

ACCEPTED MANUSCRIPT • OPEN ACCESS

TIST-Net: style transfer in dynamic contrast enhanced MRI using spatial and temporal information

To cite this article before publication: Adam George Tattersall *et al* 2024 *Phys. Med. Biol.* in press <https://doi.org/10.1088/1361-6560/ad4193>

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is © 2024 The Author(s). Published on behalf of Institute of Physics and Engineering in Medicine by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 4.0 licence, this Accepted Manuscript is available for reuse under a CC BY 4.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by/4.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

TIST-Net: Style transfer in dynamic contrast enhanced MRI using spatial and temporal information

Adam G. Tattersall^{1,2}, Keith A. Goatman², Lucy E. Kershaw¹,
Scott I. K. Semple¹, Sonia Dahdouh²

¹University of Edinburgh, Edinburgh, UK

²Canon Medical Research Europe, Edinburgh, UK

E-mail: adam.tattersall@ed.ac.uk

March 2024

Abstract.

Objective Training deep learning models for image registration or segmentation of dynamic contrast enhanced (DCE) MRI data is challenging. This is mainly due to the wide variations in contrast enhancement within and between patients. To train a model effectively, a large dataset is needed, but acquiring it is expensive and time consuming. Instead, style transfer can be used to generate new images from existing images.

In this study, our objective is to develop a style transfer method that incorporates spatio-temporal information to either add or remove contrast enhancement from an existing image.

Approach We propose a Temporal Image-to-Image Style Transfer Network (TIST-Net), consisting of an auto-encoder combined with convolutional long short-term memory (LSTM) networks. This enables disentanglement of the content and style latent spaces of the time series data, using spatio-temporal information to learn and predict key structures. To generate new images, we use deformable and adaptive convolutions which allow fine grained control over the combination of the content and style latent spaces. We evaluate our method, using popular metrics and a previously proposed contrast weighted structural similarity index measure (CW-SSIM). We also perform a clinical evaluation, where experts are asked to rank images generated by multiple methods.

Main Results Our model achieves state-of-the-art performance on three datasets (kidney, prostate and uterus) achieving an SSIM of 0.91 ± 0.03 , 0.73 ± 0.04 , 0.88 ± 0.04 respectively when performing style transfer between a non-enhanced image and a contrast-enhanced image. Similarly, SSIM results for style transfer from a contrast-enhanced image to a non-enhanced image were 0.89 ± 0.03 , 0.82 ± 0.03 , 0.87 ± 0.03 . In the clinical evaluation, our method was ranked consistently higher than other approaches.

Significance TIST-Net can be used to generate new DCE-MRI data from existing images. In future, this may improve models for tasks such as image registration or segmentation by allowing small training datasets to be expanded.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

Keywords: Style Transfer, Spatio-temporal Information, Content/Style Disentanglement, Dynamic Contrast Enhanced (DCE) - MRI

1. Introduction

Deep learning models are increasingly popular due to their impressive performance across multiple image processing tasks such as image segmentation, object detection and classification. There have been continuous improvements made due to advancements in architecture, computational performance and access to larger datasets. Having access to large datasets is beneficial as it allows training with images that are similar to images seen at inference, as well as avoiding any overfitting. However, medical imaging datasets are often small with little variability compared to natural image datasets such as ImageNet (Deng et al. 2009). This can make it difficult to train robust models. For medical image processing tasks, deep learning consistently achieves state-of-the-art performance, usually using augmentation to provide the models with additional varying data.

Datasets containing dynamic contrast enhanced (DCE) - MRI are typically small with a limited number of annotations outlining organs of interest. DCE-MRI is a type of quantitative imaging used to monitor microvascular perfusion (Ingrisch and Sourbron 2013). Multiple T1 weighted images are rapidly taken over a few minutes along with a contrast agent injection, which causes a rapid increase in intensity. DCE-MRI data contains motion and contrast enhancement which is tissue dependent.

Quantitative analysis of DCE-MRI is often performed, which usually requires the manual selection of voxels in each frame that represent the tissue of interest. However, the selected voxels in a single frame may not represent the target tissue over the temporal dimension due to motion between frames. Obtaining manually annotated images is time consuming and difficult, hence fully annotated DCE-MRI datasets are scarce. This reduces the diversity of available data to train segmentation and registration models.

When training a model using DCE-MRI data, it was found that strategically using data from the whole sequence with varying contrast enhancements led to better performing models (Tattersall et al. 2023a). However, the datasets are rarely annotated and lack diversity. Augmentation can be used to increase the size and diversity of the training dataset which can play a key role in training a robust model (Shorten and Khoshgoftaar 2019). By artificially expanding the dataset through augmentation, models are more capable of handling real-world scenarios where data can differ from the training set. Augmentation acts as a form of regularisation, effectively preventing overfitting and enhancing the model’s ability to generalise well to unseen examples (Shorten and Khoshgoftaar 2019). However, common augmentation techniques may not be enough to provide a model trained using DCE-MRI data with a diverse enough set of examples during training. This is due to the lack of availability of annotations for non contrast enhanced (CE) images, and the widely varying levels of contrast enhancement in the images. Instead, an approach which can generate new images by adjusting the contrast enhancement effect would enable a diverse dataset to be created. The generated image would remain structurally identical whilst the contrast enhancement in the image would be changed to ensure that any available ground truth can still be utilised. Such

TIST-Net: Style transfer in dynamic contrast enhanced MRI

a method would allow for robust models to be trained using smaller datasets enlarged using synthetic images.

Style transfer is a technique that aims to combine a content and a style image resulting in an image which has a combination of characteristics from both images. The content image provides the underlying structure and arrangement of objects, whilst the style image contributes the artistic patterns, colours, and textures. Within medical imaging, style transfer has been used to generate new CT images from MRI or MRI from CT (Yang et al. 2019; Reaungamornrat et al. 2022). It has also been used to improve image quality by denoising low dose CT images by translating them into a high dose counterpart (Wolterink et al. 2017). Within the scope of DCE-MRI, the following constraints need to be met. (1) The structure of the content image needs to be preserved. This allows for any existing annotations to be used. (2) The change in intensities needs to be localised and tissue dependent. (3) The characteristics of the image need to be preserved such as the noise expected in MRI.

In this work, we proposed a model which used auto-encoders with convolutional long short-term memory (LSTM) networks (Chao et al. 2018) to learn spatio-temporal information. Adaptive convolutions (AdaConv) were used to combine content and style latent spaces alongside deformable convolutions to allow the model to adapt the receptive field to account for local geometric variations. This gave the model the ability to decide which areas of the image should have specific style attributes.

Quantitatively evaluating style transfer can be difficult with unregistered data. Current metrics such as peak signal to noise ratio (PSNR), structural similarity index measure (SSIM) (Zhou Wang et al. 2004) and multiscale (MS)-SSIM (Z. Wang, Simoncelli, and Bovik 2003) primarily focus on the pixel-level similarity between a pair of images. However, in DCE-MRI data, there is a large variety of contrast enhancement which is tissue dependent. This makes evaluating synthetic DCE-MRI a difficult task. A contrast-weighted (CW)-SSIM (Tattersall et al. 2023b) was previously proposed which separated the measurement of overall content and localised style. In this work, we also explored and validated this metric to measure its effectiveness for evaluating generated images.

We compared our model with other style transfer models using standard metrics (PSNR, SSIM and MS-SSIM), CW-SSIM as well as completing a rigorous clinical evaluation of our generated images to correlate the quantitative analysis with experts' opinions.

2. Related Work

Using neural networks for style transfer was first proposed by Gatys, Ecker, and Bethge 2016 and has been successfully built upon. Previously, it was a slow iterative process with low quality style transfer, but the use of feed-forward networks (Ulyanov et al. 2016) to speed up the process, as well as creating perceptual losses (Johnson, Alahi, and Fei-Fei 2016) to give more visually pleasing results has led to new impressive models that have generated realistic images (Karras et al. 2021).

2.1. Approaches to Style Transfer

There have been multiple style transfer approaches to generate new images such as learning mappings between images or content/style disentanglement. Pix2Pix (Isola et al. 2017) and CycleGAN (Zhu et al. 2017) are methods that learn mappings between images. Pix2Pix used a conditional GAN which consists of a generator that maps an input image to a desired output and a discriminator to classify between real and fake images. Pix2Pix has been used to generate new medical images such as CT pelvis images from MRI (Maspero et al. 2018). Although high quality images can be generated using Pix2Pix, it requires paired and registered images to train effectively. To circumvent this, Zhu et al. 2017 proposed CycleGAN which can be trained using unpaired data. In this approach, generators learn mappings between two domains along with discriminators to classify between real and fake images. A cycle consistency loss is also used to ensure that translated images can be reversed. A disadvantage to both of these approaches is that they can struggle to generate diverse outputs when new style images are introduced at inference time as the model learns to map between domains directly. Galli et al. 2023 proposed a method using a CycleGAN architecture which translated the appearance of DCE-MRI breast data between two datasets in an attempt to increase the size of available datasets to improve lesion classification.

