



**HAL**  
open science

# Standardization of Multicentric Image Datasets with Generative Adversarial Networks

Clément Hognon, Florent Tixier, Olivier Gallinato, Thierry Colin, Dimitris Visvikis, Vincent Jaouen

► **To cite this version:**

Clément Hognon, Florent Tixier, Olivier Gallinato, Thierry Colin, Dimitris Visvikis, et al.. Standardization of Multicentric Image Datasets with Generative Adversarial Networks. IEEE Nuclear Science Symposium and Medical Imaging Conference 2019, Oct 2019, Manchester, United Kingdom. hal-02447807

**HAL Id: hal-02447807**

**<https://hal.science/hal-02447807>**

Submitted on 21 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Standardization of Multicentric Image Datasets with Generative Adversarial Networks

Clément Hognon, Florent Tixier, Olivier Gallinato, Thierry Colin, Dimitris Visvikis and Vincent Jaouen

**Abstract**—Due to the sensitivity of medical images to acquisition parameters, multicentric image studies often suffer from a lack of homogeneity in terms of statistical characteristics, also known as the *center effect*. There is therefore a clear need for image-based standardization techniques.

In this paper, we propose a two-step machine learning based framework in which multicentric, heterogeneous images are translated to match the statistical properties of a standard domain. We apply our standardization model to a publicly available multicentric dataset, where we show that we reduce cross-domain while preserving within-domain variability.

## I. INTRODUCTION

A growing number of publications advocate for the validation of medical imaging techniques using datasets from multiple centers [1], [2]. In general, methods developed for a given center cannot indeed directly be applied to another because of the so-called *center effect* [1]. In positron emission tomography (PET), standardization of uptake values between centers and protocols is a long lasting concern [3]. In magnetic resonance imaging (MRI), multiple acquisitions of the same patient using identical imaging systems may lead to variations induced by subtle changes in the acquisition setup [2].

Particular caution must be observed when pooling multicentric data to increase statistical significance. In multicentric radiomics studies, *post-hoc* standardization of extracted features prior to their pooling was shown to increase predictive modelling accuracy [1]. The feature extraction step (e.g. histogram binning, quantization, etc.) is however sensitive to the center effect [4]. It seems difficult to quantify the impact of such *a posteriori* feature standardization on a model’s predictive power. There is therefore a clear need for image-based standardization techniques that would reduce center-induced variability while preserving the properties (e.g. diagnostic power) of individual images.

In this paper, we are interested in standardizing multicentric image datasets directly, prior to transformation or data reduction such as e.g. feature extraction. We propose a machine learning based framework in which multicentric, heterogeneous images are translated to match the statistical properties of a standard dataset, such as a template reference image. To this end, we are inspired by recent developments in image synthesis. Major improvements were achieved in this field in the last two years by image-to-image translation frameworks

based on conditional generative adversarial networks (cGANs) [5], [6], [7]. cGANs are now being applied with success to medical imaging for various applications such as PET attenuation correction or image segmentation [8]. In this paper, we propose to formulate image standardization as a within-modality image synthesis problem, where the *source domain* is a multicentric dataset and the *target domain* is the standard one.

The proposed standardization approach is a two-step process. First, given a reference healthy template image, a standard domain is created through data augmentation. A multicentric dataset is then matched to the target domain using unpaired image-to-image translation with short training time to preserve structures that are not present in the target healthy domain [7]. In a second step, we apply paired image-to-image translation [6] by bootstrapping on the output of the first network. Doing so, we achieve robust image standardization while preserving structures of clinical value.

We apply our standardization model to a publicly available MRI multicentric dataset, where we show that we reduce cross-domain while preserving within-domain variability.

