A Generative Framework for Predicting Myocardial Strain from Cine-Cardiac Magnetic Resonance Imaging

Nina Cheng[1], Rodrigo Bonazzola[1], Nishant Ravikumar[1(✉)], and Alejandro F. Frangi[1,2]

[1] CISTIB, Centre for Computational Imaging and Simulation Technologies in Biomedicine, School of Computing and LICAMM, School of Medicine, University of Leeds, Leeds, UK
N.Ravikumar@leeds.ac.uk

[2] Department of Cardiovascular Sciences, and Department of Electrical Engineering, KU Leuven, Leuven, Belgium

Abstract. Myocardial strain is an important measure of cardiac performance, which can be altered when ejection fraction (EF) and other ventricular volumetric indices remain normal, providing an additional indicator for early detection of cardiac dysfunction. Cardiac tagging MRI is the gold standard for myocardial strain quantification but requires additional sequence acquisition and relatively complex post-processing procedures, which limit its clinical application. In this paper, we propose a framework for learning a joint latent representation of cine MRI and tagging MRI, such that tagging MRI can be synthesised and used to derive myocardial strain, given just cine MRI as inputs. Specifically, we use a multi-channel variational autoencoder to simultaneously learn features from tagging MRI and cine MRI, and project the information from these distinct channels into a common latent space to jointly analyse the multi-sequence data information. The inference process generates tagging MRI using only cine MRI as input, by conditionally sampling from the learned latent representation. Finally, automated tag tracking was performed using a cardiac motion tag tracking network on the generated tagging MRI, and myocardial strain was estimated. Experiments on the UK Biobank dataset show that our proposed framework can generate tagging images from cine images alone, and in turn, can be used to estimate myocardial strain effectively.

Keywords: Cardiac tagging MRI · Cardiac cine MRI · Myocardial strain estimation · Convolutional neural network · Machine learning

1 Introduction

Myocardial strain is used to quantitatively assess local myocardial deformation and is an important indicator regional cardiac function. Cardiac magnetic resonance tagging

N. Ravikumar and A. F. Frangi—Joint last authors.

G. Yang et al. (Eds.): MIUA 2022, LNCS 13413, pp. 482–493, 2022.
https://doi.org/10.1007/978-3-031-12053-4_36

(CMR-tagging) is a non-invasive imaging technique, and the current gold standard for quantification of local measures of myocardial motion/deformation such as strain and strain rate [1]. Strain reflects the rate of change in the length of the myocardium (typically along the circumferential or radial directions) across the cardiac cycle. Strain rate refers to the strain per unit time [2]. CMR-tagging sequences use selective radiofrequency pulses (e.g. SPAMM, DANTE) to superimpose the myocardium with tags or grids of tags that are subsequently tracked as the heart deforms across the cardiac cycle. The clinical application of tagging MRI, however, has been hindered by the need for acquiring additional sequences which increases scan time, cumbersome manual/semi-automatic post-processing steps required to derive regional strain measurements, and limited validation. Consequently, CMR-tagging has not been widely adopted in clinical settings due to the lack of rapid analysis techniques. Although several methods for tag-tracking have been proposed previously [3–6], most approaches rely on manual/semi-automatic landmark localisation and/or segmentation to derive strain measures by tracking superimposed tags across the cardiac cycle. We propose to synthesise tagged-CMR from cine-CMR images and use the synthesised images to quantify strain. This will help reduce scan time and facilitate myocardial strain quantification directly from routinely acquired cine-CMR images.

Recently, machine learning methods, especially convolutional neural networks (CNN) in deep learning [7], have shown promise in the field of medical image analysis and understanding, including automated cine-CMR image analysis, enabling detection and diagnosis of cardiovascular diseases, extraction of quantitative clinical indices and biomarkers, among others. Deep learning-based generative models such as generative adversarial networks (GANs) and variational autoencoders (VAEs) [8] have been explored extensively for image synthesis applications. We propose to tackle the limitations inhibiting wider clinical adoption of CMR-tagging imaging for strain quantification by learning to synthesise tagged-CMR images from cine-CMR images, in a patient-specific manner. We formulate this image synthesis problem in a probabilistic manner, as one of learning a joint latent representation across both types of images for subjects within a population. To this end, we utilise a Bayesian approach, namely, a multi-channel sparse variational autoencoder (mcVAE) [9] to learn a joint latent space given both channels of information for each subject in a population, i.e. their respective tagged-CMR and cine-CMR images. The latent representation learned using mcVAE subsequently enables generation of a tagged-CMR image for new/unseen subjects during inference, given just their cine-CMR image as input.

