Exploring sequence transformation in magnetic resonance imaging via deep learning using data from a single asymptomatic patient

Journal of Physics Communications

PAPER

Exploring sequence transformation in magnetic resonance imaging via deep learning using data from a single asymptomatic patient

James A Grant-Jacob[1], Chris Everitt[2], Robert W Eason[1], Leonard J King[3] and Ben Mills[1]

[1] Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, United Kingdom
[2] Department of Medical Physics, University Hospital Southampton, Southampton, SO16 6YD, United Kingdom
[3] NIHR Southampton Biomedical Research Centre, University Hospital Southampton, SO16 6YD, United Kingdom

E-mail: J.A.Grant-Jacob@soton.ac.uk

Keywords: deep learning, optics, MRI, medical imaging

OPEN ACCESS

RECEIVED
19 March 2021

REVISED
3 September 2021

ACCEPTED FOR PUBLICATION
8 September 2021

PUBLISHED
27 September 2021

Abstract

We investigate the potential for deep learning to create a transfer function from T1 to T2 magnetic resonance imaging sequences using data collected from an asymptomatic patient. Neural networks were trained on images of a human left hand, and then applied to convert T1 images to T2 images for the associated right hand. Analysis showed that the most accurate neural network considered the features in the surrounding $\sim$1 cm when converting to T2, hence indicating that the neural network was able to identify structural correlations between the sequences. However, some small features measuring $<$2 mm differed, and grid patterning was evident from the images. While using deep learning for sequence transformations could enable faster processing and diagnosis and in turn reduce patient waiting times, additional work, such as synergising physics-based modelling with neural networks, will likely be required to demonstrate that deep learning can be used to accurately create T2 characteristics from T1 images. In addition, since the present work was conducted using data collected from a single patient, further example datasets collected from patients with a range of different pathologies will be required in order to validate the proposed method.

Introduction

Medical imaging has been a fundamental technique for identifying injuries and health conditions for over 100 years. Key inventions include the invention of x-ray imaging for medical purposes in 1896 by John Hall-Edwards [1, 2], pioneering portable radiography vans by Marie Curie in World War I [3], and computed tomography (CT) [4, 5]. Other imaging techniques include positron emission tomography scanning (PET) [6], ultrasound [7] and magnetic resonance imaging (MRI) [8, 9]. Depending on the imaging technique, the cost of such processes can be high, due to the initial expense of the machinery, cost of maintenance and the expertise needed to run and evaluate results [10]. A single MRI scan in England performed under the National Health Service (NHS) can cost between £53 and £617 (2015–2016) depending on the length (resolution and size of imaging region) and type of scan [11], with larger volumes of imaging and multiple sequences requiring more time and thus expenditure. By way of example, in this work, imaging a human hand via MRI took approximately 4–7 min for each sequence. Therefore, there is much interest in the development of methods for enabling faster turnaround of MRI patients to produce a more efficient and less costly service. In England, 264,520 MRI scans were carried out in the 12 months up to the end of January 2021, with the median waiting time for an MRI scan being approximately 2.5 weeks [12]. Since multiple sequences are often carried out per MRI appointment, methods that can reduce the waiting times via increasing patient throughput, such as by reducing the number of sequences administered out per patient appointment, would be beneficial.

Owing to the increased computer graphics processing power in recent years, deep learning algorithms have gained noticable attention [13–15]. These algorithms are of interest owing to their ability to carry out identification [16, 17], whether it be in sound waves, such as for voice recognition [18] or bird audio classification [19, 20], or image processing for facial recognition [21, 22], large-scale video classification [23],

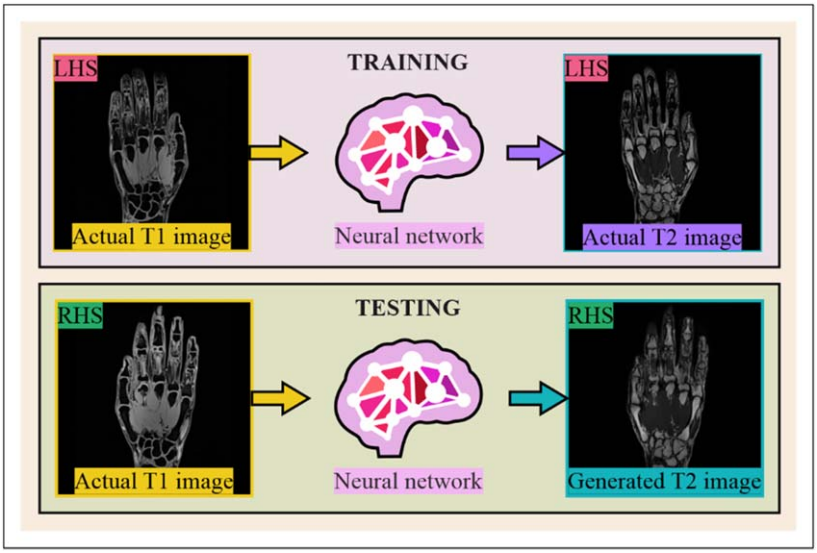

Figure 1. Schematic of the concept for using deep learning to transform T1 VIBE sequence images into T2 SPACE sequence images, via training on images from the left hand and testing on images from the right hand to generate T2 SPACE sequence images.

pollution identification [24–26], and laser materials processing [27–29]. Beyond classification algorithms, deep learning has been used for transferring one image domain to another, such as from a sketch to a photograph [30], for colourising black and white photographs [31], for transforming low-magnification images into high-magnification images [32, 33], and for transforming scattering patterns into images [34]. Deep learning has been applied to a wide range of medical imaging [35–39], including chest x-ray labelling [40], medical ultrasound [41], PET image enhancement [42], attenuation correction for MRI PET [43], radiation therapy [44], dental tomography [45], bone x-ray [46], lung abnormality detection [47], pancreas CT segmentation [48], and CT reconstruction [49]. Further discussions as well as comparisons of image-to-image generation deep learning neural networks for medical imaging have been carried out by a variety of authors [50–53].

Each medical imaging technique offers distinct advantages and disadvantages, specifically in terms of which features can be identified on the image output. For example, some bone fractures may not be visible on plain radiographs, but might be visible on CT, while different MRI sequences can be utilised to display different tissue characteristics including fat and water content. T1 VIBE (volumetric interpolated breath-hold examination) magnetic resonance images with water excitation show fat in the human body as low signal intensity and water as intermediate signal, whereas T2 SPACE (sampling perfection with application optimized contrasts using different flip angle evolution) images show both fat and water as high signal. As such, work has progressed on being able to synthesise one modality from another, such as MRI to CT scan [54, 55]. Additional applications of deep learning for MRI include accelerating magnetic resonance [56], deep learning Bloch equation simulations (DeepBLESS) for rapid and accurate T1 estimation [57], MRI brain extraction in 3D [58], super-resolution musculoskeletal MRI [59], MRI finger printing [60, 61], random forest magnetic resonance for MRI synthesis [62] and, most relevant to this work, generation of T2 sequence images from the associated combination of T1 with under-sampled T2 sequence images [63].

As shown by the concept in figure 1, we aimed to use deep learning neural network architecture to generate T2 SPACE sequence images directly from T1 VIBE sequence images by training a neural network on the left hand coronal, sagittal and axial T1 sequence images, to produce an equivalent T2 SPACE sequence output image. The neural network was then tested on the right hand T1 VIBE sequence images to produce T2 SPACE sequence images. Our primary objective for this work was to demonstrate accurate image-to-image generation for MRI sequence images from a single asymptomatic patient, and then establish technical analytical methods for evaluating the effectiveness of the trained network, which could potentially be applied to other studies involving more patients.

A commonly used neural network architecture for paired (input and output are for example, same image slice in space but different sequence) image-to-image transformations is the Pix2Pix cGAN (conditional generative adversarial neural network) model [64] that has a 'U-Net' based architecture for the generator [65], and a convolutional 'PatchGAN' classifier for the discriminator [66], which penalizes structure on the scale of the image patches. Olut *et al* [67] discuss their results of using Pix2Pix for Magnetic Resonance Angiography (MRA) for generating non-existent MRA from T1- and T2-weighted MRI images, which could be a valuable tool in retrospective subject evaluation of vascular anatomy and related diseases. Zhou *et al* [68] explore the

generation of one MRI sequence from hybrid fusion of two other sequences using Pix2Pix, such as fusing T1- and T2-weighted MRI images of the brain to form fluid attenuated inversion recovery (FLAIR)-weighted images. The Pix2Pix model has been used by Shin *et al* [69] to generate MRI T1-weighted images from segmented labelled images (label-to-MRI), to generate synthetic abnormal MRI images with brain tumours, which could be used for expanding datasets with pathological findings.

An image transformation neural network that utilises unsupervised learning, which has the benefit of not requiring paired training, is CycleGan [70]. This method uses a cycle-consistent adversarial neural network, and has been employed in work for image transformation of MRI T1-weighted sequence images to fractional anisotropy (FA) [71] and computed tomography (CT) images synthesis from MRI images [72, 73]. Another type of unsupervised image transformation neural network is UNIT (unsupervised image-to-image translation network) [74], which unlike CycleGan, uses a shared latent space. Implementation of UNIT in medical imaging has been demonstrated for brain image transformation of electroencephalogram (EEG) to functional magnetic resonance imaging (fMRI) images [75], as well as for non-contrast and contrast enhanced CT scans of kidneys [76].

While the work referenced above consisted of using neural networks on data from multiple patients, smaller datasets using single patients have also been investigated in the field of medical imaging. Training on the perfusion MRI data of a single acute stroke patient in order to predict the final infarct of the same patient has been explored by Debs *et al* [77], while comparisons of deep learning methods trained and tested on single patient ECG for seizure detection were performed by Turner *et al* [78] and a pilot study of estimating full-dose PET images from low-dose PET images of the whole body were performed by Kaplan *et al* [79] using training data from one patient and then tested on another patient.

In this work, we initially explore three types of neural network, for generating T2 SPACE sequence images from T1 VIBE sequence images of a single asymptomatic patient, and use the network with the most accurate results for a range of novel digital pathology analysis.

