



Multi-phase Liver-Specific DCE-MRI Translation via A Registration-Guided GAN

Jiyao Liu¹, Yuxin Li², Nannan Shi³, Yuncheng Zhou², Shangqi Gao²,
Yuxin Shi³, Xiao-Yong Zhang¹, and Xiahai Zhuang^{2(✉)}

¹ Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China

² School of Data Science, Fudan University, Shanghai, China
zhuangdata@aliyun.com

³ Department of Radiology, Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

<https://zmiclab.github.io/projects.html>

Abstract. In the diagnosis of liver lesions, Gd-EOB-DTPA-enhanced magnetic resonance imaging (MRI) at the hepatobiliary phase (GED-HBP) is particularly valuable. However, the acquisition of GED-HBP is more costly than that of a conventional dynamic contrast-enhanced MRI (DCE-MRI). This paper introduces a new dataset and a novel application of image translation from multi-phase DCE-MRIs into a virtual GED-HBP image (v-HBP) that could be used as a substitute for GED-HBP in clinical liver diagnosis. This is achieved by a generative adversarial network (GAN) with an auxiliary registration network, referred to as MrGAN. MrGAN bypasses the challenges from intra-sequence misalignments as well as inter-sequence misalignments. Additionally, MrGAN incorporates a pre-trained shape consistency network to promote local generation in the liver region. Extensive experiments demonstrated the superiority of our MrGAN over other state-of-the-art methods in terms of quantitative, qualitative, and clinical evaluations. We outlook the utility of our new dataset will extend to other problems beyond lesion detection due to the improved quality of the generated image. Code can be found at <https://github.com/Jy-studio/MrGAN.git>.

Keywords: Liver DCE-MRI · Image translation · Image registration

1 Introduction

Gadolinium ethoxybenzyl diethylenetriaminepentaacetic acid (Gd-EOB-DTPA) is a liver-specific contrast enhancement agent presently used for the diagnosis of liver lesions in MRI [15]. As shown in Fig. 1, in contrast to the conventional contrast agent Gadopentetic acid (Gd-DTPA), Gd-EOB-DTPA enhances MRI

Contributed equally.

at the hepatobiliary phase with higher sensitivity and accuracy in the detection of small hepatocellular carcinoma (HCC) [13]. However, acquiring GED-HBP is costly, attributed to the long acquisition time and expensive contrast agents [9]. Therefore, developing substitutes using virtual images of GED-HBP (referred to as v-HBP) using advanced image translation technology can be practically valuable in clinics.

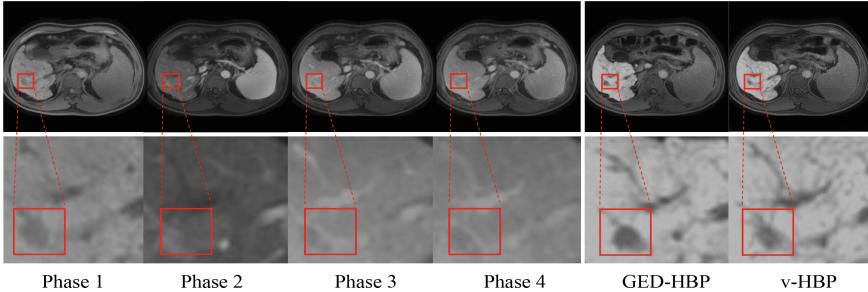


Fig. 1. A case from our dataset in which the HCC regions are marked with red boxes by radiologists. Phase 1~4 are 4 phases of DCE-MRIs. HBP and v-HBP are a pair of real and virtual GED-HBP. We can note that misalignment of details exists among all these images.

1.1 Related Works

Image-to-image translation aims to translate images from one domain to another. Deep convolutional generative adversarial networks (DCGAN) [10] produce better training results by replacing a fully connected layer with a fully convolutional layer. Pix2pix is a supervised image-to-image translation model proposed by Isola et al. [6] using a conditional GAN model. Due to the potential unavailability or limited accessibility of paired data in various scenarios, DiscoGAN [7], CycleGAN [23] and DualGAN [20] have been proposed to address the challenging unsupervised image-to-image translation task. However, these methods tend to generate numerous rich but unrealistic details and artifacts.

