UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento Di Ingegneria Dell'Informazione

Curriculum: Biongegneria Per Le Neuroscienze

Development and Validation of an Experimental Protocol for Facial Emotion Recognition

Thesis supervisor

Prof. Mattia Veronese

Thesis co-supervisor

Prof. Andrea Loreggia

Candidate

Abdalla El Soudany

Academic Year 2023/2024

Contents

List of Figures

Abstract

Emotions are integral to human behavior, yet they are susceptible to external influences, such as visual stimuli, which can introduce biases and misperceptions of reality. For these reasons, in fields like bio-engineering and neuroscience, understanding emotion dynamics is crucial. However, before investigating how emotions can be manipulated, an accurate measurement of emotional responses is essential. This thesis focuses on Facial Emotion Recognition (FER) as a tool to objectively quantify emotional reactions based on facial expressions. The aim is to develop and validate an experimental protocol for FER. The study is divided into two phases. The first phase involved developing and validating an Online Task to present a visual stimulus and collect quiz responses on the emotions experienced. The dataset of the images used as stimuli was selected from the Open Affective Standardized Image Set (OASIS), using a custom Matlab code to capture the most impactful emotional states. The validation of the Online Task was conducted in two phases. In the reduced version, 51 students (24 female, 24 male, 3 other) were asked to classify their emotions as positive, negative, or neutral after viewing gender-specific stimuli. Accuracy rates were high for both positive (F 92.31%, M 81.30%) and negative emotions (F 83.62%, M 94.23%), while neutral emotions were less accurately detected (F 50%, M 42.70%). While in the complete version, 34 participants (16 female, 17 male, 1 other) categorized emotions into Ekman's seven basic emotions. The results showed strong accuracy for happiness (F 83.58%, M 67.82%), sadness (F 60.87%, M 86.36%), and disgust (F 100%, M 76.67%). Neutral emotions were also detected with moderate accuracy (F 57.58%, M 74.07%), while other emotions like surprise, anger, and fear had lower accuracy rates. In the second phase, an

Offline FER Algorithm was developed using a modified version of the face-api-js library, which analyzed and classified facial expressions into emotional categories by extracting timestamps and probability values for each emotion from the video recordings. The Area Under Curve (AUC) of the emotional signals, which track the probability evolution of each basic emotion over time, was calculated to determine the predominant emotion after each stimulus. The validation of the algorithm revealed significant accuracy for detecting happiness, while other emotions showed more inconsistent results. The findings validate FER as a biomarker for emotional responses, creating a reliable experimental protocol. Future works, focusing on integrating the online task with the FER algorithm, would allow for a real-time emotion monitoring in various applications, such as healthcare, education, privacy, law, and human-computer interaction.

Synopsis

Emotions play a central role in human behavior, however they are susceptible to external factors, such as visual stimuli. These factors can shape how emotions are experienced, creating biases and leading to potential ”misperceptions” of reality. In fields like neuroscience and bio-engineering, investigating and describing these dynamics of emotional manipulation through stimuli is of great interest. However, before studying how emotions are controlled, it is essential to first measure emotional responses accurately. Facial Emotion Recognition (FER) offers a powerful tool for this purpose. Indeed, by analyzing facial expressions it is possible to objectively measure emotional reactions, which are indicators of internal emotional states. Recent developments of machine learning approaches have improved the accuracy of recognizing and classifying emotions, making FER a reliable method for emotion assessment. The aim of this thesis is to develop and validate an experimental protocol for FER. To do this, the study was divided into two different phasis: The first one was about the development and validation of an Online Task to present visual stimuli (images) and collect quiz responses on the emotions experienced after the exposure. The images were carefully selected from the Open Affective Standardized Image Set (OASIS) using a custom Matlab code, which identified the most impactful images based on arousal and valence values for each basic emotion. The Online Task was divided into two different versions. A reduced version in which data was collected from 51 university students (24 female, 24 male, 3 other). Participants were shown gender-specific stimuli (with the “other” group receiving a mix of both) and asked to classify their post-exposure emotions as positive (accuracy: F 92.31%, M 81.30%), negative (accuracy: F 83.62%, M 94.23%), or neutral (accuracy: F 50%, M 42.70%). On the other hand, the in complete version, data was collected from 34 participants (16 female, 17 male, 1 other) with an expansion of the three emotional categories to Ekman's seven basic emotions, with results showing notable accuracy for happiness (accuracy: F 83.58%, M 67.82%), sadness (accuracy: F 60.87%, M 86.36%), disgust (accuracy: F 100%, M 76.67%), and neutral (accuracy: F 57.58%, M 74.07%). while other emotions, such as surprise, anger, and fear, had lower accuracy rates. The second phase was about the development and validation of an Offline FER Algorithm to analyze and classify facial expressions into emotional categories. Using a modified version of the face-api-js library, adapted specifically for this study, timestamps and probability values for each emotion were extracted from each video recording. The Area Under Curve (AUC) of basic emotions signals (the ones representing the evolution over time of the probability of capturing each basic emotion) was used to determine the predominant emotion after each stimulus, which was then used to validate the FER algorithm. Among all the seven basic emotions, happiness showed significant accuracy results, while other emotions were less consistent. By completing these two phases, this thesis establishes FER as a biomarker for quantifying emotional responses, setting the stage for future research on how external stimuli influence emotional perception. Furthermore, this pipeline can be implemented in future as an online application for portable devices, integrating the two phases (task and FER) to provide real-time emotional feedback. Indeed, it presents a robust free access protocol for measuring emotions. The online task was validated effectively, demonstrating its technical reliability. The FER algorithm showed strong results primarily for happiness, although its overall performance may have been influenced by factors such as the image dataset, experimental conditions, and participant demographics.

CHAPTER 1

Introduction

1.1 Background and Motivation

1.1.1 The Science of Emotion

What emotions are and why are they so important in fields like Bio-engineering for Neuroscience? To answer these questions, it is necessary to take a few steps back and investigate the science of emotions over the course of history until today.

In ancient Egypt, emotions were deeply linked to *maat* (the principle of harmony and balance of the universe). Ancient Egyptians used to believe that the heart was not only the physical center of the body but also the core of emotional responses and moral judgment, governing a person's thoughts, feelings, and ethical decisions (see Hornung 1999 [1]). Similarly, in ancient Greece, Plato saw emotions as chaotic and irrational forces that needed to be subdued by reason to preserve personal and social harmony. He believed that unchecked emotions could disrupt rational thinking and lead to disorder. To maintain a well-ordered society, emotions had to be controlled by the rational part of the soul. On the other side, Aristotle had a more nuanced view. He saw emotions not in opposition to reason and inherently disruptive, like Plato's thought, but in alignment with it. He argued that emotions, when guided by rational judgment, helped individuals to respond appropriately to different situations. For Aristotle, ethical behavior involved finding the right balance of emotion, shaped by reason, to act virtuously (see Nussbaum 2001 [2]).

Moving eastward, in ancient China, Confucianism saw emotions as vital for promoting social harmony. Properly managing and expressing emotions was key to moral behavior and maintaining order, emphasizing self-restraint and appropriate emotional responses to strengthen relationships and community stability. In contrast, Taoism and Buddhism had a different perspective on emotions. Both philosophies saw emotions as disruptions or distortions that cloud the true understanding of reality. Taoism emphasized flowing with the natural order of life, suggesting that being overly attached to emotions could prevent one from achieving the Tao (inner "Path" to harmony and balance). Similarly, Buddhism viewed emotions as sources of suffering, and the path to enlightenment required transcending emotional attachments and desires to see the world clearly and without illusions (see Ames & Hall 2003 [3]).

