Computer Communications 89–90 (2016) 178–190

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition

Leandro Y. Mano[a], Bruno S. Faiçal[a], Luis H.V. Nakamura[a,b], Pedro H. Gomes[e,*], Giampaolo L. Libralon[c], Rodolfo I. Meneguete[b], Geraldo P.R. Filho[a], Gabriel T. Giancristofaro[a], Gustavo Pessin[d], Bhaskar Krishnamachari[e], Jó Ueyama[a,e]

[a] University of São Paulo (USP), São Carlos, São Paulo, Brazil
[b] Federal Institute of São Paulo (IFSP), Catanduva, São Paulo, Brazil
[c] Federal Institute of São Paulo (IFSP), São Carlos, São Paulo, Brazil
[d] Vale Institute of Technology (ITV), Belém, Pará, Brazil
[e] University of Southern California (USC), Los Angeles, CA, USA

ARTICLE INFO

Article history:
Available online 15 March 2016

Keywords:
Smart Home
E-health
Internet of Things
Face recognition
Emotion detection

ABSTRACT

Currently, there is an increasing number of patients that are treated in-home, mainly in countries such as Japan, USA and Europe. As well as this, the number of elderly people has increased significantly in the last 15 years and these people are often treated in-home and at times enter into a critical situation that may require help (e.g. when facing an accident, or becoming depressed). Advances in ubiquitous computing and the Internet of Things (IoT) have provided efficient and cheap equipments that include wireless communication and cameras, such as smartphones or embedded devices like Raspberry Pi. Embedded computing enables the deployment of Health Smart Homes (HSH) that can enhance in-home medical treatment. The use of camera and image processing on IoT is still an application that has not been fully explored in the literature, especially in the context of HSH. Although use of images has been widely exploited to address issues such as safety and surveillance in the house, they have been little employed to assist patients and/or elderly people as part of the home-care systems. In our view, these images can help nurses or caregivers to assist patients in need of timely help, and the implementation of this application can be extremely easy and cheap when aided by IoT technologies. This article discusses the use of patient images and emotional detection to assist patients and elderly people within an in-home healthcare context. We also discuss the existing literature and show that most of the studies in this area do not make use of images for the purpose of monitoring patients. In addition, there are few studies that take into account the patient's emotional state, which is crucial for them to be able to recover from a disease. Finally, we outline our prototype which runs on multiple computing platforms and show results that demonstrate the feasibility of our approach.

1. Introduction

The number of elderly people is rapidly increasingly in several countries, including the USA and Brazil [13,14]. These people tend to live alone and are prone to have more diseases than young people. In addition, they are often treated at home after surgical treatment or after being discharged from hospital. While having a "24/7 nurse" is ideal, most people cannot afford this type of service. This often means that technology is used to monitor patients and elderly people while they are recovering at home. These systems can monitor them and issue alerts to nurses and/or relatives whenever necessary.

In view of this, the use of Health Smart Homes (HSH) is extremely important [29,30]. The concept of HSH has emerged from a combination of telemedicine, domotics products and information systems; it can be defined as a Smart Home that is equipped with specialized devices for remote healthcare. These devices are mainly sensors and actuators that can take action whenever a critical situation is detected. This has been exploited in several works [28,37], where the authors propose technological solutions

* Corresponding author. Tel.: +1 3103101429.
E-mail addresses: leandroyukiomano@usp.br (L.Y. Mano), bsfaical@icmc.usp.br (B.S. Faiçal), nakamura@ifsp.edu.br (L.H.V. Nakamura), pdasilva@usc.edu, pedrohenriquegomes@gmail.com (P.H. Gomes), glibralon@ifsp.edu.br (G.L. Libralon), meneguette@ifsp.edu.br (R.I. Meneguete), geraldop@icmc.usp.br (G.P.R. Filho), gabrielg@usp.br (G.T. Giancristofaro), gustavo.pessin@itv.org (G. Pessin), bkrishna@usc.edu (B. Krishnamachari), joueyama@icmc.usp.br (J. Ueyama).

http://dx.doi.org/10.1016/j.comcom.2016.03.010
0140-3664/

to help caregivers monitor people in need. The aim is to give elderly people who are suffering from a disease, some independence so that they can live their lives in a more self-reliant manner.

During our research, it was found that most research projects are concentrated on improving the daily lives of patients by providing them with specific gadgets, such as alert alarms. This can be seen in [17,31,38]. In this sense, Intel Corporation has developed an ultrasound-based system [17] to monitor elderly people at home as part of their Internet of Things (IoT) system suite. This proprietary system detects patient's falls and irregular movements, and requires them to wear a wristband; however, it cannot recognize pain or any type of emotional response which is a crucial factor during the treatment at home. Few works make use of cameras, image processing, emotional recognition and novelty detection through captured images and facial expressions. Hence, we seek to investigate the use of cameras as a means of improving in-home healthcare and providing patients with a higher degree of comfort during their care outside of the hospital.

This paper offers an in-home healthcare solution by making use of images and facial expressions to monitor patients and elderly people that have special needs. Capturing facial expressions over a certain period of time can give an idea of to what extent the patient is feeling pain and can enable a nurse/relative to decide whether help is required or not. Similarly the patient's face can register his/her emotional state, which is also a crucial factor during the treatment. Emotion plays an important role in recovery from disease and this is examined by authors of [11] and [27]. Likewise, from the images of the captured face, our system is capable of determining who entered a given room and for how long. If the patient does not leave the bedroom at the usual time, the system may discover that special help is needed. A particular kind of configuration can also be shaped to each person in the house according to the person's disease and particular needs.

It is important to emphasize that facial expressions are widely researched and employed to classify people emotional states. Such facial expressions have already been studied to determine the emotional health of an individual which, in turn, can be used as an important symptom for the diagnosis of various diseases [15,18,24,33] such as: schizophrenia, depression, autism, bipolar disorder. We highlight that the above mentioned diseases can be caused by the excessive time on negative emotions, lack of emotional expressions or the instability of expressed emotions [15,24,33]. In addition to the use of emotional health to diagnose diseases, emotional states can influence social interaction and also help us in making decisions/measures (e.g. suggest to have a break or a cup of coffee, if he/she is stressed).

While having a camera in the home raises problems about privacy, we believe these can be overcome since they be turned off whenever a resident wishes. It should be noted that the use of cameras and image processing can be less intrusive as there is no need for people to be fitted with any special item (such as wireless-based emergency buttons) to ask for special help. The purpose of our approach is to detect any new feature in the house through the use of cameras and facial expressions. It should also be stressed that there is no need to have someone watching the whole video stream all day long as the detection of new features is carried out automatically without any human intervention.

We developed our prototype and carried out experiments with various people to recognize each of them and classify their emotions on the basis of their facial expressions. This involved making use of our IoT platform for in-home healthcare that we constructed by combining Wireless Sensor Networks (WSN) and smartphones. The smartphones are capable of issuing warnings to designated people whenever required. Our sensor nodes are scattered around the house and are connected to cameras that capture the residents' faces.

The remainder of this paper is structured as follows: Section 2 highlights the current work in the field and discusses our contribution to the state-of-the-art. Following this, Section 3 outlines our existing in-home healthcare platform which was built through a combination of smartphones and WSN. Section 4 outlines how our platform will be used in order to implement the in-home healthcare system based on images. Section 5 evaluates aspects of our prototype and the experiments that were conducted to validate our approach. Finally, Section 6 wraps up this paper with some final remarks and points out potential future directions for this research study.

2. Current studies on IoT-based Smart in-home healthcare

This section examines the advances made in HSH based on IoT technologies with a view to describing and explaining the main challenges of this research study. The authors of [28] review the emerging concept of Health "Smart" Home and picture potential applications of these systems, such as social alarm systems and remote monitoring. The paper also discusses prominent technologies that are being deployed and reports pilot projects of e-care systems already in use by hospitals and residences. The advances of in-home healthcare [29] are also due to the application of Ambient Intelligence (AmI) [10] and Artificial Intelligence (AI) [8,9]. The combination of modern sensing equipments, with advanced data processing techniques and wireless networks culminates in the creation of digital environments that improves the daily life of residents.

Most of the studies [30] in the literature focus on the use of body sensors and specific devices when providing assistance to elderly people with special needs. The main concern in this case is the privacy of the other residents. A few studies [2] use cameras for improving the analysis of the environment (e.g. detecting possible hazardous situations) while providing a system that is non-intrusive. The two different approaches (sensor-based and camera-based) clearly have advantages and disadvantages and should be used with specific objectives. This work argues that the use of a camera-based approach enables the emotional state of the patients to be analyzed and can provide a non-intrusive system that is not heavily dependent on attached sensors, unlike most of the current solutions. The proposed solution differs from all the other state-of-the-art systems and introduces new features for treating patients.

