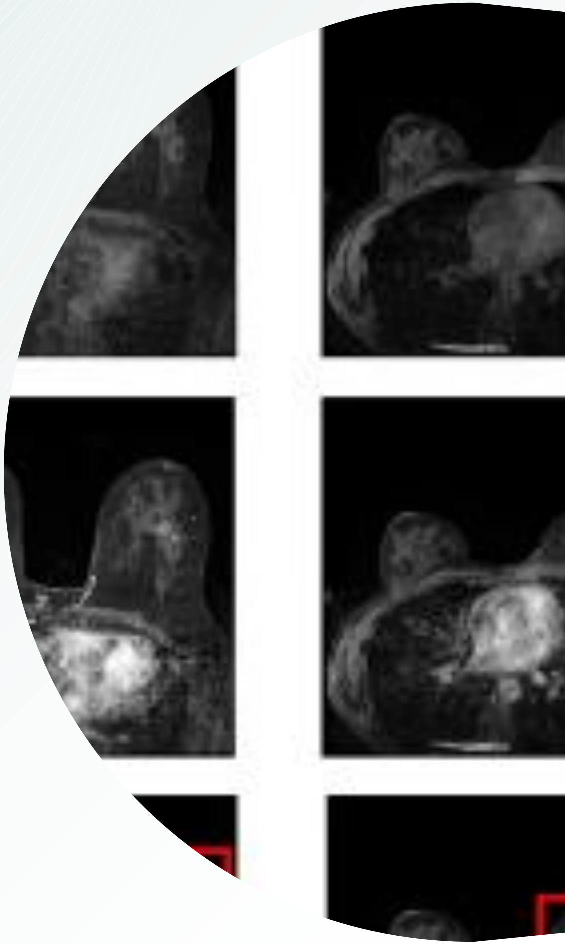
UNIVERSITÄT DES SAARLANDES

A U-Net Ensemble for breast lesion segmentation in DCE MRI

Roa’a Khaled, Joel Vidal, Joan C Vilanova, Robert Martí

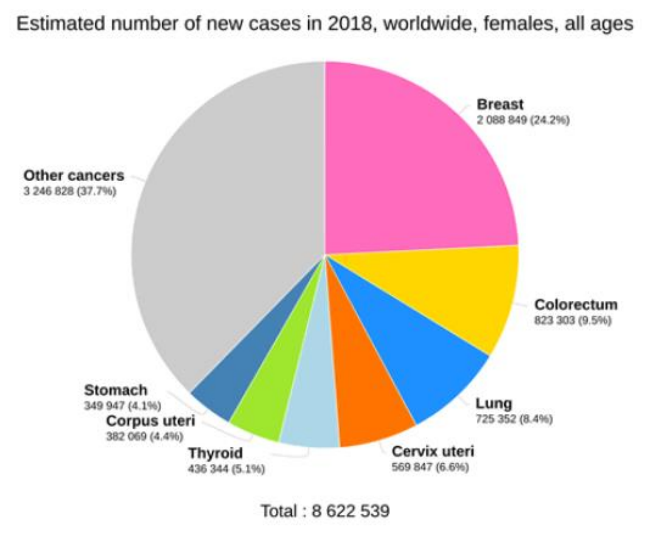
Seminar: Biomedical Image Analysis

Nazlıgül Keske

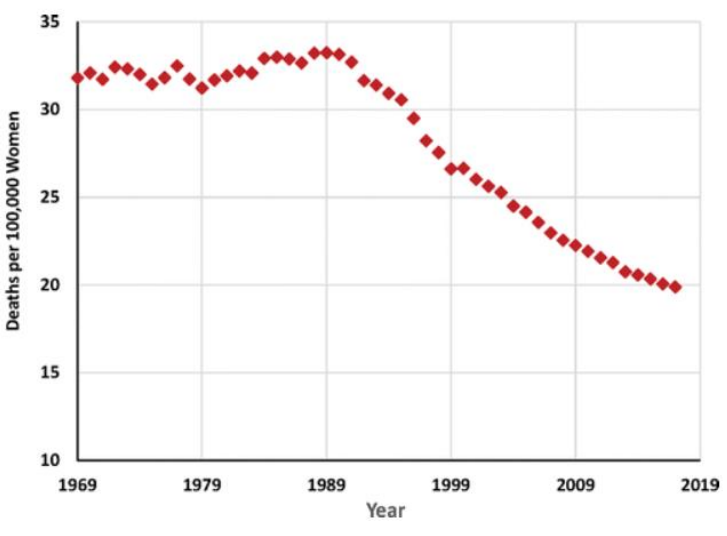


OUTLINE

1. 1. Breast cancer



(Khazaei et al., 2018)



(Hendrick et al., 2021)

- ❑ **Early detection is key**

1. 2. Imaging-based detection methods

Ultrasound

Mammography

Magnetic resonance Imaging (MRI)

Tomosynthesis
Elastphotoacoustics
Optical imaging

- **Mammography:** Sensitivity (true positive rate) decline from 75% to 50% in middle aged patients (Sree, 2011)
- **Ultrasound:** Operator dependent and having an increased false positive rate(Sree, 2011)
- **MRI:** Higher sensitivity than mammography and ultrasound in dense breasts, no radiation

⇩

Dynamic contrast enhanced MRI (DCE-MRI)

1. 2. DCE-MRI

T1 Mapping

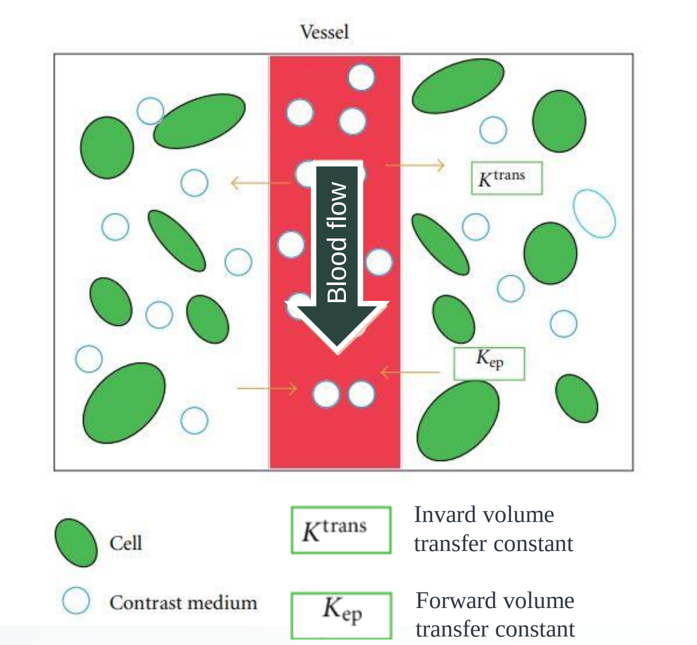
+

Dynamic scan

→ Estimation of T1 → Tissue gadolinium concentration + Arterial Input function

- Regional blood flow
- Size and number of blood vessels
- Permeability

Cancer



(Chikui, T. et al. 2012)

2. Premise of the study

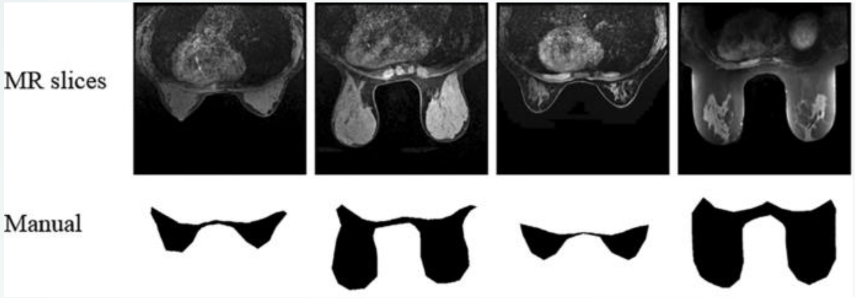
DCE-MRI analysis:

- Time consuming
- Requires experienced radiologists
- Large amount of 4-D information

Automated feature extration is required:

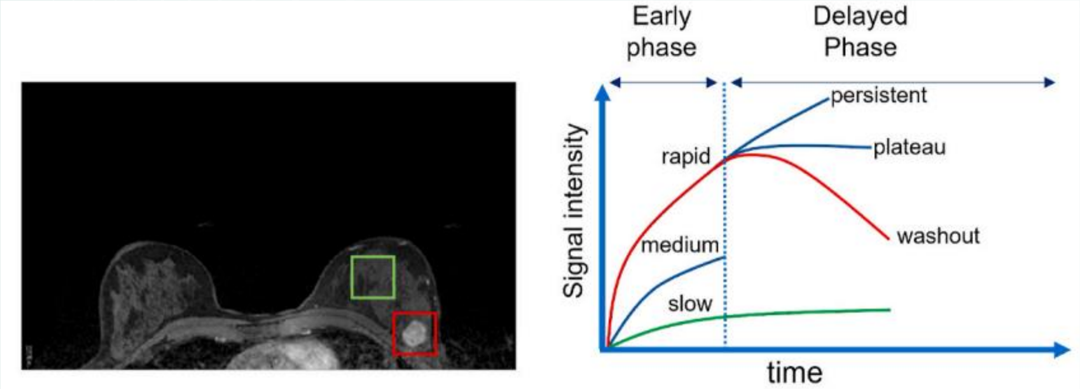
- Lesion morphology
- Texture
- Enhancement kinetics

Accurate segmentation is the first step for feature extraction



2.1. Previsous approaches for breast tissue segmentation

Semi-Automated methods	• Radiologists need to define lesion regions first
Traditional machine-learning methods	• Gubern-Mérida et al., 2015 – Segmentation by kinetic features • Hu et al., 2015 – Kinetic features of each voxel in ROI

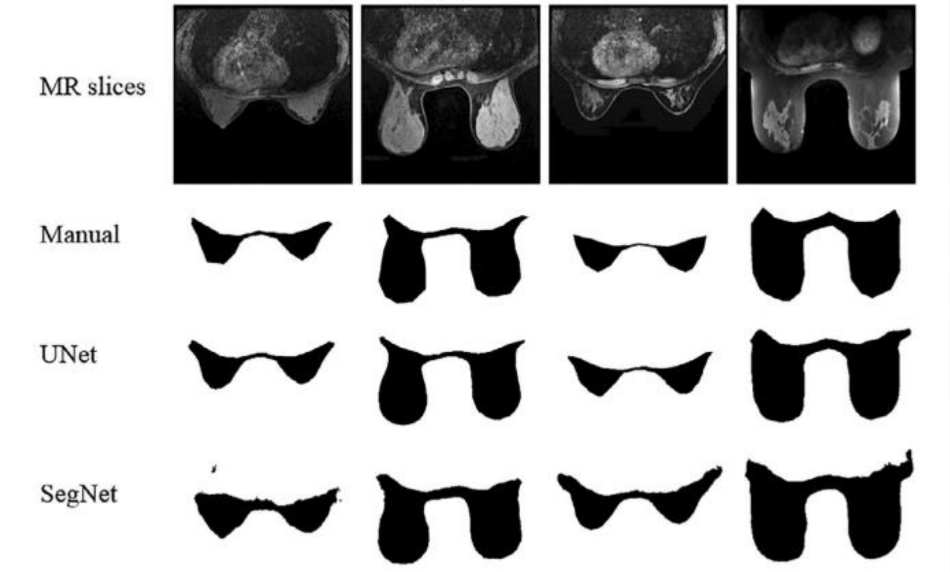


Left figure: *Second post-contrast Volume from DCE-MRI.*
Right figure: *Time Intensity Curves (TIC) of normal tissue (in green) and a malignant lesion (in red)*

