



# Detecting brain lesions in suspected acute ischemic stroke with CT-based synthetic MRI using generative adversarial networks

Na Hu<sup>1,2#^</sup>, Tianwei Zhang<sup>3#</sup>, Yifan Wu<sup>4</sup>, Biqu Tang<sup>2</sup>, Minlong Li<sup>2,5</sup>, Bin Song<sup>1</sup>, Qiyong Gong<sup>2</sup>, Min Wu<sup>2</sup>, Shi Gu<sup>3</sup>, Su Lui<sup>2^</sup>

<sup>1</sup>Department of Radiology, West China Hospital of Sichuan University, Chengdu, China; <sup>2</sup>Huaxi MR Research Center (HMRRRC), Functional and Molecular Imaging Key Laboratory of Sichuan Province, Department of Radiology, West China Hospital of Sichuan University, Chengdu, China;

<sup>3</sup>Department of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China; <sup>4</sup>Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA; <sup>5</sup>Department of Radiology, Zigong Fourth People's Hospital, Zigong, China

**Contributions:** (I) Conception and design: N Hu, T Zhang, B Song, Q Gong, M Wu, S Gu, S Lui; (II) Administrative support: N Hu, T Zhang, B Tang, M Li, B Song, Q Gong, M Wu, S Gu, S Lui; (III) Provision of study materials or patients: N Hu, T Zhang, B Tang, M Li, B Song, Q Gong, M Wu, S Gu, S Lui; (IV) Collection and assembly of data: N Hu, T Zhang, Y Wu, B Tang, M Li; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Min Wu. Huaxi MR Research Center (HMRRRC), Functional and Molecular Imaging Key Laboratory of Sichuan Province, Department of Radiology, West China Hospital of Sichuan University, No. 37, Guoxue Alley, Chengdu 610041, China. Email: wuminscu@scu.edu.cn; Shi Gu. Department of Computer Science and Engineering, University of Electronic Science and Technology of China, No. 2006, Xiyuan Avenue, Chengdu 610041, China. Email: gus@uestc.edu.cn; Su Lui. Huaxi MR Research Center (HMRRRC), Functional and Molecular Imaging Key Laboratory of Sichuan Province, Department of Radiology, West China Hospital of Sichuan University, No. 37, Guoxue Alley, Chengdu 610041, China. Email: lusuwcums@tom.com.

**Background:** Difficulties in detecting brain lesions in acute ischemic stroke (AIS) have convinced researchers to use computed tomography (CT) to scan for and magnetic resonance imaging (MRI) to search for these lesions. This work aimed to develop a generative adversarial network (GAN) model for CT-to-MR image synthesis and evaluate reader performance with synthetic MRI (syn-MRI) in detecting brain lesions in suspected patients.

**Methods:** Patients with primarily suspected AIS were randomly assigned to the training (n=140) or testing (n=53) set. Emergency CT and follow-up MR images in the training set were used to develop a GAN model to generate syn-MR images from the CT data in the testing set. The standard reference was the manual segmentations of follow-up MR images. Image similarity was evaluated between syn-MRI and the ground truth using a 4-grade visual rating scale, the peak signal-to-noise ratio (PSNR), and the structural similarity index measure (SSIM). Reader performance with syn-MRI and CT was evaluated and compared on a per-patient (patient detection) and per-lesion (lesion detection) basis. Paired *t*-tests or Wilcoxon signed-rank tests were used to compare reader performance in lesion detection between the syn-MRI and CT data.

**Results:** Grade 2–4 brain lesions were observed on syn-MRI in 92.5% (49/53) of the patients, while the remaining syn-MRI data showed no lesions compared to the ground truth. The GAN model exhibited a weak PSNR of 24.30 dB but a favorable SSIM of 0.857. Compared with CT, syn-MRI led to an increase in the overall sensitivity from 38% (57/150) to 82% (123/150) in patient detection and from 4% (68/1,620) to 16% (262/1,620) in lesion detection ( $R=0.32$ , corrected  $P<0.001$ ), but the specificity in patient detection decreased from 67% (6/9) to 33% (3/9). An additional 75% (70/93) of patients and 15% (77/517) of lesions missed on CT were detected on syn-MRI.

**Conclusions:** The GAN model holds potential for generating synthetic MR images from noncontrast CT data and thus could help sensitively detect individuals among patients with suspected AIS. However, the image similarity performance of the model needs to be improved, and further expert discrimination is strongly recommended.

<sup>^</sup> ORCID: Na Hu, 0000-0001-6362-0969; Su Lui, 0000-0003-3541-1769.

**Keywords:** Acute ischemic stroke (AIS); generative adversarial network (GAN); imaging diagnosis; CT-to-MR image synthesis

Submitted Aug 04, 2021. Accepted for publication Nov 26, 2021.

doi: 10.21037/atm-21-4056

View this article at: <https://dx.doi.org/10.21037/atm-21-4056>

## Introduction

Imaging detection of acute ischemic stroke (AIS) requires immediacy, availability, and sensitivity given the narrow therapeutic window and devastating consequences to the brain (1). Hypodensities on brain computed tomography (CT) scans are often the first to be recognized as suspected stroke lesions but require further identification (2). Comparatively, stroke lesions are more apparent on magnetic resonance imaging (MRI) sequences, among which T2-weighted fluid-attenuated inversion recovery (FLAIR) reflects the underlying pathologies earlier than CT and for a greater duration than diffusion-weighted sequences (3). However, MRI is time-consuming and lacks the availability of CT, and neither modality meets all the above-mentioned requirements. In addition, trade-offs are still involved after a scan, e.g., between scan completion and report turnaround (4) or in medically underserved areas lacking advanced imaging support (5). Dilemmas in modality selection have long persisted in emergency practice.

Consequently, the idea of integrating the strengths of CT and MRI into a single examination was conceived. Specifically, CT could be adopted for brain scans to meet the immediacy and availability requirements, and the properties of MRI could be exploited for lesion searching to reach the sensitivity demands. This assumption is especially meaningful for developing countries such as China, where imaging resources are unevenly distributed. In crowded urban hospitals, timely detection of potential patients can help reduce the turnaround time of the AIS workflow (6). In remote rural hospitals, images presenting with apparent MR features could serve as an alternative to advanced imaging when MRI is not available or the radiologist lacks sufficient experience (7).

As deep learning techniques continue to advance in medical image analysis, CT-to-MR image synthesis has been proposed as a potential solution for detecting lesions, to which generative adversarial networks (GANs) offer an approach (8). GANs are deep generative models of realistic images, typically consisting of two deep neural networks. One, the generator, synthesizes realistic images, and the

other, the discriminator, classifies images as either real or fake (synthetic). Both networks are trained simultaneously but conversely until the discriminator is “fooled” half the time, indicating that the generator has produced the most plausible images possible (9). Performing well in many image synthesis tasks, GANs are delivering on the promise of generating realistic but fake images (10). Recent studies have employed GANs to map brain images from one imaging modality to another, e.g., between different MR sequences (11,12), between MR scanners of different field strengths (13), and between MRI and CT (8,9,14). These preliminary studies have indicated that GAN models can generate multiple contrast-weighted images in a single brain scan, suggesting the possibility of CT-to-MR image synthesis in stroke imaging. However, it is uncertain whether human readers could use noncontrast CT-based outputs to search for suspected brain lesions.

