Multi-label and Multimodal Classifier for Affective States Recognition in Virtual Rehabilitation

Jesús Joel Rivas, María del Carmen Lara, Luis Castrejón, Jorge Hernández-Franco, Felipe Orihuela-Espina, Lorena Palafox, Amanda Williams, Nadia Bianchi-Berthouze, and Luis Enrique Sucar, *Senior Member, IEEE*

Abstract—Computational systems that process multiple affective states may benefit from explicitly considering the interaction between the states to enhance their recognition performance. This work proposes the combination of a multi-label classifier, Circular Classifier Chain (CCC), with a multimodal classifier, Fusion using a Semi-Naive Bayesian classifier (FSNBC), to include explicitly the dependencies between multiple affective states during the automatic recognition process. This combination of classifiers is applied to a virtual rehabilitation context of post-stroke patients. We collected data from post-stroke patients, which include finger pressure, hand movements, and facial expressions during ten longitudinal sessions. Videos of the sessions were labelled by clinicians to recognize four states: tiredness, anxiety, pain, and engagement. Each state was modelled by the FSNBC receiving the information of finger pressure, hand movements, and facial expressions. The four FSNBCs were linked in the CCC to exploit the dependency relationships between the states. The convergence of CCC was reached by 5 iterations at most for all the patients. Results (ROC AUC) of CCC with the FSNBC are over 0.940 ± 0.045 ($mean \pm std.deviation$) for the four states. Relationships of mutual exclusion between engagement and all the other states and co-occurrences between pain and anxiety were detected and discussed.

Index Terms—Affective states, affective states' dependency relationships, multi-label classification, multimodal classification, classifier chains, facial expressions, finger pressure, hand movements, posture, Semi-Naive Bayesian classifier, stroke, virtual rehabilitation.

1 Introduction

Affective computing systems that recognize multiple affective states may benefit from explicitly considering and exploiting the underlying dependency relationships between the affective states to enhance automatic recognition performances. This is critical now that affective recognition systems are used more and more in different fields of application, and in naturalistic everyday contexts. One of these fields of application is neuro-rehabilitation and healthcare in general. Specifically, for virtual rehabilitation, affective-aware platforms, could help post-stroke patients to perform their rehabilitation exercises by addressing their affective needs.

Virtual rehabilitation platforms provide opportunities for monitoring the multiple affective states experienced by the patients while interacting with the system. The automatic recognition of the patients' affective states can be useful for controlling virtual scenarios to leverage empathic and motivating interactions with the patients, to promote the adherence to the therapy [1], [2], [3], [4], and for adapting the exercise to not just physical by also psychological capabilities [5], [6]. Unfortunately, the correct detection of affective, physical, and/or cognitive states of the patients is still a challenge when applied to real data.

Some affective states tend to co-exist, while others do not; indeed, others are mutually exclusive. For example, in chronic pain rehabilitation, patients co-experience anxiety, and pain often expressed in the form of protective behaviour [7]. Another example, in the context of music, there exists songs that can elicit mixed emotions of relaxed-calm-sad, but it is quite improbable that songs elicit the emotions of surprise and quietness, or relaxed and angry at the same time [8].

Given the existence of such co-occurrences and mutually exclusive relationships between affective states, these relationships could be explicitly leveraged to support automatic recognition [9], [10]. A factor that could contribute to the robustness of the automatic recognition of affective states is when the underpinning computational models consider the interactions of the affective states modelled by the system.

This work proposes a multi-label classifier, combined with a multimodal classifier, that capitalizes on the dependency relationships between multiple affective states involved in an affective computing application. One of these applications is the automatic recognition of multiple affective, physical, and/or cognitive states involved in

- *Jesús Joel Rivas is with the Department of Computer Science, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, 72840 Mexico.*
  E-mail: jjoelrivas@gmail.com, jrivas@inaoep.mx
- *María del Carmen Lara is with the Department of Psychiatry of the Faculty of Medicine, Benémerita Universidad Autónoma de Puebla (BUAP), Puebla, Mexico.*
- *Luis Castrejón is with the Integral School of Physiotherapy, Benémerita Universidad Autónoma de Puebla (BUAP), Puebla, Mexico.*
- *Jorge Hernández-Franco is with the Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico City.*
- *Felipe Orihuela-Espina is with the Department of Computer Science, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, Mexico.*
- *Lorena Palafox is with the Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico City.*
- *Amanda Williams is with the Clinical, Educational and Health Psychology, University College of London (UCL), London.*
- *Nadia Bianchi-Berthouze is with the UCL Interaction Centre, University College of London (UCL), London.*
- *Luis Enrique Sucar is with the Department of Computer Science, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, Mexico.*

the rehabilitation of patients after stroke. In this case, the combination of the proposed classifiers aims to leverage the automatic recognition of the patients' states in a virtual rehabilitation platform, by considering the dependency relationships between the states[1].

The performance of the proposed multi-label classifier combined with the multimodal classifiers was assessed on a cohort of eight post-stroke patients for 10 virtual rehabilitation sessions, each patient over a month at the rehabilitation centre of a hospital. This dataset includes finger pressure (from a sensor we called PRE sensor), hand movements (from a sensor we named MOV sensor), and facial expressions (from a sensor we called FAE sensor) of the patients, gathered using a virtual rehabilitation platform called Gesture Therapy [1], [2]. For each of the eight patients, four states: tiredness, anxiety, pain, and engagement, were registered (through the labelling of the psychiatrists) while they participated in the virtual rehabilitation sessions.

Our proposal uses as base classifier, a derivation from Naive Bayes classifier, named Semi-Naive Bayesian classifier (SNBC) [12], for its efficiency, simplicity and because it tackles dependent features [13]. An advantage of the Bayesian approaches is that their models are interpretable, and this characteristic was useful for computing the conditional probability tables (CPTs) presented in subsection 5.5 for detecting the mutual exclusion and co-occurrences between the states.

The base classifier SNBC was used to build the Multiresolution SNBC (MSNBC) [14] and then to create the late Fusion of the three sensors (PRE, MOV, and FAE) using SNBC (called hereafter FSNBC). Finally, the dependency relationships between the states were exploited by a multi-label classifier named Circular Classifier Chain (CCC).

To evaluate the performance of CCC using the FSNBC, the dataset of post-stroke patients mentioned above was used in three experiments. The first experiment had the purpose of evaluating the convergence of CCC for the data of each patient. The second experiment allowed us to evaluate the CCC performance and compared it against the performance of the multi-label classifiers: Binary Relevance (BR) and Classifier Chains (CC) [15] (used as baselines). The three classifiers incorporated the FSNBC as the base classifier. Finally, the third experiment permitted to analyze whether the affective states ordering within the CCC could generate different results. Additionally, the conditional probability tables (CPTs) created automatically by CCC using the FSNBC, in the second experiment, were analyzed for trying to determine the dependency relationships between the states, that were captured by CCC using the FSNBC.

The contributions of this research are:

1) a novel architecture formed by a multi-label classifier called Circular Classifier Chain (CCC) combined with a set of multimodal classifiers called Fusion using a Semi-Naive Bayesian classifier (FSNBC) (one FSNBC for representing one of the states), to explicitly capture and leverage the dependency relationships between multiple affective states during the automatic recognition process;
2) a scheme for detecting the dependencies between affective states. In the application of rehabilitation of post-stroke patients, it was detected mutual exclusion between engagement and all the other states (tiredness, anxiety, and pain), and the co-occurrence of pain and anxiety for the patients during the rehabilitation sessions. The analysis of the emerged dependency relationships between the states shows that the relationships captured by the architecture reflect the clinical literature; and
3) a late fusion process in the multimodal classifier FSNBC that not only takes into account the predicted classes from each of the modalities involved in the problem but also includes correlations with the predicted classes of the other affective states.

1. This work is an extension of our research presented at the International Conference of Affective Computing and Intelligent Interaction ACII 2019 [11]. In this extended version, we included experimental validation by incrementing the number of post-stroke patients (from five to eight) in the longitudinal study; we extended the analysis of the convergence of the proposed multi-label classifier, Circular Classifier Chain (CCC); and we did further analysis to evaluate the impact of the affective states ordering within the CCC.

2 Related Work

2.1 Dependency Relationships between Affective States

Relationships of co-occurrence and mutual exclusion between emotions have been exposed in studies where emotions were elicited through video clips [16], [17]. Emotions such as anger and disgust were difficult to induce independently [16]. Their results also suggested that one emotion could trigger another; for example, anger may induce anxiety [16]. Video clips that provoked contentment, amusement also elicited happiness, but it never occurred that the videos that induced anger could induce levels of happiness at the same time [17]. Similarly, when music has been used to elicit emotions, some music can induce mixed emotions of calm-relaxed-sad, but it is improbable that it provokes the emotions of quietness and surprise, or relaxed and angry at the same time [8]. In the context of health, chronic pain patients exhibit protective behaviour during exercise in response to their anxiety, fear towards, and low confidence in such movements [5]. Additionally, the relationship between pain and protective behaviour is mediated by anxiety rather than being directly linked. This suggests that in some cases, emotional expressions that may be perceived as pain may indeed be a consequence of another emotional state, anxiety in this case, or a mixture of the two [7]. This highlights that not only relationships between states exist, but that also exists a directional dependency.

