A Review of GAN-Based Brain Image Synthesis Using MRI and CT for Deep Learning Applications

Ahmed S. El-Hossiny[1] and Mostafa El-Hussien Ibrahem [1]*

ABSTRACT

The generation of synthetic medical images, particularly for brain imaging using magnetic resonance imaging (MRI) and computed tomography (CT), has gained significant attention due to its potential to enhance diagnostic accuracy, reduce healthcare costs, and minimize patient exposure to radiation. Generative Artificial Intelligence (AI) models, especially Generative Adversarial Networks (GANs), have demonstrated exceptional promise in addressing key challenges in synthetic image generation. This review examines the role of GANs in producing synthetic brain images from magnetic resonance imaging (MRI) and computed tomography (CT) data, emphasizing their ability to generate realistic and diverse datasets for the training of advanced machine learning algorithms. Particular attention is given to issues such as the necessity for large, well-annotated datasets, the impact of paired versus unpaired data, dataset size, and the effectiveness of various GAN architectures and other deep learning (DL) techniques in brain modality translation. Furthermore, we compare the performance of different models using widely adopted metrics, including mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM), based on literature published between 2017 and 2023.

Keywords: Deep learning, Generative Adversarial Networks, Magnetic resonance imaging, Paired data, Unpaired data, Synthetic computed tomography

I. INTRODUCTION

The advent of synthetic image generation and image-to-image translation techniques has reshaped the landscape of medical imaging, particularly with the integration of artificial intelligence (AI) methodologies such as machine learning and deep learning (DL). These innovations have enabled the creation of advanced medical applications and services, facilitating progress in areas that traditionally faced limitations due to data scarcity and technological constraints. Generative models have emerged as powerful approaches for tackling a range of complex medical imaging challenges, including denoising, reconstruction, segmentation, simulation, detection, and classification. Their ability to synthesize highly realistic medical images offers a promising avenue to mitigate the persistent shortage of annotated datasets.

Practical applications of synthetic imaging methods are extensive. These include enhancing the image quality of low-dose computed tomography (CT) and positron emission tomography (PET) scans, reconstructing CT volumes from multi-view X-ray images, integrating multiple imaging modalities through fusion processes, generating synthetic 3T magnetic resonance imaging (MRI) scans from 1.5T data, and producing synthetic CT images from MRI scans or vice versa for various anatomical regions such as the brain, spinal cord, chest, and hip [1], [2].

[1] Biomedical and Systems Engineering Department, Higher Institute of Engineering, El-Shorouk Academy, Cairo, 11837, Egypt

* Corresponding Author

Several factors motivate the adoption of synthetic imaging techniques, including the reduction of healthcare costs, minimization of patient waiting times, and mitigation of risks associated with specific imaging modalities. For example, the generation of synthetic images can improve diagnostic quality when confronted with noise, artifacts, low radiation doses, or device limitations. Additionally, synthetic modalities may be employed to avoid the adverse effects of certain imaging systems or to extract enhanced diagnostic information from available data.

Developers in the field of medical AI continuously strive to design generative models capable of producing synthetic medical images that are both accurate and computationally efficient. To achieve clinically meaningful outcomes, two critical aspects must be addressed during implementation:

A. Selection of the Optimal Dataset:

The foundation of any model lies in the quality and representativeness of the training dataset. High-quality, comprehensive datasets enable models to generate realistic outputs. Metadata such as patient demographics, imaging modality, acquisition parameters, and image resolution are essential in this context. Datasets may be acquired directly from hospitals and imaging centers or sourced from established repositories such as The Cancer Imaging Archive (TCIA), FreeSurfer, and OpenNEURO. In some instances, hybrid approaches combining institutional and public datasets are employed to maximize data diversity. Notably, Figure 1 paired datasets are generally preferred over unpaired datasets, as they facilitate more accurate prediction of synthetic images by providing direct reference mappings. Data augmentation techniques may also be applied under controlled conditions to expand training datasets while minimizing the risk of overfitting.

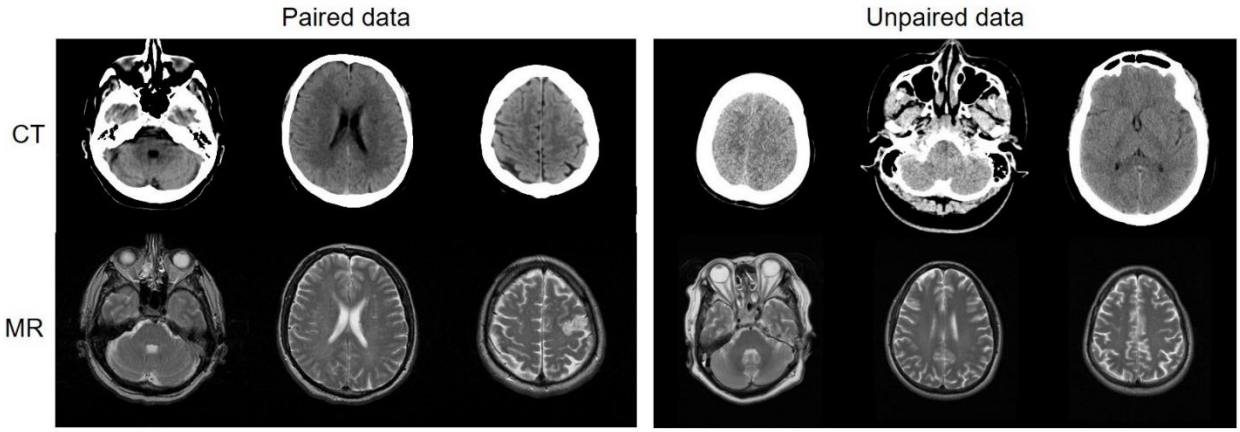

Figure 1 The difference between each paired and unpaired dataset for the brain region, images from the Institutional Review Board (IRB) [3].

