

2019.08.01

# **Review of segmentation Network Architecture**

Youngtaek Hong, PhD

---

# Cascaded FCN - Automatic Liver and Tumor Segmentation of CT and MRI Volumes Using Cascaded Fully Convolutional Neural Networks

---

# Automatic Liver and Tumor Segmentation of CT and MRI Volumes Using Cascaded Fully Convolutional Neural Networks

Patrick Ferdinand Christ<sup>a,1</sup>, Florian Ettlinger<sup>a,1</sup>, Felix Grün<sup>a</sup>, Mohamed Ezzeldin A. Elshaer<sup>a</sup>, Jana Lipková<sup>a</sup>, Sebastian Schlecht<sup>a</sup>, Freba Ahmaddy<sup>a</sup>, Sunil Tataavarty<sup>a</sup>, Marc Bickel<sup>a</sup>, Patrick Bilic<sup>a</sup>, Markus Rempfler<sup>a</sup>, Felix Hofmann<sup>b</sup>, Melvin D'Anastasi<sup>b</sup>, Seyed-Ahmad Ahmadi<sup>b</sup>, Georgios Kaissis<sup>a</sup>, Julian Holch<sup>b</sup>, Wieland Sommer<sup>b</sup>, Rickmer Braren<sup>a</sup>, Volker Heinemann<sup>b</sup>, Bjoern Menze<sup>a</sup>

<sup>a</sup>*Technical University of Munich, Arcisstrasse 21, 80333 Munich*

<sup>b</sup>*LMU Hospital Grosshadern, Marchioninistrasse 15, 81377 Munich, Germany*

2017.02.23에 발표된 후 국제 의료영상 심포지움인 **ISBI 2017**에 accept된

**Cascaded FCN**에 대해 리뷰하도록 하겠습니다.

# Abstract

- 간(liver)과 간에 위치한 병변(hepatic lesions)을 추출(segmentation)하는 것은 정확한 임상 진단과 컴퓨터의 도움을 받는 결정지원시스템(computer-aided decision support systems)을 완성하는 것의 중요한 발판이 됩니다.
- 이 논문에서는 cascaded fully convolutional neural networks (CFCNs)을 통해 자동적으로 CT와 MRI의 간과 병변을 추출(segment)해주는 방법을 제시합니다.
  1. 두 개의 FCNs으로 이루어진 CFCNs는 우선 두번째 FCN의 ROI input으로 쓰일 간을 추출하는 FCN을 학습합니다.
  2. 두번째 FCN은 오직 첫번째 FCN에서 예측한 liver ROIs 내에서만 병변을 segment합니다.
- 검증(validation) 결과로 각 volume당 100초 미만의 시간 안에 간(liver)에 대한 Dice score가 94% 이상을 달성했습니다.

# Image intensity distribution

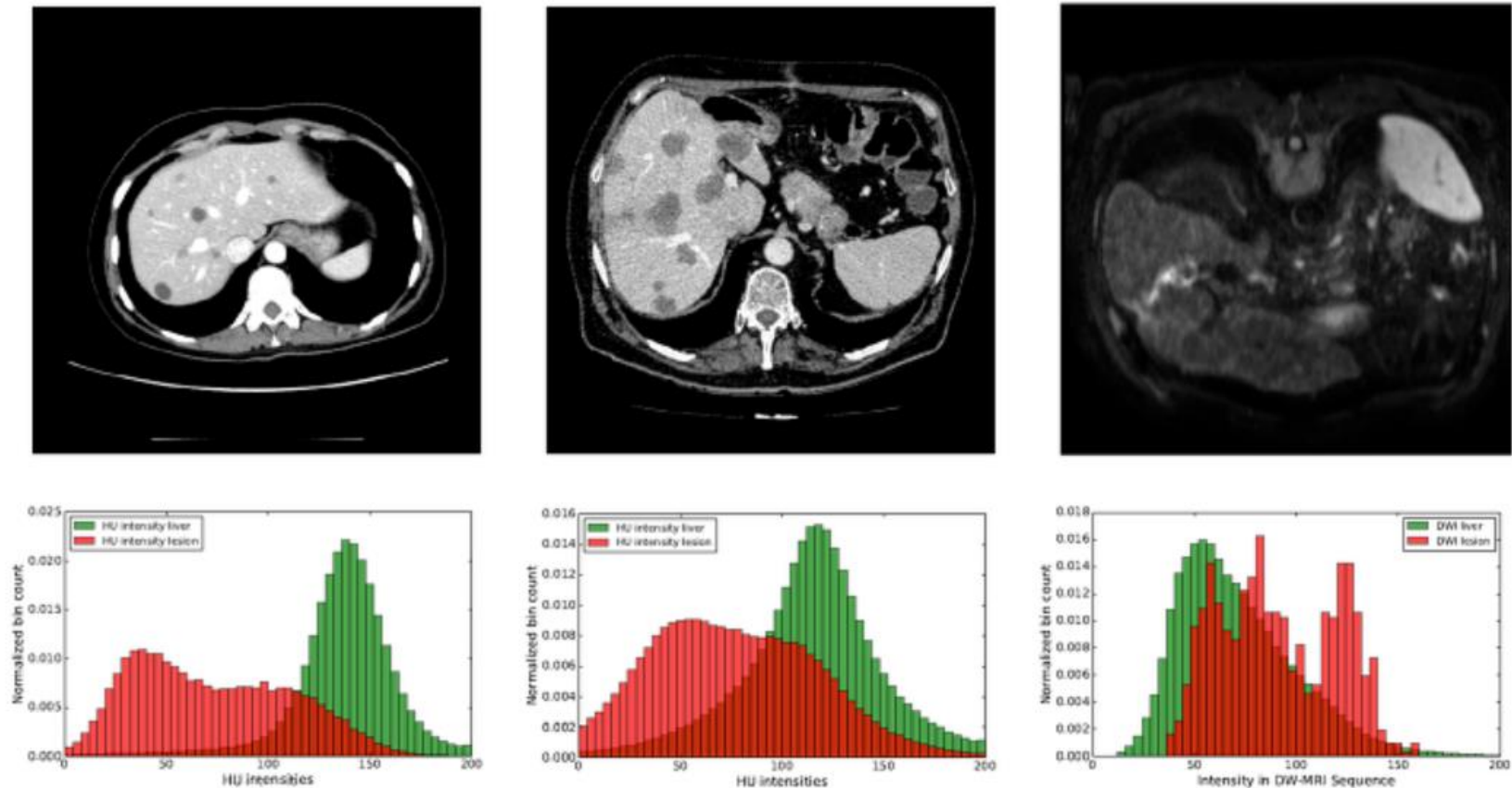


Figure 1: Liver and liver lesions slices in CT and diffusion weighted DW-MRI as well as the corresponding histogram for liver and lesions pixels in the respective modality. The shape, size and level of contrast vary for different lesions. As the histograms indicate, there is a significant overlap between liver and lesion intensities, leading to a low overall contrast.

# Abstract

- Figure 1에서는 두 가지 문제를 확인할 수 있습니다.
  1. 병변의 모양, 크기 그리고 대조의 정도가 다양합니다.
  2. 간과 병변 사이의 값(intensities)의 차이가 크지 않아서 전체적으로 낮은 대조를 갖습니다.
- 이 논문에서는 Figure 1에서 확인한 두 가지 문제점을 갖고 있는 segmentation task에서 좋은 성능을 보이는 모델을 제시합니다. 논문의 세 가지 contribution은 아래와 같습니다.
  1. 간의 병변(hepatic liver lesions)을 추출(segment)하는 난이도 높은 task에 좋은 성능을 보이는 fully convolutional CNN을 학습하고 적용했습니다.
  2. 상당히 높은 성능을 보이는 cascaded fully convolutional neural network (CFCN)을 제시합니다. CFCN은 CT 단면에 대해서 간과 병변을 연속적으로 segment합니다.
  3. 다른 양식(modality)을 가진 데이터와 다양한 실생활 데이터셋에 대해서도 좋은 성능을 보입니다 (generalization).

# Abstract

- 우선 필요없는 장기와 물체를 제거하기 위해 [-100, 400] 사이의 값을 갖도록 합니다. (-100 미만의 값은 모두 -100으로, 400을 초과하는 값은 모두 400으로 / CT가 아닌 MRI의 경우, HU Windowing 대신에 N4 Bias Correction)
- 이후 histogram equalization을 통해 대조(contract)를 높여줍니다. 이를 통해 이전보다 더 높은 대조를 갖게 되므로 비정상 간 조직(abnormal liver tissue)을 효과적으로 구분할 수 있습니다.
- elastic deformation, traslation, rotation, Gaussian noise와 같은 data augmentation을 해줍니다.

