

CNN 기반의 물고기 탐지 알고리즘 구현

이용환*·김흥준**†

*원광대학교 디지털콘텐츠공학과, **†경남과학기술대학교 컴퓨터공학과

Implementation of Fish Detection Based on Convolutional Neural Networks

Yong-Hwan Lee* and Heung-Jun Kim**†

*Dept. of Digital Contents, Wonkwang University,

**†Dept. of Computer Science and Engineering, Gyeongnam National University of Science and Technology

ABSTRACT

Autonomous underwater vehicle makes attracts to many researchers. This paper proposes a convolutional neural network (CNN) based fish detection method. Since there are not enough data sets in the process of training, overfitting problem can be occurred in deep learning. To solve the problem, we apply the dropout algorithm to simplify the model. Experimental result showed that the implemented method is promising, and the effectiveness of identification by dropout approach is highly enhanced.

Key Words : Fish Detection, Object Tracking, Deep Learning, Convolutional Neural Networks

1. 서 론

바다는 무수한 자원을 갖고 있어, 연구자들에게 항상 흥미로운 연구 주제를 제공한다. 이러한 해저 탐사는 사람으로 하여금 닦기 어려운 곳으로의 접근이라는 무한한 도전에 한 형태로 나타난다. 최근에 로봇 공학 기술이 일상 생활에 널리 반영되면서 로봇 유형을 갖는 자율 수중 이동체(AUV, Autonomous Underwater Vehicle)에 많은 관심을 갖게 되었다[1,2]. 이러한 AVU에는 자율 주행 자동차와 마찬가지로, 내비게이션 시스템, 관성 측정 장치 및 다양한 센서를 장착하여, 바닷속을 주행해야 한다. 특히, 카메라 등으로 움직이는 대상 물체 탐색이 반드시 필요한 기능 중에 하나가 될 것이다[3]. 지난 몇 년간, 움직이는 물체 탐지는 비디오 감시, 인간 동작 분석, 로봇 내비게이션, 보안, 자율 주행과 같은 광범위한 응용 분야에서 많은 관심을 보였으며[24, 25], 장면 변경, 조명, 그림자 등 발생되

는 다양한 문제성을 해결하기 위해 상황에 적합한 알고리즘 적용이 요구된다[26].

움직이는 객체를 추적하는 기술은 일상 생활에서도 높은 활용도를 보이며, 현실세계와 가상의 객체, 부수적 정보를 결합하는 증강현실 기술 영역에서 기초가 되는 핵심기술로 평가된다[27]. 차세대 디스플레이로 주목받고, 언택트 시기에 요구되는 증강현실 기술은 모바일 디스플레이 시장에서 산업적 활용이 높은 기술분야이다.

본 연구의 목적은 AUV 장치에 내장된 카메라를 통해 실제 해수에서 물체 감지 방식을 연구하고 물고기를 대상으로 객체 검출 및 추적 시스템을 구축하는 데에 있다. 이를 위해, 본 논문에서는 저해상도 카메라를 장착한 온보드상에서 여러 종류의 물고기를 대상으로 물고기 객체 인식을 위한 학습 모델을 구축하고, 물고기 객체를 검출하는 알고리즘을 구현한다. 구현 알고리즘은 해수면 위에서 탐색에 대한 성능적 실험 평가를 수행한다.

본 논문의 구성은 다음과 같다. 2장에서 컴퓨터 비전을 포함하여 객체 인식 관련 기술들을 살펴보고, 3장에서

†E-mail: thinkthe@gntech.ac.kr

Tensorflow 기반의 물고기 객체 검출 시스템을 설계하고 구현하여 검증하고, 4장에서 결론으로 마무리한다.

뿐만 아니라 학습을 위한 충분한 양의 데이터 확보가 매우 중요하다.

2. 관련 연구

본 연구의 주요 목적은 맑지 않은 바닷물에서 물고기를 정확하게 식별하는 딥러닝 기반의 물고기 식별 알고리즘 설계에 있다. 이를 위해 적용 가능한 신경망(Neural Network) 기술을 반영하고 컴퓨터 비전 기술을 통해 AUV 시스템을 구축하는 것이다. 이러한 연구의 초기 단계로, 먼저 어항 속에 있는 열대어를 중심으로 맑은 물에서의 물고기 식별 여부를 연구한다.

