Features Fusion Using Belief Functions Theory for ARDS Prediction

Aline Taoum[1,2], Farah Mourad-Chehade[1], and Hassan Amoud[2]
[1]Laboratoire de Modélisation et Sûreté des Systèmes, Institut Charles Delaunay, Université de Technologie de Troyes, Troyes, France
[2]Azm Platform for Research in Biotechnology and its Applications, EDST, Lebanese University, Tripoli, Lebanon
Email: {aline.taoum, farah.chehade}@utt.fr, hassan.amoud@ul.edu.lb

***Abstract*—Information fusion techniques are at high interest with the increase in dimensionality of the data being handled. They are applied in different applications, such as in the biomedical domain. This paper proposes an information fusion model that predicts the occurrence of ARDS using vital signs. This model uses features fusion based on the belief functions theory. Different linear and nonlinear parameters are first extracted from the signals, and a parameters selection procedure is proposed to select only pertinent parameters. These parameters are then used to construct mass functions in the belief functions framework. Afterwards, the prediction is performed in real-time by combining all the constructed mass functions. Results present the effectiveness of the belief theory predicting ARDS using the MIMIC II public database.**

***Index Terms*—acute respiratory distress syndrome, belief functions theory, features fusion, linear and non-linear parameters**

I. Introduction

Interest in information fusion has rapidly increased in the last decade, especially for biomedical applications. One of these applications is the monitoring of the health state of patients based on the recorded physiological data [1]-[3]. The data fusion can be performed on different levels, that are data, decision and feature level fusion [4]. Feature fusion has demonstrated its effectiveness and advantage over data and decision fusion techniques. It consists of extracting hidden features that characterize each source data and then a single model that combines these features and classifies them is determined.

Research does not cover all the serious pathologies that can affect the health and the autonomy of patients. One of the neglected health condition is the Acute Respiratory Distress Syndrome (ARDS) [5], [6]. ARDS is a fatal lung condition characterized by insufficiency of gas exchange with the blood and may lead to death [7]. It is diagnosed using the ratio of the partial pressure of oxygen PaO_2 to the fraction of inspired air FiO_2. A severe or moderate ARDS is diagnosed if PaO_2/FiO_2<200 mmHg [8]. Having a slow recovery, ARDS patients often develop neuromuscular weakness and neuropsychiatric problems that delay the ability to return to normal life routine by months or permanently in some cases [9], [10]. This paper aims to predict the occurrence of ARDS using a feature fusion model.

Several feature extraction methods exist in literature to extract parameters from raw data. Parameters can be linear or non-linear in both temporal and frequency domains [11]. Even though features are more informative indicators than raw data, noisy and overlapping set of features can distort the performance of a good classifier. Hence, the selection of a proper set of features becomes an essential step in feature fusion before performing the feature combination [12], [13]. Among all the techniques used to address this fusion problem, the theory of belief functions presented by Shafer [14] has attracted significant attention due to its effectiveness in combining information and dealing with imperfect evidence.

The use of the belief functions theory for the problem of information fusion was first presented in 1981 [15] and then applied in [16]-[19]. In contrast to the Bayesian theory, the belief functions theory allows each source to contribute information in different levels of detail. A priori probabilities are only assigned when an information is provided. In fact, it allows an explicit representation of total ignorance by assigning the entire mass to the frame of discernment. In addition, a new evidence is attributed to a source observation using conditioning rules. Then, the evidences provided from different sources are combined using one of the combination rules from [20]. The conditioning rules, also known as discounting, follow from those described by Mercier in [21].

The main objective of this paper is to develop a surveillance model that relies on features extraction and belief functions theory to predict the occurrence of ARDS in real-time. This is done by analyzing physiological signals from the subjects. Different linear and nonlinear parameters are extracted from the signals. Then, a selection procedure is handled in order to select only pertinent parameters and thus enhance the performance. Finally, the parameters are transformed to a measure of evidence, known as mass functions that represent the state of beliefs about a given state. These mass functions are discounted and combined by the belief functions theory.

This paper is organized as follows: Section 2 first presents the parameters extraction procedure, and then describes the belief functions model. Section III

Manuscript received May 13, 2019; revised November 6, 2019.

doi: 10.18178/ijsps.7.4.107-112

illustrates the obtained results that are interpreted in Section IV. Finally, Section V presents a summary of the paper and proposes some perspectives.

