

CNN을 이용한 능동 소나 표적/비표적 분류

김동욱[†], 석종원^{**}, 배건성^{***}

Active Sonar Target/Non-target Classification using Convolutional Neural Networks

Dongwook Kim[†], Jongwon Seok^{**}, Keunsung Bae^{***}

ABSTRACT

Conventional active sonar technology has relied heavily on the hearing of sonar operator, but recently, many techniques for automatic detection and classification have been studied. In this paper, we extract the image data from the spectrogram of the active sonar signal and classify the extracted data using CNN(convolutional neural networks), which has recently presented excellent performance improvement in the field of pattern recognition. First, we divided entire data set into eight classes depending on the ratio containing the target. Then, experiments were conducted to classify the eight classes data using proposed CNN structure, and the results were analyzed.

Key words: Convolutional Neural Networks, Active Sonar Classification, Spectrogram, Target/Non-target

1. 서 론

소나 기술은 다양한 목적으로 수중 환경의 표적들을 탐지하고 식별하는 데에 사용되어 왔다. 그 중 능동 소나는 송신부에서 특정 신호를 송신하고 수신부에서 표적에 반사된 신호를 수신하여 표적의 탐지 및 식별을 수행하는 기술이다. 해양환경의 다양한 변수들이 신호를 왜곡시키는 현상으로 인해 능동소나의 해석에는 음탐사의 청각을 많이 의존하고 있다. 그러나 인공지능 기술이 발전하면서 자동으로 표적을 식별하는 연구가 활발히 수행되고 있다[1-4]. 앞서 수행된 연구에서 능동소나를 식별하기 위해 다양한 방법들이 사용되었는데, 다중 방위 기반의 데이터

를 HMM(Hidden Markov Model)을 사용하여 식별한 연구[1]가 있었고, 하이라이트 모델을 이용해 합성된 표적신호를 SVM(Support Vector Machine)과 인공신경망을 사용하여 식별한 연구[2]와 FrFT(Fractional Fourier Transform)특징기반의 특징을 추출하여 BPNN(Back Propagation Neural Net)을 학습시켜 표적을 식별한 연구[3] 등이 있었다.

최근 식별연구에서는 딥러닝이 다양한 분야에서 좋은 성능을 보여주며 각광을 받고 있다[5-7]. 딥러닝은 성능이 좋을 뿐 아니라 네트워크의 일부가 특징 추출을 수행하여 사용자가 따로 특징추출과정을 수행하지 않아도 된다는 장점이 있다. 그 중 CNN(Convolutional Neural Networks)은 동물의 시각정

※ Corresponding Author: Jong Won Seok, Address: (51140) Changwondaehak-ro 20, Uichang-gu, Changwon, Korea, TEL: +82-55-213-3836, E-mail: jwseok@changwon.ac.kr

Receipt date: May 15, 2018, Revision date: Jun. 30, 2018
Approval date: Jul. 19, 2018

[†] School of Electronics Eng., Graduate School, Kyungpook National University (E-mail: cloud741@ee.knu.ac.kr)

^{**} Dept. of Information & Communication Eng., Changwon National University

^{***} School of Electronics Eng., Graduate School, Kyungpook National University
(E-mail: ksbae@ee.knu.ac.kr)

※ This research is financially supported by Changwon National University in 2017~2018

보처리과정을 모델링한 것으로 영상인식 분야에서 좋은 성능을 보이고 있는 기술이다[7]. 본 논문에서는 음향신호인 능동소나 표적 신호를 시간-주파수영역 2차원 데이터인 스펙트로그램으로 변환하고 윈도우를 사용해 이미지화 한 후, 별도의 특징추출과정을 수행하지 않고 CNN을 사용하여 식별하는 연구를 진행하였다.

본 논문의 구성은 다음과 같다. 2장에서 소나신호를 식별하기 위한 CNN구조와 표적의 스펙트로그램 이미지를 추출하는 방법에 대해 소개하고, 3장에서 데이터를 구성한 방법과 실험과정을 자세하게 서술한다. 그리고 4장에서 결론을 맺는다.

