Detection of Fetal Kicks Using Body-Worn Accelerometers During Pregnancy: Trade-offs Between Sensors Number and Positioning

Marco Altini[1], Patrick Mullan[2], Michiel Rooijakkers[1], Stefan Gradl[2], Julien Penders[1], Nele Geusens[3], Lars Grieten[3], and Bjoern Eskofier[2]

***Abstract*— Monitoring fetal wellbeing is key in modern obstetrics. While fetal movement is routinely used as a proxy to fetal wellbeing, accurate, noninvasive, long-term monitoring of fetal movement is challenging. A few accelerometer-based systems have been developed in the past few years, to tackle common issues in ultrasound measurement and enable remote, self-administrated monitoring of fetal movement during pregnancy. However, many questions remain unanswered to date on the optimal setup in terms of body-worn accelerometers as well as signal processing and machine learning techniques used to detect fetal movement. In this paper, we systematically analyze the trade-offs between sensor number and positioning, the presence of reference accelerometers outside of the abdominal area and provide guidelines on dealing with class imbalance. Using a dataset of 15 measurements collected employing 6 three-axial accelerometers we show that including a reference accelerometer on the back of the participant consistently improves fetal movement detection performance regardless of the number of sensors utilized. We also show that two accelerometers plus a reference accelerometer are sufficient for optimal results.**

I. INTRODUCTION

Being able to monitor fetal wellbeing during pregnancy is one of the main challenges of obstetrics today, since birth outcomes are strongly linked to the development of fetal conditions during pregnancy and not only during labour [1]. Thus, different methods to monitor fetal wellbeing during pregnancy were introduced in the past years [2]. One of such fetal wellbeing monitoring techniques, is monitoring of fetal movement. Pregnant women can start feeling fetal movement as early as during the first trimester. Absence of maternal perception of fetal movement is one symptom of fetal death, and a reduction in fetal movement is an alarming sign of fetal compromise [3]. Additionally, fetal movement is considered one of the fundamental expressions of early neural activity as it is generated spontaneously by the central nervous system [4] and it is therefore often considered a good proxy of fetal wellbeing.

Standard clinical practice for fetal wellbeing and movement monitoring relies on different methods, which can be grouped into active, passive and self-reported. Active methods such as ultrasound rely on high frequency sound waves being used to generate an image of the fetus and can be used only for a limited amount of time. While no negative associations were found between diagnostic ultrasound exposure during pregnancy and birth outcomes, safety concerns still require further investigation [5], [6]. Another common method for fetal monitoring is continuous cardiotocography (CTG), which requires bulky infrastructure and can only be used in the hospital environment with trained personnel, for short periods of time [7], [8]. The inability of these methods to monitor fetal movement outside of sporadic spot checks in the hospital environment is one of the major causes of concern and motivations behind the development of other passive methods for home-monitoring, such as accelerometer based solutions. Stillbirths are a major issue around the world today, in both developing and developed countries [15], further motivating the need for better and cheaper monitoring tools.

An alternative passive method which has been investigated in the past few years is the possibility to use on-body accelerometers to monitor fetal movement. Accelerometer based systems are safe, inexpensive, can potentially be used autonomously in the home environment and showed promising results in preliminary studies [9], [10], [11], [12], [13]. Finally, fetal movement can also be self-reported by the mother using a so called kick-chart, showing inconsistent results in literature (sensitivity between 37% and 88%) for different reasons [14]. For example fetal movement itself is not well-defined, possibly leading some researchers to count certain types of fetal movements instead of others. Secondly, the time windows in which the mothers perception has to match the ultrasound images is not consistent.

Miniaturized wearable sensors including on-board accelerometers can provide a way to investigate passively and safely fetal movement inside [10] or outside [9] the hospital settings. Additionally, advances in signal processing and machine learning techniques, recently provided higher accuracy in determining fetal movement using on-body accelerometers, compared to preliminary studies using simpler threshold-based methods [11], [12]. Accelerometer-based systems could replace kick charts in unsupervised free-living settings, providing a more objective and consistent quantification of fetal movement while freeing pregnant women from this task.

While a few different accelerometer-based solutions have been proposed in literature in the past few years to monitor fetal movement, inconsistencies between study protocols, accelerometer placement, number of sensors and signal processing techniques make it hard to understand what setup is best and what are the trade-offs between alternative methods.

