

Assessment of Pain Using Facial Pictures Taken with a Smartphone

Mohammad Adibuzzaman¹, Colin Ostberg¹, Sheikh Ahamed¹, Richard Povinelli², Bhagwant Sindhu³, Richard Love⁴, Ferdaus Kawsar¹, Golam Mushih Tanimul Ahsan¹

¹Department of MSCS, Marquette University, Milwaukee, WI, USA

²Department of EECE, Marquette University, Milwaukee, WI, USA

{mohammad.adibuzzaman, colin.ostberg, sheikh.ahamed, richard.povinelli, ferdaus.kawsar, golammushihtanimul.ahsan}@marquette.edu

³Department of Occupational Science & Technology, UW Milwaukee, WI, USA, {Sindhu}@uwm.edu
International Breast Cancer Research Foundation {richardibcrf@gmail.com}

Abstract—Timely and accurate information about patients' symptoms is important for clinical decision making such as adjustment of medication. Due to the limitations of self-reported symptom such as pain, we investigated whether facial images can be used for detecting pain level accurately using existing algorithms and infrastructure for cancer patients. For low cost and better pain management solution, we present a smart phone based system for pain expression recognition from facial images. To the best of our knowledge, this is the first study for mobile based chronic pain intensity detection. The proposed algorithms classify faces, represented as a weighted combination of Eigenfaces, using an angular distance, and support vector machines (SVMs). A pain score was assigned to each image by the subject. The study was done in two phases. In the first phase, data were collected as a part of a six month long longitudinal study in Bangladesh. In the second phase, pain images were collected for a cross-sectional study in three different countries: Bangladesh, Nepal and the United States. The study shows that a personalized model for pain assessment performs better for automatic pain assessment and the training set should contain varying levels of pain representing the application scenario.

Keywords—Automatic pain assessment; remote monitoring; quality of life.

I. INTRODUCTION

In excess of 8 million individuals globally die each year from cancer and three-quarters of these are reported to suffer from pain [1]. A primary barrier to provision of adequate symptom treatment is failure to appreciate the intensity of the symptoms—most commonly pain—patients are experiencing [2]. One difficulty for health care providers in helping patients with chronic conditions like cancers is having accurate, complete, and timely information about symptoms, daily information if possible. In particular, failure to use (repeatedly) validated symptom assessment tools prevents communication between patients and health-caregivers to bring attention to symptoms' issues [3]. The usual way to obtain such information is to ask office-visiting patients standard questions about their symptoms and their intensities. For patients with cancer the most widely used questionnaires for this task are the Edmonton Symptom

Assessment Survey (ESAS) or the Brief Pain Inventory [4] [5]. Common practice is to have patients provide answers on paper to these instruments when they are seen in doctors' offices. This practice of course means that the data obtained only cover the particular situation the patient is in at that time. For example the patient may have taken extra pain medicines because of the appointment trip and wait in the doctor's office. A more abbreviated symptom assessment strategy in doctor-patient encounters is simply to have patients verbally report their current level of pain on a visual analogue 1-10 scale; sometimes a picture of this scale with figure faces showing different levels of distress is used. However, this is a one-time and one-item assessment strategy.

In an ideal situation, to monitor patients more completely and know every day how patients feel and then of course to make adjustments in treatments, such as types, amounts and timing of pain medicines, it would be good to have data from such questionnaires every day. In studies where patients have home computers, such daily assessments reported by email and the resultant treatment adjustments, are associated with increased quality of life and survival in terminally-ill patients with lung cancer. In settings where hospice programs are available, patient and family satisfaction is clearly related to intensity of monitoring and consequent associated prompt adjustments of symptomatic managements. This intensity of hospice monitoring is almost always greater than patients have had in their regular care, which has been usually and mostly based on patient office visits face-to-face. Management through phone contact, or email contact is usually limited, mostly because doctors are uncomfortable with their command of the full picture of the problems they are managing, but also because the practice of medicine has historically been based on face-to-face encounters. All of these issues are magnified in low- and middle income countries where limited access to care, sub-optimal quality of care and usually no hospice care at all, are the norms.

