

IF2bNet: 화재 감지를 위한 설명 가능 AI 기반 최적화된 딥러닝 아키텍처

진원¹, 송미화²

¹ 세명대학교 소방방재학과 학부생

² 한성대학교 컴퓨터공학부 교수

wlsdnjsxx@semyung.ac.kr, mhsong@hansung.kr

IF2bNet: An Optimized Deep Learning Architecture for Fire Detection Based on Explainable AI

Won Jin¹, Mi-Hwa Song²

¹Department of Fire and Disaster Prevention Engineering, Semyung University

²Division of Computer Engineering, Hansung University

요 약

센서 기반의 자동화재탐지설비의 역할을 지원할 목적으로, 합성곱 신경망 기반의 AI 화재 감시 장비등이 연구되어왔다. ai 기반 화재 감지에 사용되는 알고리즘은 전이학습을 주로 이용하고 있고, 이는 화재 감지에 기여도가 낮은 프로세스가 내장되어 있을 가능성이 존재하여, 딥러닝 모델의 복잡성을 가중시키는 원인이 될 수 있다. 본 연구에서는 이러한 모델의 복잡성을 개선하고자 다양한 딥러닝 및 해석 기술들을 분석하였고, 분석 결과를 토대로 화재 감지에 최적화된 아키텍처인 “IF2bNet”을 제안한다. 구현한 아키텍처의 성능을 비교한 결과 동일한 성능을 내면서, 파라미터를 약 0.1 배로 경량화 하여, 복잡성을 완화하였다.

1. 서론

건축물에서의 화재 방지를 위한 다양한 연구가 진행되어왔으며 건물 내부 설비와 연동된 합성곱 신경망 (CNN) 알고리즘 기반의 화재 감시 장비 등이 연구되었다[1]. 기존 CNN 연구는 보편적인 객체를 탐지하는데 초점을 맞춘 전이학습 기반의 알고리즘을 사용하는 것이 현재 동향이고, 이는 알고리즘 내에 화재와 관련이 적은 부가적인 프로세스가 내장되어 있을 가능성을 시사한다. 본 연구에서는 화재 감시 기능에 최적화된 고유한 합성곱 신경망 아키텍처를 개발하고자 하며, 이를 도출하고자 사용하는 기술로 설명가능한 인공지능 기반의 Rule Extraction [2][3]을 활용, 최적화된 아키텍처 “IF2bNet”을 제안한다.

2. 연구방법론

Convolutional Neural Network (CNN)은 이미지 인식 및 분류에 사용되는 인공 신경망의 유형으로, 이미지에서 특징을 추출하고 분류하는 데 사용된다.

이미지에서 특징을 추출하는 프로세스를 화재 관점에서 최적화된 형태로 도출하고, 화재 데이터를 학습시킨 뒤, 전이학습 모델과의 성능차이를 비교하여 모델의 성능을 평가하고자 하며 구체적인 System Overview는 Fig 1 과 같다.

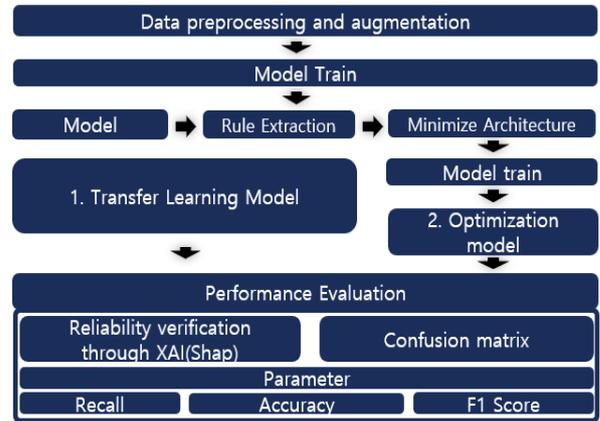


Fig 1. System Overview

Rule Extraction 은 모델이 학습한 규칙을 특정 조건에 따라 단순화시키는 과정으로, 추출된 간결한 규칙은 모델의 핵심적인 동작을 나타낸다. Fig 2 에서 의사코드를 보인다.

```

def extract_rules(model):
    layer_weights = [layer.get_weights() for layer in model.layers]
    rules = []
    for weights in layer_weights:
        for w in weights:
            if np.mean(w) > 0.5: # Setting thresholds (example)
                rules.append("Contribute to fire detection if it exceeds 0.5")
    return rules
  
```

Fig 2. Pesudo Code for Rule Extraction

본 연구에서는 Xception[4]과 같은 전이학습 알고리즘을 활용하여 CNN 모델을 구축한 후, 모델에 학습된 가중치 및 특징 맵을 분석하여 각 레이어의 화재 인식에 대한 기여도를 계산한다. 이를 통해 모델의 어떤 레이어가 중요하게 여겨졌는지를 식별한다. 기여도가 높은 레이어부터 차례대로 추출하여, 추출한 레이어를 기반으로 모델을 새롭게 구현한다[4]. 구체적인 과정은 Fig 3 와 같다.

