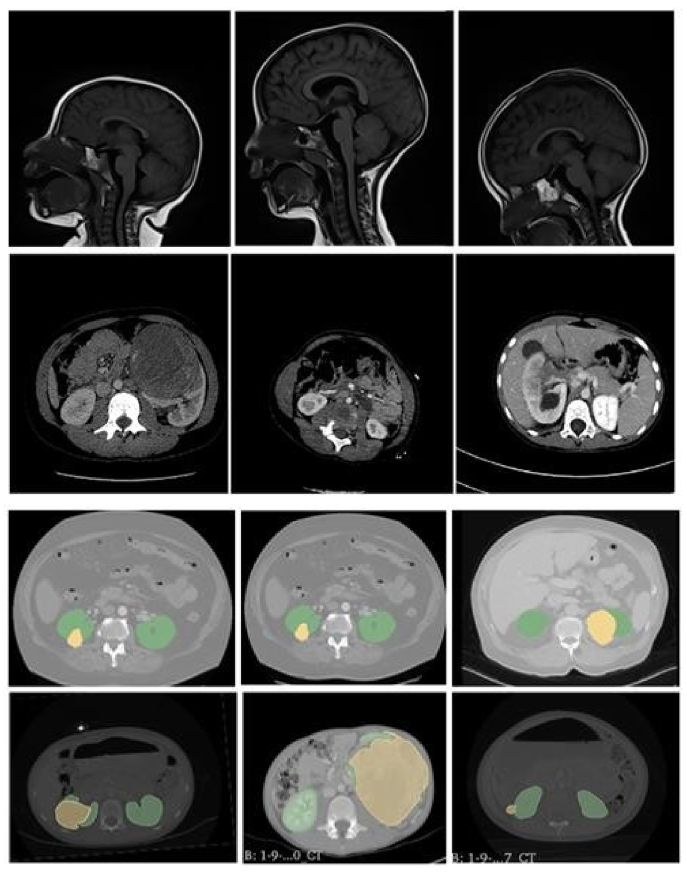
DeePRAC PROJECT
(Deep learning for segmentation of CT images to improve surgery in Pediatric RenAl Cancers)

Giammarco La Barbera[1], Pietro Gori[1], Isabelle Bloch[1,5]
Alessandro Delmonte[2], Sabine Sarnacki[2,3], Haithem Boussaid[4], Laurence Rouet[4]

1. LTCI, Télécom Paris, Institut Polytechnique de Paris, France
2. IMAG2, Imagine Institute, Université de Paris, France
3. Université de Paris, Pediatric Surgery Department, Necker Enfants-Malades Hospital, APHP, France
4. Philips Research Paris, Suresnes, France
5. Sorbonne Université, CNRS, LIP6, Paris, France



