Synthetic CT Generation from MRI Using Improved DualGAN

Denis Prokopenko\textsuperscript{1,2} \hspace{1cm} Denis.Prokopenko@skoltech.ru
Joël Valentin Stadelmann\textsuperscript{2} \hspace{1cm} Joel.Stadelmann@philips.com
Heinrich Schulz\textsuperscript{3} \hspace{1cm} Heinrich.Schulz@philips.com
Steffen Renisch\textsuperscript{3} \hspace{1cm} Steffen.Renisch@philips.com
Dmitry V. Dylov\textsuperscript{1} \hspace{1cm} D.Dylov@skoltech.ru

\textsuperscript{1} Skolkovo Institute of Science and Technology, Moscow, Russian Federation
\textsuperscript{2} Philips Innovation Labs RUS, Moscow, Russian Federation
\textsuperscript{3} Philips GmbH Innovative Technologies, Hamburg, Germany

Editors: Under Review for MIDL 2019

Abstract

Synthetic CT image generation from MRI images is necessary to create radiotherapy plans without the need of co-registered MRI and CT scans. Our model based on GAN with cycle consistency permits unpaired image-to-image translation. Perceptual loss function term and coordinate convolutional layer were added to improve the quality of translated images. The proposed architecture was tested on paired MRI-CT dataset, where the synthetic CTs were compared to corresponding original CT images. The MAE between the synthetic CT images and the real CT scans is 63 HU computed inside of the true CTs body shape.

Keywords: Machine Learning, Deep Learning, Radiology, Computed Tomography, Magnetic Resonance Imaging.

1. Introduction

Radiotherapy (RT) requires a personalised preparation of the treatment plan and, especially, a pre-treatment assessment of the radiation dose. The method is based on two examinations: magnetic resonance imaging (MRI) and computed tomography (CT). The MRI helps to locate the tumour and to outline its shape due to superior soft tissue contrast (Karlsson et al., 2009). The CT scan helps to obtain an electron density map from which the very radiation plan is derived. Following the MRI and CT procedures, their relative alignment, and dose calculation, the patient then undergoes the RT (Coy and Kennelly, 1980; Khuntia et al., 2006). In the modern approach of synthetic CT generation, a patient has to participate only in an MRI procedure, with the advantage of having an error-free registration between the CT and the MRI volumes (Jonsson et al., 2010).

This paper focuses on the unpaired GAN-based approach to the MRI-to-CT image translation. The proposed approach allows the unpaired translation and enables a generation of synthetic CT images from the T1-weighted MRI scans. Our contributions include the upgrade of the cycle consistency with an additional loss function term: the perceptual loss function term that enables a comparison of the high-level image representations obtained with the VGG-16 pre-trained model (Johnson et al., 2016). We also include a coordinate
Synthetic CT Generation from MRI Using Improved DualGAN

covolutional layer (Liu et al., 2018), which helps to spatially localise convolutional properties (Hofmann et al., 2008). The proposed architecture was trained on the brain data with a set of almost overlapping MRI and CT scans. We present meaningful synthetic CT scans and report their quality using the mean absolute error, the peak signal-to-noise ratio, and the structural similarity index.

2. Methods

The DualGAN architecture (Yi et al., 2017) consists of two image generators, which form a cycle, and two discriminators, see Figure 1. The first generator \( G_{MRI \rightarrow CT} \) is trained to translate MRI images to CT images; the second generator \( G_{CT \rightarrow MRI} \) translates images from CT to MRI domain. The discriminators distinguish generated MRI or CT images from real ones and force generators to make more realistic images during translation.

![Figure 1: DualGAN architecture.](image)

The Perceptual loss function builds on the idea of a feature matching, where high-level representations of two images are compared by mean squared error. Although extracted features by pre-trained VGG-16 are not optimised for tomographic images, their use is still relevant because the features extracted in the identical way for both compared images.

The CoordConv layer allows taking into account the spatial information of the images by concatenating two additional slices with \( i \) and \( j \) coordinates it the tensor representation of the image. The CoordConv layer in application to the MRI-to-CT translation helps to distinguish black pixels of MRI image, which could represent either a bone or air.

3. Experiments

Three medical datasets were obtained to train and test the considered methods of unpaired MRI to CT translation. Each set was divided into a train and test parts in 7:3 ratio. The volumes were normalised and preprocessed.

The CPTAC Phase 3 dataset includes the MRI T1-weighted images of 7 patients. Each 3D volume of a patient contains 22 – 24 slices in the axial anatomical plane comprised between the lower nose and the top of the head region (CPTAC, 2018; Clark et al., 2013).

The head-and-neck cancer dataset consists of CT scans of 61 patients. 3D volumes include 61 - 94 slices from the shoulders, neck and lower part of a head up to the eyes (Vallièrèses et al., 2017; Vallièrèses et al., 2017).

The private dataset contained CT images, masks and paired MRI T1-weighted images of 10 patients. The volumes consist of 66 – 137 slices in the axial plane, which are comprised between the teeth and the top of the head region. Train part was used in an unpaired way.
All models were trained separately on CPTAC with Head-and-neck cancer datasets together and the train part of the private dataset. The qualitative results and quantitative estimations of the performance were obtained on the test part of the private dataset.

In this work, four different architectures were considered and compared: DualGAN, DualGAN with the coordinate convolutional layer, DualGAN with the perceptual loss function term and DualGAN with perceptual loss function term and coordinate convolutional layer. All of them work with three-channel images allowing the application of VGG network. The initial images were copied into three channels to allow the application of the VGG-16 pre-trained network. Training parameters were kept the same for all models.

Performance estimations were calculated using one channel images. The quantitative performance results of different translation configurations can be seen in Table 1, while qualitative results are presented in Figure 2, which shows by column the real MRI slices, the synthetic CT images, the real paired CT slices and the difference between synthetic and original CTs.

Table 1: Performance comparison of different MRI to CT translation configurations.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>MAE, HU ↓</th>
<th>PSNR, dB ↑</th>
<th>SSIM ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>DualGAN</td>
<td>67.63</td>
<td>16.72</td>
<td>0.777</td>
</tr>
<tr>
<td>DualGAN, CC</td>
<td>90.3</td>
<td>15</td>
<td>0.7115</td>
</tr>
<tr>
<td>DualGAN, VGG</td>
<td>67.47</td>
<td>16.6</td>
<td>0.7717</td>
</tr>
<tr>
<td>DualGAN, VGG, CC</td>
<td>63.13</td>
<td>17.37</td>
<td>0.7901</td>
</tr>
</tbody>
</table>

Figure 2: Translation made by DualGAN with VGG loss function term and CC layer.

4. Conclusions

The presented model achieves MAE of 63 HU and SSIM of 0.7901, which compares favourably to scientific literature (Wolterink et al., 2017). The translation architecture transforms the initial image retaining the structural information. While the DualGAN with the perceptual loss function term and the DualGAN with the coordinate convolutional layer are not themselves superior to the DualGAN model, their triad combination showed meaningful improvement. Besides, our model works with the unpaired images, and the visual examination confirmed that the use of the perceptual loss term and the CoordConv layer enhances the appearance of the resulting images. The error of translation increases with the complexity of the inner structures for instance in the nose and teeth regions.
5. Acknowledgements

Data used in this publication were generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC).

References


Synthetic CT Generation from MRI Using Improved DualGAN