II. METHODS

The model proposed in this paper is a real-time approach that monitors the ongoing health state of patients using their physiological signals. Let us consider a subject s having I recorded physiological signals. Then, $x_{s,i}(l)$ denote these signals recorded at the time l, with $i \in I$ and a segment of these signals going from time a to time b is defined by $x_{s,i}(a{:}\,b)$. The proposed model takes as input the recorded signals in real-time t, $x_{s,i}(1{:}\,t)$, and gives as output a decision whether the subject is going to develop ARDS. Thus, two classes of subjects are defined as follows,

$$\begin{cases} +1, \text{ ARDS, including subjects who will develop ARDS,} \\ -1, \text{"non} - \text{ARDS", including subjects not developing ARDS.} \end{cases}$$

In the following, a detailed description of the parameters extraction is provided. Then, the belief functions theory is described. Finally, the parameters selection procedure is presented as well as the optimization of the model's characteristics.

A. Parameters Extraction

As mentioned previously, different linear and non-linear parameters are extracted from the segments $x_{s,i}(t-\tau+1{:}\,t)$ in real-time t. The linear parameters are the mean (μ), the standard deviation (σ), the skewness (Sk) and the kurtosis (Kt). The nonlinear parameters include the sample entropy ($SampEn$) and both factors from the detrended fluctuation analysis (DFA_1 and DFA_2). These parameters are described and extracted from the signals in our previous work in [11]. Furthermore, we propose a new extraction of nonlinear parameters from the recurrence quantification analysis (RQA).

The RQA studies the recurrences of a dynamical system presented by its phase space trajectory. The phase space is a multi-dimensional space that illustrates all possible states of a system [22]. It is constructed by time-lagged copies of the initial signal using a time delay t_d and an embedding dimension M. Generally, t_d and M are estimated using the average mutual information [23] and the algorithm of the false nearest neighbors [24], respectively. Then, the M-dimensional phase space is projected into a 2-dimensional illustration, known as the recurrence plot (RP) [25]. This latter is shown by a square matrix with elements corresponding to times at which a state is repeated. Different parameters can be extracted from the recurrence plot, such as the recurrence rate (R_eR_a) which is the percentage of recurrence points in an RP and the entropy ($ENTR$) of the probability distribution of its diagonal line lengths.

Since abnormalities are better detected within short segments, we propose in this paper to consider a fixed window of length τ taken at the end of the segments $x_{s,i}(t-\tau+1{:}\,t)$. An optimization procedure is performed in the following to find the optimal value of τ. Let $p_{s,j}(t), j \in J$, be the set of parameters extracted at time t from all the segments $x_{s,i}(t-\tau+1{:}\,t), i \in I$, of a subject s. When dealing with people's health, it is evident that normal is a relative state and each patient interacts differently for a new situation. Thus, a normalization of the extracted parameters is proposed in this work. It consists of computing the ratio of the extracted parameter in real-time $p_{s,j}(t)$ to the parameter of an initial segment assumed to be normal $p_{s,j}^{(0)}$. In the following $p_{s,j}$ will denote the normalized parameters.

B. Belief Functions Theory

In this section, we review the basic concepts of the belief functions theory and present related functions. The theory of belief functions has been developed by Dempster [20] and Shafer [14]. It is often interpreted as an extension of the Bayesian theory of probabilities. The belief functions theory assigns an evidence for each of the subsets of the total set of states, rather than for each of the individual states as the Bayesian theory does. Moreover, the functions of the belief theory allow one to attribute a confidence measure to an event being observed. This is represented by discounting rules performed on the provided information. In the following, the basic concept of the theory of belief functions is defined, as well as the specific notations used in this paper.

1) Main concept

Given the previously described problem, the belief functions theory operates on a frame of discernment Ω that consists of the possible states of subjects, $\Omega = \{+1, -1\}$. The set of all possible subsets of Ω is defined by 2^Ω,

$$2^\Omega = \{\emptyset,\ \{+1\},\ \{-1\},\ \{+1,-1\}\}$$

The $\emptyset$ represents the impossibility, where the state of the subject is neither "+1" nor "-1" and $\{+1,-1\}$ represents the ambiguity, that is the states of "+1" and "-1" are so close and the decision is difficult to be generated.