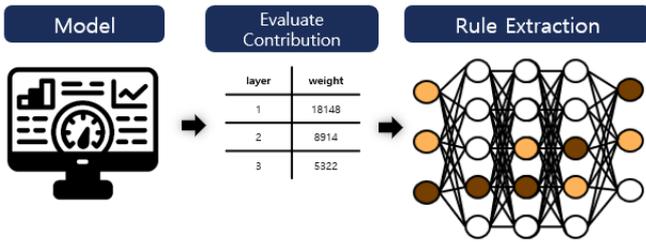


Fig 3. Extract Optimization Model

3. 실험 결과

모델의 성능을 보다 정확히 평가하기 위하여 Test dataset 기반의 모델 평가 결과를 시각화하였다. Table 1 은 모델의 정확도 추이를 나타내고, Fig 4,5 는 모델의 Confusion Matrix 를 나타낸다.

Table 1. Model Evaluation Indicator Analysis

Name	Class	Accuracy	Precision	Recall	F1_score
FCPNet	Fire	0.98	0.95	0.99	0.97
	Non_fire		0.99	1.00	1.00
	Smoke		1.00	0.95	0.97
Xception	Fire	0.88	0.85	0.83	0.84
	Non_fire		0.87	0.90	0.88
	Smoke		0.94	0.92	0.93

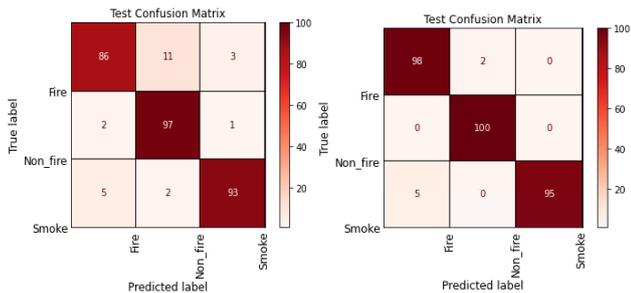


Fig 4, 5. Xception (L) and IF2bNet (R) Confusion Matrix

최소한의 노드로 경량화 된 최적화 모델도 기존의 전이학습 모델과 흡사한 성능지표를 보여주었다. 최적화 모델과 전이학습 모델의 세부적인 파라미터는 그림 5 와 같다. 전이학습에 사용된 Xception 모델의 파라미터는 약 2000 만, IF2bNet 모델의 파라미터는 약 270 만으로 로 기존 전이학습 모델에 약 0.1 배로 경량화 된 형태로도 동일한 성능을 도출하였다.

Table 2. Xception Model Parameters

Layer (type)	Shape	Parameters	Total Parameters
Xception	(7, 7, 2048)	20861480	20,867,627
GlobalAveragePooling2D	(2048)	0	
Dropout	(224, 224, 128)	0	
Dense	(224, 224, 128)	6147	

Table 3. IF2bNet Model Parameters

Layer (type)	Shape	Parameters	Total Parameters
InputLayer	(224, 224, 3)	0	2737795
Conv2D	(224, 224, 64)	1792	
BatchNormalization	(224, 224, 64)	256	
Conv2D	(224, 224, 128)	73856	
BatchNormalization	(224, 224, 128)	512	
Conv2D	(224, 224, 256)	295168	
BatchNormalization	(224, 224, 256)	1024	
Conv2D	(224, 224, 256)	590080	
BatchNormalization	(224, 224, 256)	1024	
Conv2D	(224, 224, 256)	590080	
BatchNormalization	(224, 224, 256)	1024	
Conv2D	(224, 224, 256)	590080	
BatchNormalization	(224, 224, 256)	1024	
Dense	(224, 224, 3)	771	

4. 결론 및 향후계획

본 연구에서는 화재 감지에 기여도가 적은 프로세스가 모델에 내장될 가능성이 있어, 모델의 복잡성을 증가시킬 수 있는 기존 전이학습 모델의 문제점을 보완하기 위해 Rule Extraction 을 기반으로 화재 감지에 최적화된 경량화 딥러닝 아키텍처를 제안하였다. 아키텍처는 다양한 딥러닝 및 설명 가능 인공지능 알고리즘을 분석한 결과를 토대로 구성되었으며, 학습 결과 파라미터를 약 0.1 배로 줄이면서도 동등한 성능을 유지하였다.

여전히 딥러닝 모델은 그 구조의 복잡성으로 인해, 내부 동작을 완전히 이해하기 어렵고, 따라서 프로세스를 명확하게 이해하기 위해서는 다양한 기술과 방법을 사용하여, 모델을 해석할 필요성이 있다. 이러한 해석을 통해, 동작 원리를 더욱 깊이 파악하고, 모델의 신뢰성을 높일 수 있는 방향으로 향후 연구를 기획할 예정이다.

참고문헌

- [1] 유수빈, 서승현, 유지나, 김재현. “CNN을 활용한실내화재 이미지 분석 알고리즘개발”. 한국통신학회. 2022
- [2] 최형규. “설명가능한 인공지능(Explainable AI : XAI) 연구 동향과 시사점”, 한국인공지능학회, pp4 ~ 13. 2021
- [3] Leilani H. Gilpin, David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter and Lalana Kagal. “Explaining Explanations: An Overview of Interpretability of Machine Learning”. Computer Science and Artificial Intelligence Laboratory Massachusetts Institute of Technology Cambridge. 2018
- [4] Macha, D., Kozielski, M., Wróbel, Ł., & Sikora, M. “RuleXAI—A package for rule-based explanations of machine learning model.”. Department of Computer Networks and Systems, Silesian University of Technology, Łukasiewicz Research Network. 2022
- [5] Francois Chollet.“Xception: Deep Learning with Depth wise Separable Convolutions”. Proceedings of the IEEE Conference on Computer Vision and pattern Recognition. 2017