Multi-source Cardiac MRI Segmentation with Multivariate Mixture Probabilistic Model Combining Statistical Models of Deformation and American Heart Association Model *

No Author Given

No Institute Given

Abstract. As the current medical Imaging(MRI) is one of the most commonly utilized technique to aid the diagnosis in clinics, the cardiac magnetic resonance images (CMR) is also broadly recorded for assessing the structure and the function of the heart. As the different MRI methods stores different aspects of information of the cardiac, the multivariate mixture model is especially important for analyzing multi-source(MS) MRI images. In data pre-processing, the sequences are augmented into same size for the convenience of computation and the generalization compatity of the model, using the balanced-Steady State Free Precession(bSSFP) sequences for the computation of deformation velocity in applying Statistical Models of Deformation (SMOD) on the cine sequences.That is to compatively use it on small data as well as using Bayes inference in solving the problem with general data not limited to similar features, considering induction with uncertainty. And comprehensively, the LGE images with myocardium infraction(MI), making the segmentation more effective with recognizable brightness difference from images while the T2 weighted images gives essential information of some anomalies, such as scars, injuries and etc. And for better evaluation with dice and ACD,comparing with GD, the results are assessed in 8 section for each source(instead of the standard American Heart Association Model(AHA), given the standard four series of basal, midventricular, apical and apex levels). And specifically, in EM process, the assumed distribution for training data is inverse gamma and in analysing the test sequences, the uncertainty is also considered with Bayes theory, constructing the cdf of the haar like features extracted as Gaussian Mixtured. Future direction considers application and extension into patient profile can be developed not only with statistical process but also mechanistic interpretative tensor models.

Keywords: Multivariate Mixture Model · haarlike features · AHA · Sophistical Models of Deformations.

* Supported by MICCAI2020 anonymous submission

1 Introduction

The CMR of the patients have been utilized as the methods to both diagnosis the disease of heart or assess and monitor the status of the hearts. Although other methods as CT and PET scans are also frequently taken in hospitals and clinics, the MRI advantages itself in avoiding exposing the body to radiation of either X-rays or ionization process[1] although longer and louder scans are usually unavoidable. As mentioned in abstract, the three different MRI sources has their separate features.

To be more specifical, the bSSFP can usually depict better information of blood pool and myocardium with its better image contrast. And as one real-time MRI technique, it records the imaging of contracting cardiac continuously. Accordingly, they are used in computation of deformation velocity to augment the dataset into same and larger size on time domain especially. Note that, the registration is critical so that the motion artifact can be denoised to avoid invalid images. Nonrigid registration [2] is conducted through SPM here along with linear transformations. In our statistical augmentation model[3], the shape variability is modelled with the training set space after the PCA with eigenvalues and thus reduce the dimension with fewer non-zeros in each column as well as keep the shape transforming the images between atlas space and image space effectively with inverse and forward by the sign of the velocity field which is significantly more convenient than using displacement. Meanwhile, the LGE sequences mainly gives us clear infraction information. In some work [4], the cardiac magnetic resonance with LGE not only is effective in diagnosis for its high specificity but also of good diagnostic performance(especially being associated with one other CMR sequence method, for instance T1 or T2-weighted imaging.) According to one another study [5],the accuracy of LGE can be improved from 75 percent to 85 percent if th other CMR sequence is also positive for myocarditis after the LGE. Guarantueed the independence of each pixel at different location, the model is based on the intensity distribution of each single figure with the haar like features extracted. As the main aim is the segmentation of edema and scars(although in this paper the segmentation is applied globally on all of the 5 classfications, we still conduct the segmentation with ROI similar as the training GD), the detection of the boundary combining haar-like [6] features is considered with the significant difference of the intensity. Such feature largely reduce the training time. As we exclude the discontinuous/too small clusters of pixels to be of one tissue, only horizontal, vertical and diagnal haar-like features are extracted. Unlike applying the haar transformation and inverse transformation, main aim here is to use it as feature, thus simple masks are extracted consider keeping the search direction only. (As our windows are made of 3*3, thus both the even sized type and odd sized the type are possible type as well as the label catogories are composed of 5 different levels. Instead of applying the standard extraction, we still apply the greedy algorithm avoiding the problem when more than 2 color lables appearing in one windows.) And taking them as subtypes, the segments are detected after training applying the multivariate mixture model(MvMM)[7].And due to the indenpence of the pixels,

model is considered to be naive Bayes, which multiplies the possibility of each different feature subtype instead of construct high dimension distribution. For the training dataset, the distribution is simulated under inverse gamma function while applying the segmentation onto the testing data, the unkown distribution is appximated under Gaussian process with the EM [8]method.And the optimization is conducted the dice and ACD separately. T2 weighted images gives most accurate segmentation of the scars[9], which is also in accordance with the Finally, the evaluation of the test data segmentation is computed in the 24 different segments(8 for each source of image deriving from AHA model.[10])

2 Methods and Materials

The CMRI images series constitutes of 6 sub-sequences of each source of each patient(after augmented),from whom 25 patients are for training and 20 for testing without cross-validation. Each training image is of 511*511 pixels while testing image is of 515*515 pixels. The unit time for the imaging is 1 second.