## II. METHOD

Let  $\mathcal{A}$  be a source multicentric domain and  $\mathcal{B}$  a target reference domain. For instance,  $\mathcal{A}$  may be a clinical dataset containing heterogeneous images from multiple centers (possibly showing pathologies), while  $\mathcal{B}$  may contain healthy template subjects from homogeneously-calibrated acquisitions. Our objective is to learn an implicit mapping  $\varphi : \mathcal{A} \rightarrow \mathcal{B}$  conditioned on an input image  $I$  such that translated images  $J = I \circ \varphi$  shows reduced center effect while preserving the clinical characteristics of  $\mathcal{A}$ .

The proposed approach is a two-step image translation process. We first learn a mapping  $\varphi_U : \mathcal{A} \rightarrow \mathcal{B}$  using an unpaired image-to-image translation network between  $\mathcal{A}$  and  $\mathcal{B}$ . To this end, we use a CycleGAN architecture trained with limited number of epochs [7]. CycleGAN can learn domain transfer without paired training data due to a *cyclic loss* where translated images  $\varphi_U(\mathcal{A})$  are back-translated to their original domain, thereby enforcing similarity to the original data. The motivation behind this early stopping stems from our observation that, during unpaired learning training, statistical properties tend to be transferred to the source domain earlier than the warping of image structures. This latter effect is obviously undesired for image standardization. Further training would then lead to structural alterations of features of interest [9], followed by complete overfitting towards the healthy target domain with growing number of epochs.

C. Hognon, D. Visvikis and V. Jaouen are with LaTIM, UMR1101, Inserm, Université de Bretagne Occidentale, Brest, France. C. Hognon is also with Sophia Genetics, Bordeaux, France.

O. Gallinato and T. Colin are with Sophia Genetics, Bordeaux, France.

F. Tixier is with Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, USA.

The second step of our approach consists in learning another mapping  $\varphi_P$  using a paired image-to-image translation network between domain  $\mathcal{A}$  and the previously translated multicentric dataset  $\varphi_U(\mathcal{A})$ . To this end, we use a *pix2pix* network architecture [6]. As the target domain  $\varphi_U(\mathcal{A})$  also contains structures of domain  $\mathcal{A}$ , this second learning step is less prone to alter structural information that we aim at preserving in the final standardized dataset. This second network is the one that we ultimately exploit for standardizing domain  $\mathcal{A}$ .

### III. EXPERIMENTS AND RESULTS

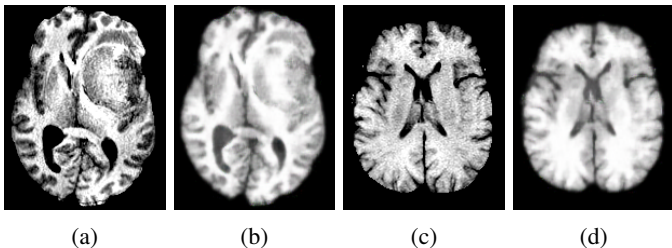


Fig. 1: Example standardization results of the BRATS dataset (a,c) original images (b,d) standardized images.

We conducted experiments using the publicly available BRATS MRI segmentation challenge dataset [10]. This dataset is constituted of 210 multiparametric MR images of high grade glioblastoma multiforme patients and corresponding tumor masks that were acquired in various clinical centers. Acquisitions were performed using a wide range of unknown MR parameters, leading to statistically heterogeneous images. We focused on the standardization of the T1-weighted images. As a target domain, we chose the brainweb T1 image, a well-studied synthetic template of a healthy brain [11].

To reduce training class imbalance between the target set constituted of a single image and the source BRATS dataset, we performed data augmentation through rigid rotations and image scaling. We then performed unpaired learning of the first-step mapping  $\varphi_U$  using the augmented set as the target domain. Two different sets of 50 BRATS subjects were randomly selected for the first and second step of our network training. The CycleGAN network was trained for 20 epochs and the *pix2pix* network for 200 epochs using a single NVIDIA Geforce GTX1080 GPU. Images were processed in 2D using axial brain slices.

