

# CBAM을 적용한 YOLOv4-P5 기반 화재 감지 시스템

유다연, 신연순\*

동국대학교

yoo121603@naver.com, \*ysshin@dongguk.edu

## Fire Detection System Based on YOLOv4-P5 with CBAM

Yu Da yeon, Shin Youn Soon\*

Dongguk Univ.

### 요약

화재 발생 시 재산 및 인명 피해를 일으킬 가능성이 크기 때문에, 화재를 감지하고 대비하는 것은 매우 중요하다. 본 논문에서는 화재 감지 시스템의 성능 향상을 위해 CBAM을 적용한 YOLOv4-P5 모델을 제안한다. 기존의 YOLOv4 모델을 네트워크 Scaling을 통해 확장한 YOLOv4-P5 모델에 이미지의 공간 및 채널 정보를 모두 고려하여 이미지의 특징을 추출하는 CBAM을 적용하여 화재 감지 precision을 향상시킬 수 있다. 실제 실험을 통해 본 논문에서 제안한 YOLOv4-P5-CBAM 모델이 기존의 YOLOv4 모델보다 14.45% 향상된 성능을 보이는 것을 확인하였다.

### I. 서론

화재는 사회적 손실과 복구비용이 매우 크고 인명 피해가 날 가능성이 높기 때문에 초기에 발견하는 것이 중요하다. 따라서 화재를 정확하게 신속하게 감지하고, 확대 상황을 예측하여 화재에 대비할 수 있어야 한다.

본 논문은 기존의 YOLOv4[1] 모델의 성능을 향상시켜 정확한 화재 감지를 목표로 한다. 모델의 성능 개선을 위해 CBAM(Convolutional Block Attention Module)[2]을 사용하였다. 이를 사용하면 이미지의 공간은 물론 채널 정보까지 모두 고려하여 이미지 특징을 추출하여 precision을 향상시킬 수 있다.

### II. 본론

#### 2.1 YOLOv4

YOLOv4는 CNN을 기반으로 한 객체 검출 딥러닝 모델로, 입력 이미지에서 객체를 검출하는 과정을 거쳐 객체의 위치와 클래스를 예측한다. Backbone Network(CSPDarknet-53)에서 입력 이미지의 특징을 추출하고 다양한 크기의 특징 맵을 생성한다. Neck Network에서는 Path Aggregation Network와 Spatial Pyramid Pooling Network 등을 사용하여 특징 맵을 연결하고 처리한다. Head Network에서는 처리된 특징 맵을 입력으로 받아서 객체 검출을 수행한다. 먼저 입력 이미지를 일정한 크기로 분할한 후, 분할된 이미지에 대해 앵커(Anchor) 박스라는 미리 지정된 사각형들을 적용하여 객체의 위치를 예측하고, 이를 토대로 객체의 클래스와 확률을 예측한다.

Scaled-YOLOv4[3]는 YOLOv4에 네트워크 scaling을 적용한 것으로, scaling의 상한과 하한을 바탕으로 한 YOLOv4-large, YOLOv4-tiny가 있다.

#### 2.2 CBAM

CBAM은 이미지 분류 및 객체 감지 모델에서 성능 향상을 위해 사용되는 딥러닝 기술 중 하나로, 이미지의 공간과 채널 정보를 모두 고려하여

이미지 특징을 추출하며 Channel Attention Module(CAM)과 Spatial Attention Module(SAM)로 이루어져 있다. 먼저 CAM에서는 입력 받은 특징 맵을 Max Pooling과 Average Pooling 한 후 Multi-layer Perceptron 레이어를 통과한다. 통과 후 나온 벡터들을 더한 후 시그모이드 함수를 통과하면 채널 어텐션이 계산된다. 이는 서로 다른 특징 맵들 간의 관계를 고려하여 어떤 특징 맵이 중요한지에 대한 확률이다. 이를 입력 특징 맵 F에 곱셈하여 F'를 생성한다. 그 후 SAM을 통과한다. 이전 채널 어텐션을 통해 새롭게 구한 F' 특징 맵을 입력 받아 채널 축에 대해 Max Pooling과 Average Pooling 한 후 연결한다. 그 후 픽셀 중 어느 부위에 집중할지를 선택하기 위해 곱셈곱 레이어를 통과하고 시그모이드 함수를 적용해서 spatial 어텐션을 구한다. 이를 입력 받은 F' 특징 맵에 곱셈하여 새로운 특징 맵 F''을 생성한다.