The belief functions rely on the modeling of the evidence provided from each extracted parameter. This modeling is represented by a mass function for each parameter j, denoted by $m_j(\cdot)$ [26]. Therefore, the mass function $m_j(\omega, s, t)$ is the measure of the belief attributed to the subset $\omega \in 2^\Omega$ for a subject s at time t. The mass functions satisfy the following conditions

$$\begin{cases} \mathrm{m_j}(\omega, \mathrm{s}, \mathrm{t}) \to [0,1], \text{for } \forall\, \omega \in 2^\Omega, \\ \mathrm{m_j}(\emptyset, \mathrm{s}, \mathrm{t}) = 0, \\ \sum_{\omega \in 2^\Omega} \mathrm{m_j}(\omega, \mathrm{s}, \mathrm{t}) = 1. \end{cases} \quad (1)$$

In order to create the mass functions, we propose to use the distributions of parameter values for each class. To do this, we consider a training set of subjects; then we select the last segments of the signals of length τ for an ARDS subject to extract the parameters and segments from the beginning of the signals for non-ARDS subjects. This is due to the fact that the instability of ARDS subjects is included in the last part of their signals

preceding the occurrence of ARDS; whereas, the stability is best guaranteed in the beginning of the recordings for non-ARDS subjects. Then, for each parameter j, probability distribution functions are estimated for ARDS parameter values only to represent $\{+1\}$, non-ARDS parameter values for $\{-1\}$ and both sets to define the total subset $\{+1,-1\}$ distribution. We add the latter set of classes to allow the modeling of ambiguity. The distribution functions are denoted by $Q_{j,\omega}(\cdot)$ for parameter j according to non-empty subset $\omega \in 2^{\Omega}$. Then, the mass functions for a new given parameter $p_{s,j}(t)$ at the real-time t is written as follows:

$$\mathrm{m_j}(\omega,\mathrm{s},\mathrm{t}) = \frac{\mathrm{Q_{j,\omega}}\left(\mathrm{p_{s,j}(t)}\right)}{\sum_{\omega' \in 2\Omega} \mathrm{Q_{j,\omega'}}\left(\mathrm{p_{s,j}(t)}\right)}, \omega \in 2^{\Omega}, \omega \neq \emptyset \quad (2)$$

2) Discounting

The discounting operation considers the reliability of the information provided from each source [27]. It is necessary when working with real data, especially in the case of biomedical signals. This rule transforms the mass functions into less informative ones according to the degree of reliability of each parameter. In this work, the contextual discounting is performed because it considers the reliability of every parameter according to each state [21], [28]. Let $\alpha_{j,\{+1\}}$ and $\alpha_{j,\{-1\}}$ be the discounting rates of a parameter j according to the states "+1" and "-1", respectively. Then, the discounted mass functions ${}^{\alpha}m_j(\cdot)$ becomes

$$\begin{cases} {}^{\alpha}\mathrm{m_j}(\{+1\},\mathrm{s},\mathrm{t}) = (1-\alpha_{\mathrm{j},\{-1\}})\,\mathrm{m_j}(\{+1\},\mathrm{s},\mathrm{t}), \\ {}^{\alpha}\mathrm{m_j}(\{-1\},\mathrm{s},\mathrm{t}) = (1-\alpha_{\mathrm{j},\{+1\}})\,\mathrm{m_j}(\{-1\},\mathrm{s},\mathrm{t}), \\ {}^{\alpha}\mathrm{m_j}(\Omega,\mathrm{s},\mathrm{t}) = \mathrm{m_j}(\Omega,\mathrm{s},\mathrm{t}) + \alpha_{\mathrm{j},\{-1\}}\mathrm{m_j}(\{+1\},\mathrm{s},\mathrm{t}) \\ \qquad\qquad + \alpha_{\mathrm{j},\{+1\}}\mathrm{m_j}(\{-1\},\mathrm{s},\mathrm{t}), \\ {}^{\alpha}\mathrm{m_j}(\emptyset,\mathrm{s},\mathrm{t}) = 0 \end{cases} \quad (3)$$

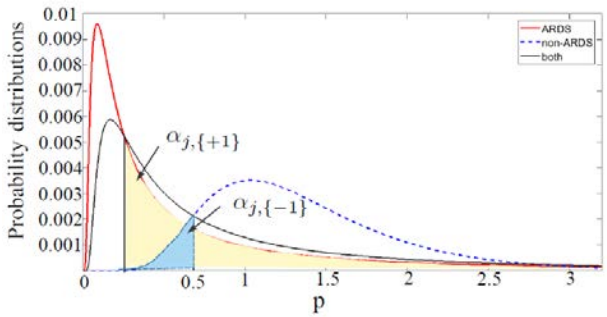

Figure 1. An example for computing the discounting rates.