During the Middle Ages, the study of emotions evolved. In medieval Western philosophy, emotions were largely framed by Christian theology. Thinkers like Augustine of Hippo viewed emotions as both gifts and challenges; they could lead to sin if uncontrolled but also played a role in spiritual connection with God when properly directed (see Knuuttila 2004 [4]). Meanwhile, in the Islamic world, philosophers like Ibn Sina took a more scientific approach, integrating ancient Greek thought with Islamic teachings. He saw emotions as physiological responses that could affect mental health and behavior, emphasizing that they needed to be understood and measured to maintain a balanced mind (see Gutas 2001 [5]). This perspective was groundbreaking, as it combined religious, philosophical, and early medical views, seeing emotions not only as moral forces but as biological phenomena that could influence cognition.

The combination of a spiritual and scientific approach to studying emotions was carried forward during the Renaissance by Descartes, the key figure of this period. He introduced a mechanistic view of emotions, which he called "passions of the soul." He argued that emotions were essential to human experience but arose from physiological reactions in the body, which could be understood and controlled by the rational mind. Descartes identified six primary emotions (love, hatred, desire, joy, and sadness) which he believed were foundational to all other emotional experiences (see Descartes 1649 [6]).

In the Enlightenment, the role of emotions in human decision-making and perception became a central concern. Hume, one of this era's most influential thinkers, argued that reason alone is not enough to drive human behavior, while emotions (or "passions") are the true motivators of action. Hume suggested that emotions shape our perceptions of the world and play a fundamental role in our moral judgments. However, they also introduce biases, as emotions can cloud rational thought and distort reality (see Hume 1739 [7]).

In the late 19th century, the science of emotions was officially born, particularly with the groundbreaking work of Charles Darwin, Carl Lange, and William James. Darwin argued that emotions are universal and can be measured through observable expressions, such as facial movements. He suggested that these expressions evolved for survival, serving as non-verbal communication between individuals (see Darwin 1872 [8]). Darwin's work laid the foundation for the idea that emotions are biologically rooted and can be studied scientifically, marking a significant shift from the earlier philosophical discourse to a more empirical approach.

At the same time, Carl Lange, alongside William James, contributed to the development of the James-Lange theory of emotion. This theory proposed that emotions arise from physiological changes in response to external stimuli. Their work highlighted the measurable link between bodily reactions and emotional experience, suggesting that emotions could be quantified through physical responses such as facial expressions, heart rate, and other physiological markers (see James 1884 [9] and Lange 1885 [10]).

Meanwhile, Sigmund Freud introduced a psychological dimension to emotions, emphasizing the deep, often unconscious, influence they may have on human behavior. Freud argued that unexpressed or inadequately managed emotions could lead to neuroses and mental illness. He saw emotions as powerful forces that, if repressed, could distort reality and decision-making, leading to psychological conflict and bias in perception (see Freud 1923 [11]).

In the first half of the 20th century, several pioneers in neuroscience advanced the understanding of emotions through neurological models. Walter Cannon challenged the James-Lange theory by proposing the Cannon-Bard theory, suggesting that emotions and physiological responses occur simultaneously, not sequentially. He argued that the thalamus sends signals to both the cortex (for emotional experience) and the autonomic nervous system (for physiological reaction) at the same time (see Cannon 1927 [12]). Building on this, John Fulton explored the role of the frontal lobes in regulating emotional behavior, helping establish the importance of the brain's prefrontal regions in emotional control (see Fulton 1935 [13]). James Papez expanded these ideas by introducing the "Papez circuit", proposing that a specific neural circuit involving the hypothalamus, thalamus, hippocampus, and cingulate cortex governs emotional experiences (see Papez 1937 [14]). However, at a certain point, all this neuroscientific interest in the study of emotions waned (see LeDoux 2000 [15]).

1.1.2 The Mind-Body Problem

The mind-body dualism, which views the mind and body as distinct entities, has shaped discussions about the relationship between mental processes and physical states in the body (see Descartes 1641 [16]). Historically, this separation framed much of neuroscience's approach to understanding emotions. However, the mid-20th century saw significant changes, particularly with the rise of the cognitive revolution. The mind began to be treated as an information-processing system, akin to software, with the body as the hardware (see Bunge 2010 [17] and Pylyshyn 1984 [18]). This approach effectively bypassed the traditional mind-body problem, as it no longer sought to explain the interaction between mind and body but rather treated cognitive processes as computational operations independent of physical states of the body. Cognitive science shifted focus from emotional processes to more easily quantifiable cognitive functions like perception, memory, and attention, processes that could be conceptualized in computational terms (see LeDoux 2000 [15]).

1.1.3 The Dark Cloud of Emotion Subjectivity

While cognitive science successfully bypassed the complexities of consciousness by focusing on unconscious, objective cognitive processes (see Kihlstrom 1987 [19]), emotion research remained anchored to the subjective experience of feelings. Emotional researchers struggled to explain how conscious emotional experiences arise, without resolving their underlying neural mechanisms (see Damasio 1994 [20]). In addition, at that time, much of the emotion research relied on animal models, as direct experimentation on human subjects was often ethically and practically unfeasible. Using animals, like rodents, allowed scientists to study the neural circuits behind emotional behaviors such as fear and anxiety in controlled settings (see Davis 1992 [21]). However, proving the existence of subjective emotional states in these animals posed significant challenges, since researchers had to infer emotions based on behavioral and physiological responses, without verbal confirmation from the subjects. This led to a degree of uncertainty and hindered the development of emotion theories. As a result, discussions about emotions frequently circled back to the issue of subjective feelings, even though measuring such states in non-human subjects was nearly impossible (LeDoux 2000 [15]).

At the same time, cognitive scientists (see Neisser 1967 [22] and Gardner 1987 [23]) emphasized that their field was not equipped to explore subjective experiences like emotions, which were seen as difficult to quantify. This sidelined emotion research in both psychology and neuroscience, as researchers focused on more objective and measurable cognitive processes. Consequently, the study of emotions became less prominent, contributing to what LeDoux [15] referred to as a "dark cloud of subjectivity" that overshadowed emotional studies and stifled progress in the field during the mid-20th century.

1.1.4 The Limbic System

In the early 1950s, the concept of the limbic system emerged as a seemingly comprehensive solution to the question of how the brain produces emotions. Paul MacLean overcame the "Papez circuit" theory, that proposed the existence of an "emotion circuit" in the brain, and expanded this model by incorporating additional brain regions, such as the amygdala and the septum, into what he termed the "limbic system". MacLean used to believe that the limbic system governed emotional experience and behavior (see MacLean 1949 [24] and 1952 [25]).

The limbic system theory proposes that a group of interconnected brain structures plays a central role in regulating emotions. These structures include the hippocampus, which is involved in memory formation; the hypothalamus, which helps maintain homeostasis and controls emotional responses; the thalamus, which processes sensory information; the amygdala, which is key in processing fear and other emotional reactions; and the cingulate gyrus, which is involved in emotional regulation and decision-making. These regions work together to generate, process, and regulate emotional experiences, contributing to our understanding of emotional behavior and its physiological basis.

This theory quickly gained traction in neuroscience due to its simplicity and appeal, as it offered a unified explanation for how the brain processes emotions. It was thought to provide a solid framework for understanding emotional responses, and for a time, it seemed that the mystery of emotions in the brain had been largely solved. However, the widespread acceptance of the limbic system concept may have unintentionally slowed further investigation into the neural mechanisms of emotions. Researchers began to believe that emotion was now well understood and shifted their attention to more "complex" cognitive processes, such as perception, memory, and decision-making (see LeDoux 1987 [26] and 1991 [27]).

