IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 March 2024, accepted 12 April 2024, date of publication 17 April 2024, date of current version 24 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3390245

RESEARCH ARTICLE

HLSNC-GAN: Medical Image Synthesis Using Hinge Loss and Switchable Normalization in CycleGAN

YANG HENG, MA YINGHUA, FIAZ GUL KHAN, AHMAD KHAN, AND ZENG HUI
Department of Computer Science, COMSATS University Islamabad, Abbottabad Campus, Abbottabad 22010, Pakistan
Corresponding authors: Ma Yinghua (1257179128@qq.com) and Fiaz Gul Khan (fiazkhan@cuiatd.edu.pk)

ABSTRACT In the field of medical image analysis, MRI and CT, among other multimodal medical images, play crucial roles. To overcome the limitations of image acquisition, researchers have proposed medical image synthesis techniques, including both traditional methods and deep learning approaches. In this study, we introduce a universal framework based on cycleGAN for generating CT images from MRI data.This framework incorporates a hinge loss function to establish mappings between different modalities and enhance structural consistency between input and output images. We also employ a switchable normalization technique to improve model stability and reduce manual intervention. These enhancements result in the generation of higher-quality synthetic images while avoiding gradient issues and mode collapse.The results of this research demonstrate significant progress in medical image synthesis. Compared to existing methods, our model exhibits superior performance in quantitative evaluation metrics while maintaining better diversity and structural consistency. This indicates that our framework holds promise in medical image synthesis and can provide valuable support in areas such as disease prediction and treatment.

INDEX TERMS CycleGAN, image translation, generative adversarial network, image synthetic, medical images.

I. INTRODUCTION

Medical imaging technology plays a crucial role in clinical diagnosis and treatment. They not only help doctors accurately judge the type and grade of diseases but also improve diagnostic efficiency and save time. Four types of medical imaging are commonly used, namely MRI [1], CT [2], PET (positron emission tomography) [3], and SPECT (single photon emission computed tomography) [4]. Among them, MRI and CT are the most commonly used images in radiation therapy plans. They are widely used in various imaging tasks, such as image segmentation [5], classification [6], detection [7], and image registration [8]. Medical imaging can help reveal internal anatomical structures, for example, MRI can distinguish malignant tumors from soft tissue organs at risk, and CT provides detailed information about bone structure. However, medical imaging has some challenges, such as cost, radiation dose, time, and specific imaging methods. Especially in the head, brain, pelvic,and neck areas, CT imaging has limited ability to reveal soft tissue contrast, which poses a challenge to accurately outline the contours of organs at risk and target areas. In addition, CT scans involve the use of X-ray radiation, and repeated scans can expose patients to ionizing radiation, which can cause radiation damage to the human body. Therefore, the number of scans is strictly controlled. In contrast, MRI is an alternative method that can produce noise radiation, is lower in cost, can provide more soft tissue details, and produces satisfactory results in organ visualization compared to CT. Recently, research has increasingly focused on generating CT images from MRI. Previously, the network structure of synthetic CT was relatively simple, but it demonstrates superior performance compared to traditional synthetic CT generation methods. Compared with traditional atlas-based methods, existing studies have proposed using different frameworks to generate sCT [9], [10], [11], [12], [13] From a broader perspective,

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad J. Abdel-Rahman.

© 2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

the generation of synthetic CT can be considered as a style transfer task of magnetic resonance imaging. Generative adversarial networks (GAN) [14] have received widespread attention in the field of medical imaging. Existing studies have used GAN to solve style transfer problems in medical imaging [15], [16], [17], [18], [19] Although previous MR-based CT synthesis methods have shown encouraging results, most existing methods are applied to training on registered images from two or more domains. However, it is very difficult to obtain one-to-one paired images with the same content but different styles. To address the challenge of generating CT images from MRI, Zhu et al. [20] proposed cycleGAN, an upgraded version of GAN, with two generators and two discriminators. CycleGAN can achieve mutual conversion of images between two domains without the need for data pairing or registration. Similar methods, such as DualGAN [21] and discoGAN [22], have also shown considerable performance. Recently, researchers have begun to apply cycleGAN to the medical field to generate synthetic CT images. For example, Wolterink et al. [23] successfully generated higher-quality head and neck sCT images using cycleGAN. However, in the CycleGAN model, the lack of direct constraints between input and output images may affect structural consistency. To improve this problem, Yang et al. [24] introduced an additional structural consistency loss, and Liu et al. [25] adjusted the loss function of cycleGAN by merging the pseudo-cycle consistency module and domain control module, thereby improving the accuracy of generated images. In addition, Kang et al. [26] proposed a cycleGAN with perceptual loss, and Sun et al. [27] proposed a DU-CycleGAN model using U-Net as the discriminator, both of which achieved good synthetic performance.

Although methods based on cycleGAN have made progress in unsupervised image generation, challenges still exist in improving the overall quality of the generated images. To address this issue, our study introduces hinge loss and switchable normalization in cycleGAN (HLSNC-GAN), which combines these mechanisms to tackle key challenges in medical image synthesis. Firstly, hinge loss is introduced to enhance the mapping relationship between different modes and ensure the consistency between input and output images. Secondly, the introduction of switchable normalization (SN) in the generator reduces manual intervention within the model. Additionally, the embedding of deep residual network modules and the use of the root mean square propagation (RMSprop) optimizer help maintain image details and enhance the stability and convergence of the model during training. Lastly, we implemented an enhanced convolutional neural network (CNN) architecture, convolutional layer with SN and leaky ReLU activation (Conv2d-SN-LReLu), in the discriminator to improve its classification efficiency. The main contributions of our proposed HLSNC-GAN method are as follows:

- Hinge loss function: The introduction of the hinge loss function enforces structural consistency between input MR and synthetic CT images, proving more effective than traditional adversarial loss techniques.
- Switchable normalization (SN): The introduction of switchable normalization into the network, including differentiable normalization layers, allows the model to automatically learn the normalization method for each layer based on the input data, enhancing the stability of model learning and reducing manual intervention.
- RMSprop optimizer: The use of the RMSprop optimizer improves the quality and stability of the generated images.
- Gradient penalty: To preserve the diversity of the generated samples and avoid excessive punishment of outliers, the gradient penalty is weakened in the experiments.
- MR-CT conversion: HLSNC-GAN is used for MR-CT conversion of medical images. It can be trained with unpaired images in unsupervised learning and is also applicable to paired and unpaired medical image datasets.
- Enhanced image details: By introducing resnet into the generator and utilizing skip connections, the diversity and quality of the generated synthetic images are enhanced.
- Activation function: The use of the leakyRelu activation function ensures that parameters do not scale during backpropagation, helping to avoid mode collapse and gradient explosion or vanishing issues.

The organization of this thesis is as follows: section II reviews related literature, section III introduces our methodology, section IV presents experiments and results, section V discusses potential limitations and future research directions, and section VI concludes the study.

II. RELATED WORK

Various researchers have proposed medical image synthesis techniques using both conventional and deep learning methods in recent times. The primary drawback of conventional image synthesis models is that they are not generalized models since humans must explicitly define the rules for synthesizing images from one domain to another. In contrast, deep learning allows for the application of different pairs of image modalities to the same architecture with a little adjustment, thereby increasing the architecture's generalizability with minimal human involvement. The frameworks for medical image synthesis in the deep learning subfield can be broadly categorized into three main groups: u-net [28], autoencoders (AE) [29], and generative adversarial networks (GAN) [30]. These three methods are not vastly different from one another but differ primarily in their level of architectural complexity. Among these three categories, GAN-based models are widely used in medical image synthesis. GAN-based Models: With the development of image synthesis technology, various deep learning models have been developed, with the GAN model being the most

popular in the field, especially for image generation due to its unique synthesis capabilities. GAN [14] proposed in 2014, with the goal of using noise and real images as input data to create new images. Recently, the GAN model has aroused great interest among researchers in the medical field, particularly in medical image conversion, and has achieved gratifying results. GANs can be used to train and synthesize realistic images not only in paired data sets [31] but also in unpaired data sets [27], resulting in satisfactory experimental results. For simplicity we can divide our literature review in medical image synthesis into two subsections that are 1) GANs for paired medical images and 2) GANs for unpaired medical images. In both these two subcategories of GANs, different researchers have proposed different GAN-based variants in image-to-image translation tasks by modifying either of these I. Improving the training method or loss function of GANs; II. Applying different neural networks to GANs; III. Combination of I and II.

GANs for paired medical images: Emami et al. [32] presented a novel GAN model for generating CT images from corresponding T1 MR images. They employed a residual network as the generator and cross-validation to assess the effectiveness of their approach. Their model was found to be superior to the CNN model, as it better-preserved details and abnormal regions of the synthetic images. Kazemifar et al. [33] also proposed an improved GAN method to synthesize MRI and CT images. They used MI (mutual information) U-Net as the loss function, a convolutional layer with a ReLU (Rectified linear unit), and a discriminator network consisting of two fully connected layers that use meta-cross entropy as the activation/loss function. The experiment achieved good results. Emami et al. [34] proposed an attention-GAN model to generate sCT with brain MRI as input. They introduced batch normalization in the model's discriminator to prevent overfitting. Sun et al. [35] proposed a stacked GAN (sGAN) model based on residual network (ResNet) and FCN. They used batch normalization (BN) layers in the generator (G) and discriminator (D) to speed up the training of the network and convergence. Liu et al. [36] proposed a GAN model using a ResNet as a generator for synthetic CT in MR-only brain radiation therapy. Gotoh et al. [37] proposed a novel cGAN model using ''U-Net'' as the generator and ''PatchGAN'' as the discriminator. BN (Batch normalization layers) are introduced in the generator and discriminator to control gradient explosion and prevent gradients from disappearing during training. Furthermore, to generate sCT images and assess the accuracy of dose estimation, Tang et al. [38] developed a generative adversarial network (GAN) with a ''U-Net'' shaped encoder-decoder architecture that incorporates certain specific image translation techniques to convert T1-weighted MRI into sCT image. Emami et al. [39] presented a SA-GAN architecture for synthesizing CT images from MRI. The generator in their model consists of two parallel streams, one for reconstruction loss and the other for structural segmentation loss, which are jointly optimized with the GAN loss. A fusion layer is employed to merge the outputs of the two streams to produce a pseudo-CT image. The performance of SA-GAN is compared against a CNN model, and the results indicate that SA-GAN outperforms CNN in generating both consistent (e.g., bone) and inconsistent (e.g., rectal gas) regions. In addition, the conditional GAN (cGAN) is a general framework utilized for image-to-image translation tasks that rely on paired datasets. Qi et al. [40] conducted a study utilizing paired datasets of MR and CT images to investigate how different MR sequences affect the precision of deep learning-based synthetic CT (sCT) generation in the intricate head and neck region. Their findings indicate that the cGAN model, which incorporates multiple MR sequences as input, achieved the highest level of accuracy, outperforming the U-net network. Ranjan et al. [41] utilized a cGAN model and presented a modified version that incorporates residual blocks within the generator network instead of conventional convolutional blocks. This residual block architecture comprises two convolutional layers, a normalization layer, an activation layer, and a skip connection that bypasses the two convolutional layers. According to their experimental results, the proposed method shows promising performance when trained on paired MRI and CT scan datasets. Cusumano et al. [42] generated synthetic CT scans for low-field MR-guided radiotherapy for lung cases by modifying the internal network of the cGAN model. They used a 2D layer architecture with skip connections and employed patchGAN as the discriminator network. In 2017, Isola et al. [43] proposed the pix2pix network model, which is also frequently used in the field of paired image-to-image generation [44], [45], [46], [47]. Aljohani and Alharbe [48] proposed using a deep pix2pix model similar to patchGAN, U-net, and vanilla GAN loss as a generator to create high-quality synthetic brain MR images. The experimental results demonstrated that the original pix2pix GAN model outperformed other models in terms of synthetic image quality. Phukan et al. [49] suggested the application of ResNet-DCGAN to produce chest X-ray images of COVID-19 patients. Specifically, they utilized ResNet50 as the discriminator in the deep convolutional neural network (DCNN) model and substituted the early adam optimization algorithm with the RAdam optimization algorithm. According to their findings, the synthetic images generated by their proposed approach exhibited superior performance compared to state-of-the-art DCGAN models on paired datasets. Ellis et al. [50] experimentally demonstrated that 3D DCGAN models exhibit greater image variability in synthetic CT images but at the expense of image quality. Huang et al. [51] introduced a medical super-resolution generative adversarial network (SRGAN) with attention-based denoising capability, referred to as AID-SRGAN, for enhancing the resolution of medical images. Initially, a realistic degradation model was proposed considering multiple degradation factors. Subsequently, attention

