

# Attention BiFPN 기반 화재 검출 딥러닝 알고리즘에 대한 연구

김대일\*, 안명수\*\*, 임혜연\*\*\*, 강대성\*\*\*\*

## A Study on Deep Learning Algorithm for Fire Detection based on Attention BiFPN

Daell Kim\*, MingShou An\*\*, Hye-Youn Lim\*\*\*, and Dae-Seong Kang\*\*\*\*

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.RS-2023-00247045)

### 요약

최근 빈번하게 발생하는 화재를 인공지능 및 머신러닝 기법을 활용하여 조기에 감지하고 효과적인 대응에 초점을 맞춘 연구들이 활발히 진행되고 있다. 화재 영상에서 복잡한 배경 및 환경 요인의 불확실성으로 인해 불꽃 또는 연기 영역을 더욱 정확하게 검출할 수 있는 기법이 요구된다. 본 논문에서는 Attention 기반 BiFPN 구조의 ResNet 알고리즘을 제안한다. 화재영역을 정확하게 검출하기 위하여 어텐션 메커니즘과 척도 불변 BiFPN 구조를 유기적으로 결합하여 객체와 배경정보를 분리함과 동시에 서로간의 연관성 유지를 통해 불필요한 정보를 제거하고, 필요한 정보만을 강조함으로써 불꽃 및 연기 영역검출 결과가 객체 크기와 상관없이 일관된 정확도를 유지하는 방법을 제안한다. 실험을 통해 불꽃은 98%의 정확도와 98.67%의 정밀도, 연기는 96.67%의 정확도와 96.67%의 정밀도의 화재 검출에 대한 평가 결과를 확인하였다.

### Abstract

Recent research has been actively focused on utilizing artificial intelligence and machine learning techniques to detect fires early and respond effectively, given the frequent occurrences of fires. There is a demand for techniques that can accurately detect flame or smoke areas in fire images due to the uncertainty of complex backgrounds and environmental factors. In this paper, to accurately detect fire areas, we propose a method that seamlessly combines attention mechanisms and scale-invariant Bidirectional Feature Pyramid Network(BiFPN) structures to separate object and background information while maintaining their correlations. It removes unnecessary information and emphasizes only the necessary information to maintain consistent accuracy in flame and smoke areas regardless of their sizes. Evaluation results showed that fires could be detected with 98% accuracy and 98.67% precision for flames and 96.67% accuracy and 96.67% precision for smoke.

### Keywords

fire detection, neural network, attention mechanism, bi-directional pyramid, feature pyramid network

- \* 특허청 수석심사관  
- ORCID: <https://orcid.org/0009-0000-5771-9188>  
\*\* 주식회사 스마트뱅크 연구부소장  
- ORCID: <https://orcid.org/0000-0002-1180-8916>  
\*\*\* 동아대학교 전자공학과 조교수  
- ORCID: <https://orcid.org/0000-0002-5189-461X>  
\*\*\*\* 동아대학교 전자공학과 교수(교신저자)  
- ORCID: <https://orcid.org/0000-0003-0186-2430>

- Received: May 17, 2024, Revised: Aug. 27, 2024, Accepted: Aug. 30, 2024  
· Corresponding Author: Dae-Seong Kang  
Dept. of Dong-A University, 37 NaKdong-Daero 550, beon-gil saha-gu, Busan, Korea  
Tel.: +82-51-200-7710, Email: [dskang@dau.ac.kr](mailto:dskang@dau.ac.kr)

## 1. 서론

화재는 활성 산소, 연소물질 및 연소열에 의해 발생하는 열적 반응으로, 불꽃, 연기 및 열에 의해 형성된다. 화재는 인간과 동물의 안전에 심각한 위협을 초래하며, 건물 등 재물의 손상, 환경 오염 및 생명을 위협할 수 있다. 화재의 원인은 인간이 의도하지 않은 자연변화의 요소로 인한 발생, 또는 인간의 고의적 행동으로 인해 발생할 수 있다. 이러한 화재는 생태학적 피해뿐 아니라 종종 인간의 생명, 동물의 생명, 건강 및 재산에 치명적인 영향을 미친다. 경제발전과 더불어 난방 및 취사 연료의 형태가 전기, 가스, 유류 등으로 변화되었고 에너지의 다양화와 각종 위험물질의 증가에 따른 화재 발생은 날로 증가하고 있으며, 최근 발생한 화재를 원인별로 보면 전기, 담배, 방화, 불장난, 불티, 가스, 유류 등이 주요 원인으로 되어있다[1].

