Optimization of Wavelets for Time Series Classification with Support Vector Machine

Karndeep Rai-Bhatti
Department of Electrical and Computer Engineering, 251C
University of California, San Diego
La Jolla, CA, United States
kraibhat@eng.ucsd.edu

William Argus
Department of Electrical and Computer Engineering, 251C
University of California, San Diego
La Jolla, CA, United States
wargus@eng.ucsd.edu

***Abstract*—Robust real time labeling of time series data in the form of electrocardiograms relies upon effective feature extraction. Careful selection of features can reduce the computational complexity and increase the prediction accuracy, particularly in instances where training data is limited. This work was conducted with the goal of defining a new approach to extract low dimensional frequency-dependent features from time series ECG data for use in classification. Wavelet multi-resolution analysis is often utilized to extract a feature set from time-series data. The novelty of the approach outlined in this paper lies in the way in which the wavelet is defined. Rather than selecting a predefined wavelet, a custom wavelet is found via particle swarm optimization. Once selected, wavelet features undergo dimensional reduction by principal component analysis before being used to train a one-vs-one support vector machine classifier. The dataset consisted of six classes, representing six beat types present in the ECG data of forty-eight patients, taken from the MIT-BIH arrhythmia database. Two optimization schemes were implemented, record based and beat based, resulting in two different optimized wavelets. Both wavelets were evaluated with 10-fold cross validation, using a train-test split of 20/80 selected at random, as well as a patient specific scheme where only initial data, 300 beats or 60 beats, was used for training. The wavelet optimized from the record based scheme achieved 98.79%, 97.54%, and 94.58% classification accuracy, respectively on the three evaluation scenarios. The wavelet optimized from the beat based scheme achieved 98.46%. 97.80%, and 94.05%, respectively on the three evaluation scenarios.**

I. INTRODUCTION

An electrocardiogram is a measure of the electrical activity of the heart with respect to time, and is commonly used in cardiovascular diagnosis. Heart arrhythmias are conditions where the heartbeat deviates from what is normally expected, either beating too slow, too fast, or irregularly [1]. These irregularities translate to identifiable differences in the electrical activity recorded by the electrocardiogram. Clearly, identifying heartbeat irregularities is of critical importance in healthcare. Developing a method for automatic classification of heartbeats will allow for efficient and timely detection, reducing the workload on medical professionals. As the method deals with critical health data classification, it must be robust to a large diverse set of data and be able to achieve high levels of accuracy.

The first step to achieving robust and accurate classification of ECG signals is the extraction and selection of meaningful features from the signal. ECG are measured as a discrete sampling of voltage over time, with a distinct waveform representing one heartbeat at regular intervals. A single waveform represents one heartbeat, and the shape of the waveform is what determines what class of heartbeat (either normal or some type of arrhythmia), the waveform represents. Features are extracted from an ECG signal either in frequency domain, time domain, obtaining spectral or morphological features, respectively. Once features are extracted and selected, classification can be performed with a variety of methods, with most modern methods using either support vector machines (SVM), or neural networks (NN) [2, 3, 4].

There have been many other works published on methods to extract features from ECG signals. The work of Min et al utilized a Discrete Wavelet Transform (DWT) for feature extraction from ECG signals with the goal of labeling beats as one of five different classes. This work focused on analyzing different methods for both dimensional reduction of the extracted features and classification following dimensional reduction. For dimension reduction, Principal Component Analysis, Linear Discriminant Analysis, and Independent Component Analysis were used, and for classification, Support Vector Machine, Neural Network, and a Probabilistic Neural Network were trained and tested. This method achieved 99.97% accuracy using 10-fold cross validation [5].

Acharya et al focused on exploring different methods for feature extraction, specifically Discrete Wavelet Transform, Empirical Mode Decomposition, and Discrete Cosine

Transform (DCT). The extracted features were then subjected to dimensional reduction for feature selection by applying Locality Preserving Projection and then selecting a predetermined number of the highest ranked components. To classify these features, this work used a K-Nearest Neighbor (kNN) classifier, achieving 98.5% accuracy through the DCT feature extraction technique which proved to be the best performing of the three [6].

