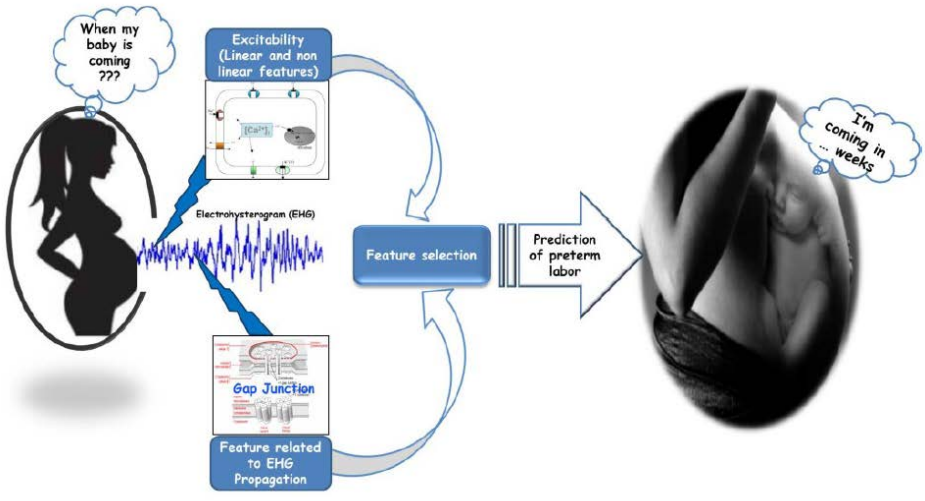
Selection of EHG parameter characteristics for the classification of uterine contractions

Dima Alamedine

Par **Dima ALAMEDINE**

Selection of EHG parameter characteristics for the classification of uterine contractions

Thèse présentée en cotutelle pour l'obtention du grade de Docteur de l'UTC



Soutenue le 21 juillet 2015
Spécialité : Biomedical engineering

D2201

Université Libanaise
École Doctorale
Sciences et Technologies
Doyen

utc
Université de Technolo
Compiègne

COTUTELLE THESIS

To obtain the degree of Doctor issued by

Sorbonne University, Université de technologie de Compiègne
Doctoral School « Sciences pour l'Ingénieur »

and

Lebanese University
Doctoral School « Sciences et Technologie »

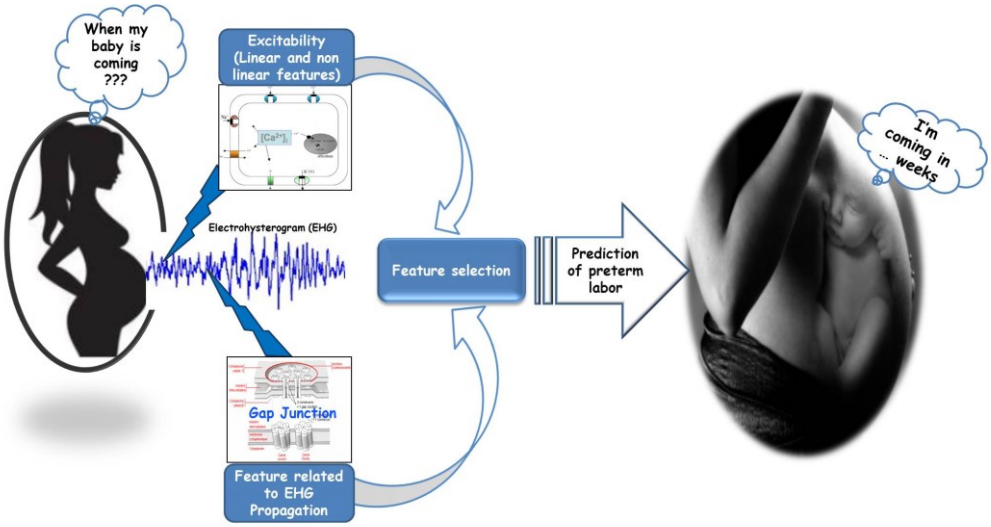
Specialty: Biomedical engineering

Presented and publicly defended by

ALAMEDINE Dima

21-7-2015

Selection of EHG parameter characteristics for the classification of uterine contractions



Jury Member

Régine LE BOUQUIN JEANNES	Prof. , Université de Rennes 1	**Reviewer**
Imad H. ELHAJJ	Associate Prof., American university of beirut	**Reviewer**
Oussama BAZZI	Prof., Lebanese University	**Examiner**
Sofiane BOUDAOUD	Associate Prof., Université de Technologie de Compiègne	**Examiner**
Catherine MARQUE	Prof., Université de Technologie de Compiègne	**Supervisor**
Mohamad KHALIL	Prof., Lebanese University	**Supervisor**

To…

Résumé Français

Titre:

Sélection de paramètres caractéristiques des EHG pour la classification des contractions utérines

Contenu :

Certaines femmes souffrent de complications de la grossesse qui peuvent aboutir à un accouchement prématuré, avant 37 semaines de gestation. Selon l'Organisation mondiale de la Santé (WHO), le taux de mortalité périnatale est généralement autour de 7 pour 1000 naissances dans la partie la plus développée du monde [1]. Une des causes principales de mortalité et de morbidité néonatales est la naissance d'enfants avant terme. Ces enfants nés avant terme présentent un risque élevé de mortalité ainsi que des problèmes de santé et de développement [2]. Un objectif principal de la surveillance de la grossesse est de maintenir le bien-être de la mère et du fœtus et de garder ce dernier in utero aussi longtemps que nécessaire pour la naissance d'un enfant en bonne santé. Par conséquent, la détection précoce d'un accouchement prématuré (AP) est importante pour sa prévention. A cet effet, de bons indicateurs de travail prématuré sont nécessaires.

Pour maintenir le fœtus in utero aussi longtemps que nécessaire, une surveillance de la contractilité utérine est essentielle pour distinguer les contractions normales de la grossesse, qui sont inefficaces de celles qui sont efficaces et pourraient entraîner la dilatation du col de l'utérus et causer une naissance prématurée.

Malgré l'augmentation des connaissances et de la compréhension des phénomènes impliqués dans le début du travail prématuré, les méthodes actuellement utilisées en obstétrique ne sont pas assez précises pour une détection précoce de menaces d'accouchement prématuré. La mesure de pression intra-utérine, seule méthode directe précise pour mesurer la force des contractions utérines, est invasive et ne peut clairement pas être utilisée pendant la grossesse. La tocographie externe, non invasive, est la méthode la plus largement utilisée pour la surveillance des contractions utérines pendant la grossesse. Cependant, elle ne permet pas de caractériser l'efficacité des contractions. Elle ne permet que de détecter le nombre de contractions pendant un intervalle de temps donné. Il a été démontré que ce paramètre n'est pas un bon indicateur pour

prédire un accouchement prématuré. Des tests biologiques, tels que la fibronectine, ont été cliniquement utilisés pour le pronostic des accouchements prématurés, bien qu'ils possèdent une valeur prédictive faible [3]. Même la mesure de la dilatation cervicale n'est pas un indicateur fiable d'une menace d'accouchement prématuré.

Nous avons besoin d'une méthode non-invasive et plus fiable pour la détection précoce et la prévention des menaces d'accouchement prématuré, car ce problème est clairement un domaine d'intérêt en santé publique. Ce diagnostic précoce permettait une administration plus rapide d'agents tocolytiques (inhibiteurs de la contractilité utérine) et un maintien du foetus in utero, ce qui réduirait la mortalité et la morbidité périnatales. Une des méthodes prometteuses pour surveiller l'efficacité des contractions utérines pendant la grossesse est l'analyse de l'activité électrique du muscle utérin, l'électrohystérogramme (EHG), qui a commencé dans les années 1950 et a été développé dans les années 1980. L'EHG est le signal recueilli par des électrodes positionnées sur l'abdomen des femmes enceintes. Il représente l'activité électrique associée à la contraction mécanique de l'utérus [4]. De nombreuses études ont prouvé qu'il est représentatif de l'activité électrique utérine recueillie en interne sur l'utérus [5,6]. De multiples travaux menés depuis plus de 15 ans, concluent qu'il est possible d'utiliser l'EHG pour détecter une menace d'accouchement prématuré. L'efficacité des contractions utérines est liée à une augmentation de deux phénomènes physiologiques : l'excitabilité cellulaire et la synchronisation de l'utérus (propagation de l'activité électrique) [5,7]. Ces phénomènes peuvent être analysés au moyen de l'EHG grâce à des analyses univariée (analyse d'un seul signal à la fois pour l'excitabilité) et bivariée (couplage de deux signaux pour la propagation).

Plusieurs outils de traitement du signal EHG, nouvellement développés, permettent l'analyse de l'excitabilité et de la propagation de l'activité électrique utérine (paramètres fréquentiels, analyse de complexité, propagation linéaire ou nonlinéaire) pour trouver des informations spécifiques qui différencient les contractions de la grossesse de celles du travail. Un grand nombre de paramètres ont à ce jour été extraits du signal EHG par de nombreux chercheurs, en utilisant différents protocoles d'enregistrement et différentes populations de femmes enceintes. En ce qui concerne la classification des signaux, la complexité des calculs requis pour le diagnostic augmente avec le nombre de paramètres mis en jeu. La réduction de la dimension de paramètres grâce à l'élimination des paramètres non pertinents et bruités est très importante en reconnaissance des

formes. De plus, les deux facteurs, excitabilité utérine et propagation de l'activité, sont tous les deux importants car ils doivent évoluer tous les deux pour passer d'une contraction inefficace de grossesse (faible excitabilité, contraction locale) à une contraction efficace d'accouchement (excitabilité forte, propagation à tout l'utérus en un court intervalle de temps). Des études antérieures, fondées uniquement sur l'étude de l'excitabilité, ont donné des résultats intéressants, mais pas suffisamment fiables pour la détection précoce de l'accouchement prématuré.

Par conséquent, la première contribution de notre étude sera de combiner les deux types d'informations (excitabilité et propagation) par des approches monovariée et bivariée simultanées, pour différencier deux types de contractions utérines : les contractions de grossesse et celles d'accouchement.

La seconde contribution sera de tester sur une même population de femmes enceintes (grossesse, travail), quels outils de traitement du signal, récemment développés pour l'analyse monovariée et bivariée de l'EHG, donnent la meilleure discrimination entre les contractions de grossesse et d'accouchement. A cet effet plusieurs méthodes de sélection de paramètres extraits de l'EHG seront analysées.

Cette sélection de paramètres se fera à partir d'une base de données de signaux recueillis suivant une méthode normalisée, sur des femmes enceintes dans différentes situations physiologiques (grossesse normale, accouchement), grâce à un système multi-électrodes permettant d'enregistrer 16 EHG monopolaires simultanés. Ces 16 signaux seront ensuite analysés par les approches monovariée et bivariée, pour extraire les paramètres représentatifs de l'excitabilité et la propagation de l'activité électrique utérine. Beaucoup d'études ont utilisé seulement des signaux issus des électrodes positionnées sur l'axe vertical médian de l'abdomen [8] pour l'analyse monovariée (une seule voie). D'autres études ont travaillé sur les caractéristiques liées à la propagation de l'EHG grâce au couplage entre toutes les voies (analyse bivariée). Des études précédentes ont ainsi tenté d'appliquer une classification multivoie [9-14]. Certaines études se sont intéressées à une analyse multivariée, afin d'extraire les informations d'excitabilité et de propagation de toutes les voies et combinaisons de voies disponibles, ce qui conduit à une très grande dimension de recherche. Par conséquent, une troisième contribution de ce travail portera sur la sélection des voies et des combinaisons de voies les plus pertinentes dans une optique de classification grossesse/accouchement.

En outre, afin d'augmenter le rapport signal/bruit des EHG, toutes les études précédentes ont porté sur des signaux bipolaires, obtenus par différence entre deux signaux recueillis par deux électrodes plus ou moins proches. Si cette différenciation se justifie pour une approche monovariée, le fait de différencier les signaux diminue cependant la résolution spatiale et conduit à un biais pour l'étude bivariée de la propagation entre deux voies adjacentes. Des travaux récents ont permis de développer une méthode de débruitage des EHG monopolaires, afin d'obtenir sans différenciation un rapport signal/bruit suffisant pour envisager le traitement de ces signaux [15]. Ces outils s'appuient sur une combinaison de CCA (Combination of Canonical component Analysis) et d'EMD (Empirical Mode Decomposition). De ce fait, en plus d'une approche classique basée sur l'étude des signaux bipolaires, nous proposerons dans ce travail d'appliquer les approches monovariée et bivariée, ainsi que la sélection de paramètres et de voies, sur des signaux monopolaires débruités.

Ce manuscrit est donc organisé comme suit:

- **Chapitre 1:** contient toutes les informations essentielles pour la bonne compréhension de l'anatomie et de la physiologie de l'activité utérine nécessaires à ce travail. Nous définirons l'accouchement prématuré et les problèmes qui y sont reliés, et nous présenterons certaines méthodes utilisées en pratique obstétricale courante pour détecter l'accouchement prématuré. Ensuite, nous présenterons une bibliographie des différentes études d'excitabilité et de la propagation faites à partir de signaux EHG (approches monovariée et bivariée). À la fin de ce chapitre, nous décrirons le système multi-électrodes pour l'enregistrement d'EHG utilisé dans notre travail, ainsi que les bases de données utilisées.

- **Chapitre 2:** présente le travail effectué sur la sélection de paramètres de l'EHG. Nous décrirons dans ce chapitre l'ensemble de paramètres choisis dans la littérature pour l'analyse monovariée et bivariée de l'EHG [16-26]. Puis nous présentons les différentes méthodes de sélection de paramètres, qui sont décomposées en deux types « filter » et « wrapper ». Dans la première partie de ce chapitre, nous présenterons un algorithme de sélection proposé pour les paramètres calculés sur l'EHG original et sur différentes bandes de fréquence, en utilisant la technique de sélection de paramètre, de type filter,

nommé Jeffrey Divergence (JD). Cette méthode de sélection, basée sur la mesure de similarité ou dissimilarité des histogrammes obtenus, pour un paramètre donné, pour les deux classes (grossesse et travail).

Dans la deuxième partie, nous testerons plusieurs méthodes de sélection de paramètres afin de sélectionner les paramètres les plus pertinents pour discriminer les contractions de grossesse de celles d'accouchement. Quatre méthodes de type de « filter » sont utilisées : Jeffrey divergence (JD) [27], « F-score » [28], «Relieff » [29], «mutual information based on clustering» (MI)[30] et sept méthodes de type « wrapper » : sélection séquentielle croissante « sequential Forward Selection» (SFS) [31], sélection séquentielle arrière « Sequential backward Selection» (SBS)[32], « Plus-l minus-r selection» (LRS)[33], recherche bidirectionnelle «Bidirectional search» (BDS) [33], «Sequential Forward Floating sequential» (SFFS) [34], algorithmes génétiques «Genetic Algorithm» (GA) [35] et «Binary Particle swarm optimization» (BPSO)[36]. Nous utiliserons le classificateur KNN pour les méthodes de type « wrapper », et deux méthodes de répartition de données pour l'apprentissage de ce classifieur: « Holdout » et « KFOLD». Après la partie sélection de paramètres, nous essayerons de valider chaque sous-ensemble de paramètres sélectionné, en calculant pour chacun le pourcentage de classification correcte obtenu sur une population d'EHG différente de celle utilisée pour la sélection.

Dans ce chapitre, la première partie sera appliquée uniquement sur des signaux bipolaires. Les autres études seront appliquées sur les signaux bipolaires et monopolaires.

Les résultats obtenus montrent que l'utilisation des méthodes basées sur la sélection de paramètres permet de mettre en évidence un groupe de paramètres, combinant analyse monovariée et bivariée, et qui démontre sa capacité à discriminer les contractions de grossesses de celles d'accouchement, avec de meilleures performances pour les signaux EHG bipolaires.

- **Chapitre 3:** dans ce chapitre, nous présenterons tout d'abord le travail effectué sur la sélection de voies (approche monovariée) en utilisant les paramètres linéaires et non linéaires. Nous testerons deux méthodes de type « filter » (F-score et relieff) pour sélectionner les voies appropriées, puis deux méthodes de type « wrapper » (Genetic algorithm et binary particle swarm optimization) avec le classifieur KNN et deux

méthodes de répartition de données (Holdout et KFOLD) pour sélectionner les meilleurs paramètres à partir des voies sélectionnées. En outre, une partie de validation sera effectuée en calculant le pourcentage de classification correcte obtenu (en utilisant une autre population d'EHG) à partir des sous-ensembles de paramètres sélectionnés par sélection de voie suivie de la sélection de paramètres.

La dernière partie de ce chapitre présentera les résultats de la sélection des combinaisons de voies (approche bivariée) en utilisant les paramètres liés à la propagation de l'EHG. Nous utilisons ici la même procédure que celle utilisée dans l'approche monovariée, pour la sélection de voies, suivie de la sélection de paramètres.

Les résultats obtenus montrent que, pour l'analyse monovariée, les canaux bipolaires et monopolaires qui offrent une meilleure capacité de discrimination entre les contractions de grossesse et de travail sont ceux positionnés sur l'axe vertical médian de l'abdomen de la femme enceinte. L'utilisation des signaux d'EHG bipolaires pour l'analyse monovariée donne de meilleurs résultats que celle des EHG monopolaires, particulièrement en utilisant l'ensemble des paramètres linéaires et non linéaires extraits des voies Vb7, Vb8 et Vb9. De plus, l'approche bivariée, sélection des combinaisons de voies suivie de sélection de paramètre, montre que l'utilisation des EHG monopolaires pour l'étude de la propagation améliore les résultats de classification entre les contractions de grossesse de celles de travail et permet d'atteindre des résultats similaires à ceux obtenus avec les signaux bipolaires.

Les résultats obtenus dans cette thèse nous ont permis d'écrire 3 articles de revues publiés (2 revues internationales et 1 nationale), 2 conférences internationales, 4 conférences nationales.

References

[1] Neonatal and perinatal mortality: country, regional and global estimates; ReportWHO 2006. http://whqlibdoc.who.int/publications/2006/9241563206_eng.pdf

[2] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero, "Epidemiology and causes of preterm birth," *The lancet*, vol. 371, no. 9606, pp. 75–84, 2008.

[3] J.D. Iams, "Prediction and early detection of preterm labor," *Obstet. Gynecol*., vol. 101, no. 2, pp. 402-12, Feb. 2003.

[4] H. Alvarez and R. Caldeyro, “Contractility of the human uterus recorded by new methods,” *Surgery, Gynecology & Obstetrics*, vol. 91, no. 1, pp. 1–13, 1950.

[5] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchene, "Uterine electromyography: a critical review," *Am J Obstet Gynecol,* vol. 169, pp. 1636-1653, Dec 1993.

[6] S. Mansour, D. Devedeux, G. Germain, C. Marque, and P. J. Duchene, “Uterine EMG spectral analysis and relationship to mechanical activity in pregnant monkeys,” *Med. Biol. Eng. Comput.*, vol. 34, no. 2, pp. 115–121, Mar. 1996.

[7] R. E. Garfield and W. L. Maner, "Physiology and electrical activity of uterine contractions," *Seminars in Cell & Developmental Biology,* vol. 18, pp. 289-295, 2007.

[8] C. K. Marque, J. Terrien, S. Rihana, and G. Germain, “Preterm labour detection by use of a biophysical marker: the uterine electrical activity,” *BMC Pregnancy Childbirth*, vol. 7, no. Suppl 1, p. S5, Jun. 2007.

[9] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, “A multisensor data fusion approach for improving the classification accuracy of uterine EMG signals,” *in Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 93–96.

[10] B. Moslem, M. O. Diab, C. Marque, and M. Khalil, “Classification of multichannel uterine EMG signals,” in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*,pp. 2602–2605, 2011.

[11] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, “Classification of multichannel uterine EMG signals by using unsupervised competitive learning,” *in Signal Processing Systems (SiPS), 2011 IEEE Workshop on*, 2011, pp. 267–272.

[12] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, “Combining multiple support vector machines for boosting the classification accuracy of uterine EMG signals*,” in Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 631–634.

[13] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, “Classification of multichannel uterine EMG signals using a reduced number of channels,” *in Mechatronics and its Applications (ISMA), 2012 8th International Symposium on*, pp. 1–4, 2012.

[14] B. Moslem, M. Diab, M. Khalil, and C. Marque, “Combining data fusion with multiresolution analysis for improving the classification accuracy of uterine EMG signals,” *EURASIP J. Adv. Signal Process*., vol. 2012, no. 1, pp. 1–9, Aug. 2012.

[15] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, “Combination of Canonical Correlation Analysis and Empirical Mode Decomposition Applied to Denoising the Labor Electrohysterogram,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2441–2447, Sep. 2011.

[16] T. Jérémy, S. Thora, M. Catherine, K. Brynjar, and others, “Synchronization between EMG at different uterine locations investigated using time-frequency ridge reconstruction: comparison of pregnancy and labor contractions,” *EURASIP J. Adv. Signal Process*., vol. 2010, Article ID 242493, 2010.

[17] J. Sikora, A. Matonia, R. Czaba´nski, K. Horoba, J. Jezewski, and T. Kupka, “Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals,” *Archives of Perinatal Medicine*, vol. 17, no. 2, pp. 97–103, 2011.

[18] C. Marque, H. Leman, M. L. Voisine, J. Gondry, and P. Naepels, “Traitement de l’électromyogramme utérin pour la caract´erisation des contractions pendant la grossesse,” *RBMNews*, vol. 21, no. 9, pp. 200–211, Dec. 1999.

[19] G. Fele-ˇZorˇz, G. Kavˇsek, ˇZ. Novak-Antoliˇc, and F. Jager, “A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and preterm delivery groups,” *Medical and Biological Engineering and Computing*, vol. 46, no. 9, pp. 911–922, 2008.

[20] M. O. Diab, C. Marque, and M. A. Khalil, “Classification for uterine EMG signals: comparison between AR model and statistical classification method,” *International Journal of Computational Cognition*, vol. 5, no. 1, pp. 8–14, 2007.

[21] A. Diab, M. Hassan, C. Marque, and B. Karlsson, “Quantitative performance analysis of fourmethods of evaluating signal nonlinearity: application to uterine EMG signals,” *in Proceedings of the 34th Annual International IEEE EMBS Conference*, San Diego, Calif, USA, September 2012,pp. 1045–1048.

[22] B. Moslem, M. Khalil, M. O. Diab and C. Marque, “Detrended fluctuation analysis of uterine electromyography,” *in First Middle East Conference on Biomedical Engineering, MECBME11*, Sharjah, UAE, 2011.

[23] K. Ansari-Asl, F. Wendling, J. J. Bellanger, and L. Senhadji, “Comparison of two estimators of time-frequency interdependencies between nonstationary signals: application to epileptic EEG,” *in 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004. IEMBS ’04, 2004, vol. 1, pp. 263–266.

[24] M. Hassan, A. Alexandersson, J. Terrien, C. Muszynski, C. Marque, and B. Karlsson, “Better pregnancy monitoring using nonlinear propagation analysis of external uterine electromyography”, *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 4, 2013,pp. 1160–1166.

[25] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, “Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings,” *Presented at the 32 nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Buenos Aires, Espagnol, Aug.2010.

[26] A.DIAB, “Study of The Nonlinear Properties And Propagation Characteristics Of The Uterine Electrical Activity During Pregnancy And Labor”, Ph.D. dissertation, Thèse de l’université de Technologie de Compiègne, 2014

[27] Y. Ma, X. Gu, and Y. Wang, “Histogram similarity measure using variable bin size distance,” *Comput. Vis. Image Underst*., vol. 114, no. 8, pp. 981–989, Aug. 2010.

[28] S. Ding, “Feature Selection Based F-Score and ACO Algorithm in Support Vector Machine,” *in Second International Symposium on Knowledge Acquisition and Modeling*, 2009. KAM ’09, 2009, vol. 1, pp. 19–23.

[29] I. Kononenko, “Estimating attributes: analysis and extensions of RELIEF,” *in Machine Learning: ECML-94*, 1994, pp. 171–182.

[30] H. Liu, Y. Mo, J. Wang, and J. Zhao, “A new feature selection method based on clustering,” *in 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2011, vol. 2, pp. 965–969.

[31] L. Ladha and T. Deepa, “Feature selection methods and algorithms,” *Int. J. Comput. Sci. Eng*., vol. 3, no. 5, pp. 1787–1797, 2011.

[32] A. W. Whitney, “A direct method of nonparametric measurement selection,” *Comput. IEEE Trans*. On, vol. 100, no. 9, pp. 1100–1103, 1971.

[33] A. R. Webb, Statistical Pattern Recognition. *John Wiley & Sons*, 2003.

[34] P. Pudil, J. Novovičová, and J. Kittler, “Floating search methods in feature selection,” *Pattern Recognit. Lett*., vol. 15, no. 11, pp. 1119–1125, Nov. 1994.

[35] M. Melanie, “An introduction to genetic algorithms,” *Camb. Mass. Lond. Engl. Fifth Print.*, vol. 3, 1999.

[36] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi, “Application of the PSOSVM model for recognition of control chart patterns,” *ISA Transactions*, vol. 49, no. 4, pp. 577–586, 2010.

Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of tables

Chapter 1

Chapter 2

Chapter 3

General Introduction

Some women suffer from complications of their pregnancy that may end with a preterm delivery, that is before 37 weeks of gestation. According to the World Health Organization (WHO), the perinatal mortality rate is typically around 7 per 1,000 births in the most developed part of the world [1]. A large part of the causes for mortality and morbidity is that children are born before their time. Children born preterm present high risk of mortality, as well as health and development problems [2]. A primary aim of pregnancy monitoring is to maintain the wellbeing of both mother and fetus and to keep the latter in utero as long as needed for a healthy birth. Therefore, the early detection of a preterm labor (PL) is important for its prevention and, for that purpose, good indicators of preterm labor are needed.

To maintain the fetus in utero as long as necessary, monitoring of uterine contractility is essential to differentiate the normal contractions of the pregnancy, which are ineffective, from those, effective, that could cause premature cervical dilation of uterus and preterm birth.

Despite increased knowledge and understanding of the phenomena involved in the onset of labor, the methods currently used in obstetrics are not precise enough for an early detection of preterm birth threats. The measurement of intrauterine pressure, the only direct method that permits presently to accurately measure the force of uterine contractions, is invasive and cannot clearly be used during pregnancy. Tocographie external, non-invasive, is the most widely used method for monitoring of uterine contractions during pregnancy. However, it does not permit to characterize the efficiency of contractions. It only permits to detect the number of contractions over a given time interval. It has been shown that this parameter is not a good predictor of premature labor. Biological tests, such as fibronectin, have been clinically used for the prognosis of premature births, although they have a low predictive value [3]. Even the measurement of cervical dilatation is not a reliable indicator of preterm labor. Indeed, a high proportion of women with cervical dilation during pregnancy will give birth at term, even without administration of tocolytic agents (inhibitors of uterine contractility).

We need a non-invasive and more reliable method for the early detection and prevention of preterm birth threats because this problem is clearly a field of interest for public health. This

earlier diagnosis would permit an earlier administration of tocolytics agents and therefore a longer maintain of the fetus in utero, with associated reduction of perinatal mortality and morbidity. One of the most promising methods for monitoring uterine activity began in the 1950s and was developed in the 1980s. It is based on the study of the electrical activity of the uterus (electrohysterogram, EHG) as recorded on the mother's abdomen [4]. The EHG is the signal recorded on the abdominal surface, which represents the electrical activity triggering the mechanical contraction of the myometrium. It has been demonstrated to be representative of the uterine electrical activity recorded internally [5, 6]. As the trigger of the mechanical contraction, its analysis is a promising method for accurate early recognition of preterm labor risk [4]. The contraction efficiency is related to an increase in two physiological phenomena: cellular excitability and spread of the electrical activity [5, 7] that could be monitored by means of the EHG.

Several processing tools of the EHG signal, recently developed, permit the analysis of excitability and of uterine electrical activity synchronization (frequency parameters, complexity analysis, linear and nonlinear propagation) in order to extract specific information that differentiates pregnancy and labor contractions. A large number of features have thus been extracted from the EHG signal by many different researchers, by using very different population and recording protocols. For a classification task, the complexity of calculations required for diagnostic purposes increases with the number of features in play. The reduction of feature dimensionality through the elimination of irrelevant and noisy features is thus very important in pattern recognition. Furthermore, most of previous studies used only a monovariate approach to study the excitability (one EHG lead processed at a time). They gave interesting results, but not reliable enough for an early detection of premature birth.

Therefore, the first contribution of our study is to combine the two approaches (monovariate and bivariate analyses) for a diagnosis based on the two kinds of information (excitability and propagation).

Then the second step will be to identify, which features developed for monovariate and bivaraite analyses of the EHG, permit the best discrimination between the two classes of signals: pregnancy and labor. For this purpose several methods for the selection of features extracted from the EHG will be analyzed and tested.

This feature selection will be made on a population from which multiple EHG have been recorded in a standardized way during pregnancy and labor. These recordings have been made by using a multi-electrode system providing the simultaneous recording of 16 monopolar EHG channels. These 16 signals will then be analyzed by processing the monovariate and bivaraite features chosen to extract the parameters representing the excitability and the propagation of the electrical activity of the uterus. Many studies used only the bipolar EHG signal corresponding to electrodes positioned in the median vertical axis of the woman's abdomen [8] for monovariate study. Other studies calculated the features related to EHG propagation by studying the coupling between all possible channels (bivariate analysis). Some studies tried to apply multichannel classification [9-14]. Using all the possible features extracted to characterize excitability and propagation, from all channels and possible combination of channels, lead to a very large dimension of search. Therefore, the third contribution of our work will be the selection of the most relevant channels and combinations channels for diagnosis purpose.

Finally, in order to increase the signal/noise ratio (SNR), all previous EHG studies have processed only bipolar signals. This bipolarization can be justified for the monovariate approach. But, for the bivariate approcah, bipolarization decreases the spatial resolution and leads to a bias when studying propagation between adjacent channels. Recently, a study was developed to filter Monopolar signals [15]. This method is based on the combination of canonical component analysis (CCA) and EMD (Empirical Mode Decomposition). Thus, in addition to the classic use of bipolar signals, we will try in our work to apply the monovariate and bivariate approaches on these denoised monopolar signals.

This manuscript is thus organized as follows:

- **Chapter 1:** contains all the essential information to understand the anatomical and physiological concepts of uterine activity required for this work. We define preterm labor and its problems and present some existing methods used in the current obstetrical practice to detect preterm labor. Then we will present a bibliography concerning the monovariate and bivariate approaches developed to characterize EHG in terms of excitability and propagation. And finally, we will describe the multichannel standardized recording protocol used in our work, as well as the different Databases.

- **Chapter 2:** presents the work done for the EHG feature selection. We will describe in this chapter the set of parameters selected from the literature, for monovariate or bivariate analysis of the EHG [16-26]. And then, we will present the different feature selection methods developed for data mining, which are decomposed into two types "Filter" and "wrapper".

  In the first part of this chapter we will present a proposed algorithm for the selection of parameters computed on the original EHG and on different frequency bands, using the feature selection technique, of type filter, named Jeffrey Divergence method (JD). This method is based on the measurement of the distance between the 2 histograms computed, for a given feature, from the pregnancy and the labor classes.

  In the second part we will test several feature selection methods, in order to select the most pertinent features that can discriminate between labor and pregnancy contractions. Four methods of type filter are used: Jeffrey divergence (JD) [27], F-score [28], Relieff [29], mutual information based on clustering (MI) [30]; and 7 feature selection methods of type wrapper: sequential Forward Selection (SFS) [31], Sequential backward Selection (SBS)[32], Plus-l minus-r selection (LRS)[33], Bidirectional search (BDS) [33], Sequential Forward Floating sequential (SFFS) [34] , Genetic Algorithm (GA) [35], Binary particle swarm optimization (BPSO)[36]. We use the classifier KNN for these wrapper methods and the two data split: holdout and KFOLD. Following the feature selection step, we will try to validate the selected feature subsets by calculating the percentages of correct classification on a different EHG population.

  These methods will be applied on bipolar EHGs in the first part of this chapter and on bipolar and monopolar EHGs in the second part.

- **Chapter 3:** We present at the beginning of this chapter the channel selection process developed for the monovariate approach (linear and nonlinear features). We will first apply two filter methods (F-score and relieff) to select the relevant channels; and then from the selected channels, we will use two feature selection methods of type wrapper (genetic algorithm and binary particle swarm optimization) using the classifier KNN and the two datasplit (Holdout and Kfold) to select the best features. A validation step will then be presented by calculating the percentages of correct classification (using another

population of EHG) of the subsets of feature obtained after channel selection followed by feature selection.

The final part of this chapter contains the same kind of approach, but applied to the bivariate analysis: channel combination selection using the bivariate features, using here the same procedure than for channels selection followed by feature selection in the monovariate approach.

A general conclusion and perspectives will finally be presented.

The results obtained in this thesis have been published in several journal papers and conferences.

List of author's publications

International journal papers

D. Alamedine, M. Khalil, and C. Marque, "Comparison of Different EHG Feature Selection Methods for the Detection of Preterm Labor," ***Computational and Mathematical Methods in Medicine***, vol. 2013, pp. 1-9, 2013.

D. Alamedine, A. Diab, C. Muszynski, B. Karlsson, M. Khalil, and C. Marque, "Selection algorithm for parameters to characterize uterine EHG signals for the detection of preterm labor," ***Signal Image Video Process***., pp. 1–10, Jun. 2014.

National journal paper

D. Alamedine, M. Khalil, and C. Marque, "Parameters extraction and monitoring in uterine EMG signals. Detection of preterm deliveries," ***IRBM journal*****,** vol. 34, no. 4–5, pp. 322–325, Nov. 2013.

International conference papers

D. Alamedine, C. Marque, and M. Khalil, "Binary particle swarm optimization for feature Selection on uterine electrohysterogram signal," **in** ***Advances in Biomedical Engineering (ICABME), 2013 2nd International Conference IEEE 2013***, Tripoli, Lebanon, pp. 125–128, 11-13 September 2013. (**Oral presenttaion)**

D. Alamedine, M. Khalil, and C. Marque, "Feature Selection Techniques in Uterine Electrohysterography Signal," **in** ***XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*****,** ***Medicon 2013,*** Sevilla, Spain, pp. 779–782, 25-28 September 2013. (**Oral presentation)**

National conference papers

D. Alamedine, M. Khalil, C. Marque, "Comparison of Feature selection for Monopolar and Bipolar EHG signal ", ***Recherche en Imagerie et Technologies pour la Santé, RITS 2015***, Dourdan, France, 25-27 mars 2015**. (Oral presentation)**

D. Alamedine, M. Khalil, C.Marque, "Filter and wrapper methods for feature Selection in Uterine EMG Signals", ***LAAS 2014.*** (**Poster presentation)**

D. Alamedine, M. Khalil, C. Marque, "Parameters Extraction and Monitoring in Uterine EMG Signals. Detection of Preterm Deliveries", ***Recherche en Imagerie et Technologies pour la Santé, RITS 2013***, Bordeaux, France, 8-11 avril 2013. (**Poster presentation)**

D. Alamedine, M. Khalil, C.Marque, "Detection of Preterm Deliveries using EHG signal. Proposition of a New Parameters Selection Method", ***LAAS 2013.*** (**Poster presentation)**

References

[1] Neonatal and perinatal mortality: country, regional and global estimates; ReportWHO 2006. http://whqlibdoc.who.int/publications/2006/9241563206_eng.pdf

[2] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero, "Epidemiology and causes of preterm birth," *The lancet*, vol. 371, no. 9606, pp. 75–84, 2008.

[3] J.D. Iams, "Prediction and early detection of preterm labor," *Obstet. Gynecol*., vol. 101, no. 2, pp. 402-12, Feb. 2003.

[4] H. Alvarez and R. Caldeyro, "Contractility of the human uterus recorded by new methods," *Surgery, Gynecology & Obstetrics*, vol. 91, no. 1, pp. 1–13, 1950.

[5] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchene, "Uterine electromyography: a critical review," *Am J Obstet Gynecol,* vol. 169, pp. 1636-1653, Dec 1993.

[6] S. Mansour, D. Devedeux, G. Germain, C. Marque, and P. J. Duchene, "Uterine EMG spectral analysis and relationship to mechanical activity in pregnant monkeys," *Med. Biol. Eng. Comput.*, vol. 34, no. 2, pp. 115–121, Mar. 1996.

[7] R. E. Garfield and W. L. Maner, "Physiology and electrical activity of uterine contractions," *Seminars in Cell & Developmental Biology,* vol. 18, pp. 289-295, 2007.

[8] C. K. Marque, J. Terrien, S. Rihana, and G. Germain, "Preterm labour detection by use of a biophysical marker: the uterine electrical activity," *BMC Pregnancy Childbirth*, vol. 7, no. Suppl 1, p. S5, Jun. 2007.

[9] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, "A multisensor data fusion approach for improving the classification accuracy of uterine EMG signals," *in Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 93–96.

[10] B. Moslem, M. O. Diab, C. Marque, and M. Khalil, "Classification of multichannel uterine EMG signals," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*,pp. 2602–2605, 2011.

[11] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, "Classification of multichannel uterine EMG signals by using unsupervised competitive learning," *in Signal Processing Systems (SiPS), 2011 IEEE Workshop on*, 2011, pp. 267–272.

[12] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, "Combining multiple support vector machines for boosting the classification accuracy of uterine EMG signals*," in Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 631–634.

[13] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, "Classification of multichannel uterine EMG signals using a reduced number of channels," *in Mechatronics and its Applications (ISMA), 2012 8th International Symposium on*, pp. 1–4, 2012.

[14] B. Moslem, M. Diab, M. Khalil, and C. Marque, "Combining data fusion with multiresolution analysis for improving the classification accuracy of uterine EMG signals," *EURASIP J. Adv. Signal Process*., vol. 2012, no. 1, pp. 1–9, Aug. 2012.

[15] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, "Combination of Canonical Correlation Analysis and Empirical Mode Decomposition Applied to Denoising the Labor Electrohysterogram," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2441–2447, Sep. 2011.

[16] T. Jérémy, S. Thora, M. Catherine, K. Brynjar, and others, "Synchronization between EMG at different uterine locations investigated using time-frequency ridge reconstruction: comparison of pregnancy and labor contractions," *EURASIP J. Adv. Signal Process*., vol. 2010, Article ID 242493, 2010.

[17] J. Sikora, A. Matonia, R. Czaba´nski, K. Horoba, J. Jezewski, and T. Kupka, "Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals," *Archives of Perinatal Medicine*, vol. 17, no. 2, pp. 97–103, 2011.

[18] C. Marque, H. Leman, M. L. Voisine, J. Gondry, and P. Naepels, "Traitement de l'électromyogramme utérin pour la caract´erisation des contractions pendant la grossesse," *RBMNews*, vol. 21, no. 9, pp. 200–211, Dec. 1999.

[19] G. Fele-˘Zor˘z, G. Kav˘sek, ˘Z. Novak-Antoli˘c, and F. Jager, "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and preterm delivery groups," *Medical and Biological Engineering and Computing*, vol. 46, no. 9, pp. 911–922, 2008.

[20] M. O. Diab, C. Marque, and M. A. Khalil, “Classification for uterine EMG signals: comparison between AR model and statistical classification method,” *International Journal of Computational Cognition*, vol. 5, no. 1, pp. 8–14, 2007.

[21] A. Diab, M. Hassan, C. Marque, and B. Karlsson, “Quantitative performance analysis of fourmethods of evaluating signal nonlinearity: application to uterine EMG signals,” *in Proceedings of the 34th Annual International IEEE EMBS Conference*, San Diego, Calif, USA, September 2012,pp. 1045–1048.

[22] B. Moslem, M. Khalil, M. O. Diab and C. Marque, “Detrended fluctuation analysis of uterine electromyography,” *in First Middle East Conference on Biomedical Engineering, MECBME11*, Sharjah, UAE, 2011.

[23] K. Ansari-Asl, F. Wendling, J. J. Bellanger, and L. Senhadji, “Comparison of two estimators of time-frequency interdependencies between nonstationary signals: application to epileptic EEG,” *in 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004. IEMBS ’04, 2004, vol. 1, pp. 263–266.

[24] M. Hassan, A. Alexandersson, J. Terrien, C. Muszynski, C. Marque, and B. Karlsson, “Better pregnancy monitoring using nonlinear propagation analysis of external uterine electromyography”, *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 4, 2013,pp. 1160–1166.

[25] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, “Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings,” *Presented at the 32 nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Buenos Aires, Espagnol, Aug.2010.

[26] A.DIAB, “Study of The Nonlinear Properties And Propagation Characteristics Of The Uterine Electrical Activity During Pregnancy And Labor”, Ph.D. dissertation, Thèse de l’université de Technologie de Compiègne, 2014

[27] Y. Ma, X. Gu, and Y. Wang, “Histogram similarity measure using variable bin size distance,” *Comput. Vis. Image Underst*., vol. 114, no. 8, pp. 981–989, Aug. 2010.

[28] S. Ding, “Feature Selection Based F-Score and ACO Algorithm in Support Vector Machine,” *in Second International Symposium on Knowledge Acquisition and Modeling*, 2009. KAM ’09, 2009, vol. 1, pp. 19–23.

[29] I. Kononenko, “Estimating attributes: analysis and extensions of RELIEF,” *in Machine Learning: ECML-94*, 1994, pp. 171–182.

[30] H. Liu, Y. Mo, J. Wang, and J. Zhao, “A new feature selection method based on clustering,” *in 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2011, vol. 2, pp. 965–969.

[31] L. Ladha and T. Deepa, “Feature selection methods and algorithms,” *Int. J. Comput. Sci. Eng*., vol. 3, no. 5, pp. 1787–1797, 2011.

[32] A. W. Whitney, “A direct method of nonparametric measurement selection,” *Comput. IEEE Trans*. On, vol. 100, no. 9, pp. 1100–1103, 1971.

[33] A. R. Webb, Statistical Pattern Recognition. *John Wiley & Sons*, 2003.

[34] P. Pudil, J. Novovičová, and J. Kittler, “Floating search methods in feature selection,” *Pattern Recognit. Lett*., vol. 15, no. 11, pp. 1119–1125, Nov. 1994.

[35] M. Melanie, “An introduction to genetic algorithms,” *Camb. Mass. Lond. Engl. Fifth Print.*, vol. 3, 1999.

[36] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi, “Application of the PSOSVM model for recognition of control chart patterns,” *ISA Transactions*, vol. 49, no. 4, pp. 577–586, 2010.

Chapter 1: Clinical problem: detection of Preterm Labor by means of uterine electromyography

1.1 Introduction

For most pregnant women the period of pregnancy occurs without difficult and the woman gives birth at term that is between 37 and 40 weeks of pregnancy. However, for the others, this period may end prematurely and induce serious complications, when the woman gives birth before 37 weeks of pregnancy. This is called premature/preterm labor (or preterm birth). Preterm birth is a major cause of neonatal morbidity and mortality. The medical, physiological and socioeconomic consequences of these prematurity are important. Indeed some days more in utero can improve the maturation of the fetus and hence its viability at birth. The early detection of a preterm labor (PL) is important for its prevention and, for that purpose, good indicators of preterm labor are needed.

One of the most promising biophysical markers of PL is the electrical activity of the uterus, the electrohysterogram (EHG). It permits to monitor the efficiency of uterine contractions during pregnancy. The EHG is the signal recorded on the abdominal surface, which represents the electrical activity triggering the mechanical contraction of the myometrium. The Goal of this thesis is the early detection of preterm labor from the analysis of EHG during pregnancy.

The purpose of this chapter is to present an overview of the various aspects to be considered in this study. We first give background on the physiologic mechanisms of uterus contractility, related to the maintenance of pregnancy and to labor induction. We then describe the mechanical and electrical aspects of uterine contraction. Concerning the problem of premature birth (causes, consequences…) we summarize the actual knowledge as well as some methods that have been used previously for the detection of preterm labor. It is known that uterine contractility depends on the excitability of uterine cells and also on the propagation of electrical activity to the whole uterus. Therefore, we present here an overview of the different excitability and propagation analysis that have been done from EHG signals. At the end of this chapter, we give an

introduction on the different multichannel recordings used in previous work and then we describe the multichannel system for EHG recording and the experimental protocol used in this thesis. We conclude this chapter by presenting the different goals of this work.

1.2 Anatomy and physiology of the uterus

1.2.1 Uterus structure

The uterus is a hollow muscular organ involved in the female reproductive system in which the fetus is developing during pregnancy. The anatomy of the uterus includes three portions: the fundus, which corresponds to the upper portion, the corpus which is the main part of the uterus including uterine cavity, and the narrow, lower section named the cervix.

The uterus is located above the vagina, midway between the bladder and the rectum. The non-pregnant uterus weighs 50 to 70 g and measures approximately 7.5 cm in length, 4 to 5 cm in width at its upper portion, and 2 to 3 cm in thickness [1] (Figure 1.1).

The uterine wall, which is thick, is formed of three layers (Figure 1.1): endometrium, perimetrium and myometrium [2]. The endometrium is the inner layer that lines the uterus. It consists of glandular cells that produce secretions. This membrane thickens to prepare the uterus for implantation of a fertilized egg. The perimetrium is the outer layer enveloping the body of the uterus and part of the cervix. The middle layer is the myometrium and forms the larger part of the uterine wall. It is composed of three layers of smooth muscles. This layer has an active role during pregnancy. It increases both by hypertrophy of the existing cells and by multiplication of the cell number. During the last stage of gestation, the smooth cells reach a maximum length of 300 μm and a maximum width of 10 μm [3]. The interaction of myosin and actin filaments produces the contractions of smooth muscle cells. When delivery occurs, the electrical activity generated by the smooth muscle cells in the myometrium, produces rhythmic contractions, which lead to birth.

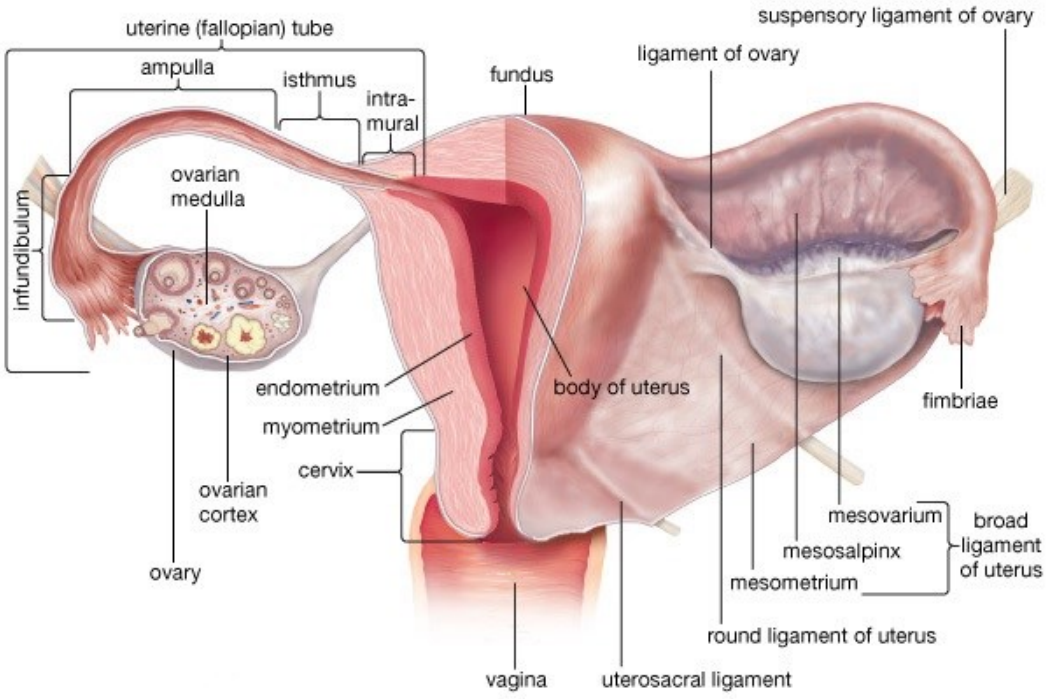


Figure 1.1: Female reproductive system [4].

1.2.2 Mechanical activity of the uterus

The gravid uterus includes a phase of relative quiescence during most of the pregnancy, followed by a period of activity leading to birth. During the active phase, the Intra Uterine Pressure (IUP) permitted to evidence two types of pregnancy contractions:

- Contractions of low IUP amplitude, which have a very local influence, named Low Amplitude High Frequency (LAHF) contractions. They occur during the first trimester of pregnancy and appear with a frequency about 1/min.

- Contractions of higher IUP amplitude but of lower frequency of appearance (from 1/day at the beginning to 1/hour) that appear at mid-pregnancy. These contractions are called Braxton Hicks contractions. Their influence extends to a larger portion of the uterus. During the last weeks of pregnancy, Braxton Hicks contractions become stronger and more frequent.

Then, when reaching the final term, the cervix starts to soften and dilate and contractions progressively evolve in amplitude and frequency.

At the beginning of labor, the propagation of electrical activity increases significantly. The contractions partially propagated of the end of pregnancy disappear and are replaced by labor contractions. These intense and frequent contractions propagate to the entire uterus and induce the opening of the cervix and expulsion of the fetus [5].

1.2.3 Electrical activity of the uterus

Each uterine contraction occurs after the generation and propagation of electrical activity in the myometrium cells [6, 7, 8]. Several authors have studied the changes in the electrical activity of the uterus during pregnancy to understand the changes in the characteristics of the myometrial contraction occurring before delivery. It has been shown that the electrical activity depends on two parameters related to the contractile process: the excitation and the propagation of the electrical activity. The evolution of uterine contractions, from weak and inefficient during pregnancy to strong and efficient during labor, is therefore related to an increase in cellular excitability as well as to an increase in the synchronization of the entire uterus [6]. Giving birth occurs after regular and efficient uterine contractions, which cause dilation of the cervix and push the baby out.

1.2.3.1 Cellular excitability

The electrical activity can be characterized using two types of potential: the resting potential and the action potential. The difference between the negative inside and the positive outside of an inactive cell corresponds to the resting potential. When recording the electrical activity of a membrane, the resting potential is unstable. It is formed of a slow wave of small amplitude, responsible for the electrical base lines. Above a certain threshold of variation of the resting potential, the action potentials are generated. The action potential is due to sudden variations in the permeability of the cell membrane, and corresponds to a reduction of the positivity of the external potential to the inner one. In the uterine contractions, these action potentials are often grouped by bursts. During pregnancy, the physiological electrical activity is composed of discontinuous bursts of action potentials (figure 1.2). This inconstant electrical activity has the consequence of the existence of irregular uterine contractions, of low intensity and localized to certain parts of myometrium. On the other hand term and labor uterine electrical activity is composed of regular bursts with several peaks of action potential (figure 1.2) and propagated to

the entire myometrium in a short time, leading to regular intense contractions [9]. These latter contractions induce delivery generating cervical dilation and expulsion of the fetus.

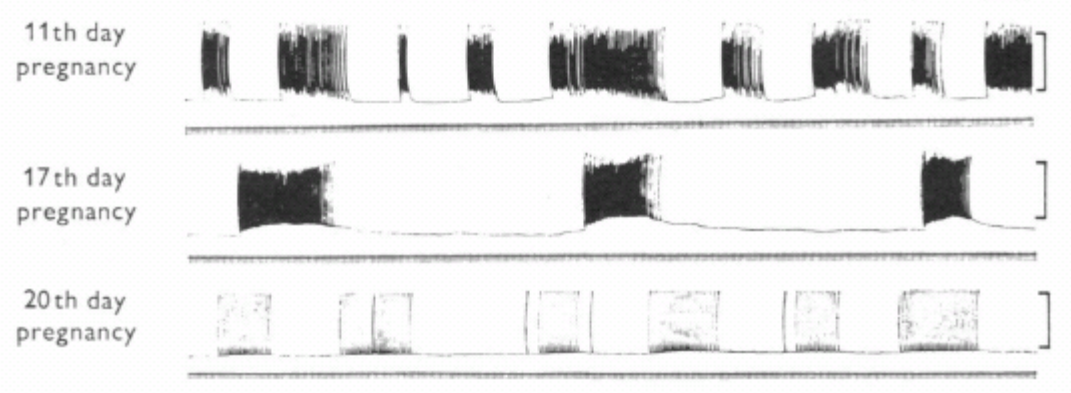


Figure 1.2: Electrical activity of rat uterus at different terms of pregnancy [10].

1.2.3.2 Uterine synchronization

The propagation of the electrical activity in the uterine muscle is ensured by a local electric potential propagation between active cells to their inactive neighbours, electrically coupled through local ionic currents [11]. Therefore, myometrial cell can either be excited by the action potentials from a neighboring cell (it's called "pacefollower") or trigger its own potential (it's named "pacemaker"). However, myometrial cells may alternatively be "pacemaker" or "pacefollower" cells. Furthermore, there is no fixed pacemaker zone in the uterus. Any cell can generate a burst of activity. The "pacemaker" cell can change from one contraction to another. In addition, the inter-cell electrical coupling is improved by the presence of gap junctions (GJ). These gap junctions are areas where the membranes of two adjacent cells form pores allowing an electrical coupling [5, 6]. They are low resistance zones between myometrium cells that permit them to communicate. It appears that the gap junctions have a significant role in the development of a synchronous electrical activity when approaching delivery. During gestation, the number of junctions is much low. The GJ are created in large number few hours before delivery, ensuring the development of a synchronized muscle activity (figure 1.3) [12]. The uterine coordination is much less efficient during pregnancy than for the labor, permitting the pregnancy to proceed smoothly.

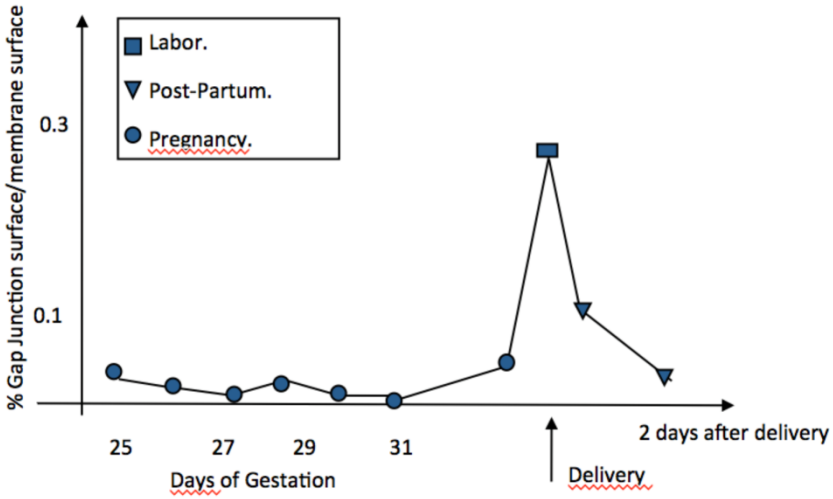


Figure 1.3: Evolution of GJ surface during Pregnancy, Labor and Post-Partum [12].

1.3 Problem of preterm labor

Preterm birth, that is, birth before the 37th week of pregnancy, remains a major problem in obstetrics. In Europe and other developed countries, the incidence of preterm birth is between 5 and 12% [13]. Preterm birth accounts for 75% of perinatal mortality and about 50% of infant long-term mortality [13]. Infants born prematurely are at high risk of mortality as well as health and development problems [13, 14]. For the infant, the risks are:

- neurological development
- lung problems
- psychological problems
- mental retardation

Despite significant technological advances in the development of intensive care in perinatology, these problems persist. Premature births also cause many affective and economic costs, for families and Society. Despite increased knowledge and understanding of the phenomena involved in the onset of premature labor, and despite improvements in obstetric care to attempt to reduce the incidence of premature births, the number of premature births has increased in recent years in most industrialized countries [13].

