

## 채널 단위 주의 메커니즘을 활용한 실시간 소형 객체 탐지 성능 향상 기법

오 승 민<sup>1</sup> · 김 진 술<sup>2\*</sup><sup>1</sup>전남대학교 지능전자컴퓨터공학과 박사과정<sup>2</sup>전남대학교 지능전자컴퓨터공학과 교수

# Performance Enhancement of Real-Time Small Object Detection Using Channel-Wise Attention Mechanism

Seungmin Oh<sup>1</sup> · Jinsul Kim<sup>2\*</sup><sup>1</sup>Doctor's Course, Department of Intelligence Electronic and Computer Engineering, Chonnam National University, Gwangju 61186, Korea<sup>2</sup>Professor, Department of Intelligence Electronic and Computer Engineering, Chonnam National University, Gwangju 61186, Korea

### [요 약]

최근 단일 단계 기반 실시간 객체 탐지 모델은 우수한 성능을 보이고 있으나, 소형 객체는 탐지 정확도가 낮은 한계를 가진다. 본 논문에서는 실시간 임베디드 시스템에서 소형 객체 탐지 성능 향상을 위해 채널 주의 메커니즘 기반의 EC-YOLO(efficient channel you only look once)를 제안한다. 기존 YOLOv8(you only look once)은 다중 스케일 특징 피라미드 구조를 사용하지만 대형 객체에 편향되는 경향이 있어 이를 보완하기 위해 채널 단위 특징 재조정 기법과 고해상도 P2 검출 헤드를 적용하였다. 제안 모델은 백본과 넥에 경량 Efficient Channel Module(ECM)을 통합하여 소형 객체 공간정보를 강화한다. 차량 브랜드 로고 데이터셋 실험 결과, EC-YOLO는 YOLOv8 대비 mAP@50 기준 최대 2.7% 향상과 추론 속도 개선을 보였다.

### [Abstract]

Recent models for single-stage real-time object detection have achieved high performances; however, their accuracy for small objects remains limited. This study proposes EC-YOLO(efficient channel you only look once), a channel attention-based model designed to improve small object detection in real-time embedded systems. Although YOLOv8 employs a multi-scale feature pyramid structure, it tends to be biased toward large objects. To address this issue, we introduce a channel-wise feature recalibration mechanism and a high-resolution P2 detection head. The proposed model integrates a lightweight efficient channel module (ECM) into the backbone and neck to enhance feature representation for small objects. Experiments on a vehicle logo dataset show that EC-YOLO achieves up to a 2.7% improvement in mAP@50 compared to YOLOv8 while also improving inference speed.

**색인어** : 소형 객체 탐지, YOLOv8, 채널 주의 메커니즘, 실시간 객체 검출, 특징 재조정**Keyword** : Small Object Detection, YOLOv8, Channel Attention, Real-Time Detection, Feature Recalibration<http://dx.doi.org/10.9728/dcs.2026.27.4.1105>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 19 February 2026; Revised 06 March 2026

Accepted 12 March 2026

\*Corresponding Author, Jinsul Kim

Tel: +82-62-530-0407

E-mail: jsworld@jnu.ac.kr

## I. 서론

객체 탐지 기술은 영상 또는 이미지 내에서 객체의 위치와 종류를 동시에 식별하는 컴퓨터 비전 기술로서 자율주행, 지능형 감시 시스템, 드론 기반 영상 분석, 산업 자동화 등 다양한 디지털 콘텐츠 응용 분야에서 폭넓게 활용되고 있다[1]. 최근 합성곱 신경망(CNN; convolutional neural network)의 발전과 대규모 데이터셋의 등장으로 객체 탐지 정확도는 크게 향상되었으며, 특히 단일 단계 기반 탐지기는 이중 단계에 비해 빠른 추론 속도를 바탕으로 실시간 응용 환경에 적합한 구조로 자리매김하였다[2].

그러나 일반 객체 탐지 기술의 발전과 달리, 소형 객체 탐지는 여전히 해결이 필요한 연구 과제로 남아 있다. 소형 객체는 입력 영상에서 차지하는 픽셀 비율이 매우 낮아 특징 표현이 제한적이며, 깊은 계층을 통과하는 과정에서 공간 정보가 점차 감소하는 특성을 가진다. 이와 관련하여 어텐션 메커니즘을 통해 특징 수준, 공간 위치, 출력 채널 간의 관계를 통합하는 프레임워크를 제안하여 소형 객체의 표현력을 개선하고자 하였다[3],[4]. 또한, 다중 스케일 특징 피라미드(FPN; feature pyramid network) 구조를 사용하는 최신 모델들조차 상대적으로 표현력이 강한 대형 객체에 편향되는 경향이 보고되고 있다. 그 결과, 복잡한 배경 환경이나 저해상도 조건에서 소형 객체는 배경에 묻혀 검출되지 않거나 오검출되는 문제가 빈번히 발생한다. 이에 따라, FPN 내 인접 계층 간의 상향식과 하향식 연결이 소형 객체 탐지에 미치는 영향을 분석하고, 데이터셋의 객체 분포에 따라 융합 비율을 조절하는 인자의 중요성을 강조하였다.

최근 YOLO 계열 모델은 앵커 프리 개념을 도입하여 다양한 객체 크기에 대한 적응성을 개선하고 정확도와 추론 속도 간의 균형을 향상시켰다[5]. 특히 YOLOv8은 구조 단순화와 학습 안정성 개선을 통해 실시간 객체 탐지 분야에서 우수한 성능을 보이고 있다. 그러나 특징 추출 단계에서 채널 간 중요도를 세밀하게 반영하지 못하는 한계가 존재하며, 채널 차원의 정보 재보정을 하지 않을 경우 미세한 구조적 특징을 갖는 소형 객체는 배경 특징에 의해 억제될 수 있다. 이는 네트워크가 공간 스케일뿐 아니라 채널 차원에서도 정보 불균형을 학습할 가능성을 내포한다.

