Program Digest, 2019 IEEE International Ultrasonics Symposium (IUS)
Glasgow, Scotland, October 6-9, 2019

Prediction of embryo implantation by machine learning based on ultrasound strain imaging

F. Sammali*, C. Blank†*, T.H.G.F. Bakkes*, Y. Huang*, C. Rabotti*, B.C. Schoot†‡, M. Mischi*
*Dept. of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
†Dept. of Reproductive Medicine, Ghent University Hospital, Ghent, Belgium
‡Dept. of Gynaecology, Catharina Hospital, Eindhoven, the Netherlands
Contacts: f.sammali@tue.nl, m.mischi@tue.nl

***Abstract*—Because of the trend to postpone childbirth, the rate of couples dealing with infertility is rapidly increasing and approaching 20%. *In-vitro* fertilization (IVF) represents the only reproduction option in Europe for 2.5 million couples. However, its success rate remains below 30%. There is clear consensus on a major involvement of uterine contractions in IVF failure, especially during and after embryo transfer. Quantitative and non-invasive measurement of uterine (peristaltic) activity, combined with accurate interpretation and classification methods, can provide an important contribution towards improved IVF success rates. Therefore, this study investigates the use of machine learning for probabilistic classification of the uterine activity, as either favorable or adverse to embryo implantation. The results obtained in 16 patients undergoing an IVF cycle confirm the ability to predict successful embryo implantation by ultrasound uterine motion analysis combined with machine learning.**

***Index Terms*—machine learning, in-vitro fertilization, uterine motion, speckle tracking, feature selection, medical ultrasound**

I. Introduction

In developed countries, infertility problems affect about 20% of couples [1]. Due to the trend to postpone childbirth, the prevalence of fertility problems is rapidly increasing. Against this situation, advanced *in-vitro* fertilization (IVF) technologies are developing. It is estimated that in Europe IVF represents the only reproduction option for over 2.5 million couples [2]. In spite of major efforts to improve IVF, its success rate remains below 30% [3]. As a result, many couples undergo repeated IVF treatments with the hope of a successful fertilization. There is increasing evidence of a major involvement of uterine contraction in IVF failure [4]–[7], especially during and after embryo transfer (ET), when the embryo may need quiescence to implant in the uterine cavity.

Unfortunately, the role of uterine contractions in IVF failure is not yet understood and the value of pharmacological treatments modulating the uterine contractility in IVF treatment remains to be established. The lack of quantitative measurement tools has represented an important limiting factor hampering a thorough characterization of the uterine activity outside pregnancy. It is therefore evident that objective and reproducible quantitative of uterine contractions is essential.

In previous studies, we have proposed new methods to quantitatively and non-invasively measure the mechanical activity of human non-pregnant uteri by transvaginal ultrasound (TVUS) motion analysis [8]. A set of amplitude-, frequency-, and energy-related features provided a good understanding of the uterine activity in the natural menstrual cycle and showed the ability to discriminate between the active and quiescent states of the uterus. This promising results motivated us towards the assessment of their clinical value in the context of IVF, especially for improving the prediction of embryo implantation prior to embryo transfer.

In this study, we aim at investigating the use of probabilistic classification of the measured uterine activity by machine learning, with the aim to improve the success rate of IVF, especially in relation to ET. In particular, we aim at determining a set of features such that their combination (amplitude-, frequency-, and energy-related features) by machine learning enables us to classify the uterine contraction characteristics as either favorable or adverse to embryo implantation before ET.

This work was supported by the Dutch NWO-TTW HTSM grant 13901.

II. Materials and Methods

A. Ultrasound data acquisition

Uterine activity was measured by B-mode transvaginal ultrasound (TVUS) in 16 women undergoing IVF treatment. In particular, we focus on the measurements taken one hour before ET (ET1). Patients were divided in two groups, namely ongoing pregnancy (OP) and non-ongoing pregnancy (n-OP), based on follow-up TVUS examination. During standard recording sessions, 4-min TVUS scans were performed by an US scanner WS80A (Samsung Medison, Seoul, South Korea) equipped with a transvaginal V5-9 probe. The scans were acquired at 5.6-MHz central frequency and 30 frames/s (sampling frequency $fs_{(US)}$), amply sufficient to meet Nyquist condition given the limited bandwidth of uterine motion (f $<$ 0.066 Hz) [9].

