025, Accepted: Sep. 17, 2025

• Corresponding Author: Hae-Yeoun Lee

Dept. of Computer Software Engineering, Kumoh National Institute of Technology, Korea

Tel.: +82-54-458-7548, Email: haeyeoun.lee@kumoh.ac.kr

1. 서 론

원자력 발전 분야는 높은 안전성과 엄격한 품질 보증 기준이 요구되는 산업 분야로, 설계와 건설 및 장기간 운영 유지보수 과정에서 많은 문서가 생산되고 검증되어야 한다[1]. 이러한 문서는 발전소의 구조적 건전성과 안전성을 보증하는 핵심 자료로 활용된다.

원자력 발전소에서 관리해야 할 주요 관심사 중 하나는 배관감육(Pipe wall thinning)이다. 배관감육이란 발전소의 배관이 장기간 운영됨에 따라 내부 유체의 흐름이나 부식, 침식 등으로 인해 배관 두께가 점차 얇아지는 현상으로, 배관의 구조적 안전성을 저해할 수 있어 엄격한 관리가 요구된다. 따라서 배관의 내구성을 평가하는 데 중요한 화학조성 데이터를 정확히 확보하는 것은 필수적이다[2].

원전 설계 및 건설 단계에서 생산되는 다양한 문서 중, 원자력 발전소에 공급되는 부품 및 장비의 품질을 증명하는 품질증빙서류(QVD, Quality Verification Document)가 있다. 화학조성 데이터는 보통 이 품질증빙서류 내 재료시험성적서(CMTR, Certified Material Test Report)에 포함된다. 그러나 국내 원자력 발전소는 대부분 디지털 전환 이전에 건설되었으며, 안전성 등의 이유로 여전히 문서가 원본 또는 스캔 형태로 보관되는 경우가 많다. 품질증빙서류 역시 예외가 아니며, 수천 페이지에 달하는 한 세트의 문서 중 실제 분석에 필요한 부분은 약 20~25%에 불과하다.

더욱이 해당 페이지에서도 화학조성 정보는 일반적인 표 형식이 아닌 표 안의 표 형태, 즉 ‘내포된 테이블(Nested Table)’ 같은 복잡한 구조로 존재하는 경우가 많다. 이로 인해 작업자는 방대한 문서 더미에서 분석 대상 페이지만을 먼저 수동으로 선별한 후, 다시 해당 페이지 내에서 국소적인 테이블 영역을 육안으로 확인하고 데이터를 수기로 입력해야 하는 이중의 비효율성에 직면하고 있다.

최근 4차 산업 혁명을 필두로 한 디지털 전환(DX, Digital Transformation) 기술, 특히 인공지능(AI, Artificial Intelligence) 기반 문서 이해 기술의 발전은 이러한 노동 집약적 업무 환경을 혁신할 돌파구를 제공하고 있다[3]. 이에 본 연구에서는 앞서 언급된

수작업의 비효율적인 문제를 해결하기 위해, 다음과 같은 2단계 자동화 딥러닝 파이프라인을 제안한다.

첫 번째 ‘페이지 분류’ 단계는 CNN(Convolutional Neural Network) 기반의 이미지 분류 모델을 사용하여, 방대한 품질증빙서류에서 분석이 필요한 화학조성 관련 페이지만을 신속하게 선별한다. 두 번째 ‘관심 영역 검출’ 단계에서는 선별된 페이지에만 객체 탐지(Object detection) 모델을 적용하여, 비정형적이고 복잡한 구조 속에서도 화학조성 정보가 담긴 특정 테이블 영역을 정밀하게 검출한다.

본 논문의 구성은 다음과 같다. 2장에서는 이미지 분류 딥러닝 기술과 객체 탐지 기반 문서 구조 분석 및 이해 관련 선행연구를 종합적으로 검토한다. 3장에서는 제안하는 2단계 자동화 딥러닝 파이프라인을 상세히 기술한다. 4장에서는 제안 기법의 유효성을 검증하기 위한 실험 설계 및 정량적 성능 평가 결과를 분석한다. 마지막으로 5장에서는 본 연구의 결론을 요약하고, 한계점과 향후 연구 방향을 논의한다.

II. 관련 연구

2.1 이미지 분류 딥러닝 기술

딥러닝 기반 이미지 분류는 이미지 내의 객체를 인식하고 범주화하도록 모델을 학습시키는 기술로, 다양한 분야에서 뛰어난 성능을 입증해 왔다. 특히 이미지 처리에 특화된 CNN 모델은 이 분야의 발전을 주도하였다[4].

문서 분류 분야에서의 초기 접근법은 텍스트의 비율, 컬럼 구조, 폰트 크기 통계 등 수동으로 설계된 특징에 의존하여 특정 양식의 문서를 분류하는데 한정되었다[5]. 이러한 방식은 새로운 종류의 문서에 적용하기 위해 특징을 다시 설계하고 조정해야 하는 명백한 한계가 있었다.