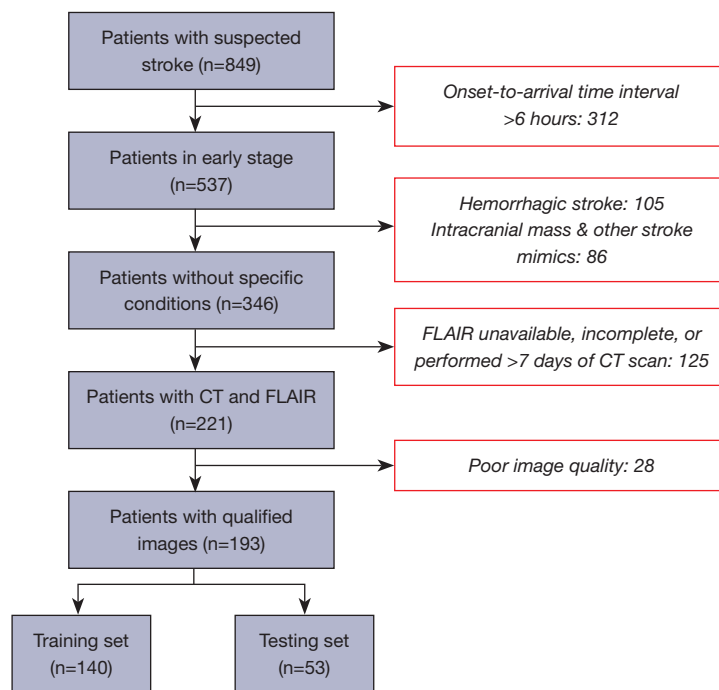
The purpose of this study was twofold: to develop a GAN model for CT-to-MR image synthesis for patients with suspected AIS and to evaluate the feasibility of detecting suspected brain lesions using synthetic MRI (syn-MRI). We hypothesized that detection sensitivity would be improved when using syn-MRI over CT.

We present the following article in accordance with the Standards for Reporting Diagnostic Accuracy (STARD) reporting checklist (available at <https://atm.amegroups.com/article/view/10.21037/atm-21-4056/rc>).

## Methods

### *Study participants*

We retrospectively collected emergency head noncontrast CT and follow-up FLAIR images from patients primarily with suspected AIS, who were collected consecutively from September 2015 to November 2018 in our hospital. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and approved by the Ethics Committee of Biomedical Research, West China Hospital of Sichuan University (No. 2020-394). Individual consent for this retrospective analysis was waived.



**Figure 1** Flowchart of patient inclusion and exclusion. CT, computed tomography; FLAIR, fluid-attenuated inversion recovery.

Patients were initially selected if they had a sudden onset of at least one of the following symptoms: numbness or weakness of the face or any limb, loss of vision, speech difficulties, dizziness, loss of balance or coordination, confusion, and severe headache. Typically, in-hospital patients undergo follow-up MRI within 1–4 weeks of symptom onset when the clinical status is stable. All MRI protocols routinely involve FLAIR but not necessarily diffusion-weighted sequences. Timing and protocols of the follow-up MRI were decided by the doctors in charge. The exclusion criteria were a time interval from onset to hospital arrival beyond 6 hours; hemorrhagic stroke, intracranial mass, or other clinically identified stroke mimics (e.g., hypoglycemia and metabolic encephalopathy); FLAIR imaging unavailable, incomplete, or performed beyond 7 days of the CT scan; and poor image quality (evident artifacts and structural distortions). *Figure 1* shows the flowchart of patient inclusion and exclusion. The sample size was determined by referring to similar studies (mean, 156; range, 140–170) (15–17). Finally, 193 patients were randomly assigned to either the training (n=140) or testing (n=53) set.

### Image acquisition

Baseline noncontrast head CT scans were performed

using a 128-slice dual-source scanner with the following parameters: 70 kVp, 150 mAs, and 5-mm slice thickness. At follow-up, axial T<sub>2</sub>-weighted FLAIR images were obtained using a 3T scanner with a 16-channel head coil and the following parameters: repetition time (TR) =11,000 ms, echo time (TE) =120 ms, inversion time (TI) =2,800 ms, field of view (FOV) =230 mm × 185 mm, matrix =240×140, and slice thickness =6 mm. Both scanners were dedicated for emergency imaging.

### Training of the GAN model

#### Image preprocessing

Raw data in Digital Imaging and Communications in Medicine (DICOM) format collected from the hospital were first converted to the Neuroimaging Informatics Technology Initiative (NifTI) format using Python 3.5.0 (RRID: SCR\_008394; <https://www.python.org>) and SimpleITK 1.2.0 (<https://simpleitk.org>). Format conversion preserved the metadata intrinsically related to the images but removed those associated with patient privacy and other redundancies. Next, to generate paired image sets for model training, spatial normalization was conducted using the SPM8 normalization algorithm (RRID: SCR\_007037; <https://www.fil.ion.ucl.ac.uk/spm>)

and the Clinical Toolbox for SPM8 (RRID: SCR\_014096; <https://www.nitrc.org/projects/clinicaltbx>). The CT and FLAIR images were coregistered to the standard Montreal Neurological Institute (MNI) space. These paired images were well aligned spatially and temporally and composed of 26 pairs of continuous slices (voxel size = 0.45 mm × 0.45 mm × 7 mm, matrix = 408×481) for each patient. A stroke window was set to standardize and optimize the image contrast on CT, with a window length of 8 HU and a window level of 32 HU. Third, the skulls of the paired images were stripped using a mask built on the CT volume with a threshold of 135 HU. Fourth, the upper and lower 7 image slices were removed to reduce noise, followed by in-plane resampling with a matrix size of 256×256 and a slice number of 12. The voxel intensities were finally normalized to zero mean and unit standard deviation using Z-score standardization.

### GAN Architecture

A 3D-CT2MR GAN model was built on the pix2pix GAN framework (18) with the following improved formulas.

$$L_{GAN}(G, D) = -E_{x \sim CT} [D(G(x))] + E_{y \sim MRI} [D(y)] - \lambda_{gp} E_{\hat{x}} \left[ \left( \|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1 \right)^2 \right] \quad [1]$$

$$L_{L1}(G) = E_{x \sim CT, y \sim MRI} [\|y - G(x)\|_1] \quad [2]$$

$$L = \arg \min_G \max_D L_{GAN}(G, D) + \lambda_{L1} L_{L1}(G) \quad [3]$$

The generator network,  $G(\cdot)$ , takes the real CT volumes as input and generates realistic MRI volumes. Conversely, the discriminator network,  $D(\cdot)$ , takes the real  $[D(y)]$  or realistic  $[D(G(x))]$  MRI as input and determines whether the input is real or fake. To accelerate the training process stably, a gradient penalty term is added to the objective function in Eq. [1] according to the Wasserstein GAN. In Eq. [1],  $\hat{x}$  is sampled uniformly along a straight line between the paired real and generated images, and  $\lambda_{gp}$  is the weight of the gradient penalty term. We also combined the GAN loss with an additional L1 loss in Eq. [2] to force the generator output to match the real images at the pixel level.  $\lambda_{L1}$  is used to control the weight of the L1 loss, and the final adversarial loss is given in Eq. [3].

The generator follows the encoder-decoder architecture, which consists of two convolutional layers for downsampling, six residual blocks, and two transposed convolutional layers for upsampling. On the top of the

generator network, we added a tanh function as the output layer to generate the synthetic images. To stabilize the training and regularize the model, we used instance normalization in all layers except the output layer. Following the pix2pix GAN, we leveraged the PatchGANs for the discriminator network and only penalized the architecture at the patch scale. The discriminator consists of five convolution blocks, each of which is composed of a 3D convolution layer and a Leaky ReLU activation function with a negative slope of 0.01. Another 3D convolution layer was introduced as the output layer instead of the fully connected layer. No normalization layers were used in the discriminator network. The network structure of the generator and the discriminator is shown in *Table 1*.

### Implementation

The GAN model was trained using a graphics processing unit (GPU) (NVIDIA Titan Xp, 12 GB memory, USA). Five discriminator updates followed by one generator update were performed during the model training. The batch size was set to 1, given the available GPU memory. The Adam optimizer was used for training both the generator and discriminator with momentum  $\beta_1=0.5$  and  $\beta_2=0.999$ . We trained the model for a total of 200,000 epochs, which took 33 hours in all. The learning rate was fixed to  $1e^{-4}$  for the first 100,000 epochs and then halved every 10,000 epochs thereafter. We set  $\lambda_{gp}=10$  in Eq. [1] and  $\lambda_{L1}=100$  in Eq. [2]. It should be mentioned that no extra data augmentation was conducted in the training process.

