Patch-based V-Net for malignant breast lesion segmentation in 3T DCE-MRI

Fedon Vocaturo M.[(1)], Altabella L.[(1)], Esposito P. G. [(1)], Cardano G.[(2)], Cardobi N.[(2)], Montemezzi S.[(2)], Cavedon C.[(1)]

1 - Medical Physics Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy
2 - Radiology Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy

110° Congresso Nazionale SIF – 9-13 settembre 2024

Introduction

In medical imaging, accurately segmenting lesions is vital for further analysis, such as monitoring tumour progression, advanced analysis and extracting radiomic features.

Image segmentation can be performed in different ways:

- manually
- semiautomatically
- automatically

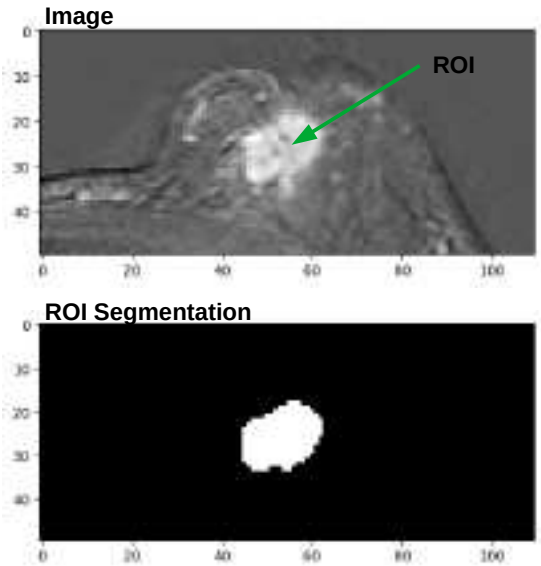


Figure: example of segmentation

Manual segmentation challenges

- high variability $\rightarrow$ undermines reproducibility
- time consuming procedure $\rightarrow$ small labelled datasets

Introduction: automatic segmentation

State of the art

Increasing use of U-shaped convolutional neural networks.

- **Contraction path:** captures context by progressively reducing input size and increasing the receptive field.
- **Expansion path:** reconstructs detailed spatial information, returning back to the input size.

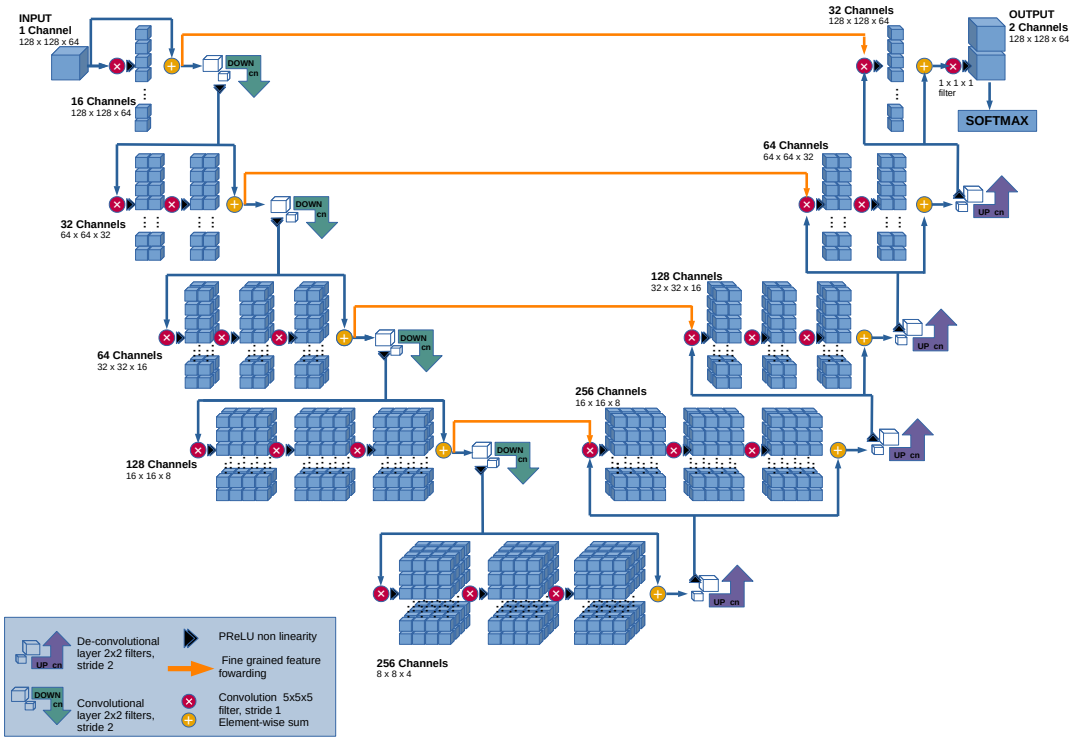


Figure: V-Net architecture (3D generalisation of U-Net)