B. Feature extraction

A set of amplitude-, frequency-, and energy-related features, such as standard deviation (SD), mean frequency (MF) and median frequency (MDF), and unnormalized first moment (UFM), were extracted from the motion signals derived by dedicated US-speckle tracking [8].

For each recording session, the considered features were extracted from the distance and strain signals derived along the longitudinal and transversal directions of the uterus in the subendometrial layer as shown in Fig. (1).

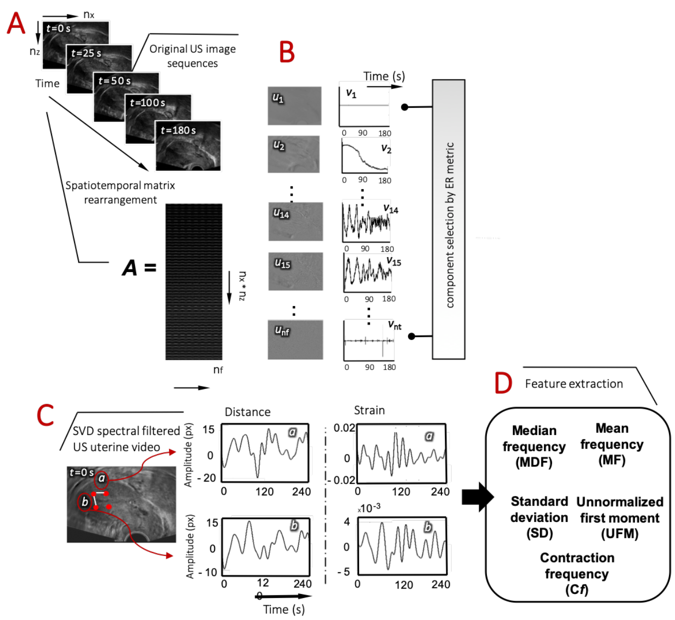

Fig. 1. Feature extraction by TVUS speckle tracking: (A) the unfiltered TVUS video is decomposed in a spatiotemporal matrix (Casorati matrix); (B) SVD filtering is implemented by selecting the desired components according to the ER metric; (C) speckle tracking is then applied to the filtered TVUS loop for the estimation of motion (distance and strain signals) along the (a) longitudinal and (b) transversal direction; (D) Amplitude-, frequency-, and energy-related features extracted from distance and strain signals along the longitudinal (a) and transversal (b) direction.

Motion tracking was implemented by a dedicated speckle tracking algorithm [8] after enhancing uterine motion by dedicated SVD filtering, aimed at suppressing uncorrelated signals that affect speckle tracking performance. Prior to SVD filtering, the speckle size was regularized by implementing a Wiener deconvolution filter according to [10].

SVD decomposition was applied to the input TVUS image sequences. Mathematically, SVD of a data matrix A, in which all spatial locations are located along the rows and the temporal dimension is represented as columns, can be described as

$$A = U \Delta V^{\top}. \tag{1}$$

where **U** and **V** represent the spatial and temporal singular vector, respectively, while $\boldsymbol{\Delta}$ contains the singular values along its diagonal. A proper range of singular values was selected by adopting an energy ratio metric described as:

$$\mathrm{ER}_i = \frac{\sum_{f=f_1}^{f_2} |X_i(f)^2|}{\sum_{f=0}^{f_s/2} |X_i(f)^2|}, \tag{2}$$

where, f_s is the TVUS frame rate ($fs_{(TVUS)}$), $X_i(f)$ represents the Fast-Fourier Transform (FFT) of each singular vector $\vec{V_i}(t)$, with i indicating the specific singular-value number. From each $\vec{V_i}(t)$, two energies were obtained by integration of the power spectrum over the entire frequency band [$f_0 =$ 0 Hz $\;f_{s(US)}/2 =$ 15 Hz] and over the interval [$f_1 =$ 0.0083 Hz $\;f_2 =$ 0.083 Hz], which was chosen to reflect the uterine bandwidth in the IVF cycles according to [5]. Only a range of subsequent singular values representing over 50% of the total energy was selected to enhance the uterine motion in the TVUS loop while suppressing undesired components. By selecting a proper range of singular values according to the ER metric in (2), we separated k = min(n_s, n_t) sources. Through multiplication of the kth column of U with the kth singular value and the kth row of V, individual singular components can be constructed as