mechanisms were integrated into the denoising module to improve its robustness. Finally, the model achieved high-resolution reconstruction by processing low-resolution radiographic images. Furthermore, in the process of medical image translation, it is often challenging to maintain the image structure. To address this issue, Li et al. [52] proposed a frequency-guided diffusion model (FGDM). FGDM is a hybrid diffusion model that uses denoising diffusion GAN as the backbone network and utilizes frequency-domain filters to guide the diffusion process, accurately preserving structural information during image transformation. This method possesses zero-shot learning capabilities, requiring training only on the target domain data and can be directly applied to source-to-target domain conversion. The FGDM method was evaluated on the task of transforming head-neck and lung cone-beam CT (CBCT) to CT, demonstrating significant advantages. Lei et al. [53] proposed a lesion attention-conditioned generative adversarial network (LAC-GAN) to synthesize retinal images with realistic lesion details. The generator utilizes residual blocks to handle randomly generated gaussian noise for synthesizing images. Moreover, a lesion feature attention mechanism based on the random forest (RF) method is introduced to focus on pathological information. Additionally, multiple discriminators with shared weights are employed to enhance model performance through affine transformations. Experimental results demonstrate that this method contributes to improving the performance of disease detection models.

GANs for unpaired medical images several image syntheses models have been developed for unsupervised model training, which can use unpaired datasets as input target data, among which cycle-consistent adversarial networks (CycleGAN) is the most prominent. Kearney et al. [54] proposed an attention-aware cyclegan model for unpaired MR-to-CT image conversion. According to the results, unpaired cycleGANs are a better choice for MR-to-CT image conversion based on paired images. Boni et al. [55] introduced an enhanced version of cycleGAN, known as augCGAN, for generating synthetic CT images in an unpaired fashion. They evaluated the mean absolute error (MAE) between actual and synthetic CT images across multiple patient-specific sequences, with the goal of improving the overall generalization performance of MR-to-CT synthesis. Liu et al. [25] suggested employing a multi-cycle GAN for synthesizing CT images from MRI scans of the head and neck. In their study, a z-Net was devised as a generator to incorporate rules for skip connections, and a pseudo-cycle consistent module was developed to ensure generation consistency, thereby enhancing the precision of detail transformation. To ensure the quality of image generation, Gu et al. 62000 proposed a Dual3D&PatchGAN model based on transfer learning. This model enables bidirectional conversion between two classes of images without the need for paired datasets. It also addresses the issue of shadows that occur during the image generation process in traditional dualGAN. Compared to other medical image generation models, it exhibits superior generative performance. Additionally, Abu-Srhan et al. [56] proposed a paired-unpaired unsupervised attention-guided generative adversarial network (uagGAN) model to generate more accurate and clearer images. The authors initially pre-trained the uagGAN model using paired datasets, and simultaneously employed a novel composite loss function during the pre-training phase to enhance the model's generative performance. The composite loss includes both adversarial loss from wasserstein GAN and non-adversarial losses (content loss and L1 loss). Subsequently, the model was retrained using unpaired datasets. Experimental results demonstrated better performance compared to existing image transformation models. Kang et al. [26] proposed a deep learning-based approach based on a cycle-consistent GAN framework for unpaired MR-to-CT image translation. They introduced a perceptual loss to cycleGAN to enhance the performance of synthetic CT generation. Moreover, their experiments demonstrated that the use of cycleGAN with perceptual loss achieved superior results in sCT generation when trained with weakly paired deformed CT (dCT)-MRI for MRgRT, outperforming u-net. Sun et al. [57] introduced a multi-discriminator-based cycleGAN (MD-CycleGAN) model to generate high-quality synthetic CT images from MRI scans. The model utilized DenseNet as the primary generator architecture and was evaluated using the quadruple cross-validation method. Experimental results indicate that the proposed approach outperforms some of the most advanced existing models. Sun et al. [27] suggested a novel approach using the cycleGAN model from MR images to generate CT in medical applications. The researchers enhanced the accuracy of the discriminator by utilizing a u-Net network and introduced a content-aware feature recombination (CARAFE) method instead of a fixed kernel for all samples, resulting in superior performance compared to existing state-of-the-art methods. Wang et al. [58] proposed a method for better and more efficient MRI-to-CT synthesis. A new framework for image-to-image conversion called cycleCUT has been proposed. The study incorporates a contrastive learning module, inspired by the contrastive unpaired translation (CUT) network, into the cycleGAN framework. The experimental outcomes confirm the practicability of the suggested cycleCUT model in training on image data that is not aligned. Furthermore, compared to cycleGAN and CUT models, cycleCUT is optimal in terms of the structural details of the generated sCT images. On the other hand, obtaining matched MRI and CT images in the field of medicine is constrained by factors such as cost, radiation dose, and modality limitations. To address this limitation, [59] proposed the DC-cycleGAN model, which synthesizes medical images from unpaired data. The discriminator of DC-cycleGAN introduces dual contrastive loss, indirectly establishing constraints between real source images and synthesized images by using samples from the source domain as negative samples and compelling the synthesized

images to diverge from the source domain. Additionally, the author integrated cross-entropy and the structural similarity index (SSIM) to simultaneously consider the brightness and structure of the samples during image synthesis. Compared to other cycleGAN-based methods for medical image synthesis, DC-cycleGAN demonstrates promising results.Furthermore, in the domain of medical image transformation, prompt diagnosis and treatment of diabetic retinopathy play a crucial role in preventing vision loss. To enhance diagnostic accuracy beyond previous techniques, Alwakid et al. [60] proposed the utilization of the contrast-limited adaptive histogram equalization (CLAHE) filtering algorithm combined with enhanced super-resolution generative adversarial networks (ESRGAN) for image enhancement, achieving an accuracy rate of 98.7%. Experimental findings demonstrated that the integration of CLAHE and ESRGAN can augment the model's performance and learning capabilities. Additionally, within the realm of medical imaging, training an effective model using image synthesis techniques is challenging due to limited lesion data. Addressing this concern, Iskandar et al. [61] introduced a novel approach employing generative adversarial networks (GANs) to tackle the scarcity and limited diversity of fetal ultrasound image data. The proposed method involves the adoption of diffusion super-resolution GAN and transformer-based GAN, leveraging publicly accessible datasets to generate synthetic fetal ultrasound brain plane images. The study's results underscore the promising potential of GAN-based techniques in generating realistic and high-resolution ultrasound images. To tackle the issue of limited image quality in generating medical images using generative adversarial networks (GANs) on minority-class datasets, Huang and Jafari [62] proposed balanced adversarial generative adversarial networks (BAGAN). However, BAGAN exhibits instability when dealing with similar image categories. To mitigate this, a supervised autoencoder with intermediate embedding models was introduced to untangle the latent vectors of labeled data. By leveraging an enhanced autoencoder initialization, the BAGAN architecture with gradient penalty (BAGAN-GP) was constructed. This method effectively addressed the instability issue of the original BAGAN, resulting in faster convergence and the production of high-quality generated images. Moreover, to address the problems of poor visual quality and structural distortion in image super-resolution reconstruction models, Ma and Li [63] introduced a deep gradient-guided method based on generative adversarial networks. This approach incorporates a gradient branch for feature propagation and fusion with the image branch to prevent edge distortion. Improved multi-scale residual blocks are employed to capture multi-scale information, and WGAN-GP is utilized to stabilize network training. In comparison to SRGAN and ESRGAN algorithms, this method effectively prevents structural distortion and enhances the quality of generated images. The model's computational complexity is 23.7 GFLOPs, surpassing other methods.

In the domain of medical image transformation, GAN-based models outperform traditional translation methods. Using unpaired datasets for model training has obvious advantages compared to using paired image data, this is especially true in the field of medical imaging, where data scarcity poses a major challenge due to cost constraints. Additionally, due to patient privacy protection, paired datasets are extremely rare.While using existing GAN techniques to generate medical images, several issues may arise. Firstly, the generated images may lose details, resulting in poor image quality. Secondly, the diversity of generated images may be insufficient, leading to high similarity among them. Thirdly, excessive punishment for insufficient diversity may result in generated images that are too smooth or unrealistic. Additionally, inadequate model stability may cause overfitting, gradient disappearance, and gradient explosion problems.

In recent years, various researchers have proposed a variety of medical image synthesis techniques using both traditional and deep learning methods Table 1. One popular approach among these methods is to utilize GAN-based models for image synthesis, especially in the case of CT and MRI, the two most widely used modalities. Despite significant advancements made by deep GAN models, the generation of high-quality medical images still faces several limitations. For instance, GAN models based on U-net, such as Pix2Pix, require highly curated training data where the training data must be paired. Another GAN variant, the LSGAN model, excessively penalizes outliers during the training process, resulting in insufficient sample generation and a lack of ''diversity'' and mode collapse.Furthermore, many GAN-based variants lack constraints on the mapping relationship between different modalities, thus unable to generate high-quality synthetic images. Additionally, networks with too many layers and parameters lead to loss of image details, and parameters are scaled during backpropagation, resulting in vanishing gradients and affecting the stability of model learning.

This study aims to propose a novel framework that utilizes a generalized cycleGAN approach to generate CT images from MRI data. Our proposed framework effectively handles unpaired training data, overcoming limitations imposed by scarce paired data. To enhance correlation between various modalities and maintain structural consistency between input and output images, we introduce a novel loss function for the generator. Additionally, we employ switchable normalization techniques to enhance the stability of model training, accelerate convergence, and reduce manual intervention. Moreover, we utilize a new adaptive learning rate method in the optimizer to effectively prevent overfitting on the training data, enhancing fidelity and robustness of the generated images. We also alleviate excessive penalty on outliers during the model training process, enhancing the model's learning capability and increasing the diversity of generated samples. Furthermore, we utilize skip connections in the generator ResNet architecture to preserve more image details, further

TABLE 1. Comparison of some GAN based image synthesis models. (Peak Signal-to-Noise Ratio : PSNR).

