

# LH 하자보수 통합관리 솔루션에 적용된 AI 성능평가 방안

The Method of AI performance evaluation applied to  
LH defect repair integrated management solution

양동석

이영미



연구관리 2024-107호

## LH 하자보수 통합관리솔루션에 적용된 AI 성능평가 방안

지 은 이 양동석, 이영미  
발 행 인 정창무  
발 행 처 한국토지주택공사 토지주택연구원  
주 소 (34047) 대전 유성구 엑스포로 539번길 99  
홈페이지 <http://lhri.lh.or.kr>

전화번호 042-866-8575  
이 메 일 [blue@lh.or.kr](mailto:blue@lh.or.kr)

이 출판물은 우리 공사의 업무상 필요에 의하여 연구 검토한 기초자료로서 공사나 정부의 공식적인 견해와  
관계가 없습니다.

우리 공사의 승인 없이 연구내용의 일부 또는 전부를 다른 목적으로 이용할 수 없습니다.

## LH 하자보수 통합관리솔루션에 적용된 AI 성능평가 방안

(The Method of AI performance evaluation applied to LH defect repair integrated management solution)

양동석·이영미

## 참여연구진

### 연구책임

**양동석** LH 토지주택연구원 연구위원

### 연구진

**이영미** LH 토지주택연구원 연구원

### 연구심의위원

**이영은** LH 토지주택연구원 실장(심의위원장)

**이수옥** 국토연구원 선임연구위원

**홍성기** 한국갤럽 상무이사

**이형철** LH 건설관리본부 준공품질관리팀 팀장

**이상훈** LH 토지주택연구원 연구위원

**송상훈** LH 토지주택연구원 연구위원

### 자문위원 (가나다순)

**송성현** AIFactory 대표이사

**김 윤** 지오비전 대표이사

**양동국** 디케이앤티 이사

**최하나** AltheNutrigene 팀장

**윤기병** (前) 원광대학교 교수



## ■ 연구의 배경 및 목적

- 인공지능 기술은 하드웨어의 발전과 학습 알고리즘 등 소프트웨어의 고도화를 배경으로 현재 여러 분야에서 활용되어지고 있음
- 공사에서는 공공주택 하자보수와 관련하여 “모바일 하자보수 통합관리솔루션 개발 용역”에 부분적으로 인공지능을 적용하여 진행중임(‘23.12~’24.12)
- 공사의 경우 해당 업무와 관련하여 인공지능 적용사례가 없을 뿐만 아니라 학습이나 평가 등에 따른 누적 데이터나 성능 기준 등이 제시되지 않아 실효성 여부 판단에 어려움이 있음
- 특히, 인공지능의 특성상 최적의 모델을 생성하고 많은 수의 고품질 데이터를 통한 학습을 함으로써 실효성 있는 성능 확보가 가능한 특징이 있음
  - 적합한 성능지표를 수립함으로써 신뢰성 및 실효성을 제시할 수 있음
  - 개발 후에도 최신의 데이터를 수집하고 학습함으로써 성능을 향상시킬 수 있어야 함
  - 학습에 필요한 데이터 확보뿐만 아니라 성능테스트에 요구되는 데이터 구축도 필요
- 이에 따라, “하자보수 통합관리 솔루션”에 적용된 인공지능 기술의 성능을 확인할 수 있는 객관적인 성능지표 개발이 요구됨
  - 인공지능이 적용된 하자보수 업무의 경우 어떤 측면에서 실효성을 확보할 수 있는지 객관적인 판단 기준 제시 필요
  - 시스템 운영 및 업그레이드 시 목표 기준으로 활용
- 본 연구는 “LH 모바일 하자보수 통합관리 솔루션”에 적용된 인공지능 기술에 대한
  - ① 인공지능 성능을 평가할 수 있는 지표 개발
  - ② 인공지능 성능평가 운영 및 발전 방안 제시 등을 목적으로 함

## ■ 연구의 내용 및 결과

### ○ 인공지능의 정의 및 성능지표

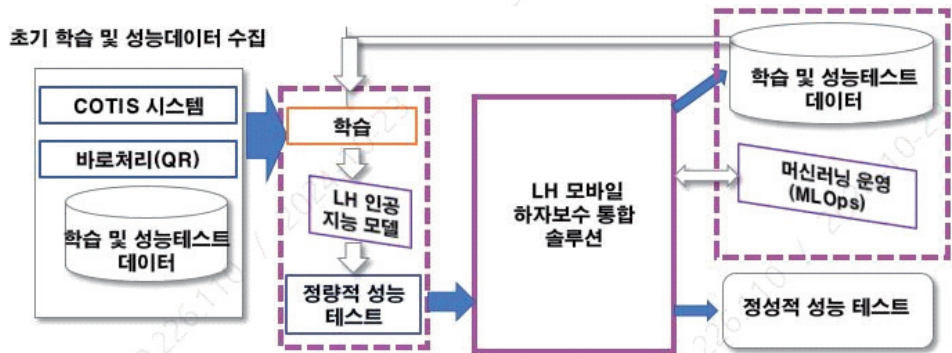
- 인공지능은 인간처럼 학습하고 문제를 해결하려는 방식인데 최근에 적용 중인 인공지능 90% 이상이 딥러닝 기술이며 본 연구에서의 인공지능은 딥러닝 기술을 의미함
- 인공지능의 주요 성능평가 지표들 중 본 연구에서 다루고자 하는 분류모델의 주요 지표로는 정확도, 정밀도, 재현율, F1-Score등이 있음
- 인공지능 성능지표 관련 주요 연구문헌들을 살펴보고 본 연구의 목표인 “LH 하자보수 통합 솔루션”에 적용된 인공지능 기술의 관리 및 평가에 적합한 성능지표 수립 방향 및 조건 등을 검토함
- 하자보수관련 인공지능 적용 사례 등을 조사하고 성능지표 수립 방향성을 검토함

### ○ LH의 하자보수 통합관리솔루션의 인공지능 적용 내용 및 주요 특성 검토

- LH의 유지보수 관련 프로그램은 입주품질관리시스템인 “바로처리(QR)”, 업무효율성을 높이고자 개발된 “바로픽스”, 고객편의성을 위한 “카카오톡 AS” 등 세 가지 프로그램으로 하자보수 업무를 수행하고 있으나 여러 문제점이 내재 중임
- LH 하자보수 관련 업무 처리시 개선되어야 할 점들로 주거생애 전체 서비스 수행, 편리한 하자유형 등록, 표준화된 하자유형 분류체계, 신속하고 효율적인 업무처리 등을 해결하기 위해 “LH 모바일 하자보수 통합관리솔루션(하자보수 통합관리플랫폼)” 개발 시행
- 본 연구에서는 해당 솔루션의 구성 및 인공지능 적용 내용을 조사하고 하자보수 관련 적용 모델 및 데이터 특성들을 검토함

### ○ LH의 하자보수 통합관리솔루션의 인공지능 성능지표 개발

- LH 하자보수 통합솔루션에 적용된 인공지능모델에 대한 성능지표를 수립하기 위해 성능지표 추진체계 및 LH 성능평가 체계를 제시함
- 성능평가는 개발된 인공지능 모델의 안전성과 신뢰성을 평가하여 실용성을 검토할 수 있음
- 특히, 성능평가 후 결과에 따라 인공지능 모델을 변경하거나 새로운 학습데이터를 추가하여 재학습 후 개선된 모델로 문제를 해결하는 단계가 추가됨
- LH형 인공지능 성능평가를 위한 정량적 지표(정밀도, 재현율, mAP, F1 score)와 정성적 지표(편의성, 신속성, 경제성, 안전성, 신뢰성)를 수립하고 기준을 제시



[그림 1] LH 하자보수 통합솔루션에 적용된 인공지능 성능평가 체계도

○ LH의 하자보수 통합관리솔루션의 발전방안 및 성능평가 운영안 제시

- LH의 신뢰성 높은 하자 이미지 데이터를 확보함으로써 그에 따른 다양한 데이터 기반 의사결정 지원 서비스에 대한 방안 제시함
- 사용자 인터페이스 고도화로 편리성 및 직관성을 높일 수 있는 방안 제시
- 사용자 입력 오류를 해결 할 수 있는 방안 제시
- 3D 설계 도면 사용 고도화 방안 및 설계 및 시공 단계에서 상호 보완 적용 방안 등을 제시함
- 평가 및 학습 운영체계의 자동화 방안 제시

## ■ 연구의 한계 및 제언

- 인공지능의 적용이 활발히 진행되고 있는 가운데 명확한 성능평가지표에 따른 해당 소프트웨어의 실효성 및 신뢰성 등의 판단에 대한 연구가 부족한 실정임
- 본 연구는 LH 하자보수 솔루션에 관련하여 인공지능 기술을 적용한 부분에 대한 성능평가지표를 개발하고 발전 방안을 제시하는 연구로써 향후 수립한 지표들에 대한 성능평가 등의 시스템 검증이 요구됨

## 주제어

하자보수, 인공지능, 성능평가, 솔루션, 개발, 방향, LH



---

# 차례 Contents

## 제1장 서론

- 1. 연구의 배경 및 목적 ————— 9
- 2. 연구의 범위 및 방법 ————— 11

## 제2장 인공지능 성능지표 및 활용사례 조사

- 1. 인공지능 정의 및 성능지표 ————— 13
- 2. 인공지능 성능지표 관련 문헌 연구 ————— 21
- 3. 하자보수 관련 인공지능 활용 사례 ————— 28

## 제3장 LH 하자보수 통합관리솔루션 인공지능 성능평가 방안

- 1. LH 하자보수 인공지능 기능 및 구성 ————— 41
- 2. LH 하자보수 인공지능 성능지표 수립 ————— 59

## 제4장 LH 하자보수 통합관리솔루션 인공지능 성능평가 운영 및 발전 방안

- 1. LH 하자보수 통합관리솔루션 인공지능 성능평가 운영 ————— 67
- 2. LH 하자보수 통합관리솔루션 인공지능 발전방안 ————— 65

## 제5장 결론

- 1. 본 연구의 성과 ————— 77
- 2. 연구의 한계 및 제언 ————— 79

---

## 표 차례 List of Tables

[표2-1] 인공지능 성능지표 관련 참고 학술 논문	21
[표2-2] 연구 흐름도 및 과제 관련도	27
[표2-3] 텍스트 분류 모델 구조	31
[표3-1] LH가 운영중인 하자보수 관련 프로그램	42

---

## 그림 차례 List of Figures

[그림2-1] 본 연구에서 의미하는 인공지능	15
[그림2-2] 예측 값과 실제 값을 비교하기 위한 오차행렬 구성	18
[그림2-3] POS-VISION	32
[그림2-4] 하수관로 인공지능 결함탐지 알고리즘 구조도	32
[그림2-5] 스캔을 통한 3D 영상의 예	33
[그림2-6] 효성중공업 ‘아모르’ 운영 프로세스	34
[그림2-7] 콜봇+챗봇 AI 엔진 서비스 프로세스 플로우	35
[그림2-8] 대우건설 ‘인공지능 계약문서 분석시스템 바로답 BaroDAP	37
[그림2-9] 삼성물산 ‘헤스티아 2.0’	38
[그림2-10] 로타벤치마크(LoTa-Bench)	39
[그림2-11] OPEN Ko-LLM 리더보드 시즌2	40
[그림3-1] 하자공종별 발생건수의 예	43
[그림3-2] 사용자 편의성을 위한 하자유형의 분류체계	44
[그림3-3] 하자유형 분류가 어려운 예	45
[그림3-4] LH 모바일 하자보수 통합관리 플랫폼 구성도	46
[그림3-5] LH 모바일 하자보수 통합관리 플랫폼의 시스템 구성도	46
[그림3-6] LH 모바일 하자보수 통합관리 플랫폼에서의 AI 적용	47
[그림3-7] 하자보수 유형별 발생건 수	48
[그림3-8] 하자보수 부위자재별 발생 건수	49
[그림3-9] 같은 유형을 다른 유형으로 분류된 하자 유형	49
[그림3-10] 하자 유형 임의 분류 및 육안 확인 어려운 이미지	50
[그림3-11] AI 모델 구축 프로세스	51
[그림3-12] COTIS 하자 이미지 자료의 하자유형 분류체계 개선 내용	52
[그림3-13] 바로처리(신축아파트)에서 생성된 하자유형 분류개선	52
[그림3-14] 하자유형에 따른 AI 모델 선정	53

[그림3-15] 하자유형 분류 보안을 위한 텍스트 분류 모델 구성	54
[그림3-16] 텍스트를 통한 하자유형 분류모델 선정	55
[그림3-17] 등록된 하자관련 텍스트 학습 데이터	55
[그림3-18] 텍스트 데이터의 하자 유형별 데이터 분포비율	55
[그림3-19] 하자보수 자동 스케줄링 활용데이터	56
[그림3-20] 하자보수 자동 스케줄링 생성 프로그램 프로세스	57
[그림3-21] 보수공사 세부내역서의 예	58
[그림3-22] 내역서 자동생성 프로세스 구성	58
[그림3-23] 인공지능 성능평가 체계도	59
[그림3-24] LH 하자보수 통합솔루션에 적용된 인공지능 성능평가 체계도	60
[그림3-25] LH 하자유형 분류에 적용된 인공지능 성능평가 예시	63
[그림3-26] LH 인공지능모델의 하자유형별 세부 정확도	64
[그림3-27] LH 하자유형 텍스트 분류 성능테스트 결과	65
[그림4-1] 고품질 학습, 평가, 테스트 데이터 확보 방안	68
[그림4-2] LH 하자보수 인공지능 학습 및 성능평가 체계	69
[그림4-3] 데이터 기반의 의사결정 지원	70
[그림4-4] 데이터 기반의 하자보수 관련 의사결정 지원	71
[그림4-5] LH 모바일 하자보수 솔루션 화면	72
[그림4-6] 이미지 촬영시 촬영 가이드 제시 및 재촬영 요구의 예	73
[그림4-7] LH 하자보수 모바일 통합솔루션에서 제공중인 3D 적용 기능	74



# 제1장 서론

## 1. 연구의 배경 및 목적

### 1.1 연구배경

- 인공지능(AI) 기술은 하드웨어의 발전과 학습 알고리즘 등 소프트웨어의 고도화를 배경으로 현재 여러 분야에서 적극적으로 활용되고 있음
- 세계적으로 인공지능을 적용한 다양한 솔루션들이 공공 업무에 활발하게 도입되고 있음
  - 미국과 중국을 중심으로 인공지능 경쟁이 치열하게 진행되고 있음
  - 인공지능 기반 기술 확보, 인력양성, 인프라 및 생태계 조성 등에 많은 자본 투입 중
  - 공공 영역에 인공지능을 적용할 경우 세계 GDP가 1.93% 높게 성장 가능<sup>1)</sup>
- 공사에서는 공공주택 하자보수와 관련하여 “모바일 하자보수 통합관리솔루션 개발 용역”에 부분적으로 인공지능을 적용하여 진행중임(‘23.12~’24.12)
  - 하자 이미지 데이터 레이블링 및 딥러닝을 통한 하자분류 AI 개발
  - 하자처리 스케줄 고도화를 위한 AI 개발
- 공사의 경우 해당 업무에 인공지능 적용사례가 없을 뿐만 아니라 학습이나 테스트 등과 관련된 누적 데이터나 성능 기준 등이 제시되지 않아 실효성 여부 판단이 어려움
- 특히, 인공지능의 특성상 최적의 모델을 생성하고 많은 수의 고품질 데이터를 통한 학습을 함으로써 실효성 있는 성능 확보가 가능함
  - 적합한 성능지표를 수립함으로써 신뢰성 및 실효성을 제시할 수 있음
  - 개발 후에도 최신의 데이터를 수집하고 학습함으로써 성능을 향상시킬 수 있어야 함
  - 인식 오류 등에 대한 전처리 작업 강화나 알고리즘 수정 등의 평가가 요구됨
  - 학습에 필요한 데이터 확보뿐만 아니라 성능테스트에 요구되는 데이터 구축도 필요

1) “국내외 AI 활용 현황과 공공적용”, IITP, 2017

- 이에따라, “하자보수 통합관리 솔루션”에 적용된 인공지능 기술의 성능을 확인할 수 있는 객관적인 성능지표 개발 및 효율적 운영이 요구됨
  - 인공지능이 적용된 하자보수 업무의 경우 어떤 측면에서 실효성을 확보할 수 있는지 객관적인 판단 기준 제시 필요
  - 시스템 운영 및 업그레이드 시 목표 기준으로 활용

## 1.2 연구목적

- 본 연구는 LH 하자보수 통합관리솔루션에 적용된 인공지능 기술에 대한
  - ① 인공지능 성능을 평가할 수 있는 지표 개발
  - ② 인공지능 성능평가 운영 및 발전 방안 제시 등을 목적으로 함

## 2. 연구의 범위 및 방법

### 2.1 연구범위

- 본 연구에서는 “LH 하자보수 통합솔루션”에 적용된 인공지능 기술을 대상으로 제한함
- 성능지표의 경우 기존에 제시된 여러 성능지표 중 LH 하자보수 관련 업무에 최적화된 지표를 찾고자 함
- 특히, 성능지표의 적용 및 기능 발전 방안을 중심으로 연구를 수행함

### 2.2 연구 방법

#### ■ 문헌 및 인터넷 조사를 통한 연구 진행

- 인공지능 정의 및 성능지표 검토
  - 인공지능 관련 연구 결과와 발간물 조사
- 하자보수관련 인공지능 활용 사례 조사
  - 문헌 및 인터넷을 통한 활용 사례 조사

#### ■ 인터뷰 및 방문조사를 통한 연구진행

- 인공지능 전문기업 방문 및 전문가 인터뷰 및 시사점 정리
  - 하자보수와 관련된 인공지능 활용 사례 구축, 성능지표, 테스트 및 운영 등의 전문가 인터뷰

#### ■ 전문가 그룹 및 관련기관 연계를 통한 연구진행

- 관련분야 전문가 및 전문기관과의 정보교류를 통한 연구완성도에 기여
  - 인공지능 관련 전문기관 자문
- 관련 실무부서와 유기적 협력체제 유지
  - 주택품질기준팀 및 하자보수 관련 부서와 담당자 인터뷰 및 워크숍을 통한 협력체제 강화
  - 연구 성과가 유효하게 적용될 수 있도록 수시로 연구결과 협의



## 제2장 인공지능 성능 지표 및 활용 사례 조사

### 1. 인공지능의 정의 및 성능지표

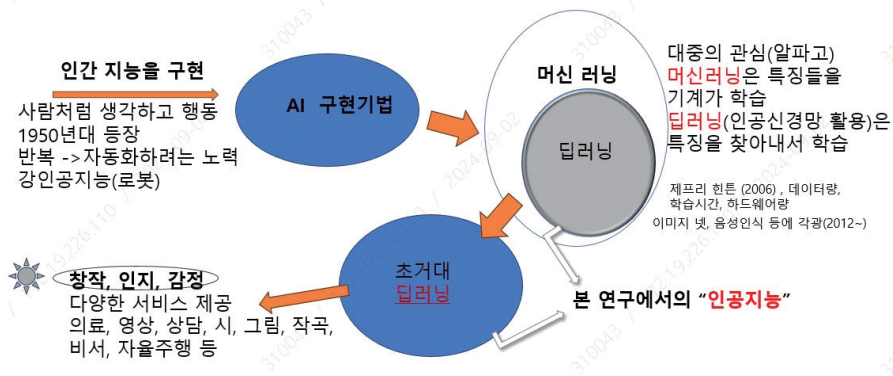
#### ■ 인공지능의 정의<sup>2)</sup>

- 인공지능(Artificial Intelligence)이란 인간이 어떤 객체를 대상으로 인지하고, 판단하고, 사고하는 방식을 컴퓨터로 구현하고자 하는 노력으로 탄생 된 개념
- 곧 인간처럼 학습하고 문제를 해결하려는 프로그램을 연구하면서 시작되었는데 이 중 한 분야인 머신러닝은 프로그램 없이 기계가 스스로 학습할 수 있도록 하는 기술을 의미함
- 머신러닝은 알고리즘을 설계하고 프로그래밍하기 어렵고 복잡하거나 불가능한 경우에 활용하는 기법임
  - 스팸 이메일 처리, 문자인식, 이미지 처리 등에 활용
  - 1980년대 이후로 전문가시스템, 통계적 방법, 의사결정트리 등 다양한 기술들을 활용하여 널리 사용됨
- 이러한 머신러닝 기법 중에서 인간의 뇌를 모델로 한 인공신경망과 많은 수의 데이터들을 가지고 학습하여 문제를 해결하는 기술을 딥러닝이라 할 수 있음
  - 딥러닝은 제프리 힌튼이 인공신경망(Artificial Neural Network)을 “딥러닝”이라는 명칭으로 바꾸고 이미지 인식 분야에 월등한 효과를 얻어내면서부터 여러 분야에 활용되기 시작함(2006년)
  - 인간의 신경세포에서 수집된 정보들이 뇌세포로 전달되고 뇌세포들간의 상호작용(학습)으로 그에 따른 정보 인지나 제어 값들을 얻어내는 것처럼 딥러닝 모델에서도 다양한 문제들에 대한 데이터들을 받아들이고 매개변수들 간의 연결강도 조정(학습하는 과정)을 통해, 입력된 문제에 대한 결과 값들을 도출해 낼 수 있음
  - 특히, 인간이 학습에서 어떠한 내용에 대해 반복 학습하듯 딥러닝 모델에서도 결과

2) 양동석 외(2024), “LH 업무 효율화를 위한 딥러닝 개발방향 연구”

값에 대한 올바른 정보를 피드백(Feedback)시켜 줌으로써 각 층의 노드들 간의 연결 강도를 더욱 세밀하게 조정함으로써 신뢰성 높은 결과를 얻을 수 있음