To the best of our knowledge this is the first study to investigate synthesis of tagged-CMR images from cine-CMR images. Only one other study has investigated tagged-MR image to cine-MR synthesis [10], for images of the throat. In Liu, et al. the authors proposed a dual-cycle constrained bijective VAE-GAN to synthesise cine-MR images from their paired tagged-MR images for each subject. Additionally, we go beyond just image synthesis to demonstrate that the tagged-CMR images synthesised from cine-CMR images can be used to quantify myocardial strain. The synthesised tagged-CMR images are used to quantify circumferential and radial strain across the cardiac cycle using a cardiac motion tracking network proposed in [11]. The proposed approach was trained and validated on data from the UK Biobank population imaging database.

2 Dataset and Pre-processing

2.1 Dataset

We utilised data from the UK Biobank [12] throughout this study, including two cardiac MR sequences for each subject, namely, cine-CMR and tagged-CMR. Participants were recruited to the initial assessment visit from 2006 to 2010, the first repeat assessment visit from 2012 to 2013, and the first imaging visit in 2014. The rationale and protocol for cardiac MRI examinations are described in [13], and this study was reviewed and approved by the Northwestern Research Ethics Committee (REC reference number: 06/MRE08/65). Scans were acquired using a 1.5 T scanner to obtain steady-state short-axis free precession cine-CMR and tagged-CMR sequences. All participants gave written informed consent.

2.2 Data Preprocessing

For cine-CMR images, we used a CNN model [14] to identify and crop regions of interest (ROI) around the heart. The purpose of cropping is to reduce the computational time and resources required to train our model. As the number of short-axis slices in the CMR data from the UK Biobank generally varies between 7 and 14 slices, the short-axis image stack was resampled to volume of 15 slices using cubic B-spline interpolation, with an isotropic resolution of 1 mm^3, then adjusted each slice of data at a common resolution of 128 $\times$ 128 pixels. Finally, we normalise the intensities in each image to a range of 0 to 1.

For the tagged-CMR images, we transformed the ROI coordinates obtained in the cine-CMR image space and cropped them. Coordinate transformation converts 2D image coordinates and 3D patient space coordinates, i.e. converts ROI coordinates to 3D world coordinates using the cine image, and then converts back to 2D coordinates using the tagging image. Following cropping, the tagged-CMR images were pre-processed in the same way as the cine-CMR images, including spatial resampling and intensity normalisation.

The coordinate transformation process is shown in Fig. 1. The key to determining the ROI in the tagging image is that the 3D spatial coordinates of the two sequence images are the same, and the process performs coordinate transformation on all samples in the experiment.

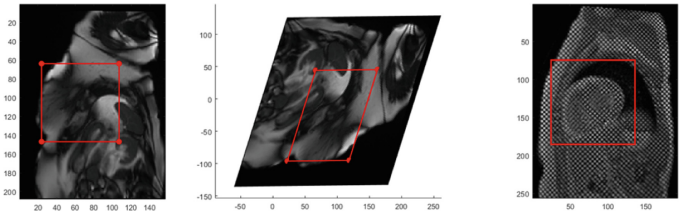

Fig. 1. Left: 2D cine image ROI; Intermediate: 3D cine image ROI; Right: 2D tagging image ROI. The CMR images were reproduced with permission from UK Biobank.

3 Methodology

3.1 System Overview

The deep learning framework in this paper consists of the following steps: (a) After multi-sequence image pre-processing, the mcVAE is trained using cine-CMR and tagged-CMR images to learn a joined latent space and to reconstruct each channel of information (i.e. each CMR sequence); (b) during inference, the trained mcVAE model is used to synthesise tagged-CMR images from unseen (test) cine-CMR images; (c) using the synthesised tagged-CMR images from (b), cardiac motion is tracked using a ResNet CNN model pre-trained on synthetic data [15]. Finally strain analysis is performed using the predicted motion trajectories, and radial and circumferential strains are estimated (Fig. 2.).

The neural network is trained using a large amount of synthetic data generated from natural images and verified by a cardiac phantom model with known strain. The procedure pre-trains it with the landmarks defined at the first time point (ED), then predicts the landmark at subsequent time points and uses it to compute the motion trajectory and the resulting deformation field. The network outputs the predicted motion path and finally calculates the strain from the deformation field.

