

측면주사 소나 영상을 이용한 딥러닝 기반 기뢰 탐지

문종현¹⁾ · 박시언¹⁾ · 김재환²⁾ · 김창형³⁾ · 김대열⁴⁾ · 손채봉^{*,1,2)}

¹⁾ 광운대학교 전자통신공학과

²⁾ 광운대학교 방위사업학과

³⁾ 주식회사 제이스택업

⁴⁾ 경남대학교 인공지능학과

Deep Learning based Mine Detection using Side-Scan Sonar Image

Jonghyeon Mun¹⁾ · Sion Park¹⁾ · Jaehwan Kim²⁾ · Changhyung Kim³⁾ ·
Daeyeol Kim⁴⁾ · Chaebong Sohn^{*,1,2)}

¹⁾ Department of Electronics and Communications Engineering, Kwangwoon University, Korea

²⁾ Department of Defense Acquisition Program, Kwangwoon University, Korea

³⁾ Jstackup Co., Ltd., Korea

⁴⁾ Department of Artificial Intelligence, Kyungnam University, Korea

(Received 18 June 2024 / Revised 24 January 2025 / Accepted 21 February 2025)

Abstract

This paper compares and analyzes the performance of deep learning-based object detection models for mine detection, particularly Faster R-CNN and YOLOv5, using side-scan sonar images. Additionally, it proposes effective data augmentation method to enhance the generalization performance of mine detection models. Performance evaluation based on the structure and size of each model indicates that the two-stage model, Faster R-CNN, is more suitable for precise search tasks, while the one-stage model, YOLOv5, offers faster processing speed, making it advantageous for rapid mine detection in large maritime areas. This contributes significantly to improving the efficiency of maritime boundary missions and operations utilizing autonomous underwater vehicles, thereby making a substantial impact on naval operations and mine counter measures strategies.

Key Words : Mine Hunting(기뢰탐색), Side-Scan Sonar Image(측면주사 소나 영상), Object Detection(객체 탐지), Data Augmentation(데이터 증강), Autonomous Underwater Vehicle(자율무인잠수정)

1. 서론

기뢰는 수중에 설치되어 적의 선박이나 잠수함을 파괴하거나 이동을 제한하기 위해 설계된 수중 무기체

* Corresponding author, E-mail: cbsohn@kw.ac.kr
Copyright © The Korea Institute of Military Science and Technology

계이다. 과거 세대의 지뢰는 목표물과의 물리적 접촉이 있어야만 폭발했던 반면, 최근 개발된 기뢰들은 다양한 센서를 통해 자기장 변화, 음향 신호, 또는 압력 변화 등을 감지하고 목표물이 근접했을 때 자동으로 폭발한다^[1]. 이러한 기뢰의 존재는 해양 경계 임무 및 작전을 수행하는 데 있어 상당한 전술적 위협을 가하며, 이에 대응하기 위한 대기뢰전(MCM, Mine Counter Measures)은 해군의 주요 임무 중 하나이다.

MCM 체계에서 기뢰탐색(Mine Hunting)은 개별 기뢰를 정확히 식별하고 위치를 확인하여 제거하는 전술이다. 현대의 기뢰는 금속뿐만 아니라 플라스틱, 섬유 강화 플라스틱, 그 밖의 비금속 재료로 만들어지기도 하며, 이러한 재질은 일반적인 자기장 감지 장비로는 탐지하기 어렵다^[2]. 또한, 기뢰가 설치된 지역의 해양 환경 조건, 기뢰의 종류와 상태, 그리고 주변에 다른 위협 요소가 존재하는지 여부 등을 정확히 파악해야만 한다. 이러한 어려움 때문에, 기뢰탐색 작업은 고도로 전문화된 장비와 기술, 그리고 숙련된 인력을 필요로 하며, 인명 피해를 최소화할 수 있는 안전하고 효율적인 기술이 요구된다.

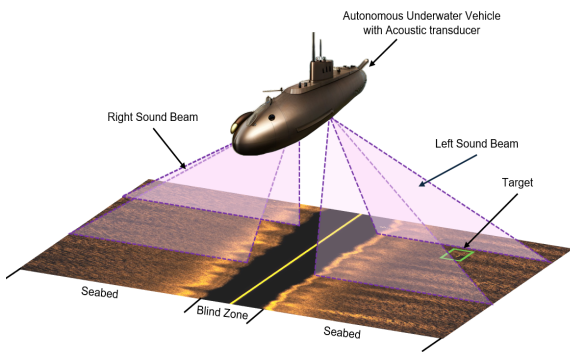


Fig. 1. AUV with side-scan sonar

이러한 요구를 충족하기 위해 자율무인잠수정(AUV, Autonomous Underwater Vehicle)은 기뢰탐색 작업에서 중요한 역할을 담당하고 있다. AUV는 사람이 탑승하지 않고도 독립적으로 해저를 탐색할 수 있는 자율적인 시스템을 갖추고 있으며, 고해상도 소나, 광학 카메라, 자기장 감지기 등 다양한 센서를 통해 수중 환경을 정밀하게 조사한다^[3]. 이러한 장비는 AUV를 해저의 기뢰를 정밀하게 탐지하고 식별하는 데 매우 유용하게 만들며, 위험한 환경에서 잠수부를 대신하여 기뢰를 안전하게 탐색하고 위치를 파악할 수 있도록

한다. 또한, AUV는 프로그래밍이 된 경로를 따라 자동으로 항해할 수 있으므로, 광범위한 지역을 체계적이고 효율적으로 탐색할 수 있다^[4]. 이는 기뢰탐색 작업을 더 안전하고, 빠르며, 비용 효과적으로 만들어 해양 경계 임무 및 작전의 효율성을 크게 향상시킨다.