$$\mathrm{A}_k = \lambda_k U_k V_k^{\top}(t). \tag{3}$$

A filtered image is eventually created by taking the sum of all desired and consecutive singular components.

After SVD filtering, speckle tracking was implemented by block matching for tracking over time four blocks manually positioned on the subendometrial layer (junctional zone), known to be the most contractile part of the uterus [11]. The match between corresponding blocks in consecutive frames was determined by the minimum of the sum of absolute differences (SAD). The block size for SAD was adaptively determined as twice the speckle size calculated after SVD filtering for each patient. The block matching was then accelerated by a diamond search (DS) strategy, maintaining similar tracking accuracy as the typical full-grid search (FGS) method with the benefit of a reduced computational time. Median filtering was then applied on the displacement of neighboring blocks (matching coordinates) to improve the tracking robustness and accuracy of the method. The defined four blocks were coupled in pairs in order to estimate distance and strain signals along the longitudinal and transversal direction.

All features, with the exception of Cf, were extracted from the FFT of the recorded signals; in particular, the SD was calculated by applying Parsevals theorem. Moreover, the features were extracted in the frequency band [$f_1 =$ 0.0083 Hz $\;f_2 =$ 0.083 Hz] to improve their quality and reinforce their connection with the uterine motion in a stimulated cycle [5]. The same frequency band was used for the ER metrics calculation. A zero-crossing detector was employed in the time domain for the estimation of Cf in terms of number of contractions per minute.

C. Machine learning

Three machine-learning models were implemented to classify successful and unsuccessful embryo implantation using the extracted TVUS features. Features were selected based on classification performance using a correlation based filter [12] and wrapper method based on forward feature selection [13] to determine the optimum feature set leading to the best classification of successful and unsuccessful pregnancies for the three different machine learning methods, i.e., support vector machine (SVM), K-nearest neighbor (KNN), and Gaussian mixture model (GMM). All features were ranked based on

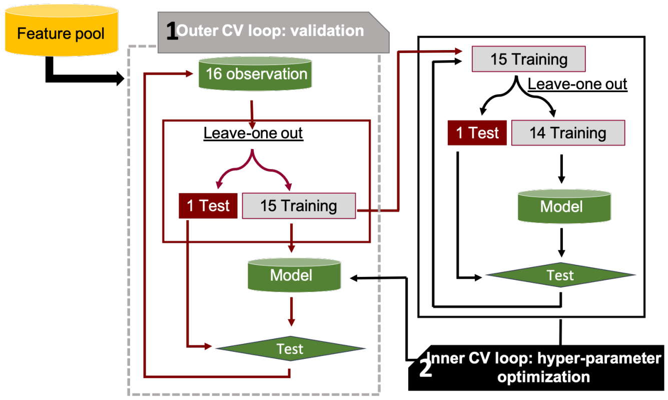

Fig. 2. Overview of the machine-learning framework with nested cross validation (CV). The entire nested cross-validation loop takes place after selecting a feature set by correlation filtering. (1) The outer loop is used for validation of the model selected by the inner cross-validation loop (2). The original dataset (16 observations) is divided in a training set (15 observations) and a test set (1 observation) to implement a leave-one-out validation loop. (2) The inner cross-validation loop is used for tuning the hyperparameters of the classifier. To this end, the training set (15 observations), generated by the outer loop, is further divided in a training set (14 observations) and a test set (1 observation) in a Leave-one-out fashion.

the obtained Pearsons correlation coefficient and the top 10 features were selected and used as the best subset of features. Forward feature selection started by loading one feature at the time. The feature giving the best performance was chosen first; additional features were then added incrementally. At each step, the combination of features providing the largest improvement was chosen. The forward feature selection process stopped when the addition of a new feature did not result in improved classification performance. The proposed classifiers were then tested and trained in a nested cross-validation loop, in the interest of reducing over-fitting.

