

사이드 스캔 소나 영상에서 수중물체 자동 탐지를 위한 컨볼루션 신경망 기법 적용

The application of convolutional neural networks for automatic detection of underwater object in side scan sonar images

김정문,¹ 최지웅,^{1†} 권혁종,¹ 오래근,¹ 손수욱²

(Jungmoon Kim,¹ Jee Woong Choi,^{1†} Hyuckjong Kwon,¹ Raegun Oh,¹ and Su-Uk Son²)

¹한양대학교, ²국방과학연구소

(Received December 27, 2017; revised February 20, 2018; accepted March 29, 2018)

초 록: 본 논문은 사이드 스캔 소나 영상을 컨볼루션 신경망으로 학습하여 수중물체를 탐색하는 방법을 다루었다. 사이드 스캔 소나 영상을 사람이 직접 분석하던 방법에서 컨볼루션 신경망 알고리즘이 보강되면 분석의 효율성을 높일 수 있다. 연구에 사용한 사이드 스캔 소나의 영상 데이터는 미 해군 수상전센터에서 공개한 자료이고 4종류의 합성수중물체로 구성되었다. 컨볼루션 신경망 알고리즘은 관심영역 기반으로 학습하는 Faster R-CNN(Region based Convolutional Neural Networks)을 기본으로 하며 신경망의 세부사항을 보유했던 데이터에 적합하도록 구성하였다. 연구의 결과를 정밀도-재현율 곡선으로 비교하였고 소나 영상 데이터에 지정한 관심영역의 변경이 탐지성능에 미치는 영향을 검토함으로써 컨볼루션 신경망의 수중물체 탐지 적용성에 대해 살펴보았다.

핵심용어: 사이드 스캔 소나, 관심영역 기반 컨볼루션 신경망, 수중물체탐지, 정밀도-재현율

ABSTRACT: In this paper, we have studied how to search an underwater object by learning the image generated by the side scan sonar in the convolution neural network. In the method of human side analysis of the side scan image or the image, the convolution neural network algorithm can enhance the efficiency of the analysis. The image data of the side scan sonar used in the experiment is the public data of NSWC (Naval Surface Warfare Center) and consists of four kinds of synthetic underwater objects. The convolutional neural network algorithm is based on Faster R-CNN (Region based Convolutional Neural Networks) learning based on region of interest and the details of the neural network are self-organized to fit the data we have. The results of the study were compared with a precision-recall curve, and we investigated the applicability of underwater object detection in convolution neural networks by examining the effect of change of region of interest assigned to sonar image data on detection performance.

Keywords: Side scan sonar, Faster R-CNN (Region based Convolutional Neural Networks), Underwater object detection, Precision-recall

PACS numbers: 43.60.Lq, 43.30.Zk, 43.30.Vh

1. 서 론

사이드 스캔 소나는 음파를 이용하여 해저를 2차

원의 영상으로 보여주는 장비이다. 사이드 스캔 소나로 생성한 영상은 일반적인 사진인 광학으로 생성한 영상과 다르다. 사이드 스캔 소나 영상(이하 ‘소나 영상’)은 음파가 물체에 부딪혀 직접 반사되는 부분에 의해 만들어지는 하이라이트와 물체에 의해 음파가 도달하지 않는 부분에 생기는 그림자로 표현되고

†Corresponding author: Jee Woong Choi (choijw@hanyang.ac.kr)
Department of Marine Science and Convergence Engineering,
Hanyang University, 55 Hanyangdaehak-Ro, Sangnok-Gu, Ansan,
15588, Republic of Korea
(Tel: 82-31-400-5531, Fax: 82-31-406-6255)

“이 논문은 2017년도 수중음향학 학술발표회에서 발표하였던 논문입니다.”

수중에 존재하는 여러 가지 소음으로 인해 음파를 사용하는 사이드 스캔 소나 영상의 해상도가 낮다는 특성이 있다.^[1] 그러므로 소나 영상은 눈으로 볼 수 없는 해저를 영상화하여 볼 수 있지만 음파의 특성 때문에 광학으로 촬영한 영상보다는 표현의 제한이 있다. 소나 영상을 활용하는 이유는 수중에서 광파는 전달에너지의 감쇠가 심한 반면에 음파의 경우는 좋은 전달 매체로 작용하기 때문이다.^[2] 따라서 음파를 이용한 소나 영상은 하이라이트와 그림자를 확인하여 물체의 존재 유무 및 종류를 구분할 수 있으므로 광범위한 해저탐색 분야에 활용도가 높다.^[3]

한편, 인공지능에 관한 연구는 다양한 분야에서 활발히 진행 중이며, 그중 한 예로 2010년부터 매년 열리고 있는 영상 인식 분야의 알고리즘을 평가하기 위한 ‘이미지넷 대용량 영상인식대회(ILSVRC: ImageNet Large Scale Visual Recognition Challenge)’가 있다. 이 대회에서 Olga Russakovsky는 알고리즘의 학습 및 테스트용으로 사용된 데이터셋을 사람에게 실험하였으며, 2012 ~ 2014년도에 사용되었던 데이터셋 중 1,500개의 무작위 샘플을 사람에게 분류하도록 실험한 결과 에러율이 5.1%였다.^[4] 그 후 Kaiming He의 연구결과 인공지능을 통한 알고리즘의 에러율이 4.9%를 달성하면서 인공지능 영상 인식이 사람의 인식률을 넘어섰으며, 이를 계기로 인공지능을 이용한 영상 인식이 다양한 분야에서 많은 관심을 받고 있다. 영상 인식 분야에서 주로 활용되는 알고리즘은 컨볼루션 신경망^[5]을 응용한 것이며, 이것은 표지판 인식,^[6] 얼굴 인식,^[7] 손글씨 인식^[8] 등 여러 종류의 영상 인식 분야에서 주목할 만한 성능을 보여주고 있다.^[9]

일반적인 광학 영상을 인식하는 컨볼루션 신경망을 사이드 스캔 소나로 생성한 소나 영상의 인식에도 적용한다면 비슷한 성능을 얻으리라고 기대할 수 있다. 본 연구의 목적은 컨볼루션 신경망을 소나 영상 인식에 적용함으로써 수중물체를 효과적으로 탐지하는 것이다. 그러나 컨볼루션 신경망은 빛으로 표현되는 광학 영상에서 학습이 잘 되고 물체를 잘 인식할 수 있으나, 음파로 표현되는 소나 영상에서도 유사한 효과를 얻을 수 있는지 검증할 필요가 있다. 또한 광학 영상과 소나 영상의 차이가 컨볼루션

신경망의 적용에 어떠한 영향을 미치는지 확인하는 것을 검토사항으로 하였다.

