

V-Net : Fully Convolution NN for Volumetric Medical Image Segmentation

 link <https://arxiv.org/pdf/1606.04797>

V-Net NN 구조를 활용하여 한번에 3D medical img Sementation (Volumetric Segmentation)

3D CNN = V-net?

00 Abstract

- 3D 의료 이미지(MRI) → CNN으로 한 번에
 - 기존 연구는 2D, slice-wise(한 장씩)
- **Dice coefficients**에 기반한 새로운 **objective function** → foreground(관심)/background(배경) 간의 불균형 보완
- **Data Augment** (의료 전문가 annotation 필요 - 데이터 부족)
 - **random non-linear transformation** : 데이터 왜곡
 - **histogram matching** : 명암 등 조절 - 조건의 다양성

01 Introduction and Related Work

도전 과제 : **prostate(전립선) MRI 데이터의 Segmentation**

- Appearance의 다양성 : MRI마다 다른 모양 (변형 deformation 또는 강도 분포 차이 variations of the intensity(밝기) distribution)
- 왜곡 : artefact(장비 오작동, 촬영 중의 변수 - 움직임, 보철물 등), field inhomogeneity (MRI 에서 자기장의 불균등 - intensity 차이 발생)

- 전립선의 정확한 크기/부피 측정 → 진단에 도움
- 전립선과 background의 정확한 경계 파악 → 치료 계획에 도움

CNN

- CV(classification, segmentation, obj detection)
- 계층적(hierarchical)인 표현 처리가 가능하고, hand-crafted feature에 의존하지 않기 때문

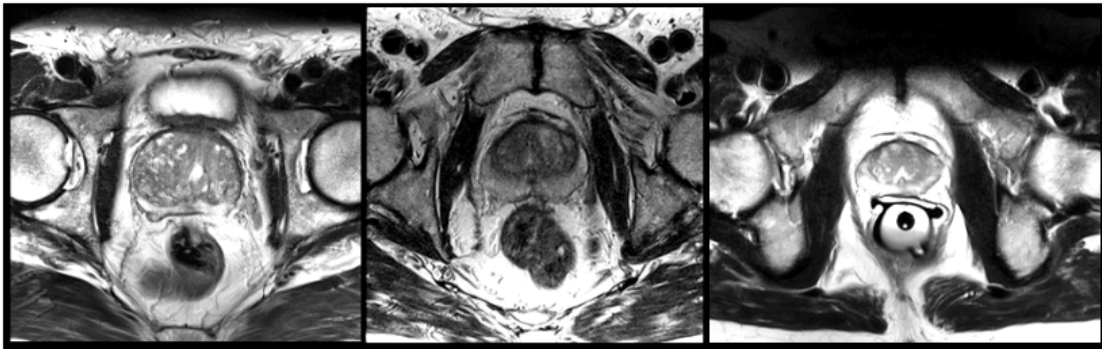


Fig. 1. Slices from MRI volumes depicting prostate. This data is part of the PROMISE2012 challenge dataset [7].

: Slice-wise. MRI, CT - 단층으로 촬영

Segmentation

- 장기 구조 파악, visual augmentation, diagnosis, intervention 등 여러 방면에 사용
- 기존의 CNN-Segmentation = PATCH-WISE
 - local을 다루기 때문에 정확도 낮음 - 다른 기술 결합/새로운 시도로 성능 상승 미미
 - 계산의 비효율성
- 3D 이미지를 다루는 경우가 많다

⇒ Volumetric Segmentation이 큰 도움이 될 것

Fully Convolutional Network Trained end-to-end

- input 크기 상관없이 입력 가능하며, 픽셀 단위의 예측으로 segmentation에 적합
- input-output 과정에서 중간 단계 없이 한 번에 학습. 효율
- (기존 연구) 다양한 구조의 모델로 segmentation mask 예측하도록 학습 - 이미지 전체에서 관심 부분을 구분하는 작업
 - VGG 출력을 deconv → RGB이미지 변환
 - 3 FCN 사용하여 분류로 선행학습된 모델을 segmentation에 맞춰 refine

이 연구에서는..