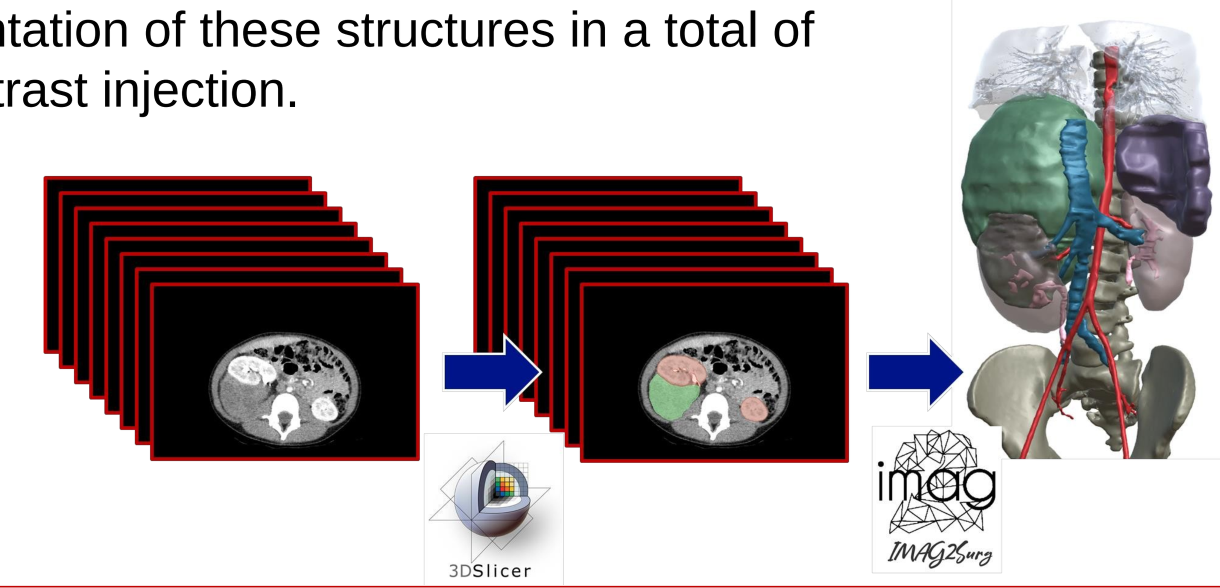
INTRODUCTION

The aim of the project is the automatic construction of an **individual 3D model** for each pediatric patient to help in surgery.
We have created a "time-line" (in importance) for structures to be segmented, starting from **kidneys and renal tumors.**

The first step was therefore the creation of the database and the manual segmentation of these structures in a total of **80 pediatric** pre-operative abdominal-visceral **CT images** with early arterial contrast injection.

Challenges in pediatric data-sets

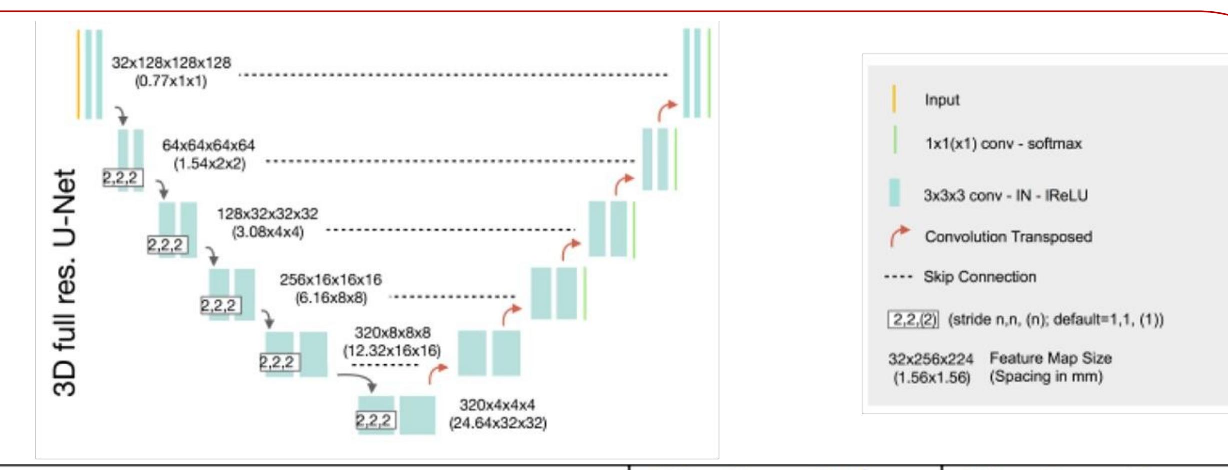
- Anatomical structures **highly heterogeneous** in terms of size.
- **High variability** in terms of pose and movements artifacts.
- **Limited in number** of images and hardly available in open access.
- **Differences between adults and children data-sets:** in terms of relative size between organs, variability among subjects and development of tumors.



STATE-OF-THE-ART

While the **literature is poor** on our specific problem **for children**, there are interesting works on adults, particularly around the MICCAI KiTS19 challenge [1] in which No-newUNet [2] was the winner.

We have been able to reproduce this network from scratch and to achieve the same results on the adult database (using 210 images). From these we attempted **transfer learning technique** on our pediatric data-set, using also ad-hoc and time-consuming data augmentation, like we did previously in [3].