One more practical way to make obtaining such more detailed symptom information possible and usable by physicians, is to put the questionnaires on a cell phone software platform, which the patient or his/her attendant could then complete at home and send by phone each day to a doctor's records/office. We have developed such a system [6] and are now trying to scale this up into a tele-home hospice system in rural Bangladesh. Our experience has highlighted two broad issues; first particularly with respect to chronic pain which characterizes the situation for patients with incurable cancers, there is an apparent "dis-connect" between what patients report about their pain levels using the standard instruments and their affect. Patients often report high pain levels while smiling. Second, many patients cannot use these instruments because of limited cognitive abilities or other medical conditions—for example patients who are seriously ill in medical intensive care units. These issues have led us to investigate whether we could reproducibly and accurately record and quantitate patients' pain levels using cell phone camera images of facial expression. In this communication we address three broad questions in our investigation: First, can in fact facial images be used to reproducibly assess pain intensity among cancer patients? Second, what algorithm can define pain intensity most accurately? Third, what system design issues arise in this work?

Facial expression for quantitative pain assessment has its roots in psychology. There have been historical concerns about the objectivity of self-report assessments and their susceptibility to behavioral bias [7]. As a result, there is ongoing work to identify universal cues for pain expression. Prkachin et al. showed that indices of facial expression change due to variations of pain [8]. Ekman and Friesen's [9] Facial Action Coding System (FACS) has been used to identify universal Action Units corresponding to pain expression. It has been shown that for cold, pressure, ischemia and electric shock there are significant changes in four types of actions - brow lowering, eye orbit tightening, upper-lip raising/nose wrinkling and eye closure [10]. Prkachin and Solomon defined a 'Prkachin and Solomon Pain Intensity' (PSPI) measure as the sum of the intensities of these four actions [11]. Therefore, it appears that there are multiple facial expression components (i.e. action units) that together comprise the facial expression depicting the intensity of pain experienced by individuals.

We conclude that a few specific action units correspond to pain intensities for most pain expressions. For acute and chronic pain, the change in different action units might be of different magnitudes. A large amount of data for a particular target application (e.g. pain monitoring for chronic pain) improves the accuracy of pain intensity predictions using facial expressions [11]. While it is comparatively easy to identify changes in action units due to acute pain, it is more difficult to find the exact change in action units due to chronic pain. In this context, instead of using the FACS, we use principal component analysis to extract information that would give us reasonable variance across a given data set for pain expression. The fact that the principal component

analysis gave that the four core action units comprises 0.30 or greater fraction of pain expression across all pain tests [12] supports this claim. Therefore, we chose to use a principal component based method for detection of pain expression. Eigenface method is such a method. Each of the Eigenfaces corresponds to different levels of variance in the training data set.

II. DESIGN ISSUES

In the beginning of our work, we spent several weeks in clinics and hospitals and in home visits with cancer patients in Bangladesh learning about patient needs and health care professional challenges. Physicians expressed significant interest in having real time "usual day and activity" symptom data on their patients. Subsequently we spent similar durations of time in the field working on deployment of our system and helped us identify the design issues for such systems.

A. Availability of Mobile Network

Among poor rural Bangladeshi patients we were encouraged to find that among 45 patients surveyed, 43 had access to a cell phone. Additionally we found that these patients were served by good data networks. Most image processing techniques require very high computing power and it is difficult to use a smart phone for this purpose. The availability of good mobile data networks made it possible for us to use the cloud for photo images using advanced software such as Matlab

B. Smile for the camera

Common practice in affective computing is that when a person has a photo image of their face taken for assessment of affective state, he or she is asked to 'pose for the camera' or 'smile for the camera' [14]. To address the biases this usual practice likely creates, we elected to take two facial expression pictures in each studied patient. The first (candid) image was taken without giving any instructions. The second image was taken after providing specific directions. We instructed the patient to make a facial expression that reflected their current pain level. We defined that image as the 'acted pain' and the image without instruction as the 'real pain' for the purpose of this paper. In our training data set to be described below (dataset subsection), the images were randomly selected from 'acted pain' and 'real pain' images.

