Improving style transfer in dynamic contrast enhanced MRI using a spatio-temporal approach

Adam G. Tattersall[1,2], Keith A. Goatman[1], Lucy E. Kershaw[2], Scott I. K. Semple[2] and Sonia Dahdouh[1]

[1] Canon Medical Research Europe Ltd. [2] University of Edinburgh

Purpose

Dynamic contrast enhanced (DCE) MRI involves rapidly acquiring T1w images whilst injecting a gadolinium contrast agent (CA). The CA causes a contrast enhancement (CE) in areas of uptake within the images, and the enhancement is **time varying and tissue dependent.**

Obtaining manually annotated ground truth for DCE-MRI data is **time consuming and difficult**. Fully annotated datasets are therefore scarce, which reduces the diversity of available data to train deep learning models.

Current style transfer methods struggle due to the localised nature of contrast enhancement. Additionally, they do not utilise any spatial information.

We propose a new augmentation method that combines autoencoders to disentangle content and style with **convolutional LSTMs** to model predicted latent spaces along time. We also use **adaptive convolutions** to tackle the localised nature of contrast enhancement in DCE-MRI.

Method

We use a structure composed of encoders and decoders to first disentangle content and style latent spaces and recombine them to generate a new image.

We use additional content encoders to predict latent spaces for images from multiple time points. Then, we pass the predicted content latent spaces to a convolutional LSTM which has been set up to be bi-directional. We use a convolutional LSTM as this allows modelling of the temporal information from a sequence of images.

Figure 1 shows our global architecture. Our model takes an input *I* of 5 images/volumes to content encoders $E^c_{0...4}$ and style encoders $E^s_{0...4}$. Generators G_{CE} and G_{NCE} construct images from latent spaces *z* predicted by the content and style encoders.