2.1 Stochastical Models of Deformations

For each sources (1 for bSSFP, 2 for LGE and 3 for T2-weighted) of each patient i, current subsequences l are registered to the first figure,the diffeomophy transformation encodes the velocity variable Vsil for each image. And the mean of the velocity filed is denoted as:$\mathbf{Vavg} = \mathbf{\Sigma_{L=2}^{l}(I(s,i,L) - I(s,i,L-1))/(l-1)}$, Using the eigenvectors to decompose each image:

$$\mathbf{W^T * I(s,i,L) = W^T * D}$$

,where W is the left eigen vector and D is the diagonal matrix with eigen values or

$$\mathbf{I(s,i,L) * V = D * V}$$

,where V is the right eigenvector and thus, the velocity in the atlas space can be sampled as the random velocity fields:

$$\mathbf{Vg = Vavg + V * x * diag(D)}$$

,where x is an array with random numbers $N(0,\sigma)$. In our simulation, the σ is set to be 1. When generalize new data, inverse it back to image space and through changing the sign of the velocity field(considering the image with size n):

$$\mathbf{Vavg = Vg - V * x * diag(D)}$$

, Simply subtract the previous image sets and get the inverse image from randomized velocity field.

2.2 Haar-like feature extraction

As introduced in the introduction, the haar-like feature[11] we are interested in are three types: horizontal, vertical and diagonal. In each image, a moving window of size 3*3. All possible features includes:

$$[\mathbf{1}, -\mathbf{1}]; [-\mathbf{1}, \mathbf{1}]; [\mathbf{1}, *; -\mathbf{1}, *]; [-\mathbf{1}, *; \mathbf{1}, *];$$
; -1,[*,1;-1,*]; [-1,*;* 1],[*,-1;1,*];
-1, 1; [-1, 1,-1];
,;*,-1,*;*,*,1; [-1,*,*;*,1,*;*,*,-1];
,1,;*,-1,*;*,1,*;
,-1,;*,1,*;*,-1,*;
,,1;*,-1,*;1,*,*; [*,*,-1;*,1,*;-1,*,*];

,where positive one will product on the brighter pixel(with higher intensity) while the negative is used to multiply the darker one.(Consider our label values are [500 600 200 1220 2221], the threshold for the difference is 100.)note that the search order according to the original algorithm should be sum the right bottom corner and left upper corner and then subtract right upper corner and left bottom corner. And after searching the whole image, the mask also is required to be configured again whether the boundary is continuous i.e. no single feature block exists.

2.3 Multivariate mixture model[12] combine Bayes model

Consider the process same for each source type s: Each patient i, a time series of l subsequences and k feature subtypes is obtained:Haari,l,s(k,x-1:x+1,y-1:y+1) = P, where P is the intensity value multiplied by haar-like mask at location(x,y) and the distribution is computed with the unit as 3*3 window. For the training data, we approximate the probability distribution with inverse gamma function: Initial the parameter with

$$\alpha_{\mathbf{0}} = \beta_{\mathbf{0}} = \mathbf{gamma_0} = \mathbf{1}$$
$$\Lambda = Gamma(2, 1), \mu_0 = 0,$$
$$\mathrm{W}_0 = \frac{1}{2\sqrt{\pi}} \sqrt{\frac{\lambda 0}{\lambda}} * \frac{\beta^{\alpha_0}}{\beta^{\alpha}} * \frac{\Gamma(\alpha)}{\Gamma(\alpha_0)}$$
$$\mathrm{W}_j = N(x_i|\phi_{j,1}, diag(\phi_j, 2))$$
$$\phi = W_0 * Q + \Sigma(W_j * \delta_{\phi_j})$$
$$\mathrm{Q} = \beta^{\alpha} * exp(-\frac{\beta}{x}) \frac{}{\Gamma(\alpha) x^{-\alpha-1}}$$

, And thus futher, the Cdf of Φ is applied as the distribution of feature intensities. We are then possible to iterate through the maximization of expectation likelihood to apply it on the test data with uncertainty combining the Bayes theory. With the atlas prior probability $P_A(k_x)$ stands for the probability of certain location(x,y) belongs to patient i 's l subsequence image' s k th feature and the proportion of label l, π_k, the prior of the location (x,y) belongs to label l is then:

$$\pi_{\mathbf{kx}} = \mathbf{p}(\mathbf{s}(\mathbf{x}) = \mathbf{k}|\theta) = \frac{\pi_{\mathbf{k}} \mathbf{P_A}(\mathbf{k_x})}{\mathbf{Z}}, \mathbf{Z} = \mathbf{\Sigma}(\pi_{\mathbf{k}} \mathbf{P_A}(\mathbf{k_x}))$$

,and thus the expectation for the training data is

$$\mathbf{P_{kx}} = \frac{\mathbf{Phi} * \pi_{\mathbf{kx}}}{\mathbf{Z_2}}, \mathbf{Z_2} = \mathbf{\Sigma}(\mathbf{Phi} * \pi_{\mathbf{kx}})$$

, And to simulate and approximate the test data, the mean and variance is computed for further iteration: after the initialization:

$$\tau_{\mathbf{ikc}} = \frac{\mathbf{P_{ikcx}}}{\mathbf{Z_3}}, \mathbf{Z_3} = \mathbf{\Sigma}(\mathbf{P_{ikcx}})$$
$$\mu_{ikc} = \frac{P * P_{ikcx}}{Z_4}, Z_4 = \Sigma(P * P_{ikcx})$$
$$\sigma^2_{ikc} = \frac{\Sigma(P - \mu_{ikc})^2 * P_{ikcx}}{Z_5}, Z_5 = \Sigma(P_{ikcx})$$

