Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients

Xin Tie[1,2], Sai-Kit Lam[1], Yong Zhang[3], Kar-Ho Lee[4], Kwok-Hung Au[4], Jing Cai[1]

1. *The Hong Kong Polytechnic University, Hong Kong SAR, China*
2. *Nanjing University, Nanjing, China*
3. *Xiamen University, Xiamen, China*
4. *Queen Elizabeth Hospital, Hong Kong SAR, China*

Corresponding Author:
Jing Cai, PhD
Department of Health Technology and Informatics
The Hong Kong Polytechnic University
Kowloon, Hong Kong
Tel: (852) 8645 0423
Email: jing.cai@polyu.edu.hk

Short Title: Pseudo-CT using multi-parametric MRI

Purpose: To develop and evaluate a novel method for pseudo-CT generation from multi-parametric MR images using multi-channel multi-path generative adversarial network (MCMP-GAN).

Methods: Pre- and post-contrast T1-weighted (T1-w), T2-weighted (T2-w) MRI, and treatment planning CT images of 32 nasopharyngeal carcinoma (NPC) patients were employed to train a pixel-to-pixel MCMP-GAN. The network was developed based on a 5-level Residual U-Net (ResUNet) with the channel-based independent feature extraction network to generate pseudo-CT images from multi-parametric MR images. The discriminator with 5 convolutional layers was added to distinguish between the real CT and pseudo-CT images, improving the non-linearity and prediction accuracy of the model. Eight-fold cross-validation was implemented to validate the proposed MCMP-GAN. The pseudo-CT images were evaluated against the corresponding planning CT images based on mean absolute error (MAE), peak signal-to-noise ratio (PSNR), Dice similarity coefficient (DSC) and Structural similarity index (SSIM). Similar comparisons were also performed against the multi-channel single-path GAN (MCSP-GAN), the single-channel single-path GAN (SCSP-GAN).

Results: It took approximately 20 hours to train the MCMP-GAN model on a Quadro P6000, and less than 10 seconds to generate all pseudo-CT images for the subjects in the test set. The average head MAE between pseudo-CT and planning CT was 75.7±14.6 Hounsfield Unit (HU) for MCMP-GAN, significantly (p-values<0.05) lower than that for MCSP-GAN (79.2±13.0 HU) and SCSP-GAN (85.8±14.3 HU). For bone only, the MCMP-GAN yielded a smaller mean MAE (194.6±38.9 HU) than MCSP-GAN (203.7±33.1 HU), SCSP-GAN (227.0±36.7 HU). The average PSNR of MCMP-GAN (29.1±1.6) was found higher than that of MCSP-GAN (28.8±1.2) and SCSP-GAN (28.2±1.3). In terms of metrics for image similarity, MCMP-GAN achieved the highest SSIM (0.92±0.02) but did not show significantly improved bone DSC results in comparison with MCSP-GAN.

Conclusions: We developed a novel multi-channel GAN approach for generating pseudo-CT from multi-parametric MR images. Our preliminary results in NPC patients showed that the MCMP-GAN method performed apparently superior to the UNet-GAN and SCSP-GAN, and slightly better than MCSP-GAN.

Keywords: deep learning, multi-parametric MRI, pseudo-CT, radiation therapy, nasopharyngeal carcinoma

Introduction

Magnetic Resonance Imaging (MRI)-only radiotherapy is an emerging technology in which all radiotherapy tasks are carried out using MRI as the sole imaging modality[1, 2]. MRI-only radiotherapy can decrease the number of scans, reduce overall cost[3] and minimize patient exposure to ionizing radiation. Furthermore, MRI offers excellent soft tissue contrast, improving tumor visualization as compared to computed tomography (CT) images[4, 5]. More and more evidences showed that the accurate delineation in MRI-guided radiotherapy could provide better results in the treatment planning, including improved dosimetry, in multiple cases of cancers[6]. One of the key challenges in MRI-only radiotherapy is that MR images do not contain information about tissue electron density which is crucial for radiation dose calculation. To overcome this challenge, MR images need to be converted to CT images for the purpose of radiation dose calculation, so-called “pseudo-CT”, or “synthetic-CT”. To date, a number of methods have been proposed for CT synthesis, which can be generally classified into three categories[2, 4, 7]: segmentation-based, atlas-based and learning-based methods.

The segmentation-based method[8-14] first classifies MR image voxels into a small number of bulk densities (often 3-4 tissue types), and then assigns corresponding CT values to each tissue type. In most cases, water equivalent and bony structures were segmented, while other types were dependent on the purpose and subjects. This method is straightforward, but with prominent disadvantages[4, 7]. For example, the ultra-short echo-time (UTE) MR sequence, which is widely used in segmentation-based methods, suffers from long acquisition time. Low signal-to-noise (SNR) ratio and partial volume effects can lead to bone segmentation errors[15]. Manual bone segmentation is impractical due to signal void of bone in conventional MRI.

In the atlas-based method[16-20], a database comprising of co-registered CT and MRI is first established. Then a new set of MR images is matched to the data atlas via deformable image registration[7]. Finally, the deformation is applied to the corresponding co-registered CT to generate the pseudo-CT. The accuracy of the atlas-based method is highly dependent on the registration quality in the MR/CT database[21]. To address it, Burgos et al.[19] proposed an iterative multi-atlas framework, combining structure-guided registration and image synthesis to build a high-quality database, which actually complicated the workflow.

The learning-based method directly builds relationship between CT- and MRI-based prior knowledge. Some groups[15, 22-28] employed conventional machine-learning methods, such as Gaussian Mixture Model (GMM), structure random forest (SRF), etc. Recently, deep learning methods[21, 29-33] have been exploited for pseudo-CT generation, showing superior performance to the atlas-based and conventional machine-learning methods[21]. For instance, Nie et al.[33] utilized fully convolutional neural network (FCN) as a generator for 3D pseudo-CT and added an adversarial network to produce realistic CT images in their work[29]. The adversarial network further improved the model in building the non-linear relationship between these two modalities, making the pseudo-CT images more realistic[30]. Emami et al.[30] trained a conditional generative adversarial network (cGAN) comprised of residual FCN as the generator and convolutional neural network (CNN) as the discriminator to address the issues of performance degradation and gradient vanishing in deeper network. Lei et al.[34] developed a dense CycleGAN-based model to produce pseudo-CT, making use of dense blocks and a novel distance loss function, which were employed to capture multi-scale information and resolve the blur and misclassification problems, respectively. In general, the deep learning-based methods achieved

91 better performance than the atlas-based methods with lower reconstruction errors[21] and dosimetric
92 errors[4].

93 Previous methods mostly utilize a single MRI type as input to generate pseudo-CT. However,
94 studies have shown that a single MRI type may be insufficient to accurately distinguish different tissue
95 types[11]. Methods of multi-parametric MR-to-CT conversion have also been demonstrated and are
96 typically handled using an early-fusion strategy[35], in which the concatenation layer stacks the multi-
97 parametric MR images. For instance, Maspero et al.[1] utilized multi-contrast Dixon-reconstructed MRI
98 as the input and cGAN as the training network for pelvic pseudo-CT generation. Leynes et al.[36] used
99 multi-parametric MRI patch input in 3D CT synthesis with three channels: proton density zero-echo-
100 time image, Dixon fractional fat and water images, respectively. This method is straightforward to
101 apply, but has limitations in handling modalities[37] whose complex relationships cannot be simply
102 modelled by the early fusion layer[35]. Recently, Chartsias et al.[38, 39] proposed a novel multi-input multi-
103 output model, which incorporated the modality-invariant latent representation for the retention of
104 modality specific features. The max-fusion strategy of the latent representations encoded from the
105 various inputs provided better synthetic results than those obtained from the unimodal models.

106 Inspired by Chartsias' work, in this study we developed a novel deep learning model with the late-
107 fusion network for better use of the multi-parametric MRI images to generate more realistic pseudo-
108 CT images. Our model is a multi-channel multi-path generative adversarial network, labeled as
109 MCMP-GAN. It was developed on the basis of a generative network, characterizing not only multi-
110 channel inputs, but also multi-path architecture. To investigate our model, especially with regard to
111 the effectiveness of the multi-path strategy, we compared MCMP-GAN to other models, including a
112 multi-channel model with the concatenation layer merging the input MRIs (i.e., multi-channel single-

path GAN, labeled as MCSP-GAN), and a single-channel single-path GAN model, labeled as SCSP-GAN. To our best knowledge, our work for the first time quantitatively investigated the impact of multi-modal inputs on image quality of pseudo-CT. The most common deep learning method to handle multi-parametric MRI thus far is to concatenate the MR images at the input, wherein each channel corresponds to each MR image volume. Although there exist some multi-input synthesis models[35, 38-40], they have not yet been used for pseudo-CT application.