광학 영상 인식분야 활용되는 컨볼루션 신경망을 소나 영상에서도 동일하게 활용하기 위해서는 소나 영상의 특성을 잘 이해하고 적합하게 적용해야 한다. 광학 영상과 소나 영상은 대상 물체를 표현하는데 차이점이 있다. 광학 영상은 광원과 카메라의 위치에 따라 다양한 표현이 가능하고 렌즈를 조절하여 물체를 뚜렷하게 기록할 수 있으며 주로 빛을 받는 물체의 형태와 색상에 의해 표현된다.^[10] 반면에 소나 영상은 음원과 수신부가 같은 위치에 있고 배경 소음에 의한 간섭이 존재하므로 광학 영상보다는 표현이 제한적이며 물체에 의한 하이라이트와 그림자로 표현된다. 그리고 광학 영상은 화각에 따라 한 번에 일정 범위를 기록할 수 있지만, 소나 영상은 데이터 획득 방법의 차이로 인해 여러 연속적인 신호를 연결해야 사람이 인지할 수 있는 수준의 영상이 만들어진다. 따라서 본 연구에서는 컨볼루션 신경망을 적용할 때 소나 영상과 광학 영상의 차이점이 결과에 어떠한 영향을 미치는지 확인하였다. 또한 본 연구에서 사용한 컨볼루션 신경망은 관심영역(Region Of Interest, ROI)을 어떻게 지정하느냐에 따라서 탐지 성능의 효과가 달라질 수 있기 때문에 광학 영상과 소나 영상을 컨볼루션 신경망에 적용한 결과의 차이점을 살펴본 후 컨볼루션 신경망의 관심영역 변경이 소나 영상의 학습 결과에 미치는 영향을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서 사이드 스캔 소나를 통해 영상을 생성하는 이론에 대해 다루었고, 3장에서 컨볼루션 신경망이 영상을 학습하고 탐지하는 원리를 설명하였으며, 4장에서 컨볼루션 신경망의 학습에 사용한 소나 및 광학 영상의 데이터셋과 실험 세부 설정에 대해 설명하였다. 5장에서 실험의 결과를 분석한 뒤 6장에서 본 논문의 결론을 제시한다.

II. 사이드 스캔 소나의 영상 생성

사이드 스캔 소나는 음파를 이용하여 넓은 해저 표면을 시각화하기에 매우 효과적이기 때문에 해저

지형 조사 및 수중물체 탐색 등의 목적으로 활용된다.^[3] 사이드 스캔 소나의 시스템은 조절 및 표시장치, 센서장치(수중 예인체), 예인케이블로 구성된다. 주파수 대역은 해저면을 영상화할 수 있는 1 kHz~1 MHz의 고주파수 대역을 사용한다.^[10] 사이드 스캔 소나의 전형적인 빔 형태는 Fig. 1(a)와 같이 수평적으로 좁고 Fig. 1(b)와 같이 수직적으로 넓다.^[3] 그러므로 하나의 소나 빔을 신호처리하면 가로방향으로 긴 한 줄의 영상을 얻을 수 있고 영상들을 계속 연결하면 사람이 인지할 수 있는 영상이 생성된다.

사이드 스캔 소나의 신호처리 중 핵심은 시간변화이득(Time Varied Gain, TVG)이다. 수중 예인체에 가까운 해저에서 수신되는 반사파는 상대적인 에너지가 높은 상태이고 시간에 비례하여 멀리서 오는 반사파는 더 많은 전달손실이 발생하므로 에너지가 상대적으로 낮다. 따라서 시간에 따라 이득을 조절해야 동일한 수준의 에너지로 신호를 표시할 수 있

다.^[10] 수신된 에너지의 세기에 따라 색을 입히고, 동일한 과정으로 처리된 자료를 연결하면 수중 예인체가 이동한 경로의 음파주사 영역 전체에 대한 영상을 만들 수 있다. 사이드 스캔 소나의 영상은 단지 신호의 강도로만 표시되기 때문에 광학 영상처럼 물체의 고유색깔이 나타나지 않고 시스템 설정 색깔에서 채도의 차이만 있는 영상이 된다.^[3]

사이드 스캔 소나의 영상은 Fig. 2에서 보는 것처럼 수주(water column)와 해저면(seabed)으로 나눌 수 있다. 수주의 넓이는 음파가 최단의 해저면에 도달하기 전까지 통과한 거리를 뜻하며, 해저면 영상의 밝기차이는 반사된 에너지 차이로 인해 발생한다. Fig. 2에서 밝기가 강한 부분(하이라이트)의 우측에 어두운 부분이 있는 것은 해저면의 굴곡이나 수중 물체에 의해 음파가 도달하지 못하는 음영지대임을 뜻한다. 즉 해저면 언덕이나 물체의 그림자이다. 실제 해저면을 보다 정확하게 표현하기 위해서는 영상

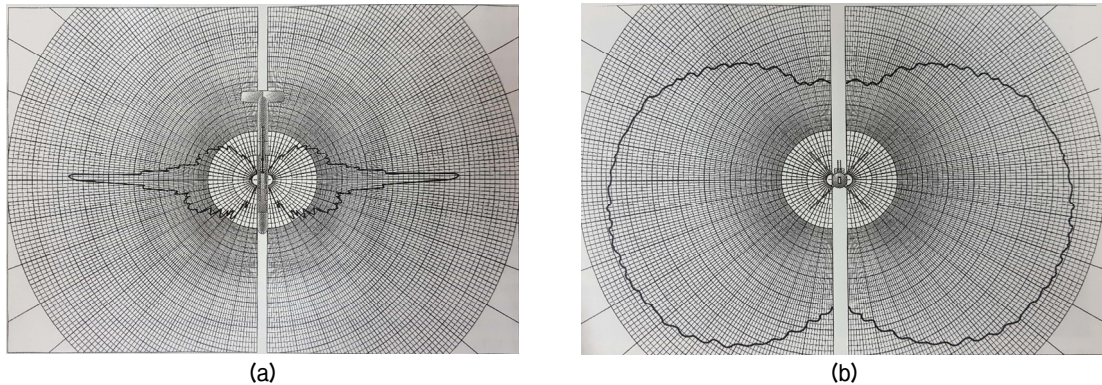


Fig. 1. A typical beam pattern for side scan sonar^[3], (a) horizontal beam pattern, (b) vertical beam pattern.

