Pakistan Journal of Science (Vol. 69 No. 1 March, 2017)

AN ALGORITHM FOR FACIAL EXPRESSION BASED AUTOMATIC DECEPTIVE PAIN DETECTION

S. Shafiq, H. Tauseef, M. A. Fahiem and S. Farhan

Department of Computer Science, Lahore College for Women University, Lahore, Pakistan
Corresponding author's email: humaiftikhar@hotmail.com

ABSTRACT: Facial expressions provide a primary source for detecting pain in humans. In present study an algorithm was designed to distinguish between real and posed pain expressions using a single image as input. Geometrical features of human face were used for pain detection. Different features from the extracted feature set were selected using different feature selection algorithms. Six different classification models were used to access the proposed algorithm. Training and testing was used to evaluate the selected classifiers. Algorithm developed provided accuracy (83%), sensitivity (82.9%) and specificity (78.1%) using all features with Instance-based k nearest neighbour classifier using cross validation.

Keywords: Deceptive Pain, Facial Expression Classification, Feature Extraction and Geometrical Features.

(Received 18-11-2016 Accepted 13-03-2017)

INTRODUCTION

Human behavior and responses related to pain and distress are used to analyze face expressions (Wang *et al.*, 2015; Lucey *et al.*, 2008 and Brahnam *et al.*, 2007), body or head movements (Hammal *et al.*, 2014; Ionescu *et al.*, 2004 and Esola *et al.*, 1996) and sound signals (Yang *et al.*, 2013; Sheinkopf *et al.*, 2012 and El Ayadi *et al.*, 2011). Human face is a major source of social interaction enriched with social information (Tomlinson *et al.*, 2010). People react in a different way and facial expressions differ for different intensities of pain (Gordon *et al.*, 2014). Mostly children or mentally impaired patients hide feeling of pain (Boerner *et al.*, 2013). There are special facial expressions that occur while experiencing real pain and some special expressions that occur while experiencing posed pain (Littlewort *et al.*, 2007).

Untrained human observers are not able to detect the deception level accurately. Human observers are trained to distinguish between real and posed pain expressions. Observing the timings and duration of specific pain expression is one of the methods for detecting fake pain (Wasner and Brock, 2008).

Some methods are based on the posed facial expressions obtained under controlled experimental conditions. Computer Expression Recognition Toolbox (CERT) is a tool to distinguish between real and posed facial expression (Bartlett *et al.*, 2014).

Pain intensity prediction can be done using self-report method (Sikka, 2014). Facial features get deformed when a person is suffering from pain. The focus is usually only on non-rigid deformation caused by pain (Rathee and Ganotra, 2015). Some contribution in the field of pain detection is made by using Scale Invariant Feature Transform (SIFT) and Speed Up Robust Feature (SURF) (Singh, 2015). Another approach is applied to learn the intensity rating function of posed pain expression. To rate the intensity of facial expression of posed pain multi domain comparative learning is used (Werner *et al.*, 2014).

Generally, frame by frame images from videos are used for detection of pain expression and fake pain expression (Rathee and Ganotra, 2015; Singh, 2015; Bartlett *et al.*, 2014 and Werner *et al.*, 2014). In present study, geometrical features extracted from single image were used to distinguish between real and posed pain.

MATERIAL AND METHODS

An algorithm for facial expression analysis was developed to distinguish between real and posed pain expression. The images were randomly taken from UNBC-McMaster shoulder pain database (Lucey *et al.*, 2011). Single image was taken as input. Researchers at McMaster University and University of Columbia captured videos of participant's faces suffering from shoulder pain were screened for selection. The database contained a total of 200 video sequences containing spontaneous facial expression comprising 48398 Facial Action Coding System (FACS) coded frames and 66 point Active Appearance Model (AAM) landmarks. Total images (n=365) were extracted from the database of both male and female (age 35-70 years). Colored images 352x240 pixels size with .jpg extension were used.

The proposed algorithm consisted of three main phases; pre-processing, feature extraction and classification. In the pre-processing phase, face detection was done by decomposing face into the regions of interest. In the feature extraction phase a feature set was formed using different features extracted from the multiple regions of interest. Selected features were then used to classify the data using different classifiers.