기존 연구들은 이러한 문제를 해결하기 위해 네트워크 깊이를 확장하거나 복잡한 특징 융합 구조를 도입하는 방식을 제안하였다[6]. 그러나 이러한 접근은 파라미터 수 증가와 연산량 확대를 수반하여 실시간 임베디드 환경에서의 적용성을 제한하는 요인이 된다. 따라서 정확도 향상과 연산 효율성을 동시에 만족하는 경량 구조 설계가 요구된다.

이에 본 논문에서는 채널 단위 주의 메커니즘을 활용한 실시간 소형 객체 탐지 성능 향상 기법을 제안한다. 제안 모델은 YOLOv8을 기반으로 하여 백본과 넥 영역에 경량 ECM을 통합함으로써 정보성이 높은 채널을 동적으로 강조하고 불필요한 채널 정보를 억제한다. 또한 고해상도 P2 검출 헤드를

추가하여 저해상도 단계에서 손실되는 공간 정보를 보완함으로써 소형 객체의 표현력을 강화하였다.

본 논문의 주요 기여는 다음과 같다.

- 1) 채널 단위 특징 재보정을 통해 소형 객체의 표현력을 강화한 EC 기반 구조를 제안하였다.
- 2) 고해상도 P2 검출 헤드를 추가하여 소형 객체의 공간 정보 보존을 개선하였다.
- 3) 정확도 향상과 추론 속도 개선을 동시에 달성하여 실시간 임베디드 환경에 적합한 구조를 제시하였다.

이후 본 논문의 구성은 다음과 같다. II장에서는 관련 연구를 정리하고, III장에서는 제안하는 네트워크 구조를 설명한다. IV장에서는 실험 환경 및 성능 비교 결과를 제시하며, V장에서 결론을 맺는다.

## II. 소형 객체 탐지를 위한 관련 기술 동향

소형 객체 탐지의 성능 향상을 위해 다양한 구조적 개선 기술들이 제안되어 왔다. 기존 연구는 크게 특징 추출 단계의 개선과 다중 스케일 특징 융합 구조의 최적화라는 두 가지 방향으로 구분할 수 있다.

먼저 특징 추출 능력을 향상시키기 위해 잔차 연결과 주의 메커니즘을 결합한 연구가 다수 보고되었다. 잔차 주의 메커니즘을 통해 특징 맵 내의 중요한 정보를 강조함으로써 멀티스케일 객체에 대한 표현력을 개선하였다[7].

이후 트랜스포머 기반 구조가 도입되면서 글로벌 문맥 정보를 반영하는 방식이 제안되었으며, 이를 통해 복잡한 배경 환경에서 객체 구분 능력을 향상시키려는 시도가 이루어졌다[8]. 그러나 이러한 접근은 네트워크 구조를 복잡하게 만들고 연산량을 증가시키는 한계를 가진다.

다중 스케일 특징 융합을 위한 연구도 활발히 진행되었다. FPN 기반 구조는 서로 다른 해상도의 특징을 통합하여 다양한 크기의 객체를 탐지하도록 설계되었으며, 이후 상-하향 경로를 보강하거나 특징 융합 가중치를 조정하는 방식이 제안되었다[9],[10]. 이러한 방법은 중-대형 객체 탐지 성능 개선에는 효과적이었으나, 소형 객체의 경우 저해상도 단계에서 이미 공간 정보가 손실되는 문제가 완전히 해결되지는 않았다.

실시간 객체 탐지의 표준 모델인 YOLO 계열 모델을 기반으로 한 소형 객체 탐지 연구 역시 지속적으로 제안되어 왔다[11]. 그림 1에서 확인할 수 있듯이, 최근 YOLO 기반 모델들은 소형 객체 검출력을 높이기 위해 표준 구조에 잔차 주의나 트랜스포머 기반 구조를 통합하여 강화하는 추세이다. 구체적으로 하이브리드 트랜스포머 헤드(HTH; hybrid transformer head)를 도입하여 전역적 문맥 정보를 반영하고, 합성곱 주의 특징 융합 모듈을 통해 채널 및 공간 차원의 가중치를 동적으로 조절함으로써 mAP를 개선하였다[11],[12]. 그러나 이러한 방식은 파라미터 수 증가와 추론 속도 저하를 수반하여 실시간 응용 환경에서의 적용성을 제한하는 요인이 된다.

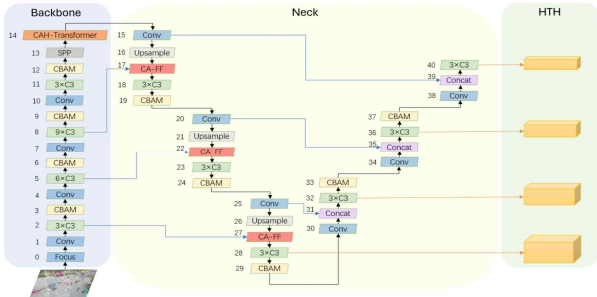


그림 1. 소형 객체 탐지를 위한 HTH-YOLOv5 아키텍처[11]  
 Fig. 1. HTH-YOLOv5 architecture for small object detection [11]

한편, 채널 차원의 경량 주의 메커니즘으로는 ECA-Net이 제안되었으며, 이는 1차원 합성곱을 활용하여 채널 간 상호작용을 효율적으로 학습하는 구조를 제시하였다[13]. 해당 방법은 파라미터 증가를 최소화하면서도 채널 중요도 재보정이 가능하다는 장점을 가진다. 그러나 이를 소형 객체 탐지에 특화하여 백본 및 검출 헤드 구조와 결합한 연구는 제한적으로 보고되고 있다.