Experimental methods

Data collection

The MRI dataset was acquired for the purposes of this study on a single asymptomatic healthy volunteer following informed consent, to establish the utility and efficacy of the method prior to extension to an enlarged data set for any further more in-depth extended study. The MRI examination was performed under standard clinical conditions in accordance with routine safety protocols following a screening questionnaire to exclude contra-indications. Imaging was performed on a Skyra 3.0-T MRI scanner (Siemens, Erlangen, Germany) at the University Hospital Southampton NHS Foundation Trust. Two MRI sequences were obtained in the coronal plane for each hand on a single subject; T1 VIBE (TR/TE msec,13.5/6; flip angle 10 degrees; water excitation; acquisition matrix 512 $\times$ 512; voxel size 0.6 mm $\times$ 0.6 mm $\times$ 0.6 mm; acquisition time 4 min 19 seconds for each hand, and T2 SPACE (TR/TE msec,1500/127; acquisition matrix 512 $\times$ 512; voxel size 0.6 mm $\times$ 0.6 mm $\times$ 0.6 mm; acquisition time 7 min 5 seconds for each hand. Each set of coronal images were reformatted in the sagittal and axial plane, creating 223 sagittal, 76 coronal and 491 axial image slices. In the analysis section, the reformatted sagittal, coronal and axial image planes correspond to X-Z, Y-Z and X-Y image planes, respectively.

Neural networks

Three types of neural networks were used. These were image-to-image (Pix2Pix), image-to-image-to-image model CycleGan, and unsupervised image-to-image (UNIT). A learning rate of 0.0002 was implemented for all neural networks, which were all trained for 15 epochs for consistency of comparison. The Pix2Pix and CycleGan were trained using an NVIDIA RTX 2080 graphics processing unit (GPU) for a total of 1 h 35 min and 4 h 14 min, respectively, while the UNIT neural network was trained using an NVIDIA QUADRO P6000 and took ~80 h.

Critically, only left hand coronal, sagittal and axial images, each with a resolution of 512 $\times$ 512 pixels, were used in training. After training, each neural network was tested on the right hand T1 images. All 790 images of the left hand (all coronal, sagittal and axial) were used for training the neural network. Data augmentation (i.e. shifting, cropping, rotating and resizing) to increase the amount of training data was not performed in case there was some spatial dependence on the MRI and generation of the T2 images. However, this is something that should be explored in future work through use of a larger sample size, as such analysis may provide insight into the trained neural network, such as the effect of image generation on location or orientation within the 3D volume.

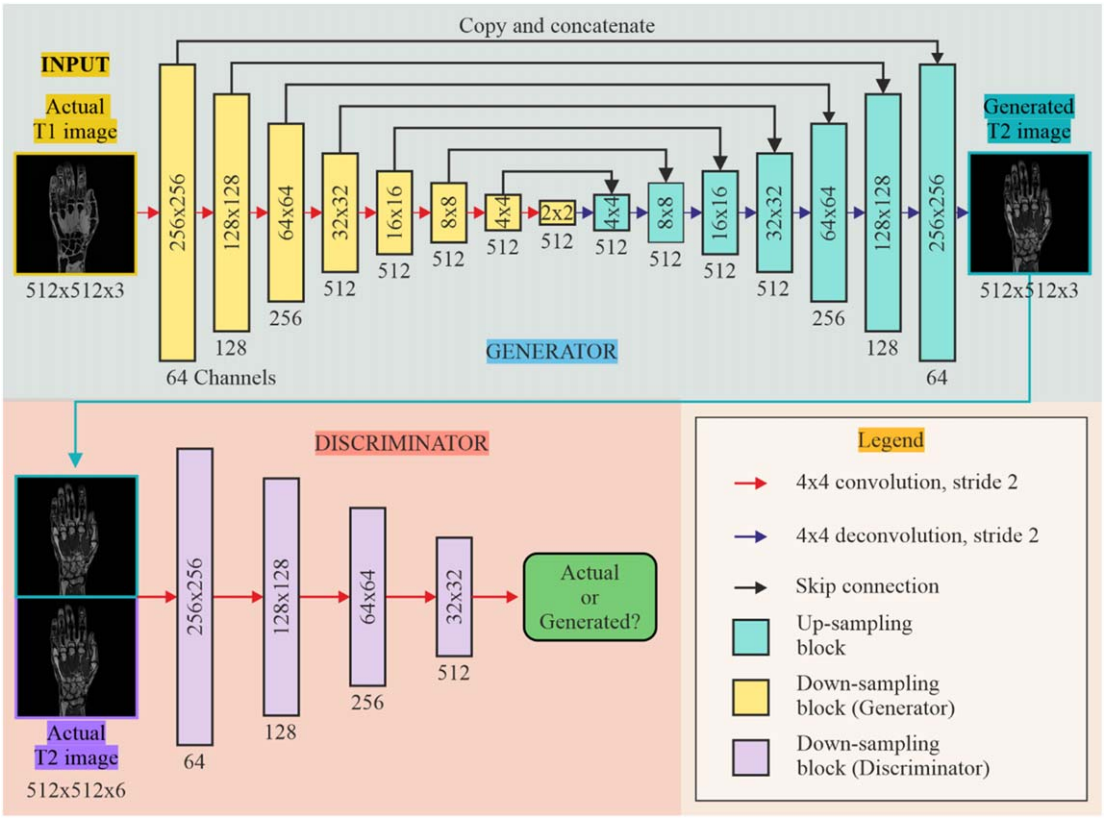

Figure 2. Diagram illustrating the Pix2Pix network architecture.

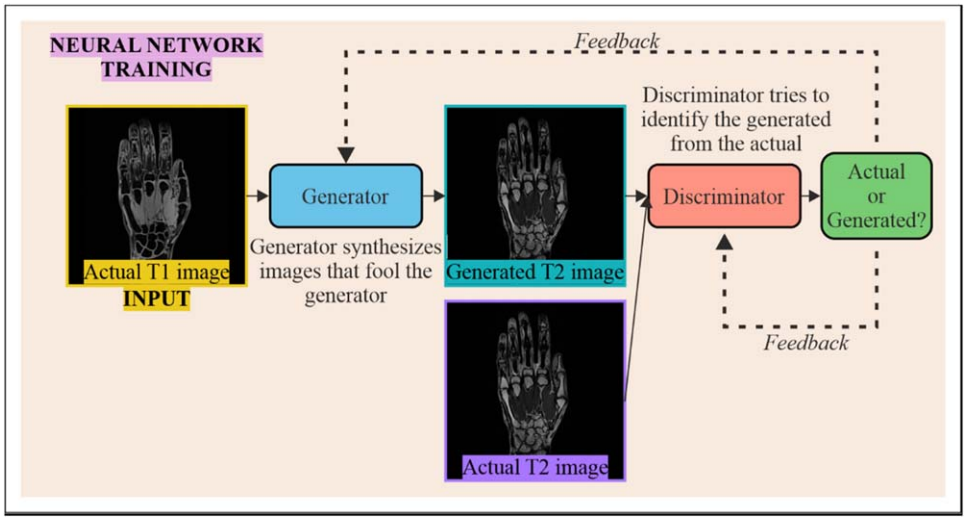

Figure 3. Diagram illustrating the process for training the Pix2Pix neural network.

Pix2Pix neural network architecture

A diagram of the Pix2Pix neural network architecture used in this work is shown in figure 2. The coloured rectangular boxes represent blocks of multi-channel feature maps, with each map's dimension indicated inside and number of channels indicated below. The U-Net architecture of the generator has a contracting path (yellow boxes) and expansive path (cyan boxes), with skip connections between each centre symmetric layer. The red and blue arrows represent down-sampling and up-sampling convolutions, respectively, while black arrows represent skip connections. Each skip connection concatenates thefeature maps from the expansive path with the equivalent layer feature maps from the contracting path. The contracting path consists of convolutional blocks for down-sampling, in which convolutional filters of size 4×4 with stride 2 are applied to the feature map to double the number of feature channels, followed by a batch-normalisation layer and a rectified linear

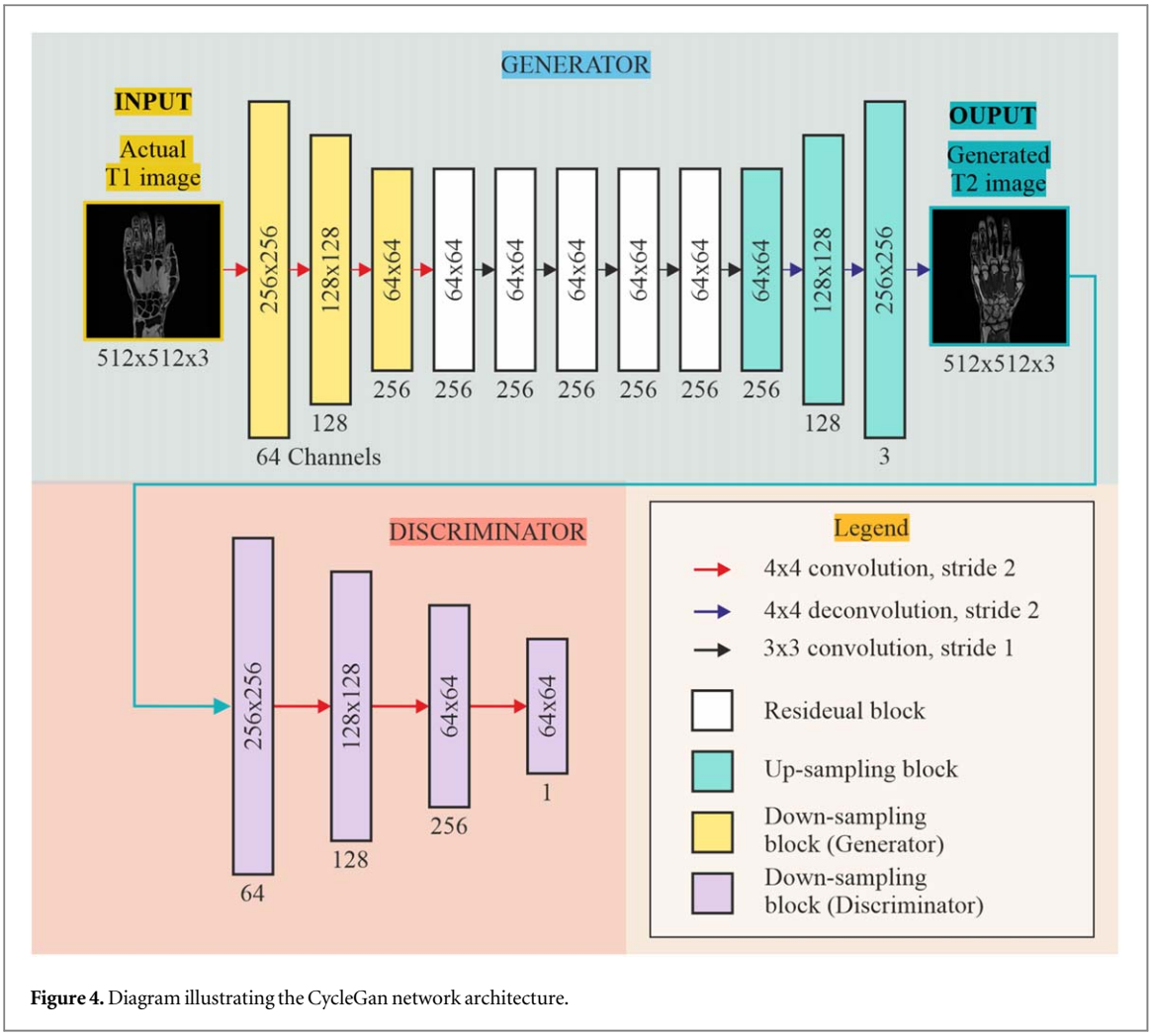

Figure 4. Diagram illustrating the CycleGan network architecture.

unit (ReLU) of 0.2. The expansive path consists of up-sampling, in which 4 × 4 convolution filters with stride of 2 are applied to the feature map to halve the number of channels, followed by a batch-normalisation layer and a ReLU of 0.2, then a concatenation with the equivalent layer feature maps from the contracting path. The discriminator also has blocks of convolutional layers with a 4 × 4 convolution filter and stride of 2, followed by a batch normalisation layer and a ReLU of 0.2. An L1 loss (defined as the least absolute deviations) between the generated images and the actual experimental images is produced, such that the L1 loss should be as small as possible for the generator, to allow it to fool the discriminator. The L1 loss of the neural network at the end of training was 0.011, while the GAN loss was 0.69 and the discriminator loss was 0.67. At the start of training, the neuron weightings for the generator and discriminator were randomly initialised.

