3D Deep Learning for Anatomical Structure Segmentation in Multiple Imaging Modalities

Barbara Villarini, Hykoush Asaturyan, Sila Kurugol, Onur Afacan, Jimmy D. Bell, E. Louise Thomas

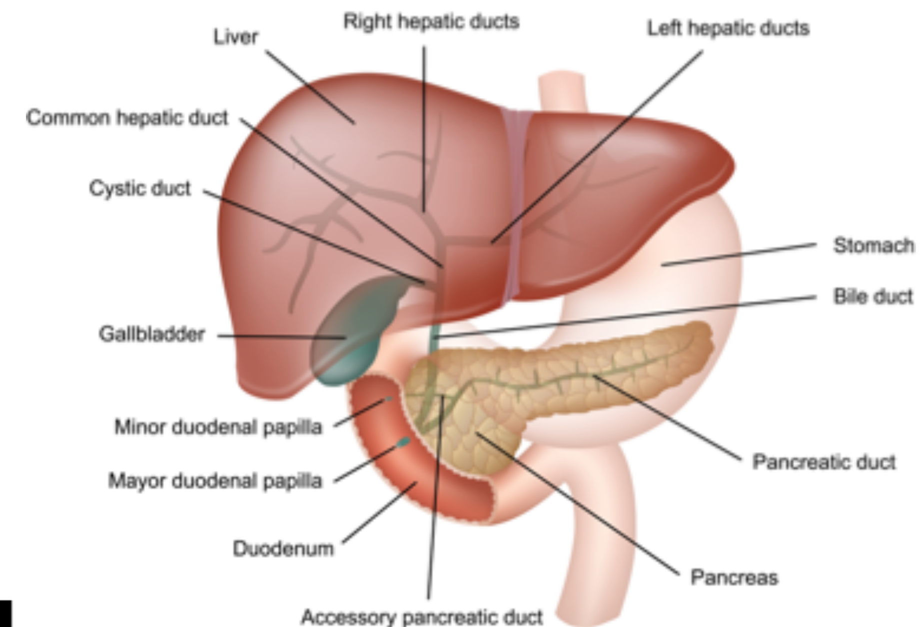
Outline

- **Introduction to problem, motivation and challenges**
- **Contributions**
- **Methodology**
  - Training Stage
    - 3D Detection and localisation
    - 3D Segmentation
  - Testing Stage
- **Evaluation and results**
- **Conclusion and future works**

Problem & Motivation

- Medical imaging using non-invasive techniques has rapidly evolved, providing detailed images of anatomy in the human body.
- Rising high numbers of patients undergo MRI or CT scans, with millions of abdominal images acquired in the UK and the EU.
- Computational methods that analyse a large medical image dataset and automatically extract information provides unique opportunities to answer fundamental clinical and scientific questions.
- Begin the process of subject stratification according to organ morphology.
- Improve the analysis and detection of disease and treatment planning performed by medical care service, including radiologists and clinicians.

Challenges

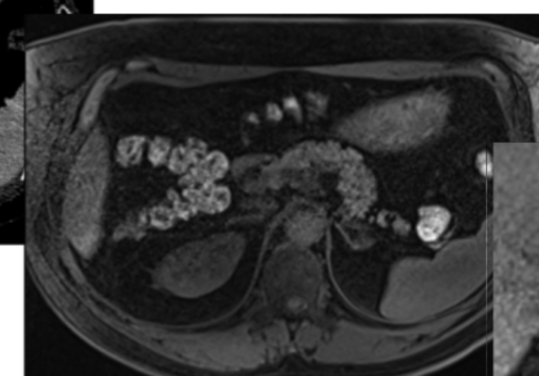


- **Organs with high structural variability**: different size, structure and location.

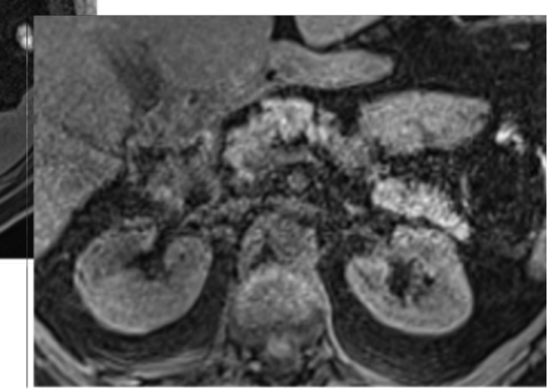
- **Differences in imaging modalities**, such as CT and MRI scans.

- **Difference in image quality** and presence of artefacts.

CT scan



MRI (a)



MRI (b)

- **Dataset limitations**: very few datasets are publicly available, especially MRI of abdominal organs. It is extremely difficult to perform direct comparison with other methodologies.

Contributions

- Novel and robust automated 3D deep learning approach for automatic quantitative organ and muscle segmentation in medical image volumes of multiple modalities.

