Automatic Recognition of Pain, Anxiety, Engagement and Tiredness for Virtual Rehabilitation from Stroke: A Marginalization Approach

Jesús Joel Rivas*†, Lorena Palafox‡, Jorge Hernández-Franco‡, María del Carmen Lara§, Nadia Bianchi-Berthouze¶, Felipe Orihuela-Espina* and Luis Enrique Sucar*

*Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, México
Email: jrivas@ccc.inaoep.mx, f.orihuela-espina@inaoep.mx, esucar@inaoep.mx

†Universidad de Carabobo (UC), Valencia, Venezuela
Email: jrivas@uc.edu.ve

‡Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico City, Mexico
Email: jhfranco@medicapolanco.com

§Benémerita Universidad Autónoma de Puebla (BUAP), Puebla, Mexico
Email: carmen.lara@correo.buap.mx

¶University College of London (UCL), London, United Kingdom
Email: n.berthouze@ucl.ac.uk

***Abstract*—Virtual rehabilitation taps affective computing to personalize therapy. States of anxiety, pain and engagement (affective) and tiredness (physical or psychological) were studied to be inferable from metrics of 3D hand location -proxy of hand movement- and fingers' pressure relevant for upper limb motor recovery. Features from the data streams characterized the motor dynamics of 2 stroke patients attending 10 sessions of motor virtual rehabilitation. Experts tagged states manifestations from videos. We aid classification contributing with a marginalization mechanism whereby absent input is reconstructed. With the hand movement information absent, marginalization statistically outperformed a base model where such input is ignored. Marginalized classification performance was (Area below ROC curve: $\mu \pm \sigma$) 0.880 ± 0.173 and 0.738 ± 0.177 for each patient. Marginalization aid classification sustaining performance under input failure or permitting different sensing settings.**

***Index Terms*—Multimodal systems; marginalization; virtual rehabilitation; stroke; probabilistic graphical models.**

1. Introduction

Affective computing has potential applications in virtual rehabilitation to adjust the therapies to the patients' specific needs. However, automatically recognizing the affective state that someone is experiencing represents a challenge for computer systems. Some computational models make use of information from different sensors [1]. Face expressions, body posture, hand gesticulations, and voice tone are external expressions that can help to estimate the individual's affective state; however there are also sensors that can measure internal physiological reactions (brain electrical activity, heart rate, blood pressure, etc) due to changes of the autonomic nervous system activity [2]; but these alternative sensors can be obstructive for the user's free movement and are not always available in everyday life settings. In machine learning systems, at training phase, all sensors could be used to build a model to predict the presence or absence of an affective state; and afterwards, at using phase, some of the sensors could be marginalized (i.e., not used) to fit the different real-life context. For example, in some situations, the video camera may not be usable because of privacy issues. This probable can be addressed by referring to the concept of marginalization in probability theory, which is used when we have a joint probability distribution $P(X_1, X_2, \cdots, X_n)$ and we want to calculate the marginal distribution of a variable or a subset of them, to obtain an expression or value, where the rest of the variables are marginalized, i.e., they are not present any more. For example, let's suppose we have three random variables: X, Y and Z and their joint probability distribution is $P(X, Y, Z)$, then we want to obtain the joint marginal distribution of X and Y, so $P(X, Y) = \sum_Z P(X, Y, Z)$ is calculated and the variable Z is not longer present in the result of $P(X, Y)$, i.e., Z has been marginalized. The marginalization idea can be used to reduce the need of some sensors. It consists in setting as absent any sensor(s) and using the others, and the information previously learned from the missing sensor(s),

to obtain as much information as possible to attenuate the loss that the system can suffer at the classification stage.

We hypothesized that a multimodal computational model based on probabilistic graphical models incorporating sensors marginalization strategies produces better affective states recognition performances than models which totally ignore and do not use any information, at testing phase, of an indicated (marginalized) sensor (called hereafter trivial marginalization). This model based on probabilistic graphical models would have the flexibility of leaving aside any of the sensors in the everyday use of the model, but trying not to lose all the information of the marginalized sensor. In this work, the development of a multimodal computational model based on probabilistic graphical models is proposed for the automatic recognition of affective states within a virtual rehabilitation platform. This model must have the capacity of sensor marginalization, which has social impact because it could be used in more conventional spaces, such as home. The identification of the patient's affective state can subsequently help to adjust the system to specific needs and to promote engagement with the rehabilitation exercises.

