
CAP 5516

Medical Image Computing

(Spring 2025)

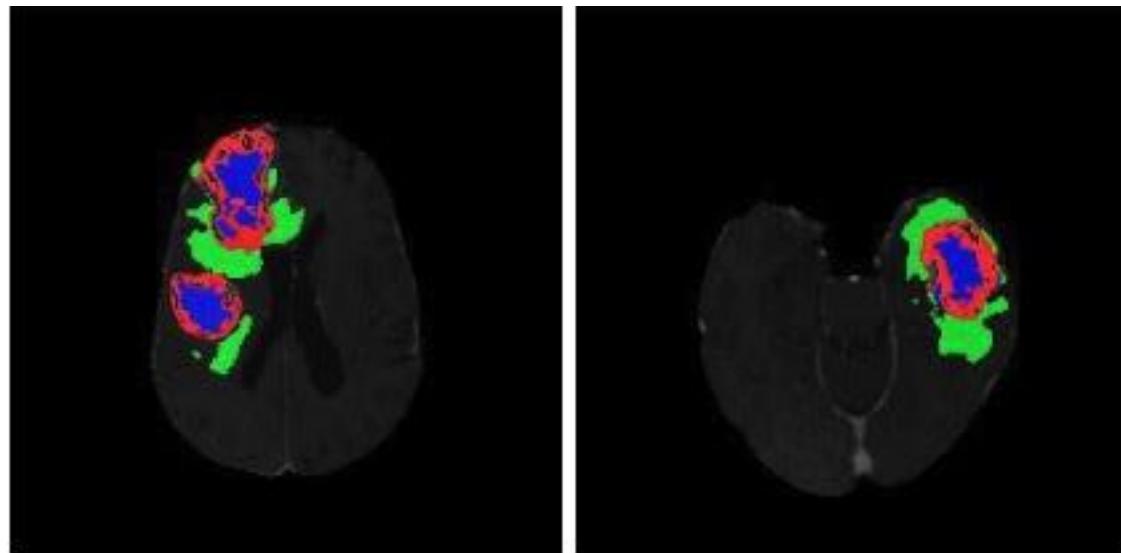
Dr. Chen Chen
Associate Professor

Center for Research in Computer Vision (CRCV)
University of Central Florida
Office: HEC 221

Email: chen.chen@crcv.ucf.edu

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

Lecture 10: Medical Image Segmentation



Medical Image Segmentation

- 1 Problem Definition
- 2 Evaluation Metrics
- 3 Datasets
- 4 Traditional Methods
- 5 Deep Learning based Methods

Introduction to Image Segmentation

- **Image Segmentation**

- Group pixels into regions that share some similar properties

Superpixels
(Ren ICCV 2003)



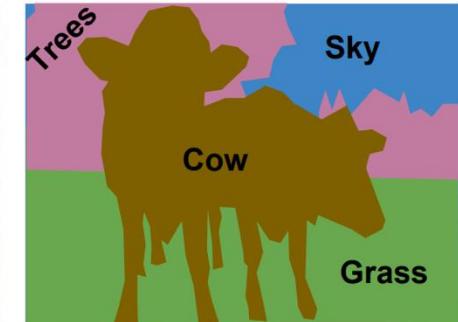
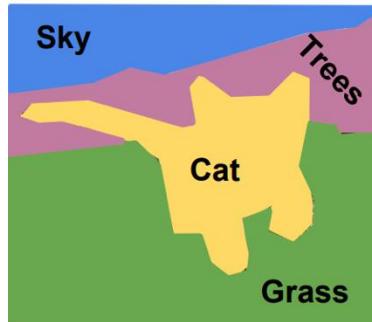
- **Segmenting images into meaningful objects**

- Object-level segmentation: accurate localization and recognition



Introduction to Image Segmentation

- Semantic Segmentation
 - Labelling every pixel in an image
- A key part of Scene Understanding



Source: http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf

Applications

- Autonomous navigation
- Assisting the partially sighted
- Medical diagnosis
- Image editing
- ⋮



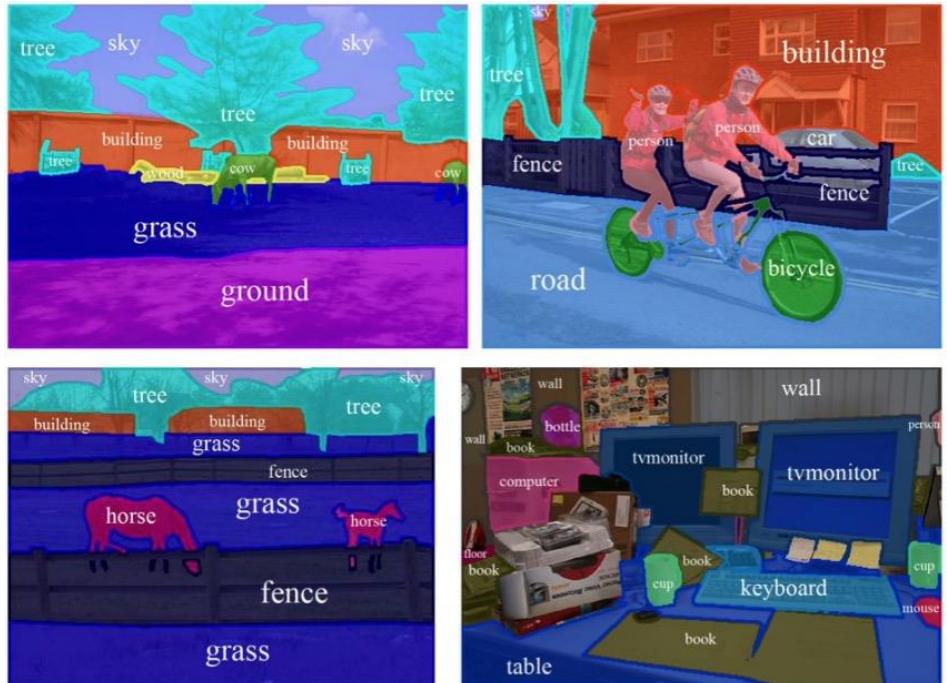
Image editing and composition (Xu, 2016)



Source: http://www.cs.toronto.edu/~tingwuwang/semantic_segmentation.pdf

Semantic segmentation vs. instance segmentation

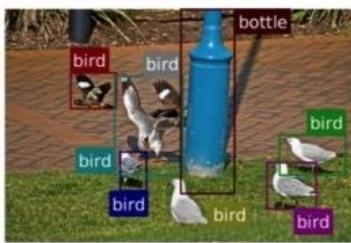
- Semantic segmentation
 - Label every pixel: recognize the class of every pixel
 - Do not differentiate instances



Mottaghi et al, "The role of context for object detection and semantic segmentation in the wild", CVPR 2014 .

Semantic segmentation vs. instance segmentation

- Instance segmentation
 - Detect instances, categorize and label every pixel
 - Labels are **class-aware** and **instance-aware**



Object detection



Semantic Segm.



Instance segm.

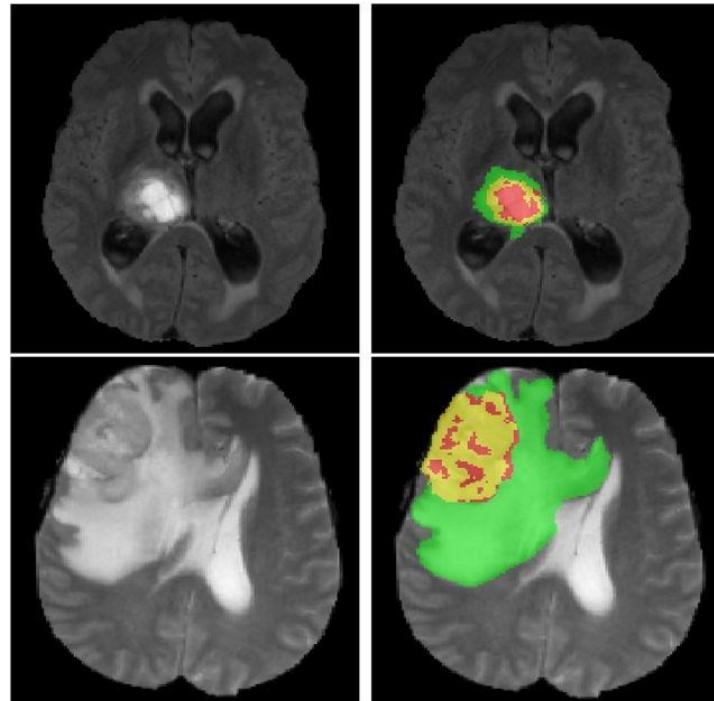


Ground truth

1 Medical Image Segmentation

For Medical

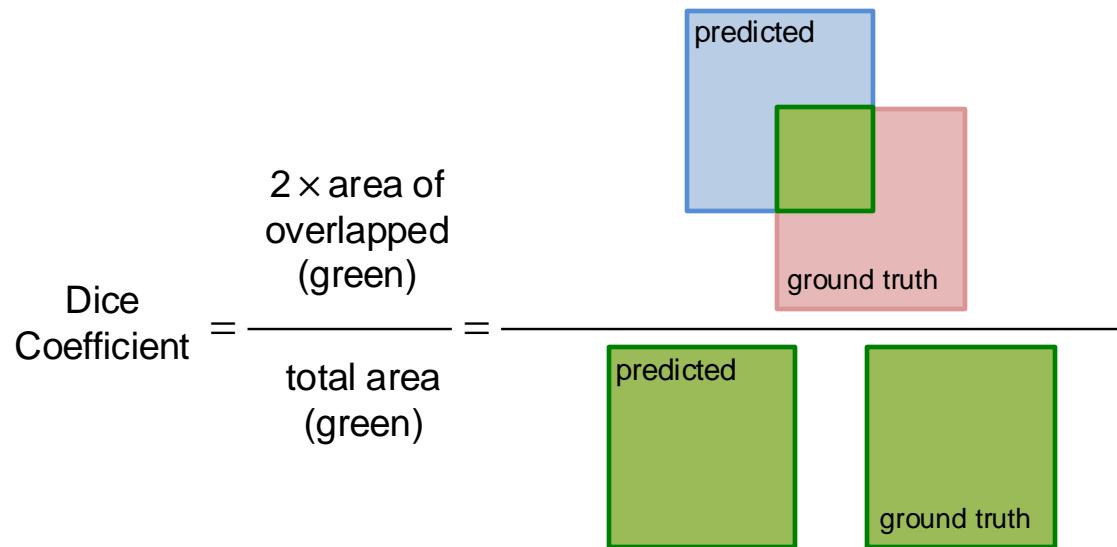
Use computer image processing technology to analyze and process 2D or 3D images to achieve segmentation of human organs, soft tissues and diseased bodies.



2 Evaluation Metrics

Dice index

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



Interpretation: Values range from 0 (no overlap) to 1 (perfect overlap).

2 Evaluation Metrics

Jaccard index(IoU)

Ratio of the intersection to the union of the predicted and true segmentation.

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

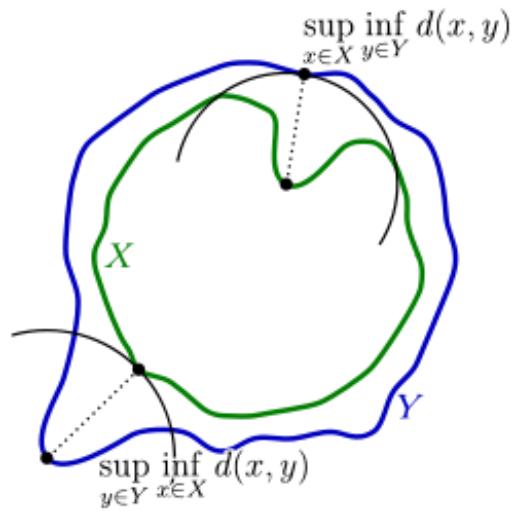
Interpretation: A lower value compared to Dice; also ranges from 0 to 1.

2 Evaluation Metrics

Hausdorff distance

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

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



Definition: Measures the maximum distance of a point in one set to the nearest point in the other set.

Interpretation: Indicates boundary mismatches; lower values indicate closer alignment.

2 Evaluation Metrics

HD95

Standard Hausdorff Distance is **highly sensitive to outliers**. A single misplaced pixel in the prediction can result in a very large HD, even if the rest of the segmentation is perfect.

HD95 addresses this by using the **95th percentile** of distances instead of the maximum:

- Compute all distances between boundary points of A and B .
- Sort these distances and take the **95th percentile value**.
- This ignores the worst 5% of distances (outliers), making the metric more robust.

2 Evaluation Metrics

HD95

When to Use HD95

- **Boundary-sensitive tasks:** Tumor resection, organ-at-risk segmentation.
- **High-stakes applications:** Radiation therapy, surgical navigation.
- **Research benchmarking:** Standard in medical segmentation challenges (e.g., BraTS, KiTS).

2 Evaluation Metrics

Metric	Focus	Strengths	Weaknesses
Dice (DSC)	Overlap	Robust to class imbalance.	Inensitive to boundary errors.
HD95	Boundary accuracy	Captures spatial alignment.	Computationally intensive.
IoU	Overlap	Intuitive for pixel-level accuracy.	Similar limitations to Dice.

2 Evaluation Metrics

<https://docs.monai.io/en/stable/metrics.html>

```
from monai.metrics import compute_hausdorff_distance
hd95 = compute_hausdorff_distance(y_pred, y_true, percentile=95)
```

Generalized Dice Score

```
monai.metrics.compute_generalized_dice(y_pred, y,
                                         include_background=True, weight_type=square, sum_over_classes=False)
```

Mean IoU

```
monai.metrics.compute_iou(y_pred, y, include_background=True,
                           ignore_empty=True)
```

2 Evaluation Metrics

Precision and Recall:

- **Precision:** Proportion of correctly predicted positive pixels over all predicted positives.
- **Recall (Sensitivity):** Proportion of correctly predicted positive pixels over all actual positives.

Additional Metrics:

- **Accuracy:** Less common in segmentation due to class imbalance.
- **Volume Similarity:** Particularly useful in 3D medical segmentation.

