
CAP 5516

Medical Image Computing

(Spring 2022)

Dr. Chen Chen

Center for Research in Computer Vision (CRCV)

University of Central Florida

Office: HEC 221

Address: 4328 Scorpius St., Orlando, FL 32816-2365

Email: chen.chen@crcv.ucf.edu

Web: <https://www.crcv.ucf.edu/chenchen/>

Lecture 9

Medical Image Segmentation

Medical Image Segmentation

- Problem Definition
- Evaluation Metrics
- Datasets
- Traditional Methods
- Deep Learning based Methods

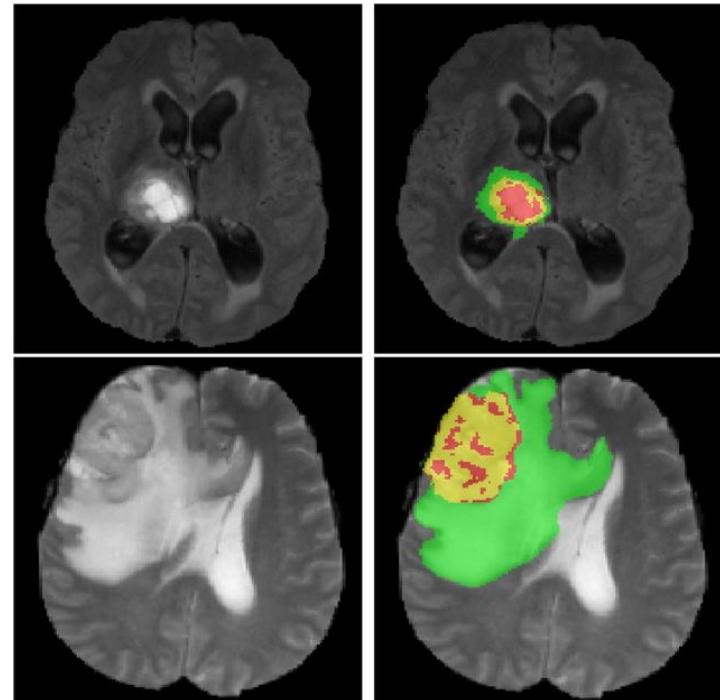
1 Problem Definition

Semantic Segmentation :

- Label each pixel in the image with a category label
- Do not differentiate instances, only care about pixels

For Medical

Use computer image processing technology to analyze and process 2D or 3D images to achieve segmentation of human **organs, soft tissues** and diseased bodies.



1 Problem Definition

The process of medical image segmentation can be divided into the following stages:

1. Obtain medical imaging data set
2. Preprocess the images
3. Use appropriate medical image segmentation method
4. Performance evaluation.

2 Evaluation Metrics

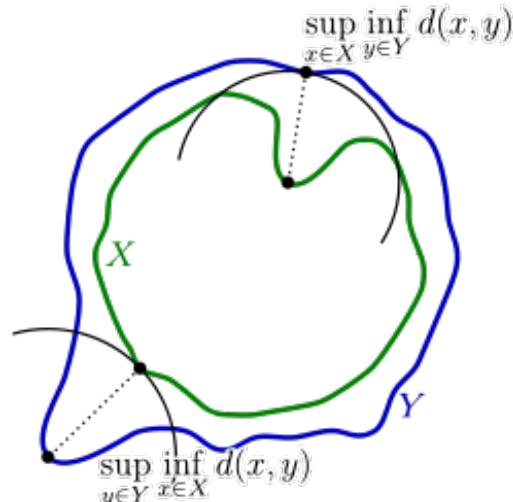
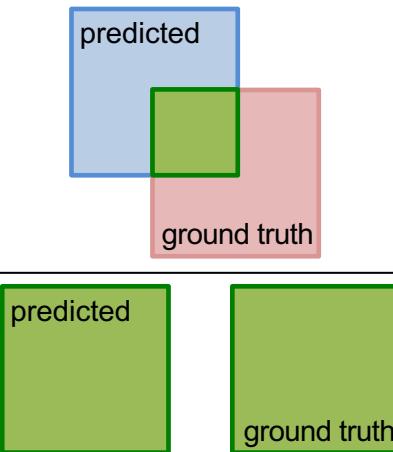
Dice index

$$Dice(A, B) = 2 \frac{|A \cap B|}{|A| + |B|}$$

Jaccard index(IoU)

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

$$\text{Dice Coefficient} = \frac{2 \times \text{area of overlapped (green)}}{\text{total area (green)}} =$$



Hausdorff distance

$$H(X, Y) = \max(h(X, Y), h(Y, X))$$

where

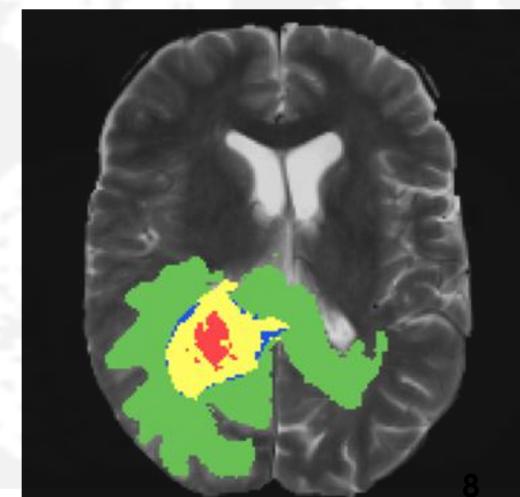
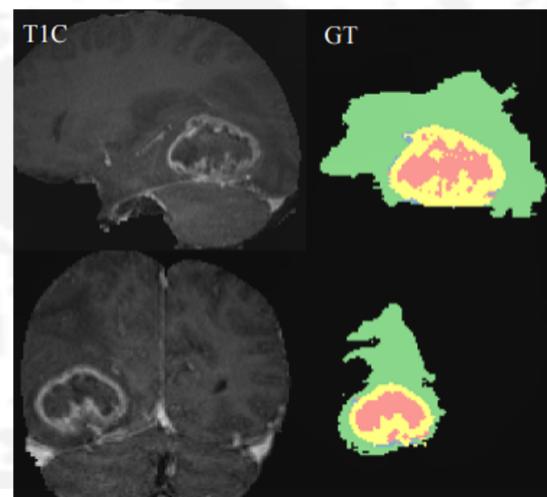
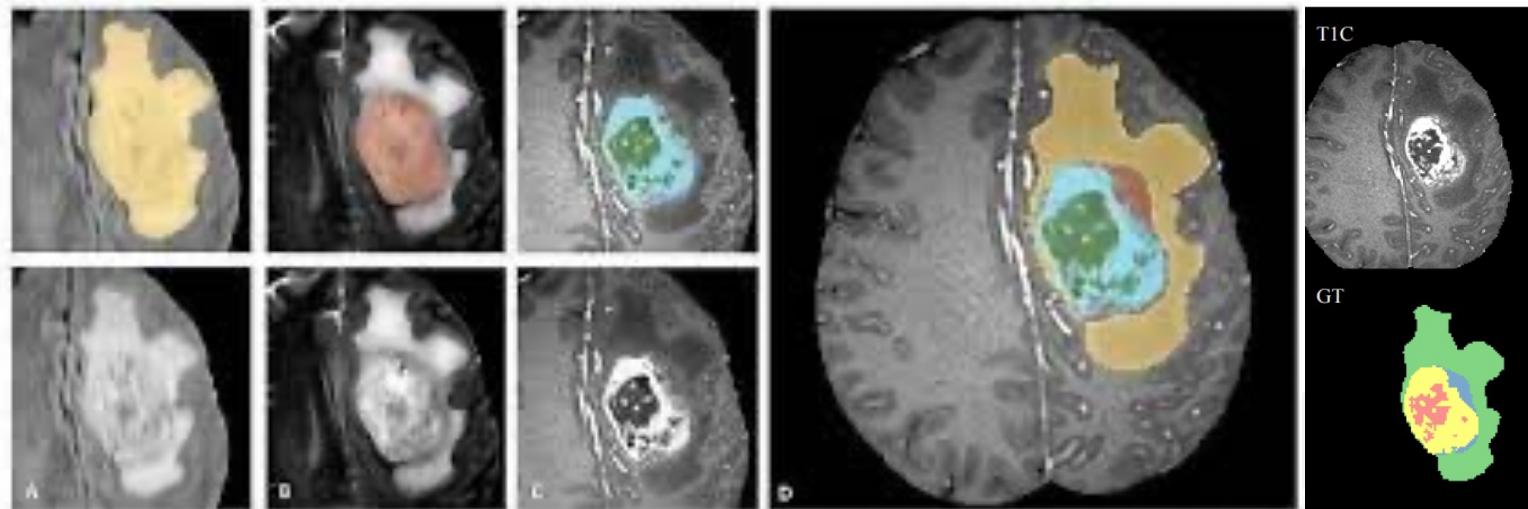
$$h(X, Y) = \max \left(\min_{x \in X} \|x - y\| \right)$$
$$h(Y, X) = \max \left(\min_{y \in Y} \|y - x\| \right)$$

3 Datasets

Data Set	Modalities	Objects	URL
MSD	MRI, CT	Various	http://medicaldecathlon.com/
BraTS	MRI	Brain	https://www.med.upenn.edu/sbia/brats2018/data.html
DDSM	Mammography	Breast	http://www.eng.usf.edu/cvprg/Mammography/Database.html
ISLES	MRI	Brain	http://www.isles-challenge.org/
LiTS	CT	Liver	https://competitions.codalab.org/competitions/17094
PROMISE12	MRI	Prostate	https://promise12.grand-challenge.org/
LIDC-IDRI	CT	Lung	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
OASIS	MRI, PET	Brain	https://www.oasis-brains.org/
DRIVE	Funduscopy	Eye	https://drive.grand-challenge.org/
STARE	Funduscopy	Eye	http://homes.esat.kuleuven.be/~mblaschk/projects/retina/
CHASEDB1	Funduscopy	Eye	https://blogs.kingston.ac.uk/retinal/chasedb1/
MIAS	X-ray	Breast	https://www.repository.cam.ac.uk/handle/1810/250394?show=full
KITS21	CT	Kidney	https://kits21.kits-challenge.org/
HVS MR2016	CMR	Heart	http://segchd.csail.mit.edu/

3 Datasets

Brain Tumor Segmentation (BraTS)



3 Datasets

Digital retinal images for vessel extraction (DRIVE)

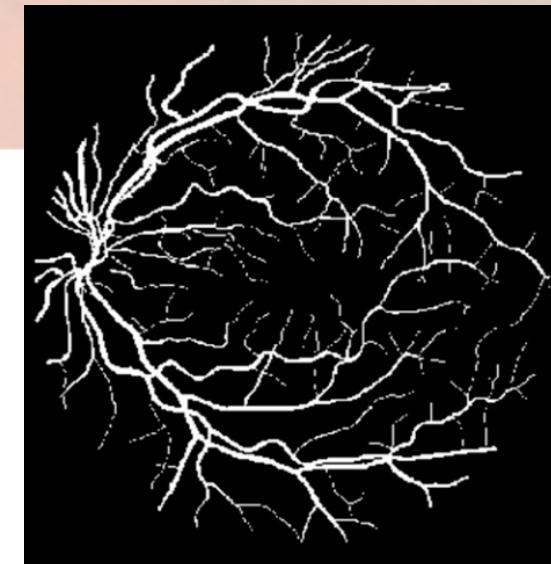
DRIVE



(a)



(b)



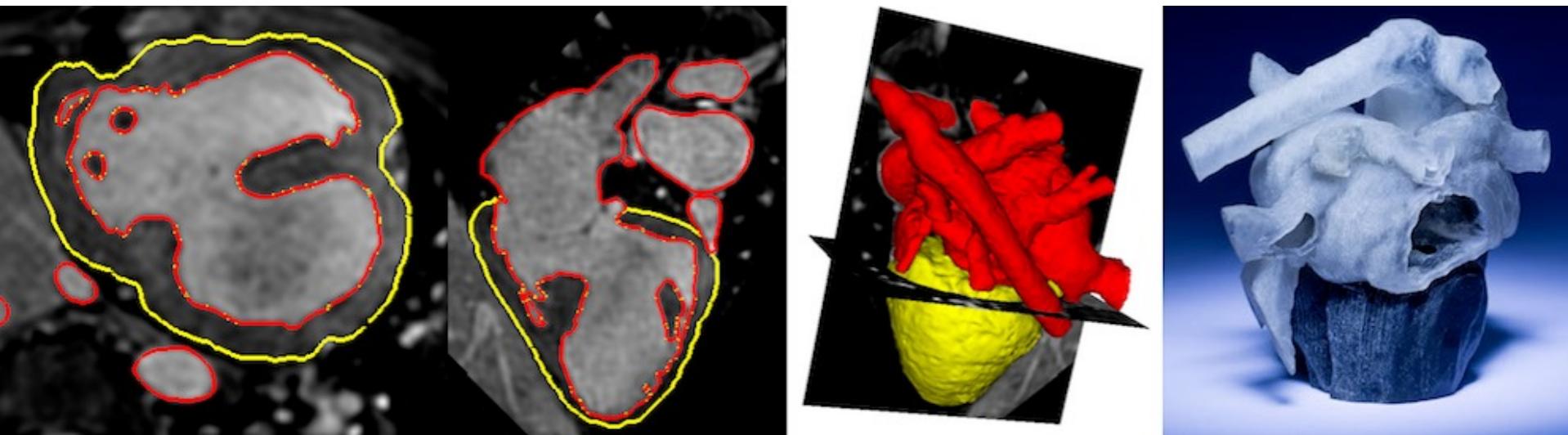
(c)

Figure 1. Digital retinal images for vessel extraction (DRIVE) sample diagram and manual labeling sample. (a) The blood vessels in retinal RGB image; (b) manual annotation 1 of sample; (c) manual annotation 2 of sample.

