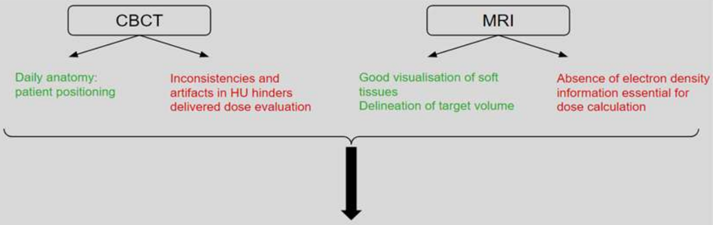
Guiding Unsupervised MRI-to-CT and CBCT-to-CT synthesis using Content and style Representation by an Enhanced Perceptual synthesis (CREPs) loss

C. Hémon[1], V. Boussot[1], B. Texier[1], JL. Dillenseger[1], JC. Nunes[1]

[1] LTSI-INSERM-UMR 1099, Univ. Rennes, CHU Rennes, CLCC Eugène Marquis, F-35000 Rennes, France

INTRODUCTION

This study aims to investigate the generation of sCT for use in external radiotherapy, with the goal of facilitating online radiotherapy guidance or enabling MRI-only workflow.



Generation of **synthetic CT** (sCT) from CBCT and MRI

- Supervised generation: requires **perfect registration**, **source of uncertainties [1][2].**
- Solution : **unsupervised** learning, no need for registration [1][2].
- Leverages unlabeled data to extract valuable information and patterns. **Better robustness** but **less accurate generation** for large dataset.
- Unsupervised learning encounter challenges related to **non-convergence** and **instability** attributed to the **lack** of an **explicit loss function.**

Despite the instability inherent in the absence of ground truth, this study proposes integrating our enhanced **perceptual loss (PL) [3]** as an additional pre-trained discriminator to **stabilize the training.**

The Content and style Representation by an Enhanced Perceptual synthesis (CREPs) loss is used for both tasks as follows:

- In MRI, use of style loss to impose **constraints** on the generator's **output**, ensuring it produces **CT-like styles.**
- In CBCT, the generation challenge is approached as a **style transfer problem**, involving the use of **CBCT content** while incorporating the **style** of the **CT.**

MATERIALS AND METHODS

- **Database**
  - Brain : **180** pairs CT/MRI or **180** pairs CT/CBCT
  - Pelvis : **180** pairs CT/MRI or **180** pairs CT/CBCT

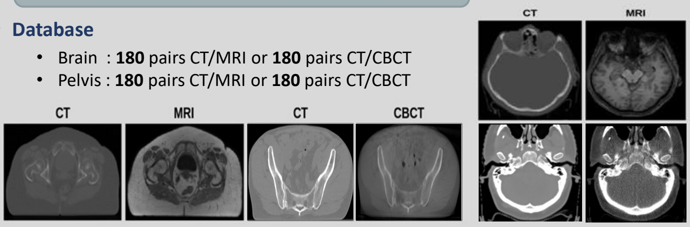


Fig 1 : Composition of the database

1) Preprocessing

- Patch size:

MR	CBCT
Brain : 66 x 168 x 168 **Pelvis** : 32 x 224 x 224	**Brain** : 168 x 168 **Pelvis** : 224 x 224

- Intensity range **scaled down** by a factor of **1000** ⇨ **[-1,3]**, **zero padding** and for CBCT **histogram matching** to **reference CT** scan (at training-time) and to **training CT** with highest mutual information (at test-time)

2) Deep learning model (trained separately for each location and modality)

- Customized unsupervised 3D cGAN [4] (6-blocks Resnet + PatchGAN) for MR-to-CT synthesis.
- Analogous 2D cGAN for CBCT-to-CT synthesis.

➢ **Loss function :**

The sCT generation problem is considered as a style transfer problem [3], based on the idea of independence between style and content..

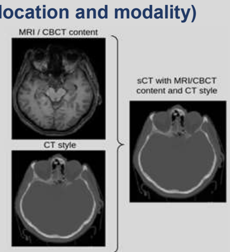


Fig 2: Illustration of the content and style transfer

- **Style loss:**

$$Gram_j^{\phi}(I) = \frac{flat(\phi(i))flat(\phi(j))^T}{C_j * H_j * W * j} \quad l_{style}^{\phi,j}(CT, sCT) = \left\| Gram_j(CT) - Gram_j(CT) \right\|_1$$

- **Content loss:**

$$l_{content}^{\phi,j}(CBCT, sCT) = \left\| \phi_j(CBCT) - \phi_j(sCT) \right\|_1$$

- Creation of a new PL (**CREPs** loss) based on ConvNext-Tiny [5] (Fig. 3)
- Training time is divided by 2 compared to VGG and significantly reduces the required VRAM space (23 GB) by 13 GB (Table 2).

