Multimodal Multilevel Fusion for Sequential Protective Behavior Detection and Pain Estimation

Md Taufeeq Uddin, Shaun Canavan
University of South Florida, Tampa, FL, USA
mdtaufeeq@mail.usf.edu, scanavan@usf.edu

Abstract— **In this paper, we present our approach to the FG 2020 EmoPain Challenge for tasks 2 (pain estimation) and 3 (protective behavior detection) from multimodal movement data. We propose to perform sequential protective behavior detection and pain estimation using human movement information. First, we predict the existence of pain, and then use this information along with the multimodal movement data for protective behavior detection. Finally, this information is fused to estimate level of pain. In this work, we apply both early fusion (feature fusion including metadata, modalities, exercises and probabilities) and post-fusion (decision fusion). The proposed approach is encouraging, as it outperforms the baseline, with high margin for both pain estimation and protective behavior detection on the EmoPain challenge 2020 dataset.**

I. Introduction

To improve the quality of life for patients with chronic pain (e.g. chronic lower back pain), rehabilitation is necessary [8]. In observing this pain, Keefe et al. [13] identified guarding, rubbing, bracing, grimacing, and sighing as five distinct categories. Along with this, Sullivan et al. [21] identified the functional category known as protective behavior, which we detect in this work. People that suffer from this often perform safety-seeking behavior (i.e. protective) and this behavior can also be measured using muscle activity [2], which is a motivating factor behind this work.

From a clinical standpoint, pain assessment is a critical task for patient well-being [20]. Coupled with the inability of some patients to describe their pain [9], automatic methods to detect pain become vital. Considering this, there has been encouraging progress made to develop automatic solutions to detecting pain [2], [17], [18], [24]. Aung et al. [1] investigated the protective behavior, guarding, which is commonly found in those with chronic lower back pain (CLBP). They used an ensemble of random forests on posture and velocity data, showing encouraging results on motion capture and electromyographic data [5]. Olugbade et al. [16] detected levels of pain (low, medium, and high) by fusing body motion and muscle activity and training a Support Vector Machine (SVM). They showed that a multimodal approach led to higher classification compared to body motion or muscle activity alone. Wang et al. [23] also detected protective behavior using MoCap and EMG data. They investigated using a dual-stream long short-term memory (LSTM) network, as well as stacked LSTM. Their results suggest that using a stacked LSTM approach can facilitate the development of wearable technology to support pain rehabilitation. Along with pain detection, motion data has also been successfully applied to predicting stress and meditation states [22], further showing the utility of this type of data for helping improve quality of life. Aside from motion data, another interesting modality, for detecting pain, is physiological data (e.g. blood pressure, respiration). Fabiano et al. [10] showed fusing physiological data can be used to accurately detect pain. They investigated the BP4D+ dataset [25], showing robustness to multiple machine learning algorithms.

Motivated by these works, we propose a sequential approach to detecting protected behavior, as well as estimating pain. To facilitate the detection of protected behavior, we propose to fuse human movement data along with pain detection results. Given this detection, we then propose to fuse movement data, the results of pain detection, and protective behavior detection to estimate the level of pain. To show the effectiveness of the approach, we conducted an ablation study of the modalities available in the emoPain movement dataset [2] by exploring the modalities separately and by combining them (early fusion). We also explore decision fusion and show that this type of fusion can preserve the competitive performance of the model while improving classification diversity [11]. The main contributions, of this work, can be summarized as follows.

1. This work proposes a multimodal, multilevel fusion approach for sequential protective behavior detection and pain estimation using human movement data.
2. An ablation study is conducted showing the effectiveness of the proposed approach. We analyze the modalities separately to compare against our proposed fusion approach.
3. Proposed approach outperforms baseline results on the EmoPain Challenge 2020 [3] validation and test sets, for protective behavior detection, and pain estimation.