문서의 구조적 유사성을 분류하기 위해 수동 특징 추출 과정 없이 원본 이미지 픽셀로부터 직접 계층적 특징을 학습하는 CNN 모델이 사용되었다. 이 접근법은 문서의 종류를 결정하는 것이 문자 단위의 세부 정보가 아닌 전체적인 레이아웃이라는

점에 착안하여, 입력 이미지를 저해상도로 다운샘플링하여 계산 효율성을 높이면서도 문서의 구조적 특징은 보존하는 방식을 사용했다[6].

최근에는 이미지넷(ImageNet)과 같은 대규모 이미지 데이터셋으로 사전 학습된 VGG(Visual Geometry Group), ResNet(Residual Network) 등의 모델을 미세 조정하는 전이학습(Transfer learning)이 표준 방법론으로 자리 잡았다. 전이학습은 적은 데이터로도 높은 성능을 달성할 수 있어 다양한 분야에 적용되고 있다. 국내에서는 주로 선박, 철도 부품, 의료 영상, 건축폐기물 등 실물 객체 분류에 이 방법론이 활발히 적용되고 있으나, 문서 페이지의 전체적인 시각적 레이아웃을 기반으로 한 분류 연구는 여전히 제한적이다[7]-[10]. 특히 품질증빙서류와 같은 산업용 비정형 문서를 대상으로 한 시각 기반 페이지 분류 접근은 국내외적으로도 활발하지 않은 실정이다.

2.2 문서 구조 분석 및 이해

스캔된 문서 이미지는 컴퓨터에게 단순한 픽셀 집합일 뿐 텍스트, 표, 그림 등 의미론적 구조를 직접적으로 인지할 수 없다. 이러한 한계로 인해 단순히 페이지 단위로 순차적으로 정보를 추출하면, 다만 문서에서 글 순서가 뒤섞이거나 표와 본문 텍스트가 구분 없이 추출되는 문제가 발생한다. 따라서 문서에서 정보를 추출하기 전에 문단, 표, 제목 등 각 구성 요소의 위치와 유형을 식별하는 문서 구조 분석 기술이 필수적으로 요구된다[11].

초기에는 주로 규칙 기반 방법론이 사용되었으나, 정해진 규칙을 벗어나는 다양한 양식에 대응하기 어렵다는 명백한 한계가 있었다[12]. 이후 SVM(Support Vector Machine)과 같은 전통적인 기계 학습 방법이 도입되었지만, 딥러닝 기술의 발전은 이 분야의 패러다임을 완전히 바꾸어 놓았다.

CNN은 이미지 전체를 단일 클래스로 분류하는 데 적합하지만, 한 페이지 내의 여러 객체(표, 문단 등)를 동시에 탐지하는 데는 한계가 있다. 이러한 문제를 해결하기 위해 객체 탐지 모델이 개발되었다. 이 모델들은 정확도를 중시하는 2단계 탐지 모델(Two-stage detectors)과 속도를 중시하는 1단계 탐지 모델(One-stage detectors)로 나뉜다. 2단계 모델인

R-CNN(Region-based CNN), Faster R-CNN은 후보 영역 제안과 분류를 순차적으로 수행하며, 1단계 모델인 YOLO(You Only Look Once) 와 SSD(Single Shot MultiBox Detector)는 이 과정을 하나의 네트워크에서 동시에 처리하여 실시간 응용에 강점이 있다 [13]-[16]. 최근에는 1단계 모델의 정확도 또한 크게 향상되어 널리 활용된다.

객체 탐지 기술은 문서 레이아웃 분석에서 효과적으로 활용되어 왔다. 대표적으로 Faster R-CNN을 이용해 복잡한 표 영역을 먼저 검출하고 내부 구조를 분석하거나, YOLOv5 모델을 계층적으로 적용하여 넓은 영역에서 특정 영역으로 점차 범위를 좁혀가는 연구가 있었다[17]-[19]. 이러한 기술의 성능을 극대화하기 위해, 대규모 문서 데이터셋으로 사전 학습(Pre-training)한 후 목표 도메인의 데이터셋으로 파인튜닝(Fine-tuning) 하는 접근법이 주로 사용된다.

최근에는 다단계 학습(Multi-stage training) 전략이 성능 향상을 위한 효과적인 방법으로 주목받고 있다. 펍레이넷(PubLayNet) 데이터셋을 제안한 연구에서는 일반 이미지 데이터셋으로 바로 특정 도메인 문서를 학습하는 것보다, 대규모 문서 데이터셋으로 중간 단계 사전학습을 수행할 때 성능이 유의미하게 향상됨을 보여주었다[20]. 이는 일반 이미지만 학습한 모델이 겪는 도메인 불일치 문제를 중간 학습 단계가 완화함을 시사한다.