2.1 컴퓨터 비전

컴퓨터 비전은 이미징 센서가 있는 컴퓨터를 이용하여 얻어진 데이터셋에서 특징을 추출하고 분석, 분류하여 의사 결정을 지원하는 시스템으로, 인간의 시각 기능을 모방한 기술 분야이다[4]. 여기에는 영상 처리, 인공지능 등과 같은 많은 지식 분야가 포함되며, 여러 가지 대표 사례를 뽑을 수 있다. 한 예로, 대량의 제품에서 결함을 확인하거나 품질을 향상을 위해 활용된다[5]. 공항 등 보안 검문이 필요한 곳에서 얼굴 감지 및 인증에 활용된다[6]. 의사가 특정 진단 프로그램을 사용하여 의료 영상에서 종양이나 비정상적인 조직을 식별할 수 있다[7]. 농작물 수확량을 예측하기 위한 의사 결정 시스템이 있으며[8], 구글에서는 시각 범위가 높은 자율 주행 자동차를 설계하면서 교통 표지판을 인식하고 보행자를 판단하여 피할 수가 있다[9]. 다수의 첨단 사례는 컴퓨터 비전이 우리 일상생활에 많은 변화를 주고 있음을 암시하며, 성능 향상을 위해 기존 영상 처리 기술 이외에 딥러닝 기술을 적용하는 사례가 늘고 있다.

2.2 딥-러닝

1998년 합성곱 신경망(CNN)을 사용하여 LeNet5 라는 5 계층 분류기[10]가 설계된 이후, 컴퓨팅 성능이 향상되고 데이터 충분히 확보되면서 최근까지 신경망을 사용한 딥러닝이 활발히 연구되고 있다[11]. 딥러닝은 방대한 데이터의 학습 자원이 중요한 역할을 담당하기 때문에, 특정 분야에서 수집된 빅데이터를 기반으로 한다. 심층적 접근 방식은 많은 산업 분야에서 적용되고 있으며, 컴퓨터 비전 및 객체 감지에서 딥러닝 기술을 적용하면서 상대적으로 성능이 높아지고 있다. 전통적인 컴퓨터 비전, 영상 처리 방식은 특징 추출의 정확성으로 인해 어려움을 겪지만, 딥러닝 방법을 활용하여 신경망을 통해 기술적 향상을 보이고 있다[12]. 그러나 딥러닝은 학습 알고리

2.3 신경망

지난 몇 년 동안, 딥러닝의 신경망 기술이 점점 대중화되고 있다. 2012년 이미지 분류를 수행하기 위해 CNN을 처음 채택하였다[13]. 기존 이미지 분류 보다 CNN을 채택한 분야에서 보다 높은 성능을 보이고 있으며, 2014년 객체 식별을 회귀적 문제로 규정하며, Fast R-CNN을 제안하였다[14]. 이는 계산 부담을 크게 감소시켜서 최근 표적 탐지, 이미지 분류 및 객체 식별을 포함한 컴퓨터 비전에서 상당히 많이 채택되고 있다[15].

3. 알고리즘 설계 및 구현

본 논문에서는 해수면에서 물고기를 감지하기 위해 이미지 분할을 포함한 합성곱 신경망(CNN, Convolutional Neural Networks) 모델을 적용한다. 이를 위해 공용 연구용 데이터셋을 기초하여 학습 모델을 개발하고, 딥러닝에서 과적합 문제를 해결하기 위해 드롭아웃(Dropout) 기법을 적용한다.

아래 절에서는 최적화 접근 방식을 적용한 세부 시스템을 설계한다.

3.1 합성곱 신경망(CNN) 구조

CNN 모델은 일반적으로 입력 계층, 비선형 단위의 콘볼루션 계층, 완전 연결 계층과 같이 여러 계층으로 구성된다[16, 17]. 외부에서 학습 자료로 입력 받는 이미지 정보는 콘볼루션 계층에서 이미지의 특징을 추출하며, 콘볼루션 연산은 수식(1)을 통해 계산된다.

