Medical Image Analysis 87 (2023) 102814

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Virtual high-resolution MR angiography from non-angiographic multi-contrast MRIs: synthetic vascular model populations for in-silico trials

Yan Xia [a,*], Nishant Ravikumar [a], Toni Lassila [a], Alejandro F. Frangi [a,b,c,d]

[a] Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), School of Computing, University of Leeds, Leeds, UK
[b] Leeds Institute for Cardiovascular and Metabolic Medicine (LICAMM), School of Medicine, University of Leeds, Leeds, UK
[c] Medical Imaging Research Center (MIRC), Cardiovascular Science and Electronic Engineering Departments, KU Leuven, Leuven, Belgium
[d] Alan Turing Institute, London, UK

ARTICLE INFO

Keywords:
MR angiography
Multi-contrast MRI
Medical image synthesis
Conditional generative adversarial net

ABSTRACT

Despite success on multi-contrast MR image synthesis, generating specific modalities remains challenging. Those include Magnetic Resonance Angiography (MRA) that highlights details of vascular anatomy using specialised imaging sequences for emphasising inflow effect. This work proposes an end-to-end generative adversarial network that can synthesise anatomically plausible, high-resolution 3D MRA images using commonly acquired multi-contrast MR images (e.g. T1/T2/PD-weighted MR images) for the same subject whilst preserving the continuity of vascular anatomy. A reliable technique for MRA synthesis would unleash the research potential of very few population databases with imaging modalities (such as MRA) that enable quantitative characterisation of whole-brain vasculature. Our work is motivated by the need to generate digital twins and virtual patients of cerebrovascular anatomy for in-silico studies and/or in-silico trials. We propose a dedicated generator and discriminator that leverage the shared and complementary features of multi-source images. We design a composite loss function for emphasising vascular properties by minimising the statistical difference between the feature representations of the target images and the synthesised outputs in both 3D volumetric and 2D projection domains. Experimental results show that the proposed method can synthesise high-quality MRA images and outperform the state-of-the-art generative models both qualitatively and quantitatively. The importance assessment reveals that T2 and PD-weighted images are better predictors of MRA images than T1; and PD-weighted images contribute to better visibility of small vessel branches towards the peripheral regions. In addition, the proposed approach can generalise to unseen data acquired at different imaging centres with different scanners, whilst synthesising MRAs and vascular geometries that maintain vessel continuity. The results show the potential for use of the proposed approach to generating digital twin cohorts of cerebrovascular anatomy at scale from structural MR images typically acquired in population imaging initiatives.

1. Introduction

Magnetic resonance angiography (MRA) images are clinically used to evaluate vascular anatomy accurately, but not included in standard scanning protocols because of the long acquisition time required. This study focuses on synthesising subject-specific MRA images, given other available multi-contrast but non-angiographic MR images (e.g. T1/T2/PD-weighted MRIs) for the same subject. A reliable technique for synthesising MRA images from other multi-contrast MR images is valuable because — (i) there are few large databases with imaging modalities (such as MRA) that enable quantitative characterisation of whole-brain vasculature. To the best of our knowledge, very few publicly available datasets exist that provide MRA images of the brain for multiple individuals, the largest of which is the Information eXtraction from Images (IXI) database containing 600 subject images;[1] and (ii) no extensive public database exists that provides image-derived anatomical models of brain vasculature, ready for use in studies investigating cerebrovascular haemodynamics. The largest public database of whole-brain vascular models is the BraVa database (Wright et al., 2013), which comprises just 61 samples of healthy adults. Previous studies that have simulated and assessed blood flow in cerebral vasculature have been limited in sample size, ranging from tens (typical) (Larrabide et al., 2015; Holmgren et al., 2021) to a few hundred (rare) (Cebral et al., 2011). This paucity of data has thus limited the size of in-silico studies conducted thus far to assess cerebrovascular

* Corresponding author.
E-mail address: y.xia@leeds.ac.uk (Y. Xia).

[1] http://brain-development.org/ixi-dataset/

https://doi.org/10.1016/j.media.2023.102814
Received 11 February 2022; Received in revised form 4 April 2023; Accepted 8 April 2023
Available online 20 April 2023

haemodynamics, thereby limiting the degree to which inter-patient variability has been characterised and leveraged to improve our understanding of the interplay between anatomy, physiology and pathological processes.

As different modality/contrast images are acquired from the same anatomical structure using different imaging settings, underlying correlation between those images and mappings can be learned to transform to each other. Here we use the term "modality" to denote different imaging techniques (such as CT, MR and PET, etc.), whilst the "contrast" indicates different contrast sequences (such as T1/T2/PD-weighted) in multi-contrast MRI scenario. Early work on cross-contrast image synthesis within the community utilised non-linear regression techniques, based on random forests for example, to predict the intensities of an image from the desired contrast, given single or multiple images from complementary contrasts as predictors (inputs) (Jog et al., 2014, 2017). Fully convolutional neural networks have also been investigated for multi-contrast image synthesis (Chartsias et al., 2017b; Joyce et al., 2017; Dewey et al., 2018, 2019; Mehta and Arbel, 2018). For instance, Chartsias et al. (2017b) and Joyce et al. (2017) proposed similar encoder–decoder neural networks for multi-contrast MRI synthesis that learned a shared modality-invariant latent representation from all input contrasts, which could be transformed into the target image via a learned decoder.

Inspired by recent advances in generative adversarial networks (GANs) and image-to-image translation techniques, cross-modality synthesis of missing or corrupted images has received increased attention. Adversarial learning has been demonstrated to yield high synthesis quality in cross-modality imaging applications such as CT to PET synthesis (Ben-Cohen et al., 2017; Bi et al., 2017), MR to CT synthesis (Wolterink et al., 2017a,b; Hiasa et al., 2018; Nie et al., 2018), CT to MR synthesis (Chartsias et al., 2017a), and retinal vessel map to image synthesis (Costa et al., 2017; Zhao et al., 2018). It is worth mentioning that studies regarding multi-contrast MRI synthesis also leverage GANs, such as in Dar et al. (2019) and Dalmaz et al. (2022). Techniques developed for multi-modal/contrast image synthesis leverage the complementary information and underlying correlations present in multi-source images of a given anatomical structure/organ. This is conceptually well-grounded as each modality or contrast captures specific characteristics of the underlying anatomy (i.e. provides additional information), enabling learning of more holistic representations that enhance image synthesis quality (Jog et al., 2014, 2015; Chartsias et al., 2017b; Jog et al., 2017; Joyce et al., 2017; Dewey et al., 2018; Mehta and Arbel, 2018; Wei et al., 2018; Olut et al., 2018; Dewey et al., 2019; Hagiwara et al., 2019; Lee et al., 2019; Li et al., 2019; Sharma and Hamarneh, 2019; Dar et al., 2020; Zhou et al., 2020; Lee et al., 2020). Several studies have demonstrated the ability of GANs and adversarial learning to interpolate the manifold of multi-contrast data with high-fidelity and synthesise high-quality images of the contrast of interest (Wei et al., 2018; Olut et al., 2018; Dewey et al., 2019; Hagiwara et al., 2019; Lee et al., 2019; Li et al., 2019; Sharma and Hamarneh, 2019; Dar et al., 2020; Lee et al., 2020; Zhou et al., 2020; Yurt et al., 2021). Important examples include the hybrid-fusion network (Hi-Net) (Zhou et al., 2020) that is based on adversarial learning and simultaneously explores the contrast-specific properties within each image and the correlations across contrasts, which are integrated using a layer-wise fusion strategy to effectively fuse information from multiple contrasts within different feature layers; and the collaborative generative adversarial network (CollaGAN) (Lee et al., 2020), which synergistically learns the joint image manifold of multiple MRI contrasts using multiple cycle consistency specifically for multiple inputs. These methods were shown to outperform several state-of-the-art synthesis methods, both quantitatively and qualitatively. Alternatively, Özbey et al. (2022) proposed an adversarial diffusion model with a cycle-consistent architecture devised with coupled diffusive and non-diffusive modules that bilaterally translate between two modalities and facilitate training on unpaired datasets. They showed that their model outperformed competing GAN and diffusion models.

Despite the success of multi-contrast MR image synthesis, several types of images are challenging to generate, including MRA images, which capture details of vascular anatomy. Thus far, few studies have explored cross-domain MRA synthesis using other multi-contrast MR images as inputs/predictors. Olut et al. (2018) extended the well-established pix2pix architecture for MRA synthesis by incorporating steerable filters of the target image and a Huber loss to highlight vesselness properties. However, the model was built on 2D images, and cannot guarantee vascular topology and continuity in 3D. Fujita et al. (2020) designed to output the weighted average of a single convolution and the U-net classification results at 5 raw 3D-quantification using an interleaved Look-Locker acquisition sequence with T2 preparation pulse (3D-QALAS) data and constructed a mapping function to convert the 3D-QALAS data to their corresponding MRA images. The authors compared their model with a simple linear combination model. The training and test datasets used in their study were five types of 3D-QALAS sequence images, which capture information of inflow effect and thus ease the difficulty of synthesising MRA images.

To the best of our knowledge, this is the first study to synthesise anatomically plausible, high-resolution 3D MRA images, whilst preserving the continuity of vascular anatomy using a combination of commonly acquired image sequences such as T1, T2 and PD-weighted MR contrasts. We demonstrate the performance of our approach across comprehensive experiments, including comparisons with state-of-the-art methods developed for multi-contrast MR image synthesis and importance assessment of various structural MR contrasts as inputs/predictors to drive MRA synthesis. Furthermore, we demonstrate the similarity in flow-derived parameters between geometries extracted from the original images and their synthesised counterparts highlights preservation of key vascular morphological characteristics. We show that the proposed approach, once pre-trained, can effectively use the small data regime to retrain and fine-tune the model to drive MRA synthesis through transfer learning, when the target domain is significantly different from the training domain. We also apply the proposed approach to completely unseen data that involves a single subject scanned at multiple centres across different scanners (no MRA acquisitions were included in the scanning protocol in this dataset), without fine-tuning and transfer learning. These results are promising and pave the way to create digital twin cohorts whereby in-silico trials of vascular implants can be evaluated at scale that no previous conventional trial has undertaken.

2. Methods

This study proposes an end-to-end GAN-based image translation framework that generates MRA images from existing multi-contrast structural MR images. To synthesise visually appealing, high resolution 3D MRA images, we proposed a dedicated generator and discriminator that leverage the shared and complementary features of multi-source images. The generator contains several residual blocks, where all normalisation layers are conditioned and modulated on available source images concatenated at the input level (i.e., T1, T2 and/or PD-weighted MRIs), to leverage complementary features from multi-contrast sequences and ensure that fine details are effectively propagated through image generation pathway. A multi-scale discriminator was employed to ensure the recovery of both global and local spatial features. All generators and discriminators used in this work are 3D models in order to leverage inter-slice context and 3D anatomical information for synthesising globally coherent images. We considered deep feature consistencies with feature matching and perceptual loss to faithfully synthesise vessel structures and preserve their continuity. We designed new loss terms in feature space for emphasising reconstruction of vasculature properties. The generator, discriminator and the loss functions will be discussed in subsequent sections. The proposed network architecture is illustrated in Fig. 1.

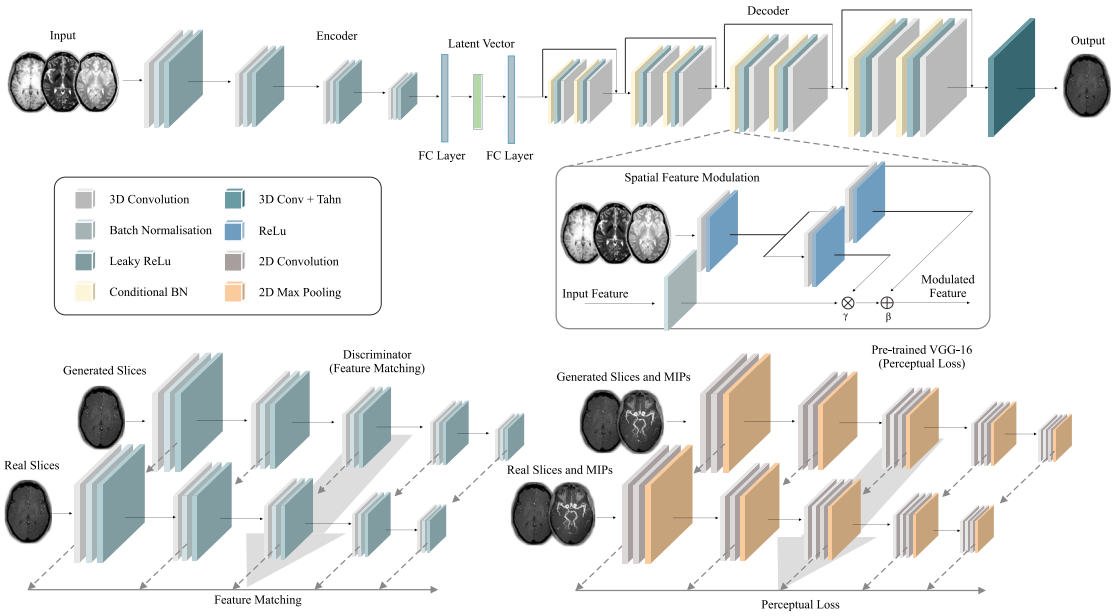


Fig. 1. Structure of the proposed GAN network for multi-contrast MRA image synthesis. The generator and discriminator used in this work are 3D models in order to leverage inter-slice context for generating globally coherent images. The generator contains several residual blocks. All normalisation layers are modulated with auxiliary, complementary information extracted from multi-contrast sequences and ensure that fine details are effectively propagated through the image generation pathway. A multi-scale discriminator is employed to ensure the recovery of both global and local spatial features. The feature consistencies with feature matching loss and perceptual loss were designed in 3D volume and 2D projection domain from different perspective views for emphasising reconstruction of vasculature properties.