종합하면, 기존 연구는 네트워크 깊이 확장이나 복잡한 특징 융합 구조를 통해 소형 객체 탐지 성능을 향상시키고자 하였으나, 연산 효율성과 실시간성을 동시에 확보하는 데에는 한계가 존재한다. 특히 채널 단위 정보 재보정을 활용하여 경량 구조 내에서 소형 객체의 특징 표현력을 강화하는 접근은 충분히 연구되지 않았다. 이에 본 논문에서는 채널 단위 주의 메커니즘을 활용하여 특징 추출 단계에서 정보성이 높은 채널을 동적으로 강조하고, 고해상도 검출 구조를 결합함으로써 소형 객체 탐지 성능을 개선하고자 한다.

### III. 채널 단위 특징 재보정 및 고해상도 분기를 활용한 EC-YOLO 설계

#### 3-1 기존 YOLOv8 구조 분석 및 한계

YOLOv8은 단일 단계 기반의 앵커 프리 객체 탐지 모델로 입력, 백본, 넥, 헤드로 구성된다. YOLOv8의 기본 구조는 그림 2와 같다.

입력 영상은 640×640 해상도로 정규화되며, 특징 추출은 CSP(cross stage partial) 기반 구조를 통해 수행된다. 백본에는 CBS(convolution+batchNorm+SiLU), C2F(CSP Bottleneck with two convolutions), SPPF(spatial pyramid pooling-fast) 모듈이 포함되어 다중 수용 영역과 효율적인 그래디언트 흐름을 제공한다.

넥 영역은 FPN와 경로 집계 네트워크 구조(PAN; path aggregation network)를 결합하여 서로 다른 해상도의 특징을 상향 및 하향 경로로 융합하며, 헤드에서는 각 스케일별로 분류와 경계 상자 회귀를 수행한다.

분류 손실은 이진 교차 엔트로피(BCE; binary cross-entropy)로 정의되며 다음과 같이 표현된다.

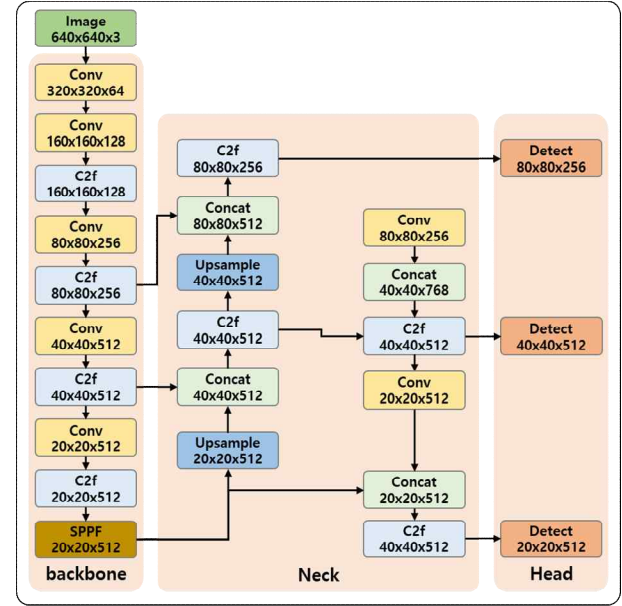


그림 2. YOLOv8 기반 객체 탐지 네트워크 구조  
 Fig. 2. Overall architecture of the YOLOv8 object detection network

$$L_{cls} = -a[y \log(p) + (1-y) \log(1-p)] \quad (1)$$

여기서  $y$ 는 정답 레이블,  $p$ 는 모델이 예측한 클래스 확률, 그리고  $a$ 는 클래스 불균형을 보정하기 위한 가중치 계수를 의미한다.

경계 상자 회귀에는 CIoU(complete intersection over union) 손실이 사용되며 다음과 같이 정의된다.

$$L_{CIoU} = 1 - IoU + \frac{p^2(b, b^{gt})}{c^2} + av \quad (2)$$

여기서  $IoU$ 는 예측 박스와 정답 박스 간의 교집합 대비 합집합 비율,  $p^2(b, b^{gt})$ 는 예측 박스 중심점과 정답 박스 중심점 간의 유클리드 거리,  $c$ 는 두 박스를 포함하는 최소 외접 박스의 대각선 길이,  $v$ 는 종횡비 차이를 반영하는 항,  $a$ 는 종횡비 항의 영향도를 조절하는 가중치 계수를 의미한다.

그러나 이러한 구조는 소형 객체 탐지에 있어 다음과 같은 한계를 가진다. 첫째, 32배 다운샘플링 단계(P5)에서 공간 해상도가 급격히 감소하여 소형 객체의 표현력이 크게 저하된다. 둘째, 채널 차원의 중요도 재보정이 충분히 이루어지지 않아 배경 특징이 상대적으로 강조될 수 있다. 셋째, 소형 객체에 특화된 고해상도 검출 분기가 존재하지 않는다.

따라서 채널 차원의 정보 재보정과 고해상도 특징 활용을 동시에 고려한 구조 개선이 필요하다.

#### 3-2 제안하는 EC-YOLO 구조

본 논문에서는 상기 문제를 해결하기 위해 ECM과 고해상도 P2 검출 헤드를 통합한 EC-YOLO 구조를 제안한다.

1) ECM 모듈

ECM은 채널 간 상호작용을 경량 구조로 학습하기 위한 모듈로, 전역 평균 풀링(GAP; global average pooling)을 통해 입력 특징 맵을 채널 벡터로 변환한다. GAP는 다음과 같이 정의된다.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_c(i, j) \tag{3}$$

여기서  $F_c(i, j)$ 는 채널  $c$ 에서 위치  $(i, j)$ 의 특징 값,  $H$ 와  $W$ 는 입력 특징 맵의 높이와 너비,  $z_c$ 는 채널  $c$ 에 대한 전역 평균 풀링 결과로 얻어진 채널 통계값을 의미한다.

