Recognition of Cardiac MRI Orientation via Deep Neural Networks and a Method to Improve Prediction Accuracy

Houxin Zhou

November 16, 2022

Abstract

In most medical image processing tasks, the orientation of an image would affect computing result. However, manually reorienting images wastes time and effort. In this paper, we study the problem of recognizing orientation in cardiac MRI and using deep neural network to solve this problem. For multiple sequences and modalities of MRI, we propose a transfer learning strategy, which adapts our proposed model from a single modality to multiple modalities. We also propose a prediction method that uses voting. The results show that deep neural network is an effective way in recognition of cardiac MRI orientation and the voting prediction method could improve accuracy.

1 Introduction

When medical images were stored, they may have different image orientations. In the further segmentation or computing, this difference may affect the results, since current deep neural network (DNN) systems generally only take the input and output of images as matrices or tensors, without considering the imaging orientation and real world coordinate. So it is crucial to recognize it before the further computing. This work is aimed to provide a study of the Cardiac Magnetic Resonance (CMR) image orientation, for fitting the coordinate system of human reality, and to develop an efficient method for recognition of the orientation.

Deep neural network has performed outstandingly in computer vision and gradually replaced the traditional methods. DNN also take a important role in medical image processing, such as image segmentation [3] and myocardial pathology analysis [2]. For CMR images, standardization of all the images is a prerequisite for further computing tasks based on DNN-based methodologies.

Most studies in the field of medical image processing have only focused on the further computing, so they have to spend a lot of manpower to do the preprocess. If we can auto adjust the images, it will save lots of time. Nevertheless, recognizing the orientation of different modality CMR images and adjusting them into standard format could be as challenging as the further computing tasks [1]. In a broad sense, recognition orientation is also a kind of image classification task, so DNN is of no doubts an effective way to solve this problem. In this work, we still use DNN as our main method.

In most image classification problem like ImageNet, we do some transformation to the image but these transformation do not change the label, for example we rotate a dog image and it’s still a dog image. However, the orientation could be changed if we do transformation like flipping to the images. In this work, we utilize this character to built a predicting model. Combining DNN and predicting model, we built a framework for recognition of Orientation.
This work is aimed at designing a DNN-based approach to achieve orientation recognition for multiple CMR modalities. Figure 1 presents the pipeline of our proposed method. The main contributions of this work are summarized as follows:

Figure 1: The pipeline of our proposed method. Firstly, we do some transformation to the image (See 2.1 and 2.3). Then, the images are taken as input to the DNN model and generate orientations. Finally, we apply inverse transformation to these orientations and vote to get the result.

1. We propose a scheme to standardize the CMR image orientation and categorize all the orientations for classification.
2. We present a DNN-based orientation recognition method for CMR image and transfer it to other modalities.
3. We propose a predicting method to improve the accuracy for orientation recognition.

2 Method

In this section, we introduce our proposed method for orientation recognition. Our proposed method is based on Deep Neural Network which was proved effective in image classification. In CMR Image Orientation Categorization, we improved the predicting accuracy by the following four steps. Firstly, we apply invertible operators to the image to get another 7 images. Then we predict these images and get 8 orientations. Finally, we use inverse transformation to these orientations and then vote to get the result.

2.1 CMR Image Orientation Categorization

Due to different data sources and scanning habits, the orientation of different CMR images may be different, and the orientation vector corresponding to the image itself may not correspond correctly. This may cause problems in tasks such as image segmentation or registration. Taking a 2D image as an example, we set the orientation of an image as the initial image and set the four corners of the image to \( \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \). Then the orientation of the 2D MR image may have the following 8 variations, which is listed in Table 1.
For each image $X_t$ from dataset, the target is to find the correct orientation from 8 classes. We denote the correct orientation of image $X_t$ as $i_t$ and denote the correctly adjusted image $X_t$ as $Y_t$. If we view each orientation as a function $f$, we can get a function set $\{f_i, i = 0 \ldots 7\}$ and $(X_t, Y_t, i_t)$ satisfy the equation $f_{i_t}(Y_t) = X_t$. In the following, function set $\{f_i, i = 0 \ldots 7\}$ is referred to as $F$.