It has been proved that a week more in utero reduces the risk of neurosensory deficiencies. It is crucial to develop efficient methods to predict preterm labor. Once preterm labor is detected, the treatment consists on administrating the Mother tocolytic agents (inhibitors of uterine

contractility), to relax the uterus delay the delivery. The earlier the administration, the more efficient the treatment is. But methods currently used in obstetric are neither precise nor reliable enough to detect early the threat of premature birth. We thus need a reliable method for the early detection and prevention of premature delivery threat.

1.4 Methods for monitoring pregnancy

One of the aims of pregnancy monitoring is to differentiate normal pregnancy contractions, which are inefficient to those, efficient, which could cause a dilatation of cervix, thus inducing a premature birth. The current methods used in obstetric for pregnancy monitoring, are based on different indicators, most of them being inefficient for a reliable detection of preterm labor:

Intrauterine pressure (IUP), the mechanical effect of uterine contraction, is the only direct method providing precise information concerning uterine contractions efficiency. A catheter is inserted into the uterine cavity and connected to a pressure sensor, giving information on the duration, amplitude and frequency of appearance of the contractions [15]. Despite the accurate information provided by IUP, its major drawback is its invasiveness. IUP cannot clearly be used during pregnancy as it requires rupture of the membranes in order to insert the catheter into the amniotic sac. Therefore, it can increase the risk of infection or accidental labor induction [9].

Tocography being external and non-invasive, is the most widely used method for monitoring uterine contractions during pregnancy. This device contains a force sensor placed on the mother's abdomen, usually over the uterine fundus. This sensor measures the deformation of the abdomen as a result of a contraction [15, 16]. Opposite to the IUP measurement, external tocography is widely used for the non-invasive monitoring of pregnancy. But it does not give an accurate information on the mechanical effect of contractions. It only permis to detect the number of contractions over a given time interval (usually 10mn). However, it is not possible to characterize the efficiency of contractions by means of external tocography. It has been shown that this parameter is not a good predictor of preterm delivery.

Biological tests, such as fibronectin, have been clinically used for the prognosis of premature births [17], although they have a low predictive value.

Clinicians also monitor several other indicators, like: cervical dilation and effacement, vaginal bleeding, or ruptured membranes [18]. But these indicators have had limited success in reducing premature birth. For example, the measurement of cervical dilation is not a reliable indicator of preterm labor. Indeed, a high proportion of women with cervical dilation during pregnancy, give birth to term, even without administration of tocolytic agents. In this context, a test called Bishop score, is made from different parameters evaluated by vaginal examination [19] to foresee impending birth. However, since these parameters are subjective and have a high variability within and between observers, this test has a low predictive value.

A noninvasive technique named light-induced auto fluorescence (LIF) has been proposed for labor monitoring [20]. It attempts to evaluate optically the concentration of collagen in cervical tissue. The result shows that this technique has the capability for estimating the cervical status (softening, ripening), which could give information for preterm labor prediction. This technique is not used in routine practice.

By using the measurement of cervix length via endovaginal ultrasonography, good predictive values were obtained only after the appearance of symptoms of preterm labor [21]. Therefore, a limited degree of success is obtained when using this technique to detect preterm labor. Additionally, the measurement of the cervical length using this technique is not reliable because it is influenced by the varying amount of urine in the bladder [17].

Magnetomyography (MMG) measures the magnetic fields associated with the uterine action potentials. It is a noninvasive technique. A device based on 151 magnetic senso array were used to MMG recording of spontaneous uterine activity [22]. Due to the high cost of this equipment, its use remains limited to research.

1.5 A new method for Preterm labor detection: Uterine Electromyography

A contraction of the uterine muscle occurs due to the generation of electrical activity in a given uterine cell that spreads to other, neighboring cells. The evolution of uterine contractions, from weak and inefficient during pregnancy to strong and efficient during labor, is therefore related to an increase in cellular excitability as well as to an increase in the synchronization of the entire uterus [6].

One of the promising methods to monitor the efficiency of uterine contractions during pregnancy is the analysis of the electrical activity of the uterine muscle, the electrohysterogram (EHG) [23], recorded on the mother's abdomen [24]. EHG consists of the summation of the electrical activity generated by the active uterine muscle cells, plus the noise related to corrupting electrical and mechanical activities.

EHG, recorded externally, has been demonstrated to be representative of the uterine electrical activity when recorded internally [5, 25]. According to these results, we can expect that this non-invasive recording of the uterine EMG will provide information on the excitability of myometrial cells and on the propagation of the electrical activity. Therefore, the EHG analysis has been proposed as a new technique to assess the contractile activity of the uterus during pregnancy and to give accurate information on the physiology of uterine contraction [5, 26].

The EHG, recorded by using surface electrodes placed on the abdomen, is characterized by a low frequency activity (0.03 to 0.1 Hz), with a superimposed activity of higher frequency (FW, fast Wave: 0.3 to 2 Hz). The low-frequency signal is considered as resulting from mechanical disturbances induced to the deformation of the abdomen under the effect of contractions [27]. Conversely, FW (then parted into two components FWL - Fast Wave Low- and FWH - Fast Wave High) is related to uterine contractions. The comparison between contractions during pregnancy and labor showed that their energy is mainly in the 0.2 - 3 Hz frequency band [28]. It has been shown that there is nevertheless a difference in frequency content between the two types of contraction (pregnancy and labor). Indeed, there is a shift towards higher frequencies of contractions, as contraction efficiency progresses [28].

During the last 15 years, many research teams have worked on the possible detection of preterm labor by means of external EHG recording and processing [23, 29]. The most important studies based on excitability and propagation parameters are presented below.

1.5.1 Parameters extracted from EHG

As previously presented, uterine contractility depends on the excitability of uterine cells and on the propagation of the electrical activity to the whole uterus [5]. Many teams have tried to extract from EHG signals features related to these two physiological phenomena, in order to get information relevant to labor detection or preterm labor prediction. The EHG signal has been studied mainly after segmentation of the electrical bursts corresponding to the mechanical contractions of the uterus [30].

Several tools of EHG signal processing have been used first to analyze the excitability (linear method and nonlinear method) from processing only one EHG lead (monovariate analysis), and then the propagation of uterine electrical activity, by using correlation between 2 EHG leads (Bivariate analysis). Linear methods, in both time and frequency domains, were the first used to extract features from one EHG signal. Recently, much attention has been paid to the use of nonlinear analysis techniques for the EHG characterization. Additionally in the last few years, several teams focused on the study of uterine synchronization based on multi-electrode EHG recordings. They studied EHG linear and/or non-linear dependencies in order to discriminate between labor and pregnancy contractions.

The paragraph below aims to expose a bibliographic study of the tools of EHG signal processing used to present.

1.5.1.1 Excitability analysis - Linear parameters

The first parameters used for the characterization of uterine contractility/excitability were extracted from the time domain [20, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37]. Then several EHG analysis and characteristics extraction were made in other domains. These include frequency representation, through Fourier transform and time-frequency representation or wavelet transform.

In the frequency domain, the power spectral density (PSD) of the electrical activity has been used in several studies. By using PSD, several variables were calculated, such as peak frequency, [9, 29, 38, 39, 40, 41, 42, 43, 44, 45], mean frequency [46], median frequency [33, 44, 47, 48] and the ratios of the energies contained in several frequency bands, or the relative energy, have also been used for the characterization of the efficiency of contractions [30].

Among the earliest approaches was that by Marque et al. [28] who computed the relative energy in two specific frequency bands, extracted from the power spectrum density (PSD). They noticed a shift toward higher frequency when comparing late pregnancy to labor contractions. Then, they applied to each selected EHG burst, two temporal features (duration, relative magnitude), fourteen frequency features calculated from the PSD (relative energy in 10 bands, peak frequency, median frequency, frequency kurtosis and skewness coefficients), and four features extracted from the instantaneous frequency IF (minimum value and its time of occurrence, maximum value and its time of occurrence) [33, 41]. Using the cross-correlation matrix, they chose to retain only the least correlated variables (correlation coefficient <80%) giving thus 12 features: duration of the contraction, relative magnitude, *B2* (0.3–0.6Hz), *B3* (0.6–0.9Hz), *B4* (0.9– 1.2Hz), *B5* (1.2–1.5Hz), high-frequency bands BHF (1.5– 3Hz), kurtosis coefficient, maximum IF value, relative associated time, minimum IF value and relative associated time.

Vinken et al. indicated that peak frequency in both human and animal studies may be the most predictive of true labor [43]. Additionally, a study [29] shows that there are statistically significant differences in the mean values of peak frequency and deviations in the EHG recordings during long-term work (TL) and non-working time (TN) and also between preterm labor (PTL) and non-preterm labor (PTN).

Maner et al. [39] also used the peak frequency to detect preterm labor computed from EHG recorded on 99 women 48h, 24h, 12h, and 8h from term delivery, and 6 days, 4 days, 2 days, and 1 day from preterm delivery. They noticed that the PSD peak frequency increased as the measurement-to-delivery interval decreased. This result is obtained by several studies [38, 39, 40, 42, 49].

Sikora et al. [44] also used one temporal feature, named intensity of the contraction (average value of the rectified EHG in 1-min interval) and also 3 features extracted from the PSD (power, median frequency and peak maximum) computed from EHG recorded on 27 women in normal labor, 21 women with symptoms of threatened preterm labor and 14 women in the first labor period. They conclude that power, median frequency and peak frequency permit to discriminate labor and pregnancy contractions (using a Lagrangian Support Vector Machines), with a better efficiency than when using features extracted from the mechanical effect recorded externally (external tocography).

Moslem et al [48] used four frequency parameters (mean frequency, peak frequency, median frequency and the 95%-limit frequency) extracted from 50 contractions randomly selected from each class (labor vs. pregnancy). After comparing their classification performances (pregnancy vs. labor) using receiver operating characteristic (ROC) curve analysis, they indicated that median frequency is the best frequency parameter that can be used for discrimination between pregnancy and labor contractions.

Some authors have also used time–frequency methods such as wavelet decomposition [50, 51] to characterize EHG. They extracted relevant parameters, after the decomposition into details coefficients. Diab et al. calculated the variances on five selected detail levels [51]. In this study, wavelet decomposition was applied on real and simulated bursts obtained after autoregressive modeling of EHG for pregnancy and labor. Arora et al. [50] computed the relative wavelet energy of each detail of a 4 level decomposition. These values formed the input of different classifiers. Both works indicate that wavelet decomposition provides a pertinent approach for discriminating pregnancy and labor EHG bursts. Lu et al. [52] used the wavelet packet decomposition to discriminate between preterm signals and term signals.

Wavelet packet decomposition has also been used to reduce the width of the frequency band analysis. A study has decomposed each contraction, selected from EHG recorded during pregnancy and labor, on packets of 3 decomposition levels [53]. Then, the authors computed the relative energy for each packet. The results showed a decrease in the amount of energy in the low-frequency bands associated with an increase in other bands that represent the high frequencies throughout pregnancy. Furthermore, they observed a noticeable difference in the energy distribution between two studied classes (pregnancy and labor).

In [54] they calculated the Normalized Wavelet Packets Energies after decomposition of the EHG into a 3-level wavelet packet tree. A principal component analysis (PCA) was applied on the values of the normalized energies of the 8 packets of the third level in order to reduce dimensionality. Then these data were used as inputs to an artificial neural network (ANN) classifier with a Gaussian radial basis function (RBF) to classify labor and pregnancy contractions. The results of this study show that the high percentage of correctly classified data obtained indicates that Normalized Wavelet Packets Energies can be used to classify EHG signals into two classes of contractions (pregnancy vs. labor).

1.5.1.2 Excitability analysis - NonLinear parameters

In an attempt to improve the results obtained using linear methods, and as the EHG, like other biomedical signals, seems to present some nonlinear characteristics, several measures have been proposed for detecting nonlinear characteristics in the EHG.

Radomski et al. [55] have applied sample entropy to identify regularity in uterine EMG. The evaluation of this method was performed on segmenting contractions from signals recorded from 66 pregnant women divided into several groups (26 term labors, 20 preterm labors and 6 post-partum periods). This study shows that the analysis of signals using this method is better than the approximate entropy used by other and allows discrimination between different dynamic states of uterus.

Vrhovec et al. [56] show that we can use sample entropy to evaluate the progress of labor.

Ivancevic et al. [57] have presented a review on the nonlinear parameters used for the prediction of preterm birth. One of the studied methods of interest is the Lyapunov exponent that studies the stability of the signal and its sensitivity to initial conditions.

Moslem et al. [58] used detrended fluctuation analysis (*DFA*) to study the law, which governs the evolution of longterm correlations in EMG uterine. *DFA* was applied to EHG recorded on 11 women (5 pregnancies and 6 labors). Based on this method, the results indicated that pregnancy contractions are less correlated than labor contractions. They stated that this result is important for the detection of premature birth threats.

Diab et al. [59] compared the performances of several nonlinear methods (time reversibility, sample entropy, Lyapunov exponent and delay vector variance) on synthetic signals. The aim was to test their sensitivity to the change of signal complexity, with or without noise. Then, they applied these methods on contractions recorded during pregnancy and labor. Results on synthetic signals showed that time reversibility are less sensitive to changes in noise. On real signals, the results show a clear superiority of time reversibility in classifying pregnancy and labor EHG.

1.5.1.3 Synchronization analysis – propagation/correlation parameters

Another approach, the bivariate analysis, has been developed to characterize the uterine synchronization, by means of the analysis of bipolar EHG propagation. These recently used methods can be divided into two types: phase synchronization (mean phase coherence and phase entropy) [60]; linear (R^2) and nonlinear (H^2) correlations [61]. They used these methods to characterize the synchronization between 2 EHGs, recorded at different locations on the abdomen, and study either its evolution during normal pregnancy, or its ability to discriminate pregnancy and labor contractions. The results indicated that nonlinear correlation between contractions during labor is stronger than during pregnancy. Results also show an increase in phase synchronization when passing from pregnancy to labor [61].

A study [62] compared several coupling methods, two nonlinear methods -nonlinear correlation coefficient (H^2), General synchronization (H) and one linear method, Granger causality (GC). They try to test the sensitivity of these methods to some characteristics of signal (nonstationarity, frequency band) or signal recording choice (bipolar or monopolar recording) in order to improve the performance of the coupling detection methods to improve the classification of contractions of pregnancy and labor EHG bursts. They found that using a combination of two preprocessing steps (windowing-preprocessing step, filtration step to retain only the low frequency band of the EHG (FWL)), leads to an increasing of H^2 performance. H method performance is highly influenced by the nonlinearity of EHG signals and therefore requires further investigation. Monopolar recordings are better than bipolar one.

1.5.1.4 Combination of parameters

Other authors tested simultaneously different methods on the same population, in order to study which feature can separate EHG leading to term and preterm deliveries. Fele-Žorž et al. [47] compared four linear methods (root mean square value, peak and median frequency, autocorrelation zero crossing) and three nonlinear methods (Lyapunov exponent, correlation dimension and sample entropy). For the pretreatment of each signal, they defined three band-pass Butterworth filters (0.08–4, 0.3–4 and 0.3–3 Hz). Their results indicate that with the 0.3–3 Hz frequency band, the median frequency and sample entropy gives the best statistical differences between term and preterm delivery registered before the 26th week ($p = 0.03$, $p =$

0.035), among all term and all preterm delivery records ($p = 0.012$, $p = 0.011$), and between term delivery registered before and after the 26^{th} week ($p \leq 0.001$).

Some studies have also compared linear methods with bivariate propagation methods. Terrien et al. [46], on internal uterine EMG recorded on pregnant women, compared the mean frequency, the minimal and maximal frequencies and the frequency of the maximal energy obtained either from the EHG or from 2 (or 3) ridges extracted from the EHG (representing 3 EHG frequency components separately). They also computed the nonlinear correlation coefficient between the signals reconstructed from the 3 ridges extracted from two EHG bursts recorded at different locations for a given contraction. They found that the mean frequency of the low (FWL)- and high-(FWH) frequency components as well as nonlinear correlation coefficient computed from the low-frequency component (FWL) gives the best results in terms of labor prediction.

Lucovnik et al. [45] compared propagation velocity, peak frequency and two clinical methods (Bishop score, Transvaginal cervical length) in order to test the accuracy of these methods for the diagnosis of preterm labor. The results revealed that the combination of the velocity and the maximum frequency gives the best results and permits to predict preterm delivery within 7 days.

In [63] Fergus et al. focused on the classification of term and preterm labors. An open dataset, containing 300 records (38 preterm and 262 term) are used in this study. Four features are computed from EHG signal (root mean squares, peak frequency, median frequency and sample entropy. To validate the choosing features, the discriminant capabilities of each feature are determined using principal component analysis (PCA). The results of PCA indicate that the choosing features have discriminant capabilities. Additionally in this study they used additionally features (clinical features like: age, parity (number of previous births), abortions, weight, hypertension, diabetes, placental position, first and second trimester bleeding, funnelling and smoking (new dataset containing 19 preterm records and 150 term records). Several machine-learning classifiers are used in this study in four approaches (see [63]). The results show that oversampled TPEHG dataset (Data in the minority class is generated using oversampling), with combined additional features using the POLYC classifier gives the best result.

Baghamoradi et al. [64] used sample entropy and thirty cepstral coefficients (the first ten cepstral coefficients of each channel) as features to classify term and preterm delivery labors. They

compare separately the average classification rate of thirty cepstral features and that of subset of features selected in several channel through sequential forward selection (fifth and eights cepstral coefficients from the first channel (S1) and tenth coefficient from the third channel (S3)) and finally that of sample entropy. The result of this study showed that selection of optimum features from informative channels using sequential forward selection can improve the average classification rates.

1.5.2 EHG Recording

In order to record EHG signals, different numbers and locations of electrodes positioned on the woman's abdomen had been used. The influence of the number and position of the recording electrodes has been already used. In several studies, only a limited number of electrodes (2, 3 and 5) were used.

1.5.2.1 Electrode number and position

Marque et al. [23] used bipolar Ag/ AgCl surface electrodes (8 mm diameter, 2.5 cm spacing) for recording EHG. A reference electrode was positioned on the women's hip. They indicate that the best electrode position was the median vertical axis of the woman's abdomen.

In [65], EHG was recorded with an electrode pair positioned in the middle of the median axis near the umbilicus and another one positioned 5cm left of the middle electrodes. They justified this choice by indicating that at extremities there are more visceral tissues between the skin and the uterus than on the middle of the median axis of woman's abdomen, impeding thus a correct uterine EMG recording.

In order to identify a suitable electrode configuration, Rabotti et al [66], placed 11 electrodes on the woman's abdomen and then calculate the average SNR in each electrode. The highest average SNR was obtained on the lower vertical median line of the abdomen, especially on the region immediately below the umbilicus. They describe these results by two physiological phenomena. On the one hand, the distance between the electrode on the skin and the signal source in the myometrium in the vertical median line on the abdomen is reduced with respect to the lateral sides [23]. On the other hand, due to the uterus movement during contraction, the position of the uterus relative to the abdominal wall in the region surrounding the umbilicus remains constant even during contractions [5]. Therefore, according to these results they used

four unipolar Ag/AgCl electrodes placed on the abdomen and a reference electrode was placed on the right hip.

Randomski et al. [55] used a system containing two bipolar electrode pairs, one Tocographic probe and a reference electrode attached on the left hip. The distance between the electrodes forming the bipolar channels was fixed at 5 cm. The electrodes were attached in the vertical median axis of the woman's abdomen because they indicate that this position provides a suitable signal/noise ratio due to a closer contact and during contractions, more invariant position of the uterus in relation to the abdominal wall [67].

Fele-zorz et al. [47], Baghamoradi et al [64] and [63], used four AgCl electrodes to record EHG. These electrodes were placed in two horizontal rows, symmetrically under and above the umbilicus, spaced 7 cm apart. Three bipolar EHG were used in these studies.

Terrien et al. [68] used two electrodes placed on woman's abdomen (inter-electrode distance: 2.1cm, on the uterine median axis, midway between the fundus and the symphysis. A reference electrode was placed on the hip of the women.

Lucovnik et al. [45], used four electrodes positioned around the umbilicus in a form of square shape, 2.5 cm electrode-electrode vertical and horizontal inter-electrode distances (measured from center to center). For EHG recording they use differential, bipolar electrode pairs.

Several studies indicated that this small number of electrodes used in several studies was not suitable for a proper study of the propagation in this complex environment represented by the uterine muscle and abdominal anatomy of a pregnant woman. To this end, Karlsson et al [69] said that it is important to investigate the electrical activity concurrently at different locations on the woman's abdomen. For this reason, they study the propagation of the uterine electrical activity recorded on women during labor by using a 16 (4 x 4) monopolar electrode grid positioned on the woman's abdomen (interelectrode distance: 2.1 cm). They put the third electrode column of the grid on the uterine median axis, and the 10-11h electrodes on the middle of the uterus. Two reference electrodes were placed on each hip of the women. In order to increase the signal to noise ratio, they used the vertical bipolar signals (BPi), giving thus a rectangular matrix of size 3 x 4. Rabotti et al [70] used a multichannel EHG recording composed

of 64 high density electrodes in order to estimate invasively of the conduction velocity of the EHG-action potentials.

1.5.2.2 Channel selection

A comparison between the classification performances of several channels has been done in [71, 72, 73, 74, 75, 76] in order to find which channel gives the best signals classification between pregnancy and labor EHG. Sixteen electrodes were used in these studies, based on the vertical bipolar signals (Vbi). They tried to benefit from this multichannel recording, thus they proposed a new classification approach based on multi-channel analysis of uterine EMG signals, by using different types of classifier. The aim was to compare the 12 channels in order to evidence if some channels could better classify EHG burts. Each channel provided individual classification and the 12 decisions were merged to obtain a final decision. A fusion rule based on the weighted sum of 12 decisions showed that this approach could be a solution to the problem of classification of the EHG by using different classifiers to evaluate the classification performance of each channel from 40 contractions chosen from each class (pregnancy vs. labor). They tested Support vector Machines (SVM) with a Gaussian Radial Basis Function (RBF) kernel [71], Radial Basis Function network (RBF) [72], Competitive Neural Network (CNN) [73]. The results indicated different channels giving either the highest (respectively Vb5, Vb4, Vb10) or the lowest (respectively Vb7, Vb8, Vb7) weighs.

In [75] they tried to find the best combination of channels that can provide the highest classification accuracy. However, the result indicates that the highest classification can be obtained by the use of 4 bipolar channels (Vb10+Vb2+Vb1+Vb4).

In [76] the multichannel classification is combined with other signal processing techniques, such as multiresolution analysis to improve the classification performance. The classifier used in this study is Support Vector Machine (SVM). The results indicate that, for the classification of pregnancy and labor, Vb3 gave the highest predictive value while Vb7 got the worst.

1.5.2.3 Electrode configuration

All previously conducted studies on the EHG have processed bipolar signals, in order to increase the signal/noise ratio (SNR). In this configuration, the signals recorded by two close electrodes are subtracted from each other in order to generate a single bipolar signal. The advantage of this configuration is to reduce the common mode noise and to obtain a signal having a correct SNR. But this configuration introduces a bias for studying the propagation, as two adjacent bipolar signals can share a same electrode. Monopolar EHG could be more interesting to get rid of this bias and increase the spatial resolution when processing signals. For this reason, Hassan et al [77] developed a powerful method to filter monopolar EHG, and thus obtain a correct SNR permitting to study the propagation of the electrical uterine activity. These tools are based on combination of canonical component analysis (CCA) and on EMD (Empirical Mode Decomposition). The SNR obtained by this method CCA-EMD is larger than with bipolar EHG and with monopolar EHG filtered by other methods. Therefore, this denoising method allows us to obtain usable monopolar signals directly to study the synchronization of uterine activity.

1.5.3 Multichannel system for standardized EHG recording: the TMSi system

As indicated above the multichannel EHG recording seems suitable for a proper study of the propagation in this complex environment represented by the uterine muscle and abdominal anatomy of a pregnant woman [69]. As in a labor room, the placement of a large number of electrodes for measuring EHG takes time and is difficult to perform, a project started in Iceland in 2009 in order to create a new design that reduces the inconvenient of multiple electrode positioning, by defining a standard position of recording electrodes. The new design involves a guide with holes, guiding the placement, as well as adhesion of the electrodes by means of a double-coated adhesive sheet. A simple frame ensures correct positioning of the guide onto the adhesive sheet, while positioning the electrodes. With about 2 cm inter-electrode distance, the 16 electrodes (8mm in diameter) are arranged in a 4x4 matrix. This system permitted us to standardize the acquisition of EHG signals within the framework of a previous european project (BioModUE_PTL, http://www.erasysbio.net/index.php?index=268), involving partners from 4 countries (France, Iceland, the Netherlands, Slovenia).

The standardization of EHG measurements has been initiated by using a multielectrode system composed of a 16 channels device, commonly used for investigating sleep disorders (Embla

A10). During the BioModUE_PTL, a new multielectrode system has been used (Porti 32, TMSi) that offers recording of up to 32 channels (Figure 1.4). Based on this system, we standardized the protocol using reusable Ag/AgCl electrodes (8mm diameter), positionned as a 4x4 matrix placed on the woman's abdomen (the third electrode column of the grid being set on the uterine median axis, and the 10-11h electrode on the middle of the uterus), with two reference electrodes on each of her hips. The collected signals are fed into an amplifier and then to an A/D converter (PORTI 32). Then, by using an optical fiber and a USB cable, the signals are collected by a PC where they can be saved on disk or uploaded to an online database. Another signal called TOCO (output of a tocographic probe) is also recorded using the probe placed above the electrode matrix (figure 1.4). This signal is further used for EHG bursts segmentation.

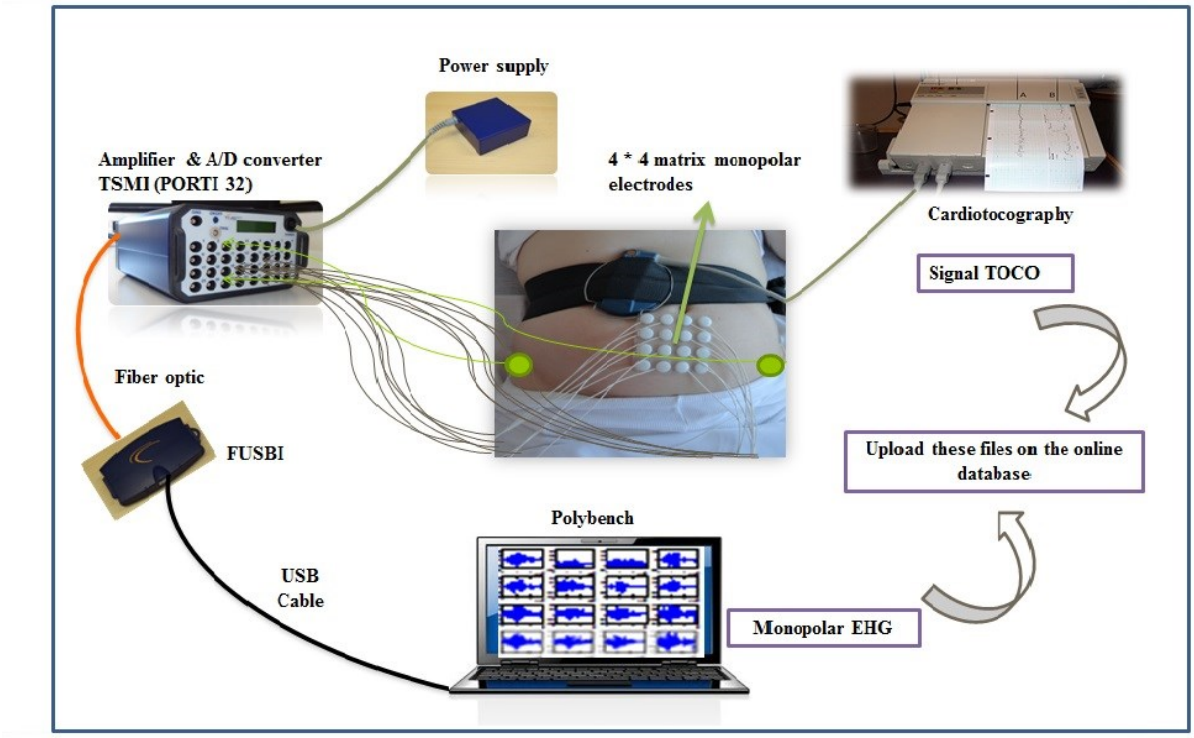


Figure 1.4: The new multichannel system: 4x4 electrode matrix placed on the women abdomen with two reference electrodes on the hip and the tocographic probe.

Before the placement of the electrode matrix and of the two reference electrodes, the skin is carefully prepared using an abrasive paste and alcoholic solution. After the recording, we followed the pregnant women in order to label the signals as either pregnancy or labor. When the woman gave birth within 24 hours, the signal was labeled "labor". If the delivery occurred later, the signal was labeled "pregnancy".

The measurements were performed at two hospitals in France and Iceland. In Iceland, the measurements were performed at the Landspitali University Hospital, using a protocol approved by the relevant ethical committee (VSN02-0006-V2). In France, the measurements were performed at the Center for Obstetrics and Gynecology (Amiens), using a protocol approved by the relevant ethical committee (ID-RCB 2011-A00500-41)).

To increase the signal to noise ratio, we calculated the vertical bipolar signals (Vbi), obtaining thus 12 bipolar signals (Figure 1.5). The monopolar and bipolar bursts of uterine electrical activity that correspond to contractions were manually segmented, based on the TOCO signal recorded simultaneously. The tocodynamometer paper trace was digitalized in order to facilitate the segmentation of the uterine contractions (Figure 1.5).

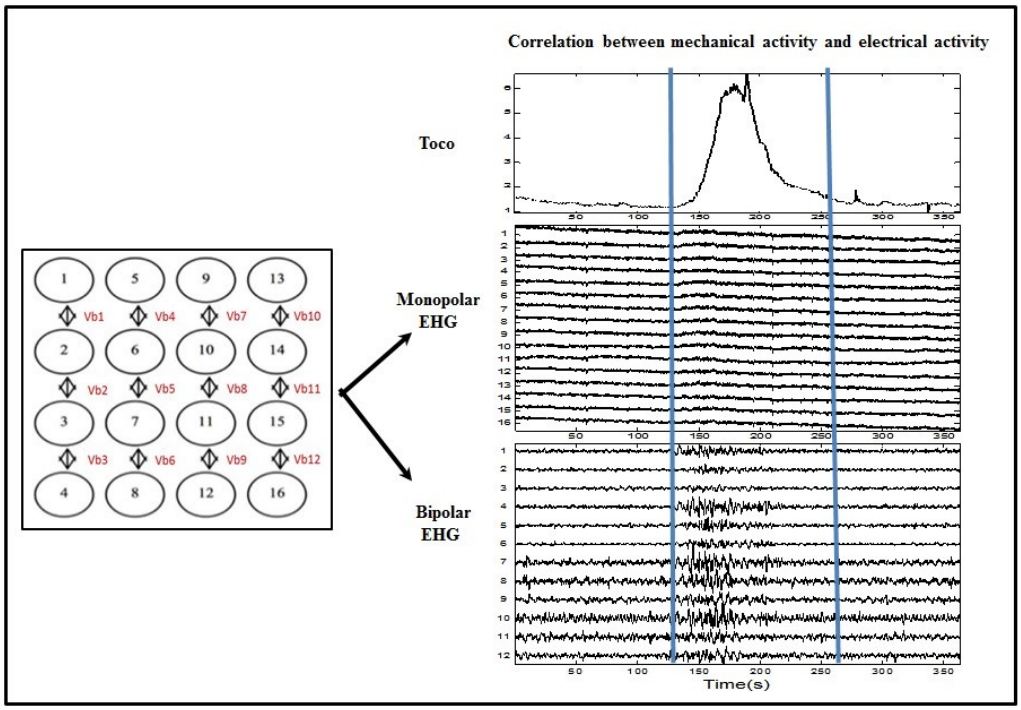


Figure 1.5: Digitized tocodynamometer paper (Top), monopolar signals (middle), corresponding bipolar signals (bottom). The blue lines define the beginning and the end of the contraction according to TOCO.

Recently monopolar burst were denoised by using the powerful EHG filtering methods developed by [77] previously presented, permitting thus to compute all parameters from monopolar and bipolar EHGs.

1.6 Thesis aim and work

1.6.1 Thesis goal

The general aim of this thesis is the early detection of Premature Threat of Labor (PTL) from the analysis of the excitability and of the propagation of the uterine electrical activity. As presented above, several signal-processing tools have been developed by many teams for the analysis of the excitability and the propagation of the uterine electrical activity (linear parameters, nonlinear parameters and parameters related to the EHG propagation) to find specific information that differentiates pregnancy and labor contractions. Therefore, a large number of parameters have been extracted from the EHG in numerous studies, based on very different populations and recording protocols. These different data make difficult the comparison of their pertinence for PTL diagnosis purpose. A large number of parameters increases the computational complexity in a diagnostic objective. Reducing the size of the parameter set through the elimination of irrelevant and/or noisy parameters is very important for pattern recognition. A multichannel system composed of 16 channels is used in our work. Extracted parameters from these channels (using monovariate analysis) or combinations channels (using bivariate analysis) provide a large dimension of features causes a complex classification. Therefore a channels and combinations channels selection is performing in our thesis. Furthermore, the constitution of a statistically significant signal database is very important to develop diagnostic tools (pregnancy monitoring, detection of Premature Delivery Threats).

This thesis work is thus based on several parts:

- Experimental part: Consists in collecting signals from different populations of pregnant women (normal pregnancies or at risk), with 16 electrodes (TMSi system) attached to the woman's abdomen with the standardized protocol, in order to create a database sufficient for the development of diagnostic tools. These recordings are made at the CGO (Gynecologic/Obstetric department of the Amiens hospital, the reference center for risk pregnancy); the already available database has thus been enriched by recordings made on pregnant women during normal or risk pregnancies.

- Feature selection part: A literature review has been done in our thesis, to define the first set of relevant parameters to be tested, for monovariate or bivariate analysis. The reduction of feature

dimensionality through the elimination of irrelevant and noisy features is very important. The main aim of our work is to compare all the parameters of interest, extracted from the bibliography, computed on the same EHG signals corresponding to a standardized recording database, in order to select the most pertinent parameters, which can discriminate between pregnancy and labor contractions. This will permit a reduction of the parameter number used for diagnosis purpose. Therefore a feature selection step is applied in our work, on the set of parameters selected after the literature review, and on the whole EHG signals corresponding to the standardized recording database. Several methods have been used for feature selection. The first one, developed in this work, is based on the measurement of the Jeffrey divergence (JD) distance between the histograms of the parameters of the 2 classes, pregnancy and labor. The other methods are existing Data Mining methods extracted from the literature. Dimensionality reduction classical methods are also used in order to compare these results with those of feature selection methods.

- Channels and combination of channels selection: In continuation of the previous studies [71-75], we will also try to select the most relevant channels (monovariate analysis) and the best combination of channels (bivariate analysis) by applying channel selection methods to the whole available channels collected by our matrix.

We will also test in these two last parts (feature selection, Channels and combination of channels selection) the clinical diagnosis potential of different configurations, by comparing the results obtained by using either monopolar or bipolar EHG.

1.6.2 The used databases

Two databases are used in our work:

- Database 1, correspond to signals recorded on 22 Icelandic women, during the first step of standardization: 11 women during pregnancy (33–39 weeks of gestation) and 11 during labor (39–42 weeks of gestation), and signal recorded on 3 French women: 2 women during pregnancy (33–39 weeks of gestation) and 1 during labor (39–42 weeks of gestation) (Table 1.1).. After the manual segmentation of EHG bursts, we obtained 106 pregnancy bursts and 106 labor bursts corresponding to identified contractions (see women information in Appendix A, Table A.1 and Table A.2). The sampling frequency is 200 Hz.

- Database 2, correspond to signals recorded during and after the BioModUE_PTL project on: 41 Icelandic women: 27 during pregnancy and 14 during labor; on 3 French women: 2 during pregnancy and 1 during labor; on 3 Slovenian women: 3 during pregnancy (Table 1.1). After manual segmentation of EHG bursts, we obtained 290 pregnancy contractions and 189 labor bursts corresponding to contractions (see women information in Appendix A, Table A.3 and Table A.4).

The monopolar bursts were denoised by using the EHG filtering methods developed by [77]. Then, from these denoised monopolar signals we calculated the vertical bipolar EHG signals (Vbi), for the whole pregnancy and labor contractions. The signal sampling is different in these 2 databases. Some contractions were sampled at 200 Hz and the other at 512 Hz, when using the TMSi recording system. We thus downsampled by a factor of 2 these latter one, obtaining thus a new sampling frequency of 256 Hz. This database is divided in two parts. The first part is used for classifier training and testing and the second for the validation (see table 1.2).

Table 1.1: Number of women and contractions used in database 1 and database 2

	Database 1		**Database 2**	
	Pregnancy	**Labor**	**Pregnancy**	**Labor**
Icelandic women	11	11	27	14
French women	2	1	2	1
Slovenian woman			3	
Number of contractions	106	106	290	189

For the feature selection step and the monovariate analysis, we first used only the bipolar channels Vb7 and the CH10 monopolar signals that correspond to the classical positions mostly used in the literature. For the bivariate analysis we used the mean value of feature on the whole combination of available channels. Then, for the channel selection and channel combination selection, we used the whole available channels (12 for bipolar and 16 for the monopolar configurations) and the whole available channel combination (132 for bipolar and 240 for monopolar).

Table 1.2: Number of women and contractions used in database 2 (split data for the classification into training, testing and validation part)

	Part 1		**Part 2**		
Sampling frequency (Hz)	200		512/2=256 Hz		
Type	Pregnancy	labor	Pregnancy	labor	
Number of women for training and testing part	19	10	1	1	
Number of women for validation part	11	4	1	0	
Number of contractions for training and testing part	233	131	7	8	**Total number of contractions:** 240 pregnancy 139 labor
Number of contractions for validation part	48	50	2	0	**Total number of contractions:** 50 pregnancy 50 labor

1.7 Discussion and Conclusion

Preterm birth remains a major problem in obstetrics. Therefore**,** it has been a topic of interest for many researchers. In this chapter we have presented an overview of knowledge about premature birth, uterine physiology, as well as the two phenomena that control the generation of uterine contraction. Different methods for monitoring pregnancy are described. According to this chapter, we can conclude that one of the most promising methods for monitoring uterine activity is based on the study of the electrical activity of the uterus as recorded on the mother's abdominal surface, the electrohysterogram (EHG).

We provided an overview on recording mode of EHG and signal processing, allowing analysis of excitability (linear method and nonlinear method) and propagation of uterine electrical activity. The new multichannel EHG recordings using in our study was described in this chapter.

We notice that a large number of parameters have been extracted from the EHG by many researches, in order to find pertinent information to detect preterm birth. Additionally, on these multiple studies, the parameters being computed from different signal databases, obtained with different recording protocols, it is sometimes difficult to compare their results in order to choose the “best” parameter for preterm labor detection. Therefore, the selection of pertinent features is a very important problem in pattern recognition. We need reliable feature selection methods to reduce the number of features, through the elimination of irrelevant and noisy features. For this reasons, in chapter 2 we will present several methods of feature selection to select the most significant feature subset extracted from multiple studies in order to identify, on a given population, the most pertinent ones and to discriminate between pregnancy and labor contractions.

As the electrode position seems also important for diagnosis purpose, we will also present the results of channel selection (for monovariate analysis) and of channel combination selection (for bivariate analysis) in chapter 3.

References

[1] H. Ellis, “Anatomy of the uterus,” *Anaesthesia & Intensive Care Medicine*, vol. 6, no. 3, pp. 74-75, Mar. 2005.

[2] E. M. Ramsey, “Anatomy of the human uterus,” *The Uterus*, pp. 18–40, 1994.

[3] A.I. Csapo, “Smooth muscle as a contractile unit,” *Phys. Rev. Suppl.*, vol. 5, pp. 7-33, Jul. 1962.

[4]“uterus anatomy,” *Encyclopedia Britannica*. [Online]. Available: http://www.britannica.com/EBchecked/topic/620603/uterus. [Accessed: 15-Jan-2015].

[5] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchene, "Uterine electromyography: a critical review," *Am J Obstet Gynecol,* vol. 169, pp. 1636-1653, Dec 1993.

[6] R. E. Garfield and W. L. Maner, "Physiology and electrical activity of uterine contractions," *Seminars in Cell & Developmental Biology,* vol. 18, pp. 289-295, 2007.

[7] S. Rihana, J. Santos, S. Mondie, and C. Marque, "Dynamical analysis of uterine cell electrical activity model," *Conf Proc IEEE Eng Med Biol Soc,* vol. 1, pp. 4179-4182, 2006.

[8] S. Rihana, E. Lefrancois, and C. Marque, "A two dimension model of the uterine electrical wave propagation," *Conf Proc IEEE Eng Med Biol Soc,* vol. 2007, pp. 1109-1112, 2007.

[9]D. Schlembach, W. L. Maner, R. E. Garfield, and H. Maul, "Monitoring the progress of pregnancy and labor using electromyography," *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 144, pp. S33-S39, 2009.

[10] H. Kuriyama & H. Suzuki. "Changes in electrical properties of rat myometrium during gestation and following hormonal treatments*,* " *Journal Of Physiology*, vol. 260, pages 315–333, 1976.

[11] S. Mansour, "Etude de l'electromyographie uterine: caractérisation, propagation, modélisation des transferts " Génie biomédical, Université de Technologie de Compiègne 1993.

[12] R. E. Garfield, D. Merrett & A. K. Grover. "Gap junction formation and regulation in myometrium, "*American Journal of Physiology Cell Physiology*, vol. 239, no. 5, pages C217–228, 1980.

[13] R. L. Goldenberg, J. F. Culhane, J. D. Iams, and R. Romero, “Epidemiology and causes of preterm birth,” *The lancet*, vol. 371, no. 9606, pp. 75–84, 2008.

[14] K. Costeloe and EPICure Study Group, “EPICure: facts and figures: why preterm labour should be treated,” *BJOG Int. J. Obstet. Gynaecol*., vol. 113, pp. 10–12, 2006.

[15] R. E. Garfield, K. Chwalisz, L. Shi, G. Olson, and G. R. Saade, "Instrumentation for the diagnosis of term and preterm labour," *J Perinat Med,* vol. 26, pp. 413-436, 1998.

[16] M. Cabaniss and M. Ross, *fetal monitoring interpretation*, second ed., 2010.

[17] J.D. Iams, "Prediction and early detection of preterm labor," *Obstet. Gynecol*., vol. 101, no. 2, pp. 402-12, Feb. 2003.

[18] R.K. Creasy, "Preterm birth prevention: where are we?," *American journal of obstetrics and gynecology*, vol. 168, no. 4, pp. 1223-1230, Apr. 1993.

[19] E. H. Bishop, "Pelvic Scoring for Elective Induction," *Obstetrics & Gynecology,* vol. 24, pp. 266-268, 1964.

[20] R. E. Garfield, G. Saade, C. Buhimschi, I. Buhimschi, L. Shi, S. Q. Shi, and K. Chwalisz, "Control and assessment of the uterus and cervix during pregnancy and labour," *Hum Reprod Update*, vol. 4, pp. 673-95, Sep-Oct 1998.

[21] R. Romero, R. Gomez, and W. Sepulveda, "The uterine cervix, ultrasound and prematurity," *Ultrasound. Obstet. Gynecol*., vol. 2, no. 6, pp. 385-8, Nov. 1992.

[22] H. Eswaran, H. Preissl, J. D. Wilson, P. Murphy, S. E. Robinson, and C. L. Lowery, "First magnetomyographic recordings of uterine activity with spatial-temporal information with a 151-channel sensor array," *Am J Obstet Gynecol*, vol. 187, pp. 145-51, Jul 2002.

[23] C. K. Marque, J. Terrien, S. Rihana, and G. Germain, "Preterm labour detection by use of a biophysical marker: the uterine electrical activity," *BMC Pregnancy Childbirth*, vol. 7, no. Suppl 1, p. S5, Jun. 2007.

[24] H. Alvarez and R. Caldeyro, "Contractility of the human uterus recorded by new methods," *Surgery, Gynecology & Obstetrics*, vol. 91, no. 1, pp. 1–13, 1950.

[25] S. Mansour, D. Devedeux, G. Germain, C. Marque, and P. J. Duchene, "Uterine EMG spectral analysis and relationship to mechanical activity in pregnant monkeys," *Med. Biol. Eng. Comput.*, vol. 34, no. 2, pp. 115–121, Mar. 1996.

[26] M. Doret, J. C. Pasquier, C. Gharib, and P. Gaucherand, "L'électromyogramme utérin : principes et intérêt pour le diagnostic de travail prématuré," Journal de Gynécologie Obstétrique et Biologie de la Reproduction, vol. 37, pp. 24-32, 2008.

[27] C. Sureau, "Etude de l'activité électrique de l'utérus au cours du travail," *Gynecol. Obstet*., vol. 555, pp. 153-175, Apr-May. 1956.

[28] C. Marque, J. Duchêne, S. Leclercq, G. Panczer, and J. Chaumont, "Uterine EHG processing for obstetrical monitoring," *IEEE Trans. Biomed. Eng*., vol. 33, no. 12, pp. 1182-7, Dec. 1986.

[29] W. L. Maner and R. E. Garfield, "Identification of human term and preterm labor using artificial neural networks on uterine electromyography data," *Ann Biomed Eng,* vol. 35, pp. 465-473, Mar 2007.

[30] H. Maul, W.L. Maner, G.R. Saade, and R.E. Garfield, "The physiology of uterine contractions," Clinics in perinatology, vol. 30, no. 4, pp. 665-676, Dec. 2003.

[31] C. Sureau, J. Chavinié, and M. Cannon, "L'electrophysiologie uterine," *Bull Fed Soc Gynecol Obstet*, vol. 17, pp. 79-140, 1965.

[32] J. Gondry, C. Marque, J. Duchêne, and D. Cabrol, "Electrohysterography during pregnancy: preliminary report," *Biomed. Instrum. Technol.*, vol. 27, no.4, pp. 318-24, 1993.

[33] C. Marque, H. Leman, M. L. Voisine, J. Gondry, and P. Naepels, "Traitement de l'électromyogramme utérin pour la caractérisation des contractions pendant la grossesse," *RBM-News,* vol. 21, pp. 200-211, 1999.

[34] C. Buhimschi and R. E. Garfield, "Uterine contractility as assessed by abdominal surface recording of electromyographic activity in rats during pregnancy," *Am J Obstet Gynecol,* vol. 174, pp. 744-753, Feb 1996

[35] H. Maul, W. L. Maner, G. Olson, G. R. Saade, and R. E. Garfield, "Non-invasive transabdominal uterine electromyography correlates with the strength of intrauterine pressure and is predictive of labor and delivery," *J Matern Fetal Neonatal Med,* vol. 15, pp. 297-301, May 2004.

[36] J. Zietek, K. Horoba, J. Jezewski, A. Matonia, J. Sikora, and T. Kupka, "Analysis of Bioelectrical Uterine Activity for Detection of Threatening Premature Labour," in *14th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics,* vol. 20, A. Katashev, Y. Dekhtyar, and J. Spigulis, Eds.: Springer Berlin Heidelberg, 2008, pp. 469- 472.

[37] S. Q. Shi, W. L. Maner, L. B. Mackay, and R. E. Garfield, "Identification of term and preterm labor in rats using artificial neural networks on uterine electromyography signals," *Am J Obstet Gynecol,* vol. 198, pp. 235 e231-234, Feb 2008.

[38] C. Buhimschi, M. B. Boyle, G. R. Saade, and R. E. Garfield, "Uterine activity during pregnancy and labor assessed by simultaneous recordings from the myometrium and abdominal surface in the rat," *Am. J. Obstet. Gynecol.*, vol. 178, no. 4, pp. 811–822, 1998.

[39] W. L. Maner, R. E. Garfield, H. Maul, G. Olson, and G. Saade, "Predicting term and preterm delivery with transabdominal uterine electromyography," *Obstet. Gynecol.*, vol. 101, no. 6, pp. 1254–1260, 2003.

[40] R. E. Garfield, W. L. Maner, H. Maul, and G. R. Saade, "Use of uterine EMG and cervical LIF in monitoring pregnant patients," *BJOG,* vol. 112 Suppl 1, pp. 103-108, Mar 2005.

[41] H. Leman, C. Marque, and J. Gondry, "Use of the electrohysterogram signal for characterization of contractions during pregnancy," *IEEE Trans Biomed Eng,* vol. 46, pp. 1222-1229, Oct 1999.

[42] M. Doret, R. Bukowski, M. Longo, H. Maul, W. L. Maner, R. E. Garfield, and G. R. Saade, "Uterine electromyography characteristics for early diagnosis of mifepristone-induced preterm labor," *Obstet. Gynecol.*, vol. 105, no. 4, pp. 822–830, 2005.

[43] M. P. Vinken, C. Rabotti, M. Mischi, and S. G. Oei, "Accuracy of frequency-related parameters of the electrohysterogram for predicting preterm delivery: a review of the literature," *Obstet. Gynecol. Surv.*, vol. 64, no. 8, pp. 529–541, 2009.

[44] J. Sikora, A. Matonia, R. Czabanski, K. Horoba, J. Jezewski, and T. Kupka, "Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals," *Arch. Perinat. Med.*, vol. 17, no. 2, pp. 97–103, 2011.

[45] M. Lucovnik, W. L. Maner, L. R. Chambliss, R. Blumrick, J. Balducci, Z. Novak-Antolic, and R. E. Garfield, "Noninvasive uterine electromyography for prediction of preterm delivery," *Am. J. Obstet. Gynecol*., vol. 204, no. 3, pp. 228.e1–228.e10, Mar. 2011.

[46] J. Terrien, T. Steingrimsdottir , C. Marque, B. Karlsson, "Synchronization between EMG at different uterine locations investigated using time-frequency ridge reconstruction: comparison of pregnancy and labor contractions," *EURASIP J. Adv. Signal Process.*, vol. 2010, 2010.

[47] G. Fele-Žorž, G. Kavšek, Ž. Novak-Antolič, and F. Jager, "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups," *Med. Biol. Eng. Comput.*, vol. 46, no. 9, pp. 911–922, 2008.

[48] B. Moslem, B. Karlsson, M. O. Diab, M. Khalil, and C. Marque, "Classification performance of the frequency-related parameters derived from uterine EMG signals," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 3371–3374.

[49] C. Buhimschi, M. B. Boyle, and R. E. Garfield, "Electrical activity of the human uterus during pregnancy as recorded from the abdominal surface.," *Obstet. Gynecol.*, vol. 90, no. 1, pp. 102–111, 1997.

[50] S. Arora and G. Garg, "A Novel Scheme to Classify EHG Signal for Term and Pre-term Pregnancy Analysis," *Int. J. Comput. Appl.*, vol. 51, no. 18, pp. 37–41, 2012.

[51] M. O. Diab, C. Marque, and M. A. Khalil, "Classification for uterine EMG signals: Comparison between AR model and statistical classification method," *Int. J. Comput. Cognit*, vol. 5, no. 1, 2007.

[52] N. Lu, J. Wang, I. McDermott, S. Thornton, M. Vatish, and H. Randeva, "Uterine electromyography signal feature extraction and classification," *Int. J. Model. Identif. Control*, vol. 6, no. 2, pp. 136–146, Jan. 2009.

[53] B. Moslem, M. Khalil, C. Marque, and M. O. Diab, "Energy distribution analysis of uterine electromyography signals," *J. Med. Biol. Eng.*, vol. 30, no. 6, pp. 361–366, 2010.

[54] M. O. Diab, B. Moslem, M. Khalil, and C. Marque, "Classification of uterine EMG signals by using Normalized Wavelet Packet Energy," in *Electrotechnical Conference (MELECON), 2012 16th IEEE Mediterranean*, 2012, pp. 335–338.

[55] D. Radomski, A. Grzanka, S. Graczyk, and A. Przelaskowski, "Assessment of Uterine Contractile Activity during a Pregnancy Based on a Nonlinear Analysis of the Uterine Electromyographic Signal," *Inf. Technol. Biomed.*, pp. 325–331, 2008.

[56] J. Vrhovec, "Evaluating the progress of the labour with sample entropy calculated from the uterine EMG activity," *Elektrotehniski Vestn.-Electrotech. Rev.*, vol. 76, no. 4, pp. 165–170, 2009.

[57] T. Ivancevic, L. Jain, J. Pattison, A. Hariz, and others, "Preterm birth analysis using nonlinear methods," vol. 1, no. 3, pp. 160–170, 2008.

[58] B. Moslem, M. Khalil, M. O. Diab, C. Marque, and C. Bioingénierie, "Detrended fluctuation analysis of uterine electromyography," in *First Middle East Conference on Biomedical Engineering, MECBME11*, Sharjah, UAE , 2011.

[59] A. Diab, M. Hassan, C. Marque, and B. Karlsson, "Quantitative performance analysis of four methods of evaluating signal nonlinearity: Application to uterine EMG signals," Presented at the 34th Annual International IEEE EMBS Conference. San Diego, USA, 2012.

[60] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, "Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings," in *Conference proceedings:... 32nd Annual InternationalConference of the IEEE Engineering in Medicine and Biology Society,* Buenos Aires, 2010, vol. 2010, p. 4642.

[61] M. Hassan, J. Terrien, C. Muszynski, A. Alexandersson, C. Marque, and B. Karlsson, "Better Pregnancy Monitoring Using Nonlinear Correlation Analysis of External Uterine Electromyography," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 4, pp. 1160–1166, 2013.

[62] A. Diab, "Study of The Nonlinear Properties And Propagation Characteristics Of The Uterine Electrical Activity During Pregnancy And Labor", Ph.D. dissertation, Thèse de l'université de Technologie de Compiègne, 2014.

[63] P. Fergus, P. Cheung, A. Hussain, D. Al-Jumeily, C. Dobbins, and S. Iram, "Prediction of Preterm Deliveries from EHG Signals Using Machine Learning," *PLoS ONE*, vol. 8, no. 10, p. e77154, Oct. 2013.

[64] S. M. S. Baghamoradi, M. Naji, and H. Aryadoost, "Evaluation of cepstral analysis of EHG signals to prediction of preterm labor," in *Biomedical Engineering (ICBME), 2011 18th Iranian Conference of*, 2011, pp. 81–83.

[65] J. Terrien, S. Rihana, J. Gondry, and C. Marque, "Modeling the effects of the electrodes position on the surface EMG characteristics," *Model. Control Biomed. Syst. 2006*, p. 171, 2006.

[66] C. Rabotti, M. Mischi, J. O. E. H. van Laar, G. S. Oei, and J. W. M. Bergmans, "Estimation of internal uterine pressure by joint amplitude and frequency analysis of electrohysterographic signals," *Physiol. Meas.*, vol. 29, p. 829, 2008.

[67] S. Graczyk, J. Jezewski, J. Wrobel, and A. Gacek, "Abdominal electrohysterogram data acquisition problems and their source of origin," in *, IEEE Engineering in Medicine and Biology Society, 1995 and 14th Conference of the Biomedical Engineering Society of India. An International Meeting, Proceedings of the First Regional Conference*, 1995, pp. PS13–PS14.

[68] J. Terrien, C. Marque, J. Gondry, T. Steingrimsdottir, and B. Karlsson, "Uterine electromyogram database and processing function interface: An open standard analysis platform for electrohysterogram signals," *Comput. Biol. Med.*, vol. 40, no. 2, pp. 223–230, Feb. 2010.

