

## Improving On-board Aircraft Target Detection Model: Utilizing Shape Information and Mixed Precision Quantization

Junhyeong Bak\*, Yong-ho Choi\*, Jong-Won Moon\*, Bora Whang\*, Ingu Park\*\*

\*Research Engineer, IIR Seeker R&D, LIG Nex1 Co., Ltd., Yongin, Korea

\*\*Chief Research Engineer, IIR Seeker R&D, LIG Nex1 Co., Ltd., Yongin, Korea

### [Abstract]

The object detection technology in the field of computer vision is utilized to detect aircraft targets using electro-optical systems. Recent advancements in object detection, driven by deep neural networks and hardware accelerators, have led to significant performance improvements, which can also be applied to aircraft target detection. However, for successful integration into weapon systems, it is essential to achieve high detection reliability and real-time processing in on-board environments. This paper presents two strategies to address these challenges. These strategies focus on improving model training by utilizing object shape information and optimizing the model with mixed precision computation. The experimental results demonstrate that the methods improve the performance of the aircraft target detection model and enhance its applicability to weapon systems.

▶ **Key words:** Object detection, Aircraft target detection, Deep neural network, Data augmentation, Label assignment, NPU, Quantization

### [요 약]

항공 표적을 전자 광학 시스템으로 탐지하는 데에는 컴퓨터 비전 분야의 객체 탐지 기술이 이용된다. 최신의 객체 탐지 기술은 심층 신경망 기술과 이를 가속하는 하드웨어의 발전으로 성능이 비약적으로 향상되었는데, 이는 항공 표적 도메인에도 적용될 수 있다. 그러나 발전된 기술을 실제 무기 체계에 적용하기 위해서는 온-보드(On-board) 환경에서 높은 탐지 신뢰도와 실시간성을 달성할 수 있어야 한다. 본 논문에서는 이에 적합한 방법들을 두 가지 전략으로 구분하여 소개한다. 전략은 객체 형상 정보를 이용한 모델 훈련 개선과 복합 정밀도 연산 기반의 최적화로 요약될 수 있다. 실험 결과는 논문의 방법이 항공 표적 탐지 모델의 성능을 개선하고 무기 체계 적용성을 향상시킬 수 있음을 보여준다.

▶ **주제어:** 객체 탐지, 항공 표적 탐지, 심층 신경망, 데이터 증강, 라벨 할당, NPU, 양자화

- First Author: Junhyeong Bak, Corresponding Author: Junhyeong Bak  
\*Junhyeong Bak (junhyeong.bak@lignex1.com), IIR Seeker R&D, LIG Nex1 Co., Ltd.  
\*Yong-ho Choi (yongho.choi@lignex1.com), IIR Seeker R&D, LIG Nex1 Co., Ltd.  
\*Jong-Won Moon (jongwon.moon@lignex1.com), IIR Seeker R&D, LIG Nex1 Co., Ltd.  
\*Bora Whang (bora.hwang@lignex1.com), IIR Seeker R&D, LIG Nex1 Co., Ltd.  
\*\*Ingu Park (ingu.park2@lignex1.com), IIR Seeker R&D, LIG Nex1 Co., Ltd.
- Received: 2024. 10. 08, Revised: 2024. 11. 18, Accepted: 2024. 11. 27.

## I. Introduction

공중에서 아군의 전력을 공격하거나 각종 비전투 작전으로 적을 지원할 수 있는 항공 표적은 우선적으로 탐지되어야 하는 주요 군사 표적이다. 이를 탐지하기 위한 수단으로는 주로 레이더나 중적외선 대역의 전자 광학 시스템을 이용하는 데, 전자 광학 시스템은 영상 정보를 이용하므로 컴퓨터 비전 분야의 객체 탐지 기술이 적용된다.

객체 탐지 기술은 다양한 산업 분야에서 높은 활용도를 가지고 있어 경쟁적으로 연구가 이루어지고 있는 분야이다. 최근에는 심층 신경망 기술을 적용하여 성능이 비약적으로 발전했는데[1], 이후로 소개된 연구들도 대부분 심층 신경망 기술을 다루고 있다[2, ... 10].

심층 신경망을 이용하는 기술은 하드웨어 측면의 발전도 이루어지고 있다. 기존에는 고성능 그래픽 처리 장치로 연산을 가속하던 것을 신경망 연산에 특화된 전용 연산 장치로 가속하는 기술이 등장하였다[11]. 이러한 연산장치의 높은 연산 효율은 심층 신경망 모델을 저전력-저성능 컴퓨팅 하드웨어에서 직접 구동하는 온-보드 추론 기술을 실현할 수 있도록 만들어 주었다[12].

관련된 기술의 발전에도 불구하고 실전에 투입 가능한 온-보드 항공 표적 탐지 기술을 개발하는 것은 여전히 도전적인 문제이다. 무기 체계는 오탐지로 인한 작전 실패가 아군에게 치명적인 피해를 초래할 수 있으므로 다른 산업 분야보다 높은 탐지 신뢰도가 요구된다[13]. 그러나 도메인의 특수성으로 인한 훈련 데이터의 부족과 전장 환경에서 운용 가능한 하드웨어들의 낮은 성능은 탐지 신뢰도를 개선하는 데 걸림돌이 되고 있다.