이후 1차원 합성곱을 이용하여 인접 채널 간 상호작용을 모델링하며, 시그모이드 함수를 통해 채널 가중치를 산출한다.

$$s = \sigma(\text{Conv1D}(z)) \tag{4}$$

여기서  $z$ 는 GAP를 통해 얻은 채널 벡터,  $\text{Conv1D}(z)$ 는 채널 간 관계를 학습하기 위한 1차원 합성곱 연산,  $\sigma$ 는 시그모이드 활성화 함수,  $s$ 는 각 채널에 대한 가중치 벡터를 의미한다.

최종적으로 원 특징 맵에 채널 가중치를 곱하여 다음과 같이 정의된다.

$$F' = s \odot F \tag{5}$$

여기서  $F$ 는 입력 특징 맵,  $s$ 는 채널 가중치 벡터,  $\odot$ 는 채널 단위의 요소별 곱을 의미하여,  $F'$ 는 가중치가 반영된 출력 특징맵을 나타낸다.

이와 같은 구조는 파라미터 증가를 최소화하면서도 정보성이 높은 채널을 강조하고 불필요한 채널을 억제하는 효과를 가진다. 본 연구에서는 ECM을 백본과 neck의 주요 블록 뒤에 삽입하여 채널 중요도를 동적으로 재조정하였다. 제안하는 ECM의 구조는 그림 3과 같다.

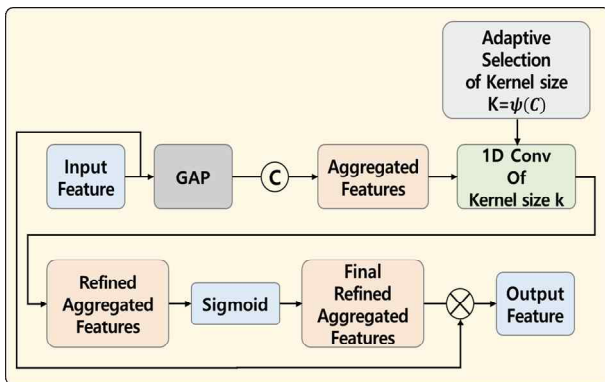


그림 3. ECM 구조  
Fig. 3. Structure of the ECM

2) 고해상도 P2 검출 헤드 추가

기존 YOLOv8은 8배, 16배, 32배 단계의 특징 맵을 활용하여 객체를 검출한다. 그러나 32배 다운샘플링된 P5 단계에서는 입력 해상도 640×640 기준 20×20 크기의 특징 맵만 생성된다. 이 경우 32픽셀 이하의 소형 객체는 1픽셀 수준으로 축소되어 공간 정보가 충분히 유지되지 못한다.

이러한 한계를 해결하기 위해 본 연구에서는 4배 다운샘플링 단계의 P2 특징 맵을 활용하는 고해상도 검출 분기를 추가하였다. P2 단계의 특징 맵은 160×160 해상도를 유지하며, 이는 소형 객체의 세밀한 구조 정보를 보존하는 데 유리하다.

입력 해상도 640×640 기준 각 단계의 특징 맵 해상도는 표 1에 정리하였다.

표 1. 다운샘플링 단계별 특징 맵 해상도 비교(입력 640×640 기준)  
Table 1. Feature map resolution comparison under 640×640 input

Stage	Down-sampling	Feature Map Size	YOLOv8	EC-YOLO
P2	4×	160 × 160	Not used	Used
P3	8×	80 × 80	Used	Used
P4	16×	40 × 40	Used	Used
P5	32×	20 × 20	Used	Removed

표 1에서 확인할 수 있듯이, 제안 모델은 기존 P5 검출 분기를 제거하고 P2 분기를 추가하였다. 이는 소형 객체 탐지에 불리한 저해상도 단계를 축소하고, 고해상도 특징을 적극적으로 활용하기 위한 구조적 선택이다.

결과적으로 제안 구조는 채널 단위 특징 재보정과 고해상도 검출 분기를 동시에 통합함으로써 소형 객체 탐지 성능을 향상시키면서도 실시간 추론 속도를 유지하도록 설계되었다.

3) 제안된 EC-YOLO 구조

제안하는 EC-YOLO의 전체 네트워크 구조는 그림 3에 나타내었다. 기존 YOLOv8 구조를 기반으로 하되, 백본과 neck 영역에 ECM을 통합하고 P2 고해상도 검출 분기를 추가한 것이 핵심적인 차이점이다. YOLOv8 구조와 제안하는 EC-YOLO의 차이는 표 2에 정리하였다.

표 2. EC-YOLO와 YOLOv8 구조 비교  
Table 2. Architectural comparison between YOLOv8 and proposed EC-YOLO

Component	YOLOv8	Proposed EC-YOLO
Backbone	CSP-based backbone	CSP backbone + ECM
Neck	FPN + PAN	FPN + PAN + ECM
Channel Attention	Not applied	ECM
Detection Heads	P3, P4, P5	P2, P3, P4
Large Object Branch	P5 (32×)	Removed
Real-Time Inference	Yes	Yes