Alternatively, content/style disentanglement methods such as MUNIT (Huang, Liu, et al. 2018) have been proposed. These methods aim to predict two latent spaces which describe the content and style of the image. This type of method can allow for better control of the translation process and the preservation of the content image whilst transferring style. Content/style disentanglement methods usually use an auto-encoder based architecture. Encoders typically encode content and style latent spaces whilst the decoders combine the latent spaces to generate an image. Discriminators can also be used for adversarial training. Lee et al. 2020 proposed DRIT++ which is an extension of MUNIT. DRIT++ aimed to enhance diversity by introducing disentanglement at the domain level, and improving attribute manipulation. A key issue with these approaches is the large computational requirement. Content/style disentanglement methods have been used with DCE-MRI data such as the method proposed by Cai et al. 2023 who used content and style decoders along with a mapping network to increase style diversity.

There has also been some work using style transfer with videos. Early work processed frames independently however, it was found that this created videos that

TIST-Net: Style transfer in dynamic contrast enhanced MRI

6

flicker and produced false discontinuities (D. Chen et al. 2017). To alleviate this, D. Chen et al. 2017 proposed a method to ensure temporal consistency between image frames. Their method uses three components: a style sub-network, a flow sub-network and a mask sub-network. The style sub-network is an auto-encoder which performs the style transfer. The flow sub-network estimates the correspondence between consecutive image frames and warps the features and finally, the mask sub-network that regresses a mask to features in adjacent time frames. This ensures that the features of objects in each image that are similar can be reused. Using this approach allows a smooth style transfer between image frames, but it is prone to errors when there is large motion between frames as well as propagating errors over time leading to inconsistent style transfer and blurriness.

2.2. Injecting Style

Style transfer methods aim to transform images from one domain to another whilst preserving content. However, they often fail to generate diverse and realistic images. To achieve this, methods which inject style into the content image has been proposed as an effective solution as it helps preserve the structure of the generated image whilst creating an image that looks similar to the style image.

A popular method for injecting style is Adaptive Instance Normalization (AdaIN) proposed by Huang and Belongie 2017. As shown in Figure 1a, AdaIN normalises the activations in a neural network based on the statistics (mean and standard deviation) of the style image. This allows the network to transfer the style of the style image to the content image by matching their statistical properties, but this only uses global style information. In the case of contrast transfer for DCE-MRI data, CE is variable between the tissues, however, AdaIN adds CE to the whole image, making it unsuitable for this task as the contrast enhancement is localised to specific tissues.

Spatially-Adaptive Normalization for Generative Networks (SPADE), proposed by Park et al. 2019 is a method that builds on AdaIN and introduces spatially-adaptive normalisation by incorporating semantic information for more controlled and semantically consistent style transfer (Figure 1b). Usually, a semantic segmentation map is used to predict the normalisation parameters for each pixel location. This improves the control over the style generation process.

Another approach to injecting style into the generator network is through an Adaptive Convolution (AdaConv) (Chandran et al. 2021). In AdaConv, the convolution filters are learned dynamically from the style latent code to create a set of convolution filters which are then applied to the content image (Figure 1c). While this comes at a cost of increased computation, AdaConv captures global and local information to predict parameters to combine the style and content latent spaces.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

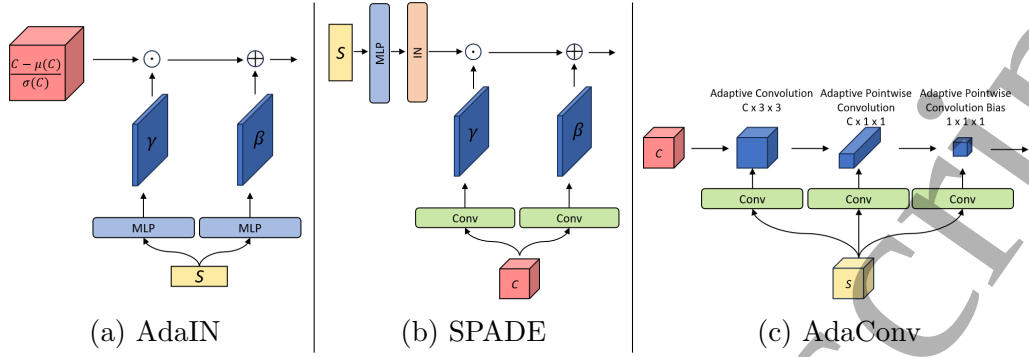


Figure 1: Methods to combine content (C) and style (S): (a) AdaIN, (b) SPADE and (c) AdaConv. MLP, Conv and IN denote multilayer perceptron, convolutional and instance normalisation layers, respectively. μ represents the mean, σ represents the standard deviation, γ is a learnable scaling parameter and β is a learnable bias parameter.

3. Materials and Methods

3.1. Data

Three DCE-MRI datasets were used. The first contains 2D kidney DCE-MRI (Lietzmann et al. 2012) from 13 patients with 375 images acquired continuously at a temporal resolution of 1.6 s, a spatial resolution of 384x348 and pixel sizes 1.08x1.08 mm. The second contains 3D prostate DCE-MRI (Lemaître et al. 2015) from 20 patients with 40 volumes acquired continuously at a temporal resolution of 6 s, a spatial resolution of 256x192x16 and voxel sizes 1.12x1.12x3.5 mm. The third dataset contains 3D uterus DCE-MRI (Reavey et al. 2021) from 36 patients with 150 volumes acquired continuously at a temporal resolution of 2.45 s, a spatial resolution of 192x192x30 and voxel sizes 2.08x2.08x4mm. Each dataset was split at a patient level with an 80:20 ratio, training was done using five-fold cross validation.

3.2. Style Transfer - Global Architecture

As in Tattersall et al. 2023b), we used a structure composed of encoders E , decoders D and discriminators Dis . A sequence of images was passed into encoders to predict content and style latent spaces. Successful disentanglement led to content latent spaces containing information representing the structures of the image and style latent spaces containing information representing the modality (MRI) and any contrast enhancements. The content latent spaces were then passed into a bi-directional (Graves and Schmidhuber 2005) convolutional LSTM (Figure 2) which allowed for the modelling of temporal information from a sequence of images. The convolutional LSTM used spatial invariance by extracting relevant features from the images regardless of their spatial position. By using a bi-directional LSTM, the model was able to learn information from past and future events.

To combine content and style, Tattersall et al. 2023b used AdaConv to convolve a kernel (adaptive convolution), pointwise kernel (adaptive pointwise convolution) and bias predicted from a style latent space over a content latent space. This allowed for local, spatial information to be used when combining content and latent spaces. However, we found that the generated images exhibited smoothening effects, especially in areas of contrast enhancement. To improve this, we proposed TIST-Net (Temporal Image-to-image Style Transfer) which used deformable convolutions proposed by Dai et al. 2017 to offset the adaptive convolution. The offset is simply predicted by passing a feature map, in our case the content latent space, through a convolutional layer. This contained information to decide how much a kernel should be deformed. By using this, it allows for the receptive field of a convolutional kernel to adapt and account for local geometric variations in the input data. Figure 3 shows the role of deformable convolutions with AdaConv. By offsetting the kernel, the model decided which areas of the content image should gain particular style attributes which is necessary when generating images with contrast enhancement at various time points.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

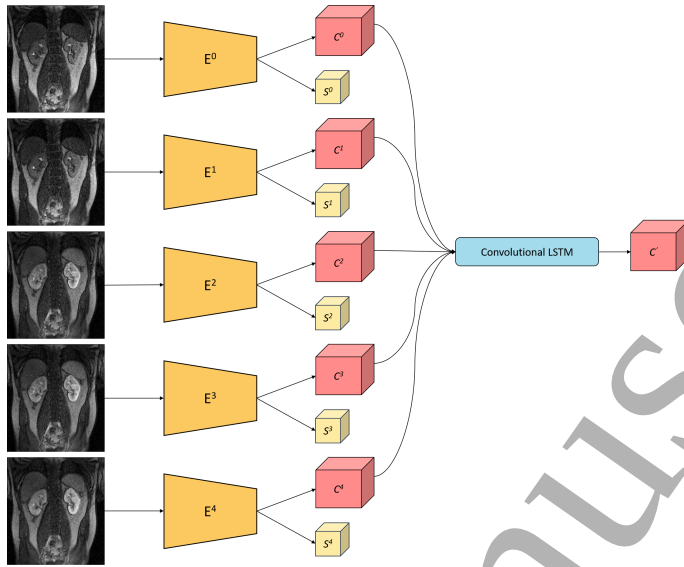


Figure 2: Encoding of images which FL were FL passed into encoders E^i to encode content C^i and style S^i , where i is between 0 and 4. Content spaces FL were FL passed into a convolutional LSTM as a sequence. We then output a new predicted content latent space C'

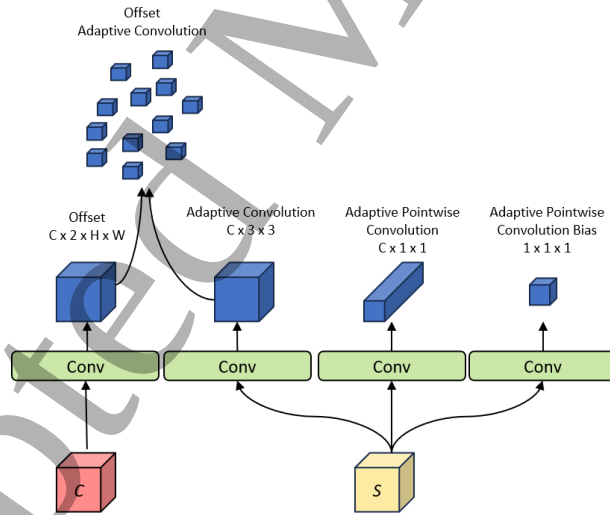


Figure 3: A style latent space is passed through three convolutional layers. An adaptive convolution, adaptive pointwise convolution, and adaptive pointwise convolution bias is predicted from each of the layers. An offset is also predicted from a content latent space (of size channel (C), height (H), Width (W)). The offset is then used to deform the adaptive convolution. These predictions are used to combine the content and style latent spaces. Note, this figure shows an example for 2D style transfer. For 3D an additional dimension is used for depth.