Figure 3 shows a schematic of the Pix2Pix neural network training process, which involves single generator and single discriminator. An actual T1 image was used as input to the generator network (along with a 2D array of noise), which produced a generated T2 image. In line with nomenclature for this field, the experimentally measured images are referred to as actual, and the images produced by the neural network are referred to as generated. During training, the discriminator received either the actual T1 and actual T2, or the actual T1 and generated T2, and had to identify which combination was actual and which was generated. At the same time, the generator was trained to fool the discriminator, by generating images that were visually similar to the real images. The motivation for such adversarial training is to reach a Nash equilibrium [80], where the generated images are indistinguishable from the actual images. At this point, the generator network can be used to convert any T1 image into the associated T2 image.

CycleGan architecture

A diagram of the CycleGan generator and discriminator architecture is shown in figure 4. As per Pix2Pix, a 512 × 512 × 3 image is fed into the neural network and down-sampled, in which the channels increase and size of the feature map decreases using blocks of 4 × 4 convolutional filters with a stride of 2, followed by

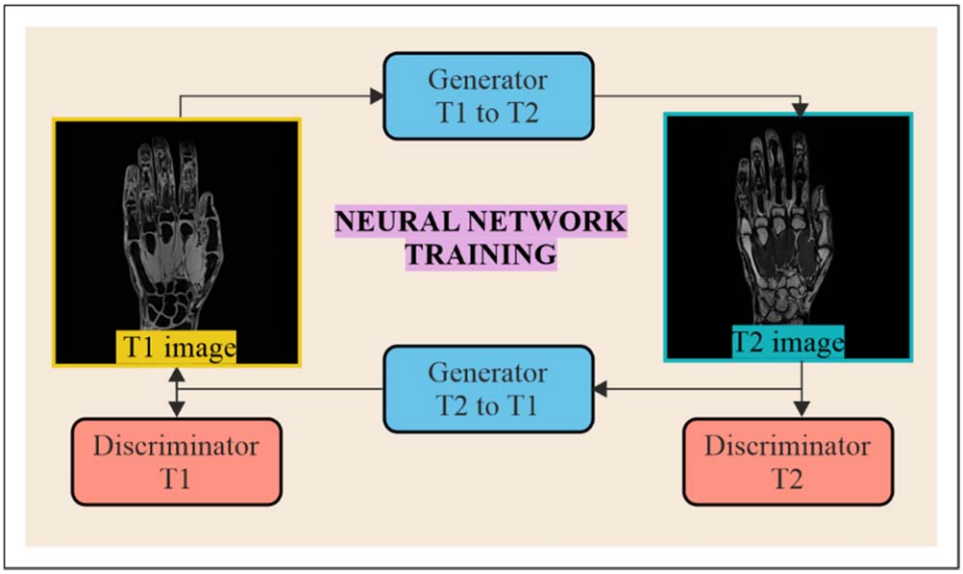

Figure 5. Diagram illustrating the process for training the CycleGan neural network.

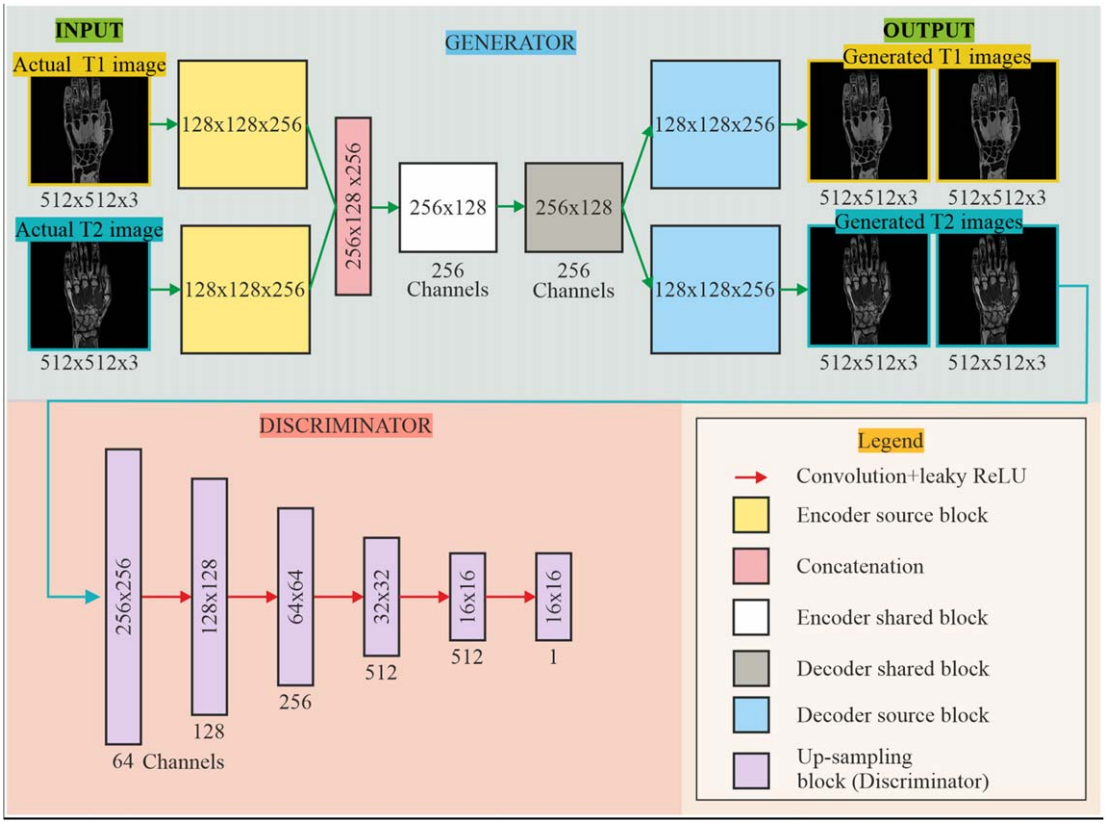

Figure 6. Diagram illustrating the UNIT network architecture.

batch-normalisation and a ReLU activation function. Then the image is passed into 5 residual blocks consisting of 3 $\times$ 3 convolutional filters with a stride of 1, each followed by a batch-normalisation and ReLU. The image is then up-sampled and the number of channels decreased using several blocks consisting of 4 $\times$ 4 convolutional filters with a stride of 2, batch normalisation and a ReLU activation function, apart from the last layer which has a Tanh activation function. The discriminator consists of 4 layer blocks, with convolutional filters with size 4 $\times$ 4 and a stride of 2, with a ReLU following the 1–3 convolutional filters and a batch normalisation following 2–3 convolutional filters.

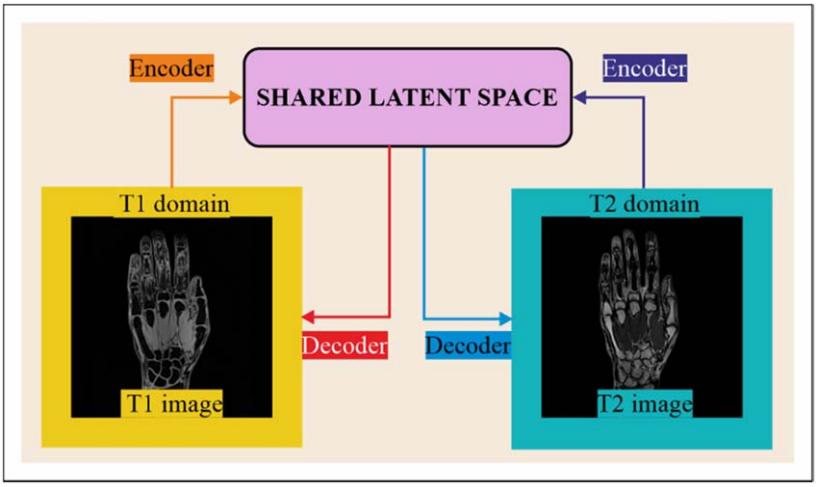

Figure 7. Diagram illustrating the shared latest space for training the UNIT neural network.

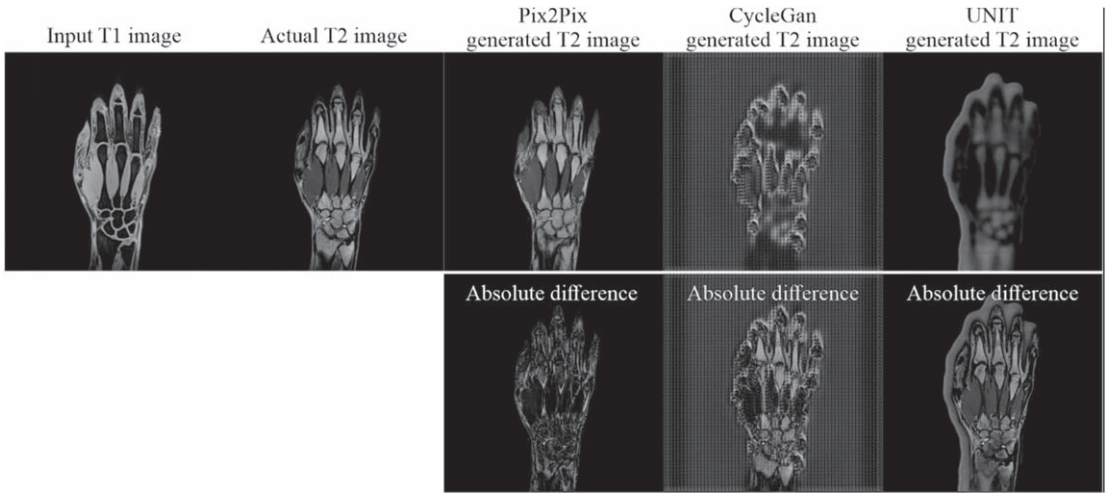

Figure 8. Capability of the trained neural network for generating a T2 image of the right hand from a single asymptomatic patient, showing input T1 image (1st column), actual T2 images (2nd column), Pix2Pix generated T2 images (3rd column), CycleGan generated T2 images (4th column) and UNIT generated T2 images (5th column), for the same coronal view of the centre of the hand. The absolute differences between the generated and actual T2 images are displayed in the 2nd row.

