Automatic Cardiac View Classification of Echocardiogram

J. H. Park[1], S. K. Zhou[1], C. Simopoulos[2], J. Otsuki[2], and D. Comaniciu[1]
[1]Integrated Data Systems Department, Siemens Corporate Research, Princeton NJ 08540
[2]Ultrasound Division, Siemens Medical Solutions, Mountain View, CA 94039

Abstract

we propose a fully automatic system for cardiac view classification of echocardiogram. Given an echo study video sequence, the system outputs a view label among the pre-defined standard views. The system is built based on a machine learning approach that extracts knowledge from an annotated database. It characterizes three features: 1) integrating local and global evidence, 2) utilizing view specific knowledge, and 3) employing a multi-class Logit-boost algorithm. In our prototype system, we classify four standard cardiac views: apical four chamber and apical two chamber, parasternal long axis and parasternal short axis (at mid cavity). We achieve a classification accuracy over 96% *both of training and test data sets and the system runs in a second in the environment of Pentium 4 PC with 3.4GHz CPU and 1.5G RAM.*

1. Introduction

There is a growing need to automate the process of cardiac ultrasound image analysis that involves many tasks such as cardiac view classification, wall motion analysis, measurement computation, automatic placement of Doppler gate over the valves, etc. Among all these tasks, cardiac view classification is the first step to achieve automation of the other tasks.

For example, for wall motion analysis [16], we require the cardiac view knowledge to regularize the motion analysis. For automatic placement of Doppler gate, we also need to know the cardiac view beforehand because each view shows different valves. For instance, apical four chamber view shows tricuspid valve and mitral valve and apical two chamber view shows only mitral valve.

There are several challenges in building an automatic system for cardiac view classification.

- *Within-view variation.* The image appearance in echocardiogram belonging to the same cardiac view characterizes significant variations, making it difficult to achieve high view classification accuracy. The variations arise from speckle noise inherent to ultrasound imaging modality, patient individuality, instrument difference, sonographer dependency, etc.
- *Between-view variation.* Apart from severe within-view appearance variations, how to characterize the between-view variation is another challenge too. Ideally, the global view template should provide maximum information about the characteristic chamber structure belonging to the view in a consistent fashion while discriminating different views. Designing global view templates is very crucial.
- *Structure localization.* To reduce the variation of the global view template, we confront the challenge of localizing the chambers (such as ventricles and atria) as their positions are unknown. This is an object detection/recognition problem which is an active research question in the computer vision literature. To robustly localize the individual chamber, we utilize information at a local scale.

We propose a fully automatic system for cardiac view classification (CardiacVC) of echocardiogram. This system employs a machine learning approach which extracts knowledge from annotated databases.

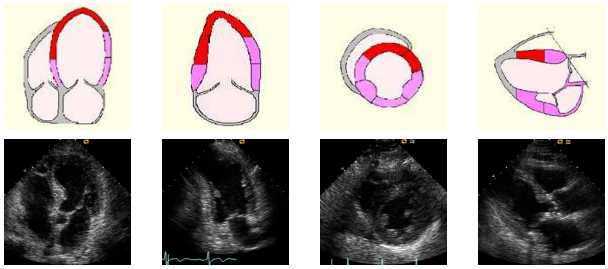
It possesses the following features:

1. *Integration of local and global structure.* It integrates evidence from both local and global scales. It utilizes local information to anchor the representative chamber such as left ventricle (LV) which is present in all four views. In the CardiacVC system, four LV detectors are used to extract local information because the system deals with four views. The global information is used when designing the global templates. This approach reduces within-view variation by aligning the global heart structure.
2. *LV Detector Dependent global view classification.* We design view-specific global templates to accommodate view characteristics based on aligning the representative chamber into canonical positions. Therefore,

for a given LV detector associated with a particular view, we use it to bootstrap examples from training sequences belonging to all of the four views and based on those examples to learn a multi-class classifier. Therefore, for each view, we have learned a LV detector dependent(LV-DD) multi-class classifier.

3. *Information fusion.* Given that each view has an LV detector and a multi-class classifier and hence produces its own result, there is a need to combines these results into one final output through intelligent information fusion.

Currently, we are focusing on four standard cardiac views: apical four chamber (A4C) and apical two chamber (A2C), parasternal short axis(SAX) at mid cavity and parasternal long axis (LAX), but the proposed approach is scalable to handle more than four views. Figure 1 illustrates the four cardiac views.



