Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems

Değer Ayata¹ · Yusuf Yaslan¹ · Mustafa E. Kamasak¹

Received: 11 July 2017 / Accepted: 17 December 2019 / Published online: 7 January 2020
© The Author(s) 2020

Abstract

Purpose The purpose of this paper is to propose a novel emotion recognition algorithm from multimodal physiological signals for emotion aware healthcare systems. In this work, physiological signals are collected from a respiratory belt (RB), photoplethysmography (PPG), and fingertip temperature (FTT) sensors. These signals are used as their collection becomes easy with the advance in ergonomic wearable technologies.

Methods Arousal and valence levels are recognized from the fused physiological signals using the relationship between physiological signals and emotions. This recognition is performed using various machine learning methods such as random forest, support vector machine and logistic regression. The performance of these methods is studied.

Results Using decision level fusion, the accuracy improved from 69.86 to 73.08% for arousal, and from 69.53 to 72.18% for valence. Results indicate that using multiple sources of physiological signals and their fusion increases the accuracy rate of emotion recognition.

Conclusion This study demonstrated a framework for emotion recognition using multimodal physiological signals from respiratory belt, photo plethysmography and fingertip temperature. It is shown that decision level fusion from multiple classifiers (one per signal source) improved the accuracy rate of emotion recognition both for arousal and valence dimensions.

Keywords Physiological data · Emotion recognition · Multi-sensor data fusion

1 Introduction

Cloud computing and wearable technologies have started to transform traditional healthcare systems. The focus of healthcare systems has been shifting from treatment to prevention with efficient and reliable personalization. The data collected through wearable sensors can be enormous, and can bring new opportunities in many fields such as medical care, fitness, aging, and finance [1]. Wearable computers can monitor various physiological signals such as heart rate, blood pressure, temperature, location and movements [2]. These physiological signals have been used for various

e-health applications. E-health is the healthcare practice supported by electronic processes and communications [3]. A common practice in e-health is using wearable devices and mobile applications as sensors. Fitness and health monitoring are amongst the most popular applications for self-monitoring health devices. Self-monitoring devices combined with e-health applications may enhance personalized preventive medicine. For high-quality service, scalability and availability of data is a key issue that can be supplied by cloud computing systems.

One of our motivations is to design an emotion recognition system that can be incorporated into a health monitoring system for elderly people living alone. An increasing number of elderly people require continuous monitoring of their health and emotion states. Seafarers, long-distance drivers and pilots may also benefit from the proposed system.

Assisting health monitoring systems with emotional data may also improve overall service quality. The proposed framework is a wearable integrated cloud computing based emotion-aware healthcare system. The framework collects respiration, plethysmograph and fingertip temperature

✉ Değer Ayata
deger.ayata@aiatus.com; ayatadeğer@itu.edu.tr

Yusuf Yaslan
yyaslan@itu.edu.tr

Mustafa E. Kamasak
kamasak@itu.edu.tr

¹ Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey

data from the subjects, and sends these signals to the cloud via mobile networks. Furthermore, it uses machine learning algorithms to process the collected data and to predict the physiological or psychological conditions of subjects. Sharing this information with a contextual receiver, e.g., a healthcare organization is also possible with the proposed framework. There are various studies on emotion recognition using physiological signals. Picard et al. used peripheral neurons system including heart rate variations, skin conductivity and respiration [4]. Li et al. also proposed an emotion recognition technique based on a method that uses three-channel bio-sensors to measure users' electrocardiogram, galvanic skin response and photoplethysmography [5].

