

일반논문 (Regular Paper)

방송공학회논문지 제28권 제6호, 2023년 11월 (JBE Vol.28, No.6, November 2023)

<https://doi.org/10.5909/JBE.2023.28.6.801>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 컴퓨터 그래픽스 기반의 합성 데이터를 이용한 항공표적 탐지 모델 학습 방법

박준형<sup>a)\*</sup>, 최용호<sup>a)</sup>, 문종원<sup>a)</sup>, 박인구<sup>a)</sup>, 이기남<sup>a)</sup>

### Method for Aerial Target Detection Model Training with Computer-Graphics-Based Synthetic Data

Junhyeong Bak<sup>a)\*</sup>, Yong-ho Choi<sup>a)</sup>, Jong-Won Moon<sup>a)</sup>, Ingu Park<sup>a)</sup>, and Ki-Nam Lee<sup>a)</sup>

#### 요약

중적외선 영상을 이용한 항공표적의 탐지는 군사적인 효용성이 높은 기술이지만, 데이터 확보의 문제로 인하여 딥러닝 기술을 적용하는 데 어려움이 있다. 본 연구에서는 이를 극복하기 위하여 컴퓨터 그래픽스 기반의 합성 데이터를 이용하는 항공표적 탐지 모델 학습 방법을 제안한다. 학습은 컴퓨터 그래픽스로 합성한 가시광 영역의 항공표적 데이터로 사전학습을 수행하고, 실제로 획득한 중적외선 영역의 데이터를 학습하도록 구성된다. 사전학습은 실제 중적외선 영상과 유사한 특성의 영상을 생성할 수 있는 데이터 증강 기법을 포함한다. 모델 학습과 평가에 이용하는 실제 항공표적 데이터는 중적외선 카메라를 이용하여 직접 획득하였다. 실험 결과는 제안하는 방법이 자체 데이터 셋에 대하여 mAP는 6.9%, mAR은 7.5% 향상시킬 수 있음을 보여준다.

#### Abstract

The detection of aerial targets using mid-wave infrared image is a highly valuable military technology, but it is difficult to apply deep learning due to data acquisition problems. To address this issue, We propose a training method for an aerial target detection model with computer-graphics-based synthetic data. The training process is structured to do pre-training with computer-graphics-based visible-band aerial target data and fine tuning through actual infrared-band data. The pre-training includes a data augmentation technique capable of generating an image with similar characteristics to an actual infrared image. The actual aerial target data for model training and evaluation have been directly acquired through an infrared camera. Experimental results show that the proposed method can improve mAP by 6.9% and mAR by 7.5% on our own dataset.

Keyword : Object Detection, Infrared Image, Computer Graphics, Data Synthesis, Data Augmentation

Copyright © 2023 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

## I. 서론

표적 탐지 기술은 무인화 및 지능화된 무기체계를 개발 하는데 필요한 핵심기술로, 우리 군에서도 많은 관심을 가지고 있다. 표적이 되는 대상은 무기체계에 따라 다양한데, 그 중에서도 전략적으로 가치가 높은 것은 항공표적이다. 항공표적은 군의 주요 자산을 정찰하거나 파괴할 수 있고 높은 살상력을 지닌 무기를 투발할 수 있으므로 반드시 포착하고 차단할 대상이기 때문이다.

최근에는 딥러닝과 같은 혁신적인 기술이 탐지 분야에서 우수한 성능을 발휘하고 있기 때문에 항공표적의 탐지에도 이를 적용하고자 하는 시도가 있다. 그러나 항공표적은 딥러닝 모델을 학습하는 데 필요한 충분한 양의 데이터를 확보하기 어렵다. 또한 실제 무기체계는 주-야간 구분 없이 운용이 가능하고 항공기 엔진의 열에너지를 감지하는 데 적합한 중적외선 영역의 영상을 주로 이용하는 특수성 때문에 데이터 확보의 어려움이 심화된다.

본 연구에서는 이를 극복하기 위하여 컴퓨터 그래픽스 기반의 학습 데이터를 이용하는 항공표적 탐지 모델 학습 방법을 제안한다. 제안하는 방법에서 컴퓨터 그래픽스 기반의 학습 데이터는 고화질의 환경맵에 항공기 3차원 모델을 배치하고 렌더링하여 가시광 영역의 영상으로 합성한다. 환경맵은 사실적이고 다양한 하늘의 묘사를 포함하므로 양질의 데이터를 합성하는데 도움을 준다. 3차원 모델은 여러 종류의 항공기 모델을 사용하고, 카메라의 시야 안에서 무작위하게 배치하므로 다양한 형상이 나타난다. 탐지 모델은 일반적인 딥러닝 기반 탐지 모델을 이용할 수 있으며, 중적외선 영상을 모델 신경망에 입력하기 위한 영상 전처리를 포함한다. 탐지 모델의 학습은 합성 데이터로 사전학습을 수행한 후, 실제로 획득한 중적외선 영역의 데이터를 학습하도록 구성된다. 사전학습은 실제 중적외선 영상과 유사한 특징을 가지는 영상을 생성할 수 있는 데이터 증강

기법을 포함한다. 이를 통하여 탐지 모델은 실제 데이터에 나타나는 특징을 검출할 수 있는 능력을 사전학습 단계에서도 일부 학습할 수 있다. 모델 학습과 평가에 필요한 실제 항공표적 데이터는 보유하고 있는 중적외선 카메라를 통하여 직접 획득하였다.