The discounting rates $\alpha_{j,\omega}$, $\omega = \{+1\}$ or $\omega = \{-1\}$, are estimated in this paper from the distribution functions $Q_{j,\omega}(\cdot)$ as follows

$$\alpha_{\mathrm{j},\omega} = \int_{\mathrm{D_{j,\omega}}} \mathrm{Q_{j,\omega}(p)dp} \quad (4)$$

where $D_{j,\omega} = \{p \,|Q_{j,\omega}(p) < Q_{j,\omega'}(p), \forall\, \omega' \in 2^{\Omega}, \omega' \neq \omega\}$. In other words, $\alpha_{j,\omega}$ is computed for all parameter values where a $Q_{j,\omega'}(p)$ for any subset $\omega' \neq \omega$ is higher than the actual distribution $Q_{j,\omega}(p)$. Fig. 1 illustrates an example of the estimated probability distribution for ARDS subjects in red straight line, the one for non-ARDS subjects in blue dashed line and the one of all subjects in black straight line. It also presents the computation of the error rates for the states of $\{+1\}$ in yellow and of $\{-1\}$ in blue.

3) Combination

The mass functions can be combined to yield a new mass function $m(\cdot)$, by a combination rule. The classical one is the Dempster's rule of combination [20], also named *orthogonal sum*, noted by $m_{\oplus}(\cdot)$ and defined by:

$$\mathrm{m_{\oplus}}(\omega,\mathrm{s},\mathrm{t}) = \frac{\sum_{\cap\omega'^{(\mathrm{j})}=\omega} \prod_{\mathrm{j}\in\mathrm{J}} {}^{\alpha}\mathrm{m_j}(\omega'^{(\mathrm{j})},\mathrm{s},\mathrm{t})}{1-\mathrm{K}} \quad (5)$$

with $K = \sum_{\cap\omega'^{(j)}=\emptyset} \prod_{j\in J} {}^{\alpha}\mathrm{m_j}(\omega'^{(j)},\mathrm{s},\mathrm{t})$.

K is a normalization term that measures the degree of conflict between the mass functions. $K = 1$ corresponds to total contradiction between mass functions; whereas $K = 0$ implies the absence of conflict. This combination rule leads a more informative mass intersecting all modeled information.

4) Decision making

Finally, a decision is made on the health state of a subject s using the pignistic transformation defined by:

$$BetP(\omega,s,t) = \sum_{\omega' \in 2^{\Omega}, \omega \subseteq \omega'} \frac{m_{\oplus}(\omega',s,t)}{|\omega'|}, \forall \omega \in \Omega \quad (6)$$

In this transformation, the masses of the subsets are injected in the masses of singletons. Thus, the mass of $\{+1,-1\}$ subset is divided between $\{+1\}$ and $\{-1\}$. Then, the state having the highest mass is selected at time t.

Since the analysis is performed in real-time, an alert can be generated from the first positive decision or after a succession of positive decisions. In the latter case, a threshold must be defined as being the needed number of successive positive decisions to generate ARDS alert. Therefore, an optimization of this threshold is proposed in this paper to find the optimal successions of positive decisions. Different values of possible successive positive decisions are considered and the performance indexes are computed from a training dataset. Then, the threshold that gives the higher performance is selected.

The performance indexes are the sensitivity (Se) and the specificity (Sp), also known as the true positive rate and the true negative rate, respectively. They are computed as follows

$$Se = \frac{\text{Number of correctly identified ARDS subjects}}{\text{Total number of ARDS subjects}}$$

$$Sp = \frac{\text{Number of correctly identified non-ARDS subjects}}{\text{Total number of non-ARDS subjects}}$$

C. Parameters Selection

The high dimensionality of the input data to the model increases the time of computations and may decrease the performance of the model. Hence, the selection of the proper set of features is an essential step in feature fusion models. In this work, a ranking of the parameters is proposed based on the discounting rates, also called error

rates, estimated from the distributions functions and then a sequential forward selection is performed [29].

Hence, all the parameters are ranked by computing the mean of the conditional error rates $\alpha_j = (\alpha_{j,\{+1\}} + \alpha_{j,\{-1\}})/2$. Then, the selection procedure starts from the top ranked parameter, thus having the lowest error rate, and adds sequentially one parameter at a time. This procedure stops whenever an added parameter shows a decrease in the performance.

