Recognition of emotions in the elderly through facial expressions: Emotions biofeedback as a therapy support tool

Arianne Sarmento Torcate[1], Maíra Araújo Santana[1], Juliana Carneiro Gomes[1], Ingrid Bruno Nunes[1], Flávio Secco Fonseca[1], Gisele Machado Magalhães Moreno[2] and Wellington Pinheiro dos Santos[3*†]

[1]Polytechnic School of the University of Pernambuco, Recife, Brazil.
[2]Institute for Astronomy, Geosciences and Atmospherical Sciences, University of São Paulo, São Paulo, Brazil.
[3]Department of Biomedical Engineering, Federal University of Pernambuco, Recife, Brazil.

*Corresponding author(s). E-mail(s): wellington.santos@ufpe.br;
Contributing authors: ast@ecomp.poli.br; mas2@ecomp.poli.br; jcg@ecomp.poli.br; ibn@ecomp.poli.br; fsf2@ecomp.poli.br; gisellemoreno@usp.br;
†These authors contributed equally to this work.

Abstract

Purpose: Data provided by the World Health Organization indicate that the number of people aged 60 years or more in the world population is gradually increasing. Moreover, projections clarify that in 2030 there will be 1.4 billion elderly people and 2.1 billion in 2050. Along with the natural aging process of any human being, changes in perception and cognition can cause damage that contributes to elderly people having difficulties both to recognize and to express emotions through the face. Few researches in the literature address the recognition of emotions in elderly, whether they are affected by dementia processes or not. In consequence of this, few databases are made available for carrying out work and experiments. Not being able to express and recognize emotions

[*]Corresponding author: Wellington Pinheiro dos Santos, ORCID 0000-0003-2558-6602

through the face can contribute to the elderly having difficulties in communicating important messages, which may even compromise their physical integrity. **Methods:** Therefore, this work aims to develop an application for Emotion Recognition in the elderly through Facial Expressions. For this, we first used Haar cascade Frontal Face for face detection and implemented a Convolutional Neural Network to classify emotions, using FER2013 database for training, validation and testing. In a second part of the methodology, in order to assess the performance of the algorithm in this context, we applied the developed model to recognize emotions in static images of elderly people. **Results:** As a result, the accuracy achieved by the developed model was 0.6375. From the images tested, for 52.63% of them the model was able to detect the face and identify some emotion. On the other hand, in 47.37% of the images, the model had difficulty both in detecting the face and in identifying emotions. **Conclusion:** Finally, the findings and discussions exposed in this work are promising, we also found and shared limitations and related them to our goals for future works. The possibility of developing intelligent systems that support emotion recognition in elderly population emerges as a valuable tool, representing an alternative to promote not only quality of life for the elderly themselves, but also for the entire support network around them.

Keywords: Recognition of Emotions, Elderly, Facial Expressions, Affective Computing, Emotions biofeedback

1 Introduction

1.1 Problem characterization

Population aging is a global phenomenon that has been growing steadily in recent years. About half of the current population of elderly people over 75 years old suffers from physical and/or mental disabilities, with dementia being one of the major challenges affecting the quality of life of the elderly and also of their caregivers (Dantcheva, Bilinski, Nguyen, Broutart, & Bremond, 2017). It is important to clarify that along with the aging process, changes in perception and cognition can damage the recognition of facial emotions (C.D. Ferreira & Torro-Alves, 2016). Ochi and Midorikawa (2021) emphasize that in the literature the relationship between emotion recognition and cognitive function during aging is still unclear. However, works in the literature (C.D. Ferreira & Torro-Alves, 2016; Grondhuis, Jimmy, Teague, & Brunet, 2021; Ko et al., 2021) highlight at least four possible causes: (i) impairment of brain structures responsible for processing emotions; (ii) issues related to the natural aging process, such as wrinkles and folds that mask the emotion displayed, thus making the process of interpreting emotions difficult; (iii) the socio-emotional selectivity theory, which states that the elderly would have greater preservation in the recognition of positive emotions; and, finally, iv) atrophy of the facial skeleton, loss of soft tissue and poor muscle positioning.

Basic emotions can be understood as involuntary physiological responses shared by human beings, visually distinguishable and shaped by lifelong experiences (Bomfim, Ribeiro, & Chagas, 2019). Therefore, the ability to express and recognize emotions through facial expressions is considered a fundamental stage of basic communication (Ko et al., 2021) and an essential skill to get along in society (Bomfim et al., 2019). Not being able to express emotions such as anger, sadness or disgust can result in social isolation (Grondhuis et al., 2021) or negatively affect non-verbal communication. In this sense, by decreasing the ability to express emotions, older people may have difficulty communicating important messages such as discomfort associated with treatments and others complications.

With this problem in mind, technology has become a strong ally to increase the quality of life of the elderly (Dornelles & Corrêa, 2020), whether they have any cognitive impairment or not. Although the technological barrier is still a very present point in older generations, current devices and software have broken paradigms and have been adapted to achieve this audience (Zilidis & Zilidou, 2018). Basically, what we have is a context in which the increase in life expectancy generates more demands for assistance, care and trained professionals, which are required for longer and longer periods of time. These very specific health resources become unfeasible when we talk about large-scale policies. Iancu and Iancu (2017) state that assistive technologies are the only solution for this problem. In addition to supporting possible responses to health, assistive technologies seek to integrate these elderly people in a society that is sometimes distant and in which isolation has become not only social but also digital.

It is in this digitalized world that Affective Computing emerges. Uniting Psychology, Cognitive Sciences and Computer Science. This research area seeks to integrate the human emotional side to machines. Either through the recognition of emotions (Han, Zhang, Cummins, & Schuller, 2019) or even in the development of devices and studies that deal with the human aspect. It is precisely from the perspective of affective computing that we will conduct this work.

The main goal of this research is to develop an application capable of recognizing emotions in elderly people through facial expressions. For the development of the experiments, we used Google Colab and Python programming language, along with libraries such as OpenCV and Tensor Flow. Face detection was performed with the Haarcascade Frontal Face (Viola & Jones, 2001). For classification, we chose the Convolutional Neural Networks (CNN or ConvNet) algorithm. For training, validation and testing of the model, we used the Facial Expression Recognition 2013 (FER2013) (Goodfellow et al., 2013) database.

This work is structured as follows: in addition to the Introduction, Section 2 with its respective subsections present important theoretical references for understanding the present research topic. In Section 3 related works are reported. Section 4 presents the methodological approach adopted to carry out the experiments. Section 5 describes the results obtained in detail. Finally, in Section 6 we highlight the final considerations, as well as the perspectives for future work.

1.2 Background

This section presents the main theoretical references that support the realization and understanding of this research. Thus, in subsection 2.1 we briefly talk about aging, cognitive deficits and dementias. In subsection 2.2 we conceptualize emotions and their importance for any human being. Subsection 2.3 presents the field of study of Affective Computing. In subsection 2.4 we bring an explanation about the recognition of emotions and make a specific topic for facial expressions in subsection 2.4.1. Finally, in order to logically limit the theoretical framework to the scope of this study, we end with subsection 2.5, where we present a discussion on the recognition of emotions in the elderly through facial expressions.

1.2.1 Aging, Cognitive Deficits and Dementias

According to the WHO (2018), the number of people aged 60 years or more in the population is gradually increasing, where in 2030 projections indicate that there will be 1.4 billion elderly people and in 2050 it will reach 2.1 billion. There are those who see this milestone from two perspectives. The first of them we can understand based on Tavares et al. (2017) explanation, where the authors point out that being able to achieve this longevity is a success for humanity since the elderly can contribute to society through knowledge, skills and experiences. On the other hand, the second perspective exposed by Rudnicka et al. (2020) states that population aging is the most important medical and social demographic problem in the world. The WHO (2021) emphasizes that it is necessary to deconstruct these ideas and attitudes related to age, as it can lead to discrimination and directly affect the way the elderly and society face aging.

According to Hayashi et al. (2021), as a result of the aging process, people have cognitive, biological and psychological changes. Where about half of the current elderly population over 75 years old suffers from physical and/or mental disabilities, with dementia being one of the great challenges that affect the quality of life of the elderly and their caregivers (Dantcheva et al., 2017). Cognitive deficit can be understood as the difficulty the brain has to learn, concentrate and remember, but it does not change the elderly's quality of life (Hayashi et al., 2021). Dementia, on the other hand, is characterized by a gradual impairment of cognitive function that interferes with social and professional activities (de Sousa Silva, dos Santos Zanetti, de Barros, de Souza, & Barreto, 2021), causing damage to the quality of life of the elderly and their families.

Alzheimer's Disease (AD) is the most common form of dementia (Dantcheva et al., 2017; Guo et al., 2020; Reale, Gonzales-Portillo, & Borlongan, 2020; Torcate et al., 2020), affecting 47 million people worldwide (Cavalli et al., 2020). In addition to AD, Vascular, Parkinson, Senile, frontotemporal and other dementia also stand out. The WHO (2018) clarifies that current trends, such as technology, should be used to devise strategies aimed to promote health and quality of life of the elderly population. For example, games are already being used to motivate elderly people to practice physical exercise, contributing to their motivation and engagement (Crespo, Idrovo, Rodrigues, & Pereira, 2016). Virtual Reality (VR) applications are already being used to assist in the rehabilitation of stroke survivors (Cameirão, Pereira, & i Badia, 2017) and personalized therapies based on social robotics (Agres et al., 2021) are also being applied. Based on this, several approaches and contributions have emerged in the field of Artificial Intelligence, mainly focusing on the Affective Computing sub-area.