Purpose of the study

$\rightarrow$ Development of an automated segmentation pipeline powered by AI

Materials and methods: dataset

Records from a cohort of 131 women with malignant breast lesions who underwent neoadjuvant chemotherapy

- imaging protocol: 3T DCE-MRI (THRIVE), fat saturated (SPAIR, TE=2 ms, TR=shortest), before and after contrast agent injection, with a temporal resolution of 90 s
- spacing: median [0.89, 0.89,0.95] mm, minimum [0.7,0.7,0.5] mm, maximum [1,1,1.1] mm
- manual segmentations provided by an experienced radiologist

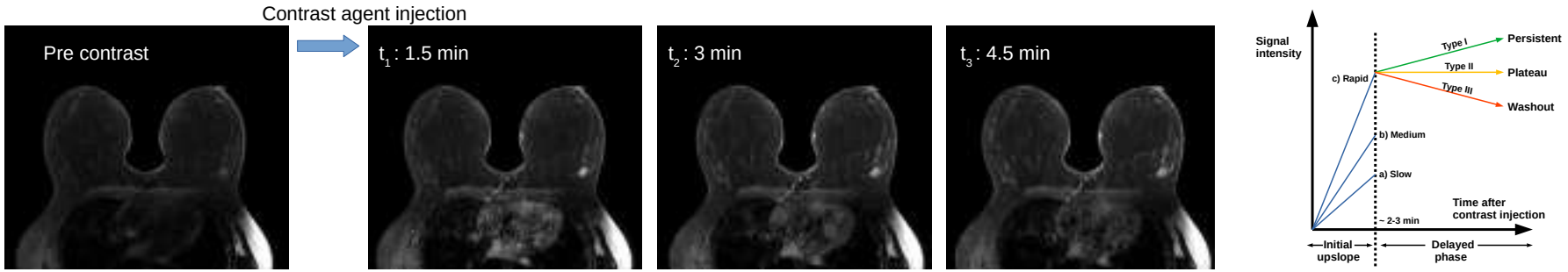


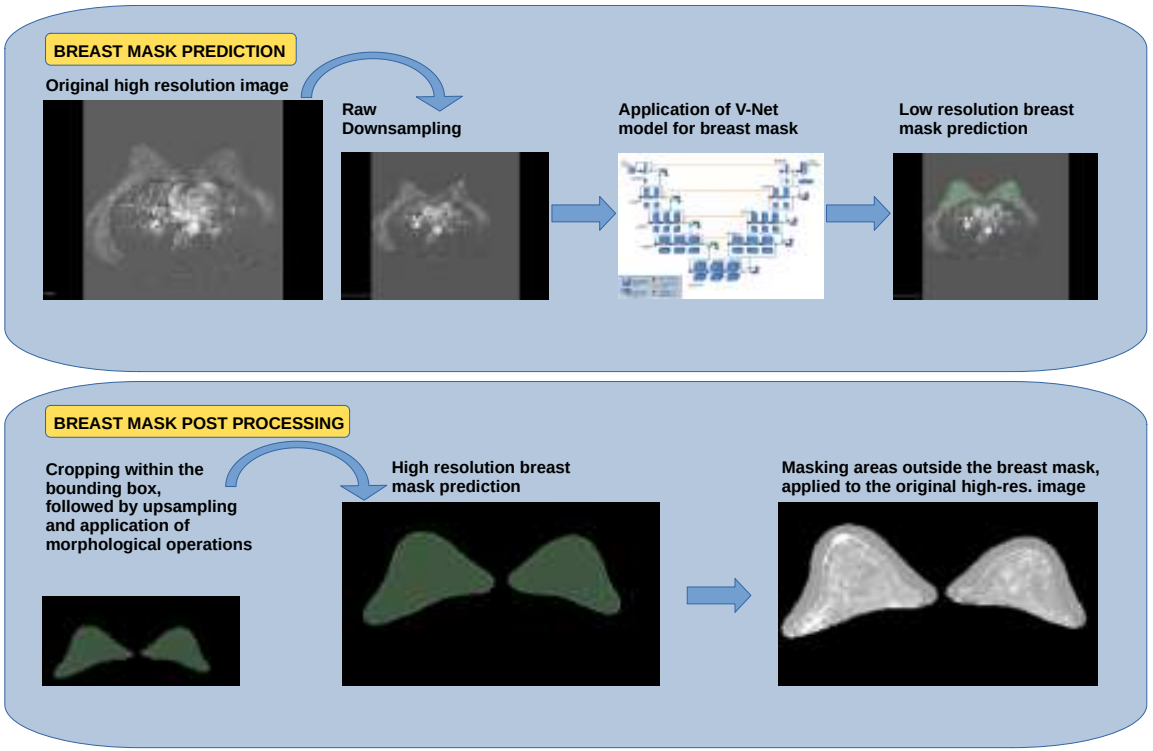
Figure: example of DCE-MRI slices and TIC showing contrast agent uptake.