C. Personalized Model

In the first phase of our work, our goal was to identify pain levels from facial photographs for individuals followed longitudinally. There is a significant amount of variance in pain levels with facial expressions among individuals. As a result, we hypothesized that a training database with multiple images from a single person would eliminate this variance. Furthermore, individualized training databases for pain intensity would accommodate for different total numbers of images from individuals.

TABLE I. SIZE OF THE DATA SET

Longitudinal Study			
Subject	Training Set	Test Set	Total
A	6	8	14
B	36	80	116
C	36	124	160
D	6	6	12
E	36	78	114
F	6	32	38
Total			454
Cross Sectional Study			
Location	Training Set	Test Set	
Bangladesh	454	131	
Nepal	454	311	
United States	454	71	
Total		513	

III. DATA

A. Data Collection Architecture

In our first phase longitudinal study, we collected images of patients with advanced breast cancer in rural Bangladesh. The patient attendant used a mobile phone to take an image using our software. The software automatically uploaded the image once it was taken. Images were uploaded using PHP, Javascript and Wamp server.

B. Dataset

Our protocol for the longitudinal pilot study in a small number of patients was approved by the Institutional Review Board (IRB) at Marquette University and by The Bangladesh Medical Research Council in Bangladesh. All patients provided written informed consent. Each patient was given a Nokia X6 phone with internet provided by Grameen Phone, the largest mobile service provider in Bangladesh. The patients (all women) were aged between 35 and 48 years.

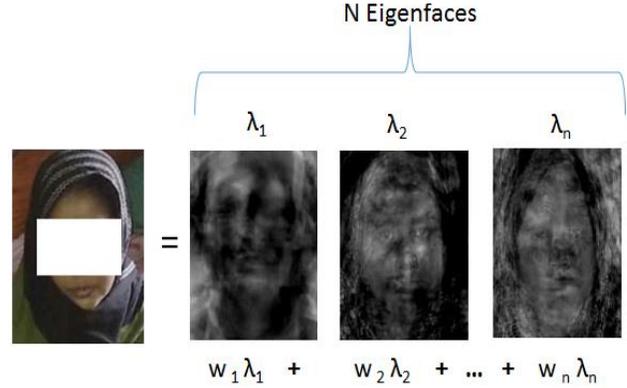
We recruited 14 patients. Each subject and the attendant of the subject were trained how to take the pictures using the camera at the health center.

The key aspects of the training and image creation are:

- A doctor would take a picture of the subject.
- The subjects were told about the Visual Analog Scale (VAS), which describes the pain intensity with 0 being the no pain and 10 being maximum pain possible.

From among the 14 patients, 6 lived longer than 3 months and regularly provided a total of 454 usable images. In the second phase of our work we conducted a cross sectional study. The protocol for this study was approved at Marquette University and by the responsible ethical review boards in Bangladesh, Nepal and Rapid City South Dakota in the United States. In this study we recruited patients

FIGURE 1. LINEAR COMBINATION OF N EIGENFACES FOR ONE IMAGE.



presenting for a clinic visit with advanced cancer and at that single visit obtained two facial images as noted above—one candid and one after specific instructions. Table I shows that we obtained usable photographs for 131 Bangladeshi, 311 Nepali, and 71 American Indian patients. 36 randomly selected images were used as the training set for each subject during the longitudinal study. For the cross sectional study, the entire dataset of the longitudinal study (454 images) was used as the training set.

IV. METHODS

A. Face detection

While subjects in both phases of our work were asked to take images of only the face, most of the images contained significant extraneous background around the face. We used Picasa 3.0 image viewing software to create images containing only the face. After the face portion was extracted, each image was resized to 160 times 120 pixels.