Very few works have addressed the dependency relationships between emotions and multidimensional classification [10]. Olugbade et al. [5] has modelled pain, anxiety, and confidence recognition as parallel and co-present expressions by building independent recognition models for the three states. However, her work has not taken advantage of their relationship. The works of [18] and [10] do exploit the dependency relationships between emotions. In [18], a Bayesian network is used to learn the relations of co-occurrence and mutual exclusion between pairs of emotions; but this is somewhat limited because it does not include the dependency relationships between more than two emotions simultaneously. In [10], a three-layer Boltzmann restrictive machine is used to detect the dependency relationships between more than one pair of emotions; but the problem is the computational cost involved in training and making inferences in a Boltzmann machine. In our proposal, we tackle the dependency relationships between two or more affective states by using Circular Classifier Chains [19] where the predicted classes of the previous affective states in the chain are incorporated as additional feature inputs to the succeeding classifiers. Our core classifier is the SNBC, which maintains the efficiency and simplicity of the Naive Bayesian classifiers [13].

2.2 Modalities: Finger Pressure, Hand Movements, and Facial Expressions

Computational models for automatic affect recognition benefit from including information from several sensors to improve classification rates [20], [21]. Complementarity between some sensors' signals may lead to an increase in recognition performance [20]. Most of the research in affective computing has mainly considered three kinds of modalities: visual (facial expressions), audio (vocalization), and text information (written communication) [21]. Around 10% of the research has addressed the integration of modalities related to body movements and other modalities like facial expressions [20], [21]. In particular, even less research has looked on the use of touch in combination with the modalities mentioned above; indeed, touch has been studied quite in isolation [22]. Recently, Filntisis et al. (2019) [23] studied the body movements and facial expressions of children for recognizing manifestations of affective states, to promote emphatic child-robot interaction. They used Deep Neural Networks (DNN) for processing each modality (body and facial), and made a late fusion through a fully connected layer where the scores obtained through the modalities were combined. Some works have explored affect recognition combining the modalities of hand gestures and facial expressions [24], [25], [26], [27]. Computational models included Hidden Markov Models (HMM) and fuzzy logic for studying the hand gestures and the facial expressions, respectively [24]. Other alternatives have been Bayesian Networks for each modality [25], and combinations of HMM, Adaboost, and Random Forest [26]. With respect to the fusion, the results of each modality (facial and hand) have been combined at late fusion using a weighted sum of the two modalities [24], [25], [26], or using only sum or product rule [25], [26].

A growing body of work has been exploring the automatic expression of not-lab induced pain (for a review on pain datasets see [28]), exploring a combination of several modalities [29], [30], [31] but not hand or touch movement. In addition, none of these works has attempted to directly exploit the relation of pain with other emotional states to improve recognition performances.

To our knowledge, the proposal of studying the combination of finger pressure and hand movements with facial expressions is novel. Additionally, we have not found studies about that combination of modalities in relation to pain, anxiety, engagement, and tiredness, together with the dependency relationships between these states. For rehabilitation therapies of post-stroke patients, finger pressure and hand movements recovery of an impaired upper limb is a specific target to achieve [2], so it is relevant to include these modalities in a virtual rehabilitation system [1], [2], [32], [33]. This is even more important if we aim to incorporate patients' affective state automatic recognition functions into a virtual rehabilitation platform [14], [34].

3 Dataset of Spontaneous Affective States of Post-stroke Patients

A dataset of post-stroke patients was collected and labelled by clinicians to develop and assess the performance of the proposed classifiers, which exploits the dependency relationships between the patient's affective states. The dataset consists in data from post-stroke patients that participated in virtual rehabilitation sessions using a computational platform called Gesture Therapy (GT) [1], [2]. The first version of this dataset, including 5 patients, was presented in [14], where only finger pressure and hand movements data were used. The dataset was then used in [11] by considering all modalities available, finger pressure, hand movements, and facial expressions. For this current paper, the dataset was extended to include 6 new patients, as described below. Unfortunately, 3 of them (P06, P09, and P10) could not be used for the training and evaluation of the multi-label automatic recognition task because their facial expression features could not be extracted. We briefly describe the GT system used to collect the data from patients, and then we describe the full dataset.

3.1 The Virtual Rehabilitation Platform: Gesture Therapy

Gesture Therapy (GT) [1], [2] is an upper limb rehabilitation platform for post-stroke patients consisting of a set of virtual reality serious games for doing the therapeutic exercises (Fig. 1).

GT integrates five interacting modules [2]:

1) Physical System (the hardware elements), composed of a personal computer, a webcam, and a device created in our lab, the gripper (see Fig. 1), which is held by the patient;
2) Tracking System, for tracking the hand movements through the gripper's colour ball, and for detecting the finger pressure exerted on the gripper's pressure sensor (PRE sensor);
3) Simulated Environment, to display the serious games, and to control the interaction with the patient;
4) Trunk Compensation Detector, to identify whether the patient is making compensatory movements; and
5) Adaptation System for real-time dynamical adjustment of the difficulty levels of the games to the patient's requirements and progress. The automatic recognition of the patient's affective states will be useful to provide real-time customization of the system not only to the physical needs but also to the psychological ones.

GT also includes capabilities for recording a video (through the webcam), for registering the patient's session.

The video is used, in this study, to capture the patient's facial expressions (FAcial Expression: FAE sensor) and for the labelling of the affective states as described in the next session. On the other hand, the tracker system estimates, at each video frame, the 3D coordinates of the hand movements (MOV sensor), and the finger pressure (PRE sensor) value, and conveys these values to the simulated environment. These tracked pieces of information (movements and pressure) are used for real-time customization of the game.

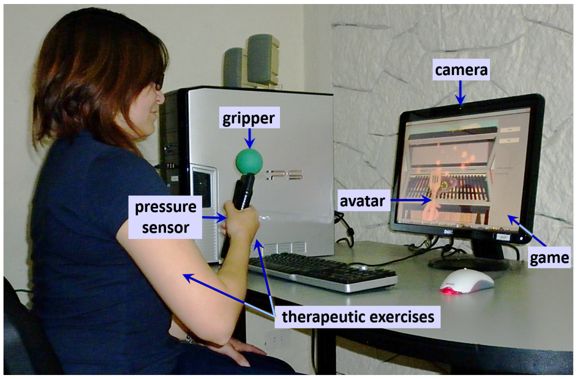


Fig. 1. A volunteer participating in the Gesture Therapy platform. The participant is holding the gripper with the right hand. The gripper has a frontal sensor for registering finger pressure. The webcam follows the gripper's colour ball to control an avatar (in this example, it is the kitchen palette) in the game.

3.2 Patient Recruitment and Data Collection

Eleven post-stroke volunteer patients were recruited from the Instituto Nacional de Neurología y Neurocirugía (INNN), Mexico City, and the Rehabilitation Centre of the Hospital Universitario de Puebla, Benemérita Universidad Autónoma de Puebla, in Mexico. The demographic information is summarized in Table 1. The patients received a brief explanation of the research before giving their consent to participate and agreeing that their rehabilitation sessions could be video recorded, and their hand movement and facial expressions data could be used for scientific purposes. After that, patients attended the virtual rehabilitation sessions using the GT during ten longitudinal sessions over a month approximately (each session was performed on a different day, at most 3 sessions per week). All the sessions were supervised by a qualified occupational therapist that had previous experience with the GT platform. Patients played 5 games for at most 3 minutes each. The precise amount of time was decided by the therapist. GT recorded a frontal video of the patient for each game at 15 frames per second, where the spontaneous facial expressions, the upper torso postures, and the hand movements could be tracked. It also recorded instantaneous hand location proxied by the gripper's ball and the gripping strength exerted by the fingers at 15 Hz synchronized with the video frames. The stream of 3D coordinates of the hand motions and finger pressure at each video frame, and the information of the facial expressions, were used as independent variables. Data were labelled, frame by frame, by a group of psychiatrists considering four spontaneous patients' states, tiredness, anxiety, pain, and engagement [14]. The states were considered as the dependent variables. This set of states had been decided through discussion with a group of clinicians formed by a therapist, psychiatrists, and an affective computing expert involved in the project. They considered that these states are relevant and critical during post-stroke rehabilitation to provide support to the patients and adjust the therapy to their needs.