Melgani et al published a method that focused on the classification aspect of ECG classification. Feature extraction was done using Discrete Wavelet Transform and feature selection varied across several test set-ups from using PCA to doing no dimensional reduction. The method's novel approach was using Particle Swarm Optimization (PSO) to improve their SVM classifier. To demonstrate the performance, the PSO-SVM was compared to SVM, kNN, and Radial Basis Function (RBF) classifiers. The PSO-SVM achieved accuracy of 89.72%; better than the SVM, kNN, and RBF-NN classifiers which achieved, respectively, 85.98%, 83.70%, and 82.34%. One critical part of this work is the small number of samples used to train the classifier, which represents a realistic use-case scenario for ECG classification [7].

Liu et al, which this paper's method is most closely related to, used Biorthogonal (6,8) wavelet for feature extraction, and then PCA for dimension reduction. This paper showed that by choosing the 12 most significant principle components, 95% of the feature information can be retained, and the data is reduced to 4.76% of its original size. The reduced features are then used by an SVM for classification of beat type. This work defined two different classification scenarios for ECG classification. The first scenario, designated as beat-based classification, split the dataset up according to each beat. In other words, data was not separated by person and the training and testing was done on a randomized split of all data. The second scenario, designated record-based classification, split the dataset by individual, meaning that all the data collected from one person was either used in training or testing, which was aimed at improving the robustness of the method for classifying ECG data of previously unseen individuals. The beat-based scenario achieved classification accuracy of 99.70%, and the record-based scenario achieved accuracy of 81.47% [8].

Based on the previous research, there is a need for an ECG classification method that can be readily applied to ECG data from new patients. Some of the aforementioned literature managed to achieve over 99% classification accuracy in testing, however the testing methods used to achieve these results commonly had one or both of the following deficiencies. Firstly, their training data was selected randomly from the entirety of the dataset. This is an unrealistic scenario since it implies that in order to accurately classify a person's ECG signal, data from the entirety of the signal must be labeled and used for training, which defeats the purpose of automatic ECG classification. The second problem was the large amount of training data required for accurate results, as this results in time consuming and computationally costly training. This paper aims to improve on these deficiencies by developing a method for accurate ECG classification that requires training on only a small amount of ECG beat samples, all taken from the beginning of an ECG dataset for a person. This would allow the beginning of an ECG to be hand labeled and then used to train the classifier, which will then automatically classify the remainder of the ECG data for that individual.

In order to achieve this goal, the approach presented here is derived from the method developed by Liu et al., with the difference being the wavelet used for feature extraction will be randomly initialized, and then optimized using Particle Swarm Optimization (PSO). The wavelet will be optimized under two different scenarios, beat-based and record-based train-test splits of the data. Three testing scenarios for evaluation are defined to be a 20/80 randomly selected train-test split of all the data, a 300/rest test-train split from a single individual, and a 60/rest test-train split on a single individual. The accuracy of feature extraction, selection, and classification using each of the two optimized wavelets is then compared to the accuracy when using the Biorthogonal (6,8) wavelet used by Liu et al.

II. METHODS AND APPROACH

A. *Dataset*

This dataset consists of ECG readings from 48 patients that exhibit various types of arrhythmia. Each patient's dataset is taken from a 30 minute ECG recording and band-pass filtered to remove noise outside of 0.1 to 100 Hz. The ECG signal is sampled at 360 Hz and then divided into 0.7 second segments each containing 1 beat. By dividing the dataset in this way, each segment is considered 1 sample, that contains 1 beat, and has 252 discrete measurements in the units of mV. Each sample was then labeled, with five different beat types considered: one normal label (N), and four labels for different types of arrhythmia (A, L, R, V).

