

3D Faster R-CNN 을 이용한 의료영상 비전분석

엄종석*

*한성대학교 컴퓨터공학부

e-mail : jsum@hansung.ac.kr

Vision Analysis of Medical Images Using 3D Faster R-CNN

Jong-Seok Um*

*Division of Computer Engineering, Hansung University

요 약

머신 러닝 알고리즘 중 하나인 CNN 은 2 차원 영상 학습에 특화되어 3 차원 의료 영상에 사용하기는 부적합하며, 이에 본 연구에서는 3D-CNN 을 이용하여 의료 영상을 학습할 수 있는 모델을 구현하였다. 학습의 정확성을 높이기 위해 Region Proposal 을 적용한 Faster R-CNN 을 사용하였으며, 실험 결과 기존 CNN 알고리즘으로는 불가능했던 3 차원 의료 영상의 비전 분석이 가능하게 되었다.

1. 서론

현재까지 개발된 모든 머신 러닝 알고리즘은 ANN 에서 시작되었다. ANN 은 인간의 뇌에 있는 뉴런과 시냅스의 연결 형태를 프로그램으로 구현한 것이며, 적당한 input 과 output 을 가지는 데이터에 적용이 가능하다. 그러나 ANN 은 영상과 같이 많은 수의 input 이 있는 데이터에는 사용하기 부적합하며, 이를 위해 개발된 알고리즘이 CNN[1]이다.

CNN 의 기본 작동과정은 영상에 filter 를 convolution 하여 feature 들을 추출하는 과정의 연속이다. 이 때 사용되는 filter 들은 학습이 가능한 parameter 들로 구성되어 있으며, gradient descent algorithm 을 사용하여 오류가 줄어드는 방향으로 지속 갱신된다. 영상에 대한 세밀한 분석을 원하는 경우 보다 더 많은 수의 filter 와 layer 를 사용하면 되는데, 모델이 크고 복잡해질수록 모델의 학습시간과 난이도 또한 증가하게 된다. 따라서 좋은 CNN 모델을 구축하기 위해서는 적절한 hyper-parameter 를 결정하는 것도 중요하지만, 학습이 올바른 방향으로 빠르게 진행될 수 있도록 학습 최적화 알고리즘을 적절하게 적용하는 것이 핵심이다.

2. 3D Convolutional Neural Network

현재 사용되는 대표적인 의료 영상으로는 컴퓨터 단층 촬영 영상(computed tomography, CT) 이나 자기공명 영상(magnetic resonance imaging, MRI), 양전자 방출 단층 촬영 영상(Positron Emission Tomography, PET) 등과 같은 것들이 있으며 이러한 영상들은 모두 3 차원 입체 영상이다.

기존의 CNN 모델의 경우 2D 영상에 대해서는 우수한 성능을 보이지만 3D 영상에 대해서는 그다지 좋은 성능을 보이지 않는데, 이는 CNN 알고리즘 자체가 2 차원 영상을 처리하기 위해 고안되어 있기

때문이다. 만약 CNN 에 3 차원 영상을 학습시키려 하는 경우 depth 에 관한 정보는 대부분 잃어버리게 되며, 정교한 특징 검출이 불가능하게 된다.

이러한 문제점들은 기존의 CNN 알고리즘에서 한 차원 확장된 3D-CNN 알고리즘[2]을 이용하면 해결이 가능하다. CNN 의 convolution 은 row 와 column 방향으로만 처리되지만, 3D-CNN 은 depth 방향까지 3 차원으로 convolve 되며 이로 인해 depth 정보를 잃어버리지 않고 모든 feature 를 추출할 수 있게 된다. 또한 Region Proposal 을 통해 물체가 있을 만한 영역을 미리 찾아서 제안된 영역에서만 결절 여부를 검사하는 3D Faster R-CNN[3]을 적용하여 그 정확성을 높였다. 3D Faster R-CNN 의 구조는 아래 그림 1 과 같다.