현재, 화재 인식 기술은 고도로 발전하고 있으며, 인공지능과 센서 기술의 발전으로 높은 정확성과 빠른 응답 시간을 갖추고 있다. 다양한 기술들이 화재를 조기에 감지하고 효과적으로 대응하는 데 기여하고 있다. 특히 인공지능 및 딥러닝 기술의 발전이 심화함에 따라 딥러닝 및 컴퓨터 비전 기술을 활용하여 화재의 시각적 특성을 자동으로 인식하는 시스템이 개발되고 있다. 이러한 기술은 높은 정확성과 낮은 오류율을 보여주며, 화재 조기 경보 시스템에서 중요한 역할을 하고 있다. 이미지 및 비디오를 대상으로 하는 화재 검출 알고리즘은 전통적인 화재탐지기가 화재 근처에 접근해야 만이 작동하는 단점을 효과적으로 극복하였지만, 영상의 복잡한 배경 및 환경 요인의 불확실성으로 인해 여전히 극복해야 할 다양한 과제들이 존재한다[2][3].

이를 해결하기 위하여 최근에는 Attention 메커니즘[4] 기반의 신경망 구조를 활용한 방법들이 제안되었다. Attention 메커니즘은 입력 영상에서의 특정 부분을 강조해서 보는 형태로 사용되며, 일종의 자원 분배 방안으로 정보 과부하 문제를 해결하는 수단으로 사용될 수 있다. 이 때문에 많은 연구자들은 컴퓨터 비전 분야에서 현저성을 바탕으로 하는 시각 주의력 모델을 제시하고 국부적인 저차원 시각

적 특징을 추출하여 현저성 관심 영역을 검출하는데 활용하였다. 이처럼 Attention 메커니즘은 신경망 구조에서 쉽게 접할 수 있는 구성 부분이 되었고, 자막 생성, 텍스트 분류, 기계번역, 영상 분석 등에서 널리 사용되고 있다[5].

본 연구는 다음과 같은 구성으로 진행된다. 2장에서는 Attention 메커니즘, 피라미드 특징추출[6][7] 등 관련 이론에 대해서 기술한다. 3장에서는 제안하는 화재 검출을 위한 Attention 메커니즘 기반 양방향 피라미드 구조의 네트워크에 대해 소개한다. 4장에서 제안하는 방법의 유효성과 성능을 입증하기 위한 평가결과를 기술하고, 5장에서 결론을 맺는다.

## II. 관련 이론

### 2.1 Attention 메커니즘

딥러닝에서 Attention 메커니즘은 NLP[8]를 비롯한 기계 번역 분야에서 사용이 활발하며, 컴퓨터 비전에서는 컨볼루션 뉴럴 네트워크 아키텍처를 보완하거나 대체하여 사용하는 방안을 중심으로 연구되고 있다. 컨볼루션 뉴럴 네트워크의 Attention 개념은 대부분 VQA(Visual Question Answering)[9]이나 IC(Image Captioning)[10]처럼 멀티 모달간의 관계를 이용한 특징 선택에 많이 사용되었다. 대표적인 방법으로는 논문[2]에서와 같이 영상의 어떤 부분을 관심영역으로 보고 문장을 생성했는지 효과적으로 시각화 할 수 있는 방법을 제안하였다. 즉, 이 관심 영역에 대하여 Attention이 적용된 것이다. 이와 같이 이미지 분류, 이미지 검출 등 입력 영상에 따라 어떤 영역을 관심 깊게 볼 것인가에 따라 Attention을 적용 가능하다. 최근에 Residual 기술과 접목하여 이미지 관련 분야에서도 다양한 네트워크가 연구되고 있고 모델의 구성요소로 Attention이 발전하고 있다. Attention은 상호보완적 관계를 이루며 딥러닝의 성능 향상에 유의미한 결과들을 보여주고 있다. ImageNet classification, COCO detection, VOC detection와 같은 여러 벤치마크[11] 및 ResNet, WideResNet, ResNext, MobileNet와 같은 여러 모델[12]에서 그 성능을 입증하였다.

우선, Attention 모듈의 구조를 보면 시그모이드로 일반화된 어텐션 맵으로부터 컨볼루션 특징을 추출하여 element-wise 곱해진다. Attention의 성능은 적은 연산량으로도 큰 성능 향상을 이끌어내는 것으로 중요한 역할을 맡고 있다. attention map은 channel-wise / spatial-wise로 분해하여 계산한다. 대표적인 Attention 모듈로 BAM(Bottleneck Attention Module)과 CBAM(Convolutional Block Attention Module)이 있다 [13]. 이것들은 아주 단순한 pooling 및 convolution으로 이루어져 있다. self-attention을 모듈화하여 어떤 CNN에도 쉽게 연결할 수 있게 하였다. 또한, 기존 네트워크와 함께 end-to-end training이 가능하다. 그림 1은 BAM의 구조가 추가된 일반적인 딥러닝 모델이다.

후속 연구로 진행된 CBAM은 모듈의 pooling, spatial과 channel attention의 결합 방식의 변형으로 BAM보다 더 높은 성능을 보여준다. 그림 2는 CBAM의 개념도이다.

