The AffectMove Challenge: some machine learning approaches

Gérard Dray, Pierre-Antoine Jean, Yann Maheu, Jacky Montmain, Nicolas Sutton-Charani

The AffectMove Challenge : some machine learning approaches

1st Gerard Dray
EuroMov Digital Health in Motion
Univ Montpellier, IMT Mines Ales
Alès, France
ORCID : 0000-0003-1525-5682

2nd Pierre-Antoine Jean
IMT Mines Ales
Alès, France
firstname.lastname@mines-ales.fr

3rd Yann Maheu
IMT Mines Ales
Alès, France
firstname.lastname@mines-ales.org

4th Jacky Montmain
EuroMov Digital Health in Motion
Univ Montpellier, IMT Mines Ales
Alès, France
ORCID : 0000-0003-0918-5788

5th Nicolas Sutton-Charani
EuroMov Digital Health in Motion
Univ Montpellier, IMT Mines Ales
Alès, France
ORCID : 0000-0002-3065-0712

***Abstract*—This paper describes some machine learning methods that we have implemented to participate in the AffectMove challenge which aims to develop technologies for classification of body movements in the areas of physical rehabilitation of chronic pain, mathematical problem solving and interactive dance contexts. The methods and results obtained are presented as well as some futureworks.**

***Index Terms*—Body movements signals processing, Features extraction, Supervised machine learning methods**

I. Introduction

As the organizers of the AffectMove Challenge describe it [1] "The AffectMove challenge is based on 3 naturalistic datasets on body movement, which are fundamental components of everyday living both in the execution of the actions that make up physical functioning as well as in rich expression of affect, cognition, and intent [2] [3]. The datasets were built according to automatic detection requirements for chronic pain physical rehabilitation, maths problem solving, and interactive dance contexts respectively." In this paper, we only deal with the first two sets of data, with reference to the two challenge tasks: Task 1: "Protective Behaviour Detection based on Multimodal Body Movement Data" and Task 2: "Detection of Reflective Thinking based on Body Movement Data". We chose to analyze the data as a two-class supervised classification problem. For each task of the challenge we present the data, their preprocessing (including features extraction), then the considered machine learning methodsand the obtained results. We end with a conclusion and directions for future work.

II. Task 1: Protective Behaviour Detection based on Multimodal Body Movement Data

The aim of this task is to advance continuous detection of protective behaviours, i.e., bodily-expressed pain behaviours, in people with chronic musculoskeletal pain. As described in [1] "Chronic pain is a major healthcare challenge and technology that would be able to assess pain behaviour could support the delivery of personalised therapies in the long-term and the self-directed management of the condition with the aim of improving engagement in valued everyday activities [4]. The participants in this task had to construct a model for classifying continuous protective behaviors, present or absent throughout exercise of a person with chronic pain, based on the position of the joints of the whole body and back muscle activity. Ground truth for the type of exercise is also available, but it is not used as an input."

A. Data description

The anonymised 3D full-body joint positions and concomitant back muscle activity data for 19 people with chronic low back pain from the EmoPain dataset were provided [5]. The data were given with corresponding protective behaviour labels obtained from clinician observers. The data were given in training, validation, and test partitions which contain instances from 10, 4, and 5 people with chronic pain respectively. The test partition did not include the protective behaviour labels. AffectMove task 1 dataset is composed of 9 exercises (stand on one leg, sit still, reach forward, sit-to-stand, stand-to-sit, stand still, bend, walk and a other movement events) with the body movement and muscle activity measures. Each repetition of an exercise by a person is a 180-frame segment (3 seconds) and the label (label 0 where there is no protective behavior, 1 otherwise) for each segment is based on continuous labelling provided by expert raters. Movement data is the 3D full-body joint positions where the skeleton is represented by 17 joints and muscle activity data (electromyography) were collected for 4 muscle groups (cf. Fig. 1).

B. Exploring the classifiers space

The first submission relies on a combination of motion capture (MoCap) pre-processing data and a classification ap-

978-1-6654-0021-3/21/$31.00 ©2021 IEEE

proach. Three main steps constitute the whole pipeline (cf. Fig. 2):

- Posture Features Extraction. The coordinates of the skeleton joints are used to compute the feature vectors which represent human postures.
- Benchmarking classification approaches. Several classification methods have been evaluated and Random Forest was selected as the most efficient one.
- Fine-tuning of the selected model. A hyper-parameters optimization step (number of trees, maximum depth of tree, etc.) has been performed thanks to a grid search.