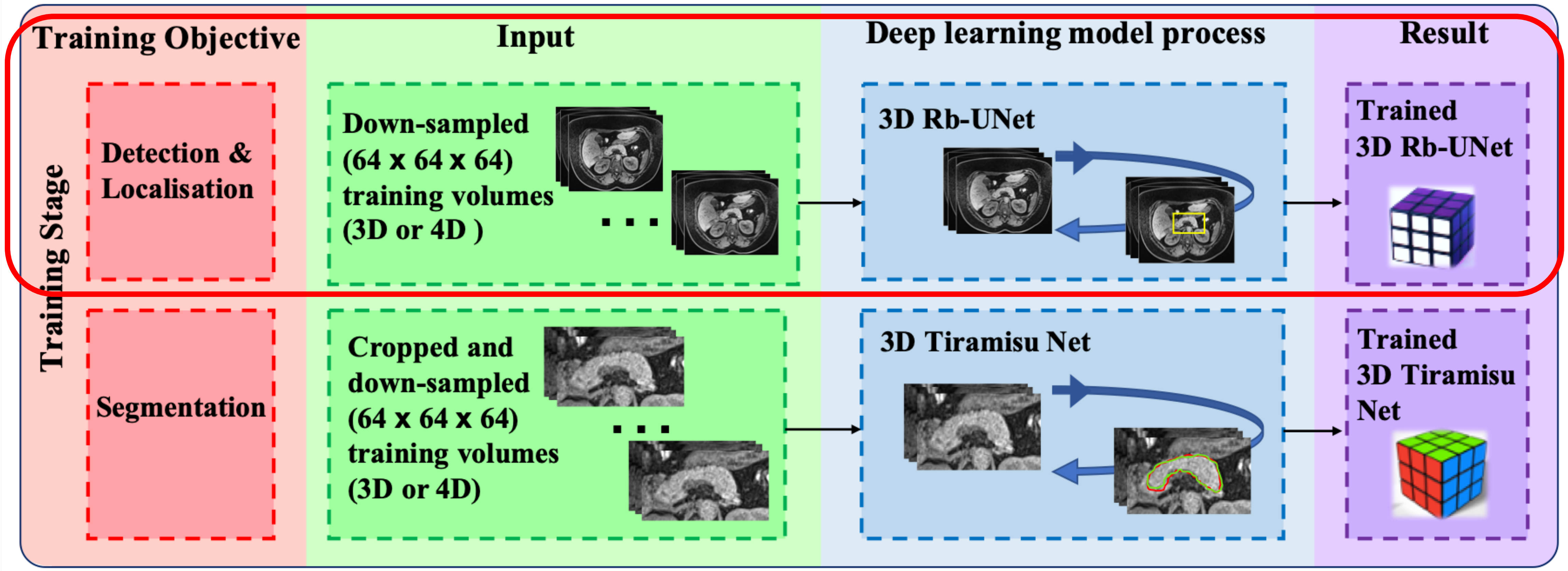
- Employ volumetric information instead of 2D feature learning and is modular, scalable and generalisable.

- Evaluation on six different datasets of MRI, DCE-MRI and CT modality, targeting distinct abdominal structures including the pancreas, liver, kidneys and iliopsosas muscles.

- Quantitative results outperform or are comparable with the state-of-the-art, demonstrating high statistical stability.