2 Evaluation Metrics

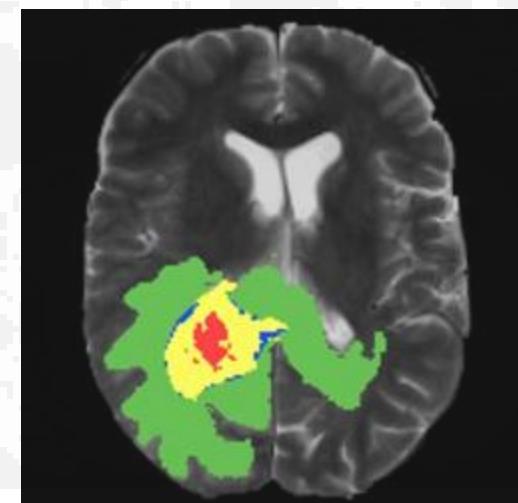
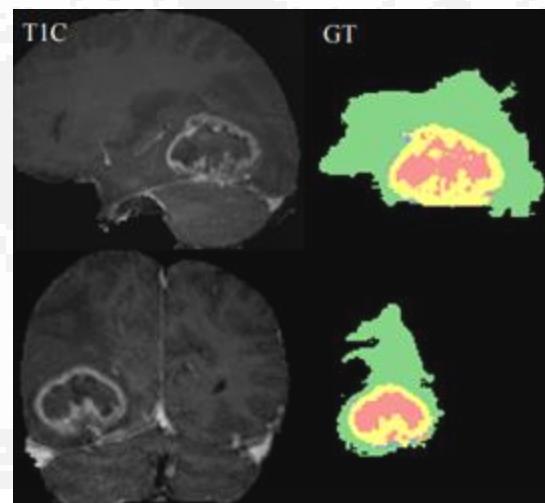
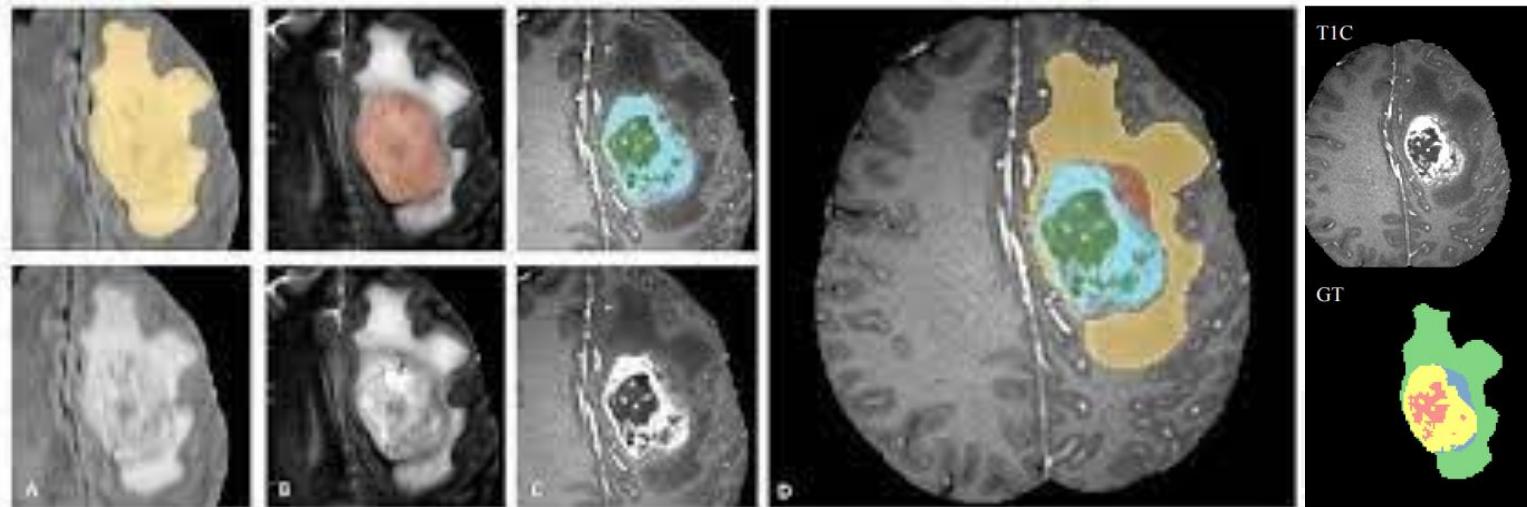
- Each metric reflects different aspects of segmentation performance.
- Trade-offs: For example, a model might have a high Dice score but still miss fine boundary details captured by Hausdorff distance.

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/
HVSMR2016	CMR	Heart	http://segchd.csail.mit.edu/

3 Datasets

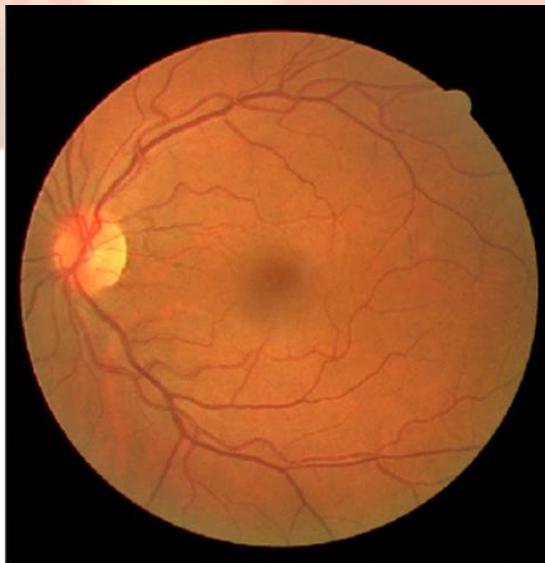
Brain Tumor Segmentation (BraTS)



3 Datasets

Digital retinal images for vessel extraction (DRIVE)

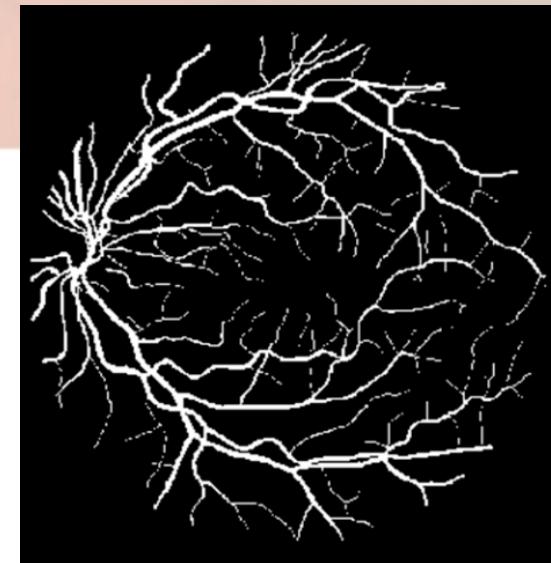
DRIVE



(a)



(b)



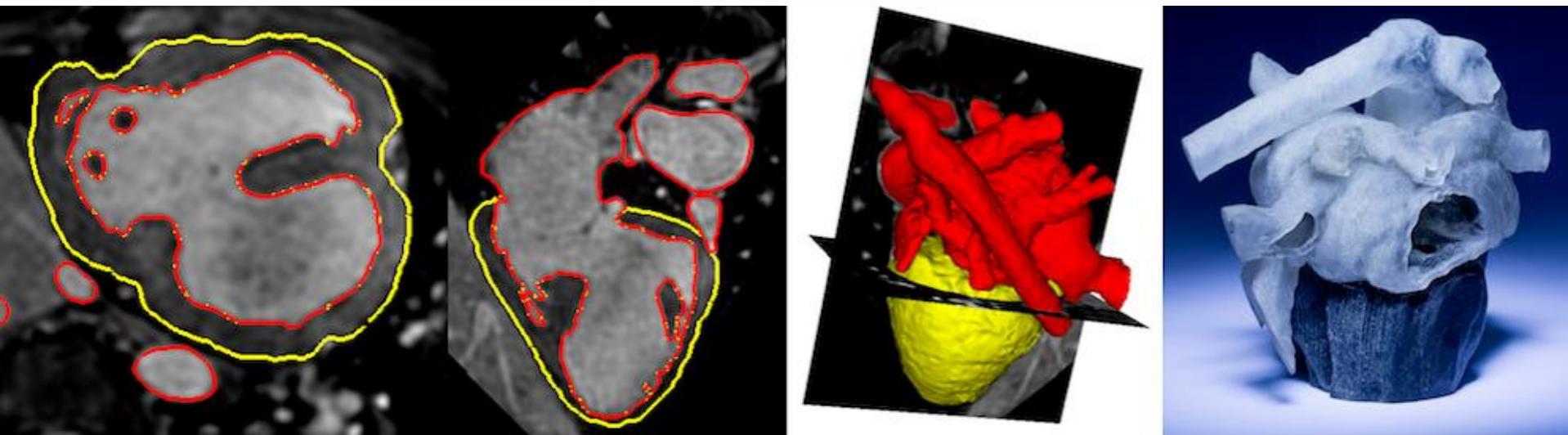
(c)

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

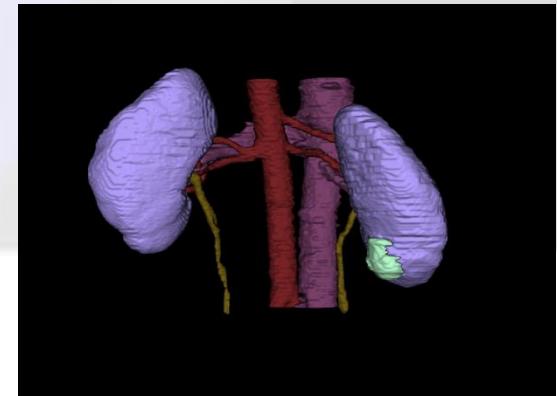
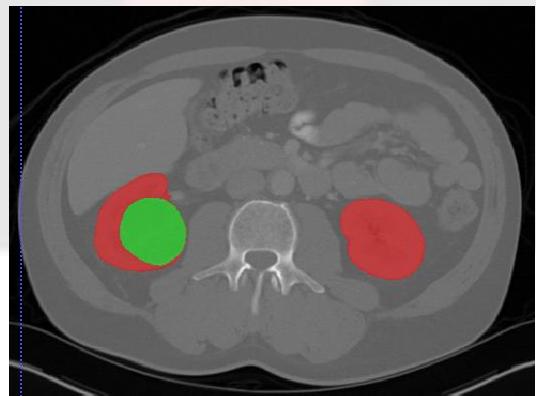
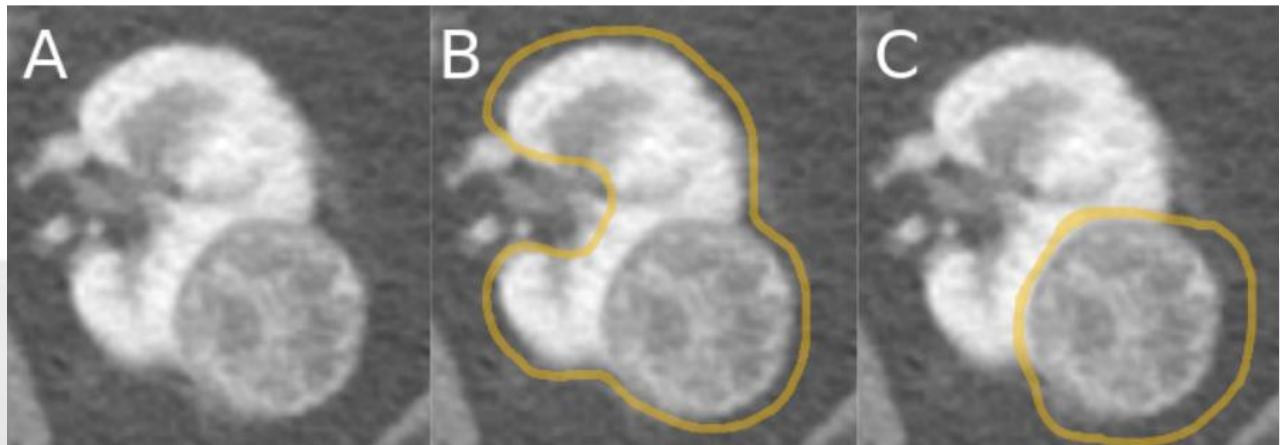
HVSMR 2016

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



3 Datasets

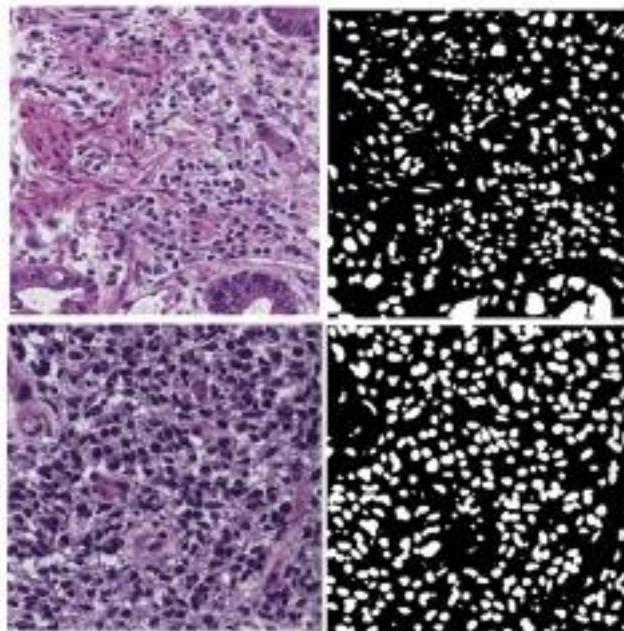
KiTS19/21



KiTS19 Challenge

3 Datasets

- Nucleus Segmentation Challenge
Nuclear segmentation in digital microscopic tissue images

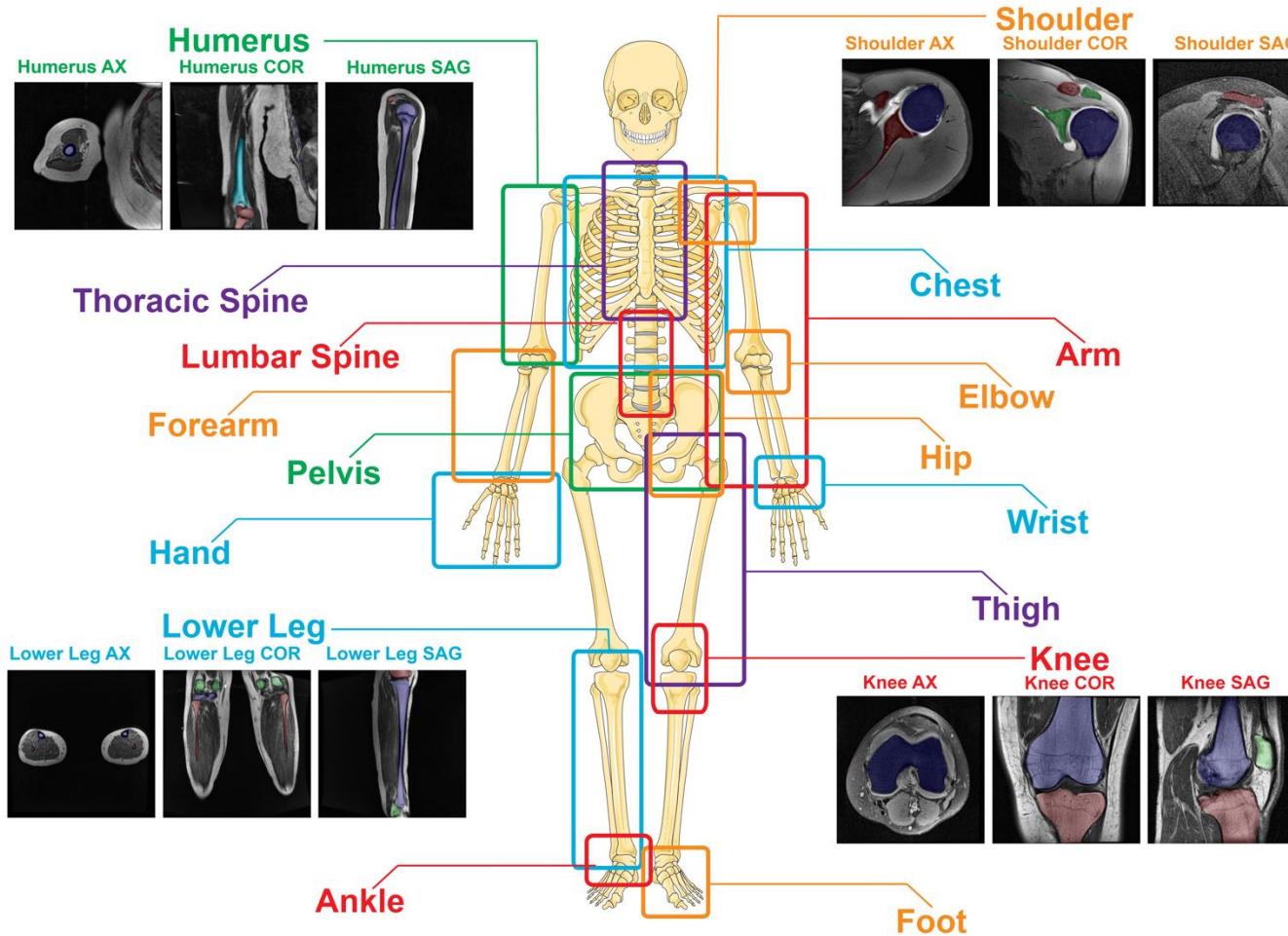


N. Kumar et al., "[A Multi-organ Nucleus Segmentation Challenge](#)," in *IEEE Transactions on Medical Imaging*

<https://monuseg.grand-challenge.org/>

3 Datasets

- SegmentAnyBone: A Universal Model that Segments Any Bone at Any Location on MRI



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.