B. Choice of Model Architecture:

The selection of a suitable DL architecture is pivotal in determining predictive accuracy and clinical utility. A recent review of 84 studies on MRI-to-CT synthesis for radiotherapy identified U-Net and GAN-based frameworks as the most effective, though performance varied with anatomical region and model configuration. Prominent GAN variants include pix2pix (employing a U-Net generator and PatchGAN discriminator), conditional GANs (cGANs) that incorporate label information, deep convolutional GANs (DCGANs) based on convolutional architectures, Wasserstein GANs (WGANs) with enhanced loss functions, and CycleGANs designed for unpaired image-to-image translation. Each of these architectures incorporates two core components: a generator (responsible for producing synthetic data) and a discriminator (tasked

with distinguishing real from synthetic data). The iterative competition between these components progressively improves the realism of the generated images. New architectures continue to evolve, often by modifying generator–discriminator structures, increasing model depth, or refining loss functions [4].

Choosing a DL model that fits the data well and produces reliable predictions is considered the most important part because it depends on the programmer's skills and the information he possesses. DL model examples, such as generative adversarial network (GAN/GANs) and their models that extend on it, like pixel-to-pixel (pix2pix) (where generator depend on U-Net and discriminator use PatchGAN)[5], conditional or controllable generative adversarial network (cGAN) (which is extends on classing the input labels) [6], deep convolutional generative adversarial network (DCGAN) (extends on CNN architecture), or models that scale up GANs architecture, such as Wasserstein generative adversarial network (WGAN) (discriminator training on data at less time than generator, by enhancing loss function) [7], CycleGAN (a generative model that translates images from source to target) [8],[9],[7].

Anyone who works on GANs has two main components. The first part is called the generator (G) which learns to generate new random data. The second part called the discriminator (D) learns to classify the generator's fake and real data, those processes or game between D and G rebate in this time generator is updating from impact of the loss function, until the discriminator can’t be able to distinguish fake data. The target of any generative model is to provide the best measurable results and outcomes closest to realism. Alternatively, a new model is designed based on the previously mentioned methods by changing the nature of both generator and discriminator with an increasing number of the model’s layers and functions, enhancing the model's loss function of generator and discriminator.

A review was conducted to identify studies that utilized brain data in the development of computational models. To ensure a comprehensive search, records were retrieved from ten scientific databases, including AAAI, EKB, Elsevier, IEEE, Google Schoolar, PubMed, ResearchGate, Springer, MDPI, and Wiley. In addition, six relevant study registers were searched. This process yielded a total of 66 records. The keywords used in the search for these papers were GAN, brain CT to MRI, MRI to CT, and synthesis medical images. During the search process, we received to number over fifty papers and publications. After careful consideration, we excluded the ones that did not specialize in the brain and head region and Anonymous, as well as those that were published before 2017

Prior to the screening phase, 16 records were removed. Of these, eight were identified as duplicates, while the remaining eight were excluded for other reasons, such as incomplete metadata or irrelevant document types. The remaining 60 records were subjected to title and abstract screening to assess their initial relevance to the review objectives.

Following the screening process, 15 studies were excluded due to a lack of alignment with the research scope. The remaining 45 full-text articles were then assessed for eligibility based on predefined inclusion and exclusion criteria. During this stage, nine studies were excluded because they did not utilize brain data to build the models, which was a fundamental requirement for inclusion. An additional six studies were excluded due to the use of unconventional methods that yielded poor results or lacked scientific rigor. Ultimately, 30 studies satisfied all inclusion criteria and were included in the final synthesis of this review. The detailed selection and screening procedure are visually represented in the flow diagram (Figure 2).

In this review, we analyze 30 representative studies focusing on the synthesis of CT images from MRI and MRI from CT for brain imaging applications. The motivations identified in these works include reducing risks associated with ionizing radiation, lowering financial and temporal costs, leveraging the strengths of each imaging modality (e.g., MRI for soft tissue characterization, CT for bone imaging and electron density), and overcoming modality-specific limitations in certain clinical scenarios [10]. These factors, detailed in subsequent sections, highlight the clinical and technical significance of synthetic brain imaging in modern healthcare