(a) (b) (c) (d)

Figure 1. illustrations of heart structure and example images. (a) A4C, (b) A2C, (c) SAX, and (d) LAX views.

2. Previous Works

To the best of our knowledge, four papers [1, 15, 11, 3] have been published so far to directly deal with automatic view classification of echocardiogram.

Ebadollahi *et al.*[3] suggested a part-based representation approach to recognizing the cardiac view. This method first detects heart chambers in cardiac echo images using the cavity detection algorithm proposed in [2]. Each view is represented by the constellation of the detected heart chambers, which is coded as Markov Random Fields (MRF) [8]. Finally, the energy vectors computed by matching a test image to the models are fed into Support Vector Machine to determine the final classification view. This method, however, does not guarantee good performance when cavities are falsely detected and/or missed which might frequently happen in noisy or zoomed-up echocardiogram.

In [15], Zhou *et al.* proposed a novel algorithm to tackle cardiac view classification using multi-class object detection approach. Unlike the other conventional approaches for multiple object detection that trains multiple binary classifiers (detectors), only one multi-class object detector is learned using the LogitBoosting algorithm [4] by including not only the positives corresponding to the cardiac views but also the background negatives. However, this method possibly yields contradicting detection results in an image. It needs a sophisticated method to handle the contradicting detection results to guarantee the high classification accuracy.

Otey *et al.*[11] proposed a two-level hierarchical classification approach combined with a simple dimensionality deduction approach. At the top level, it classifies an input sequence into either apical class or paresternal class, and then it further classifies the sequence into one of the four final views at the second level. They showed that this approach achieved 92.7% classification accuracy in testing.

Aschkenasy *et al.* [1] proposed a multiscale elastic registration algorithm [6] based on a continuous model of both images and deformation maps. In this algorithm, multiscale template images are constructed to represent the views in a spline domain using a third-order direct B-spline transform filter [13]. Both deformation energy and similarity between the warped image and its template image are used to classify the input image. This algorithm achieved 90% classification accuracy for training samples and 82.2% accuracy using the "leave-one-out" strategy. This method, however, needs proper templates to cover all the variations of the views. It is also sensitive to appearance variations introduced by translation, scale and rotation.

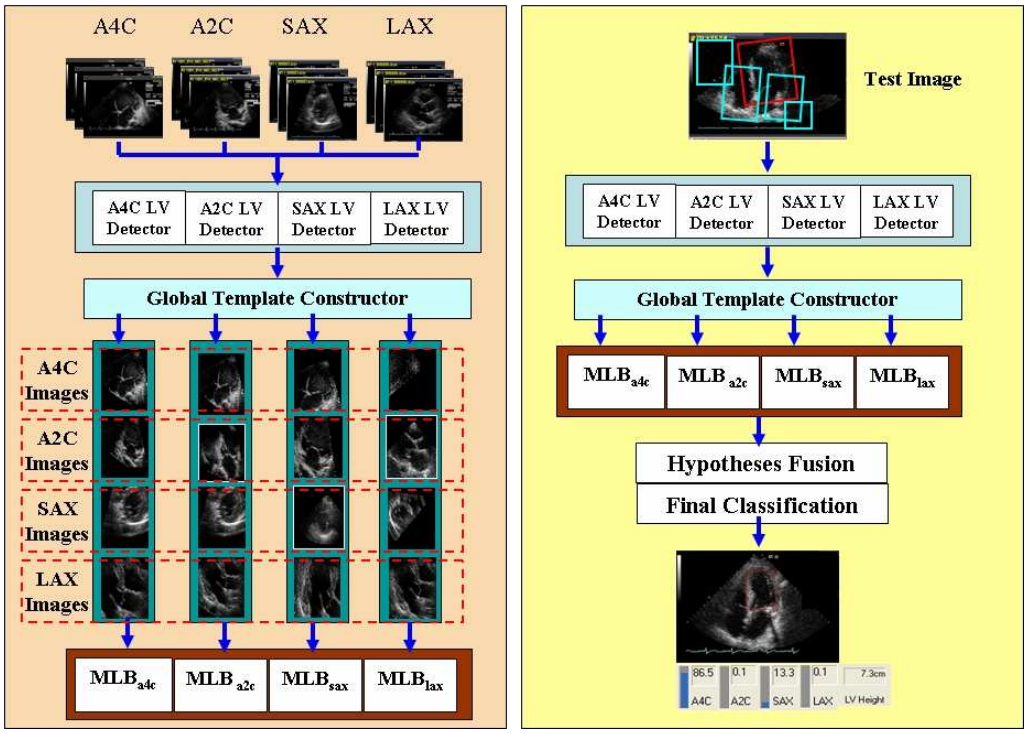
3. Algorithm Overview

The algorithm flow of the CardiacVC system is illustrated in Figure 2. It can be divided into two parts: (a) *offline model training* and (b) *system integration for on-line CardiacVC*.