AUV에 탑재된 소나 센서는 기뢰탐색 작업의 핵심 기술 중 하나로, 수중에서 음향 반사를 활용해 기뢰의 존재와 위치를 정확히 파악하는 데 활용된다. 이 센서들은 물리적 접촉 없이 해저면과 수중 객체에서 반사된 음파를 캡처하여 깊은 물에서도 선명한 이미지를 생성할 수 있어, 시야가 제한된 환경에서도 시각적 정보를 제공한다. 특히, Fig. 1과 같이 측면주사 소나는 양방향으로 광범위한 영역을 스캔할 수 있으며, 해저 지형과 기뢰를 포함한 수중 객체의 상세한 이미지를 분석가들에게 제공함으로써 데이터의 정확한 해석을 지원한다^[5].

컴퓨터 비전 분야의 객체 탐지 기술은 이러한 이미지 데이터를 통해 이미지 내 객체의 형태, 크기 및 패턴을 인식함으로써 자동 객체 식별을 가능하게 한다. 전통적인 기뢰 탐지 방법은 주로 이미지 처리 기법이나 통계적 모델링 기법에 의존하여 특정한 특징을 수동으로 설계하는 방식이기 때문에, 데이터의 복잡한 패턴을 충분히 반영하지 못할 수 있다^[6]. 반면, 딥러닝 기반의 객체 탐지 기술은 이미지 데이터를 통해 이미지 내 객체의 다양한 형태, 크기 및 패턴을 학습하여 더욱 정확하고 빠르게 작업을 수행할 수 있다. 특히 합성곱 신경망(CNN, Convolutional Neural Network)은 이미지 처리와 객체 탐지에서 널리 사용되는 딥러닝 알고리즘으로, 이미지 내 시각적 계층을 효과적으로 학습하여 각 객체의 형태와 특징을 인식할 수 있는 능력을 갖추고 있어, 다양한 크기와 형태를 가진 객체를 효과적으로 식별하고 분류하는 데 중요한 역할을 한다^[7]. CNN 모델은 여러 계층을 통해 이미지에서 저수준의 특징부터 고수준의 특징까지 추출하며 객체 탐지의 정확도를 크게 향상시킨다. 이러한 딥러닝 기반의 객체 탐지 모델은 대량의 이미지 데이터를 통해 학습하여, 다양한 형태와 크기를 가진 객체를 신속하고 정확하게 탐지할 수 있다.

AUV에서 측면주사 소나를 통해 수집된 고해상도의 이미지 처리 과정은 상당한 양의 컴퓨팅 리소스를 요구한다. 이는 결과적으로 AUV의 전력 소모에 영향을 미치며, 적은 연산량을 통한 효율적인 처리 과정은 장기간 운용의 가능성을 높이는 데 핵심적이다^[8]. 더불어

어, 높은 정확도로 기뢰와 비기뢰를 구분하여 인명 피해를 최소화하며, 기뢰탐색 작업의 신뢰성을 보장해야 한다. 이를 위해, 다양한 환경과 기뢰 형태에 적응할 수 있는 일반화 성능을 갖춘 모델 개발이 필수적이며, 이는 대량의 학습 데이터를 필요로 한다. 이러한 요구 사항들은 AUV를 이용한 기뢰탐색 시스템의 설계와 구현 전략에 있어 핵심적인 요소로 작용한다^[9].

본 논문에서는 딥러닝 기반 객체 탐지 모델들의 구조 및 크기가 측면주사 소나 영상을 활용한 기뢰 탐지 성능에 미치는 영향을 비교하고 평가한다. 또한, 대량의 측면주사 소나 영상을 효과적으로 증강할 수 있는 데이터 증강 기법을 제안하며, 이를 통해 객체 탐지 모델을 학습함으로써 기뢰 탐지 성능 개선에 미치는 효과를 평가한다.

2. 딥러닝 기반 객체 탐지 모델

딥러닝 기반의 객체 탐지 기술은 주로 객체의 위치를 결정하는 영역 제안(Region Proposal) 과정과 해당 객체가 어떤 카테고리에 속하는지 결정하는 분류(Classification) 과정을 포함한다. 이 두 과정을 동시에 수행할지, 혹은 순차적으로 진행할지에 따라, 객체 탐지 모델은 크게 One-stage와 Two-stage 구성으로 나뉜다. Fig. 2에서 보여주듯이, 현대의 딥러닝 기반 객체 탐지 모델은 백본(Backbone), 넥(Neck), 헤드(Head)의 세 가지 주요 네트워크 구성 요소로 이루어져 있다^[10].