1.2.2 Emotions

Although it seems simple, defining what emotions are is complex, as this term is often used in everyday life (Paxiuba & Lima, 2020). Bomfim et al. (2019) states that basic emotions can be understood as involuntary physiological responses, which are visually distinguishable and also molded according to our life experiences. The research carried out by Madeira (2011) explains that emotion is not a single variable or entity that can be easily identified, but it is a process that has distinct elements that can be related, such as sensations, facial expressions, body movement, voice, breathing, and heart rate (de Oliveira & Jaques, 2013).

In this context, emotions can be understood as one of the most important and fundamental daily experiences for the regulation of social interaction and interpersonal functioning (C.D. Ferreira & Torro-Alves, 2016), in addition to directly guiding our choices, preferences and decision-making (Chaturvedi et al., 2021), being the basis for our motivation and essential for verbal and non-verbal communication (Dorneles, Barbosa, & Barbosa, 2020; Marosi-Holczberger et al., 2012).

There are at least two ways to classify emotions, which is through i) Discrete (or categorical) Model and ii) Two-Dimensional Model (de Santana et al., 2020). In the Discrete Model, emotional states are represented and can be categorized by six basic emotions, which are: Sadness, Joy, Disgust, Anger, Fear and Surprise (Bomfim et al., 2019; de Oliveira & Jaques, 2013; Han, Zhang, Pantic, & Schuller, 2021; Khateeb, Anwar, & Alnowami, 2021). On the other hand, the Two-Dimensional Model is based on two main categories, which are: Valence and Excitement. The valence dimension represents the degree of likeability or dislike of a signal and is usually measured on a continuous scale ranging from positive (nice) to negative (unpleasant) (Bhattacharya, Gupta, & Yang, 2021). While the Excitement dimension measures the degree of intensity of an emotional state (Han et al., 2021). The Figure 1 below illustrates the two-dimensional model, considering valence and excitation.

It is important to emphasize that each emotion has a particular definition, characteristic, facial expression and peculiar way of manifesting itself in each individual. Finally, emotions can be understood as brief psychophysiological phenomena, which are adaptive in relation to environmental changes. Therefore, they prepare the organism to act and respond to stimuli (Madeira, 2011). Knowing that emotions are brief, involuntary and with different physiological patterns, there are several methods to recognize the transmitted emotion (Okada, 2018). Moreover, according to Marosi-Holczberger et al. (2012), it is precisely because there are more and more different methods for recognizing emotions that this field of research is constantly growing.

1.2.3 Affective Computing

It is known that emotions play a vital role in our daily life activities, influencing communication and personal and social development of any individual (Khateeb et

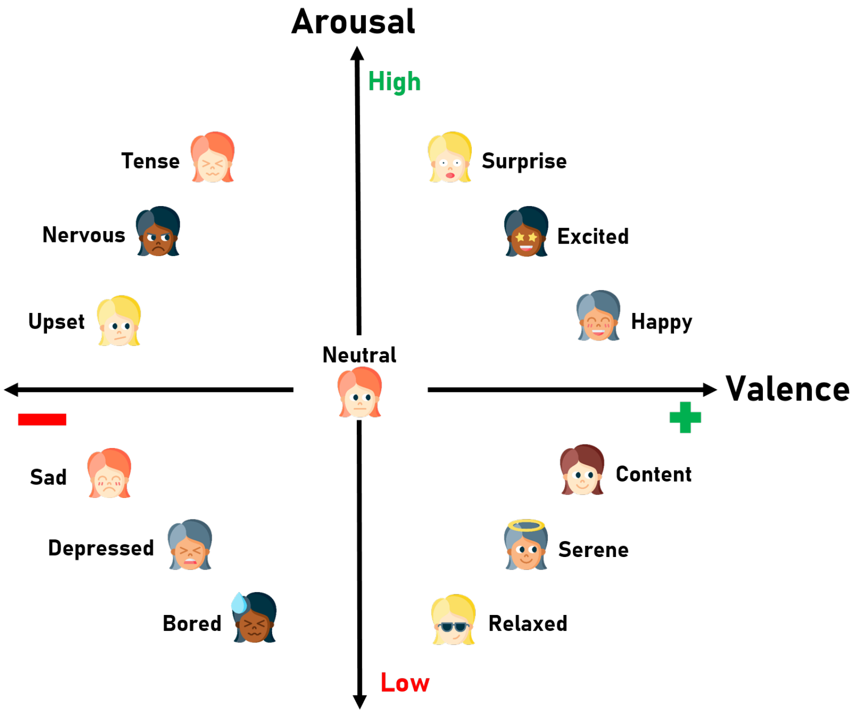

Figure 1 Two-Dimensional Model, also known as Circumplex Model of Affect, proposed by Russell (1980). The representation of the emotions happiness and sadness are located oppositely in two-dimensional space. Neutral emotion represents an intermediate state, that is, the subject does not feel a preominance of valence or manifestation of affective state.

al., 2021). The subarea of Artificial Intelligence that studies emotion is called Affective Computing (AC) (Paxiuba & Lima, 2020). Nalepa, Kutt, and Bobek (2019) explains that the term "Affective Computing" was proposed by Rosalind Picard, in 1997. And, it is an interdisciplinary field of study with other areas of knowledge (such as Biomedical Engineering, Psychology, Computer Science and others) and that it seeks to develop computational models and emotion recognition methods in order to improve human-computer interaction (HCI).

For a long time, recognizing human emotions through the computer has been one of the challenges and main focus of Affective Computing (Bhattacharya et al., 2021; de Sousa, Costa, Pires, & Araújo, 2020). Briefly, it can be understood that AC studies how computers can recognize, interpret, model and express emotions (and other human psychological aspects) (Khateeb et al., 2021; Picard, 2000). Tan, Šarlija, and Kasabov (2021) explains that the idea is that computer systems are not only able to identify the emotional states of users, but may also be able to generate responses that humans perceive as emotional or affective.

According to Sousa, Costa, Pires, and Araújo (2016), AC field of study can be viewed from two perspectives. The first is that AC studies the synthesis of emotions in a machine, that is, when you want to insert human emotions in the machine. The second perspective investigates the recognition of human emotions by machines during the interaction between human and computer. In this context, it is important to emphasize that sentiment analysis and affective computing play an essential role in enabling the emotional intelligence of machines (Han et al., 2021).

Knowing the interdisciplinary potential of this area, the ability to automatically recognize emotions has valuable implications for a variety of applications (Bhattacharya et al., 2021), for example, in games that aim to improve the user experience (Setiono, Saputra, Putra, Moniaga, & Chowanda, 2021), on education (Jaques, Nunes, Isotani, & Bittencourt, 2012), in analyzing language-based emotions to measure consumer satisfaction (Ren & Quan, 2012), stress detection to improve the health of individuals (Greene, Thapliyal, & Caban-Holt, 2016), and Emotion analysis of customer telephone complaints (Gong, Dai, Ji, Wang, & Sun, 2015).

1.2.4 Emotion Recognition

Considering that AC seeks to make the emotion existing in communication between people also present in the interaction between humans and computers, one of the main objectives is to investigate the recognition of emotions.Therefore, emotions recognition can be defined as a process of automatic perception of human affection (Han et al., 2021). Abdullah, Ameen, Sadeeq, and Zeebaree (2021) add that the recognition of emotions is a dynamic process that aims to identify the emotional state of individuals, where the emotions corresponding to each person's actions are different.

However, for emotion recognition to happen, human data must be collected. It is worth mentioning that there are several ways for a person to express/communicate their emotions, which can be either verbal or non-verbal (Abdullah et al., 2021; de Oliveira & Jaques, 2013). According to González and McMullen (2020), data can originate from different sources, such as Galvanic Skin Response (GSR), Facial Expressions, Electroencephalographic (EEG) signals, tone of voice, electrocardiograms (ECG), Eye Tracker and others. The authors also clarify that each data source has specific methodologies and application contexts. For example, the EEG is most commonly used to provide insights into emotions, while skin conductance is known to be the most reliable source for assessing stress and arousal.

It is important to highlight that collecting spontaneous data related to emotions is a challenging task due to the duration of emotions, which are usually brief. Knowing this, in research environments it is necessary to place the individual in situations where emotions are aroused/stimulated by using odors (Meska, Mano, Silva, Pereira, & Mazzo, 2020), visual or auditory stimuli, which can be images, videos or music (Vicencio-Martínez, Tovar-Corona, & Garay-Jiménez, 2019) and others. In the literature, it is pointed out that the recognition of emotions can be performed in two ways. The first is unimodal, that is, using only one data source, such as voice (Leão, Bezerra, Matos, & Nunes, 2012). The second way is multimodal, using combined data sources, such as voice and GSR (de Oliveira & Jaques, 2013; Han et al., 2021).

1.2.5 Facial Expressions

Facial expressions are considered one of the most powerful, immediate and natural ways for human beings to communicate their emotions and intentions (de Sousa et al., 2020). Namba, Matsui, and Zloteanu (2021) explain that facial expressions can be understood as affective signals that convey social information about an individual's experience of an emotional event. The ability to recognize and express emotions through the face is a fundamental stage of basic communication (Ko et al., 2021) and an important social skill (Grondhuis et al., 2021).