We show the global architecture for our method in Figure 4. Discriminators were used to predict if the generated image was real or fake. We also highlight the losses

used to ensure disentanglement.

The L1 loss was used as the cycle consistency loss between the original and reconstructed images and each of the initially predicted and reconstructed latent spaces. This ensured that important information such as structure was not lost during the reconstruction and translation process. This was calculated between the original and reconstructed images as well as between the initially predicted and reconstructed latent spaces. The mean square error (MSE) loss was used as an adversarial loss to predict if an image was real or generated. This helped the model to generate images that were realistic. We also used two perceptual losses proposed by Johnson, Alahi, and Fei-Fei 2016. The content, style and generated images were passed through a pretrained model P to predict feature maps. These feature maps were used as inputs to each of the losses. The first was a feature reconstruction loss which calculated the squared, normalised Euclidean distance (Eq. 1) between the feature maps of the content c and generated g image. f is the dimension of the feature map. This loss encouraged the style transfer model to generate images with similar structural features. The second was a style reconstruction loss which penalised the differences in style between the feature maps of the style s and generated image. The squared Frobenius norm of the difference between the Gram matrices of the input feature maps was used (Eq. 2). Minimising this loss encouraged the model to generate images which had similar style patterns and texture to the style image.

$$L_{feature} = \frac{1}{f} \|P(g) - P(c)\|_2^2 \quad (1)$$

$$L_{Frob} = \|Gram(P(g)) - Gram(P(s))\|_F^2 \quad (2)$$

For the 2D image generation, we used a VGG-16 network pretrained on ImageNet (Russakovsky et al. 2015) which was converted to grayscale. For the 3D data we used MedicalNet (S. Chen, Ma, and Zheng 2019) which is a pretrained multi-modal, multi-tissue model with a ResNet-18 backbone.

We applied the same weighting as shown in MUNIT (Huang, Liu, et al. 2018) to each of the losses used during training. The cycle consistency loss between the original and reconstructed image had a weighting of 10 whilst the remaining losses had an equal weighting of 1. Some methods of style transfer such as the one proposed by Gatys, Ecker, and Bethge 2016 aimed to balance a style and a content loss to generate an image. In our work, we do not need to balance the losses, we only need to minimise each of the losses.

3.3. Implementation

We used PyTorch to implement our approach and trained our approach with the Adam optimiser (Kingma and Ba 2015), with a learning rate of 0.001 and batch size of 8 for 2D data and 5 for 3D data. Early stopping was used with a patience of 20. Our experiments

TIST-Net: Style transfer in dynamic contrast enhanced MRI

11

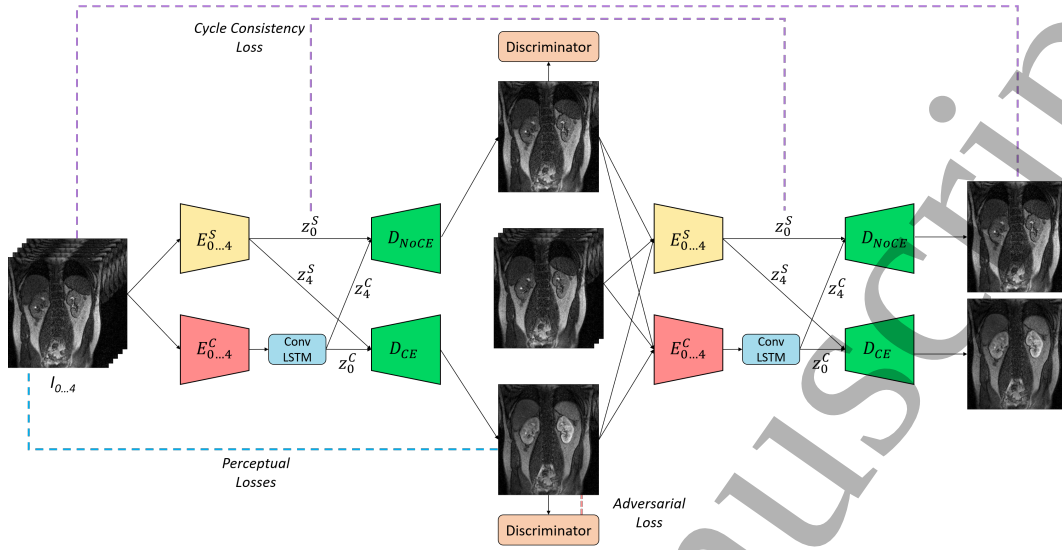


Figure 4: Our model takes an input I of 5 images/volumes to content encoders $E_{0..4}^c$ and style encoders $E_{0..4}^s$. Decoders D_{CE} and D_{NCE} construct images from latent spaces z_i predicted by the content and style encoders: CE represents the decoder which generated contrast enhanced images whilst NCE represents the decoder which generated non-contrast enhanced images. We also show the losses used: cycle consistency, perceptual and adversarial. For clarity of the figure, we have only shown one example for each of the losses.

were run on an RTX Titan GPU. To generate new images, we swapped the predicted style latent spaces from input images with and without contrast enhancement.

3.4. Contrast Weighted (CW) - SSIM

Evaluating generated DCE-MRI for content and style resemblance can be difficult as there can be varying levels of contrast enhancement in an image depending on the time of acquisition after contrast agent injection. Methods such as PSNR and SSIM primarily focus on the pixel level differences of the whole image. Here we evaluate a previously proposed metric, CW-SSIM, constructed by applying a contrast-based weighting to SSIM. To enable weighting by contrast enhancement, we highlighted regions where there was a change in intensity caused by contrast enhancement. Using this information, we were able to weight areas of the image depending on the distance from the contrast enhancement. We began by taking a DCE-MRI series of length n and subtracting the first image (I_0) from each of the images (I_t) and taking an average (Eqn. 3) to find the areas of contrast enhancement (ACE). We then applied a threshold T (Eqn. 4) to determine which voxels were contrast enhanced. For our datasets, an empirical threshold of 20 was found as a good compromise between highlighting the areas of expected contrast enhancement and emphasising noise.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

12

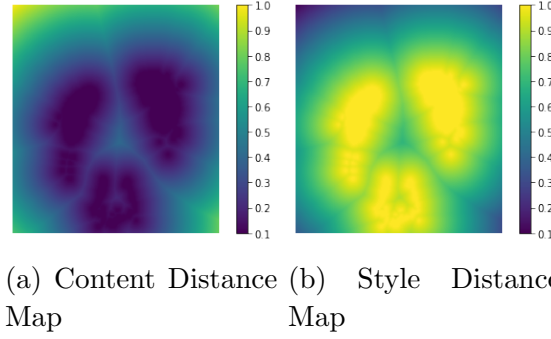


Figure 5: An example of two distance maps (content and style) calculated from one kidney DCE-MRI sequence.

$$ACE(I_n, T) = f \left(\frac{1}{n-1} \sum_{t=1}^{n-1} |I_t - I_0|, T \right) \quad (3)$$

where

$$f(a, T) = \begin{cases} 0, & \text{if } \forall a_i \in a, a_i < T \\ 1, & \text{if } \forall a_i \in a, a_i \geq T \end{cases} \quad (4)$$

Next, we calculated two distance maps, *dist_map*; one for content and one for style evaluation. For the content distance map, we calculated the shortest euclidean distance from each voxel to a contrast enhanced voxel. The distances were normalised between 0.1 and 1 so that each voxel in the image can contribute to the CW-SSIM. To evaluate style, we inverted the distance map so that a voxel has a higher weighting when it is closer to a contrast enhanced voxel. x is the generated image and y is either the content or style image.

$$\text{CW-SSIM}(x, y, \text{dist_map}) = \text{SSIM}(x \cdot \text{dist_map}, y \cdot \text{dist_map}) \quad (5)$$

To evaluate the metric, we performed a series of tests on 100 non-contrast enhanced and 100 contrast enhanced kidney images to ensure our metric evaluates the correct areas of the image. To evaluate the style CW-SSIM, we took an image and modified the intensity values in the regions we expect to see real contrast enhancement by adding or removing intensity values using annotations of the kidney. Similarly for the content CW-SSIM, we took an image and applied warps with increasing amplitude. To warp the image we use sine waves to offset pixels in the image. For each image, we calculated the CW-SSIM and compared it to the SSIM and MS-SSIM.

For the experiments involving the change in intensity, we expected to see a high content CW-SSIM whilst the style CW-SSIM decreased. Similarly, for the experiments that have warped images, we expected to see the content CW-SSIM decrease as the warp increases whilst the style CW-SSIM remains high. We expected the SSIM and MS-SSIM to decrease for each experiment.

3.5. Clinical Evaluation

To evaluate the style transferred images, we conducted a clinical evaluation. To do this, we created a website so that experts could login and visually assess the images. A user was presented with one question at a time. For each question an original image was shown and was noted to the user. We also showed images generated from our method, the generated images from the method proposed in Tattersall et al. 2023b, MUNIT, CycleGAN and StyleGAN3. We also showed the style image used in the style transfer.