A feasibility pilot is introduced here whereby 2 post-stroke patients with upper limb impairment attended longitudinally 10 rehabilitation sessions over 4 weeks, employing a virtual rehabilitation platform while their hand movements and fingers' pressure information (experimental units) were captured and supplied as inputs to the proposed computational model. Since people suffering from stroke have stronger differences in their motor disabilities and hence behavioural expression, personalization is often patient-based. We hence decided to develop independent classifiers for each patient. From the classifiers point of view, the observations correspond to the local hand movements and fingers' pressure behaviour of the patient during the affective episodes. In other words, the sample size is not 2 patients, but the amount of affective episodes for each patient as labelled by experts. The registered affective states were manifested spontaneously by the patients whilst they attended the virtual rehabilitation program. At this stage of our research, we are only trying to recognize the patient's affective states, so we do not intent to control the rehabilitation platform yet, for being adjusted to the patient's needs at this moment. Preliminary experiments reveal the feasibility of the marginalization proposal and its effectiveness showing better performances than those obtained with trivial marginalization. The proposed marginalization-approach-based computational model represents a contribution to the affective computing field. The new model could favour intelligent and empathic human computer interactions that address the constraints and variety of everyday life setting.

This paper is organized as follows, section 2 summarizes previous works in machine learning and learning using privileged information approaches. Section 3 describes the methodological framework which includes the information of the movement/pressure feature vector, the design of the proposed classification model and the design of a preliminary experiment to evaluate the feasibility of the proposed marginalization solution. Section 4 highlights the results achieved; and finally the discussion, conclusions and future work are contained in section 5.

2. Background

Support vector machine (SVM) is generally used in affective recognition systems [3]. The Naïve Bayes classifier has been studied and compared with other classifiers and has performed more effectively than sophisticated rules [4], [5]. Our proposal includes a classifier derived from Naïve Bayes, i.e. Semi-Naïve Bayes [6], for its efficiency, simplicity and because it deals with dependent features [7].

One alternative for marginalization is the paradigm of learning using privileged information which was first introduced by Vladimir Vapnik and Akshay Vashist [8] and implemented with SVM. The potential of this paradigm can be considered for its possibility to register several sensors as privileged information during the training phase, so the information of these sensors will not be available during the testing phase. One of the applications has been the implicit tagging of emotional videos, in which the observer's physiological responses are analyzed to label segments of the video with the emotions the user experienced [9]. In Wang et al. [10], the implicit video emotion tagging and the recognition of emotions were studied. Features were created from EEG signals and from audio-visual information of the videos used to induce emotions. Through canonical correlation analysis (CCA) two feature spaces were created, one for EEG and the other for the video; two SVMs were trained separately over each feature space. The results for valence and arousal classifications were better than the ones obtained from classical SVM classifications. The drawback of this approach is that it does not provide a solution when there are more than two sensors. Chen et al. [11] developed an SVM model with similarity restrictions in the mapping functions to capture the relationship between EEG signals, multiple user peripheral physiological signals (EOG, EMG, ECG, GSR, RSP, TEMP and PLET [1]) and the features of video content. In this case the EEG signals and the different peripheral physiological signals represent the privileged information to implicit video tagging, i.e., during the testing phase (video tagging) only video features are available.

Bayesian networks inherently allow marginalization of features and, for this reason, are useful for learning using privileged information [12]. Basically, Wang et al. [12] studied Bayesian structures of three general nodes: the class node y which represents the emotion variable, the available information node x and the privileged information node x^*. They also studied all possible connections and directions of the arcs between these nodes.

1. EOG: Electrooculography; EMG: Electromyography; ECG: Electrocardiography; GSR: Galvanic Skin Response; RSP: Respiration; TEMP: Skin temperature and PLET: Plethysmograph.