2.1. Previsous approaches for breast tissue segmentation

Semi-Automated methods	• Radiologists need to define lesion regions first
Traditional machine-learning methods	• Gubern-Mérida et al., 2015 – Segmentation by kinetic features • Hu et al., 2015 – Kinetic features of each voxel in ROI
Deep-learning methods	• Badrinarayanan et al., 2017 - SegNet • Ronneberger et al., 2015 – U-Net

2.1. Previsous approaches for breast tissue segmentation:



Data	Model	Dice Coefficient
DCE dataset 1		
	UNet	0.92 ± 0.07
	SegNet	0.84 ± 0.11
DCE dataset 2		
	UNet	0.87 ± 0.06
	SegNet	0.80 ± 0.06

(Zhang, L. et al. 2019)

2.1. Previsous approaches for breast tissue segmentation

Semi-Automated methods	• Radiologists need to define lesion regions first
Traditional machine-learning methods	• Gubern-Mérida et al., 2015 – Segmentation by kinetic features • Hu et al., 2015 – Kinetic features of each voxel in ROI
Deep-learning methods	• Badrinarayanan et al., 2017 - SegNet • Ronneberger et al., 2015 – U-Net

- Shi et al., 2015 – ConvLSTM – 4 layer U-Net
- Piantadosi et al., 2019 - Three Time Points approach (3 TP) of U-Net
- Zhang et al., 2019 - mask-guided hierarchical learning (MHL) U-Net framework

2.1. Previsous approaches for breast tissue segmentation

Semi-Automated methods

- Radiologists need to define lesion regions first

Traditional machine-learning methods

- Gubern-Mérida et al., 2015 – Segmentation by kinetic features
- Hu et al., 2015 – Kinetic features of each voxel in ROI

Deep-learning methods

- Badrinarayanan et al., 2017 - SegNet
- Ronneberger et al., 2015 – U-Net

- Shi et al., 2015 – ConvLSTM – 4 layer U-Net
- Piantadosi et al., 2019 - Three Time Points approach (3 TP) of U-Net
- Zhang, J. et al., 2019 - mask-guided hierarchical learning (MHL) U-Net framework

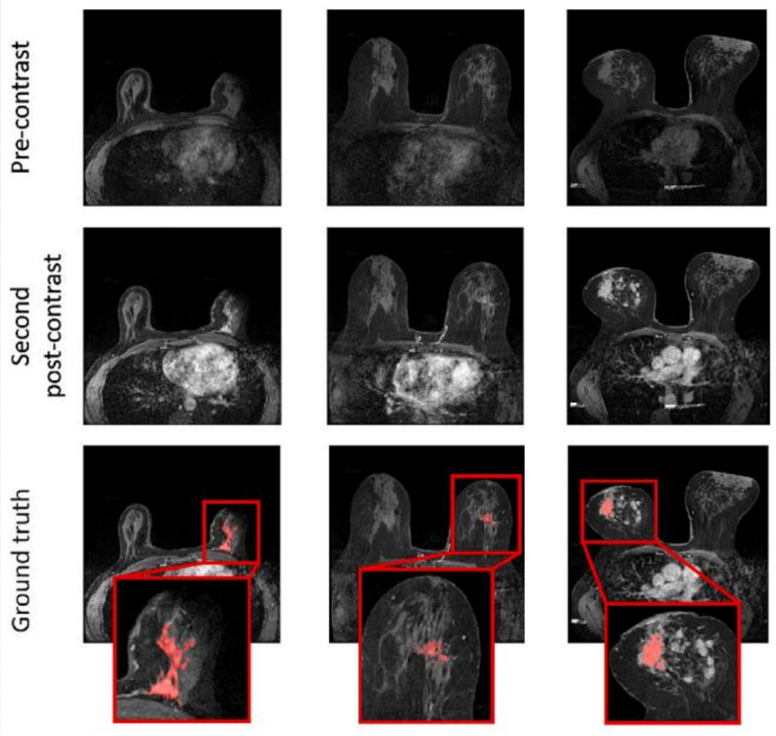
☆ Original 4-layered U-Net is selected and modified in this study

2.2. Premise of the study

Aim: Developing an effective and accurate method for segmenting breast lesions in DCE-MRI using a modified deep learning model, U-Net ensemble.

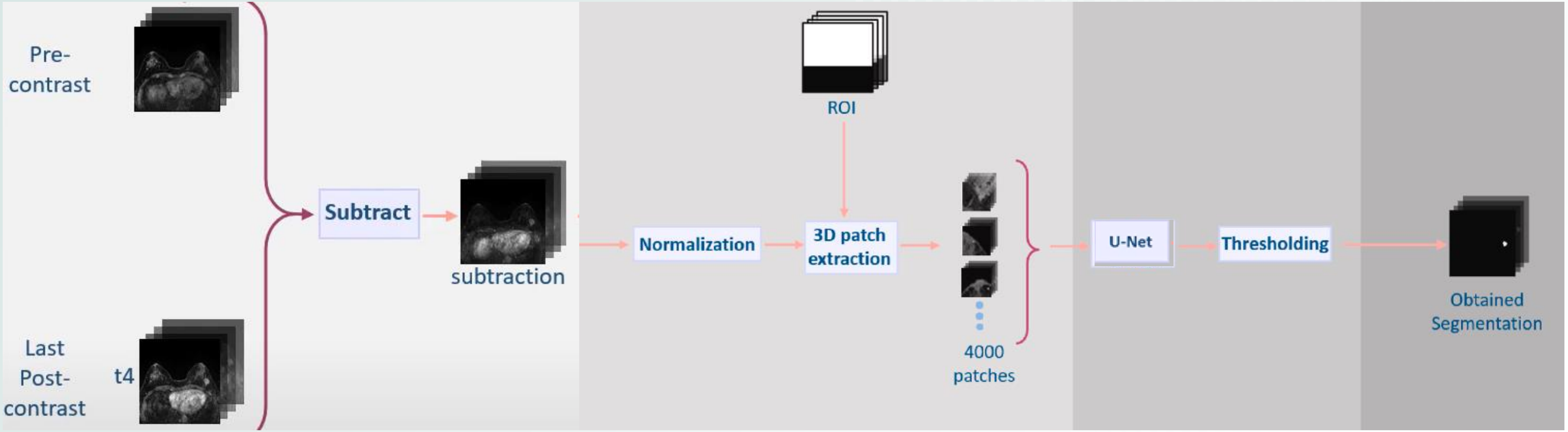
- Improving previous approach, Region of Intrest (ROI) guided, 3D patch based U-Net framework
- Incorporation of residual basic blocks
- Integrating three distinct models

3. 1. Data



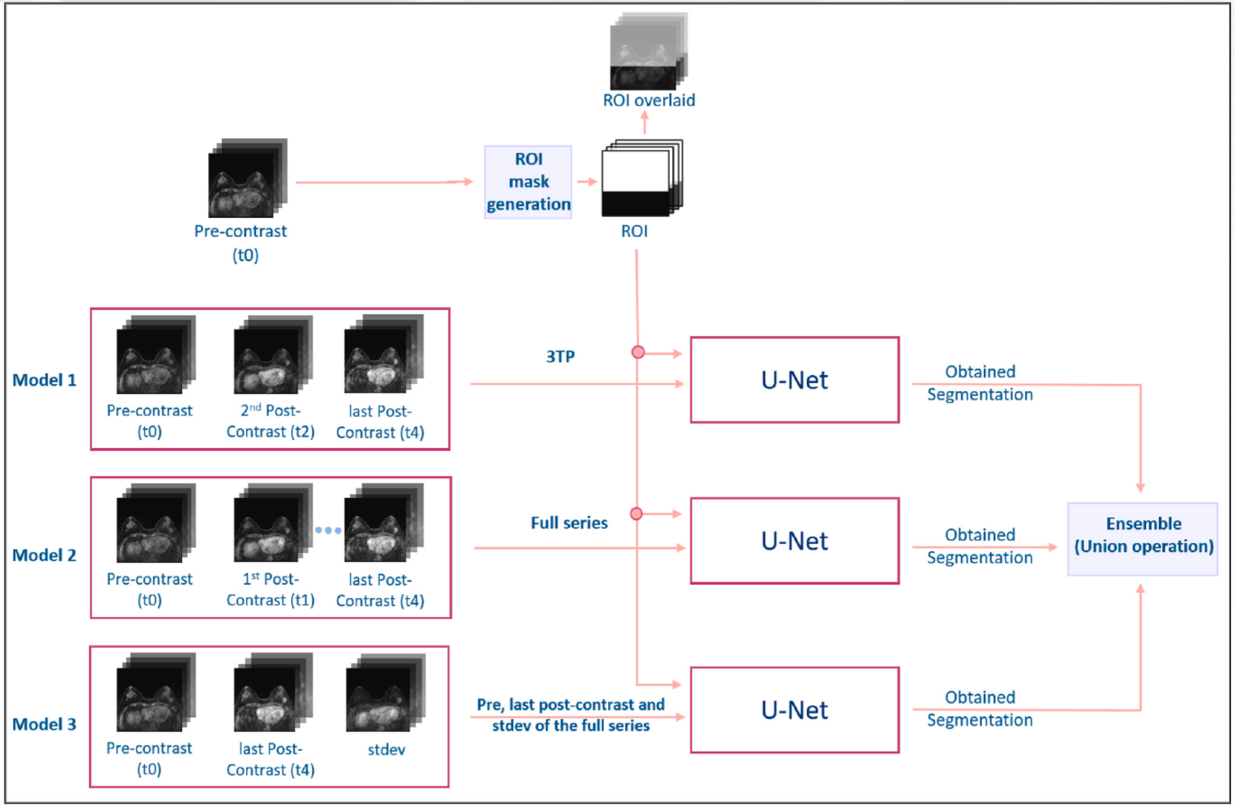
- 46 cases from the TCGA-BRCA
- All cases have at least one lesion (verified by needle biopsy)
- Complex and challenging: various sizes, shapes, locations and intensity enhancements also multiple lesions