3.2 Multi-channel Variational Autoencoder

Our method is based on a multi-channel sparse variational autoencoder [9], where two encoder/decoder network pairs are trained simultaneously, using on one of the two sequences (i.e. cine-CMR and tagged-CMR) as input channels to each pair. The two encoder-decoder network pairs share a latent space enabling a joint latent representation to be learned for the input channels/sequences. The network architecture is shown in Table 1.

In our mcVAE network, each subject's data $\boldsymbol{x}$ includes two information channels from the two CMR sequences, and the latent space $\boldsymbol{s}$ is represented by an l -dimensional vector shared between each data point x. The generative process for the observed channels of information can be described by,

$$\boldsymbol{s} \sim p(\boldsymbol{s}),$$

$$\boldsymbol{x}_c^{\sim} p(\boldsymbol{x}_c|\boldsymbol{s}, \theta_c), \quad c \, in \, 1, \, \ldots, \, C, \tag{1}$$

where, $p(\boldsymbol{s})$ is the prior distribution, and $\boldsymbol{x}_c \sim p(\boldsymbol{x}_c|\boldsymbol{s}, \theta_c)$ represents the likelihood distribution of the observation. Each likelihood function belongs to distribution family $\mathcal{P}$, which is parameterised by parameter set $\theta = \{\theta_1, \ldots, \theta_c\}$.

The inference process can determine the common latent space, and each channel generate the observation data from this latent space. $p(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)$ represents the posterior distribution of the problem, and variational inference is often used to calculate the approximate posterior. Each channel contributes information about the distribution of latent variables, and the posterior distribution $q(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)$ is approximated by a single

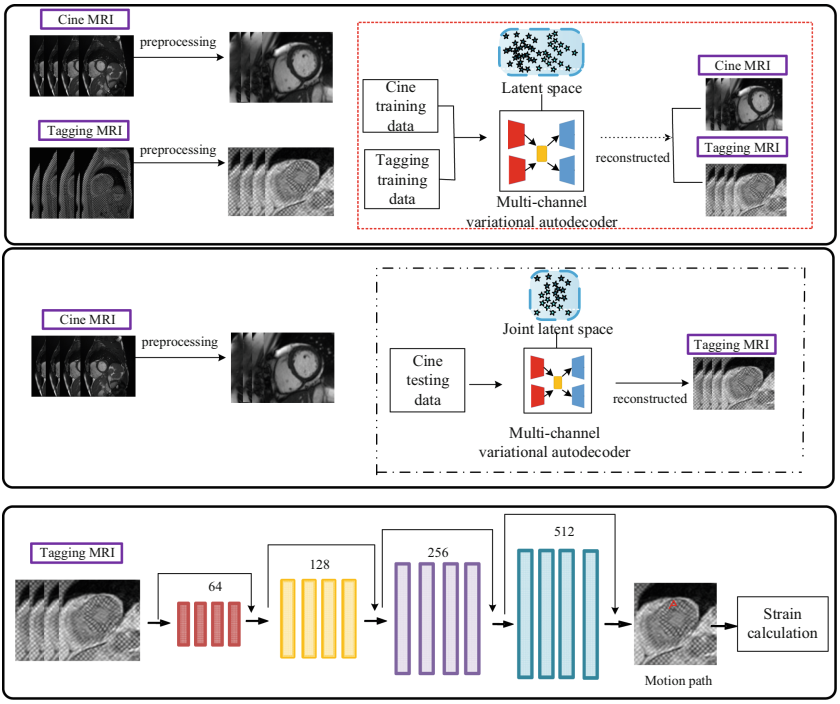

Fig. 2. Overview of machine learning framework that joint cine MRI and tagging MRI to estimate tagging MRI and myocardial strain using cine MRI alone.

channel and parameter c. Since each channel provides different approximations, constraints are imposed to force each q $(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)$ to be as close as possible to the target posterior distribution. The distribution is measured using Kullback-Leibler (KL) divergence, specified as follows:

$$arg \min_{q \in Q} \mathbb{E}_c[D_{KL}(q(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)||p(\boldsymbol{s}|\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_c, \theta))], \tag{2}$$

where the $q(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)$ belong to a distribution family $\mathcal{Q}$ parametrized by the parameters $\phi = \{\phi_1, \phi_2, \ldots, \phi_c\}$. The quantity $\mathbb{E}_c$ is the average over all channels. Minimising Eq. (3.2) is equivalent to the optimization of the following formula:

$$\mathcal{L}(\theta, \phi, \boldsymbol{x}) = \mathbb{E}_c[L_c - \mathcal{D}_{KL}(q(\boldsymbol{s}|\boldsymbol{x}_c, \phi_c)||p(\boldsymbol{s}))], \tag{3}$$

where, $L_c = \mathbb{E}_{q(\boldsymbol{s}|\mathbf{x}_c, \phi_c)} \sum_{i=1}^{C} \ln p(\boldsymbol{x}_i|\mathbf{s}, \theta_i)$ is the expected log-likelihood of decoding each channel from the latent representation. After learning the joint latent space, L_c can reconstruct multi-channel information only from one single channel, that is, only use the coding information of a single channel to reconstruct other channels, or both channels.

In our study, once trained, the mcVAE network was used to reconstruct tagged-CMR images given only cine-CMR images as inputs. The synthesised tagged-CMR images were subsequently used to calculate myocardial strain.

Table 1. Network architecture for each channel of a multi-channel variational autoencoder.

Channel	Encoder	Decoder
Tagging	Input 128 * 128 * 25	Latent variables, FC, Reshape
	Five times: 3 * 3 conv, ReLU, Dropout 0.15	Five times: 3 * 3 ConvTranspose2d, LeakyReLU (alpha 0.2), Dropout 0.15
	FC output layer, Latent variables	Tanh activation, Reconstructed tagging images
Cine	Input 128 * 128 * 10	Latent variables, FC, Reshape
	Five times: 3 * 3 conv, ReLU, Dropout 0.15	Five times: 3 * 3 ConvTranspose2d, LeakyReLU (alpha 0.2), Dropout 0.15
	FC output layer, Latent variables	Tanh activation, Reconstructed Cine images

3.3 Cardiac Motion Tag Tracking

We use a CNN-based approach [15] to track cardiac motion in synthesised tagged-CMR images. First, we created a synthetic training set comprising 1 million patches and trained a ResNet (ResNet-18) model on the same to learn to predict spatial position vectors of the estimated motion path, enabling tag tracking automatically.

Synthetic images for the training process were randomly sampled from ImageNet [16], and the process for creating the synthetic training set is the same as that proposed in [15]. The tag tracking network uses a modified version of ResNet (ResNet-18), with spatiotemporal (2 + 1)D convolutional layers [17] and CoordConv channels [18] for each convolutional block. The last layer of the network is a fully connected layer with linear activations, which outputs the spatial position vector of the estimated motion path. The optimization process uses stochastic gradient descent, the learning rate is adjusted using cosine annealing methods [19], 90% of the data is used for model training and 10% is used for model inference, and 10 patches are randomly selected per image in the synthetic image, so the training contains 1 million patches. The error between the predicted motion path and the real motion path is measured using the mean squared error (MSE) loss function.

Strain analysis needs to manually mark the intersection points in the end diastole (ED) time point. The model predicts and obtains the motion paths of other time points in all time ranges, essentially predicting the landmark displacement at subsequent time points, and calculates the strain through these paths of all points. Calculate strain by first fitting the deformation map using a Gaussian radial basis function (RBF) [19] with the shape parameter is twice of the tracking points spacing, then solving for the analytical derivative of the RBF in order to calculate the Green-Lagrange strain tensor, the formula for the Green-Lagrange strain is as follows:

$$\epsilon(t) = \frac{1}{2}\left(\frac{L_t^2 - L_0^2}{L_0^2}\right) \tag{4}$$

where L_t represents the segment length at any frame t, and L_0 represents the initial length.

4 Experiments and Results

In this study, we trained a mcVAE for jointly learning cine MRI and tagging MRI, which can generate tagging MRI from cine MRI, and realize the estimation of myocardial strain in the inference stage.

4.1 Experimental Setup

We use total of 535 subjects' (13375 slice pairs) images in our experiments, 60% of which were selected for training, 20% for validation, and 20% for evaluation. Cardiac cine and tagging MRI images from the UK Biobank database were used in our experiments, and we resized these images to 128 × 128 for a fair comparison. Our learning framework is implemented using the PyTorch deep learning toolbox [20], and the training takes about 9 h on NVIDIA V100 GPUs. After obtaining the generated tagging images, the tag points are obtained using a pretrained tag tracking network, and the radial and circumferential strains are calculated based on the displacement of the tag points, which takes about 0.1 s. Specifically, we use the Adam optimizer for training. The learning rate is set to 5e−4 and the weight decay is set to 1e−5.