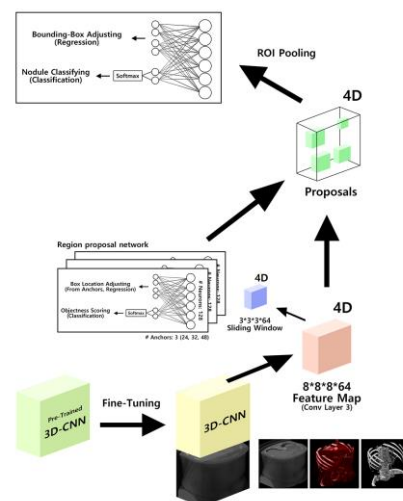


그림 1. Faster 3D R-CNN

3. 구현

먼저 머신 러닝 알고리즘으로는 앞서 소개한 3D-CNN 알고리즘을 사용하였다. 학습 데이터는 2016년 Grand-Challenge 사이트에서 개최된 LUNA16 챌린지[4]의 데이터를 사용하였다. 이 데이터는 888명의 폐 결절 환자의 3D CT 데이터와 이들에 대해 4명의 영상 전문의가 판독한 정보로 구성되어 있다. 영상 판독 데이터의 경우 의심 환부가 모두 나열되어 있으며, 여기서 실제로 결절인 경우 class 1, 결절이 아닌 경우 class 0으로 분류되어 있다. 이 때 class가 0인 데이터와 1인 데이터의 비는 735418:1557로 데이터의 분포 불균형이 매우 심한데, 이 문제는 class 0인 데이터의 선택적 사용과 class 1인 데이터의 over sampling을 통해 보완하였다. Over sampling은 원본 데이터를 그대로 이용하지 않고 영상의 회전, x, y, z축 flipping, offset jittering 등을 통해 최대한 다른 데이터처럼 보이게 하였다.

3D-CNN 모델의 layer는 input layer를 제외한 7개로 구성하였으며, 자세한 내용은 아래 그림 2와 같다.

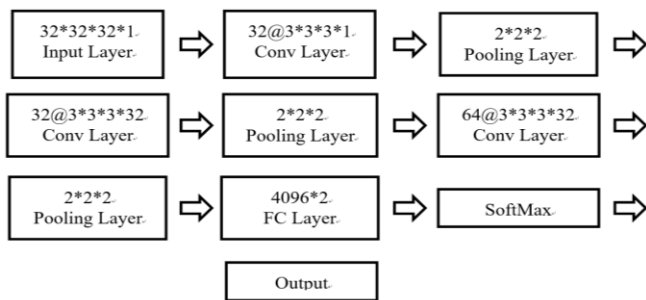


그림 2. 3D-CNN 모델의 Layer 구조도

처음에는 2D-CNN의 성능이 우수한 듯 보이나, Epoch 20 즈음에 들어서자 Error가 폭발하여 학습이 중단된다. 반면 3D-CNN 모델의 경우 꾸준히 학습이 진행되고 있는 것을 확인할 수 있다. 3D-CNN의 경우 2D-CNN보다 parameter가 많으므로 기본적으로 학습 속도가 느린데, 높은 learning rate를 주어도 gradient가 exploding되지 않고 학습이 잘 되는 특성을 보이므로 위 모델에서 learning rate를 높여준다면 더욱 좋은 성능을 보일 것이다.

4. 결론

3D Faster R-CNN 알고리즘을 이용하여 의료 영상의 비전을 효율적으로 분석할 수 있는 모델을 구현하였으며, 이를 이용하면 응급 환자의 병변을 빠르게 판독할 수 있다.

참고문헌

- [1] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document

recognition" Proceedings of the IEEE, v. 86, pp. 2278- 2324, 1998.

- [2] Shuiwang Ji, Wei Xu, Ming Yang, and Kai Yu, "3D convolutional neural networks for human action recognition", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 35, no. 1, pp. 221-231, 2013.

- [3] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards real-Time Object Detection with Region Proposal Networks, NIPS15, 2015

- [4] A.A.A. Setio, A. Traverso, T. Bel and et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. In arXiv:1612.08012, 2016.