As seen in figure 5, the CycleGAN neural network consists of two generators and two discriminators, where one generator takes in T1 images and generates T2 images, which are then passed into the corresponding discriminator along with actual T2 images, which tries to correctly classify the images as actual or generated. Likewise, the other generator takes in T2 images and generates T1 images, which are then passed into the corresponding discriminator along with actual T1 images, which tries to correctly classify them as actual or generated.

UNIT architecture

The inputs to the UNIT neural network generator are T1 and T2 images and the output is four images. Two of the output images are transformed images, T1 to T2 and T2 to T1, while the other two are self-reconstructed images, T1 to T1 and T2 to T2. Moreover, the UNIT network consists of a generator that has two input images, one from each domain, T1 and T2, and both of these are fed into their own encoder to give $128 \times 128 \times 256$ activations in each block (see figure 6). The outputs of these encoders are concatenated along one dimension to form a $256 \times 128 \times 256$ output. Following this, the output is sent into a shared encoder block and then into a shared decoder block, again with $256 \times 128 \times 256$ activations. Finally, the output is sent to either a T1 decoder block or a T2 decoder block, where the output has the same dimensions as the input images $512 \times 512 \times 3$. Then a T1 discriminator takes in real and generated T1 images and evaluates whether they are realistic, and

likewise a T2 discriminator takes in real and generated T2 images and evaluates whether they are realistic. This is done via a series of convolutional filters and leaky ReLUs [81].

As shown in figure 7, a pair of corresponding images in two different domains, T1-weighted and T2-weighted, can be mapped to the same latent code in a shared-latent space, using encoders mapping the images to the latent codes and the decoders mapping the latent codes to the images.

Results and discussion

The three trained neural networks were tested on actual T1 images of the right hand, and the generated T2 images were compared to the actual T2 experimental images from the single asymptomatic patient. Figure 8 shows a 512 × 512 1 channel 8-bit input T1 image (first column), actual T2 image (second column), Pix2Pix generated T2 image (third column), CycleGan generated T2 image (fourth column) and UNIT generated T2 image (fifth column) of the same coronal view of the centre of the hand. Since the pixel intensity of an MRI sequence image is related to specific characteristics (i.e. of bone, fat etc), normalising the image (such as to the maximum value) would change the intensities and thus characteristics displayed. The absolute difference between the T2 generated images and the actual T2 images are displayed below their respective generated image (one minus the other), such that the higher intensity (whiter value) signal corresponds to greater difference in pixel intensity value.

It is evident upon looking at figure 8 that the Pix2Pix, under the same learning rate and number of epochs, has more accurately generated the actual T2 image, compared with the CycleGan and UNIT neural networks. Indeed, the characteristics and associated colour and shape seem to be preserved. This is perhaps due to the multiple concatenation connections and greater depth of the Pix2Pix neural network. The absolute difference image between the generated T2 image for Pix2Pix is visibly darker compared with the CycleGan and UNIT absolute difference images, indicating a greater accuracy of image generation for Pix2Pix. Error analysis on the images also indicates this, and is detailed below.

The normalised root mean square error (NRMSE) was calculated by taking the mean of the squared difference between the intensity value of each pixel in the generated image (0–255 intensity range) and that of the actual experimental image (0–255 intensity range),

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(G_i - I_i)^2}$$

$$NMRSE = \frac{RMSE}{I_{i\,\max} - I_{i\,\min}}$$

where N is the number of data points (pixels), G_i is the generated pixel value and I_i is the actual pixel value, with $I_{i\max}$ being the maximum pixel value and $I_{i\min}$ being the minimum pixel value of the actual image. The lower the value of NMRSE, the smaller the difference between the generated and actual images. The mean NRMSE and standard deviation for all the generated images is 0.0396 ± 0.0175 for Pix2Pix, 0.0844 ± 0.0596 for CycleGan, and 0.0473 ± 0.0207 for UNIT.

To further quantify the performance of the image generation by the Pix2Pix neural network, we determine the Peak Signal-to-Noise Ratio (PSNR) of all the images generated, defined as,

$$PSNR = 10\log_{10}\left(\frac{\max^2(I,\ G)}{\frac{1}{N\times M}\sum_{M,N}(I(m,\ n) - G(m,\ n))^2}\right)$$

where N and M are the total number of rows and columns of pixels in the images, m and n are the pixels in each row and column, and max(I, G) is the maximum intensity value of the actual ground-truth image I and the generated image G. The mean PSNR (greater number means greater accuracy of image generation) and standard deviation for all the generated images is 40.8 ± 13.7 for Pix2Pix, 25.3 ± 4.3 for CycleGan, and 27.5 ± 4.6 for UNIT.

In addition, we have also calculated the Structural Similarity Index Measurement (SSIM) between the generated and the actual images, which assesses the visual impact of image luminance, contrast and structure, defined as,

$$SSIM(I,\ G) = \frac{(2\mu_I\mu_G + C_1)(2\sigma_{IG} + C_2)}{(\mu_I^2 + \mu_G^2 + C_1)(\sigma_I^2 + \sigma_G^2 + C_2)}$$

where μ_I is the average of I, μ_G is the average of G, σ_I^2 is the variance of I, σ_G^2 is the variance of G, σ_{IG} is the covariance of I and G, $C_1 = (0.01\,L)^2$ and $C_2 = (0.03\,L)^2$, such that L is the dynamic range of the pixel values.

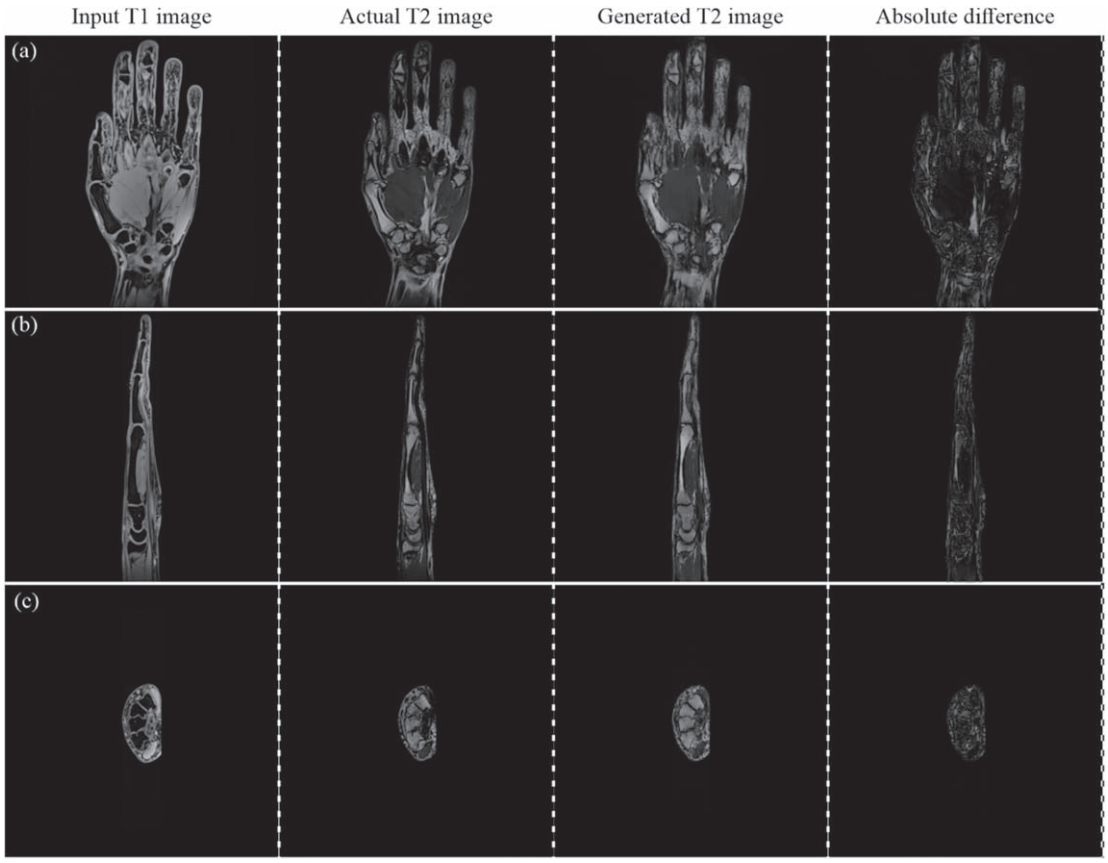

Figure 9. Capability of the trained neural network for generating a T2 image of the right hand, showing T1 images (1st column), actual T2 images (2nd column), generated T2 images (3rd column) and absolute difference between actual and generated (4th column), for (a) coronal, (b) sagittal (along the middle finger) and (c) axial planes.

Table 1. NRMSE, PSNR and SSIM of the images generated by the three trained neural networks.

Neural network	NRMSE	PSNR (dB)	SSIM
Pix2Pix	0.0396 ± 0.0175	40.8 ± 13.7	0.9676 ± 0.0359
CycleGan	0.0844 ± 0.0596	25.3 ± 4.3	0.3921 ± 0.3612
UNIT	0.0473 ± 0.0207	27.5 ± 4.6	0.8230 ± 0.0650

These results are presented in table 1 below. The mean SSIM and standard deviation for all the generated images is 0.9676 ± 0.0359 for Pix2Pix, 0.3921 ± 0.3612 for CycleGan, and 0.8230 ± 0.0650 for UNIT.

The lower the NRMSE, the higher the PSNR value and the higher the SSIM value (max 1), the higher accuracy of generated images compared with the actual experimental images. Visually, and as shown by the error numbers, the Pix2Pix is clearly the most accurate neural network trained under the conditions (learning rate and epochs) described in this work.

