

컨볼루션 신경망의 앙상블 모델을 활용한 마스트 영상 기반 잠수함 탐지율 향상에 관한 연구

정미애¹⁾ · 마정목^{*1)}

¹⁾ 국방대학교 국방과학학과

A Study on the Improvement of Submarine Detection Based on Mast Images Using An Ensemble Model of Convolutional Neural Networks

Miae Jeong¹⁾ · Jungmok Ma^{*1)}

¹⁾ Department of Defense Science, National Defense University, Korea

(Received 17 February 2020 / Revised 23 March 2020 / Accepted 30 March 2020)

ABSTRACT

Due to the increasing threats of submarines from North Korea and other countries, ROK Navy should improve the detection capability of submarines. There are two ways to detect submarines : acoustic detection and non-acoustic detection. Since the acoustic-detection way has limitations in spite of its usefulness, it should have the complementary way. The non-acoustic detection is the way to detect submarines which are operating mast sets such as periscopes and snorkels by non-acoustic sensors. So, this paper proposes a new submarine non-acoustic detection model using an ensemble of Convolutional Neural Network models in order to automate the non-acoustic detection. The proposed model is trained to classify targets as 4 classes which are submarines, flag buoys, lighted buoys, small boats. Based on the numerical study with 10,287 images, we confirm the proposed model can achieve 91.5 % test accuracy for the non-acoustic detection of submarines.

Key Words : Submarine Mast(잠수함 마스트), CNN(컨볼루션 신경망), Transfer Learning(전이학습), Ensemble(앙상블)

1. 서론

잠수함의 중요성과 가치는 무기체계로서 본격적인

역할을 시작한 1차 세계대전 이래 다양한 해상전에서 입증되어 왔다. 2차 세계대전 당시 영국 수상으로 재임하였던 윈스턴 처칠은 “전쟁기간 중 정말 나를 두렵게 했던 단 하나의 무기”로 유보트를 언급하였으며, 1982년 포클랜드 해전에서 아르헨티나의 단 1척의 잠수함은 영국의 기동함대가 34일 동안 방어적인 대잠

* Corresponding author, E-mail: jxm1023@gmail.com
Copyright © The Korea Institute of Military Science and Technology

전을 수행하도록 강요하였다^[1]. 이처럼 잠수함은 은밀성, 공세성, 억제 및 방어수단으로서의 유용성을 바탕으로 현대해상전의 주도권을 좌우할 수 있는 전략적인 무기체계이다.

한편 북한은 1960년대부터 위스키급 잠수함을 보유하기 시작하여 상어급, 연어급 등 70~80여척의 잠수함(정)을 보유 중이며, 이를 활용하여 우리 해역을 침범해 적대 행위를 할 가능성이 항상 도사리고 있다. 실제로 북한은 1996년 상어급 잠수함 강릉침투, 1998년 유고급 잠수정 속초침투, 2010년 연어급 잠수정 천안함 폭침 등을 자행한 바 있다. 뿐만 아니라, 경쟁적으로 잠수함 전력 증강에 박차를 가하고 있는 주변국의 잠수함 또한 유사시 제해권 확보와 해상교통로 교란을 위해 우리 해역에서 활동할 가능성이 존재한다. 대한민국 해군은 이같은 북한 및 주변국 잠수함의 위협에 대비하기 위하여 잠수함을 조기에 탐지할 수 있는 능력 확충이 절실히 필요하다.

잠수함의 탐지를 위해서는 크게 음향, 비음향 탐지방법이 사용되는데, 음향 탐지방법은 수중에서는 잘 전달되지 않는 전자파 대신 음파를 이용하여 잠수함을 탐지하는 방법이다. 그러나 음파도 수중에서의 굴절, 산란으로 인해 잘 전달되지 않을 수 있고, 더불어 한반도 주변해역의 복잡한 해양환경은 음파를 통한 잠수함 탐지를 더욱 어렵게 만들고 있다. 특히, 독특한 굴절률을 가진 음향메타물질과 같이 잠수함의 은밀성을 더욱 보장하는 기술들이 접목된다면 잠수함의 음향 탐지는 더욱 힘들어질 것이 분명하다.

비음향 탐지방법은 수상함이나 항공기에서 운용하는 탐지레이더, 적외선 탐지장비, 당직자의 쌍안경 및 육안으로 잠수함을 탐지하는 방법으로 잠망경심도(Periscope depth)에서 항해하면서 마스트세트(Mast set)를 운용하는 잠수함을 탐지하기 위해 사용되며, 음향 탐지방법을 보완하기 위해 필수적으로 수행되는 방법이다. 특히, 축진지 충전이나 해상관측, 위치보정, 공격직전 목표물 확인 등을 위해 수면 가까이로 부상하는 것이 꼭 필요한 디젤잠수함^[2] 탐지하는데 유용하며, 부상하는 시기가 잠수함이 가장 취약한 시기이므로 수상함 및 항공기는 잠수함을 탐지하기 위해 이를 효과적으로 활용하여야 한다.

그러나 비음향 탐지방법 또한 제약사항들이 있는데, 마스트세트를 운용하는 잠수함을 탐지할 수 있는 시간이 짧고, 잠수함인지 판단하는데 수상함 근무자의 직관과 경험에 크게 의존한다는 것이다. 특히, 잠수함

의 비음향 탐지수단으로 유용하게 여겨지는 방법 중 하나인 쌍안경이나 육안으로 전방을 감시하여 탐지하는 방법은 대부분 수상함 근무 경력이 짧은 해상병이 그 임무를 수행하고 있어 효과적인 탐지를 더욱 어렵게 하며, 그마저도 저출산 시대로 인한 국방개혁으로 점차 병력을 감축해 나갈 것을 고려하면 임무를 수행해야 할 인원이 갈수록 부족해질 전망이다.