TABLE 1
Cohort Demographics

Patient	Age [years]	Gender	Stroke date	Therapy onset	Paretic side	No. of sessions
P01	55	M	Apr, 2014	May, 2014	Left	6
P02	57	F	May, 2014	Sep, 2014	Left	10
P03	67	M	Jan, 2013	Jul, 2016	Right	10
P04	44	M	Dec, 2016	Jul, 2017	Right	8
P05	41	M	Nov, 2015	Jul, 2017	Right	5
**P06*	76	F	Oct, 2018	Nov, 2018	Left	5
P07	65	M	Nov, 2018	Dec, 2018	Left	10
P08	64	F	Dec, 2018	Dec, 2018	Right	10
**P09*	58	F	Nov, 2018	Jan, 2019	Right	10
**P10*	33	M	Dec, 2018	Feb, 2019	Right	10
P11	65	M	Mar, 2019	Apr, 2019	Left	10

* Data from patients P06, P09, and P10 were not included in the multi-label automatic recognition because the facial recognition system could not obtain the required features from their videos.

3.3 Feature Vectors

Feature vectors were created with a sliding window (of a predefined size) over consecutive frames of the respective data [14] (see subsection 4.2). This procedure yielded a feature vector for each step forward of the sliding window. The feature vector for finger pressure (from PRE sensor) contains 3 features (*averages of* the data contained in the sliding window): *pressure* (*Pres*), *pressure speed* (*PresSpe*) and *pressure acceleration* (*PresAce*). For hand movements, the feature vector (from MOV sensor) contains 5 features (*averages of* the data contained in the sliding window): *speed* (*Spe*), *acceleration* (*Ace*) and *differential location* by the axes: x (*DifLx*), y (*DifLy*), z (*DifLz*). Finally, for the facial expressions, 20 features from each frame of the patients' frontal video were extracted [35]. These features by frame represent distances of geometrical figures over the eyebrows, the eyes, and the mouth, and some angles over the eyebrows [36] (see Fig. 2). More precisely, the feature vector (from FAE sensor) contains 20 averaged features (*averages of* the data contained in the sliding window) for $F1$ (*avF1*), $F2$ (*avF2*), ..., $F20$ (*avF20*) (see Fig. 2 for the meaning of each feature number). All the feature vectors have four binary tags (from the set $\{-1, 1\}$), one for each state (tiredness, anxiety, pain, and engagement), indicating the presence (1) or the absence (-1) of the state. Since the data were labelled frame by frame, the corresponding tag was generated as the majority label in the sliding window. Data of the three sensors and the class tags were synchronized through the associated frames.

4 Classifiers: SNBC, MSNBC2, FSNBC, and CCC

The following classifiers were assembled to obtain the final model, which includes the affective states' relationships to improve their automatic recognition. The base model was

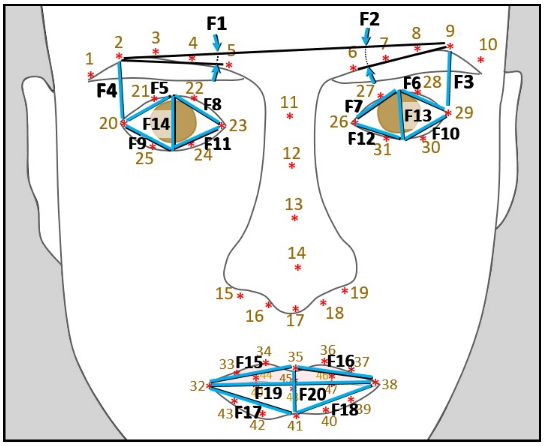


Fig. 2. Facial Expressions Features: $F_i,\ i \in \{1, 2, \cdots, 20\}$. F_1 and F_2 represent angles over the eyebrow, and $F_i,\ i \in \{3, \cdots, 20\}$ constitutes a distance of a geometrical figure over the eyebrows, the eyes, and the mouth, at a video frame.

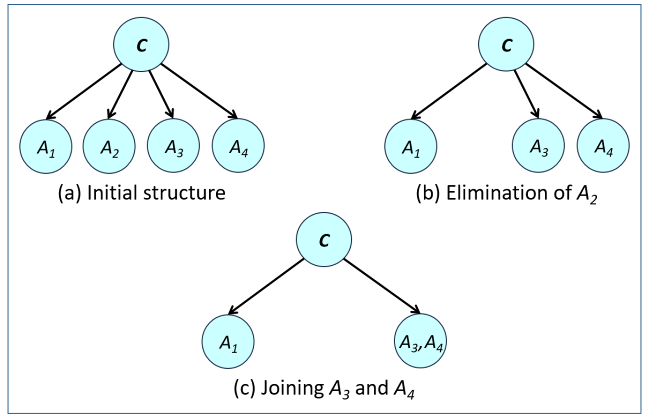


Fig. 3. Example of the process of structural improvement to obtain a Semi-Naive Bayesian (SNB) model: (a) An original Naive Bayes (NB) model with 4 features (all of them are assumed independent in NB), (b) Feature A_2 is eliminated because the mutual information value between A_2 and the class C is close to zero, (c) Features A_3 and A_4 are joined into one, as they are considered dependent based upon the conditional mutual information value between A_3 and A_4 given the class C.

the SNBC, which was fundamental to build all the other models. Then, for processing each sensor, i.e., each modality for affective states recognition, we used the Multiresolution Semi-Naive Bayesian classifier (MSNBC) (with a modification, and we call the new model, MSNBC2). The MSNBC2 for each sensor estimated the presence or the absence of an affective state, and a late Fusion using SNBC (FSNBC) was implemented to combine the individual sensors' predictions to recognize the occurrence of the affective state. There were as many FSNBCs as affective states, each one for recognizing one affective state independently. These FSNBCs were linked in a Circular Classifier Chain (CCC), which integrated the interactions of the affective states to enhance the final recognition.

4.1 Semi-Naive Bayesian Classifier (SNBC)

Semi-Naive Bayesian classifier (SNBC) is based on the Naive Bayes classifier (NBC) [12], [37]. Given a feature vector $\vec{A} = (A_1, A_2, \cdots, A_n)$ and given a sample $\vec{s}_a = (a_1, a_2, \cdots, a_n)$, the decision rule of NBC for a two-class problem (the class variable C takes values in {-1, 1}), is expressed as:

$$class(\vec{s}_a) = \arg\max_{c \in \{-1,1\}} (Prob(C = c) \prod_{i=1}^{n} Prob(A_i = a_i | C = c)) \tag{1}$$

The naive assumption that all features A_i are independent given the class C supports the multiplication in (1) [12]. To address a more generic and realistic situation, the SNBC executes a structural improvement [12], [38], [39] to remove and/or join features (to eliminate redundant or irrelevant features and/or join dependent features). The structural improvement (Fig. 3) employs mutual information and conditional mutual information calculations [40] between the features and the class to make the improvements. After each operation of elimination or join of features, the new structure is tested to determine whether classification performance is improved. The process is repeated until all features have been analyzed.

4.2 Multiresolution Semi-Naive Bayesian Classifier 2 (MSNBC2)

Multiresolution Semi-Naive Bayesian classifier (MSNBC) is a binary classifier to explore the occurrence of an affective state of interest in the trace over time [41]. The classifier operationalizes several odd-size sliding windows (starting from 3) concentric to a current frame. These parallel sliding windows are shifted simultaneously over the trace to calculate several features in the environment of the current frame (neighbourhood). There is a SNBC associated with each window to discriminate the presence or not of the affective state in the corresponding window. (Fig. 4). The name multiresolution is used because the windows represent several concurrent resolutions at the current frame of the trace. Therefore, the associated SNBCs constitute simultaneous sliding estimators at different resolutions. MSNBC represents an ensemble of SNBCs with a late (decision level) fusion process by majority vote. Each SNBC receives the features coming from a different window size and infers the presence or not of the affective state of interest. Finally, in the fusion stage, the presence or not is decided through the majority vote of the SNBCs. Since the input features were numeric values, we employed a discretization process called Proportional k-interval discretization (PKID) [42]. This process tries to match the number of intervals with the amount of values within each interval [42].

A modification was made to MSNBC replacing the majority voting with a SNBC in the late fusion module, and the resulting classifier was called MSNBC2. Fig. 5, part a), shows the architecture of MSNBC2.

4.3 Late Fusion using SNBC (FSNBC)

There is an independent MSNBC2 for each sensor (PRE, MOV, and FAE) to predict the occurrence of an affective state. Then, the predicted class labels (1 –presence– or -1 –absence–) of the three MSNBC2 are fused using a SNBC (FSNBC) [43]. FSNBC is a binary classifier that represents a multimodal affective states recognizer (Fig. 5, part b)).

