MR to CT synthesis using GANs : a practical guide applied to thoracic imaging

Arthur Longuefosse, Baudouin Denis de Senneville, Gaël Dournes, Ilyes Benlala, François Laurent, Pascal Desbarats, Fabien Baldacci

MR to CT synthesis using GANs : a practical guide applied to thoracic imaging

Arthur Longuefosse[1], Baudouin Denis De Senneville[2], Gaël Dournes[3], Ilyes Benlala[3], François Laurent[3], Pascal Desbarats[1], and Fabien Baldacci[1]

[1] LaBRI, Université de Bordeaux, Talence, France
[2] Institut de Mathématiques de Bordeaux Université de Bordeaux, Talence, France
[3] Service d'Imagerie Médicale Radiologie Diagnostique et Thérapeutique, CHU de Bordeaux, France

Abstract. In medical imaging, MR-to-CT synthesis has been extensively studied. The primary motivation is to benefit from the quality of the CT signal, i.e. excellent spatial resolution, high contrast, and sharpness, while avoiding patient exposure to CT ionizing radiation, by relying on the safe and non-invasive nature of MRI. Recent studies have successfully used deep learning methods for cross-modality synthesis, notably with the use of conditional Generative Adversarial Networks (cGAN), due to their ability to create realistic images in a target domain from an input in a source domain. In this study, we examine in detail the different steps required for cross-modality translation using GANs applied to MR-to-CT lung synthesis, from data representation and pre-processing to the type of method and loss function selection. The different alternatives for each step were evaluated using a quantitative comparison of intensities inside the lungs, as well as bronchial segmentations between synthetic and ground truth CTs. Finally, a general guideline for cross-modality medical synthesis is proposed, bringing together best practices from generation to evaluation.

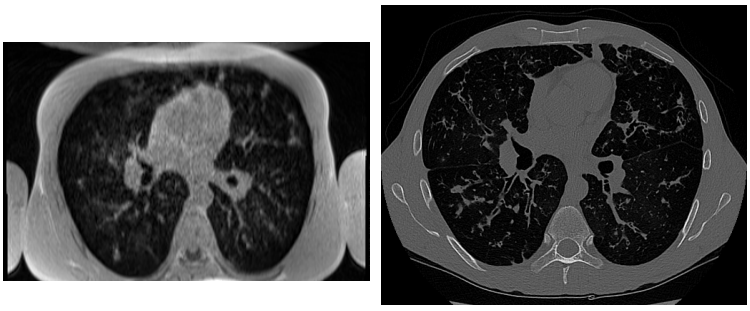
Keywords: GAN · CT Synthesis · Lung.

1 Introduction

In clinical practice, computed tomography (CT) is typically used to diagnose lung conditions. However, this modality exposes patients to ionizing radiation, which may have negative effects on their health. Recently, lung MRI with ultrashort or zero echo-time (UTE/ZTE) has shown promise for high-resolution structural imaging of the lung [1, 2]. However, the appearance of images obtained using this technique is substantially different from those obtained using CT, notably imaging texture, blurring, and noise which has limited its adoption in clinical practice (cf Figure 1). The generation of CT images from MRI may be a good alternative and could improve patient diagnosis by providing high quality images to radiologists based solely on the safe and non-invasive nature of MRI. Over recent years, deep learning approaches, particularly generative adversarial

networks (GANs) [3], have been extensively studied for image synthesis in medical imaging. This type of network is made of a generator and a discriminator, and is able to produce high quality synthetic data similar to a given dataset by learning a complex non-linear relationship between MR and CT. Previous research on cross-modality synthesis has used GANs to synthesize images in several different regions of the body, such as the brain [4, 5], pelvic region [6], and also in the lungs using Dixon MRI [7]. Many studies have been conducted on the development of specific GAN models, including unpaired methods based on cycleGAN [8] and paired methods based on pix2pix [9]. In addition, research also focused on the development of various loss functions, such as cycle consistency [8], feature-matching [12], perceptual [21], and contrastive loss [22]. However, most state-of-the-art studies are limited to these developments and do not properly address the full range of steps involved in medical translation tasks, such as preprocessing and robust evaluation.