The users were not told which approach each image has come from, or which image is the original style image. All of the images were shuffled for each question so that the order of the images were different each time. The users were asked to rank each image according to the following questions:

- (i) Compare the structures in each of the images to the original image. Rank in order (1 being best) each image with the closest structures to the original image, regardless of contrast enhancement.
- (ii) Each of these images have been generated to add (or remove) contrast enhancement to the original image. Rank in order (1 being best).
- (iii) Rank each image in order (1 being best) of general image quality (free from artefacts, realistic noise characteristics etc.).

In total, 16 experts with an average of 11 years of MRI experience completed our questionnaire. The participants had varying expertise such as MR physicists, radiologists, clinical scientists, medical physicist, neurologist and a researcher in medical image analysis. In total, 25 images from each dataset (kidney, prostate and uterus) were evaluated by at least three different observers for each image. If an observer could not decide between images, they were allowed to give them the same score. On average, each user took 26 minutes to complete the study. In addition to evaluating our images, we also used this to evaluate CW-SSIM, by studying how the experts' qualitative evaluation correlates with the quantitative results of the style transferred images.

To test for significant differences between the results of our proposed method and the other style transfer methods, we computed the Kruskal-Wallis test (Kruskal and Wallis 1952). This is a suitable choice as the data does not follow a normal distribution. Additionally, the Kruskal-Wallis test works well with multiple groups that contain ranking data.

4. Results

4.1. Style Transfer - Qualitative Results

Figures 6, 7 and 8 qualitatively highlight the good results of our method on 2D and 3D datasets. We show an example of adding or removing contrast enhancement from an image along with their corresponding metrics: SSIM, content CW-SSIM and style CW-SSIM. We compare TIST-Net to the method proposed in Tattersall et al. 2023b and other popular style transfer methods, namely MUNIT, CycleGAN and StyleGAN3. We did not compare against DRIT++ due to the large computational requirements.

In Figure 6, we can see the results using kidney DCE-MRI. When we compare TIST-Net (Figure 6g) to StyleGAN3 (Figure 6e), our results are sharper and contain structures that better resemble the content image. MUNIT (Figure 6c) and CycleGAN (Figure 6d) both struggle to transfer the style of the style image whereas our method can. A similar trend is shown with the uterus data. For the prostate dataset, each method struggled to add contrast enhancement to the image.

Figures 9, 10 and 11 show results for performing style transfer between two different patients for each dataset. The results for methods such as MUNIT and CycleGAN shows artefacts from the style image in the generated image. TIST-Net generates images that has the structure of the content image whilst having the style of the style image.

4.2. Style Transfer - Quantitative Results

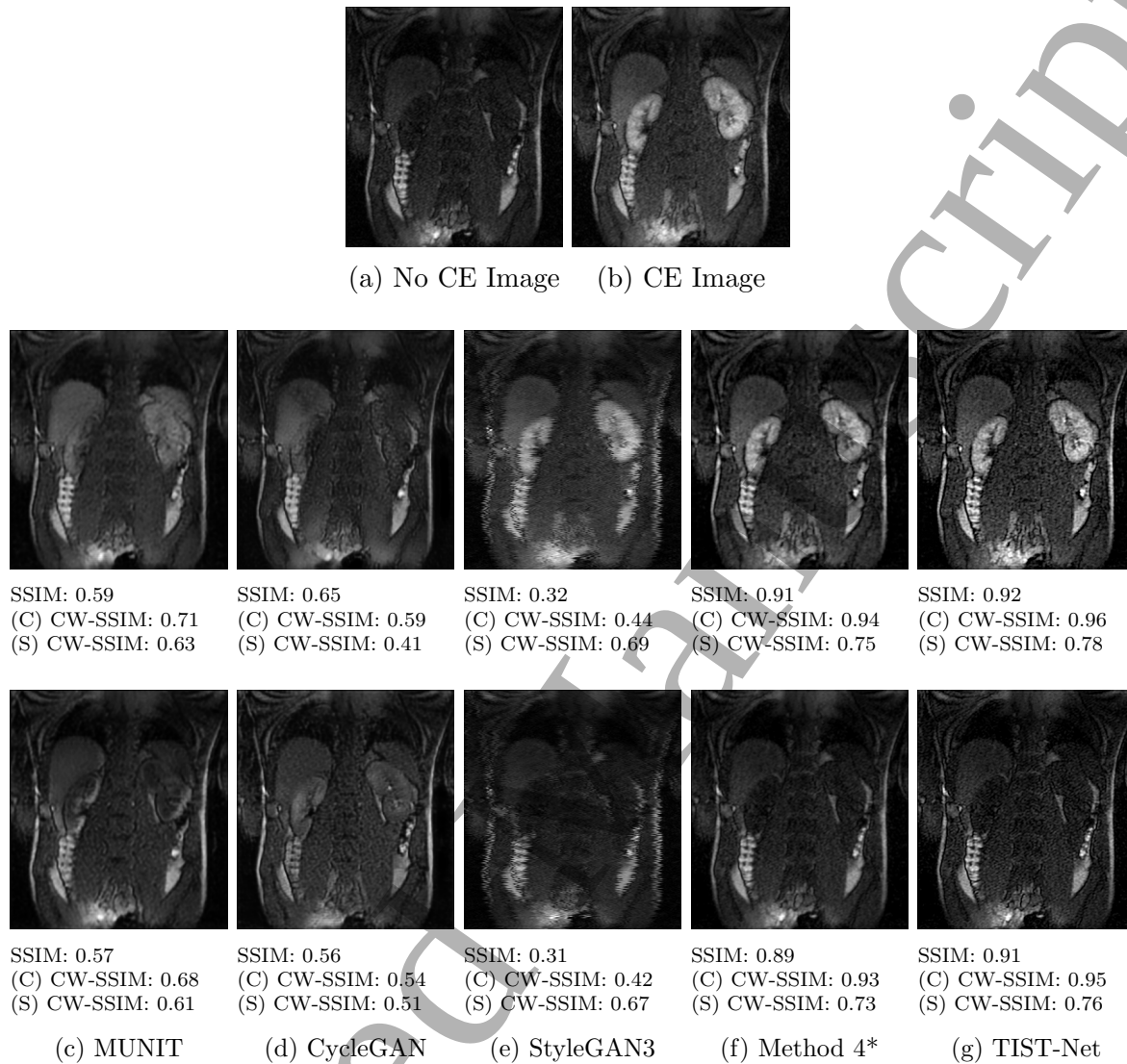
Tables 1, 2 and 3 show our quantitative results: PSNR between the style (image we want to transfer style from) and generated image, SSIM and MS-SSIM between the content (image we want to take structure from) and generated image and finally, our proposed weighted SSIMs. We compare our proposed method with the method proposed in Tattersall et al. 2023b and other popular style transfer methods, namely MUNIT, CycleGAN and StyleGAN3. For each style transfer direction and metric, our method consistently outperforms the other approaches.

4.3. CW-SSIM Evaluation Results

Figures 12 and 13 show some examples of the modifications made to the images with the respective scores. Tables 4 and 5 show the quantitative results from each of the tests. When we evaluate the style CW-SSIM, we can see that it decreases for both directions (when increasing or decreasing intensity values), whilst the content CW-SSIM remains high. The SSIM and MS-SSIM both decrease as the style changes. For the content evaluation tests, as the intensity of the warp increases, the score decreases. The SSIM and the MS-SSIM both decrease as the intensity of the warps increases. The style CW-SSIM decreases slightly as the areas of contrast enhancement are warped, but the score remains high. These results show that the proposed method can separate the evaluation of content and style from the generated images.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

