

# 수중 소나 영상 학습 데이터의 왜곡 및 회전 Augmentation을 통한 딥러닝 기반의 마커 검출 성능에 관한 연구

## Study of Marker Detection Performance on Deep Learning via Distortion and Rotation Augmentation of Training Data on Underwater Sonar Image

이 언 호<sup>1</sup> · 이 영 준<sup>2</sup> · 최 진 우<sup>3</sup> · 이 세 진<sup>†</sup>

Eon-Ho Lee<sup>1</sup>, Yeongjun Lee<sup>2</sup>, Jinwoo Choi<sup>3</sup>, Sejin Lee<sup>†</sup>

**Abstract:** In the ground environment, mobile robot research uses sensors such as GPS and optical cameras to localize surrounding landmarks and to estimate the position of the robot. However, an underwater environment restricts the use of sensors such as optical cameras and GPS. Also, unlike the ground environment, it is difficult to make a continuous observation of landmarks for location estimation. So, in underwater research, artificial markers are installed to generate a strong and lasting landmark. When artificial markers are acquired with an underwater sonar sensor, different types of noise are caused in the underwater sonar image. This noise is one of the factors that reduces object detection performance. This paper aims to improve object detection performance through distortion and rotation augmentation of training data. Object detection is detected using a Faster R-CNN.

**Keywords:** Deep Learning, Data Augmentation, Object Detection, Underwater Sonar Image

### 1. 서 론

지상 환경에서의 자율 주행 연구는 광학 카메라에 기반을 두고, GPS 등 다양한 센서를 같이 사용하여 주변 환경에서 얻어지는 랜드마크로 로봇의 위치를 추정한다. 수중 환경에서도 로봇의 위치 추정을 위해 광학 카메라로 데이터를 얻기도 하지만, 물의 탁도나 빛의 산란에 의해 가시거리가 제한된다. 그래서 수중 환경에서는 외부 요인으로 인해 생기는 한계점이

적은 초음파 센서를 사용하여, 수중 환경에서 랜드마크를 인식해 로봇의 위치를 추정할 수 있다<sup>[1-3]</sup>. 또한, 지상 환경과는 다르게 수중 환경 속에서는 지속적으로 추출될 수 있는 랜드마크가 적다. 그래서 [Fig. 1]과 같은 인공 마커를 제작하여 수중 환경에 설치하는 방식으로 로봇의 위치를 추정할 수 있는 강인하고 지속적인 랜드마크를 얻는다<sup>[4]</sup>.

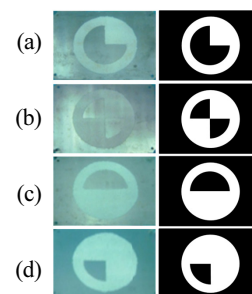
본 논문에서는 수중 초음파 센서를 이용하여 설치된 인공 마커의 이미지 데이터를 얻고, 이 데이터를 Faster R-CNN으로

Received : Dec. 7. 2018; Revised : Jan. 7. 2019; Accepted : Jan. 10. 2019

※ This research was supported by a grant from Endowment Project of "Development of fundamental technologies on underwater environmental recognition for long-range navigation and intelligent autonomous underwater navigation" funded by Korea Research Institute of Ships and Ocean engineering (PES9390).

1. PhD. Candidate, Mechanical Engineering, Kongju National University, Cheonan, Korea (Eonho@smail.kongju.ac.kr)
2. Researcher, Korea Research Institute of Ships and Ocean Engineering, Daejeon, Korea (leeyongjun@kriso.re.kr)
3. Senior Researcher, Korea Research Institute of Ships and Ocean Engineering, Daejeon, Korea (jwchoi@kriso.re.kr)

† Associate Professor, Corresponding author: Division of Mechanical & Automotive Engineering, Kongju National University, Cheonan, Korea (sejiny3@kongju.ac.kr)



[Fig. 1] Artificial marks for imaging sonar: (a) ID-1, (b) ID-2, (c) ID-3, (d) ID-4

인공 마커의 이미지를 분류 및 위치 추정을 한다<sup>15</sup>. 본 논문의 구성은 다음과 같다. 3장에서는 Faster R-CNN의 분류 성능 향상을 위해 데이터를 Augmentation하는 방법에 대해 설명한다. 4장에서는 실험 결과를 설명하며, 5장에서 결론을 맺는다.

## 2. 관련 연구

### 2.1 수중 초음파센서

초음파 센서는 물체의 재질에 따라 반사, 흡수 및 투과 등의 물리적인 현상에 의해 데이터 수집 중에 심한 데이터 노이즈가 생기며, 저화질의 영상 정보로 얻어진다<sup>6</sup>. 그럼에도 불구하고 수중 환경에서는 광학 카메라보단 초음파 센서를 선호하게 된다. 그 이유는 물의 탁도나 수심의 깊이에 따른 가시거리 제한이 없기 때문이다. 이러한 수중 초음파 센서에서 얻는 소나 이미지를 활용한 기존 연구는 표적 인식과 분류를 위해 딥러닝 모델로 특징들을 학습하여 SVM으로 분류하는 방식의 연구와 SAS (Synthetic aperture sonar)을 이용하여 표적 분류를 하는 연구가 있었다<sup>7,8</sup>. 또한, 소나 데이터에서 특정한 선 스펙트럼을 추출하여 SVM으로 물체 감지를 하는 연구가 있었다<sup>9</sup>. 선 스펙트럼을 딥러닝 모델 중 Autoencoder를 이용하여 물체 감지하는 연구도 있었다<sup>10</sup>.

