



HAL
open science

Synthetic tumor insertion using one-shot generative learning for cross-modal image segmentation

Guillaume Sallé, Pierre-Henri Conze, Nicolas Bousson, Julien Bert, Dimitris Visvikis, Vincent Jaouen

► **To cite this version:**

Guillaume Sallé, Pierre-Henri Conze, Nicolas Bousson, Julien Bert, Dimitris Visvikis, et al.. Synthetic tumor insertion using one-shot generative learning for cross-modal image segmentation. IEEE Nuclear science symposium and medical imaging conference 2021, Oct 2021, Virtual, Japan. hal-03666614

HAL Id: hal-03666614

<https://hal.science/hal-03666614v1>

Submitted on 12 May 2022

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.

Synthetic tumor insertion using one-shot generative learning for cross-modal image segmentation

Guillaume Sallé, Pierre-Henri Conze, Nicolas Boussion, Julien Bert, Dimitris Visvikis and Vincent Jaouen

Abstract—Unpaired cross-modal translation with cyclic loss is being increasingly used for a large variety of medical imaging applications such as e.g. segmentation. However, finer-scale details like tumors may be lost during translation, which is a critical limitation in oncological imaging. In this paper, we propose to address the problem of vanishing tumors for cross-modal segmentation. First, we propose a new method to insert realistic tumors in 3-D images using a deep generative model trained on a single 2-D image. Second, we leverage the proposed model using a new unpaired-then-paired two-stage I2I architecture to better penalize the suppression of tumors in cross-modal segmentation. In our experiments, we validate our model on the ongoing MICCAI crossMoDa tumor segmentation challenge, where we demonstrate superior performance over CycleGAN-based models.

I. INTRODUCTION

II. INTRODUCTION

THE domain shift (DS) refers to differences in data properties during machine learning (ML) model training and deployment [1]. DS is a major challenge for medical imaging applications as it reduces the generalization power of ML models. For this reason, numerous efforts have recently been devoted to the mitigation of DS in medical imaging through advanced processing techniques, referred to as domain adaptation (DA) methods [2]. A popular application of DA in medical imaging is cross-modal image segmentation, whereby a model is trained on a different domain (i.e. different modality, vendor, etc.) than the one seen during deployment [3].

Deep image-to-image (I2I) translation architectures such as CycleGAN-based models [4] allow to learn complex relationships between source and target image distributions in an *unpaired* setting, i.e. without requiring co-registration between target and source images. Due to this appealing property, unpaired approaches are being increasingly used for DA in medical imaging [3], [5], [6]. CycleGAN-based models have however raised concerns due to their tendency to learn a global intensity mapping that may not preserve image features at smaller scales [7], [8]. This drawback is a critically limiting factor in cancer imaging, where a major risk is to hinder the visibility of the pathology (i.e. tumors) in the target domain and thus to reduce the potential interest of these approaches in oncology.

In this paper, we propose to address the problem of vanishing tumors in unpaired image-to-image translation for cross-modal segmentation by combining two contributions. First,

we develop a new method to insert plausible tumors in 3D images using a modified single-image SinGAN model [9]. The SinGAN model has seldomly been used in medical imaging and to our knowledge never in the context of 3D imaging and fake pathology synthesis. Second, we exploit the proposed model in a new two-stage I2I network with pixel loss to better penalize the suppression of tumors in the context of unpaired cross-modal segmentation.

In our experiments, we validate our model by showing improved performances over CycleGAN-based cross-modal segmentation models.

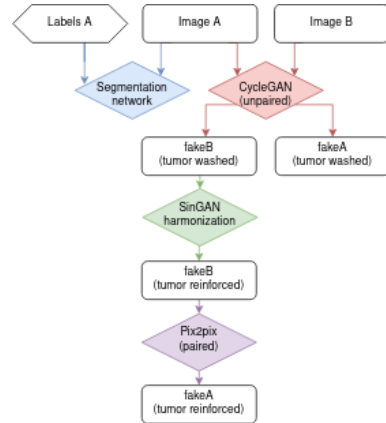


Fig. 1: Training workflow for cross domain segmentation

III. METHOD

Let \mathcal{A} be a source image domain and \mathcal{B} a target image domain both containing tumors. Let $\omega \in \Omega$ the subspace of tumor pixels of the shared physical image space Ω and $\varphi_{AB}: \mathcal{A} \rightarrow \mathcal{B}$ the desired mapping function between the two domains. We assume that an approximate mapping φ_{AB}^0 between \mathcal{A} and \mathcal{B} can be learned using a CycleGAN model with sufficient accuracy in $\Omega \setminus \omega$ but with poor accuracy in ω , i.e. with vanishing tumors in domain \mathcal{B} .