이에 본 연구는 4차 산업혁명의 중심기술인 딥러닝 중에서도 이미지 인식에 최적화된 CNN(Convolutional Neural Network) 모델을 비음향 탐지에 활용하여 잠망경을 비롯한 스노클(Snorkel), 통신, ES(Electronic Support) 등의 마스트를 운용중인 잠수함을 탐지하는 방안을 최초로 제시하였다. 또한 수면 위로 올라오는 마스트는 소형표적이므로 분류 정확도가 낮을 수 있는데, 이를 보완하기 위해 여러 CNN 모델을 결합한 앙상블(Ensemble) 모델을 활용하여 성능을 향상시키는 방안을 모색하였다. 2장에서는 잠수함 탐지방법을 이론적으로 고찰하였으며, CNN 모델을 활용하여 잠수함을 탐지하고자 한 선행된 연구들을 소개하였다. 3장에서는 본 연구에서 제시한 잠수함 분류 모델을 소개하였고, 4장에서는 데이터를 수집한 후 3장에서 소개한 잠수함 분류 모델에 따라 여러 CNN 모델의 적합한 하이퍼파라미터(Hyperparameter)를 탐색하고, 다양한 앙상블 모델을 활용하여 성능의 향상 여부를 확인하였다. 5장에서는 연구결론 및 제한사항을 제시하였다.

2. 기존 연구 동향

2.1 잠수함 탐지에 대한 이론적 고찰

2.1.1 잠수함 음향 탐지의 한계

음파는 수중에서 전자파보다 훨씬 빠르고 멀리 전달되기 때문에 이를 이용한 대잠감시가 우수한 탐지 수단임은 분명하지만, 잠수함을 탐지하기 위한 유일한 수단으로서는 많은 제약이 있다^[3].

그 이유를 살펴보면, 첫째, 음파는 수중에서 확산과 감쇠로 인한 전달손실을 가지므로 전달되면서 음의 강도가 점차 감소되어 탐지거리가 짧아진다. 둘째, 음파는 Fig. 1과 같은 음속의 수직구조에 의해 혼합층(Mixed layer), 수온약층(Thermocline), 심해층(Deep water)을 형성하는데, 각 층별로 음파전달 양상이 달라 음파가 전달되지 않는 음영구역을 발생시킬 수 있다. 셋째, 물리적 화학적 특성이 서로 다른 해수가 만나

생성되는 수온전선 및 와동류로 인해 음파전달 특성이 급격하게 변화하여 탐지거리가 감소될 수 있다. 넷째, 파도소리, 어류에 의한 소리, 유체소음 등의 각종 배경소음으로 인해 표적소음의 탐지가 제한될 수 있다. 이처럼 음파는 수중에서 물체를 탐지하는데 우수한 수단이나 언급한 제약사항들을 가지며, 이를 보완할 수 있는 수단이 필요하다.

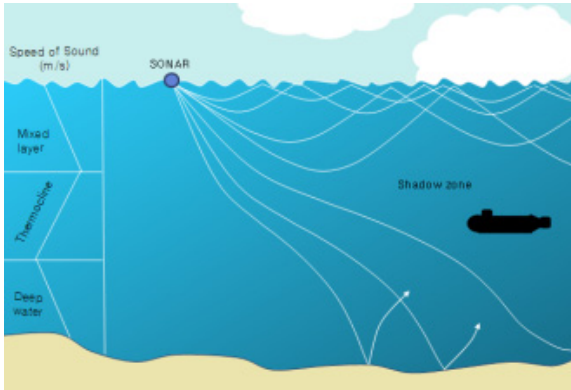


Fig. 1. Sonic structure and acoustic propagation

2.1.2 잠수함 비음향 탐지의 가능성

디젤잠수함은 축전지 충전을 위해서 반드시 주기적으로 수면 가까이로 부상하여 스노클을 하여야 한다. 잠수함은 수중항해시 축전지에 저장된 전력으로 추진 전동기를 구동하고, 이 전동기에 연결된 프로펠러를 회전시킴으로써 추진력을 얻게 된다. 그러나 축전지는 항해시 저속으로 기동하더라도 약 2-4일이 경과하면 모두 방전되며, 이를 충전하기 위해서는 디젤엔진을 구동시켜야 하는데, 디젤엔진 구동을 위해서 반드시 공기를 공급해 주어야 하므로 공기를 흡입하기 위해 수면 가까이로 부상해야 하는 것이다. 이때 잠수함은 수면 위로 완전히 부상하지 않고, 수면 밑 일정 깊이에서 스노클 마스트만을 수면 위로 내밀어 공기를 흡입한다. 이러한 특성은 북한 및 주변국 디젤잠수함도 동일한데, 과거 북한의 침투사례를 분석하였을 때도 북한 잠수함은 NLL 북방에서부터 잠항하여 우리해역으로 진입후 축전지 충전을 00% 도달 시 야간을 틈타 0시간 이상 스노클을 하였음이 밝혀진 바 있다^[6].

한편, 잠수함은 스노클 외에도 수면 가까이로 부상하여 마스트세트를 운용하는 경우가 있는데, 그 대표적인 경우가 잠망경(Periscope)을 운용할 때이다. 잠망경은 공격잠망경(Attack periscope)과 탐색잠망경(Search

periscope)으로 구분되는데, 공격잠망경은 공격 직전 목표물의 방위 등을 최종적으로 확인하기 위해 사용하며, 탐색잠망경은 해상 관측용 또는 GPS(Global Positioning System)를 통한 위치보정을 위해 사용한다. 잠수함이 수상함에 대한 공격을 한다면 수중에서 수동소나(Passive sonar)를 이용하여 목표물의 방위·속도를 파악하게 되는데, 북한잠수함에 설치된 수동소나는 그 성능이 저급하여 수중에서의 공격 문제해결이 제한되므로 잠망경으로 최종 목표물을 확인할 확률이 매우 높을 것으로 판단된다^[6]. 이를 뒷받침해 주듯이 북한의 과거 침투 사례에서도 일몰 직후에 우리해역에서 잠망경을 운용하여 공격 목표나 해안을 정찰하였음이 드러난 바 있다.