### 2.2 물체 감지

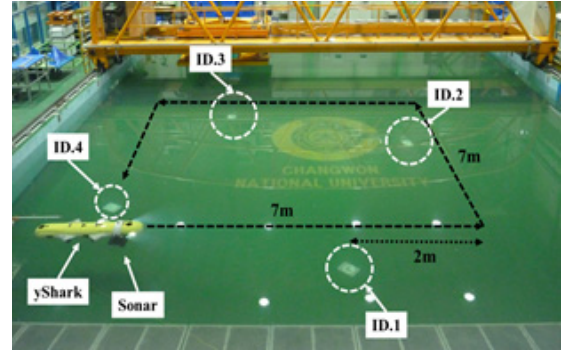
데이터를 분석하여 물체 인식 및 위치 추정을 물체 감지 연구라고 한다. 이 연구는 AlexNet이후로 성능이 2배 이상 향상했다<sup>11</sup>. 물체 감지 모델로 대표적인 R-CNN은 Selective search로 후보 영역을 찾아 AlexNet으로 학습하여 분류 및 감지하는 구조이었다<sup>12</sup>. 물체 감지 성능이 약 2배 향상했으나 모델의 구조의 한계로 연산 시간이 오래 걸리는 단점 있었다. 이러한 구조를 바꾸어, 먼저 입력 값을 CNN으로 학습을 한 후, Selective search를 통해 후보 영역을 찾아 Spatial Pyramid Pooling을 이용한 연구인 SPP-Net과 RoI Pooling을 이용한 Fast R-CNN이 개발되었다<sup>13,14</sup>. 이를 통해 연산 시간을 줄이고 입력 데이터의 크기의 제한을 제거하였다. 하지만 Fast R-CNN을 실시간으로 사용하기 힘들어, 후보 영역을 찾는 Selective search 기법을 Region Proposal Network (RPN)으로 대체하여 1장의 이미지를 연산하는데 0.2초대로 줄인 Faster R-CNN을 개발하였고 본 논문에서 이 모델을 이용하여 물체 감지를 수행하였다.

## 3. 데이터 Augmentation

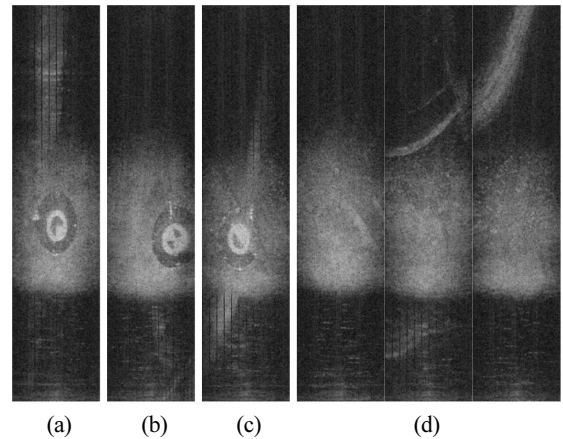
딥러닝 모델을 이용하여 데이터를 학습할 때, 보유하고 있



[Fig. 2] DIDSON sensor

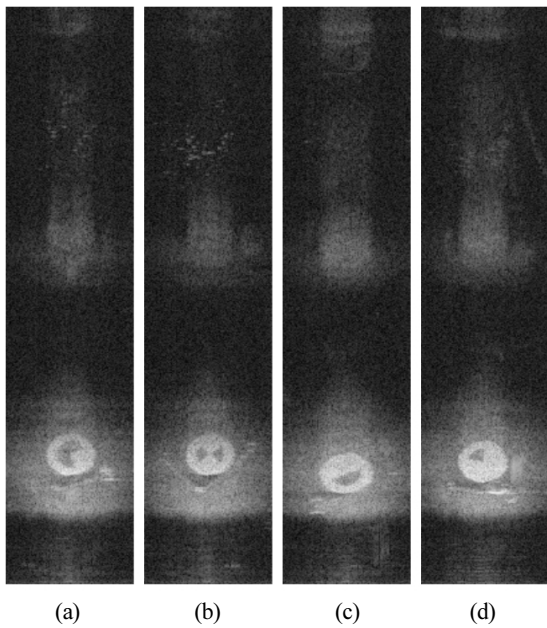


[Fig. 3] Experimental environments



[Fig. 4] Sonar images of dynamic data set: (a) ID-1, (b) ID-2, (c) ID-3, (d) Background

는 데이터가 적어서 데이터의 일반화 특성이 떨어지는 현상이나, 학습 모델의 과적합(Overfitting)이 발생할 경우 학습 데이터를 Augmentation을 하여 문제를 해결한다. 보통 Flip, Rotation, Scale, Crop 등의 기법으로 데이터를 Augmentation을 한다<sup>15</sup>. 본 논문에서는 Rotation과 Scale 기법으로 데이터를 변형하였다. 본 논문에서 이용된 이미지 데이터는 [Fig. 2]와 같은 DIDSON 센서를 이용하여, [Fig. 1]와 같은 인공 마커를 수집한 데이터로 동적 데이터와 정적 데이터세트를 구성하였다. 여기서 동적 데이터라 함은 센서가 움직이는 상황에서 마커의 소나 이미지를 획득하였기 때문이고, 정적 데이터는 센서가 움직이지 않고 있을 때에 획득한 소나 이미지 세트이다. 동적 데이터 이미지는 센서가 움직이는 상황에서 얻었기 때문에 상대적으로 정적 데이터 이미지에 비해 위치에 따라 왜곡 현상이 발생된다.