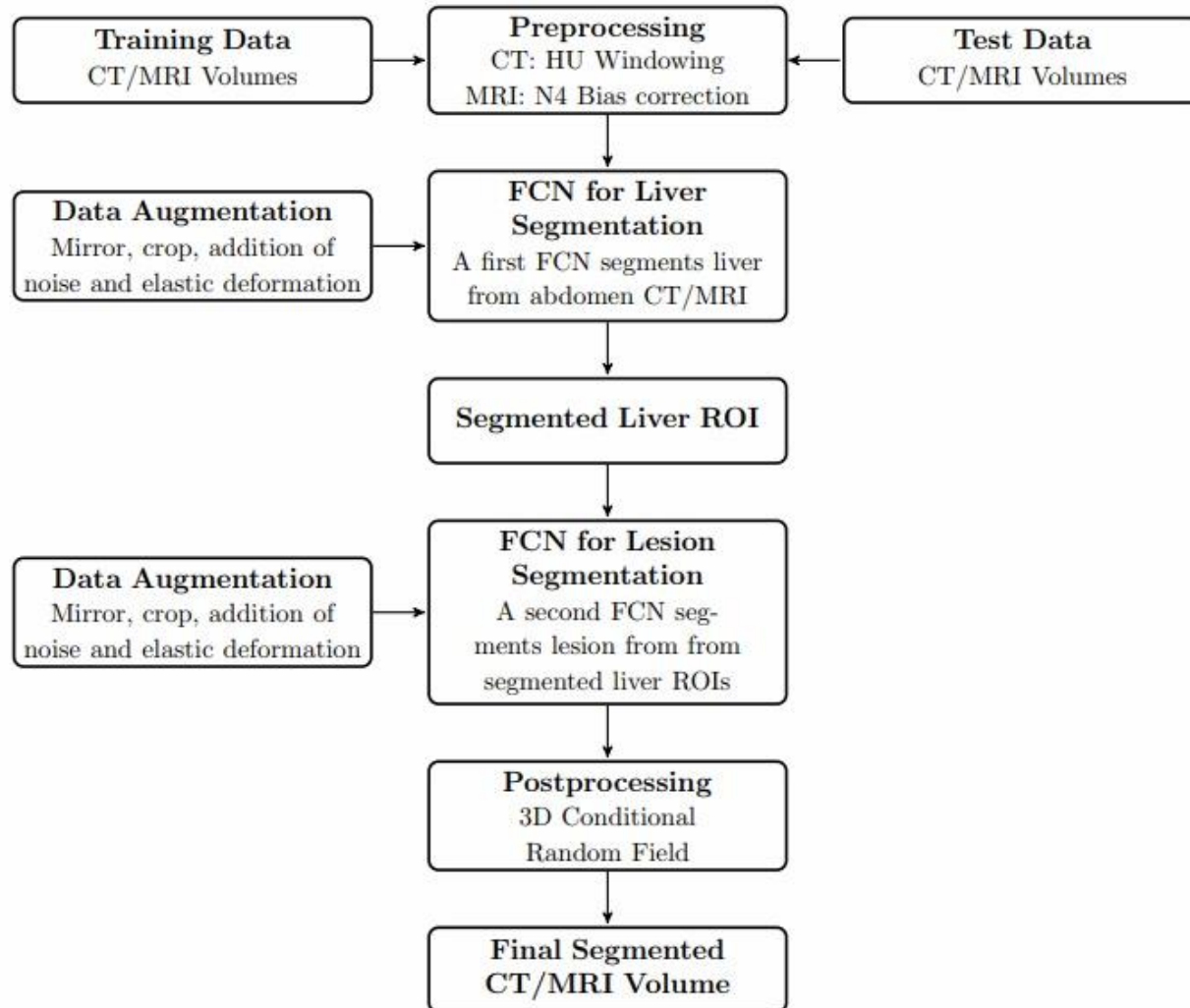
# Data preprocessing



Figure 3: Overview of the applied preprocessing steps. The raw CT slices (left) are windowed to a Hounsfield Unit range of -100 to 400 HU to neglect organs and tissues that are not of interest. The HU-windowed slice (middle) is further processed using a histogram equalization to allow further contrast enhancement of abnormal tissue (right).



# Flow Chart



# Two cascade U-Net

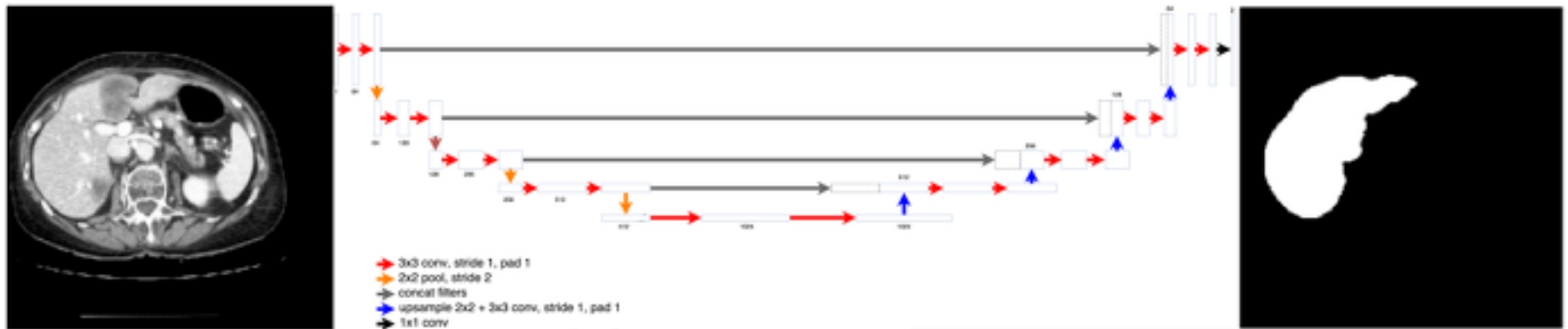


Figure 5: Step 1 of Cascaded FCN: The first U-Net learns to segment livers from a CT slice.

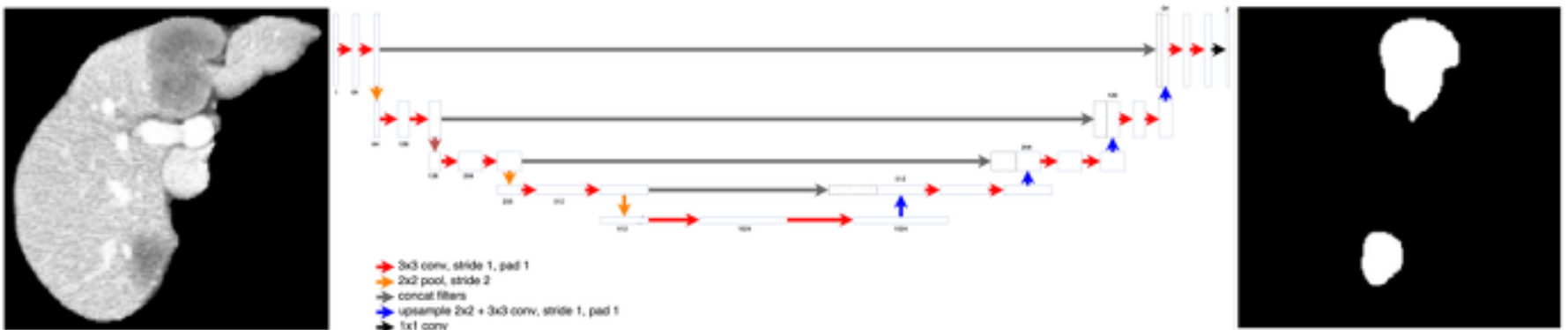


Figure 6: Step 2 of Cascaded FCN: The second U-Net learns to segment lesions from a liver segmentation mask segmented in step 1 of the cascade

# Effect of class balancing

- FCN을 학습할 때, 각각의 class의 pixel-wise frequency에 따라 적절한 class balancing을 해주는 것이 중요합니다.
- 이 논문에서는 class balancing을 거친 결과와 거치지 않은 결과를 비교함으로써 class balancing의 중요성을 강조합니다. CFCN에서는 일반적인 cross entropy loss function에 weighting factor 를 추가해줌으로써 class balancing을 해결했습니다.

$$L = -\frac{1}{n} \sum_{i=1}^N \omega_i^{class} \left[ \hat{P}_i \log P_i + (1 - \hat{P}_i) \log(1 - P_i) \right]$$

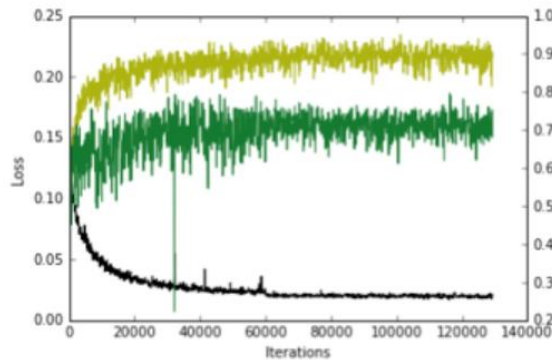
<Term Explanation>

※  $P_i$  = voxel i가 foreground에 속할 확률

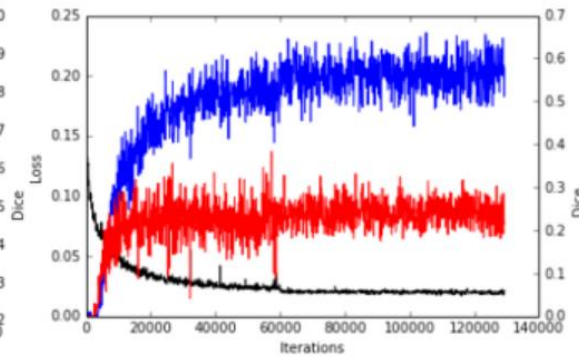
$\hat{P}_i$  = ground truth

$$w_i^{class} = \begin{cases} \frac{\sum_i 1 - \hat{P}_i}{\sum_i \hat{P}_i} & (\hat{P}_i = 1) \\ 1 & \text{otherwise} \end{cases}$$