The part of offline model training consists of three modules: (i) collecting training data and annotation, (ii) training LV detectors, and (iii) training multi-class view classifiers. We collect training samples of the four views and annotate LV endocardium using a contour.

We train an LV detector for each view class (four LV detectors in total). To train the LV detectors, we employ the object detection approach proposed in [12], which incorporates the Haar-wavelet type local features [10] and boosting learning technique. This method has been proven to be very efficient for object detection in real-time environment [14, 15, 5, 12, 7].

The system has four LV-DD global view classifiers. Each LV-DD global view classifier is trained using the training data collected by applying the according LV detector to all the training data, not only for the corresponding view but also the other views. For the LV-DD view classifiers, we also use the same Haar-wavelet type local features, but employ the multi-class LogitBoost (MLBoost) algorithm proposed in [4, 15].



(a) (b)

Figure 2. (a) The algorithm flow of the model training. (b) The algorithm flow of the on-line CardiacVC system.

The online CardiacVC system is roughly divided into three processes: (i) view-specific LV scanning (exhaustive searching), (ii) LV-DD global view classification, and (iii) fusion of classification outputs. To be specific, given an input cardiac video sequence, we first detect LV candidates by applying the LV detectors: one LV candidate per LV detector. Using the detected LV structures, we construct corresponding global templates and feed them into corresponding LV-DD multi-class view classifiers. We arrive at the final classification by combining the multiple classification results from the view classifiers.

4. Train Individual Components

4.1. Global view template

In this section, we will discus about the template design for LV-DD global view classifiers. The template should be designed to represent each view correctly and to minimize the intra-shape difference.

We design the global template based on the LV structure because it is relatively stable in echocardiogram. Since the LV cavity is the anchoring chamber used in our system, our template design is based on aligning the LV endocardial wall to the canonical location.

Figure 3 shows the three templates for global view classification used in the current CardiacVC system. As shown in Figure 3, A2C and A4C share the same template. The LV box in Figure 3 is a tight bounding box of the LV endocardial contour. Empirical evidence shows that it is not sensitive how to design each template as long as it contains all the structures of the heart as the heart structures are roughly spatially aligned.

In the training phase, we assume the knowledge of the LV endocardium, either manually traced for training of LV detectors or automatically inferred for global view classifiers as shown later. In the online system, the LV location is automatically computed.

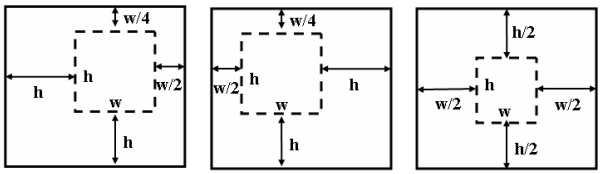


Figure 3. Template layout for the global view classification. From left to right: A4C/A2C, SAX, and LAX.

4.2. Multi-class Boosting

We employed the multi-class LogitBoost (MLBoost) algorithm proposed by Friedman *et al.* [4] and the feature selection approach proposed in [14] to build a global view classifier. MLBoost is a generalized version of two-class LogitBoost, which is another interpretation of Ada-boosting

using the forward additive logistic regression. The LogitBoost algorithm uses quasi-Newton steps [9] to fit an additive symmetric logistic model that maximizes the multinomial likelihood. In each iteration, it finds a $f_j(s)$ to satisfy Eq. (1) by using quasi-Newton step, and inserts it to the target function

$$El(F + f) \geq El(F), \tag{1}$$

where F and f stand for the target functions and a new function to be found at the current iteration, and $El(\cdot)$ denotes the expected log-likelihood. The MLBoost algorithm has the interpretation that it increases the classification accuracy for training data by adding a new function.