This work was funded by Bloom Technologies

[1]M. Altini, M. Rooijakkers and J. Penders are with Bloom Technologies, Diepenbeek, BE altini.marco@gmail.com

[2]P. Mullan, S. Gradl and B. Eskofier are with the Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuernberg, DE

[3]N. Geusens and L. Grieten are with the Department of Future Health, Ziekenhuis Oost-Limburg, Genk, BE

Finally, differences in reference methods and evaluation metrics make it impossible to compare different studies and determine the efficacy of each technique.

In this paper, we evaluate performance improvements in fetal kicks detection when using multiple accelerometers positioned on different locations on the body, using a dataset comprising comprising 15 measurements from 6 on-body accelerometers, including one reference accelerometer placed on the back. In particular, we show that two accelerometers plus a reference accelerometer are sufficient for optimal results and that a reference accelerometer is necessary to discriminate maternal movement regardless of the number of sensors used. We also discuss several points related to trade-offs and design choices relevant in the context of developing a fetal movement detection algorithm (window size, class imbalance, choice of classifier, performance metrics), in order to provide a clear framework and ease comparisons with future works.

II. RELATED WORKS

Related works can be grouped according to different criteria: number of sensors used, presence of a reference accelerometer placed outside of the abdominal area and data analysis techniques used. Most studies involved one single accelerometer placed on the abnormal area and reported rather low sensitivity and specificity [16]. Comparison between studies is challenging due to the different reference and evaluation methods. However, single-accelerometer systems typically reported detection rates around 50%, deemed insufficient by the researchers themselves [16]. The rationale behind the addition of a reference accelerometer is that by monitoring maternal movement artifacts using an accelerometer placed outside of the abdominal area, fetal movement should be separable from maternal movement and therefore detected more accurately [9]. However, accelerometer placement should be outside of the abdominal and thoracic area, since accelerometer placement on the upper thoracic area was still able to detect fetal movement and was therefore unusable as reference [17]. The difference in movement detection performance when including or excluding the reference accelerometer is not reported by the previous studies, and often used as post-processing signal to discard data more than to inform the classification process [11], [9].

Data analysis techniques used up to today mainly focused on feature extraction by means of time (e.g. the magnitude of the acceleration vector [18]) and frequency domain signal processing techniques [16], [9], [17] and only recently touched machine learning techniques such as using Support Vector Machines to classify a set of features into a binary problem (movement vs no-movement). While determining optimal features is a necessary first step, thresholding on a single feature provided poor results and combining multiple features and machine learning methods has the potential for more accurate fetal movement detection [11]. In the context of using supervised learning methods to classify movements and non-movements, an additional challenge arises. Fetal movement occurs only for a short percentage of the time during a measurement, therefore proper methods such as downsampling of the majority class (i.e. no-movement) need to be employed. However, evaluation of the method should be performed on the entire data stream and not only on chunks of data pre-selected by the researchers, as reported in [11]. Other design choices concern the windows size on which to compute features, the choice of classifier and possibly feature selection method, performance metrics used to evaluate the system and finally the reference system used to validate fetal movement detection algorithms.

Most studies relied on ultrasound as reference for fetal movement. While ultrasound is the clinical standard, limitations apply, even during research studies. For example, with fetal growth it becomes impossible to fully display the fetus given the limited field of vision of the ultrasound probe, starting at approximately week 20. While this is not a problem during hospital checkups, moving and re-positioning the probe while trying to measure small accelerations as reflected on the pregnant women abdomen is impractical and can easily introduce noise. In this study we used maternal perception and expert annotations as reference. While there are limitations also in maternal perception, there is no trustworthy reference for fetal movement. By analyzing the algorithms performance and trade offs with respect to the same reference, we could get a better understanding of the influence of different sensor number, positioning and data analysis methods in effectively detecting fetal movement.

III. DATA ACQUISITION

A. Accelerometers Data and Reference

Fifteen recordings were collected from 6 pregnant women at different time points during pregnancy, all from week 30 onwards. Measurements were performed using the Porti7 device from Twente Medical Systems International (TMSi). The Porti7 is a 32 channel analog-to-digital converter able of sampling up to 2048 Hz with a resolution of 22 bit. To reduce computational complexity the signals were downsampled to 128 Hz before data analysis. Accelerometer data were also bandpass filtered between 1 and 20 Hz with a second order butterworth IIR filter since fetal movement is expected to be in this frequency band [11]. All pregnant women were given a handheld toggle which they were advised to press when feeling fetal movement. The output of the button was always used as reference for fetal movements. The experimenter manually annotated fetal movements as a pre-processing step, by locating accelerometer movements anticipating button triggers. Finally, an experience midwife also collected reference movement data by visually analyzing the pregnant women abdomen during the measurement. Both references were combined and only fetal kicks were considered movement in this study to ensure more consistency across participants and annotations. During data collection, pregnant women had to lie down. Five accelerometer sensors were positioned on the abdomen with the navel serving as central marker. The sixth sensor was placed on the back. Exact positions of the accelerometers can be seen in Fig. 1.