3. Methods

3.1. Hand movements and fingers' pressure Dataset

In a previous work [16] a dataset was constructed, which contains the records of the rehabilitation sessions of 2 stroke patients that attended therapies to recover the mobility of their upper limb. The patients (an extroverted man and an introverted woman, as judged by the psychiatrists who contributed to this study) participated in 45-minute average sessions that took place in different days (max 3 per week) in a period of 4 weeks. The virtual rehabilitation platform *Gesture Therapy* (GT) [17] was used in each session, and the 3D movements of the affected hand, the fingers' pressure and a frontal video, of the patient, were recorded synchronously at 15 Hz. To record the hand movements and the fingers' pressure, two sensors monitored through the gripper of the GT system were used: one to continuously track the 3D position of the hand (called hereafter MOV sensor, and it tracked the gripper's coloured ball using the computer camera) and the other to track the pressure exerted by the fingers (called hereafter PRE sensor, integrated in front of the gripper) (see figure 1). The patient's frontal video was used by three psychiatrists to tag the frames in which the patient exhibited one or more of the 4 states: tiredness (physical or psychological), anxiety, pain and engagement. Each of the states was rated with binary values, 1 for the presence of the state and -1 for the absence of the state.

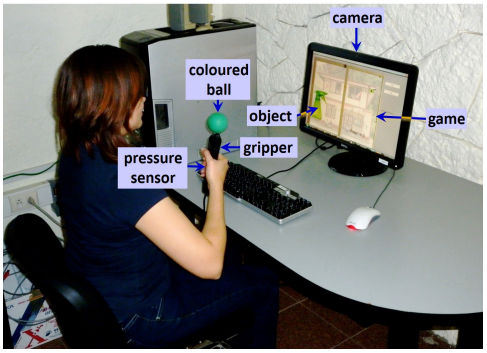


Figure 1. Demonstration of the GT Platform. The gripper, held here with the right hand, is monitored with a tracking system that uses the computer camera to locate the position of the hand (coloured ball), and controls an object in the virtual environment. The object is represented, in this case, by the aerosol and when the patient's fingers press the pressure sensor, the bottle sprays the insecticide to kill a mosquito. As the user interacts with the rehabilitation game, the 3D hand location, the gripping force and a frontal video are recoded at 15 Hz.

3.2. Multiresolution Semi-Nave Bayesian classifier (MSNB)

One difficulty for detecting emotions is their sudden appearance, generated by some stimulus, and the fact that their duration is highly variable [13]; in some cases relatively short [14]. No consensus has been reached on how long the emotions last [14]. In a previous work [15] a binary classifier was proposed to explore the appearance of affective states of interest on the trace over time. The classifier operationalizes several odd-size windows W (starting from 3) concentric to a current point p_i that shift simultaneously over the trace, and which it becomes possible to calculate, in the current point environment, several features $(A_{j1}, A_{j2}, ..., A_{jb};\ j \in \{1, 2, ..., 5\},\ j$ = identification of SNB classifier, b = amount of features) that may help to discriminate the presence of the affective state (see figure 2). This classifier was called Multiresolution Semi-Nave Bayesian classifier (MSNB) because the windows represent several simultaneous resolutions at the current point p_i of the trace. The classifier represents an ensemble of Semi-Naïve Bayesian classifiers (SNB) [6] with a late (decision level) fusion process by majority vote. Each SNB receives the features coming from a different window size $|W|$ and infers the presence (1) or not (-1) of the affective state of interest (variable $C_{cj} \in \{-1, 1\}$, where C_{cj} = estimated class through classifier j, $j \in \{1, 2, ..., 5\}$). At the end, at the fusion stage, the presence or not is decided (in variable $C \in \{-1, 1\}$), by means of the majority vote of the SNBs. In figure 2, part b), the architecture of MSNB is shown. Because of the good results obtained in the previous work [15], we decided to use MSNB as the base classifier for the computational model to this proposal.