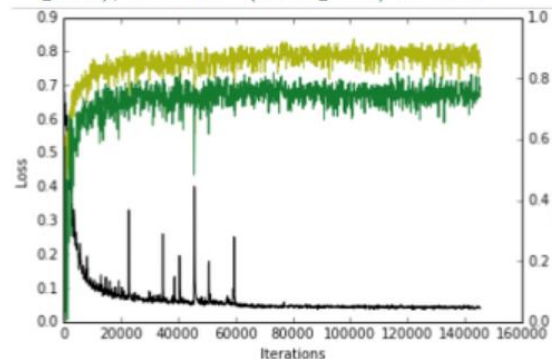
# Test of class weight



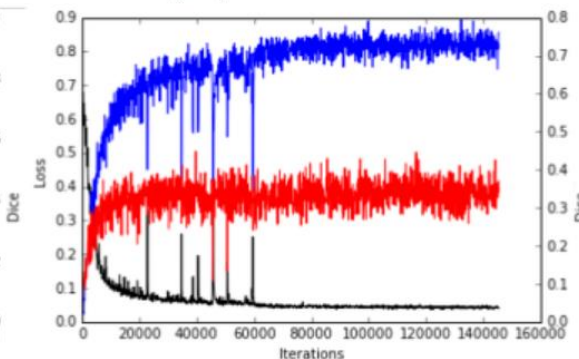
(a) AlexFCN architecture without class balancing: Loss (black), Training Dice (light green), Test Dice (dark green) of Liver



(b) AlexFCN architecture without class balancing: Loss (black), Training Dice (blue), Test Dice (red) of Lesion

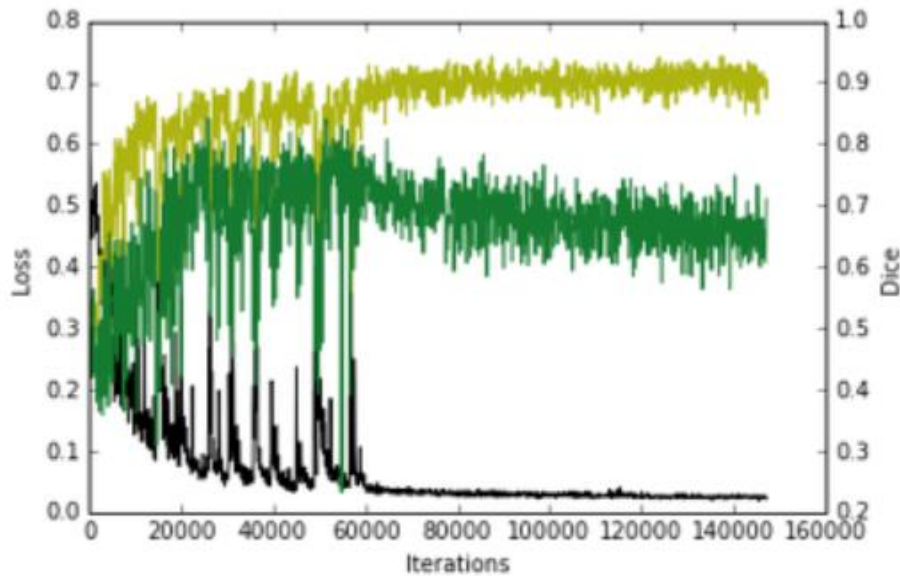


(c) AlexFCN architecture with class balancing: Loss (black), Training Dice (light green), Test Dice (dark green) of Liver

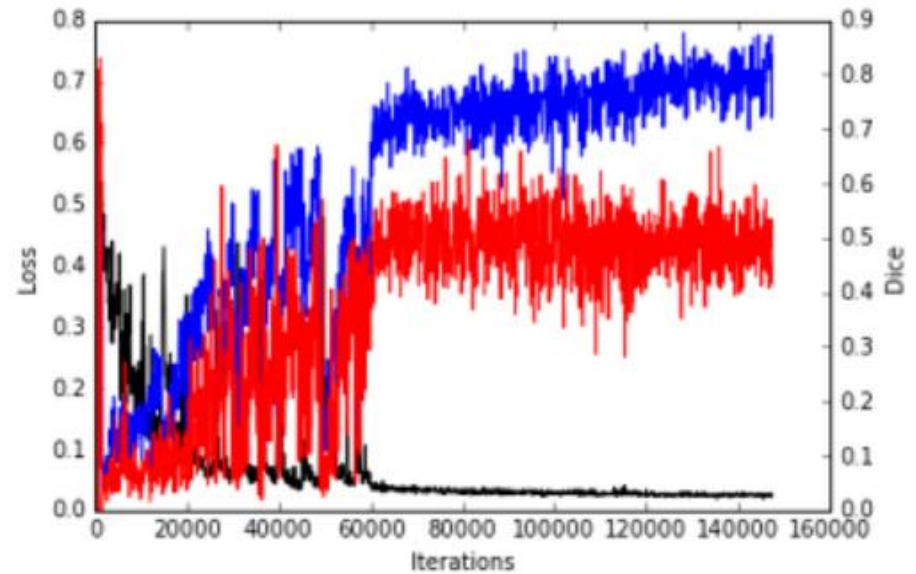


(d) AlexFCN architecture with class balancing: Loss (black), Training Dice (blue), Test Dice (red) of Lesion

# U-Net training result



(e) U-Net architecture with class balancing: Loss (black), Training Dice (light green), Test Dice (dark green) of Liver



(f) U-Net architecture with class balancing: Loss (black), Training Dice (blue), Test Dice (red) of Lesion



# Result figures

A) CT Slice

B) Single UNet

C) CFCN

초록색 영역 : 정확하게 간을 예측한 영역  
노란색 영역 : 간을 잘못 예측한 영역 (false negative + false positive)  
파란색 영역 : 정확하게 병변을 예측한 영역  
빨간색 영역 : 병변을 잘못 예측한 영역 (false negative + false positive)

D) CT Slice

E) CFCN

F) CFCN + 3DCRF

# Result table

Approach	Dataset	VOE [%]	RVD [%]	ASD [mm]	MSD [mm]	DICE [%]
U-Net as in [18]	3DIRCAD	39	87	19.4	119	72.9
Cascaded U-Net	3DIRCAD	12.8	-3.3	2.3	46.7	93.1
Cascaded U-Net + 3D CRF	3DIRCAD	10.7	-1.4	1.5	24.0	94.3
Li et al. [5] (liver-only)	3DIRCAD	9.2	-11.2	1.6	28.2	
Chartrand et al. [34] (semi-automatic)	3DIRCAD	6.8	1.7	1.6	24	
Li et al. [6] (liver-only)	3DIRCAD					94.5
Cohen et al. [35] (liver-only)	Own Clinical CT					89
Cascaded U-Net	MR-DWI	23	14	5.2	135.3	87
Cascaded U-Net	Clinical CT	22	-3	9.5	165.7	88
Cascaded U-Net + 3D CRF	Clinical CT	16	-6	5.3	48.3	91

Table 1: Quantitative segmentation results of the liver on the 3DIRCADb dataset and other clinical CT and MR-DWI datasets. Scores are reported as presented in the original papers.

# Conclusion

- CFCN은 기존의 segmentation이 갖고 있던 치명적인 문제점인 false positive lesions segmentation을 해결하려 노력한 모델입니다.
- 병변을 segment하기 이전에 병변이 포함된 조직(ex - 간)을 우선적으로 segment한 후 얻어낸 특정 영역에 한해서만 예측이 이루어지도록 하는 방법을 통해 성능을 향상시켰습니다.
- End-to-end로 학습되지 않는다는 한계가 있지만 메모리 소모가 큰 3D 이미지가 아닌 2D 단면을 가지고도 비교적 높은 성능을 보였다는 장점이 있습니다.
- CFCN과 3D volume 기반의 segmentation 기법을 합치는 방법을 통해 더욱 성능을 높일 수 있을 것으로 기대됩니다.