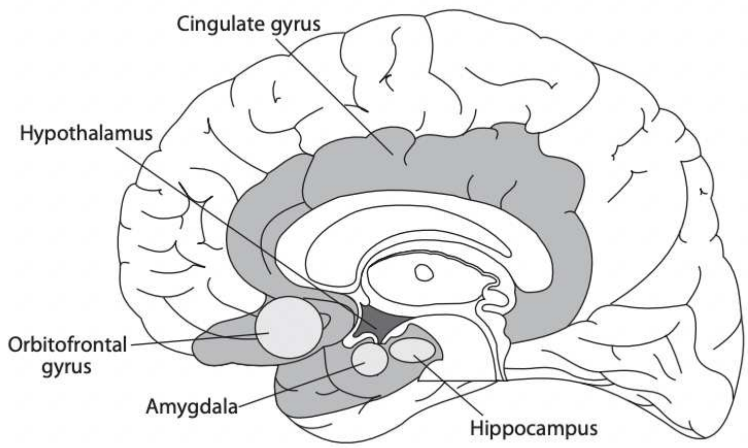

Figure 1.1: Anatomical illustration of key areas of the Limbic System.

1.1.5 Cognitive Revolution for Emotion Studies and Vice Versa

The cognitive revolution, which originally toned down the importance of emotion research, paradoxically offers the tools to resurrect it by providing methods to study emotions independently of subjective feelings. Cognitive science's focus on unconscious processing mechanisms allows for an examination of how the brain handles emotional information without first needing to resolve the origin of conscious emotional experiences (Bargh & Chartrand 1999 [28]). Contrary to the traditional belief that emotions require conscious awareness, research has shown that emotional responses, much like cognitive functions, can occur through unconscious processes (see LeDoux 1996 [30]). This shift in perspective opens new doors for emotion research, enabling scientists to study the neural and behavioral underpinnings of emotional responses regardless of the elusive question of conscious experience and without being constrained by subjective interpretations.

Conversely, emotion research has much to offer to cognitive science. Indeed, a purely cognitive approach that disregards the influence of emotions, motivations, and similar factors presents a limited and unrealistic portrayal of the mind (see Damasio 1994 [20] and Panksepp 1998 [29]). Emotions are not extraneous to cognition; rather, they are intertwined with it. The mind is neither purely cognitive nor purely emotional: it is a combination of both, a computational part shaped by desires, fears, and other affective states (see LeDoux 2000 [15]).

Integrating emotion research into cognitive frameworks would provide a more complete understanding of the mind, helping to move beyond the view of the brain merely as an information-processing system devoid of motivational and emotional content.

Towards the end of the twentieth century, following an integrated approach of emotion and cognition, research on emotions experienced a resurgence. This thesis will adopt a "neuro-engineering" experimental perspective to analyze emotional responses going even further than the traditional concept of the "limbic system" and highlighting the complexity underlying a seemingly simple emotional reaction. Finally, the importance of developing a robust and objective pipeline for measuring emotions will be demonstrated, after giving an engineering definition of emotions (see LeDoux 1996 [30] and Pessoa 2008 [31]).

1.2 Emotional Responses to Visual Stimuli

From a "neuro-engineering" experimental perspective, in which the focus is mainly on the technical aspects behind the cognitive processes (without losing sight of the importance of subsequently integrating this information with a psychological point of view), the whole emotion process can be studied when a visual stimulus (such as an image or a video) is detected by a sensory organ and begins a processing mechanism, triggering a cascade of neurological activities that bring to an emotional response (see LeDoux 1996 [30] and Pessoa 2008 [31]).

The initial step in processing visual stimuli involves sensation, the process by which the visual system detects light and converts it into neural signals. The light entering the eyes is focused by the lens onto the retina, where photoreceptor cells (rods and cones) transduce it into electrical signals (see Purves et al. 2012 [32]). Subsequently, a visual evoked potential and functional MRI techniques have shown that these signals are transmitted via the optic nerve to the lateral geniculate nucleus in the Thalamus (see Kandel et al. 2013 [33]).

Joseph LeDoux's theory of the "emotional pathways" explains how the brain processes emotional responses to stimuli (see LeDoux 1996 [30]). When a visual stimulus, such as an image, is detected, it can follow two distinct neural routes in the brain, leading to different types of responses:

The "low road" is faster and more primitive, where the sensory information bypasses the cortex and goes directly from the thalamus to the amygdala, which triggers the emo-

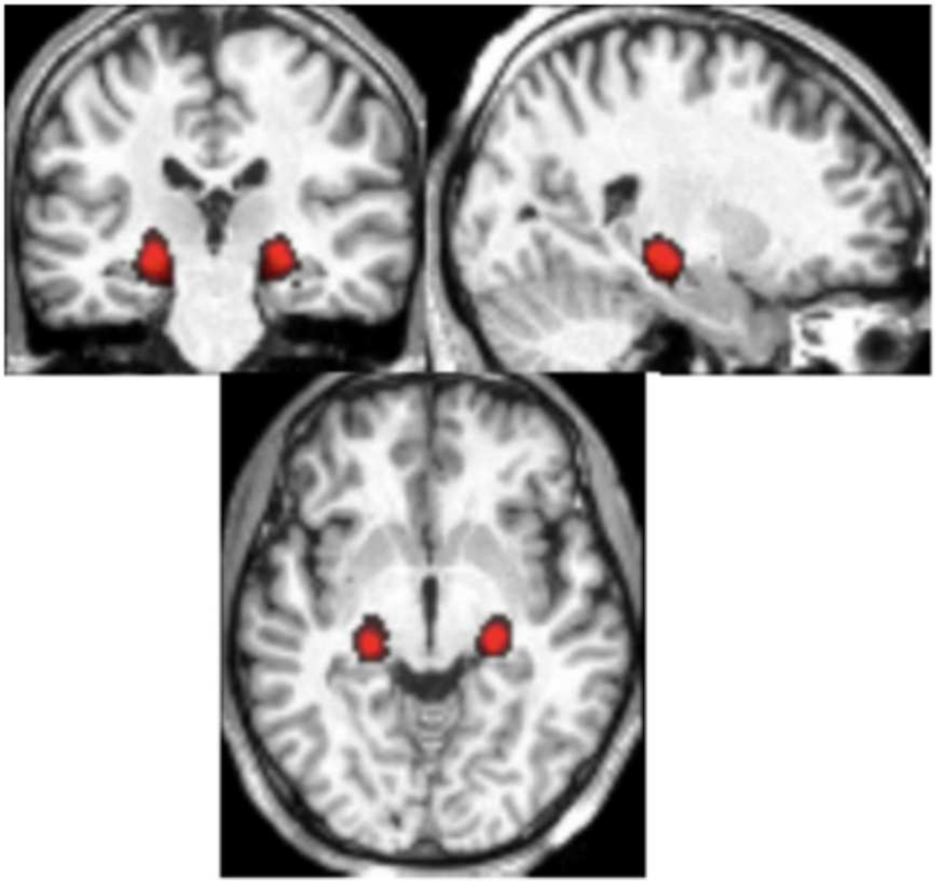

Figure 1.2: Structural template for Lateral Geniculate Nucleus (LGN) in the Thalamus overlaid on the normalized (MNI 152) anatomical image.

tional response. This allows for an immediate, unconscious emotional reaction, such as the ”fight or flight” response, even before undergoing the second step called perception, the process by which a stimulus is being fully processed and understood. This “low road” pathway is essential for survival, as it allows quick reactions to potentially threatening stimuli without the delay of complex cognitive processing.

On the other hand, the ”high road” is slower and more refined. In this route, the sensory information is first processed in the sensory cortex before being relayed to the amygdala. This allows for more detailed analysis and a conscious understanding of the stimulus (aka perception step) before having an emotional response.