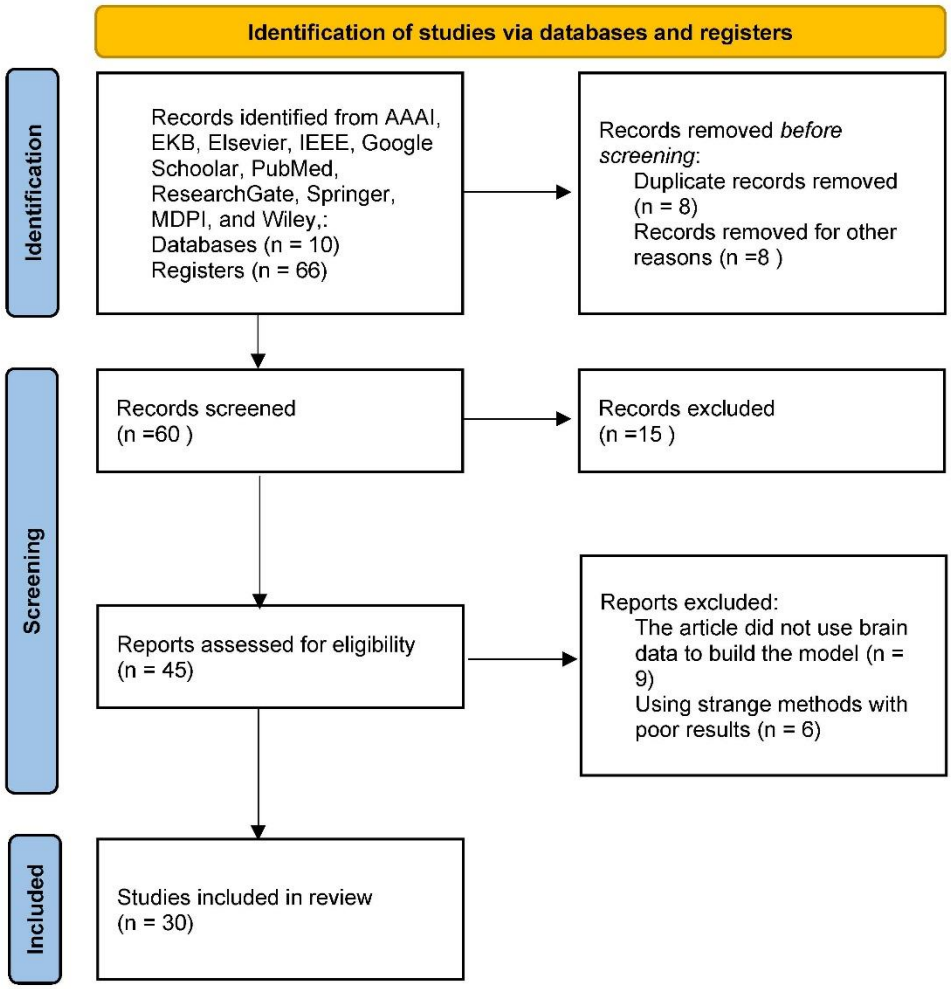

Figure 2 Flow diagram of the retrieval and selection process

II. BACKGROUND

Figure 3, provides an insightful representation of the trends in research on synthetic brain images over recent years, with a clear dominance of projects focused on converting CT to MRI images. Analysis of the data reveals that, on average, three to four synthetic brain imaging studies are published each year. Notably, the creation of synthetic CT images appears to be more prevalent than synthetic MRI images, suggesting a higher level of research activity in this area. This trend may be attributed to several factors, including the relative complexity of converting MRI data or the focus on reducing radiation exposure associated with CT scans, especially in regions where access to MRI equipment may be limited due to economic constraints. Additionally, the scarcity of research aimed at converting CT to MRI images further emphasizes the challenges and limitations associated with this process.

The bar graph on the left side of Figure 3 illustrates the annual distribution of research papers from 2017 to 2023, categorized by: MR to CT, CT to MR, and Both. It is evident that the majority of studies focus on the translation from MR to CT, with the highest output occurring in 2020, where six papers were published, followed by a slightly reduced number of publications in 2019 (four papers). In contrast, the CT to MR category shows significantly fewer papers, with only a few notable studies emerging in 2019, 2020, 2022 and 2023. The Both categories, addressing both translation directions in a single study, remains the least explored, as reflected by the small number of papers published in this area, with a brief peak in 2019.

The pie chart on the right side of Figure 3 provides a comprehensive breakdown of the total number of papers across all years, highlighting the dominance of the MR to CT translation, which accounts for 70% of the research output. Studies focusing on the translation from CT to MR represent only 17% of the total publications, and those addressing both translation directions comprise 13%. This distribution underscores the growing interest in converting MR to CT images, likely due to the clinical utility of CT scans in emergency and skeletal imaging, whereas the translation from CT to MRI, though an emerging area, remains less explored. These trends suggest that while MR to CT conversion is the focal point of most research efforts, there is a need for further exploration into the challenges of CT to MR conversion and the potential benefits of combining both modalities in a single framework.

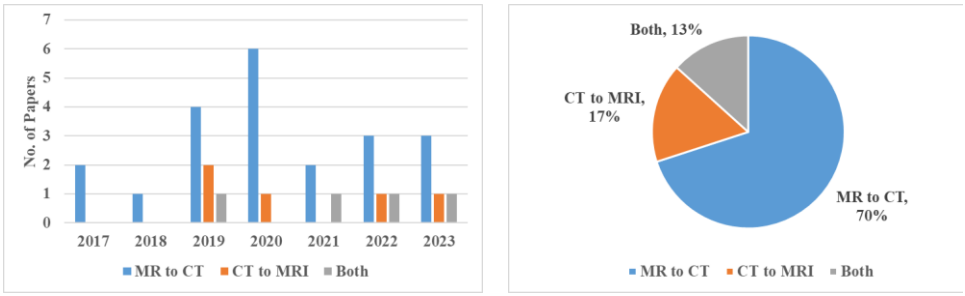

Figure 3 Distribution of research papers on synthetic brain images across different modalities.

A. MEDICAL BACKGROUND

The utilization of synthetic CT and MRI imaging has been widely reported in the literature, primarily due to its potential to mitigate patient risk, lower financial expenditure, and shorten diagnostic delays. A key motivation for generating synthetic CT images from MRI data is the elimination of exposure to ionizing radiation [10], while preserving the anatomical benefits of CT, including optimal visualization of osseous structures [11], and quantitative electron density estimation [12], [13]. Conversely, synthesizing MRI data from CT images is often motivated by the lower acquisition cost and faster clinical availability of CT, which serves as an efficient preliminary diagnostic modality. Nevertheless, MRI offers superior tissue contrast and diagnostic precision, particularly for pathologies such as neoplasms and cerebrovascular events [14], [15]. The adoption of MRI can be constrained in rural or resource-limited healthcare facilities lacking access to MRI equipment [16], and its higher procedural cost compared to CT, coupled with contraindications for individuals bearing metallic implants or pacemakers [3], further limits its applicability.

B. SOFTWARE BACKGROUND

The most widely used methods for brain image translation are pix2pix, cGAN, CycleGAN, DCGAN, and U-Net. Spatial papers and articles focused on forming new structures or combining convolutional layers to achieve the best results. Let's take two examples to illustrate this point about the changing on CycleGAN loss function:

$$L_{cycle}(G,F) = Ex_{\sim pdata}(x)\left|\left(G(x)\right) - x\right|_1 + Ey_{\sim pdata}(y)\left|G\left(F(y)\right) - y\right|_1$$

$$L_{DC\text{-}cycle}(G,F,D_X,D_Y) = L_{GAN}(G,D_Y,X,Y) + L_{GAN}(F,D_X,Y,X) + \alpha L_{cycle}(G,F)\ [8]$$

dual contrast CycleGAN (DC-cycleGAN), which is used to synthesize MRI images from CT and vice versa with two generators and two Discriminators like CycleGAN and consist of DC loss. DC-cycleGAN loss function:

$$L_{DC-cycle}(G,F,D_X,DY) = L_{GAN}(G,D_Y,X,Y) + L_{GAN}(F,D_X,Y,X) + \beta L_{DC}(D_Y,X,Y) + \beta L_{DC}(D_X,Y,X) + \alpha L_{scycle}(G,F)$$

$$L_{scycle}(G,F) = SSIM\ (F(G(x)),x) + SSIM(F\left(G(y)\right),y)$$

Where β and α control the equation weight, the method achieves better results than cycleGAN, where the results for MRI synthesis are MAE = 0.04559, PSNR = 26.68858, and SSIM = 0.82622 for the DC-cycleGAN model, and MAE = 0.09155, PSNR = 20.63825, and SSIM = 0.71670 for the ordinary cycleGAN model [17]. Also, DC-cycleGAN achieves higher results than CT synthetic.