Tian, Kanade, and Cohn (2005) reinforce that facial expressions are facial changes in response to a internal emotional states or social communications. Moreover, A. Ferreira, Teixeira, and Tavares (2013) highlights seven universally recognized facial expressions of emotions, which can be identified as: Disgust, Surprise, Fear, Anger, Happy, Sadness and Contempt (Figure 2).

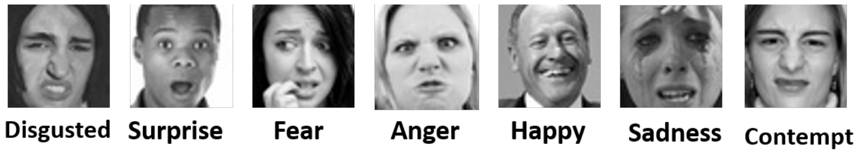

Figure 2 Images from the FER2013 (Goodfellow et al., 2013) database to illustrate the seven universally recognized facial expressions of emotions.

In the research carried out by A. Ferreira et al. (2013), two important scales are presented to understand the intensity of emotions in facial expressions. The first is the Intensity Scale, which basically defines the muscular contraction of an expression that starts in the neutral state until the peak emotion, that is, maximum expression of the emotion. The authors present an example adopted by Ekman, Friesen, and Hager (2002), which divides the scale into five categories, as shown in Figure 3, where A means the degree of intensity of the traits, B is the minimum, C means marked/pronounced, D is severe/extreme and E means the maximum.

Figure 3 The Intensity Scale measures in five categories the intensity of facial features to express emotions, ranging from the lowest degree (A), which is when the emotion starts in the neutral state, to the maximum degree (E), which represents when the face reached the peak of the desired emotion (Ekman et al., 2002; A. Ferreira et al., 2013).

The second scale presented by A. Ferreira et al. (2013) it called the energy scale, created by Trnka and Stuchlíková (2011). Such as may be seen in Figure 4, the authors created three intensities, corresponding to Little, Medium and Very, where emotions are organized in order, from those that need less energy to those that need more energy to manifest.

Finally, it is worth noting that the recognition of emotional facial expressions is directly related to non-verbal social behavior and the adaptation of human beings to

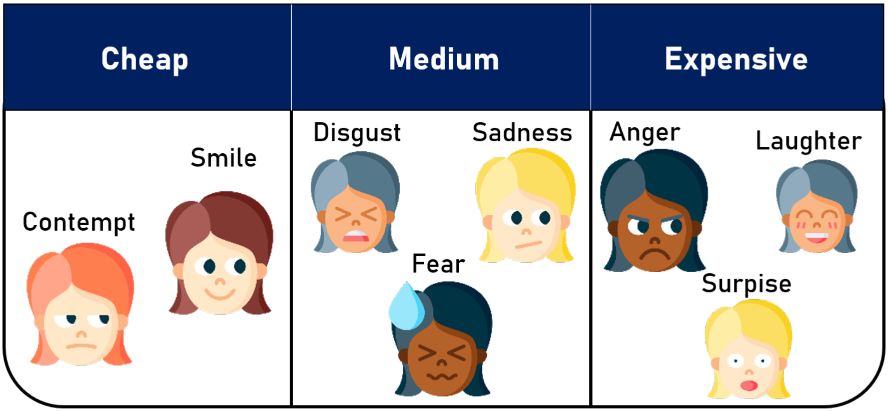

Figure 4 The Energy Scale illustrates and organizes the emotions facials on a scale of Low, Medium and High, highlighting which of them are they emotions most need energy to manifest, as well as those that need it least (A. Ferreira et al., 2013; Trnka & Stuchlíková, 2011).

different contexts (Nozima, Demos, & Souza, 2018). Recognizing emotions through the face consists of identifying patterns of facial features, such as the shape of the mouth, face, distance between eyes and others (de Sousa et al., 2020).

1.2.6 Recognition of Emotions in the Elderly through Facial Expressions

As already mentioned, recognizing emotional facial expression is an important social skill (Grondhuis et al., 2021). However, the ability to express and recognize emotions through facial expressions changes with age (Ko et al., 2021), especially with old age. Chuang et al. (2021) explain that facial recognition of emotions is one of the essential components of social cognition, being crucial for people to be able to recognize and express emotions (Ochi & Midorikawa, 2021). Ma et al. (2019) clarify that most of the research that covers automatic recognition of emotions is focused on adults and, to a lesser extent, the elderly, even though these are a significant part of the population and that it is constantly growing.

Studies in the literature (Fölster, Hess, & Werheid, 2014; Grondhuis et al., 2021; Ko et al., 2021) list possible causes of natural changes related to aging that contribute to elderly people having difficulties to express and recognize emotions through the face, such as i) wrinkles, folds and wear of facial muscles (Fölster et al., 2014; Grondhuis et al., 2021) and ii) atrophy of the facial skeleton, loss of soft tissue and poor muscle positioning (Ko et al., 2021). These changes contribute for the elderly to have a decline in the recognition/perception of emotions, especially low intensity negative ones (C.D. Ferreira & Torro-Alves, 2016), such as Fear, Anger and Sadness (Ko et al., 2021; Micillo, Stablum, & Mioni, 2021).

Elderly people who have Alzheimer's Disease (AD) (Ladislau, Guimarães, & Souza, 2015), Parkinson's (Nozima et al., 2018), Depression (Bomfim et al., 2019) and others diseases also have compromised brain structures that make up the system responsible for processing emotions, which results in damage to the recognition of emotional expressions on the face. In addition, the poor performance in recognizing the emotional facial expression of these elderly people is related to the progression of the disease.

In this scenario, Dantcheva et al. (2017) emphasizes that the recognition of emotions in the elderly through facial expressions is essential, as they have lost their cognitive and verbal communication skills, having difficulties in communicating important messages, such as pain, discomfort associated with treatments and other complications. For example, not being able to signal negative emotions such as anger, sadness and fear can reduce the quality of interpersonal communication (Grondhuis et al., 2021) and even put their own physical integrity at risk, as they are not able to signal risky situations.

As technology has become a strong ally to assist in the treatment of elderly patients with dementia or not, helping them to have a better quality of life (Dornelles & Corrêa, 2020), there is still a need to develop intelligent systems that support the recognition of emotions in this audience. However, it is worth noticing that studies in the literature are fully dedicated to the recognition of emotions in young people and adults, and few studies address the recognition of emotions in the elderly such as stated by Ma et al. (2019). The authors also explain that aging causes many changes in the shape and appearance of the face, so it is important that systems are developed for the automatic recognition of emotions specifically for this audience. However, it is worth mentioning that there are few datasets built to recognize emotions in the elderly. Furthermore, this lack of data is even greater in relation to the recognition of emotions in elderly people with dementias, such as AD.

1.3 Related Works

In the field of facial expressions, Grondhuis et al. (2021) presents possible answers about why the recognition of emotions in elderly people is more difficult than in young people. Initially, he suggests two hypotheses, relating this difficulty either to (i) excess wrinkles or (ii) atrophy of facial muscles, very common in this age group. For the comparative effect between these two variables (wrinkles/folds vs facial muscles), Generative Adversarial Networks were used in the image treatment. Young individuals have been aged, as well as older people have been artificially rejuvenated. With a database of 28 people, they found that emotions on younger faces when artificially wrinkled are 16% less likely to be identified. However, emotions on the faces of naturally elderly people were 50% less likely to be right. In contrast to this, older faces, artificially rejuvenated, had 74% less chance of getting right compared to young natural ones. The results then suggest that wrinkles and marks on the face of individuals have a much smaller impact than facial musculature on their ability to express emotions.

The research carried out by Micillo et al. (2021) makes an interesting analysis of issues related to the perception and expression of emotions in both elderly and young individuals. The author states that even though the visual recognition of emotions is widely used in the literature, elderly people rarely participate in research. As a way to investigate the interpretation of neutral, happy and angry emotions, 55 elderly people aged between 60 and 85 years and 52 young people aged between 20 and 30 years were tested. At first, the authors did not see a clear relationship for age variations in the perception of correct emotions, contrary to previous studies that said that older adults tend to have greater variability in responses. However, a detail came to light regarding the response time of the 2 groups. Both age groups have a longer response time when subjected to stimuli from emotional faces than to neutral faces. Images of younger people also influenced time and were classified more quickly by the elderly. Images of elderly people did not differ in terms of classification time.

The study carried out by Kuruvayil and Palaniswamy (2021) seeks to investigate the challenges regarding the extraction of facial features. Some of these challenges are partial occlusion, pose and lighting variations. Therefore, the authors propose a modeling of an emotion recognition system based on machine learning and deep learning that aims to generalize well in natural obstructions. The training was carried out using a large volume of data. One challenge was the scarcity of images with the desired characteristics in the basic emotion recognition datasets. The proposed system, ERMOPI (Emotion Recognition using Meta-learning through Occlusion, Pose and Illumination), was trained for 5 basic emotions with facial images having 5 occlusion categories, 7 head poses and 5 lighting levels. The results were 90% accuracy for CMU Multi-PIE images (dataset) and 68% accuracy for AffectNet images (dataset).