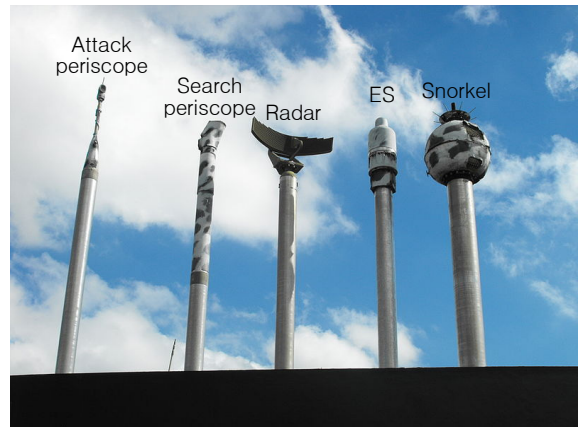


Fig. 2. Submarine mast^[6]

이처럼 잠수함은 Fig. 2와 같이 스노클 마스트, 잠망경 등으로 구성된 마스트세트를 운용한다. 이때가 잠수함에게는 가장 취약한 시기이며, 1997년 서해 소흑산도 근해 잠망경으로 보이는 물체를 발견했다는 어민 신고로 중국의 밍급 잠수함을 탐지하여 부상시킨 사례도 있다^[7]. 따라서 수상함 및 항공기에서는 레이더 뿐만 아니라 쌍안경 및 육안을 이용하여 이를 적극적으로 탐지할 필요가 있다. 하지만, 서론에서 언급한 것과 같이 판단할 수 있는 시간이 짧고 수상함 근무자의 경험과 판단의 개입이 커 효과적으로 탐지하는 것이 어렵다. 따라서 수상함에서 운용하고 있는 TV카메라로 획득하는 영상에 4차 산업혁명의 대표적인 기술인 딥러닝 기술을 적용한다면 잠수함 마스트의 수면위 노출을 효과적으로 포착할 수 있을 뿐만 아니라 수상함 근무자의 경험적 요소나 외부 환경요

소의 영향을 받지 않고 효율적으로 잠수함을 탐지할 수 있을 것으로 판단된다.

2.2 CNN 모델과 잠수함 탐지

본 연구에서 잠수함의 마스트를 탐지하기 위해 적용한 모델은 딥러닝 알고리즘 중 이미지 분야를 다루기에 최적화된 CNN(Convolutional Neural Network) 모델이다. CNN은 크게 컨볼루션층(Convolution layer)과 풀링층(Pooling layer), 완전연결층(Fully-connected layer)으로 구성되어 있다. 컨볼루션층은 컨볼루션 연산을 통해 이미지의 특징을 추출해 내는 층인데, 사용하는 필터(Filter)에 따라 원본 이미지에서 다양한 특징을 추출해 낼 수 있다. 풀링층(Pooling layer)은 최대풀링(Max pooling), 평균풀링(Average pooling), 최소풀링(Min pooling)을 통해 이미지 차원을 축소함으로써 필요한 연산량을 감소시키고, 이미지의 강한 특징만을 선별한다^[8]. 완전연결층(Fully-connected layer)은 컨볼루션층의 3차원 출력값을 1차원 벡터로 만들어 출력층이 이미지의 최종 분류를 할 수 있도록 판단하는 값을 제공한다.

CNN 모델과 같은 딥러닝 기술을 활용하여 잠수함을 탐지하려는 연구는 주로 음향 탐지분야에서 이루어졌다. Kim et al.^[9]은 능동소나에서 잠수함 또는 기뢰를 바위와 같은 비표적과 구분하기 위해 CNN 모델 중 하나인 Alexnet 모델을 일부 변경하여 학습하면서, 데이터의 양에 따른 모델 성능 변화를 확인하였다. Kim et al.^[10]은 능동소나에 의한 표적 이미지를 데이터 확장하고, 생성된 이미지가 표적을 포함하는 비율에 따라 CNN 모델을 활용하여 표적/비표적을 분류하는 연구를 진행하였다. 그 외에 잠수함을 탐지하는 연구로 Kim et al.^[11]은 Faster R-CNN을 이용하여 해양에서의 선박을 검출 후 항공모함, 잠수함, 컨테이너선 등 7가지로 분류하였다. 그러나 이 연구의 목적은 해상에서 충돌을 피하기 위해 항해하는 다양한 물표들을 검출하고 분류하는 것으로 잠수함 또한 완전히 부상항해하는 잠수함을 검출하는 것이다. 따라서 은밀성을 유지하고자 하는 잠수함을 탐지하는 본 연구와는 목적과 추후 모델 활용 방법이 다르다.

CNN 모델의 앙상블을 실험한 연구로 Park et al.^[12]은 산림곤충 데이터세트에 대해 CNN 모델의 앙상블 결합 방법에 따른 성능 분석에 관한 연구를 진행하였으며, B. S. Kim^[13]은 재활용쓰레기 데이터세트에 대해 CNN 모델을 활용하여 미세조정(Fine tuning), 앙상블 단계를

거치면서 성능이 향상됨을 보이고, 이 모델이 기존에 사용되었던 모델인 SVM(Support Vector Machine)에 비해 성능이 우수함을 확인하였다. 이에 본 연구는 잠수함 비음향 탐지분야에 최초로 CNN 모델을 적용시켜 잠수함의 마스트와 항해 중 유사하게 보일 수 있는 표적들을 분류하고, 이 모델들을 결합한 앙상블 모델을 활용함으로써 잠수함 비음향 탐지용 향상방안을 제시하고자 한다.