2. 능동소나 표적식별

2.1 Convolutional Neural Networks

본 논문에서는 능동 소나 이미지 데이터를 분류하기 위해 CNN을 사용하였다[7]. CNN은 동물의 시각 정보 처리과정을 모델링한 신경망 구조이다. 시각 정보가 세포에 들어오면 모든 신경세포에 자극이 전달되는 것이 아니라 해당 수용영역의 세포에서 자극을 받아들이는데 이를 모델링하여 만든 네트워크가 CNN이다. Fig. 1은 실험에서 사용한 네트워크의 구조를 보여준다. 128×128 크기의 이미지를 입력으로 하고 3개의 합성곱층과 8크기의 출력을 가진다. 플랫폼은 3번째 합성곱층의 출력을 1차원으로 펼친 것으로 출력층과 가중치로 연결되어 있다. 활성화 함수(activation function)로는 ReLU(Rectified Linear Unit)를 사용하였고, Dropout을 적용하였다.

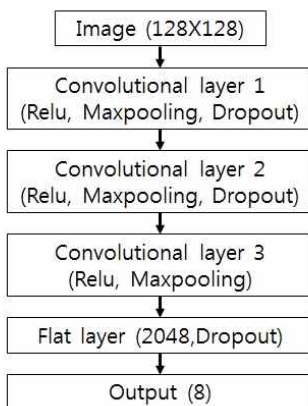


Fig. 1. CNN architecture for active sonar classification.

2.2 이미지 데이터 생성

능동 소나 표적신호의 스펙트로그램을 사용하여 8비트 그레이 이미지 데이터를 생성하였다. Fig. 2는 표적의 스펙트로그램 이미지의 예를 보여준다. 샘플링주파수 31.25kHz, 중심주파수 3.9kHz, 대역폭 400Hz, 펄스폭 50ms의 LFM(Linear Frequency Modulation)신호에 의해 얻어진 4초길이의 능동소나신호를 사용하였다[8]. 해당 이미지 전체를 네트워크의 입력으로 사용하기에는 표적신호에 해당하는 영역에 비해 너무 넓으므로 효율이 좋지 않다. 이에 따라, 윈도우를 사용하여 표적을 포함하는 128X128크기의 이미지를 추출하였다. Fig. 3은 추출한 이미지의 예를 보여준다. 윈도우는 10픽셀씩 이동하며 이미지를 추출하며, 추출하는 과정에서 이미지는 표적신호에 해당하는 영역을 모두 포함하는 경우가 있고, 표적신호의 일부만 포함하는 경우도 있으며, 표적신호를 포함하지 않는 경우도 있다.

Fig. 3의 (a)는 표적신호를 모두 포함하는 이미지이고, (b)는 위상단에 표적신호의 전반부 일부만 포함하는 이미지, (c)는 위하단에 표적신호의 후반부 일부만 포함하는 이미지이다. 또한 Fig. 4는 표적신호를 포함하지 않는 이미지의 예를 보여준다. Fig.



Fig. 2. Example of a spectrogram image of a target.

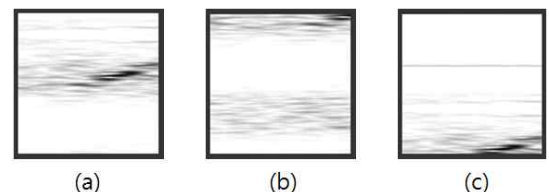


Fig. 3. Examples of target images extracted using windows. (a) Full target (b) Front part of target (c) Back part of target

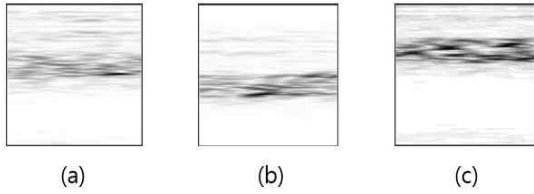


Fig. 4. Examples of non-target images extracted using windows. (a) Noise (b) Clutter (c) Clutter

4의 (a)는 잡음의 이미지를 보여주고, (b)는 단일 클러터의 이미지, (c)는 연속된 클러터의 이미지를 보여준다.