백본 네트워크는 입력 이미지로부터 고수준의 특징을 추출하는 역할을 하며, VGGNet, ResNet, DarkNet 등이 이에 해당한다. 이러한 백본 네트워크는 특히 작은 객체 탐지와 같이 어려운 상황에서도 높은 탐지 정확도와 연산 효율성을 달성하기 위해 중요하다^[11]. 넥 네트워크는 백본 네트워크로부터 추출된 이미지나 피쳐맵의 크기를 다양하게 재구성하거나 재조정하여 헤드 네트워크로 전달하는 역할을 한다. 넥 네트워크는 피쳐맵의 피라미드 구조를 활용하여 멀티 스케일 객체 인식 성능을 향상시키는 FPN(Feature Pyramid Networks), PAN(Path Aggregation Network) 등과 같은 넥 네트워크들이 널리 활용되고 있다^[12]. 헤드 네트워크는 최종적으로 객체의 정확한 위치와 카테고리를 결정한다. 각 객체에 대한 최종 분류와 위치 조정을 담당하며, 이 네트워크에 따라 Two-stage 객체 탐지 모델과 One-stage 객체 탐지 모델로 나뉜다.

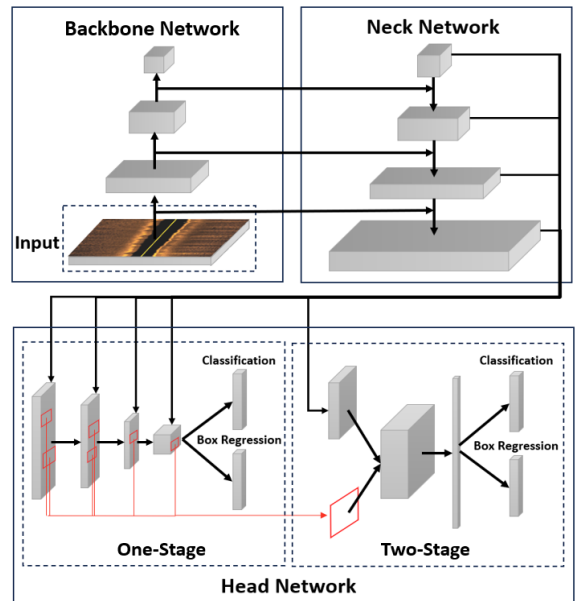


Fig. 2. Object detection model architecture

Two-stage 구성의 객체 탐지 모델은 객체 후보 영역을 먼저 제안하고 이후에 분류하는 과정을 순차적으로 진행하는 방식으로, 대표적으로 R-CNN과 이에 개선된 형태인 Fast R-CNN, Faster R-CNN이 이에 해당한다^[13]. 이 방식은 높은 정확도를 제공하지만, 상대적으로 연산속도가 느리다. 반면, One-stage 구성의 모델은 영역 제안과 분류 과정을 동시에 수행하여 연산속도를 크게 향상시킨다. SSD, YOLO 등이 대표적인 예로, 빠른 처리 속도를 자랑하지만, Two-stage 모델보다 정확도가 다소 낮을 수 있다^[14].

이처럼 딥러닝 기반의 객체 탐지 기술은 다양한 모델 구조와 접근 방식을 통해 상황별로 최적화된 성능과 효율성을 제공한다. 이러한 방식 중, Two-stage 모델인 Faster R-CNN과 One-stage 모델인 YOLOv5는 각각의 특징을 지닌다. 본 논문에서는 측면주사 소나 영상을 이용한 기뢰 탐지 성능을 평가하기 위해 이 두 모델을 선정하였다.

2.1 Faster R-CNN

Faster R-CNN은 대표적인 Two-stage 객체 탐지 모델 중 하나로, 기존 모델들이 별도의 알고리즘을 사용하여 후보 영역을 제안하는 방식에서 발전하여, RPN (Region Proposal Network)을 통해 네트워크를 단일화하고 통합함으로써 더욱 효율적인 객체 탐지가 가능

하게 되었다. RPN은 입력 이미지에서 객체가 있을 가능성이 높은 영역을 동적으로 선택하며, 이를 통해 기존의 R-CNN 계열의 Two-stage 객체 탐지 모델의 처리 속도를 개선하고 후보 영역 제안 과정에서 발생하는 지연을 최소화했다.

이 모델은 백본 네트워크로써 주로 ResNet50이나 ResNet101 같은 구조를 사용한다^[15]. 이들 구조는 깊은 레이어를 통해 높은 수준의 추상화와 복잡한 특징을 학습할 수 있으며, 이는 성능 향상에 기여하지만, 동시에 계산 비용도 증가시킨다. 백본 네트워크에서 추출된 피쳐맵은 다양한 스케일의 객체를 감지할 수 있도록 넥 네트워크인 FPN을 통해 다양한 해상도의 피쳐맵과 결합된다. FPN은 저해상도의 고수준 피쳐맵을 고해상도의 저수준 피쳐맵과 순차적으로 결합하는 상향식 결합 방식으로 피쳐맵을 생성하고, 이를 헤드 네트워크로 전달하여 최종 객체 탐지에 활용된다. 객체 탐지의 정확도를 더욱 높이기 위해, 헤드 네트워크에서는 전달받은 후보 영역들을 분류하고, 객체의 정확한 위치를 회귀 방식으로 조정한다. 이 과정에서 사용되는 손실 함수는 분류 오류와 위치 오차를 동시에 최소화하는 방식으로 설계하여, 전체 모델의 성능을 최적화한다.