Chen et al. proposed an emotion-aware mobile cloud computing framework [6]. They proposed an infrastructure with a mobile terminal where physiological data were collected. A local cloudlet preprocessed the data to increase quality or reduce data size. Finally, a remote cloud where the analysis of all data was performed to achieve high accuracy of emotion recognition. For each user, a personal classifier is built based on historical data associated with the user's emotion labels [6]. Chen et al. also proposed a wearable-computing assisted system for health monitoring and remote medical care. Their framework had three components which are collaborative data collection via wearables, sentiment analysis, and affective interactions. In addition, they proposed a cloud based approach to be able to make computation intensive analysis of data faster, and allocate resources dynamically [7]. Tokuno et al.'s study can provide insight about the relationship between emotion recognition and healthcare. They did not involve a cloud layer in their work but investigated the usage of emotion recognition for military healthcare. They detected emotional changes from natural speaking voice and determine if the changes are related to mental stress. They represented four different emotion classes. This study found out that the duration of soldiers' field stay affected their values for the "joy" and "sorrow" classes. Their results demonstrated that emotion recognition can be used for mental status monitoring and mental healthcare [8]. Lu et al. proposed a prototype cloud computing based healthcare information system for elderly people. Their simulation environment consisted of a three-way data transferring pipeline centered on a cloud sharing service. Hospital, healthcare service, and homestay were defined as the three sources of elderly healthcare information. However, the prototype does not involve any smart physiological signal generation or a mobile device terminal. They concerned only with transferring and storing data on cloud [9]. Ji et al. proposed a mobile healthcare system based on cloud computing techniques. The system used a middleware on the user side to provide a plug-and-play environment for different combinations of wireless sensors and mobile

terminals to utilize different communication protocols. Architecture also included an emergency first aid system [10]. Henia et al. studied human emotion recognition using electrocardiogram, galvanic skin response, skin temperature and respiration volume [11].

Koelstra et al. [12] presented a multimodal dataset (DEAP) for the analysis of human affective states. The dataset contains 32 EEG channels and physiological signals. Rather than EEG we aimed to focus on different modalities and decrease number of EEG channels/sensors as the collection of 32 channels EEG signals has substantial limitations in practical use.

In our first study on this topic, we investigated emotion recognition only from GSR signals [13, 14]. Later, we used PPG and GSR signals together and proposed a data fusion based emotion recognition method for music recommendation engines [15]. This wearable music recommendation framework utilizes not only the user's demographics but also his/her emotion state at the time of recommendation.

In this paper, we used RB, PPG and FTT signals simultaneously. In the mentioned previous studies, the combination of RB, PPG, and FTT signals has never been used together in a similar framework. Also, we have used decision level fusion as opposed to feature level fusion [15].

2 Emotion Representation

Two models are commonly used for emotion representation by psychologists: the dimensional model and the categorical (discrete) model. According to the dimensional model, people's emotions can be represented with a limited number of independent affective dimensions. Two important dimensions are arousal and valence. Arousal points out the intensity of emotion and valence marks the polarity of emotion as either negative or positive [16–19].

Arousal is related to activity in both our mind and body. It indicates the intensity of being awake and alert. Organs are stimulated with perception, and arousal is related to the influence of stimulation. When human beings are stimulated, they become aroused. At higher degrees of stimulation, humans are more aroused. Arousal indicates how calming/soothing or exciting/agitating the stimulus is. The arousal-nonarousal scale measures how energized or soporific one feels. Stress, anxiety, anger, and fear are high-arousing emotions whereas soothing and calming are low-arousal emotions [18]. The valence term indicates the emotional value that is associated with a certain stimulus and used to characterize/categorize specific emotions. Valence broadly refers to the positive and negative character of emotion or some of its aspects. Anger and fear are amongst the emotions with "negative valence" whereas joy has "positive valence".

The valence-arousal dimensional model, represented in Fig. 1, is widely used for emotion representation [13, 14, 18, 20].

3 Physiological Signals

3.1 Respiratory Belt (RB)

Respiratory belt (RB) (or respiration belt) is a sensor that captures the breathing activity of subjects. It can be worn on either thoracic or abdominal areas. The amount of stretch in the elastic part is measured and recorded for monitoring respiration rate or depth of breath. Emotionally, respiration rate typically decreases with relaxation. On the other hand, negative oriented emotions generally cause irregularity in the respiration pattern. In extreme tense cases, respiration may cease for a short duration.

3.2 Photo Plethysmography (PPG)

Plethysmograph measures the changes in volume within an organ or the whole body using the plethysmography technique. Photo plethysmography (PPG) is a process of applying a light source and measuring the transmitted or reflected light. The changes in light intensity are associated with small variations in blood perfusion of the tissue and provide information on the cardiovascular system, in particular, the pulse rate [21, 22]. The sensor system typically consists of a red and infrared (IR) source and detectors.