Author	Method description / used of study	Dataset (unpaired/paired)	Generator(G) / Discriminator(D)	(Function of optimization/ Activation/ loss/ Normalization): foO/foA/foL/foN	Evaluation Metrics
[26]	Generating sCT images from weakly paired MRI using CycleGAN with perceptual loss	Seoul National University College of Medicine (90 paired images)	G: U-Net D: ResNet	foA: ReLU foO: Adam foL: adv/perceptual/cyclic/L2 foN: spectral N	PSNR=26.3 SSIM=0.90
[27]	DU-CycleGAN for CT image synthesis from 3D MR	MICCAI 2020 challenge (75 paired images)	G: U-Net D: UNet-like	foL: adv/consistency foN: Z-score standardization	PSNR=27.12 SSIM=0.84
[41]	Conditional GAN model for sCT Generation from T2-MRI	The whole brain atlas (367 paired images)	G: multiple ResNet blocks D: VGG	foA: ReLU/sigmoid foO: Adam foL: adv/L1/soft-max foN: IN	PSNR=21.422 SSIM=0.823
[51]	AID-SRGAN, denoise and generate high-resolution radiographs	MURA (paired) (40,005 musculoskeletal radiographs)	G: AID and SR network D: VGG16	foL: L1/SSIM foO: Adam	PSNR=31.90 SSIM=0.9271
[52]	FGDM Denoised DiffusionGAN (DDGAN) for generate sCT from CBCT	Head and neck/lung datasets (paired 99 images/18 patients)	G: DCNN D: CNN	foL: Structural detail/Adv foO: Adam foN: linearly	PSNR=27.8 SSIM=0.908
[54]	Attention-aware CycleGAN model for generate sCT from MRI	Radiation Oncology, University of California (5560 unpaired slices)	G: U-Net-128 D:attention-gated "PatchGAN"	foA: LReLU/ReLU/tanh foO: Adam foL: adv foN: Synchronized BN	PSNR=62.35 SSIM=0.778
[63]	DGGGAN for image super-resolution reconstruction	DIV2K (800 images)	G: Gradient branch/ Multi-scale residual block D: CNN based on VGG	foL: Adv/gradient/perceptual/pixel foO: Adam foN: BatchNorm	PSNR=30.48 SSIM=0.8601

improving the diversity and quality of synthesized samples. Finally, we use the leaky ReLU activation function to prevent parameter scaling during backpropagation, effectively preventing mode collapse and gradient explosion or vanishing issues.

III. MODEL

The development based on cycleGAN for unsupervised image generation has made significant progress, but there is still considerable room for improvement in enhancing the overall quality of generated images. Many details are lost during the image transformation process, which may lead to an overly strong discriminator and excessive penalization of outliers, resulting in insufficient sample diversity and mode collapse. Additionally, as the number of parameters and network layers increases, various models based on enhanced GANs have been developed. However, parameter scaling during backpropagation leads to the vanishing gradients problem in the later stages of training, affecting the stability of model learning.

This study aims to generate high-quality CT images from unpaired MRI data and proposes a new model based on cycleGAN called HLSNC-GAN. In the generator of HLSNC-GAN, we introduce a novel hinge loss function, combined with cycle consistency and L1 norm expectation constraints, to establish better mode mapping relationships and ensure consistency between input and output images. Furthermore, we introduce a switchable normalization (SN) mechanism, allowing normalization layers to automatically adjust batch sizes based on input data, thereby enhancing the stability of model learning and improving overall performance. The residual convolutional network in the generator preserves more feature details through skip connections, increasing sample diversity, aiding in preventing parameter scaling issues, and stabilizing model learning. Additionally, we moderately relax the gradient penalty in the model to avoid excessive punishment of discrete data, thereby improving the quality of generated images. These improvements collectively address the limitations of early GAN-based image-to-image transformation techniques and significantly enhance the quality of synthetic CT images generated from unpaired MRI data. Our model is applicable to both paired and unpaired MRI-CT image datasets, providing a more reliable foundation for medical image synthesis.

A. CYCLEGAN-BASE

CycleGAN (Cycle-consistent GAN) is a framework that integrates the concept of dual learning into the original GAN model. CycleGAN is capable of transforming images between different domains. In 2016, Microsoft Research

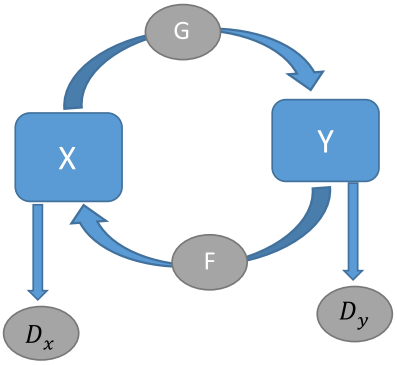

FIGURE 1. Cycle-GAN.

Asia (MSRA) introduced dual learning as an unsupervised reinforcement learning method for machine translation. It can address the problem of lacking paired annotations in massive data. Dual learning operates on a dual task mechanism similar to bilingual translation, utilizing feedback signals between the original task and the dual task to train the model. CycleGAN consists of two generators and two discriminators, forming a pair of symmetrical GAN networks. It involves two mapping relationships, as illustrated in the diagram (FIGURE.1).

While the original GAN and cGAN/Pix2Pix can also be used for tasks such as image translation and style transfer, the key difference between them and the cycleGAN-based HLSNC-GAN lies in the requirement of input as paired datasets. In contrast, the HLSNC-GAN framework of cycleGAN does not have this restriction and can be trained and generate desired images even with unpaired data. In the following sections, we will provide a detailed presentation of our optimization strategies. We will showcase the research model of HLSNC-GAN based on the cycleGAN framework and describe the architecture of the generator and discriminator in the HLSNC-GAN network. Furthermore, we will explain the loss functions, optimizers, and normalization techniques utilized in the training process of the HLSNC-GAN framework. The following FIGURE 2 illustrates the schematic diagram of HLSNC-GAN.

B. HLSNC-GAN'S GENERATOR

The generator is a key component in the HLSNC-GAN network architecture (Table 2), which is based on the cycle-GAN framework and consists of two generators (FIGURE 3): $G_{\text{Generator}}$ and $F_{\text{Generator}}$. $G_{\text{Generator}}$ is responsible for mapping input MR images to CT images in the target domain, while $F_{\text{Generator}}$ is responsible for mapping CT images back to MR images. Each generator is composed of convolutional layers, resNet modules (FIGURE 4), and deconvolutional layers. Both $G_{\text{Generator}}$ and $F_{\text{Generator}}$ contain 9 residual blocks (Table 3), with each block comprising convolutional layers, SN normalization, and ReLU activation functions. The models also incorporate skip connections to preserve and propagate fine-grained details of the images. During the training process, $G_{\text{Generator}}$ aims to generate high-quality synthesized CT' images by minimizing the difference between the generated CT' images and the real CT images. Similarly, $F_{\text{Generator}}$ strives to generate high-quality synthesized MR' images by minimizing the difference between the generated MR' images and the real MR images. To increase the depth and expressive power of the network, the two generators leverage the stacking of residual blocks to capture the relationship between input MR images and CT images. Through the operations of convolutional and deconvolutional layers, $G_{\text{Generator}}$ and $F_{\text{Generator}}$ gradually extract features from the images and preserve and propagate these features to subsequent layers via skip connections to maintain the integrity of image details.The weights of $G_{\text{Generator}}$ and $F_{\text{Generator}}$ are updated using the backpropagation algorithm to better adapt to the image features in the target domain. Through this training process, $G_{\text{Generator}}$ learns the mapping relationship between input MR images and target CT images, resulting in the generation of high-quality synthesized CT' images. Similarly, $F_{\text{Generator}}$ learns the mapping relationship between input CT images and target MR images, generating high-quality synthesized MR' images. As a result, the entire model achieves bidirectional image translation. Additionally, the hinge Loss function is employed to measure the feature distance between generated images and real images, encouraging the generators to generate more realistic and structurally consistent images. By minimizing the Hinge Loss function, the generators learn how to generate CT images that match the input MR images, ensuring structural consistency and high-quality synthesis.

TABLE 2. HLSNC-GAN's generator network structure.

Layer	Filter size	Stride	Padding	Output dimensions
Input layer				1 x 256 x 256
Conv2d-SN-ReLu	5x5	1	valid	64 x 252 x 252
Conv2d-SN-ReLu	3x3	2	same	128 x 126 x 126
Conv2d-SN-ReLu	3x3	2	same	256 x 63 x 63
Resnet Block (9 times)	3x3	1	valid	256 x 63 x 63
ConvTranspose-SN-ReLu	3x3	2	same	128 x 126 x 126
ConvTranspose-SN-ReLu	3x3	2	same	64 x 252 x 252
ConvTranspose-Tanh	5x5	1	valid	1 x 252 x 252
Output layer				1 x 256 x 256

C. HLSNC-GAN'S DISCRIMINATOR

The discriminator is another crucial component in the HLSNC-GAN network architecture, consisting of two discriminators (FIGURE 5): Discriminator D_1 (D_{CT}) and discriminator D_2 (D_{MR}). D_{CT} is responsible for assessing the authenticity of the synthesized CT' images generated by $G_{\text{Generator}}$, while D_{MR} assesses the authenticity