[69] B. Karlsson, J. Terrien, V. Gudmundsson, T. Steingrimsdottir, and C. Marque, "Abdominal EHG on a 4 by 4 grid: mapping and presenting the propagation of uterine contractions," in *11th Mediterranean Conference on Medical and Biomedical Engineering and Computing 2007*, pp. 139–143., 2007.

[70] C. Rabotti, M. Mischi, S. G. Oei, and J. W. Bergmans, "Noninvasive estimation of the electrohysterographic action-potential conduction velocity," *Biomed. Eng. IEEE Trans. On*, vol. 57, no. 9, pp. 2178–2187, 2010.

[71] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, "A multisensor data fusion approach for improving the classification accuracy of uterine EMG signals," in *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 93–96.

[72] B. Moslem, M. O. Diab, C. Marque, and M. Khalil, "Classification of multichannel uterine EMG signals," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*,pp. 2602–2605, 2011.

[73] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, "Classification of multichannel uterine EMG signals by using unsupervised competitive learning," in *Signal Processing Systems (SiPS), 2011 IEEE Workshop on*, 2011, pp. 267–272.

[74] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, “Combining multiple support vector machines for boosting the classification accuracy of uterine EMG signals,” in *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 631–634.

[75] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, “Classification of multichannel uterine EMG signals using a reduced number of channels,” in *Mechatronics and its Applications (ISMA), 2012 8th International Symposium on*, pp. 1–4, 2012.

[76] B. Moslem, M. Diab, M. Khalil, and C. Marque, “Combining data fusion with multiresolution analysis for improving the classification accuracy of uterine EMG signals,” *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, pp. 1–9, Aug. 2012.

[77] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, “Combination of Canonical Correlation Analysis and Empirical Mode Decomposition Applied to Denoising the Labor Electrohysterogram,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2441–2447, Sep. 2011.

Chapter 2: Feature selection for classification of Electrohysterograms

2.1 Introduction

One of the most common ways to detect preterm labor is to extract features from EHG signals in order to find information pertinent for this detection. Numerous types of features have been extracted from the electrohysterogram (EHG) on different studies. Using temporal, linear and nonlinear parameters, the number of available features becomes very large. Additionally, the EHG databases are recorded from different populations and with different recording protocols. It is sometimes difficult to compare their results and to select the best features, before applying a classification between Labor and Pregnancy. The selection of pertinent features is a very important problem in pattern recognition. Therefore, we need reliable feature selection methods to reduce the number of features through the elimination of irrelevant and noisy features. For this purpose, in this chapter several methods for feature selection are presented, proposed and tested. These features are divided into three types: linear, nonlinear features and features related to the EHG propagation. We will firstly give a background on feature selection techniques based on “Filter” and “Wrapper” methods. Then, two other parts will be presented in this chapter. In the first part, we will propose a selection method based on the measurement of histogram similarity or dissimilarity, for a given feature, between the two classes (pregnancy and labor) using the Jeffrey divergence method. In the second part, several methods for feature selection will be implemented to choose the best subsets for the classification of labor and pregnancy contractions. The percentage of correct classification for each subset is calculated for result comparison. The comparison of the results from all these feature selection methods will be done by using monopolar and then bipolar EHG signals, in order to compare the recording configuration influence.

2.2 Feature selection methods

2.2.1 Introduction

As computer databases and technologies are moving fast, we need a high capacity in computers and high level programming tools to analyze data [1]. In classification techniques, the huge number of features causes problems in pattern recognition. One of the most important and used techniques in data preprocessing for data mining is "Feature Selection" [2, 3].

Feature selection is generally defined as a search process that is used to find a "relevant" subset of features F_{Best} from those of the starting subset of features F (Figure 2.1). The notion of relevance of a chosen subset of features always depends on the objective function and system criteria. The objective function evaluates candidate subsets and returns a measure of their "goodness". This measure is used by the search strategy to select new candidates.

Assume that $F = \left\{ f_1, f_2, \dots, f_i, \dots, f_m \right\}$ is a set of features of size m, where m represents the total number of original features and $i = 1 \dots m$. J is an evaluation function of a chosen subset of features. We assume that the greatest value of J will be obtained for the best subset of features. The aim of the selection is to find a subset $F_{Best} = \left\{ f_{s_1}, f_{s_2}, \dots, f_{s_j}, \dots, f_{s_P} \right\}$ of size P ($P < m$), $s_j \in \{1 \dots m\}$, and $F_{Best} \subseteq F$ such as :

$$J(F_{Best}) = \text{Max } J(SC) \tag{1}$$

where SC is a candidate subset of features, $SC \subseteq F$ and the size of SC can be a number, either already defined by the user, or controlled by one of the generation methods of subsets which will be described in the next section.

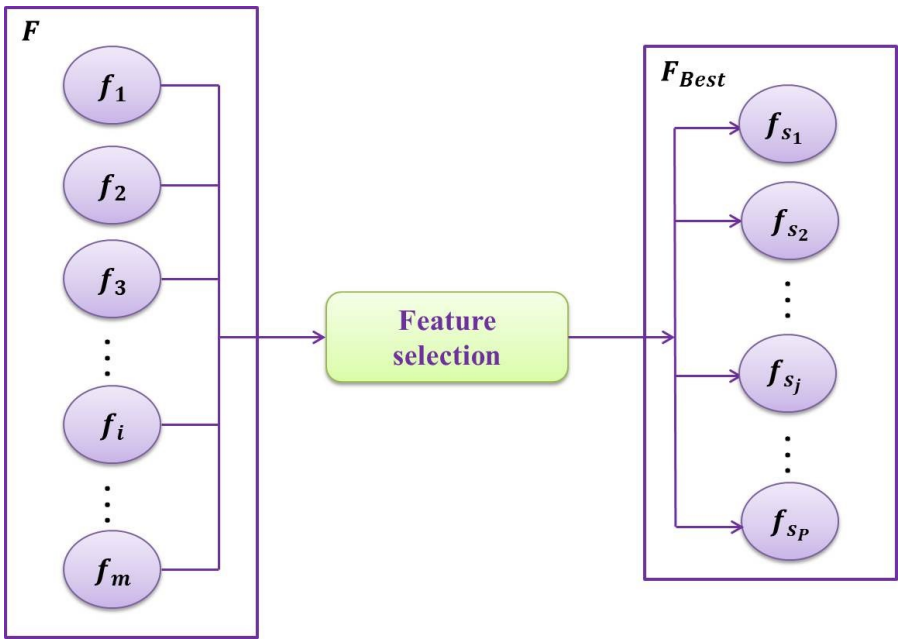


Figure 2.1: Feature Selection technique

Therefore, feature selection has several advantages. It reduces the number of features, removes irrelevant, redundant, or noisy features. Additionally, it can significantly improve the mining performances of learning algorithms such as the speed of learning and the predictive accuracy.

2.2.2 Feature selection process

A general procedure proposed by [4] for a feature selection method is shown Figure 2.2. In this typical feature selection algorithm there are four steps: subset generation, subset evaluation, stopping criterion, and result validation.

Subset generation is a research strategy [5] that is used to select candidate feature subsets for evaluation. Then, an evaluation criterion measures the goodness of this candidate subset, and then compares it with the previous best one in order to determine whether this subset is suitable or not. If the new candidate subset is found to be better, it replaces the previous best subset. These two processes are repeated until a given stopping criterion is satisfied. Finally, a validation part is used in order to check the validity of the selected best subset [1].

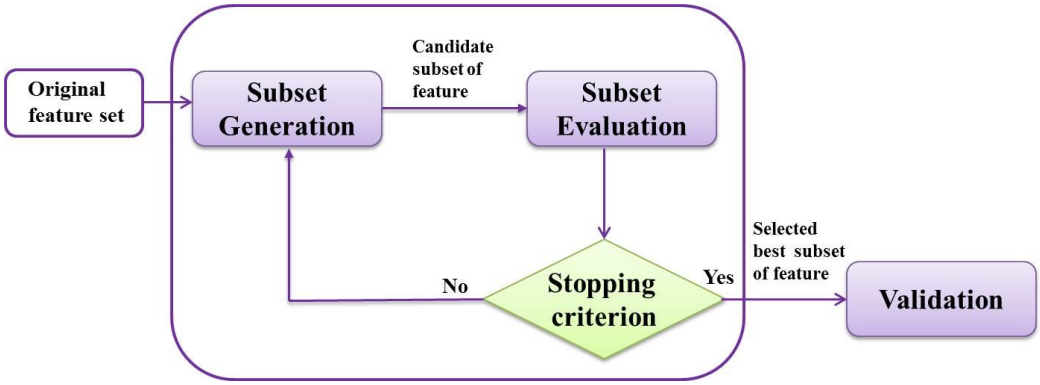


Figure 2.2: Feature Selection Process (adapted from [4]).

2.2.2.1 Search procedures

The subset generation step in feature selection process usually falls into three categories [1]:

- *Exhaustive or complete search:* examination of all combinations of feature subsets is performed to select the "best" subset of features. In this case, the optimum subset is selected. This method requires too much computing and time. The major problem with this approach is that the number of combinations increases exponentially with the number of features.
- *Heuristic search:* selection is directed under certain guidelines. The algorithms that use heuristic search are generally iterative algorithms. Features are iteratively added and removed using heuristic search. Therefore, search can start with an empty set and successively add features (i.e., sequential forward selection), or start with a full set and successively remove features (i.e., sequential backward selection), or start with both ends and add and remove features simultaneously (i.e., bidirectional selection). Algorithms based on heuristic search are simple to implement and fast in producing results but they can complete the process with some risks of losing optimal subsets.
- *Random search:* The algorithm used in this search starts with a randomly selected subset and may generate the next subset by two ways: one based on sequential search and the other based on the generation of a subset in a random manner. The use of random research helps to avoid local optima in the search space.

2.2.2.2 Evaluation

Evaluation is an important part in feature selection. It is used to determine which subset gives the optimal solution. We can distinguish 3 categories for evaluation:

- *Filter method*: This technique is generally used for the selection of features. It evaluates the relevance of features only by examining their intrinsic properties [6,7]. It is considered as a preprocessing step (filtering) before the learning process, this means that the evaluation of this technique is independent from the classifier [8]. In the majority of cases, a feature relevance score is calculated, and low-scoring features are removed [9]. The best subset of features obtained by this technique is presented as input to the classification algorithm [7]. The procedure of the filter model is illustrated Figure 2.3.

The most important measures used in the literature as score or evaluation criteria are [1]:

- Distance measures: for a two-class classification problem, a feature *A* is more scored than another feature *B* if *A* leads to a greater difference between the conditional probabilities of the two classes than *B* [1]. Euclidean or other distances can be used here.
- Information measures or entropy: these measurements determine the information gain of a given feature. The Feature *A* is scored more than *B* if the information gain is greater for *A* than for *B* [1].
- Measures of dependences or correlations: this describes the ability to predict the value of a variable from another. In feature selection for classification, we seek to know the strength of feature associated with the class [1]. If the correlation between the feature A with class C is greater than the correlation of the feature B with C, then the feature A is preferable to the feature B [1].

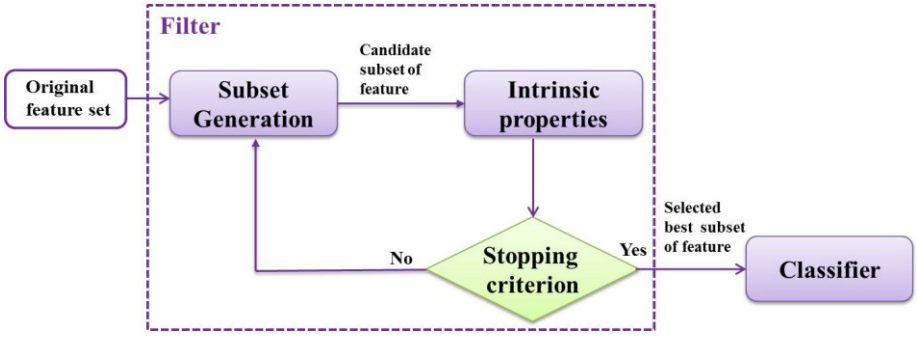


Figure 2.3: Filter based feature Selection Process

- *Wrapper method:* The main disadvantage of the filter methods is that they ignore the influence of the selected feature on the performance of the classifier. In [10], Kohavi and John proposed the wrapper method to solve this problem. This technique evaluates a subset of features after using the classification performances (learning and testing) [11]. It searches for a good subset using a learning algorithm. The procedure of the wrapper model is illustrated Figure 2.4.
- *Embedded method:* This technique achieves the selection process inside the induction algorithm [6]. The search for an optimal subset of features is included in the classifier construction and can be regarded as a search in the combined space of feature subsets and assumptions [7].

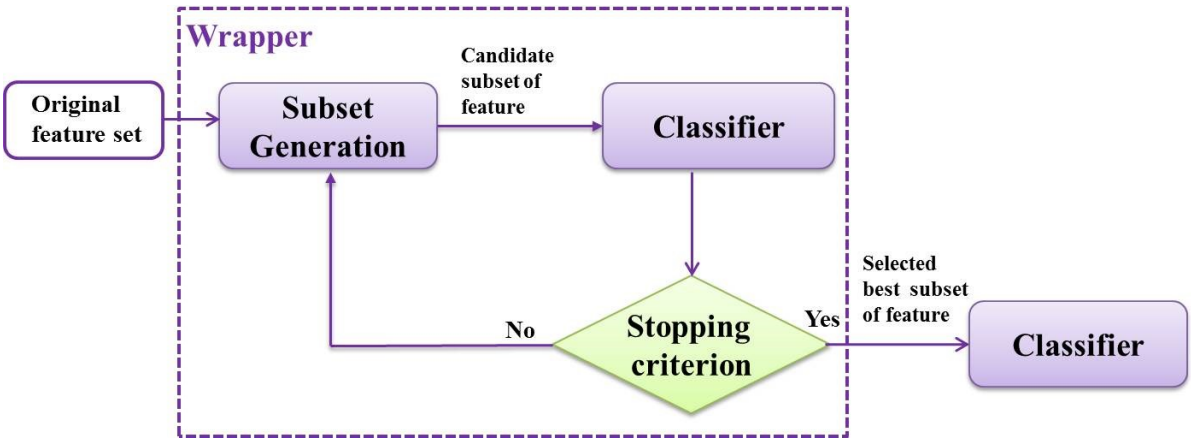


Figure 2.4: Wrapper based feature Selection Process

In our work we use several wrapper and filter methods. Table 2.1 presents the advantages and disadvantages of these two techniques.

2.2.2.3 Stopping Criteria

Some criteria must be defined to stop the search process. Some frequently used stopping criteria are [1, 12]:

- The search is completed
- A threshold is reached, like minimum number of features or maximum number of iterations.
- There is no more precision improvement, in other words, when there is no possibility of finding a subset better than the current subset.

Some stopping criteria commonly used are based on the order of features. They are classified according to some relevance scores. Those having the highest scores will be selected and used by a classifier (filter method).

Table 2.1: Advantage and disadvantage of filter and wrapper methods [11]

Method	**Advantage**	**Disadvantage**
Filter	• Fast execution • Lower computational cost	• No interaction with a classifier
Wrapper	• Interaction with a classifier	• High computational cost • Slow execution • Risk of overfitting

2.2.3 Feature selection methods selected from the literature

The filter and wrapper methods used in our work are presented below:

Four filter methods:

- Jeffrey divergence (JD)
- F-score
- Relieff
- Mutual information based on clustering

Seven wrapper methods:

- Sequential Forward Selection (SFS)
- Sequential backward Selection (SBS)
- Plus-l minus-r selection (LRS)
- Bidirectional search (BDS)
- Sequential Forward Floating sequential (SFFS)
- Genetic Algorithm (GA)
- Binary particle swarm optimization (BPSO)

Filter methods are subdivided in two categories (Table 2.2): Univariate filtering methods and multivariate filtering methods. The difference between these two categories is the relationship

between features. Univariate filtering methods consist of selecting features by adding, deleting or comparing one feature at each iteration. Therefore, the selection of each feature is independent from the other selected ones. On the other hand, multivariate filtering methods have been introduced using the independency between features. Therefore in this method, we consider all features simultaneously [7, 11, 13].

Wrapper method can be divided into two categories (Table 2.3) according to the search method. These two categories are: Deterministic search algorithm and, randomized search algorithms [7].

Table 2.2: Filter methods

Filter method	
Univariate filtering methods	**Multivariate filtering methods**
- Jeffrey divergence (JD) - F-score	- Relieff - Mutual information based on clustering (MI)

Table 2.3: Wrapper methods

Wrapper method	
Deterministic	**Randomized**
- Sequential Forward Selection (SFS) - Sequential backward Selection (SBS) - Plus-l minus-r selection (LRS) - Bidirectional search (BDS) - Sequential Forward Floating sequential (SFFS)	- Genetic Algorithm (GA) - Particle swarm optimization (BPSO)

2.2.3.1 Filter method

A- Jeffrey divergence (JD)

Histogram is an important tool that represents the statistical characteristics of data. The histogram is used to represent the distribution of each feature extracted from the signal to be classified. To compare two histograms, a well-known technique named "Jeffrey divergence" has been developed to compute the similarity for classification [14]. Therefore, we proposed to measure similarity or dissimilarity for a given feature between the two histograms of the two classes using the Jeffrey divergence method. This method consists of measuring the bin-by-bin distances of two corresponding histograms. The Jeffrey divergence is the symmetric form of the Kullback–Leibler divergence, one of the classical and efficient tools for histogram distance estimation [15]. This Jeffrey divergence method is presented by equation 2 [16]:

$$D_{Je}(H,G) = \sum_{z} \left(h_z \log \frac{h_z}{g_z} + g_z \log \frac{g_z}{h_z} \right) \quad (2)$$

H and G are the two histograms of a given feature in class 1 and class 2. If B are the bins (we used B=10 in our work, we can define H=$\{h_z\}$ and G=$\{g_z\}$ *as* two histograms with the bin index $z \in \{1,2,\ldots,B\}$.

Therefore, this distance will be then used to select the discriminant features. A lower distance means a larger similarity, and a larger distance means a lower similarity [16]. Indeed, the greater the distance between the feature histograms of the two classes, the most discriminant should be the given feature.

Algorithm 2.1 presents the algorithm of the Jeffrey divergence method. The input of this method is the histogram for each feature and the different classes. H presents the histograms of m features in class 1 and G presents the histograms of m features in class 2. The output of this algorithm $Fbest$ is the subset of features that have the greater distances. In our work, we selected the features with distance value higher than mean value of all distances (Threshold).

Algorithm 2.1: *Jeffrey divergence (JD)*

1. Input:

$F = \{f_1, f_2, \dots, f_i, \dots, f_m\}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{y_1, \dots, y_O\}$, *C* is *vector of classes corresponding to each observation*

$H = \{h_{z1}, \dots, h_{zi}, \dots, h_{zm}\}$, *H :histogram of features of classe 1*

$G = \{g_{z1}, \dots, g_{zi}, \dots, g_{zm}\}$, *G :histogram of features of classe 2*

$z \in \{1,2, \dots, B\}$, *B is the number of bins*

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3. Initialization :

Start with an empty set of features $F_{Best} = \emptyset$

4. for i=1 :m

Compute $D_{Jei}(h_{zi}, g_{zi})$ according to equation (2)

end for

$D_{Je} = \{D_{Je1}, \dots, D_{Jei}, \dots, D_{Jem}\}$

for i=1 :m

  if $D_{Jei} > Threshold$

  Add features f_i to $Fbest$

  end if

end for

4. Return $Fbest$

B - F-score

F-score is a novel filter method which calculates the discriminative ability of each feature [8, 17]. The value of F-score $Fscore(i)$ is calculated for each feature using the following equation:

$$Fscore(i) \equiv \frac{\left(\bar{f}_i^{(+)} - \bar{f}_i\right)^2 + \left(\bar{f}_i^{(-)} - \bar{f}_i\right)^2}{\frac{1}{n_+ - 1}\sum_{l=1}^{n_+}\left(f_{l,i}^{(+)} - \bar{f}_i^{(+)}\right)^2 + \frac{1}{n_- - 1}\sum_{l=1}^{n_-}\left(f_{l,i}^{(-)} - \bar{f}_i^{(-)}\right)^2} \quad (3)$$

where $\bar{f}_i$, $\bar{f}_i^{(+)}$ and $\bar{f}_i^{(-)}$ are the average of the i^{th} feature of the whole, positive (class 1) and negative (class 2) datasets. $f_{l,i}^{(+)}$ and $f_{l,i}^{(-)}$ are the i^{th} feature of the lth positive and negative instance respectively. Features with higher F-score value are selected. In our work, we selected the features with a F-score value higher than the mean value of all F-score values (threshold) [8]. The pseudo code of this method is given in algorithm 2.2.

Algorithm 2.2: *F-score*

1. Input:

$F = \{f_1, f_2, \dots, f_i, \dots, f_m\}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{y_1, \dots, y_O\}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3. Initialization:

Start with an empty set of features $F_{Best} = \emptyset$

4. for i=1 :m

Compute $Fscore(i)$ using equation (3)

end for

$Fscore = \{Fscore(1), \dots, Fscore(i), \dots, Fscore(m)\}$

for i=1 :m

  if $Fscore(i) > Threshold$

    Add features f_i to $Fbest$

  end if

end for

5. Return $Fbest$

C - Relief

Kira et al [18] developed an algorithm named Relief method. This method can estimate the quality of features through a type of nearest neighbor algorithm that selects neighbors (instances) from the same class and from the different classes. From the whole instances, a random sample (*NS*) is firstly chosen. An instance l_j is selected randomly. "Near Hit" and "Near Miss" of the instance l_j are calculated using the measure of euclidean distance. "Near Hit (*H*)" is the nearest instance that belongs to the same class of the selected instance (having the minimum Euclidean distance with this instance). "Near Miss (*M*)" is the nearest instance that belongs to a different class of the selected instance. Therefore, weights (*W*) for each feature (f_i) are estimated based on these Near Hits, and the Near Misses according to the following equation:

$$W_i = W_i - \frac{diff(f_{l,i}, f_{H,i})^2}{NS} + \frac{diff(f_{l,i}, f_{M,i})^2}{NS} \quad (4)$$

Where:

- $f_{l,i}$ is the value of the feature f_i for the instance l
- $f_{H,i}, f_{M,i}$ are the neighborhoods H and M for the feature f_i

- *If* f_i is :
  - nominal, $diff(f_{l,i}, I_{i,t}) = \begin{cases} 0 \; if \; value \; of \; f_{l,i} \; and \; I_{i,t} \; are \; equal \\ 1 \; if \; the \; values \; are \; different \end{cases}$
  - Numerical, $diff(f_{l,i}, I_{i,t}) = \frac{Valeur(f_{l,i}) - Valeur(I_{i,t})}{max\ (f_i) - min\ (f_i)}$,

    $max\ (f_i)$ is the maximum value of the feature f_i

    $Valeur(f_{l,i})$ is the value of the feature f_i for the instance l.

This adjustment process of weights is repeated for *NS* instances. The Relief algorithm produces weights for each feature, and selected only the features having weights greater than or equal to a threshold. The Relief pseudo code is presented below in algorithm 2.3.

Algorithm 2.3: *Relief*

1. Input:

$F = \{f_1, f_2, \dots, f_i, \dots, f_m\}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{y_1, \dots, y_O\}$, *C* is *vector of classes corresponding to each observation*

NS =Number of sample by default equal to O

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m),\ s_j \in \{1 \dots m\}$

3. Initialization :

Start with an empty set of features $F_{Best} = \emptyset$

Initialize all weights W_i =0

4. for j=1 :NS

Randomly selecting an instance (observation) l_j of the set of observations

Find their Near Hit *(H)* and Near Miss *(M)*

for i= 1 :m

Compute $W_i$using equation (4)

end for

end for

$W = \{W_1, \dots, W_i, \dots, W_m\}$

for i= 1 : m

if $W_i > Threshold$

Add features f_i to $Fbest$

end if

end for

5. Return $Fbest$

An improved Relief method has been proposed by Kononenko [19], which consists of choosing K nearest neighbors instead of one. This new algorithm named "Relieff" has been shown to be more robust to noisy features [19, 20, 21]. The pseudo code of this method is presented in Algorithm 2.4. When obtained the weight for each feature, we select in our work the features with weight value larger than the average value of all weights.

Algorithm 2.4: *Relieff*

1. Input

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

NS =Number of sample by default equal to O

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3. Initialization:

Start with an empty set of features $F_{Best} = \emptyset$

Initialize all weights W_i =0

4. for j=1 :NS

Randomly selecting an instance l_j of the set of observations

Find their q Near Hit $H = \{ H_1, H_2, \dots, H_t, \dots, H_q \}$

and their q Near Miss $M = \{ M_1, M_2, \dots, M_t, \dots, M_q \}$

for i= 1 :m

$$W_i = W_i - \sum_{t=1}^{q} \frac{diff\left(f_{l,i}, f_{H_t,i}\right)^2}{NS * q} + \sum_{C \neq class(l_j)} \frac{\frac{P(C)}{1 - P\left(Class(l_j)\right)} \sum_{t=1}^{q} \frac{diff\left(f_{l,i}, f_{M_t,i}\right)^2}{NS}}{NS * q}$$

% $P(C)$ is the probability of the class C which belongs to the "near miss" $f_{M_t,i}$

end for

end for

$W = \{W_1, \dots, W_i, \dots, W_m\}$

for i= 1 : m

if $W_i > Threshold$

Add features f_i to $Fbest$

end if

end for

5. Return $Fbest$

D – Mutual information based on clustering

Mutual information gives us an idea about the amount of information between two variables. The mutual information of two variables $X$and Y is given by:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y)\, log \frac{p(x,y)}{p(x)p(y)} \tag{5}$$

Mutual information is used in feature selection methods. It is a filter method that permits to know the quantity of information contained in the features that is related to classes. Let F be the initial set consisting of m features and C represents the classes. The goal is to find a subset $Fbest$ containing P features that maximizes the mutual information $I\ (Fbest, C)$.

In [22] Lui et al. proposed a new method of feature selection based on clustering. They used the mutual information $I\ (f; C_u)$ as a distance measurement between a candidate cluster f and the class cluster C_u. Therefore, the distance S_b between the selected cluster Y and class cluster C_u is represented as:

$$S_b\ (C_u, Y) = \sum I\ (y; C_u) \quad ,\ y \in Y \tag{6}$$

A maximum value of S_b gives only the best features that represent the classes. It is important to take into account the redundancy between the set of selected features Y and the candidate features f. Therefore the distance between f and Y ($S(f)$) is the sum of distance between f and all $y \in Y$:

$$S(f) = \sum_{y \in Y} CR\ (y, f) \tag{7}$$

Where $CR\ (y, f)$ is the coefficient of relevance.

$CR\ (y, f) = \frac{I(y;f)}{H(y)}$; $H(y)$ is the entropy

$$CR\ (y, f) = \begin{cases} 0\ means\ that\ y\ and\ f\ are\ independent \\ 1\ means\ a\ wide\ dependency \end{cases}$$

Likewise, the within cluster distance $S_w\ (Y)\ of\ Y$ can be obtained accumulatively:

$$S_w\ (Y \cup f) = S_w(Y) + S\ (f) \tag{8}$$

The size of Y is an important aspect in features selection. According to the explanation above, the evaluation function J for the candidate feature f is:

$$J(f) = \frac{S_b(C_u, Y) + I(f; C_u)}{|Y| + S_W(Y) + S(f)} \tag{9}$$

Where $|Y|$ is the number of features in Y.

The pseudo code of features selection based on clustering [22] is presented in algorithm 2.5.

Algorithm 2.5: Mutual information with clustering

1. Input

$F = \{f_1, f_2, \dots, f_i, \dots, f_m\}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{y_1, \dots, y_O\}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3. Initialization:

Start with an empty set of features ,$Y = \emptyset$, $F_{B\ est} = \emptyset$, $C_u = C$

4.for i= 1 :m

$S_{b,i}(C_u, f_i) = I(f_i; C)$

end for

choose f that have max $(I(f; C))$; $F = F \backslash \{f\}$; $Y = \{f\}$; $S_w = 0$; $J = \emptyset$

While $|Y| \leq m$

for each candidate feature $f \epsilon F$

Calculate $J(f)$

end for

choose f that have max $(J(f))$

$J = J \cup \{J(f)\}$

$Y = Y \cup \{f\}$; $F = F \backslash \{f\}$

$S_w = S_w + S(f)$

$S_b = S_b + I(f; C_u))$

end while

Compute variances of *J*

$Fbest$ is the subset corresponds to the first value of variance J < mean of all values of variance J between parameter additions operations

5. Return $Fbest$

The best subset of features selected by this method corresponds to a negligible velocity of *J,* where the velocity stabilizes close to a constant velocity between parameter additions operations i.e. between subsets. We use as threshold the mean of all values of variance *J*. The best subsets

corresponds thus to the first value of variance J < mean of all values of variance J between parameter additions operations.

2.2.3.2 Wrapper Method

A- Sequential Forward Selection (SFS)

Sequential forward selection (SFS) was proposed by [23]. It is a sequential search algorithm for feature selection [24] developed for data mining and classification. The pseudo code of this method is presented in algorithm 2.6 [25].

Algorithm 2.6: *SFS*

1.Input:

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{ f_{s1}, \dots, f_{sj}, \dots, f_{sP} \}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3.Initialization:

Start with an empty set of features $F_{Best} = \emptyset$, $Y = \emptyset$, $E = \emptyset$

4. While $|Y| \leq m$

$f_{add} = \underset{f \notin Y}{\text{argmax}}\ J(Y \cup f)$

$E = E \cup \{J(Y \cup f_{add})\}$

$Y = Y \cup f_{add}$

End while

F_{Best} is the subset corresponding to min(E)

5. Return $Fbest$

SFS begins with an empty subset. The value of the criterion function (J) is calculated for each feature by using a classifier. The feature presenting the best classification performance is selected (f_{add}) and then added to the subset (Y). The next step consists of adding sequentially the feature f_{add} that has the highest criterion function $J(Y \cup f_{add})$ when combined with the set of features Y that have already been selected. To avoid to precise a number of features as threshold, in our work this method is implemented starting from a single parameter subset, all the way to the full parameter set in order to plot and analyze the classification error along the complete procedure. Then we choose the combination of features that have the minimum classification error.

B- Sequential backward Selection (SBS)

Sequential backward Selection was a method proposed by [26]. SBS works in a way opposite to SFS. Starting from the full set of features (all features), at each iteration, the worst feature will be sequentially removed. The SBS algorithm is presented in the pseudo code 2.7. To avoid to precise a number of features as threshold, in our work this method is implemented using all combination (from one parameter to all parameters) subsets and then plot and analyze the classification error along the complete procedure. At the end, we choose the combination of features that have the minimum classification error.

Algorithm 2.7: *SBS*

1.Input:

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of features* , f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3.Initialization:

Start with an empty set of features $F_{Best} = \emptyset$, $Y = F$, $E = \emptyset$

4.While $|Y| \leq 1$

$f_{remove} = \underset{f \in Y}{\text{argmax}}\ J(Y \backslash f)$

$E = E \cup \{J(Y \backslash f_{remove})\}$

$Y = Y \backslash f_{remove}$

End while

F_{Best} is the subset which corresponds to min(E)

5.Return $Fbest$

C- Plus-l minus-r selection (LRS)

Plus-L minus-R search (LRS) [27] is a generalization and a combination of two methods SFS and SBS. LRS attempts to compensate the weaknesses of the SFS and SBS with some backtracking capabilities [28]. In this algorithm we use two values *L* and *R* respectively corresponding to the number of parameters to add (L) and remove (R).

If L>R, LRS starts with an empty set and adds L features. Then it moves towards the next step, where it removes R features. This procedure (Add L features and remove R features) is repeated to attain the maximum possible number of parameters and then plot and analyze the classification error of the full procedure instead of fixing a number of features as threshold.

If R>L, LRS starts with a full set and removes R features. Then it moves towards the next step, where it adds L features. This procedure (Removing R features and adding L features) is repeated to attain the minimum possible number of parameters and then plot and analyze the classification error of the full procedure instead of fixing a number of features as threshold.
Algorithm 2.8 presents the pseudo code of LRS.

Algorithm 2.8: *LRS*

1.Input:

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of feature,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{ f_{s1}, \dots, f_{sj}, \dots, f_{sP} \}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3.Initialization:

Start with an empty set of features $F_{Best} = \emptyset$, $E = \emptyset$

Initialize the two Numbers *L* and *R*

if *L* >R then

Start with an empty set *Y={Φ}* and go to step 4.1

else

Start with the full set $Y = F$ and go to step 4.2

4. While $|Y| \leq m$

1. Repeat *L* times

$f_{add} = \underset{f \notin Y}{\text{argmax}}\ J(Y \cup f)$

$Y = Y \cup f_{add}$

$E = E \cup \{J(Y \cup f_{add}\}$

2. Repeat *R* times

$f_{remove} = \underset{f \in Y}{\text{argmax}}\ J(Y \backslash f)$

$Y = Y \backslash f_{remove}$

$E = E \cup \{J(Y \backslash f_{remove})\}$

End while

F_{Best} is the subset which correspond to min(E)

5. Return $Fbest$

In our work we apply LRS method starting from an empty set and repeatedly adding three (L=3) and removing 2 features (R=2) [28]. Then we choose the best subset of selected features that corresponds to the minimum classification error based on the plot analysis.

D- Bidirectional search (BDS)

Bidirectional search BDS [27] is a parallel implementation of SFS and SBS. To guarantee that SFS and SBS converge to the same solution, BDS respects the two conditions:

- Features already selected by SFS are not removed by SBS
- Features already removed by SBS are not selected by SFS

The number of selected features must be fixed because BDS uses the two directions for searching [28]. In our work the number of selected features is equal to the half of the number of full set of features. This will allow plotting the classification error along the full possible number of operations for better visualization and analysis. The obtained plots are used to select the best SFS and SBS subsets with minimum classification error. The algorithm of BDS is outlined in the pseudo code 2.9.

Algorithm 2.9: *BDS*

1.Input:

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbestf = \{ f_{s1}, \dots, f_{sj}, \dots, f_{sP} \}$, *P: size of features selected* , $(P < m),\ s_j \in \{1 \dots m\}$

$Fbestb = \{ f_{s1}, \dots, f_{sj}, \dots, f_{sr} \}$, *r: size of features selected* , $(r < m),\ s_j \in \{1 \dots m\}$

3.Initialization:

Start with an empty set of features $F_{Best} = \emptyset$, $E_f = \emptyset$, $E_b = \emptyset$

Start SFS with $Y_F = \emptyset$

Start SBS with $Y_B = F$

4. While $|Y_F|\ and\ |Y_B| \leq m/2$

1. $f_{add} = \underset{\substack{f \notin Y_F \\ f \in Y_B}}{\operatorname{argmax}}\ J(Y_F \cup f)$

$Y_F = Y_F \cup f_{add}$

$E_f = E_f \cup \{J(Y \cup f_{add})\}$

2. $f_{remove} = \underset{\substack{f \notin Y_F \\ f \in Y_B}}{\operatorname{argmax}}\ J(Y_B \backslash f)$

$Y_B = Y_B \backslash f_{remove}$

$E_b = E_b \cup \{J(Y \backslash f_{remove})\}$

End while

F_{Bestf} is the subset which corresponds to $\min(E_f)$

F_{Bestb} is the subset which corresponds to $\min(E_b)$

5. Return F_{Bestf} , F_{Bestb}

E- Sequential Forward Floating sequential (SFFS)

The SFFS method is introduced by Pudil et al. [29]. It is an extension to Plus-L minus-R selection (LRS) with flexible backtracking abilities. This method begins by an empty set, and then a subset is generated by adding a feature. After this forward step, a conditional deletion step is examined. Therefore, SFFS performs backward step, and conditionally permanently removes the least significant feature from the current subset if and only if the deletion of this feature improves the objective function. We will repeat this step as long as an improvement occurs, otherwise it puts back the last removed feature and loop back to the forward step [30, 31]. The method requires a desired number of features used as threshold in the input of the algorithm. In our work we specified the maximum number of features that corresponds to the full features set size. This will allow plotting the classification error along the full possible number of operations for better visualization and analysis. The resulting plots are used to select the best subset with minimum classification error. The pseudo code of SFFS is presented in Algorithm 2.10.

Algorithm 2.10: *SFFS*

1.Input:

$F = \{f_1, f_2, \dots, f_i, \dots, f_m\}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{y_1, \dots, y_O\}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{f_{s1}, \dots, f_{sj}, \dots, f_{sP}\}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \dots m\}$

3.Initialization:

Start with an empty set of features $F_{Best} = \emptyset$, $Y = \emptyset$, $E = \emptyset$

4. While $|Y| \leq m$

1. $f_{add} = \underset{\substack{f \notin Y \\ f \in F}}{\operatorname{argmax}} J(Y_F \cup f)$

$Y = Y \cup f_{add}$

$E = E \cup \{J(Y \cup f_{add})\}$

2. $f_{remove} = \underset{f \in Y}{\operatorname{argmax}} J(Y \backslash f)$

If $J(Y \backslash f_{remove}) > J(Y)$ then

$Y = (Y \backslash f_{remove})$

$E = E \cup \{J(Y \backslash f_{remove})\}$

$F = F \backslash f_{remove}$

Go to step 2

else

Go to step 1

End while

F_{Best} is the subset which corresponds to $\min(E)$

5. Return $Fbest$

F- Genetic Algorithm

Genetic Algorithms (GA) can be specified as population-based and algorithmic search heuristic methods that mimic the natural biological evolution [28, 32]. GA is an iterative algorithm based on the reproduction and the evolution of natural individuals. Individuals (“chromosomes “) representing a better solution to the target problem are given more chances to "reproduce" than the chromosomes associated to a poorer solution [33]. Therefore, with GA we try to optimize a given function (objective function) in a search space of individuals. For that, we define an evaluation function (fitness) related to the objective function and applied to each individual, or chromosome.

GA is based on several steps. First, an initial population (N chromosomes) is created randomly and evaluated using the fitness function. At each generation, the algorithm selects a group of individuals in the current population, called parents. Therefore, the new population of chromosomes (solution candidates) is created by the process of selecting individuals according to their level of fitness combined with genetic functional, such as crossover and mutation. GA stops when the condition of the problem is satisfied (Algorithm 2.11).

GA creates three types of children for the next generation (population):

- *Elite children:* are the chromosomes in the current generation with the best fitness values. These chromosomes automatically survive to the next generation.
- *Crossover children*: are created by combining the vectors of a pair of parents. Crossing can build an individual (or a solution) which is the mixing of several solutions. In cross over, the chromosomes exchange gene sequences between them.
- *Mutation children:* are created by introducing random changes, or mutations, of the value of a gene in a single parent.

GA has been used in the field of feature selection [28, 34]. Chromosomes are the bit strings (defined below) and a gene is the feature. The length of each chromosome is determined by the total number of features. A gene value “1” means that the feature is selected. If it is ’0’, the feature is not selected. Table 2.4 presents the value of each parameter used for GA in our work.

Algorithm 2.11: *GA*

1. Input:

$F = \{ f_1, f_2, \ldots, f_i, \ldots, f_m \}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \ldots, y_O \}$, *C* is *vector of classes corresponding to each observation*

Output:

$Fbest = \{ f_{s1}, \ldots, f_{sj}, \ldots, f_{sP} \}$, *P: size of features selected* , $(P < m)$, $s_j \in \{1 \ldots m\}$

3.Initialization :

Number of chromosome *N*

Length of each chromosome equal to total number of features *m*

Initialize bit strings of all chromosome randomly

Set the number of generation *K* and other *GA parameters*

4.While number of generation is less than *K*

1. Compute the fitness value of each chromosome

   Compute mean fitness of all fitness value in the generation

   Select best fitness in the generation

   Select best chromosome in the generation

2. Create new population by applying genetic operators :

   a-Performs the selection process of chromosome (based on their fitness)

   -Use the cross over operator

   -Use the Mutation operator

   b- Select Elite children

   c-Replaces the current population with the children to form the next generation

3. Number of generation =Number of generation+1

End while

5. When the limit number of generation *K* is reached, we obtain an optimal solution, the best chromosome $Fbest$ (best subset of feature selection)

Table 2.4: Parameters used in GA [34]

GA parameters	**Value**
Population size	100
Length of chromosome	total number of features
Population type	Bitstrings
Number of generations	100
Crossover	Arithmetic Crossover
Crossover Probability	0.8
Mutation	Uniform Mutation
Mutation Probability	0.1
Selection process	Tournament of size 2
EliteCount	2
Fitness Function	KNN (Percentages of correct classification)

G- Binary particle swarm optimization

Particle Swarm Optimization (PSO) was developed by Eberhart and Kennedy in 1995. It is a population-based stochastic optimization technique that was inspired by the social behavior of bird flocking or fish schooling [35]. PSO uses a number of particles (the swarm) moving around in the search space in order to achieve the best solution. At each iteration, each particle tries to search for the best position in a multidimensional space and adjust its position according to its own experience and the experiences of its particles neighbors. We assume that our search space is n-dimensional and that each particle is a point in this space. The position of the i^{th} particle of the swarm is represented as $X_i = (x_{i1},...x_{id},... x_{in})$. Each particle has a best previous position $pbest_i=(p_{i1},...,p_{id},... p_{i,n})$, which corresponds of the best fitness value (in our case best classification given by a classifier fed with the selected features). The global best particle among all the particles in the population is represented by $gbest = (p_{g1},...,p_{gd},...,p_{gn})$. The velocity (the rate of the position change) of the i^{th} particle is denoted by $V_i= (v_{i1},...v_{id},...,v_{in})$. The particles velocity and position are manipulated according to the following two equations:

$$v_{id}^{k+1} = w\, v_{id}^{k} + c_1 r_1^{k}\left(p_{id}^{k} - x_{id}^{k}\right) + c_2 r_2^{k}\left(p_{gd}^{k} - x_{id}^{k}\right) \quad (10)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (11)$$

where:

- w is the inertia weight
- $c_1$and c_2 are positive constants
- $r_1$and r_2 are two random values in the range [0, 1]
- v_{id}^{k} the current velocity in iteration k of i^{th} particle in a d -dimensional space
- v_{id}^{k+1} the new velocity in iteration k+1 of i^{th} particle in a d -dimensional space
- p_{id}^{k} the best previous position in iteration k of i^{th} particle in a d -dimensional space
- p_{gd}^{k} the global best particle among all the particles in the population space in iteration k
- x_{id}^{k} the current position in iteration k of i^{th} particle in a d -dimensional space
- x_{id}^{k+1} the new position in iteration k+1 of i^{th} particle in a d -dimensional space

Kennedy and Eberhart also proposed a binary particle swarm optimization (BPSO) in order to solve optimization problems with discrete valued parameters [36]. In BPSO, the position of each particle is represented as binary strings. By comparing PSO and BPSO we found that they have a common velocity equation, and a different particle position equation computed as follows:

$$S\left(v_{id}^{k+1}\right) = \frac{1}{1 + e^{-v_{id}^{k+1}}} \tag{12}$$

$$x_{id}^{k+1} = \begin{cases} 1 \; if \; r_3 < S\left(v_{id}^{k+1}\right) \\ 0 \qquad Otherwise \end{cases} \tag{13}$$

where $\mathrm{S}\left(v_{id}^{k+1}\right)$ is the sigmoid function and r_3 is a random number in the range [0, 1].

BPSO has been widely used recently in the literature for feature subset selection [37]. In this case, the length of a binary string of each particle is equal to the length of the total number of features, and each particle presents a candidate for subset selection. If the bit included in the binary strings has a value of "1" thus the feature is selected, otherwise the feature is not selected. The following steps of BPSO are presented in algorithm 2.12.

Algorithm 2.12: *BPSO*

1.Input:

$F = \{ f_1, f_2, \dots, f_i, \dots, f_m \}$, *m: size of features,* f_i *is a vector of values for feature i of all observation O*

$C = \{ y_1, \dots, y_O \}$, *C* is *vector of classes corresponding to each observation*

2. Output:

$Fbest = \{ f_{s1}, \dots, f_{sj}, \dots, f_{sP} \}$, *P: size of features selected* , $(P < m),\ s_j \in \{1 \dots m\}$

3. Initialization:

Length of particle equal to total number of features *(n=m)*

Initialize all particles positions and velocities randomly

Set the number of iterations *K* and other *BPSO parameters*

4.For k=1:K

A- Calculate the fitness value $F(X_i)$ of each particle. Fitness represents the percentages of correct classification.

B- Compare the fitness of each particle to its best fitness so far ($pbest_i^k$ of last iteration k):

$if\ \ F(X_i^{k+1}) > F(pbest_i^k)\ then\ F(pbest_i^{k+1}) = F(X_i^{k+1})\ and\ pbest_i^{k+1} = X_i^{k+1}$

$Else\ F(pbest_i^{k+1}) = F(pbest_i^k)\ and\ pbest_i^{k+1} = pbest_i^k$

C- Determine the global best position $gbest^{k+1}$ from all $pbest_i^{k+1}$. Then compare $gbest^{k+1}$ with $gbest^k$:

$if\ \ F(gbest^{k+1}) > F(\ gbest^k)\ then\ \ gbest = gbest^{k+1}$

$Else\ gbest = gbest^k$

D- Update the position and the velocity of each particle according to equation (10) and (13).

End for

5. When the limit number of iterations *K* is reached, we obtain an optimal solution $Fbest$ (best subset of feature selection)

In our work, the parameters for the BPSO were chosen classically as: number of particles is 30 particles; length of each particle is equal to the maximum number of features; K=100 iterations [37]. The acceleration constant $c_1$and c_2 were set to 2 [37]. We also used a linear descending inertia weight [37].

To compute the fitness (percentage of correct classification), the classical classifier KNN is used in our work because it is simple and efficient. The best feature subset chosen by the BPSO algorithm is defined as the one giving the maximum percentages of correct classification after 100 iterations (1 run).

2.3 Signal characterization and feature extraction

Following the literature studies done on the EHG signal in chapter 1, several pertinent features have been used in our work.

2.3.1 Monovariate analysis: Linear features

For this monovariate analysis, we computed the selected features on the EHG of the bipolar channel Vb7 and on the EHG of the related monopolar channel CH10. These channels indeed correspond to our reference position (located on the median vertical axis of the uterus).

2.3.1.1 Features related to power spectral density

Several frequency parameters have been extracted from the power spectral density (PSD), $S_x(f)$. In our work, we use the Welch Periodogram method to calculate the power spectral density of each burst [42]. Eleven frequency features are extracted from this PSD: mean frequency *MPF* [38], Peak Frequency *PF* [39, 40, 41], deciles *D1…D9* [42], which contain the median frequency *D5* [40, 42, 43]. Deciles correspond to the frequencies *D1…D9* that divide the power spectral density into 10 parts, each part contains 10% of the total energy.

$$\int_{D_{P-1}}^{D_P} S_x(f)df = 0.1\int_0^{f_{max}} S_x(f)\,df \tag{14}$$

2.3.1.2 Features extracted from wavelet packet decomposition

The decomposition of signals into orthonormal bases (discrete wavelet transform), is based on the theory of multi-resolution analysis (AMR). The basic principles of multi-resolution analysis were laid in 1989 by Mallat [44]. This theory proves that we can analyze a signal by decomposing it into approximation and detail coefficients [45]. The approximations correspond to smoothed versions of the low pass filtered signal, in which changes are gradually rubberized or attenuated. The details contain only the information of high frequencies or discontinuities.

The multiresolution analysis involves projecting the signal x in a series of orthogonal approximation subspace V_j and details subspaces W_i. Changing the resolution of a signal is used to process only the most interesting details for a given task.

The information lost between two approximations is collected into detail signals, which then reflect the behavior of signals at different resolutions. The signal projections subspaces are wholly characterized by the two filters used (high pass and low pass). These filters allow the

rapid calculation of the coefficients of the discrete wavelet transform using an iterative algorithm.

The approximation of a function *f* at resolution 2^{-j} is defined as its orthonormal projection on V_j. To calculate this projection, we need an orthonormal basis of V_j. An orthonormal basis is constructed by dilating and translating a single function called scaling function. The function allows passing from one approximation space to another [46], that is to say, from one scale to another. The approximation subspace V_0 is defined by using the scale of $\phi_{0,0}(t)$ and its translates. The subspace V_j are generated by dilating this function by a factor of 2^j and their translated [46]:

$$\{\phi_{j,k}(t) = 2^{-j/2}\phi_{0,0}(2^{-j}t - k)\}\ k\epsilon\mathbb{Z} \tag{15}$$

The approximation coefficients for signal *x* are defined by:

$$a_x(j,k) = \int_{-\infty}^{+\infty} x(t)\,\phi_{j,k}(t)dt \tag{16}$$

We define $\{W_i\}$, the ensembles such as:

$$W_j \oplus V_j = V_{j-1} \tag{17}$$

W_j represents the details spaces. These spaces are orthogonal complementary of the approximations spaces V_j. One of the major evidence of this approach is the existence of $\psi(t)$, called the mother wavelet and denoted $\psi_{0,0}(t)$, constructed from the scaling function and such as:

$$\{\psi_{j,k}(t) = 2^{-j/2}\psi_{0,0}(2^{-j}t - k)\}\ k\epsilon\mathbb{Z} \tag{18}$$

We define the detail coefficients at level *j* by:

$$d_x(j,k) = \int_{-\infty}^{+\infty} x(t)\,\psi_{j,k}(t)dt \tag{19}$$

The approximation of the signal *x* to the resolution 2^{-j}, corresponding to its projection in V_j, and the associated detail signal, corresponding to its projection in W_j, are presented by the two equations:

$$A_j x(t) = \sum_k a_x(j,k)\, \phi_{j,k}(t) \tag{20}$$

$$D_j x(t) = \sum_k d_x(j,k)\, \psi_{j,k}(t) \tag{21}$$

Authors have used wavelet decomposition to characterize the non-stationary characteristics of the EHG. Relevant features are extracted from the detail coefficients, after the decomposition. In [48], Diab et al. calculated the variances on five selected detail levels: 2, 3, 4, 5 and 6 (named *W1*, *W2*, *W3*, *W4* and *W5*). In our work, we used the wavelet symlet 5, a choice based on the study referenced in [47]. This study compared several types of wavelets. The results have shown that the symlet 5 appears to be the most appropriate wavelet for the analysis of EHG signals for detection and classification purposes. After decomposition of each EHG burst into detail coefficients, we calculate the variances on same detail levels as previously proposed in [48]. The choice of the details depends on the sampling frequency of the signal [48]. The sampling frequency used in our work is 200 Hz, down-sampled by a factor of 12 to obtain a new frequency of 16.67 Hz. Therefore the detail coefficients bandwidth are: Detail 2 [2.08- 4.17 Hz], Detail 3 [1.04-2.08 Hz], Detail 4 [0.52-1.04 Hz], Detail 5 [0.26-0.52 Hz], and Detail 6 [0.13-0.26 Hz] (Figure 2.5). These selected details contain more than 96% of the signal energy and cover the frequency band of interest. Other sampling frequency used in our work is 256 Hz, down-sampled by a factor of 12 and then decompose each EHG into detail coefficients, the 5 detail bandwidths obtained are very close of the ones obtained above.

In [49], wavelet packet decomposition has also been used to reduce the width of the frequency band analysis. Therefore, they decomposed each contraction, selected from EHG recorded during pregnancy and labor, on packets of 3 decomposition levels. In that case, the details as well as the approximations can be split [49]. To minimize the number of decompositions, the signals are first down-sampled by 32 (initial sampling frequency 200 Hz) because the main EHG frequencies lie between 0.1 and 3Hz [49]. After this decomposition, they computed the relative energy for each packet.

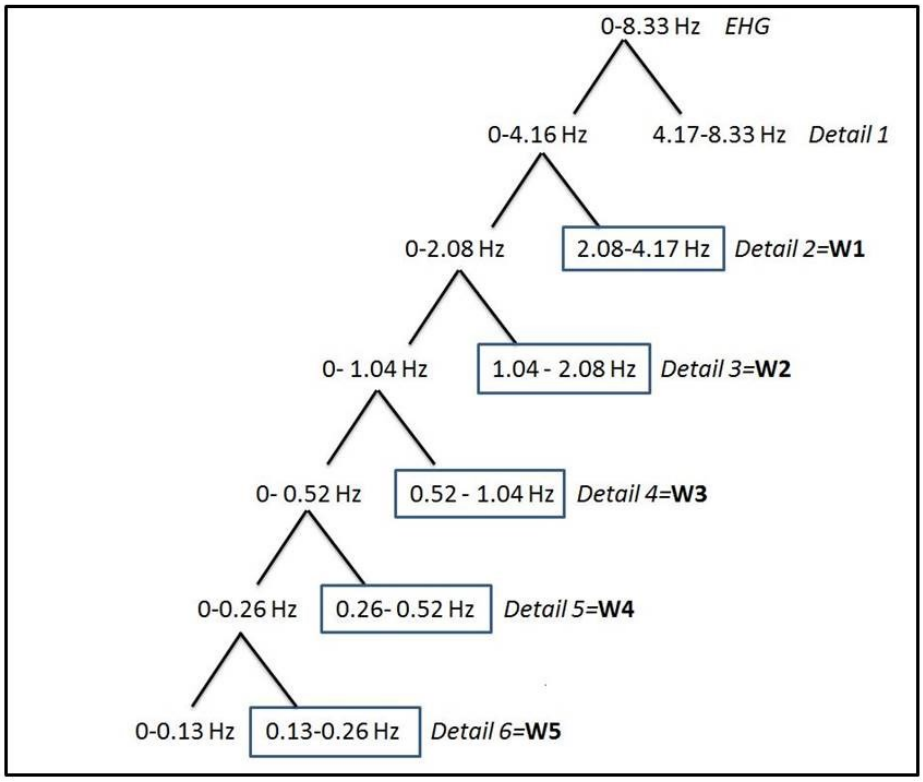


Figure 2.5: Wavelet decomposition

2.3.2 Monovariate analysis: Nonlinear features

2.3.2.1 Time reversibility

A time series is reversible if the probabilistic properties are unchanging with respect to time reversal. The time irreversibility is a good indication of nonlinearity. To calculate the Time Reversibility (*Tr*) characteristic of the signal *x* we have used the equation 22 described in [50]:

$$Tr(\tau) = \frac{1}{N-\tau} \sum_{d=\tau+1}^{N} \left(x(d) - x(d-\tau)\right)^3 \quad (22)$$

Where *x* is the time series or signal, *N* is the signal length and τ is the time delay.

2.3.2.2 Lyapunov exponent

The Lyapunov exponent (*LE*) studies the stability and the sensitivity to initial conditions of the system. It measures the rate of trajectory separation between adjacent tracks in the phase space explained below [50] [51]. In our study we used the equation of *LE* described in [50] and represented by:

$$\lambda = \lim_{t\to\infty} \lim_{\| \Delta_{d_0} \| \to 0} \left(\frac{1}{t}\right) \log \left(\| \Delta_{d_t} \| / \| \Delta_{d_0} \| \right) \quad (23)$$

where $\| \Delta_{d_0} \|$ represents the Euclidean distance between two states of the system at an arbitrary time t_0, and $\| \Delta_{d_t} \|$ corresponds to the Euclidean distance between the two states of the system at a later time *t*.

Phase space:

The phase space is an abstract mathematical space that is considered in the system dynamic. It's constructed using a time delay and an embedding dimension. In order to attain the success of reconstructing the attractor from finite data [52] the choice of the appropriate time delay τ and embedding dimension *m,* is very important.

The reconstructed trajectory, *X*, can be expressed as a matrix where each row is a phase-space vector [53]:

$$X = [X1X2 \ldots . XM]^T \quad (24)$$

where X_i is the state of the system at discrete time *i*.