The work developed by Zhang, Tjondronegoro, and Chandran (2014) presents a pertinent question, which is: "Does arousal have a greater correlation with a categorized emotion than valence?". From this, the authors propose a framework to evaluate the performance of different texture features merged with geometric distance features to represent facial expressions in a continuous valence-excitation space. Texture features include discriminative subsets of three texture descriptors: local binary patterns - LBP, scale invariant trace transformation - SIFT and Gabor filter outputs that showed state-of-the-art facial expression recognition performance. Correlations between the dimensions of emotion and categorized emotions are investigated from four aspects: spatial distribution, change, similarity (Bhattacharyya distance) and correlation between predicted outcomes and corresponding fundamental truths.

Following an approach similar to the one proposed in this work, Mehendale (2020) work uses convolutional neural networks to recognize emotions in faces. Through an architecture divided into 2 parts, the first network removes the background from the images and a second one extracts features from the facial expressions. The background removal step, as well as the division into different steps of the method, helped to deal with various face position and camera distance issues. For training, testing and validation of the model 3 datasets as well as the author's own images were widely used. The Caltech faces, The CMU database and NIST database datasets achieved accuracy of 85%, 78% and 96% respectively. Finally, for the author, the use of larger datasets, such as the CMU with 750.000 images, showed low precision due to overfitting. Very small datasets, likewise, did not perform well, reaching an ideal number of 2000 to 10000 images for the best performance of the model.

2 Materials and Methods

2.1 Proposal

The methodological path adopted in this research consists of two parts, as shown in Figure 5. Initially, in part 1, the pre-processing of the FER2013 database was carried out. It is important to highlight that the referred database already has a basic processing, as mentioned in subsection 4.1, where the authors resized the images to 48x48, so that the face was positioned centrally in the image. Also, all images were already in grayscale. For our work, it was necessary to transform the pixel list into an array, using numpy. After that, the pixels were on a scale from 0 to 255, for better processing by the neural network, we normalized the pixels to a scale from 0 to 1. After pre-processing, we chose Convolutional Neural Networks (CNN) to be the classifier of this research. The choice was due to the fact that this model is the most used in the literature to work with images. To decide which parameters and configurations to use in CNN, five experiments were carried out with different architectures, as presented in subsection 4.2.1. To train the models, we split the data into 80% for training, 10% for testing, and 10% for validation. It is important to clarify at the end of each epoch the validation of the model in real time was performed. While the test set was used after the model was fully trained. To identify and select the best architecture, we analyzed performance based on evaluative metrics (subsection 4.2.2) such as Accuracy, Precision, Recall, F1 Score and Zero One Loss.

After selecting the best architecture for us to use, we ran part 2 of the methodology. In this step, we randomly chose 19 images of elderly people by Google Images, from the Creative Commons License filter. After obtaining the test images, the Haarcascade Frontal Face (subsection 4.3) was applied to detect the face and, subsequently, the CNN classifier to attribute emotions to the images. It is important to clarify that the experiments were performed on Google Colab, using Python programming language, along with several libraries, such as Keras and TensorFlow. In addition, the experiments ran on a server with adequate capacity (Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz 2.10GHz, RAM: 128GB and 64-bit OS).

2.2 Datasets

The database used to perform the experiments was the Facial Expression Recognition 2013 (FER-2013), created by Pierre Luc Carrier and Aaron Courville and introduced in the ICML 2013 - Challenges in Representation Learning, by Goodfellow et al. (2013). Briefly, the data in this database was created from searches using the Google Images API. Then, it was searched for faces that matched a set of 184 words related to emotions, such as Happy, Raging, Fear and others. These words were also combined with gender, age or ethnicity, at the end of this process, 600 strings were obtained to use as a query for facial image research.

The returned images were kept for the processing stage, where OpenCV was used to add bounding boxes around each face in the collected images. After that, the images were resized to 48x48 pixels and converted to scaled in gray. Altogether, the base comprises a total of 35.887 images and comprises seven classes of emotions. In

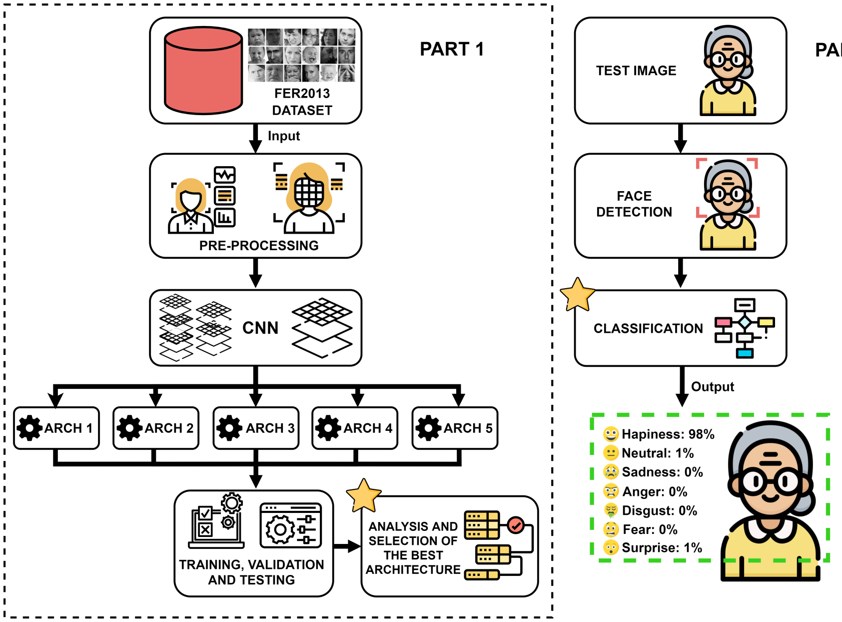

Figure 5 Diagram with the steps of this work. In part 1, the pre-processing of the FER2013 base was carried out. Subsequently, we chose the convolutional neural networks to be the classifier and tested five architectures. Tests with the architectures went through training, validation and testing. Based on the metrics analysis, we choose the best architecture to use in the classifier. In part 2, we chose 19 images of elderly people for testing, we used the frontal face haarcascade to detect the face and the CNN to classify emotions.

the Figure 6 it is possible to see the emotion classes included in the dataset, as well as the number of images per class.

It is important to highlight that FER2013 is currently considered the largest facial expression database publicly available for researchers who wish to train machine learning models, mainly Deep Neural Networks (DNNs). In addition to what has already been said, it is worth noting that the winners of the FER2013 challenge obtained 71.2% accuracy in the test set, using Convolutional Neural Networks and Support Vector Machine (SVM).

2.3 Classification

In this approach we tested different CNN architectures to perform classification. Convolutional Neural Networks are deep learning intelligent methods. In this method, the network learns the patterns of a given input from its processing through several layers. The more layers that are inserted, the deeper the network becomes. In the context of image processing, an input image is subjected to successive convolution and pooling processes. At the end of the network there is a data synthesis layer (fully connected), resulting in the output data.

Figure 7 illustrates this network. Each input image is initially subjected to convolution operations. This mathematical operation works like a filtering process,

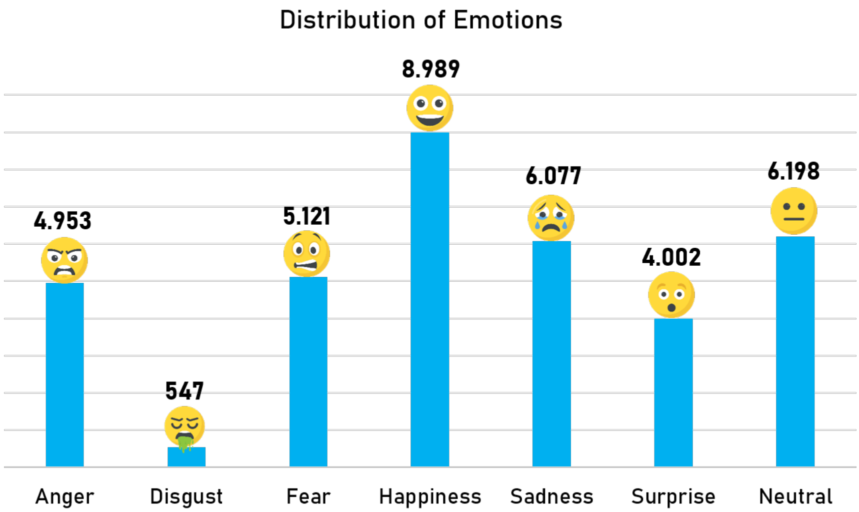

Figure 6 Distribution of Emotions in the FER2013 database.

highlighting or fading information of interest. The quantity and orientation of the filters are defined from the neighborhood considered for the convolution. After each convolution layer, the resulting images are subjected to pooling layers. These layers performs the downsampling process which consists of reducing the image size (e.g. four pixels are replaced by just one). This process is essential since it is responsible for reducing memory consumption during the execution of the algorithm. It is particularly helpful in cases where the number of layers in the network is very large.

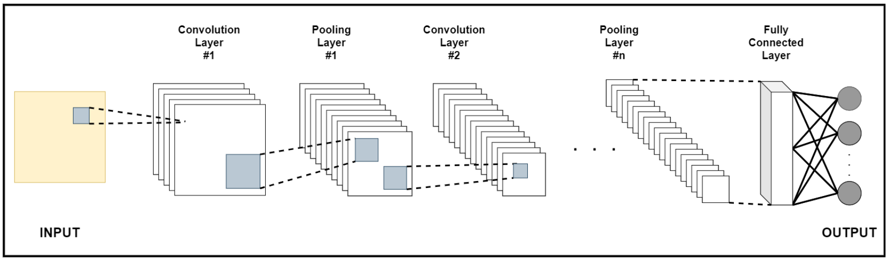

Figure 7 General Convolutional Neural Networks Architecture.

