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-03666614
Submitted on 12 May 2022
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.

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 $\Omega$ and $\varphi_{AB}: \mathcal{A} \to \mathcal{B}$ the desired mapping function between the two domains. We assume that an approximate mapping $\varphi_{AB}^0$ 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 $\omega$, 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 $D_i$ at a given scale $i$ 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 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 $\tilde{J}(\omega) = \varphi_{AB}(J(\omega)) = (G_N \circ G_{N-1} \circ \ldots \circ G_0)(I_k(\omega))$. 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 $S_A$ on modality $\mathcal{A}$ with ground truth labels and we seek to obtain the best achievable mapping $\varphi_{BA}(B)$ between $\mathcal{B}$ and $\mathcal{A}$ to reduce the domain shift at inference on $\mathcal{B}$. To this end, we learn an approximate mapping $\varphi_{AB}$ with CycleGAN (with vanishing tumors). We then enforce tumor presence in $\mathcal{B}$ by naively pasting the tumors of $\mathcal{A}$ to $\varphi_{AB}(\mathcal{A})$ that we then harmonize using the proposed SinGAN-based tumor harmonization model $\mathcal{B}'$. Finally, we learn a more efficient tumor-preserving mapping $\varphi_{BA}(\mathcal{B}')$ 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 $\varphi_{BA}(\mathcal{B})$ before segmentation by the $S_A$ 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 $S_A$. 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 $S_A$ 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.

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