

Development of Gas Type Identification Deep-learning Model through Multimodal Method

Seo Hee Ahn[†] · Gyeong Yeong Kim^{**} · Dong Ju Kim^{***}

ABSTRACT

Gas leak detection system is a key to minimize the loss of life due to the explosiveness and toxicity of gas. Most of the leak detection systems detect by gas sensors or thermal imaging cameras. To improve the performance of gas leak detection system using single-modal methods, the paper propose multimodal approach to gas sensor data and thermal camera data in developing a gas type identification model. MultimodalGasData, a multimodal open-dataset, is used to compare the performance of the four models developed through multimodal approach to gas sensors and thermal cameras with existing models. As a result, 1D CNN and GasNet models show the highest performance of 96.3% and 96.4%. The performance of the combined early fusion model of 1D CNN and GasNet reached 99.3%, 3.3% higher than the existing model. We hoped that further damage caused by gas leaks can be minimized through the gas leak detection system proposed in the study.

Keywords : AI, Deep Learning, Multimodal, Gas Detection, Gas Identification

멀티모달 방식을 통한 가스 종류 인식 딥러닝 모델 개발

안 서 희[†] · 김 경 영^{**} · 김 동 주^{***}

요 약

가스 누출 감지 시스템은 가스의 폭발성과 독성으로 인한 인명 피해를 최소화할 핵심적인 장치이다. 누출 감지 시스템은 대부분 단일 센서를 활용한 방식으로, 가스 센서나 열화상 카메라를 통한 검출 방식으로 진행되고 있다. 이러한 단일 센서 활용의 가스 누출감지 시스템 성능을 고도화하기 위하여, 본 연구에서는 가스 센서와 열화상 이미지 데이터에 멀티모달형 딥러닝을 적용한 연구를 소개한다. 멀티모달 공인 데이터셋인 MultimodalGasData를 통해 기존 논문과의 성능을 비교하였고, 가스 센서와 열화상 카메라의 단일모달 모델을 기반으로 네 가지 멀티모달 모델을 설계 및 학습하였다. 이를 통해 가스 센서와 열화상 카메라는 각각 1D CNN, GasNet 모델이 96.3%와 96.4%의 가장 높은 성능을 보였다. 앞선 두 단일모달 모델을 기반한 Early Fusion 형식의 멀티모달 모델 성능은 99.3%로 가장 높았으며, 또한 기존 논문의 멀티모달 모델 대비 3.3% 높았다. 본 연구의 높은 신뢰성을 갖춘 가스 누출 감지 시스템을 통해 가스 누출로 인한 추가적인 피해가 최소화되길 기대한다.

키워드 : 인공지능, 딥러닝, 멀티모달, 가스탐지, 가스식별

1. 서 론

가스 누출은 가연성 가스로 인한 폭발이나 독성 가스로 인명 피해가 발생할 위험이 있다. 화학물질안전원의 통계자료에 따르면, 9년간(2014.01~2023.02)의 산업현장 내 화학물

질 사고 757건 중 가스 누출 사고는 601건으로 79.4%의 가장 높은 비율을 차지하였다[1]. 해당 상황으로 인해 가스를 활용하는 산업체와 연구계에서는 가스 안전에 대한 다양한 조치를 추가하였고, 대표적으로 가스 누출 감지 시스템의 도입이 있다. 한국가스안전공사에서 조사한 가스 사고 발생 현황에 따르면(Table 1), 가스 누출 감지 시스템의 도입이 본격화되었던 2018년도 148건에서 2022년도 73건으로 약 51%의 가스 사고가 줄어들었다[2]. 특별히 누출을 통해 이차적으로 발생할 수 있는 사고인 폭발, 화재 등의 수가 많이 줄어들었다. 한편, 가스의 특성으로 인해 누출에 따른 폭발, 화재, 중독, 산소결핍, 파열 등은 심각한 물질, 인적자원의 피해를 초래할 수 있다는 심각성을 고려한다면, 현재 줄어든 가스 사고 건수에 만족할 수 없다. 이를 통해 가스 누출에 대한 조기 감지의 중요성이 높다고 할 수 있고, 이러한 가스 누출 사고에 대비하기

* 이 논문은 2023년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구이며(No. 2022R1A6A1A03052954), 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.RS-2023-00231158, 비전기술을 활용한 편물 검단 및 환편기 예지 보전 원격제어 통합모니터링 플랫폼).

[†] 정 회 원 : 포항공과대학교 인공지능연구원 연구원

^{**} 비 회 원 : 포항공과대학교 인공지능연구원 팀장

^{***} 비 회 원 : 포항공과대학교 인공지능연구원 연구부장

Manuscript Received : August 28, 2023

First Revision : November 17, 2023

Accepted : November 20, 2023

* Corresponding Author : Dong Ju Kim(kkb0320@postech.ac.kr)

Table 1. Gas Incident Status(2018~2022)

	누출	폭발	화재	중독	산소결핍	파열	계
2018	35	39	33	10	0	26	143
2019	29	34	30	8	0	17	118
2020	18	30	26	6	2	16	98
2021	16	17	22	4	1	18	78
2022	14	25	17	6	0	11	73

위해 보다 안전하고 미세한 누출에도 효과적으로 감지할 수 있는 시스템이 필요하다.

가스 누출 감지 시스템을 위한 다양한 연구가 진행되어왔으며, 대표적으로 가스 센서, 열화상 카메라, 적외선 카메라 등을 통한 방법이 등장하였다[3]. 가스 센서는 실내의 가스 농도 값을 토대로 특정 임계치가 넘어갈 때 곧바로 위험 여부를 확인할 수 있다. 열화상/적외선 카메라의 경우 지역적인 파이프라인에서의 가스 누출 시, 누출에 따른 온도 변화와 압력 변화의 특성을 가시화할 수 있다는 점에서 실제적인 누출 확인에 용이하다. 현재 연구되고 상용화된 기술과 제품들은 대부분 가스 내에 한가지 특성으로 누출을 판단하는 단일모달형 연구로 정의할 수 있다. 그러나 단일모달형 연구들은 실제적인 상용화 수준까지 왔지만 기대보다 낮은 가스 인식률과 미검출 사례가 발생한다. 가스 센서의 경우 공기 중 가스 비율이 매우 낮은 경우 식별이 불가하고, 센서가 위치한 부분에 따라 매우 지역적으로 측정한다는 점에서 비효율적이다. 단일 열화상이나 적외선을 통한 광학적 가스 이미징(Optical Gas Imaging, OGI)의 경우에도, 가스의 존재는 식별하지만, 가스 종류, 누출량에 대한 정보는 알기 어렵고 화질이 낮아 오검출할 수 있다. 이러한 한계점은 추후 실제 가스 누출 시에 심각한 문제를 초래할 수 있는 점을 고려할 때 신뢰할 수 있는 정도의 높은 정확도를 갖춘 연구를 요구한다. 따라서 가스 누출 검출 시 여러 단일 모달의 정보를 종합적으로 고려하여 데이터에 대한 분석력이 향상된 누출 감지 시스템을 적용할 필요가 있다.