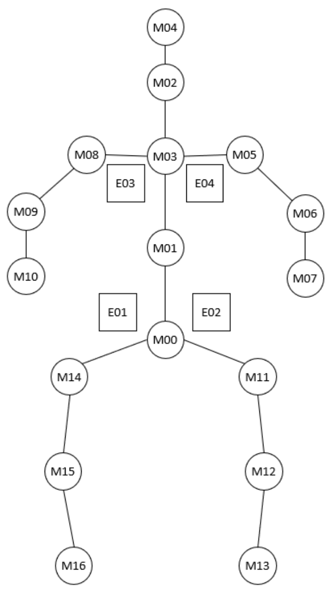

Fig. 1. Skeleton represented by 17 joints (M - -) and 4 EMG sensors (E - -).

The first step is the **data processing** where only MoCap data is used. The process exploits spatial features computed from 3D skeleton coordinates, without including the time information in the computation, in order to make the system independent of the speed of movement. For each skeleton frame, a posture feature vector is computed. Each joint is represented by J_i, a three-dimensional vector. The person can be found at any place within the coverage area of the MoCap capture, and the coordinates of the same joint may assume different values. It is necessary to compensate this effect, by using a proper features computation process. A straightforward solution, proposed by [6], is to compensate the position of the skeleton by centering the coordinate space in one skeleton joint. Considering the skeleton composed by 17 joints, J_0 being the hips center and J_3 the shoulders center (cf. Fig. 1), the *i*th joint feature d_i is the distance vector between J_i and J_0, normalized by the distance between J_3 and J_0.

$$d_i = \frac{J_i - J_0}{||J_3 - J_0||}, i = 1, 2, ..., 16 \tag{1}$$

For a repetition, represented by 180 frames, the length of the feature vector to characterize it is, $16 \times 180 = 2880$.

The second step consists of a **classifier selection** as exposed in [7]. As multiple classifier types are available, selecting the right classifier according to its performance is a crucial task. Six kinds of classifiers from the Scikit-learn API [8] have been explored:

- SVM (*Support Vector Machines*) with various kernels (linear, poly and rbf)
- Random Forest
- Ridge Regression
- k-Nearest Neighbors
- Multi-Layer Perceptron
- Passive Aggressive classifier

Classifiers have been trained with their default parameters (according to Scikit-learn). Random Forest [9] performs best and will be used in the next process.

The last step is a **classifier optimization** following a grid search on training data with a 3-fold cross validation based on accuracy performances. For a Random Forest Classifier, there are several different hyperparameters that can be adjusted. Four main parameters have been investigate: the number of trees in the forest, the maximum depth of each tree, the minimum number of samples required to split leaf node and the minimum number of samples required to be at a leaf node. The resulting best hyperparameters are as follows: *RandomForestClassifier(n_estimators=2000, max_depth=85, min_samples_leaf=1, min_samples_split=2).*

C. Metrics, scores and conclusion for the task 1

Four metrics were used by organizers of the challenge, F1-score for both classes, Matthews correlation coefficient (MCC) and Accuracy. Table I shows scores for the first submission of this article.

TABLE I
TASK 1 : SUBMISSION SCORES

F1-score (class 0)	F1-score (class 1)	Accuracy	MCC
0.89	0.30	0.81	0.23

III. TASK 2: DETECTION OF REFLECTIVE THINKING BASED ON BODY MOVEMENT DATA

As described in [1] "The purpose of this task is to continuously detect reflective thinking in children during maths problem-solving activities. Understanding mathematical ideas such as angles and shapes is a key part of basic education and digital learning technology that promotes the use of body movement as well as further recognizes critical learning moments (e.g., reflective thinking) could support learning of abstract mathematical ideas which may otherwise be challenging to relate to. The participants in this task had to build a model for classification of reflective thinking continuously as 'observed' or 'not observed' while a child solves maths problems, based on joint positions."