15

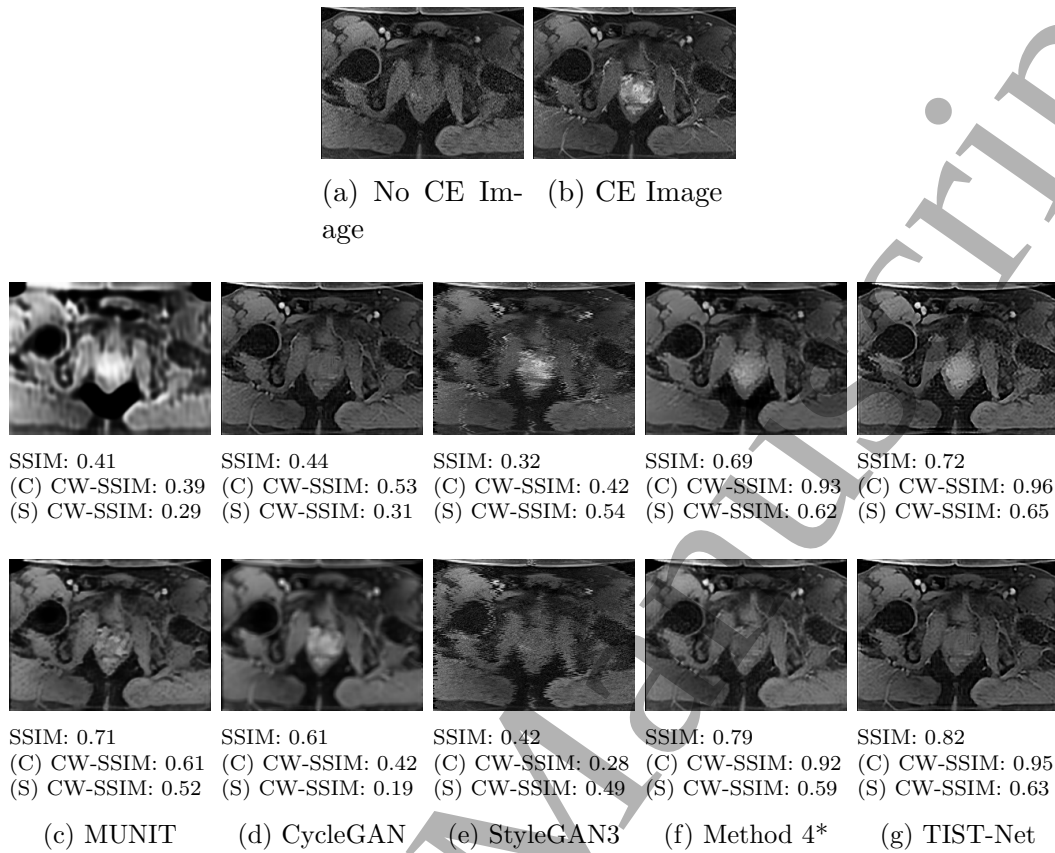


* Method 4 is the method proposed by Tattersall et al. 2023b

Figure 6: Example results from different style transfer approaches ((c) MUNIT, (d) CycleGAN, (e) StyleGAN, (f) method in Tattersall et al. 2023b), (g) TIST-Net) with the kidney data. The first row (a and b) shows the input images. The second row shows the results when (a) is used as the content image and (b) as the style. The third row shows results when (b) is used as the content image and (a) as the style. We also show scores given by the SSIM, content (C) CW-SSIM and style (S) CW-SSIM.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

16

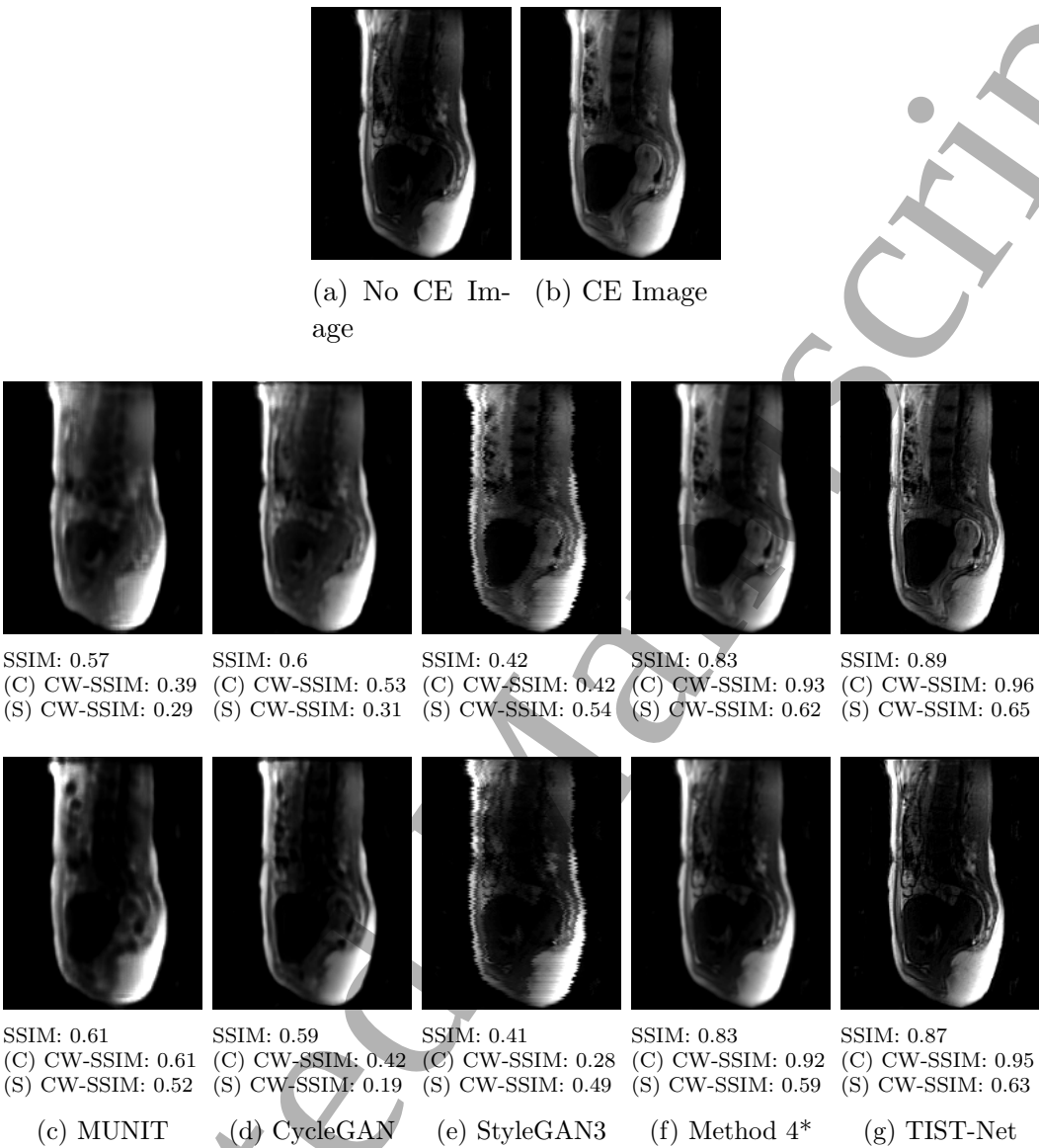


* Method 4 is the method proposed by Tattersall et al. 2023b

Figure 7: Example results from different style transfer approaches with the prostate data (a 2D slice is shown from the 3D volume). The first row (a and b) shows the input images. The second row shows the results when (a) is used as the content image and (b) as the style. The third row shows results when (b) is used as the content image and (a) as the style. We also show scores given by the SSIM, content (C) CW-SSIM and style (S) CW-SSIM.

4.4. Clinical Evaluation Results

Tables 6, 7 and 8 show the results from our clinical user study. Each table shows the mean and standard deviation (std) of the ranks given by the participants for each question, style transfer direction and approach. Our method outperforms the other methods and was even ranked higher than the real image shown to the user for having the closest structure. We performed the Kruskal-Wallis test with a significance level of 0.05 and found that there was a significant difference between the scores given to our method and the other methods. Additionally, we also found that there was no significant difference when we stratified the observers by experience.

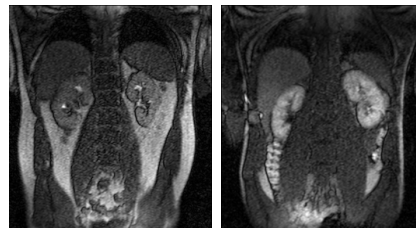


* Method 4 is the method proposed by Tattersall et al. 2023b

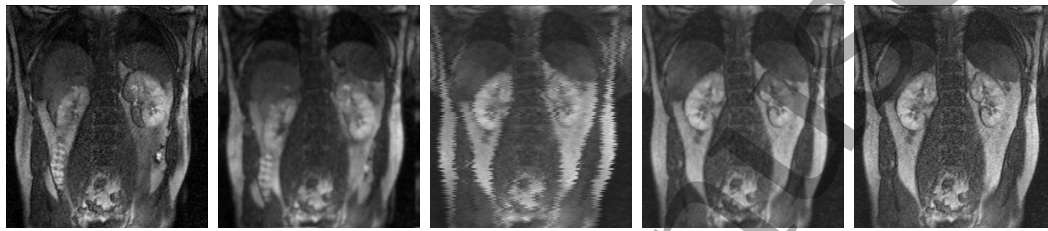
Figure 8: Example results from different style transfer approaches with the uterus data (a 2D slice is shown from the 3D volume). The first row (a and b) shows the input images. The second row shows the results when (a) is used as the content image and (b) as the style. The third row shows results when (b) is used as the content image and (a) as the style. We also show scores given by the SSIM, content (C) CW-SSIM and style (S) CW-SSIM.

TIST-Net: Style transfer in dynamic contrast enhanced MRI

18



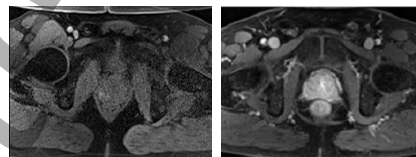
(a) No CE Im- (b) CE Image
age



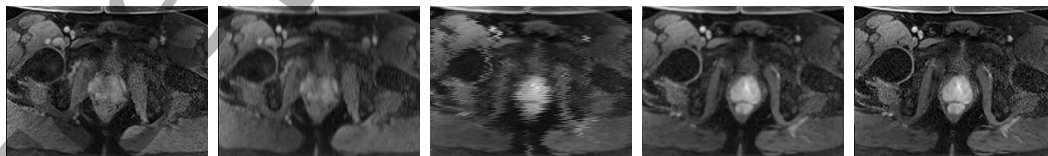
(c) MUNIT (d) CycleGAN (e) StyleGAN3 (f) Method 4* (g) TIST-Net

* Method 4 is the method proposed by Tattersall et al. 2023b

Figure 9: Example results of performing style transfer between two different patients for each method with the kidney data. (a) was used as the content image and (b) was used as the style image.