본 연구에서는 일곱 가지 가스 센서와 열화상 이미지 데이터를 활용하여 가스 누출을 탐지할 수 있는 딥러닝 기반의 멀티모달(Multimodal) 모델을 개발하였다. 본 모델을 개발하고자 (1) 공인 데이터셋인 MultimodalGasData[4]를 데이터셋으로 선정하였고, (2) 열화상 이미지와 가스 센서 데이터를 각각 활용하여 센서 데이터 기반 세 가지 모델과 열화상 이미지 기반 네 가지 모델의 성능을 각각의 데이터별로 비교 및 분석하였다. 이후 (3) 단일 데이터의 성능이 가장 좋았던, 1D-CNN과 CNN을 활용하여 네 가지의 데이터 융합(Fusion) 방식을 적용한 멀티모달 모델의 성능을 비교 및 분석하였다. 이를 통해 Early fusion 중 한 방식으로 구현된 Fusion I 모델에서는 99.21%의 정확도로 기존 단일모달 모델의 성능(96%)보다도 더 높은 성능임을 확인하였다. 본 연구를 통한 멀티모달 데이터 융합 모델이 단일모달 모델보다 더 안정적이고 신뢰할 수 있는 점과 동일한 데이터 셋의 기존 멀티모달 융합 모델[3]보

다 더 높은 성능을 가진 점에서 의의가 있다. 본 연구에서 개발한 기술이 실제 현장에서 유해가스 누출 시 피해를 최소화하는데 적용되길 기대한다.

2. 관련 연구

가스 탐지 시스템을 위한 연구는 크게 두 가지로 나누어져 진행되고 있다. 가스 센서를 활용한 농도 측정 방법과 가스 누출 지점을 파악 후 열화상/적외선 카메라를 활용하여 이미지 및 영상을 활용하는 방법이다. 가스 센서를 활용하는 경우 가스 농도를 수집하여 모니터링하며, 일정 농도 이상일 때 부저를 울리거나 알람을 전송하여 누출에 대비한다[5]. 그 외에 딥러닝 모델을 통해 가스 농도의 추세를 통해 2~5분 후의 예상되는 농도 값에 따라 위험도를 판별하고 해당 장소에 있는 사용자에게 알릴 수 있는 시스템에 대해 개발하는 연구도 있다[6]. 다양한 가스 센서의 연구와 모니터링 시스템의 연구가 진행되고 있으나, 일차적으로 가스 센서를 통한 농도 값의 신뢰성이 검증되지 않는 경우가 많았고, 실제 실험을 할 수 없는 관계로 예측한 가스 값이 나올 거라는 부분이 검증하기 까다로운 문제가 존재한다.

열화상/적외선 활용 가스 탐지는 가스 누출 시점과 지점을 파악하기 위한 연구가 주를 이룬다. 특별히 실험을 통한 누출될 위치에 카메라를 설치하여 누출 전과 후에 대해 비교할 수 있는 데이터를 수집하고, 이에 따라 검출(Detection)이나 구획화(Segmentation), 분류(Classification) 방식 등을 통해 가스 누출을 자동으로 인식하도록 개발한다. Jihao Shi의 연구에서는 탄화수소 누출의 탐지를 목적으로 적외선 카메라를 설치하여 영상을 수집하였다[7]. 가스 누출의 여부를 판단하기 위해 딥러닝 기반의 검출기인 Faster R-CNN, SSD detector를 활용하였고, 다양한 백본망들을 통해 비교분석하여 최종적으로 mAP@0.5 IoU가 0.98의 성능을 보였다. 그러나 탄화수소에 대한 누출감지 연구로 다른 가스의 경우에 동일하게 적용할 수 없는 한계를 지닌다. Rui Xiao는 음향신호를 기반으로 웨이블릿 변환과 SVM(Support Vector Machine)을 활용하여 가스 파이프라인의 누출 감지 방법을 제안한다[8]. 먼저 수집된 음향신호를 웨이블릿 변환으로 전처리한 후 Relif-F 알고리즘과 SVM으로 누출 여부 및 심각도를 분류하며 95.6%의 정확도를 보였다. 해당 모델은 누출 여부와 심각도의 측정이 가능하다는 의의가 있으나, 정확한 누출 위치와 가스 종류를 파악할 수 없는 한계를 지닌다. 이처럼 가스 누출 탐지를 위해 다양한 모델이 연구되고 있다. 하지만 가스 누출 탐지에 멀티모달 모델을 접목한 연구는 찾아본 결과 전무하였다. 이는 다양한 도메인의 데이터를 동일한 환경에서 수집해야 하는 어려움이 있기 때문으로 분석된다.

선례를 찾아보기 힘든 만큼 본 연구에서 활용한 멀티모달 오픈 데이터셋[4]이 연구적인 큰 의의가 있으며, 적용한 데이터셋을 기반으로 가스 누출 탐지를 연구는 다음과 같다.

Narkhede et al.은 가스 센서와 열화상 이미지를 활용하여 단일모달 모델과 멀티모달 모델의 성능을 비교 분석하였다[9]. 4가지 가스 상황을 분류할 수 있는 딥러닝 기반 모델을 만들고자 각 상황에서의 가스 센서값과 열화상 이미지를 수집 및 가공하였다. 가스 센서 데이터와 열화상 이미지를 각각 LSTM 모델과 CNN 모델을 구축하여 학습하였고, 분류 정확도가 각각 83%, 93%다[9]. 이후 개발한 모델을 기반한 두 가지 융합 방식(Early fusion, Late fusion)을 적용하여 멀티모달 모델을 구축하였다. Early fusion 모델의 경우에는 단일모달 데이터들의 특징 벡터들을 융합하여 분류기(Classifier)를 거쳐 예측하였으며, Late fusion는 단일 모달 모델 각각의 분류기를 거친 예측값을 Maximum, Average로 나누어 융합하였다. 두 Fusion 모델의 정확도는 96%로 기존 단일 모델의 성능보다 높게 나옴을 통해 멀티모달로써의 높은 안정성을 검증하였다. 해당 연구에서는 수집한 데이터에 접근할 수 있는 점과 단일모달과 멀티모달을 구분하여 개발하였다는 점에서 본 연구의 취지에 일치하다. 단, 최종 분류 정확도인 96% 값은 실제 기술로 적용하기에 부적합한 수준이기 때문에 본 연구에서는 동일한 데이터셋으로 고도화된 모델과 멀티모달 방식을 적용하였다.

3. 데이터셋 구성

본 연구에서는 멀티모달 형식의 공인 데이터셋인 Multi-modalGasDate[4]를 활용하였다. 본 데이터셋은 4가지 상황(Mixture, NoGas, Perfume, Smoke)에서 동일한 순간에 수집된 가스 센서와 열화상 이미지가 각각 매칭되어 구성되어 있고, 수집 과정은 Fig. 1과 같다. 일반 환경(NoGas)에서 30분마다 가스 분출 간격을 15초씩 늘리며 세 가지 상황(Mixture, Perfume, Smoke)을 구성하였으며, 2초 간격으로 총 6,400건(상황별 1,600건)의 데이터를 수집하였다[4].

가스 센서 데이터는 총 7가지의 가스 센서(MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, MQ135)의 측정값을 포함한 데이터로, 상황별 범주는 Table 2와 Fig. 2와 같다. 특별히 센서에 따라 네 가지의 상황이 일정 부분 구분할 수 있을 것으로 추측했으나, 센서값에 따른 네 가지 상황의 분류가 쉽지 않다.