Measurements taken from a plethysmograph can be used to compute other physiological parameters such as heart rate

(HR), inter-beat periods, and heart rate variability (HRV) that are highly correlated with the emotional states.

3.3 Fingertip Temperature (FTT)

Fingertip temperature (FTT) is a sensitive signal that can be used to monitor the state of relaxation. In a relaxed mood, vessels are dilated and the fingertip becomes warmer. In anxious or tense moods, the vessels shrink and fingertip becomes cooler. The change in fingertip temperature can be recorded and used for emotion recognition.

4 Data Analysis and Emotion Recognition

4.1 Dataset

In our study, we have used an existing DEAP multimodal dataset [12]. This dataset contains peripheral physiological (Galvanic Skin Response, Respiratory Belt, Plethysmograph, and Fingertip Temperature) and EEG signals of 32 participants for the analysis of human affective states. In this dataset, each subject was shown 40 different 1-min long excerpts of music videos. Subjects' EEG and peripheral physiological signals were recorded as they watched the videos. After watching each video, subjects rated their levels of arousal and valence. The total signal record time for each video is 63 s. The data was downsampled to 128 Hz, electrooculogram (EOG) artifacts were removed, a bandpass frequency filter from 4.0 to 45.0 Hz was applied, and the data was segmented into 60 s trials and a 3 s pre-trial. A sampling frequency of 128 Hz means 8064 data points per channel.

In this study, we have focused on only RB, PPG, and FTT signals. We have not considered EEG and GSR signals in order to decrease number of sensors. We have extracted and used signals from RB, PPG, and FTT among the recorded channels.

4.2 Feature Extraction

In order to represent physiological signals, each record has been divided into varying length moving windows and statistical features have been extracted in the temporal domain. Each 63 seconds signal has been divided into smaller windows (e.g. 21 segments with 3 s per window) as shown in Fig. 2.

Each feature is extracted from signal points using statistical functions as shown in Eq. 1.

$$f_i^j = \varphi_i(\mathbf{x}^j) = \varphi_i(x_1^j, x_2^j, x_3^j, \dots, x_K^j) \quad (1)$$

where

Fig. 1 Valence—arousal model for emotion representation

Fig. 2 Feature extraction using time windows

φ_i : function from $\mathbb{R}^K$ to $\mathbb{R}$, where φ_i stands for one of features listed in Table 1.

x_k^j : input signal for sensor j for sample k where $j \in \{1, 2, 3\}$ and $k \in \{1, 2, \dots, K\}$

f_i^j : feature extracted from input signal in a sub window for sensor j

$\mathbf{x}^j$: input vector that consists of sub window points for sensor j

N : Number of total sample signal points

K : Number of sample points extracted from input signal in a signal sub window where $K \leq N$

Table 1 lists extracted features and their corresponding formulas [14]. From this Table, arithmetic mean value,

maximum value, minimum value, standard deviation, variance, skewness, kurtosis, median, zero crossings, entropy, mean energy, moments, and change in signal values have been considered as features. Features have been extracted from each window, and their values across consecutive windows have been concatenated for each subject and each video. Concatenated features form the feature vector.

4.3 Feature Selection

Feature selection is a process in which features with more contribution to classification performance are selected and the rest are dropped. Hence, the feature selection process reduces the dimension of the feature vector and eliminates

Table 1 List of 22 features extracted from data and their formula. In this table sgn , $\oplus$, $P(\cdot)$ denotes sign, xor operator and probability respectively. S_b and S_w denotes the feature scatter between and within emotions