3.3. Feature extraction

From the data acquisition process three series were obtained: $Mser$, $Pser$ and $Aser$, representing the hand movements series (from MOV sensor), the fingers' pressure series (from PRE sensor) and the annotation series respectively. Each sample of the dataset was generated with the feature extraction from the consecutive points considered in same size windows W that were shifted over $Mser$, $Pser$ and $Aser$ synchronously. From points stream of $Mser$, 5 features were obtained: (averages, in points of W, of) speed, acceleration and differential location along the x, y, z axes. From points stream of $Pser$, 3 features were generated: (averages, in points of W, of) pressure, pressure speed and pressure acceleration. The classes consisted of 4 binary labels (of the set $\{-1, 1\}$), one for each affective state, indicating presence (1) or absence (-1) of the respective affective state. Each class label was obtained as the majority label assigned by the experts to the video frames ($Aser$ series) considered in W. Video frames were synchronized with the motions of the upper limb. Regardless of windows sizes, the number of samples for each one is the same because all windows are concentric and, start and stop synchronously. A total of 5826 samples were obtained, for each window size, for $P1$ and 8935 for $P2$. For patient $P2$, the affective state of *pain was not observed* in any of her videos.

3.4. Proposed model

To address the marginalization of a sensor (MOV or PRE), the proposed computational model, employs the

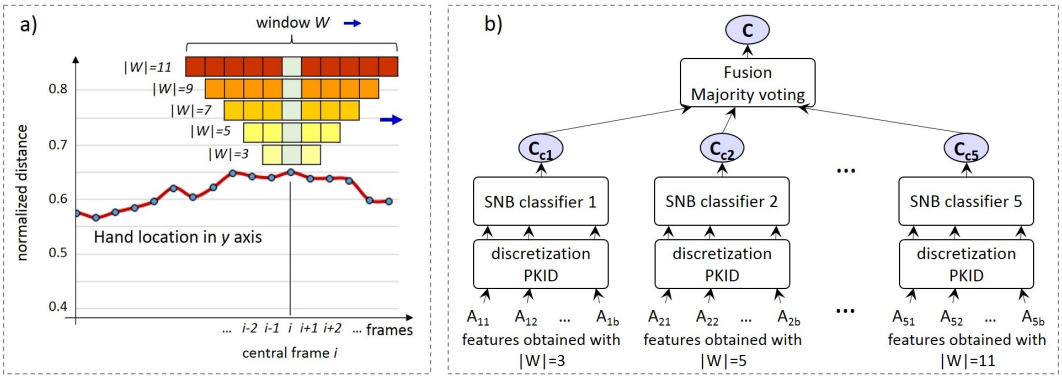


Figure 2. Multiresolution Semi-Nave Bayesian classifier (MSNB). In a), we present an example of the process of multiresolution with the use of windows, of different odd size, $|W|$ = 3,5,7,9,11; for the hand movements series in y axis. This example, for reasons of simplicity, shows the series in one of the axes; however the complete hand movements series is in 3D. Similarly, these windows are shifted on fingers' pressure series and over the labels series synchronously. Each window W represents a surrounding area of the current point p_i of the series. A semi-Nave Bayesian (SNB) [6] model is built, for each window W, to estimate the presence (1) or not (-1) of the affective state in W. In b) the architecture of the classifier is presented. Features $(A_{j1}, A_{j2}, ..., A_{jb}$; $j \in \{1, 2, ..., 5\}$, j = identification of SNB classifier, b = amount of features) from several concentric odd-size window W with respect to a point p_i of the series, are supplied and discretized with PKID method [5]; then SNB classifiers independently decide whether or not the affective state exists (variables $C_{cj} \in \{-1, 1\}$, where C_{cj} = estimated class value through classifier j) for that window size $|W|$. These inferences are received by the late fusion module where, by majority voting, finally decides (in variable $C \in \{-1, 1\}$) whether the affective state is present or not around p_i.