B. Eigenface

The Eigenface method is based on principal component analysis (PCA) which identifies the principal components represented by the eigenvectors corresponding to the highest eigenvalues. The method identifies the most significant features represented in an image and disregards the others. Figure 1 shows some sample Eigenfaces from the training database. Each of the Eigenfaces is a combination of all the images in the training database corresponding to different Eigenvalues. Here we only show Eigenfaces to preserve the confidentiality of the subjects.

C. Classification of weight vectors

For the classification of the weight vectors, we applied three approaches. These approaches used Euclidean distance, angular distance and support vector machine respectively. Different distant measures were needed for the high dimensions of the weight vectors. The dimension of the weight vector is equal to the number of images minus one in the training set. Consequently, for a training set of 36 images, we have 35 Eigenfaces and the weight vector was of

35 dimensions. Angular distance works better (mean absolute error decreases) than Euclidean distance in high dimensional space. We also used support vector machine to improve the sensitivity and specificity for each pain class.

V. RESULTS

The results of our system were evaluated in terms of two performance measure: the mean absolute error and, sensitivity and specificity analysis for the three pain classes, low (L), medium (M) and high (H). Pain level between 1 and 4 was termed as low, between 5 and 7 was considered as medium and between 8 and 10 was defined as high. This classification into three categories is similar to the Brief Pain Inventory which has been proposed and validated across different cultures [5]. As a consequence of insufficient data for subjects A, D, and F (Table I), the results are shown only for subjects B, C and E.

A. First phase--longitudinal study

1) Mean Absolute Error

We have six different training sets for the six subjects from the first phased longitudinal study. These training sets have a randomized combination of 36 images when available (subjects B, C, E) of 'acted' and 'real' pain.

TABLE II. MEAN ABSOLUTE ERROR FOR A 10 FOLD CROSS VALIDATION FOR THE LONGITUDINAL STUDY.

Cross Val	Subject B		Subject C		Subject E	
	Angular	SVM	Angular	SVM	Angular	SVM
1	0.95	1.07	0.71	0.88	1.06	0.64
2	1.02	1.142	0.71	0.77	1.01	0.67
3	0.79	0.81	0.75	0.8	1.04	0.68
4	1	1.01	0.8	0.78	0.98	0.66
5	1.12	0.97	0.83	0.83	0.98	0.72
6	1.07	0.86	0.707	0.94	1.22	0.66
7	0.88	0.94	0.82	0.87	1.09	0.62
8	0.83	0.91	0.73	0.92	1.12	0.75
9	0.92	0.73	0.78	0.82	1.04	0.53
10	1.04	1.05	0.79	0.78	0.96	0.63
Mean	0.96 ±	0.94 ±	0.76 ±	0.84 ±	1.05 ±	0.66 ±
± SD	0.10	0.12	0.04	0.06	0.08	0.05

The training sets in this setup are referred as 'personalized training database' in this paper. There are two reasons for this. As indicated earlier such a training set would eliminate the individual differences in pain expression. Second, our 'gold standard' or 'ground truth', the pain level provided by the subjects, is objective but with a behavioral bias. A personalized training database would eliminate that behavioral bias. Each subject's images were tested against the training set of the corresponding subject. With the personalized training database, we tested the classification algorithm with three distant measures: Euclidean distance, angular distance and support vector machine. Euclidean distance gave poor results and is not reported here. The mean absolute error for angular distance and support vector machine is shown in Table II. Subjects

A, D and F had only six images in the training set. As a result, the method did not work well and the results for subjects B, C and E are reported here. A 10 fold cross validation was performed

2) Sensitivity Analysis

For reproducible use in clinical settings it would be optimal to reduce the mean absolute error for pain level assessment. It is also important that the input and output pain distributions are similar or it may be possible that the system is always giving the same pain level as output but the mean absolute error is low. From a machine learning perspective, the system would perform well when the input pain level distribution is similar to that of the training data set. For a robust clinical decision support system we want the system performing well independent of the input pain level distribution. The sensitivity and specificity of each class (low, medium and high) of a 10 fold cross validation for subjects B, C, and E are shown in Table III.

