

# 능동소나 스펙트로그램 이미지와 CNN을 사용한 표적/비표적 식별

## Target/non-target classification using active sonar spectrogram image and CNN

김 동 욱\*, 석 종 원\*\*★, 배 건 성\*

Dong-Wook Kim\*, Jong-Won Seok\*\*★, Keun-Sung Bae\*

### Abstract

CNN (Convolutional Neural Networks) is a neural network that models animal visual information processing. And it shows good performance in various fields. In this paper, we use CNN to classify target and non-target data by analyzing the spectrogram of active sonar signal. The data were divided into 8 classes according to the ratios containing the targets and used for learning CNN. The spectrogram of the signal is divided into frames and used as inputs. As a result, it was possible to classify the target and non-target using the characteristic that the classification results of the seven classes corresponding to the target signal sequentially appear only at the position of the target signal.

### 요 약

CNN(Convolutional Neural Networks)은 동물의 시각정보처리과정을 모델링한 신경망으로 다양한 분야에서 좋은 성능을 보여주고 있다. 본 논문에서는 CNN을 사용하여 능동소나 신호의 스펙트로그램을 분석하고, 표적과 비표적을 식별하는 연구를 수행하였다. 데이터를 표적이 포함된 비율에 따라 8클래스로 구분하고, CNN의 학습에 사용하였다. 신호의 스펙트로그램을 프레임별로 나누어 입력으로 사용한 결과, 표적신호의 위치에서만 표적신호에 해당하는 7개 클래스의 식별 결과가 순차적으로 나타나는 특성을 사용하여 표적과 비표적을 식별해낼 수 있었다.

*Key words* : Sonar signal processing, Active sonar, Target classification, Convolutional Neural Networks, Spectrogram

### 1. 서론

소나(sound navigation and ranging)신호처리 기법은 다양한 목적으로 수중 환경의 표적들의 탐지 및 식별에 사용되어 왔다. 특히 군사 분야에서 선박 및 잠수함의 저소음화 기술이 발달함에 따라 능동소나 신호 처리의 중요성은 크게 증가하였다. 또한 탐지 및 식별의 효율성을 위해 자동화 기술이 요구되고 있다[1-3].

딥러닝은 높은 수준의 추상화를 통해 데이터를 표현하기 위해 다수의 비선형기법을 조합하여 구성된 신경망을 학습하는 알고리즘으로 최근 다양한 식별연구에서 좋은 성능을 보여주며 주목되고

\* School of Electronics Engineering, Kyungpook National University

\*\* Dept. of Information & Communication Eng., Changwon National University

★ Corresponding author

E-mail : jwseok@changwon.ac.kr, Tel : +82-55-213-3836

※ Acknowledgment

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

Manuscript received Dec. 7, 2018; revised Dec. 21, 2018; accepted Dec. 24, 2018

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

있는 분야이다[4-5]. 딥러닝은 일반적으로 은닉층의 수가 2층 이상인 신경망을 학습하는 알고리즘을 뜻한다. 그 중 CNN(Convolutional Neural Networks)은 동물의 시각정보처리과정을 모델링한 신경망으로 이미지분류에서 좋은 성능을 보여주고 있다[6].

본 논문에서는 CNN을 사용하여 해상에서 수집된 능동소나 신호의 스펙트로그램 이미지를 분석하여 표적을 식별하는 연구를 수행하였다. 학습을 위해서는 충분한 데이터가 필요하나 소나 데이터의 수집은 신호 획득의 절차가 복잡하고 시간적 공간적 제약이 있어 제한적일 수밖에 없다. 이에 데이터 확장(data augmentation)기법을 적용하여 보유한 데이터의 양을 증가시켜 학습에 사용하였다[7]. 데이터 확장의 과정에서 획득한 데이터를 표적이 포함된 비율에 따라 8클래스로 구분하고 스펙트로그램 이미지를 분석한 결과 표적신호를 비표적신호와 구분할 수 있었다.