The authors of [30] propose a wireless architecture for Personal Area Network (WPAN) that takes advantage of the benefits offered by an intelligent environment by using information from different sensors. The goal of the system is to ensure safety and provide services to users through monitoring techniques. The system is based on image processing and can provide solutions that are integrated with other control devices. In this case, a sensor based on image processing has a combination of filters, frame rate comparisons and other algorithms that act intelligently in the environment. This enables the detection of movements and patterns of the activities of a resident, such as speed and direction. It should be noted, however, that no attempt is made to analyze the emotional state of the residents in order to assess the state of their health.

The authors of [2] set out a solution for assisting elderly and vulnerable people to reduce or mitigate accidents experienced during the night. The system is based on night vision and operates through a minimum of interactions with the patients. It is algorithmically based on the notion of causality and spatio-temporal reasoning and the system alerts and provides explanations to caregivers about past accidents or accidents that may occur. The results of the system (alerts and explanations) are based on a set of causal rules that refer to the places and times of activities, leading to a likely scenario within the house.

Table 1
Summary of the papers identified in the literature.

Related work references	Use of IoT sensors?	Exploit image processing?	Propose infra for IoT with Cloud?	Identify each individual patient?	Consider emotional aspects?
Romero et al., 2009 [30]	√	X	√	X	X
Augusto et al., 2007 [2]	X	√	X	X	X
Helal et al., 2005 [16]	√	X	√	X	X
Abowd et al., 2002 [1]	√	X	√	X	X
Our SAHHc system	√	√	√	√	√

Another important work is being developed by the University of Florida's Mobile and Pervasive Computing Laboratory, creating the Gator Tech Smart House (GTSH) [16], an experimental laboratory and a real environment aimed at validating smart home technologies and systems. The goal of GTSH is to create supportive environments, such as homes, that can help its residents and is able of mapping the physical world with remote monitoring and intervention services, conducting research and development activities designed to assist elderly and people with special needs in maximizing independence and maintain a high quality of life. Similarly, Georgia Institute of Technology developed the Aware Home Research Initiative (AHRI) [1]. Research conducted by this research group investigate how new technologies can impact people's lives in the home environment on three main areas: (i) health and well-being; (ii) media and digital entertainment; and (iii) sustainability. The Aware Home has enabled research in many areas, such as: provision of data to the house, innovative sensing and control infrastructure, and a variety of applications to assist residents. The health and well-being of residents is one area of research that has received significant efforts focusing on technologies that enable a better aging, and help caregivers of children with disabilities such as autism. The Aware Home is a residential laboratory installed in a two-story identical, each designed with: kitchen, dining room, living room, 3 bedrooms, 2 bathrooms, and a laundry room. Although experiments in both papers, are made by short duration (1–10 days due legal implications), these studies have achieved satisfactory results enabling real residents to use the developed prototypes [1,16,19,20].

Table 1 summarizes the related literature detailed below. We provide the following comparison items to ensure that an overall view is given: (**i**) use of IoT sensors while monitoring patients/elderly people; (**ii**) adoption of image processing techniques as a way of monitoring people; (**iii**) proposal of an IoT-based infrastructure with pluggable sensors plus integration with Cloud; (**iv**) proposal of an approach to identify individual patients as means of providing customized treatment, and; (**v**) employment of user's emotions for identifying patient's well-being.

To the best of our knowledge, none of the proposals in the literature take into account the patient's emotional side and hence they fail to exploit a considerable amount of "rich" information available in facial expressions and body language. This includes detecting pains and other emotional related issues.

3. Our Smart in-home healthcare system

We propose a system called **S**mart **A**rchitecture for In-**H**ome **H**ealth**c**are: SAHHc. SAHHc is based on the use of IoT for smart and individualized monitoring of elderly patients in their homes. The main features of SAHHc are the scalability of its components and the ability to infer the emotions or feelings that patients experience (such as pain or the onset of depression), allied with individualized monitoring.

SAHHc is composed of two elements: (i) Sensors; and (ii) the Decision Maker. The sensor elements are devices distributed in the environment; they can be found in large numbers and their

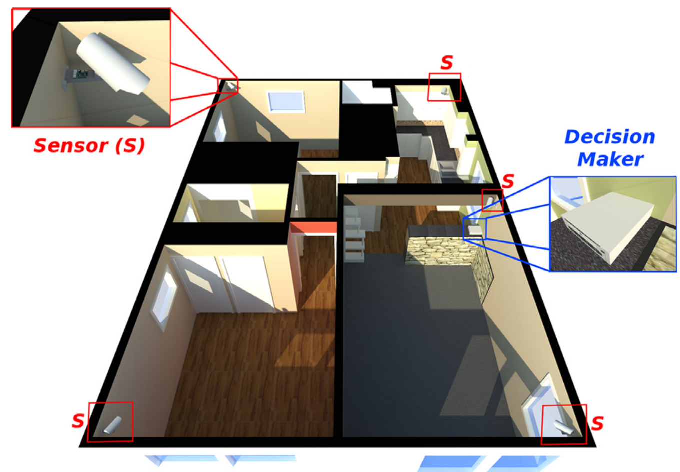


Fig. 1. Operating scenario of SAHHc.

purpose is to collect information (for instance, capture images) on the patient's health and send this to the Decision Maker element. The Decision Maker element processes the received information and makes decisions aimed at preserving the patients' health or giving them appropriate treatment. Initially, the Decision Maker element is defined as a single one (one machine) and also works as the network manager, but its functions can be split between several decision makers for scalability purpose. Furthermore, the Decision Maker serves as an interface between the patient and the helper. Fig. 1 shows the elements in a scenario where the proposed system is used.

Given the features of the environment where this type of application has to be implemented, all the devices must run on the power supply of the household and use battery as a backup source. In addition, it is assumed that the residence has an internal communication network that can be used by the system. Although wireless networks ensure that the SAHHc can be implemented with ease and mobility, a wired network can also be used.

The structure of SAHHc consists of simple and dedicated devices that act as sensors and collect the information about the patient's health and the environment (if necessary), and a decision maker with more powerful computing resources, e.g. a workstation. The sensor elements are able to detect the presence of people and also capture their images. When a person enters a room the device registers this activity and collects the person's image, which is processed to capture the person's facial characteristics and, finally, identify that person. If the identified person is not a patient, the behavior of this individual is ignored. On the other hand, if the system recognizes a patient, the activities are monitored. Moreover, in this case, the current and future captured images from the patient's face are used for detecting his/her emotional state [3,21].

Our underlying model for developing SAHHc was borrowed from [40] in which new features can be deployed at run-time according to the target needs and environment. We designed SAHHc

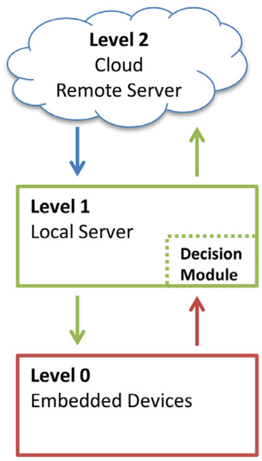


Fig. 2. Configurable elements (levels) of the SAHHc architecture.

with configurable elements (or levels) in which the so-called levels are deployable whenever required.

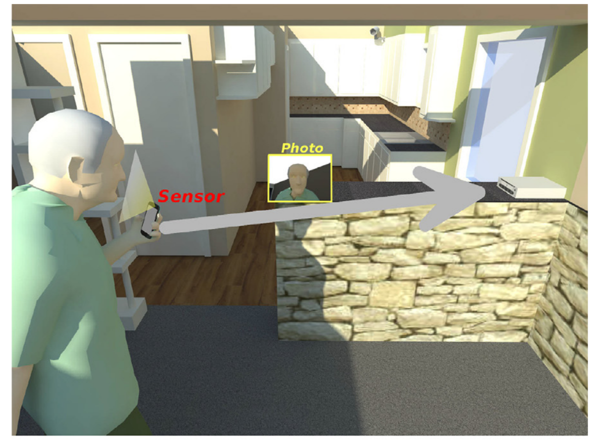
Fig. 2 presents the configurable elements (levels) of the SAHHc architecture. In a nutshell, SAHHc is divided into three levels: Level 0 (with the embedded devices) performs the face recognition, pre-processing and identification the patients being monitored as well as their facial expressions. Level 1 (local server) will carry out the monitoring of the house and monitoring of the health of patients. Hence, it needs to have more processing power and the storage availability. It also performs the entire control of the whole system. At Level 2 (remote server), the image processing in the Cloud is carried out when the local server is overloaded and unbalanced. Thus, it provides a load balancing scheme and ensures that SAHHc architecture is scalable. A detailed explanation of the IoT architecture is presented in Section 4.3.

