

## A CNN-based Water Pipe Leakage Classification Model Using Frequency-domain Feature Learning

Chanhee Kwak\*, Seong-Min Lee\*\*

\*Associate Professor, Dept. of AI Convergence Engineering, Kangnam University, Yongin, Korea

\*\*Student, Dept. of Industry Data-science, Kangnam University, Yongin, Korea

### [Abstract]

Leakage from aging infrastructure causes significant water losses and economic costs. Conventional acoustic detection methods are limited by their dependence on expert experience and susceptibility to external noise. While sensor-based data and machine learning methods have been introduced to address these issues, existing models treating frequency spectra as independent features suffer from class imbalance and limited capacity to capture global patterns, resulting in poor generalization. In this study, we design a hybrid architecture that combines 1D-CNN and 2D-CNN and validate its performance using the AI-Hub leak detection dataset. The proposed model leverages 1D-CNN to extract local patterns along the frequency axis and 2D-CNN to capture inter-channel interactions, thereby overcoming the shortcomings of previous approaches. Experimental results demonstrate that the hybrid architecture achieves an accuracy of 0.9587, a Macro F1 score of 0.9487, and an MCC of 0.9456, confirming its potential applicability to a wide range of frequency-based sensor data analysis tasks.

▶ **Key words:** Water pipeline, Leak detection, Frequency spectrum, CNN, Sensor data

### [요 약]

상수도관의 노후화와 환경적 요인으로 발생하는 누수는 수자원 손실과 경제적 부담을 초래하는 대표적 문제다. 기존의 청음식 탐지 방식은 전문가 경험에 크게 의존하고 외부 소음에 취약해 효율성이 떨어진다. 이를 보완하기 위해 센서 기반 계측 데이터와 기계학습 기법이 활용되고 있으나, 주파수 스펙트럼을 독립적인 특성으로만 처리하는 기존 모델은 클래스 불균형과 전역적 패턴 학습의 한계로 일반화 성능에 제약이 있었다. 본 연구에서는 1D-CNN과 2D-CNN을 결합한 하이브리드 구조를 설계하고, AI-Hub 누수 감지 데이터로 성능을 검증하였다. 제안된 구조는 1D-CNN으로 주파수 축의 국소 패턴을, 2D-CNN으로 채널 간 상호작용을 동시에 학습해 기존 방법의 한계를 보완하였으며, Accuracy 0.9587, Macro F1 score 0.9487, MCC 0.9456를 기록해 다양한 주파수 기반 센서 데이터 분석에도 확장 가능성을 확인하였다.

▶ **주제어:** 상수관로, 누수 탐지, 주파수 스펙트럼, 컨볼루션 신경망, 센서 데이터

- 
- First Author: Chanhee Kwak, Corresponding Author: Chanhee Kwak  
\*Chanhee Kwak (chk@kangnam.ac.kr), Dept. of AI Convergence Engineering, Kangnam University  
\*\*Seong-Min Lee (sbqlsbql1234@naver.com), Dept. of Industry Data-science, Kangnam University
  - Received: 2025. 10. 17, Revised: 2025. 11. 24, Accepted: 2025. 12. 02.

## I. Introduction

최근 상수도 시스템은 고도 경제성장기에 집중적으로 건설된 시설의 노후화와 더불어, 기온 변화, 외부 하중 등 다양한 환경적 요인에 의해 누수 발생 위험에 노출되고 있다[1]. 지하에 매장된 상수도관은 다양한 요인에 의해 손상될 수 있으며, 특히 소규모 누수의 경우는 누수 위치를 발견하기 쉽지 않아 탐지까지의 시간이 지체되고 장기간 많은 수자원 손실을 발생시킨다[2]. 이러한 누수는 단순히 수자원 손실에 그치지 않는다. 장기간 방치될 경우 수도사업자와 소비자 모두의 경제적 부담이 늘어나고, 공급 가능한 수자원을 비효율적으로 소모하게 되어 환경적 피해를 초래한다[3].

상수도관 누수를 탐지하는 대표적인 방법은 청음식 누수 탐지 방식으로 센서를 통해 파이프에서 발생하는 누수음을 탐지하여 배관에서 누수 현상이 발생하였는지를 판단한다. 하지만 이 방식은 지하에 매설된 상수도관을 현장 전문가가 장비를 활용하여 직접 탐사하기에, 작업자의 탐사 능력에 의존하고 누수음 탐지 중 외부 소음의 간섭으로 인한 정확한 판단 제한, 그리고 누수가 발생한 정확한 위치 파악에 어려움이 발생하여 시간적·경제적 비용이 과도하게 소모되는 한계가 있다[3].