닥레이넷(DocLayNet) 연구에서는 펍레이넷이 학술 논문에 국한되어 레이아웃 다양성이 부족한 한계를 지적하였다. 닥레이넷은 재무 보고서, 기술 매뉴얼, 특허 등 약 12,000페이지와 80,000개 이상의 객체 레이블을 포함하며, 문단, 표, 제목, 그림 등 주요 구성 요소로 주석 처리되어 있다[21]. 학습, 검증, 테스트용으로 분할되어 있어 객체 탐지 및 문서 구조 분석 모델 학습과 평가에 활용 가능하며, 공개 라이선스로 제공된다. 이러한 구성 덕분에 모델은 처음 보는 이질적 문서에서도 높은 일반화 성능을 보일 수 있으며, 중간 학습 단계에서 활용하면 특정 도메인에 특화되기 전에 문서의 보편적 구조를 효과적으로 학습할 수 있다.

이와 같이 사전 학습과 다단계 파인튜닝을 결합한 접근법은 문서 레이아웃 분석에서 높은 잠재력을 보였다. 그러나 기존 연구들은 대부분 학술 논문

이나 일반 비즈니스 문서와 같이 비교적 정형화된 데이터를 대상으로, 표 검출이나 문단 분류 등 단일 과업에 집중하는 경향이 있었다.

반면, 원자력 발전소의 품질증빙서류와 같은 비표준화된 산업 문서는 스캔 품질의 불균일성, 다양한 서식, 복잡한 내포 테이블 구조 등 특수한 문제를 포함한다. 이러한 문서에 최신 객체 탐지 기반 접근법을 성공적으로 적용하고 최적화하는 실증적 연구는 아직 충분하지 않다.

III. 제안하는 방법

본 연구에서는 선행 연구와 도메인 특수성을 고려하여, 언급된 산업적 난제들을 해결하기 위해 설계된 통합적 페이지 분류 및 관심 영역 검출 파이프라인을 제안한다. 이 파이프라인은 수천 페이지에 달하는 품질증빙서류에서 화학조성 테이블 이미지를 효과적으로 자동 추출하는 2단계 파이프라인으로 구성되어 있다. 또한, 계산 비용이 높은 객체 탐지 모델을 모든 페이지에 적용하는 비효율을 피하고자, 분석 대상을 점진적으로 좁혀나가는 계층적 접근법을 사용한다.

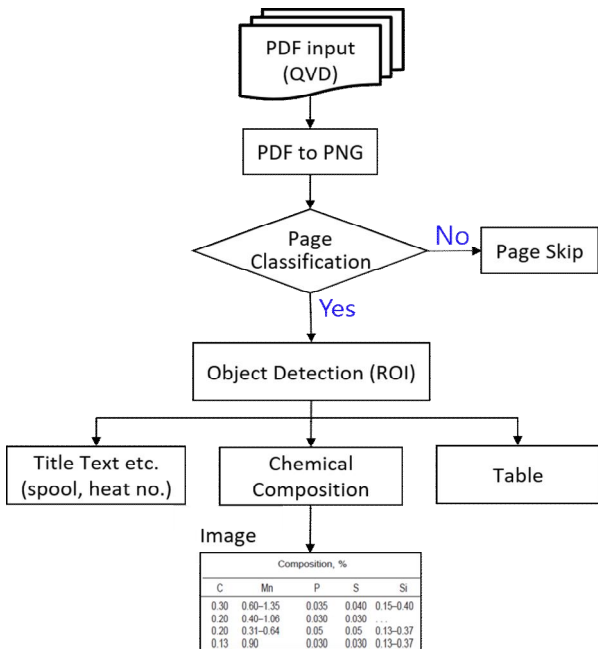


그림 1. 화학조성 테이블 추출 프로세스
Fig. 1. Process of chemical composition table extraction

제안하는 방법의 전체적인 처리 과정은 그림 1과 같다. 먼저, 1단계 페이지 분류 모델이 입력된 전체 문서 세트에서 화학조성 관련 페이지만을 신속하게 선별한다. 이후, 2단계 관심 영역 검출 모델은 선별된 페이지에만 적용되어, 화학조성 정보가 포함된 테이블 영역의 위치를 경계 상자로 정밀하게 특정한다. 최종적으로 검출된 영역은 이미지 파일로 저장되며, 이는 수작업 데이터 입력 과정을 대체하거나 추후 문서의 디지털 변환 작업의 입력 자료로 활용될 수 있다.

3.1 1단계: 관심 페이지 분류

품질증빙서류의 약 75-80%는 분석과 무관한 페이지이므로, 본 단계에서는 VGG16BN과 ResNet50과 같이 검증된 이미지 분류 모델을 활용하여 관심 페이지만 신속하게 필터링한다.