2.3.1 Tested Architectures

Table 1 shows the different configurations that were changed in the architecture experiments. Changes refer to the number of 2D Convolutional layers, Batch Size, Number of Epochs, Dropout and Activation function. The Conv2D convolution layer is responsible for extracting the input features, the number of features was different in the five architectures, but the kernel size of 3x3 was applied to all. The Batch Size has also been modified in all architectures, this is an important hyperparameter and refers to the number of examples used in an iteration, that is, it controls the number of training samples to be worked on. Another parameter that differs in architectures is the number of Epochs, which refers to the number of complete passes through the training dataset. In order to avoid overfitting, we tested different values for the Dropout function, which is responsible for turning off some neurons along with their connections randomly during training, however, during the predictions the neurons are kept active. The activation function determines the output of each neuron, in the experiments we tested three types: Relu, Elu and Softmax. Relu and Elu were used in the Conv2D layers, while Softmax was used only in the last dense layer of CNN, in order to return and assign probabilities to the output ratings.

Table 1 Table with the parameters of the convolution layers (Conv2D), Batch Size, Epochs, Dropout and Activation function that were modified and tested in CNN architectures. It is worth clarifying that each model had a typical single execution, where the expected error had a variation around 10%.

	Tested configurations				
Architectures	Conv2D	Batch Size	Epochs	Dropout	Activation Function
Architecture 1	64, 64, 128, 128, 256, 256, 512, 512	64	100	0.5	Relu and Softmax
Architecture 2	32, 32, 64, 64, 128, 128, 256, 256	16	100	0.2	Elu and Softmax
Architecture 3	64, 64, 128, 128, 256, 256, 512, 512	30	50	0.3	Relu and Softmax
Architecture 4	64, 64, 128, 128, 256, 256, 512, 512	32	70	0.6	Elu and Softmax
Architecture 5	20, 30, 40, 50, 60, 70, 80, 90	256	100	0.2	Relu and Softmax

Some configurations were performed on all architectures. As an example, at the end of two 2D convolution layers both Batch Normalization and Pooling were applied using the Max Pooling technique, with a Pool Size of 2x2 and Strides of 2x2. Pooling layer was applied to reduce the size of input data (feature map) and regularize the network, thus contributing to reduce memory cost and improve processing. We also use a Flatten layer to transform the resulting matrix from the other convoluted layers into a linear array, with a single dimension (1D). In addition, all experiments had four dense layers, where the last dense layer with Softmax activation function had 7 outputs corresponding to emotions classes of Anger, Disgust, Fear, Happiness, Sadness, Surprise and Neutral.

2.3.2 Analysis and Selection of the Best Architecture

To evaluate the performance of the tested architectures, we considered five evaluative metrics, they are: Accuracy, Precision, Recall, F1 Score and Zero One Loss. We can understand accuracy as the most used metric to evaluate algorithm performance. This metric indicates the overall performance of the model, in other words, it is the probability of true positives and true negatives among all results (Barbosa, Gomes, de Santana, de Lima, et al., 2021; Barbosa, Gomes, de Santana, Jeniffer, et al., 2021; Commowick et al., 2018; de Lima, da Silva-Filho, & dos Santos, 2014, 2016; de Souza, dos Santos Lucas e Silva, dos Santos, & de Lima, 2021; Gomes et al., 2020, 2021; Macedo, Santana, dos Santos, Menezes, & Bastos-Filho, 2021; Pereira et al., 2021). Precision, also known as positive predictive value, can be understood as a metric that indicates the ratio of positive predictions that are actually positive (Barbosa, Gomes, de Santana, de Lima, et al., 2021; Barbosa, Gomes, de Santana, Jeniffer, et al., 2021; Commowick et al., 2018; de Lima et al., 2014, 2016; de Souza et al., 2021; Gomes et al., 2020, 2021; Macedo et al., 2021; Pereira et al., 2021). Generally, its value varies, approaching 1 when values are positive and 0 when values tend to be false positive (Arpaci, Huang, Al-Emran, Al-Kabi, & Peng, 2021). Recall (also known as sensitivity) is the rate of true positives and seeks to assess the ability of the algorithm to successfully detect results classified as positive. F1 Score (or F-Measure) can be understood as a harmonic average between accuracy and Recall/Sensitivity. In other words, it's a metric that allows us to view Precision and Recall together (Chicco & Jurman, 2020; Lipton, Elkan, & Narayanaswamy, 2014). Zero One Loss is a common loss function that is widely used in classification tasks. Its operation is simple, basically, this measure attributes a cost (loss) to the failure to guess the correct class. In some situations, different types of misclassification have different costs associated with them (Domingos & Pazzani, 1997). Generally, the value assigned by this zero one loss metric is 0 for a correct classification and 1 for an incorrect classification ("Encyclopedia of Machine Learning", 2010). Having knowledge of the metrics used, the Table 2 presents the results obtained.

Table 2 Performance of tested architectures evaluated based on metrics, such as Accuracy, Precision, Recall, F1 Score and Zero One Loss.

Metric	Architectures				
	Architecture 1	Architecture 2	Architecture 3	Architecture 4	Architecture 5
Accuracy	0.6297	0.6375	0.6090	0.2454	0.5784
Precision	0.6301	0.6428	0.6090	0.2454	0.5784
Recall	0.6297	0.6374	0.5078	0.1428	0.5426
F1 Score	0.6234	0.6337	0.5989	0.0967	0.5756
Zero One Loss	0.3702	0.3624	0.3909	0.7545	0.4215

Among the architectures tested, the one that best stood out in terms of accuracy (0.6375), precision (0.6428), Recall (0.6374), F1 Score (0.6337) and Zero One Loss (0.3624) was architecture 2, followed by architecture 1. On the other hand, as it is noticeable, the worst performance was obtained by architecture 4, with outliers regarding accuracy (0.2454), Precision (0.2454), Recall (0.1428), F1 Score (0.0967) and Zero one Loss (0.7545), followed by architecture 5. The architecture 3 presented mediated values, as it is possible to visualize, the accuracy (0.6090), Precision (0.6090), Recall (0.5078), F1 Score (0.5989) and Zero One Loss (0.3909) are concentrated among the best and worst results obtained. Finally, based on the analysis, the architecture selected to carry out the classifications was architecture 2.

2.4 Face Detection

For the face detection step, we use the Haarcascade Frontal Face, introduced by Viola and Jones (2001). Haar cascade is a method of detecting objects in images or videos, based on machine learning (Rudinskaya & Paringer, 2020). By proposing this method, Viola and Jones (2001) highlight three main contributions. The first one of these is a new form of image representation, called an integral image that allows the detector to quickly calculate the resources used. The second is an AdaBoost based algorithm, which selects a small number of features from a larger data set to produce more efficient classifiers. The third is to combine the classifiers in a cascade fashion to allow background regions of the image to be quickly discarded, while the detector focuses attention on promising regions of the image.

Verma and P Renukadevi (2021) explain that the Haar cascade works on the basis of positive images (covers the class of objects you want to detect) and negative (contains images that the classifier does not recognize). That is, Haar characteristics are weighted according to their success in accepting positive samples and rejecting negative samples (Varley, Cristina, Bonnici, & Camilleri, 2021). Khairuddin, Shahbudin, and Kassim (2021) states that this method already contains pre-trained classifiers and that they are available in the .xml format file to perform different types of recognition, such as detecting eyebrows, smile, nose, face, mouth and others (Anand, 2021; Verma & P Renukadevi, 2021). Finally, Haar Cascade is one of the most successful cascading techniques in this context (Varley et al., 2021), standing out for being fast and achieving good detection rates (Kasinski & Schmidt, 2010; Viola & Jones, 2001).

3 Results

The Figure 8 presents the confusion matrix regarding the recognition of emotions in facial expressions using CNN, generated from the prediction of the test dataset (corresponding to 3.589 images). The X (horizontal) axis of the matrix represents predicted emotions and the Y (vertical) axis represents the correct classification of actual emotions.

In a second moment, in order to verify the effectiveness of the classification model in the researched context, tests were performed to recognize emotions in static images of the elderly. As already mentioned, these images were obtained from Google, using the Creative Commons Licenses type of license. It's possible to view the results in the Figure 9. Along with the classification of emotions, the probability of the emotion belonging to other classes of emotions is also exposed.

The images in Figure 10 evidence an important finding that deserves to be discussed. In some tested images of elderly people two faces were detected and there was difficulty for the classifier to identify emotions. We believe that what may have influenced this failure is that the data used to train the classifier as well as the face detector are not specific to the context of elderly people.