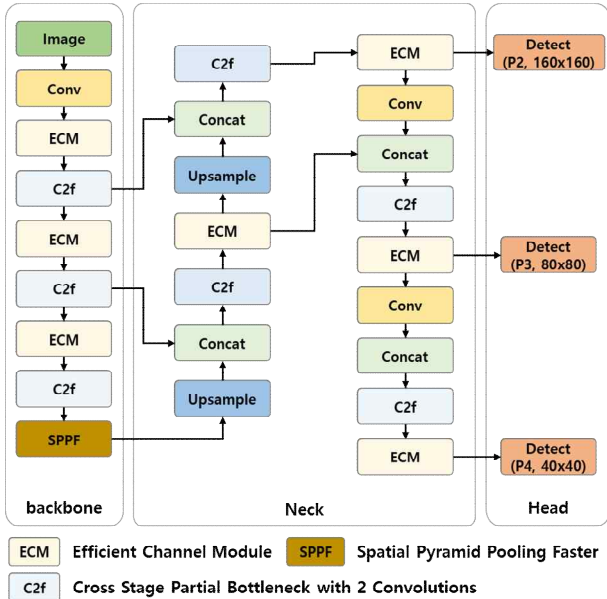


그림 4. 제안하는 EC-YOLO 전체 네트워크 구조  
 Fig. 4. Overall architecture of the proposed EC-YOLO

표 2에서 확인할 수 있듯이, 제안 모델은 기존 구조를 유지하면서도 채널 단위 특징 재보정과 고해상도 검출 분기를 통합하였다. 따라서, 제안 모델은 다음과 같은 구조적 특징을 가진다.

첫째, 백본의 주요 특징 추출 블록 뒤에 ECM을 삽입하여 채널 단위 특징 중요도를 동적으로 재조정한다. 이를 통해 소형 객체의 미세한 채널 특징이 배경 특징에 의해 억제되는 현상을 완화하였다.

둘째, 넥 영역의 상·하향 경로에도 ECM을 통합하여 다중 스케일 특징 융합 과정에서 채널 차원의 정보 균형을 유지하도록 설계하였다.

셋째, 기존 32배 다운샘플링 기반의 P5 검출 분기를 제거하고, 4배 다운샘플링 기반의 P2 검출 분기를 추가하여 소형 객체에 대한 공간 해상도를 강화하였다.

그림 4에서 확인할 수 있듯이, ECM은 백본과 넥의 여러 단계에 분산 배치되며, P2, P3, P4 검출 헤드가 각각 독립적으로 객체 분류 및 위치 회귀를 수행한다. 이러한 구조는 채널 차원과 공간 차원의 개선을 동시에 달성하도록 설계되었다.

#### IV. 실험 및 결과 분석

##### 4-1 실험 환경 및 데이터셋

본 연구의 실험은 Intel Core i5-13400 CPU, NVIDIA GeForce RTX 3060 Ti GPU, 8GB VRAM, 32GB RAM 환경에서 수행하였다. 운영체제는 Windows 10 64-bit이며,

딥러닝 프레임워크는 Python 3.9, PyTorch 1.11.0, Torchvision 0.12.0을 사용하였다.

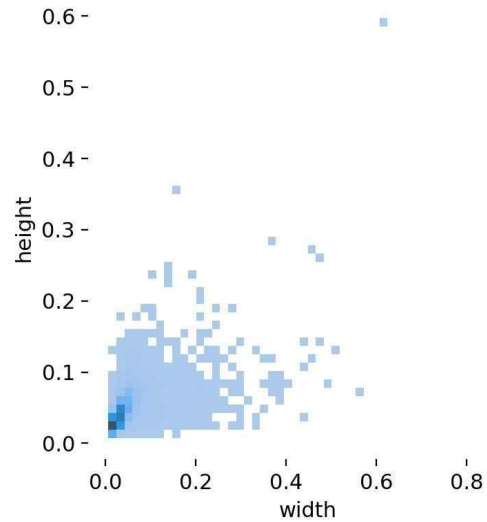


그림 5. 객체 크기 분포도  
 Fig. 5. Distribution of object widths and heights

제안 모델의 성능 검증을 위해 차량 브랜드 로고 데이터셋을 활용하였다. 해당 데이터셋은 총 32개 클래스이며, 학습 데이터 5,056장, 검증 데이터 786장, 테스트 데이터 695장으로 구성된다.

그림 5는 객체의 너비와 높이 분포를 나타낸 것이다. 그림 5에서 확인할 수 있듯이, 대부분의 객체가 좌측 하단 영역에 밀집되어 있으며 이는 32×32 픽셀 이하의 소형 객체 비율이 매우 높음을 의미한다. 이러한 데이터 특성은 소형 객체 탐지 성능을 평가하기에 적합한 환경을 제공한다.

##### 4-2 평가 지표

정량적 평가를 위해 정밀도(Precision), 재현율(Recall), F1-score 및 mAP를 사용하였다. 정밀도, 재현율, F1-score는 각각 다음과 같이 정의된다.

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (8)$$

mAP는 Precision-Recall 곡선 아래 면적으로 계산되며, mAP는 모든 클래스의 AP(average precision) 평균으로 정의된다.

4-3 정량적 성능 비교

먼저 YOLOv8의 다양한 모델 크기(n, s, m, l, x)에 대한 성능을 비교하여 기준 성능을 설정하였다. 결과는 표 3에 정리하였다.

표 3. YOLOv8 모델 크기별 성능 비교

Table 3. Performance comparison of YOLOv8 variants

Model	Parameters (M)	mAP@50	mAP@50-95	Inference Time (ms)
YOLOv8n	3.0	0.553	0.244	1.4
YOLOv8s	11.1	0.662	0.277	2.0
YOLOv8m	25.8	0.699	0.299	3.9
YOLOv8l	43.6	0.715	0.310	4.3
YOLOv8x	68.1	0.741	0.348	6.0

표 3에서 확인할 수 있듯이, 모델 규모가 증가할수록 mAP는 향상되지만 추론 시간 또한 증가한다. 이는 정확도와 연산 비용 간의 trade-off 관계를 보여준다.