MSNB as base classifier to infer the affective state of interest from the features of the respective sensor. Then the information inferred from each MSNB is supplied to a Semi-Nave Bayesian (SNB) classifier at late (decision level) fusion. The complete model is called fusion-using-SNB (FSNB). The FSNB architecture is presented in figure 3, part a). The proposed strategy for marginalization consists in estimating the features' values of the marginalized sensor through simple linear regression from the features' values of the other sensor. The Pearson correlation coefficient was calculated between each feature of the available sensor and each feature of the missing sensor, choosing the relation that had the highest value. Subsequently, the corresponding linear models were created and with them the values of the features of the missing sensor were produced (see figure 3, parts b1 and b2). Therefore, FSNB models received two types of outputs from MSNB: outputs obtained from the available sensor and outputs generated over estimations for the missing sensor.

3.5. Experimental design

An experiment was carried out involving FSNB models to which a sensor was absent: MOV or PRE, one at a time. FSNB models were generated for each patient and for each affective state, so 4 models for $P1$ and 3 for $P2$ (remember $P2$ did not exhibited pain state during the labelling). The corresponding MSNBs were constructed using the odd window sizes of 3 to 11, i.e., $|V|$ = 3,5,7,9,11 (see section 3.2); so 5 different window sizes were used. Internal validation was performed using the stratified 10-fold cross-replication mechanism.

The marginalization process was evaluated as follows: One sensor at the time was marginalized and the effectiveness of the marginalization strategy was investigated. Three models were compared: (a) The model with the two sensors available, i.e. the complete model without marginalization, (b) the model where a sensor is marginalized following the proposed strategy; and (c) the model where only the other sensor, the available one, is considered (trivial marginalization). Accuracy, F-measure, and ROC AUC metrics [2] were used to compare performances. The results of the three models were contrasted to see if the marginalization model (b) achieve a better performance than the model (c) and how far is from model (a).

The experimental hypothesis is that by estimating the feature values of the missing sensor from the feature values of the available sensor using simple linear regression, the FSNB model achieves better recognition performances than the model of trivial marginalization. The objective of the experiment is to validate the results of the proposed marginalization strategy with respect to the results of the trivial marginalization, and the results of FSNB developed with all the sensors (MOV and PRE).

4. Results

The results of the proposed marginalization strategy when MOV sensor was marginalized, called FSNB-$\widehat{MOV}$-PRE (FSNB with MOV estimations, and PRE), are presented in Table 1 for both patients and for all the affective states. The results are summarized as $\mu \pm \sigma$ across the 10 folds, and are compared with the respective results of trivial marginalization, which corresponds to FSNB-PRE where PRE sensor was only considered, and with the results of

2. TP: true positive, TN: true negative, FP: false positive and FN: false negative; Accuracy = (TP+TN)/(TP+FP+TN+FN); F-measure = 2(precision*sensitivity)/(precision+sensitivity); ROC AUC = (Sensitivity + Specificity)/2.