Feature	Formula	Feature	Formula
Minimum	$\min\{x_k\}$	Skewness	$\sum_{k=1}^K \frac{(x_k - \mu)^3}{(K-1)SD^3}$
Maximum	$\max\{x_k\}$	Kurtosis	$\sum_{k=1}^K \frac{(x_k - \mu)^4}{(K-1)SD^4}$
Arithmetic mean (μ)	$\frac{1}{K} \sum_{k=1}^K x_k$	Median	$\frac{1}{2} \left(\frac{K}{2} + \frac{K+1}{2} \right) \text{ or } \left(\frac{K+1}{2} \right)$
Mean absolute value	$\frac{1}{K} \sum_{k=1}^K x_k $	Moment m ($m \in \{3, 4, 5, 6\}$)	$\frac{1}{K} \sum_{k=1}^K x_k^m$
Mean energy	$\frac{1}{K} \sum_{k=1}^K x_k^2$	Entropy	$\sum_{k=1}^K P(x_k) * \log_2 P(x_k)$
Root mean square	$\sqrt{\frac{1}{K} \sum_{k=1}^K x_k^2}$	First degree difference (diff_1)	$\frac{1}{K-1} \sum_{k=1}^K x_{k+1} - x_k $
Standard deviation (std)	$\sqrt{\frac{1}{K} \sum_{k=1}^K (x_k - \mu)^2}$	Second degree difference (diff_2)	$\frac{1}{K-2} \sum_{k=1}^K x_{k+2} - x_k $
Zero crossing	$\sum_{k=1}^K \text{sgn}(x_k) \oplus \text{sgn}(x_{k-1})$	Normalized k th degree difference ($k \in \{1, 2\}$)	diff_k/std
Mean absolute deviation	$\frac{1}{K} \sum_{k=1}^K x_k - \mu $	Maximum scatter difference (msd)	$\max_{w \geq 0} w^T (S_b - S_w) w$

irrelevant, noise, and redundant features. In this study, we have used mRMR (minimum redundancy maximum relevance) algorithm.

Before application of mRMR, the values of features are normalized to $[-1, 1]$ range using min-max normalization. Minmax is a normalization approach that linearly transforms f value to f_{norm} value using Eq. 2 where f_{max} and f_{min} are the maximum and minimum values in the feature vector $\mathbf{f}$.

$$f_{norm,i} = 2 \frac{f_i - f_{min}}{f_{max} - f_{min}} - 1 \quad i = 1, \dots, L \quad (2)$$

mRMR uses mutual information (MI) for feature selection [23, 24]. Mutual information difference (MID) and mutual information quotient (MIQ) are the two most used mRMR schemes. It selects a feature subset that characterizes the statistical property of target classification.

In order to find an appropriate feature set, different combinations of features have been selected as feature set and relationship between arousal and valence has been studied. We have obtained four main feature sets (FS) based on ranking scores. We have assigned features to feature sets according to their mRMR ranking scores. Feature Set K (FS- K) contains top ranking K features for $K \in \{10, 14, 18, 22\}$. Hence, FS- K_1 is covered by FS- K_2 (FS- K_1 FS- K_2) when $K_1 < K_2$.

Figure 3 shows feature scores of the mRMR algorithm and Table 2 lists the features included within each feature subset.

4.4 Emotion Recognition Using Classifiers

The main task of classification is to find the label or category of an input represented by its feature vector. The classification process has a learning (training) step. In the training step, the classifier parameters are selected based on the training set.

Since we represent emotion with arousal and valence dimensions we have two categorical output values. Labeling the samples is required in supervised machine learning methods. In our study, we have labeled arousal and valence levels as low and high classes based on their values. In the DEAP dataset, arousal and valence values are in the (1–9) range. Rating values greater than or equal to 4.5 are assigned to label high, whereas values lower than 4.5 are assigned to label low.

$$\begin{aligned} M^j &= \Phi(\mathbf{f}^j) = \Phi(f_1^j, f_2^j, f_3^j, \dots, f_L^j) \\ &= \Phi(\varphi_1(\mathbf{x}^j), \varphi_2(\mathbf{x}^j), \varphi_3(\mathbf{x}^j), \dots, \varphi_L(\mathbf{x}^j)) \end{aligned} \quad (3)$$

M^j : Machine learning output for sensor $j \in \{1, 2, 3\}$

Φ : Classifier that is fed with features

Fig. 3 Feature Scores: each feature mRMR is shown for respiratory belt signal Arousal label

Table 2 Feature subsets

Feature subset	Features
Feature 10 (FS-10)	Second order difference, variance, moment 4, skewness, kurtosis, first order difference, moment 6, normalized second degree difference, zero crossing, arithmetic mean
Feature 14 (FS-14)	FS-10, mean absolute energy, minimum, maximum, mean absolute deviation
Feature 18 (FS-18)	FS-14, moment 5, mean absolute value, median, maximum scatter difference
Feature 22 (FS-22)	FS-18, normalized first degree difference, root mean square, moment 3, standard deviation