Materials and methods: segmentation pipeline

The procedure employed V-Net and was divided into two phases:

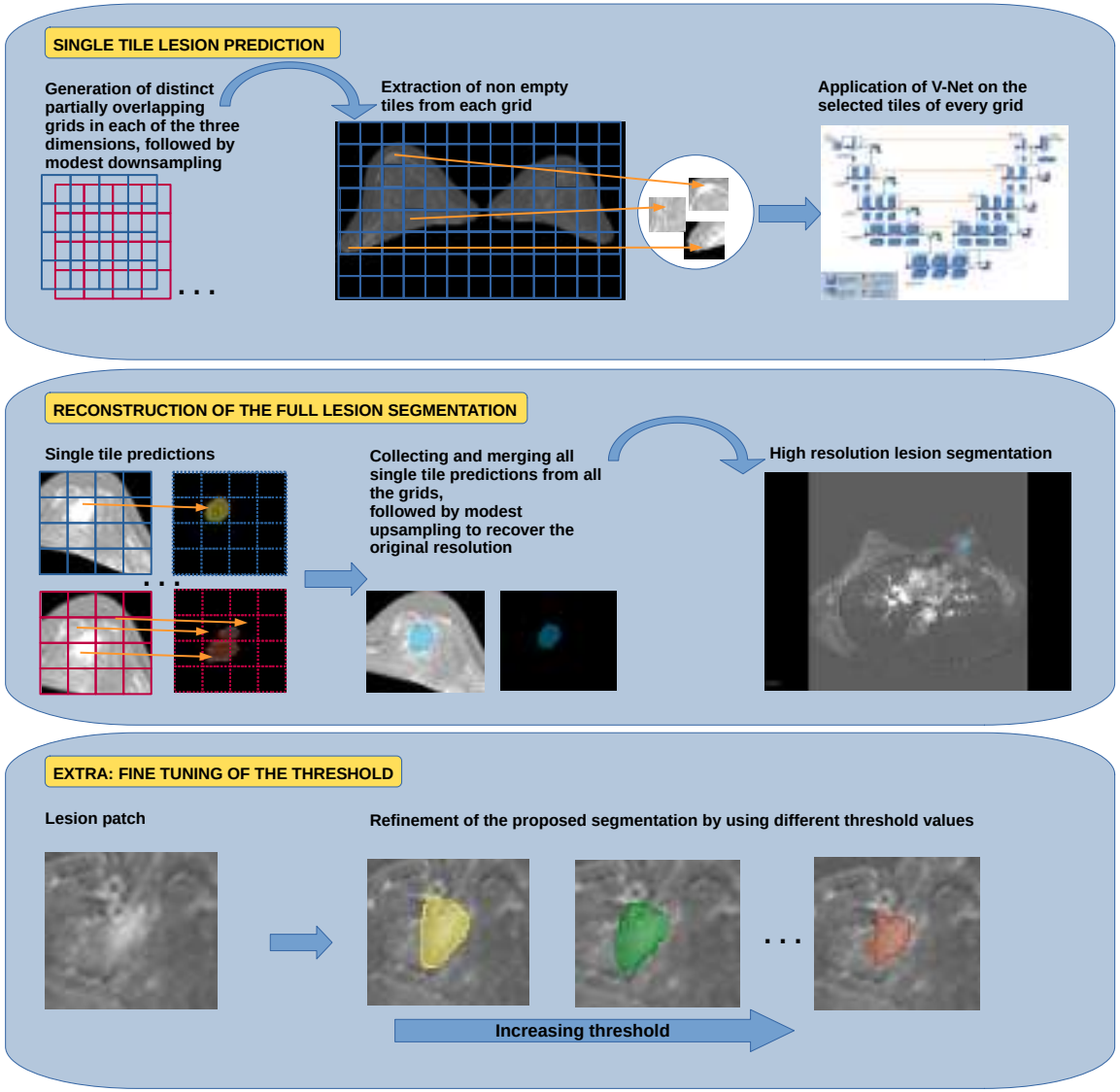
1. breast segmentation;
2. lesion segmentation.