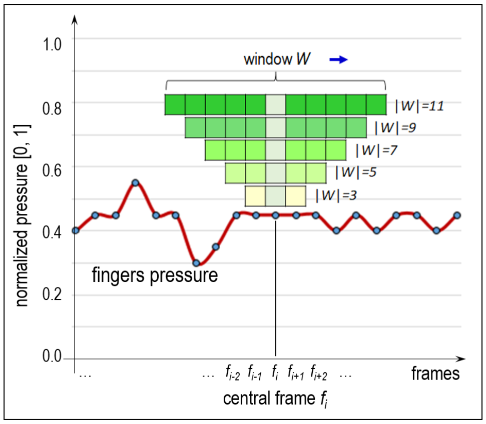


Fig. 4. Multiresolution process using several odd-size sliding windows (3, 5, 7, 9, and 11) concentric to a current frame f_i. The sliding windows are shifted simultaneously over the trace. Exemplification corresponds to the trace of finger pressure at each video frame during a segment of a rehabilitation session. A SNBC is trained for each window size to infer the presence (1) or absence (-1) of the affective state into consideration. Then, each of the 5 SNBC models (a model for each one of the 5 windows) returns its prediction for the class label in each sample f_i of the series, and the MSNBC2 makes a late fusion using a SNBC to assign the final class label (1 or -1) to f_i.

4.4 Circular Classifier Chain (CCC)

There are as many FSNBCs as affective states, each one for recognizing one state independently. These FSNBCs are connected into a Circular Classifier Chain (CCC), linking their predicted class labels between them to incorporate the dependency relationships between the states. The FSNBCs are the base classifiers of the multi-label classifier CCC.

CCC [19] is an extension of the multi-label classifier Classifier Chains (CC) [15]. CCC addresses the problem of defining the class variables' ordering in the chain. CC is related to Binary Relevance (BR), an approach that consists of q base binary classifiers for classifying q class variables, where each one is independently trained to predict the occurrence of a class variable. CC incorporates class interactions to the BR approach through a strategy of creating a chain where each classifier includes as additional features the predicted class labels of the previous classifiers in the chain (except for the first classifier) [15]. A drawback to CC is that the class variables' ordering is decided at random, and this has effects on the classification rates [15], [44].

CCC consists of q base binary classifiers (in our case, the FSNBCs, one for each affective state) linked circularly in a chain, creating a ring architecture (see Fig. 6). As in CC, each classifier at succeeding positions $2, 3, \ldots, q$ aggregates as inputs the predicted class labels of its previous classifiers. The circular configuration is generated after the first "cycle" or iteration when the predicted class labels of the classifiers at positions $2, \ldots, q$ are entered as additional features to the first one in the chain. The propagation of the predicted class labels continues to the succeeding classifiers $(2, 3, \ldots, q)$, and this mechanism is repeated for N iterations or until convergence.

5 Experimental Results

Three experiments were carried out to evaluate the performance of CCC using the FSNBC as the base classifier:

Experiment 1: Evaluate the convergence of CCC for the data of each patient.

Experiment 2: Evaluate the CCC performance and compare it against the performance of the multi-label classifiers: Binary Relevance (BR) and Classifier Chains (CC) [15].

Experiment 3: Analyze whether the affective states ordering within the CCC could generate different results in the multi-label metrics.

Additionally, the conditional probability tables (CPTs) created automatically by CCC using the FSNBC (they captured the relationships automatically), in the second experiment, were analyzed for determining the dependency relationships between the states.

5.1 Experimental Setups

CCC models were independently trained for each patient to predict the occurrence of the four states (T–tiredness, A–anxiety, P–pain, and E–engagement) in the multi-label classification scheme. This led to 8 CCC models, one for each patient. Each CCC involved the development of 4 FSNBCs, one for each state of the patient. Similarly, BR and CC models were independently developed for each patient, leading to 8 BR and 8 CC models, with their corresponding 4 FSNBCs for each one. Therefore, the three multi-label classifiers BR, CC, and CCC, were implemented using the FSNBC as the base classifier for all of them.

The performance of CCC was evaluated against the baselines BR and CC, using several multi-label classification metrics [45]: Global accuracy ($GAcc$), Mean accuracy ($MAcc$), Multi-label accuracy ($MLAcc$) and $F1$.

The notation to describe the metrics for multi-label classification is as follows:

r: number of samples in the data set;

q: number of classes. Each sample u has q class variables (each class variable takes one label: 1 –presence– or -1 –absence–);

$c_{u,j}$: jth true class label in sample u;

$c'_{u,j}$: jth label predicted by the multi-label classifier for sample u;

$\vec{c}_u$: vector of true class labels in sample u;

$\vec{c}'_u$: vector of class labels predicted by the multi-label classifier for sample u.

$$GAcc = \frac{1}{r}\sum_{u=1}^{r}\bigwedge_{j=1}^{q}(c'_{u,j} = c_{u,j}) \tag{2}$$

where $\bigwedge_{j=1}^{q}$ is the logical AND operator.

$$MAcc = \frac{1}{q}\sum_{j=1}^{q} Acc_j = \frac{1}{q}\sum_{j=1}^{q}\frac{1}{r}\sum_{u=1}^{r}\delta(c'_{u,j}, c_{u,j}) \tag{3}$$

where Acc_j is the calculation of accuracy for the class j and $\delta(c'_{u,j}, c_{u,j}) = 1$ if $c'_{u,j} = c_{u,j}$ and 0 otherwise.

$$MLAcc = \frac{1}{r}\sum_{u=1}^{r}\frac{|\vec{c}'_u \cap \vec{c}_u|}{|\vec{c}'_u \cup \vec{c}_u|} \tag{4}$$

where in $|\vec{c}'_u \cap \vec{c}_u|$ we count the number of coincidences of the two vectors (predicted and true) and in $|\vec{c}'_u \cup \vec{c}_u|$ we count the number of labels covered by those vectors.

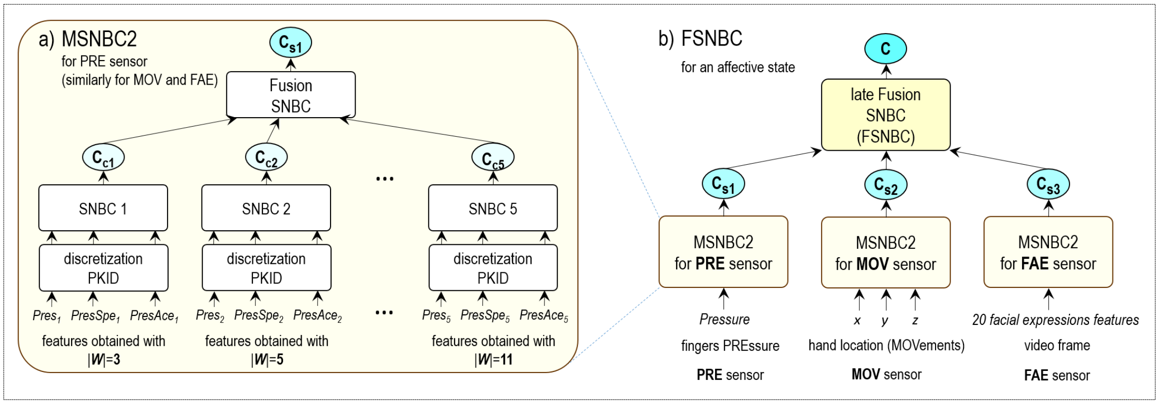


Fig. 5. Multiresolution Semi-Naive Bayesian classifier 2 (MSNBC2) and late Fusion using SNBC (FSNBC). a) MSNBC2 is a binary classifier that combines a set of parallel sliding windows W of different odd sizes, $|W|$ = 3,5,7,9,11; all concurrently centred around the same frame of the respective sensor. PKID is a discretization method called Proportional k-interval discretization [42] to handle the numeric features. b) FSNBC is a binary classifier which contains a MSNBC2 for each sensor (PRE, MOV, and FAE) and makes a late fusion using SNBC. FSNBC is the multimodal affective states recognizer for an affective state. Acronyms meanings: $C,\ C_{ck},\ C_{sj}$ = class of the respective classifier for the same affective state, e.g. anxiety; $Pres$ = *pressure*, $PresSpe$ = *pressure speed* and $PresAce$ = *pressure acceleration*.

