Classification of Magnetomyography Signals using Discrete Wavelet Transform and Genetic Algorithm

[1]T. Ananda Babu, [2]Dr. P. Rajesh Kumar
[1]Research Scholar, [2]Professor & Chairman BoS
[1]Department of Electrical and Communication Engineering,
[1]Andhra University College of Engineering (A), Andhra University, Visakhapatnam, India

Abstract : The classification of multichannel uterine magnetomyography signals attempted using discrete wavelet transform and genetic algorithm. The uterine magnetomyographic signals analyzed in this research for the detection of term labor. The MMG signals of Physionet mmgdb database decomposed with six level discrete wavelet transform. The features extracted are energy, waveform length, standard deviation, entropy and variance from the discrete wavelet transform coefficients. Significant features selected using genetic algorithm. The features are fed to different classifiers for the labor assessment. The performance of classifier calculated by using different mother wavelets. The support vector machine classifier trained with GA selected features is good for classifying the pregnancy and labor with an accuracy of 95.9425%. The experimental results obtained will be helpful in term labor monitoring.

Index Terms **- Discrete wavelet transform, Labor prediction, Uterine magnetomyography, Genetic Algorithm (GA)**

I. INTRODUCTION

The labor prediction is the most difficult and important problem in both normal and premature pregnancies. The difficulties aroused at the labor onset will lead to the risk of infants and mothers. A better understanding of parturition process could help in reducing these difficulties. The myometrium transition from non-labor to labor state can be observed by the uterine magnetomyography signals. The SARA (SQUID Array for Reproductive Assessment) system used to record uterine contractions during the labor [1]. The substantial increase in myometrium physiological activity can be observed 48 hours prior to the active labor [2]. The Hilbert-Wavelet duo can be used to find out the parameters that can extract and characterize the uterine contractions [3-7].

The literature available for uterine MMG signals had the focus on quantifying the uterine contractions. The works on labor diagnosis employed uterine EMG (electro myography)/EHG (electrohysterography) signals in most of the cases. The prediction of labor achieved through the classification of term and preterm records in [8]. The wavelet transform features used to train the artificial neural network for classifying EHG signals [9]. Diab et al. [10] employed unsupervised classification method for detecting preterm deliveries. The advanced neural network used by Fergus et al. [11] is better for classifying the TPEHGDB database comprises of term and preterm records. Moreover, the wavelet features of EMG/EHG signals used for labor prediction unlike the wavelet features of MMG signals [3-7]. In our previous studies, the features used for classification purpose only [12-15].

The classifier performance improved with the relevant features that are selected from a pool of features. Feature selection is used as a preprocessing step prior to the actual classification. The main aim of it is to remove feature redundancy and identify features relevance. Mutual information (statistical dependency) between features and features and their actual classes used for feature selection [16]. Genetic algorithm is one such type of algorithm (evolutionary) for feature selection [17]. The discrete wavelet transform (for feature extraction) and genetic algorithm (for feature selection) are adapted for this study on MMG signals.

The Physionet mmgdb database is a publicly available database from which twenty four signals used for this work. First the signals were divided into two groups depends on their time prior to delivery (labor and pregnancy). Six level decomposition performed on these signals using discrete wavelet transform. The transform coefficients further used to emulate the features i.e. variance, standard deviation, waveform length, energy per waveform length and entropy from each level. Different mother wavelets (sym5, sym8, db4, db8 and coif5) used for the decomposition. The number of features for classification can be reduced by selecting only the relevant features. Feature section technique (genetic algorithm) used as the feature subset selection method that can be applied to choose the best subsets for labor assessment. The selected features then applied to four different classifiers for classifying labor and pregnancy signals.

II. MATERIALS AND METHODS

2.1 Data Acquisition and Preprocessing

Physionet mmgdb database is a publicly available database for uterine magnetomyography signals [18]. The signals were recorded using SARA (SQUID Array for Reproductive Assessment) system at University of Arkansas for Medical Sciences, Little Rock, USA [7]. The MMG records digitized (250Hz) and sampled with 32 Hz frequency. Band pass filtering (0.1-1 Hz) applied and the signals are further processed with notch filter (0.25-0.35 Hz). These are mandatory for removing maternal and fetal cardiac signals and suppressing maternal breathing from the actual signals. Each record lasts between 10-20 minutes and contains multichannel signals. Based on the assumption that the uterine activity increases 48 hours prior to the delivery [2], the twenty four records divided into two groups [19] (before 48 hours- labor and after 48 hours- pregnancy).

