3.3 obtain 4-class classifier via patch learning

Phase 3 aims to obtain a 4-class classifier via patch learning. Global model for constructing a 4-class classifier would naturaly be KT-TFCM considering the existing historical cluster centroids. The preliminary result of clustering would roughly be 4 classes, i.e., bone, air, soft tissue and fat tissue, after applying KT-TFCM to MR feature data. Owing to the fact that mDixon sequences are expert in reflecting fat tissue and soft tissue, the gained clusters of fat tissue and soft tissue are fairly accurate. (先暂时说明mr的soft也反映得好)

Actually, the results of leftover cluster are quite rough, i.e, the air cluster is the mixture of pure air and bone and the bone cluster is the mixture of bone and soft tissue. (先暂时说明他俩就是这么混合起来的吧，为啥会混合呢？)

In order to gain the final 4-class classifier to generate precise synthetic CT image, patch models are reconstructed in these mixture parts. Benefiting from the prior knowledge of subject’s CT image, we can acquire labeled examples. However, it is practically imfeasible for radiologist to label the whole data manually because of huge time consumption. Thus, semi-supervised classification can be used in patch model and Lapsvm is ideal. With numerous unlabeled data and a limit amount of mark data, which is practically feasible for radiologist, it’s capable of obtaining a final 4-class classifier.

3.THE PROPOSED PL-TS-SC METHOD

The proposed PL-TS-SC method consists of five phases and these phases can be summarized in four main procedures in general, as shown in Fig. In procedure 1, Phase 1 extracts features of four types of MR data to form seven-dimensional MR feature data for each subject. In procedure 2, Phase 2 obtains the referenced class prototypes of 4 key tissue types via conventional FCM from prior knowledge of referenced MR data. In procedure 3, we obtain multiple 4-class classifiers via patch learning. Phase 3 constructs global model as KT-TFCM and pick up appropriate patch districts regarding to error that impacts most. Phase 4 reconstructs patch model as Lapsvm, which is semi-supervised classification, for patch districts picked to form a final 4-class classifier together with global model. In procedure 4, phase 5 organically combines multiple 4-class classifier to generate target synthetic CT image. Next, we detail each phase as follows.

3.1 Phase 1: Generate MR Feature Data for Each Subject With Given MR Images.

Feature extraction is of vital importance in machine learning and pattern recognition. The quality of features has a crucial impact on generalization performance. Considering the unavoidable noise in acquiring MR image, we adopt a principle of convolutional kernel to extract local texture features, which learns from the convolution layer in convolutional neural network (CNN). We extract local texture features from four types of abdominal MR datas, i.e., fat, water, in-phase(IP) and opposed-phase(OP), as the input to our PL-TS-SC method for each subject.

Except for above features, position feature is adopted as well for better distinguish with the consideration of the similarity of signals of air and bone in MR data. The scan voxel size and pixel slices of 3D MR image are 0.98×0.98×5 mm3 and 512×512 pixel slices of Z-axis respectively. Considering isotropism, we design a grid partition strategy of 5×5 voxels assembled. Each voxel position feature is determined by spacing of size 4.9×4.9×5 mm3. Thus position feature (x, y, z) can be expressed as the indexes of grid, 1≤x≤103, and 1≤y≤103. Combining all the features we obtain seven-dimension MR feature data as the input data to our method.

3.2 Phase 2: Obtain The Referenced Class Prototypes Of Four Key Tissue Types.

To generate final synthetic CT image from MR image as accurate as possible, we need referenced prior knowledge, i.e., referenced class prototypes of bone, air, fat tissue and soft tissue. Thus, paired CT and MR images are necessary and each pair needed to deformably registered before.

The work and data flows of Phase 2 are shown in Fig. Suppose there are n pairs CT and MR images to be referenced, we take one pair as example for detail. For reference pair 1, bone class centroid can be first determined since we can recognize bone data in MR image with CT data whose Hounsfield Unit (HU) is 300, and takes the average value of bone data as bone class centroid. Then we apply FCM to the leftover data, i.e., data without bone, for clustering and obtain three clusters centroids later. At this point, we obtain four class centroids of pair 1. And n referenced pairs produce n class centroids, class prototypes are the average value of class centroids for every class. Thus we obtain the reference class prototypes of four key tissue types.

3.3 Phase 3: Construct Global Model Via KT-TFCM

Phase 3 aims to obtain a 4-class classifier via patch learning. Global model for constructing a 4-class classifier would naturaly be KT-TFCM considering the existing historical referenced cluster centroids. The preliminary result of clustering would roughly be 4 classes, i.e., bone, air, soft tissue and fat tissue, after applying KT-TFCM to MR feature data. Owing to the fact that mDixon sequences are expert in reflecting fat tissue and soft tissue, the gained clusters of fat tissue and soft tissue are fairly accurate.

3.4 Phase 4: Reconstruct Patch Model Via Lapsvm

Actually, the results of leftover cluster are quite rough, i.e., the air cluster is the mixture of pure air and bone and the bone cluster is the mixture of bone and soft tissue. In order to gain the final 4-class classifier to generate precise synthetic CT image, patch models are reconstructed in these mixture parts. Benefiting from the prior knowledge of subject’s CT image, we can acquire labeled examples. However, it is practically infeasible for radiologist to label the whole data manually because of huge time consumption. Thus, semi-supervised classification can be used in patch model and Lapsvm is ideal. With numerous unlabeled data and a limit amount of mark data, which is practically feasible for radiologist, it’s capable of obtaining a final 4-class classifier.

3.5 Phase 5: Synthesis Target CT Image Through Multiple 4-Class Classifiers

We assemble multiple results of 4-class classifiers via the strategy of voting to decide voxel type. Considering numerous data in MR feature data, it’s infeasible to directly take the entire data as input to our method because of huge time consumption. Thus, we propose sampling-KNN mechanism, which involves randomly sampling the MR feature data and using K nearest neighbor (KNN) to restitute the whole prediction results, in our method to accelerate the whole process. Sampling-size denotes as ss.

Specific CT value are set to 380, -700, 98 and 32 corresponding with bone, air, fat tissue and soft tissue with the referring to XX to reconstruct a synthetic CT image.

4. EXPERIMENT RESULTS

4.1 Setup

In this section, we assess the effectiveness of the proposed PL-TL-SC method for generating synthetic CTs. Thus, ten subjects were recruited using a protocol approved by the University Hospitals Cleveland Medical Center Institutional Review Board. A deformable registration was performed to wrap the CT image to the MR image using OpenREGGUI, an open-source image registration package.

Moreover, three existing methods are in comparison with our method, i.e., the all-water method (AW), fuzzy C-means clustering (FCM) and transfer FCM (TFCM).

The parameters involved in our PL-TS-SC method are shown in Table X. Three metrics, i.e., the mean absolute prediction deviation (MAPD), the root mean square error (RMSE), and R, are used to evaluate the effectiveness of our method.

We adopt Leave-one-out strategy to generate ultimate result. That is, we treat one subject as test dataset and the other leftover subjects as train dataset. Each result for generating synthetic CT of one subject are assembled from results of the remainder classifiers, which excludes the classifier of the test one.

Our experimental studies were carried out on a computer with an Intel i5-4590 3.3 GHz CPU, 8 GB of RAM, Microsoft Windows 10 (64 bit), and MATLAB 2017a.

4.2 Analysis and Discussion