2.2 YOLOv5

YOLOv5는 이미지 내의 객체의 위치 추정과 분류를 동시에 진행하는 One-stage 객체 탐지 모델로, YOLO 시리즈의 최신 버전 중 하나이다. 이 모델은 고성능 및 고속 처리 능력을 갖춘 다양한 크기의 버전을 제공하여, 컴퓨팅 자원과 처리 요구 사항에 따라 선택적으로 사용할 수 있다.

YOLOv5의 백본 네트워크인 CSPDarknet^[16]은 피쳐맵을 두 부분으로 분할하고, 한 부분은 직접 전달하며 다른 부분은 여러 레이어를 거쳐 처리한 후 재결합하는 구조를 가진다. 이 구조는 모델의 파라미터 수를 최소화하면서도 효과적인 특징 추출을 수행함으로써 연산 효율성을 향상시킨다. YOLOv5의 넥 네트워크인 PAN은 FPN의 상향식 결합 방식에 하향식 결합을 추가하여 보다 작은 객체에 대한 정보 손실을 최소화한다. YOLOv5의 헤드 네트워크는 각 객체의 위치와 클래스를 결정하는 역할을 하며, 객체의 위치와 클래스 확률을 예측하는 여러 레이어의 합성곱 신경망으로 구성되어 있다. 각 레이어는 특정 스케일에서 최적화되어 객체를 감지하고, 이를 통해 모델은 다양한 크기

의 객체를 효과적으로 처리할 수 있다. 또한, 헤드 네트워크는 객체의 신뢰도 점수를 산출하여 탐지의 정확도를 높이며, NMS(Non-Maximum Suppression) 알고리즘을 적용하여 중복 탐지를 제거함으로써 최종 탐지 결과의 정확성을 보장한다.

본 논문에서는 백본 네트워크로 ResNet50과 ResNet101을 사용하는 Two-stage 모델인 Faster R-CNN과 One-stage 모델인 YOLOv5의 세 가지 버전('m' for medium, 'l' for large, 'x' for extra large)을 평가 모델로 선정하였다. 모든 모델을 동일한 하드웨어 환경에서 학습 및 평가를 진행하여, 각 모델의 구조와 크기가 객체 탐지 정확도와 처리 속도에 미치는 영향을 평가한다.

3. 측면주사 소나 영상

3.1 측면주사 소나 데이터 세트

본 논문에서는 딥러닝 기반의 객체 탐지 모델의 학습 및 평가를 위해 포르투갈 해군 공병 잠수과견대가 포르투갈 해안을 따라 900-1800 kHz 측면주사 소나 센서를 탑재한 AUV로 촬영한 측면주사 소나 이미지를 데이터 세트로 사용한다^[17]. 이 데이터 세트는 실제 운영 환경에서 수집된 이미지로 구성되어 있어 자동화된 기뢰 탐지 시스템 개발을 위해 설계된 알고리즘 평가에 대한 기준을 마련해준다. Fig. 3은 해당 데이터 세트의 측면주사 소나 이미지를 보여준다.

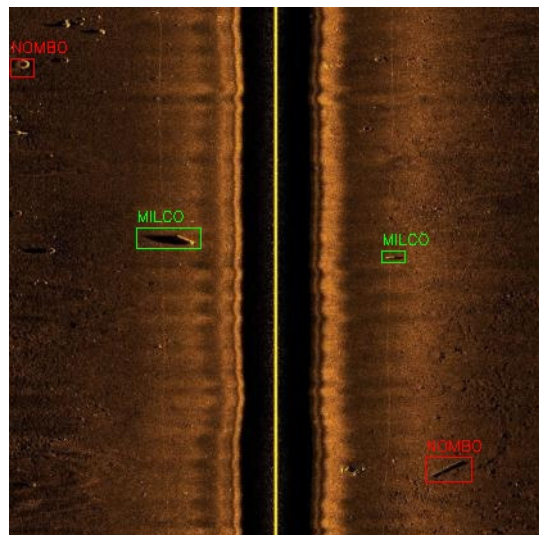


Fig. 3. Example of side-scan sonar image

원본 이미지의 크기는 416×416 또는 1024×1024 해상도를 가지며, 이미지 내 객체는 기뢰의심물체(MILCO, Mine-Like Contacts)와 비기뢰물체(NOMBO, Non-Mine-like Bottom Objects)로 구분된다. 각 객체의 위치는 이미지 좌표로 변환되어 경계 상자(Bounding Box) 형태로 객체 탐지 모델의 학습 및 평가에 레이블로써 사용된다.

3.2 학습 데이터 증강

이 데이터 세트는 Table 1과 같이 2010년부터 2021년까지 수집된 측면주사 소나 이미지로 구성되어 있으며, 432개의 MILCO와 235개의 NOMBO가 존재한다. 이러한 객체를 포함하는 이미지는 총 1,170개 중 304개로, 이는 전체 데이터 세트의 약 26%를 차지한다. 일부 이미지 처리 알고리즘의 개발과 적용에 있어서 데이터 세트의 크기가 클수록 알고리즘의 성능이 향상되는 경향이 있다^[18].