,the optimal π_k is then computed through the coeficient $C_x = \Sigma(P_A(k_x)) * \pi_k$
$\pi_k = \frac{\Sigma_x P_{kx}}{\Sigma_x (P_A(k_x)/C_x)}$
and the parameter updated as follows :

$$\mu = \frac{\lambda_\mathbf{0} * \mu_\mathbf{0} + \mathbf{P}}{(\lambda_\mathbf{0} + \mathbf{1})}$$
$$\lambda = (\lambda_0 + 1), \alpha = \alpha_0 + 0.5$$
$$\beta = \beta_0 + 0.5 * (\lambda_0 * \mu_0^2 - \lambda * \mu^2)^2$$

,thus the Q and W can be updated again and finally the Φ, leading to the approximation of hidden data[1]12:

$$\mathbf{P_{ikcx}} = \frac{\mathbf{\Phi} * \tau_{\mathbf{ikc}}}{\mathbf{P}} * \mathbf{P_{kx}}$$

Note that, the uncertainty of approximation with pdf and cdf are different, and the result in this paper is computed with cdf. The difference computation can be found in Appendix A.

2.4 Dice and ACD Evaluation with American Heart Association derived section

As the same metrices applied in previous work[13], the Dice and ACD here is computed with 8 segments for each subsequence. The quadrant is evenly segmented into 8 part with 45 degree for each(of 360 in all. As the common domain of of the three different sources of CMRI depends on how they crossed each other(spatially here assuming no uncertainty in time domain consider the simultaneous source.)Thus, The segments are also differently rotated for the three space separately with 90 degree difference from bSSFP to CMR and then to the T2.(The coordinate map can be found in the Appendix B).

2.5 algorithm

For biomedical model, The ability to capture extensive data about individual patients are especially important to the model, and thus this paper focuses on the comprehensive application of multivariable variable mixture model from not only full utilities and associating with uncertainty to get good performance but also the extension with stochastic augmentation of the dataset and more complete evaluation considering the robust and versatility of the model.(pipeline of the module can be found in the figure 1).
Considering each patient's CMRI sequences X from training data, preprocess the images with registration and transformation both rigidly and unrigidly denoted as Xi.

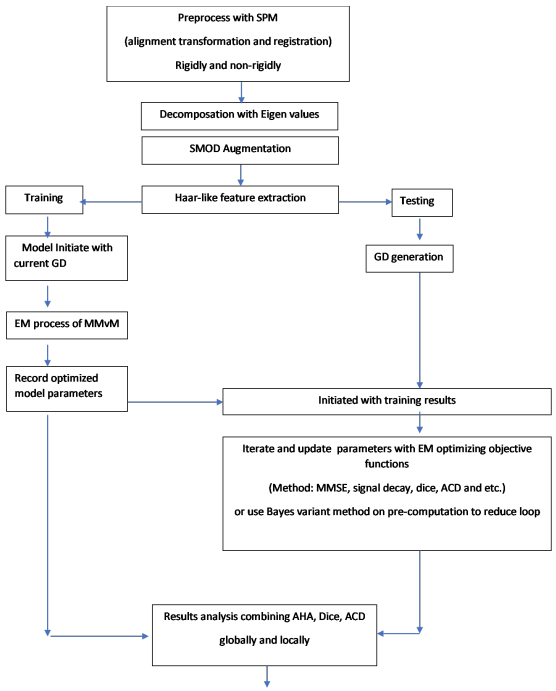

Fig. 1. Model pipeline

Decompose each images with eigenvectors through left sigular matrix W'*D or or right singular matrix D * V (It might be possible to largely compress images into sparse matrix even as long as the singular matrix is kept not to be digressed.)

Map current training images into atlas space and compute the average velocity Vavg, and forward them through the random velocity field computation(by columns):

Vg initialized with Vavg: $Vg = Vavg$;

Scan each column, add randomized values following normal distribution: $forc = 1tosize(Vavg, 2)Vg(:, c) = Vg(:, c) + W * x * diag(D)$, where $x\ N(0, 1)$

generalize new image through inverse it back to image space: $Vavg = Vg - W' * x * diag(D)$;

extract the inversed image from the Vavg and inverse it back: $I(i, l) = Vavg * (l - 1) - \Sigma_2^{l-1}(V(i, L) - V(i, L - 1)) + I(i, l - 1)$

Get the distribution(3*3 non overlaped windows) of the augmented training dataset with inverse gamma function after haar-like feature extractions (not detailedly listed following fomula in chapter 2.3 to get Phi)

Considering computation with Bayes inference, follow the formula again down to the end of iteration repeating EM process. Alternatively, dynamically training with optimization can also be conducted with comparing the distribution iterated with GD either directly computing distance (MMSE, or the ACD in the evaluation and etc.) or through other metrices or likelihood of objective functions.

Finally, for each patient i, each source of CMRI s, each subsequnece l, evaluate the segmentation result with GD. Meanwhile, take down the parameters for achieving best result.