4 Discussion

As can be seen on Figure 8, the model did not perform well in the ratings of Anger, Disgust and Fear emotions. While the emotion Anger is mostly correctly classified (with 267), but still, the model misclassified 231 images, where a significant part was confused with Fear (72), Sad (68) and Neutral (55). With regard to the Disgust emotion, it is important to highlight that it has a low number of representation in the database images, both for training, as well as for validation and testing, this justifies the small number for this class in the matrix of confusion, where out of 52 images, the model was able to correctly classify 23 and confused a significant number of

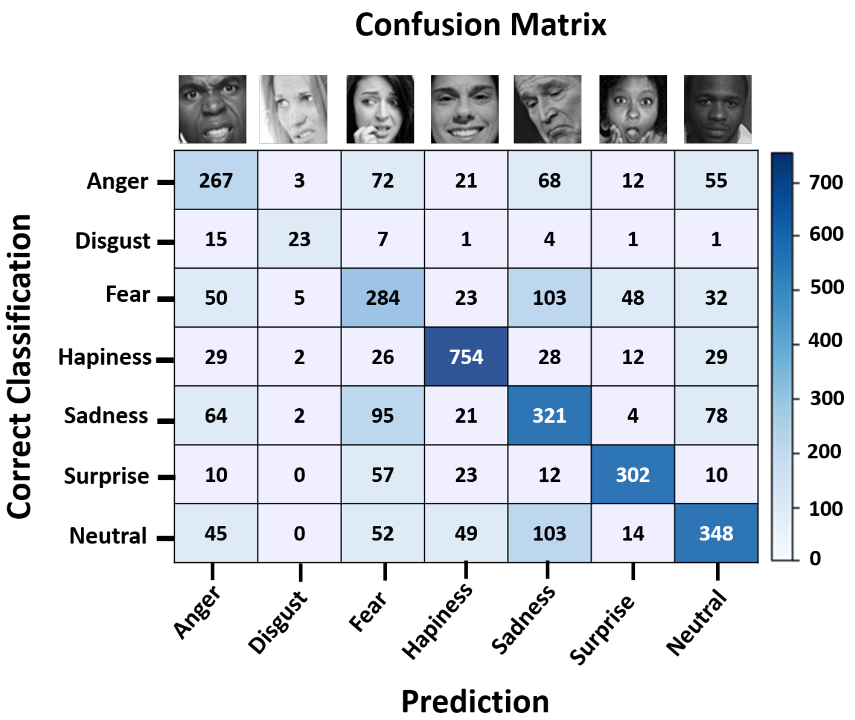

Figure 8 Results presented through the Confusion Matrix. The X (horizontal) axis of the matrix represents predicted emotions and the Y (vertical) axis represents the correct classification of actual emotions.

15 images with rabies. As for the Fear emotion, although the classification was better than the two aforementioned emotions (anger and disgust), out of 545 images, 284 of them were correctly classified, while a considerably high number of 103 was classified as Sad emotion.

On the other hand, the model managed to perform better in the classification of emotions as Happy, Sad, Surprised and Neutral. For Happy class, 754 images were correctly classified and 126 of them were wrongly distributed in the other classes. For sadness, 321 images were correctly classified, however, a high number of 264 images were attributed to other classes of emotions. For the Surprised class, 302 images were correctly classified, but if you look closely, for example, the model tends to classify images of Surprise as Fear (with 57 images). Finally, the Neutral class has 348 correctly classified images and 263 misassigned to other emotion classes.

This result was possibly influenced by the unbalancement of the database, as some classes have much less images than others. As exposed in the confusion matrix (Figure 8), the classification performance was worse for the disgust, anger and fear classes, exactly the classes that have a smaller amount of images.

The results obtained by the test image ratings reinforce what is exposed in the confusion matrix. Where, for example, the Neutral emotion also tends to be classified/confused by the model with the sad emotion. While the Happy emotion has

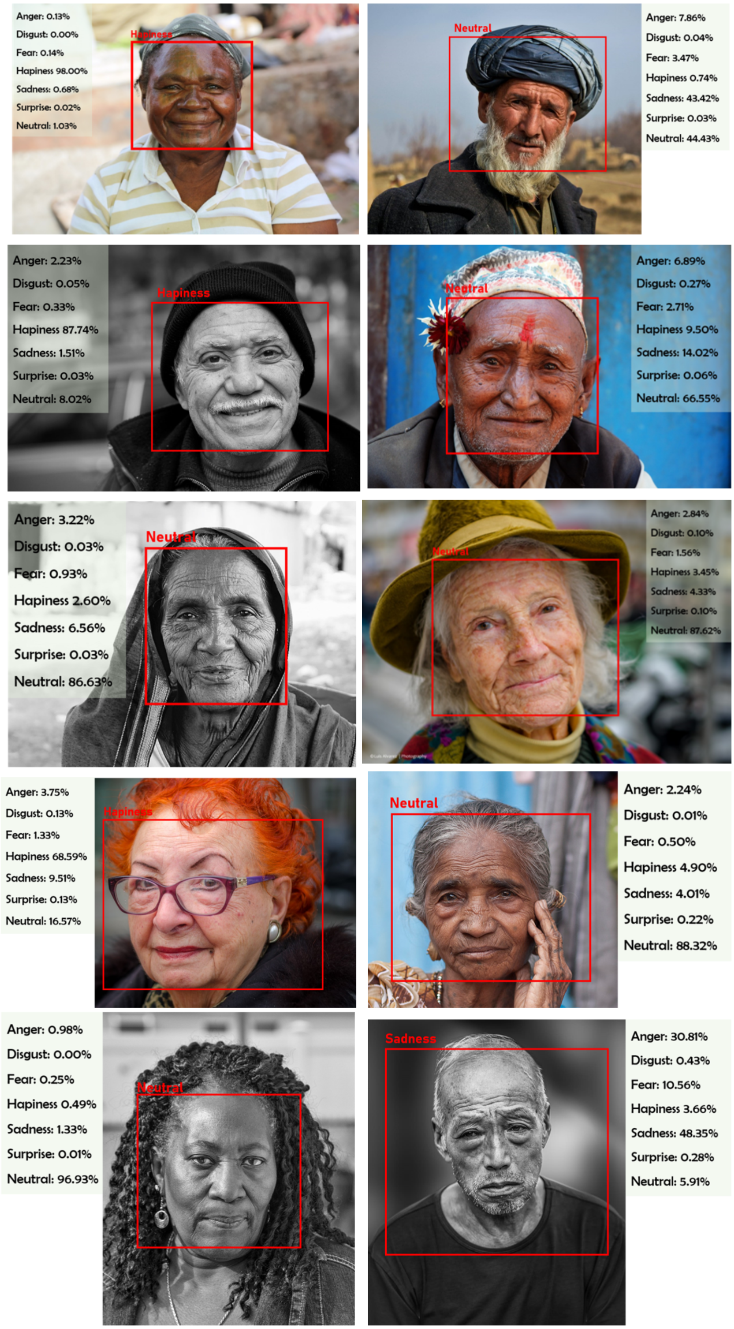

Figure 9 Result of recognition of emotions in static images of elderly people. In all, of the 19 test images we used, the model was able to detect the face and attribute emotions to 52.63% of them.

Figure 10 Face detection errors and attribution of emotion by the classifier in elderly images. In all, of the 19 images we used for testing, the model had difficulties in detecting the face and correctly attributing emotions to 47.37% of them.

chances to belong to the Neutral emotion. These facts can be better observed in the probabilities presented along with each image in Figure 9.

Although good results have been achieved, we believe that errors can be even greater when testing the classifier in the real context of the elderly. Thus, considering that this audience has a compromise in brain and facial structures that are important for the emotions processing and expression, it is believed that the classifier may present more difficulties in assertiveness to recognize emotions, as well as the detector to identify the face. Therefore, we emphasize the importance of having another data source, such as voice, electroencephalographic signals, galvanic skin response, and others, and also using databases specific to the researched context.

5 Conclusion

Studies in the literature that address the recognition of emotions in the elderly, with or without dementia, still represent a small portion close to the problem and need for this support. As a result, there is still a scarcity of specific datasets for this context, especially those that take into account changes in form, face, changes in non-verbal behavior patterns and others. The scarcity of works also directly impacts the limited variety of publicly available databases. This also affects the quality of available databases, which, in most cases, are heavily unbalanced.

A point that is worth noting is that the results obtained in the initial experiments are considerably good. However, the emotion recognition classifier had difficulties to attribute emotions. The face detector classifier also found difficult to identify the faces of some of the elderly. It is believed that this problem was caused by the datasets we used, which are not specific for the elderly context.

However, being aware of these difficulties and limitations reported here, as future works, we intend to mine public databases in the specific context of elderly people to validate the proposed models in more elderly-specific contexts. The variability of training images can also be improved by inserting new images through the database combination. The acquisition of images of the elderly should also be encouraged, whether they are healthy or affected by pathologies, such as dementia.

We firmly believe that this study, as well as the proof of concept we presented here, has contributed to demonstrate that it is possible to build intelligent systems for recognizing emotions in the facial expressions of elderly people in real time. Therefore, therapists from different specialties, especially art therapists, can have a biofeedback of emotions while operating the therapeutic procedures, being able to adjust or discard approaches based on the patient's emotional feedback.

Acknowledgement

This study was financed in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES-Brazil, Finance Code 001, Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq-Brazil, and by Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, Facepe-Brazil, under the code IBPG-0013-1.03/20.

Conflict of Interest

All authors declare they have no conflicts of interest.

Compliance with ethical standards

This is an observational study. The Research Ethics Committee of the Federal University of Pernambuco has confirmed that no ethical approval is required.

References

Abdullah, S.M.S.A., Ameen, S.Y.A., Sadeeq, M.A., Zeebaree, S. (2021). Multimodal emotion recognition using deep learning. *Journal of Applied Science and Technology Trends*, *2*(02), 52–58.