Training Stage

Overview of Methodology



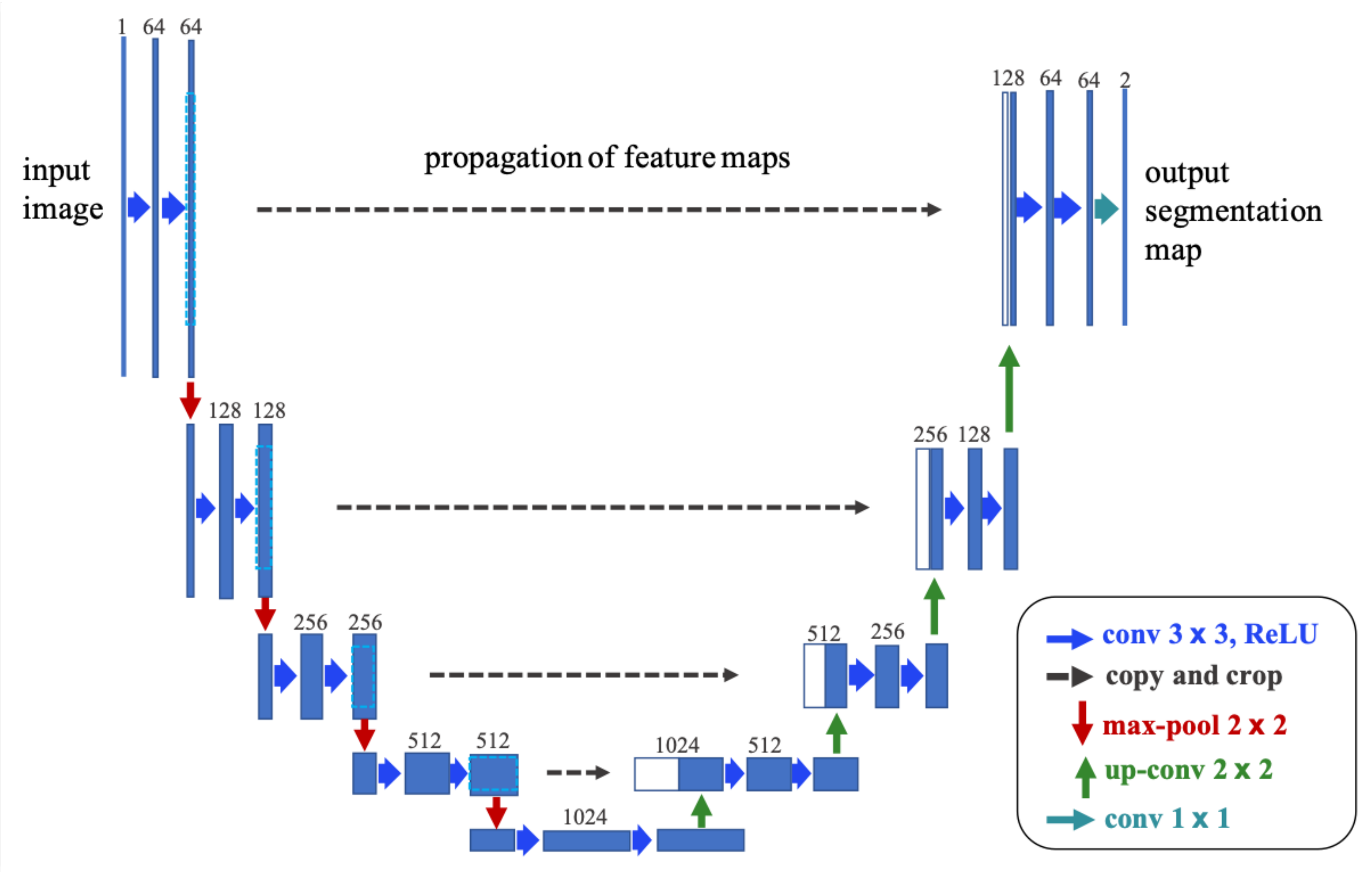
The training stage simultaneously develops:

- **Detection and Localisation**: 3D Rb-UNet localises target organ.
- **Segmentation**: segmentation network, 3D Tiramisu, predicts labels that correspond to “organ” and “non-organ” tissue.

Training Stage

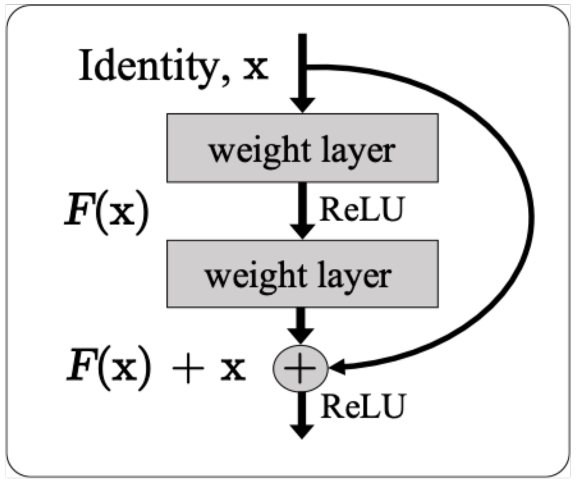
Detection and Localisation: Rb-UNet

U-Net



Three sections in U-Net:

- Downsampling (encoder)
- Bottleneck
- Upsampling (decoder)