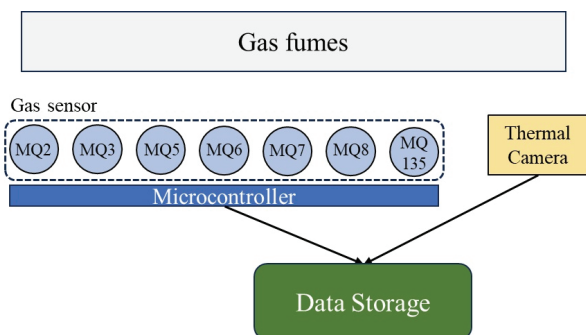


Fig. 1. Process of Data Collection

Table 2. Sensor Value Range by Gas Sensor Type (min, max)

		Mixture	No Gas	Perfume	Smoke
MQ2	min	502	555	592	510
	avg	593	748	745	623
	max	734	809	824	801
MQ3	min	383	512	515	337
	avg	419	529	528	372
	max	460	539	543	451
MQ5	min	344	370	364	291
	avg	407	431	440	740
	max	522	525	596	411
MQ6	min	311	328	341	328
	avg	376	425	429	368
	max	466	517	524	413
MQ7	min	361	506	464	555
	avg	460	606	617	580
	max	796	689	715	609
MQ8	min	220	417	419	541
	avg	315	637	634	584
	max	716	783	794	667
MQ135	min	302	387	389	275
	avg	403	474	481	308
	max	482	584	589	379

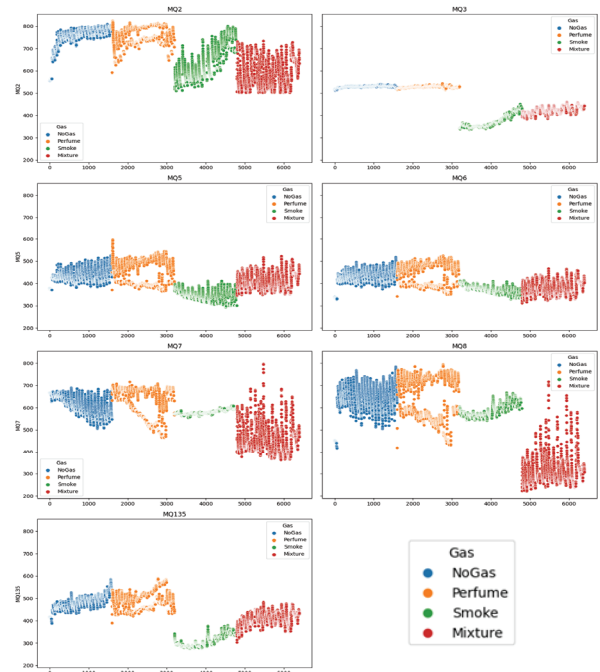


Fig. 2. Distribution of Sensor Values by 4 Gas Classes (Mixture, NoGas, Perfume, Smoke)

특히 NoGas, Perfume의 분포와 평균치를 통해 두 상황이 유사한 것으로 분석하였다. 이로써 단순 규칙 기반의 분류 방식은 활용하기 어려운 점 또한 알 수 있다.

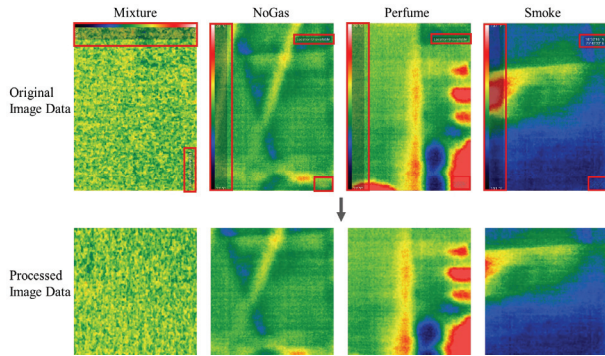


Fig. 3. Thermal Image Data and Result of Preprocess

열화상 이미지 데이터는 206 × 156의 크기이며 36°의 시야로 -40℃부터 330℃의 온도 범위 내 측정이 가능한 소형 열화상 카메라를 활용하여 수집하였다(Seek Thermal Imaging Camera). 네 가지 상황에 대한 대표적인 열화상 이미지는 Fig 3의 원본 사진과 같다. 수집된 이미지에는 카메라 로고, 촬영 위치 정보, 온도 스펙트럼과 온도 값이 상황별로 다른 형태로 포함되어 있음을 확인하였다. 이러한 특징들은 4가지 상황과 무관한 정보임에도 해당 정보에 편향되게 학습이 될 위험이 있고, 다른 환경에서 촬영된 데이터에 대해서 적용하지 못할 수 있다. 이에 따라 관련한 특징을 Crop한 뒤 일정한 크기(224×224)로 조정(Resize)하여 전처리하였고, 전처리 전/후 이미지는 Fig. 3과 같다.

4. 멀티모달형 가스식별 모델 개발

가스 종류 인식 모델은 먼저 가스 센서와 열화상 이미지에 대한 단일모달 네트워크를 개발하였고, 이후 단일모달에서 가장 높은 성능의 네트워크를 중심으로 멀티모달 모델을 개발 및 분석하였다. 가스 센서 데이터의 경우 7가지 센서값을 하나의 입력부로 데이터 체계를 구성하였고, 열화상 이미지는 전처리 이미지를 활용하되, 전처리의 효과를 확인하기 위해 원본 이미지도 적용하였다. 라벨의 경우 가스, 열화상 데이터에 공통적으로 네 가지의 상황을 각각 ('Mixture':0, 'NoGas':1, 'Perfume':2, 'Smoke':3) 숫자로 매핑하여 진행하였다. Table 3과 같이 구축된 데이터셋을 기반으로 단일모달과 멀티모달 모델을 학습하였다. 모델의 성능 비교를 위해 학습(검증)/평가 데이터셋을 4:1의 비율로 나눈 뒤 학습 데이터셋은 4:1로 나누어 동일하게 모든 모델의 학습과 검증에 적용하였다. 이때 상황별 데이터가 균일하기 때문에 각 데이터셋 내의 네 가지 상황의 데이터 비율은 유지하여 분리하였다. 모델들은 Python 내 Keras 프레임워크를 통해 개발하였으며, 독자적으로 개발한 모델들의 경우, 동일한 초매개변수로 설정하여 진행되었다. (조기종료, Batch 정규화 적용, Dropout: 0.3, Optimizer: Adam, Activation: ReLu, Batch size: 32, Epoch: 200, Eardly Stopping Patience: 12, Learning rate : 0.001) 모델 학습의 경우 동일한 GPU (NVIDIA Quadro P5000)와 64GB RAM을 활용하였다.

Table 3. Dataset for Multimodal

Label	Mixture(0)	NoGas(1)	Perfume(2)	Smoke(3)
Gas sensor	[206,392, 344,311,395, 222,302]	[555,515, 377,338,666, 451,416]	[592.517, 370.341.632, 419.389]	[518,341, 365,400,572, 572,332]
Thermal Image				

4.1 단일모달형 가스식별 모델

가스 센서 데이터는 단순 수치 데이터의 형식을 지니고 있다. 해당 데이터의 특성을 고려하여 RNN 계열의 LSTM(Long Short Term Memory)[10]과 1D CNN(One-Dimensional Convolutional Neural Network)[11], 다중 신경망을 통한 분류기 네트워크로 쓰이는 DNN(Deep Neural Network)으로 총 세 가지의 모델을 비교 분석하였다. 다양한 실험을 거쳐 모델을 개발하였으며, 개발한 1D CNN 기반 네트워크의 경우 Fully Connected Layer(3층)를 분류기로 활용하였으며, Softmax 함수를 활성화 함수로 적용하였다. LSTM 모델은 64개의 units으로 이루어진 순환층을 학습한 후 Fully Connected Layer를 거쳐 최종 예측을 수행하고, DNN 모델은 Fully Connected Layer(4층)를 기반으로 분류기 위주의 학습 후 예측한다. 구체적인 모델의 구조는 Fig. 4를 통해 확인할 수 있다.