3.2. Framework of U-Net with 3D patching method

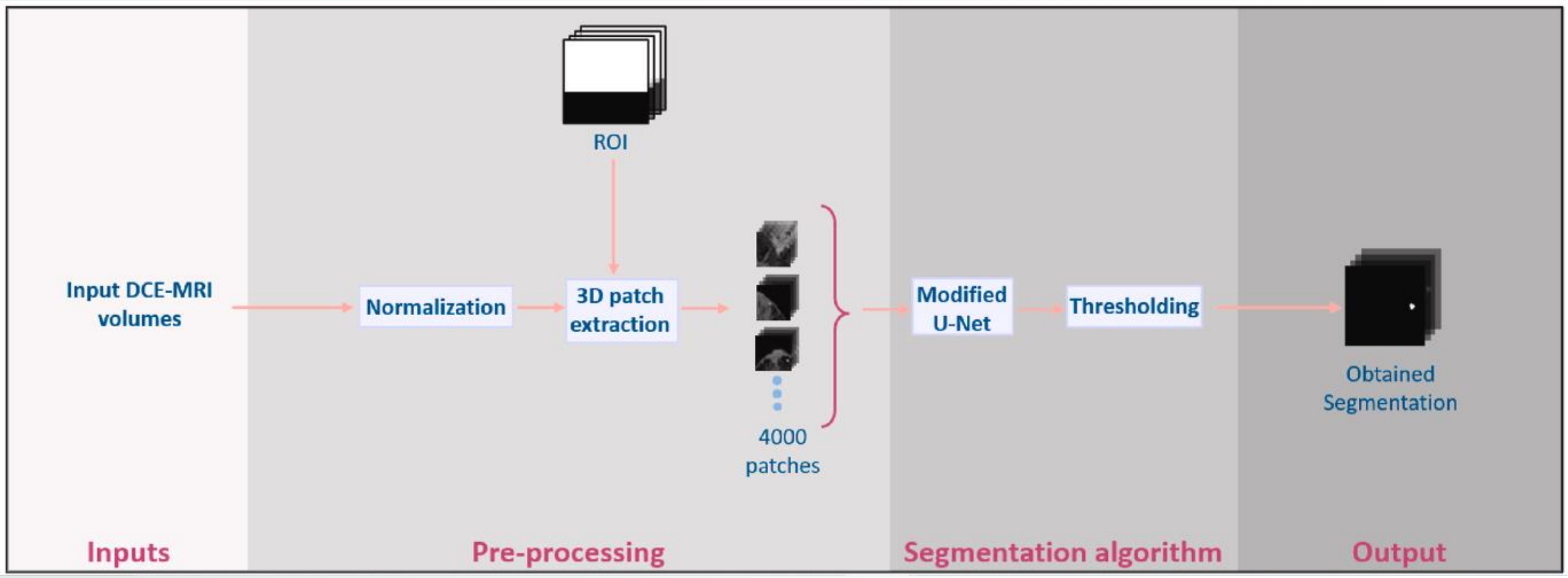


(Khaled et al., 2021)

3.3. Framework of proposed method with Modified U-Net

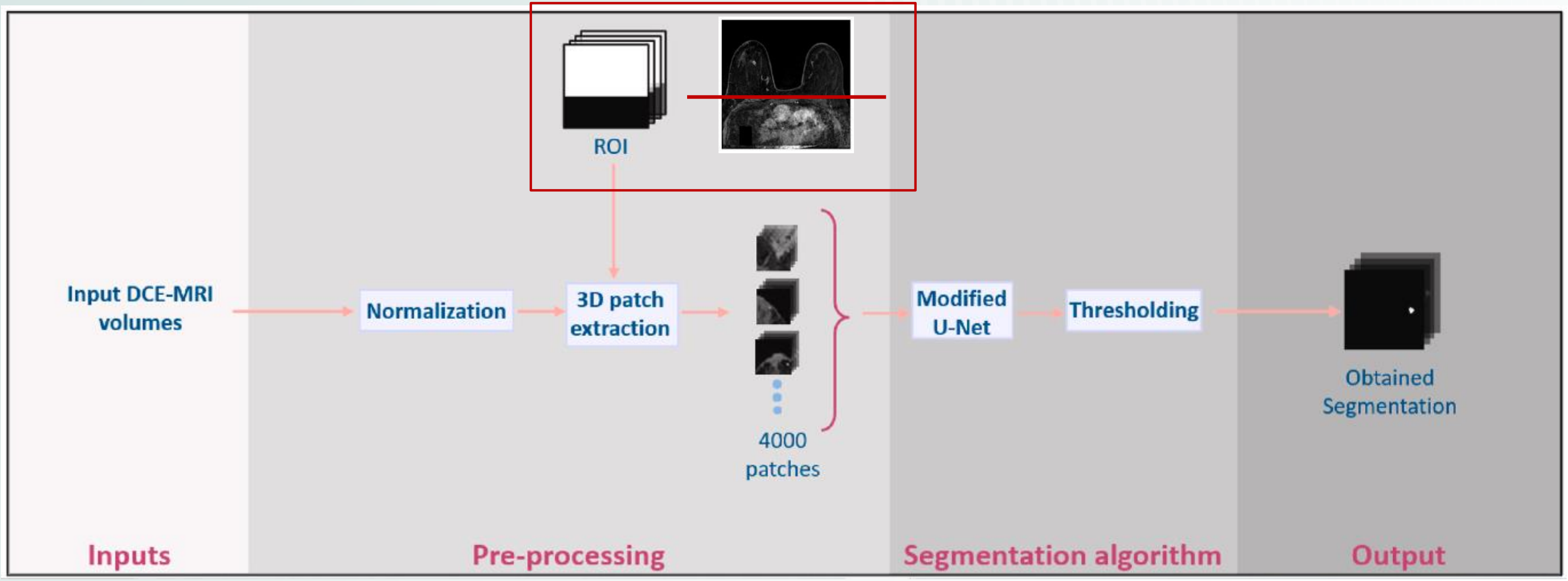


3.3. Framework of proposed method with Modified U-Net

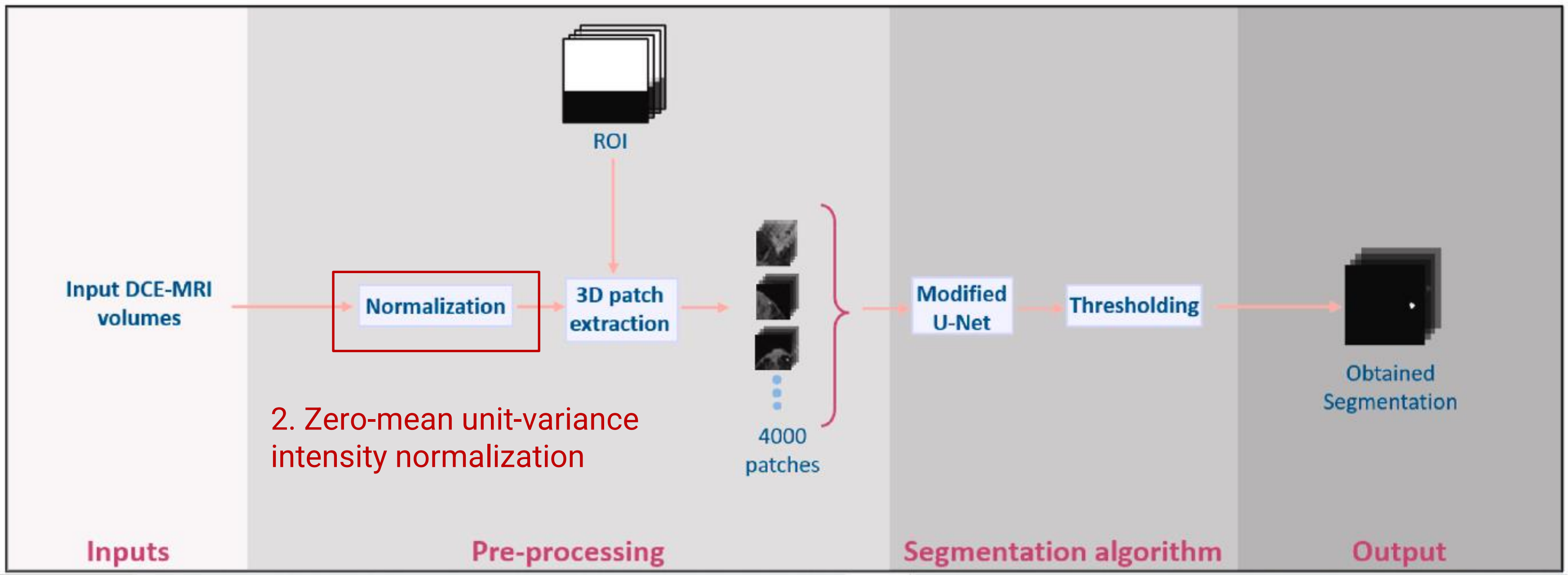


3.3. Framework of proposed method with Modified U-Net

1. Breast ROI masks generation

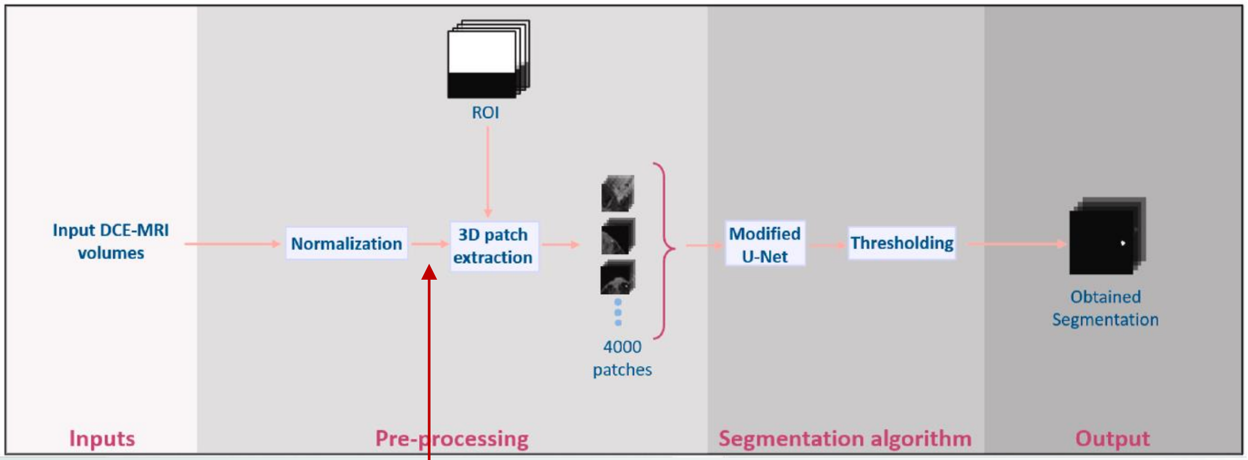


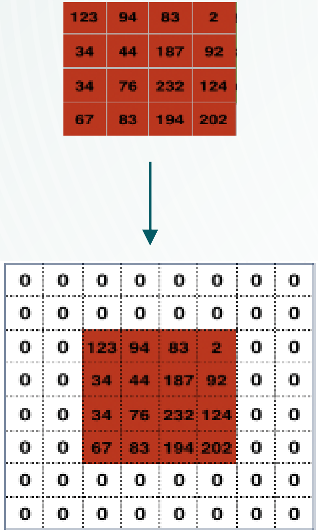
3.3. Framework of proposed method with Modified U-Net