$\mathbf{f}^j$: feature vector for sensor j

L : number of features where $L \in \{10, 14, 18, 22\}$

Ψ : voting function, which returns the most repeated label

Random forest (RF), support vector machine (SVM) and logistic regression (LR) algorithms are used as classifiers for emotion recognition. Each classifier Φ is fed with extracted features as shown in Eq. 3. Best results are obtained with RF classifier which is an ensemble method with which classification is performed using multiple decision trees [25]. A decision tree is a non-parametric supervised learning method that predicts the value of a target variable by learning decision rules from the data. Decision tree partitions the dataset into groups as homogeneous as possible in terms of the variable to be predicted. Entropy and information gain are used to process attribute selection [26]. Forest of decision trees is constructed using a random subset of features, thus RFs have the ability to cope with high dimensional data.

During the training of the RF classifiers, many parameters are optimized to achieve the highest performance on training data. The feature extraction (window size, window overlap, etc.), feature selection (number of features), and other hyper-parameters of the classifiers are optimized in this process. The optimum parameters are listed in Table 3.

4.5 Decision Level Fusion

The main objective of employing fusion is to produce a fused result that provides the most detailed and reliable information possible. Fusing multiple information sources also produces a more efficient representation of the data. Lahat et al. [27, 28] defines data fusion as the analysis of

several datasets such that different datasets can interact and inform each other. Data fusion techniques combine data from different sources together. Data fusion is categorized as feature level and decision level. Feature level fusion requires the extraction of different features from the source data, before the features are merged. A feature-level fusion scheme integrates unimodal features before learning concepts as given in Eq. 4.

$$\mathbf{f} = [\mathbf{f}^1, \mathbf{f}^2, \mathbf{f}^3] \quad (4)$$

$$M = \Phi(\mathbf{f})$$

The main advantage of this scheme is the use of only one learning stage and the benefit of mutual information from data. Decision-level fusion, or fusion of classifiers, consists of processing the classification results of prior classification stages by combining the results from multiple algorithms to yield a final fused decision. The main goal of this procedure is to take advantage of the redundancy of a set of independent classifiers to achieve higher robustness by combining their results [29] as in Eqs. 5 and 6.

$$M = \Psi(\Phi(f^1), \Phi(f^2), \Phi(f^3)) \quad (5)$$

$$M = \Psi(M^1, M^2, M^3) \quad (6)$$

M : global machine learning model

Ψ : voting function, which returns the most repeated M^i label

In our study, we used decision-level fusion. In decision-level classification fusion, we classified the modalities individually and then combined the classifier outputs as shown in Fig. 4. These results are obtained using the optimum random forest parameters listed in Table 3.

Table 3 Selected classifier and optimized hyper-parameters for RB, PPG, and FTT signals

Data source	Classifier	Feature set	Window duration	Window overlap
RB	Random forest	FS-14	3 s	Overlap
PPG	Random forest	FS-10	8 s	No-overlap
FTT	Random forest	FS-10	5 s	Overlap

Fig. 4 Emotion recognition pipeline

Table 4 Single modality and decision level fusion accuracy results

Signal Source	Arousal Accuracy %	Valence Accuracy %
Respiratory belt	71.32	70.62
Photo plethysmography	68.59	70.23
Finger tip temperature	69.68	67.73
Fused	73.08	72.18

5 Emotion Recognition Results and Discussion

In order to determine the performance of the emotion recognition system robustly, we have used leave one subject out-cross validation (LOSO-CV) method. In LOSO-CV, the data is partitioned into 32 subsets, where each subset includes an individual subject's data. Of the 32 subject subsets, 31 subsets are used for training the model and the remaining one is used for testing. This process is repeated 32 times with a different test subset. Accuracies from 32 tests are averaged to determine the performance of the emotion recognition system. We have conducted LOSO based tests both for arousal and valence dimensions.

We performed tests using each physiological signal separately. The accuracy values corresponding to different physiological signals are listed in Table 4.