Generative model for medical image analysis has received a lot of attention recently. However, different from the synthesis of single-class medical images (e.g., brain MRI [19], lesion area patches [5]), liver MRI has multi-class anatomy (i.e., liver, spleen, spine, and so on). Therefore, as illustrated in [22], the attention-aware generator may extract more information specific to the region of interest, which improves the performance of tumor detection. Similar thoughts can be found in TarGAN [3], which incorporate target masks to enable the generator to focus on the local translation of the target area.

A straightforward strategy to construct paired training images is to register these multi-modality MRIs. However, there always exist small misalignments, *i.e.* intra-sequence misalignments among 4-phase DCE-MRIs as well as inter-sequence misalignments between DCE-MRIs and GED-HBP, leading to blurred

and corrupted local details which can be critical for assessing pathologies. Arar et al. [1] introduced a multi-modal registration method for natural images, but their work focused only on registration and did not discuss the relationship between registration and image translation. Beyond the above difficulties, the long scanning time of GED-HBP results in minimal liver morphology changes but significant differences in the intestinal region, which might cause more artifacts.

1.2 Contribution

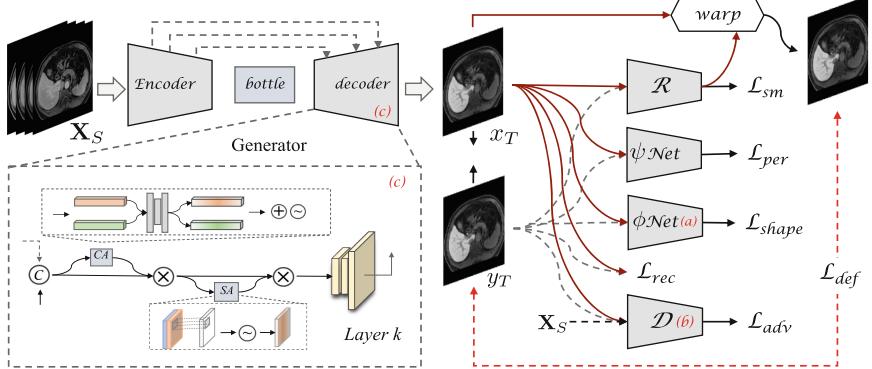


Fig. 2. The framework of our MrGAN method. We use an encoder-decoder model as the generator. \mathcal{R} denotes the auxiliary registration network and ψ Net calculates the perceptual similarity loss. (a) is a pre-trained pseudo-label segmentation network (ϕ Net). (b) is the discriminator of our MrGAN model. (c) presents a CBAM layer in the decoder, where CA and SA denote the channel and spatial attention, respectively.

To address the aforementioned issues, we design MrGAN to efficiently and effectively integrate the information from the 4-phase DCE-MRIs to generate v-HBP. To the best of our knowledge, our work is the first to achieve image translation between different contrast agents using multiple phases of DCE-MRI. We train the generator with an auxiliary registration network that adapts to the intra-sequence and inter-sequence misalignments, thereby seeking the optimal solution for both translation and registration tasks. Since the paired DCE-MRIs and GED-HBP may exhibit minimal changes in liver morphology but significant differences in the intestinal region, our method applies shape consistency through a pre-trained segmentation network, enabling a more prominent local generation in the liver region.

Our contribution is three-fold: 1) We propose a new dataset and a novel application by translating multi-phase DCE-MRIs to generate v-HBP as practical substitutes for GED-HBP. 2) We develop a multi-phase, registration-guided GAN, referred to as MrGAN, which addresses both intra-sequence and inter-sequence misalignments. 3) We validate the proposed MrGAN using clinical data with promising results.