2. Zero-mean unit-variance intensity normalization

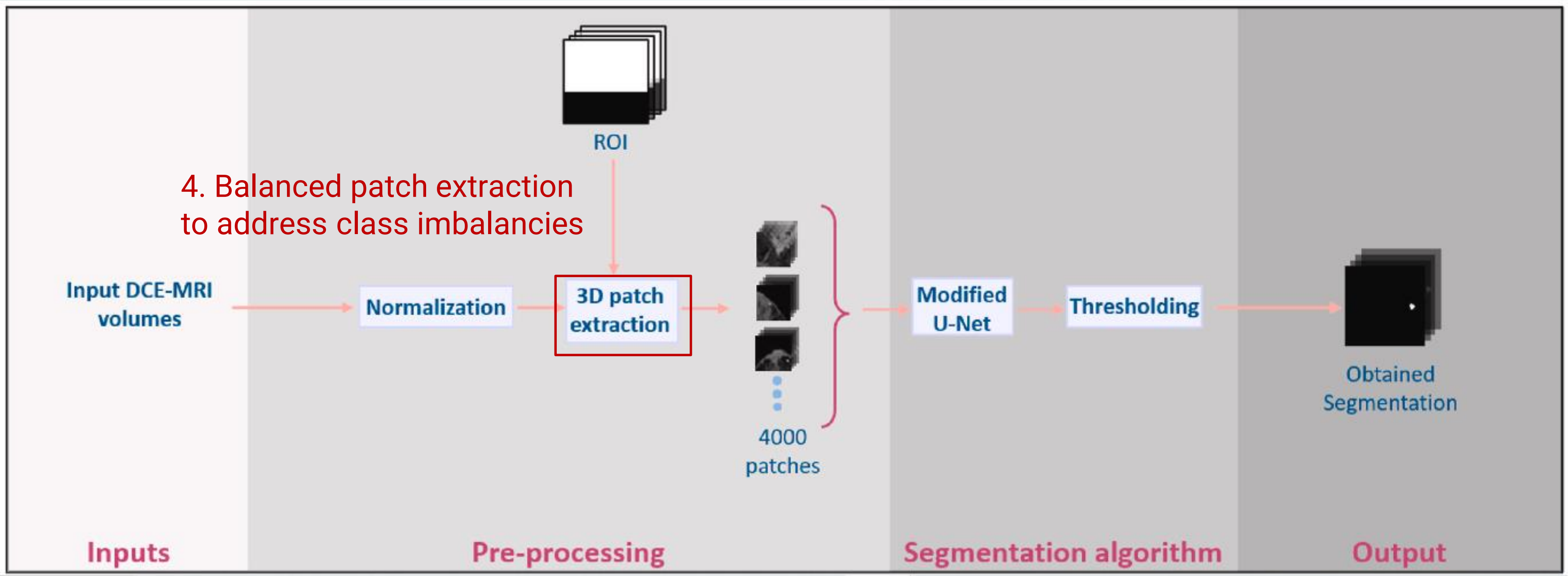
3.3. Framework of proposed method with Modified U-Net





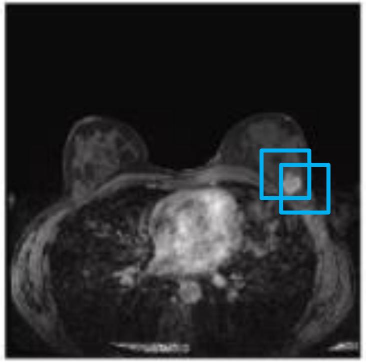
3. Zero padding with padding width equal to half of the patch size

3.3. Framework of proposed method with Modified U-Net

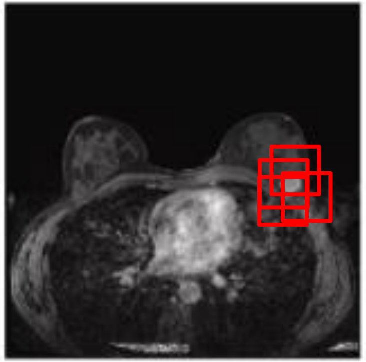


3.3. Balanced Patch Extraction

- Obtaining coordinates of positive voxels from the ground truth
- Random shifting of patch centers
- Generating more positive patches



Uniform patch sampling

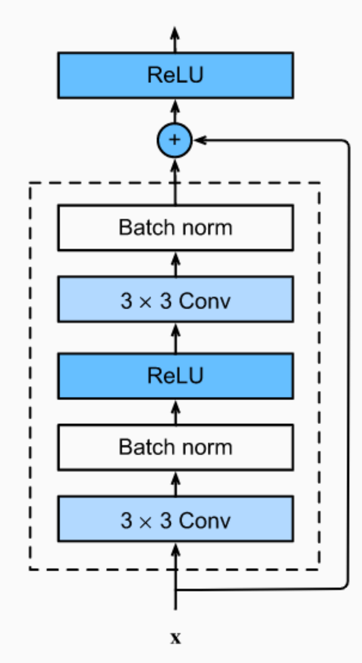


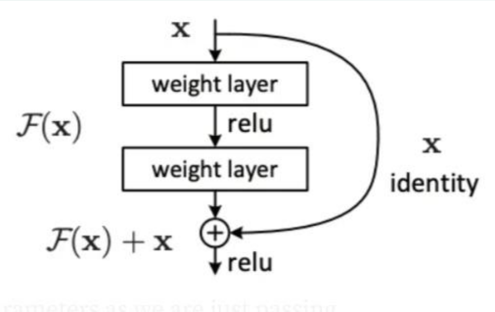
Balanced patch sampling

Per case:

- *2000 positive patches*
- *2000 negative patches*

3.4. Residual basic blocks



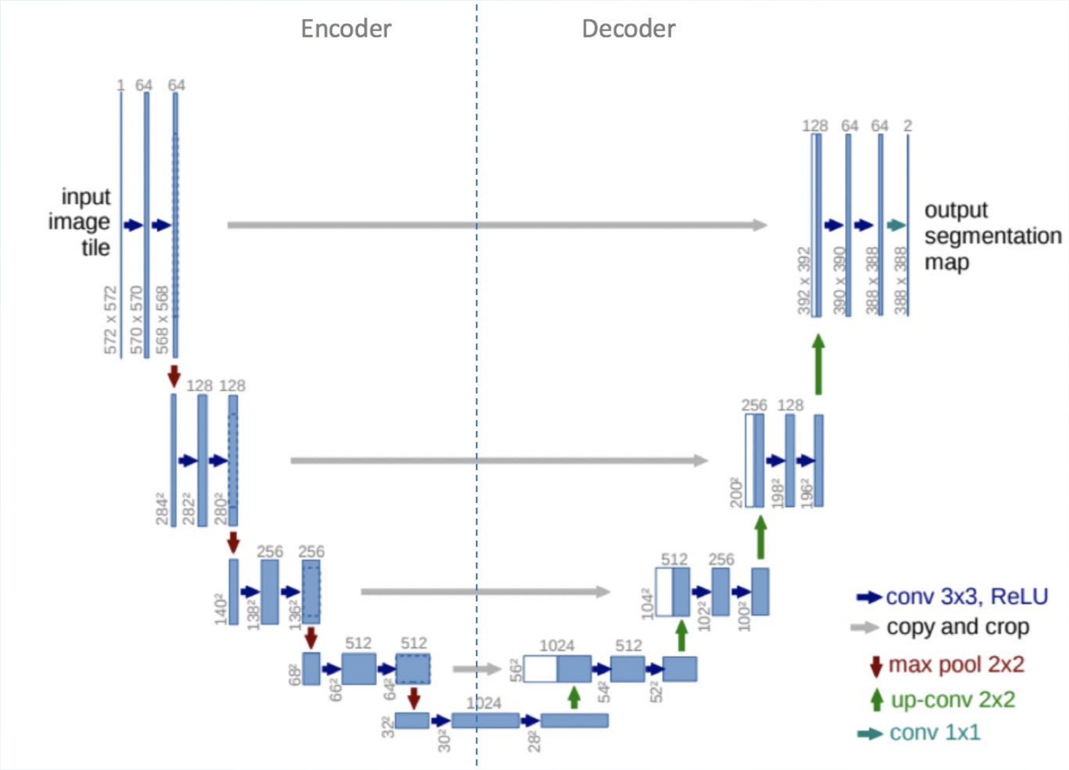


(Sahoo, 2022)

- Batch normalization layer then, a non-linear activation function.
- Layer fed directly into the layers about 2–3 hops next
- Skipping connections convenient in deep networks.