- 어떠한 문제를 딥러닝으로 표현하고 해결하기 위해서는 그 문제에 적합한 각 층들의 노드 수를 구성하고 계산할 수 있어야 하는데 그에 따른 엄청난 계산량이 요구되며, 또한 정확한 결과값을 도출하기 위해서는 수많은 학습데이터도 필요함
- 초거대 인공지능(초거대 딥러닝)은 하드웨어 발전을 기반으로 기존에 수행하지 못하던 계산 처리 능력을 확보함으로써 딥러닝 모델의 노드 수를 인간의 뇌세포 크기만큼 (860억 이상의 뇌세포수) 노드 수를 늘리고 각 노드의 연결을 통해 복잡한 문제를 해결할 수 있는 모델임
  - 초거대딥러닝은 언어생성모델(LLM: Large Language Model)이라고도 부르는데 이는 인간의 언어를 중심으로 학습을 시키고 그에 따른 결과 값도 언어로 생성해 제시할 수 있음을 의미(ChatGPT 3.5의 경우 인터넷상의 45 TB(Tera Byte) 데이터를 가지고 학습)
  - 초거대 딥러닝은 생성형 AI라고도 부르는데 기존 정보를 단순 보여주는 것이 아니라 자신이 최적의 결과를 만들어서 제시한다는 의미로 다양한 서비스들 제공
  - ChatGPT와 미드저니를 이용하여 하루만에 동화책을 완성하여 판매
  - ImageGPT는 4억장의 그림을 학습하여 텍스트 입력만으로 그림을 생성해 줌
  - 언어 번역, 질문 응답, 챗봇 등 다양한 자연어 처리 작업에 사용될 수 있음
  - 객체 감지, 이미지 분석 등에서 복잡한 패턴과 관계를 학습하여 정확성과 견고성을 향상시킬 수 있음
  - 구글의 Gemini Ultra의 경우 멀티모달 AI(multimodal AI)는 텍스트, 이미지, 영상, 음성 등 다양한 데이터 모달리티를 함께 고려하여 서로간의 관계성을 학습 및 표현할 수 있음
  - GPT-3를 통해 상담서비스 및 소설 창작이 가능함
  - LG AI 연구원에서는 엑사원을 활용하여 뉴욕 패션위크에서 적용한 옷을 디자인 함
- 현재 활발히 적용되고 있는 인공지능 기술은 딥러닝 또는 초거대딥러닝이 90%이상 차지하고 있는 실정으로 본 연구에서의 인공지능은 그러한 딥러닝 활용 기술로 정의하고자 함



[그림2-1] 본 연구에서 의미하는 인공지능

## ■ 주요 인공지능(딥러닝) 모델

- ANN(인공신경망 : Artificial Neural Network) : 인공신경망은 인공 뉴런들을 연결하여 연결강도(결합세기, 가중치, 매개변수 값)에 따라 문제를 해결하는 모델을 의미
  - 다수의 입력데이터를 받는 입력층(Input Layer), 은닉층(Hidden Layer), 출력층(Output Layer)로 구성됨
  - 출력층의 값으로 예측 결과를 얻어냄
  - 학습과정에서 파라미터의 최적값을 찾기 어려움
  - 은닉층이 많으면 학습 성능이 올라가지만 그만큼 연산량이 기하급수적으로 증가됨
- DNN(심층신경망 : Deep Neural Network)은 은닉층을 2개 이상 구성하여 학습하는 모델
  - 많은 데이터의 반복학습, 사전학습과 오류역전파 알고리즘을 통해 널리 사용
  - DNN을 응용한 모델로 CNN, RNN, LSTM, GRU 등이 있음
- CNN(합성곱신경망: Convolution Neural Network)은 데이터의 특징을 먼저 추출하고 특징들의 패턴을 파악하는 학습 모델임
  - Convolution Layer와 Pooling Layer를 복합적으로 구성함
  - Convolution은 데이터의 특징을 추출하는 과정으로 하나의 압축과정이며 파라미터의 개수를 효과적으로 감소시키는 역할을 함
  - Pooling 은 Convolution 과정을 거친 레이어 사이즈를 줄이고 노이즈를 제거하며, 일관적인 특징을 제공
  - 정보추출, 문장분류, 얼굴인식 등의 분야에서 활발하게 사용됨

- RNN(순환신경망: Recurrent Neural Network)은 순환구조를 통하여 과거의 학습한 내용을 현재 학습에 반영하는 구조의 모델
  - 반복적이고 순차적인 데이터에 대한 학습에 유리
  - 음성웨이블 폼을 파악하거나 텍스트 앞 뒤 성분 분석에 활용

## ■ 인공지능의 학습 방법

- 인공지능 모델을 데이터를 가지고 학습하는 방법으로 크게 지도학습, 비지도학습, 강화학습 등 세 가지로 구분됨
- 지도 학습(Supervised learning)은 학습하고자 하는 데이터에 대하여 정답을 설정하고 이를 학습하도록 함
  - 예를들어 하자보수 이미지 데이터를 제시할 때 해당 하자유형이 어떤 것인지 알려준 후 학습
  - 지도학습은 어느 유형에 속하는 지를 구분하는 분류 문제, 연속적인 값을 추정하는 추론문제 등에 적용됨
- 비지도학습(Unsupervised learning)의 경우는 데이터를 학습하는데 있어 해당 데이터에 대한 특정한 값을 제시하지 않고 학습을 수행시킴
  - 데이터들 간의 특징이나 특성을 학습하여 적절한 학습 후 제시된 문제에 대하여 데이터를 분류하거나 추천하는 결과를 제시함
  - 반드시 원하는 특성이나 구조를 학습한다고 보장할 수 없는 단점이 있음
- 강화학습(Reinforcement learning)의 경우는 최적의 값을 반복하여 학습하는 방법임
  - 인공지능 모델의 추론 결과를 가지고 해당 내용에 대한 결과를 개선하는 방향으로 다시 학습을 수행
  - 대표적인 적용 사례로 바둑이나 체스 게임 등을 들 수 있음
- 인공지능 모델은 올바른 결과를 도출하기 위해 학습 과정에서 학습데이터를 반복하여 학습
  - 전체 학습데이터셋을 학습하여 모델 내부의 파라미터를 변경하는 일련의 과정을 1 Epoch라고 함
  - 일반적으로 Epoch 횟수가 증가할수록 모델의 학습성능이 향상되지만 한계점을 지나칠 경우 오히려 성능이 낮아지는 과적합(Overfitting) 현상이 발생할 수 있음



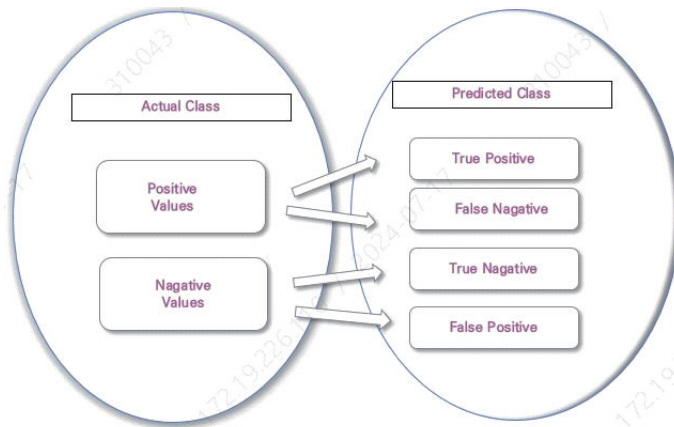
- 학습데이터를 활용하여 모델을 학습(훈련)시킬 경우 모델이 학습데이터를 어느정도 예측하는 지 측정 함(학습손실(Train Loss))
  - 예측하지 못하는 학습 손실 값을 줄이는 방향으로 모델을 개발
  - 학습이 완료된 모델에 대해 학습데이터 대신 검증데이터를 사용하여 테스트 결과 예측 정도를 나타내는 값을 검증손실(Validation Loss)이라 함
  - 모델의 일반화 정도를 나타낼 때 검증손실값을 활용하며, 검증손실과 학습손실 값이 크게 차이가 나지 않는 것이 이상적임
  - 성능평가는 검증손실값 등을 활용하여 모델의 우수성을 평가함

## ■ 인공지능의 성능 평가

- 인공지능으로 개발된 모델의 결과 값이 올바른 결과를 도출하는 지를 평가함으로써 해당 모델의 성능을 가늠해 볼 수 있음
- 인공지능 기술을 활용한 소프트웨어를 실제 적용하기 위해서는 적절한 성능평가가 요구됨
  - 도로의 신호인식, 질병진단 등 안전성과 비용적인 측면에서 민감한 결과를 판단해야 할 경우 오류가 발생할 경우 큰 문제점을 일으킴
  - 특히, 이미지 판별의 경우 입력의 범위를 한정하는 것이 어려움(촬영방법이나 외부적 노이즈 등의 변수등이 다를 경우 등)
  - 특정 오류를 수정하기 위해 재학습한 후 성능이 오히려 저하될 경우도 있어서 성능평가의 어려움이 있음
- 모델 학습과정에서 기본적인 인공지능 학습테스트 기법들이 있음
  - 메타모픽 테스트 : 인공지능의 경우 입력에 대한 출력의 정답을 알 수 없기 때문에 (정확한 테스트 데이터를 만들 수 없을 경우) 입력의 변화에 대해서 예측되는 출력의 변화를 기준으로 테스트를 수행
  - 뉴런커버리지 테스트 : 충분한 데이터셋으로 테스트를 수행할 수 있도록 가능한 많은 뉴런들이 활성화 되도록 데이터를 만들어 수행
  - 최대 안전 반경 테스트 : 입력데이터에 따라 출력이 크게 변경되는 현상을 막기 위한 크게 변하지 않는 범위를 구할 때 사용
  - 커버리지 검증 테스트 : 절대적으로 엄수해야 하는 조건에 대해 모든 입력에 대해 만족함을 검증할 때 사용

## ■ 인공지능 분류모델의 성능평가 지표

- 분류모델의 성능을 평가하기 위해서는 정확도(Accuracy), 정밀도(Precision), 재현도(Recall), F1 Score 지표 등을 활용함
- 오차행렬(Confusion Matrix)이란 학습을 통한 예측 성능을 측정하기 위해 예측 값(Predictive values)과 실제값(Actual values)을 비교하기 위한 표임
  - TP(True Positive) : 참을 참이라 예측한 경우
  - TN(True Negative) : 거짓을 거짓이라 예측한 경우
  - FP(False Positive) : 거짓을 참이라 예측한 경우
  - FN(Fase Negative) : 참을 거짓이라 예측한 경우



[그림2-2] 예측 값과 실제 값을 비교하기 위한 오차행렬 구성

- **정확도(Accuracy)**는 가장 직관적인 적합성 평가 지표임
  - 판별한 전체 값들 중 TP와 TN이 차지하는 비율을 의미함
  - 식은  $(TP + TN) / (TP + TN + FP + FN)$ 로 나타낼 수 있으며 높을수록 성능이 우수하며 모델 훈련 과정에서 최적화 목적함수로 사용할 수 있음
  - 예측한 값 중 참과 거짓을 현실 값에 맞게 명확하게 예측한 것들을 의미함
  - 분류 모델을 평가하기에 가장 단순한 지표이나 불균형한 클래스를 가진 데이터 셋을 평가하기는 어렵다는 단점이 있음(예: 10%의 참과 90%의 거짓으로 이루어진 데이터셋의 경우 모두 거짓으로 분류했을 때도 정확도가 90%로 측정됨)

- 정밀도(Precision)는 결과 예측값 중 참이라 예측한 경우 실제로 참일 경우의 비율을 의미함
  - $TP / (TP + FP)$  로 나타냄
  - PPV(Positive Predictive Value)라고도 함
  - Positive로 검출된 결과가 얼마나 정확한지를 의미
  - 높을수록 성능이 우수하며 예를 들어 예측한 값이 A라고 가정했을 때 실제로 A일 비율을 의미
- 재현율(Recall)은 민감도(Sensitivity) 또는 TPR(True Positive Rate)와 같은 의미로 실제로 참인 경우 중 예측값도 참인 경우의 비율을 의미함
  - $TP / (TP + FN)$
  - 높을수록 좋은 모델이며, 실제 A 중에 A라고 예측한 비율을 의미함
  - 모델의 실용성과 관련된 척도임(실제 Positive 클래스를 놓치지 않고 도출해낼 수 있는지를 나타냄)
- 특이도(Specificity)는 실제로 거짓인 경우 중 예측 값도 거짓인 경우의 비율을 의미함
  - $TN / (FP + TN)$
  - TNR(True Negative Rate) 또는 selectivity 라고 표현함
  - 재현율(Recall)과 반대의 개념
- 위양성율(Fall out)은 FPR(False Positive Rate) 실제로 거짓인 경우 중 예측 값은 참인 경우의 비율을 의미함
  - $FP / (FP + TN)$  으로 나타냄
  - 값이 낮을수록 우수한 모델임
- F1 Score는 정밀도와 재현율의 조화평균을 의미함
  - 정확도 대신 F1 score를 쓰는 경우는 데이터가 불균형할 때임
  - $2 * (Precision * Recall / (Precision + Recall))$
  - 정확도와 재현율을 동시에 높일 수는 없기 때문에 적절한 선을 찾기 위해서 조화평균으로 사용
  - 재현율을 높이기 위해서는 참으로 판단하는 기준(Threshold)를 낮추어 약간의 근거만으로 참으로 판단하도록 조정하는 기법을 활용할 수 있으나 그에 따라 위양성율이 증가할 수 있게 됨
  - F1-score는 0과 1사이 값이며 1에 가까울수록 분류 성능이 좋음을 나타냄

- ROC Curve(Receiver Operating Characteristic Curve) 는 분류 모델(이진 분류기) 과의 성능비교 및 적절한 Threshold를 찾고자 할 때 사용함
  - ROC 곡선의 경우 데이터셋과 별개로 일반적인 상황에서 모델 성능을 낙관적으로 평가함
  - True Positive Rate(TPR)와 False Positive Rate(FPR) 를 x, y축에 표시
  - 곡선의 휨정도가 좌상단에 붙을수록 분류 성능이 우수함
  - 랭킹, 추천, 광고 분야에서 데이터셋과 별개로 안정적인 모델 성능을 평가하기에 유리함
- P-R 곡선(Precision-Recall Curve)는 정밀도와 재현율 간에는 trade-off 관계가 있음에 따라 decision threshold를 통해 그러한 관계를 조절가능하고 정밀도와 재현율 사이의 관계를 표현하는 것이 Precision-Recall Curve 임
  - Precision과 Recall 값을 x, y 축에 표시함
  - 모델간의 비교를 하기 위해 각 모델의 P-R 곡선을 비교함
  - 클래스간 불균형이 심할 때 그에 따른 모델의 성능을 반영함
  - 이상 탐지분야에서 P-R 곡선을 이용하면 더 정확한 평가를 할 수 있음
- AP(Average Precision)는 분류 모델간의 정량적 성능을 비교하기 위해 P-R 곡선의 아래쪽 면적을 AP(Average Precision)으로 정의함
  - 그림 수식
  - multi-class 분류 모델에서는 각 클래스의 AP를 평균한 mAP(mean Average Precision)을 정의하여 각 모델을 비교 평가함

## 2. 인공지능 성능지표 관련 문헌 연구

[표 2-1] 인공지능 성능지표 관련 참고 학술 논문

제목	연도	수행기관	주요내용
인공지능 시스템의 성능평가 표준 :ISO/IEC TR 24029-1	2023	한국전기전자학회 논문지 (이성수)	인공지능 시스템의 성능을 평가하기 위해 국제표준인 ISO/IEC TR 2409-1의 인공지능 성능지표들에 대해 기술되어 본 과제에서 개발하고자 하는 AI 성능지표의 방향성을 제시하고 있음
데이터 기반의 충돌 성능 목표 수립 지표 개발	2022	한국자동차공학회 춘계학술발표대회 (서육환 외)	차량 충돌 성능 관련 안전성 관련 연구를 기반으로 32개의 요인들을 정의하고 안전등급지수요인의 중요도를 평가하여 안전등급지수를 제시함, 기술적 관점 뿐만 아니라 기업전략 관점의 충돌 성능 목표 제시하고 있음
지반정보 변화객체 탐지·추출시스템 개발	2021	한국측량학회 (김광수 외)	지하공간 통합지도의 갱신을 위하여 변화객체를 찾아내는 시스템을 설계하고 구현된 시스템의 성능을 평가하는 지표로 변화 객체 탐지율 및 추출률을 사용함. 본 연구에서 적용된 AI 성능지표 개발에 참고
실사와 BIM 기반 컴퓨터 이미지의 딥러닝 객체인식 성능비교 연구	2021	연세대학교 (최재진)	BIM 모델로부터 추출한 이미지 데이터와 하자의 특징을 사용해 이미지 데이터 셋을 구축하는 방법을 제안하고 해당 데이터 셋으로 학습을 통해 객체인식 시스템을 개발함으로써 시간과 비용 절감을 기대함. 본 연구의 성능지표 데이터 구축 방향에 참고
Transformer 기반 공공임대 주거하차 공사유형 예측 연구	2023	성균관대학교 (유지석)	신속한 하차처리를 위해 하차 접수단계에서 향후 예상되는 공사유형에 대한 정확한 식별이 필요함에 따라 텍스트임베딩 및 딥러닝 기법을 공사유형 분류 모델로 적용하여 높은 성능을 증명함. 본 과제의 AI 분류 기능 성능지표 개발에 참고
공동주택하차 자동분류 및 특성 분석을 위한 HCLA 하차 관리모델	2022	한양대학교 (김별)	공동주택 하차발생에서 잠재하차를 예측하여 관리하기 위해서 하차의 개별적 및 역학적 특성을 정량화하여 분석하여 중점 관리대상을 선정하는 하차관리 모델 제안함 본 연구에서 향후 발전방향에 참고
스마트 AI 기술을 활용한 주택유지보수 공사 통합관리솔루션 연구	2024	토지주택연구원 (이상훈 외)	LH 공동주택 유지보수공사 관리에 사용중인 모바일 앱들의 통합방안과 업무효율을 높일 수 있는 스마트기술개발하여 통합앱 고도화방안 제시 본 연구에서 AI 하차보수 적용 발전방안등에 활용

■ 인공지능 시스템의 성능 평가 표준 : ISO/IEC TR 24029-1 (2023, 이성수)

- 인공지능 시스템의 성능 지표를 보간(Interpolation)과 분류(Classification) 카테고리 분류

- Interpolation 카테고리에 해당하는 성능지표는 인공지능 시스템이 예측한 값이 실제 값과 어느 정도 근접한지 그 성능을 평가하는 지표
  - Classification 카테고리는 인공지능 시스템이 분류한 종류가 실제 종류와 얼마만큼 일치하는지를 평가하는 지표
- ISO/IEC TR 24029-1은 인공지능 시스템의 성능지표를 정량화하여 객관적인 수치로 나타내고 이를 비교할 수 있음
- Interpolation은 Root Mean Square Error(RMSE), Max Error(MAXERR), Actual/Predicted Correlation(PCC) 3개의 지표를 제시함
- RMSE는 예측 오차의 표준편차를 의미하며 RMSE가 작을수록 인공지능 시스템이 예측한 값과 실제 값이 차이가 없음을 의미함
  - Max Error는 예측 오차의 최대값을 의미하여 이는 예측한 값과 실제 값의 최대 차이를 비교
  - Actual/Predicted Correlation(PCC)는 -1에서 1사이의 값을 가지며 1에 가까울수록 예측값과 실제값의 상관관계가 높아서 예측값이 실제값과 거의 정비례하고 실제 값과 근접한 주위에 모여서 분포함을 의미함
- 시사점 : 본 연구에서는 수행하고자 하는 하자보수 관련 인공지능 성능지표는 분류 카테고리에 해당됨에 따라 그에 따른 인공지능 성능지표를 선정하고자 함

#### ■ 데이터 기반의 충돌 성능 목표 수립 지표 개발 (2022, 서욱환 외)

- 상품성, 지역 특성, 수익성 등을 고려한 효율적인 안전 등급 지수를 산출하는 모델을 전문가 기반의 모델과 시장 기반의 모델로 구성함
- 데이터에 대한 민감도 분석 및 다양한 특성을 반영하여 합리적이고 전략적인 성능 목표 수립 지표 개발
  - 안전 등급 지수 요인을 개발하고 그에 따른 중요도를 검토하여 안전지수 성능지표 제시함
- 안전 등급 지수 요인을 개발하고 그에 따른 중요도를 검토하여 안전지수 성능지표 제시함
- 충돌 성능 목표 설정을 위해 고려하는 요소와 중요도 제시가 가능
  - 평가 결과를 바탕으로 한 신차 개발 목표 적절성 비교 가능
  - 고객기반 모델링을 통해 Market Share 및 Cost 범위 예측이 가능함
  - 개발 투자비 적절성 평가 가능

- 충돌 안전성 데이터 분석에 머신러닝 모델을 활용
  - 영향을 주는 주요 설계 인자 도출 및 민감도 분석 가능
  - 주요 설계 인자에 대한 집중관리를 통해 효율적인 충돌 성능 관리 가능
- 시사점 : 인공지능 성능지표 개발 방향 수립 시 참고 가능
  - 개발프로그램의 단순한 성능지표 뿐만 아니라 업무담당자 및 사용자 측면에서의 성능지표 개발 필요
  - 주요 성능지표 관리를 통한 효율적인 시스템 운영 방법 모색 필요

#### ■ 지반정보 변화 객체 탐지·추출시스템 개발 (2021, 김광수 외)

- 지하공간 통합지도는 지하시설물, 지하구조물, 지반정보로 구성되며 주기적으로 갱신 필요
  - 변화된 지반정보만을 탐지 및 추출하는 시스템 개발
  - 객체 비교기, 변화 객체 탐지기, 과거 데이터 관리기, 변화 객체 추출기, 변화 객체 저장기 모듈 구성
- 시스템 성능 평가를 위해 객체 탐지율 및 추출률 지표를 선정
  - 탐지율은 재현율(Recall)의 개념으로 정의함
  - 탐지율은 (변화객체로 탐지된 수/ 변화객체 수)로 계산하고 추출률은 (실제 변화된 객체 수/ 변화객체로 타지된 수)로 계산함
  - 테스트하기 위한 데이터를 생성(속성의 추가 및 변경)하여 실험군을 정의
  - 기존 방법과 개발된 시스템의 갱신 시간 및 장점 비교
- 서로 다른 지방자치단에서 구축한 다수의 지반정보 지도에 적용하여 다양한 변화 객체를 찾아내어 적용 함으로써 안정성, 효율성, 확장성을 확보하여야 함
- 시사점 : 기본 인공지능 성능지표를 대체할 수 있는 지표를 제시하고, 테스트 환경을 세밀하게 구축 함(테스트 데이터 속성변경 및 갱신 기간 등)

## ■ 실사와 BIM 기반 컴퓨터 이미지의 딥러닝 객체 인식 성능 비교 연구 (2022, 최재진)

- BIM 모델로부터 추출한 이미지 데이터와 하자의 특징을 사용해 데이터셋을 구축하는 방법을 제안함
  - 하자 관련한 데이터 셋을 구축하기 어려워 모델의 학습과 테스트가 어려움
  - 건설 현장 BIM 모델로부터 생성된 렌더링 이미지와 하자가 포함된 이미지로부터 추출한 하자 객체로부터 이미지를 생성
- 인공지능 성능 비교를 하기 위해 성능 평균 정확도 (mAP(mean average precision)) 사용
- 실제 학습용 데이터를 구축하기 힘들 경우 근접한 데이터를 구축하여 학습
- 시사점 : 본 연구에서 적용하고자 하는 이미지 인식 인공지능 모델의 성능지표로서 성능평균정확도(mAP)를 사용한 점과 학습 및 테스트 데이터를 임의로 생성하여 구축함