이미지넷으로 사전 학습된 두 모델은 이미지의 질감, 모서리, 형태와 같은 근본적인 시각적 특징을 추출하는 능력이 뛰어나며, 이러한 범용 특징 추출 능력은 일반 사물 이미지뿐만 아니라 문서 내의 표나 텍스트 블록 같은 구조를 인식하는 데에도 효과적으로 전이될 수 있다[22][23].

전이학습 형태로 기존 모델의 마지막 완전 연결 계층(Fully-connected layer) 분류기를 제거하고 관심/비관심 × 가로/세로를 출력하는 4 클래스 분류기로 교체하였다. 페이지 방향(가로/세로)을 함께 분류하는 것은 후속 객체 탐지 모델의 입력을 정렬하기 위함이다. 이후, 문서 이미지의 고유한 특징에 모델 전체가 미세하게 적응하도록, 모든 계층을 0.0001의 낮은 학습률(Learning rate)로 파인튜닝을 하였다.

모델 학습 및 검증을 위해 학습 10,504장, 검증 933장, 시험 4,008장으로 구성된 대규모 데이터셋을 구축했으며, 전체 데이터셋의 클래스별 분포는 그림 2와 같다.

3.2 2단계: 화학조성 테이블 영역 검출

1단계에서 선별된 페이지 내에서 화학조성 테이블이 존재하는 영역을 추출하는 단계로서 속도와

정확도 측면에서 뛰어난 성능을 보이는 YOLOv8 모델을 활용하였다. 모델이 탐지해야 할 관심 영역은 정보의 활용 목적에 따라 다음과 같이 3개의 클래스(Class)로 정의한다.

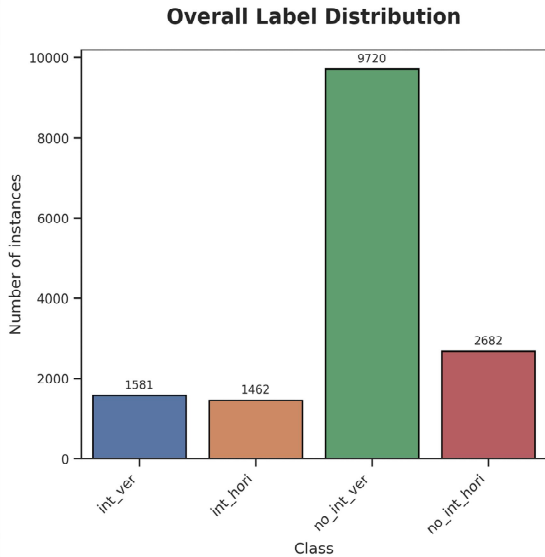


그림 2. 페이지 분류 데이터셋의 클래스별 분포
Fig. 2. Class distribution of the page classification dataset

첫째, Title 영역은 문서의 종류를 식별할 수 있는 제목 부분으로서 향후 문서 양식 세분화의 기반 정보로 활용된다. 둘째, Chemical 영역은 탄소(C), 크롬(Cr), 몰리브덴(Mo) 등 화학 원소와 함량 값이 집중적으로 제시되는 핵심 테이블 부분이다. 마지막으로, Table 영역은 Chemical 영역을 포함하면서 Heat No. 등 부가적인 정보를 포함하는 더 넓은 범위의 테이블로, Chemical 영역의 오탐지를 보완하고 향후 배관 컴포넌트와 화학조성을 연결하는 데 활용된다.

모델 학습에서는 품질증빙서류 문서의 고유한 특성을 고려하여 다단계 파인튜닝 전략을 적용하였다. 사전 학습된 YOLOv8 모델은 일반 객체를 담은 저해상도 COCO(Common Objects in Context) 데이터셋으로 학습되어, 글자와 테이블 위주의 고해상도 문서 이미지와는 도메인 및 해상도 불일치 문제가 발생한다.

이러한 문제를 해결하고 모델의 일반화 성능을 높이기 위해, 그림 3과 같이 2단계 파인튜닝 전략을 설계하였다. 첫 번째 단계에서는 문서 레이아웃 분석에 특화된 공개 데이터셋인 닥레이넷

으로 중간 학습을 수행하며, 이 과정은 일반 객체만 인식하던 모델이 최종 목표인 품질증빙서류를 학습하기 전에 문서의 보편적인 구조와 레이아웃 특성을 이해하도록 돕는 중요한 징검다리 역할을 한다. 이를 통해 도메인 불일치를 완화하고, 해상도 차이에 따른 성능 저하를 줄이며, 이후 단계에서의 학습 효율을 높일 수 있다.