II. Sequential Protective Behavior and Pain Prediction

In this work, we propose a multimodal, multilevel fusion model to detect protective behavior and to estimate pain intensity levels sequentially. Using metadata, mocap, EMG, and exercises (i.e. feature vector in V_1), we first detect the existence of pain. We then merge the outcome of the pain detection model with the feature vectors (e.g. F_1) in V_1, which creates new feature vectors in V_2. In this step, we use V_2 to detect protective behavior (PB). Our final task is pain estimation in which we use the combination of the outcome

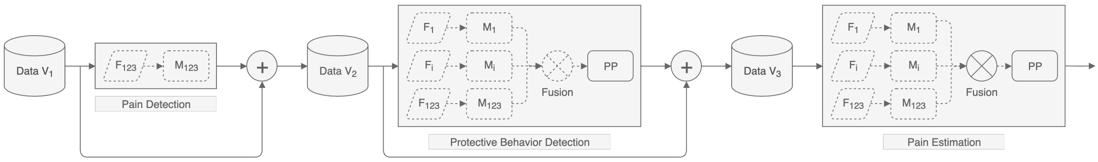

Fig. 1. Sequential protective behavior detection and pain estimation framework. Here, F, M, and PP indicate features (modalities), models, and post-processing. For example, M_i indicates a model (e.g. random forests) trained on feature set F_i (see Table I).

TABLE I

FEATURE VECTOR COMBINATIONS. EACH ROW REPRESENTS ONE FEATURE VECTOR. EACH COLUMN REPRESENTS ONE VERSION OF THE FEATURE VECTOR. MORE PRECISELY, IN (FEATURE VECTOR) VERSION V_1, WE HAVE 7 DIFFERENT COMBINATION OF METADATA, MOVEMENT DATA, EXERCISES, WHICH CONSTRUCT 7 UNIQUE FEATURE VECTORS. TO CONSTRUCT V_2, WE APPEND PAIN DETECTION PROBABILITIES (PD_P) TO THE V_1. FINALLY, IN V_3, PB DETECTION PROBABILITIES (PB_P) IS APPENDED TO THE V_2. NOTE THAT MD, ANG, ENG, EX INDICATE METADATA (E.G. PATIENT OR HEALTY SUBJECT, NORMAL OR DIFFICULT EXERCISE), ANGLES, ENERGIES, EXERCISES, RESPECTIVELY.

Index	V_1	V_2	V_3
1	F_1 = {MD, ang, ex}	{F_1, PD_p}	{F_1, PD_p, PB_p}
2	F_2 = {MD, eng, ex}	{F_2, PD_p}	{F_2, PD_p, PB_p}
3	F_3 = {MD, EMG, ex}	{F_3, PD_p}	{F_3, PD_p, PB_p}
4	F_{12} = {MD, ang, eng, ex}	{F_{12}, PD_p}	{F_{12}, PD_p, PB_p}
5	F_{23} = {MD, eng, EMG, ex}	{F_{23}, PD_p}	{F_{23}, PD_p, PB_p}
6	F_{13} = {MD, ang, EMG, ex}	{F_{13}, PD_p}	{F_{13}, PD_p, PB_p}
7	F_{123} = {MD, ang, eng, EMG, ex}	{F_{123}, PD_p}	{F_{123}, PD_p, PB_p}

of the PB detection model with feature vectors in V_2 as input feature vector (see Figure 1).

A. Pain Detection Model

An XGBoost classifier [7] is trained using feature vector F_{123} (fusion of metadata (subject type, exercise difficulty type), exercises, mocap, and EMG data) in V_1 as shown in Table I. We extract the class probabilities of pain and neutral classes from the XGBoost pain detection model and merge the probabilities with each feature vector in V_1 to get new feature vectors in V_2. Knowing the existence of pain may boost the performance of the PB detection model since people are likely to show protective behavior in the presence of pain, although there is no direct relationship between them [15] [19].

B. Protective Behavior Detection Model

For protective behavior detection, we apply multimodal, multilevel fusion by performing both feature fusion and decision fusion. We build a set of models using several combinations of input data, using the combinations shown in Table I. For instance, to build model M_i (e.g. M_1), we use i^{th} feature vector from V_2 by combining F_i (e.g. F_1) and pain detection probabilities (PD_p). For protective behavior detection, we use the XGBoost classifier, in which we put more weight on the protective behavior class (positive class) because the protective behavior class is less representative in the dataset. The weight parameter W of XGBoost classifier is optimized by