The architectural features are not tightly-coupled and this allows us to configure it according to specific needs. This kind of flexibility ensures that new elements can be easily plugged into the architecture whenever required. For example, if the patient begins to use new devices (e.g. tablets, smartphones and smartwatches) that are able to collect information about his/her health, these new devices can be instantiated in the SAHHc system. If the new device is able to collect a new kind of patient data, it is also necessary for the Decision Maker element to receive an instance for processing this type of data and possibly compose this data with other information that is already being used.

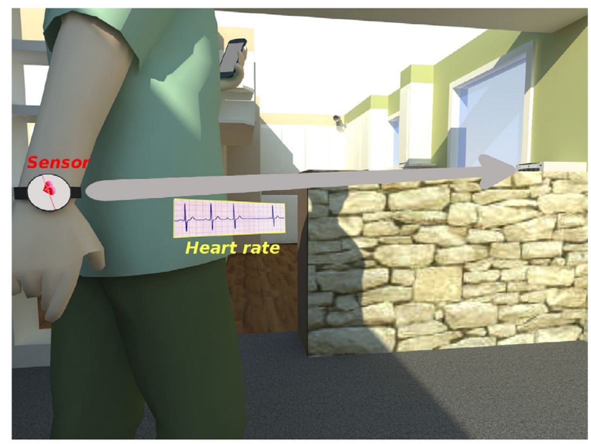
As shown in Fig. 3(a), it is usual for devices such as smartphones, to have front cameras so that they can be used to capture the facial image of the user. Thus, these devices can be prepared (for example, by installing a SAHHc application) to collect and transmit images whenever the Decision Maker element is within reach via the communication network. In addition, the devices have other sensors which may be useful for monitoring patients, for example, the accelerometer and gyroscope for detecting an accident or sudden illness.[1]

Additionally, smartwatches and clothes with wearable sensors can be included in SAHHc as a means of collecting physiological information (e.g. body temperature and heartbeat rate) and transmitting this kind of information to the Decision Maker (see Fig. 3(b)). Thus, with a processing module for this information, they can be made use of with the captured images by other devices, in the interests of the patient's health.

[1] A sudden illness is any clinical symptom that characterizes a rapid loss of consciousness, i.e. the main functions of the individual. The problem might occur suddenly even among apparently healthy people.



(a) Sensor nodes as smartphones



(b) Sensor nodes as smartwatches

Fig. 3. Integration of different sensor type in SAHHc.

The operation of the system and tasks carried out by the Decision Maker element are based on gathered information, and are designed to ensure that the measures taken for the patient's health are in compliance with the treatment protocol. If the data does not contradict the patients expected behavior or health, the Decision Maker element does not take any action. On the other hand, if the Decision Maker finds signs of unexpected behavior, it acts on behalf of the patient.

In the standard structure of SAHHc, the Decision Maker element acts by communicating with a patient's relative or another person responsible for the patient's health (e.g. a nurse or caregiver), making the information available. The Decision Maker element can also be expanded to include other kinds of actions, as well as different actions for new situations or issues detected and issuing warnings through smartphone notifications, SMS or phone calls.

Although the design of SAHHc is based on components and totally scalable, a simple standard setting is used as a basis for the rest of this paper. This standard setting consists of sensor nodes with motion detection and cameras and a single Decision Maker (Fig. 4).

Initially, the sensor elements detect the presence (based on a motion sensor) of some person in a room in the house and collect images (by means of image sensors) containing the person's face. After this, the collected data is sent to the Decision Maker to identify the user and decide whether or not this individual should be monitored. If the individual is a patient, the Decision Maker uses current data to classify the patient's emotion and implement the decision rule. This monitoring cycle is repeated by sensor elements

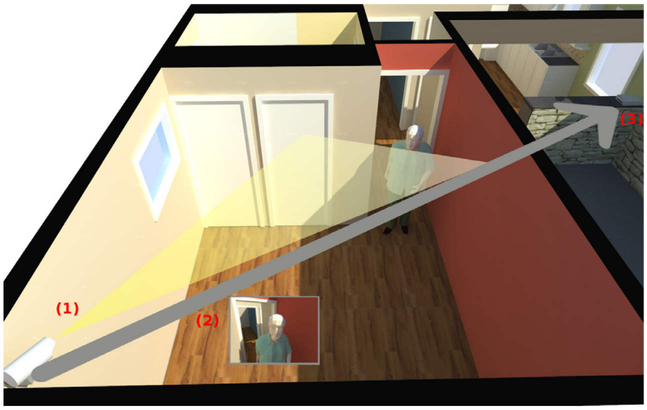


Fig. 4. Standard SAHHc setting.

each time interval configured in SAHHc. It should be underlined that this methodology which is employed for the monitoring cycle, where the system checks whether each collected image is in fact from the patient, allows individualized patient monitoring and prevents someone else from entering the room and being mistakenly monitored after an initial identification.

The best way of positioning the cameras and motion sensors is a challenge that can be exploited to optimize the accuracy of our system. Authors of [42] propose a hybrid architecture to solve this problem, which employs some cameras with attached presence detectors and a few stand-alone cameras. The use of smart cameras with pan/tilt/zoom features and robotic cameras, like Rovio [41], can assist the image capturing process, especially in the case of elderly and depressed patients who tend to look down.

Last but not the least, in our system, the fault tolerance mechanisms can be given through long-lasting batteries that can provide energy during power outages. Note that the set of batteries can be recharged as soon as the energy is back and restored. Our platform can also be equipped with additional batteries to deal with the problem of battery depletion. Finally, our wireless sensors system is not intended to be a pure Low power and Lossy Network (LLN), and hence it is connected to the power grid and it does not often rely on batteries.

4. Visual-based resident tracking for our smart in-home healthcare

Computational vision has been used in many systems since it allows the acquisition of huge amounts of multidimensional data related to the monitored environment. For example, it can be used to learn the activities of daily living (ADL) of elderly people to analyze health problems or cognitive disruption [4].

In the current proposal, computational vision provides a pervasive layer with the resident, and allows non-intrusive monitoring. Additionally, the architecture can use other devices for information (such as smartphones) to improve the accuracy of the currently available data or provide better alternative forms of communication for the resident.

Another important issue to consider is the emotional state of the resident. In the case of elderly people or drug addicts undergoing treatment, the careful monitoring of emotional states on a regular basis can prevent depression. For this reason, the FaceTracker computational vision system [34] is employed in two distinct ways: (i) to identify the users that have health problems and need monitoring; and (ii) to recognize the nature of their emotion. From these two perspectives, the FaceTracker vision system is a valuable support to our architecture. In addition, the fact that different techniques can be reused for distinct purposes highlights the flexibility of our proposal.

4.1. Ensuring resident's health by means of FaceTracker

Motor Expressions which are also known as Expressive Reactions are responsible for communicative behavior of an individual [23,35]. Thus, this article explores the use of the components related to motor expressions - or more precisely, the face of the user – and enables computing systems to determine this behavior and what emotion the user is feeling at a particular moment. It should be stressed that motor expressions are features that people from different ethnicities, regions and cultures share in common. In the proposed work, they are used as mechanisms to interpret the resident's emotional state/well-being and needs.

The task of automatic facial expression analysis can be divided into three key stages: face detection, facial feature extraction and classification. Face detection is a processing stage to automatically find the facial region for the input images or sequences. Detecting a face in a complex scene is a non-trivial problem: head motion, occlusion, changing illumination, presence of hair, glass or jewelry are examples of possible complications to the system [39].

After the face detection, it is time to extract and represent the facial changes produced by expressions. Several types of perceptual cues to the emotional states are displayed in the face: relative displacements of features (raised eyebrows), textural changes in the skin surface (furrowing of the brow), changes in skin hue (blushing) and others. Depending on how the face and its expression are modeled, the features have to be designed so as to condense this information or a part of it into a set of numbers to form the basis for the classification, and thus decide the quality of the final results [39].

The proposed extraction approach for face identification and expression analysis is based on geometric features. The geometric facial features model the shape and locations of facial components (including mouth, eyes, eyebrows and nose) through the use of feature points and geometric elements such as angles, distances or areas used to represent the geometry of the face. The classification of facial expressions is the last stage and, techniques based on Machine Learning (ML) are employed to tackle this problem.

The FaceTracker [34] system can handle partial occlusions of the face and uses an optimization strategy for local experts-based deformable model fitting. Deformable model fitting employs a linear approximation to determine how the shape of a non-rigid object deforms. It registers a parametrized model for an image, by fitting the landmarks into consistent locations of the analyzed object. With the aid of a reference model of the face, which consists of 66 feature points, the algorithm seeks to align the elements of the face of a new user with the feature points of the current reference model to fit a model for the current face. This represents another advantage of our architecture since the same features required for emotion recognition, are used for human identification, thus simplifying the analysis carried out by AI techniques.

For this reason alterations were made to the FaceTracker algorithm [21] so that it would only map a subset of the original 66 facial points. The purpose of this simplification is to reduce computational costs by eliminating possible redundancy. As a result, the elements took account of the proposed facial representation model and the facial features which are intrinsically related to movements, which, in the view of psychologists, are caused by emotional expressions [6,7].