The output of the MLBoost algorithm is a set of response functions $F_j(x)$, one for each class.

$$F_j(x) = \sum_m f_{jm}(x). \tag{2}$$

The posterior probability of x being the j^{th} view is given by

$$p(j|x) = \frac{\exp(F_j(x))}{\sum_{i=1}^{J} \exp(F_i(x))} \tag{3}$$

These response functions $F_j(x)$ share the same so-called weaker learners (or weak classifiers) that are weighted differently for different views. The weak learners are selected and their coefficients are learned incrementally during boosting. We associate each weak learner with a local image filter by assuming $f_{jm}(x)$ is a piecewise constant function of the filter response. We use the same local gradient features (Haar wavelet style), and the same feature selection approach used in training of LV detectors. Hence, boosting operates as a feature selection oracle.

Figure 4 presents the MLBoost algorithm. We are given N training images from J classes (N_j training data points for the j^{th} class). The training data set is denoted as $\{(I_i, y_i)\}_{i=1}^{N}$, where I_i represents i^{th} training image, and is y_i a J-dimensional class indicator vector of I_i. Suppose that we generate M filters. Let us define a matrix $X_{N\times M}$ whose i^{th} row contains the M filter response collected from the image I_i. Since each filter is considered as a weak classifier (WC), the main goal of the training is to construct a classifier by selecting good filters among the huge filter pool for classification.

As mentioned earlier, for each LV detector, we learn a multi-class classifier. We denote the four classifiers by $p(j|x;k); k \in \{a4c, a2c, sax, lax\}$.

4.3. LV Detector-Dependent View Classifier

As discussed earlier, we trained four LV detectors, one per each view. Using each LV detector, we collect a full set of training data for a global view classifier. Each LV detector is applied to not only its corresponding view but also the other views to anchor the LV structure. The LV detector anchors true LV structure for the correct view, and it provides false positives for the other views. It is because of the similarity of LV structures along the views or the weakness of the LV detectors. We train the LV detectors not to selective because we are not able to obtain global evidence if the LV detector detects no LV structure.

LogitBoost (J classes)

1. Start with weights $w_{ij} = 1/N$, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, J$, $F_j(x) = 0$, and $p_j(x) = 1/J \; \forall j$.
2. Repeat for $m = 1, 2, ..., M$:
   - Repeat for $j = 1, 2, \ldots, J$:
     - Compute working responses and weights in the j^{th} class
     $$z_{ij} = \frac{y^*_{ij} - p_j(x_i)}{p_j(x_i)(1 - p_j(x_i))}; \tag{4}$$
     $$w_{ij} = p_j(x_i)(1 - p_j(x_i)). \tag{5}$$
     - (Δ) Fit the function $f_{mj}(x)$ by a weighted least-squares regression of z_{ij} to x_i with weights w_{ij}.
   - Set $f_{mj}(x) \leftarrow \frac{J-1}{J}(f_{mj}(x) - \frac{1}{J}\sum_{k=1}^{J} f_{mk}(x))$, and $F_j(x) \leftarrow F_j(x) + f_{mj}(x)$.
   - Update $p_j(x) \propto \exp(F_j(x))$.
3. Output the classifier $\arg\max_j F_j(x)$.

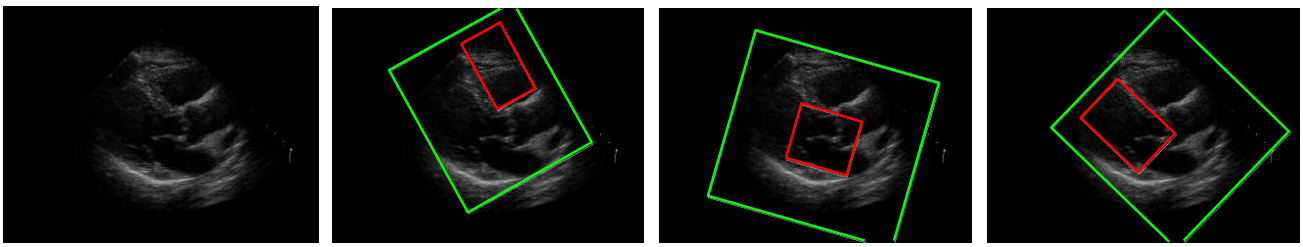
Figure 4. The multi-class LogitBoost algorithm [4].