In this paper, we present a general guideline for image-to-image translation applied to thoracic MR to CT synthesis, covering key topics such as pre-processing steps, data normalization and quantization, and the importance of an adapted resampling before registering the input. We review the different types of GANs and losses and compare their performances in thoracic image-to-image translation. A quantitative evaluation of the different models and parameters is presented, using traditional metrics as well as a comparison of the segmentations of airways in synthesized CT images versus ground truth CT images, to help identify the factors that have the biggest impact on the performance of medical image-to-image translation. Overall, our evaluation helps to provide a better understanding of the different models and parameters used for medical image-to-image translation and can serve as a useful reference for researchers and practitioners in this field.



(a) UTE MR (b) CT

Fig. 1: Visual comparison of thoracic UTE MR and CT modalities of the same patient at a corresponding axial slice. The CT scan shows a higher signal quality, greater contrast and sharpness, and fewer artifacts compared to the MRI.

2 Methods

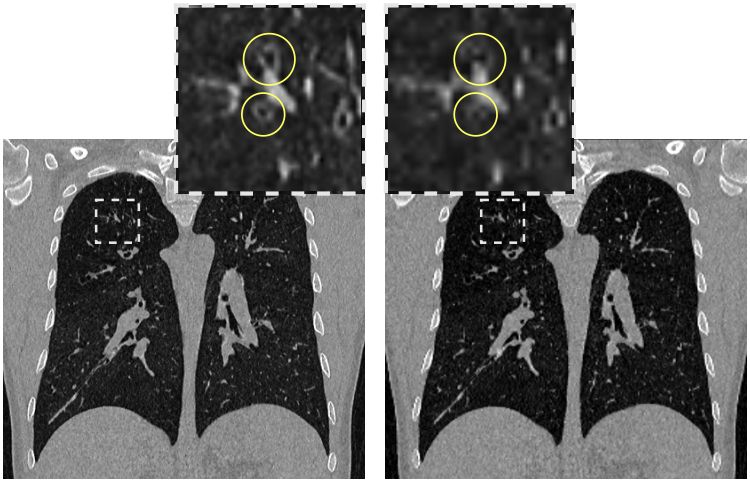
2.1 Data acquisition

The dataset used in this study consists of UTE MR and CT thoracic images of 110 patients. Both modalities were acquired on the same day, from 2018 to 2022. CT images were obtained using a Siemens SOMATOM Force, in end-expiration, with sharp filters. The parameters used were a DLP of 10 mGy.cm and a SAFIRE iterative reconstruction. UTE MR images were acquired using the SpiralVibe sequence on a SIEMENS Aera scanner, with the following parameters: TR/TE/flip angle=4.1ms/0.07ms/5°. Since the slice plane is encoded in Cartesian mode, the native acquisition was performed in the coronal plane with field-of-view outside the anterior and posterior chest edges to prevent aliasing. It should be noted that resolutions, voxel spacings, and fields of view are not identical in CT and MR images. In addition, modalities may have been taken at different points in the respiratory cycle. To obtain a paired dataset, an adequate resampling and a deformable registration will thus be required between CT and MR volumes.

2.2 Preprocessing

Resampling In multimodal registration, it is typically advised to use the image with the highest resolution as the fixed image and the image with the lower resolution as the moving image, since a higher level of detail and accuracy in the fixed image can help improve the performance of the registration process. In our case, we have to register the CT volume, with a voxel size of $0.6 \times 0.6 \times 0.6$ mm^3, on the MRI, with a voxel size of $1 \times 1 \times 1$ mm^3, which implies a resampling of the CT to the MRI resolution, and thus a loss of information, as shown in Figure 2. To avoid this issue, we propose to upsample the MRI voxel size to the CT voxel size, allowing to keep the initial resolution of the CT, which implies a better convergence of the registration algorithm as well as better performances for the GAN. The two modalities are therefore resampled on a common grid of $0.6 \times 0.6 \times 0.6$ mm^3 using tricubic interpolation. For comparison purposes, CT and MR volumes were also resampled on a $1 \times 1 \times 1$ mm^3 grid.