제안 모델의 성능 향상을 보다 명확히 분석하기 위해 YOLOv8과 EC-YOLO를 직접 비교하였다. 표 4에서 확인할 수 있듯이, 제안 모델은 모든 모델 규모(n, s, m, l, x)에서 기존 YOLOv8 대비 일관된 성능 향상을 보였다. 특히 mAP@50 기준으로 최대 약 2.7%의 향상이 나타났으며, mAP@50-95 기준으로는 최대 약 5.1%, F1 Score는 최대 약 6.9%의 성능 개선을 보였다. 또한 추론 시간 역시 전반적으로 감소하여 최대 약 21.6%의 속도 향상을 보이며 정확도와 연산 효율성을 동시에 향상시킨 것을 확인할 수 있다. 이러한 결과는 제안한 구조가 단일 성능 지표의 개선에 국한되지 않고 다양한 평가 지표에서 일관된 개선 경향을 보인다는 점에서 의미가 있다.

표 4. 제안 EC-YOLO와 YOLOv8 성능 비교

Table 4. Performance comparison between YOLOv8 and proposed EC-YOLO

Model	mAP@50	mAP@50-95	F1 Score	Inference Time (ms)
YOLOv8n	0.553	0.244	0.58	1.4
EC-YOLO-n	0.567 (+2.5%)	0.251 (+2.9%)	0.62 (+6.9%)	1.2 (-14.2%)
YOLOv8s	0.662	0.277	0.68	2.0
EC-YOLO-s	0.68 (+2.7%)	0.291 (+5.1%)	0.70 (+2.9%)	1.7 (-15.0%)
YOLOv8m	0.699	0.299	0.69	3.9
EC-YOLO-m	0.707 (+1.1%)	0.309 (+3.3%)	0.71 (+2.9%)	3.1 (-20.5%)
YOLOv8l	0.715	0.310	0.72	4.3
EC-YOLO-l	0.722 (+1.0%)	0.321 (+3.5%)	0.74 (+2.8%)	3.8 (-11.6%)
YOLOv8x	0.741	0.348	0.73	6.0
EC-YOLO-x	0.756 (+2.0%)	0.361 (+3.7%)	0.75 (+2.7%)	4.7 (-21.6%)

또한 동일한 실험 환경에서 모든 모델 규모에 대해 일관된 성능 향상이 관찰되었으며, 이는 제안 구조가 특정 실험 조건에 의존하기보다는 구조적 개선에 의해 안정적인 성능 향상을 제공함을 시사한다. 객체 탐지 모델의 학습 과정에서는 초기 가중치나 데이터 샘플링에 따라 일부 성능 변동이 발생할 수 있으나, 본 연구에서는 다양한 모델 규모에서 동일한 개선 경향이 반복적으로 나타났다는 점에서 제안 방법의 효과가 안정적으로 나타난 것으로 판단된다.

일반적으로 ECM과 고해상도 P2 검출 헤드의 추가는 연산량 증가 요인이 될 수 있다. 그러나 본 연구에서는 기존 YOLOv8 구조에서 사용되던 32배 다운샘플링 기반의 P5 검출 분기를 제거함으로써 저해상도 단계에서 수행되는 고채널 특징 연산을 줄였다. 특히 P5 단계는 높은 채널 수를 가진 특징 맵에서 검출 연산이 수행되기 때문에 전체 연산량에 큰 영향을 미치는 구조적 요소이다. 반면 ECM은 1차원 합성곱 기반의 경량 채널 주의 모듈로 설계되어 추가되는 파라미터와

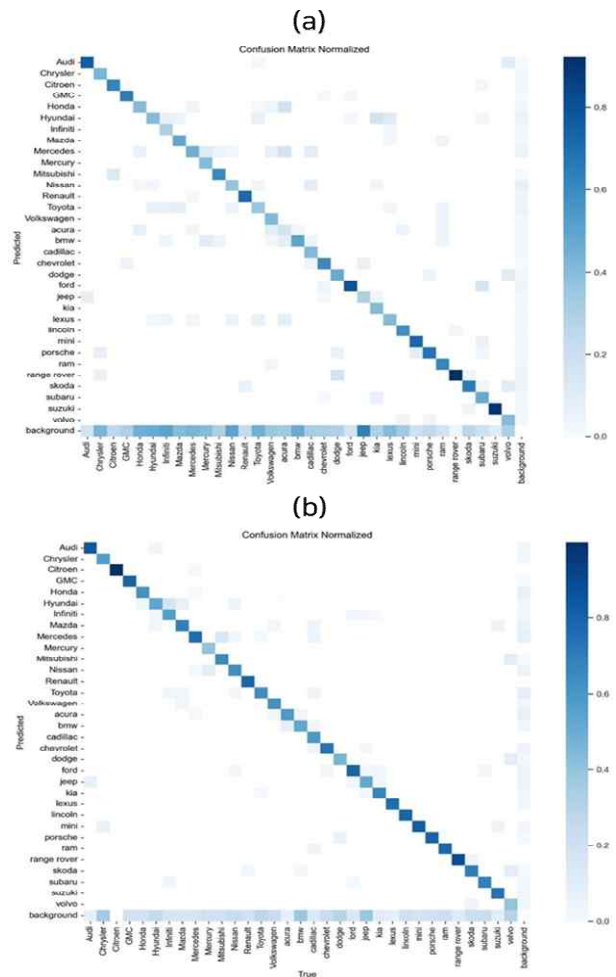


그림 6. YOLOv8(a)과 제안한 EC-YOLO 모델(b)의 정규화된 혼동 행렬 비교

Fig. 6. Normalized confusion matrix comparison between YOLOv8(a) and the proposed EC-YOLO(b)

연산 오버헤드가 제한적이다. 이러한 구조적 변화는 불필요한 고채널 연산을 감소시키면서도 중요한 채널 정보를 강조하도록 하여 전체 연산 구조를 보다 효율적으로 구성하게 하며, 결과적으로 추론 속도 향상에 기여한 것으로 분석된다.