3. 제안한 잠수함 분류 모델

본 연구는 진술적인 상황에서 잠망경을 비롯한 마스트를 올리고 항해하는 잠수함을 탐지하는 것을 주목표로 한다. 추가적으로는 잠망경을 올리는 것 외에 부상한 잠수함 또한 식별 가능하여야 할 것이며, 수상함에서 잠수함 마스트로 오인 가능한 깃대부이(Flag buoy), 등부표(Lighted buoy), 소형 어선·보트(Small boat)와 구별 가능할 것을 목표로 진행하였다.

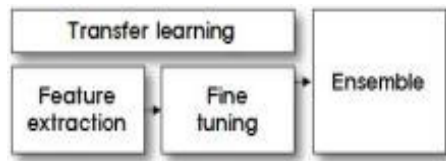


Fig. 3. Proposed submarine classification model

분류 모델 구축단계는 데이터 수집, 이미지 분류 모델 학습 및 평가단계로 구성된다. 수집된 데이터들은 Fig. 3과 같이 3단계를 거쳐 학습된다. 1, 2단계는 사전 학습된(Pre-trained) 모델을 사용하여 전이학습(Transfer learning)을 실시하였는데, 사전 학습된 모델이란 대규모 이미지 분류 문제를 위해 대량의 데이터세트에서 미리 훈련되어 저장된 네트워크로, 일반적으로 ILSVRC (Imagenet Large Scale Visual Recognition Competition)에서 100만장이 넘는 이미지 데이터세트인 이미지넷 (Imagenet)을 이용하여 학습한 모델이 사용된다. 이 모델에서 학습된 파라미터(Parameter)와 네트워크의 구조를 이용하면 상대적으로 작은 데이터세트만으로도 이미지 분류 모델을 학습시킬 수 있다. 여기서 파라미터란 가중치(Weight)와 편향(Bias)을 의미하는데, 딥러닝에서는 학습하고자 하는 데이터를 가중치와 편향을 포함한 가설함수로 나타낸 후 가설함수가 실제데이터

와 얼마나 다른지를 나타내는 손실함수를 최소화하는 방향으로 가중치와 편향을 업데이트하게 된다. 즉, 파라미터란 기계가 학습하면서 결정하는 값으로 하이퍼 파라미터(Hyperparameter)와는 구분해서 이해되어야 하는데, 하이퍼파라미터란 학습률(Learning rate), 유닛(Unit)의 개수, 활성화 함수, 학습세대(Epoch), 배치(Batch) 등 기계가 자동으로 결정하는 값이 아니라 연구자가 설정해주어야 하는 값이다.

본 연구의 1, 2단계에서 사용한 전이학습은 경우에 따라 미세조정(Fine tuning)과 같은 의미로 사용되기도 하지만, 본 연구에서는 특성추출(Feature extraction)과 미세조정 2가지로 세분화하여 1단계에서는 특성추출을 사용하고, 2단계에서 미세조정을 사용하였다. 특성추출은 컨볼루션층의 파라미터는 고정시키고 완전연결층의 파라미터를 재학습시키는 것을 의미하고, 미세조정은 완전연결층과 더불어 컨볼루션층의 마지막 몇 개층 파라미터를 재학습시키는 것을 의미한다. 여러 개의 컨볼루션층 중 컨볼루션을 처음 수행하는 층은 이미지의 가로 모서리, 세로 모서리 등의 일반적인 특성을 추출하고 마지막 층으로 갈수록 이미지에 특화된 특성을 추출한다^[14]. 따라서 미세조정을 함으로써 잠수함 마스트에 특화된 특성을 추출하여 완전연결층만 재학습하는 특성추출 대비 모델 성능이 향상되는지 확인하였다. 3단계는 앙상블 모델을 사용하였다. 단일 CNN 모델이 이미지 내에서 모든 특성들을 추출해 내지 못하는 경우 다수의 모델을 결합하는 앙상블은 뛰어난 하나의 모델보다 우수한 성능을 발휘할 수 있다^[12]. 본 연구는 마스트를 올리고 항해하는 잠수함의 탐지율을 극대화하기 위하여 2단계에서 사용한 우수한 모델들을 결합한 앙상블 모델을 사용함으로써 이미지 분류 모델의 성능이 향상되는지를 확인하였다.

본 연구는 GPU 활용을 위해 많이 이용되는 Google colabory를 사용하였으며, 그에 따라 Ubuntu 18.04.3 LTS, GPU Persistence-M CUDA Version 10.1, Python 3.6.9, Tensorflow 1.15.0, Keras 2.2.5 실험환경에서 수행하였다.

4. 잠수함 분류 모델 구축 및 실험

제시한 모델에 따른 데이터 수집 및 실험 결과는 다음과 같다.

4.1 데이터 수집

데이터는 연구 목표에 부합하기 위하여 Fig. 4와 같이 마스트를 올리고 항해하는 잠수함, 부상한 잠수함, 깃대부이, 등부표, 소형 어선·보트의 데이터를 구글 및 네이버 이미지를 검색하여 총 2,395장을 수집하였다. 데이터는 수상함의 TV카메라가 배울확대기능을 가지는 것을 고려하여 표적 크기가 큰 것부터 작은 것까지 다양하게 수집하였다. 부상한 잠수함의 경우 표적이 이미지의 절반 정도 비율을 차지하였으나, 최대한 원거리에서부터 표적을 분류하는 기능을 구현하기 위해 중점적으로 포함시킨 Fig. 5와 같은 작은 표적의 데이터는 표적이 이미지의 수천분의 일 크기 수준이었다.