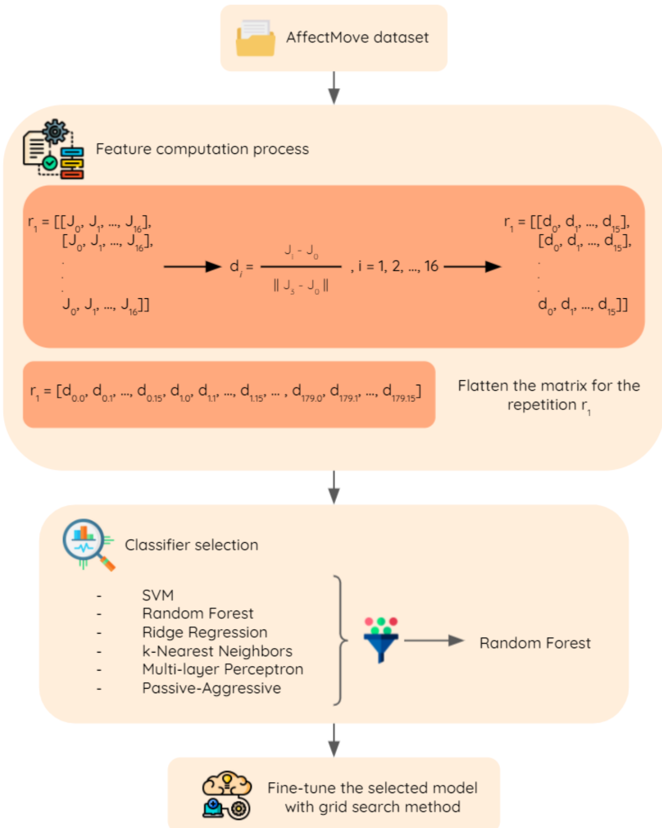

Fig. 2. Pipeline overview. Flatten operation allows to obtain the feature vector for one repetition.

A. Data description

The 3D full-body joint positions for 24 children from the weDraw-1 Movement dataset [10] were provided. The data is accompanied with corresponding reflective thinking labels based on expert observer annotation. Labels for the corresponding math problem-solving activities have also been included. The data was given in training, validation, and test partitions which contain instances from 13, 5, and 6 children respectively. The test partition did not include the reflective thinking labels or the activity type.

B. Method and results

As we will be using a cross-validation method, we have chosen to group together the training and validation sets. Files of less than 4.5 seconds have been eliminated, the classification rule for these recordings will be to classify them as "0" for "reflective thinking not observed". Each trial is then resampled at 25Hz and 4 seconds (from the 5th to 104th sample) are retained. At the end of this processing, our training set consists of 2849 trials: 2458 of class "0" for "reflective thinking not observed" and 391 of class "1" for "reflective thinking observed" ; Each trial consisting of 51 lines (3D position for 17 joints) of 100 samples. As we can see, our training set is strongly unbalanced. To overcome this problem, we applied the ADASYN (Adaptive Synthetic Sampling) method which is fully described in [12]. The aim of the ADASYN algorithm is to improve the class balance by synthetically creating new examples from the minority class via a linear interpolation between the existing minority class examples. ADASYN is an extension of the SMOTE (Synthetic Minority Oversampling TEchnique) method [11], creating more examples near the border between the two classes than inside the minority class. After this step, the learning set consists of 4897 trials of 5100 features : 2458 of class "0" and 2439 of class "1". In order to reduce the number of features we applied Common Spatial Pattern (CSP) method [13] which consists of a binary data-driven supervised data projection of a signal by maximizing the variance of the positive class while minimizing the variance of the negative one. These filters can be used in the construction of feature vectors, or other analyzes, where it is useful to remove as much noise as possible from a signal. This method is conventionally used to process electroencephalogram (EEG) signals, but we thought it would be useful in our study, the movements of the subjects being spatially reduced and very close. For the classification we have tried several classifiers (Linear Discriminant Analysis LDA, Logistic Regression LR, Naives Bayes NB, Cubic Support-Vector Machines SVM, Ensemble Sub-space k-Nearest Neighbours ESkNN [14]) and several numbers of CSPs. We used a 5 folds cross-validation technique to assess the accuracy of the classification (cf. Table II).

TABLE II
TASK 2 : ACCURACY ON TRAINING SET

#CSP	LDA	LR	NB	SVM	ESkNN
2	0.635	0.635	0.621	0.493	0.555
4	0.670	0.674	0.667	0.494	0.672
6	0.712	0.712	0.696	0.722	0.789
8	0.734	0.737	0.726	0.782	0.858
10	0.763	0.766	0.751	0.899	0.894
12	0.768	0.771	0.759	0.849	0.918
14	0.778	0.778	0.776	0.875	0.923
16	0.780	0.780	0.775	0.878	0.930
18	0.784	0.782	0.788	0.888	0.935
20	0.791	0.790	0.800	0.905	0.945
22	0.798	0.797	0.799	0.909	0.945
24	0.798	0.800	0.804	0.914	0.948
26	0.812	0.810	0.803	0.916	0.951
28	0.817	0.814	0.815	0.924	0.953
30	0.816	0.812	0.813	**0.923**	0.954
32	0.814	0.814	0.816	0.934	0.955
34	0.818	0.816	0.815	0.933	0.958
36	0.816	0.815	0.820	0.935	0.958
38	0.817	0.817	0.817	**0.938**	**0.961**
40	0.818	0.819	0.817	0.937	0.958
42	0.817	0.819	0.821	0.936	0.960

The three bold results in table II were selected and submitted to the challenge (cf. Table III).