제안하는 방법은 합성 기반의 학습 데이터 확보 과정에서 전문적인 지식이나 정교한 작업이 필요하지 않기 때문에 항공표적 탐지 모델 개발에 투입되는 시간과 비용을 줄일 수 있다. 뿐만 아니라 충분한 데이터 확보가 어려운 표적 유형에 대한 탐지 능력을 획득하는데 도움을 줄 수 있다. 본 논문의 실험결과는 제안하는 방법으로 학습된 모델과 그렇지 않은 모델의 정량 및 정성적인 탐지 성능 비교를 통하여 이를 뒷받침한다. 논문의 결론에서는 앞선 실험결과를 바탕으로 연구 결과를 고찰하고 연구의 한계점과 향후 연구방향에 대하여 논의한다.

## II. 관련연구

### 1. 딥러닝 기반의 탐지 모델

딥러닝 기술은 심층 신경망을 통하여 컴퓨터 스스로 특징 또는 패턴을 학습하는 방법으로, 뛰어난 일반화 능력을 가지고 있는 것으로 알려져 있다. 컴퓨터가 스스로 학습하는 능력은 새로운 특징에 대한 모델 개발을 용이하게 한다는 이점도 가지고 있다. 탐지 분야에서 딥러닝 기술은 R-CNN<sup>[1]</sup>이 등장하면서 본격적으로 연구되었다. R-CNN은 객체가 있을 것으로 예상되는 영역을 먼저 선택하고, 각 영역의 카테고리를 개별적으로 분류하는 2단 구조를 제안하였다. 객체의 위치까지 신경망을 통하여 추정하는 기술은 Region proposal network (RPN)<sup>[2]</sup>가 등장하면서 가능하게 되었다. RPN은 미리 설정해둔 다수의 Anchor를 통하여 객체 후보를 검출한 뒤, 유력한 객체의 경계상자만 남게 하는 후처리를 통하여 객체의 위치를 검출할 수 있다. 객체의 위치 추정과 분류를 단일 신경망으로 추정하는 1단 구조는 이후 등장한 YOLO<sup>[3]</sup>에서 제안되었다. YOLO는 간략화 된 구조를 가지고 있으므로 실시간 객체 탐지 분야에서 주로 연구되고 있는 방식이다.

a) LIG넥스원, PGM탐색기연구소(PGM Seeker R&D)

‡ Corresponding Author : 박준형(Junhyeong Bak)

E-mail: junhyeong.bak@lignex1.com

Tel: +82-31-326-9478

ORCID: <https://orcid.org/0009-0009-4006-2357>

· Manuscript September 26, 2023; Revised November 7, 2023;

Accepted November 7, 2023.

초기의 딥러닝 기반 탐지 모델 연구들은 객체 위치를 추정하기 위하여 RPN의 Anchor 방식을 참고하였다. 그러나 Anchor의 크기는 예상되는 탐지 대상의 크기를 통계적으로 분석하여 설정하기 때문에 너무 크거나 작은 객체는 검출하기 어렵다는 문제가 있다. 또한 RPN의 후처리 과정은 객체 후보의 개수가 많은 경우 처리속도를 증가시키는 병목으로 작용하기도 한다. 이러한 문제에 착안하여 Corner Net<sup>[4]</sup>은 Anchor-free 방식을 제안하였다. 같은 방식을 적용한 최신 탐지 모델인 YOLOX<sup>[5]</sup>가 State-of-the-art 성능을 보여주고 있는 만큼, 향후에는 Anchor-free 방식을 적용한 연구가 더욱 늘어날 것으로 기대된다.

## 2. 관련 데이터 셋

항공표적 데이터는 항공기 영상 데이터의 한 종류로 볼 수 있다. 항공기는 지상에서 쉽게 포착할 수 없기 때문에 데이터 셋 구축이 어려운 대상이다. 대표적인 대규모 영상 데이터 셋인 COCO<sup>[6]</sup>의 경우에도 항공기를 포함하는 영상은 약 3,000장에 불과하다.

데이터 획득의 어려움에도 불구하고 가시광 영역에 대하여는 FGVC-Aircraft 데이터 셋<sup>[7]</sup>이나 Airbus aircraft detection 데이터 셋<sup>[8]</sup>과 같이 많은 연구자들에게 알려진 데이터 셋이 존재한다. FGVC-Aircraft 데이터 셋은 항공 스포팅 영상을 수집하여 구축하였는데, 대상의 측면을 망원 카메라로 근접 촬영한 영상이 대부분이므로 탐지 분야 연구에 사용되기는 어렵다. Airbus aircraft detection 데이터 셋은 다수의 항공기가 주기된 지역의 위성영상으로 구축하였는

데, 항공기의 상부만을 볼 수 있으므로 항공표적과 유사한 특징을 거의 갖고 있지 않다.

중적외선 영역에 대하여는 Hui 등<sup>[9]</sup>이 공개한 데이터 셋이 있다. 해당 데이터 셋은 항공표적으로서 적합한 특징을 갖추고 있다. 그러나 영상의 해상도가 256 x 256으로 작고, 영상 불균일로 인하여 데이터의 품질이 낮다는 문제가 있다. 또한 연속되는 프레임으로 구성된 22개의 데이터 시퀀스만을 가지고 있어 데이터의 다양성도 부족하다. 언급한 데이터 셋들이 가지는 문제는 그림 1의 예시로 쉽게 이해할 수 있다.