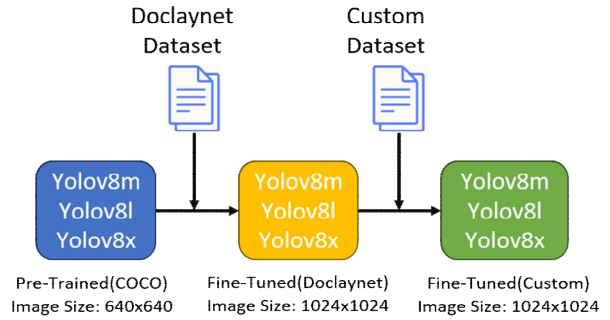


그림 3. 제안하는 2단계 파인튜닝 전략 흐름도
Fig. 3. Flowchart of the proposed 2-stage fine-tuning strategy

두 번째 단계에서는 원자력발전소 품질증빙서류를 직접 라벨링하여 구축한 맞춤형(Custom) 데이터셋으로 최종 파인튜닝을 진행한다. 맞춤형 데이터셋은 학습 1,080장, 검증 271장, 시험 230장으로 구성되었으며, 각 클래스별 데이터 분포는 그림 4와 같다. 이러한 2단계 파인튜닝을 통해, 모델은 먼저 일반적인 문서 구조를 학습하고 이어서 맞춤형 데이터셋의 고유한 특성과 문서 내 다양한 배치 패턴까지 반영하여 최종 탐지 성능을 향상시킬 수 있으며, 실제 문서 환경에서의 일반화 가능성도 높일 수 있다.

IV. 실험 및 결과

4.1 1단계: 페이지 분류 성능 평가

3.1절에서 제안한 페이지 분류 모델의 성능을 평가하기 위해, 사전 학습된 VGG16BN과 ResNet50 두 가지 모델을 대상으로 성능 비교 실험을 수행하였다. 모델의 분류 성능은 정확도(Accuracy)를 핵심 지표로 사용하였으며, 각 클래스별 성능을 세밀하게 분석하기 위해 정밀도(Precision)와 재현율(Recall)을 추가로 측정하였다.

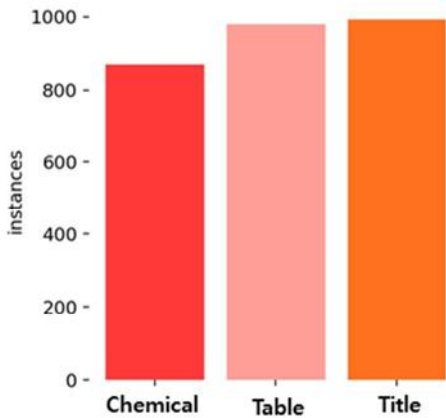


그림 4. 객체 탐지 데이터셋의 클래스별 분포

Fig. 4. Class distribution of the custom dataset for object detection

실험 결과, VGG16BN 99.4%, ResNet50 99.3%로 두 모델 모두 99% 이상의 매우 높은 분류 정확도를 보였다. 표 1은 각 모델이 시험 데이터셋의 4개 클래스에 대해 예측한 정밀도와 재현율을 보여준다. 두 모델 모두 모든 클래스에서 96% 이상의 높은 정밀도와 재현율을 기록하여 특정 클래스에 치우치지 않는 안정적인 분류 성능을 보였음을 확인하였다.

표 1. 페이지 분류 모델의 클래스별 성능 평가 결과
Table 1. Performance evaluation results of the page classification models for each class

Model	Class	Precision	Recall
VGG16BN	int_ver	0.960	0.988
	int_hori	0.994	0.979
	no_int_ver	0.998	0.995
	no_int_hori	0.995	0.999
ResNet50	int_ver	0.969	0.984
	int_hori	0.990	0.984
	no_int_ver	0.997	0.996
	no_int_hori	0.996	0.997

이는 분류하고자 하는 관심 페이지가 고유의 양식 등 다른 문서와 구별되는 명확한 시각적 특징을 가지고 있어, CNN 기반 모델이 이를 효과적으로 학습했기 때문으로 분석된다.

4.2 2단계: 관심 영역 검출 성능 평가

1단계에서 선별된 페이지를 대상으로 3.2절에서 설계한 2단계 파인튜닝 전략을 적용하여 관심 영역

검출 실험을 진행하였다. 실험에는 YOLOv8m, YOLOv8l, YOLOv8x 모델을 사용하였다. 모델의 성능을 정량적으로 평가하기 위해 정밀도, 재현율, 그리고 이 둘을 종합적으로 고려하는 mAP(mean Average Precision)를 평가지표로 사용하였다.

성능 평가 결과, 표 2와 같이 시험한 세 모델 모두 mAP50-95 기준 0.85 이상의 우수한 성능을 보였다. 이러한 높은 성능은 COCO 데이터셋과의 도메인 및 해상도 불일치 문제를 닥레이넷을 통한 중간 학습으로 효과적으로 완화했기 때문으로 판단된다.