3. 5. U-Net Architectures

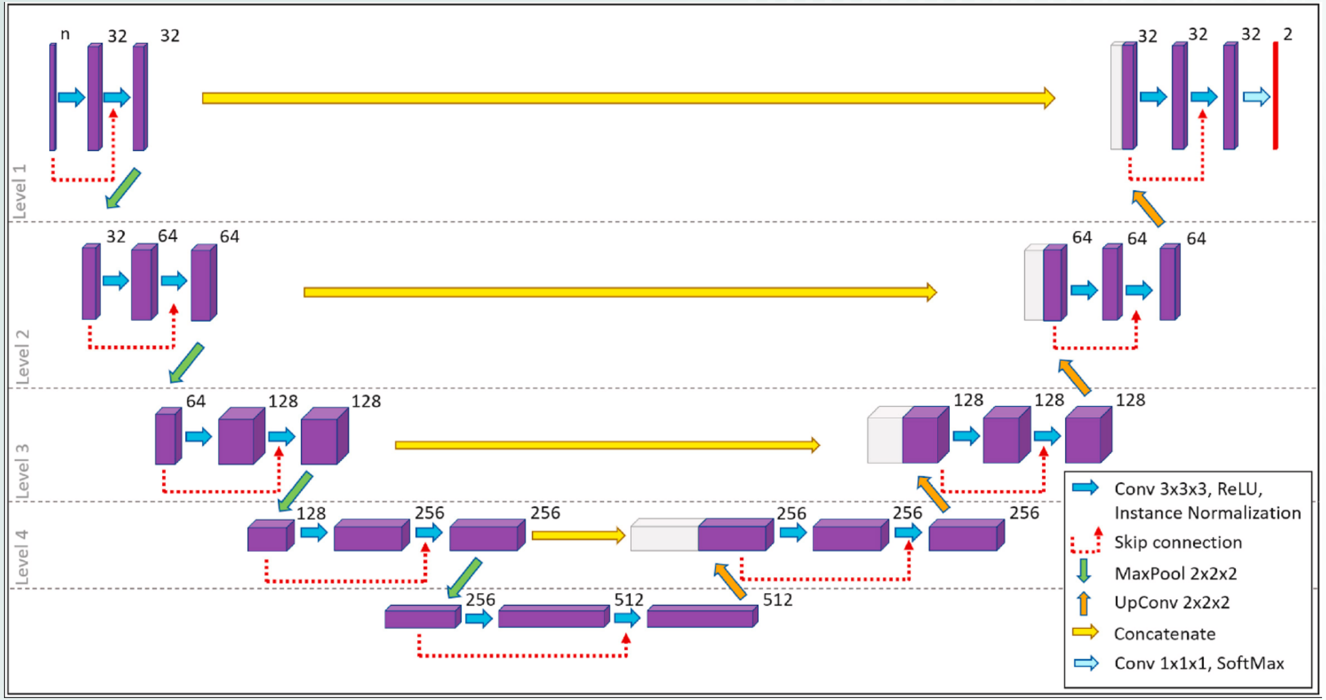
- Original:



(Ronneberger, O. et al., 2015)

3. 5. U-Net Architectures

- ❑ Modified



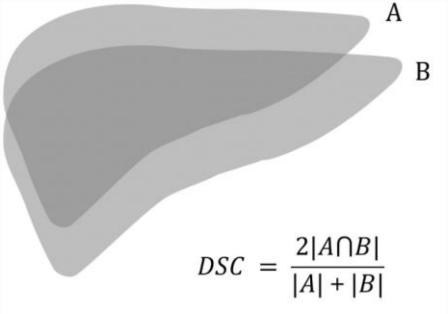
4. Experiments

Experiment 1: Volume Re-sampling	***Experiment 2:*** Input Volumes	***Experiment 3:*** U-Net Architecture	***Experiment 4:*** Detection evaluation and complemented GT	***Experiment 5:*** Comparison with non-linear methods

- ✓ 5 fold cross-validation
- ✓ 20 epochs per fold
- ✓ 9 cases for testing set (10 cases for the last fold).
- ✓ Remainings are randomly shuffled ⇨ 80% for training and 20% for testing

4. 1. Metrices for performance assesments

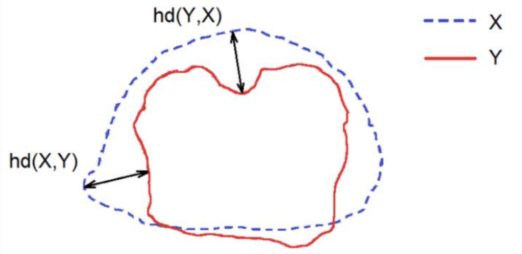
Dice Similarity Coefficient (DSC): Measures the overlap between two sets of data, which are sets of pixels or voxels representing segmented regions in two different images.



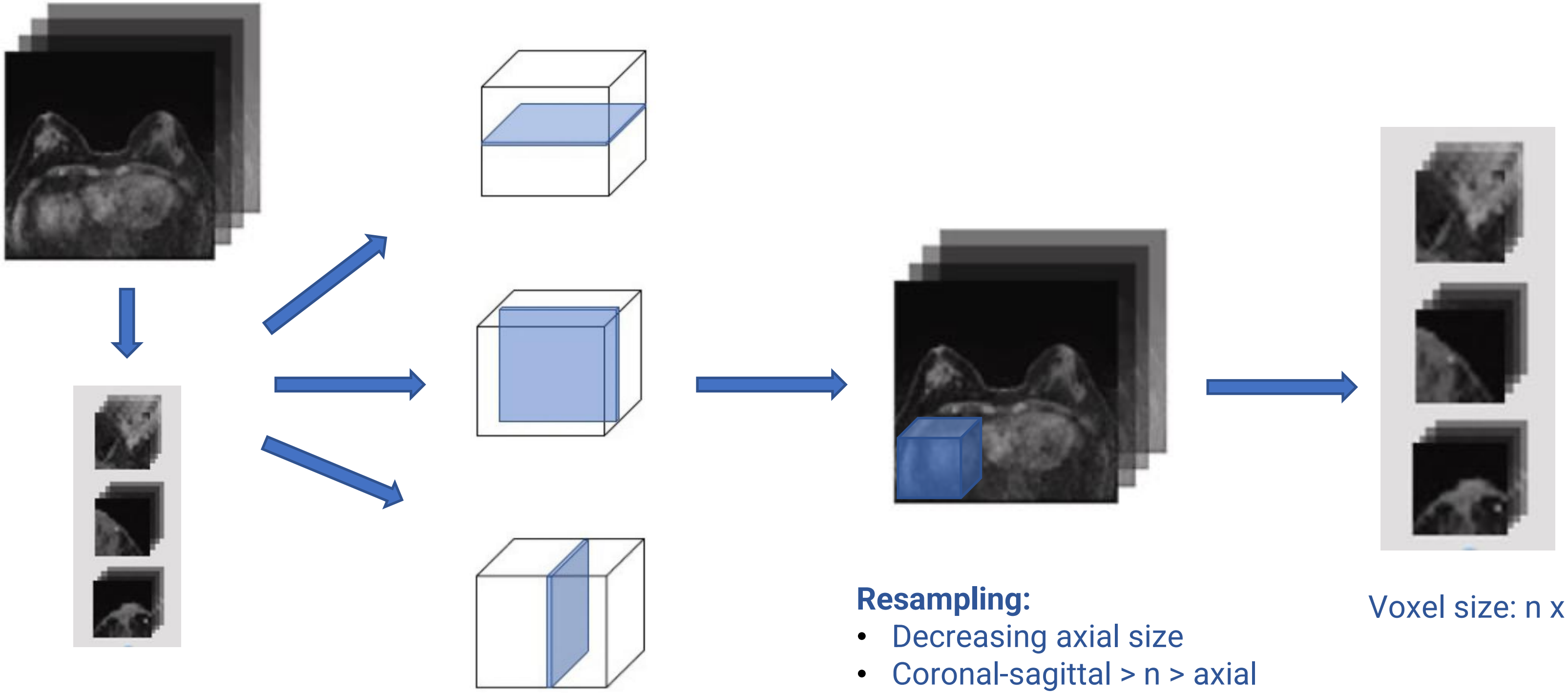
DSC: Dice similarity coefficient

Hausdorff Distance (HD): Measure of the dissimilarity between two sets of points. It quantifies the maximum distance between each point in one set and its closest point in the other set.

$$\tilde{\delta}_H(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$



Experiment 1: Volume re-sampling



Resampling:
- Decreasing axial size
- Coronal-sagittal > n > axial

Voxel size: n x n x n

3D Patches with various sizes

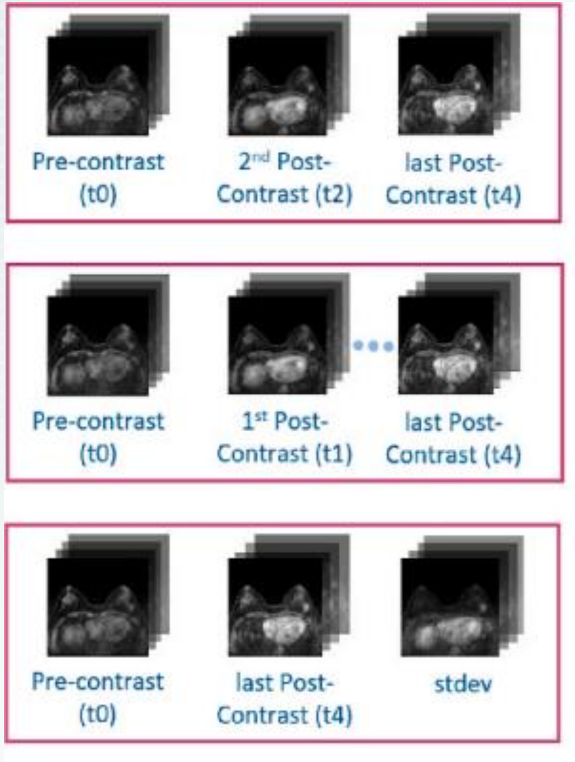
Results of volume resampling

Voxel Size	*DSC*	*DSC_L*
Original (non-isotropic)	**0.649±0.258**	**0.719 ±0.250**
Isotropic 1	0.580±0.274 (p=.008)	0.693±0.267 (p=.200)
Isotropic 2	0.579±0.260 (p=.003)	0.633±0.277 (p< .001)

⇨ Isotropic volumes did not improve the performance

Experiment 2: Input volumes

1. Pre and last post contrast
2. Pre, last post contrast substaraction
3. 3 TP (Pre, second post, last post contrast)
4. Full series (Pre, all post contrasts)
5. Pre, last post, stdev of full series
6. Ensemble (Majority voting) of 3, 4 and 5
7. Ensemble (Union) of 3, 4 and 5