Table 1. Summary of the side-scan sonar dataset

Date	Images (with objects/without objects)	MILCO	NOMBO
2010	345 (28/317)	22	12
2015	120 (118/2)	238	175
2017	93 (19/74)	28	2
2018	345 (112/452)	95	46
2021	48 (27/21)	49	0
Total	1170 (304/866)	432	235

특히 딥러닝과 같은 많은 머신러닝 알고리즘의 경우 일반적으로 더 많은 데이터가 있을수록 성능이 향상되고, 모델의 일반화 능력도 높아진다. 또한, 객체 탐지 모델의 학습 데이터가 객체가 없는 이미지를 다수 포함하는 경우, 모델은 해당 이미지에 대한 편향을 발생시키며 실제로 존재하는 객체를 놓칠 수 있는 미탐지 현상이 증가할 수 있다. 더불어, 이러한 객체가

없는 이미지는 학습 과정에서 불필요한 계산량을 추가하게 되어 학습 시간을 증가시키고 모델의 학습 효율성을 저하시킨다.

따라서 본 논문에서는 측면주사 소나 영상을 이용한 딥러닝 기반의 기뢰 탐지 모델 학습에 효과적인 데이터 증강 기법을 제안한다.

3.2.1 단일 표본 데이터 증강 기법

측면주사 소나 이미지는 대칭적 특성을 갖는다. 이 특성을 활용하여 Fig. 4과 같이 단일 표본 데이터 증강 기법인 좌우 반전(Horizontal Flip)과 상하 반전(Vertical Flip)을 학습 단계에서 임의의 이미지와 해당 레이블에 적용한다. 이를 통해 다양한 위치와 방향을 갖는 객체를 포함한 데이터를 생성하여 학습 데이터의 다양성을 확장한다.

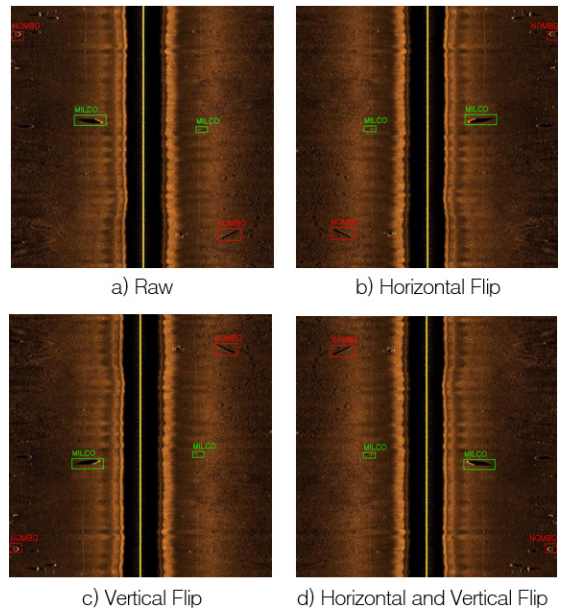


Fig. 4. Single-sample data augmentation method

3.2.2 이미지 합성 기반 데이터 증강 기법

Fig. 5와 같이 측면주사 소나 이미지 중앙의 사각지대(Blind Zone)를 기준으로 객체가 없는 이미지의 해저 영역과 객체가 있는 이미지의 해저 영역을 자르고 서로 붙여 객체가 존재하는 새로운 이미지를 생성한다. 이러한 이미지 합성 기반 데이터 증강 기법을 통해 여러 이미지를 조합하여 새로운 학습 데이터를 대량으로 생성할 수 있다.

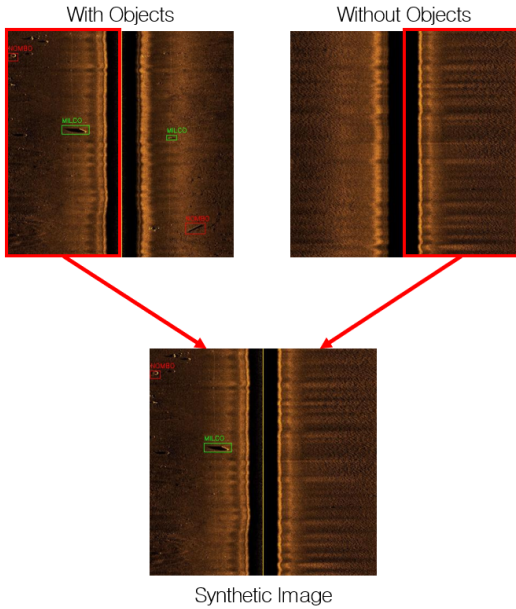


Fig. 5. Generation of sonar image with objects

4. 실험 방법 및 결과

4.1 데이터 세트 구성

본 논문에서는 모델 학습을 위해 Table 2와 같이 원본 학습 데이터와 5배수 증강한 학습 데이터를 각각 사용한다. 증강한 학습 데이터는 원본 학습 데이터로부터 객체가 존재하는 새로운 이미지를 생성하였으며, 클래스 간의 비율을 유지하였다. 모델 검증 및 평가에는 학습 데이터에 포함되지 않는 데이터를 통해 각 딥러닝 기반 기뢰 탐지 모델의 성능을 비교한다.