본 논문의 구성은 다음과 같다. I 장에서는 서론에 대해 서술 하고, II 장의 본문에서 CNN의 개요와 능동소나 식별을 위한 CNN에 대해 설명한다. III 장에서는 실험을 통해 스펙트로그램 이미지를 분석한 결과를 제시하고, IV 장에서 결론을 맺는다.

## II. 본론

### 1. CNN 개요

CNN은 동물의 시각 정보처리과정을 모델링한 신경망이다. 시각 정보가 입력되면 모든 신경세포에 자극이 전달 되는 것이 아니라 해당 수용영역의 세포에서 자극을 받아들여지게 되는데 이를 신경망 구조로 표현하여 영상처리에 적합하게 설계된 신경망이다[6].

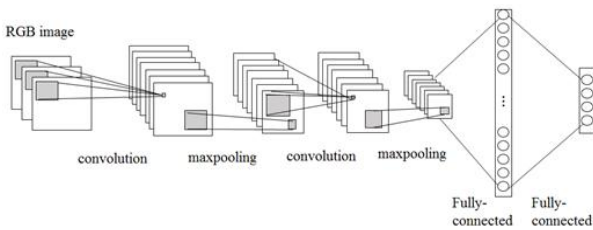


Fig. 1. Example of CNN architecture.  
그림 1. CNN구조의 예

그림 1은 CNN구조의 예를 보여준다. CNN은 크게 컨볼루션층과 풀링층, 풀링층으로 구분할 수 있

다. 컨볼루션층은 세포에 전달되는 자극을 표현한 층이며, 입력 영상의 일부를 필터와 가중합하여 다음 층에 전달하는 역할을 한다. 실제 연산은 이미지와 필터의 상호상관(cross correlation)으로 이루어지며 입력  $I$ 와  $W_f \times H_f$  크기의 필터  $F$ 에 대한 출력  $O$ 의 관계는 식 (1)과 같이 나타낼 수 있다[8].

$$O(m,n) = \sum_{w=0}^{W_f-1} \sum_{h=0}^{H_f-1} I(m+w,n+h)F(w,h) \quad (1)$$

필터의 크기에 따라 출력크기는 입력이미지보다 작아지는데 이를 보완하기 위해 입력이미지에 제로 패딩을 해줄 수 있다. 그리고 stride를 통해 컨볼루션 연산 시 필터의 시프트 크기를 설정하여 연산량을 조정할 수 있다. 하나의 축에 대해 입력의 크기가  $X$ , 필터의 크기는  $K$ , 제로 패딩의 크기는  $P$ , stride의 크기는  $S$ 일 때, 출력의 크기  $Y$ 는 식 (2)와 같이 나타낼 수 있다[9].

$$Y = \frac{X - K + 2P}{S} + 1 \quad (2)$$

식 (2)의 결과가 정수가 아닐 경우 신경망이 제대로 동작할 수 없으므로 신경망의 설계 시 반드시 고려하여야 한다.

풀링 과정은 일종의 서브샘플링 과정이며, 대표적으로 맥스 풀링(max pooling)을 많이 사용한다. 맥스 풀링은 필터 내에서 가장 큰 값만 전달하고 나머지는 무시하여 샘플링을 수행한다.

마지막으로 플랫폼은 일반적인 신경망의 은닉층과 같은 구조를 가지고 있으며, 컨볼루션 과정을 통해 추출한 특징을 식별하는 역할을 한다.

### 2. 능동소나 식별을 위한 CNN

능동소나 식별을 CNN에 적용하기 위해 표적신호를 이미지화하였다. 표적신호는 LFM(Linear Frequency Modulation)신호를 송신하여 획득한 4초길이의 능동소나신호이며, 그림2와 같이 신호의 스펙트로그램을 8비트 그레이 이미지로 변환하여 CNN의 입력으로 사용하였다. 이미지는 5166x513의 크기이며, 그림 3과 같이 윈도우를 수평 및 수직방향으로 이동하여 이미지를 추출하는 데이터 확장기법을 사용하여 128x128크기의 이미지데이터를 생성하였다[10].

