Students: Biel CASTAÑO, David FAGET. Team IberDeep (nº5) bielca98@gmail.com, david.faget_cano@ens-paris-saclay.fr

Harmonizing CT Images via Physics-based Deep Neural Networks

Authors: Mojtaba Zarei, Saman Sotoudeh-Paima, Cindy McCabe, Ehsan Abadi, Ehsan Samei

MATHÉMATIQUES VISION APPRENTISSAGE
école normale supérieure paris–saclay

In medical imaging, the precision of quantifying biomarkers is crucial, yet it faces challenges due to variations in image renditions. These variations can stem from differences in CT scan reconstruction kernels and dose levels, introducing biases that compromise the reliability of studies.

Existing techniques such as parametric empirical Bayes, ComBat and physics-based ones are constrained by strict assumptions or incomplete solutions to variability. Moreover, the use of deep neural networks is limited by they need of large amounts of training data and high quality references.

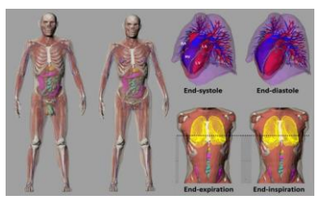
This study overcomes these limitations by using a generative adversarial network (GAN) model with a generator informed by the scanner's modulation transfer function (MTF) and leveraging virtual imaging trials.

• Methods

Materials

Training images and ground truth:

40 XCAT phantoms with different attributes



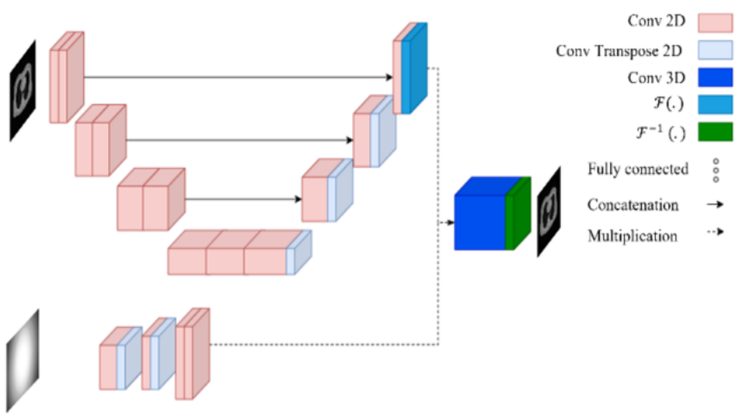
CT simulator with various dose levels, kernels and artifacts

10 with emphysema

30 with healthy lungs or cancer nodules

DNN architecture and training

Generator architecture (U-Net+CNN):



$$I_h(x,y) = H_\theta(I_o(x,y), MTF_{2d}),$$

Texture loss:

$$Loss_t(I_h, I_r) = \sum_{i=1}^{N} MSE\left(\frac{G(VGG_i(I_h))}{\|G(VGG_i(I_h))\|}, \frac{G(VGG_i(I_r))}{\|G(VGG_i(I_r))\|}\right),$$

Generator loss:

$$Loss_H(I_h, I_r) = \lambda_d \times Loss_d(I_h, I_r) + \lambda_t \times Loss_t(I_h, I_r) + \lambda_a \times Loss_a(I_h, I_r),$$

Adversarial loss:

$$Loss_a(I_h, I_r) = \mathbb{E}\left[D_\theta(I_{h,o})\right],$$

Distance loss:

$$Loss_d(I_h, I_r) = \begin{cases} 0.5(I_h - I_r)^2, & |I_h - I_r| \le \delta \\ \delta(|I_h - I_r| - 0.5 * \delta) & |I_h - I_r| > \delta' \end{cases}$$

• Evaluation and Results

Evaluation and Validation Procedure

- Specific models for emphysema and lung cancer, with their CT scans, were isolated for testing.
- The harmonized images' quality was assessed using SSIM, NRMSE, and PSNR metrics, alongside evaluations of biomarker quantification accuracy and morphological radiomic feature analysis.