$$W = \sqrt{N_c/P_c}, \quad (1)$$

where N_c and P_c represent the total number of training samples belonging to the neutral class and positive (e.g. protective behavior (PB)) class, respectively. It can be inferred by merging the information from Figure 1 and Table I that we build 7 classifiers for PB detection using V_2 (the combination of feature vector F_i in V_1 in Table I and pain detection probabilities). We then, for two different submissions (submission 2 and 3), perform two different types of late decision fusion such as *averaging* and *majority voting*, to incorporate diversity [11] in the classification results, which could be useful to deal with domain shift [14]. In the case of averaging, we first select k number of competitive models (in terms of PB detection performance) from the 7 trained PB detection models, and then we extract the probabilities of a given sample belong to PB class from those k models and compute the mean of the probabilities. If the mean probability is greater than or equal to a threshold t, we classify the sample as PB, otherwise, we classify the sample as neutral. On the other hand, in the case of majority voting, after making the inference using the 7 PB detection models for a given sample, we perform a vote on top of the inference and we finally classify the sample as PB if at least 4 classifiers classify the sample as PB, otherwise, we classify the sample as neutral. Note that we train and perform inference using our classifiers at each time step (frame), while the EmoPain challenge 2020 makes inference at the segment level using a sliding window of 180 timesteps with 75% overlapping time steps (frames). Hence, we post-process the results we obtain in each timestep to mimic the challenge evaluation setting. We use the same segmentation approach as the challenge, to map 180 class labels (for 180 samples) to 1 label.

C. Pain Estimation Model

We perform multimodal, multilevel fusion for pain estimation as well, using V_3 which is the combination of V_1, the output of the pain detection model and PB detection model (see Figure 1 and Table I). Note that out of 7 PB detection models, we only use the outcome of the PB detection model M_{123} since M_{123} is trained on all sources of information including metadata, mocap (angle, energy), EMG, exercise and pain detection probabilities. Note that the level of pain (low and high) classes (compared to neutral) are less representative in the dataset; as a consequence, we perform oversampling of the training dataset using synthetic minority over-sampling technique (SMOTE) [6]. For pain estimation, we train both random forests [4] and XGBoost [7] classifiers using all possible (7) combination of modalities (see Table I). Thus, we have 14 pain estimation models in total. Next, we perform late fusion via decision aggregation. To do so, we select k competitive models (in terms of pain recognition performance) from all trained (14) pain estimation models. More precisely, we extract the output (numeral pain level) of k models and compute the mean of the outputs ($output_m$). Then, we use thresholds t_1, t_2 to get the final pain estimation as highlighted in Eqn. 2 and Eqn. 3.

$$output_m = \Big[\sum_{i=1}^{k} output_i\Big]/k \tag{2}$$

$$ePain = \begin{cases} 0 & \text{if } 0 \leq output_m < t_1 \\ 1 & \text{if } t_1 \leq output_m \leq t_2 \\ 2 & \text{if } t_2 < output_m \leq 3 \end{cases} \tag{3}$$

Note that our approach operates at each time step (frame). Hence, to get exercise instance level pain estimation (to mimic EmoPain challenge evaluation setting), we perform post-processing in which we apply majority voting over estimated pain for all of the time steps (samples) that belong to respective exercise.

III. Experiments and Analysis

Dataset. The emoPain dataset [2] is a multimodal pain dataset that contains visual, body movement and EMG data. The authors captured both pain intensity and protective behavior of subjects (both healthy and patient with CLBP) aged in between $[19, 67]$ (mean: 50.5 years). The subjects performed a set of physical exercises (e.g. walking, bend down, one leg stand, sit-to-stand) to elicit pain and protective behavior. The public release version of the dataset contians observer report of pain intensity and protective behavior. Note that in case of labeling pain intensity for the movement portion of the dataset, a subset of the samples was labeled as "Not reported". In our experiment, we transformed those samples (in training dataset) to "neutral" to augment the training dataset. For detailed information about the dataset, we refer the readers to Aung et al. [2].

Model validation and performance evaluation. In this work, we used the evaluation metrics Matthew's correlation coefficient (MCC) [12], F1-score (F1-score for each class, and mean F1-score), and accuracy. MCC is quite useful in this problem since it can take care of both positive and negative classes. During pain detection and PB detection, we performed leave-one-subject-out (LOSO) validation, and we extracted the detection results of the leave-out subject which was later used to train models (e.g. pain detection results $\rightarrow$ PB detection model, and pain detection and PB detection results $\rightarrow$ pain estimation model). For comparison with the baseline, we reported results on the validation and test partitions of the dataset. For the test set, we reported results of submitted models only (3 PB detection and 3 pain estimation models) as we did not have access to the test set.