2 Methodology

As depicted in Fig. 2, MrGAN has 4 steps: 1) We use a set of pseudo labels of liver segmentation to train an anatomy network ϕ as shape priors. 2) A generator \mathcal{G} translates multi-phase images input into v-HBP. 3) A discriminator \mathcal{D} ensures good image fidelity and contains the right target characteristics. 4) An auxiliary registration network \mathcal{R} guides the generator to address misalignment problem. Conventional GANs in image-to-image translation (*e.g.* pix2pix [6]) combine reconstruction loss and adversarial loss, but are not effective in this multi-phase images translation task, due to the issue of misalignment. Distinct from them, MrGAN introduces a deformable loss for alleviating the misalignment problem, a smoothness loss for minimizing the gradient of the deformation field, a shape consistency loss and a perceptual loss enabling more realistic global details and prominent local liver regions. The followings are the details of our MrGAN.

2.1 Standard Generative Adversarial Network

Generative Adversarial Network. The generator \mathcal{G} is an encoder-decoder network, which is responsible for translating 4 input images \mathbf{X}_S from source modality S into a v-HBP x_T approximating to the image y_T from the target modality T . \mathcal{G} consists of a set of CONV-InstanceNorm-LeakyReLU blocks. After each channel-wise concatenation in the decoder, a Convolutional Block Attention Module (CBAM) layer is added [18] in the decoder, as Fig. 2(c) shows, where channel attention (CA) and spatial attention (SA) blocks are utilized to facilitate channel-wise and spatial-wise feature recalibration, respectively. \mathcal{D} distinguishes whether the v-HBP is real or fake, as Fig. 2(b) shows. We minimize conditional and patch GAN loss as [6], which leads to higher quality image generation compared to the traditional GAN loss, and define adversarial loss as

$$\min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}_{adv}(\mathcal{G}, \mathcal{D}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\log \mathcal{D}(\mathbf{X}_S, y_T)] + \mathbb{E}_{\mathbf{X}_S, y_T} [\log (1 - \mathcal{D}(\mathbf{X}_S, \mathcal{G}(\mathbf{X}_S)))] \quad (1)$$

Reconstruction Loss. As the generator is tasked to not only fool the discriminator but also approximate the ground truth output in an L_1 sense, we use L_1 loss to reconstruct the target image. Furthermore, as our method does not require pixel-level accuracy, a Gaussian blur kernel is incorporated to alleviate the L_1 loss, which is given respectively as

$$\min_{\mathcal{G}} \mathcal{L}_{L_1}(\mathcal{G}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\|y_T - \mathcal{G}(\mathbf{X}_S)\|_1], \min_{\mathcal{G}} \mathcal{L}_{L_1^*}(\mathcal{G}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\|y_T^* - \mathcal{G}(\mathbf{X}_S)^*\|_1], \quad (2)$$

where $\mathcal{G}(\mathbf{X}_S)^*$ and y_T^* are blurred $\mathcal{G}(\mathbf{X}_S)$ and y_T , respectively. We add these two loss functions together as $\mathcal{L}_{rec}(\mathcal{G}) = \mathcal{L}_{L_1}(\mathcal{G}) + \mathcal{L}_{L_1^*}(\mathcal{G})$.

2.2 Registration-Guided Generative Adversarial Network

Beyond the standard GANs, our MrGAN introduces deformation loss, smoothness loss, shape consistency loss, and perceptual loss to address intra-sequence and inter-sequence misalignment issues in multi-phase image translation.

Deformable Registration Network. Inspired by [8], a registration network and GAN are trained simultaneously, in order to alleviate the misalignment problem of intra-sequence and inter-sequence. \mathcal{R} predicts the deformable vector field (DVF) between v-HBP and GED-HBP. Affine registration and a vector momentum-parameterized stationary velocity field (vSVF) are implemented to get better transformation regulation. The registration network is based on U-net [11] and the deformation loss is defined as

$$\min_{\mathcal{G}, \mathcal{R}} \mathcal{L}_{def}(\mathcal{G}, \mathcal{R}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\|\mathcal{y}_T - \mathcal{G}(\mathbf{X}_S) \circ \mathcal{R}(\mathcal{G}(\mathbf{X}_S), y_T)\|_1], \quad (3)$$

where, $R(\mathcal{G}(\mathbf{X}_S), y_T)$ is the deformation field and symbol \circ represents the warp operation. Moreover, a smoothness loss [2] is defined to evaluate the smoothness of the deformation field and minimize its gradient, namely,

$$\min_{\mathcal{R}} \mathcal{L}_{sm}(\mathcal{R}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\|\nabla \mathcal{R}(\mathcal{G}(\mathbf{X}_S), y_T)\|_2^2]. \quad (4)$$

The smoothness loss regularization can avoid the deformation field from being too large to cause possible foldings or unrealistic details in the output images.