Analysis

Since Pix2Pix produced the most accurate image generation, the results were further explored. Additional examples of the Pix2Pix neural network image generation is shown in figure 9, which presents the input T1 image (first column), actual T2 image (second column), generated T2 image (third column), and absolute difference (forth column) for (a) coronal view at the centre of the hand, (b) sagittal view down the middle finger and below, and (c) axial view of the wrist. The fourth column shows the absolute difference between the actual and generated images.

The neural network was able to generate visually similar images to the actual sequence type. As shown in figure 9, muscle on the generated T2 images is of similar intermediate signal intensity to the actual T2 images,

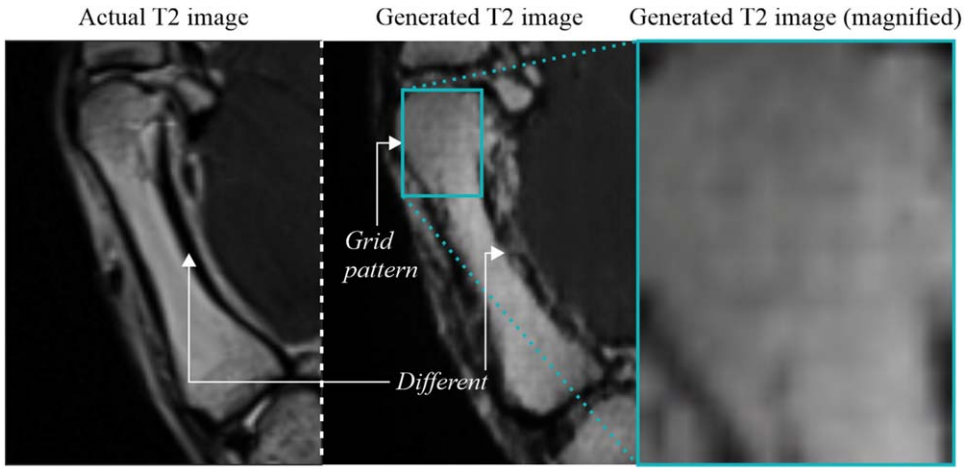

Figure 10. Actual T2 image (left), generated T2 image (middle) and magnified generated T2 image (right), for the metacarpal bone of the thumb. An artefactual grid pattern can be seen on the generated image.

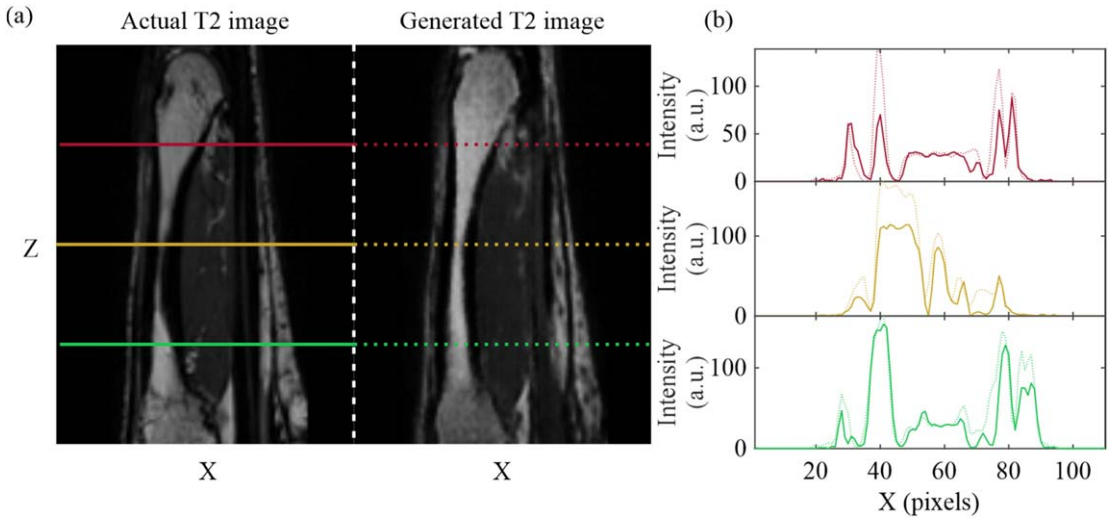

Figure 11. (a) Actual T2 image (left), generated T2 image (right) and (b) slice through the actual and generated images, indicated by the red, yellow and green solid and dotted lines, respectively.

Table 2. NRMSE, PSNR and SSIM for all the Pix2Pix generated images in each plane.

Image plane	NRMSE	PSNR (dB)	SSIM
Coronal	0.0650 ± 0.0188	26.4 ± 2.4	0.8448 ± 0.0212
Sagittal	0.0315 ± 0.0081	34.9 ± 5.2	0.9759 ± 0.0053
Axial	0.0440 ± 0.0191	48.8 ± 15.7	0.9768 ± 0.0192

and bone marrow is of similar high signal. Specific features present in the actual images, such as localised regions of high signal are also present in the actual and generated images, for example at the 5th carpometacarpal joint between the hamate and the base of the 5th metacarpal in figure 9(a).

As shown in table 2, the NRMSE between the generated and the actual T2 image for coronal, sagittal and axial images shown in figure 9 is 0.0650 ± 0.0188, 0.0315 ± 0.0081 and 0.0440 ± 0.0191, respectively. The mean PSNR and standard deviation for the images in the coronal plane, sagittal plane and the axial plane is 26.4 ± 2.4 dB, 34.9 ± 5.2 dB, and 48.8 ± 15.7 dB, respectively. The values of the SSIM are also shown in table 2, with a mean value of 0.8448 ± 0.0212 for the images in the coronal plane, 0.9759 ± 0.0053 for the images in the sagittal plane, 0.9768 ± 0.0192 for the images in the axial plane. The differences in the NRMSE, PSNR and SSIM between the 3 planes can be understood due to the area of the image in which the features of the

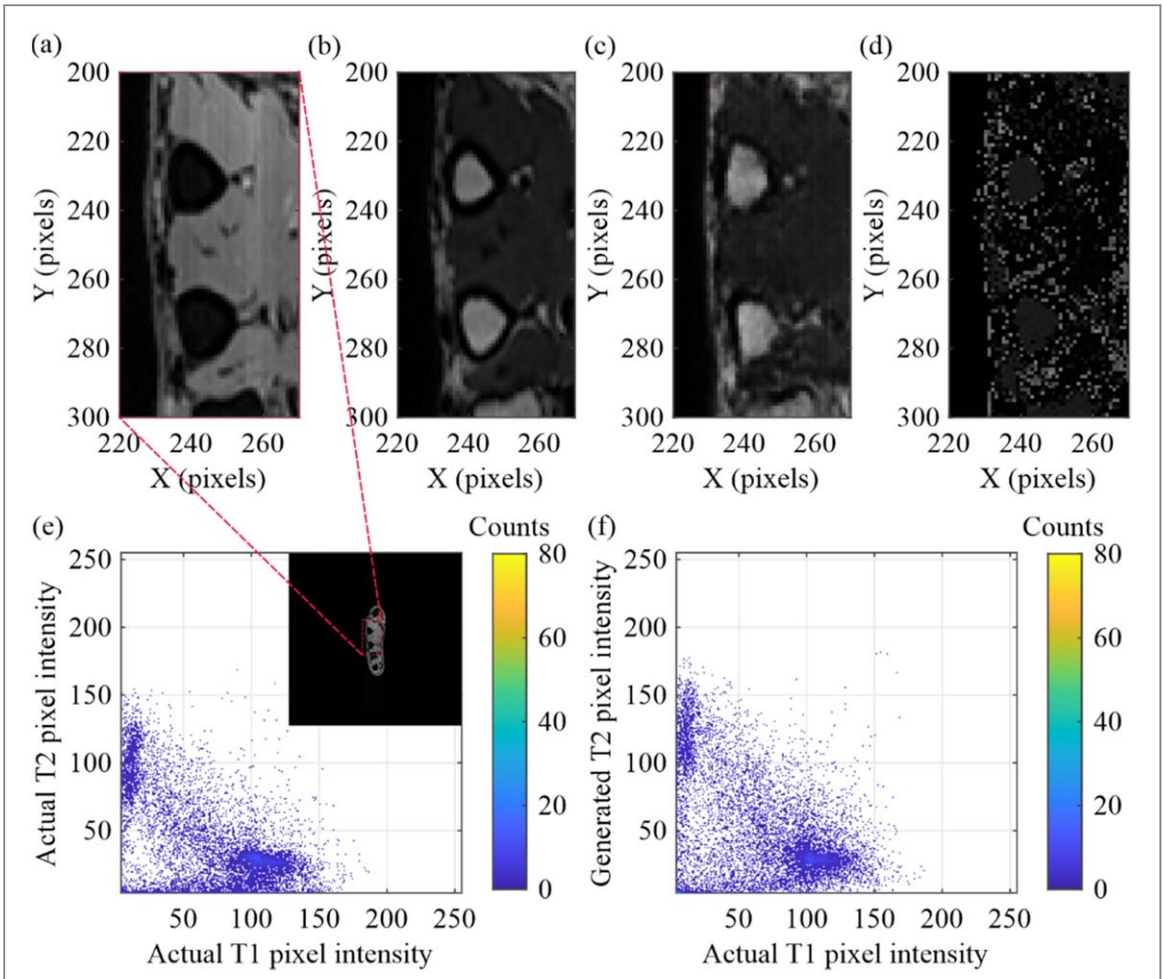

Figure 12. Signal intensity dependence analysis of the trained neural network for generating a T2 image of the wrist showing (a) T1 input image fed into the network, (b) actual image, (c) generated image, (d) colour-transformed image that uses the most probable data transformation, (e) transformation histogram of intensity for actual T1 and actual T2 pixel intensities values (0–255) and (f) transformation histogram of signal intensity for actual T1 and generated T2 pixel intensities values (0–255).

hand are present, with coronal plane containing more hand structure than the sagittal and the axial. It is important to realise that actual T2 and generated T2 images would not be expected to be identical, due to small movements in the patient's hand position and orientation between sequence measurements. In addition, disparities between the patient's left and right hands, such as different degrees of muscle mass or bone density, may also have resulted in an imperfect prediction. Additional training data could help reduce such an effect.