B. *Method overview*

The flowchart outlining the proposed optimization method is shown in figure 1. This begins by segmenting ECG signals beatwise, as described above, with each sample having length 252. To begin the particle swarm optimization, particles were initialized with each having randomly defined wavelet parameters and velocities. These wavelets were then used in Wavelet Mult-Resolution Analysis (WMRA) on the data to extract features. These features underwent dimensionality reduction via PCA, and then were used to train low-dimensional wavelet feature SVMs, following the method proposed by Liu et al. From the trained SVMs, a classification error was tabulated for each particle, and used to update all of the particle's wavelet parameters and velocities. This process was repeated until the error converged to a global minimum or another termination condition was met.

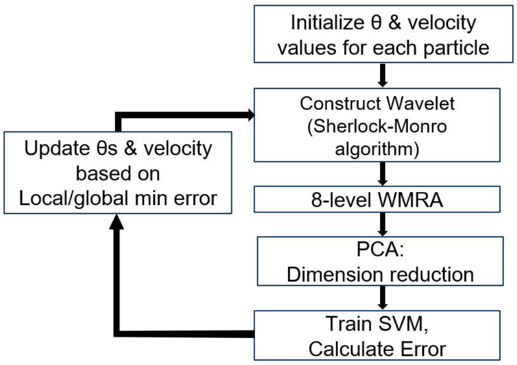

Figure 1: Flow chart of the wavelet optimization technique. Particle swarm optimization was used to minimize the classification error of a Low-dimensional Wavelet feature based SVM. This is done by altering the parameters of the wavelet used in the transformation.

C. Low-Dimensional Wavelet Feature based SVM

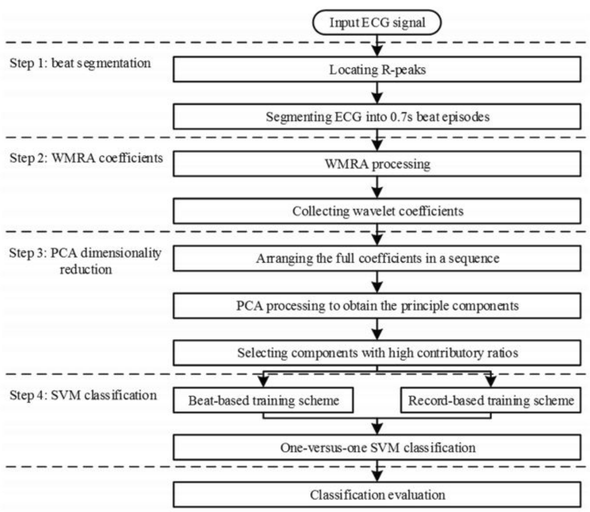

Figure 2: Flow of the ECG classification framework proposed by Liu et al.

In order to optimize the wavelets, a process for generating the classification error used in optimization is required. This was achieved with the low-dimensional wavelet feature based support vector machine proposed by Liu et al [8]. Their procedure consists of segmenting individual beats from a time series and applying an 8-level wavelet decomposition using the bior6.8 wavelet, iterating on the low-passed signal. The principle components were then calculated, and data was projected onto the first 12 principle components. 12 principal components were chosen by the developers of the method since adding more than 12 principal components did not result in an increase in accuracy.

These compact feature vectors were then used to train and test a one-vs-one SVM classifier. This classifier works by constructing $n(n-1)/2$ hyperplanes from n classes, and then uses a voting system amongst all the hyperplanes [9]. Each hyperplane gets a vote for which class the input data should be labeled as, and the class receiving the most votes is outputted as the label for that input. The process is described mathematically below, with $\phi(\boldsymbol{C}, \boldsymbol{I})$ being the output label given a vector of classes $\boldsymbol{C}$ and data input $\boldsymbol{I}$, $\nu(p, \boldsymbol{C}, \boldsymbol{I})$ being the vote of hyperplane p, and $\delta(p, \boldsymbol{I})$ being the classification of hyperplane p, given data $\boldsymbol{I}$.

$$\phi(C, I) = \sum_{p=1}^{n(n-1)/2} \nu(p, C, I) \quad (1)$$

$$\nu(p, C, I) = 1 \; \textit{if} \; \delta(p, S) = C, \; 0 \; \textit{otherwise} \quad (2)$$

The specified wavelet, decomposition depth, and number of principal components were determined experimentally. For the purposes of this paper, all variables besides the wavelet coefficients were reused and held constant.