sc-cycleGAN, which was developed for unpaired synthetic CT images from MRI, was a model developed with two generators, G_{CT} and G_{MR}, and two discriminators, D_{CT} and D_{MR}. This method aims to provide better results. The CycleGAN loss function for this model is: $L_{cycle}\ (G_{CT}, G_{MR}) = |G_{CT}(G_{MR}(I_{CT})) - I_{CT}|_1 + |G_{MR}(G_{CT}(I_{MR})) - I_{MR}|_1$ and sc-cycleGAN loss function:

$$L(G_{CT}, G_{MR}, D_{CT}, D_{MR}) = L_{adv}(G_{CT}, D_{CT}) + L_{adv}(G_{MR}, D_{MR}) + \lambda_1 L_{cycle}(G_{CT}, G_{MR}) + \lambda_2 L_{struct}(G_{CT}, G_{MR})$$

Where: $L_{adv}(G_{CT}, D_{CT}) = D_{CT}(G_{CT}(I_{MR}))_2 + (1 - D_{CT}(I_{CT}))_2$ and $L_{adv}(G_{MR}, D_{MR}) = D_{MR}(G_{MR}(I_{CT}))_1\ (1 - D_{MR}(I_{MR}))_2$ results that sc-cycleGAN achieves compared to cycleGAN MAE = 127.59, PSNR = 24.41, and SSIM = 0.773 to MAE = 143.78, PSNR = 23.73, and SSIM = 0.558 [12].

There are also other examples of improvement models not CycleGAN, like Auto-GAN implemented by an autoencoder network and GAN generator [18], MCMP-GAN with five layers consisting of convolutional levels based on U-Net connect with generator and based on the DCGAN method of producing synthetic images [19], and JP-DAM-GAN (jigsaw puzzle-Discriminating Attenuated model- Generative Adversarial Networks) where that model is based on the pix2pix method to convert between MRI modalities [20].

III. RESULTS AND DISCUSSION

In evaluating the performance and efficiency of generative models, it is essential to employ both qualitative and quantitative assessment methods. Qualitative evaluation is typically conducted through visual inspection by radiologists or specialized medical professionals, who assess the clinical relevance and realism of the generated images. Complementing this, quantitative analysis involves the use of mathematical equations and performance metrics, which provide a more objective measure of the model's accuracy and effectiveness in generating synthetic images. This combined approach ensures a comprehensive evaluation of the model's capabilities and its potential for clinical applications.

A. MATHEMATICAL EVALUATION METRICS

Evaluation of synthetic image model performance can be achieved by applying measurement metrics as quantitative indicators of the results. Commonly used equations include mean absolute error (MAE) (1), mean square error (MSE) (2), peak signal-to-noise ratio (PSNR) (3), and structural similarity index measure (SSIM) (4). There are also equations such as dice similarity coefficient (DSC) (5), root mean square error (RMSE) (6), relative global error (ERGAS) (7), universal quality index (UQI) (8), and spatial correlation coefficient (SCC) (9).

$$\mathbf{MAE} = \frac{\mathbf{1}}{\mathbf{N \times M}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{j=1}}^{\mathbf{M}} \left| \mathbf{I(i,j)} - \check{\mathbf{I}}\mathbf{(i,j)} \right|_{\mathbf{i}} \tag{1}$$

$$\mathbf{MSE} = \frac{\mathbf{1}}{\mathbf{N \times M}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{M}} \left(\mathbf{I(i,j)} - \check{\mathbf{I}}\mathbf{(i,j)} \right)_{\mathbf{i}}^{\mathbf{2}} \tag{2}$$

$$\mathbf{RMSE} = \sqrt{\mathbf{MSE}} \tag{3}$$

$$\mathbf{PSNR} = \mathbf{10} \log\left(\frac{\mathbf{MAX^2}}{\mathbf{MSE}} \right) = \mathbf{20} \log\left(\frac{\mathbf{MAX}}{\sqrt{\mathbf{MSE}}} \right) \tag{4}$$

$$\mathbf{SSIM} = \left(\frac{(\mathbf{2}\,\boldsymbol{\mu}_{\mathbf{I}}\,\boldsymbol{\mu}_{\check{\mathbf{I}}} + \mathbf{C_1})(\mathbf{2}\,\boldsymbol{\sigma}_{\mathbf{I}\check{\mathbf{I}}} + \mathbf{C_2})}{(\boldsymbol{\mu}_{\mathbf{I}}^{\mathbf{2}} + \boldsymbol{\mu}_{\check{\mathbf{I}}}^{\mathbf{2}} + \mathbf{C_1})(\boldsymbol{\sigma}_{\mathbf{I}}^{\mathbf{2}} + \boldsymbol{\sigma}_{\check{\mathbf{I}}}^{\mathbf{2}} + \mathbf{C_2})} \right) \tag{5}$$

$$\mathbf{DSC} = \frac{\mathbf{2} \times |\mathbf{A} \cap \mathbf{B}|}{|\mathbf{A}| + |\mathbf{B}|} \tag{6}$$

$$\mathbf{ERGAS} = \mathbf{100}\, \frac{\mathbf{R}_{\mathbf{synthetic\ CT}}}{\mathbf{R}_{\mathbf{actual\ CT}}} \sqrt{\frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{i=1}}^{\mathbf{n}} \left(\frac{\mathbf{RMSE^2}}{\boldsymbol{\mu}_{\check{\mathbf{I}}}^{\mathbf{2}}} \right)_{\mathbf{i}}} \tag{7}$$

$$\mathbf{UQI} = \frac{\mathbf{4}\boldsymbol{\sigma}_{\mathbf{I}\check{\mathbf{I}}}\,\boldsymbol{\mu}_{\mathbf{I}}\,\boldsymbol{\mu}_{\check{\mathbf{I}}}}{(\boldsymbol{\mu}_{\mathbf{I}}^{\mathbf{2}} + \boldsymbol{\mu}_{\check{\mathbf{I}}}^{\mathbf{2}})(\boldsymbol{\sigma}_{\mathbf{I}}^{\mathbf{2}} + \boldsymbol{\sigma}_{\check{\mathbf{I}}}^{\mathbf{2}})} \tag{8}$$