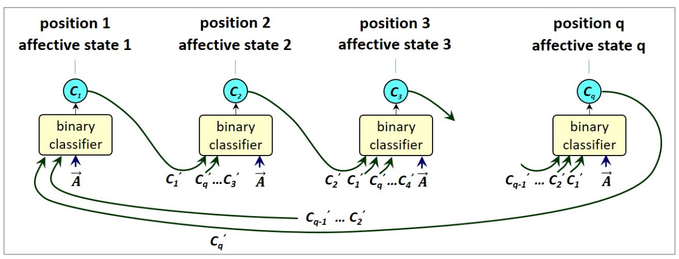


Fig. 6. Circular Classifier Chain (CCC). At the first iteration, the predicted class labels $C_j^{\prime}, j \in \{1, 2, \ldots, q-1\}$ are propagated as classifier chains (CC). $\vec{A}$ is the feature vector (in our case, the feature vectors of PRE, MOV, and FAE). Then, for the second iteration, the classifier at position 1 receives the predicted class labels from the last classifier (the one at position q) and the other classifiers (positions $2, 3, ..., q-1$). After that, the propagation process continues to the succeeding classifiers in the chain. The process is repeated until convergence or until CCC reached a maximum number N of iterations.

$$F1 = \frac{1}{r} \sum_{u=1}^{r} \frac{2\,|\vec{c}_u^{\,\prime} \cap \vec{c}_u|}{|\vec{c}_u^{\,\prime}| + |\vec{c}_u|} \tag{5}$$

The internal validity of the BR, CC, and CCC models was established using the stratified ten fold cross-validation mechanism across all the rehabilitation sessions.

The class variables' ordering for CC, and initially for CCC, was defined considering the BR results of the area under the curve (AUC) of the class variables. They were sorted in decreasing order according to these AUC results, with the rationale that the class variables with worse outcomes should be at the last positions so they could receive more information from the class variables of the preceding positions. The ordering for each patient is shown in Table 2.

TABLE 2
CLASS VARIABLES' ORDERING

Patient	Decreasing order according to BR results
P01, P02	tiredness $\geq$ anxiety $\geq$ engagement $\geq$ pain
P03, P11	tiredness $\geq$ anxiety $\geq$ pain $\geq$ engagement
P04	pain $>$ anxiety $>$ tiredness $>$ engagement
P05	tiredness $>$ engagement $>$ anxiety $>$ pain
P07, P08	pain $>$ anxiety $>$ engagement $>$ tiredness

5.2 Experiment 1: Convergence of CCC

CCC was run with $N = 8$ iterations. This number was determined based on previous work [19]. The convergence of CCC was analyzed, observing the results of the four multi-label classification metrics at each iteration of CCC. The behaviour of the system was similar for all the patients, CCC converged to a fixed value. For patients P1, P2, P3, and P4, the convergence was achieved at iteration 2. For patients P5 and P8, at iteration 3. And, for patients P7 and P11, at iterations 4 and 5, respectively. Fig. 7 shows the convergence process for patient P11.

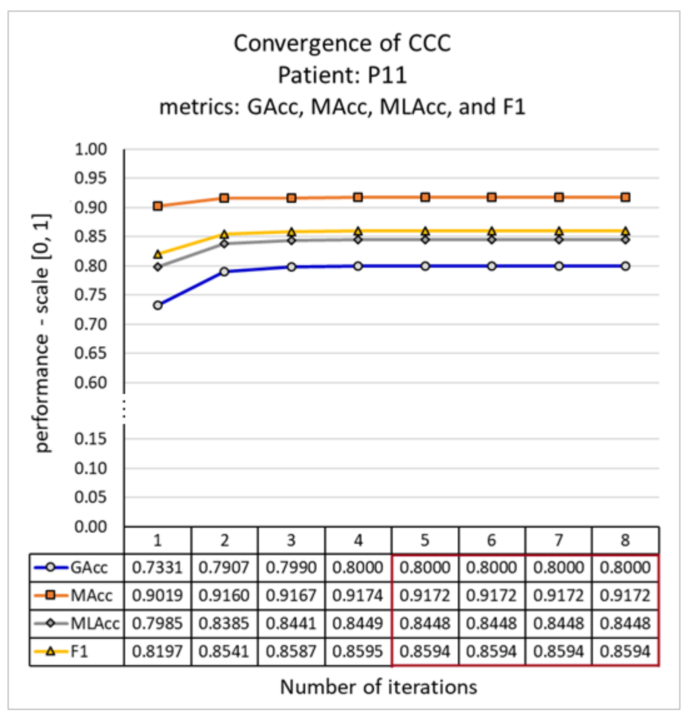


Fig. 7. Convergence of CCC - Patient P11. The number of iterations of CCC was set to $N = 8$. The four metrics, $GAcc$, $MAcc$, $MLAcc$, and $F1$, reached a fixed value at the fifth iteration. Therefore, CCC converged to a fixed value at iteration fifth for patient P11.

5.3 Experiment 2: Performance Comparison of BR, CC, and CCC

Table 3 summarizes the classification results, $mean \pm std.\,deviation$ (across the 8 patients and across the 10 folds of the cross-validation), of BR, CC, and CCC. The experiments were done with a laptop Intel Core i7-8750H CPU, 8th Gen. 2.20 GHz, 16 GB RAM with an operating system Windows 64 bits. CCC using the FSNBC during training (with 10 fold cross-validation) had an execution time of ($mean \pm std.\,deviation$) 124.09 ± 30.46 sec. which corresponds to 2.07 ± 0.51 min. The number of data samples was ($mean \pm std.\,deviation$) 7225.75 ± 3763.03 across the eight patients. The best results for each metric are highlighted in bold type. Results reveal that CCC outperformed BR and CC for all the metrics. Significant differences (Friedman test, for $GAcc$, $MAcc$, $MLAcc$, $F1$ $\chi^2(2) = 20.000, p < 0.05$, with post hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction, $p < 0.017$) were obtained for CCC.

TABLE 3
Performance comparisons between BR, CC, and CCC ($mean \pm std.\,deviation$) across the 8 patients and the 10 folds of the cross-validation. CCC was run with $N = 8$ iterations, and the convergence had been achieved by iteration 5 at most for all the patients. The best results for each metric are highlighted in bold type.

Classifier	$GAcc$	$MAcc$	$MLAcc$	$F1$
BR	0.813 ± 0.124	0.942 ± 0.039	0.859 ± 0.094	0.874 ± 0.086
CC	0.860 ± 0.098	0.952 ± 0.034	0.892 ± 0.075	0.903 ± 0.069
CCC	**0.905 ± 0.069** ‡	**0.962 ± 0.028** ‡	**0.927 ± 0.054** ‡	**0.933 ± 0.050** ‡

‡ means significant differences between CCC and CC, and between CCC and BR (Friedman test, for $GAcc$, $MAcc$, $MLAcc$, $F1$ $\chi^2(2) = 20.000, p < 0.05$, post hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction, $p < 0.017$).

Table 4 shows the average AUC results for each state across the eight patients and across the ten folds. Average AUC results of CCC were significantly higher than the ones of BR and CC (Friedman test, for tiredness $\chi^2(2) = 20.000, p < 0.05$, for anxiety $\chi^2(2) = 20.000, p < 0.05$, for pain $\chi^2(2) = 17.590, p < 0.05$, and for engagement $\chi^2(2) = 18.200, p < 0.05$, with post hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction, $p < 0.017$).

TABLE 4
Performance comparisons between BR, CC, and CCC ($mean \pm std.\,deviation$) across the 8 patients and the 10 folds of the cross-validation, using ROC AUC for each state. CCC was run with $N = 8$ iterations, and the convergence had been achieved by iteration 5 at most for all the patients. The best results for each state are highlighted in bold type.

Classifier	Tiredness	Anxiety	Pain	Engagement
BR	0.927 ± 0.059	0.928 ± 0.054	0.920 ± 0.058	0.903 ± 0.076
CC	0.946 ± 0.047	0.939 ± 0.048	0.927 ± 0.048	0.924 ± 0.065
CCC	**0.957 ± 0.040** ‡	**0.957 ± 0.036** ‡	**0.940 ± 0.045** ‡	**0.940 ± 0.058** ‡

‡ means significant differences between CCC and CC, and between CCC and BR (Friedman test, for tiredness $\chi^2(2) = 20.000, p < 0.05$, for anxiety $\chi^2(2) = 20.000, p < 0.05$, for pain $\chi^2(2) = 17.590, p < 0.05$, and for engagement $\chi^2(2) = 18.200, p < 0.05$, post hoc analysis with Wilcoxon signed-rank tests with Bonferroni correction, $p < 0.017$).

5.4 Experiment 3: Evaluation of the Affective States Ordering within the CCC

To analyze whether the affective states ordering within the CCC can generate different results, CCC was assessed with all different permutations of the 4 states, and permutation tests [46] were applied. The assessment of CCC was made considering the 24 permutations of the 4 states on the data of each patient, so the analysis was carried out for each patient. For every permutation, CCC was run with $N = 8$ iterations, and the results of the multi-label classification metrics were observed at iteration 8, where CCC had converged. After the results of the 24 permutations were obtained, the permutation tests [46] were performed. There were no significant differences in the results of the multi-label metrics for patients P1, P2, P3, and P5, but for the rest of the patients, P4, P07, P08, and P11, indeed there were. It could be expected that the results for each metric were equal for all the permutations of a patient, but there were some differences. Therefore, we calculated how large the differences were. The average differences ($mean \pm std.\,deviation$) across all the patients, between the highest result minus the lowest result, within the permutations of each patient were $GAcc$: 0.106 ± 0.069, $MAcc$: 0.026 ± 0.017, $MLAcc$: 0.057 ± 0.037, and $F1$:0.040 ± 0.026. The main differences were generated by $GAcc$, which requires the exact match between the multi-label vectors that are compared, each time (see Eq. 2).