본 논문에서는 앞서 식별한 제한 사항들을 완화하여 항공 표적 탐지 모델의 성능을 개선하는 방법을 소개한다. 본 논문의 방법은 크게 세 가지로 요약된다.

- 1) **증강 방법 개선:** 객체 형상 정보를 이용하는 개선된 훈련 데이터 증강 기법을 적용하여 부족한 훈련 데이터를 사실적이고 정교하게 증강한다.
- 2) **라벨(Label) 할당 방법 개선:** SimOTA[7]를 이용한 라벨 할당에 객체 형상 정보로부터 획득한 무게 중심 정보를 반영하여 손실 계산을 최적화한다.
- 3) **복합 정밀도 연산 기반의 양자화:** 양자화를 통한 모델 최적화 시 복합 정밀도를 적용하여 정밀도 변화로 인한 성능 하락을 완화한다.

1과 2의 방법은 부족한 훈련 데이터로부터 획득할 수 있는 유용한 정보를 늘려 모델 훈련을 개선하는 전략으로, 실험을 통하여 YOLOX-M[7] 모델의 평균 정밀도

(Average precision, AP)와 평균 재현율(Average recall, AR)을 약 10% 향상시킬 수 있음을 확인하였다. 3의 방법은 저성능 하드웨어에서의 동작을 최적화할 때 탐지 신뢰도의 하락을 완화하는 전략으로, 실험을 통하여 단일 정밀도 모델보다 높은 AP, AR 및 동작 속도를 동시에 달성할 수 있음을 확인하였다.

논문의 구성은 다음과 같다. 우선 2장에서 심층 신경망 기반 객체 탐지 기술과 이에 대한 최적화 기술의 기존 연구들을 소개한다. 3장에서는 제안하는 방법을 상세히 설명하고, 이어지는 4장에서는 이들 방법이 유효함을 확인할 수 있는 실험 결과를 제시한다. 마지막 5장에서는 제안하는 방법을 통한 성능 개선이 무기 체계 적용성을 향상시키는 데 도움이 되는지를 평가한다. 그리고 한계점을 바탕으로 향후 연구 방향에 대하여도 제시한다.

## II. Related Works

### 1. Object Detection via Deep Neural Network

심층 신경망을 적용한 객체 탐지 기술에 대한 선구적인 연구는 2014년에 소개된 R-CNN[1]으로 알려져 있다. R-CNN은 이미지에서 후보 영역을 추출한 후, 각 후보 영역에 합성곱 신경망을 개별적으로 적용하여 객체를 분류하도록 고안되었다. 이러한 방식은 영역별로 합성곱 신경망을 이용하므로 동작 속도가 느린데, 이후 등장한 Fast R-CNN[2]은 신경망 내에서 후보 영역을 공유하는 방식으로 속도 문제를 개선하였다. 후보 영역 추출에도 신경망을 적용하는 본격적인 신경망 기반 탐지 기술은 RPN(Region Proposal Network)이라는 이름으로 Faster R-CNN[3]에서 소개되었다. 이 기술은 검출과 분류를 위한 신경망이 구분되는 2단 탐지기의 원형으로 여겨지고 있다.

검출과 분류를 단일 신경망에서 수행하는 1단 탐지기는 YOLO[4]라는 이름으로 2016년에 소개되었다. 이 방식은 기존의 후보 영역 기반 탐지 모델들보다 탐지 신뢰도는 다소 떨어지지만 동작 속도 측면에서 유리하다는 장점이 있다. YOLO를 제안한 Redmon은 모델을 YOLOv3 까지 발전시켰는데, 이후에는 Ultralytics[5]와 같은 다른 연구자들이 이를 계승하고 있다.

YOLO의 발전 과정에서 특기할 만한 변화는 앵커-프리(Anchor-free) 구조의 도입이다. 해당 구조를 적용한 대표적인 초기 모델로는 FCOS[6]가 있는데, FCOS는 앵커 박스(Anchor box)를 정의하지 않고 객체의 중심을 기준으로 경계 상자를 직접 예측할 수 있음을 보여주었다. 이

러한 방식은 앵커의 개수가 감소하여 연산량이 줄어들고 앵커 박스의 사전 정의가 필요하지 않아 일반화 성능을 향상시키는 데 유리하다는 이점이 있다. 이후로 등장한 YOLO 파생 모델 중 YOLOX[7]는 FCOS에서 제안한 개념과 SimOTA 라벨 할당 방법을 조합하여 탐지 신뢰도와 동작 속도를 큰 폭으로 향상시키는 데 성공하였다.

본 논문의 연구 주제인 항공 표적 탐지에 중점을 둔 연구는 신경망 기반 탐지기의 발전과 함께 계속되어왔다. YOLOv4[8]가 등장한 시점에는 Hou 등[9]이 이를 연구했는데, 얇은 특징맵(Feature map)의 추출을 강화하고 다중 스케일 확장 합성곱 신경망과 반-잔차 블록을 이용하는 구조를 도입하여 작은 목표물의 검출 능력을 개선하였다. 유사한 전략을 Ultralytics의 YOLOv8에 적용한 연구는 Liu 등[10]이 제안하였다. 이 연구도 얇은 특징맵의 추출을 강화하였으며, 수용 영역을 유연하게 조정할 수 있는 합성곱 신경망을 도입하여 작은 크기나 중간 크기 목표물의 검출 능력을 개선하였다.