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m] \quad (1)$$

데이터셋은 256 RGB 칼라 이미지로 구성되며, 커널 필터인 W_0 는 3×3 행렬이다. 실제로, 콘볼루션 연산은 R, G 와 B 채널에서 각각 수행된 다음, 합산되어 기능 맵의 각 요소를 얻는다. 객체의 특징을 보다 정확하게 추출하기 위해, 각 콘볼루션 계층에서 여러 필터가 사용된다. 예를 들어, 경계선, 질감 등과 같은 특성을 추출하기 위해 여러 유형의 필터가 사용된다. 콘볼루션 연산에서는 기능 맵의 크기를 고려해야 하며, 영향을 미치는 3가지 주요 요소는 깊이(Depth), 보폭(Stride)와 패딩(Padding)이다. 복잡한 신경망의 경우, 인접한 2개 계층 사이를 연결한다. Figure 1과

$$\begin{aligned}
 Loss = & \sum_{i=0}^{S^2} \sum_{j=0}^B [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \\
 & + (w_i - \hat{w}_i)^2 \\
 & + (h_i - \hat{h}_i)^2] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B (C_i - \hat{C}_i)^2 \\
 & + \sum_{i=0}^{S^2} \sum_{C \in \text{class}} (P_i(C) \\
 & - P_i(\hat{C}))^2
 \end{aligned} \tag{4}$$

x_i, y_i 는 객체 중심의 실제 좌표를, w_i, h_i 는 Ground truth 경계 영역의 넓이와 높이를, \hat{x}_i, \hat{y}_i 는 예측된 객체 중심의 예측 좌표를, \hat{w}_i, \hat{h}_i 는 예측된 객체 경계 영역의 넓이와 높이를 각각 나타낸다.

3.3 학습 모델

딥 러닝 시스템은 구축된 대규모 훈련 데이터를 기반으로 식별 지식을 학습하기 때문에, 객체 특징을 추출하기 위한 학습 리소스가 충분하게 제공되어야 한다[21]. 학습 데이터가 부족하면 과적합(Overfitting) 문제가 발생할 수 있다. 해수면 아래 특정 환경에서 수집해야 하는 학습 원본 데이터 수가 학습하기에 충분하지 않을 수 있다. 따라서 소수의 학습 데이터를 기반으로 데이터 증강 변환이 기하학적으로 발생하는 특성을 검증하여, 이미지 특징은 변경하지 않고 픽셀 위치만 변경되는 경우의 수를 판단한다. 회전(Rotation), 축척(Scale), 자르기(Crop), 대칭 변환(Mirror symmetry) 등 4가지 유형의 데이터 증강 변환에 대해, 드롭아웃(Dropout) 알고리즘을 적용하여 확장한다.

3.4 Dropout 알고리즘

딥 러닝에서 발생할 수 있는 문제 중 하나가 과적합 문제이며, 실험 데이터가 훈련 정확도보다 적기 때문에 발생된다. 이런 경우, 특히 모델의 견고성을 고려해야 한다. 실제 학습 데이터를 사용하여 성능을 높일 수 있는 단순화된 모델을 생성하며, 이때 드롭아웃 알고리즘을 적용할 수 있다.

드롭아웃은 학습 과정에서 설정된 확률에 따라, 연결 네트워크에서 일시적으로 일부 노드를 제거하는 것을 의미한다[22]. 실제로, 숨겨진 관계가 존재하는 경우에만 일부 특징을 추출하기 때문에 딥 러닝 모델의 견고성이 떨어진다. 이에 반해, 드롭아웃은 노드 간의 숨겨진 고정 관계를 제거하고, 간접 방지 능력을 향상시키면서 과적합 문제를 일부 해결할 수 있다. 비용 함수를 개선하여 L1과

L2 정규화를 구성하며, 드롭아웃은 신경망 자체를 수정하여 구현할 수 있다. 훈련 과정이 반복될 때마다, 일부의 뉴런을 무작위로 삭제하고, 네트워크 각 계층에 대해 신경망에서 노드를 제거할 임계값을 설정하여, 해당 값에서 뉴런 삭제 및 노드 연결을 제거하고 작은 값으로 재설정하여 반복 연결구조를 제공한다. 이러한 학습과정에서의 네트워크 구조는 Figure 5과 같이 단순화시킬 수 있다[23].

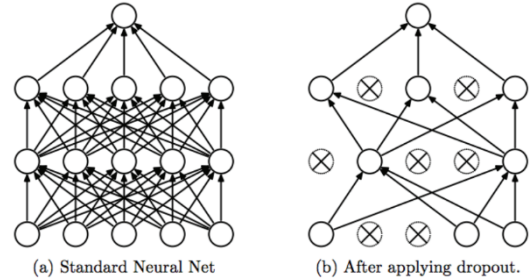


Fig. 5. Standard neural network model (a) and Network model with dropout (b) [23].