Table 1: Orientation categorization of 2D CMR Images. Here, sx, sy and sz respectively denote the size of image in X-axis, Y-axis and Z-axis.

| No. | Operation                                        | Image | Correspondence of coordinates                |
|-----|--------------------------------------------------|-------|-----------------------------------------------|
| 0   | initial state                                    | 1 2   | Target[x,y,z]=Source[x,y,z]                   |
| 1   | horizontal flip                                  | 2 1   | Target[x,y,z]=Source[sx-x,y,z]                |
| 2   | vertical flip                                    | 3 4   | Target[x,y,z]=Source[x,sy-y,z]                |
| 3   | Rotate 180° clockwise                            | 4 3   | Target[x,y,z]=Source[sx-x,sy-y,z]             |
| 4   | Flip along the upper left-lower right corner     | 1 3   | Target[x,y,z]=Source[y,x,z]                   |
| 5   | Rotate 90° clockwise                             | 3 1   | Target[x,y,z]=Source[sx-y,x,z]                |
| 6   | Rotate 270° clockwise                            | 2 4   | Target[x,y,z]=Source[y,sy-x,z]                |
| 7   | Flip along the bottom left-top right corner      | 4 2   | Target[x,y,z]=Source[sx-y,sy-x,z]             |

2.2 Deep Neural Network

Suppose given image $X_t$, $X_t$ is then normalized. We denote the processed $X_t$ as $X'_t$. CNN take the $(X'_t, i_t)$ as input. In the proposed framework, the orientation recognition network consists of 3 convolution layers and 2 fully connected layers. The orientation predicted is denoted as $\hat{O}_i$. We use the standard categorical loss to calculate the loss between predicted orientation $\hat{O}_i$ and orientation label $O_i$. The orientation loss is formulated as below:

$$L_{\text{orientation}} = \sum_{i=1}^{8} O_i \log(\hat{O}_i) \quad (1)$$

2.3 Improved Prediction Method

As we can see in 2.1, we regard label $i_t$ as function $f_{i_t}$. It can be easily proved that $f_i \circ f_j \in F$ for any $i$ and $j$, so we can not only view $f_i$ as a function but an operator whose define domain and value domain are both $F$. For convenience, we denote the operator $f_i$ as $g_i$. Surprisingly, we can prove that operator $g_i$ is a surjection in $F$ for any $i$, which can
be simply explained by the following matrix $A$. In matrix $A$, $A_{ij} = k$ means $g_i(f_j) = f_k$.

$$A = \begin{pmatrix}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 0 & 3 & 2 & 5 & 4 & 7 & 6 \\
2 & 3 & 0 & 1 & 6 & 7 & 4 & 5 \\
3 & 2 & 1 & 0 & 7 & 6 & 5 & 4 \\
4 & 6 & 5 & 7 & 0 & 2 & 1 & 3 \\
5 & 7 & 4 & 6 & 1 & 3 & 0 & 2 \\
6 & 4 & 7 & 5 & 2 & 0 & 3 & 1 \\
7 & 5 & 6 & 4 & 3 & 1 & 2 & 0
\end{pmatrix}$$

Because $F$ is a finite set, $g_i$ is a injection and invertible. For any operator $g_i$, it exist an inverse operator and we denote it as $g_i^-$. Operator $g_i^-$ is also a surjection and injection in $F$, which can be simply explained by the following matrix $A^-$. In matrix $A^-$, $A^-_{ij} = k$ means $g_i^-(f_j) = f_k$. For simplification, we omit $f$ and use $g_i^-(j) = k$ to express the results above.

$$A^- = \begin{pmatrix}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\
1 & 0 & 3 & 2 & 5 & 4 & 7 & 6 \\
2 & 3 & 0 & 1 & 6 & 7 & 4 & 5 \\
3 & 2 & 1 & 0 & 7 & 6 & 5 & 4 \\
4 & 6 & 5 & 7 & 0 & 2 & 1 & 3 \\
5 & 7 & 4 & 6 & 1 & 3 & 0 & 2 \\
6 & 4 & 7 & 5 & 2 & 0 & 3 & 1 \\
7 & 5 & 6 & 4 & 3 & 1 & 2 & 0
\end{pmatrix}$$

Based on the above premise, we built the predicting method by the follow 4 steps. Figure 1 shows the method by a flow chart.

1. Apply $(f_0, f_1, \ldots, f_7)$ to the image $X_t$ to get 8 images. We denote these images as $(X_{t0}, X_{t1}, \ldots, X_{t7})$

2. Take these 8 images as input to DNN and get 8 orientations $(i_{t0}, i_{t1}, \ldots, i_{t7})$.

3. Apply $(g_0^-, g_1^-, \ldots, g_7^-)$ to these 8 labels and get another 8 labels $(g_0^-(i_{t0}), g_1^-(i_{t1}), \ldots, g_7^-(i_{t7}))$.