## ■ Transformer 기반 공공임대 주거하자 공사유형 예측 연구 (2022, 유지석 외)

- 신속한 하자 처리를 위해서 최초 하자 접수시 '공사유형'에 대한 식별이 중요
  - Pre-trained Transformer 기반의 텍스트 임베딩 및 딥러닝 기법을 건축 분야 14개 공사유형에 대한 분류 모델에 적용
  - 기존 모델에 비해 91.5%의 성능 개선 증명
- BERT 임베딩 및 딥러닝 기반의 Multi-Class 분류 모델의 성능평가를 위해 정확도(Accuracy), 정밀도(Precision)와 재현율(Recall)은 데이터 불균형이 존재하는 데이터에서 필수적으로 검토 해야 하는 성능지표임
  - 정밀도는 특정 클래스로 분류한 데이터 중 실제로 그 클래스에 해당하는 데이터의 비율을 나타냄
  - 재현율은 실제 특정 Class 의 데이터(True Positive + False Positive) 중에서 예측이 맞은 데이터(True Positive)의 비율을 의미함
  - 성능평가 지표로서 F1-score(정밀도와 재현율의 조화평균 계산값)로 설정함
- 시사점 : 본 연구에서 적용하고자 하는 하자보수 유형분류 모델과 성능지표에 직접적으로 참고 가능



■ 공동주택 하자 자동분류 및 특성 분석을 위한 HCLA 하자 관리모델 개발(2022, 김별)

- 건설관련 하자 관리 및 방지의 중요성 증대되고 있으며 잠재하자를 예측 관리 필요
  - 하자의 개별적 및 역학적 특성을 정량화하여 분석 필요
  - 데이터의 입력(하자분류작업) 및 분석(하자관리전략 도출) 과정의 자동화를 통해 관리주체의 접근성을 확보하고, 정확성 및 전문성을 향상
  - 딥러닝 모델중 CNN(Convolution Neural Network)을 활용하여 자동 분류모델 제안
- 하자관리 업무 생산성 제고를 위해 LDA(Loss Distribution Approach) 활용
  - 하자 리스크의 프로파일을 확인하여 하자보수책임기간별 중점관리 하자 인지
  - ARM(Association Rule Mining)을 활용 각 연관성 분석을 통한 중점관리 하자 패키지 도출
  - 하자관리 우선 순위 설정하고 잠재하자/누락하자 예측으로 효율적인 관리 유도
  - 하자 항목에 대한 표준체계를 구성하고 지속적인 피드백 환경의 하자관리 플랫폼 환경을 구축
  - 하자의 공종, 위치, 요소, 유형을 모두 고려하여 하자의 손실을 정량화하고 하자 간 생성 구조를 명확히 규명
  - 하자를 사전에 예방하는데 효과적인 관리를 통하여 사회적 손실을 절감
- HCLA(Hybrid-CNN-LDA-ARM) 하자 관리모델로 분류 및 평가를 함으로써 관리 우선순위를 설정하고 잠재하자/누락하자들을 예측
- 작업 평균 F1-Score 점수가 90.72로 뛰어난 성능 제시
- 하자 항목에 대한 표준체계를 구성하고 지속적인 피드백 환경의 하자관리 플랫폼 환경을 제안
  - 공종, 위치, 요소, 유형을 고려 하자의 손실을 정량화하고 하자간 생성 구조를 정의
  - 하자의 세부요인과 개별적 및 역학적 특성을 통합적으로 고려한 관리 우선순위 패키지 제공
- 시사점 : 하자보수 관련 인공지능 적용 및 발전방안에 활용 가능함

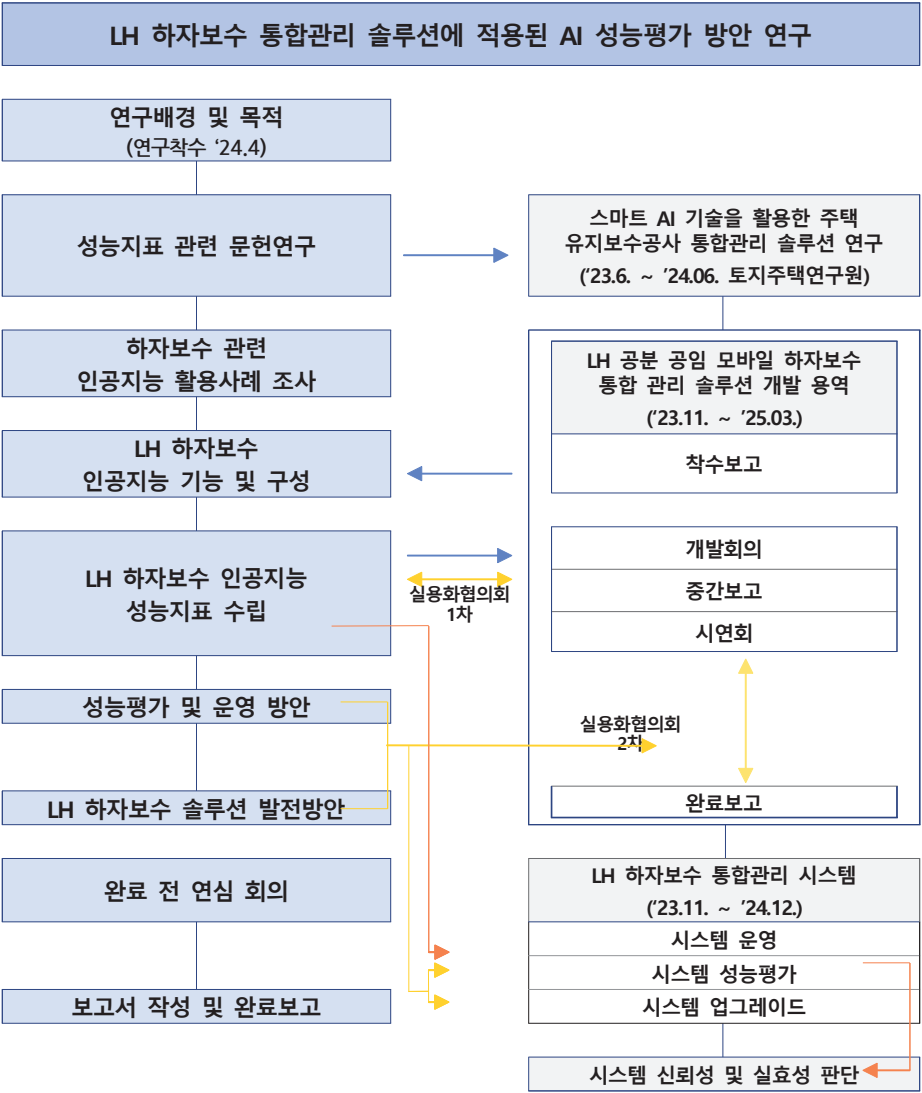
■ 스마트 AI 기술을 활용한 주택 유지보수 공사 통합관리솔루션 연구 (2024, 이상훈 외)

- LH 하자보수 관련한 앱들의 문제점을 파악하고 통합 구축방안 제시
  - 모바일 기술, AI 기술, 증강현실 기술 등 3가지 영역으로 구분하여 통합앱 구축을 위한 세부 기술요소를 검토함
  - 하자분류 가능성 검증을 위해 학습데이터와 테스트데이터를 이용하여 검증(하자유형 11개분류, 1300개 이미지 데이터로 학습 및 검증)
- 지속가능하고 효율적인 앱을 구축하고 운영하기 위한 방안 제시
  - 사용자 참여율을 주기적으로 분석하고 활용할 수 있는 기능 제공 필요
  - 하자인식, 분류, 견적서 산출 등을 AI 기술을 활용하여 자동화하고 데이터 기반 의사결정 지원체계 도입 필요
- 시사점 : 본 연구에서 평가하고자 하는 하자유형 분류 모델에 대한 시뮬레이션을 수행함에 따라 발생하는 문제점을 분석 및 통합관리솔루션 발전방안에 활용가능

■ 선행연구와의 차별성

- 선행연구들은 AI 성능지표를 개발하기 위한 방향성 및 기준으로 참고할 수 있음
- 해당 연구 중 LH에서 수행 중인 하자보수 통합관리 솔루션 개발 용역에 적용된 AI 기능들과 유사한 내용이 일부 있으나, 하자보수에 적용된 AI 모델이나 데이터에 적합한 구체적인 성능지표를 제시하고 있지 않음
- 특히, 그림에서처럼 기존 문헌 및 개발사례 등에서 도출한 시사점과 LH 용역사업인 “LH 공분·공임 모바일 하자보수 통합관리솔루션 개발”과 관련하여 개발업체와 협업을 통해 실용성 있는 연구를 수행할 수 있도록 함
- 본 연구에서는 “LH 하자보수 모바일 통합관리 솔루션”에 적용된 인공지능에 적합한 성능지표를 개발함에 차별화가 있으며, 신뢰성 및 실효성 판단의 객관적인 기준을 제시하는 데 의의가 있음
  - 또한, 향후 해당 시스템 업그레이드 시 요구되는 성능목표 설정 및 학습데이터에 대한 운영 방안을 수립하여 즉각적인 활용이 기대됨

[표 2-2] 연구 흐름도 및 과제 관련도



### 3. 하자보수 관련 인공지능 활용 사례

#### 1) 하자보수 관련 인공지능 모델

##### ■ 이미지 분류(Classification) 기법과 모델

###### ○ 전통적인 기법

- K-최근접 이웃 (K-Nearest Neighbors, KNN): 데이터 포인트의 K개의 가장 가까운 이웃을 기준으로 클래스를 결정하는 간단한 기법
- 서포트 벡터 머신 (Support Vector Machines, SVM): 데이터 포인트를 최적의 경계로 분리하여 분류하는 모델

###### ○ 심층 학습 기반 모델

- 합성곱 신경망 (Convolutional Neural Networks, CNN): 이미지 데이터를 처리하는 데 최적화된 구조로, 여러 층의 합성곱 필터를 사용하여 특징을 추출
- LeNet: 초기 CNN 구조로 손글씨 인식에 주로 사용
- AlexNet: 2012년 ImageNet 대회에서 큰 성과를 낸 모델, 깊은 층과 활성화 함수인 ReLU를 사용
- VGGNet: 심플한 구조의 깊은 CNN, 작은 필터 크기를 사용해 더 깊은 네트워크를 구현
- ResNet: 잔차 연결(residual connections)을 통해 매우 깊은 네트워크를 효과적으로 학습할 수 있게 함
- Inception: 다양한 크기의 필터를 동시에 적용하여 특징을 추출하는 구조

###### ○ 전이 학습 (Transfer Learning)

- 사전 훈련된 모델 사용: ImageNet과 같은 대규모 데이터셋에서 학습한 모델을 특정 작업에 맞게 조정하는 방식. 예를 들어, ResNet, Inception, EfficientNet 등의 모델을 Fine-tuning 할 수 있음

###### ○ 최신 기법

- Vision Transformers (ViT): 이미지 데이터를 패치 단위로 나누어 처리하는 변환기 기반 모델, CNN과는 다른 접근 방식을 사용하여 이미지 분류에 효과적임
- EfficientNet: 모델 크기, 연산량, 성능의 균형을 맞춘 최적화된 CNN 구조

## ■ 이미지 분석(Segmentation)기법과 모델

- 합성곱 신경망(CNN, Convolutional Neural Networks): 이미지 분류, 객체 감지, 이미지 세분화 등 다양한 이미지 관련 작업에 널리 사용됨, CNN은 필터를 사용해 이미지의 특징을 추출
  - YOLO(You Only Look Once): 실시간 객체 감지에 강력한 성능을 보이는 모델로, 이미지를 한 번만 보는 방식으로 빠른 속도임
  - R-CNN(Region-based CNN): 이미지 내 특정 객체를 찾는 데 사용되며, 선택적 검색을 통해 후보 영역을 추출한 후 CNN을 적용
  - Mask R-CNN: R-CNN의 확장으로, 객체 감지 뿐만 아니라 픽셀 수준의 분할(masking)도 수행
  - U-Net: 주로 생물학적 이미지의 세분화에 사용되는 모델로, 인코더-디코더 구조를 가지고 있어 정확한 세분화 성능
  - VGGNet: 이미지 분류 대회인 ImageNet에서 뛰어난 성능을 보였던 모델로, 깊은 네트워크 구조를 가지고 있음
  - ResNet(Residual Network): 매우 깊은 네트워크를 효과적으로 학습할 수 있도록 설계된 모델로, 잔차 연결을 통해 기울기 소실 문제를 해결
  - Generative Adversarial Networks (GANs): 이미지를 생성하는 데 사용되며, 두 개의 네트워크(생성자와 판별자)가 서로 경쟁하여 높은 품질의 이미지를 생성

## ■ 객체 탐지(Object Detection)기법과 모델

- 전통적인 방법
  - Haar Cascades: 특징 기반 접근 방식을 사용하여 객체를 탐지, 얼굴 인식에 주로 사용
  - HOG (Histogram of Oriented Gradients): 객체의 형태를 특징으로 잡아내는 방법 SVM과 함께 사용되어 객체를 탐지
- 딥러닝 기반 모델
  - R-CNN (Regions with CNN features): CNN을 사용하여 이미지에서 후보 영역을 추출한 후, 각 영역에 대해 분류기를 적용
  - Fast R-CNN: R-CNN의 개선된 버전으로, 후보 영역을 한 번만 추출하고, 그 위에서 CNN을 실행하여 속도를 높임

- Faster R-CNN: Region Proposal Network(RPN)를 도입하여 후보 영역을 더 빠르게 생성
- YOLO (You Only Look Once): 전체 이미지를 한 번만 처리하여 객체를 탐지하는 모델로, 실시간 객체 탐지에 적합, 각 버전마다 성능이 개선됨
- SSD (Single Shot Multibox Detector): 다양한 크기의 객체를 동시에 탐지할 수 있도록 여러 스케일의 피쳐 맵을 활용하여 속도와 정확성을 동시에 갖춘 모델
- 경량 모델
  - MobileNet SSD: 모바일 및 엣지 디바이스에서 실행할 수 있도록 최적화된 경량 객체 탐지 모델
  - Tiny YOLO: YOLO의 경량 버전으로, 속도는 빠르지만 정확도는 상대적으로 낮음
- 변환기 기반 모델
  - DETR (Detection Transformer): Transformer 아키텍처를 사용하여 객체 탐지를 수행. 객체 탐지 문제를 세트 예측 문제로 변환하여 혁신적인 접근 방식을 제공
- 분할 모델
  - Mask R-CNN: Faster R-CNN에 픽셀 수준의 분할 기능을 추가하여 객체 탐지뿐만 아니라 세분화도 수행

## ■ 텍스트 분류 모델

- 텍스트 분류 모델은 건설 산업, 특히 건축물의 하자보수 관리에서 중요한 역할을 하는 자연어 처리(NLP)기반의 AI 모델임. 이 모델은 건축물의 하자 및 결함을 기술한 텍스트 데이터를 분류하여 문제의 성격과 범위, 우선순위 등을 자동으로 파악하는 데 사용
  - 하자보수 텍스트 분류 모델은 주로 건설 현장에서 발생한 문제나 하자 신고서 등에서 발생하는 다양한 텍스트 데이터를 분류하는 데 사용됨. 텍스트를 분석하고, 문제의 종류, 위치, 우선순위 등을 파악하여 자동으로 분류하거나 우선순위를 매김
- 하자보수 텍스트 분류 모델의 구조
  - 하자보수 텍스트 분류 모델은 자연어 처리 기술을 활용하여 하자보수와 관련된 텍스트 데이터를 자동으로 분석하고, 해당 정보를 여러 카테고리로 분류하는 작업을 수행함

[표 2-3] 텍스트 분류 모델 구조

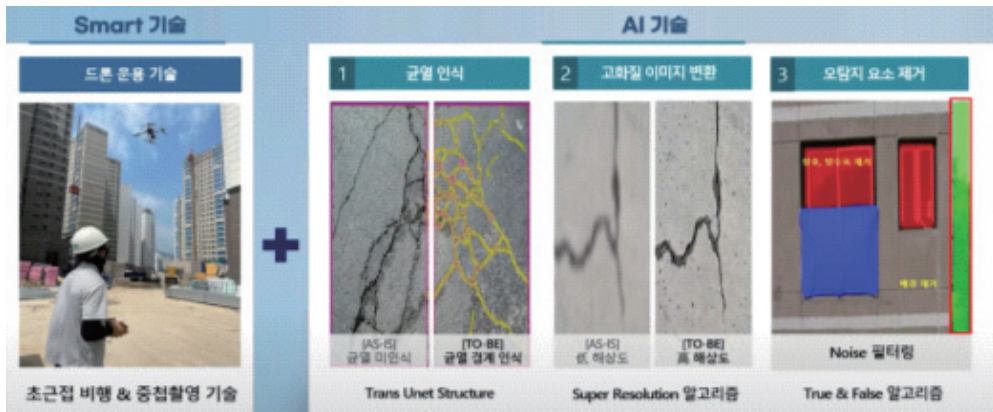
① 데이터 전처리	<ul style="list-style-type: none"> <li>• 하자 보수 관련 텍스트 데이터를 모델에 적합한 형태로 변환</li> <li>- 불용어 제거, 형태소 분석(예: "배수구 막힘" -&gt; "배수구", "막힘"), 정규화(예: 날짜, 시간, 특수 문자 제거), 단어 벡터화(Word2Vec, TF-IDF 등)를 포함</li> </ul>
② 모델 학습	<ul style="list-style-type: none"> <li>• 주로 사용되는 알고리즘</li> <li>- Naive Bayes: 텍스트 분류에서 기본적으로 자주 사용되는 통계적 방법</li> <li>- SVM (Support Vector Machine): 텍스트 분류의 정확도를 높이는 데 효과적인 기법</li> <li>- 딥러닝 모델: 최근에는 BERT, GPT와 같은 Transformer 기반 모델을 사용</li> <li>- LSTM: 문맥을 고려한 순차적 텍스트 분석이 필요한 경우 LSTM(Long Short-Term Memory)을 사용</li> </ul>
③ 하자 유형 및 우선순위 분류	<ul style="list-style-type: none"> <li>• 하자 유형(예: 전기, 배관, 구조적 결함)과 우선순위(긴급, 중간, 낮음)를 분류</li> <li>- 다중 분류 문제로서 각 하자에 대해 여러 개의 클래스를 분류할 수 있음</li> <li>- 모델은 하자 내용을 분석하여 해당 내용이 어떤 유형의 문제인지 파악</li> </ul>
④ 하자 이력 관리 및 예측	<ul style="list-style-type: none"> <li>• 모델은 기존의 하자 데이터를 기반으로, 향후 발생할 수 있는 하자 예측</li> <li>- 과거 하자 이력을 바탕으로 특정 지역이나 자재에서 반복적으로 발생하는 하자 유형을 예측하고, 사전 점검이나 보수 계획을 추천</li> </ul>

## 2) 하자 보수와 관련하여 인공지능(AI)을 활용한 사례

### ■ 자동 결함 탐지 사례

- 영상 분석: 드론이나 고해상도 카메라를 이용하여 건물 외벽, 도로, 교량 등의 이미지를 촬영하고, CNN과 같은 딥러닝 모델을 사용해 균열, 결함, 부식 등을 자동으로 탐지
  - 포스코이앤씨는 드론 기반의 AI 균열 관리 솔루션 'POS-VISION'을 개발, 이 시스템은 고화질 영상 장비 장착한 드론으로 외벽을 촬영해 균열의 크기, 길이, 위치를 상세 분석함. 특히 0.1mm 이상의 작은 균열도 감지할 수 있어 하자로 판별되는 기준을 충족하는 균열을 적시에 보수할 수 있음
  - POS-VISION은 외벽 품질 문제를 예방하기 위해 개발되었으며, 최근 9개 프로젝트에 적용되어 보수 작업의 효율성을 높임. 드론 운영기술에 균열 인식, 고화질 이미지 변환, 오탐지 요소 제거능력을 탑재한 인공지능을 접목함으로써 공사 담당자는 건물 외벽의 상세정보를 탐지할 수 있음<sup>3)</sup>

3) <http://newsroom.posco.com>를 참고하여 정리



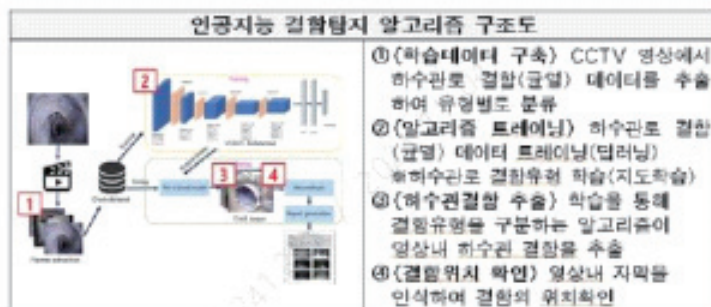
[그림 2-3] POS-VISION  
(출처: POSCO NEWSROOM, 2024.03.13.)

- 서울 디지털재단은 AI를 활용하여 결함유형 및 위치를 판별하는 ‘인공지능 기반 하수관로 결함탐지 시스템’을 개발, 기존에는 CCTV 영상자료를 검사관이 육안으로 판독해야 했으나 인공지능이 CCTV를 판독해 결함 유형 및 위치를 자동으로 식별함
  - 기존에는 하수관로 CCTV 영상 9000km를 검사관이 수작업으로 판독
  - 10종의 결함 유형을 도출하고 5000개의 CCTV 영상을 학습시켜 AI 모델 개발

○ 사람의 육안검사를 컴퓨터비전 기술 기반 인공지능이 검사 대체



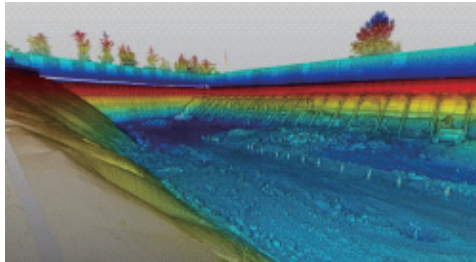
○ 인공지능 기반 하수관로 결함탐지 알고리즘 개요



[그림 2-4] 하수관로 인공지능 결함탐지 알고리즘 구조도  
(출처: 인디포커스, 서울시, 2021.06.23.)