4.2 Qualitative Evaluations

Figure 3 shows the generated results of the multi-channel VAE, which shows that the proposed framework successfully generates tagging MRI images from only cine MRI images, which is consistent with the target original tagging MR images. The generated images achieve visually pleasing results with good structural consistency with respect to the original images. For quantitative evaluation, we employed evaluation metrics widely used in images: root mean square error (RMSE) [21], structural similarity index measure (SSIM) [22], and peak signal-to-noise ratio (PSNR) [23]. Table 2 lists the numerical comparisons of the test dataset results of simultaneously inputting cine and tagging images to generate cine and tagging images, as well as reconstructing tagging images using only cine images, and reconstructing cine images using only tagging images. We note that the generation results with a single channel are inferior to the two channels generation results, which is to be expected.

Table 2. Quantitative numerical comparisons of generated results.

	RMSE	SSIM	PSNR
Tagging	0.14 ± 0.01	0.84 ± 0.21	28.13 ± 4.85
Cine	0.21 ± 0.04	0.69 ± 0.27	23.91 ± 6.68
Only cine to tagging	0.17 ± 0.02	0.72 ± 0.24	24.57 ± 5.84
Only tagging to cine	0.31 ± 0.10	0.65 ± 0.29	20.44 ± 7.51

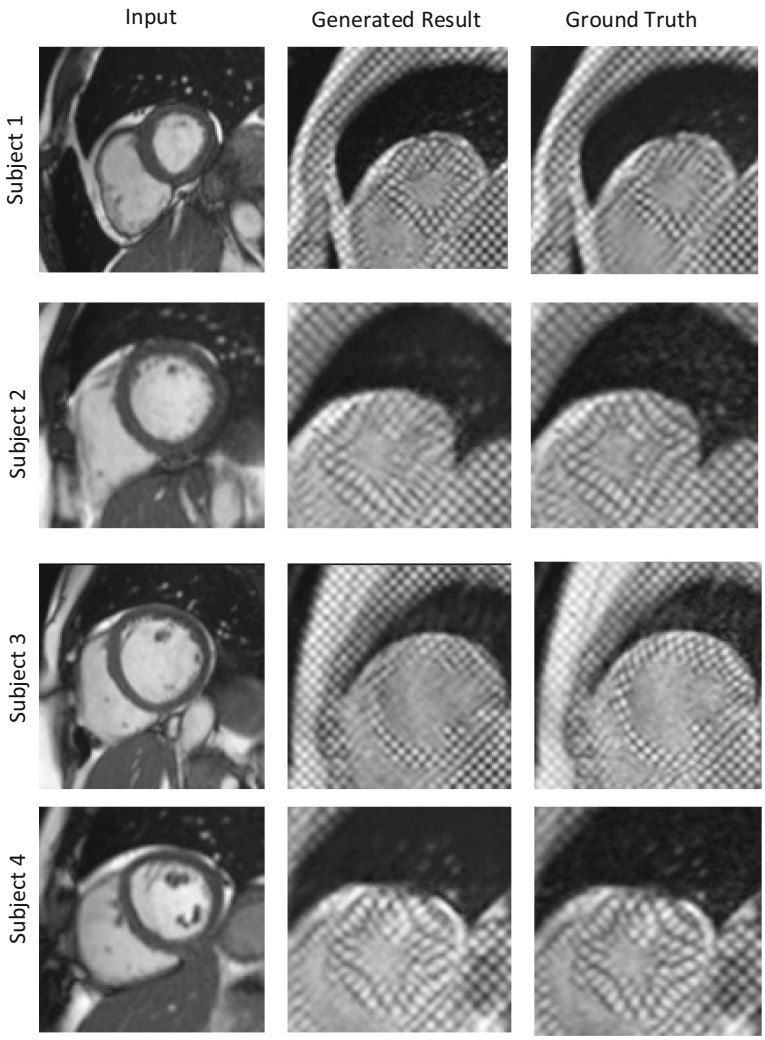

Fig. 3. Examples comparing generated and original tagged-CMR images from different subjects. The CMR images were reproduced with permission from UK Biobank.