In summary, we developed a 3D-CT2MR GAN framework to learn to synthesize MR images from CT data based on the pix2pix image-to-image translation network (18). The raw data were preprocessed to generate paired CT and MR images, which were fed into the GAN model. The generator translated the images from baseline CT to syn-MRI, while the discriminator distinguished between the syn-MRI and follow-up FLAIR images. Training loss was improved to maintain both the global anatomywise and the local lesionwise consistency during image translation. Once Nash equilibrium was achieved, model training was accomplished (*Figure 2*). The trained model required 52.24 seconds to perform CT-to-MR translation for each image series, averaging 52.23 seconds for preprocessing and the remaining 0.01 seconds for synthesis. The source code is publicly reposted at GitHub ([https://github.com/ZtwZx/CT2MRI\\_3DGAN](https://github.com/ZtwZx/CT2MRI_3DGAN)).

**Table 1** Network structure of the generator and the discriminator

Network components	Layer information	Output size
<b>Generator</b>		
Input		1×256×256×12
Input layer	Conv3D (N64, K7×7×7, S1, P3) + IN + ReLU	64×256×256×12
Downsampling	Conv3D (N128, K4×4×4, S2, P1) + IN + ReLU	128×128×128×6
	Conv3D (N256, K4×4×4, S2, P1) + IN + ReLU	256×64×64×3
Bottleneck	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
	Residual block: Conv (N256, K3×3×3, S1, P1) + IN + ReLU	256×64×64×3
Upsampling	DeConv (N128, K4×4×4, S2, P1) + IN + ReLU	128×128×128×6
	DeConv (N64, K4×4×4, S2, P1) + IN + ReLU	64×256×256×12
Output layer	DeConv (N1, K7×7×7, S1, P3) + IN + ReLU	1×256×256×12
<b>Discriminator</b>		
Input		1×256×256×12
Hidden layer	Conv3D (N64, K4×4×4, S2, P1) + Leaky ReLU	64×128×128×6
Hidden layer	Conv3D (N128, K4×4×4, S2, P1) + Leaky ReLU	128×64×64×3
Hidden layer	Conv3D (N256, K4×4×3, S2×2×1, P1) + Leaky ReLU	256×32×32×3
Hidden layer	Conv3D (N512, K4×4×3, S2×2×1, P1) + Leaky ReLU	512×16×16×3
Hidden layer	Conv3D (N1,024, K4×4×3, S2×2×1, P1) + Leaky ReLU	1,024×8×8×3
Output layer	Conv3D (N1, K3×3×3, S1, P1)	1×8×8×3

N, number of output channels; K, kernel size; S, stride size; P, padding size; IN, instance normalization.

### Quality evaluation of the syn-MR images

In the testing set, CT images were translated to syn-MR images via the model. The quality of the syn-MR images was evaluated visually by a neuroradiologist with 9 years of experience, examining whether the synthetic images introduced new structural distortions or hallucinogenic areas that were not visible on the ground-truth FLAIR images. Neither significant structural changes nor hallucinogenic signal intensities were ultimately found.

### Brain lesion reference standard

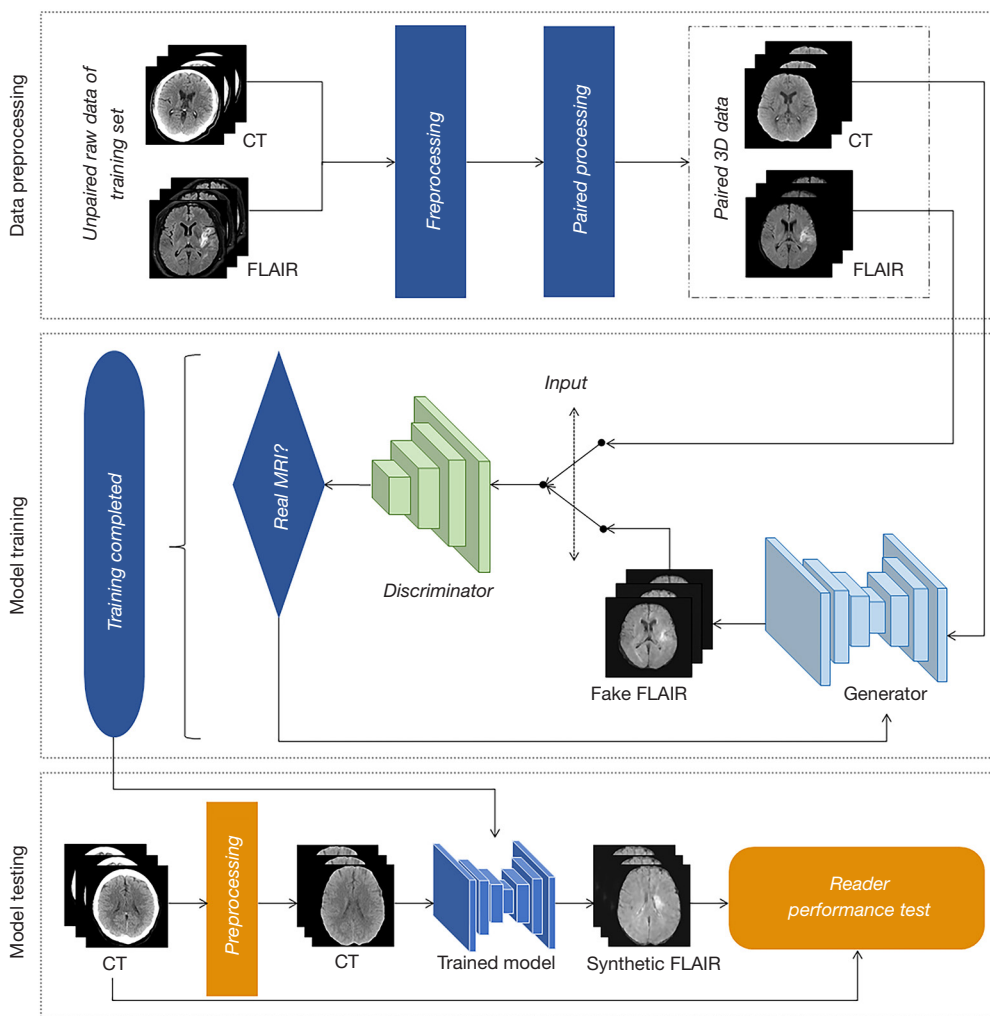
In the testing set, hyperintense brain lesions on FLAIR images were manually segmented by the neuroradiologist using ITK-SNAP 3.8.0-beta (RRID: SCR\_002010; [https://](https://www.itksnap.org)

[www.itksnap.org](https://www.itksnap.org)). Another neuroradiologist with 12 years of experience examined the segmented images and corrected the tracings, if necessary. The manual segmentation was set as the reference standard for its irreplaceability in imaging diagnosis. We excluded hyperintense nodes smaller than 5 mm, symmetric periventricular hyperintensities, and chronic infarcts. On a per-patient basis, patients with any segmented lesions were labeled “positive”; otherwise, they were labeled “negative”. Both neuroradiologists were blinded to the clinical information of the patients.

### Evaluation of image similarity between syn-MRI and ground-truth FLAIR

The image similarity of hyperintense brain lesions was





**Figure 2** Training and testing of the generative adversarial network model. CT, computed tomography; FLAIR, fluid-attenuated inversion recovery; 3D, three-dimensional; MRI, magnetic resonance imaging.

evaluated between the syn-MR and ground-truth FLAIR images in the testing set. First, the neuroradiologist qualitatively investigated the overall image similarity for each patient. A visual rating scale was applied, accounting for the location and extent of these lesions on syn-MRI relative to the FLAIR images: Grade 1, negative on syn-MRI but positive on FLAIR imaging; Grade 2, partially overlapping (<50%); Grade 3, largely overlapping (50–70%); and Grade 4, almost to completely overlapping (≥70%) or negative on both syn-MRI and FLAIR imaging.

Second, two quantitative image quality metrics were calculated, namely, the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) (19,20). The PSNR measures the ratio between the maximum

possible signal power and the power of the distorting noise that affects the image quality. The SSIM is used to estimate the perceived image quality by calculating the similarity between the original and generated images. It is designed based on a large number of image factors, i.e., the luminance, contrast, and structure, to better suit the workings of the human visual system than the PSNR.