표 기반의 정량적 성능 비교와 함께 모델의 분류 특성을 분석하기 위해 혼동 행렬을 시각적으로 비교하였다. 그림 5는 YOLOv8과 제안한 EC-YOLO 모델의 정규화된 혼동 행렬을 나타낸다. 기존 YOLOv8 모델의 경우 일부 차량 브랜드 간 오분류가 나타나는 것을 확인할 수 있으며, 특히 유사한 형태의 로고를 가지는 클래스 간에서 혼동이 발생한다. 반면 제안한 EC-YOLO 모델은 대부분의 클래스에서 대각선 성분이 더욱 뚜렷하게 나타나며 클래스 간 오분류가 감소하는 경향을 보인다. 이는 ECM 모듈이 채널 간 상호작용을 효과적으로 학습하여 특징 표현력을 향상시키고 차량 로고와 같은 소형 객체의 세부 특징을 보다 정확하게 반영한 결과로 해석할 수 있다. 이러한 결과는 표 4에서 나타난 mAP 및 F1-score 향상과 일관된 경향을 보이며 제안한 구조가 실제 분류 성능 개선에도 기여함을 확인할 수 있다.

#### 4-4 정성적 성능 비교

다음은 YOLOv8x 모델과 EC-YOLO-x에 대한 정성적 성능을 비교하여 살펴본다.

그림 7은 Grad-CAM(gradient-weighted class activation map) 기반의 정성적 시각화를 통해 모델의 활성화 분포를 비교한 결과이다. 기존 YOLOv8은 배경 영역에도 비교적 높은 활성도를 보이는 반면, EC-YOLO는 차량 로고 영역에 보다 집중된 활성도를 나타내는 경향을 보인다. 이러한 결과는 채널 단위 특징 재보정이 객체 영역에 대한 특징 표현을 강화하는 방향으로 작용할 가능성을 시사한다.

그림 8은 실제 도로 환경에서의 검출 결과를 비교한 것이다. 기존 YOLOv8은 전경의 객체는 탐지하였으나 원거리 소형 객체를 검출하지 못하였다. 반면 EC-YOLO는 원거리 소형 객체까지 성공적으로 탐지함을 확인할 수 있다.

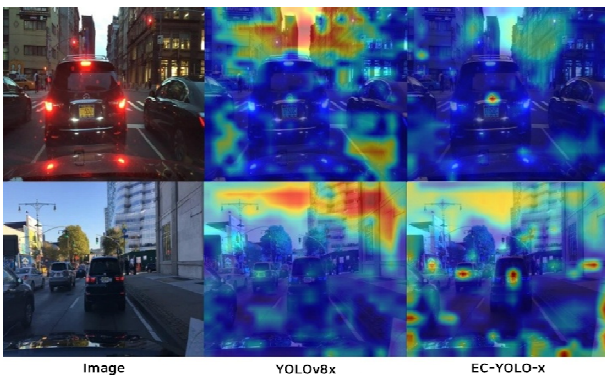


그림 7. Grad-CAM 시각화 비교  
Fig. 7. Grad-CAM comparison between YOLOv8x and EC-YOLO-x

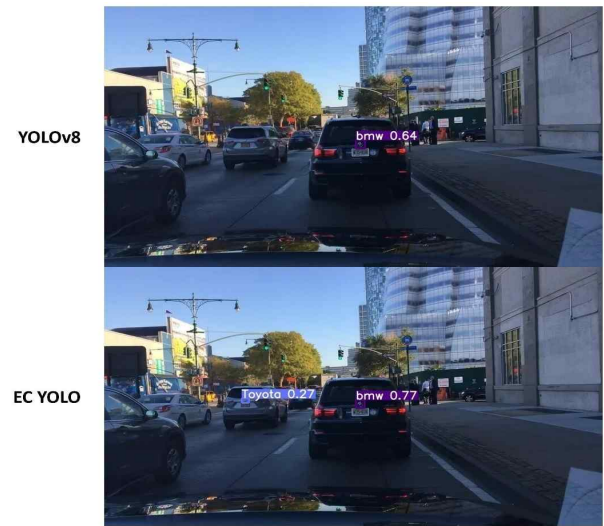


그림 8. 검출 결과 비교  
Fig. 8. Detection result comparison between YOLOv8 and EC-YOLO

## V. 결 론

본 논문에서는 소형 객체 탐지의 구조적 한계를 해결하기 위해 채널 단위 주의 메커니즘 기반의 실시간 소형 객체 탐지 기법을 제안하였다. 기존 YOLOv8 구조는 정확도와 추론 속도 측면에서 우수한 성능을 보이지만, 소형 객체 탐지에 있어 공간 정보 손실과 채널 차원의 정보 불균형이라는 구조적 한계를 가진다.

이를 해결하기 위해 본 연구에서는 백본 및 넥 영역에 ECM을 통합하여 채널 단위 특징을 동적으로 재조정하고, 4배 다운샘플링 기반의 P2 고해상도 검출 분기를 추가하였다. 또한 소형 객체 표현에 불리한 P5 분기를 제거함으로써 고해상도 특징 활용을 강화하였다.

차량 브랜드 로고 데이터셋을 활용한 실험 결과, 제안 모델은 기존 YOLOv8 대비 모든 모델 규모에서 mAP@50 및 mAP@50-95와 F1-score가 향상되었으며, 특히 EC-YOLO-s는 YOLOv8s 대비 mAP@50 기준 최대 약 2.7%의 성능 향상을 보였다. EC-YOLO-x는 YOLOv8x 대비 추론 시간 기준 약 21.6% 단축되어 정확도와 연산 효율성을 동시에 개선했음을 확인하였다. 또한 Grad-CAM 시각화를 통해 제안 모델이 객체 영역에 보다 집중된 활성화 패턴을 보이는 경향을 확인하였다.