Fig. 1 shows standardization results for two BRATS images issued from different source groups (“TCIA04” and “CBICA\_AXO”). Images prior to standardization present heterogeneous visual characteristics. After standardization, regional gray levels become more homogeneous while images maintained their individual structural characteristics. To assess these visual observations more quantitatively, we studied the distribution of intensities within the whole tumor region across three centers (TCIA02: 34 patients, TCIA03: 12 patients, TCIA04: 11 patients). Fig. 2 shows the median kernel density estimates (KDEs) of the whole tumor gray level values for these three sets before and after standardization. After standardization, intensity distributions become more consistent

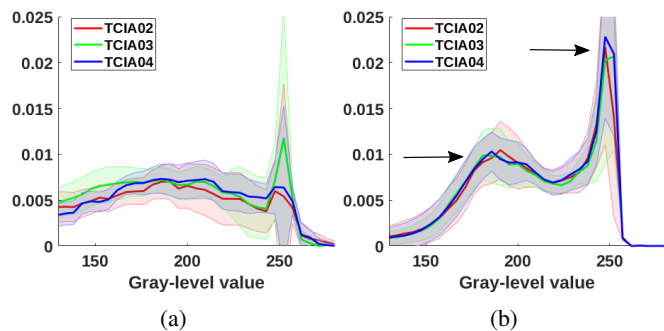


Fig. 2: Kernel density estimations of graylevel values within tumor masks for 3 centers. Curves show median $\pm$ standard deviation. (a) before standardization. (b) after standardization.

across centers while within-domain variance appears to be maintained. The post-standardization KDEs also exhibit two well-distinct modes that are related to sub-regions of the tumors (arrows in Fig. 2b). This could suggest a potential interest of the proposed approach for image segmentation. We also studied the variability of commonly extracted radiomics features (not shown here for space reasons). A similar behavior was obtained, with preserved within-center and reduced cross-center variability.

### IV. CONCLUSION

We have proposed a generic two-step framework for multicentric image-based standardization with conditional generative adversarial networks. The method can be used to standardize heterogeneous datasets with as little as one reference image while simultaneously preserving relevant clinical structures of interest from the source set. Promising results were obtained on an open MRI dataset, although the method is not modality dependent. Future works will focus on further validating the approach in the context of multimodal image segmentation and radiomics using modalities such as PET and computed tomography.

### REFERENCES

- [1] F. Orhac *et al.*, “A postreconstruction harmonization method for multicenter radiomic studies in PET.” *JNM*, 2018.
- [2] L. G. Nyúl *et al.*, “New Variants of a Method of MRI Scale Standardization,” *IEEE TMI*, 2000.
- [3] J. W. K. Jr., “SUV: standard uptake or silly useless values,” *JNM*, 1995.
- [4] T. Upadhyaya *et al.*, “Comparison of radiomics models built through machine learning in a multicentric context with independent testing: Identical data, similar algorithms, different methodologies,” *IEEE TRPMS*, vol. 3, no. 2, pp. 192–200, 2019.
- [5] I. Goodfellow *et al.*, “Generative adversarial nets,” in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [6] P. Isola *et al.*, “Image-to-image translation with conditional adversarial networks,” in *IEEE CVPR*, 2017, pp. 1125–1134.
- [7] J.-Y. Zhu *et al.*, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *IEEE CVPR*, 2017, pp. 2223–2232.
- [8] A. F. Frangi, S. A. Tsafaris, and J. L. Prince, “Simulation and synthesis in medical imaging,” *IEEE TMI*, vol. 37, no. 3, p. 673, 2018.
- [9] J. Cohen *et al.*, “How to cure cancer (in images) with unpaired image translation,” *MIDL*, 2018.
- [10] B. H. Menze *et al.*, “The multimodal brain tumor image segmentation benchmark (BRATS),” *IEEE TMI*, vol. 34, no. 10, pp. 1993–2024, 2015.
- [11] Cocosco *et al.*, “Brainweb: Online interface to a 3D MRI simulated brain database,” in *NeuroImage*, 1997.