**2. Optimization of Object Detection Model**

심층 신경망 기반 탐지 모델을 활용하는 여러 분야 중에는 높은 실시간성과 저전력-저성능 환경에서 실행 가능한 모델이 요구되는 경우가 있다. 이러한 조건을 충족하기 위해서는 모델 최적화를 통하여 연산 복잡도와 컴퓨팅 자원의 점유율을 낮추어야 한다.

탐지 모델의 최적화는 다양한 접근법이 존재한다. 우선적으로 고려할 수 있는 접근법은 신경망의 구조를 개선하는 것이다. 신경망 구조 개선과 관련된 선구적인 연구로는 MobileNet[14]의 논문이 잘 알려져 있다. 논문은 깊이별 분리 합성곱 구조를 통하여 모델의 파라미터 수와 연산 복잡도를 낮추는 방법을 제안하였는데, 실험 결과로 MobileNet 구조를 적용한 탐지 모델의 성능을 제시하였다. 그러나 MobileNet 구조를 통한 최적화는 현재의 기준으로는 탐지 신뢰도가 크게 손실되기 때문에 최신의 연구에서는 거의 사용되지 않는다. 보다 진보된 방법으로는 추론 정확도에 기여가 적은 합성곱 채널을 배제하는 채널 가지치기 기법이 있으며, 탐지 모델에 이를 적용하는 연구는 Zhang 등[15]에 의하여 소개되었다.

크기가 작은 신경망 모델의 훈련을 최적화하는 접근법도 존재한다. 대표적으로 지식 증류 기법이 있는데, 이는 큰 크기의 신경망 모델에서 학습한 지식을 작은 크기의 신경망 모델이 효율적으로 학습하도록 돕는 기술이다. 탐지 모델에 이를 적용하는 연구는 Chen 등[16]에 의하여 소개되었다.

최근 주목받고 있는 접근법으로는 모델 연산의 정밀도를 낮추는 일명 양자화 방법이 있다. 일반적으로 심층 신경망 모델은 32비트 부동 소수점으로 연산을 수행하는데, 양자화를 거치면 이를 4비트까지도 낮출 수 있다. 양자화 기술을 탐지 모델에 적용하여 성능을 최적화한 최신 연구로는 Wang 등[17]의 연구가 알려져 있다. 현재 양자화 방법은 인공지능 연산 전용 하드웨어인 Neural processing unit(이하 NPU)가 등장하면서 이전보다 많은 관심을 받고 있다. NPU는 구조상 정수 연산에 최적화되어 있어 양자화된 모델의 연산을 가속하는 데 적합하기 때문이다. 양자화된 탐지 모델의 NPU 기반 가속에 대한 연구는 본 논문의 저자가 참여한 Hur 등의 연구[12]에서도 다루어졌다.

**III. Methods**

**1. Improvement of Data Augmentation**

**1.1 Configuration of Data Augmentation**

항공 표적은 데이터 도메인의 특수성으로 인하여 다량의 데이터를 확보하기 어렵다. 이를 극복하기 위해서는 강력한 데이터 증강 방법이 필수적이다. 따라서 증강 방법은 여러 가지 기법들을 조합하여 구성한다. 이때 증강 기법은 검출기 관점에서 사실성을 고려하여 검출기 외부에 대한 증강, 데이터를 혼합하는 증강, 검출기 내부에 대한 증강으로 분류한 후 Fig. 1의 순서대로 적용한다.

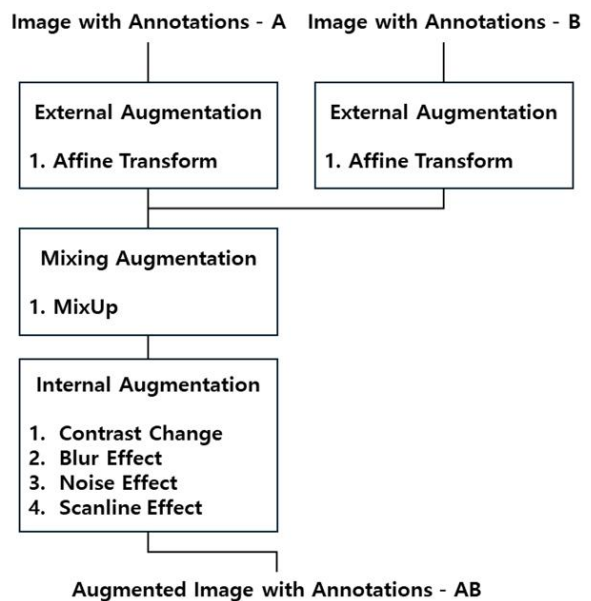


Fig. 1. Flowchart of The Augmentation Method

사용하는 증강 기법들은 주요 탐지 모델들을 통하여 잘 알려진 기법들인데, 예외적으로 스캔라인(Scanline) 효과는 적외선 영상에서 주로 발생하는 현상을 반영한 기법이다. 스캔라인은 검출기의 물리적 감도가 균일하지 않아 발생하는 세로선 효과로, 해당 효과를 증강한 결과는 Fig. 2와 같다.