Attention은 딥러닝에서 주로 cross-modal feature selection 등의 역할을 하며 다양하게 활용되었다. 하지만 최근에 self-attention이 일반적인 CNN에서도 연결성이 높아 여러 활용 방안이 연구되었다. 그 대표적인 모듈이 BAM과 CBAM이며, 적은 파라미터의 양으로도 기존의 모델에 적용하였을 때 유의미한 성능 향상을 보여주기 때문에 활용 가치가 매우 높다.

## 2.2 피라미드 구조

FPN(Feature Pyramid Network)는 CVPR 2017에 발표된 'Feature Pyramid Networks for Object Detection' 논문에서 제안되었으며, 컴퓨터 비전의 다양한 분야에서 사용되고 있다.

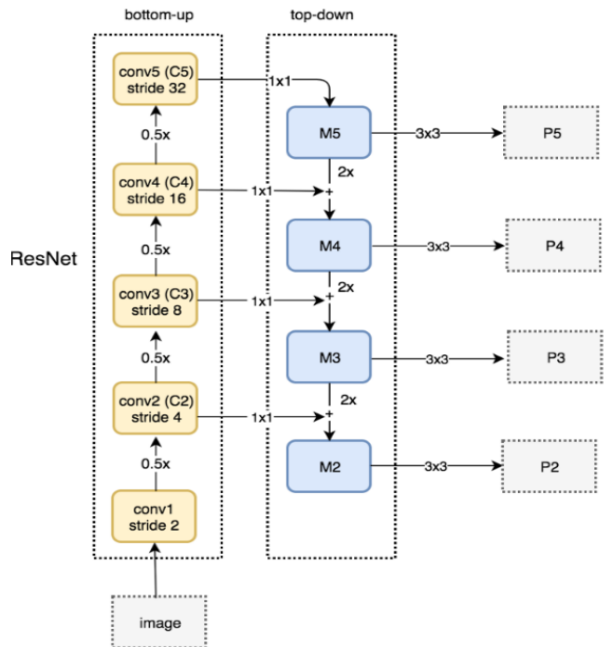


그림 3. ResNet에서의 FPN 구조  
Fig. 3. Structure of FPN in ResNet

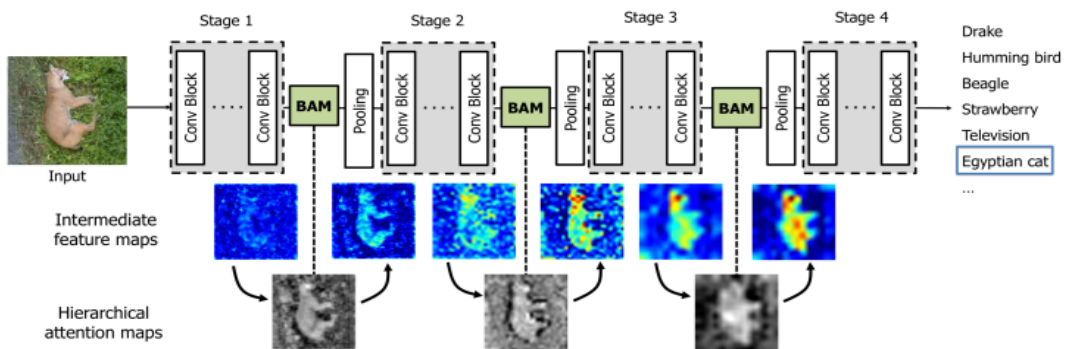


그림 1. BAM 딥러닝 모델  
Fig. 1. Deep learning model with BAM

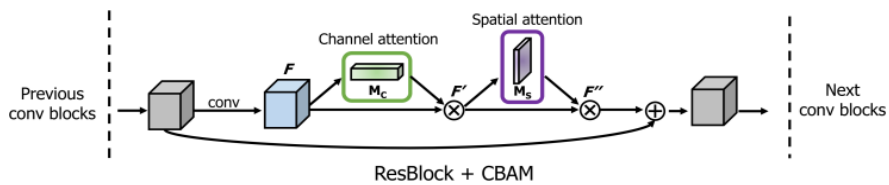


그림 2. CBAM 개념도  
Fig. 2. Conceptual map of CBAM

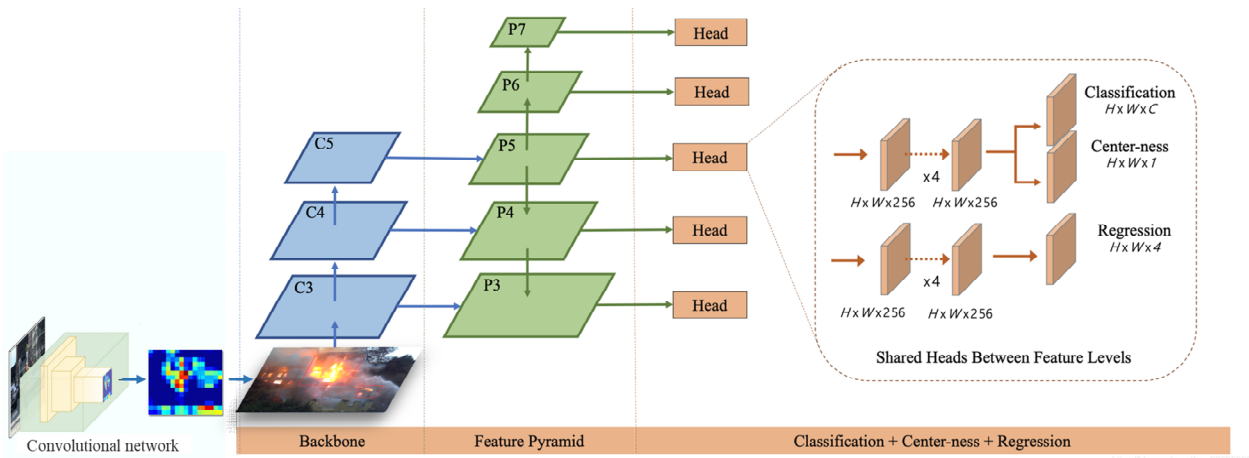


그림 4. Fast R-CNN을 위한 FPN  
Fig. 4. FPN for fast R-CNN

이는 객체 검출의 Upsampling과 Skip Connection으로 Scale-Invariant 문제를 해결하고 이미지의 해상도를 높여 정확도를 올리는 방식이다. 즉, 저해상도의 이미지와 고해상도의 이미지를 더하여 압축된 특징 맵에 해상도를 보완하는 방식이다. 임의의 크기의 단일 스케일 이미지를 입력으로 하며, backbone은 ResNets를 사용한다.

