Intelligent wristbands for the automatic detection of emotional states for the elderly

Jaime A. Rincon[1], Angelo Costa[2], Paulo Novais[2], Vicente Julian[1], Carlos Carrascosa[1]

[1] Universitat Politècnica de València. D. Sistemas Informáticos y Computación
{jrincon, vinglada, carrasco@dsic.upv.es}
[2] ALGORITMI Center, University of Minho, Portugal
{acosta, pjon@di.uminho.pt}

Abstract. Over the last few years, research on computational intelligence is being conducted to detect emotional states of people. This paper proposes the use of intelligent wristbands for the automatic detection of emotional states to develop an application which allows to monitor older people in order to improve their quality of life. The paper describes the hardware design and the cognitive module that allows the recognition of the emotional states. The proposed wristband also integrates a camera that improves the emotion detection.

Keywords: Wearable Devices, Emotional Models, Ambient Assisted Living

1 Introduction

With the start of mobile technology every day our devices are connected to the world-wide network (WWW). This connection allows us to share our photos, send messages, use the social network, etc... However, in recent years we can see how different devices can be connected to the network with the aim to control and to monitorize our environment. This new technology called Internet of Things (IoT), allows the creation of applications in different environments. This has been made possible by the emergence of several smaller and more powerful devices. These devices are increasing applications in domains such as smart homes, smart cities, healthcare and robotics.

The elderly people care is one of the most interesting applications, since globally, the elderly population is increasing, according to demographic projections [1]. According to the OMS, the amount of people aged over 60 is expected to double between 2000 and 2050 [2]. As less developed countries start to evolve, this trend is onset immediately [2]. A common societal issue that emerges from a rapid elderly population growth is the exponential demand or care (medical and otherwise).

These care demands can be met by using the new technology in IoT and embedded systems, these devices can be used as wearable devices. These devices can obtain information related to the person movement into environment and acquire some bio-signals such as electrocardiogram (ECG), heartbeat and skin resistance. These data can be used to know not only the good state of health, if not, can be used to know the emotional state. Some studies [3], [4] associate the emotional change with health problems.

In this work we propose to detect the emotional state of older people in an Ambient Intelligence (AmI) application with the help of wearables. The application would be used by the caregivers to try to improve the activities to be done with the elderly and by the way, improving their quality of life. Thus, the main goal of the proposed system is to use the knowledge of the emotional state of older people to maintain a state as close as possible to happiness or to detect unwanted situations from an emotional point of view. For this purpose, an intelligent wristband has been designed in such a way that it integrates the necessary components to take biometric measurements to infer the person's emotional state.

The rest of the paper is structured as follows. Section 2 presents the related work. Section 3 describes the proposed system, which has been divided into the hardware description and the cognitive service. Section 4 briefly presents a case study. Finally, some conclusions are shown in section 5.

2 Related Work

Other efforts that are in the same or related domains (emotional detection through body signals) of our project are presented in this subsection. they represent the most advanced methods and technologies, being reference points of best, and sometimes, worst practices. These research projects provide cues to what will be the next developments in the domain.

In terms of new approaches to the detection of emotions through body sensoring, [5] present a novel approach to detecting emotions through ECG. Their initial results show over 90% accuracy in detecting emotions in a controlled environment. The feature extraction (in this case the emotion) goes through an unorthodox process of data processing, they have altered the classical signal processing and data classification structure. The authors have first implemented a quantization method that compares the incoming signal to a dataset, meta-classifying them. The authors justify this approach by assuming that early data processing stage can constrain the data. Then, they compress the ECG metadata using an ECG dataset as reference. Finally, they classify the ECG using probability methods. The main issue of this project is that the number of training individuals is low (only 24) and the emotional identification lacks nuance to detect muted emotions.