## 3. 중적외선 영상 합성

중적외선 영상을 직접 촬영하는 대신 합성을 통하여 획득하는 방법은 컴퓨터 시뮬레이션 기술의 발전과 함께 꾸준히 연구되어 왔다<sup>[10,11]</sup>. 그러나 중적외선 영상을 합성하기 위한 시뮬레이션에는 정확한 물질 정보(종류, 방사율 두께 등), 환경 정보(온도, 습도, 대기 투과율 등), 검출기 특성(파장대역 등)이 고려되어야 한다. 또한 엔진 열, 배기가스 열, 공력가열까지 모사하기 위해서는 각각에 대한 모델링 및 시뮬레이션을 수행하고 그 결과를 활용해야 한다. 이러한 일련의 과정에는 분야별 전문지식과 함께 복잡하고 정교한 작업이 수반된다.

최근에는 Generative adversarial networks (GAN)를 이용하여 가시광 영상으로부터 중적외선 영상을 생성하는 연구도 시도되고 있다<sup>[12,13]</sup>. 그러나 GAN 기반 결과물은 작은 객체와 같이 세부적인 묘사는 무시되는 경향이 있고 영상

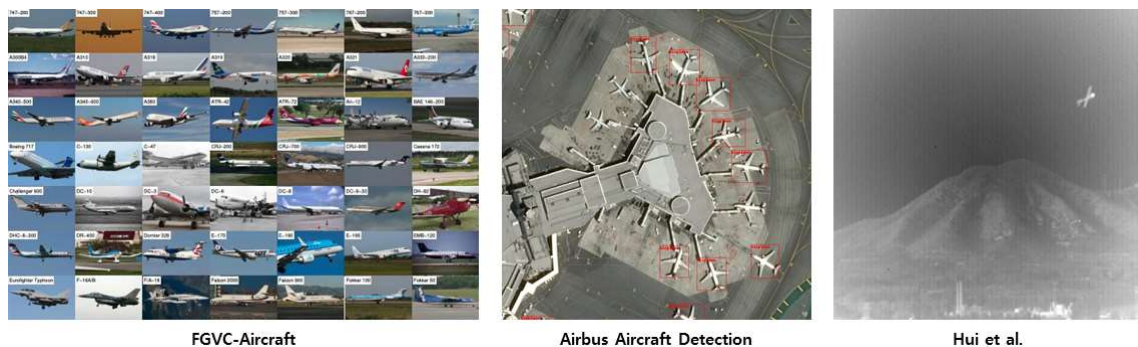


그림 1. 항공기와 관련된 기존 데이터 셋의 예시  
 Fig. 1. Examples of previous dataset related with aircraft

의 품질이 떨어진다. 또한 물리적인 해석에 기반 하지 않으므로 데이터의 사실성도 부족하다.

### III. 데이터 셋 구축

#### 1. 중적외선 카메라를 이용한 항공표적 데이터 획득

비행 중인 실제 항공표적은 항공기가 자주 통행하는 지역의 지상에 설치한 중적외선 카메라를 이용하여 동영상으로 촬영하였다. 촬영에 이용한 카메라는 FLIR 사의 A6751<sup>[14]</sup> 제품이다. 동영상은 FLIR 사의 자체 형식으로 저장되기 때문에 업체에서 제공하는 Application programing interface (API)를 이용하여 일정 프레임 단위의 RAW 영상으로 추출하였다. 영상의 해상도는 640 x 512이며, 검출한 열 에너지를 정밀하게 나타내기 위하여 16 bit 정밀도를 사용한다.

주석 작업은 LabelImg<sup>[15]</sup>라는 소프트웨어를 사용하였다. 시각화한 RAW 영상은 육안으로는 빠르게 표적을 식별하기 어렵기 때문에 영상의 명암을 개선한 상태에서 작업하였다. 주석은 탐지 분야에서 주로 통용되는 COCO<sup>[6]</sup> 데이터 셋의 형식으로 변환하였으며, 카테고리는 항공기의 양력 발생 방식에 따라 고정익, 회전익으로 구분하였다. 두 카테고리 외에 항공표적으로 오인될 수 있는 새나 날벌레 등은 비행 생물체라는 별도 카테고리를 부여해주었다.

중적외선 카메라를 이용하여 획득한 전체 영상의 개수는 학습용 3,184개, 평가용 797개이다. 영상 데이터의 예시는 그림 2와 같으며, 데이터의 구성은 표 1과 같다. 크기를 분

류하는 기준은 COCO<sup>[6]</sup>의 기준을 따른다. 따라서 픽셀 면적이 32 x 32 이하이면 소형 표적이고, 96 x 96 이상이면 대형 표적이다. 그 중간은 중형표적이다. 전체적인 구성을 살펴보면 제트 여객기가 가장 널리 운용되는 항공기이므로 고정익의 수가 회전익의 수보다 월등히 많은 것을 확인할 수 있다.