Materials and Methods

Patient data

This study included 32 nasopharyngeal carcinoma (NPC) patients from Queen Elizabeth Hospital (QEH) of Hong Kong who had both MR and CT scans for radiotherapy treatment planning. Three MRI datasets, pre-contrast T1-weighted (T1-w) MRI, post-contrast T1-w MRI with fat-saturation, and T2-weighted (T2-w) MRI, were used as input images for the MCMP-GAN model. All MR images were acquired with proper immobilization in a 1.5T clinical MRI scanner (Avanto, Siemens, Germany). The T1-w MR images were acquired using the spin echo (SE) MR sequence with the following parameters: repetition time (TR): 562-739 ms; echo time (TE): 13-17 ms; matrix: 256-320; slice-thickness: 3.3-4.0 mm; voxel size 0.75-0.94 mm. The T2-w MR images were acquired using the short tau inversion recovery (STIR) MR sequence with the following parameters: TR: 7640 ms; TE: 97 ms; inversion time (TI): 165 ms; matrix: 320; slice-thickness: 4.0 mm; voxel size 0.75 mm. The CT images were performed on a Brilliance Big Bore (Philips, USA) scanner with the following parameters: tube current: mostly 264 mA, tube voltage: 120 kVp, slice thickness: 3 mm and pixel spacing: 1.0-1.2 mm.

MR and CT images were acquired within the same day. The MR/CT pairs were co-registered using the affine registration algorithm in MIM Maestro (MIM Software Inc., Beachwood, OH, USA). All MR and CT images were resampled to an isotropic voxel of 1.0×1.0×1.2 mm^3 and cropped to 240×192 before further preprocessing. A binary head mask excluding outer air was extracted from CT images via thresholding and Canny edge detection for each patient and was used in model training. All MR images were corrected for signal inhomogeneity using a N4 bias correction algorithm[41] and then normalized using a histogram-matching technique[42]. The standard intensity space was determined by the MRI fed to the standardization model. If new MR images were inputted, the model could map them to the same scale[42]. The pre-set parameters, such as cutoff values and landmark locations, were all set to the default values as in Github (https://github.com/loli/medpy). In CT images, regions outside the masks were set to -1000 HU.

Network architecture

Figure 1 shows the architecture of the proposed MCMP-GAN. The input layer includes three channels, corresponding to three input MR images respectively. The generative network was built based on the U-Net proposed by Ronneberger et al.[43], consisting of a contracting path and an expanding path. The contracting path is split into three training paths, wherein each channel has its own feature extraction network. These independent encoding paths were designed to separately extract the image characteristics from each input MRI dataset and to avoid the loss of unique features that otherwise would be merged in the low level. Despite the independent encoding paths for each input MRI dataset, the entire network was trained simultaneously. In the decoder, the outputs of each residual block are concatenated to the feature maps within the same depth level from the encoder via long skip

connections. The extra feature maps copied from each encoding path make it easier for the extending path to recover the image information which is lost during the down-sampling.

Furthermore, the skip connections rendered the network more flexible[21], i.e., the network could skip the coarse features from high level if the fine features were sufficient to generate high-quality images. Instead of the regular convolutional block, the residual convolutional block was used in the MCMP-GAN. The residual blocks prevented performance degradation and gradient vanishing when the neural network was very deep[44]. The identity maps, where 2D convolution with a kernel size of 1×1 was used to adjust the number of filters, added the block input to the output. Each residual block contained two convolutional layers with a kernel size of 3×3, both of which were batch normalized[46] and activated by ReLU. Unlike some UNet-like architectures, the max-pooling layers were replaced by the convolutional layers with strides of 2, which avoided the excessive loss of information, achieving a better performance, especially in the deep convolutional GAN (DCGAN)[47]. The structures of each encoding path were the same. While in the extending path, each residual block had a 3×3 kernel following a 5×5 kernel with a dilation rate of 2 which amplified the receptive field on the concatenated features. In the final layer, a 1×1 convolutional layer was used to project the feature maps to the corresponding CT images.

The detailed parameters and output size of each step are shown in **Table 1**. "×3" means the total number of feature extraction networks which were trained independently along each encoding path. Additionally, Dropout layers[48] were added as an option in the residual blocks to prevent overfitting and improve performance in the validation. The dropout ratio was set to the default value of 0.5.

The discriminator consists of four convolutional layers with a kernel size of 5×5 and strides of 2, followed by batch-normalization layers and 'LeakyReLU'[49] (alpha=0.2) activation layers (see details

in **Table 2**). The derivative of 'LeakyReLU' in the negative part is a small fraction, unlike 'ReLU' which is zero. The final layer is a 3×3 convolutional layer with only one filter. The output of the discriminator is the validity of the input CT images. The discriminator is real (*validity=1*) for planning CT and is fake (*validity=0*) for generated CT. The benefits of the adversarial network have been shown by Emami[30], Nie[29] and Ledig[50], which can be summarized as follows: (1) it prevents the generated images from blurring and preserve better details, especially for edge features; (2) the accuracy of pseudo-CT within bone regions is increased; and (3) the discriminator detects patch features in both real and fake images, mitigating mis-registration problem caused by the imperfect alignment between the multi-parametric MRI and CT.

Implementation details

The proposed model was implemented in *Keras* (https://github.com/fchollet/keras). The loss function was similar to that of least square GAN (LSGAN) which has been shown better than the cross-entropy loss function by providing better image quality and performing more consistently [51]. The objective function is defined as below:

The generator loss is

$$\min_G L(\mathrm{G}) = \frac{1}{2} E_{z\sim P_G(z)}[(\mathrm{D}(\mathrm{G}(\mathrm{z}))-1)^2] + \lambda E_{\mathrm{z}\sim \mathrm{P}_G(z),\mathrm{x}\sim \mathrm{P}_{data}} \| \mathrm{G}(\mathrm{z}) - x \|_1 \quad (1)$$

and the discriminator loss is

$$\min_D L(\mathrm{D}) = \frac{1}{2} E_{x\sim P_{data}}[(\mathrm{D}(\mathrm{x})-1)^2] + \frac{1}{2} E_{\mathrm{z}\sim P_G(\mathrm{z})}[(\mathrm{D}(\mathrm{G}(\mathrm{z}))-0)^2] \quad (2)$$

where G is the generator, D is the discriminator, and z is the input of the generative network, sampled from the probability distribution of the MR data (P_G). G(z) is the generated output, and x is the reference output of the G, sampled from the probability distribution of the CT data (P_{data}). LS loss

prevents blurring of the images, but may lead to sharped images and introduce artifacts[52, 53]. L_1 reconstruction loss helps to produce more realistic images with less artifacts. The weighting factor (λ) measuring the significance of reconstruction error was set to 10.

The optimization used in our model was Adam[54] with the learning rate of 2e-4 and momentum term (β_1) of 0.5. It stabilizes training in the learning process[47]. The stochastic optimization method randomly selects the subsets from the training data and updates the parameters, so-called mini-batch. Batch size of 5 was used for training in our study. The weight initiators were randomly sampled from a truncated normal distribution[55] centered at 0 with the standard deviation of $\sqrt{2/(\mathrm{fan_{in}} + \mathrm{fan_{out}})}$ ($\mathrm{fan_{in}}$ and $\mathrm{fan_{out}}$ are the number of input units and output units in the weight tensor, respectively). The initial biases were set to "zero". To avoid overfitting, we used early stopping at the end of the learning process. Before training, data augmentation was performed artificially. The samples from the training set were randomly selected to flip horizontally and vertically, or rotate in some certain angles. Eight-fold cross validation was implemented, where each group had 4 subjects. At each validation fold, seven groups (28 patients) were used for training the model and the remaining group (4 patients) was used for validation.