Fig. 5 shows the graphical elements of the proposed representation. This is based on a reduced subset of 33 facial points: eight for mapping the mouth, six for each of the eyes, three for each eyebrow and the chin, two for the nostrils and two for demarcating the lateral extremities of the face near the eyes. The points are represented by red dots. As well as this, eight areas are mapped to model the shape of the eyes, the mouth and facial regions related to emotional muscular movements, and these can be

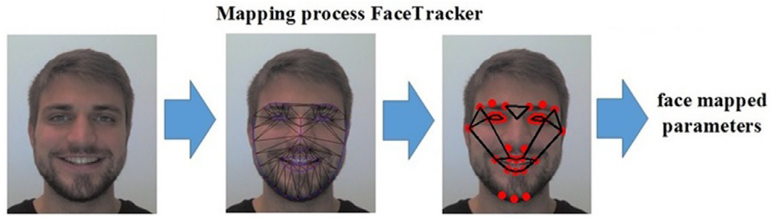


Fig. 5. The mapping process employed consists of using FaceTracker, then proceed to the creation of our proposed simplified facial representation. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

identified by the geometric regions bounded by the black line segments. In addition, the distances and angles are obtained for all possible combinations of points, to show the line connecting two distinct points with the horizontal axis. This creates a representation with a dimensionality of $D_1 = 2 \cdot 33 + 8 + 2 \cdot 528 = 1130$ attributes.

4.2. Artificial intelligence for person identification and emotion recognition

The proposed approach applies person identification and emotion recognition while monitoring the patient. This is carried out by employing distinct AI paradigms such as k-Nearest Neighbor (kNN), Decision Tree or Logistic Regression, Fuzzy Logic, Bayesian Networks and Support Vector Machine (SVM) [26].

The emotion recognition task is based on the psychological and cognitive science literature where it is stated that there are two basic views on the representation of emotions: categorical and continuous. In the categorical representation, different emotions are mapped into distinct categories. The most well-known example of this description is the set of six basic emotions, and the facial expressions related to them, which are innate and culturally uniform. All other emotional categories are then built up from combinations of these basic emotions. This approach is supported by the cross-cultural studies conducted by [5], which claim that humans perceive certain basic emotions conveyed by facial expressions in the same way, regardless of culture.

The main advantage of a representation of categories is that people use this categorical scheme to describe observed emotional displays in daily life. The labeling scheme based on these categories is very intuitive and thus matches people's experience. Within the discrete models, the most widespread example is proposed by [5], which lists the following basic emotions - happiness, sadness, fear, anger, disgust and surprise - as enabling humans to adapt to different everyday situations.

An alternative to the categorical description is the “continuous view”, in which emotions are described as points in multidimensional space, that rely on continuous scales or dimensional bases. An affective state is characterized in terms of a small number of latent dimensions rather than a small number of discrete emotional categories. From this standpoint, affective states are not discrete and independent of each other. Instead, they are systematically related to one another. These dimensions include evaluation, activation, control and power, as well as other factors.

The evaluation and activation dimensions are expected to reflect the main aspects of emotions. The evaluation dimension measures how a human feels, ranging from positive to negative emotions. The activation dimension measures whether humans are more or less likely to take action when in an emotional state, ranging from active to passive behavior. The most common dimensional approach is Russell's circumplex model, shown in Fig. 6 [32], which argues that all emotions lie in a continuous two-dimensional space where the dimensions are valence (how positive or negative the emotion is) and arousal (the energy or excitation level associated with the emotion).

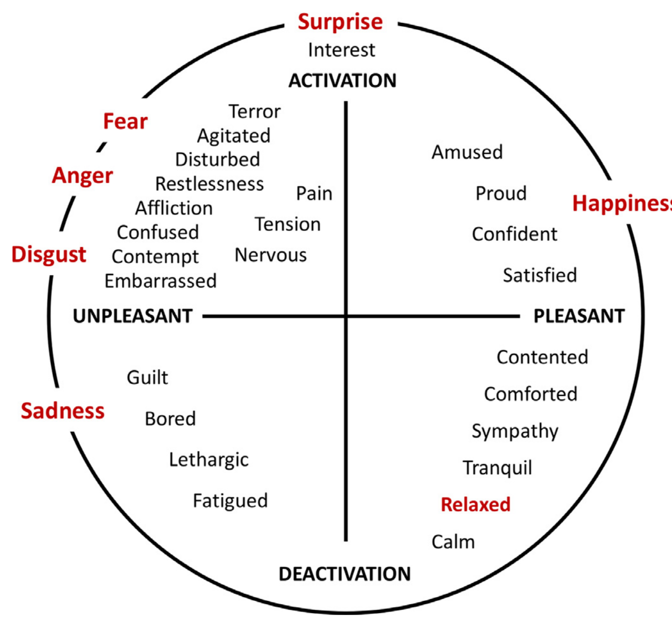


Fig. 6. Russell's circumplex model [32].

It should be noted that both representations of emotions have been used in various computational applications and achieved a good performance. As a result, our proposed approach employs a categorical representation of emotions.

According to studies conducted by [25], emotions can be identified by humans in distinct ways. These authors found three groups of emotions: (i) happiness and surprise, which are easy to identify; (ii) anger and sadness, which are more difficult; and (iii) fear and disgust, which are even harder to recognize.

On the basis of the findings of this study and our view that the disgust emotion is not of value to our practical approach, we decided not to include its recognition in the proposed system to make it easier to identify other emotions. This in turn, can help in the development of a real time application since the AI algorithms will more readily identify the emotions in question. For this reason, the emotions that were analyzed are displayed in Fig. 7. A neutral state is also taken into account, since it is used as a reference-point for the detection of the emotional states.

As well as the individual analysis of the previously mentioned AI algorithms for person identification and emotion recognition, a set of classification algorithms was also produced and evaluated. The reason for this is that the selection of a single classifier means that a significant amount of potentially useful information is rejected. For this reason, the concept of Ensemble of Classifiers (EC) has been found to be an ideal solution for the development of high performance systems in the area of pattern recognition [36]. EC is based on the premise that the combination of classifiers, that are not subject to coincident failures, allows the errors of a classifier to be corrected through the output obtained from all the other components. In this way, it can lead to an improved performance in

Fig. 7. Set of basic emotions [5] considered in the proposed approach. From left to right: a neutral state, joy and sadness (top); fear, anger and surprise (bottom).

terms of a better generalization and/or improvement in the final degree of accuracy.

The development of the proposed EC is based on the ML techniques already mentioned in this paper (kNN, Decision Tree, Fuzzy Logic, Bayesian Networks and SVM). All the ML algorithms, as well as the EC module were implemented to evaluate the progress and accuracy obtained by the face classification module. The methodology employed to measure the performance of the selected algorithms in both patient identification and emotion recognition is described in Section 5.

4.3. Discussion of our IoT-based environment

SAHHc's architecture is based on INCAS (INCidents-Aware System) [12], the main objective of which is to anticipate and detect incidents that may occur as well as to inform the user, caregiver or anybody else concerned about the incident/accident in the house. INCAS has three basic modules - the mobile module, web service module and sensor module (Fig. 8). The mobile module communicates with the web service module through Wi-Fi or cellular network (e.g. 3G or 4G). An Android application configures the user's preferences, such as the way he/she would like to be informed about an incident. These alert options can involve either playing an audio, starting a vibration or displaying an animated image. Thus, the user can choose a configuration that best suits his/her particular disability e.g. visual or hearing impairments. In addition, the application shows the user a list of rooms that can be monitored. Once a room has been selected for monitoring, a message is sent to the web service module which stores this information in the database. To avoid security issues, when this application starts, it requests a user authentication and also all the messages exchanged between the modules are transferred by means of the Secure Socket Layer (SSL).

The web service module is connected to a central computer at the user's house. This module needs an Internet connection to receive requests from the user's mobile device and a wireless connection via Wi-Fi to receive the data collected by the sensors. At this stage, there is an abstraction and data processing module, where these data are stored, analyzed and inferred, and we are thus in a position to know if the events captured by the sensors are really an incident.

The sensor module controls and manages the sensors. The hardware components consist of a Raspberry Pi model B, an USB wireless network interface, a camera and a simple presence sensor. Raspbian (official Raspberry Pi's operating system) was used for the software components. FaceTracker was used for monitoring video signals from the WebCam and a Python program to read the presence sensor's data from the Raspberry GPIO (General Purpose Input/Output) pins.

The INCAS architecture allows environmental monitoring to be carried out through motion sensors, gas sensors, humidity, and temperature. In this way, it can inform the handicapped people or carers about possible accidents that could occur at a particular location. However, the initial prototype is restricted to environmental monitoring and is unable to make inferences about the patient's behavior. New components have been added to the SAHHc architecture to abstract more information about a particular patient or resident of the new residence and enable recognition of the face from the facial expressions of a given patient, as can be seen in Fig. 9.