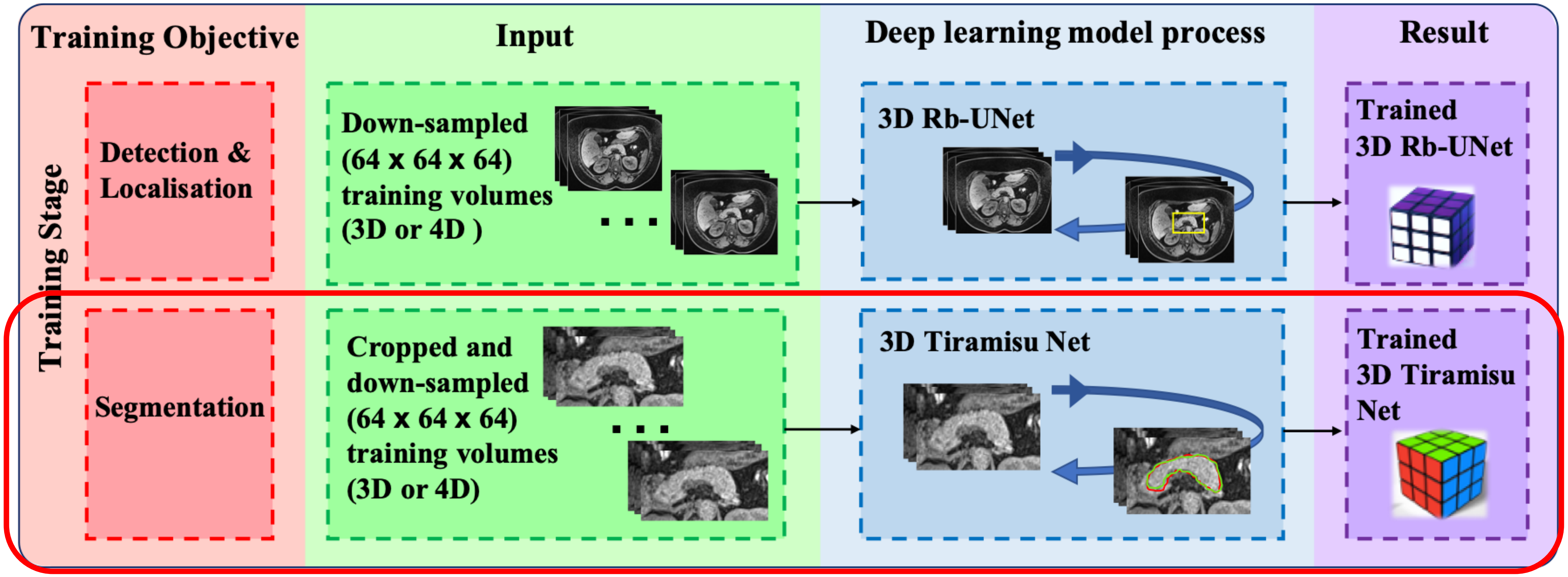


Residual Block (Rb) Learning

- Advantage of alleviating vanishing gradient problem.
- A residual block connects the input of a convolutional layer in the U-Net architecture at each scale to the output of the corresponding layer.

Training Stage

Overview of Methodology



The training stage simultaneously develops:

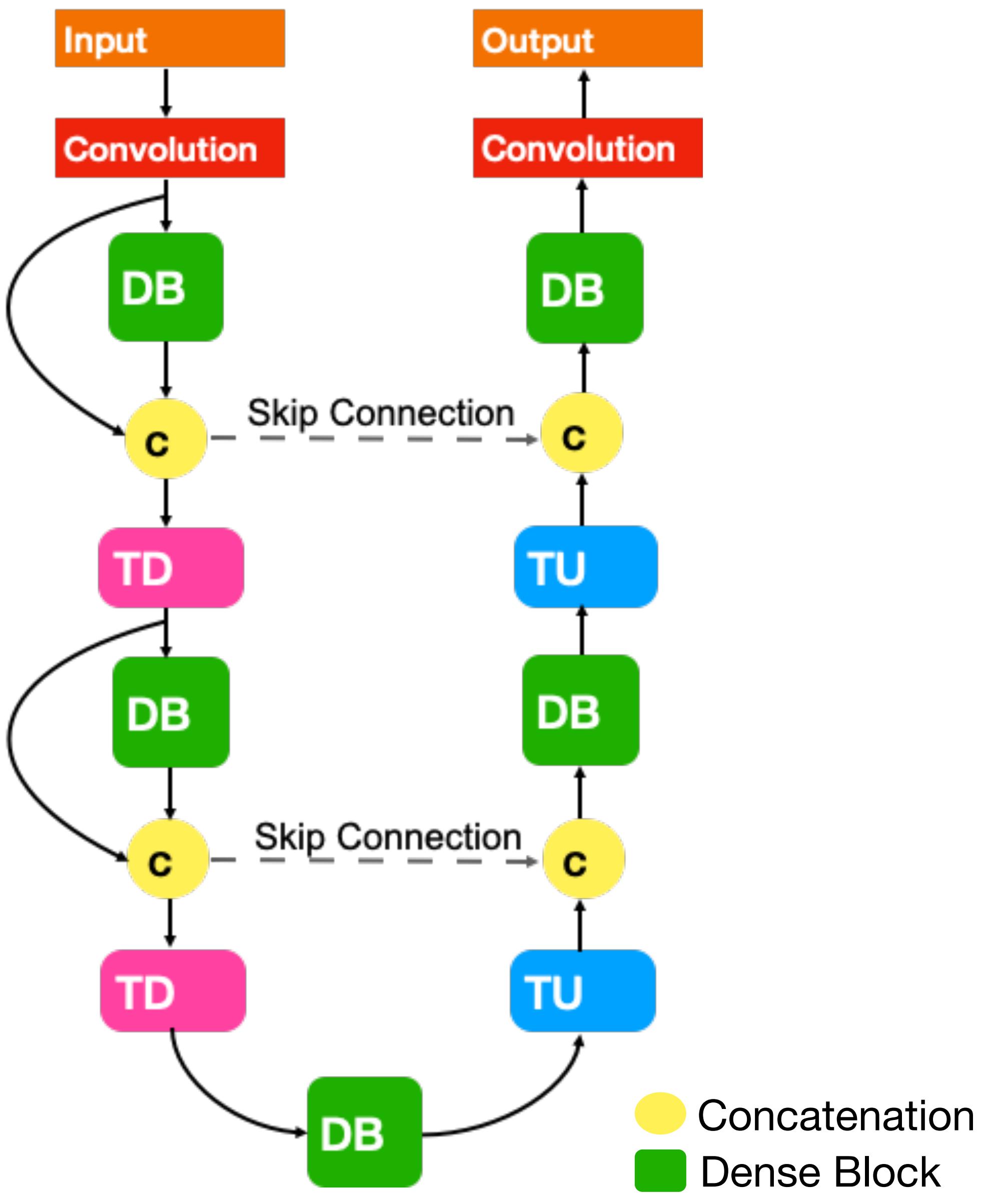
- **Detection and Localisation**: 3D Rb-UNet localises target organ.
- **Segmentation**: segmentation network, 3D Tiramisu, predicts labels that correspond to “organ” and “non-organ” tissue.