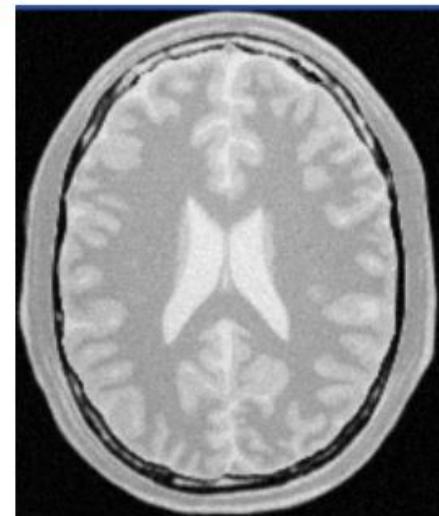
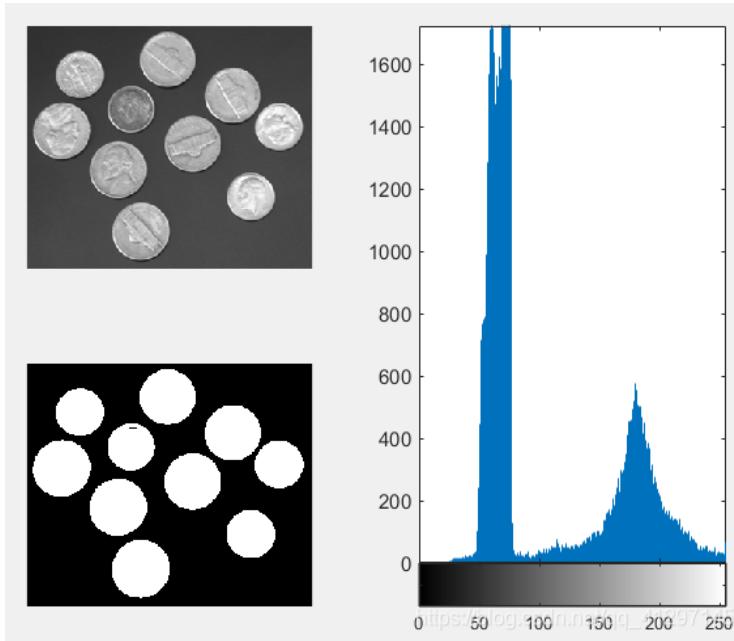


Figure 3. Original image.



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

4 Traditional Methods

Edge Detection-based Method

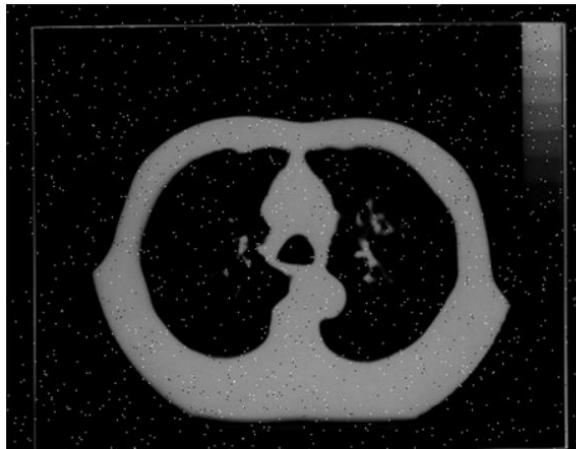


Fig.1. Original lungs CT image with salt-and-pepper noise.

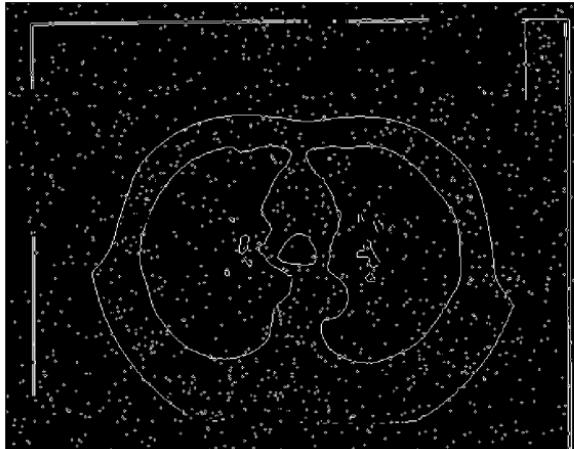


Fig.3. Lungs CT image processed by Sobel detector.



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

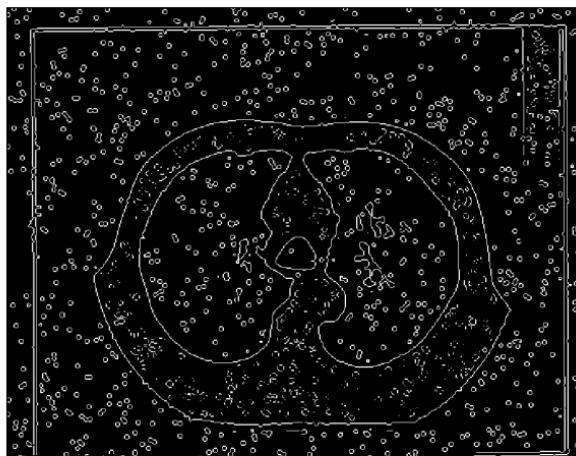


Fig.2. Lungs CT image processed by Laplacian of Gaussian operator.



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



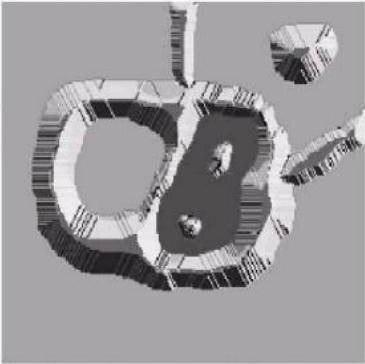
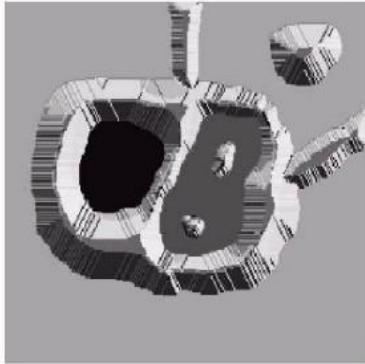
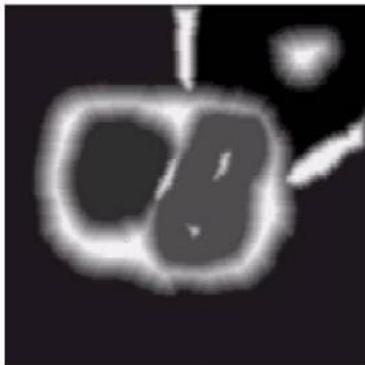
Fig.6. Lungs CT image processed by the novel morphological edge detector.

4 Traditional Methods

Region-based Method

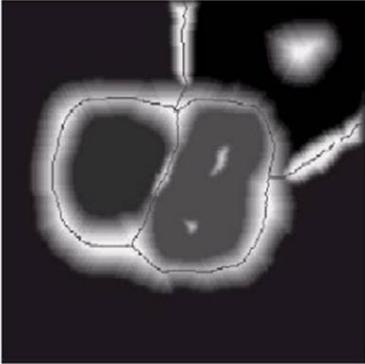
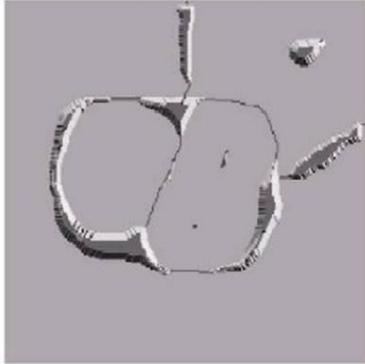
a
b
c
d

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



e
f
g
h

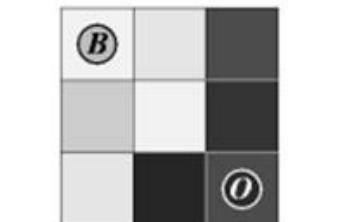
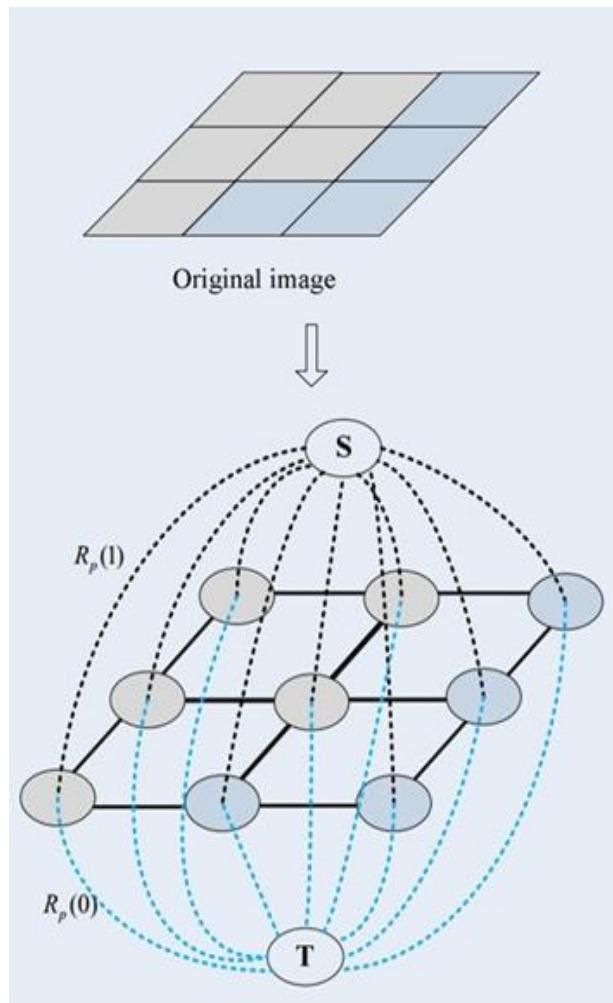
FIGURE 10.44
(Continued)
(e) Result of further flooding.
(f) Beginning of merging of water from two catchment basins (a short dam was built between them). (g) Longer dams. (h) Final watershed (segmentation) lines. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)



Watershed Method

4 Traditional Methods

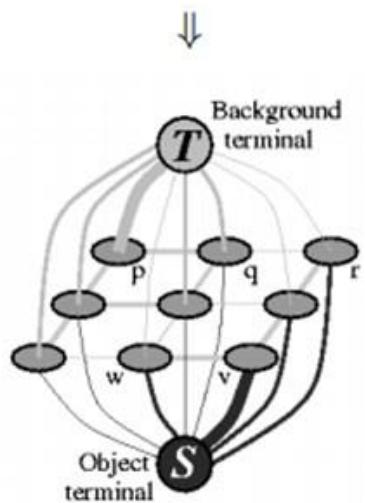
Segmentation based on Graph Theory



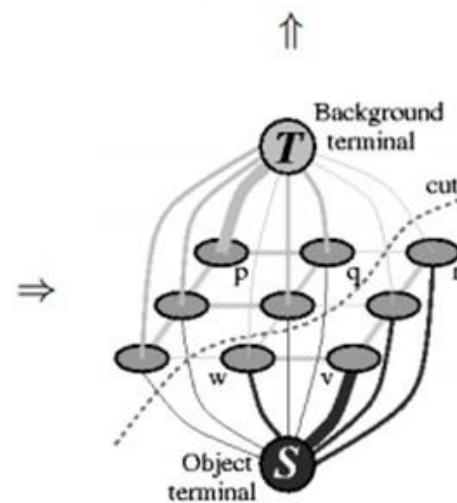
(a) Image with seeds.



(d) Segmentation results.



(b) Graph.



(c) Cut.

3x3 image segmentation of Graph Cut

5 Deep Learning based Methods

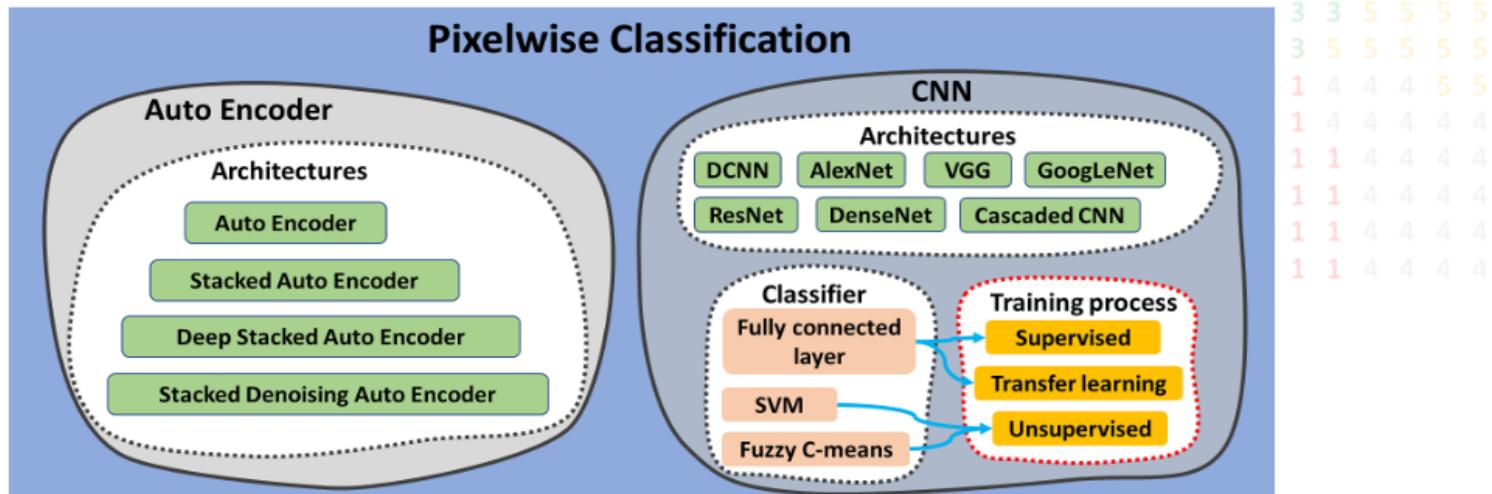


Fig. 1 The network components of the pixelwise classification methods.