이러한 한계점을 극복하기 위해, 최근에는 배관 굴착 없이 관로에서 발생하는 진동, 음향, 압력 신호를 측정할 수 있는 비파괴 센서 기반 자동화 탐지 기법이 주목받고 있다. 이때 센서를 통해 수집된 신호는 단순 계측에서 나아가 Fast Fourier Transform(FFT)을 통해 주파수 영역으로 변환되어 이를 이용한 기계학습 기반 누수 탐지 연구가 새로운 대안으로 제안되고 있다[2]. 센서 데이터 기반 누수 탐지 기법은 시간과 비용 절감 효과를 바탕으로, 기존 전문가 중심의 탐지를 데이터 기반 누수 탐지 프로세스로 대체하는 흐름을 형성하고 있다[4].

하지만 머신러닝 기반 모델들은 FFT된 주파수 스펙트럼의 bin을 개별 feature로만 취급하여 이웃 bin 간 국소적 패턴이나 추세를 반영하지 못하고 스펙트럼 전역에 걸쳐 미약하게 분포하는 소규모 누수 신호의 연속적 변화 식별에 한계가 있다. 또한 머신러닝 모델 중 가장 많이 사용되는 트리 기반 모델은 분할에 따라 특정 feature에 과도하게 의존하기 때문에, 주파수 스펙트럼의 전역적 분포 특성을 충분히 반영하지 못하고 일부 변수에 편향된 의사결정을 내리는 문제가 있다[5].

본 연구는 위와 같은 머신러닝 기반 누수 탐지 방법의 한계를 극복하고자 CNN을 이용한 누수 탐지 모델을 제안

한다. 제안하는 모델은 1D-CNN과 2D-CNN을 단계적으로 활용하는 구조로, 1D-CNN을 통해 주파수 스펙트럼의 연속적 변화와 국소적 패턴을 추출하고 2D-CNN을 이용하여 보다 넓은 주파수 구간에서 나타나는 특성을 학습한다. 제안하는 모델 구조는 머신러닝 기반 모델이나 단독으로 사용되는 CNN 모델에서는 놓치기 쉬운 주파수 스펙트럼의 미세한 변화와 구간 간 상관 패턴을 동시에 반영할 수 있다. 이러한 접근은 기존 머신러닝 기반 방식이 가진 표현력의 한계를 넘어, 주파수 영역 특성을 정밀하게 학습하는 새로운 탐지 방법론을 제시한다.

## II. Preliminaries

### 1. Traditional approaches

Choi 외는 누수 탐지를 위해 남양주시 상수도관에 하이드로폰 센서 두 개를 설치하여 심야 시간대에 두 센서로부터 누수음을 동시에 계측하였다[6]. 계측된 신호는 GPS 1PPS 이용해 시간을 동기화 하였고 두 센서의 신호를 교차상관법으로 분석하였다. 실험 결과 약 100m 구간에서 분당 3L 규모의 누수를 성공적으로 탐지하였고 교차상관법 기반 추정위치가 실제 누수 지점과 일치함을 확인하여 제안된 시스템이 실시간 원격 누수 탐지에 활용 가능성을 입증하였다. 다만, 해당 실험은 실제 물 사용량이 거의 없는 심야 시간대 및 약 100m 구간의 관로에서만 수행되어, 장거리 관로나 대규모 상수도망 전체에 대한 적용성에는 한계를 보였다.

Kim 외는 기존 상수관망의 유량 자료만으로는 누수 여부를 정확히 판별하기 어렵다는 점을 지적하여 이를 개선하기 위해 실제 정음시의 상수관망의 DMA 블록 유량 자료와 모의 유량 데이터를 사용하여 가변적 유량측정 기반 적응 칼만필터 알고리즘을 제안하였다[7]. 실험은 적응 칼만필터를 이용하여 기존 유량 시계열 데이터를 기반으로 예측 유량값을 생성하고 실제 계측된 유량값과 잔차의 크기를 계산하여 이상 유량을 탐지하였으며, 또한 잔차 크기에 따라 샘플링 간격을 자동으로 조정하여, 잔차가 큰 경우에는 짧은 주기로 계측해 이상 유량을 정밀하게 감시하고 잔차가 작은 경우에는 긴 주기로 계측함으로써 불필요한 데이터 축적을 줄였다. 실제 정음시 MG4·MG8 블록에 적용하여 실험해 본 결과 각각 30건과 16건의 이상 유량 구간을 검출하였으며, 전체 데이터 사용량을 기존 대비 85~92% 절감하는 결과를 보여주었다. 다만 제안된 기법은 잔차에 기반하여 이상 유량을 탐지하였으나, 실제 누수