2.2 Feature Extraction

The main aim of feature extraction step is to process the raw signals for getting informative parameters which are used at classification step instead of the original data. This is the mandatory and decisive process in medical diagnosis as the parameters can facilitate human interpretations at emergency situation. Wavelet transforms can analyze the data in the time - frequency domain. Discrete Wavelets Transform (DWT) is defined as the filter banks that can process the bio-medical signals for numerous applications [20]. The MMG signal filtered (low-pass then high-pass) to get low frequency components (approximate coefficients) and high frequency components (detailed coefficients). The approximate coefficients further processed (up to required level) as detailed and approximation coefficients. The decomposition using DWT up to six levels illustrated in figure 1. The i^{th} level approximate and detailed coefficients are labelled as cAi and cDi. Most of the studies on MMG detect uterine contractions (locations, duration and their peak values) by applying DWT. In this work, DWT applied to each channel to decompose the signal up to six levels. The DWT coefficients further used to emulate the features i.e. variance, standard deviation, waveform length, energy per waveform length and entropy from each level. We chose the wavelets for this research have the history in classifying physiological signals [15].

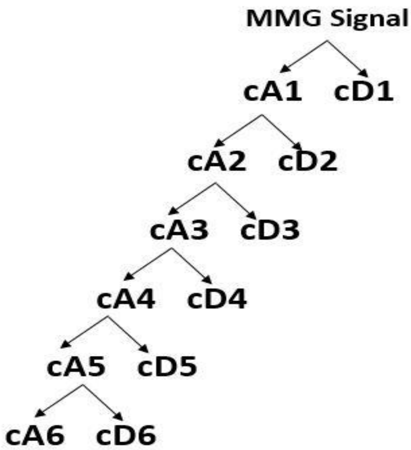

Figure 1. Discrete wavelet transform tree with six decomposition levels

2.3 Genetic Algorithm

The Genetic Algorithm (GA) is the most efficient method of feature selection based on the Darwin theory *survival of the fittest* [17]. It imitates the natural evolution by modelling a dynamic population of solutions for a given problem. The members of the population are called as chromosomes, used to encode the features. Each chromosome leads to a model that can be built by using these encoded features. The training data is used to quantify the criterion function which is served as the fitness function. During the evolution process, crossover and mutation operations are performed on the chromosomes. The algorithm allows survival and reproduction of fittest chromosomes, so that the criterion function minimized/maximized efficiently in next generations. This process stops when the maximum number of generations are reached or the desired fitness value is achieved.

The feature selection using GA influenced by many factors. The crossover and mutation parameters should be chosen carefully to prevent early convergence to homogeneous population occupying a local minimum. The success of GA also depend on the choice of initial population. So care should be taken while choosing the crossover and mutation probabilities that will be suitable for the specific application. Mutual Information between features and labels measures whether the features are dependent on their class labels or not. The mutual information as defined in [16] used as fitness function for the present work.

2.4 Classification

The Naïve-Bayes (NB) used in different applications (medical diagnosis, spam mail filtering and weather prediction) and is based on Bayes theorem. KNN classifier with Mahalanobis distance metric (k=5) used in the present work. Support vector machine (SVM) classifier with RBF kernel in our present work. The artificial neural networks (ANNs) used to predict the pre-term labor by classifying the uterine contractions. The weights and bias values of our neural network are update by using scaled conjugate gradient backpropagation function. Supervised classification employed in this work and the feature space is manually divided as 2/3 of feature space (16 recordings) for train set and the remaining for test set (8 recordings). To predict term delivery, the classifiers capability was tested by the classification of labor and pregnant signals.

III. RESULTS AND DISCUSSION

The performance metrics for the validation are accuracy, precision and false positive rate. These are derived from the true positives, true negatives, false positives and false negatives as illustrated in Table 1.