Evaluation metrics

Performance of MCMP-GAN was evaluated by comparing the generated pseudo-CT images against the planning CT images (as references) to determine the mean absolute error (MAE), peak signal-to-noise ratio (PSNR), Dice similarity coefficient (DSC), and structure similarity index (SSIM). MAE is defined as:

$$MAE = \frac{1}{a}\sum_{A} | \mathrm{CT}_{real} - CT_{pseudo} | \quad (3)$$

where *a* is the total number of voxels within the head region that was delineated previously. The lower the MAE, the higher the accuracy of the pseudo-CT images. MAE was measured for the entire head region, and for the bony structure only. For the latter, *a* is the total number of voxels of bony structure which was segmented using a threshold of 200 HU on the planning CT images. The PSNR is defined as:

$$PSNR = 10 \cdot \log_{10}(\frac{R^2}{MSE}) \tag{4}$$

where *MSE* is the mean square error, defined as $\mathrm{MSE} = \frac{\sum_{A_x,A_y} (\mathrm{CT_{real}} - \mathrm{CT_{pseudo}})^2}{A_x \cdot A_y}$, in which A_x and A_y are the row and column of the image respectively; R is the maximal fluctuation of the input image. The larger the PSNR, the lower the reconstruction error. DSC and SSIM are commonly used metric for similarity measures and their calculations were performed as usual. Their expressions are defined below:

$$DSC = \frac{2 \times \mathrm{bone_{real}} \cap \mathrm{bone_{pseudo}}}{|\mathrm{bone_{real}}| \cdot |\mathrm{bone_{pseudo}}|} \tag{5}$$

$$SSIM = \frac{(2\mu_x\mu_y + \mathrm{C_1})(2\delta_{xy} + \mathrm{C_2})}{(\mu_x^2 + \mu_y^2 + \mathrm{C_1})(\delta_x^2 + \delta_y^2 + \mathrm{C_2})} \tag{6}$$

Where bone with subscript represents the bone segmentation maps with threshold 200 HU extracted from real CT and pseudo-CT images respectively. By default, $\mathrm{C_1}$ and $\mathrm{C_2}$ are expressed as $\mathrm{C_1} = (0.01 \cdot \mathrm{R})^2$, $\mathrm{C_1} = (0.03 \cdot \mathrm{R})^2$.

Comparison models

To evaluate the MCMP-GAN model, especially to investigate the impact of multi-channel input and independent feature extraction network in the contracting path, we also implemented a SCSP-GAN model and a MCSP-GAN model for comparison. The SCSP-GAN and MCSP-GAN have the optimization method and training strategy as those of MCMP-GAN, with only slight differences in

architecture as detailed below. The SCSP-GAN was comprised of the single channel residual U-Net and 5-layer CNN. The post-contrast T1-w MR images were used as the single input to the SCSP-GAN network; and unlike the generator of MCMP-GAN, the single extraction network was utilized in the contracting path to capture the image characteristics from high to low resolution. The discriminator was the same as that of MCMP-GAN. Both Maspero[1] and Emami[30] developed the single channel GAN, which outperformed the regular CNN methods. Here, we borrowed their ideas (the elaborate descriptions were shown in Isola et al.[53]), constructed a model with the similar architecture, but incorporated the residual blocks, identical to what we did in the MCMP-GAN model.

The MCSP-GAN model was built based on the architecture of SCSP-GAN. The concatenation layer was added between the input layer and the first residual block to stack the input multi-parametric MRIs along the channel. In the generative network of MCSP-GAN, the images were fused at the input, which meant the information from each type of MR cannot be disentangled in the deeper layers.

Results

Pseudo-CT images of MCMP-GAN

Approximately 7000 samples (after data augmentation) were used in the model training. With a mini-batch size of 5, it took about 100 epochs for the model to converge, resulting in a model training time of 20 hours on a Quadro P6000 workstation. There were approximately 350 images in the testing dataset. Once the model was trained, the pseudo-CT images were generated within a few seconds. **Table 3** summarizes the measurements for all patients. The average MAE was 75.7±14.6 HU and the mean PSNR was 29.1±1.6 for the entire head region. For bony structure only, the average MAE was

194.6±38.9 HU, indicating that the prediction accuracy for the bone is still challenging. As respect to image similarity metrics, MCMP-GAN achieved 0.86±0.03 for bone DSC and 0.92±0.02 for SSIM.

Figure 2 shows example pseudo-CT images generated using MCMP-GAN, along with the multi-parametric input MR images and the reference planning CT images, as well as the difference maps between the pseudo-CT and the reference planning CT. It can be seen that the difference between the reference planning CT and pseudo-CT was minimal in the soft tissues, but apparent in the bone regions, especially at the edges of the bony structure. Large differences were also observed at the interface between air and bone, shown in the regions of maxillary sinus, which were highlighted in the colored boxes in the fourth row. These large differences were presumably caused by the following reasons: (1) CT values in regions between two abut tissue types are discrete, not continuous. Neural network may have difficulty to build localized discrete function to handle this situation. As a result, large gradient changes may cause errors in these regions. (2) There were residual registration errors between MRI and CT images, which caused wrong learning models in the imperfectly aligned regions.

Model comparison

The average MAE and PSNR were 75.7±14.6 HU and 29.1±1.6 for MCMP-GAN, as compared to 85.8±14.3 HU and 28.2±1.3 for SCSP-GAN. **Table 4** summarizes the average evaluating metrics for all subjects, along with the p-values comparing MCMP-GAN and other networks. It can be seen that MCMP-GAN performed slightly better than MCSP-GAN and significantly better than SCSP-GAN, yielding the lowest overall and bone MAE and largest PSNR.

Figure 3 shows the representative results obtained from MCMP-GAN, and SCSP-GAN respectively. The blurs and large errors occurred in the areas with complex details in the SCSP-GAN,

but decreased in the MCMP-GAN. At the interfaces between bone and tissues, the errors of the MCMP-GAN results were slightly smaller than those of the UNet-GAN and SCSP-GAN results, while in the air cavities, the MCMP-GAN performed apparently better than the other methods. For instance, the ethmoidal sinuses contained fine details, which was a great test for the proposed model and others. In the first row, only the pseudo-CT generated via MCMP-GAN preserved more details, similar to the real CT. However, the pseudo-CT images obtained by SCSP-GAN lost some details and were blurry. Another example was that the obvious errors, highlighted in the red circles in the third row, were only found in the pseudo-CTs produced by the SCSP-GAN, but were not present in those generated by MCMP-GAN. At the interfaces between the maxillary sinuses and surrounding bony structures, the pseudo-CT from MCMP-GAN succeeded to depict the borders, but the pseudo-CTs obtained via SCSP-GAN failed, as shown in the red and green boxes in the second row. The yellow boxes (2nd row) showed the reconstruction of the sphenoid sinus: only the pseudo-CT generated by MCMP-GAN held the comparatively complete information.

The MCSP-GAN yielded the average MAE of 79.2±13.0 HU and mean PSNR of 28.8±1.2 across the entire FOV of head. The quantitative comparison showed that MCMP-GAN performed slightly but significantly better than MCSP-GAN with lower MAE (p-values<0.05), higher PSNR (p-values<0.05) and higher SSIM (p-values<0.05). For bony structure, the MAE of MCMP-GAN was also significantly smaller than that of MCSP-GAN (p-value<0.05). However, the bone DSC didn't show an improved result in MCMP-GAN, which was probably due to the rough bone segmentation maps extracted from the pseudo- and real CT.

Figure 4 shows the visual comparison of pseudo-CT images obtained via MCMP-GAN and MCSP-GAN, and zooms in the marked details below the CT images. The enlarged regions in the first

309 row illustrated that the pseudo-CT generated by MCMP-GAN was more similar to the real CT in the
310 maxillary sinuses, while the pseudo-CT generated by MCSP-GAN showed large errors within and at
311 the border of the sinuses, as shown in the colored boxes. The clear blurs and large errors in the petrous
312 temporal bone, enhanced in the red boxes in the second row, occurred in the pseudo-CT produced by
313 MCSP-GAN, but did not appear noticeably in the MCMP-GAN output.
314

315 ## Discussion

316 Pseudo-CT generation is a key component in MR-only radiotherapy treatment planning, and has
317 been proven a challenging task due to various reasons including, but not limited to, low signal of bony
318 structure and no signal of air cavity in MR images, MR image distortion, image misalignment, etc. In
319 this study we demonstrated a novel deep learning-based MCMP-GAN model for generating pseudo-
320 CTs from multi-parametric MR images. This is the first work focusing on the impact of the multi-
321 channel input on the quality of pseudo-CT images, as well as on using independent feature extraction
322 network to produce pseudo-CT images. Our results showed that overall MCMP-GAN outperformed
323 other comparing methods: MCSP-GAN, SCSP-GAN, and UNet-GAN.