2.1. Conditional GANs

Image-to-image translation architecture (Isola et al., 2017) (also known as the pix2pix model) has been shown to effectively map to target images y from statistically-dependent observed source images x. Two sub-models, a generator G and a discriminator D, are trained simultaneously. G aims to generate images that can fool the discriminator D, and D tries to classify whether generated images are real, from the domain, or fake, generated by the generator G. The optimisation of G and D can be reformulated as:

$$\mathcal{L}_{\text{cGAN}} = \mathbb{E}_{\mathbf{x}_m,y}\left[\log D\left(\mathbf{x}_m, y\right)\right] + \mathbb{E}_{\mathbf{x}_m}\left[\log\left(1 - D\left(\mathbf{x}_m, G\left(\mathbf{x}_m\right)\right)\right)\right], \tag{1}$$

where $\mathbf{x}_m : m = 1, 2, \ldots, M$ indicating the concatenation of M contrast structural MRI sequences in our case.

The reconstruction loss, e.g., Euclidean (L2) distance or L1 distance between the target and synthesised images, is integrated into the objective function as:

$$\mathcal{L}_{L1} = \mathbb{E}_{\mathbf{x}_m,y}\left[\left\|y - G\left(\mathbf{x}_m\right)\right\|_1\right]. \tag{2}$$

The loss function combines Eqs. (1) and (2). However, although impressive results were obtained for synthesising photographic images, the native image-to-image translation models may be unstable and prone to failure for synthesising medical images, which are more complex, noisy, higher dimensional and contain richer quantitative information. Our generative model is inspired by recent advances of GAN architectures and is able to generate HR anatomically plausible 3D MRA sequences, faithfully reproducing vessel structures and preserving continuity. The following subsections outline the detailed architecture of the proposed model.

2.2. Generator

The generator is an encoder–decoder model, where the encoder aims to extract the feature of source images, and the decoder maps the underlying representation into MRA images. The encoder comprises several 3D convolutional layers (kernel size = 3 × 3 × 3, padding = 1 × 1 × 1, stride = 2 × 2 × 2), instance normalisation and activation LeakyReLU layers. The decoder follows a full pre-activation residual network architecture (i.e. BatchNorm-ReLu-convolution) (He et al., 2016). The model consists of the residual blocks, followed by nearest neighbour up-sampling layers. The residual block contains two convolutional layers (kernel size = 3 × 3 × 3, padding = 1 × 1 × 1, stride = 2 × 2 × 2), and a learned residue of input is added to the output.

To ensure shared feature maps and relevant fine details are effectively propagated through the decoder, we adapted conditional batch normalisation (CBN), which has been exploited in several previous studies (De Vries et al., 2017; Chen et al., 2019; Park et al., 2019; Xia et al., 2020). CBN suggests a new conditioning mechanism to incorporate auxiliary information (such as labels, embedding, masks or input latent vectors) into the image generation through batch normalisation, to guide image generation process with respect to certain constraints. In this study, we used available structural MRIs concatenated as multi-channel input (i.e., a concatenation of T1, T2 and/or PD-weighted MRIs) to the CBN module, to consistently encode and emphasise shared spatial features and fine structural details, which contributes to enhanced performance in MRA generation regarding high quality and anatomical plausibility. We first normalise the extracted features to zero mean and unit standard deviation in each CBN layer. Next, the normalised features are modulated/de-normalised using the affine transformation whose scale and shift parameters are learned from multi-channel input. Specifically, in the batch normalisation, input feature batch $n_{b,w,h,c} \in \mathbb{R}^{B\times W\times H\times C}$ ($b \in B, w \in W, h \in H$, and $c \in C$ denote the batch size, width, height, and channel of the feature map, respectively, and $N = B \times W \times H$) is normalised in a channel-wise manner:

$$n'_{b,w,h,c,} = \gamma_{w,h,c}\left(\mathbf{x}_m\right) \times \frac{n_{b,w,h,c} - \mu_c}{\sigma_c + \epsilon} + \beta_{w,h,c}\left(\mathbf{x}_m\right), \tag{3}$$

with

$$\mu_c = \frac{1}{N}\sum_{b,w,h} n_{b,w,h,c}, \quad \sigma_c^2 = \frac{1}{N}\sum_{b,w,h}\left(n_{b,w,h,c} - \mu_c\right)^2, \tag{4}$$

where $\mathbf{x}_m$ denotes multi-channel input stack, $\gamma(\cdot)$ and $\beta(\cdot)$ represent the spatial dimension-dependent functions and ϵ is a number to avoid

division by zero. To leverage contextual information contained across slices, 3D convolutional kernels (kernel size = 3 × 3 × 3, padding = 1 × 1 × 1 and stride = 1 × 1 × 1) were suggested to learn γ and β. The modulation parameters of the CBN were learned simultaneously through the GAN training. The structure of the proposed CBN is illustrated in Fig. 1.

2.3. Discriminator

Rather than using a deeper network that would increase the network capacity and cause overfitting, we employed a multi-scale discriminator (Durugkar et al., 2017; Nguyen et al., 2017; Wang et al., 2018) operating at two-scale pyramid structure. The input images to the first discriminator are down-sampled by a factor of two for the second discriminator. The network is treated in a multi-task manner. The discriminator with the larger receptive field can guide the generator to synthesise globally coherent images, and the smaller scale one can encourage the generator to capture local finer details. The discriminator is built using five blocks with a 3D convolution (kernel size = 4 × 4 × 4, padding = 2 × 2 × 2 and stride = 2 × 2 × 2), instance normalisation and LeakyReLU with slope 0.2 applied in each block in the discriminator.

2.4. Loss functions

By adopting the least-squares loss function (Mao et al., 2017) and the multi-scale discriminator, our GAN training loss can be formulated as:

$$\mathcal{L}_{\text{GAN}}\left(G, D_k\right) = -\mathbb{E}_{\mathbf{x}_m, y}\left[\left(D_k\left(\mathbf{x}_m, y\right) - 1\right)^2\right] - \mathbb{E}_{\mathbf{x}_m}\left[D_k\left(\mathbf{x}_m, G\left(\mathbf{x}_m\right)\right)^2\right], \quad (5)$$

where $k \in [0, 1]$ denotes the discriminator operating at two different scales.

As MRA specifically targets vasculature imaging, accurately and faithfully synthesising vascular structures is crucial. Inspired by the fact that the MIP is commonly used to process 3D MRA images to highlight vascular trees over background tissues, we removed the pixel-wise L1 loss that struggles with capturing high-frequency details and designed new loss terms for emphasising vascular properties by minimising the statistical difference between the intermediate feature representations of the target images and the synthesised outputs in both the 3D volume and 2D projection domains, as shown in Fig. 2. First, a feature matching loss (Salimans et al., 2016; Wang et al., 2018) based on the discriminator was used to match the feature consistency between real volumetric MRA data and synthesised image data, in multiple intermediate layers of D_k:

$$\mathcal{L}_{\text{FM}_{3D}}\left(G, D_k\right) = \mathbb{E}_{\mathbf{x}_m, y}\sum_{i=1}^{T}\frac{1}{N_i}\left[\left\|D_k^i\left(\mathbf{x}_m, y\right) - D_k^i\left(\mathbf{x}_m, G\left(\mathbf{x}_m\right)\right)\right\|_1\right], \quad (6)$$

where i means the ith layer features in D, N_i is the number of features in each layer, T is the total number of layers. Second, we introduced a perceptual loss (Johnson et al., 2016) and measured perceptual differences between real and generated images in both 3D volume domain and 2D projection domain through high-level features extracted from a pre-trained VGG-16 network V:

$$\mathcal{L}_{\text{PER}_{3D}}\left(G, V\right) = \mathbb{E}_{\mathbf{x}_m, y}\sum_{j=1}^{F}\frac{1}{M_j}\left[\left\|V^j\left(y\right) - V^j\left(G\left(\mathbf{x}_m\right)\right)\right\|_1\right], \quad (7)$$

$$\mathcal{L}_{\text{PER}_{2D\_ax}}\left(G, V\right) = \mathbb{E}_{\mathbf{x}_m, y}\sum_{j=1}^{F}\frac{1}{M_j}\left[\left\|V^j\left(p_{\text{ax}}\left(y\right)\right) - V^j\left(p_{\text{ax}}\left(G\left(\mathbf{x}_m\right)\right)\right)\right\|_1\right], \quad (8)$$

$$\mathcal{L}_{\text{PER}_{2D\_co}}\left(G, V\right) = \mathbb{E}_{\mathbf{x}_m, y}\sum_{j=1}^{F}\frac{1}{M_j}\left[\left\|V^j\left(p_{\text{co}}\left(y\right)\right) - V^j\left(p_{\text{co}}\left(G\left(\mathbf{x}_m\right)\right)\right)\right\|_1\right], \quad (9)$$

$$\mathcal{L}_{\text{PER}_{2D\_sa}}\left(G, V\right) = \mathbb{E}_{\mathbf{x}_m, y}\sum_{j=1}^{F}\frac{1}{M_j}\left[\left\|V^j\left(p_{\text{sa}}\left(y\right)\right) - V^j\left(p_{\text{sa}}\left(G\left(\mathbf{x}_m\right)\right)\right)\right\|_1\right], \quad (10)$$

where j indicates the jth layer features, M_j and F represent the number of features in layer j and the total number of layers in V, respectively. The MIP algorithm projects the largest voxel values of 3D MRA data along the axial, coronal and sagittal direction path on a 2D projection image $p_{\text{ax}}(\cdot)$, $p_{\text{co}}(\cdot)$ and $p_{\text{sa}}(\cdot)$, respectively. These 2D projection views of inherently 3D structures can reveal important vascular structures due to varying projection angles. The feature-based loss terms presented in Eq. (6)–(10) aim to encourage the generator to synthesise HR anatomically plausible 3D MRA images resembling the appearance of target images and preserve the topology and continuity of vascular structures simultaneously. The final composite objective combines the adversarial loss, the feature matching loss, and the perceptual loss as:

$$\begin{aligned}\mathcal{L}_{\text{Final}} = \min_G\Bigg(&\left(\max_{D_1, D_2}\sum_{k=1,2}\mathcal{L}_{\text{GAN}}\left(G, D_k\right)\right)\\ &+\lambda_1\sum_{k=1,2}\mathcal{L}_{\text{FM}_{3D}}\left(G, D_k\right) + \lambda_2\mathcal{L}_{\text{PER}_{3D}}\left(G, V\right)\\ &+\lambda_3\mathcal{L}_{\text{PER}_{2D\_ax}}\left(G, V\right) + \lambda_4\mathcal{L}_{\text{PER}_{2D\_co}}\left(G, V\right) + \lambda_5\mathcal{L}_{\text{PER}_{2D\_sa}}\left(G, V\right)\Bigg), \quad (11)\end{aligned}$$

where λ_1, λ_2, λ_3, λ_4, and λ_5 control the weighting of the feature matching loss and perceptual loss, to the adversarial loss. The conditioned G and D models are optimised by $\mathcal{L}_{\text{Final}}$ to infer an MRA image y from available source images $\mathbf{x}_m$, such as T1, T2 and PD-weighted MR images.

3. Experimental setup

3.1. Datasets

We evaluated the proposed method using three separate neuroimaging datasets: the IXI dataset,[2] the MIDAS dataset[3] and the SIMONS dataset.[4]

In the IXI dataset, images were acquired on a 1.5T and a 3T scanner with the following parameters. T1-weighted images: TE = 4.6 ms, TR = 9.6/9.8 ms, flip angle = 8°. T2-weighted images: TE = 100 ms, TR = 5725.8/8178.3 ms, flip angle = 90°. PD-weighted images: TE = 8 ms, TR = 8178.3 ms, flip angle = 90°. MRA images: TE = 5.7/6.9 ms, TR = 16.7/20 ms, flip angle = 16°/25°. T1, T2, PD-weighted contrast sequences have spatial resolution = 0.94 × 0.94 × 1.2 mm^3 with matrix size = 256 × 256 × 150, and the MRA sequences have resolution of 0.5 × 0.5 × 0.8 mm^3 with a matrix size of 512 × 512 × 100. Images from 486 subjects were used and divided into a split of 340/50/96 for training/validation/test sets, respectively. All structural MRIs were registered onto the corresponding MRA images based on mutual information performed via FSL (Jenkinson and Smith, 2001).

In the MIDAS dataset, images were acquired on a 3T unit under the protocols with the following parameters. T1-weighted images: TE = 4.38 ms, TR = 1700 ms, flip angle = 8°. T2-weighted images: TE = 80 ms, TR = 7730 ms, flip angle = 180°. T1 and T2 weighted images were reconstructed at 1.0 × 1.0 × 1.0 mm^3. The MRA images were acquired using Siemens ALLEGRA 3.0T MRI scanner with imaging parameters (TR = 35.0 ms, TE = 3.56 ms, flip angle = 22°), which are very different from the imaging protocol used acquired the IXI dataset. Thus, we fine-tuned and evaluated the proposed generative model pre-trained on IXI data. The aim was to demonstrate that our approach, once pre-trained, can effectively use the small data regime to retrain and fine-tune the model to drive MRA synthesis. In total, 53 subjects were chosen for training, and 25 subjects for testing the model. Both IXI and MIDAS data used in this study are non-contrast enhanced TOF MRAs.