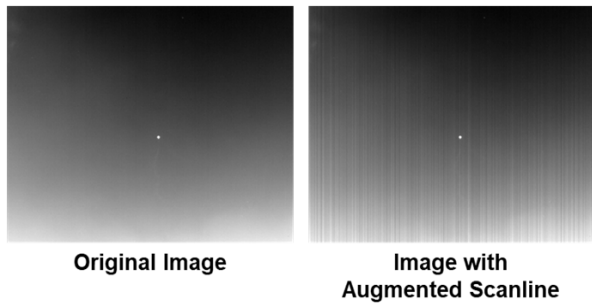


Fig. 2. The Effect of Scanline Augmentation

증강 기법 중 아핀(Affine) 변환과 MixUp은 객체 형상 정보를 활용하도록 개선된 방법을 이용한다. 이는 1.2와 1.3에서 상세히 서술한다.

1.2 Improvement of Affine Transform

경계 상자 정보만을 이용하는 기존의 아핀 변환은 객체를 회전할 때 경계 상자의 왜곡이 발생할 수 있다는 문제가 있다. 이는 객체와 함께 회전된 경계 상자의 꼭지점이 객체의 외곽이 아닐 수 있기 때문에 발생하는 현상이다. 일반적인 탐지 모델의 훈련에서는 이러한 왜곡은 무시되는 경향이 있다. 그러나 적외선 항공 표적 도메인에서는 표적 객체의 형상으로 인하여 왜곡이 심한 경향이 있으므로 반드시 왜곡을 제거하는 것이 유리하다. 이를 위하여 본 논문의 방법은 아핀 변환 수행 시 객체의 형상 정보인 객체 분할 정보(Segmentation)를 대신 변환한 후 새롭게 객체의 외곽을 판단하여 경계 상자를 추출한다. 기존 아핀 변환과 개선된 방법을 비교한 결과는 Fig. 3과 같다.

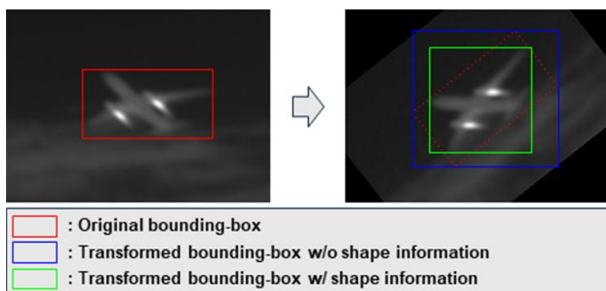


Fig. 3. Comparison of a Conventional Affine Transform and the Improved Method

1.3 Improvement of MixUp

두 영상의 전체를 알파 블렌딩(Alpha blending)으로 융합하는 기존의 MixUp은 두 영상의 평균 영상을 도출하기 때문에 영상 정보가 모호해지는 경향이 있다. 이는 융합을 여러 번 수행할수록 심화되는 현상이기 때문에 MixUp은 데이터 당 한 번만 수행하는 것이 일반적이다. 그러나 전체 데이터에서 차지하는 비중이 적은 다중 표적 데이터를 증강을 통하여 획득하기 위해서는 MixUp을 여러 번 수행하는 것이 필요하다.

본 논문에서는 한 영상에 다른 영상의 객체만을 정교하게 융합하는 MixUp을 적용하여 위 문제를 해결한다. 융합시 객체 영역을 정교하게 선택하는 것은 객체 형상 정보인 객체 분할 정보를 이용하여 수행할 수 있으며, 객체 가장자리가 자연스럽게 처리되도록 융합하는 것은 포아송 블렌딩(Poisson blending)을 이용한다[18]. 개선된 MixUp은 MixUp 수행 시 배경부의 사실성은 훼손되지 않으며, 객체 영역도 비교적 높은 사실성을 유지한다. 또한 한 영상 위에 MixUp을 여러 차례 수행하여도 영상이 모호해지지 않는다. 기존 MixUp과 개선된 MixUp의 비교는 Fig. 4를 통하여 확인할 수 있다.

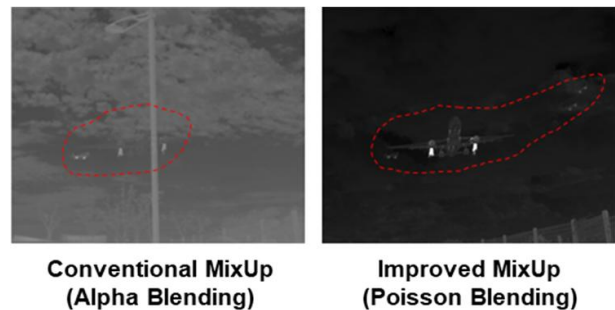


Fig. 4. Comparison of a Conventional MixUp and the Improved Method

2. Improvement of Label Assignment

YOLOX[7]를 통하여 소개된 SimOTA는 대표적인 추정-라벨 할당 방법 중 하나로, 라벨과 앵커의 공간적 근접성을 고려하여 할당 후보를 제한한다. 이는 객체와 공간적으로 가까운 앵커들만 활성화함으로써 손실 계산 과정의 연산량을 줄이고 할당 정확도를 높여준다. 이때 근접성을 판단하는 기준은 객체 경계 상자의 중심점을 이용한다. 그러나 본 논문에서는 추가적인 기준으로 객체 형상 정보로부터 획득한 무게 중심을 동시에 이용한다. 추가적인 기준은 표적 기체의 몸통과 같이 실제 객체 영역의 핵심 위치에 있는 앵커들에게 라벨을 할당하는 데 유리하다. 이는 Fig.