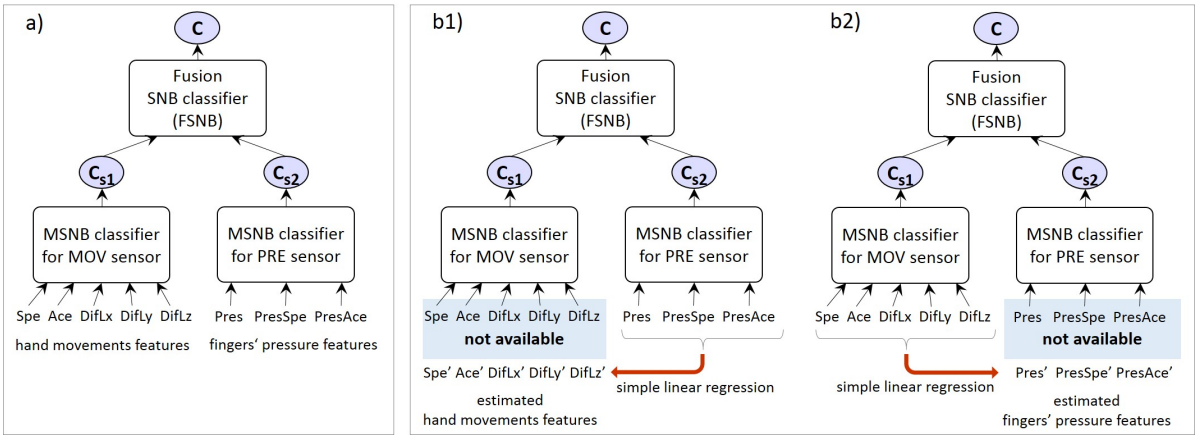


Figure 3. Proposed model. In a) FSNB classifier is presented. Its components are MSNB as the base classifier, one for each sensor, and a decision fusion module. In this case, a MSNB classifier receives the features of hand movements sensor (MOV sensor) and the other receives the features of fingers' pressure (PRE sensor). Each MSNB classifier estimates the presence (1) or absence (-1) of an affective state (variables $C_{sj} \in \{-1, 1\}$, where C_{sj} = estimated class value through classifier from sensor j, $j = 1, 2$). All the MSNB's estimations C_{sj} are fused using SNB classifier at decision module, to infer finally (in variable $C \in \{-1, 1\}$) whether the affective state is present or not. Hand movements features are represented as: (average of) speed (Spe), acceleration (Ace), differential location x (DifLx), differential location y (DifLy) and differential location z (DifLz); fingers' features are indicated as: (average of) pressure (Pre), pressure speed (PresSpe) and pressure acceleration (PresAce). In b1) the MOV sensor is not available and each of its feature values were estimated through simple linear regression from the most associated feature of PRE; these features estimations of MOV are used by first MSNB classifier to obtain the class C_{s1}. In b2) the opposite case is shown when PRE is not available.

FSNB-MOV-PRE where the two sensors are available. In all states of $P1$, except engagement, and in the average over states, the ROC area results of FSNB-$\widehat{MOV}$-PRE are between the ROC area results of FSNB-PRE and FSNB-MOV-PRE. For $P1$, the results for anxiety state of FSNB-$\widehat{MOV}$-PRE were the same as the base case: FSNB-PRE. In all states of $P2$ and in the average over states, the ROC area results of FSNB-$\widehat{MOV}$-PRE are between the ROC area results of trivial marginalization and FSNB-MOV-PRE.

The models FSNB-$\widehat{MOV}$-PRE obtained results significantly higher than those of FSNB-PRE, using the ROC area of affective states (Wilcoxon test of sign ranges: $W = -3.823$, $p < 0.05$).

When PRE sensor was marginalized (model FSNB-MOV-$\widehat{PRE}$), the results were the same as trivial marginalization, i.e. as those of FSNB-MOV, where MOV sensor was only considered. In this case it is evident that the estimation of the feature values of PRE sensor from the features of MOV sensor was not good enough to give information about the original values of PRE sensor. This result reveals that PRE is more difficult to estimate by simple linear regression from MOV.

Average results for $P1$ were higher than average results for $P2$. The best recognized states through FSNB-$\widehat{MOV}$-PRE for $P1$ were pain and engagement; and for $P2$ was tiredness.

5. Discussion and Conclusions

The problem of the absence of a sensor at testing phase has been explored. The performance of the chosen strategy, in a two sensor problem, was mixed, for MOV sensor the estimation of its features, in general, contributed to improve performance with respect of trivial marginalization; but the results for PRE sensor did not leveraged the performance. Although the proposed marginalization strategy is simple, it produces for both patients in MOV sensor marginalization, average results which overcome the corresponding ones of trivial marginalization, increasing the mean value and decreasing the standard deviation.

The conjunction of information from different sensors can contribute to the recognition of affective states, but this is affected by relations of complementarity, redundancy or noise between sensors. For example, for engagement of patient P1, the MOV sensor introduces noise and its presence alters the classification performance decreasing the average results values and increasing its standard deviation. For anxiety of P1, the results suggest that MOV sensor is absolutely necessary (complements the PRE sensor), and the proposed marginalization strategy could not replace its values with estimates ones.