From the Table 1, Accuracy can be defined as the proportion of both labor and pregnancy signals that are correctly classified.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} * 100\% \qquad (1)$$

Precision gives the ratio of labor signals that are identified from both types of signals.

$$Precision = \frac{TP}{TP + FP} * 100\% \qquad (2)$$

False positive rate (FPrate) measures the misclassified labor signals to the total number of signals that are not predicted as belongs to labor class.

$$FPrate = \frac{FP}{TN + FP} * 100\% \qquad (3)$$

Table 1. Confusion matrix of the labor and pregnancy classes

		Actual Class: Labor	Actual Class: Pregnancy
Predicted Class	Labor	TP True Positive	FP False Positive
	Pregnancy	FN False Negative	TN True Negative

The features that are extracted using different wavelets db4, db8, sym5, sym8 and coif5 are selected using the GA technique. These features are fed to the classifiers Naïve-Bayes, ANN, KNN and SVM classifiers and the classification results are recorded in the table 2.

Table 2. The classification results (in %) for the features selected using Genetic Algorithm

Wavelet	Classifier	Accuracy	Precision	FPrate
Sym8	ANN	56.2130	57.6061	42.3939
	NB	55.3677	55.8049	44.1951
	KNN	89.1801	89.1869	10.8131
	SVM	94.3364	94.9088	10.1824
Sym5	ANN	49.5351	49.5312	50.4688
	NB	49.5351	49.4886	50.5114
	KNN	90.6171	90.6224	09.3776
	SVM	92.8149	93.7130	12.5740
Db8	ANN	53.5080	53.5265	46.4735
	NB	54.0997	54.2267	45.7733
	KNN	88.9265	89.1160	10.8840
	SVM	89.2646	91.1560	17.6880
Db4	ANN	56.7202	57.0572	42.9428
	NB	50.4649	50.4975	49.5025
	KNN	87.7430	87.9644	12.0356
	SVM	94.5055	95.0457	09.9085
Coif5	ANN	58.1572	58.3875	41.6125
	NB	53.5926	54.2229	45.7771
	KNN	91.0397	91.0476	08.9524
	SVM	95.9425	96.2441	07.5117

It can be observed from the table 2, the SVM classifier on GA selected features performed better compared to other classifiers. The SVM has the highest and lowest accuracy values for coif5 (95.9425%) and db8 features (89.2646%) respectively. The highest discrimination accuracy achieved again for coif5 features i.e. 91.0397% and lowest achieved for db4 features (87.7430%) in case of KNN classifier. The highest accuracies for ANN and Naïve-Bayes (NB) are very low (58.1572% for coif5 and 55.3677% for sym8 respectively) which resembles their inability in classifying the labor and pregnancy signals. The precision of SVM classifier is most among all classifiers similar to accuracies. The higher precision for SVM classifier observed for coif5 wavelet (96.2441%) closely followed by db4 wavelet (95.0457%). The highest and lowest precision values achieved for KNN classifier in case of coif5 and db4 wavelets with 91.0476% and 87.9644% respectively. The lower FP rates achieved by SVM classifier across all wavelets except db8 and sym5 features. Here the KNN classifier outperformed SVM classifier well with smaller FPrates.

Table 3 depicts the classifiers accuracies for Naïve-Bayes, ANN, KNN and SVM classifiers compared with two techniques. Columns represents the direct classification of features (DWT) [15], and the technique employed for features selection GA (DWT+GA). All the accuracy values are compared for different wavelets db4, db8, sym5, sym8 and coif5.

For sym8 wavelet, the accuracy increased from 80% with direct classification to 89.1801% using GA technique in case of KNN classifier. The SVM achieved direct classification accuracy of 84.2857% and improved with DWT+GA technique to 94.3364%. In case of sym5 wavelet, the accuracy increased using the DWT+GA selected features for KNN (72.5% to 90.6171%) and SVM classifiers (85% to 92.8149%). When the features are extracted using db4 wavelet, the KNN and SVM classifiers performance improved with the feature selection technique. The performance of classifiers improved for db8 and coif5 features with the DWT+GA technique. Moreover, the SVM classifier wins the race with 95.9425% for coif5 features selected using GA. Since the conditional independence does not hold for MMG signals, the NB and ANN classifiers performance did not improved with the selected features also.