[Fig. 5] Sonar images of static data set: (a) ID-1, (b) ID-2, (c) ID-3, (d) ID-4

[Fig. 3]과 같은 수중 환경에서 무인 잠수정이 검은 점선 영역을 평균 0.15 m/s의 속력으로 자유롭게 이동하며 인공마커 및 수조에 대한 수중 이미지의 동적 데이터를 수집하였다. 정적 데이터는 동적 데이터와 달리, 고정된 한 자리에서 인공 마커의 대한 수중 이미지를 수집하였다.

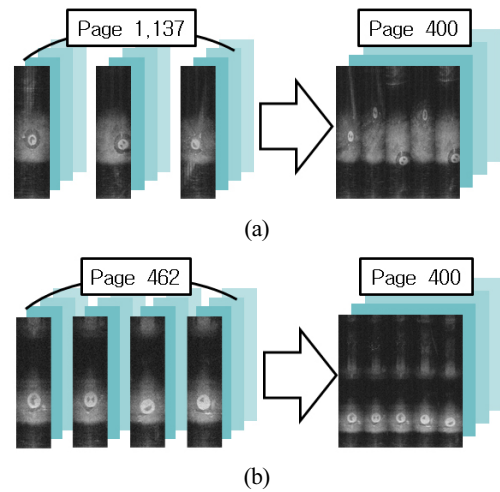
동적 데이터는 [Fig. 4]과 같이 인공 마커 ID-1, ID-2, ID-3과 마커가 없는 이미지(배경) 데이터로 크기는 96×512 픽셀이고, 총 11,144장으로 구성된다. 정적 데이터는 [Fig. 5]와 같이 인공 마커 ID-1, ID-2, ID-3, ID-4로 모두 마커가 있는 이미지 데이터로 구성되었고, 크기는 96×512 픽셀이며 총 462장이다. 또한, 본 논문에서는 학습 데이터에 사용된 동적 데이터와 정적 데이터도 테스트 데이터로 사용되었다.

### 3.1 학습 데이터

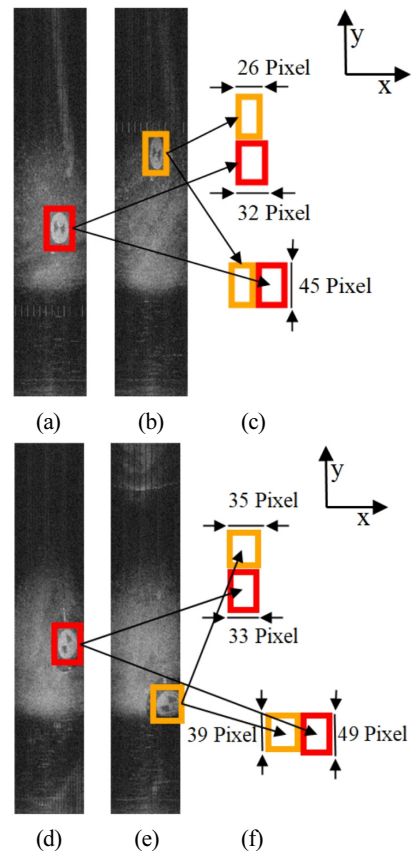
학습 데이터는 데이터 종류별로 마커가 있는 이미지 데이터만 다시 분류하였다. 이렇게 분류된 96×512 픽셀 이미지를 [Fig. 6]과 같이 랜덤으로 5개씩 선출하여 하나의 이미지로 합쳤다. 한 이미지 데이터로 재 생성된 480×512 픽셀 크기의 이미지 데이터를 각 400장씩 만들었다.

#### 3.1.1 Distortion Augmentation

데이터를 Scale 변형 기법으로 Augmentation을 할 때에는 이미지 전체의 크기를 일괄적으로 5-10%를 축소하거나 확대 시켜준다. 그러나 본 논문에서는 한 방향으로 주된 변형이 일

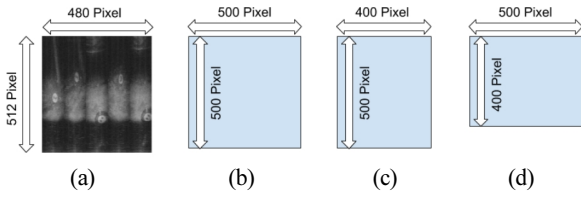


[Fig. 6] Regenerated sonar image by combining 5 images: (a) Dynamic data set, (b) Static data set



[Fig. 7] Partial distortion of sonar image: (a) Marker in the center, (b) Marker on the edge, (c) Compare the wide size of two markers, (d) Marker in the center, (e) Marker on the edge, (f) Compare the high size of two markers