Fast R-CNN에서는 FPN을 통하여 다양한 크기의 feature map을 추출하고, 가장 해상도가 높은 feature map을 이용하여 Region Proposal 단계를 수행한다.

### III. 제안하는 방법

컴퓨터 비전 기술과 딥러닝 기술의 발전에 따라 인간의 시각 시스템을 모방하여 구성하는 방법들이 주목받고 있다. 즉, 입력장치로부터 입력되는 수많은 시각 정보 중 현실점에서 가장 유용하다고 생각되는 정보를 인간의 상향식 주의 시각에 기반하여 선택하는 시각 정보 선택 기법에 의한 딥러닝 신경망 구현 방법이다. 이러한 방법으로 현재 가장 많이 사용되는 메커니즘은 Attention 메커니즘에 의한 주의력 강화 방식이다. 하지만 이러한 학습 방법은 여전히 실시간 처리 및 검출 정확도 등 면에서 개선이 필요하다. 신속하고 정확하게 화재 영역을 검출하기 위하여 본 논문에서는 Attention 메커니즘 기반의 양방향(Bi-directional) 피라미드 특징맵 구조[14]의 ResNet 알고리즘을 제안한다. 제안하는 알고리즘의 구조도는 그림 5에서 보여주며, 이는 특징 추출

네트워크(backbone), short connection 네트워크, 양방향 피라미드 특징 네트워크로 구성된다.

그림 5에서처럼 특징 추출 네트워크는 상하 라인이 대칭되는 구조이다. 상하 라인은 short connection에 의해 의미(semantic)정보를 공유하며, 인간의 시각 계통에서 데이터를 교류 및 처리하는 방식을 모방하여 시물레이션하며, 결과를 출력하기 전에 양방향 피라미드 구조를 도입하여 얇은 층의 텍스처(texture) 정보와 깊은 층의 텍스처 정보를 결합하여 다단계 멀티 스케일 객체 인식을 구현한다. 본 네트워크 모델의 출력은 두 프레임의 인진화 saliency 이미지이다. 위 라인 네트워크를 통해 출력되는 결과는 화재 영역에 대한 saliency 이미지이며, 그림에서 흰색 부분은 화재 영역이고, 검은색 부분은 배경이다. 아래 라인 네트워크에서 출력되는 결과는 배경 영역에 대한 saliency 이미지이며, 흰색 부분은 배경이고, 검은색 부분은 화재 영역이다.