The detailed functioning of the proposed algorithm is given in **Figure 1**. The image was preprocessed for better results. Preprocessing included detection of face. The original image of 352 x 240 pixels was reduced to 186 x 235 pixels. Face was detected using Viola-Jones algorithm (Viola and Jones, 2004).

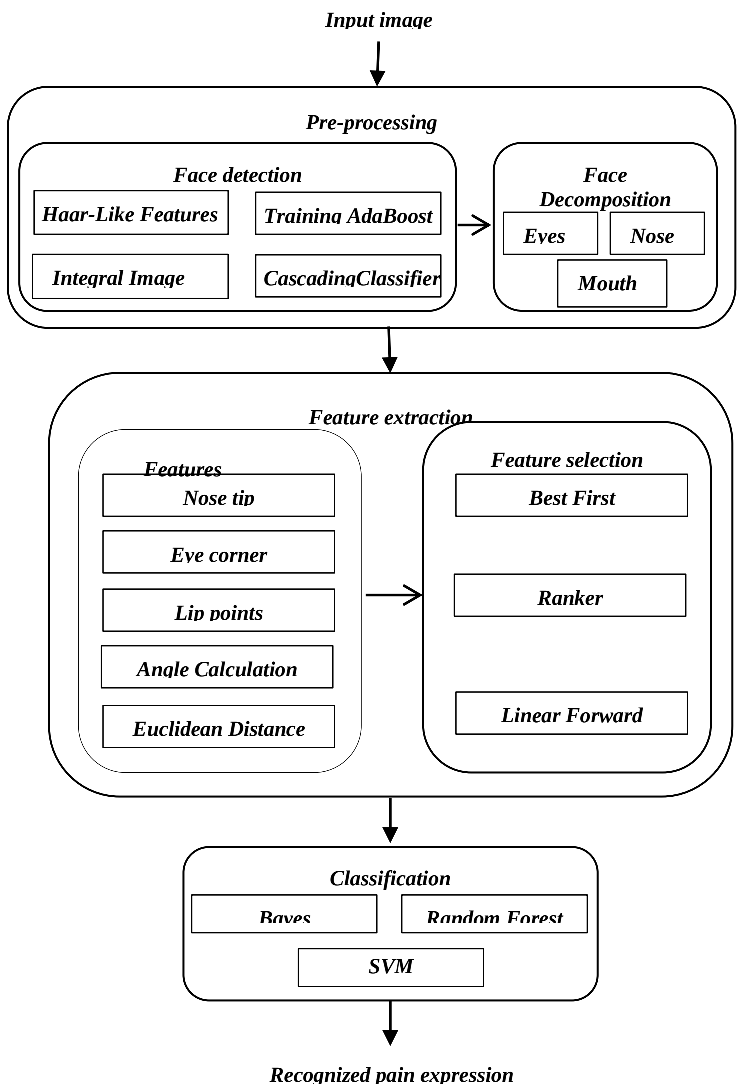

Fig-1: Detailed Diagram of the Proposed Algorithm

Using Haar-like features the similarities of human face i.e. eyes nose, mouth etc. were matched. A fixed size window was slided around the original image detecting the face and no-face region. It produced large number of samples to classify as face or no-face. Integral image was used to evaluate rectangular features. The sum of pixels above and to the left of (x,y), resulted in the integral image at location (x,y). AdaBoost was used for training boosted classifier using (**Equation 1**) $X_y(a) = \sum_{y=1}^{Y} x_y\ (a)$.

There were certain numbers of features in every stage of cascading. If it was classified as face it move on to the next stage and if it was classified as no-face on any of the stage it immediately discarded and did not move on to the next stage.

After detection of facial region, face was decomposed into eyes, nose and mouth. Eyes, mouth and nose were also detected using voila-johns algorithm. The decomposed face was then used for features extraction. Features extraction used edge information. A vector was assigned to each pixel that pointed to the closest edge pixel. The component x and y of vector was used to detect the eyes and mouth. Intensity information was used for eye center localization. Hue channel of lips were used for the detection of lips corner. (Asteriadis *et al.*, 2007). The geometrical features that were calculated included nose tip, eye corner and lip corner.