와 특정 시간대의 비정상적 수요 변동이 모두 잔차 신호로 유사하게 표현되어 두 현상을 명확히 구별하기 어렵다는 한계가 있다. 이에 따라 계절·시간대별 수요 패턴을 반영한 보완 연구가 필요함을 지적하였다.

## 2. Machine learning based approaches

앞서 살펴본 전통적인 누수 탐지 기법은 외부 소음에 취약하고 누수와 수요 변동을 구분하기 어려우며, 장거리 관로 적용성에도 한계가 있었다. 이러한 문제를 극복하기 위해 최근에는 계측된 신호를 데이터 기반으로 해석하고 자동 분류할 수 있는 머신러닝 기법이 도입되었다.

Choi 외는 기존 방식은 전문가 의존성과 소음 간섭으로 한계가 있으므로, 센서 기반 자동 분류 모델 개발 필요성을 제기하였다[4]. AI-Hub에서 제공하는 상수관로 누수 감지 데이터 중 FFT를 이용해 변환된 주파수 영역의 데이터만 활용해 RBF커널 기반 SVM 누수 분류 모델을 제안하였으며, 실험 결과 Accuracy 0.8478, F1 score 0.8099, Matthews Correlation Coefficient(MCC) 0.7986을 달성하였다. 다만 저자들은 Accuracy가 상대적으로 높게 나타난 원인을 데이터셋 내 클래스 불균형에 의해 과대평가된 결과라고 지적하였다. 즉, 다수 클래스에 대한 편향으로 Accuracy는 높게 산출될 수 있으나, 소수 클래스에서의 분류 성능 저하는 F1 score와 MCC에서 반영되어 상대적으로 낮게 나타나는 한계가 확인되었다.

Lee 외는 AI-Hub 상수관로 누수 감지 데이터를 활용하여 Logistic Regression, KNN, Random Forest, MLP 등 다양한 머신러닝 알고리즘의 성능을 비교하였다[8]. FFT로 변환된 주파수 데이터를 입력으로 정상, 옥외 누수, 옥내 누수 세 가지의 클래스로 분류하였으며, Site, Sensor ID, Leak date 등 모델 성능을 왜곡시킬 수 있는 변수들은 제외하였다. 실험 결과 정규화 미적용 KNN 모델이 가장 우수하여 Macro Accuracy 93.56%, Macro F1 score 93.14%의 성능을 보였다. 저자들은 과거 연구에서 보고된 99% 이상의 과도하게 높은 정확도가 Sensor ID와 Leak date와 같은 변수들을 포함했기 때문임을 지적하며, 데이터셋 변수 특성에 대한 면밀한 검토 필요성을 강조하였다.

Seo 외는 기존 하드웨어 기반 기법은 정밀하지만 비용과 시간이 많이 소요되어 대규모 상수관망에는 한계가 존재하여 머신러닝과 빅데이터 기반의 저비용·고효율 소프트웨어적 접근을 통한 누수 감지가 필요하다고 주장하였다[9]. AI-HUB의 상수관로 누수 감지 데이터셋을 사용하여 옥외 누수, 옥내 누수, 기기계·전기음, 환경음, 정상음 5가지의 레이블로 분류 하였으며, 입력 특징으로는 누수 확률

과 누수 레벨, 0~5,120Hz 범위에서 추출한 513개의 주파수 데이터, 그리고 센서가 일정 시간 동안 반복 측정하여 기록한 20개의 감지 정보를 포함해 총 535개의 특성을 사용하였다. 6가지 트리기반 머신러닝 모델인 Random Forest, AdaBoost, GBM, XGBoost, LightGBM, CatBoost의 성능을 비교하였고, 전통적인 Random Forest나 AdaBoost에 비해 최신 부스팅 계열 모델들인 XGBoost, GBM, LightGBM, CatBoost가 좋은 성능을 보여주었다. 그 중에서도 LightGBM 모델은 정확도 0.96, F1 score 0.948을 기록하며 가장 뛰어난 성능을 보여주었다.