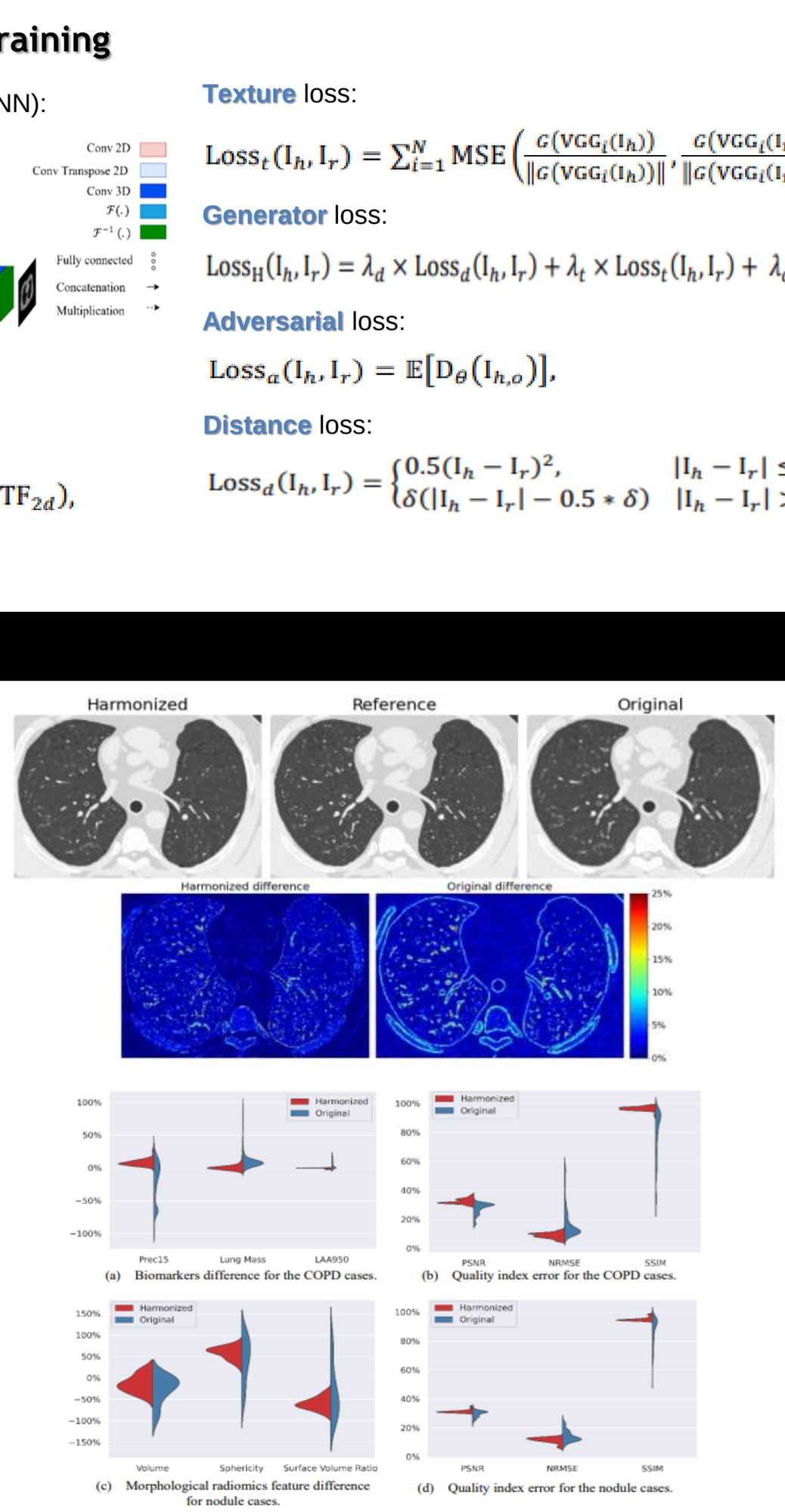
Results

Table 1: Quantitative results on the COPD test set. All metrics were measured on a single 2D slice.

	NRMSE (%)	SSIM (%)	PSNR (dB)
Harmonized	**8.9±2.0**	**96.4±1.4**	**32.4±1.8**
Original	17.1±10.3	77.7±17.7	27.5±4.0
	$\Delta_{LAA\ -950}$	$\Delta_{\ Prec\ 15}$	$\Delta_{\ Lung\ mass}$ (mm^3)
Harmonized	**-1.5±1.8**	**13.6±9.3**	**0.1±0.3**
Original	9.9±12.3	-36.0±53.4	0.2±0.4

Table 2: Quantitative results on the test set with lung nodule inserted. The first three metrics were measured on a single 2D slice, and radiomics features (Volume, Surface to Volume ratio, and Sphericity) were measured on the 3D slices.

	NRMSE (%)	SSIM (%)	PSNR (dB)
Harmonized	**11.5±1.9**	**95.1±1.3**	**31.3±1.1**
Original	14.2±4.0	87.4±9.1	29.6±2.3
	$Volume_{NRMSE(\%)}$	Surface to volume ratio $_{NRMSE(\%)}$	$Sphericity_{\ NRMSE(\%)}$
Harmonized	**2.1±1.3**	**14.2±3.0**	16.5±**3.8**
Original	3.4±2.5	14.8±6.9	**16.3**±8.45



(a) Biomarkers difference for the COPD cases.
(b) Quality index error for the COPD cases.
(c) Morphological radiomics feature difference for nodule cases.
(d) Quality index error for the nodule cases.

Figure 1: On the top we observe visual differences between harmonized image and original image. On the bottom quantitative outcomes are presented.

• Limitations and Conclusion

Broadening testing to include real patient data from diverse clinical settings could significantly improve the model's evidence of effectiveness and its generalization across patient populations and healthcare environments.

Overall, the paper marks a notable progress in medical imaging by introducing a physics-informed GAN for CT image harmonization. Addressing limitations and extending validation could amplify its clinical reliability and diagnostic accuracy.

• Bibliography

[1] Johnson, W. E., Cheng, L., and Rabinovic, A. Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics 8, 1 (2007), 118–127.

[2] Morris, C. N. Parametric empirical bayes inference: theory and applications. Journal of the American statistical Association 78, 381 (1983), 47–55.

[3] Zarei, M., Sotoudeh-Paima, S., McCabe, C., Abadi, E., and Samei, E. Harmonizing CT images via physics-based deep neural networks. 124631Q.