We propose to learn the transfer of the tumor content of $I \in \mathcal{A}$ to the cycleGAN-generated image $J = \varphi_{AB}^0(I)$ with realistic harmonization with respect to the style of domain \mathcal{B} using a modified SinGAN model [9]. SinGAN is a multi-stage generative approach where patches from N given image scales are used to train a series of generators within a single image. A global image generator is decomposed into N scale-specific generators, each of them responsible for learning the distribution D_l at a given scale l in a coarse-to-fine fashion. We harmonize a naively pasted tumor on a given slice by selecting one generator scale $n \leq N$ and using every generator linked to

The authors are with LaTIM, UMR Inserm 1101, Université de Bretagne Occidentale, IMT Atlantique, Brest, France. J. Bert and N. Boussion are also with CHRU Brest University Hospital, Brest, France.

Corresponding address: guillaume.salle@univ-brest.fr

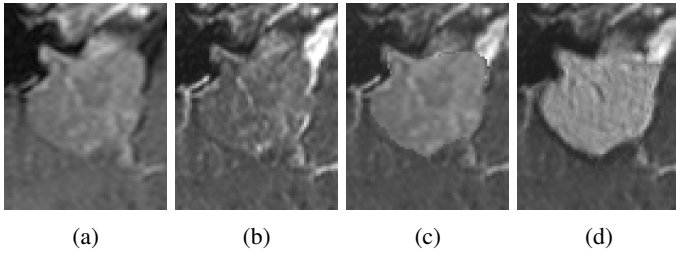


Fig. 2: Fake tumor insertion in T2-weighted MR imaging. (a) initial T1-w image (b) vanishing tumor after CycleGAN translation (c) naively pasted tumor on T2-w (d) proposed harmonized result using our modified SinGAN architecture.

the scales above. The tumor is then inferred by the successive generators on all tumor slices using corresponding masks from \mathcal{A} and slice image from \mathcal{B} to produce a globally harmonized, volumetric, tumor-augmented image $\tilde{J} \in \mathcal{B}'$, where \mathcal{B}' is the output tumor-augmented domain $\tilde{J}(\omega) = \varphi_{AB}(I(\omega)) = (G_N \circ G_{N-1} \circ \dots \circ G_n)(I_k(\omega))$. To this end, the SinGAN original implementation was adapted to accept varying masks and source images during inference with pre-trained generators.

We leverage our SinGAN-based fake tumor insertion method in the context of unpaired cross-modal translation, where the target modality \mathcal{B} does not possess segmentation labels. The proposed workflow is shown in Fig. 1. We first train a segmentation network \mathcal{S}_A on modality \mathcal{A} with ground truth labels and we seek to obtain the best achievable mapping $\varphi_{BA}(\mathcal{B})$ between \mathcal{B} and \mathcal{A} to reduce the domain shift at inference on \mathcal{B} . To this end, we learn an approximate mapping φ_{AB}^0 with CycleGAN (with vanishing tumors). We then enforce tumor presence in \mathcal{B} by naively pasting the tumors of \mathcal{A} to $\varphi_{AB}^0(\mathcal{A})$ that we then harmonize using the proposed SinGAN-based tumor harmonization model \mathcal{B}' . Finally, we learn a more efficient tumor-preserving mapping $\varphi_{B'A}(\mathcal{B}')$ using a conditional GAN with pixel-loss [10] between \mathcal{B}' and \mathcal{A} . The unlabelled images to segment from \mathcal{B} are then passed through $\varphi_{B'A}(\mathcal{B})$ before segmentation by the \mathcal{S}_A network.