Training Stage

Segmentation

Fully Convolutional Dense-Net (Tiramisu)

➢ Extend DenseNet architecture to fully convolutional networks (FCNs) to mitigate excessive feature maps.



Layer
Batch Normalization
ReLu
3 x 3 Convolution
Dropout = 0.2

Transition Down (TD)
Batch Normalization
ReLu
1 x 1 Convolution
Dropout = 0.2
2 x 2 Max Pooling

Transition Up (TD)
3 x 3 Transposed convolution Stride = 2

Architecture
Input m = 3
DB (4 Layers) + TD, m = 112
DB (5 Layers) + TD, m = 192
DB (7 Layers) + TD, m = 304
DB (10 Layers) + TD, m = 464
DB (12 Layers) + TD, m = 656
DB (15 Layers) , m = 896
TU + DB (12 Layers), m = 1088
TU + DB (10 Layers), m = 816
TU + DB (7 Layers), m = 578
TU + DB (5 Layers), m = 384
TU + DB (4 Layers), m = 256
1 x 1 Convolution, m = 2 (or 3)
Softmax

$$W_{cross-entropy} = -\frac{1}{n}\sum_{i=1_t}^{N} w_i^c\,[\hat{p}_i \log p_i + (1-\hat{p}_i)\log(1-p_i)]$$

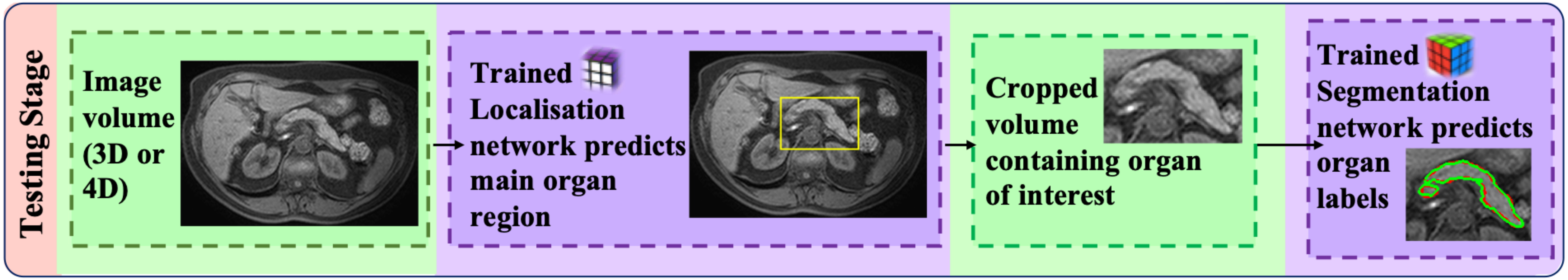
Testing Stage

- Fully trained 3D Rb-UNet performs a coarse segmentation of the target organ in an (unseen) image volume.

  - ➢ A minimum organ bounding box is generated.

- Cropped image volume containing the main organ region processes through the fully trained 3D Tiramisu model.

  - ➢ 3D Tiramisu model performs finer and more detailed voxel-wise predictions of target "organ" or "non-organ".