Table 3. The classification accuracy (%) for the features classified directly and selected using genetic algorithm

Wavelet	Classifier	DWT	DWT+ GA
Sym8	ANN	54.2857	56.2130
	NB	55.7143	55.3677
	KNN	80.0000	89.1801
	SVM	84.2857	94.3364
Sym5	ANN	42.1429	49.5351
	NB	48.5714	49.5351
	KNN	72.5000	90.6171
	SVM	85.0000	92.8149
Db8	ANN	46.0714	53.5080
	NB	53.5714	54.0997
	KNN	71.4286	88.9265
	SVM	77.1429	89.2646
Db4	ANN	46.4286	56.7202
	NB	51.4286	50.4649
	KNN	74.6429	87.7430
	SVM	78.5714	94.5055
Coif5	ANN	52.8571	58.1572
	NB	52.8571	53.5926
	KNN	76.7857	91.0397
	SVM	82.1429	95.9425

The KNN classifier shows greater improvement (13.6% on average) than its counterpart (here SVM with 12% on average) in accuracy for DWT+GA feature selection technique. The SVM classifier has the highest discrimination accuracy (95.9425%) when the coif5 features are selected using the GA technique closely followed by db4 and sym8 features. To conclude, the GA is one suitable feature selection technique for the feature selection.

The selection of mother wavelet plays an important role for the feature extraction stage. The classifier performance improved by the suitable wavelet and illustrated in the figure 2. The line chart in figure 2 resembles the SVM classifier accuracy for different wavelets against DWT and DWT+GA techniques. The direct classification of sym5 features (85%) had the higher accuracy while the DWT+GA technique had that for coif5 features (95.9425%). It can be noted from the figure 2 that the SVM classifier has the higher accuracy of 85% for the sym5 wavelet in DWT method. The improvement in accuracy for sym8 and db8 features is approximately equal while the db4 performance improved a lot with the feature selection technique (17.4%). The feature selection with GA helped in improving the classifier accuracy, as evident from the figure and the coif5 features had the highest accuracy of 95.9425%.

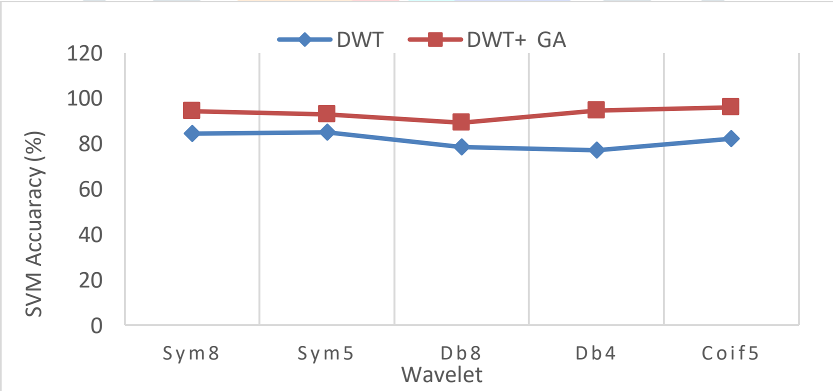

Figure 2. Accuracy values of SVM classifier for different wavelets

The impact of feature selection technique on classifier performance is illustrated in figure 3.