3 Datasets

HVS MR 2016

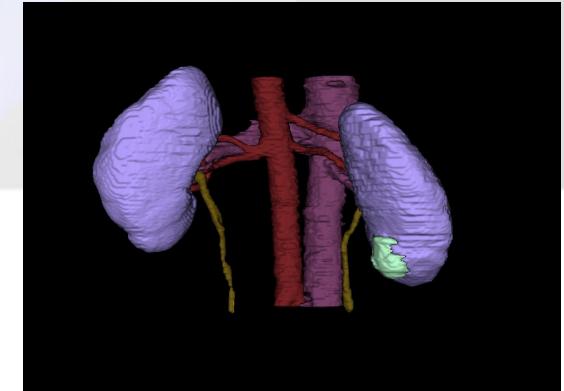
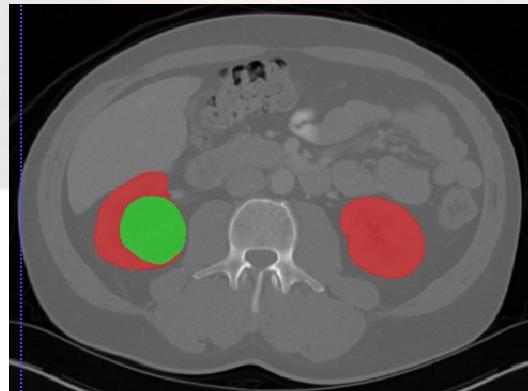
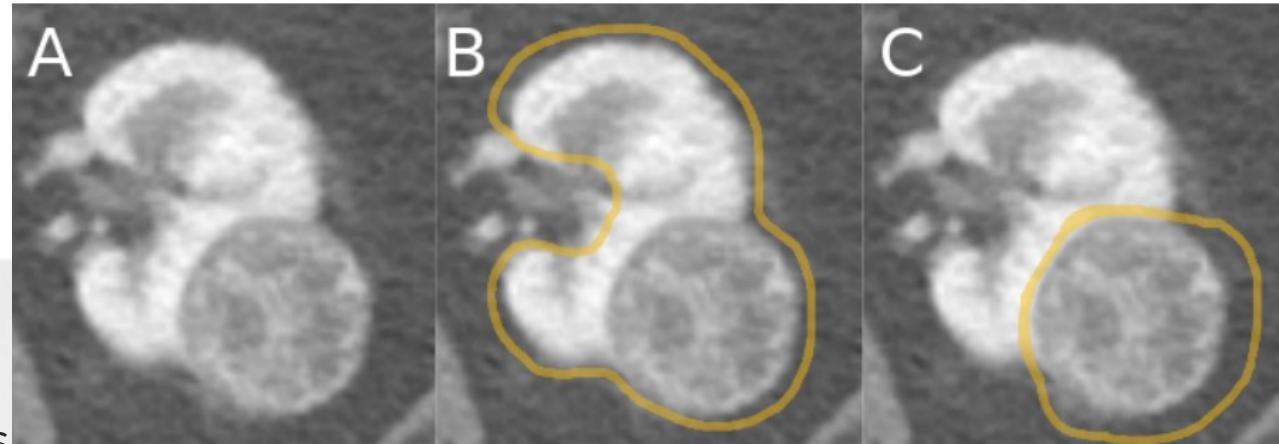
MICCAI Workshop on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease.



3 Datasets

KiTS19/21

Kidney Tumor Segmentation challenge is a competition in which teams compete to develop the best system for automatic semantic segmentation of renal tumors and surrounding anatomy using CT scans.



KiTS19 Challenge

4 Traditional Methods

- **Threshold-based segmentation**
- **Edge-based segmentation**
- **Region-based segmentation**
- Segmentation based on Graph Theory
- Segmentation based on Energy Function

4 Traditional Methods

Threshold-based Method

- Simple calculation, high efficiency
- Only consider the characteristics of the pixel point grayscale value itself, generally do not consider space characteristics
- more sensitive to noise, and the robustness is not high.

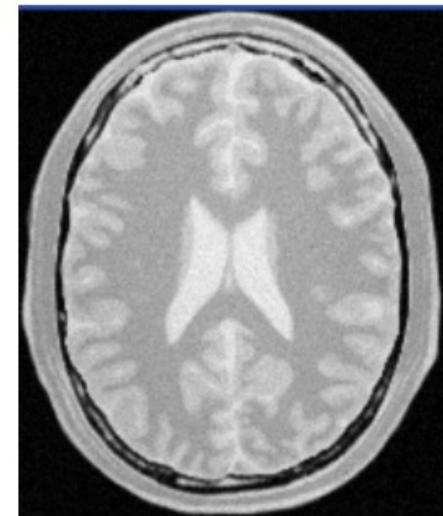
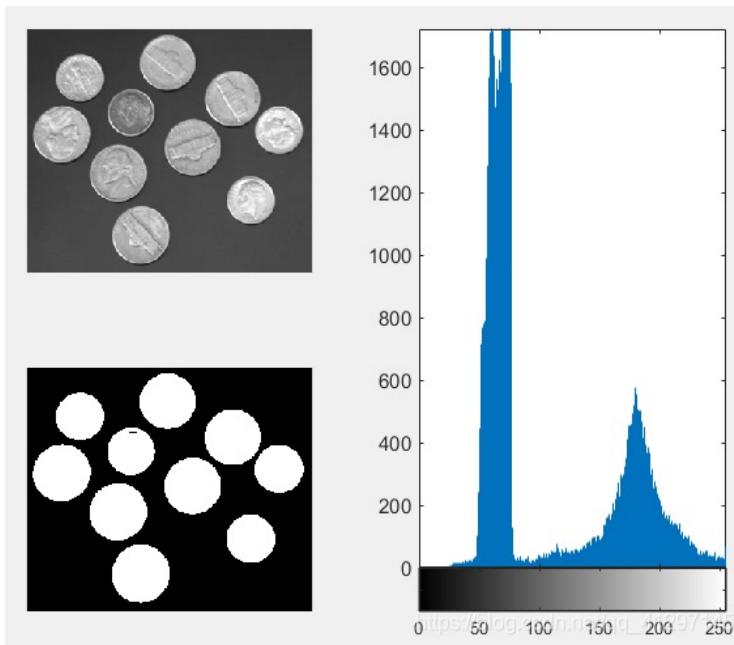


Figure 3. Original image.



Figure 5. The same image after the proposed threshold-based level set method.

4 Traditional Methods

Edge Detection-based Method

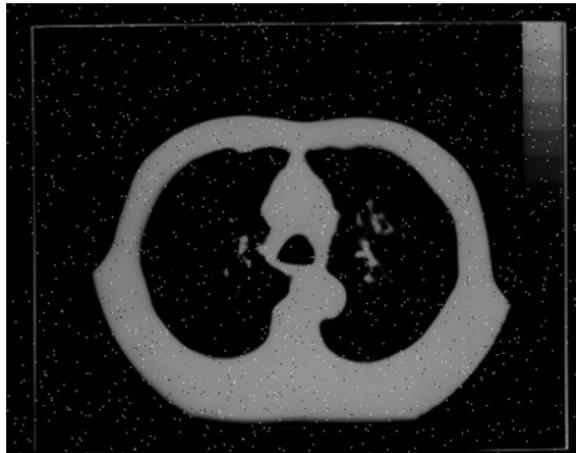


Fig.1. Original lungs CT image with salt-and-pepper noise.

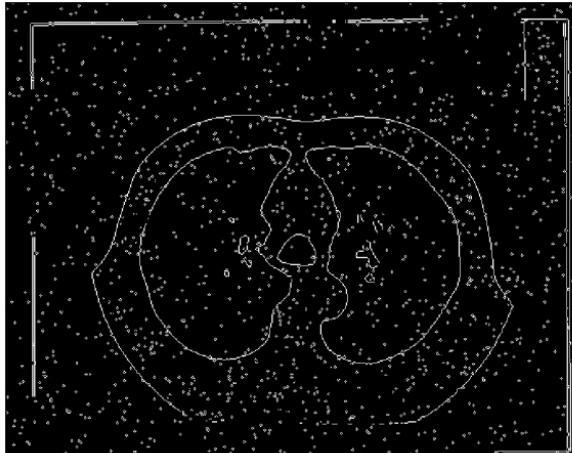


Fig.3. Lungs CT image processed by Sobel detector.



Fig.5. Lungs CT image processed by dilation residue edge detector.

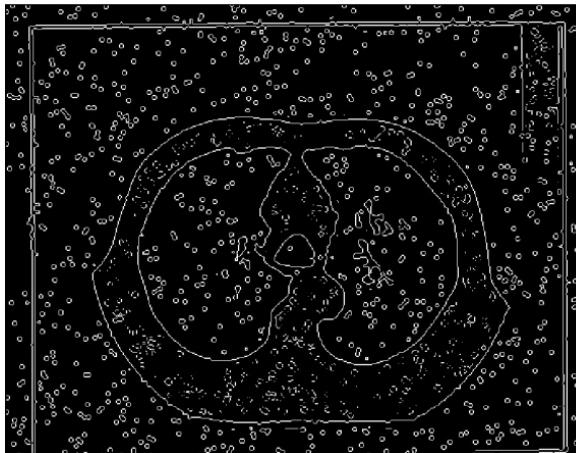


Fig.2. Lungs CT image processed by Laplacian of Gaussian operator.



Fig.4. Lungs CT image processed by morphological gradient operation.



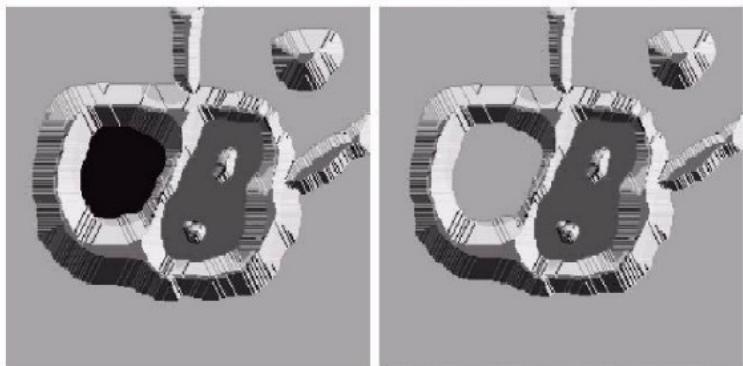
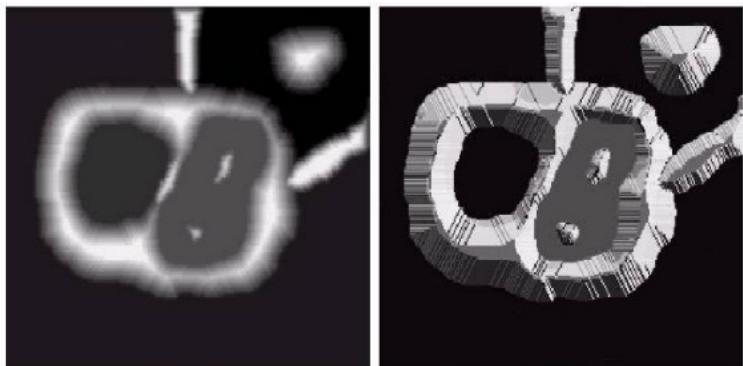
Fig.6. Lungs CT image processed by the novel morphological edge detector.

4 Traditional Methods

Region-based Method

a
b
c
d

FIGURE 10.44
(a) Original image.
(b) Topographic view.
(c)–(d) Two stages of flooding.



e
f
g
h

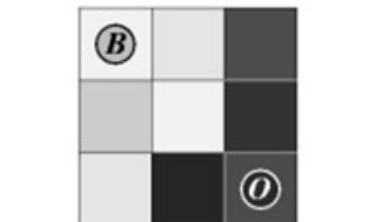
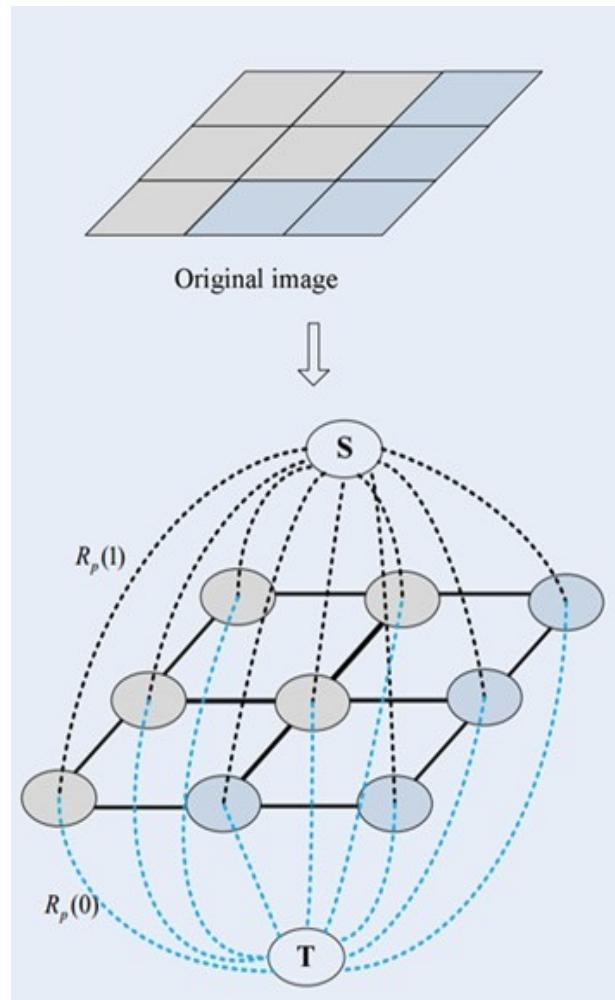
FIGURE 10.44
(Continued)
(e) Result of further flooding.
(f) Beginning of merging of water from two catchment basins (a short dam was built between them). (g) Longer dams. (h) Final watershed (segmentation) lines. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)

Watershed Method

Cigla, C.; Alatan, A.A. Region-based image segmentation via graph cuts. In Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008; pp. 2272–2275.