D. Orthonormal Wavelet Parameterization

The set of orthonormal wavelets was chosen as a source of potential candidates due to its ease of parameterization. A pair of orthonormal decomposition filters of length 2N can be factored into a series of N rotation and delay matrices as seen in equation 3 .

$$H_p(z) = \begin{pmatrix} c_0 & s_0 \\ -s_0 & c_0 \end{pmatrix} \prod_{i=1}^{N-1} \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \begin{pmatrix} c_i & s_i \\ -s_i & c_i \end{pmatrix} \quad (3)$$

Thus, in order to find the optimal length 2N orthonormal wavelet, it is only necessary to optimize over an N dimensional vector space corresponding to the rotation angles. Furthermore, each dimension is bounded between 0 and 2π radians due to the periodicity of the sine and cosine functions, further limiting the search.

Given a collection of angles, the wavelet can be constructed using the recursive formula derived by Sherlock and Monro [10].

E. Particle Swarm Optimization

PSO is a population-based optimization method used to search a vector space for a global minimum. The approach begins by defining a number of particles, with randomly selected parameters and velocities . A swarm size of 40 particles was used for all of the experiments. The vector space that is being optimized over consists of the angle parameter space used to define an orthonormal wavelet. Each angle is drawn from a uniform distribution between 0 and 2π radians. For simplicity, a fixed vector space dimensionality of 8 was selected, to match the size of the reference wavelet, Bior6.8.

The randomly generated angles for each particle are then used to generate a pair of decomposition low-pass and high-pass filters using the recursive Sherlock-Monro algorithm. An SVM is then trained for each particle as described above, utilizing features defined by each particle's filter pair. For the experimentation, each classifier was trained on the same 500 beats. Error was calculated by testing the classifiers on a set of 10,000 samples and taking the fraction of incorrect predictions.

$$v_i(t+1) = wv_i(t) + c_1 \cdot r_1(t)(p_{bi}(t) - p_i(t)) + c_2 \cdot r_2(t)(p_g(t) - p_i(t)) \quad (4)$$

$$p_i(t+1) = p_i(t) + v_i(t) \quad (5)$$

The error of each particle served as the metric that the PSO sought to minimize. The angle and velocity parameters of each particle were updated using equations 4 and 5, where $v_i(t)$ is the velocity of the i^{th} particle at time t. The variable $p_i(t)$ is the position of the i^{th} particle in the parameter space(angle vector), p_{bi} and p_g are the best positions found by all particles in the past and of the current generation respectively. c_1, c_2, and w are constants. For experimentation, c_1 and c_2 were set to 1.3, and the inertia was bound by 0.1 and 1.1. Finally, $r_1(t)$ and $r_2(t)$ are random variables drawn for a uniform distribution ranging from 0 to 1 at time t.

Optimization was terminated following either 200 iterations or 30 iterations where a new global minimum is not found under the assumption of convergence.

F. *Calculating Error within Optimization Loop*

The methodology for the testing and training that occurs within the optimization loop in order to find the error at each iteration is described as follows. As mentioned previously, two different wavelets were obtained with two different optimization methods, in order to determine which produced the most accurate classification. Accordingly, two different train and test set splits were utilized. The first , beat-based training, consisted of training on 500 beats randomly selected across all individuals and testing on 10,000 additional randomly selected samples. The second configuration, record-based training, also had 500 training beats and 10,000 testing beats, with the added restriction that the patients used in the training set were excluded from the testing set. The reason for the relatively small amount of training data during optimization is due to the large computational expense of PSO combined with the limited computational resources of personal computers. Cross validation was also not performed during optimization for the same reason.