The SIMON dataset is a multi-centre study involving a single volunteer for multiple observations across networks. MRIs were performed with the three leading MRI vendors (Siemens, Philips, GE). The core

[2] http://brain-development.org/ixi-dataset/
[3] http://insight-journal.org/midas/community/view/21/
[4] http://fcon_1000.projects.nitrc.org/indi/retro/SIMON.html

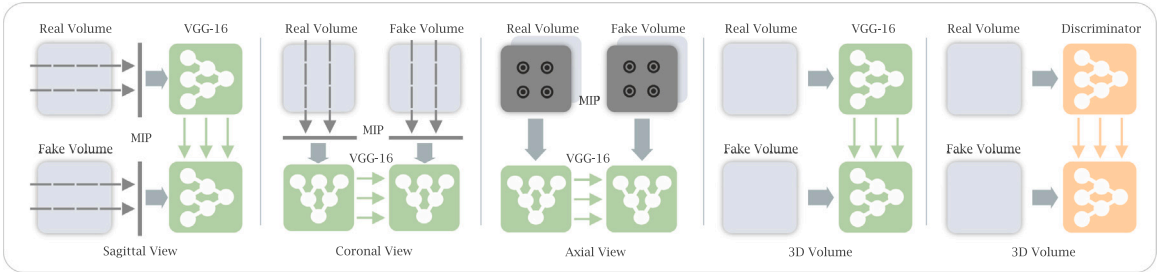


Fig. 2. To faithfully synthesise vessel structures and preserve the topology and continuity, several feature consistencies with feature matching loss and perceptual loss were designed in 3D volume and 2D projection domain from different perspective views (sagittal, coronal and axial) for emphasising reconstruction of vasculature properties.

protocol includes: an isotropic T1 weighted scan with voxel size 1.0 × 1.0 × 1.0 mm^3 and with an acceleration factor of 2, and an interleaved proton density/T2-weighted scan with resolution 0.9 × 0.9 × 3.0 mm^3, fat saturation, and an acceleration factor of 2. TR, TE, and other parameters were all chosen to obtain images of similar quality in terms of contrast and resolution. The detailed parameters are available on the CDIP website (www.cdip-pcid.ca).

3.2. Implementation details

The Adam optimiser with an initial learning rate of 1×10^{-4} was used for optimising the loss function for both the generator and discriminator. The decay rates of the gradient estimates' first and the second momentum were set to 0 and 0.9, respectively. The relative weighting factors of the feature matching loss and the perceptual loss to the GAN loss in Eq. (11) were initialised as $\lambda_1 = 300$, $\lambda_2 = 10$, $\lambda_3 = 10$, $\lambda_4 = 10$, and $\lambda_5 = 10$ and adjusted to $\lambda_1 = 30$ and the rest unchanged after 10 epochs. All the hyperparameters were selected on the validation set through a grid search for our method, Vox2vox, HiNet, CollaGAN and Pix2pix. For MRA-Net and sGAN, we directly used the hyperparameters reported in the original papers as they addressed the same problem of synthesising TOF MRA images from non-angiographic MRIs in these studies. For the MIDAS dataset, the model was initialised with the weights of the previously trained network on the IXI data and fine-tuned for another 20 epochs. The aim was to demonstrate that our approach, once pre-trained, can effectively use the small data regime to retrain and fine-tune the model to drive MRA synthesis. For the IXI and MIDAS datasets, the source images (i.e., T1, T2 and PD-weighted MRIs) and the target HR MRA images were resized to 512 × 512 × 100 and then were divided into small patches with a size of 256 × 256 × 32.

3.3. Competing methods

To demonstrate the effectiveness and superiority of the proposed method, we compared it with six state-of-the-art methods that were either specifically designed to synthesise MRA images (i.e., sGAN Olut et al., 2018 and MRA-Net Fujita et al., 2020) or are closely related to multi-contrast MR image synthesis (i.e., Pix2pix Isola et al., 2017, Vox2vox Cirillo et al., 2020, HiNet Zhou et al., 2020 and CollaGAN Lee et al., 2020). To emphasise vesselness properties, sGAN Olut et al., 2018 extends the standard GAN architecture and introduces an additional loss term that compares steerable filter responses of the target and generated images through a Huber loss function. MRA-Net (Fujita et al., 2020) combines the output of a single convolution and a U-net model, and constructs a mapping function that converts five raw 3D-QALAS images to their corresponding MRA images. Both HiNet (Zhou et al., 2020) and CollaGAN (Lee et al., 2020) are recent adversarial learning-based fusion networks dedicated to multi-modal MR image synthesis, leveraging complementary information from multiple source contrasts. We also compared our method with the popular Pix2pix model (Isola et al., 2017) and Vox2vox model (Cirillo et al., 2020) as 2D and 3D baseline image translation networks, respectively. We conducted all the statistical analysis using Wilcoxon signed rank test, which is a nonparametric statistical significance test and that does not assume the data to be normally distributed. The computed p-value < 0.001 indicates statistical significance in performance between two competing methods.

3.4. Cerebral vessel segmentation

To segment vascular structures in 3D MRA images, we used a global thresholding method that is based on a sum of lognormal distributions for voxel intensities (Allenby et al., 2021). All voxel intensities within an MRA volumetric data were divided into 50 bins spanning the whole greyscale value range. The six parameters for the lognormal distribution were determined by minimising normalised error to the log voxel intensity of histogram bin 3–48 with a certain margin (to avoid overexposed/underexposed errors). A threshold value T to segment vascular trees can be computed for each subject as:

$$T = x\left[\frac{\partial^3 f_{1+2}}{\partial x^3} = 0\right] + \left(\text{mode}\left(f_2\right) - x\left[\frac{\partial^3 f_{1+2}}{\partial x^3} = 0\right]\right) \times \left(\frac{f_2\left(\text{mode}\left(f_2\right)\right)}{f_{1+2}\left(\text{mode}\left(f_1\right)\right)}\right)^{k_T}, \tag{12}$$

where f_1, f_2 denote the lognormal distribution of background and intravascular tissue respectively, and $x\left[\frac{\partial^3 f_{1+2}}{\partial x^3} = 0\right]$ indicates a critical point of inflexion change between the modes of f_1 and f_2, and $k_T = 1.73$ denotes an empirically pre-defined weight parameter. Once the global threshold value T was determined, 3D vasculature can be segmented from background cranial tissue.

4. Results

4.1. Synthesis quality evaluation

We first evaluated the proposed model using all available source images (i.e., T1, T2 and PD-weighted MRIs) from the IXI dataset as inputs to the generative model. Visual comparison of the generated 2D MRA axial slices to the ground-truth MRA slices from three representative subjects is shown in Fig. 3a, alongside their respective error maps in the rightmost column. For better visualisation, all images were skull-stripped using the Brain Extraction Tool (BET) (Smith, 2002) from the FMRIB Software Library (FSL) (Smith et al., 2004). Fig. 3a shows that MRA images synthesised using the proposed approach are visually similar to the actual (ground-truth) MRA, reflected by the in-plane axial slice error maps, indicating that both fine-grained extravascular and intravascular tissues are captured in the synthesised images. Fig. 3b shows results obtained for the same subjects in the two other orthogonal (sagittal and coronal) views, demonstrating multi-planar consistency of the generated MRA images. Results from the sagittal and coronal views better highlight the ability of the proposed method to preserve 3D vascular structure and continuity. These advantages can

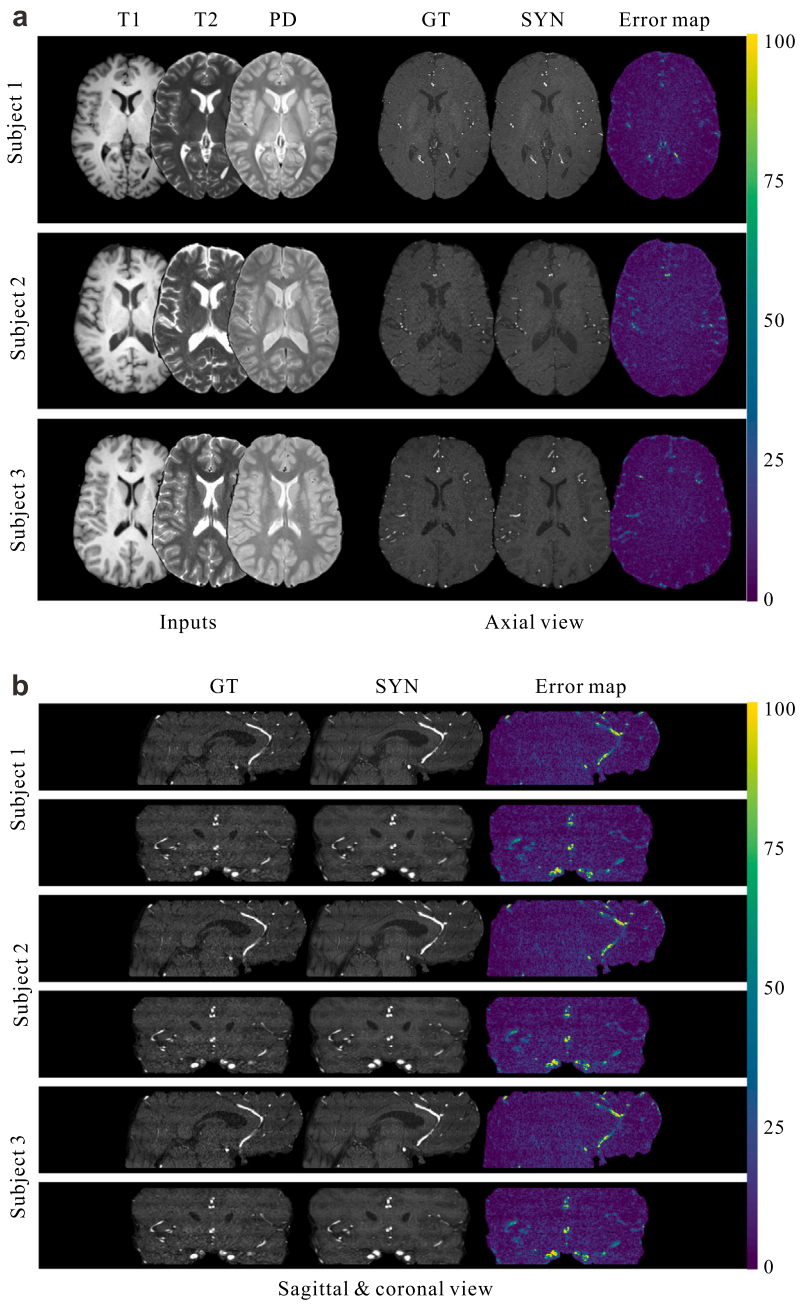


Fig. 3. Qualitative evaluation of the proposed method. (a) Visual comparison of the generated 2D MRA axial slices to the ground-truth MRA slices from three representative subjects. (b) Results of the same subjects in the sagittal and coronal views. The comparison is also augmented by the error maps shown in the rightmost column.

be attributed to the use of a 3D generative model and feature matching across multiple orthogonal views in the proposed approach, designed to emphasise vascular features and preserve 3D vascular structures.

Maximum intensity projections (MIPs) of the authentic and synthesised MRA images for two subjects from the test set of the IXI database are shown in Fig. 4a. Transverse MIPs were computed across three projection thicknesses, namely, 32 mm, 48 mm, and 64 mm. Higher projection thickness indicates more sequential in-plane axial slices covered from the viewpoint to the projection plane, thus leading to more vasculature being visible in the projection. As can be seen, the proposed model can accurately reproduce the representation of cerebral vessels and preserve topology and continuity, regardless of projection thickness. Zoomed-in ROI regions of the images marked with yellow rectangles are also displayed, along with projected T1, T2 and PD weighted images to highlight the correlation and difference between the MRA images with inflow effect and input source sequences that are not designed to emphasise inflow effect. These results indicate that the proposed approach is capable of capturing small, low-contrast vessels in the periphery of the cerebral hemispheres, in addition to the more central major cerebral arteries. Fig. 4b highlights whole-brain vasculature in the original and synthesised MRA images (in two orthogonal views), via a 3D rendering of the vessels extracted from their respective images. The rendered vascular structures show good similarity between the original and synthesised MRA images, as reflected by the side-by-side comparison between the former and the latter.