표 1. 중적외선 카메라로 획득한 데이터의 표적 구성  
Table 1. Composition of data acquired through a mid-wave infrared camera

	Small	Medium	Large	Total
Fixed wing	2,814	897	163	3,874
Flywheel	188	0	0	188
Flying Creatures	99	0	0	99

#### 2. 컴퓨터 그래픽스를 이용한 항공표적 데이터 합성

컴퓨터 그래픽스 소프트웨어인 Blender<sup>[16]</sup>를 이용하여 환경맵에 3차원 항공기 모델을 배치하고 렌더링을 수행하면 항공표적 데이터를 합성할 수 있다. 환경맵은 사실적인 가상공간을 빠르게 구축할 수 있다는 장점이 있는데, 본 연구에서는 Poly Haven<sup>[17]</sup>에서 공개한 환경맵 40개를 이용하였다. 공개된 환경맵은 고해상도의 파노라마 사진으로 제작되었으며, 사실적이고 다양한 형태의 하늘과 광원을 포함하므로 양질의 데이터를 생성하는데 도움을 준다. 3차원 항공기 모델은 웹상에 공개된 항공기 모델 8종을 사용하였다. 3차원 모델은 텍스처의 사용을 고려하지 않았다.

환경맵이 설정된 공간에 3차원 모델과 카메라를 배치하

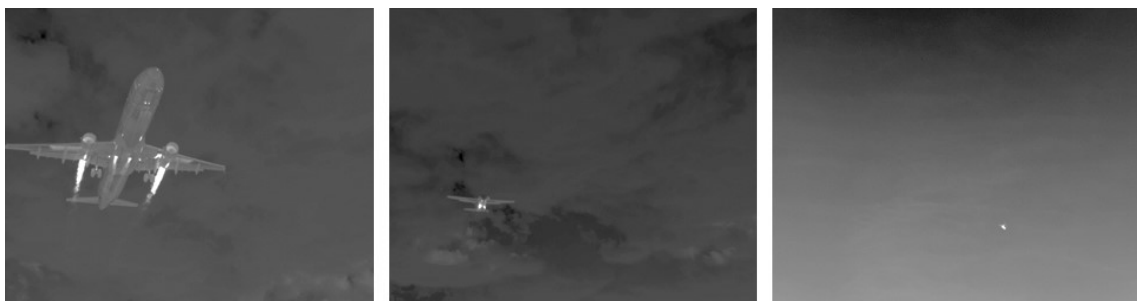


그림 2. 중적외선 카메라로 획득한 영상의 예시  
Fig. 2. Examples of images acquired through an mid-wave infrared camera

고 렌더링을 수행하는 절차는 다음과 같다. 먼저, 카메라를 가상공간의 원점에 배치하고 3차원 모델을 카메라 시야 내 무작위 위치  $\{x, y, z\}$  [m]에 무작위 자세  $\{\phi, \theta, \psi\}$  [deg]로 배치한다. 물체가 카메라 시야 범위 내에 존재하면서 카메라와의 거리가 50m 이상 1,500m 이하가 되도록 하는  $\{x, y, z\}$ 의 결정 방법은 수식 (1)과 같다. 수식 (1)에서  $\{I_{hor}, I_{ver}\}$ 는 영상의 해상도이며,  $\{P_{hor}, P_{ver}\}$ 는 픽셀 피치 크기,  $F$ 와  $L$ 은 각각 초점거리와 3차원 모델의 길이이다. 무작위 자세의 범위는  $\{\phi, \theta, \psi\} \in [0, 360]$  이다.

$$x \in \left[ -\frac{I_{hor} - LC_{hor}}{2C_{hor}}, \frac{I_{hor} - LC_{hor}}{2C_{hor}} \right] \text{ and}$$

$$y \in [50, 1500] \text{ and}$$

$$z \in \left[ -\frac{I_{ver} - LC_{ver}}{2C_{ver}}, \frac{I_{ver} - LC_{ver}}{2C_{ver}} \right] \quad (1)$$

,where  $\{C_{hor}, C_{ver}\} = \left\{ \frac{F}{yP_{hor}}, \frac{F}{yP_{ver}} \right\}$

카메라와 모델의 배치가 완료되면 카메라 위치를 기준으로 카메라와 모델을 3차원 변환하여 시점을 변경한다. 이후 렌더링을 수행하면 변환된 시점의 영상을 얻을 수 있다. 카메라와 3차원 모델의 배치부터 렌더링까지의 일련의 과정들은 그림 3과 같이 나타낼 수 있으며, Blender에서 제공하는 API로 자동화하였다. 렌더링 엔진은 Blender의 Cycles 엔진을 사용하였고 렌더링으로 생성되는 영상의 해상도와 정밀도는 중적외선 카메라로 획득한 영상과 동일하게 설정하였다. 합성된 영상에 대한 주석 작업은 3차원 모델에 대

한 마스크를 이용하여 자동화하였다. 마스크는 Blender의 Compositor 기능을 이용하면 렌더링 과정에서 같이 추출할 수 있다.

컴퓨터 그래픽스를 이용하여 합성한 전체 영상의 개수는 학습용 17,479개, 평가용 920개이다. 데이터의 구성은 표 2와 같다. 새로운 환경맵과 3차원 모델을 추가한다면 데이터의 수량의 크게 늘릴 수 있다.