- 롯데건설 AI 단열 설계 검토 프로그램 ‘인스캐너(INScanner)’<sup>4)</sup>
  - 이 모델은 콘크리트 벽체, 단열재, 창, 문 등의 건축 요소를 인식하고 분류하며, 사용자는 기존 도면을 업로드하면 AI 모델이 단열재 누락 여부를 분석하고 검출함
  - 롯데건설은 컴퓨터 비전 알고리즘을 사용하여 1000장 이상의 단열 설계 도면을 학습하고 있으며, 지속적인 데이터 추가를 통해 정확도를 높이고 있음
  - 스캐너로 모인 공간 정보는 공사현황 분석 및 공정관리에도 활용되며, 롯데건설은 구조물 형태의 변화 등을 정밀 관리하여 기술안전을 강화하고 시공오차를 관리하여 하자 없는 시공을 위해 이 장비를 지속적으로 적용할 예정
  - 단열재의 누락이나 미비로 인한 결로 및 곰팡이 등의 하자를 예방하기 위해, 기존의 수작업 검토 과정을 AI 기술로 대체, 설계 단계에서 발생하는 오류를 지속적으로 체크하고, 검토 시간을 크게 단축시켜 설계 품질을 향상시킴

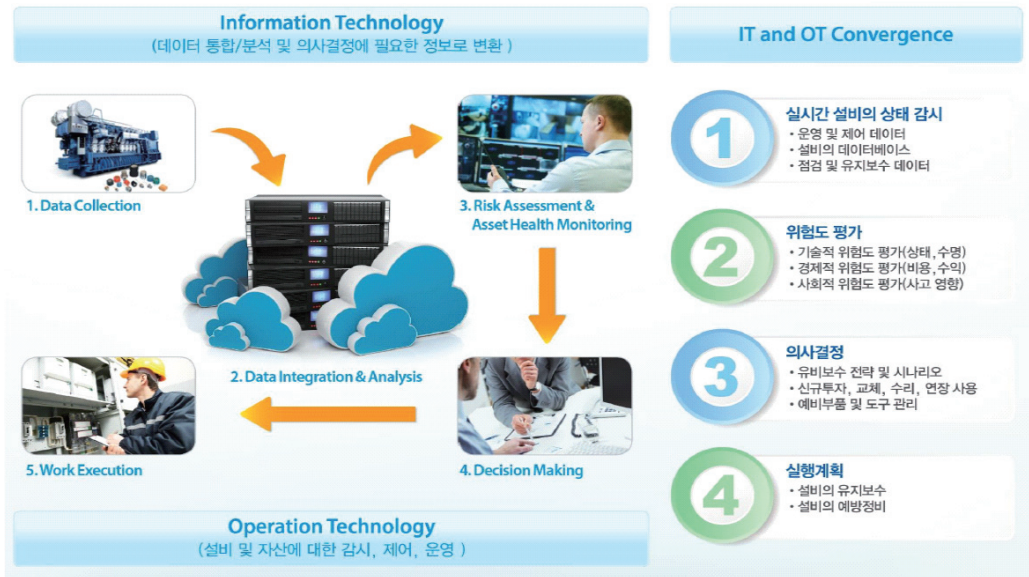


[그림 2-5] 스캔을 통한 3D 영상의 예  
(출처: 롯데건설)

## ■ 예측 유지보수 사례

- 데이터 분석: IoT 센서를 통해 수집한 데이터를 분석하여 장비의 상태를 모니터링하고, 고장이 발생하기 전에 유지보수가 필요할 시점을 예측, 인공지능 기법을 활용해 과거 데이터를 기반으로 고장 패턴을 학습함
  - 효성중공업이 1982년부터 지금까지 40년간 축적한 15,000건의 데이터를 개발, 제작 및 유지관리 하면서 축적한 빅데이터로 상태평가, 수명예측, 유지보수 및 교체 의사결정이 가능한 인공지능 기술 개발

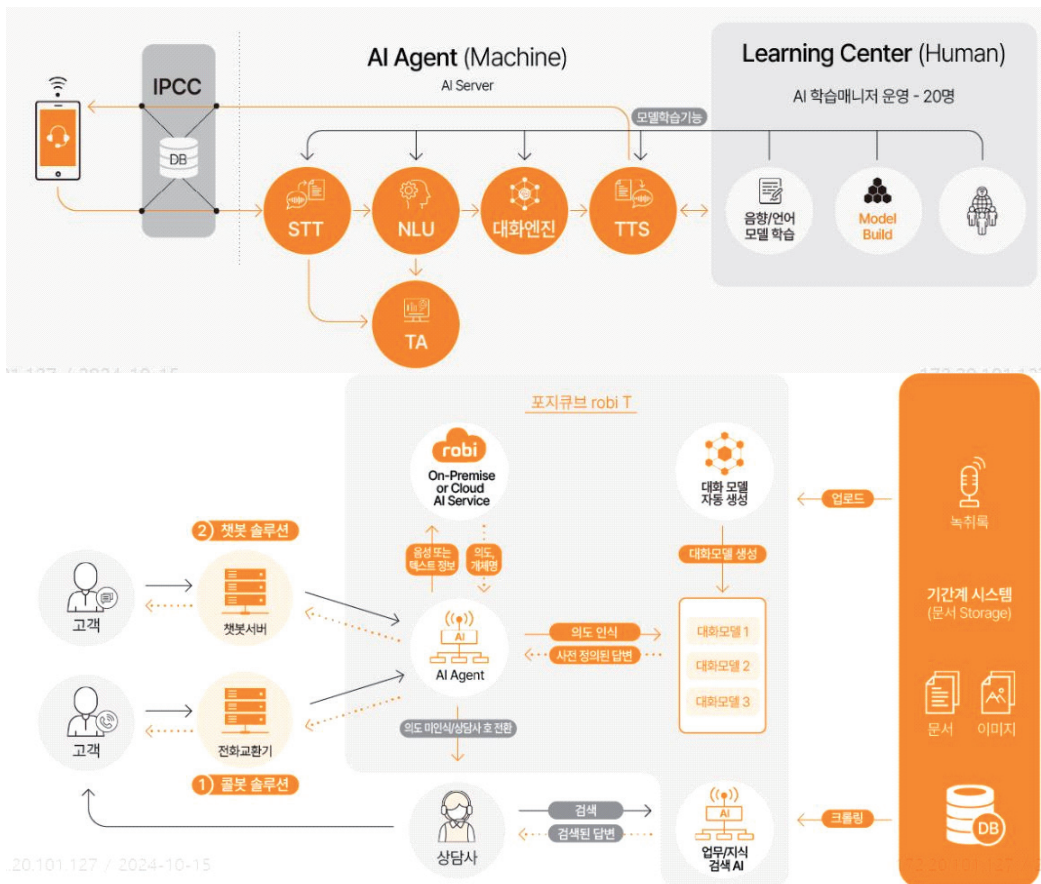
4) <http://n.news.naver.com>를 참고하여 정리



[그림 2-6] 효성중공업 '아모르' 운영 프로세스  
(출처: 기술과 혁신, 2022.09.10.)

## ■ 하자관련 인공지능 자연어 처리(NLP) 사례

- 건축물의 하자 신고서나 불만 제기 텍스트를 자동으로 분류하여 우선순위 선정
  - "전기 배선 문제", "벽에 금이 갔다", "배수구 막힘" 등의 하자 사항을 분류하고, 그에 따른 보수 작업을 긴급도에 따라 우선순위를 정함
- 고객 불만 분석: 고객의 하자 관련 불만이나 피드백을 수집하고, 텍스트 분석을 통해 문제의 유형과 빈도를 분석하여 빠르게 대응함,
  - 아파트아이 (컨텍센터 OS): 콜봇과 챗봇을 하나의 AI 엔진으로 동작하여 통합 상담 관리가 가능, 운영위탁을 통해 AI 학습센터를 운영하여 매일 전문 모니터링 요원이 AI 상담원의 오류를 검토하고, 매주 콜 데이터와 신규 도메인 특화용어를 학습시킴



[그림 2-7] 콜봇+챗봇 AI 엔진 서비스 프로세스 플로우 (출처: 포지큐브 홈페이지)

○ 자동화된 하자 분류 시스템

- AI 기반 모델을 사용하여 건설 현장에서 발생하는 다양한 하자 문제를 자동 분류함
  - 건설 현장 로그, 검수 보고서 또는 고객 피드백을 텍스트로 분석하여 하자 유형(전기, 배관, 구조적 결함 등)이나 위치 정보(벽, 바닥, 천장 등)을 추출

○ 하자 트래킹 및 관리 시스템

- 하자 발생에 대한 자동 추적과 관리 시스템을 통해, 하자 처리 상황을 실시간으로 추적함. 이를 통해 하자 처리의 상태나 보수 완료 여부를 기록하고, 이후의 예방 조치나 추가 점검을 자동으로 추천할 수 있음

■ 자동화된 보고서 작성

○ AI 기반 문서 생성: 하자 발생 시 AI를 통해 관련 데이터를 수집하고 분석하여 자동으로 보고서를 작성, 이 과정에서 템플릿을 활용해 문서를 생성

- 대우건설의 바로답(BaroDAP)<sup>5)</sup>은 인공지능(AI) 기반의 입찰안내서(ITB)와 해외 프로젝트 계약 문서를 AI를 활용해 정밀하게 분석하고, 프로젝트 수행을 지원하는 시스템이며, 고객이 대우건설에 대해 궁금한 점이나 요청 사항을 간편하게 해결할 수 있도록 지원함
  - '바로답'은 사전에 학습한 3억개 이상의 단어를 바탕으로 ITB의 문서 구조를 빠르게 분석해 공종별 카테고리에 맞춰 분류하고, 각 카테고리에 맞는 체크리스트를 제공해 발주처의 요구사항이나 제약사항을 파악할 수 있도록 도움
  - 다양한 분야 전문가들의 지식과 경험이 집약된 기존 'Rule 기반 지식관리 시스템'을 대규모 언어모델(LLM, Large Language Model) 기반으로 고도화하여 특히 건설산업에 최적화된 검색증강생성(RAG, Retrieval-Augmented Generation) 기술을 AI 에이전트(Agent) 형태로 구현해 복잡한 계약 문제를 다양한 관점에 분석할 수 있도록 함

① AI 챗봇: 고객의 질문에 실시간으로 응답하는 AI 챗봇 기능을 제공하여, 24시간 언제든지 상담을 받을 수 있음

② 정보 검색: 대우건설의 다양한 서비스, 프로젝트, 최신 소식 등을 쉽게 찾아볼 수 있는 기능이 포함

③ 고객 맞춤형 서비스: 고객의 요청에 따라 맞춤형 정보를 보다 나은 상담제공

5) <http://www.cnbnews.com>를 참고하여 정리

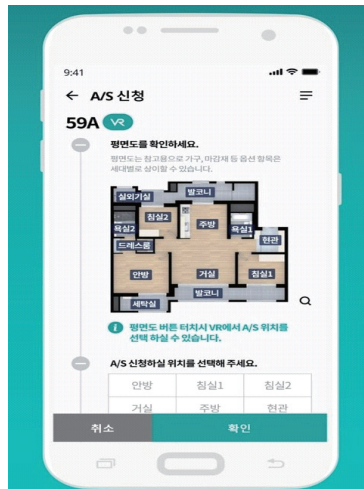
- ④ 통계 및 분석: 고객의 문의 패턴을 분석하여, 자주 묻는 질문이나 필요로 하는 정보에 대한 통계 데이터를 수집하고 이를 바탕으로 서비스 개선에 활용



[그림 2-8] 대우건설 '인공지능 계약문서 분석시스템 바로답 BaroDAP'  
(출처: 대우건설)

## ■ 인공지능 모델을 활용한 시뮬레이션과 최적화

- 구조적 시뮬레이션: AI를 활용한 시뮬레이션 기법을 통해 구조물의 하자 발생 가능성을 예측하고, 최적의 설계 방안을 제시
  - 삼성물산 '헤스티아2.0'
    - 삼성SDS는 메타버스 솔루션 기업 “평행공간”과 협업해 3차원 공간정보기반의 디지털 트윈 기술을 적용함. 업계 최초로 가상현실(VR)기능을 도입
    - 입주민은 VR을 통해 동일한 평형의 내부를 실제처럼 확인가능하고 가구배치나 인테리어를 하기위해 길이 측정도 가능함
    - 사진이 올라오면 1차적으로 AI가 분류 및 업체를 지정하고, 접수부위에 따라 담당 엔지니어를 실시간으로 분류함
    - 고객이 하자접수를 할 때 위치를 지정하면 좌표 값으로 저장이 돼서 AS를 처리하는 작업자들도 앱을 열면 위치를 정확하게 확인 할 수 있음
    - 인공지능(AI) 기반 이미지 분석 기술을 통해 고객의 AS 요청을 자동 분석하며, 처리 완료 후엔 완료 사진으로 결과를 확인하는 것이 가능



[그림 2-9] 삼성물산 '헤스티아 2.0'  
(출처: www.wowtv.co.kr)

## ■ 인공지능 모델의 프로젝트 관리 적용 예

- AI 기반 일정 관리: 하자 보수 일정과 자원을 효율적으로 관리하기 위해 AI 알고리즘을 활용하여 프로젝트의 진척 상황을 모니터링하고, 일정 지연을 예측

- 포스코이앤씨는 건설현장에서의 문제 해결을 신속하게 지원하기 위해 Quality AI System을 개발, 이 시스템은 일반적인 AI와 달리 건설 관련 전문 지식과 법규를 기반으로 하여, 직원들이 직면한 다양한 현장 문제에 대해 즉각적인 정보를 제공하는 데 중점을 두고 있음

- 사용자가 “외벽 커튼 월 유리공사에 대한 품질관리 주의사항”을 물으면, AI는 관련 국가 건설 기준코드와 품질관리 기준을 제시. 예를 들어 KCS 41 55 09 유리공사의 설치 공법과 관련된 내용을 인용하여 구체적인 지침을 제공함

- ① 주요 특징: 일반적인 AI가 건설 용어와 최신 법규 개정 키워드 검색에 취약한 점을 보완하기 위해, 포스코이앤씨는 방대한 건설 관련 지식 데이터베이스를 구축, 이를 통해 직원들은 필요한 정보를 빠르게 찾을 수 있음

- ② 즉각적인 추천: 시스템을 통해 직원들은 건설 현장 문제의 개선 및 조치 방안, 관련 법규, 표준 시방서 등을 즉시 추천받을 수 있음. 추천된 정보는 출처와 링크가 함께 제공되어 확인이 가능함

- ③ AI와 전문가 협업: AI 답변 외에도 사내 전문가(마스터, 엑스퍼트)와 연결하여 AI와 사람이 협업할 수 있는 플랫폼을 제공, 이를 통해 더 정확하고 신뢰성 있는 정보를 얻을 수 있음



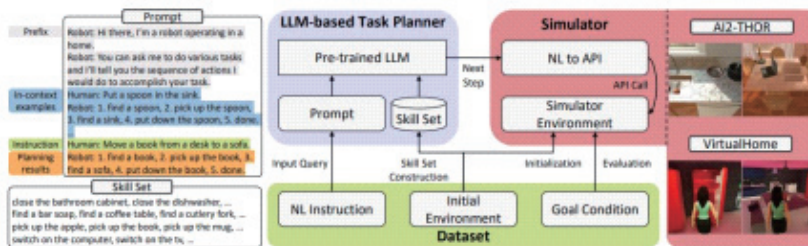
○ 설계 문서 관리 및 분석에 활용되는 자연어 처리 모델

- Adobe Sensei: Adobe의 ai 플랫폼으로, 설계 문서 및 이미지에서 정보를 추출하고 분석하는 기능을 제공
- IBM Watson의 NLP: 자연어 처리 기술을 활용해 계약서, 설계 문서 등을 분석하여 규정 준수 여부를 평가하는데 사용
- Kira Systems의 NLP: 건축 관련 법적 및 관리 업무를 분석 및 자동화하여 리스크를 사전에 파악하며 프로젝트의 일정 초과, 비용 초과, 변경 사항 등을 자동으로 추적 가능하여 적시에 문제를 파악하고 대응함, 계약 위반, 일정 지연, 품질 문제 등에서도 미리 예측하고 법적 대응 방안을 도움

■ 초거대 AI 모델 성능 평가 사례

○ ETRI, LLM 33종 절차 생성 성능 평가 '로타벤치마크(LoTa-Bench)'

- 대형 언어 모델(LLM)을 기반으로 생성하여 절차생성 AI의 성능을 자동으로 평가할 수 있는 로타벤치마크(LoTa-Bench)은 사용자가 말로 기술한 작업 명령에 따라 생성한 작업 프로세스를 실행하고, 목표 달성 가능성을 자동으로 평가함
- 기존에는 여러 사람이 직접 성능 평가에 시간과 노력을 기울이고 평가자들의 주관적 판단이 포함되었으나 이 모델로써 시간과 비용을 최소화하고 객관성을 확보함



[그림 2-10] 로타벤치마크(LoTa-Bench)

○ 가계정보자원관리원(NIA) 'Open Ko-LLM 리더보드'

- 국내 최대 개방형 한국어 LLM 평가 체계이며, 대형 언어 모델(LLM)의 성능 평가를 위해 국가정보자원관리원(NIA)와 업스테이지가 공동 운영하는 공개형 평가임
- 기존의 5개 평가 지표를 넘어 9개의 새로운 성능 지표를 장착하여 새로운 지표에는 추론 능력, 감성 분석, 무해성 등이 포함되어 모델의 언어 이해 능력과 윤리적 측면까지 대조하여, LLM의 다양한 성능을 더욱 다양하게 평가함

성능지표	주요내용	출처
추론능력	· (Winogrande) 글 속 대명사에 대한 독해 평가를 통해 추론능력 측정	허깅페이스 시즌1
산술추론능력	· (GSM8K, Grade School Math 8K) 대규모 초등학교 수학 문제를 바탕으로 산술 연산 분야 추론능력 평가	허깅페이스 시즌1
정보추출능력	· (Ko-GPQA) 대학원 수준의 Google 검증 Q&A 벤치마크	허깅페이스 시즌2
지시 이행 능력	· (Ko-IFEval) 주어진 지시를 얼마나 잘 이해하고 따르는지를 평가	허깅페이스 시즌2
감성 평가	· (Ko-EQ-Bench) 대화 맥락에서 다양한 감정과 사회적 상호작용을 이해하는 능력 검증	한국어 LLM 리더보드 시즌2
정보정확성	· (KoNAT-Knowledge) 국가 구성원들의 공통 지식에 대해 LLM이 얼마나 잘 알고 있는지 평가	23 NIA 인증지능 학습용 데이터 구축 사업
소셜 얼라인먼트	· (KoNAT-Social-Value) 국가 구성원들의 사회적 가치관에 대해 LLM이 얼마나 일치하는지 평가	
무해성	· (Ko-Harmlessness) 사회적으로 유해할 수 있는 영역에 대해 LLM이 얼마나 무해한지 평가	
도움 적정성	· (Ko-Helpfulness) LLM이 사용자 의도에 따라 권리의 유용성을 얼마나 잘 판단할 수 있는지 평가	

[그림 2-11] OPEN Ko-LLM 리더보드 시즌2

○ 카카오 '평선 챗 벤치(FunctionChat Bench)'

- 대형 언어 모델(LLM)의 성능을 평가하며 특히 대화형 AI 모델에서 명령 실행 능력과 절차 이해를 평가하는 데 초점을 맞춰 평가 기준임
- 평선콜(Function Call, 함수호출)은 AI 언어모델이 자체적으로 수행할 수 없는 동작을 지시하거나 사전에 학습하지 않는 정보를 실시간 정보를 받아올 수 있도록 언어모델과 API 등의 외부 도구를 연결하는 기술을 의미
- 데이터셋은 함수 이름과 인자추출의 정확성, 함수 호출 결과 전달의 정확성, 누락 정보 인지를 통한 추가 질의 발생 여부, 호출 가능한 함수와의 관련성 감지 등을 평가하는 기준 항목으로 구성



# 제3장 LH의 하자보수 통합관리솔루션

## 인공지능 성능평가 방안

### 1. LH의 하자보수 인공지능 활용 기능 및 구성

#### 1.1 LH의 하자보수 통합솔루션 구성

##### ■ LH 하자보수 업무처리 시스템 현황

- LH가 유지보수해야 할 공동주택(장기임대, 공공임대 등)은 매해 증가추세에 있고 유지보수 물량은 한 해 약 88만건에 이르고 있으며, 예산 또한 1조원에 다다르고 있음<sup>6)</sup>
- 공사는 입주품질관리시스템인 “QR(바로처리)”, 업무효율성을 높이고자 개발된 바로픽스, 고객편의성을 위한 카카오톡 AS 등 세 가지 프로그램으로 하자보수 업무를 수행하고 있으나 다음과 같은 문제점이 있음
  - 바로픽스의 경우 대상이 LH 담당자와 협력사만 사용함으로써 사용률 및 효율이 낮음, 또한 바로픽스(COTIS)는 유지관리 업무외 11개의 다른 업무를 병행하고 있어서 유지관리업무에 최적화되기 어려움
  - 바로처리의 경우 입주시 점검에 효율적이나, 담당자들이 활용하는 바로픽스와 연계가 되어있지 않은 관계로 데이터 연동 문제 및 업무 효율 저하
  - 카카오톡 AS의 경우 고객방문일정관리, AS진행정보 제공 등 사용자 편의성을 도모하고 있으나 바로픽스와 연계되지 않아 업무담당자와 보수업체 간의 업무 효율 저하
  - 공공클라우드를 통해 바로처리와 카카오톡 AS에서 발생하는 정보(고객, 업무담당자)는 외부망에 저장되고 있어 내부에서는 이를 활용할 수 있는 분석이 되고 있지 않음
  - 특히, 바로처리의 경우 하자 접수부터 처리 완료까지의 사진(보수전,중,후)을 앱에 저장한 후 COTIS 시스템에 이중 등록하는 중복작업을 수행 중임
  - 바로처리의 경우 하자접수 기능만 지원하고 있어서 하자처리 진행상황이나 관련 문의 등의 쌍방 의사소통을 할 수 없음
  - 하자보수 분야의 경우 고객이 접수시에 공종 및 부위자재 등 선택하기가 어렵고 담

6) 이상훈외(2024), “스마트 AI 기술을 활용한 주택 유지보수 공사 통합관리솔루션 연구”, 토지주택연구원

당사 역시 고객의 입력한 항목을 다시 확인하여 재분류하고 보수업체를 배정하는 단 순작업을 반복하고 있는 실정임

- 하자보수 관련 분석을 통하여 하자보수 시기를 예측하고 미리 서비스할 수 있는 체 계를 갖추고 있지 못함

[표 3-1] LH가 운영중인 하자보수 관련 프로그램 [출처: 이상훈외(2024)]