The theory of Joseph LeDoux’s two emotional processing pathways, the ”high road” (slow, conscious processing) and ”low road” (fast, unconscious processing), has been supported by various studies (most of them concentrated on fear emotion study) using neuroimaging techniques such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG). Bechara et al. (see

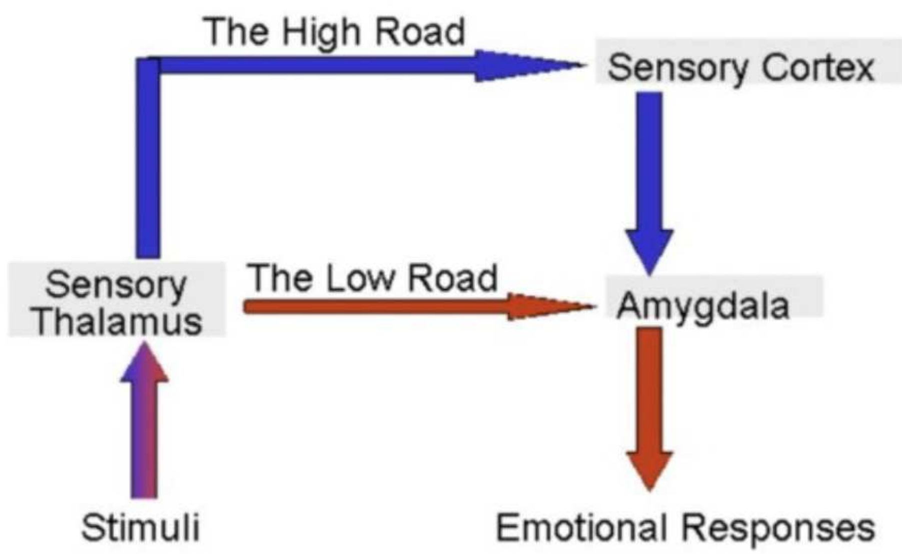

Figure 1.3: Schematic representation of Joseph LeDoux's theory of the two emotional processing pathways.

Bechara et al. 1995 [35]) used fMRI to demonstrate how rapid emotional responses (low road) are mediated by the direct route from the thalamus to the amygdala, which allows for immediate reactions to stimuli like fear. In contrast, the high road involves the cortical regions, which process the stimulus more slowly and consciously before sending signals back to the amygdala to modulate the response. Moreover, studies analyzing the temporal dynamics of emotional processing using EEG and MEG show that low road responses can occur as quickly as 15-30 milliseconds, particularly for fear-related stimuli (see Ohman 2001 [36]). The high road, involving more complex cognitive evaluation, typically takes 250-500 milliseconds or more, depending on the complexity of the stimulus (see Pessoa 2010 [37]). Through neuro-engineering and computational modeling, researchers have simulated the temporal dynamics and pathways involved in emotional processing. Marois and Ivanoff (see Marois 2005 [38]) used neural network simulations to confirm that direct connections between the thalamus and amygdala (low road) are significantly faster in generating an emotional response, while the thalamus - sensory cortex - amygdala route (high road) allows for more detailed analysis of the stimulus. These models have helped quantify the temporal characteristics of each pathway and confirm the results from imaging analysis.

1.3 Biases and Misperceptions in Emotion Recognition

Bias in emotional response affects how individuals perceive and react emotionally to a stimulus. These biases arise when the cognitive processing of stimuli leads to systematic deviations from an objective interpretation of reality (see Pessoa 2008 [31] and Vuilleumier 2005 [39]). For instance, an individual subjected to a fearful visual stimulus may react more intensely, due to pre-existing biases, leading to an exaggerated emotional reaction that distorts the actual threat level (see Barrett & Kensinger 2010 [40]). Pre-existing biases in emotional responses can be attributed to several factors. These factors mainly depend on individual experiences and personality traits. Biases are often the result of past emotional conditioning, where repeated exposure to specific stimuli elicits a learned emotional response over time (see De Houwer & Hermans 2010 [41]). In fact, it has been demonstrated that an individual who has experienced a past trauma may develop a bias toward interpreting neutral stimuli as threatening, leading to exaggerated fear responses due to their past experiences (see Gross 2014 [42]).

On the other hand, some personality traits, such as neuroticism, have also been linked to emotional biases. Indeed, it has been proven that individuals with higher levels of neuroticism are more prone to perceive stimuli (even the neutral ones) negatively. This is due to their predisposition to experience anxiety and emotional instability (see Kensinger 2007 [43]).

Besides individual experiences and personality traits, biases in emotional response depend also on cultural and environmental factors, as societal norms can shape how individuals react emotionally to similar stimuli (see Pessoa 2008 [31]). Cultural norms dictate which emotions are considered appropriate in a specific situation. For instance, in collectivist cultures, such as many East Asian societies, emotional restraint is valued, leading individuals to suppress emotions like anger to maintain social harmony, creating a bias in interpreting stimuli (see Mesquita & Walker 2003 [44]). Conversely, individualistic cultures, such as those found in the United States or Western Europe, may encourage more open emotional expression, resulting in stronger emotional reactions (see Kitayama et al. 2000 [45]).

Environmental factors, such as socioeconomic and educational conditions, may also contribute to emotional response biases. For example, individuals exposed to chronic stress or trauma, such as those in impoverished environments, may develop a heightened sensitivity to negative stimuli. This can result in a "negativity bias," where threatening or stressful stimuli are perceived as more significant or immediate than they objectively are, increasing the likelihood of suffering from anxiety disorders (see Norris et al. 2011 [46]).

Additionally, biases can emerge from cognitive processes, such as selective attention, where emotionally charged stimuli are given priority over neutral ones during the perception process. This means that emotionally significant input, particularly stimuli linked to human survival (such as fear or danger), is more likely to be captured by an individual's focus. As a result, due to this mechanism of "emotional filtering," an emotionally biased interpretation may override a more objective assessment of a situation (see Vuilleumier 2005 [39]).

1.4 Importance of Measuring Emotional Responses

Emotional response biases, as explained in the previous paragraph, can distort how a stimulus is perceived and shape an individual's decision process in ways that deviate from objective reality (Pessoa 2008 [31]). In fields like marketing and politics, emotional biases are strategically exploited to shape consumer behavior and public opinion.

Emotional manipulation in marketing involves using positive or negative emotional triggers to influence buying decisions. For example, brands may design advertising specifically to evoke feelings of nostalgia or happiness aimed at connecting consumers emotionally to their products, increasing brand loyalty and impulsive purchases, and thus increasing profits (see Lerner et al. 2015 [47]). Moreover, other brands may use fear emotion, such as emphasizing limited-edition or limited-time offers, to lead to irrational decision-making, leveraging consumer anxiety about losing the offer (see Murray et al. 2018 [48]).

In politics, the classic example concerns how fear and anger are often used to manipulate voters. Politicians exploit emotional biases by amplifying concerns regarding national security, immigration, economic stability, or health. This tactic biases voters toward candidates who promise safety and stability, regardless of the more nuanced policy considerations (Brader 2006 [49] and Iyengar & Westwood 2015 [50]).

Another unsettling aspect of political manipulation is emotional polarization, a process by which politicians, during political campaigns, amplify the divides between groups. In this way, they ensure partisan loyalty and reduce the voter's capacity for critical thinking and handling open rational dialogues. This phenomenon tends to deepen group identities and create an "us vs. them" mentality, making it hard to consider alternative perspectives (see Lodge & Taber 2013 [51]).

Moreover, the advent of social media and algorithmic targeting has emphasized these mechanisms by enabling political campaigns to direct emotionally charged content specifically to individuals predisposed to emotional triggers. This not only reinforces pre-existing biases but also deepens the polarization, as people are forced to submit to similar content chosen by the social media algorithm (see Bail et al. 2018 [52]).

These emotional manipulation strategies are particularly effective because they target automatic unconscious cognitive processes, bypassing the rational deliberative processes that typically guide decision-making (see Lodge & Taber 2013 [51]). Indeed, it's been demonstrated how moral and political judgments are frequently based on "intuitions" (fast and unconscious emotional reactions) rather than rational and conscious thoughts (see Haidt 2001 [53]).

To control emotional biases and prevent its dangerous unconscious manipulation, it's crucial to quantify how stimuli elicit emotional responses. The emotion quantification allows uncovering how biases shape behavior and identify the manipulation tactics (see Lerner et al. 2015 [47]).

