

3D CNN 기반 전립선 MRI 영상 분할 기술

문주혁, 최환, 이세호, 장원동, 김창수
고려대학교

{jmun, hwanc, seholee, wdjang, cskim}@mcl.korea.ac.kr

3D CNN-Based Segmentation of Prostate MR images

Juhyeok Mun, Hwan Choi, Se-Ho Lee, Won-Dong Jang, Chang-Su Kim
Korea University

요 약

본 논문에서는 남성의 하반신을 촬영한 MRI 영상으로부터 전립선을 분할하는 알고리즘을 제안한다. 우선 3 차원 입체 영상을 학습하기 위해 3D 컨볼루션 계층(convolutional layer) 및 3D 풀링 계층(pooling layer)에 기반한 네트워크를 제안한다. 다음으로 네트워크의 최후단에 해당하는 전연결 계층(fully connected layer)의 강인한 학습을 돕는 잡음 계층을 제안한다. 잡음 계층은 네트워크의 학습 파라미터 혹은 출력 영상에 가우시안 잡음을 더함으로써 드롭 아웃과 같이 훈련 영상에 대한 과적합(overfitting)을 막고 테스트 영상에 강인한 네트워크의 학습을 돕는다. 마지막으로 실험을 통해 제안하는 기법이 기존 기법에 비해 우수한 분할 성능을 보임을 확인한다.

1. 서론

전립선 분할 기술은 3D MRI 영상에서 전립선으로 판단되는 영역을 분할한다. 분할된 영역에 근거하여 계산된 전립선의 크기는 전립선 비대증 등과 같은 진단에 이용된다. 전문가가 직접 영상을 분석하여 전립선의 크기를 구할 수도 있지만, 한 장에 20 분 이상의 시간이 소요되는 힘든 작업이다. 이러한 노고를 줄일 수 있도록, 본 논문에서는 MRI 영상에서 자동으로 전립선을 분할하는 알고리즘을 제안한다.

전립선 분할 기술은 다음과 같은 두가지 문제가 있다. 첫번째는 영상 간의 큰 편차이다. 전립선 MRI 영상은 촬영 시 사용되는 직장 내 코일에 의해 영상이 왜곡되어 큰 편차를 갖는다. 두번째는 훈련 영상의 부족이다. 전립선 MRI 영상은 영상 내에 환자의 정보를 포함하고 있기 때문에, 온라인 상에서 공유하기 위해서는 환자의 동의를 얻고 개인정보를 삭제해야 한다. 따라서 안정적으로 CNN 을 훈련시킬 수 있는 수준의 전립선 MRI 영상의 대규모 데이터베이스가 존재하지 않는다.

본 논문에서는 알지만 효율적인 구조를 갖는 네트워크를 제안하고 과적합을 막기 위한 잡음 계층을 제안한다. 우선, 3D 컨볼루션 계층과 3D 풀링 계층 및 전연결 계층을 이용하여 네트워크를 구축하고 [1,2]에서 제안하는 특징 전달 기법을 적용하여 효과적 학습이 가능한 기본 네트워크를 구축한다. 이에 잡음 계층을 추가함으로써 강인한 훈련이 가능한 네트워크를 제안한다. 실험을 통해 제안 기법이 우수함을 보인다.

2. 제안하는 기법

2.1 3D CNN 기반 네트워크

잡음 계층을 제안하기에 앞서 기본이 되는 모델을 소개한다. 3 차원 필터를 이용하는 3D 컨볼루션 계층과 3D 풀링 계층은 MRI 영상에서 3 차원 입체 정보를 학습하여 강인한 특징 벡터를 추출한다. 이어지는 3D 디컨볼루션 계층(deconvolutional layer)은 컨볼루션의 반대 연산을 수행하는 계층으로, 특징 벡터에 압축되어 있는 전립선 영역에 대한 정보를 실제 영상으로 복원한다. 전연결 계층은 복원된 영상을 많은 수의 파라미터로 정련한다. 그림 1 은 제안하는 알고리즘에서 사용된 네트워크를 도시한다.

2.2 잡음 계층

CNN 알고리즘에서 안정적인 성능을 보장하기 위해서 가장 중요한 것은 네트워크를 강인하게 훈련시키는 것이다. 그 동안 강인한 훈련을 위한 방법으로 수 많은 연구가 이루어 졌고 실험적으로 효과가 증명된 배치 정규화 기법[4]과 드롭 아웃 기법[5]이 널리 이용되었다. 기법의 특성상 정규화 기법은 컨볼루션 계층의 중간에 적용되고 드롭 아웃은 전연결 계층에 적용된다.

본 논문에서는 드롭 아웃과 같이 전연결 계층에 적용할 수 있는 강인한 훈련 기법을 제안한다. 3D 디컨볼루션 계층을 여러번 거쳐 복원된 영상은 많은 파라미터에 의해 목표 영상과 매우 비슷하게 출력된다. 따라서 반복 학습을 거칠 수록 네트워크의 파라미터는 훈련 영상에 과적합되는 문제가 있다.