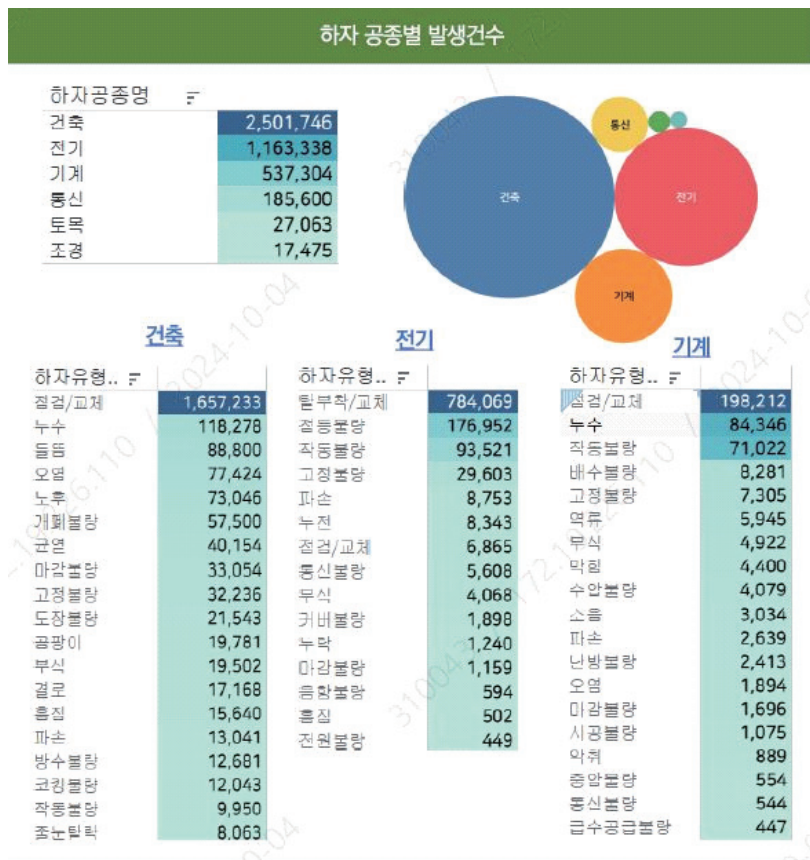
구분	QR(바로처리)	바로픽스(COTIS)	카카오톡 AS
개발 기간	'20.3월~'21.1월	(1차) '18.6월~'19.9월	(1차) '15.3월~'15.06월 (2차) '20.8월~'20.12월
목적	- 입주 후 민원 저감 - 처리 결과 즉시 피드백, 처리 결과 원격 확인/평가	- COTIS 연계(모바일COTIS) - 유지관리시스템 업무효율성 향상을 목적으로 모바일 앱 개발(일반사용자용 X)	- 고객 방문 일정관리 - AS 진행정보 제공 - 양방향 문자서비스
사용자	전체	하자민원접수(관리소)/ 유지보수업체	사용자 / 유지보수업체
대상	장기 임대 (Q+웰컴데이 ~ 보수확인평가)	전 지구	전 지구
기타	장기 임대 확대 적용 중 COTIS 자동 이관 구현 예정	-사용률 저조 -최초 방문 예정일 콜센터 조율	(3차) 실시간 스케줄링 가능 구현 예정

## ■ LH 하자보수 처리 시 개선 사항

- 입주 후 하자보수처리(바로처리) 뿐만 아니라 주거 생애 전체에 걸쳐 하자 관련 서비 스를 처리 해주는 통합 앱이 요구됨
- COTIS 시스템 등록되는 하자 접수 건수는 1개월 평균 약 50만 건이 생성되고 있으 며, 바로처리 시스템은 1년 약 16만 건이 접수되고 있는 실정임
  - 하자접수 및 처리에 신속하고 효율적인 업무 처리가 요구됨
- 사용자들이 하자보수 신청 시 해당 하자 건에 대한 공간과 부위자재, 하자유형 등을 입력해야하는 어려움 해결 필요
  - 그림에서처럼 2023년 3월에서 8월까지 6개월동안 바로픽스(COTIS)에 축적된 하 자보수의 경우 건축>전기>기계>통신 순으로 많이 발생함
  - COTIS와 바로처리시스템에서 수집되는 이미지 해상도의 차이가 커서 인식 프로그 램을 개발하는데 어려움이 발생함에 따라 수집하는 단계에서 수집 이미지의 기준을 제시할 필요가 있음
  - 수집 이미지에 그림자가 있을 경우 하자유형 분류에 오류가 발생하기 때문에 촬영

시 가이드라인을 제공하고 데이터 검증단계(시공사 또는 유지관리업체 등)가 제시되어야 함

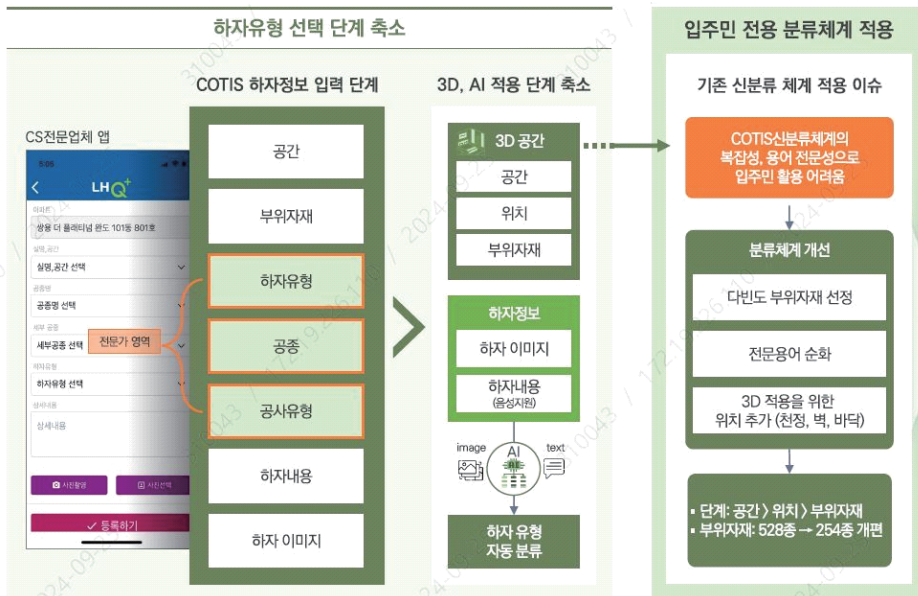
- 하자유형 분류 프로그램을 고도화하기 위해서는 다양한 학습데이터 수집 체계가 마련되어야 하며, 그에 따른 프로그램 성능향상 업그레이드가 주기적으로 이루어질 수 있어야 함
- 이미지에 대한 자동 레이블 작업이 이루어질 수 있어야 함



[그림3-1] 하자공종별 발생건수의 예(참고: LH 내부자료)

- 사용자 편의를 위한 하자 유형 설정 단계 축소 필요
  - 하자유형을 선택하기 위한 기존 바로픽스(COTIS)의 분류체계의 경우 공간을 선택하고 부위자재를 분류한 후 하자유형 3,580 종류로 복잡하게 구성되어 있어 입주인이 하자공간이나 유형을 선택하는데 어려움이 있음
  - 용어의 전문성 보다는 사용자 편의성에 맞춘 유형 분류체계 확보가 요구됨

- 복잡하고 다양한 유형을 분별하기 쉬운 체계로 단순화 할 필요가 있음
- 그림에서처럼 사용자가 공종이나 공사유형을 모르더라도 공간과 위치 등을 선택함으로써 자동으로 분류체계를 설정하고 그에 따른 정보를 수집할 수 있어야함



[그림3-2] 사용자 편의성을 위한 하자유형의 분류체계 확보(참고: LH 내부자료)

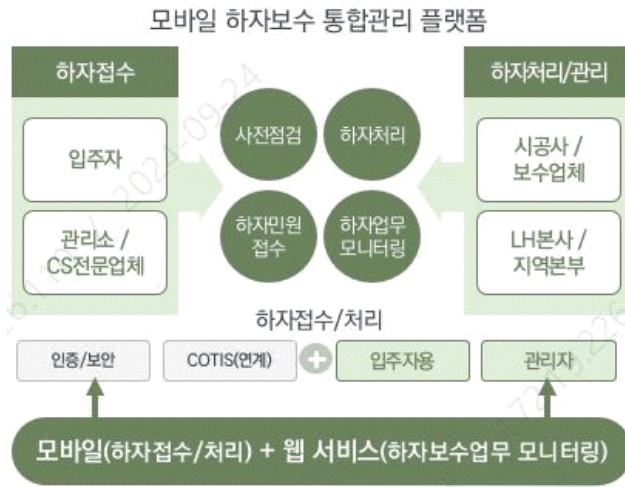
- 하자 유형을 이미지만으로는 확인할 수 없는 문제 해결 필요
  - 소음, 점검, 탈부착불량, 작동불량, 점등불량, 마감불량, 역류 등 이미지만으로는 유형을 분류할 수 없을 뿐만 아니라 입주민이 이미지 등록이 어려운 유형들이 있음
  - 사용자들이 편리하게 선택할 수 있는 방안 마련 필요
  - 해당 유형들을 학습할 수 있는 데이터 확보 및 분류 알고리즘 개발 필요

하자유형명	≡
점검/교체	1,889,585
탈부착/교체	918,931
작동불량	281,940
접촉불량	179,975
노후	77,558
마감불량	40,630
코킹불량	17,715
배수불량	14,479
방수불량	12,685
통신불량	9,570
소음	8,817
누전	8,651
줄눈탈락	8,063
손잡이불량	7,833
역류	6,998
신규(추가)설치	6,047
막힘	5,417
장금불량	5,225
수압불량	4,868

[그림3-3] 하자유형 분류가 어려운 예

#### ■ LH 모바일 하자보수 통합관리솔루션 개발

- 이러한 문제점을 해결하고자 공사는 ‘23.11 ~ ’24.12(14개월) 기간 동안 “LH 공분·공임 모바일 하자보수 통합관리솔루션 개발” 용역을 시행중임
  - 사용자 및 업무확대에 따른 접근이 용이하고 사용성이 향상된 고객중심 서비스 제공 및 실시간 업무모니터링을 목표로하는 플랫폼 개발을 목표로 함
  - 고객편의성을 도모하기 위해 모바일에서의 하자보수 접수를 지원하고 관리자는 웹 서비스를 통한 모니터링을 효과적으로 수행
  - 기존에 사용하고 있던 세 가지 프로그램을 통합하여 사용자 편의성을 높이고 데이터의 연동을 통한 업무 효율화 및 신속하고 정확한 업무 처리 효과 기대
  - 하자보수 관련 통합된 데이터 확보로 향후 예측 및 진단에 관련한 다양한 서비스 도출이 용이하며, 하자보수 관련 인공지능 모델 개발 시 고품질의 학습 및 테스트 데이터 제공이 가능함



[그림3-4] LH 모바일 하자보수 통합관리 플랫폼 구성도

- 해당 플랫폼은 입주자와 관리소/CS전문업체의 하자접수를 통하여 시공사/보수업체와 LH본사/지역본부의 업무담당자들이 하자처리, 사전점검, 민원접수, 모니터링 등의 종합적인 기능을 하나의 통로로 수행할 수 있도록 함
- 모바일 하자보수 통합관리 플랫폼의 시스템 구성은 그림에서처럼 기존모바일 서비스에서 수행하고 있는 하자신청과 접수, 조회, 완료보고 등의 업무와 입주자들의 모바일 서비스 및 관리 모니터링 기능 등을 통합하여 제공함

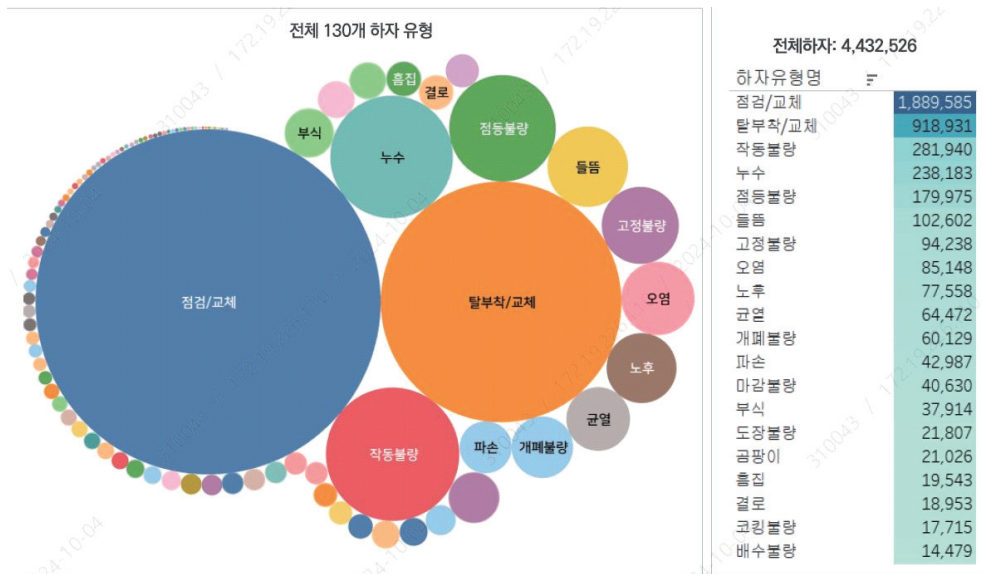


[그림3-5] LH 모바일 하자보수 통합관리 플랫폼의 시스템 구성도  
(출처: LH 내부자료)





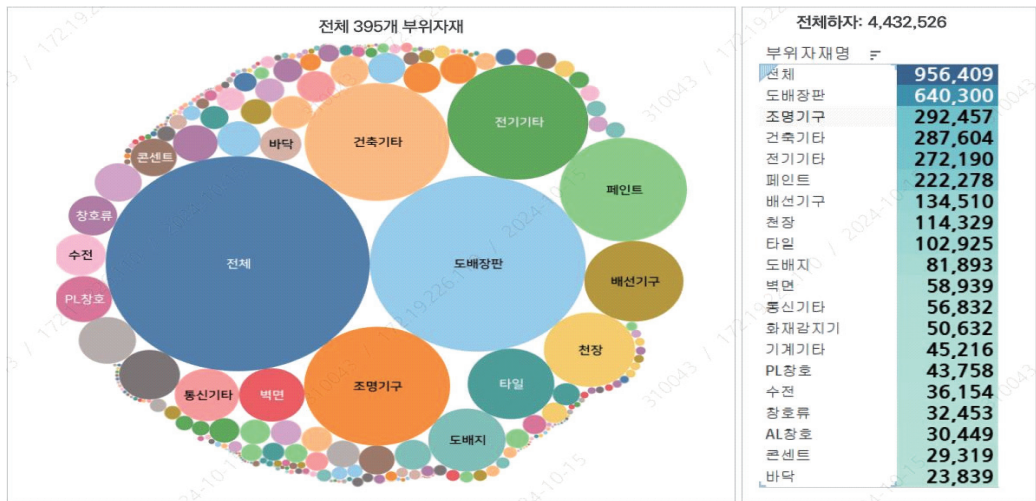
- 기계의 경우 점검교체> 누수> 작동불량> 배수불량> 고정불량> 역류> 부식> 막힘 등의 순으로 하자가 발생함
- 하자공사 종류별 발생 건수는 퇴거공사> 조명설비공사> 도배공사> 창호공사> 방수공사> 소방설비공사> 장기수선> 교체공사 등의 순으로 나타나고 있음
- 전체하자 중 130개 하자 유형 중 주요 하자유형 발생건수는 그림에서처럼 점검교체> 탈부착/교체> 작동불량> 누수> 점등불량> 들뜸> 고정불량 등의 순으로 나타나고 있음



[그림3-7] 하자보수 유형별 발생건 수 (출처: LH 내부자료)

- 부위자재별 하자 발생건수는 그림에서처럼 전체 395개의 부위자재 중 도배장판> 조명기구> 건축기타> 전기기타> 페인트> 배선기구> 천장 등의 순으로 나타나고 있음
- 특히, 기존 COTIS에 등록된 하자 이미지의 경우 전체 하자 이미지 4,432,526건 중 81%에 해당하는 3,579, 843건이 이미지를 분류할 수 없는 유형(점검교체, 탈부착교체, 작동불량, 점등불량, 노후 등)으로 나타나고 있어서 인공지능 모델을 학습하거나 테스트할 수 있는 데이터로 활용이 어려움





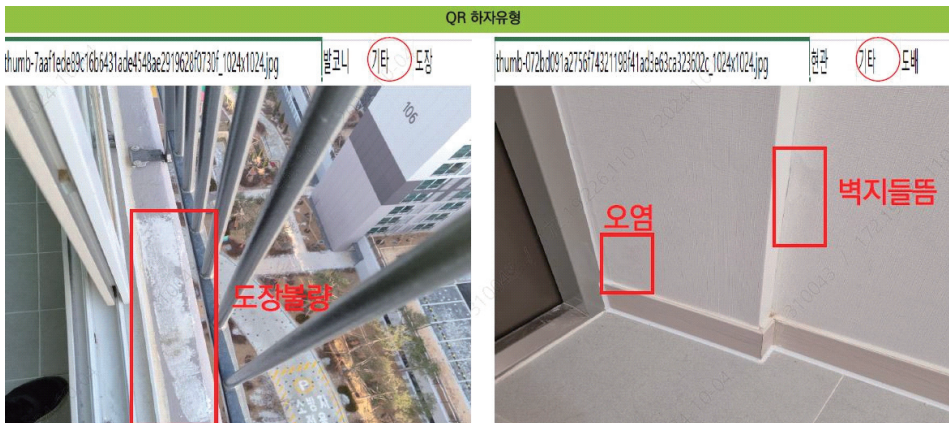
[그림3-8] 하자보수 부위자재별 발생 건수 (출처: LH 내부자료)

- COTIS에 등록된 하자 이미지의 경우 분석한 결과 분류할 수 없는 이미지가 많고 실제 하자유형과 맞지 않은 경우와 그림에서처럼 하자유형이 유사하지만 다르게 분류되어 있는 경우가 많이 발생하고 있음



[그림3-9] 같은 유형을 다른 유형으로 분류된 하자 유형(예:도배불량)

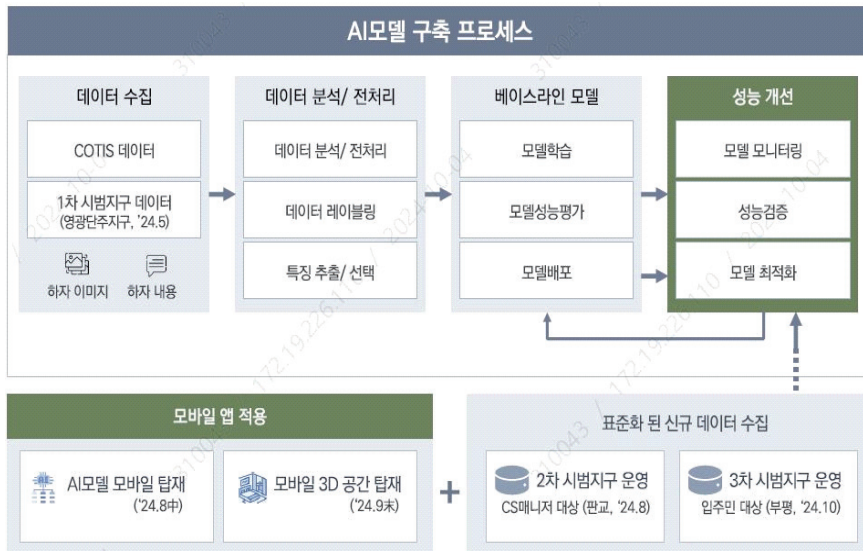
- 바로처리(신축)의 경우 근경과 원경이미지로 구분하여 등록을 하였으나 실제 근경이 미지에 원경이미지로 등록된 경우가 많고 육안으로 구분하기 어려운 경우의 이미지 도 많이 발생함
- 바로처리 데이터의 경우 입주민이나 매니저가 판단한 하자유형이 작성되어 있으나 사람마다 하자유형을 판단하는 기준이 달라서 일관성 없는 분류가 많음



[그림3-10] 하자 유형 임의 분류 및 육안 확인 어려운 이미지의 예

### ■ 인공지능 모델 구축

- 하자보수 유형 분류 인공지능 모델을 구축하기 위해서 데이터를 수집하고 데이터 전처리를 통한 학습을 한 후 성능평가를 통한 모델 최적화 작업 후 최종 모델을 구축함
  - 데이터 수집은 COTIS에 등록된 이미지 데이터를 기반으로 하고 추가로 '24.5월 영광단주 지구의 자료를 수집하여 학습 및 테스트를 수행함
  - 데이터 분석 및 전처리 단계에서 이미지 촬영 오류로 인한 분류가 불가능하거나 분류할 수 없는 데이터(소음, 통신불량, 누전 등)들을 제외하고 학습에 필요한 데이터 레이블링 작업 수행
  - 학습된 모델을 성능평가를 수행하고 모델 모니터링을 통한 성능개선을 통하여 모델을 최적화한 후 배포하는 단계 수행
  - 특히, 분류모델 인식에 따른 사용자 시범 적용을 '24년 10월에 수행하고 그에 따른 개선 방안을 모색할 예정임



[그림3-11] AI 모델 구축 프로세스  
(출처: LH 내부자료)

## ■ 하자유형 분류 체계 개선

### ○ 하자유형 분류 모델 학습을 위한 하자유형 분류체계 개선

- 그림에서처럼 바로픽스(COTIS)에 저장된 하자 이미지의 경우 그림에서처럼 이미지를 분류할 수 없는 75종을 제외한 55종의 유형을 1차로 구분하고 유사성을 감안하고, 데이터가 부족한 27종을 제외하여 2차로 17종의 하자유형 분류체계를 개선함
- 입주품질관리시스템인 바로처리(QR)에 저장된 하자 이미지의 경우에는 신축아파트의 하자에 관련되어 기존 하자유형 보다는 다양하지 않음
- 이에 따라, 데이터가 부족하여 하자분류가 어려운 22종의 하자유형을 제외하고 39종의 하자유형에서 이미지 분류가 가능한 17종을 선택하고,
- COTIS에서 선정된 하자유형 분류체계에 매칭하여 신축 하자 이미지 분류 유형은 11개로 선정하여 학습 및 테스트 수행함