IV. CONCLUSION

This article presents our contributions to the first two tasks of the AffectMove challenge : Task 1: "Protective Behaviour Detection based on Multimodal Body Movement Data" and Task 2: "Detection of Reflective Thinking based on Body Movement Data". For each of these tasks, we present the

TABLE III
TASK 2 : SUBMISSION SCORES

		Train	Test			
Classifier	#CSP	Acc.	F1-score (class 0)	F1-score (class 1)	Acc.	MCC
ESkNN	38	0.961	0.918	0.232	0.853	0.220
SVM Cubic	30	0.923	0.874	0.206	0.783	0.082
SVM Cubic	38	0.938	0.885	0.192	0.799	0.084

(Acc. : Accuracy)

complete process that we implemented to deal with these classification problems and the obtained results on the train and test sets.

Due to lack of time we could not implement a diagnostic-oriented method for task 1, but we describe it in the next section. We hope to be able to implement and present it during the AffectMove workshop.

V. FUTUREWORKS

Protective behavior detection based on multimodal body movement data defines a diagnosis problem and should be formalized as such. Protective behaviors correspond to counteractions for eliminating or minimizing the effects of physical or muscular disturbances. The earlier the detection, the better the anticipation of corrective actions. Fault detection and diagnosis have been largely studied in the literature related to industrial systems' reliability but can naturally be transposed to other application domains. Fault analysis of systems (technological, medical, human systems etc.) consists of detecting failures as early as possible in order to be able to anticipate and minimize their future effects. [15] distinguishes two kinds of reasoning for solving diagnostic tasks: abnormal-operation-oriented or abductive reasoning and normal-operation-oriented reasoning. Abnormal-operation-oriented diagnosis uses knowledge about how the system's components are affected by some specific faults in order to trace those faults: all possible failures that could occur are hypothesized, their effects are predicted and the counteractions for eliminating or minimizing these effects are designed. Normal-operation-oriented diagnosis uses knowledge about how normal components work to detect deviations from normality in observed behavior, from which a minimal set of faults is hypothesized. [16] proposed a logical theory of model-based diagnosis, also referred to as consistency-based diagnosis. The analysis is aimed at obtaining consistency between the observations and the model by removing assumptions about the behavior of some component [17]. This theory was extended and formalized in [18] and many refinements have since been proposed ([19]; [20]; [21]). In a short way, modeling relationships between variables provides a set of analytical redundancy relationships (ARR) that enables assessing consistency between observations online. Each ARR can be seen as a virtual sensor that does not measure any physical quantity but is dedicated to anomaly detection when checking expected consistency between variables. Model-based diagnosis is part of our further research works w.r.t the AffectMove challenge. The idea is to complete the position and EEG features with ARRs to diagnose abnormal behaviors more efficiently. We will proceed in several stages. First, the position features space will be completed with kinematics and dynamics features. We will use the CusToM software [22] to automatically implement kinematics and inverse dynamics methods: from the Mocap X, Y, Z positions of 17 points of the human body, the angles at the joints, the joint torques, the muscular efforts and the muscular activations will be computed. The first task will consist of formatting our .txt data into .c3d with the BTK toolkit [23]. Once the file obtained, we will have to differentiate the subjects, indeed musculoskeletal modeling is particularly sensitive to height and weight of the subject, then we will adjust the model by choosing thresholds and methods specific to the data provided for the challenge. The results of the machine learning part will be then applied upon this augmented features space. The interest of our approach is also to have a visual understanding of movements: we can observe some peculiarities of movements by representing raw data in 3D, that is a very useful aspect in terms of man/machine interaction and AI explainability. In the second step of our approach, we are also working on a more biomechanics-based approach. Since joint torques, muscular efforts and muscular activations are available through CusToM, we propose to establish the biomedical constraints that provide relationships between forces and torques. To each of these biomechanical relationships, an ARR that enables online checking of the consistency between the observations may be attached. We expect that the additional knowledge the biomechanical variables (forces and torques) and the related ARRs provide will improve the classifiers results that have been proposed for the AffectMove challenge.