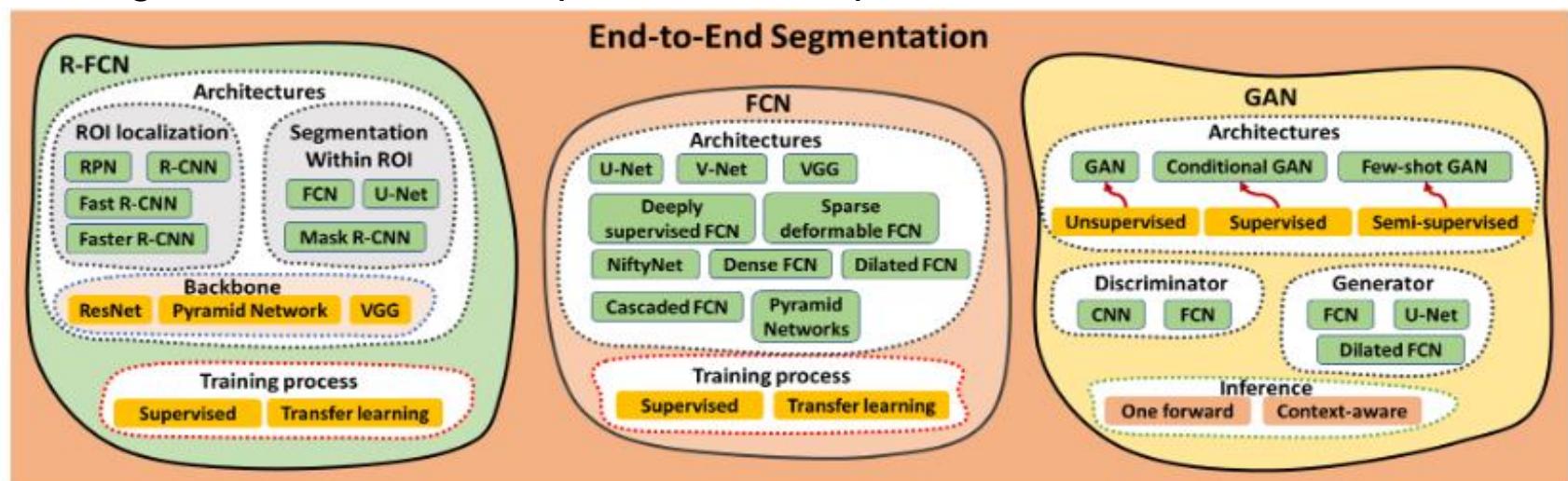


Fig. 2 The network components of the end-to-end segmentation methods.

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					

height



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

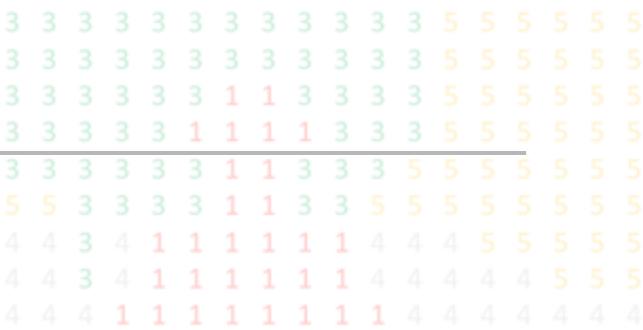
Overview (continued)

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



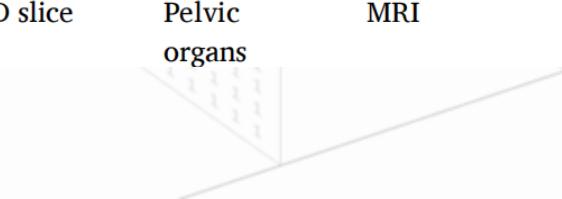
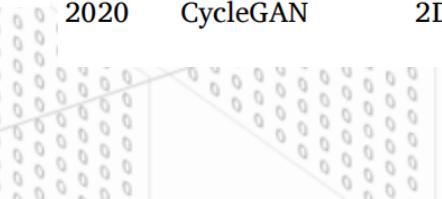
5 Deep Learning based Methods

Overview of GAN Method



Year	Network	Dimension	Site	Modality	Year	Network	Dimension	Site	Modality
2015	SCAN	2D slice	Chest	X-rays	2018	Conditional Generative Refinement Adversarial Networks	2D slice	Brain	MRI
2017	Multi-connected adversarial networks	2D slice	Brain	Multi-modality MRI					
2017	Dilated GAN	2D slice	Brain	MRI	2018	SegAN	2D slice	Brain	MRI
2017	Conditional GAN	2D slice	Brain tumor	MRI	2018	MDAL	2D slice	Left and Right-Ventricular	Cardiac MRI
2017	GAN	2D patch	Retinal Vessel	Fundoscopic				Whole body	X-ray
2017	Adversarial Image-to-Image Network	3D volume	Liver	CT	2018	TD-GAN	2D slice	Thorax	CT
					2019	U-Net-GAN	3D volume	Nuclei	Histopathology Images
					2019	Conditional GAN	2D slice		CT
2017	Adversarial FCN-CRF Nets	2D slice	Mass	Mammograms	2019	Distance-aware GAN	2D slice	Chest	CT
2018	GAN	Not specified	Brain tumor	MRI	2019	Shape Constraint GAN	3D volume	Head & Neck	CT/MRI
2018	Few-shot GAN	3D patch	Brain	MRI	2019	Shape Constraint GAN	3D volume	Abdomen	CT
2018	Context-aware GAN	2D cropped slices	Cardiac	MRI	2020	CycleGAN	2D slice	Pelvic organs	MRI

height



5 Deep Learning based Methods

FCN adapting classifiers for dense prediction

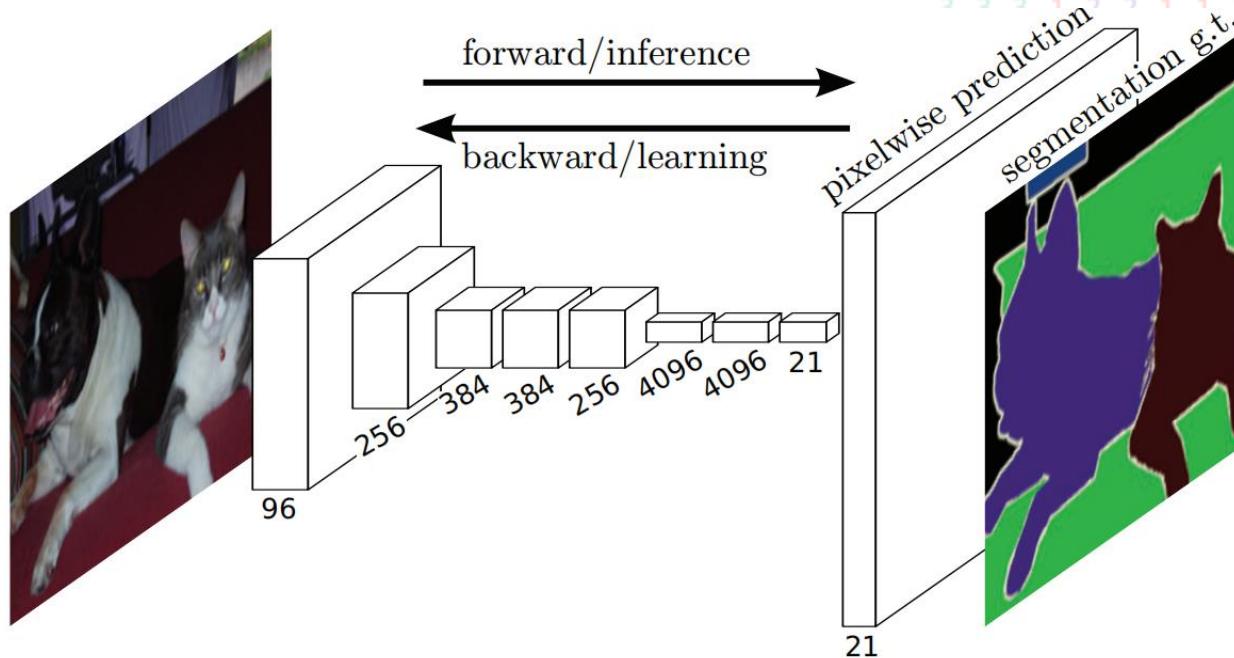


Figure 1. Fully convolutional networks can efficiently learn to make dense predictions for per-pixel tasks like semantic segmentation.

5 Deep Learning based Methods

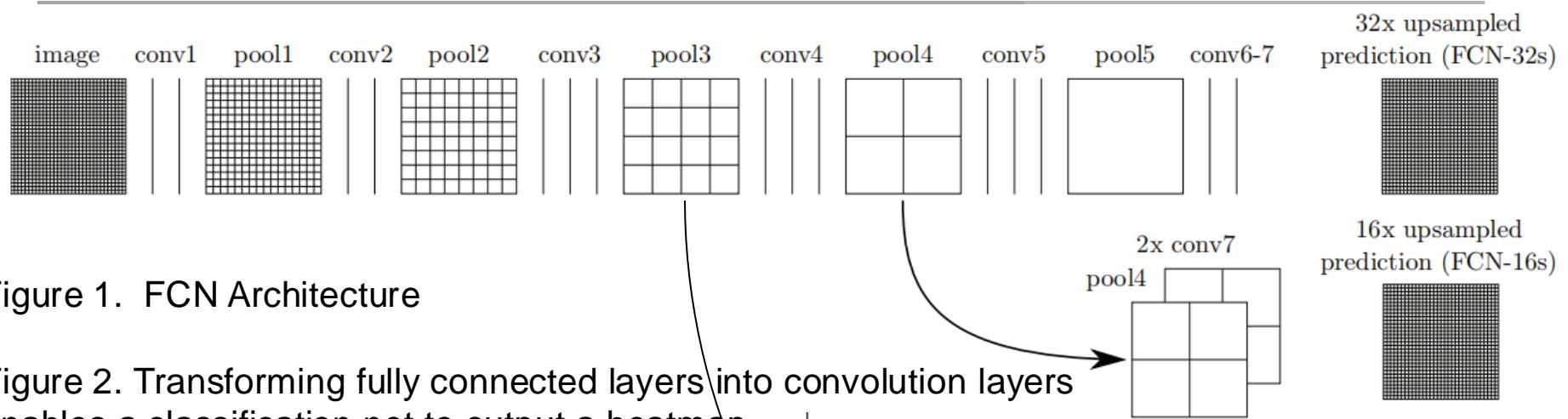
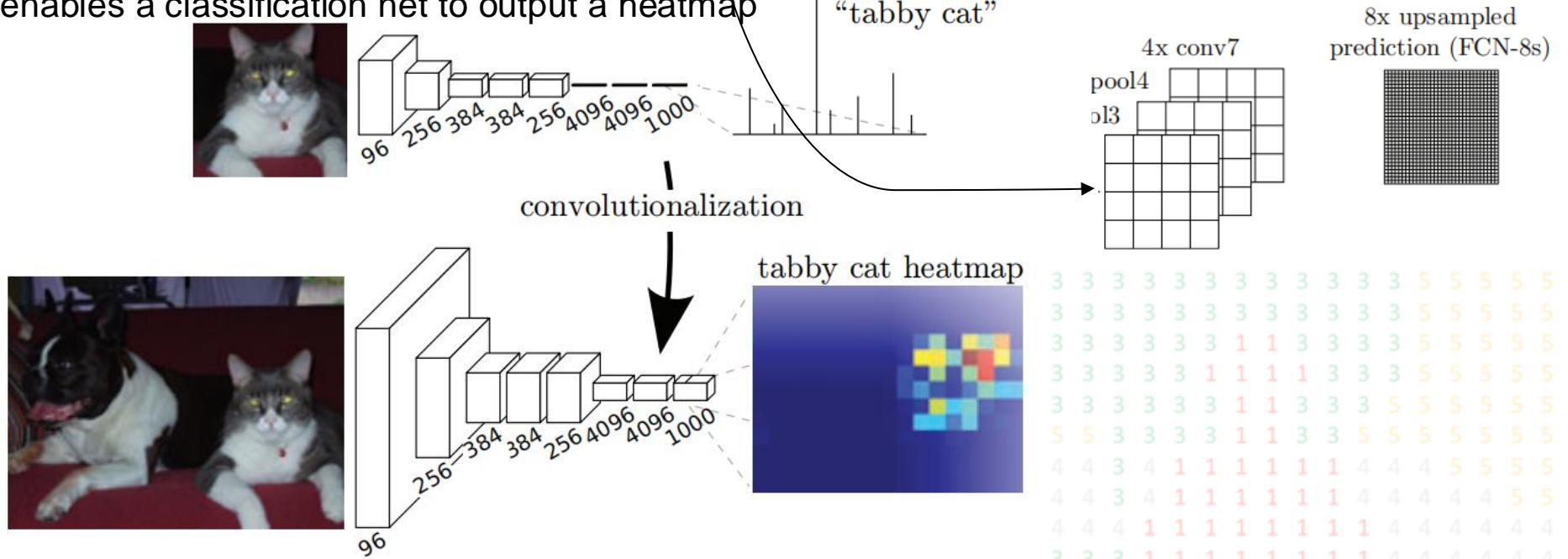