The formulas for the two metrics are as follows:

$$PSNR = 10 \log_{10} \frac{\max^2(y(x), G(x))}{\frac{1}{N} \sum \|y(x) - G(x)\|_2^2} \quad [4]$$

$N$  is the total voxel number of each image, and  $\max(y(x), G(x))$  is the maximum intensity value between the ground-truth image  $y(x)$  and the generated image  $G(x)$ . The

PSNR value for favorable image quality is often above 30 dB.

$$SSIM = \frac{(2\mu_{y(x)}\mu_{G(x)} + c_1)(2\sigma_{y(x)G(x)} + c_2)}{(\mu_{y(x)}^2 + \mu_{G(x)}^2 + c_1)(\sigma_{y(x)}^2 + \sigma_{G(x)}^2 + c_2)} \quad [5]$$

Here,  $\mu_{y(x)}$  and  $\mu_{G(x)}$  are the means of images  $y(x)$  and  $G(x)$ , respectively,  $\sigma_{y(x)}$ ,  $\sigma_{G(x)}$  and  $c_1$  are their variances. The positive constants  $c_1$  and  $c_2$  are included to avoid a null denominator. The SSIM values range from  $-1$  to  $1$ , with the ideal value close to  $1$ .

### Reader performance test with the CT and syn-MR images

In the testing set, reader performance was assessed with the CT and syn-MR images. The readers were asked to detect lesions including parenchymal hypodensities on CT and hyperintensities on syn-MRI, with exclusion criteria similar to those described in the *Brain Lesion Reference Standard* subsection. Three readers with beginning to intermediate levels of expertise participated in this blinded test, including a CT technician with 1.5 years of experience, a second-year radiology resident, and a junior radiologist with 4 years of experience. They were trained together on lesion detection and segmentation using ITK-SNAP and then tested under the same ambient conditions.

Either CT or syn-MR image series were randomly displayed via in-house software. First, on a per-patient basis (patient detection), the readers aimed to detect all the series with lesions. They were required to label each series rapidly and independently, using “positive” once the first lesion was recognized or “negative” if none was found. Following each labeling, a confidence rating was assigned using a sliding scale (0–100), with higher points indicating more self-confidence. One month later, on a per-lesion basis (lesion detection), the readers were asked to segment all the lesions. We recorded the labels, detection time, and self-confidence ratings in patient detection and the segmented regions in lesion detection. Self-confidence was defined as the reader’s satisfaction with the judgment in whether the patients had any positive lesions.

### Statistical analysis

Statistical analyses were performed with IBM SPSS Statistics 22.0 (IBM Corp, Armonk, NY, USA; RRID: SCR\_019096). All the data from the testing set were included for statistical analysis. True positives, true negatives, false positives, and false negatives were counted

for patient detection, and true positives, false positives, and false negatives were counted for lesion detection. The metrics of the readers’ performance were calculated, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F measure (21) for patient detection and sensitivity, PPV, and F measure for lesion detection. The overall reader performance was calculated by combining the results of all the readers (see [Appendix 1](#) for calculation methods). Detection time and self-confidence in patient detection and performance metrics in lesion detection were compared between the use of syn-MRI and CT. Paired *t*-tests or Wilcoxon signed-rank tests were used for these comparisons. Spearman correlation analysis was used to assess the correlations between similarity ratings and time-related clinical characteristics and between reader performance metrics and clinical and lesion characteristics.  $P < 0.05$  was considered to indicate a significant difference. Multiple comparisons were corrected with the Bonferroni method.

## Results

### Study participants

Patient characteristics are shown in *Table 2*. Age, sex, National Institute of Health Stroke Scale (NIHSS) score, onset-to-CT time, and CT-to-MRI time were matched between sets. Among the entire patient cohort, 66.3% (128/193) had large artery atherosclerosis, 14.5% (28/193) had small artery occlusion, 9.8% (19/193) had cardioembolism, and 9.3% (18/193) were cryptogenic. Regarding the vascular territories, 83.4% (161/193) involved the anterior circulation, 10.9% (21/193) involved the posterior circulation, 2.1% (4/193) involved both, and 3.6% (7/193) lacked specific findings. After the CT scan, 86.5% (167/193) only received supportive care, and the others additionally underwent intravenous thrombolysis (7.8%, 15/193), mechanical thrombectomy (3.6%, 7/193), or both (2.1%, 4/193).

### Reference lesions

Among 50 patients in the testing set, 540 hyperintense lesions (median size, 1.7 cm<sup>2</sup>; interquartile range, 0.7–5.7 cm<sup>2</sup>) were identified on FLAIR imaging; the remaining 3 patients showed no visible lesions. In terms of lesion size, 83.9% (453/540) were smaller than 10 cm<sup>2</sup>, 9.8% (53/540) measured 10–40 cm<sup>2</sup>, and 6.3% (34/540) were larger than 40 cm<sup>2</sup>.

**Table 2** Patient characteristics

Characteristics	Training set (n=140)	Testing set (n=53)	P value
Age (years), median [IQR]	67 [56–77]	72 [59–77]	0.43
Sex, male, n (%)	88 (62.9)	39 (73.6)	0.16
NIHSS score, median [IQR]	5 [3–14]	6 [2–14]	0.81
Onset-to-CT time (hours), median [IQR]	3 [2–4]	3 [2–4]	0.12
CT-to-MRI time (days), median [IQR]	2 [1–3]	2 [1–3]	0.50

IQR, interquartile range; NIHSS, National Institute of Health Stroke Scale.

**Table 3** Similarity rating scores between syn-MRI and ground-truth FLAIR

Grade	Number of patients
1	4
2	20
3	17
4	12 <sup>†</sup>

<sup>†</sup>, including 3 patients with no lesions on either syn-MRI or ground-truth FLAIR. Syn-MRI, synthetic MRI; FLAIR, fluid-attenuated inversion recovery.

### Image similarity between syn-MR and ground-truth FLAIR images

The similarity ratings for image quality are shown in *Table 3*. In the testing set, 92.5% of patients (49/53) had hyperintense Grade 2–4 brain lesions, including 3 with no lesions on the ground-truth images and 46 with lesions overlapping partially to completely. Grade 1 lesions were observed in the remaining patients (7.5%, 4/53), including 3 with small- to medium-sized infarcts in the frontal lobe or basal ganglia and 1 with a large infarct in the middle cerebral artery territory (onset-to-CT time, 2–5 hours; CT-to-MRI time, 1–2 days). None of the Grade 1 patients showed positive findings on the baseline CT (*Figure 3*). The quantitative evaluation showed that our model achieved a PSNR value of 24.30 dB and an SSIM value of 0.857.

The effects of imaging timing on image similarity were assessed with Spearman correlation analysis, showing no correlations in the similarity ratings with either the onset-to-CT time ( $P=0.51$ ) or the CT-to-MRI time ( $P=0.07$ ).