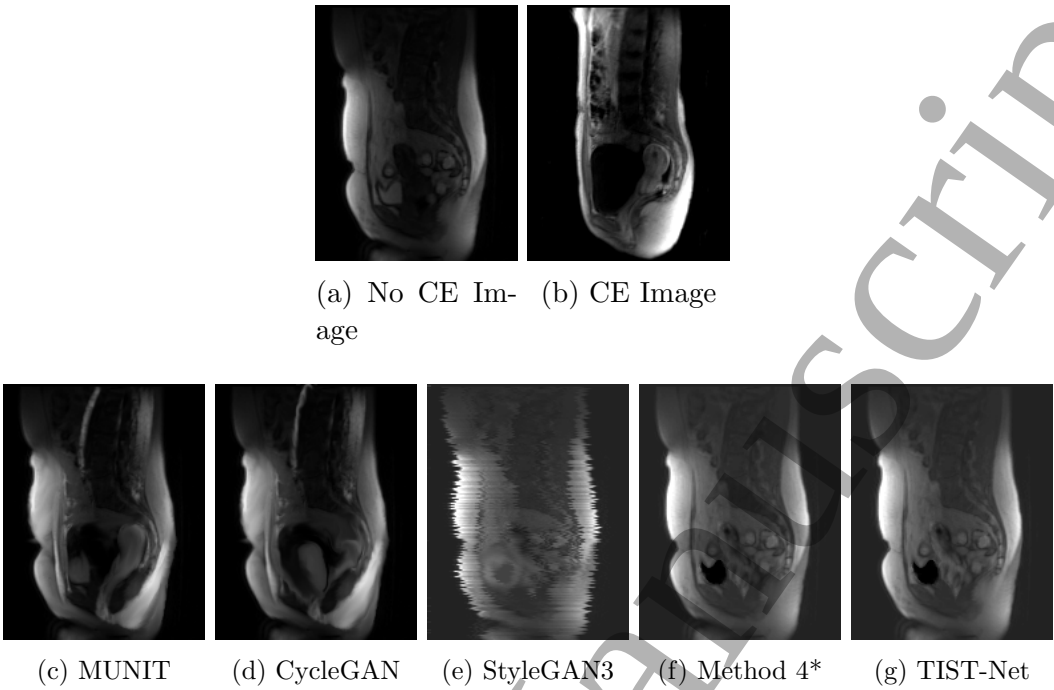
(a) No CE Im- (b) CE Image
age



(c) MUNIT (d) CycleGAN (e) StyleGAN3 (f) Method 4* (g) TIST-Net

* Method 4 is the method proposed by Tattersall et al. 2023b

Figure 10: Example results of performing style transfer between two different patients for each method with the prostate data. (a) was used as the content image and (b) was used as the style image.



* Method 4 is the method proposed by Tattersall et al. 2023b

Figure 11: Example results of performing style transfer between two different patients for each method with the uterus data. (a) was used as the content image and (b) was used as the style image.

Table 1: Quantitative results for each approach when adding or removing CE in the kidney.

Metric	Direction	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
PSNR	Non-CE to CE	70.6 ± 3.2	66.5 ± 2.7	52.3 ± 3.9	81.3 ± 2.7	83.1 ± 1.9
	CE to Non-CE	71.5 ± 3.4	66.8 ± 2.8	51.9 ± 3.8	78.5 ± 2.6	82.4 ± 1.6
SSIM	Non-CE to CE	0.61 ± 0.05	0.66 ± 0.06	0.32 ± 0.06	0.89 ± 0.04	0.91 ± 0.03
	CE to Non-CE	0.58 ± 0.04	0.58 ± 0.05	0.34 ± 0.07	0.88 ± 0.03	0.89 ± 0.03
MS SSIM	Non-CE to CE	0.55 ± 0.04	0.43 ± 0.07	0.34 ± 0.04	0.82 ± 0.03	0.87 ± 0.03
	CE to Non-CE	0.52 ± 0.04	0.45 ± 0.05	0.37 ± 0.05	0.79 ± 0.04	0.83 ± 0.04
Content CW-SSIM	Non-CE to CE	0.73 ± 0.03	0.61 ± 0.05	0.47 ± 0.05	0.95 ± 0.02	0.96 ± 0.03
	CE to Non-CE	0.69 ± 0.03	0.57 ± 0.05	0.41 ± 0.05	0.94 ± 0.02	0.96 ± 0.03
Style CW-SSIM	Non-CE to CE	0.64 ± 0.03	0.41 ± 0.05	0.67 ± 0.05	0.74 ± 0.02	0.79 ± 0.02
	CE to Non-CE	0.62 ± 0.03	0.49 ± 0.05	0.64 ± 0.05	0.72 ± 0.02	0.81 ± 0.04

* Method 4 is the method proposed by Tattersall et al. 2023b

TIST-Net: Style transfer in dynamic contrast enhanced MRI

20

Table 2: Quantitative results for each approach when adding or removing CE in the prostate.

Metric	Direction	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
PSNR	Non-CE to CE	40.6 $\pm$ 4.3	52.2 $\pm$ 3.8	43.5 $\pm$ 5.2	72.1 $\pm$ 3.6	77.2 $\pm$ 2.3
	CE to Non-CE	68.7 $\pm$ 4.1	61.2 $\pm$ 3.4	51.9 $\pm$ 6.1	68.9 $\pm$ 4.1	70.4 $\pm$ 2.2
SSIM	Non-CE to CE	0.41 $\pm$ 0.06	0.46 $\pm$ 0.05	0.36 $\pm$ 0.09	0.71 $\pm$ 0.04	0.73 $\pm$ 0.04
	CE to Non-CE	0.69 $\pm$ 0.04	0.59 $\pm$ 0.05	0.43 $\pm$ 0.08	0.78 $\pm$ 0.03	0.82 $\pm$ 0.03
MS SSIM	Non-CE to CE	0.49 $\pm$ 0.04	0.38 $\pm$ 0.06	0.28 $\pm$ 0.05	0.76 $\pm$ 0.04	0.82 $\pm$ 0.03
	CE to Non-CE	0.47 $\pm$ 0.04	0.39 $\pm$ 0.05	0.32 $\pm$ 0.06	0.71 $\pm$ 0.03	0.75 $\pm$ 0.03
Content CW-SSIM	Non-CE to CE	0.38 $\pm$ 0.04	0.52 $\pm$ 0.05	0.41 $\pm$ 0.05	0.92 $\pm$ 0.03	0.94 $\pm$ 0.03
	CE to Non-CE	0.63 $\pm$ 0.04	0.46 $\pm$ 0.04	0.32 $\pm$ 0.06	0.93 $\pm$ 0.04	0.95 $\pm$ 0.04
Style CW-SSIM	Non-CE to CE	0.34 $\pm$ 0.05	0.28 $\pm$ 0.06	0.52 $\pm$ 0.07	0.59 $\pm$ 0.05	0.66 $\pm$ 0.04
	CE to Non-CE	0.53 $\pm$ 0.04	0.21 $\pm$ 0.05	0.47 $\pm$ 0.06	0.57 $\pm$ 0.04	0.64 $\pm$ 0.04

* Method 4 is the method proposed by Tattersall et al. 2023b

Table 3: Quantitative results for each approach when adding or removing CE in the uterus.

Metric	Direction	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
PSNR	Non-CE to CE	71.5 $\pm$ 2.9	64.9 $\pm$ 3.1	51.8 $\pm$ 3.6	80.8 $\pm$ 2.4	83.6 $\pm$ 2.1
	CE to Non-CE	72.2 $\pm$ 2.8	67.3 $\pm$ 2.4	52.3 $\pm$ 3.4	79.3 $\pm$ 2.2	83.1 $\pm$ 1.4
SSIM	Non-CE to CE	0.57 $\pm$ 0.05	0.58 $\pm$ 0.06	0.39 $\pm$ 0.06	0.82 $\pm$ 0.04	0.88 $\pm$ 0.04
	CE to Non-CE	0.60 $\pm$ 0.06	0.58 $\pm$ 0.04	0.38 $\pm$ 0.05	0.82 $\pm$ 0.05	0.87 $\pm$ 0.03
MS SSIM	Non-CE to CE	0.52 $\pm$ 0.04	0.41 $\pm$ 0.06	0.29 $\pm$ 0.03	0.79 $\pm$ 0.04	0.88 $\pm$ 0.05
	CE to Non-CE	0.53 $\pm$ 0.06	0.42 $\pm$ 0.06	0.36 $\pm$ 0.07	0.81 $\pm$ 0.06	0.85 $\pm$ 0.05
Content CW-SSIM	Non-CE to CE	0.74 $\pm$ 0.04	0.64 $\pm$ 0.04	0.49 $\pm$ 0.06	0.94 $\pm$ 0.04	0.96 $\pm$ 0.05
	CE to Non-CE	0.67 $\pm$ 0.05	0.58 $\pm$ 0.06	0.42 $\pm$ 0.04	0.92 $\pm$ 0.04	0.95 $\pm$ 0.06
Style CW-SSIM	Non-CE to CE	0.62 $\pm$ 0.03	0.43 $\pm$ 0.06	0.65 $\pm$ 0.04	0.73 $\pm$ 0.04	0.78 $\pm$ 0.03
	CE to Non-CE	0.64 $\pm$ 0.04	0.48 $\pm$ 0.04	0.66 $\pm$ 0.04	0.73 $\pm$ 0.04	0.79 $\pm$ 0.02

* Method 4 is the method proposed by Tattersall et al. 2023b

Table 4: Mean results and standard deviation of the content (C) CW-SSIM evaluation.