Figure 1. FCN Architecture

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



5 Deep Learning based Methods

U-Net

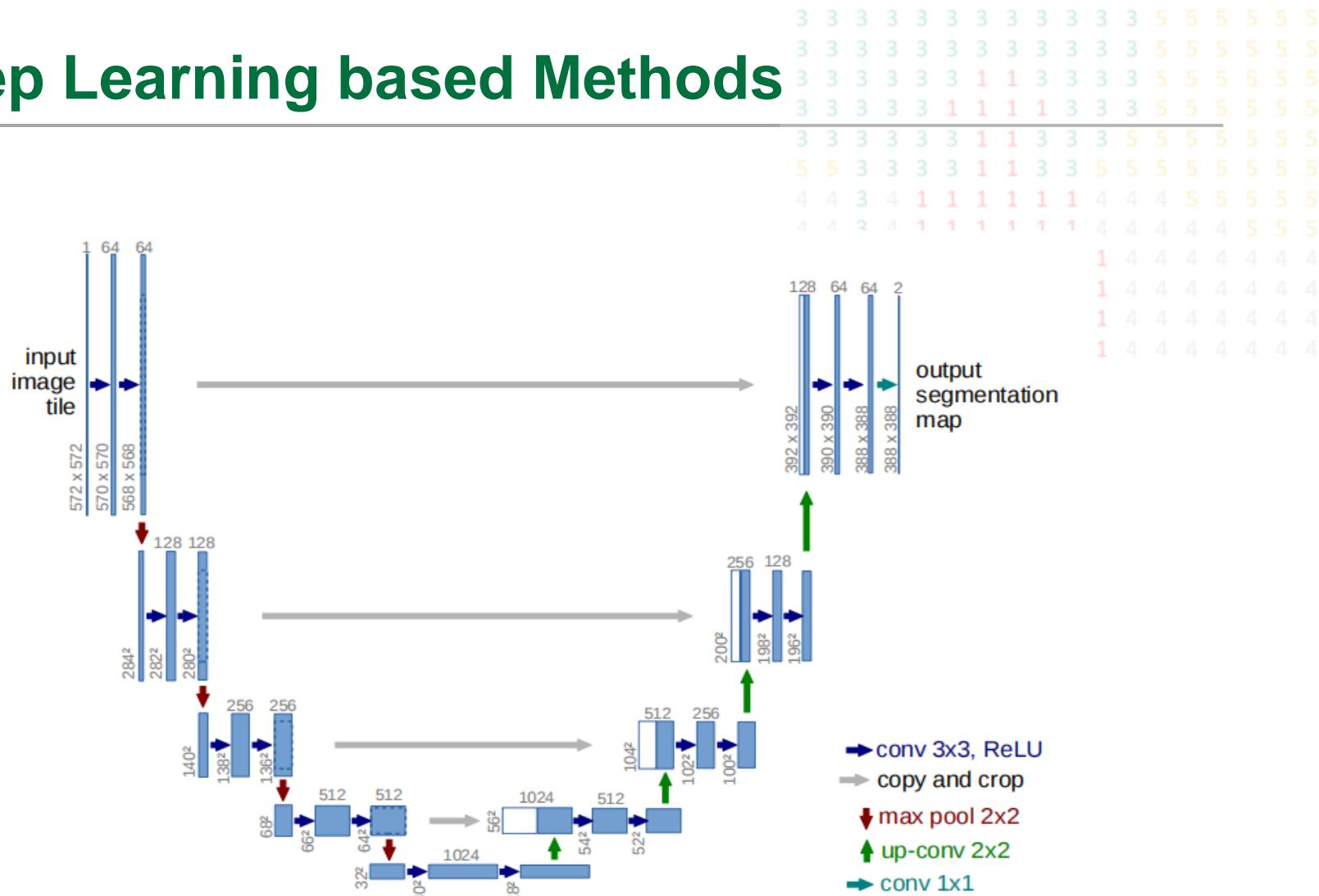


Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows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

5 Deep Learning based Methods

- Upsampling

1. Interpolation-Based Upsampling:

- **Nearest Neighbor:** Replicates the nearest pixel value, resulting in blocky artifacts but is very fast.
- **Bilinear (or Bicubic):** Uses linear (or cubic) interpolation between neighboring pixels to smooth the transition. These methods are non-learnable but computationally efficient.

5 Deep Learning based Methods

- Upsampling

2. Transposed Convolution (Deconvolution):

- This is a learnable upsampling technique where the network learns the appropriate kernels to “reverse” the downsampling process.
- It often involves inserting zeros between pixels (sometimes called “fractionally strided convolution”) followed by a standard convolution to produce a higher resolution feature map.

5 Deep Learning based Methods

- Upsampling

3. Unpooling:

- Used in conjunction with max-pooling layers. During pooling, indices of the maximal values are stored. In unpooling, these indices are used to place the values back in their original positions, often followed by a convolution to fill in the gaps.

5 Deep Learning based Methods

U-Net

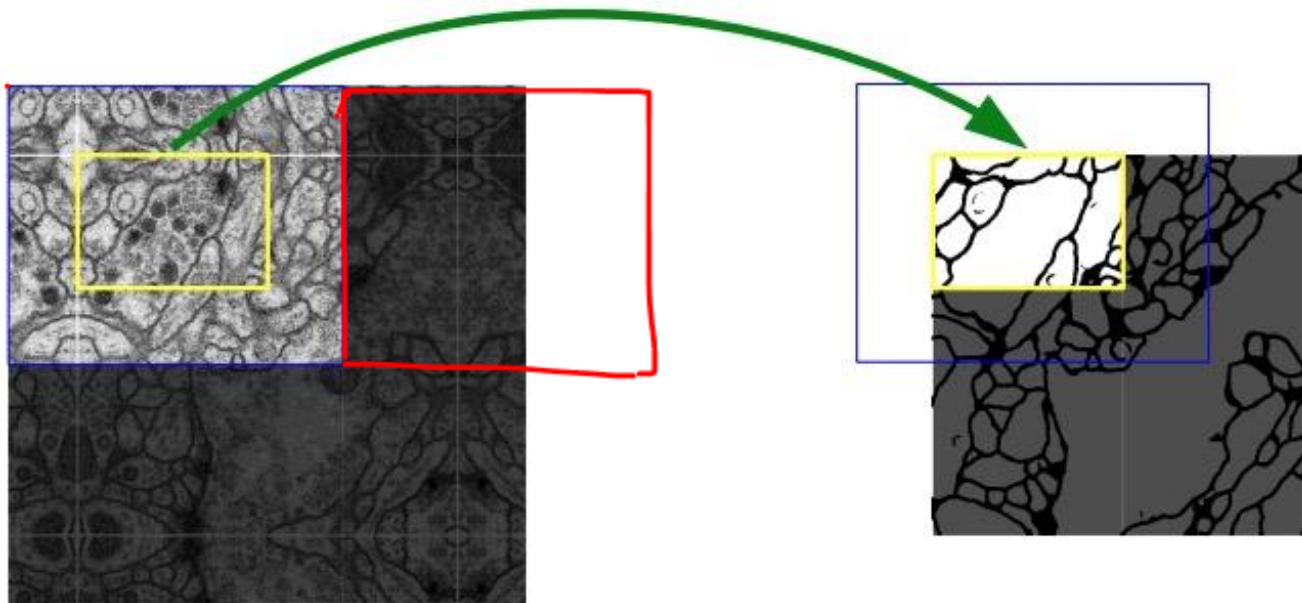


Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring

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

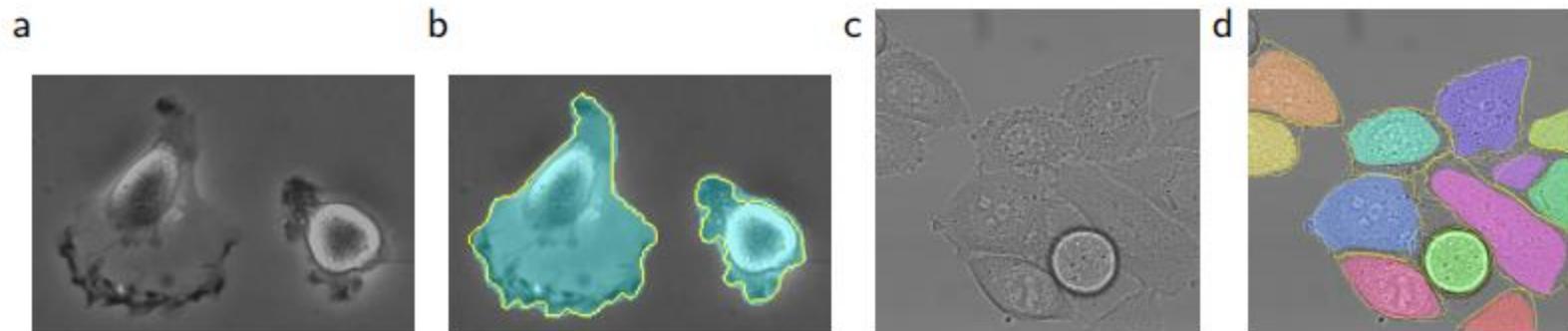


Fig. 4. Result on the ISBI cell tracking challenge.

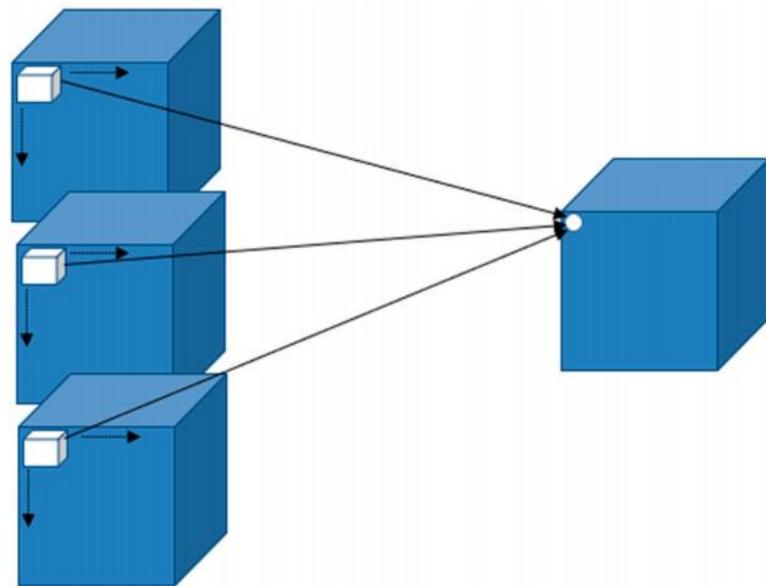
Table 2. Segmentation results (IOU) on the ISBI cell tracking challenge 2015.

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

Why U-net performs well in medical image segmentation?

5 Deep Learning based Methods

3D U-Net



3D CNN

2D CNN

$$\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$$

Cov Kernel: 3×3



3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

Cov layer

5 Deep Learning based Methods

3D U-Net

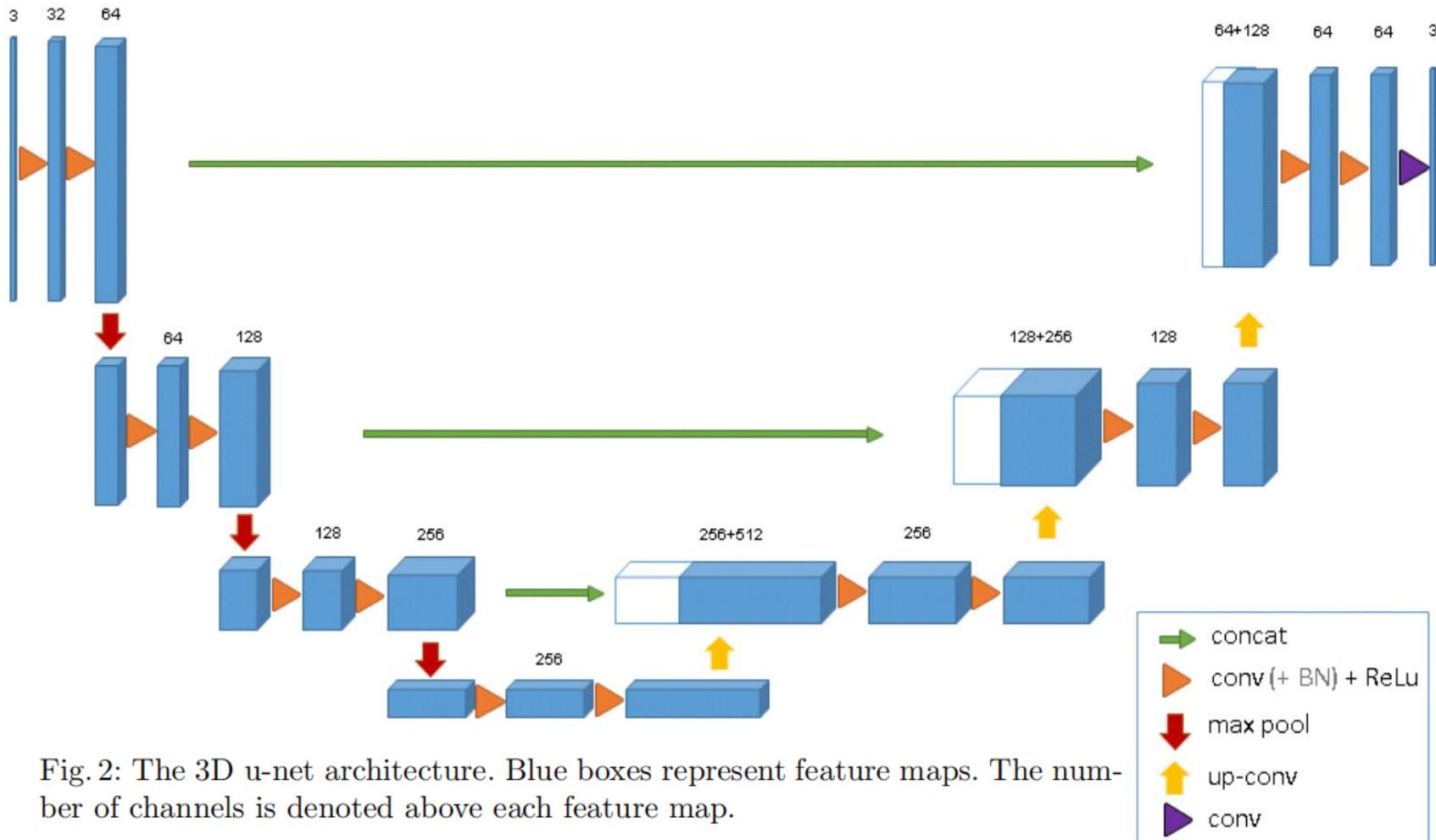
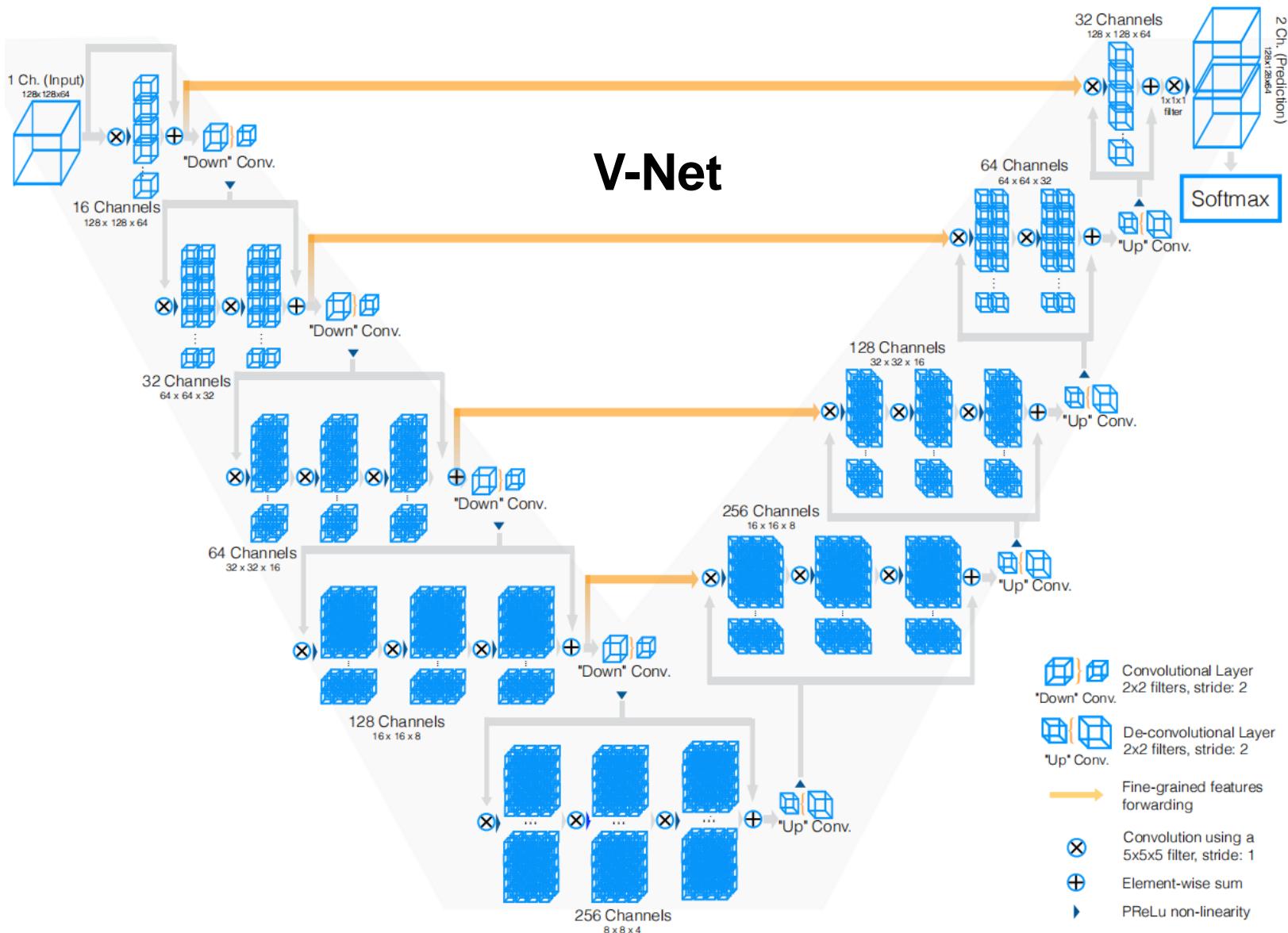


Fig. 2: The 3D u-net architecture. Blue boxes represent feature maps. The number of channels is denoted above each feature map.