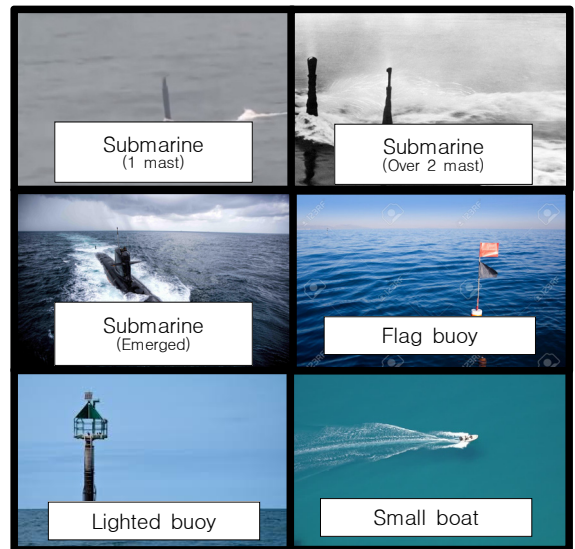


Fig. 4. Example images of data set(Google image)

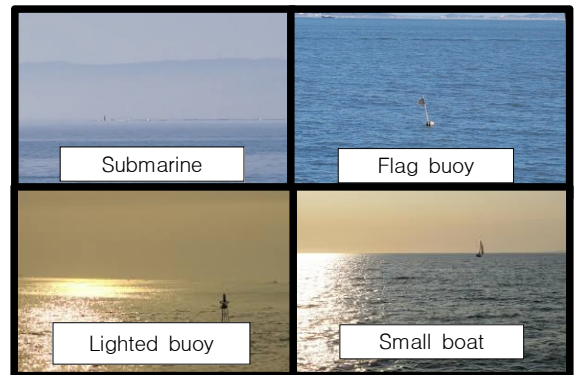


Fig. 5. Example images of small target(Google image)

Table 1. Data set information

| Class | Training (Data augmentation) | Validation | Test | Total (Data augmentation) |
|--------------|------------------------------|------------|------|---------------------------|
| Submarine | 783 (3,915) | 87 | 50 | 920 (4,052) |
| Flag buoy | 178 (890) | 22 | 50 | 250 (962) |
| Lighted buoy | 320 (1,600) | 36 | 50 | 406 (1,686) |
| Small boat | 692 (3,460) | 77 | 50 | 819 (3,587) |
| Total | 1,973 (9,865) | 222 | 200 | 2,395 (10,287) |

수집한 데이터는 검증을 위해 일반적으로 7:3, 8:2, 9:1의 비율로 훈련데이터(Training data), 검증데이터(Validation data)로 나눌 수 있는데 본 연구에서는 훈련데이터의 수량 확보를 위해 9:1의 비율로 나누었고, 평가데이터(Test data)는 테스트 정확도의 정확성을 위하여 Table 1과 같이 4개 클래스가 동일한 50개의 데이터를 가지도록 별도로 구성하였다. 함수함의 테스트 이미지 50장은 마스크 1개 22장, 마스크 2개 이상 12장, 함교탑까지 포함한 사진 13장으로 구성하였으며, 표적의 크기 또한 다양하게 구성하였다.

딥러닝에서 훈련데이터에 대해서는 예측을 잘 하지만 새로운 데이터에 대한 예측 정확도가 떨어지는 과대적합(Overfitting)이 발생할 수 있는데, 이를 방지하기 위한 유용한 방법 중 한 가지는 하나의 데이터를 회전, 이동 등의 방법으로 데이터 확장(Data augmentation)하는 것이다. 본 연구에서는 훈련데이터 1,973장에 대해 회전(Rotation), 좌우이동(Width_shift), 상하이동(Height_shift), 좌우반전(Horizontal_flip), 상하반전(Vertical_flip)의 데이터 확장을 실시하였다.

4.2 이미지 분류 모델 학습 및 평가

4.2.1 특성추출(Feature extraction)

특성추출단계에서 사용한 모델은 VGG16, VGG19,

ResNet50, ResNet101이다. VGG16^[15]은 2014년 ILSVRC에서 준우승한 모델로 Fig. 6과 같이 5개의 풀링층을 포함한 13개의 컨볼루션층과 3개의 완전연결층으로 구성되어 있으며, 3×3의 균일한 크기를 가진 필터와 패딩(Padding)을 이용해 이미지의 크기가 유지될 수 있도록 컨볼루션을 수행한다. 이전 CNN 모델에 비해 작은 필터를 깊은 네트워크에서 반복적으로 사용하여 효과는 동일하게 하되 파라미터 수를 적게 사용하였다는 특징이 있다. VGG19는 VGG16의 컨볼루션층을 16개로 구성한 모델이다.

ResNet^[16]은 2015년 ILSVRC에서 우승한 모델로 기존의 CNN 분야에 있었던 학습층이 과다하게 깊어지면 훈련 정확도가 떨어지는 문제를 Residual net을 사용하여 해결했다. Fig. 7은 Plain net과 Residual net의 차이를 보여주는데 Plain net에서는 출력을 $H(x)$ 로만 표현하고 있지만, Residual net에서는 입력인 x 를 다시 출력층에 제공해 줌으로써 출력을 $H(x) = F(x) + x$ 로 표현하고 있다. 즉, 입력 x 를 제공받기 때문에 입·출력 간의 오차 $F(x)$ 만 추가로 학습하면 출력인 $H(x)$ 를 얻게 되는 것이다. 이로 인해 $H(x)$ 전체를 학습하는 이전 CNN 모델들에 비해 학습층 깊이를 깊게 할 수 있었고, 그 층의 수에 따라 본 연구에서 사용한 ResNet50, ResNet101 외에도 ResNet18, ResNet34, ResNet152 등이 있다.