Monitoring and making inferences about the possible status of the patient is not the only feature of SAHHc however. It also enables the sensors to recognize the patients and analyze their feelings through a certain facial expression, for example, an expression of pain. This platform also allows the doctor or the carer to have more information about the status of a particular patient in real time through this information which is accessed via a smartphone.

Analyzing the SAHHc from a top-down view in Fig. 9, we can see that the Caregiver or a Doctor can use an mobile alert application via an Android smartphone to obtain more information about the status of a particular patient in real time through messages exchanged using web services. When the local server is started,

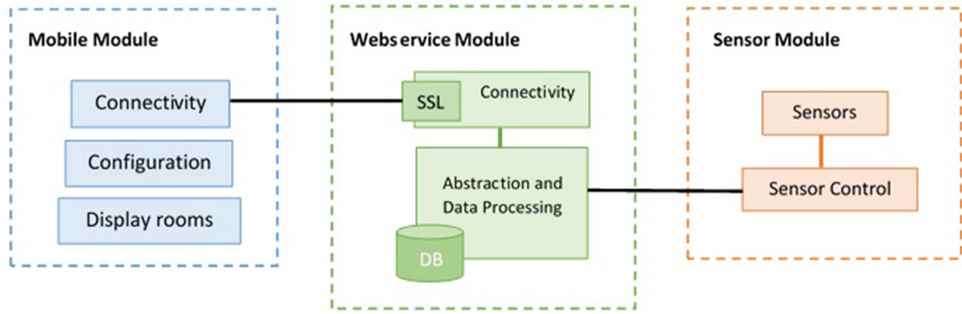


Fig. 8. INCAS extension abstract overview.

the web services are executed to collect the information from the database and make them available to the users. At the same time, the decision module starts the system Bootstrapper that will decide in which level the image processing will occur, it depends on the system workload that was registered by the Workload Analysis.

So SAHHc is divided into three levels (the higher is the level the higher is the computer power): Level 0 is always executed in the Raspberry where from time to time a camera takes a picture of the patient. This image capture is supported by a presence sensor, but the facial recognition can or cannot be executed according to the workload analysis. When the workload is low, the FaceTracker in the Level 0 performs the face recognition preprocessing and allows to recognize the patient being monitored as well as the facial expressions. Thus, the Level 0 is always responsible for capturing images, but the processing is performed only in cases where there are few patients and rooms to be monitored. Level 1 will carry out the monitoring of the house and monitoring the health of many patients at home, so it needs to have more processing power and the storage performs all the system control. At Level 2 the image processing in the cloud is carried out when the local server is overloaded, thus determining a balancing architecture workload and allowing the SAHHc architecture to be scalable. This level can often be used when the environment is a clinic or hospital.

The workload processing balance initializes the analysis to determine which level should be used, and if the selected level is Level 0, the processing flow will be in executed in the Raspberry Pi. Otherwise the system checks whether the local server is overloaded or not. If it is not overloaded, the processing will occur locally, i.e. the location server may receive the images from Raspberry via the NFS (Network File System) protocol, for performing the identification of the face and facial expressions. However, if the local server is overloaded, the architecture passes the processing to the Cloud, and thus balances the whole load. The transfer of images from the local server to the cloud was carried out by using the SFTP (SSH File Transfer Protocol) protocol that allows a reliable transfer of information. Any connection with the given database was made through the TCP (Transmission Control Protocol) protocol. Monitoring and making inferences about the possible status of the patient is not the only feature of SAHHc. It also enables the sensors to recognize the patients and analyze their feelings through a certain facial expression, for example, an expression of pain. All levels record the facial recognition results in the database and the information is automatically available through the web services by HTTPS connections.

5. Evaluation and discussion

5.1. Accuracy of the patient identification and emotion classification

In this paper, pictures of emotional facial expressions from the Extended Cohn-Kanade (CK+) [22] free-access database were used to validate the methodology employed for patient identification and emotion recognition. CK+ database shows facial expressions of adult actors, 69% female and 31% male. Among the sample population, 81% are European or American, 13% African-American and 6% belong to other ethnic groups. This database has 593 facial expressions, from 123 actors. However, only a subset of CK+ database was used for the experiments. Taking into account that the CK+ data set is composed by a sequence of images obtained from videos, there are many images that are very similar (almost identical). In this sense, the subset we have employed in the experiments is the CK+ with the removal of those very similar images, also this removal helps the creation of a balanced subset of relevant images which can improve the generalization performance of the ML classifiers investigated. In the current evaluation, we have taken into account frontal pictures alone for the emotion recognition.

For both experiments (i.e. patient identification and emotion classification), the performance of algorithms was analyzed separately by means of a k-fold cross-validation technique where $k = 10$, which provides us a more accurate estimate.

In ensuring that an approach was adopted that is similar to the way humans interpret the features of the face, multiple photos of the same individual were used for the patient identification model. Thus, different facial representations (neutral, happy, sad, afraid, angry and surprised) of the same individual is used to identify that person; using multiple facial images makes it possible to encode a unique configuration for each individual. The selected subset of CK+ database includes all the existing expressions of different actors. Thus, for the experiment conducted to identify the patient, 20 different individuals were studied, with 100 images each, making a total 2000 pictures, in order to achieve good degree of accuracy.

Table 2 shows the degree of accuracy with regard to the identification of the patients. The results demonstrate that the SVM algorithm outperforms the other solutions. Although other algorithms also achieve high rates of accuracy, such as kNN and EC, with 99% and 99.70%, respectively, SVM had a rate of accuracy of 99.75% in the experiments. Furthermore, although the EC achieved a rate of accuracy that is very close to SVM, Ensemble employs a large number of algorithms which significantly increases its computational costs which are higher than when a single ML algorithm is used. For this reason, we chose the SVM in our proposal for the identification of patients, since it achieves the most accurate rate and has lower computational costs than the other techniques used.

In the case of the emotion classification tests, the same 20 individuals were used as they had formed the sample in the patient identification experiment. Five images were randomly selected for each individual, making a total of 100 images for each run. The results demonstrate that the Ensemble technique enables a more precise classification to be made than the use of simple classifiers. This can be seen in Fig. 10 that shows the results of the classifications. In addition, it is possible to confirm a smaller dispersion