We collected training images for the global view classifier by cropping according to the pre-defined template layouts shown in Figure 3. The images in each column just below the box of **Global Template Constructor** in Figure 2-(a) are the training images collected by applying each LV detector. Using these training images, we build four LV-DD global view classifiers using MLBoost algorithm. Let us denote the classifiers as MLB_{a4c}, MLB_{a2c}, MLB_{sax} and MLB_{lax}. MLB_{a4c} implies global view classifiers trained using the training data set provided by A4C LV detector and the others follow the same rule.

4.4. LV-DD Approach Vs. LV-DI Approach

The question might be raised why the system needs four view classifiers instead of one canonical view classifier which can be trained using the training data collected only applying each LV detector to its own view data (the diagonal of 4×4 image matrix in Figure 2-(a)).

If a trained LV detector always detects LV for its view and detects no LV for the other views, the view classification problem can be solved very easily. We can only depend on the LV detectors without the global view classi-



(a) Input image (b) A2C LV detector (c) SAX LV detector (d) LAX LV detector

Figure 5. The red box represents an LV candidate provided by each LV detector, and the green box represents global template for view classification. A4C LV detector fails to detect LV candidate

fiers. However, in reality, the trained LV detectors possibly yields false positives for the other views. If an LV detector is trained to decrease the false positives for the other views, it possibly increases the missed detection ratio for the correct view which is more malicious than the false positives. The false positives in the other views are inevitable in real situations.

Let us discuss further about the LV-DD view classification approach and the LV detector independent (LV-DI) view classification approach. Supposed that all the four LV detectors provide LV candidates given an input echo sequence and four global templates are constructed based on them. The one LV-DI approach may yield a good classification results only for a global template corresponding to the correct view. This approach, however, may provid random classification results for the other global templates because the pattern of the templates were not used in classifier training.

Therefore, the LV-DI approach possibly yields one good solution and three random solutions. By combining one true classification result and the three random classification results, it is hard to anticipate a good final classification result. This observation inspired us to propose LV-DD view classification approach. In this method, we learn four global view classifiers by considering the false positives from other views.

Figure 5 provides an example of LAX view to support the reason why the LV-DI approach is not robust enough. In this example, A4C LV detector fails to detect LV candidate while the other LV detectors provide their own best LV candidates which are shown from Figure 5-(b) to Figure 5-(d). The inner box represents a detected LV and the outer box represents the window to crop an image according to Figure3 to feed into a global view classifier.

Table 1 shows the comparison of the probabilities of global view classification produced by LV-DI approach and LV-DD approach using the three test images, I_{a2c}, I_{sax} and I_{lax}, provided by the three LV detectors. I_{a4c} is not used because A4C LV detector detects no LV structure in the video sequence.

The probabilities of I_{a4c} are set to be zeros. As shown in Table 1, both methods yield very good results for I_{lax} by classifying it to LAX by almost 100% accuracy. However, the LV-DI approach yields wrong result for I_{a2c} and I_{sax} by classifying it to SAX by 100% and 73% of probabilities respectively. The proposed LV-DD approach, however, classifies it correctly by 100% accuracy. The test sequence is finally classified as SAX using the LV-DD view classification approach as shown in the last row of Table 1.

5. Online CardiacVC System

In the run time, we classify an input echocardiogram video through three stages: 1) LV detection, 2) global view classification using four LV-DD multi-view classifiers, and 3) final cardiac view classification by integrating the classification results.

In the first stage, we employed the learned LV detectors (one LV detector per each view) to localize the LV candidate region. We used only the ED frame (and its neighboring frames if necessary) for classifying the query echo video. As shown in Figure 2-(b), each LV detector is applied to the test image by sliding a window on the ED frame from the top-left corner to bottom-right corner by changing the location, width, height and orientation. In Figure 2-(b), the blue boxes represent the cropped images to be fed into the LV detectors, and the red box represent the ground truth LV box. The box that yields the maximum detector score is used to construct the global template for global view classification. Therefore, we obtain four LV candidates, one per view, and subsequently four global templates that are denoted as I_{a4c}, I_{a2c}, I_{sax}, and I_{lax}, respectively.

In the next stage, each LV-DD multi-class view classifier is applied to its corresponding cropped global template. Finally, given four classification results, we use the following fusion strategy to arrive at a final classification result (eg, the total probability law).