전체 하자유형 130에서  
이미지 분류 할 수 없는  
75종을 제외한 55종 선택

COTIS 하자유형(55종)	
누수	
오염	
곰팡이	
결로	
균열	
콘크리트균열	
고정불량	
이음불량	
도배불량	
들뜸	
마감불량	
벗겨짐	
도장불량	
박리	
연결불량	
이격	
구멍	
파손	
손잡이불량	
창금불량	
손잡이미설치	
시트지불량	
코킹불량	
줄눈탈락	
실리콘불량	
개폐불량	
흡입	
타일파손	
...	
포장불량	
녹발생	
접착불량	
무너짐	
표면요철	

28종

하자분류를 위한  
데이터 부족

27종

COTIS간 유사 하자유형	
곰팡이, 오염, 녹발생, 균열	
결로, 부식, 변색, 먼지불량	
결로, 도배불량	
곰팡이, 누수	
백시멘트불량	
콘크리트철근노출, 타일불량	
들뜸	
도배불량, 이격	
도장불량	
도배불량, 누수, 구배불량	
들뜸, 도배불량, 균열	
누락, 들뜸	
곰팡이, 벽지들뜸, 균열	
도장불량	
개폐불량, 구멍, 누수, 도배불량, 단차	
구멍, 콘크리트 균열	
손잡이미설치, 개폐불량	
손잡이미설치, 손잡이불량, 부식, 노후	
손잡이불량, 개폐불량	
벗겨짐	
실리콘불량, 이음불량, 줄눈탈락, 백시멘트불량	
코킹불량	
손잡이불량	
파손, 들뜸, 도배불량	

선택된 55개의 하자유형과 유사한  
유형들을 그룹핑하여 하자유형 재분류

SI 모델 하자유형(17종)	
곰팡이(오염)	
곰팡이(오염)	
곰팡이(오염)	
곰팡이(오염)	
균열	
균열	
도배불량	
도배불량	
도배불량	
도배불량, 시트지 벗겨짐, 타일 들뜸, 콘센트 불량	
도배불량, 시트지 벗겨짐	
도배불량, 시트지 벗겨짐, 거울 벗겨짐	
도장불량, 문틀 도장불량	
도장불량	
도장불량	
배관 멀어짐, 배관구멍, 석고구멍	
배관 멀어짐, 배관구멍, 석고구멍	
배관구멍, 석고구멍	
손잡이 불량	
손잡이 불량	
시트지 벗겨짐	
실리콘 구멍	
실리콘 구멍	
실리콘 구멍	
연결고리 불량, 문 불량, 손잡이 불량	
찍힘	
타일파손	

SI모델의 성능을 최적화하기 위한  
하자유형 분류 재구성  
기존의 하자 유형과의 매핑 적용

[그림3-12] COTIS 하자 이미지 자료의 하자유형 분류체계 개선 내용  
(출처: LH 내부자료)

전체 39종의 하자유형에서  
이미지 분류가 가능한 17종의  
유형 분류

QR 하자유형(39종)	
이음불량	
크랙	
면발량	
들뜸	
접착불량	
들뜸	
줄눈불량	
코킹불량	
오염	
틈새	
벗겨짐	
흡입	
파손	
고정불량	
찍힘	
뒹김	
패임	
기타	
훼손	
물음	
수평불량	
변색	
개폐불량	
태움불량	
작동불량	
...	
단차	
벗겨짐	
금감	
깨짐	
쿠킹불량	
하자없음	
부식	
구배불량	

하자분류를 위한 이미지 데이터양  
이 부족한 것과, 이미지 분류가  
불가능한 유형 제외

22종

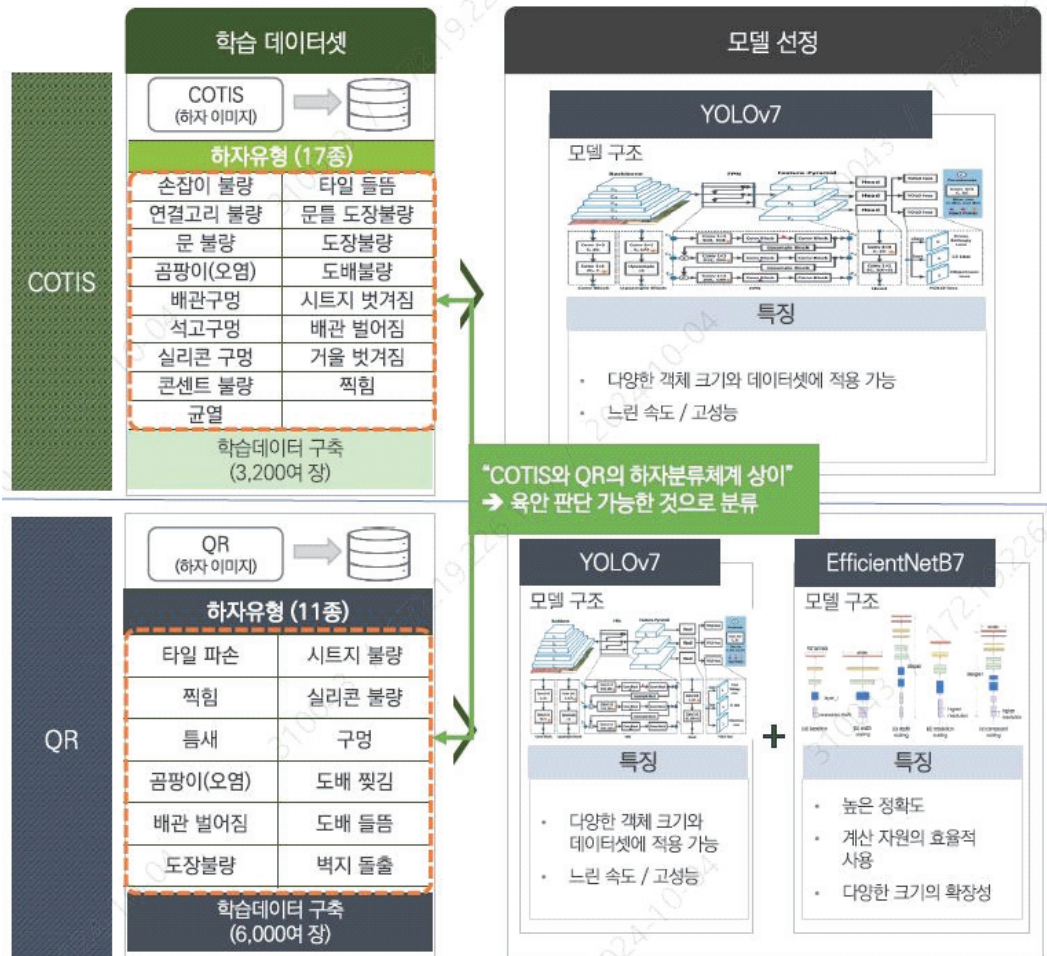
SI 모델 하자유형(11종)	
도배 뒹김	
균열	
도장불량	
도배 들뜸, 시트지 불량, 도배 뒹김	
시트지 불량, 도배 뒹김	
벽지 들뜸	
실리콘 불량	
실리콘 불량	
곰팡이(오염)	
틈새	
시트지 불량, 도배 뒹김	
타일파손, 찍힘	
타일 파손	
배관 멀어짐	
찍힘	
도배 뒹김, 시트지 불량	
찍힘	

신속 이미지 분류 모델의 총 11종  
하자유형 확정  
기존 하자유형과 매핑 적용

[그림3-13] 바로처리(신축아파트)에서 생성된 하자유형 분류개선  
(출처: LH 내부자료)



- 인공지능의 분류모델 중 다양한 객체와 실시간 객체 탐지에 유용한 YOLO(You Only Look Once) 계열의 YOLOv7<sup>7)</sup> 모델을 선정함
- COTIS와 바로처리에 등록된 하자 이미지의 경우 분류체계가 다름



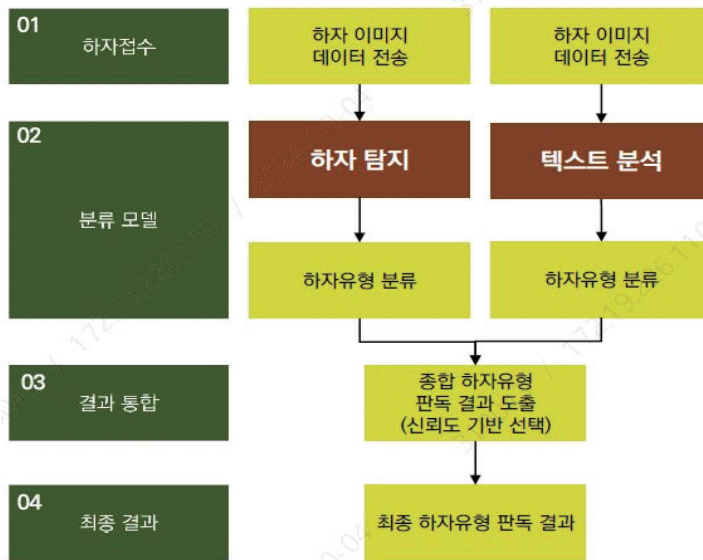
[그림3-14] COTIS 및 바로처리(QR)에 등록된 하자유형에 따른 AI 모델 선정

## 2) 하자 텍스트 분류 모델

- 등록된 이미지로 하자 유형 분류가 어려울 경우 하자 이미지와 함께 등록된 하자관련 내용의 텍스트를 분석하여 하자유형 분류를 보완할 수 있도록 함

7) 이미지를 그리드로 분할하고, 각 그리드 셀에서 객체의 경계상자(bounding box)와 클래스를 예측하는 방식

- 하자유형 분류시 이미지 분류의 오류를 해결하고자 해당 이미지와 연계된 텍스트 분석을 통한 하자유형 분류 보완
- 이미지 데이터에 대한 하자유형 분류 결과 신뢰 값과 텍스트 분석으로 도출된 하자 유형 신뢰값을 비교하여 높은 값의 하자유형을 선택함



[그림3-15] 하자유형 분류 보완을 위한 텍스트 분류 모델 구성

- 텍스트를 분석하기 위해서 딥러닝 모델 중 한국 자연어 처리를 위해 학습된 언어모델 중 KLUE-BERT 모델을 선정함
  - 언어모델의 경우 어떤 데이터를 가지고 학습하느냐에 따라서 특정 분야의 성능 차이가 있어서 하자 민원의 데이터들과 유사한 “청원” 데이터들로 구축된 KLUE-BERT 모델이 적합할 것으로 보여 선정됨
  - 그림에서처럼 COTIS 하자유형에 등록된 텍스트 데이터의 경우 민원빈도가 낮은 유형의 경우 학습데이터를 구축하기 어려움
  - 학습데이터 구축에 있어서 특수문자나 숫자, 불용어(카카오톡, 사진첨부, 기업명 등) 등 분석에 의미 없는 데이터들을 제거하는 사전작업 수행
  - 데이터의 일관성을 높이고 분석 정확도를 위해 단어의 표준화 작업수행 (예: 화재경보기(화재경보기, 경보, 싸이렌))

모델 선정		
텍스트 하자 분류 모델은 한국어 자연어 처리를 위해 구축된 모델 중 '하자 민원'과 유사한 '청원' 데이터를 학습한 KLUE-BERT 모델을 선정하였음		
모델명	개발	학습 데이터
KorBERT	ETRI	<ul style="list-style-type: none"> <li>뉴스/백과사전</li> <li>23GB</li> </ul>
KoBERT	SKT	<ul style="list-style-type: none"> <li>위키피디아</li> <li>50M</li> </ul>
HanBERT	투블록 AI	<ul style="list-style-type: none"> <li>일반/특허문서</li> <li>70GB</li> </ul>
KoreALBERT	삼성 SDS	<ul style="list-style-type: none"> <li>위키피디아/나무위키/뉴스/책</li> <li>줄거리 요약 등</li> <li>43GB</li> </ul>
KLUE-BERT	klue project	<ul style="list-style-type: none"> <li>모두의 말뭉치/CC-100-Kor/나무위키/뉴스/청원 등</li> <li>63GB</li> </ul>
KRBERT	서울대	<ul style="list-style-type: none"> <li>위키피디아/뉴스</li> </ul>
DistillKoBERT	개인(박창원)	<ul style="list-style-type: none"> <li>위키피디아/나무위키/뉴스 등</li> </ul>
KcBERT	개인(이준범)	<ul style="list-style-type: none"> <li>네이버 뉴스의 댓글/대댓글</li> </ul>
KcELECTRA	개인(이준범)	<ul style="list-style-type: none"> <li>네이버 뉴스의 댓글/대댓글</li> </ul>
KoBigBird	개인(박창원)	<ul style="list-style-type: none"> <li>위키피디아/뉴스/모두의 말뭉치/Common Crawl 등</li> </ul>

모델 학습 데이터	
COTIS 하자유형	학습 데이터 수
점검/고제	163,730
작동불량	159,526
고장불량	140,963
누수	111,098
틀림	97,301
흔집	93,653
오염	79,611
개폐불량	77,254
마감불량	68,339
점등불량	53,517
기타	1
금형	8
전기시설불량	8
불질하자	7
콘크리트연배수불량	7
덜개암설치	6
설비개통불량	5
경계재불량	5
가스누출	4
동선추가/연결	3
덜개간격과다	1
수직불량	1
합계	1,403,139

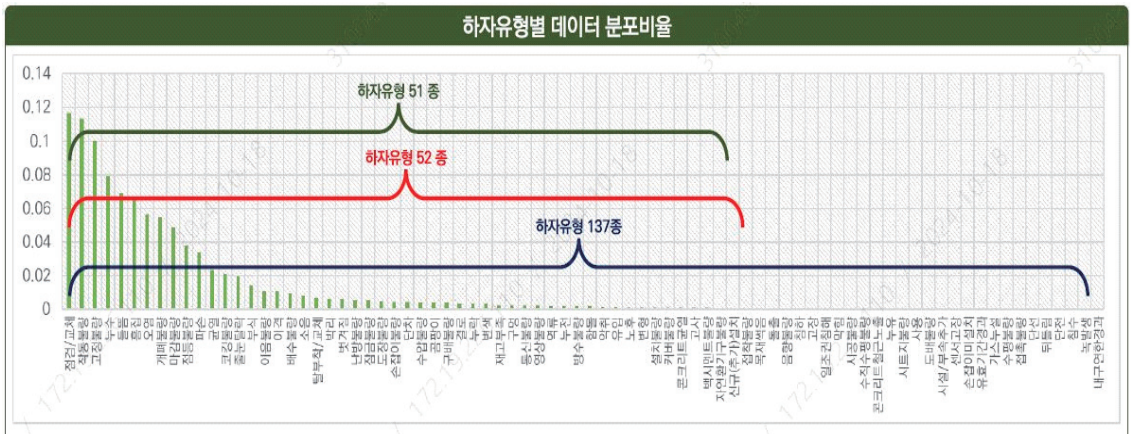
**결과 요약**

- 텍스트 분류 모델은 '하자 내용'에 포함된 텍스트를 활용하여 모델 학습을 진행하였음.
- 하자유형별로 민원 발생 빈도가 낮은 하자유형과 높은 하자유형이 존재하였으며, 학습 데이터의 양에 큰 차이가 있음을 파악하였음

[그림3-16] 텍스트를 통한 하자유형 분류모델 선정

[그림3-17] 하자보수 이미지와 함께 등록된 하자관련 텍스트 학습 데이터

- 특히, 그림에서처럼 텍스트 데이터의 경우 전처리 작업을 통한 137종의 하자유형 중에 51종의 하자유형이 전체의 99.13%를 차지하고 있음
- 51종 제외 나머지 데이터의 양이 없는 관계로 학습 불균형 초래를 해결하기 위해 51종의 하자유형으로 모델을 학습시킴



[그림3-18] 텍스트 데이터의 하자 유형별 데이터 분포비율

### 3) 하자 보수 일정 스케줄링

- 하자보수 관련 업무효율화를 위한 인공지능기술 적용 일환으로 모바일을 통해 접수된 하자내용과 유사한 과거의 하자처리 내용을 기반으로 자동 스케줄링 기능을 제공
  - 하자유지보수업체에서는 보수 전 사진을 촬영한 후 보수 소요일/인력수 등을 산정한 후 가능일정을 고객과 함께 결정함
  - 이에 필요한 데이터들로는 기존에 등록된 하자처리 상세내역서에서 그림에서처럼 보수지시일, 보수완료일자, 하자내용, 작업자, 보수투입일자 등의 자료를 활용함
  - 결과의 신뢰성 및 정확성을 확보하기 위해서는 보수기간, 투입인력수, 보수가능일자 추가 등이 요구됨
  - 현재는 투입인력 수, 업체 작업 일정 등에 대한 데이터가 없는 관계로 이를 가상으로 데이터를 만들어 “하자보수 가능일자” 도출 기능을 구현함
  - 향후, 통합플랫폼에 그러한 정보를 입력하고 데이터베이스화함으로써 그러한 데이터를 기반으로 해당 프로그램의 개선 및 신뢰성을 높일 계획임

활용 데이터	
과거 하자처리 데이터	
• (신)공간	• 보수지시일
• (신)부위자재	• 보수예정일자
• (신)하자유형	• 보수완료일자
• 하자내용	• 보수완료보고일자
보수업체 데이터(가상생성)	
• 보수업체명	• 보수투입일자
• 작업자명	• 보수철수일자

[그림3-19] 하자보수 자동 스케줄링 활용데이터

- 그림에서처럼 하자보수 자동 스케줄링 생성 프로그램의 프로세스는 다음과 같음
  - 하자 민원정보가 접수되었을 때 AI 유사도 기반으로 과거 하자보수내역 중 해당민원과 가장 유사한 정보를 조회함
  - 해당 보수내역을 중심으로 보수기간 및 투입인력 수를 산정
  - 각 보수 업체의 작업일정을 고려하여 보수가능일정을 추천함





[그림3-20] 하자보수 자동 스케줄링 생성 프로그램 프로세스  
(출처: LH 내부자료)

#### 4) 하자 보수 일정 스케줄링

- 하자보수 업체에서 보수작업을 완료했을 때 보수공사에 대한 내역서를 작성시 이를 자동으로 작성하도록 인공지능 기술을 적용
  - 세부내역서 해당공사와 유사한 일위대가를 연동하여 자동 산정하는 기능 제공
  - 그림에서처럼 규격, 단위, 수량, 자재비, 노무비, 경비 등의 데이터들을 활용함
  - 그림에서처럼 세부내역서 정보에서 해당 보수공사와 유사한 데이터들을 조회하여 일위대가를 연동하여 하자접수정보, 하자처리내역, 일위대가 등을 자동으로 생성할 수 있는 기능을 제공
- 현재 내역서는 수기로 기록되었거나 시스템에 등록이 되어있지 않아서 향후, 관련 데이터들의 DB 구축이 요구됨
  - 일위대기 세부내역서는 지역본부별로 상이한 형태이기 때문에 데이터 취합을 위한 구조화가 요구됨(입력 및 출력 표준화 양식 구성 필요)
  - 보수업체의 작업일정 및 보수시작/종료일자, 투입인력 수 등 DB화가 요구됨

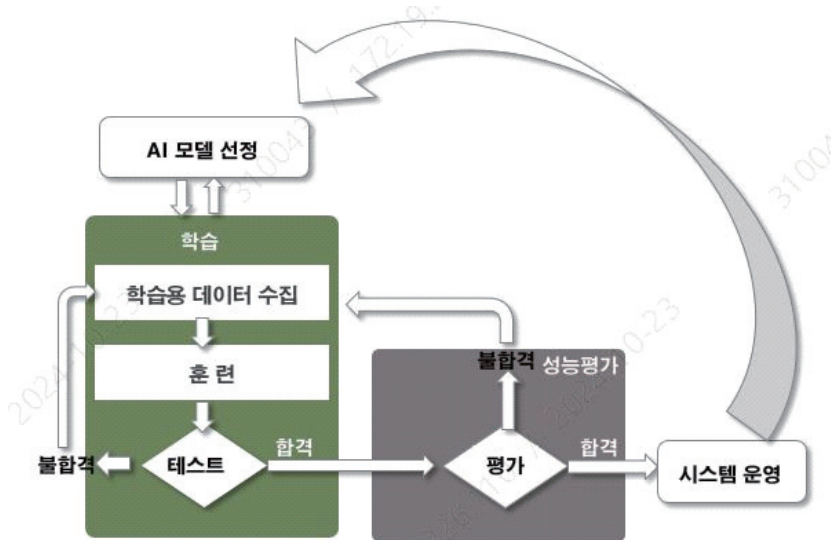


## 2. LH의 하자보수 인공지능 성능지표 수립

### 2.1 LH 하자보수 인공지능 성능지표 평가 체계

#### ■ 성능지표 추진 체계

- 인공지능 기술을 적용한 소프트웨어의 성능을 측정하기 위해서는 기존 소프트웨어의 테스트와는 다르게 귀납적인 방법을 사용함
  - 기존 소프트웨어의 경우 연역적 방법으로 테스트함으로써 테스트에 따라 올바르게 작성되었는지 확인 가능함
  - 반면에 인공지능 모델을 적용했을 경우 테스트의 결과가 일부 거짓으로 도출되었어도 학습이 올바르게 이루어지지 않았다고 할 수 없음



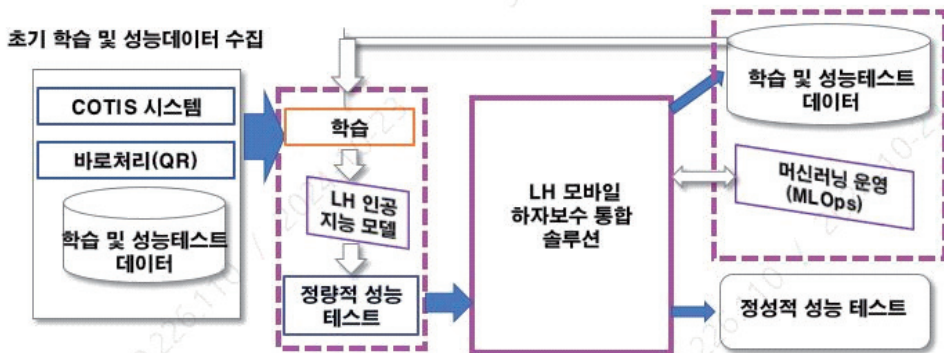
[그림3-23] 인공지능 성능평가 체계도

- 학습용 데이터셋(또는 학습데이터)는 학습에 사용하는 데이터와 정답을 나타내는 데이터의 집합을 의미함
- AI 모델 선정 후 학습용 데이터 셋을 준비하는 데 이때 “훈련용 데이터셋”과 “테스트용 데이터셋”을 구분하여 준비해야 함
- 훈련용데이터셋과 테스트용 데이터셋을 반복하여 최종 학습이 완료된 모델을 가지고 성능평가를 하게 됨
  - 인공지능 성능을 체크 하기에 앞서 품질 좋은 학습 데이터 확보가 중요함