Shape Consistency. Since the long scanning time of GED- HBP results in significant differences in the intestinal region and we pay more attention on the liver region, we add a shape consistency loss [3]. A segmentation network ϕ [11], as shown in Fig. 2(a), is pre-trained to regularize the result into a desired realistic shape and more prominent liver region. The shape consistency loss is defined as

$$\min_{\mathcal{G}} \mathcal{L}_{sp}(\mathcal{G}) = \mathbb{E}_{\mathbf{X}_S, y_T} [\|\phi(\mathcal{G}(\mathbf{X}_S)) - \phi(y_T)\|_2], \quad (5)$$

where $\phi(\cdot)$ is the feature map of the segmentation network pre-trained.

Perceptual Similarity. The generated images generally suffer from missing fine details in the space [16] with only pixel-wise losses. Perceptual loss [21] is applied for feature-level comparison to provide additional constraints, calculated by extracting the intermediate features of pre-trained VGG-19 as follows,

$$\min_{\mathcal{G}} \mathcal{L}_{per}(\mathcal{G}) = \mathbb{E}_{\mathbf{X}_S, y_T} \sum_l \frac{1}{H_l W_l C_l} \|\psi_l(\mathcal{G}(\mathbf{X}_S)) - \psi_l(y_T)\|_2, \quad (6)$$

where, $\psi_l(\cdot)$ is the l^{th} layer map of feature extractor.

Finally, we combine all the loss functions above as follows,

$$\min_{\mathcal{G}, \mathcal{R}} \max_{\mathcal{D}} \mathcal{L}_{total}(\mathcal{G}, \mathcal{R}, \mathcal{D}) = \mathcal{L}_{adv} + \lambda_1 \mathcal{L}_{rec} + \lambda_2 (\mathcal{L}_{def} + \mathcal{L}_{sm}) + \lambda_3 \mathcal{L}_{sp} + \lambda_4 \mathcal{L}_{per}. \quad (7)$$

These techniques can result in better translation. The weights, $\lambda_1, \lambda_2, \dots, \lambda_4$, balance the regularization effect of these techniques.

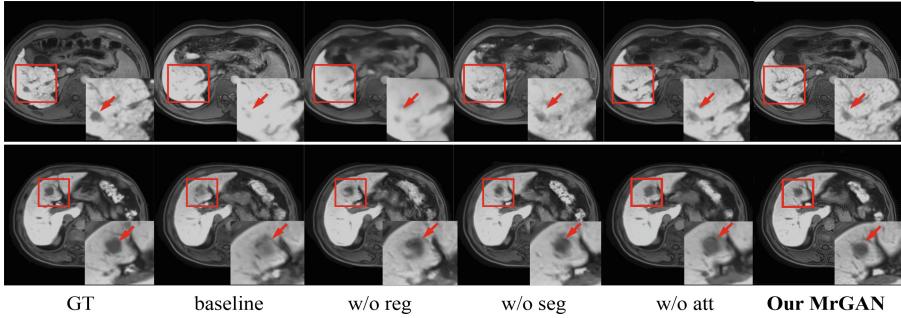


Fig. 3. Ablation study: qualitative results with ground truth (GT) in different settings

Table 1. Ablation study: quantitative results under different settings. To measure image quality, we compute a perceptual metric, *i.e.*, LPIPS [21], between feature maps from GED-HBP and v-HBP. **Bold** denotes the best performance in terms of mean values.