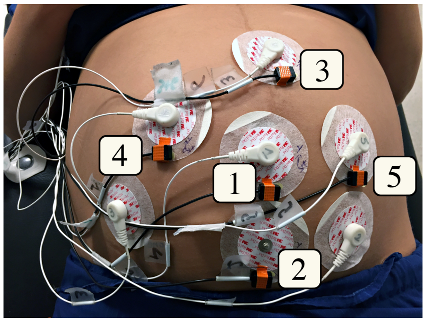

Fig. 1. On-body accelerometers placement for the 5 accelerometers placed on the abdomen. The sixth accelerometer, placed on the back, is not visible. Also visible are electrodes used to acquire ExG data, not used in this study.

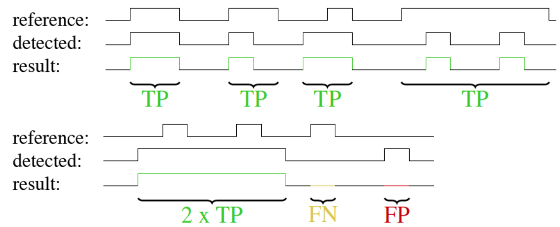

Fig. 2. Graphical example of our evaluation strategy. TP = true positives, FN = false negatives, FP = false positives.

IV. DATA ANALYSIS

Several design choices need to be made when developing a method to detect fetal movement using body-worn accelerometers, from feature computation to selecting the proper performance metrics. In this section, we provide an overview of the design choices and validation techniques used before we could analyze trade-offs between sensors number and positioning.

A. Features

We computed features over 0.5 seconds non-overlapping windows of accelerometer bandpass filtered data. We computed low-complexity time domain features to possibly enable easy implementation on an embedded device. Features were: mean, standard deviation, interquartile range, correlation between axis and correlation with the reference sensor over all axis. Each feature was computed per axis and per sensor, for a total of 83 parameters. We chose 0.5 seconds windows given the short duration of fetal kicks. Longer time windows showed an averaging out of the signal during our exploratory data analysis.

B. Features Selection, Class Imbalance and Classification

We chose random forests as classifiers in order to exploit a few advantages. During training, random forests pick a subset of the available features at each iteration, therefore exploiting information present in the many features included in this study without having to reduce the feature space using feature selection techniques. Additionally, using random forests we can better deal with class imbalance, since similarly to selecting subset of features at each iteration, we can also select subsets of the majority class at each iteration, therefore being able to train our model on balanced data without discarding relevant information. We did not choose a 1:1 ratio to reduce class imbalance but determined the optimal ratio by cross-validating and optimizing the F-score. Our optimal balance included all data from the minority class (kicks) and one third of the majority class data. Finally, random forests are composed of classification trees, and therefore do not require feature normalization.

C. Performance Metrics and Validation Method

Models were derived and validated using leave one participant out cross-validation and a binary classification problem distinguishing fetal kicks and non-fetal kicks (e.g. non-movement, noise, etc.). Given the binary classification problem and many comparisons, we chose the F-score as a single metric representative of the main outcomes of interest, i.e. sensitivity (the proportion of kicks that are correctly identified as such) and positive predictive value (PPV, proportion of detected kicks that are actually kicks). The F-score was computed as $2 \times \frac{Se \times PPV}{Se + PPV}$.The agreement between accelerometer-detected kicks and manual annotations was not determined on a window by window basis, since kicks and annotations can last different durations and include delays. Thus, performance metrics were determined according to the strategy depicted in Fig. 2. Finally, performance metrics were computed on the entire data stream for all participants during cross-validation, and not only on the subsets of data used for training, in order to provide more realistic results.