어나도록 하였다. 그 이유는 [Fig. 7]과 같이 데이터를 수집하는 과정에서 이미지 내에 인공 마커의 위치에 따라 마커의 크기와 모양이 한 방향으로만 왜곡되기 때문이다. 이러한 왜곡



[Fig. 8] Distorted data for augmentation: (a) Regenerated sonar image, (b) 500×500 Pixel, (c) 400×400 Pixel, (d) 500×400 Pixel

현상이 일어나도 강인하게 동일한 마커를 분류되도록 학습 데이터를 [Fig. 8]과 같이 Distortion Augmentation을 하였다. 선행 연구로 Distortion Augmentation의 효과를 확인해보았다<sup>[6]</sup>. 학습 데이터에 Distortion Augmentation을 포함시켜, 물체 감지의 성능 결과를 [Table 1]에 정리하였고, 학습 데이터에 Distortion Augmentation을 포함 하지 않고 실험한 결과를 [Table 2]에 정리하였다. 이 실험에서는 학습 데이터와 테스트 데이터를 모두 동적 데이터로 사용하였다. [Table 1]을 보면, Distortion Augmentation한 성능 결과 정밀도는 ID-1: 1, ID-2: 1과 ID-3: 0.92의 값을 얻고, 재현율은 ID-1: 0.71, ID-2: 0.92와 ID-3: 0.75의 값을 얻는다. Distortion Augmentation을 하지 않은 [Table 1]을 보면, 정밀도는 ID-1: 0.93, ID-2: 0.46과 ID-3: 0.77의 값을 가지고, 재현율은 ID-1: 0.83, ID-2: 0.70과 ID-3: 0.57의 값을 갖는다. Distortion Augmentation을 통해 ID-2와 ID-3의 정밀도를 각 0.54와 0.15 향상시켜주었고, ID-1을 제외한 ID-2와 ID-3의 재현율도 각 0.22와 0.15 향상시켜 Distortion Augmentation의 효과를 확인 할 수 있었다.

[Fig. 7] (a)와 (d)는 이미지 내에 중심부에서 인공 마커가 위치한 이미지이고, [Fig. 7] (b)와 (e)는 [Fig. 7] (a)와 (d)보다 인공 마커가 상단 혹은 하단에 위치한 이미지이다. 마커의 크기를 [Fig. 7] (a)와 (d)는 빨간색 사각형으로 표시하고, [Fig. 7] (b)와 (e)는 주황색 사각형으로 표시하였다. [Fig. 7] (a)와 (b)에 표시된 사각형의 크기를 비교하면, 높이는 45 픽셀로 길이 차이는 없었지만, 너비의 길이가 각 32 픽셀과 26 픽셀로 한 방향으로만 약 20%의 변형이 생겼다. [Fig. 7] (d)와 (e)에 표시된 사각

[Table 1] Result of object detection with distortion augmentation

	ID-1	ID-2	ID-3
Precision	1.0	1.0	0.92
Recall	0.71	0.92	0.75

[Table 2] Result of object detection without distortion augmentation

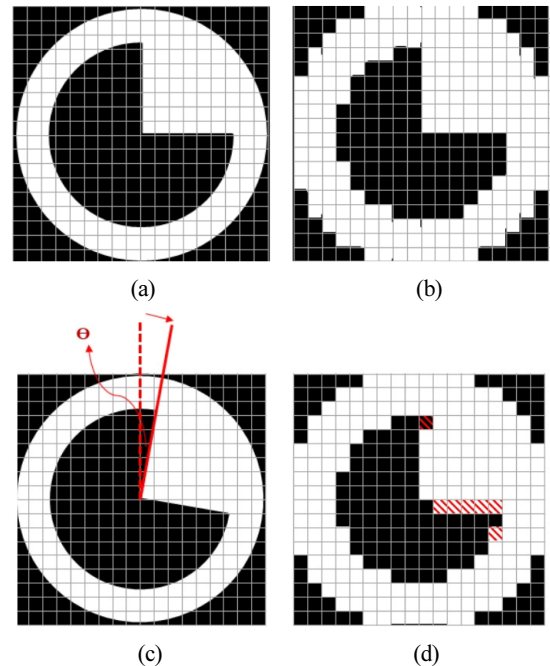
	ID-1	ID-2	ID-3
Precision	0.93	0.46	0.77
Recall	0.83	0.70	0.57

형의 크기를 비교하면, 높이는 각 49 픽셀과 39 픽셀로 약 20%의 변형이 생겼고, 너비의 길이는 각 33 픽셀과 35 픽셀로 약 6%의 변형이 생겼다. [Fig. 7] (a)와 (d)는 빨간색 사각형을 기준으로 양의 y축으로 이동하거나, x축 방향으로 마커의 이미지가 이동하게 되면 마커의 너비의 길이가 변하는 경향을 가진다. [Fig. 7] (a)와 (d)는 빨간색 사각형을 기준으로 음의 y축으로 마커의 이미지가 이동하게 되면, 마커의 높이가 변하는 경향을 관찰할 수 있었다.