Methods	Quantitative Metrics			
	MAE ↓	PSNR↑	SSIM↑	LPIPS↓
baseline	0.104 ± 0.037	18.100 ± 2.187	0.632 ± 0.055	0.220 ± 0.042
MrGAN w/o att	0.102 ± 0.040	18.812 ± 2.31	0.671 ± 0.064	0.198 ± 0.046
MrGAN w/o seg	0.103 ± 0.037	18.760 ± 2.152	0.679 ± 0.057	0.199 ± 0.044
MrGAN w/o reg	0.104 ± 0.045	18.581 ± 2.086	0.651 ± 0.054	0.208 ± 0.042
MrGAN	0.096 ± 0.035	19.108 ± 2.134	0.685 ± 0.065	0.186 ± 0.045

3 Experiments and Results

3.1 Experiments

Dataset. Our dataset of Multi-Phase Liver DCE-MRI images includes 131 patients. These samples were obtained within a month, using conventional Gd-DTPA and Gd-EOB-DTPA agents, and scanned by Siemens 3.0T Skyra or Philips Ingenia 3.0T, consisting of 26 to 78 slices. The dataset includes non-enhanced MR images, hepatic arterial phase, portal venous phase, delayed phase of DCE-MRI, and GED-HBP, as shown in Fig. 1. DCE-MRI and GED-HBP scans were obtained at 0, 25, 60, and 180s post-injection of Gd-DTPA and at 20 min post-injection of Gd-EOB-DTPA, respectively. Radiologists provided the pathological gold standard for 50 HCC cases.

Data Preprocessing. The dataset underwent intra-subject 3D image registration followed by inter-sequence registration using an affine transformation and multi-scaled FFD [12] transformation to coarsely align the DCE-MRI to the GED-HBP. Due to the limited size of the dataset, we used the validation set as the test set. The proposed dataset was divided into a training and test set with

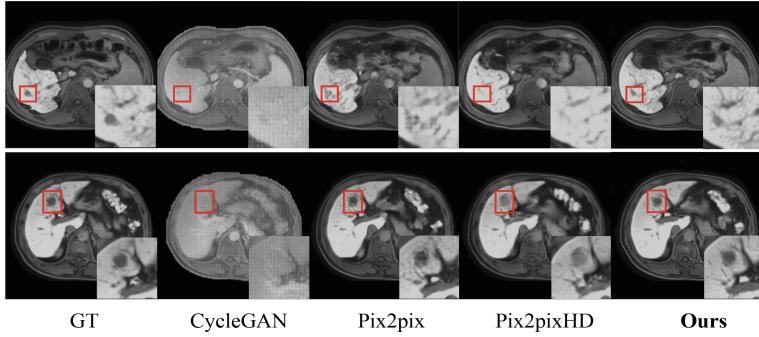


Fig. 4. Visualized results of different methods on the task of image translation.

Table 2. Quantitative results of different models on the task of image translation.

Metrics				
Methods	MAE↓	PSNR↑	SSIM↑	LPIPS↓
cycleGAN [23]	0.221 ± 0.039	11.588 ± 0.847	0.182 ± 0.056	0.605 ± 0.027
pix2pix [6]	0.124 ± 0.035	18.080 ± 2.053	0.621 ± 0.054	0.240 ± 0.044
pix2pixHD [17]	0.119 ± 0.034	18.353 ± 1.990	0.625 ± 0.058	0.216 ± 0.034
Our MrGAN	0.096 ± 0.035	19.108 ± 2.134	0.685 ± 0.065	0.186 ± 0.045

a ratio of 4:1. We expanded the dataset by splitting the 3D volumes into 2D slices, resulting in a total of 2703 slices. All MRIs were resized into 256×256 by resampling to $1 \times 1 \text{ mm}^2$. The min-max normalization overcame the intensity differences among subjects and removed the pixels with extreme values. We used data augmentation during training by applying a uniform random affine transformation with elements in the affine matrix sampled from a normal distribution $\mathcal{N}(0, 0.02)$ and a random translation of 5px.

Training. The proposed method is developed on PyTorch and trained on one NVIDIA RTX 2080Ti. We use the Adam optimizer with an initial learning rate of 10^{-4} with momentum parameters of $\beta_1 = 0.5$ and $\beta_2 = 0.999$. The network is trained for 100 epochs with a batch size of 1. We set the loss trade-off hyper-parameters $\lambda_1, \lambda_2, \dots, \lambda_4$ to 5, 1, 1, and 0.5, respectively.