Fig. 5 shows examples of acquired T1, T2, and PD-weighted structural MRIs and the corresponding generated MRA images, when the authentic MRA images were not acquired or missed in the IXI dataset. The high quality of synthesised MRA images is maintained with preserved vessel continuity. This demonstrates that the proposed method

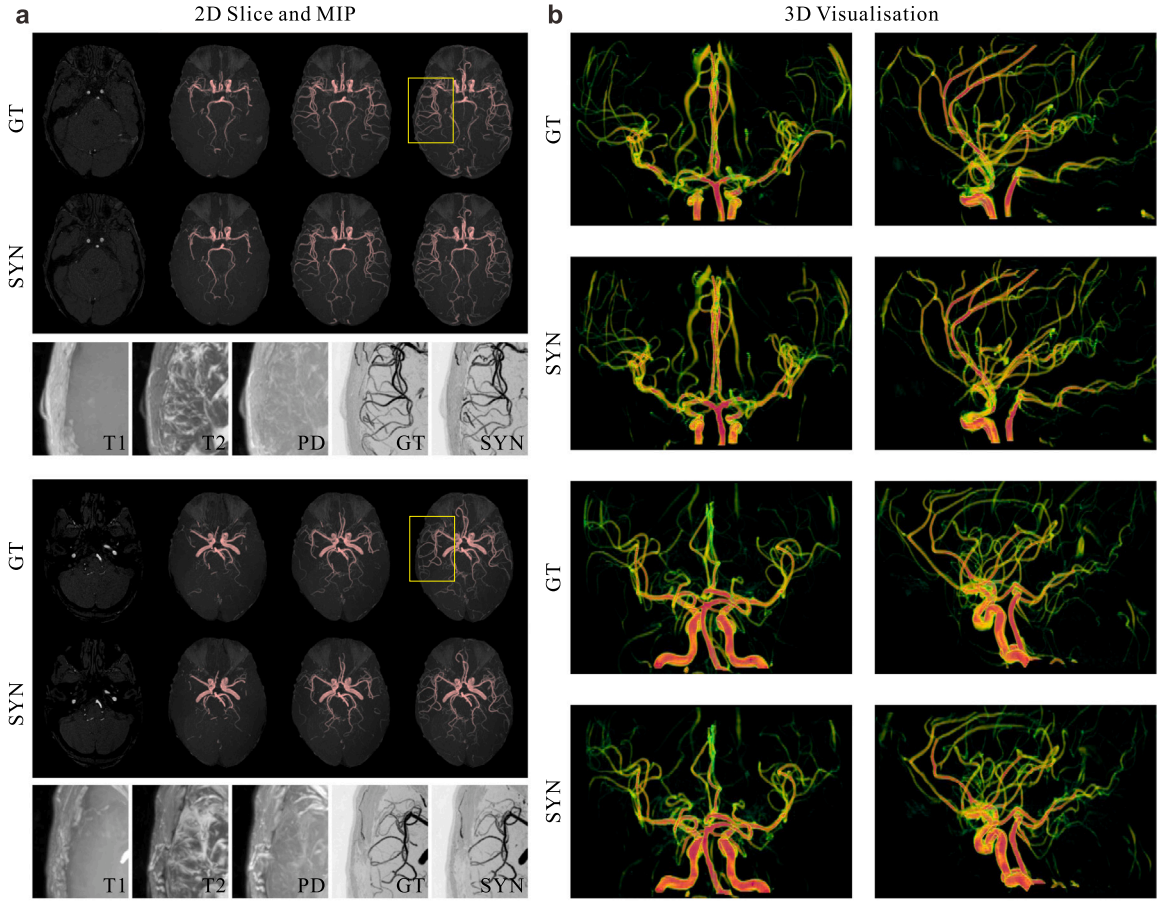


Fig. 4. (a) Two examples of axial slices and the corresponding transverse maximum intensity projections (MIPs) across three projection thicknesses. Zoomed-in ROI regions of the images marked as the yellow rectangles are displayed below, along with projected T1, T2 and PD weighted images to reveal the correlation and difference between the MRA images and input source sequences that are not designed to emphasise inflow-effect. (b) 3D visualisation of the synthesised MRA images from two different perspective views, yielding comparable visibility of intracranial vascular branches, compared with the actually acquired MRA images (i.e., GT).

can potentially serve as an image imputation to synthesise MRA images from existing MRI databases that MRA contrast is missed or corrupted, to prevent partially acquired subject data from being entirely discarded when analysing a given cohort.

Histograms and lognormal distributions of the voxel intensities of extravascular (f_1, yellow lines) and intravascular regions (f_2, green lines) for two representative cases are presented in Fig. 6a. The sum of the lognormal distributions for voxel intensities was used to compute a global threshold for each subject according to Eq. (12) and then to segment vessels from surrounding intracranial tissues. Segmentation results are shown in Fig. 6b, with the surface-to-surface distance computed between vascular surface meshes extracted from the synthesised and the reference images, over three ROIs. No apparent differences are found in the intensity distributions between the synthesised and ground-truth images. Most surface distance errors are smaller than the in-plane spacing of MRA images (i.e., 0.5 mm/pixel). We also computed two voxel-wise vascular morphological indices, namely, voxel distance from vessel centrelines and voxel distance from the nearest carotid vessel (i.e., the nearest centrepoint at the bottom of the volume image), presented in Fig. 6c. As can be seen, there is a good agreement regarding these morphological properties between vessels derived from the real and generated MRA images.

4.2. Importance of different MR contrasts for MRA synthesis

To evaluate the impact of using different combinations of structural MR images as inputs to drive MRA synthesis, and correspondingly, on the final quality of MRA images synthesised using our approach, we conducted a series of data ablation experiments using various combinations of T1-, T2- and PD-weighted MR images as inputs to train our model. We trained different network instances using several possible combinations of T1, T2, and PD inputs and generated subject-specific MRA images. In total, there were five different models: three T1, T2, PD-weighted uni-contrast models, one bi-contrast model with T1 and T2 as input, and one tri-contrast model with a concatenation of all available source contrasts. The Wilcoxon T-test was used to evaluate statistical significance when comparing results from all investigated models. Visual comparison of the MIPs computed using the real and synthesised MRA images from various contrast combinations is shown in Fig. 7a for two samples from the IXI dataset. The quality of images synthesised and their similarity to their respective ground-truth (real) images was evaluated for each model and quantified in 3D across the test subjects from the IXI database using two metrics: structural similarity index measurement (SSIM) and peak signal-to-noise ratio (PSNR), shown in Fig. 7b. We also evaluated the quality of 2D MIPs derived from the synthesised MRA images and their similarity to their respective ground-truth MIPs, denoting as 2D PSNR and 2D SSIM in the figures. Images with higher SSIM and PSNR values indicate higher image quality. Several morphometric properties of vasculature were characterised: voxel distance from centreline, distance to nearest vessel surface, voxel distance from the nearest carotid vessel. The three morphometries yield a consistent relationship. A voxel-specific polynomial regression can be fitted to measure the consistency between three voxel-wise morphological indices, i.e., centreline distance, base

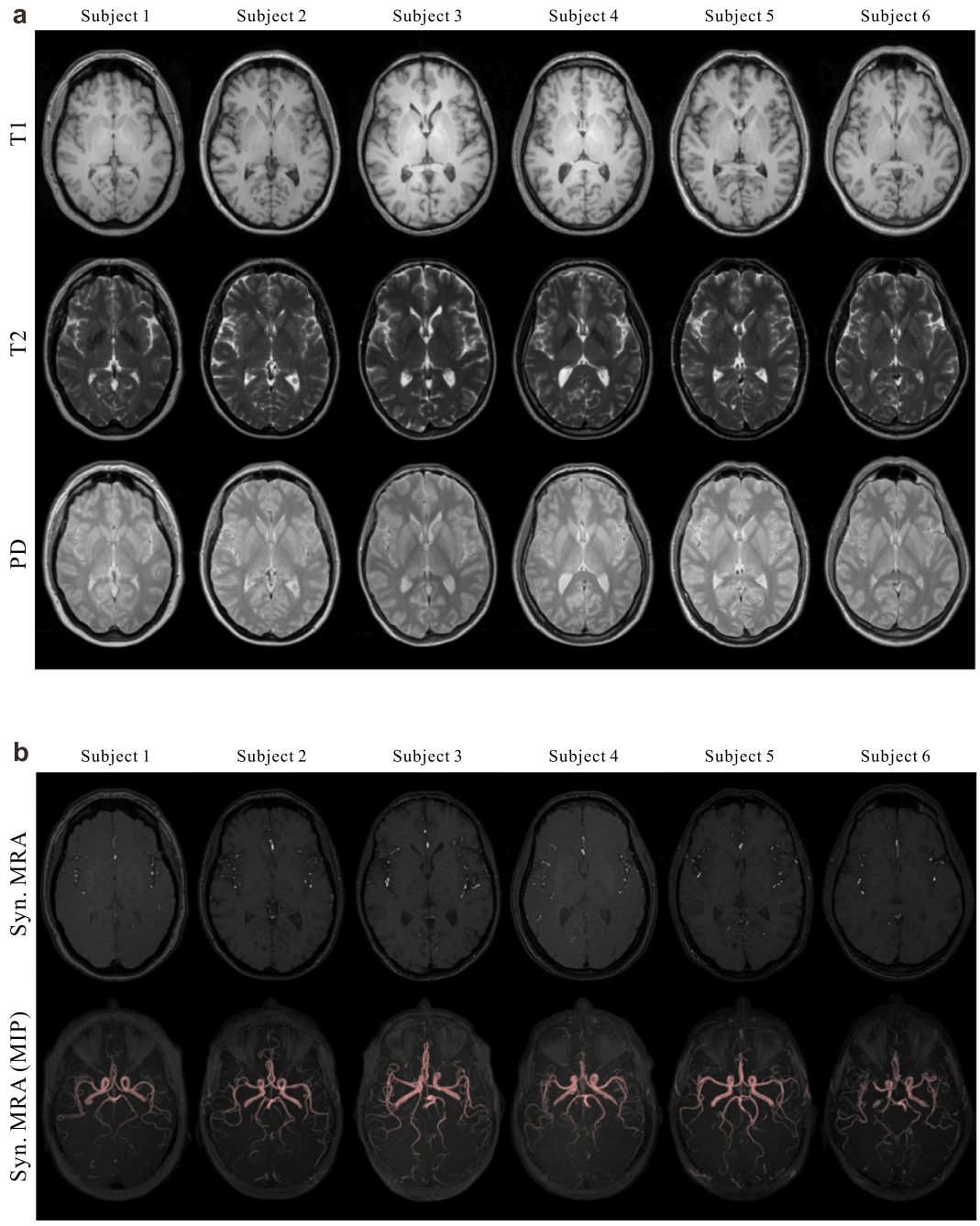


Fig. 5. Examples of the MRA images (synthesised) that were not acquired in the IXI dataset. (a) Acquired T1, T2, and PD-weighted structural MRIs are usually included in the imaging protocol. (b) Corresponding generated MRA images, where the authentic MRA images were not acquired in this dataset. This demonstrates that the proposed method can serve as an image imputation for missing/corrupted MRA images to prevent partially acquired subject data from being entirely discarded when analysing any given cohort.

distance and edge distance. We computed median regressions of the test subjects and presented the results in Fig. 7c. Two additional morphological indices – maximum vessel length and volume – were also calculated from the segmentation masks and plotted in Fig. 7d. We fitted a 2D Gaussian distribution to the GT and derived morphological indices and quantified the distance of the resulting probability distributions using the Bhattacharyya distance.

Overall, we found that the multi-contrast synthesis model trained using T1, T2 and PD sequences yields the best performance, both qualitatively and quantitatively. The SSIM and PSNR values of MRA images synthesised using the multi-contrast model are significantly higher than those obtained for images synthesised using the other models investigated ($p < 0.001$). This indicates that all three MR contrasts provide complementary and valuable information for MRA synthesis. Fig. 7c and d show that the three-contrast model yields the closest results regarding the distance of median voxel regression and morphological characteristics such as vessel volume and length. The computed Bhattacharyya distance of the resulting probability distributions between the reference and the derived vessel length is 0.063 for three-input model, which is smaller compared with 0.075 for PD model and 0.109 for T1+T2 model. The Bhattacharyya distances for predicted vessel volume are 0.066, 0.095 and 0.068 for three-input, T1+T2 and PD models, respectively. When using a combination of T1 and T2 contrasts as input, the results are superior to those using T1 or T2 information alone. This improvement is also statistically significant ($p < 0.001$). For three uni-contrast models, we found T1 contrast alone