Evaluation

Datasets

- **216** T2-weighted abdominal **MRI** scans using a Philips Intera 1.5T scanner - **Pancreas** annotated.
- **132** T2-weighted abdominal **MRI** scans using Siemens Trio 3T scanner - **Pancreas** annotated.
- **82** abdominal contrast-enhanced **CT** 3D scans using Philips and Siemens MDCT scanners - **Pancreas** annotated.
- **30** T2-weighted abdominal **MRI** scans using Siemens Trio 3T scanner - **Liver** annotated.
- **30** T2-weighted abdominal **MRI** scans using Siemens Trio 3T scanner - **Iliopsoas muscle** annotated.
- **60** 4D **DCE-MRI** scans acquired at 3T for six minutes after injecting Gadavist (gadobutrol) - **Kidney** annotated.

Evaluation

Quantitative Metrics

- **Jaccard Index (J)**: measures the overlap between the segmentation outcome and desired outcome often referred as the size of the intersection between two sets (i.e. the segmentation (S) and ground-truth (G)) divided by the size of the union between these two sets.

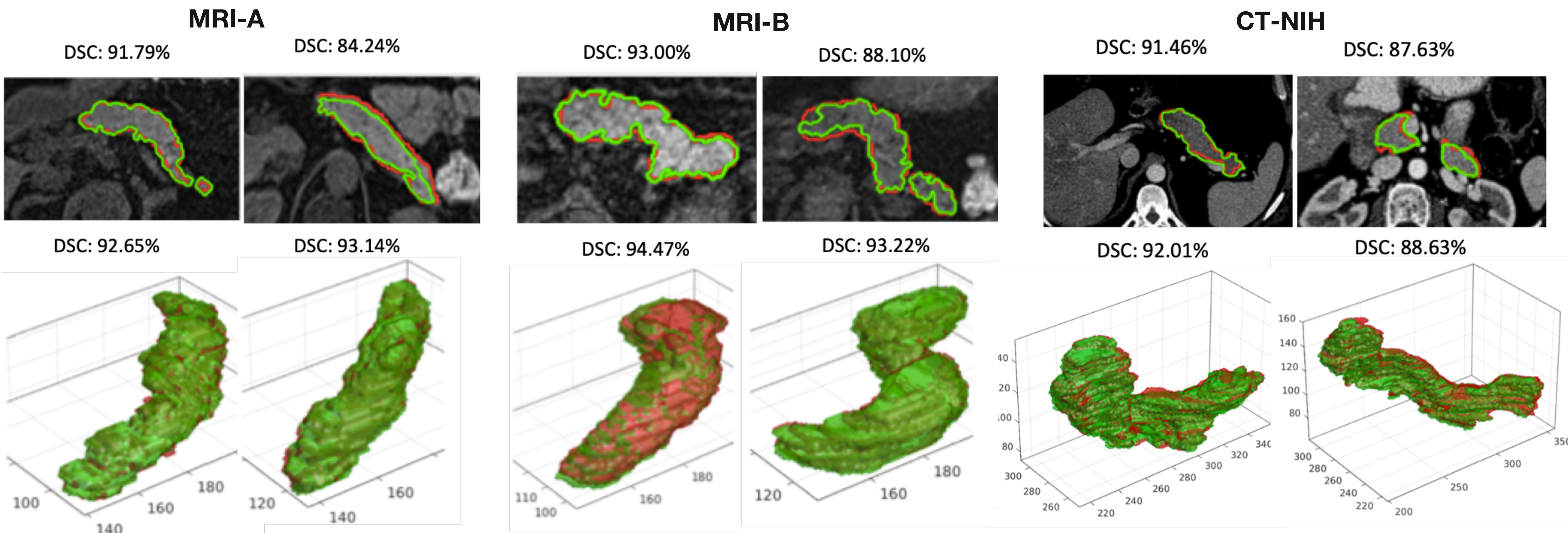
$$JI = \frac{|G \cap S|}{|G| \cap |S|}$$

- **Dice Similarity Coefficient (DSC)**: describes twice the number of elements common to both sets (i.e. segmentation (S) and ground-truth (G)) divided by the sum of the number of elements in each of these sets.