5.5 Dependency Relationships between the States

Since late fusion in the FSNBC for predicting the class labels of a certain state S, creates a SNBC model which receives as inputs the predicted class labels from PRE, MOV, and FAE for S, and the predicted class labels from the other states; it is possible to analyze the conditional probability tables (CPTs) of the other states given state S. CPTs can provide us information about the relationships of the states that SNBC considers to assess the presence (1) or the absence (-1) of the state S. These CPTs were generated after the eight iterations of CCC, from experiment 2. It should be noted that in this experiment, the affective states were sorted in decreasing order according to the AUC results obtained for each state in BR (see Subsection 5.1). Although, each patient has his/her own decreasing order of the AUC results for the four states as the initial order for CCC (Table 2), the results related to the dependency relationships between the affective states, represented in CPTs, were similar for all the patients. Fig. 8 shows the CPTs of each state. The information is organized in blocks a), b), c), and d) corresponding to the CPTs of tiredness, anxiety, pain, and engagement, respectively. Each table entry indicates $mean \pm std.\,deviation$ across the eight patients. From the analysis of the tables, we can detect the following relationships of the states for the eight patients:

1) **Tiredness**: When tiredness is present, engagement is not, nor anxiety or pain.
2) **Anxiety**: When there is anxiety, there is no engagement, and there is no tiredness either. But when anxiety is not present, pain is not present either.
3) **Pain**: When there is pain, there is anxiety too; but there is neither engagement nor tiredness.
4) **Engagement**: When engagement is present, none of the other states (tiredness, anxiety, and pain) is present.

This evidence establishes that there is a mutual exclusion between engagement and all the other states and co-occurrences between pain and anxiety for the eight patients during their rehabilitation sessions. Such results are in line with the literature on chronic pain, suggesting that anxiety mediates pain rather than vice versa [7]. The results of the dependency relationships were similar for all the patients, no matter the initial order of the four states for CCC. The dependency relationships between the affective states were automatically learned by CCC using the FSNBC as the base classifier, and then they were used for the classification process.

a)

$P(A\|T) =$		Tiredness: No (-1)	Tiredness: Yes (1)
Anxiety	-1	0.397±0.093	**0.977**±0.029
	1	0.603±0.093	0.023±0.029

$P(P\|T) =$		Tiredness: No (-1)	Tiredness: Yes (1)
Pain	-1	0.672±0.091	**0.985**±0.029
	1	0.328±0.091	0.015±0.029

$P(E\|T) =$		Tiredness: No (-1)	Tiredness: Yes (1)
Engagement	-1	0.590±0.088	**0.985**±0.011
	1	0.410±0.088	0.015±0.011

b)

$P(T\|A) =$		Anxiety: No (-1)	Anxiety: Yes (1)
Tiredness	-1	0.567±0.145	**0.987**±0.014
	1	0.433±0.145	0.013±0.014

$P(P\|A) =$		Anxiety: No (-1)	Anxiety: Yes (1)
Pain	-1	**1.000**±0.000	0.499±0.098
	1	0.000±0.000	0.501±0.098

$P(E\|A) =$		Anxiety: No (-1)	Anxiety: Yes (1)
Engagement	-1	0.425±0.140	**0.981**±0.026
	1	0.575±0.140	0.019±0.026

c)

$P(T\|P) =$		Pain: No (-1)	Pain: Yes (1)
Tiredness	-1	0.689±0.119	**0.987**±0.025
	1	0.311±0.119	0.013±0.025

$P(A\|P) =$		Pain: No (-1)	Pain: Yes (1)
Anxiety	-1	0.675±0.026	0.000±0.000
	1	0.325±0.026	**1.000**±0.000

$P(E\|P) =$		Pain: No (-1)	Pain: Yes (1)
Engagement	-1	0.613±0.085	**0.989**±0.016
	1	0.387±0.085	0.011±0.016

d)

$P(T\|E) =$		Engagement: No (-1)	Engagement: Yes (1)
Tiredness	-1	0.664±0.108	**0.985**±0.012
	1	0.336±0.108	0.015±0.012

$P(A\|E) =$		Engagement: No (-1)	Engagement: Yes (1)
Anxiety	-1	0.328±0.105	**0.968**±0.048
	1	0.672±0.105	0.032±0.048

$P(P\|E) =$		Engagement: No (-1)	Engagement: Yes (1)
Pain	-1	0.653±0.044	**0.989**±0.016
	1	0.347±0.044	0.011±0.016

Fig. 8. Conditional Probability Tables (CPTs) generated by SNB models at the late fusion of the FSNBC, after the eight iterations of CCC. T = Tiredness, A = Anxiety, P = Pain, and E = Engagement. Each CPT entry represents $mean \pm std.\, deviation$ across the eight patients. The highest value for each CPT is highlighted in bold type. Part a) corresponds to tiredness CPTs, b) to anxiety CPTs, c) to pain CPTs, and d) to engagement CPTs.

6 Discussion

The Circular Classifier Chain (CCC) in conjunction with the Fusion using a Semi-Naïve Bayesian classifier (FSNBC), can automatically extract information about the mutual exclusion and co-occurrences between some affective, physical, and/or cognitive states that are involved in the rehabilitation of post-stroke patients. The model sought to leverage the automatic recognition of the post-stroke patients' states, tiredness, anxiety, pain, and engagement, by considering the dependency relationships between these states. In this extended version of the conference paper [11], we expanded the dataset from five to eleven post-stroke patients with ten rehabilitation sessions, collected at the rehabilitation centre of the hospital, to each of them. All the new data were labelled frame-by-frame over the four states by psychiatrists. Unfortunately, for three of the new patients, only finger pressure and hand movements tracking data are available. While the data from these three patients are still very valuable for the research community to explore these two modalities as in [14], only the patients with data from the full set of modalities were used in this paper. The results presented in the conference paper [11] with five patients were confirmed on the data collected from the three new post-stroke patients with full multimodal data (Table 1).

CCC using the FSNBC identified mutual exclusion between engagement and all the other states and identified co-occurrence of pain and anxiety for the eight patients during the rehabilitation sessions. Such results are in line with the literature on chronic pain, suggesting that anxiety mediates pain rather than vice versa [7]. The mutual exclusion of engagement with tiredness, pain, and anxiety is naturally due to the barriers that the latter three pose to the first. It could also be that up to a certain level of physical and psychological demand; engagement may distract from such unpleasant states. As it has been shown in other studies, engaging games appears to provide better coping-capabilities within a limited range of fear and physical demand [6].

CCC using the FSNBC allows the prediction of the patient's considered states since all the patients' states in our dataset were recognized with results at least of 0.905 ± 0.069 ($mean \pm std.\, deviation$) on the multi-label metrics. The classification rates were at least 0.940 ± 0.045 ($mean \pm std.\, deviation$) when using the area under the curve for each of the four states. When the dependency relationships between tiredness, anxiety, pain, and engagement were not taken into account in BR, $GAcc$, for instance, obtained 0.813 ± 0.124 ($mean \pm std.\, deviation$) (Table 3).

Additionally, we have obtained a late fusion process in the FSNBC (for predicting the class labels of an affective, physical, and/or cognitive state) that not only takes into account information from each modality (i.e., finger pressure, hand movements, and facial expressions) but also includes correlations with the predicted class labels of the FSNBCs of the other states. In this way, the fusion process not only analyzes the modalities involved in the recognition but also learns the correlation between the states.

An advantage of using a Bayesian approach in FSNBC was the possibility of explicability and interpretability of the results generated in the late fusion. This characteristic was useful for determining which dependency relationships between the states were automatically detected by CCC using the FSNBC (Subsection 5.5). Further developments and evaluations can be done concerning the performance of CCC using the FSNBC against other multi-label classifiers, especially classifiers associated with neural networks and deep learning. However, the results of CCC using the FSNBC were in the order of 0.905 ± 0.069 ($mean \pm std.\, deviation$).

It should be noted that CCC using the FSNBC converged in 5 iterations at most for all the patients. Therefore, the system does not require an extensive process of iterations for getting the convergence. This tendency was confirmed with the data of the new patients incorporated into the dataset.