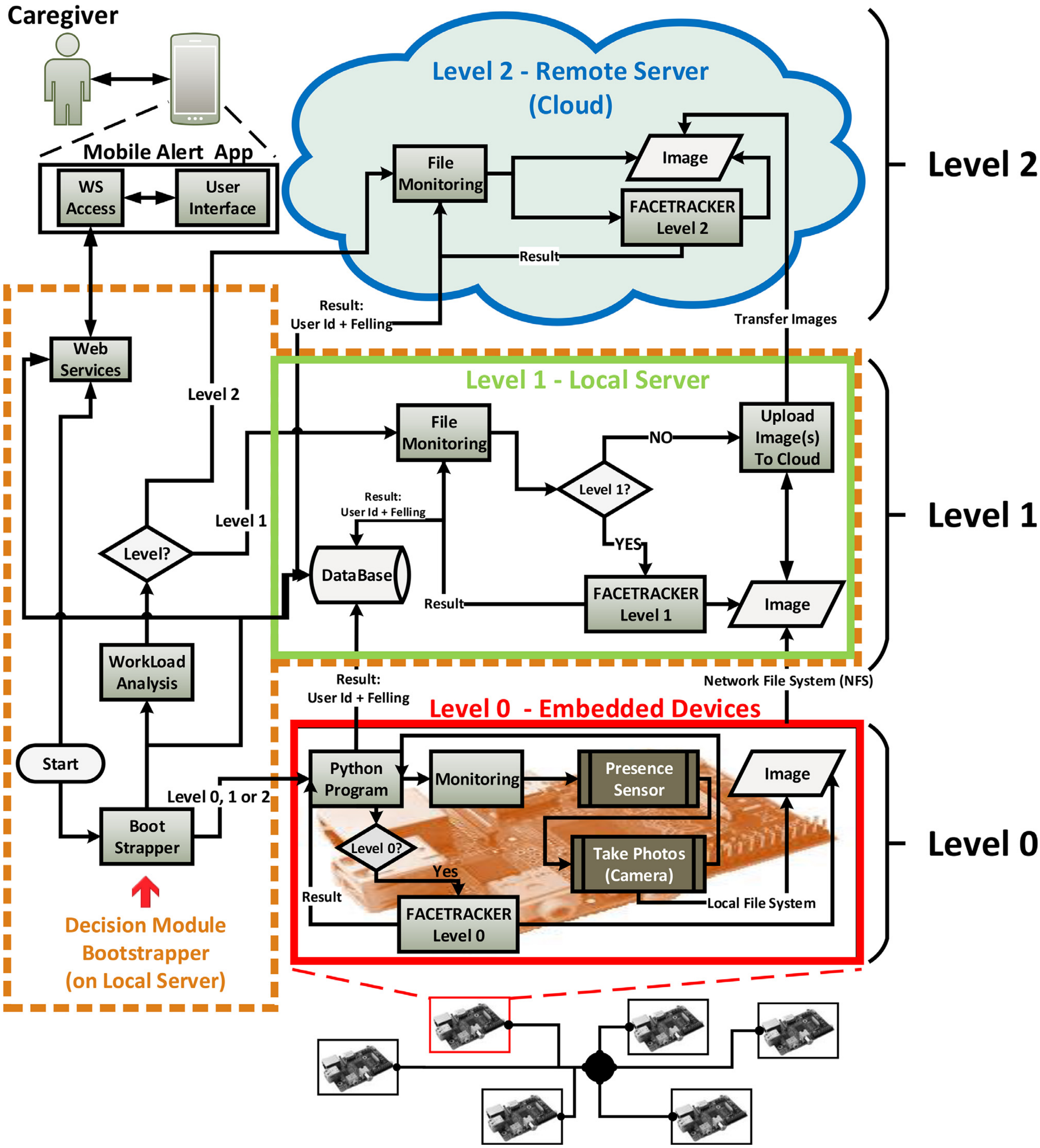


Fig. 9. Abstractions of the three levels which consisting architecture SAHHc: (i). Level 0 (red rectangle) is responsible for capturing images, (ii). Level 1 (green rectangle) is responsible for monitoring of the house and monitoring the health of many patients at home, (iii). Level 2 (blue cloud) is responsible for image processing in the cloud is carried out when the local server is overloaded. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the results obtained by the Ensemble, which shows a greater stability in their execution.

A statistical evaluation was performed to validate both results. Initially, the Shapiro–Wilk method was employed to verify the normality of the distribution error; it leads to parametric (t-test) or non-parametric tests (Wilcoxon–Mann–Whitney). Results from Shapiro–Wilk tests showed some p-values lower than 0.05, hence, it leaded us to refute the hypothesis of normality, considering 95% of confidence. In light of this, non-parametric tests must be carried out for next analysis.

The pairwise comparisons resulting from the Wilcoxon rank sum test are displayed in Table 3 for patient identification tests and in Table 4 for emotion classification. With regard to the procedure for the patient identification, the results show that the SVM algorithm has a statistically significant difference compared with other used techniques, except for kNN and Ensemble. However, as stated earlier, the use of SVM is more suitable for our proposal because it has a higher average rate of accuracy and lower computational cost. On the other hand with regard to emotion classification, the results show that the Ensemble has a statistically

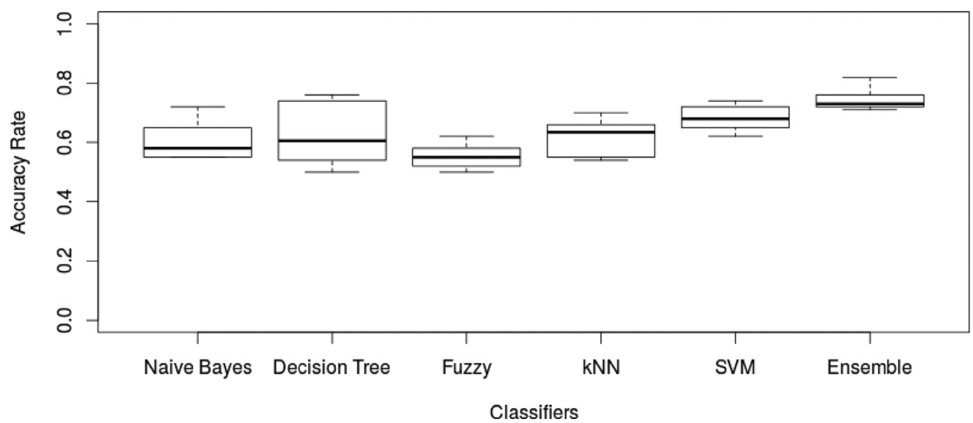


Fig. 10. Accuracy values for emotion recognition experiment.

Table 2
Accuracy values for patient identification experiment—best accuracy was achieved by SVM.

Accuracy table (%)–Person identification						
	N. Bayes	Dec. Tree	Fuzzy	kNN	SVM	Ensemble
Test 01	89.50	95.00	81.00	99.50	99.50	99.50
Test 02	92.50	93.00	86.00	99.50	100.00	100.00
Test 03	90.00	94.00	85.00	99.50	100.00	100.00
Test 04	91.00	93.50	80.00	97.50	100.00	100.00
Test 05	88.00	96.00	83.00	99.00	100.00	99.50
Test 06	86.00	92.50	84.00	99.00	99.50	99.50
Test 07	85.00	96.00	84.00	99.00	100.00	100.00
Test 08	87.50	95.00	81.00	98.50	99.50	100.00
Test 09	84.00	93.50	85.00	99.50	99.00	99.00
Test 10	86.00	94.00	82.50	99.00	100.00	99.50
Mean	87.95	94.00	83.00	99.00	99.75	99.70

Table 3
P-values of the pair-wise comparison with the Wilcoxon rank sum test for person identification model.

Wilcoxon rank sum test–Person identification					
	N. Bayes	Dec. Tree	Fuzzy	kNN	SVM
Dec. Tree	0.001	-	-	-	-
Fuzzy	0.005	0.001	-	-	-
kNN	0.001	0.001	0.001	-	-
SVM	0.001	0.001	0.001	0.072	-
Ensemble	0.001	0.001	0.001	0.072	1.000

Table 4
P-values of the pair-wise comparison with the Wilcoxon rank sum test for classifications of emotion model.

Wilcoxon rank sum test–Classification of emotion					
	N. Bayes	Dec. Tree	Fuzzy	kNN	SVM
Dec. Tree	1.000	–	–	–	–
Fuzzy	0.236	0.552	–	–	–
kNN	1.000	1.000	0.226	–	–
SVM	0.119	1.000	0.003	0.236	–
Ensemble	0.005	0.236	0.003	0.003	0.044

significant difference from other classifiers, except for Decision Tree. However, given the higher average degree of accuracy and the smallest dispersion of results displayed by the EC (see Fig. 10), we believe that it is the most efficient approach.

5.2. Computational performance of the implemented architecture

Another important factor that must be analyzed is the performance of the proposed methodology. This analysis allows us to find out the expected behavior and to investigate efficient ways of making new implementations, by not overloading the network nodes and preventing possible bottlenecks.

Two scenarios were defined for the next experiments, both implementing the same architecture, but employing the methodology proposed in different network elements: (i) the Sensor nodes and (ii) the Decision Maker node. In the first scenario, where the identification and classification of a patient's emotions is run in the Sensor nodes, the results allow us to find out the minimum interval between each image that is captured, since before capturing a new image, it is necessary to process the last captured image and transmit the results to the Decision Maker node. However, this approach restricts the performance because the Sensor devices have very limited resources. In the second scenario, the Sensor nodes only capture images and transmit them to the Decision Maker node, which is responsible for processing all the images and making the appropriate decisions. Although the Decision Maker node has more computing resources available, since there are several sensor nodes, the Decision Maker will make processing competitive and may be overloaded if the workload is too heavy.

Comparisons of the performance can be fairer, if it is assumed that the images are already stored in the elements that will employ the methodology. In this way, the run times for each experiment can be determined to measure the load level and allow them to be compared. 30 replications were carried out for each configuration at 10 s intervals to achieve a better statistical analysis of the results.

The experiments were planned to simulate real life situations more closely and represent the behavior of the architecture in these situations. Thus, in the first scenario (OnBoard—**Level 0**), where the Sensor nodes process the captured images, there is no competition for processing (it is assumed that the nodes are not employed in any other application) and regardless of the number of Sensor nodes present in the architecture; it is also possible to assume that the execution time is equal for all the nodes (since all the sensor nodes have the same microprocessor). On the other hand, in the second scenario (in the local server, the Decision Maker element processes all the images captured by Sensor nodes—**Level 1**), there may be competition for processing, with a maximum level equal to the number of sensor nodes that compose the architecture. In the worst case scenario, where the images are received by the Decision Maker at the same time and the various instances of the proposed methodology are initiated immediately, 1, 2, 4 and 8 were taken as the number of concurrent images being processed.