능동 소나 표적신호는 자료의 수집이 쉽지 않아 데이터가 충분하지 않은 경우가 많다. 딥러닝모델을 학습시킬 때, 데이터가 부족할 경우 모델이 주어진 학습데이터에 과적합하여 학습데이터 외의 실제 데이터에 대한 성능을 기대하기 어려워진다. 본 논문에서는 이를 해결하기 위해 데이터 확장기법을 사용하였다[7]. 데이터 확장기법은 과적합현상을 방지하기 위해 변형된 학습데이터를 생성하여 데이터의 양을 증가시키는 기법이다. 이미지 분류의 문제에서는 학습데이터에 이동, 회전, 반전 등의 방법을 적용하여 데이터 확장기법을 수행한다. 원본 데이터와 확장된 데이터를 모두 모델의 학습에 사용하면 과적합문제를 해결할 수 있고, 모델의 성능도 향상된다. 능동 소나의 스펙트로그램 이미지에서는 표적이미지가 회전, 반전되는 경우는 찾아보기 쉽지 않으므로 이동 방법을 적용한다. 이미지를 추출하는 윈도우를 시간축 및 주파수축 상에서 10픽셀씩 이동시키면 추출하는 과정에서 하나의 표적에 대해 시간-주파수축 상에서 위치가 다른 수십 장의 이미지가 생성되어 데이터 확장효과를 얻을 수 있다.

3. 실험

3.1 실험데이터

데이터 셋은 19개의 서로 다른 펄스 반복 구간에서 얻어진 빙포밍데이터이다. 전체 데이터 중 표적을 포함하는 7개 빔의 데이터를 사용하여 구성하였다. 표적이 포함된 데이터를 세분화하여 7개의 클래스로 구분하고, 표적이 포함되지 않은 데이터를 1개 클래스로 두어 8개 클래스를 설정하였다. 클래스를 구분한 기준은 추출한 이미지 내에 포함된 표적의 비율이

다. 이 비율을 결정하는 것은 표적에 대응하는 가상의 선분이며, 이 선분은 스펙트로그램에서 보이는 LFM 표적신호의 양 끝점을 기준으로 정하였다. Fig. 5는 스펙트로그램 이미지에 나타난 표적의 이미지와 선분을 정하기 위한 끝점을 보여준다. 또한 Fig. 6은 클래스를 나누기 위해 생성한 선분과 윈도우를 보여준다. 선분은 Fig. 5의 끝점에 의해 정의되고, 윈도우는 스펙트로그램이미지 위를 시간축과 주파수축으로 10픽셀씩 이동하며 이미지를 추출한다. 추출된 이미지 내에 포함된 선분의 비율을 표적의 비율 X 로 정의하여, Table 1과 같이 클래스를 구분하였다. C1~C7의 데이터는 표적데이터를 부분데이터로 세분화한 것이고, C8은 비표적데이터이다. C1은 추출한

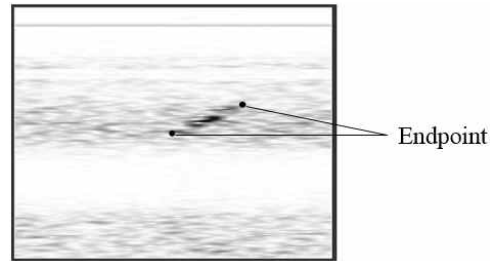


Fig. 5. Endpoint of the target.

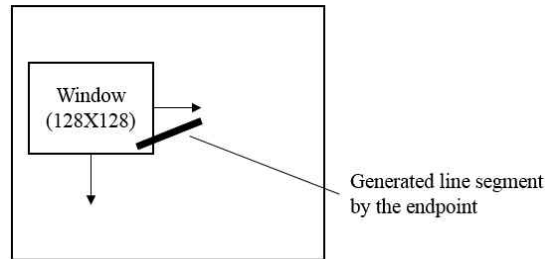


Fig. 6. Line segment to represent the percentage of the target in the window.

Table 1. Rate of target by class

Class	Rate of target
C1	$X = 100\%$
C2	$75\% \leq X < 100\%$, Front parts
C3	$50\% \leq X < 75\%$, Front parts
C4	$25\% \leq X < 50\%$, Front parts
C5	$75\% \leq X < 100\%$, Back parts
C6	$50\% \leq X < 75\%$, Back parts
C7	$25\% \leq X < 50\%$, Back parts
C8	$X = 0\%$, Non target

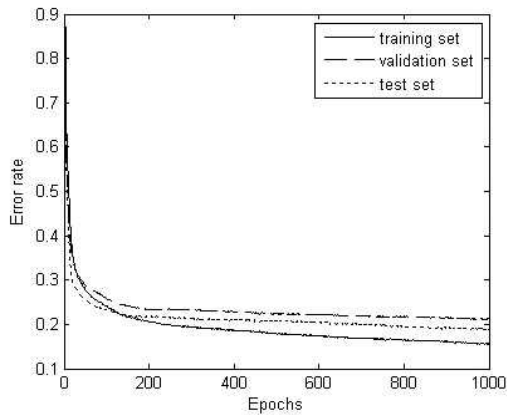


Fig. 7. Error rate of data sets for training epoch.