표 2. 컴퓨터 그래픽스로 합성한 데이터의 구성  
 Table 2. Composition of data synthesized with computer-graphics

	Small	Medium	Large	Total
Fixed wing	3,236	5,448	2,937	11,621
Flywheel	3,443	3,315	20	6,778

#### IV. 딥러닝 기반 탐지 모델 학습 방법

탐지 모델은 일반적인 딥러닝 기반의 탐지 모델을 수정하지 않고 사용할 수 있다. 그러나 일반적인 탐지 모델은 색상 채널당 정밀도가 8 bit인 영상에 대하여 동작하도록 설계되었다. 반면 획득한 중적외선 영상의 정밀도는 16 bit이다. 따라서 모델의 입력 단에는 정밀도를 낮추기 위한 영상 전처리를 추가한다. 중적외선 영상의 정밀도를 낮추는 과정에서 발생하는 명암 정보의 손실은 영상의 정규화를 통하여 최소화한다. 정규화를 대신하여 다양한 종류의 히스토그램 평활화를 적용하는 것도 고려할 수 있

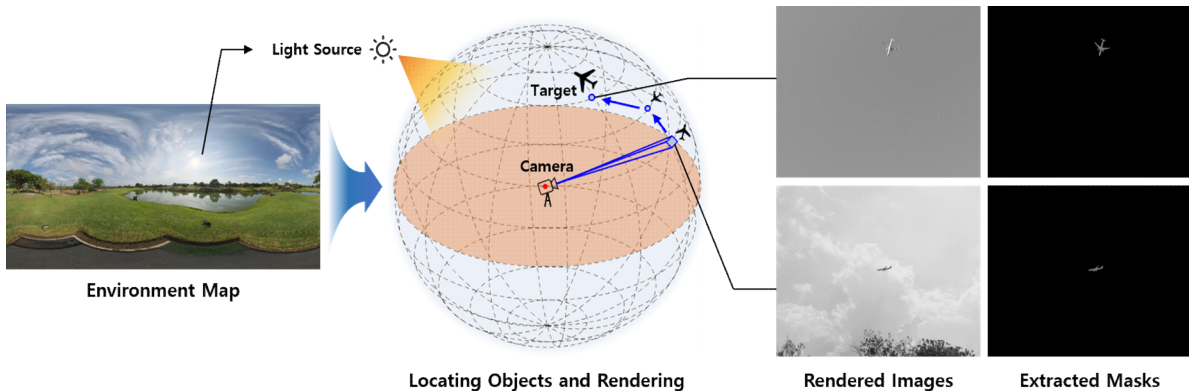


그림 3. 컴퓨터 그래픽스를 이용한 항공표적 영상 합성 방법  
 Fig. 3. Method of computer-graphics-based aerial target image synthesis

다. 그러나 자체적인 실험에서는 이러한 기법들의 이점을 발견하지 못하였다.

탐지 모델의 학습은 합성된 가시광 영역의 항공표적 데이터로 사전학습을 수행하고, 실제로 획득한 중적외선 영역의 데이터를 학습하도록 구성된다. 사전학습에 사용하는 합성 데이터는 가상의 가시광 영상이기 때문에 실제로 획득한 데이터와는 많은 차이를 가지고 있다. 따라서 실제 중적외선 영상과 유사한 특징을 가지는 영상이 많이 발생하게 하는 데이터 증강을 수행한다. 데이터 증강이 포함하는 영상처리는 가우시안 잡음, 블러, 색 반전, 명암 변조이다. 색 반전은 적용 유무가 무작위이며 다른 영상처리는 무작위의 세기로 적용된다. 각 영상처리의 역할을 정리하면, 가우시안 잡음과 블러는 합성 영상에서 잘 나타나지 않는 영상의 품질 저하를 반영한다. 색 반전은 표적이 배경보다 어두운 경우와 밝은 경우를 모두 발생시킨다. 명암 변조는 표적과 배경간의 명암 특성을 정규화 기반의 영상 최댓값, 최솟값의 변화로 다양화한다. 그러나 명암 손실로 인하여 표적 탐지가 과도하게 어려워지는 것을 방지하기 위하여 변조의 범위의 제한한다. 범위는 영상 정밀도가  $B$  bit인 영상의 변조 후 최댓값  $newMax$ 와 최솟값  $newMin$ 이 수식 (2)를 만족하면 된다.

결과적으로, 데이터 증강을 통하여 생성될 수 있는 영상 중 일부는 정규화 된 실제 중적외선 영상과 구별되는 특징이 상당히 감소하게 된다. 이는 그림 4의 예시로 확인할 수 있다.

## V. 실험 및 결과

### 1. 실험의 설계

무기체계의 실제 운용 관점을 고려하면, 탐지 모델은 다양한 크기의 표적을 빠르게 탐지할 수 있어야한다. 따라서 딥러닝 기반 탐지 모델은 이러한 조건에 부합하는 1단 구조의 Anchor-free 기반 탐지 모델인 YOLOX<sup>[5]</sup>를 이용하였다. 모델 학습을 위한 설정은 모델의 논문에서 제안한 것과 대부분 동일하다. 그러나 실험의 통제를 단순화하기 위하여 YOLOX의 방법과 달리 학습률의 점진적 증가와 감소가 한 주기만 이루어지도록 OneCycleLR 학습률 스케줄러<sup>[18]</sup>를 사용하였다. 학습 epoch는 학습 데이터의 양을 고려하여 사전학습은 100 epoch, 실제 데이터에 대한 학습은 30 epoch로 설정하였다.