### Overall reader performance

The overall reader performance with CT and syn-

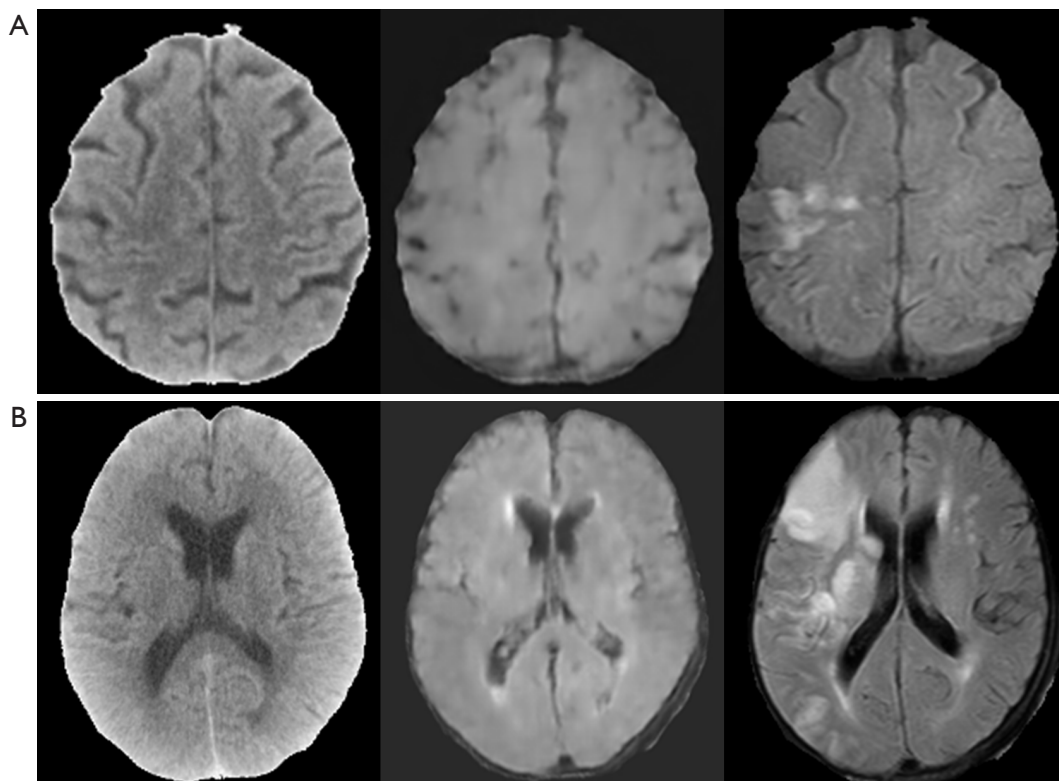
MRI is shown on a per-patient (patient detection) and per-lesion (lesion detection) basis in *Table 4*; individual reader performances are shown in *Appendix 1*. In patient detection, 159 pairs of CT and syn-MRI series were labeled. A total of 93 patients were improperly labeled on CT as “negative”, but 75% (70/93) were corrected with syn-MRI. *Figure 4* shows examples of brain lesions on both imaging modalities. Compared with CT, syn-MRI led to an increase in the sensitivity from 38% (57/150; 95% CI: 30–46%) to 82% (123/150; 95% CI: 75–88%), in the NPV from 6% (6/99; 95% CI: 3–13%) to 10% (3/30; 95% CI: 8–65%), and in the F measure from 0.56 to 0.88. The PPV with syn-MRI [95% (123/129; 95% CI: 90–98%)] was similar to that with CT [95% (57/60; 95% CI: 85–99%)]; however, the specificity decreased from 67% (6/9; 95% CI: 31–91%) to 33% (3/9; 95% CI: 9–69%). The median detection time was 9 seconds with syn-MRI versus 8 seconds with CT, and the self-confidence was 96 points with syn-MRI versus 100 points with CT. No differences were found in the detection time ( $P=0.29$ ) or self-confidence ( $P=0.14$ ).

In lesion detection, 15% (77/517) that were unseen on CT were detected on syn-MRI. Overall, compared with CT, syn-MRI resulted in an increase in the sensitivity from 4% (68/1,620; 95% CI: 3–5%) to 16% (262/1,620; 95% CI: 14–18%) ( $R=0.32$ , corrected  $P<0.001$ ), in the PPV from 31% (68/216; 95% CI: 25–38%) to 48% (262/548; 95% CI: 44–52%) ( $R=0.34$ , corrected  $P=0.002$ ), and in the F measure from 0.07 to 0.27 ( $R=0.48$ , corrected  $P=0.007$ ) (*Figure 5*).

### Effects of other factors on reader performance

When detecting lesions on syn-MRI, the overall sensitivity ( $R_s=0.17$ ,  $P=0.04$ ) and PPV ( $R_s=0.22$ ,  $P=0.02$ ) were positively, albeit mildly, correlated with the baseline NIHSS score, and the PPV was negatively correlated with the onset-to-CT time ( $R_s=-0.22$ ,  $P=0.02$ ). No correlations were





**Figure 3** Examples of Grade 1 similarity between synthetic MRI and ground-truth FLAIR imaging. (A) An 80-year-old female with an acute medium-sized infarction in the right frontal lobe (onset-to-CT time, 2 hours; CT-to-MRI time, 2 days). (B) A 76-year-old male with an acute large-scale infarction in the right middle cerebral artery territory (onset-to-CT time, 2 hours; CT-to-MRI time, 2 days). Head CT (left) and synthetic MRI (middle) of both patients fail to show any definite infarcts, but FLAIR (right) confirms the acute lesions as hyperintensities. MRI, magnetic resonance imaging; FLAIR, fluid-attenuated inversion recovery; CT, computed tomography.

found between the other metrics (sensitivity, PPV, and F measure) and the CT-to-MRI time ( $P=0.07$ – $0.26$ ). Stratified by treatment, the overall sensitivity with syn-MRI was the highest among patients who received supportive care [20% (234/1,173; 95% CI: 18–22%)] (Table 5).

The lesion number on follow-up MRI was positively correlated with the overall PPV ( $R_s=0.28$ ,  $P=0.002$ ) but negatively correlated with the F measure ( $R_s=-0.29$ ,  $P=0.01$ ), and no association was seen with sensitivity. The overall performance metrics, including sensitivity ( $R_s=0.35$ ,  $P<0.001$ ), PPV ( $R_s=0.43$ ,  $P<0.001$ ), and F measure ( $R_s=0.24$ ,  $P=0.04$ ), were mildly to moderately correlated with the maximum lesion size on follow-up MRI. Stratified by lesion size, lesions larger than  $40\text{ cm}^2$  were the most sensitively detected [62% (63/102, 95% CI: 52–71%)]. Improvements in the overall sensitivity were observed within all strata, with the greatest effect found for lesions measuring  $8$ – $10\text{ cm}^2$  (Table 5). Regarding the vascular territories, lesions

involving the anterior circulation [17% (238/1,410, 95% CI: 15–19%)] were more sensitively detected than those involving the posterior circulation [11% (24/210, 95% CI: 8–17%)] (Table 5).

## Discussion

In this work, we proposed a GAN model for CT-to-MR brain image synthesis in patients with suspected AIS and evaluated the feasibility of detecting suspected brain lesions with syn-MRI. Image similarity was evaluated between syn-MRI and the ground truth, and the detection performance of readers was tested. To our knowledge, deep-learning methods for noncontrast CT-to-MR image synthesis have not been well established for AIS detection. Our GAN model generated syn-MRI from CT images in 52 seconds on average, and the readers took an extra 9 seconds for patient detection. The GAN-based method synthesized