5의 예시로도 확인할 수 있다. 훈련을 통하여 이러한 위치의 앵커들도 표적을 추정하게 만드는 것은 기존 기준에 의하여 활성화되는 앵커들의 추정을 보완하는 데 도움을 줄 수 있다.

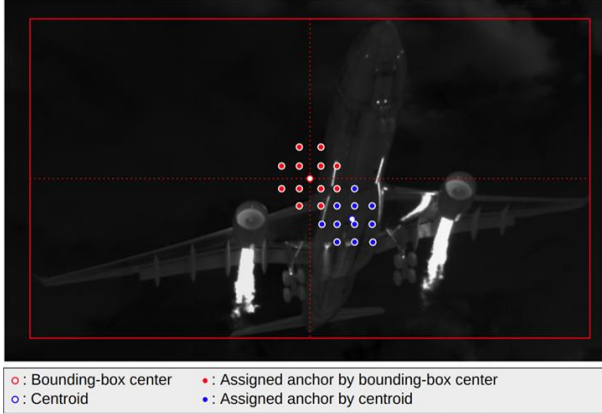


Fig. 5. Visualization Result of the Improved SimOTA

기존 기준과 추가적인 기준을 모두 이용하는 근접성 판단 수식은 다음과 같다. 수식에서  $(x_{anchor}, y_{anchor})$ 는 앵커의 좌표,  $(x_{center}, y_{center})$ 는 객체 중심점의 좌표,  $(x_{centroid}, y_{centroid})$ 는 무게 중심의 좌표를 의미하고  $d$ 는 유효 반경을 의미한다.

$$(c_1 \wedge c_2) \vee (c_3 \wedge c_4)$$

$$\begin{aligned} c_1 &: x_{center} - d \leq x_{anchor} \leq x_{center} + d \\ c_2 &: y_{center} - d \leq y_{anchor} \leq y_{center} + d \\ c_3 &: x_{centroid} - d \leq x_{anchor} \leq x_{centroid} + d \\ c_4 &: y_{centroid} - d \leq y_{anchor} \leq y_{centroid} + d \end{aligned}$$

### 3. Mixed Precision Quantization

훈련된 모델을 온-보드 환경에서 실행 가능한 모델로 변환하기 위해서는 양자화를 통한 최적화가 필요하다. 양자화는 일반적으로 16비트, 8비트, 4비트 정수를 이용하는데, 정밀도가 낮을수록 실시간성은 향상되지만 탐지 신뢰도의 하락은 커지는 상충 관계가 있다. 그러나 정밀도가 미치는 영향이 신경망의 모든 부분에서 동일하지는 않다. 따라서 정밀도가 중요한 부분에는 16비트 양자화를, 그렇지 않은 부분에는 8비트, 4비트 양자화를 적용하는 복합 정밀도 양자화를 수행하면 최적의 양자화 결과를 얻을 수도 있다.

따라서 본 논문에서는 다음의 규칙으로 정밀도를 결정한다. 먼저 신경망 헤드(Head)의 말단 레이어(Layer)들은 최종 추정 결과와 직결되는 부분이므로 16비트 정밀도를 적용한다. 나머지 레이어들 중 연산 텐서(Tensor)의 크기

가 임계값 보다 큰 레이어들은 양자화를 통한 연산 복잡도 감소의 이득이 크다고 판단하여 4비트를 적용한다. 해당 레이어들은 대체로 깊은 신경망의 중앙에 위치하므로 낮은 정밀도 연산의 영향이 희석될 가능성도 높다고 볼 수 있다. 나머지 레이어들은 모두 8비트를 적용한다.

## IV. Experiments

### 1. Experimental Environment

본 논문의 방법을 적용할 탐지 모델로는 YOLOX-M[7]을 선정하였다. 모델의 훈련과 평가에 사용한 데이터는 FLIR A6701[19]로 직접 촬영한 40,906 장의 640 x 512 크기 중적외선 영상이며, Anylabeling[20]을 이용하여 라벨링(Labeling)하였다. 라벨링 수행 시 카테고리는 고정익기와 회전익기 두 가지로 구분하였고, 객체 분할 정보는 MobileSAM[21]을 이용하여 반자동으로 취득하였다. 훈련 데이터의 양이 COCO[22]와 같은 대규모 이미지 데이터 세트와 비교하여 적기 때문에 모델 훈련 에포크(Epoch)는 30으로 설정하였다. 훈련의 마지막 에포크에는 데이터의 사실성을 훼손할 수 있는 증강은 수행하지 않도록 설정하였다. 기타 설정은 YOLOX의 것을 준용하였다. 최적화 이후의 온-보드 실험은 Hailo-8 NPU[11]를 1개 탑재한 Raspberry Pi 5[23] SBC(Single board computer)에서 수행하였다.

### 2. Benchmark Methods

탐지 신뢰도는 COCO에서 사용되는 공식 평가 기준 및 도구를 이용하여 측정하였다. COCO에서 AP와 AR은 모든 카테고리에 대한 평균으로 계산하며, IoU=50:05:95를 적용한다.  $AP_{small}$ ,  $AR_{small}$ 은 객체의 픽셀 면적이  $32^2$ 보다 작은 객체에 대한 AP, AR이며,  $AP_{large}$ ,  $AR_{large}$ 은 그 반대이다. 동작 속도와 소비 전력은 탑재 NPU의 제조사인 Hailo에서 제공하는 벤치마크 도구로 측정하였다. 벤치마크 도구는 동작 속도를 초당 처리한 영상 프레임의 수(Frame per second, FPS)로 측정하고, 소비 전력을 동작 시간 동안의 평균 전력(Average Watt)으로 측정한다.