제안하는 방법과의 대조군은 사전학습을 수행하지 않는 모델과 사전학습 시 증강기법을 사용하지 않는 모델로 설정하였다. 대조군과의 정량적인 탐지 성능의 비교는 Precision, Recall, F1-score, Average precision (AP), Average recall (AR)을 COCO API<sup>[19]</sup>로 산출하여 수행하였다. AP와 AR 계산 시 Intersection over union (IoU) 임계점은 [0.50:0.05:0.95]로 설정하였으며, 카테고리 전체에 대한 평균까지 수행한 경우 mAP와 mAR로 표기하였다. 정성적인 탐지 성능의 비교는 탐지한 표적의 경계상자를 시각화하여 수행하였다.

$$newMin \in [0, 2^B - Max + Min - 1], newMax \in [newMin + Max - Min, 2^B - 1] \quad (2)$$



그림 4. 증강된 합성 영상의 예시

Fig. 4. Examples of augmented synthetic images

## 2. 정량적 결과

정량적 결과는 표 3과 같다. 제안하는 방법을 전부 적용한 모델은 종합적인 성능 지표인 mAP와 mAR에 대하여 가장 높은 점수를 기록하였으며, 증강기법만 적용하지 않은 모델은 그 다음을 기록하였다. 제안하는 방법의 적용으로 인한 모델의 성능 향상 정도는 mAP는 최대 6.9%, mAR은 최대 7.5%이다.

크기와 카테고리를 구분하여 확인한 성능도 대체로 동일한 경향이 나타났다. 주목할 점은 제안하는 방법을 전부 적용한 경우에 한하여 소형 표적과 회전익 표적에 대한 성능 향상이 뚜렷하다는 것이다. 특히 회전익 표적에 대한 Precision 성능 향상의 경우에는 제안하는 방법 중 증강기법만 적용하지 않았을 때 1.1%, 제안하는 방법을 전부 적용하였을 때 12.3%를 기록하며 큰 차이를 보였다. 이는 증강기법이 소형 표적과 회전익 표적에 대하여 더욱 사실적인 결과를 생성하기 때문으로 추정할 수 있다.

크기 별 성능 비교에서 소형 표적에 대한 mAR의 향상 정도는 최대 7.7%로 최대 6.3%인 mAP의 향상 정도보다 큰데, 이는 무기체계 관점에서 주목할 만하다. 표적이 먼 위치에서 접근하여 작은 크기로 인식되는 초기 탐지 단계에는 의심되는 표적을 놓치지 않는 능력이 더욱 중요하기 때문이다.

## 3. 정성적 결과

정성적 결과는 그림 5와 같다. 그림은 이해를 돕기 위하여 확대 및 크롭이 수행되었음을 밝힌다. 제안하는 방법을 모두 사용한 경우는 노란색 경계상자로 표시하였는데, 초록색으로 표시한 참값과 대체로 가장 유사한 결과를 보여주고 있다. 반면에 빨간색으로 표시한 제안하는 방법을 사용하지 않은 방법은 상당히 큰 경계상자 오차를 보여주고 있다. 심지어는 경계상자를 중복으로 검출하거나 표적이 아닌 대상을 검출하는 경우도 확인할 수 있다. 주목할 점은

표 3. 탐지 성능의 정량적인 비교  
 Table 3. Quantitative comparison of detection performance

		w/o proposed	Proposed w/o augmentation	Proposed
Total scores				
mAP		0.348	0.361	0.372
mAR		0.398	0.409	0.428
Scores by size				
Small	mAP	0.320	0.327	0.340
	mAR	0.375	0.385	0.404
Medium	mAP	0.797	0.813	0.812
	mAR	0.826	0.839	0.840
Large	mAP	0.849	0.892	0.883
	mAR	0.871	0.900	0.888
Scores by category				
Fixed wing	Precision	0.636	0.638	0.643
	Recall	0.967	0.974	0.976
	F1-Score	0.767	0.771	0.775
	AP	0.636	0.638	0.643
	AR	0.673	0.677	0.681
Flywheel	Precision	0.374	0.378	0.420
	Recall	0.901	0.881	0.911
	F1-Score	0.528	0.529	0.575
	AP	0.374	0.378	0.420
	AR	0.453	0.460	0.502