**Table 4** Overall reader performance with CT and synthetic MRI in the testing set

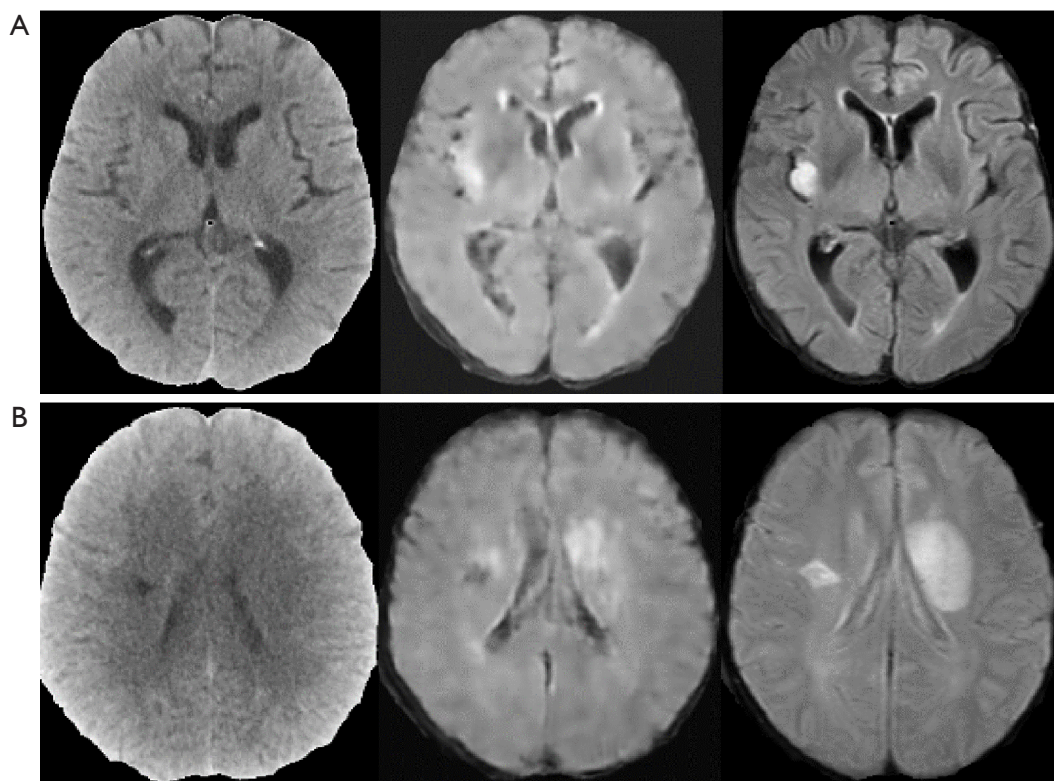
Performance metrics	CT (n=159)	Syn-MRI (n=159)	R value	P value
Patient detection				
True positive	57	123	n/a	n/a
True negative	6	3	n/a	n/a
False positive	3	6	n/a	n/a
False negative	93	27	n/a	n/a
Sensitivity (%) <sup>†</sup>	38 (57/150) [30, 46]	82 (123/150) [75, 88]	n/a	n/a
Specificity (%) <sup>†</sup>	67 (6/9) [31, 91]	33 (3/9) [9, 69]	n/a	n/a
PPV (%) <sup>†</sup>	95 (57/60) [85, 99]	95 (123/129) [90, 98]	n/a	n/a
NPV (%) <sup>†</sup>	6 (6/99) [3, 13]	10 (3/30) [8, 65]	n/a	n/a
F measure	0.56	0.88	n/a	n/a
Detection time (s) <sup>*</sup>	8 (6, 18)	9 (4, 18)	-0.06	0.29
Self-confidence <sup>*</sup>	100 (82, 100)	96 (85, 100)	-0.08	0.14
Lesion detection				
True positive	68	262	n/a	n/a
False positive	148	286	n/a	n/a
False negative	1,552	1,358	n/a	n/a
Sensitivity (%) <sup>†</sup>	4 (68/1,620) [3, 5]	16 (262/1,620) [14, 18]	0.32	<0.001 <sup>‡</sup>
PPV (%) <sup>†</sup>	31 (68/216) [25, 38]	48 (262/548) [44, 52]	0.34	0.002 <sup>‡</sup>
F measure	0.07	0.27	0.48	0.007 <sup>‡</sup>

<sup>†</sup>, data in parentheses are the numerator and denominator, with 95% CI in brackets. <sup>\*</sup>, data are the median, with interquartile range in parentheses. <sup>‡</sup>, corrected with the Bonferroni method. True negatives were not counted at the lesion level. Syn-MRI, synthetic magnetic resonance imaging; PPV, positive predictive value; NPV, negative predictive value; n/a, not applicable.

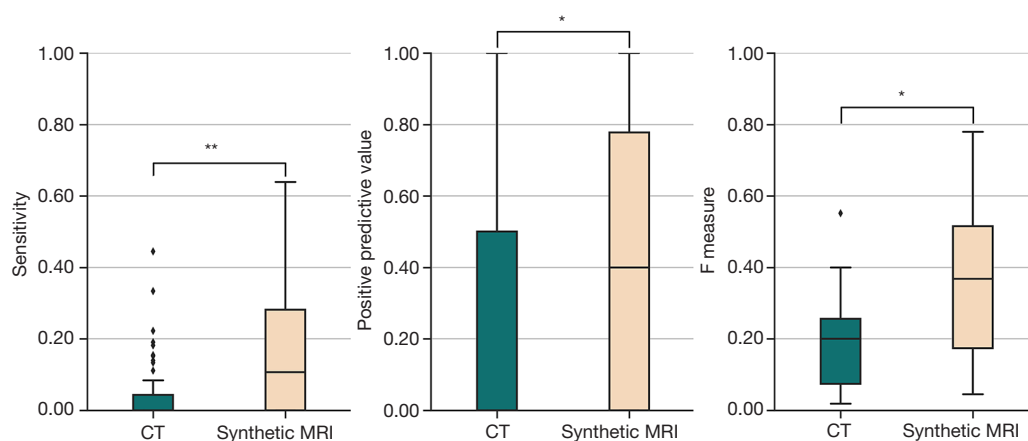
targeted lesions for 92.5% of the patients with suspected AIS, with a favorable SSIM value but a weak PSNR value. For all the readers, an additional 75% of patients and 15% of lesions missed on CT were detected on syn-MRI. Compared with that using CT, the overall sensitivity using syn-MRI rose to 82% in patient detection and 16% in lesion detection, but the specificity dropped to 33% in patient detection. These findings suggest that GAN-based syn-MRI might help readers increase detection sensitivity for potential individuals among suspected AIS patients, but model optimization for lesion enhancement and expert discrimination for definite diagnosis are needed.

Many studies on stroke image analysis have attempted to improve lesion visibility on CT, such as lesion-to-tissue contrast amplification (22,23), voxelwise comparison (16,17,24,25), and texture analysis (26-30). Despite heterogeneous methodologies and sensitivities (31),

substantial evidence indicates that certain radiological features, even if they are subtle or invisible, could be captured to improve human perception of the underlying pathologies. Comparatively, our method has strengths in design concept, bias and risk control, and knowledge preparation. First, most methods assist in formal interpretation as contrast amplifiers (22,23) or density searchers (24,25). Our model targets early detection but not final diagnosis and probably better serves nonexperts onsite than radiologists in-door. Second, patients with suspected rather than diagnosed stroke were included; unlike those with a definite diagnosis (8,17), use of this population helps avoid a major source of bias and approaches real-world detection scenarios. Third, neither transhemispheric comparisons (31) nor asymmetry corrections (17) were needed, potentially relieving the misdiagnosis of symmetrical strokes and the overdiagnosis of other



**Figure 4** Examples of patient detection using synthetic MRI versus CT. (A) A 72-year-old man with a small acute infarct in the right insula. Head CT (left) fails to show any abnormal parenchymal hypodensities. In contrast, synthetic MRI (middle) shows regional hyperintensity in the right insula, consistent with FLAIR (right). (B) A 48-year-old woman with bilateral infarcts of different phases, an acute lesion in the left basal ganglia and a chronic lesion (lacune) in the right corona radiata. Head CT (left) shows the chronic but not the acute lesion, while synthetic MRI (middle) shows both, which are confirmed on FLAIR (right). MRI, magnetic resonance imaging; CT, computed tomography; FLAIR, fluid-attenuated inversion recovery.



**Figure 5** Overall reader performance in lesion detection. The overall sensitivity ( $P<0.001$ ), positive predictive value ( $P=0.002$ ), and F measure ( $P=0.007$ ) were significantly improved when using synthetic MRI versus CT. \*,  $P<0.05$ ; \*\*,  $P<0.01$ . MRI, magnetic resonance imaging; CT, computed tomography.

**Table 5** Stratification analysis of the overall sensitivity in lesion detection

Stratification factors	CT (n=159)	Syn-MRI (n=159)
Treatment		
Thrombolysis	0 (1/318) [0, 2]	6 (20/318) [4, 10]
Thrombectomy	5 (4/81) [2, 13]	10 (8/81) [5, 19]
Thrombolysis + thrombectomy	0 (0/48) [0, 9]	0 (0/48) [0, 9]
Supportive care	5 (63/1,173) [4, 7]	20 (234/1,173) [18, 22]
Lesion size (cm <sup>2</sup> )		
0–2	1 (11/870) [1, 2]	6 (49/870) [4, 7]
2–4	5 (13/243) [3, 10]	18 (43/243) [13, 23]
4–6	10 (12/120) [6, 17]	23 (27/120) [16, 31]
6–8	7 (5/72) [3, 16]	26 (19/72) [17, 38]
8–10	2 (1/54) [0, 11]	28 (15/54) [17, 42]
10–20	3 (3/87) [1, 10]	23 (20/87) [15, 34]
20–40	13 (9/72) [6, 23]	36 (26/72) [26, 48]
≥40	14 (14/102) [8, 22]	62 (63/102) [52, 71]
Involved artery territory		
Anterior circulation	4 (60/1,410) [3, 5]	17 (238/1,410) [15, 19]
Posterior circulation	4 (8/210) [2, 8]	11 (24/210) [8, 17]

Data are sensitivity (%), with the numerator and denominator in parentheses and 95% CI in brackets. Syn-MRI, synthetic magnetic resonance imaging.

contralateral pathologies. Fourth, this algorithm dispenses with clinical judgments as prior knowledge input (17) and thus is emergency-adapted and reader-friendly, particularly in detection tasks with inadequate information.