For an N-point time series $\{x_1, x_2, \ldots, x_N\}$, each X_i is given by:

$$X_i = \left[x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\right] \quad (25)$$

where τ and m represent respectively, the time lag or reconstruction delay and the embedding dimension. Therefore, X is an M×m matrix, and the constants m, M, τ, and N are linked by:

$$M = N - (m-1)\tau \quad (26)$$

As indicated above, the good reconstruction of the phase space depends on the choice of the two parameters, τ and *m*. We have therefore chosen an optimal value for each of these parameters. For the time delay, τ, we use the first local minimum of the average mutual information between the set of measurements *X(i)* and *X(i + τ)*. Mutual information measures the general dependence of two variables [54].

For estimated the minimum embedding dimension, *m*, we used an algorithm proposed by Kennel et al. [55]. The algorithm is based on the idea that, when passing from dimension *m* to

dimension *m+1*, one can differentiate points on the orbit that are true neighbors from those that are false ones. A false neighbor is a point in the data set that is identified as a neighbor solely because the attractor is viewed in a too small embedding space. When the embedding dimension is large enough, all neighbors of every attractor point in the multivariate phase space will be true neighbors [52].

2.3.2.3 Sample entropy

To identify the regularity of EHG signals, the sample entropy (*SE*) is used. In our work, we have used the sample entropy described in [43]. They indicated that a least predictable time series have a higher sample entropy. Consider a time series *x* which represents EHG signal of length *N* and patterns a_j (*0, ...,m-1*) of length *m*, with $m < N$, and $a_j(i)=x(i+j)$; *i=0, ..., m-1; j=0, ...,N-m.* The time series x in a time *t =ts, x* (*ts, ..., ts + m-1*) is a match for a given pattern a_j, if $| x(ts + i) - a_j(i) | <= r$ for each $0<=i<m$. Sample entropy is then computed as follows using equation 27:

$$SE_{m,r}(x) = \begin{cases} -\log\left(\dfrac{C_m}{C_{(m-1)}}\right) & : C_m \neq 0 \wedge C_{m-1} \neq 0 \\ -\log\left(\dfrac{N-m}{N-m-1}\right) & : C_m = 0 \vee C_{m-1} = 0 \end{cases} \quad (27)$$

where the four parameters *N, m, r* and C_m represent respectively the length of the time series, the length of sequences to be compared, the tolerance for accepting, and the number of matching pattern (within a margin for *r*) that is constructed for each *m.*

The value of *m* is chosen equal to 2; this value is determined by the method of the False Nearest Neighbors (FNN). The value of *r* equals 0.2 according to the literature [43].

2.3.2.4 Detrended fluctuation analysis

In [56], Moslem et al. studied the fluctuation of a time series x, EHG signal of length N. This method consists of computing a new integrated series $X(k)$ from the original series x by using equation 28:

$$X(k) = \sum_{i=1}^{k} [\, x(i) - \bar{x} \,] \quad (28)$$

where $\bar{x}$ is the average of $x(i)$ over the whole points. $\bar{x}$ is calculated by using equation 29 :

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x(i) \tag{29}$$

In the next step, we divide the data into boxes of same length *n*. In each box, a polynomial function of degree *a*, is then used to interpolate the sequence (Figure 2.6).

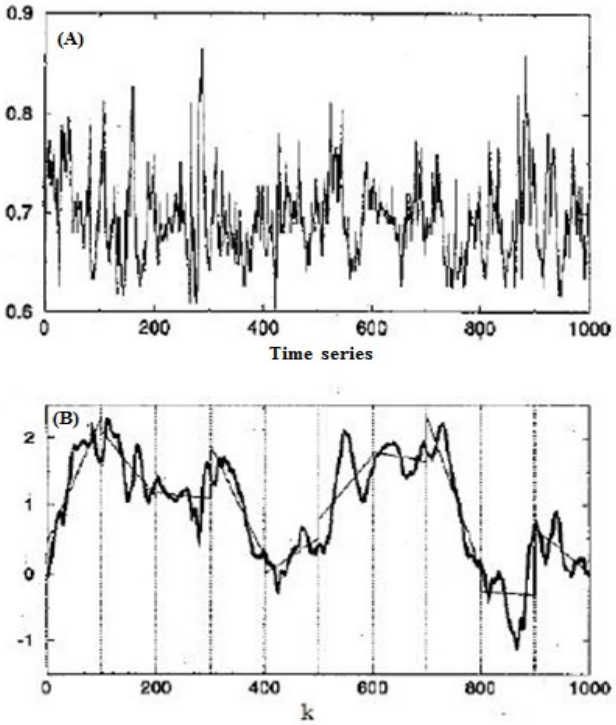


Figure 2.6: (A) Time series. (B) Integrated series $X(k)$.Vertical dotted lines represent windows of length n. Solid lines are the trends estimated for each window by the least square method [56]

A linear fit is normally used, although quadratic, cubic, or higher order polynomials can be used in the fitting procedure. The fluctuation function $F(n)$ is then computed according to equation 30:

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[X(k) - X_n(k)]^2} \tag{30}$$

Characteristically, $F(n)$ increases with the box size *n*. Under these conditions, if the time series is self-similar, a relationship indicates the presence of power law scaling $F(n) \sim n^{\alpha}$. The scaling exponent alpha α can be estimated by using a linear fit on the log–log plot of $F(n)$ versus *n*. The value of α is the correlation properties of the time series.

Finally, we note that α can have different values: for uncorrelated data (ie, white noise), α = 0.5.

If the value of α is lower than 0.5, this indicates that the correlations in the signal are anti-persistent. These correlations become persistent when α is greater than 0.5.

2.3.2.5 Variance entropy

Recent studies [57] used the variance entropy (*VarEn*) to characterize signals but it has never concerned the EHG. We are interested in our work to use variance entropy, because this parameter combines the variance with sample entropy via inverse-variance weighting. For a time series *x*, *VarEn* is defined as equation 31:

$$VarEn(x,m,r) = \frac{\sum_{l=1}^{p} SE_{m,r}(x_l) \,.\, w_l}{\sum_{l=1}^{p} w_l} \tag{31}$$

where x_l is the *l-th* segment of *x*, w_l is inverse variance of x_l, *r* is the tolerance for accepting and *p* is the number of sliding windows. *p* is not fixed because the length of the signals in our data base depends on EHG burst durations.

The sliding window, of size equal to 50, slides over time with a step of 45 (step size), leading to an overlap between the sliding windows equal to 5. The choice of window size and step size were made empirically after several trials. *p* therefore depends on window size.

Because variance entropy combines the variance with sample entropy via inverse-variance weighting, the number of windows is very important and can significantly affect the results. *p* must be neither too high nor too small. A too large *p* value induces large computation time and does not give a precise result. A too small *p* value limits detection of variability.

2.3.3 Bivariate analysis: Features related to EHG propagation

To analyze the synchronization between two EHGs, we choose several methods. For this bivariate analysis, we computed features over all combination of available channels.

2.3.3.1 Correlation coefficient

A- Linear correlation (R^2)

Linear correlation coefficient is calculated for two time series *x(t)*and *y(t)* in the time domain as:

$$R^2 = \max_{\tau} \frac{cov^2(x(t), y(t+\tau))}{var(x(t))var(y(t+\tau))} \tag{32}$$

where *var*, *cov*, and τ denote, respectively, variance, covariance, and time shift between the two time series [58].

B- Nonlinear correlation (H^2)

To calculate H^2, a scatter plot of *y (*Vb8) versus *x (*Vb7) is studied. All values of *x* are divided into bins. We calculate for each bin the *x* value of the midpoint p_u and the average of *y,* q_u calculated from the same bin interval. The regression curve is approximated by linking the obtained points (p_u, q_u) by straight-line segments. Then, H^2 between the two signals *x* and *y* is computed as [59]:

$$H_{Y/X}^2 = \frac{\sum_{k=1}^{N} x(k)^2 - \sum_{k=1}^{N} \big(y(k) - f(x_u)\big)^2}{\sum_{k=1}^{N} y(k)^2} \tag{33}$$

where $f(x_u)$ is the linear piecewise approximation of the nonlinear regression curve, the index u is the number of bins and the index k is a sample of the signal. The nonlinear correlation is an asymmetric measure, in the sense that $H^2(x|y) \neq H^2(y|x)$.

C- General synchronization (H)

Let us reconstruct delay vectors x_n and y_n from the time series measured in two systems x and y. $x_n = \big(x_n, \dots, x_{n-(m-1)\tau}\big)$, $y_n = \big(y_n, \dots, y_{n-(m-1)\tau}\big)$ where $n = 1, \dots . N$; m is the embedding dimension and τ is the delay time. The time indices of the k nearest neighbors of x_n and y_n are respectively $r_{n,j}$ and $s_{n,j}$, $j = 1, \dots . k$.

The squared mean Euclidean distance for each x_n to its k neighbors is:

$$R_n^{(k)}(x) = \frac{1}{k} \sum_{j=1}^{k} \Big(x_n - x_{r_{n,j}}\Big) \tag{34}$$

The y-conditioned squared mean Euclidean distance is:

$$R_n^{(k)}(x|y) = \frac{1}{k} \sum_{j=1}^{k} \Big(x_n - x_{s_{n,j}}\Big)^2 \tag{35}$$

Consequently, a measure of nonlinear interdependence can be identified according to [60]:

$$H^{(k)} = \frac{1}{N} \sum_{n=1}^{N} \log \frac{R_n(x)}{R_n^{(k)}(x|y)} \tag{36}$$

where $R_n(x)$ is the average distance of a vector x_n to all the other vectors.

If x and y are independent, this measure is close to zero, while it is positive if nearness in y also implies nearness in x for equal time partners. The nonlinear interdependence is an asymmetric measure, in the sense that $H(x|y) \neq H(y|x) so\ H(x|y) > H(y|x), if\ x \to y$.

D- Bivariate piecewise stationary signal pre-segmentation (bPSP) and filtration:

In [61], Diab tried to improve the performance of the methods H^2 and H in order to improve the classification of contractions between pregnancy and labor, after testing the sensitivity of these methods to some characteristics of signal (nonstationarity, frequency band) or signal recording (bipolar or monopolar recording). Therefore, he applied H^2 and H on real EHGs segmented by using the bPSP algorithm. After this segmentation, he found that the performance of these methods is improved slightly but the difference obtained between pregnancy and labor remains non-significant. Therefore, he decided to retain only the low frequency band of the EHG (FWL), which is supposed to be more related to the propagation of EHG, with keeping the windowing-preprocessing step. The combination of these two preprocessing steps (Filtered-Windowed-H^2 (FW_H^2) and Filtered-Windowed-H (FW_H) gave the best results, with a clear increase from pregnancy to labor.

2.3.3.2 Phase synchronization

The phase synchronization principle ($\varphi_{e,f}$) consists in a phase locking between two systems described as [62]:

$$\varphi_{e,f}(t) = \left| e\Phi_x(t) - f\Phi_y(t) \right| \leq C \tag{37}$$

where e and f are integers indicating the ratios of possible frequency locking. In this work and according to the literature [63], we assume e=f=1 for simplicity. $\Phi_x(t)$ and $\Phi_y(t)$ are the unwrapped phases of the signals x and y, and C is a constant. In our study we apply a phase synchronization called "mean phase coherence" ($\gamma_{e,f}$) used in [59], represented by equation 38:

$$\gamma_{e,f} = \sqrt{\langle \cos \varphi_{e,f}(t) \rangle^2 + \langle \sin \varphi_{e,f}(t) \rangle^2} \tag{38}$$

where <> indicates average over time.

2.4 Preliminary study: proposition of an algorithm for the selection of parameters using the Jeffrey Divergence Distance

Our proposed method is divided into two parts. The first part consists of implementing and computing all the features already extracted from the EHG that have been published in the literature (presented above). These features were computed, for each existing contraction in the selected database, both on the original EHG and on different frequency bands obtained using wavelet packet decomposition. After obtaining the two matrices of histograms for the two classes (Pregnancy and labor), we proposed to measure histogram similarity or dissimilarity for a given feature between the two classes by using the Jeffrey Divergence method. Therefore, the discriminant and pertinent features have the greater distance between the feature histograms of the pregnancy and labor classes.

2.4.1 Calculated Features and their histograms

As this study took place at the beginning of the thesis, as a preliminary step, we used only here the database1 (described chapter 1 section 1.6.2), only available at this time, composed of 106 pregnancy (Class1) and 106 labor contractions (Class2). We use only the two bipolar signals Vb7 and Vb8, because Vb7 is a reference recording position that has been used for a long time in our research. It is located on the median vertical axis of the uterus. The signal energy in this area remains high throughout the pregnancy as well as during labor. At the time of this study, the algorithm permitting to denoise the monopolar signals was not available yet. We thus used here for monovariate analysis the bipolar Vb7 and for bivariate analysis the two bipolar Vb7 and Vb8. For each contraction of each group, we apply the following steps:

Step 1: Normalize, filter between 0.1 and 3Hz and down-sample (the original sampling frequency is 200 Hz, down-sampled by a factor of 32, we obtain a new sampling frequency = 6.25 Hz) the two signals (Vb7 and Vb8) that correspond to this contraction.

Step 2: Decompose the EHG bipolar channels (Vb7 and Vb8) using wavelet packet transform into three levels. We obtain 15 packets as shown in Figure 2.7. Then we process the packets that contain more than 1% of the total energy. In this case, we obtain 9 signals: 8 subsignals from the selected packets and the original signal (Vb7 or Vb8). In our study, the chosen packets correspond to the following frequency bands [0–3.125 Hz], [0–1.56 Hz], [0–0.78 Hz], [0.78–1.56 Hz], [0–0.39 Hz], [0.39–0.78 Hz], [0.78–1.17 Hz], [1.17–1.56 Hz].

Normalize, filter between 0.1 and 3Hz and down-sample the original signal (the original sampling frequency is 200 Hz, down-sampled by a factor of 12 to obtain a new sampling frequency = 16.67Hz) [64]. As we must calculate in this part the variance of the obtained details after decomposition of each EHG burst, so we did this pretreatment, because the choice of the details depend on the sampling frequency of the signal (as was described in the section 2.3.1.2), in order to correspond to the same frequency bands as the one selected in [48].

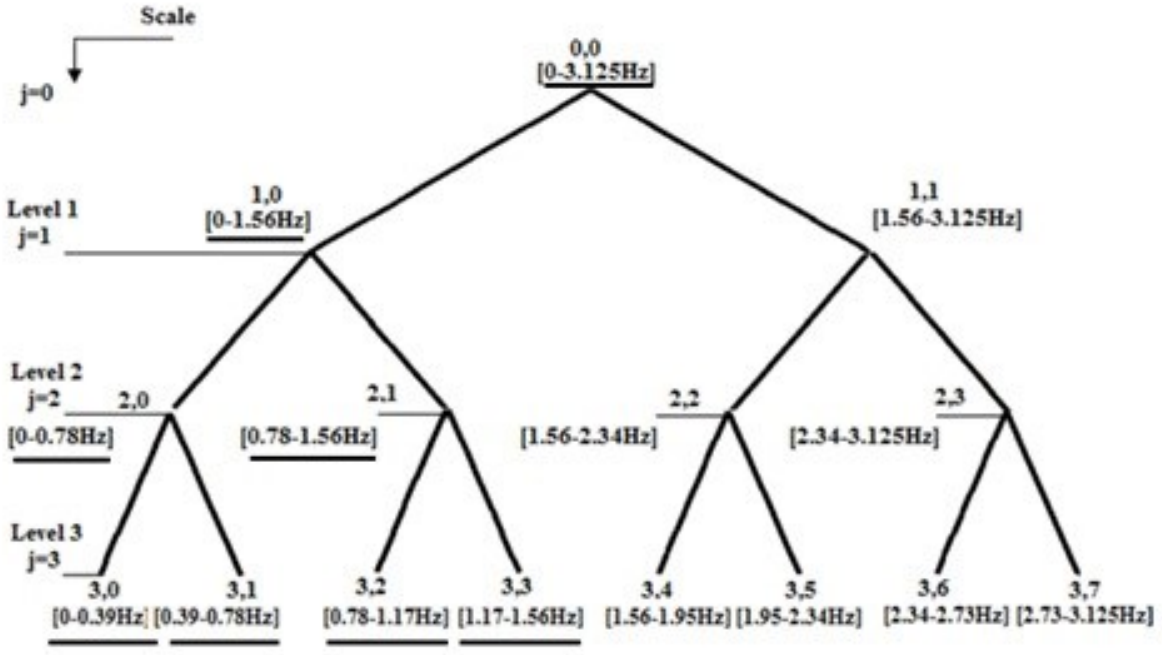


Figure 2.7: Tree of the wavelet packet transform [49]. The underlined packets are the ones selected in this study.

Step 3: Calculate the propagation features between the 8 packets of Vb7 and Vb8 and between the original EHG. Recall that these parameters are: Linear (R^2), nonlinear correlation coefficient (H^2), and the Phase synchronization (γ).

Step 4: For each subsignal and for the original signal, calculate the nonlinear features on the eight packets of the decomposition of Vb7 and on the original EHG. These methods are: Time reversibility (*Tr*), Lyapunov exponent (*LE*), sample entropy (*SE*), Detrended Fluctuation Analysis (*DFA*), and Variance entropy (*VarEn*).

Step 5: Calculate the variances on the following details levels after wavelet decomposition 2,3,4,5 and 6 (*W1, W2, W3, W4* and *W5*) of each packet and on the original EHG.

Step 6: For each signal, compute the frequency features: deciles (*D1, D2, D3, D4, D5, D6, D7, D8* and *D9*), mean frequency (*MPF*), Peak Frequency (*PF*), from the PSD of each packets and from the original EHG.

Steps 7: Group in a matrix these values calculated on the eight packets and the original EHG. Thus we obtain for a given contraction, a matrix with dimensions 9x25. The 25 columns correspond to the 25 features and the 9 rows these values computed from the eight packets and the original EHG.

Step 8: As the Database contains 106 contractions in each class, we obtain 106 matrices of dimensions 9x25 by class. We then compute from these 106 matrices (and for each class) the histogram for each feature. We will obtain 2 matrices of 9x25 histograms (9 signals, 25 features, 2 classes).

2.4.2 Features selection method based on Jeffrey divergence

After obtaining the two matrices of histograms for labor and pregnancy classes, we measure the distance between the two histograms of the 2 classes for a given feature. To measure this distance, we use the Jeffrey Divergence method explained in section 2.2.3.1-A. After calculating these distances, we obtain finally a distance matrix of dimension 9x25. Figure 2.8 presents the distance matrix in the form of a color matrix array, the red color representing the maximum distance value of the distribution and the blue color its minimum. Columns of this matrix present the features and the rows are the eight packets plus the original EHG.

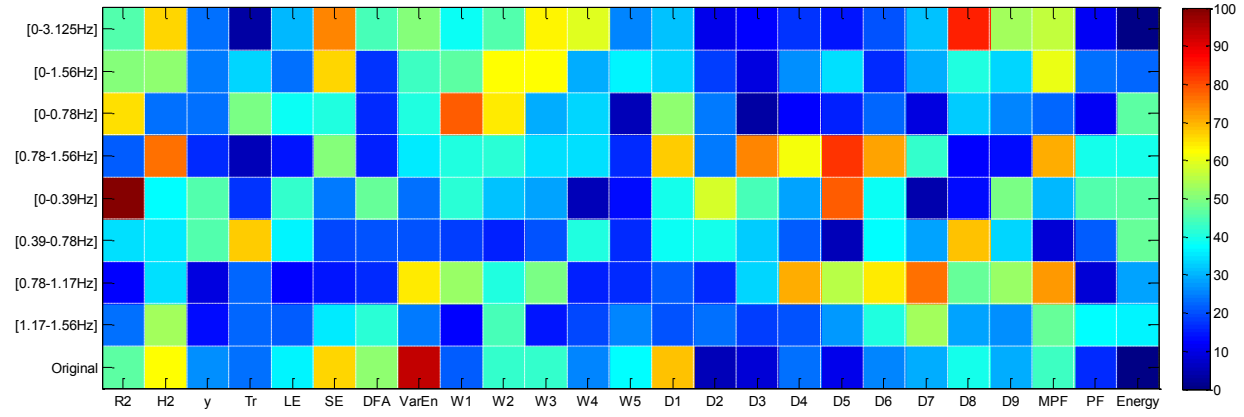


Figure 2.8: Matrix colors presenting the distances between histograms of features, on different packets and on original EHG

The goal of our study is to select the most discriminant features; therefore we apply two thresholds on the matrix of distance in order to select the parameters associated to the larger distances. We calculate the mean and standard deviation of the distance distribution. The first threshold (threshold1) is equal to the mean+1*standard deviation of the distance distribution; the second threshold (threshold2) is the mean+2*standard deviations of the same distribution. We use the second threshold to select the most discriminant parameters. We therefore made no assumption about the number of features that should be selected but kept only the most discriminant ones, in a statistical point of view.

2.4.3 Results on the selection using the Jeffrey Divergence method

2.4.3.1 Features computed on the original EHG

As most of previous works have been done on the whole signal characterization, in order to compare our results to those presented in the literature, we first present the results obtained from the original EHG.

Table 2.5 presents the features selected by applying the two thresholds on the matrix of distance: threshold1= mean + std and threshold2 = mean+ 2*std.

By applying threshold1, we obtained four selected parameters (H^2, *SE, VarEn and D1*), while when applying threshold2, we selected only one most discriminant parameter, *VarEn*.

Table 2.5: Features selected for the best discrimination between pregnancy and labor contractions, for the original signal

Features Selected (Threshold1)	Features Selected (Threshold2)	Histogram distance	Evolution from pregnancy to labor
Nonlinear correlation coefficient (H^2)		0.3306	Decrease
Sample entropy (*SE*)		0.3530	Increase
Variance entropy (*VarEn*)	Variance entropy (*VarEn*)	0.4989	Increase
Decile 1 (*D1*)		0.3654	Decrease

2.4.3.2 Features computed on the packets after wavelet decomposition and on the EHG

This analysis permits us to see which packet (that is which frequency band) has the most influence in discriminating classes. Figure 2.9 shows the selection matrix obtained when processing the selected packets (Packets (0,0)-Packet (3,3)) as well as the original EHG. The results are presented as a black and white matrix of dimensions 9x25, the rows present the selected packets plus the original EHG, and the columns present the features. A white color indicates that the feature has been selected to discriminate between pregnancy and labor. These results are obtained by applying Threshold1 on the distance matrix. For example, the first feature, linear correlation coefficient R^2, can discriminate between pregnancy and labor when applied to the two packets (2, 0) and (3, 0).

Figure 2.10 and Table 2.6 presents the best features on packets and original EHG for discrimination between pregnancy and labor when applying Threshold2 on the distance matrix. These selected features are more discriminant than the ones selected with Threshold1 because Threshold2 selects the features that give a larger distance between pregnancy and labor classes. Indeed, the larger the distance, the more discriminant is the related feature. This table also presents the evolution of these feature values from pregnancy to labor.

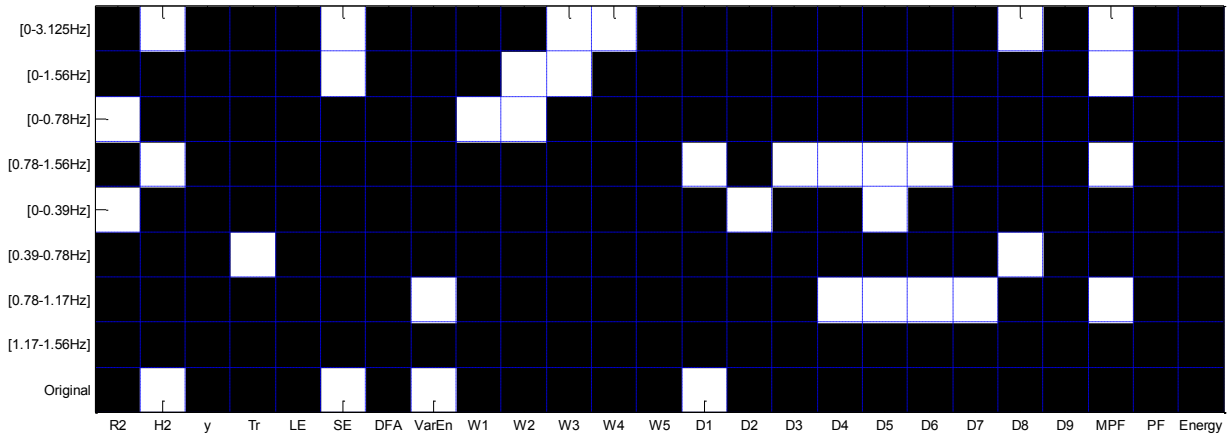


Figure 2.9: Selection matrix, presenting the best features on packets and original EHG for the discrimination between pregnancy and labor. These features are selected by applying threshold1.

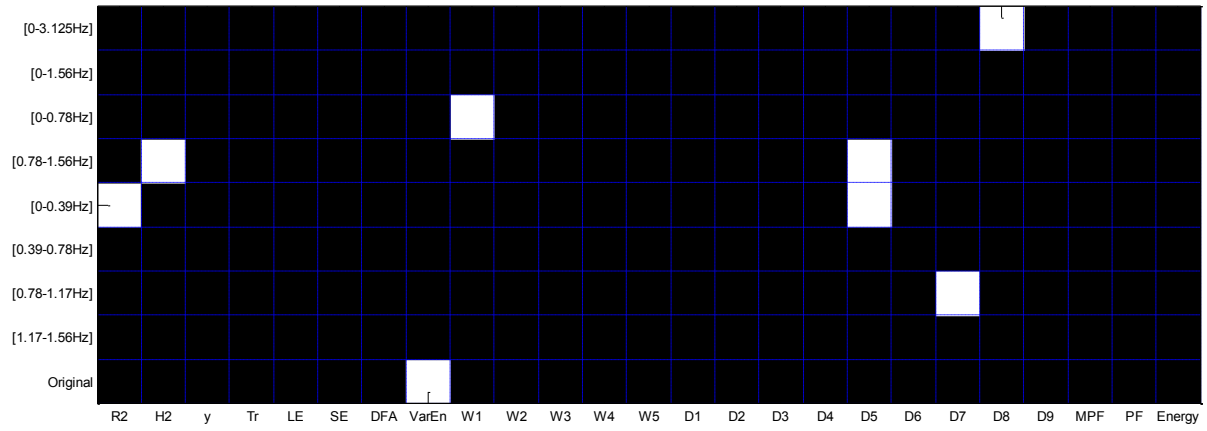


Figure 2.10: Selection matrix, presenting the best features on packets and original EHG for the discrimination between pregnancy and labor. These features are selected by applying threshold2.

Table 2.6: Features selected for the best discrimination between pregnancy and labor contractions, for packets and original signal. These features are selected by applying threshold2.

Features Selected (Threshold2)	**Packets and Original**	**Histogram distance**	**Evolution from pregnancy to labor**
Linear correlation coefficient (R^2)	[0-0.39Hz]	0.5325	Decrease
Nonlinear correlation coefficient (H^2)	[0.78-1.56Hz]	0.4051	Decrease
Variance entropy (*VarEn*)	Original	0.4989	Increase
Variances on the details levels 2 *(W1)*	[0-0.78Hz]	0.4160	Decrease
Decile 5(*D5*)	[0.78-1.56Hz] [0-0.39Hz]	0.4417 0.4165	Decrease Decrease
Decile 7(*D7*)	[0.78-1.17Hz]	0.4096	Decrease
Decile 8(*D8*)	[0-3.125Hz]	0.4494	Increase

2.4.5 Discussions

This part is the preliminary work of our thesis, where we set up all signal processing tools after adjustments (linear parameters, nonlinear parameters and parameters related to the EHG

propagation). Then, a selection method has been applied to choose the best parameters to classify contractions in the uterine electrohysterography (EHG) signal for the detection of preterm labor. Additionally, a Frequency band enhancement EHG characterization has also been extensively studied.

According to the literature, our results indicate that nonlinear correlation coefficient (H^2), sample entropy (*SE*), variance entropy (*VarEn*), and decile 1 (*D1*) are the most discriminating when applied to the original EHG.

The frequency features (deciles *D1*...*D8* and *MPF*), also permit the discrimination between pregnancy and labor on several packets (Figure 2.9). The evolution of frequency features values from pregnancy to labor indicates an increase in the high frequency bands, for example *D8* in the packet [0-3.125Hz] increases from 0.3852 to 0.5085, as well as a small decrease in the low frequency bands, for example *D1* in the packet [0.78-1.56Hz] decreases from 1.0575 to 1.0146. This is in agreement with the works done on the whole signal by different teams, who evidenced a shift of the original EHG content towards higher frequency when going from pregnancy to labor [65].

The propagation features, H^2 and R^2, provide discrimination between pregnancy and labor on several packets and/or on the original EHG (Figure 2.9). Both features decrease on passing from pregnancy to labor contractions. This is in disagreement with previous work done by our team [66] [67]. The explanation is that, in this work, we do not correct H^2 and R^2 by using surrogates, in order to get free from the evolution of the frequency content (proved by the shift of the frequency content towards higher frequency, evidenced by the increase in *MPF*) that induces a bias in R^2 and H^2 values [66] [67]. We did this in order to test H^2 on a wider database than previously used, and because the correction of this feature takes a lot of computing time. Indeed, the main objective of our work is to find the best way to discriminate easily between pregnancy and labor. A recent study [61] found that the use of two simpler preprocessing steps (windowing-preprocessing step, filtration step to retain only the low frequency band of the EHG (FWL)), leads to an increasing of H^2 performance. *H* requires further investigation because the performance of this method is also influenced by the nonlinearity of EHG signals. The two propagation parameters H^2 and *H* with two preprocessing steps will be computed in the following study.

The nonlinear features *Tr*, *SE* and *VarEn*, permit to discriminate between pregnancy and labor on several packets and/or on the original EHG (Figure 2.9). The parameters *SE* and *VarEn* increase from pregnancy to labor, demonstrating an increase in EHG non-linearity in agreement with the bibliography, with the exception of *VarEn* that decreases in the packet [0.78-1.17Hz]. This could be explained by the fact that *VarEn* is sensitive to the number of points. *SE* is very sensitive to sampling frequency [68] as well as *Tr*. When decreasing the number of points used to compute *Tr, Tr* decreases. Here *Tr* is discriminating for the packet (3,1) where the number of points is small, explaining thus why *Tr* decreases. Four features extracted from variance of the wavelet decomposition *(W1*, *W2*, *W3* and *W4*) give discrimination between labor and pregnancy on several packets when using Thereshold1 (Figure 2.9). After analysis of their frequency content, it can be noticed that theses wavelet variances all correspond to the [0.39-1.56Hz] frequency band, where most changes seem to appear between pregnancy and labor contractions. Only *W1* is selected when applying Threshold2. These variance features may be retained if we want to get a better discrimination between pregnancy and labor classes.

2.5 Feature selection from bipolar and monopolar EHG signals

2.5.1 Calculating Features

In this part, 26 features are calculated for each existing contraction in the selected database 2 (explained in chapter1 section 1.6.2). This database is composed of 290 pregnancy contractions and 189 labor contractions from which we processed either bipolar or monopolar signals.

For the monovariate approach, linear and nonlinear features are calculated only from the bipolar signal Vb7 (our reference position and from the monopolar channel CH10 (channel corresponding to one of the electrodes used to compute the bipolar signal Vb7). Features related to the EHG propagation are computed over all combinations of bipolar and monopolar channels.

To compute features, for each contraction of each group, we apply the following steps:

Step 1: Calculate the nonlinear features on bipolar signal Vb7. These features are: Time reversibility (*Tr*), Lyapunov exponent (*LE*), sample entropy (*SE*), Detrended Fluctuation Analysis (*DFA*), and Variance entropy (*VarEn*).

Step 2: Compute the frequency features: deciles (*D1, D2, D3, D4, D5, D6, D7, D8* and *D9*), mean frequency (*MPF*), Peak Frequency (*PF*), from the PSD of bipolar signal Vb7.

Step 3: Normalize, filter between 0.1 and 3Hz and down-sample the bipolar signal Vb7 that correspond to this contraction (sampling frequency for most contractions equal 200 and for other 256, both down-sampled by a factor of 12 to obtain a new sampling frequency equal 16.67Hz and 21.33 respectively). Decompose this signal using wavelet decomposition. Then, calculate the variances on the following details levels after wavelet decomposition 2,3,4,5 and 6 (*W1, W2, W3, W4* and *W5*).

Step 4: To compute the features related to the EHG propagation, between two bipolar channels Vbi and Vbj each feature takes both bipolar channels as input. We took all possible combinations of two bipolar channels among all the available bipolar channels *n.* Therefore we obtain n^2 values of coupling for each feature. As the coupling between the same bipolar channels is forced to zero, we obtain (n^2-n) values of coupling of each feature (here n=12, so 132 coupling values of each feature are obtained). These features are: Linear (R^2), nonlinear correlation coefficient (H^2), and the Phase synchronization (γ), Filtered-Windowed-H^2 (FW_H^2) and Filtered-Windowed-H (FW_H). The features R^2, H^2 and γ are calculated without pretreatment while FW_H^2 and FW_H are calculated after segmentation and filtering the EHG to take the lower frequency bands. All calculations described above are done after segmentation. Finally, we take the mean of each feature over all segments to obtain one value for each feature corresponding to one contraction.

Step 5: For a given contraction, we obtain 26 values. As we have 479 contractions (189 bipolar labor contractions and 290 bipolar pregnancy contractions), we obtain a matrix with dimension 479x26. The 26 columns correspond to the 26 features, and the 479 rows correspond to bipolar signals.

For monopolar database to compute the monovariate and bivariate features, we apply the same five steps presented above but using CH10 (instead of Vb7) and *n*=16 monopolar channels (instead of the 12 bipolar channels). Therefore we obtain a matrix with dimension 479 x 26. The 26 columns correspond to the 26 features, and the 479 rows correspond to monopolar channels.

2.5.2 Feature selection methods

In this study, we tested several methods for feature subset selection to choose the best subsets for classifying labor and pregnancy contractions. These methods are applied to monopolar as well as bipolar EHG signals.

The tested methods are first four features selection methods of type filter (Jeffrey divergence, F-score, relieff and mutual information with clustering). These methods rank features in terms of their importance. Then, the selected features by each method correspond to features that have a rank larger than a specific threshold.

Then the other methods are seven features selection methods of type wrapper: sequential forward selection (SFS), sequential backward selection (SBS), Plus-l minus-r selection (LRS), Bidirectional search (BDS), Sequential Forward Floating sequential (SFFS), Genetic Algorithm (GA) and Binary particle swarm optimization (BPSO). These methods evaluate a subset of features by its classification performance, using a learning algorithm. We used in our work the classical classifier KNN. For the learning step, the data are split by using two cross validation algorithms:

- Holdout method: is the simplest kind of cross validation. The data is randomly split into two sets, called the training set and the testing set (in our work 70% of the data set are used for classifier training and the remaining 30% for testing).
- KFOLD cross validation: the data set is divided into k subsets, and the holdout method is repeated k times. Each time, one of the k subsets is used as the test set and the other k-1 subsets are used as the training set. Then the average error in all k trials is calculated. In our work we use k=10.

The best features subset chosen by sequential methods is defined as the one giving the minimum classification error after trying all sequential combinations. We thus used 50 iterations for each sequential method, and then we choose the subset that has the minimal error between all 50 iterations. For each iteration, data is randomly split using either Holdout or KFOLD method.

The best feature subset obtained by using genetic algorithm is an optimal solution obtained when the limit number of generations K=100 is reached.

The best feature subset chosen by the BPSO algorithm is defined as the one giving the maximum percentages of correct classification after 100 iterations (1 run). Then, to evaluate the performances and variability of BPSO, we performed multiple runs (200 runs). Then we choose the subset of the run that has the maximum percentage of correct classification between all the 200 runs. In each run, dataset is randomly split using either Holdout or KFOLD method.

2.5.3 Results

2.5.3.1 Results of Features selection

For the whole tested methods, we used a dataset of dimensions 379 x 26 (139 labor contractions and 240 pregnancy contractions), except for the Jeffrey Divergence methods that requires an equal number of contractions in each class when measuring the bin-by-bin distances of corresponding histogram bins. We thus used in that case 139 labor contractions and 139 pregnancy contractions extracted from the 240 available.

A- Features Selection using JD Distance

After calculating the distances between every two corresponding feature histograms for the 26 features using bipolar contractions, we obtain a distance vector of dimension 26, as presented Figure 2.11-A. The red color represents the maximum distance value of the distribution and the blue color its minimum. Each point of the horizontal axis represents a different feature. The distance values for each feature is presented in Appendix B.1 (Table B.1.1). Figure 2.11-B shows the selection vector obtained after applying the threshold = mean (all values of distance) on the distance vector. Features having value of distances higher than the mean value of all distances are selected. A white color indicates that these features have been selected as being discriminants between pregnancy and labor.

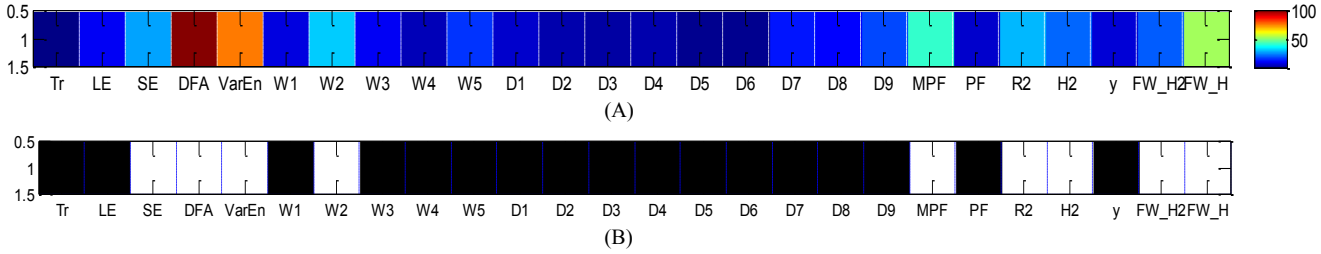


Figure 2.11: (A) Color vector representing the distribution of distances between features from bipolar EHGs. (B) Selection vector representing the most discriminating features for the discrimination between bipolar pregnancy and labor EHG bursts.

Figure 2.12 presents the same processing described above but by using monopolar pregnancy and labor EHGs. The distance value of each feature is presented in Appendix B.1 (Table B.1.1). The two subsets of features obtained after applying Jeffrey Divergence on bipolar and monopolar data set are presented table 2.7.

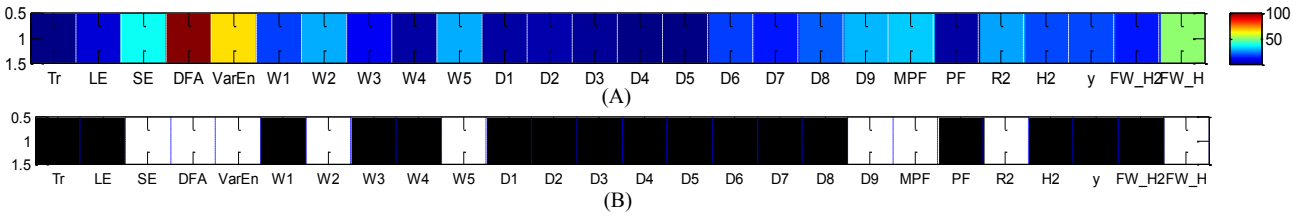


Figure 2.12: (A) Color vector representing the distribution of distances between features from monopolar EHGs. (B) Selection vector representing the most discriminating features for the discrimination between monopolar pregnancy and labor EHG bursts.

B- Features Selection using F-score

After calculating F-score for each feature, we obtain 26 F-score-values. The F-score values of each feature are presented Appendix B.1 (Table B.1.2). Figure 2.13-A and Figure 2.14-A present the distribution of F-score values from bipolar and monopolar EHGs respectively. Features having values of F-score higher than threshold = mean (all values of F-score) are selected. Therefore, 6 features are selected from bipolar (Figure 2.13-B) and monopolar datasets (Figure 2.14-B). These two subsets of features are presented table 2.7.

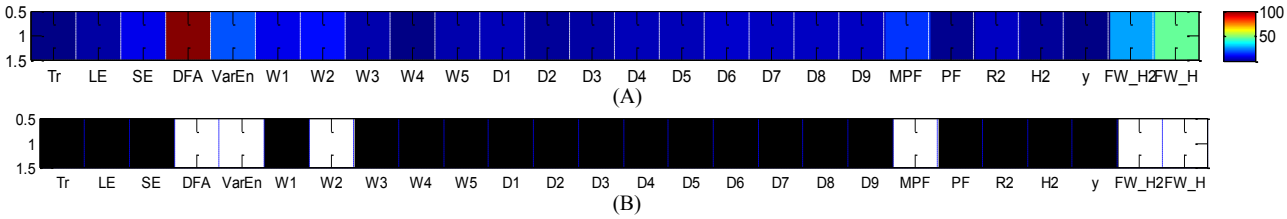


Figure 2.13: (A) Color vector representing the distribution of F-score values for bipolar dataset. (B) Selection vector representing the best features for the discrimination between pregnancy and labor contractions using bipolar EHGs.

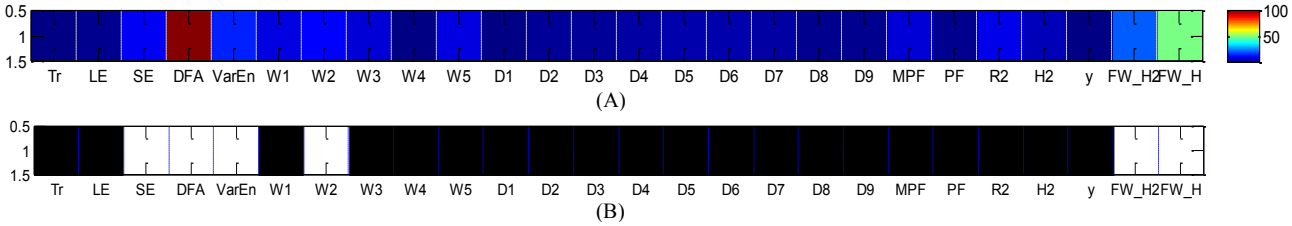


Figure 2.14: (A) Color vector representing the distribution of F-score values for monopolar dataset. (B) Selection vector representing the best features for the discrimination between pregnancy and labor contractions using monopolar EHGs.

C- Feature Selection using Relieff

Relieff method calculates the weight of each feature. 26 weights are obtained corresponding to the 26 features. The values of feature weights are presented Appendix B.1 (Table B.1.3). The

distribution of weights is presented in Figure 2.15-A and Figure 2.16-A for respectively bipolar and monopolar datasets. After applying a threshold = mean (all values of weights), 8 features (Figure 2.15-B) and 9 features (Figure 2.16-B) are selected from bipolar and monopolar dataset respectively. The two obtained subsets are presented table 2.7.

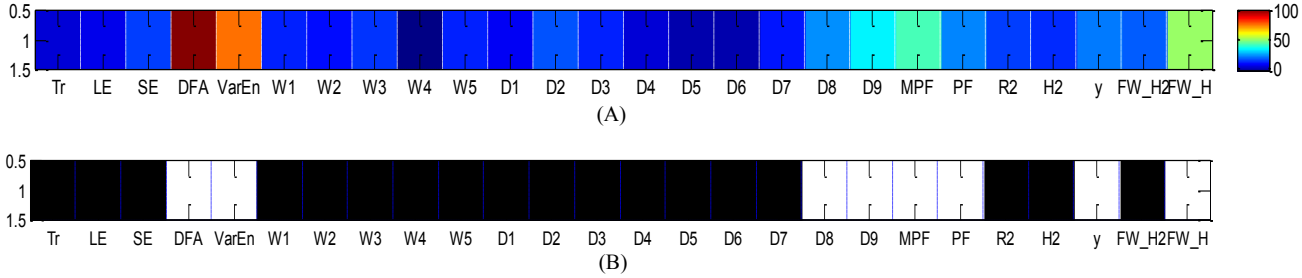


Figure 2.15: (A) Color vector representing the distribution of weight values for bipolar dataset. (B) Selection vector representing the best features for the discrimination between pregnancy and labor contractions using bipolar EHGs.

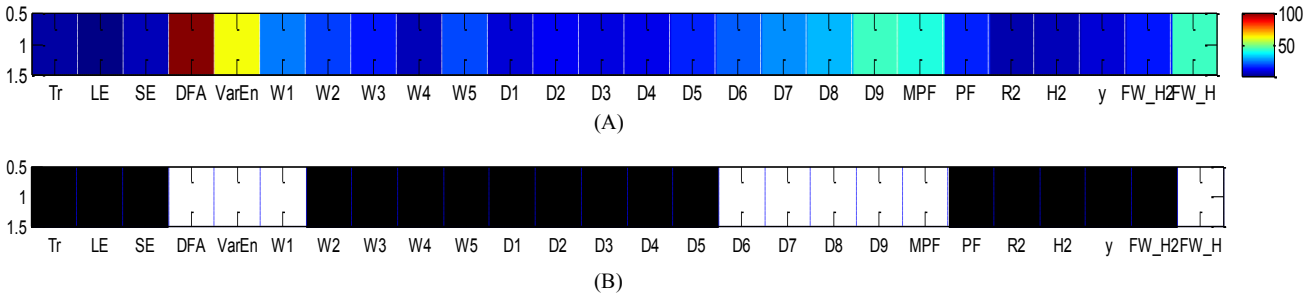
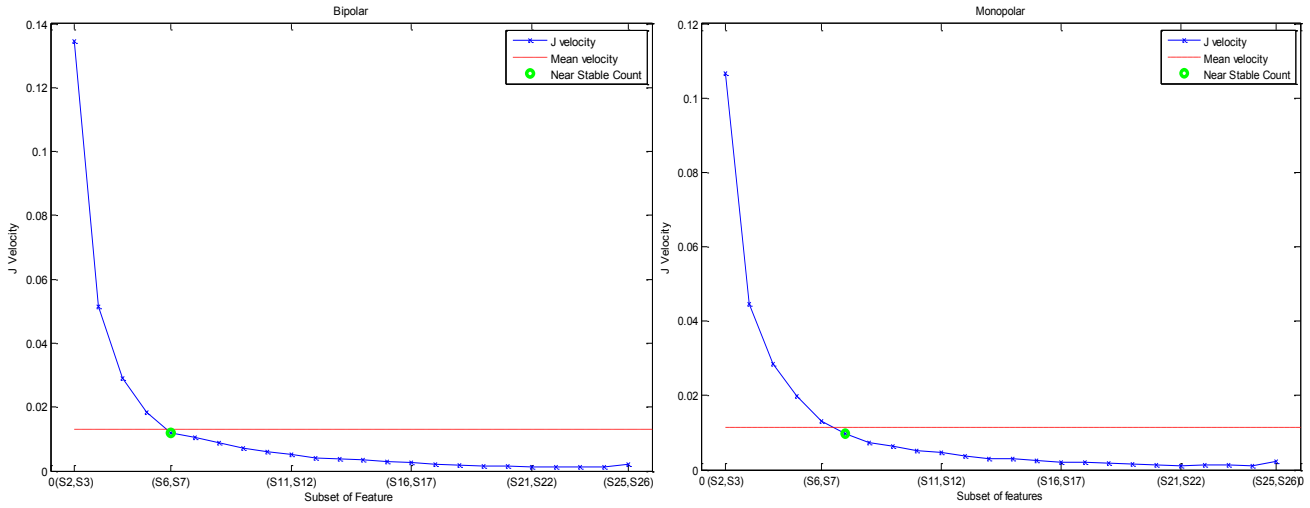


Figure 2.16: (A) Color vector representing the distribution of weight values for monopolar dataset. (B) Selection vector representing the best features for the discrimination between pregnancy and labor contractions using monopolar EHGs.

D- Feature Selection using mutual information with clustering

The best subset of features selected using mutual information with clustering is found when the variation of J is approximatively negligible. Figure 2.17-A and Figure 2.17-B present the variation of J using bipolar and monopolar datasets respectively. In figure 2.17-A, we notice that the variation of J (J velocity) calculated between the two subset of features S6 and S7 (S6: subset corresponds to the six added features and S7: subset corresponds to the seven added features) is the first which is lower than the mean value of all variances. Therefore the best subset selected corresponds to the subset S6. In figure 2.17-B the best subset selected is the subset S7. The two obtained subsets are presented table 2.7.



(A) (B)

Figure 2.17: (A) J velocity using bipolar contraction. (B) J velocity using monopolar contractions.

Table 2.7: Subsets of features selected using several filter methods, for the best discrimination between pregnancy and labor contractions, as well as their related computation time.

Methods	**Selected feature Subset**		**Time (s)**	
	Bipolar	**Monopolar**	**Bipolar**	**Monopolar**
Jeffrey divergence (JD)	[DFA, VarEn, FW_H, MPF,W2, R^2, SE, H^2, FW_H^2]	[DFA, VarEn, FW_H, SE, MPF, D9,W5, W2, R^2]	0.09	0.08
F-Score	[DFA, FW_H, FW_H^2, VarEn, MPF, W2]	[DFA, FW_H, FW_H^2 , VarEn, W2, SE]	0.005	0.009
Relieff	[DFA, VarEn, FW_H, MPF, D9, D8, PF, y]	[DFA, VarEn, FW_H, D9, MPF, D8, D7, W1, D6]	0.35	0.36
Mutual information with clustering	[DFA, MPF, FW_H, SE, VarEn, FW_H^2]	[DFA, VarEn, FW_H, SE, W5, MPF, FW_H^2]	0.08	0.08

E- Feature Selection using sequential methods

Figure 2.18 presents the evolution of the procedure for each sequential method applied on the monopolar dataset and using KFOLD data split. In this figure each evolution is followed by

means of three curves. The first curve is the variation of the criterion (classification error) corresponding to each combination of features sequentially generated (third curve). The operation direction for each method is represented in the second curve by a value of either "-1" for feature removal or "+1" for feature addition. The red circle in the first and second curves corresponds to the lowest classification error obtained for each features subset. We choose the feature subset with the smallest number of features corresponding to this lowest classification error.

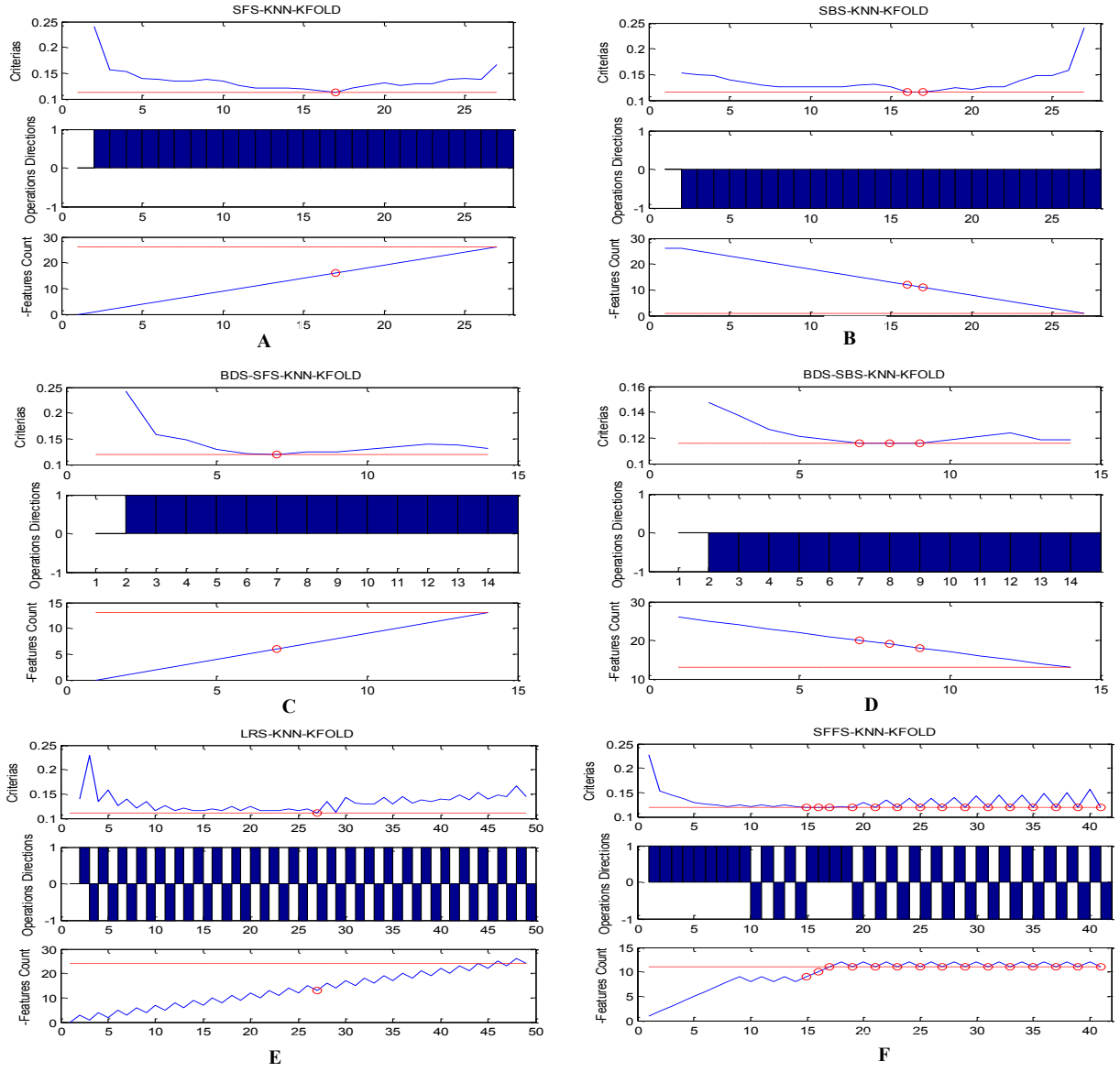


Figure 2.18: Evolutions of the procedures for the sequential methods using KNN-KFOLD with the monopolar dataset. (A) SFS-KNN-KFOLD. (B) SBS-KNN-KFOLD. (C) BDS-SFS-KNN-KFOLD. (D) BDS-SBS-KNN-KFOLD. (E) LRS-KNN-KFOLD. (F) SFFS-KNN-KFOLD.

The other sequential methods evolutions on bipolar using KNN-KFOLD, monopolar using KNN-Holdout and bipolar using KNN-Holdout are presented Appendix B.2.

Table 2.8 presents the selected feature subsets obtained from these sequential methods using the classifier KNN and the two data splits KFOLD and Holdout when applied to the bipolar and monopolar EHGs. Each selected feature subset corresponds to the one that gives the maximum percentage of correct classification calculated from the classification error.

Table 2.8: Selected feature subsets using sequential methods on bipolar and monopolar datasets as well as their related computation time.