Since this work has only involved data from a single asymptomatic patient, in this section we focus on demonstrating a wide variety of analytical methods, which could be applied to larger studies and even point of care research. More specifically, we perform a range of digital pathology techniques, such as image cross-section analysis and sectional image generation to understand the resolution limitations of the generated images and the amount of information required from the input image for successful output image generation. On closer inspection of the images, it is clear that although there is similarity in the signal intensity and shape of the actual and generated images, as visualised in figure 10, small structures can be different in shape (see for example, labelled regions on images). It is also evident that the generated images contain a grid pattern, most likely a result of the discriminator neural network resolution. It is necessary for both faults to be corrected for future work if neural network transformations are to be applied to clinical practice. Further to this, on comparing the signal intensity of slices through the actual and generated images of the metacarpal bone of the thumb, as shown in figure 11, it is evident that the generated signal intensities are generally higher, and the peaks generally broader, as compared to the actual T2 images.

To allow a more detailed analysis of the neural network effectiveness, a small region from the axial plane is presented in figure 12, which shows (a) actual T1 image, (b) actual T2 image, and (c) generated T2 image. Since the training was carried out on the left hand and testing was carried out on the right hand, the network was not purely memorising features. Of interest, therefore, is the numerical relationship between the actual T1 image and actual T2 image, compared with the actual T1 image and generated T2 image. This analysis is shown in figures 12(e) and (f) respectively, which show scatter plots showing the signal intensity values of T1 image pixels

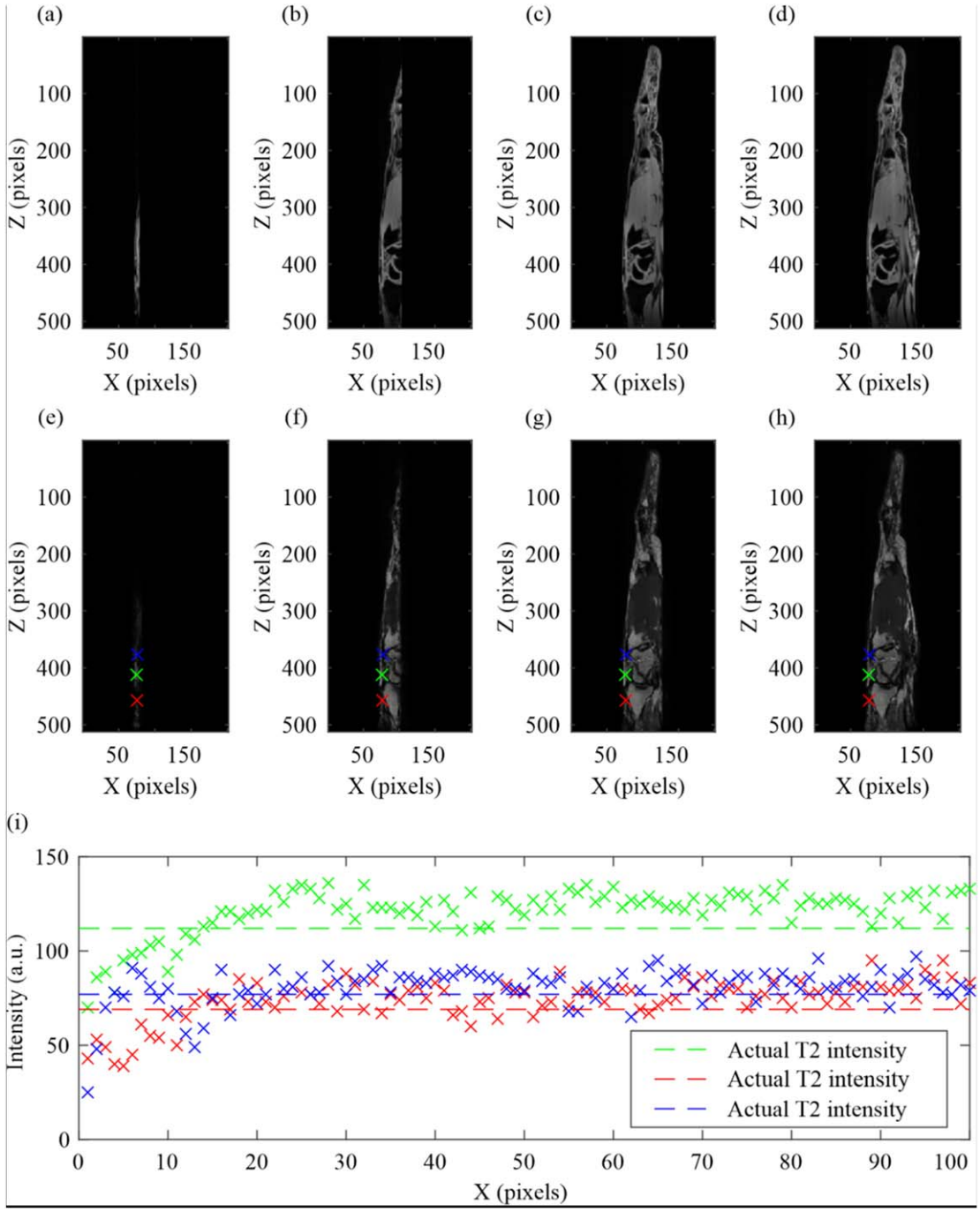

Figure 13. Adjacent pixel dependence analysis of the trained neural network for generating a T2 image of the right hand, showing input T1 image (top row) and generated T2 image (middle row), for ((a) and (e)) 5-pixel width, ((b) and (f)) 25-pixel width, ((c) and (g)) 50-pixel width, and ((d) and (h)) 75-pixel width window of a sagittal image slice, where the rest of the image outside of the window has been set to zero. (i) Pixel intensity for 3 different × points in the generated images (position shown in (e)) as a function of window width in pixels in X. Actual T2 signal intensity for these positions is indicated by the dashed line.

against the intensity values of the associated T2 image pixels. These scatter plots show that the transformation of intensity from T1 to T2 is a one-to-many mapping, but that the neural network has indeed, to a degree, replicated this transformation. The images imply that the neural network is processing the structure and intensity when predicting the associated T2 image, rather than simply changing the intensity of each image pixel. As a further comparison, figure 12(d) shows the result of computing T2 via a one-to-one transformation, formed by taking the most probable T2 image signal intensity value for each T1 image pixel from figure 12(e). The fact that this simple computation produces an output that bears almost zero resemblance to the actual T2 is further evidence that the neural network has indeed learned to recognise features within the images.

To quantify the sizes of features used by the neural network when making a T2 prediction, the effect of adjacent pixels on the generated T2 images was explored. This was achieved by analysing the predicted T2 output

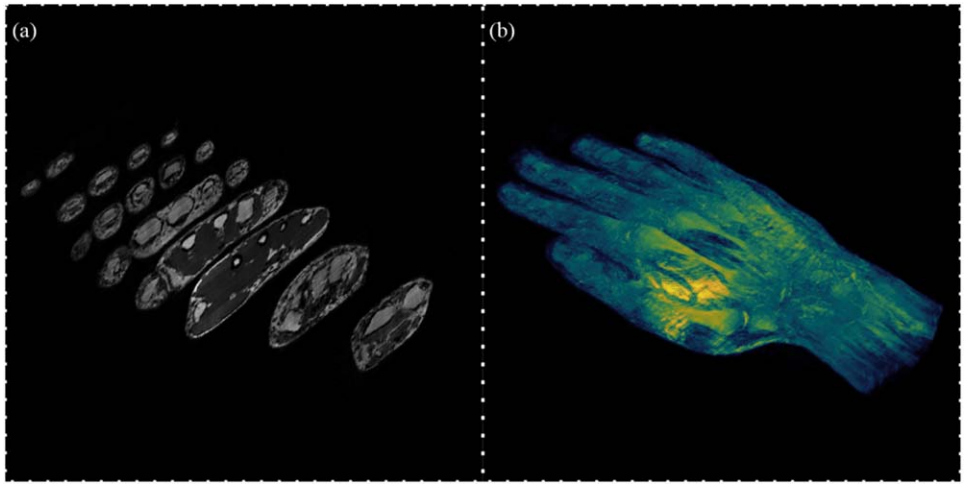

Figure 14. Generation of a 3D MRI T2 volume, generated from a neural network directly from T1 images, showing (a) eight generated T2 axial image slices along the hand and (b) a combination of all generated axial T2 image slices with a false colour map (blue indicates low signal intensity, yellow indicates high signal intensity) to enhance visual clarity.

whilst restricting the amount of information present in the T1 image, through setting signal intensity values of pixels in specific regions of the T1 image to zero. This method can be observed in figures 13(a)–(d), which shows the T1 image with varying amounts of information present, via setting the width of the observable window from 1-pixel to 100-pixels wide. Figures 13(e)–(h) show the associated neural network predictions for these restricted images, which clearly show that as the T1 image is uncovered, the generated T2 image similarly uncovers. Figure 13(i) shows the intensity values of three pixels (indicated by the red, green, and blue crosses in figure 13(e)), as a function of the width of the observable window. On the plot, the horizontal dashed lines show the intensity value at the same pixel positions on the actual T2 image. The fact that the intensity values on the generated T2 image do not match the actual value until approximately an observable window of approximately 20 pixels, implies that the neural network requires at least 20 pixels of information to make a strong prediction. Based on the scales in the images, 20 pixels corresponds to 1.01 cm, and hence this figure provides evidence that the neural network considers the structural information present in the surrounding $\sim$1 cm when transforming a T1 image into a T2 image.

Finally, as shown in figure 14, the neural network was applied to all T1 image slices in the axial plane to generate T2 images for the entire volume. Figure 14(a) shows eight axial planes of generated T2, whilst figure 14(b) shows the result of combining all axial images together to produce a complete and generated T2 image volume.

A richer data collection from a variety of individuals, which includes abnormalities, along with exploring open datasets, such as the multimodal Brain Tumour Image Segmentation Benchmark (BRATS) [82] and OpenfMRI [83], will enable rigorous investigation of such methods. More specifically, our previous work on Pix2Pix for accurate image generation of ($\sim$5-pixel resolution) has involved approximately 100 different shapes, and so increasing the dataset to include, not just different image planes but additional people/objects would be beneficial. In addition, we found that variation is as important as number. Indeed, we have often used 1000s of randomly created shapes in order to cover as many possible angles/sizes and variations as possible, and this is where phantoms could be useful. Since the area of the reconstruction was limited to $\sim$1 cm, it could be useful to train using phantoms of smaller than that size in order to potentially allow the neural network to generate accurate characterisations on a smaller scale. Since we found the surrounding structure to be important to the image generation, this could further aid in examination of any links between tissue boundaries and abnormalities, or other cues of surrounding areas. Moreover, perhaps moving to 3D data and using a 3D U-net architecture [84] could aid in more accurate characteristic generation if there is a reliance on surrounding structure for characteristic generation accuracy, since this would add in an extra dimensional constraint to the training and learning. This could be done along with examination of the convolutional filters and their activations, to allow for greater understanding of the transformation function and quantifying the neural networks capabilities. Determining the limitations could unlock methods for improving the image generation, either through additional data or via physics-guided networks, where one incorporates physics-based modelling [85]. This should ideally be done using 3D image volume data, and where possible similar to previous image segmentation approaches [86], in which one could add physics branches to the architecture that consists of the MRI physics parameters or physics-based static-equation sequence simulations to help the networks encode a

more accurate representation of the physics underlying MRI. Furthermore, one could use a CycleGAN [70] network to cycle between the physics-guided neural network generated images and the actual images, in which one varies the input parameters and thus the T1-weighted and the T2-weighted sequence images. This could perhaps create a neural network that has an understanding of the underlying MRI physics and produce more accurate characteristic generation.