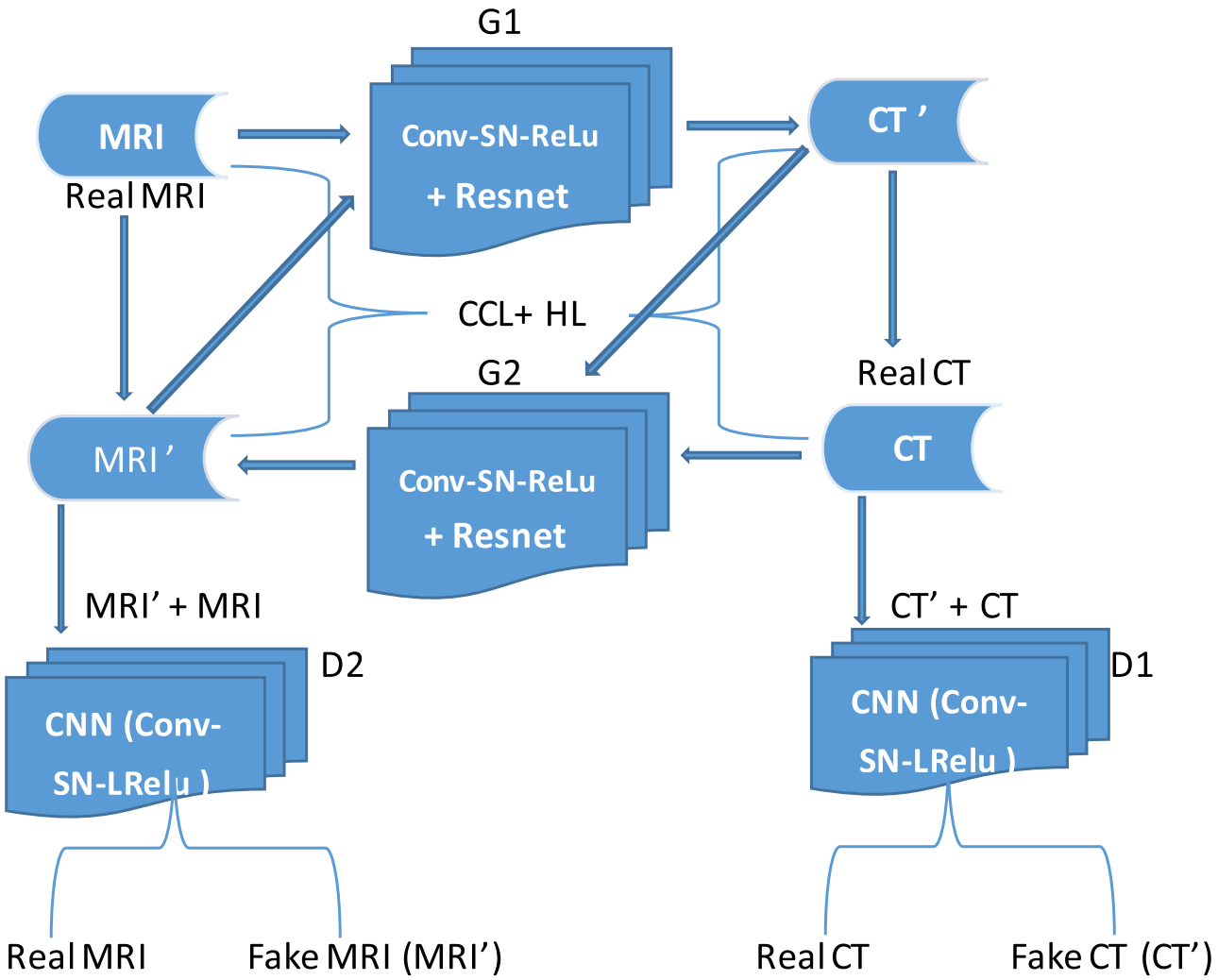

FIGURE 2. For this purpose, deep learning model frameworks such as cnn and resnet, as well as cycleGAN models in the python language, will be utilized. Our proposed HLSNC-GAN model is based on cycleGAN and introduces a new hinge loss function (HL) in the generator, along with switchable normalization (SN) in both the generator and discriminator. In the model training section, G_1 and G_2 are generation networks for MRI→sCT and CT→sMRI, respectively. The path from MRI to CT′ consists of real MRI→sCT′ → reconstructed MRI, while the path from CT to MRI′ consists of real CT→sMRI′ → reconstructed CT. Furthermore, D_1 and D_2 are CT and MRI discriminative networks. D_1 classifies synthetic CT′ and real CT, whereas D_2 classifies synthetic MRI′ and real MRI.

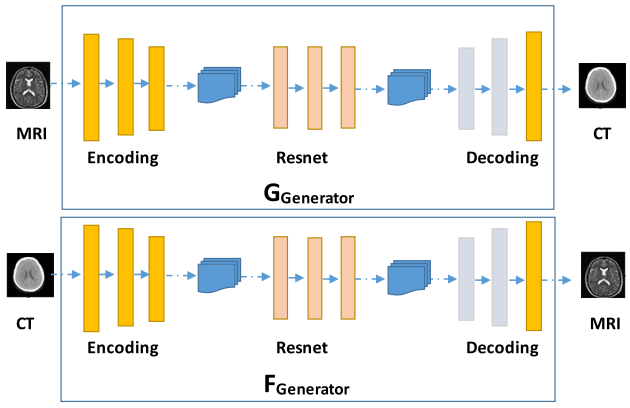

FIGURE 3. HLSNC-GAN's generator.

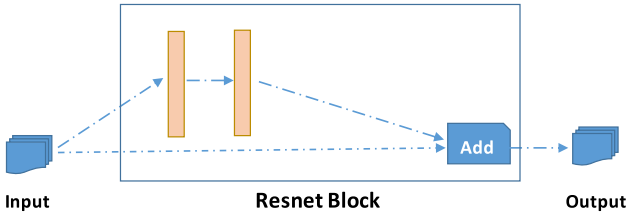

FIGURE 4. HLSNC-GAN's resnet block.

of the synthesized MR' images generated by $F_{\text{Generator}}$. Both discriminators employ switchable normalization (SN), a differentiable normalization layer that allows the model to automatically learn the normalization method for each layer, thereby enhancing the stability of model learning.In D_{CT}, it takes CT images as input and extracts features through convolutional layers, SN normalization, and leakyReLU activation functions. The objective of D_{CT} is to accurately discriminate the differences between the generated synthesized CT' images and the real CT images and classify them as real or synthesized images. Similarly, in D_{MR}, it takes MR images

TABLE 3. HLSNC-GAN's resnet block network structure.

Layer	Filter size	Stride	Padding	Output dimensions
Input layer				256 x 63 x 63
Conv2d-SN-ReLu	3x3	1	valid	256 x 63 x 63
Conv2d-SN	3x3	1	valid	256 x 63 x 63
Output layer				256 x 63 x 63

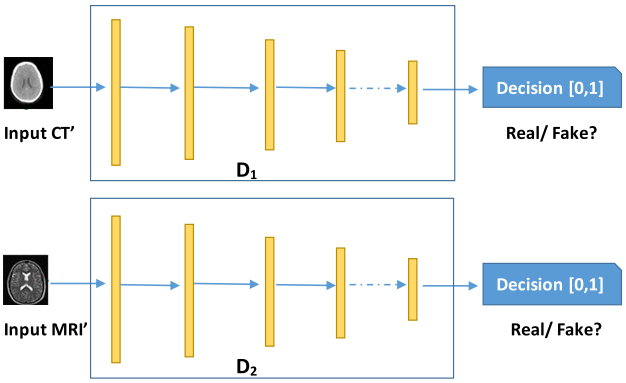

FIGURE 5. HLSNC-GAN's discriminator.

as input and extracts features through convolutional layers, SN normalization, and leakyReLU activation functions. The objective of D_{MR} is to accurately discriminate the differences between the generated synthesized MR' images and the real MR images and classify them as real or synthesized images.The training objective for discriminator D_1 and D_2 is to enhance their discriminative capabilities by minimizing the classification errors of the discriminators on the synthesized and real images. This enables the discriminators to better distinguish between the generated synthesized images and the real images. Table 4 describes the hierarchical structure of the discriminator network.By utilizing the switchable normalization method and the corresponding network architecture, D_1 and D_2 effectively determine the authenticity of the synthesized images generated by $G_{\text{Generator}}$ and $F_{\text{Generator}}$, respectively. This, in turn, promotes the learning of the generators and improves the quality of the generated images.

D. HLSNC-GAN LOSS

The loss functions used in the HLSNC-GAN model include adversarial loss, cycle consistency loss,Identity Loss and hinge loss. The descriptions of these loss functions and their combined losses are provided below:

1) ADVERSARIAL LOSS

The adversarial loss is employed to ensure that the generated images closely resemble real images, motivating the generator to produce realistic synthesized images. For the G and D_1,

TABLE 4. HLSNC-GAN's discriminator network structure.

Layer	Filter size	Stride	Padding	Output dimensions
Input layer				2 x 256 x 256
Conv2d-LeakyReLu	4x4	2	2	64 x 128 x 128
Conv2d-LeakyReLu	4x4	2	same	128 x 64 x 64
Conv2d-LeakyReLu	4x4	2	same	256 x 63 x 63
Conv2d-LeakyReLu	4x4	1	valid	512 x 63 x 63
Conv2d-sigmoid	4x4	1	same	1 x 30 x 30
Output layer				1 x 30 x 30

the adversarial loss can be expressed as (Equation 1):

$$L_{\text{adv}}(G, D_1) = E[\log(D_1(CT))] + E[\log(1 - D_1(G(MR)))] \quad (1)$$

Here, CT represents the real CT images, and MR represents the real MR images. The objective of the G is to minimize the ability of D_1 to accurately classify the generated CT' images, while the objective of discriminator D_1 is to maximize its ability to classify the generated and real images.

Similarly, for the F and D_2, the adversarial loss can be expressed as (Equation 2):

$$L_{\text{adv}}(F, D_2) = E[\log(D_2(MR))] + E[\log(1 - D_2(F(CT)))] \quad (2)$$

The objective of the F is to minimize the ability of discriminator D_2 to accurately classify the generated MR' images, while the objective of D_2 is to maximize its ability to classify the generated and real images.

2) CYCLE CONSISTENCY LOSS

The cycle consistency loss is used to enforce the consistency between the G and F, ensuring that the image transformations through both generators are reversible. The cycle consistency loss can be expressed as (Equation 3):

$$L_{\text{cycle}}(G, F) = E[\|F(G(MR)) - MR\|] + E[\|G(F(CT)) - CT\|] \quad (3)$$

Here,||.|| represents the pixel-level difference measurement, and MR and CT represent the real MR images and CT images, respectively. The objectives of the G and F are to minimize the differences between the cyclically transformed images and the original input images.

3) IDENTITY LOSS

The objective of the identity loss function is to minimize the difference between the generated composite images and the original input images, ensuring consistency between the input and output images. The following expression represents the

identity loss (Equation 4):

$$L_{\text{identity}}(G, F) = E[\|G(MR) - MR\|] + E[\|F(CT) - CT\|] \quad (4)$$

Here, G(MR) represents the composite result obtained by mapping the MR image using G to the CT image, and F(CT) represents the composite result obtained by mapping the CT image using F back to the MR image. MR and CT denote the real MR image and CT image, respectively.

4) HINGE LOSS

Hinge loss aims to reduce the difference between generated images and real images by optimizing the adversarial process between the generator and the discriminator, thereby improving the quality of synthesized images. Specifically, the hinge loss penalizes the distance between images generated by the generator and real images, prompting the generator to produce outputs closer to real images. In generative adversarial networks, the generator tries to generate realistic images, while the discriminator tries to distinguish between real and generated images. The hinge loss function optimizes this process as follows:

1. For the discriminator, the hinge loss function can be defined as: (Equation 5)

$$L_D = \max(0, 1 - D(x)) + \max(0, 1 + D(G(z))) \quad (5)$$

where $D(x)$ is the output of the discriminator for the real image x, $G(z)$ is the image generated by the generator based on random noise z, and $D(G(z))$ is the output of the discriminator for the generated image. This loss function encourages the discriminator to output values close to 1 for real images and close to -1 for generated images.

2. For the generator, the hinge loss function can be defined as: (Equation 6)

$$L_G = -\min(0, -1 + D(G(z))) \quad (6)$$

This loss function encourages the generator to produce images that make the discriminator's output close to 1, i.e., generate more realistic images.

In this way, hinge loss helps to reduce the difference between generated images and real images, thereby improving the quality of synthesized images.