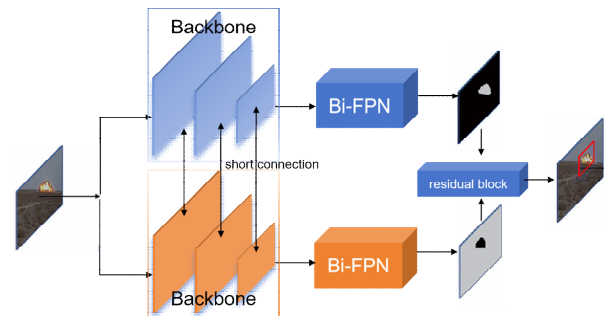


그림 5. 제안하는 화재 검출 모델  
Fig. 5. Proposed model of fire detection



### 3.1 특징 추출 네트워크

모델의 층이 깊어질수록 ResNet의 residual block 구조는 퇴화 문제를 잘 해결할 수 있기 때문에 본 논문에서 제안하는 방법은 병렬 ResNet 구조의 교차 연결을 이용하여 인간의 두 눈의 기능을 시뮬레이션하여 각각 객체와 배경 정보를 학습한다. 특징 추출 네트워크 구조는 다음 그림에서 보여준다.

그림 6에서, 특징 추출 네트워크의 베이스 라인은 2개의 병렬 교차하는 ResNet 네트워크를 사용한다. ResNet 네트워크는 다층의 residual block으로 구성되었으며, 다층의 멀티 스케일 특징을 추출할 수 있다. 얇은 층의 특징은 컨볼루션 처리 회수가 적으며, 완전한 공간 정보를 보류하여 saliency 객체의 에지를 검출하는데 효과적이다. 그러나 고도의 컨볼루션 연산을 진행한 후 얻은 깊은 층의 특징들은 풍부한 의미 정보를 가지고 있어 현저성 객체를 효율적으로 검출하고, 네트워크의 일반화 능력을 강화할 수 있다. 인간의 시각 시스템은 극히 강한 정보 교류 및 데이터 처리 능력을 갖고 있다. 본 논문에서 제안하는 알고리즘은 short connection을 통해 정보 교류 능력을 시뮬레이션하여 에지 정보와 의미 정보를 강화시키는 목적을 달성할 수 있다. 이를 통해 출력되는 강화된 특징맵  $C_2, C_3, C_4, C_5$ 를 각각 양방향 피라미드 네트워크에 입력으로 사용되어 화재 검출의 멀티 스케일 예측 능력을 향상시킬 수 있다.

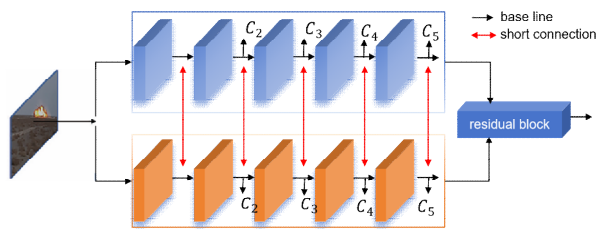


그림 6. 특징 추출 네트워크  
Fig. 6. Feature extraction network

### 3.2 Short connection

Short connection 네트워크에서는 우선 상위 라인에서의 특징에 dilated 컨볼루션 연산을 통해 특징의 receptive field를 확대하여 화재 영역의 에지 정보를

저장한다. 다음으로 Attention 메커니즘 구조(CSA)를 도입하여 듀얼 채널의 ResNet으로 하여금 객체의 대량의 정보를 갖는 채널과 공간적 정보에 포커싱한다. 일반적인 컨볼루션과 달리, dilated 컨볼루션은 dilated 확률 지표를 증가, 즉 컨볼루션 연산중에 0을 채워넣는 과정을 추가하여 동일 해상도 특징맵에서 컨볼루션의 receptive field를 확대하고 정보 손실을 효율적으로 해결할 수 있다.

### 3.3 Bi-FPN 구조

전통적인 FPN(Feature Pyramid Network) 구조는 컨볼루션 네트워크에서 더 얇은, 즉 입력층에 보다 가까울수록 특징맵은 높은 해상도(High resolution)를 가지며, 가장자리, 곡선 등과 같은 저수준 특징(Low-level feature)을 보유하고 있다. 반대로 더 깊은 층에서 얻을 수 있는 특징맵은 낮은 해상도(Low resolution)를 가지며, 질감과 물체의 일부분 등 클래스를 추론할 수 있는 고수준 특징(High-level feature)을 가지고 있다. 하지만 깊은 층의 특징에는 색상, 텍스처, 공간 위치 정보 등 객체의 얇은 층의 특징을 가지고 있지 않아 위치 정보 검출 정확도가 떨어질 수 있다. 이와 같은 기존의 FPN 구조의 단점을 보완하기 위하여 본 논문에서는 기존의 FPN의 top-down 구조에 down-top 구조를 융합하여 4장의 라벨링 샘플과 scale이 같은 [0, 1]로 정규화된 예측 특징맵을 출력한다.

제안하는 방법은 다음 그림에서와 같이 구성되며, 그림 7에서 TCN(Transpose Convolution Network)은 deconvolution을 통해 특징맵의 upsampling을 수행하며, CD(Channel Downsample)을 통해 입력된 특징맵의 파라미터를 반으로 줄여준다.