Pain detection. We trained an XGBoost classifier to classify pain from no pain (neutral) using feature set F_1 in V_1 (see Table I). The class weight W was optimized (Eqn. 1) and we set the number of estimators, and learning rate to 200, and 0.1, respectively. The proposed model was able to classify pain samples from neutral samples with 100% accuracy with the inclusion of metadata and exercises with angles, energies and EMG data. Knowing the existence of pain is likely to boost the performance of PB detection models as well as pain estimation models as it is well-known that there is a co-occurance relationship between pain and protective behavior [2]. Hence, we extracted the class probabilities of the XGBoost-based pain detection model to feed in to the PB detection and pain estimation models.

Protective behavior detection. All 7 XGBoost classifiers for PB detection had 200 trees (estimators), class weight of $W = 4$ (Equation 1) and learning rate of 0.1. The parameter values were selected empirically to trade-off between training time and predictive performance. As can be seen in Table II, we obtained competitive results using M_2, M_3, M_{23}, and M_{123} models. As a result, for our final submission, we used the M_{123} model and two fusion models. The submitted models outperformed (with high margin) the baseline results provided by the emoPain Challenge 2020 (see Table II). For submission 2 (mean fusion), we selected the M_2, M_3, M_{23}, and M_{123} models as the top $k(= 4)$ models, to obtain the final outcome. For submission 3, we selected all 7 models and performed majority voting to achieve the final outcome. The goal of the fusion models is to incorporate diversity [11] in terms of classification. From Table II, we observe that, on validation set, fusion preserved the competitive model performance while incorporating diversity in terms of classification, while on test set (unknown dataset), fused models outperformed single model, which validates our assumption that on unknown shifted samples fusion could be better choice than single model, as well as PD detection improving the PB detection results due to co-occurance relation.

Pain estimation. We trained 7 random forest and XGBoost classifiers using each feature vector in V_2 in Table I and the method described in Section III. To alleviate class imbalance issues, we oversampled the training partition of the dataset using SMOTE [6]. For both classifiers, we set the number of estimators (trees) to 200, and for XGBoost classifier, we set learning rate to 0.1. To preserve performance and to incorporate classification diversity, we performed a late fusion of top k pain estimation models. For submission 1 (pain estimation), we fused random forests based pain estimation models M_{23} and M_{123} trained on feature vectors $\{F_{23}, PD_p, PB_p\}$ and $\{F_{123}, PD_p, PB_p\}$ using the approach described in Section III. We set threshold $t_1 = 0.9$ and $t_2 = 1.5$. For submission 2, we selected $k = 8$ models for fusion, where we used $\{F_2, PD_p, PB_p\}$, $\{F_3, PD_p, PB_p\}$, $\{F_{23}, PD_p, PB_p\}$, and $\{F_{123}, PD_p, PB_p\}$ to train XGBoost

TABLE II

PB DETECTION RESULTS. F1_{M}, F1_{N}, AND F1_{PB} REPRESENT MEAN F1 SCORE, F1 SCORE OF NEUTRAL AND PB CLASSES, RESPECTIVELY. NOTE THAT PB DETECTION MODELS WERE TRAINED USING V_2, FOR EXAMPLE, M_1 WAS TRAINED USING $\{F_1, PD_p\}$ IN V_2 AND SO ON.

	Validation set					Test set				
Model	MCC	F1_{m}	F1_{n}	F1_{pb}	Accuracy	MCC	F1_{m}	F1_{n}	F1_{pb}	Accuracy
Baseline [3]	-	0.48	0.96	-	0.46	-	0.57	0.9	0.25	0.83
M_1	0.29	0.62	0.91	0.32	0.84	-	-	-	-	-
M_2	0.427	0.68	0.93	0.44	0.88	-	-	-	-	-
M_3	0.411	0.68	0.94	0.43	0.89	-	-	-	-	-
M_{12}	0.282	0.62	0.91	0.32	0.85	-	-	-	-	-
M_{23}	**0.465**	**0.72**	0.95	**0.49**	0.91	-	-	-	-	-
M_{13}	0.283	0.61	0.91	0.32	0.84	-	-	-	-	-
M_{123} (submission 1)	0.45	**0.72**	**0.96**	0.48	**0.93**	0.59	0.78	0.96	0.61	0.92
Mean fusion (submission 2)	0.45	0.71	0.95	0.48	0.91	**0.66**	0.81	0.96	**0.67**	0.93
Majority voting (submission 3)	0.42	0.7	0.95	0.45	0.9	0.63	0.81	**0.97**	0.65	**0.94**

TABLE III

PAIN ESTIMATION RESULTS. F1_{M}, F1_{N}, F1_{LP}, AND F1_{HP} REPRESENT F1 SCORES - MEAN, NEUTRAL, LOW, AND HIGH LEVEL PAIN CLASSES.