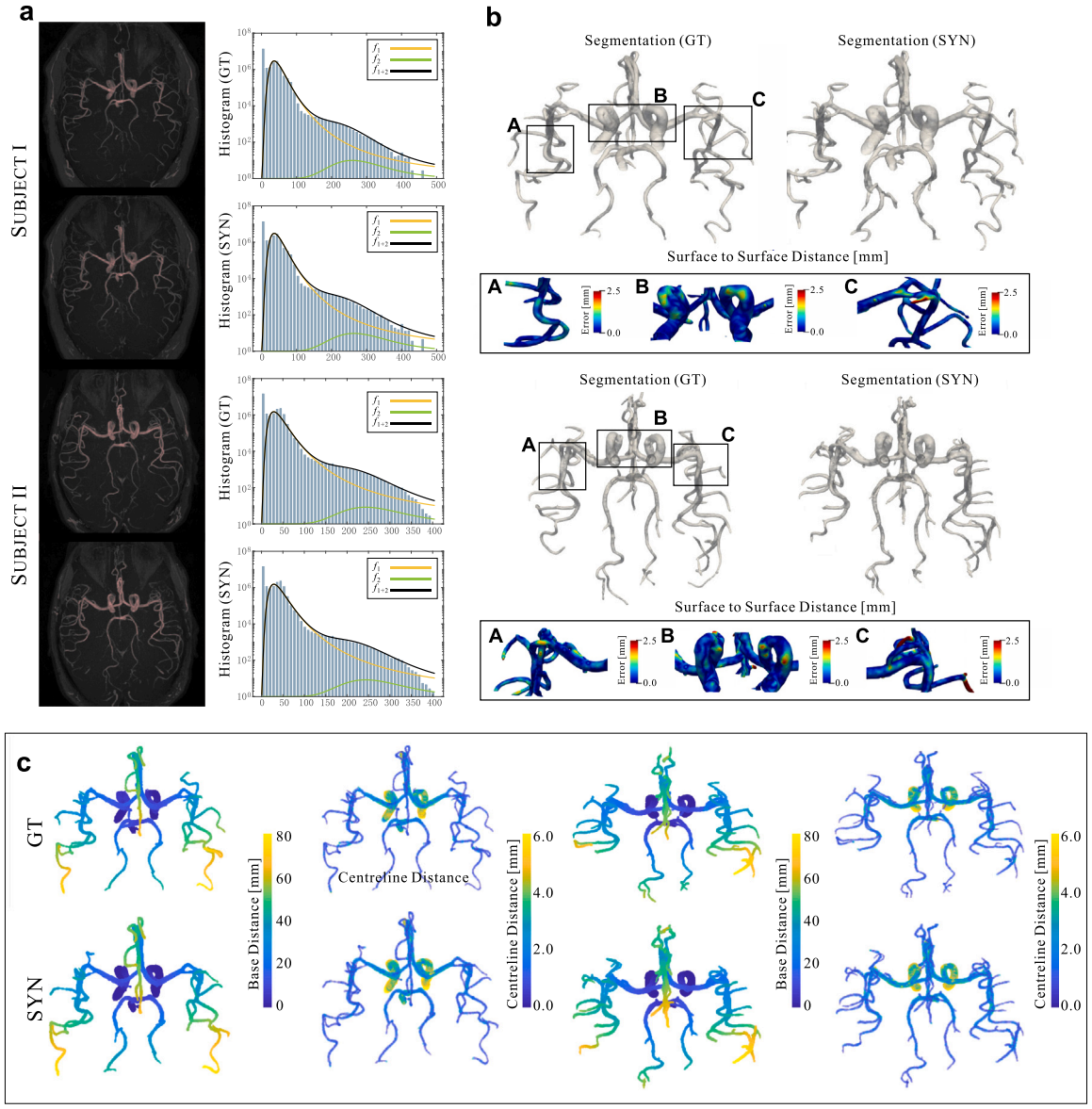


Fig. 6. Cerebral vessel segmentation results and derived morphometries. (a) Histograms and lognormal distributions of the voxel intensities of extravascular (f_1, yellow lines) and intravascular regions (f_2, green lines) for two representative cases. The sum of the lognormal distributions for voxel intensities was further used to compute a global threshold and then segment vessels from surrounding intracranial tissues. (b) Segmentation results with surface-to-surface distance computed between the synthesised images and the reference over three ROIs. Most surface distance errors are smaller than the in-plane spacing of MRA images (i.e., 0.5 mm/pixel). (c) Two voxel-specific morphometries of vasculature, i.e., voxel distance from centreline and voxel distance from the nearest carotid vessel.

yields the highest error and fails to preserve the continuity of vascular structures in the synthesised MRA images, as can be seen in Fig. 7a. This indicates that T1 contrast contains the least correlated and redundancy information to the MRA sequence. All other input combinations yielded statistically significant improvements in the quality of MRA images synthesised, relative to using only T1 images. We also observed that the PD contrast improves the visibility of small vessel branches in the peripheral regions than other single-input models.

4.3. Comparisons with state-of-the-art

We compared the performance of the proposed method against six state-of-the-art approaches for multi-contrast-driven MRA image synthesis. We trained the competing techniques with three-contrast inputs and synthesised subject-specific MRA images. A visual comparison of the synthesised slices and MIPs for an example test case from the IXI dataset can be seen in Fig. 8a, including Pix2pix, CollaGAN, HiNet, sGAN, MRA-Net, Vox2vox and the proposed method. Here, Pix2Pix, CollaGAN, HiNet and sGAN models are 2D, while MRA-Net, Vox2vox and the proposed method are 3D models. All six state-of-the-art methods are also quantitatively compared with our approach using the SSIM and PSNR metrics, summarised in Fig. 8b. Scatter plots of SSIM measurements for the proposed method against the state-of-the-art techniques are included in Fig. 9. Overall, although the 2D models synthesise good quality images visually in 2D axial slices, capturing fine-grained structures, they are less successful at preserving the continuity of vessels in 3D as can be seen from their corresponding MIPs. The 3D MRA-Net and Vox2vox outperform their 2D counterparts in this regard and are able to produce adequate continuity in 3D vessel structures in the main cerebral artery branches. However, a large proportion of the smaller, low-contrast vessels towards the periphery of the cerebral hemispheres (with increasing branching degree) are still missed by both models. In contrast, our approach can retain key visual cues and provide significantly better intracranial vasculature

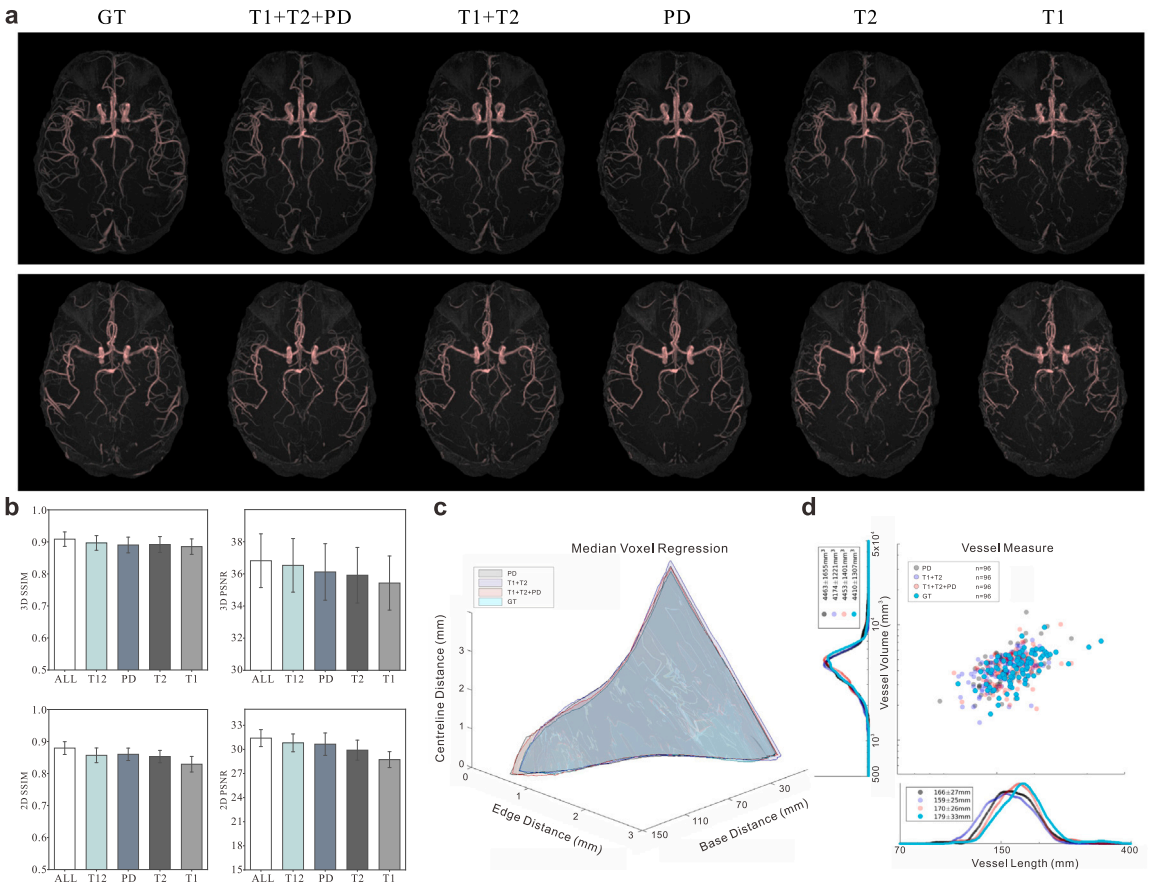


Fig. 7. Importance assessment on various contrast inputs to the MRA synthesis model. (a) Visual comparison of the MIPs computed from the real and synthesised MRA images from various contrast combinations. (b) The bar plots of structural similarity index measurement (SSIM) and peak signal-to-noise ratio (PSNR) were computed from 3D volumes and 2D MIPs from the test subjects. (c) Median voxel-specific polynomial regressions of test subjects fitted to measure the consistency between three morphometries, i.e., centreline distance, base distance and edge distance. (d) Comparison of two additional morphometric features, i.e., vessel maximum length and volume, were derived from the segmentation masks.

visualisations by emphasising vesselness properties through the dedicated architecture and loss functions, and yields the most comparable results in terms of visual quality to the ground-truth images. Besides qualitative improvements, the SSIM and PSNR values of the proposed method also produce statistically significant improvements over the state-of-the-art generative models, considering a significance level of $p < 0.001$ (as shown in Fig. 8b).

Additionally, we utilised an evaluation metric proposed in Sajjadi et al. (2018) to quantify synthesis performance across different generative models. The method defines a novel notion of precision and recall for generative models to compare a learned distribution Q to the actual data distribution P. The quantity precision intuitively measures the quality of samples generated from Q, while recall measures the proportion of P that is covered by the generator. We measured these quantities for the proposed model and the competing methods, and the resulting precision–recall distribution (PRD) curves are shown in Fig. 8c. These PRD curves relate intuitively to the traditional precision and recall metrics used for evaluating classification algorithms. We also plotted pairs of F_8 and $F_{1/8}$ scores that are estimated from each PRD curve, to understand the bias towards precision vs recall. As can be seen, the proposed model yields high probability under the actual data distribution and is interpreted as a higher value in precision and recall. Thus, it is less penalised by the F_8 and $F_{1/8}$ scores than its counterparts.

4.4. Ablation study

This section presents the ablation study, in which we systematically assessed the effectiveness and the contribution of each component in the proposed method. The comparison was made between four variants and the proposed method. These four variants correspond to progressively remove or replace one of the proposed components, namely, replacing multi-scale discriminator with a single discriminator (Variant 1), removing the feature matching term (Variant 2), replacing the perceptual loss in 3D and 2D domain with the L1 distance loss between real and synthetic images (Variant 3), and replacing the CBN with the standard batch normalisation (Variant 4) in turn. For fair comparison, we trained these variant networks using the same epochs as the proposed method. Results from each of these network configurations are then compared with the proposed model.

Fig. 10 illustrates a visual comparison of the images generated in this ablation study. The proposed method with all the components included produces images not only most visually comparable to that of the GT, but also perceptually appealing and anatomically plausible. Whilst using a single discriminator still retains the high-quality synthesised images, removing the feature-based loss terms presented in Eq. (6)–(10) results in a significant drop in performance, particularly for the perceptual loss (cf. Variant 3 in Fig. 10). Last, replacing the CBN with standard batch normalisation yields the worst synthesis results as the shared spatial features and fine structural details of multi-channel input cannot be effectively propagated through the generation pathway, which is crucial to generate high-resolution MRA images. Quantitative results in Table 1 confirm this observation and indicate that the single discriminator yields slightly inferior results than its multi-scale counterpart, while removing or replacing feature-based loss terms leads to a huge performance degradation in (e.g., an SSIM of