다만, 해당 연구의 높은 성능은 모델 성능을 왜곡시킬 수 있는 메타데이터(예: 사이트 번호, 센서 번호 등)를 그대로 활용한 결과일 가능성이 있으며, 이는 실제 환경에서의 일반화 성능을 과대평가하게 만들 수 있다[8].

## 3. Overcoming the limitations of machine learning approaches

기존 머신러닝 기반 누수 탐지 연구들은 일정 수준 이상의 성능을 달성하였으나, 데이터셋의 클래스 불균형에 취약하여 소수 클래스의 신호를 안정적으로 분류하기 어렵다는 한계가 지적된다.

Byun은 분류 모델은 지도학습 모델을 사용하며 정상 조건과 누출 조건 모두에서 대규모 데이터 세트가 필요하다는 점을 언급하며 누수 모니터링과 탐지 문제를 해결하기 위해 공공 데이터 기반 딥러닝 모델을 제작해 기존 머신러닝 기법들과 성능을 비교하였다[10]. 제안한 딥러닝 모델은 512차원 주파수 스펙트럼을 입력받아, 32차원의 Dense layer와 Relu를 교차 적용하는 다층 신경망 구조로 설계되었다. 입력층을 제외한 각 layer에는 과적합 방지를 위해 Dropout을 0.5로 설정하였고, 마지막에는 Softmax를 사용하여 정상, 옥내 누수, 기계·전기음 3가지의 클래스에 대한 출력층을 두었다. 실험 결과, 기존 머신러닝 기법인 SVM과 KNN은 각각 0.8006, 0.9168의 정확도를 보여주었으나 제안된 딥러닝 모델은 0.9384의 높은 정확도를 보여주었다. 저자는 제안한 모델이 경량 구조를 사용했기 때문에 클라우드뿐만 아니라 IoT 기기나 스마트폰 수준의 하드웨어에서도 상용화 가능성이 높다고 주장하였다. 또한 빠른 진단과 실시간 감시로 대규모 누출·연쇄 오염사고와 같은 환경 피해를 줄일 수 있는 효과적인 대안인 점을 강조하였다.

Choi 외는 기존의 머신러닝 기반 누수 탐지 기법들이 상관법이나 MFCC와 같은 신호처리 기반 특징 추출에 의존함에 따라, 잡음 환경과 데이터 불균형 상황에서 일반화

성능이 제한적이라는 한계를 지적하였다[11]. 이러한 사유로 이들은 노이즈 환경에서도 안정적으로 동작할 수 있고, 특징 설계 과정이 자동화된 학습 기반 접근법의 필요성을 제기하였다. 실험은 수도관에 설치된 진동 센서로부터 수집한 음향 신호를 FFT 변환하여 2차원 스펙트럼으로 재구성한 뒤 2D-CNN 모델을 적용하였으며, 정상, 실내 누수, 실외 누수, 전기·기계음, 환경음의 다섯 클래스를 대상으로 실험을 수행하였다. 그 결과, CNN 모델은 F1-score 94.82%, MCC 94.47%로 기존 SVM 기반 방식 대비 현저히 높은 성능을 달성함으로써, 딥러닝 기반 접근이 기존 전문가 의존적·특징 설계 기반 방식의 한계를 극복할 수 있음을 입증하였다.

### III. The Proposed Scheme

#### 1. Dataset

본 연구는 AI-HUB에서 공개된 상수관로 누수 감지 데이터를 기반으로 각 샘플은 FFT를 이용하여 변환된 후 10Hz bin으로 분할되어있는 10~5120Hz의 주파수 스펙트럼 512개를 이용하였다. 해당 데이터는 광주광역시 및 전라남도 고흥군 지역의 상수도관에 설치된 누수 감지 센서를 통해 총 70,384 건이 수집되었다. 센서는 관로 내 진동 신호를 측정하고, 취득된 데이터는 현장 누수 탐사 전문가의 판독 결과에 따라 정상(Normal), 옥외누수(Out), 옥내 누수(In), 환경음(Other), 기계·전기음(Noise)의 5개 클래스로 라벨링되었다. 각 클래스 별 개수는 Table 1과 같다.