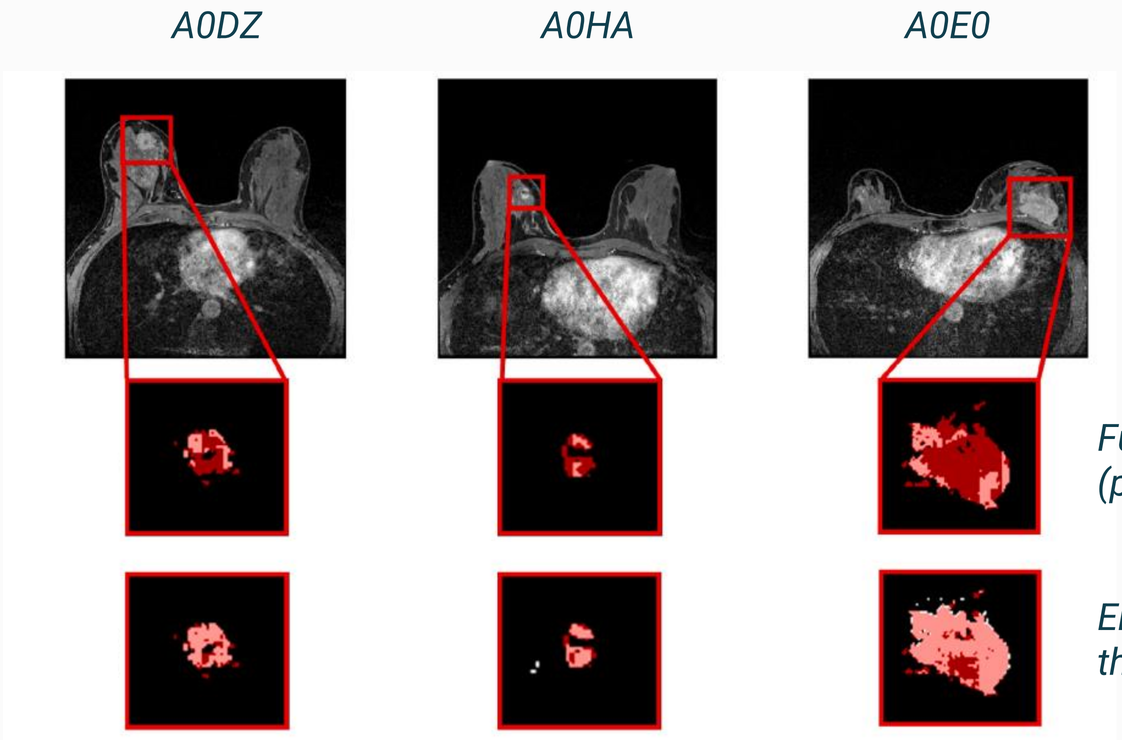


Experiment 2: Input volumes

1. Pre and last post contrast
2. Pre, last post contrast substaraction
3. 3 TP (Pre, second post, last post contrast)
4. Full series (Pre, all post contrasts)
5. Pre, last post, stdev of full series
6. Ensemble (Majority voting) of 3, 4 and 5
7. Ensemble (Union) of 3, 4 and 5

Input Model	DSC_L	HD_L
1	0.652 ± 0.283	23.281
2	0.649 ± 0.248	16.291
3	0.716 ± 0.241	11.996
4	0.730 ± 0.222	11.879
5	0.719 ± 0.250	13.098
6	0.728 ± 0.243	11.561
7	**0.790 ± 0.172**	**11.256**

Results: Input volumes



- *Ground truth: Red*
- *Output segmentation: White*

Full series (pre, all posts)

Ensemble of three methods

Experiment 3: U-Net architecture

Basic U-Net with four levels along with a ***cross-entropy*** loss function.

Two hierarchical basic U-Nets with ***dice-sensitivity-like*** and ***dice-like loss***

U-Net architecture with ***residual blocks***

Results: U-Net architecture

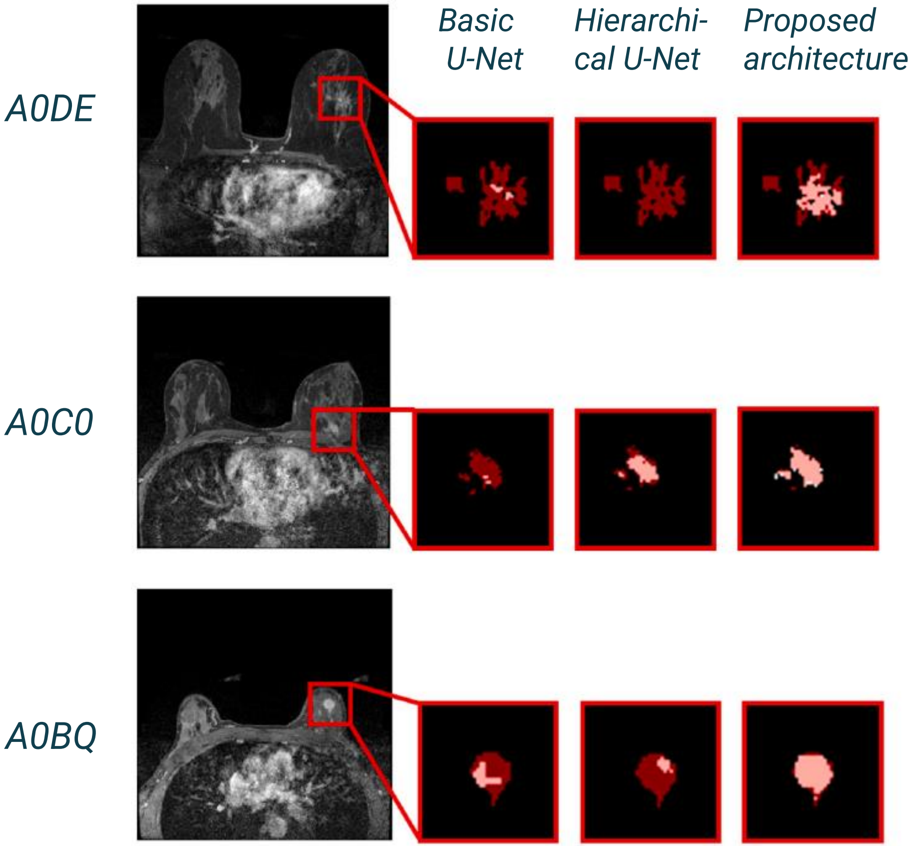
Basic U-Net with four levels along with a ***cross-entropy*** loss function.

Two hierarchical basic with ***dice-sensitivity-like*** and ***dice-like loss***

U-Net architecture with ***residual blocks***

Network	***DSC***	***DSC_L***	***HD***	***HD_L***	**FPR**
Single Model Basic U-Net	0.551 ±0.286 (p=.001)	0.590±0.302 (p<.001)	123.939 (p <.001)	16.369 (p=.001)	1.21E-04 (p=.910)
Single Model Two hierarchical U-Nets	0.615±0.266 (p=.277)	0.669±0.275 (p=.137)	212.123 (pp=.320)	15.885 (p=.034)	1.29E-04 (p=.729)
Single Model U-Net with residual blocks	**0.649 ± 0.258**	**0.719 ± 0.250**	226.873	**13.097**	1.24E-04

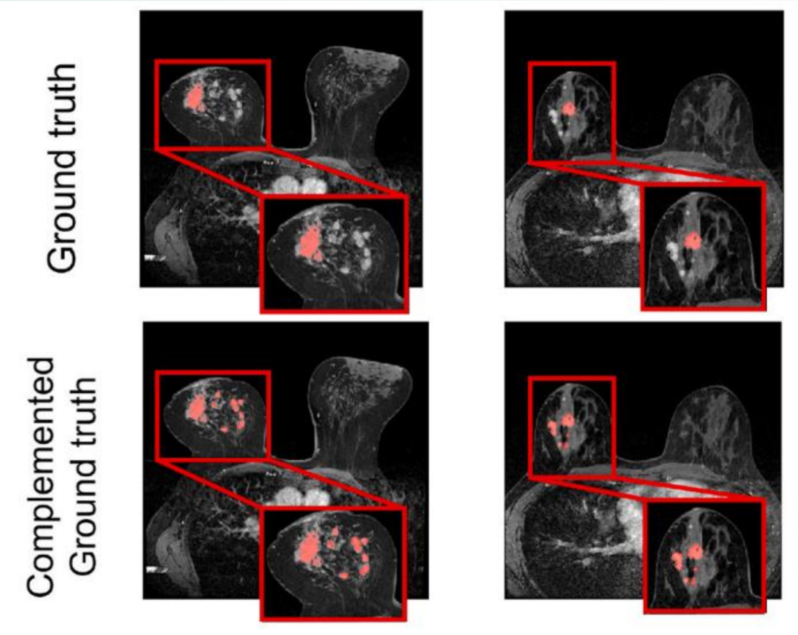
Results: U-Net architecture



- ❑ *Ground truth: Red*
- ❑ *Output segmentation: White*

Experiment 4: Detection evaluation and complemented GT

- Only one lesion per case (primary lesion)
- Incomplete ground truth annotations affects the results → lowers dice values.
- Annotations were complemented with the help of an experienced radiologist (11 of 46 cases)



Results: Detection evaluation and complemented GT

Data	*DSC*	DSC_L	*HD*	HD_L	FPR
Before complementing	0.674±0.224	0.790±0.172	278.548	11.256	1.73E-04
After complementing	**0.680 ± 0.221**	**0.802 ± 0.156**	**275.013**	**11.042**	**1.71E-04**

⇩

Improvement was significant only marginally

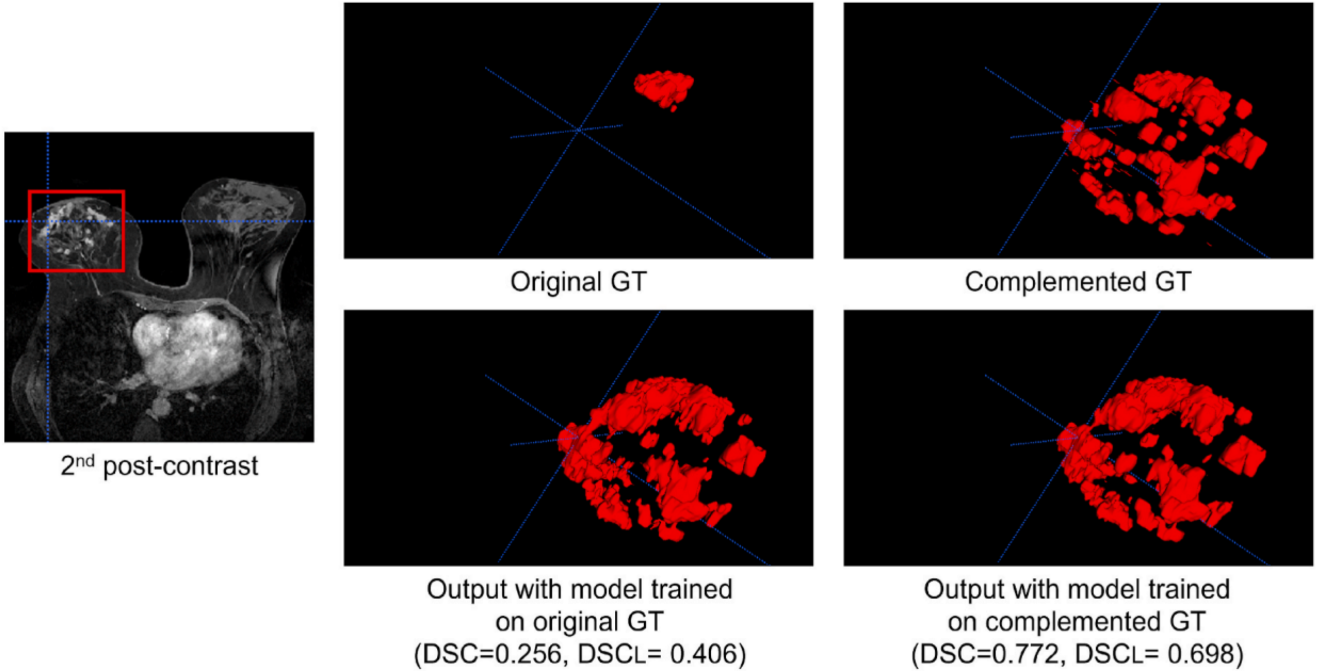
Results: Detection evaluation and complemented GT

Data	Detected Lesions (I >= 0.2)	Detected Lesions (I >= 0.5)	FP lesions (before filtering)	FP lesions (after filtering)	Mean I
Before complementing	97.8% (45/46 lesions)	89.1% (41/46 lesions)	1736	382	0.811
After complementing (all 46 cases)	78.8% (89/113 lesions)	69% (78/113 lesions)	1834	426	0.761
After complementing (excluding case A0B1)	89.6% (60/67 lesions)	82.1% (55/67 lesions)	1809	424	0.785

Complicated multiple lesions with many small lesions.