G. *Evaluation Method*

In order to test the optimized wavelet, this paper utilized the same method proposed by Liu et al of WMRA, dimensionality reduction (PCA) and classification (SVM) that was used within the optimization loop. The two wavelets obtained from beat-based and record-based optimizations, respectively, were compared to each other, and to the standard bior6.8 wavelet used by Liu et al. The metric used for comparison was classification accuracy: number of correctly labeled samples divided by total number of samples. Evaluation used three different testing scenarios.The first scenario used was a 20/80 randomized train test split across data from all patients. The second and third scenarios were done on a patient specific basis where initial data - either the first 300 or the first 60 beats - were used for training and the remainder of the data from that patient was used for testing. The goal of the first scenario was to test the accuracy of the wavelets in a standard ECG classification evaluation. The second and third scenarios represent more realistic and consequently more difficult testing approaches in ECG classification, as described in the introduction of this paper, where there is only a small amount of ECG data from the beginning available for classification.

H. *Dataset and Data Availability*

The data was sourced from the MIT-BIH Arrhythmia Database developed specifically for investigating ECG rhythm analysis. The data is available at physionet.org. Additionally, the required library to read the data, the WFDB toolbox for Matlab, can also be found on their website.

III. RESULTS

A. *Optimization*

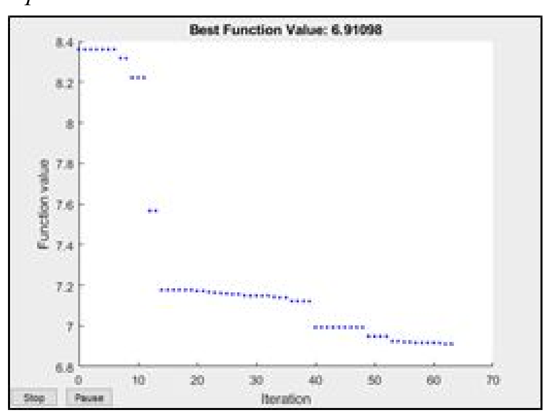

Figure 3: Plot of number of iterations and the minimal classification error found up to that point,beat-based optimization .

By utilizing PSO, significant improvements in prediction accuracy were made while using a limited training size. In both beat-based and record-based training, an approximate 1.5% increase in maxima global accuracy was observed. Both optimizers converged following approximately 60 iterations of optimization.

Figure 4 shows the wavelets found via the optimization. It should be noted that the optimized wavelets are a great deal wider than the reference bior6.8 wavelet and do not match the typical expectations of high and low pass filters.

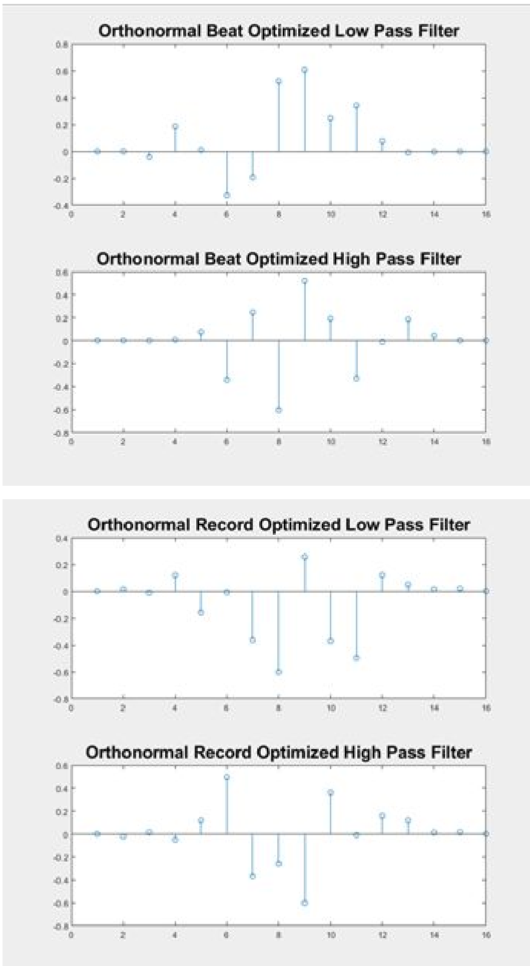

Figure 4: Top: Low pass and high pass filters generated using the beat based optimization scheme. Bottom: Low pass and high pass filters generated using the record based optimization scheme.