표 2. 관심 영역 검출 모델 성능 평가 결과

Table 2. Performance evaluation of the object detection models

Model	Precision (Box)	Recall (Box)	mAP 50	mAP 50-95
YOLOv8m	0.987	0.995	0.990	0.870
YOLOv8l	0.983	0.993	0.990	0.857
YOLOv8x	0.990	0.997	0.990	0.853

실제로 동일한 YOLOv8x 모델에서 닥레이넷 중간 학습 단계를 생략하고 맞춤형 데이터셋으로만 직접 학습시켰을 경우, mAP50-95 성능은 0.792로 2단계 파인튜닝을 적용했을 때(0.853)보다 낮게 나타났다. 맞춤형 데이터셋의 규모가 제한적임을 고려할 때, 닥레이넷을 활용한 중간 학습은 모델이 특정 데이터에 과적합되는 것을 방지하고 문서의 일반적인 구조적 특징을 학습하는 데 기여한 것으로 판단된다. 시험 데이터셋에서의 성능 향상폭은 상대적으로 제한적이지만, 이러한 중간 학습 전략은 실제 환경에서 접할 수 있는 다양한 문서 양식에 대한 모델의 일반화 성능 향상에도 긍정적인 영향을 미칠 것으로 기대된다. 실제 시험 데이터셋에 대한 추론 결과, 그림 5와 같이 3가지 관심 영역이 모두 정확하게 추출되었으며, 보안상 민감한 내용에 대해서는 블러링(Blurring) 처리하여 도시하였다.

V. 결론 및 향후 과제

본 연구에서는 원자력 발전 분야의 방대한 품질 증빙서류에서 화학조성 정보를 자동으로 추출하기 위한 딥러닝 기반 2단계 자동화 파이프라인을 설계하고 그 실효성을 검증하였다.

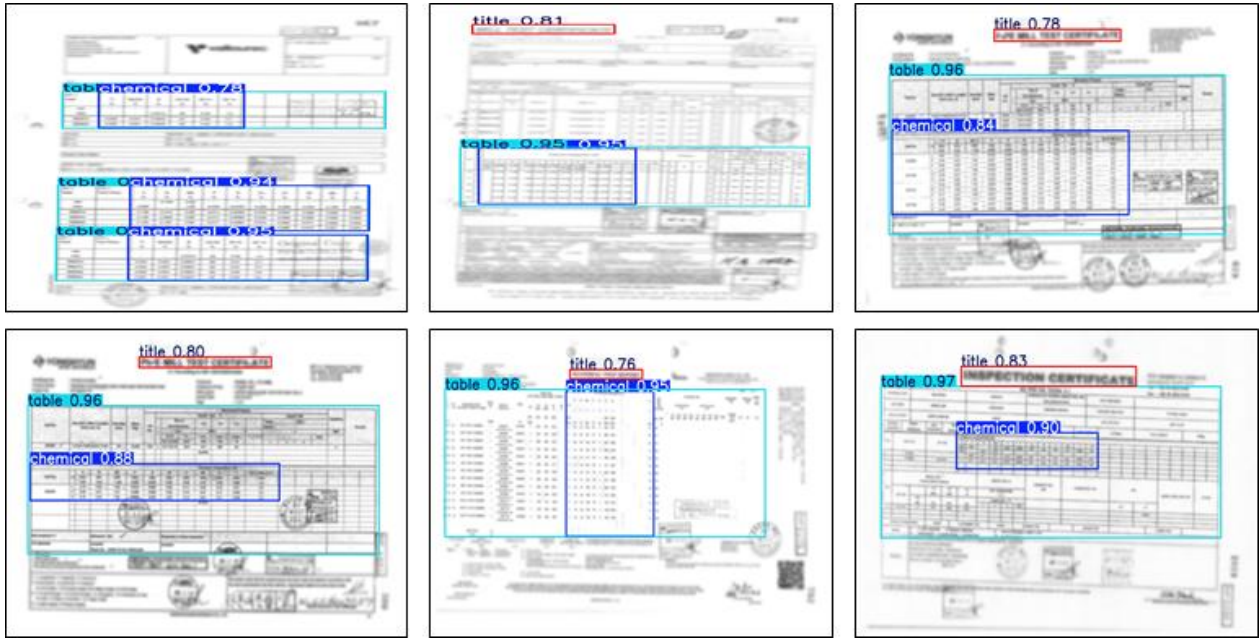


그림 5. 객체 탐지 모델의 시험 데이터셋 추론 결과

Fig. 5. Inference results of the object detection model on the test dataset

페이지 분류와 관심 영역 검출의 계층적 구조를 통해 정보 검색 효율성과 정확성을 높였으며, 실험 결과 주요 모델이 분석 대상 페이지와 관심 영역을 높은 정확도로 선별함을 확인하였다.