4.3 Tag Tracking

Our tag tracking method uses a Resnet-18 neural network trained on a synthetic dataset of tagged images. This approach has been shown to accurately track tag intersections when applied to in vivo data. The points to be tracked are first selected in the ED time frame of the tagging images and fed into the neural network. The network outputs the Lagrangian displacement vector of the tracking point at each time point, and the network can track any point in the image, not just along the tagged line or the intersection of

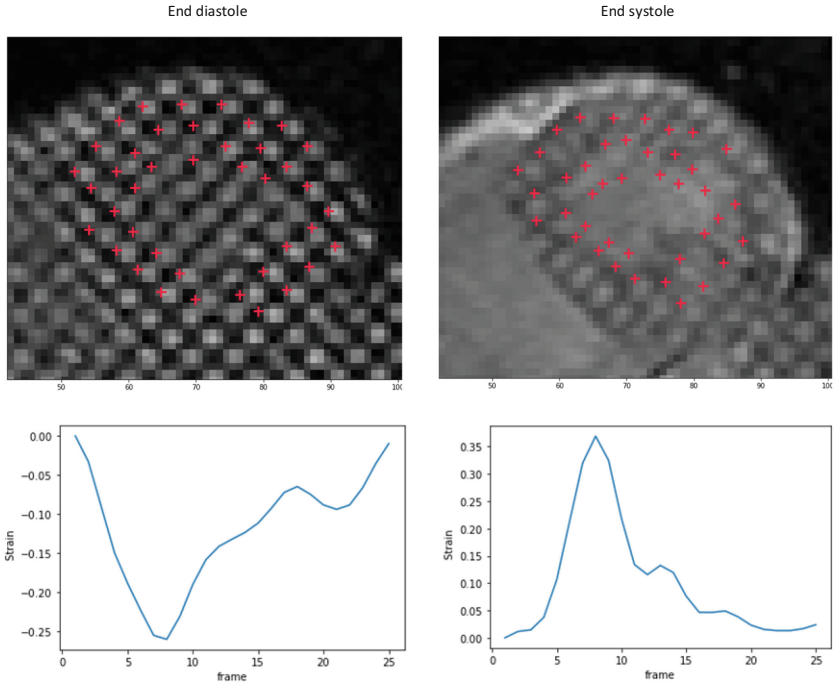

Fig. 4. Examples of tag tracking estimated during end-diastole (ED) and end-systole (ES) (top row) and strain calculations for the entire cine; circumferential strain (bottom left) and radial strain (bottom right).

the tagged lines. Figure 4 shows an example of tag detection and tracking produced at ED and ES from the generated images, along with strain estimates (circumferential and radial) at all time frames.

4.4 Strain Analysis

To validate the generative model, we performed strain analysis on the original tagging images and those generated from cine images, calculating circumferential and radial strains. Bland-Altman analysis [24] was used to quantify the agreement of two measurements, the result is shown in Fig. 5. These results show that the mean difference between circumferential strain and radial strain is close to zero, and most cases are within 95% agreement. Some outliers indicate cases of large error, which require further investigation that we will undertake in future work. The results obtained for radial strain were worse than circumferential strain, which is consistent with previous studies showing reduced accuracy in calculating radial strain using tagging MRI [25].

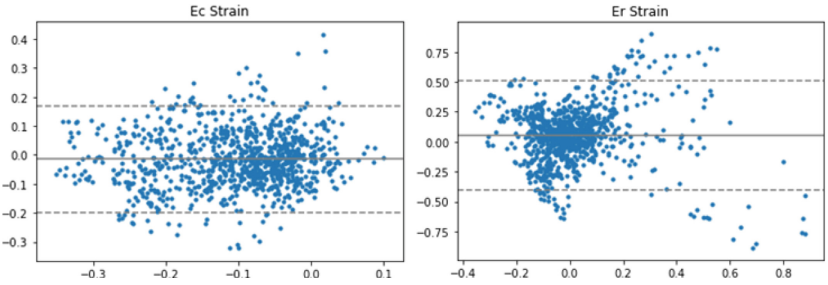

Fig. 5. Bland-Altman plot of end-systolic strain. The strain values obtained from the generated tagging images are compared with the strain values of the original images. Left plot shows mean circumferential strain, right plot shows mean radial strain, solid line indicates mean difference; dashed line indicates 95% concordance limit (mean ± 1.96 * Standard Deviation [SD]).

In addition to this, we randomly select the original and generated images of ten test results, and using the semi-automatic approach described previously for tag tracking on the original and generated images, calculated the strain, and plotted the errorbars for the estimated strains across the cardiac cycle. The results are shown in Fig. 6, and the circumferential and radial strains of the original and generated images are consistent.