생성한 데이터는 그림 3과 같이 이미지를 자르는 과정에서 표적이 모두 포함되는 경우도 있고, 표적의 일부만 포함하는 경우도 있다. 생성한 데이터가 표적신호를 포함하는 비율에 따라 식별이 가능한지 결과를 분석하기 위해 표 1과같이 데이터의 클래스를 구분하였다.

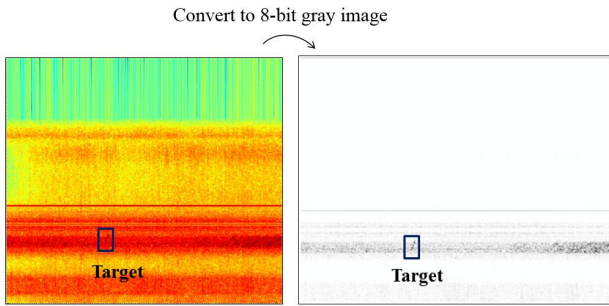


Fig. 2. Spectrogram and gray image of active sonar signal.  
그림 2. 능동소나 신호의 스펙트로그램과 이미지화 결과

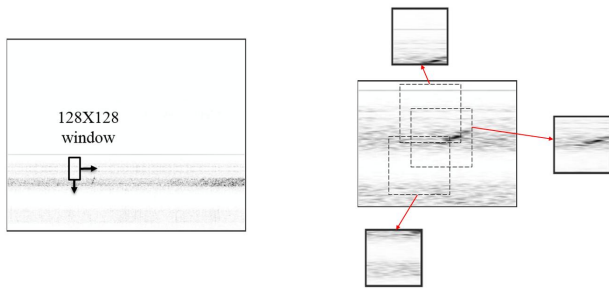


Fig. 3. Data generation using data augmentation.  
그림 3. 데이터확장을 적용한 데이터 생성

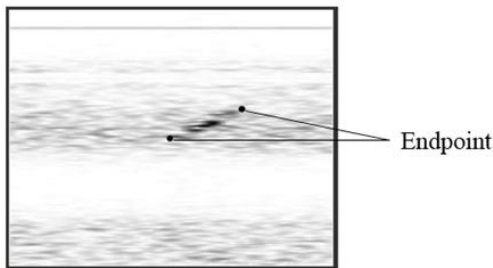


Fig. 4. Endpoint to define the ratio of inclusion of the target.  
그림 4. 표적을 포함하는 비율을 정의하기 위한 끝점

데이터에 포함된 표적신호의 비율은 그림 4와 같이 스펙트로그램에서 표적신호에 해당하는 이미지의 끝점을 설정하고 가상의 선분을 생성한 후 데이터가 포함된 선분의 비율을 계산하여 정의하였다 [10]. 표 1에서 C2, C3, C4의 경우는 표적신호의 전반부를 포함하는 데이터이고, C5, C6, C7은 표적신

Table 1. Classes according to the ratio of the target.

표 1. 표적의 비율에 따른 클래스

Class	Ratio of target	Example
C1	$X = 100\%$	
C2	$75\% \leq X < 100\%$ (Front)	
C3	$50\% \leq X < 75\%$ (Front)	
C4	$25\% \leq X < 50\%$ (Front)	
C5	$75\% \leq X < 100\%$ (Back)	
C6	$50\% \leq X < 75\%$ (Back)	
C7	$25\% \leq X < 50\%$ (Back)	
C8	$X = 0\%$ (Non Target)	

호의 후반부를 포함하는 데이터이다.

그림 5는 실험에 사용된 CNN 모델을 보여준다. 128x128크기의 입력층과 4개의 컨볼루션층, 2개의 풀랫층, 그리고 8개 유닛의 출력층으로 구성되어 있다. 각 층의 활성화함수로는 ReLU(Rectified Linear Unit)를 사용하고 있으며, drop out기법을 사용하고 있다[11-12].