324 Comparison between MCMP-GAN and MCSP-GAN showed that MCMP-GAN made better use
325 of multi-parametric MR images and had higher accuracy in pseudo-CT. Instead of stacking the multi-
326 parametric MR images at the input, we trained the independent feature extraction network for each
327 encoding path in the contracting path; while in the extending path, the feature maps were fused with
328 those in the contracting path in the same depth level, so-called feature fusion. **Figure 5** shows the
329 intermediate convent outputs (output of the level 4 at the encoder). The feature maps from each
330 encoding path were clearly different. Based on the similarity of the features extracted from each type

of MR, they can be divided into two groups: shared features and independent features. The shared features represented the similar images characteristics, which were probably more beneficial to the CT synthesis. At the same time, the independent features were still retained, increasing the total number of feature maps at each level and further helping recover the spatial details of the images during the upsampling. By comparison, if we concatenated MRI at the low-level stage, the independent features might be lost at the higher level. Multi-parametric MR images included unique and complementary characteristics. Stacking them like handling RGB images decreased the utilization of each weighted MRI. Another benefit of our network was a more flexible architecture which could handle the data-deficiency issue among the multi-parametric MRIs. Assuming that the cases of T1-w MRI and CT were quite abundant while those of T2-w MRI and CT were not as rich. In the MCSP-GAN, lots of T1-w images could not be used in the training because they did not have the corresponding T2-w images. However, in the MCMP-GAN, these T1-w images could serve as the samples in the pre-training stage. In the training stage, the pretrained weights in the encoder can be transferred to the T1-w encoding path.

The MCMP-GAN model yielded an overall MAE of 75.7±14.6 HU, lower than those reported by Nie[29] (92.5±13.9HU) and by Emami[30] (89.30±10.25HU). Nie et al.[29] extended the generative model to three dimension, which required more GPU memory and computation time. Emami et al.[30] incorporated ResNet (residual network) into FCN, and achieved exciting results in GAN compared to CNN methods. The loss functions in our work and theirs were both the combination of the least square loss and reconstruction error. Emami used FCN without long skip connections between the feature maps in the contracting path and those in the extending path. Instead of the regular ResNet, we constructed the ResUNet, in which the copy layers were added to help recover the spatial information.

Not only that, the separate feature extraction network at the encoder further increased the number of feature maps in the same depth level, improving the utilization of multi-channel inputs.

Another progress in pseudo-CT generation was CycleGAN model for MR-to-CT translation using unpaired data[32]. They achieved low MAE of 73.7±2.3 HU and high PSNR of 32.3±0.7, and demonstrated that the model trained using unpaired data outperformed the model trained using paired data. Considering that this method avoided the misalignment between MR and CT, the highly accurate results with less artifacts and blurs were understandable. However, in Jin's work[5], they pointed out that the images obtained from CycleGAN using unpaired data had poor anatomical definitions compared with those generated from the model trained with paired data. Additionally, the voxel-wise loss for paired data played a more significant role in providing the realistic images with less blurs. Unquestionably, Wolterink et al. [32] presented very exciting results, but it was still hard to prove whether the model using unaligned data was really superior to the GAN model on paired data. Extending the CycleGAN model to multi-channel CycleGAN is undoubtedly an interesting topic, which can reduce the need for paired data and realize the many-to-one or many-to-many mappings. Almahairi et al. [56] proposed the Augmented CycleGAN to handle it and examined its feasibility on several image datasets. In future work, we will try to introduce the Augmented CycleGAN to the MR-to-CT translation task, in hope of further improving the accuracy of generated CT images and strengthening its availability in clinical work.

In multi-modal segmentation, some papers presented novel networks for the late-fusion approaches, which also give us some new ideas for future work. In Nie's [40] late-fusion FCN, each modality image had a separate network to capture features, which were fused in the high-level layers for the final infant brain segmentation. Dolz et al.[35] incorporated the inception modules and hyper-

dense connectivity into the multi-path U-Net to better account for the complex and non-linear relationship among different modalities in ischemic stroke lesion segmentation. In some cases, a huge network with the complicated architecture and so many training parameters may cause overfitting and leave heavy burden on the GPU. In the future, we will consider adding new modules into the network to improve the complexity, and at the same time avoid overfitting.

Another point that will be explored in future study is whether the feature extraction network in the contracting path can improve the robustness of the model. In **Figure 3**, the pseudo-CT images generated by SCSP-GAN suffer from serious errors in certain regions, but the MCMP-GAN performs apparently better in reducing these errors, which indirectly proves the improved robustness of the network. To better validate it, our preliminary idea is to randomly add some noise in one of the MR weighted images and examine the quality of pseudo-CT. This is straightforward but simplistic. In addition, whether the proposed model can efficiently reduce the impact of the misalignment errors between MR and CT, and the intra-registration errors among MRIs are still unknown. Discussion about these interesting topics will be part of our future study.

GAN models often suffer the gradient vanishing problem during the training process, which may influence the convergence of the network. The optimization method mentioned above was similar to that utilized in deep convolutional GAN[47], but the instability and vanishing gradient were still not well resolved. One of the potential applications of pseudo-CT is the MR-based treatment planning which can be completed without extra scan of CT. In the future, the modification and dosimetric analysis of our model will be further discussed.

There are limitations in our work. First, our method was based on 2D MR-to-CT translation, not 3D. Considering more training parameters compared to the single-channel model, it’s anticipated that

in 3D MR-to-CT, balancing the computational memory, network architecture and accuracy of results will be the primary task. Second, the size of the training samples may not be large enough. One of the superiorities of the U-Net was its ability to handle a small-size dataset and utilize data augmentation to improve efficiency of data exploitation[43]. Third, there were residual registration errors between multi-parametric MR images which may have contributed to the discrepancies between real CT and pseudo CT. Our results implied that the overall MAE may be significantly affected by the discrepancy in bony regions as a result of image misalignment. It can be reasonably expected that the performance of MPMC-GAN could be even better if the registration errors can be reduced by using more sophisticated deformable image registration algorithms, or by using simultaneous multi-parametric MRI techniques such as magnetic resonance fingerprinting (MRF). In 2013, Ma[57] first introduced MRF that permitted the quantification of the tissue properties, such as T1 relaxation time, T2 relaxation time, and proton density, in a time-efficient acquisition. The signal evolution curves obtained from certain MR sequence were matched to the best corresponding MRF dictionary entry and the highly accurate quantitative maps were generated[58]. Another technical breakthrough was MAGiC[59] (MAGnetic resonance image Compilation) which allowed the acquisition of multi-contrast images in a single scan, including T1-w, T2-w, PD-w and some contrasts that would not be generated in conventional MRI.

Conclusion

In this work, we developed and evaluated a novel deep learning-based MCMP-GAN model for generating pseudo-CT images using multi-parametric MR images as the inputs. The preliminary results showed that the proposed MCMP-GAN model overall performed better than MCSP-GAN and SCSP-GAN.

419

420 **ACKNOWLEDGEMENTS**

421 This work is partly supported by funding NIH R01CA226899, GRF 151021/18M and GRF
422 151022/19M.

423

Tables

Figure Legends

Figure 1: The architecture of the proposed MCMP-GAN model.

Figure 2: Axial, coronal and sagittal view of the representative pseudo-CT images. Each one is accompanied with the real CT, corresponding MRIs and difference maps. The image types that each column represents have been indicated at the bottom of the figure.

Figure 3: Comparison of MCMP-GAN and SCSP-GAN in representative patients.

Figure 4: Comparison of MCMP-GAN and MCSP-GAN in representative patients.

Figure 5: The impact of independent feature extraction in the encoder. Each column corresponds to the intermediate convent outputs of one channel. From left to right: (a) pre-contrast T1-w, (b) post-contrast T1-w, and (c) T2-w.

Table 4: Detailed training parameters of the generative network of MCMP-GAN.