	$p(k)$	LV-DD view classification				LV-independent view classification			
		$p(a4c\|I_k)$	$p(a2c\|I_k)$	$p(sax\|I_k)$	$p(lax\|I_k)$	$p(a4c\|I_k)$	$p(a2c\|I_k)$	$p(sax\|I_k)$	$p(lax\|I_k)$
I_{a4c}	0	0	0	0	0	0	0	0	0
I_{a2c}	1/3	0	0	1	0	0	0	0	1
I_{sax}	1/3	0	0	0.73	0.27	0	0	0	1
I_{lax}	1/3	0	0	0	1	0	0	0.01	0.99
$p(k\|I)$		0	0	0.57	0.43	0	0	0.002	0.998

Table 1. Comparison of view classification results between LV-DD approach and LV-independent approach

$$p(j|I) = \sum_{k\in K, p(k)>0} p(j|I_k;k)p(k) = \frac{1}{|k|}\sum_{k\in K, p(k)>0} p(j|I_k;k), \quad (6)$$

where $K = \{a4c, a2c, sax, lax\}$ and $|k|$ is the number k's to meet the condition of $k \in K, p(k) > 0$. Other fusion strategy can be applied too. In the above, the prior probability is assumed uniform. For instance, $p(k) = 1/4$ if all the global templates, I_k's, exist, and $p(k) = 1/3$ if only three of them exist as shown in Table 1. In practice, we can use the prior information extracted from the LV appearance.

Once the view is classified, we can determine the LV location accordingly and calculate measurements about the LV, such as the LV height. Such measurements provides useful feedback to the sonographer for probe adjustment toward better image quality.

6. Experiments

6.1. Cardiac View Classification and LV Localization

For the training purpose, we collected total 1080 video sequences with the LV endocardium annotated by experts (478 for A4C view, 310 for A2C view, 175 for SAX view, and 117 for LAX view). An LV endocardium is represented as an open contour with 17 landmarks for A4C, A2C and PSAX, and represented as a closed contour with 18 landmarks for SAX.

For test purpose, we collected 223 video sequences which were not used in training (96 for A4C view, 61 for A2C view, 28 for SAX view, and 38 for LAX view). The test dataset contains diverse sequences including not only the sequences which are similar to the typical view structures shown in Figure 1 but also some sequences which are quite different from the typical structures.

Table 2 presents the confusion matrix of the view classification results both of the training data and the test data computed using the proposed LV-DD global view classification approach. On the average, we achieved 96.4% accuracy on the training dataset and 96.3% accuracy on the test database, which are quite consistent.

Training Data	A4C	A2C	SAX	LAX
A4C(478)	97.9%	1.7%	0.4%	0%
A2C(310)	3.9%	95.2%	0.6%	0.3%
SAX(175)	0.6%	1.1%	97.7%	0.6%
LAX(117)	0 %	2.6%	2.6%	94.9%
Test Data	A4C	A2C	SAX	LAX
A4C(96)	97.9%	2.1%	0%	0%
A2C(61)	3.3%	93.5%	1.6%	1.6%
SAX(28)	3.6%	0%	96.4%	0%
LAX(38)	0%	0%	2.6%	97.4%

Table 2. The confusion matrix of view classification results computed using the proposed LV-DD approach

Table 3 shows the classification results of the test dataset using the LV-DI approach. This table shows that the LV-DI approach yields relatively good classification results for A4C and A2C sequences, but very poor results for SAX and LAX. It can be interpreted as follows. The LV structure of A4C and A2C are similar to each other and these two views share the same global template prototype for view classification. Therefore, we can assume that the classification results using A2C LV detector and A4C LV detector are relatively correct but the classification results using SAX LV detector and LAX LV detector might be random. When we compute the final classification by integrating the four classification results, the two correct classification results can dominate the other random classification results. However, SAX view and LAX view can have only one correct classification result and three random classification results. It might yield more classification errors compared to A4C view and A2C view. Refer to Table 1 to see one of the examples where a PLAX view is misclassified to PSAX because of this reason.

Test data	A4C	A2C	SAX	LAX
A4C(96)	93.8%	3.1%	3.1%	0.0%
A2C(61)	1.6%	93.4%	3.4%	1.6%
SAX(28)	3.6%	3.6%	89.2%	3.6%
LAX(38)	10.5%	0.0%	39.5%	50.0%

Table 3. The confusion matrix of view classifier results computed using LV-DI approach for test data.