Table 2. Dataset configuration

Dataset		Images	MILCO	NOMBO
Train Dataset	Raw	936	343	189
	Augmented	4680	1721	952
Validation Dataset		117	44	22
Test Dataset		117	45	24

4.2 실험 환경 및 평가지표

본 논문의 실험에서 사용된 하드웨어 및 소프트웨어 구성은 Table 3과 같으며, 동일한 실험 환경에서 모든 모델의 학습 및 평가를 진행하였다.

Table 3. Experiment environment

CPU	Intel® Core™ i7-12700 12-core Processor
RAM	32GB
GPU	NVIDIA GeForce RTX 3060 Ti
OS	Ubuntu 18.04
Tools	Python 3.10.3, Pytorch 2.2.2

평가지표는 객체 탐지 알고리즘에서 널리 사용되는 각 클래스에 대한 AP(Average Precision), mAP(mean Average Precision), 런타임 측정을 통해 측면주사 소나 영상을 이용한 기뢰 탐지 모델의 성능을 평가한다¹⁹⁾.

객체 탐지 분야에서는 객체를 판단하기 위한 기준으로 IoU(Intersection over Union)를 이용한다. IoU는 식 (1)과 같이 예측 영역과 객체의 실제 영역 간의 교집합 면적(Area of Intersection)과 합집합 면적(Area of Union)의 비율로 계산한다. 본 논문에서는 IoU를 0.5로 설정하여 AP값을 측정한다.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union} \quad (1)$$

mAP는 여러 클래스에 대한 AP의 평균값으로 나타낸다. AP는 특정 클래스에 속할 것으로 예측될 확률(Confidence score)의 변화에 따라 측정된 모델의 정밀도와 재현율을 사용하여 그린 Precision-Recall 곡선 아래의 면적을 계산하여 얻어지는 지표이다. 높은 AP값은 모델이 높은 정밀도와 재현율을 모두 달성하고 있는 것을 의미하며, 좋은 탐지 성능을 가진 모델을 나타낸다.

$$AP = \int_0^1 P(R)dR \quad (2)$$

여기서, P는 정밀도이고, R은 재현율이다.

정밀도는 참으로 예측한 결과 중에서 얼마나 많은 실제 참인 객체가 있는지를 식 (3)과 같이 나타낸다.

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{All\ Detections} \quad (3)$$

여기서, TP(True Positive)는 올바르게 예측된 객체의

수를 나타내며, FP(False Positive)는 잘못 예측된 객체의 수를 나타내며 오탐지를 의미한다.

재현율은 실제로 참인 모든 객체를 얼마나 많이 찾아냈는지를 식 (4)과 같이 나타낸다.

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{All\ Ground\ Truths} \quad (4)$$

여기서, FN(False Negative)은 잘못 예측된 대상의 수를 나타내며 미탐지를 의미한다.

런타임 지표는 객체 탐지 모델이 입력데이터를 처리하고 결과를 출력하는 데 걸리는 시간, 즉 추론 속도를 나타낸다. 이는 실제 운용 환경에서의 객체 탐지 모델 적용 시 중요한 요소 중 하나로, 특히 실시간 응

용 프로그램이나 제한된 시간 내에 결과를 제공해야 하는 경우에 높은 추론 속도가 필수적이다. 런타임은 모델의 복잡성, 사용된 하드웨어의 성능, 입력데이터의 해상도 등에 의해 영향을 받는다.

따라서, 본 논문에서는 동일한 시스템 환경에서 입력 이미지 크기를 416 × 416으로 고정하여, 런타임 지표를 통해 모델의 추론 속도를 평가 및 비교한다. 또한, 클래스별 AP와 mAP를 모델의 예측 결과를 통해 평가하여, 모델의 성능을 정량적으로 측정한다.

4.3 실험 결과 및 분석

Table 4는 학습 데이터와 모델의 종류에 따른 기뢰 탐지 정확도에 대한 평가 결과를 보여준다. 데이터 증강을 실시한 결과, 전반적으로 모델 성능이 향상되었

Table 4. Dataset and model combination experiment result

Model	Raw Dataset			Augmented Dataset		
	AP _{MILCO}	AP _{NOMBO}	mAP	AP _{MILCO}	AP _{NOMBO}	mAP
YOLOv5-m	0.698	0.710	0.694	0.743	0.692	0.717
YOLOv5-l	0.744	0.696	0.720	0.768	0.701	0.735
YOLOv5-x	0.759	0.759	0.759	0.781	0.735	0.758
Faster R-CNN-ResNet50	0.788	0.737	0.763	0.826	0.728	0.777
Faster R-CNN-ResNet101	0.827	0.741	0.784	0.814	0.793	0.804