- 또한, 성능체크에 가장 중요한 것은 학습데이터와 별개로 성능체크 데이터를 준비하여야 함
- 성능평가는 개발된 인공지능 모델의 안전성과 신뢰성을 평가하여 실용성을 검토할 수 있음
- 특히, 성능평가 후 결과에 따라 인공지능 모델을 변경하거나 새로운 학습데이터를 추가하여 재학습 후 개선된 모델로 문제를 해결하는 단계가 추가됨
- 인공지능을 적용하여 결과 값을 도출할 때 이에 대한 명확한 참과 거짓을 알 수 없기 때문에 성능지표 역시 분명하게 구분하기 어려움

#### ■ LH 하자보수 통합솔루션 인공지능 성능평가 체계

- LH 하자보수 통합 솔루션에 적용된 인공지능의 성능평가 체계는 그림과 같음
  - 초기 학습 및 테스트에 요구되는 데이터의 경우 기존에 COTIS에 등록되어 있는 하자이미지와 바로처리 시스템의 신규데이터들을 활용함
  - 초기 학습 데이터의 경우 촬영 품질이나, 하자유형의 레이블 작업이 이루어져 있지 않아 많은 데이터 확보가 어려운 실정
  - 그에 따라, 전체 하자유형에 대한 분류가 어렵기 때문에 LH 모바일 하자보수 통합솔루션 시스템 운영에 따른 기능적 개선이 요구됨
  - 초기 모델의 성능평가는 정량적인 평가지표를 활용하여 측정하고 향후 정성적 지표를 포함하여 평가를 함으로써 실용성 및 효과 파악 가능



[그림3-24] LH 하자보수 통합솔루션에 적용된 인공지능 성능평가 체계도

- 인공지능 모델의 신뢰성 및 실용성 평가를 위해서는 기존 성능평가 지표외에 정성적인 성능지표가 요구됨

- 사용자 사용 편의성 지표, 업무 효율성 지표(시간, 비용)
- 반응속도, 안정성, 정확성, 재현성, 신뢰도, 만족도 등의 지표 필요
- LH 하자보수 통합솔루션에 적용된 모델의 경우 지속적인 개선을 위해 고품질의 학습 데이터뿐만 아니라 테스트를 위한 데이터 확보를 할 수 있어야 하며,
- 확보된 데이터로 재학습을 수행하고 성능평가를 한 후 배포하는 자동체계(MLOps<sup>8)</sup>)를 갖출 수 있어야 함
- 특히, 지속적인 재학습에도 성능이 나이지지 않을 경우 새로운 인공지능 모델의 변경도 고려할 수 있음

## 2.2 LH 하자보수 인공지능 성능지표 수립

### ■ LH 인공지능 성능평가 정량적 지표 수립 방향

- 인공지능 분류모델의 경우 실용성 및 신뢰성을 확보하려면
  - 일반적인 기준 정확도가 85% 이상이면 실용성이 있다고 할 수 있음[Devlin, j. et al.(2019)][Vaswani, A., et al(2017)]
  - 질문/응답 시스템에서 85% 이상의 정확도가 사용자의 만족도에 부합
  - 결과의 오류가 심각한 문제를 일으키는 솔루션의 경우 예를 들어 의료분야 같은 경우 위험성이 높아 정확도가 90%이상 요구된다고 할 수 있으며 95~99%정확도가 이상적임[ Ardila, D. et al(2019)][Litman, T.(2020)]
  - 유방암 진단에서 AI기반 도수들은 90% 이상의 정확도를 달성했을 때 의사보조도구로 인정받기 시작
  - 자율주행 차량의 시스템에 적용된 AI의 경우 99.9%이상의 신뢰도가 요구됨
  - 데이터의 불균형이나 사용 목적에 따라 단순한 정확도 외에 F1 점수, 정밀도, 재현율 등을 혼합하여 활용해야 함
  - LH 하자유형 분류에 적용된 인공지능 모델의 경우 정밀도뿐만 아니라 사용자의 선택을 할 수 있는 기능을 만족하기 위해서는 재현율도 고려를 해야 함
- 하자유형을 분류하는 인공지능 모델의 성능지표 중 다음 지표들을 활용

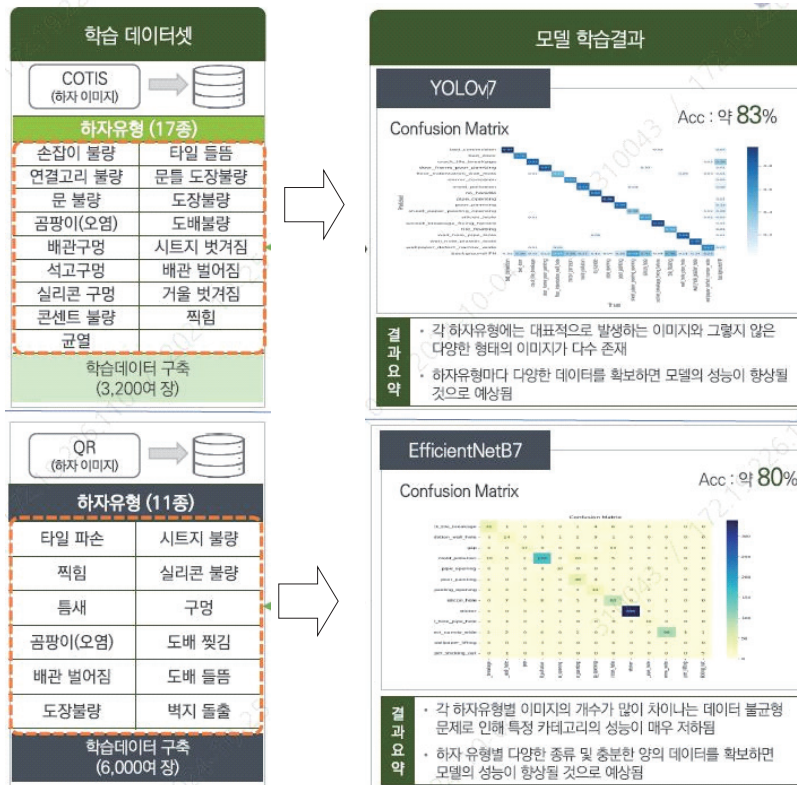
8) Machine Learning Operating System : 데이터를 수집하고 분석하는 단계와 모델을 학습하고 배포하는 단계까지 전체 과정 곧 AI 생애 주기를 총괄하는 시스템을 의미함

- 정밀도(Precision)는 모델이 True로 예측한 것 중 실제 True인 비율을 의미
- 재현율(Recall) 실제 True label 중 모델이 True로 예측한 비율
- mAP(mean Average Precision) 는 각 class별 AP의 평균을 나타낸 값으로 모델의 정확도를 측정하는 지표
- F1-Score는 분류 모델의 정밀도와 재현율의 조화평균을 의미하는 지표(0과 1사이의 값으로 1에 가까울수록 성능이 우수함)

## ■ LH 하자보수 인공지능 성능 평가 및 개선 사항

- 하자유형의 분류의 경우 복잡하고 다양한 클래스로 구성이 되어 높은 성능을 확보하기가 어려움
  - 그림에서처럼 COTIS에 등록된 하자유형 17종에 대해 학습데이터 3,200개를 기반으로 YOLOv7 모델을 학습 후 정확성 지표로 표현된 학습 결과임(정확도 : 83%)
  - 바로처리에 등록된 하자유형의 경우 11종의 유형에 대해 6,000개의 학습데이터를 기반으로 EfficientNetB7 모델을 학습한 후 정확도를 측정한 결과임(정확도:89%)
  - LH 하자보수 유형의 경우 분류체계를 개선한다고 하여도 130 여종의 다양한 하자 유형에 대해 정확도만으로 성능을 체크하기에는 어려움이 있고 각각의 하자유형에 대한 충분한 학습데이터 확보가 요구되고 있는 실정임
  - 특히, 사용자가 자동 분류된 하자유형을 선택하기 위해서는 높은 수준의 재현율이 요구되는 만큼 그에 따른 성능지표가 요구됨





[그림3-25] LH 하자유형 분류에 적용된 인공지능 성능평가 예시

- LH 하자유형 분류의 성능평가를 위해서는 체계적인 접근이 요구되어 짐
  - 효과적인 이미지 분류 모델을 학습하기 위해서는 각각의 클래스에 대한 충분한 학습 데이터가 요구되어지고 있는 반면에 LH 하자보수 유형분류에서는 각각의 클래스에 대한 데이터가 충분히 확보되어지고 있지 않음
  - 하자유형의 분류체계도 표준화 되어 있지 않아서 성능에 영향을 줄 수 있음
  - 테스트 결과 클래스별 모델 성능이 다르게 나타나고 있어서 정확성 지표만으로는 실용성 및 신뢰성을 평가하기가 어려움
  - 정확도(Accuracy)만으로 성능을 평가하기 위해서는 테스트 데이터 구축에 있어서 각 하자 유형별로 적정 비율로 구성할 수 있어야 함

COTIS(구축) 모델의 종합 성능 "83%"				QR(신축) 모델의 종합 성능 "80%"		
72% 하자유형 : 모델 성능 80~90% 이상 - 48% 하자유형 : 모델 성능 90% 이상 - 24% 하자유형 : 모델 성능 80% 이상				58% 하자유형 : 모델 성능 80~90% 이상 - 25% 하자유형 : 모델 성능 80% 이상 - 33% 하자유형 : 모델 성능 70% 이상		
번호	하자유형	모델성능(mAP)	모델 성능(정확도)	번호	하자유형	모델 성능(정확도)
0	전체	0.815	0.835	0	전체	0.807
1	손잡이 불량	0.86	0.8	1	타일 파손,균열	0.65
2	연결고리 불량	0.97	0.93	2	찍힘	0.62
3	문 불량	0.84	0.75	3	틈새	0.53
4	곰팡이	0.85	0.84	4	곰팡이, 오염	0.74
5	배관구멍	0.94	0.94	5	배관 벌어짐	1.00
6	석고구멍	0.95	0.95	6	도장불량	0.74
7	실리콘 구멍	0.55	0.59	7	시트지 불량	0.61
8	콘센트 불량	0.98	0.98	8	실리콘 불량	0.74
9	균열	0.88	0.96	9	배관 구멍	0.83
10	타일 들뜸	0.55	0.83	10	도배 찢김	0.83
11	문틀 도장불량	0.89	0.94	11	벽지 들뜸	0.71
12	도장불량	0.91	0.9			
13	도배불량	0.8	0.77			
14	시트지 벗겨짐	0.58	0.66			
15	배관 벌어짐	0.99	1			
16	거울 벗겨짐	0.78	0.81			
17	찍힘	0.54	0.55			

[그림3-26] LH 인공지능모델의 하자유형별 세부 정확도  
(출처: LH 내부자료)

## ■ LH 인공지능 성능평가 정량적 지표

- 하자유형에 대한 텍스트 분류 모델의 학습결과에서처럼 하자유형간 빈도수가 낮거나 하나의 민원에 대한 복합 하자유형이 있을 경우 성능 평가가 어려움
  - 그림에서처럼 학습데이터가 다른 클래스에 비해 극히 적을 경우 인식이 어려움
  - 하나의 민원에 여러 하자유형이 존재할 경우 인식 오류 발생( 알고리즘 개선 필요)
  - 하자민원 내용과 감독관이 판독한 최종 하자유형과의 불일치에 따른 학습 어려움
- 이에 따라, LH 하자보수 통합관리솔루션에 적용된 인공지능의 정량적 성능평가 지표는 정밀도와 재현율 및 F-Score를 선정함
  - 각 클래스별 정밀도의 경우 85% 이상의 성능을 나타내고, 재현율의 경우는 90% 이상(사용자 선택 용이성)을 성능을 확보할 경우 실용성을 확보할 수 있을 것으로 판단됨
  - 해당 정량적 성능평가 결과뿐만 아니라 정성적 성능평가 결과를 추가하여 실효성 판단을 할 수 있을 것으로 판단됨
  - 특히, 이러한 성능 확보를 위해서는 표준화된 분류체계 및 충분한 학습데이터를 확보할 수 있어야 할 것임



모델 테스트 결과					
번호	하자유형	Precision	Recall	F-Score	Support
1	점검/교체	0.6761	0.6400	0.6576	40,916
2	작동불량	0.6855	0.7686	0.7247	39,405
3	고정불량	0.6457	0.6549	0.6503	35,096
4	누수	0.7294	0.8427	0.7820	28,118
5	들뜸	0.7158	0.7169	0.7164	23,924
64	누유	0.5143	0.7013	0.5934	77
65	시트지불량	0.5000	0.0411	0.0759	73
* 성능저하요인					
67	도배불량	0	0	0	59
70	손잡이미설치	0	0	0	59
71	유효기간경과	0.5000	0.5111	0.5055	45
72	가스누설	0.6667	0.0500	0.0930	40
132	덜개간격과다	0	0	0	1
133	수직불량	0	0	0	1
134	Accuracy			0.6911	350,790
135	Macro avg	0.2798	0.2496	0.2509	350,790
136	Weighted avg	0.6846	0.6911	0.6857	350,790

[그림3-27] LH 하자유형 텍스트 분류 성능테스트 결과

## ■ LH 인공지능 성능평가 정성적 지표

- 인공지능의 정량적 성능지표만으로는 적용된 기술의 실효성 및 신뢰성을 판단하기가 어려움
- 이에 따라, 사용자 설문조사를 통한 편의성, 신속성, 경제성, 안전성, 신뢰성 등 5 개의 인공지능 적용시 정성적 성능 평가지표를 제안함
  - 해당 지표는 5점 척도(매우불만족, 불만족, 보통, 만족, 매우만족)로 하자유형 분류 기능에 대한 사용자 설문을 통해 조사[Reichheld, F. F.(2003)], [Nielsen, J.(1993)], [Sauro, J., & Lewis, J. R.(2016)]
    - 4.0 이상 : 대체로 높은 만족도로 간주되며 실용성이 있다고 평가
    - 3.5~4.0 : 실용성은 있으나 개선이 필요한 수준
    - 3.5 미만 : 실용성이 부족하다고 판단
    - 1.0~2.0 : 명백히 불만족
  - 설문 대상자는 민원 당사자인 주거 사용자(편의성, 신속성)와 업무 담당자(경제성, 안전성, 신뢰성)를 대상으로 항목을 구분하여 실시함
  - 편의성은 해당 기능(하자유형 자동분류)를 활용했을 때의 편리함의 척도를 의미

- 신속성은 해당 기능의 반응 속도를 나타내고 있음
- 경제성은 해당 기능이 없을 경우 인적 시간적 비용 절감을 의미
- 안전성은 해당 기능의 오류 발생 여부를 의미
- 신뢰성은 해당 기능에 대해 신뢰 할 수 있는가를 의미함
- 5점 척도 결과 3.8 이상 나타나면 실용성이 있다라고 판단할 수 있음  
[Parasuraman, A., et (1988)]
  - 공공서비스에서 3.8~4.0 정도가 일반적으로 만족도의 실용성 기준으로 간주함
  - 하자보수 처리에 있어 다양한 사용자를 대상으로 하기 때문에 4.0 이상이면 서비스가 효과적으로 나타남을 의미함
- 사용자 대상의 설문을 통한 정성적 성능지표 값 확보하여 적용된 인공지능 모델에 대한 정성적 평가를 함으로써 정량적 성능평가를 보완할 수 있음
- 사용자 설문조사를 수행은 대상자에 대해 시스템이 자동으로 실행되어 결과 값을 도출할 수 있도록 시스템 업그레이드에 기능 구현이 요구됨

# 제4장 LH의 하자보수 인공지능 성능 평가 운영 및 발전 방안

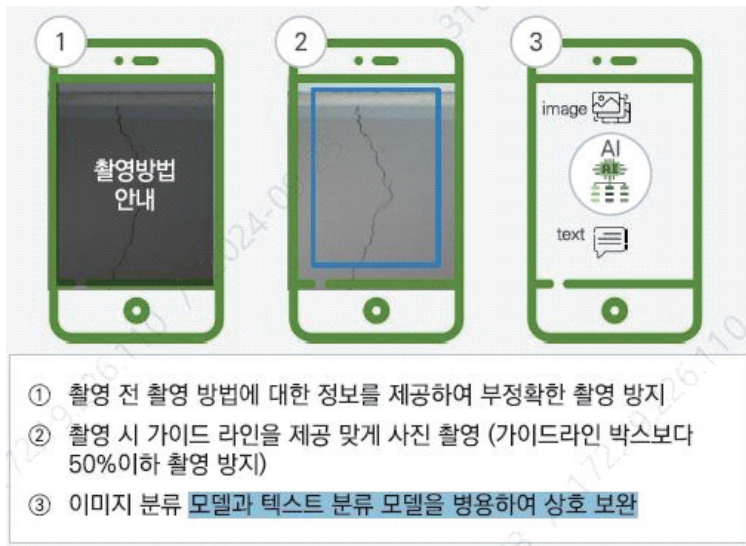
## 1. LH 하자보수 인공지능 성능평가 운영

### ■ 고품질의 학습데이터, 테스트데이터, 성능평가 데이터 구축 필요

- 기존 COTIS에 구축된 하자 데이터의 경우 학습데이터로 활용하는데 어려움
  - 이미지 크기가 다양하고 복잡하며 손상으로 인한 품질 낮은 이미지 데이터(데이터 가공에 많은 비용이 발생함)로 학습에 어려움 발생
  - 하자과 관련 없는 이미지, 이미지와 다른 하자로 등록 되어 있는 문제점(학습 효율성 저하)
  - 이미지의 원경과 근경이 기준 없이 등록되고, 여러하자가 중복된 이미지 등 기본적인 촬영 기준 없음
  - 하자유형의 원천데이터와 그에 따른 메타데이터(하자정보)의 연결성 부족함
- 고품질 학습 및 성능 평가 데이터 구축
  - 하자유형은 다양하고 복잡할뿐만 아니라 새로 개발된 공법이나 재료들로 인해 새로운 유형의 하자가 발생됨에 따라 개발된 모델의 재학습이 요구됨
  - 정확한 하자유형 분류 및 이미지를 축적시킴으로써 고품질의 학습 데이터 구축이 요구되어짐
  - 학습 데이터와 별개로 반드시 성능평가 데이터는 구분하여 구축한 후 운영할 수 있어야 함
  - 솔루션 운영 후 하자유형 표준분류체계 개선 후 재학습 필요
- LH 모바일 하자보수 통합솔루션을 통하여 이미지 생성 가이드라인 및 사용자 편의성 제공필요
  - 사용자가 하자보수 이미지를 촬영하기에 앞서 해당 이미지의 위치 및 자재를 선정하여 이에 대한 정보를 자동 획득할 수 있어야 함(필요시 공중 정보까지)
  - 촬영하는 방법에 대해 명확하게 가이드라인을 제시
  - 가이드라인 및 기준에 맞게끔 촬영되었을 때 자동유형분류 결과를 보여주고 관련 유

형이 없을 경우 재촬영과 추가 하자유형에 대한 자세한 텍스트 요구(텍스트도 자동으로 제공 필요)

- 사용자가 촬영한 이미지에 대한 자동분류 결과가 85% 이하의 정확율<sup>9)</sup>로 인식이 되고 사용자가 그에 대해 선택확인을 하였을 경우 그 이미지는 시스템에 등록됨과 동시에 해당 하자 분류 유형의 예비 학습데이터로 동시에 등록함
- 85% 이상의 정확율로 분류된 경우 재학습의 필요성이 낮음<sup>10)</sup>에 따라 재학습데이터 구성조건을 85% 이하로 함
- 그림에서처럼 관리소와 LH 지역본부 담당자의 하자유형 현황화(정확한 하자유형 분류)를 마친 후 해당 하자의 경우 실제 분류모델의 학습데이터로 등록함
- 학습데이터로 등록시 10% 정도는 성능평가 데이터로 재분류하여 등록 필요
- 학습데이터와 성능평가 데이터는 분류모델 학습시 재활용됨



[그림4-1] 고품질 학습, 평가, 테스트 데이터 확보 방안

#### ■ 하자보수 데이터 품질관리 방안 마련

- 하자보수 데이터(학습용 데이터)의 계획, 구축, 운영·활용 세가지 생애주기 각 단계에서 품질관리를 수행할 수 있어야 함

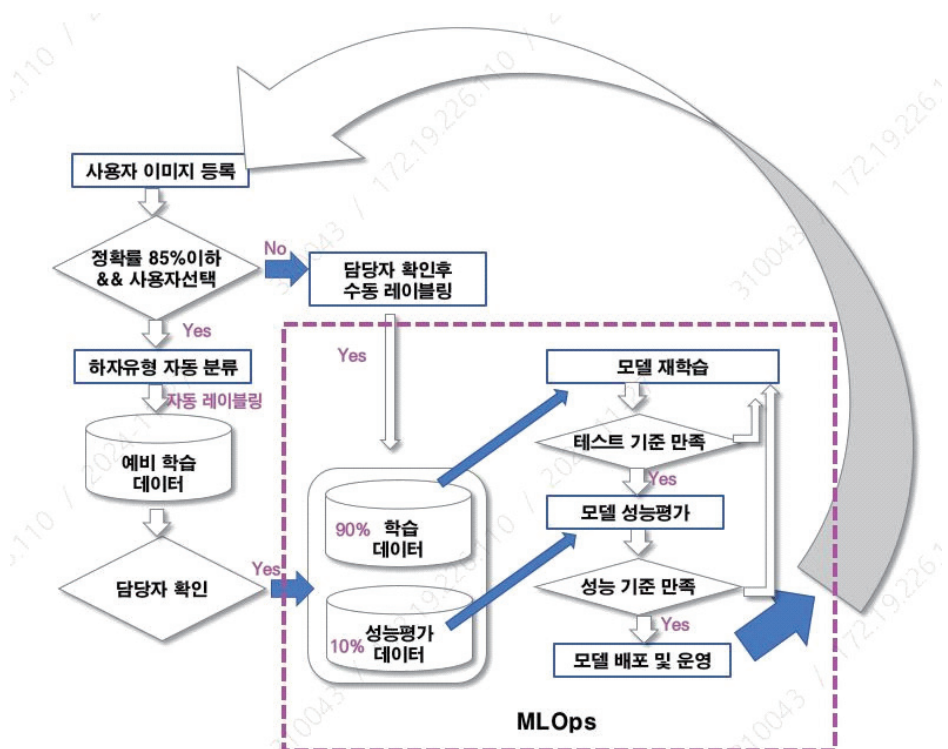
9) LH 하자보수 통합솔루션에서 제공하는 정확율

10) 앞 장에서 언급한 실용성 측면

- 구축데이터 품질관리는 원시데이터, 라벨링 데이터 자체의 품질을 검사하고, 오류를 개선할 수 있어야 함(앞에서 제시한 학습데이터 선별에 따른 검사 필요)
- 구축프로세스 품질관리는 구축과정(수집, 정제, 라벨링 등)에서 발생하는 품질관리를 모니터링하고 문제점을 해결해야 함
- 개방데이터 품질관리는 구축된 데이터를 학습용 데이터셋으로써 지속적인 관리를 수행하고 개방하게 되는 학습데이터에 대한 품질 이슈가 발생시 해결해야 함