5) TOTAL LOSS

The combined loss for G and F is obtained by integrating the adversarial loss, cycle consistency loss, and hinge loss. The total loss can be defined as (Equation 7):

$$\begin{aligned} L_{\text{total}}(G, F, D_1, D_2) = {} & L_{\text{adv}}(G, D_1) + L_{\text{adv}}(F, D_2) \\ & + \lambda \cdot L_{\text{cycle}}(G, F) \\ & + \lambda \cdot L_{\text{identity}}(G, F) \\ & + \lambda \cdot L_D + \lambda \cdot L_G \end{aligned} \quad (7)$$

Here, λ is a weight parameter used to control the importance of cycle consistency loss and hingeall loss in the overloss function.

E. HLSNC-GAN RMSPROP

In deep learning, an optimizer is an algorithm used to update model parameters in order to minimize the loss function. Different optimizers have their own advantages and disadvantages. In the HLSNC-GAN framework, we use the RMSProp optimizer to update the weights of the generator and discriminator to minimize the loss function between them. Specifically, we update the weights of the discriminator based on its gradients and update the weights of the generator based on its gradients. By using the RMSProp optimizer, our framework can automatically adjust the learning rate based on the magnitude of gradients of the generator and discriminator, thus adapting to the update requirements of the model's parameters. This improves the stability and convergence speed of training. Moreover, RMSProp controls the variance of parameter updates by using an exponential moving average of squared gradients, reducing the impact of early gradients on parameter updates. This helps improve the stability of the model and reduce oscillations during training. Additionally, since the gradients of the generator and discriminator are often sparse, the RMSProp optimizer can effectively handle this case and ensure appropriate updates to sparse gradients, thereby improving the performance of the model.

When using the RMSProp optimizer for the calculation of generating CT images from MR images, it is important to understand the process of parameter updates. The following are the calculation formulas for the RMSProp optimizer:

When given the model's parameters θ, the loss function $L(\theta)$, the learning rate α, and the exponential decay rate β, the parameter update process of the RMSProp optimizer is as follows:

$$g_t = \nabla(L(\theta)) \quad (8)$$

Updating the accumulated squared gradient v_i:

$$v_t = \beta \cdot v_{t-1} + (1 - \beta) \cdot g_t^2 \quad (9)$$

where β is a decay rate parameter between 0 and 1, typically chosen to be a larger value such as 0.9. Then, calculate the update quantity $\nabla\theta_t$ for the parameters.

$$\nabla\theta_t = -\frac{\alpha}{\varepsilon + \sqrt{v_t}} \cdot g_t \quad (10)$$

Here, α is the learning rate, θ is a small constant (e.g., 10^{-8}) used to avoid division by zero. Finally, update the model parameters θ:

$$\theta_{t+1} = \theta_t + \nabla\theta_t \quad (11)$$

First, for each model parameter θ_i, initialize the variable v_i, which accumulates squared gradients, to 0. During each training iteration, calculate the gradient of the model parameters θ.

By repeating the above steps, the RMSProp optimizer can adaptively adjust the learning rate and update the model parameters to better accommodate the gradient variations of different parameters. In the calculation of generating CT images from MR images, the RMSProp optimizer

continuously updates the parameters of the generator and discriminator to minimize their respective loss functions. This helps optimize our model and improve the generation capability of the generator and the discrimination capability of the discriminator.

F. HLSNC-GAN NORMALIZATION

Switchable normalization (SN) is a mechanism that dynamically selects various normalization techniques encompass methods like instance normalization, batch normalization, and layer normalization. based on the current task and data. In image synthesis, different normalization techniques can have varying impacts on the quality of the synthesized images. By utilizing the switchable normalization mechanism, the model can dynamically choose the most suitable normalization method, thereby improving the quality and consistency of the synthesized images. The principle of this mechanism is as follows:

- Combination of multiple normalization methods:The switchable normalization mechanism typically combines multiple normalization methods, such as batch normalization (BN), layer normalization (LN), and instance normalization (IN). Each of these normalization methods has its advantages and disadvantages and is suitable for different scenarios and data distributions.
- Dynamic weight allocation:In the switchable normalization mechanism, each normalization method is assigned a weight, and these weights are learnable parameters. During the training process, the model automatically adjusts these weights based on the current task and data characteristics to select the most appropriate normalization strategy.
- Improving model adaptability: Since different normalization methods perform differently in various scenarios, dynamically selecting the best normalization method allows the switchable normalization mechanism to make the model better adapt to different data distributions and task requirements, thereby improving the quality and consistency of the synthesized images.
- Reducing training difficulty: In traditional methods, selecting the appropriate normalization method may require extensive experimentation and tuning. The switchable normalization mechanism, by automatically selecting the best method, reduces the difficulty and time required for model tuning.

By dynamically choosing the most suitable normalization method for the current task and data, the switchable normalization mechanism enhances the model's adaptability and the quality of the synthesized images. In our model framework for generating CT images from MR, the implementation steps for switchable normalization are as follows:

1) Selection steps:
   - Input features:X (size:Batch $\times$ Channels $\times$ Height $\times$ Width)
   - Calculate statistical information: Mean: mean $=$ mean(X, axis $=$ (0, 2, 3)) (calculate mean along the channel dimension) Standard deviation: std $=$ std(X, axis $=$ (0, 2, 3)) (calculate standard deviation along the channel dimension)
   - Learn switchable parameters: Use fully connected layer or convolutional layer to map input features to a switchable parameter vector s, size: Channels
   - Calculate weights for switchable parameters: Weights: w $=$ softmax(s) (normalize switchable parameters using the softmax function to ensure the weights sum up to 1)
   - Select normalization method: Based on weights w, choose the normalization method:
     Batch normalization: y $=$ batch_norm(X, mean, std) (apply batch normalization to each feature channel)
     Instance normalization: y $=$ instance_norm(X) (apply instance normalization to each sample)
     Layer normalization: y $=$ layer_norm(X, mean, std) (apply layer normalization to each feature channel)
     No normalization: $y = X$
2) Normalization step: Apply the selected normalization method y to the input features X.

The formula for applying normalization during the process of generating CT images from MR images is as follows: For MR image $x_j(i, w, h)$ and CT image $y_j(i, w, h)$, where i represents the image index, w represents the image width, and h represents the image height, we can use switchable normalization to normalize the feature representations of MR and CT images. MR image normalization:

$$v_d(x_j) = \frac{1}{I_j WHC} \sum_{i=1}^{I_j} \sum_{w=1}^{W} \sum_{h=1}^{H} (x_j(i, w, h) - \min(x_j)) \tag{12}$$

$$\tau_d(x_j) = \sqrt{\frac{1}{I_j WHC} \sum_{i=1}^{I_j} \sum_{w=1}^{W} \sum_{h=1}^{H} (x_j(i, w, h) - \min(x_j))^2 + \epsilon} \tag{13}$$

CT image normalization:

$$v_d(y_j) = \frac{1}{(I_j \cdot W \cdot H \cdot C)} \sum_{i=1}^{I_j} \sum_{w=1}^{W} \sum_{h=1}^{H} (y_j(i, w, h) - \min(y_j)) \tag{14}$$

$$\tau_d(y_j) = \sqrt{\frac{1}{(I_j \cdot W \cdot H \cdot C)} \sum_{i=1}^{I_j} \sum_{w=1}^{W} \sum_{h=1}^{H} (y_j(i, w, h) - \min(y_j))^2 + \epsilon} \tag{15}$$

where: - x_j and y_j are feature maps in channel j of MR and CT images, respectively. - $v_d(x_j)$ and $v_d(y_j)$ are the means of the feature maps in channel j of MR and CT images, respectively.

- $\tau_d(x_j)$ and $\tau_d(y_j)$ are the standard deviations of the feature maps in channel j of MR and CT images, respectively. - I_j is the number of images. - W is the width of the images. - H is the height of the images. - C is the number of channels in the images. - $x_j(i, w, h)$ represents the pixel value of the i-th MR image at position (w, h). - $y_j(i, w, h)$ represents the pixel value of the i-th CT image at position (w, h). - $\min(x_j)$ and $\min(y_j)$ are the minimum values of the feature maps in channel j of MR and CT images, respectively. - ϵ is a small constant to avoid division by zero.

IV. EXPERIMENT AND RESULTS

A. HARDWARE CONFIGURATION

The experiments in this study were conducted in a hardware environment with the following specifications:

- Processor: Intel Core i9-10900K, 3.7 GHz
- Memory: 32 GB DDR4 RAM
- Graphics Card: NVIDIA GeForce RTX 3080, 10 GB GDDR6X
- Storage: 1 TB SSD + 2 TB HDD
- Operating System: Linux

B. DATA

In this study, we used two open public datasets: an unpaired dataset from Kaggle [64] and a paired dataset from related research [41], [59].

Unpaired dataset: It contains a total of 4974 cross-sectional CT and MRI scans of the brain. For the CT scans, each patient is represented by approximately 30 slices with a thickness of 5 mm. For the MRI scans, only the axial cross-sections were selected, and the images include various age groups with different slice orientations, including tumors. This dataset was randomly divided into 3486 MR-CT image pairs for the training set and 1488 pairs for the test set, with all images resized to 512 × 512 pixels.

Paired dataset: Obtained from the related literature [41], [59], it consists of a total of 367 pairs of 2D slices of T2-weighted MRI and CT images, most of which have a slice thickness of 3 mm. This dataset was randomly divided, with 290 MR-CT scan image pairs in the training set and the remaining 55 pairs used for the test set. Additional images were used for the validation set to adjust parameters, with all images also resized to 512 × 256 pixels. The image formats for both datasets include PNG and JPG.

FIGURE 6 presents several examples of experimental data to provide a more intuitive understanding of the characteristics of the datasets.

C. PERFORMANCE EVALUATION METRICS

The proposed HLSNC-GAN model will be empirically tested to generate synthetic CT images, and its performance will be evaluated mainly by using the PSNR [65], and SSIM [66]. Currently in the medical field, these evaluation metrics are widely used to evaluate the synthesis of medical images.

1) PEAK SIGNAL TO NOISE RATIO (PSNR)

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I^2}{\mathrm{MSE}}\right) \tag{16}$$

The expression MAX^2 refers to the squared value of the maximum pixel in the image. A greater PSNR value indicates a stronger resemblance between rCT and sCT.

2) STRUCTURAL SIMILARITY INDEX MEASURE (SSIM)

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{17}$$

SSIM, also known as the structural similarity index metric, is utilized to gauge the resemblance between two images. The calculation involves comparing an uncompressed, undistorted image with a distorted image. The SSIM values range from 0 to 1, with a value of 1 signifying an exact match between the two images. In the SSIM calculation, μ_x and μ_y represent the average value of the MR and the CT image. The variances of the MR and CT images are denoted by σ_x^2 and σ_y^2, respectively. σ_{xy} signifies the covariance of the MR and CT images. The constants $C1$ and $C2$ are defined as $C1 = (K1 \cdot L)^2$ and $C2 = (K2 \cdot L)^2$, respectively, with default values of $K1 = 0.01$ and $K2 = 0.03$. L represents the dynamic range of pixel values.