Agres, K.R., Schaefer, R.S., Volk, A., van Hooren, S., Holzapfel, A., Dalla Bella, S., ... others (2021). Music, computing, and health: A roadmap for the current and future roles of music technology for health care and well-being. *Music & Science*, *4*, 2059204321997709.

Anand, D. (2021). Implementation and analysis of sentimental analysis on facial expression using haar cascade methods. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(2), 2787–2793.

Arpaci, I., Huang, S., Al-Emran, M., Al-Kabi, M.N., Peng, M. (2021). Predicting the covid-19 infection with fourteen clinical features using machine learning classification algorithms. *Multimedia Tools and Applications*, *80*(8), 11943–11957.

Barbosa, V.A.d.F., Gomes, J.C., de Santana, M.A., de Lima, C.L., Calado, R.B., Bertoldo Júnior, C.R., ... others (2021). Covid-19 rapid test by combining a random forest-based web system and blood tests. *Journal of Biomolecular Structure and Dynamics*, *2021*, 1–20.

Barbosa, V.A.d.F., Gomes, J.C., de Santana, M.A., Jeniffer, E.d.A., de Souza, R.G., de Souza, R.E., dos Santos, W.P. (2021). Heg.ia: An intelligent system to support diagnosis of covid-19 based on blood tests. *Research on Biomedical Engineering*, *2021*, 1–18.

Bhattacharya, P., Gupta, R.K., Yang, Y. (2021). Exploring the contextual factors affecting multimodal emotion recognition in videos. *IEEE Transactions on Affective Computing*.

Bomfim, A.J.d.L., Ribeiro, R.A.d.S., Chagas, M.H.N. (2019). Recognition of dynamic and static facial expressions of emotion among older adults with major depression. *Trends in psychiatry and psychotherapy*, *41*, 159–166.

Cameirão, M.S., Pereira, F., i Badia, S.B. (2017). Virtual reality with customized positive stimuli in a cognitive-motor rehabilitation task. *2017 international conference on virtual rehabilitation (icvr)* (pp. 1–7).

Cavalli, E., Battaglia, G., Basile, M.S., Bruno, V., Petralia, M.C., Lombardo, S.D., ... others (2020). Exploratory analysis of ipscs-derived neuronal cells as predictors of diagnosis and treatment of alzheimer disease. *Brain sciences*, *10*(3), 166.

Chaturvedi, V., Kaur, A.B., Varshney, V., Garg, A., Chhabra, G.S., Kumar, M. (2021). Music mood and human emotion recognition based on physiological signals: a systematic review. *Multimedia Systems*, 1–24.

Chicco, D., & Jurman, G. (2020). The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC genomics*, *21*(1), 1–13.

Chuang, Y.-C., Chiu, M.-J., Chen, T.-F., Chang, Y.-L., Lai, Y.-M., Cheng, T.-W., Hua, M.-S. (2021). An exploration of the own-age effect on facial emotion recognition in normal elderly people and individuals with the preclinical and demented alzheimers disease. *Journal of Alzheimer's Disease*, *2021*(Preprint), 1–11.

Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., ... others (2018). Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. *Scientific reports*, *8*(1), 1–17.

Crespo, A.B., Idrovo, G.G., Rodrigues, N., Pereira, A. (2016). A virtual reality uav simulation with body area networks to promote the elders life quality. *2016 1st international conference on technology and innovation in sports, health and wellbeing (tishw)* (pp. 1–7).

Dantcheva, A., Bilinski, P., Nguyen, H.T., Broutart, J.-C., Bremond, F. (2017). Expression recognition for severely demented patients in music reminiscence-therapy. *2017 25th european signal processing conference (eusipco)* (pp. 783–787).

de Lima, S.M., da Silva-Filho, A.G., dos Santos, W.P. (2014). A methodology for classification of lesions in mammographies using zernike moments, ELM and SVM neural networks in a multi-kernel approach. *2014 ieee international conference on systems, man, and cybernetics (smc)* (pp. 988–991).

de Lima, S.M., da Silva-Filho, A.G., dos Santos, W.P. (2016). Detection and classification of masses in mammographic images in a multi-kernel approach. *Computer methods and programs in biomedicine*, *134*, 11–29.

de Oliveira, E., & Jaques, P.A. (2013). Classificação de emoções básicas através de imagens capturadas por webcam. *Revista Brasileira de Computação Aplicada*, *5*(2), 40–54.

de Santana, M.A., Gomes, J.C., de Souza, G.M., Suarez, A., Torcate, A.S., Fonseca, F.S., ... dos Santos, W.P. (2020). Reconhecimento automático de emoçoes a partir de sinais multimodais e inteligência artificial. *Anais do IV Simpósio de Inovação em Engenharia Biomédica-SABIO 2020*, 43.

de Sousa, A.L., Costa, S.W.S., Pires, Y., Araújo, F. (2020). Fourface: Uma ferramenta de reconhecimento de expressões faciais. *Brazilian Journal of Development*, *6*(10), 81667–81675.

de Sousa Silva, B., dos Santos Zanetti, I., de Barros, P.P., de Souza, L.T., Barreto, L.B. (2021). Diagnósticos diferenciais das deficiências cognitivas em idosos. *Revista Eletrônica Acervo Científico*, *26*, e7565–e7565.

de Souza, R.G., dos Santos Lucas e Silva, G., dos Santos, W.P., de Lima, M.E. (2021). Computer-aided diagnosis of Alzheimers disease by MRI analysis and evolutionary computing. *Research on Biomedical Engineering*, *37*(3), 455–483.

Domingos, P., & Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning*, *29*(2), 103–130.

Dorneles, S.O.D.S.O., Barbosa, D.N.F., Barbosa, J.L.V. (2020). Sensibilidade ao contexto na identificação de estados afetivos aplicados à educação: um mapeamento sistemático. *RENOTE*, *18*(1).

Dornelles, V.J., & Corrêa, E.A. (2020). Atividade de lazer e novas tecnologias em pessoas com doença de alzheimer. *Revista MotriSaúde*, *2*(1).

Ekman, P., Friesen, W.V., Hager, J.C. (2002). Facial acts coding system: The manual. *Salt lake.* Research Nexus.

Encyclopedia of machine learning. (2010). In C. Sammut & G.I. Webb (Eds.), (pp. 1031–1031). Boston, MA: Springer US.

Ferreira, A., Teixeira, P.M., Tavares, P. (2013). Estudos de expressões faciais para animação 3d. *Edição IPCA ISBN: 978-989-97567-6-2 Design editorial· Cláudio Ferreira.*

Ferreira, C.D., & Torro-Alves, N. (2016). Reconhecimento de emoções faciais no envelhecimento: uma revisão sistemática. *Universitas Psychologica*, *15*(5).

Fölster, M., Hess, U., Werheid, K. (2014). Facial age affects emotional expression decoding. *Frontiers in psychology*, *5*, 30.

Gomes, J.C., Barbosa, V.A.d.F., Santana, M.A., Bandeira, J., Valença, M.J.S., de Souza, R.E., ... dos Santos, W.P. (2020). Ikonos: An intelligent tool to support diagnosis of covid-19 by texture analysis of x-ray images. *Research on Biomedical Engineering*, *2020*, 1–14.

Gomes, J.C., Masood, A.I., Silva, L.H.d.S., da Cruz Ferreira, J.R.B., Júnior, A.A.F., dos Santos Rocha, A.L., ... dos Santos, W.P. (2021). Covid-19 diagnosis by combining rt-pcr and pseudo-convolutional machines to characterize virus sequences. *Scientific Reports*, *11*(1), 1–28.

Gong, S., Dai, Y., Ji, J., Wang, J., Sun, H. (2015). Emotion analysis of telephone complaints from customer based on affective computing. *Computational intelligence and neuroscience*, *2015*.

González, E.J.S., & McMullen, K. (2020). The design of an algorithmic modal music platform for eliciting and detecting emotion. *2020 8th international winter conference on brain-computer interface (bci)* (pp. 1–3).

Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., ... others (2013). Challenges in representation learning: A report on three machine learning contests. *International conference on neural information processing* (pp. 117–124).

Greene, S., Thapliyal, H., Caban-Holt, A. (2016). A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health. *IEEE Consumer Electronics Magazine*, *5*(4), 44–56.

Grondhuis, S.N., Jimmy, A., Teague, C., Brunet, N.M. (2021). Having difficulties reading the facial expression of older individuals? blame it on the facial muscles, not the wrinkles. *Frontiers in Psychology*, *12*.

Guo, T., Zhang, D., Zeng, Y., Huang, T.Y., Xu, H., Zhao, Y. (2020). Molecular and cellular mechanisms underlying the pathogenesis of alzheimers disease. *Molecular neurodegeneration*, *15*(1), 1–37.

Han, J., Zhang, Z., Cummins, N., Schuller, B. (2019). Adversarial training in affective computing and sentiment analysis: Recent advances and perspectives [review article]. *IEEE Computational Intelligence Magazine*, *14*(2), 68-81.

10.1109/MCI.2019.2901088

Han, J., Zhang, Z., Pantic, M., Schuller, B. (2021). Internet of emotional people: Towards continual affective computing cross cultures via audiovisual signals. *Future Generation Computer Systems*, *114*, 294–306.