Breast segmentation: reduces the impact of confounding enhanced organs, image sizes, and computational burden.



Materials and methods: segmentation pipeline

Lesion Segmentation: applies V-Net to individual high-resolution tiles to ensure high quality with consistent spacing in each direction.



Material and methods: training parameters

Various settings were explored during training (using PyTorch v2.0.1), including different background patch percentages and threshold tuning.

	Breast segmentation	**Lesion segmentation**			
		lesion + 30% bg.	lesion +50% bg.	lesion +75% bg.	lesion +100% bg.
input size	(128, 128, 64)	patches of (32, 32, 32) extracted from the full image of (416, 256, 192)			
image spacing	(3.3, 3.3, 3.1) mm	(1,1,1) mm			
loss function	BCE	BCE			
optimiser	SGD (momentum=0.9)	SGD (Nesterov momentum)			
learning rate	0.01	0.001	0.005	0.005	0.005
batch size	16	64	128	128	128
epochs	150	100	100	100	70
pre-processing	min-max normalisation, 0 padding to fill standard size	min-max normalisation, 0 padding for values outside the breast mask			
augmentation	no	use of patches			
post-processing	upsampling, binary dilation [3,3,3], Gaussian smoothing [5,5,5], eventual removal of small disconnected components ($< 10\%\ V_{max\ comp.}$)	no			
binarisation threshold	0.5	0.5	0.5	0.5	0.3-0.6

Table: Training parameters for the segmentation models

Segmentation evaluation

To assess interoperator variability and model performance, it's essential to define segmentation comparison methods.

Def. of segmentation

Considering the medical volume as a set of points $X = \{x_1, ..., x_n\}$, the segmentation performed by the operator $j = \{1, 2, ...\}$ can be represented as the partition $S_j = \{S_j^L, S_j^B\}$ of X associated with the membership function $f_j^i(x) = \begin{cases} 1 & \text{if } x \in S_j^i \\ 0 & \text{if } x \notin S_j^i \end{cases}$ for $i = L, B$,

where S^L represents the identified lesion and S^B is the background.

Similarity measures to compare segmentations

Dice Similarity Coefficient (overlap based)
Relative Size Difference (volume based)

$$DSC = 2 \cdot \frac{|S_1^L \cap S_2^L|}{|S_1^L| + |S_2^L|} \quad , \quad RSD = \frac{|S_2^L|}{|S_1^L|} - 1 \tag{1}$$

Results: breast segmentation

Breast segmentation is a straightforward task, showing promising results even with a small dataset. The overlap is substantial, and larger borders are preferred to avoid losing potential information in the breast area.

DSC	0.835 $\pm$0.005
RSD	0.22 $\pm$ 0.03
accuracy	0.962 $\pm$ 0.01
specificity	0.965 $\pm$ 0.02
precision	0.77 $\pm$ 0.01
sensitivity	0.927 $\pm$ 0.009

Table: Mean evaluation metrics summary over 5 folds.