3.2 Results

Ablation study. We conducted the ablation experiments to verify the effectiveness of our three key components including the registration network, segmentation network, and attention block. Our experiment settings included: 1) the baseline GAN without three key components (baseline), 2) the MrGAN without attention block (MrGAN w/o att), 3) the MrGAN without segmentation network (MrGAN w/o seg), 4) the MrGAN without registration network (MrGAN

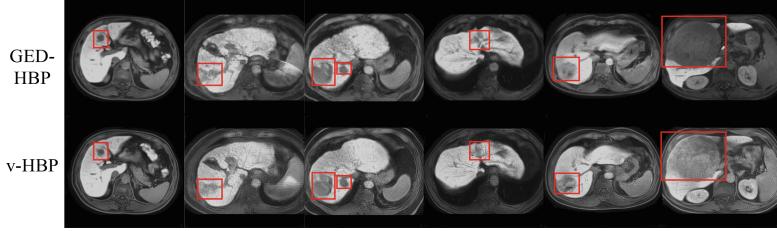


Fig. 5. Clinical evaluation: qualitative results in HCC diagnosis.

w/o reg), and 5) the proposed MrGAN. Figure 3 and Table 1 present both qualitative and quantitative results. Our key components outperformed the baseline method in all measurements, particularly in the lesion regions highlighted by the red boxes. Our registration-guided method also performed better than other components, which was more valid for bypassing the misalignment problem.

As shown in Table 1 and Fig. 3, training with additional modules enhanced edge information and produced more realistic results. The registration network yields a substantial improvement in the qualitative metric. This is because in this setting the registration network R implicitly performs both the alignment tasks. Visual results of v-HBP by MrGAN were most similar to the ground truth, with sharper textures and more details, particularly in red-boxed regions.

Comparison to Image Translation Methods. To demonstrate the superiority of our proposed MrGAN method, we compared our method with three state-of-the-art image translation methods, including CycleGAN, pix2pix, and pix2pixHD [17]. We trained the MrGAN by minimizing the total loss in (7).

Table 3. Results of hypothesis testing. The alternative hypotheses are $H_0^{(1)} : \#\{x_T\} > \#\{y_T\}$, $H_0^{(2)} : \#\{x_T\} < \#\{y_T\}$ and $H_0^{(3)} : \#\{x_T\} \neq \#\{y_T\}$ respectively.

Hypothesis			
Methods ($\alpha = 0.05$)	$H_0^{(1)}$ (Greater)	$H_0^{(2)}$ (Less)	$H_0^{(3)}$ (Two-sided)
Wilcoxon Rank Sum test (p -value)	0.9703	0.1729	0.3458
Mann-Whitney U test (p -value)	0.9322	0.0823	0.1646
Sign test (p -value)	/	/	0.1250

Table 2 reports the quantitative results of compared methods. One can see that our MrGAN outperforms the other three methods. We showed qualitative results from two HCC patients in Fig. 4. The highlighted demonstrates that our MrGAN method significantly outperforms other methods in the liver area, particularly the lesion area. Instead, other GAN-based methods generated numerous artifacts that are absent in the original image, which hinders radiologists' accu-

rate diagnosis. This suggests MrGAN can not only generate realistic v-HBP but also properly retain liver and lesion structure.

Clinical Evaluation. We greatly focused on the evaluations from radiologists, particularly regarding the performance of the liver region in our results. For the purpose of diagnosis, we performed the Wilcoxon Rank Sum test, Mann Whitney U test, and sign test [4] to test our null hypothesis [14], namely no significant difference between the v-HBP by MrGAN and the GED-HBP in HCC diagnosis.

- 1) $\#\{x_T\} > \#\{y_T\}$: The v-HBP outperforms the GED-HBP.
- 2) $\#\{x_T\} = \#\{y_T\}$: There is no significant difference between them.
- 3) $\#\{x_T\} < \#\{y_T\}$: This is the opposite case of the first hypothesis.

where $\#$ refers to the number of cases in each class. And the resultant p -values with different statistics are shown in Table 3.