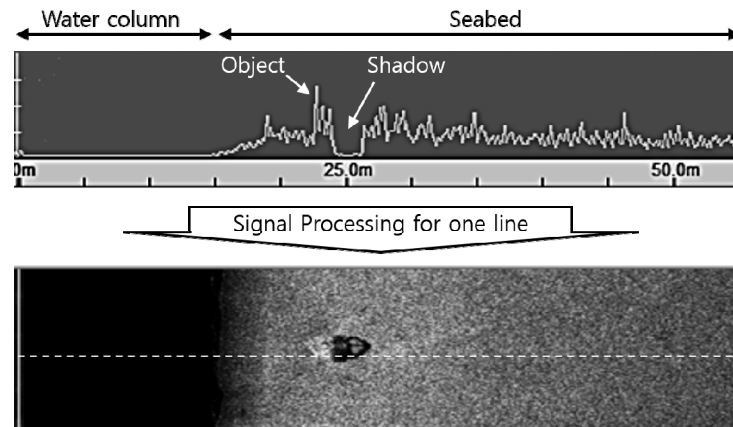


Fig. 2. Signal processing for one line.

자료를 생성할 때의 왜곡을 보정해야 한다. 자료 왜곡의 원인은 예인 선박 속도의 변화와 수중 예인체 고도에 기인하는 주사 방향 압축 현상 등을 들 수 있다.^[3] 왜곡된 자료를 실제와 가까운 해저면으로 표현하기 위해서 속도와 빔면거리를 보정하여 신호를 처리한다. 본 논문에서는 최종 보정된 소나 영상을 컨볼루션 신경망의 학습에 활용하였다.

III. 컨볼루션 신경망을 활용한 영상 학습

본 연구에서는 컨볼루션 신경망 알고리즘 중에서 관심영역 기반으로 물체를 학습하고 탐지하는 Faster R-CNN(Region based Convolutional Neural Networks)을 사용하였다. Faster R-CNN은 R-CNN, Fast R-CNN에 비해서 RPN(Region Proposal Networks)을 개선하면서 정확도 및 처리속도를 높인 알고리즘이다.^[11] 이 알고리즘의 적용 과정은 전처리 단계, 학습 단계, 테스트 단계로 구분된다. 전처리 단계에서는 데이터를 학습 단계에 적용할 수 있도록 사용자의 직접적인 작업이 진행된다. 학습시킬 영상 데이터에서 목표 물체를 관심영역으로 지정하며, 관심영역의 위치정보는 영상 내의 픽셀 값을 이용하여 저장된다. 그리고 해당 관심영역의 클래스를 설정하여 다른 영상에서도 동일한 목표 물체로 인식될 수 있도록 한다. 전

처리가 완료된 데이터는 영상 파일, 영상의 관심영역, 관심영역의 클래스로 구성된다. 학습 단계에서는 학습 데이터를 이용하여 알고리즘의 레이어에 설정된 값으로 영상의 특징을 추출하고 특징 값과 관심영역을 서로 비교하면서 신경망의 파라미터를 계속 조정한다. 파라미터가 특징 값과 관심영역을 연결하는 최적의 값으로 설정하는 것이 학습 단계의 목표이고 신경망의 각 레이어의 설정에 따라 파라미터는 다를 수 있다. 테스트 단계에서는 학습 데이터와 별도로 준비된 테스트 데이터를 사용하여 알고리즘의 성능을 평가한다. 학습 단계에서 최적화된 파라미터는 테스트 데이터를 이용하여 학습된 관심영역과 비교하여 특징 값이 얼마나 일치하는지 확률 값으로 나타내는데 사용된다.

본 연구에서 사용한 컨볼루션 신경망은 Table 1에서와 같이 입력 레이어, 컨볼루션 레이어, ReLU(Rectified Linear Unit)레이어 등으로 구성되어 있으며 각 레이어의 역할은 다음과 같다. 우선 입력 레이어에서 영상 데이터의 값이 행렬로 입력되어진 후, 컨볼루션 레이어에서 임의로 정해진 필터와 컨볼루션된다. 이 값들 중 음수는 ReLU 레이어에서 0으로 설정된다. 이후 Dropout 레이어에서 일정한 비율의 값을 랜덤하게 추출하고 나머지만 다음 레이어로 보내지며^[12] Max pooling 레이어에서는 앞 레이어의 값 중 일정한

Table 1. Layers of convolutional neural network.

Order	Type	Size	Stride	Padding
1	Input	16×16×3	、	、
2	Convolution	7×7@96	1	3
3	ReLU	max(0,x)		
4	Dropout	0.2 rate random drop		
5	Max pooling	5×5	1	2
6	Convolution	7×7@96	1	3
7	ReLU	max(0,x)		
8	Dropout	0.2 rate random drop		
9	Max pooling	14×14	6	4
10	Fully Connected	64×384	、	、
11	ReLU	max(0,x)		
12	Dropout	0.5 rate random drop		
13	Fully Connected	2×64	、	、
14	Softmax	2×1	、	、
15	Classification	Object, Background		

행렬구역의 최대값을 샘플링한다. 이 단계까지가 입력된 영상의 주된 특징을 필터를 이용하여 추출하는 단계에 해당한다, 다음 단계부터는 물체의 클래스를 분류하는 기능이 수행된다.^[13] 클래스 분류의 첫 단계인 Fully Connected 레이어는 이전 레이어까지의 과정에서 추출된 영상 특징 값과 설정된 클래스를 이용하여 파라미터를 최적화하며, Softmax 레이어는 영상의 특징과 클래스가 일치하는 정도를 수치로 나타낸 값들을 정규화하고, 분류(Classification) 레이어는 높은 값을 획득한 클래스와 정규화 값을 출력하며 최종적으로 물체를 분류한다.