Method	**Selected parameter Subset**		**Percentage of correct classification**		**Time (s)**	
	Bipolar	**Monopolar**	**Bipolar**	**Monopolar**	**Bipolar**	**Monopolar**
SFS-KNN-KFOLD	[SE, DFA, W4, MPF, H^2]	[Tr, SE, DFA, VarEn, W1, W2, W3, D1, D2, D3, D4, D5, D6, MPF, PF, FW_H]	88.13	88.65	6.41	6.23
SBS-KNN-KFOLD	[SE, DFA, W4, D8, MPF, H^2, y]	[DFA, W3, W4, D5, D6, MPF, PF, H^2, y, FW_H^2, FW_H]	87.86	88.39	6.33	6
BDS-SFS-KNN-KFOLD	[DFA, W2, W3, D2, D4, D6, MPF, H^2]	[SE, DFA, D1, D5, MPF, PF]	88.13	88.13	6.64	6.67
BDS-SBS-KNN-KFOLD	[Tr, SE, DFA, VarEn, W1, W2, W4, D3, D4, D6, D8, D9, MPF, H^2, y]	[Tr, DFA, VarEn, W1, W2, W4, D1, D3, D5, D6, D7, D8, MPF, PF, R^2, y, FW_H^2, FW_H]	87.07	88.39	6.68	6.66
LRS-KNN-KFOLD	[DFA, W1, W3, D2, D3, MPF, H^2]	[Tr, SE, DFA, VarEn, W1, W2, W3, D3, D4, D5, PF, R^2, FW_H^2]	88.13	88.92	30.01	30.08
SFFS-KNN-	[SE, DFA, D2, D3, MPF, R^2]	[Tr, SE, DFA, D2, D4, D5, MPF, PF,	88.13	88.13	14.91	14.99

KFOLD		FW_H^2]				
SFS-KNN-Holdout	[Tr, DFA, W1, W2, W3, D1, D4, D8, MPF, R^2, FW_H]	[Tr, DFA, VarEn, W1, W2, W4, D1, D2, D4, D5, D8, R^2]	93.81	95.58	1.20	1.20
SBS-KNN-Holdout	[DFA, W5, MPF, FW_H^2]	[SE, DFA, VarEn, W4, W5, D7, D8, PF, R^2, FW_H^2]	95.58	93.81	1.23	1.26
BDS-SFS-KNN-Holdout	[DFA, D3, D8, H^2]	[SE, DFA, VarEn, W4, D5, D8, MPF, R^2]	93.81	94.69	1.31	1.34
BDS-SBS-KNN-Holdout	[DFA, VarEn, W1, W3, D1, D2, D7, D8, D9, MPF, PF, H^2, y]	[SE, DFA, VarEn, W4, D5, D6, D7, D8, MPF, R^2, H^2, y, FW_H^2]	92.92	94.69	1.32	1.34
LRS-KNN-Holdout	[Tr, DFA, W1, W2, D1, D2, D3, D4, D5, D6, D7, D8, PF, R^2, H^2]	[DFA, VarEn, W1, W2, D1, D2, D4, D8, MPF, R^2, y, FW_H]	93.81	95.58	5.79	5.90
SFFS-KNN-Holdout	[Tr, SE, DFA, W1, W2, W3, MPF, R^2, H^2, FW_H^2]	[Tr, DFA, W2, D4, R^2, H^2]	94.69	93.81	3.07	2.67

F- Features Selection using Genetic Algorithm

Figure 2.19 presents the evolution of the GA-KNN-KFOLD selection algorithm for the bipolar dataset. This figure shows the best and mean classification error percentage for the population at every generation step. The best subset obtained by GA is an optimal solution obtained when the limit number of generation *K* (*K*=100) is reached. The other figures corresponding to the procedure for GA-KNN-Holdout-bipolar, GA-KNN-KFOLD-monopolar and GA-KNN-Holdout-monopolar are presented Appendix B.3.

Table 2.9 presents the results of the selected feature subsets obtained after applying GA-KNN-KFOLD and GA-KNN-Holdout to the bipolar and monopolar EHGs. These subsets correspond

to the ones that gives the maximum correct classification percentage (calculated form the classification error percentage).

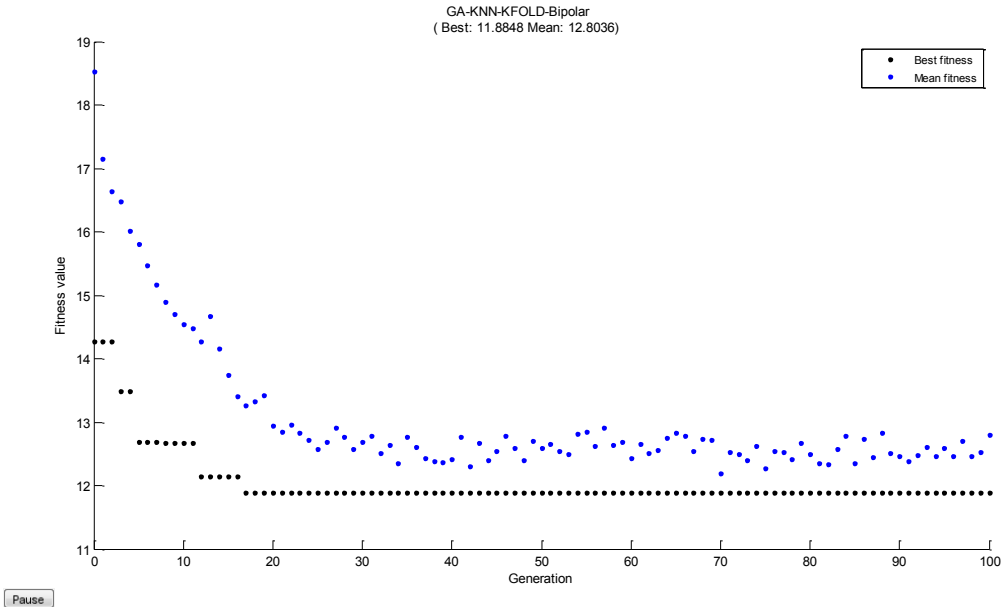


Figure 2.19: Evolution of the procedure of GA-KNN-KFOLD on bipolar dataset

Table 2.9: Selected features subset using Genetic algorithm (GA) on bipolar and monopolar datasets, as well as their related computation time.

Methods	**Selected parameter Subset**		**Percentage of correct classification**		**Time (s)**	
	Bipolar	**Monopolar**	**Bipolar**	**Monopolar**	**Bipolar**	**Monopolar**
GA-KNN-KFOLD	[Tr, SE, DFA, W1, W4, D1, D3, D4, D5, D8, MPF, R^2, H^2, y]	[DFA, W1, W2, W3, W4, D5, D6, MPF, R^2, FW_H]	88.12	88.14	145.56	161.52
GA-KNN-Holdout	[SE, DFA, VarEn, W1, W2, W4, W5, D1, D5, MPF, R^2, H^2, FW_H^2, FW_H]	[DFA, VarEn, D3, D4, D7, D8, PF, R^2, FW_H]	93.81	94.69	21.24	25.67

G- Features Selection using binary particle swarm optimization

Table 2.10 presents the selected feature subsets obtained after applying BPSO-KNN-KFOLD and BPSO-KNN-Holdout to the bipolar and monopolar EHGs. Each selected feature subset corresponds to the one that give the maximum percentage of correct classification after 200 runs.

Table 2.10: Selected features subset using binary particle swarm optimization *(BPSO) on bipolar and monopolar datasets, as well as their related computation time.*

Methods	**Selected feature Subset**		**Percentage of correct classification**		**Time (s)**	
	Bipolar	**Monopolar**	**Bipolar**	**Monopolar**	**Bipolar**	**Monopolar**
BPSO-KNN-KFOLD	[Tr, SE, DFA, W1, W2, W4, D3, D4, D5, D6, D8, MPF, R^2, H^2, y]	[SE, DFA, VarEn, W1, D1, D2, D3, D4, D5, D6, MPF, PF, H^2, y, FW_H]	88.16	88.42	44.19	43.72
BPSO-KNN-Holdout	[Tr, DFA, W4, D4, D5, D6, D8, D9]	[Tr, DFA, VarEn, W3, W4, D1, D2, D5, D6, MPF, H^2, y, FW_H]	95.58	96.46	6.06	6.04

2.5.3.2 Validation Part

A validation part was applied on the remaining contractions (50 labor and 50 pregnancy contractions) not used during features selection and training steps, in order to evaluate the performances of the selected subsets presented above. Therefore, two datasets are used of dimension 100*26 for monopolar dataset and 100*26 for bipolar dataset. From the results presented above, the features selection methods applied on bipolar and monopolar EHGs give:

- 8 subsets of features selected by filter methods
- 32 subsets of features selected by wrapper methods (using the classifier KNN and the data split KFOLD or Holdout)

We evaluated the performances of these 40 selected subsets by calculating for each of them the mean ± standard deviation of the percentages of correct classification using 500 repetitions. We used for this validation the same classifiers and data split already used in the selection step. For the wrapper methods, the evaluation will use the data split used for feature selection part. For the

filter methods we will evaluate the subsets performance by using KNN-KFOLD and KNN-Holdout.

The results of this validation step, presented Appendix B.4 table B.4.1 contain the mean ± standard deviation of the percentages of correct classification after 500 repetitions for each subset. Table 2.11 presents the feature subsets with the highest mean of the percentage of correct classification for bipolar and monopolar EHGs. For bipolar signals, we retained the 2 subsets selected by GA-KNN-KFOLD and BPSO-KNN- KFOLD, because they present very close mean correct classification percentages (85.06 ± 1.63 and 85.05 ± 1.69). For monopolar signals, we retained only the subset selected by BDS-SBS-KNN- Holdout that presents the highest mean of correct classification (74.88 $\pm$ 7.10).

Table 2.11: Subsets of features selected for bipolar and monopolar dataset with highest Mean ± STD (over 500 repetitions) for the percentage of correct classification using KNN

Datasets	**Methods**	**Selected feature Subset**	**Mean $\pm$ standard deviation of percentage of correct classification (500 repetitions)**
Bipolar	GA-KNN-KFOLD	[Tr, SE, DFA, W1, W4, D1, D3, D4, D5, D8, MPF, R^2, H^2, y]	85.06 $\pm$ 1.63
	BPSO-KNN- KFOLD	[Tr, SE, DFA, W1, W2, W4, D3, D4, D5, D6, D8, MPF, R^2, H^2, y]	85.05 $\pm$ 1.69
Monopolar	BDS-SBS-KNN-Holdout	[DFA, VarEn, W1, W3, D1, D2, D7, D8, D9, MPF, PF, H^2, y]	74.88 $\pm$ 7.10

In order to compare the results of classification with and without selection, table 2.12 presents the mean of the percentage of correct classification for bipolar and monopolar EHGs using all features combinations without selection.

Comparing tables 2.11 and 2.12, we can notice that the results (percentages of correct classification) obtained when using the selected feature subsets are always higher than those obtained when using the whole features, regardless the data splits methods or the datasets (monopolar and bipolar EHGs).

Table 2.12: Mean ± STD of the percentage of correct classification using KNN of 500 repetitions using the feature set on bipolar and monopolar EHG

	Datasets	**Subset**	**Mean ± standard deviation of percentage of correct classification (500 repetitions)**
Kfold	**Bipolar**	[MPF, PF, D1, D2... D9, W1, W2, W3, W4, W5, Tr, LE, SE, DFA, VarEn, R^2, H^2, y, FW_H^2, FW_H]	71.89 ± 2.01
	Monopolar		68.94 ± 1.86
Holdout	**Bipolar**		71.9 ± 7.26
	Monopolar		67.87± 7.48

In order to evaluate the performance of classification for the three retained subsets, corresponding to the the highest mean of the percentage of correct classification for bipolar and monopolar EHGs, we use the three classical statistical measures:

- Sensitivity (true positive ratio) = $\frac{True\ positive}{True\ positive+False\ negative}$
- Specificity (true negative ratio) = $\frac{True\ negative}{True\ negative+False\ positive}$
- Accuracy = $\frac{True\ positive+True\ negative}{True\ positive+False\ negative+\ True\ negative+False\ positive}$

Table 2.13 presents the results of Mean ± standard deviation of sensitivity, specificity and accuracy of the 3 selected subsets using, for each subset, KNN classifier and the data split similar to the one used in the selection step.

Table 2.13: Performance of the retained feature subsets for bipolar and monopolar datasets.

Dataset	Method	Feature Subset	Performance of classification (Mean± std deviation over 500 repetitions)		
			Sensitivity	Specificity	Accuracy
Bipolar	GA-KNN-KFOLD	[Tr, SE, DFA, W1, W4, D1, D3, D4, D5, D8, MPF, R^2, H^2, y]	92.22 ± 2.38	78.8 ± 3	84.99±1.73
	BPSO-KNN-KFOLD	[Tr, SE, DFA, W1, W2, W4, D3, D4, D5, D6, D8, MPF, R^2, H^2, y]	91.89 ± 2.51	78.9 ± 2.71	84.92 ± 1.77
Monopolar	BDS-SBS-KNN-Holdout	[DFA, VarEn, W1, W3, D1, D2, D7, D8, D9, MPF, PF, H^2, y]	80.84 ±9.35	68.64±11.7	74.35 ± 6.94

2.5.4 Discussion

- Selection method:

We can notice from the results (Tables 2.7, 2.8, 2.9 and 2.10) that the different feature selection methods give different feature subsets. Therefore, 20 selected feature subsets are obtained from bipolar EHGs and 20 other from monopolar EHGs, with different number of features from one subset to another (Table B.5.1, Appendix B.5).

We can notice also from this Table, that the wrapper methods take more computation time than the filter methods, due to the use of a classifier.

When comparing the subsets selected by wrapper methods from bipolar EHGs to the ones selected from monopolar EHGs, we notice that the highest percentage of correct classification (96.46 %) is obtained from the monopolar signals.

- Selected features:

Due to the variability of the features selected in the different subsets by the different methods, it is very important to highlight the most repetitive selected features. Indeed, these features are expected to be the most pertinent. Table B.5.2 (Appendix B.5) presents the number of

appearance of each feature over the 20 selected feature subsets for monopolar and bipolar EHGs. We can see from this table that only one feature (*DFA*) is always selected on monopolar and bipolar case, whatever the method. This monovariate nonlinear feature is thus confirmed to be of great interest for pregnancy/labor comparison. This confirms the observation made by different teams concerning the interest of taking into account the nonlinear characteristics of EHG for diagnostic purposes [45, 50]. *DFA* increases from pregnancy to labor on monopolar as well as on bipolar EHGs (see appendix B.6 table B.6.1), demonstrating an increase in EHG non-linearity. Contrary to the *DFA*, the *LE* is not selected on monopolar and bipolar case, whatever the applied method. This feature should not be used for diagnosis purpose. Some other features, selected several times are also of possible great interest for diagnosis. For example the *MPF* is selected 17 over 20 times on bipolar EHG. *MPF* increases when passing from pregnancy to labor on monopolar as well as bipolar EHGs (see appendix B.6 table B.6.1). This is in agreement with different studies, which evidenced a shift of the EHG content toward higher frequency when going from pregnancy to labor [65].

Concerning the bivariate parameters, H^2 presents the highest rank of this feature family for bipolar signals and *FW_H* the highest for monopolar signals. This demonstrates a discriminating power of these features. But when we look at their pregnancy/labor evolutions, they both decrease when going from pregnancy to labor, which is an unexpected result. Some further work needs to be done to explain the results.

- Validation step:

According to the validation studies developed in order to test the performance of the selected subsets (Appendix B.4, Table B.4.1), we noticed that, the most feature subsets selected by wrapper methods gave best performance than features subsets selected by filter methods. We thus conclude that wrapper methods take more time in the selection than filter methods, but they improve the performance of classification of pregnancy and labor contractions.

Additionally, according to these validation studies, we notice that the results of classification with selection is better than without selection on bipolar and monopolar signal. From the results of classification with selection, we noticed that the best performance obtained for bipolar EHGs is about 85 % of correct classification (Table 2.11). The best result obtained for monopolar EHGs is about 75% (Table 2.11). The results presented in table 2.13 indicate that the sensitivity is higher than the specificity for the 3 selected features subsets whatever the bipolar or

monopolar EHGs. This may be due to the fact that, in study we use a different number of labor and pregnancy contractions for the feature selection part, from which the classifiers can learn.

Furthermore, when comparing the results obtained with these 3 subsets we can notice that the two subsets selected from bipolar EHGs give better results also in terms of sensitivity, specificity and accuracy than the subset selected from monopolar EHGs. We can thus conclude from this study that bipolar EHGs should be used rather than monopolar EHGs because they give higher classification performance. A more complete comparison of monopolar vs bipolar EHGs use will be presented chapter 3.

References

[1] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *Knowl. Data Eng. IEEE Trans. On*, vol. 17, no. 4, pp. 491–502, 2005.

[2] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artif. Intell.*, vol. 97, no. 1, pp. 245–271, 1997.

[3] H. Liu and H. Motoda, *Feature extraction, construction and selection: A data mining perspective*. Springer Science & Business Media, 1998.

[4] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, no. 3, pp. 131–156, 1997.

[5] P. Langley and others, "Selection of relevant features in machine learning". *Defense Technical Information Center*, 1994.

[6] S. F. Pratama, A. K. Muda, Y.-H. Choo, and N. A. Muda, "A Comparative Study of Feature Selection Methods for Authorship Invarianceness in Writer Identification," *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 4, pp. 467–476, 2012.

[7] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[8] S. Ding, "Feature Selection Based F-Score and ACO Algorithm in Support Vector Machine," in *Second International Symposium on Knowledge Acquisition and Modeling, 2009. KAM '09*, 2009, vol. 1, pp. 19–23.

[9] Hall, M. A. "Correlation-based Feature Subset Selection for Machine Learning". *PhD Thesis*, University of Waikato, 1999.

[10] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1, pp. 273–324, 1997.

[11] V. Bolón-Canedo, N. Sánchez-Maroño, and A. Alonso-Betanzos, "A review of feature selection methods on synthetic data," *Knowl. Inf. Syst.*, vol. 34, no. 3, pp. 483–519, Mar. 2013.

[12] H. Chouaib, " Sélection de caractéristiques: méthodes et applications," Ph.D. dissertation, Thèse de l'université Paris Descartes, 2011. [in French].

[13] W. Megchelenbrink, E. Marchiori, and P. Lucas, "Relief-Based feature selection in bioinformatics: detecting functional specificity residues from multiple sequence Alignments," *Radboud Univ. Nijmegen*, 2010.

[14] A. Rajaei, E. Dallalzadeh, and L. Rangarajan, "Symbolic representation and classification of medical X-ray images," *Signal Image Video Process*., pp. 1–11.

[15] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis*., vol. 40, no. 2, pp. 99–121, 2000.

[16] Y. Ma, X. Gu, and Y. Wang, "Histogram similarity measure using variable bin size distance," *Comput. Vis. Image Underst*., vol. 114, no. 8, pp. 981–989, Aug. 2010.

[17] H.-H. Hsu, C.-W. Hsieh, and M.-D. Lu, "Hybrid feature selection by combining filters and wrappers," *Expert Syst. Appl*., vol. 38, no. 7, pp. 8144–8150, Jul. 2011.

[18] K. Kira and L. A. Rendell, "A practical approach to feature selection," *in Proceedings of the ninth international workshop on Machine learning*, 1992, pp. 249–256.

[19] I. Kononenko, "Estimating attributes: analysis and extensions of RELIEF," *in Machine Learning: ECML-94*, 1994, pp. 171–182.

[20] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Mach. Learn*., vol. 53, no. 1–2, pp. 23–69, 2003.

[21] M. Robnik-Šikonja and I. Kononenko, "Comprehensible interpretation of relief's estimates*," in Machine Learning: Proceedings of the Eighteenth International Conference on Machine Learning (ICML'2001), Williamstown, MA, USA. San Francisco: Morgan Kaufmann*, 2001, pp. 433–40.

[22] H. Liu, Y. Mo, J. Wang, and J. Zhao, "A new feature selection method based on clustering," *in 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD),* 2011, vol. 2, pp. 965–969.

[23] T. Marill and D. M. Green, "On the effectiveness of receptors in recognition systems," *Inf. Theory IEEE Trans. On*, vol. 9, no. 1, pp. 11–17, 1963.

[24] D. W. Aha and R. L. Bankert, "A comparative evaluation of sequential feature selection algorithms," *in Learning from Data, Springer*, 1996, pp. 199–206.

[25] L. Ladha and T. Deepa, "Feature selection methods and algorithms," *Int. J. Comput. Sci. Eng*., vol. 3, no. 5, pp. 1787–1797, 2011.

[26] A. W. Whitney, "A direct method of nonparametric measurement selection," *Comput. IEEE Trans*. On, vol. 100, no. 9, pp. 1100–1103, 1971.

[27] A. R. Webb, *Statistical Pattern Recognition*. John Wiley & Sons, 2003.

[28] M. Sabeti, R. Boostani, S. D. Katebi, and G. W. Price, "Selection of relevant features for EEG signal classification of schizophrenic patients," *Biomed. Signal Process. Control*, vol. 2, no. 2, pp. 122–134, Apr. 2007.

[29] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994.

[30] Z. Jiang, K. Yamauchi, K. Yoshioka, K. Aoki, S. Kuroyanagi, A. Iwata, J. Yang, and K. Wang, "Support vector machine-based feature selection for classification of liver fibrosis grade in chronic hepatitis C," *J. Med. Syst*., vol. 30, no. 5, pp. 389–394, 2006.

[31] D. Ververidis and C. Kotropoulos, "Fast and accurate sequential floating forward feature selection with the Bayes classifier applied to speech emotion recognition," Signal Process., vol. 88, no. 12, pp. 2956–2970, Dec. 2008.

[32] M. Melanie, "An introduction to genetic algorithms," *Camb. Mass. Lond. Engl. Fifth Print*., vol. 3, 1999.

[33] D. Whitley, "A genetic algorithm tutorial," *Stat. Comput*., vol. 4, no. 2, pp. 65–85, Jun. 1994.

[34] B. Oluleye, L. Armstrong, L. Jinsong, and D. Diepeveen, "Zernike moments and genetic algorithm: Tutorial and application," *Br. J. Math. Comput. Sci*., vol. 4, no. 15, pp. 2217–2236, 2014.

[35] R. Eberhart and J. Kennedy, "New optimizer using particle swarm theory," in *Proceedings of the 6th International Symposium on Micro Machine and Human Science (MHS '95)*, pp. 39– 43, October 1995.

[36] J. Kennedy and R. C. Eberhart, "Discrete binary version of the particle swarm algorithm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4104–4108, October 1997.

[37] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi, "Application of the PSOSVM model for recognition of control chart patterns," *ISA Transactions*, vol. 49, no. 4, pp. 577–586, 2010.

[38] J. Terrien, T. Steingrimsdottir, C. Marque and B. Karlsson, "Synchronization between EMG at different uterine locations investigated using time-frequency ridge reconstruction: comparison of pregnancy and labor contractions," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 2010.

[39] W. L. Maner, R. E. Garfield, H. Maul, G. Olson, and G. Saade, "Predicting term and preterm delivery with transabdominal uterine electromyography," *Obstetrics & Gynecology*, vol. 101, no. 6, pp. 1254–1260, 2003.

[40] J. Sikora, A. Matonia, R. Czabanski, K. Horoba, J. Jezewski, and T. Kupka, "Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals," *Archives of perinatal medicine*, vol.17, no.2, pp.97-103, 2011.

[41] J M. Lucovnik, W. L. Maner, L. R. Chambliss, R. Blumrick, J. Balducci, Z. Novak-Antolic, and R. E. Garfield, "Noninvasive uterine electromyography for prediction of preterm delivery," *Am. J. Obstet. Gynecol*., vol. 204, no. 3, pp. 228.e1–228.e10, Mar. 2011.

[42] C. Marque, H. Leman, M. L. Voisine, J. Gondry, and P. Naepels, "Traitement de l'électromyogramme utérin pour la caractérisation des contractions pendant la grossesse," *RBM-News*, vol. 21, no. 9, pp. 200–211, Dec. 1999.

[43] G. Fele-Žorž, G. Kavšek, Ž. Novak-Antolič, and F. Jager, "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups," *Medical and Biological Engineering and Computing*, vol. 46, no. 9, pp. 911–922, 2008.

[44] G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions* on, vol. 11, pp. 674-693, 1989.

[45] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection", *Information Theory, IEEE Transactions on*, Vol. 38, No. 2, 1992, pp. 713–718

[46] F. Truchetet, *Ondelettes pour le signal numérique*: Hermes Sciences Publicat., 1998.

[47] M. Khalil, "Une approche de la détection et de la classification dans les signaux non stationnaires. Application a l'EMG utérin," Ph.D. dissertation, Thèse de l'université de Technologie de Troyes, 1999. [in French].

[48] M. O. Diab, C. Marque, and M. A. Khalil, "Classification for uterine EMG signals: Comparison between AR model and statistical classification method, "*International Journal of computational cognition*, vol. 5, no. 1, pp.8-14, 2007.

[49] B. Moslem, M. Khalil, C. Marque, and M. O. Diab, "Energy distribution analysis of uterine electromyography signals," J. Med. Biol. Eng., vol. 30, no. 6, pp. 361–366, 2010.

[50] A. Diab, M. Hassan, C. Marque, and B. Karlsson, "Quantitative performance analysis of fourmethods of evaluating signal nonlinearity: application to uterine EMG signals," *in Proceedings of the 34th Annual International IEEE EMBS Conference*, San Diego, Calif, USA, September 2012,pp. 1045–1048.

[51] T. Ivancevic, L. Jain, J. Pattison, A. Hariz, and others, "Preterm birth analysis using nonlinear methods," *Recent patents on biomedical engineering,* vol. 1, no. 3, pp. 160-170, 2008.

[52] C.K. Lee, H.G. Jo, and S.K. Yoo, "Non-linear Analysis of Single Electroencephalography (EEG) for Sleep-Related Healthcare Applications," *Healthc Inform Res*, vol. 16, no. 1, pp. 46-51, Mar. 2010.

[53] M.T. Rosenstein, J.J. Collins, and C.J. De Luca, "A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets," Physica D-Nonlinear Phenomena, vol. 65, no. 1-2, pp. 117-134, May. 1993.

[54] H.D.I. ABARBANEL, "Analysis of Observed Chaotic Data," *in Rev. Mod. Phys*., New York, London: Springer, 1st Ed., 1996.

[55] M.B. Kennel, R. Brown, and H.D.I. Abarbanel, "Determining Embedding Dimension for Phase-Space Reconstruction Using a Geometrical Construction," Physical Review A, vol. 45, no. 6, pp. 3403-3411, Mar. 1992.

[56] B. Moslem, M. Khalil, M. O. Diab, C. Marque,"Detrended fluctuation analysis of uterine electromyography", *Presented at the First Middle East Conference on Biomedical Engineering, MECBME11*, Sharjah, UAE, Feb.2011.

[57] M. Hu and H. Liang, "Variance entropy: a method for characterizing perceptual awareness of visual stimulus," *Applied Computational Intelligence and Soft Computing*, vol. 2012, pp.1-6, 2012.

[58] K. Ansari-Asl, F. Wendling, J. J. Bellanger, and L. Senhadji, "Comparison of two estimators of time-frequency interdependencies between nonstationary signals: application to epileptic EEG," *in 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2004. IEMBS '04*, 2004, vol. 1, pp. 263–266.

[59] M. Hassan, A. Alexandersson, J. Terrien, C. Muszynski, C. Marque, and B. Karlsson, "Better pregnancy monitoring using nonlinear propagation analysis of external uterine electromyography", *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 4, 2013,pp. 1160–1166.

[60] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals*," Prog. Neurobiol*., vol. 77, no. 1, pp. 1–37, 2005.

[61] A. Diab, "Study of The Nonlinear Properties And Propagation Characteristics Of The Uterine Electrical Activity During Pregnancy And Labor", Ph.D. dissertation, Thèse de l'université de Technologie de Compiègne, 2014.

[62] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, "Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings," Presented at the 32 nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires, Espagnol, Aug.2010.

[63] J. Bhattacharya, "Reduced degree of long-range phase synchrony in pathological human brain," *Acta Neurobiol. Exp. (Warsz.),* vol. 61, no. 4, pp. 309–318, 2001.

[64] D. Alamedine, M. Khalil, and C. Marque, "Comparison of Different EHG Feature Selection Methods for the Detection of Preterm Labor," *Comput. Math. Methods Med*., vol. 2013, pp. 1-9, 2013

[65] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchêne, "Uterine electromyography: a critical review.," *Am. J. Obstet. Gynecol.*, vol. 169, no. 6, p. 1636, 1993.

[66] J. Terrien, M. Hassan, G. Germain, C. Marque, and B. Karlsson, "Nonlinearity testing in the case of non Gaussian surrogates, applied to improving analysis of synchronicity in uterine contraction," in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, 2009, pp. 3477–3480.

[67] T. Jérémy, M. Catherine, G. Guy, and K. Brynjar, “Sources of bias in synchronization measures and how to minimize their effects on the estimation of synchronicity: application to the uterine electromyogram,” 2009.

[68] A. Diab, M. Hassan, B. Karlsson, and C. Marque, “Effect of decimation on the classification rate of non-linear analysis methods applied to uterine EMG signals,” *IRBM*, vol. 34, no. 4, pp. 326–329, 2013.

Chapter 3: Channel and channel combination selection for monovariate and bivariate analysis on EHG

3.1 Introduction

A multichannel system was used in order to record simultaneous EHG at different locations. In our study, we used a 4x4 electrode matrix positioned on the woman's abdomen to record 16 monopolar EHGs. However, using simultaneously multiple sensors entails a complex processing of the recorded signal, through the large number of features obtained in this situation. Multichannel classification has been already applied to EHG [1-6]. A comparison between channels showed that some channels give better classification than others. Therefore, it is interesting to select the best channels to evidence the differences between pregnancy and labor contractions.

In our work we used monovariate (processing only one EHG channel at a time) and bivariate (measuring the coupling between 2 EHG channels) analysis. Using all channels, for the monovariate, or all combinations of channels, for the bivariate analysis, leads to a large dimension of features for one contraction. The aim of this chapter is, 1) the selection of the best channels, for the monovariate, 2) the selection of best channel combinations, for the bivariate analysis, that provide the most useful information to discriminate between pregnancy and labor classes. For this purpose, Relieff [7] and F-score method [8] (a filter type feature selection methods), are applied in order to select the best channels or the best channel combinations for discriminating these two classes. After channel selection, wrapper type feature selection methods, named Binary particle swarm optimization [9,10] and Genetic algorithm (GA) [11], are used in order to select the best features (monovariate linear and nonlinear features) from the selected channels, and to select the best features related to the EHG propagation (bivariate analysis) from the selected channel combinations. This step is very important to ease classification problem. After the selection part, we validated the performance of the selection process on a population of EHG bursts different from the ones used for the selection process. Furthermore, in this chapter the channels and

channel combinations selection followed by feature selection are applied to monopolar and then bipolar EHG signals, for both monovariate and bivariate analysis.

3.2 Channel selection for EHG monovariate analysis

3.2.1 Calculated Features

In this part of work, we use only linear and nonlinear features that are the suitable features for monovariate analysis. The chosen linear features are: mean frequency (*MPF*) [12], Peak Frequency (*PF*) [13], deciles (*D1*…*D9*) [14] which contain the median frequency *D5* [14, 15], features extracted from wavelet decomposition (*W1*…*W5*) [16]. The nonlinear features are: Time reversibility (*Tr*) [17], Lyapunov exponent (*LE*) [17], Sample Entropy (*SE*) [15], Variance entropy (*VarEn*) [18] and Detrended fluctuation analysis (*DFA*) [19]. Therefore, 21 features (Linear and nonlinear) have been extracted from each EHG burst. See section 2.5.1 for more details about the features calculation.

EHG signals are recorded using the multichannel system permitting the simultaneous recording of 16 channels of EHG or 12 bipolar EHG described in section 1.5.3. We compute these 21 features from each burst corresponding to either the 16 monopolar or the 12 bipolar channels, for the 290 pregnancy and the 189 labor contractions of Database 2 (section 1.6.2). Therefore, we obtain 2 three dimensional matrices: the first one corresponds to the values of features extracted for all pregnancy and labor contractions (479 contractions) for the whole monopolar channels CHi (479*16*21). The second one corresponds to the features extracted for all pregnancy and labor contractions corresponding to each bipolar channel Vbi (479*12*21).
A part of this dataset is used for the selection part (379 over the 479 contractions: 240 for pregnancy class, 139 for Labor class). The other part is used for validation (50 pregnancy and 50 labor contractions).

3.2.2 Channel selection followed by feature selection

Computing features for the whole channels leads to a large searching dimension (number of channels*number of features) and induces classification problems due to the large number of inputs. Therefore, in this part, we aimed to select the channels with the best discrimination ability between the two classes, according to each given feature. For this purpose, the feature selection methods F-score and Relieff, described in section 2.2.3.1-B and 2.2.3.1-C respectively, are

adapted for monopolar (CHi) or bipolar (Vbi) channel selection. Due to the huge number of channels and parameters, we choose to use a filter type feature selection method of two different categories (Fscore (univariate), Relieff (multivariate)) due to their fast execution time.

After this channel selection step, two feature selection methods, Binary Particle Swarm Optimization (BPSO) and Genetic algorithm are used in order to select the best features from the selected channels. The choice of these methods relied to the results obtained in chapter 2 (BPSO and GA were from the best methods obtained in chapter 2).

3.2.3 Results

3.2.3.1 Channel selection using Relieff and F-score

Two datasets are used: the first one is the monopolar dataset (dimension 379*16*21) and the second is the bipolar dataset (dimension 379*12*21). For the F-score and Relieff methods, used for channel selection, the selection is done by computing the best discrimination ability between the two classes (labor and pregnancy) using a given feature. As in our work we extracted 21 features from the EHG, 21 subsets of channels selection are obtained for each method. The channels will finally be selected according to their frequency of selection over the 21 possible features.

All the results for the bipolar and monopolar channels selection are presented Appendix C-1. The results of bipolar channels selection using the F-score and Relieff method applied to each feature are summarized Table 3.1-A and 3.1-B respectively (see the details of selection in Appendix C.1, table C.1.1).

We can notice from Table 3.1-A that the bipolar channel Vb8 (selected 17 over 21 times), and Vb7 (selected 15 over 21 times) are the most repetitive channels. From Table 3.1-B the bipolar channel Vb7 (selected 17 over 21 times), Vb8 and Vb9 (selected 14 over 21 times) are the most repetitive channels. Figure 3.1-(B, C) presents the location of these selected bipolar channels.

Table 3.1: Bipolar channels appearance after channel selection using F-score (A) and Relieff (B)

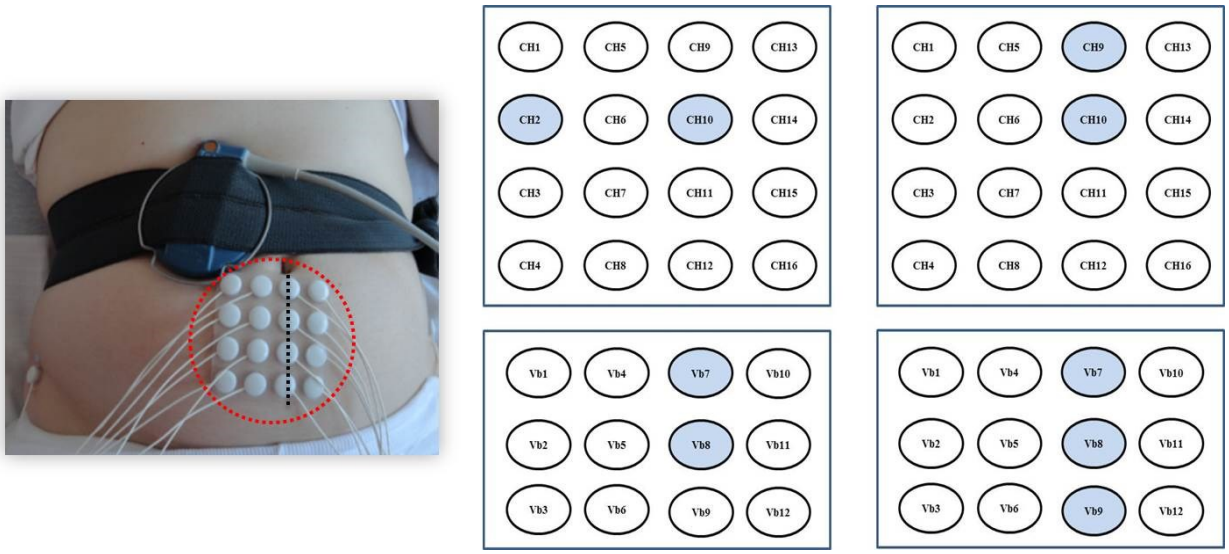
(A) **F-score**		(B) **Relieff**	
Bipolar Channel (sorted)	**Number of appearance**	**Bipolar Channel (sorted)**	**Number of appearance**
Vb8	17	Vb7	17
Vb7	15	Vb9	14
Vb5	14	Vb8	14
Vb9	11	Vb2	12
Vb4	10	Vb5	11
Vb1	10	Vb11	9
Vb2	8	Vb10	7
Vb11	6	Vb4	7
Vb3	6	Vb3	7
Vb12	5	Vb12	6
Vb10	5	Vb6	4
Vb6	3	Vb1	3

The same channel selection process is applied in order to select the monopolar channels (CHi).The results summarized Table 3.2-A show that the CH2 and CH10 (selected 14 over 21 times) are the most repetitive channels, while in Table 3.2-B the most repetitive channels are CH9 and CH10 (selected 18 over 21 times). Figure 3.1-(B, C) presents the location of these selected monopolar channels. For more information of the selection see Appendix C.1, table C.1.2.

Table 3.2: Monopolar channels appearance after channel selection using F-score (A) and Relieff (B).

(A) F-score		**(B) Relieff**	
Monopolar Channel (sorted)	**Number of appearance**	**Monopolar Channel (sorted)**	**Number of appearance**
CH 10	14	CH10	18
CH 2	14	CH9	18
CH 9	12	CH12	13
CH 5	12	CH5	13
CH 16	11	CH1	12
CH 12	11	CH16	11
CH 11	11	CH^2	11
CH 15	10	CH15	9
CH 1	8	CH14	9
CH 14	7	CH11	8
CH 3	7	CH8	6
CH 13	6	CH4	6
CH 6	5	CH13	4

CH 8	4	CH7	2
CH 4	4	CH6	2
CH 7	3	CH3	2



(A) (B) (C)

Figure 3.1: (A) Position of the 16 monopolar electrodes. (B) Blue circles: Monopolar and Bipolar channels selected by F-score. (C) Blue circle: Monopolar and Bipolar channels selected by Relieff.

3.2.3.2 Feature selection using BPSO and GA on the selected channels

Binary Particle Swarm Optimization and genetic algorithm are applied to the features of the selected channels, in order to reduce the dimension of features. They both use a fitness function, here the percentage of correct classification obtained by the K-nearest neighbors classifier (KNN). The data set is split in two for the learning phase: Holdout (70% of the data set for classifier training and the remaining 30% for testing) and KFOLD cross validation (K=10). The results will be presented as the percentages of correct classification using the classifier KNN and the data split KFOLD or Holdout.

BPSO and GA using KNN-KFOLD and KNN-Holdout are applied to our four obtained datasets (see Table 3.3). The dataset obtained after bipolar channel selection using F-score (DL1) contains the features extracted from the contractions, for the 2 bipolar channels selected. In this case the particle length of BPSO and length of chromosome of GA are equal to 42 (2 bipolar

channels*21 features). The second dataset corresponds to the bipolar channel selection using Relieff (DL2) that contains the features extracted from the contractions for the 3 bipolar channels selected. In this case the particle length of BPSO and length of chromosome of GA are equal to 63 (3 bipolar channels *21 features). The two other datasets contain the features extracted from the 2 monopolar channels CH2, CH10 selected by F-score (DL3=379*42) and the 2 monopolar channels CH9 and CH10 selected by relieff (DL4=379*42). In this case the particle length of BPSO and the length of chromosome of GA are equal to 42.

Table 3.3: Dataset obtained after bipolar and monopolar channel selection using F-score and Relieff

Channel selection method	**Bipolar**		**Monopolar**	
	Channels	**Dataset**	**Channels**	**Dataset**
F-score	Vb7, Vb8	DL1= 379*42	CH2, CH10	DL3= 379*42
Relieff	Vb7, Vb8, Vb9	DL2= 379*63	CH9, CH10	DL4= 379*42

Table C.2.1 in Appendix C.2 presents the selected features subsets obtained from BPSO and GA by using the classifier KNN and the two data splits KFOLD and Holdout when applied to our 4 bipolar and monopolar channels selection datasets (DL1, DL2, DL3 and DL4). The number of selected features varies from one method to another as indicated in table C.2.2 (Appendix C). Table 3.4 presents an extract of the 3 results (subsets S8, S12 and S16 - see appendix C.2, Table C.2.1) that gave the maximum percentage of correct classification on bipolar (DL2) and monopolar channels selection datasets (DL3, DL4). These percentages are 97.35% and 96.46% respectively.

Table 3.4: Subsets of features selected from the bipolar and monopolar channels with the highest percentage of correct classification using KNN

Datasets	Methods	Dataset	Selected feature subset	Number of features	Percentage of correct classification	Time (s)
Bipolar	**Bipolar, Relieff, BPSO-KNN-Holdout**	DL2	S8	23	97.35	7.01
Monopolar	**Monopolar, F-score , BPSO-KNN-Holdout**	DL3	S12	20	96.46	6.13
	Monopolar, Relieff, BPSO-KNN-Holdout	DL4	S16	21	96.46	6.11

3.2.3.3 Validation

The results presented above gave 16 subsets of features selected using BPSO and GA with the classifier KNN and the two data split KFOLD and Holdout, when applied to our 4 bipolar and monopolar channel selection datasets. We try in this part to test the performance of each selected feature subset on the part of the data base that was kept apart for validation (50 labor and 50 pregnancy contractions). These contractions were not used during the feature selection step which results are presented above.

Appendix C.3 (Table C.3.1) presents the mean ± standard deviation of the percentages of correct classification of 500 repetitions for each subset. Table 3.5 presents only the subsets of features selected on bipolar and monopolar EHG with the highest mean ± standard deviation of the percentage of correct classification. Additionally in order to evaluate the classification performance of these two selected feature subsets, we calculate the Mean ± standard deviation of sensitivity, specificity and accuracy of these subsets using KNN classifier and the data split corresponding to each subset (Table 3.6).

Table 3.5: Selected feature subset from the bipolar and monopolar channel selection with the highest Mean± STD (over 500 repetitions) of the percentage of correct classification using KNN.

Dataset	Method	Selected feature subset	Number of features	Percentage of correct classification
Bipolar	**Bipolar, Relieff, GA-KNN-KFOLD**	S5	36	78.07 ± 1.13
Monopolar	**Monopolar, Relieff, BPSO-KNN-Holdout**	S16	21	72.33 ± 6.88

Table 3.6: Performance of the selected feature subset from the bipolar and monopolar selected channels giving the highest Mean ± STD (over 500 repetitions) of the percentage of correct classification using KNN.

Datasets	Methods	Performance of classifier (Mean + std dev. over 500 repetitions)		
		Sensitivity	**Specificity**	**Accurancy**
Bipolar	**Bipolar, Relieff , GA-KNN-KFOLD**	84.89 ±2.19	71.65±2.79	77.97 ± 1.11
Monopolar	**Monopolar, Relieff, BPSO-KNN-Holdout**	67.05 ±12.57	76.75 ± 10.71	71.45 ± 6.77

3.2.4 Discussion

In this first part of this chapter, we presented the results of bipolar and monopolar channels selection followed by feature selection, permitting to evidence at best the differences between pregnancy and labor contractions. A comparison between the results of bipolar and monopolar channels selection using two different methods (F-score and Relieff) followed by a feature selection step, on these selected channels, using two methods (binary particle swarm optimization and genetic algorithm) has been performed to choose the methods and the subsets allowing the best discrimination between labor and pregnancy contractions. Finally, as the final validation step, we compared the performance of the obtained subsets, by using labor and pregnancy contractions not used in the selection steps.

After channels selection, we notice that the bipolar channels Vb7, Vb8 and Vb9 are the most repetitive channels using the relieff method (respectively Vb7 and Vb8 for F-score). For

monopolar channels the result indicates that the most repetitive channels are CH9, CH10 using relieff method (respectively CH2, CH10 for F-score). The position of these selected monopolar and bipolar channels are in the median vertical axis of the abdomen (except for CH2). This is in agreement with the literature that indicates that the best electrode position was the median vertical axis of the woman's abdomen [20]. This best result could be explained by the fact that this position corresponds to the location where the electrodes are closer to the uterus (due to the specific morphology of pregnant woman's abdomen) [21]. Furthermore, during contraction, the uterus tends to become more circular in its transversal section, and to tip over the abdominal wall. This effect gets the electrodes closer to the median vertical axis. This electrode location is therefore confirmed as the best position to obtain good results in a monovariate approach.

It is clear from Table 3.4 that the percentage of correct classification obtained using BPSO-KNN-Holdout algorithm is higher using bipolar channels (97.35%) rather than the monopolar ones (96.46%).

Additionally, from the validation study, we notice from Table 3.5 that the feature subset selected, S5 (see the subsets in appendix C.2 Table C.2.1), by using the bipolar channels (3 bipolar channels selected Vb7, Vb8, Vb9) has given a higher percentage of correct classification than the monopolar one. Finally, the classification performances expressed in terms of Sensitivity, Specificity and Accuracy (Table 3.6) are better for bipolar signals rather than for monopolar.

We can conclude from these results that for Monovariate analysis, the best results are obtained by using bipolar signals recorded on the median vertical axis of the abdomen.

3.3 Channel combination selection for EHG bivariate analysis

3.3.1 Calculated Features

Five features related to EHG propagation are used for the bivariate analysis. These features are: Linear (R^2) [22], nonlinear correlation coefficient (H^2) [23], Phase synchronization (γ) [24], Filtered-Windowed-H^2 (FW_H^2) [25], and Filtered-Windowed-H (FW_H) [25]. The details of calculation of these features have already been described in chapter 2 section 2.5.1.

Each feature related to EHG propagation is computed between a combination of two different channels (bipolar or monopolar). Our multichannel system being composed of 16 monopolar

channels or 12 bipolar channels, we thus have 132 possible bipolar channel combinations (number of combinations=number of channels^2-number of channels) and 240 possible monopolar channel combinations. Two data matrices of three dimensional are thus obtained when computing the 5 features for the whole combinations: the first one corresponds to the values of the features extracted for all pregnancy and labor contractions (479 contractions) for the monopolar combinations (479*240*5); the second one corresponds to the features extracted for all pregnancy and labor contractions corresponding to bipolar combinations (479*132*5).
As during the previous study, these datasets are parted in two: 379 contractions for the selection steps, 100 contractions for the validation.

3.3.2 Channel Combination selection followed by feature selection

Due to this huge number of channel combinations, available to compute propagation/coherence features, select channel combinations with the best discrimination ability between pregnancy and labor contractions, should be very important in order to reduce the large number of inputs of the classifier.

We used here the same methodology as the one described in section 3.2.2 in order to select the best channel combinations (for bipolar Vb(i,j) and monopolar CH(i,j) signals) followed by feature selection, once the channel combinations selected.

3.3.3 Results

3.3.3.1 Results of channel combinations using Relieff and F-score

Monopolar and bipolar datasets are used in this section. The first monopolar dataset is a matrix of dimension 379*240*5 and the other bipolar dataset is of dimension 379*132*5 (as explained in section 3.3.1).

F-score and Relieff methods are used for channel combinations selection (bipolar or monopolar) by calculating the best discrimination ability between labor and pregnancy contractions using a specific feature. In our work 5 features related to EHG propagation are extracted from our EHG, thus 5 subsets of selected channel combinations are obtained for each method using each dataset (bipolar or monopolar). All the results of channel combination selection are presented in Appendix C.4.

Figure 3.2 (A-B-C-D-E) presents the results of bipolar channel combinations selection using F-score for the 5 features. The results are presented as a black and white matrix of dimensions 12x12. A white color indicates that the bipolar channel combination has been selected as discriminant between pregnancy and labor. Figure3.2-F presents the number of appearance over the 5 features for each bipolar channel combination. This figure is in the form of a color matrix of dimension 12x12. The red color in this figure represents the maximum number of appearance over the 5 features of bipolar channel combination and the blue color its minimum number (0= blue, 5= red). The results for bipolar channel combinations selection using Relieff, as well as the results of monopolar channel combinations selection using F-score and Relieff are presented in appendix C.4 (Figure C.4.2, Figure C.4.1, and Figure C.4.3 respectively). The number of appearance over the five features for each of them are presented appendix C.4 Table C.4.1 (Bipolar signals) and Table C.4.2 (Monopolar signals).

The channel combinations will finally be retained according to their number of appearance over the five features for each selection method. We first sorted the selected channel combinations following their appearance number (for bipolar and for monopolar signals separately). We then retained the first 10% of the number of combinations (13 for bipolar and 24 for monopolar combinations) corresponding to the higher appearance number over the 5 features. When the appearance number of the last selected channel combination is equal to the number of appearance of the following ones, we take into account all following channel combinations having the same appearance number (for example the number of selected bipolar channel combinations using F-score is equal to 17 instead of 13, as presented in Appendix C.4 table C.4.1). Figure 3.3 (A, B) presents the most repetitive bipolar channel combinations retained for F-score and relieff respectively (A white color indicates that the bipolar channel combination has been retained). The most repetitive monopolar channel combinations retained for F-score and relieff are presented in figure 3.4 (A, B).

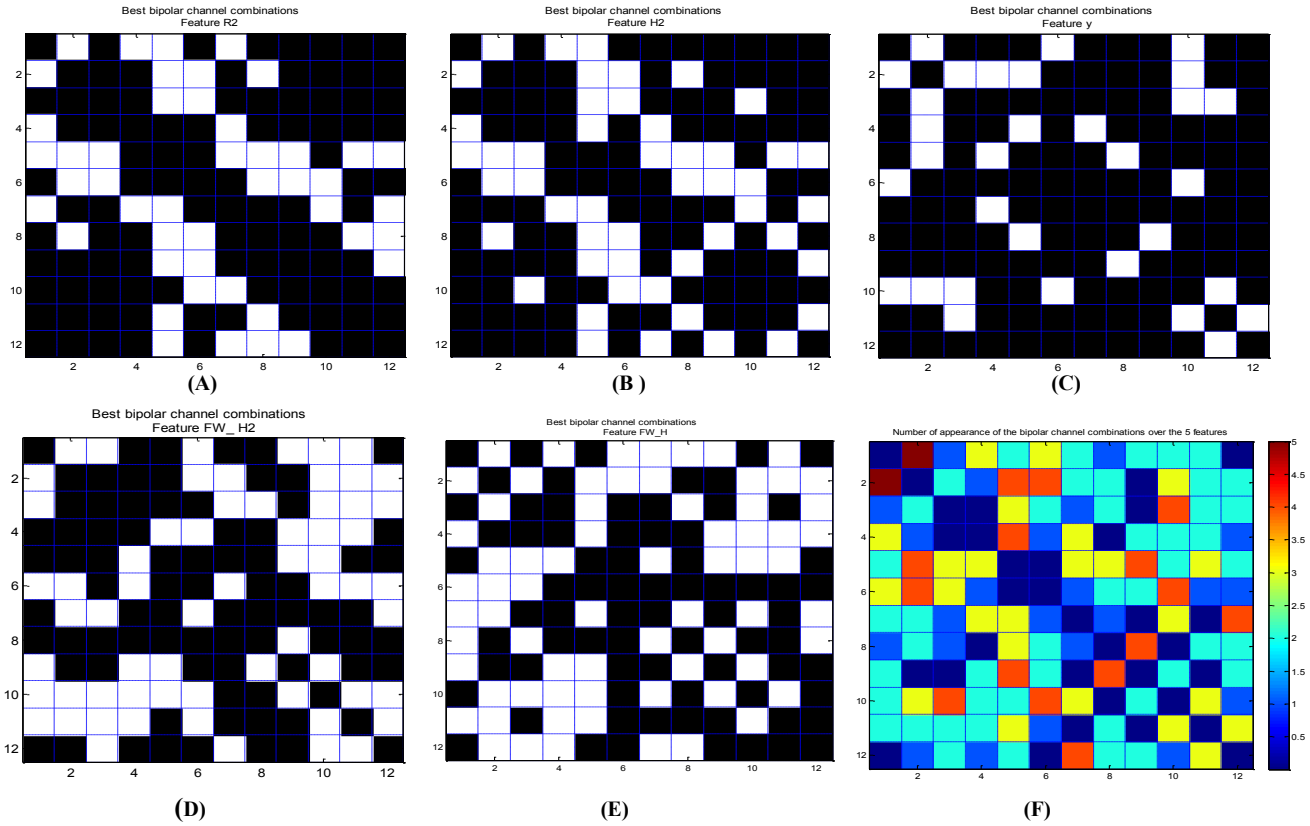


Figure 3.2: Bipolar channel combinations selection by F-score using five features related to EHG propagation: R^2(A), H^2 (B), y (C), FW_H^2 (D) and FW_H (E). Each axis (vertical and horizontal) represents the channel numbers. A white square indicates that the related combination is selected for the given feature. (F) Number of appearance of the bipolar channel combinations over the 5 features.

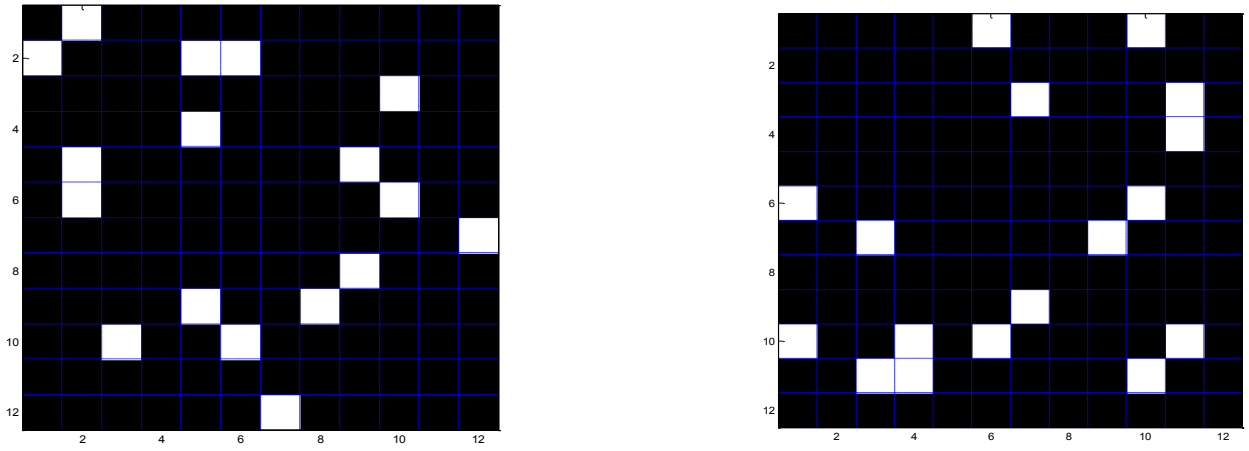
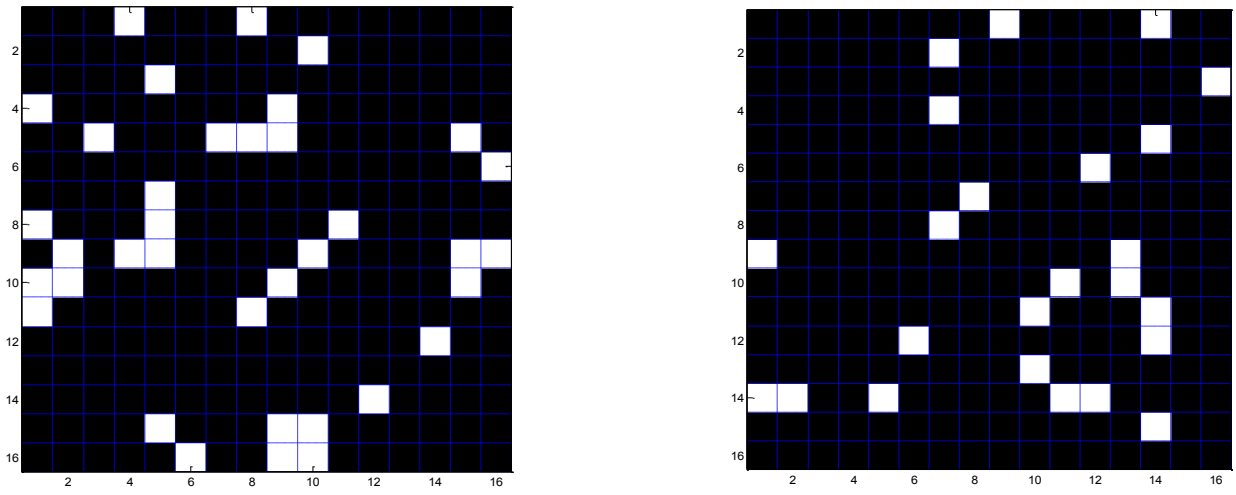


Figure 3.3: Most repetitive bipolar channel combinations retained for F-score (A) and relieff (B) in the 5 features related to EHG bivariate analysis.