B. *Classification*

Once the optimized wavelets were generated, their performance was quantified and compared to the baseline standard Bior6.8. Liu et. al tested a large set of wavelets and found Bior6.8 to consistently outperform the others in classification tasks. For this reason, it was chosen to serve as a point of reference.

Train/Test Split	Bior6.8	Record Based Optimization	Beat Based Optimization
Random 20/80	96.20%	98.79%	98.46%
300 beats/ Remainder	97.30% (4%)	97.54% (3.85%)	97.80% (3.69%)
60 beats/ Remainder	94.35% (7.97%)	94.58% (7.86%)	94.05% (9.04%)

Table 1: Comparison of the performance of Bior6.8 and the two optimized wavelets from the 3 tests. For Random 20/80 an average prediction accuracy is listed following 10 fold cross validation. For the remaining two tests, average accuracy and standard deviation across individuals are listed.

For each wavelet, three tests were performed to quantify their performance. Each test utilized the wavelet feature based SVM framework discussed above. The first test consisted of randomly dividing 20% of all of the patient data into a training set, and using the remainder for testing.This is a common training scheme utilized by many authors and tends to overestimate the classifier's real-world performance due to over fitting. Under this condition, both optimized wavelets outperformed the standard wavelet by over 2% accuracy. The remaining tests consisted of training on fewer and fewer data points in a more realistic structure. An SVM was independently trained for each patient using the first 300 beats and 60 beats, roughly 5 minutes and 1 minute of data, respectively. In this temporally structured training method, the three wavelets performed identically, with none of them having a clear advantage.

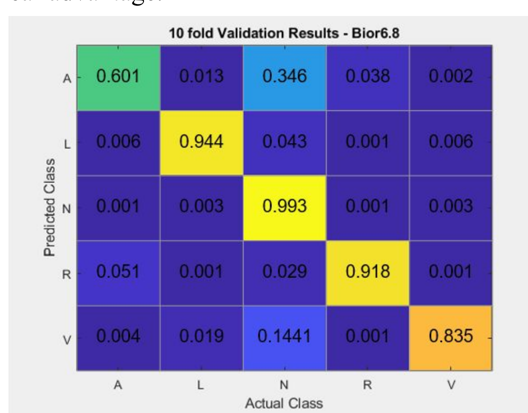

Figure 5: Bior6.8 Average confusion matrix following 10 fold cross validation

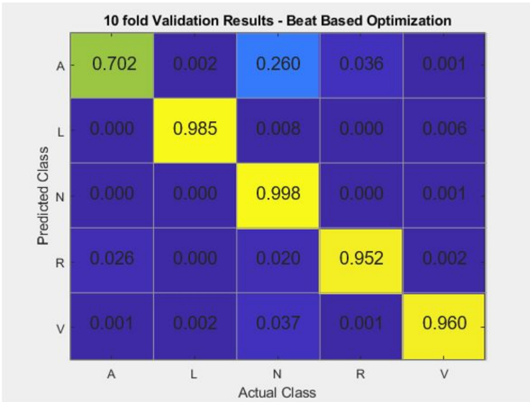

Figure 6: Average Confusion matrix following 10 fold cross validation utilizing the wavelet optimized under the Beat Based scheme. Beat-Based wavelet exhibits increased accuracy compared to Bior6.8, while still being outperformed by the record-based counterpart.