# M-Net: A Convolutional Neural Network for Deep Brain Structure Segmentation

---

## **M-NET: A Convolutional Neural Network for Deep Brain Structure Segmentation**

by

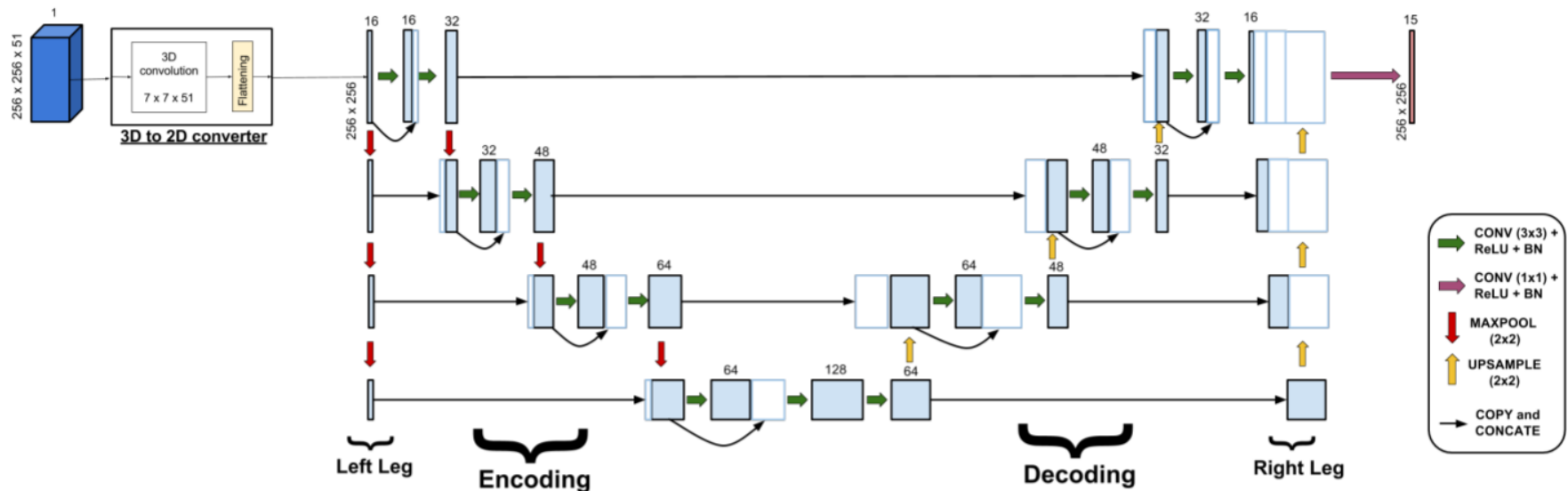
Raghav Mehta, Jayanthi Sivaswamy

in

*IEEE International Symposium on Biomedical Imaging 2017  
(ISBI-2017)*

Melbourne, Australia

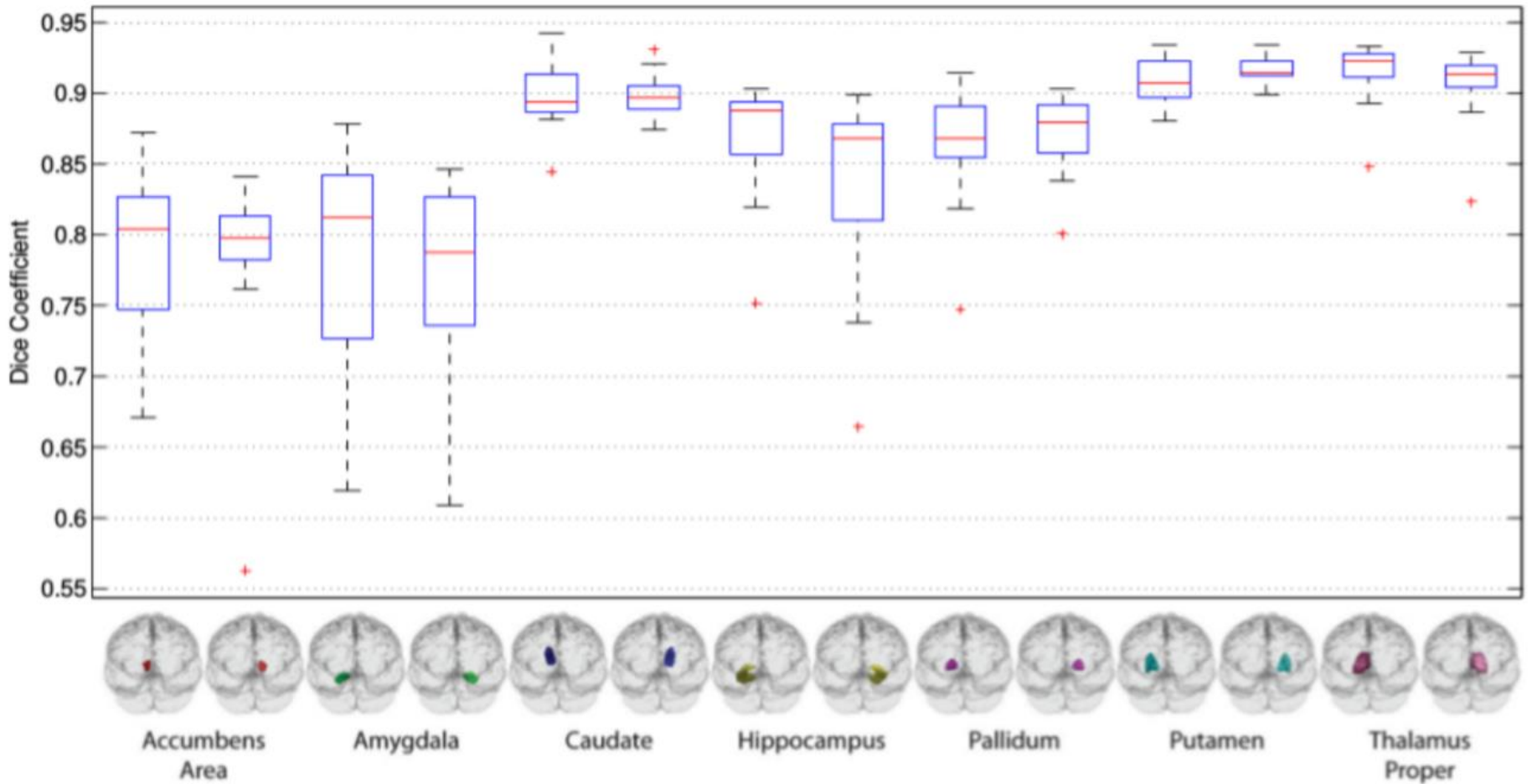
# Abstract

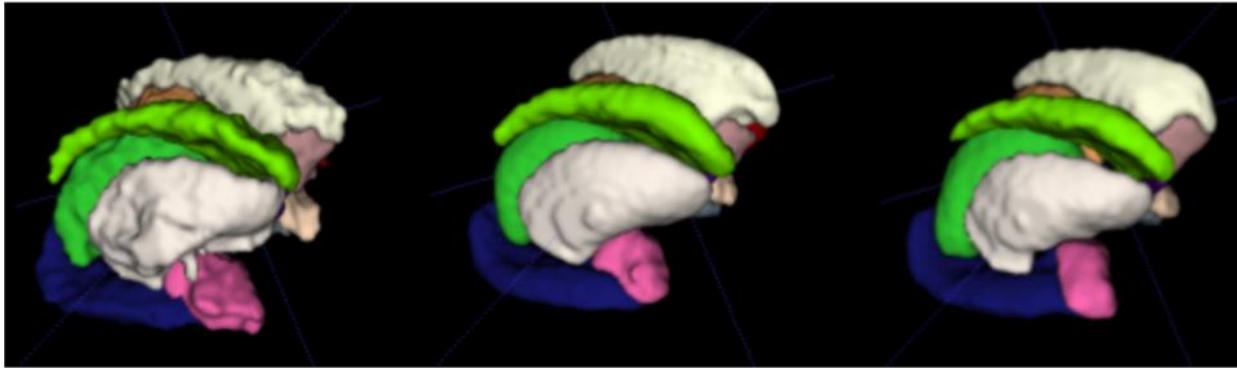


**Fig. 1:** Schematic representation of the *M-net* CNN architecture. Solid blue boxes represent multi-channel feature maps. Blue framed boxes represent copied feature maps. Number of channels is denoted on the top of the box.

**Table 1:** Quantitative comparison of performance on the IBSR dataset. Reported Dice coefficient values for a structure are averaged over the values for the 2 hemispheres.

	Freesurfer	FSL-FIRST	RF + MRF	FCN + MRF	MS-CNN + RW	<i>M-net</i>
Amygdala	0.69	0.70	0.62	0.64	0.67	<b>0.73</b>
Caudate	0.82	0.83	0.78	0.78	<b>0.87</b>	<b>0.87</b>
Hippocampus	0.77	0.81	0.59	0.71	<b>0.82</b>	<b>0.82</b>
Pallidum	0.71	0.76	0.62	0.75	0.80	<b>0.82</b>
Putamen	0.81	0.84	0.77	0.83	0.88	<b>0.90</b>
Thalamus	0.86	0.88	0.80	0.87	<b>0.90</b>	<b>0.90</b>
Accumbens Area	0.69	0.73	0.60	0.63	0.69	<b>0.75</b>
<b>Overall</b>	0.76	0.79	0.69	0.75	0.80	<b>0.83</b>





**Fig. 4:** Qualitative comparison of segmentation results, using 3D rendering, for the Diencephalon dataset. From left to right: Freesurfer, FSL-FIRST and *M-net*

# Automatic Segmentation of LV and RV in Cardiac MRI

---

## Automatic Segmentation of LV and RV in Cardiac MRI

Yeonggul Jang<sup>1</sup>, Yoonmi Hong<sup>2(✉)</sup>, Seongmin Ha<sup>2</sup>, Sekeun Kim<sup>2</sup>,  
and Hyuk-Jae Chang<sup>2,3</sup>