All these considerations lead immediately to a key question: is it possible to measure the emotional response? Over time, several researchers have asked this question and the answer depends on the type of the emotions expressed to be measured.

1.4.1 Traditional Methods to Measure Emotions

In psychology, the most common traditional method for measuring emotional responses is through verbal expression. This approach consists of directly asking participants to verbally describe the emotions they are experiencing after exposure to a stimulus. The responses are acquired using structured or unstructured questionnaires and interviews. The most widely used structured questionnaire is the Positive and Negative Affect Schedule (PANAS) (see Watson & Tellegen 1985 [54]). This format uses predefined fixed-order questions that allow participants to rate their emotional states on a Likert scale (see Likert 1932 [55]). This evaluation of theirs is used to assess two dimensions of the emotional experience: positive affect (such as joy or excitement) and negative affect (such as sadness or anger).

In addition to structured questionnaires, interviews are a valid alternative to measure verbally an emotional response. Verbal interviews, which can be closed or open-ended, allow for a more detailed exploration of emotions (see Diener et al. 1985 [56]). However, open-ended interviews that offer richer data suffer from being less standardized and more difficult to be compared to other open-ended interviews or closed questionnaires (see Patton 2002 [57]).

In general, verbal self-report methods implicitly assume that individuals are completely accurate and honest in reporting their emotions, which opens the door to many biases. The first one regards the inevitable subjective influence of both the interviewer and the participant (see Diener et al. 1985 [56]). Another important bias regards the social pressure, which may push participants to modify their responses to align with societal norms. This happens particularly with emotions deemed negative or undesirable. For instance, in a workplace setting, subjects may under-report feelings of stress, sadness, or frustration to avoid seeming unprofessional (see Bradburn 1969 [58]). Moreover, in contexts where the emotional response is analyzed at a long distance from the stimulus, the Memory Recall Error should be considered. The latter regards the fact that emotions can be fleeting or distorting over time and so individuals may struggle to accurately remember or articulate how they felt at that moment (see Watson & Tellegen 1985 [54]). Another problem to face is the cognitive and linguistic limitation. Indeed, verbal self-report methods often require a developed vocabulary and cognitive ability to accurately express complex emotions (see Barrett et al. 2007 [59]).

To overcome the linguistic and cognitive limitations associated with the verbal self-report methods, the Self-Assessment Manikin (SAM) was developed (see Lang 1980

[60]). SAM is a non-verbal and visual tool that allows participants to rate their emotional experiences using illustrated representations across three core dimensions. Valence measures pleasantness, i.e., how positive or negative an emotion feels, represented with images ranging from a smiling face (high valence) to a sad face (low valence). Arousal measures the intensity of emotional activation, going from a calm and relaxed figure (low arousal) to an excited and explosive figure (high arousal). Lastly, Dominance measures the sense of control in a situation, represented by varying figure sizes, going from large (high dominance, indicating control) to small (low dominance, indicating submission). SAM's reliance on images rather than language makes it particularly effective for individuals with language barriers or cognitive impairments, such as children or non-native speakers (see Bradley & Lang, 1994 [61]).

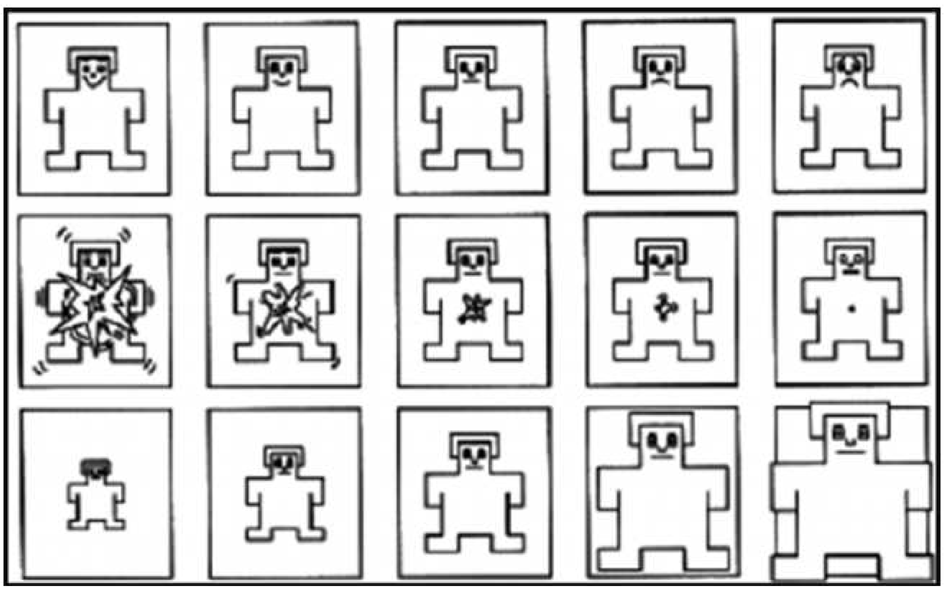

Figure 1.4: Representation of the Self-Assessment Manikin (SAM). From top to bottom, the figure illustrates the levels of Valence, Arousal, and Dominance.

Besides traditional methods for measuring emotional responses through verbal expression, it's possible to observe and measure the physiological expression of emotions as changes in autonomic responses. One of the earliest methods consists of monitoring heart rate and blood pressure and considering their variability as an indicator of emotional arousal (see James 1884 [9] and Cannon 1927 [12]). Another widely used measure is skin conductance, also known as the Galvanic Skin Response (GSR), which monitors changes in sweat gland activity in response to emotional arousal. The monitoring is done by measuring the electrical conductance of the skin, which increases as sweat production grows. This method is particularly useful in stress or anxiety studies (see Jung 1906 [64] and Lykken 1959 [65]). Finally, pupil dilation observation can be used as a biomarker to monitor changes in emotional response after exposure to a stimulus.

Indeed, it has been proven that pupils tend to dilate in response to positive or negative emotional stimuli. This automatic involuntary response can help track an individual's unconscious emotional reactions (see Hess 1965 [66]).

In addition to verbal and physiological expression methods, the behavioral expression of emotions is another critical area in emotional research. These methods are based on observing and interpreting outward signs of emotional states, such as facial expressions, body language, or speech patterns. Regarding facial expression analysis through manual coding, one of the earliest and most robust methods for assessing behavioral emotional responses is the Facial Action Coding System (FACS) (see Ekman & Friesen, 1978 [67]). FACS is a comprehensive framework that identifies and categorizes specific facial muscle movements (known as Action Units), which correspond to various emotions. Each Action Unit represents a distinct facial movement (for example, eyebrow or lips raising), that can be linked to an emotion response (for example, happiness or sadness). The whole system is built on Charles Darwin's theory of universality of facial emotion expressions (Darwin 1872). This manual process is highly detailed and was considered one of the most reliable methods for assessing facial expressions of emotion until the rise of automated facial recognition systems (Cohn & Ekman 2005 [69]). Another key method in behavioral expression research is the observation of body language, such as posture, gestures, and other non-verbal cues. For instance, open body language is often associated with positive emotions, while a closed-off posture may signal uneasiness or stress (Mehrabian 1968 [70]). Finally, also speech tone and rate can provide a valuable index of an individual's emotional state. For instance, raised speech tone or rate can indicate excitement or anxiety (see Scherer 1986 [71]).

1.4.2 Technological Approaches to Measure Emotions

In recent decades, technological advancements have significantly transformed the way emotional responses are measured. The first step is the automation of verbal expression of emotions. While traditionally, verbal reports of emotional states were acquired by structured or unstructured questionnaires and interviews to be subsequently analyzed manually, today these processes are automated by using text analysis algorithms that interpret emotional states from questionnaires or interviews. For instance, sentiment analysis algorithms evaluate text data to give scores to categorize emotional valence (whether the text is positive, negative, or neutral) by analyzing word choice, syntax, and context (see Pang & Lee 2008 [81]). These types of systems are widely used in psychological studies, social media monitoring, and customer feedback analysis with the great advantage of offering scalability and objectivity compared to traditional manual coding (see Liu 2012 [82]).