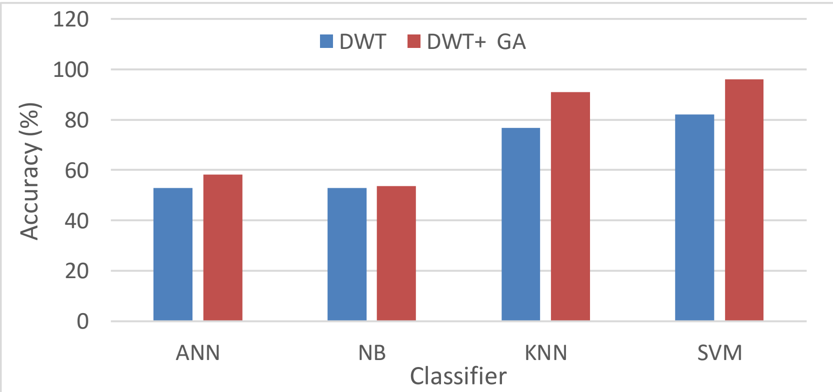

Figure 3. The comparision of classifier accuracies for DWT and DWT+GA techniques

Figure 3 represents the classifiers accuracy for coif5 wavelet. It can be observed from the figure that classifier performance imporved with the GA technique. The improvement is very less in case of NB classifier. The KNN and SVM classifier performance improved at the same rate using the genetic algorithm. The classification of term/preterm EHG signals yields reasonable accuracies for [8-10], though these works studied on uterine contractions only. Fergus et al. [11] examined the entire EHG signal and extract the wavelet features. Hassan et al. [21] analysed the term delivery records, similar to the present work to differentiate term labor and term non labor groups. The whole MMG signal considered in this work to extract the DWT features. The relevant features are selected by using genetic algorithm to discriminate the labor and pregnancy groups. The accuracy achieved for DWT+GA technique is 95.9425% using coif5 features by SVM classifier. The experimental results may provide a fundamental method for classifying the pregnancy and labor signals and the labor prediction.

IV. CONCLUSION

The work is aimed to reduce the features that extracted from the Physionet (mmgdb database) MMG signals. Six level DWT and genetic algorithm used for this purpose. The classification results of direct method are compared with the ones obtained by using the selected features of genetic algorithm. The conclusion drawn from the results is that KNN classifier shows highest increment (13.6% on average) in classification performance where the SVM classifier has the highest discrimination accuracy (95.9425%) when the features are selected using the GA technique. In future, the authors wanted to test the present methods on a larger MMG database. Further, the features are optimized using swarm intelligent techniques. It will enable the authors to use most robust and relevant cross-validation techniques than the present simple technique. In future, empirical mode decomposition features and the features related to the synchronization and propagation of uterine contractile activity, will be included.

REFERENCES

[1] Eswaran H, Preissl H, Wilson JD, Murphy P, Robinson SE, Lowery CL. First magnetomyographic recordings of uterine activity with spatial temporal information with a 151-channel sensor array. American Journal of Obstetrics & Gynecology, Volume 187, Issue 1, 145 – 151.

[2] Eswaran H, Preissl H, Wilson JD, Murphy P, Lowery CL. Prediction of labor in term and preterm pregnancies using non-invasive magnetomyographic recordings of uterine contractions. Am J of Obstet Gynecol. 2004, 190[6]:1598–1603.

[3] Eswaran H, Govindan RB, Furdea A, Murphy P, Lowery CL, Preissl HT. Extraction, Quantification and Characterization of Uterine Magnetomyographic Activity - A Proof of Concept Case Study. European journal of obstetrics, gynecology, and reproductive biology. 2009;144(Suppl 1):S96-100. doi:10.1016/j.ejogrb.2009.02.023.

[4] Furdea A, Eswaran H, Wilson JD, Preissl H, Lowery CL, Govindan RB. Magnetomyographic recording and identification of uterine contractions using Hilbert-wavelet transforms. Physiol Meas. 2009 Oct;30(10) 1051-1060. doi:10.1088/0967-3334/30/10/006. PMID: 19738317; PMCID: PMC2965828.

[5] Furdea A, Preissl H, Lowery CL, Eswaran H, Govindan RB. Conduction Velocity of the Uterine Contraction in Serial Magnetomyogram (MMG) Data: Event Based Simulation and Validation. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference. 2011;2011:6025-6028. doi:10.1109/IEMBS.2011.6091489.

[6] Govindan RB, Siegel E, Mckelvey S, Murphy P, Lowery CL, Eswaran H. Tracking the Changes in Synchrony of the Electrophysiological Activity as the Uterus Approaches Labor Using Magnetomyographic Technique. Reproductive Sciences. 2015; 22(5): 595 - 601. doi: 10.1177/ 1933719114556484.