3) Training

- Batch size of **8** for MR and **128** for CBCT
- AdamW optimizer (weight decay = 10^{-3}), constant **lr** scheduler (10^{-4}), **no optimization** of network **hyper-parameters** on the validation set.
- **200** epochs for MR and **100** epochs for CBCT, epoch selected with the best results on subset validation in terms of **MAE sCT/CT**.

- **Style + Content** loss for **CBCT** and **only Style** loss for **MR** because there are discernible differences in content between the **MR** and **CT**
- **CREPs** loss (ConvNext) improve generation in terms of **MAE**, **PSNR** and **SSIM**.

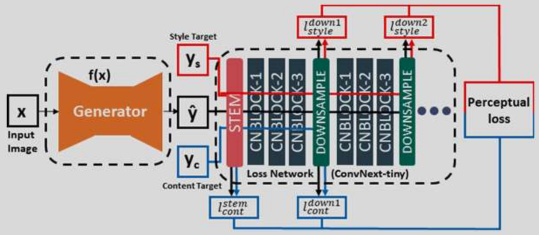


Fig. 3: ConvNext-based Perceptual Loss definition diagram. The new PL is the weighted sum of the content (in blue) and style (in red) loss functions

- **Evaluation on another database [6]**

Table 1. Voxel-wise evaluation of sCT generation with the different architectures on the different structures (External body, air, bones).

Metric / Architecture	MAE(HU)	PSNR(dB)	SSIM
	External body		
VGG	68.6	26.9	0.67
ConvNext	58.5	27.0	0.78

Table 2. Training time and memory space used for the different architectures based on 3D GAN.

Loss functions	Training time(Hours)	VRAM space used(GB)
VGG	15	35.6
ConvNext	7	22.3

4) Postprocessing

- **Median** of **overlapping patches**, removes padding and intensity rescaling (**x1000**)

5) Results

Table 3: Metrics on the challenge dataset

	MAE (HU)		GPR (%)	
Task	1	2	1	2
Max	119.97	139.41	100.0	100.0
Min	50.34	56.78	81.50	85.48
Mean	71.27	93.56	99.24	98.55
Std	13.60	19.01	2.22	2.44

CONCLUSION

- In summary, this generation method of sCTs illustrates their potential utility in dosimetry, whether in the context of MRI-only scenarios [7], dose accumulation (for comparing planned and delivered doses), or even in the realm of online adaptive radiotherapy.
- An intriguing possibility to enhance the quality of sCT involves the integration of a post-processing step for style transfer, where the content is derived from sCT and the closest CT style is determined using mutual information.
- Exploring the development of a 3D CREPs loss tailored to medical images presents an interesting direction, offering the potential for a specialized loss metric designed specifically for the domain of medical imaging.

ACKNOWLEDGEMENTS

The present work was funded by CominLabs CEMMTAUR. This research was partially supported by a PhD scholarship grant from Elekta AB, by Rennes Métropole and by ANR as part of the VATSop project.

REFERENCES

[1] Spadea, M.F., Maspero,M et al.: Deep learning based synthetic-CT generation in radiotherapy and PET: A review. Medical Physics 48(11), 6537–6566 (2021).
[2] Boulanger, M. et al.: Deep learning methods to generate synthetic CT from MRI in radiotherapy: A literature review. Physica Medica 89, 265–281 (Sep 2021).
[3] Johnson, J. et al. Perceptual Losses for Real-Time Style Transfer and Super-Resolution (2016)
[4] Goodfellow, I.J. et al. : Generative Adversarial Networks (2014)
[5] Liu, Z. et al. A ConvNet for the 2020s (2022), arXiv:2201.03545
[6] Hémon , C et al. Unsupervised CBCT-to-CT synthesis using 3D GAN with a New ConvNext-based Perceptual Loss for Head and Neck Cancer Radiotherapy. ISBI 2023
[7] Tahri, S et al., A high-performance method of deep learning for prostate MR-only radiotherapy planning using an optimized Pix2Pix architecture,264 Physica Medica 103 (2022) 108–11