	Validation set						Test set					
Model	MCC	F1_{m}	F1_{n}	F1_{lp}	F1_{hp}	Accuracy	MCC	F1_{m}	F1_{n}	F1_{lp}	F1_{hp}	Accuracy
Baseline [3]	0.02	0.31	0.39	0.09	0.44	0.34	-	-	-	-	-	-
M_1 (random forests)	0.62	0.51	1	0.53	0	0.85	-	-	-	-	-	-
M_2 (random forests)	0.83	0.72	1	0.82	0.33	0.93	-	-	-	-	-	-
M_3 (random forests)	0.82	0.72	1	0.81	0.34	0.93	-	-	-	-	-	-
M_{12} (random forests)	0.64	0.52	1	0.57	0	0.86	-	-	-	-	-	-
M_{23} (random forests)	0.86	0.74	1	0.86	0.37	0.94	-	-	-	-	-	-
M_{13} (random forests)	0.63	0.52	1	0.55	0	0.85	-	-	-	-	-	-
M_{123} (random forests)	0.9	**0.79**	1	0.91	0.47	0.96	-	-	-	-	-	-
M_1 (XGBoost)	0.63	0.52	1	0.53	0.03	0.85	-	-	-	-	-	-
M_2 (XGBoost)	0.74	0.65	1	0.66	0.30	0.89	-	-	-	-	-	-
M_3 (XGBoost)	0.79	0.71	1	0.76	0.36	0.92	-	-	-	-	-	-
M_{12} (XGBoost)	0.62	0.52	1	0.49	0.06	0.84	-	-	-	-	-	-
M_{23} (XGBoost)	0.78	0.69	1	0.73	0.34	0.91	-	-	-	-	-	-
M_{13} (XGBoost)	0.64	0.53	1	0.56	0.03	0.86	-	-	-	-	-	-
M_{123} (XGBoost)	0.87	0.78	1	0.87	0.48	0.95	-	-	-	-	-	-
Submission 1	**0.91**	**0.79**	1	**0.92**	0.46	**0.97**	-	-	1	0.25	0.14	0.45
Submission 2	0.86	0.77	1	0.86	0.44	0.95	-	-	1	0.26	0.21	0.45
Submisison 3	0.87	**0.79**	1	0.87	**0.49**	0.95	-	-	1	**0.27**	**0.27**	0.45

and random forest classifiers (Figure 1). For submission 3, we selected 3 random forest-based models (trained on $\{F_2, PD_p, PB_p\}$, $\{F_{23}, PD_p, PB_p\}$, $\{F_{123}, PD_p, PB_p\}$) and 2 XGBoost based models (trained on $\{F_3, PD_p, PB_p\}$, $\{F_{123}, PD_p, PB_p\}$) for fusion. Our assumption is that under uncertainty (unknown test dataset), the fusion of models is likely to perform better than a single model. As it can be seen in Table III, our method produced an MCC value of 0.91 that outperformed the baseline results by 0.89. It also resulted in an accuracy of 0.97, outperforming the baseline by 0.63. Note that MCC score reported in the validation set was computed in multiclass setting. On test set, MCC was computed for each class by emoPain (challenge 2020) organizer, our proposed model obtained MCC of $(1, 0.28, 0.21; mean = 0.5)$, $(1, 0.29, 0.25; mean = 0.51)$, and $(1, 0.3, 0.3; mean = 0.53)$ for neutral, low level and high level pain, respectively, using submission 1, submission 2 and submission 3. Notice that pain estimation results on test set are comparatively worse than results on validation set. One potential reason could be domain shift [14], where the test set has a different distribution compared to the training and validation sets.

IV. CONCLUSION

We have detailed our proposed approach to the EmoPain challenge 2020 at IEEE FG 2020 for tasks 2 and 3. Our approach is a sequential multimodal, multilevel approach to detecting protected behavior, as well as pain estimation. The proposed approach outperforms the baseline, showing the utility of the proposed approach for pain detection. We have also shown that the proposed fusion is necessary, as it outperforms single modalities for the task. An interesting finding, from this work, is that the incorporation of exercise type improves the performance of models. This can partially be explained by the association between exercise type and certain body movements. As rehabilitation is important for those with chronic pain, this work can help facilitate automatic approaches for helping with this, and ultimiately improving quality of life.