이미지가 완전한 표적을 포함하는 경우이고, C2,3,4는 표적의 전반부의 일부를 포함하는 경우, C5,6,7은 표적의 후반부의 일부를 포함하는 경우이다. 그리고 C8은 이미지에 표적이 포함되지 않고, 잡음이나 클러터만 존재하는 경우이다. 표적데이터를 부분데이터로 세분화한 이유는 부분데이터가 표적 및 비표적 데이터와 구분이 가능한지 확인하고 추후 연구에서 활용하기 위해서이다. 8개 클래스 데이터는 펄스 반복 구간에서 무작위로 빔을 선택하는 방법을 사용하여 학습데이터, 비교데이터, 테스트 데이터로 나누어지며, 각각 11100개, 11200개, 11100개의 데이터로 구성된다.

3.2 실험 환경

실험은 파이썬 3.5버전의 텐서플로우 라이브러리를 사용하여 진행하였다[9]. 합성곱층에 사용된 필터는 3×3의 크기로, Max pooling의 크기는 2×2로 설정하였다. 활성화함수로는 ReLU함수를 사용하였고[10],

Drop out은 첫 번째와 두 번째 합성곱층과 풀러층에 적용하였는데, 합성곱층에서 Drop out 확률은 0.8로, 풀러 층에서의 Drop out 확률은 0.5로 설정하였다 [11]. 특징맵은 첫 번째 합성곱층에서 8개, 두 번째 합성곱층에서 16개, 세 번째 합성곱층에서 8개 사용하였다. 학습률은 10^{-4} 으로 설정하였고, 배치크기는 100개, 최대 학습 횟수는 1000회로 설정하였다. 네트워크는 softmax cross entropy함수를 cost 함수로 사용하여 학습된다. 학습 횟수에 따른 오류율은 Fig. 7에 나타나 있다. 비교 데이터의 오류율은 970번째 학습에서 21.10%로 가장 낮았고, 이 때 학습데이터와 테스트 데이터의 오류율은 각각 15.75%와 19.03%였다.

3.3 실험 결과

Table 2는 클래스별 인식률을 보여준다. C1의 경우 인식률은 62.78%이고 C2와 C5에 대한 오거부율은 각각 14.89%와 20.04%이다. 그리고 C2와 C5에 대한 오인식률은 4.30%와 8.58%이다. C2의 경우는 C1과 C3로 오거부한 경우가 4.30%, 23.50%이다. 비율이 인접한 클래스들 사이의 오거부율 및 오인식률은 그렇지 않은 클래스들 사이의 결과보다 상대적으로 높다. 그 이유는 C1은 표적을 100%포함하는 데이터이고, C2와 C5는 각각 표적의 전반부와 후반부를 포함하는 75%이상 100%미만의 데이터인데, 이 데이터들 중 비율이 90%이상 100%미만의 데이터와 비율이 100%인 데이터는 눈으로 직접 구별하기에도 쉽지 않기 때문이다. 마찬가지로 비율이 75%이상 80%미만인 데이터와 70%이상 75%미만의 데이터들도 클래스는 다르지만 정확한 구별이 쉽지 않으므로 오거부율과 오인식률이 높게 나타난다. 그리고 대부분

Table 2. Results of classification (%)

	C1	C2	C3	C4	C5	C6	C7	C8
C1	62.78	14.89	1.29	0	20.04	0	0	1.01
C2	4.30	71.51	23.50	0.07	0.38	0	0	0.23
C3	0.66	5.38	71.74	21.56	0.54	0.06	0	0.06
C4	0	0	3.68	96.05	0	0.07	0.07	0.14
C5	8.58	0.69	0.23	0	73.41	16.93	0	0.15
C6	0	0	0	0	6.82	75.21	17.42	0.55
C7	0	0	0	0	0	4.09	95.54	0.37
C8	1.32	0	0.25	0	1.56	0.25	0	96.63

Table 3. Classification rates of target and nontarget data (%)

	Target	Non-target
Target	99.67	0.33
Non-target	3.37	96.63

의 오인식은 C1~C7의 표적데이터들 사이에서 발생하였고, 표적데이터들인 C1~C7과 비표적데이터인 C8의 오인식률은 모두 1.56%이하의 낮은 결과를 보여준다.