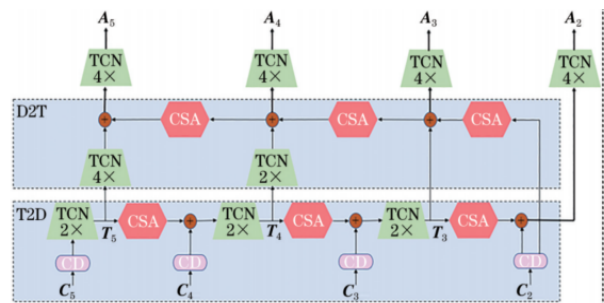


그림 7. Bi-FPN 구조  
Fig. 7. Structure of Bi-FPN

그림 8에서와 같이 Bi-FPN는  $C_2, C_3, C_4, C_5$  특징맵을 입력하여 top-down 컨볼루션 네트워크의 forward 연산을 수행하여 높은 층의 많은 정보를 가진 특징맵을 2배로 upsampling한  $T_5, T_4, T_3$ 을 얻는다. 유효 채널과 공간 정보에 대한 학습을 강화하기 위하여  $T_5, T_4, T_3$ 에 다시 Attention 연산을 하여 인접한 2층의 특징맵을 융합한다. top-down 네트워크의 구조는 FPN 네트워크중의 구조와 같은 원리로 구성된다. 다만 FPN의 구조는 깊은 층에서 얇은 층으로의 정보 전달로 기존 층의 특징과 깊은 층의 특징만 포함되므로, 보다 정확한 검출 결과를 획득하기 힘들다. 제안하는 방법에서는 위 방법에 down-top의 구조를 결합하여 얇은 층의 보다 많은 텍스처 및 공간 위치 정보를 포함하고 있는 특징맵  $C_2$ 에서부터 시작하여 Attention 모듈을 통과하여 더 깊은 층의 특징과 결합하여 얇은 층에서부터 깊은 층으로의 특징 정보를 축적한다. 따라서 본 논문에서 제안하는 Bi-FPN 구조는 각각의 출력층에서 완전한 심층 의미 정보와 얇은 층의 텍스처 위치 정보를 이용하여 보다 정확한 객체 분할 효과를 달성할 수 있다. 이렇게 예측된  $A_2, A_3, A_4, A_5$ 을 가중평균하면 최종 전경예측 결과와 배경 예측 결과를 획득할 수 있다.

## IV. 실험 결과

### 4.1 실험 데이터셋

본 논문에서 제안한 모델의 성능을 검증하기 위하여 AI Hub에서 제공하는 화재 발생 예측 영상 데이터셋[15]을 이용하여 실험을 진행한다. 상기 데이터 셋은 화재 발생 전에 발생하는 연기를 촬영한 동영상으로부터 화재 발생을 사전 예측하는 AI 기술 개발을 위한 화재 발생 예측 AI학습용 데이터이다. 구축된 데이터는 화재현 이미지 740,000장, 유사현 이미지 660,000장, 무관현 이미지 300,000장으로 구성되었다. 이 중 70%는 학습데이터, 15%는 테스트데이터, 15%는 검증데이터로 구분하여 모델을 학습시켰다.

### 4.2 성능지표

본 논문에서 제안하는 방법을 통한 불꽃 및 연기에 대한 검출 결과를 객관적으로 평가하기 위하여 정확도(Accuracy), 정밀도(Precision), 재현율(Recall)을 통해 평가한다. 정확도는 가장 직관적으로 모델의 성능을 나타낼 수 있는 평가지표로서, 객체를 객체라고 검출한 경우, 객체가 아닌 것을 아닌 것으로 검출한 경우를 고려한 지표이다. 정밀도는 객체를 검출한 결과가 실제 객체와 일치하는지를 나타내는 성능지표이고, 재현율은 실제 객체를 기준으로 얼마나 검출했는지에 대한 검출율을 나타내는 성능지표이다.

$$AC = \frac{T_P + T_N}{T_P + F_N + T_N + F_P} \quad (1)$$

$$P = \frac{T_P}{T_P + F_P} \quad (2)$$

$$R = \frac{T_P}{T_P + F_N} \quad (3)$$

상기 수식에서  $AC$ 는 정확도,  $P$ 는 정밀도,  $R$ 은 재현율,  $T_P$ 는 화재(불꽃, 연기)에 대해서 정확히 검출한 수(True positive),  $T_N$ 는 화재가 아닌 것을 아니라고 검출한 수(True negative),  $F_P$ 는 오 검출된 화재 수(False positive),  $F_N$ 는 미검출된 화재 수(False negative)를 나타낸다.