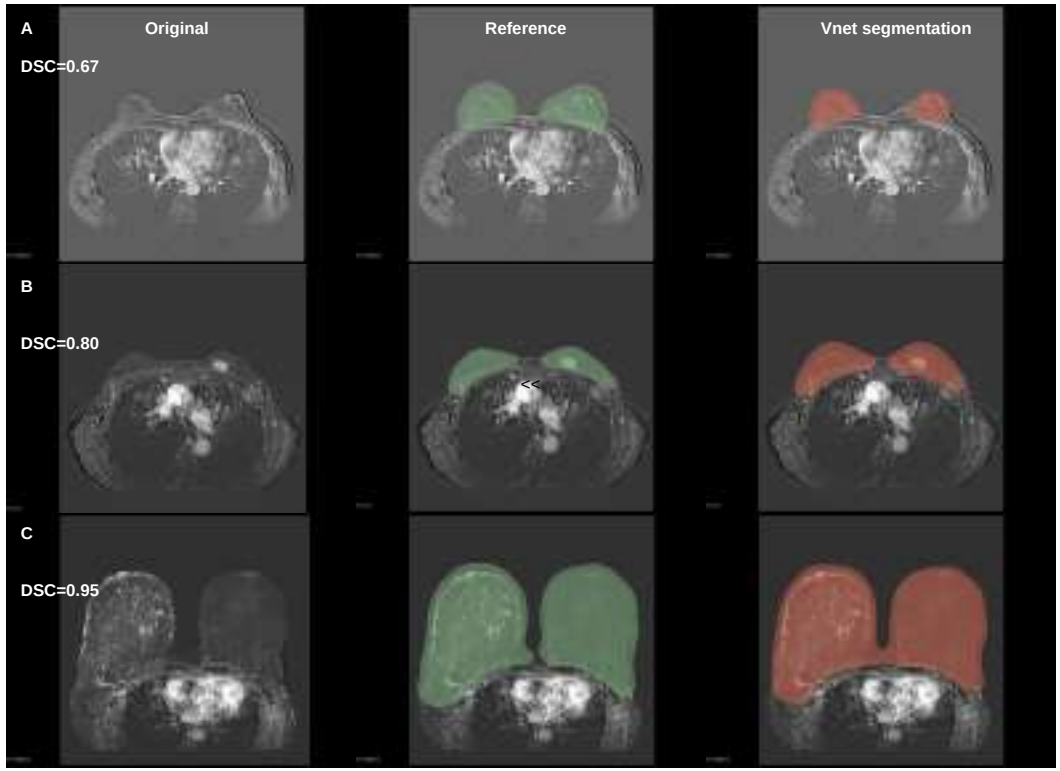


Figure: Examples of the obtained breast masks, axial view

Results: lesion segmentation

Among the various settings explored, preference was given to using all image data (100% bg) and applying an intermediate threshold (0.5) to balance true and false positives.

	median	IQR
DSC	0.71	0.52-0.82
RSD	0.003	-0.276-0.300
accuracy	0.9998	0.9998-0.9999
specificity	0.9999	0.9999-1.0000
precision	0.74	0.64-0.85
sensitivity	0.81	0.53-0.92

Table: Performances of the selected model

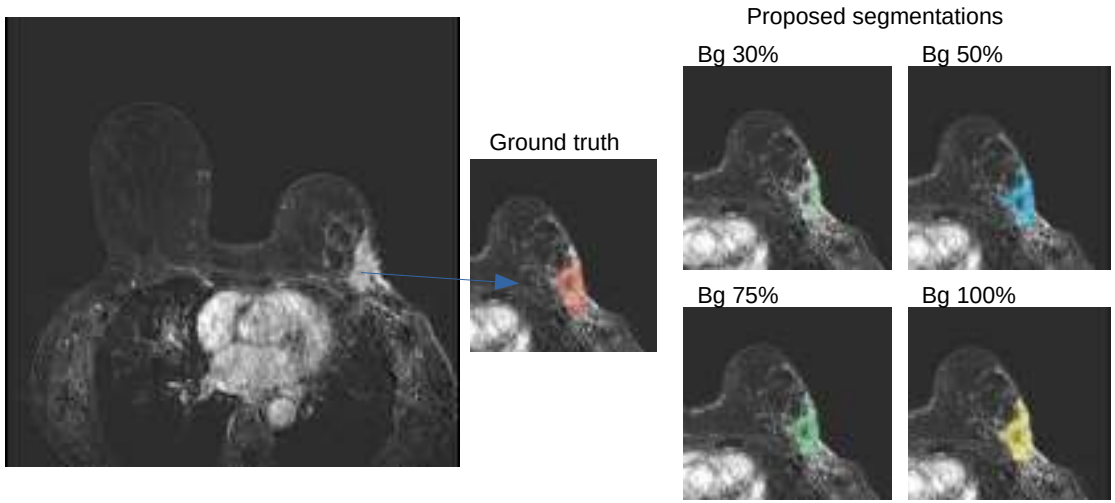


Figure: Example of proposed segmentation by the models obtained using different percentages of background patches.