4. The labels which occur most in $(g_1^- (i_{t0}), g_2^- (i_{t1}), \ldots, g_7^- (i_{t7}))$ is the final result.

3 Experiment

3.1 Experiment Setup

We evaluate orientation recognition network on the MyoPS dataset. The MyoPS dataset provides the three-sequence Cardiac Magnetic Resonance (LGE, T2 and C0) and three anatomy masks, including myocardium (Myo), left ventricle (LV), and right ventricle (RV), some of the three-sequence Cardiac Magnetic Resonance is shown as Figure 2. MyoPS further provides two pathology masks (myocardial infarct and edema) from the 45 patients. For the simplified orientation recognition network, we train model for single modality on the MyoPS dataset, then transfer the model to other modalities. For each sequence, we resample each slice of each 3d image and the corresponding labels to an in-plane resolution of $1.367 \times 1.367$ mm.
We divide slices into three sub-sets, i.e., the training set, validation set and test set, at the ratio of 50%, 30% and 20%. Three sub-sets don’t have slices from same patient. Then, for each standard 2d image, we apply all function from $F$ to it to expand dataset. For training set, image slices are cropped or padded to $256 \times 256$ for the orientation recognition network, and apply random augmentation. For test set and validation set, the images are only resized to $256 \times 256$.

### 3.2 Orientation Recognition Network

During each training iteration, a batch of the three-channel images $X'$ is fed into the orientation recognition network. Then, the network outputs the predicted orientation network.

Figure 3 and Table 2 shows the training process and accuracy of the three sequences on test set. The results show that the model get quite high accuracy in three modalities. However, the size of test data is small, random factors influence the result heavily. In the following, we redivide the data, retrain the model, and analyses sensitivity of the model.

| Modality | Accuracy | Description      |
|----------|----------|------------------|
| C0       | 1.0      | Pre-train        |
| LGE      | 1.0      | Transfer learning|
| T2       | 1.0      | Transfer learning|

Figure 3: Training images
3.3 Sensitivity analysis

When using deep learning method to solve the problem, the volume of data is always the most important factor. In medical image processing, it is difficult to get a lot of data, so analysing the sensitivity of model is necessary. We redivided the data set into training set and test set 5 times, and the proportion of training set was 60%, 50%, 40%, 30% and 20% respectively. In each divided data set, we retrain the model in training set and compute accuracy in test set.

The accuracy variation is shown in Table 3. There is not a distinct difference among the accuracy while the volume of data decrease. Therefore DNN is a suitable method in CMR orientation recognition with high accuracy and low sensitivity.

Table 3: Accuracy of different training data size, modalities and prediction methods. The left column is the ratio of training data. Improved prediction means the method we proposed in 2.3 while Direct prediction means inputing the resized image to the DNN directly.

| Improved prediction | Direct prediction |
|---------------------|-------------------|
|                     | C0     | LGE    | T2  | C0     | LGE    | T2  |
| 60%                 | 1.0    | 1.0    | 1.0 | 1.0    | 1.0    | 0.998 |
| 50%                 | 1.0    | 1.0    | 1.0 | 1.0    | 0.991  | 0.999 |
| 40%                 | 0.974  | 1.0    | 1.0 | 0.967  | 0.996  | 0.996 |
| 30%                 | 1.0    | 0.979  | 1.0 | 1.0    | 0.975  | 0.997 |
| 20%                 | 0.982  | 0.960  | 0.986 | 0.978 | 0.953 | 0.983 |

3.4 Comparison between improved prediction and direct prediction

Table 3 shows the difference of accuracy between improved prediction and direct prediction. We can find that improved prediction always have higher accuracy in our experiment. However, it is not inevitable, because the vote may make the original correct decision wrong. Sometimes, the improved prediction have a lower accuracy, but in an average sense, the improved prediction is better than direct prediction.

4 Conclusion

DNN model get quite high accuracy in recognition of CMR image orientation and transfer learning make it easy to be transferred to other modalities. Thanks to the data expansion and augmentation, the model only need a few data. The improved prediction we proposed further increase the accuracy. We are sure that DNN model combining with transfer learning and improved prediction can be used in other recognition of orientation tasks.

References

[1] Ke Zhang and Xiahai Zhuang. Recognition and Standardization of Cardiac MRI Orientation via Multi-tasking Learning and Deep Neural Networks. In Xiahai Zhuang and Lei Li, editors, Myocardial Pathology Segmentation Combining Multi-Sequence Cardiac
[2] Xiahai Zhuang. Multivariate mixture model for cardiac segmentation from multi-sequence mri. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 581–588. Springer, 2016.

[3] Xiahai Zhuang. Multivariate mixture model for myocardial segmentation combining multi-source images. IEEE transactions on pattern analysis and machine intelligence, 41(12):2933–2946, 2019.