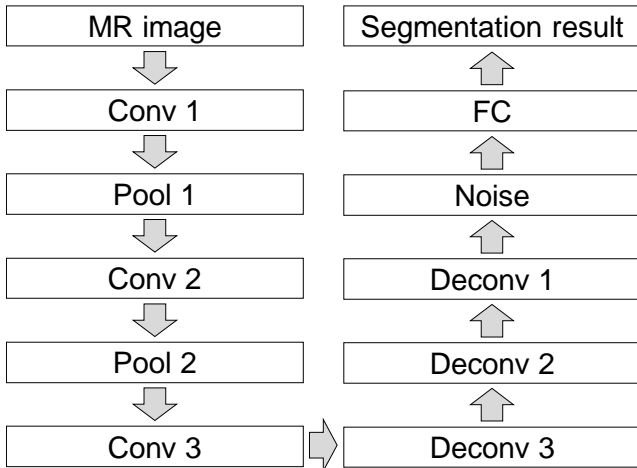


그림 1. 제안하는 전립선 분할 네트워크 구조: Conv 는 3D 컨볼루션 계층 3 개로 이루어진 계층, Pool 은 3D 풀링 계층, Deconv 는 3D 디컨볼루션 계층 1 개와 컨볼루션 계층 2 개로 이루어진 계층, FC 는 전연결 계층, Noise 는 제안하는 잡음 계층을 나타낸다.

제안하는 잡음 계층은 복원된 영상에 가우시안 잡음을 첨가함으로써 랜덤한 오차를 유발하여 네트워크가 훈련용 데이터베이스에 과적합되는 것을 막는다.

2.3 훈련 목표 함수 및 훈련 방법

본 논문에서는 3D MRI 영상의 입체 정보를 효과적으로 학습하기 위해 코사인 유사도를 목표 함수로 이용한다. 코사인 유사도(cosine similarity)는 정보 이론 분야에서 고차원의 정보를 비교할 때 유용하다고 알려져있다[3,6]. 코사인 유사도에 근거하여 분할된 영상과 목표 영상의 유사도를 비교한다. 식 (1)은 코사인 유사도를 나타낸다. N 은 입력 영상의 크기, p_i, q_i 는 각각 분할 결과와 목표 영상 내에서 각 점이 전립선일 확률을 0에서 1 사이의 값으로 나타낸다.

$$S_c = \frac{\sum_{i=1}^N p_i q_i}{\sqrt{\sum_{i=1}^N p_i^2 \sum_{i=1}^N q_i^2}} \quad (1)$$

이를 p_j 에 대해 미분하면 다음과 같다.

$$\frac{\partial S_c}{\partial p_j} = \frac{q_j \sum_{i=1}^N p_i^2 \sum_{i=1}^N q_i^2 - p_j \sum_{i=1}^N p_i q_i \sum_{i=1}^N q_i^2}{(\sum_{i=1}^N p_i^2 \sum_{i=1}^N q_i^2)^{\frac{3}{2}}} \quad (2)$$

제안하는 네트워크를 훈련시키는 방법으로는 Adam 그라디언트 하강(gradient descent) 기법을 이용한다. 식 (2)로 계산되는 목표 함수의 그라디언트를 역전파(backpropagation) 방법으로 네트워크의 파라미터에 전달한다. 그 결과 네트워크의 파라미터는 코사인 유사도를 최대화하는 방향으로 갱신된다.

3. 실험 결과

본 논문의 실험에서는 PROMISE12 데이터베이스[7]를 이용하였다. PROMISE12 데이터베이스는 50 장의 MRI 영상과 각 영상에서의 전립선 영역 정보를 제공한다. 실험에서는 50 장의 영상 중 45 장을 훈련에 사용하고 나머지 5 장을 테스트 영상으로 이용하였다. 표 1 은 정량적 실험 결과를 나타낸다. 정량적 지표는 다이스 유사 계수(Dice similarity coefficient, DSC)를 이용하였다. 다이스 유사 계수는 식 (3)과 같다. 모든 실험은 NVIDIA GTX Titan X GPU 에서 이루어졌다.

$$DSC = \frac{2 \sum_{i=1}^N p_i q_i}{\sum_{i=1}^N (p_i^2 + q_i^2)} \quad (3)$$

표 1. 정량적 성능 비교

Network	DSC	Train time	Iteration
Proposed (base)	0.8648	2 hours	3000
Proposed (noise)	0.8661	2 hours	3000
V-Net[2]	0.5966	20 hours	10000

실험 결과 제안하는 네트워크를 이용한 알고리즘이 짧은 학습 시간에 비하여 높은 성능을 보임을 나타낸다. 비교 알고리즘[2]의 네트워크는 깊이가 매우 깊은 네트워크로 완전한 훈련을 위해서는 100,000 회 이상의 학습을 필요로 하여 1 주일 이상의 시간이 소요된다. 반면 제안하는 네트워크는 단순하고 얇은 구조를 갖기 때문에, 단 3,000 회의 학습만으로 짧은 시간 내에 데이터베이스 학습을 완수한다.

4. 결론

본 논문에서는 MRI 영상에서 전립선의 입체 정보를 학습하고 전립선을 자동으로 분할하는 알고리즘을 제안하였다. 규모가 적고 영상 내 편차가 큰 데이터베이스를 학습하기 위해 알지만 효과적인 네트워크를 제안하였다. 또한 강인한 학습을 돕기위한 잡음 계층을 제안하였다. 실험 결과는 제안하는 기법이 짧은 학습 시간에도 불구하고 높은 정확도를 보여 비교 알고리즘보다 성능이 우수함을 확인하였다.

5. 참고 문헌

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.

[2] F. Milletari, N. Navb, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *MICCAI*, 2016.

[3] J. Mun, W.D. Jang, and C.S. Kim, "Comparison of Objective Functions in CNN-Based Prostate Magnetic Resonance Image Segmentation," in *ICIP*, 2017.

[4] S. Ioffe, and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *ICML*, 2015.

[5] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *JMLR*, Jun. 2014.

[6] A. Singhal, "Modern Information Retrieval: A Brief Overview," in *IEEE Data Eng. Bull.*, 2001.

[7] PROMISE12. [Online]. Available: <https://promise12.grand-challenge.org>.