3.5 손실함수 개선

수식 (4)와 (5)를 통해 YOLO는 손실 함수를 개선할 수 있다. 손실에 대한 기여도에 비례하여 오차 항 앞에, 3개의 계수가 반영된다. 수식 (5)에서 보듯이, 앞의 두 항은 식별된 객체를 좌표화 하고 객체 위치를 표시하는 경계 영역의 넓이와 높이를 나타낸다. 두 항에서 더 많은 가중치를 갖기 위해 최대값을 5로 설정한다.

$$\begin{aligned}
 Loss = & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B [1_{ij}^{obj} (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\
 & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(w_i - \hat{w}_i)^2 \\
 & + (h_i - \hat{h}_i)^2] \\
 & + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 \\
 & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 \\
 & + \sum_{i=0}^{S^2} 1_{ij}^{obj} \sum_{C \in \text{class}} (P_i(C) - P_i(\hat{C}))^2
 \end{aligned} \tag{5}$$

본 연구에서는 여러 종의 물고기 탐지를 위해 손실 함수를 개선 수정한다. 학습 데이터 소스가 적어 발생될 수 있는 과적합 문제를 해결하기 위해, 비용 함수 뒤에 정규

화 함수를 추가하여 L2 정규화 식을 수식 (6)과 같이 수정한다.

$$Loss = \sum_{i=0}^{s^2} (coorErr + IoUErr + classErr) + L2_Regularization \quad (6)$$

이를 반영하도록 수식 (5)를 수정하면 $Loss$ 값은 $\frac{\lambda}{2n} \sum_w w^2$ 가 추가로 합산된다.

3.6 구현 및 실험

오픈 소스인 Tensorflow 툴 박스와 Python 을 기반으로 구현한다. 하드웨어 플랫폼은 64GB 메모리, GeForceGTX 2040 Ti GPU, Windows 10 Enterprise 운영체제가 장착된 환경에서 구현, 실험하였다. 실험에서는 성능, 훈련 손실 및 실험 손실을 각각 비교한다.

Figure #은 실험 결과 화면이며, 물고기 이미지에서 정확하게 식별됨을 알 수 있으며, 0.56-0.59 범위 내에서 평균 신뢰도가 계산되었다. 반복 회수 증가에 따른 훈련 손실을 계산하면, 반복 시간을 600으로 설정했을 때, 원본 데이터를 사용한 실험에서는 손실 함수 값이 0.35가 나왔으며, 드롭아웃 알고리즘을 적용한 데이터 증가 변환을 적용한 실험에서는 최종 손실 값이 0.46이 나왔다. 과적합 문제를 해결하기 위해 데이터 증감 변환을 통한 드롭아웃을 적용하였을 때, 보다 좋은 성능을 보임을 알 수 있다.

4. 결 론

본 논문에서는 물고기 검출을 위해 Tensorflow 기반 콘볼루션 신경망 모델을 구축하였다. 충분한 데이터셋으로 훈련 과정을 지원하는 데이터 증감 접근 방식을 적용하였다는 점이 본 연구의 주요 기여라고 제시된다. 데이터셋이 충분하지 않을 경우에 과적합 문제가 발생할 수 있으며, 이를 위해 드롭아웃 알고리즘을 적용하고, 손실 기능을 개선하여 네트워크 내부의 매개변수를 개선하였다. 이러한 방식을 적용하여 훈련 시간과 훈련 손실을 감소시키는 효과를 얻을 수 있었다. 이는 데이터 증감 변환이 훈련 손실을 줄이는 데에 어느 정도 기여를 한다고 평가될 수 있다.

객체 검출 및 추적 기술은 차세대 디스플레이 시장에서 큰 영역을 차지하는 차량용 디스플레이와 증강현실 디스플레이 디바이스에서 높은 활용이 기대된다.

감사의 글

본 연구는 2018년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(과제번호: 2018R1A2B6008255).