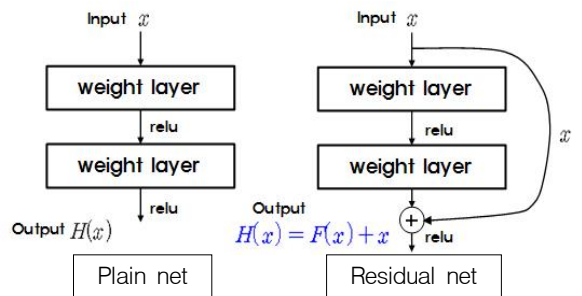


Fig. 7. Comparison of plain net and residual net



Fig. 6. Structure of VGG16 model

Table 2. Best results for each model in feature extraction

| Hyperparameter | VGG16 | VGG19 | ResNet101 | ResNet50 |
|-----------------------------------|--------|--------|-----------|----------|
| Fully-connected layer number | 1 | 1 | 1 | 1 |
| Drop out ratio(%) | 50 | 50 | 50 | 50 |
| Drop out layer number | 1 | 2 | 2 | 2 |
| Batch size | 32 | 32 | 32 | 32 |
| Fully-connected layer unit number | 1024 | 1024 | 1024 | 1024 |
| Activation function | relu | relu | relu | relu |
| Regularization | - | L1, L2 | L1, L2 | L1, L2 |
| Learning rate | 1e-5 | 1e-4 | 1e-5 | 1e-5 |
| Maximum epoch | 100 | 200 | 200 | 200 |
| Optimum epoch | 96 | 200 | 194 | 199 |
| Train accuracy(%) | 86.70 | 77.25 | 81.34 | 86.80 |
| Train loss | 0.7798 | 4.1075 | 2.3053 | 1.4102 |
| Validation accuracy(%) | 88.34 | 88.75 | 84.30 | 80.27 |
| Validation loss | 0.7019 | 3.0145 | 2.3196 | 1.5897 |
| Test accuracy(%) | 86.5 | 84.5 | 79.5 | 79.0 |

각 모델의 하이퍼파라미터는 Table 2와 같이 배치 32, 학습세대 100, 완전연결층 1개층, 드롭아웃(Drop out) 비율 50 %로 학습하였으며, 학습률을 10^{-4} , 10^{-5} , 10^{-6} 등으로 변화시키고 최종 완전연결층의 유닛을 512, 1024, 2048 등으로 지정하면서 실험하였다. 실험 결과 우수한 테스트 정확도를 얻을 수 있었던 하이퍼 파라미터 값과 그때 학습 결과로 생성된 파라미터 값은 2단계 모델에서 다시 사용하였다. 이때 검증손실(Validation loss)이 훈련손실(Train loss)에 비해 더 이상 떨어지지 않을 때는 과대적합이 발생한 것으로 판단하여 드롭아웃 층을 2개 층으로 추가하고, 가중치 규제(Weight regularization) L1과 L2를 적용하였다. L1은 손실함수 값을 정의할 때 에러 값에 가중치 절대 값에 비례하는 값을 추가하는 것이고, L2는 가중치 제곱에 비례하는 값을 추가하는 것인데, 가중치 규제의 적용은 모델의 고차항을 나타내는 특성을 0에 가깝게 만들어줌으로써 모델을 단순화시켜 과대적합을 최소화시킨다^[17]. 이때는 손실함수 최소값이 커지기 때문에 훈련손실이 학습세대를 거듭하면서 충분히 줄어

들 수 있도록 학습세대를 200으로 적용하였다. 실험은 하이퍼파라미터를 변경하면서 VGG16 6가지, VGG19 9가지, ResNet101 6가지, ResNet50 10가지 총 31가지의 경우에 대해 실시하였고, 모델별로 얻은 가장 높은 테스트 정확도는 Table 2와 같이 각각 VGG16 86.5 %, VGG19 84.5 %, ResNet101 79.5 %, ResNet50 79.0 %였다.

4.2.2 미세조정(Fine tuning)

미세조정단계에서는 1단계에서 완전연결층을 학습시켜 우수한 테스트 정확도를 나타낸 모델들을 선택하여 컨볼루션층의 마지막 2-3개층을 재학습토록 설정하고, 학습률을 10^{-4} , 10^{-5} , 10^{-6} 등으로 변화시키면서 학습세대 100까지 학습시켜 성능이 향상되는지 확인하였다. 이처럼 완전연결층을 먼저 학습 후 컨볼루션층을 학습시킴으로써 비교적 작은 데이터셋으로도 큰 오차의 전과없이 컨볼루션층을 학습시킬 수 있다^[14]. 미세조정 후 테스트 정확도는 VGG16은 3개의 컨볼루션층을 학습률 10^{-4} 로 학습하였을 때 89.5 %, VGG19

역시 3개의 컨볼루션층을 학습률 10^{-5} 로 학습하였을 때 89.0 %를 달성하였으며, ResNet101은 2개의 컨볼루션층을 학습률 10^{-6} 로 학습하였을 때 86.0 %, ResNet50 역시 동일 조건에서 81.0 %로 모델별로 2.0~6.5 %가 향상되었다. 모델별 특성추출단계와 미세조정단계에서의 테스트 정확도의 차이는 Table 3에서 확인할 수 있다.

Table 3. Best results for each model in feature extraction and fine tuning

| Model | Feature extraction test accuracy(%) | Fine tuning test accuracy(%) |
|-----------|-------------------------------------|------------------------------|
| VGG16 | 86.5 | 89.5 |
| VGG19 | 84.5 | 89.0 |
| ResNet101 | 79.5 | 86.0 |
| ResNet50 | 79.0 | 81.0 |

4.2.3 앙상블(Ensemble)

앙상블단계는 여러 개의 다른 모델의 예측값을 결합하여 더 좋은 성능을 얻는 것을 목표로 한다. 본 연구에서는 미세조정단계에서 우수한 결과를 가진 모델들을 병렬로 구성하였고, 이를 통해 여러 개의 서로 다른 모델들의 예측을 거친 값을 Fig. 8과 같이 평균하여 최종 평가한다.