열화상 이미지 데이터의 경우 이미지 특징 추출에 적합한 CNN(Convolution Neural Network) 기반의 VGG16[12], ResNet[13], GoogLeNet[14]을 백본망으로 하여 설계한 모델 (scratch 학습)과 직접 설계한 GasNet 모델을 학습 및 비교분석 하였다(Fig. 5). 개발한 네트워크는 VGG16 형식의 stacking network 형식을 지닌다. 다양한 초매개변수의 조정 등을 거쳐 최종적으로 특징 추출을 위해 6층, 분류를 위해 3층으로 구성하고, 커널 크기는 3×3으로 지정하였다.

4.2 멀티모달형 가스식별 모델

가스 센서에서의 1D CNN과 열화상 이미지에서의 GasNet을 기반으로 멀티모달 모델을 구축하였다. 멀티모달은 단일 모달 네트워크의 결합할 층의 위치와 결합하는 전략에 따라서 크게 3가지 결합 방식(Early fusion, Late fusion, Hybrid fusion)으로 나눌 수 있다[15]. Early fusion은 데이터의 특징 벡터를 결합함으로써 모델의 학습 초기 상호 작용을 활용하는 특징 기반 융합방식이며, Late fusion은 단일모달 모델 각각의 예측값을 결합하여 최종 예측에 활용하는 분류단계의 융합 방식이다[16, 17]. Hybrid fusion은 Early fusion과 Late fusion 방식을 모두 적용하도록 설계된 모델이다[18].

본 연구에서는 Early fusion 기반의 모델(Fusion I), Late fusion 기반의 두 가지 모델(FusionII, FusionIII), Hybrid fusion 기반의 모델(FusionIV)로 총 네 가지의 멀티모달 모델을 구축하여 비교 분석하였다. Fig. 6은 각 멀티모달 모델의 네트워크

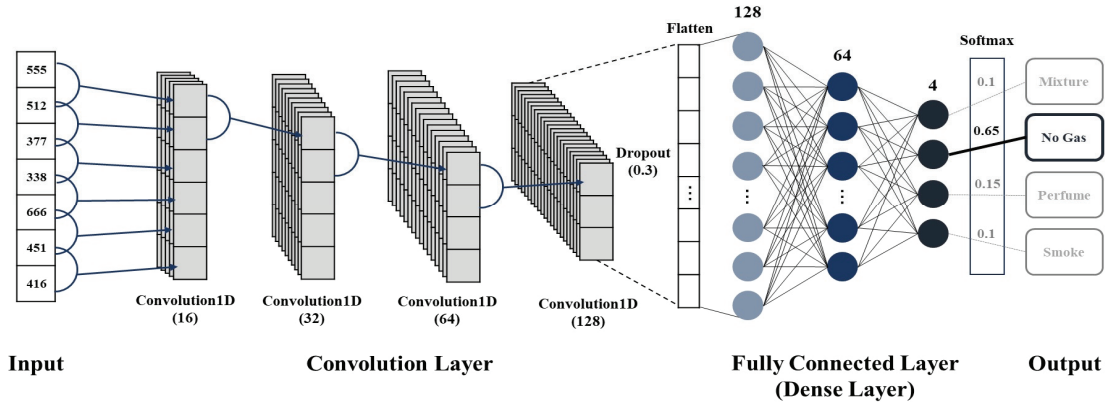


Fig. 4. 1D CNN Model Network Structure

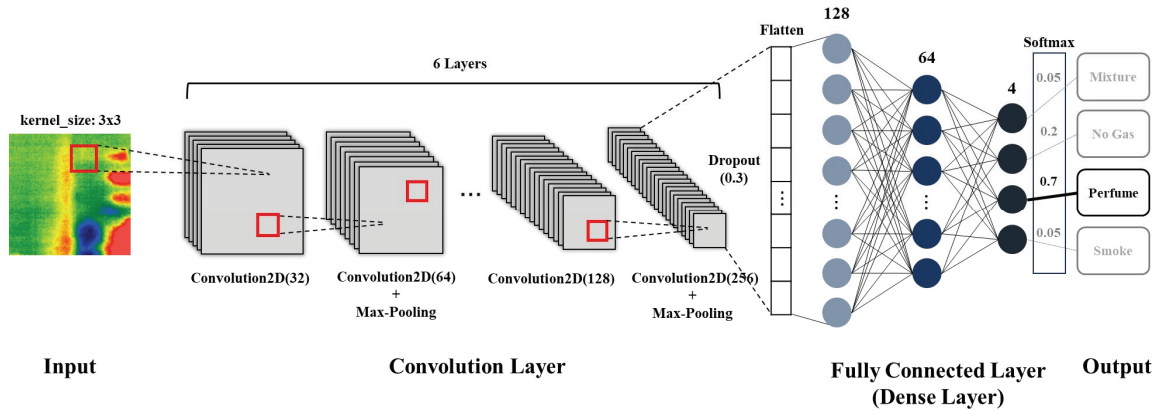


Fig. 5. GasNet Model Network Structure

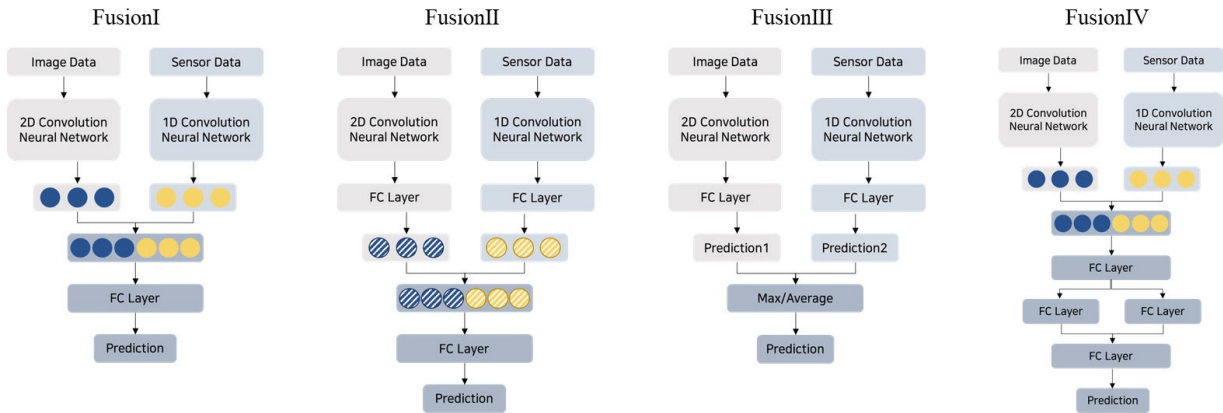


Fig. 6. Multimodal Model Network Structure

구조도이며, Figs. 7-10은 학습의 전반적인 흐름을 파악할 수 있는 멀티모달 모델들의 pseudo code이다.

Fusion I: Fusion I 모델의 경우, 1D CNN, GasNet 네트워크에서 분류기 이전의 최종 특징 벡터를 결합하여 Fully Connected Layer를 거쳐 예측하는 모델이다. 각 데이터의 특징 벡터를 결합하여 종합적인 정보를 기반으로 분류하기에 다른 모델보다 단일모달 모델의 의존도가 낮다. 기존 각 네트워크에서의 특징 벡터가 각각 2304, 384의 길이를 가짐에 따라 합

치면서 2688의 길이를 분류할 수 있도록 하였다.