4 Traditional Methods

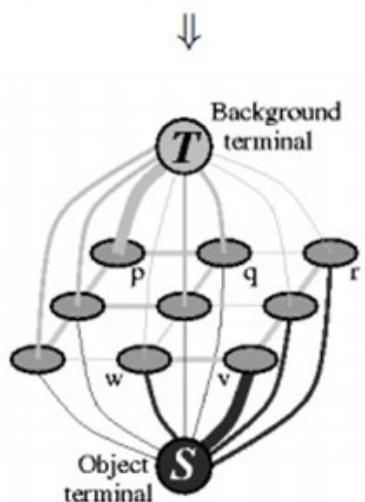
Segmentation based on Graph Theory



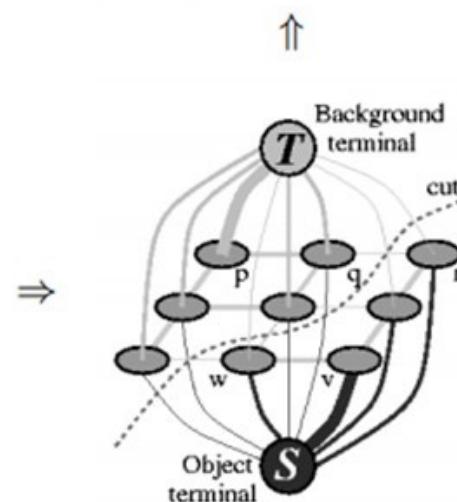
(a) Image with seeds.



(d) Segmentation results.



(b) Graph.



(c) Cut.

3x3 image segmentation of Graph Cut

5 Deep Learning based Methods

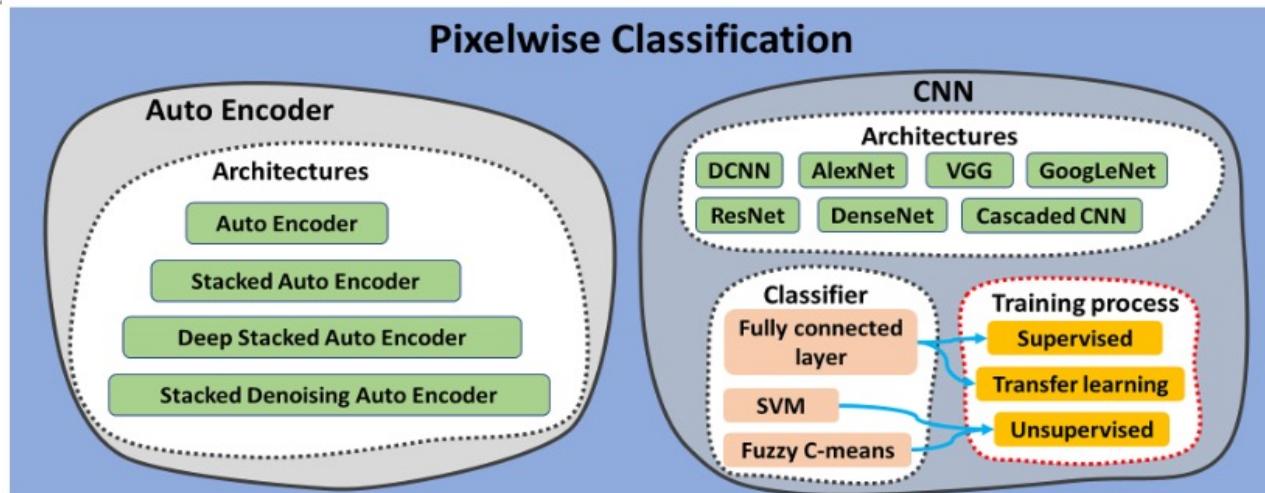


Fig. 1 The network components of the pixelwise classification methods.

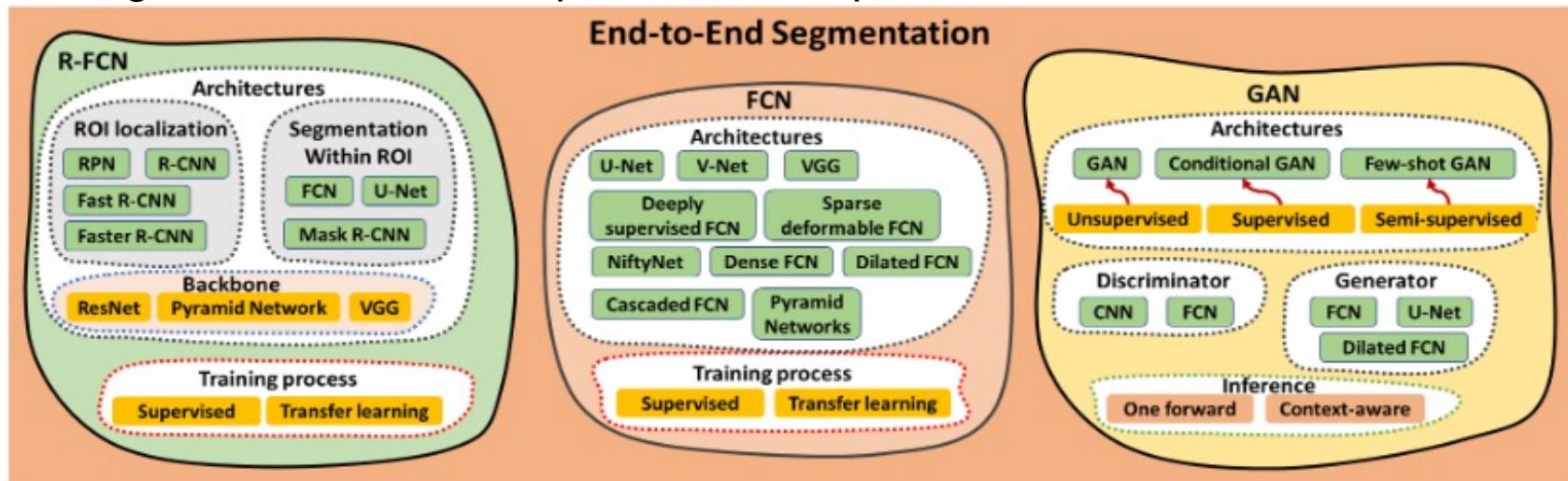


Fig. 2 The network components of the end-to-end segmentation methods.

Fu Y, Lei Y, Wang T, et al. A review of deep learning based methods for medical image multi-organ segmentation[J]. *Physica Medica*, 2021, 85: 107-122.

5 Deep Learning based Methods

Overview

Year	Network	Dimension	Site	Modality	Year	Network	Dimension	Site	Modality
2017	Deep deconvolutional neural network (DDNN)	2D slice	Brain	CT	2018	2D and 3D CNN	2D slice, 3D volume	Artery / vein	CT
2017	3D CNN	3D patch	Brain lesion	MRI	2018	3D ConvNets	3D volume	Brain	MRI
2015	Multi-level DCNN	2D patch	Pancreas	CT	2018	CNN with specific fine-tuning	2D slice, 3D volume	Brain, abdomen	Fetal MRI
2016	Holistically Nested CNN	2D patch	Pancreas	CT	2018	2D and 3D DCNN	2D slice, 3D volume	Whole body	CT
2017	3D CNN	3D patch	Chest	CT	2019	Deep fusion Network	2D slice	Chest	CXR
2017	3D DCNN	Not specified	Abdomen	CT	2019	DCNN	2D slice	Abdomen	CT
2017	CNN	3D patch	Head & Neck	CT	2019	2.5D CNN	2.5D patch	Thorax	CT
2017	Fuzzy-C-Means CNN	3D patch	Lung nodule	CT	2019	Cascaded CNN	2D slice	Head & Neck	CT
2017	DCNN	2D Slice	Body, Chest, Abdomen	CT	2019	2D and 3D CNN	2D slice, 3D volume	Thorax	CT
2018	Fusion Net	2D patch	100 ROIs	HRCT	2019	U-Net Neural Network	3D patch	Lung	CT
2018	DCNN	2D patch	Spinal lesion	CT					
2018	DCNN	2D slice	Malignant pleural mesothelioma	CT					

5 Deep Learning based Methods

Overview (continued)

Year	Network	Dimension	Site	Modality
2015	U-Net	2D slice	Neuronal structure	Electron microscopic
2016	3D U-Net	3D volume	Kidney	Microscopic
2017	Dilated FCN	2D slice	Abdomen	CT
2017	3D FCN Feature Driven Regression Forest	3D patch	Pancreas	CT
2017	2D FCN	2.5D slices	Whole body	CT
2018	Foveal Fully Convolutional Nets	N.A.*	Whole body	CT
2018	DRINet	2D slice	Brain, abdomen	CT
2018	3D U-Net	3D volume	Prostate	MRI
2018	Dense V-Net	3D volume	Abdomen	CT
2018	NiftyNet	3D volume	Abdomen	CT
2018	PU-Net, CU-Net	2D slice	Pelvis	CT
2018	Dilated U-Net	2D slice	Chest	CT
2018	3D U-JAPA-Net	3D volume	Abdomen	CT
2018	U-Net	2D slice	Pelvis	CT
2018	Multi-scale Pyramid of 3D FCN	3D patch	Abdomen	CT
2018	Shape representation model constrained FCN	3D volume	Head & Neck	CT
2018	Hierarchical Dilated Neural Networks	2D slice	Pelvis	CT
2018	Dense 3D FCN	3D volume	Abdomen	MRI
2018	3D FCN	3D patch	Head & Neck	CT
2019	Dilated FCN	2D slice	Lung	CT
2019	Dense-U-Net	2D slice	Head & Neck	Stained colon adenocarcinoma dataset
2019	2D and 3D FCNs	2D slice and 3D volume	Pulmonary nodule	CT
2019	Dedicated 3D FCN	3D patch	Thorax/abdomen	DECT
2019	2D FCN (DeepLabV3 +)	2D slice	Pelvis	MRI
2019	2D FCN	2D patch	Pulmonary vessels	CT

5 Deep Learning based Methods

Overview (continued)

Year	Network	Dimension	Site	Modality
2019	Dual U-Net	2D slice	Glioma Nuclei	Hematoxylin and eosin (H&E)-stained histopathological image
2019	Consecutive deep encoder-decoder Network	2D slice	Skin lesion	CT
2019	U-Net	2D slice	Lung	HRCT
2019	3D U-Net	3D volume	Chest	CT
2019	3D U-Net with Multi-atlas	3D volume	Brain tumor	Dual-energy CT
2019	Triple-Branch FCN	Not specified	Abdomen/torso	CT
2019	2.5D Deeply supervised V-Net	2.5 patch	Prostate	Ultrasound
2019	Group dilated deeply supervised FCN	3D volume	Prostate	MRI
2019	3D FCN	3D volume	Arteriovenous malformations	Contract-enhanced CT
2019	3D FCN	3D volume	Left ventricle	SPECT
2019	DeepMAD	2.5D patch	Vessel wall	MRI
2019	3D U-Net	3D volume	Head & Neck	CT
2019	OBELISK-Net	3D volume	Abdomen	CT
2019	OAN-RC	2D slice	Abdomen	CT
2019	Multi-stage 3D FCN	3D volume	Head & Neck	CT
2019	2D/3D FCN	3D patch	Abdomen	CT
2020	U-Net	3D patch	Abdomen	CT
2020	2.5D U-Net	2.5D patch	Body	CT
2020	3D Attention U-Net	3D patch	Pancreas/Abdomen	CT
2020	3D U-Net	3D patch	Thoracic/Abdomen	CT
2020	3D U-Net	3D volume	Head & Neck	CT

5 Deep Learning based Methods

FCN adapting classifiers for dense prediction

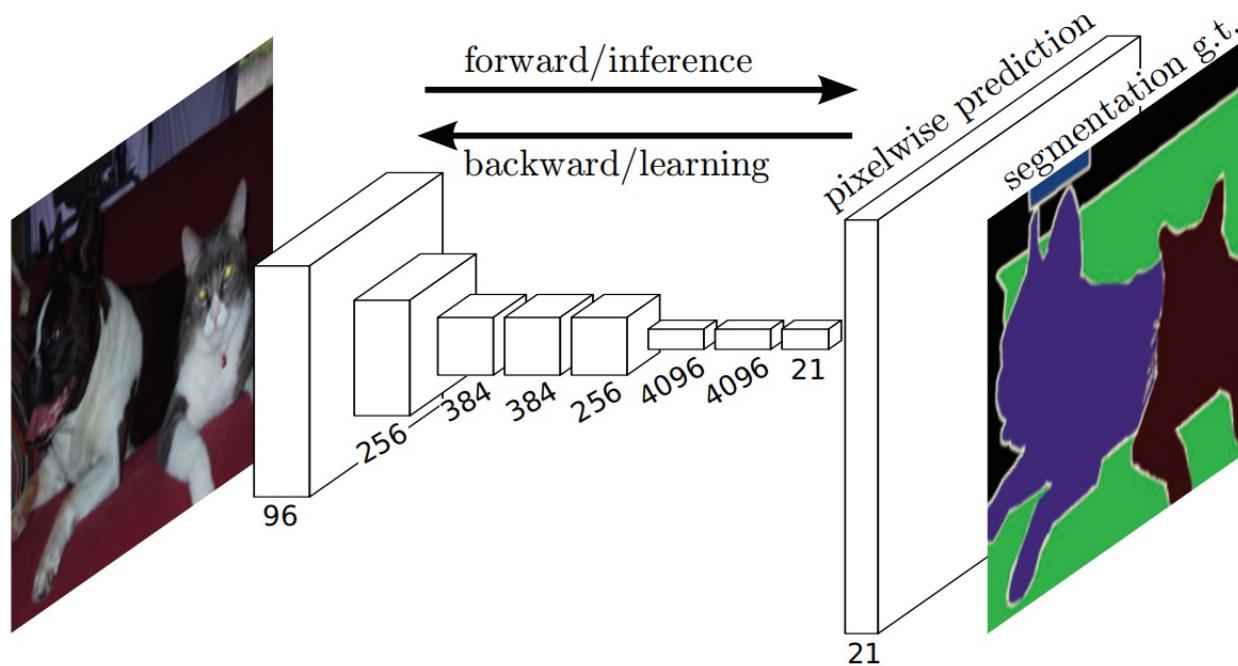


Figure 1. Fully convolutional networks can efficiently learn to make dense predictions for per-pixel tasks like semantic segmentation.

Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

5 Deep Learning based Methods

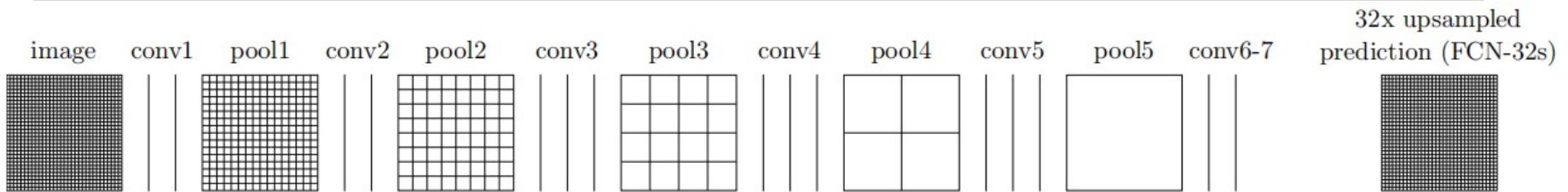
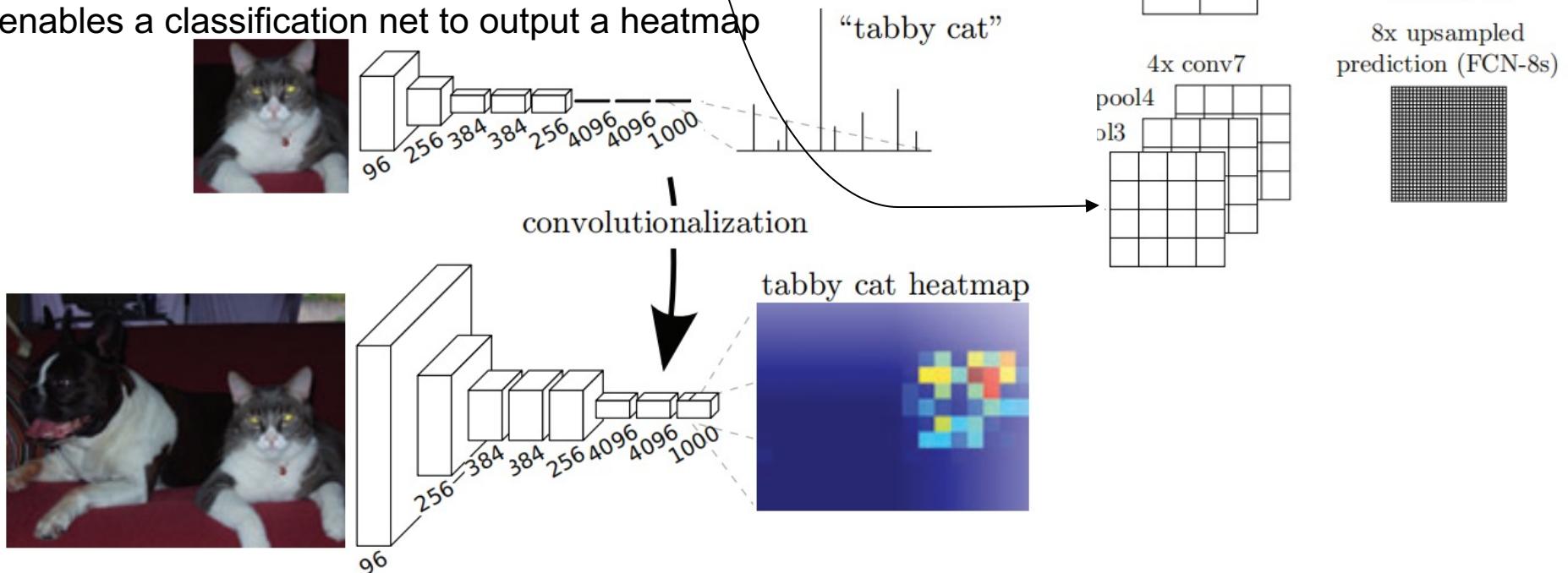


Figure 1. FCN Architecture

Figure 2. Transforming fully connected layers into convolution layers enables a classification net to output a heatmap



Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

5 Deep Learning based Methods

U-Net

Combines spatial information from the down-sampling path with the up-sampling path to retain good spatial information