ACKNOWLEDGMENT

This material is based on work that was supported in part by an Amazon Machine Learning Research Award.

References

[1] M. H. Aung, N. Bianchi-Berthouze, P. Watson, and A. d. C. Williams. Automatic recognition of fear-avoidance behavior in chronic pain physical rehabilitation. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 158–161, 2014.

[2] M. S. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh, et al. The automatic detection of chronic pain-related expression: requirements, challenges and the multimodal emopain dataset. *IEEE transactions on affective computing*, 7(4):435–451, 2015.

[3] N. Berthouze, M. Valstar, A. Williams, J. Egede, T. Olugbade, C. Wang, H. Meng, M. Aung, N. Lane, and S. Song. Emopain challenge 2020: Multimodal pain evaluation from facial and bodily expressions. *arXiv preprint arXiv:2001.07739*, 2020.

[4] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[5] M. Cella, M. Aung, H. Meng, N. Bianchi-Berthouze, B. Romera-Paredes, A. Singh, M. Shafizadehkenari, A. Williams, P. Watson, A. Kemp, and S. Brooks. Identifying pain behaviour for automatic recognition. In *International Association for the Study of Pain*, 2012.

[6] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[7] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[8] A. C. de C Williams, C. Eccleston, and S. Morley. Psychological therapies for the management of chronic pain (excluding headache) in adults. *Cochrane database of systematic reviews*, (11), 2012.

[9] 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)*, pages 689–696. IEEE, 2017.

[10] D. Fabiano and S. Canavan. Emotion recognition using fused physiological signals. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 42–48. IEEE, 2019.

[11] Z. Gong, P. Zhong, and W. Hu. Diversity in machine learning. *IEEE Access*, 7:64323–64350, 2019.

[12] G. Jurman, S. Riccadonna, and C. Furlanello. A comparison of mcc and cen error measures in multi-class prediction. *PloS one*, 7(8), 2012.

[13] F. J. Keefe and A. R. Block. Development of an observation method for assessing pain behavior in chronic low back pain patients. *Behavior therapy*, 1982.

[14] T. Lesort, V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, and N. Díaz-Rodríguez. Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. *Information Fusion*, 58:52–68, 2020.

[15] T. Olugbade, N. Bianchi-Berthouze, and A. C. d. C. Williams. The relationship between guarding, pain, and emotion. *Pain reports*, 4(4), 2019.

[16] T. A. Olugbade, M. H. Aung, N. Bianchi-Berthouze, N. Marquardt, and A. C. Williams. Bi-modal detection of painful reaching for chronic pain rehabilitation systems. In *Proceedings of the 16th International Conference on Multimodal Interaction*, pages 455–458, 2014.

[17] T. A. Olugbade, N. Bianchi-Berthouze, N. Marquardt, and A. C. Williams. Pain level recognition using kinematics and muscle activity for physical rehabilitation in chronic pain. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 243–249. IEEE, 2015.

[18] T. A. Olugbade, A. Singh, N. Bianchi-Berthouze, N. Marquardt, M. S. Aung, and A. C. D. C. Williams. How can affect be detected and represented in technological support for physical rehabilitation? *ACM Transactions on Computer-Human Interaction (TOCHI)*, 26(1):1–29, 2019.

[19] 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 *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–7. IEEE, 2019.

[20] J. Stephenson. Veterans' pain a vital sign. *JAMA*, 281(11):978–978, 1999.

[21] M. J. Sullivan, P. Thibault, A. Savard, R. Catchlove, J. Kozey, and W. D. Stanish. The influence of communication goals and physical demands on different dimensions of pain behavior. *Pain*, 125(3):270–277, 2006.

[22] M. T. Uddin and S. Canavan. Synthesizing physiological and motion data for stress and meditation detection. In *ACII Workshops and Demos*, pages 244–247, 2019.

[23] 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*, pages 225–230, 2019.

[24] C. Wang, M. Peng, T. A. Olugbade, N. D. Lane, A. C. D. C. Williams, and N. Bianchi-Berthouze. Learning bodily and temporal attention in protective movement behavior detection. *arXiv preprint arXiv:1904.10824*, 2019.

[25] Z. Zhang, J. M. Girard, Y. Wu, X. Zhang, P. Liu, U. Ciftci, S. Canavan, M. Reale, A. Horowitz, H. Yang, J. Cohn, Q. Ji, and L. Yin. Multimodal spontaneous emotion corpus for human behavior analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3438–3446, 2016.