Figure 7: Average Confusion matrix following 10 fold cross validation utilizing the wavelet optimized under the Record Based scheme. This Wavelet outperformed all others in specificity, only showing major misclassification in one group.

The confusion matrices from the Random 20/80 train and test split evaluation scenario are seen above in figures 5, 6 and 7. Across both of the optimized wavelets, as well as the Bior6.8 wavelet, the by-class results of classification accuracy are greater than 95% except for the prediction of arrhythmia type A. Beats that are classified as type A are in fact normal heart beats and are thus misclassified 21.3%, 26%, and 18.7% of the time for the Bior6.8, beat-based optimized wavelet, and record-based optimized wavelet, respectively. It should however be noted the specificity for all categories increases when comparing the optimized wavelets with Bior6.8.

IV. DISCUSSION

The method of optimizing a wavelet with particle swarm optimization produced better results than the standard on the first test: the conventional randomized training and testing data split. This demonstrates that although the Bior6.8 wavelet was identified by Liu et. al as the best standard wavelet for feature extraction, there is still potential for improvement in feature extraction space of the ECG classification problem. Though this first test does not necessarily translate into guaranteed improvement in real work results, it does show that the theory behind the PSO approach proposed in this paper is valid.

The second and third tests demonstrated that the optimized wavelet does not demonstrate any significant improvement over the Bior6.8 wavelet in terms of real world training and testing scenarios. Despite lack of improvement, there are positive takeaways from these two tests. The first is that the optimized wavelet did not give worse results than the Bior6.8 wavelet, implying that there may be an upper limit for the classification accuracy in the Liu et al method. The second positive takeaway comes when the training and testing scenario during the optimization feedback loop is considered. As previously mentioned, the amount of training data in this loop was small, and cross-validation was not implemented, as doing so would have prevented the PSO from successfully converging due to the computational limits of the training hardware. This means that with access to better hardware, the PSO could potentially achieve better results when optimizing the wavelets, translating into better classification results.

One cause of error that is consistent across all three wavelets is the misclassification of normal heartbeats as arrhythmias of type A. This is seen in the confusion matrices in the results section and is the only significant and consistent error. As it is preventalent across all three wavelets, this error is likely due to limitations in the SVM classifier. Improvements could be made by utilizing the same PSO optimization scheme proposed in this paper, but with different methods of dimensional reduction and classification.

Despite only equaling the results of the Bior6.8 wavelet in a realistic scenario, both of the optimized wavelets achieved accuracy over 94% in the most stringent test case. By subjecting the method proposed in this paper to realistic testing scenarios, as opposed to optimal scenarios with very large training sets that are used in many related works to produce accuracies over 99%, the authors hope to show that strong ECG signal classification results can be achieved even in realistic scenarios.

V. CONCLUSION

In creating a method for optimizing a wavelet used for feature extraction in an SVM-based classification process, this work has aimed to improve the labeling of heartbeats from an ECG signal. The signal is first segmented into beats that are 252 samples in length each. Then the wavelets that will be used for feature extraction are optimized with Particle Swarm Optimization, under 2 different schemes, beat-based and record-based, to produce two different, optimized wavelets. In

PSO, the wavelet to be optimized is used to find the wavelet coefficients of each beat, these coefficients undergo dimensional reduction with PCA selecting the 12 most significant components, and are then used to train a one-vs-one SVM classifier. The error of the accuracy of the classifier is used as feedback in the optimization loop to iteratively improve the wavelet. The same method of wavelet decomposition, PCA, and then SVM is used for testing once the optimized wavelets are found. Compared to the bior6.8 wavelet, both the wavelet optimized in a beat-based scheme and the wavelet optimized in a record-based scheme gave over 2% higher classification accuracy on the standard evaluation of a randomized train-test split of the data. In the two real world testing scenarios of training only on the first 300 and first 60 beats, respectively, the two optimized wavelets performed equal to the bior6.8 wavelet. This shows that PSO is a viable method for obtaining wavelets to be used in ECG feature extraction. It is also worth noting that a consistent and common source of error across all three wavelets is the misclassification of a normal heartbeat as a class of arrhythmia. In future work, with more computational resources, the optimization process can be modified to train on a larger amount of data during each optimization loop, as well as use cross-validation during each optimization loop, to increase the robustness of the optimized wavelet. Another area that should be investigated is the consistent misclassification of normal beats as type A arrhythmia, as solving this problem, potentially by implementing a different classifier instead of SVM, can greatly improve the classification accuracy of the method as a whole.