이 실험은 Table 4와 같이 VGG16, VGG19, ResNet50, ResNet101의 4개의 모델을 조합할 수 있는 총 13가지에 대하여 수행하였고, 그 결과로 VGG16, VGG19, ResNet101의 3개 모델을 조합한 경우와 4개의 모델 모두를 조합한 경우에서의 테스트 정확도가 91.5 %로 가장 높았으며, 모델의 복잡도를 고려하여 3개의 모델을 조합한 경우를 최적의 모델로 평가하였다. 2개의 모델을 조합하였을 경우에는 단일 모델의 테스트 정확도보다 우수한 정확도를 달성하지 못하였다. 그러나 3개의 모델을 조합하였을 경우에는 VGG16, VGG19 대비 성능이 낮은 ResNet을 조합하였음에도 ResNet50 90.5 % ResNet101 91.5 %로 성능이 향상됨을 확인하였다. 이는 ResNet이 VGG16과 VGG19가 추출해 내지 못하는 특성을 추출해 내기 때문인 것으로 판단된다. 4개의 모델 모두를 조합하였을 때 성능이 향상되지 않은 이유는 나머지 3개의 모델에 비하여 ResNet50에서의 테스트 정확도가 81.0 %로 다소 낮고, 성능이

낮은 모델이라 하더라도 다양한 모델을 앙상블 할 경우 성능이 향상되는 경우가 있지만, 이미 ResNet50과 유사한 구조인 ResNet101이 포함되어 있기 때문에 ResNet50이 구조의 다양성 측면에서 기여하지 못한 것으로 판단된다.

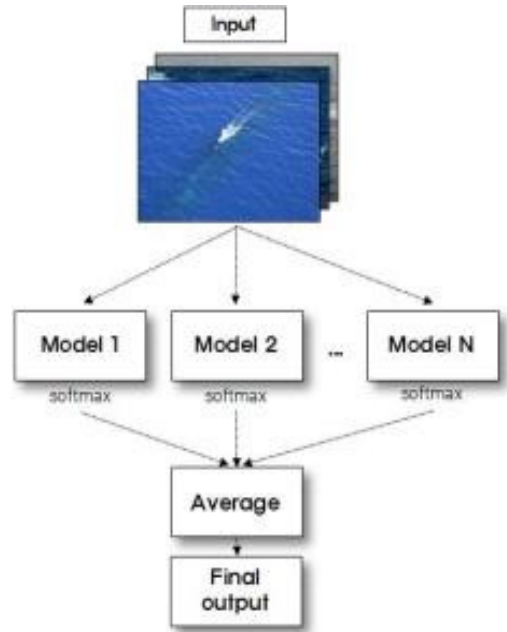


Fig. 8. A diagram of ensemble model

Table 4. Results for various model's ensemble

| Ensemble model | Test acc(%) |
|-----------------------------------|-------------|
| VGG16, VGG19, ResNet101, ResNet50 | 91.5 |
| VGG16, VGG19, ResNet101 | 91.5 Best |
| VGG16, VGG19, ResNet50 | 90.5 |
| VGG19, ResNet101, ResNet50 | 88.0 |
| VGG16, ResNet101, ResNet50 | 89.5 |
| VGG16, VGG19 | 89.5 |
| VGG16, ResNet101 | 89.0 |
| VGG16, ResNet50 | 89.0 |
| VGG19, ResNet101 | 89.0 |
| VGG19, ResNet50 | 89.0 |
| ResNet101, ResNet50 | 83.0 |

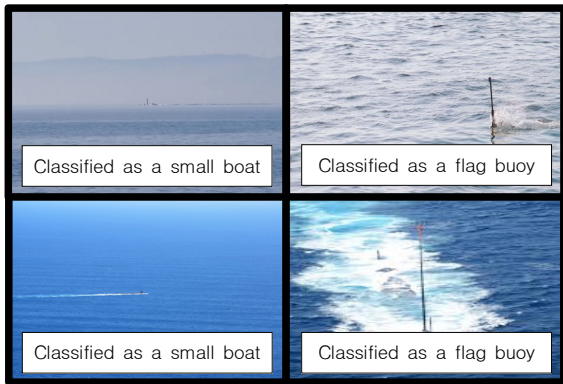


Fig. 9. Submarine images classified as other classes

최적의 모델에서 클래스별 테스트 정확도는 잠수함이 91.2 %였으며, 깃대부이 77.2 %, 등부표 98.2 %, 소형 어선·보트 100 %였다. 잠수함이 다른 표적으로 오분류된 이미지는 Fig. 9와 같고, 오분류된 이미지와 정상분류된 이미지간의 뚜렷한 특성 차이는 찾을 수 없었다. 깃대부이가 다른 클래스에 비해 정확도가 낮은 이유 또한 다른 클래스에 비해서도 현저하게 부족한 데이터 때문인 것으로 판단된다. 이는 보다 많은 데이터 확보를 통해 개선될 수 있을 것이라 판단된다.