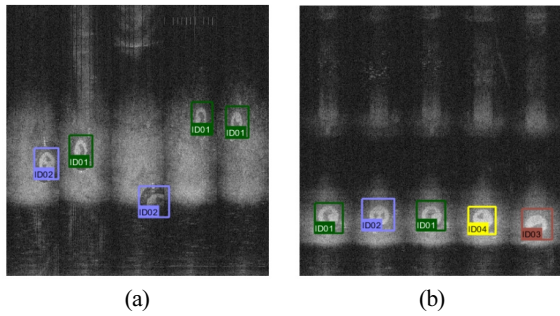
480×512 픽셀 이미지를 500×500 픽셀 크기로 변형시켜 이미지 데이터를 정사각형으로 만들고, 너비와 높이를 순차적으로 400 픽셀로 줄여 변형시켰다. 400장의 480×512 픽셀의 이미지 데이터를 500×500 픽셀, 400×500 픽셀과 500×400 픽셀로 변형시켜 총 1,200장씩 동적 데이터와 정적 데이터를 각각 학습 데이터로 준비하였다.

### 3.1.2 Rotation Augmentation

이전 연구에서는 ‘3.1.1 Distortion Augmentation’장에서 준비된 데이터를 90°씩 회전시켜주었다. 이 실험 결과에서 다른 환경의 데이터를 분류하는데, 마커 식별 부분의 이미지가 0°, 90°, 180°, 270°로 회전 시킨 이미지와 비슷한 경우 마커의 분류를 했지만, 그 외의 회전된 마커 이미지는 분류 하지 못했다. 본 논문에서는 회전 각도를 3°씩 하여 이전 연구보다 조밀하게 회전 Augmentation을 하였다. 이때 적절한 각도 3°를 찾기



[Fig. 9] Changing images of pixel units with the rotation of markers in the image: (a) Rotated marker by 0°, (b) Pixelated ‘(a)’, (c) Rotated marker by  $\theta$ °, (d) Pixelated ‘(c)’



[Fig. 10] Ground truth of training data: (a) Dynamic data set, (b) Static data set

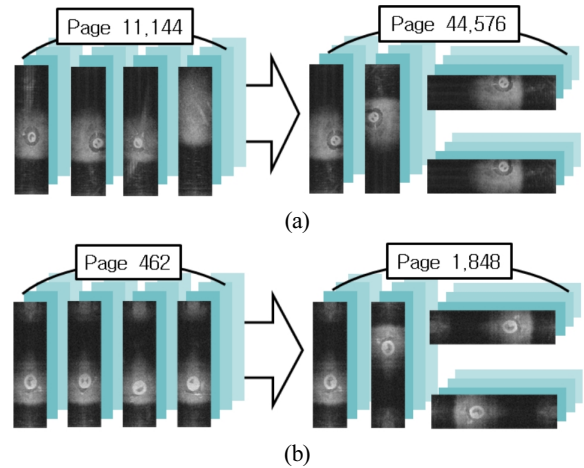
위해 [Fig. 9]와 같은 방법으로 계산하였다. [Fig. 9] (a)는 이미지 내에 마커 이미지를 픽셀 단위로 분할된 이미지이고, [Fig. 9] (b)는 [Fig. 9] (a)의 이미지를 픽셀화한 이미지이다. [Fig. 9] (b)를 보면, 분할된 픽셀 안에 마커 이미지가 조금이라도 담겨 있으면, 해당하는 픽셀에 계속적으로 이미지 정보가 있는 것으로 표시된다. [Fig. 9] (c)처럼 일정한 회전 각도  $\theta$ 만큼 마커 중심에서 마커 이미지를 회전 시킬 때, [Fig. 9] (d)처럼 마커의 이미지가 픽셀 단위로 변화가 생기는 회전 각도를 찾는다. [Fig. 9] (d)는 [Fig. 9] (b)와 비교하여 픽셀 단위로 변한 부분을 빨간색 대각선으로 표시하였다. 동적 데이터와 정적 데이터의 이미지 내에 마커의 반지름  $r$ 은 13-25 픽셀이었다. 마커가 중심에서 회전한다고 가정했을 때, 픽셀 단위로 이미지가 변하게 되는 각도  $\theta$ 는 식 (1)에 의해  $2.29^\circ$ - $4.40^\circ$ 의 사이 값이 된다.

$$Rotation\ Angle(\theta) = \arctan\left(\frac{1}{r}\right) \quad (1)$$

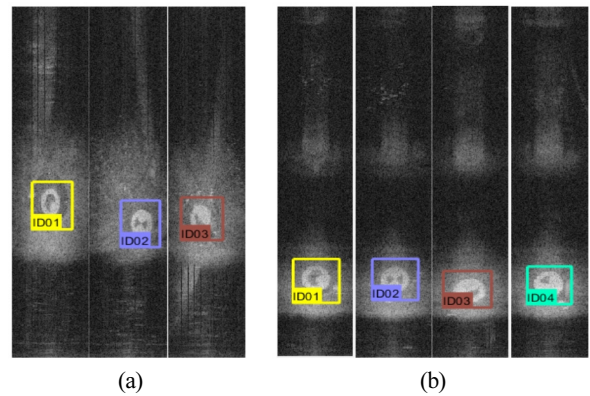
본 논문에서는 ‘3.1.1 Distortion Augmentation’장에서 준비된 데이터를  $3^\circ$ 씩 회전시켜, 동적 데이터와 정적 데이터를 각 144,000장씩 준비하였다. 준비된 데이터들 [Fig. 10]과 같이 마커의 위치와 종류를 확인하여 수작업으로 학습 데이터로 만들었다.