본 연구는 소형 객체 탐지에서 채널 단위 특징 재보정의 중요성을 실험적으로 입증하였으며, 복잡한 네트워크 확장 없이도 경량 구조 기반으로 성능 개선이 가능함을 제시하였다. 이는 실시간 임베디드 시스템 및 엣지 환경에서 활용 가능한 객체 탐지 모델 설계에 유용한 구조적 방향성을 제공한다.

향후 연구에서는 모델의 파라미터 수 및 연산량을 추가적으로 최적화하여 초저전력 환경에 적합한 경량화를 수행하고,

위성 영상 분석, 산업 결합 검출, 드론 기반 원거리 탐지 등 다양한 소형 객체 응용 분야로 확장할 계획이다.

### 감사의 글

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신 기획평가원-지역지능화혁신인재양성사업의 지원을 받아 수행된 연구임(IITP-2026-RS-2022-00156287).

### 참고문헌

[1] J. E. Gallagher and E. J. Oughton, "Surveying You Only Look Once (YOLO) Multispectral Object Detection Advancements, Applications, and Challenges," *IEEE Access*, Vol. 13, pp. 7366-7395, 2025. <https://doi.org/10.1109/ACCESS.2025.3526458>

[2] Z. Tian, C. Shen, H. Chen, and T. He, "FCOS: A Simple and Strong Anchor-Free Object Detector," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 4, pp. 1922-1933, April 2022. <https://doi.org/10.1109/TPAMI.2020.3032166>

[3] X. Dai, Y. Chen, B. Xiao, D. Chen, M. Liu, L. Yuan, and L. Zhang, "Dynamic Head: Unifying Object Detection Heads with Attentions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville: TN, pp. 7369-7378, June 2021. <https://doi.org/10.1109/CVPR46437.2021.00729>

[4] Y. Gong, X. Yu, Y. Ding, X. Peng, J. Zhao, and Z. Han, "Effective Fusion Factor in FPN for Tiny Object Detection," in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa: HI, pp. 1159-1167, January 2021. <https://doi.org/10.1109/WACV48630.2021.00120>

[5] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver: Canada, pp. 7464-7475, June 2023. <https://doi.org/10.1109/CVPR52729.2023.00721>

[6] Z. Guo, C. Wang, G. Yang, Z. Huang, and G. Li, "MSFT-YOLO: Improved YOLOv5 Based on Transformer for Detecting Defects of Steel Surface," *Sensors*, Vol. 22, No. 9, pp. 3467-3481, May 2022. <https://doi.org/10.3390/s22093467>

[7] J. Zhang, Q. Qi, H. Zhang, Q. Du, F. Wang, and X. Shi, "Residual Attention Mechanism and Weighted Feature Fusion for Multi-Scale Object Detection," *Multimedia Tools and Applications*, Vol. 82, No. 26, pp. 40873-40889, 2023. <https://doi.org/10.1007/S11042-023-14997-8>

[8] S. Fang, X. Lu, Y. Huang, G. Sun, and X. Liu, "Dynamic Multi-Headed Self-Attention and Multiscale Enhancement Vision Transformer for Object Detection," *Multimedia Tools And Applications*, Vol. 83, No. 25, pp. 67213-67229, 2024. <https://doi.org/10.1007/s11042-024-18234-8>

[9] F. Yang, J. Zhou, Y. Chen, J. Liao, and M. Yang, "MSF-YOLO: A Multi-Scale Features Fusion-Based Method for Small Object Detection," *Multimedia Tools and Applications*, Vol. 83, No. 22, pp. 61239-61260, 2024. <https://doi.org/10.1007/s11042-023-17818-0>

[10] L. Jiao, C. Kang, S. Dong, P. Chen, G. Li, and R. Wang, "An Attention-Based Feature Pyramid Network for Single-Stage Small Object Detection," *Multimedia Tools and Applications*, Vol. 82, No. 12, pp. 18529-18544, 2023. <https://doi.org/10.1007/s11042-022-14159-2>

[11] Z. Liu, X. Gao, Y. Wan, J. Wang, and H. Lyu, "An Improved YOLOv5 Method for Small Object Detection in UAV Capture Scenes," *IEEE Access*, Vol. 11, pp. 14365-14374, 2023. <https://doi.org/10.1109/ACCESS.2023.3241005>

[12] Y. Liu, G. He, Z. Wang, W. Li, and H. Huang, "NRT-YOLO: Improved YOLOv5 Based on Nested Residual Transformer for Tiny Remote Sensing Object Detection," *Sensors*, Vol. 22, No. 13, pp. 4953-4968, 2022. <https://doi.org/10.3390/S22134953>

[13] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle: WA, pp. 11531-11539, June 2020. <https://doi.org/10.1109/CVPR42600.2020.01155>

### 오승민 (Seungmin Oh)



2019년 : 한국나사렛대학교  
디지털콘텐츠학과(공학사)

2021년 : 전남대학교  
ICT융합시스템학과(공학석사)

2021년~현 재: 전남대학교 지능전자컴퓨터공학과 박사과정

※ 관심분야 : 딥러닝, 인공지능

### 김진술(Jinsul Kim)



2001년 : Computer Science from  
University of Utah, Salt  
Lake City, Utah,  
USA(공학사)

2005년 : 한국과학기술원  
정보통신공학(공학석사)

2008년 : 한국과학기술원  
정보통신공학(공학박사)

2005년~2008년: 한국전자통신연구원 IPTV 인프라 기술, 융·  
복합 방송/통신 분야 연구원

2009년~2012년: 나사렛대학교 멀티미디어학과 교수

2012년~현 재: 전남대학교 지능전자컴퓨터공학과 교수

※관심분야 : QoS/QoE 예측/분석/관리, 모바일 미디어 처리/  
통신, 클라우드 컴퓨팅 디지털 미디어 및 네트워  
크 지능