Towards Automated Pain Detection in Children using Facial and Electrodermal Activity

Xiaojing Xu[1], Büşra Tuğçe Susam[2], Hooman Nezamfar[3], Damaris Diaz[4], Kenneth D. Craig[5], Matthew S. Goodwin[6], Murat Akcakaya[2], Jeannie S. Huang[4], Virginia R. de Sa[7]

[1] Department of Electrical and Computer Engineering, UC San Diego, La Jolla, CA, USA, xix068@ucsd.edu
[2] Department of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA
[3] Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA
[4] Rady Childrens Hospital and Department of Pediatrics, UC San Diego, CA, USA
[5] Department of Psychology,University of British Columbia Vancouver, BC, Canada
[6] Department of Health Sciences, Northeastern University, Boston, MA, USA
[7] Department of Cognitive Science, UC San Diego, La Jolla, CA, USA

Abstract. Accurately determining pain levels in children is difficult, even for trained professionals and parents. Facial activity and electrodermal activity (EDA) provide rich information about pain, and both have been used in automated pain detection. In this paper, we discuss preliminary steps towards fusing models trained on video and EDA features respectively. We demonstrate the benefit of the fusion with a special test case involving domain adaptation and improved accuracy relative to using EDA and video features alone.

Keywords: Automated Pain Detection, EDA, Facial Action Units, GSR

1 Introduction

Accurate pain assessment in children is basic to safe and efficacious pain management. Under-estimation leads to patient suffering and inadequate care while over-estimation leads to adverse side-effects, including opioid addiction [1]. The most widely used method to assess clinical pain is patient self-report [2]. However, this method is subjective and vulnerable to social biases, and requires substantial cognitive, linguistic, and social competencies [2]. Objective pain estimation is required for appropriate pain management in the clinical setting.

In previous work, features extracted from facial action units (AUs) and EDA signals have both been used to automatically detect pain events using machine learning methods [3–6].

2 Methods

2.1 Participants

Forty-two pediatric research participants (30 males, 12 females) aged 13[10,15] (median [25%, 75%]) years and primarily Hispanic (79%) who had undergone

medically necessary laparoscopic appendectomy were recruited for a study examining automated assessment of children's post-operative pain using video and body sensors. Children and their parents provided assent and parental consent prior to study evaluations.

2.2 Experimental Design and Data Collection

Data were collected over 3 visits (V): (V1) within 24 hours after appendectomy in hospital; (V2) in hospital one calendar day after V1; and (V3) a follow-up visit in an outpatient lab up to 42 days postoperatively. At each visit, videos (60 fps at 853x480 pixel resolution) of the patient's face and EDA responses (using Affectiva Q sensor) were recorded while manual pressure was exerted at the surgical site for 10 seconds (equivalent of a clinical examination). Participants rated their pain level using a 0-10 Numerical Rating Scale, where 0 = no pain and 10 = worst pain ever. Following convention for recognizing clinically significant pain [7], videos and EDA with ratings of 0-3 were labeled as no pain, and videos with ratings of 4-10 were labeled as pain. We obtained 22 pain samples from V1, 8 pain and 8 no pain samples from V2, and 22 no pain samples from V3.

2.3 Feature Extraction

Video Features: Each 10-second video was processed with iMotions software which automatically estimates the log probabilities of 20 AUs (AU 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 18, 20, 23, 24, 25, 26, 28, 43) and 3 head pose indicators (yaw, pitch and roll) from each frame. We then applied 11 statistics (mean, max, min, standard deviation, 95th, 85th, 75th, 50th, 25th percentiles, half-rectified mean, and max-min) to each AU over all frames to obtain 11×23 features.

EDA Features: EDA signals were trimmed to 30 seconds (10 seconds before, during, and after the pressure respectively), smoothed by a 0.35 Hz FIR low pass filter, down-sampled to 1 Hz, and normalized with z-score normalization. We then applied timescale decomposition (TSD) with standard deviation metric, and computed the mean, SD, and entropy of each row of each TSD, to obtain a feature vector of length 90 [5].

2.4 Machine Learning Models

Support Vector Machine (SVM): A linear SVM was used to obtain a pain score as well as a pain prediction for each sample using video/EDA features after PCA. The number of principal components was chosen using cross-validation. In other ways, training was as in [5] (Inputs were normalized with z-score normalization over the full dataset).

Linear Discriminant Analysis (LDA): LDA was used to differentiate between pain and no pain using pain scores from the SVMs. Inputs were either one single pain score from one SVM, or a fusion of pain scores from both SVMs.