The proposed marginalization strategy can be generalized for studying the marginalization in a system of more than two sensors. When a specific sensor is marginalized, its feature values can be estimated through the values of the remaining sensors features using linear regression.

Sensor marginalization can give us an alternative to use systems of automatic affective states recognition in everyday life. An issue to consider is which sensor can be marginalized or which sensor is absolutely necessary, i.e., if the marginalization of a sensor leads to a small increase in error, this may be acceptable compare to the cost or the side effect that using that sensor may lead to. However, if the increase in error is not acceptable, then that sensor cannot be marginalized. Another issue is that decision may differ between the affective states that are crucial and the ones that

TABLE 1. CLASSIFICATION RESULTS ($\mu \pm \sigma$, THROUGH THE 10 FOLDS OF CROSS-VALIDATION) IN FSNB-MOV-PRE, FSNB-$\widehat{MOV}$-PRE AND FSNB-PRE.

		Patient $P1$					**Patient** $P2$		
Affective state	Method	Accuracy	F-Measure	ROC area	Affective state	Method	Accuracy	F-Measure	ROC area
tiredness	FSNB-MOV-PRE	0.951 ± 0.070	0.957 ± 0.058	0.945 ± 0.079	tiredness	FSNB-MOV-PRE	0.921 ± 0.092	0.915 ± 0.106	0.921 ± 0.092
	FSNB-$\widehat{MOV}$-PRE	0.868 ± 0.178	0.899 ± 0.126	0.869 ± 0.164		FSNB-$\widehat{MOV}$-PRE	0.907 ± 0.112	0.900 ± 0.122	0.907 ± 0.112
	FSNB-PRE	0.860 ± 0.186	0.894 ± 0.130	0.859 ± 0.178		FSNB-PRE	0.900 ± 0.108	0.893 ± 0.117	0.900 ± 0.108
anxiety	FSNB-MOV-PRE	0.926 ± 0.109	0.938 ± 0.088	0.924 ± 0.112	anxiety	FSNB-MOV-PRE	0.760 ± 0.081	0.796 ± 0.054	0.749 ± 0.094
	FSNB-$\widehat{MOV}$-PRE	0.858 ± 0.172	0.895 ± 0.117	0.852 ± 0.180		FSNB-$\widehat{MOV}$-PRE	0.720 ± 0.095	0.764 ± 0.073	0.710 ± 0.101
	FSNB-PRE	0.858 ± 0.172	0.895 ± 0.117	0.852 ± 0.180		FSNB-PRE	0.661 ± 0.089	0.736 ± 0.063	0.642 ± 0.105
pain	FSNB-MOV-PRE	0.956 ± 0.141	0.950 ± 0.158	0.971 ± 0.090	pain	FSNB-MOV-PRE	–	–	–
	FSNB-$\widehat{MOV}$-PRE	0.894 ± 0.193	0.913 ± 0.178	0.889 ± 0.188		FSNB-$\widehat{MOV}$-PRE	–	–	–
	FSNB-PRE	0.872 ± 0.254	0.869 ± 0.313	0.839 ± 0.296		FSNB-PRE	–	–	–
engagement	FSNB-MOV-PRE	0.917 ± 0.163	0.922 ± 0.157	0.916 ± 0.167	engagement	FSNB-MOV-PRE	0.624 ± 0.090	0.635 ± 0.109	0.624 ± 0.089
	FSNB-$\widehat{MOV}$-PRE	0.909 ± 0.180	0.916 ± 0.168	0.909 ± 0.184		FSNB-$\widehat{MOV}$-PRE	0.596 ± 0.151	0.560 ± 0.198	0.596 ± 0.151
	FSNB-PRE	0.906 ± 0.189	0.867 ± 0.284	0.918 ± 0.162		FSNB-PRE	0.570 ± 0.170	0.529 ± 0.227	0.570 ± 0.170
average	FSNB-MOV-PRE	0.937 ± 0.122	0.942 ± 0.119	0.939 ± 0.115	average	FSNB-MOV-PRE	0.768 ± 0.150	0.782 ± 0.147	0.765 ± 0.152
	FSNB-$\widehat{MOV}$-PRE	0.882 ± 0.175	0.906 ± 0.144	0.880 ± 0.173		FSNB-$\widehat{MOV}$-PRE	0.741 ± 0.175	0.741 ± 0.196	0.738 ± 0.177
	FSNB-PRE	0.874 ± 0.196	0.904 ± 0.169	0.867 ± 0.205		FSNB-PRE	0.710 ± 0.187	0.719 ± 0.211	0.704 ± 0.192

are less crucial for a specific problem.