5 Deep Learning based Methods



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

Results of V-Net

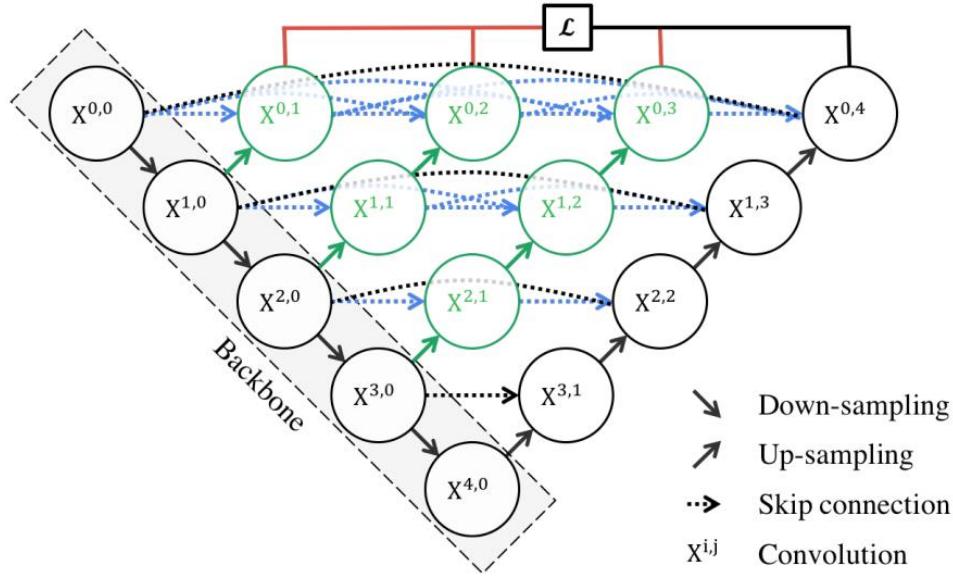
Table 2. Quantitative comparison between the proposed approach and the current best results on the PROMISE 2012 challenge dataset.

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

5 Deep Learning based Methods

Other U-Net Variants: Unet++

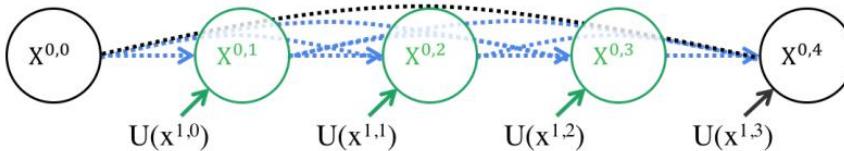
$$x^{i,j} = \begin{cases} \mathcal{H}(x^{i-1,j}), & j = 0 \\ \mathcal{H}\left(\left[x^{i,k}\right]_{k=0}^{j-1}, \mathcal{U}(x^{i+1,j-1})\right), & j > 0 \end{cases}$$



Feature Fusion

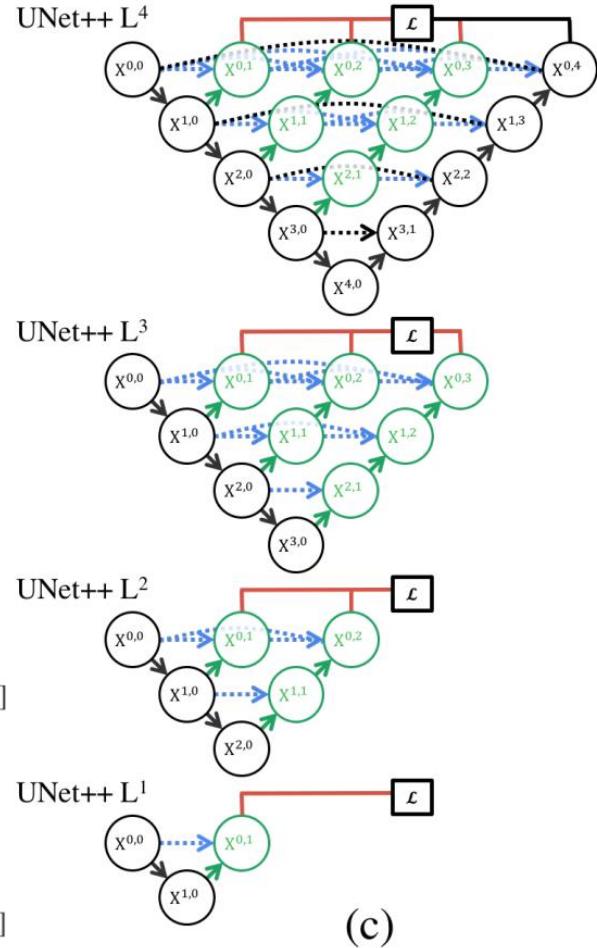
(a)

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



(b)

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



5 Deep Learning based Methods

Results of Unet++

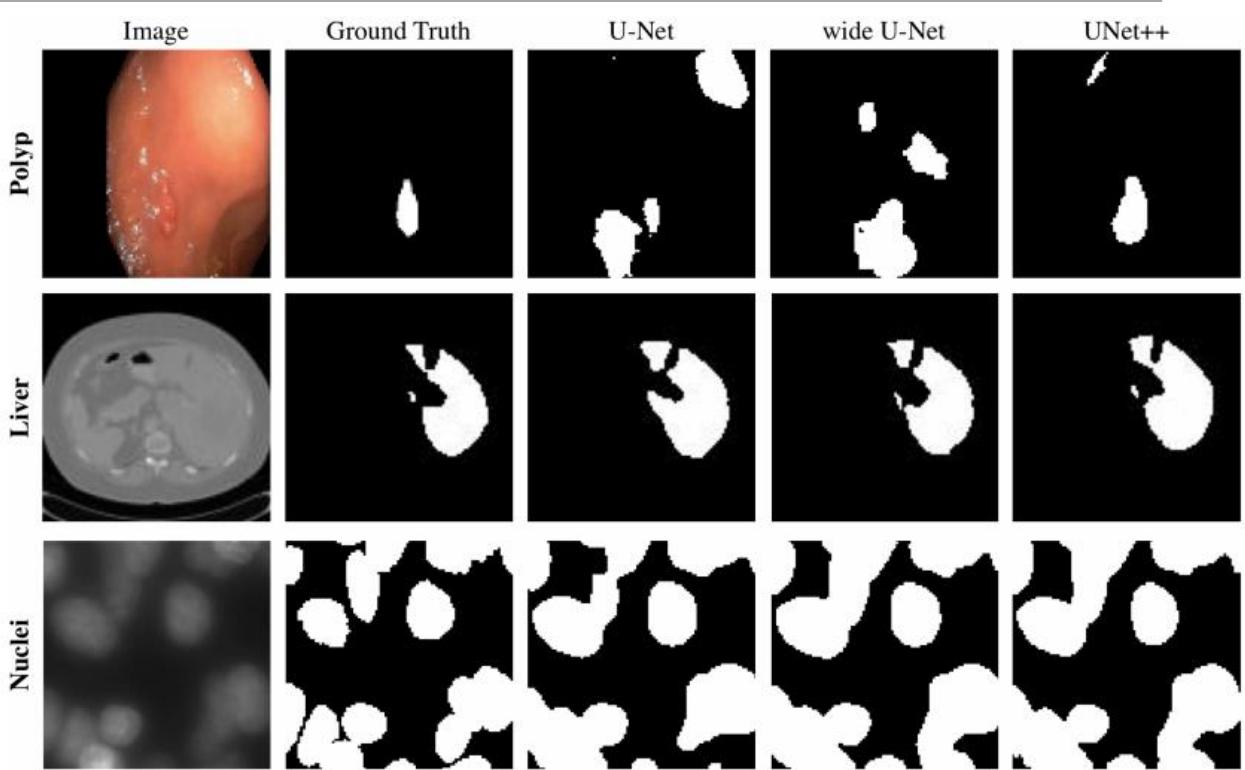


Fig. 2: Qualitative comparison between U-Net, wide U-Net, and UNet++, showing segmentation results for polyp, liver, and cell nuclei datasets (2D-only for a distinct visualization).

$$\mathcal{L}(Y, \hat{Y}) = -\frac{1}{N} \sum_{b=1}^N \left(\frac{1}{2} \cdot Y_b \cdot \log \hat{Y}_b + \frac{2 \cdot Y_b \cdot \hat{Y}_b}{Y_b + \hat{Y}_b} \right)$$

Loss Function

5 Deep Learning based Methods

Results of Unet++

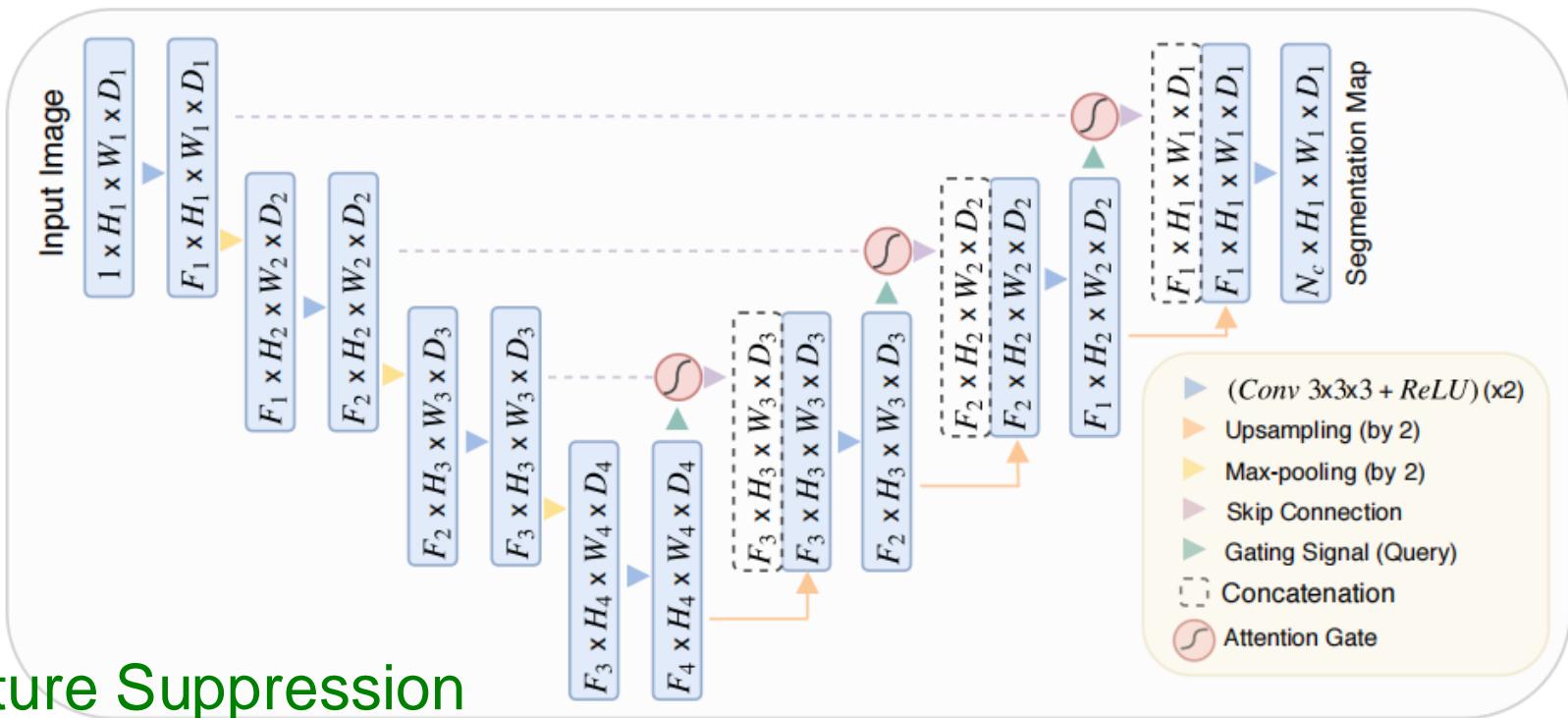
Table 3: Segmentation results (IoU: %) for U-Net, wide U-Net and our suggested architecture UNet++ with and without deep supervision (DS).

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.