Fusion II: 1D CNN, GasNet을 학습한 후에 Fully Connected Layer를 통해 분류 작업을 수행하는 중간층을 결합하는 모델이다. 분류 중간 과정에서 결합하여 최종 예측을 수행함으로써 Fusion I보다 단일모달 모델에 대한 의존도가 비교적 높다. 각각의 모델을 4의 길이로 분류 후 융합하여 이를 최종 예측하였다.

Fusion III: 앙상블 모델 구조로도 활용되며, 가스 센서, 이미지 모델을 학습 시에 결합하는 것이 아닌 각각 완료한 학습 결

과인 예측값을 결합하는 모델이다. 두 모델의 예측 확률값에 대해 최댓값이나 평균으로 예측하는 방식이 일반적이며 비교해본 결과 최댓값을 활용한 예측이 더 높은 성능을 보였다. 데이터 하나의 예측값 중 최댓값을 결과로 가져오는 방식으로써 단일모달 모델에 대한 의존도가 가장 높고 예측 결과에 대하여 두 모델이 보완될 수 있다는 강점이 있다. 그러나 두 도메인 간의 연계적인 특징을 학습하지 못하는 한계점을 가진다.

FusionIV: Fusion I 모델과 FusionII 모델을 융합한 하이브리드 형식으로 설계하였으며, 두 번의 fusion을 거쳐 멀티모달 데이터를 학습한다. 단일모달 모델(ID CNN/ GasNet)을 통해 데이터의 특징을 학습한 후 Early fusion을 진행한다. 이후 Fully Connected Layer를 거쳐 예측 학습 뒤 다시 두 모델로 분리되어 각각의 분류 층을 학습하고, Late fusion을 통해 최종 예측 수행한다. 두 개의 융합 방식을 적용함으로써 단일 모델만의 독자적인 특징과 연계적인 특징을 모두 학습한다는 강점을 가지지만, 모델이 복잡하여 두 특징이 학습되기 위한 세밀한 전략이 요구되며 학습 시간이 오래 소요되는 한계점이 존재한다.

FusionI Model

Input: X_{img}, X_{sensor}
Output: Probability of Classes
Image Model: $img_model \leftarrow CNN2D(X_{img})$
Sensor Model: $gas_model \leftarrow CNN1D(X_{sensor})$
Concatenate Model Outputs:
 $concat_model \leftarrow Concatenate([img_model.output, gas_model.output])$
 $concat_model \leftarrow FC\ Layer(concat_model)$
Training Loop:
for $epoch$ **in** 1 to 100 **do**
 $history \leftarrow concat_model.fit([img_train, gas_train], label_train,$
 $batch_size = 32, validation_data = ([img_val, gas_val], label_val))$
if $earlystopper.should_stop(history)$ **then**
break
end if
Parameter Update: basis on Back propagation
 $concat_model, img_model, gas_model \leftarrow backward()$
end for
 $probability = concat_model.predict([img_test, sensor_test])$

Fig. 7. Pseudo Code of Fusion I Model

FusionII Model

Input: X_{img}, X_{sensor}
Output: Probability of Classes
Image Model:
 $img_model \leftarrow CNN2D(X_{img})$
 $img_model \leftarrow FC\ Layer(img_model)$
Sensor Model:
 $gas_model \leftarrow CNN1D(X_{sensor})$
 $gas_model \leftarrow FC\ Layer(gas_model)$
Concatenate Model Outputs:
 $concat_model \leftarrow Concatenate([img_model.output, gas_model.output])$
 $concat_model \leftarrow FC\ Layer(concat_model)$
Training Loop:
for $epoch$ **in** 1 to 100 **do**
 $history \leftarrow concat_model.fit([img_train, gas_train], label_train,$
 $batch_size = 32, validation_data = ([img_val, gas_val], label_val))$
if $earlystopper.should_stop(history)$ **then**
break
end if
Parameter Update: basis on Back propagation
 $concat_model, img_model, gas_model \leftarrow backward()$
end for
 $probability = concat_model.predict([img_test, sensor_test])$

Fig. 8. Pseudo Code of FusionII Model

FusionIII(Max) Model

Input: X_{img}, X_{sensor}
Output: Probability of Classes
Image Model:
 $img_model \leftarrow CNN2D(X_{img})$
 $img_model \leftarrow FC\ Layer(img_model)$
Sensor Model:
 $gas_model \leftarrow CNN1D(X_{sensor})$
 $gas_model \leftarrow FC\ Layer(gas_model)$
Training Loop:
for $epoch$ **in** 1 to 100 **do**
 $history \leftarrow img_model.fit(img_train, label_train, batch_size = 32,$
 $validation_data = (img_val, label_val))$
 $history \leftarrow sensor_model.fit(sensor_train, label_train, batch_size = 32,$
 $validation_data = (sensor_val, label_val))$
if $earlystopper.should_stop(history)$ **then**
break
end if
Parameter Update: basis on Back propagation
 $img_model, gas_model \leftarrow backward()$
end for
 $img_probability = img_model.predict(img_test)$
 $sensor_probability = sensor_model.predict(sensor_test)$
 $max_probability = maximum(img_probability, sensor_probability)$

Fig. 9. Pseudo Code of FusionIII Model

FusionIV Model

Input: X_{img}, X_{sensor}
Output: Probability of Classes
Image Model: $img_model \leftarrow CNN2D(X_{img})$
Sensor Model: $gas_model \leftarrow CNN1D(X_{sensor})$
Concatenate Model Outputs:
 $concatenate \leftarrow Concatenate([img_model.output, gas_model.output])$
 $feature_output \leftarrow FC\ Layer(concatenate)$
 $img_predictor \leftarrow FC\ Layer(feature_output)$
 $sensor_predictor \leftarrow FC\ Layer(feature_output)$
 $decision_concat \leftarrow Concatenate([img_predictor, sensor_predictor])$
 $concat_model \leftarrow Model([img_model.input, gas_model.input], decision_concat)$
Training Loop:
for $epoch$ **in** 1 to 100 **do**
 $history \leftarrow concat_model.fit([img_train, gas_train], imglabel_train,$
 $batch_size = 32, validation_data = ([img_val, gas_val], imglabel_val))$
if $earlystopper.should_stop(history)$ **then**
break
end if
Parameter Update: basis on Back propagation
 $concat_model, img_model, gas_model \leftarrow backward()$
end for
 $probability = concat_model.predict([img_test, sensor_test])$

Fig. 10. Pseudo Code of FusionIV Model

5. 모델 결과 및 분석

5.1 평가 방법 및 기준

단일모달 모델과 멀티모달 모델의 성능은 혼돈행렬을 기반으로 분류 정확도를 평가한다. 혼돈행렬은 실제 라벨값과 예측값을 비교하며 이진 분류의 경우 Fig. 11과 같다. True Positive(TP)는 예측값과 실제값이 모두 참일 경우를 의미하며, False Positive(FP)는 예측값은 참이지만 실제값이 거짓인 경우이다. False Negative(FN)는 예측값이 거짓일 때 실제값이 참인 경우를 의미하며, True Negative(TN)는 예측값과 실제값이 모두 거짓이다. 혼돈행렬을 기반으로 네 가지의 지표를 가지고 모델의 평가가 가능하며, 지표는 다음과 같이 계산된다. Precision(정밀도)은 모델이 참으로 예측값 중 실제값이 참인 비율이며, Recall(재현율)은 실제값이 참인 경우 중 예측값이 참인 비율이다. 두 성능지표를 모두 반영한 지표는 F1-

		Actual values	
		True	False
Predicted values	True	True Positive(TP)	False Positive(FP)
	False	False Negative(FN)	True Negative(TN)

Fig. 11. Confusion Matrix

Score이며, 정밀도와 재현율의 조화평균으로 계산한다. 마지막으로 Accuracy(정확도)의 경우 전체의 예측값 중 실제 예측값과 일치한 비율로 정의한다. 본 연구에서는 선행 논문의 참조를 종합하여 정확도와 F1-score를 모델들의 성능지표로 평가한다. 이때 클래스별 데이터 개수가 동일한 점을 토대로 F1-score는 macro average를 적용한다.