Different measurements were made from these features, such as length and width of left eye and right eye. For length and width calculation Euclidean Distance was used (**Equation 2**) $\sqrt{(\boldsymbol{a_1}-\boldsymbol{a_2})^2+(\boldsymbol{b_1}-\boldsymbol{b_2})^2}$ **.**

Eye and mouth opening and lip stretching were important features considered for pain detection. The normal eye opening was indicated using a threshold value. Values below the threhsold indicated the tight closing of the eyes. While in pain the lips of the person were stretched indicated by an increased distance between the lip corner points.

The angle between the nose tip and lip corner was calculated that determined the intensity of change in facial expression. The angle between the nose tip and left eye was calculated using (**Equation 3**) $\cos^{-1} {}^{base}/_{hyp}$.

The feature set was formed after extraction consisted of eighteen features; nose tip, both eyes corners, lips corners, length of both eyes, width of both eyes, mouth opening, lip stretch, nose to lip corners angles and nose to both eyes corners angles. Feature selection was done using different searching algorithms. The best features were selected from a set of given features. The searching algorithms used include Best First, Ranker and Linear Forward Selection. The distances and angles were then used for classification. Different classifiers were trained for the classification of dataset. The classifiers chosen for classification included instance-based k nearest neighbour (IBK), sequential minimal optimization (SMO), random forest, kstar, hyperpipes and Jrip. The images were classified using expressions as pain, no pain and fake pain.

Figure 2 illustrates the detailed image processing steps performed using the proposed algorithm. Euclidean distances were calculated for the extracted feature points. Angles were calculated using the extracted feature points. The extracted features and the calculated measurements were then used for classification and expression recognition of pain using different classifiers.