본 연구의 핵심 기여는 수천 페이지에 달하는 비정형 문서에서 소수의 핵심 정보 영역을 정확히 추출할 수 있는 자동화 파이프라인을 구축한 점이다. 이를 통해 기존 수작업 중심의 정보 검색을 대체하고, 배관감육 평가에 필요한 데이터를 신속하고 정확하게 제공할 수 있는 방법론을 제시하였다. 또한, 제안한 파이프라인을 통해 핵심 정보 영역을 자동으로 추출함으로써, 이후 데이터베이스화 및 분석 과정의 효율성을 향상시킬 수 있는 가능성을 확인하였다. 나아가, 데이터베이스 구축을 위한 전처리 자동화의 첫 단계를 실현하였으며, 향후 문자인식 기술과 연계하면 완전한 데이터베이스 자동화를 구현할 수 있는 핵심 기반 기술로 활용될 수 있다는 점에서 중요한 의미가 있다.

향후 연구에서는 본 연구에서 제안한 자동 검출 파이프라인을 전처리 모듈로 활용하여, 추출된 테이블 이미지로부터 문자 정보를 인식하고 정형화된 데이터베이스로 변환하는 후속 연구를 수행하고자 한다. 이를 통해 정보 검색부터 데이터베이스화까지 이어지는 전 과정을 자동화하는 통합 솔루션을 개

발하여, 원자력 분야 문서의 디지털 전환과 데이터 기반의 안전 관리 체계 구축에 기여할 것이다.

References

- [1] S. G. Lee and Y. S. Jung, "Quality Assurance System for Nuclear Power Plant Construction", *Journal of the Korean Society of Civil Engineers*, Vol. 43, No. 4, pp. 6-11, Apr. 1995.
- [2] H. Yun, S. J. Moon, and Y. J. Oh, "Development of Wall-Thinning Evaluation Procedure for Nuclear Power Plant Piping—Part 1: Quantification of Thickness Measurement Deviation", *Nuclear Engineering and Technology*, Vol. 48, No. 4, pp. 820-830, Mar. 2016. <https://doi.org/10.1016/j.net.2016.01.020>.
- [3] D. Baviskar, S. Ahirrao, V. Potdar, and K. Kotecha, "Efficient Automated Processing of the Unstructured Documents Using Artificial Intelligence: A Systematic Literature Review and Future Directions", *IEEE Access*, Vol. 9, pp. 72894-72936, Apr. 2021. <https://doi.org/10.1109/ACCESS.2021.3072900>.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Communications of the ACM*,

- Vol. 60, No. 6, pp. 84-90, Jun. 2017. <https://doi.org/10.1145/3065386>.
- [5] C. Shin and D. Doermann, "Document Image Retrieval Based on Layout Structural Similarity", Proc. of the International Conference on Image Processing, Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, Vol. 2, pp. 606-612, Jan. 2006.
- [6] L. Kang, J. Kumar, P. Ye, Y. Li, and D. Doermann, "Convolutional Neural Networks for Document Image Classification", Proc. of the 22nd International Conference on Pattern Recognition, Stockholm, Sweden, pp. 3168-3172, Aug. 2014. <https://doi.org/10.1109/ICPR.2014.546>.
- [7] W. Y. Baek and S. G. Kang, "Ship Classification Method using Two-Stage CNN Model", Journal of the Korean Institute of Information Technology, Vol. 21, No. 8, pp. 203-210, Aug. 2023. <https://doi.org/10.14801/jkiit.2023.21.8.203>.
- [8] H. J. Lee, S. M. Lee, D. Y. Kwak, Y. J. Suh, and D. J. Kim, "Development and Compression of ResNet-based Classification Model for Key Components of Railway Vehicles", Proc. of the Korean Society of Information Technology Conference, Jeju, Korea, pp. 145, May 2024.
- [9] M. J. Cho, J. M. Choi, and Y. H. Kim, "Diabetic Retinopathy Prediction Model through Classification of Retinal Images", Journal of the Korean Institute of Information Technology, Vol. 22, No. 11, pp. 21-28, Nov. 2024. <https://doi.org/10.14801/jkiit.2024.22.11.21>.
- [10] J. K. Sung, M. C. Yang, K. G. Moon, and Y. G. Kim, "A Comparison of Image Classification System for Building Waste Data based on Deep Learning", The Journal of The Institute of Internet, Broadcasting and Communication, Vol. 23, No. 3, pp. 199-206, Jun. 2023. <https://doi.org/10.7236/JIIBC.2023.23.3.199>.
- [11] N. Subramani, A. Matton, M. Greaves, and A. Lam, "A Survey of Deep Learning Approaches for OCR and Document Understanding", arXiv preprint arXiv:2011.13534, Feb. 2021. <https://doi.org/10.48550/arXiv.2011.13534>.
- [12] J. L. Fisher, S. C. Hinds, and D. P. D'Amato, "A Rule-Based System for Document Image Segmentation", Proc. of the 10th International Conference on Pattern Recognition, Atlantic City, NJ, USA, pp. 567-572, Jun. 1990. <https://doi.org/10.1109/ICPR.1990.118166>.
- [13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, pp. 580-587, Jun. 2014. <https://doi.org/10.1109/CVPR.2014.81>.
- [14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6, pp. 1137-1149, Jun. 2017. <https://doi.org/10.1109/TPAMI.2016.2577031>.
- [15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 779-788, Jun. 2016. <https://doi.org/10.1109/CVPR.2016.91>.
- [16] W. Liu, et al., "SSD: Single Shot MultiBox Detector", Proc. of the European Conference on Computer Vision, Amsterdam, Netherlands, pp. 21-37, Oct. 2016. https://doi.org/10.1007/978-3-319-46448-0_2.
- [17] D. S. Lee and S. K. Kwon, "Methods of Classification and Character Recognition for Table Items through Deep Learning", Journal of Korea Multimedia Society, Vol. 24, No. 5, pp. 651-658, May 2021. <https://doi.org/10.9717/kmms.2020.24.5.651>.
- [18] S. J. Kim, D. H. Yu, S. Y. Hwang, and M. H. Kim, "Development of Intelligent OCR Technology to Utilize Document Image Data", Proc. of the KICS Conference, Pyeongchang, Korea, pp. 212-215, Jun. 2022.
- [19] S. Y. Kim, J. W. Park, S. M. Kim, Y. Na, and Y. J. Jang, "Multi-book Label Detection Model