(A) **(B)**

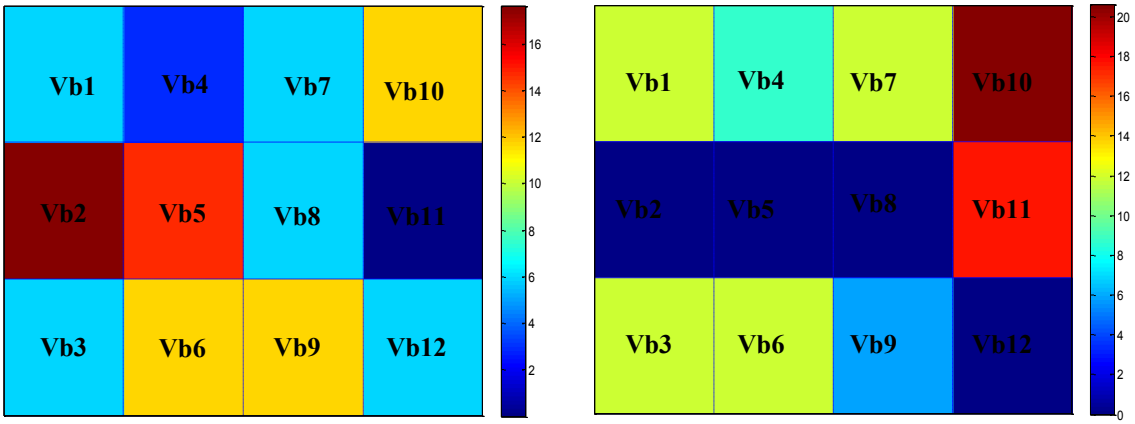
Figure 3.4: Most repetitive monopolar channel combinations retained for F-score (A) and relieff (B) in the 5 features related to EHG bivariate analysis

Table 3.7 contains the name of the retained bipolar and monopolar channel combinations over the 5 features for F-score and Relieff (presented in Figure 3.3 and figure 3.4).

We then computed, for each channel (Bipolar, Vbi, or Monopolar, CHi) the number of times it appears in the channel combinations retained. Figure 3.5 presents two color matrices of dimension 12x12 representing the topologic distribution of the number of bipolar channels associated to each retained bipolar channel combination for F-score (figure 3.5-A) and relieff (Figure 3.5-B). The two 16x16 color matrices corresponding to the retained monopolar channel combinations for F-score and relieff are presented in Figure 3.6-A and Figure 3.6-B respectively.

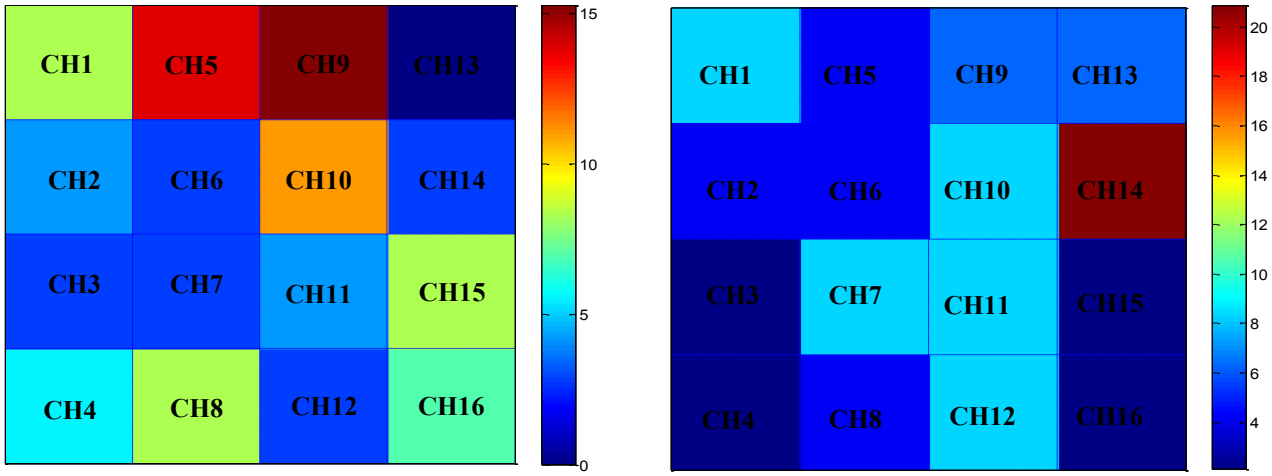
Table 3.7: Name of the retained bipolar and monopolar channel combinations over the 5 features related to EHG bivariate analysis for F-score and Relieff

Methods	Bipolar Channel combinatons	Monopolar Channel combinatons
F-score	['Vb(2,1)', 'Vb(1,2)','Vb(12,7)', 'Vb(10,6)','Vb(10,3)','Vb(9,8)', 'Vb(9,5)','Vb(8,9)','Vb(7,12)','Vb(6,10)','Vb(6,2)','Vb(5,9)','Vb(5,2)','Vb(4,5) ','Vb(3,10)','Vb(2,6)','Vb(2,5)']	['Ch(10,2)','Ch(8,1)','Ch(2,10)','Ch(1,8)', 'Ch(16,10)','Ch(16,9)','Ch(16,6)','Ch(15,10)', 'Ch(15,9)','Ch(15,5)','Ch(14,12)','Ch(12,14)', 'Ch(11,8)','Ch(11,1)','Ch(10,15)','Ch(10,9)', 'Ch(10,1)','Ch(9,16)','Ch(9,15)','Ch(9,10)', 'Ch(9,5)','Ch(9,4)','Ch(9,2)','Ch(8,11)', 'Ch(8,5)','Ch(7,5)',‘Ch(6,16)','Ch(5,15)', 'Ch(5,9)','Ch(5,8)','Ch(5,7)','Ch(5,3)', 'Ch(4,9)','Ch(4,1)','Ch(3,5)','Ch(1,4)']
Relieff	['Vb(11,10)','Vb(11,4)','Vb(11,3)', 'Vb(10,11)','Vb(10,6)','Vb(10,4)', 'Vb(10,1)','Vb(9,7)','Vb(7,9)', 'Vb(7,3)','Vb(6,10)','Vb(6,1)', 'Vb(4,11)','Vb(3,11)','Vb(3,7)', 'Vb(1,10)', 'Vb(1,6)']	['Ch(15,14)','Ch(14,12)','Ch(14,11)', 'Ch(14,5)','Ch(14,2)','Ch(14,1)', 'Ch(13,10), 'Ch(12,14)','Ch(12,6)','Ch(11,14)','Ch(11,10) ','Ch(10,13)','Ch(10,11)','Ch(9,13)','Ch(9,1)', 'Ch(8,7)','Ch(7,8)','Ch(6,12)','Ch(5,14)', 'Ch(4,7)'Ch(3,16)','Ch(2,7)','Ch(1,14)', 'Ch(1,9)']



(A) **(B)**

Figure 3.5: Color matrix of dimension 12x12 representing the distribution of the number of retained bipolar channels combinations using F-score (A) and relieff (B)



(A) **(B)**

Figure 3.6: Color matrix of dimension 16x16 representing the distribution of the number of retained monopolar channel combinations for F-score (A) and relieff (B*)*

3.3.3.2 Results of feature selection using BPSO and GA

The same procedure as used in section 3.2.3.2 is used here in order to select the best subset of features from the selected channel combinations, by means of the two feature selection methods BPSO and GA. As we have obtained four subsets of channel combinations, we used four datasets in this part (Table 3.8). The datasets DC1 and DC2 contain the 5 features extracted, for the 379 EHG bursts, from the 17 retained bipolar channel combinations for F-score (379*85) and the 17 retained for Relieff (379*85) respectively. Therefore, the particle length of BPSO and length of chromosome of GA are equal to 85 (17 bipolar channel combinations*5 features). The monopolar datasets DC3 and DC4 correspond to the features extracted, for the 379 EHG bursts, from the 36 retained monopolar channel combinations for F-score (379*180) and the 24 retained monopolar channel combinations for Relieff (379*120). The particle length for BPSO and length of chromosome of GA for these two last datasets are equal to 180 and 120 respectively. The selected feature subsets obtained after applying BPSO and GA (using the classifier KNN and the two data split KFOLD and Holdout) to these 4 bipolar and monopolar datasets (DC1, DC2, DC3 and DC4) are presented in appendix C.5 (table C.5.1). Table C.5.2, appendix C.5, presents the number of features selected for each subset obtained and the time of calculation of each method.

Table 3.8: Datasets obtained for the retained bipolar and monopolar channel combinations, for F-score and Relieff

Methods for channel selection	Bipolar		Monopolar	
	Number of Bipolar channel combinations	**Dataset**	**Number of monopolar channel combinations**	**Dataset**
F-score	17	DC1= 379*85	36	DC3= 379*180
Relieff	17	DC2= 379*85	24	DC4= 379*120

Table 3.9 presents the two feature subsets giving the maximum percentage of correct classification from bipolar and monopolar channel combination (Appendix C.5, Table C.5.1).

Table 3.9: Subsets of features selected from the bipolar and monopolar selected channel combinations with highest percentage of correct classification using KNN

Datasets	Methods	Dataset	Selected feature subset	Number of features	Percentage of correct classification	Time (s)
Bipolar	**Bipolar, F-score, BPSO-KNN-Holdout**	DC1	SC4	45	95.58	7.68
Monopolar	**Monopolar, Relieff , GA-KNN-Holdout**	DC4	SC14	35	94.69	27.17

3.3.3.3 Validation

We try in this part to evaluate the performances of these 16 selected subsets (Appendix C.5, Table C.5.1) on the remaining contractions (50 labor contractions and 50 pregnancy contractions), not used during the feature selection steps. We use the classifier KNN and the data split KFOLD or holdout (depending to the data split used in feature selection for each obtained subset). We computed for each subset the mean ± standard deviation of the percentages of correct classification (over 500 repetitions). The results obtained in this validation step are presented in Appendix C.6 (Table C.6.1). The two subsets of feature selected on monopolar and bipolar EHG giving the higher mean ± standard deviation of the percentages of correct

classification are presented table 3.10. Table 3.11 presents the results of Mean ± standard deviation of sensitivity, specificity and accuracy for these 2 subsets.

Table 3.10: Selected feature subset from the bipolar and monopolar selected channel combinations with highest Mean ± STD of the percentage of correct classification using KNN of 500 repetitions

Datasets	Methods	Selected feature subset	Number of features	Mean ± standard deviation of percentage of correct classification (500 repetitions)
Bipolar	**Bipolar, F-score, GA-KNN-KFOLD**	SC1	25	76.01 ±1.33
Monopolar	**Monopolar, F-score, GA-KNN-KFOLD**	SC9	55	74.85±2.26

Table 3.11: Performance of the selected feature subset from the bipolar and monopolar selected channel combinations with highest Mean ± STD of the percentage of correct classification using KNN of 500 repetitions

Datasets	Methods	Performance of classifier		
		Sensitivity	**Specificity**	**Accurancy**
Bipolar	**Bipolar, F-score, GA-KNN-KFOLD**	76.76±2.72	76.21±02,87	76.02±1.34
Monopolar	**Monopolar, F-score, GA-KNN-KFOLD**	85.18±3.06	65.84 ±4.11	74.56 ±2.42

3.3.4 Discussion

In the second part of this study, we presented two methods (F-score and Relieff) to select the best bipolar and monopolar channel combinations permitting to discriminate at best pregnancy and labor classes, for a bivariate approach. After this combination selection, two feature selection methods were used to select the best features related to EHG bivariate analysis from the selected channel combinations. Additionally, we applied these selection steps on bipolar and monopolar EHG signals in order to see which is the best to use for the bivariate analysis.

It is clear from Figure 3.3 and table 3.7, for bipolar channels, from Figure 3.4 and table 3.7, for monopolar channels, as well as from figure 3.5 and 3.6, that the number and location of bipolar

and monopolar channels selected in the obtained subsets vary from one method to the other. The reason is that these two filter methods originate from different categories. Relieff selects the most discriminant channel combinations with some degree of integration of channel combinations dependency. Therefore this method considers all channel combinations simultaneously. While F-score method does not take into account the dependency between channel combinations. It selects channel combinations by adding, deleting or comparing one channel combination at a time.

It is clear from the results obtained either for the feature selection (Table 3.9) of for the validation (Tables 3.10 and 3.11) steps, that the results obtained from Bipolar or Monopolar datasets present similar performances in terms of percentage of correct classification and accuracy. The sensitivity is slightly higher for the monopolar signals and the specificity slightly higher for the bipolar ones. These results differ from the monovariate analysis where the best results were always obtained from bipolar signals. Monopolar signal use seems thus to be more pertinent for bivariate analysis.

3.4 Discussion and Conclusion

According to the results obtained in the monovariate analysis, we notice that the bipolar and monopolar selected channels are positioned on the vertical axis of the women's abdomen (except for one channel, CH2). The percentage of correct classification for bipolar channel selection followed by features selection is higher than monopolar one in the selection as well as on the validation steps. The two subsets corresponding to this higher percentage (S8 and S5) are obtained from the dataset D2 that contain the features selected from the 3 bipolar channels Vb7, Vb8 and Vb9 by using Relieff.

For the Bivariate analysis, channel combination selection, the bipolar and monopolar combinations selected correspond to different channels for the 2 methods (F-score, Relieff). No specific channels or channel configuration appear to give always the best results. Additionally, the classification performances obtained for this bivariate analysis with bipolar signals are similar to the ones obtained with monopolar signals even if the highest percentage of correct classification is obtained from the bipolar channel combinations, selected here by F-score.

Relieff performing better for monovariate analysis, and F-score for the bivariate one, concerning the selection methods, no one of the two methods, appear to give always the best results for all

the selection steps.

Concerning channel location:

We can conclude that, for the monovariate analysis, the bipolar as well as the monopolar channels offer better discrimination ability between labor and pregnancy when positioned in the median vertical axis of the woman's abdomen. This is fortunately the standard position defined in our standardized recording protocol. Nevertheless, no conclusion can be given concerning the best channels or the best combination of channels to use for this bivariate analysis, no trend appearing for the channels (Bipolar or monopolar) selected for the bivariate analysis.

Concerning recording configuration:

From the monovariate analysis, the higher percentage of correct classification obtained in the selection part and validation part with the bipolar channels, permits us to conclude that using bipolar EHG for monovariate study gives better results than using monopolar EHG. The differentiation induced by the bipolarization removes indeed the common low frequency noise and enhances thus the high frequency content of the signal. As an increase in the excitability of uterine cells is associated to a shift of FWh (higher frequency content of the EHG) towards higher frequencies [26], using bipolar signal permits to focus on this FWh content of the signal, increasing thus the excitability part of the EHG.

From bivariate analysis, the hypothesis was that using monopolar signals permit to get free from the bias induced by using a common monopolar electrode to create two adjacent bipolar channels. We expected thus better results for bivariate analysis when using monopolar channels. We can conclude from this bivariate study that the use of monopolar EHG gives similar results than the ones obtained from bipolar signals. As the Monopolar signals induce no bias, we can go on using monopolar signals for bivariate analysis.

No conclusion can be given at this time concerning the best method to use for channel selection whatever the situation: monovariate/bivariate, bipolar/monopolar. Further studies should be developed to conclude on this aspect.

References

[1] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, "A multisensor data fusion approach for improving the classification accuracy of uterine EMG signals," in *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 93–96.

[2] B. Moslem, M. O. Diab, C. Marque, and M. Khalil, "Classification of multichannel uterine EMG signals," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*,pp. 2602–2605, 2011.

[3] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, "Classification of multichannel uterine EMG signals by using unsupervised competitive learning," in *Signal Processing Systems (SiPS), 2011 IEEE Workshop on*, 2011, pp. 267–272.

[4] B. Moslem, M. Khalil, M. O. Diab, A. Chkeir, and C. Marque, "Combining multiple support vector machines for boosting the classification accuracy of uterine EMG signals," in *Electronics, Circuits and Systems (ICECS), 2011 18th IEEE International Conference on*, 2011, pp. 631–634.

[5] B. Moslem, M. O. Diab, M. Khalil, and C. Marque, "Classification of multichannel uterine EMG signals using a reduced number of channels," in *Mechatronics and its Applications (ISMA), 2012 8th International Symposium on*, pp. 1–4, 2012.

[6] B. Moslem, M. Diab, M. Khalil, and C. Marque, "Combining data fusion with multiresolution analysis for improving the classification accuracy of uterine EMG signals," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, pp. 1–9, Aug. 2012.

[7] I. Kononenko, "Estimating attributes: analysis and extensions of RELIEF," *in Machine Learning: ECML-94*, 1994, pp. 171–182.

[8] S. Ding, "Feature Selection Based F-Score and ACO Algorithm in Support Vector Machine,"*in Second International Symposium on Knowledge Acquisition and Modeling, 2009. KAM '*09, 2009,Vol. 1, pp.19-23.

[9] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," *IEEE International Conference on Systems, Man, and Cybernetics*, October 1997, vol. 5, pp. 4104–4108.

[10] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi, "Application of the PSOSVM model for recognition of control chart patterns," *ISA Transactions*, vol. 49, no. 4, pp. 577–586, 2010.

[11] B. Oluleye, L. Armstrong, L. Jinsong, and D. Diepeveen, "Zernike moments and genetic algorithm: Tutorial and application," *Br. J. Math. Comput. Sci.,* 2014.Vol. 4, no. 15, pp. 2217–2236.

[12] T. Jérémy, S. Thora, M. Catherine, K. Brynjar, and others, "Synchronization between EMG at different uterine locations investigated using time-frequency ridge reconstruction: comparison of pregnancy and labor contractions," *EURASIP J. Adv. Signal Process.*, vol. 2010, Article ID 242493, 2010.

[13] J. Sikora, A. Matonia, R. Czaba´nski, K. Horoba, J. Jezewski, and T. Kupka, "Recognition of premature threatening labour symptoms from bioelectrical uterine activity signals," *Archives of Perinatal Medicine*, vol. 17, no. 2, pp. 97–103, 2011.

[14] C. Marque, H. Leman, M. L. Voisine, J. Gondry, and P. Naepels, "Traitement de l'électromyogramme utérin pour la caract´erisation des contractions pendant la grossesse," *RBMNews*, vol. 21, no. 9, pp. 200–211, Dec. 1999.

[15] G. Fele-ˇZorˇz, G. Kavˇsek, ˇZ. Novak-Antoliˇc, and F. Jager, "A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and preterm delivery groups," *Medical and Biological Engineering and Computing*, vol. 46, no. 9, pp. 911–922, 2008.

[16] M. O. Diab, C. Marque, and M. A. Khalil, "Classification for uterine EMG signals: comparison between AR model and statistical classification method," *International Journal of Computational Cognition*, vol. 5, no. 1, pp. 8–14, 2007.

[17] A. Diab, M. Hassan, C. Marque, and B. Karlsson, "Quantitative performance analysis of fourmethods of evaluating signal nonlinearity: application to uterine EMG signals," in *Proceedings of the 34th Annual International IEEE EMBS Conference*, San Diego, Calif, USA, September 2012,pp. 1045–1048.

[18] M. Hu and H. Liang, "Variance entropy: a method for characterizing perceptual awareness of visual stimulus," *Applied Computational Intelligence and Soft Computing*, vol. 2012, pp.1-6, 2012.

[19] B. Moslem, M. Khalil, M. O. Diab and C. Marque, "Detrended fluctuation analysis of uterine electromyography," in *First Middle East Conference on Biomedical Engineering, MECBME11*, Sharjah, UAE, 2011.

[20] C. K. Marque, J. Terrien, S. Rihana, and G. Germain, "Preterm labour detection by use of a biophysical marker: the uterine electrical activity," *BMC Pregnancy Childbirth*, vol. 7, no. Suppl 1, pp. S5, Jun. 2007.

[21] J. Terrien, S. Rihana, J. Gondry, and C. Marque, "Modeling the effects of the electrodes position on the EMG characteristics: consequences for the monitoring of uterine contractions.," *in IFAC (International Federation of Automatic Control) MCBMS (Modelling and Control in Biomedical Systems)*, Reims, Septembre 2006.

[22] K. Ansari-Asl, F. Wendling, J. J. Bellanger, and L. Senhadji, "Comparison of two estimators of time-frequency interdependencies between nonstationary signals: application to epileptic EEG," *in 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2004. IEMBS '04*, 2004, vol. 1, pp. 263–266.

[23] M. Hassan, A. Alexandersson, J. Terrien, C. Muszynski, C. Marque, and B. Karlsson, "Better pregnancy monitoring using nonlinear propagation analysis of external uterine electromyography", *IEEE Transactions on Biomedical Engineering*, Vol. 60, No. 4, 2013,pp. 1160–1166.

[24] M. Hassan, J. Terrien, A. Alexandersson, C. Marque, and B. Karlsson, "Improving the classification rate of labor vs. normal pregnancy contractions by using EHG multichannel recordings," *Presented at the 32 nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Buenos Aires, Espagnol, Aug.2010.

[25] A.DIAB, “Study of The Nonlinear Properties And Propagation Characteristics Of The Uterine Electrical Activity During Pregnancy And Labor”, Ph.D. dissertation, Thèse de l’université de Technologie de Compiègne, 2014.

[26] D. Devedeux, C. Marque, S. Mansour, G. Germain, and J. Duchêne, “Uterine electromyography: a critical review,” Am. J. Obstet. Gynecol., vol. 169, no. 6, pp. 1636-53, Dec. 1993.

General conclusion and perspectives

We have presented in this thesis several approaches in order to improve the early prediction of term or preterm labor by analysis of EHG signals. Our main aim is the selection of the best features obtained from both monovariate and bivariate analyses of the uterine electrical activity to classify uterine contractions, in order to improve the performance of previously tools used for the clinical diagnosis of preterm labor. Our secondary aim is the selection of the best recording condition, addressed in terms of recording derivation (bipolar/monopolar) and channel location (electrode position).

Using abdominal electrodes to detect electrohysterogram EHG is not new, and the analysis of the uterine excitability and the propagation of activity processing a large number of EHGs with different tools (temporal representation, frequency, time-frequency, non-linear characteristics) have been the goal of many studies during the last 20 years. These studies used very different populations and various recording protocols. Additionally, using this large number of signal processing tools increases the computational complexity in a diagnostic objective. In addition, all previously conducted studies on EHG have used only bipolar derivation (in order to increase the SNR). The recording of EHG began using a small number of electrodes (giving usually 1 or 2 bipolar derivations). But due to recent hardware developments, a large number of electrodes are now used. Therefore extract information from these channels makes complexity for a diagnostic purpose even greater. This thesis has thus mainly addressed the following points:

- Enrich the already available database by new recordings made on pregnant women to test our work on a larger standardized database.
- Extract from a bibliographic review, the most representative features to study, from the EHG, uterine excitability (monovariate analysis) and propagation (bivariate analysis).
- Using several feature selection methods to select, from the previously define set of features, the most relevant features computed on our EHG database. Comparing the classification results with and without feature selection to quantify the gain provided by this selection.

- Selecting the most discriminating channels using linear and nonlinear features (monovariate analysis).
- Selecting the most discriminating channel combinations using features related to EHG propagation (bivariate analysis).
- Testing if discrimination is more efficient using the two types of information (excitability, propagation).
- Testing our work using monopolar signals instead of bipolar signals currently used to see if the use of monopolar signals (specially for the bivariate analysis) improves the results.

The results of our work are presented in the second and third chapter that investigated respectively the selection of relevant features (excitability and propagation) and then the channel and channel combination selection for monovariate and bivariate analysis on uterine electrical activity.

Chapter 2 presents the study done for the selection of relevant feature subsets, combining monovariate and bivariate features. For this purpose, several features selection methods were applied in this work. First, an algorithm for the selection of parameters using the feature selection method named Jeffrey Divergence was developed in order to select the best parameters computed both on the original EHG and on different frequency bands obtained using wavelet packet decomposition. Our results indicate that, 13 linear parameters (*W1*, *W2*, *W3*, *W4*, *D1*… *D8* and *MPF*), 3 nonlinear parameters (*Tr*, *SE*, *VarEn*) and 2 propagation parameters (R^2, H^2) demonstrate ability to discriminate pregnancy and labor contractions. Some of them are more efficient when extracted from the original EHG; others are more efficient when computed from different frequency bands.

Due to complexity of the calculation presented in the part above, we decided to use only the original EHG in the second part of this work. We thus tested several feature selection methods to choose the best subsets for classifying labor and pregnancy contractions. Two types of feature selection were used in this part (filter and wrapper). For wrapper method, we used the classifier KNN and two classical datasplit (KFOLD and Holdout). The results of each subset were compared using the percentage of correct classification. Additionally, we compared the results of feature selection in this part by using bipolar and then monopolar EHG signals.

After extraction of the most pertinent features from bipolar then monopolar EHGs, the result showed that the monovariate nonlinear feature (*DFA*) is always selected, whatever the derivation (monopolar/bipolar), and the feature selection method. *DFA* increases passing form pregnancy to labor on monopolar as well as on bipolar EHG. These results are in agreement with different studies concerning the interest of the nonlinear characteristics of EHG for pretem labor detection. Opposite to *DFA*, the feature *LE* was not selected neither on bipolar nor on monopolar EHGs, which makes it a very bad candidate for discrimination between pregnancy and labor. Additionally, the results showed that some other features, selected several times among the whole methods, are also of considerable potential interest for the diagnosis. For example on Bipolar EHG, *MPF* was selected 17 over 20 times. The pregnancy/labor evolutions of this parameter, showed an increasing when going from pregnancy to labor, which is in agreement with different studies that evidenced a shift of the EHG spectral content towards higher frequencies when going from pregnancy to labor.

For the bivaraiate features, the results show that H^2 was the most pertinent bivariate feature, on bipolar signals, and *FW_H*, on monopolar signals. The results showed a decrease in both H^2 and *FW_H* when passing from pregnancy to labor, which is one of the unexpected results. Some additional work must be done to explain these results. One possible hypothesis is that the synchronization of the uterine muscle, occurring during labor, is not solely related to an electrical phenomenon. The electrical propagation through the diffusion process could be mainly local, as evidenced by the recent work of Rabotti et al. [1]. But the main global synchronization of the whole uterine muscle should be due to the hydrodynamic-stretch activation mechanism, recently proposed by Young [2]. This new hypothesis is presently under investigation in our lab.

From the validation part and after calculating the percentage of correct classification on all obtained feature subsets, using bipolar and monopolar EHG, the results show that the most feature subsets selected by wrapper methods give better results for the classification between pregnancy and labor contractions, due to use of a classifier within the method. Additionally, the result show that using bipolar EHG gives a better classification than using monopolar. Therefore, we can conclude that, when mixing monovariate and bivariate approach, using bipolar is better than monopolar. Furthermore, we compared the results of classification with and without feature

selection. The obtained results demonstrated that the classification with selection is better than without selection specially when using bipolar signal.

Finally, we presented the results concerning the selection, from the multichannel system, of the most relevant channels (monovariate analysis) and the best combination of channels (bivariate analysis) for classification of pregnancy/labor contractions. The same procedure was used for channel selection for the monovariate analysis, and channel combination selection for the bivariate analysis. Two methods of feature selection of type filter (Relieff and F-score method) have been adapted to select channels and combination channels. We then applied two feature selection methods of type wrapper (Binary particle swarm optimization and Genetic algorithm (GA)) in order to select the best monovariate features from the selected channels, and the best bivariate features from the selected channel combinations. We then tested the performance of our obtained results in a validation step.

For the monovariate analysis, results show that the most pertinent selected bipolar and monopolar channels are positioned on the vertical axis of women's abdomen. Additionally, the percentages of correct classification obtained after bipolar channel and feature selections were higher, for selection and validation as well, than the ones obtained for monopolar channel and feature selections.

For the bivariate analysis, from the results obtained for channel combination selection, we were not able to evidence any specific channel location giving the best results. Additionally, the percentage of correct classification obtained in selections and validation part were similar for bipolar or monopolar channel combinations selection followed by feature selection, proving that for bivariate analysis, monopolar signals behave better than for monovariate analysis.

We can conclude from these results that using bipolar and monopolar channels positioned in the median vertical axis of women's abdomen gives good discrimination ability between the two classes of contractions (pregnancy and labor). Luckily, this position is in agreement with our recording protocol. While for bivariate analysis, we did not found any specific best channels location to be used for the channel combinations.

For the monovariate analysis, we can conclude from this work that using bipolar EHG gives better results than using monopolar ones, for the discrimination between pregnancy and labor.

These results are linked to the bipolarization, which removes the common low frequency noise and improves the high-frequency content of the signal. An increase in the excitability of uterine cells is associated to a shift of FWh (higher frequency content of the EHG) towards higher frequencies. Therefore, using bipolar signal allows focusing on this FWh content of the signal, increasing thus the excitability part of the EHG. Use of bipolar signals makes thus easier to evidence this increase.

For the bivariate analysis, we can conclude from this work that using bipolar EHG gives similar results than using monopolar ones, for the discrimination between pregnancy and labor contractions. As the monopolar signals induce no bias (opposite to the bipolar signals that use a common monopolar electrode to create two adjacent bipolar signals), we can advice to use monopolar signals for bivariate analysis.
Another way to process (that we did not have time to test in our work) should be to use bipolar signals to compute the monovariate features (thanks to an increase in SNR) and monopolar signals to compute the bivariate features (thanks to a decrease bias and an increased spatial resolution). These tests should be done soon in our lab.

No definite conclusion can be given at the moment regarding the best method to use for channel selection regardless of the experimental situation: monovariate/bivariate, bipolar/monopolar. Our results only permitted us to evidence that our reference position (Vb7, CH9, CH10) corresponds to one of the best identified for monovariate analysis. Therefore, other studies shoul be done for the bivariate analysis.

The results presented in this thesis can be improved by further works:

- Combine several types of methods for feature selection in order to increase the selection performance.
- Increase the actual database of EHG signals by recordings made on pregnant women during different physiological and pathological situations (normal and risk pregnancies, term and preterm labor) in order to test our results on a larger database.
- Use the linear and nonlinear features extracted from the selected channel and the features related to EHG propagation selected from channels combinations as input for a classifier in order to increase the percentage of classification between labor and pregnancy classes.

- Apply for the validation step of these selections, other classification methods on the selected features to classify pregnancy and labor contractions.
- Develop tools to test the new hypothesis of uterine synchronization. This could be done through a graph theory approach, combined twith a multiscale multiphysic (electrical/mechanical) biophysical model of the uterine contractile activity.

Our final aim should be to be able to integrate all these knowledge in reliable detection tools of the early signs of preterm labor.

References

[1] C. Rabotti and M. Mischi, "Propagation of electrical activity in uterine muscle during pregnancy: a review," *Acta Physiol*., vol. 213, no. 2, pp. 406–416, 2015.

[2] R. C. Young, "Myocytes, myometrium, and uterine contractions," *Ann. N. Y. Acad. Sci*., vol. 1101, no. 1, pp. 72–84, 2007.

Appendix A: Database

Table A.1: Information for women in pregnancy (database 1)

Woman	Record	Place of Record	Weight	Height	Term (*)	Placenta (**)	Number of contractions	Term at delivery
W1	R1	Island	71	1.75	32+5	Post	6	41+2
	R2	Island	71	1.75	35+5	Post	6	
	R3	Island	71	1.75	37+5	Post	4	
W2	R1	Island	61	1.75	35+5	Ant	6	40+2
	R2	Island	61	1.75	39+2	Ant	5	
	R3	Island	61	1.75	40+1	Ant	5	
W3	R1	Island	62	1.65	33+1	Post	3	39
W4	R1	Island	48 - 50	1.6	29+5	Post	2	41+4
	R2	Island	48 - 50	1.6	31+5	Post	2	
W5	R1	Island	75	1.72	36	Ant	2	40+1
	R2	Island	75	1.72	38	Ant	3	
W6	R1	Island	70 - 75	1.76	33	Post	4	40+6
	R2	Island	70 - 75	1.76	35	Post	2	
	R3	Island	70 - 75	1.76	37	Post	1	
	R4	Island	70 - 75	1.76	38	Post	2	
W7	R1	Island	56	1.63	40+2	Fundus	9	41+1
W8	R1	Island	100	1.78	32+4	Post	6	40+5
W9	R1	Island	92.4	1.8	35	Ant	1	39+5
	R2	Island	92.4	1.8	37	Ant	3	
	R3	Island	92.4	1.8	38	Ant	1	
	R4	Island	92.4	1.8	39	Ant	2	
W10	R1	Island	67	1.64	34+1	Post	4	38
	R2	Island	67	1.64	36+1	Post	5	
	R3	Island	67	1.64	37+1	Post	3	

W11	R1	Island	76.2	1.7	33+6	xxx	3	xxx
	R2	Island	76.2	1.7	35+6	xxx	4	xxx
	R3	Island	76.2	1.7	36+6	xxx	3	xxx
W12	R1	France	52	1.5	34	Post	4	39
W13	R1	France	57	1.5	36+1	Ant	4	41+3

*: Week + Day

**: Post: posterior, Ant: anterior, Sup: Superior

Table A.2: Information for women in labor (database 1)

Woman	**Record**	**Place of Record**	**Weight**	**Height**	**Term (*)**	**Placenta (**)**	**Number of contractions**	**Term at delivery**
W14	R1	Island	xxx	xxx	37	xxx	28	37
W15	R1	Island	xxx	xxx	39+1	Ant	17	39+1
W16	R1	Island	62	1.63	39+3	Fundus	4	39+3
W17	R1	Island	63.4	1.63	39+3	Ant	7	39+3
W18	R1	Island	57.5	1.66	41+5	Post - lat/left	4	41+5
W19	R1	Island	89	1.7	42+2	Ant	1	xxx
W11	R1	Island	76.2	1.7	37	post/sup	4	xxx
W5	R1	Island	75	1.72	40+1	Ant	1	40+1
W20	R1	Island	109	xxx	39+6	Post	1	39+6
W21	R1	Island	xxx	xxx	40+3	Post	20	40+3
W22	R1	Island	xxx	xxx	42+1	Post	16	42+1
W23	R1	France	62	1.6	41+4	xxx	3	xxx

*: Week + Day

**: Post: posterior, Ant: anterior, Sup: Superior

Table A.3: Information for women in pregnancy (database 2)

Woman	Record	Place of Record	Weight	Height	Term (*)	Placenta (**)	Number of contraction	Term at delivery
W1	R1	Iceland	71	1.75	32+5	Post	7	41+2
	R2	Iceland	71	1.75	35+5	Post	8	
	R3	Iceland	71	1.75	37+5	Post	3	
W2	R1	Iceland	61	1.75	35+5	Ant	7	40+2
	R2	Iceland	61	1.75	39+2	Ant	5	
	R3	Iceland	61	1.75	40+1	Ant	6	
W3	R1	Iceland	62	1.65	33+1	Post	4	39
W4	R1	Iceland	48 - 50	1.6	29+5	Post	2	41+4
	R2	Iceland	48 - 50	1.6	31+5	Post	2	
	R3	Iceland	48 - 50	1.6	34+5	Post	1	
W5	R1	Iceland	75	1.72	36	Ant	2	40+1
	R2	Iceland	75	1.72	38	Ant	3	
W6	R1	Iceland	70 - 75	1.76	33	Post	4	40+6
	R2	Iceland	70 - 75	1.76	35	Post	2	
	R3	Iceland	70 - 75	1.76	38	Post	4	
W7	R1	Iceland	56	1.63	40+2	Fundus	8	41+1
W8	R1	Iceland	100	1.78	32+4	Post	7	40+5
W9	R1	Iceland	92.4	1.8	35	Ant	5	39+5
	R2	Iceland	92.4	1.8	37	Ant	5	
	R3	Iceland	92.4	1.8	38	Ant	6	
	R4	Iceland	92.4	1.8	39	Ant	2	
W24	R1	Iceland	105	1.72	33+3	ant	1	38+1
	R2	Iceland	105	1.72	35+5	ant	5	
	R3	Iceland	105	1.72	37+5	ant	3	
W10	R1	Iceland	67	1.64	34+1	Post	6	38
	R2	Iceland	67	1.64	36+1	Post	7	
	R3	Iceland	67	1.64	37+1	Post	9	

W25	R2	Iceland	76.2	1.64	34+5	ant/sup	5	xxx
W11	R1	Iceland	76.2	1.7	33+6	xxx	6	xxx
	R2	Iceland	76.2	1.7	35+6	xxx	4	
	R3	Iceland	76.2	1.7	36+6	xxx	3	
W26	R1	Iceland	60	1.69	38+0	Ant	13	xxx
W27	R1	Iceland	82.9	1.7	34	Post	1	40+1
	R2	Iceland	85.1	1.7	36	Post	2	
	R3	Iceland	86	1.7	37+2	Post	4	
	R4	Iceland	87	1.7	40	Post	4	
W28	R1	Iceland	68	1.68	33+2	Ant	9	39+1
W29	R1	Iceland	69.5	1.67	30+3	Ant, high	3	38+5
	R2	Iceland	71.7	1.67	35+3	Ant, high	4	38+5
W30	R1	Iceland	95.8	1.62	34	Ant, high	1	39
W31	R1	Iceland	109.9	1.76	37+3	Ant, high	1	41+1
	R2	Iceland	110.9	1.76	38+3	Ant, high	1	
	R3	Iceland	110.9	1.76	39+3	Ant, high	2	
	R4	Iceland	111	1.76	40+2	Ant, high	9	
W32	R1	Iceland	92	1.68	36+5	Post, high	1	38+5
W33	R1	Iceland	85.5	1.68	31+4	Ant, high	10	40
	R2	Iceland	86	1.68	37+2	Ant, high	9	
	R3	Iceland	86.5	1.68	38+2	Ant, high	2	
W34	R1	Iceland	78	1.63	38+5	Post, high	1	41+6
W35	R1	Iceland	113.3	1.73	36	Ant, more on right	3	38+5
W36	R1	Iceland	65.5	1.69	38+2	Ant, high	6	40+3
W37	R1	Iceland	76	1.68	36+4	Ant	4	40+5
W38	R1	Iceland	88	1.76	36+3	Ant, high	1	40+6
	R2	Iceland	89	1.76	38+5	Ant, high	1	
	R3	Iceland	90	1.76	40+5	Ant, high	1	

W39	R1	Iceland	82	1.67	33+2	Ant, high	11	40
	R2	Iceland	83	1.67	35+6	Ant, high	9	
	R3	Iceland	84	1.67	37+3	Ant, high	9	
	R4	Iceland	85	1.67	39+4	Ant, high	9	
W40	R1	Slovenia	69	1.60	41+2	Ant	4	xxx
W41	R2	Slovenia	74	1.66	38+2	xxx	1	xxx
W42	R1	Slovenia	99	164	36+6	Ant	3	xxx
W43	R1	France	xxx	xxx	37+6	xxx	2	xxx
W44	R1	France	55	160	32+3	Ant	7	xxx

*: Week + Day

**: Post: posterior, Ant: anterior, Sup: Superior

Table A.4: Information for women in labor (database 2)

Woman	Record	Place of Record	Weight	Height	Term (*)	Placenta (**)	Number of contraction	Term at delivery
W16	R1	Iceland	62	1.63	39+3	Fundus	4	39+3
W17	R1	Iceland	63.4	1.63	39+3	Ant	7	39+3
W19	R1	Iceland	89	1.7	42+2	Ant	22	xxx
W24	R1	Iceland	105	1.72	38+1	Ant	10	38+1
W11	R1	Iceland	76.2	1.7	37	Post/Sup	5	xxx
W5	R1	Iceland	75	1.72	40+1	Ant	1	40+1
W20	R1	Iceland	109	xxx	39+6	Post	3	39+6
W21	R1	Iceland	xxx	xxx	40+3	Post	26	40+3
W22	R1	Iceland	xxx	xxx	42+1	Post	11	42+1
W45	R1	Iceland	xxx	xxx	40+2	Ant	37	40+2
W46	R1	Iceland	xxx	xxx	39+1	Post	25	39+1
W47	R1	Iceland	xxx	xxx	xxx	xxx	25	xxx
W48	R1	Iceland	95	1.63	39	Ant	1	39
W29	R1	Iceland	72	1.67	38+5	Ant,high	4	38+5
W49	R1	France	xxx	1.62	37+6	Fundus	8	xxx

*: Week + Day

**: Post: posterior, Ant: anterior, Sup: Superior

Appendix B: Feature selection results

B.1- Feature selection results using filter method

Table B.1.1: Distance value for each feature obtained using Jeffrey Divergence method on bipolar and monopolar dataset. Selected features are indicated. in bold font.

Bipolar dataset		**Monopolar dataset**	
Feature names	**Distance (Sorted)**	**Feature names**	**Distance (Sorted)**
'DFA'	**1.6427**	**'DFA'**	**1.3859**
'VarEn'	**1.2384**	**'VarEn'**	**0.9086**
'FW_H'	**0.8810**	**'FW_H'**	**0.7100**
'MPF'	**0.6930**	**'SE'**	**0.5310**
'W2'	**0.5357**	**'MPF'**	**0.4541**
'R^2'	**0.5043**	**'D9'**	**0.4179**
'SE'	**0.4722**	**'W5'**	**0.4070**
'H^2'	**0.3705**	**'W2'**	**0.4045**
'FW_H^2'	**0.3495**	**'R^2'**	**0.3991**
'D9'	0.3224	'D8'	0.3005
'W5'	0.2822	'y'	0.2688
'D7'	0.2319	'H^2'	0.2675
'D8'	0.2121	'W1'	0.2608
'W3'	0.1912	'D6'	0.2571
'LE'	0.1813	'D7'	0.2004
'W1'	0.1561	'FW_H^2'	0.1967
'y'	0.1365	'W3'	0.1564
'PF'	0.1298	'LE'	0.1146
'D1'	0.1246	'D2'	0.0654
'W4'	0.0907	'W4'	0.0537
'D2'	0.0722	'PF'	0.0521
'D4'	0.0689	'D1'	0.0474
'D3'	0.0617	'D3'	0.0383
'D6'	0.0324	'D4'	0.0074
'D5'	0.0172	'D5'	0.0065
'Tr'	0.0001	'Tr'	0.0001

Table B.1.2: F-score value for each feature obtained using F-score method on bipolar and monopolar dataset. Selected features are indicated.in bold font.

Bipolar contractions		**Monopolar contractions**	
Feature names	**F-score (Sorted)**	**Feature names**	**F-score (Sorted)**
'DFA'	**0.6728**	**'DFA'**	**0.6750**
'FW_H'	**0.3226**	**'FW_H'**	**0.3369**
'FW_H^2'	**0.1907**	**'FW_H^2'**	**0.1481**
'VarEn'	**0.1359**	**'VarEn'**	**0.1045**
'MPF'	**0.1163**	**'W2'**	**0.0893**
'W2'	**0.0932**	**'SE'**	**0.0802**
'SE'	0.0753	'R^2'	0.0771
'W1'	0.0703	'W1'	0.0706
'D6'	0.0496	'W5'	0.0684
'D7'	0.0467	'W3'	0.0634
'R^2'	0.0461	'MPF'	0.0602
'D8'	0.0434	'H^2'	0.0413
'D9'	0.0433	'D6'	0.0314
'D5'	0.0398	'D5'	0.0310
'D4'	0.0364	'D4'	0.0288
'D1'	0.0353	'D7'	0.0260
'W3'	0.0345	'D2'	0.0238
'W5'	0.0286	'D3'	0.0231
'D3'	0.0268	'D9'	0.0163
'LE'	0.0236	'LE'	0.0139
'H^2'	0.0194	'D8'	0.0138
'D2'	0.0160	'PF'	0.0112
'PF'	0.0095	'y'	0.0086
'Tr'	0.0039	'D1'	0.0085
'W4'	0.0017	'W4'	0.0061
'y'	0.0015	'Tr'	0.0038

Table B.1.3: Weight value of each feature obtained using Relieff method on bipolar and monopolar dataset. Selected features are indicated in bold font.

Bipolar contractions		**Monopolar contractions**	
Feature names	**Weight (Sorted)**	**Feature names**	**Weight (Sorted)**
'DFA'	**0.0450**	**'DFA'**	**0.0498**
'VarEn'	**0.0340**	**'VarEn'**	**0.0309**
'FW_H'	**0.0228**	**'FW_H'**	**0.0221**
'MPF'	**0.0193**	**'D9'**	**0.0220**
'D9'	**0.0152**	**'MPF'**	**0.0206**
'D8'	**0.0107**	**'D8'**	**0.0154**
'PF'	**0.0103**	**'D7'**	**0.0136**
'y'	**0.0098**	**'W1'**	**0.0127**
'FW_H^2'	0.0086	**'D6'**	**0.0112**
'D2'	0.0079	'W5'	0.0102
'R^2'	0.0071	'W2'	0.0096
'SE'	0.0069	'PF'	0.0082
'W3'	0.0065	'D5'	0.0079
'H^2'	0.0059	'FW_H^2'	0.0076
'W5'	0.0057	'W3'	0.0075
'W1'	0.0055	'D2'	0.0060
'D3'	0.0055	'D4'	0.0056
'D7'	0.0050	'D3'	0.0053
'W2'	0.0045	'D1'	0.0049
'D1'	0.0040	'y'	0.0048
'LE'	0.0034	'H^2'	0.0034
'D4'	0.0023	'SE'	0.0034
'Tr'	0.0022	'W4'	0.0033
'D6'	0.0004	'R^2'	0.0027
'D5'	0.0003	'Tr'	0.0023
'W4'	-0.0016	'LE'	0.0005

B.2- Feature selection using Sequential methods

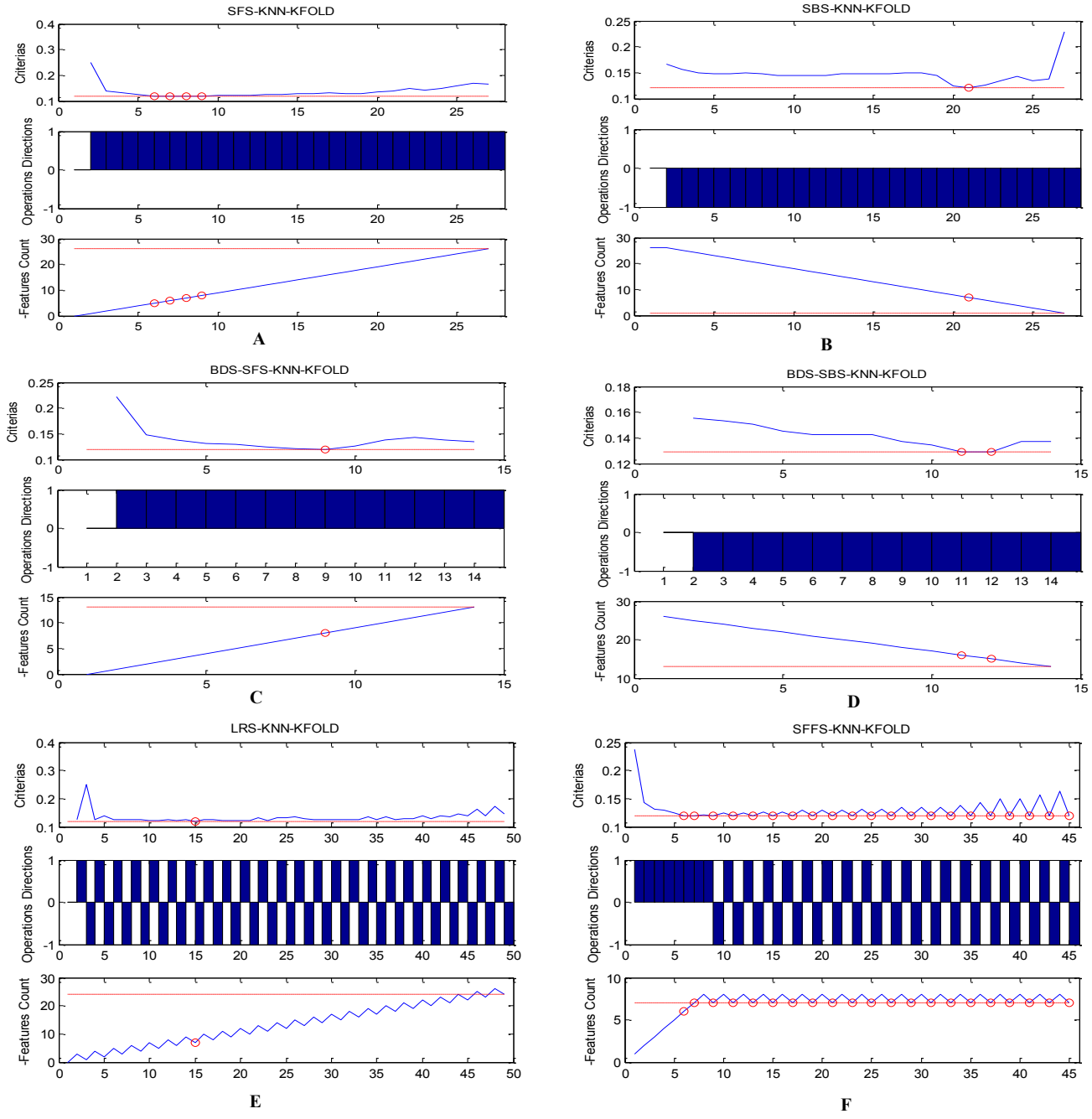


Figure B.2.1: Procedure of sequential methods using KNN-KFOLD with the bipolar dataset. (A) SFS-KNN-KFOLD. (B) SBS-KNN-KFOLD. (C) BDS-SFS-KNN-KFOLD. (D) BDS-SBS-KNN-KFOLD. (E) LRS-KNN-KFOLD. (F) SFFS-KNN-KFOLD.

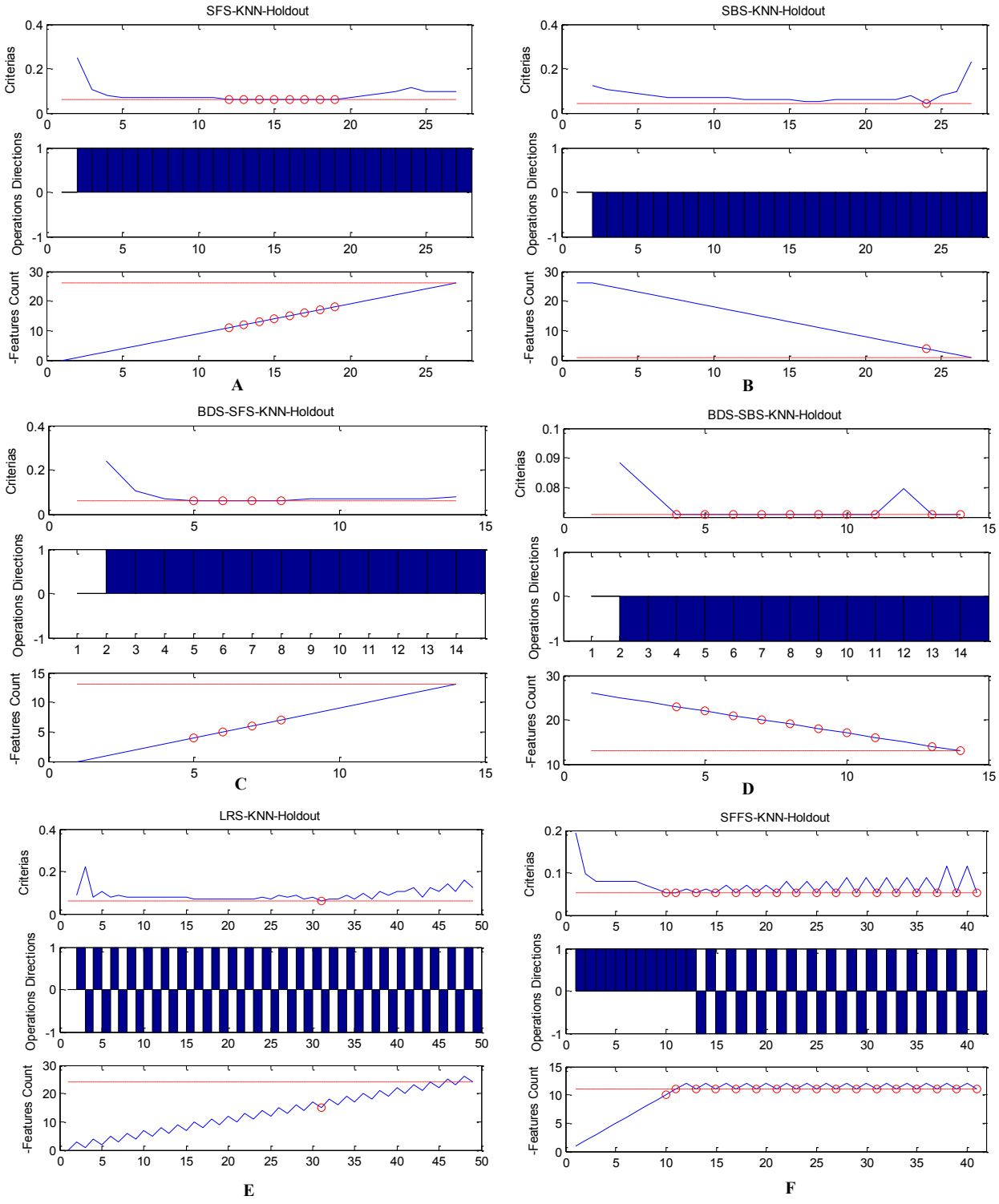


Figure B.2.2: Procedure of sequential methods using KNN-Holdout with the bipolar dataset. (A) SFS-KNN-Holdout. (B) SBS-KNN- Holdout. (C) BDS-SFS-KNN- Holdout. (D) BDS-SBS-KNN- Holdout. (E) LRS-KNN- Holdout. (F) SFFS-KNN- Holdout.