REFERENCES

[1] Aya, F. K., Mohamed, I. O. & Inas, A. Y. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. *Expert Systems with Applications* 42, 8361–8368 (2015).

[2] Zadeh, Ataollah & Khazaee, Ali & Ranaee, Vahid. (2010). Classification of the electrocardiogram signals using supervised classifiers and efficient features. Computer methods and programs in biomedicine. 99. 179-94. 10.1016/j.cmpb.2010.04.013.

[3] Abibullaev, Berdakh & Seo, Hee. (2011). A New QRS Detection Method Using Wavelets and Artificial Neural Networks. Journal of medical systems. 35. 683-91. 10.1007/s10916-009-9405-3

[4] Li, Pengfei & Wang, Yu & He, Jiangchun & Wang, Lihua & Tian, Yu & Zhou, Tian-shu & Li, Tianchang & Li, Jing-song. (2016). High-Performance Personalized Heartbeat Classification Model for Long-Term ECG Signal. IEEE Transactions on Biomedical Engineering. 64. 1-1. 10.1109/TBME.2016.2539421.

[5] Martis, Roshan & Acharya, U Rajendra & Lim, Choo. (2013). ECG beat classification using PCA, LDA, ICA and discrete wavelet transform. Biomedical Signal Processing and Control. 8. 437-448. 10.1016/j.bspc.2013.01.005.

[6] Acharya, U Rajendra & Fujita, Hamido & Adam, Muhammad & Oh, Shu Lih & K Sudarshan, Vidya & Tan, Jen Hong & Koh, Joel En Wei & Hagiwara, Yuki & Chua, Kuang & Poo, Chua & Tan, Ru San. (2016). Automated Characterization and Classification of Coronary Artery Disease and Myocardial Infarction by Decomposition of ECG Signals: A Comparative Study. Information Sciences. 377. 10.1016/j.ins.2016.10.013.

[7] Melgani, Farid & Bazi, Yakoub. (2008). Classification of Electrocardiogram Signals With Support Vector Machines and Particle Swarm Optimization. Information Technology in Biomedicine, IEEE Transactions on. 12. 667 - 677. 10.1109/TITB.2008.923147.

[8] Qin, Qin & Li, Jianqing & Zhang, Li & Yue, Yinggao & Liu, Chengyu. (2017). Combining Low-dimensional Wavelet Features and Support Vector Machine for Arrhythmia Beat Classification. Scientific Reports. 7. 10.1038/s41598-017-06596-z.

[9] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[10] B. G. Sherlock and D. M. Monro, "On the space of orthonormal wavelets," in *IEEE Transactions on Signal Processing*, vol. 46, no. 6, pp. 1716-1720, June 1998, doi: 10.1109/78.678504.

TEAM MEMBER CONTRIBUTIONS

The team member contributions are as follows. Karneep worked on the code for the wavelet transform implementation, the SVM and PSO implementation, and the wavelet optimization and testing. William worked on the code for the loading and partitioning of data for model training, the implementation of cross validation, implementation of PCA, and the literature review. Both presentations and the report were a joint effort, with the work divided equally between the two group members, and each group member proof-reading and amending the other’s sections as necessary. Numerically the total work divide was 50%-50%.

GITHUB REPOSITORY:

https://github.com/Karndeep-UCSD/Wavelet-Optimization-for-Time-Series-Classification

https://github.com/warg15/Wavelet-Optimization-for-Time-Series-Classification

Both repositories contain the same code.