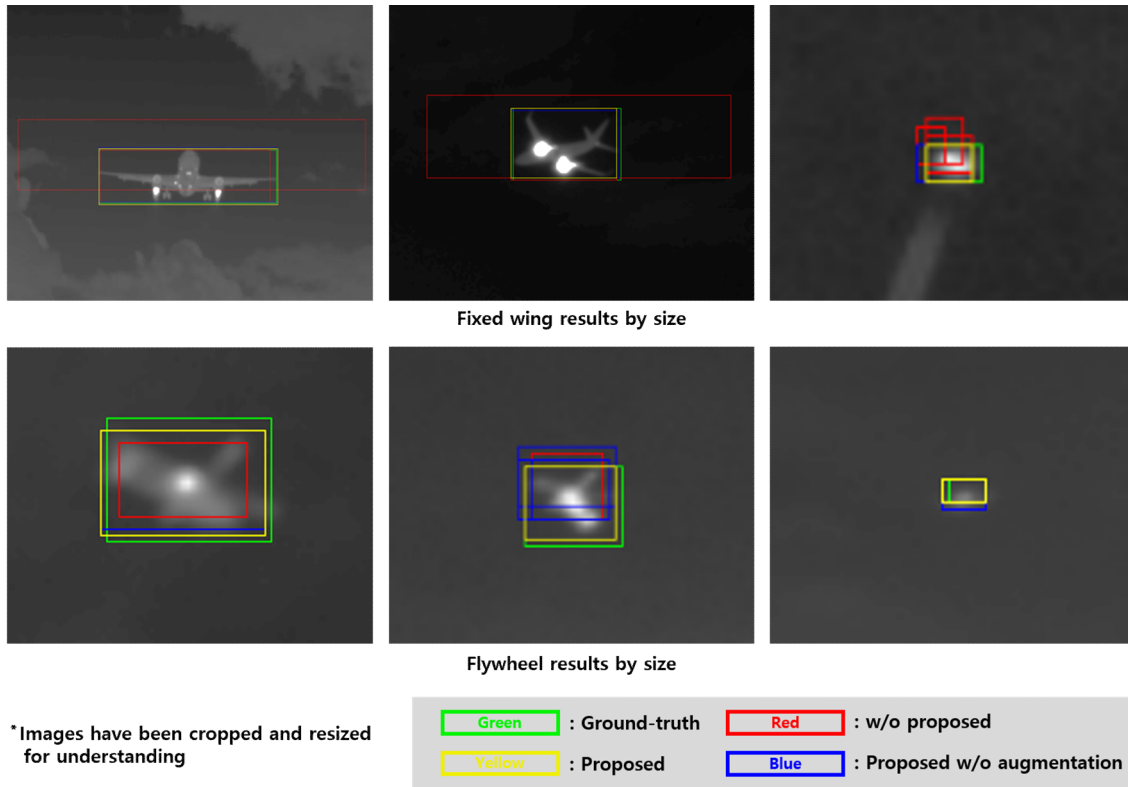


그림 5. 탐지 성능의 정성적 비교  
 Fig. 5. Qualitative comparison of detection performance

중적외선 영역 항공표적 데이터의 특성상 데이터 주석 작업이 까다로워 참값이 다소 부정확한 측면이 있는데, 제안하는 방법으로 학습된 모델은 그것보다 정교하게 표적을 검출하기도 하였다. 이는 사전학습에 이용한 합성 기반 학습 데이터가 정교한 참값을 가지고 있어 도움을 준 것으로 보인다.

## VI. 결론

중적외선 영상을 이용한 항공표적의 탐지는 군사적인 효용성이 높은 기술이지만 데이터 확보의 문제로 인하여 최신의 딥러닝 기술을 적용하는 데 어려움이 있다. 본 논문에서는 이를 극복하기 위하여 컴퓨터 그래픽스 기반의 학습 데이터를 이용하는 항공표적 탐지 모델 학습 방법을 제안하였다. 또한 모델 학습과 평가에 필요한 실제 항공표적 데

이터를 중적외선 카메라를 이용하여 직접 획득하였다.

제안하는 방법은 합성 기반의 데이터 확보 과정에서 전문적인 지식이나 정교한 작업이 필요하지 않다. 그럼에도 실제 데이터를 충분히 확보하기 어려운 유형의 대상에 대한 탐지 성능을 향상시키는 데 도움을 줄 수 있다. 실험결과는 자체 데이터 셋에 대하여 mAP는 0.372 mAR은 0.428 까지 달성할 수 있음을 보여주었다. 이는 제안하는 방법을 적용하지 않았을 때 보다 mAP는 6.9%, mAR은 7.5% 개선된 수치이다. 추후 학습 데이터를 더 많이 확보한다면 더 큰 성능 향상도 기대할 수 있을 것이다.

그러나 제안하는 방법에서 사용하는 사전학습은 획득한 항공표적의 특징 검출 능력을 온전히 전달할 수 있는 방법은 아니다. 향후 연구에서는 도메인 적응 기법을 적용하여 이러한 문제를 완화하기 위한 시도를 해볼 수 있을 것이다. 다만, 도메인 적응 기법의 경우 도메인 간에 공통적으로 나타나는 특징에 집중하는 경향이 있는 것으로 알려



저 있다. 그러나 중적외선 영상에서는 표적의 높은 열 에너지에 집중하면 쉽게 탐지가 가능한 대상이 많았다. 그러므로 도메인 적응 기법을 적용하는 연구에서는 중적외선 영상의 고유 특징을 검출할 수 있는 능력은 잃지 않도록 하는 방법을 강구해야 할 것으로 보인다. 더불어 본 연구는 지상 배경이 나타나는 항공표적은 다루지 않고 있다. 항공표적을 지상이 아닌 공중에서 바라볼 경우 이러한 조건에서도 동작할 수 있어야 한다. 이는 무기체계의 운용 방식에 따라 중요한 요구조건이 될 수 있으므로 추가 연구를 통하여 다루고자 한다.