III. Results

A. Materials

In order to validate the proposed model, the multi-parameter intelligent monitoring of intensive care II (MIMIC II) database is considered [30], [31]. It is a publicly available database collected over a seven years period in intensive care units. It contains mainly two types of data sets: the monitor waveforms and the clinical data. First, the clinical database was considered to select ARDS and non-ARDS subjects. ARDS subjects start with a ratio of PaO_2/FiO_2>200 mmHg then this ratio decreases to less than 200 for at least 12 hours; whereas non-ARDS subjects have PaO_2/FiO_2>200 mmHg over the recordings length. Then, the selected subjects are matched to the waveform database, and their signals are selected. Four time series are considered for each subject, that are the heart rate (HR), the respiratory rate (RR), the arterial oxygen saturation (SpO_2) and the mean airway blood pressure (MABP). These time series have a sampling frequency of one sample per minute.

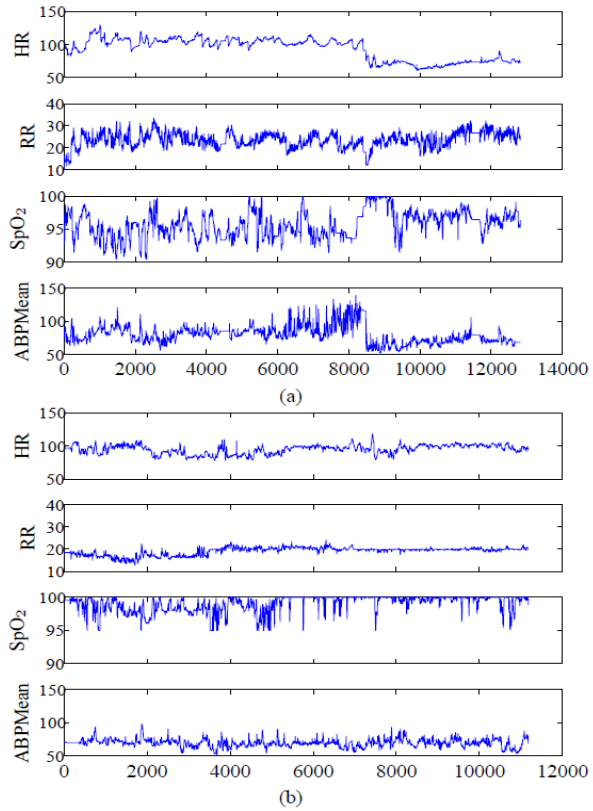

Figure 2. An example of the four signals for (a) an ARDS subjects and (b) a non-ARDS subjects.

Among the selected ARDS subjects, all the subjects that started their time series records after ARDS diagnosis or ended their records before ARDS diagnosis are excluded from the study. This leads to 50 ARDS subjects and 135 non-ARDS subjects. In order to obtain a more accurate model, the number of ARDS and non-ARDS subjects has to be equivalent. Thus, the non-ARDS group is reduced leading to 50 non-ARDS subjects. Fig. 2 illustrates an example of the extracted time series for an ARDS subject in (a) and a non-ARDS subject in (b), where the x-axis presents the time in minutes. The validation of the model is performed using a 5-fold cross validation repeated 10 times.

B. Statistical Analysis

In this section, a statistical analysis is performed to identify parameters that show significant difference between ARDS and non-ARDS groups. For each type of parameters for each signal, values were compared between both groups using the two-sample F-test from Matlab 2017. A parameter is considered significantly different between the two groups if the p-value<0.05. Table I presents the significant parameters for each signal type.

TABLE I. The Most Significant Parameters for Each Type of Signals

Signal	Parameters
HR	$\mu, Sk, Kt, SampEn, DFA_1, R_eR_a, ENTR$
RR	$\sigma, Sk, Kt, SampEn, DFA_1, R_eR_a, ENTR$
SpO_2	$\sigma, Sk, Kt, SampEn, DFA_1, DFA_2\ R_eR_a$
MABP	$Sk, DFA_1, DFA_2\ R_eR_a$

TABLE II. Performance of the Belief Model Over Training and Test Sets

Model's phase	Training set		Test set	
	Se (%)	Sp (%)	Se (%)	Sp (%)
BF without discounting, nor selection	67.5	65	58	60
BF without discounting	65	75	47.5	72.5
Complete BF model	68	75.5	62	66

C. Performance of the Model

In this section, the complete belief functions model is tested on the selected subjects by performing 5-fold cross validations. Then, a comparison is made between the different phases of the proposed model, such as the selection of parameters and the discounting. Table II presents the results of each of the model's phases over both training and testing sets. When considering all the parameters for the construction and the combination of mass functions, sensitivities of 67.5% and 58% and specificities of 65% and 60% are obtained respectively over the training and test sets. The parameter selection phase has enhanced the specificity but decreased the sensitivity with (Se = 65%, Sp = 75%) over the training set and (Se = 47.5%, Sp = 72.5%) over the test set. finally,

the discounting of mass functions lead to a sensitivity of 68% and a specificity of 75.5% for the training set and Se = 62%, Sp = 66% for the test set.