[7] Escalona-Vargas D, Govindan RB, Furdea A, Murphy P, Lowery CL, Eswaran H. Characterizing the Propagation of Uterine Electrophysiological Signals Recorded with a Multi-Sensor Abdominal Array in Term Pregnancies. Young RC, ed. PLoS ONE. 2015;10(10):e0140894. doi:10.1371/journal.pone.0140894.

[8] Maner WL, Garfield RE, Identification of Human Term and Preterm Labor using Artificial Neural Networks on Uterine Electromyography Data, Ann Biomed Eng. 2007 Mar; 35(3): 465–473. Published online 2007 Jan 17. doi: 10.1007/s10439-006-9248-8.

[9] Marque, C. K., Terrien J, Rihana S, Germain G, Preterm labour detection by use of a biophysical marker: the uterine electrical activity, BMC Pregnancy Childbirth. 2007/06/19, 7 Suppl 1, p. S5. doi: 10.1186/1471-2393-7-s1-s5.

[10] Diab, M. O., Marque, C. and Khalil, M. (2009) 'An unsupervised classification method of uterine electromyography signals: Classification for detection of preterm deliveries', Journal of Obstetrics and Gynaecology Research, 35(1), pp. 9–19. doi: 10.1111/j.1447-0756.2008.00981.x.

[11] Fergus, P., Idowu, I.O., Hussain, A.J., & Dobbins, C. (2016). Advanced artificial neural network classification for detecting preterm births using EHG records. Neurocomputing, 188, 42-49. doi: 10.1016/j.neucom.2015.01.107.

[12] T. A. Babu and P. R. Kumar, "Characterization and classification of uterine magnetomyography signals using KNN classifier," 2018 Conference on Signal Processing and Communication Engineering Systems (SPACES), Vijayawada, 2018, pp. 163-166. doi: 10.1109/SPACES.2018.8316337.

[13] T Ananda Babu, P Rajesh Kumar, Features Extraction and Classification of Uterine Magnetomyography Signals, International Journal of Current Engineering and Scientific Research, Volume 5, Issue 3-2018.

[14] T Ananda Babu, P Rajesh Kumar, Prediction of Term Labor using Wavelet Analysis of Uterine Magnetomyography Signals, Proceedings of International Conference on Computational Intelligence and Data Engineering, Lecture Notes on Data Engineering and Communication Technologies 28, pp 29-37, https://doi.org/10.1007/978-981-13-6459-4_4.

[15] Ananda Babu T, & P Rajesh Kumar, Dr., Comparison of different feature extraction methods for the analysis of uterine magnetomyography signals to predict term labor. International Journal of Engineering & Technology,7(3.29), 1-7. doi: http://dx.doi.org/10.14419/ijet.v7i3.29. 18449.

[16] Pohjajalainena J, Rasanen O, Kadioglub S, Feature Selection Methods and Their Combinations in High-Dimensional Classification of Speaker Likability, Intelligibility and Personality Traits, Computer Speech & Language · November 2013, DOI: 10.1016/j.csl.2013.11.004.

[17] Bazarghan, M., Amandi, R., Jaberi, Y., Abedi, M., Shahbazi, A., & Karami, E. (2012). Automatic ECG Beat Arrhythmia Detection. CoRR, abs/1209.0167. arXiv:1209.0167v3 [cs.NE] 7 Nov 2012.

[18] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation 101(23):e215- e220[Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/ full/101/23/e215]; 2000 (June 13).

[19] T Ananda Babu, P Rajesh Kumar, Separation of Uterine MMG Records of Term Pregnancies using Linear and Nonlinear Features, International Journal of Pure and Applied Mathematics, special issue, Volume 118 No. 9, 2018, 697-702.

[20] M. Akay, Wavelet applications in medicine, *IEEE Spectrum*, vol. 34, no. 5, pp. 50–56, May 1997.

[21] Hassan M, Terrien J, Muszynski C, Alxanderson A, Marque C, Karlsson B., Better pregnancy monitoring using nonlinear correlation analysis of external uterine electromyography, IEEE Trans Biomed Eng. 2013 Apr;60(4):1160-6. doi: 10.1109/TBME.2012.2229279.