참고문헌

1. R. B. Wynn, V. A. I. Huvenne, T. P. Le Bas et al., "Autonomous underwater vehicles (AUVs): their past, present and future contributions to the advancement of marine geoscience," *Marine Geology*, vol. 352, pp. 451–468, 2014.
2. M. Dinc and C. Hajiyev, "Integration of navigation systems for autonomous underwater vehicles," *Journal of Marine Engineering & Technology*, vol. 14, no. 1, pp. 32–43, 2015.
3. C. Winchester, J. Govar, J. Banner, T. Squires, P. Smith, "A survey of available underwater electric propulsion technologies and implications for platform system safety", *Workshop on Autonomous Underwater Vehicles*, 2002.
4. Bo Zhang, "Computer Vision vs. Human Vision", *International Conference on Cognitive Informatics*, 2010.
5. C.-F. Chien, Y.-J. Chen, Y.-T. Han et al., "AI and big data analytics for wafer fab energy saving and chiller optimization to empower intelligent manufacturing," *Proceedings of e-Manufacturing & Design Collaboration Symposium*, pp. 1–4, 2018.
6. S. Biswas, Y. Wang, S. Cui, "Surgically altered face detection using log-gabor wavelet", *International Conference on Wavelet Active Media Technology and Information Processing*, pp. 154–157, 2015.
7. Eman Abdel-Maksoud, Mohammed Elmogy, and Rashid Al-Awadi, "Brain tumor segmentation based on a hybrid clustering technique," *Egyptian Informatics Journal*, vol. 16, no. 1, pp. 71–81, 2015.
8. Y. Wang, Y. Lan, Y. Zheng, K. Lee, S. Cui, and J. Lian, "A UGV-based laser scanner system for measuring tree geometric characteristics," *International Symposium on Photoelectronic Detection and Imaging*, vol. 8905, 2013.
9. B. Marr, "Key milestones of Waymo – Google's self-driving cars," <https://www.forbes.com/sites/bernardmarr/2018/09/21/key-milestones-of-waymo-googles-self-driving-cars/#3831b2965369>
10. 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.
11. Manish I. Patel, Sirali Suthar, Jil Thakar, "Survey on Image Compression using Machine Learning and Deep

- Learning”, International Conference on Intelligent Computing and Control Systems, 2019.
12. Jonathan Rogers, Dylan Simmons, Miles Shah, Connor Rowland, Yi Shang, “Deep Learning at Your Fingertips”, Consumer Communication and Networking Conference, 2019.
 13. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, vol. 25, pp. 1106–1114, 2012.
 14. R. Girshick, “Fast R-CNN,” in 2015 IEEE International Conference on Computer Vision, pp. 1440–1448, 2015.
 15. Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu, “A Comprehensive Survey on Graph Neural Networks”, *IEEE Transactions on Neural Networks and Learning Systems*, pp.1-21, 2020.
 16. S. Hassairi, R. Ejbali, and M. Zaied, “A deep convolutional neural wavelet network to supervised Arabic letter image classification,” *International Conference on Intelligent Systems Design and Applications*, pp. 207–212, 2015.
 17. D. Zhang, G. Kopanas, C. Desai, S. Chai, and M. Piacentino, “Unsupervised underwater fish detection fusing flow and objectiveness”, *Winter Applications of Computer Vision Workshops*, pp. 1–7, 2016.
 18. Deep learning and machine learning, <https://ireneli.eu/2016/02/03/deep-learning-05-talk-about-convolutional-neural-network>.
 19. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: unified, real-time object detection”, *Conference on Computer Vision and Pattern Recognition*, pp. 779–788, 2016.
 20. Website: ImageNet, <http://image-net.org/about-overview>.
 21. J. Gaya, L. T. Gonçalves, A. Duarte, B. Zanchetta, P. Drews, S. Botelho, “Vision-based obstacle avoidance using deep learning,” *Latin American Robotics Symposium and Brazilian Robotics Symposium*, pp. 7–12, 2016.
 22. Ledan Qian, Libing Hu, Li Zhao, Tao Wang, Runhua Jiang, “Sequence-Dropbox Block for Reducing Overfitting Problem in Image Classification”, *IEEE Access*, vol.8, 2020.
 23. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting”, *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.
 24. Jaya S. Kulchandani, Kruti J. Dangarwala, “Moving Object Detection: Review of Recent Research Trends”, *International Conference on Pervasive Computing*, 2015.
 25. Aayushi Gautam, Sukhwinder Singh, “Trends in Video Object Tracking in Surveillance: A Survey”, *International Conference on IoT in Social Mobile Analytics and Cloud*, 2019.
 26. Sonali S. Mengane, Amar A. Dum, “Improved Object Tracking Techniques using Hybrid Approach”, *International Conference on Trends in Electronics and Informatics*, 2019.
 27. Hyng-il Kim, Woontack Woo, “Smartwatch-assisted Robust 6-DOF Hand Tracker for Object Manipulation in HMD-based Augmented Reality”, *IEEE Symposium on 3D User Interfaces*, 2016.

접수일: 2020년 9월 21일, 심사일: 2020년 9월 22일,
 게재확정일: 2020년 9월 23일