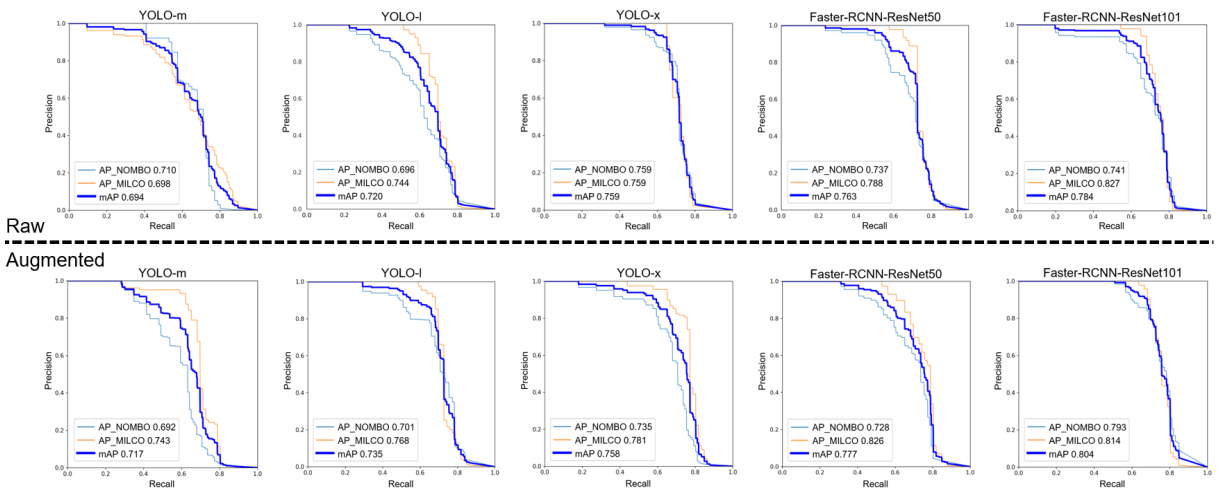


Fig. 6. Precision-recall curve for raw dataset(top) and augmented dataset(bottom)

다. 특히, Faster R-CNN-ResNet101 모델은 원본 데이터 세트로 학습된 모델과 비교하여 mAP가 0.02 증가하였다. 이는 제한한 데이터 증강 기법이 딥러닝 기반의 기뢰 탐지 모델의 일반화 능력을 효과적으로 개선한 것을 알 수 있다.

모델의 종류에 따른 기뢰 탐지 정확도 측면에서는 Two-stage 모델인 Faster R-CNN이 One-stage 모델인 YOLOv5보다 높은 mAP를 달성했다. 이러한 결과는 Faster R-CNN이 RPN을 통해 후보 영역 제안 단계를 거치면서 보다 정밀한 객체 탐지 과정을 수행하기 때문이다. 이를 통해 객체 탐지 모델의 구조적 차이가 성능에 영향을 미치는 것을 확인할 수 있다.

Table 5는 각 모델의 파라미터 수(Params)와 런타임을 보여주며, YOLOv5가 Faster R-CNN보다 추론 속도가 월등히 빠른 것을 확인할 수 있다. 모델의 크기가 클수록 성능이 향상되는 경향을 보이는 반면, 성능 향상과 모델의 크기 및 연산속도는 트레이드 오프 관계에 있다. 이는 고성능의 모델이 더 많은 계산 리소스와 처리 시간을 요구함을 의미하며, 이러한 특성은 특히 AUV와 같은 제한된 하드웨어 리소스를 가진 환경에서 중요한 고려 사항이 된다.

Table 5. Number of model parameters and runtime

Model	Params(M)	Runtime(ms)
YOLOv5-m	21.2	31.9
YOLOv5-l	46.5	58.5
YOLOv5-x	86.7	98.7
Faster R-CNN-ResNet50	41.1	470.8
Faster R-CNN-ResNet101	60.0	530.6

그리고 성능 평가에서 우수한 모델은 Fig. 6에서 볼 수 있듯이 곡선이 우측 상단에 가까이 위치하며, 높은 정밀도 값을 유지하면서 재현율이 증가하는 경향을 보인다. 이는 모델이 많은 양의 양성 예측을 정확하게 처리하면서도 실제 양성 케이스를 잘 놓치지 않고 찾아내고 있음을 의미한다. 이와 같은 성능은 특히 기뢰 탐지와 같은 안전과 신뢰성이 중요한 응용 분야에서 매우 중요하다.

마지막으로 Table 4와 Fig. 6을 통해 알 수 있듯이, 기뢰에 대한 AP보다 비기뢰에 대한 AP가 전반적으로

낮은 것을 확인할 수 있다. 이는 비기뢰의 종류, 형태 및 크기가 매우 다양하여 일반화하기 어려운 특성 때문에 발생하는 현상으로 보이며, 이러한 문제를 해결하기 위해서는 비기뢰에 대한 세분화된 분류 체계를 개발하여 각각의 특징을 보다 명확하게 학습할 수 있는 데이터 수집 및 가공 과정을 적용할 필요가 있다.

5. 결론

본 논문에서는 딥러닝 기반 객체 탐지 모델의 구조와 크기가 측면주사 소나 영상에서의 기뢰 탐지 성능에 미치는 영향을 평가 및 분석하였으며, 데이터 증강 기법을 적용하여 모델의 일반화 능력을 개선함으로써 기뢰탐색 작업의 정확성을 향상시키는 방법을 제안하였다. 데이터 증강 기법의 적용 결과, 모델들의 mAP가 전반적으로 증가하였다. 이는 측면주사 소나 영상의 대칭성을 활용한 좌우 반전, 상하 반전, 그리고 이미지 합성을 포함한 데이터 증강 기법들이 기뢰탐색 모델의 일반화 능력을 개선하는 데 효과적인 것을 입증한다. 이러한 기법들은 모델이 다양한 환경과 상황에서 기뢰를 정확하게 식별하고 분류할 수 있도록 돕고, 실제 해양 작전 환경에서의 적용 가능성을 높인다.