$$\mathbf{SCC} = \frac{\sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{M}} (\mathbf{I(i,j)} - \boldsymbol{\mu}_{\mathbf{I}})\,(\check{\mathbf{I}}\mathbf{(i,j)} - \boldsymbol{\mu}_{\check{\mathbf{I}}})}{\sqrt{\sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{M}} (\mathbf{I(i,j)} - \boldsymbol{\mu}_{\mathbf{I}})^{\mathbf{2}} \sum_{\mathbf{i=1}}^{\mathbf{N}} \sum_{\mathbf{i=1}}^{\mathbf{M}} \left(\check{\mathbf{I}}\mathbf{(i,j)} - \boldsymbol{\mu}_{\check{\mathbf{I}}}\right)^{\mathbf{2}}}} \tag{9}$$

Error metrics: MAE, MSE, and RMSE. where MAE is used to measure the average distance between the values of synthetic image I and the actual image $\check{\mathbf{I}}$ (1), **i** and **j** refers to the index on the image pixels, N and M refer to dimensions (x, y) of the image, where MSE is similar to MAE but different, the MSE equation calculates the square of average differences, RMSE the root main square for MSE, smaller distance identify smaller error in the results.

PSNR (4) is used for evaluating the quality of reconstructed or generated images; MAX refers to the maximum possible pixel value of the image (for example, 255 for 8-bit images($\mathbf{2^n - 1}$)). SSIM or SSI (5) is used for evaluating the quality of images by measuring the structural similarity between a reference image and a synthetic image, where I refers to the synthetic image and $\check{\mathbf{I}}$ refers to an actual image, average or mean (μ), standard division ($\boldsymbol{\sigma}$), $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$ where L = MAX value and $K_1 = 0.01$ and $K_2 = 0.02$.

By default, the SSIM index ranges from -1 to 1, where 1 indicates perfect similarity between the images, higher SSIM values suggest higher perceived image quality. The DSC (6) is commonly used to measure specific tissue between synthetic and actual images with ranges from 0 to 1. When the result is close to one, it is close to being the same as the actual image, where A represents the pixels in the predicted mask and B represents the pixels in the ground truth mask. ERGAS (7) was used to measure the quality between synthetic and actual image, where $\left(\mathbf{R_{synthetic\ CT}/R_{actual\ CT}}\right)$ define spatial resolution ratio between the synthesized and actual image. UQI (8) has the same function as the SSIM equation, but it aims to capture both local and global image characteristics, where the output value is between 1 and -1. SCC (9), also known as the spatial correlation index, measures the similarity in spatial patterns between images, and the results range between 1 and -1 [1], [21].

B. DISCUSSION AND COMPARISON

Table I, Table II, and Table III, present a detailed summary of CT and MRI modalities, the specific methodologies adopted, and the corresponding outcomes. The literature encompassed in these tables covers the period from 2017 to 2023, with a marked increase in brain-focused studies emerging from 2019 onwards, as depicted in Figure 3, underscoring the expanding research activity and notable advancements in this domain. The literature search strategy involved the use of keywords such as "GAN," "brain CT to MRI," "MRI to CT," and "synthetic medical images," applied across scholarly databases including Google Scholar and the Egyptian Knowledge Bank (EKB). The initial retrieval process identified over fifty relevant articles and publications. Following a rigorous screening process, only studies specifically addressing the brain and head region and published from 2017 onwards were retained; publications with anonymous authorship and those outside the specified anatomical focus or date range were excluded. This section offers an in-depth review and critique of selected works deemed particularly significant. The ensuing discussion emphasizes key findings, drawing on the comparative data presented in the preceding tables to encapsulate the core insights identified throughout the paper.

TABLE I OVERVIEW OF PAPERS ON CT SYNTHESIS FROM MRI (MRI→ CT) FOR BRAIN IMAGES, RESULTS FOCUSED ON MAE, PSNSR, AND SSIM, TABLE SORTED BY PUBLICATION DATE.

	Author & Date	Data Quantity	Data Type	DL model	Best Results	Highlights
1	Bowen Xin, [13] (2023)	270 patients	Paired	3D pix2pix GAN	MAE = 74.28	providing accurate synthetic CT images
2	Dr. Teodor Stanescu [21] (2023)	50 patients	Paired	Nine CycleGAN models	Model with GAN + L1 and ResNet_3 generator structure is the better with PSNR = 28.76 & SSIM = 0.90	CT synthesis to minimizing exposure in cancer radiation therapy
3	Lotte Nijskens [22] (2023)	95 patients	Unpaired	pix2pix (cGAN)	T1wGd has the best results of MAE = 69.6 ± 12.2 for RC & 71.0 ± 12.2 for LC	give a good identification for effect of use a different set of MRI images
4	Jiffy Joseph [23] (2022)	26 cancer patients	Paired	WGAN	MAE = 48.39, PSNR = 31.09, & SSIM = 0.9265 dB	Target to eliminates the side effects of radiation
5	Dalal Rajni Rajnish, [10] (2022)	-----	Unpaired	CycleGAN	-----	This paper focuses on explaining how synthetic image is done
6	Chun-Chieh Wang, [24] (2022)	50 patients	Paired	Pix2pix cGAN	PSNR = 29.47	providing a solution for generating synthetic CT images
7	Longfei Zhoua [25] (2021)	41 patients	Paired	U-Net, U-Net++, & Pix2Pix	U-Net++ model is better MAE = 0.082 PSNR = 67.9	U-Net++ Modified method from U-Net