	Level	Conv Layer	Filter	Stride	Padding	Output
Input	Level 0	The encoding network is trained independently in each channel				240×192×1 (×3)
Encoding	Level 1	Conv1_1	3×3 – 3×3 / 32 (×3)	(1,1) – (1,1)	same	240×192×32 (×3)
		Conv1_2	3×3 – 3×3 / 32 (×3)	(1,1) – (1,1)	same	240×192×32 (×3)
	Level 2	Conv2_1	3×3 – 3×3 / 64 (×3)	(2,2) – (1,1)	same	120×96×64 (×3)
		Conv2_2	3×3 – 3×3 / 64 (×3)	(1,1) – (1,1)	same	120×96×64 (×3)
	Level 3	Conv3_1	3×3 – 3×3 / 128 (×3)	(2,2) – (1,1)	same	60×48×128 (×3)
		Conv3_2	3×3 – 3×3 / 128 (×3)	(1,1) – (1,1)	same	60×48×128 (×3)
	Level 4	Conv4_1	3×3 – 3×3 / 256 (×3)	(2,2) – (1,1)	same	30×24×256 (×3)
		Conv4_2	3×3 – 3×3 / 256 (×3)	(1,1) – (1,1)	same	30×24×256 (×3)
		Conv5_1	3×3 – 3×3 / 512 (×3)	(2,2) - (1,1)	same	15×12×512 (×3)
	Level 5	**Concatenate**				15×12×1536
Decoding		Conv5_2	5×5 – 3×3 / 512	(1,1) – (1,1)	same	15×12×512
	Level 4	Conv6	5×5 – 3×3 / 512	(1,1) – (1,1)	same	30×24×512
	Level 3	Conv7	5×5 – 3×3 / 256	(1,1) – (1,1)	same	60×48×256
	Level 2	Conv8	5×5 – 3×3 / 128	(1,1) – (1,1)	same	120×96×128
	Level 1	Conv9	5×5 – 3×3 / 64	(1,1) – (1,1)	same	240×192×64
Output		Conv10	1×1 / 1			240×192×1

442 Table 5. Detailed training parameters of the adversarial network of MCMP-GAN.

443

	Level	Conv Layer	Filter	Stride	Padding	Activation	Output
Input	Level 0	Concatenate the input MRI (the label) with the generated CT					240×192×2
Encoding	Level 1	Conv1	5×5 / 64	(2,2)	same	LeakyReLU	120×96×64
	Level 2	Conv2	5×5 / 128	(2,2)	same	LeakyReLU	60×48×128
	Level 3	Conv3	5×5 / 256	(2,2)	same	LeakyReLU	30×24×256
	Level 4	Conv4	5×5 / 512	(2,2)	same	LeakyReLU	15×12×512
Output		Conv5	3×3 / 1	(1,1)	same		15×12×1

Table 6: Summery of evaluating metrics for each subject.

Patient	Head MAE (HU)	Bone MAE (HU)	PSNR	Bone DSC	SSIM
Patient 01	74.7	176.2	29.4	0.88	0.92
Patient 02	67.5	189.1	29.3	0.87	0.92
Patient 03	73.7	192.7	29.2	0.87	0.91
Patient 04	73.5	209.0	29.1	0.86	0.91
Patient 05	81.0	221.5	28.5	0.83	0.90
Patient 06	85.7	225.9	28.2	0.84	0.89
Patient 07	67.6	177.8	30.1	0.88	0.93
Patient 08	67.2	163.2	30.0	0.87	0.93
Patient 09	76.1	212.1	28.7	0.82	0.92
Patient 10	91.3	242.9	27.5	0.83	0.92
Patient 11	88.4	205.5	28.3	0.82	0.89
Patient 12	80.9	187.8	28.6	0.88	0.92
Patient 13	46.1	116.9	32.9	0.89	0.96
Patient 14	72.7	176.9	29.5	0.87	0.92
Patient 15	69.2	167.8	29.3	0.85	0.93
Patient 16	81.5	202.5	28.4	0.85	0.91
Patient 17	83.7	218.8	27.8	0.86	0.91
Patient 18	81.1	187.4	28.1	0.84	0.91
Patient 19	84.5	268.4	27.8	0.80	0.90
Patient 20	73.8	177.8	29.4	0.86	0.92
Patient 21	85.0	246.5	28.0	0.82	0.91
Patient 22	46.1	110.5	33.7	0.90	0.96
Patient 23	82.8	218.2	28.1	0.85	0.91
Patient 24	88.5	215.9	27.6	0.87	0.91
Patient 25	82.3	221.8	28.2	0.82	0.90
Patient 26	77.1	164.8	29.5	0.87	0.92
Patient 27	89.0	233.8	27.5	0.86	0.90
Patient 28	71.0	193.9	29.3	0.85	0.92
Patient 29	69.5	160.1	30.0	0.88	0.93
Patient 30	64.1	166.8	30.9	0.87	0.92
Patient 31	84.8	194.9	28.5	0.85	0.90
Patient 32	63.3	179.9	30.1	0.88	0.93
Mean ± std	75.7±14.6	194.6±38.9	29.1±1.6	0.86±0.03	0.92±0.02

449 Table 4: Summary of evaluating metrics for all subjects and for comparing models.

450

	Head MAE (HU)	Bone MAE (HU)	PSNR	DICE	SSIM
MCMP-GAN	75.7±14.6	194.6±38.9	29.1±1.6	0.86±0.03	0.92±0.02
MCSP-GAN	79.2±13.0	203.7±33.0	28.8±1.2	0.85±0.04	0.91±0.02
p-value	<0.05	<0.05	<0.05	0.07	<0.05
SCSP-GAN	88.6±14.3	230.1±36.7	27.9±1.3	0.83±0.03	0.89±0.02
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

451

452

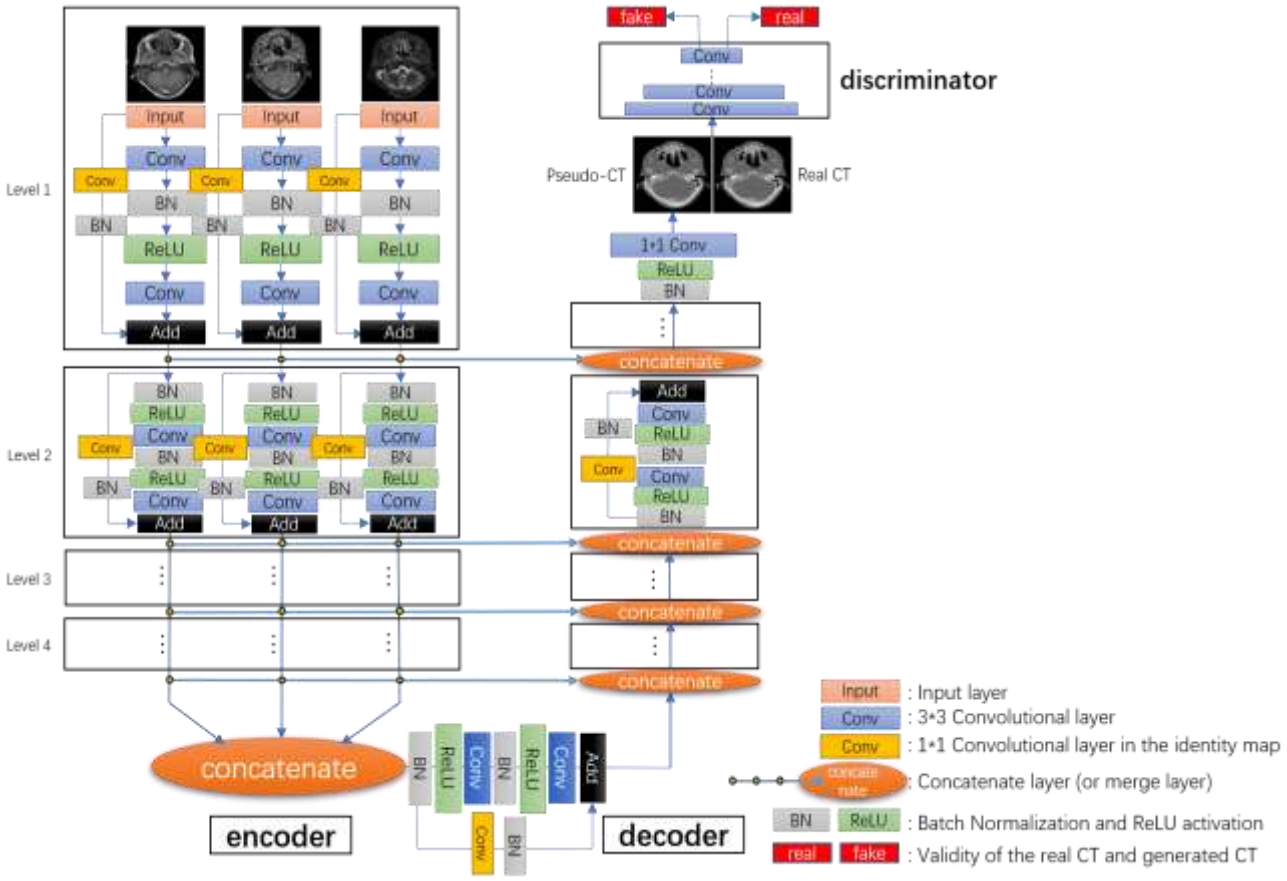


Figure 2: The architecture of the proposed MCMP-GAN model.