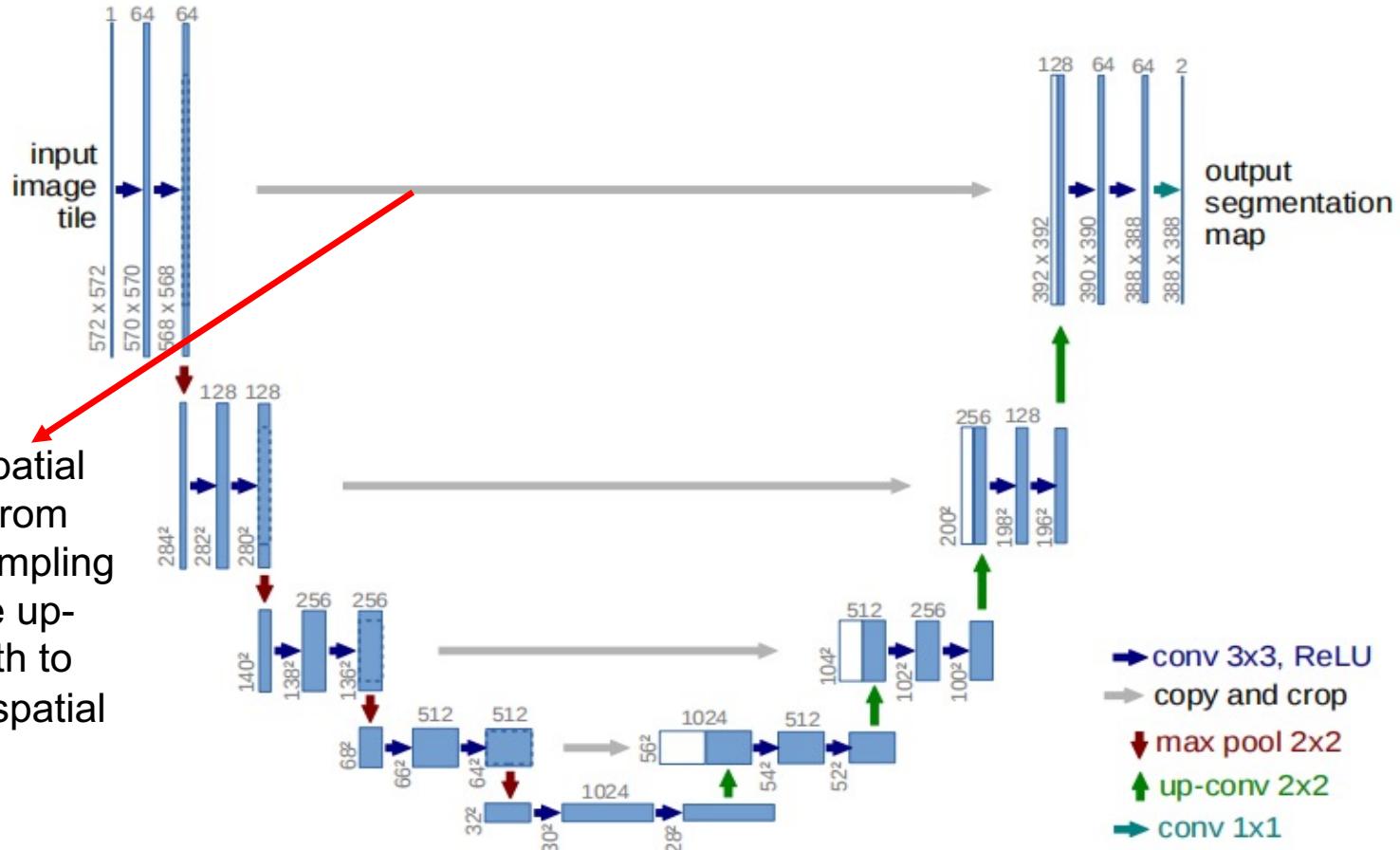


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

5 Deep Learning based Methods

U-Net

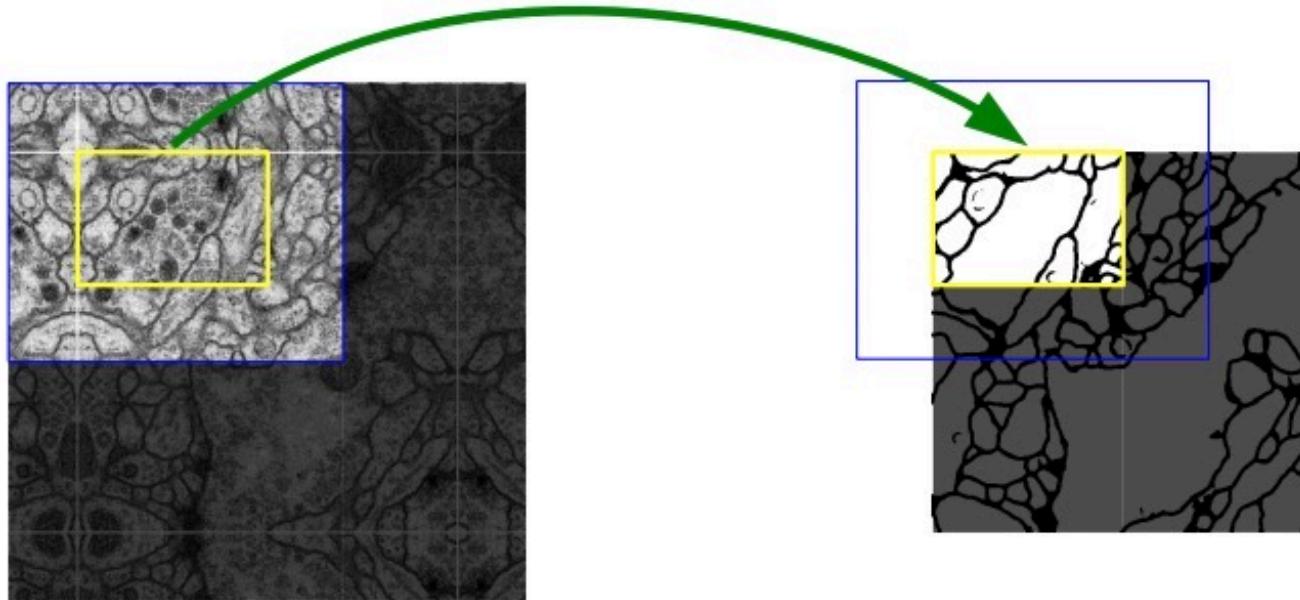


Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

5 Deep Learning based Methods

Results of U-Net

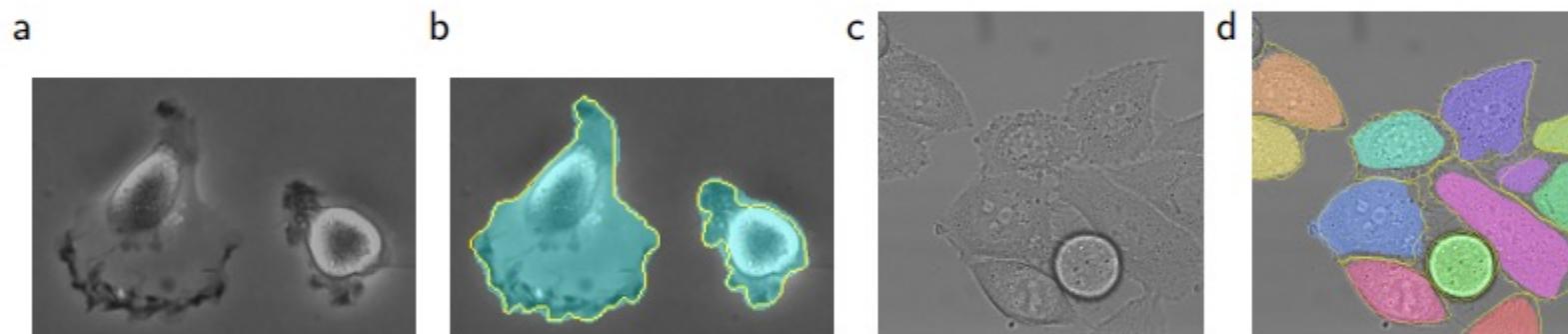


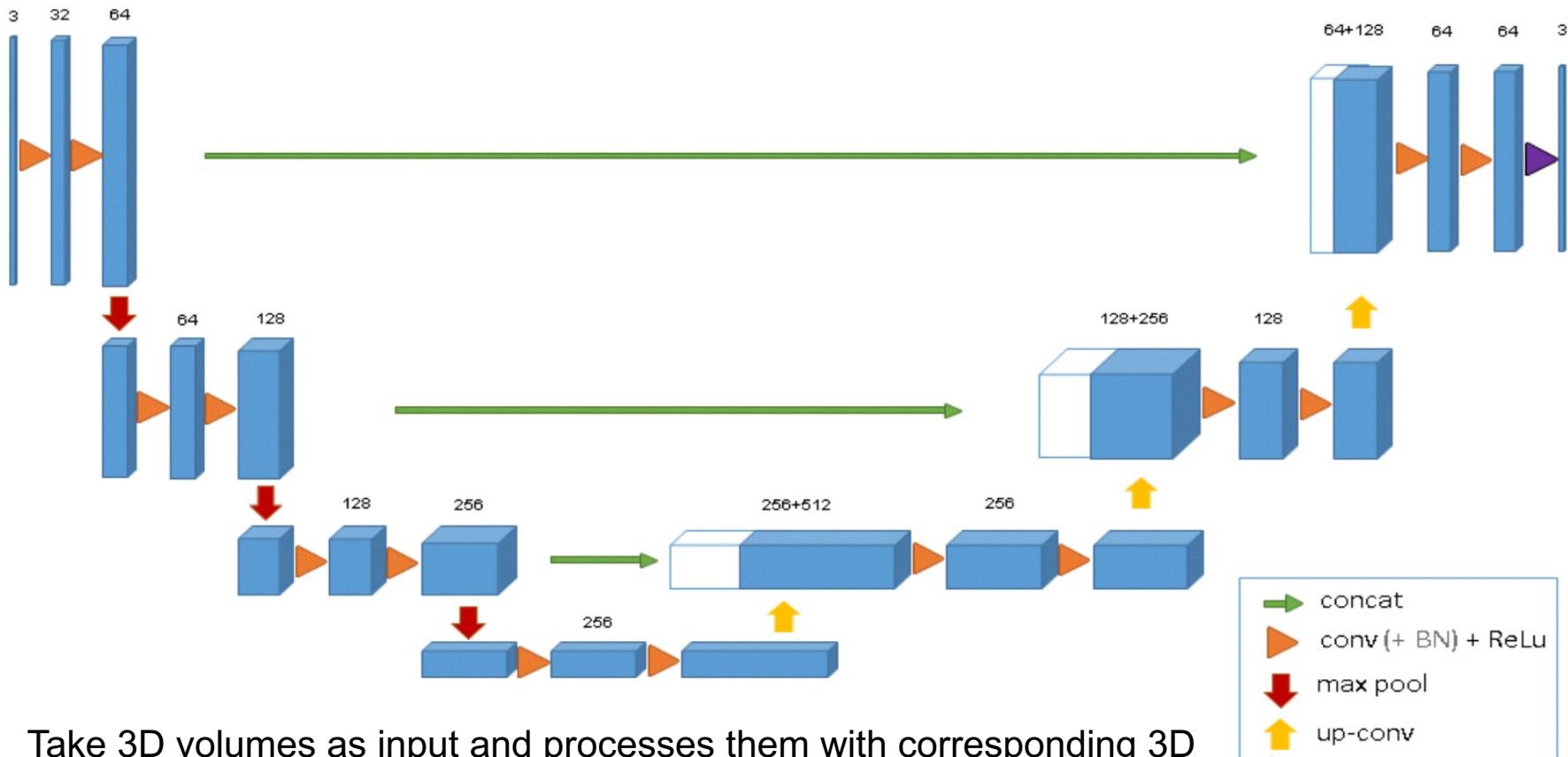
Fig. 4. Result on the ISBI cell tracking challenge.

Table 2. Segmentation results (IOU) on the ISBI cell tracking challenge 2015.

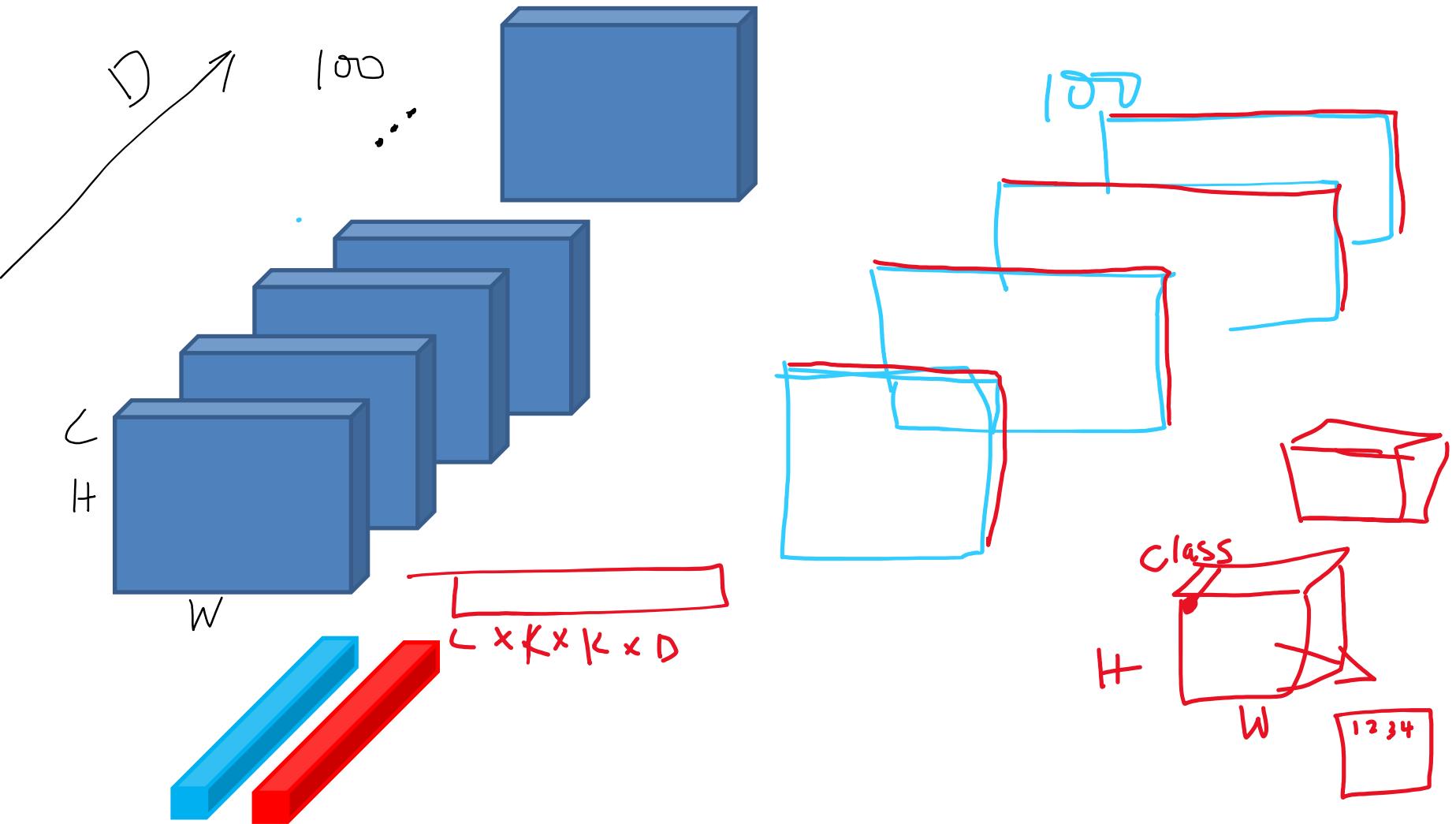
Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best 2015	0.83	0.46
u-net (2015)	0.9203	0.7756

5 Deep Learning based Methods

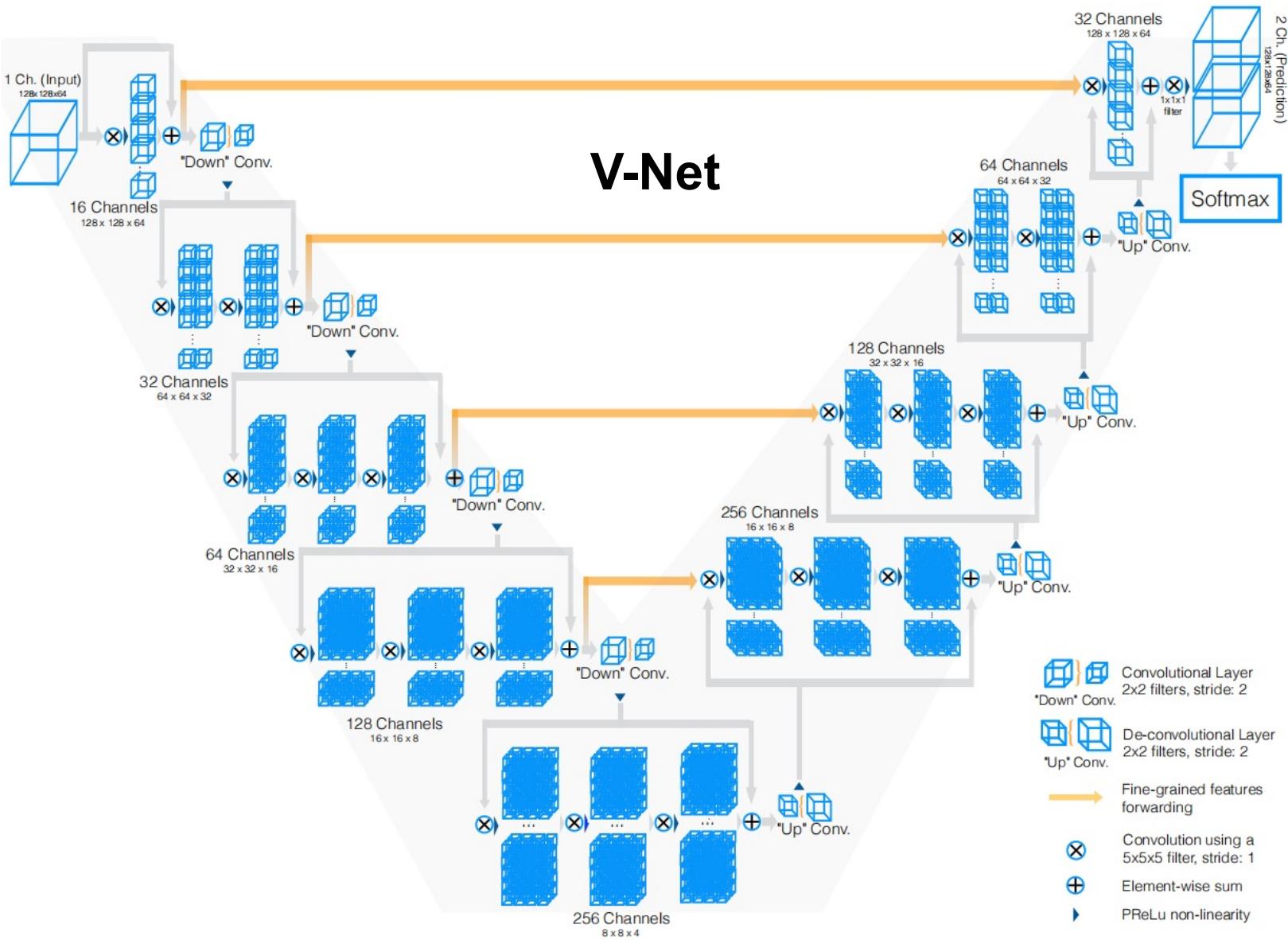
3D U-Net



Take 3D volumes as input and processes them with corresponding 3D operations, in particular, 3D convolutions, 3D max pooling, and 3D up-convolutional layers



5 Deep Learning based Methods



5 Deep Learning based Methods

Results of V-Net

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

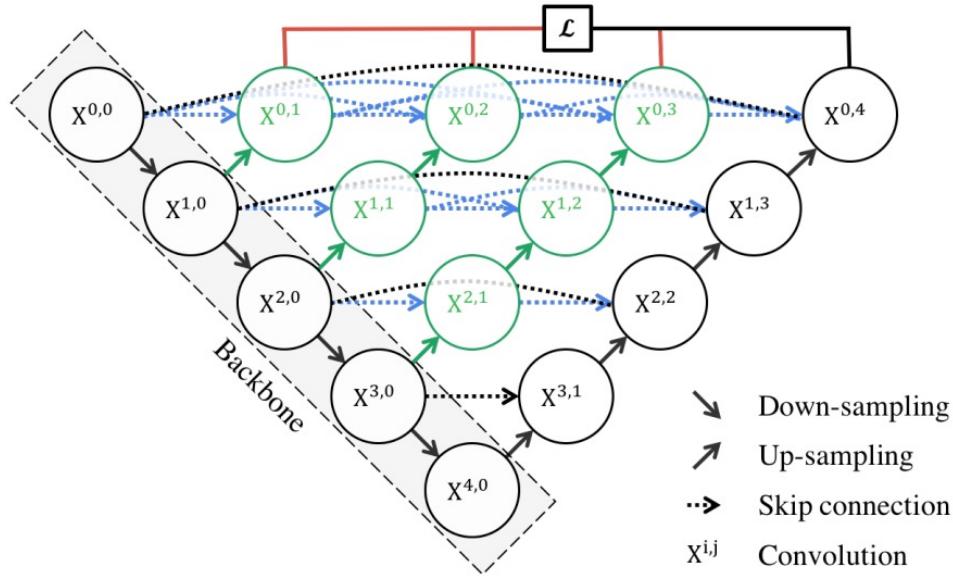
Algorithm	Avg. Dice	Avg. Hausdorff distance	Score on challenge task
V-Net + Dice-based loss	0.869 ± 0.033	5.71 ± 1.20 mm	82.39
V-Net + mult. logistic loss	0.739 ± 0.088	10.55 ± 5.38 mm	63.30
Imorphics [18]	0.879 ± 0.044	5.935 ± 2.14 mm	84.36
ScrAutoProstate	0.874 ± 0.036	5.58 ± 1.49 mm	83.49
SBIA	0.835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

Milletari F, Navab N, Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C]//2016 fourth international conference on 3D vision (3DV). IEEE, 2016: 565-571.