Using decision level fusion described in Sect. 4.5, accuracy improved from 69.86 to 73.08% for arousal, and from 69.53 to 72.18% for valence. Results indicate that using multimodal physiological signal sources improved accuracy rate compared to using modalities separately. This improvement is shown in Fig. 5.

Recognizing arousal and valence values directly from biosensors is a challenging task. We showed that there is relationship between RB, PPG and FTT signals and arousal and valence. Results indicate that decision level fusion of multi modal data yields an improvement in the accuracy rate in both arousal and valence dimensions.

This work has certain limitations. Firstly, the number of subjects used in the tests should be increased for more robust results. Also, the system performance is dependent on the handcrafted feature set. We have used basic descriptive statistics to extract features from signals. However, a broader feature set can also be considered and the effect of each feature also can be analyzed.

An application of the emotion recognition system proposed in this work is an emotion-aware health monitoring system for elderly people that live alone. An increasing number of elderly people require continuous monitoring of their health and emotional states. A remote monitoring system may allow an early action in case of dangerous

Fig. 5 Accuracy of both arousal and valence improved by using data from all modalities simultaneously via decision level fusion

situations for elderly persons' lives (depression, anxiety, state before a heart attack or a stroke, etc.).

Seafarers, long-distance drivers, and pilots may also benefit from an emotion aware monitoring system. Pilots must control their emotions to ensure flight safety [30]. Similarly, drivers are expected to keep control of their emotions, and know how to deal with emotions. Such a monitoring system can be integrated to an airplane, car/truck or ship cabin. Alternatively, pilots, seafarers, and long-distance drivers may be screened before the beginning of each task. Screening operators of critical systems is an active research area [31–34].

6 Conclusion

This study demonstrated a framework for emotion recognition using multimodal physiological signals from respiratory belt, photo plethysmography and fingertip temperature. It is shown that decision level fusion from multiple classifiers (one per signal source) improved the accuracy rate of emotion recognition both for arousal and valence dimensions.

The results also indicate that emotion recognition from physiological signals is still a challenging task. Although, emotion recognition accuracy is improved with three input sources in this study, more sources might be required for more robust and reliable recognition such as EEG, ECG, etc. However, this brings a trade-off between an ergonomic data collection system and emotion recognition accuracy. We preferred the RB, PPR and FTT as these physiological signals are easy to collect with relatively compact and ergonomic wearable devices.

The proposed emotion recognition system can be used for emotion aware health monitoring. This may enable systems