Results: interoperator variability

A subset of 72 cases was segmented by 2 different radiologists:

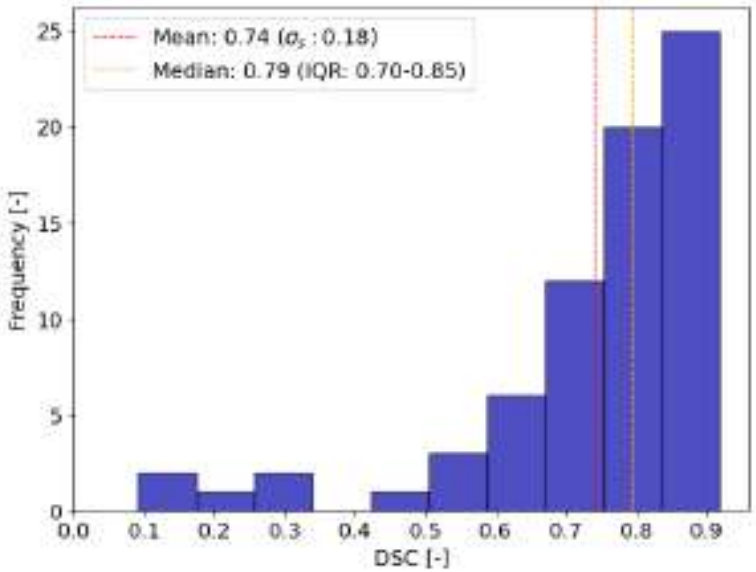


Figure: overlap-based comparison.

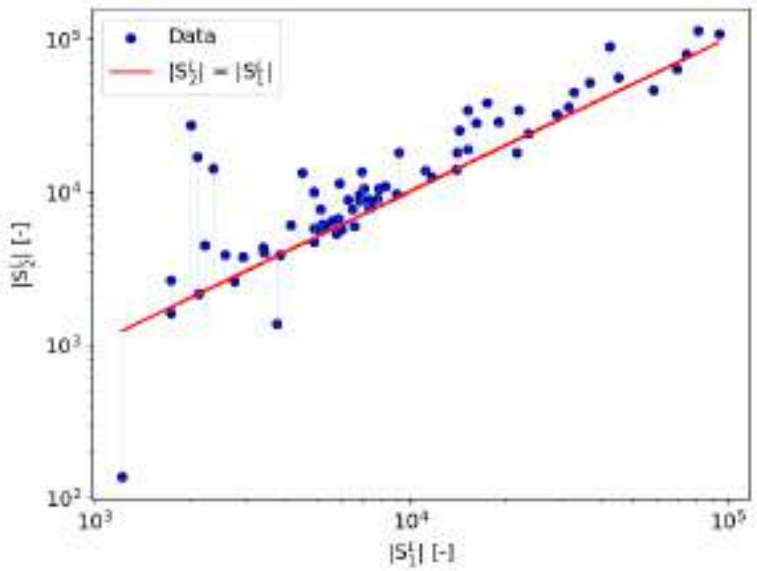


Figure: size comparison.

metric	median	IQR	min	max
DSC	0.79	[0.70-0.85]	0.09	0.92
RSD	0.2	[0.1-0.5]	-0.9	12.5

Table: comparison between two operators

Limitations and considerations

Main limitations of the study

- Small dataset size
- Lack of external validation

In some cases, the agreement was suboptimal due to:

- false positives detected at lymph nodes in the axillary area;
- detection of lesions missed by radiologists.

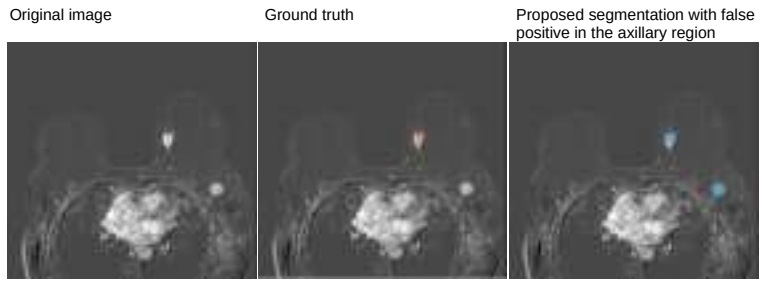


Figure: Example of segmentation with false positive findings in the axillary area.