Hayashi, C.L., Vasconcellos, T.H.F., de Oliveira, D.V., do Nascimento Júnior, J.R.A., Franco, M.F., Nogueira, J.K.A. (2021). Treinamento de dupla tarefa associado a estratégias cognitivas de associação e imagem mental: impacto no equilíbrio, cognição e na saúde mental de idosos. *Research, Society and Development*, *10*(10), e449101018675–e449101018675.

Iancu, I., & Iancu, B. (2017). Elderly in the digital era. theoretical perspectives on assistive technologies. *Technologies*, *5*(3), 60.

Jaques, P.A., Nunes, M.A.S., Isotani, S., Bittencourt, I. (2012). Computaçao afetiva aplicada a educaçao: Dotando sistemas tutores inteligentes de habilidades sociais. *Anais do workshop de desafios da computação aplicada à educação* (pp. 50–59).

Kasinski, A., & Schmidt, A. (2010). The architecture and performance of the face and eyes detection system based on the haar cascade classifiers. *Pattern Analysis and Applications*, *13*(2), 197–211.

Khairuddin, M., Shahbudin, S., Kassim, M. (2021). A smart building security system with intelligent face detection and recognition. *Iop conference series: Materials science and engineering* (Vol. 1176, p. 012030).

Khateeb, M., Anwar, S.M., Alnowami, M. (2021). Multi-domain feature fusion for emotion classification using deap dataset. *IEEE Access*, *9*, 12134–12142.

Ko, H., Kim, K., Bae, M., Seo, M.-G., Nam, G., Park, S., … Lee, J.-Y. (2021). Changes in facial recognition and facial expressions with age. *Sensors*.

Kuruvayil, S., & Palaniswamy, S. (2021). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *Journal of King Saud University-Computer and Information Sciences*.

Ladislau, R., Guimarães, J.G., Souza, W.C.d. (2015). Percepção de expressões faciais emocionais em idosos com doença de alzheimer. *Psicologia: Reflexão e Crítica*, *28*, 804–812.

Leão, L.P., Bezerra, J.S., Matos, L.N., Nunes, M.A.S.N. (2012). Detecção de expressões faciais: uma abordagem baseada em análise do fluxo óptico. *REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS*, *2*(5), 472–489.

Lipton, Z.C., Elkan, C., Narayanaswamy, B. (2014). Thresholding classifiers to maximize f1 score. *ArXiv*, 1402–1892.

Ma, K., Wang, X., Yang, X., Zhang, M., Girard, J.M., Morency, L.-P. (2019). Elder-react: a multimodal dataset for recognizing emotional response in aging adults. *2019 international conference on multimodal interaction* (pp. 349–357).

Macedo, M., Santana, M., dos Santos, W.P., Menezes, R., Bastos-Filho, C. (2021). Breast cancer diagnosis using thermal image analysis: A data-driven approach based on swarm intelligence and supervised learning for optimized feature selection. *Applied Soft Computing*, *2021*, 107533.

Madeira, M.A.M. (2011). *Reconhecimento de emoções faciais na esquizofrenia* (Unpublished doctoral dissertation). Universidade da Beira Interior.

Marosi-Holczberger, E., Prieto-Corona, D.M.B., Yáñez-Téllez, M.G., Rodríguez-Camacho, M.A., Rodríguez-Camacho, H., Guerrero-Juárez, V. (2012). Quantitative spectral eeg assessments during affective states evoked by the presentation of the international affective pictures. *Journal of behavior, health & social issues (México)*, *4*(2), 23–35.

Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (ferc). *SN Applied Sciences*, *2*.

Meska, M.H.G., Mano, L.Y., Silva, J.P., Pereira, G.A., Mazzo, A. (2020). Reconocimiento de emociones para ambiente clínico simulado con uso de olores desagradables: estudio cuasiexperimental. *Revista Latino-Americana de*

Enfermagem, *28*.

Micillo, L., Stablum, F., Mioni, G. (2021). Do the young and the old perceive emotional intervals differently when shown on a younger or older face? *Cognitive Processing*, 1–9.

Nalepa, G.J., Kutt, K., Bobek, S. (2019). Mobile platform for affective context-aware systems. *Future Generation Computer Systems*, *92*, 490–503.

Namba, S., Matsui, H., Zloteanu, M. (2021). Distinct temporal features of genuine and deliberate facial expressions of surprise. *Scientific reports*, *11*(1), 1–10.

Nozima, A.M.M., Demos, B., Souza, W.C.d. (2018). Ausência de prejuízo no reconhecimento de expressões faciais entre indivíduos com parkinson. *Psicologia: Teoria e Pesquisa*, *34*.

Ochi, R., & Midorikawa, A. (2021). Decline in emotional face recognition among elderly people may reflect mild cognitive impairment. *Frontiers in Psychology*, *12*.

Okada, H.K.R. (2018). *Detecção automática de fases temporais de emoção em vídeos a partir de características da face* (Unpublished doctoral dissertation). Universidade Federal do Amazonas.

Paxiuba, C.M., & Lima, C.P. (2020). An experimental methodological approach working emotions and learning using facial expressions recognition. *Brazilian Journal of Computers in Education*, *28*, 92–114.

Pereira, J., Santana, M.A., Gomes, J.C., de Freitas Barbosa, V.A., Valença, M.J.S., de Lima, S.M.L., dos Santos, W.P. (2021). Feature selection based on dialectics to support breast cancer diagnosis using thermographic images. *Research on Biomedical Engineering*, *37*(3), 485–506.

Picard, R.W. (2000). *Affective computing*. MIT press.

Reale, M., Gonzales-Portillo, I., Borlongan, C.V. (2020). Saliva, an easily accessible fluid as diagnostic tool and potent stem cell source for alzheimers disease: Present and future applications. *Brain research*, *1727*, 146535.

Ren, F., & Quan, C. (2012). Linguistic-based emotion analysis and recognition for measuring consumer satisfaction: an application of affective computing. *Information Technology and Management*, *13*(4), 321–332.

Rudinskaya, E., & Paringer, R. (2020). Face detection accuracy study based on race and gender factor using haar cascades. *Information Technology and Nanotechnology*, *2667*.

Rudnicka, E., Napierala, P., Podfigurna, A., Meczekalski, B., Smolarczyk, R., Grymowicz, M. (2020). The world health organization (who) approach to healthy ageing. *Maturitas*, *139*, 6–11.

Russell, J.A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.

Setiono, D., Saputra, D., Putra, K., Moniaga, J.V., Chowanda, A. (2021). Enhancing player experience in game with affective computing. *Procedia Computer Science*, *179*, 781–788.

Sousa, A.L.d., Costa, S.W.d.S., Pires, Y.P., Araújo, F. (2016). Reconhecimento de expressões faciais e emocionais como método avaliativo de aplicações computacionais. *Encontro Regional de Computação e Sistemas de Informação-ENCOSIS*.

Tan, C., Šarlija, M., Kasabov, N. (2021). Neurosense: Short-term emotion recognition and understanding based on spiking neural network modelling of spatio-temporal eeg patterns. *Neurocomputing*, *434*, 137–148.

Tavares, R.E., Jesus, M.C.P.d., Machado, D.R., Braga, V.A.S., Tocantins, F.R., Merighi, M.A.B. (2017). Healthy aging from the perspective of the elderly: an integrative review. *Revista Brasileira de Geriatria e Gerontologia*, *20*, 878–889.

Tian, Y.-L., Kanade, T., Cohn, J.F. (2005). Facial expression analysis. *Handbook of face recognition* (pp. 247–275). Springer.

Torcate, A.S., de Santana, M.A., Moreno, G.M.M., Suarez, A., Gomes, J.C., dos Santos, W.P., ... de Souza, G.M. (2020). Intervenções e impactos da musicoterapia no contexto da doença de alzheimer: Uma revisão de literatura sob a perspectiva da computação afetiva. *Anais do IV Simpósio de Inovação em Engenharia Biomédica-SABIO 2020*, 31.

Trnka, R., & Stuchlíková, I. (2011). Anger coping strategies and anger regulation. *Re-constructing emotional spaces: From experience to regulation*, 89–103.

Varley, P.A., Cristina, S., Bonnici, A., Camilleri, K.P. (2021). As plain as the nose on your face? *Visigrapp (4: Visapp)* (pp. 471–479).

Verma, P., & P Renukadevi, H. (2021). Face recognition using open cv. *Annals of the Romanian Society for Cell Biology*, *25*(6), 6272–6281.

Vicencio-Martínez, A.A., Tovar-Corona, B., Garay-Jiménez, L.I. (2019). Emotion recognition system based on electroencephalography. *2019 16th international conference on electrical engineering, computing science and automatic control (cce)* (pp. 1–6).

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 ieee computer society conference on computer vision and pattern recognition. cvpr 2001* (Vol. 1, pp. I–I).

WHO, W.H.O. (2018). *Ageing.* Available in: $https://www.who.int/health-topics/ageingtab=tab_1$.

WHO, W.H.O. (2021). *Ageing and health.* Available in: $https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.$

Zhang, L., Tjondronegoro, D., Chandran, V. (2014). Representation of facial expression categories in continuous arousal–valence space: feature and correlation. *Image and Vision Computing*, *32*(12), 1067–1079.

Zilidis, G., & Zilidou, V. (2018). The use of new technologies addressing social exclusion and improving the quality of life of the elderly. *Interscientific Health Care*, *10*(4).