IV. Discussion

This paper presents a model for the prediction in real-time of ARDS from time series data. The proposed model is based on the extraction of parameters and the belief functions theory. The time series data are the heart rate, the respiratory rate, the peripheral arterial oxygen saturation and the mean airway blood pressure. Different linear and non-linear parameters are extracted, such as mean, standard deviation, skewness, kurtosis, sample entropy, both factors of the detrended fluctuation analysis, the recurrence rate and the entropy from the recurrence quantification analysis. The MIMIC II database was considered to validate the proposed model.

Identifying a pathology can be done using just one signal [32], but most of the serious diseases occur with concurrent abnormalities in multiple physiological signals [1]. Despite the intense research on ARDS, there remains a lack of the characterization of this condition using features. ARDS is associated in its development with many cardiovascular and pulmonary complications, such as abnormal tension or blood pressure, abnormal heart beat, deficiency in oxygen delivery and abnormal respiration [33]-[35]. Therefore, the extracted signals are the heart rate that represent the heartbeat, the respiratory rate that reflects any abnormalities in respiration, the peripheral arterial oxygen saturation that measures the oxygen levels in the blood and the mean airway blood pressure that reflects the tension.

Hence, in this paper, different linear and nonlinear parameters are extracted from the time series data. Each of these parameters reflects the presented characteristics in the signals. Linear parameters provide information on the general distributions of the data. For instance, the standard deviation reflects the cyclic components that are responsible on the variability in the segments [36]. The skewness reflects the acceleration and deceleration capacity of the time series according to the sign of the skewness [37]. The kurtosis measures the concentration of the data around the mean [13]. However, nonlinear parameters measure hidden features in the segments. The sample entropy measures the amount of complexity in the data. The detrended fluctuation analysis measures the roughness of signals [38]. The RQA parameters measures the similarity in time in the signals. As shown in the statistical test, there exist different linear and nonlinear parameters that show significant difference between both groups.

From the obtained results, the reduction of the number of parameters, by performing a selection procedure, improved the specificity of the model. This way only parameters having low error rates are included in the model and the combination that presents the best local accuracy is considered. Then, the inclusion of the conditional reliability of each source has led to an enhancement in the accuracy of the model over the training and the test sets.

V. Conclusion

This paper proposes a model based on belief functions theory to predict ARDS in real-time. Noninvasive vital signs are included in the study for the reason of the facility of their acquisition and the link between these signals and the risk factors associated with ARDS. Linear and non-linear parameters are extracted since they can provide information about the properties of a signal more than the signal itself. The belief functions model considers imprecision and unreliability of information sources. It assigns masses to each subset according to each parameter. Then, these masses are discounted and combined according to a measure of confidence of each parameter. This model is then extended to reduce the dimension of the input by performing a parameter selection procedure. This model has achieved high performances in both ARDS and non-ARDS groups.

Further work must be done to extract more parameters from the time series data. Moreover, it would be interesting to establish the relationships between the changes in time series and the cardiovascular mechanism. A further study on the belief functions model will also be done, to propose a new learning algorithm for the construction of mass functions.

Acknowledgment

The authors would like to thank Université de Technologie de Troyes, France and the Lebanese University, Lebanon for their funding of this research.

References

[1] A. R. M. Forkan and I. Khalil, “A clinical decision-making mechanism for context-aware and patient-specific remote monitoring systems using the correlations of multiple vital signs,” *Comput. Methods Programs Biomed*., vol. 139, pp. 1-16, 2017.

[2] A. Lorenz and R. Oppermann, “Mobile health monitoring for the elderly: Designing for diversity,” *Pervasive Mob. Comput.*, vol. 5, no. 5, pp. 478-495, October 2009.

[3] K. C. Tseng, B. S. Lin, L. D. Liao, Y. T. Wang, and Y. L. Wang, “Development of a wearable mobile electrocardiogram monitoring system by using novel dry foam electrodes,” *IEEE Syst. J.*, vol. 8, no. 3, pp. 900-906, September 2014.