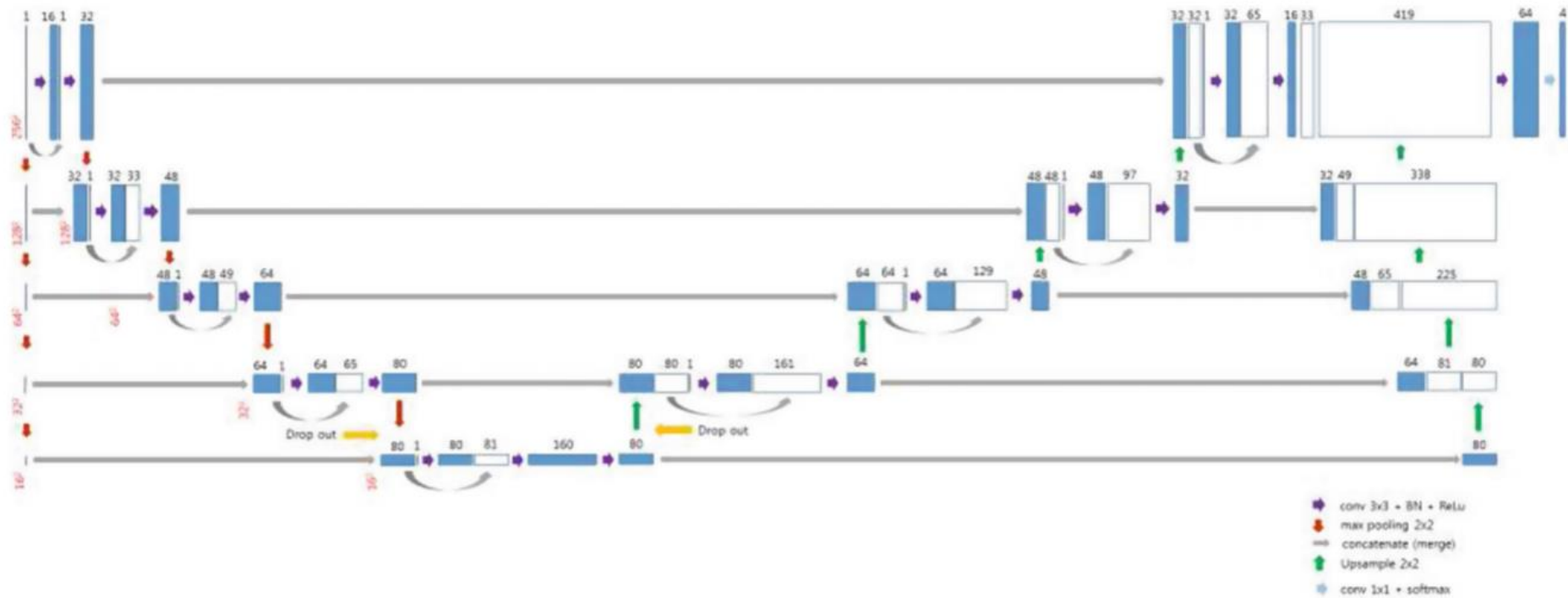
<sup>1</sup> Brain Korea 21 PLUS Project for Medical Science,  
Yonsei University, Seoul, South Korea  
jygl722@gmail.com

<sup>2</sup> Integrative Cardiovascular Imaging Research Center,  
Yonsei University College of Medicine, Seoul, South Korea  
yoonmhong@gmail.com

<sup>3</sup> Division of Cardiology, Severance Cardiovascular Hospital,  
Yonsei University College of Medicine, Seoul, South Korea



# Abstract

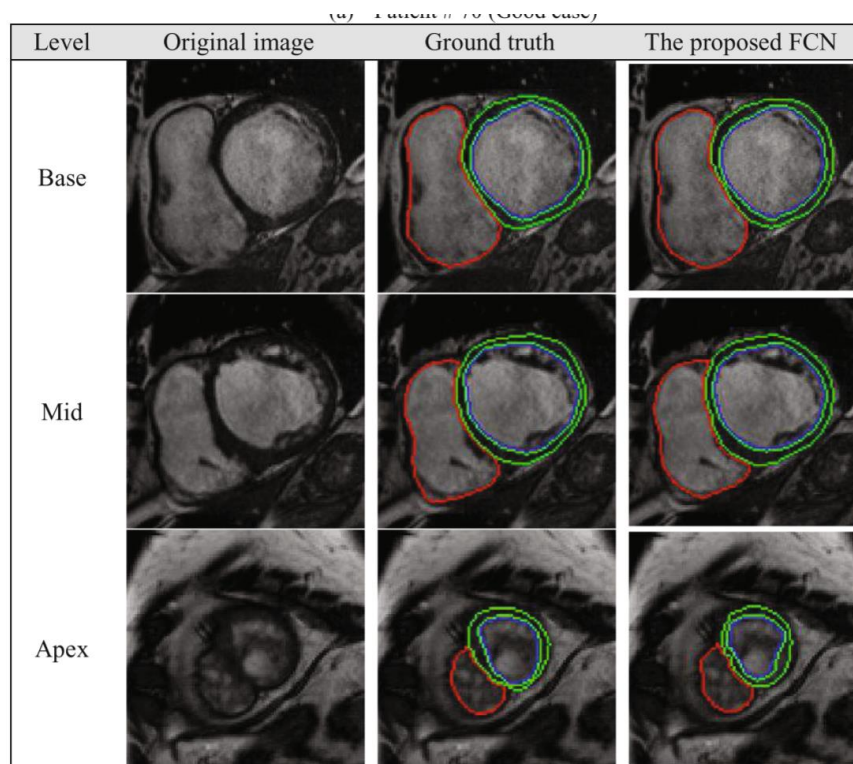
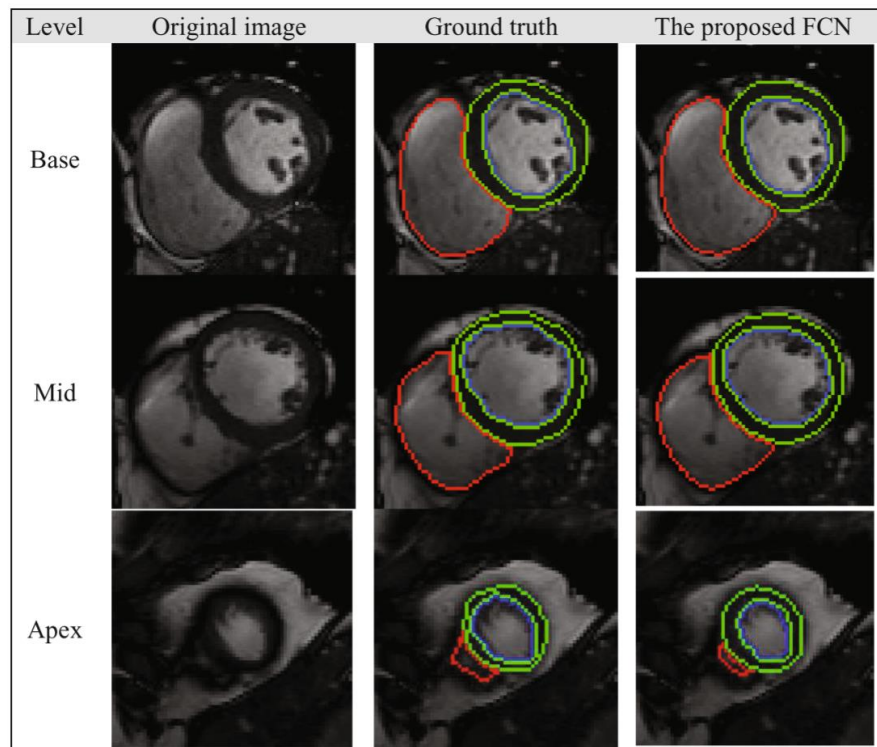


# Results

**Table 2.** Cross-validation (CV) results of our model on the 100 cases (training datasets are 80, test datasets are 20). Values correspond to the mean and standard deviation.

		RV		LV-Myo		LV	
		ED	ES	ED	ES	ED	ES
DSC	CV#1	0.934 ± 0.03	0.837 ± 0.09	0.880 ± 0.02	0.903 ± 0.02	0.960 ± 0.02	0.924 ± 0.05
	CV#2	0.924 ± 0.03	0.853 ± 0.07	0.866 ± 0.03	0.882 ± 0.04	0.959 ± 0.01	0.919 ± 0.04
	CV#3	0.917 ± 0.03	0.851 ± 0.10	0.861 ± 0.05	0.884 ± 0.04	0.956 ± 0.02	0.898 ± 0.09
	CV#4	0.918 ± 0.06	0.843 ± 0.09	0.853 ± 0.04	0.873 ± 0.04	0.958 ± 0.02	0.911 ± 0.08
	CV#5	0.932 ± 0.04	0.891 ± 0.04	0.892 ± 0.03	0.900 ± 0.03	0.970 ± 0.02	0.920 ± 0.05
	Avg	0.925 ± 0.04	0.855 ± 0.08	0.870 ± 0.04	0.888 ± 0.04	0.961 ± 0.02	0.915 ± 0.06
	Total	0.890 ± 0.07		0.879 ± 0.04		0.938 ± 0.05	
HD (mm)	CV#1	12.49 ± 5.87	15.54 ± 5.05	9.23 ± 5.93	8.68 ± 3.91	7.29 ± 4.36	7.15 ± 4.27
	CV#2	14.63 ± 5.55	16.20 ± 5.89	10.88 ± 9.02	11.86 ± 6.78	6.94 ± 4.50	9.52 ± 4.10
	CV#3	15.61 ± 7.62	15.34 ± 5.57	9.52 ± 5.27	10.49 ± 3.80	8.73 ± 5.66	8.78 ± 4.69
	CV#4	15.30 ± 7.60	16.98 ± 8.80	9.68 ± 6.21	11.83 ± 6.28	7.49 ± 8.29	9.33 ± 5.94
	CV#5	11.82 ± 4.91	14.64 ± 7.49	7.84 ± 6.50	9.22 ± 5.36	3.99 ± 2.83	8.90 ± 6.55
	Avg	12.60 ± 6.02	14.78 ± 6.36	9.47 ± 6.64	10.05 ± 5.27	6.27 ± 4.74	8.28 ± 4.67
	Total	13.69 ± 6.30		9.76 ± 6.02		7.27 ± 4.83	