클래스별 주파수 스펙트럼의 특성을 비교하기 위해 Fig 1과 같이 각 클래스의 주파수 구간별 평균 진폭 값을 산출하여 시각화하였다. 시각화 결과, 저주파 영역에서는 클래스별로 뚜렷한 차이가 관찰되었다. 예를 들어, 환경음과 기계·전기음은 500Hz 이하에서 높은 진폭을 보이며, 정상 데이터는 전 구간에서 타 클래스에 비해 낮은 진폭을 보인다. 또한 옥내 누수와 옥외 누수는 공통적으로 유사한 패턴을 보였으나, 세부적인 피크의 형태와 크기에서 차이를 보였다. 한편, 1,000Hz 이상의 고주파 영역에서는 전체적으로 진폭이 낮아지는 경향을 보여졌으나, 4,000Hz 대역의 주파수 영역에서는 클래스별 미세한 패턴 차이가 여전히 존재하였다.

Table 1. Class-wise Sample Counts in the Dataset

Leak Type	Count
normal	22166
out	19731
in	14932
other	7897
noise	5658

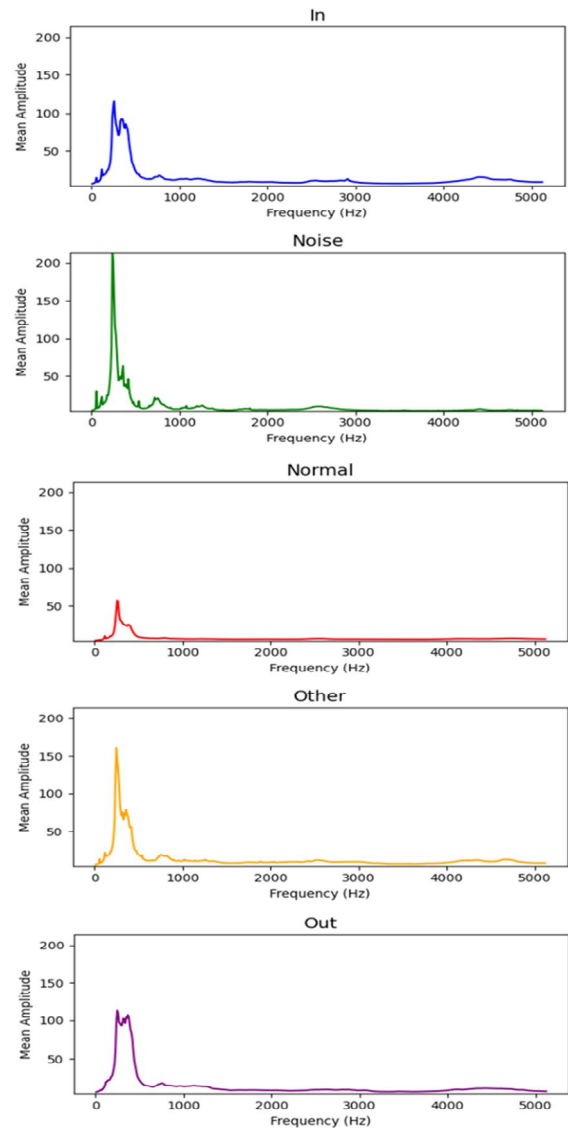


Fig. 1. Comparative Mean Frequency Spectra of Leak Types

#### 2. Model architecture

본 연구에서 제안하는 모델은 1D-CNN과 2D-CNN을 계층적으로 연결한 구조로, Fig. 2. 와 같다. 입력된 512차원 주파수 스펙트럼은 3단계의 1D-CNN과 MaxPooling을 통해 국소 주파수 패턴을 추출한다. 이 과정에서 1D-CNN의 kernel size는 5, padding 값을 2로 설정하여 입력 시