Metric	Warped 1	Warped 2	Warped 3
(C) CW-SSIM	0.78 $\pm$ 0.04	0.63 $\pm$ 0.05	0.41 $\pm$ 0.03
(S) CW-SSIM	0.95 $\pm$ 0.03	0.87 $\pm$ 0.04	0.76 $\pm$ 0.06
SSIM	0.84 $\pm$ 0.03	0.61 $\pm$ 0.06	0.44 $\pm$ 0.06
MS-SSIM	0.87 $\pm$ 0.04	0.58 $\pm$ 0.03	0.47 $\pm$ 0.04

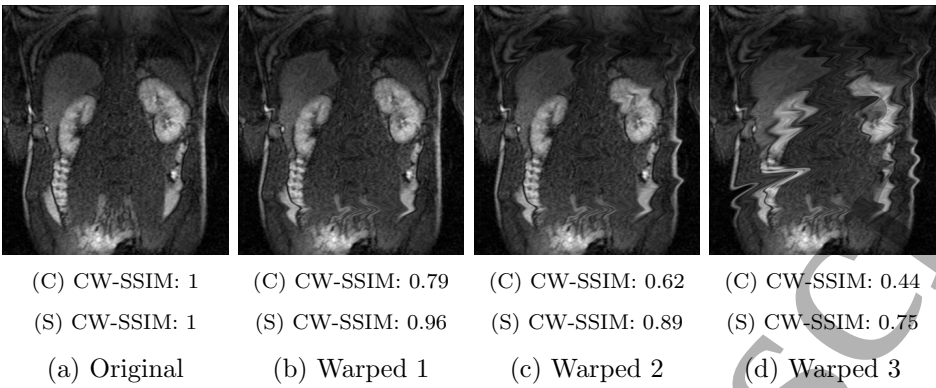


Figure 12: An experiment to evaluate the content CW-SSIM. Here we have an original image and images with an increasingly large amount of warps added to them. The CW-SSIM score is also shown.

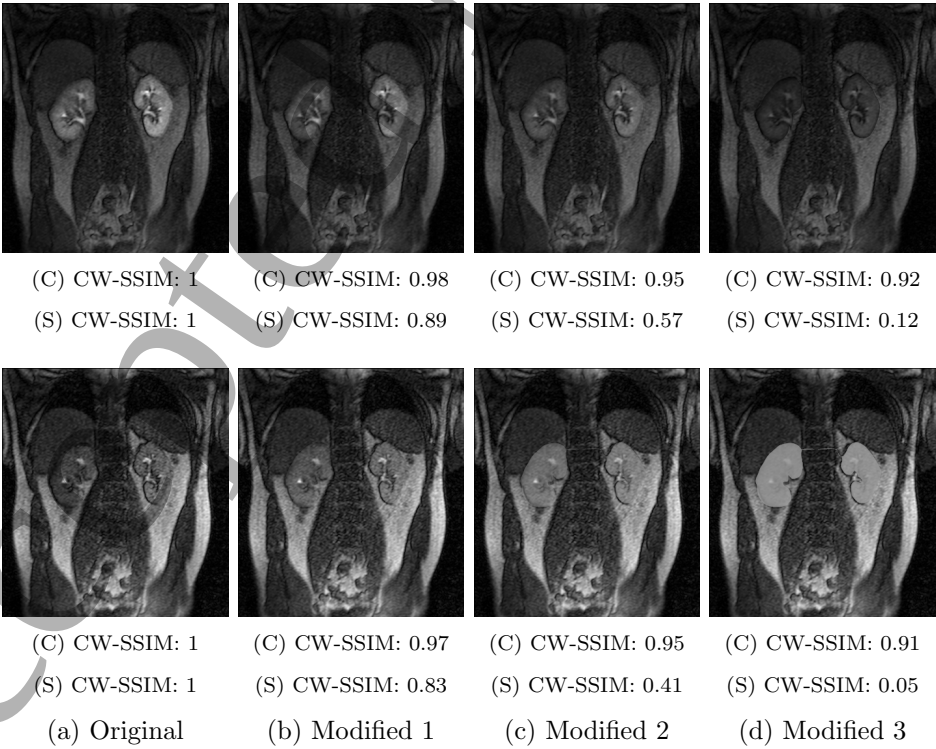


Figure 13: An experiment to evaluate the content CW-SSIM. Here we have an original image and images with intensity either being added or removed. The CW-SSIM score is also shown.

Table 5: Mean results and standard deviation of the style (S) CW-SSIM evaluation.

Metric	Modified 1	Modified 2	Modified 3
(C) CW-SSIM	0.96 ± 0.04	0.94 ± 0.05	0.89 ± 0.05
(S) CW-SSIM	0.85 ± 0.06	0.49 ± 0.09	0.08 ± 0.06
SSIM	0.89 ± 0.03	0.82 ± 0.04	0.76 ± 0.04
MS-SSIM	0.88 ± 0.05	0.84 ± 0.03	0.79 ± 0.04

Table 6: Mean rank and standard deviation of the kidney data for each question in the user study.

Question	Original Image	Style Image	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
1	CE	2.83 ± 0.76	4.95 ± 0.82	5.04 ± 0.46	5.33 ± 0.71	1.93 ± 0.54	1.30 ± 0.62
	NoCE	2.79 ± 0.74	4.57 ± 0.74	5.00 ± 0.40	5.35 ± 0.72	1.79 ± 0.56	1.15 ± 0.47
2	CE	2.83 ± 0.76	4.64 ± 0.81	5.04 ± 0.50	5.39 ± 0.66	1.87 ± 0.55	1.21 ± 0.56
	NoCE	2.79 ± 0.74	4.72 ± 0.78	4.99 ± 0.48	5.34 ± 0.67	1.90 ± 0.47	1.23 ± 0.55
3	CE	2.90 ± 0.74	4.79 ± 0.81	4.95 ± 0.45	5.33 ± 0.68	1.80 ± 0.56	1.27 ± 0.61
	NoCE	2.79 ± 0.72	4.67 ± 0.76	4.97 ± 0.50	5.40 ± 0.70	1.87 ± 0.59	1.28 ± 0.64

* Method 4 is the method proposed by Tattersall et al. 2023b

Table 7: Mean rank and standard deviation of the prostate data for each question in the user study.

Question	Original Image	Style Image	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
1	CE	2.83 ± 0.76	4.61 ± 0.72	4.99 ± 0.38	5.35 ± 0.67	1.93 ± 0.55	1.27 ± 0.59
	NoCE	2.87 ± 0.77	4.72 ± 0.84	5.01 ± 0.50	5.48 ± 0.64	1.83 ± 0.58	1.31 ± 0.65
2	CE	1.43 ± 0.69	4.61 ± 0.73	5.01 ± 0.48	5.39 ± 0.62	1.97 ± 0.47	1.86 ± 0.80
	NoCE	1.40 ± 0.63	4.65 ± 0.80	5.05 ± 0.44	5.32 ± 0.70	2.08 ± 0.71	1.84 ± 0.69
3	CE	2.75 ± 0.74	4.74 ± 0.82	4.97 ± 0.37	5.47 ± 0.63	1.87 ± 0.50	1.31 ± 0.65
	NoCE	2.77 ± 0.71	4.70 ± 0.77	5.03 ± 0.44	5.38 ± 0.70	1.87 ± 0.52	1.27 ± 0.57

* Method 4 is the method proposed by Tattersall et al. 2023b

Table 8: Mean rank and standard deviation of the uterus data for each question in the user study.

Question	Original Image	Style Image	MUNIT	CycleGAN	StyleGAN3	Method 4*	TIST-Net
1	CE	2.83 ± 0.78	4.68 ± 0.79	5.04 ± 0.46	5.43 ± 0.67	1.98 ± 0.54	1.30 ± 0.64
	NoCE	2.80 ± 0.75	4.72 ± 0.80	5.02 ± 0.46	5.37 ± 0.64	1.91 ± 0.55	1.30 ± 0.63
2	CE	2.80 ± 0.70	4.69 ± 0.79	5.00 ± 0.48	5.44 ± 0.65	1.91 ± 0.54	1.29 ± 0.62
	NoCE	2.73 ± 0.75	4.73 ± 0.80	4.96 ± 0.40	5.39 ± 0.63	1.95 ± 0.57	1.25 ± 0.58
3	CE	2.72 ± 0.72	4.81 ± 0.84	4.94 ± 0.49	5.36 ± 0.65	1.90 ± 0.51	1.44 ± 0.73
	NoCE	2.81 ± 0.75	4.71 ± 0.78	5.07 ± 0.50	5.46 ± 0.63	1.91 ± 0.53	1.32 ± 0.68

* Method 4 is the method proposed by Tattersall et al. 2023b

5. Discussion

In this work, we proposed TIST-Net, a method for style transfer that utilised temporal information by using convolutional LSTMs. This enabled the model to learn changes in structure caused by contrast enhancement across the temporal dimension. We improved upon classical style transfer by using AdaConv combined with deformable convolutions. The deformable convolutions predict an offset to adapt the AdaConv kernel to change the receptive field of the model. The receptive field of the model can then account for local geometric variations. This enabled better fine-grained control over the transfer of style. TIST-Net can be used as a method of augmentation for DCE-MRI data to generate new images and fully utilise any available annotations in a dataset.