Complexity of the dataset and the lesions that were added

Results: Detection evaluation and complemented GT



Experiment 5: Comparison with non-learning methods (Fuzzy c-means)

Data	***DSC***	***DSC_L***	***HD***	***HD_L***	**FPR**
Before complementing	0.102±0.092	**0.560 ± 0.340**	365.340	21.439	3.57E-03
After complementing	$\boxed{\mathbf{0.112 \pm 0.105}}$	0.549 1±0.33	**360.923**	**21.164**	**3.55E-03**

Experiment 5: Comparison with non-learning methods (Fuzzy c-means)

Data	Detected Lesions (I >= 0.2)	Detected Lesions (I >= 0.5)	FP lesions (before filtering)	FP lesions (after filtering)	Mean I
Before complementing	97.8% (45/46 lesions)	82.6% (38/46 lesions)	113 234	12 520	0.730
After complementing (all 46 cases)	96.5% (109/113 lesions)	78.8% (89/113 lesions)	113 425	12 711	0.687
After complementing (excluding case A0B1)	94% (63/67 lesions)	70.1% (47/67 lesions)	112 195	12 547	0.675

Number of FP lesion detections increases (25-fold of Modified U-Net)

Comparison with existing deep-learning methods

	Architecture	2D/ 3D	Number of cases (public/private)	Inputs	Loss function	Evaluation criteria and score	Scanner
Zhang et al., 2019b [34]	U-Net	2D	1246 slices (private)	2nd post-contrast (lesion bounding boxes)	Cross-entropy	DSC = 0.91	–
	U-Net	3D	158 cases (private)	2nd post-contrast (lesion bounding boxes)	Cross-entropy	DSC = 0.92	
El Adoui et al., 2019 [15]	U-Net	2D	5452 slices (private)	Post-contrast	Cross-entropy	IoU = 0.761 4	1.5T Siemens
	SegNet	2D	5452 slices (private)	Post-contrast	Cross-entropy	IoU = 0.688 8	
Piantadosi et al., 2019 [21]	U-Net	2D	35 case (256x128x80) (private)	Pre-contrast, 2 min post-contrast, and 6 min post-contrast	Dice	DSC = 0.612 4	1.5T Siemens
Zhang et al., 2019a [33]	Two hierarchical U-Nets	3D	272 cases (private)	Pre-contrast, post-contrast, and subtraction (breast mask guided)	First stage: Dice-sensitivity-like Second stage: Dice-like	DSC = 0.72	1.5T GE and 3.0T Siemens
Our proposed work	Ensemble of 3 U-Nets with ResNet basic blocks	3D	46 cases (public)	Ensemble (Union) of 3 models, each with different inputs (ROI mask guided)	Cross- entropy	DSC = 0.680 (0.802 for primary lesions only)	1.5T GE

Comparison with existing deep-learning methods

	Architecture	2D/ 3D	Number of cases (public/private)	Inputs	Loss function	Evaluation criteria and score	Scanner
Zhang et al., 2019b [34]	U-Net	2D	1246 slices (private)	2nd post-contrast (lesion bounding boxes)	Cross-entropy	DSC = 0.91	–
	U-Net	3D	158 cases (private)	2nd post-contrast (lesion bounding boxes)	Cross-entropy	DSC = 0.92	
El Adoui et al., 2019 [15]	U-Net	2D	5452 slices (private)	Post-contrast	Cross-entropy	IoU = 0.761 4	1.5T Siemens
	SegNet	2D	5452 slices (private)	Post-contrast	Cross-entropy	IoU = 0.688 8	
Piantadosi et al., 2019 [21]	U-Net	2D	35 case (256x128x80) (private)	Pre-contrast, 2 min post-contrast, and 6 min post-contrast	Dice	DSC = 0.612 4	1.5T Siemens
Zhang et al., 2019a [33]	Two hierarchical U-Nets	3D	272 cases (private)	Pre-contrast, post-contrast, and subtraction (breast mask guided)	First stage: Dice-sensitivity-like Second stage: Dice-like	DSC = 0.72	1.5T GE and 3.0T Siemens
Our proposed work	Ensemble of 3 U-Nets with ResNet basic blocks	3D	46 cases (public)	Ensemble (Union) of 3 models, each with different inputs (ROI mask guided)	Cross- entropy	DSC = 0.680 (0.802 for primary lesions only)	1.5T GE

Lesion bounding boxes are used as input

Fully automatic

5. Limitations

1. Dataset had a small number of cases
2. No healthy cases
3. All scans used in this study were acquired with the same scanner
4. ROI could also include surrounding confounding organs in few cases

5. Limitations - Future work

1. **Bigger dataset**
   Dataset had a small number of cases

2. **Include healthy cases**
   No healthy cases

3. **Effects of different scanners to segmentation accuracy**
   All scans used in this study were acquired with the same scanner

4. **Breast masks instead of breast ROIs**
   ROI could also include surrounding confounding organs in few cases

Limitations - Future work

- ❑ Using more complex models
- ❑ Representing of the 4D data in a single volume as a compressed input
- ❑ Effective capture of TIC for each voxel

⇩

★ Decrease the computational requirements

★ Enhance the classification performance

5. References

- Khazaei, Z. *et al.* (2018) “The incidence and mortality of cancer in Eastern Mediterranean Regional Office (EMRO) and its relationship with human development index (HDI): An ecological study,” *Asian Pacific Journal of Environment and Cancer*, 1(2). Available at: https://doi.org/10.31557/apjec.2018.1.2.59.
- Hendrick, R.E., Helvie, M.A. and Monticciolo, D.L. (2021) “Breast cancer mortality rates have stopped declining in U.S. women younger than 40 Years,” *Radiology*, 299(1), pp. 143–149. Available at: https://doi.org/10.1148/radiol.2021203476.
- Sree, S.V. (2011) “Breast Imaging: A survey,” *World Journal of Clinical Oncology*, 2(4), p. 171. Available at: https://doi.org/10.5306/wjco.v2.i4.171.
- Chikui, T. *et al.* (2012) “The principal of dynamic contrast enhanced MRI, the method of pharmacokinetic analysis, and its application in the head and Neck Region,” *International Journal of Dentistry*, 2012, pp. 1–10. Available at: https://doi.org/10.1155/2012/480659.
- Zhang, L. *et al.* (2019) “Automated Deep Learning Method for whole-breast segmentation in diffusion-weighted breast mri,” *Journal of Magnetic Resonance Imaging*, 51(2), pp. 635–643. Available at: https://doi.org/10.1002/jmri.26860.
- Khaled, R., Vidal, J. and Martí, R. (2021) “Deep learning based segmentation of breast lesions in DCE-MRI,” Pattern Recognition. ICPR International Workshops and Challenges, pp. 417–430. Available at: https://doi.org/10.1007/978-3-030-68763-2_32.

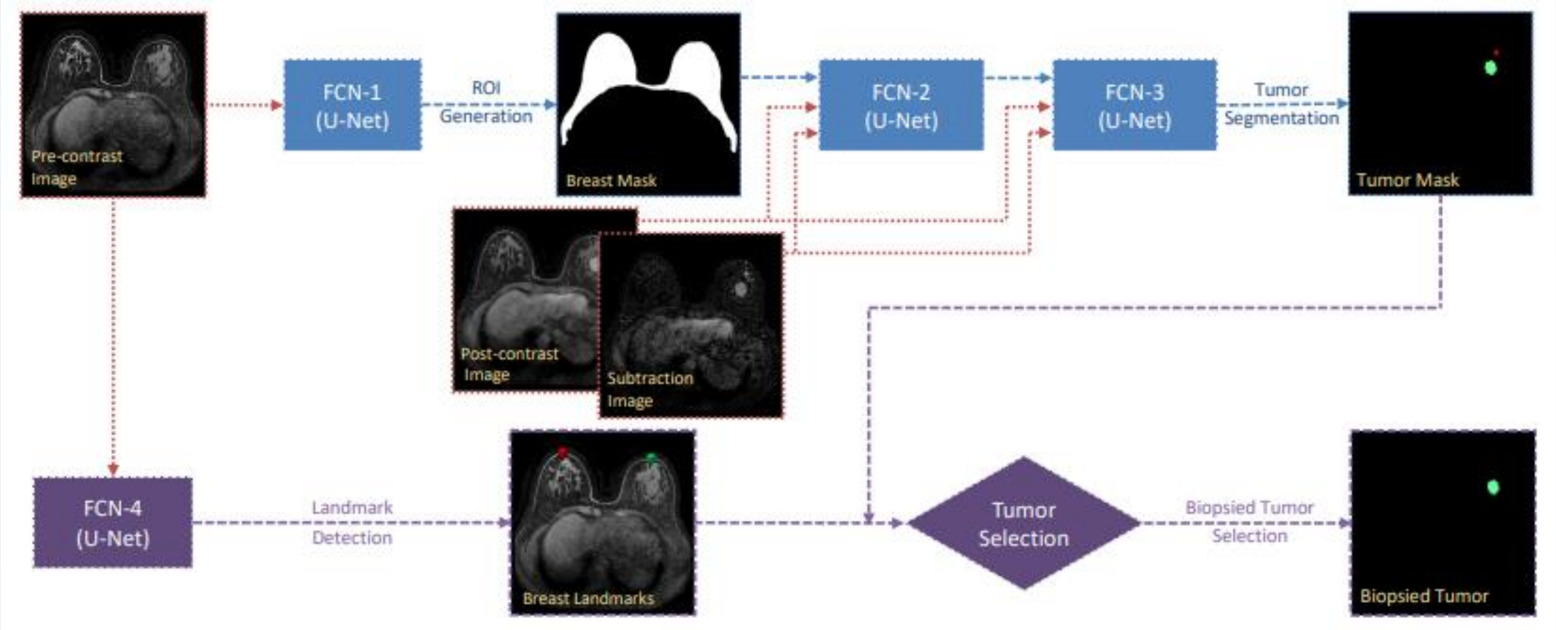
5. References

- Ronneberger, O., Fischer, P. and Brox, T. (2015) “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *Lecture Notes in Computer Science*, pp. 234–241. Available at: https://doi.org/10.1007/978-3-319-24574-4_28.
- A. Gubern-M´erida, R. Martí, J. Melendez, J. Hauth, R. Mann, N. Karssemeijer, B. Platel, Automated localization of breast cancer in DCE-MRI, Med. Image Anal. 20 (2014) 265–274, https://doi.org/10.1016/j.media.2014.12.001
- L. Hu, Z. Cheng, M. Wang, Z. Song, Image manifold revealing for breast lesion segmentation in DCE-MRI, Bio Med. Mater. Eng. 26 (2015) S1353–S1360, https://doi.org/10.3233/BME-151433.
- Badrinarayanan, A. Kendall, R. Cipolla, SegNet: a deep convolutional encoderdecoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 2481–2495, https://doi.org/10.1109/TPAMI.2016.2644615.
- Zhang, J. et al. (2019) “Hierarchical convolutional neural networks for segmentation of breast tumors in MRI with application to radiogenomics,” IEEE Transactions on Medical Imaging, 38(2), pp. 435–447. Available at: https://doi.org/10.1109/tmi.2018.2865671.
- Sahoo, S. (2022) Residual blocks-building blocks of Resnet, Medium. Towards Data Science. Available at: https://towardsdatascience.com/residual-blocks-building-blocks-of-resnet-fd90ca15d6ec (Accessed: February 24, 2023).