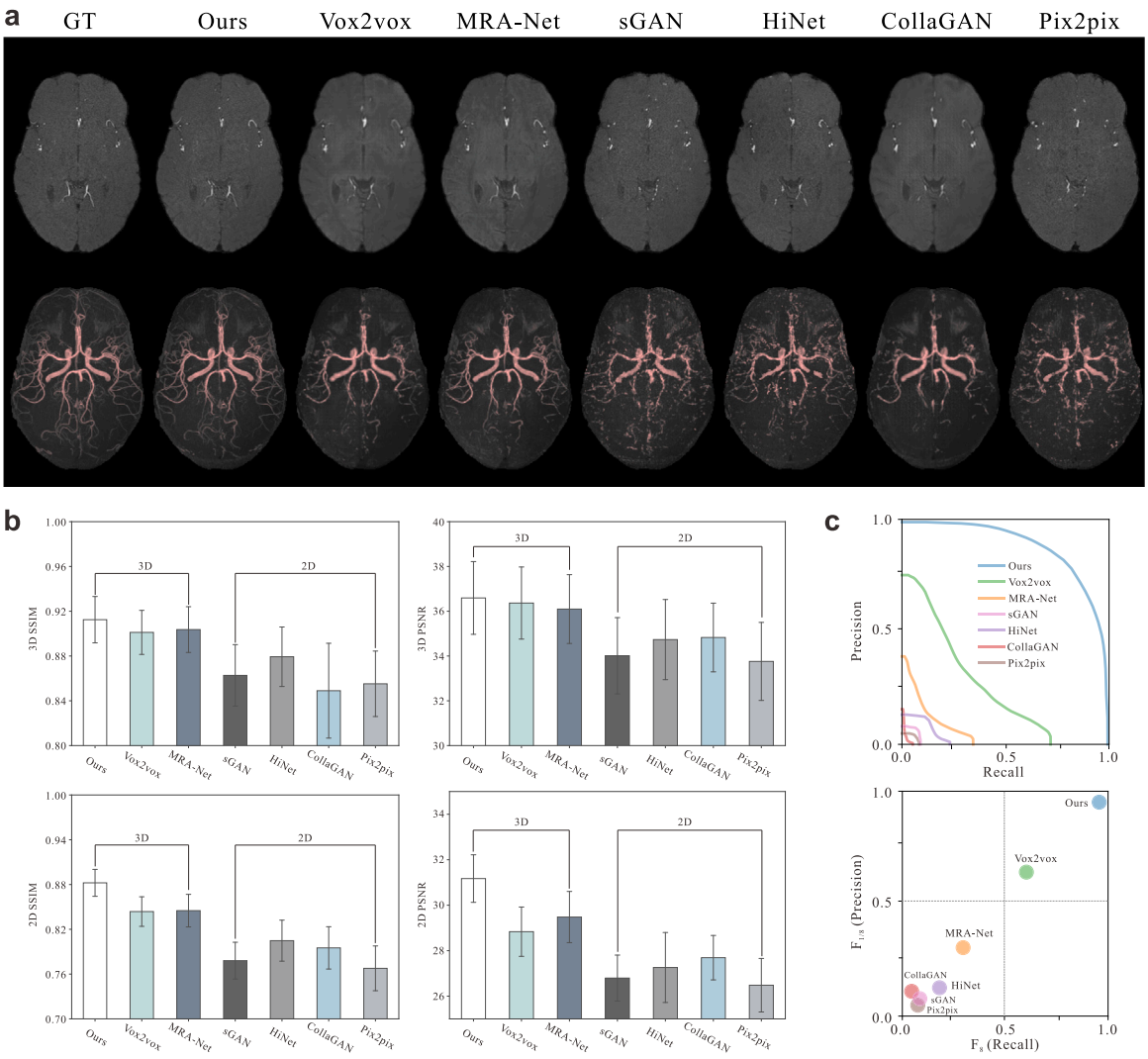


Fig. 8. Demonstrations of the proposed method against six state-of-the-art approaches for multi-contrast MRA synthesis. (a) Qualitative comparison of the synthesised slices and MIPs for a representative example, including Vox2vox (Cirillo et al., 2020), MRA-Net (Fujita et al., 2020), sGAN (Olut et al., 2018), HiNet (Zhou et al., 2020), CollaGAN (Lee et al., 2020) and Pix2pix (Isola et al., 2017) and the proposed approach. (b) Summary of the SSIM and PSNR computed on the test subjects for the competing methods. (c) A novel notion of precision and recall for generative models to compare a learned distribution to a true data distribution (Sajjadi et al., 2018). Here, the precision recall distributions (PRD) curves for the competing methods and pairs of F_8 and $F_{1/8}$ distilled from each PRD curve are plotted.

Table 1
Quantitative results of the ablation study for four variants that correspond to replacing multi-scale discriminator with a single discriminator (Variant 1), removing the feature matching term (Variant 2), replacing the perceptual loss in 3D and 2D domain with the L1 distance loss between real and synthetic images (Variant 3), and replacing the CBN with the standard batch normalisation (Variant 4) in turn.

Metrics	SSIM	PSNR
Proposed	**0.882 ± 0.019**	**31.17 ± 1.04**
Variant 1	0.870 ± 0.024	30.91 ± 1.23
Variant 2	0.864 ± 0.026	29.87 ± 1.20
Variant 3	0.841 ± 0.029	28.72 ± 1.24
Variant 4	0.820 ± 0.026	26.93 ± 1.18

0.870 ± 0.024 to 0.841 ± 0.029). We can see that replacing the CBN with standard batch normalisation further degrades the quality of the images synthesised, which is reflected by an SSIM decrease of 0.841 ± 0.029 to 0.820 ± 0.026. The progressively decreasing PSNR values also demonstrate the degradation of image quality.

4.5. Cross-database evaluation

We fine-tuned and evaluated the proposed generative model (pre-trained on the IXI data) on the MIDAS dataset. The primary aim of this experiment was to demonstrate that our approach, once pre-trained, can perform effectively in the small data regime through transfer learning. The model was initialised with the weights of the network pre-trained on the IXI data and fine-tuned using data from 53 subjects in the MIDAS dataset. To leverage complementary information of multiple source images, we used all available contrasts in the MIDAS dataset for fine-tuning our model. This included T1 and T2 images concatenated to form multi-channel inputs and the original MRA images available in MIDAS for each subject as the target to guide fine-tuning of the network. For comparison, the proposed method also were trained directly on MIDAS data as a reference/benchmark model. Fig. 11 shows a test sample from the MIDAS dataset, including MIPs of the original (top row), synthesised images with the transferred model (middle row), synthesised images with the directly trained model (bottom row) across four projection thicknesses (32 mm, 48 mm, 64 mm and 96 mm). As can be seen, the transferred model yields overall similar performance to that of the IXI data when using two-contrast input and is able to

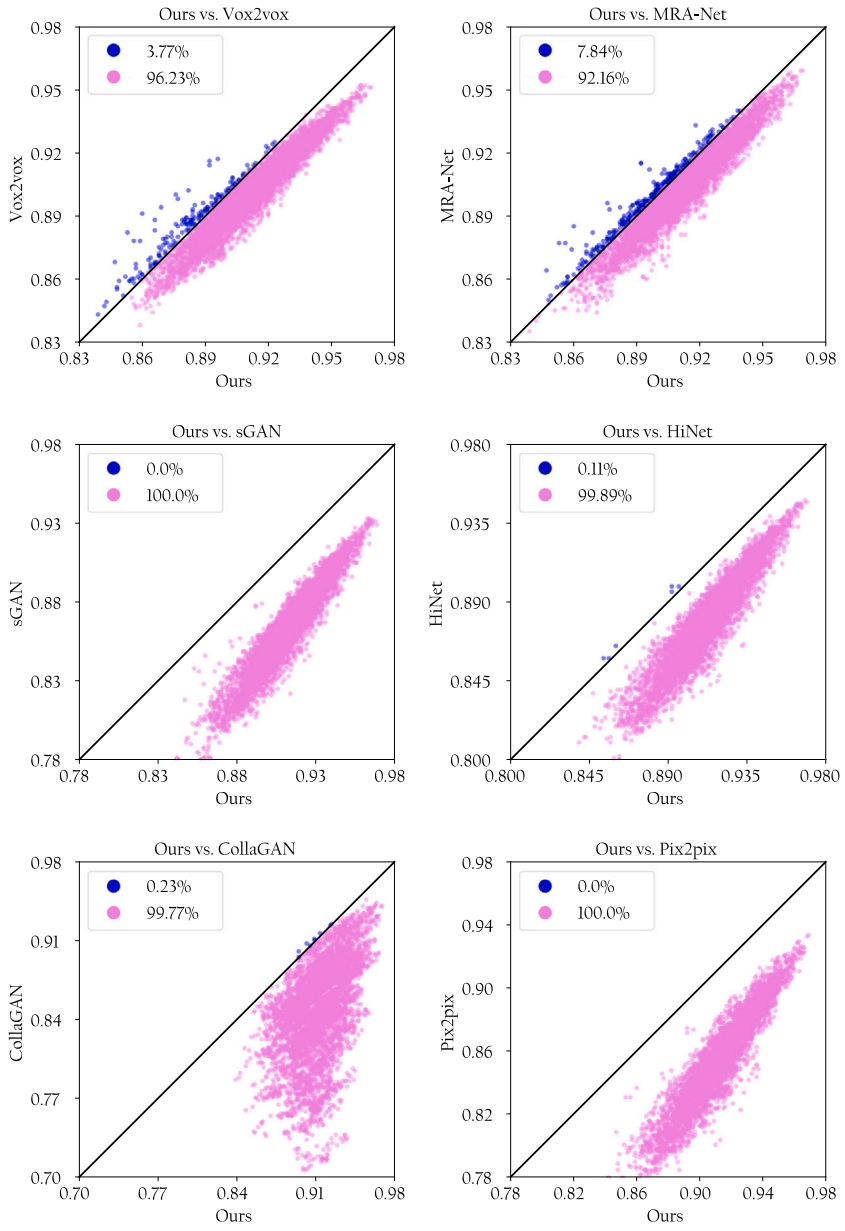


Fig. 9. Scatter plots of SSIM measurements for the proposed method against six state-of-the-art techniques for multi-contrast MRA synthesis. Each point represents a value computed in the axial slice in the test set. The proportion of test samples in which either method yields superior performance is noted in figure legends. The proposed model yields the best performance and shows the superior results over Vox2vox (Cirillo et al., 2020), MRA-Net (Fujita et al., 2020), sGAN (Olut et al., 2018), HiNet (Zhou et al., 2020), CollaGAN (Lee et al., 2020) and Pix2pix (Isola et al., 2017) for 96.23%, 92.16%, 100.0%, 99.89%, 99.77% and 100.0% of test samples, respectively.

accurately capture 3D vascular structures whilst preserving continuity, as demonstrated by the MIP images estimated across varying projection thicknesses. We observed a marginal drop in performance in the quality of MRA images synthesised for test samples in MIDAS, relative to IXI, particularly in the peripheral areas of the cerebral hemispheres, resulting in reduced detail being captured for small arterial branches in these areas. This may be attributed to the absence of PD images in the MIDAS dataset, which was used as an additional input MR contrast in the IXI dataset. This result is consistent with those obtained in the data ablation experiment conducted using the IXI dataset (discussed previously).

Fig. 12 presents a quantitative evaluation of the quality of synthesised MRA images for 25 test subjects from the MIDAS dataset. As previously, the SSIM and PSNR metrics are used to evaluate the quality of synthesised MRA images in 3D and their corresponding MIPs in 2D, relative to the original (real) images. These quantitative results are consistent with results obtained in the data ablation study when two inputs contrasts were used in the IXI dataset; see Fig. 7b. Across the datasets, the transferred model yields variability of merely 0.72% and 0.85% in mean values of SSIM and PSNR between the IXI and MIDAS datasets. This demonstrates the adaptation ability of the proposed approach, even in the presence of a small dataset. We also observed that the directly trained model shows a huge performance degradation compared to the transferred model, due to the limited training data in MIDAS dataset.

To demonstrate the ability of the proposed approach to generalise to completely unseen data, we also evaluated the pre-trained model on the SIMON dataset that involves a single subject scanned at multiple centres across the three leading MRI vendors (Siemens, Philips, GE), without fine-tuning and transfer learning. The common MR sequences such as T2 and PD were acquired but MRA acquisitions were not included in the scanning protocol in this dataset. Fig. 13a shows

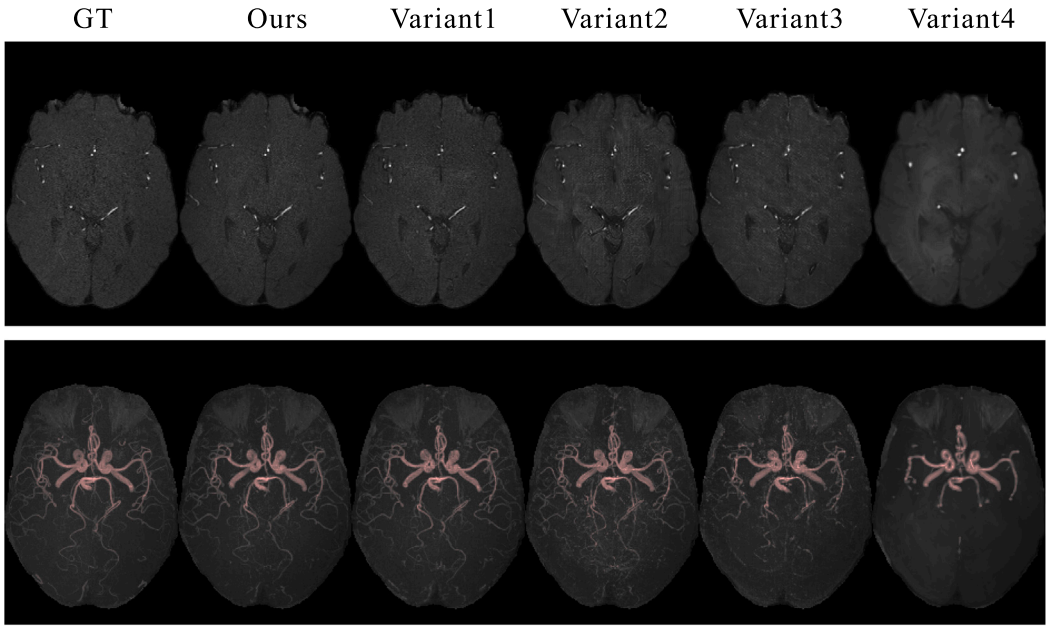


Fig. 10. An overview on results of the ablation study. Visual comparison of four variants that correspond to replacing multi-scale discriminator with a single discriminator (Variant 1), removing the feature matching term (Variant 2), replacing the perceptual loss in 3D and 2D domain with the L1 distance loss between real and synthetic images (Variant 3), and replacing the CBN with the standard batch normalisation (Variant 4) in turn.