- fully convolutional neural network (trained end-to-end) MRI volumes 처리에 적합하게 강화
- slice-wise 대신 volumetric convolution 사용
- Dice coefficient maximisation → 새로운 objective function
- 더 높은 정확성 on prostate MRI test volume

02 Method

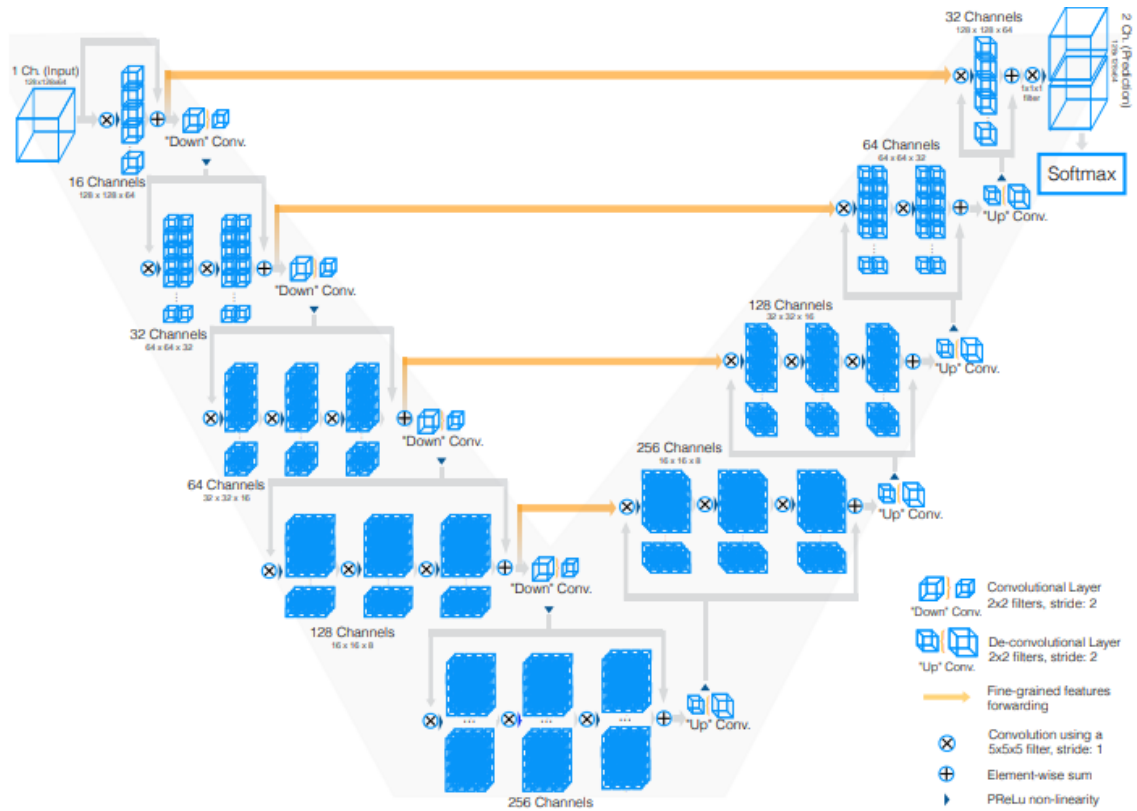
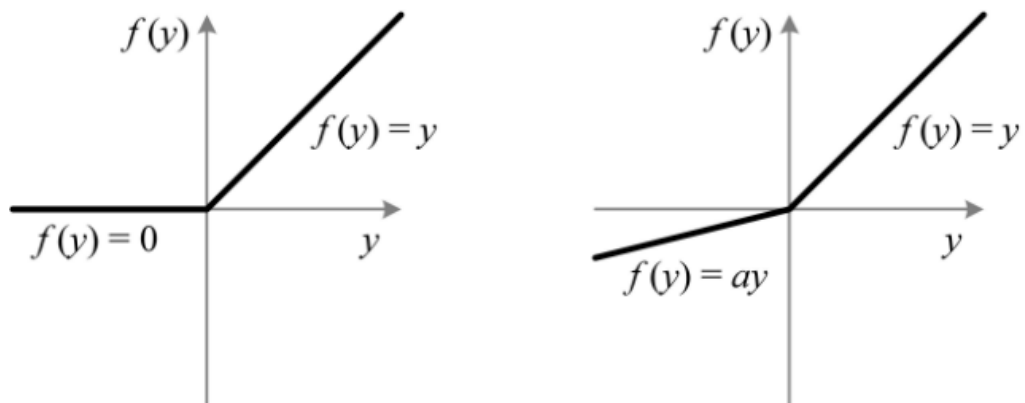


Fig. 2. Schematic representation of our network architecture. Our custom implementation of Caffe [5] processes 3D data by performing volumetric convolutions. Best viewed in electronic format.