IV. 데이터 구성과 실험

실험에 사용한 소나 영상 데이터는 미 해군 수상전센터(NSWC : Naval Surface Warfare Center)에서 공개한 자료로 2003년에 해상실험으로 획득했으며, 소나의 주파수와 같은 세부 정보는 미공개 되어있다. 하지만 사이드 스캔 소나의 신호처리과정을 거쳐서 나온 영상 데이터이므로 실험에 사용 가능하다. 총 4종류의 합성수중물체가 포함되며 종류별 70장 내외의 영상이고 총 282장을 사용하였다.^[14]

Fig. 3에서와 같이 Target 1은 원형물체이며 하이라이트 부분이 넓지만, Target 2는 Target 1 보다 작은 원형물체이며 하이라이트 부분이 좁다. Target 3과 Target 4는 각각 마름모꼴과 각뿔이 눌려져 있는 형태로 수중 예인체의 스캔 방향에 따라 모양이 다르게 보일

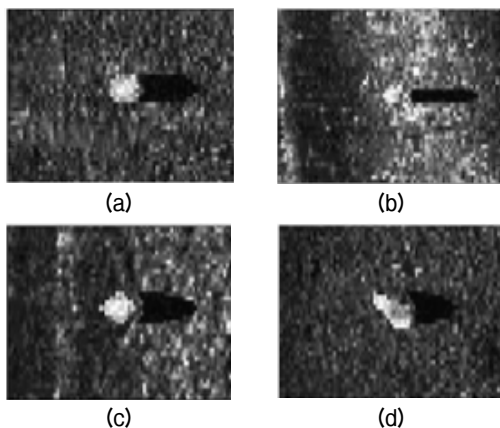


Fig. 3. Target types of sonar image data, (a) Target 1, (b) Target 2, (c) Target 3, (d) Target 4.

수 있지만, 본 논문에 사용된 자료는 해저면으로부터 수중 예인체까지의 고도가 거의 일정하여서 각 영상 데이터마다 Target들의 크기 비율도 일정하였다. 수중물체가 설치되었던 지역의 해저환경은 Fig. 4 처럼 수중물체가 잘 구별되는 평평한 지역부터 잘 구별되지 않는 암반과 해초가 있는 지역까지 차이가 있다. Fig. 4(a)는 암반 해저면, (b), (c)는 각각 켈프, 해초가 뒤덮인 해저면을 사이드 스캔 소나로 탐색한 결과이다.

실험에 사용한 광학 영상 데이터는 소나 영상의 수중물체와 비슷한 물체를 선정하려고 시도하였다. Fig. 5의 (a)와 (b)는 각각 테니스공과 골프공으로 주변에서 접할 수 있는 물체들을 직접 촬영하였다. 대체로 영상에서 물체들이 차지하는 비율을 일정하게 촬영하였고, 촬영 환경을 조정할 수 있는 물체는 비슷한 색상의 다른 물체도 영상 데이터에 포함되도록 구성하였다. 광학 데이터는 종류별 80장 내외의 영상이고 총 164장을 사용하였다.

소나 및 광학 영상은 컨볼루션 신경망에 적용할 때 3가지의 데이터셋으로 구분하였다. 총 데이터 수량의 60 %는 학습 데이터, 20 %는 교차검증 데이터, 나머지 20 %는 테스트 데이터로 구분하였다. 각 데이터가 섞이지 않도록 하는 것은 효과적인 학습과 정확한 테스트를 위해서 필요하다. 교차검증 데이터

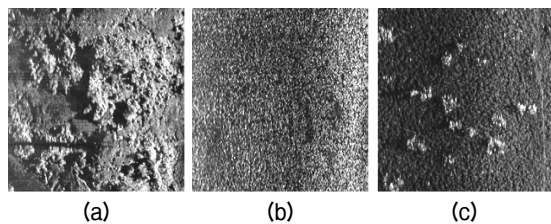


Fig. 4. Seabed image types by side scan sonar, (a) Rocky Bottom, (b) Kelp, (c) Sea Grass.



Fig. 5. Target types of optical image data, (a) Tennis ball, (b) Golf ball.

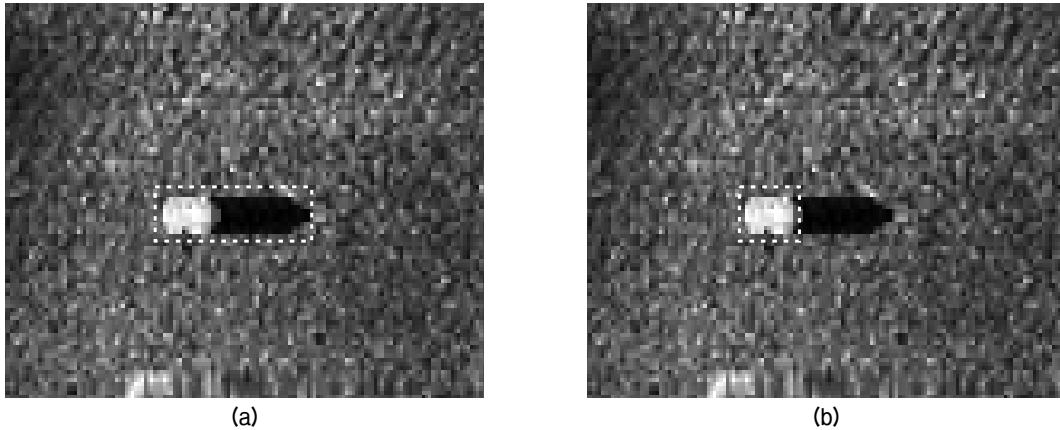


Fig. 6. Settings of ROI, (a) highlight + shadow, (b) highlight only.