Technique	Dice Score Kidney	Dice Score Tumor
Direct Inference (weights frozen)	20.83 (35.55)	18.29 (35.73)
Fine-Tuning (first 2 blocks encoder and last 2 decoder)	53.38 (25.84)	51.05 (31.76)
Fine-Tuning (entire decoder)	81.75 (7.18)	75.79 (23.24)
Fine-Tuning (entire network)	84.99 (6.38)	81.08 (23.01)
Training (entire network)	**91.79 (2.81)**	**91.09 (6.95)**

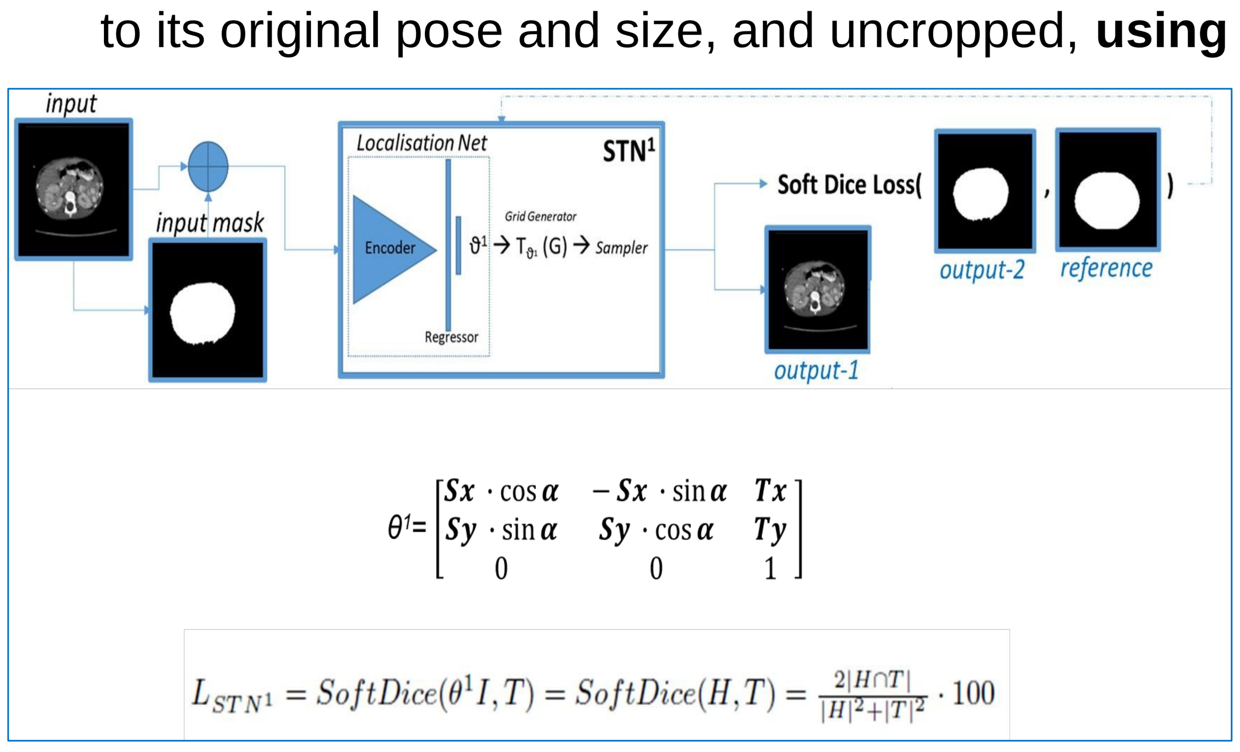
Problems encountered