### 3.2 테스트 데이터

$96 \times 512$  픽셀 크기의 동적 데이터와 정적 데이터를 각 11,144장과 462장 준비 후, [Fig. 11]과 같이  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ 로 회전시켜 각 44,576장과 1,848장으로 준비하였다. 동적 데이터에는 마커 ID-1, ID-2, ID-3과 마커가 없는 이미지(배경) 데이터 모두를 사용하였고, 마커의 위치와 종류를 [Fig. 12] (a)와 같이 수작업으로 기입하였다. 또한, 정적 데이터도 마커의 위치와 종류를 [Fig. 12] (b)와 같이 직접 확인하여 기입하였고, 마커의 종류는 ID-1, ID-2, ID-3, ID-4를 사용하여 얻은 이미지 데이터이다.



[Fig. 11] Preparing test data set for object detection: (a) Dynamic data set, (b) Static data set



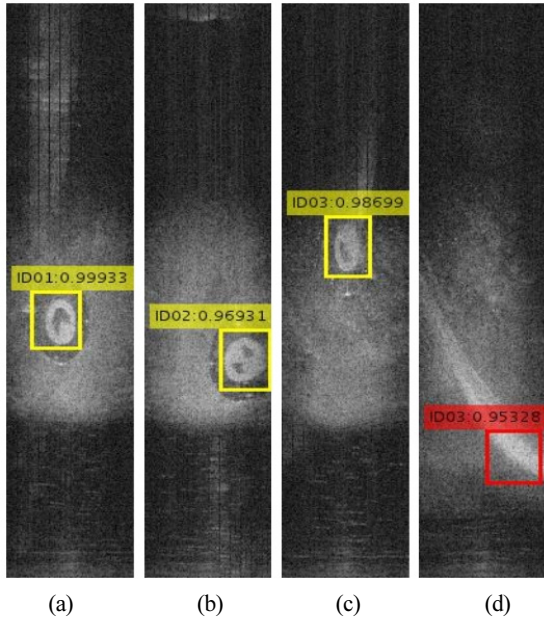
[Fig. 12] Ground truth of Test data: (a) Dynamic data set, (b) Static data set

## 4. 실험 결과

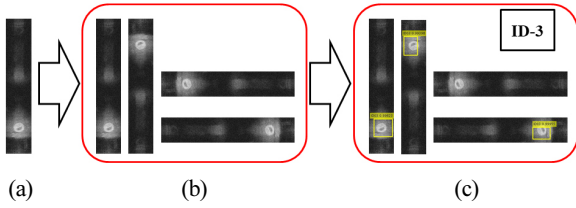
본 논문은 마커 검출 성능을 Faster R-CNN으로 하였다. Faster R-CNN의 구조는 4개의 Convolution layer와 4개의 Pooling layer로 설정하였다. 4개의 Convolution layer의 커널 크기는  $3 \times 3$ 으로 했고, 커널의 수는 순차적으로 8개, 16개, 32개, 64개로 설정하였다.

테스트 데이터의 결과는 [Fig. 13]과 같이 이미지 내의 마커에 바운딩 박스를 표시하고, 마커의 종류와 정답일 확률을 0-1의 사이 값으로 표시했다. 바운딩 박스의 색깔은 마커의 종류와 위치가 일치할 때, 노란 색으로 표시되고, 마커의 종류나 위치가 불일치할 때에는 빨간색으로 표시된다.

추가적으로 [Fig. 14]는 마커를 분류 할 때, 한 장의 이미지 데이터에서  $0^\circ$ ,  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ 로 회전된 데이터를 하나의 세트로 묶어 준 후, 각 이미지의 결과가 동일한 결과가 50%이상 넘는지 확인하여, 한 장의 이미지 데이터에 대해 마커를 분류하였다.



[Fig. 13] Test result images: (a) True positive (ID-1), (b) True positive (ID-2), (c) True positive (ID-3), (d) False positive (background)



[Fig. 14] Process for classifying an image: (a) An image data, (b) Rotated images data set, (c) Classified image data set

[Table 3]은 학습 데이터와 테스트 데이터의 종류에 따른 인공 마커 ‘ID-1’, ‘ID-2’, ‘ID-3’에 대한 Average precision을 정리한 것이다. 여기서 인공 마커를 ‘ID-1’, ‘ID-2’, ‘ID-3’만 이용한 이유는 동적 데이터와 정적 데이터에 공통적으로 사용된 인공 마커를 이용하여 실험 결과를 비교 분석하기 위해서이다. ( $^{\circ}$ ) 안의 각도는 학습 데이터를 만들 때, 회전 Augmentation의 각도를 기입한 것이다. ‘\*’의 표시가 있는 것은 [Fig. 14]의 과정을 거쳐 하나의 이미지 데이터를 분석할 때,  $0^{\circ}$ ,  $90^{\circ}$ ,  $180^{\circ}$ ,  $270^{\circ}$ 로 회전 시킨 이미지 4장을 사용한 경우이다. ‘\*’의 표시가 없는 것은 마커를 분류할 때,  $0^{\circ}$ 에 해당하는 이미지 데이터만 사용한 것을 의미한다.