B. Second phase--cross sectional study

Our findings from the first phase analyses could benefit from validation across a large numbers and different populations. Because of the individual differences in pain expression and behavioral bias from self-reported pain level data, a decrease in the system performance with non-longitudinal and different population data was expected. For the cross sectional study we had one image for each subject with a total of 513 subjects. We experimented with different dataset for the training. We found that when we used the entire dataset for the longitudinal study (454 images for six subjects) as the training database, we had a mean absolute error of 2.91. The sensitivity and specificity analysis is given in Table IV.

VI. DISCUSSION AND FINDINGS

A. Personalized model works better

The classification accuracy using the method works much better for the longitudinal study when we use the images of one person over a long time. Table II shows that we had a mean absolute error less than 1 for the longitudinal study. This proves the subjectivity of pain expression and reflects the behavioral bias for pain expression. We also found for the Eigenface method, angular distance and SVM gave similar result when used with the longitudinal dataset, but angular distance was better for the cross-sectional study (Table IV).

B. Distribution of pain level in the training set

The sensitivity and specificity varied much across different subjects and different training database. The primary reason for that is the lack of images of the representing class (low, medium and high) in the training dataset. For example, the sensitivity was very high for the class 'low' for subject C whereas for subject E, the sensitivity was high for the 'medium' pain level across each

cross validation (Table III). Analysis of the percentage of images with low and medium pain level in the training database explains this result. We had a high percentage of images with low pain level for subject C in the training set, resulting in better accuracy for low pain level for subject C (Figure 2).

C. Application scenario

It has been shown that pain measurement can be used in clinical settings for the improvement of the quality of life for cancer patients with pain using pain assessment tools such as Brief Pain Inventory (BPI) [5]. Consequently, automatic pain detection into three categories: low, medium and high has application in clinical settings. The primary goal is accurate and timely intervention for cancer patients with pain. One of the barriers to that is inadequate measurement of pain levels and such systems are of importance to address this problem.

VII. CONCLUSION

Automatic emotion detection from facial images is a challenging research problem and a significant amount of work has been done in this area during the last twenty years [15] [16]. The success of these application specific systems depends on narrowing down the context of the application and collecting enough data from specific settings [11]. In this paper we showed that a smart phone based tool can be used for remote monitoring of pain intensity for long term pain management with appropriate training dataset. Most work for pain detection involves the detection of pain or no pain. But in a clinical setting pain intensity is very important. The use of a mobile phone for pain intensity detection might reduce the healthcare costs and allow assessments in otherwise un-evaluable patients.

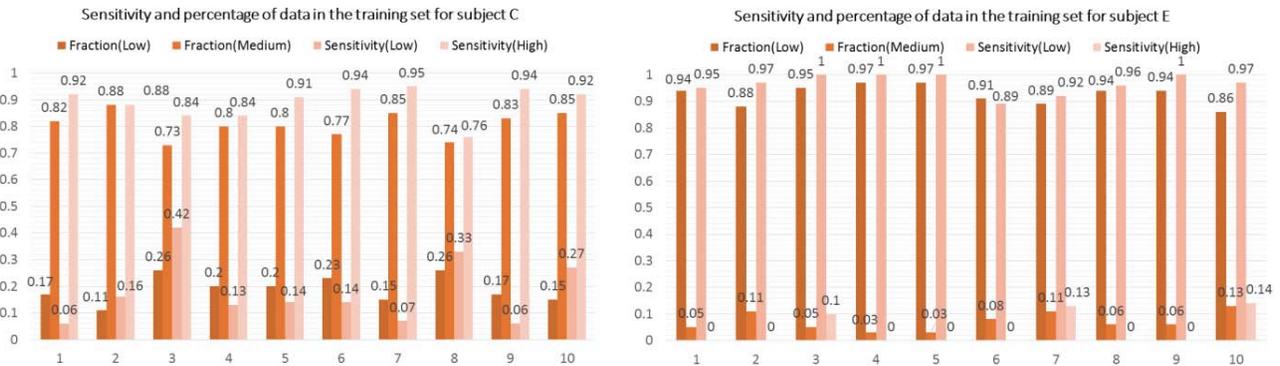
TABLE III. MEAN SENSITIVITY AND SPECIFICITY OF A 10 FOLD CROSS VALIDATION FOR THE LONGITUDINAL STUDY.