### 3. Performance Based on Model Training

제안하는 모델 훈련 개선 전략의 효과를 정량적으로 확인한 결과는 Table. 1과 같다. 논문의 방법을 하나씩 추가할 때마다 탐지 성능은 모든 지표에서 상승하였음을 표에서 확인할 수 있다. 논문의 방법 중 성능을 가장 큰 폭으로

Table 1. Quantitative Performance Based on Model Training Methods

Method	AP	AP <sub>small</sub>	AP <sub>large</sub>	AR	AR <sub>small</sub>	AR <sub>large</sub>
YOLOX-M	43.0	42.6	86.2	48.1	47.5	89.5
+ Affine Transform Improvement	45.4	44.7	92.9	49.8	49.1	94.8
+ MixUp Improvement	46.5	45.6	94.4	52.1	51.5	95.8
+ Label Assignment Improvement	47.5	46.6	94.5	52.8	52.1	95.9
Ultralytics YOLOv8-M	42.6	41.6	95.7	46.8	46.1	97.0
Ultralytics YOLO11-M	43.5	42.4	95.6	47.9	47.2	97.1

향상시킨 것은 개선된 아핀 변환이다. 특히 아핀 변환은 기존 방법의 문제가 부각되는 큰 객체의 탐지 성능을 비약적으로 향상시켰다. 다른 방법 중 개선된 MixUp은 AR의 향상이 두드러지는데, 이는 놓치는 표적이 크게 감소함을 의미한다. 배경의 클러터(Clutter)와 구분이 어려운 객체들은 모델의 검출 대상에서 제외되기 쉬운데, 개선된 MixUp을 통한 다중 표적 증강이 이를 완화하여 놓치는 표적이 감소한 것으로 추정된다. 마지막으로 적용된 방법인 라벨 할당 개선은 작은 객체에 대하여 유의미한 성능 향상이 종합적인 성능 향상을 이끌어냈다. 다만 큰 객체에 대하여는 상대적으로 효과가 적었는데, 이는 제안하는 방법으로 추가 활성화될 수 있는 앵커들이 없어도 큰 객체는 충분히 정확한 탐지가 가능하기 때문으로 추정된다.

개선된 YOLOX-M 모델과 비교할 수 있는 탐지 모델의 성능도 Table. 1에서 확인할 수 있다. 비교 모델은 Ultralytics[5]가 공개한 YOLOv8-M, YOLO11-M이며, 각각은 본 논문의 훈련 데이터로 재훈련되었다. 실험 결과를 보면 개선된 YOLOX-M은 비교 모델보다 종합적으로 우수한 성능을 보여주었는데, 특히 작은 객체에 대하여 격차가 컸다. 큰 객체의 경우에는 비교 모델의 성능이 소폭 우위를 차지했지만, 항공 표적의 특성상 전체 데이터에서 큰 객체가 차지하는 비중이 적으므로 종합적인 성능에 미친 영향은 미미했다.

훈련 개선 전략의 효과를 정성적으로 확인한 결과는 Fig. 6의 시각화 결과와 같다. 비교는 논문의 방법을 적용하지 않은 모델과 모두 적용한 모델에 대하여 수행하였다. 표시된 글자는 카테고리 추정 결과이며, 괄호 안의 숫자는 객체성과 카테고리 추정 확률을 곱하여 획득한 추정 신뢰도를 퍼센트로 나타낸 것이다. 논문의 방법을 모두 적용한 모델은 전체적으로 표적에 매우 근접한 결과를 추정하는 것을 확인할 수 있다. 반면 논문의 방법을 적용하지 않는 모델은 경계 상자 오차가 상대적으로 크게 나타나고 오검출이 발생하는 경우도 확인할 수 있었다. 추정 신뢰도 또한 논문의 방법을 모두 적용했을 때의 결과가 그렇지 않은 경우의 결과보다 더 높은 것을 확인할 수 있다.