### 4.3 실험 결과 분석

본 논문에서 제안한 모델의 성능을 검증하기 위하여 실내, 실외 불꽃 발생 영상(1080p) 각각 50장씩 선별하여 정확하게 추론하는지에 대한 판별실험을 하였다. 또한, 화재와 유사한 비화재 영상 각각 50개를 통해 모델의 오검출율에 대한 실험도 병행하여 진행하였다. 연기 발생 영상에 대해서도 똑같은 방법으로 실험을 진행했다. 제안하는 방법의 구동 환경은 Intel Core i7, RAM 16GB, GTX2080Ti GPU를 이용하여 수행하였다.

표 1은 3종의 화재 영상으로부터 불꽃 영역을 검출하여 제안하는 방법이 불꽃을 정확하게 판단한 횟수와 비화재 영상으로부터 불꽃으로 검출한 횟수를 기록하였다. 영상 1종당 모델이 추론한 장면들 중에서 임의로 50장씩 선별하여 화재 검출 정확도를 표에서 나타내었다.

마찬가지로, 표 2는 화재영상과 비화재영상에서 연기를 검출한 결과로, 각 영상에서 50장씩 선별하여 검출 정확도를 기록하였다. 실험과정에서 기록한 값들을 통해 식 (1)~(3)을 이용하여 평가지표들을 구하여 표 3과 같은 결과를 얻었다. 불꽃 검출에 대한 정확도 98%, 정밀도 98.67%, 재현율 97.33%, 기록하였으며, 연기 검출에 대한 정확도는 96.67%, 정밀도 96.67%, 재현율 96.67%를 기록하였다.

표 1. 불꽃 및 비불꽃 비디오에 대한 검출 결과

Table 1. Results of flame detection for flame and non-flame videos

Flame video	Detection ( $T_p$ )	Non-flame video	Detection ( $F_p$ )
Video#1	49/50	Video#1'	0/50
Video#2	50/50	Video#2'	1/50
Video#3	49/50	Video#3'	3/50

표 2. 연기 및 비연기 비디오에 대한 검출 결과

Table 2. Results of flame detection for smoke and non-smoke videos

Smoke video	Detection ( $T_p$ )	Non-smoke video	Detection ( $F_p$ )
Video#4	47/50	Video#4'	3/50
Video#5	49/50	Video#5'	0/50
Video#6	49/50	Video#6'	2/50

표 3. 평가 지표 결과

Table 3. Evaluation results

Videos	Accuracy	Precision	Recall
Flame video	98%	98.67%	97.33%
Smoke video	96.67%	96.67%	96.67%

그림 8에서는 본 논문에서 제안하는 방법에 의한 불꽃 및 연기에 대한 검출 결과를 보여준다, 그림에서 불꽃 영역에 대한 검출 바운딩박스는 빨간색으로, 연기 영역에 대한 검출 결과는 노란색 바운딩 박스로 표기하였다.



그림 8. 검출 결과 예시

Fig. 8. Examples of detection result

## V. 결 론

본 논문에서 영상의 복잡한 배경 및 환경 요인의 불확실성으로 인해 불꽃 또는 연기 영역을 정확하게 검출하지 못하거나 오검출 하는 문제점을 해결하고자 화재 영역을 돋보이게 할 수 있는 어텐션 메커니즘과 이미지의 해상도를 높여 척도 불변의 FPN 구조를 적용하여 정확도를 올리는 방법을 제안하였다. 특징 추출 네트워크로는 모델의 퇴화 문제를 잘 해결할 수 있는 병렬 교차 구조 연결 방식의 ResNet을 이용하여 객체와 배경 정보를 분리하여 학습하고 Short connection 네트워크를 통해 서로간의 연관성을 유지시켰다. 입력층에 가까울수록 특징맵이 높은 해상도를 가지고, 층이 깊어질수록 낮은 해상도를 가지는 FPN을 기반으로 할 경우 위치 정보 검출 정확도가 떨어질 수 있다는 구조적 단점을 보완하기 위하여 본 논문에서는 기존의 FPN의 top-down 구조에 down-top 구조를 융합하여 사용하였다.

본 논문에서 제안한 모델의 성능을 검증하기 위하여 AI-Hub의 화재 발생 예측 영상 데이터셋을 이용하여 학습하는 과정에서 높은 정확도에 수렴하였으며, 학습된 모델을 불꽃 영상과 연기 영상에 적용한 검출결과에서도 높은 정확도와 정밀도 및 재현율을 보였다.

향후 지속적인 연구를 통해 데이터셋의 보강, 모델의 검출 결과에 대한 확률적 파라미터 값 등, 추가 조건 조합을 통하여 정확도와 정밀도를 향상시킬 수 있는 방향으로 추가적인 연구를 수행할 계획이다.