#### ■ 학습 및 성능평가 운영체계의 자동화 방안

- MLOps를 통하여 자동으로 학습데이터와 평가데이터를 구축하고 그에 따른 재학습과 배포단계를 자동으로 수행함
  - 6개월 단위의 재학습을 통한 분류모델 개선과 이를 배포하여 성능 개선이 요구됨
- 1년의 운영 및 성능평가를 통하여 분류모델의 개선이 이루어지지 않을 경우, 새로운 학습모델 및 알고리즘 적용을 검토해야 함



[그림4-2] NH 하자보수 인공지능 학습 및 성능평가 운영 체계

## 2. LH의 하자보수 인공지능 발전 방안

### ■ 데이터 기반 의사결정 지원

- 하자보수 데이터의 축적에 따른 분석을 통해 사전 서비스 수행 및 데이터 기반 의사결정을 지원함
  - 기존의 경우 정확한 하자유형의 분류에 따른 데이터가 축적되지 않아서 그에 따른 분석이나 예측 서비스 등이 이루어지지 않음
  - 정확하고 신뢰성 높은 데이터들이 실시간으로 축적을 하고 이러한 데이터들을 분석함으로써 하자유형 예측 및 사전 서비스 등을 수행할 수 있음
  - 하자발생 사전예측에 따른 하자보수 서비스에 대한 유용 자원들을 효과적으로 배치함으로써 시간 및 비용 절감을 할 수 있음
  - 하자보수 비용 분석에 따른 장기수선관리비 산정 등 신뢰성 높은 결과들을 도출하여 활용 가능함

데이터 기반 의사결정  
현안의 체계적 해결



[그림4-3] 데이터 기반의 의사결정 지원

- LH 콜센터와 연계하여 주택 및 시공하자 관련 하자통계 내용들을 활용한 대시보드 적용 및 CEO 의사결정 지원(최근 5년간 하자유형 시계열 분석)
  - ① 고객 콜센터의 최근 유형별 접수 민원을 그래프로 표현
  - ② 고객 콜 센터 주요 내용 표시
  - ③ 주택 및 시공관련 하자 정보 → 기간 및 유형별로 그래프로 표시 → 관련부서 CEO 지시사항 등에 활용



[그림4-4] 데이터 기반의 하자보수 관련 의사결정 지원(참고문헌: 양동석(2024))

## ■ 사용자 인터페이스의 고도화

- 사용자 입장에서 직관적으로 하자보수 요청에 따른 진행 상황을 파악할 수 있도록 사용자 인터페이스 고도화 필요
  - 모바일 환경에서 사용자는 사용이 편리하고 직관적인 사용자 인터페이스를 선호하  
니만큼 사용 대상을 구분하여 메뉴 및 인터페이스를 구분하는 것이 요구됨
  - 특히, 하자보수 앱의 경우 자주 찾는 프로그램이 아니다 보니 사용자가 요구한 하자  
건에 대한 직관적이고 직접적인 정보 제공이 요구됨
  - 진행상황의 경우 단계별 처리 기간 대비 진행 상태를 일목요연하게 볼 수 있도록 정  
보 제공이 요구됨
  - 하자 이미지를 등록시 유사한 하자 유형의 서비스 처리 내용(유사사례, 사전방지 대  
책, 사후처리 단계 등)들을 확인할 수 있도록 게시판 정보 제공 필요함
- 음성인식을 통한 자동 텍스트 입력 및 AI(LLM 모델)를 통한 하자유형 자동입력이 가  
능하도록 기능 제공
  - AI 챗봇을 이용하여 사용자가 음성으로 하자 내용을 입력하면 자동으로 텍스트 저  
노한 및 하자유형이 입력되는 기능 제공 필요
  - 하자내용에 대한 입력과정을 데이터베이스화하여 향후 관련 내용을 학습하여 효율  
적인 자동 입력기능 제공이 요구됨

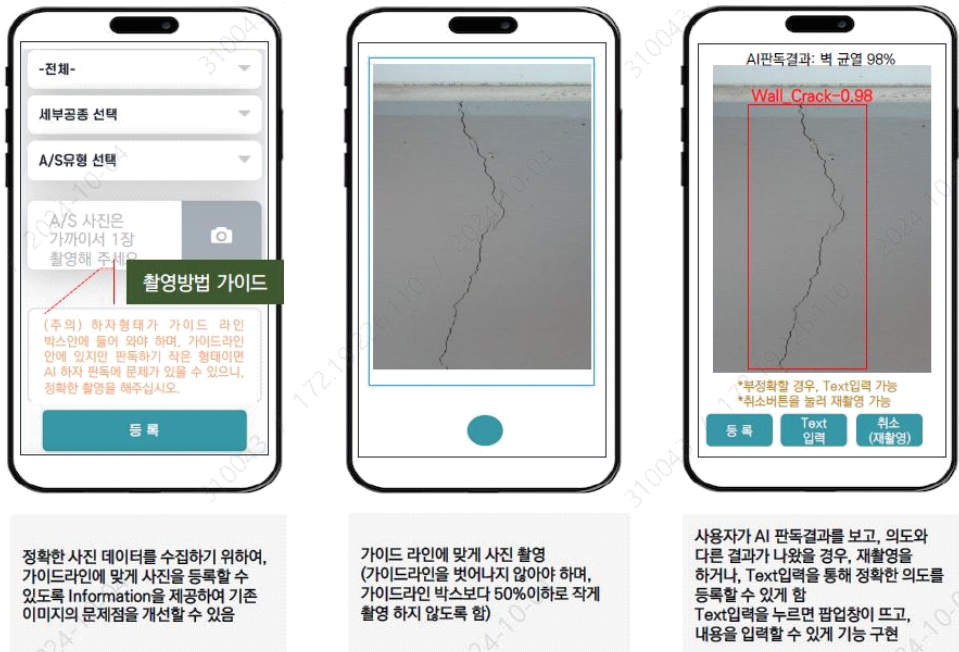


[그림4-5] LH 모바일 하자보수 솔루션 화면

#### ■ 사용자 입력 오류 최소화 방안 마련 필요

- 바로처리(QR코드) 프로그램에서 입력 받아 등록된 데이터들의 경우 특징을 명확히 나타낼 수 있게끔 이미지를 등록하는데 어려움이 있는 것으로 나타남
- 특히, 신축의 경우 준공된지 1년 이내의 하자유형이 많은 관계로 하자의 분류가 육안으로 판별하기 어려움
- 통합솔루션에서 제시한 사용자들이 이미지를 등록할 때에 좀더 세밀하고 명확한 이미지를 등록함으로써 유형 자동분류의 신뢰성을 높일 필요가 있음
  - 사진 등록시 표준화된 가이드를 제시하고 일정 수준에 도달하지 않을 시 재촬영을 할 수 있도록 함
  - 촬영결과에 따른 자동 하자유형 분류 결과를 보여주고 그에 따른 유형이 없을 경우 재촬영 및 텍스트 입력 요청
  - 사용자 등록 이미지 하자유형과 담당자 검토내용이 일치할 경우 학습데이터 및 평가 데이터로 9대1로 선별해서 등록





[그림4-6] 이미지 촬영시 촬영 가이드 제시 및 재촬영 요구의 예

#### ■ 3D 설계 도면 사용 고도화

- 현재 일부 단지의 3D 설계도가 제공되어 사용자가 하자보수 유형을 선택하기에 앞서 관련 위치나 공종, 부의자재 등의 부분적인 선정에 도움을 주고 있음
- 향후, 전체적인 3D 설계도면 제공 및 관련 부의자재 등 세밀한 부분 사용자가 선택할 수 있도록 지원이 요구됨
- 사용자는 해당 위치의 부의자재 클릭만으로 자동으로 관련 정보 등이 등록되고 해당 부분만 촬영하여 등록하는 편의성 제공 필요



3D 적용 화면 예시

[그림4-7] LH 하자보수 모바일 통합솔루션에서 제공중인 3D 적용 기능

#### ■ 하자보수 AI 기능의 설계 및 시공 단계에서의 상호 보완 적용이 요구됨

- 하자보수에 등록된 자료들에 대한 분석으로 반복 하자 자재에 대해 시공이나 설계단계에서 검토할 수 있음
  - 현재 등록된 자재에 대한 정보가 없음에 따라 하자건에 대한 관련 자재의 분석이 이루어지고 있지 않음
  - 3D 도면 등록시 관련 부재에 대한 정보도 등록을 하고 그에 따른 하자정보를 함께 등록하여 분석할 수 있어야 함
  - 분양자료(주택전시관 또는 안내책자)와 다른 시공 결과물에 대한 차이에 따른 하자 분쟁의 잠재적 요인을 사전 방지
- 시공시 반복되는 하자 발생 부위에 대한 이미지와 하자발생 이미지를 연계하여 실시간으로 시공 품질을 모니터링하고 하자를 사전 발견 조치할 수 있는 서비스 구현이 요구됨

- 하자 분석에 있어 시공 자체의 문제점(외국인 기술자의 숙련도 부족 등)을 사전에 미리 방지할 수 있도록 관련 교육을 시행할 수 있는 기반 마련 필요

## ■ 하자 유형분류의 정확성 및 신뢰성 고도화 방안

- 공종별, 부위자재별, 공간별 분류모델 활용
  - 다양한 하자유형에 대해 공종별로 또는 부위자재별로 구분하여 하자유형을 축소한 후 그에 대한 분류를 할 수 있는 인공지능모델로 학습을 진행한 후 각 모델들을 병합하여 운영하는 방안 검토 필요
  - 표준화한 하자유형 분류체계 개발 필요
- 고품질 학습 및 성능 평가 데이터 구축
  - 하자유형은 다양하고 복잡할뿐만 아니라 새로 개발된 공법이나 재료들로 인해 새로운 유형의 하자가 발생됨에 따라 개발된 모델의 재학습이 요구됨
  - 정확한 하자유형 분류 및 이미지를 축적시킴으로써 고품질의 학습 데이터 구축이 요구되어짐
  - 학습 데이터와 별개로 반드시 성능평가 데이터는 구분하여 구축한 후 운영할 수 있어야만 함
- 텍스트 분류모델의 경우 공간, 부위자재 등 사용자가 선택했던 정보를 함께 학습시킴으로써 정확도 및 신뢰성 확보 필요
- 인공지능의 개발된 모델은 실제 학습된 구성 내용을 인간이 이해하기 어렵고, 오류가 발견되었을 때 그것의 수정이 용이하지 않음
- 음성인식을 통한 사용자와의 대화 내용에 따른 하자유형 선택의 과정에서 나온 “대화 내용(질의 및 답변)”을 DB화하여 향후 AI 모델 학습데이터로 활용 필요



# 제5장 결론

## 1. 본 연구의 성과

### ■ 인공지능의 정의 및 성능지표

- 인공지능은 인간처럼 학습하고 문제를 해결하려는 방식인데 최근에 적용 중인 인공지능 90% 이상이 딥러닝 기술이며 본 연구에서의 인공지능은 딥러닝 기술을 의미함
  - 딥러닝은 인간의 뇌를 모델로 한 인공신경망과 많은 수의 데이터들을 가지고 학습하여 그에 따른 결과를 도출하는 기술이라 할 수 있음
- 인공지능의 주요 성능평가 지표들 중 본 연구에서 다루고자 하는 분류모델의 주요 지표로는 정확도, 정밀도, 재현율, F1-Score 등이 있음
- 인공지능 성능지표 관련 주요 연구문헌들을 살펴보고 본 연구의 목표인 LH 하자보수 통합 솔루션에 적용된 인공지능 기술의 관리 및 평가에 적합한 성능지표 수립 방향 및 조건 등을 검토함

### ■ 하자보수관련 인공지능 활용 사례

- 하자보수관련 활용되고 있는 주요 인공지능 모델에 대해 검토함
- 드론이나 고정상도 카메라를 이용하여 건물외벽, 도로, 교량 등의 이미지를 촬영하고 딥러닝 모델을 사용해 균열, 결함 부식등의 하자 자동 탐지 사례
- IoT 센서를 통해 수집한 데이터를 분석하여 장비 상태를 분석하고 고장이 발생하기 전에 유지보수가 필요한 시점을 예측하는 사례
- 하자관련 고객불만에 대한 텍스트 분석을 통해 문제의 유형과 빈도를 신속하고 효율적으로 대응한 사례
- 인공지능을 활용하여 시물레이션을 통한 구조물 하자발생 가능성 예측 및 최적설계 방안 제시한 사례
- 하자보수 관련해서 다양한 분야에 인공지능 기술이 적용되고 있음을 살펴볼 수 있으며 각 사례의 경우 구체적인 성능지표 수립하고 있지 않으나 업무효율성 측면에서 개발 모델의 실효성을 검토하고 있는 실정임

### ■ LH 하자보수 통합관리솔루션의 인공지능 적용 내용 및 주요 특성 검토

- LH의 유지보수 관련 프로그램은 입주품질관리시스템인 “바로처리(QR)”, 업무효율성을 높이고자 개발된 바로픽스, 고객편의성을 위한 카카오톡 AS 등 세 가지 프로그램으로 하자보수 업무를 수행하고 있으나 여러 문제점이 내재 중임
- LH 하자보수 관련 업무 처리시 개선되어야 할 점들로 주거생애 전체 서비스 수행, 편리한 하자유형 등록, 표준화된 하자유형 분류체계, 신속하고 효율적인 업무처리 등을 해결하기 위해 “LH 모바일 하자보수 통합관리솔루션(하자보수 통합관리플랫폼)” 개발 시행
- 본 연구에서는 해당 솔루션의 구성 및 인공지능 적용 내용을 조사하고 하자보수 관련 적용 모델 및 데이터 특성들을 검토함

### ■ LH 하자보수 통합관리솔루션의 인공지능 성능지표 개발

- LH 하자보수 통합솔루션에 적용된 인공지능모델에 대한 성능지표를 수립하기 위해 성능지표 추진체계 및 LH 성능평가 체계를 제시함
- LH 인공지능 성능평가를 위한 정량적 지표(정밀도, 재현율, mAP, F1 score)와 정성적 지표(편의성, 신속성, 경제성, 안전성, 신뢰성)를 수립하고 기준을 제시

### ■ LH 하자보수 통합관리솔루션의 성능평가 운영 및 발전 방안 제시

- 신뢰성 높은 하자 이미지 데이터를 확보함으로써 그에 따른 다양한 데이터기반 의사결정 지원 서비스에 대한 방안 제시
- 사용자 인터페이스 고도화로 편리성 및 직관성을 높일 수 있는 방안 제시
- 사용자 입력 오류를 해결할 수 있는 방안 제시
- 3D 설계 도면 사용 고도화 방안 및 설계 및 시공 단계에서 상호 보완 적용 방안 등을 제시함
- 하자 유형분류의 정확성 및 신뢰성 고도화 방안 제시
- 평가 및 학습운영체계의 자동화 방안 제시

## 2. 연구의 한계 및 제언

- 인공지능의 적용이 활발히 진행되고 있는 가운데 명확한 성능평가지표에 따른 해당 소프트웨어의 실효성 및 신뢰성 등의 판단에 대한 연구가 부족한 실정임
- 본 연구는 LH 하자보수 솔루션에 관련하여 인공지능 기술을 적용한 부분에 대한 성능평가지표를 개발하고 발전 방안을 제시하는 연구로써 향후 수립한 지표들에 대한 성능평가 등의 시스템 검증이 요구됨
- 제안된 성능지표 및 운영방안, 기능 발전 방안 등은 “LH 하자보수 통합관리 플랫폼”에 적용될 것으로 기대됨





## 참고문헌 Reference

- 김광수와2(2021), “지반정보 변화객체 탐지·추출 시스템 개발”, 한국측량학회, 한국측량학회지, 39(6), p515-523, 2021.12.
- 김기연(2021), “AI를 활용한 항공보안 엑스선 검색장비의 위해물품 자동탐지 성능 평가 시험법 개발”, 경기대학교 학위논문(석사), 2021.2.
- 김동명 외(2022), “인공지능과 딥러닝의 기초”, 소음·진동, 제32권 제2호
- 김별(2022), “공동주택 하자 자동분류 및 특성 분석을 위한 HCLA 하자 관리모델”, 한양대학교 대학원 학위논문, 2022.
- 김의중(2023), “알고리즘으로 배우는 인공지능, 머신러닝, 딥러닝”, 미리어드 스페이스, 2023. 8김진혁외1(2012), “수영장 청소 로봇 개발 및 청소성능지표”, 한국로봇학회, 로봇학회논문지, 7(4), p.243-251, 2012.12
- 김현정 외(2023), “인공지능을 위한 머신러닝과 딥러닝 with 파이썬”, (주)도서출판 길벗, 2023
- 딜로이트AI연구소(2023), “인공지능(AI) 활용서: 6대 산업별 활용사례”, Delloitte Insight, 2023. 8
- 라카이푸, 천차우판(2023), “AI 2041”, 한빛비즈, 2023. 1
- 매그너스 에크만(2023), “ 딥러닝 초보자를 위한 엔비디아 가이드북”, 에이콘, 2023
- 박병섭(2023), “인공지능과 딥러닝 개발환경 구축 및 응용”, 복두출판사, 2023.8
- 박향아외3(2021), “머신러닝을 활용한 디지털 트윈 기반 ESS 최적 운영 기술 개발”, 대한전기학회, 대한전기학회학술대회논문집, p766-767, 2021.7.
- 백성복(2023), “AI 표준화 및 시험인증”, TTA Journal vol.205
- 사토나오토외(2023), “인공지능 소프트웨어 품질 보증을 위한 테스트 기법”, 주식회사 제이펍, 2023. 3
- 서옥환외4(2022), “데이터 기반의 충돌 성능 목표 수립 지표 개발”, 한국자동차공학회, 한국자동차공학회 춘계학술대회, p580-584, 2022.6.
- 연합뉴스 기사, 강원 분만 취약지 산모 전용주택 ‘안심스테이’ 첫 삽, 2020.11.20.

오마이건설뉴스, 롯데건설, 3차원 레이저 스캐너 도입...현장 정밀사공 'UP'

오인섭(2018), “교통 빅 데이터를 활용한 고속도로 성능평가 방법론 개발”, 아주대학교, 학위논문(박사), 2018.

양동석 외(2024), “LH 업무 효율화를 위한 딥러닝 개발 방향 연구”, LHRI 연구보고서, 2024.

유지석(2021), “Transformer 기반 공공임대 주거하자 공사유형 예측 연구”, 성균관대학교 대학원 학위논문, 2021

이상훈 외(2024), “스마트 AI 기술을 활용한 주택유지보수 공사 통합관리 솔루션 연구”, LHRI 연구보고서, 2024.

이성수(2023), “인공지능 시스템의 성능평가 표준: ISO/IEC TR24029-1”, 한국전기전자학회 논문지, 제27권 제3호, p350-354, 2023.09.

이성주외3(2022), “인공지능 모델 평가 방식에 따른 도금공정 운영 최적화”, 한국산학기술학회 논문지, 23(12), p917-923, 2022.12.

이원진(2022), “인공지능 딥러닝의 역사와 현황, 그리고 미래 방향”, 대한치과의사협회지, 제 60권 제5호, p298~314

인공지능 기사, AI 언어모델 ‘평션 콜(Function Call)’ 성능 평가 위한 데이터셋...카카오 국내 최초 구축, 오픈소스로 공개

천안국(2023), “인공지능 파이썬으로 머신러닝과 딥러닝”, INFINITYBOOKS, 2023

최재진(2022), “실사와 BIM 기반 컴퓨터 이미지의 딥러닝 객체인식 성능 비교 연구”, 연세대학교 대학원 학위논문(석사), 2022

카카오, AI 언어모델 ‘평션콜’ 성능 평가 벤치마크 데이터셋 구축

카카오, AI 언어모델 성능 평가 데이터셋 오픈소스로 공개

카카오, AI 언어모델 성능 평가 데이터셋 구축 및 오픈소스 공개

포스코이앤씨, 건설지식 집약 '퀄리티 AI 시스템' 개발

한국금융신문 기사, “AI야, 프로젝트 분석해줘” 대우건설, 계약문서 분석시스템 ‘바로답 AI’ 개발

한국과학기술정보통신부 외(2022), “인공지능 학습용 데이터 품질관리 가이드라인 v2.0”, 진한엠앤비, 2022.5.

한양대학교법학연구소(2023), “인공지능과 법”, 2023. 2

헤럴드경제 기사, 롯데건설, 청량리4구역에 3차원 레이저 스캐너 도입

Amodei, D., et al.(2016), “Deep Speech 2:End-to-End Speech Recognition in English and Mandarin,” International Conference on Machine Learning(ICML)

Ardila, D. et al.(2019), “End-to-end Lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature medicine, 25(6), 954-961, 2019.

BBS 기사, 대우건설, AI 기반 계약문서 분석시스템 ‘바로답 AI’ 개발,

CNB뉴스 기사, 대우건설, ‘바로답 AI’ • 공개·방대한 계약문서 실시간 분석,  
2024.11.01.

Devlin, J., et al.(2019),“ BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL-HLT, 2019.

ISO/IEC TR 29119-11:2020,“Software and Systems Engineering-Software Testing - Part 11:Guidelined on the Testing of I-Based Systems”,  
<https://www.iso.org/standard/79016.html>

Kevin P. Murphy(2022),“Machine Learning: An Introduction”, Massachusetts institute of Technology, 2022.

Reichheld, F. F.(2003),“The one number you need to grow,” Harvard Business Review, 2003.

Parasuraman, A., Zeithaml, VI A., & Berry, L. L.(1988),“ SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality”, Journal of Retailing, 1988.

VasWani, A., et al.(2017),“Attention is all you need”, Advances in neural information systems(NeurlPS), 2017.