5. 결 론

본 연구에서는 한반도에서 활동하고 있는 북한 및 주변국 잠수함에 대한 비음향 탐지율을 향상시키기 위해 CNN 모델을 활용하여 잠수함 마스트와 수상함에서 향해 중 유사한 표적으로 식별될 수 있는 깃대부이, 등부표, 소형 어선·보트에 대한 분류를 진행하였다. 이때 최대한 원거리에서부터 표적을 분류하는 기능을 구현하기 위해 작은 표적을 포함하여 학습하였고, 모델의 성능을 극대화하기 위해 특성추출, 미세조정, 앙상블을 단계적으로 진행하면서 다양한 하이퍼파라미터와 여러 경우의 앙상블 모델을 실험하였다. 결과적으로 VGG16, VGG19, ResNet50, ResNet101 4가지 모델 전부를 앙상블 하였을 때와 VGG16, VGG19, ResNet101 모델을 앙상블 하였을 때 91.5 %의 정확도를 달성하였다. 이는 단일모델에서 완전연결층만 재학습한 모델의 최고 테스트 정확도인 86.5 % 대비 5.0 % 향상된 값이며, 컨볼루션층의 일부까지 미세조정된 모델의 최고 테스트 정확도인 89.5 % 대비 2.0 % 향상

된 값이다. 앙상블 실험을 통해 성능이 낮은 모델이라 하더라도 성능이 우수한 모델과 병합함으로써 모델 전체의 성능을 향상시킬 수 있음을 확인하였다. 또한 이것은 모델의 구조가 기존 모델과 다를 때 효과가 확실함을 알 수 있었다.

본 연구의 제한사항은 보다 다양한 CNN 모델을 사용하지 못하였다는 점이다. 본 연구에서 사용한 모델 외에도 Inception_V3, DenseNet 등 활용할 수 있는 모델이 다양하지만 그러한 모델들을 다양한 하이퍼파라미터로 학습을 시도하였음에도 1학습세대에서부터 검증 정확도가 크게 저하되고 테스트 정확도 또한 낮은 문제를 보였고, 가중치 규제, 드랍아웃 층 수 조절 등의 조치에도 개선되지 않았다. 이는 학습해야 할 데이터는 단순한데 비해 모델이 복잡하여 과대적합이 발생한 것으로 추정된다. 완전연결층 유닛 수를 512로 구성하였을 때의 VGG16, Inception_V3, DenseNet의 파라미터수를 확인해 보았을 때 VGG16 2천8백만여개 (Trainable : 천3백만여개, Non-trainable : 천5백만여개), Inception_V3 4천8백만여개 (Trainable : 2천6백만여개, Non-trainable : 2천2백만여개), DenseNet 3천3백만여개 (Trainable : 2천6백만여개, Non-trainable : 7백만여개)로 VGG16에 비해 Inception_V3와 DenseNet의 모델 파라미터 수가 커 과대적합이 발생하기 쉬운 조건임을 알 수 있었다. 이러한 모델들의 기본 구조는 유사하게 설계하되 깊이를 조절하여 학습한다면 검증/테스트 정확도가 향상될 수도 있으리라 기대되지만 본 연구에서는 별도로 진행하지 않았다. 추후에는 지속적인 데이터 확보, 추가적인 모델과 최적의 하이퍼파라미터 탐색 등을 통해 보다 정확도가 향상된 모델을 구축해 나가야 할 것이다.

References

- [1] S. H. Lee, "The Value and Implications of Submarine in the Battle of the Sea," Lighthouse of Officer, Vol. 26, pp. 79-94, 2012.
- [2] S. G. Lee, et al., "Understanding Submarines 100 Questions 100 Answers," ROKN Submarine Fleet Command, 2014.
- [3] S. S. Byun, "A Study on the Non-acoustic Anti-submarine Detection Against Submarine Threats," Maritime Strategy, Vol. 153, pp. 49-91, 2012.

- [4] J. W. Park, "Comprehensive Analysis of Enemy Penetration Forces," Commander Republic of Korea Fleet, 2018.
- [5] I. K. Lee, "Development of Anti-submarine Warfare against North Korean Submarine Threats," Maritime Strategy, Vol. 161, pp. 157-195, 2014.
- [6] Saberwyn, "Onslow Masts," April 19, 2010, https://commons.wikimedia.org/wiki/File:Onslow_masts.jpg. (accessed Jan. 2, 2020.)
- [7] Y. W. Yoo, "Representative of Maritime Patrol Aircraft, P-3C," August 10, 2011, http://bemil.chosun.com/site/data/html_dir/2011/08/10/2011081001045.html?related_all. (accessed Jan. 10, 2020.)
- [8] Solaris, "Deep Learning to Learn by Tensorflow," Youngjin.com, Seoul, 2018.
- [9] J. H. Kim, D. S. Choi, H. S. Lee, J. W. Lee, "Target Classification of Active Sonar Returns based on Convolutional Neural Network," Journal of the Korea Institute of Information and Communication Engineering, Vol. 21, No. 10, pp. 1909-1916, 2017.
- [10] D. W. Kim, J. W. Seok, K. S. Bae, "Target/non-target Classification Using Active Sonar Spectrogram Image and CNN," Institute of Korean Electrical and Electronics Engineers, Vol. 22, No. 4, pp. 1044-1049, 2018.
- [11] K. H. Kim, S. J. Hong, B. H. Choi, L. H. Kim, E. T. Kim, "Ship Detection Using Faster R-CNN in Maritime Scenarios," Information and Control Symposium, pp. 158-159, 2017.
- [12] S. W. Park, J. C. Kim, D. Y. Kim, "A Study on Classification Performance Analysis of Convolutional Neural Network Using Ensemble Learning Algorithm," Journal of Korea Multimedia Society, Vol. 22, No. 6, pp. 665-675, 2019.
- [13] B. S. Kim, "A Study on Trash Classification using Convolution Neural Network," M. E. Thesis, University of Ulsan, 2018.
- [14] C. Francois, "Deep Learning with Python," Gilbut Publishing, Seoul, 2018.
- [15] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-scale Image Recognition," International Conference on Learning Representations, pp. 1-14, 2015.
- [16] K. He, X. Zhang, S. Res, J. Sun, "Deep Residual Learning for iMage Recognition in Computer Vision and Pattern Recognition," The IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778, 2016.
- [17] C. W. Kim, "Daeson's Deep Learning Foundation," Digital Books, Seoul, 2018.