Figure 6 shows view classification results along with LV localization for some of the selected test samples. The red box represents the LV location computed by the corresponding LV detector. The view classification results are overlaid on bottom-left corner of the images. There are four bars to represent the probabilities of A4C, A2C, PSAX and PLAX from left to right, and the actual probabilities are represented using blue bars. The longer blue bar, the higher probability. As shown in the experiment results, the proposed CardiacVC performs consistently well regardless of noise, LV diversity and missing heart structure. It also runs approximately 1 second using Pentium 4 PC with 3.4 GHz CPU and 1.5G RAM.

7. Conclusion

In this paper, we proposed an automatic system for cardiac view classification of echocardiogram. This system is built based on the data-driven machine learning techniques along with two key approaches. The first is to incorporate the local information (LV structure) and global information (global view by anchoring the LV structure). To obtain the local information of LV structure, we built an LV detector per each view by treating it as a binary classification problem. The second is to employ the view-specific knowledge for global view classification to properly incorporate the local information to global information. In other words, rather than training one LV-independent global view classifier, we built four LV-DD global view classifiers, MLB_{a4c}, MLB_{a2c}, MLB_{sax} and MLB_{lax}.

As shown in the experimental results, the proposed CardiacVC system performs very well for the four standard views, A4C, A2C, SAX and LAX even though some of the test data are very challenging because of the noise and missing parts. It achieves a classification accuracy over 96% both of training and test data sets and the system runs in a second using a Pentium 4 PC. Even though the current prototype system deals with only four views but it can be easily generalized to classify more views.

Future direction of the CardiacVC system is to handle not only the predefined views but also the negative class, *aka, non of the predefined views*. The current CardiacVC system always classifies an input cardiac sequence to one of predefined views even though the sequence is not from the views. One of the possible solutions is to use the method proposed in [15], which trains classifiers by including negative class along with the predefined classes. However, including the negative class makes the problem more complicated because the space of the class is infinite.

References

[1] S. V. Aschkenasy, C. Jansen, R. Osterwalder, A. Linka, M. Unser, S. Marsch, and P. Hunziker. Unsupervised image classification of medical ultrasound data by multiresolution ela stic registration. *Ultrasound in Medicine and Biology*, 32(7):1047–1054, 2006. 2

[2] D. R. Bailes. The use of the gray level sat to find the salient cavities in echocardiograms. *Journal of Visual Communication and Image Representation*, 7(2):169–195, 1996. 2

[3] S. Ebadollahi, S.-F. Chang, and H. Wu. Automatic view recognition in echocardiogram videos using parts-based representation. In *CVPR*, volume II, pages 2–9, 2004. 2

[4] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. *Ann. Statist.*, 28(2):337–407, 2000. 2, 3, 4

[5] B. Georgescu, X. S. Zhou, D. Comaniciu, and A. Gupta. Database-guided segmentation of anatomical structures with complex appearance. In *CVPR*, pages 429–436, 2005. 2

[6] J. Kybic and M. Unser. Fast parametric elastic image registration. *IEEE Transactions on Medical Imaging*, 12(11):1427–1442, 2003. 2

[7] S. Li and Z. Zhang. Floatboost learning and statistical face detection. *PAMI*, 26(9):1112–1123, September 2004. 2

[8] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, 2001. 2

[9] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 1999. 4

[10] M. Oren, C. Papageorgiou, S. P., O. E., and T. Poggio. Pedestrian detection using wavelet templates. In *CVPR*, pages 193–199, 1997. 2

[11] M. Otey, J. Bi, S. Krishna, B. Rao, J. Stoeckel, A. S. Katz, J. Han, and S. Parthasarathy. Automatic view recognition for cardiac ultrasound images. In *in Proceedings of Int'l Workshop on Computer Vision for Intravascular and Intracardiac Imaging*, pages 187–194, 2006. 2

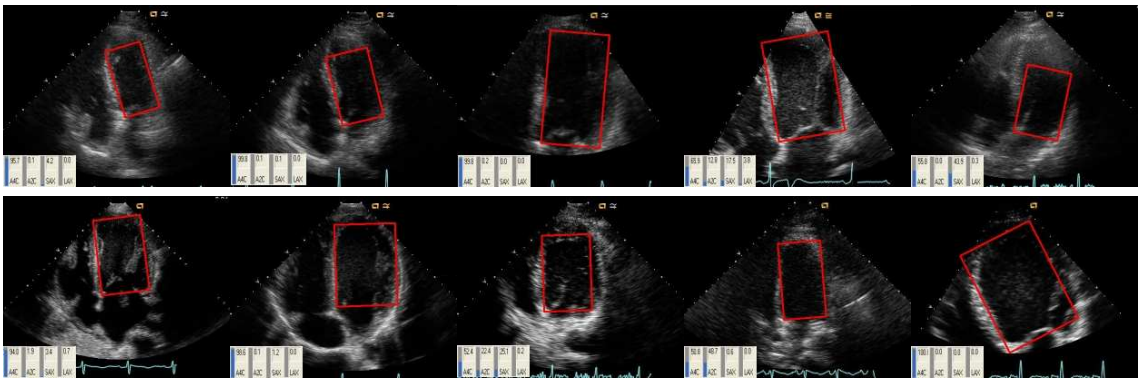
[12] Z. Tu. Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In *ICCV*, 2005. 2