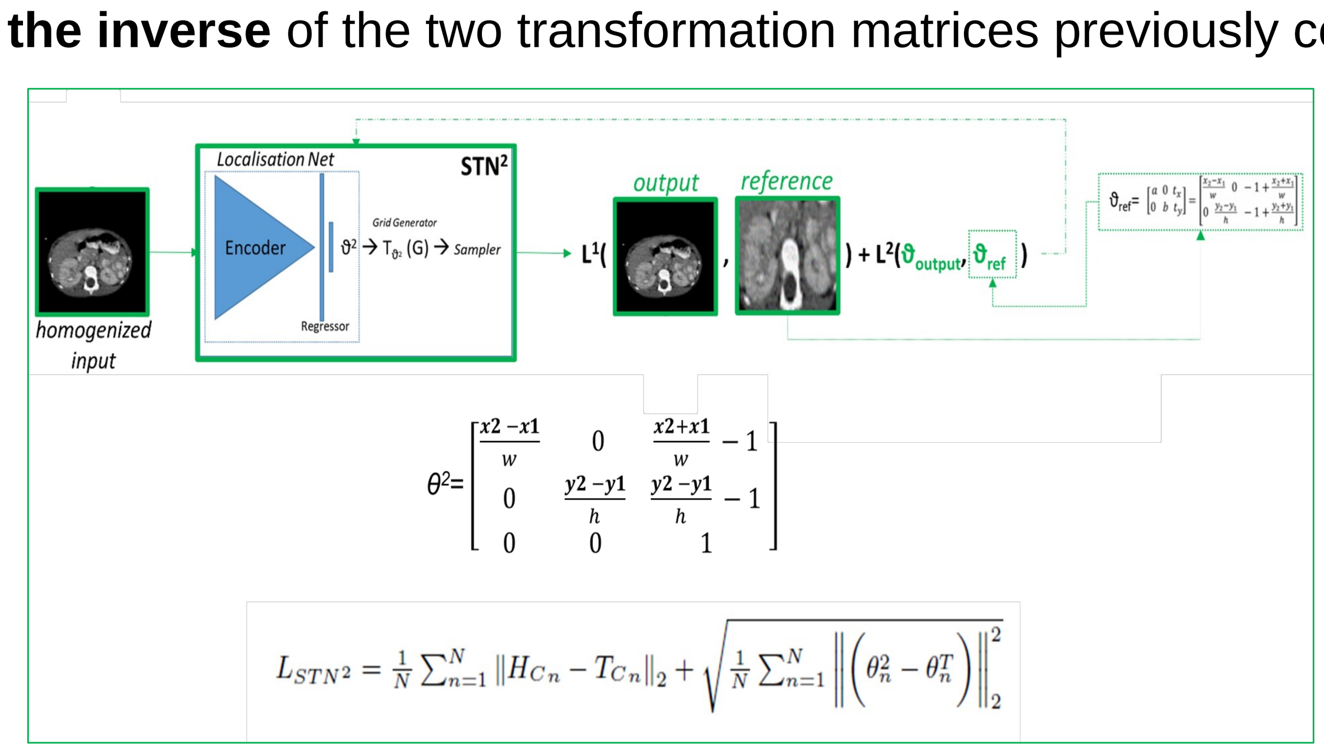
- **Difficulty in transferring information** because of the different morphology of structures.
- **CNNs implement only translational equivariance by construction (no rotation and scaling):** strong use of data augmentation is required, with high training times (7 days) and large space required (GPU of 16 Gb).
- Having a **sufficient database** manually segmented **takes months of work** (4 months for ours).