Multimodal registration Accurate alignment of images from different modalities often requires non-rigid registration, especially in parts of the body subject to severe periodic deformations, such as cardiac and respiratory motions. Edge-alignment methods seems particularly well suited for multimodal medical registration since they don't rely on input landmarks and can overcome differences in intensity and contrast between modalities, by focusing on boundary information. In our dataset, a rigid translation is estimated to ease convergence, before the EVolution algorithm [14] is employed to estimate the elastic deformation, a patch-based approach that includes a diffusion regularization term and a similarity term that favors edge alignments. To prevent physically implausible folding of the volumes during the registration process, a diffeomorphic transformation is ensured by minimizing the inverse consistency error ([15], [16]).



(a) CT $0.6 \times 0.6 \times 0.6$ mm^3 (b) CT $1 \times 1 \times 1$ mm^3

Fig. 2: Visual comparison of identical CT slices at different voxel spacing, with zoomed regions that highlight bronchi, circled in yellow. Pulmonary bronchi are nearly indistinguishable in the 1 mm^3 version due to the lower resolution of the image.

Intensity normalization CT and MR modalities have fundamental differences that must be taken into account when normalizing intensity values. CT intensity values are defined in Hounsfield units (HU) and have a physical meaning, whereas MR intensity values strongly depend on acquisition parameters. Therefore, methods used for intensity normalization must be tailored to the specific characteristics of each modality.
In our study, CT intensities are cropped to [-1000; 2000] HU window to remove irrelevant values from the table or background and rescaled to [-1; 1] using the same window limits. MR intensity inhomogeneities, also known as bias field, are first corrected using the popular N4 bias field correction algorithm [11]. MR values are then normalized using z-score, i.e. zero mean and unit variance, cropped to $[-3\sigma\ ;\ 3\sigma]$ to remove outliers, σ being the standard deviation, and rescaled to [-1; 1] based on minimum and maximum intensities. Nyul histogram matching [17] was also considered, but the findings of Reinhold et al. [18] indicated that the synthesis process was robust to the choice of MR normalization method used. As a result, we opted for a traditional Z-score normalization approach.

Field of view standardization As shown in Figure 1, modalities may have different fields of view (FOV). Due to the use of a narrow beams of X-rays to produce images, CT field of view is typically limited to a small area of the body, whereas MRI allows capturing a wider field of view. In our dataset, patients may also be in a different position depending on the modality. This is reflected

by the visibility of the arms on the MRI, as opposed to the CT image. To uniformize FOV, which can be useful to speed up calculations and guide training, one common approach is to identify a region of interest (ROI) using segmentation methods. Few methods for lung segmentation in MRI have been developed due to the lower signal and contrast, as well as the lack of data. On the other hand, many effective methods are available for CT, such as the U-Net R-231 convolutional network [19]. The CT volume being registered on the MRI, it is then possible to apply the segmentation of the lungs from the CT on the MRI, allowing to obtain the same FOV on both modalities. All axial slices are then either cropped or zero-padded to 512×512, depending on the CT lung mask size.

Impact of intensity quantization In this study, we also investigate the impact of the bit depth of input medical images on the performance of a GAN in lung MR to CT translation. We create two datasets, one in line with most of the current state-of-the-art papers in the field with 8-bit images, and another dataset with 16-bit images, and evaluate the GAN's performance on each dataset. This allows us to determine whether using higher bit-depth images can improve the performance of the GAN for thoracic CT synthesis.

2.3 Image-to-image translation

Conditional generative adversarial networks (cGANs), are a variant of GANs trained with additional constraints on a specific input image and have demonstrated significant potential for image-to-image translation tasks. CGANs are typically divided into two main categories: unpaired methods, often based on the CycleGAN model [8], designed for image-to-image translation without the need for corresponding pairs of images, and paired methods, based on the pix2pix model, using corresponding pairs of images.
Since the introduction of the cycle consistency loss in CycleGAN, many unpaired methods have been developed, including NICE-GAN [20], a decoupled network training method that uses the discriminator to encode the image of the target domain. As for the paired methods, pix2pix improvements are described in the method pix2pixHD [12], which is no longer dependent on the pixel-wise loss, but on a new feature matching loss, as well as a multi-scale discriminator and a perceptual loss. SPADE [13] also enhanced the performance of paired methods by injecting class-specific information into the generator network. This model introduces a spatially adaptive normalization based on the inputs, that improves the performance and reliability of the generator, allowing synthesized images that are conditioned on the input class. The SPADE architecture can be integrated in other models, such as pix2pixHD, to apply additional constraints on the inputs and guide training.
Finally, recent works on paired image-to-image translation developed a new type of bidirectional contrastive loss, called PatchNCE loss [22], that assesses the similarity between two images based on the mutual information from embedded patches, unlike GANs discriminator that only evaluates the realism of a synthesized image. This contrastive loss produces a smooth and interpretable loss

