Diffuse myocardial diseases can be diagnosed using $T_1$ mapping technique based on $T_1$ relaxation times from MRI data. The $T_1$ relaxation parameter is acquired through pixel-wise fitting of the MRI signal. Hence, pixels misalignment resulted by cardiac motion leads to an inaccurate $T_1$-mapping. Therefore, registration is needed. However, due to the intensity differences between the different time-points, recent unsupervised deep-learning approaches based on minimizing the mean-squared-error (MSE) between the images cannot be utilized directly. To overcome this challenge, we propose a new double-stage method, in which a style-transfer is used to harmonize the signal intensities over time, followed by an unsupervised deep-learning based minimization of the MSE between the images. We evaluated our approach on a publicly available cardiac T1 mapping database of 210 subjects. Our approach achieved the best median model-fitting $R^2$ compared to baseline methods (0.9794, vs. 0.9651/0.9744/0.9756) and $T_1$ values which are much closer to the the expected myocardial $T_1$ value. Furthermore, both metrics have less variability compared to the other methods.

1 Introduction

$T_1$ relaxation time is a key source of soft tissue contrast in MRI. Mapping of each pixel $T_1$ relaxation time, can depict relatively small variations within the cardiac muscle, highlight tissue pathology such as acute myocardial infarction, chronic scar tissue, or detect fatty infiltration. (14)

Creating $T_1$ mapping (Fig.1) requires a time series of aligned images in which each pixel describes the same tissue across time. Nonetheless, during image acquisition there are inevitable cardiac motion, respiratory motion and involuntary patient motion (15). Therefore there is a great need of image registration before curve fitting (4). In recent years, additional to traditional and to machine learning methods, deep learning-based methods have emerged (2). Usually, the standard metric for deep learning registration, where the images for alignment differ in their appearance, is Mutual Information (MI) (1,5,13,16). Unfortunately, this metric is far behind within-contrast metrics such as Normal Cross Correlation (NCC) and Mean Square Error (MSE) in terms of accuracy (6,7).

To tackle the motion-related challenges involved in cardiac T1 mapping, we propose changing the style of all the time series images to the same style using a style transformer approach (11). Then, we
Figure 1: Schematic description of $T_1$ mapping for a single pixel. (a) Myocardial images at 11 sequential time points (displayed at their absolute value). (b) Fitting an inversion recovery curve of the magnetization $M_z$ over different time points $t$ and extracting the corresponding $T_1$ and $M_0$ parameters. (c) Displaying $T_1$ mapping for all the pixels in the image.

Figure 2: (a) Illustration of the training process. First, harmonization of the fixed and moving images using the style network, and afterward, registration of the harmonized images using voxelmorph. (b) Illustration of the inference time - First, harmonization of the fixed and moving images using the style network, calculating the deformation field for the harmonized pair images and applying it to the original moving image.

We propose using the harmonized images as input to the within-contrast registration network based on MSE loss (2) that fits cases of images with similar contrast and intensity distributions.

## 2 Methods

Our proposed method consists of two sequential stages. The first stage is a preparation stage, adjusting the color style of each image in the MRI sequence according to the first image color style using StyleGAN architecture (9; 11). Once all the images consist of the same color style, the registration can be learned using voxelmorph (2) based on the mean square error indices as the loss function. The voxelmorph architecture is used for pairwise registration. Aligning a moving image $I_M$ to a fixed image $I_F$ by yielding an optimized deformation field $\phi$ and a warped image $I_M \circ \phi$.

During the network training, firstly, for each patient, every image in the N length sequence images $I_M^t$, $\{ t \in 1, \cdots, N \}$ harmonized to the first time point image $I_F^0$ as a reference style, using StyleGAN as the style transfer network. Following, the voxelmorph network was trained on the harmonized images, while for each patient, the moving image is one of $I_M^{t,styled}$, $\{ t \in 1, \cdots, N \}$ and the fixed image is $I_F^0$. The above training process is demonstrated in Fig.2(a). In inference time as demonstrated in Fig.2(b), the deformation field is calculated according to the harmonized moving image, but the yielded deformation field is applied on the original moving images.