Following the classical methods of acquisition,signal processing and classification we have the [6] that present an approach to the detection of emotional features through the use of ECG and GSR. They have used the Matching Pursuit algorithm and a Probabilistic Neural Network for the detection of the emotions. The authors have restricted the quantity of emotions to 4, being: scary, happy, sad, and peaceful. These are lifted from the Pleasure Arousal Dominance (PAD) model. The authors used music as activator on 11 students and affirm a high level of accuracy in emotion detection in most of the cases (over 90%). They have discovered that the GSR has little impact for emotions detections. They affirm that the detection of emotions was clear in terms of the arousal, and far less significant in the other fields. [7] shows the ATREC project (a military development) that accesses the stress levels of military personnel through the usage of body sensors. They have discovered that it is possible to determine a high level of valence markers and alert levels from ECG and GSR palaced in the throat, which are directly related to stress levels. Furthermore, their tests have revealed that the speech, GSR (on the hands/arms) or skin temperature provide little additional information about the current emotions. [8,9] presents a study that presents a great accuracy in emotions detections, from the combination of ECG with forehead biosignals. The authors found that actions like frowning or facial movements convey a high level of information about the emotion the subject is feeling. Contradicting the low significance of the usage of the GSR is the [10] that have found that each sensor is related to an emotional field. GSR is related to extremes of valence emotion while ECG is related to emotionally active states (subtle displays of emotion). The authors have determined that direct data fusioning is worst than give different tasks to the sensors. They have found that the activation (when differing from neutral state) should be performed by the GSR, while the classification should be done with the data of the ECG.

As it can be observed from this short sample, there are different approaches (even conflicting ones) to the detection of human emotions with minimal intrusion. Two things are clear from their research, ECG is crucial for the detection and classification of emotions, and that using various sensors may improve the classification accuracy or help detecting trigger events.

3 Description

In this section we describe the proposal of the bracelet-camera that incorporates a camera in a wristband to monitor the elderly. In current years, the use of wearable devices has been growing, devices such as *Samsung*[3] with the *Gear Fit, Gear S2, Gear S3,* or *Apple*[4] with the *Apple Watch* are only some examples. These devices can measure heart rate beat or hand movement using the IMU (Inertial Measurement Unit). Based on these devices and using the current technology in embedded systems, it is possible to create new smart bracelets. These new devices have been seen to create applications in IoT, since these new devices are smaller and more powerful.

The use of these devices has many fields of application, the most common being in sport, nevertheless, these devices can be used to monitor older people. However, these devices have a problem, as they all need a smartphone to work properly. This device is normally designed to communicate with this smartphone via Bluetooth communication. This means that the smartphone processes and

[3] http://www.samsung.com
[4] http://www.apple.com

analyses the signals and records user information such as number of slopes per day, sleep analysis, etc. In recent years, new devices with different communication protocols such as WIFi, Bluetooth and LoRa have appeared. All these features are advantageous for the creation of applications for the monitoring of elderly people. This monitoring can be through the acquisition of signals such as electrocardiography (ECG), photoplethysmography (*PPG*), respiratory rate, etc.

However, it is possible to introduce other types of sensors such as recessed cameras. These cameras can be used as a small device for videoconferencing, emotion recognition or as a tool to know if the person has fallen. However, most of these devices do not take into account the introduction of a camera with new technologies, so it might be possible to create a wristband that has a camera and uses it as another input. A camera can be used in different areas, as a way to recognize emotions, video conferencing and if the bracelet detects that the person has fallen, you can send a message to the caregiver and see the image on your terminal.

To make this application possible, it is necessary to use different technologies not only in hardware, but also in a service that analyses the information sent by the different devices. This service needs to have the capability of pattern recognition, image analysis, emotion analysis, stress detection, etc.