5.2 단일모달 모델 결과 분석(가스 센서)

모델 학습 과정에서 LSTM과 DNN이 1분 내외, 1D CNN은 2분 정도로 비교적 빠른 학습 속도를 보이며 세 모델 모두 1초 이내로 예측한다. 이는 센서 데이터의 구조가 단순하고 모델 학습 시 계산량이 다른 데이터에 비해 비교적 적어 수행 시간이 빠르다. 모델별 정확도와 F1-score 값은 Table 4와 같다. 먼저, 1D CNN 모델이 96.3%의 정확도로 가장 높은 성능을 보였고, 이후에 DNN, LSTM 모델 순으로 집계된다. 두 번째로 기존 논문에서 적용하였던 LSTM 모델보다 본 연구를 통해 개발한 LSTM 모델은 6.5% 높게, 1D CNN 모델은 14.3% 높게 나왔다. LSTM 모델의 경우 입력하는 데이터가 큰 경우에 강인한 모델로 알려져 있으나, 해당 모델에서는 입력부가 7개로 작은 편이기에 LSTM보다 1D CNN 모델이 더 잘 나오게 된 점을 확인하였다. 또한 DNN모델보다 1D CNN 모델이 더 잘 나온 점을 통해 가스 센서들이 독립적인 성격을 가지지 않고 연계적인 성격을 가지고 있다는 점을 추론할 수 있다. 실제로 센서들 각각이 측정하는 가스의 성격이 겹치는 경우가 있기에 연계될 수 있다는 근거가 된다. F1-score의 경우에도 다른 모델과 달리 1D CNN 모델이 기존 연구보다 6.3% 높게 나온 점을 토대로 오류에 대한 예측도 감소하여 전반적인 예측에 있어 안정적인 모델이다. 또한 LSTM[9] 모델의 경우 NoGas와 Perfume 사이에서 218개의 오분류와 이외 소수의 다양한 오분류가 발생하는 반면, 1D CNN 모델의 경우 NoGas와 Perfume 사이에서만 132개의 오분류가 존재한다. 이를 토대로 기존 성능에 비해 분류 예측 성능이 향상되었음을 확인하였다. 이는 이를 기반으로 가스 센서에 대한 단일 모델로 1D CNN 모델을 선정하였다.

Table 4. Comparison of Unimodal Model(Gas Sensor)

	Accuracy	F1-score
LSTM(3 layers)	0.8852	0.8848
1D CNN	0.9625	0.9625
DNN	0.8969	0.8969
LSTM[9](5 layers)	0.82	0.90

5.3 단일모달 모델 결과 분석(열화상 이미지)

열화상 이미지 단일 모달의 경우 평균적으로 10분 내외로 학습하며, 30초 이내로 예측한다. 학습 데이터가 이미지인 점에서 센서 단일모달 모델에 비해 오랜 수행 시간이 소요된다. 각각의 모델에 대한 분류 정확도와 최종적인 손실 값은 Table 5와 같다. 먼저 직접 구축한 GasNet이 96.4%로 가장 좋은 성능이다. 이는 기존 연구의 가장 높은 성능보다도 3.41% 높은 수치로써 의미가 있다. 또한 F1-score의 경우에도 기존 연구보다 3.4% 높은 점을 토대로 비교적 안정적인 예측 모델이다. 또한 오분류 사례에 대해서도 비교해 본 결과 기존 연구 모델의 85개보다 약 45% 감소한 46개로 예측 성능의 개선을 확인하였다. 구축한 데이터셋의 경우 GasNet 모델의 깊이가 일정 기준보다 깊어지는 경우 성능이 낮아진다. 이는 열화상 이미지가 기존 RGB로 진행되는 이미지와 비교해 이미 특징이 향상된 상태이기 때문으로 판단된다. 기존의 네트워크들은 형상(Shape)에 대한 특징을 학습하는 데에 최적화된 모델이지만, 열화상 이미지의 경우 형상보다도 픽셀마다의 값 그 자체에 더 중요도를 갖추고 있다.

이에 대한 근거를 검증하고 이미지 내 가스와 무관한 물체가 학습되진 않았는지 확인하고자 GasNet에 대한 Grad-CAM[19]을 적용하였다. Grad-CAM은 학습한 네트워크를 통해 이미지 내 활성화된 위치를 가시화하는 모델로, 중요한 특징으로 판단될수록 밝은색을 띠는 특징을 가진다. 데이터셋 중에는 직접 가스 분사하는 과정에서 사람의 손도 찍힌 이미지가 포함되어 있었고, 이러한 데이터들을 중심으로 활성화맵을 확인하였다. Fig. 12는 두 장의 Perfume과 Smoke의 Grad-CAM 적용 결과 이미지다. 왼쪽부터 차례대로 원본 이미지, Grad-CAM 필터, 원본과 필터를 씌운 이미지를 출력한 결과다. 결과에서 손이 나타나는 부위가 어두운색으로써 실제 학습할 때 활성화되지 않고 학습된 것을 볼 수 있다. 예측 원인은 다른 외부 요인(손, 물체 등)이 아닌 공기 부분의 온도 값을 기반으로 상황을 예측하는 것으로 분석된다. 본 모델을 기반으로 이미지 분류모델에 대한 신뢰성을 확보하고, 이후 멀티모달 가스식별 모델의 개발에 GasNet을 활용하였다.

또한 전처리의 효과를 확인하기 위해 전처리 전/후에 대한 모델 성능과 Grad-CAM 결과도 분석하였다. 실제로 전처리 전 분류 성능이 97.3%로 나와 전처리 후의 성능보다 좋은 것

Table 5. Comparison of Unimodal Model(Thermal Image)

	Accuracy	F1-score
GoogLeNet	0.9344	0.9341
ResNet(34)	0.9563	0.9562
VGG16	0.9187	0.9185
GasNet (processed image)	0.9641	0.9641
GasNet (original image)	0.9773	0.9773
CNN[9]	0.93	0.93