- extract feature, reduce resolution(해상도 감소 - 특징 요약) 맞춰 네트워크 구성
- 우측: compression path / 좌측 : decompression path(original size가 될 때까지)
 - 좌측 부분
 - 단계별로 해상도가 줄어듦 - feature map은 두 배로 늘어남
 - 각 스테이지에서 residual function (잔차 함수, output→input 단계에 생기는 변화) 학습
 - (a) conv & 비선형 변환 적용
 - (b) 다음 단계에 더하기 (residual connection)
 - 네트워크가 학습해야 할 것이 residual 뿐 → convergence 빨라져 안정적인 학습이 가능
 - Volumetric kernel (5×5×5) → 특징 추출
 - (compression path, 해상도 축소) Convolution (2×2×2, stride 2) → 데이터 사이즈 절반으로 축소

- max-pooling \Rightarrow convolution으로 대체
 - 학습중의 memory footprint 감소
 - no switches : Pooling layer - 역전파 과정에서 어떤 값이 선택되었는지 기록록하는 switch 저장이 필요함
 - un-pooling(switch 기반의 복원) 보다 de-convolution이 해석적으로 나옴
- PReLU non linearity (오른쪽)



◦ 우측 부분

- 해상도 확장 (데이터를 다시 복원) - De-convolution
 - 낮은 해상도에서 얻은 피쳐 맵을 통해, 공간적 맥락 보충
 - 복원 중, 좌측 네트워크에서 압축한 정보와 새로 추출한 정보를 결합 (horizontal connection, 세밀한 정보 손실을 막는다)
- output two-channel volumetric segmentation : 각각 foreground/background 를 볼륨 기반으로 segmentation
- 마지막 conv($1 \times 1 \times 1$) \rightarrow input과 같은 사이즈의 출력 \rightarrow softmax(foreground/background 각각의 확률로 변환)
- de-convolution이후 입력 사이즈 두 배 \rightarrow 커널은 절반만 사용하도록 함

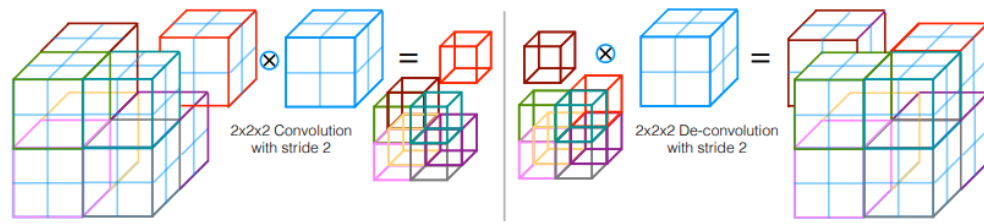


Fig. 3. Convolutions with appropriate stride can be used to reduce the size of the data. Conversely, de-convolutions increase the data size by projecting each input voxel to a bigger region through the kernel.

데이터 축소, 복원 과정

- 가장 내부의 레이어 : 데이터를 한번에 캡처 (Receptive Capture)
 - poorly visible anatomy
 - 전체 파악 → global constraints

Table 1. Theoretical receptive field of the $3 \times 3 \times 3$ convolutional layers of the network.

Layer	Input Size	Receptive Field	Layer	Input Size	Receptive Field
L-Stage 1	128	$5 \times 5 \times 5$	R-Stage 4	16	$476 \times 476 \times 476$
L-Stage 2	64	$22 \times 22 \times 22$	R-Stage 3	32	$528 \times 528 \times 528$
L-Stage 3	32	$72 \times 72 \times 72$	R-Stage 2	64	$546 \times 546 \times 546$
L-Stage 4	16	$172 \times 172 \times 172$	R-Stage 1	128	$551 \times 551 \times 551$
L-Stage 5	8	$372 \times 372 \times 372$	Output	128	$551 \times 551 \times 551$

03 Dice Loss Layer

two-columns → soft-max → 각각 foreground/background에 속할 확률

이 논문의 과제와 같이 매우 국소 부위를 구분하는 문제에서는 local minima에 갇히기 쉬움
 ⇒ 편향이 커지면서 foreground(비교적 적은 부분)을 missing 하거나 partially detected 할 가능성이 커짐