In particular, as shown in Fig. (2), our adopted strategy consisted of first performing forward-feature selection, then tuning the hyperparameters of the classifiers in an inner cross-validation loop and, finally, validating the output model on new data in an outer cross-validation loop.

A full grid search was adopted for hyperparameter optimization. Validation of the classifiers was performed in a leave-one-out fashion. Accuracy (Acc), sensitivity (Se), and specificity (Sp) were used as performance metrics.

III. Results

A. Classification results

Table I reports the performance of different feature combinations in terms of accuracy, sensitivity, and specificity using the three adopted classifiers. The iteration number indicates the number of selected features along the forward selection process. The best performance of each machine-learning method using its optimum feature set is highlighted in light gray color. A smaller feature set is preferred to reduce the model complexity and the risk of overfitting.

TABLE I
PERFORMANCE OF DIFFERENT FEATURE COMBINATIONS USING THE THREE ADOPTED CLASSIFIERS ONE HOUR PRIOR TO EMBRYO TRANSFER (ET1). (DL) AND (DT) SUBSCRIPTS: FEATURE EXTRACTED FROM THE DISTANCE SIGNAL ALONG THE LONGITUDINAL AND TRANSVERSAL DIRECTION, RESPECTIVELY; (SL) AND (ST) SUBSCRIPTS: FEATURE EXTRACTED FROM THE STRAIN SIGNAL ALONG THE LONGITUDINAL AND TRANSVERSAL DIRECTION, RESPECTIVELY.

Classifier: Support vector machine (SVM)				
Iteration	**Feature**	**Acc (%)**	**Se (%)**	**Sp (%)**
1	Cf_{DL}	81.3%	57.1%	100%
2	Cf_{DT}	87.5%	71.4%	100%
3	MDF_{SL}	81.3%	71.4%	88.9%
4	Cf_{ST}	75.0%	71.4%	77.8%
5	MF_{DT}	81.3%	71.4%	88.9%
6	SD_{DL}	68.8%	57.1%	77.8%
Classifier: K-Nearest Neighbours (KNN))				
Iteration	**Feature**	**Acc (%)**	**Se (%)**	**Sp (%)**
1	Cf_{DL}	87.5%	71.4%	100%
2	Cf_{DT}	87.5%	71.4%	100%
3	Cf_{ST}	87.5%	71.4%	100%
4	MF_{ST}	93.8%	85.7%	100%
5	MF_{DT}	81.3%	71.4%	88.9%
Classifier: Gaussian mixture model (GMM)				
Iteration	**Feature**	**Acc (%)**	**Se (%)**	**Sp (%)**
1	Cf_{DL}	81.3%	71.4%	88.9%
2	MF_{DT}	81.3%	71.4%	88.9%
3	SD_{DL}	81.3%	71.4%	88.9%
4	Cf_{DT}	68.8%	57.1%	77.8%

IV. Discussion and Conclusions

In this study, we assessed the clinical value of uterine motion analysis by speckle tracking for the prediction of IVF success (ongoing pregnancy) by machine learning. To this end, we explored the classification of uterine motion characteristics using frequency-, amplitude-, and energy-related features extracted form TVUS motion signals.

A crucial task for pregnancy prediction consists of identifying effective feature combinations able to distinguish between patients with ongoing and those with non-ongoing pregnancy. This task was performed by a machine-learning model (classifier), and forward-feature selection was employed to find the optimum feature set. Three classifiers were employed and compared on the same database of 16 patients undergoing IVF treatment. Optimization of the classifier hyperparameters was performed in a nested cross-validation loop.

When an individual feature was used, Cf_{DL} (contraction frequency extracted from the distance along the longitudinal direction by TVUS speckle tracking) provided the best accuracy performance on KNN (87.5%) followed by the SVM and GMM (81.3%).