# Result figures



# 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation

---

## 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation

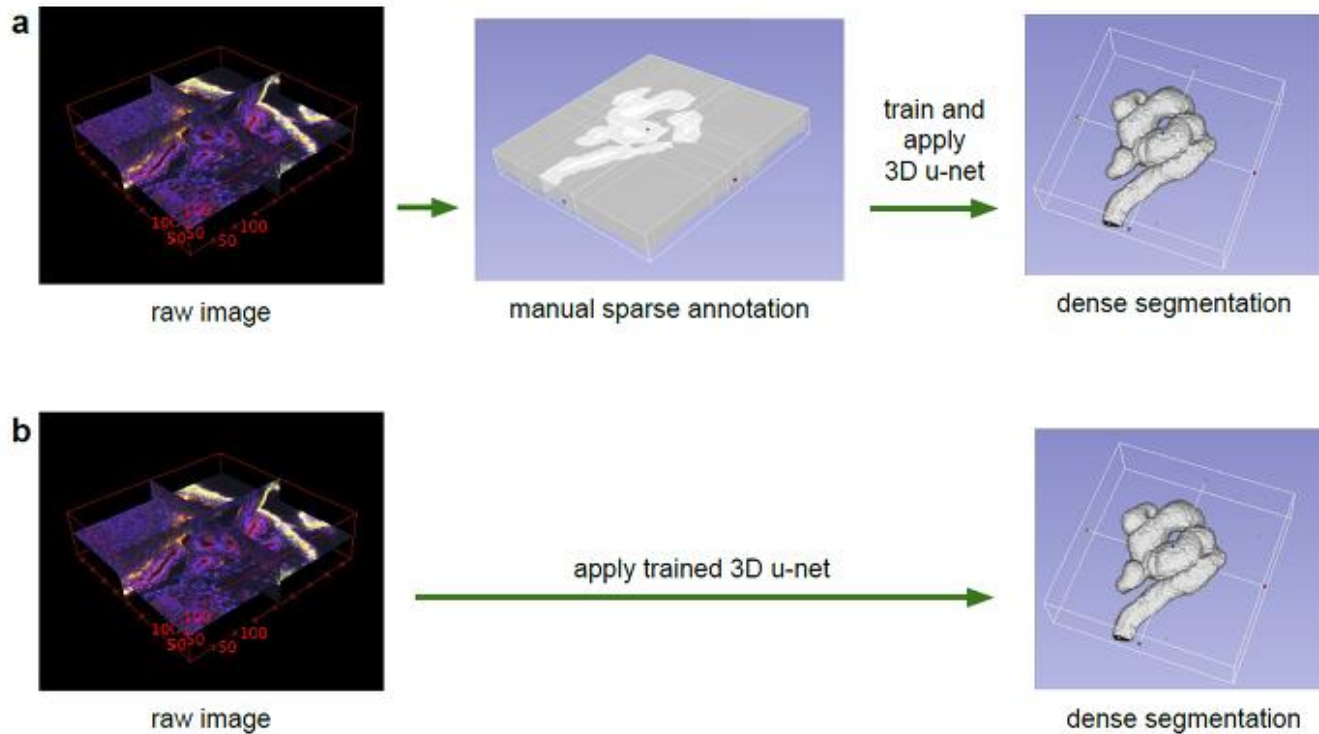
Özgün Çiçek<sup>1,2</sup>, Ahmed Abdulkadir<sup>1,4</sup>, Soeren S. Lienkamp<sup>2,3</sup>, Thomas Brox<sup>1,2</sup>, and Olaf Ronneberger<sup>1,2,5</sup>

# Abstract

- 이 논문에서는 띄엄띄엄(sparse) 레이블링된 입체 이미지(sparsely annotated volumetric images)로부터 학습하는 3D segmentation 모델을 제시합니다.
- 여기서 띄엄띄엄 레이블링된 입체 이미지가 강조된 것은 segmentation task의 ground truth는 매우 만들기 어렵고 원하는 데이터를 얻기 어려우므로 비교적 적은 수의 데이터만 가지고 학습할 수 있기 때문입니다.
- 이전에 제시된 u-net에서 2D 연산을 3D로 확장시킨 네트워크입니다.
- 학습 과정에서 효율적인 data augmentation을 위해 elastic deformation을 가합니다.
- End-to-end로 학습될 뿐만 아니라 pre-trained network이 요구되지 않습니다.

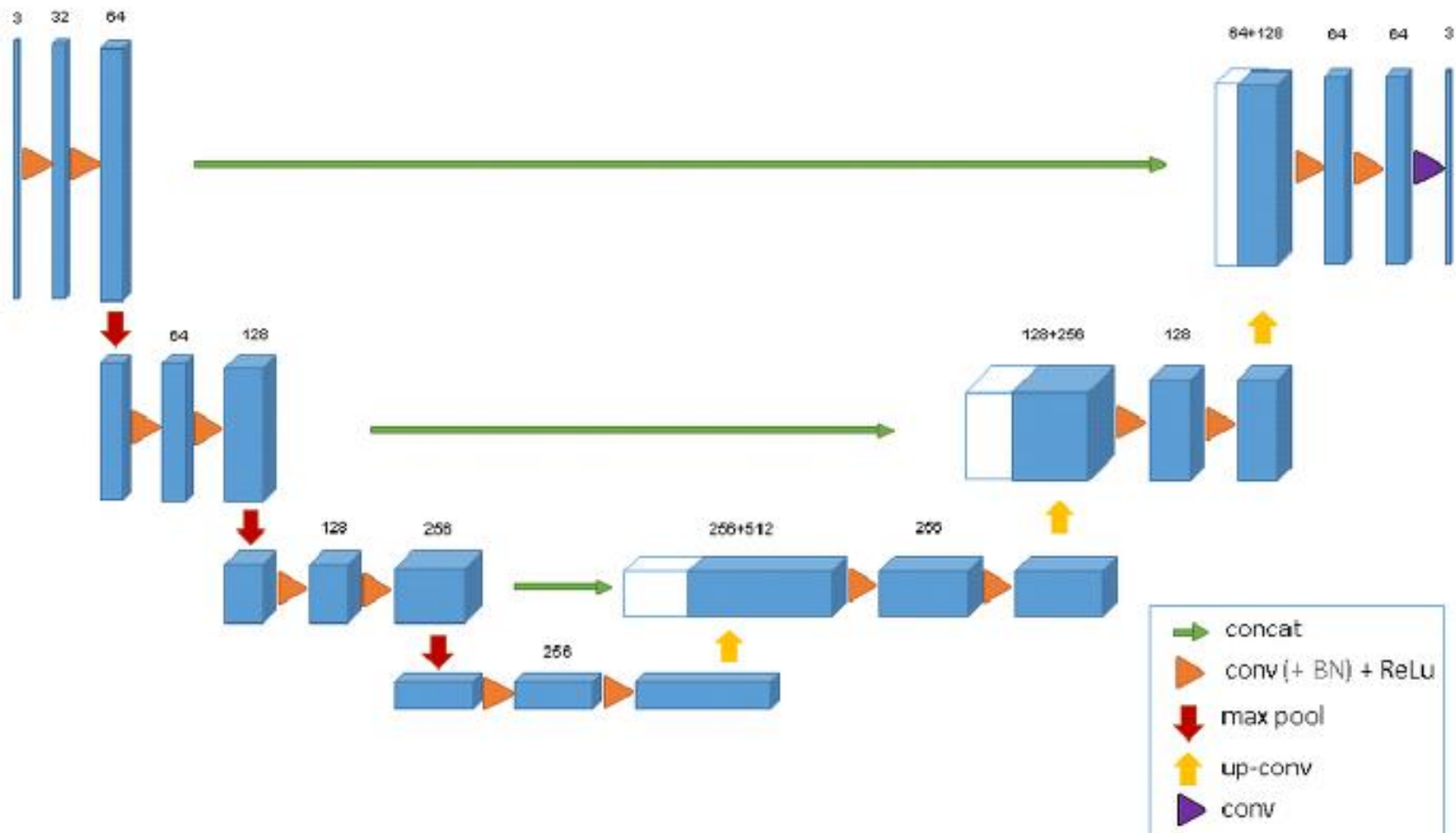
- 의학 데이터 분석에서 3D 데이터(volumetric data)는 많이 존재합니다. 하지만 이러한 데이터를 annotate하는 것은 매우 어려운 일입니다. 왜냐하면 컴퓨터 화면에서는 오직 2D 단면만 시각화 할 수 있기 때문입니다.
- 또한 이웃하는 단면들은 서로 거의 동일한 정보를 갖고 있기 때문에 이를 하나하나 annotation하는 것은 매우 비효율적입니다. 즉, 3D 데이터의 전체 단면에 대해서 annotate하는 것은 매우 효율적인 방법이 아니며 크고 질 좋은 데이터셋을 만드는 것이 어렵습니다.
- 이 논문에서는 3D 데이터의 전체 단면이 아닌 일부 단면에 대해서만 annotate된 데이터를 필요로 하는 네트워크를 제시합니다. 3D U-Net은 아래의 그림과 같은 두 가지 방법으로 사용될 수 있습니다.