또한, Two-stage 모델인 Faster R-CNN과 One-stage 모델인 YOLOv5를 비교 분석하였으며, 이들의 성능은 처리 속도와 정확도라는 두 가지 관점에서 평가되었다. Faster R-CNN은 높은 mAP를 달성하며 정확도 측면에서 우수한 결과를 보였으나, 처리 속도 측면에서는 YOLOv5가 더 뛰어난 성능을 나타냈다. 이러한 결과는 특정 딥러닝 모델의 절대적인 우수성을 주장하기보다는, 기뢰탐색 임무의 특성과 AUV의 연산 성능 및 운영 조건에 따라 적절한 모델이 선택되어야 함을 시사한다. 예를 들어, 중소형 AUV와 같이 배터리 용량과 연산 능력이 제한적인 환경에서는 YOLOv5와 같은 경량화된 모델이 적합할 수 있다. 특히, 광범위한 해역을 탐사하며 빠른 처리가 요구되는 임무에서는 One-stage 모델의 장점이 더욱 부각될 수 있다. 반면, 대형 AUV와 같이 강력한 연산 자원을 사용할 수 있는 환경이나 고도의 정확도가 필수적인 임무에서는 Faster R-CNN과 같은 정밀 탐지 모델이 적합하다.

따라서, 본 논문은 기뢰탐색을 위한 딥러닝 모델의 최적화 및 선택 과정에서 필요한 구조적 고려 사항을

제시하였으며, 향후 AUV의 자원 조건과 임무 요구 사항을 기반으로 보다 구체적이고 실질적인 적용 가능성 평가의 기초 자료로 활용될 수 있을 것으로 기대된다.

References

- [1] N. E. Youngblood and Jo Mackiewicz, "The development of mine warfare: a most murderous and barbarous conduct," Bloomsbury Publishing USA, 2006.
- [2] J. M. Bachkosky, et al., "Unmanned vehicles(UV) in mine countermeasures(U)," Naval Research Advisory Committee USA, 2000.
- [3] S. T. Tripp, "Autonomous underwater vehicles(AUVs): a look at Coast Guard needs to close performance gaps and enhance current mission performance," USCG Res. Develop. Center Studies USA, 2006.
- [4] Veronika Yordanova, et al., "Coverage path planning for mine countermeasures: Adapting track orientation," In: OCEANS 2019-Marseille. IEEE, pp. 1-7, 2019.
- [5] B. W. Flemming, "Side-scan sonar: a practical guide," The international hydrographic review, 1976.
- [6] Xiaofeng Li, et al., "Deep-learning-based information mining from ocean remote-sensing imagery," National Science Review, pp. 1584-1605, 2020.
- [7] Laith Alzubaidi, et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," Journal of big Data 8, pp. 1-74, 2021.
- [8] Jing Yan, et al., "Energy-Efficient Data Collection Over AUV-Assisted Underwater Acoustic Sensor Network," in IEEE Systems Journal, Vol. 12, No. 4, pp. 3519-3530, 2018.
- [9] Narcís Palomeras, et al., "Automatic Target Recognition for Mine Countermeasure Missions Using Forward-Looking Sonar Data," in IEEE Journal of Oceanic Engineering, Vol. 47, No. 1, pp. 141-161, 2022.
- [10] Jianyuan Guo, et al., "Hit-detector: Hierarchical trinity architecture search for object detection," Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 11405-11414, 2020.
- [11] Huilan Luo, et al., "Small object detection network based on feature information enhancement," Computational Intelligence and Neuroscience, 2022.
- [12] Sara Bouraya, and Abdessamad Belangour, "Deep learning based neck models for object detection: a review and a benchmarking study," International Journal of Advanced Computer Science and Applications, 2021.
- [13] Ross Girshick, "Fast R-CNN," In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448, 2015.
- [14] Demetris Demetriou, et al., "Real-time construction demolition waste detection using state-of-the-art deep learning methods; single-stage vs two-stage detectors," Waste Management, pp. 194-203, 2023.
- [15] Peiji Chen, et al., "Computerized-assisted scoliosis diagnosis based on faster R-CNN and resnet for the classification of spine X-ray images," Computational and Mathematical Methods in Medicine, 2022.
- [16] Qingqing Xu, et al., "Effective face detector based on YOLOv5 and superresolution reconstruction," Computational and mathematical methods in medicine, pp. 1-9, 2021.
- [17] N. P. Santos, et al., "Side-scan sonar imaging data of underwater vehicles for mine detection," Data in Brief, 2024.
- [18] Sen Qiu, et al., "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," Information Fusion, pp. 241-265, 2022.
- [19] Di Feng, et al., "A review and comparative study on probabilistic object detection in autonomous driving," IEEE Transactions on Intelligent Transportation Systems, pp. 9961-9980, 2021.