	Author & Date	Data Quantity	Data Type	DL model	Best Results	Highlights
8	Faeze Gholamiankhah [26] (2021)	86 patients	Paired	ResNet, & GAN	ResNet MAE = 114.1±27.5 GAN MAE = 161.3±38.1	Target is to eliminates concerns about using CT
9	Heran Yang [12] (2020)	45 patients	Unpaired	sc-cycleGAN	MAE = 127.59, PSNR = 24.41, & SSIM = 0.773	The paper tried to generate new cycleGAN model and compare it with other Previous attempts
10	Matteo Maspero [27] (2020)	60 patients	Paired	cGAN	(MAE) = 61 ± 14	development more effective synthetic for pediatric radiation therapy planning
11	Bin Tang [28] (2020)	37 patients	Paired	U-Net	MAE = 60.52 ± 13.32 & PSNR = 49.23 ± 1.92	eliminating the need for CT scanning in the radiotherapy workflow
12	Xin Tie [19] (2020)	32 patients	Paired	MCMP-GAN	MAE = 75.7 ± 14.6, SSIM = 0.92 ± 0.02, & PSNR = 29.1 ± 1.6	generating accurate pseudo-CT images from multi-parametric MRI for nasopharyngeal carcinoma (NPC) patients
13	Mengke Qi [29] (2020)	45 patients	Unpaired	cGAN	Best results with four channel cGAN MAE = 69.98 ± 12.02, SSIM = 0.85 ± 0.03, & PSNR = 29.39 ± 1.29	Compare between U-Net and cGAN and give a good identification for effect of use a different set of MRI images
14	Bing Cao [18] (2020)	16 patients	Paired	Auto-GAN	SSIM = 0.9592	Auto-GAN is a multi-modality method
15	Haitao Wu1 [30] (2019)	222 images	Paired	UC-GAN	MAE = 76.7 ± 4.5 & PSNR = 46.1 ± 1.5	Paper compares UC-GAN with cycleGAN
16	Mariana Ferreira [31] (2019)	42 patients	Paired	Pix2pix (cGAN)	MAE_{Body} = 87 ± 11 of 2D sCT Generation	Work on 2D and 3D sCT
17	Bodo Kaiser [32] (2019)	60 patients	Paired	Pix2pix, U-Net, & context-aware	Pix2pix achieve best results: MAE = 136.9 & PSNR = 46.77	Target to reduce the need for CT scans
18	Yang Lei [33] (2019)	24 patients	Paired	CycleGAN	MAE = 55.7 & PSNR = 26.6	based on dense cycle-consistent generative adversarial networks
19	Hajar Emami [34] (2018)	15 patients	Paired	cGAN	Average MAE= 89.30±10.25, SSIM = 0.83±0.03, & PSNR= 26.64±1.17	Target to generating synthetic CT images
20	Xiao Han [35] (2017)	18 patients	Paired	DCNN	MAE = 84.8 HU	DCNN model near U-net architecture
21	Jelmer M. Wolterink [36] (2017)	24 patients	Paired & Unpaired	CycleGAN	Unpaired: MAE = 73.7 ± 2.3 & PSNR = 32.3 ± 0.7, Paired: MAE = 89.4 ± 6.8 & PSNR = 30.6 ±0.9	Target to synthesizing quality CT without using Unpaired data

TABLE II OVERVIEW OF PAPERS ON MRI SYNTHESIS FROM CT (CT→ MRI) FOR BRAIN IMAGES, RESULTS FOCUSED ON MAE, PSNSR, AND SSIM, TABLE SORTED BY PUBLICATION DATE.

	Author & Date	Data Quantity	Data Type	DL model	Best Results	Highlights
1	Jake McNaughton [15] (2023)	181 patients	Paired	Eight models: UNet V1, V2, Patch, 2D, ++, Attention, Transformer, & CycleGAN	U-Net better results with MAE = 18.29 ± 6.61, PSNR = 21.571 ± 2.724, & SSIM = 0.882 ± 0.030	The paper test U-Net models and compares them with the original structures.
2	Na Hu1 [16] (2022)	26 patients	Paired	pix2pix	PSNR = 24.30, & SSIM = 0.857	Diagnosing strokes
3	Wen Li [14] (2020)	------	------	U-Net with loses functions models, Pix2Pix, & CycleGAN	U-Net with L1+L2 loss achieved the best results: MAE = 74.19, PSNR = 32.44, and SSIM = 0.9440	Comparison between methods to find best model
4	Cheng-Bin Jin [3] (2019)	202 patients	Paired & unpaired	MRI-GAN	Paired: MAE =20.34, PSNR =64.28, & SSIM =0.24 Unpaired: MAE =22.94, PSNR =63.77, & SSIM =0.22	MRI-GAN similar structure to cycleGAN
5	Jonathan Rubin [37] (2019)	94 patients	paired	cGAN	-----	Results measured by FCN values.

TABLE III OVERVIEW OF BOTH MODALITY (CT↔ MRI) PAPERS OF DATA, MODEL AND RESULTS, RESULTS MEASURED BY MAE, PSNSR, AND SSIM EQUATIONS, TABLE SORTED BY PUBLICATION DATE.

	References & Date	Data Quantity	Data Type	DL model	Best Results	Highlights
1	Jiayuan Wang [11] (2023)	367 images of CT & MRI	Paired	ADC-cycleGAN	MRI to CT synthesis: MAE = 0.11005, PSNR = 19.04385, SSIM = 0.68551 CT to MRI synthesis: MAE = 0.11080, PSNR = 20.12068, SSIM = 0.65568	Target to high-quality image synthesis
2	Jiayuan Wang [17] (2022)	-----	Unpaired	DC-cycleGAN	MAE = 0.04559, PSNR = 26.68858, & SSIM = 0.82622	contain code resources and comparison with other models
3	Alaa Abu-Srhana [38] (2021)	367 paired & 840 unpaired images	Paired & unpaired	uagGAN	Results of paired–unpaired uagGAN models: MRI to CT synthesis: PSNR = 34.786, SSIM = 0.739 CT to MRI synthesis: PSNR = 31.821, SSIM = 0.603	Overcoming the challenges that faces the paired and unpaired methods by creating a new paired-unpaired method
4	Denis Prokopenko [39] (2019)	78 patients	Unpaired	DualGAN	CT to MRI synthesis: MAE = 60.83, PSNR = 17.21, & SSIM = 0.8 CT to MRI synthesis: MAE = 37.99, PSNR = 23.31, & SSIM = 0.78	Target to get best result from DualGAN

a. MRI TO CT

From Table I, we can note that paper No. 4 achieved the best measurable values on all the following: SSIM, PSNR, and MAE. In their study, they used a dataset from the MVR Cancer Center in the DICOM format. Challenges still need to be solved. Paper hinted at it: masking the CT images to filter the noises still has some imperfections due to the CT and MRI differences in data nature. Papers number (No.) 8, 9, and 17 achieved poor numerical results in terms of MAE compared to the rest of the papers. Paper 8 explained that the reason for these limitations was that they worked on one type of data and that the performance of their project could be measured by comparing it with methods such as U-Net. For paper No. 9, their focus was on creating a research paper that would overcome CycleGAN and provide visual inspections. That was mentioned in detail in the “SOFTWARE BACKGROUND” paragraph. They succeeded in their development compared to the ordinary CycleGAN model, but they were not able to achieve high MAE values because of the nature and quality of the training data used.