는 학습단계를 거친 파라미터들을 이용하여 컨볼루션 신경망의 과적합과 부적합을 줄여주는 역할을 한다. 테스트 데이터는 테스트 결과를 신뢰하기 위해서 학습 및 교차검증 단계에서 사용되지 않아야 한다. 전처리 단계에서 Fig. 6(a)처럼 소나 영상 데이터에는 관심영역을 물체의 하이라이트 부분과 그림자를 같이 지정하였고, 광학 영상 데이터에는 물체의 형태부분으로 관심영역을 지정하였다.^[11] 이는 소나 이미지 데이터와 광학 영상 데이터에서 물체의 형태를 표현하는 요소가 서로 다르기 때문에 차이를 둔 것이다. 소나 영상 데이터 4종류, 광학 영상 데이터 2종류를 컨볼루션 신경망에 각각 학습시켰다. 학습 데이터와 교차검증 데이터 모두 200회(epoch) 학습하였다.

V. 실험결과 고찰

5.1 정밀도-재현율 곡선에 의한 알고리즘 진단

진단방법은 Fig. 7과 같이 정밀도-재현율(precision-recall) 곡선을 사용하였다.^[15] 곡선 아래의 면적을 평균 정밀도(Average Precision, AP)라고 하며, 이 지표는 ILSVRC에서 인공지능 알고리즘을 진단하여 순위를 매길 때 사용되는 지표이다. Fig. 8과 같이 TP(True Positives)는 정확히 탐지한 실패적(실패)을, FP(False Positives)는 실패적이라고 오인한 허위표적(false target)을, TN(True Negatives)는 실패적(true target)이 아니라고 정확히 제외한 허위표적(실패)을, FN(False Negatives)는 탐지

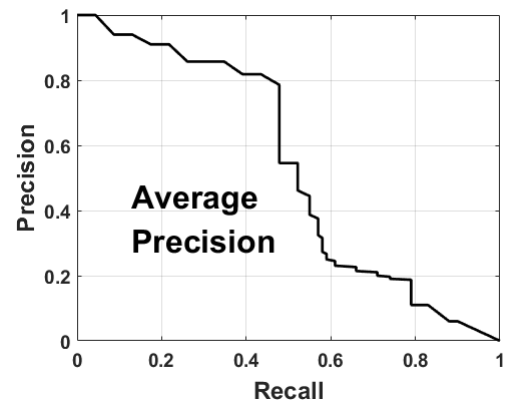


Fig. 7. Average precision.

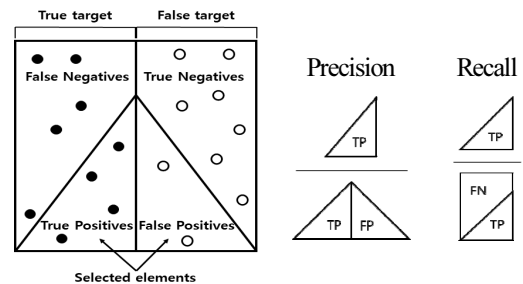


Fig. 8. Precision and recall (TP: true positives, TN: true negatives, FP: false positives, FN: false negatives).^[15]

하지 못한 실패적(실패)을 각각 나타낸다. 정밀도는 탐색 결과 중에서 허위표적에 대해 실패적(실패)이 포함되어 있는 비율이고, 재현율은 전체 실패적(실패) 중에서 탐색 결과에 실패적(실패)이 포함되어 있는 비율이다. 정밀도와 재현율은 각각 0에서 1까지의 값으로 나타나며, Fig. 7과 같이 정밀도-재현율 곡선의 아래 면적의 최대도 1이다. 따라서 알고리즘을 평가한 결과의 평균 정밀

도가 1이라면 허위표적이 포함되지 않았으며 실표적을 전부 탐지한 것이다. 허위표적이 포함되거나 실표적을 전부 탐지하지 못했을 경우에는 평균 정밀도가 1보다 낮다.

정밀도와 재현율을 분리하여 하나의 지표로만 놓고 본다면 자료 해석의 오류를 일으킬 수 있다. 예를 들어 찾아야 할 수증물체가 10개가 있다고 가정할 때, 탐지결과 2개의 물체가 확인되었고 2개 모두 실표적이라면 정밀도는 1이고 재현율은 0.2이다. 반대로 탐지결과 40개의 물체가 포함되어 있는데 그중에서 10개는 실표적이고 나머지 30개는 허위표적이라면 정밀도는 0.25이고 재현율은 1이다. 따라서 정밀도나 재현율 중 하나만 지표로 사용한다면 왜곡된 진단을 할 수 있다. 그러므로 먼저 알고리즘의 종합적인 성능을 평가할 수 있는 평균 정밀도를 확인한 후 세부적인 사항으로 정밀도와 재현율을 해석하는 것이 타당하다.

5.2 소나 영상과 광학 영상의 컨볼루션 신경망 적용 결과에 대한 고찰

소나 영상과 광학 영상을 각각 컨볼루션 신경망으로 학습하고 테스트한 결과는 Table 2의 (a), (b)와 같다. 각 클래스들의 평균 정밀도를 평균한 값을 mAP(mean Average Precision)라고 한다. 소나 영상의 각 클래스들(Target 1 ~ 4) mAP는 0.74이고, 광학 영상의 각 클래스들(테니스공, 골프공) mAP는 0.92이다.

Table 2. Average precision of optical image and sonar image: (a) optical image, (b) sonar image (highlight+shadow), (c) sonar image (highlight only).

Class		Average precision
(a) Optical image	Tennis ball	0.9421
	Golf ball	0.8990
(b) Sonar image (highlight+shadow)	Target 1	0.6963
	Target 2	0.7749
	Target 3	0.6154
	Target 4	0.8667
(c) Sonar image (highlight only)	Target 1	0.9982
	Target 2	0.7719
	Target 3	0.6818
	Target 4	0.8832

따라서 소나 영상과 광학 영상의 mAP는 각 클래스에 대한 탐지성능이 양호한 수준이라고 할 수 있다. 다시 말하면, 자체 구성한 컨볼루션 신경망이 소나 및 광학 영상을 학습하고 탐지하는 데 잘 작동한다는 것을 보여준다.