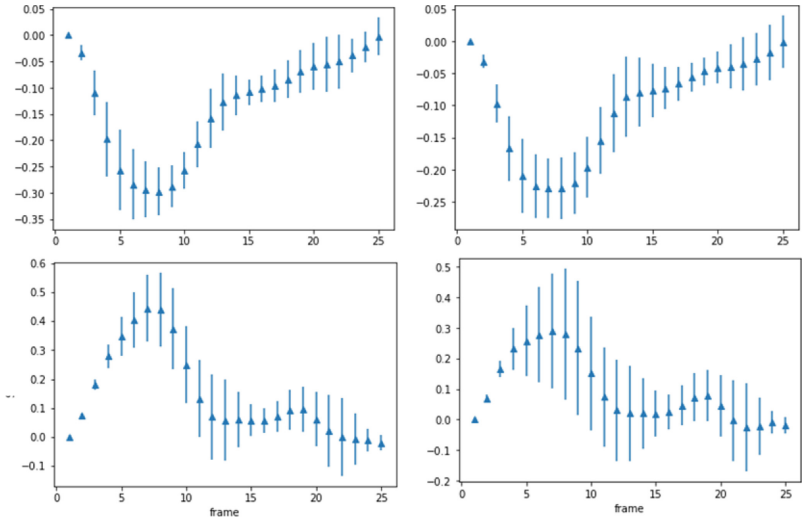

Fig. 6. Circumferential and radial strain errorbar for the original and generated images of the ten test results, top left: original images circumferential strain; top right: synthetic images circumferential strain; bottom left: original images radial strain; bottom right: synthetic images radial strain.

5 Conclusion

We propose a novel deep learning framework to jointly learn tagging MRI and cine MRI to generate tagging MRI and estimate myocardial strain using cine MRI alone. The generative process uses a multi-channel variational autoencoder to project data from different sequences into a common latent space data, which enables the reconstruction of one channel's information using only the encoded information from the other channel, i.e. of tagging-MRI from cine-MRI. After obtaining the resulting tagging MR images, automatic tag tracking was performed using a cardiac motion tag tracking neural network, and radial and circumferential strain was estimated across the full cardiac cycle, achieving comparable results to strains quantified using the original tagging MR images. Estimating myocardial strain from tagging images is a challenging problem, and our designed framework facilitates strain quantification using just cine-MR images as input. To the best of our knowledge, this proof-of-concept study is the first of its kind and opens up new avenues for research in myocardial strain quantification.

Acknowledgements. This research was conducted using data from the UK Biobank under access application 11350. AFF is funded by the Royal Academy of Engineering (INSILEX CiET1819\19), Engineering and Physical Sciences Research Council (EPSRC) programs TUSCA EP/V04799X/1, and the Royal Society Exchange Programme CROSSLINK IES\NSFC\201380, and was supported partly by China Scholarship Council Studentship with the University of Leeds.

References

1. Ibrahim, E.-S.H.: Myocardial tagging by cardiovascular magnetic resonance: evolution of techniques–pulse sequences, analysis algorithms, and applications. J. Cardiovasc. Magn. Reson. **13**(1), 1–40 (2011). https://doi.org/10.1186/1532-429X-13-36
2. Sutherland, G.R., Di Salvo, G., Claus, P., D'hooge, J., Bijnens, B.: Strain and strain rate imaging: a new clinical approach to quantifying regional myocardial function. J. Am. Soc. Echocardiogr. **17**(7), 788–802 (2004). https://doi.org/10.1016/j.echo.2004.03.027
3. Wu, L., Germans, T., Güçlü, A., Heymans, M.W., Allaart, C.P., van Rossum, A.C.: Feature tracking compared with tissue tagging measurements of segmental strain by cardiovascular magnetic resonance. J. Cardiovasc. Magn. Reson. **16**(1), 1–11 (2014). https://doi.org/10.1186/1532-429X-16-10
4. Götte, M.J., Germans, T., Rüssel, I.K., Zwanenburg, J.J., Marcus, J.T., van Rossum, A.C., et al.: Myocardial strain and torsion quantified by cardiovascular magnetic resonance tissue tagging: studies in normal and impaired left ventricular function **48**(10), 2002–2011 (2006). https://doi.org/10.1016/j.jacc.2006.07.048
5. Guttman, M.A., Prince, J.L., McVeigh, E.R.: Tag and contour detection in tagged MR images of the left ventricle **13**(1), 74–88 (1994). https://doi.org/10.1109/42.276146
6. Young, A.A., Kraitchman, D.L., Dougherty, L., Axel, L.: Tracking and finite element analysis of stripe deformation in magnetic resonance tagging. IEEE Trans. Med. Imaging **14**(3), 413–421 (1995). https://doi.org/10.1109/42.414605
7. O'Shea, K., Nash, R.: An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458 (2015)
8. Pu, Y., Gan, Z., Henao, R., Yuan, X., Li, C., Stevens, A., et al.: Variational autoencoder for deep learning of images, labels and captions. In: 29th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain, p. 29 (2016). https://doi.org/10.48550/arXiv.1609.08976