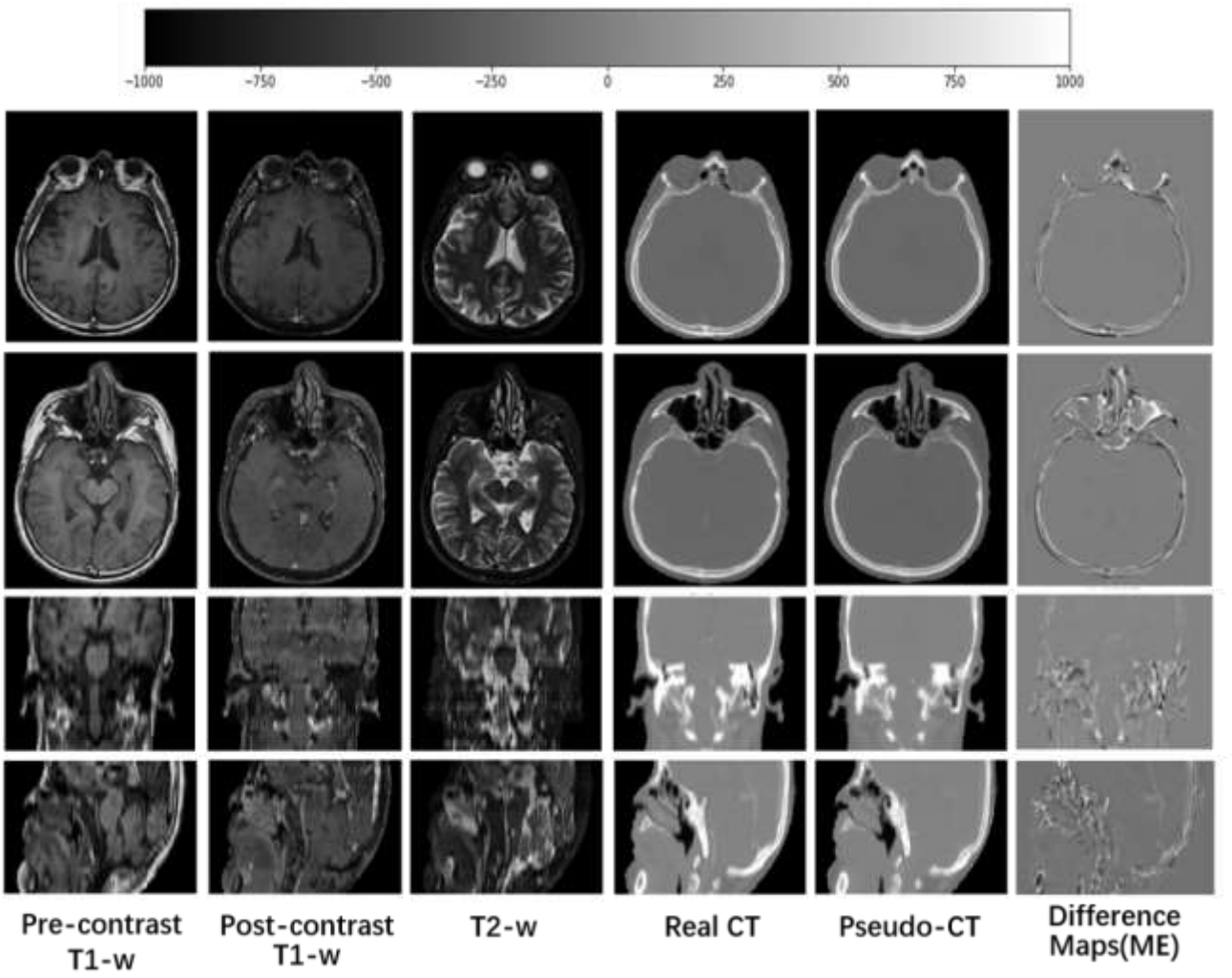


Figure 2: Axial, coronal and sagittal view of the representative pseudo-CT images. Each one is accompanied with the real CT, corresponding MRIs and difference maps. The image types that each column represents have been indicated at the bottom of the figure.

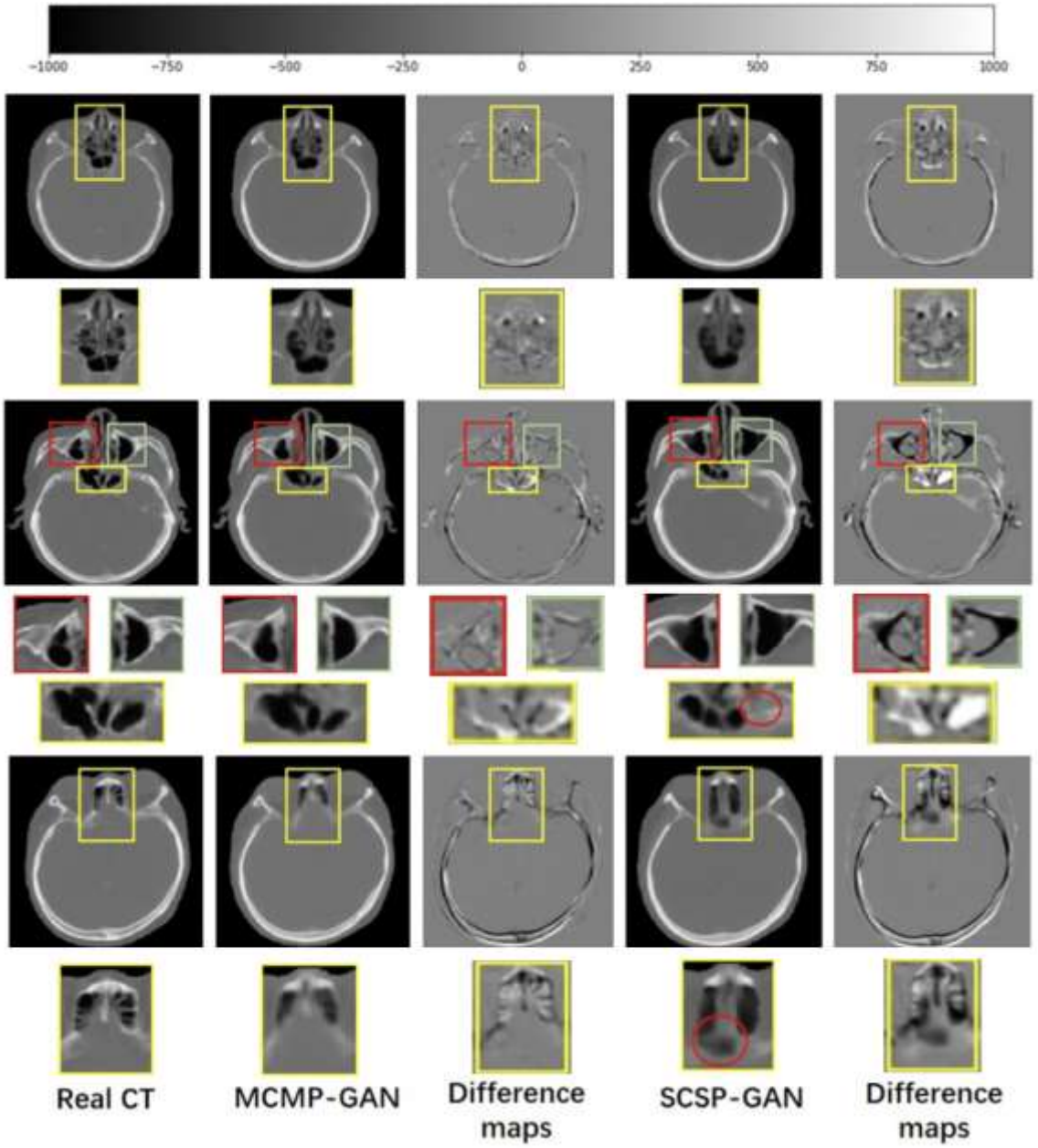
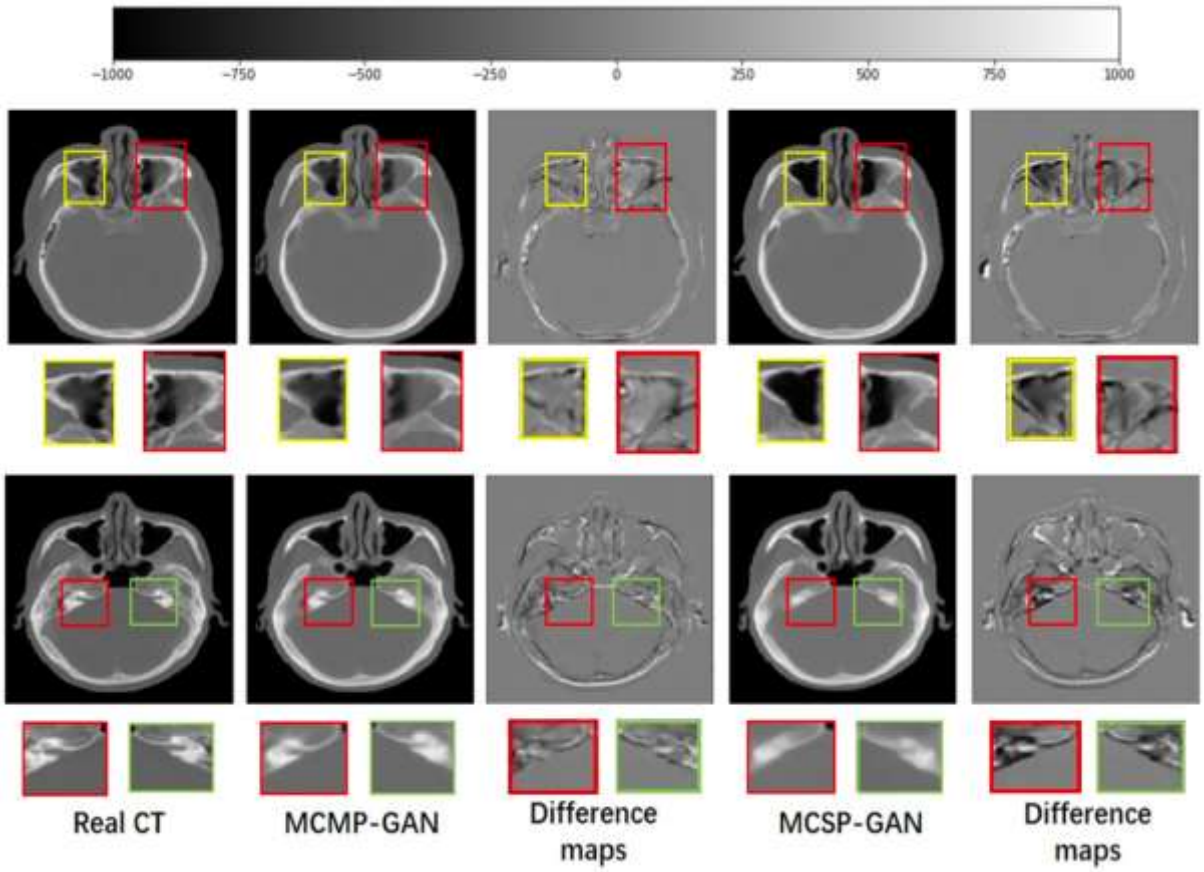