In addition, the automation of measurement techniques of physiological expressions of emotions, which were once manually recorded and analyzed, is now done with digital sensors and computer algorithms, that improve accuracy, scalability, and real-time analysis leading to more precise data interpretation (see McDuff et al. 2016 [72]). For instance, heart rate (see McDuff et al. 2016 [72]), blood pressure (see Healey & Picard 2005 [73]), and galvanic skin response (see Picard & Healey 1997 [74] and Setz et al. 2010 [75]) monitoring, which was traditionally done using medical instruments to assess emotional arousal, is now automated through wearable devices like smartwatches and fitness trackers. These devices allow the continuous monitoring of emotional arousal with greater sensitivity and precision. Moreover, technological tools, such as eye-tracking systems, have also improved the measurement of pupil dilation, used as a biomarker for emotional response (see Bradley et al. 2008 [76]).

When it comes to behavioral expression, technological approaches have also automated the traditional observation methods. For instance, a computer vision system, equipped with sensors and cameras, can capture three-dimensional body movements and process them through algorithms to classify the emotional state from the body posture (see Karg et al. 2013 [77]). These systems can be used both in research fields, such as social robotics and human-computer interaction (to detect and respond to human emotions in real-time) (see Zeng et al. 2009 [78]), or in clinic areas, where this technology can monitor patient body language to assess some emotional states (e.g., stress levels) without intrusive questioning or manual observation (Salah et al. 2011 [79]). Finally, also the analysis of speech patterns and tone can be automated. Indeed, modern machine learning algorithms can analyze thousands of vocal parameters, such as pitch variation and speech rate, to classify emotional states with a high degree of accuracy (see Cowie et al. 2001 [80]).

Although the automation of emotional expression analysis, explored in this section, is advancing significantly and allowing for more precise and scalable measurement, one of the most impactful innovations is Facial Emotion Recognition.

1.5 Facial Emotion Recognition Systems

Facial Emotion Recognition (FER) systems offer several key advantages over other emotion recognition technologies. First, FER systems are non-invasive and based only on visual data (e.g., from a camera) to assess emotions. This makes them much more accessible and user-friendly compared to methods like skin conductance or heart rate measurement, which often require the use of cumbersome wearable sensors (see Calvo & D'Mello, 2010 [84]). Second, facial expressions are one of the most direct and observable indicators of emotional states, making them a very powerful tool for studying emotion (see Ekman 1993 [83]). Indeed, unlike verbal reports, facial expressions are often unconscious and involuntary, offering a more authentic emotional reaction to be measured in real-time (see Zeng et al. 2009 [78] and Calvo & D'Mello 2010 [84]). Third, facial expressions are universally recognized across cultures. Indeed, FER systems can be applied globally without the limitations of language or cultural barriers (Matsumoto & Hwang 2011 [85]).

In general, FER systems tend to have higher accuracy if compared to other emotion recognition technologies, thanks to FER's ability to capture spontaneous emotional responses directly from facial expressions (see Mandryk & Inkpen 2004 [86]). For instance, in commercial applications, some software like FaceReader has demonstrated high accuracy rates (89%) for static facial expressions, especially when analyzing well-defined emotions such as happiness or sadness (see Lewinski et al. 2014 [87] and Stöckli et al. 2017 [88]). Besides, one of the most important key advantages is the detection of micro-expressions. Research has shown that FER can capture complex and subtle emotional cues, which are often unconscious or involuntary (see Mollahosseini et al. 2017 [89]). Thanks to all the features explained above, facial emotion recognition has critical applications in fields like healthcare, where it is increasingly being utilized for monitoring and diagnosing emotional disorders, such as depression, anxiety, or autism spectrum disorders. Indeed, these detailed non-invasive systems can help identify emotional irregularities by analyzing small changes in facial expressions, allowing doctors to track the patient's emotional states (see Ekman 1992 [90] and Calvo & D'Mello 2010 [84]).

FER systems use computer vision and machine learning algorithms to analyze facial expressions and identify the emotional states they represent. This analysis process is based on several key steps: face detection, feature extraction, and emotion classification.

The first step in any FER system is detecting faces in an image or video. Early FER systems relied heavily on Haar cascades or other traditional computer vision techniques to detect faces (see Viola & Jones, 2001 [91]). However, modern systems based on deep learning use convolutional neural networks (CNNs) to detect faces with higher accuracy and efficiency. CNNs are trained on big datasets and are able to recognize face shapes even in dynamic situations (see Shih-Chung Hsu et al. 2017 [92]).

Once a face is detected, the next step is the feature extraction. Relevant features can be the position of key facial landmarks (eyes, eyebrows, nose, mouth) or the global shape of the face (see Khan 2018 [93]). A common approach to feature extraction is landmark-based detection (e.g., detecting the 68 key points of a face). This step captures the subtle movements and shapes that correspond to different emotions. For example, the stretching of the landmarks of the mouth corners is an indicator of emotion responses like surprise or happiness (Ekman & Friesen 1978 [67] and Sariyanidi et al. 2015 [94]).

The final step is emotion classification, which consists of classifying the extracted features into specific emotions. The emotion classification algorithm typically uses supervised learning models like Support Vector Machines (SVMs) (see Bartlett et al. 2006 [95]), Random Forests (see Arriaga et al. 2017 [96]), or deep neural networks (DNNs) (Mollahosseini et al. 2016 [97]). These models are trained on labeled datasets, where each face is annotated with a corresponding emotional label (e.g., happy, sad, angry, etc.) and their goal is to map the facial features to the closest matching emotional category.

1.6 Aim

The aim of this thesis is to develop and validate a robust experimental protocol for Facial Emotion Recognition (FER) that accurately measures emotional responses elicited by visual stimuli. This will allow FER to be established as a biomarker for quantifying emotional responses, facilitating further exploration into the mechanisms of emotional manipulation and misperception of reality. Besides, this free-access protocol can also be used in research and clinical areas (such as those explored previously). This pipeline is very fast and immediate on one side and provides accurate online results on the other.

CHAPTER 2

Methodology

2.1 Overview of the Experimental Design

The development and data collection process were structured into two main phases. The first phase focused on the creation of an online task aimed at acquiring participants' self-reported emotional responses to visual stimuli via a quiz. The quiz responses served as a subjective measure of the participants' perceived emotional states and to test whether the stimuli are mapped correctly within the OASIS dataset.The second phase involved the collection of video recordings of participants while they viewed the same set of stimuli and completed the quiz. The video recordings are meant to be analyzed by an offline Facial Emotion Recognition (FER) algorithm. The quiz responses of the second phase were used as the ground truth for the evaluation of the objective measures of the FER algorithm, providing a baseline for comparison.