In future work, we will study the proposed model with more patients and considering placing a Bayesian network at late (decision level) fusion. It will be useful to know the performance of this classifier and compare it with the results obtained so far with FSNB. The architecture where a Bayesian network is placed at decision level, would favour to tackle the problem of marginalization due to Bayesian networks mechanisms.

Acknowledgments

The authors would like to thank to red UbiSalud from the Mexican CONACYT (U0003-2015-1-253669) and Scholarship No. 434867 from the Mexican CONACYT.

References

[1] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Transactions on affective computing*, vol. 1, no. 1, pp. 18–37, 2010.

[2] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.

[3] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using EEG," in *International Conference on Augmented Cognition*. Springer, 2013, pp. 492–501.

[4] D. J. Hand and K. Yu, "Idiot's Bayes - not so stupid after all?" *International statistical review*, vol. 69, no. 3, pp. 385–398, 2001.

[5] Y. Yang and G. I. Webb, "Proportional k-interval discretization for naive-bayes classifiers," in *12th European Conference on Machine Learning (ECML)*. Springer, 2001, pp. 564–575.

[6] M. J. Pazzani, "Searching for dependencies in Bayesian classifiers," in *Learning from Data*. Springer, 1996, pp. 239–248.

[7] L. E. Sucar, *Probabilistic Graphical Models: Principles and Applications*. Springer, 2015.

[8] V. Vapnik and A. Vashist, "A new learning paradigm: Learning using privileged information," *Neural Networks*, vol. 22, no. 5, pp. 544–557, 2009.

[9] M. Soleymani and M. Pantic, "Human-centered implicit tagging: Overview and perspectives," in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2012, pp. 3304–3309.

[10] S. Wang, Y. Zhu, L. Yue, and Q. Ji, "Emotion recognition with the help of privileged information," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 189–200, 2015.

[11] S. Chen, S. Wang, C. Wu, Z. Gao, X. Shi, and Q. Ji, "Implicit hybrid video emotion tagging by integrating video content and users' multiple physiological responses," in *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, 2016, pp. 295–300.

[12] S. Wang, M. He, Y. Zhu, S. He, Y. Liu, and Q. Ji, "Learning with privileged information using Bayesian networks," *Frontiers of Computer Science*, vol. 9, no. 2, pp. 185–199, 2015.

[13] P. Verduyn, P. Delaveau, J.-Y. Rotgé, P. Fossati, and I. Van Mechelen, "Determinants of emotion duration and underlying psychological and neural mechanisms," *Emotion Review*, vol. 7, no. 4, pp. 330–335, 2015.

[14] K. R. Scherer, "What are emotions? and how can they be measured?" *Social science information*, vol. 44, no. 4, pp. 695–729, 2005.

[15] J. J. Rivas, "Clasificador semi-naïve Bayes con multiresolución para la estimación de estados afectivos: Aplicación en rehabilitación virtual," Master's thesis, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), 2015.

[16] J. J. Rivas, F. Orihuela-Espina, L. E. Sucar, L. Palafox, J. Hernández-Franco, and N. Bianchi-Berthouze, "Detecting affective states in virtual rehabilitation," in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2015 9th International Conference on*. IEEE, 2015, pp. 287–292.

[17] L. E. Sucar, F. Orihuela-Espina, R. L. Velazquez, D. J. Reinkensmeyer, R. Leder, and J. Hernández-Franco, "Gesture Therapy: An upper limb virtual reality-based motor rehabilitation platform," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 22, no. 3, pp. 634–643, 2014.