5 Deep Learning based Methods

Other U-Net Variants: Attention U-Net



Feature Suppression

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

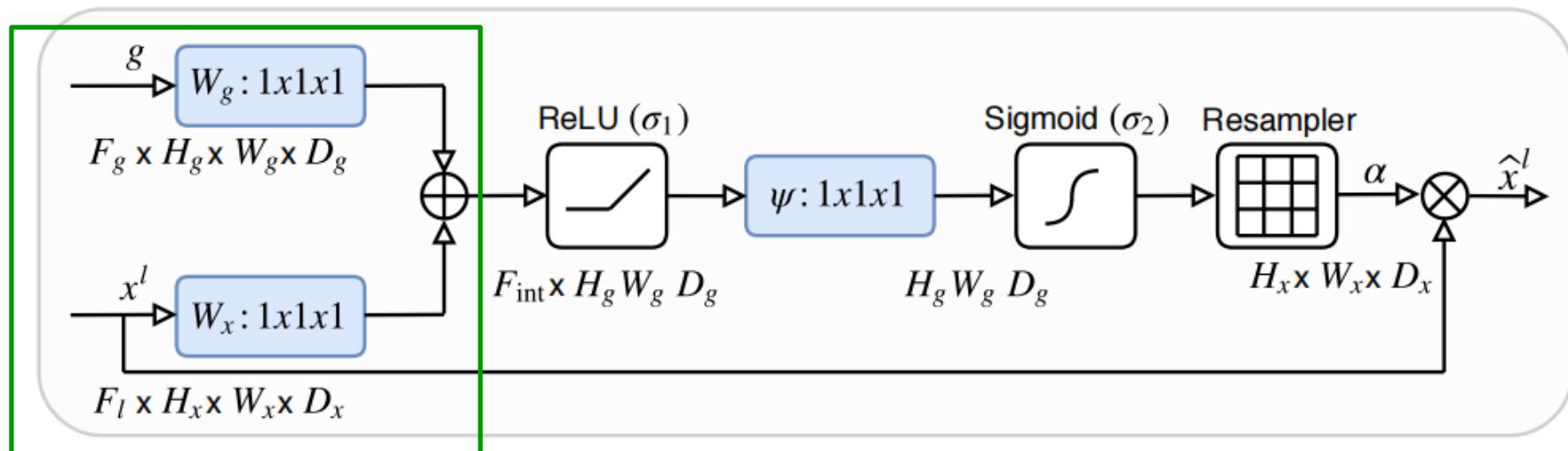
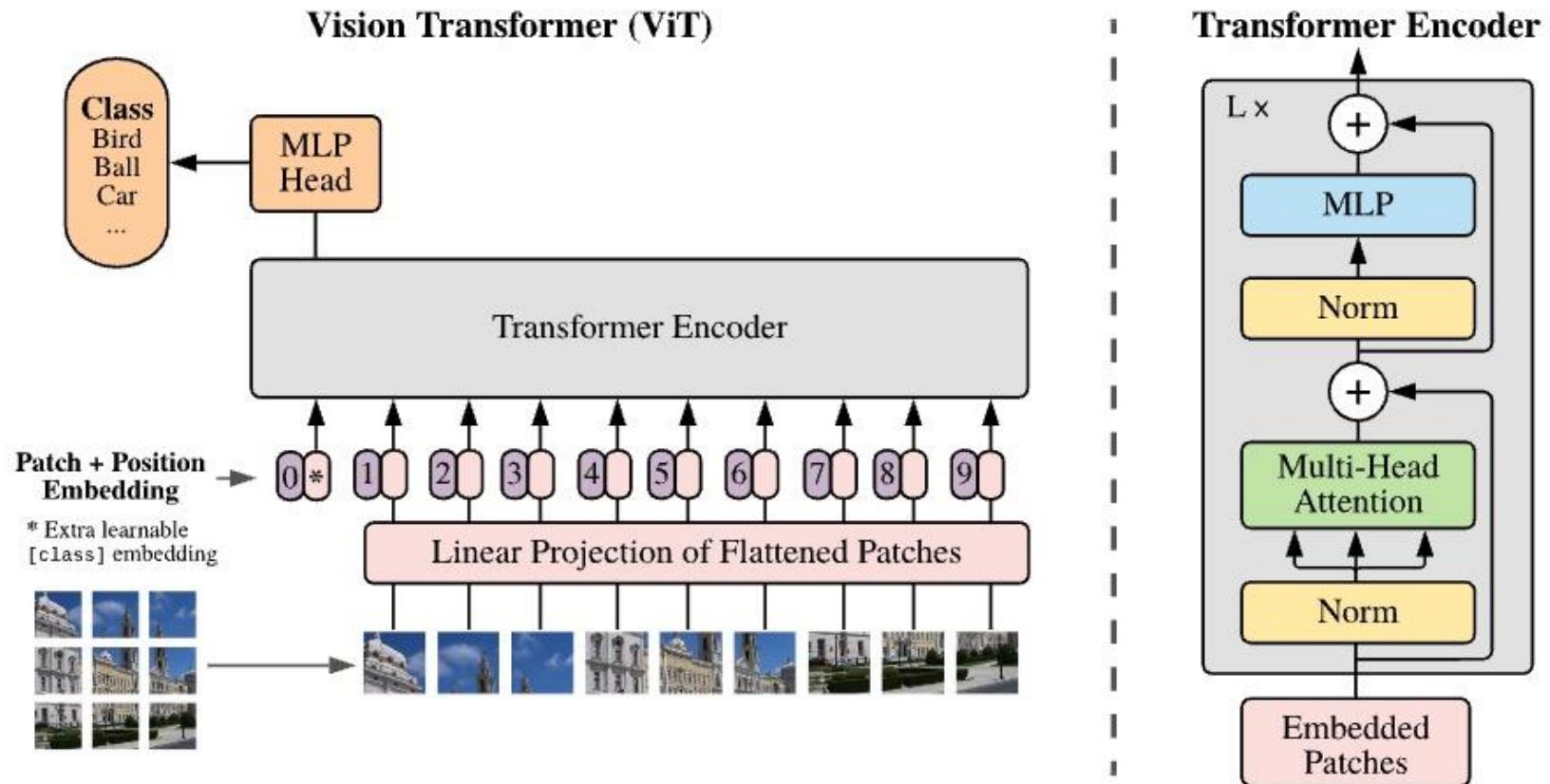


Figure 2: Schematic of the proposed additive attention gate (AG). Input features (x^l) are scaled with attention coefficients (α) computed in AG. Spatial regions are selected by analysing both the activations and contextual information provided by the gating signal (g) which is collected from a coarser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

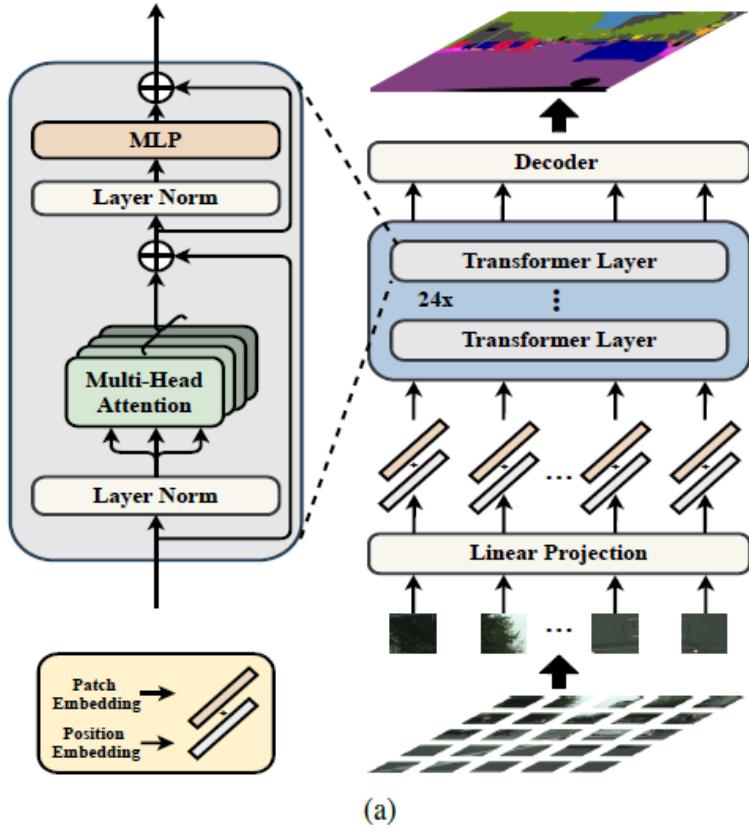
Vision Transformer



Alexey, D. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv: 2010.11929*.

5 Deep Learning based Methods

Vision Transformer



Adjust image data: patches
>> sequence (convolution)

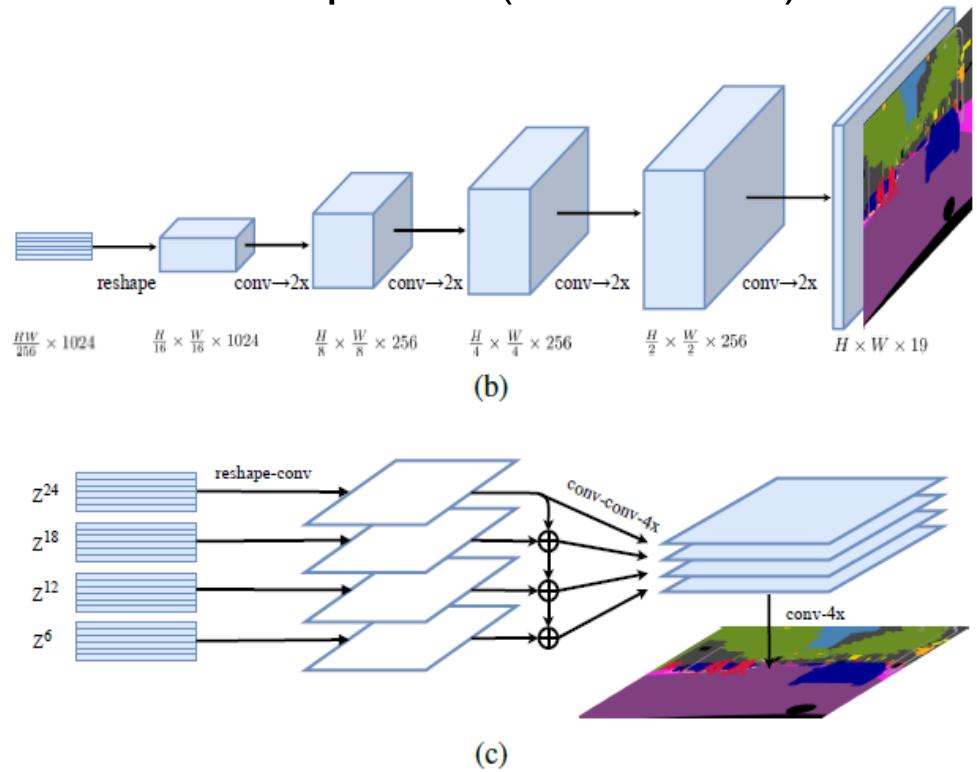


Figure 1. Schematic illustration of the proposed **SEgmentation TTransformer** (SETR) (a). We first split an image into fixed-size patches, linearly embed each of them, add position embeddings, and feed the resulting sequence of vectors to a standard Transformer encoder. To perform pixel-wise segmentation, we introduce different decoder designs: (b) progressive upsampling (resulting in a variant called SETR-PUP); and (c) multi-level feature aggregation (a variant called SETR-MLA).

5 Deep Learning based Methods

Medical image segmentation + Transformer

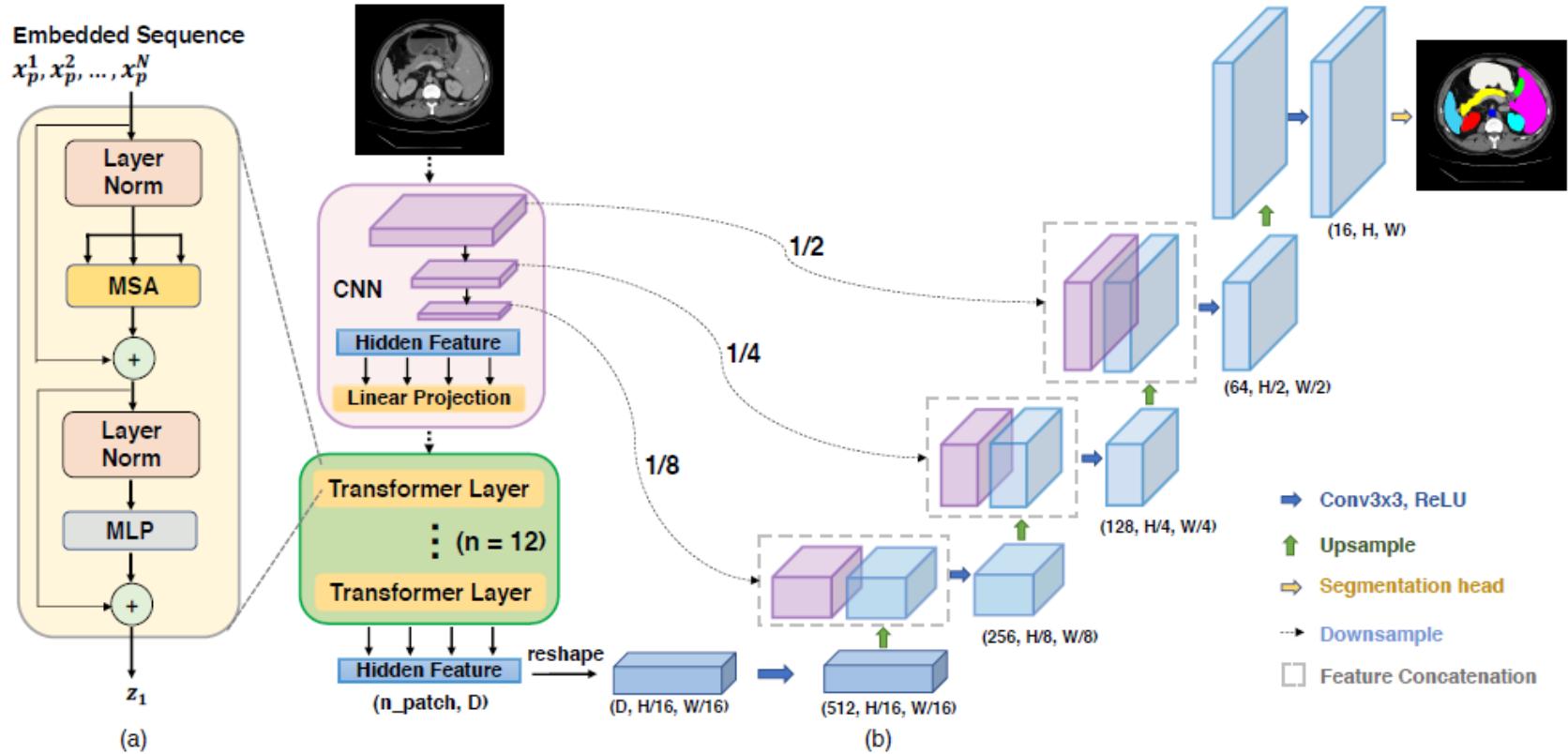


Fig. 1: Overview of the framework. (a) schematic of the Transformer layer; (b) architecture of the proposed TransUNet.