#### 2.3 제안 모델

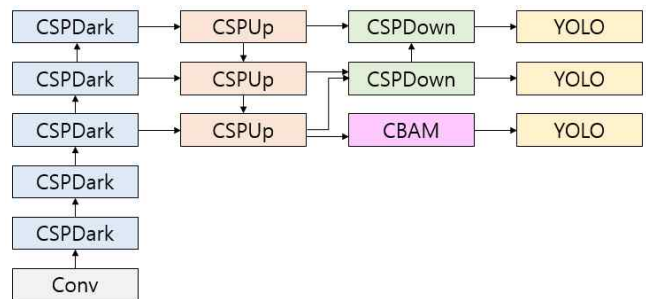


그림 1. Architecture of YOLOv4-P5-CBAM

본 논문에서 제안하는 YOLOv4-P5-CBAM 모델은 YOLOv4를 확장한 YOLOv4-P5 모델에 그림 1과 같이 CBAM을 추가한 것이다. YOLOv4-P5는 YOLOv4-large 중 하나로, CSP가 적용되고 input size와 stage의 수에 대해 복합 scaling을 수행하는 모델이다. CBAM은 이미지의 색상 정보를 통해 특징을 추출하는 CAM과 공간 정보를 통해 특징을

추출하는 SAM이 순서대로 결합한 형태이다. 이를 CAM layer와 SAM layer로 구성하여, 첫 번째 YOLO layer 앞에 추가하였다.

## 2.4 실험

제안한 모델의 성능을 평가하기 위해 동일한 조건으로 4개의 모델에 대해 실험을 진행하였다. Google Colab에서 진행하였고, 사용된 하이퍼파라미터는 Table 1과 같다. 학습 데이터는 github에서 다운받은 450개의 화재 및 비화재 이미지를 이용하였다.

Table 1. Hyperparameter

Configuration	Value
Batch	64
Subdivision	16
Momentum	0.949
Decay	0.0005
Learning Rate	0.001
Activation	Mish

## 2.5 결과

모델의 성능 평가를 위해 AP50 값을 이용하였다. YOLOv4-CBAM 모델은 YOLOv4에 CBAM을 추가한 것으로, 제안한 모델과 동일하게 첫 번째 YOLO layer 앞에 CBAM을 추가하였다. YOLOv4 모델과 YOLOv4-P5 모델에 CBAM을 추가했을 때 Table 2에서 보이는 것처럼 AP50 값이 각각 44.15%에서 47.77%로 3.62%p 증가, 48.63%에서 50.53%로 1.9%p 증가한 것을 확인할 수 있다.

Table 2. Results

Model	Backbone	AP50
YOLOv4	CSPDarknet-53	44.15
YOLOv4-CBAM	CSPDarknet-53	47.77
YOLOv4-P5	CSP-P5	48.63
<b>YOLOv5-P5-CBAM</b>	<b>CSP-P5</b>	<b>50.53</b>

## III. 결론

본 논문에서는 YOLOv4에 CBAM을 적용하여 화재를 감지하는 모델 (YOLOv4-P5-CBAM)을 제안했다. YOLOv4-P5-CBAM은 50.53%의 AP50 값을 보이며 기존의 YOLOv4 모델보다 14.45% 향상된 성능을 보였다. 제안한 모델을 화재 감지 시스템에 적용한다면 화재를 더욱 정확하게 감지하여 재산 및 인명 피해를 줄일 수 있을 것이다.

## ACKNOWLEDGMENT

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업 지원을 받아 수행되었음. (2023-0-00049)

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 인공지능융합 혁신인재양성사업(IITP-2023-RS-2023-00254592)의 연구결과로 수행되었음.

## 참고 문헌

- [1] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934, 2020.
- [2] Woo, Sanghyun, et al. "Cbam: Convolutional block attention module." Proceedings of the European conference on computer vision (ECCV) pp. 3 - 19, 2018.
- [3] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "Scaled-yolov4: Scaling cross stage partial network." Proceedings of the IEEE/cvf conference on computer vision and pattern recognition. 2021.