Conclusion

To conclude, using data from a single asymptomatic patient, we have utilised neural networks to generate T2 SPACE sequence images from T1 VIBE images and subsequently performed a range of pathologies on the results to aid in understanding the capabilities of the deep learning methodology. Analysis showed that the Pix2Pix neural network was able to generate the similar signal through consideration of the surrounding $\sim$1 cm of image information. Specifically, regions of high signal intensity were correctly observed in a generated T2 image, such as between the hamate and the base of the 5th metacarpal. However, some signal intensities were 50% less than the actual and the resolution of some structures in the generated images was not as high as those in the actual. In addition, there were regions of artefactual grid patterning on the images. To resolve these issues and improve the transformation accuracy, a larger and more varied training dataset could be used, i.e. images from more than one person and area of the body. Additional work will be needed to prove that the neural network can accurately create T2 characteristics from T1 images, such as via synergising physics-based modelling with deep learning to understand and enhance the MRI technique. Also, other T2 sequences such as the STIR (short tau inversion recovery) sequence that nulls the signal from fat could be investigated, since for clinical diagnosis alternative weighted sequences would allow information on the characteristics involved, and thus assist in further quantification of the accuracy of the generated images. Furthermore, imaging could be carried out on parts of the body in which there are known abnormalities, such that they are not identified in one weighted type (e.g. T1-VIBE), but present in another (e.g. T2-STIR), to further evaluate the neural network. This initial work could be the basis for faster processing and diagnosis, and in turn, reduce the patient waiting times, in addition to being a useful mathematical tool for assisting in fundamental research of magnetic resonance imaging.

Acknowledgments

BM was supported by an EPSRC Early Career Fellowship (EP/N03368X/1) and EPSRC grant (EP/T026197/1).

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5258/SOTON/D1634.

Ethics

No ethical approval was required.

Conflicts of Interest

The authors declare no conflict of interest.

ORCID iDs

James A Grant-Jacob https://orcid.org/0000-0002-4270-4247
Robert W Eason https://orcid.org/0000-0001-9704-2204
Ben Mills https://orcid.org/0000-0002-1784-1012

References

[1] Frankel R I 1996 Centennial of Röntgen's discovery of x-rays *West. J. Med.* **164** 497–501
[2] Hoheisel M 2006 Review of medical imaging with emphasis on x-ray detectors Nucl. instruments methods *Phys. Res. Sect. A-Accelerators Spectrometers Detect. Assoc. Equip.* **563** 215–24
[3] Coppes-Zantinga A R and Coppes M J 1998 Marie Curie's contributions to radiology during world war I *Med. Pediatr. Oncol.* **31** 541–3

[4] Cnudde V and Boone M 2013 High-resolution x-ray computed tomography in geosciences: a review of the current technology and applications *Earth-Science Rev.* **123** 1–17
[5] Hounsfield G N 1979 Computer reconstructed x-ray imaging *Philos. Trans. R. Soc. A* **292** 223–32
[6] Schrevens L, Lorent N, Dooms C and Vansteenkiste J 2004 The role of PET scan in diagnosis, staging, and management of non-small cell lung cancer *Oncologist* **9** 633–43
[7] Wells P N T and Liang H-D 2011 Medical ultrasound: imaging of soft tissue strain and elasticity *J. R. Soc. Interface* **8** 1521–49
[8] Erasmus L J, Hurter D, Naude M, Kritzinger H G and Acho S 2004 A short overview of MRI artefacts *South African J. Radiol.* **8** a127
[9] Scheltens P H *et al* 1998 White matter changes on CT and MRI: an overview of visual rating scales *Eur. Neurol.* **39** 80–9
[10] Rua T, Malhotra B, Vijayanathan S, Hunter L, Peacock J, Shearer J, Goh V, McCrone P and Gidwani S 2019 Clinical and cost implications of using immediate MRI in the management of patients with a suspected scaphoid fracture and negative radiographs results from the SMaRT trial *Bone Joint J.* **101-B** 984–94
[11] Care, Department of Health and Social 2016 *Information on how NHS providers spent money to provide healthcare to patients NHS Reference Costs 2015-16* (https://gov.uk/government/publications/nhs-reference-costs-2015-to-2016)
[12] NHS England 2020 NHS Diagnostic Waiting Times and Activity Data Monthly Report 1–19 (https://england.nhs.uk/statistics/wp-content/uploads/sites/2/2016/06/DWTA-Report-November-2016.pdf)
[13] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A 2015 Going deeper with convolutions *Proc. of the IEEE conf. on Comput. Vision and pattern Recognit. (Boston, MA, 7–12 June 2015)* (Picastaway, NJ: IEEE) pp 1–9
[14] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436
[15] Serre T, Wolf L, Bileschi S, Riesenhuber M and Poggio T 2007 Robust object recognition with cortex-like mechanisms *IEEE Trans. Pattern Anal. Mach. Intell.* **29** 411–26
[16] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in neural information processing systems* **60** 1097–105
[17] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition arXiv: 1409.1556
[18] Qian Y, Bi M, Tan T and Yu K 2016 Very deep convolutional neural networks for noise robust speech recognition *IEEE/ACM Trans. Audio, Speech, Lang. Process.* **24** 2263–76
[19] Ruff Z J, Lesmeister D B, Duchac L S, Padmaraju B K and Sullivan C M 2019 Automated identification of avian vocalizations with deep convolutional neural networks *Remote Sens Ecol and Conserv* **6** 79–92
[20] Joly A, Goëau H, Glotin H, Spampinato C, Bonnet P, Vellinga W-P, Champ J, Planqué R, Palazzo S and Müller H 2016 LifeCLEF 2016: multimedia life species identification challenges *International Conference of the Cross-Language Evaluation Forum for European Languages* vol 9822 (Berlin: Springer) pp 286–310
[21] Mollahosseini A, Chan D and Mahoor M H 2016 Going deeper in facial expression recognition using deep neural networks *IEEE winter conference on applications of computer vision (WACV) (Lake Placid, NY, 7–10 March 2016)* (Picastway, NJ: IEEE) pp 1–10
[22] Ranjan R, Sankaranarayanan S, Bansal A, Bodla N, Chen J-C, Patel V M, Castillo C D and Chellappa R 2018 Deep learning for understanding faces: machines may be just as good, or better, than humans *IEEE Signal Process Mag.* **35** 66–83
[23] Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R and Fei-Fei L 2014 Large-scale video classification with convolutional neural networks *IEEE Conf on Comput Vision and Pattern Recognit (Columbus, OH, 23–28 June 2014)* (Picastaway, NJ: IEEE) pp 1725–32
[24] Grant-Jacob J A, Mackay B S, Baker J A G, Heath D J, Xie Y, Loxham M, Eason R W and Mills B 2018 Real-time particle pollution sensing using machine learning *Opt. Express* **26** 27237–46
[25] Grant-Jacob J A, Jain S, Xie Y, Mackay B S, McDonnell M D T, Praeger M, Loxham M, Richardson D J, Eason R W and Mills B 2019 Fibre-optic based particle sensing via deep learning *J. Phys.: Photonics* **1** 44004
[26] Grant-Jacob J A, Xie Y, Mackay B S, Praeger M, McDonnell M D T, Heath D J, Loxham M, Eason R W and Mills B 2019 Particle and salinity sensing for the marine environment via deep learning using a Raspberry Pi *Environ. Res. Commun.* **1** 035001
[27] Mills B, Heath D J, Grant-Jacob J A and Eason R W 2018 Predictive capabilities for laser machining via a neural network *Opt. Express* **26** 17245–53
[28] Heath D J, Grant-Jacob J A, Xie Y, Mackay B S, Baker J A G, Eason R W and Mills B 2018 Machine learning for 3D simulated visualization of laser machining *Opt. Express* **26** 4984–8
[29] Mills B, Heath D J, Grant-Jacob J A, Xie Y and Eason R W 2018 Image-based monitoring of femtosecond laser machining via a neural network *J. Phys.: Photonics* **1** 15008
[30] Liu Y-C, Chiu W-C, Wang S-D and Wang Y-C F 2017 Domain-Adaptive generative adversarial networks for sketch-to-photo inversion *IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP) (Tokyo, Japan, 25–28 September 2017)* (Picastaway, NJ: IEEE) pp 1–6
[31] Poterek Q, Herrault P-A, Skupinski G and Sheeren D 2020 Deep learning for automatic colorization of legacy grayscale aerial photographs *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **13** 2899–915
[32] Rivenson Y, Göröcs Z, Günaydin H, Zhang Y, Wang H and Ozcan A 2017 Deep learning microscopy *Optica* **4** 1437–43
[33] Grant-Jacob J A, Mackay B S, Baker J A G, Xie Y, Heath D J, Loxham M, Eason R W and Mills B 2019 A neural lens for super-resolution biological imaging *J. Phys. Commun.* **3** 065004
[34] Grant-Jacob J A, Praeger M, Loxham M, Eason R W and Mills B 2020 Lensless imaging of pollen grains at three-wavelengths using deep learning *Environ. Res. Commun.* **2** 075005
[35] Suzuki K 2017 Overview of deep learning in medical imaging *Radiol. Phys. Technol.* **10** 257–73
[36] Lundervold A S and Lundervold A 2019 An overview of deep learning in medical imaging focusing on MRI *Z. Med. Phys.* **29** 102–27
[37] Greenspan H, Van Ginneken B and Summers R M 2016 Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique *IEEE Trans. Med. Imaging* **35** 1153–9
[38] Milletari F, Navab N and Ahmadi S-A 2016 V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation *IEEE Fourth International Conference on 3D Vision (3DV) (Stanford, CA, 25–28 October 2016)* (Piastaway, NJ: IEEE) pp 565–71
[39] Gibson E *et al* 2018 NiftyNet: a deep-learning platform for medical imaging *Comput. Methods Programs Biomed.* **158** 113–22
[40] Baltruschat I M, Nickisch H, Grass M, Knopp T and Saalbach A 2019 Comparison of deep learning approaches for multi-label chest x-ray classification *Sci. Rep.* **9** 6381
[41] Liu S, Wang Y, Yang X, Lei B, Liu L, Li S X, Ni D and Wang T 2019 Deep learning in medical ultrasound *Analysis: A Review Engineering* **5** 261–75
[42] Wang Y, Yu B, Wang L, Zu C, Lalush D S, Lin W, Wu X, Zhou J, Shen D and Zhou L 2018 3D conditional generative adversarial networks for high-quality PET image estimation at low dose *Neuroimage* **174** 550–62