trajectory, which makes it easier to evaluate the convergence of the training process and determine the number of epochs needed. This is a common challenge with GANs since their traditional loss functions tend to be noisy and provide no clear indication of training progress.

2.4 Assesment of the MR to CT synthesis

In order to evaluate the performance of generative models, past research has proposed several extrinsic evaluation measures, most notably Inception Distances (FID [23], KID [24]), which compare the generated images to a set of real images and assess their quality and similarity. Such measures have been proven to be insensitive to global structural problems [25], and may not be sufficient for the evaluation of medical image translation.
Traditional image processing metrics, such as MSE, PSNR, SSIM, are the state-of-the-art reference metrics for evaluating synthetic images. They can provide information on how well the model preserves spatial structure and content of the original images, but are still highly sensitive to noise and distortion, and may not accurately reflect the visual quality of an image with low-level artifacts [26]. Our assumption is that task-specific metrics are required to accurately quantify synthesized images and evaluate the performance of a model, by taking into account the structure and semantics of the images. In our dataset, we defined them as Dice score, precision, and sensitivity between synthesized and ground truth bronchial tree segmentations, by using NaviAirway [27], a bronchiole-sensitive airway segmentation pipeline designed for CT data. This allows us to accurately quantify false positives and false negatives at the bronchial level for each synthetic CT image. A qualitative evaluation conducted by radiologists or other medical experts can also be valuable, to ensure that the translation has preserved overall fidelity with the ground truth, as well as key diagnostic information.

3 Experiments and results

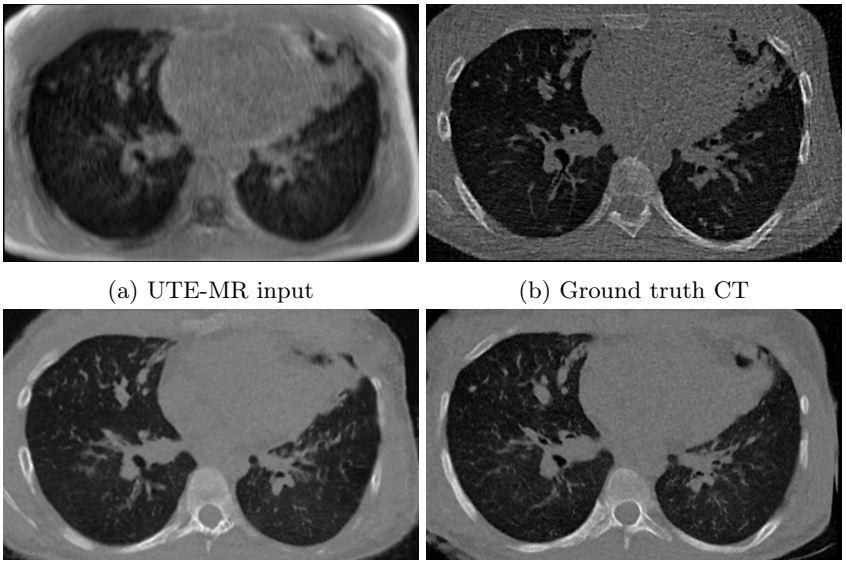
The initial dataset of 110 MR-CT thoracic images is split in a training set of 82 patients, and testing set of 28 patients. Although 3D GANs allow perception of volumetric and neighborhood spatial information, they involve an excessive computational cost and a reduction of the number of samples, which can be challenging to implement for some datasets. Therefore, we choose to train the models on the 2D axial slices of the CT and MR volumes, and define datasets that will allow us to assess the impact of each preprocessing step :