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

Medical image segmentation + Transformer

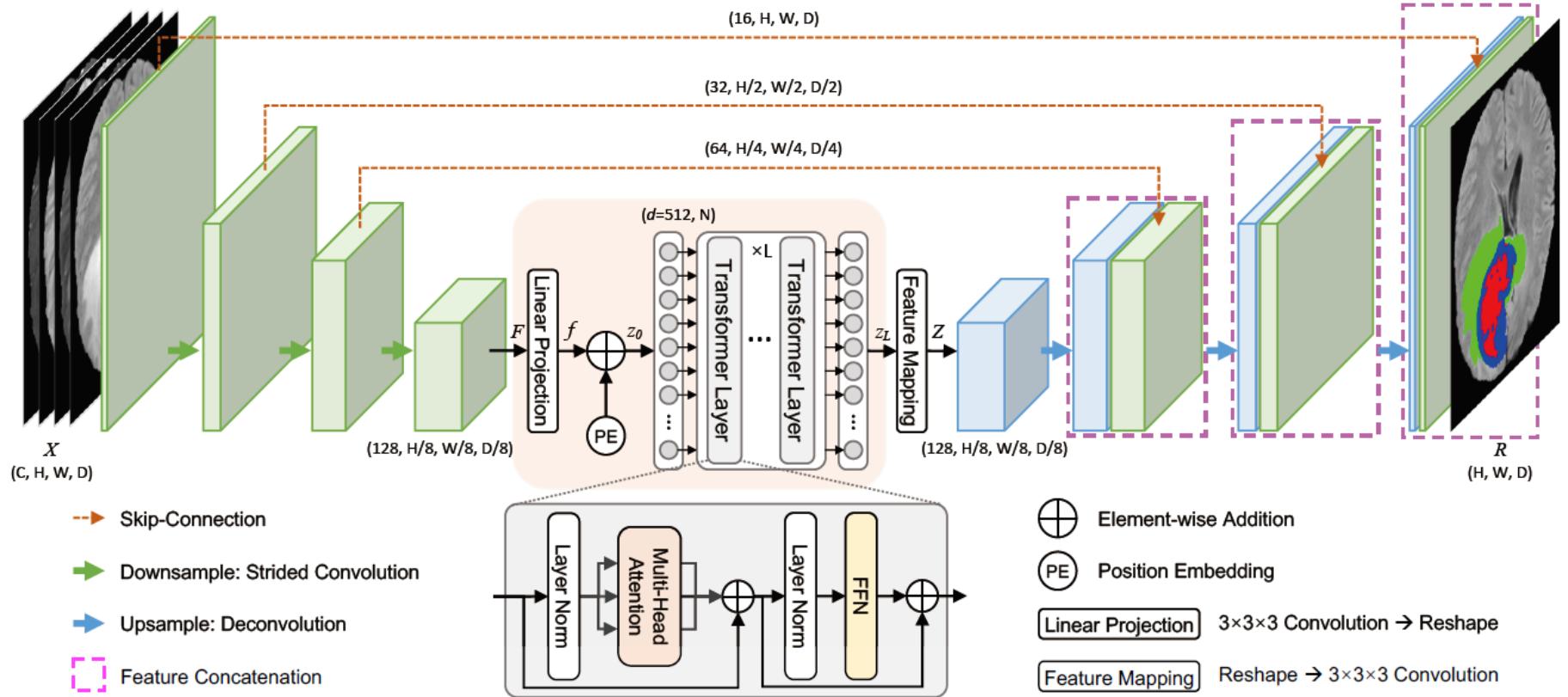


Fig. 1. Overall architecture of the proposed TransBTS.

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

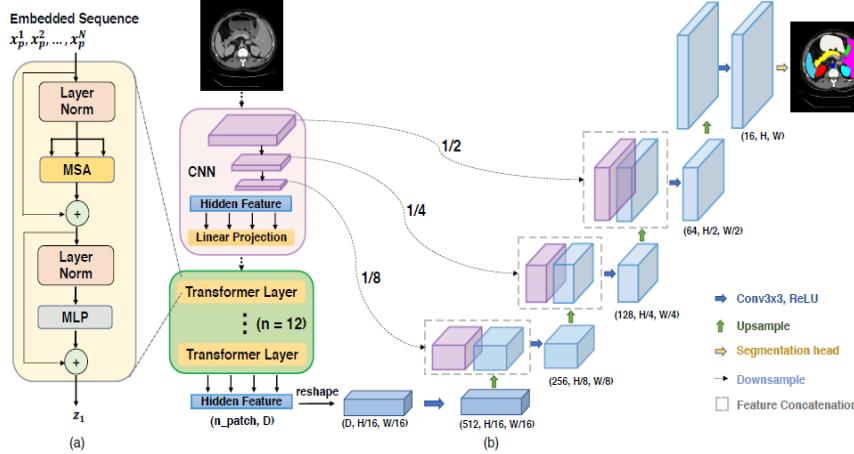


Fig. 1: Overview of the framework. (a) schematic of the Transformer layer; (b) architecture of the proposed TransUNet.

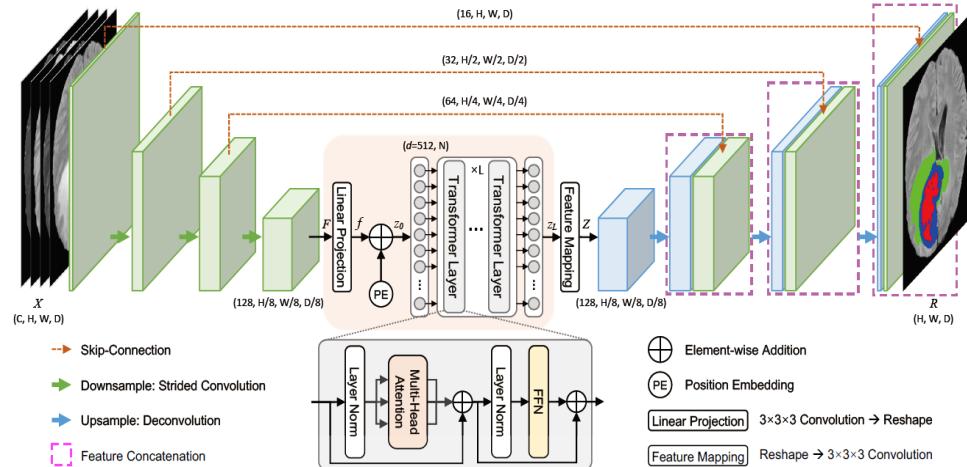


Fig. 1. Overall architecture of the proposed TransBTS.

- 2D slice-by-slice
- Load Pretrain model

VS
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 arXiv:2103.04430, 2021.

5 Deep Learning based Methods

Medical image segmentation + Transformer

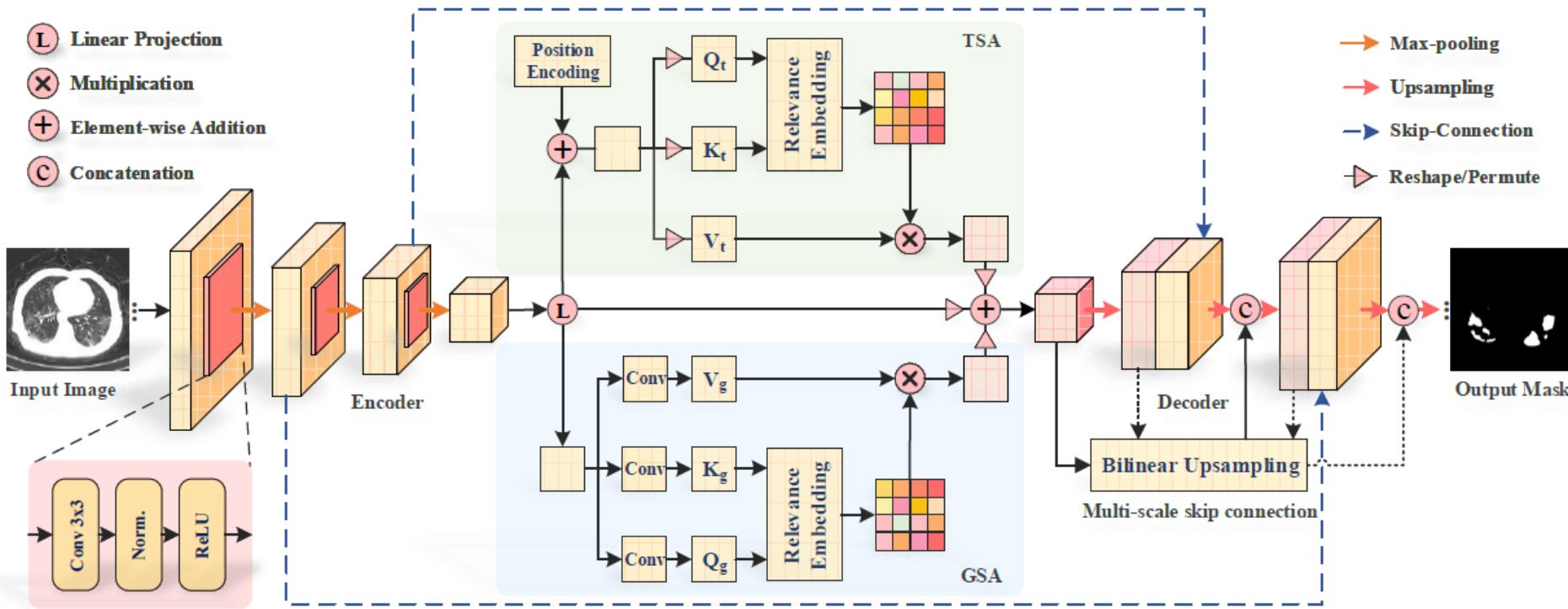
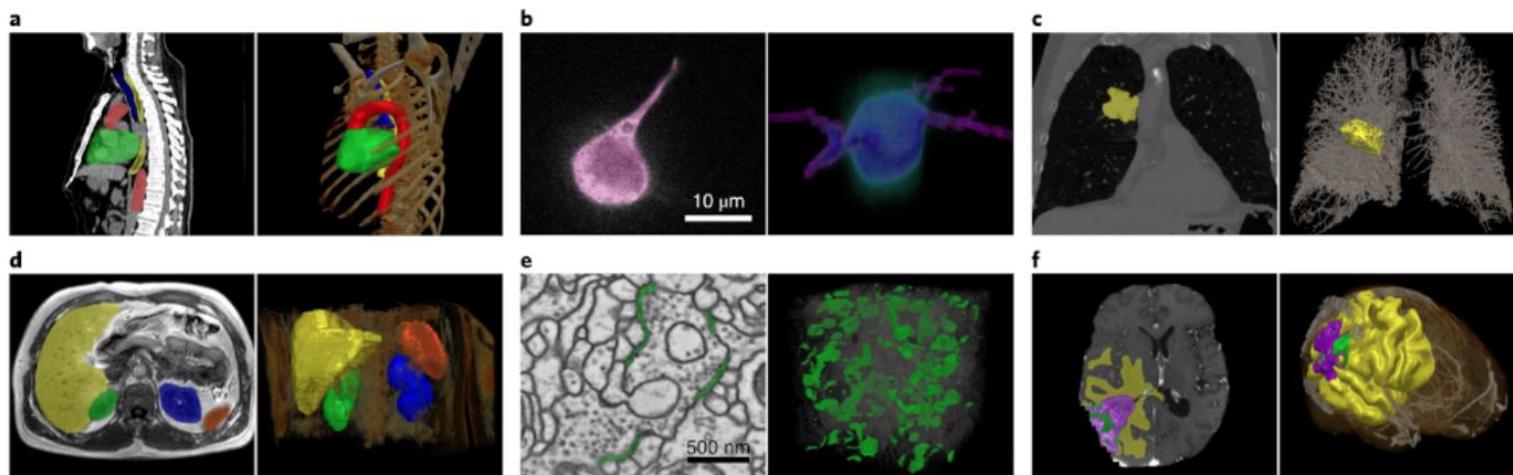
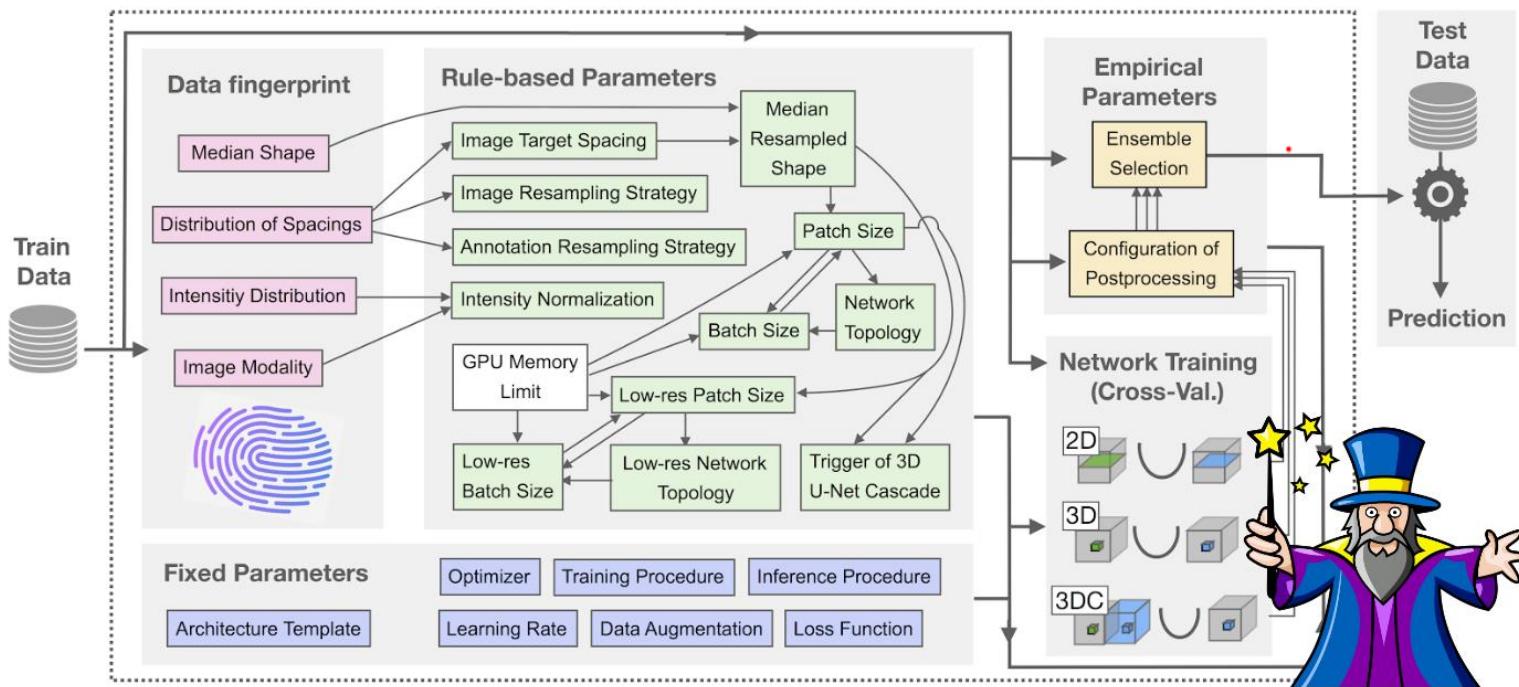


Fig. 2. Illustration of the proposed TransAttUnet for automatic medical image segmentation. (a) Both TSA and GSA mechanisms are embedded into the SAA module to model the long-range interactions and global spatial relationships. (b) The multi-scale skip connections between decoder blocks are designed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

-
- nnUNet
 - nnU-Net is a semantic segmentation method that automatically adapts to a given dataset. It will analyze the provided training cases and automatically configure a matching U-Net-based segmentation pipeline.
 - nnU-Net was evaluated on 23 datasets belonging to competitions from the biomedical domain.

<https://github.com/MIC-DKFZ/nnUNet>



<https://github.com/MIC-DKFZ/nnUNet>

Thank you!

Question?