5 Deep Learning based Methods

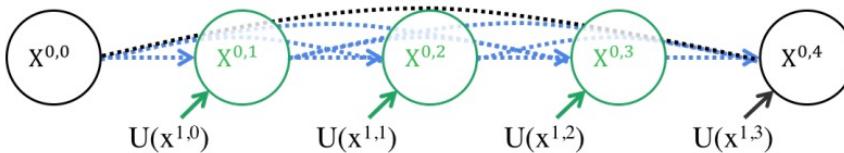
Other U-Net Variants: Unet++

$$x^{i,j} = \begin{cases} \mathcal{H}(x^{i-1,j}), & j = 0 \\ \mathcal{H}\left(\left[x^{i,k}\right]_{k=0}^{j-1}, \mathcal{U}(x^{i+1,j-1})\right), & j > 0 \end{cases}$$



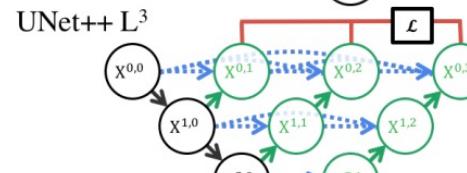
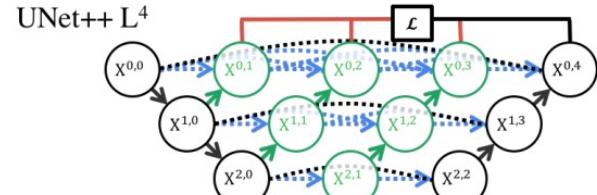
(a)

$$x^{0,1} = H[x^{0,0}, U(x^{1,0})] \quad x^{0,2} = H[x^{0,0}, x^{0,1}, U(x^{1,1})] \quad x^{0,3} = H[x^{0,0}, x^{0,1}, x^{0,2}, U(x^{1,2})]$$



(b)

$$x^{0,4} = H[x^{0,0}, x^{0,1}, x^{0,2}, x^{0,3}, U(x^{1,3})]$$



(c)

5 Deep Learning based Methods

Results of Unet++

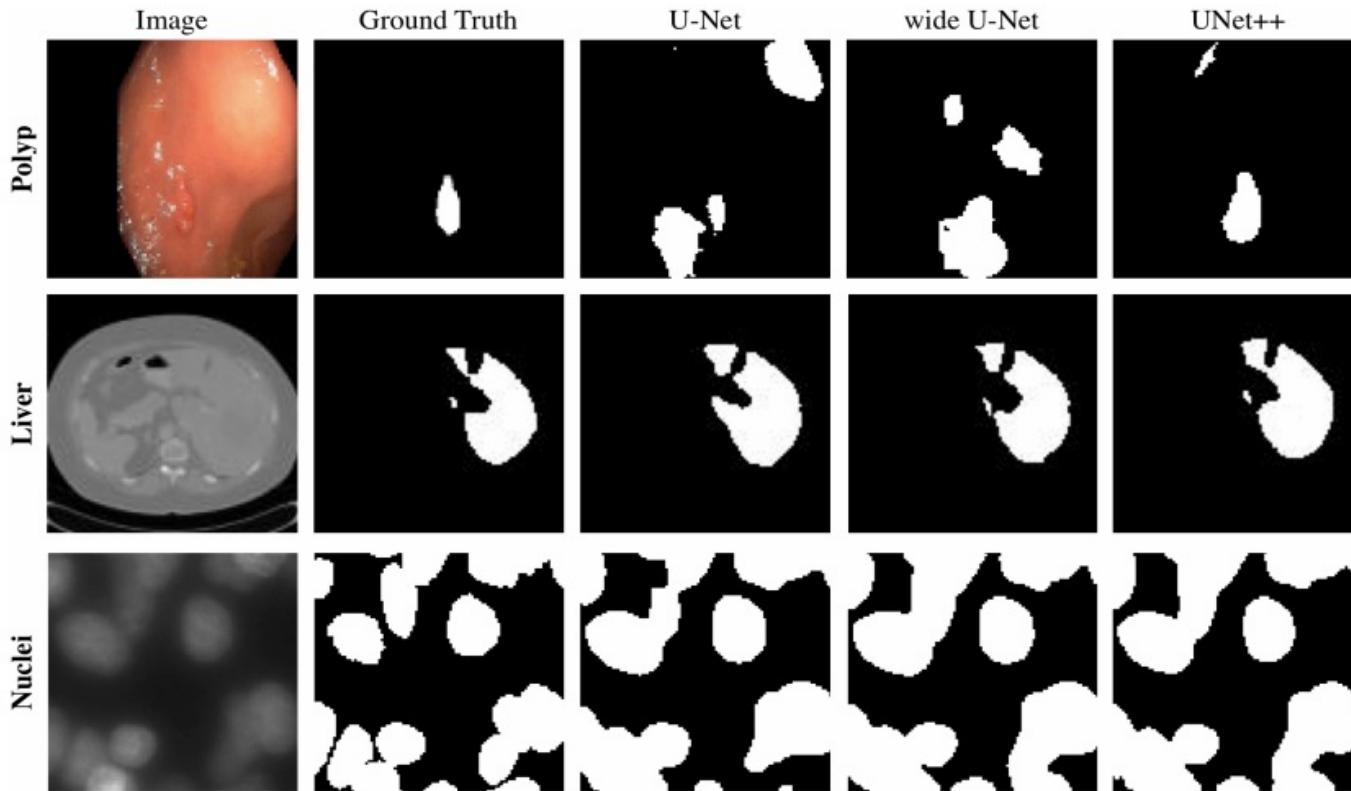


Fig. 2: Qualitative comparison between U-Net, wide U-Net, and UNet++, showing segmentation results for polyp, liver, and cell nuclei datasets (2D-only for a distinct visualization).

Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]//Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Cham, 2018: 3-11.

5 Deep Learning based Methods

Results of Unet++

Table 3: Segmentation results (IoU: %) for U-Net, wide U-Net and our suggested architecture UNet++ with and without deep supervision (DS).

Architecture	Params	Dataset			
		cell nuclei	colon polyp	liver	lung nodule
U-Net [9]	7.76M	90.77	30.08	76.62	71.47
Wide U-Net	9.13M	90.92	30.14	76.58	73.38
UNet++ w/o DS	9.04M	92.63	33.45	79.70	76.44
UNet++ w/ DS	9.04M	92.52	32.12	82.90	77.21

UNet++ with deep supervision achieves an average IoU gain of 3.9 and 3.4 points over U-Net and wide U-Net.

Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation[M]//Deep learning in medical image analysis and multimodal learning for clinical decision support. Springer, Cham, 2018: 3-11.

5 Deep Learning based Methods

Other U-Net Variants: **Attention U-Net**

- Attention in U-Net is a method to highlight only the relevant activations during training

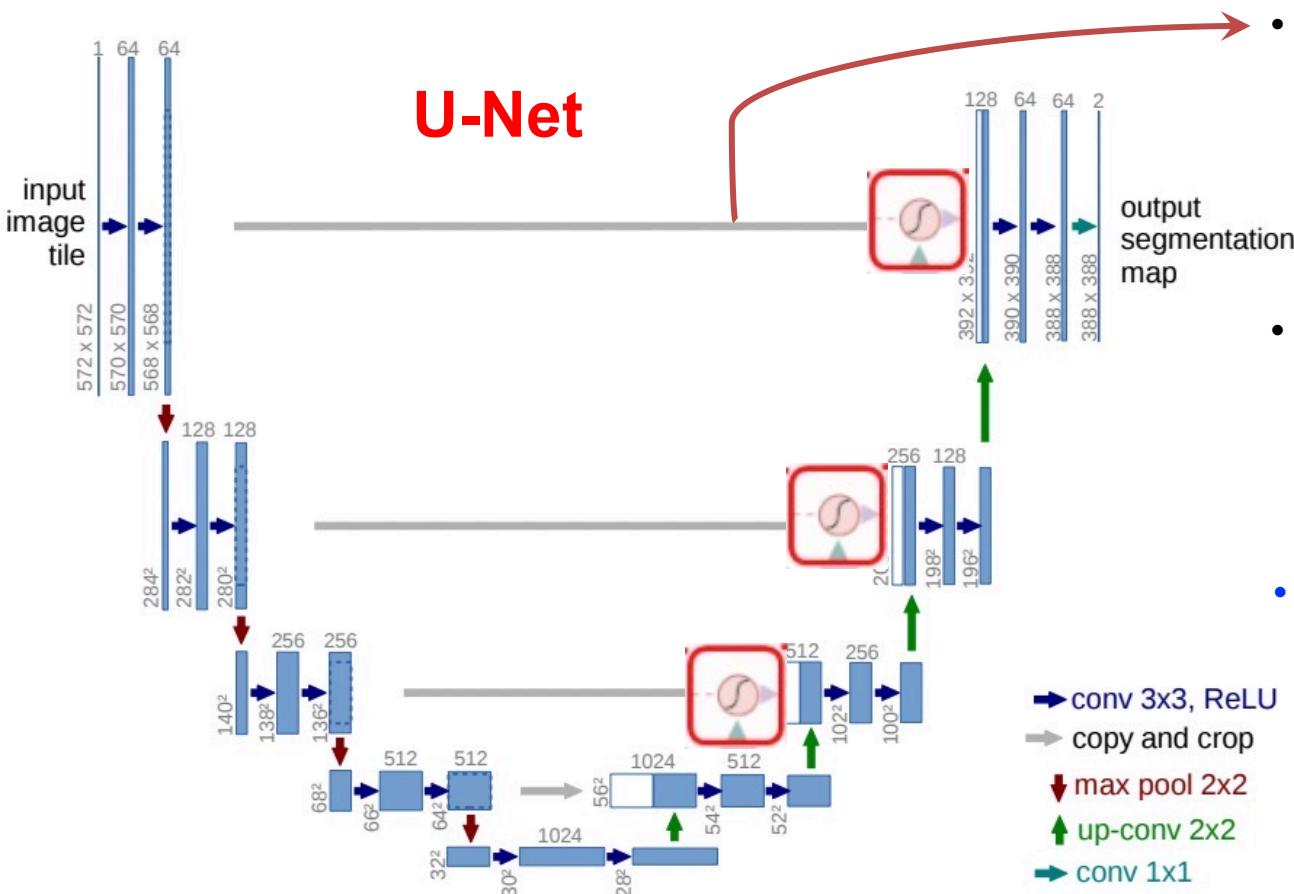
Soft Attention

- Weighting different parts of the image
- Relevant parts of image get large weights and less relevant parts get small weights
- During training, the weights also get trained making the model pay more attention to relevant regions
- Basically, it adds weights to pixels based on the relevance

Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv:1804.03999, 2018.

5 Deep Learning based Methods

Why adding Attention to U-Net



- Combines spatial information from the down-sampling path (i.e. encoder) with the up-sampling path (i.e. decoder) to retain good spatial information
- But this process brings along the poor feature representation from the initial layers of the encoder
- Use soft attention at the skip connections will actively suppress activations at irrelevant regions

Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv:1804.03999, 2018.

5 Deep Learning based Methods

Other U-Net Variants: Attention U-Net

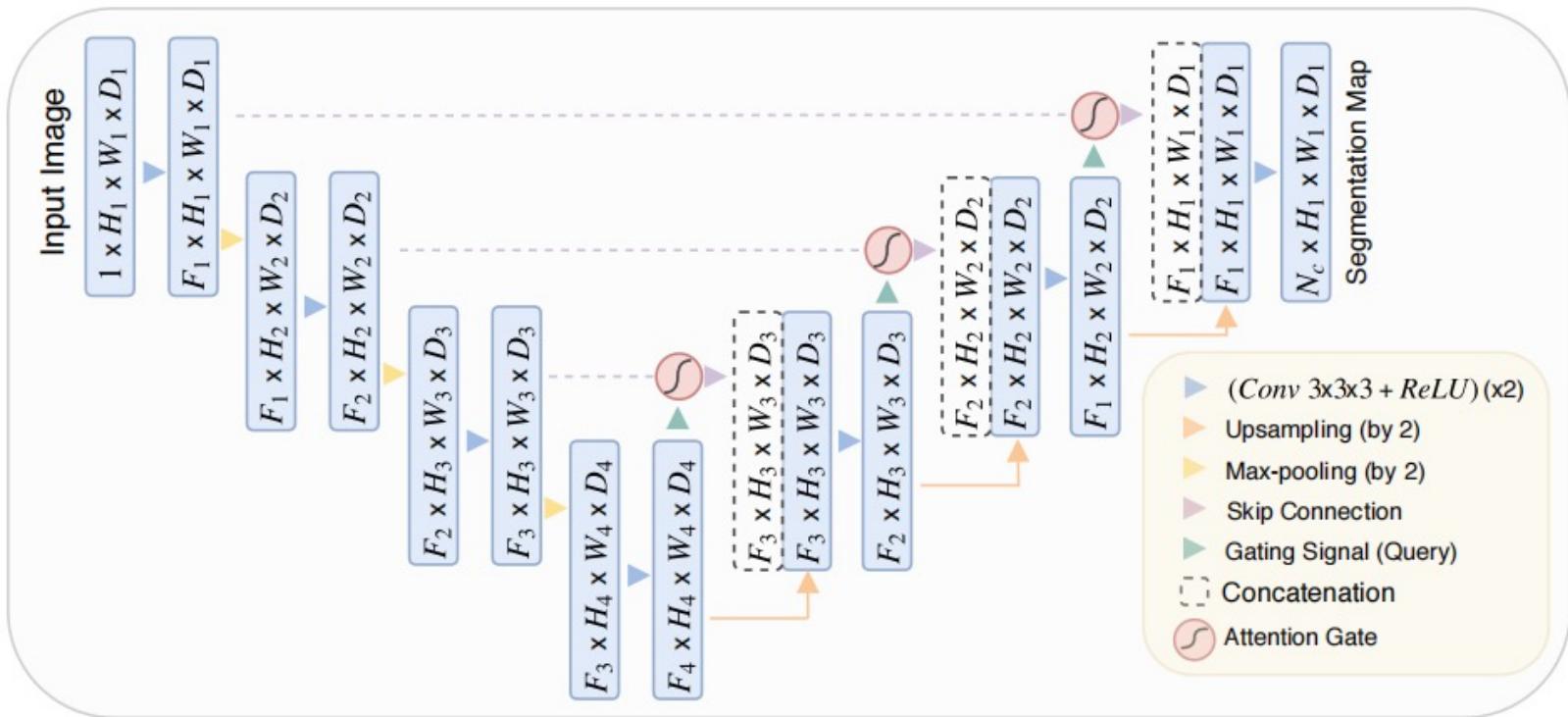


Figure 1: A block diagram of the proposed Attention U-Net segmentation model. Input image is progressively filtered and downsampled by factor of 2 at each scale in the encoding part of the network (e.g. $H_4 = H_1/8$). N_c denotes the number of classes. Attention gates (AGs) filter the features propagated through the skip connections. Schematic of the AGs is shown in Figure 2. Feature selectivity in AGs is achieved by use of contextual information (gating) extracted in coarser scales.

Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv:1804.03999, 2018.

5 Deep Learning based Methods

Attention U-Net

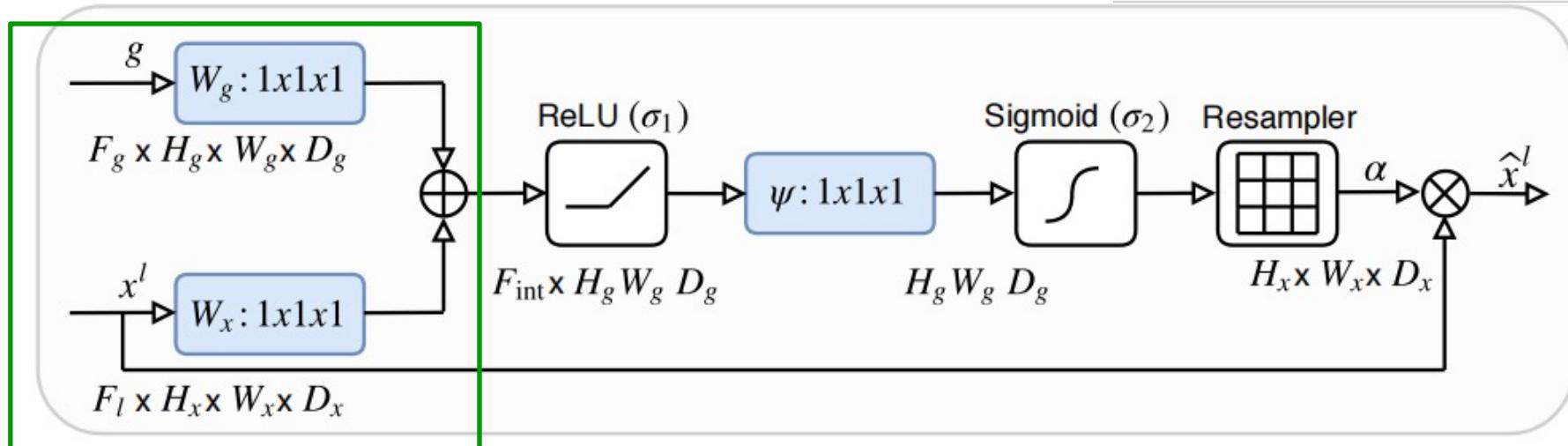
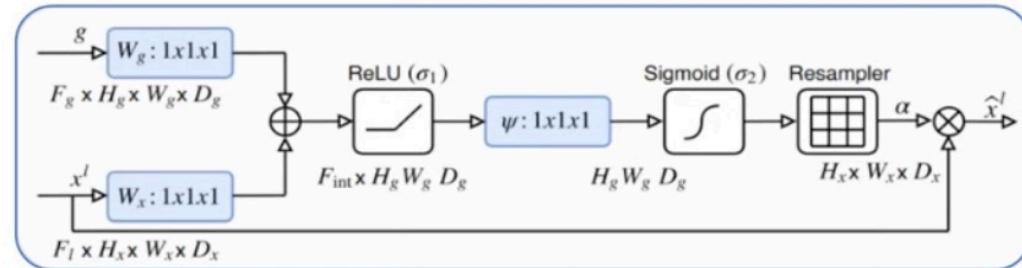


Figure 2: Schematic of the proposed additive attention gate (AG). Input features (x^l) are scaled with attention coefficients (α) computed in AG. Spatial regions are selected by analysing both the activations and contextual information provided by the gating signal (g) which is collected from a coarser scale. Grid resampling of attention coefficients is done using trilinear interpolation.

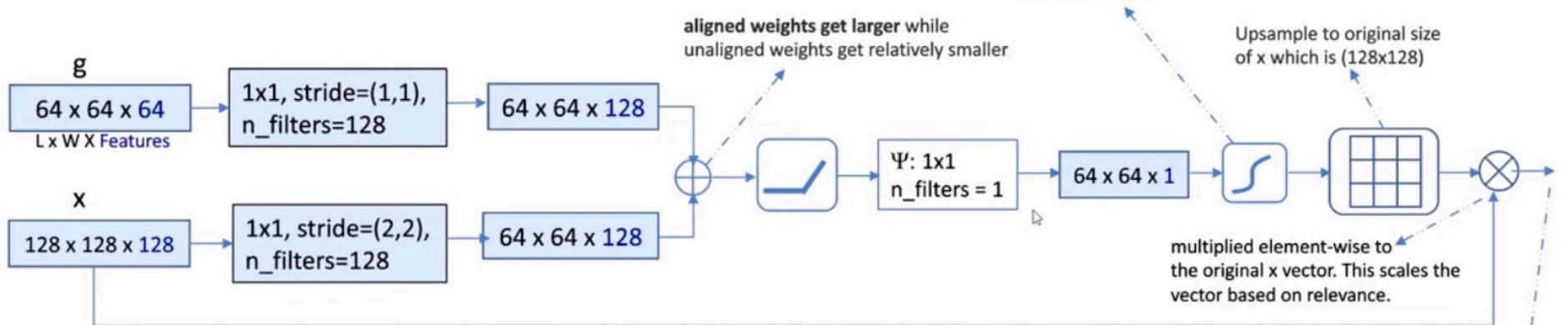
Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv:1804.03999, 2018.

5 Deep Learning based Methods

Attention U-Net



Scales all weights to
between 0 and 1 (1 being
more important)



Credit: <https://www.youtube.com/watch?v=KOF38xAvo8I>

5 Deep Learning based Methods

Attention U-Net

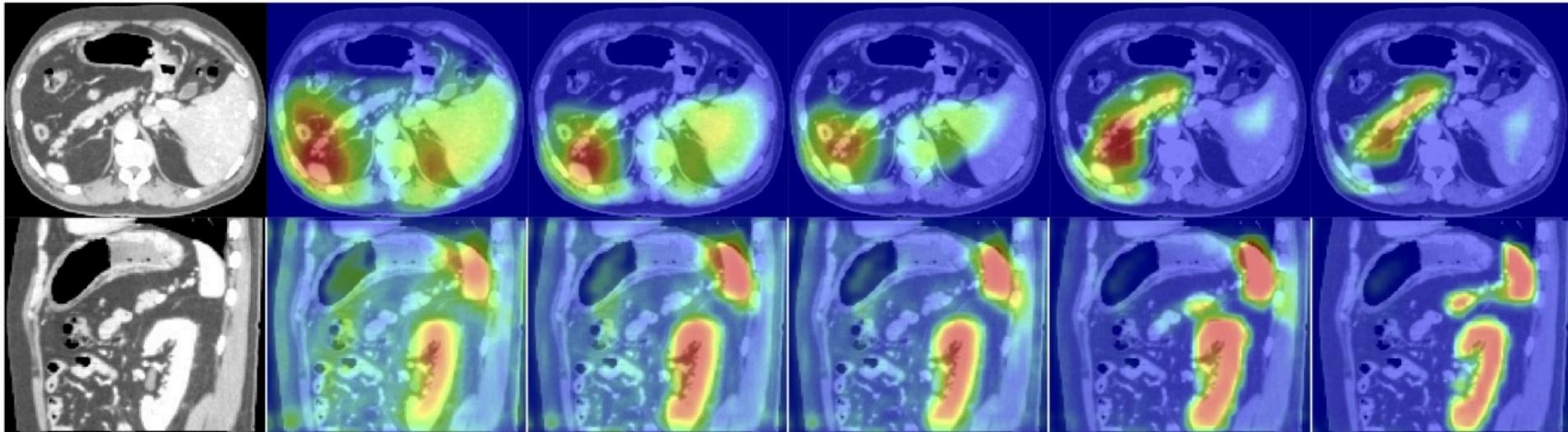
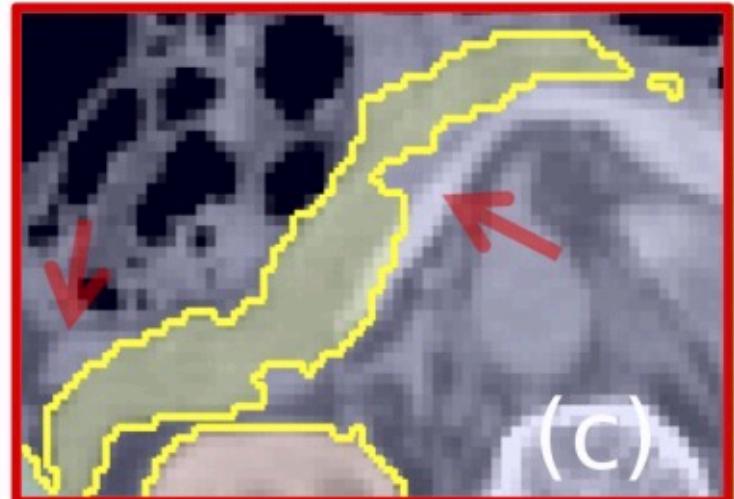
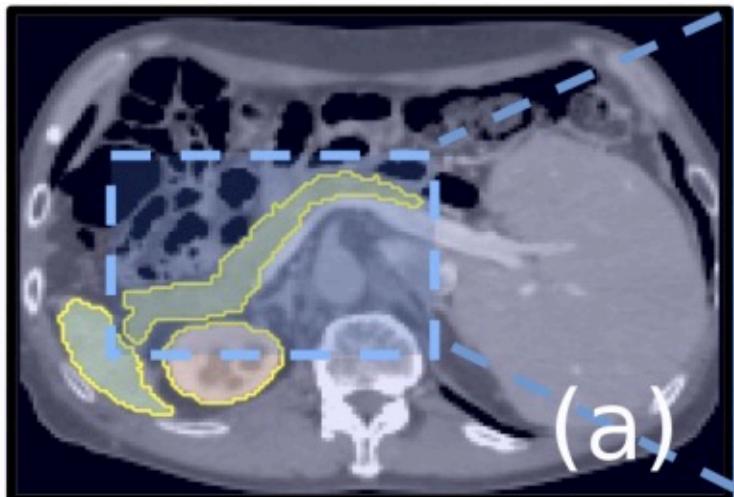


Figure 4: The figure shows the attention coefficients ($\alpha^{l_{s2}}, \alpha^{l_{s3}}$) across different training epochs (3, 6, 10, 60, 150). The images are extracted from sagittal and axial planes of a 3D abdominal CT scan from the testing dataset. The model gradually learns to focus on the pancreas, kidney, and spleen.

5 Deep Learning based Methods

Attention U-Net



The ground-truth pancreas segmentation (a) is highlighted in blue (b).

(c) U-Net model prediction

(d) Attention U-Net model prediction

Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: Learning where to look for the pancreas[J]. arXiv preprint arXiv:1804.03999, 2018.

5 Deep Learning based Methods

Attention U-Net

Table 1: Multi-class CT abdominal segmentation results obtained on the *CT-150* dataset: The results are reported in terms of **Dice score (DSC)** and mesh surface to surface distances (S2S). These distances are reported only for the pancreas segmentations. The proposed Attention U-Net model is benchmarked against the standard U-Net model for different training and testing splits. Inference time (forward pass) of the models are computed for input tensor of size $160 \times 160 \times 96$. Statistically significant results are highlighted in bold font.

Method (Train/Test Split)	U-Net (120/30)	Att U-Net (120/30)	U-Net (30/120)	Att U-Net (30/120)
Pancreas DSC	0.814 ± 0.116	0.840 ± 0.087	0.741 ± 0.137	0.767 ± 0.132
Pancreas Precision	0.848 ± 0.110	0.849 ± 0.098	0.789 ± 0.176	0.794 ± 0.150
Pancreas Recall	0.806 ± 0.126	0.841 ± 0.092	0.743 ± 0.179	0.762 ± 0.145
Pancreas S2S Dist (mm)	2.358 ± 1.464	1.920 ± 1.284	3.765 ± 3.452	3.507 ± 3.814
Spleen DSC	0.962 ± 0.013	0.965 ± 0.013	0.935 ± 0.095	0.943 ± 0.092
Kidney DSC	0.963 ± 0.013	0.964 ± 0.016	0.951 ± 0.019	0.954 ± 0.021
Number of Params	5.88 M	6.40 M	5.88 M	6.40 M
Inference Time	0.167 s	0.179 s	0.167 s	0.179 s