본 논문에서 학습 데이터는 테스트 데이터의 일부를 가공하여 준비되었기 때문에, 같은 환경에서 수집된 데이터를 학습 데이터와 테스트 데이터로 사용한 [Table 3] (1-8)은 상대적으로 유의미한 실험 결과 값으로 보기 힘들었다. 또한, [Table 3] (9-12)는 제대로 물체 감지가 되지 않아 실험 결과 비교에서 제외하였

[Table 3] Average precision of object detection

Index	Training (Rotation angle ( $^{\circ}$ ))	Test (Data set (*))	Average precision		
			ID-1	ID-2	ID-3
(1)	Static ( $90^{\circ}$ )		0.99	1.0	0.95
(2)	Static ( $3^{\circ}$ )		1.0	1.0	1.0
(3)	Static ( $90^{\circ}$ )*		0.96	1.0	1.0
(4)	Static ( $3^{\circ}$ )*		1.0	1.0	1.0
(5)	Dynamic ( $90^{\circ}$ )		0.81	0.95	0.95
(6)	Dynamic ( $3^{\circ}$ )		0.95	0.88	0.92
(7)	Dynamic ( $90^{\circ}$ )*		0.94	0.98	0.99
(8)	Dynamic ( $3^{\circ}$ )*		1.0	0.99	1.0
(9)	Static ( $90^{\circ}$ )	Dynamic	0.27	0.21	0
(10)	Static ( $3^{\circ}$ )	Dynamic	0	0.06	0
(11)	Static ( $90^{\circ}$ )	Dynamic*	0.20	0.35	0
(12)	Static ( $3^{\circ}$ )	Dynamic*	0.08	0.15	0
(13)	Dynamic ( $90^{\circ}$ )	Static	0.82	0.62	0.53
(14)	Dynamic ( $3^{\circ}$ )	Static	1.0	1.0	0.63
(15)	Dynamic ( $90^{\circ}$ )	Static*	1.0	1.0	0.55
(16)	Dynamic ( $3^{\circ}$ )	Static*	1.0	1.0	0.99

고, 분류가 되지 않은 이유를 ‘5. 결론’에서 고찰하였다. 그러므로 ‘4.1 Rotation Augmentation’과 ‘4.2 테스트 데이터의 세트화’에서 실험 결과의 비교는 [Table 3] (13-16)을 주로 서술하였다.

#### 4.1 Rotation Augmentation

회전 Augmentation의 각도를 조밀하게 하여 학습 데이터를 만들었을 때, 물체 감지 성능이 향상 되었다. [Table 3] (13)과 (14)를 보면, ID-1, ID-2와 ID-3의 인식률이 18%, 38%와 10%씩 크게 상승하였다. 또한, [Table 3] (15)와 (16)을 보면, ID-1과 ID-2는 그대로 유지되었고, ID-3의 인식률은 44%로 크게 상승하였다. 이를 통해 학습 데이터에 회전 Augmentation을 할 때, 회전 각도를 조밀하게 하는 것이 마커의 인식률 향상에 긍정적인 영향을 주는 것을 알 수 있었다.

#### 4.2 테스트 데이터의 세트화

[Fig. 14]와 같이 테스트 데이터의 기본 이미지인  $0^{\circ}$  이미지 외에  $0^{\circ}$  이미지를  $90^{\circ}$ ,  $180^{\circ}$ ,  $270^{\circ}$ 로 회전시킨 이미지를 추가적으로 사용하여 하나의 세트에 묶어, 물체 감지를 하면 검출 성능이 향상되었다. [Table 3] (13)과 (15)를 비교하면, ID-1은 18%, ID-2는 38%, ID-3은 2%로 검출 성능이 전체적으로 크게 상승하였다. [Table 3] (14)과 (16)을 비교하면, ID-1과 ID-2는 그대로 유지되었고, ID-3의 인식률이 36%로 크게 상승하였다.

마커 분류 시 하나의 데이터로만 분류한 결과 값보다 [Fig. 14]의 과정을 거쳐, 4장의 이미지를 하나의 테스트 데이터 세트로 분류한 결과 값이 전체적으로 인식률을 향상 시켜주는걸 확인할 수 있었다.

## 5. 결 론

본 논문에서는 수중 소나 이미지에 생기는 부분 왜곡에 관한 노이즈도 강력하게 마커를 검출하기 위해 학습 데이터의 Augmentation 방법을 제안하였다. 또한, 테스트 데이터의 세트화를 통해 물체 감지 성능을 향상시켰다. 조밀한 각도로 회전 Augmentation을 하는 것은 물체 감지 성능 향상에 도움이 되었다. 또한, 이미지 데이터를 세트화 하여 마커 검출에 적용하는 방식으로 마커 검출에 안정성을 주어, 마커의 미 검출이나 오답에 대한 결과 값을 보정하여 마커의 인식률을 향상시켰다.