V. RESULTS AND DISCUSSION

Results for different sensors number configurations and for the additional reference accelerometer are shown in Fig. 3. We first report results when no reference accelerometer on the back was used. Mean sensitivity and PPV for the case of one single sensor were 0.51 and 0.51, with the exception of sensor 6 (placed on the back) that resulted in sensitivity 0.0, PVV 0.0, highlighting how this placement is optimal to detect maternal movement instead of fetal movement. Mean sensitivity and PPV for the case of two sensors were 0.63 and 0.54 respectively, while for three sensors sensitivity was 0.69 and PPV was 0.57. When using four sensors mean sensitivity was 0.70 and PPV was 0.58. Finally, using all five sensors together provided the same performance (sensitivity 0.70, PPV 0.58). Including a reference accelerometer consistently improved detection performance, as shown in Fig. 3. In particular, when adding the reference accelerometer to a single sensor, we obtained sensitivity 0.57 and PPV 0.56. When adding the reference accelerometer to a two sensors system, we obtained sensitivity 0.68 and PPV 0.61. Results improved marginally when moving to three sensors, with sensitivity 0.70 and specificity 0.63 and more consistently when moving to four sensors, with sensitivity 0.75 and PPV 0.65. Including all sensors did not improve results compared to the four sensors case.

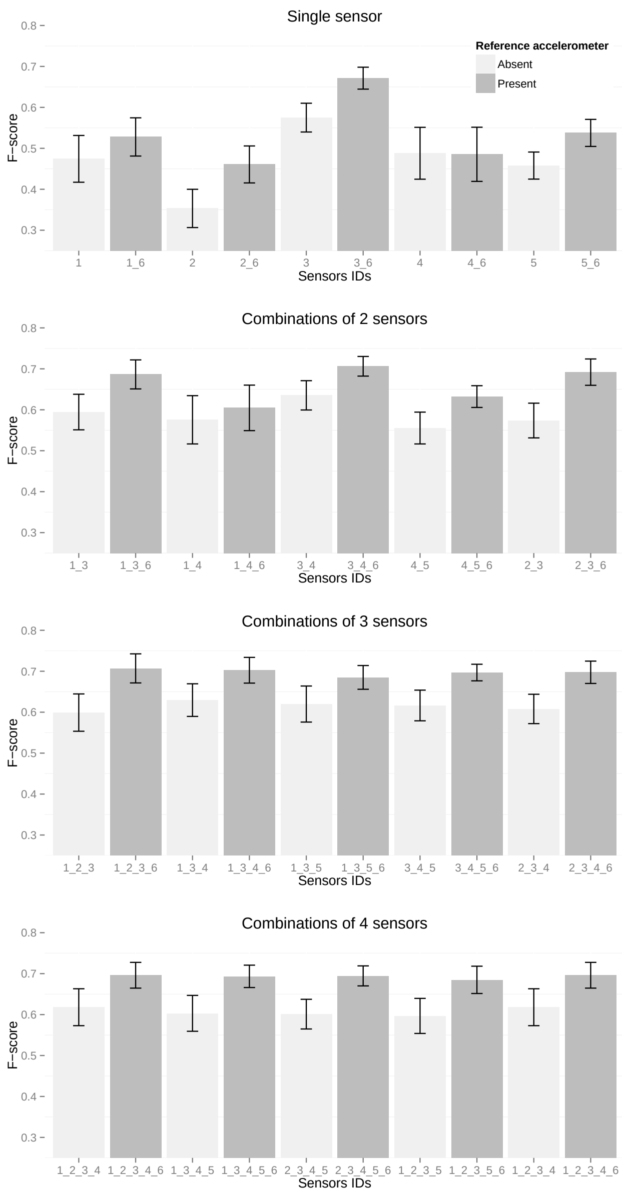

Fig. 3. F-score for different combinations of accelerometer sensors. Results when including a reference accelerometers on the back are shown in darker gray. Only the five best combinations are listed for clarity.

These results, combined with Fig. 3, show how two accelerometers and a reference accelerometer on the back can provide optimal results, when including a sensor above the navel (e.g. sensor 3 in this study). However, for consistent high performance across different locations, three accelerometer in addition to a reference accelerometer seem to be preferable. This setup is similar to the one used in [11].

VI. CONCLUSIONS

In this paper we analyzed trade-offs between sensors number and positioning as well as provided insights on how to handle imbalanced datasets using random forests and how to cross-validate performance of fetal movement detection algorithms to provide more realistic results. Comparing sensor number, positioning and the presence of a reference accelerometer on the same dataset allowed us to make meaningful comparisons and determine performance differences in detecting fetal kicks under different conditions. Future work will investigate the possibility to explore different machine learning tools to add temporal dependency to consecutive time windows as well as include different types of fetal movements and input signals.