Figure: Example of a proposed segmentation discovering fragments missed by the radiologist

Conclusions

Main results

In summary, despite the main challenge of reduced sample size, the pipeline demonstrates promising results, including:

- reliable segmentation performance
- enhanced segmentation speed (>5 times faster)

This suggests the potential use as a segmentation baseline for future radiomic studies, ensuring reproducibility.

Further perspectives involve:

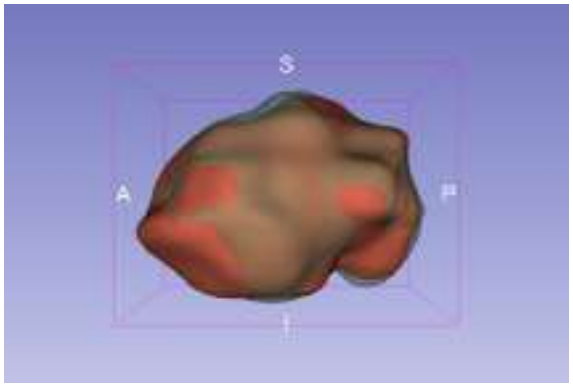
- validation with external data;
- model improvements (changes in architecture to capture also global information, data augumentation, etc.);
- and more...

Thank you for your attention!

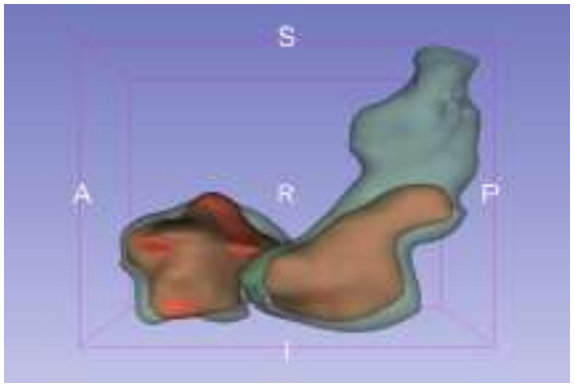
Backup slides

Some additional content...

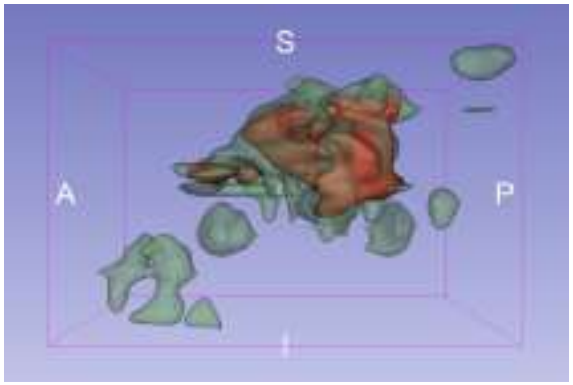
Examples of manual segmentation



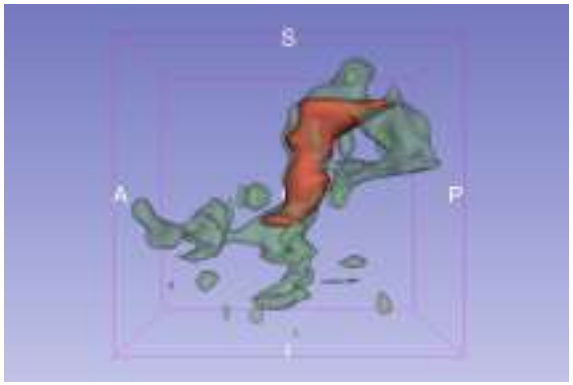
(a) optimal agreement
($DSC = 0.90$, $RSD = -0.07$)



(b) modest agreement
($DSC = 0.66$, $RSD = -0.91$)



(c) poor agreement
($DSC = 0.58$, $RSD = -1.3$)



(d) very poor agreement
($DSC = 0.33$, $RSD = -1.9$)

Figure: Examples of segmentation differences

V-Net implementation

The model is based on the V-Net architecture, with a few targeted modifications.