2.2 Stimuli Dataset Selection

In order to select appropriate visual stimuli for this study, it has been considered several image datasets commonly used for similar studies, such as International Affective Picture System (IAPS) and the Nencki Affective Picture System (NAPS). IAPS, considered the "gold standard", is one of the most widely used affective image databases in psychology and neuroscience. It was developed by Lang et al. in 1997 [98] to provide a robust tool for studying emotion and attention through a standard emotionally evocative set of pictures. Each image is rated on two primary dimensions: valence (ranging from pleasant to unpleasant) and arousal (ranging from calm to excited). This dataset has been validated across different populations, contributing to its popularity, and allowing for cross-study comparisons (see Lang et al., 1997 [98]). However, despite these strengths, the IAPS suffers from certain limitations that reduce its adoption for this experimental protocol. The first one is about the labels of emotions that are often generalized into basic emotional categories, such as positive, neutral, and negative. This leaves a little room for studies of more specific emotions, such as happiness, sadness, or surprise (see Mikels et al. 2005 [99]). The second limitation is about licensing restrictions (see Lang et al., 2008 [100]). Indeed, IAPS is non open source and therefore it needs institutional access or special permission to be used. This would greatly limit the proposed experimental setting because the latter claims to be free to access for anyone. Besides, NAPS dataset is an open-source alternative to IAPS. It was developed by Marchewka et al. in 2014 [101] to provide a large set of standard images rated on emotional dimensions like those of IAPS (valence, arousal, and dominance). This allows studies to be conducted on a higher quantity of stimuli (see Marchewka et al., 2014 [101]). However, some researches have criticized NAPS for not providing images that trigger intense emotions. Furthermore, NAPS categorizes images according to basic emotions, but it may lack diversity in some "near" emotional states like disgust or fear. This may lead to potential biases in studies where these emotions are crucial (see Balsamo 2020 [102]). Fortunately, the search for the ideal dataset has led to the exploration of a final dataset that allows to overcome some of the above limitations: the Open Affective Standardized Image Set (OASIS).

2.2.1 Open Affective Standardized Image Set (OASIS)

The Open Affective Standardized Image Set (OASIS) dataset is widely used for studies of emotional responses to visual stimuli. It consists of a collection of carefully selected and standardized images, representing a large range of emotional expressions. OASIS dataset images are selected to evoke a variety of basic emotions, including happiness, sadness, anger, fear, disgust, surprise, and neutrality. Each image is rated for valence and arousal, going from 1 to 7 (see Kurdi 2017 [103]). This dataset was chosen for several reasons. First, it is open access, this overcomes the licensing restrictions, allowing researchers from various disciplines, such as psychology, computer science, and engineering, to use the dataset for their experimental protocols (see Redies 2020 [104] and Jović 2022 [105]). Moreover, OASIS has been validated across different populations in various cultural contexts, making it a reliable tool to study the emotional responses on a global scale. Additionally, OASIS is able to overcome the limited emotional range of IAPS (particularly in evoking extreme emotions, such as disgust or fear) (Kurdi et al. 2017 [103]). On the other hand, NAPS, while offering free access and a bigger range of images, it has been criticized anyway for not being able to differentiate between emotions (Marchewka et al. 2014). In contrast, OASIS provides a more balanced emotional range, offering clear and consistent ratings for both arousal and valence. This makes the dataset flexible and easily controllable. Moreover, both IAPS and NAPS contain images that may not always represent a real context in which individuals typically experience a specific emotion (Marchewka et al. 2014 [101]) while OASIS is more realistic. Indeed, the 900 images of the OASIS dataset were collected from various sources, including publicly available image databases, advertisements, and media sources. Subsequently they got curated, standardized, and divided into 4 main categories (animal, objects, persons, and scenes) to ensure a big range of emotional response, providing a high-quality set of images that can be used in emotion research. Furthermore, OASIS reports for each image a mean and standard deviation values of valence and arousal. These values are results of dataset validation research. Another added benefit is that the mean and SD values of valence and arousal are reported for both male and female identifying participants. This allowed us to conduct the experiment dividing by gender and investigating any differences related to gender.

2.3 Development of the Online Task: Structure and Data Collection Procedures

The entire dataset of 900 images from OASIS was evaluated, and a MATLAB code, specifically implemented for this purpose, was used to process, and select the images to be adopted as a visual stimulus. The MATLAB code performed several tasks, including separating the images based on gender (male and female) and identifying specific areas within the valence-arousal space that correspond to distinct emotions.These areas were defined as rectangular regions based on Russell's Circumplex Model of emotion (see Russell 1980 [106]):

- Happiness: [5.5 - 7; 4 - 6.5]
- Sadness: [1 - 3; 2 - 4]
- Surprise: [4 - 6; 4.5 - 7]
- Fear: [1 - 3; 4.4 - 6]
- Anger: [1 - 3; 5.5 - 7]
- Disgust: [1 - 3; 4 - 6]
- Neutrality: [4 - 6; 3 - 5]

For instance, in the case of happiness, valence range is going from 5.5 to 7 and arousal from 4 to 6.5. This will form a rectangle in the upper-right quadrant of the valence-arousal graph, representing the happiness area.

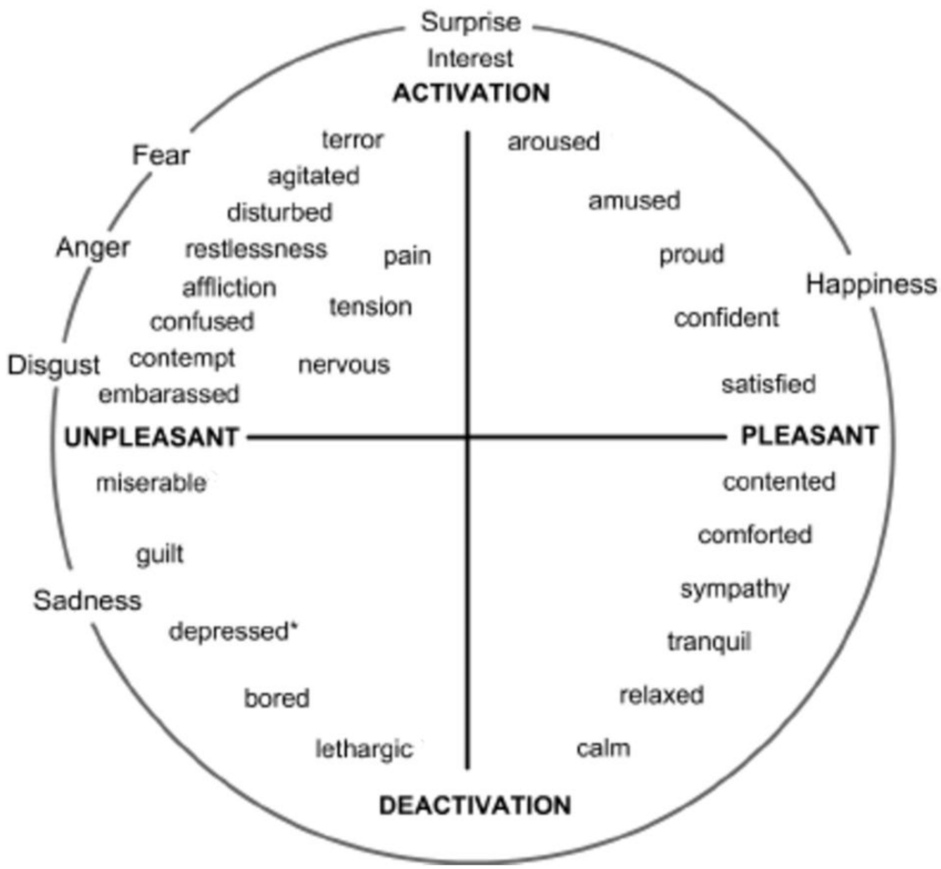

Figure 2.1: Representation of the Circumplex Model of Emotion.