Table 3 presents all testing results, meaning that we cannot reject the null hypothesis at a significance level of 0.05. (See examples in the Fig. 5.) This indicates that there is no statistically significant difference between the v-HBP by MrGAN and the GED-HBP in clinical diagnosis.

4 Conclusion

This work introduces a novel application by translating multi-phase DCE-MRIs into a virtual image, toward replacing GED-HBP which is too costly to be widely available in reality. This is achieved via a new framework based on a registration-guided GAN, which addresses the common misalignment problem, and the shape consistency method, which has effectively exploited shape information. Our MrGAN results in more realistic global details and prominent local liver regions in v-HBP. Experimental results also demonstrated that our MrGAN outperformed existing methods in quantitative, qualitative, and clinical metrics. In the future, we plan to explore the potential of the additional clinical diagnostic benefits of generating GED-HBP beyond HCC detection, such as focal nodular hyperplasia(FNH), hepatocellular adenoma(HCA), and so on.

References

1. Arar, M., Ginger, Y., Danon, D., Bermano, A.H., Cohen-Or, D.: Unsupervised multi-modal image registration via geometry preserving image-to-image translation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13410–13419 (2020)
2. Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V.: Voxelmorph: a learning framework for deformable medical image registration. IEEE Trans. Med. Imaging **38**(8), 1788–1800 (2019)
3. Chen, J., Wei, J., Li, R.: TarGAN: target-aware generative adversarial networks for multi-modality medical image translation. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12906, pp. 24–33. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87231-1_3

4. Conover, W.J.: Practical nonparametric statistics, vol. 350. John wiley & sons (1999)
5. Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., Greenspan, H.: Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* **321**, 321–331 (2018)
6. Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1125–1134 (2017)
7. Kim, T., Cha, M., Kim, H., Lee, J.K., Kim, J.: Learning to discover cross-domain relations with generative adversarial networks. In: *International Conference on Machine Learning*, pp. 1857–1865. PMLR (2017)
8. Kong, L., Lian, C., Huang, D., Yanle, H., Zhou, Q., Li, Z.: Breaking the dilemma of medical image-to-image translation. *Adv. Neural. Inf. Process. Syst.* **34**, 1964–1978 (2021)
9. Purysko, A.S., Remer, E.M., Veniero, J.C.: Focal liver lesion detection and characterization with gd-eob-dtpa. *Clin. Radiol.* **66**(7), 673–684 (2011)
10. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434* (2015)
11. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
12. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Trans. Med. Imaging* **18**(8), 712–721 (1999)
13. Sano, K., et al.: Imaging study of early hepatocellular carcinoma: usefulness of gadoxetic acid-enhanced mr imaging. *Radiology* **261**(3), 834–844 (2011)
14. Shan, H., et al.: Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. *Nat. Mach. Intell.* **1**(6), 269–276 (2019)
15. Verloh, N., et al.: Assessing liver function by liver enhancement during the hepatobiliary phase with gd-eob-dtpa-enhanced mri at 3 tesla. *Eur. Radiol.* **24**, 1013–1019 (2014)
16. Wang, C., Chang, X., Wang, C., Tao, D.: Perceptual adversarial networks for image-to-image transformation. *IEEE Trans. Image Process.* **27**(8), 4066–4079 (2018)
17. Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional gans. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8798–8807 (2018)
18. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) *ECCV 2018. LNCS*, vol. 11211, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1
19. Yang, W., et al.: Predicting CT image from MRI data through feature matching with learned nonlinear local descriptors. *IEEE Trans. Med. Imaging* **37**(4), 977–987 (2018)
20. Yi, Z., Zhang, H., Tan, P., Gong, M.: Dualgan: unsupervised dual learning for image-to-image translation. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2849–2857 (2017)

21. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–595 (2018)
22. Zhao, J., et al.: Tripartite-gan: synthesizing liver contrast-enhanced mri to improve tumor detection. *Med. Image Anal.* **63**, 101667 (2020)
23. Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)