Subject	Angular						SVM					
	Sensitivity			Specificity			Sensitivity			Specificity		
	L(0-4)	M(5-7)	H(8-10)									
B	0.18	0.91	NaN	0.91	0.18	1	0.18	0.89	NaN	0.89	0.18	1
C	1	0	NaN	0	1	1	0.97	0.04	NaN	0.04	0.97	1
E	0.11	0.88	NaN	0.88	0.21	1	0.24	0.97	NaN	0.97	0.24	1
Mean ± SD	0.43±0.45	0.60±0.44	NaN	0.60±0.44	0.46±0.45	1 ± 0	0.46±0.37	0.60±0.43	NaN	0.63±0.43	0.46±0.37	1 ± 0

TABLE IV: SENSITIVITY AND SPECIFICITY FOR THE CROSS-SECTIONAL STUDY.

Angular						SVM					
Sensitivity			Specificity			Sensitivity			Specificity		
L	M	H	L	M	H	L	M	H	L	M	H
0.55	0.39	0.02	0.40	0.58	0.99	0	1	0	1	0	1

FIGURE 2. RATIOS OF THE NUMBER OF IMAGES FOR THE TWO DIFFERENT CLASSES (LOW AND MEDIUM) AND THE SENSITIVITY FOR EACH CLASS FOR THE 10 FOLD CROSS VALIDATION DURING THE LONGITUDINAL STUDY.



Further work is needed to address the issues of appropriate training set for target application, selection of the right algorithms. The usability of such systems with patients with chronic pain and the effect on the system performance due to ‘candid’ image and ‘acted’ image also needs to be investigated.