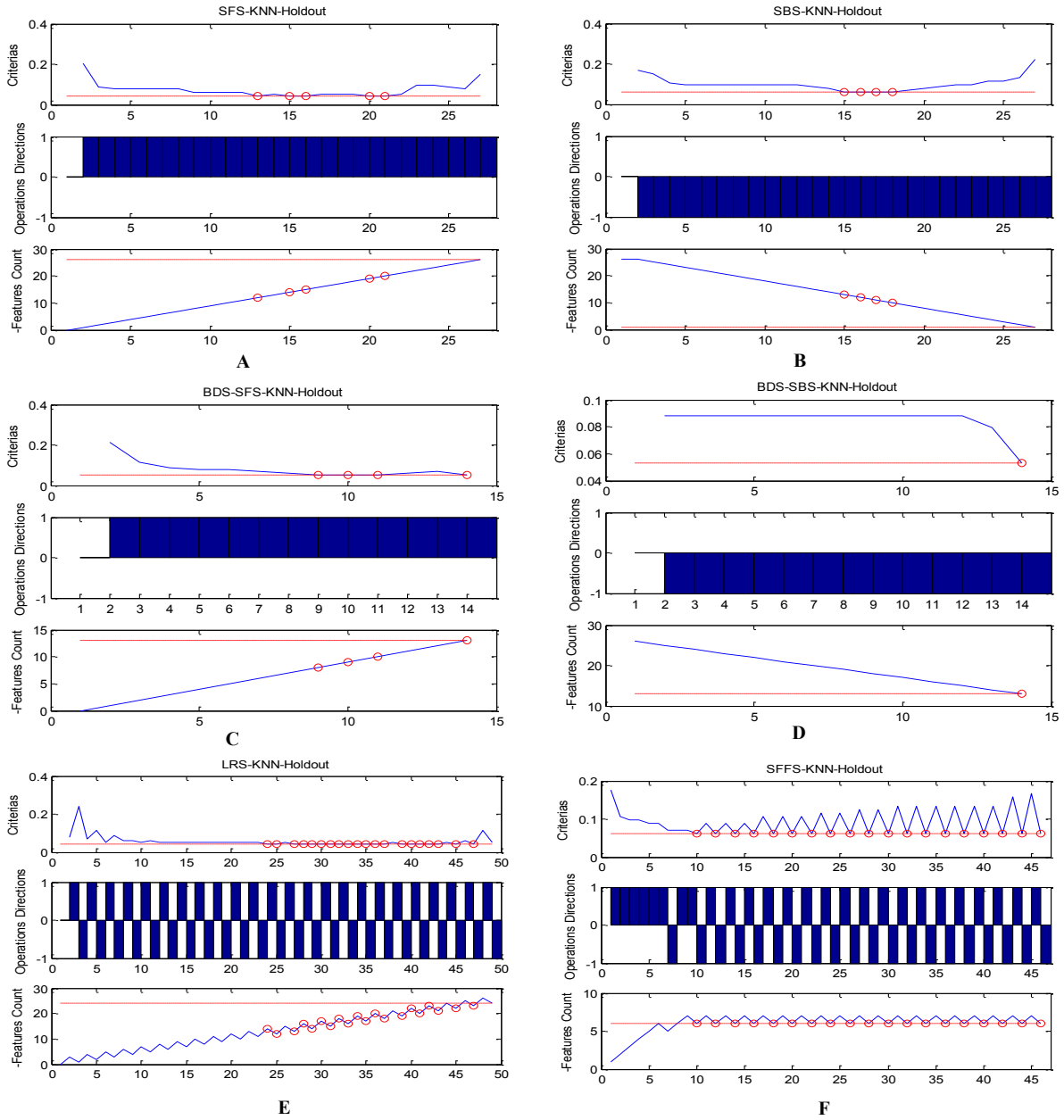


Figure B.2.3: Procedure of sequential methods using KNN-Holdout with the monopolar dataset. (A) SFS-KNN-Holdout. (B) SBS-KNN- Holdout. (C) BDS-SFS-KNN- Holdout. (D) BDS-SBS-KNN- Holdout. (E) LRS-KNN- Holdout. (F) SFFS-KNN- Holdout.

B.3- Process of Feature Selection Using genetic algorithm

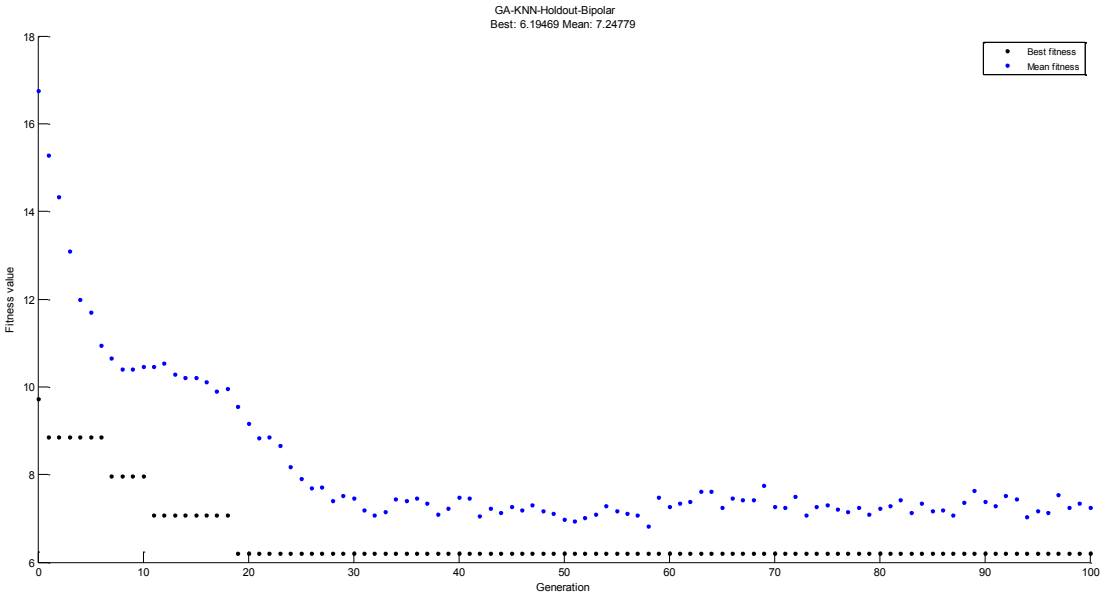


Figure B.3.1: Procedure of GA-KNN-Holdout for the bipolar dataset

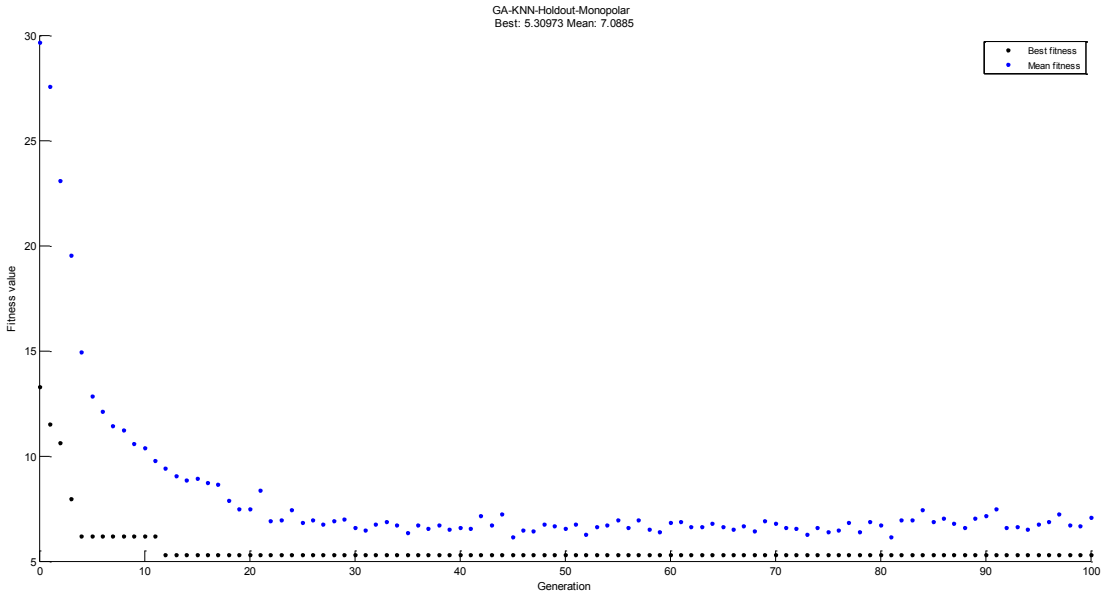


Figure B.3.2: Procedure of GA-KNN-Holdout for the monopolar dataset

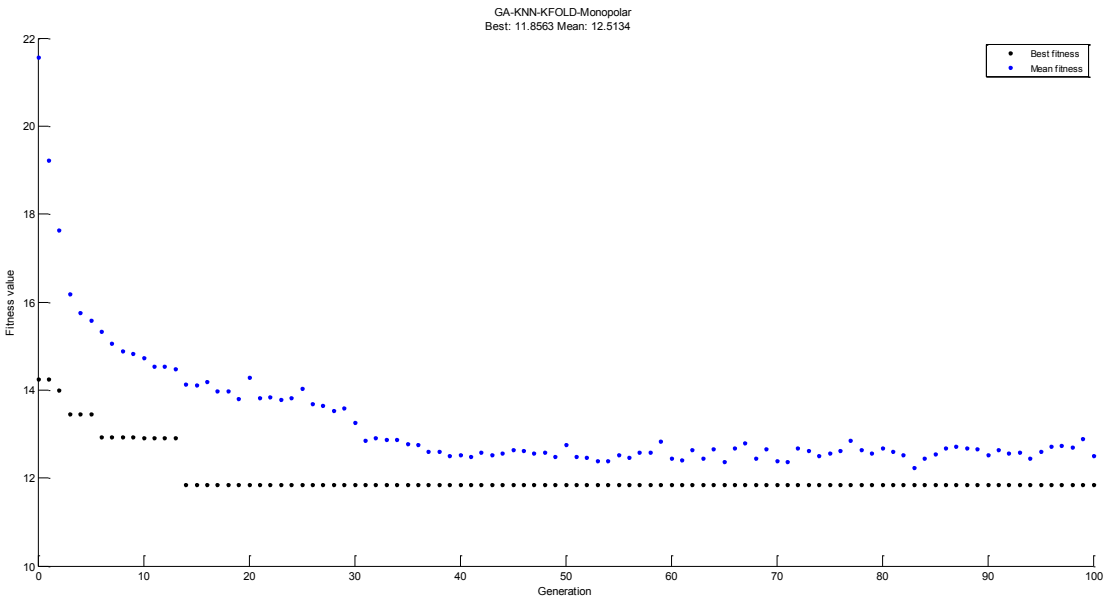


Figure B.3.3: Procedure of GA-KNN-KFOLD for the monopolar dataset

B.4- Validation of the selected feature subsets

Table B.4.1: Mean ± STD of the percentage of correct classification using KNN of 500 repetitions for the selected features subset.

Methods	**Selected feature Subset**		**Mean ± standard deviation of percentage of correct classification (500 repetitions)**	
	Bipolar	**Monopolar**	**Bipolar**	**Monopolar**
Kl-divergence seuil – KFOLD	[DFA, VarEn, FW_H, MPF, W2, R^2, SE, H^2, FW_H^2]	[DFA, VarEn, FW_H, SE, MPF, D9, W5, W2, R^2]	73.76± 2.39	65.39±1.83
F-Score-KFOLD	[DFA, VarEn, W2, MPF, FW_H^2,FW_H]	[SE, DFA, VarEn, W2, FW_H^2, FW_H]	68.39 ± 2.16	61.29 ± 1.97
Relieff – KFOLD	[DFA, VarEn, D8, D9, MPF, PF, y, FW_H]	[DFA, VarEn, W1, D6, D7, D8, D9, MPF, FW_H]	66.90 ± 2.16	62.22± 2
Mutual information with clustering - KFOLD	[SE, DFA, VarEn, MPF, FW_H, FW_H^2]	[SE, DFA, VarEn, W5, MPF, FW_H, FW_H^2]	71.41 ±1.93	63.99 ± 2.27
Kl-divergence seuil -Holdout	[DFA, VarEn, FW_H, MPF, W2, R^2, SE, H^2, FW_H^2]	[DFA, VarEn, FW_H, SE, MPF, D9, W5, W2, R^2]	72.66±7.61	63.66 ± 7.29
F-Score-Holdout	[DFA, VarEn, W2, MPF, FW_H^2,FW_H]	[SE, DFA, VarEn, W2, FW_H^2, FW_H]	68.72±7.48	62.79±7.06
Relieff-Holdout	[DFA, VarEn, D8, D9, MPF, PF, y, FW_H]	[DFA, VarEn, W1, D6, D7, D8, D9, MPF, FW_H]	65.26±7.73	61.93±6.81
Mutual information with clustering - Holdout	[SE, DFA, VarEn, MPF, FW_H, FW_H^2]	[SE, DFA, VarEn, W5, MPF, FW_H, FW_H^2]	70.91 ±7.64	63.91± 7.28
SFS-KNN-KFOLD	[SE, DFA, W4, MPF, H^2]	[Tr, SE, DFA, VarEn, W1, W2, W3, D1, D2, D3, D4, D5, D6, MPF, PF, FW_H]	82.33 ± 1.54	65± 2.20

SBS-KNN-KFOLD	[SE, DFA, W4, D8, MPF, H^2, y]	[DFA, W3, W4, D5, D6, MPF, PF, H^2, y, FW_H^2, FW_H]	84.13±1.76	68.35±1.84
BDS-SFS-KNN-KFOLD	[DFA, W2, W3, D2, D4, D6, MPF, H^2]	[SE, DFA, D1, D5, MPF, PF]	78.44 ± 1.70	66.40±1.87
BDS-SBS-KNN-KFOLD	[Tr, SE, DFA, VarEn, W1, W2, W4, D3, D4, D6, D8, D9, MPF, H^2, y]	[Tr, DFA, VarEn, W1, W2, W4, D1, D3, D5, D6, D7, D8, MPF, PF, R^2, y, FW_H^2, FW_H]	74.99 ± 2.12	67.74±1.65
LRS-KNN-KFOLD	[DFA, W1, W3, D2, D3, MPF, H^2]	[Tr, SE, DFA, VarEn, W1, W2, W3, D3, D4, D5, PF, R^2, FW_H^2]	81.11 ± 1.65	72.21±1.33
SFFS-KNN-KFOLD	[SE, DFA, D2, D3, MPF, R^2]	[Tr, SE, DFA, D2, D4, D5, MPF, PF, FW_H^2]	81.35 ± 1.34	70.15±1.82
GA-KNN-KFOLD	[Tr, SE, DFA, W1, W4, D1, D3, D4, D5, D8, MPF, R^2, H^2, y]	[DFA, W1, W2, W3, W4, D5, D6, MPF, R^2, FW_H]	85.06 ± 1.63	66.69±2
BPSO-KNN-KFOLD	[Tr, SE, DFA, W1, W2, W4, D3, D4, D5, D6, D8, MPF, R^2, H^2, y]	[SE, DFA, VarEn, W1, D1, D2, D3, D4, D5, D6, MPF, PF, H^2, y, FW_H]	85.05 ± 1.69	68.12±1.73
SFS-KNN-Holdout	[Tr, DFA, W1, W2, W3, D1, D4, D8, MPF, R^2, FW_H]	[Tr, DFA, VarEn, W1, W2, W4, D1, D2, D4, D5, D8, R^2]	68.52±8.04	70.23±7.13
SBS-KNN-Holdout	[DFA, W5, MPF, FW_H^2]	[SE, DFA, VarEn, W4, W5, D7, D8, PF, R^2, FW_H^2]	72.29±6.88	70.11±7.72
BDS-SFS-KNN-Holdout	[DFA, D3, D8, H^2]	[SE, DFA, VarEn, W4, D5, D8, MPF, R^2]	74.03±7.05	72.27±7.79
BDS-SBS-KNN-Holdout	[DFA, VarEn, W1, W3, D1, D2, D7, D8, D9, MPF, PF, H^2, y]	[SE, DFA, VarEn, W4, D5, D6, D7, D8, MPF, R^2, H^2, y, FW_H^2]	71.94±7.94	74.88 ±7.10
LRS-KNN-Holdout	[Tr, DFA, W1, W2, D1, D2, D3, D4, D5, D6, D7, D8, PF, R^2, H^2]	[DFA, VarEn, W1, W2, D1, D2, D4, D8, MPF, R^2, y, FW_H]	72.88±7.51	68.45±7.33

SFFS-KNN-Holdout	[Tr, SE, DFA, W1, W2, W3, MPF, R^2, H^2, FW_H^2]	[Tr, DFA, W2, D4, R^2, H^2]	81±6.53	68.57±7.08
GA-KNN-Holdout	[SE, DFA, VarEn, W1, W2, W4, W5, D1, D5, MPF, R^2, H^2, FW_H^2, FW_H]	[DFA, VarEn, D3, D4, D7, D8, PF, R^2, FW_H]	73.81±7.58	66.84±6.95
BPSO-KNN-Holdout	[Tr, DFA, W4, D4, D5, D6, D8, D9]	[Tr, DFA, VarEn, W3, W4, D1, D2, D5, D6, MPF, H^2, y, FW_H]	67.29±8.26	68.19±7.22

B.5- Number of features contained in the selected subsets and appearance of each feature in all selected subsets

Table B.5.1: Number of features contained in each selected subset and time of calculation of each method of feature selection

Methods	Number of Selected features		Time (s)	
	Bipolar	Monopolar	Bipolar	Monopolar
Jeffrey divergence	9	9	0.09	0.08
F-Score	6	6	0.005	0.01
Relieff	8	9	0.35	0.36
Mutual information with clustering	6	7	0.08	0.08
SFS-KNN-KFOLD	5	16	6.41	6.23
SBS-KNN-KFOLD	7	11	6.33	6
BDS-SFS-KNN-KFOLD	8	6	6.64	6.67
BDS-SBS-KNN-KFOLD	15	18	6.68	6.66
LRS-KNN-KFOLD	7	13	30.01	30.08
SFFS-KNN-KFOLD	6	9	14.91	14.99
SFS-KNN-Holdout	11	12	1.20	1.20
SBS-KNN- Holdout	4	10	1.23	1.26
BDS-SFS-KNN- Holdout	4	8	1.31	1.34
BDS-SBS-KNN- Holdout	13	13	1.32	1.34
LRS-KNN- Holdout	15	12	5.79	5.90
SFFS-KNN- Holdout	10	6	3.07	2.67
GA-KNN- KFOLD	14	10	145.56	161.52
BPSO-KNN-KFOLD	15	15	44.19	43.72
GA-KNN- Holdout	14	9	21.24	25.67
BPSO-KNN-Holdout	8	13	6.06	6.04

Table B.5.2: Feature appearance in the 20 selected feature subsets for bipolar and monopolar EHGS.

Feature appearance_ bipolar		**Feature appearance _monopolar**	
Feature	**Number of Appearance**	**Feature**	**Number of Appearance**
'DFA'	20	'DFA'	20
'MPF'	17	'VarEn'	15
'H^{2}'	13	'MPF'	14
'D8'	10	'FW_H'	12
'SE'	10	'D5'	12
'W2'	9	'R^{2}'	11
'W1'	9	'SE'	11
'R^{2}'	8	'PF'	9
'D4'	7	'W2'	9
'D3'	7	'D8'	8
'W4'	7	'D6'	8
'VarEn'	7	'D4'	8
'Tr'	7	'W4'	8
'FW_H'	6	'W1'	8
'y'	6	'FW_H^{2}'	8
'FW_H^{2}'	6	'D1'	7
'D6'	5	'Tr'	7
'D5'	5	'y'	6
'D2'	5	'D2'	6
'D1'	5	'H^{2}'	5
'W3'	5	'D7'	5
'D9'	4	'D3'	5
'PF'	3	'W3'	5
'D7'	2	'W5'	3
'W5'	2	'D9'	2
'LE'	0	'LE'	0

B.6- Mean ± standard deviation (STD) of features values

Table B.6.1: Mean ± standard deviation (STD) of features on monopolar and bipolar EHGs

Feature name	Bipolar EHG		Monopolar EHG	
	Mean ± STD (Pregnancy)	**Mean ± STD (Labor)**	**Mean ± STD (Pregnancy)**	**Mean ± STD (Labor)**
Tr	$-6.10*10^{-6} \pm 0.0001$	0.03 ± 0.37	$-2.23*10^{-6} \pm 1.85*10^{-5}$	-0.02 ± 0.19
LE	5.82 ± 0.73	5.77 ± 0.56	5.76 ± 0.73	5.72 ± 0.62
SE	0.98 ± 0.15	1.06 ± 0.11	0.99 ± 0.14	1.06 ± 0.12
DFA	1 ± 0.19	1.24 ± 0.15	1.03 ± 0.19	1.27 ± 0.15
VarEn	0.08 ± 0.12	0.16 ± 0.11	0.08 ± 0.15	0.15 ± 0.13
W1	0.01 ± 0.01	0.02 ± 0.03	0.01 ± 0.02	0.02 ± 0.02
W2	0.03 ± 0.03	0.05 ± 0.04	0.03 ± 0.04	0.05 ± 0.04
W3	0.09 ± 0.05	0.11 ± 0.07	0.08 ± 0.05	0.11 ± 0.08
W4	0.30 ± 0.11	0.27 ± 0.09	0.27 ± 0.11	0.24 ± 0.09
W5	0.45 ± 0.13	0.42 ± 0.15	0.46 ± 0.11	0.40 ± 0.14
D1	0.14 ± 0.02	0.14 ± 0.02	0.13 ± 0.02	0.13 ± 0.03
D2	0.16 ± 0.04	0.17 ± 0.04	0.15 ± 0.03	0.15 ± 0.05
D3	0.18 ± 0.05	0.19 ± 0.05	0.17 ± 0.04	0.18 ± 0.07
D4	0.21 ± 0.05	0.22 ± 0.08	0.19 ± 0.05	0.21 ± 0.10
D5	0.23 ± 0.06	0.26 ± 0.13	0.22 ± 0.07	0.24 ± 0.13
D6	0.27 ± 0.07	0.32 ± 0.19	0.26 ± 0.11	0.29 ± 0.17
D7	0.33 ± 0.16	0.40 ± 0.25	0.32 ± 0.20	0.37 ± 0.27
D8	0.44 ± 0.39	0.57 ± 0.43	0.48 ± 0.61	0.55 ± 0.55
D9	0.78 ± 1.02	1.07 ± 0.89	0.91 ± 1.45	1.05 ± 1.13
MPF	0.39 ± 0.25	0.54 ± 0.26	0.41 ± 0.36	0.52 ± 0.35
PF	0.18 ± 0.07	0.19 ± 0.10	0.16 ± 0.05	0.17 ± 0.10
R^2	0.16 ± 0.09	0.12 ± 0.06	0.19 ± 0.12	0.14 ± 0.09
H^2	0.20 ± 0.1	0.17 ± 0.08	0.24 ± 0.12	0.19 ± 0.10
y	0.27 ± 0.08	0.26 ± 0.07	0.29 ± 0.09	0.27 ± 0.08
FW_H^2	0.48 ± 0.10	0.54 ± 0.09	0.50 ± 0.11	0.56 ± 0.08
FW_H	3.77 ± 0.57	3.24 ± 0.63	3.78 ± 0.58	3.23 ± 0.63

Appendix C: channel selection followed by feature selection

C.1- Channel selection

Table C.1.1: (A) Bipolar channels selection by F-score. (B) bipolar channels selection by Relieff. The first line in these Tables corresponds to the feature names. Each column corresponds to the subset of bipolar channels selected for each feature (gray color).

Tr	LE	SE	DFA	Var-En	W1	W2	W3	W4	W5	D1	D2	D3	D4	D5	D6	D7	D8	D9	MPF	PF
Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1
Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2
Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3
Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4
Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5
Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6
Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7
Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8
Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9
Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10
Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11
Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12

(A)

Tr	LE	SE	DFA	Var-En	W1	W2	W3	W4	W5	D1	D2	D3	D4	D5	D6	D7	D8	D9	MPF	PF
Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1	Vb1
Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2	Vb2
Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3	Vb3
Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4	Vb4
Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5	Vb5
Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6	Vb6
Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7	Vb7
Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8	Vb8
Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9	Vb9
Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10	Vb10
Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11	Vb11
Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12	Vb12

(B)

Table C.1.2: (A) Monopolar channels selection by F-score. (B) Monopolar channels selection by Relieff. The first line in these tables corresponds to the feature names. Each column corresponds to the subset of monopolar channels selected for each feature (gray color).

Tr	LE	SE	DFA	Var-En	W1	W2	W3	W4	W5	D1	D2	D3	D4	D5	D6	D7	D8	D9	MPF	PF
CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1
CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2
CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3
CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4
CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5
CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6
CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7
CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8
CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9
CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10
CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11
CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12
CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13
CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14
CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15
CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16

(A)

Tr	LE	SE	DFA	Var-En	W1	W2	W3	W4	W5	D1	D2	D3	D4	D5	D6	D7	D8	D9	MPF	PF
CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1	CH1
CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2	CH2
CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3	CH3
CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4	CH4
CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5	CH5
CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6	CH6
CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7	CH7
CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8	CH8
CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9	CH9
CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10	CH10
CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11	CH11
CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12	CH12
CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13	CH13
CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14	CH14
CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15	CH15
CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16	CH16

(B)

C.2- Feature selection after channel selection

Table C.2.1: Subsets of features obtained from BPSO and GA using monopolar and bipolar channels selection

Methods[(1)]	**Dataset**[(2)]	**Selected feature Subset**	**Name of Selected feature Subset**	**Percentage of correct classification**
Bipolar , F-score , GA-KNN-KFOLD	DL1	['SE-Vb7', 'DFA-Vb7', 'DFA-Vb8', 'VarEn-Vb7', 'W1-Vb7', 'W1-Vb8', 'W2-Vb7', 'W3-Vb8', 'W4-Vb7', 'W4-Vb8', 'W5-Vb8', 'D1-Vb8', 'D2-Vb7', 'D2-Vb8', 'D4-Vb8', 'D5-Vb7','D6-Vb8', 'MPF-Vb7']	S1	89.20
Bipolar , F-score , GA-KNN-Holdout	DL1	['SE-Vb7', 'SE-Vb8', 'DFA-Vb7','DFA-Vb8', 'W1-Vb7', 'W3-Vb8', 'W4-Vb7', 'W4-Vb8' , 'W5-Vb7', 'D2-Vb7', 'D3-Vb7', 'D5-Vb8', 'PF-Vb7']	S2	94.69
Bipolar , F-score , BPSO-KNN-KFOLD	DL1	['Tr-Vb7', 'Tr-Vb8', 'DFA-Vb7', 'DFA-Vb8', 'VarEn-Vb7', 'W1-Vb8', 'W3-Vb7', 'W3-Vb8', 'W4-Vb8','D1-Vb7', 'D5-Vb8', 'D7-Vb7', 'MPF-Vb7', 'PF-Vb8']	S3	87.89
Bipolar, F-score, BPSO-KNN-Holdout	DL1	['LE-Vb8', 'SE-Vb7', 'DFA-Vb7', 'DFA-Vb8', 'VarEn-Vb7', 'VarEn-Vb8', 'W1-Vb7', 'W4-Vb7', 'W4-Vb8','W5-Vb7', 'D1-Vb7', 'D1-Vb8', 'D2-Vb7', 'D4-Vb8', 'D6-Vb8', 'D8-Vb7', 'D9-Vb8']	S4	95.58
Bipolar, Relieff, GA-KNN-KFOLD	DL2	['Tr-Vb7', 'LE-Vb7', 'LE-Vb8', 'LE-Vb9', 'SE-Vb7', 'SE-Vb8', 'SE-Vb9', 'DFA-Vb7', 'DFA-Vb8', 'DFA-Vb9', 'VarEn-Vb7', 'VarEn-Vb9', 'W1-Vb7', 'W2-Vb7', 'W2-Vb8', 'W2-Vb9', 'W3-Vb8', 'W3-Vb9', 'W4-Vb7', 'W4-Vb8', 'W4-Vb9', 'W5-Vb8', 'D1-Vb9', 'D2-Vb7', 'D2-Vb8', 'D3-Vb8', 'D5-Vb7', 'D5-Vb8', 'D5-Vb9', 'D6-Vb7', 'D6-Vb9', 'D7-Vb8', 'D7-Vb9', 'MPF-Vb7', 'MPF-Vb8', 'PF-Vb9']	S5	89.45

Bipolar, Relieff, GA-KNN-Holdout	DL2	['Tr-Vb7', 'Tr-Vb9', 'SE-Vb7', 'SE-Vb8', 'SE-Vb9', 'DFA-Vb8', 'VarEn-Vb7', 'VarEn-Vb9', 'W2-Vb7', 'W2-Vb9', 'W3-Vb7', 'W3-Vb8', 'W4-Vb9', 'W5-Vb7', 'W5-Vb8', 'D1-Vb7', 'D2-Vb9', 'D3-Vb9', 'D4-Vb9', 'D5-Vb8', 'D5-Vb9', 'D6-Vb9', 'D8-Vb8','D9-Vb7', 'MPF-Vb7']	S6	94.69
Bipolar, Relieff, BPSO-KNN-KFOLD	DL2	['LE-Vb8', 'LE-Vb9', 'SE-Vb7','DFA-Vb7', 'DFA-Vb8', 'DFA-Vb9', 'VarEn-Vb7', 'VarEn-Vb8', 'VarEn-Vb9', 'W1-Vb7', 'W1-Vb8', 'W1-Vb9', 'W2-Vb7', 'W2-Vb8','W3-Vb8','W3-Vb9', 'W4-Vb7', 'W4-Vb8', 'W5-Vb8', 'D1-Vb7', 'D1-Vb9', 'D2-Vb7', 'D2-Vb8', 'D2-Vb9', 'D3-Vb7', 'D5-Vb9', 'D6-Vb7', 'D6-Vb8', 'D7-Vb7','D7-Vb8', 'D9-Vb7', 'MPF-Vb7']	S7	89.47
Bipolar, Relieff, BPSO-KNN-Holdout	DL2	['Tr-Vb9', 'LE-Vb7', 'LE-Vb8','SE-Vb7', 'SE-Vb8', 'DFA-Vb9', 'VarEn-Vb7', 'VarEn-Vb8', 'W1-Vb8', 'W2-Vb7', 'W3-Vb7', 'W3-Vb8', 'W3-Vb9', 'W4-Vb7', 'W5-Vb8', 'D1-Vb9', 'D3-Vb9', 'D4-Vb8', 'D4-Vb9', 'D5-Vb8',' D6-Vb8', 'D6-Vb9', 'D8-Vb8']	S8	97.35
Monopolar,F-score, GA-KNN-KFOLD	DL3	['SE-CH2', 'SE-CH10', 'DFA-CH2', 'DFA-CH10', 'W3-CH2', 'W4-CH10', 'D1-CH10', 'D2-CH10', 'D3-CH2', 'D3-CH10', 'D4-CH2', 'D4-CH10', 'D5-CH10', 'D6-CH2', 'D6-CH10', 'D7-CH2', 'MPF-CH10', 'PF-CH10']	S9	89.16
Monopolar, F-score, GA-KNN-Holdout	DL3	['Tr-CH10', 'LE-CH2', 'LE-CH10', 'SE-CH10', 'DFA-CH^2', 'DFA-CH10', 'W1-CH2', 'W1-CH10', 'W3-CH2', 'W4-CH2', 'D1-CH10', 'D2-CH^2', 'D2-CH10', 'D3-CH2' , 'D3-CH10', 'D4-CH2', 'D4-CH10', 'D5-CH2', 'D6-CH10', 'D7-CH2', 'D7-CH10', 'D8-CH2', 'D9-CH10', 'MPF-'	S10	94.69

		CH2', 'PF-CH2']		
Monopolar, F-score, BPSO-KNN-KFOLD	DL3	['Tr-CH2', 'SE-CH2', 'SE-CH10', 'DFA-CH2', 'DFA-CH10', 'W1-CH2', 'W3-CH2', 'W3-CH10', 'W4-CH2', 'W4-CH10', 'D1-CH2', 'D1-CH10', 'D4-CH2', 'D4-CH10', 'D5-CH2', 'D6-CH2', 'MPF-CH10']	S11	88.68
Monopolar, F-score, BPSO-KNN-Holdout	DL3	['Tr-CH2', 'LE-CH10', 'SE-CH10', 'DFA-CH2', 'VarEn-CH10', 'W1-CH2', 'W1-CH10', 'W2-CH2', 'W2-CH10', 'W3-CH10', 'W4-CH2', 'W5-CH10', 'D1-CH2', 'D3-CH2','D4-CH10', 'D6-CH2', 'D6-CH10', 'D7-CH10', 'PF-CH2', 'PF-CH10']	S12	96.46
Monopolar, Relieff, GA-KNN-KFOLD	DL4	['Tr-CH9', 'Tr-CH10', 'LE-CH10', 'SE-CH10', 'DFA-CH9', 'DFA-CH10', 'VarEn-CH9', 'W1-CH9', 'W3-CH9', 'D1-CH10', 'D6-CH10', 'D8-CH9', 'D8-CH10']	S13	89.73
Monopolar, Relieff, GA-KNN-Holdout	DL4	['LE-CH10', 'SE-CH9', 'DFA-CH9', 'W3-CH9', 'W5-CH9', 'D1-CH9', 'D1-CH10', 'D3-CH10', 'D4-CH9', 'D6-CH10', 'D7-CH9', 'D8-CH9', 'D9-CH9', 'D9-CH10', 'PF-CH9', 'PF-CH10']	S14	94.69
Monopolar, Relieff, BPSO-KNN-KFOLD	DL4	['Tr-CH9', 'Tr-CH10', 'SE-CH9','SE-CH10', 'DFA-CH10', 'VarEn-CH9', 'W2-CH9', 'W2-CH10', 'W4-CH9', 'W4-CH10', 'D1-CH10', 'D2-CH10', 'D3-CH10', 'D4-CH10', 'D5-CH9', 'D5-CH10', 'D6-CH10', 'D7-CH10', 'MPF-CH9', 'PF-CH10']	S15	89.74
Monopolar, Relieff, BPSO-KNN-Holdout	DL4	['Tr-CH9', 'SE-CH9', 'SE-CH10', 'DFA-CH10', 'VarEn-CH9', 'W1-CH10', 'W2-CH10', 'W3-CH9', 'W3-CH10', 'D1-CH9', 'D1-CH10', 'D2-CH9', 'D3-CH10', 'D4-CH10', 'D6-CH9', 'D6-CH10', 'D7-CH9', 'D8-CH9', 'D9-CH9', 'PF-CH9', 'PF-CH10']	S16	96.46

(1)Method names:

- Recording configuration: Bipolar/Monopolar

- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

(2)Dataset:

- DL1: Dataset contain the 21 features extracted from the contractions (number contractions equal to 379), for the 2 bipolar channels Vb7 and Vb8 (selected using F-score). DL1= 379*42.
- DL2: Dataset contain the 21 features extracted from the contractions (number contractions equal to 379), for the 3 bipolar channels selected Vb7, Vb8 and Vb9 (selected using Relieff). DL2= 379*63.
- DL3: Dataset contain the 21 features extracted from the contractions (number contractions equal to 379), for the 2 monopolar channels CH2 and CH10 (selected using F-score) DL3= 379*42.
- DL4: Dataset contain the 21 features extracted from the contractions (number contractions equal to 379), for the 2 monopolar channels CH9 and CH10 (selected using Relieff). DL4= 379*42.

Table C.2.2: Number of features contained in the selected subsets after feature selection for the selected channels and time of calculation of each method of feature selection

Methods[(1)]	**Selected feature Subset**[(2)]	**Number of features**	**Percentage of correct classification**	**Time (s)**
Bipolar, F-score, GA-KNN-KFOLD	S1	18	89.20	148.48
Bipolar, F-score, GA-KNN-Holdout	S2	13	94.69	20.74
Bipolar, F-score, BPSO-KNN-KFOLD	S3	14	87.89	43.81
Bipolar, F-score, BPSO-KNN-Holdout	S4	17	95.58	6.14
Bipolar, Relieff ,GA-KNN-KFOLD	S5	36	89.45	173.52
Bipolar, Relieff ,GA-KNN-Holdout	S6	25	94.69	25.27
Bipolar, Relieff ,BPSO-KNN-KFOLD	S7	32	89.47	48.5
Bipolar, Relieff ,BPSO-KNN-Holdout	S8	23	97.35	7.01
Monopolar ,F-score , GA-KNN-KFOLD	S9	18	89.16	147.60
Monopolar,F-score, GA-KNN-Holdout	S10	25	94.69	24.67
Monopolar,F-score, BPSO-KNN-KFOLD	S11	17	88.68	43.30
Monopolar,F-score, BPSO-KNN-Holdout	S12	20	96.46	6.13

Monopolar,Relieff, GA-KNN-KFOLD	S13	13	89.73	150.77
Monopolar,Relieff, GA-KNN-Holdout	S14	16	94.69	23.25
Monopolar,Relieff, BPSO-KNN-KFOLD	S15	20	89.74	42.77
Monopolar, Relieff, BPSO-KNN-Holdout	S16	21	96.46	6.11

(1)Method names:

- Recording configuration: Bipolar/Monopolar
- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

(2)Selected feature subset: correspond to the subset obtained in table C.2.1

C.3- Validation of the selected features subset after channel selection

Table C.3.1: Mean ± STD of the percentage of correct classification using KNN of 500 repetitions for the selected features subset from the bipolar and monopolar selected channels

Methods(1)	**Selected feature subset(2)**	**Mean ± standard deviation of percentage of correct classification (500 repetitions)**
Bipolar, F-score , GA-KNN-KFOLD	S1	75.8980 ±1.9440
Bipolar, F-score ,GA-KNN-Holdout	S2	73.9000 ±7.0158
Bipolar, F-score , BPSO-KNN-KFOLD	S3	74.9180 ± 1.5073
Bipolar, F-score , BPSO-KNN-Holdout	S4	69.3133 ±7.7964
Bipolar, Relieff , GA-KNN-KFOLD	S5	78.0740 ± 1.1345
Bipolar, Relieff , GA-KNN-Holdout	S6	69.9733 ±7.9606
Bipolar, Relieff , BPSO-KNN-KFOLD	S7	74.0420± 2.1102
Bipolar, Relieff , BPSO-KNN-Holdout	S8	72.1133 ± 7.2614
Monopolar, F-score , GA-KNN-KFOLD	S9	69.6260 ± 1.3964
Monopolar, F-score , GA-KNN-Holdout	S10	66.6533 ± 7.8138
Monopolar, F-score ,BPSO-KNN-KFOLD	S11	70.2200 ±1.6737
Monopolar, F-score , BPSO-KNN-Holdout	S12	68.7867 ±7.3342
Monopolar, Relieff , GA-KNN-KFOLD	S13	69.3360 ± 1.9968
Monopolar , Relieff , GA-KNN-Holdout	S14	67.0133 ±7.8997

Monopolar, Relieff , BPSO-KNN-KFOLD	S15	70.8720 ± 1.8511
Monopolar, Relieff, BPSO-KNN-Holdout	S16	72.3333 ±6.8820

(1)Method names:

- Recording configuration: Bipolar/Monopolar
- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

(2)Selected feature subset: correspond to the subset obtained in table C.2.1

C.4- Channel combinations selection

Table C.4.1: Bipolar channels combination selection using F-score and Relieff

F-score		Relieff	
Bipolar combinations Name	**Number of appearance**	**Bipolar combinations Name**	**Number of appearance**
'Vb(2,1)'	5	'Vb(11,10)'	5
'Vb(1,2)'	5	'Vb(11,4)'	5
'Vb(12,7)'	4	'Vb(11,3)'	5
'Vb(10,6)'	4	'Vb(10,11)'	5
'Vb(10,3)'	4	'Vb(10,6)'	5
'Vb(9,8)'	4	'Vb(10,4)'	5
'Vb(9,5)'	4	'Vb(10,1)'	5
'Vb(8,9)'	4	'Vb(9,7)'	5
'Vb(7,12)'	4	'Vb(7,9)'	5
'Vb(6,10)'	4	'Vb(7,3)'	5
'Vb(6,2)'	4	'Vb(6,10)'	5
'Vb(5,9)'	4	'Vb(6,1)'	5
'Vb(5,2)'	4	'Vb(4,11)'	5
'Vb(4,5)'	4	'Vb(3,11)'	5
'Vb(3,10)'	4	'Vb(3,7)'	5
'Vb(2,6)'	4	'Vb(1,10)'	5
'Vb(2,5)'	4	'Vb(1,6)'	5
'Vb(12,11)'	3	'Vb(12,10)'	4
'Vb(11,12)'	3	'Vb(11,2)'	4
'Vb(11,10)'	3	'Vb(11,1)'	4
'Vb(11,5)'	3	'Vb(10,12)'	4
'Vb(10,11)'	3	'Vb(10,3)'	4
'Vb(10,7)'	3	'Vb(10,2)'	4
'Vb(10,2)'	3	'Vb(9,10)'	4
'Vb(8,5)'	3	'Vb(9,8)'	4
'Vb(7,10)'	3	'Vb(9,1)'	4
'Vb(7,5)'	3	'Vb(8,9)'	4
'Vb(7,4)'	3	'Vb(7,6)'	4

'Vb(6,3)'	3	'Vb(4,2)'	4
'Vb(6,1)'	3	'Vb(3,12)'	4
'Vb(5,11)'	3	'Vb(3,10)'	4
'Vb(5,8)'	3	'Vb(3,2)'	4
'Vb(5,7)'	3	'Vb(2,11)'	4
'Vb(5,4)'	3	'Vb(2,10)'	4
'Vb(5,3)'	3	'Vb(2,3)'	4
'Vb(4,7)'	3	'Vb(2,1)'	4
'Vb(4,1)'	3	'Vb(1,11)'	4
'Vb(3,5)'	3	'Vb(1,9)'	4
'Vb(2,10)'	3	'Vb(1,2)'	4
'Vb(1,6)'	3	'Vb(12,11)'	3
'Vb(1,4)'	3	'Vb(12,3)'	3
'Vb(12,9)'	2	'Vb(11,12)'	3
'Vb(12,8)'	2	'Vb(10,9)'	3
'Vb(12,5)'	2	'Vb(9,6)'	3
'Vb(12,3)'	2	'Vb(9,2)'	3
'Vb(11,8)'	2	'Vb(7,2)'	3
'Vb(11,4)'	2	'Vb(6,11)'	3
'Vb(11,3)'	2	'Vb(6,9)'	3
'Vb(11,2)'	2	'Vb(6,7)'	3
'Vb(11,1)'	2	'Vb(5,4)'	3
'Vb(10,9)'	2	'Vb(4,10)'	3
'Vb(10,5)'	2	'Vb(4,5)'	3
'Vb(10,4)'	2	'Vb(2,9)'	3
'Vb(10,1)'	2	'Vb(2,7)'	3
'Vb(9,12)'	2	'Vb(2,4)'	3
'Vb(9,10)'	2	'Vb(12,4)'	2
'Vb(9,6)'	2	'Vb(11,6)'	2
'Vb(9,4)'	2	'Vb(10,5)'	2
'Vb(9,1)'	2	'Vb(9,12)'	2
'Vb(8,12)'	2	'Vb(9,11)'	2
'Vb(8,11)'	2	'Vb(8,6)'	2
'Vb(8,6)'	2	'Vb(8,5)'	2
'Vb(8,2)'	2	'Vb(8,4)'	2

'Vb(7,2)'	2	'Vb(6,8)'	2
'Vb(7,1)'	2	'Vb(6,5)'	2
'Vb(6,9)'	2	'Vb(5,10)'	2
'Vb(6,8)'	2	'Vb(5,6)'	2
'Vb(5,12)'	2	'Vb(5,1)'	2
'Vb(5,10)'	2	'Vb(4,3)'	2
'Vb(5,1)'	2	'Vb(3,9)'	2
'Vb(4,11)'	2	'Vb(3,6)'	2
'Vb(4,10)'	2	'Vb(3,4)'	2
'Vb(4,9)'	2	'Vb(3,1)'	2
'Vb(3,12)'	2	'Vb(1,5)'	2
'Vb(3,11)'	2	'Vb(1,3)'	2
'Vb(3,8)'	2	'Vb(12,9)'	1
'Vb(3,6)'	2	'Vb(12,8)'	1
'Vb(3,2)'	2	'Vb(12,6)'	1
'Vb(2,12)'	2	'Vb(12,5)'	1
'Vb(2,11)'	2	'Vb(12,2)'	1
'Vb(2,8)'	2	'Vb(12,1)'	1
'Vb(2,7)'	2	'Vb(11,9)'	1
'Vb(2,3)'	2	'Vb(11,8)'	1
'Vb(1,11)'	2	'Vb(11,7)'	1
'Vb(1,10)'	2	'Vb(11,5)'	1
'Vb(1,9)'	2	'Vb(9,4)'	1
'Vb(1,7)'	2	'Vb(9,3)'	1
'Vb(1,5)'	2	'Vb(8,12)'	1
'Vb(12,10)'	1	'Vb(8,11)'	1
'Vb(12,4)'	1	'Vb(8,7)'	1
'Vb(12,2)'	1	'Vb(8,3)'	1
'Vb(11,6)'	1	'Vb(8,2)'	1
'Vb(10,12)'	1	'Vb(7,11)'	1
'Vb(8,7)'	1	'Vb(7,8)'	1
'Vb(8,3)'	1	'Vb(7,4)'	1
'Vb(8,1)'	1	'Vb(6,12)'	1
'Vb(7,8)'	1	'Vb(6,4)'	1
'Vb(7,6)'	1	'Vb(5,12)'	1

'Vb(7,3)'	1	'Vb(5,11)'	1
'Vb(6,12)'	1	'Vb(5,8)'	1
'Vb(6,11)'	1	'Vb(5,3)'	1
'Vb(6,7)'	1	'Vb(4,12)'	1
'Vb(6,4)'	1	'Vb(4,8)'	1
'Vb(4,12)'	1	'Vb(4,7)'	1
'Vb(4,6)'	1	'Vb(4,6)'	1
'Vb(4,2)'	1	'Vb(3,5)'	1
'Vb(3,7)'	1	'Vb(2,12)'	1
'Vb(3,1)'	1	'Vb(2,8)'	1
'Vb(2,4)'	1	'Vb(1,12)'	1
'Vb(1,8)'	1	'Vb(12,7)'	0
'Vb(1,3)'	1	'Vb(10,8)'	0
'Vb(12,6)'	0	'Vb(10,7)'	0
'Vb(12,1)'	0	'Vb(9,5)'	0
'Vb(11,9)'	0	'Vb(8,10)'	0
'Vb(11,7)'	0	'Vb(8,1)'	0
'Vb(10,8)'	0	'Vb(7,12)'	0
'Vb(9,11)'	0	'Vb(7,10)'	0
'Vb(9,7)'	0	'Vb(7,5)'	0
'Vb(9,3)'	0	'Vb(7,1)'	0
'Vb(9,2)'	0	'Vb(6,3)'	0
'Vb(8,10)'	0	'Vb(6,2)'	0
'Vb(8,4)'	0	'Vb(5,9)'	0
'Vb(7,11)'	0	'Vb(5,7)'	0
'Vb(7,9)'	0	'Vb(5,2)'	0
'Vb(6,5)'	0	'Vb(4,9)'	0
'Vb(5,6)'	0	'Vb(4,1)'	0
'Vb(4,8)'	0	'Vb(3,8)'	0
'Vb(4,3)'	0	'Vb(2,6)'	0
'Vb(3,9)'	0	'Vb(2,5)'	0
'Vb(3,4)'	0	'Vb(1,8)'	0
'Vb(2,9)'	0	'Vb(1,7)'	0
'Vb(1,12)'	0	'Vb(1,4)'	0

Table C.4.2: Monopolar channels combination selection using F-score and Relieff

F-score		Relieff	
Channel combinations Name	**Number of appearance**	**Channel combinations Name**	**Number of appearance**
'Ch(10,2)'	5	'Ch(15,14)'	5
'Ch(8,1)'	5	'Ch(14,12)'	5
'Ch(2,10)'	5	'Ch(14,11)'	5
'Ch(1,8)'	5	'Ch(14,5)'	5
'Ch(16,10)'	4	'Ch(14,2)'	5
'Ch(16,9)'	4	'Ch(14,1)'	5
'Ch(16,6)'	4	'Ch(13,10)'	5
'Ch(15,10)'	4	'Ch(12,14)'	5
'Ch(15,9)'	4	'Ch(12,6)'	5
'Ch(15,5)'	4	'Ch(11,14)'	5
'Ch(14,12)'	4	'Ch(11,10)'	5
'Ch(12,14)'	4	'Ch(10,13)'	5
'Ch(11,8)'	4	'Ch(10,11)'	5
'Ch(11,1)'	4	'Ch(9,13)'	5
'Ch(10,15)'	4	'Ch(9,1)'	5
'Ch(10,9)'	4	'Ch(8,7)'	5
'Ch(10,1)'	4	'Ch(7,8)'	5
'Ch(9,16)'	4	'Ch(6,12)'	5
'Ch(9,15)'	4	'Ch(5,14)'	5
'Ch(9,10)'	4	'Ch(4,7)'	5
'Ch(9,5)'	4	'Ch(3,16)'	5
'Ch(9,4)'	4	'Ch(2,7)'	5
'Ch(9,2)'	4	'Ch(1,14)'	5
'Ch(8,11)'	4	'Ch(1,9)'	5
'Ch(8,5)'	4	'Ch(16,14)'	4
'Ch(7,5)'	4	'Ch(16,13)'	4
'Ch(6,16)'	4	'Ch(16,9)'	4
'Ch(5,15)'	4	'Ch(16,7)'	4
'Ch(5,9)'	4	'Ch(16,5)'	4
'Ch(5,8)'	4	'Ch(16,3)'	4

'Ch(5,7)'	4	'Ch(15,7)'	4
'Ch(5,3)'	4	'Ch(15,2)'	4
'Ch(4,9)'	4	'Ch(15,1)'	4
'Ch(4,1)'	4	'Ch(14,16)'	4
'Ch(3,5)'	4	'Ch(14,15)'	4
'Ch(1,4)'	4	'Ch(14,13)'	4
'Ch(16,15)'	3	'Ch(14,9)'	4
'Ch(16,12)'	3	'Ch(14,8)'	4
'Ch(15,16)'	3	'Ch(13,14)'	4
'Ch(15,14)'	3	'Ch(13,12)'	4
'Ch(14,15)'	3	'Ch(13,9)'	4
'Ch(14,13)'	3	'Ch(13,5)'	4
'Ch(13,14)'	3	'Ch(13,4)'	4
'Ch(12,16)'	3	'Ch(13,1)'	4
'Ch(12,9)'	3	'Ch(12,13)'	4
'Ch(12,4)'	3	'Ch(9,14)'	4
'Ch(11,2)'	3	'Ch(9,8)'	4
'Ch(10,16)'	3	'Ch(9,6)'	4
'Ch(10,6)'	3	'Ch(8,14)'	4
'Ch(9,12)'	3	'Ch(8,9)'	4
'Ch(9,11)'	3	'Ch(7,16)'	4
'Ch(9,8)'	3	'Ch(7,4)'	4
'Ch(9,7)'	3	'Ch(7,2)'	4
'Ch(9,3)'	3	'Ch(6,9)'	4
'Ch(8,9)'	3	'Ch(6,7)'	4
'Ch(8,4)'	3	'Ch(6,3)'	4
'Ch(7,9)'	3	'Ch(5,16)'	4
'Ch(6,10)'	3	'Ch(5,13)'	4
'Ch(4,12)'	3	'Ch(5,8)'	4
'Ch(4,8)'	3	'Ch(5,4)'	4
'Ch(3,9)'	3	'Ch(4,13)'	4
'Ch(3,1)'	3	'Ch(3,2)'	4
'Ch(2,11)'	3	'Ch(2,15)'	4
'Ch(2,1)'	3	'Ch(2,14)'	4
'Ch(1,15)'	3	'Ch(2,13)'	4

'Ch(1,11)'	3	'Ch(2,12)'	4
'Ch(1,10)'	3	'Ch(2,3)'	4
'Ch(1,3)'	3	'Ch(1,13)'	4
'Ch(1,2)'	3	'Ch(16,12)'	3
'Ch(16,14)'	2	'Ch(16,2)'	3
'Ch(16,13)'	2	'Ch(16,1)'	3
'Ch(16,7)'	2	'Ch(15,13)'	3
'Ch(16,5)'	2	'Ch(13,16)'	3
'Ch(16,2)'	2	'Ch(13,15)'	3
'Ch(16,1)'	2	'Ch(13,3)'	3
'Ch(15,12)'	2	'Ch(12,16)'	3
'Ch(15,11)'	2	'Ch(12,8)'	3
'Ch(15,8)'	2	'Ch(12,7)'	3
'Ch(15,7)'	2	'Ch(12,3)'	3
'Ch(15,3)'	2	'Ch(12,2)'	3
'Ch(15,1)'	2	'Ch(11,12)'	3
'Ch(14,16)'	2	'Ch(11,4)'	3
'Ch(14,9)'	2	'Ch(11,2)'	3
'Ch(14,4)'	2	'Ch(11,1)'	3
'Ch(14,1)'	2	'Ch(10,7)'	3
'Ch(13,16)'	2	'Ch(10,6)'	3
'Ch(13,12)'	2	'Ch(10,2)'	3
'Ch(13,11)'	2	'Ch(9,16)'	3
'Ch(13,10)'	2	'Ch(9,5)'	3
'Ch(13,8)'	2	'Ch(9,2)'	3
'Ch(13,7)'	2	'Ch(8,12)'	3
'Ch(13,4)'	2	'Ch(8,6)'	3
'Ch(13,3)'	2	'Ch(8,5)'	3
'Ch(12,15)'	2	'Ch(7,15)'	3
'Ch(12,13)'	2	'Ch(7,12)'	3
'Ch(12,11)'	2	'Ch(7,10)'	3
'Ch(12,10)'	2	'Ch(6,10)'	3
'Ch(12,8)'	2	'Ch(6,8)'	3
'Ch(12,7)'	2	'Ch(6,5)'	3
'Ch(12,2)'	2	'Ch(5,9)'	3

'Ch(12,1)'	2	'Ch(5,6)'	3
'Ch(11,15)'	2	'Ch(5,2)'	3
'Ch(11,13)'	2	'Ch(4,14)'	3
'Ch(11,12)'	2	'Ch(4,11)'	3
'Ch(11,10)'	2	'Ch(4,5)'	3
'Ch(11,9)'	2	'Ch(4,2)'	3
'Ch(11,7)'	2	'Ch(4,1)'	3
'Ch(11,4)'	2	'Ch(3,14)'	3
'Ch(11,3)'	2	'Ch(3,13)'	3
'Ch(10,13)'	2	'Ch(3,12)'	3
'Ch(10,12)'	2	'Ch(3,9)'	3
'Ch(10,11)'	2	'Ch(3,6)'	3
'Ch(10,8)'	2	'Ch(2,16)'	3
'Ch(9,14)'	2	'Ch(2,11)'	3
'Ch(9,6)'	2	'Ch(2,10)'	3
'Ch(9,1)'	2	'Ch(2,9)'	3
'Ch(8,15)'	2	'Ch(2,5)'	3
'Ch(8,13)'	2	'Ch(2,4)'	3
'Ch(8,12)'	2	'Ch(1,15)'	3
'Ch(8,10)'	2	'Ch(1,11)'	3
'Ch(8,7)'	2	'Ch(1,4)'	3
'Ch(8,6)'	2	'Ch(16,11)'	2
'Ch(7,15)'	2	'Ch(16,10)'	2
'Ch(7,13)'	2	'Ch(16,4)'	2
'Ch(7,12)'	2	'Ch(15,11)'	2
'Ch(7,11)'	2	'Ch(15,8)'	2
'Ch(7,8)'	2	'Ch(15,6)'	2
'Ch(7,6)'	2	'Ch(15,4)'	2
'Ch(7,4)'	2	'Ch(14,10)'	2
'Ch(7,3)'	2	'Ch(14,4)'	2
'Ch(6,9)'	2	'Ch(13,11)'	2
'Ch(6,8)'	2	'Ch(13,6)'	2
'Ch(6,7)'	2	'Ch(13,2)'	2
'Ch(6,4)'	2	'Ch(12,11)'	2
'Ch(6,3)'	2	'Ch(12,10)'	2