As an extension of [11], we carried out an analysis of the effect of the affective states ordering within the CCC. The results revealed that the order might have a small average effect on the multi-label metrics of $MAcc$, $MLAcc$, and $F1$, in the order of 0.05 ± 0.03 ($mean \pm std.\, deviation$), and for

$GAcc$, in the order of 0.106 ± 0.069 ($mean \pm std.\,deviation$) for some patients. In this case, 4 states were involved, and all the permutations could be evaluated. Further studies have to be carried out to determine why the order of the affective states in the chain generated different results in some patients and not in others. They could be due not only to physical differences but also to the ability of the person with their condition. In the literature on chronic pain, personal differences exist in people's ability to manage their condition. People suffering from strong anxiety do present less capability to engage in physical activity.

7 Conclusions

This work extended our research presented at the International Conference of Affective Computing and Intelligent Interaction (ACII 2019) [11]. In this extended version, we included experimental validation by incrementing the number of post-stroke patients in the longitudinal study (from five to eleven patients, although data of 8 patients were used in the experiments); we analyzed the convergence of the proposed multi-label classifier, Circular Classifier Chain (CCC); and we evaluated the impact of the affective states ordering within the CCC.

The relationships of mutual exclusion and co-occurrences between tiredness, anxiety, pain, and engagement, in the rehabilitation of some post-stroke patients, were studied using the proposal of two classifiers, the multi-label classifier Circular Classifier Chain (CCC), combined with the multimodal classifier Fusion using a Semi-Naïve Bayesian classifier (FSNBC) used as the base classifier of CCC. The synergy between these two classifiers boosted the automatic recognition of the patients' mentioned states by considering the dependency relationships between the states. CCC using the FSNBC is simple, efficient, and capable of addressing the dependency relationships between states. The dependency relationships between the states were automatically learned by CCC using the FSNBC, which converged at 5 iterations at most for all the patients.

CCC using the FSNBC provided a scheme (through the CPTs created by the classifiers) for the automatic detection of the dependencies between the affective states involved in an affective computing system. This scheme can be useful for several applications in the area. CCC using the FSNBC detected the relationships of mutual exclusion between engagement and all the other states, and co-occurrences between pain and anxiety for the eight patients during their rehabilitation sessions. Moreover, CCC using the FSNBC enhanced the automatic recognition of the states in a multi-label classification approach, outperforming CC and BR significantly (both CC and BR using the FSNBC as the base classifier too).

A particular purpose of this study was the assessment of CCC as an extension of CC for dealing with the problem of the class variables' ordering in CC. Comparisons can be made with other multi-label classifiers (including, for instance, some concerning neural networks), although an advantage of the Bayesian approach is the possibility of explicability and interpretability of the results.

The incorporation of this multi-label classifier combined with the multimodal classifier in virtual rehabilitation platforms for post-stroke patients could leverage intelligent and empathic interactions, as well as real-time personalization of exercise plans to promote adherence to rehabilitation exercises.

As future work, the problem of determining why the order of the affective states in the chain generated different results in some patients and not in others has to be analyzed in depth. Another problem to be addressed consists in dealing with missing sensors since this problem is common in the naturalistic everyday use of computational models.

Acknowledgments

This project was supported by a grant from the Secretaría de Educación, Ciencia, Tecnología e Innovación de la Ciudad de México. SECITI/042/2018 (INGER-DI-CRECITES-001-2018) "Red Colaborativa de Investigación Traslacional para el Envejecimiento Saludable de la Ciudad de México" (RECITES), and a grant from the project FORDECyT 296737 "Consorcio en Inteligencia Artificial". We also acknowledge the support for the Scholarship No. 434867 from the Mexican Research Council CONACYT. Jesús Joel Rivas is also grateful to God (Father, Son, and Holy Spirit) for giving him inspiration and encouragement during this project.

References

[1] L. E. Sucar, R. Leder, J. Hernández, I. Sánchez, and G. Azcárate, "Clinical evaluation of a low-cost alternative for stroke rehabilitation," in *Proceedings of the 11th International Conference on Rehabilitation Robotics (ICORR 2009)*. IEEE, 2009, pp. 863–866.

[2] 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," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 634–643, 2014.

[3] S. M. Ávila Sansores, "Adaptación en línea de una política de decisión utilizando aprendizaje por refuerzo y su aplicación en rehabilitación virtual," Master's thesis, Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, México., 2013.

[4] L. E. Sucar, S. M. Ávila-Sansores, and F. Orihuela-Espina, "User modelling for patient tailored virtual rehabilitation," in *Foundations of Biomedical Knowledge Representation*. Springer, 2015, pp. 259–278.

[5] T. Olugbade, A. Singh, N. Bianchi-Berthouze, N. Marquardt, M. Aung, and A. Williams, "How can affect be detected and represented in technological support for physical rehabilitation?" *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 1, 2019.

[6] A. Singh, N. Bianchi-Berthouze, and A. C. Williams, "Supporting everyday function in chronic pain using wearable technology," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 3903–3915.

[7] T. Olugbade, A. Williams, and N. Bianchi-Berthouze, "The relationship between guarding, pain, and emotion." *PAIN Report*, vol. In press, 2019.

[8] K. Trohidis, G. Tsoumakas, G. Kalliris, and I. P. Vlahavas, "Multi-label classification of music into emotions." in *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)*, vol. 8, 2008, pp. 325–330.

[9] S. Wang, Z. Wang, and Q. Ji, "Multiple emotional tagging of multimedia data by exploiting dependencies among emotions," *Multimedia Tools and Applications*, vol. 74, no. 6, pp. 1863–1883, 2015.

[10] S. Wang, J. Wang, Z. Wang, and Q. Ji, "Multiple emotion tagging for multimedia data by exploiting high-order dependencies among emotions," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2185–2197, 2015.

[11] J. J. Rivas, F. Orihuela-Espina, L. E. Sucar, A. Williams, and N. Bianchi-Berthouze, "Automatic recognition of multiple affective states in virtual rehabilitation by exploiting the dependency relationships," in *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII 2019)*. IEEE, 2019, pp. 1–7.

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

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

[14] J. J. Rivas, F. Orihuela-Espina, L. Palafox, N. Berthouze, M. del Carmen Lara, J. Hernández-Franco, and E. Sucar, "Unobtrusive inference of affective states in virtual rehabilitation from upper limb motions: A feasibility study," *IEEE Transactions on Affective Computing*, 2018.

[15] J. Read, B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification," in *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2009, pp. 254–269.

[16] P. Philippot, "Inducing and assessing differentiated emotion-feeling states in the laboratory," *Cognition and emotion*, vol. 7, no. 2, pp. 171–193, 1993.

[17] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition & emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[18] Z. Wang, S. Wang, M. He, Z. Liu, and Q. Ji, "Emotional tagging of videos by exploring multiple emotions' coexistence," in *Proceedings of the 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG 2013)*. IEEE, 2013, pp. 1–6.

[19] J. J. Rivas, F. Orihuela-Espina, and L. E. Succar, "Circular chain classifiers," in *Proceedings of Machine Learning Research. 9th International Conference on Probabilistic Graphical Models*, vol. 72, 2018, pp. 380–391.

[20] S. K. D'mello and J. Kory, "A review and meta-analysis of multimodal affect detection systems," *ACM Computing Surveys (CSUR)*, vol. 47, no. 3, pp. 43:1–43:36, 2015.

[21] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98–125, 2017.

[22] Y. Gao, N. Bianchi-Berthouze, and H. Meng, "What does touch tell us about emotions in touchscreen-based gameplay?" *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 19, no. 4, pp. 1–30, 2012.

[23] P. P. Filntisis, N. Efthymiou, P. Koutras, G. Potamianos, and P. Maragos, "Fusing body posture with facial expressions for joint recognition of affect in child–robot interaction," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4011–4018, 2019.

[24] T. Balomenos, A. Raouzaiou, S. Ioannou, A. Drosopoulos, K. Karpouzis, and S. Kollias, "Emotion analysis in man-machine interaction systems," in *Proceedings of the International Workshop on Machine Learning for Multimodal Interaction*. Springer, 2004, pp. 318–328.

[25] H. Gunes and M. Piccardi, "Bi-modal emotion recognition from expressive face and body gestures," *Journal of Network and Computer Applications*, vol. 30, no. 4, pp. 1334–1345, 2007.

[26] ——, "Automatic temporal segment detection and affect recognition from face and body display," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 1, pp. 64–84, 2009.

[27] M. Gavrilescu, "Recognizing emotions from videos by studying facial expressions, body postures and hand gestures," in *2015 23rd Telecommunications Forum Telfor (TELFOR)*. IEEE, 2015, pp. 720–723.

[28] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss, and R. Picard, "Automatic recognition methods supporting pain assessment: A survey," *IEEE Transactions on Affective Computing*, 2019.

[29] J. Egede, M. Valstar, and B. Martinez, "Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation," in *2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)*. IEEE, 2017, pp. 689–696.