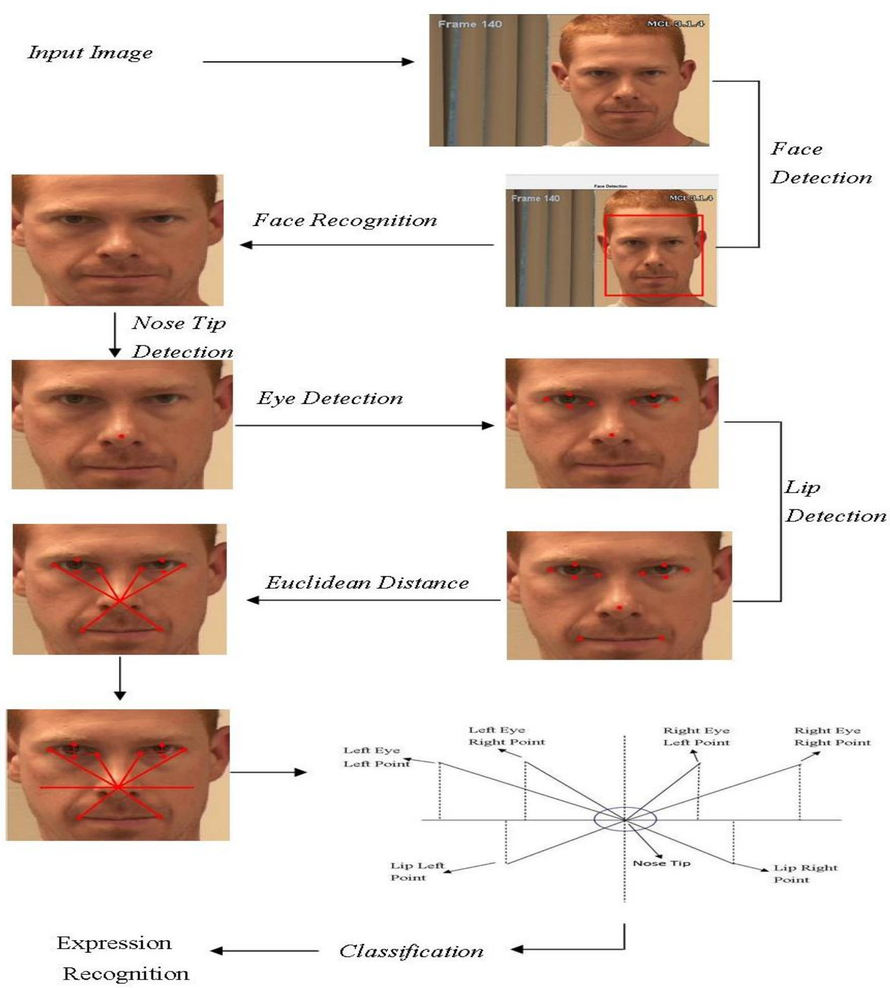

Fig-2: Detailed Image Processing Steps of the Proposed Algorithm

RESULTS AND DISCUSSION

The proposed algorithm correctly classified 83% of the instances when 15 folds cross validation was used. Specificity of 78% and sensitivity of 83% was achieved when IBK classifier was used with all features. When the best features were selected using genetic search; an accuracy of 72.55% was achieved using SMO classifier, sensitivity was 76.2% and specificity 44.2 %. When best features were selected using other searching algorithm including Best First, Ranker and linear search, the accuracy of 70.38%, sensitivity of 70.4% and specificity of 47.8% was achieved using SMO classifier (Table 1).

When best features were selected using Ranker, Exhaustive search and Linear Forward Selection, the accuracy of 70.65% was achieved using 10 fold cross validation with Random Forest classifier. Specificity of 42% and sensitivity of 70.7% was achieved. When best features were selected using genetic search, an accuracy of 72.82%, sensitivity of 72.8% and specificity of 49.3% was achieved using hyperpipes classifier. An accuracy of 81.2% was achieved using all features. Sensitivity and specificity of 81.3% and 74.2% were achieved respectively. Tabular representation of results obtained for accuracy, sensitivity and specificity achieved using different classifiers for 10 folds is given for reference (Table 2).

When the best features were selected using Genetic Search, an accuracy of 74.18%, sensitivity of 74.2% and specificity of 40.9% were achieved using 7 folds cross validation using the SMO classifier. Accuracy of 81.5% was achieved when IBK classifier was used with all features. Sensitivity of 81% and specificity of 78.2% were achieved. When features were selected using best first search and ranker search, an accuracy of 71.19%, specificity of 41.5% and sensitivity of 71.2% were achieved (Table 3).

When the percent split method was used, an accuracy of 79.2% was achieved with k-star classifier. Sensitivity of 79.2% and specificity of 70.7% were achieved. When the features were selected using genetic search, an accuracy of 74.4% was achieved with HyperPipes classifier. Sensitivity of 74.4% and specificity of 46.6% was achieved. When the features were selected using Best First and Linear Forward Selection an accuracy of 73%, sensitivity of 72.8%, and specificity of 30.1% was achieved using the Kstar classifier (Table 4).

It was evident from the results that the proposed algorithm achieved highest accuracy using all features. Accuracy of all features was greater than accuracy achieved by selection of features with genetic search and best first search i.e., 83% > 72.55% > 70.38 % using 15 fold cross validation.

Results achieved using the proposed research were compared to relevant well known researches conducted in the field. The comparison was made keeping in view four factors namely; input type, types of features, feature set size and accuracy. It was concluded that the proposed algorithm yielded better results using a smaller feature set. Also, the proposed algorithm used a single image for classification of deceptive pain with a feature set size of 18 to achieve an accuracy of 85%. On the contrary most of the approaches used video as an input and used multiple frames for classification that increased the processing complexity. Using multiple frames required using a very large feature set as compared to the proposed algorithm.

Some contribution in pain detection was made by using SIFT and SURF. SIFT produced the accuracy of 75.79% and SURF produced the accuracy of 72.63 (Singh, 2015). A feature set size of 290 was used for SIFT and 302 for SURF. The proposed algorithm's performed better as classification accuracy was higher (83%) using only a very few features (18) extracted from a single image.

Some researchers modeled the facial feature deformation using a technique called Thin Plate Spline (TPL). Distance Metric Learning (DML) technique was adopted to map the data into higher discriminative space. Accuracy of 96% by applying DML methodology and 79% without applying DML technology was achieved (Rathee and Ganotra, 2015). DML implicitly assumes that the examples in each class have a unimodal distribution. The margin constraints are designed to learn a distance metric under which all pairs of similarly labeled inputs are closer than all pairs of differently labeled inputs. Although the accuracy achieved by Rathee and Ganotra (2015) was better than that of the proposed algorithm but the proposed algorithm had the advantage that it used a very small feature set (18 features) extracted from a single image.