한편, 소나 및 광학 영상의 mAP를 서로 비교해보면 소나 영상의 mAP가 0.18정도 낮은 값을 보이므로 그 차이에 대해서 고찰하였다. 실험에 사용한 데이터들을 비교해보면 광학 영상은 소나 영상보다 다양한 색상과 복잡한 형태로 되어 있다. 반면에 소나 영상은 하이라이트와 그림자로 구성되어 있는 단순한 형태로 볼 수 있음에도 불구하고 mAP가 상대적으로 낮았다. 이 원인을 검토한 결과 소나 영상에서 컨볼루션 신경망의 학습에 저해되는 요소인 물체의 그림자를 관심영역에 포함했기 때문이라고 추정할 수 있다. 사이드 스캔 소나의 특성상 수중 예인체로부터의 거리에 따라 물체에 의해 생성되는 그림자의 길이는 변하지만 하이라이트 부분은 수중 예인체와의 거리와 상관없이 일정하다. 따라서 소나 영상에 대한 탐지성능을 더 향상시키기 위해서는 관심영역에서 그림자 부분을 제외하는 것이 타당한 것으로 판단된다.

5.3 소나 영상의 변경된 관심영역에 대한 실험 결과

소나 영상의 관심영역을 Fig. 6(b)처럼 하이라이트 부분만으로 지정한 후 컨볼루션 신경망에 학습 및 테스트한 결과는 Table 2의 (c)와 같다. 각 클래스들의 mAP는 0.83이다. 이 결과는 관심영역을 하이라이트와 그림자로 지정했던 Table 2의 (b)보다 mAP가 0.09 상승하였다. 5.2장의 소나 영상과 광학 영상의 컨볼루션 신경망 적용 결과에 대한 고찰에서 추정하였던, 소나 영상의 그림자가 컨볼루션 신경망의 학습에 저해되는 요소라는 가설이 적절하다는 것을 확인하였다.

알고리즘의 전체적인 성능은 하이라이트와 그림자를 관심영역으로 지정한 데이터보다 하이라이트만 관심영역으로 지정한 데이터에 대해서 더 양호하다는 것을 평균 정밀도로써 알 수 있다. 세부적인 성능은 정밀도-재현율 곡선 그래프를 통하여 확인하

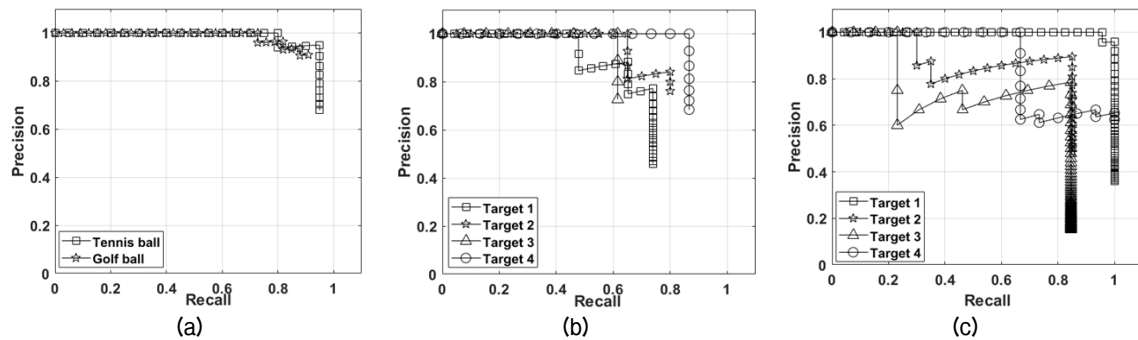


Fig. 9. Precision-recall curves, (a) optical image, (b) sonar image (highlight+shadow), (c) sonar image (highlight only).

Table 3. Precision and recall of sonar image.

Class		Highlight + shadow	Highlight only
Precision	Target 1	0.7727	0.9583
	Target 2	0.8421	0.8947
	Target 3	1	0.7857
	Target 4	1	0.6522
Average		0.9037	0.8227
Recall	Target 1	0.7391	1
	Target 2	0.8000	0.8500
	Target 3	0.6154	0.8462
	Target 4	0.8667	1
Average		0.7553	0.9241

였다.

Fig. 9(b)는 소나 영상의 관심 영역으로 하이라이트와 그림자를 같이 지정한 경우이며, Fig. 9(c)는 하이라이트만 지정한 경우의 정밀도-재현율 곡선 그래프이다. 각 클래스에서 재현율이 최대일 때 정밀도의 최댓값을 Table 3과 같이 확인할 수 있다. 관심영역을 하이라이트와 그림자로 지정한 데이터와 관심영역을 하이라이트로만 지정한 데이터는 정밀도와 재현율에서 서로 상반된 결과를 보인다. 하이라이트와 그림자를 관심영역으로 지정한 데이터의 정밀도는 약 0.90이므로 하이라이트로만 관심영역을 지정한 데이터보다 약 0.08 높다. 반대로 하이라이트만 관심영역으로 지정한 데이터의 재현율은 약 0.92로 하이라이트와 그림자를 모두 관심영역으로 지정한 데이터보다 약 0.17 높다.

Table 3에서 살펴본 정밀도와 재현율의 차이가 탐지결과에서 어떻게 나타나는지 Fig. 10과 같이 확인할 수 있다. Fig. 10(a)에서 하이라이트와 그림자를 관

심영역으로 지정한 왼쪽 그림은 오탐지를 나타내는 False Positive가 있었지만 하이라이트만 관심영역으로 지정한 오른쪽 그림은 False Positive가 없었다. Fig. 10(b)에서 왼쪽 그림에는 False Positive뿐만 아니라 탐지하지 못한 실패적을 나타내는 False Negative도 있었지만, 오른쪽 그림에는 실패적만 정확하게 탐지하였다. Fig. 10(c)에서는 오른쪽 그림이 실패적을 모두 정확하게 찾아냈지만 오탐지는 왼쪽 그림보다 많았다. 종합하면, 컨볼루션 신경망이 하이라이트만 관심영역으로 지정하여 학습한 데이터에 대해서 하이라이트와 그림자까지 관심영역으로 지정하여 학습한 데이터보다 허위표적이 많이 포함되어 정밀도가 약 8% 떨어지지만 전체 실패적 중 약 92% 탐지하였다.