[4] B. V. Dasarathy, *Decision Fusion*, Los Alamitos, CA: IEEE Computer Society Press, 1994.

[5] L. Ware and M. Matthay, “The acute respiratory distress syndrome,” *N. Engl. J. Med.*, vol. 342, no. 18, pp. 1334-1349, May 2000.

[6] K. P. Steinberg, *et al.*, “Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome,” *N. Engl. J. Med.*, vol. 354, no. 16, pp. 1671-1684, April 2006.

[7] D. G. Ashbaugh, D. B. Bigelow, T. L. Petty, and B. E. Levine, “Acute respiratory distress in adults,” *Lancet*, vol. 290, no. 7511, pp. 319-323, August 1967.

[8] V. M. Ranieri, *et al.*, “Acute respiratory distress syndrome: The Berlin definition,” *JAMA.*, vol. 307, no. 23, pp. 2526-2533, 2012.

[9] M. S. Herridge, C. M. Tansey, A. Matté, G. Tomlinson, N. Diaz-Granados, A. Cooper, C. B. Guest, C. D. Mazer, S. Mehta, T. E. Stewart, and P. Kudlow, “Functional disability 5 years after acute respiratory distress syndrome,” *N. Engl. J. Med.*, vol. 364, no. 14, pp. 1293-1304, April 2011.

[10] E. Fan, *et al.*, “Physical complications in acute lung injury survivors: A 2-year longitudinal prospective study,” *Crit. Care Med.*, vol. 42, no. 4, p. 849, April 2014.

[11] A. Taoum, M. Sabbagh, H. Amoud, and F. Mourad-Chehade, “Features extraction from vital signs to characterize acute

respiratory distress syndrome," in *Proc. IEEE 4th Middle East Conference on Biomedical Engineering*, 2018, pp. 62-66.
[12] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491-502, April 2005.
[13] A. L. Rivera, *et al.*, "Heart rate and systolic blood pressure variability in the time domain in patients with recent and long-standing diabetes mellitus," *PLoS One*, vol. 11, no. 2, p. e0148378, February 2016.
[14] G. Shafer, *A Mathematical Theory of Evidence*, Princeton: Princeton University Press, 1976.
[15] T. D. Garvey, J. D. Lowrance, and M. A. Fischler, "An inference technique for integrating knowledge from disparate sources," in *Proc. 7th International Joint Conference on Artificial Intelligence*, 1981, vol. 1, pp. 319-325.
[16] H. Wu, M. Siegel, R. Stiefelhagen, and J. Yang, "Sensor fusion using Dempster-Shafer theory," in *Proc. 19th IEEE Instrumentation and Measurement Technology Conference*, 2002, pp. 7-11.
[17] H. Wu, M. Siegel, and S. Ablay, "Sensor fusion using dempster-shafer theory II: Static weighting and Kalman filter-like dynamic weighting," in *Proc. 20th IEEE Information and Measurement Technology Conference*, 2003, pp. 907-912.
[18] É. Bossé, P. Valin, A. C. Boury-Brisset, and D. Grenier, "Exploitation of a priori knowledge for information fusion," *Inf. Fusion*, vol. 7, no. 2, pp. 161-175, June 2006.
[19] M. Morbee, L. Tessens, H. Aghajan, and W. Philips, "Dempster-Shafer based multi-view occupancy maps," *Electron. Lett.*, vol. 46, no. 5, pp. 341-343, March 2010.
[20] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," in *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Berlin/Heidelberg: Springer, 2008, pp. 57-72.
[21] D. Mercier, É. Lefèvre, and F. Delmotte, "Belief functions contextual discounting and canonical decompositions," *Int. J. Approx. Reason.*, vol. 53, no. 2, pp. 146-158, February 2012.
[22] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence, Warwick 1980*, Berlin/Heidelberg: Springer, 1981, pp. 366-381.
[23] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A*, vol. 33, no. 2, p. 1134, February 1986.
[24] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A*, vol. 45, no. 6, p. 3403, March 1992.
[25] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhysics Lett.*, vol. 4, no. 9, p. 973, 1987.
[26] P. Smets and R. Kennes, "The transferable belief model," *Artif. Intell.*, vol. 66, no. 2, pp. 191-234, April 1994.
[27] P. Smets, "Belief function: The disjunctive rule of combination and the generalized bayesian theorem," *Int. J. Approx. Reason.*, vol. 9, no. 1, pp. 1-35, 1993.
[28] D. Mercier, F. Pichon, and É. Lefèvre, "Corrigendum to 'Belief functions contextual discounting and canonical decompositions' [International Journal of Approximate Reasoning 53 (2012) 146-158]," *Int. J. Approx. Reason.*, vol. 70, pp. 137-139, 2016.
[29] P. A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*, New Jersey: Prentice hall, 1982.
[30] M. Saeed, *et al.*, "Multiparameter intelligent monitoring in intensive care II (MIMIC- II): A public-access intensive care unit database," *Crit. Care*, vol. 39, no. 5, pp. 952-960, May 2011.
[31] A. L. Goldberger, *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, June 2000.
[32] T. Klingeberg and M. Schilling, "Mobile wearable device for long term monitoring of vital signs," *Comput. Methods Programs Biomed.*, vol. 106, no. 2, pp. 89-96, May 2012.
[33] O. Gajic, *et al.*, "Early identification of patients at risk of acute lung injury: Evaluation of lung injury prediction score in a multicenter cohort study," *Am. J. Respir. Crit. Care Med.*, vol. 183, no. 4, pp. 462-470, February 2011.
[34] R. Mac Sweeney and D. F. McAuley, "Acute respiratory distress syndrome," *Lancet*, vol. 388, no. 10058, pp. 2416-2430, 2016.
[35] D. Chaudhury, J. Hasan, S. Paul, and I. Ali, "A study on clinical profile and outcome of patients with acute respiratory distress syndrome in a tertiary care hospital in North East India," *Sepsis*, vol. 13, pp. 25-29, 2017.
[36] A. J. Camm, *et al.*, "Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," *Circulation*, vol. 93, no. 5, pp. 1043-1065, 1996.
[37] A. Bauer, J. W. Kantelhardt, P. Barthel, R. Schneider, T. Mäkikallio, K. Ulm, K. Hnatkova, A. Schömig, H. Huikuri, A. Bunde, and M. Malik, "Deceleration capacity of heart rate as a predictor of mortality after myocardial infarction: Cohort study," *Lancet*, vol. 367, no. 9523, pp. 1674-1681, May 2006.
[38] P. J. Millar, A. S. Levy, C. L. McGowan, N. McCartney, and M. J. MacDonald, "Isometric handgrip training lowers blood pressure and increases heart rate complexity in medicated hypertensive patients," *Scand. J. Med. Sci. Sports*, vol. 23, no. 5, pp. 620-626, October 2013.