Another approach was to rate the intensity of facial expression of posed pain using multi domain comparative learning. The study mainly focused on comparative learning. 3D geometrical features were extracted from videos that increased the complexity. SVM classifier was used to classify the extracted features. The classification accuracy of 92.9 % was achieved using a feature set size of 16 per frame (Werner *et al.*, 2014). The proposed algorithm had the edge of using a small feature set and yet produced high accuracy rate.

Accuracy rates obtained for untrained and trained observers 51% and 55% respectively. After human observation, these videos were processed using a computer system called CERT and a correct classification rate of 85% was achieved (Bartlett *et al.*, 2014). The results were achieved using 16 features extracted from each frame of the video, hence considerably increasing the feature set size and complexity of the approach.

Table 1. Classification using Cross Validation 15 Folds Expressed as Percentage for Accuracy, Sensitivity and Specificity.

Classifier	Cross Validation (Fold 15)								
	All Features			Genetic Search			Best First		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
IBK	**83**	**82.9**	**78.1**	72.28	72.3	69	67.66	67.7	49.5
SMO	76.63	76.6	55.6	72.55	76.2	44.2	70.38	70.4	47.8
Random Forest	70.92	70.9	32	70.38	70.4	35.5	69.83	69.8	39.5
kstar	82.33	82.3	75.8	72.55	72.6	63.1	70.38	70.4	30.7
Hyperpipes	71.46	71.5	36.6	71.73	71.7	45.6	55.16	55.2	70.3
Jrip	66.03	66	40	68.75	68.8	35.9	70.38	70.4	31.9

Table: 2. Classification using Cross Validation 10 Folds Expressed as Percentage for Accuracy, Sensitivity and Specificity.

Classifier	Cross Validation (Fold 10)								
	All Features			Genetic Search			Best First		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
IBK	80.4	80.4	**75.5**	71.46	71.5	68.7	68.75	68.8	50.3
SMO	76.6	79.6	55	71.73	71.7	44.5	70.38	70.4	48.2
Random Forest	70.65	70.7	31.9	70.1	70.1	34	70.65	70.7	42
kstar	**81.25**	**81.3**	72.4	71.46	71.5	63.9	70.38	70.4	30.7
Hyperpipes	72.55	72.6	39.2	72.82	72.8	40.3	55.97	56	71
Jrip	70.38	70.4	44.8	70.38	70.4	38.1	70.38	70.4	31.9

Table: 3. Classification using Cross Validation 7 Folds Expressed as Percentage for Accuracy, Sensitivity and Specificity.

Classifier	Cross Validation (Fold 7)								
	All Features			Genetic Search			Best First		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
IBK	**81.5**	81.5	78.2	71.46	71.5	66.8	68.47	68.5	50.9
SMO	75.8	75.8	53.6	74.18	71.5	40.8	70.65	70.7	48.4
Random Forest	70.92	70.9	35.3	70.92	70.9	35.4	71.19	71.2	41.5
kstar	80.7	80.7	75.9	73.91	73.9	64.3	70.38	70.4	30.7
Hyperpipes	71.46	71.5	38.3	73.64	73.6	47.9	51.08	51.1	64.5
Jrip	69.39	67.4	41.3	68.47	68.5	35.2	70.38	70.4	31.4

Table: 4. Classification using Percent Split Expressed as Percentage Accuracy, Sensitivity and Specificity.

Classifier	Percent Split								
	All Features			Genetic Search			Best First		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
IBK	72.9	72.9	72.2	64.8	64.8	63.2	68	68	44.4
SMO	77.6	76.6	54.9	72	72	40.5	69.6	69.6	41.4
Random Forest	72	72	29.9	70.4	70.4	29.2	68	68	33.8
kstar	**79.2**	79.2	70.7	72.8	72.8	62.6	72.8	72.8	30.1
Hyperpipes	76	76	40.1	74.4	74.4	46.6	52	52	65.7
Jrip	72	72	51.9	73.3	73.6	46.6	72.8	72.8	30.1

Conclusion: It was concluded that IBK performed better classification of deceptive pain as compared to SMO, Random Forest, Kstar, Hyperpipes and Jrip. All geometrical features produced better classification accuracy as compared to selected geometrical features.