Paper No. 17 achieved reliable results compared with the rest of the models, and they explained their limitations because of the lack of data used. Not all measurements are decisive for deciding about the quality of the model, and visual inspection is a major part of the decision. Why? Because the model's ability to notice small-sized problems, such as tumors and clots, elevates it to the point of reliability, Paper No. 21 represents the effect of using paired and unpaired data mathematically with the CycleGAN model, where the unpaired data achieves better results with the mode due to its nature of dealing with unpaired data. Paper No. 13 helped us to obtain a good idea about the effects of MRI image sequences (T1, T2, T1C, and T1DixonC-water) by applying the cGAN with a 4-channels model to obtain results, which showed that T1 is the best, followed by T1C, T2, and T1DixonC-water they achieve disparity results.

b. CT TO MRI

Noted that from Table II and Table III for synthesis MRI images, synthetic MRI does not

achieve high-quality image measurements (PSNR, SSIM). The only paper that achieves high image quality was Paper No. 1 and 3 from Table II, where this paper compares many models to show that the best model, model U-Net (ordinary U-Net), achieved good results over other modifications. U-Net models and ordinary CycleGAN, and the models U-Net_L1+L2 and paired-cycleGAN from paper No. 3 achieved higher results and good visual inspection.

From the previous comparisons, it can be concluded that the U-Net models and their improvements achieved the best results in terms of quality and error rate compared to other models. The main reason for the few papers dealing with the synthesis of MRI images from CT is that MRI images contain many details that CT images may fail to capture, which subsequently leads to deficiencies in the synthesis of images. This sensitivity, especially in critical medical decisions like detecting tumors and strokes in brain regions, necessitates heightened efforts and benefits from early diagnostic capabilities.

Across all tables, it is noted that the amount of patient data used increases over time, publishing data for researchers, especially after the coronavirus disease (COVID-19), and we cannot ignore the effect of data on the results, as in some papers, the models achieved results, and the same models achieved lower or higher results in other papers related to the dataset used in building the generative model.

C. SUMMARY

The key findings from the reviewed papers and articles can be summarized as follows: The results of a model do not solely reflect the performance of the method itself but are influenced by both the model and the dataset on which it was trained. To properly assess a method's performance, it must be compared with other approaches using the same data input. Visual inspection remains crucial for evaluating the output images generated by the model. Additionally, using unpaired data with the CycleGAN model yields superior results compared to paired data. Expanding the dataset size has a positive impact on the model's predictive capability, with unpaired data again demonstrating better outcomes. In terms of MRI to CT translation, MRI T1 data has shown significant advantages and greater applicability. Lastly, models like U-Net and its variations, along with GAN-based upgrades such as the Pix2Pix generator, outperform other techniques in terms of image quality and error rates, especially in CT to MRI translation.

IV. CONCLUSION

In conclusion, significant progress has been made in the field of synthetic medical imaging, several challenges remain, particularly in the spatial synthesis of brain images. Notably, differences in the nature and complexity of CT and MRI data create limitations for MRI synthesis from CT images. Additionally, the limited availability of high-quality, paired datasets further hinders progress in this area. Addressing these issues is critical for achieving more reliable and accurate results in synthetic brain imaging projects. This review provides a comprehensive analysis of advancements in generating synthetic MRI-CT images for brain applications using generative models. It highlights the potential benefits of these techniques, such as cost-effectiveness and enhanced imaging quality. The paper underscores the importance of selecting optimal datasets and choosing the most appropriate deep learning (DL) models to ensure reliable outcomes. Challenges, including the scarcity of research focused on CT to MRI translation and the inherent difficulty in converting to MRI, were also discussed. Furthermore, the review touches on the computational aspects and types of models employed in this field. Overall, this paper offers valuable insights into the current state of synthetic MRI-CT brain imaging and provides a foundation for future research and development in this promising area.

REFERENCES

[1] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Med. Image Anal.*, vol. 58, Sep. 2019, doi: 10.1016/j.media.2019.101552.

[2] F. Shamshad *et al.*, "Transformers in medical imaging: A survey," Jan. 2023. doi: 10.1016/j.media.2023.102802.

[3] C. Bin Jin *et al.*, "Deep CT to MR synthesis using paired and unpaired data," *Sensors (Switzerland)*, vol. 19, no. 10, pp. 1–19, 2019, doi: 10.3390/s19102361.

[4] I. K. Acquah, S. Issahaku, and S. N. A. Tagoe, "A systematic review of deep learning techniques for generating synthetic CT images from MRI data," *Polish J. Med. Phys. Eng.*, vol. 31, no. 1, pp. 20–38, 2025, doi: 10.2478/pjmpe-2025-0003.

[5] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," Nov. 2016.

[6] S. O. Mehdi Mirza, "Conditional Generative Adversarial Nets," pp. 1–7, 2014.

[7] A. Thakur and M. Satish, "Generative Adversarial Networks."

[8] J. Zhu, T. Park, A. A. Efros, B. Ai, and U. C. Berkeley, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks".

[9] Mohammed Alhami, "Generative Adversarial Networks GANs: A Beginner's Guide," https://towardsdatascience.com/generative-adversarial-networks-gans-a-beginners-guide-f37c9f3b7817.

[10] T. Suryawanshi, "Medical image synthesis using GAN," no. July, 2022.

[11] J. Wang, Q. M. J. Wu, and F. Pourpanah, "An Attentive-based Generative Model for Medical Image Synthesis," 2023, doi: 10.1007/s13042-023-01871-0.