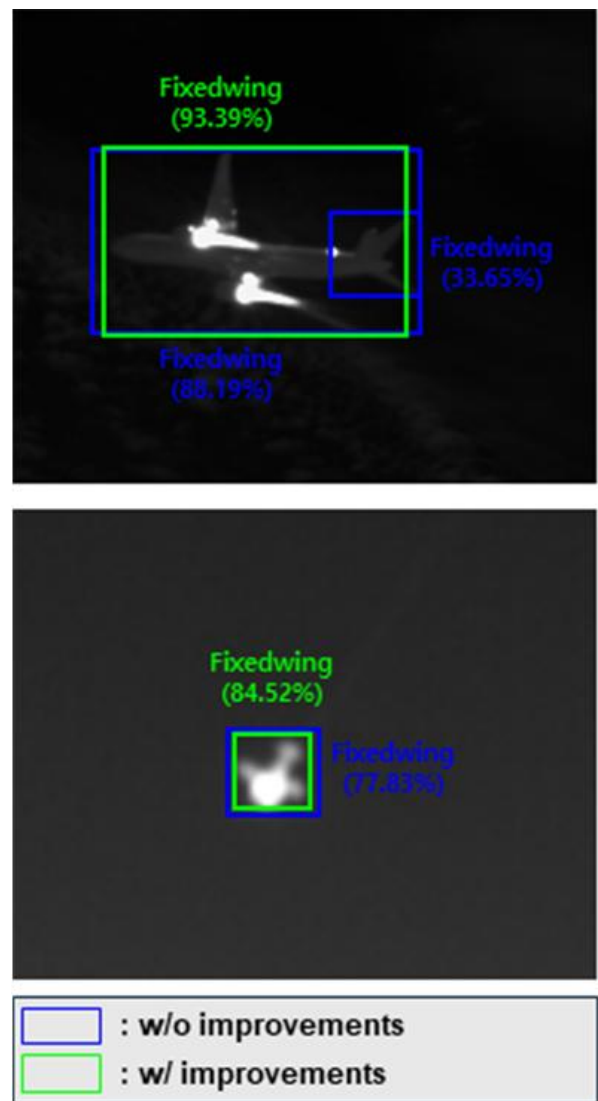


Fig. 6. Qualitative Performance Based on Model Training Methods

#### 4. Performance Based on Optimization

제안하는 최적화 전략의 효과를 정량적으로 확인한 결과는 Table. 2와 같다. 실험 결과에서 논문의 복합 정밀도 양자화 방법을 모두 수행한 모델은 탐지 신뢰도와 동작 속도 측면 모두에서 가장 높은 성능을 기록하여 최적화 효과가 입증되었다. 반면 8비트 단일 정밀도 양자화를 수행한

모델은 정반대의 결과가 나왔다. 소비 전력은 복합 정밀도를 적용한 모델이 조금 더 높은 것을 확인할 수 있는데, 차이는 약 0.1 W 수준으로 크지 않았다.

Table 2. Quantitative Performance Based on Quantization Methods

Method	AP	AR	FPS	Avg. W
YOLOX-M (INT8)	40.0	45.6	27.5	1.52
+ End-node INT16	43.1	48.4	28.5	1.61
+ Huge Tensor INT4	43.5	48.8	28.9	1.64

## V. Conclusions

본 논문에서는 항공 표적 탐지 모델의 무기 체계 적용성을 높이기 위한 성능 개선 방법을 연구하였다. 방법은 두 가지 전략으로 구성된다. 첫 번째는 부족한 훈련 데이터로부터 획득할 수 있는 유용한 정보를 늘려 모델 훈련을 개선하는 전략이다. 유용한 정보를 추가 획득하기 위하여 객체 형상 정보를 활용한다. 두 번째는 저성능 하드웨어에서의 동작을 최적화할 때 탐지 신뢰도의 하락을 완화하는 전략이다. 최적화는 양자화를 통하여 수행하는데, 복합 정밀도 연산을 적용하여 정밀도 변화로 인한 성능 하락을 완화한다.

제안하는 전략이 유효함과 각 방법이 기여하는 정도는 정량적, 정성적인 실험을 통하여 확인하였다. 논문의 방법을 적용한 YOLOX-M[7] 모델은 자체 데이터 세트에 대하여 AP는 47.5, AR은 52.8을 달성하였는데, 이는 최신 YOLO 모델인 YOLO11-M[5]이 기록한 수치보다 높다. 온-보드 실행을 위하여 최적화된 모델은 단일 정밀도로 양자화 된 모델보다 AP, AR 및 동작 속도가 모두 높게 나왔음에도 소비 전력은 약 0.1 W 증가하는 데 그쳤다. 특히 동작 속도는 실시간의 기준이 되는 30 FPS에 근접한 28.9 FPS를 대표적인 SBC인 Raspberry Pi 5[23]에서 달성하였다. 결과적으로 본 논문의 방법은 무기 체계 적용성을 판단할 때 참고할만한 주요 성능 지표를 전반적으로 개선하는 데 도움이 된다.

다만 본 연구는 다음의 한계점들을 가진다. 먼저 훈련 및 평가에 사용한 데이터의 절대적인 수량 부족으로 탐지 모델이 과적합되는 양상이 일부 관찰되었다. 해당 문제에 대하여는 컴퓨터 그래픽스 기반의 적외선 영상 합성 기술

로 부족한 데이터를 보충하는 연구를 진행하고 있다. 또한 모델 훈련 및 최적화 과정에서 최신 QAT(Quantization Aware Training)[24] 기술을 적용한다면 양자화 이후의 성능을 현재보다 개선시킬 수 있는 여지가 있다. 따라서 추후에는 항공표적 탐지 모델에 적용 가능한 QAT 기술을 연구하여 모델 최적화 효과를 더욱 향상시킬 계획이다.