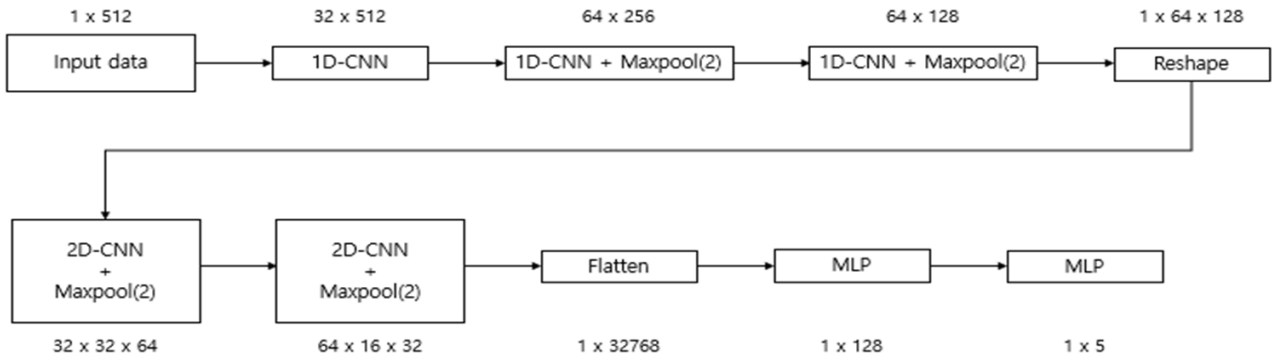


Fig. 2. Model Architecture

퀀스의 길이를 보존하면서 인접 주파수의 상관성을 학습하고 채널 수를 단계적으로 64까지 확장한다. 또한 두 번째와 세 번째 1D-CNN 층에 포함된 MaxPooling은 각 층의 출력 시퀀스 길이를 절반으로 줄여, 512차원의 입력 스펙트럼 벡터를 두 단계에 걸쳐 128차원까지 축소한다. 이를 통해 차원을 축소하면서 각 주파수 구간에서 상대적으로 큰 스펙트럼 성분을 보존할 수 있으며, 작은 진폭의 변동은 억제된다. 이러한 결과는 이후 계층이 불필요한 변동의 영향을 덜 받고, 더 안정적인 입력을 기반으로 주요 스펙트럼 특성을 학습할 수 있도록 한다.

1D-CNN은 시퀀스 축을 따라 국소적인 특징을 추출하는데 효과적이며, 출력 채널을 생성할 때 입력 채널 전체의 convolution 결과를 합산하여 채널 간 정보를 통합한다. 그러나 이러한 결합은 채널 전체를 전역적으로 합산하는 방식으로 이루어지기 때문에, 인접한 채널 간의 상호작용과 시퀀스 구간을 동시에 고려하는 지역적 패턴을 표현하는 데에는 한계가 있다.

이러한 한계를 보완하기 위해 본 연구에선 1D-CNN의 출력을 2차원 텐서로 변환하여 2D-CNN에 입력하도록 설계하였다. 2D-CNN은 커널이 시퀀스 축과 채널 축을 동시에 포함하는 2차원 receptive field에서 합성곱 연산을 수행하므로, 시퀀스와 채널 축 간의 지역적 상호작용뿐만 아니라, 1D 구조에서는 제한적이었던 채널 간 패턴까지 효과적으로 포착할 수 있다[13]. 모델의 자세한 구조는 F1g 2과 같다.

### 3. Experimental setup

본 연구는 Python 3.12와 PyTorch2.9 버전에서 Tesla T4 GPU를 사용하여 진행되었다.

학습에는 CrossEntropy와 Adam Optimizer를 사용하였으며, CrossEntropy에는 Label smoothing 계수를 0.01 추가하여 모델이 특정 클래스에 과도하게 확신하는 현상을 완화하였다. 이는 클래스 불균형과 클래스 간 스펙트럼 유사성으로 인해 발생할 수 있는 불안정한 학습을 완화하고, 보다 안정적인 확률 분포를 학습하도록 하기 위함이다. 초기 학습률은  $5e-4$ , Epoch은 30으로 설정하였다. 평가지표는 Accuracy, Macro F1, MCC를 사용하였다. Accuracy는 모델의 전체적인 예측 정확도를 나타내며, Macro F1은 특정 다수 클래스에 성능이 치우치지 않고 모든 클래스를 균형 있게 예측하는 정도를 평가한다. 추가적으로 사용한 MCC는 혼동행렬의 네 요소(TP, TN, FP, FN)를 모두 고려하여 모델의 전반적 예측 상관성을 측정하는 지표로, 클래스 불균형 상황에서도 보다 안정적인 성능 평가가 가능하다. 이러한 평가 지표들을 기반으로, 제안한 Hybrid CNN의 구조적 효과를 검증하기 위해 제안한 모델 구조에서 2D-CNN을 제거한 1D-CNN 단일 모델과 본 연구처럼 주파수 기반 5개의 클래스로 다중 분류하였던 선행 연구들의 실험 결과를 함께 비교하였다. 실험 결과는 Table 2에 정리하였다.