Figure 3: Comparison of MCMP-GAN and SCSP-GAN in representative patients.



461

462 Figure 4: Comparison of MCMP-GAN and MCSP-GAN in representative patients.

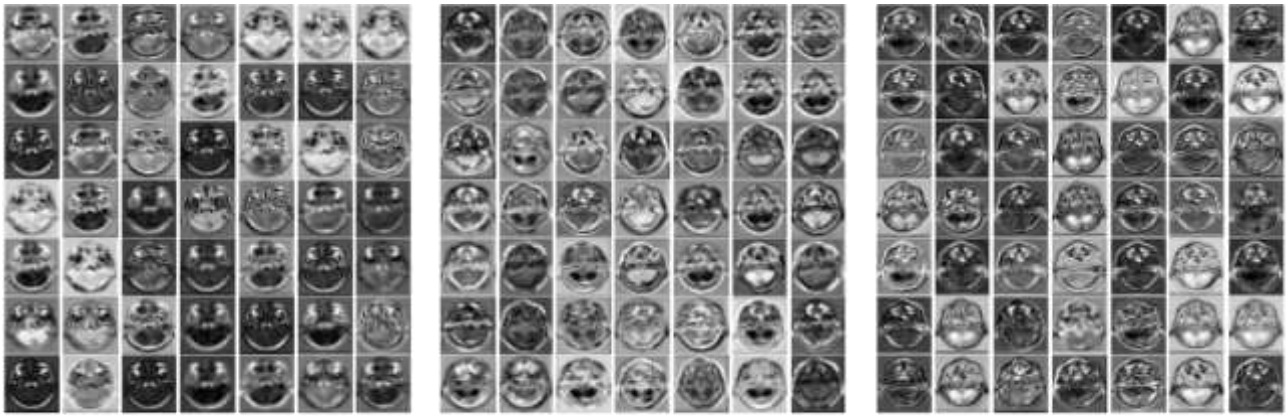


Figure 5: The impact of independent feature extraction in the encoder. Each column corresponds to the intermediate convent outputs of one channel. From left to right: (a) pre-contrast T1-w, (b) post-contrast T1-w, and (c) T2-w.

Reference

[1] Maspero M., Savenije M., Dinkla A. et al., Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy, *Phys. Med. Biol.* 2018; 63(18): 185001.

[2] Edmund J. M. and Nyholm T., A review of substitute CT generation for MRI-only radiation therapy, *Radiat. Oncol.* 2017; 12(1): 12-28.

[3] Devic S., MRI simulation for radiotherapy treatment planning, *Med. Phys*. 2012; 39(11): 6701-6711.

[4] Arabi H., Dowling J. A., Burgos N. et al., Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region, *Med. Phys.* 2018; 45(11): 5218-5233..

[5] Jin C. B., Kim H., Jung W. et al., Deep CT to MR Synthesis using Paired and Unpaired Data; 2018. arXiv preprint arXiv:1805.10790.

[6] Owrangi A. M., Greer P. B., Glide-Hurst C. K., MRI-only treatment planning: benefits and challenges, *Phys. Med. Biol.* 2018; 63(5): 05TR01.

[7] Johnstone E., Wyatt J. J., Henry A. M. et al., Systematic Review of Synthetic Computed Tomography Generation Methodologies for Use in Magnetic Resonance Imaging-Only Radiation Therapy, *Int. J. Radiat. Oncol. Biol. Phys.* 2018; 100(1): 199-217.

[8] Chin A. L., Lin A., Anamalayil S., Teo B. K., Feasibility and limitations of bulk density assignment in MRI for head and neck IMRT treatment planning, *J. Appl. Clin. Med. Phys.* 2014; 15(5): 100-111.

[9] Berker Y., Franke J., Salomon A. et al., MRI-Based Attenuation Correction for Hybrid PET/MRI Systems: A 4-Class Tissue Segmentation Technique Using a Combined Ultrashort-Echo-Time/Dixon MRI Sequence, *J. Nucl. Med.* 2012; 53(5): 796-804.

[10] Catana C., Kouwe A. V., Benner T. et al., Toward Implementing an MRI-Based PET Attenuation-Correction Method for Neurologic Studies on the MR-PET Brain Prototype, *J. Nucl. Med.* 2010; 51(9): 1431-1438.

[11] Hsu S. H., Cao Y., Huang K. et al., Invesigation of a method for generation synethetic CT models from MRI scans of the head and neck for radiation therapy, *Phys. Med. Biol.*, 2013; 58(23): 8419-8435.

[12] Keereman V., Fierens Y., Broux T. et al., MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences, *J. Nucl. Med.* 2010; 51(5): 812-818.

[13] A. Martinez-Moller, Souvatzoglou M., Delso G. et al., Tissue Classification as a Potential Approach for Attenuation Correction in Whole-Body PET/MRI: Evaluation with PET/CT Data, *J. Nucl. Med.* 2009; 50(4): 520-526.

[14] Edmund J. M., Kjer H. M., Van Leemput K. et al., A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times, *Phys. Med. Biol.* 2014; 59(23), 7501-7519.

[15] Rank C. M., Hunemohr N., Nagel A. M. et al., MRI-based simulation of treatment plans for ion radiotherapy in the brain region, *Radiat. Oncol.* 2013; 109(3): 414-418.

[16] Sjolund J., Forsberg D., Andersson M., and Knutsson H., Generating patient specific pseudo-CT of the head from MR using atlas-based regression, *Phys. Med. Biol.* 2015; 60(2): 825-839.

[17] Dowling J. A., Sun J., Pichler P. et al., Automatic Substitute Computed Tomography

Generation and Contouring for Magnetic Resonance Imaging (MRI)-Alone External Beam Radiation Therapy From Standard MRI Sequences, *Int. J. Radiat. Oncol. Biol. Phys.* 2015; 93(5): 1144-1153.