- unpaired dataset with unregistered $0.6 \times 0.6 \times 0.6$ mm^3 CT, 8-bit
- unpaired dataset with registered $0.6 \times 0.6 \times 0.6$ mm^3 CT, 8-bit
- paired dataset with registered $0.6 \times 0.6 \times 0.6$ mm^3 CT, 8-bit
- paired dataset with registered $0.6 \times 0.6 \times 0.6$ mm^3 CT, 16-bit
- paired dataset with registered $1 \times 1 \times 1$ mm^3 CT, 16-bit

The NICE-GAN model was trained using unpaired datasets, while the pix2pixHD and SPADE models were trained using paired datasets. We also evaluated the performance gain of the contrastive loss when applied to these paired methods.

Table 1: Mean squared error (MSE), cross-correlation (CC) and structural similariy index (SSIM) between synthesized CT and real CT inside the lungs.

Model	MSE	CC	SSIM
NICE-GAN	88,45 ± 9,91	0,9283 ± 0,023	0,9725 ± 0,024
NICE-GAN registered	82,41 ± 9,74	0,9406 ± 0,015	0,9776 ± 0,022
pix2pixHD	78,46 ± 13,03	0,9499 ± 0,011	0,9834 ± 0,032
pix2pixHD w/ contrast	75,51 ± 12,05	0,9557 ± 0,010	0,9900 ± 0,030
SPADE	67,82 ± 8,18	0,9635 ± 0,083	0,9915 ± 0,016
SPADE w/ contrast	**67,53 ± 7,31**	**0,9646 ± 0,088**	0,9927 ± 0,017
SPADE 8-bit w/ contrast	67,76 ± 7,70	0,9630 ± 0,096	**0,9932 ± 0,018**
SPADE 1mm^3 w/ contrast	76,36 ± 8,79	0,9505 ± 0,016	0,9830 ± 0,026



(a) UTE-MR input (b) Ground truth CT

(c) SPADE 1 mm^3 (d) SPADE $0.6 \times 0.6 \times 0.6$ mm^3

Fig. 3: Comparison between MR, synthetic CT from SPADE $1 \times 1 \times 1$ mm^3 and SPADE $0.6 \times 0.6 \times 0.6$ mm^3 and ground truth CT axial slices.

All models are trained using the same procedure and architecture defined in the respective papers, apart from pi2pixHD/SPADE dataloader and inference parts, which have been adapted to support 16-bit input and output arrays. Table 1 lists the quantitative evaluation using mean squared error, cross-correlation and

structural similarity index between synthesized CT images and ground truth CT. Calculations are constrained within the intersection between CT and synthesized CT lung masks, to avoid comparing the backgrounds and narrow the results inside the lungs. Figure 3 shows an example axial slice between input MR, ground truth CT and synthetic CT from SPADE model with different samplings. SPADE results based on CT with a voxel size of $0.6 \times 0.6 \times 0.6$ mm^3 present enhanced contrast and sharpness, and therefore allow a more accurate distinction of vessels and bronchi inside the lungs. To validate these assumptions, we performed the airways segmentation of synthesized and ground truth CT using NaviAirway [27], a bronchiole-sensitive airway segmentation pipeline designed for CT data, and computed dice score, precision, and sensitivity. To enable comparison, the SPADE $1 \times 1 \times 1$ mm^3 was resampled to the same resolution as the ground truth CT before calculating the airways segmentation.

Table 2: Dice, precision and sensitivity between synthesized and ground truth airways segmentations

Model	Dice	Precision	Sensitivity
NICE-GAN	0,590 ± 0,088	0,636 ± 0,0752	0,583 ± 0,1298
NICE-GAN-registered	0,640 ± 0,071	0,665 ± 0,069	0,642 ± 0,104
pix2pixHD	0,707 ± 0,054	0,796 ± 0,060	0,660 ± 0,102
pix2pixHD w/ contrast	0,741 ± 0,031	0,787 ± 0,052	0,715 ± 0,088
SPADE	0,733 ± 0,068	**0,829 ± 0,060**	0,681 ± 0,108
SPADE w/ contrast	**0,743 ± 0,060**	0,819 ± 0,054	0,706 ± 0,104
SPADE 8-bit w/ contrast	0,742 ± 0,055	0,802 ± 0,057	**0,719 ± 0,098**
SPADE 1mm w/ contrast	0,687 ± 0,078	0,766 ± 0,068	0,652 ± 0,120