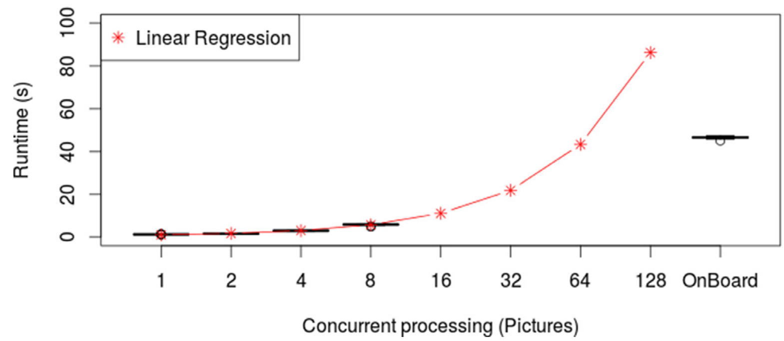


Fig. 11. Using linear regression to determine a result time longer than Raspberry Pi time. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

Fig. 11 displays the execution times obtained for each experimental setup. Although the OnBoard configuration has no competition for processing, this approach had the highest execution time of all. This is due to the limited computational resources in the Sensor nodes (in the experiments we considered an architecture with the Raspberry Pi model B for the Sensor nodes). Raspberry Pi model B has ARM® processor (one core—700 MHz of clock), 512 MB of RAM and SD Card of 8GB. The Decision Maker is hosted in a local server equipped with an Intel® i5 processor (four cores [two physical and two due to Hyper-Threading]—3 GHz of clock), 16 GB of RAM and a disk of 1 TB.

The results for the implementation in the Decision Maker node can be seen in the first four points of Fig. 11. These names also refer to the level of competition in each experiment. It can be seen that the running time grows, as the level of competition increases. However, even though the experiment provides up to eight simultaneous processes (twice the number of cores available in the desktop element that implements the Decision Maker), the longer running time observed in the Decision Maker node is about 5 times shorter than the running time of the OnBoard experiment.

On the basis of these results, we carried out a linear regression and used it to infer the level of competition necessary for the run-time, if the Decision Maker element becomes greater than the OnBoard, which suggests a saturation of the architecture. It can be assumed that the characteristics of the increase in run-time is maintained while the calculations are being made for new levels of competition. These results are indicated by asterisks in Fig. 11 with a straight line between the points, both in red.

Thus, it was found that when there is a competition involving 64 pictures, the execution time is close to that displayed by OnBoard, but is still lower. When the level of competition is equal to 128 pictures, it can be observed that the saturation of the architecture has an execution time around twice as high as in the OnBoard scenario.

This makes it possible to assume that the execution of the proposed methodology in the Decision Maker element (which has an identical implementation) is feasible with architectures that have up to 64 sensor elements. If a higher number of sensor elements is required, an approach that involves using the OnBoard scenarios will be more appropriate.

We performed a statistical evaluation aiming to verify whether the execution times from 1 to 8 pictures differ or not from the execution times for the OnBoard approach. The Wilcoxon rank sum test was employed and showed us that the sets are distinct.

These results demonstrate that the implementation of the proposed methodology in the Decision Maker element produced a satisfactory performance for small and medium-sized deployments. It is important to note that these results were produced by taking into account the number of sensor elements and this provided a shorter interval between the images of the patient that have been gathered. However, implementations with a large number of sensor elements are better suited to an OnBoard approach as a way of avoiding a possible saturation of the architecture.

5.3. Analysis of the effects of transferring images to be processed in the Cloud

An alternative strategy which the SAHHc can adopt is to use the Cloud to execute patient identification and emotion classification (**Level 2**). In this case, the architecture is not affected by the overhead when processing the images captured by the sensor elements. On the other hand, the transfer of images for processing in the cloud can cause congestion in the transmission and thus increase the time required for the dispatch of messages.

The concept of Cloud Computing, is of a system that has elastic computational resources and can thus keep the features required for a given service that can be contracted. In view of this, it can be assumed that the processing time can be set in accordance with the ongoing implementation. However, the time to send images does not form a part of the characteristics of services contracted because they depend on resources which are not controlled by the Cloud Computing provider.

The transmission time is part of the total time (which includes the time from image acquisition until the end of the proposed methodological procedure) and results in an increase in the minimum interval between the monitoring processes. With this in mind, we evaluated the impact of simultaneous transmissions that resulted in an increase in the minimum interval between the monitoring processes. For this reason, an experiment was conducted to measure the simultaneous transmission of images acquired by internal cameras. This was carried out through the SFTP communication tool. The connection for communication was established between a Decision node and a remote virtual server implemented in the Cloud.

The real server is a DELL PowerEdge R720, Intel® Xeon 32 cores (two physical processors with eight cores and HyperThreading (2x))—2 GHz of clock, 256 GB of RAM and 6 HDs SAS Hot-plug 3.5 in. of 600 GB in RAID 5. The virtual machine was hosted in this real server and used 4 VCPUs, 16GB of RAM and 20 GB of disk.

The image transmission process follows the same parallelism scheme that we conducted in the previous experiments, i.e. 1, 2, 4, 8 images are sent through the network simultaneously. The elapsed time for sending this number of images in parallel can be viewed in Fig. 12.

According to what is shown in Fig. 12, the transfer time increases linearly as the parallelism increases. In view of this, a linear regression was used to estimate the transfer time for 16, 32, 64 and 128 pictures. It can be said that the total processing time did not exceed the processing time carried out by our Onboard mode. This was in the case of (a) if a virtual cloud infrastructure is

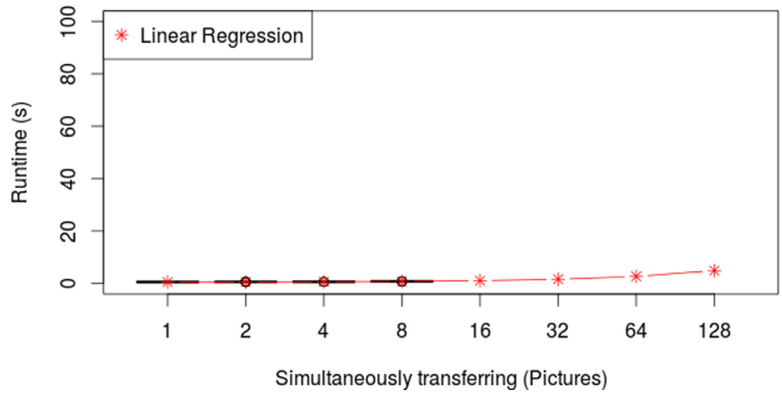


Fig. 12. Using linear regression to demonstrate the increase in the simultaneous image transmission time of the local server to the virtual server in the Cloud.

employed to obtain a similar response time for processing 32 images simultaneously (Fig. 11), or (b) one uses a server with a similar configuration to the Decision Maker element. This help us to avoid the saturation of the architecture that is implemented for processing 128 pictures. The reason for this is that the server will adapt its computing resources to keeping to the agreed response time.

We performed a statistical evaluation aiming to verify whether the transmission times of images by the Decision Maker element differ according to different number of images. Results of the statistical test showed that, apart from the simultaneous transmission of 2 and 4 images, the other transmission figures have significant differences.

Such results show that the transfer time increases significantly when compared with the growth of parallelism, which suggests that a saturation point can also be obtained. However, this approach provides increased robustness for the implementation of SAHHc since it enables the processing of a larger number of images without exceeding the time shown in the OnBoard approach.

6. Final remarks

This paper discussed the use of images and emotions for helping the healthcare in smart home environments in an automatic way through an IoT infrastructure (i.e. without the need for human intervention to detect new features). We made use of images to identify each person to ensure that the right person is tracked in the house and that a particular kind of treatment is ensured. Experiments were conducted on this front and obtained an accuracy of 99.75% when identifying each person. Account was taken of the fact that people in the house could have their own individual healthcare scheme or undergo specialist treatment. We also drew on the images to check if the emotional part of each person relies on the same IoT technology. The rate of accuracy was again around 80% and there was a good convergence obtained through our proposed Ensemble model. According to psychologists, emotions play a crucial role while a patient is trying to recover from a wide range of diseases. Hence, we believe that this is an important feature to measure while monitoring patients. All these IoT developments and experiments were conducted to show the suitability of our approach which was carried out by finding a prototype for our system in resource constrained devices. As a result, we adopted an approach that took into consideration all these features and conducted experiments to validate the whole idea in terms of performance, accuracy and statistical analysis. As future research in this area, we plan to exploit evolutionary mechanisms to ensure that the classification algorithms of facial expression can make progress while taking into account advances in the treatment of certain ailments like Parkinson's disease. Patients of Parkinson's disease do not have full control of their facial expressions particularly when the disease progresses. This can be again a good challenge for our model as it will require processing cycles and communications along with evolutionary approaches in a heterogeneous environment. Furthermore, it is worthwhile to mention that we plan to improve our emotion recognition system by taking into account different facial angles, gestures and postures.

Acknowledgments

Authors would like to thank the financial support from CNPq, CAPES and FAPESP (process ID 2014/10685-3, 2013/18859-8, 2014/19076-0 and 2014/06330-5) for funding the bulk of this research project. Jó Ueyama would like to thank Office of Naval Research Global (ONRG) for funding part of his research project.