## REFERENCES

- [1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, June 2014. DOI: 10.1109/CVPR.2014.81
- [2] R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision, pp. 1440-1448, December 2015. DOI: 10.1109/ICCV.2015.169
- [3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 6, pp. 1137-1149, June 2017. DOI: 10.1109/TPAMI.2016.2577031
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, June 2016. DOI: 10.1109/CVPR.2016.91
- [5] Ultralytics, <https://www.ultralytics.com>
- [6] Z. Tian, C. Shen, H. Chen, and T. He, "FCOS: Fully Convolutional One-Stage Object Detection," Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9627-9636, October 2019. DOI: 10.1109/ICCV.2019.00972
- [7] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO Series in 2021," arXiv preprint arXiv:2107.08430, July 2021. DOI: 10.48550/arXiv.2107.08430
- [8] A. Bochkovskiy, C.-Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint, April 2020. DOI: 10.48550/arXiv.2004.10934
- [9] X. Hou, J. Ma, and S. Zang, "Airborne Infrared Aircraft Target Detection Algorithm Based on YOLOv4-tiny," Proceedings of the 2021 International Conference on Advances in Optics and Computational Sciences, January 2021. DOI: 10.1088/1742-6596/1865/4/042007
- [10] Y. Liu, N. Li, L. Cao, Y. Zhang, X. Ni, X. Han, and D. Dai, "Research on Infrared Dim Target Detection Based on Improved YOLOv8," Remote Sensing, Vol. 16, No. 16, August 2024. DOI: 10.3390/rs16162878
- [11] Hailo-8 NPU, <https://hailo.ai/products/ai-accelerators/hailo-8-m2>

-ai-acceleration-module/

- [12] D.-H. Hur, D.-H. Park, J.-Y. Baek, J.-H. Bak, and S.-H. Bae, "Object Detection Performance Analysis between On-GPU and On-Board Analysis for Military Domain Images," *Journal of the Korea Society of Computer and Information*, Vol. 29, No. 8, pp. 157-164, August 2024. DOI: 10.9708/jksci.2024.29.08.157
- [13] H. Jung, S. Park, and D.-Y. Lee, "Challenges and Implications of Testing and Evaluation of AI Autonomous Weapon Systems," *KIDA Defense Issues & Analyses*, No. 2005, pp. 24-32, 204, August 2024.
- [14] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *arXiv preprint arXiv:1704.04861*, April 2017. DOI: 10.48550/arXiv.1704.04861
- [15] P. Zhang, Y. Zhong, and X. Li, "SlimYOLOv3: Narrower, Faster and Better for Real-Time UAV Applications," *IEEE/CVF International Conference on Computer Vision Workshop*, pp. 37-45, October 2019. DOI: 10.1109/iccwv.2019.00011
- [16] G. Chen, W. Choi, X. Yu, T. Han, and M. Chandraker, "Learning Efficient Object Detection Models with Knowledge Distillation," *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 742-751, December 2017. DOI: 10.5555/3294771.3294842
- [17] M. Wang, H. Sun, J. Shi, X. Liu, B. Zhang, and X. Cao, "Q-YOLO: Efficient Inference for Real-time Object Detection," *arXiv preprint arXiv:2307.04816*, July 2023. DOI: 10.48550/arXiv.2307.04816
- [18] P. Pérez, M. Gangnet, and A. Blake, "Poisson Image Editing," *ACM Transactions on Graphics*, Vol. 22, No. 3, pp. 313-318, July 2003. DOI: 10.1145/882262.882269
- [19] FLIR A6701, <https://www.flirkorea.com/products/a6700-mwir/?vertical=rd+science&segment=solutions>
- [20] AnyLabeling, <https://anylabeling.nl.ai/>
- [21] C. Zhang, Y. Zheng, J. Lin, R. Wang, Y. Lu, and Y. Liu, "Faster Segment Anything: Towards Lightweight SAM for Mobile Applications," *arXiv preprint arXiv:2306.14289*, June 2023. DOI: 10.48550/arXiv.2306.14289
- [22] COCO, <https://cocodataset.org/>
- [23] Raspberry Pi 5, <https://www.raspberrypi.com/products/raspberry-pi-5/>
- [24] B. Jacob, S. Kligys, B. Chen, M. Zhu, N. Tang, A. Howard, M. Adam, and D. Kalenichenko, "Quantization and Training of Neural Networks for Efficient Integer-Arithmetic-Only Inference," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2704-2713, June 2018. DOI: 10.1109/CVPR.2018.00286

## Authors



Junhyeong Bak received his B.S. degree in Electronics Engineering from Hannam University in 2020, and M.S. degree in Electrical and Computer Engineering from Inha University in 2022. He is currently a

research engineer at IIR Seeker R&D, LIG Nex1 Co., Ltd. He is interested in fast visual object detection and tracking on low-power hardware.



Yong-ho Choi received M.S. degree in the Department of Intelligent Mechatronics Engineering, Sejong University, Seoul, 05006, Korea, and the B.S. degree from the School of Intelligent Mechatronics Engineering at

Sejong University. He is currently a research engineer at IIR Seeker R&D, LIG Nex1 Co., Ltd. He is interested in fast visual object detection and tracking on low-power hardware.



Jong-Won Moon received his B.S. degree in Electronic Engineering from Paichai University in 2021, and M.S. degree in Department of Intelligent Mechatronics Engineering from Sejong University in 2023.

He is currently a research engineer at IIR Seeker R&D, LIG Nex1 Co., Ltd. He is interested in multi-object detection, tracking and on-board AI.



Bora Whang received the B.S. degree in Electronic Engineering from Dankook University in 2016. She is currently a research engineer at IIR Seeker R&D, LIG Nex1 Co., Ltd.

She is interested in integrated control and has experience in RTOS-based systems and graphical user interface application development.



Ingu Park received the B.S., in Computer Science and Engineering and M.S. degrees in Computer Engineering from Pusan National University, Korea, in 2004 and 2006, respectively.

Mr. Park is currently a chief research engineer in IIR Seeker R&D, LIG Nex1. He is interested in embedded system, and object detection & tracking using deep learning.