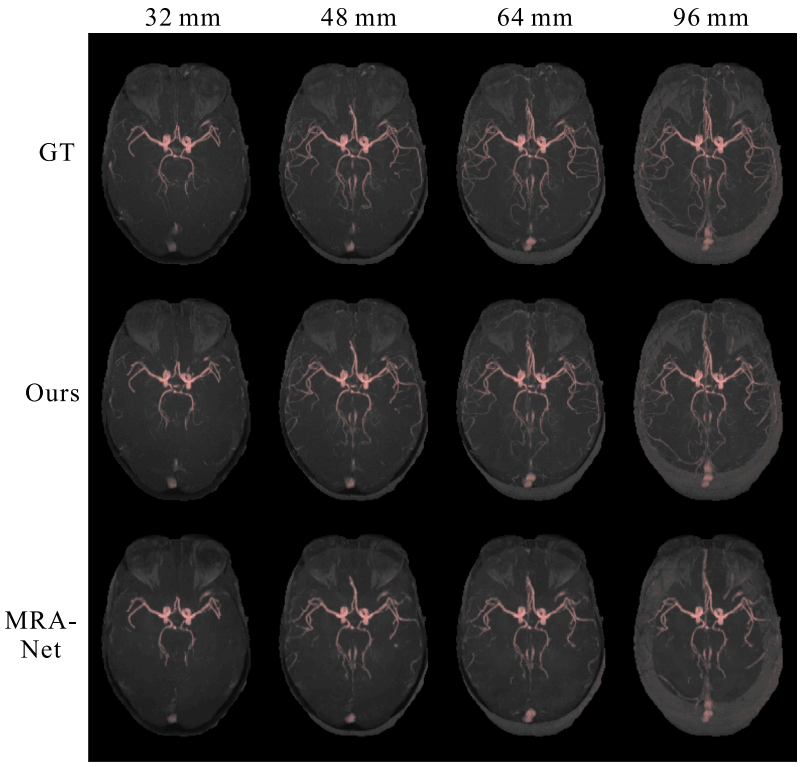


Fig. 11. Cross-database evaluation with the transferred model (i.e., pre-trained on the IXI data and then fine-tuned MIDAS data) and the directly trained model using 53 subjects from the MIDAS dataset. T1 and T2 images were concatenated as multi-channel input to generate the corresponding MRA images. (a) Visual inspection of examples of the corresponding MIPs across four projection thicknesses (32 mm, 48 mm, 64 mm and 96 mm).

examples of acquired T2/PD-weighted MRIs and synthesised MRA with the proposed method and Fig. 13b shows MIPs of the synthesised MRA images from the structural MRIs acquired from 12 sessions. The sites and scanner models associated with MRIs acquired in each session were summarised in Fig. 13c. Although no ground truth is available, we provide two quantitative measures for assessing the synthesised vascular anatomies using data from different imaging centres. The first measurement is the vessel volume. We first applied the segmentation method to extract vascular structures from 3D synthetic MRA images, followed by computing vascular volume in each case (i.e., by summing up the number of voxels in the segmentation result and multiplying by the voxel spacing). The second measurement is the Dice score

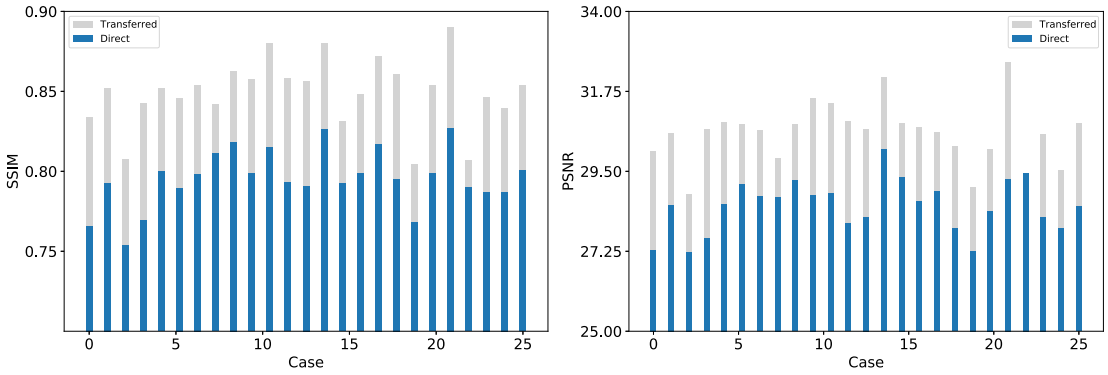


Fig. 12. Quantitative measurements of SSIM and PSNR computed from 2D projected MIPs across 25 test subjects for the transferred model and directly trained model.

with respect to the reference. The reference was obtained by applying the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm (Warfield et al., 2004) from the binary segmentations we obtained from the synthetic MRA images. Quantitative measures are shown in Fig. 13d. We found the proposed framework can still generate high-resolution, anatomically plausible MRA capturing whole-brain vasculature, even with the large slice thickness in the SIMON T2/PD-weighted images (3 mm). The results demonstrate that the proposed approach is pre-trained on the IXI data but can generalise to the SIMON data where images were acquired across multiple centres and scanners, whilst synthesising MRAs and vascular geometries that maintain vessel continuity in 3D. We highlighted the robustness of the trained model to inter-scanner variations, as demonstrated by the high synthesis quality on SIMON data. Note that the quantitative measures are affected by not only the algorithm's synthesis performance but also the downstream image processing. For example, the accuracy of the vessel volume calculation is subject to both synthesis quality and segmentation accuracy. Similarly, the measured Dice score is also subject to synthesis performance, segmentation accuracy, the reference estimation from STAPLE, alignment of each image, etc.

Discussion

MR image synthesis offers the potential to minimise scan time by acquiring only essential imaging data from which others images can be derived. Despite success on multi-contrast MR image synthesis, generating specific modalities remains challenging. For instance, MRA images highlight details of vascular anatomy using specialised imaging sequences underpinned by an endogenous or exogenous contrast mechanism. A reliable technique for synthesising MRA images from other available MR contrasts is valuable because there are few large databases with imaging modalities (such as MRA) that enable quantitative characterisation of whole-brain vasculature and is also motivated by the need to generate digital twin populations of cerebrovascular anatomy at scale for use in in-silico studies and/or in-silico trials.

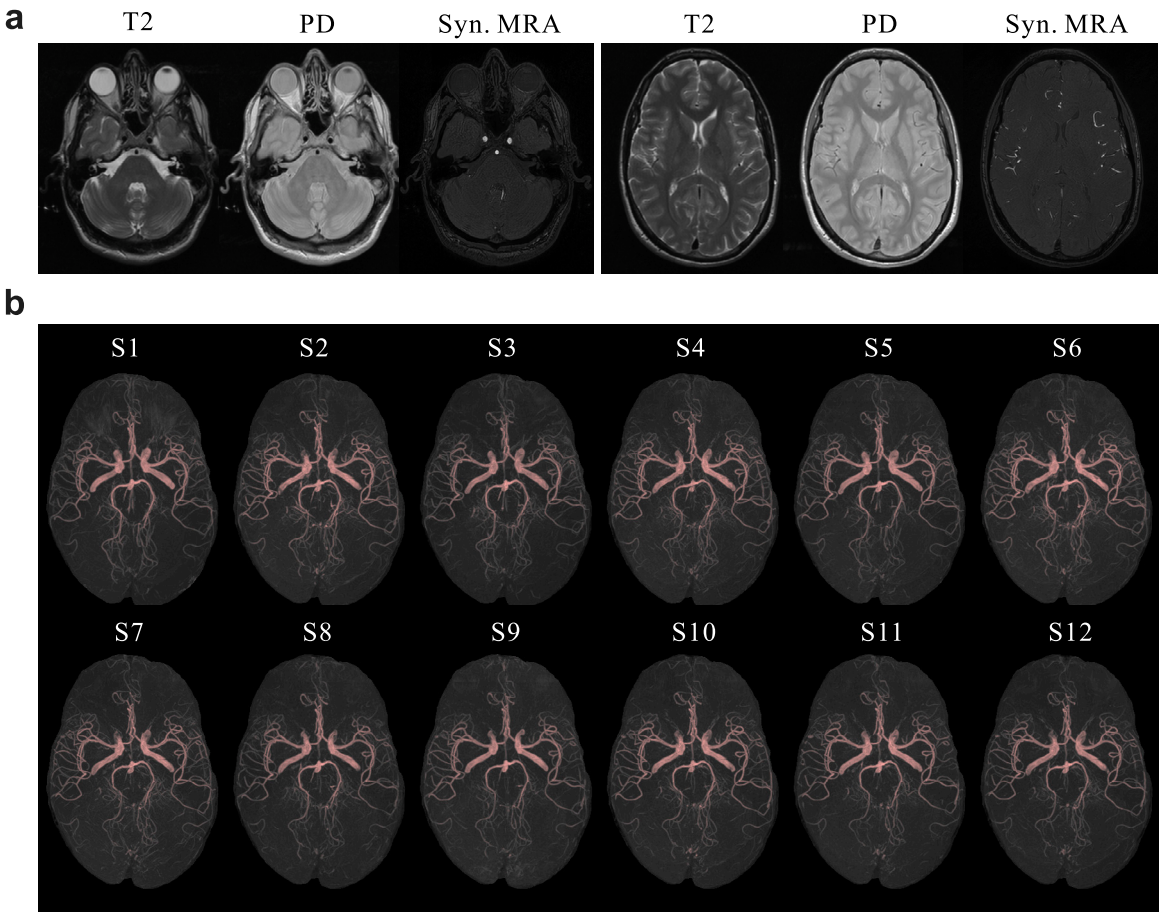
This study demonstrated that multi-contrast structural MR images of the brain (such as T1, T2, and PD) could be used to synthesise high-resolution, anatomically plausible MRA images that capture whole-brain vasculature using the proposed GAN-based image translation framework. Our approach leverages complementary information available from multi-contrast structural MR images to learn a mapping to their corresponding MRA image in a subject-specific manner. A shared latent embedding for multi-contrast MR images from the source domains (T1, T2 and PD) are learned through an encoder network and used by a dedicated generator and discriminator to synthesise MRA images. The multi-scale discriminator utilises deep feature consistency loss terms estimated slice-wise along orthogonal views (axial, sagittal and coronal) to emphasise vascular structures and synthesise MRA images whilst preserving the continuity of vessels in 3D.

A data ablation study assessed the relative importance of the structural MR contrasts used as inputs/predictors to drive MRA synthesis and revealed that incorporating T1, T2 and PD-weighted images provided the best results in terms of the quality of MRA images synthesised. T2 and PD-weighted images were better predictors of MRA images than T1. Additionally, we found that PD-weighted images contributed to better visibility of small vessel branches towards the peripheral regions than the other two MR contrasts. For the adaptability, we demonstrated that the proposed approach, once pre-trained, can effectively use the small data regime to retrain and fine-tune the model to drive MRA synthesis in the MIDAS dataset, when the target domain is significantly different from the training domain. We also demonstrated the ability of the proposed approach to generalise to completely unseen data (SIMON dataset) whilst synthesising MRAs and vascular geometries that maintain vessel continuity in 3D. We highlighted the robustness of the trained model to inter-scanner variations, as demonstrated by the high synthesis quality on SIMON data.

In this work, we selected a global thresholding method used for segmenting vasculature for subsequent analyses because the variances of each structural class (background, non-vessel and vessel) are relatively small in MRA, which makes the segmentation of the vasculature relatively easy. We applied this traditional, unsupervised segmentation method to both authentic MRA images and synthetic images for comparison. Therefore, we expect that the segmentation results and the derived vascular morphological indices in the evaluation would involve both synthesis and segmentation errors.

It is also worth to noting that the input structural MRIs typically have lower spatial resolution (e.g., $0.9 \times 0.9 \times 1.2$ mm^3), and the generated MRA images have higher resolution (e.g., $0.5 \times 0.5 \times 0.8$ mm^3), which yield comparable overall image quality and visualisation of all intracranial vessels to the actually acquired MRA images. Thus, the task tackled by the propose method involves not only image-to-image translation but also super-resolution, implicitly. The ability for generating high-resolution MRA images is attributed to: (i) replacing the traditional pixel-wise loss functions with feature-based loss terms facilitates better recovery of high-frequency image information, (ii) using the effective feature guidance by consistently encoding and emphasising shared spatial features and fine structural details of multi-channel input through the CBN module along the generation pathway.

One limitation of this study is that all the datasets investigated only involve the healthy subjects. Although we did not include patients with arterial pathologies such as aneurysm in our study, the synthesised high-resolution MRA images were $0.5 \times 0.5 \times 0.8$ mm^3 and the quality of MRA images may be sufficiently high to evaluate arterial diseases. Also, based on the observations that the synthetic MRA images exhibit similar legibility and quality as to the real MRA data, we argue that the proposed method with the dedicated modules will not adversely impact the image quality and is capable of reconstructing final details



c

	Imaging Centre	Scanner Model
S1	CHUM Campus Notre-Dame	Philips Achieva
S2	IUGM	Siemens TrioTim
S3	IRM Quebec - Mailoux	Philips Achieva
S4	McConnell Brain Imaging Center	Siemens TrioTim
S5	MRBSTTB17C	Philips Ingenia
S6	Hospital Douglas	Siemens TrioTim

	Imaging Centre	Scanner Model
S7	Anlagenzentrum	Siemens Prisma
S8	Foothills Hospital	GE DISCOVERY
S9	WCMI Uptown	GE SIGNA Pioneer
S10	Hospital for Sick Children	Siemens Prisma
S11	St. Michael's Hospital	Siemens Skyra
S12	Royal University Hospital	Siemens Skyra

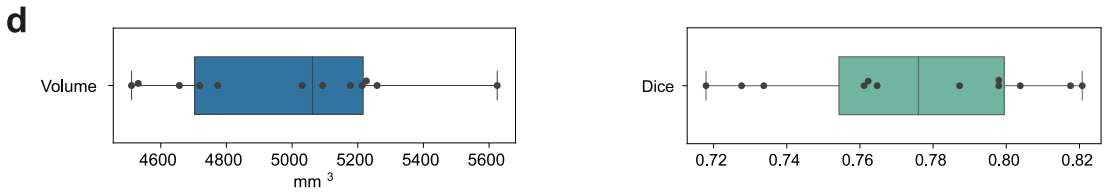


Fig. 13. Synthesised MRA images for the SIMON dataset containing a single subject scanned at multiple sites and with various scanner models. The common MR sequences such as T1, T2 and PD were acquired but MRA acquisitions were not included in the scanning protocol in this dataset. (a) Examples of acquired T2 and PD-weighted MRIs and synthesised MRA. (b) MIPs of the synthesised MRA images from the structural MRIs acquired from 12 sessions. (c) Various sites and scanner models associated with MRIs were acquired in each session. (d) Two quantitative measures on the differences of the synthesised vascular anatomies.

in pathological anatomy such as clogged vessels or aneurysms manifesting as protrusions of arterial walls. Nevertheless, future studies should assess in more detail the generalisability and performance of the proposed network on the MRA image synthesis across pathologies more specifically, for instance, investigating the applicability of the model trained on healthy volunteer data to subjects with intracranial aneurysms. The current results, whilst preliminary, show the potential for use of the proposed approach to generating digital twin cohorts of cerebrovascular anatomy at scale from structural MR images typically acquired in population imaging initiatives such as the UK Biobank. This could serve as a platform to curate libraries of whole-brain vascular geometries that may be used in the future to scale up in-silico studies and/or in-silico trials assessing cerebrovascular blood flow to previously unseen sample sizes.