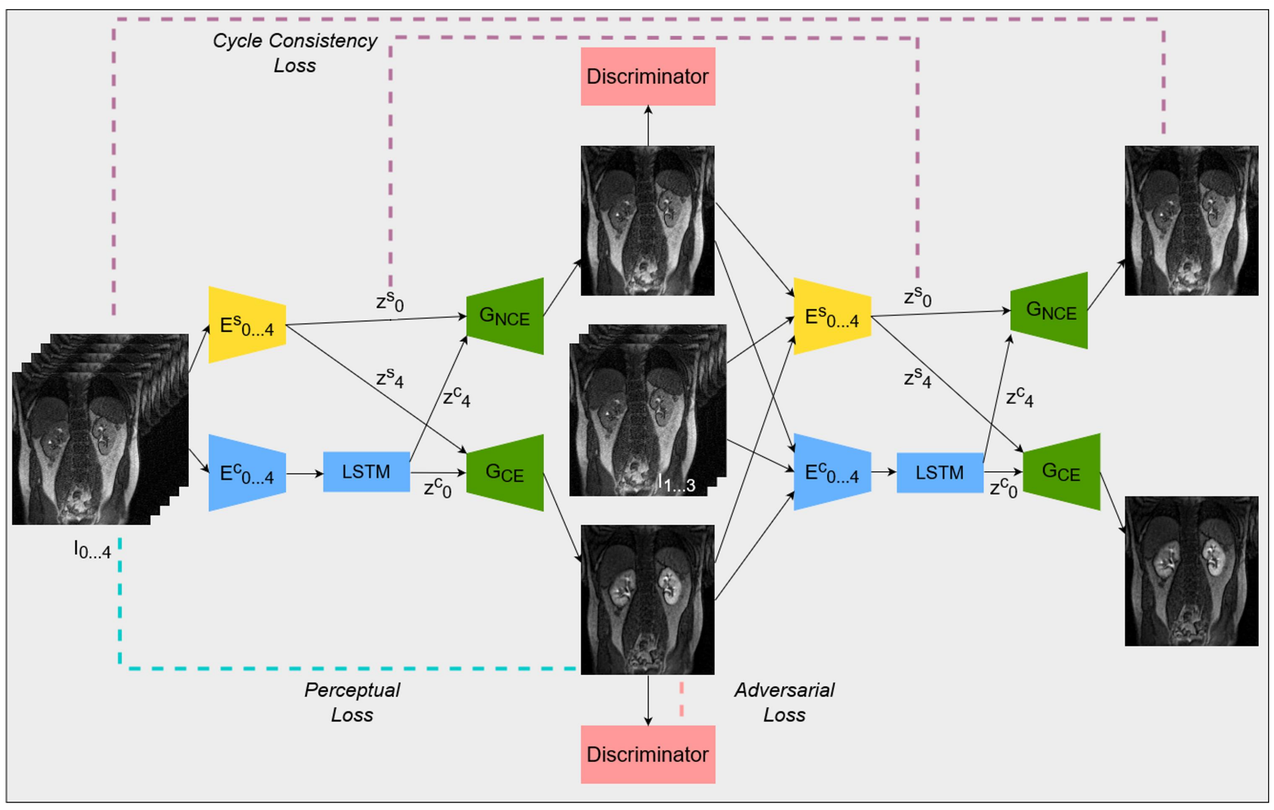

Figure 1: Global architecture of our proposed method

Results

Metric	Direction	MUNIT	CycleGAN	StyleGAN3	Our Method
PSNR	Non-CE to CE	70.6 ± 3.2	66.5 ± 2.7	52.3 ± 3.9	**81.3 ± 2.7**
	CE to Non-CE	71.5 ± 3.4	66.8 ± 2.8	51.9 ± 3.8	**78.5 ± 2.6**
SSIM	Non-CE to CE	0.61 ± 0.05	0.66 ± 0.06	0.32 ± 0.06	**0.89 ± 0.04**
	CE to Non-CE	0.58 ± 0.04	0.58 ± 0.05	0.34 ± 0.07	**0.88 ± 0.03**
MS-SSIM	Non-CE to CE	0.55 ± 0.04	0.43 ± 0.07	0.34 ± 0.04	**0.82 ± 0.03**
	CE to Non-CE	0.52 ± 0.04	0.45 ± 0.05	0.37 ± 0.05	**0.79 ± 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**
	CE to Non-CE	0.69 ± 0.03	0.57 ± 0.05	0.41 ± 0.05	**0.94 ± 0.02**
Style CW-SSIM	Non-CE to CE	0.64 ± 0.03	0.41 ± 0.05	0.67 ± 0.05	**0.74 ± 0.02**
	CE to Non-CE	0.62 ± 0.03	0.49 ± 0.05	0.64 ± 0.05	**0.72 ± 0.02**

Table 1: Quantitative results of style transfer with kidney DCE-MRI

We also evaluated our method with prostate DCE-MRI. Please see our paper for those results.

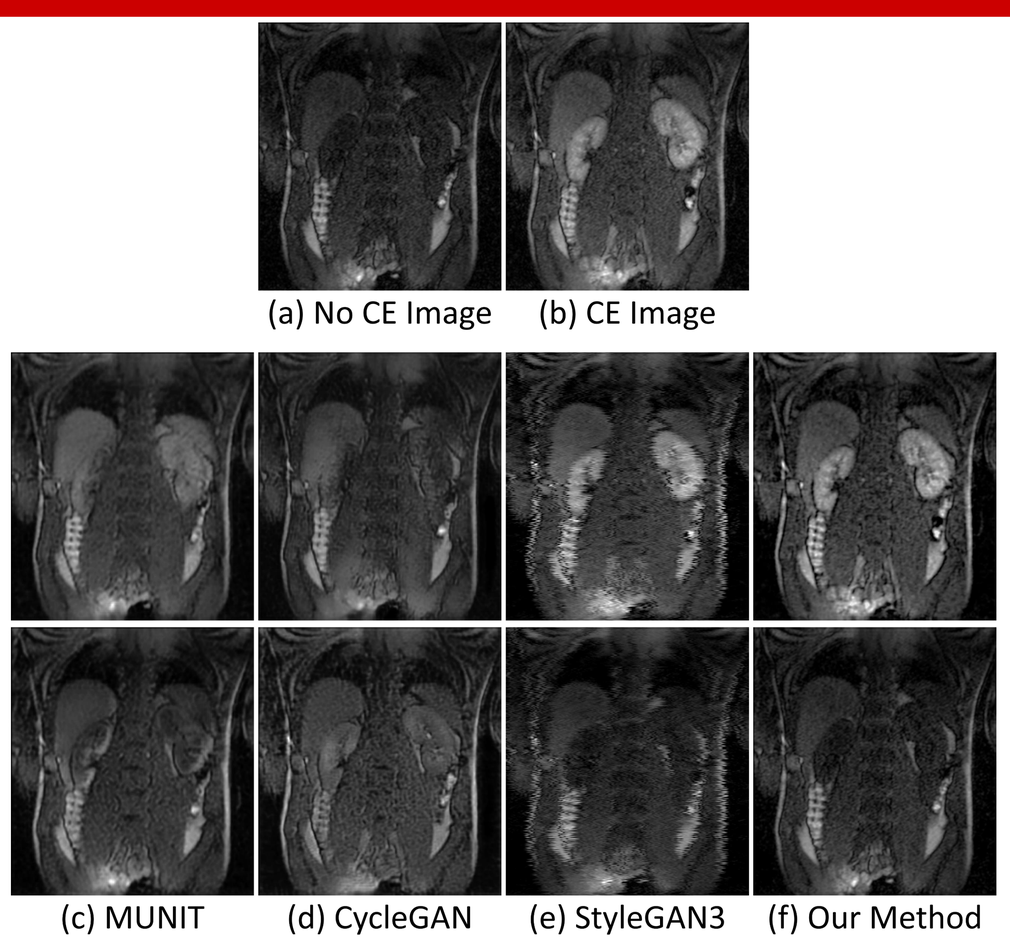

Figure 2: Results of style transfer with kidney DCE-MRI. The first row (a and b) shows the input images. The second row shows the results when (a) is used as the content image and (b) as the style. The third row shows results when (b) is used as the content image and (a) as the style.

Conclusions

- Our method learns information from time series data to accurately capture key structures from content images whilst localising the addition or removal of CE.
- We outperform other popular style transfer techniques by having sharper images, the expected noise characteristics of MRI as well as having better localised contrast enhancement.

References

[1] Huang, X., Liu, M.Y., Belongie, S., Kautz, J.: Multimodal unsupervised image-to- image translation. In: ECCV (2018)

[2] Zhu, J.Y., Zhang, R., Pathak, D., Darrell, T., Efros, A.A., Wang, O., Shechtman, E.: Toward multimodal image-to-image translation. NeurIPS (2017)

[3] Karras, T., Aittala, M., Laine, S., H¨ark¨onen, E., Hellsten, J., Lehtinen, J., Aila, T.: Alias-free generative adversarial networks. In: Proc. NeurIPS (2021)