In order to evaluate the proposed metric, a publicly available myocardial $T_1$ mapping dataset was used. (3) The dataset includes 210 patients with known or suspected cardiovascular diseases. For each patient, 5 slices at 11 time points were available with their corresponding myocardial segmentation map.

*Equal contribution
2.1 Implementation details

The implementation was based on PyTorch and ran on NVIDIA Tesla V100 GPU with 32G RAM. The style transfer network was trained for 100k iterations with $\lambda_{cycle} = 100$, $\lambda_{style} = 20$, $\lambda_{divergence} = 1$, style dimension and batch size of 4. The registration network was trained for 50k iterations with $\lambda_{smooth} = 0.003$ and a batch size of 64. For both network, ADAM optimizer was used with a learning rate of $1 \cdot 10^{-4}$.

3 Results

In order to evaluate the effectiveness of our two-staged method, the results were compared with the state of the art deep learning algorithms for medical image registration: pairwise VoxelMorph with mutual information loss [2] and pairwise SynthMorph [6], as well as with non-registered images. The four methods were evaluated comparing the pixels $R^2$ and $T_1$ value. The mean and median values are presented in Table.1 and their values distribution is presented in Fig.3.

$T_1$ value is vary according to the magnetic field, the imaging protocol, gender and the specific cardiovascular disease. [10] [12]. Moreover, according to the imaging protocol, the images are relatively aligned with no significant movements during time. Therefore, it is reasonable to deduce that the original images estimated median $T_1$ value is closer to the actual median value. For the VoxelMorph algorithm, although the mean $R^2$ was higher than the non-registered images, the median $R^2$ was lower, indicates an higher variation in the estimated $T_1$ results. The proposed method and the SynthMorph, both, have reasonably close $T_1$ value and high $R^2$ value with less variability. Nonetheless, for both criteria, our proposed method is slightly better.

![Figure 3: 2D boxplot comparing $R^2$ and $T_1$ value for different registration methods. (left) zoom out, (right) zoom in to the interquartile interval. Our StyleReg method achieves the highest median $R^2$ value while keeping the $T_1$ around the median of the original $T_1$ values. The full circles are the intersection between $T_1$ and $R_2$ medians.](image)

Table 1: Evaluation of the four algorithms according to $R^2$ and $T_1$ mean and median results.

<table>
<thead>
<tr>
<th></th>
<th>$mean R^2 \pm std$</th>
<th>$median R^2$</th>
<th>$mean T_1 \pm std [ms]$</th>
<th>$median T_1 [ms]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.9267 ± 0.1136</td>
<td>0.9756</td>
<td>1155.6256 ± 244.6994</td>
<td>1131.5202</td>
</tr>
<tr>
<td>SynthMorph</td>
<td>0.9416 ± 0.0932</td>
<td>0.9744</td>
<td>1130.248 ± 223.1061</td>
<td>1110.2073</td>
</tr>
<tr>
<td>VoxelMorph</td>
<td>0.9443 ± 0.0722</td>
<td>0.9651</td>
<td>964.8632 ± 195.3005</td>
<td>976.1892</td>
</tr>
<tr>
<td>StyleReg</td>
<td>0.9436 ± 0.0919</td>
<td>0.9794</td>
<td>1154.3392 ± 210.2422</td>
<td>1136.8539</td>
</tr>
</tbody>
</table>

4 Conclusion

Varying appearance over the different time points, impose challenging registration task. In this work We propose a two-step method based on style transfer as preprocessing harmonization step before registration. Our approach benefits the within-contrast registration while we solve a cross-contrast problem. Our experimental results on a publicly available cardiac T1 mapping of 210 patients show that our process improves regressions $R^2$ and the accuracy of the $T_1$ values, which are much closer to the myocardial expected $T_1$ values compared to state-of-the-art methods. Our method was trained
on myocardial $T_1$w images but can extend to other quantitative MRI tasks, which includes different contrasts images registration.

References