D. EXPERIMENT SETUP

In this study, we configured the HLSNC-GAN model environment based on cycleGAN, implemented using Python 3.7 and the tensorFlow framework. Here are the detailed configurations of the model:

- Generator structure: The generator consists of convolutional layers, 9 residual blocks, and deconvolutional layers, suitable for processing images with a resolution of 256 × 256. We set different filter sizes (including 7 × 7 and 3 × 3) and stride values (1 and 2) to meet the requirements of different layers.
- Discriminator structure: The discriminator uses an improved CNN network structure with a filter size of 4 × 4 and the same stride values as the generator.
- Optimizer: The model training utilizes the RMSprop optimizer with an exponential decay rate $\beta = 0.9$.
- Loss function: The hinge loss function is introduced, and switchable normalization (SN) layers are used to enhance the performance of the entire HLSNC-GAN model. During the loss calculation process, we set the λ value to 0.01 to control the weights.
- Normalization: For numerical stability, we added a small constant $\varepsilon = 10^{-8}$ in the normalization layer computation to prevent division by zero errors.
- Learning rate and batch size: The initial learning rate is set to 0.0002, and the batch size is set to 1.
- Training epochs: To ensure convergence, we conducted training for 500 epochs.
- Data division: We divided 80% of the dataset images into the training set, 15% for testing the model's

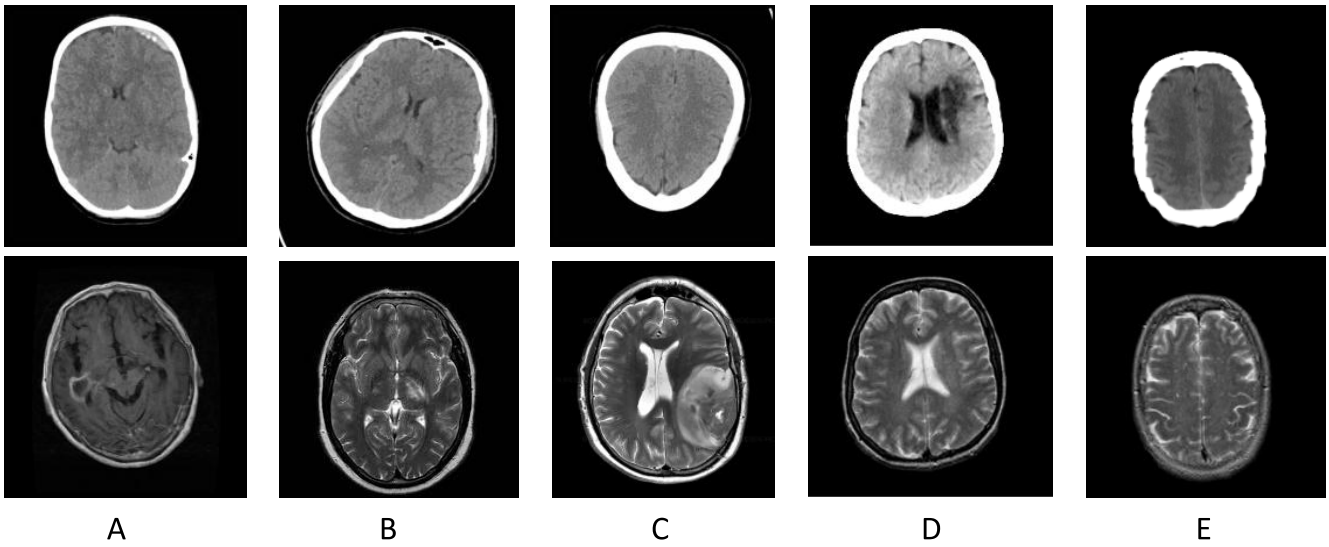

FIGURE 6. Five sets of preprocessed data were randomly selected from the two datasets. The images displayed in A, B, and C are unpaired CT and MR images selected from the Kaggle [64] dataset. The images displayed in D and E are paired CT and MR images selected from the [41] and [59] dataset. All the displayed images are denoised head medical images.

TABLE 5. In the A40B57 dataset, MR to CT images.

Method	PSNR	SSIM
Hinge+IN+RMSprop	46.133	0.818
Lsgan+SN+RMSprop	46.322	0.821
SN+RMSprop	46.776	0.854
Hinge+LN+RMSprop	47.229	0.867
Hinge+SN+Adam	48.312	0.905
Wgan+SN+RMSprop	48.623	0.913
Hinge+SN+SGD	48.703	0.915
Hinge+BN+RMSprop	49.682	0.911
HLSNC-GAN (Our)	**52.484**	**0.969**

TABLE 6. In the A40B57 dataset, CT to MR images.

Method	PSNR	SSIM
Hinge+IN+RMSprop	46.13600	0.81754
Lsgan+SN+RMSprop	46.36612	0.82324
SN+RMSprop	46.75259	0.85387
Hinge+LN+RMSprop	47.19191	0.86553
Hinge+SN+Adam	48.2192	0.90135
Wgan+SN+RMSprop	48.62404	0.91188
Hinge+SN+SGD	48.69773	0.91356
Hinge+BN+RMSprop	49.60494	0.90860
HLSNC-GAN (Our)	**51.05101**	**0.94905**

performance, and the remaining 5% as the validation set for hyperparameter tuning.

E. RESULTS

In this chapter, we will conduct a comprehensive performance comparison of our proposed HLSNC-GAN model. To begin, we will perform MRI to CT image transformations using datasets from [41] and [59], which we will denote as A40B57. We will compare different combinations of loss functions, normalization methods, and optimizers (see Table 5). Additionally, we will also perform CT to MR image transformations using the same datasets and performance evaluation methods (see Table 6). Furthermore, we will compare the performance of our model in MRI to CT (see Table 7) and CT to MRI (see Table 8) transformations with the latest models released in the field, all using the same datasets. Finally, we will employ the Kaggle dataset (as shown in Table 9) for comparisons with the latest models in MR to CT image transformation tasks.

In Table 5 and 6, we present the quantitative accuracy of different methods for generating CT and MR images on the datasets [41] and [59]. Based on the quantitative results, our experimental results indicate that our proposed HLSNC-GAN outperforms other model methods in terms of PSNR and SSIM. This is primarily attributed to the incorporation of hinge loss in the HLSNC-GAN architecture, which performs better than LSGAN and WGAN losses.

TABLE 7. In the A40B57 dataset, MR to CT images, partially referencing the research results from [41] and [59].

Method	PSNR	SSIM
[41]	21.422±3.964	0.823±0.063
U-Net	18.691	0.762
[67]	19.880	0.813
Pix2Pix	19.439	0.768
[59]	25.199	0.870
HLSNC-GAN (Our)	**52.484**	**0.969**

TABLE 8. In the A40B57 dataset, CT to MR images, partially referencing the research results from [59].

Method	PSNR	SSIM
[67]	18.66927	0.68031
[59]	26.68858	0.82622
HLSNC-GAN (Our)	**51.05101**	**0.94905**

TABLE 9. The Kaggle dataset consists of MR images transformed into CT images.

Method	PSNR	SSIM
Cycle-GAN	49.7621	0.9130
HLSNC-GAN(Our)	**50.2093**	**0.9483**

Hinge loss helps strengthen the structural consistency between input and output images, thereby enhancing image quality. Furthermore, using SN normalization yields superior results compared to common normalization methods such as IN, BN, and LN. SN allows the model to automatically adjust normalization methods based on input data, thus improving model stability and performance. Additionally, our model performs better when using the RMSprop optimizer compared to commonly used Adam and SGD optimizers. The RMSprop optimizer's adaptive learning rate feature accelerates model convergence and more effectively optimizes model parameters.

In generating CT and MR images, our model combines hinge loss, SN, and RMSprop optimizer, resulting in achieved PSNRs of 52.484 and 51.05101, and SSIMs of 0.969 and 0.94905, respectively. The superiority of HLSNC-GAN in PSNR and SSIM can be attributed to the combined effects of several factors, including improved structural consistency, adaptive normalization, and optimizer selection. Our test results demonstrate significant improvements compared to existing state-of-the-art methods.

To further validate the performance of our HLSNC-GAN model, we utilized the Kaggle dataset. Table 9 illustrates the quantitative accuracy comparison between our model and the latest research models for generating CT images on the Kaggle dataset. Our model achieved a PSNR of 50.2093 and SSIM of 0.9483. Notably, our method not only obtained favorable results on our own test set but also demonstrated advantages when compared with other recent research models. Clearly, HLSNC-GAN exhibits outstanding results compared to cycleGAN.

The traditional cycleGAN method uses L1 or L2 loss functions, which may lead to blurry synthesized images in some cases. We introduce the hinge loss function, which helps maintain the structural consistency of images, thus producing clearer and more accurate synthesized images. The hinge loss function achieves this by forcing the model to focus on important image features rather than averaging all pixel values. In traditional cycleGAN, a fixed normalization method such as batch normalization or instance normalization is usually used. However, different normalization methods may perform differently in different datasets and tasks. Our HLSNC-GAN framework introduces a switchable normalization mechanism, allowing the model to dynamically choose the most suitable normalization method based on the current task and data, thereby improving the adaptability and stability of the model. Traditional cycleGAN methods may encounter difficulties in handling multimodal medical image synthesis, especially in maintaining consistency between different modalities. The HLSNC-GAN framework effectively addresses this issue by combining the hinge loss function and switchable normalization mechanism, improving the quality and consistency of synthesized images.

We demonstrate the effectiveness of the HLSNC-GAN framework through theoretical analysis and experimental validation. By comparing with the traditional cycleGAN method, we show the advantages of HLSNC-GAN in generating high-quality, structurally consistent medical images. Furthermore, we explore the performance of HLSNC-GAN under different normalization methods and loss function choices, further confirming its application value in multimodal medical image synthesis.

V. POTENTIAL LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

A. LIMITATIONS

Despite the excellent performance of HLSNC-GAN in multimodal medical image synthesis, there are still some potential limitations and drawbacks:

- The performance of HLSNC-GAN heavily relies on the quality and diversity of the training data. In scenarios with limited data volume or poor data quality, the model may struggle to generate high-quality synthetic images.
- Due to the introduction of the hinge loss function and switchable normalization mechanism, HLSNC-GAN has a higher model complexity, which may lead to longer training times and higher computational resource requirements.
- While HLSNC-GAN performs well on specific datasets, its generalization ability to other types of medical images or different medical conditions still needs further validation.

B. COMPUTATIONAL EFFICIENCY AND SCALABILITY

- HLSNC-GAN requires relatively more computational resources, including GPU memory and processing power, which may limit its application in resource-constrained environments.
- Although HLSNC-GAN performs well on small-scale datasets, its scalability for processing large-scale datasets needs further research, especially considering the continuous growth of medical image data volumes.

C. FUTURE RESEARCH DIRECTIONS

Based on the results of this study, we propose the following future research directions:

- Explore more advanced loss functions or normalization techniques to further improve the quality and consistency of synthetic images.
- Extend the model to support image synthesis between more modalities, such as incorporating PET images into the synthesis process.

- Apply synthetic medical images to clinical scenarios, such as early diagnosis of diseases, assessment of treatment effects, and prognosis prediction.
- Use synthetic images to augment training datasets to improve the performance of other medical image analysis tasks.
- Conduct more technical validations and clinical trials to ensure the effectiveness and safety of the model.
- Collaborate with interdisciplinary teams, including radiologists, biostatisticians, and ethicists, to ensure the comprehensiveness and practicality of the research.