3.1 Hardware Description

The wristband has been manufactured using different sensors (Figure 1), which can acquire bio-signals, and a camera that can be used to take pictures of the user. Two sensors were used to acquire the bio-signs, one to detect the heartbeat and the other to measure the skin's resistance. The sensor used to obtain the heartbeat is a PPG sensor (Figure 1 (c)). A PPG is often obtained using a pulse oximeter that illuminates the skin and measures changes in light absorption. A conventional pulse oximeter monitors blood perfusion to the dermis and subcutaneous tissue of the skin. With each heart cycle, the heart pumps blood to the periphery. Although this pressure pulse is somewhat cushioned when it reaches the skin, it is enough to distend the arteries and arterioles in the subcutaneous tissue. If the pulse oximeter is connected without compressing the skin, a pressure pulse can also be seen from the venous plexus, such as a small secondary peak.

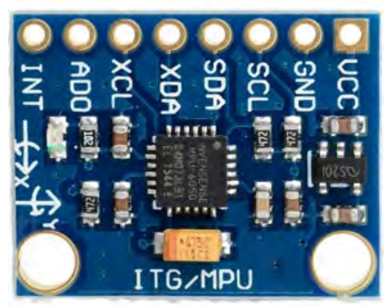
The skin is the largest organ in the human body and has different properties, one of which is the ability to vary resistance. This variation has different names such as electrodermal activity (EDA), galvanic skin response (GSR), electrodermal response (EDR) and psychogalvanic reflex (PGR). The galvanic response variation can be used either to detect stress [11], or even emotions [12]. The sensor to detect such variation can be seen in the Figure 1 (b).

At the same time a mini-camera (Figure 1(a)) is introduced that allows us to use the photos it acquires to analyze the emotional states. This camera (OV2640) has a low voltage CMOS image sensor that provides full functionality of a single-chip VGA camera (an image processor in a small package). The image is processed by a ESP-32 chip. This chip has wifi and bluetooth communication protocols, that make it especially important for IoT applications.

(a) Camera developer board

(b) GSR Sensor

(c) Photo Plethysmography Sensor



(d) Inertial Measurement Unit

Fig. 1: Sensors used in the wristband.

However, this information needs to be processed. As the chip has a low computing power, it is necessary to have a central unit that can be used to manage this information. It is for this reason that we see the need to introduce a cognitive service to analyze the information sent by the sensors and the camera. This cognitive service is explained below.

3.2 Cognitive Service

The Cognitive Service is a new tool that uses a machine learning technique to create smarter and more engaging applications. This cognitive service introduce

API to detect emotion, speech recognition, conversion of text to speech and more. Some of the most important services that can be used right now are *Microsoft Cognitive Service (formerly Project Oxford)*[5], *IBM Watson*[6], *Google*[7] and *AMAZON AWS*[8].

The cognitive service was divided into two parts, one part specialized in the recognition of emotions through image processing (sending data through the camera) and the other part in which bio-signals are used to recognize emotions (sending data through sensors). These elements are explained below.

Emotion detection through image processing The emotion classification uses *Convolutional Neural Networks* [13]. They are very similar to ordinary Neural Networks. They are made up of neurons that have learnable weights and biases. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other.

Our cognitive service can classify 7 emotions: fear, angry, upset, happy, neutral, sad and surprised. To train the model, the KDEF dataset [14] , which consists of a total of 980 images, was used. In this dataset the 70 actors, 35 women and 35 men, represent the 7 different emotions. Each image has a resolution of 562 pixels wide and 762 pixels high. Before performing the training, each image is pre-processed to detect and extract the face. Once extracted, a colour transformation is performed in grayscale. Finally, it is necessary to resize the image to 128x128, and to train the model using Tensor Flow [9].

The structure of the network was modified changing different parameters such as the number of convolution filters in the different layers. This codification allowed us to obtain the best results. In the end, the best network has the following structure as showed in the Table 1.