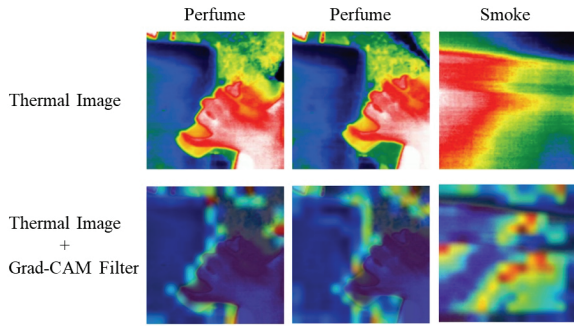


Fig. 12. Thermal Image(Preprocessed) with Grad-CAM

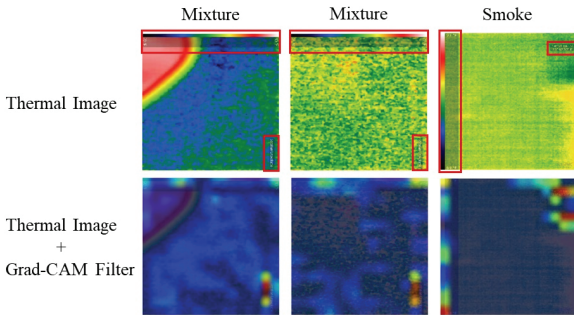


Fig. 13. Thermal Image(Original) with Grad-CAM

처럼 보인다. 이는 이미지 내에서 스펙트럼이나 온도 숫자 값 등의 단서가 학습에 반영됨에 따라 그 난도가 낮아졌기 때문으로 판단된다. 또한 Grad-CAM 결과로 표출된 Fig. 13을 통해 확인할 수 있듯이, 실제로도 스펙트럼 구간과 다른 구간에 활성화가 나타난 점을 통해 이미지 분류에 적용되었다는 점을 확인할 수 있다. 즉, 촬영하는 구도가 변경되면 언제든 원하는 분류를 할 수 없다는 한계를 지녀 최종적으로 전처리 된 이미지를 기준으로 학습된 CNN 기반의 네트워크인 GasNet을 최종 모델로 선정하였다.

5.4 멀티모달 모델 결과 분석

멀티모달 학습의 경우 이미지 단일모달 모델과 유사하게 10분 내외의 학습 시간을 보이며, 예측 시간은 또한 30초 내외로 유사하다. 단일모달의 구조를 기반으로 설계한 모델임에도 소요 시간이 증가하지 않은 점에서 다양하게 멀티모달 모델이 사용될 수 있음을 확인하였다. 네 가지 멀티모달 모델에 대한 학습 추이는 Fig. 14와 같다. 조기 종료를 적용하여 35-66 사이의 epochs로 학습하였고, 3 epochs 이상부터 학습 정확도의 min-max 차이가 0.2 이하이며, 20 epochs 이상부터 비슷한 수치의 정확도가 유지된 점에서 안정적인 학습 과정을 보인다. 학습을 거친 최종 멀티모달 모델의 분류 성능과 손실 값은 Table 6과 같다. 먼저 Fusion I 방식이 99.3%로 가장 높은 분류 성능을 보였으며, FusionIV도 유사한 성능이다. 또한 F1-score에도 동일하게 기존 연구보다 3% 높은 점을 토대로 균형잡힌 예측을 수행하고 있음을 확인하였고, 이는 오분류에도 강인한 모델임을 의미한다. 이는 두 데이터의 특징을 파악

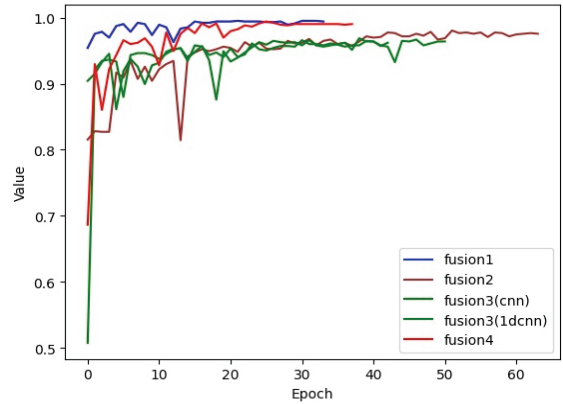


Fig. 14. Multimodal Model Training Trend

Table 6. Comparison of Multimodal Model

	Accuracy	F1-score
Fusion I	0.9930	0.9930
FusionII	0.9852	0.9852
FusionIII	0.9789	0.9788
FusionIV	0.9906	0.9906
Early Fusion[3]	0.96	0.96

하여 두 단일모달 사이의 연계점이 학습을 통해 최적화를 이룰 수 있기 때문이다. 또한 초기 단계에서 융합하여 최적화 과정에서 다른 멀티모달 모델들보다 빠르게 단일모달 모델들의 미흡한 부분들을 보완하며 학습했기에 높은 성능을 도출함으로 분석된다. 또한 기존 가스 센서에서의 96.3%와 열화상 이미지에서의 96.4%의 성능보다 대부분 더 높아졌다. 이를 통해 가스 센서와 열화상 이미지 간의 정보 간의 상관관계가 높다는 점을 확인할 수 있다. 세 번째로 기존 논문의 96% 성능과 대비하여 3.3% 성능이 높아졌다.

기존 연구 모델의 경우 50개의 오분류 사례가 발생한 것을 기반으로 본 연구 모델들의 Error를 분석하고자 Figs. 15-18으로 멀티모달 모델의 혼동행렬을 확인하였다. 이는 기존 연구의 오분류 수보다 모든 모델에서 50% 이상 감소한 수치이다. 대부분의 오분류를 포함하여, 특별히 1% 내의 모든 Error는 No Gas와 Perfume 상황 사이에서 발생함을 확인하였는데, 일부 Perfume 데이터는 No Gas로 예측하며 No Gas는 Perfume으로 예측한 경우가 발생하였다. 발생한 오분류 데이터와 Table 7과 같은 샘플을 확인하였다. 해당 샘플을 통해 알 수 있듯이, 열화상 이미지가 매우 비슷하게 보이며, 당시의 가스 센서에 대해 네트워크를 구동한 결과도 동일하게 No Gas 범주에 포함된다. Perfume에도 불구하고 No Gas와 유사한 데이터가 들어간 배경을 분석한 결과 향수라고 라벨링되어 있지만, 실제 향수를 뿌리기 이전 타이밍의 데이터로 적용됨에 따라 라벨과는 달리 No Gas 상황인 것이 타당한 일종의 데이터 Outlier로 확인하였다. 즉, 데이터상의 오류를 제외하면 본 멀티모달 모델이 네 가지의 상황에 대해 정확하게 분류하는 점을 입증할 수 있다.

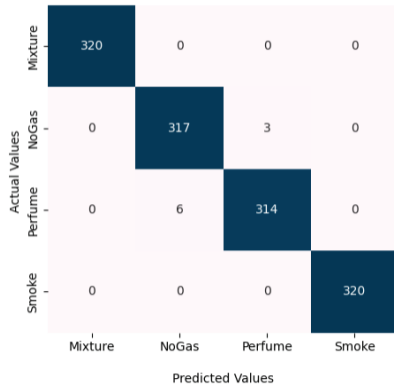


Fig. 15. Fusion I Confusion Matrix

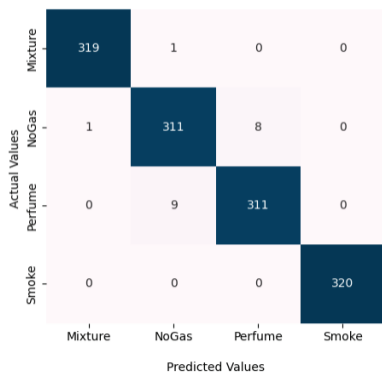


Fig. 16. Fusion II Confusion Matrix

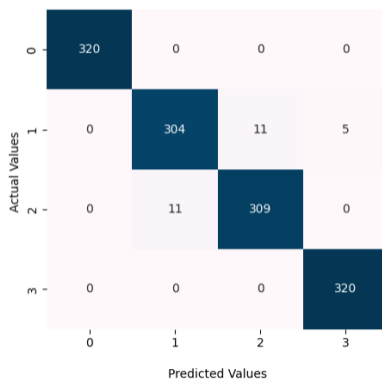


Fig. 17. Fusion III Confusion Matrix

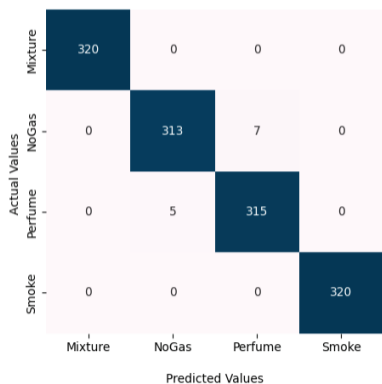
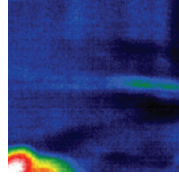
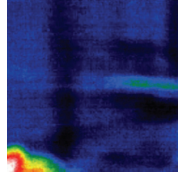