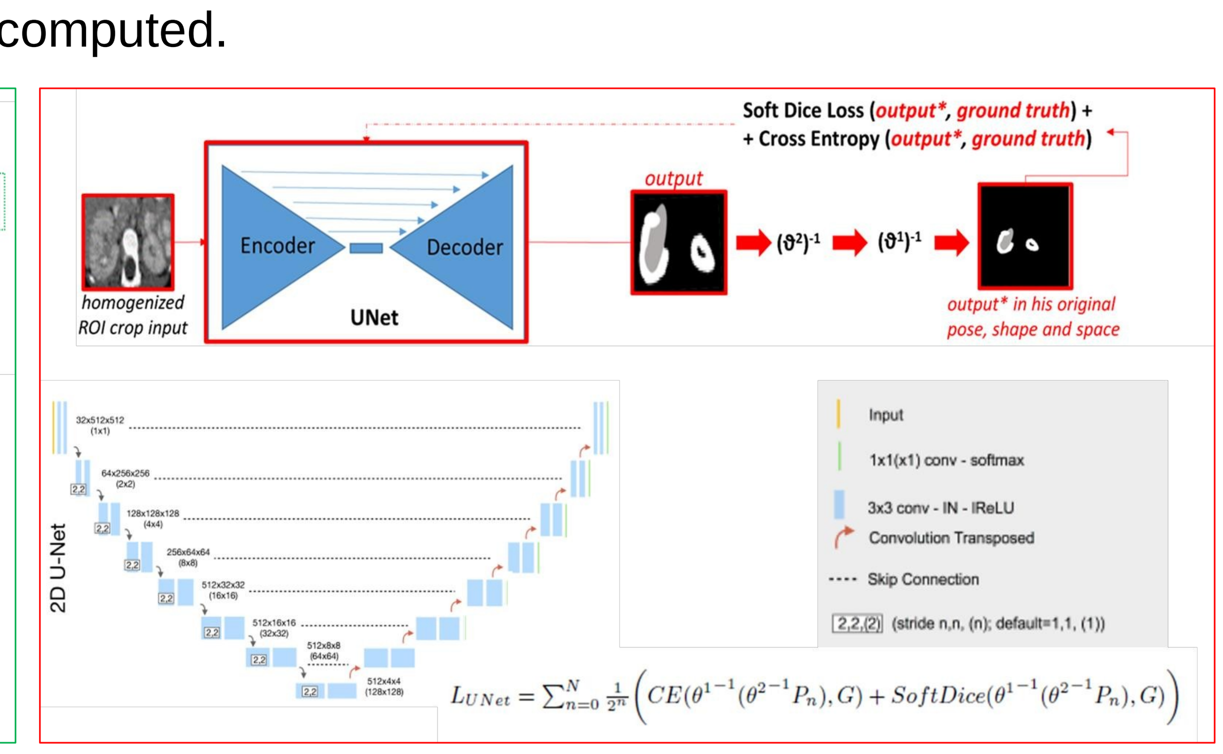
PROPOSED METHOD

We propose solutions to the above problems using **Spatial Transfomer Network (STN)** [4], a differentiable module that learns how to perform spatial manipulation on the input image in order to enhance the geometric invariance of the model. **Our architecture is composed of three sequential modules:**

- **a first STN** that deals with **homogenization of pose and size**, transforming all images to be as similar as to a chosen one;
- **a second STN** that **crops** the homogenized image **in the region of interest (ROI)**, where the structures to be segmented are present;
- **finally a segmentation network**, built as a nnUNet, in which the cropped homogenized image is given as input and **the output is then restored** to its original pose and size, and uncropped, **using the inverse** of the two transformation matrices previously computed.







PRELIMINARY RESULTS ON 2D

- The results show that the use of the STN[1] to homogenize pose and size **outperforms the baseline** with data augmentation **in time and for the tumor segmentation task**, while showing **comparable results for the kidney segmentation.**
- The combination of the two STNs does not lead to improvements compared to using STN[1] alone for the segmentation tasks. This still requires some work but it already leads to **a significant gain in time and requested memory, while not losing performance.**