[18] Burgos N., Guerreiro F., McClelland J. et al., Iterative framework for the joint segmentation and CT synthesis of MR images: application to MRI-only radiotherapy treatment planning, *Phys. Med. Biol.* 2017; 62(11): 4237-4253.

[19] Arabi H., Koutsouvelis N., Rouzaud M., Miralbell R., Zaidi H. et al., Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET-MRI-guided radiotherapy treatment planning, *Phys. Med. Biol.* 2016; 61(17): 6531-6552.

[20] Ahunbay E. E., Thapa R., Chen X., Paulson E., Li X. A., A technique to rapidly generate synthetic CT for MRI-guided online adaptive replanning: an exploratory study, *Int. J. Radiat. Oncol. Biol. Phys.* 2019; 103(5): 1261-1270.

[21] Han X., MR-based synthetic CT generation using a deep convolutional neural network method, *Med. Phys.* 2017; 44(4): 1408-1419.

[22] Korhonen J., Kapanen M., Tenhunen M. et al., A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer, *Med. Phys*. 2014; 41(1): 011704.

[23] Kapanen M. and Tenhunen M., T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning, *Acta. Oncol.* 2013; 52(3): 612-618.

[24] Johansson A., Karlsson M., Nyholm T., CT substitute derived from MRI sequences with ultrashort echo time, *Med. Phys*. 2011; 38(5): 2708-2714.

[25] Johansson A., Garpebring A., Karlsson M. et al., Improved quality of computed tomography substitute derived from magnetic resonance (MR) data by incorporation of spatial information--potential application for MR-only radiotherapy and attenuation correction in positron emission tomography, *Acta. Oncol.* 2013; 52(7): 1369-1373.

[26] Johansson A., Karlsson M., Yu J., Asklund T., Nyholm T., Voxel-wise uncertainty in CT substitute derived from MRI, *Med. Phys.*, 2012; 39(6): 3283-3290.

[27] Rank C. M., Tremmel C., Hünemohr N. et al., MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach, *Radiat. Oncol.* 2013; 8(51).

[28] Huynh T., Gao Y., Kang J. et al., Estimating CT Image From MRI Data Using Structured Random Forest and Auto-Context Model, *IEEE Trans. Med. Imag.* 2016; 35(1): 174-183.

[29] Nie D., Trullo R., Petitjean C., Ruan S., D. Shen, Medical Image Synthesis with Context-Aware Generative Adversarial Networks; 2016. arXiv preprint arXiv:1612.05362.

[30] Emami H., Dong M., S. P. Nejad-Davarani, C. K. Glide-Hurst, Generating synthetic CTs from magnetic resonance images using generative adversarial networks, *Med. Phys.* 2018; 45(8): 3627-3636.

[31] Dinkla A. M., Wolterink J. M., Maspero M. et al., MR-Only Brain Radiation Therapy: Dosimetric Evaluation of Synthetic CTs Generated by a Dilated Convolutional Neural Network, *Int. J. Radiat. Oncol. Biol. Phys.* 2018; 102(4): 801-812.

[32] Wolterink J. M., Dinkla A. M., Savenije M. et al., Deep MR to CT Synthesis using Unpaired Data; 2017. arXiv preprint arXiv: 1708.01155.

[33] Nie D., Cao X., Gao Y., Wang L., D. Shen, Estimating CT Image from MRI Data Using 3D Fully Convolutional Networks, https://link.springer.com/chapter/10.1007%2F978-3-319-46976-8_18. Accessed 2016.

[34] Lei Y., Harms J., Wang T. et al., MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks, *Med. Phys.* 2019; 46(8).

[35] Dolz J., Ayed I. B., and Desrosiers C., Dense Multi-path U-Net for Ischemic Stroke Lesion Segmentation in Multiple Image Modalities; 2018. arXiv preprint arXiv: 1810.07003.

[36] Leynes A. P.,Yang J., Wiesinger F. et al., Zero-Echo-Time and Dixon Deep Pseudo-CT (ZeDD CT): Direct Generation of Pseudo-CT Images for Pelvic PET/MRI Attenuation Correction Using Deep Convolutional Neural Networks with Multiparametric MRI, *J. Nucl. Med.* 2018; 59(5): 852-858.

[37] Srivastava N. and Salakhutdinov R., Multimodal Learning with Deep Boltzmann Machines, *J. Mach. Learn. Res*. 2014; 15(1): 2949-2980.

[38] Chartsias A., Joyce T., Giuffrida M. V., Tsaftaris S. A, Multimodal MR Synthesis via Modality-Invariant Latent Representation, *IEEE Trans Med Imaging* 2018; 37(3): 803-814.

[39] Joyce T, Chartsias A, Tsaftaris S, Robust Multi-Modal MR Image Synthesis. in Medical Image Computing and Computer-Assisted Intervention-MICCAI 2017 : 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), *20th International Conference on Medical Image Computing and ComputerAssisted Intervention, MICCAI 2017*; vol. 10435 LNCS, Springer-Verlag: 347-355.

[40] Nie D., Wang L., Gao Y., Shen D., Fully convolutional networks for multimodality isointense infant brain image segmentation, *Proc. IEEE Int. Symp. Biomed. Imag.* 2016:1342–1345.

[41] Tustison N. J., Avants B. B., Cook P. A. et al., N4ITK: improved N3 bias correction, *IEEE Trans. Med. Imag.* 2010; 29(6): 1310-1320.

[42] Nyul L.G., Udupa J. K., and Zhang X., New Variants of a Method of MRI Scale Standardization, *IEEE Trans. Med. Imag.* 2000; 19(2): 143-150.

[43] Ronneberger O., Fischer P., and Brox T. "U-Net: Convolutional Networks for Biomedical Image Segmentation; 2015. arXiv preprint arXiv: 1505.04597

[44] He K, Zhang X, Ren S, Sun J., Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016:770–778.

[45] Drozdzal M., Vorontsov E., Chartrand G., Kadoury S., Pal C., The Importance of Skip Connections in Biomedical Image Segmentation; 2016. arXiv preprint arXiv: 1608.04117.

[46] Ioffe S. and Szegedy C., Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift; 2015. arXiv preprint arXiv: 1502.03167.

[47] Radford A., Metz L., Chintala S., Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks; 2015. arXiv preprint arXiv: 1511.06434.

[48] Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R., Dropout: A Simple Way to Prevent Neural Networks from overfitting, *J. Mach. Learn. Res*. 2014; 15(1): 1929-1958.

[49] Maas A. L., Hannun A. Y., Ng A. Y., Rectifier Nonlinearities Improve Neural Network Acoustic Models, in *Proc. Int. Conf. Mach. Learn.*, 2013; 28.

[50] Ledig C., Theis L., Husz´ar F. et al., Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network; 2017. arXiv preprint arXiv: 1609.04802.

[51] Mao X., Li Q., Xie H. et al., Least Squares Generative Adversarial Networks; 2017. arXiv preprint arXiv:1611.04076.

[52] Mathieu M., Couprie C., and LeCun Y., Deep Multi-Scale Video Prediction beyond Mean Square Error; 2016. arXiv preprint arXiv:1511.05440.

[53] Isola P., Zhu J. Y., Zhou T., and Efros A. A., Image-to-Image Translation with Conditional Adversarial Networks; 2018. arXiv preprint arXiv: 1611.07004.

[54] Kingma D. P., Ba J., Adam: A Method for Stochastic Optimization; 2017. arXiv preprint arXiv: 1412.6980.

[55] He K., Zhang X., Ren S., Sun J., Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification; 2015. arXiv preprint arXiv: 1502.01852.

[56] Almahairi A., Rajeswar S., Sordoni A., Bachman P., Courville A. et al., Augmented CycleGAN: Learning Many-to-Many Mappings from Unpaired Data; 2018. arXiv preprint arXiv: 1802.10151.

[57] Ma D., Gulani V., Seiberlich N. et al., Magnetic resonance fingerprinting, *Nature* 2013; 495(7440): 187-192.

[58] Panda A., Mehta B. B., Coppo S. et al., Magnetic Resonance Fingerprinting-An Overview, *Curr. Opin. Biomed. Eng.* 2017; 3: 56-66.

[59] Tanenbaum L. N., Tsiouris A. J., Johnson A. N. et al., Synthetic MRI for Clinical Neuroimaging: Results of the Magnetic Resonance Image Compilation (MAGiC) Prospective, Multicenter, Multireader Trial, *AJNR. Am. J. Neuroradiol*. 2017; 38(6): 1103-1110.