REFERENCES

Asteriadis, S., N. Nikolaidis, I. Pitas and M. Pardas (2007). Detection of facial characteristics based on edge information. *1st Int. Conf. VISAPP*, 247-252.

Bartlett, M.S., G.C. Littlewort, M.G. Frank and K. Lee (2014). Automatic decoding of facial movements reveals deceptive pain expressions. Curr Biol, 24(7): 738-743.

Boerner, K.E., C.T. Chambers, K.D. Craig, R.R.P. Riddell and J.A. Parker (2013). Caregiver accuracy in detecting deception in facial expressions of pain in children. Pain, 154(4): 525-533.

Brahnam, S., C.-F. Chuang, R.S. Sexton and F.Y. Shih (2007). Machine assessment of neonatal facial expressions of acute pain. Decis Support Syst, 43(4): 1242-1254.

El Ayadi, M., M.S. Kamel and F. Karray (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognit, 44(3): 572-587.

Esola, M.A., P.W. McClure, G.K. Fitzgerald and S. Siegler (1996). Analysis of lumbar spine and hip motion during forward bending in subjects with and without a history of low back pain. Spine, 21(1): 71-78.

Gordon, I., M.D. Pierce, M.S. Bartlett and J.W. Tanaka (2014). Training facial expression production in children on the autism spectrum. J Autism Dev Disord, 44(10): 2486-2498.

Hammal, Z., J.F. Cohn and D.T. George (2014). Interpersonal Coordination of HeadMotion in Distressed Couples. IEEE T Affect Comput, 5(2): 155-167.

Ionescu, A., E. Buchser, B. Rutschmann, B. Najafi and K. Aminian (2004). Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation. Gait Posture, 20(2): 113-125.

Littlewort, G.C., M.S. Bartlett and K. Lee (2007). Faces of pain: automated measurement of spontaneousallfacial expressions of genuine and posed pain, 9th Int. Conf. ICMI, 15-21.

Lucey, P., J.F. Cohn, K.M. Prkachin, P.E. Solomon and I. Matthews (2011). Painful data: The UNBC-McMaster shoulder pain expression archive database, 9th Int. Conf. IEEE FG2011, 57-64.

Lucey, P., J. Howlett, J. Cohn, S. Lucey, S. Sridharan and Z. Ambadar (2008). Improving pain recognition through better utilisation of temporal information. Int.Conf. AVSP, 167.

Rathee, N. and D. Ganotra (2015). A novel approach for pain intensity detection based on facial feature deformations. J Vis Commun Image Represent, 33: 247-254.

Sheinkopf, S.J., J.M. Iverson, M.L. Rinaldi and B.M. Lester (2012). Atypical Cry Acoustics in 6-Month-Old Infants at Risk for Autism Spectrum Disorder. Autism Res, 5(5): 331-339.

Sikka, K. (2014). Facial Expression Analysis for Estimating Pain in Clinical Settings. 16th Int. Conf. ICMI, 349-353.

Singh, S.K. (2015). SIFT and SURF performance evaluation for pain assessment using facial expressions. J Biol Eng Res Rev, 2(1): 6-14.

Tomlinson, D., C.L. von Baeyer, J.N. Stinson and L. Sung (2010). A systematic review of faces scales for the self-report of pain intensity in children. Pediatrics: peds. 126(5): 1168-1198.

Viola, P. and M.J. Jones (2004). Robust real-time face detection. Int J Comput Vision, 57(2): 137-154.

Wang, S., C. Eccleston and E. Keogh (2015). The role of spatial frequency information in the recognition of facial expressions of pain. Pain, 156(9): 1670-1682.

Wasner, G.L. and J.A. Brock (2008). Determinants of thermal pain thresholds in normal subjects. Clin. Neurophysiol., 119(10): 2389-2395.

Werner, P., A. Al-Hamadi and R. Niese (2014). Comparative learning applied to intensity rating of facial expressions of pain. Int J Pattern Recogn, 28(05): 1451008.

Yang, Y., C. Fairbairn and J.F. Cohn (2013). Detecting depression severity from vocal prosody. IEEE T Affect Comput, 4(2): 142-150.