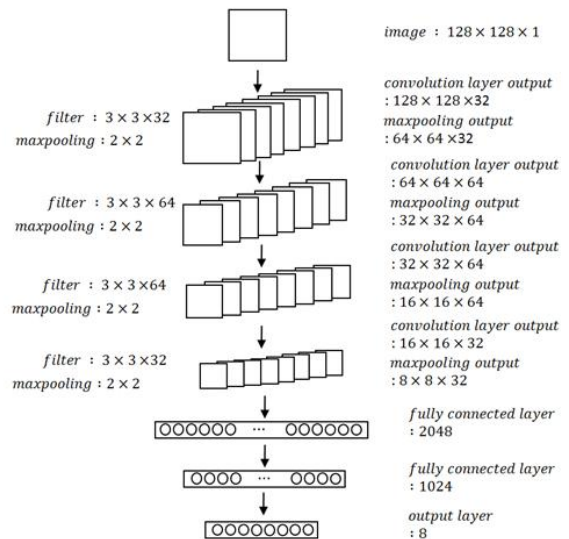


Fig. 5. CNN model using experiment.

그림 5. 실험에 사용된 CNN 모델

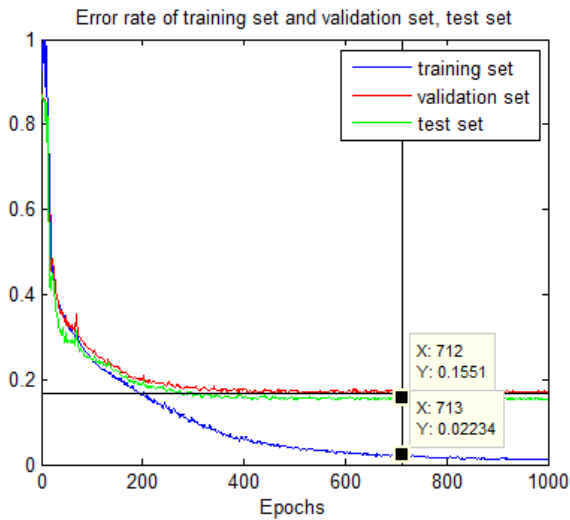


Fig. 6. Error rate depending on learning epochs.  
그림 6. 학습 횟수에 따른 오류율

### III. 실험

능동소나 표적식별을 위하여 신호의 스펙트로그램을 이미지화하여 데이터 셋을 만들고 CNN을 학습시켰다. 학습된 CNN을 검증하기 위해 신호의 스펙트로그램을 프레임별로 나누어 입력으로 사용하고 결과를 분석하였다.

실험에 사용된 데이터는 총 57개의 표적데이터이며 학습데이터, 비교데이터, 테스트데이터에 각 19개씩 사용되었다. 표적이미지의 일부를 잘라내어 데이터를 생성하는 방법으로 데이터 확장을 적용하였고, 각 11100개, 11200개, 11100개의 학습데이터, 비교 데이터, 테스트 데이터 셋을 구성하였다. 실험은 파이썬 3.5버전의 텐서플로우 라이브러리를 사용하여 진행하였다. 컨볼루션 필터는 3x3의 크기로, 맥스 풀링의 필터는 2x2의 크기로 설정하였다. 활성화함수로는 ReLU함수를 사용하였고, 컨볼루션층에서는 0.8의 확률로 drop out을 적용하였으며, 풀랫층에서는 0.5의 확률로 drop out을 적용하였다. 학습률은  $10^{-4}$ 이고, 배치크기는 100, 최대 학습 횟수는 1000회이다. 신경망은 softmax cross entropy 함수를 사용하여 학습된다. 그림 6은 학습 횟수에 따른 신경망의 오류율을 보여준다. 712회의 학습에서 비교데이터의 오류율이 가장 낮았으며, 그 때 학습데이터의 오류율은 2.09%, 테스트데이터의 오류율은 15.51%였다.

표 2는 클래스별 데이터의 식별률을 나타낸다.