Generally, our GAN model generated targeted brain lesions rapidly in this study, which showed a favorable SSIM and no distortional or hallucinogenic effects. This supports a possible application of the GAN-based method for rapid stroke detection. Notably, evaluation of image similarity also revealed the weakness of the present model. Among patients with positive lesions, the synthesized lesions of 20 patients only partially overlapped with the ground truth. This implies that some synthesized lesions tended to be smaller than those on follow-up MRI. In addition, the weak PSNR indicates that the signal of the hyperintense lesions is expected to be enhanced in the future. We could improve our framework in two ways. First, the current framework could be modified by using a more complex network and learning strategy. This would help decouple the style and content space of the paired images and address the problem

of domain-specific deformation to generate MR images with higher resolution but no structural distortions. Second, more lesion-specific constraints could be added, such as a lesion consistency loss or an auxiliary attention module. This would help ensure that the model focuses more on the preservation of stroke lesion information during training so that more tiny lesions can be detected on syn-MRI.

In four patients with no positive lesions on CT, the present model failed to generate hyperintense lesions on syn-MRI. This could be partially explained by the poor PSNR value. More importantly, it is well established that the course of AIS is dynamic. Furthermore, imaging signs could be influenced by multiple factors, such as the stroke subtype, vascular status, timing of imaging, and treatment. Although imaging timing was not related to image quality or detection sensitivity in our study, it should be noted that the early signs on CT could be very subtle or invisible, possibly carrying limited information to be learned by the model. Moreover, the time interval between our scans was relatively long, during which illness progression and

treatment might also complicate the learning process. In this context, when the syn-MRI is negative, the images should be interpreted with caution, and suspicion of AIS should not be excluded.

Using the GAN model, readers exhibited favorable performance in sensitivity relative to the other metrics and on the per-patient basis compared with the per-lesion basis. This suggests value in detecting potential patients rather than specific lesions. Conventionally, the oldest studies are placed at the top of a radiologist's reading list, which risks reporting delays for newly scanned patients with severe disease. A recent management system showed a reduced turnaround time by 17.3 minutes when the worklist was adjusted based on stroke code (32). In this sense, screening for potential patients may expedite the workflow by informing the radiologist to pick studies in a timely manner. Users of our method could be either novices (radiographers and paramedics of mobile stroke units) or rural radiologists who could contact neuroradiologists for identification via in-hospital or teleradiology systems (33).

Notably, specificity decreased in patient detection. This probably suggests that noncontrast CT or synthetic MR images might inherently provide insufficient information for patient confirmation. This could be attributed to the definitions of "sample" and "lesion", by which some, albeit not all, stroke-mimic patients were excluded, while extensive white matter hyperintensities on FLAIR images might be included. Given the low true-negative and high false-positive results, our model could inadvertently generate excessive alerts. Since imaging-to-needle time is a major contributor to delays in thrombolysis for AIS (4) and the chance of missing a devastating diagnosis should be mitigated (34), the GAN model should be preferred for ruling-in rather than ruling-out practice. Accordingly, synthetic images may provide a prealert reference for priority adjustment but may not serve as a diagnostic tool for decision making. Once informed, the radiologist should view the original images quickly but cautiously.

The correlation analysis revealed that a higher sensitivity was mildly related to worse neurologic deficits at the time of presentation. Stratification analysis showed that the patients with the highest sensitivity may have had the largest lesions. Since neurological dysfunction is closely associated with early CT signs (35) and ischemic volume (36), we speculate that large-scale hypodensities and other positive signs on CT may be an essential feature learned and transformed by the GAN model. The highest sensitivity was also found in patients receiving supportive care rather than those

receiving recanalization therapies. Among these patients, the pathophysiological changes in the former relative to the latter may better approximate the changes seen in the natural course of the disease. Given the potential effects of treatment on imaging signs, our GAN model might play a more significant role in lesion detection before but not after treatment. Additionally, the readers detected lesions of the anterior circulation more sensitively than those of the posterior circulation. This is consistent with the limitations in applying CT to the posterior fossa. Here, noncontrast CT provides suboptimal visualization of the structures due to obscuration by beam-hardening artifacts produced by the bony cranial base, and the lesions are often too small to be interpreted relative to those of the anterior circulation (37). Therefore, our model might be more beneficial for large and anterior circulation infarcts, especially with early CT signs.

Several limitations should be noted in this study. First, the original images were derived from a single center with dedicated scanners; this makes the study prone to spectrum bias and weakens the generalizability of the findings, warranting multicenter studies with more scanners from multiple vendors. Second, the time interval between the baseline and follow-up scans was relatively long, and the therapeutic approach was not controlled during this period. To reduce the influence of imaging timing and treatment on image quality, it would be ideal to obtain both emergency CT and MR images at baseline. Third, to reflect pathologies over a large time span in our sample, we preferred FLAIR over diffusion-weighted sequences as the "gold" standard reference. This may have accounted for the inclusion of nonspecific and in-evolution lesions. In our ongoing research, another GAN model is currently being trained with diffusion-weighted data obtained earlier than FLAIR images and reader performance will be compared between the models. Fourth, the current GAN model needs to be improved to enhance image similarity due to the poor signal-to-noise ratio and insufficient extent of the synthesized lesions. Modifying the current training strategy at both the image and lesion levels would be a solution. Fifth, time consumption was separately recorded in image synthesis and detection, short of the overall assessment in the real settings. Future work will involve implanting the GAN model into our hospital system and prospectively testing workflow efficiency. Sixth, performance was tested with only a small number of readers. More readers with varying levels of expertise should be considered, and related stratification analysis is further required.



## Conclusions

The GAN model holds great potential for generating synthetic MR images from noncontrast CT images, but the image similarity of the brain lesions needs to be improved by modifying the model and optimizing the imaging data for training. CT-based syn-MRI using GANs could assist in sensitively detecting potential individuals among patients with suspected AIS, although further expert discrimination is strongly recommended. When synthetic MR images are negative, they should be interpreted with caution, and stroke should not be excluded. Future GAN-based methods might be more beneficial for quickly searching for large infarcts of the anterior circulation territory before treatment and more suitable for nonradiologists for timely detection before formal interpretation than for neuroradiologists for final diagnosis.

## Acknowledgments

The authors thank Xiaodan Bao, BSc (Chengdu Second People's Hospital, Chengdu, China) for test participation and Xue Pei, BSc (West China Fourth Hospital, Sichuan University, Chengdu, China) for data collection. A part of this manuscript has been presented in the oral section of ISMRM & SMRT Annual Meeting & Exhibition on May 19, 2021.

**Funding:** This work was supported by the National Natural Science Foundation of China (grant numbers 82102000, 81671664, 61876032), National Key Research and Development Program of China (grant number 2018YFC0116400), Sichuan Science and Technology Program (grant number 2019YJ0155), and 1-3-5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (grant number ZYJC18020).

## Footnote

**Reporting Checklist:** The authors have completed the STARD reporting checklist. Available at <https://atm.amegroups.com/article/view/10.21037/atm-21-4056/rc>

**Data Sharing Statement:** Available at <https://atm.amegroups.com/article/view/10.21037/atm-21-4056/dss>

**Conflicts of Interest:** All authors have completed the ICMJE uniform disclosure form (available at <https://atm.amegroups.com/article/view/10.21037/atm-21-4056/coif>).