[30] C. R. Jonassaint, N. Rao, A. Sciuto, G. E. Switzer, L. De Castro, G. J. Kato, J. C. Jonassaint, Z. Hammal, N. Shah, and A. Wasan, "Abstract animations for the communication and assessment of pain in adults: cross-sectional feasibility study," *Journal of medical Internet research*, vol. 20, no. 8, p. e10056, 2018.

[31] C. Wang, T. A. Olugbade, A. Mathur, A. C. De C. Williams, N. D. Lane, and N. Bianchi-Berthouze, "Recurrent network based automatic detection of chronic pain protective behavior using mocap and semg data," in *Proceedings of the 23rd International Symposium on Wearable Computers*, 2019, pp. 225–230.

[32] A. Bonarini, M. Garbarino, M. Matteucci, and S. Tognetti, "Affective evaluation of robotic rehabilitation of upper limbs in post-stroke subjects," in *Proceedings of the 1st Workshop on the Life Sciences at Politecnico di Milano*, 2010, pp. 290–293.

[33] P. Kan, R. Huq, J. Hoey, R. Goetschalckx, and A. Mihailidis, "The development of an adaptive upper-limb stroke rehabilitation robotic system," *Journal of Neuroengineering and Rehabilitation*, vol. 8, no. 1, pp. 1–18, 2011.

[34] 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 *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2015)*. IEEE, 2015, pp. 287–292.

[35] A. Bandini, S. Orlandi, H. J. Escalante, F. Giovannelli, M. Cincotta, C. A. Reyes-Garcia, P. Vanni, G. Zaccara, and C. Manfredi, "Analysis of facial expressions in parkinson's disease through video-based automatic methods," *Journal of neuroscience methods*, vol. 281, pp. 7–20, 2017.

[36] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 42–55, 2012.

[37] R. O. Duda, P. E. Hart *et al.*, *Pattern classification and scene analysis*. Wiley New York, 1973, vol. 3.

[38] M. Martínez-Arroyo and L. E. Sucar, "Learning an optimal naive Bayes classifier," in *Proceedings of the 18th International Conference on Pattern Recognition, 2006. (ICPR 2006)*, vol. 3. IEEE, 2006, pp. 1236–1239.

[39] M. Martínez Arrollo, "Aprendizaje de clasificadores bayesianos estáticos y dinámicos," Ph.D. dissertation, Instituto Tecnológico y de Estudios Superiores de Monterrey, México, 2007.

[40] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.

[41] 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), Puebla, México., 2015.

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

[43] J. J. Rivas, F. Orihuela-Espina, and L. E. Sucar, "Recognition of affective states in virtual rehabilitation using late fusion with Semi-Naive Bayesian classifier," in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2019)*. ACM, 2019, pp. 308–313.

[44] E. C. Goncalves, A. Plastino, and A. A. Freitas, "A genetic algorithm for optimizing the label ordering in multi-label classifier chains," in *Proceedings of the 25th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 2013, pp. 469–476.

[45] C. Bielza, G. Li, and P. Larrañaga, "Multi-dimensional classification with Bayesian networks," *International Journal of Approximate Reasoning*, vol. 52, no. 6, pp. 705–727, 2011.

[46] M. Ojala and G. C. Garriga, "Permutation tests for studying classifier performance," *Journal of Machine Learning Research*, vol. 11, no. Jun, pp. 1833–1863, 2010.

Jesús Joel Rivas received a Ph.D degree in Computer Science from the National Institute of Astrophysics, Optics, and Electronics (INAOE), Puebla, Mexico, and a M.Sc. degree in Computer Science too from the INAOE. He has made secondments at the University College of London, Interaction Centre, London, United Kingdom. He has been working in affective computing applied to stroke rehabilitation. His research interest includes affective computing, probabilistic graphical models, and machine learning.

María del Carmen Lara is a psychiatrist who received a Ph.D. in Medical Sciences from the National University of Mexico. She is a Professor-Investigator at the Benemérita Universidad Autónoma de Puebla. Her main field of interest is the measurement of clinical phenomena.

Luis Castrejón is a medical specialist in Rehabilitation Medicine, graduated from the Medical School of the Benemérita Universidad Autónoma de Puebla, Puebla, Mexico. He got a speciality in rehabilitation medicine at the National Institute of Rehabilitation Luis Guillermo Ibarra Ibarra, Mexico City, Mexico. He made a training stay in neurological rehabilitation of the adult and the child, at La Paz hospital, Madrid, Spain. He is currently the principal of the rehabilitation medicine service of the University Hospital of the Benemérita Universidad Autónoma de Puebla. His research interest includes the rehabilitation of cerebral vascular disease and Parkinson's disease.

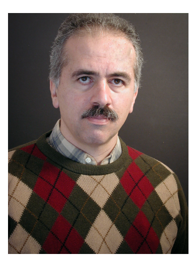
Jorge Hernández-Franco graduated as medical doctor from the Universidad Autónoma de México, Mexico City, Mexico, in 1985, got his speciality in rehabilitation medicine, in 1989, and was certified in neurological rehabilitation by the Newcastle University, Newcastle upon Tyne, U.K., in 1999. Since 1991, he is the Head of the rehabilitation ward at the National Neurology and Neurosurgery Institute MVS, Mexico City, Mexico, where he lectures on neurological rehabilitation. He has further lectured on physical therapy in neurological rehabilitation at the American British Cowdray Hospital since 1996. He is member of the editorial board of the journal *Developmental Neurorehabilitation* since 2005. Dr Hernández-Franco is the Vice-President for Mexico, Central America and the Caribbean of the World Federation for Neurologic Rehabilitation.

Felipe Orihuela-Espina received his Ph.D. degree from the University of Birmingham, Birmingham, U.K. He has been a Lecturer in the Autonomous University of the State of Mexico and a Postdoctoral Research Associate at Imperial College London and at the National Institute of Astrophysics, Optics, and Electronics (INAOE), Puebla, Mexico. He later joined INAOE as faculty, and he is currently a Reader at INAOE and a member of the National Research System. He has authored over 80 papers and carried out research stages and secondments in MGH/HST Martinos Center, UCL, and ETH Zurich, among others. His current research interests are in neuroimage understanding and interpretation.

Lorena Palafox is graduated as occupational therapist from the National Rehabilitation Institute, Mexico City, in 2009; and was certified in neurological rehabilitation by the Universidad Autónoma de Barcelona and the Guttmann Institute, Spain, in 2012. Since 2009, she has collaborated in investigations of different neurological damages, she is the only occupational therapist at the National Neurology and Neurosurgery Institute MVS, Mexico City. She has been certified in different rehabilitation techniques.

Amanda Williams is an academic and clinical psychologist at University College London, UK, and at the Pain Management Centre, University College London Hospitals. She also works as a research consultant for the International Centre for Health and Human Rights (ICHHR). She has been active in research and clinical work in pain for 30 years, with particular interests in evaluation of psychologically-based treatments for pain; in behavioural expression of pain and its interpretation by clinicians; in evolutionary understanding of pain and pain behaviour; and in pain from torture. Dr Williams has written over 200 papers and chapters, presents at national and international pain meetings, and is on the editorial boards of several major pain journals.

Nadia Bianchi-Berthouze is a Professor in Affective Computing and Interaction at the University College London, UK. She received the Laurea degree with honors in computer science in 1991 and the Ph.D. degree in science of biomedical images in 1996 from the University of Milano, Milano, Italy. Her research interests include the study of body movement, muscle activity and touch behaviour as ways to automatically recognize and steer the quality of experience of humans interacting and engaging with/through whole-body technology. She has been pioneering the analysis of affective body expressions in the context of physical rehabilitation. She was the Principal Investigator on an EPSRC funded project on Pain rehabilitation: E/Motion-based automated coaching (Emo-pain.ac.uk). She is now investigating wellbeing, movement and affect in a variety of real-life situations such as factory work, education and textile design.



Luis Enrique Sucar (Senior Member IEEE) has a Ph.D. in Computing from Imperial College, London, 1992; a M.Sc. in Electrical Engineering from Stanford University, USA, 1982; and a B.Sc. in Electronics and Communications Engineering from ITESM, Mexico, 1980. He is currently Senior Research Scientist at the National Institute for Astrophysics, Optics and Electronics, Puebla, Mexico. He has been an invited professor at the University of British Columbia, Canada; Imperial College, London; INRIA, France; and CREATE-NET, Italy. He has more than 300 publications and has directed 21 Ph.D. thesis. Dr. Sucar is Member of the National Research System and the Mexican Science Academy. He is associate editor of the Pattern Recognition journal, and has served as president of the Mexican AI Society and as member of the Advisory Board of IJCAI. In 2016 he received the National Science Prize from the Mexican government. His main research interest are in graphical models and probabilistic reasoning, and their applications in computer vision, robotics, energy, and biomedicine.