Apply the parameters in testing and update the model again evaluating the ACD(totally for the whole image) while dice (for each different labels) as in this paper until it reaches the tolerance or over the max steps. for L = 1,... to number of Haar-like features, Haari,l,s(k,x,y), consider in the 3*3 window, compare it with the data, if accordance with the criteria in Chapter 2.2 distance set as 0, in-accordance pixel distance set as 1.Compute HD in each window: $ACD(i, l, s, k) = sum(sum(HD(x - 1 : x + 1, y - 1 : y + 1)))$; Finally get the ACD mean and sum. For dice, Compute DV in each window more strictly with 1 at where their intensity I(x,y) are in same level also according to the criteria(for tissue label) in Chapter 2.2, volume crossed DV sum up 1 while not crossed do not do any operation, divide by the total volum $V1 + V2$thus compared to ACD add one sub loop: $fort = 1to5$
$Dice(i, l, s, k, t) = 2 * sum(sum(DV(x - 1 : x + 1, y - 1 : y + 1)))/(V1 + V2)$, Similarly, get the Dice mean and sum.

Evaluate the convergence and accuracy(sensitivity and specificity). $sensitivity = \frac{sum(truepredictiontruedata)}{sum(truedata)}$ $specificity = \frac{sum(falsepredictionfalsedata)}{sum(falsedata)}$

3 Preprocessing

3.1 SMOD Augmentation

As the dataset of the origin data is not of same time period,the regular preprocess is conducted through SPM12, including alignment, transformation, registration both rigidly and nonrigidly. After that, the SMOD is applied to make it into the same size(25*6*511*511).Similarly, and testing dataset are augmented with the SMOD as well and into size(20*6*515*515). One example of raw data, augmentation velocity field and after augmented figure of training dataset and test dataset separately are shown below:

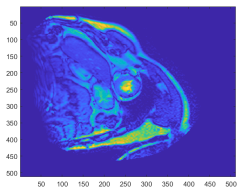

Fig. 2. Training

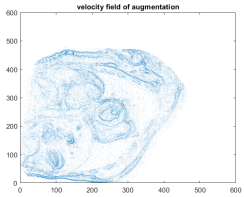

Fig. 3. TrainingVelocity

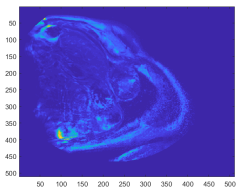

Fig. 4. AugmentedTrain

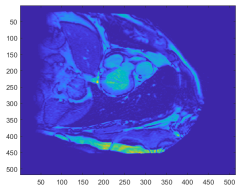

Fig. 5. Testing

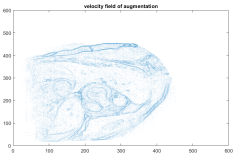

Fig. 6. TestingVelocity

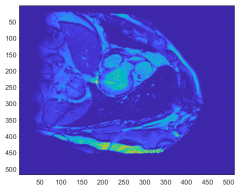

Fig. 7. AugmentedTesting

It can be seen that the training velocity field of testing data is of higher resolution which is due to the larger size of eigenvalues are kept while for this training dataset, the eigenvalue kept rather few and thus the augmentation is not with information for similar data regeneration. However,instead of reconstruction, the augmentation is based on stachastic process which strictly keep the shape of the tissue while considering losing instensity information as the principle components are based on 2 to the power level compressible on shape information usually. Thus, the GD lable is relatively important influencing the result of the segmentation and especially in classification of edema or scars. Note that the reason that we use the SMOD still is that intensity of raw data is of different for different source even comparing horizontally in CMRI only.As has

mentioned, In some of the origin data of T2 weighted source MRIs, relatively less distinctive boundaries also make the segmentation harder even with GD labels Thus, we use only bSSFP source for augmentation considering the real-time feature of it in final test.

3.2 Haar-like Feature Extraction

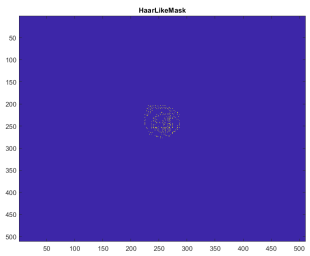

Fig. 8. HaarMask

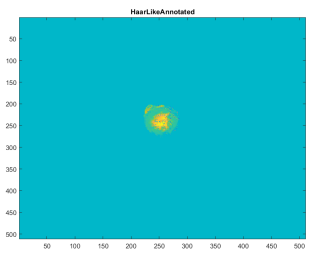

Fig. 9. HaarAnnotated

Different than simple haar feature extraction on two valued data, considering the 5 labels making comparing 2 for each time might aslo be computation consumed, we simply search and lable the haar-like feature manually. In addition, the as to eliminate single point subtracted as feature which is not useful in our case(although defined also as centered haar-like type), we also make correction after the first search on the whole image scale deleting those single points first and then with increasing radius length, some clusters with small mount of points are eliminated as well. Compared with the only annotation with addition and subtraction result in figure 8, the corrected result in figure 7 is significantly better after we apply both mask for label and correction with elimination of over small clusters,However, according to human eye, the result still seems to be with discontinuous boundary. Such disadvantage is later overcome with setting looser threshold of radius when searching.

3.3 GD annotation for test results

As the we first try the experiment of testing with the parameters recorded only after training data.(However, the result is not that satisfactory especially on with ACD metrics.)Thus, in the first round, we finally chose one best subtissue depending on the data source type only for each time series.(Some of the first round results can be found in the Appendix C mainly as heuristic process only.) In the final round, we found that the update of the parameter was not completely correct and after correcting it with making GD labels to keep the evaluation of the expectation on measure-able data with more information than the previous

training data. And to still keep the robust of the model, our augmentation is conducted still with randomization and thus guarantee both the known and unkown properties of the dataset are fully utilized.