[13] M. Unser, A. Aldroubi, and M. Eden. Fast b-spline transforms for continuous image representation and interpolation. *PAMI*, 12(3):277–285, 1991. 2

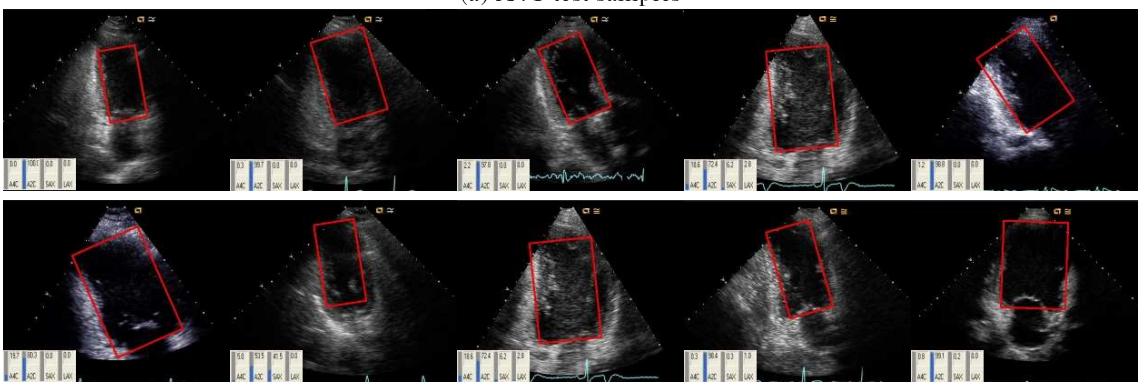
[14] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR*, pages 511–518, 2001. 2, 3

[15] S. Zhou, J. Park, B. Georgescu, J. Simopoulos, J. Otsuki, and D. Comaniciu. Image-based multiclass boosting and echocardiographic view classification. In *CVPR*, pages 1559– 1565, 2006. 2, 7

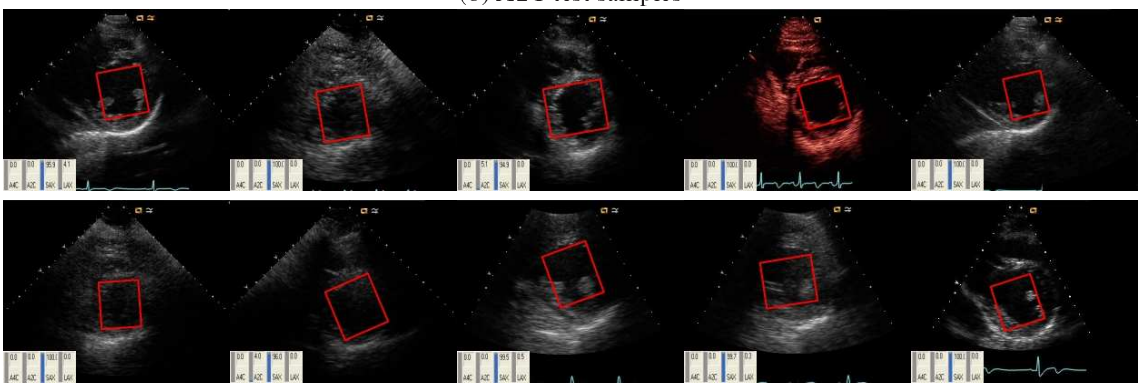
[16] X. S. Zhou, D. Comaniciu, and A. Gupta. An information fusion framework for robust shape tracking. *PAMI*, 27(1):115–129, January 2005. 1



(a) A4C test samples



(b) A2C test samples



(c) PSAX test samples

(d) PLAX test samples

Figure 6. Selected view classification results along with LV localization for test data