IV. EXPERIMENTS AND RESULTS

We focused our experiments on the ongoing *crossMoDA* cross-modal MR segmentation challenge between labelled contrast T1-weighted imaging and unlabelled high resolution T2-weighted images [11]. A nnU-Net segmentation model [12] was used to learn \mathcal{S}_A . 105 unpaired patients were used for training in each modality and 32 other were used for online validation. We compared our model to CycleGAN-based cross-modal translation followed by segmentation by \mathcal{S}_A without fake tumor insertion, using Dice similarity (DSC) and average symmetric surface distance (ASSD). Quantitative results using these indices are summarised in Table 1.

Fig. 2d shows an example fake tumor insertion on the crossMoDa dataset. A naively pasted tumor (Fig. 2c) is added instead of the "washed out" tumor obtained from T1-to-T2 CycleGAN transfer (Fig. 2b) and harmonized with the proposed modified SinGAN approach (Fig. 2d). Figure 3 shows example two segmentation results. Table 1 shows quantitative

results achieved in the validation set, with a clear impact of the proposed SinGAN-based two-stage approach. In 9/32 patients we achieved a DSC greater than 0.8 compared to only 3 for our baseline comparison.

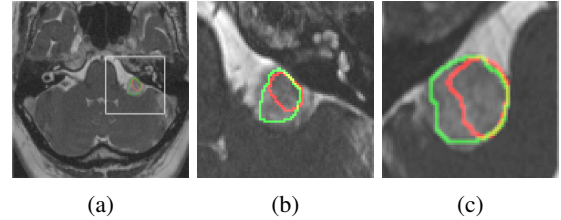


Fig. 3: Test images for vestibular Schwannoma segmentation. (a) example T2 MR image (b,c) results on two subjects. The baseline result is shown in red, ours in green.

	DSC (25%/50%/75%)	ASSD (25%/50%/75%)
Baseline	0.46 ± 0.32 (0.13, 0.61, 0.70)	8.4 ± 14.5 (4.8, 1.6, 1.0)
Ours	0.58 ± 0.31 (0.37, 0.68, 0.81)	7.5 ± 17.9 (2.1, 1.0, 0.6)

Table 1: Quantitative scores for the CrossMoDA validation set.

V. CONCLUSION

In this paper, we have proposed a new method for realistic fake tumor insertion in 3D medical images using a modified SinGAN generative model requiring a single 2D image at the training stage. The method was validated in the challenging context of cross modality image segmentation without target modality ground truth where we demonstrated superior tumor preservation over cycleGAN-based models. Future experiments will focus on validating the approach in other anatomical locations and using combined PET-MR data. Optimization, limitations and potential failure cases will also be further studied.

REFERENCES

- [1] D. C. de Castro, I. Walker, and B. Glocker, "Causality matters in medical imaging," *Nature Communications*, 2019.
- [2] C. Hognon, F. Tixier, O. Gallinato, T. Colin, D. Visvikis, and V. Jaouen, "Standardization of Multicentric Image Datasets with Generative Adversarial Networks," in *IEEE NSS-MIC 2019*.
- [3] Y. Huo, Z. Xu, H. Moon, S. Bao, A. Assad, T. K. Moyo, M. R. Savona, R. G. Abramson, and B. A. Landman, "SynSeg-Net: Synthetic Segmentation Without Target Modality Ground Truth," *IEEE transactions on medical imaging*, vol. 38, no. 4, pp. 1016–1025, 2018.
- [4] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," in *ICCV 2017*.
- [5] J. Hoffman and al., "CyCADA: Cycle-Consistent Adversarial Domain Adaptation," 2017.
- [6] A. Chatsias and al., "Adversarial Image Synthesis for Unpaired Multimodal Cardiac Data," in *Simulation and Synthesis in Medical Imaging*, 2017.
- [7] J. Jiang and al., "Tumor-Aware, Adversarial Domain Adaptation from CT to MRI for Lung Cancer Segmentation," in *MICCAI 2018*.
- [8] J. P. Cohen, M. Luck, and S. Honari, "How to Cure Cancer (in images) with Unpaired Image Translation," in *MIDL 2018*.
- [9] T. R. Shaham, T. Dekel, and T. Michaeli, "SinGAN: Learning a Generative Model from a Single Natural Image," vol. abs/1905.01164, 2019.
- [10] P. Isola *et al.*, "Image-to-image translation with conditional adversarial networks," in *IEEE CVPR*, 2017.
- [11] "CrossMoDA challenge," <https://crossmoda.grand-challenge.org/>.
- [12] F. Isensee and al., "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, 2021.