5. Conclusion

In conclusion, this work proposed an end-to-end generative adversarial network that can synthesise anatomically plausible, high-resolution 3D MRA images using commonly acquired multi-contrast MR images (e.g. T1/T2/PD-weighted MR images) for the same subject whilst preserving the continuity of vascular anatomy and key vascular morphological characteristics. We assessed the performance of our approach across comprehensive experiments, including comparisons against state-of-the-art image synthesis methods, assess the importance of various contrast sources, and explore the reproducibility of our results across databases. Combined with population imaging studies, these results pave the way to create digital twin cohorts of cerebrovascular anatomy at scale.

CRediT authorship contribution statement

Yan Xia: Conceptualization, Methodology, Software, Visualization, Validation, Formal analysis, Writing – original draft. **Nishant Ravikumar:** Writing – review & editing. **Toni Lassila:** Writing – review & editing. **Alejandro F. Frangi:** Conceptualization, Writing - Review & Editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the data used in this study are publicly accessible.

Acknowledgements

AFF acknowledges support from the Royal Academy of Engineering Chair in Emerging Technologies Scheme (CiET1819/19) and the Engineering and Physical Sciences Research Council (EPSRC), UK through TUSCA (EP/V04799X/1).

References

Allenby, M.C., Liang, E.S., Harvey, J., Woodruff, M.A., Prior, M., Winter, C.D., Alonso-Caneiro, D., 2021. Detection of clustered anomalies in single-voxel morphometry as a rapid automated method for identifying intracranial aneurysms. Comput. Med. Imaging Graph. 89, 101888.

Ben-Cohen, A., Klang, E., Raskin, S.P., Amitai, M.M., Greenspan, H., 2017. Virtual PET images from CT data using deep convolutional networks: initial results. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 49–57.

Bi, L., Kim, J., Kumar, A., Feng, D., Fulham, M., 2017. Synthesis of positron emission tomography (PET) images via multi-channel generative adversarial networks (GANs). In: Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Springer, pp. 43–51.

Cebral, J.R., Mut, F., Weir, J., Putman, C., 2011. Quantitative characterization of the hemodynamic environment in ruptured and unruptured brain aneurysms. Am. J. Neuroradiol. 32 (1), 145–151.

Chartsias, A., Joyce, T., Dharmakumar, R., Tsaftaris, S.A., 2017a. Adversarial image synthesis for unpaired multi-modal cardiac data. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 3–13.

Chartsias, A., Joyce, T., Giuffrida, M.V., Tsaftaris, S.A., 2017b. Multimodal MR synthesis via modality-invariant latent representation. IEEE Trans. Med. Imaging 37 (3), 803–814.

Chen, T., Lučić, M., Houlsby, N., Gelly, S., 2019. On self-modulation for generative adversarial networks. In: International Conference on Learning Representations. ICLR.

Cirillo, M.D., Abramian, D., Eklund, A., 2020. Vox2Vox: 3D-GAN for brain tumour segmentation. arXiv preprint arXiv:2003.13653.

Costa, P., Galdran, A., Meyer, M.I., Niemeijer, M., Abràmoff, M., Mendonça, A.M., Campilho, A., 2017. End-to-end adversarial retinal image synthesis. IEEE Trans. Med. Imaging 37 (3), 781–791.

Dalmaz, O., Yurt, M., Çukur, T., 2022. ResViT: residual vision transformers for multimodal medical image synthesis. IEEE Trans. Med. Imaging 41 (10), 2598–2614.

Dar, S.U., Yurt, M., Karacan, L., Erdem, A., Erdem, E., Cukur, T., 2019. Image synthesis in multi-contrast MRI with conditional generative adversarial networks. IEEE Trans. Med. Imaging 38 (10), 2375–2388.

Dar, S.U., Yurt, M., Shahdloo, M., Ildız, M.E., Tınaz, B., Çukur, T., 2020. Prior-guided image reconstruction for accelerated multi-contrast MRI via generative adversarial networks. IEEE J. Sel. Top. Sign. Proces. 14 (6), 1072–1087.

De Vries, H., Strub, F., Mary, J., Larochelle, H., Pietquin, O., Courville, A.C., 2017. Modulating early visual processing by language. In: Advances in Neural Information Processing Systems. pp. 6594–6604.

Dewey, B.E., Zhao, C., Carass, A., Oh, J., Calabresi, P.A., van Zijl, P.C., Prince, J.L., 2018. Deep harmonization of inconsistent MR data for consistent volume segmentation. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 20–30.

Dewey, B.E., Zhao, C., Reinhold, J.C., Carass, A., Fitzgerald, K.C., Sotirchos, E.S., Saidha, S., Oh, J., Pham, D.L., Calabresi, P.A., et al., 2019. DeepHarmony: A deep learning approach to contrast harmonization across scanner changes. Magn. Reson. Imaging 64, 160–170.

Durugkar, I., Gemp, I., Mahadevan, S., 2017. Generative multi-adversarial networks. In: International Conference on Learning Representations. ICLR.

Fujita, S., Hagiwara, A., Otsuka, Y., Hori, M., Takei, N., Hwang, K.P., Irie, R., Andica, C., Kamagata, K., Akashi, T., et al., 2020. Deep learning approach for generating MRA images from 3D quantitative synthetic MRI without additional scans. Invest. Radiol. 55 (4), 249–256.

Hagiwara, A., Otsuka, Y., Hori, M., Tachibana, Y., Yokoyama, K., Fujita, S., Andica, C., Kamagata, K., Irie, R., Koshino, S., et al., 2019. Improving the quality of synthetic FLAIR images with deep learning using a conditional generative adversarial network for pixel-by-pixel image translation. Am. J. Neuroradiol. 40 (2), 224–230.

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

Hiasa, Y., Otake, Y., Takao, M., Matsuoka, T., Takashima, K., Carass, A., Prince, J.L., Sugano, N., Sato, Y., 2018. Cross-modality image synthesis from unpaired data using CycleGAN. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 31–41.

Holmgren, M., Støverud, K.H., Zarrinkoob, L., Wåhlin, A., Malm, J., Eklund, A., 2021. Middle cerebral artery pressure laterality in patients with symptomatic ICA stenosis. PLoS One 16 (1), e0245337.

Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1125–1134.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5 (2), 143–156.

Jog, A., Carass, A., Pham, D.L., Prince, J.L., 2014. Random forest FLAIR reconstruction from T1, T2, and PD-weighted MRI. In: 2014 IEEE 11th International Symposium on Biomedical Imaging. ISBI, IEEE, pp. 1079–1082.

Jog, A., Carass, A., Pham, D.L., Prince, J.L., 2015. Tree-encoded conditional random fields for image synthesis. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 733–745.

Jog, A., Carass, A., Roy, S., Pham, D.L., Prince, J.L., 2017. Random forest regression for magnetic resonance image synthesis. Med. Image Anal. 35, 475–488.

Johnson, J., Alahi, A., Fei-Fei, L., 2016. Perceptual losses for real-time style transfer and super-resolution. In: European Conference on Computer Vision. Springer, pp. 694–711.

Joyce, T., Chartsias, A., Tsaftaris, S.A., 2017. Robust multi-modal MR image synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 347–355.

Larrabide, I., Geers, A.J., Morales, H.G., Aguilar, M.L., Rüfenacht, D.A., 2015. Effect of aneurysm and ICA morphology on hemodynamics before and after flow diverter treatment. J. Neurointerventional Surg. 7 (4), 272–280.

Lee, D., Kim, J., Moon, W.J., Ye, J.C., 2019. CollaGAN: Collaborative GAN for missing image data imputation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2487–2496.

Lee, D., Moon, W.J., Ye, J.C., 2020. Assessing the importance of magnetic resonance contrasts using collaborative generative adversarial networks. Nat. Mach. Intell. 2 (1), 34–42.

Li, H., Paetzold, J.C., Sekuboyina, A., Kofler, F., Zhang, J., Kirschke, J.S., Wiestler, B., Menze, B., 2019. DiamondGAN: unified multi-modal generative adversarial networks for MRI sequences synthesis. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 795–803.

Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Paul Smolley, S., 2017. Least squares generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2794–2802.

Mehta, R., Arbel, T., 2018. RS-Net: Regression-segmentation 3D CNN for synthesis of full resolution missing brain MRI in the presence of tumours. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 119–129.

Nguyen, T., Le, T., Vu, H., Phung, D., 2017. Dual discriminator generative adversarial nets. In: Advances in Neural Information Processing Systems. pp. 2670–2680.

Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., Wang, Q., Shen, D., 2018. Medical image synthesis with deep convolutional adversarial networks. IEEE Trans. Biomed. Eng. 65 (12), 2720–2730.

Olut, S., Sahin, Y.H., Demir, U., Unal, G., 2018. Generative adversarial training for MRA image synthesis using multi-contrast MRI. In: International Workshop on Predictive Intelligence in Medicine. Springer, pp. 147–154.

Özbey, M., Dar, S.U., Bedel, H.A., Dalmaz, O., Özturk, Ş., Güngör, A., Çukur, T., 2022. Unsupervised medical image translation with adversarial diffusion models. arXiv preprint arXiv:2207.08208.

Park, T., Liu, M.Y., Wang, T.C., Zhu, J.Y., 2019. Semantic image synthesis with spatially-adaptive normalization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2337–2346.

Sajjadi, M.S., Bachem, O., Lucic, M., Bousquet, O., Gelly, S., 2018. Assessing generative models via precision and recall. In: NeurIPS.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X., 2016. Improved techniques for training GANs. In: Advances in Neural Information Processing Systems. pp. 2234–2242.

Sharma, A., Hamarneh, G., 2019. Missing MRI pulse sequence synthesis using multi-modal generative adversarial network. IEEE Trans. Med. Imaging 39 (4), 1170–1183.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219.

Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B., 2018. High-resolution image synthesis and semantic manipulation with conditional GANs. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8798–8807.

Warfield, S.K., Zou, K.H., Wells, W.M., 2004. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Trans. Med. Imaging 23 (7), 903–921.

Wei, W., Poirion, E., Bodini, B., Durrleman, S., Ayache, N., Stankoff, B., Colliot, O., 2018. Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 514–522.

Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg, C.A., Išgum, I., 2017a. Deep MR to CT synthesis using unpaired data. In: International Workshop on Simulation and Synthesis in Medical Imaging. Springer, pp. 14–23.

Wolterink, J.M., Dinkla, A.M., Savenije, M., Seevinck, P.R., van den Berg, C., Išgum, I., 2017b. MR-to-CT synthesis using cycle-consistent generative adversarial networks. In: Proc. Neural Inf. Process. Syst. NIPS.

Wright, S.N., Kochunov, P., Mut, F., Bergamino, M., Brown, K.M., Mazziotta, J.C., Toga, A.W., Cebral, J.R., Ascoli, G.A., 2013. Digital reconstruction and morphometric analysis of human brain arterial vasculature from magnetic resonance angiography. Neuroimage 82, 170–181.

Xia, Y., Zhang, L., Ravikumar, N., Attar, R., Piechnik, S.K., Neubauer, S., Petersen, S.E., Frangi, A.F., 2020. Recovering from missing data in population imaging–Cardiac MR image imputation via conditional generative adversarial nets. Med. Image Anal. 67, 101812.

Yurt, M., Dar, S.U., Erdem, A., Erdem, E., Oguz, K.K., Çukur, T., 2021. mustGAN: Multi-stream generative adversarial networks for MR image synthesis. Med. Image Anal. 70, 101944.

Zhao, H., Li, H., Maurer-Stroh, S., Cheng, L., 2018. Synthesizing retinal and neuronal images with generative adversarial nets. Med. Image Anal. 49, 14–26.

Zhou, T., Fu, H., Chen, G., Shen, J., Shao, L., 2020. Hi-net: hybrid-fusion network for multi-modal MR image synthesis. IEEE Trans. Med. Imaging 39 (9), 2772–2781.