[43] Liu F, Jang H, Kijowski R, Bradshaw T and McMillan A B 2017 Deep learning MR imaging–based attenuation correction for PET/MR Imaging *Radiology* **286** 676
[44] Sahiner B, Pezeshk A, Hadjiiski L M, Wang X, Drukker K, Cha K H, Summers R M and Giger M L 2019 Deep learning in medical imaging and radiation therapy *Med. Phys.* **46** e1–e36
[45] Hatvani J, Horváth A, Michetti J, Basarab A, Kouamé D and Gyöngy M 2019 Deep learning-based super-resolution applied to dental computed tomography IEEE Trans. Radiat. Plasma *Med. Sci.* **3** 120–8
[46] Spampinato C, Palazzo S, Giordano D, Aldinucci M and Leonardi R 2017 Deep Learning for automated skeletal bone age assessment in x-ray images *Med. Image Anal.* **36** 41–51
[47] Bhandary A, Prabhu G A, Rajinikanth V, Thanaraj K P, Satapathy S C, Robbins D E, Shasky C, Zhang Y-D, Tavares J M R S and Raja N S M 2020 Deep-learning framework to detect lung abnormality - a study with chest x-ray and lung CT scan images *Pattern Recognit. Lett.* **129** 271–8
[48] Roth H R, Farag A, Lu L, Turkbey E B and Summers R M 2015 Deep convolutional networks for pancreas segmentation in CT imaging *Medical Imaging 2015: Image Processing* **9413** 94131G
[49] Han Y S, Yoo J and Ye J C 2016 Deep Residual Learning for Compressed Sensing CT Reconstruction via Persistent Homology Analysis arXiv: 1611.06391
[50] Yi X, Walia E and Babyn P 2019 Generative adversarial network in medical imaging: A review *Med. Image Anal.* **58** 101552
[51] Meharban M S *et al* 2021 Introduction to Medical Image Synthesis Using Deep Learning: A Review *7th International Conference on Advanced Computing and Communication Systems (ICACCS)* vol 1 *(Coimbatore, India, 19–20 March 2021)* (Picastaway, NJ: IEEE) pp 414–9
[52] Kaji S and Kida S 2019 Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging *Radiol. Phys. Technol.* **12** 235–48
[53] Singh N K and Raza K 2021 *Medical Image Generation Using Generative Adversarial Networks: A Review Health Informatics: A Computational Perspective in Healthcare* ed R Patgiri *et al* vol 932 (Berlin: Springer) pp 77–96
[54] Wolterink J M, Dinkla A M, Savenije M H F, Seevinck P R, van den Berg C A T and Išgum I 2017 Deep MR to CT synthesis using unpaired data *Int. Work. Simul. Synth. Med. Imaging* (Vol 10557) (Berlin: Springer) pp 14–23
[55] Han X 2017 MR-based synthetic CT generation using a deep convolutional neural network method *Med. Phys.* **44** 1408–19
[56] Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, Feng D and Liang D 2016 Accelerating magnetic resonance imaging via deep learning *2016 IEEE 13th International Symp. on Biom. Imaging (ISBI) (Prague, Czech Republic, 13–16 April 2016)* (Picastaway, NJ: IEEE) pp 514–7
[57] Shao J, Ghodrati V, Nguyen K and Hu P 2020 Fast and accurate calculation of myocardial T1 and T2 values using deep learning Bloch equation simulations (DeepBLESS) *Magn. Reson. Med.* **84** 2831–45
[58] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K H, Bendszus M and Biller A 2016 Deep MRI brain extraction: a 3D convolutional neural network for skull stripping *Neuroimage* **129** 460–9
[59] Chaudhari A S, Fang Z, Kogan F, Wood J, Stevens K J, Gibbons E K, Lee J H, Gold G E and Hargreaves B A 2018 Super-resolution musculoskeletal MRI using deep learning *Magn. Reson. Med.* **80** 2139–54
[60] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine J L, Duerk J L and Griswold M A 2013 Magnetic resonance fingerprinting *Nature* **495** 187–92
[61] Cohen O, Zhu B and Rosen M S 2018 MR fingerprinting deep reconstruction network (DRONE) *Magn. Reson. Med.* **80** 885–94
[62] Jog A, Carass A, Roy S, Pham D L and Prince J L 2017 Random forest regression for magnetic resonance image synthesis *Med. Image Anal.* **35** 475–88
[63] Xiang L, Chen Y, Chang W, Zhan Y, Lin W, Wang Q and Shen D 2018 Ultra-Fast T2-Weighted MR Reconstruction Using Complementary T1-Weighted Information *Med. Image Comput. Comput. Interv.* **11070** 215–23
[64] Isola P, Zhu J-Y, Zhou T and Efros A A 2017 Image-to-image translation with conditional adversarial networks arXiv:1611.07004
[65] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional Networks for Biomedical Image Segmentation *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (*Lecture Notes in Computer Science* Vol 9351) ed N Navab *et al* (Berlin: Springer) (https://doi.org/10.1007/978-3-319-24574-4_28)
[66] Li C and Wand M 2016 Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks *Computer Vision – ECCV 2016* (*Lecture Notes in Computer Science* Vol 9907) (Berlin: Springer) pp 702–16
[67] Olut S, Sahin Y H, Demir U and Unal G 2018 Generative adversarial training for MRA image synthesis using multi-contrast MRI *MEdicine. PRIME 2018* (*Lecture Notes in Computer Science* vol 11121) ed I Rekik *et al* (Berlin: Springer) pp 147-154
[68] Zhou T, Fu H, Chen G, Shen J and Shao L 2020 Hi-net: hybrid-fusion network for multi-modal MR image synthesis *IEEE Trans. Med. Imaging* **39** 2772–81
[69] Shin H-C, Tenenholtz N A, Rogers J K, Schwarz C G, Senjem M L, Gunter J L, Andriole K P and Michalski M 2018 Medical image synthesis for data augmentation and anonymization using generative adversarial networks *International Workshop on Simulation and Synthesis in Medical Imaging* (*Lecture Notes in Computer Science* vol 11037) (Berlin: Springer) pp 1–11
[70] Zhu J-Y, Park T, Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks *2017 IEEE International Conference on Computer Vision (ICCV) (Venice, Italy, 22–29 Oct. 2017)* (Picastaway, NJ: IEEE) pp 2242–51
[71] Gu X, Knutsson H, Nilsson M and Eklund A 2019 Generating diffusion MRI scalar maps from T1 weighted images using generative adversarial *Networks Image Analysis* (Vol 11482) ed M Felsberg *et al* (Berlin: Springer) pp 489–98
[72] Oulbacha R and Kadoury S 2020 MRI to CT synthesis of the lumbar spine from a pseudo-3D cycle GAN *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (Iowa City, IA, 3–7 April 2020)* (Picastaway, NJ: IEEE) pp 1784–7
[73] Wolterink J M, Dinkla A M, Savenije M H F, Seevinck P R, van den Berg C A T and Išgum I 2017 Deep MR to CT synthesis using unpaired data BT - Simulation and synthesis in medical imaging ed S A Tsaftaris *et al* (Cham: Springer International Publishing) pp 14–23
[74] Liu M-Y, Breuel T M and Kautz J 2017 Unsupervised image-to-image translation networks *Advances in Neural Information Processing Systems 30 (NIPS 2017) (Long Beach, CA)* pp 700–8 (https://proceedings.neurips.cc/paper/2017/hash/dc6a6489640ca02b0d42dabeb8e46bb7-Abstract.html)
[75] Cheng D, Qiu N, Zhao F, Mao Y and Li C 2021 Research on the modality transfer method of brain imaging based on generative adversarial network *Front. Neurosci.* **15** 255
[76] Zhu Y, Tang Y, Tang Y, Elton D C, Lee S, Pickhardt P J and Summers R M 2020 *Cross-domain Medical Image Translation by Shared Latent Gaussian Mixture Model Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* ed A L Martel *et al* (Cham: Springer International Publishing) pp 379–89

[77] Debs N, Rasti P, Victor L, Cho T-H, Frindel C and Rousseau D 2020 Simulated perfusion MRI data to boost training of convolutional neural networks for lesion fate prediction in acute stroke *Comput. Biol. Med.* **116** 103579
[78] Turner J T, Page A, Mohsenin T and Oates T 2014 Deep belief networks used on high resolution multichannel electroencephalography data for seizure detection 2014 arXiv:1708.08430 AAAI Spring Symposium Series
[79] Kaplan S and Zhu Y-M 2019 Full-dose PET image estimation from low-dose PET image using deep learning: a pilot study *J. Digit. Imaging* **32** 773–8
[80] Heusel M, Ramsauer H, Unterthiner T, Nessler B and Hochreiter S 2017 Gans trained by a two time-scale update rule converge to a local nash equilibrium (arXiv:1706.08500)
[81] Maas A L, Hannun A Y and Ng A Y 2013 Rectifier nonlinearities improve neural network acoustic models *Proc. International Conference on Machine Learning* 30, pp 3
[82] Menze B H *et al* 2015 The multimodal brain tumor image segmentation benchmark (BRATS) *IEEE Trans. Med. Imaging* **34** 1993–2024
[83] Poldrack R A and Gorgolewski K J 2017 OpenfMRI: open sharing of task fMRI data *Neuroimage* **144** 259–61
[84] Çiçek Ö, Abdulkadir A, Lienkamp S S, Brox T and Ronneberger O 2016 *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* ed S Ourselin *et al* vol 9901 (Cham: Springer International Publishing) pp 424–32
[85] Raissi M, Perdikaris P and Karniadakis G E 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations *J. Comput. Phys.* **378** 686–707
[86] Borges P, Sudre C, Varsavsky T, Thomas D, Drobnjak I, Ourselin S and Cardoso M J 2019 *Physics-Informed Brain MRI Segmentation Simulation and Synthesis in Medical Imaging* ed N Burgos *et al* vol 11827 (Cham: Springer International Publishing) pp 100–9