이 fore/back 사이 불균형 문제 완화를 위해

Dice coefficient D

$$D = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}$$

- 두 값 간의 유사도를 측정
 - 0(다름) ~ 1(같음) 사이의 값
- ⇒ dice coefficient 최대화 (완전히 같음 상태인 1로 만들기, pred = ground truth 하는 방향으로)

gradient 계산을 위해 이 식을 미분하면

$$\frac{\partial D}{\partial p_j} = 2 \left[\frac{g_j \left(\sum_i^N p_i^2 + \sum_i^N g_i^2 \right) - 2p_j \left(\sum_i^N p_i g_i \right)}{\left(\sum_i^N p_i^2 + \sum_i^N g_i^2 \right)^2} \right]$$

불균형 해결을 위해 따로 가중치를 지정할 필요가 없음

불균형 문제에 효과적 (fore<<<back 인 경우)

- multinomial logistic loss을 사용했을 때보다 좋은 성능을 보임

03-1 Training

- end-to-end
- PROMISE2012 challenge dataset : prostate scans in MRI (위의 Figure 1 참고)
 - annotated medical volumes 얻기 힘들다 - 전문가의 수작업 필요
 - robustness와 precision 위해 **데이터 증강**이 필요

- deformation : $2 \times 2 \times 2$ control-point grid, B-spline interpolation 활용
 - on-the-fly(즉시 생성) : 저장공간 절약 위함. iteration 전에 바로 생성
 - Histogram Matching : intensity distribution을 다양하게 조정
- ⇒ 데이터의 수와, 다양성을 높임
- $128 \times 128 \times 64$ voxel, $1 \times 1 \times 1.5$ mm resolution

03-2 Testing

- unseen MRI volume - feed forward manner
- 확률 0.5 기준으로 foreground/background 분류

04 Results

PROMISE2012 Challenge Dataset

- train : 50 prostate MRI volumes
- test : 30 prostate MRI volumes (ground truth 없음)
- 다양한 병원 - 임상적 다양성을 가지도록 함
- env
 - Python & Caffe custom version
 - Intel i7-5820K CPU, NVIDIA GTX 1080 GPU, 64GB mem
 - 약 48시간(30,000 iteration), test 1

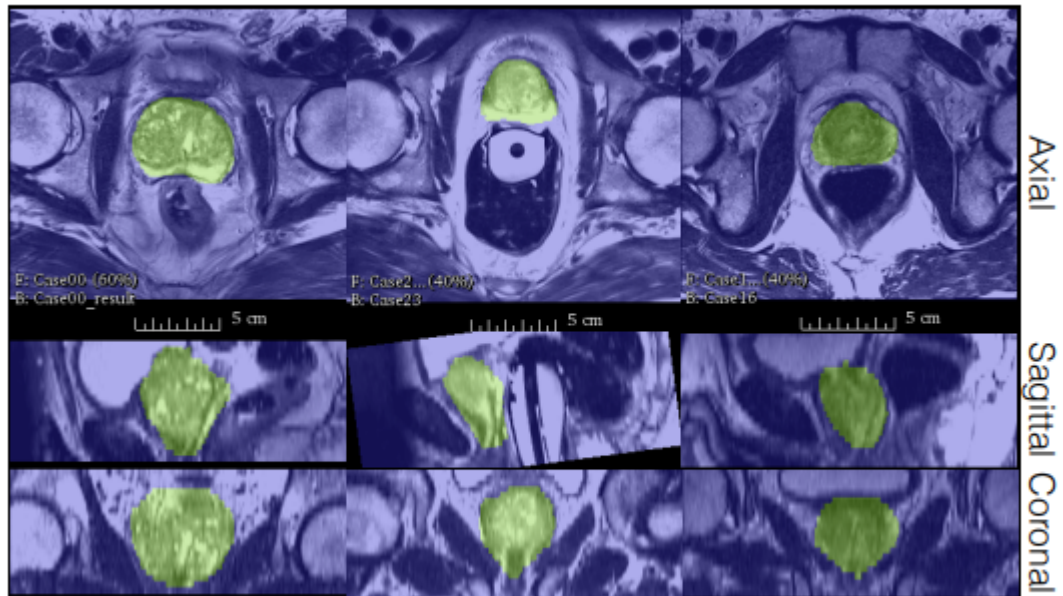


Fig. 4. Qualitative results on the PROMISE 2012 dataset [7].