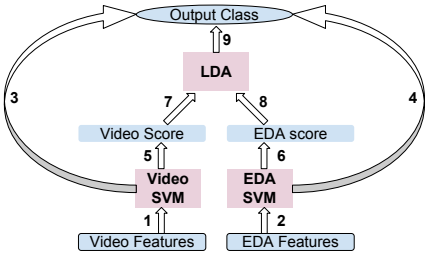

Fig. 1. Graph of Model Hierarchy

Table 1. Accuracy for Classification on V2

Video	EDA	Fusion-Feature	Fusion-Model	Video-V2	EDA-V2	Fusion-Model-V2
0.5	0.75	0.56	0.56	0.69	0.71	**0.84**

3 Preliminary Results and Discussion

In this work, we focused on V2 (in-hospital) pain v. no pain classification (a priority clinical concern), and used accuracy to evaluate model performance.

3.1 Performance Using Video/EDA Features

We first used V1 pain and V3 no pain samples to train an SVM for classification, following $1 \Rightarrow 3$ and $2 \Rightarrow 4$ in Figure 1. Table 1 shows the performance on V2 was fine for EDA, but suboptimal for video features. We hypothesized that this was due to iMotions feature sensitivity to environmental differences between V1/2 (in hospital) and V3 (in outpatient lab) [6]. One solution to this problem was to use V2 to train the model. However with only 16 datapoints in V2, results had very large variance. Likewise, training with V1/3 and V2 data together did not improve V2 performance. Consequently, we needed to solve the domain adaptation problem which learns a model from a source domain (V1/3) and performs it on a different target domain (V2).

3.2 Fusion of Video and EDA

We hoped to improve performance on V2 by combining video and EDA features. Our first simple attempt at fusion was to fit an LDA model to distinguish between pain v. no pain using the output pain scores from each of the SVM models trained with video and EDA features respectively ($1, 2 \Rightarrow 5, 6 \Rightarrow 7, 8 \Rightarrow 9$ in Figure 1). However, this method performed poorly (0.56) compared to using EDA features alone.

3.3 Training with V2 Scores

In the fusion, our LDA classifier had only two inputs: video and EDA SVM pain scores. Since the dimension of features was greatly reduced by SVM, it

became feasible to train a classifier using only V2 samples. Relative to Figure 1, we thus trained 1,2 with V1/3, and 7,8 with V2 data using cross-validation. The accuracy using video scores was 0.69, much higher than 0.5, showing the benefit of training on target domain V2, even if the features, V2 scores, were obtained from a model trained on V1/3. Finally with a fusion of SVM output scores for video and EDA, we achieved the comparative best accuracy of 0.84. For comparison, directly concatenating video and EDA features (fusion at the feature level) did not perform well (0.56).

4 Conclusion

We present preliminary results from a fusion approach to detecting pain in children. While the results demonstrate improvement with our domain adaptation fusion approach than with video or EDA features alone, we believe these results can be further improved by tailoring the two modalities to be more sensitive to their relative benefits and limitations.

Acknowledgments

This work was supported by National Institutes of Health National Institute of Nursing Research grant R01 NR013500.

References

1. B. L. Quinn, E. Seibold, and L. Hayman. Pain assessment in children with special needs: A review of the literature. *Exceptional Children*, 82(1):44–57, 2015.
2. G. Zamzmi, C.-Y. Pai, D. Goldgof, R. Kasturi, Y. Sun, and T. Ashmeade. Machine-based multimodal pain assessment tool for infants: a review. *preprint arXiv:1607.00331*, 2016.
3. K. Sikka, A. A. Ahmed, D. Diaz, M. S. Goodwin, K. D. Craig, M. S. Bartlett, and J. S. Huang. Automated assessment of childrens postoperative pain using computer vision. *Pediatrics*, 136(1):e124–e131, 2015.
4. S. Gruss, R. Treister, P. Werner, H. C. Traue, S. Crawcour, A. Andrade, and S. Walter. Pain intensity recognition rates via biopotential feature patterns with support vector machines. *PloS one*, 10(10):e0140330, 2015.
5. B. T. Susam, M. Akcakaya, H. Nezamfar, D. Diaz, X. Xu, V. R. de Sa, K. D. Craig, J. S. Huang, and M. S. Goodwin. Automated pain assessment using electrodermal activity data and machine learning (in press). In *Engineering in Medicine and Biology Society (EMBC), 2018 40th Annual International Conference of the IEEE*. IEEE, 2018.
6. X. Xu, K. D. Craig, D. Diaz, M. S. Goodwin, M. Akcakaya, B. T. Susam, J. S. Huang, and V. R. de Sa. Automated pain detection in facial videos of children using human-assisted transfer learning (in press). In *Joint Workshop on Artificial Intelligence in Health*, 2018.
7. D. Hoffman, A. Sadosky, E. Dukes, and J. Alvir. How do changes in pain severity levels correspond to changes in health status and function in patients with painful diabetic peripheral neuropathy. *Pain*, 149(2), May 2010.