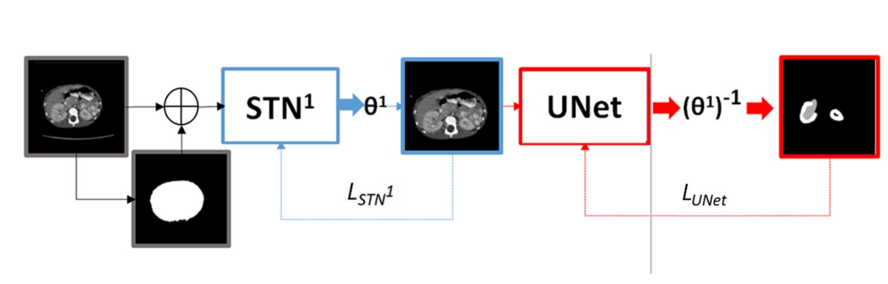
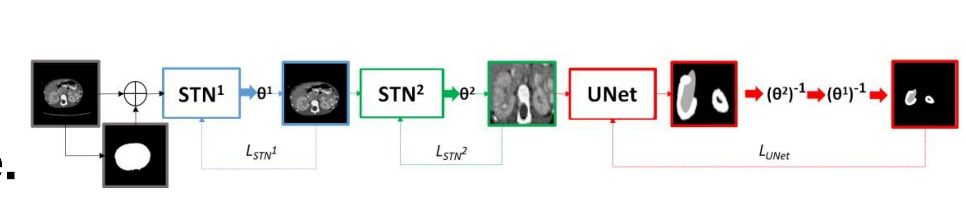


Image 128x128 with Batch Size of 32			
Architecture	Training Time	Dice Score Kidney	Dice Score Tumor
nnUNet	1h35	83.66 (7.88)	69.52 (24.61)
nnUNet (+ data augmentation)	2h15	**88.99 (3.71)**	74.18 (22.07)
STN pose-size + nnUNet	**1h45**	86.75 (6.47)	**77.31 (27.36)**
Image 512x512 with Batch Size of 12			
Architecture	Training Time	Dice Score Kidney	Dice Score Tumor
nnUNet	22h	88.07 (5.61)	78.14 (26.19)
nnUNet (+ data augmentation)	33h	**88.91 (5.08)**	85.52 (24.65)
STN pose-size + nnUNet	**25h**	88.01 (6.25)	**87.12 (23.39)**



Architecture	Input size UNet	Training Time	Memory allocated	Dice score kidney	Dice score tumor
nnUNet	512×512	22h	10.05Gb	88.07 (5.61)	78.14 (26.19)
nnUNet (+ data augmentation)	512×512	33h	10.05Gb	88.91 (5.08)	85.52 (24.65)
STN pose-size + STN crop + nnUNet	512×512	28h	10.05Gb	88.84 (7.79)	84.25 (31.15)
STN pose-size + STN crop + nnUNet	256×256	**19h30**	**3.52Gb**	86.71 (19.36)	84.15 (30.11)

FUTURE WORK

- **Improve the STN for cropping:** the drop in performance depends on the renal tumor size, sometimes bigger than 256x256, so we downsample the ROI loosing important information.
- **Extend it to 3D**: with the hope of gaining even more time and requested memory respect to the baseline and compared to the already excellent performances found in 2D.
- **Insert in STN the use of *diffeomorphism* for transfer learning** (as presented in [5]): allowing even smaller databases to achieve high performance.
- **Segment other structures** such as ureters, arteries and veins: manual segmentation is currently in progress.

[1] N. Heller, F. Isensee, K. H. Maier-Hein, X. Hou, C. Xie, F. Li, Y. Nan, G. Mu, Z. Lin, Miofei Han, et al., "The state of the art in kidney and kidney tumor segmentation in contrast-enhanced ct imaging: Results of the KiTS19 challenge," Medical Image Analysis, vol. 67, 2021. [2] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, Wirkert S., and K. H. Maier-Hein, "nnU-Net: self-adapting framework for u-netbased medical image segmentation," Bildverarbeitung f¨ur die Medizin, Springer Vieweg, Wiesbaden, 2019. [4] G. La Barbera, I. Bloch, G. Barraza, C. Adamsbaum, and P Gori, "Robust segmentation of corpus callosum in multi-scanner pediatric T1-w MRI using transfer learning," in OHBM, 2019. [4] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," in NIPS, 2015. [5] N. S. Detlefsen, O. Freifeld, and S. Hauberg, Deep Diffeomorphic Transformer Networks. CVPR IEEE, 2018.

- LTCI, Télécom Paris, Institut Polytechnique de Paris (bureau 5.B56)
- Hôpital Necker - IMAG2 (bureau Bâtiment Lavoisier)

giammarco.labarbera@telecom-paris.fr

Giammarco La Barbera