[12] H. Yang *et al.*, "Unsupervised MR-to-CT Synthesis Using Structure-Constrained CycleGAN," *IEEE Trans. Med. Imaging*, vol. 39, no. 12, pp. 4249–4261, 2020, doi: 10.1109/TMI.2020.3015379.

[13] B. Xin, A. Nicolson, H. Chourak, and G. Belous, "Team KoalAI : Locally-enhanced 3D Pix2Pix GAN for Synthetic CT Generation," pp. 1–6, 2023.

[14] W. Li *et al.*, "Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy," *Quant. Imaging Med. Surg.*, vol. 10, no. 6, pp. 1223–1236, 2020, doi: 10.21037/QIMS-19-885.

[15] J. McNaughton, S. Holdsworth, B. Chong, J. Fernandez, V. Shim, and A. Wang, "Synthetic MRI Generation from CT Scans for Stroke Patients," *BioMedInformatics*, vol. 3, no. 3, pp. 791–816, 2023, doi: 10.3390/biomedinformatics3030050.

[16] N. Hu *et al.*, "Detecting brain lesions in suspected acute ischemic stroke with CT-based synthetic MRI using generative adversarial networks," *Ann. Transl. Med.*, vol. 10, no. 2, pp. 35–35, 2022, doi: 10.21037/atm-21-4056.

[17] J. Wang, Q. M. J. Wu, and F. Pourpanah, "DC-cycleGAN: Bidirectional CT-to-MR synthesis from unpaired data," *Comput. Med. Imaging Graph.*, vol. 108, no. November 2022, 2023, doi: 10.1016/j.compmedimag.2023.102249.

[18] B. Cao, H. Zhang, N. Wang, X. Gao, and D. Shen, "Auto-GAN: Self-supervised collaborative learning for medical image synthesis," *AAAI 2020 - 34th AAAI Conf. Artif. Intell.*, pp. 10486–10493, 2020, doi: 10.1609/aaai.v34i07.6619.

[19] 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, doi: 10.1002/mp.14062.

[20] X. Z. Can Chang, Li Yao, "A Weakly GAN Combining Global Attention Network in Incomplete

Multimodal MRI Synthesis," *SSRN*, p. 14, 2023, doi: 10.2139/ssrn.4463368.

[21] T. Kutlu, "The Investigation of Generative Adversarial Networks for the Generation of Synthetic Computer Tomography Data from Magnetic Resonance Imaging Images for Applications in Adaptive Radiation Therapy for the Central Nervous System," vol. 4, pp. 88–100, 2023.

[22] L. Nijskens, C. A. T. van den Berg, J. J. C. Verhoeff, and M. Maspero, "Exploring contrast generalisation in deep learning-based brain MRI-to-CT synthesis," *Phys. Medica*, vol. 112, no. March, 2023, doi: 10.1016/j.ejmp.2023.102642.

[23] J. Joseph, C. Hemanth, P. Pulinthanathu Narayanan, J. Pottekkattuvalappil Balakrishnan, and N. Puzhakkal, "Computed tomography image generation from magnetic resonance imaging using Wasserstein metric for MR-only radiation therapy," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 6, pp. 2080–2093, 2022, doi: 10.1002/ima.22777.

[24] C. C. Wang *et al.*, "Magnetic Resonance-Based Synthetic Computed Tomography Using Generative Adversarial Networks for Intracranial Tumor Radiotherapy Treatment Planning," *J. Pers. Med.*, vol. 12, no. 3, pp. 1–17, 2022, doi: 10.3390/jpm12030361.

[25] L. Zhou *et al.*, "Synthesizing Human Brain Computed Tomography Images from Magnetic Resonance Images Based on Machine Learning," *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.3989550.

[26] F. Gholamiankhah, S. Mostafapour, and H. Arabi, "Deep learning-based synthetic CT generation from MR images: comparison of generative adversarial and residual neural networks," *Int. J. Radiat. Res.*, vol. 20, no. 1, pp. 121–130, 2022, doi: 10.52547/ijrr.20.1.19.

[27] M. Maspero *et al.*, "Deep learning-based synthetic CT generation for paediatric brain MR-only photon and proton radiotherapy," *Radiother. Oncol.*, vol. 153, pp. 197–204, 2020, doi: 10.1016/j.radonc.2020.09.029.

[28] B. Tang *et al.*, "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, 2021, doi: 10.1002/acm2.13176.

[29] M. Qi *et al.*, "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, 2020, doi: 10.1002/mp.14075.

[30] H. Wu, X. Jiang, and F. Jia, *UC-GAN for MR to CT image synthesis*, vol. 11850 LNCS, no. 1. Springer International Publishing, 2019. doi: 10.1007/978-3-030-32486-5_18.

[31] M. Ferreira, "A generative adversarial network approach to synthetic-CT creation for MRI-based radiation therapy," *Dissertation*, 2019.

[32] B. Kaiser and S. Albarqouni, "MRI to CT Translation with GANs," 2019.

[33] Y. Lei *et al.*, "MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks," *Med. Phys.*, vol. 46, no. 8, pp. 3565–3581, Aug. 2019, doi: 10.1002/mp.13617.

[34] 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, 2018, doi: 10.1002/mp.13047.

[35] X. Han, "MR-based synthetic CT generation using a deep convolutional neural network method," *Med. Phys.*, vol. 44, no. 4, pp. 1408–1419, Apr. 2017, doi: 10.1002/mp.12155.

[36] J. Wolterink, A. M. Dinkla, M. Savenije, and P. R. Seevinck, "Deep MR to CT Synthesis using Unpaired Data," no. November, 2017.

[37] J. Rubin and S. M. Abulnaga, "CT-To-MR conditional generative adversarial networks for ischemic stroke lesion segmentation," 2019. doi: 10.1109/ICHI.2019.8904574.

[38] 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, no. August, 2021, doi: 10.1016/j.compbiomed.2021.104763.

[39] D. Prokopenko, J. V. Stadelmann, H. Schulz, S. Renisch, and D. V Dylov, "Unpaired synthetic image generation in radiology using GANs," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, pp. 94–101. doi: 10.1007/978-3-030-32486-5_12.