VI. CONCLUSION

Over the past few years, various researchers have suggested diverse medical image synthesis techniques using both conventional and deep learning methods. One prevalent approach among these methods is the utilization of GAN-based models for image synthesis particularly in the case of the two most widely used modalities, CT and MRI. Despite the considerable advancements made through deep GAN models the production of superior medical images remains constrained by various limitations. For example, Pix2Pix a u-net-based GAN model requires very rigorous image data for training where the training data must be paired. Another GAN variant, the LSGAN model, over-punishes outliers during training, resulting in insufficient sample generation which leads to a lack of ''diversity'' and pattern collapse. In addition, many GAN-based variants lack constraints on the mapping relationship between different modalities and cannot generate high-quality synthetic images. Moreover, there are too many network layers and parameters, resulting in the loss of details of the synthesized image, and the parameters are scaled during the backpropagation process. This leads to the disappearance of the gradient in the later stage of training, which affects the stability of model learning. This study aims to propose a novel framework utilizing a generalized cycleGAN approach for producing CT images from MRI data. Our proposed HLSNC-GAN framework effectively handles unpaired training data, overcoming limitations posed by scarce paired data.To enhance the correlation between various modalities and preserve the structural consistency between the input and output images, we propose a novel loss function for the generator. Additionally, we adopt switchable normalization techniques to enhance model training stability, accelerate convergence, and reduce human intervention. Furthermore, we use a new adaptive learning rate method in the optimizer to effectively prevent overfitting to the training data, Enhancing the fidelity and robustness of the generated images. We also alleviate the excessive punishment of outliers during model training, enhancing the model's learning capacity and increasing the diversity of generated samples. In addition, we utilize skip connections in the generator's ResNet architecture to preserve more image details, further improving the diversity and quality of synthesized samples. Finally, we use a leaky ReLU activation function to avoid parameter scaling during the backpropagation process, effectively preventing mode collapse and gradient explosion or vanishing problems.Through these enhancements, we achieve higher quality synthesized images. The results of this study demonstrate significant progress in medical image synthesis. Our approach outperforms existing methods [41], [59] in terms of quantitative evaluation metrics, while also exhibiting superior diversity and structural consistency. These findings highlight the potential of our framework in medical image synthesis, providing valuable support for applications such as disease prediction and treatment.In this study, we used 2D medical slice images to train our HLSNC-GAN model and compare it with other existing methods. However, with the advancement of 3D and 4D medical imaging, this study has certain limitations. In future work, we aim to focus on researching and developing medical image synthesis models suitable for higher dimensions (3D/4D) to address these limitations.

ETHICAL CONSIDERATIONS AND DATA PROTECTION

Given the sensitivity and privacy concerns of medical data, this study strictly adheres to relevant laws and regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and takes appropriate measures to ensure data security and patient privacy. We also recognize the potential risks of misusing synthetic medical images, such as misleading diagnoses or insurance fraud, and thus emphasize the need for strict ethical guidelines and regulatory scrutiny when using synthetic images.

ACKNOWLEDGMENT

(Yang Heng and Ma Yinghua are co-first authors.)

REFERENCES

[1] A. Manduca, P. V. Bayly, R. L. Ehman, A. Kolipaka, T. J. Royston, I. Sack, R. Sinkus, and B. E. Van Beers, ''MR elastography: Principles, guidelines, and terminology,'' *Magn. Reson. Med.*, vol. 85, no. 5, pp. 2377–2390, May 2021.

[2] P. J. Withers, C. Bouman, S. Carmignato, V. Cnudde, D. Grimaldi, C. K. Hagen, E. Maire, M. Manley, A. Du Plessis, and S. R. Stock, ''X-ray computed tomography,'' *Nature Rev. Methods Primers*, vol. 1, no. 1, p. 18, Feb. 2021.

[3] H. Zaidi and I. El Naqa, ''Quantitative molecular positron emission tomography imaging using advanced deep learning techniques,'' *Annu. Rev. Biomed. Eng.*, vol. 23, no. 1, pp. 249–276, Jul. 2021.

[4] M. Jamoulle, G. Kazeneza-Mugisha, and A. Zayane, ''Descriptive and narrative study of long COVID cases in general practice and diagnostic value of single photon emission computed tomography (SPECT scan),'' *medRxiv*, Mar. 2022.

[5] T. A. Soomro, L. Zheng, A. J. Afifi, A. Ali, S. Soomro, M. Yin, and J. Gao, ''Image segmentation for MR brain tumor detection using machine learning: A review,'' *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 70–90, 2023.

[6] I. Wahlang, A. K. Maji, G. Saha, P. Chakrabarti, M. Jasinski, Z. Leonowicz, and E. Jasinska, ''Brain magnetic resonance imaging classification using deep learning architectures with gender and age,'' *Sensors*, vol. 22, no. 5, p. 1766, Feb. 2022.

[7] S. Singh and R. Kumar, ''Breast cancer detection from histopathology images with deep inception and residual blocks,'' *Multimedia Tools Appl.*, vol. 81, no. 4, pp. 5849–5865, Feb. 2022.

[8] A. Krishan and D. Mittal, ''Feature based CT image registration of liver cancer,'' *Proc. Inst. Mech. Eng., H, J. Eng. Med.*, vol. 236, no. 1, pp. 3–11, Jan. 2022.

[9] J. Yuan, E. Fredman, J.-Y. Jin, S. Choi, D. Mansur, A. Sloan, M. Machtay, and Y. Zheng, "Monte Carlo dose calculation using MRI based synthetic CT generated by fully convolutional neural network for gamma knife radiosurgery," *Technol. Cancer Res. Treatment*, vol. 20, Jan. 2021, Art. no. 153303382110464.

[10] W. Boukellouz and A. Moussaoui, "Magnetic resonance-driven pseudo CT image using patch-based multi-modal feature extraction and ensemble learning with stacked generalisation," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 33, no. 8, pp. 999–1007, Oct. 2021.

[11] B. O. Macaulay, B. S. Aribisala, S. A. Akande, B. A. Akinnuwesi, and O. A. Olabanjo, "Breast cancer risk prediction in African women using random forest classifier," *Cancer Treatment Res. Commun.*, vol. 28, 2021, Art. no. 100396.

[12] S. F. Khorshid and A. M. Abdulazeez, "Breast cancer diagnosis based on K-nearest neighbors: A review," *PalArch's J. Archaeol. Egypt/Egyptol.*, vol. 18, no. 4, pp. 1927–1951, 2021.

[13] S. Kaushik, M. Bylund, C. Cozzini, D. Shanbhag, S. F. Petit, J. J. Wyatt, M. I. Menzel, C. Pirkl, B. Mehta, V. Chauhan, K. Chandrasekharan, J. Jonsson, T. Nyholm, F. Wiesinger, and B. Menze, "Region of interest focused MRI to synthetic CT translation using regression and classification multi-task network," 2022, *arXiv:2203.16288*.

[14] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[15] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, "Medical image synthesis with context-aware generative adversarial networks," in *Proc. 20th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Sep. 2017, pp. 417–425.

[16] F. Zijlstra, K. Willemsen, M. C. Florkow, R. J. Sakkers, H. H. Weinans, B. C. van der Wal, M. van Stralen, and P. R. Seevinck, "CT synthesis from MR images for orthopedic applications in the lower arm using a conditional generative adversarial network," *Proc. SPIE*, vol. 10949, Mar. 2019, Art. no. 109491J.

[17] H. Liu, M. K. Sigona, T. J. Manuel, L. M. Chen, C. F. Caskey, and B. M. Dawant, "Synthetic CT skull generation for transcranial MR imaging-guided focused ultrasound interventions with conditional adversarial networks," *Proc. SPIE*, vol. 12034, Apr. 2022, Art. no. 120340O.

[18] X. Tie, S. Lam, Y. Zhang, K. Lee, K. Au, and J. Cai, "Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients," *Med. Phys.*, vol. 47, no. 4, pp. 1750–1762, Apr. 2020.

[19] Y. Liu, L. Meng, and J. Zhong, "MAGAN: Mask attention generative adversarial network for liver tumor CT image synthesis," *J. Healthcare Eng.*, vol. 2021, pp. 1–11, Jan. 2021.

[20] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2242–2251.

[21] Z. Yi, H. Zhang, P. Tan, and M. Gong, "DualGAN: Unsupervised dual learning for image-to-image translation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2868–2876.

[22] T. Kim, M. Cha, H. Kim, J. K. Lee, and J. Kim, "Learning to discover cross-domain relations with generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 1857–1865.

[23] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. Išgum, "Deep MR to CT synthesis using unpaired data," in *Proc. 2nd Int. Workshop Simulation Synth. Med. Imag.*, Springer, 2017, pp. 14–23.

[24] H. Yang, J. Sun, A. Carass, C. Zhao, J. Lee, Z. Xu, and J. Prince, "Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN," in *Proc. 4th Int. Workshop Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis.*, Springer, 2018, pp. 174–182.

[25] Y. Liu, A. Chen, H. Shi, S. Huang, W. Zheng, Z. Liu, Q. Zhang, and X. Yang, "CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy," *Computerized Med. Imag. Graph.*, vol. 91, Jul. 2021, Art. no. 101953.

[26] S. K. Kang, H. J. An, H. Jin, J.-I. Kim, E. K. Chie, J. M. Park, and J. S. Lee, "Synthetic CT generation from weakly paired MR images using cycle-consistent GAN for MR-guided radiotherapy," *Biomed. Eng. Lett.*, vol. 11, no. 3, pp. 263–271, Aug. 2021.

[27] B. Sun, S. Jia, X. Jiang, and F. Jia, "Double U-Net CycleGAN for 3D MR to CT image synthesis," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 18, no. 1, pp. 149–156, Aug. 2022.

[28] Y. Luo, L. Zhou, B. Zhan, Y. Fei, J. Zhou, Y. Wang, and D. Shen, "Adaptive rectification based adversarial network with spectrum constraint for high-quality PET image synthesis," *Med. Image Anal.*, vol. 77, Apr. 2022, Art. no. 102335.

[29] S. Y. Selçuk, O. Dalmaz, S. U. H. Dar, and T. Çukur, "Improving image synthesis quality in multi-contrast MRI using transfer learning via autoencoders," in *Proc. 30th Signal Process. Commun. Appl. Conf. (SIU)*, May 2022, pp. 1–4.

[30] S. Amirrajab, Y. Al Khalil, C. Lorenz, J. Weese, J. Pluim, and M. Breeuwer, "Label-informed cardiac magnetic resonance image synthesis through conditional generative adversarial networks," *Computerized Med. Imag. Graph.*, vol. 101, Oct. 2022, Art. no. 102123.

[31] M. Schellenberg, J. Gröhl, K. K. Dreher, J.-H. Nölke, N. Holzwarth, M. D. Tizabi, A. Seitel, and L. Maier-Hein, "Photoacoustic image synthesis with generative adversarial networks," *Photoacoustics*, vol. 28, Dec. 2022, Art. no. 100402.

[32] H. Emami, M. Dong, S. P. Nejad-Davarani, and C. K. Glide-Hurst, "Generating synthetic CTs from magnetic resonance images using generative adversarial networks," *Med. Phys.*, vol. 45, no. 8, pp. 3627–3636, Aug. 2018.