REFERENCES

- [1] R. Twycross and S. Fairfield, "Pain in far-advanced cancer," *Pain*, vol. 14, pp. 303-310, 1982.
- [2] S. Grossman and S. Nesbit, *Cancer Pain in Clinical Oncology*, 3rd Edition, Philadelphia: Elsevier, Philadelphia., 2004.
- [3] E. Au, C. Loprinzi, M. Dhodapkar, T. Nelson, P. Novotny, J. Hammack and J. O'Fallon, "Regular use of verbal pain scale improves the understanding of oncology inpatient pain intensity," *J Clin Oncol*, vol. 12, pp. 2751-2755, 1994.
- [4] E. Bruera, N. Kuehn, M. M., P. Selmsler and K. Macmillan, "The Edmonton Symptom Assessment System (ESAS): a simple method of the assessment of palliative care patients.," *J Palliative Care*, vol. 7, pp. 6-9., 1991.
- [5] C. Cleeland and K. Ryan, "Pain assessment: global use of the Brief Pain Inventory," *Annals Academy Med Singapore*, vol. 23 , no. 2, pp. 129-138, 1994.
- [6] M. M. Haque, F. Kawsar, M. Adibuzzaman, S. Ahamed and R. L. e. al., "Findings of e-ESAS: a mobile based symptom monitoring system for breast cancer patients in rural Bangladesh," in *CHI*, 2012.
- [7] K. Prkachin, "Assessing pain by facial expression: facial expression as nexus," *Pain Research & Management: The Journal of the Canadian Pain Society*, vol. 14, no. 1, pp. 53-58, 2009.
- [8] K. Prkachin and K. Craig, "Influencing non-verbal expressions of pain: Signal detection analyses," *Pain*, vol. 21, no. 4, pp. 399-409, 1985.
- [9] P. Ekman and W. Friesen, *Manual for the Facial Action Coding System*, Palo Alto: Consulting Psychologists Press, 1978.
- [10] K. Prkachin, "The consistency of facial expressions of pain: a comparison across modalities.," *Pain*, vol. 51, no. 3, pp. 297-306, 1992.
- [11] P. Lucey, J. Cohn, K. Prkachin, P. Solomon, S. Chen and I. Mathews, "Painful monitoring: Automatic pain monitoring using the UNBC-McMaster shoulder pain expression archive database," *Image and Vision Computing*, pp. 197-205, 2012.
- [12] K. Prkachin and P. Solomon, "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain," *Pain*, vol. 139, pp. 267-274, 2009.
- [13] M. Monwar and S. Rezaei, "Pain Recognition Using Artificial Neural Network," in *IEEE International Symposium on Signal Processing and Information Technology*, Vancouver, BC, 2006.
- [14] K. Boehner, R. DePaula, P. Dourish and P. Sengers, "How emotion is made and measured," *International Journal of Human-Computer Studies*, vol. 65, no. 4, 2007.
- [15] S. Calvo R. and D'Mello, "Affect Detection: An interdisciplinary Review of Models, Methods and Their Applications," *IEEE Transactions on Affective Computing*, vol. 1, no. 1, 2010.
- [16] Z. Zeng, M. Pantic, G. Roisman and T. Huang, "A survey of affect recognition methods: audio, visual and spontaneous expressions," *IEEE Transactions of pattern analysis and machine intelligence*, vol. 31, no. 1, 2009.
- [17] X. Zhu, "Face detection, pose estimation, and landmark localization in the wild," 2012.
- [18] S. Brahnam, L. Nanni and R. Sexton, "Introduction to neonatal facial pain detection using common and advanced face classification techniques," *Stud. Comput. Intel*, vol. 48, pp. 225-253, 2007.
- [19] S. Brahnam, C. Chuang, F. Shih and M. Slack, "Machine recognition and representation of neonatal facial displays of acute pain," *Artif. Intel. Med.*, vol. 36, pp. 211-222, 2006.
- [20] M. Monwar and S. Rezaei, "Appearance-based Pain Recognition from Video Sequences.," *IJCNN*, pp. 2429-2434, 2006.
- [21] M. Monwar, S. Rezaei and K. Prkachin, "Eigenimage Based Pain Expression Recognition," *IAENG International Journal of Applied Mathematics*, vol. 36, no. 2, 2006.
- [22] R. Niese, A. Al-Hamadi, A. Panning, D. Brammen, U. Ebmeyer and B. Michaelis, "Towards Pain Recognition in Post-Operative Phases Using 3D-based Features From Video and Support Vector Machines," *International Journal of Digital Content Technology and its Applications*, vol. 3, no. 4, pp. 21-31, December 2009.
- [23] D. Turk and R. Melzack, *Handbook of Pain Assessment*, New York, 2001.
- [24] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, 1991.
- [25] A. Ashraf, S. Lucey, J. Cohn, T. Chen, Z. Ambadar, K. Prkachin, S. P. and B. Eheobald, "The Painful Face - Pain Expression Recognition Using Active Appearance Models," in *ICMI*, 2007.
- [26] P. Becouze, C. Hann, J. Chase and G. Shaw, "Measuring facial grimacing for quantifying patient agitation in critical care.," *Computer Methods and Programs in Biomedicine*, vol. 87, no. 2, pp. 138-147, 2007.
- [27] B. Gholami, W. Haddad and A. Tannenbaum, "Relevance Vector Machine Learning for Neonate Pain Intensity Assessment Using Digital Imagine," *IEEE Trans. Biomed. Eng.*, 2010.
- [28] H. Hadjistavropoulos and K. Craig, "Acute and chronic low back pain: Cognitive, affective, and behavioral dimensions," *Journal of Consulting and Clinical Psychology*, vol. 62, no. 2, p. 341, 1994.
- [29] J. Raheja and U. Kumar, "Human facial expression detection from detected in capture image using back propagation neural network," *International Journal of Computer Science & Information Technology*, vol. 2, no. 1, pp. 116-123, Feb 2010.
- [30] A. Ashraf, S. Lucey, J. Cohn, T. Chen, K. Prkachin and P. Solomon, "The painful face II--Pain expression recognition using active appearance models.," *International Journal of Image and Vision Computing*, vol. 12, no. 27, pp. 1788-1796, November 2009.