표적을 포함한 비율이 인접한 클래스 사이에는 오류가 큰 것을 확인할 수 있다. 예를 들어 C2와 C3의 식별결과를 살펴보면 C2를 C3로 잘못 판단한 경우가 14.36%로 가장 컸고, C3를 C2로 잘못 판단한 경우도 8.94%에 달했다. 하지만 비율의 경계부분에서 정확하게 판단을 내리는 것은 육안으로도 힘든 경우이므로 오류가 많이 발생할 수밖에 없을 것이다. 비율이 인접하지 않은 클래스사이에서는 C3를 C1으로 잘못 판단한 경우가 가장 많았는데 1.09%의 오류율을 기록하였다. 이 오류율은 앞서 말한 비율이 인접한 클래스의 경우에 비해 현저히 낮음을 확인할 수 있다. 학습된 신경망을 사용하여 실제 스펙트로그램 영상을 입력으로 주어 식별결과를 확인해보았다.

Table 2. Result of classification experiment.









표 2. 식별실험 결과 (단위:%)

Class	C1	C2	C3	C4	C5	C6	C7	C8
C1	76.10	11.40	0.18	0	12.32	0	0	0
C2	7.68	77.65	14.36	0	0.31	0	0	0
C3	1.09	8.94	80.50	9.42	0.06	0	0	0
C4	0	0	6.13	93.81	0	0	0	0.06
C5	10.57	0.46	0.07	0	75.71	13.18	0	0
C6	0	0	0	0	8.20	80.79	11.02	0
C7	0	0	0	0	0	7.19	92.81	0
C8	0.74	1.81	0.49	0.08	0.08	0.08	0.41	96.30

전체 스펙트로그램 이미지에서 송신펄스의 주파수 범위를 포함하는 픽셀(321~448픽셀)을 정하여 128x128크기의 윈도우, 124픽셀의 오버랩으로 프레임을 나누었다. 프레임별로 식별을 수행하고 식별 결과는 표 3에 나타난 색을 통해 표시하였다. 총 19개 테스트데이터에 대해 식별을 수행하였으며, 그림 7은 식별한 결과들 중 6개 결과의 예를 보여준다. 그림에 나타난 것처럼 표적신호의 경우에는 색 띠가 표적의 비율에 따라 클래스별로 순차적으로 표시되는 것을 확인할 수 있다. 또한 비표적 신호에 대해서는 그림 7의 (a)에서 나타난 것처럼 분산된 형태의 색 띠가 나타나거나 그림 7의 (b)에서 나타난 것처럼 밀집되어 나타나더라도 표적신호와 같이 모든 클래스가 순차적으로 표시되지 않을 것을 확인할 수 있다.

Table 3. Colors for classification.

표 3. 클래스를 구분하기 위한 색

Class	Color	Class	Color
C1	 (yellow)	C5	 (green)
C2	 (magenta)	C6	 (blue)
C3	 (cyan)	C7	 (black)
C4	 (red)	C8	 (white)

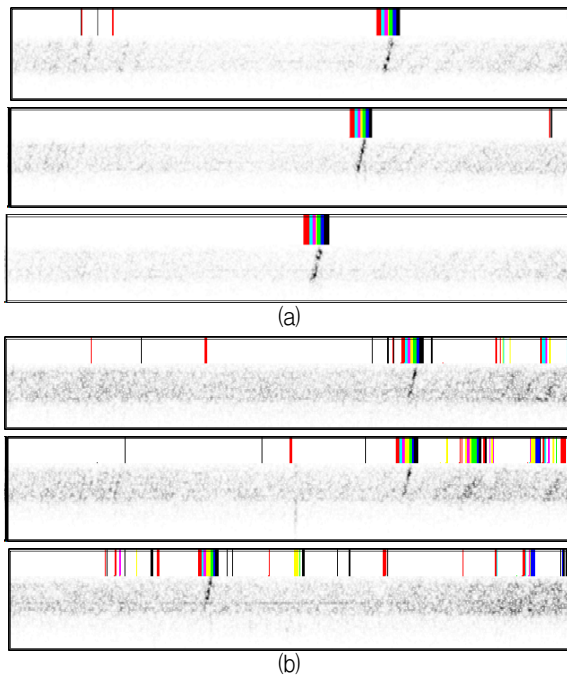


Fig. 7. Classification results of spectrogram image. (a) Examples in which a small number of misclassified data are observed. (b) Examples in which many number of misclassified data are observed.