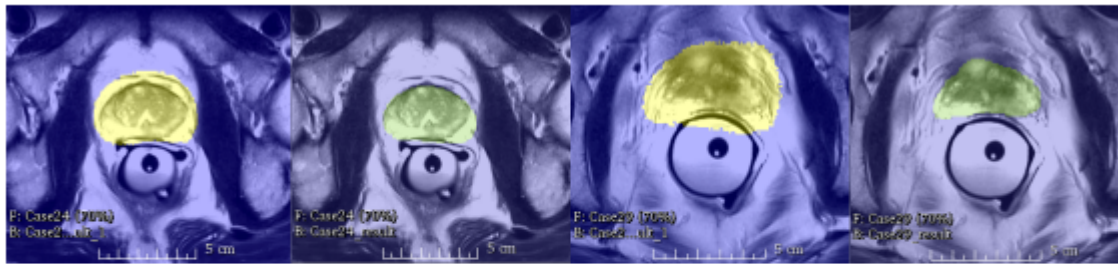


Fig. 6. Qualitative comparison between the results obtained using the Dice coefficient based loss (green) and re-weighted soft-max with loss (yellow).

- Dice-based Loss(녹색)가 Re-weighted softmax loss(노란색)보다 더 정확히 segmentation
- 전립선의 다양한 크기와 형태에 대해서도 안정적인 성능

Table 2. Quantitative comparison between the proposed approach and the current best results on the PROMISE 2012 challenge dataset.

Algorithm	Avg. Dice	Avg. Hausdorff distance	Score on challenge task
V-Net + Dice-based loss	0.869 ± 0.033	5.71 ± 1.20 mm	82.39
V-Net + mult. logistic loss	0.739 ± 0.088	10.55 ± 5.38 mm	63.30
Imorphics [18]	0.879 ± 0.044	5.935 ± 2.14 mm	84.36
ScrAutoProstate	0.874 ± 0.036	5.58 ± 1.49 mm	83.49
SBIA	0.835 ± 0.055	7.73 ± 2.68 mm	78.33
Grislies	0.834 ± 0.082	7.90 ± 3.82 mm	77.55

지표

- Dice coefficient : 예측과 실제값의 유사도
- Hausdorff distance : 예측과 실제 경계 간 최대 거리
- Challenge Score : PROMISE2012 점수
- Multinomial Logistic Loss 사용 시 Dice 값이 크게 낮아짐(0.739)
- Morphics(0.879 Dice)와 비슷한 성능, Challenge Score는 Morphics(84.36)보다 약간 낮음

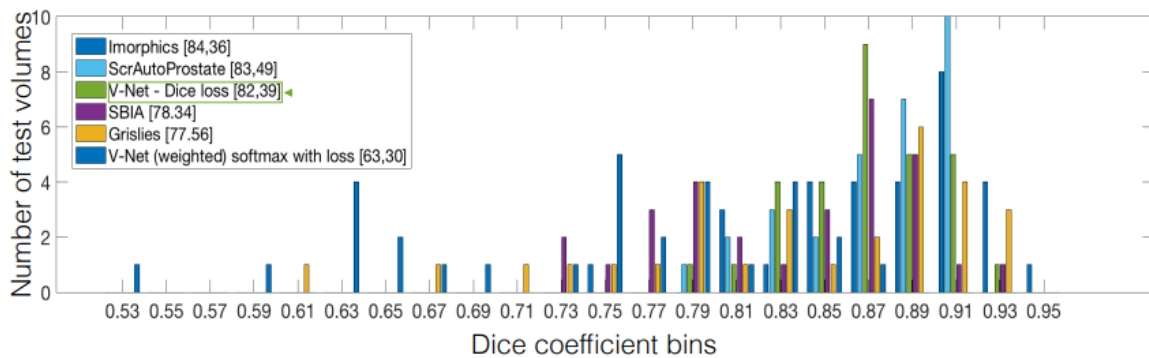


Fig. 5. Distribution of volumes with respect to the Dice coefficient achieved during segmentation.

대부분 테스트 볼륨에서 Dice 값이 0.85 이상으로 높다

05 Conclusion

- MRI prostate volumes segmentation → volumetric convolutional NN
- Dice loss layer : 데이터 불균형 - 가중치 재지정 없이 좋은 성능
- state 분류, residual function을 학습하게 하여 효율을 높임