Our qualitative evaluation showed that TIST-Net led to sharper images, better content preservation, better localised CE and realistic MRI appearance compared to previous works and state-of-the-art style transfer methods. Our model achieved good results in both style transfer directions (adding or removing CE) for both the quantitative and qualitative results. When performing style transfer between different patients, TIST-Net was able to generate images which had the structure of the content image whilst having the style characteristics of the style image. Additionally, we outperformed the other algorithms for each metric, for each style transfer direction. Our results showed that when the contrast enhancement has defined edges in the image, such as those in the kidney, it is an easier task to perform style transfer in both directions. In comparison, when there was no clear boundary to the contrast enhancement (such as in the prostate data), the task was much harder. Typically, when moving from 2D to 3D models, there can be a decrease in performance due to the increased number of parameters to learn whilst having a small amount of data. However, we found that TIST-Net performed well on 2D and 3D datasets as evidenced by our results on the uterus dataset.

CW-SSIM enabled the evaluation of the quality of the content and style of a generated image when there was contrast enhancement. When we compared it with the qualitative results from our clinical evaluation, the quality of the images was in line with each of the CW-SSIM results. By using two distance maps to weight the SSIM, we were able to evaluate the similarity in overall structure as well as the localised style between the generated image and the content and style images. An advantage to using this metric over the SSIM or MS-SSIM is that we were able to measure the content and style of a generated image separately. This allowed us to ensure that the key structures remained whilst the intensity values of the image matched those of the style image. This is important with DCE-MRI as the contrast enhancement is tissue specific rather than a global change in the image. When we compared the style CW-SSIM of the kidney and uterus images generated by our method with their corresponding style images, we could see that the high scores reflect the image quality. In comparison, the style CW-SSIM scores of the prostate data were much lower, which again reflects the quality of the generated image. By separating the content and style weighting, we lessen the

possibility of regions of the image with little or no contrast enhancement inflating the SSIM score.

A limitation to our style transfer approach was that it struggled to generate realistic images when the edges between areas of contrast enhancement were ambiguous. This was seen in the prostate data where there was no clear boundary of the contrast enhancement compared to the kidney or uterus data. It is worth noting that this struggle was shared by all the methods we compared to. We hypothesise that this lack of contrast between neighbouring tissues might lead to a sub-optimal disentanglement. Finally, similar to MUNIT, our method has large computational and memory requirements. In this work, we focused on using LSTMs to encode content latent spaces and not style latent spaces. During our preliminary work we found that the model was unable to learn from the style latent spaces effectively and it was detrimental to our model, leading to worse results. The preliminary exploration was not shown in this paper as it is outside the scope of this paper.

The results from our clinical evaluation confirmed that TIST-Net generates images that are realistic as TIST-Net was often given a higher rank than the real image. These results show that our images have similar structures to the content image, similar styles to the style image as well as having characteristics that are expected in MRI. The results also match the order given by the content CW-SSIM, further proving that this metric is suitable for evaluating synthetic images.

6. Conclusion

We proposed TIST-Net, a style transfer approach which used temporal information to predict disentangled representations of content and style. To learn spatio-temporal information, convolutional LSTMs were used which allowed better content latent spaces to be predicted from structural information through time. We also used deformable convolutions to offset AdaConv to combine content and style latent spaces. We observed an increase in performance for both adding and removing contrast enhancement compared to state-of-the-art style transfer methods.

The qualitative and quantitative analyses showed that our method outperformed state-of-the-art style transfer techniques. The results from our clinical evaluation confirmed that our method generates images that are realistic. We also evaluated CW-SSIM, to validate its viability as a metric. The results from our clinical evaluation further demonstrated that it can be used to evaluate the generated images. Using TIST-Net, we can generate new images and sequences with varying levels of contrast enhancement. Using this augmentation approach enables the use of a small number of annotations making it easier to train robust models for tasks such as image registration or segmentation.

REFERENCES

25

Acknowledgments

This work was funded by Medical Research Scotland and Canon Medical Research Europe. We would also like to thank the participants that completed the clinical evaluation.

References

Cai, Naxin et al. (2023). "Registration on DCE-MRI images via multi-domain image-to-image translation". In: *Computerized Medical Imaging and Graphics*.

Chandran, Prashanth et al. (2021). "Adaptive Convolutions for Structure-Aware Style Transfer". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Chao, Zeyi et al. (2018). "Research on real-time local rainfall prediction based on MEMS sensors". In: *Journal of Sensors*.

Chen, Dongdong et al. (2017). "Coherent Online Video Style Transfer". In: *IEEE ICCV*.

Chen, Sihong, Kai Ma, and Yefeng Zheng (2019). "Med3D: Transfer Learning for 3D Medical Image Analysis". In: *arXiv preprint arXiv:1904.00625*.

Dai, Jifeng et al. (2017). "Deformable Convolutional Networks". In: *Proceedings of the IEEE International Conference on Computer Vision*.

Deng, Jia et al. (2009). "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*.

Galli, Antonio et al. (2023). "Generative Adversarial Networks for Domain Translation in Unpaired Breast DCE-MRI Datasets". In: *Deep Learning Theory and Applications*.

Gatys, Leon, Alexander Ecker, and Matthias Bethge (2016). "A Neural Algorithm of Artistic Style". In: *Journal of Vision*. ISSN: 1534-7362.

Graves, Alex and Jürgen Schmidhuber (2005). "Framewise phoneme classification with bidirectional LSTM and other neural network architectures". In: *Neural Networks*.

Huang, Xun and Serge Belongie (2017). "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization". In: *Proceedings of the IEEE International Conference on Computer Vision*.

Huang, Xun, Ming-Yu Liu, et al. (2018). "Multimodal Unsupervised Image-to-Image Translation". In: *Computer Vision – ECCV 2018*.

Ingrisch, Michael and Steven Sourbron (2013). "Tracer-kinetic modeling of dynamic contrast-enhanced MRI and CT: A primer". In: *Journal of Pharmacokinetics and Pharmacodynamics*.

Isola, Phillip et al. (2017). "Image-to-image translation with conditional adversarial networks". In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*.

Johnson, Justin, Alexandre Alahi, and Li Fei-Fei (2016). "Perceptual Losses for Real-Time Style Transfer and Super-Resolution". In: *ECCV*.

Karras, Tero et al. (2021). "Alias-Free Generative Adversarial Networks". In: *NeurIPS*.

REFERENCES

26

- Kingma, Diederik P. and Jimmy Ba (2015). “Adam: A Method for Stochastic Optimization”. In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.
- Kruskal, William H. and W. Allen Wallis (1952). “Use of Ranks in One-Criterion Variance Analysis”. In: *Journal of the American Statistical Association*.
- Lee, Hsin-Ying et al. (2020). “DRIT++: Diverse Image-to-Image Translation via Disentangled Representations”. In: *International Journal of Computer Vision*.
- Lemaître, Guillaume et al. (2015). “Computer-Aided Detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: A review”. In: *Computers in Biology and Medicine*.
- Lietzmann, Florian et al. (2012). “DCE-MRI of the human kidney using BLADE: A feasibility study in healthy volunteers”. In: *Journal of Magnetic Resonance Imaging*.
- Maspero, Matteo et al. (2018). “Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy”. In: *Physics in Medicine and Biology*.
- Park, Taesung et al. (2019). “Semantic image synthesis with spatially-adaptive normalization”. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Reaungamornrat, Sureerat et al. (2022). “Multimodal image synthesis based on disentanglement representations of anatomical and modality specific features, learned using uncooperative relativistic GAN”. In: *Medical Image Analysis*.
- Reavey, J J et al. (2021). “Markers of human endometrial hypoxia can be detected in vivo and ex vivo during physiological menstruation”. In: *Human Reproduction*.
- Russakovsky, Olga et al. (2015). “ImageNet Large Scale Visual Recognition Challenge”. In: *IJCV*.
- Shorten, Connor and Taghi M. Khoshgoftaar (2019). “A survey on Image Data Augmentation for Deep Learning”. In: *Journal of Big Data*.
- Tattersall, Adam G. et al. (2023a). “Can a single image processing algorithm work equally well across all phases of DCE-MRI”. In: *Medical Imaging: Image Processing*.
- (2023b). “Style transfer in dynamic contrast enhanced MRI using spatial and temporal information”. In: *Simulation and Synthesis in Medical Imaging*.
- Ulyanov, Dmitry et al. (2016). “Texture networks: Feed-forward synthesis of textures and stylized images”. In: *ICML*.
- Wang, Z., E.P. Simoncelli, and A.C. Bovik (2003). “Multiscale structural similarity for image quality assessment”. In: *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*.
- Wang, Zhou et al. (2004). “Image quality assessment: from error visibility to structural similarity”. In: *IEEE Transactions on Image Processing*.
- Wolterink, Jelmer M. et al. (2017). “Generative adversarial networks for noise reduction in low-dose CT”. In: *IEEE Transactions on Medical Imaging*.
- Yang, Junlin et al. (2019). “Unsupervised Domain Adaptation via Disentangled Representations: Application to Cross-Modality Liver Segmentation”. In: *Lecture*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

Zhu, Jun-Yan et al. (2017). “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*.