Fig. 18. Fusion IV Confusion Matrix

Table 7. Comparison of Misclassification Data

	No Gas	Perfume
Gas sensor	[798, 532, 506, 484, 676, 729, 513]	[710, 527, 409, 397, 630, 588, 436]
Thermal Image		
Predicted Values	No Gas	No Gas

6. 결론 및 향후 연구 방향

본 연구에서는 가스 누출 감지 시스템을 위해 다양한 Modality인 가스 센서와 열화상 이미지 정보를 종합적으로 활용한 딥러닝 기반의 누출 검출 모델을 개발하였다. 불필요한 표시 정보로 인해 모델 학습에 미치는 영향을 최소화하기 위한 데이터 전처리를 수행하였고, 멀티모달로부터 추출된 정보를 융합하는 시점에 따라 다양한 모델을 설계 및 분석하였다. 본 연구에서 개발한 멀티모달 모델의 성능이 기존 논문의 멀티모달 성능과 단일모달(가스센서, 열화상 이미지)의 성능보다 우세함을 통해 이에 대한 가스 점검 시스템의 적용 가능성을 확인하였다. 본 연구를 토대로 더 다양하고 많은 가스 누출 사례에 대한 분류모델로 확장하고, 가스 누출의 실시간 감지 여부를 확인할 수 있는 실증형 연구를 진행하고자 한다. 멀티모달을 적용한 높은 신뢰도의 가스 누출 감지 시스템을 통해 실제 현장에서 가스 누출에 대한 피해를 최소화할 것으로 기대한다. 본 연구의 모델이 가스 누출 감지 시스템에만 적용되는 것이 아닌 화재 탐지와 같이 다양한 멀티모달이 필요한 산업현장과 일상생활에서도 근무자들의 안전을 위해 활용되길 기대한다.

References

- [1] 화학물질 사고현황 및 사례 [Internet], <https://icis.me.go.kr/search/searchType2.do>(화학물질안전원).
- [2] 가스사고 발생현황 및 사고일람표 [Internet], <https://www.kgs.or.kr/kgs/abcc/board.do>(한국가스안전공사).
- [3] A. M. Adesina, W. K. Fung, and A. Karnik, "Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches," in *Sensors*, Vol.19, No.11, pp.2548, 2019.
- [4] P. Narkhede, R. Walambe, P. Chandel, S. Mandaokar, and K. Kotecha, "MultimodalGasData: Multimodal dataset for gas detection and classification," *Data*, Vol.7, No.8, pp.112, 2022.
- [5] M. M. Khan, "Sensor-based gas leakage detector system," in *Engineering Proceedings*, Vol.2, No.1, pp.28, 2020.

[6] J. M. Ahn, G. Y. Kim, and D. J. Kim, "A real time urban gas leakage prediction and monitoring system based on deep learning," *Journal of the Korean Institute of Plant Engineering*, Vol.26, No.4, pp.45-53, 2021.

[7] J. Shi, Y. Chang, C. Xu, F. I. Khan, G. Chen, and C. Li, "Real-time leak detection using an infrared camera and Faster R-CNN technique," *Computers & Chemical Engineering*, Vol.135, pp.106780, 2020.

[8] R. Xiao, Q. Hu, and J. Li, "Leak detection of gas pipelines using acoustic signals based on wavelet transform and Support Vector Machine," *Measurement*, Vol.146, pp.479-489, 2019.

[9] P. Narkhede, R. Walambe, S. Mandaokar, P. Chandel, K. Kotecha, and G. Ghinea, "Gas detection and identification using multimodal artificial intelligence based sensor fusion," *Applied System Innovation*, Vol.4, No.1, pp.3, 2021.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, Vol.9, No.8, pp.1735-1780, 2010.

[11] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier," *Journal of Signal Processing Systems*, Vol.91, pp.179-189, 2019.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint, arXiv:1409.1556*, 2014.

[13] B. Koonce, "Convolutional neural networks with swift for tensorflow: Image recognition and dataset categorization," Springer, 1st ed., CA: Apress Berkeley, pp.51-61, 2021.

[14] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, pp.1-9, 2015.

[15] T. Baltrušaitis, C. Ahuja, and L. P. Morency, "Multimodal machine learning: A survey and taxonomy," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.41, No.2, pp.423-443, 2019.

[16] K. Gadzicki, R. Khamsehashari, and C. Zetsche, "Early vs late fusion in multimodal convolutional neural networks," in *IEEE 23rd international conference on information fusion (FUSION2020)*, Virtual, pp.1-6, 2020.

[17] K. Liu, Y. Li, N. Xu, and P. Natarajan, "Learn to combine modalities in multimodal deep learning," in *arXiv preprint, arXiv:1805.11730*, 2018.

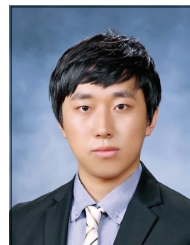
[18] P. K. Atrey, M. A. Hossain, A. E. Saddik, and M. Kankanhalli, "Multimodal fusion for multimedia analysis: A survey," *Multimedia Systems*, Vol.16, pp.345-379, 2010.

[19] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision*, Venecia, pp.618-626, 2017.



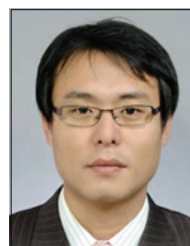
안 서 희

<https://orcid.org/0009-0008-6457-7737>
 e-mail : seohee6544@postech.ac.kr
 2022년 인천대학교 수학과(학사)
 2022년 ~ 현 재 포항공과대학교
 인공지능연구원 연구원
 관심분야 : Deep Learning, Machine Learning, Computer Vision



김 경 영

<https://orcid.org/0000-0001-6891-4130>
 e-mail : gy_kim@postech.ac.kr
 2015년 KAIST 수리과학과(학사)
 2018년 KAIST 건설및환경공학과(석사)
 2020년 ~ 현 재 포항공과대학교
 인공지능연구원 팀장
 관심분야 : Computer Vision, Face Recognition, Deep learning



김 동 주

<https://orcid.org/0009-0009-6950-4200>
 e-mail : kkb0320@postech.ac.kr
 1998년 충북대학교 전파공학과(학사)
 2000년 충북대학교 전파통신학과(석사)
 2010년 성균관대학교 전기전자 및 컴퓨터공학과(박사)
 2016년 ~ 현 재 포항공과대학교 인공지능연구원 연구부장/
 연구부교수
 관심분야 : Computer Vision, Image Processing, Deep Learnin