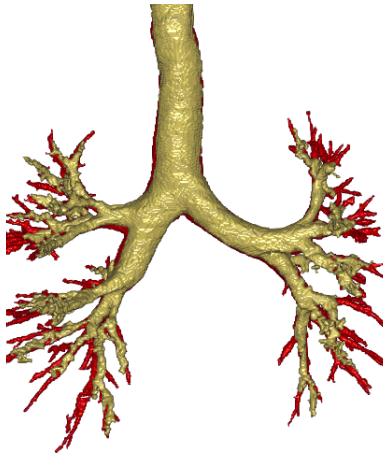


Fig. 4: Airways segmentation example based on SPADE with contrastive loss (yellow) and real CT (red) using the NaviAirway pipeline [27].

4 Discussion

Results from Table 1 based on image processing metrics and Table 2 based on the evaluation of airways segmentation are strongly correlated, with identical trends. Unpaired methods seem to benefit from the elastic registration but produce less satisfactory results than paired methods, which is in agreement with statements in state-of-the-art [10]. Paired pix2pixHD method combined with the conditional normalization layer SPADE provides better performances than the pix2pixHD method alone since it can overcome false positives and false negatives by adding constraints on the inputs. The introduction of the PatchNCE [22] contrastive loss has improved the performance of paired methods, particularly for the pix2pixHD model that tends to diverge. This addition had only a minor impact on the SPADE model, but still provided better control over convergence during training and a more accurate way to differentiate epochs. The performance of the SPADE model with a voxel size of $0.6 \times 0.6 \times 0.6$ mm^3 is significantly superior to that of the model with a voxel size of $1 \times 1 \times 1$ mm^3, both in terms of signal quality and bronchi reconstruction. These results confirm our initial hypothesis that input data should be registered based on the voxel size of the modality with the highest resolution, since downsampling the ground truth reference leads to a loss of information, especially in fine structures such as vessels and bronchi. Surprisingly, the intensity quantization in the input dataset images does not appear to have a significant impact on GAN performances; both SPADE 16-bit and SPADE 8-bit models performed similarly. The reason for this could be that our dataset is composed of highly contrasted information, such as vessels and bronchi in the lungs, and the representation in 8-bit instead of the initial 12-bit would barely impact the reconstruction using GANs. Future works will aim to confirm this hypothesis by conducting similar experiments using different medical datasets in other parts of the body.

5 Conclusion

In this paper, we present a comprehensive guide for medical image translation using GANs. We focus on the importance of data preprocessing, and its impact on performance; the benefits of using a resampling based on the modality with the highest resolution, as opposed to state-of-the-art statements, have been demonstrated. We advocate the use of contrastive loss methods, such as PatchNCE, to address one of the most significant challenges of GANs, which is assessing convergence and stability during training. In addition, we argue that traditional GAN metrics commonly used in the field, such as FID and KID, as well as standard image processing metrics, do not provide sufficient information to adequately evaluate GAN performances in medical image-to-image translation tasks. We recommend defining task-specific quantitative evaluation methods, ideally in conjunction with a qualitative evaluation by experts, in order to robustly assess the performance of a model in this context. In future work, we plan to investigate the validity of our assumptions on different datasets for

other parts of the body and provide guidance on incorporating 3D information into the training process for medical image-to-image translation.