4 Experiment Results and Analysis

4.1 Training with GD

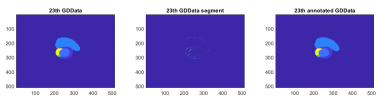

Fig. 10. TrainingGD

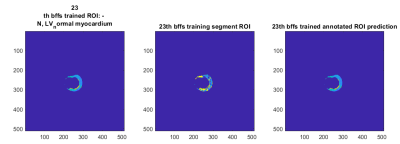

Fig. 11. Trainingbssfp

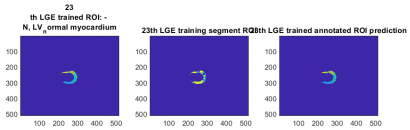

Fig. 12. TrainingLGE

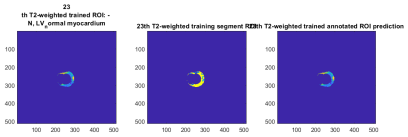

Fig. 13. TrainingT2

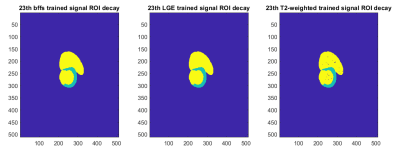

Fig. 14. Trainingsignal decay

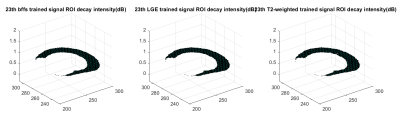

Fig. 15. Trainingsignal decay in dB

As multi-variable the mixture model consider both the measurable data with the labels and unknown distribution, when training with the dataset, the segmentation on GDData is also applied. Through segment result, it can be told clear that all the five labeled classifications can be classified validly although small amount of the prediction is not correct. In this example(training data of the 23rd patient), without exemption, the highest detected part is normal myocardium. As mentioned, the classification results is comprehensively on all the five parts, so still here we introduce this example as it gives out highest dice and lowest ACD in all the training data whcih might give clearst analysis as we conduct similar analysis on basis of it for the rest of the data. As the intensity is value given by the three different source type of CMRI, the level of intensity prediction are also different while the shape information due to the knowledge of GD is basically with no false prediction. And the difference of among them is also subtle. As T2 and bSSFP are not as sensitive as LGE MRI to the intensity(figure

10-figure12), the segment of them are also not as accurate as the LGE's. At least from the example here, the volume correctly predicted for normal myopnium is also the largest for LGE's. However, the prediction result is very close to the GD on both shape and intensity. While in decay level, the mean value can be easily achieved through figure 14 which is on dB unit (log10 scale), and the difference can be achieved through figure 13 instead, which here shows some anomalies on T2 weighted(right most). This is also consistent to some studies that T2 MRI have good performance in acute injuries.

4.2 Testing with generated GD

Similar as training, after the first round, we decide to generate GD for the testing data as well to keep the supervised part of the multivariate mixture functional as well as the unsupervised part. And here the example is the first sub-sequence of the 5th patient.

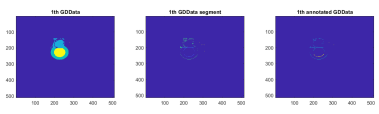

Fig. 16. TestingGD

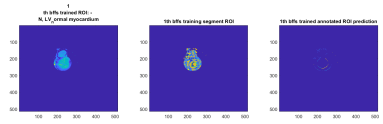

Fig. 17. testingbssfp

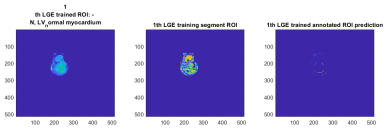

Fig. 18. TestingLGE

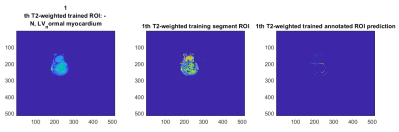

Fig. 19. TestingT2

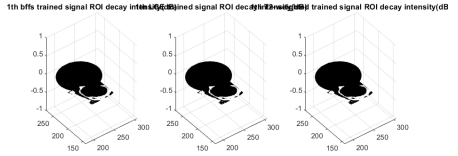

Fig. 20. TestingSignal decay in dB

As the GDData segmentation generalized using Bayes inference and moving average windows alternatively, the GD keeps the distribution approximately as the origin data while segmented into each label level considering it as a binary classification problem. As seen in the left most GD in figure 15. Comparing it with the origin data of three types simultaneously (i.e.the left most figure in fig.16-fig.18), due to the scaled level image, only relatively intensity difference and shape information can be shown while the exact absolute intensity level(both evaluation of dice and ACD can be seen in the table while the complete evaluation and results can also be downloaded [14])

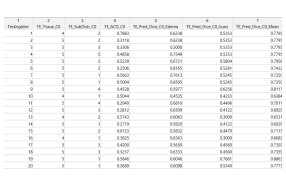

Fig. 21. TestCo

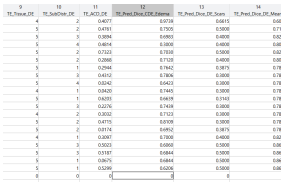

Fig. 22. TestDE

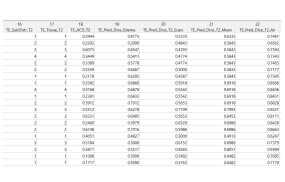

Fig. 23. TestT2

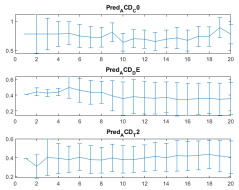

Fig. 24. ACD(Round 2)