Further in the selection procedure, different features were added into each optimum feature set leading to improved classification performance. The three classifiers selected different optimum feature sets. This suggests the optimum feature set to depend on the adopted classifier; hence, in order to achieve the best classification performance, the feature selection procedure should be carried out separately for each machine-learning model.

The highest accuracy, 93.8%, was achieved by KNN with sensitivity and specificity of 85.7% and 100%, respectively. The lower performance of the GMM classifier, unable to improve the classification over individual features, may be due to the feature distribution, which cannot be well represented by the combination of Gaussian distributions. In general, frequency-related features result to be powerful predictors for embryo implantation. In the future, a larger dataset should be realized in order to improve the accuracy, robustness, and generalizability of the classifiers.

REFERENCES

[1] G. F. Whitman-Elia and E. G. Baxley, "A primary care approach to the infertile couple." *The Journal of the American Board of Family Practice*, vol. 14, no. 1, pp. 33–45, 2001.

[2] E. C. W. Group *et al.*, "Social determinants of human reproduction," *Human Reproduction*, vol. 16, no. 7, pp. 1518–1526, 2001.

[3] A. Nyboe Andersen, V. Goossens, A. Ferraretti, S. Bhattacharya, R. Felberbaum, J. de Mouzon, K. Nygren, f. t. E. S. o. H. R. The European IVF-monitoring (EIM) Consortium, and E. (ESHRE), "Assisted reproductive technology in europe, 2004: results generated from european registers by eshre," *Human Reproduction*, vol. 23, no. 4, pp. 756–771, 2008. [Online]. Available: http://dx.doi.org/10.1093/humrep/den014

[12] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proceedings, Twentieth International Conference on Machine Learning*, T. Fawcett and N. Mishra, Eds., vol. 2, 2003, pp. 856–863.

[4] R. Fanchin and J. M. Ayoubi, "Uterine dynamics: impact on the human reproduction process," *Reproductive BioMedicine Online*, vol. 18, pp. S57 – S62, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1472648310604506

[5] R. Fanchin, C. Righini, F. Olivennes, S. Taylor, D. de Ziegler, and R. Frydman, "Uterine contractions at the time of embryo transfer alter pregnancy rates after in-vitro fertilization," vol. 13, pp. 1968–74, 08 1998.

[6] P. Lesny, S. Killick, J. Robinson, G. Raven, and S. Maguiness, "Junctional zone contractions and embryo transfer: is it safe to use a tenaculum?" *Human reproduction*, vol. 14, no. 9, pp. 2367–2370, 1999.

[7] M. M. IJland, H. J. Hoogland, G. A. Dunselman, C. R. Lo, and J. L. Evers, "Endometrial wave direction switch and the outcome of in vitro fertilization," *Fertility and sterility*, vol. 71, no. 3, pp. 476–481, 1999.

[8] F. Sammali, N. P. Kuijsters, Y. Huang, C. Blank, C. Rabotti, B. C. Schoot, and M. Mischi, "Dedicated ultrasound speckle tracking for quantitative analysis of uterine motion outside pregnancy," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 66, no. 3, pp. 581–590, 2019.

[9] C. Bulletti, D. de Ziegler, V. Polli, L. Diotallevi, E. D. Ferro, and C. Flamigni, "Uterine contractility during the menstrual cycle," *Human reproduction*, vol. 15, no. suppl_1, pp. 81–89, 2000.

[10] M. Kuenen, T. Saidov, H. Wijkstra, and M. Mischi, "Contrast-ultrasound dispersion imaging for prostate cancer localization by improved spatiotemporal similarity analysis," *Ultrasound in medicine & biology*, vol. 39, no. 9, pp. 1631–1641, 2013.

[11] M. M. Ijland and J. L. H. Evers, "Subendometrial contractions in the nonpregnant uterus: an ultrasound study," *European Journal of Obstetrics & Gynecology and reproductive Biology*, vol. 70, no. 1, pp. 23–24, 1996.

[13] A. Marcano-Cedeño, J. Quintanilla-Domínguez, M. Cortina-Januchs, and D. Andina, "Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network," in *IECON 2010-36th annual conference on IEEE industrial electronics society*. IEEE, 2010, pp. 2845–2850.