REFERENCES

[1] http://www.casapaganini.it/entimement/workshops/2021

[2] B. de Gelder (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 364(1535), 3475–3484. https://doi.org/10.1098/rstb.2009.0190.

[3] G. Pezzulo, F. Donnarumma, H. Dindo, A. D'Ausilio, I. Konvalinka and C. Castelfranchi (2019). The body talks: Sensorimotor communication and its brain and kinematic signatures, Physics of Life Reviews, Volume 28, Pages 1-21, ISSN 1571-0645, https://doi.org/10.1016/j.plrev.2018.06.014.

[4] T. A. Olugbade, A. Singh, N. Bianchi-Berthouze, N. Marquardt, M. S. H. Aung, and A. C. De C. Williams (2019). How Can Affect Be Detected and Represented in Technological Support for Physical Rehabilitation? ACM Trans. Comput.-Hum. Interact. 26, 1, Article 1 (February 2019), 29 pages. DOI:https://doi.org/10.1145/3299095.

[5] M. S. H. Aung et al. (2015). The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal EmoPain dataset. IEEE transactions on affective computing 7.4: 435-451.

[6] E. Cippitelli, S. Gasparrini, E. Gambi, and Spinsante (2016). A human activity recognition system using skeleton data from RGBD sensors. Computational intelligence and neuroscience.

[7] Q. Perez, P. A. Jean, C. Urtado, and S. Vauttier (2021). Bug or not bug? That is the question. arXiv preprint arXiv:2103.12218.

[8] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux (2013). API design for machinelearning software: experiences from the Scikit-learn project. ECMLPKDD Workshop: Languages for Data Mining and Machine Learning, pp. 108–122.
[9] L. Breiman (2001). Random forests. Machine learning, 45(1), 5-32.
[10] T. Olugbade et al. (2020). Automatic Detection of Reflective Thinking in Mathematical Problem Solving based on Unconstrained Bodily Exploration. in IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2020.2978069.
[11] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer. (2002). SMOTE: Synthetic Minority Over-sampling Technique,ournal of Artificial Intelligence Research, volume 16, pages 321–357, ISSN 1076-9757, DOI 10.1613/jair.953.
[12] H. He, Y. Bai, E. A. Garcia and S. Li (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), 2008, pp. 1322-1328, doi: 10.1109/IJCNN.2008.4633969.
[13] A. K. Keng, C. Z. Yang, W. Chuanchu, G. Cuntai and Z. Haihong (2012). Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b. Frontiers in Neuroscience,vol. 6, p.39, doi:10.3389/fnins.2012.00039.
[14] A. Gul, A. Perperoglou, Z. Khan et al. Ensemble of a subset of kNN classifiers. Adv Data Anal Classif 12, 827–840 (2018). https://doi.org/10.1007/s11634-015-0227-5
[15] D. Poole (1988). Normality and faults in logic-based diagnosis. Proc. International Joint Conference on Artificial Intelligence 1998 (1304-1310), Detroit, Michigan, USA.
[16] R. Reiter (1987). A theory of diagnosis from first principles. Artificial Intelligence 32, pp. 57-95.
[17] R. Isermann (1997). Supervision, fault-detection and fault-diagnosis methods – An introduction. Control Eng. Practice 5(5), pp. 639-652.
[18] J. De Kleer, A. Mackworth and R. Reiter (1992). Characterizing diagnoses and systems. Artificial Intelligence, 56(2-3), 197–222.
[19] P. Frank (1996). Analytical and qualitative model-based fault diagnosis, a survey and some new results. European Journal of Control, 1 (2), pp 6 28.
[20] M-O. Cordier, P. Dague, F. Lévy, J. Montmain, M. Staroswiecki, L. Travé-Massuyès (2004). Conflicts versus Analytical Redundancy Relations: A comparative analysis of the model based diagnosis approach from the Artificial Intelligence and Automatic Control perspectives. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, 34 (5).
[21] S. Gentil, J. Montmain and C. Combastel (2004). Combining FDI and AI Approaches within Causal-Model-based Diagnosis. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, 34 (5).
[22] A. Muller, C. Pontonnier, P. Puchaud, G. Dumont (2019). CusToM : a Matlab toolbox for musculoskeletal simulation. Journal of Open Source Software, 2019, 4 (33), pp.1-3.
[23] A. Barré and S. Armand (2014). Biomechanical ToolKit: Open-source framework to visualize and process biomechanical data. Computer Methods and Programs in Biomedicine, 114 (1), 2014, pp 80-87, ISSN 0169-2607.