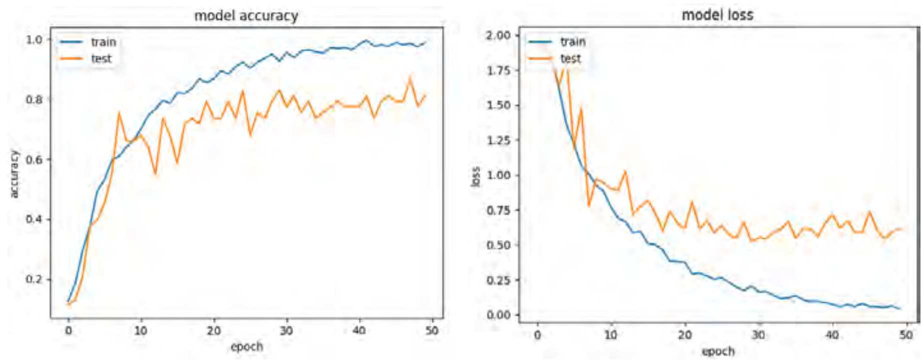
	Layer 1	**Layer 2**	**Layer 3**
Num Input Channels	3	3	3
Convolutional Layer	32	32	64
Conv Filter Size	3	3	3

Table 1: CNN Architecture

Figure 2(a) shows the Model Accuracy in the Train Data and Validation Sets, seeing how the test graph follows the training graph. This behaviour allows the pressure level between the training data set and the test data to be determined. Figure 2(b) shows the Model Loss in the Training and Validation Data Sets.

[5] https://azure.microsoft.com/en-us/services/cognitive-services/
[6] https://www.ibm.com/watson/
[7] https://cloud.google.com/
[8] https://aws.amazon.com
[9] https://www.tensorflow.org/

This graph shows the validation process between the training model and our test data.



(a) Model Accuracy on Train and Validation Datasets (b) Model Loss on Training and Validation Datasets

The web service can be used in two modes, a static mode in which the user loads the image and returns the emotion, and a dynamic mode. This second mode is used by any device that has access to web services, allowing them to send the streaming image to the web services, where it is analysed and the web services return the emotion that was detected.

Emotion detection through Bio-Signal processing To detect emotion through bio-signals we used the DEAP [12] data set, which is structured as a series of participant classifications, physiological recordings and facial video of an experiment in which 32 volunteers viewed a subset of 40 of the previous music videos. The electroencephalogram and physiological signals were recorded and each participant also rated the videos as mentioned above. In our case, to train our deep learning model, we use the GSR and PPG bio-signals. The other signals were dismissed as it is not comfortable to wear an EEG helmet or use an ECG holter to acquire these signals. The signals acquired by the wristband were filtered, we used a butterworth filter to eliminate the noise introduced by the electric field. This process is done in the web service, as the wristband does not have the necessary computational power to perform this filtering.

Figure 2 shows two signals, the input wave is the original signal that, as can be observed, has noise. This noise must be reduced, because if it does not, it may lead to a bad classification. The other signal is a filtering signal, so you can see that the noise is reduced. Once you get this new filtering signal, the next step is to train the network.

In the same way that the parameters of the network used to analyse the images were modified, the network that analyses the signal was modified to try to obtain the best results (Table2).

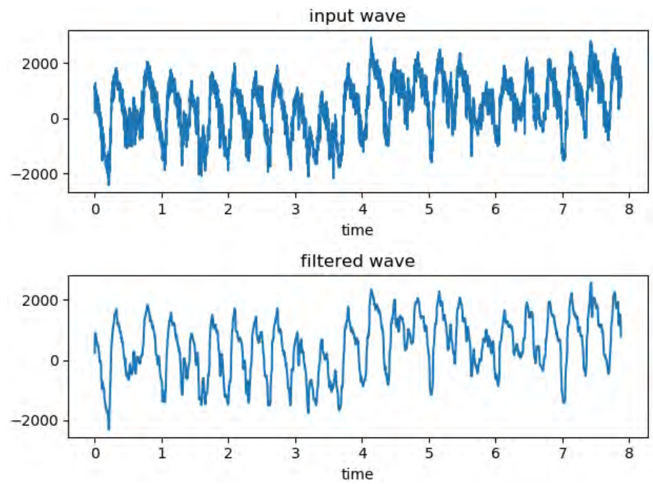


Fig. 2: Unfiltered Dataset Signal

	Layer 1	Layer 2	Layer 3
Num Input Channels	3	3	3
Convolutional Layer	32	64	64
Conv Filter Size	3	3	3