5 Deep Learning based Methods

Medical image segmentation + Transformer

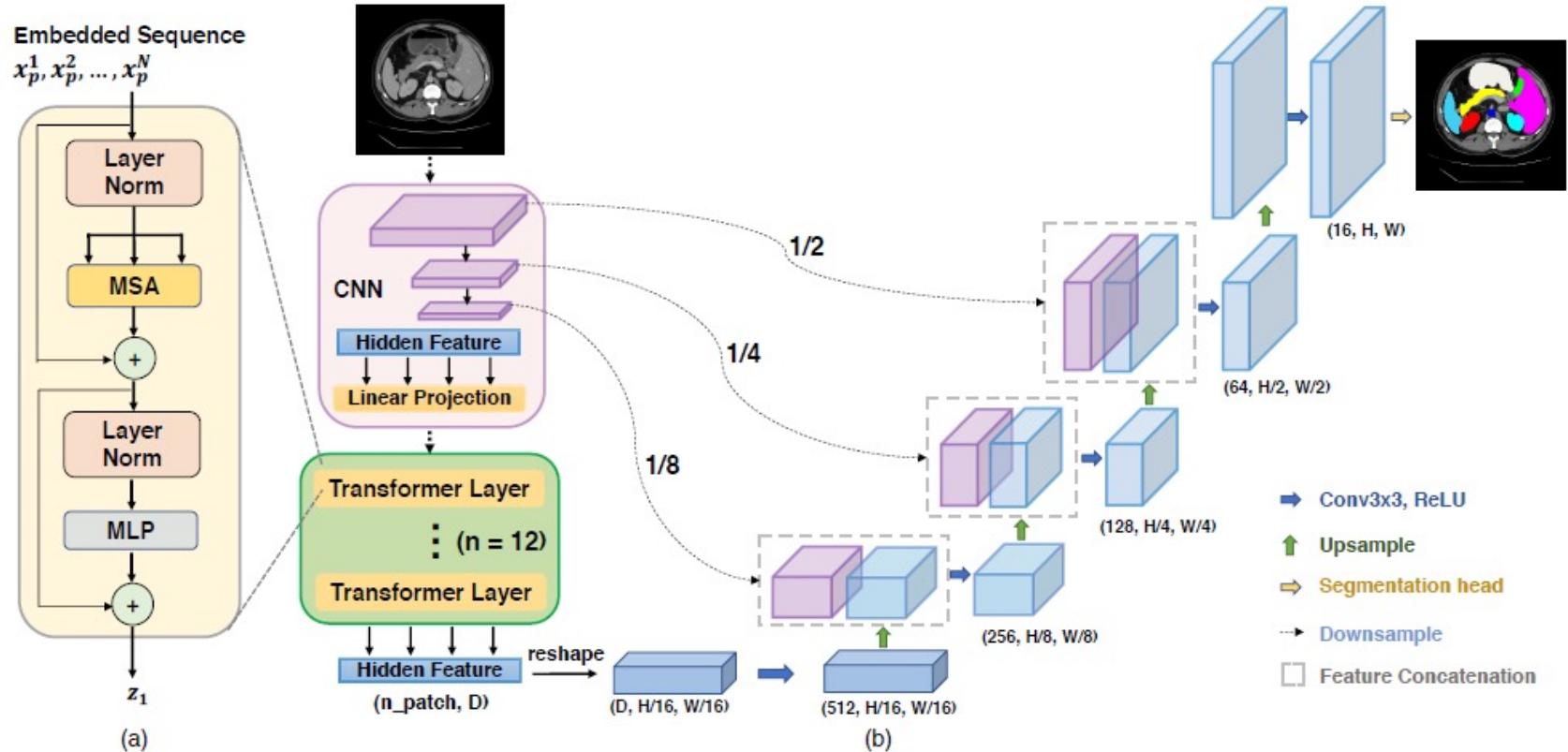
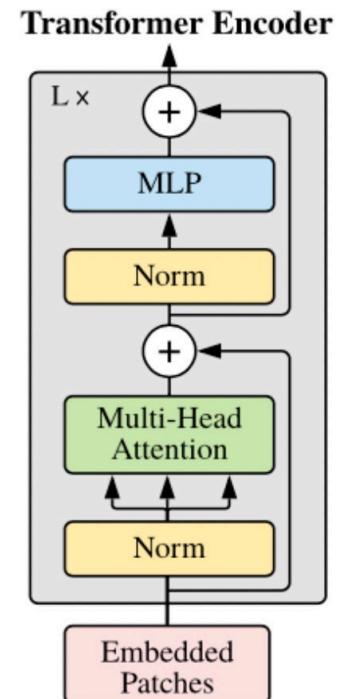


Fig. 1: Overview of the framework. (a) schematic of the Transformer layer; (b) architecture of the proposed TransUNet.

Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation[J]. arXiv preprint arXiv:2102.04306, 2021.

5 Deep Learning based Methods

Vision Transformer



<https://github.com/lucidrains/vit-pytorch>

5 Deep Learning based Methods

Medical image segmentation + Transformer

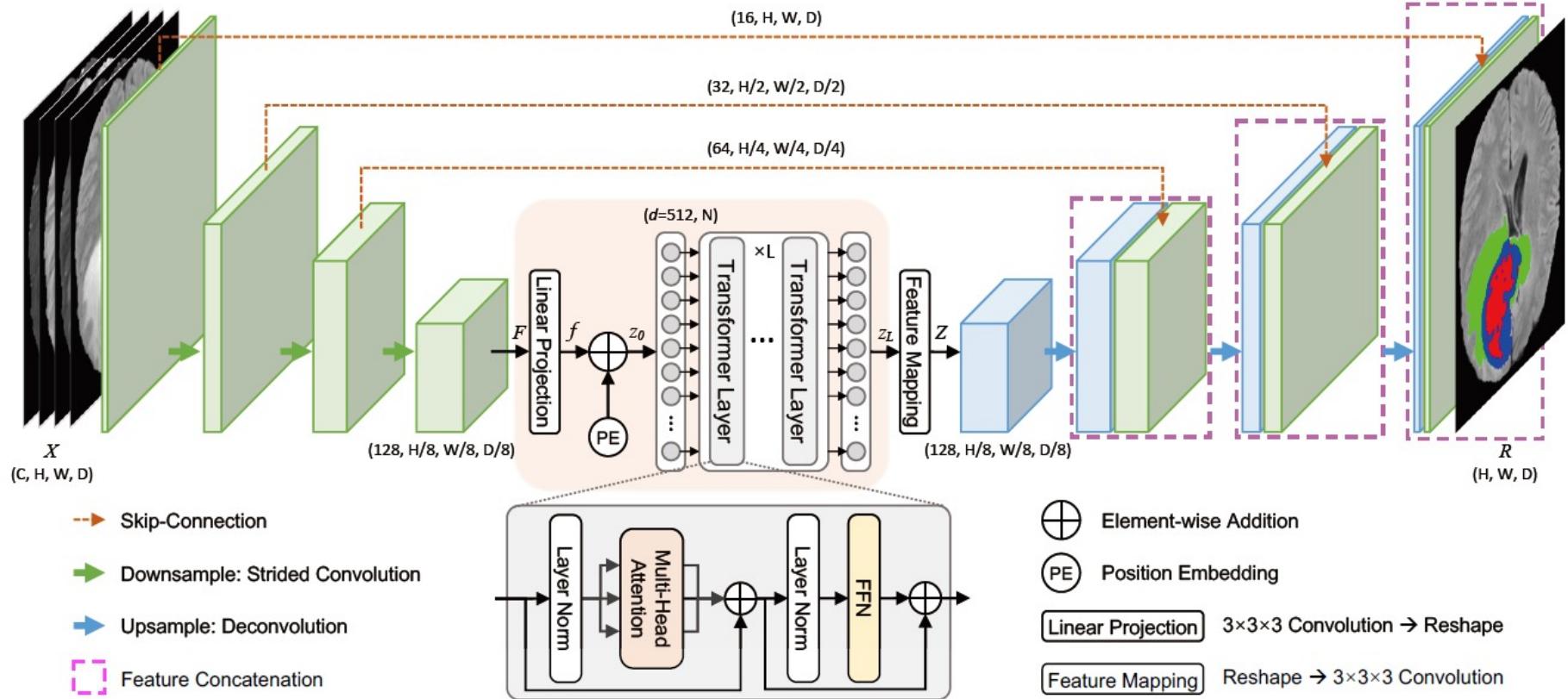


Fig. 1. Overall architecture of the proposed TransBTS.

Wang W, Chen C, Ding M, et al. TransBTS: Multimodal Brain Tumor Segmentation Using Transformer[J]. arXiv preprint arXiv:2103.04430, 2021.

5 Deep Learning based Methods

Medical image segmentation + Transformer

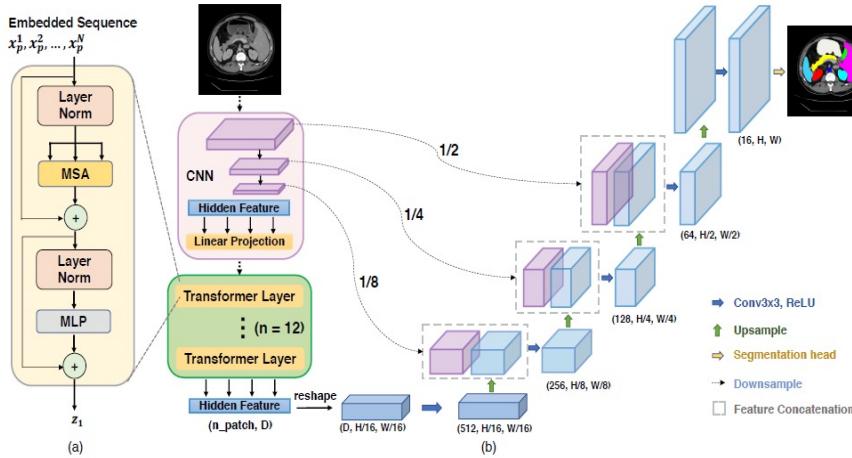


Fig. 1: Overview of the framework. (a) schematic of the Transformer layer; (b) architecture of the proposed TransUNet.

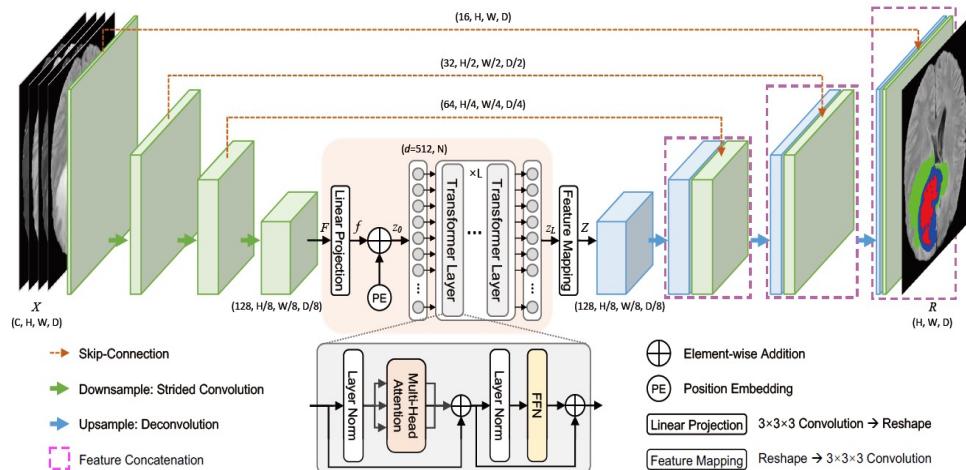


Fig. 1. Overall architecture of the proposed TransBTS.

- 2D slice-by-slice
- Use Pretrain model

VS

3D volume

No Pretraining

Chen J, Lu Y, Yu Q, et al. Transunet: Transformers make strong encoders for medical image segmentation[J]. arXiv preprint arXiv:2102.04306, 2021.

Wang W, Chen C, Ding M, et al. TransBTS: Multimodal Brain Tumor Segmentation Using Transformer[J]. arXiv preprint

5 Deep Learning based Methods

Medical image segmentation + Transformer

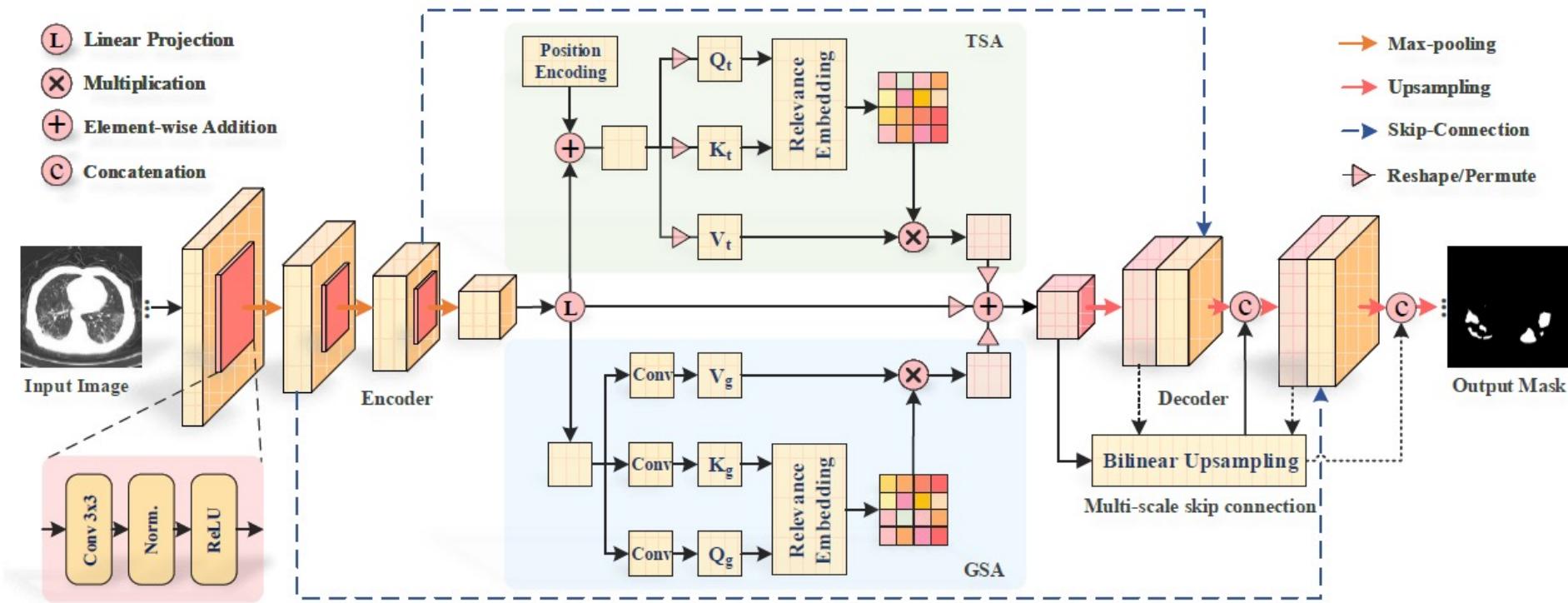


Fig. 2. Illustration of the proposed TransAttUnet for automatic medical image segmentation. (a) Both TSA and GSA mechanisms are embedded into the SAA module to model the long-range interactions and global spatial relationships. (b) The multi-scale skip connections between decoder blocks are designed to aggregate the downsampled features of varying semantic scales by progressive upsampling, concatenation, and convolution.

Chen B, Liu Y, Zhang Z, et al. TransAttUnet: Multi-level Attention-guided U-Net with Transformer for Medical Image Segmentation[J]. arXiv preprint arXiv:2107.05274, 2021.

Thank you!

Question?