- a. Semi-automated segmentation : 사용자가 segment될 각각의 volume의 몇 개의 단면만 annotate합니다. 그리고 네트워크는 dense segmentation을 예측합니다. 즉, sparsely annotated slices로 부터 dense annotated slices를 예측하도록 학습합니다.
- b. Fully-automated segmentation : 네트워크는 representative training set의 annotated 단면으로 학습을 하고 non-annotated volumes을 예측할 수 있습니다. 이는 흔히 이루어지는 segmentation task라고 생각하시면 될 것 같습니다. [Fully-annotated volumes로 학습한 후, non-annotated volumes을 예측]





- 기존의 U-Net과 큰 차이는 없기 때문에 간단하게 중요한 점만 정리하고 넘어가도록 하겠습니다.

1. Max pooling 전에 channel의 수를 2배로 늘려 bottleneck를 해결했습니다.
2. Batch normalization("BN")을 사용했습니다.
3. Sparse annotation에 대해서도 학습할 수 있도록 가중 소프트맥스 손실 함수(weighted softmax loss function)을 사용합니다.

# Cross entropy

0	1	2	3	4	5	6	7	8	9
0	0	0	0	0	0	1	0	0	0

actual probabilities, "one-hot" encoded

Cross entropy:  $-\sum Y_i' \cdot \log(Y_i)$

computed probabilities

.01	.03	.00	.04	.03	.05	0.8	.02	.01	.01
0	1	2	3	4	5	6	7	8	9

this is a "6"

$$loss = -\sum_i^N w_i * p_{gt}^i \ln(p_{pred}^i) = w[index\_gt] \ln(p_{pred}[index\_gt])$$

# Experiments and results

## 3. Implementation Details

- Data augmentation에 회전(rotation), scaling, 그리고 gray value augmentation 뿐만 아니라 smooth dense deformation도 추가하였습니다.
- Weighted cross-entropy loss에서 background에 대한 가중치(weights)는 상대적으로 줄이고 segment하고자 하는 대상에 대한 가중치는 높여서 class imbalance를 해결했습니다.  
[논문에는 기술되어 있지 않지만 주로 positive patch와 negative patch의 비율을 1:3으로 구성하여 학습합니다.]
- Optimizer로는 Stochastic Gradient Descent (SGD)를 사용했습니다.
- NVIDIA TitanX GPU를 통한 70000 training iterations에 3일이 소요되었습니다.

# Experiments and results

Table 1: Cross validation results for semi-automated segmentation (IoU)

test slices	3D w/o BN	3D with BN	2D with BN
subset 1	0.822	0.855	0.785
subset 2	0.857	0.871	0.820
subset 3	0.846	0.863	0.782
average	0.842	0.863	0.796

Table 2: Effect of # of slices for semi-automated segmentation (IoU)

GT slices	GT voxels	IoU S1	IoU S2	IoU S3
1,1,1	2.5%	0.331	0.483	0.475
2,2,1	3.3%	0.676	0.579	0.738
3,3,2	5.7%	0.761	0.808	0.835
5,5,3	8.9%	0.856	0.849	0.872

**Table 1:** 2D보다 3D를 사용하는 것이, 그리고 Batch Normalization을 사용하는 것이 더욱 좋은 결과를 보입니다.

**Table 2:** Semi-automated 학습에 사용되는 단면(slices)의 개수가 많을수록 더 좋은 성능을 보입니다.

Table 3: Cross validation results for fully-automated segmentation (IoU)

test volume	3D w/o BN	3D with BN	2D with BN
1	0.655	0.761	0.619
2	0.734	0.798	0.698
3	0.779	0.554	0.325
average	0.723	0.704	0.547

**Table 3:** Fully-automated segmentation에서도 2D보다는 3D를 사용하는 것이, 그리고 BN을 사용하는 것이 더 좋은 성능을 보입니다.

# Conclusion

- 3D U-Net은 2D input을 받고 2D 연산을 하였던 기존의 U-Net을 확장시켜 3D 데이터를 다룰 수 있을 뿐만 아니라 이를 통해 성능을 향상시켰습니다.
- 이 논문을 이후로 의료 데이터에서는 주로 2D 단면을 통해 학습하기보다는 3D patch를 통해 학습합니다.

# **3D Context Enhanced Region-based Convolutional Neural Network for End- to-End Lesion Detection (3DCE)**

---

# Abstract

- 이 논문에서는 2D 이미지로부터 얻은 feature maps을 합치는 방법을 통해 3D 정보를 활용하는 3D context enhanced region-based CNN (**3DCE**)를 제안합니다.
- 3DCE는 학습과 inference 모두 end-to-end로 가능하고 모든 종류의 병변(lesion)을 찾을 수 있는 universal lesion detector입니다.



# Abstract

- 이 논문에서는 2D 이미지로부터 얻은 feature maps을 합치는 방법을 통해 3D 정보를 활용하는 3D context enhanced region-based CNN (**3DCE**)를 제안합니다.
- 3DCE는 학습과 inference 모두 end-to-end로 가능하고 모든 종류의 병변(lesion)을 찾을 수 있는 universal lesion detector입니다.

# Introduction

- Abstract에서도 언급했듯이 3D 정보(context)는 병변과 병변이 아닌 영역을 구분하는 것에 중요한 역할을 합니다.
- 따라서 기존의 방법은 단순히 2D 이미지가 아닌 3D volume을 input으로 주었습니다. 하지만 직접적으로 3D volume을 input으로 주게 되면 메모리가 많이 소모된다는 문제(주로 작은 3D patches를 만드는 방법으로 해결)와 3D bounding box는 annotate하기 어렵다는 문제가 있습니다.

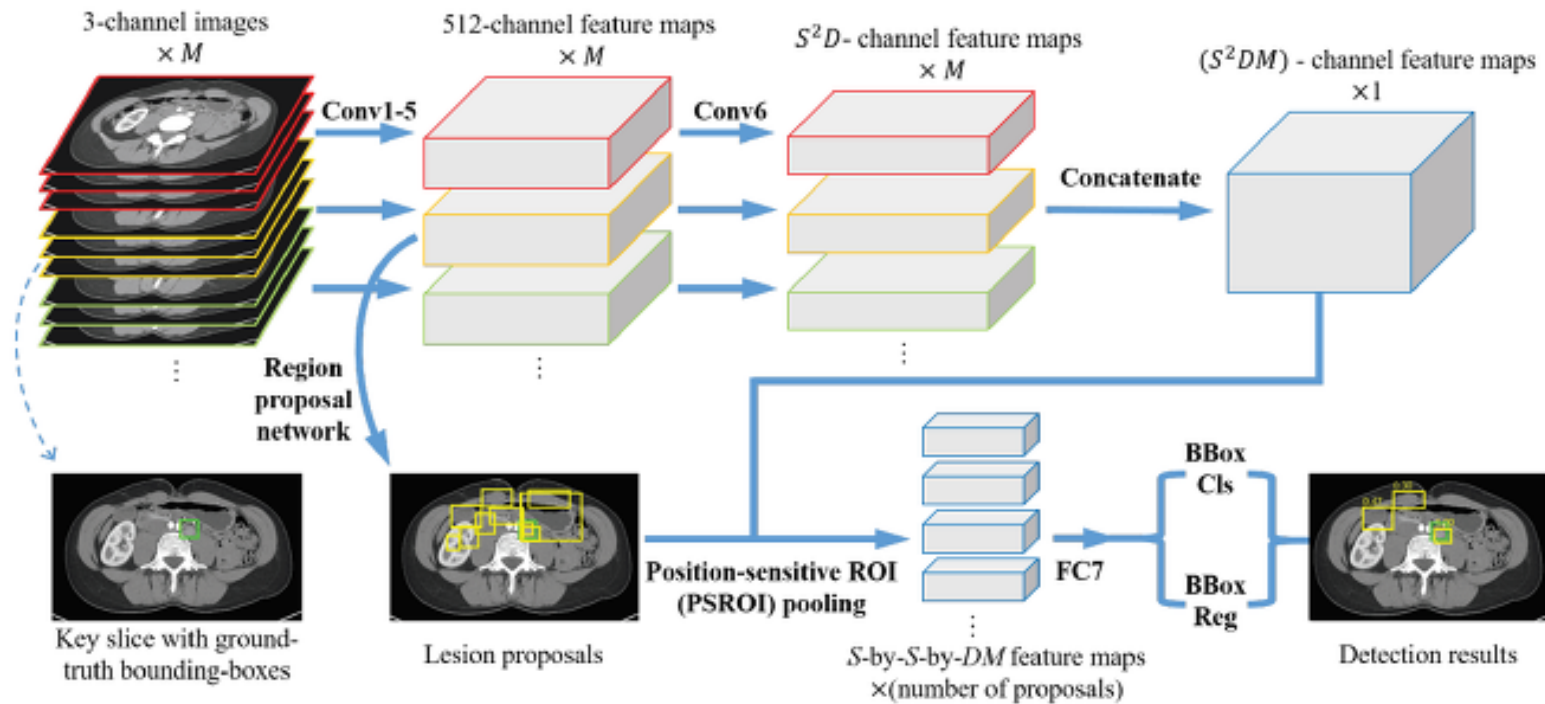
# Introduction

- 3D patch 또한 2D slice에 비하면 메모리를 많이 차지하기 때문에 연속된 3개의 단면을 네트워크의 input으로 준 방법도 제시되었습니다.
- 하지만 해당 논문에서 제시한 방법은 병변 후보(lesion candidates)들이 병변인지 최종적으로 결정하기 위해서 독립적인 CNN 분류 모델을 사용했기 때문에 two stages model로 end-to-end하지 않다는 단점이 있습니다.