Table 2: CNN Architecture

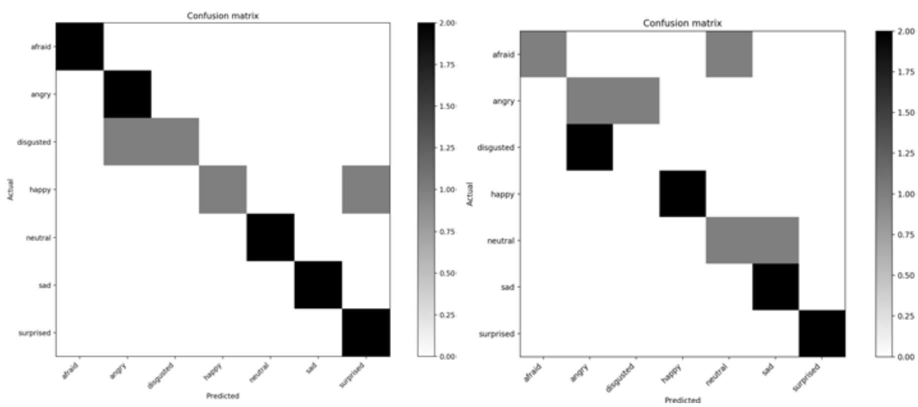
4 Case Study

Elderly people are susceptible of great emotional changes even when subjected to minor interactions, and these changes are aggravated when they suffer from cognitive disorders. The goal of this project is to counterweight environmental changes with positive *stimuli*, thus trying to achieve a neutral/positive emotional state. A case study was designed to subject the elderly people to a set of *stimuli*, in a controlled manner, so that our methods could be validated accordingly.

The case study was done in the *Centro social Irmandade de São Torcato* to test the bracelet and to evaluate if the mixture of both perception methods outputs relevant classification results of emotions. A prototype bracelet (as described in section 3.1) was used by five persons (from now referred to as users) between the ages of 65 and 70, and the test consisted in recognizing the emotions while watching a music video. The users were chosen specifically to be a representation of the elderly community. They are 3 woman and 2 men, with none to mild cognitive disorders, representing the scale of cognitive disorders that allow the bearers to use fairly complex technological devices.

The test operation consists in record the bio-signals for one minute (via the bracelet); next, send a warning vibration (through a motor in the bracelet) to grab the attention of the user, and concurrently, tracking the user's face and taking pictures. This information is synchronized with the server for immediate processing.

The server analyses the information received, and outputs the classification error and the mixture of both methods result, as these values are the main goal of this case study. Moreover, the server also outputs the classification of the user's emotional state. In a production stage the information about the emotional state can be directed to a caregiver, allowing him/her to change the activities/calendar to suit the user's or to improve their mood.



(a) Image processing. (b) Signal processing.

Fig. 3: Confusion matrix obtained from the classification of emotions

The results of this case study are showed in Figure 3(a) and Figure 3(b), which show the confusion matrix resulting from the classification of emotions through image processing and the confusion matrix resulting from the classification of emotions through signal processing, accordingly. As it can be observed, the integration of the two methods is complementary, achieving 90% accuracy through image processing and 79% through the bio-signal method.

This means that our initial hypothesis is correct and it produces relatively stable results. Although there are clear issues that may undermine the overall results (like the values of bio-signals), we believe that they can be improved by using more accurate or different collection methods (like a body vest with sensors). With these positive results, a production-ready bracelet can be produced to be used by a larger set of users, or in a group environment to test social interactions.

5 Conclusions and future work

In this work we have presented how to integrate non-invasive bio-signals and cameras for the detection of human emotional states into an intelligent wristband. In this way, the proposed wristband integrates two ways to detect human emotions to improve the emotional detection.