'Ch(6,2)'	2	'Ch(12,9)'	2
'Ch(5,16)'	2	'Ch(12,5)'	2
'Ch(5,11)'	2	'Ch(11,16)'	2
'Ch(5,10)'	2	'Ch(11,13)'	2
'Ch(5,4)'	2	'Ch(11,9)'	2
'Ch(5,2)'	2	'Ch(10,14)'	2
'Ch(4,14)'	2	'Ch(10,12)'	2
'Ch(4,13)'	2	'Ch(10,3)'	2
'Ch(4,11)'	2	'Ch(9,12)'	2
'Ch(4,7)'	2	'Ch(9,3)'	2
'Ch(4,6)'	2	'Ch(8,15)'	2
'Ch(4,5)'	2	'Ch(8,4)'	2
'Ch(4,2)'	2	'Ch(8,3)'	2
'Ch(3,15)'	2	'Ch(8,2)'	2
'Ch(3,13)'	2	'Ch(7,6)'	2
'Ch(3,11)'	2	'Ch(7,3)'	2
'Ch(3,7)'	2	'Ch(5,12)'	2
'Ch(3,6)'	2	'Ch(5,1)'	2
'Ch(2,16)'	2	'Ch(4,15)'	2
'Ch(2,12)'	2	'Ch(4,10)'	2
'Ch(2,9)'	2	'Ch(4,8)'	2
'Ch(2,5)'	2	'Ch(4,6)'	2
'Ch(2,4)'	2	'Ch(3,10)'	2
'Ch(1,16)'	2	'Ch(3,8)'	2
'Ch(1,14)'	2	'Ch(3,7)'	2
'Ch(1,12)'	2	'Ch(2,8)'	2
'Ch(1,9)'	2	'Ch(2,1)'	2
'Ch(16,11)'	1	'Ch(1,16)'	2
'Ch(16,8)'	1	'Ch(1,5)'	2
'Ch(15,13)'	1	'Ch(1,2)'	2
'Ch(15,2)'	1	'Ch(16,15)'	1
'Ch(14,8)'	1	'Ch(16,6)'	1
'Ch(14,7)'	1	'Ch(15,16)'	1
'Ch(14,3)'	1	'Ch(15,9)'	1
'Ch(14,2)'	1	'Ch(15,5)'	1

'Ch(13,2)'	1	'Ch(15,3)'	1
'Ch(13,1)'	1	'Ch(14,3)'	1
'Ch(11,16)'	1	'Ch(13,7)'	1
'Ch(11,5)'	1	'Ch(12,15)'	1
'Ch(10,7)'	1	'Ch(12,1)'	1
'Ch(10,5)'	1	'Ch(11,15)'	1
'Ch(8,16)'	1	'Ch(11,8)'	1
'Ch(8,14)'	1	'Ch(11,7)'	1
'Ch(7,16)'	1	'Ch(10,16)'	1
'Ch(7,14)'	1	'Ch(10,5)'	1
'Ch(7,10)'	1	'Ch(10,4)'	1
'Ch(7,2)'	1	'Ch(10,1)'	1
'Ch(7,1)'	1	'Ch(9,11)'	1
'Ch(6,1)'	1	'Ch(8,11)'	1
'Ch(5,14)'	1	'Ch(7,13)'	1
'Ch(5,1)'	1	'Ch(7,11)'	1
'Ch(3,14)'	1	'Ch(7,5)'	1
'Ch(3,8)'	1	'Ch(6,16)'	1
'Ch(3,4)'	1	'Ch(6,15)'	1
'Ch(3,2)'	1	'Ch(6,14)'	1
'Ch(2,15)'	1	'Ch(6,13)'	1
'Ch(2,14)'	1	'Ch(6,4)'	1
'Ch(2,13)'	1	'Ch(6,1)'	1
'Ch(2,7)'	1	'Ch(5,15)'	1
'Ch(2,6)'	1	'Ch(5,10)'	1
'Ch(2,3)'	1	'Ch(5,7)'	1
'Ch(1,13)'	1	'Ch(5,3)'	1
'Ch(1,7)'	1	'Ch(4,16)'	1
'Ch(1,6)'	1	'Ch(4,3)'	1
'Ch(1,5)'	1	'Ch(3,15)'	1
'Ch(16,4)'	0	'Ch(3,11)'	1
'Ch(16,3)'	0	'Ch(3,5)'	1
'Ch(15,6)'	0	'Ch(3,4)'	1
'Ch(15,4)'	0	'Ch(3,1)'	1
'Ch(14,11)'	0	'Ch(2,6)'	1

'Ch(14,10)'	0	'Ch(1,12)'	1
'Ch(14,6)'	0	'Ch(1,10)'	1
'Ch(14,5)'	0	'Ch(1,6)'	1
'Ch(13,15)'	0	'Ch(1,3)'	1
'Ch(13,9)'	0	'Ch(16,8)'	0
'Ch(13,6)'	0	'Ch(15,12)'	0
'Ch(13,5)'	0	'Ch(15,10)'	0
'Ch(12,6)'	0	'Ch(14,7)'	0
'Ch(12,5)'	0	'Ch(14,6)'	0
'Ch(12,3)'	0	'Ch(13,8)'	0
'Ch(11,14)'	0	'Ch(12,4)'	0
'Ch(11,6)'	0	'Ch(11,6)'	0
'Ch(10,14)'	0	'Ch(11,5)'	0
'Ch(10,4)'	0	'Ch(11,3)'	0
'Ch(10,3)'	0	'Ch(10,15)'	0
'Ch(9,13)'	0	'Ch(10,9)'	0
'Ch(8,3)'	0	'Ch(10,8)'	0
'Ch(8,2)'	0	'Ch(9,15)'	0
'Ch(6,15)'	0	'Ch(9,10)'	0
'Ch(6,14)'	0	'Ch(9,7)'	0
'Ch(6,13)'	0	'Ch(9,4)'	0
'Ch(6,12)'	0	'Ch(8,16)'	0
'Ch(6,11)'	0	'Ch(8,13)'	0
'Ch(6,5)'	0	'Ch(8,10)'	0
'Ch(5,13)'	0	'Ch(8,1)'	0
'Ch(5,12)'	0	'Ch(7,14)'	0
'Ch(5,6)'	0	'Ch(7,9)'	0
'Ch(4,16)'	0	'Ch(7,1)'	0
'Ch(4,15)'	0	'Ch(6,11)'	0
'Ch(4,10)'	0	'Ch(6,2)'	0
'Ch(4,3)'	0	'Ch(5,11)'	0
'Ch(3,16)'	0	'Ch(4,12)'	0
'Ch(3,12)'	0	'Ch(4,9)'	0
'Ch(3,10)'	0	'Ch(1,8)'	0
'Ch(2,8)'	0	'Ch(1,7)'	0

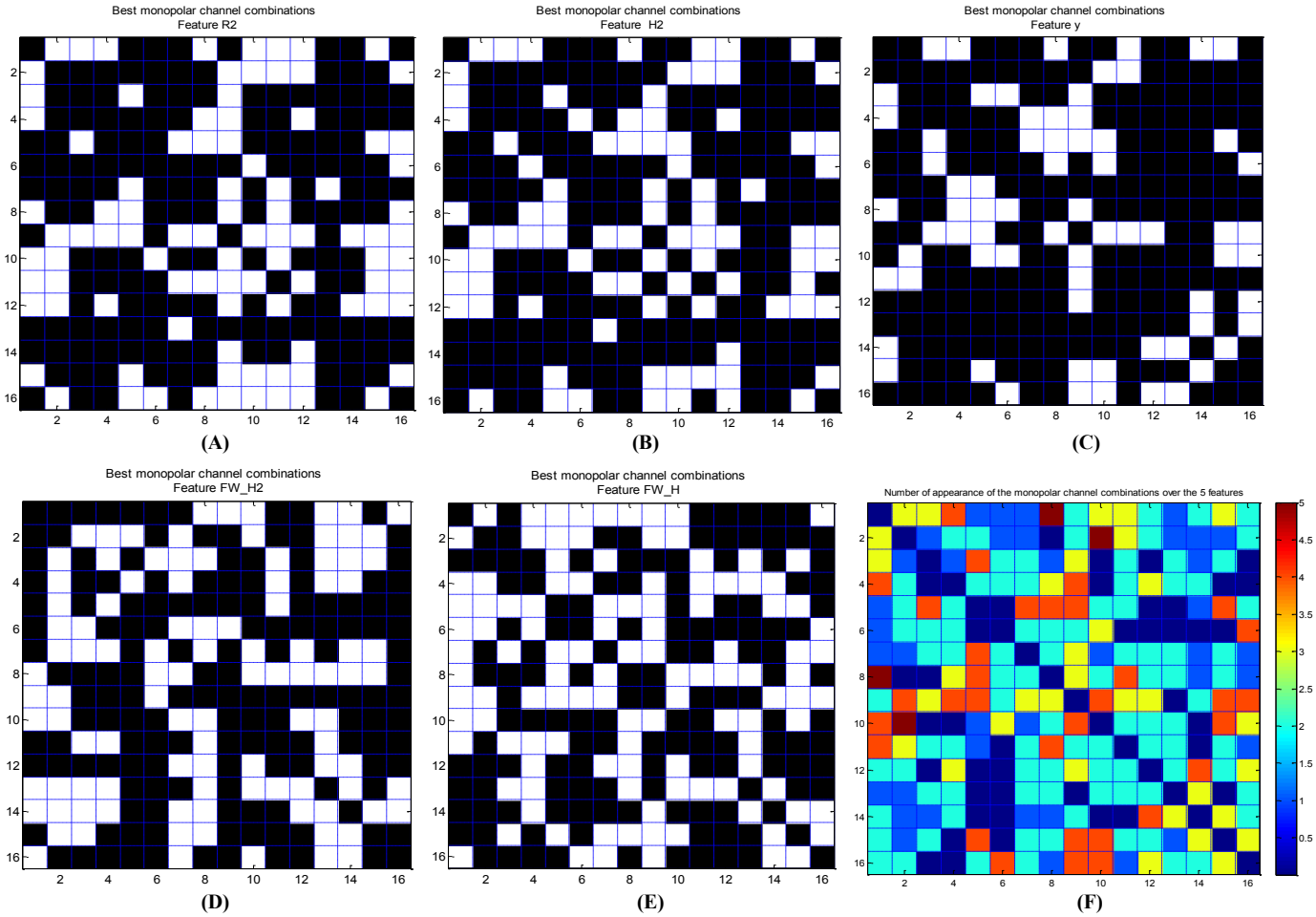


Figure C.4.1: Monopolar channel combinations selection by F-score using five features related to EHG bivariate analysis: R^2(A), H^2(B), y (C), FW_H^2(D) and FW_H (E). (F) Number of appearance of monopolar channel combinations for all the features.

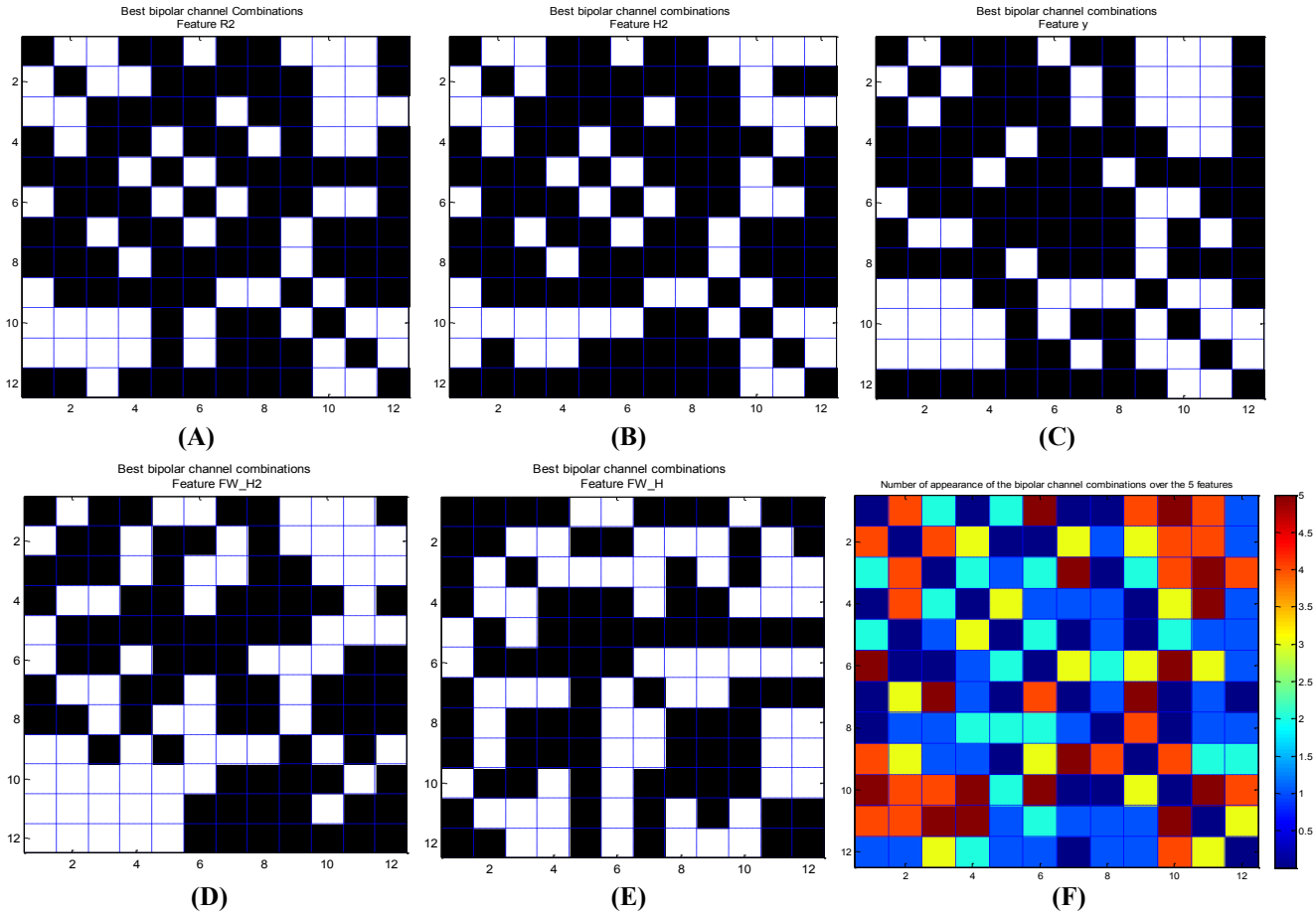


Figure C.4.2: Bipolar channel combinations selection by relieff using five features related to EHG bivariate analysis: R^2 (A), H^2(B), y (C), FW_H^2(D) and FW_H (E). (F) Number of appearance of bipolar channel combinations for all the features.

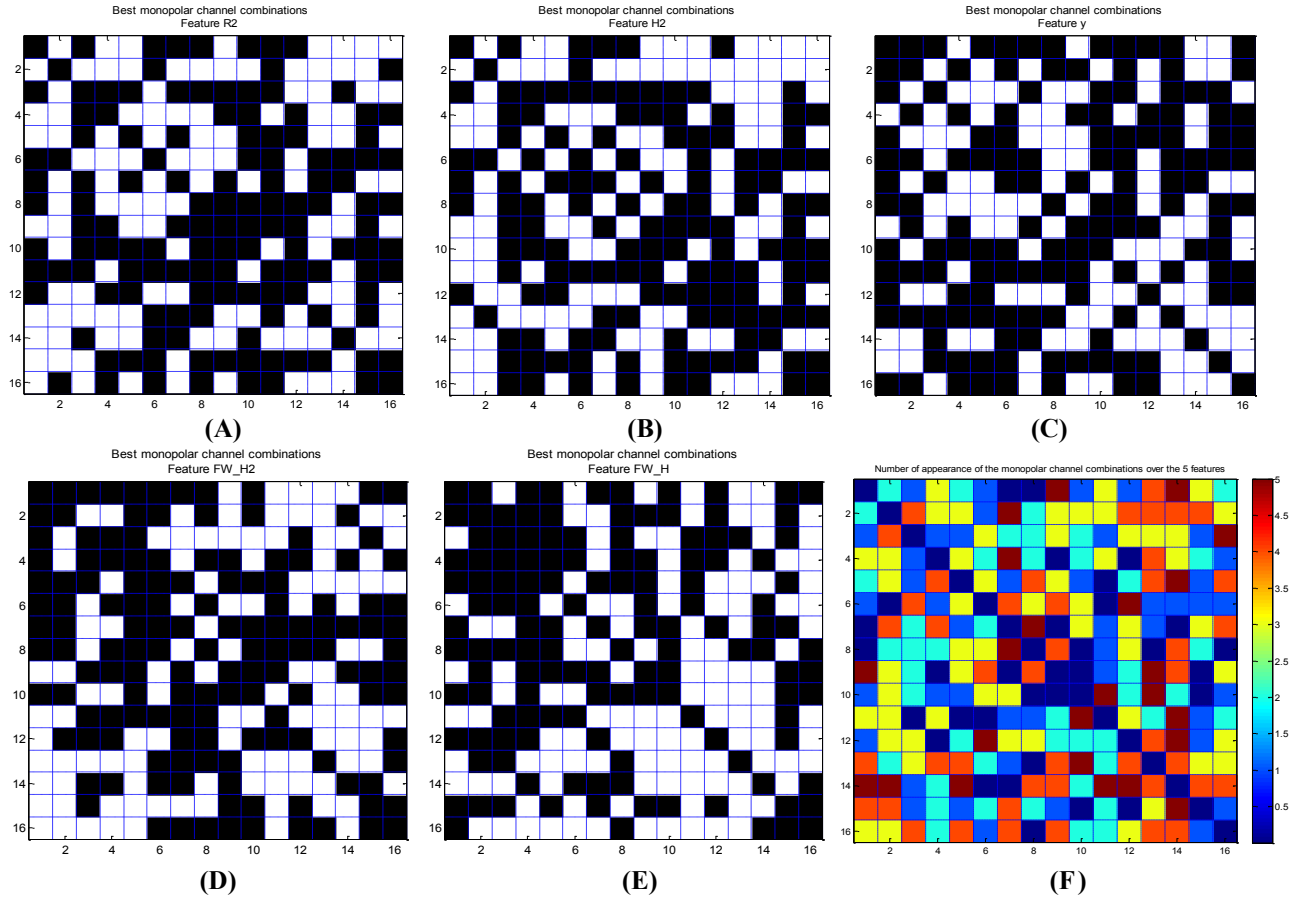


Figure C.4.3: Monopolar channel combinations selection by relieff using five features related to EHG bivariate analysis: R^2 (A), H^2(B), y (C), FW_H^2(D) and FW_H (E). (F) Number of appearance of monopolar channel combinations for all the features.

C.5- Feature selection after Channel combinations selection

Table C.5.1: Subsets of features obtained from BPSO and GA using monopolar and bipolar channel combinations selection

Methods[(1)]	**Dataset** [(2)]	**Selected feature subset**	**Name of Selected feature Subset**	**Correct classification percentage**
Bipolar, F-score, GA-KNN-KFOLD	DC1	['R^2-Vb(10,3)','R^2-Vb(9,5)','R^2-Vb(7,12)', 'R^2-Vb(6,2)','R^2-Vb(5,9)','R^2-Vb(5,2)','R^2-Vb(4,5)', 'H^2-Vb(2,1)','H^2-Vb(10,6)','H^2-Vb(10,3) , 'H^2-Vb(9,8)','H^2-Vb(6,2)','H^2-Vb(5,9)', 'H^2-Vb(2,6)' ,'H^2-Vb(2,5)','y-Vb(10,3)','y-Vb(9,5)','y-Vb(5,2)' ,'y-Vb(3,10)','FW_H^2-Vb(10,6)','FW_H^2-Vb(9,8)', 'FW_H^2-Vb(6,2)','FW_H-Vb(6,2)', 'FW_H-Vb(5,2)', 'FW_H-Vb(2,6)']	SC1	88.39
Bipolar, F-score, GA-KNN-Holdout	DC1	['R^2-Vb(2,1)','R^2-Vb(1,2)','R^2-Vb(12,7)' ,'R^2-Vb(10,6)','R^2-Vb(10,3)','R^2-Vb(9,5)','R^2-Vb(8,9)', 'R^2-Vb(7,12)','R^2-Vb(6,10)','R^2-Vb(5,2)' ,'R^2-Vb(3,10)','R^2-Vb(2,5)' ,'H^2-Vb(2,1) ','H^2-Vb(12,7)', 'H^2-Vb(10,6)','H^2-Vb(9,8)','H^2-Vb(8,9)' ,'H^2-Vb(7,12)','H^2-Vb(5,9)' ,'H^2-Vb(3,10)','H^2-Vb(2,6)' ,'H^2-Vb(2,5)','y-Vb(12,7)', 'y-Vb(10,6)','y-Vb(9,8) ','y-Vb(9,5)','y-Vb(8,9)','y-Vb(7,12)','y-Vb(6,2)','y-Vb(5,2)','y-Vb(4,5)', 'FW_H^2-Vb(9,8)','FW_H^2-Vb(7,12)', 'FW_H^2-Vb(6,10)','FW_H^2-Vb(6,2)' ,'FW_H^2-Vb(5,9)', 'FW_H^2 Vb(2,5)', 'FW_H-Vb(2,1)','FW_H-Vb(1,2)','FW_H-Vb(9,5)','FW_H-Vb(8,9)']	SC2	94.69
Bipolar, F-score, BPSO-KNN-KFOLD	DC1	['R^2-Vb(2,1)','R^2-Vb(10,6)','R^2-Vb(10,3)','R^2-Vb(9,8)','R^2-Vb(9,5)','R^2-Vb(6,2)' ,'R^2-Vb(5,2)' ,'R^2-Vb(4,5)','R^2-Vb(2,6)','R^2-Vb(2,5)','H^2-Vb(10,3)','H^2-Vb(9,5)','H^2-Vb(8,9) ','H^2-Vb(7,12)' ,'H^2-Vb(6,10)','H^2-Vb(5,2)','H^2-Vb(4,5)','H^2-Vb(2,6)','y-Vb(2,1)','y-Vb(12,7)','y-Vb(10,3)','y-Vb(6,2)','y-Vb(5,9)','y-Vb(5,2)','y-Vb(4,5)','y-	SC3	86.84

		Vb(3,10)','yVb(2,5)','FW_H^2-Vb(12,7)' ,'FW_H^2-Vb(10,6)', 'FW_H^2-Vb(9,5)', 'FW_H^2-Vb(6,10)' ,'FW_H^2-Vb(6,2)','FW_H^2-Vb(2,5)', 'FW_H-Vb(9,5)','FW_H-Vb(8,9)']		
Bipolar, F-score, BPSO-KNN-Holdout	DC1	['R^2-Vb(2,1)','R^2-Vb(10,6)', 'R^2-Vb(9,5)', 'R^2-Vb(7,12)', 'R^2-Vb(6,10)','R^2-Vb(6,2)', 'R^2-Vb(5,9)','R^2-Vb(5,2)', 'R^2-Vb(4,5)','R^2-Vb(3,10)', 'R^2-Vb(2,5)','H^2-Vb(12,7)', 'H^2-Vb(10,6)','H^2-Vb(10,3)', 'H^2-Vb(9,8)','H^2-Vb(9,5)',' H^2-Vb(7,12)','H^2-Vb(6,10)', 'H^2-Vb(5,9)','H^2-Vb(5,2)', 'H^2-Vb(4,5)','y-Vb(1,2)','y-Vb(12,7)','y-Vb(10,6)','y-Vb(10,3)','y-Vb(9,8)','y-Vb(7,12)','y-Vb(6,10)','y-Vb(3,10)','FW_H^2-Vb(2,1)', 'FW_H^2-Vb(1,2)','FW_H^2-Vb(10,6)','FW_H^2-Vb(9,8)', 'FW_H^2-Vb(7,12)','FW_H^2-Vb(6,10)','FW_H^2-Vb(6,2)', 'FW_H^2-Vb(5,9)','FW_H^2-Vb(2,6)', 'FW_H^2-Vb(2,5)', 'FW_H-Vb(9,5)','FW_H-Vb(8,9)','FW_H-Vb(6,10)', 'FW_H-Vb(6,2)', 'FW_H-Vb(5,9)','FW_H-Vb(3,10)']	SC4	95.58
Bipolar, Relieff, GA-KNN-KFOLD	DC2	['R^2-Vb(11,3)','R^2-Vb(10,4)','R^2-Vb(6,10)', 'H^2-Vb(11,4)','H^2-Vb(6,1)','y-Vb(11,10)','y-Vb(11,3)', 'y-Vb(10,11)','y-Vb(10,6)','y-Vb(6,1)', 'FW_H^2-Vb(11,10)', 'FW_H^2-Vb(11,4)','FW_H^2-Vb(10,6)', 'FW_H^2-Vb(10,1)', 'FW_H^2-Vb(4,11)', 'FW_H-Vb(10,11)','FW_H-Vb(10,4)', 'FW_H-Vb(10,1)']	SC5	82.86
Bipolar, Relieff, GA-KNN-Holdout	DC2	['R^2-Vb(11,4)','R^2-Vb(10,6)', 'R^2-Vb(10,1)','R^2-Vb(7,3)','R^2-Vb(6,10)','R^2-Vb(6,1)', 'R^2-Vb(4,11)','R^2-Vb(3,11)','R^2-Vb(3,7)','H^2-Vb(11,4)','H^2-Vb(10,1)','H^2-Vb(9,7)','H^2-Vb(6,1)' ,'H^2-Vb(4,11)','H^2-Vb(3,7)','H^2-Vb(1,6)' ,'y-Vb(11,4)','y-Vb(10,6)','y-Vb(9,7)','y-Vb(6,1)','y-Vb(4,11)', 'FW_H^2-Vb(10,11)', 'FW_H^2-Vb(10,6)', 'FW_H^2-Vb(10,4)', 'FW_H^2-Vb(10,1)', 'FW_H^2-Vb(9,7)', 'FW_H^2-Vb(7,9)','FW_H^2-Vb(6,10)', 'FW_H^2-Vb(6,1)','FW_H^2-Vb(3,11)','FW_H^2-Vb(3,7)', 'FW_H-Vb(10,4)','FW_H-Vb(10,1)', 'FW_H-Vb(7,3)', 'FW_H-	SC6	92.04

		Vb(3,7)','FW_H-Vb(1,10)' , 'FW_H-Vb(1,6)']		
Bipolar, Relieff, BPSO-KNN-KFOLD	DC2	['R^2-Vb(11,4)','R^2-Vb(11,3)','R^2-Vb(10,4)', 'R^2-Vb(10,1)','R^2-Vb(9,7)','R^2-Vb(7,9)' ,'R^2-Vb(6,1)','R^2-Vb(4,11)','R^2-Vb(3,11)' ,'R^2-Vb(1,10)' , 'R^2-Vb(1,6)','H^2-Vb(11,10)','H^2-Vb(11,4)', 'H^2-Vb(11,3)','H^2-Vb(10,6)', 'H^2-Vb(10,4)','H^2-Vb(7,9)', 'H^2-Vb(7,3)','H^2-Vb(6,1)', 'H^2-Vb(4,11)','H^2-Vb(3,11)', 'H^2-Vb(3,7)','H^2-Vb(1,10)', 'y-Vb(11,4)','y-Vb(11,3)','y-Vb(10,6)', 'y-Vb(10,1)','y-Vb(6,1)','y-Vb(4,11)','y-Vb(3,11)', 'y-Vb(1,6)', 'FW_H^2-Vb(11,10)','FW_H^2-Vb(11,4)' ,'FW_H^2-Vb(11,3)', 'FW_H^2-Vb(10,6)', 'FW_H^2-Vb(7,3)','FW_H^2-Vb(4,11)', 'FW_H^2-Vb(3,7)', 'FW_H^2-Vb(1,10)','FW_H-Vb(10,4)', 'FW_H-Vb(10,1)','FW_H-Vb(7,3)','FW_H-Vb(3,7)', 'FW_H-Vb(1,10)']	SC7	80.79
Bipolar, Relieff, BPSO-KNN-Holdou	DC2	['R^2-Vb(11,3)','R^2-Vb(10,1)', 'R^2-Vb(9,7)', 'R^2-Vb(6,10)', 'R^2-Vb(6,1)','R^2-Vb(4,11)', 'R^2-Vb(1,10)','R^2-Vb(1,6)',' H^2-Vb(11,10)','H^2-Vb(10,6)' ,'H^2-Vb(10,4)','H^2-Vb(10,1)' ,'H^2-Vb(9,7)','H^2-Vb(6,1)', 'H^2-Vb(4,11)','H^2-Vb(1,10)', 'y-Vb(11,10)','y-Vb(11,4)','y-Vb(10,4)' ,'y-Vb(10,1)','y-Vb(9,7)','y-Vb(4,11)','y-Vb(3,11)', 'y-Vb(3,7)','y-Vb(1,10)','y-Vb(1,6)', 'FW_H^2-Vb(10,11)', 'FW_H^2-Vb(9,7)','FW_H^2-Vb(7,9)', 'FW_H^2-Vb(7,3)',' FW_H^2-Vb(3,11)','FW_H-Vb(10,6)','FW_H-Vb(10,4)', 'FW_H-Vb(9,7)', 'FW_H-Vb(6,1)','FW_H-Vb(3,7)']	SC8	88.50
Monopolar, F-score, GA-KNN-KFOLD	DC3	['R^2-Ch(8,1)','R^2-Ch(16,9)','R^2-Ch(15,9)', 'R^2-Ch(14,12)','R^2-Ch(12,14)', 'R^2-Ch(11,1)','R^2-Ch(10,1)', 'R^2-Ch(9,16)','R^2-Ch(9,15)', 'R^2-Ch(9,2)', 'R^2-Ch(6,16)', 'R^2-Ch(5,8)','R^2-Ch(4,9)', 'H^2-Ch(10,2)','H^2-Ch(2,10)', 'H^2-Ch(16,9)','H^2-Ch(15,10)' , 'H^2-Ch(15,9)', H^2-Ch(14,12)','H^2-Ch(12,14)','H^2-Ch(11,8)', 'H^2-Ch(9,10)','H^2-Ch(9,2)', 'H^2-Ch(7,5)','H^2-Ch(5,15)', 'H^2-Ch(5,9)' ,'H^2-Ch(5,8)', 'H^2-	SC9	91.02

		Ch(3,5)','y-Ch(8,1)','y-Ch(16,9) ', 'y-Ch(15,10)','y-Ch(15,5)','y-Ch(14,12)','y-Ch(10,15)','y-Ch(9,16)','y-Ch(9,5)','y-Ch(9,4)','y-Ch(7,5)','y-Ch(5,9)','FW_H^2-Ch(10,2)','FW_H^2-Ch(8,1)', 'FW_H^2-Ch(2,10)','FW_H^2-Ch(14,12)', 'FW_H^2-Ch(11,8)' , 'FW_H^2-Ch(10,9)','FW_H^2-Ch(9,15)','FW_H^2-Ch(5,15)', 'FW_H^2-Ch(5,7)', 'FW_H^2-Ch(5,3)','FW_H^2-Ch(4,1)', 'FW_H^2-Ch(3,5)','FW_H^2-Ch(1,4)','FW_H-Ch(15,5)', 'FW_H-Ch(10,15)','FW_H-Ch(3,5)']		
Monopolar, F-score, GA-KNN-Holdout	DC3	['R^2-Ch(10,2)','R^2-Ch(8,1)','R^2-Ch(1,8)','R^2-Ch(16,9)','R^2-Ch(15,10)', 'R^2-Ch(15,9)','R^2-Ch(14,12)','R^2-Ch(11,1)', 'R^2-Ch(9,15)','R^2-Ch(9,4)', 'R^2-Ch(8,5)','R^2-Ch(7,5)', 'R^2-Ch(6,16)', 'R^2-Ch(5,15)', 'R^2-Ch(3,5)','R^2-Ch(1,4)', 'H^2-Ch(10,2)','H^2-Ch(8,1)', 'H^2-Ch(2,10)','H^2-Ch(1,8)' , 'H^2-Ch(16,10)','H^2Ch(16,9)','H^2-Ch(16,6)','H^2-Ch(15,10)','H^2-Ch(14,12)', 'H^2-Ch(11,8)','H^2-Ch(11,1)',' H^2-Ch(10,9)','H^2-Ch(10,1)', 'H^2-Ch(9,15)','H^2-Ch(9,10)', 'H^2-Ch(9,5)','H^2-Ch(9,2)' , 'H^2-Ch(7,5)','H^2-Ch(6,16)', 'H^2-Ch(5,9)' , 'H^2-Ch(5,7)', 'H^2-Ch(5,3)','H^2-Ch(4,9)', 'H^2-Ch(4,1)' ,'H^2-Ch(3,5)','y-Ch(1,8)','y-Ch(16,9)', 'y-Ch(16,6)' ,'y-Ch(15,5)','y-Ch(14,12)','y-Ch(11,1)' ,'y-Ch(10,15)','y-Ch(9,10)','y-Ch(9,4)','y-Ch(8,5)' ,'y-Ch(5,9)','y-Ch(3,5)','FW_H^2-Ch(16,6)' ,'FW_H^2-Ch(11,8)', 'FW_H^2-Ch(11,1)','FW_H^2-Ch(8,11)', 'FW_H^2-Ch(8,5)', 'FW_H^2-Ch(7,5)','FW_H^2-Ch(5,15)','FW_H^2-Ch(5,8)', 'FW_H-Ch(8,1)', 'FW_H-Ch(1,8)', 'FW_H-Ch(16,6)', 'FW_H-Ch(15,5)' ,'FW_H-Ch(11,8)', 'FW_H-Ch(11,1)', 'FW_H-Ch(9,4)', 'FW_H-Ch(8,11)','FW_H-Ch(6,16)', 'FW_H-Ch(5,3)']	SC10	93.81
Monopolar, F-score, BPSO-KNN-KFOLD	DC3	['R^2-Ch(10,2)','R^2-Ch(16,9)','R^2-Ch(15,9)','R^2-Ch(14,12)','R^2-Ch(11,1)', 'R^2-Ch(10,9)','R^2-Ch(10,1)', 'R^2-Ch(9,16)','R^2-Ch(9,15)', 'R^2-Ch(9,10)','R^2-Ch(9,2)', 'R^2-Ch(8,11)' ,'R^2-Ch(5,9)', 'R^2-	SC11	85.79

		Ch(5,8)','R^2-Ch(5,3)', 'R^2-Ch(4,9)','R^2-Ch(4,1)', 'R^2-Ch(3,5)','R^2-Ch(1,4)', 'H^2-Ch(10,2)', 'H^2-Ch(16,9)', 'H^2-Ch(16,6)','H^2-Ch(14,12)' ,'H^2-Ch(11,8)', 'H^2-Ch(10,15)','H^2-Ch(9,16)', 'H^2-Ch(9,15)','H^2-Ch(5,15)','H^2-Ch(5,7)','H^2-Ch(5,3)', 'H^2-Ch(4,9)','H^2-Ch(4,1)','y-Ch(2,10)','y Ch(16,10)','y-Ch(16,9)','y-Ch(16,6)','y-Ch(15,10)' ,'y-Ch(15,9)', 'y-Ch(14,12)','y-Ch(11,1)','y-Ch(10,1)','y-Ch(9,16)' ,'y-Ch(9,15)','y-Ch(8,11)','y-Ch(8,5)','y-Ch(7,5)','y-Ch(6,16)','y-Ch(5,15)','y-Ch(5,9)','y-Ch(4,9)','y-Ch(1,4)','FW_H^2-Ch(10,2)' ,'FW_H^2-Ch(8,1)', 'FW_H^2-Ch(16,10)','FW_H^2-Ch(16,9)','FW_H^2-Ch(15,9)', 'FW_H^2-Ch(15,5)', 'FW_H^2-Ch(14,12)', 'FW_H^2-Ch(11,8)' , 'FW_H^2-Ch(10,9)','FW_H^2-Ch(10,1)','FW_H^2-Ch(9,10)', 'FW_H^2-Ch(9,4)', 'FW_H^2-Ch(9,2)', 'FW_H^2-Ch(8,5)', 'FW_H^2-Ch(5,15)','FW_H^2-Ch(5,3)', 'FW_H^2-Ch(4,1)', 'FW_H^2-Ch(1,4)','FW_H-Ch(1,8)', 'FW_H-Ch(16,10)' ,'FW_H-Ch(15,10)', 'FW_H-Ch(15,5)', 'FW_H-Ch(10,15)' ,'FW_H-Ch(7,5)']		
Monopolar, F-score, BPSO-KNN-Holdout	DC3	['R^2-Ch(2,10)','R^2-Ch(1,8)', 'R^2-Ch(16,9)', 'R^2-Ch(16,6)' ,'R^2-Ch(12,14)','R^2-Ch(11,8)','R^2-Ch(11,1)','R^2-Ch(10,9)','R^2-Ch(9,15)','R^2-Ch(9,10)','R^2-Ch(9,5)','R^2-Ch(9,2)','R^2-Ch(8,5)', 'R^2-Ch(7,5)','R^2-Ch(5,9)','R^2-Ch(5,8)','R^2-Ch(5,7)','R^2-Ch(4,1)','R^2-Ch(3,5)','R^2-Ch(1,4)', 'H^2-Ch(16,10)','H^2-Ch(16,9)','H^2-Ch(15,9)','H^2-Ch(15,5)','H^2-Ch(12,14)' ,'H^2-Ch(10,15)','H^2-Ch(10,1)' ,'H^2-Ch(9,15)','H^2-Ch(9,5)','H^2-Ch(8,11)','H^2-Ch(7,5)','H^2-Ch(6,16)','H^2-Ch(5,15)','H^2-Ch(5,7)','H^2-Ch(5,3)','H^2-Ch(4,9)', 'H^2-Ch(4,1)','H^2-Ch(3,5)','y-Ch(10,2)','y-Ch(8,1)' ,'y-Ch(1,8)','y-Ch(16,9)','y-Ch(16,6)','y-Ch(15,10)', 'y-Ch(15,5)','y-Ch(12,14)','y-Ch(10,15)','y-Ch(9,10)' ,'y-Ch(9,5)','y-Ch(9,4)','y-Ch(8,5)','y-Ch(6,16)','y-Ch(5,9)','y-Ch(4,1)','FW_H^2-Ch(15,10)' , 'FW_H^2-Ch(15,5)','FW_H^2-	SC12	91.15

		Ch(12,14)' ,'FW_H^2-Ch(11,8)', 'FW_H^2-Ch(11,1)' ,'FW_H^2-Ch(10,15)','FW_H^2-Ch(10,9)', 'FW_H^2-Ch(10,1)' , 'FW_H^2-Ch(9,16)', 'FW_H^2-Ch(9,15)', 'FW_H^2-Ch(8,5)','FW_H^2-Ch(5,9)', 'FW_H^2-Ch(5,8)', 'FW_H^2-Ch(5,3)','FW_H-Ch(10,2)', 'FW_H-Ch(1,8)','FW_H-Ch(15,10)','FW_H-Ch(15,5)', 'FW_H-Ch(14,12)','FW_H-Ch(12,14)', 'FW_H-Ch(11,8)', 'FW_H-Ch(10,9)','FW_H-Ch(10,1)','FW_H-Ch(9,16)', 'FW_H-Ch(9,5)' ,'FW_H-Ch(8,11)', 'FW_H-Ch(7,5)', 'FW_H-Ch(5,9)', 'FW_H-Ch(5,8)','FW_H-Ch(4,1)', 'FW_H-Ch(3,5)','FW_H-Ch(1,4)']		
Monopolar, Relieff, GA-KNN-KFOLD	DC4	['R^2-Ch(15,14)','R^2-Ch(14,12)','R^2-Ch(14,11)', 'R^2-Ch(14,5)','R^2-Ch(14,2)', 'R^2-Ch(12,14)','R^2-Ch(12,6)' ,'R^2-Ch(11,14)','R^2-Ch(10,13)','R^2-Ch(9,13)', 'R^2-Ch(7,8)','R^2-Ch(6,12)', 'R^2-Ch(3,16)','R^2-Ch(2,7)', 'H^2-Ch(15,14)', 'H^2-Ch(14,12)','H^2-Ch(14,11)', 'H^2-Ch(14,5)','H^2-Ch(13,10)' ,'H^2-Ch(12,14)','H^2-Ch(12,6)', 'H^2-Ch(11,10)' ,'H^2-Ch(10,13)','H^2-Ch(10,11)','H^2-Ch(9,13)' ,'H^2-Ch(9,1)','H^2-Ch(7,8)',' H^2-Ch(6,12)','H^2-Ch(5,14)', 'H^2-Ch(4,7)','H^2-Ch(3,16)', 'H^2-Ch(2,7)','H^2-Ch(1,14)' , 'y-Ch(15,14)','y-Ch(14,5)','y-Ch(13,10)', 'y-Ch(12,6)','y-Ch(11,14)','y-Ch(10,13)','y-Ch(9,13)', 'y-Ch(9,1)','y-Ch(8,7)','y-Ch(7,8)','y-Ch(5,14)','y-Ch(4,7)','y-Ch(3,16)','y-Ch(2,7)','y-Ch(1,9)', 'FW_H^2-Ch(15,14)', 'FW_H^2-Ch(14,11)', 'FW_H^2-Ch(14,5)','FW_H^2-Ch(13,10)','FW_H^2-Ch(12,6)' , 'FW_H^2-Ch(11,14)',' FW_H^2-Ch(10,13)','FW_H^2-Ch(9,1)', 'FW_H^2-Ch(7,8)', 'FW_H^2-Ch(3,16)', 'FW_H^2-Ch(2,7)', 'FW_H-Ch(14,11)', 'FW_H-Ch(11,14)','FW_H-Ch(5,14)', 'FW_H-Ch(4,7)']	SC13	86.58
Monopolar, Relieff, GA-KNN-Holdout	DC4	['R^2-Ch(15,14)','R^2-Ch(14,12)', 'R^2-Ch(14,11)', 'R^2-Ch(14,5)','R^2-Ch(13,10) ','R^2-Ch(12,14)','R^2-Ch(12,6)','R^2-Ch(11,14)' ,'R^2-Ch(11,10)','R^2-	SC14	94.69

		Ch(10,11)','R^2-Ch(9,13)',' R^2-Ch(9,1)','R^2-Ch(3,16)', 'R^2-Ch(2,7)','R^2-Ch(1,14)', 'H^2-Ch(14,11)','H^2-Ch(14,2)', 'H^2-Ch(14,1)','H^2-Ch(13,10)','H^2-Ch(12,14)', 'H^2-Ch(11,14)','H^2-Ch(11,10)', 'H^2-Ch(10,11)', 'H^2-Ch(9,13)','y-Ch(14,11)', 'y-Ch(14,1)','y-Ch(12,6)' ,'FW_H^2-Ch(14,1)','FW_H^2-Ch(12,6)','FW_H^2-Ch(10,13)' , 'FW_H^2-Ch(6,12)','FW_H^2-Ch(4,7)', 'FW_H^2-Ch(2,7)', 'FW_H^2-Ch(1,14)','FW_H-Ch(14,11)']		
Monopolar, Relieff, BPSO-KNN-KFOLD	DC4	['R^2-Ch(15,14)','R^2-Ch(14,11)','R^2-Ch(14,5)',' R^2-Ch(14,1)','R^2-Ch(12,14)' ,'R^2-Ch(11,14)','R^2-Ch(11,10)' ,'R^2-Ch(9,13)',' R^2-Ch(9,1)','R^2-Ch(8,7)', 'R^2-Ch(6,12)','R^2-Ch(3,16)',' R^2-Ch(2,7)','R^2-Ch(1,9)', 'H^2-Ch(14,12)','H^2-Ch(14,11)','H^2-Ch(14,5)', 'H^2-Ch(14,2)','H^2-Ch(14,1)' ,'H^2-Ch(12,6)','H^2-Ch(11,14)','H^2-Ch(11,10)', 'H^2-Ch(10,11)', 'H^2-Ch(8,7)','H^2-Ch(4,7)','H^2-Ch(2,7)','y-Ch(15,14)','y-Ch(14,12)', 'y-Ch(13,10)','y-Ch(12,6)','y-Ch(11,10)','y-Ch(10,13)','y-Ch(9,13)','y-Ch(9,1)','y-Ch(6,12)','y-Ch(4,7)','y-Ch(1,9)', 'FW_H^2-Ch(14,11)','FW_H^2-Ch(14,5)', 'FW_H^2-Ch(14,2)', 'FW_H^2-Ch(14,1)','FW_H^2-Ch(12,6)', 'FW_H^2-Ch(4,7)', 'FW_H^2-Ch(3,16)','FW_H-Ch(11,14)', 'FW_H-Ch(6,12)', 'FW_H-Ch(2,7)']	SC15	83.42
Monopolar, Relieff, BPSO-KNN-Holdout	DC4	['R^2-Ch(15,14)','R^2-Ch(14,11)','R^2-Ch(14,5)', 'R^2-Ch(12,14)','R^2-Ch(12,6),'R^2-Ch(11,10)','R^2-Ch(10,13)','R^2-Ch(9,1)','R^2-Ch(8,7)','R^2-Ch(6,12)','R^2-Ch(5,14)','R^2-Ch(4,7)','R^2-Ch(2,7)','H^2-Ch(15,14)','H^2-Ch(14,11)','H^2-Ch(14,5)"H^2-Ch(14,2)','H^2-Ch(14,1)','H^2-Ch(13,10)','H^2-Ch(12,14), 'H^2-Ch(12,6)','H^2-Ch(10,13)' ,'H^2-Ch(9,13)','H^2-Ch(8,7)' ,'H^2-Ch(7,8)','H^2-Ch(6,12)', 'H^2-Ch(5,14)','H^2-Ch(3,16)' ,'H^2-Ch(2,7)','H^2-Ch(1,14)', 'y-Ch(14,5)','y-Ch(12,6)','y-Ch(8,7)','y-Ch(7,8)','y-Ch(6,12)','y-Ch(3,16)', 'FW_H^2-Ch(15,14)', 'FW_H^2-	SC16	91.15

		Ch(14,12)','FW_H^2-Ch(13,10)' , 'FW_H^2-Ch(12,14)', 'FW_H^2-Ch(9,13)', 'FW_H^2-Ch(9,1)', 'FW_H^2-Ch(7,8)', 'FW_H^2-Ch(3,16)', 'FW_H^2-Ch(2,7)',' FW_H-Ch(14,1)','FW_H-Ch(7,8)','FW_H-Ch(3,16)']		

(1)Method names:

- Recording configuration: Bipolar/Monopolar
- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

(2)Dataset:

- DC1: The datasets DC1 contain the 5 features extracted, for the 379 EHG bursts, from the 17 retained bipolar channel combinations for F-score (379*85).
- DC2: The datasets DC2 contain the 5 features extracted, for the 379 EHG bursts, from the 17 retained bipolar channel combinations for Relieff (379*85).
- DC3: The monopolar datasets DC3 correspond to the features extracted, for the 379 EHG bursts, from the 36 retained monopolar channel combinations for F-score (379*180)
- DC4: The monopolar datasets DC4 correspond to the features extracted, for the 379 EHG bursts, from the 24 retained monopolar channel combinations for Relieff (379*120).

Table C.5.2: Number of features contain in the selected subsets after feature selection for the selected channel combinations and time of calculation of each method of feature selection

Methods[(1)]	**Selected feature Subset**[(2)]	**Number**	**Correct classification percentage**	**Time (s)**
Bipolar, F-score, GA-KNN-KFOLD	SC1	25	88.39	161.62
Bipolar, F-score, GA-KNN-Holdout	SC2	41	94.69	27.83
Bipolar, F-score, BPSO-KNN-KFOLD	SC3	35	86.84	50
Bipolar, F-score, BPSO-KNN-Holdout	SC4	45	95.58	7.68
Bipolar, Relieff, GA-KNN-KFOLD	SC5	18	82.86	155.43
Bipolar, Relieff, GA-KNN-Holdout	SC6	37	92.04	27.47
Bipolar, Relieff, BPSO-KNN-KFOLD	SC7	44	80.79	49.93
Bipolar, Relieff, BPSO-KNN-Holdout	SC8	36	88.50	7.64
Monopolar, F-score, GA-KNN-KFOLD	SC9	55	91.02	208.50
Monopolar, F-score, GA-KNN-Holdout	SC10	71	93.81	36.34
Monopolar, F-score, BPSO-KNN-KFOLD	SC11	75	85.79	67.09
Monopolar, F-score, on BPSO-KNN-Holdout	SC12	86	91.15	13
Monopolar, Relieff, GA-KNN-KFOLD	SC13	63	86.58	215.87
Monopolar, Relieff, GA-KNN-Holdout	SC14	35	94.69	27.17
Monopolar, Relieff, BPSO-KNN-KFOLD	SC15	47	83.42	59.03
Monopolar, Relieff, BPSO-KNN-Holdout	SC16	48	91.15	8.75

[1]Method names:

- Recording configuration: Bipolar/Monopolar
- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

[2]Selected feature subset: correspond to the subset obtained in table C.5.1

C.6- Validation of the selected features subset after channel combinations selection

Table C.6.1: Mean ± STD of the percentage of correct classification using KNN of 500 repetitions for the selected features subset from the bipolar and monopolar selected channel combinations

Methods[(1)]	**Selected feature subset**[(2)]	**Mean ± standard deviation of percentage of correct classification (500 repetitions)**
Bipolar, F-score, GA-KNN-KFOLD	SC1	76.01 ± 1.33
Bipolar, F-score, GA-KNN-Holdout	SC2	75.69 ± 7.31
Bipolar, F-score, BPSO-KNN-KFOLD	SC3	75.37 ± 1.62
Bipolar, F-score, BPSO-KNN-Holdout	SC4	73.30 ± 7.27
Bipolar, Relieff, GA-KNN-KFOLD	SC5	64.56 ± 2.03
Bipolar, Relieff, GA-KNN-Holdout	SC6	64.03 ± 6.93
Bipolar, Relieff, BPSO-KNN-KFOLD	SC7	66.62 ± 1.62
Bipolar, Relieff, BPSO-KNN-Holdout	SC8	65.93 ± 7.85
Monopolar, F-score, GA-KNN-KFOLD	SC9	74.85± 2.26
Monopolar, F-score, GA-KNN-Holdout	SC10	70.41± 8.02
Monopolar, F-score, BPSO-KNN-KFOLD	SC11	70.90 ± 1.96
Monopolar, F-score, on BPSO-KNN-Holdout	SC12	67.49 ± 8.16
Monopolar, Relieff, GA-KNN-KFOLD	SC13	60.06 ± 2.10
Monopolar, Relieff, GA-KNN-Holdout	SC14	61.45 ± 9.22
Monopolar, Relieff, BPSO-KNN-KFOLD	SC15	64.18 ± 2.37
Monopolar, Relieff, BPSO-KNN-Holdout	SC16	60.88 ± 7.58

[(1)]Method names:
- Recording configuration: Bipolar/Monopolar
- Channel selection : F-score/Relieff
- Feature selection: GA-KNN-KFOLD, GA-KNN-Holdout, BPSO-KNN-KFOLD, BPSO-KNN-Holdout

[(2)]Selected feature subset: correspond to the subset obtained in table C.5.1

Summary:

One of the most promising biophysical markers of preterm labor is the electrical activity of the uterus, picked up on woman's abdomen, the electrohysterogram (EHG). Several processing tools of the EHG signal (linear, nonlinear), allow the analysis of both excitability and propagation of the uterine electrical activity in order to differentiate between pregnancy contractions, which are ineffective, from labor effective contractions that might cause preterm birth. Therefore, on these multiple studies, the parameters being computed from different signal databases, obtained with different recording protocols, it is sometimes difficult to compare their results in order to choose the "best" parameter for preterm labor detection. Additionally, this large number of parameters increases the computational complexity for diagnostic purpose. Therefore, the main objective of this thesis is to select, among all the features of interest extracted from multiple studies, the most pertinent feature subsets in order to discriminate, on a given population, pregnancy and labor contractions. For this purpose, several methods for feature selection are tested. The first one, developed in this work, is based on the measurement of the Jeffrey divergence (JD) distance between the histograms of the parameters of the 2 classes, pregnancy and labor. The other are "Filter" and "Wrapper" Data Mining methods, extracted from the literature. In our work monovariate (in one given EHG channel) and bivariate analysis (propagation of EHG by measuring the coupling between channels) are used. The EHG signals are recorded using a multichannel system positioned on the woman's abdomen for the simultaneous recording of 16 channels of EHG. Using all channels, for the monovariate, or all combinations of channels for the bivariate analysis, leads to a large dimension of parameters for each contraction. Therefore, another objective of our thesis is the selection of the best channels, for the monovariate, or best channel combinations, for the bivariate analysis, that provide the most useful information to discriminate between pregnancy and labor classes. This channel selection step is then followed by the feature selection for the channels or channel combinations selected. Additionally, we tested all our work using monopolar and bipolar signals.

The results of this thesis permits us to evidence, when processing the EHG, which channels and features can be used with the best chance of success as inputs of a diagnosis system for discrimination between pregnancy and labor contractions. This could be further used for preterm labor diagnosis.

Keyword:

EHG; Preterm labor; Parameters extraction; Feature selection; Channel selection.

Résumé:

Un des marqueurs biophysique le plus prometteur pour la détection des accouchements prématurés (AP) est l'activité électrique de l'utérus, enregistrée sur l'abdomen des femmes enceintes, l'électrohystérogramme (EHG). Plusieurs outils de traitement du signal (linéaires, non linéaires) ont déjà été utilisés pour l'analyse de l'excitabilité et de la propagation de l'EHG, afin de différencier les contractions de grossesse, qui sont inefficaces, des contractions efficaces d'accouchement, qui pourraient provoquer un AP. Dans ces études nombreuses, les paramètres sont calculés sur des bases de données de signaux différentes, obtenus avec des protocoles d'enregistrement différents. Il est donc difficile de comparer les résultats afin de choisir les «meilleurs» paramètres pour la détection de l'AP. En outre, ce grand nombre de paramètres augmente la complexité de calcul dans un but de diagnostic. Par conséquent, l'objectif principal de cette thèse est de tester, sur une population de femmes donnée, quels outils de traitement du signal EHG permettent une discrimination entre les deux types de contractions (grossesse/accouchement). Dans ce but plusieurs méthodes de sélection de paramètres sont testées afin de sélectionner les paramètres les plus discriminants. La première méthode, développée dans cette thèse, est basée sur la mesure de la distance entre les histogrammes des paramètres pour les différentes classes (grossesse et accouchement) en utilisant la méthode « Jeffrey divergence (JD)». Les autres sont des méthodes de fouille de données existantes issues de la littérature. Les EHG ont été enregistrés en utilisant un système multivoies posé sur l'abdomen de la femme enceinte, pour l'enregistrement simultané de 16 voies d'EHG. Une approche monovariée (caractérisation d'une seule voie) et bivariée (couplage entre deux voies) sont utilisées dans notre travail. Utiliser toutes les voies, analyse monovariée, ou toutes les combinaisons de voies, analyse bivariée, conduit à une grande dimension des paramètres. Par conséquent, un autre objectif de notre thèse est la sélection des voies, ou des combinaisons de voies, qui fournissent l'information la plus utile pour distinguer entre les contractions de grossesse et d'accouchement. Cette étape de sélection de voie est suivie par la sélection des paramètres, sur les voies ou les combinaisons de voies sélectionnées. De plus, nous avons développé cette approche en utilisant des signaux monopolaires et bipolaires.
Les résultats de ce travail nous permettent de mettre en évidence, lors du traitement de l'EHG, les paramètres et les voies qui donnent la meilleure discrimination entre les contractions de grossesse et celles d'accouchement. Ces résultats pourront ensuite être utilisés pour la détection des menaces d'accouchement prématuré.

Mot clés :

EHG ; Accouchements prématurés ; Extraction des paramètres ; Sélection de paramètres ; Sélection des voies.