$$DSC = \frac{2\,|G \cap S|}{|G| + |S|}$$

Evaluation and Results

Pancreas Segmentation



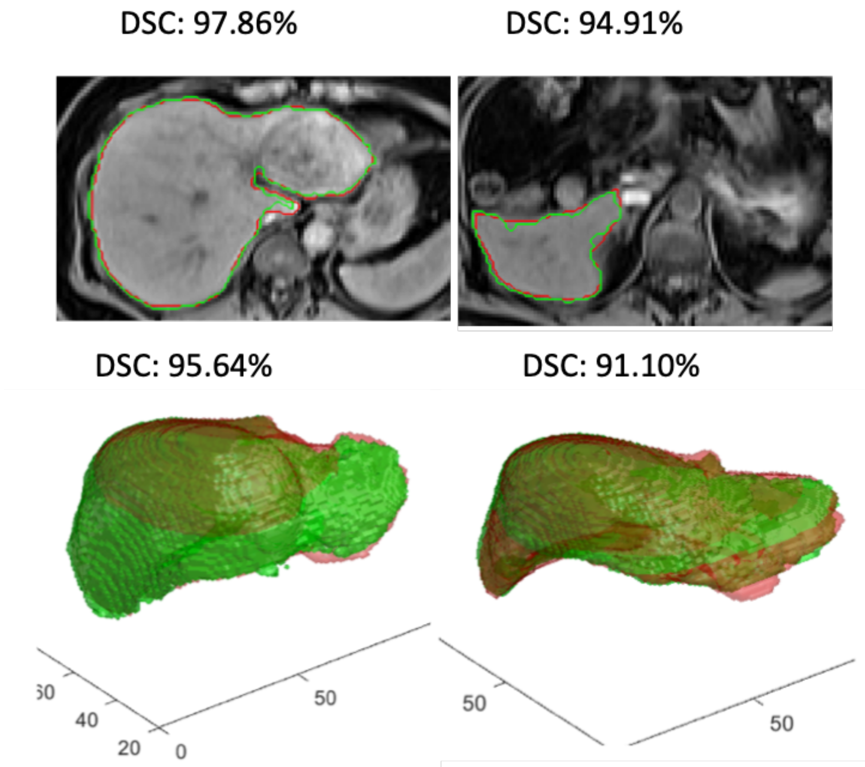
	MRI-A		MRI-B	
Method	**DSC(%)**	**JI (%)**	**DSC (%)**	**JI (%)**
2D UNet [19]	69.1±10.2	53.8±14.2	72.8±7.5	67.9±10.2
2D FCN [13]	70.2±8.5	63.5±13.5	70.9±7.7	65.4±13.5
Deeporgan [24]	44.5±25.2	32.7±29.4	50.1±22.7	44.9±12.0
Multiorgan [14]	52.6±17.1	44.1±20.7	55.8±18.6	49.9±18.7
Casc. 3D FCN [25]	65.2±10.1	52.2±15.3	69.6±11.5	61.2±15.9
Geo. descript. [26]	79.6±5.7	66.5±7.9	81.6±5.1	69.2±7.1
Hausdorff-Sine [27]	84.1±4.6	72.9±6.5	85.7±2.3	75.1±3.5
Proposed	**89.9±3.4**	**81.9± 5.6**	**90.2±5.1**	**82.6±7.8**

	CT-NIH	
Method	**DSC (%)**	**JI (%)**
2D UNet [19]	79.7±7.6	66.3±4.0
2D FCN [13]	80.3±9.0	67.1±4.7
Recurrent NN [16]	82.4±6.7	70.6±9.0
Holistically-nested CNN [12]	81.3±6.2	68.5±3.2
Cascade 3D FCN [25]	76.8±9.4	62.3±4.9
Geo. Descriptors [26]	79.3±4.4	66.1±6.2
Hausdorff-Sine [27]	83.1±5.3	71.4±7.4
Proposed	**84.7±7.9**	**74.2±11.4**