# Introduction

- 논문에서는 위 방법들에서 제기된 문제점들을 모두 해결한 3D context enhanced region-base CNNs (**3DCE**)를 제시합니다.
- 2D 네트워크에 여러 개의 이웃한 단면들이 주어진다면 각각의 feature maps이 만들어지고 이 feature map들은 최종 결정을 위해 합쳐집니다.
- 그리고 **R-FCN**(region-based fully convolutional network)을 향상시켰습니다. 이러한 방법들을 거친 3DCE는 아래의 장점들이 있습니다.
  - 1) 학습과 inference를 end-to-end로 할 수 있다.
  - 2) VGG16과 같은 2D CNN backbone과 pretrained weights를 사용할 수 있다.
  - 3) 3D보다 annotation하는 것이 쉬운 2D bounding box를 요구합니다.

# Method



**Fig. 1.** The framework of 3DCE for lesion detection.

# Method

- 3DCE에서는 널리 쓰이는 Faster R-CNN 대신에 더욱 빠르고 정확하며 메모리 효율적인 R-FCN을 이용했습니다.
- R-FCN은 물체의 클래스 정보와 위치 정보를 담고 있는 position-sensitive score maps로 구성되고 이러한 score maps는 이후의 position-sensitive region of interest (PSROI) pooling layer에서 summarize됩니다.
- 최종 bounding box에 대한 결과는 다른 pooling 연산에 의해 얻어집니다.

# Method

- 3DCE는 PSROI pooling 뒤에 3개의 새로운 레이어를 추가함으로써 R-FCN을 향상시켰습니다.
- 추가한 3개의 레이어는 다음과 같습니다: 2048D FC layer, ReLU 레이어, 그리고 classification과 bounding-box regression을 위한 두 개의 레이어.
- 3개의 추가적인 레이어덕분에 0.7%만큼 성능을 향상할 수 있었다고 합니다.



# Method

- Fig 1.에 나타나 있듯이 M개의 3-channels 이미지가 그룹화됩니다.
- 학습할 때, 중간 단면이 ground-truth bounding-box를 가지고 다른 단면들은 3D 정보를 제공합니다.
- M개의 3-channels 이미지가 (VGG16의 Conv1-5의) convolution 연산을 거쳐 M개의 feature maps을 만듭니다.
- Lesion proposals를 생성하기 위해서 오직 가운데의 단면에서 얻은 feature map만 region proposal network(RPN)로 전달됩니다.
- 나머지 feature maps들은 모두 추가적인 convolution layer (conv6)를 거친 후 concatenate되어 S2DM-channel feature map을 형성합니다.
- (S = pooled feature map의 크기 / D = 2D feature maps의 개수 / 논문에서는 S = 7, D = 10으로 설정)

# Experiments and results

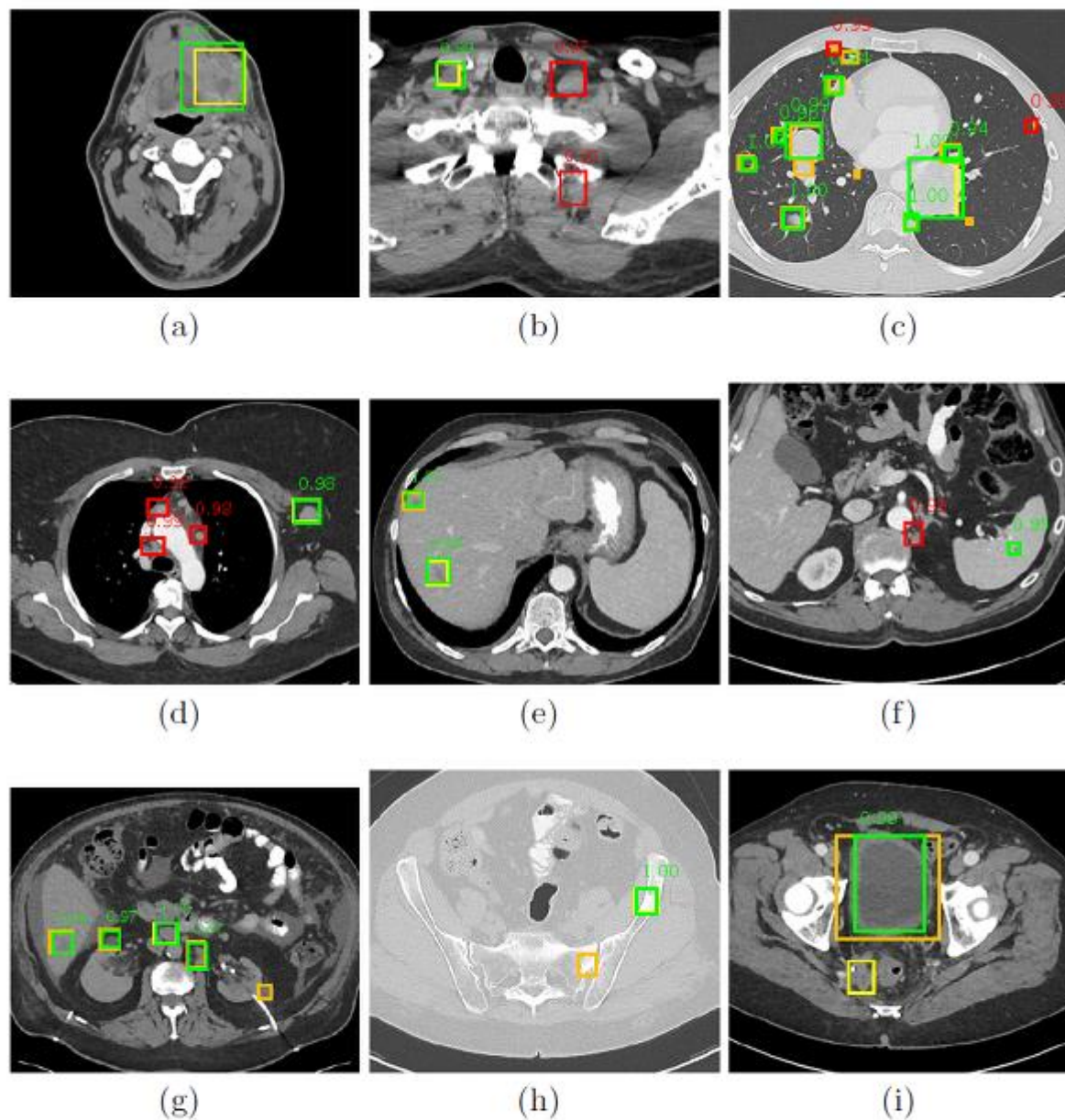
## 3. Experiments

- DeepLesion Dataset에 대한 전처리는 아래와 같습니다.

- 1) -1024 ~ 3071 HU 값을 갖는 이미지의 픽셀값들을 [0, 255]의 범위 안의 값으로 rescaling
- 2) 각각의 픽셀이 0.8mm에 해당하도록 resizing
- 3) 각 단면 간의 거리가 2mm가 되도록 설정
- 4) 데이터셋 분할 : Training (70%) - Validation (15%) - Test (15%)
- 5) Data augmentation을 사용하지 않음

**Table 2.** Sensitivity (%) at 4 FPs per image of different methods. Both the original test set and the all-lesion test set are investigated using two overlap computation criteria.

	Original test set		All-lesion test set		Inference time (ms)
	IoU	IoBB	IoU	IoBB	
No 3D context	76.51	80.16	66.90	72.47	<b>19</b>
Faster RCNN, 3 slices	80.32	85.34	71.60	80.31	32
Improved R-FCN, 3 slices	81.53	85.89	74.39	81.88	<b>19</b>
Data-level fusion, 11 slices	82.94	86.52	74.56	80.49	28
3DCE, 9 slices	83.57	87.81	<b>76.31</b>	<b>82.75</b>	56
3DCE, 27 slices	<b>84.37</b>	<b>87.85</b>	75.09	<b>82.75</b>	114



**Fig. 3.** Detection results of 3DCE with scores  $> 0.9$  on the test set of DeepLesion. Yellow, orange, green, and red boxes correspond to ground-truths in the test set, additional ground-truths in the all-lesion test set, predicted true positives, and false positives.

# Conclusion

- 3DCE는 3D patch가 아닌 2D 단면을 input으로 받으면서도 각각의 2D 단면에서 얻어낸 feature maps을 concat해주는 방법을 통해 3D 정보를 이용했다는 특징이 있습니다.
- 이러한 방법을 통해 기존의 3D 모델과는 달리 메모리적으로 효율적이고 빠르다는 장점을 보입니다. 그리고 모든 과정이 학습과 inference 과정에서 end-to-end로 진행된다는 장점도 있습니다.