5.4 수중물체 탐색 목적을 중심으로 한 고찰

수중물체를 탐색하는 것은 중요한 목적을 갖는다. 해양에서 주된 탐지 수단인 음파는 광활한 해저면과 수중환경 변화라는 조건에서 불규칙적인 신호로 수신되므로 이를 해석하는데 어려움이 크다. 그러므로 수중의 물체를 찾는 데 많은 시간과 노력이 소모되며, 찾는 대상 물체도 군사적, 비군사적인 성격에 따라 위험물이거나 고가일 경우가 많다. 따라서 찾고자하는 물체를 놓치거나 늦어질수록 그만큼 인적, 물적, 비용적인 손실이 커진다.

앞서 소나 영상의 관심영역을 변경한 후 컨볼루션 신경망으로 테스트한 결과를 비교했을 때, 허위표적과 실패적을 탐지하는 것은 어느 정도의 트레이드-오프 관계가 있다는 것을 알 수 있었다. mAP가 동일하다고 가정할 때, 정밀도가 높으면 탐지결과에서

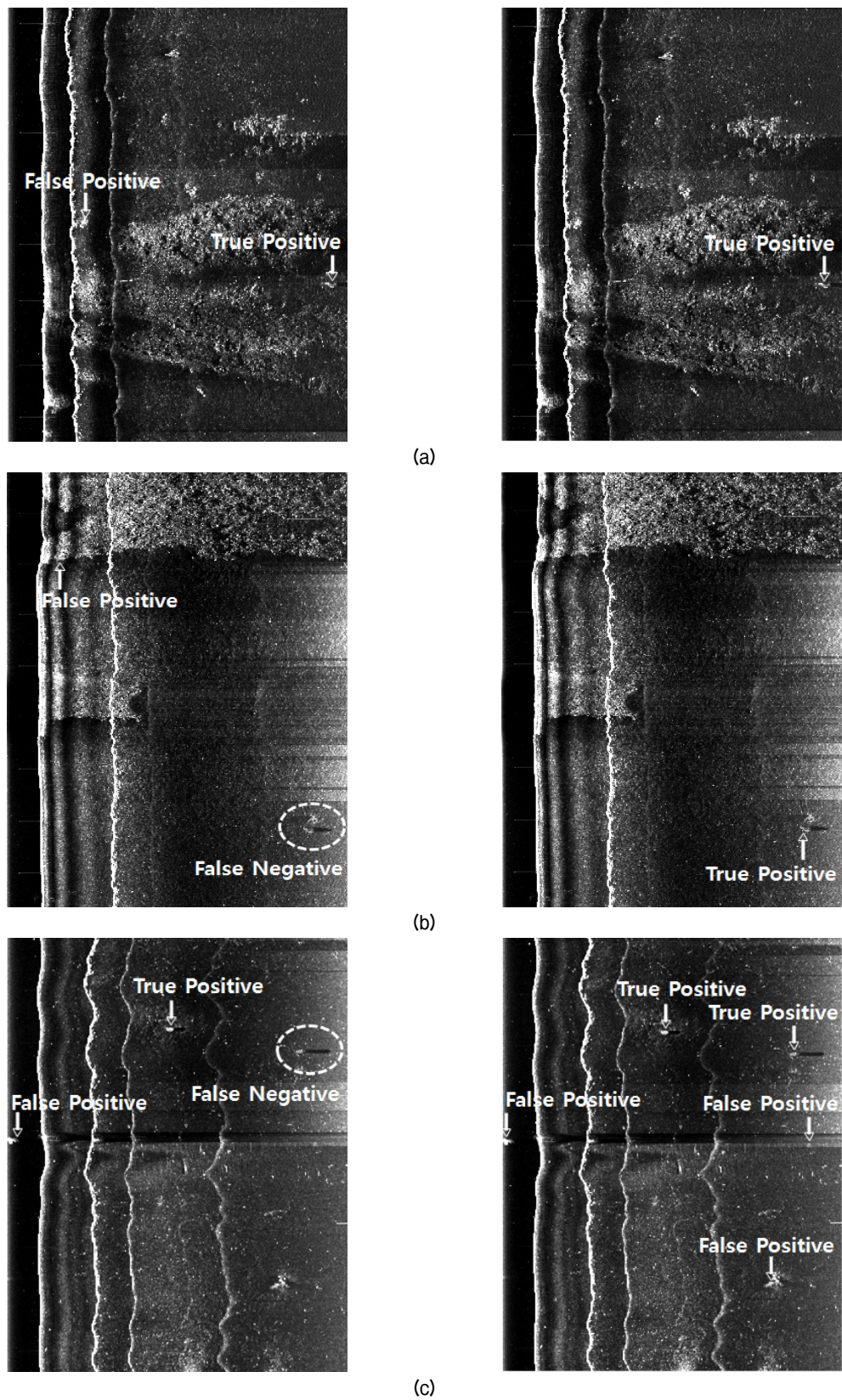


Fig. 10. Detection result (left: highlight+shadow, right: highlight only), (a) example 1, (b) example 2, (c) example 3.

허위표적을 제외하는 대신에 실패적도 포함되지 않을 가능성이 있지만, 재현율이 높으면 허위표적이 포함될 수 있는 대신에 실패적도를 놓칠 가능성이 줄어든다. 수중물체를 탐색하는 목적에서는 허위표적이 많이 발생하는 것 보다는 실패적도를 정확히 찾는 것이 중요하다. 그러므로 허위표적을 감소시킬 것인가 실패적 탐지를 늘릴 것인가를 결정할 때 재현율을 고려해야 한다.

한편, 하이라이트와 그림자를 포함하여 지정한 관심영역은 하이라이트만 지정한 관심영역 보다 mAP가 낮았지만, 정밀도-재현율 곡선의 정밀도는 더 높았다. 이 의미는 그림자도 물체의 특성이 투영된 것으로서 이를 통해 물체의 높이 및 형태를 파악할 수 있게 하므로 물체를 분류하는데 하나의 특징요소로 작용한 것으로 보인다. 다만 수중 예인체와 물체와의 거리 및 앵글에 따라서 그림자의 길이가 변하는 데다가 데이터의 수량도 제한되었기 때문에 완전하게 학습이 되지 않아서 효과가 적은 것으로 사료된다. 그림자는 변하는 요소가 컨볼루션 신경망을 통한 학습에 제한되는 요소였지만, 길이 변화에 대한 함수를 정립한다면 보다 효과적인 학습이 가능할 것이다. 즉, 물체의 하이라이트 특성과 그림자가 제대로 학습된다면 하이라이트만 관심영역으로 지정하여 학습한 것보다 허위표적을 감소시키면서 실패적도를 더 찾을 수 있는 향상된 탐지성능을 보여줄 것이라 예상된다.