그림 7. 스펙트로그램이미지 식별 결과 (a) 소수의 오분류 데이터가 관측된 예 (b) 다수의 오분류 데이터가 관측된 예

#### IV. 결론

CNN을 사용하여 능동소나 스펙트로그램 이미지를 분석하여 표적/비표적 식별을 수행하였다. 표적 신호를 포함하는 비율에 따라 7클래스의 표적 데이터와, 1클래스의 비표적 데이터를 데이터 셋으로 구성하고 CNN을 학습시켰다. 테스트데이터에 해당하는 19개 능동소나 표적 신호에 대해 스펙트로그램 이미지를 생성하고 프레임별로 식별실험을 수행하였다. 그 결과 표적신호에 대해서는 7개의 표적 클래스가 순차적으로 나타났고, 비표적신호에

대해서는 오분류 결과가 분산된 형태로 나타나거나 순차적이지 않은 색 띠의 형태로 나타나 표적을 구분해낼 수 있었다. 다만 보유한 데이터가 한정적이므로 향후 더 많은 데이터를 확보한 후 추가 검증이 필요할 것이다.

#### References

- [1] P. R. Runkle, P. K. Bharadwaj, L. Couchman, and L. Carin, "Hidden Markov models for multispect target classification," *IEEE Transactions on Signal Processing*, vol.47, no.7, pp.2035–2040, 1999. DOI:10.1109/78.771050
- [2] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, and G. J. Dobeck, "Underwater Target Classification Using Wavelet Packets and Neural Networks," *IEEE Transactions on Neural Networks*, vol.11, no.3, pp.784–794, 2000. DOI:10.1109/72.846748
- [3] J. Seok and K. Bae, "Target Classification Using Features Based on Fractional Fourier Transform," *IEICE Transactions on Information and Systems*, vol.E97.D, no.9, pp.2518–2521, 2014. DOI:10.1587/transinf.2014EDL8003
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* 521.7553, pp.436–444, 2015.
- [5] J. Schmidhuber. "Deep learning in neural networks: An overview," *Neural networks*, Vol.61, pp.85–117, 2015. DOI:10.1016/j.neunet.2014.09.003
- [6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol.86, no.11, pp.2278–2324, 1998. DOI:10.1109/5.726791
- [7] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp.1097–1105, 2012.
- [8] R. C. Gonzalez, and R. E. Woods, *Digital image processing, 4th ed.* Upper Saddle River, NJ: Prentice Hall, 2010.
- [9] F. Li, and A. Karpathy, "Convolutional neural networks," <http://cs231n.github.io/convolutional-networks/>
- [10] D. Kim, J. Seok and K. Bae, "Active Sonar Target/Non-target Classification using Convolutional

Neural Networks,” *Journal of Korea Multimedia Society*, Vol.21, No.9, pp.1062–1067, 2018.

DOI:10.6109/jkiice.2017.21.10.1909

[11] 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.

[12] 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.

---

## BIOGRAPHY

---

**Dong-Wook Kim** (Member)



2010 : MS degree in Electronics Engineering, Kyungpook National University.

2013- : Ph.D Candidate in Electronics Engineering, Kyungpook National University.

**Jong-Won Seok** (Corresponding author)



2010 : MS degree in Electronics Engineering, Kyungpook National University.

2010 : Ph.D degree in Electronics Engineering, Kyungpook National University.

2004~ : Professor, Changwon National University.

**Keun-Sung Bae** (Member)



1977 : BS degree in Electronics Engineering, Seoul National University.

1979 : MS degree in Electrical Engineering, KAIST.

1989 : Ph.D degree in Electrical Engineering, University of Florida.

1979~ : Professor, Kyungpook National University.