[Table 3] (9-12)처럼 정적 데이터로 학습하고, 동적 데이터로 테스트한 결과는 좋지 않았다. 그 이유는 정적 데이터를 취득할 때, 인공 마커와 센서가 모두 고정된 위치에 있었기 때문이다. 동적 데이터는 이동 로봇이 움직이며 데이터를 얻게 되어, 같은 마커라 해도 이미지 내에서 마커의 위치와 방향이 다르게 수집 되었지만, 정적 데이터는 마커의 이미지 데이터 간에 차이가 없었다. 그 결과 462장의 정적 데이터가 실질적으로 종류별로 하나의 데이터로 4장인 꼴이 되어, Faster R-CNN으로 학습 시, 데이터의 일반화 특성이 떨어지는 현상이 생겨, 학습 데이터로는 적합하지 못했기 때문이다.

추 후 연구에서는 이미지내의 마커 검출 결과를 접목하여 SLAM을 기반으로 한 이동 로봇의 위치 추정 및 경로 생성 연구를 할 예정이다.

## References

- [1] D. Ribas, P. Ridao, J. D. Tardós, and J. Neira, "Underwater SLAM in Man-Made Structured Environments", *Journal of Field Robotics*, vol. 25, no. 11-12, pp. 898-921, 2008.
- [2] S. Lee, "Deep Learning of Submerged Body Images from 2D Sonar Sensor based on Convolutional Neural Network", *2017 IEEE Underwater Technology(UT)*, Busan, South Korea, pp. 1-3, 2017.
- [3] Y.-S. Shin, Y. Lee, H.-T. Choi, and A. Kim, "Bundle Adjustment and 3D Reconstruction Method for Underwater Sonar Image", *Journal of Korea Robotics Society*, vol. 11, no. 2, pp. 51-59, 2015.
- [4] Y. Lee, J. Lee, and H.-T. Choi, "A Framework of Recognition and Tracking for Underwater Objects based on Sonar Images: Part 1. Design and Recognition of Artificial Landmark considering Characteristics of Sonar Images", *Journal of the Institute of Electronics and Information Engineers*, vol. 51, no. 2, pp. 182-189, 2014.
- [5] 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, Jun., 2017.
- [6] Shark Marine Technologies Inc., *Navigator: Diver Held Sonar Imaging and Navigation System*, [Online], <http://www.sharkmarine.com>, Accessed: December 5, 2018.
- [7] P. Zhu, J. Isaacs, B. Fu, and S. Ferrari, "Deep learning feature extraction for target recognition and classification in underwater sonar images", *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, Melbourne, VIC, pp. 2724-2731, 2017.
- [8] D. P. Williams, "Underwater target classification in synthetic aperture sonar imagery using deep convolutional neural networks", *2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, pp. 2497-2502, 2016.
- [9] L. Jian, H. Yang, L. Zhong, and X. Ying, "Underwater Target Recognition Based on Line Spectrum and Support Vector Machine", *International Conference on Mechatronics, Control and Electronic Engineering (MCE2014)*, pp. 79-84, 2014.
- [10] X. Cao, X. Zhang, Y. Yu, and L. Niu, "Deep Learning-Based Recognition of Underwater Target", *2016 IEEE International Conference on Digital Signal Processing (DSP)*, Beijing, China, pp. 89-93, 2016.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Neural Information Processing Systems*, pp. 84-90, Jun., 2017.
- [12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580-587, 2014.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition", *European Conference on Computer Vision*, pp. 346-361, 2014.
- [14] R. Girshick, "Fast R-CNN", *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 1440-1448, 2015.
- [15] Stanford University, *Lecture 7: Training Neural Networks, part II(Update rules, ensembles, data augmentation, transfer learning)*, [Online], <http://cs231n.stanford.edu/syllabus.html>, Accessed: April 24, 2018.
- [16] E. Lee, Y. Lee, J. Choi, and S. Lee, "Study of Marker Detection Performance via Training Data Augmentation for Partial Distortion of Underwater Sonar Image", *The Korean Society of Mechanical Engineers Annual Conference*, pp. 2176-2181, 2018.



**이 언 호**

2009 공주대학교 기계자동차공학부 기계설계  
전공(공학사)  
2014 공주대학교 기계공학과(공학석사)  
2018~현재 공주대학교 기계공학과(공학박사)

관심분야: 물체인식, 수중 로봇, 심화학습



**최 진 우**

2003 포항공과대학교 기계공학과(공학사)  
2005 포항공과대학교 기계공학과(공학석사)  
2009 포항공과대학교 기계공학과(공학박사)  
2013~현재 선박해양플랜트연구소 선임연구원

관심분야: 수중 로봇, 위치인식, SLAM



**이 영 준**

2009 충남대학교 메카트로닉스공학과(공학사)  
2014 충남대학교 메카트로닉스공학과(공학  
석사)  
2011~현재 선박해양플랜트연구소 기술원

관심분야: 소나 영상처리, 영상 소나



**이 세 진**

2003 한양대학교 기계공학과(공학사)  
2005 포항공과대학교 기계공학과(공학석사)  
2009 포항공과대학교 기계공학과(공학박사)  
2013~현재 공주대학교 기계자동차공학부  
부교수

관심분야: 이동로봇, 지도작성, 위치추정, 심화학습