REFERENCES

[1] E. Symonds, "On-line processing of the fetal electrocardiogram. a new direction for fetal monitoring." *The Journal of reproductive medicine*, vol. 32, no. 7, pp. 509–512, 1987.

[2] H. P. van Geijn, "2 developments in ctg analysis," *Baillière's clinical obstetrics and gynaecology*, vol. 10, no. 2, pp. 185–209, 1996.

[3] J. F. Pearson and J. B. Weaver, "Fetal activity and fetal wellbeing: an evaluation." *Br Med J*, vol. 1, no. 6021, pp. 1305–1307, 1976.

[4] J. I. De Vries, G. H. Visser, and H. F. Prechtl, "The emergence of fetal behaviour. i. qualitative aspects," *Early human development*, vol. 7, no. 4, pp. 301–322, 1982.

[5] K. Å. Salvesen, "Efsumb: safety tutorial: epidemiology of diagnostic ultrasound exposure during pregnancy?european committee for medical ultrasound safety (ecmus)," *European journal of ultrasound*, vol. 15, no. 3, pp. 165–171, 2002.

[6] E. Sheiner, I. Shoham-Vardi, and J. S. Abramowicz, "What do clinical users know regarding safety of ultrasound during pregnancy?" *Journal of ultrasound in medicine*, vol. 26, no. 3, pp. 319–325, 2007.

[7] Z. Alfirevic, D. Devane, G. Gyte *et al.*, "Continuous cardiotocography (ctg) as a form of electronic fetal monitoring (efm) for fetal assessment during labour," *Cochrane Database Syst Rev*, vol. 3, no. 3, 2006.

[8] R. M. Grivell, Z. Alfirevic, G. Gyte, and D. Devane, "Antenatal cardiotocography for fetal assessment," *Cochrane Database Syst Rev*, vol. 1, 2010.

[9] K. Nishihara, N. Ohki, H. Kamata, E. Ryo, and S. Horiuchi, "Automated software analysis of fetal movement recorded during a pregnant woman?s sleep at home," *PloS one*, vol. 10, no. 6, p. e0130503, 2015.

[10] B. Boashash, M. S. Khlif, T. Ben-Jabeur, C. E. East, and P. B. Colditz, "Passive detection of accelerometer-recorded fetal movements using a time–frequency signal processing approach," *Digital Signal Processing*, vol. 25, pp. 134–155, 2014.

[11] S. Layeghy, G. Azemi, P. Colditz, and B. Boashash, "Non-invasivemonitoring of fetal movements using time-frequency features of accelerometry," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4379–4383.

[12] ——, "Classification of fetal movement accelerometry through time-frequency features," in *Signal Processing and Communication Systems (ICSPCS), 2014 8th International Conference on*. IEEE, 2014, pp. 1–6.

[13] L. Minjie, T. Yongkang, L. Yunfeng, Y. Song, and D. Jingxin, "Fetal movement detection based on mems accelerometer," *une*, vol. 13, p. 15, 2016.

[14] Z. R. Hijazi and C. E. East, "Factors affecting maternal perception of fetal movement," *Obstetrical & gynecological survey*, vol. 64, no. 7, pp. 489–497, 2009.

[15] R. L. Goldenberg, E. M. McClure, Z. A. Bhutta, J. M. Belizán, U. M. Reddy, C. E. Rubens, H. Mabeya, V. Flenady, G. L. Darmstadt *et al.*, "Stillbirths: the vision for 2020," *The Lancet*, vol. 377, no. 9779, pp. 1798–1805, 2011.

[16] G. Thomas, O. T. John, M. Mostefa, B. Boualem, C. Ian, W. Stephen, F. Miguel, C. Susan, and C. Paul, "Detecting fetal movements using non-invasive accelerometers: A preliminary analysis," in *Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on*. IEEE, 2010, pp. 508–511.

[17] M. S. H. Khlif, B. Boashash, S. Layeghy, T. Ben-Jabeur, P. B. Colditz, and C. East, "A passive dsp approach to fetal movement detection for monitoring fetal health," in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*. IEEE, 2012, pp. 71–76.

[18] M. Mesbah, M. Khlif, C. East, J. Smeathers, P. Colditz, and B. Boashash, "Accelerometer-based fetal movement detection," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011, pp. 7877–7880.