Aline Taoum was born in Al Cheikh Mohamad, Lebanon, on Octover 4, 1992. She received the diploma degree in electrical engineering from the Lebanese University, Faculty of Engineering, Lebanon. She also received the Master degree in biomechanics and bioengineering from the Lebanese University, Lebanon and the University of Technology of Compiegne, France. She is currently working toward the Ph.D. degree in systems optimization and security from the University of Technology of Troyes, France and the Lebanese University, Lebanon. Her current research interests include biomedical applications, machine learning, and signal processing.

Farah Mourad-Chehade was born on January 15, 1984. She received the diploma degree in electrical engineering from the Lebanese University, Faculty of Engineering, Tripoli, Lebanon, in 2006. She also received the Master degree, in 2007, and the Ph.D. degree in 2010, in systems optimization and security from the University of Technology of Troyes (UTT), France. Since September 2011, she has been an associate professor at the UTT. She obtained her habilitation to direct research (HDR) in December, 2017. She works on multi-sensor data fusion, pattern recognition and modeling for biomedical and environmental applications. She serves as a reviewer for several journals (IEEE Transactions on Signal Processing, IEEE Transactions on Robotics, IEEE Transactions on Vehicular Technology, and Elsevier Signal Processing, etc.) and conferences (EUSIPCO, WOSSPA, GRETSI, ROADEF, etc.). She leads several projects for dependent elderly people localization, prediction of ARDS occurrence and diagnosis of multi-pathologies.

Hassan AMOUD was born in Minieh, Lebanon in 1980. He received his bachelor's degree in Electrical Engineering from the Lebanese University, Lebanon in 2002. He then received his master's degree in System Information and Technology from the University of Technology of Compiègne, France and his PhD degree in Signal Processing and Biomedical Engineering from the University of Technology of Troyes, France in 2003 and 2006, respectively. After completion of his degrees, he held a Postdoctoral Research position in the Charles Delaunay Institute at the University of Technology of Troyes for three years. In 2009, he joined the Lebanese University as an associate professor and was promoted to a full professor in 2017. His research interests include biomedical signal processing using nonlinear and kernel methods, with application to health monitoring.