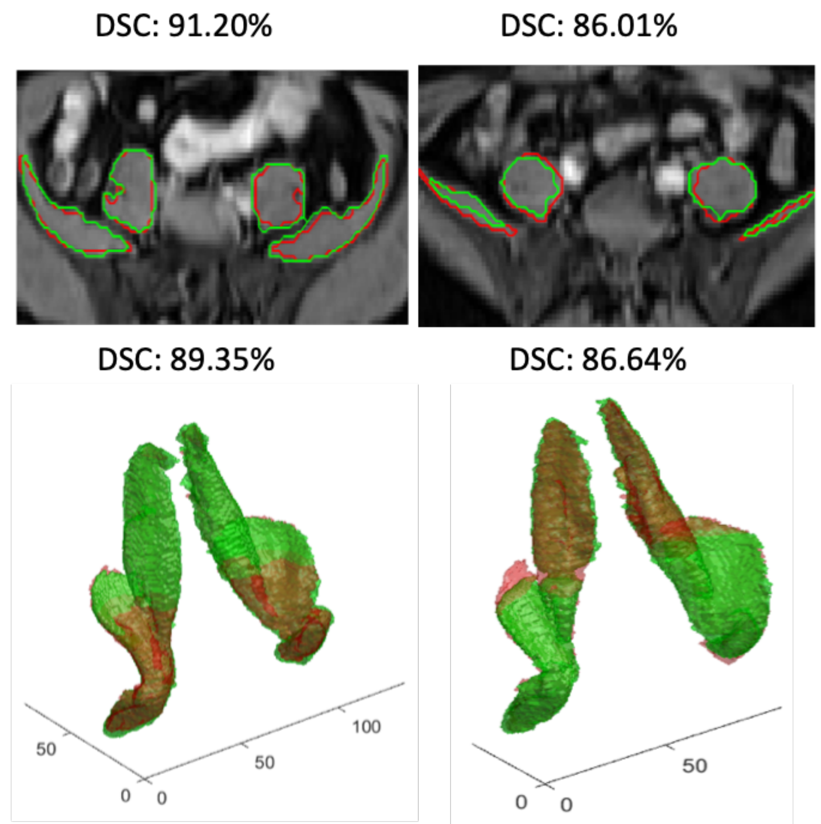
Evaluation and Results

Additional Organ & Muscle Segmentation

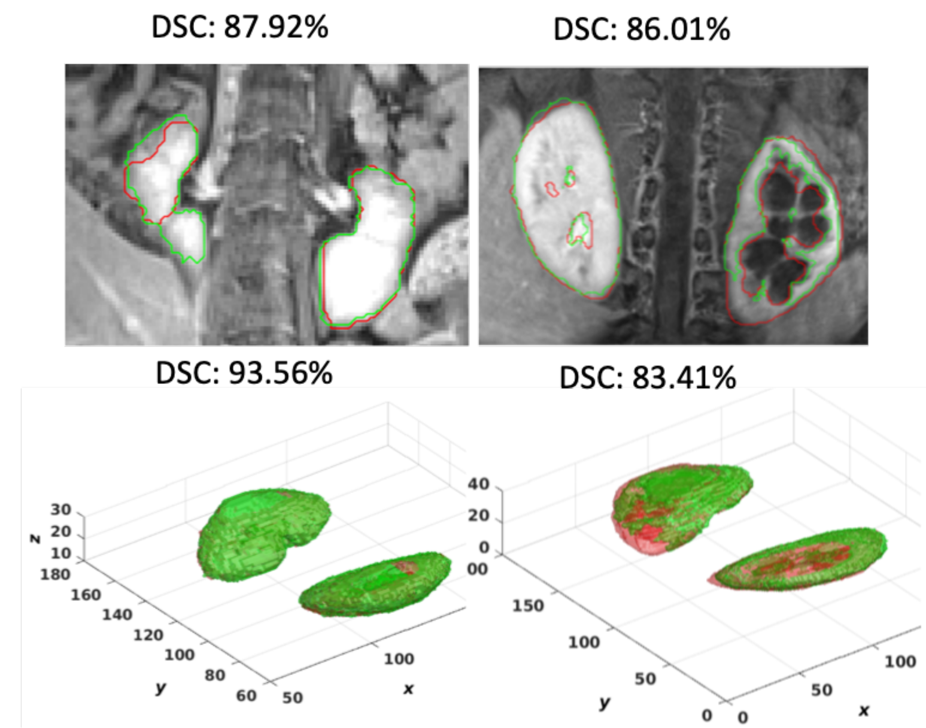
Liver MRI



Iliopsoas muscles MRI



Kidneys DCE-MRI



Anatomical Structure	Clinical Assessment	Modality	Data dimension	Dataset Size	DSC (%)
Pancreas	Early type 2 diabetes	MRI	3D	216	89.9 ± 3.4
Pancreas	Early type 2 diabetes	MRI	3D	132	90.2 ± 5.1
Pancreas	Normal	CT	3D	82	84.7 ± 7.9
Liver	Normal	MRI	3D	30	95.64 ± 1.31
Iliopsoas muscles	Normal	MRI	3D	30	88.41 ± 2.39
Kidneys	Normal	DCE-MRI	4D	30	90.48 ± 1.56
Kidneys	Hydronephrosis	DCE-MRI	4D	30	86.44 ± 3.84

Conclusion and Future Works

- Serious challenges towards developing robust segmentation methods include high organ size variations, data from different scanner modalities, protocols, and image resolution.

- We propose a 3D deep learning approach by exploiting volumetric contextual information to perform localisation and fine segmentation of the target anatomical structure.

- We achieve robust segmentation performance using CT, MRI and DCE-MRI with higher statistical stability than state-of-the-art approaches.

- The proposed approach can classify clinical measures and indicate the progression or severity of a medical condition.

- This framework can incorporate into the development of a practical medical image analysis cloud-based application, accessible to clinicians, scientists and medical care services aiming to improve detection and diagnosis of diseases.

Thank you for listening

We would like to thank **Julie Fitzpatrick** (University of Westminster) and **Karen Sarao** (Boston Children's Hospital & Harvard Medical School) for providing expert-led manual annotations of ground-truth data.

This work was supported by the **Royal Academy of Engineering under the Leverhulme Trust Research Fellowship scheme** (LTRF1920\16\26), and it was supported partially by **the Boston Children's Hospital Translational Research Program Pilot Grant 2018**, **Society of Paediatric Radiology Multi- center Research Grant 2019**, **Crohn's and Colitis Foundation of America's** (CCFA) **Career Development Award and AGA-Boston Scientific Technology and Innovation Award 2018** and by **NIDDK of the National Institutes of Health** under award number 1R21DK123569-0

UNIVERSITY OF WESTMINSTER

HARVARD MEDICAL SCHOOL

Boston Children's Hospital

CBMS Computer-Based Medical Systems 2021