Table 2. Experimental Results

Metric	Hybrid CNN (ours)	1D CNN (ours)	SVM[4]	2D CNN[11]
Loss	0.1909	0.2189	-	-
Accuracy	0.9587	0.9491	0.8478	0.9580
Macro F1	0.9487	0.9378	0.8099	0.9482
MCC	0.9456	0.9329	0.7986	0.9447

#### 4. Results

제안한 Hybrid CNN은 단일 1D-CNN 및 선행 연구 대비 Accuracy, Macro F1, MCC에서 모두 향상된 성능을 보였으며, 이는 클래스 불균형 상황에서 분류 일관성이 강화된 결과로 해석된다. 2D-CNN 단일 구조와 비교하면, 세 지표 모두 소폭 증가한 수준에 그쳤는데, 이는 해당 데이터셋이 일정 수준 이상의 성능을 비교적 쉽게 달성할 수 있는 데이터 특성을 반영한 결과로 판단된다.

기존 머신러닝 기반 선행 연구들이 주파수 스펙트럼을 개별 특징으로 사용한 반면, 본 연구는 1D-CNN으로 주파수 신호의 특징을 추출·압축한 뒤 이를 2차원 feature map으로 확장하여 채널-시퀀스 간 상관관계를 학습하도록 설계하였다. 아울러, label smoothing 적용은 클래스 불균형으로 인한 과도한 확신을 완화하여, 모델이 보다 균형적인 확률 분포를 학습하도록 유도하였다. 그 결과 제안한 하이브리드 구조는 예측의 일관성과 학습 안정성 측면에서 향상된 성능을 보여주었다.

#### IV. Conclusions

본 연구에서는 AI-HUB의 상수관로 누수 감지 데이터를 이용하여, 10Hz 간격으로 분할된 주파수 스펙트럼 기반 누수 탐지 모델을 제안하였다. 제안된 모델은 1D-CNN과 2D-CNN을 결합한 하이브리드 구조로 구성되며, 총 5개 누수 유형 (정상, 옥외 누수, 옥내 누수, 환경음, 기계·전기음)에 대해 분류 성능을 실험하였다.

연구 결과 제안한 하이브리드 모델은 Accuracy 0.9587, Macro F1 0.9487, MCC 0.9456의 결과를 보여 주었으며 1D CNN만 사용하거나 동일한 데이터셋을 사용한 선행 연구들과 비교했을 때 전반적으로 향상된 성능을 보여주었다.

본 연구는 상수관로 누수 감지라는 구체적 응용 분야에서 출발하였지만, 제안한 Hybrid CNN 구조는 푸리에 변환된 주파수 스펙트럼 입력을 다루는 일반적인 특성을 지니기 때문에, 누수 탐지 문제뿐만 아니라 다양한 주파수 기반 센서 데이터 분석에도 적용 가능할 잠재력을 가진다.

그러나 본 연구에도 몇 가지 한계가 존재한다. 첫째, 실험 데이터가 특정 환경에 국한되어, 실제 현장 환경에서의 잡음, 관 재질 차이, 수압 변동 등 다양한 외란 요인에 대한 일반화 성능은 추가 검증이 필요하다. 둘째, 모델은 스펙트럼 기반 정적 입력만을 활용하였기 때문에, 시간에 따른 연속적 변화나 누수 진행 상태를 반영하지 못한다는 제

약이 있다. 셋째, 제안한 모델과 기존 2D-CNN 간 성능 차이가 크지 않게 나타난 점은, 2D-CNN만으로도 주어진 데이터 분포에서 어떤 분류기도 넘을 수 없는 이론적 최소 오분류율인 베이스 오차에 가까운 수준의 성능에 도달했기 때문으로 해석할 수 있다[14].

향후 연구에서는 시계열 기반 동적 입력이나 멀티 센서 융합을 고려한 확장 연구가 필요하다. 또한, Lime이나 Class Activation Map(CAM) 기법을 적용하여, 모델이 주파수 스펙트럼의 어느 영역에 주목하여 판별을 수행하는지 시각화할 계획이다. 이를 통해 성능 향상뿐만 아니라, 모델이 판단의 근거를 명확히 해석할 수 있는 설명 가능한 인공지능(XAI) 체계로 확장하고자 한다.

#### ACKNOWLEDGEMENT

This research is supported by 2023 Kangnam University Research Grant.