- non linearity: PReLU
- loss function: Binary Cross Entropy
- output: single channel

Layer	operation	channels
L-1	1 x convolution	16
L-2	2 x convolution	32
L-3	3 x convolution	64
L-4	3 x convolution	128
L-5	3 x convolution	256
R-4	3 x convolution	256
R-3	3 x convolution	128
R-2	2 x convolution	64
R-1	1 x convolution	32

Table: Architecture parameters

layer type	**operation**	**kernel**	**stride**	**padding**
normal layer	convolution	5x5x5	2x2x2	2
downsampling	convolution	2x2x2	2x2x2	2
upsampling	transposed convolution	2x2x2	2x2x2	2

Table: Layer parameters

Results: lesion segmentation

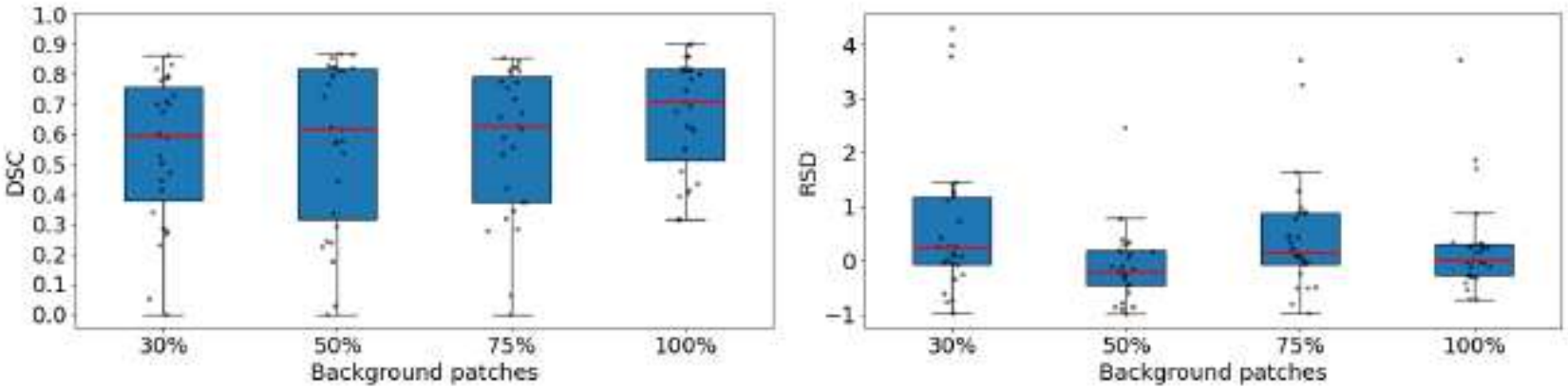


Figure: Results varying percentage of background patches.

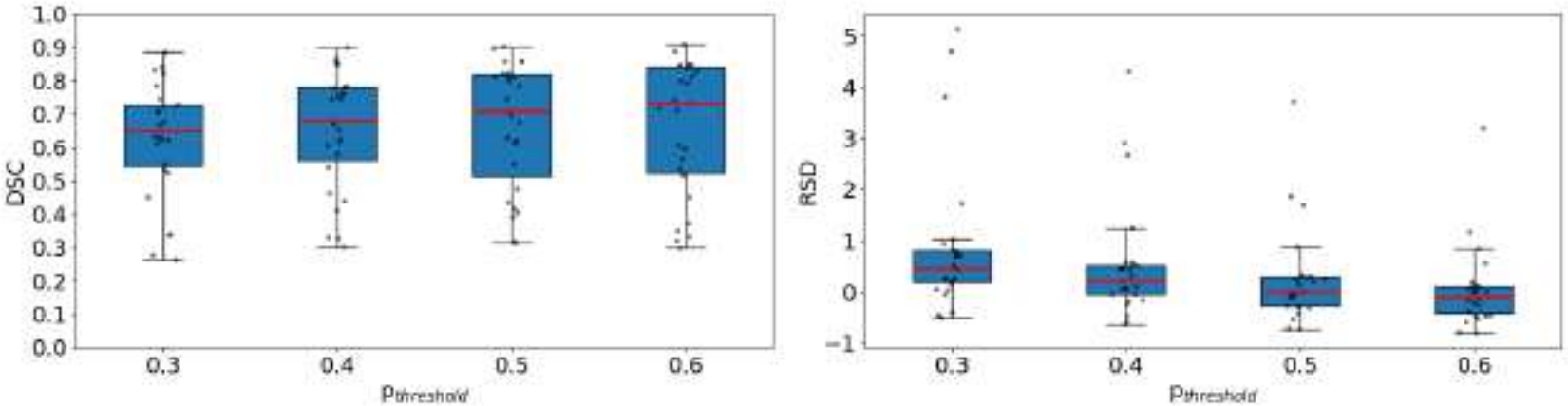


Figure: Results varying binarisation threshold for the case with 100% bg. patches.