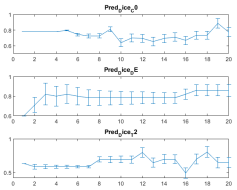

Fig. 25. Dice(Round2)

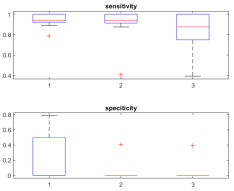

Fig. 26. Sensitivity and Speciticity(ROund2)

As the Atlas of generated GD is predicted to be quite similar by LGE and T2(see the results in centered graphs from fig16-fig18), both slightly better with ACD(p¡0.5 ttest and ANOVA tested by results in tables of figure 20-fig 22) than the results by the C0 while the evaluate with Dice are all quite similar(In addition to statistical test results, the error bar chats can also be found in fig 23-figure25). This is mainly due to some outliers in means of ACD(might introduced with the missing filling replacement of the results replacing some NAs of both DICE and ACD in the tables.) for C0 MRI segment evaluations,leading to large variance as well which can also be observed through the error bar. (Compared with round 1 results, figures in Appendix C, the ACD levels are all decreased with dice level significantly higher Meanwhile, the sensitivity is also significantly improved. Although the specificity is kept nearly unchanged, mainly with larger true negatives rate for C0 CMRI.) And according to the segmentation results, there are more edema patients than scars patients. (Most outliers happen to occur on scars prediction evaluations)

4.3 evaluation with ACD and dice example

The segmentation of the subsequence of each patient is evaluated separately for different sub-tissue with dice while as the whole value with ACD metrices.

Take patient 1's 1st subsequence as instance, the ACD decreases in the first 8's steps from 10 to 8.8 and converges. The dice for different sub-tissues follow the same increase rate, the highest is the right ventricular blood pool and last is the LV myovardial edema. Similarly it converges until the 8'th step.

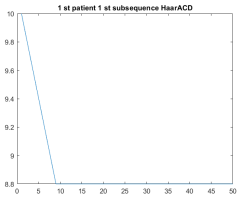

Fig. 27. ACD steps of patient 1's 1st subsequence C0 CMRI

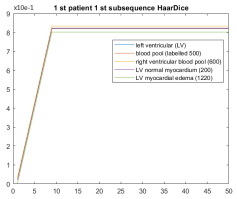

Fig. 28. dice steps of patient 1's 1st subsequence C0 CMRI

4.4 evaluation with ACD and dice combining AHA and different tissues

As the area of the cross-common-space is dependent on the degree between each two different sub-sequence MRI, we introsuce the AHA model to evaluate the ACD and Dice in different location of the pixels considering segment the quardrant evenly into 8 sections with 45 degree of 360 for each. And the degree change of the subsequence method can be easily shown with adjust of the rotation of the polar plot. According to this model, the cross angle between each subtype is relative sharp and thus, as for the ACDs on the left sides of figure 26 shows, the distribution of the center of the segments are not with large difference.

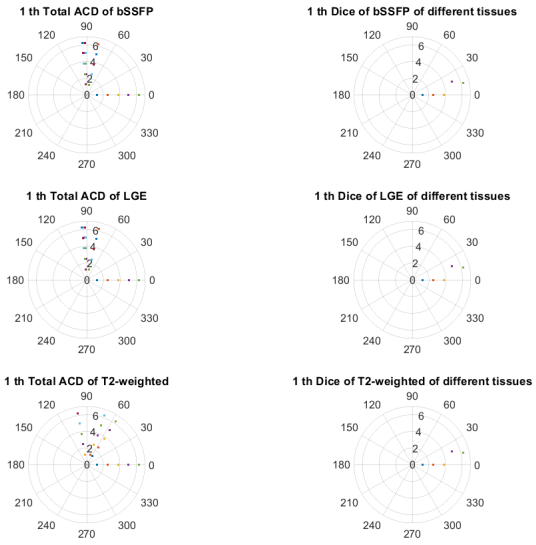

Fig. 29. ACD and dice with AHA and different sub-tissues

In this part, the ACD evaluation gives similar results for bSSFP and LGE sequences(different from the total result considering on whole image,which gives out similarity to LGE and T2-weighted in the previous error bar chart.)This seems to be more reasonable as T2-weighted MRI is usually different in intensity giving more annotations for injuries.(Which might explains the line of the dots between 30-60 and over 90 degree here. for some specific scars for instaces.) Speaks of the reults for different tissues, as it is shown as one best prediction for each sub-type, here we have identicle predictions from different subsequence for each subtype.(The whole predictions for or the 20 test patients can also be found in the table in fig20-fig 21).