THANK YOU

SUPPLEMENTARY

Mask-guided hierarchical learning (MHL) U-Net framework



(Zhang et al., 2019)

- A FCN-1 model is trained to estimate the breast mask, and two additional FCN-2, 3 models are trained to obtain rough and refined segmentation results.
- A landmark detection model is also developed to identify biopsied tumors. These components are presented in detail in the following sections.

Dice-like loss function

$$\Omega_2(\mathbf{W_3}) = 1 - \underbrace{\frac{1}{N}\sum_{n=1}^{N} \frac{2\sum_{v=1}^{V} S_{n,v}\, g(X'_{n,v}, \mathbf{W_3})}{\sum_{v=1}^{V} S_{n,v}^2 + \sum_{v=1}^{V} g(X'_{n,v}, \mathbf{W_3})^2}}_{\text{DSC}},$$

- *g(X'n,v,W3) is the estimated probability map by using the network coefficients W3 of FCN-3*
- *X'n,v is the v-th (v = 1, · · · , V) voxel of the input data X0 n with 3 channel*

- ❑ 1 minus the Dice coefficient is used to minimize the distance between the predicted segmentation and the ground truth segmentation.
- ❑ The Dice loss function is particularly useful for imbalanced datasets

Dice-sensitivity-like loss function

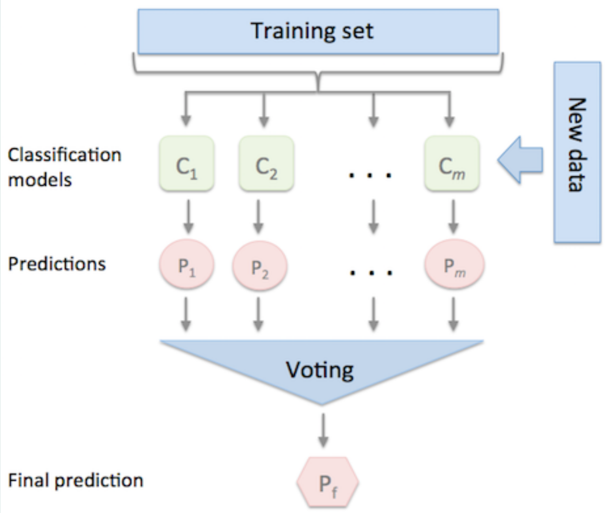
$$\Omega_1(\mathbf{W_2}) = 2 - \underbrace{\frac{1}{N}\sum_{n=1}^{N}\frac{2\sum_{v=1}^{V} S_{n,v} f(X_{n,v},\mathbf{W_2})}{\sum_{v=1}^{V} S_{n,v}^2 + \sum_{v=1}^{V} f(X_{n,v},\mathbf{W_2})^2}}_{\text{DSC}}$$

$$- \underbrace{\frac{1}{N}\sum_{n=1}^{N}\frac{\sum_{v=1}^{V} S_{n,v} f(X_{n,v},\mathbf{W_2})}{\sum_{v=1}^{V} S_{n,v}^2}}_{\text{SEN}},$$

- *f(Xn,v,W2) is the estimated probability for Xn,v by using the network coefficients W2 of FCN-2*
- *N is the number of training images in a batch.*

- ❑ The sensitivity loss function is used to penalize false negatives, which occur when the model fails to detect a positive sample.
- ❑ DSLL function that not only penalizes incorrect segmentations, but also penalizes false negatives.

Union vs. Majority voting



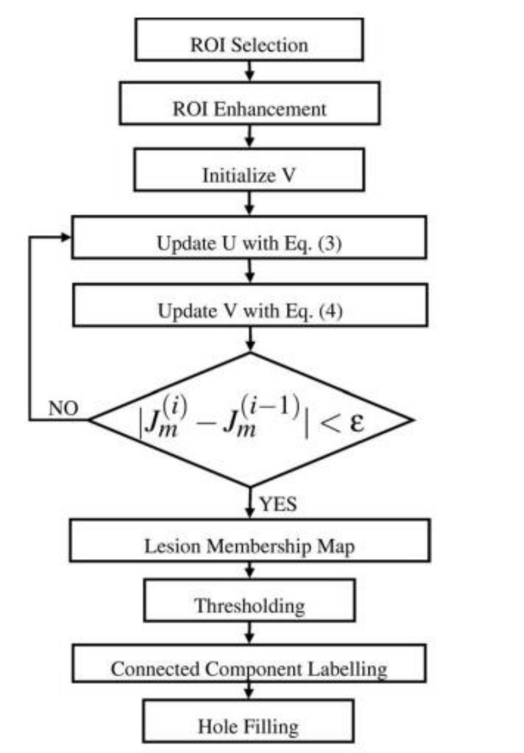
Majority voting:

- Each model predicts a class label for the image.
- Final prediction is the label that receives the most votes.

Union operation:

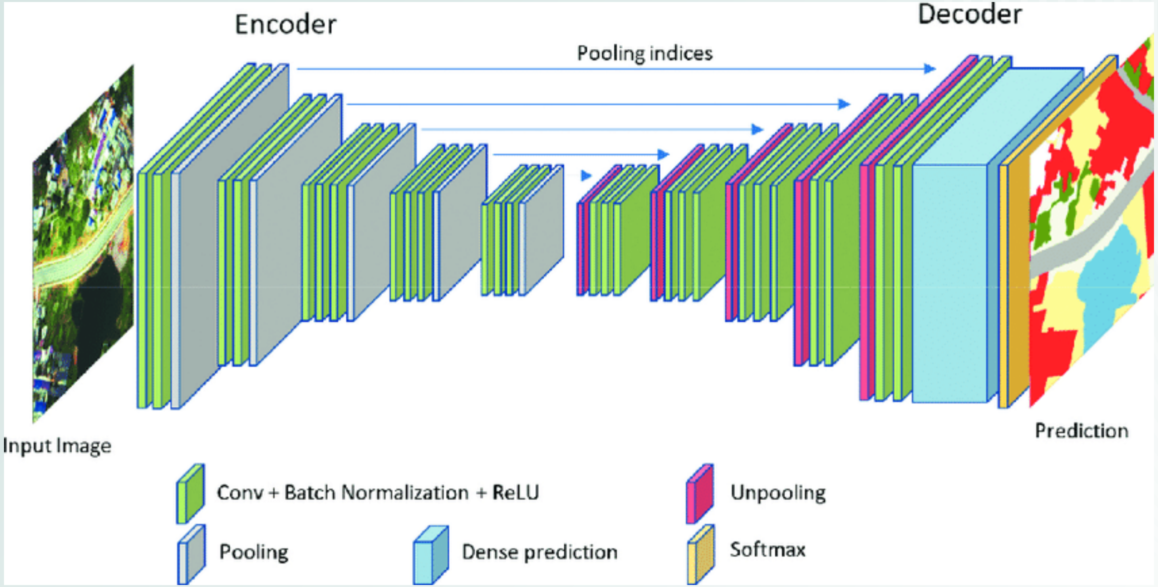
- Merging the predictions of multiple models into a single prediction by taking the union of all predicted object regions in the image
- Final segmentation will include all the labels that were assigned by any of the models

Fuzzy c-means



- Each data point is assigned a membership value to each cluster
- Belonging to multiple clusters with varying degrees of membership
- Cluster is represented by a centroid
- Iteratively updating the membership values and centroids until a convergence criterion is met
- *U: partition matrix, V: class centers*

SegNet



Badrinarayanan et al., 2017

- Pooling layers that progressively reduce, while U-Net's encoder consists of a series of convolutional layers that gradually increase the number of channels
- "unpooling" vs. "up-sampling"

Complete results: Input volumes

1. Pre and last post contrast

2. Pre, last post contrast substaraction

3. 3 TP (Pre, second post, last post contrast)

4. Full series (Pre, all post contrasts)

5. Pre, last post, stdev of full series

6. Ensemble (Majority voting) of 3, 4 and 5

7. Ensemble (Union) of 3, 4 and 5

Inputs	*DSC*	*DSC*$_L$	*HD*	*HD*$_L$	FPR
(1) Pre, last post	0.591 ± 0.279 (p = .001)	0.652 ± 0.283 (p<.001)	253.061 (p = .081)	23.481 (p = .141)	1.25E-04 (p = .001)
(2) Pre, last post, subtraction (post-pre)	0.587 ± 0.260 (p<.001)	0.649 ± 0.248 (p<.001)	250.651 (p = .036)	16.291 (p = .099)	1.42E-04 (p = .010)
(3) 3 TP (pre, 2nd post, last post)	0.654 ± 0.262 (p = .403)	0.716 ± 0.241 (p<.001)	219.712 (p<.001)	11.996 (p = .453)	1.07E-04 (p<.001)
(4) Full series (pre, all posts)	0.664 ± 0.234 (p = .572)	0.730 ± 0.222 (p = .005)	227.654 (p<.001)	11.879 (p = .135)	1.14E-04 (p<.001)
(5) Pre, last post, stdev of the full series	0.649 ± 0.258 (p = .220)	0.719 ± 0.250 (p = .004)	226.873 (p<.001)	13.098 (p = .315)	1.24E-04 (p<.001)
Ensemble (Majority voting) of (3), (4) and (5)	0.679 ± 0.258 (p = .835)	0.728 ± 0.243 (p = .002)	135.415 (p<.001)	11.561 (p = .593)	1.00E-04 (p<.001)
Ensemble (Union) of (3), (4) and (5)	0.674 ± 0.224	**0.790 ± 0.172**	278.548	**11.256**	1.73E-04