9. Antelmi, L., Ayache, N., Robert, P., Lorenzi, M.: Sparse multi-channel variational autoencoder for the joint analysis of heterogeneous data. In: International Conference on Machine Learning, Long Beach, California (2019)
10. Liu, X., Xing, F., Prince, J.L., Carass, A., Stone, M., El Fakhri, G., et al.: Dual-cycle constrained bijective VAE-GAN for tagged-to-cine magnetic resonance image synthesis. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), IEEE, France (2021). https://doi.org/10.1109/ISBI48211.2021.9433852
11. Loecher, M., Hannum, A.J., Perotti, L.E., Ennis, D.B.: Arbitrary point tracking with machine learning to measure cardiac strains in tagged MRI. In: Ennis, D.B., Perotti, L.E., Wang, V.Y. (eds.) FIMH 2021. LNCS, vol. 12738, pp. 213–222. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78710-3_21
12. Biobank U: About uk biobank (2014)
13. Petersen, S.E., Matthews, P.M., Bamberg, F., Bluemke, D.A., Francis, J.M., Friedrich, M.G., et al.: Imaging in population science: cardiovascular magnetic resonance in 100,000 participants of UK Biobank-rationale, challenges and approaches **15**(1), 1–10 (2013). https://doi.org/10.1186/1532-429X-15-46
14. Zheng, Q., Delingette, H., Duchateau, N., Ayache, N.: 3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation. IEEE Trans. Med. Imaging **37**(9), 2137–48 (2018). https://doi.org/10.48550/arXiv.1804.09400
15. Loecher, M., Perotti, L.E., Ennis, D.B.: Using synthetic data generation to train a cardiac motion tag tracking neural network. Med. Image Anal. **74**,1022–1023 (2021). https://doi.org/10.1016/j.media.2021.102223
16. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami. IEEE (2009). https://doi.org/10.1109/CVPR.2009.5206848
17. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., Paluri, M.: A closer look at spatiotemporal convolutions for action recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, USA, pp. 6450–6459 (2018)
18. Liu, R., Lehman, J., Molino, P., Petroski Such, F., Frank, E., Sergeev, A., et al.: An intriguing failing of convolutional neural networks and the coordconv solution, 31 (2018)
19. Loshchilov, I., Hutter, F.: SGDR: stochastic gradient descent with warm restarts. In: 5th International Conference on Learning Representations (ICLR 2016), Toulon, France (2016)
20. Prakash, K.B., Kanagachidambaresan, G.R. (eds.): Programming with TensorFlow. EICC, Springer, Cham (2021). https://doi.org/10.1007/978-3-030-57077-4
21. Wang, W., Lu, Y.: Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing, (2018). https://doi.org/10.1088/1757-899X/324/1/012049
22. Dosselmann, R., Yang, X.D.: A comprehensive assessment of the structural similarity index. Signal Image Video Process. **5**(1), 81–91 (2011). https://doi.org/10.1007/s11760-009-0144-1
23. Poobathy, D., Chezian, R.M.: Edge detection operators: peak signal to noise ratio based comparison. IJ Image Graph. Signal Process. **10**, 55–61 (2014). https://doi.org/10.5815/ijigsp.2014.10.07
24. Bunce, C.: Correlation, agreement, and Bland–Altman analysis: statistical analysis of method comparison studies. Am. J. Ophthalmol. **148**(1), 4–6 (2009). https://doi.org/10.1016/j.ajo.2008.09.032
25. Young, A.A., Li, B., Kirton, R.S., Cowan, B.R.: Generalized spatiotemporal myocardial strain analysis for DENSE and SPAMM imaging. Magn. Reson. Med. **67**(6), 1590–159 (2012). https://doi.org/10.1002/mrm.23142