References

[1] G.D. Abowd, A.F. Bobick, I.A. Essa, E.D. Mynatt, W.A. Rogers, The aware home: A living laboratory for technologies for successful aging, in: Proceedings of the AAAI-02 Workshop "Automation as Caregiver, 2002, pp. 1–7.
[2] J. Augusto, P. McCullagh, V. McClelland, J. Walkden, Enhanced healthcare provision through assisted decision-making in a smart home environment, in: IEEE Second Workshop on Artificial Inteligence Techniques for Ambient Inteligence, IEEE Computer Society, 2007.
[3] J.N. Bailenson, E.D. Pontikakis, I.B. Mauss, et al., Real-time classification of evoked emotions using facial feature tracking and physiological responses, Int. J. Hum. Comput. Stud. 66 (5) (2008) 303–317.
[4] T. Banerjee, J.M. Keller, M. Popescu, M. Skubic, Recognizing complex instrumental activities of daily living using scene information and fuzzy logic, Comput. Vis. Image Understand. 140 (2015) 68–82.
[5] P. Ekman, Cross-cultural Studies of Facial Expression, Academic, New York, 1973.
[6] P. Ekman, Darwin and Facial Expression: A Century of Research in Review, Ishk, 2006.
[7] P. Ekman, W.V. Friesen, J.C. Hager, in: The Facial Action Coding System, second ed., Research Nexus eBook, Salt Lake City Weidenfeld and Nicolson London, 2002.
[8] G. Filho, J. Ueyama, B. Faical, D. Guidoni, L. Villas, Residi: an intelligent decision platform for residential infrastructure using wireless sensors and actuators, in: 2015 XXXIII Brazilian Symposium on Computer Networks and Distributed Systems (SBRC), 2015, pp. 237–245, doi:10.1109/SBRC.2015.36.
[9] G.P.R. Filho, J. Ueyama, L.A. Villas, A.R. Pinto, V.P. Goncalves, G. Pessin, R.W. Pazzi, T. Braun, Nodepm: a remote monitoring alert system for energy consumption using probabilistic techniques, Sensors 14 (1) (2014) 848–867, doi:10.3390/s140100848.
[10] G.P.R. Filho, J. Ueyama, L.A. Villas, A.R. Pinto, V.P. Goncalves, S. Seraphini, An intelligent approach for improving energy efficiently in smart grids, in: International Conference on Wireless Networks (ICWN), 12, 2013, pp. 227–233.
[11] B.L. Fredrickson, R.W. Levenson, Positive emotions speed recovery from the cardiovascular sequelae of negative emotions, Cognit. Emotion 12 (2) (1998) 191–220.
[12] D. Freitas, T. Marcondes, L.H.V. Nakamura, J. Ueyama, P.H. Gomes, R. Meneguette, Combining cell phones and WSNs for preventing accidents in smart-homes with disabled people, in: NTMS'2015—New Technologies and ServicesTrack (NTMS'2015 New Technologies Track), Paris, France, 2015.

[13] V.P. Goncalves, V.P. de Almeida Neris, S. Seraphini, T.C. Dias, G. Pessin, T. Johnson, J. Ueyama, Providing adaptive smartphone interfaces targeted at elderly people: an approach that takes into account diversity among the elderly, Universal Access Inf. Soc. (2015) 1–21, doi:10.1007/s10209-015-0429-9.
[14] V.P. Gonçalves, V.P. de Almeida Néris, J. Ueyama, S. Seraphini, An analytic approach to evaluate flexible mobile user interfaces for the elderly., in: ICEIS (3), 2012, pp. 91–96.
[15] R.E. Gur, C.G. Kohler, J.D. Ragland, S.J. Siegel, K. Lesko, W.B. Bilker, R.C. Gur, Flat affect in schizophrenia: relation to emotion processing and neurocognitive measures, Schizophr. Bull. 32 (2) (2006) 279–287.
[16] S. Helal, W. Mann, J. King, Y. Kaddoura, E. Jansen, et al., The gator tech smart house: a programmable pervasive space, Computer 38 (3) (2005) 50–60.
[17] Intel Corporation, Internet-of-Things (IoT) Advances Home Healthcare for Seniors, http://goo.gl/aFye65, 2015.
[18] A. Ivanova, Humor recognition in psychiatric patients and artificial intelligence, in: 2012 AAAI Fall Symposium Series, 2012.
[19] J.A. Kientz, S.N. Patel, B. Jones, E. Price, E.D. Mynatt, G.D. Abowd, The Georgia tech aware home, in: CHI'08 Extended Abstracts on Human Factors in Computing Systems, ACM, 2008, pp. 3675–3680.
[20] O. Lehmann, M. Bauer, C. Becker, D. Nicklas, From home to world-supporting context-aware applications through world models, in: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications, 2004. PerCom 2004, IEEE, 2004, pp. 297–306.
[21] G. Libralon, R. Romero, Mapping of facial elements for emotion analysis, in: Proceedings of the Brazilian Conference on Intelligent Systems, 2014.
[22] P. Lucey, J.F. Cohn, et al., The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2010.
[23] S. Mahlke, M. Minge, Consideration of multiple components of emotions in human–technology interaction, in: Affect and Emotion in HCI, Springer, 2008.
[24] M.K. Mandal, R. Pandey, A.B. Prasad, Facial expressions of emotions and schizophrenia: a review, Schizophr. Bull. 24 (3) (1998) 399.
[25] A. Martinez, S. Du, A model of the perception of facial expressions of emotion by humans: research overview and perspectives, J. Mach. Learn. Res. 13 (2012) 1589–1608.
[26] T. Mitchell, Machine Learning, McGraw Hill, 1997.
[27] G.V. Ostir, I.-M. Berges, M.E. Ottenbacher, A. Clow, K.J. Ottenbacher, Associations between positive emotion and recovery of functional status following stroke, Psychosom. Med. 70 (4) (2008) 404–409.
[28] V. Rialle, F. Duchene, N. Noury, L. Bajolle, J. Demongeot, Health "Smart" home: information technology for patients at home, Telemed. J. e-Health 8 (4) (2002) 395–409.
[29] G. Riva, Ambient intelligence in health care, Cyber Psycholo. Behav. 6 (3) (2003) 295–300.
[30] E. Romero, A. Araujo, J.M. Moya, J.-M. de Goyeneche, J.C. Vallejo, P. Malagon, D. Villanueva, D. Fraga, Image processing based services for ambient assistant scenarios, in: Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living. Lecture Notes in Computer Science, 5518, Springer, Berlin, Heidelberg, 2009, pp. 800–807.
[31] L. Russek, Alarm for Patient Monitor and Life Support Equipment System, 1994, US Patent 5,319,355.
[32] J. Russell, A circumplex model of affect, J. Personality Soc. Psychol. 39 (1980) 1161–1178.
[33] J.E. Salem, A.M. Kring, S.L. Kerr, More evidence for generalized poor performance in facial emotion perception in schizophrenia, J. Abnorm. Psychol. 105 (3) (1996) 480.
[34] J.M. Saragih, S. Lucey, et al., Deformable model fitting by regularized landmark mean-shift, Int. J. Comput. Vis. 91 (2) (2011) 200–215.
[35] K.R. Scherer, What are emotions? and how can they be measured? Soc. Sci. Inf. 44 (4) (2005) 695–729.
[36] B. Schuller, S. Reiter, R. Muller, et al., Speaker independent speech emotion recognition by ensemble classification, in: IEEE International Conference on Multimedia and Expo, 2005. ICME 2005, 2005.
[37] J.A. Stankovic, Q. Cao, T. Doan, L. Fang, Z. He, R. Kiran, S. Lin, S. Son, R. Stoleru, A. Wood, Wireless sensor networks for in-home healthcare: potential and challenges, in: High Confidence Medical Device Software and Systems Workshop (HCMDSS), 2005.
[38] Y. Sui, C. Ahn, C. Ahn, A new smart fall-down detector for senior healthcare system using inertial microsensors, in: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, 2014, pp. 590–593, doi:10.1109/EMBC.2014.6943660.
[39] Y.L. Tian, T. Kanade, C.J. F., Handbook of Face Recognition, Springer, 2005.
[40] J. Ueyama, V.P. Pinto, E.R. Madeira, P. Grace, T.M. Jonhson, R.Y. Camargo, Exploiting a generic approach for constructing mobile device applications, in: Proceedings of the Fourth International ICST Conference on COMmunication System softWAre and MiddlewaRE, 2009.
[41] WowWee, Rovio—Mobile Webcam, http://goo.gl/TmKu5v, 2015.
[42] N. Yang, I. Demirkol, W. Heinzelman, Motion sensor and camera placement design for in-home wireless video monitoring systems, in: 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011), 2011, pp. 1–5, doi:10.1109/GLOCOM.2011.6133718.