5 Discussion and Conclusion

The multivariate mixture model is useful in modeling biomedical images sequences which are not mono-catogorized. In this paper, although the results are evaluated separately for the three different source type, bSSFP, LGE and T2-weighted so as to show the different property of them in intensity especially and to keep their specific interpretation ability in different usage in clinics and hspitals, the subsequences and 5 labels are considered as the multi-variable in each model. As the results shows, in the first round, only using labels in training data lead to bad results in dice for T2 especially. And thus we took some tests with optimization not only on parameters but also depending on subsequence. And with the trial of generating labels for the testing datasets, the final results got significantly improved. NOt only the dice are higher generally, the sensitivity of the model also is largely increased. Which also enlarges the robust of the model. As we have introduced, this is also achieved with the augmentation before training and testing, whcih is especially helpful in small dataset application(no matter for classificaation or segmentation tasks./[15]). And as the compativity to unkown data is also challenged but possible for MMvM, here, we also tried to generate GD by ourselves and again test them with augmented data as well.According to the results with the Bayes theory combined EM process, the result is valid for most datasets and although the speciticity is not so ideal considering the difficulty in computation of the unkown data with large uncertainty. As for clinical and hospital biomedical real-data cases, the ability to construct clear and interpretative profiles [16] for patients and physicians is also important as well as tracking data, the AHA combined evaluation can be considered as a highleight as well showing the evaluations of the three sub-types considering their cross in common domain, which can be adjusted with the rotation of the coordinates with the overlapped space area change. Furthermore, in the future, the ensemble method or hierarchical model might be considered. When computing with weight gradient based method different weight for each classifier for each source can be considered. (In this paper, the π_l is considered as constant after converged, computing for each source separately.) And as the complete profile of patient usually requires not only image data but also eletrode signals such as ECGs, the next step for extension is to combine ECG data and

to not only induct with stochastical process but also deduct into mechanistic models considering the tensor model which is also widely used combing imaging processing techniques as well.

AppendixA

Theorem 1. *Here we try one computation of each pixel after decomposition I_{ixk},belongs to patient i's CMRI k subsequence (K subsequence in all)and as for consider the poission process with gamma distribtuion which is after augmentation, getting masked as member of label l(1-5's the level LB) with possibility $alpha_{ixk}$, count in the 3*3 windows defined in the paper . And thus, we have the prior of being the boundary as:* $L(\lambda|Y) = \pi_{ikj=1}^{K} \frac{\lambda^{y_{ikj}*exp(-\lambda)}}{y_{ikj}!}$
$== \frac{*\Sigma_{ikj}^{K}(y_{ikj}*exp(-K*\lambda))}{\pi_{ikj}^{K}*y_{ikj}!}$
, where y_{ikj} is 1 when marked as in same lable with j'th pix in the window.(THe window is moving non-overlapped on the whole Image A_{ik}) Thus, we have the possibility, assuming the x follows gamma distribution as the conjugate prior:
$f(x) = \frac{\beta^{\alpha_{ijk}}}{\gamma(\alpha_{ijk})*x^{\alpha ijk-1}*exp(-\beta*x)}$
, and thus the posterior is thus gamma poisson as well: $\lambda^{(\Sigma_a lpha_{xk}+alpha_{xi})} + exp(-(9+\alpha_{xk})*\lambda)$
where, $k! = i$. One another alternative method is to have large samples to make prior of its mean to be flat distributed from generate a nonfinformative prior for its $sigmma^2$ from inverse gamma distribution : σ^2 $IG(a,b)$*, the pdf then becomes* $f(\sigma_2|a,b) = (\sigma^2)^{-(a+1)*exp(-beta/sigma^2)}$.
While the different distribution is just special case, the important step is to take conjugate pair for Φ and k in Dirichlet process, and with the induced integral approximating original distribution,
$p(dx) = \Sigma_j W_j * \psi(dx|phi_j)$
And for the Gaussian kernel, both the mean and the variance will be available in an explicit form for the distribution above.

Proof. Consider $\phi = (\phi_1, \phi_2)$ and $\psi(dx);\phi) = N(dx;\phi_1,\phi_2)$
for location and scale parameters ϕ_1 and ϕ_2. Then for the Gaussian kernel $k(x,x') = tau * exp(-(x-x')^2/2\lambda)$
, the kernel mean becomes :
$mu(x) = \Sigma_j(\frac{\tau*\lambda*w_j}{(\lambda^2+\phi j,2)^{0.5}*exp(-\frac{(x-\phi_{j,1})^2}{2*(\lambda^2+\phi_{j,2})})})$
and the initial variance can be expressed as,
$p*p(k) = \Sigma*(\Sigma(\frac{\tau*\lambda*W_j*W_j'}{(\lambda^2+\phi_{j,2}+\phi_{j',2})^{0.5}}) * exp(-\frac{-(\phi_{j,1}-\phi_{j,1})^2}{2*(\lambda^2+\phi_{j,2}+\phi_{j',2})})$
Take Gaussian mixture distribution as p(dx) as what did in the paper, if the function f(x) is chosen as the polynomial, $f(x) = Sigma_i a_i * x^{b_i}$
,the integral p(f) is computable in closed-form and the generic approximation properties of Gaussian mixtures, the GP prioe with mean function $m_\theta(x=0)$ and Gaussian covariance function: $k_\theta(x,x') = \tau * exp(-(x-x')^2/2/\lambda^2)$;

AppendixB Classification after segmentation through 24 section AHA(8,8,8 sections for each figure),based on cross-common-space and angle

$sections = zeros(:,:,8);$
$for x = 1:256$
$sections(x, 1:x, 1, 1) = 1;$
$sections(x + 255, 1:(257 - x), 2, 1) = 1;$
$sections(x + 255, (257 - x):256, 3, 1) = 1;$
$sections(x + 255, 256:(x + 255), 4, 1) = 1;$
$sections(x + 255, (x + 255):511, 5, 1) = 1;$
$sections(x, (512 - x):-1:(257 - x), 6, 1) = 1;$
$sections(x, 256:(512 - x), 7, 1) = 1;$
$sections(x, x:256, 8, 1) = 1;$
end