현재의 실험결과에서 C1~C7의 데이터를 표적데이터로, C8의 데이터는 비표적 데이터로 구분하여 인식률을 구해보면 99.33%이고, 표적/비표적 데이터를 구분하여 인식률을 계산하면 Table 3과 같다. 표적신호의 인식률은 99.67%로 나타났고, 따라서 오거부율은 0.33%였다. 하지만 오수락율은 3.37%로 오거부율에 비해 상대적으로 높은 편이었다.

5. 결 론

본 논문에서는 능동소나 표적신호의 스펙트로그램 이미지를 추출하고, 추출된 이미지를 식별할 합성곱 신경망을 설계하였다. 표적이 이미지에 포함되는 비율에 따라 8 클래스로 나누어 식별실험을 진행하였고, 그 결과 비슷한 비율을 가지는 클래스사이의 인식률은 비율의 차이가 큰 클래스 사이에 비해 낮은 인식률을 보였다. 마지막으로 표적/비표적 식별실험의 결과는 99.33%의 높은 성능을 보여주었고, 오거부율과 오수락율은 각각 0.33%와 3.37%로 오수락율이 높았다. 향후에는 데이터를 업데이트하여 비표적 데이터를 클러스터와 잡음 등으로 세분화하고, 오수락율을 향상시킬 수 있는 방안을 연구할 계획이다.

REFERENCE

[1] P.R. Runkle, P.K. Bharadwaj, L. Couchman, and L. Carin, "Hidden Markov Models for Multiaspect Target Classification," *IEEE Transactions on Signal Processing*, Vol. 47, Issue 7, pp. 2035-2040, 1999.

[2] T. Kim, J. Park, J. Nam, S. Lee, and K. Bae, "Synthesis and Classification of Active Sonar Target Signal Using Highlight Model," *The*

Journal of the Acoustical Society of Korea, Vol. 28, Issue 2, pp. 135-140, 2009.

- [3] J. Seok and K. Bae, "Target Classification Using Features Based on Fractional Fourier Transform," *IEICE Transactions on Information and Systems*, Vol. 97, No. 9, pp. 2518-2521, 2014.
- [4] J. Seok, "Active Sonar Target/Nontarget Classification Using Real Sea-trial Data," *Journal of Korea Multimedia Society*, Vol. 20, No. 10, pp. 1637-1645, 2017.
- [5] L. Yann, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, Vol. 521, No. 7553, pp. 436-444, 2015.
- [6] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, Vol. 61, pp. 85-117, 2015.
- [7] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012.
- [8] S. Hyung, M. Park, S. Hwang, and K. Bae, "Estimation of Target Distance Based on Fractional Fourier Transform Analysis of Active Sonar Linear Frequency Modulation Signals," *The Journal of the Acoustical Society of Korea*, Vol. 35, Issue 1, pp. 8-15, 2016.
- [9] Tensorflow, <https://www.tensorflow.org/> (accessed May, 9, 2018).
- [10] V. Nair and G.E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," *Proceedings of the 27th International Conference on Machine Learning*, pp. 807-814, 2010.
- [11] N. Srivastava, G. Hinton, and A. Krizhevsky, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *The Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1929-1958, 2014.



김 동 옥

2010년 8월 경북대학교 전자공학
부 석사
2013년 9월~현재 경북대학교 전
자공학부 박사과정
관심분야: 수중음향신호처리, 디
지털신호처리, 패턴인식 등



배 건 성

1977년 2월 서울대학교 전자공학
과 학사
1979년 2월 한국과학기술원 전기
및 전자공학과 석사
1989년 5월 University of Florida
공학박사



석 종 원

1995년 경북대학교 전자공학과
석사
1999년 경북대학교 전자공학과
박사
2004년~현재 창원대학교 정보통
신공학과 부교수

1979년~현재 경북대학교 전자공학부 교수
관심분야: 음성신호처리, 디지털신호처리, 적응필터링,
패턴인식, 소나/레이더신호처리 등

관심분야: 디지털신호처리, 수중음향신호처리, 패턴인
식, 디지털콘텐츠보호관리 등