## References

- [1] National Fire Agency, "Statistical Yearbook of Fire Department 2018", National Fire Agency Information Statistics Office, pp. 85-94, Sep. 2018.
- [2] F. M. Talaat and H. ZainEldin, "An improved fire detection approach based on YOLO-v8 for smart cities", *Neural Computing and Applications*, Vol. 35, No. 28, pp. 20939-20954, Oct. 2023. <https://doi.org/10.1007/s00521-023-08809-1>.
- [3] K. Avazov, M. K. Jamil, B. Muminov, A. B. Abdusalomov, and Y. I. Cho, "Fire detection and notification method in ship areas using deep learning and computer vision approaches", *Sensors*, Vol. 23, No. 16, pp. 7078, Aug. 2023.
- [4] W. Yin, H. Schütze, B. Xiang, and B. Zhou, "Abcnm: Attention-based convolutional neural network for modeling sentence pairs", *Transactions of the Association for computational linguistics*, Vol. 4, pp. 259-272, Jun. 2016. [https://doi.org/10.1162/tacl\\_a\\_00097](https://doi.org/10.1162/tacl_a_00097).
- [5] J. Han and D.-S. Kang, "U-OLANet : Low-Resolution One-Layer Attention Neural Networks for Image Denoising", *Journal of KIIT*, Vol. 19, No. 9, pp. 1-9, Sep. 2021. <http://dx.doi.org/10.14801/jkiit.2021.19.9.1>.
- [6] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection", In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (CVPR2017), Honolulu Hawaii, pp. 2117-2125, Jul. 2017.
- [7] S. W. Kim, H. K. Kook, J. Y. Sun, M. C. Kang, and S. J. Ko, "Parallel feature pyramid network for object detection", In *Proceedings of the European conference on computer vision* (ECCV2018), Munich, Germany, Vol. 16, pp. 234-250, Sep. 2018.
- [8] D. Hu, "An introductory survey on attention mechanisms in NLP problems", In *Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys)*, Springer International Publishing, Vol. 1038, pp. 432-448, Aug. 2019. [https://doi.org/10.1007/978-3-030-29513-4\\_31](https://doi.org/10.1007/978-3-030-29513-4_31).
- [9] S. Antol, A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. L. Zitnick, and D. Parikh, "Vqa: Visual question answering", In *Proceedings of the IEEE international conference on computer vision* (CVPR2015), Santiago, Chile, pp. 2425-2433, Dec. 2015.
- [10] Q. You, H. Jin, Z. Wang, C. Fang, and J. Luo, "Image captioning with semantic attention", In *Proceedings of the IEEE conference on computer vision and pattern recognition* (CVPR2016), Las Vegas Nevada, pp. 4651-4659, Jul. 2016.
- [11] E. Xie, J. Ding, W. Wang, X. Zhan, H. Xu, P. Sun, and P. Luo, "Detco: Unsupervised contrastive learning for object detection", In *Proceedings of the IEEE/CVF international conference on computer vision* (ICCV2021), pp. 8392-8401, Oct. 2021.
- [12] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks", *Artificial intelligence review*, Vol. 53, pp. 5455-5516, Apr. 2020. <https://doi.org/10.1007/s10462-020-09825-6>.
- [13] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module", In *Proceedings of the European conference on computer vision* (ECCV2018), Munich, Germany, pp. 3-19, Sep. 2018.
- [14] H. Guo and D. Gu, "Closely arranged inshore ship detection using a bi-directional attention feature pyramid network", *International Journal of Remote Sensing*, Vol. 44, No. 22, pp. 7106-7125,

Nov. 2023. <https://doi.org/10.1080/01431161.2023.2277166>.

[15] AI Hub, <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=176>  
[accessed : Dec. 19, 2023]

저자소개

강 대 성 (Dae-Seong Kang)



1994년 5월 : Texas A&M 대학교  
전자공학과(공학박사)  
1995년 9월 ~ 현재 : 동아대학교  
전자공학과 교수  
관심분야 : 영상처리, 패턴인식,  
인공지능

김 대 일 (Daell Kim)



2004년 2월 : 동아대학교  
전자공학과(공학석사)  
2007년 3월 ~ 2009년 2월 :  
동아대학교 전자공학과(박사과정  
수료)  
2014년 12월 ~ 현재 : 특허청  
심사관

관심분야 : 영상처리, 패턴인식, 인공지능

안 명 수 (MingShou An)



2016년 2월 : 동아대학교  
전자공학과(공학박사)  
2017년 9월 ~ 2024년 2월 :  
서안공업대학교 강사  
2024년 3월 ~ 현재 : 주식회사  
스마트뱅크 연구소 부소장  
관심분야 : 영상처리, 패턴인식,

인공지능, 빅데이터

임 혜 연 (Hye-Youn Lim)



2013년 2월 : 동아대학교  
전자공학과(공학박사)  
2013년 3월 ~ 현재 : 동아대학교  
전자공학과 조교수  
관심분야 : 영상처리, 패턴인식,  
인공지능