#### REFERENCES

- [1] S. J. Lee, and S. T. Seo, "Fundamental Study on the Aging of Waterworks Pipelines." *Journal of The Korea Society of Environmental Administration*, Vol. 24, No. 1, pp.29-36, June 2017.
- [2] C. W. Lee and D. G. Yoo. "Development of leakage detection model in water distribution networks applying LSTM-based deep learning algorithm" *Journal of The Korea Water Resources Association*, Vol. 54, No. 8, pp. 599-606, Jan 2021. <https://doi.org/10.3741/JKWRA.2021.54.8.599>
- [3] R. Liu, T. Zayed and R. Xiao. "Contrastive learning method for leak detection in water distribution networks." *npj Clean Water*, Vol. 7, No. 1, Art. no. 118, Nov 2024. <https://doi.org/10.1038/s41545-024-00406-6>
- [4] J. G. Choi, and S. B. Im. "Leak Detection and Classification of Water Pipeline based on SVM using Leakage Noise Magnitude Spectrum." *The Journal of Institute of Electronics and Information Engineers*, Vol. 60, No. 2, pp. 6-14, Feb 2023. <https://doi.org/10.5573/ieie.2023.60.2.6>
- [5] Z. Zhou, G. Hooker. "Unbiased Measurement of Feature Importance in Tree-Based Methods" *The Journal of ACM Transactions on Knowledge Discovery from Data*, Vol. 15, No. 26, pp.1-21, Jan 2021. <https://doi.org/10.1145/3429445>
- [6] S. H. Choi and D. S. Eom, "Wireless Water Leak Detection System Using Sensor Networks." *The Journal of Institute of Electronics Engineers of Korea - Signal Processing*, Vol. 48, No. 3, pp.

125-131, May 2011.

- [7] S. W. Kim, A. C. Jeong, Y. T. Lim, W. S. Yu and K. S. Jung "Leakage Detection in Water Distribution Network Using Adaptive Kalman Filter Based on Variable Flow Measurement." The Journal of Korean Society of Hazard Mitigation, Vol. 16, No. 5, pp. 47-53, Jan 2016. <https://doi.org/10.9798/KOSHAM.2016.16.5.47>
- [8] J. B. Lee, J. U. Woo, K. H. Kim, H. J. Kim, and S. H. Cho. "Development for a Water Pipe Leak Detection Model based on Artificial Intelligence Algorithms." in Proceedings of the Symposium of the Korean Institute of Communications and Information Sciences, pp.1159-1199, Nov 2024.
- [9] Y. M. Seo, K. H. Choi, Y. S. Lim, B. J. Lee, and Y. Y. Choi. "Application of Machine Learning Models for Water Pipeline Leakage Detection" The Journal of Crisisonomy, Vol. 19, No. 4, pp. 45-54, Apr 2023. <https://doi.org/10.14251/crisisonomy.2023.19.4.45>
- [10] S. W. Byun "Water Leak Detection using IoT Acoustic Sensors and Deep Learning" The Journal of Korea Academia-Industrial cooperation Society, Vol. 26, No. 4, pp. 392-398, Apr 2025. <https://doi.org/10.5762/KAIS.2025.26.4.392>
- [11] J. G. Choi, and S. B. Im. "Application of CNN Models to Detect and Classify Leakages in Water Pipelines Using Magnitude Spectra of Vibration Sound" Applied Sciences, Vol. 13, No. 5, Art. no. 2845, Feb 2023. <https://doi.org/10.3390/app13052845>
- [12] I. Daniel and A. Cominola, "Estimating irregular water demands with physics-informed machine learning to inform leakage detection", Water Research, Vol. , No. , July 2023, <https://doi.org/10.48550/arXiv.2309.02935>
- [13] S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci, and M. Gabbouj, "1-D Convolutional Neural Networks for Signal Processing Applications" in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8360-8364, May. 2019. <https://doi.org/10.1109/ICASSP.2019.8682194>
- [14] T. Ishida, I. Yamane, N. Charoenphakdee, G. Niu, M. Sugiyama, "Is the Performance of My Deep Network Too Good to Be True? A Direct Approach to Estimating the Bayes Error in Binary Classification" in Proceedings of the The Eleventh International Conference on Learning Representations (ICLR), May 2023, <https://doi.org/10.48550/arXiv.2202.00395>

## Authors



Chanhee Kwak received the B.S. degree in computer science, M.S, and Ph.D. degrees in management engineering from KAIST, Korea. Dr. Kwak joined the faculty of the Department of AI Convergence Engineering

at Kangnam University, Yongin, Korea, in 2020. He is currently an Associate Professor at the same university. His research interests include multi-modal data analysis and explainable AI.



Seong-Min Lee is currently a student for the B.S. degree in Data Science at Kangnam University, Korea. He is interested in machine learning, artificial intelligence, and data analysis.