to generate healthcare actions from the subject's temporal emotion changes.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Jhajharia, S., Pal, S. K., & Verma, S. (2014). Wearable computing and its application. *International Journal of Computer Science and Information Technology*, 5(4), 5700–5704.
- Popat, K. A. & Sharma, P. (2013). Wearable computer applications: A feature perspective. *International Journal of Engineering and Technology Innovation*, 3(1). <https://pdfs.semanticscholar.org/ec9c/33e897e38e691f60baf828303a0ae64e1899.pdf>.
- Della Mea, V. (2001). What is e-health: The death of telemedicine? *Journal of International Medical Research*, 3(2), e22.
- Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175–1191.
- Li, C., Chao, X., & Feng, Z. (2016). Analysis of physiological for emotion recognition with the irs model. *Neurocomputing*, 178, 103–111.
- Chen, M., Zhang, Y., Li, Y., Hassan, M. M., & Alamri, A. (2015). Aiwac: Affective interaction through wearable computing and cloud technology. *IEEE Wireless Communications*, 22(1), 20–27.
- Chen, M., Zhang, Y., Li, Y., Mao, S., & Leung, V. C. M. (2015). Emc: Emotion-aware mobile cloud computing in 5g. *IEEE Network*, 29(2), 32–38.
- Tokuno, S., Tsumatori, G., Shono, S., Takei, E., Suzuki, G., Yamamoto, T., Mituyoshi, S., & Shimura, M. (2011). Usage of emotion recognition in military health care. In *Defense science research conference and expo (DSR)* (pp. 1–5). IEEE
- Lu, F., & Khan, I. (2012). Sharing elderly healthcare information on cloud computing
- Zhanlin, J., Ivan, G., O'Droma, M., Xin, Z., & Xueji, Z. (2014). A cloud-based x73 ubiquitous mobile healthcare system: design and implementation. *The Scientific World Journal*. <https://www.hindawi.com/journals/tswj/2014/145803/ref/>.
- Wiem, M.B.H., & Lachiri, Z. (2016). Emotion assessing using valence-arousal evaluation based on peripheral physiological signals and support vector machine. In *2016 4th International Conference on Control Engineering & Information Technology (CEIT)* (pp. 1–5). IEEE
- Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, Ashkan, Ebrahimi, Touradj, et al. (2012). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31.
- Deger, A., Yusuf, Y., & Kamasak, M. (2016). Emotion recognition via galvanic skin response: Comparison of machine learning algorithms and feature extraction methods. In *Editorial to the Special Issue on Selected Papers from Medical Technologies Conference* (vol. 3093, p. 3129).
- Ayata, D., Yaslan, Y., & Kamasak, M. (2016). Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches. *National Congress on Medical Technologies, 2016*, 1–4.
- Ayata, D., Yaslan, Y., & Kamasak, M. E. (2018). Emotion based music recommendation system using wearable physiological sensors. *IEEE Transactions on Consumer Electronics*, 64, 196.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145.
- Mehrabian, A. (1968). Communication without words. *Psychology today*, (pp. 53–56).
- Mehrabian, A. (1980). *Basic Dimensions for a General Psychological Theory Implications for Personality, Social, Environmental, and Developmental Studies*. Gunn: Oelgeschlager.
- Lang, P.J.: International affective picture system (iaps): Affective ratings of pictures and instruction manual. Technical report (2005)
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Cambridge: The MIT Press.
- Tamura, T., Maeda, Y., Sekine, M., & Yoshida, M. (2014). Wearable photoplethysmographic sensors-past and present. *Electronics*, 3(2), 282–302.
- Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2006). Emotion assessment: Arousal evaluation using eeg's and peripheral physiological signals. *Multimedia content representation, classification and security* (pp. 530–537).
- Peng, H., Ding, C., & Long, F. (2005) Minimum redundancy maximum relevance feature selection (MRMR). *IEEE Intelligent Systems*, 20. https://www.researchgate.net/publication/237133616_Minimum_Redundancy_-_Maximum_Relevance_Feature_Selection. Accessed May 2018.
- Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), 1226–1238.
- Cho, S., & Seo, Y. (2013). Extraction of significant features using empirical mode decomposition and its application. In *Proceedings of the World Congress on Engineering and Computer Science* (vol. 1).
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Francisco: Morgan Kaufmann Publishers.
- Dana, L., Adali, T., & Jutten, C. (2014) Challenges in multimodal data fusion. In *European signal processing conference (EUSIPCO)* (pp. 101–105). IEEE
- Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal data fusion: An overview of methods, challenges, and prospects. *Proceedings of the IEEE*, 103(9), 1449–1477.
- Santiago, P., & Iriondo, I. (2012). Comparison between decision-level and feature-level fusion of acoustic and linguistic features for spontaneous emotion recognition. In *Iberian Conference on Information System and Technology (CISTI)* (pp. 1–6). IEEE.
- von dem Knesebeck, O., Mnich, E., Angermeyer, M. C., Kofahl, C., & Makowski, A. (2015). Changes in depression stigma after the germanwings crash-findings from German population surveys. *Journal of Affective Disorders*, 186, 261–265.
- Hoste, V., Van Cynthia, H., & Poels, K. (2016). Towards a framework for the automatic detection of crisis emotions on social media: A corpus analysis of the tweets posted after the crash of germanwings flight 9525. In *HUSO* (pp. 29–32).
- Bal Besikci, E., Tavacioglu, L., & Arslan, O. (2016). The subjective measurement of seafarers' fatigue levels and mental symptoms. *Maritime Policy & Management*, 43(3), 329–343.

33. Bor, R., Field, G., & Scragg, P. (2002). The mental health of pilots: An overview. *Counselling Psychology Quarterly*, 15(3), 239–256.
34. Butcher, J. N. (2002). Assessing pilots with ‘the wrong stuff’: A call for research on emotional health factors in commercial aviators. *International Journal of Selection and Assessment*, 10(1–2), 168–184.