using Object Detection and OCR", Journal of the Korean Institute of Information Technology, Vol. 21, No. 2, pp. 1-8, Feb. 2023. <https://doi.org/10.14801/jkiit.2023.21.2.1>.

- [20] X. Zhong, J. Tang, and A. J. Yepes, "PubLayNet: largest dataset ever for document layout analysis", Proc. of the 2019 International Conference on Document Analysis and Recognition, Sydney, NSW, Australia, pp. 1015-1022, Sep. 2019. <https://doi.org/10.1109/ICDAR.2019.00166>.
- [21] B. Pfizmann, C. Auer, M. Dolfi, A. S. Nassar, and P. Staar, "DocLayNet: A Large Human-Annotated Dataset for Document-Layout Analysis", Proc. of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, pp. 3845-3853, Aug. 2022. <https://doi.org/10.1145/3534678.3539043>.
- [22] Y. Yang, et al., "A comparative analysis of eleven neural networks architectures for small datasets of lung images of COVID-19 patients toward improved clinical decisions", Computers in Biology and Medicine, Vol. 139, pp. 104887, Sep. 2021. <https://doi.org/10.1016/j.combiomed.2021.104887>.
- [23] J. Plested and T. Gedeon, "Deep transfer learning for image classification: a survey", arXiv preprint arXiv:2205.09904, May. 2022. <https://doi.org/10.48550/arXiv.2205.09904>.

저자소개

이 상 훈 (Sang-Hoon Lee)



2008년 2월 : 인하대학교
기계공학부(공학사)
2020년 8월 : 경북대학교
산업대학원
기계공학전공(공학석사)
2023년 3월 ~ 현재 :
국립금오공과대학교

디지털융합공학과 박사과정
2012년 11월 ~ 현재 : 한국전력기술(주) 디지털솔루션연구소
관심분야 : Computer Vision, Deep Learning

문 성 빈 (Seong-Bin Mun)



2014년 2월 : 부산대학교
기계공학부(공학사)
2024년 2월 : 국립금오공과대학교
디지털융합공학과(공학석사)
2013년 7월 ~ 현재 :
한국전력기술(주)
디지털솔루션연구소

관심분야 : Computer Vision, Deep Learning

오 영 진 (Young-Jin Oh)



1998년 2월 : 서울대학교
원자핵공학과(공학사)
2000년 2월 : 서울대학교
원자핵공학과(공학석사)
2006년 8월 : 서울대학교
원자핵공학과(공학박사)
2006년 12월 ~ 현재 :

한국전력기술(주) 디지털솔루션연구소 연구원
관심분야 : 설비열화 예측, 데이터기반 예측진단

박 상 원 (Sang-Won Park)



2020년 3월 ~ 현재:
국립금오공과대학교
컴퓨터공학과 학사과정
관심분야 : Computer Vision,
LLMs, Data Mining

이 해 연 (Hae-Yeoun Lee)



1997년 2월 : 성균관대학교
정보공학과(학사)
1999년 2월 : KAIST 전산학과
(공학석사)
2006년 2월 : KAIST
전자전산학과(공학박사)
2008년 3월 ~ 현재 :

국립금오공과대학교 컴퓨터소프트웨어공학과 교수
관심분야 : Digital Forensics, Image Processing, IoT