NH, TZ, SG, and SL report two related patents pending (Chinese patent number 202010522445.1; PCT number PCT/CN2020/118667). NH reports that this work was supported by National Natural Science Foundation of China (grant number 82102000) and Sichuan Science and Technology Program (grant number 2019YJ0155). BS reports that this work was supported by National Key Research and Development Program of China (grant number 2018YFC0116400). SG reports that this work was supported by National Natural Science Foundation of China (grant number 61876032). SL reports that this work was supported by National Natural Science Foundation of China (grant number 81671664) and 1-3-5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (grant number ZYJC18020). The other authors have no conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Ethics Committee of Biomedical Research, West China Hospital of Sichuan University (No. 2020-394) and individual consent for this retrospective analysis was waived.

**Open Access Statement:** This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Powers WJ, Rabinstein AA, Ackerson T, et al. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: A guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2019;50:e344-418. Erratum in: *Stroke* 2019;50:e440-1.
2. von Kummer R, Dzialowski I. Imaging of cerebral ischemic edema and neuronal death. *Neuroradiology*

- 2017;59:545-53.
3. Allen LM, Hasso AN, Handwerker J, et al. Sequence-specific MR imaging findings that are useful in dating ischemic stroke. *Radiographics* 2012;32:1285-97; discussion 1297-9.
  4. Sauser K, Levine DA, Nickles AV, et al. Hospital variation in thrombolysis times among patients with acute ischemic stroke: The contributions of door-to-imaging time and imaging-to-needle time. *JAMA Neurol* 2014;71:1155-61.
  5. Berlyand Y, Raja AS, Dorner SC, et al. How artificial intelligence could transform emergency department operations. *Am J Emerg Med* 2018;36:1515-7.
  6. Yedavalli VS, Tong E, Martin D, et al. Artificial intelligence in stroke imaging: Current and future perspectives. *Clin Imaging* 2021;69:246-54.
  7. He L, Yu H, Shi L, et al. Equity assessment of the distribution of CT and MRI scanners in China: a panel data analysis. *Int J Equity Health* 2018;17:157.
  8. Rubin J, Abulnaga SM. CT-To-MR conditional generative adversarial networks for ischemic stroke lesion segmentation. In: *Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI)*, Xi'an, China. June 10-13, 2019.
  9. Kearney V, Ziemer BP, Perry A, et al. Attention-Aware Discrimination for MR-to-CT Image Translation Using Cycle-Consistent Generative Adversarial Networks. *Radiol Artif Intell* 2020;2:e190027.
  10. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. *Med Image Anal* 2019;58:101552.
  11. Dar SU, Yurt M, Karacan L, et al. I Image Synthesis in Multi-Contrast MRI With Conditional Generative Adversarial Networks. *IEEE Trans Med Imaging* 2019;38:2375-88.
  12. Yu B, Zhou L, Wang L, et al. Ea-GANs: Edge-Aware Generative Adversarial Networks for Cross-Modality MR Image Synthesis. *IEEE Trans Med Imaging* 2019;38:1750-62.
  13. Nie D, Trullo R, Lian J, et al. Medical Image Synthesis with Deep Convolutional Adversarial Networks. *IEEE Trans Biomed Eng* 2018;65:2720-30. Erratum in: *IEEE Trans Biomed Eng* 2020;67:2706.
  14. Jin CB, Kim H, Liu M, et al. Deep CT to MR Synthesis Using Paired and Unpaired Data. *Sensors (Basel)* 2019;19:2361.
  15. Gomolka RS, Chrzan RM, Urbanik A, et al. A Quantitative Method Using Head Noncontrast CT Scans to Detect Hyperacute Nonvisible Ischemic Changes in Patients With Stroke. *J Neuroimaging* 2016;26:581-7.
  16. Lisowska A, O'Neil A, Dilys V, et al. Context-aware convolutional neural networks for stroke sign detection in non-contrast CT scans. In: Hernández MV, González-Castro V, editors. *Medical image understanding and analysis*. Cham: Springer, 2017;494-505.
  17. Qiu W, Kuang H, Teleg E, et al. Machine Learning for Detecting Early Infarction in Acute Stroke with Non-Contrast-enhanced CT. *Radiology* 2020;294:638-44.
  18. Isola P, Zhu J, Zhou T, et al. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii. July 21-26, 2017.
  19. Sara U, Akter M, Uddin MS. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. *J Comput Commun* 2019;7:8-18.
  20. Setiadi DRIM. PSNR vs SSIM: imperceptibility quality assessment for image steganography. *Multimed Tools Appl* 2021;80:8423-44.
  21. Hsieh YZ, Luo YC, Pan C, et al. Cerebral small vessel disease biomarkers detection on MRI-sensor-based image and deep learning. *Sensors (Basel)* 2019;19:2573.
  22. Bier G, Bongers MN, Ditt H, et al. Accuracy of non-enhanced CT in detecting early ischemic edema using frequency selective non-linear blending. *PloS One* 2016;11:e0147378.
  23. Przelaskowski A, Sklinda K, Bargieł P, et al. Improved early stroke detection: wavelet-based perception enhancement of computerized tomography exams. *Comput Biol Med* 2007;37:524-33.
  24. Gillebert CR, Humphreys GW, Mantini D. Automated delineation of stroke lesions using brain CT images. *Neuroimage Clin* 2014;4:540-8.
  25. Srivatsan A, Christensen S, Lansberg MG. A relative noncontrast CT map to detect early ischemic changes in acute stroke. *J Neuroimaging* 2019;29:182-6.
  26. Davis A, Gordillo N, Montseny E, et al. Automated detection of parenchymal changes of ischemic stroke in non-contrast computer tomography: A fuzzy approach. *Biomed Signal Process Control* 2018;45:117-27.
  27. Maya BS, Asha T. Automatic detection of brain strokes in CT images using soft computing techniques. In: Hemanth J, Balas VE, editors. *Biologically rationalized computing techniques for image processing applications*. Cham: Springer, 2018;85-109.
  28. Peter R, Korfiatis P, Blezek D, et al. A quantitative symmetry-based analysis of hyperacute ischemic stroke lesions in noncontrast computed tomography. *Med Phys* 2017;44:192-9.

29. Rebouças Filho PP, Sarmiento RM, Holanda GB, et al. New approach to detect and classify stroke in skull CT images via analysis of brain tissue densities. *Comput Methods Programs Biomed* 2017;148:27-43.
30. Tang FH, Ng DK, Chow DH. An image feature approach for computer-aided detection of ischemic stroke. *Comput Biol Med* 2011;41:529-36.
31. Mikhail P, Le MGD, Mair G. Computational image analysis of nonenhanced computed tomography for acute ischaemic stroke: A systematic review. *J Stroke Cerebrovasc Dis* 2020;29:104715.
32. Osborne TF, Grabiell AJ, Clark RH. The benefit of a triage system to expedite acute stroke head computed tomography interpretations. *J Stroke Cerebrovasc Dis* 2018;27:1190-3.
33. Mathur S, Walter S, Grunwald IQ, et al. Improving prehospital stroke services in rural and underserved settings with mobile stroke units. *Front Neurol* 2019;10:159.
34. Grant K, McParland A, Mehta S, et al. Artificial intelligence in emergency medicine: Surmountable barriers with revolutionary potential. *Ann Emerg Med* 2020;75:721-6.
35. Scott JN, Buchan AM, Sevick RJ. Correlation of neurologic dysfunction with CT findings in early acute stroke. *Can J Neurol Sci* 1999;26:182-9.
36. Furlanis G, Ajčević M, Stragapede L, et al. Ischemic volume and neurological deficit: correlation of computed tomography perfusion with the national institutes of health stroke scale score in acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2018;27:2200-7.
37. Nouh A, Remke J, Ruland S. Ischemic posterior circulation stroke: a review of anatomy, clinical presentations, diagnosis, and current management. *Front Neurol* 2014;5:30.

**Cite this article as:** Hu N, Zhang T, Wu Y, Tang B, Li M, Song B, Gong Q, Wu M, Gu S, Lui S. Detecting brain lesions in suspected acute ischemic stroke with CT-based synthetic MRI using generative adversarial networks. *Ann Transl Med* 2022;10(2):35. doi: 10.21037/atm-21-4056