### 참 고 문 헌 (References)

- [1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 580-587, 2014.  
doi: <https://doi.ieeecomputersociety.org/10.1109/CVPR.2014.81>
- [2] R. Girshick, "Fast R-CNN," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1440-1448, 2015.  
doi: <https://doi.org/10.1109/ICCV.2015.169>
- [3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.  
doi: <https://doi.org/10.1109/CVPR.2016.91>
- [4] H. Law and J. Deng, "Cornernet: Detecting objects as paired keypoints," in Proceedings of the European conference on computer vision (ECCV), pp. 734-750, 2018.  
doi: [https://doi.org/10.1007/978-3-030-01264-9\\_45](https://doi.org/10.1007/978-3-030-01264-9_45)
- [5] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO Series in 2021," arXiv preprint arXiv:2107.08430, 2021.  
doi: <https://doi.org/10.48550/arXiv.2107.08430>
- [6] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 740-755, 2014.  
doi: [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
- [7] S. Maji, E. Rahtu, J. Kannala, M. Blaschko, and A. Vedaldi, "Fine-grained visual classification of aircraft," arXiv preprint arXiv:1306.5151, 2013.  
doi: <https://doi.org/10.48550/arXiv.1306.5151>
- [8] Airbus Aircraft Detection, <https://www.kaggle.com/datasets/airbusgeo/airbus-aircrafts-sample-dataset>
- [9] B. Hui, Z. Song, H. Fan, P. Zhong, W. Hu, X. Zhang, J. Lin, H. Su, W. Jin, Y. Zhang, and Y. Bai, "A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background," Science Data Bank, 2019.  
doi: <https://doi.org/10.11922/sciedb.902>
- [10] C. J. Willers, M. S. Willers, and F. Lapierre, "Signature modelling and radiometric rendering equations in infrared scene simulation systems," in Technologies for Optical Countermeasures VIII, Vol. 8187, pp. 173-188, 2011.  
doi: <https://doi.org/10.1117/12.903352>
- [11] H. S. Jang, D. G. Kim, K. Y. Kwon, and N. K. Ha, "Design and Implementation of Infrared Image Generation Algorithm with Three-Dimensional Conjugate Heat Transfer," The Society of Convergence Knowledge Transactions, Vol. 9, No. 4, pp. 43-54, 2021.  
doi: <https://doi.org/10.22716/sckt.2021.9.4.042>
- [12] X. Yi, H. Pan, H. Zhao, P. Liu, C. Zhang, J. Wang, and H. Wang, "Cycle Generative Adversarial Network Based on Gradient Normalization for Infrared Image Generation," Applied Sciences, Vol. 13, No. 1, p. 635, 2023.  
doi: <https://doi.org/10.3390/app13010635>
- [13] T. Hui, Y. Xu, H. Li, Q. Zhou, and J. Rasol, "Water-surface infrared small object detection based on spatial feature weighting and class balancing method," IET Image Processing, Vol. 17, No. 10, pp. 3012-3027, 2023.  
doi: <https://doi.org/10.1049/ipr2.12851>
- [14] FLIR A6751, <https://www.flirkorea.com/products/a6750-mwir/?model=29440-251&vertical=rd+science&segment=solutions>
- [15] LabelImg, <https://github.com/tzutalin/labelImg>
- [16] Blender, <https://www.blender.org/>
- [17] Poly Haven, <https://polyhaven.com/>
- [18] L. N. Smith and N. Topin, "Super-convergence: Very fast training of neural networks using large learning rates," in Artificial Intelligence and Machine Learning for Multi-domain Operations Applications, Vol. 11006, pp. 369-386, 2019.  
doi: <https://doi.org/10.1117/12.2520589>
- [19] COCO API, <https://github.com/cocodataset/cocoapi>

---

저 자 소 개



**박 준 형**

- 2020년 2월 : 한남대학교 전자공학과 학사
- 2022년 2월 : 인하대학교 전기컴퓨터공학과 석사
- 2022년 1월 ~ 현재 : LIG넥스원 PGM탐색기연구소 연구원
- ORCID : <https://orcid.org/0009-0009-4006-2357>
- 주관심분야 : 컴퓨터 비전, 딥러닝, 그래픽스, 임베디드시스템



**최 용 호**

- 2021년 2월 : 세종대학교 지능기전공학부 학사
- 2023년 2월 : 세종대학교 지능기전공학과 석사
- 2023년 2월 ~ 현재 : LIG넥스원 PGM탐색기연구소 연구원
- ORCID : <https://orcid.org/0000-0002-5953-5312>
- 주관심분야 : 영상처리, 컴퓨터 비전, 딥러닝, 몰입형 입체영상



**문 종 원**

- 2021년 2월 : 배재대학교 전자공학과 학사
- 2023년 2월 : 세종대학교 지능기전공학과 석사
- 2023년 2월 ~ 현재 : LIG넥스원 PGM탐색기연구소 연구원
- ORCID : <https://orcid.org/0009-0006-1452-3955>
- 주관심분야 : 컴퓨터 비전, 딥러닝 기반 물체 검출 및 추적, 모델 경량화



**박 인 구**

- 2006년 : 부산대학교 전자전기정보컴퓨터공학부 학사
- 2008년 : 부산대학교 컴퓨터공학과 석사
- 2008년 ~ 현재 : LIG넥스원 PGM탐색기연구소 수석 연구원
- ORCID : <https://orcid.org/0009-0009-1849-1067>
- 주관심분야 : 임베디드시스템, 영상신호처리, 딥러닝



**이 기 남**

- 1997년 : 명지대학교 기계공학과 학사
- 2013년 : 고려대학교 기계공학과 석사
- 1997년 ~ 현재 : LIG넥스원 PGM탐색기연구소 수석 연구원
- ORCID : <https://orcid.org/0009-0004-0108-2420>
- 주관심분야 : 제어시스템, 신호처리, 검출기