The main application domain of the designed wristband is the elderly care in nursing homes, but it can be used in other types of domains.

The wristband is able to recognize emotions through the processing of images and biological signals. To do this, the wristband acquires the bio-signals and images through the corresponding sensors and sends them to a cognitive service for the analysis. The proposed approach is being validated by workers and patients of a daycare centre *Centro Social Irmandade de São Torcato*. The validation is being performed through simple interactions with the patients under the supervision of caregivers.

Preliminary results show that the wristband is well accepted by the elderly people resident of the center. Future work will initially focus on the development of new tests with a larger number of users. These new tests will allow the wristbands to be used to improve the activities and tasks performed at the centre in order to avoid situations where older people have undesirable emotional states.

6 Acknowledgements

This work is partially supported by the MINECO/FEDER TIN2015-65515-C4-1-R and the FPI grant AP2013-01276 awarded to Jaime-Andres Rincon. This work is supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT - Fundacao para a Ciencia e Tecnologia within the projects UID/CEC/00319/2013 and Post-Doc scholarship SFRH/BPD/102696/2014 (A. Costa).

References

1. United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2015. (ST/ESA/SER.A/390), 2015.
2. United Nations, Department of Economic and Social Affairs, Population Division. World population prospects: The 2017 revision, key findings and advance tables. (ESA/P/WP/248), 2017.
3. Timothy W Smith, Kelly Glazer, John M Ruiz, and Linda C Gallo. Hostility, anger, aggressiveness, and coronary heart disease: An interpersonal perspective on personality, emotion, and health. *Journal of personality*, 72(6):1217–1270, 2004.
4. Laura Smart Richman, Laura Kubzansky, Joanna Maselko, Ichiro Kawachi, Peter Choo, and Mark Bauer. Positive emotion and health: going beyond the negative. *Health psychology*, 24(4):422, 2005.
5. Susana Brás, Jacqueline H. T. Ferreira, Sandra C. Soares, and Armando J. Pinho. Biometric and emotion identification: An ECG compression based method. *Frontiers in Psychology*, 9, apr 2018.
6. Atefeh Goshvarpour, Ataollah Abbasi, and Ateke Goshvarpour. An accurate emotion recognition system using ECG and GSR signals and matching pursuit method. *Biomedical Journal*, 40(6):355–368, dec 2017.
7. Fernando Seoane, Inmaculada Mohino-Herranz, Javier Ferreira, Lorena Alvarez, Ruben Buendia, David Ayllón, Cosme Llerena, and Roberto Gil-Pita. Wearable

biomedical measurement systems for assessment of mental stress of combatants in real time. *Sensors*, 14(4):7120–7141, apr 2014.

8. Mohsen Naji, Mohammad Firoozabadi, and Parviz Azadfallah. A new information fusion approach for recognition of music-induced emotions. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, jun 2014.
9. Mohsen Naji, Mohammd Firoozabadi, and Parviz Azadfallah. Emotion classification during music listening from forehead biosignals. *Signal, Image and Video Processing*, 9(6):1365–1375, dec 2015.
10. Priyanka Das, Anwesha Khasnobish, and D.N. Tibarewala. Emotion recognition employing ECG and GSR signals as markers of ANS. In *2016 Conference on Advances in Signal Processing (CASP)*. IEEE, jun 2016.
11. María Viqueira Villarejo, Begoña García Zapirain, and Amaia Méndez Zorrilla. A stress sensor based on galvanic skin response (GSR) controlled by ZigBee. *Sensors (Switzerland)*, 12(5):6075–6101, 2012.
12. Sander Koelstra, Christian Mühl, Mohammad Soleymani, Jong Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. DEAP: A database for emotion analysis; Using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, 2012.
13. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
14. Manuel G. Calvo and Daniel Lundqvist. Facial expressions of emotion (kdef): Identification under different display-duration conditions. *Behavior Research Methods*, 40(1):109–115, Feb 2008.