VI. 결 론

본 논문에서는 컨볼루션 신경망을 이용하여 수중 음향으로 나타난 영상을 인식하고 학습된 물체를 탐지하는 것을 연구하였다. 소나 영상과 광학 영상을 각각 컨볼루션 신경망에 적용한 결과 평균 정밀도로 진단한 성능 mAP가 각각 0.74, 0.92로 대체로 양호하였다. 하지만 소나 영상에 대한 성능 mAP는 광학 영상보다 0.18 낮았다. 이 원인으로 그림자의 길이 변화를 컨볼루션 신경망의 학습 저해 요인으로 추정하였으며, 소나 영상의 관심영역을 하이라이트와 그림자를 동시에 지정한 것과 하이라이트만 지정한 것을 비교하였다. 그 결과는 하이라이트만 지정한 데이터

세트의 성능 mAP가 0.83으로 높게 나왔다.

한편, 하이라이트와 그림자를 동시에 학습한 데이터에 대한 정밀도는 높았고, 하이라이트만 학습한 데이터는 재현율이 높았다. 수중물체를 탐지하는 관점에서 탐지성능을 진단하기 위해서 종합적인 지표인 평균 정밀도를 보는 것도 중요하지만 재현율의 의미를 살펴볼 필요성이 있다. 수중물체를 탐색하는 과정에서 목표물체를 놓치게 될 때의 손실을 고려한다면 재현율이 높은 알고리즘을 선택하는 것이 더 효과적이다.

머신러닝의 특성상 다량의 데이터가 필요하나 논문은 제한된 데이터로 결과를 확인한 것에 한계가 있다. 따라서 데이터가 부족할 때에는 하이라이트만으로 학습하여 탐지하고 하이라이트와 그림자를 동시에 학습한 것으로 오탐지를 줄이는 하이브리드 방법을 사용하는 것도 고려될 수 있다. 또한 소나 영상의 특징 요소인 그림자가 수중 예인체와 물체와의 거리 및 각도에 따라 변하는 관계에 대해 연구가 필요하다. 이 부분을 적용하면 소나 영상을 컨볼루션 신경망에 적용할 때 허위표적을 감소시키면서 실패적도를 더 찾을 수 있는 향상된 탐지성능을 기대할 수 있다.

감사의 글

본 논문은 2018년 해양수산부 재원으로 해양수산과학기술진흥원의 지원[위험유해물질(HNS)사고 관리기술 개발, 과제번호: 20150340], 국방과학연구소의 지원(관리번호: UD160006DD)을 받아 수행된 연구임.

References

1. S. Reed, Y. Petillot, and J. Bell, "Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information," *IEE. Proc.-Radar Sonar Navig.*, **151**, 48-56 (2004).
2. S. R. Kim, "The reason why to use acoustic waves on the sea-bottom survey" (in Korean), *J. Korean Society of Marine Engineering*, **32**, 481-489 (2008).
3. J. P. Fish and H. A. Carr, *Sound Underwater Images: A Guide to the Generation and Interpretation of Side*

Scan Sonar Data (Lower Cape Publishing Co., Orleans, 1990), pp. 11-47.

4. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Image net large scale visual recognition challenge," *Int. J. Computer Vision*, **115**, 211-252 (2015).
5. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, 1097-1105 (2012).
6. D. Ciresan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," *CVPR*, 3642-3649 (2012).
7. Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," *Advances in neural information processing systems*, 1988-1996 (2014).
8. L. Wan, M. Zeiler, S. Zhang, Y. L. Cun, and R. Fergus, "Regularization of neural networks using dropconnect," *Proc. ICML-13*, 1058-1066 (2013).
9. K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *Proc. IEEE international conference on computer vision*, 1026-1034 (2015).
10. P. Blondel, *The Handbook of Side Scan Sonar* (Springer Science & Business Media, Chichester, UK, 2010), pp. 23-34, 63-65.
11. S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE, trans. on pattern analysis and machine intelligence*, **39**, 1137-1149 (2017).
12. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. machine learning research*, **15**, 1929-1958 (2014).
13. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, USA, 2016), pp. 199-216.
14. E. Dura, Y. Zhang, X. Liao, G. J. Dobeck, and L. Carin, "Active learning for detection of mine-like objects in side-scan sonar imagery," *IEEE, J. Oceanic Engineering*, **30**, 360-371 (2005).
15. C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," *ECIR.*, **5**, 345-359 (2005).

저자 약력

▶ 김 정 문 (Jungmoon Kim)



2006년 2월 : 해군사관학교(학사)
2018년 2월 : 한양대학교 해양융합학과 석사
2018년 3월 ~ 현재 : 해군 1함대 사령부

▶ 최 지 웅 (Jee Woong Choi)



1996년 2월 : 한양대학교(학사)
1998년 2월 : 한양대학교(석사)
2002년 8월 : 한양대학교(박사)
2002년 11월 ~ 2007년 8월 : Applied Physics Lab., Univ. of Washington, USA, Research Associate
2007년 9월 ~ 현재 : 한양대학교 해양융합 공학과 교수

▶ 권 혁 중 (Hyuckjong Kwon)



2013년 2월 : 한양대학교 해양환경과학과 학사
2013년 3월 ~ 현재 : 한양대학교 석·박사 통합 과정

▶ 오 래 근 (Raegeun Oh)



2013년 2월 : 한양대학교 해양환경과학과 학사
2014년 9월 ~ 현재 : 한양대학교 석·박사 통합 과정

▶ 손 수 욱 (Su-Uk Son)



2008년 2월 : 한양대학교 지구해양과학과 학사
2010년 2월 : 한양대학교 해양환경과학과 석사
2015년 8월 : 한양대학교 해양융합과학과 박사
2015년 9월 ~ 2016년 9월 : 한양대학교 박사 후 연구원
2016년 10월 ~ 현재 : 국방과학연구소