[33] S. Kazemifar, A. M. B. Montero, K. Souris, S. T. Rivas, R. Timmerman, Y. K. Park, S. Jiang, X. Geets, E. Sterpin, and A. Owrangi, "Dosimetric evaluation of synthetic CT generated with GANs for MRI-only proton therapy treatment planning of brain tumors," *J. Appl. Clin. Med. Phys.*, vol. 21, no. 5, pp. 76–86, May 2020.

[34] H. Emami, M. Dong, and C. K. Glide-Hurst, "Attention-guided generative adversarial network to address atypical anatomy in synthetic CT generation," in *Proc. IEEE 21st Int. Conf. Inf. Reuse Integr. Data Sci. (IRI)*, Aug. 2020, pp. 188–193.

[35] H. Sun, Z. Lu, R. Fan, W. Xiong, K. Xie, X. Ni, and J. Yang, "Research on obtaining pseudo CT images based on stacked generative adversarial network," *Quant. Imag. Med. Surgery*, vol. 11, no. 5, pp. 1983–2000, May 2021.

[36] X. Liu, H. Emami, S. P. Nejad-Davarani, E. Morris, L. Schultz, M. Dong, and C. K. Glide-Hurst, "Performance of deep learning synthetic CTs for MR-only brain radiation therapy," *J. Appl. Clin. Med. Phys.*, vol. 22, no. 1, pp. 308–317, Jan. 2021.

[37] M. Gotoh, T. Nakaura, Y. Funama, K. Morita, D. Sakabe, H. Uetani, Y. Nagayama, M. Kidoh, M. Hatemura, T. Masuda, and T. Hirai, "Virtual magnetic resonance lumbar spine images generated from computed tomography images using conditional generative adversarial networks," *Radiography*, vol. 28, no. 2, pp. 447–453, May 2022.

[38] B. Tang, F. Wu, Y. Fu, X. Wang, P. Wang, L. C. Orlandini, J. Li, and Q. Hou, "Dosimetric evaluation of synthetic CT image generated using a neural network for MR-only brain radiotherapy," *J. Appl. Clin. Med. Phys.*, vol. 22, no. 3, pp. 55–62, Mar. 2021.

[39] H. Emami, M. Dong, S. P. Nejad-Davarani, and C. K. Glide-Hurst, "SA-GAN: Structure-aware GAN for organ-preserving synthetic CT generation," in *Proc. 24th Int. Conf. Med. Image Comput. Comput. Assist. Intervent.* Springer, 2021, pp. 471–481.

[40] M. Qi, Y. Li, A. Wu, Q. Jia, B. Li, W. Sun, Z. Dai, X. Lu, L. Zhou, X. Deng, and T. Song, "Multi-sequence MR image-based synthetic CT generation using a generative adversarial network for head and neck MRI-only radiotherapy," *Med. Phys.*, vol. 47, no. 4, pp. 1880–1894, Apr. 2020.

[41] A. Ranjan, D. Lalwani, and R. Misra, "GAN for synthesizing CT from T2-weighted MRI data towards MR-guided radiation treatment," *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 35, no. 3, pp. 449–457, Jun. 2022.

[42] D. Cusumano, J. Lenkowicz, C. Votta, L. Boldrini, L. Placidi, F. Catucci, N. Dinapoli, M. V. Antonelli, A. Romano, V. De Luca, G. Chiloiro, L. Indovina, and V. Valentini, "A deep learning approach to generate synthetic CT in low field MR-guided adaptive radiotherapy for abdominal and pelvic cases," *Radiotherapy Oncol.*, vol. 153, pp. 205–212, Dec. 2020.

[43] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5967–5976.

[44] J. D. S. T. Cordeiro and J. H. Saito, "Pix2pix network for fingerprint texture image synthesis," in *Proc. Anais 17th Workshop Visão Computacional (WVC)*, Nov. 2021, pp. 13–18.

[45] M. Lupión, J. F. Sanjuan, and P. M. Ortigosa, "Using a multi-GPU node to accelerate the training of Pix2Pix neural networks," *J. Supercomput.*, vol. 78, no. 10, pp. 12224–12241, Jul. 2022.

[46] M. S. K. Raghavendra and P. N. Sarappadi, "Transfer learning with pix2pix gan for generating realistic photographs from viewed sketch arts," *J. Southwest Jiaotong Univ.*, vol. 57, no. 4, pp. 197–207, Aug. 2022.
[47] B. K. S. Isaac-Medina, N. Bhowmik, C. G. Willcocks, and T. P. Breckon, "Cross-modal image synthesis within dual-energy X-ray security imagery," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2022, pp. 332–340.
[48] A. Aljohani and N. Alharbe, "Generating synthetic images for healthcare with novel deep pix2pix GAN," *Electronics*, vol. 11, no. 21, p. 3470, Oct. 2022.
[49] S. Phukan, J. Singh, R. Gogoi, S. Dhar, and N. D. Jana, "COVID-19 chest X-ray image generation using ResNet-DCGAN model," in *Proc. Adv. Intell. Comput. Commun.*, Springer, 2021, pp. 227–234.
[50] S. Ellis, O. E. M. Manzanera, V. Baltatzis, I. Nawaz, A. Nair, L. Le Folgoc, S. Desai, B. Glocker, and J. A. Schnabel, "Evaluation of 3D GANs for lung tissue modelling in pulmonary CT," 2022, *arXiv:2208.08184*.
[51] Y. Huang, Q. Wang, and S. Omachi, "Rethinking degradation: Radiograph super-resolution via aid-srgan," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, Springer, 2022, pp. 43–52.
[52] Y. Li, H.-C. Shao, X. Liang, L. Chen, R. Li, S. Jiang, J. Wang, and Y. Zhang, "Zero-shot medical image translation via frequency-guided diffusion models," 2023, *arXiv:2304.02742*.
[53] H. Lei, Z. Tian, H. Xie, B. Zhao, X. Zeng, J. Cao, W. Liu, J. Wang, G. Zhang, S. Wang, and B. Lei, "LAC-GAN: Lesion attention conditional GAN for ultra-widefield image synthesis," *Neural Netw.*, vol. 158, pp. 89–98, Jan. 2023.
[54] V. Kearney, B. P. Ziemer, A. Perry, T. Wang, J. W. Chan, L. Ma, O. Morin, S. S. Yom, and T. D. Solberg, "Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks," *Radiol, Artif. Intell.*, vol. 2, no. 2, Mar. 2020, Art. no. e190027.
[55] K. N. D. B. Boni, J. Klein, A. Gulyban, N. Reynaert, and D. Pasquier, "Improving generalization in MR-to-CT synthesis in radiotherapy by using an augmented cycle generative adversarial network with unpaired data," *Med. Phys.*, vol. 48, no. 6, pp. 3003–3010, Jun. 2021.
[56] A. Abu-Srhan, I. Almallahi, M. A. M. Abushariah, W. Mahafza, and O. S. Al-Kadi, "Paired-unpaired unsupervised attention guided GAN with transfer learning for bidirectional brain MR-CT synthesis," *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104763.
[57] H. Sun, Q. Xi, R. Fan, J. Sun, K. Xie, X. Ni, and J. Yang, "Synthesis of pseudo-CT images from pelvic MRI images based on an MD-CycleGAN model for radiotherapy," *Phys. Med. Biol.*, vol. 67, no. 3, Feb. 2022, Art. no. 035006.
[58] J. Wang, B. Yan, X. Wu, X. Jiang, Y. Zuo, and Y. Yang, "Development of an unsupervised cycle contrastive unpaired translation network for MRI-to-CT synthesis," *J. Appl. Clin. Med. Phys.*, vol. 23, no. 11, p. e13775, Nov. 2022.
[59] J. Wang, Q. M. J. Wu, and F. Pourpanah, "DC-CycleGAN: Bidirectional CT-to-MR synthesis from unpaired data," *Computerized Med. Imag. Graph.*, vol. 108, Sep. 2023, Art. no. 102249.
[60] G. Alwakid, W. Gouda, and M. Humayun, "Deep learning-based prediction of diabetic retinopathy using CLAHE and ESRGAN for enhancement," *Healthcare*, vol. 11, no. 6, p. 863, Mar. 2023.
[61] M. Iskandar, H. Mannering, Z. Sun, J. Matthew, H. Kerdegari, L. Peralta, and M. Xochicale, "Towards realistic ultrasound fetal brain imaging synthesis," 2023, *arXiv:2304.03941*.
[62] G. Huang and A. H. Jafari, "Enhanced balancing GAN: Minority-class image generation," *Neural Comput. Appl.*, vol. 35, no. 7, pp. 5145–5154, Mar. 2023.
[63] D. Ma and Y. Li, "Image super-resolution via gradient guidance and gnenrative adversarial network," *J. Comput. Electron. Inf. Manage.*, vol. 10, no. 1, pp. 12–16, Feb. 2023.
[64] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: An overlooked learning opportunity," *Int. J. Forecasting*, vol. 37, no. 2, pp. 587–603, Apr. 2021.
[65] M. H. N. Azam, F. Ridzuan, and M. N. S. M. Sayuti, "A new method to estimate peak signal to noise ratio for least significant bit modification audio steganography," *Pertanika J. Sci. Technol.*, vol. 30, no. 1, pp. 497–511, Jan. 2022.
[66] I. Bakurov, M. Buzzelli, R. Schettini, M. Castelli, and L. Vanneschi, "Structural similarity index (SSIM) revisited: A data-driven approach," *Expert Syst. Appl.*, vol. 189, Mar. 2022, Art. no. 116087.
[67] L. Kong, C. Lian, D. Huang, Z. Li, Y. Hu, and Q. Zhou, "Breaking the dilemma of medical image-to-image translation," in *Proc. 35th Conf. Neural Inf. Process. Syst.*, vol. 34, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds. Red Hook, NY, USA: Curran Associates, 2021, pp. 1964–1978. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2021/file/0f2818101a7ac4b96ceeba38de4b934c-Paper.pdf

YANG HENG is currently pursuing the Ph.D. degree with COMSATS University Islamabad, Abbottabad Campus. His research interests include machine learning, image processing, and bioinformatics.

MA YINGHUA is currently pursuing the Ph.D. degree with COMSATS University Islamabad, Abbottabad Campus. Her research interests include machine learning, bioinformatics, and image processing.

FIAZ GUL KHAN received the master's (specialization) and Ph.D. degrees from Politecnico di Torino, Italy, in 2013. He is currently a Tenured Associate Professor with the Computer Science Department, COMSATS University Islamabad, Abbottabad Campus, Pakistan. His research interests include the field of concurrent computing machine learning, artificial intelligence, and GPU computing.

AHMAD KHAN received the Ph.D. degree from the National University of Computer and Emerging Sciences (FAST-NU), Islamabad, Pakistan, in 2015. He is currently an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad (CUD), Abbottabad Campus. His research interests include computer vision, machine learning, and evolutionary algorithms.

ZENG HUI is currently pursuing the Ph.D. degree with COMSATS University Islamabad, Abbottabad Campus. His research interests include computer networks, image processing, and security encryption technology.