References

1. G. Dournes, D. Grodzki, J. Macey et al. : Quiet Submillimeter MR Imaging of the Lung Is Feasible with a PETRA Sequence at 1.5 T. Radiology. 276. 258-65. (2015)
2. G. Dournes, J. Yazbek, W. Benhassen et al. : 3D ultrashort echo-time MRI of the lung using stack-of-spirals and spherical k -Space coverages: Evaluation in healthy volunteers and parenchymal diseases: Lung MRI With 3D UTE Spiral VIBE Sequence. Journal of Magnetic Resonance Imaging. 48. (2018)
3. I. Goodfellow, J. Pouget-Abadie, Mirza, B. Xu, et al. : Generative adversarial nets. In: Advances in neural information processing systems. pp. 2672–2680 (2014)
4. J. Wolterink, A. Dinkla, M. Savenije et al. : Deep MR to CT Synthesis using Unpaired Data. (2017)
5. D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang : Medical Image Synthesis with Context-Aware Generative Adversarial Networks. 10435. 417-425. (2017)
6. Y. Lei, J. Harms, T. Wang et al. : MRI-Only Based Synthetic CT Generation Using Dense Cycle Consistent Generative Adversarial Networks. Medical Physics. (2019)
7. A. Baydoun, K. Xu, H. Yang, et al. : Dixon-based thorax synthetic CT generation using Generative Adversarial Network. Intelligence-Based Medicine. 3-4. (2020)
8. J. Zhu, T. Park, P. Isola and A. Efros : Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. 2017 IEEE (ICCV) pp. 2242-2251, (2017)
9. P. Isola, J. Zhu, T. Zhou, A. Efros : Image-to-Image Translation with Conditional Adversarial Networks. CVPR, 5967-5976. (2017)
10. S. Kaji, S. Kida : Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging. Radiol Phys Technol. 12. (2019)
11. N.J Tustison, B.B Avants, P.A Cook et al. : N4ITK: improved N3 bias correction. IEEE Transactions on Medical Imaging 29(6), 1310-1320 (2010)
12. T. Wang, M. Liu, J. Zhu, et al : High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. CVPR, 8798-8807. (2018)
13. T. Park, M.-Y. Liu, T.-C. Wang, et al : Semantic Image Synthesis with Spatially-Adaptive Normalization. CVPR, 2337-2346 (2019)
14. B. Denis de Senneville, C. Zachiu, M. Ries, and C. T. W. Moonen : Evolution: an edge-based variational method for non-rigid multi-modal image registration. Physics in Medicine and Biology, 61(20):7377 (2016)
15. Christensen, G., Johnson, H. : Consistent image registration. IEEE Transactions on Medical Imaging 20, 568–582. (2001)
16. M. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. Gleeson, S. Brady, and J. Schnabel. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. Medical Image Analysis, 16(7):1423-1435. (2012)
17. L. G. Nyúl, J. K. Udupa : New Variants of a Method of MRI Scale Standardization. IEEE Transactions on Medical Imaging 19(2), 143–150 (2000)
18. J. C. Reinhold, B. E. Dewey, A. Carass, J. L. Prince : Evaluating the impact of intensity normalization on MR image synthesis. In Medical Imaging 2019: Image Processing, Vol. 10949. (2019)
19. J. Hofmanninger, F. Prayer, J. Pan, et al. : Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European Radiology Experimental. (2020)

20. R. Chen, W. Huang, B. Huang, S. Fuchun, B. Fang : Reusing Discriminators for Encoding Towards Unsupervised Image-to-Image Translation (2020)
21. J. Johnson, A. Alahi, L. Fei-Fei : Perceptual Losses for Real-Time Style Transfer and Super-Resolution. ECCV 2016 (2016)
22. A. Andonian, T. Park, B. Russell, et al. : Contrastive Feature Loss for Image Prediction. Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 1934–1943 (2021)
23. M. Heusel, H. Ramsauer, T. Unterthiner, et al. : GANs trained by a two time-scale update rule converge to a local nash equilibrium. In Proceedings of the 31st International Conference on NIPS 6629–6640 (2017)
24. M. Bińkowski, D.J. Sutherland, M. Arbel, A. Gretton : Demystifying MMD GANs. CLR 2018 (2018)
25. A. Tsitsulin, M. Munkhoeva, D. Mottin, et al : The Shape of Data: Intrinsic Distance for Data Distributions, ICLR'2020 (2020)
26. G. Toderici, D. Vincent, N. Johnston, S. Hwang, D. Minnen, J. Shor, M. Covell : Full Resolution Image Compression with Recurrent Neural Networks. (2016)
27. A. Wang, T. Chi Chun Tam, H. Ming Poon et al. : NaviAirway: a bronchiole-sensitive deep learning-based airway segmentation pipeline for planning of navigation bronchoscopy. Transactions on Medical Imaging (2022)