AppendixC

References

Author D. Simon: Magnetic resonance imaging of brain and brain stem. In: ICD-9-CM(2009).
SPM homepage,https://www.fil.ion.ucl.ac.uk/spm/
Author J. Acero, E. Zacur, H. Xu, R. Ariga, A. Bueno-Orovio, P. Lamata and V. Grau: SMOD-Data augmentation based on Statistical Models of Deformation to enhance segmentation in 2D cine cardiac MRI. In: FIMH2019.014v4.
Author Baccouche H, Mahrholdt H, Meinhardt G, Merher R, Voehringer M, Hill S, Klingel K, Kandolf R, Sechtem U, Yilmaz A :Diagnostic synergy of non-invasive cardiovascular magnetic resonance and invasive endomyocardial biopsy in troponin-positive patients without coronary artery disease. In: Eur Heart J. 2009 Dec; 30(23):2869-79.
Author Abdel-Aty H, Boyé P, Zagrosek A, Wassmuth R, Kumar A, Messroghli D, Bock P, Dietz R, Friedrich MG, Schulz-Menger J: Diagnostic performance of cardiovascular magnetic resonance in patients with suspected acute myocarditis: comparison of different approaches. In:J Am Coll Cardiol. 2005 Jun 7; 45(11):1815-22.
Author Viola and Jones: "Rapid object detection using a boosted cascade of simple features", In: Computer cool Vision and Pattern Recognition, 2001
Xiahai Zhuang: Multivariate mixture model for myocardial segmentation combining multi-source images. In: IEEE Transactions on Pattern Analysis and Machine Intelligence (T PAMI), vol. 41, no. 12, 2933-2946, Dec 2019.
Sundberg, Rolf: "Maximum likelihood theory for incomplete data from an exponential family". Scandinavian Journal of Statistics. 1 (2): 49–58. JSTOR 4615553. MR 0381110.(1974)
Anthony Bozzo,2 Judith Marcoux,3 Mohan Radhakrishna,4 Julie Pelletier,5 and Benoit Gouletcorresponding author1: The Role of Magnetic Resonance Imaging in the Management of Acute Spinal Cord Injury. In:J Neurotrauma PMC3143391
Manuel D. Cerqueira, MD; Neil J. Weissman, MD; Vasken Dilsizian, MD; Alice K. Jacobs, MD; Sanjiv Kaul, MD; Warren K. Laskey, MD; Dudley J. Pennell, MD; John A. Rumberger, MD;Thomas Ryan, MD; Mario S. Verani, MD: Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart. In: AHA Scientific Statement
Crow, F: "Summed-area tables for texture mapping", In: Proceedings of SIGGRAPH, 18(3):207–212, 1984

Xiahai Zhuang: Multivariate mixture model for cardiac segmentation from multi-sequence MRI. In:International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.581-588, 2016.
Sebastian Farquhar, Michael Osborne, Yarin Gal: Radial Bayesian Neural Networks: Robust Variational Inference In Big Models
https://github.com/LilyHeAsamiko/MyoPS2020-
Alexandra Deis: Data Augmentation for Deep Learning, In: DataScience Journal(2019)1,2-11
Blanca Eodriguez, Jorge Corral-Acero, Francesca Margara, Maciej Marciniak,Edward, Ada Doltra, Tina Morrison, Yingjing Feng, Hongxing Luo, Alfonso Bueno-Orovio: The 'Digital Twin' to enable the vision of precision cardiology, in European Heart Journal(2020)0,1-11 doi:10.1093/eurheart/ehaa159. 2020
Chris. J. Oates, Steven Niederer, Angela Lee, Francois-Xavier Briol, Mark Girolami: Probablistic Models for Integration Error in the Assessment of Functional Cardiac Models. arXiv: 1606.06841v5

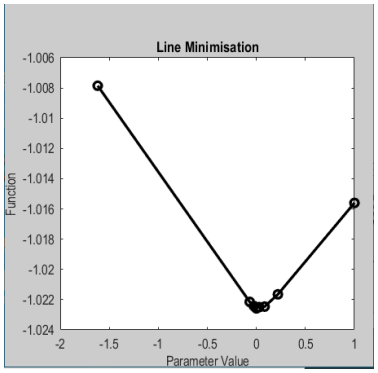

Fig. 30. Preorocess with SPM

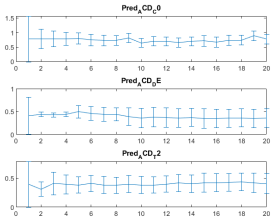

Fig. 31. ACD(Round1)

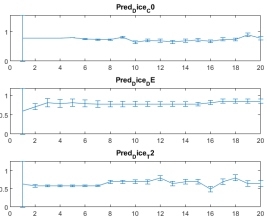

Fig. 32. Dice(Round1)

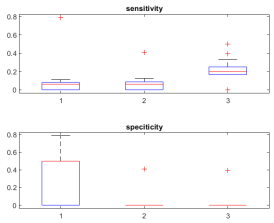

Fig. 33. Sensitivity and specitic-ity(Round1)

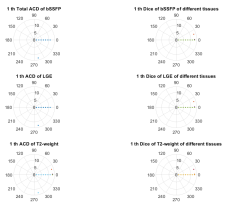

Fig. 34. Evaluation1(Round1)

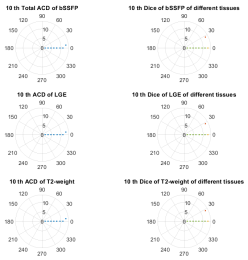

Fig. 35. Evaluation10(Round1))