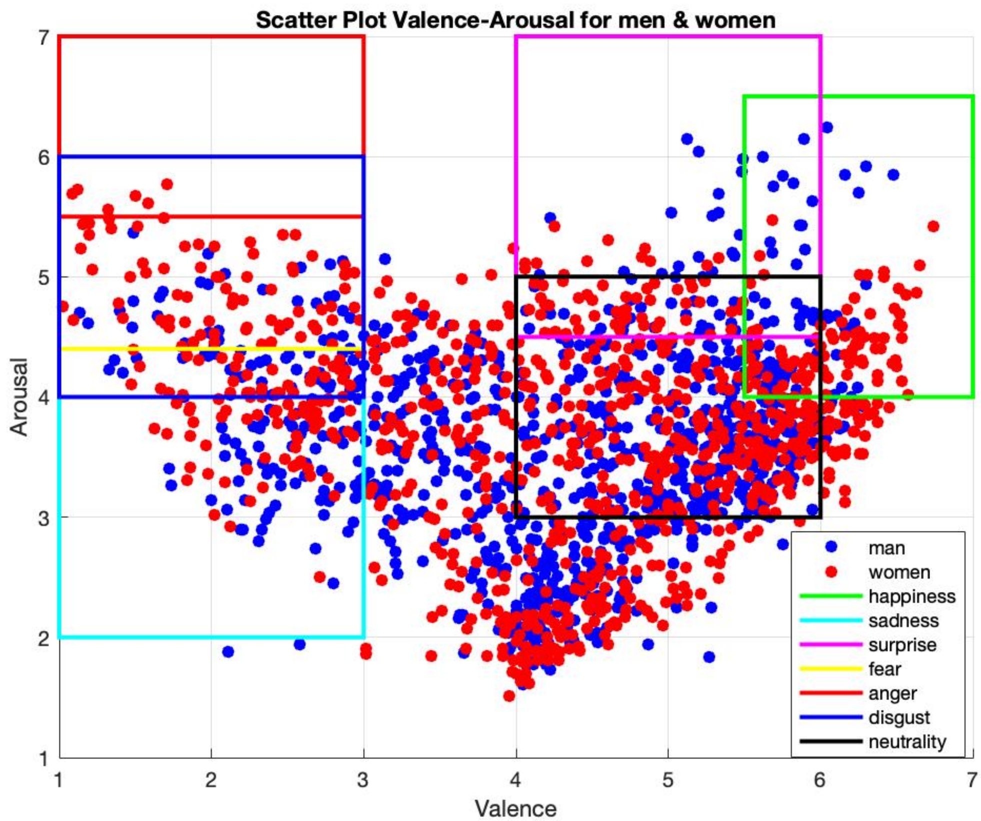

Figure 2.2: Representation of the OASIS dataset underlining the areas corresponding to the basic emotions according to the Circumplex Model of Emotion.

Finally, the MATLAB code then extracted the images within each emotional rectangle and selected those with the most significant valence and arousal values. Specifically, for each base emotion (happiness, sadness, surprise, fear, anger, disgust, and neutrality), two images were selected from each category (animals, objects, people, and scenes) for both male and female groups, ensuring a balanced and representative selection across all emotions and categories. From these 112 stimulus images (56 for each gender) a further selection was made to reduce the number of stimuli to 30 (i.e., 15 for each gender). In more detail, these 15 stimulus images are divided into 5 positive stimuli (i.e., happiness), 5 negative ones (sadness, fear, anger, and disgust) and 5 neutral ones. The stimuli that evoke surprise were excluded a priori because the surprise emotion can be considered both positive and negative and this can create ambiguity in the evaluation (see Sieun et al. 2017 [107]). Once the stimuli selection phase was over, the development of an online task was initiated. The online task aim is to present the stimuli, randomly, to participants and subsequently acquiring, for each image, an answer to a quiz that indicates the emotional state elicit immediately after exposure to the stimulus. It was decided to use Python to program a web page that performs this task via localhost. More specifically, the web page works as follows: first, it asks participants to identify their gender. This determines the package from which the stimuli to be presented are selected (in the case of identification as 'other,' the selection path will include both male and female stimuli). After that, a resting image (a white background with a black + on it) is shown for 3 seconds. Then, a stimulus is presented for 4 seconds (see La Monica et al. 2023 [108]). Finally, the participant is asked to indicate their emotional state after the exposure by answering a simple quiz. This cycle is repeated for each stimulus to be shown. At the end of the cycle, the quiz results are saved in a CSV file located in the folder with the Python code that manages the localhost. To validate this online task, the study was divided into two versions.

2.3.1 Reduced Version

In this simplified version, the goal is to collect data of emotional responses using only three categories: positive, negative, and neutral. The study involved 51 students, 24 of them identifying as female, 24 as male, and 3 as other with ages going from 22 to 28. Specifically, 33 participants were shown 15 images, while 10 images were presented to the remaining 18. This simplified version was implemented to evaluate whether the task is sufficiently valid before proceeding to evaluate a more complex emotional state.

2.3.2 Complete Version

In the complete version, the categories were expanded to include all seven of Ekman's basic emotions, making the analysis more complex. Data were collected from 34 students, 16 of them identifying as female, 17 as male, and 1 as other with ages going from 22 to 28. In this setup, 20 participants were shown 10 images, while 15 images were presented to the remaining 14. This version aimed to capture a more detailed emotional responses, providing a deeper understanding of the participants' emotional processing across a wider spectrum of emotions.

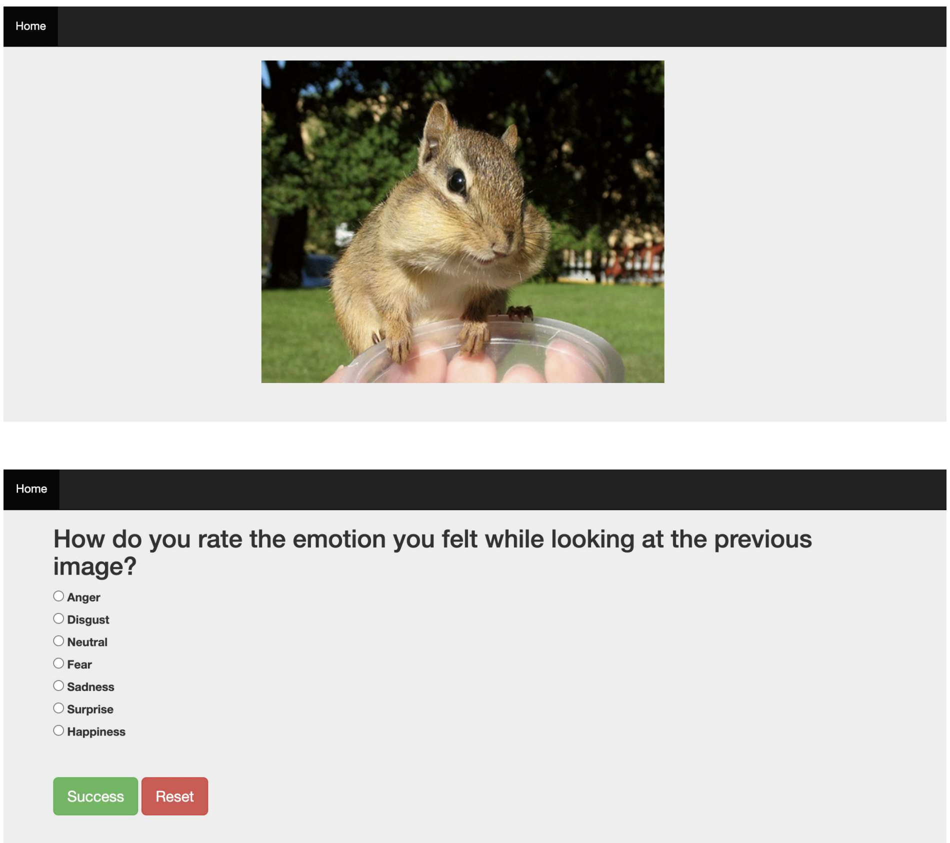

Figure 2.3: An example extracted from the Online Task.

2.3.3 Data Pre-Processing and Statistical Analysis

The data collected in the CSV files from both versions of the task (reduced and complete) were pre-processed via a MATLAB code by excluding quiz responses where the stimulus was not displayed correctly. Following this, the responses for each image were extracted, and histograms were generated to show the frequency of emotions experienced for each image. Finally, confusion matrices were calculated, using the labels provided in the OASIS dataset as the ground truth. This allowed the calculation of accuracy and precision for each emotion selected by the participants.

2.4 Development of the Offline FER Algorithm

After testing the online task, the next step consists of the development of the offline Facial Emotion Recognition (FER) algorithm. In order to select an optimal approach several alternatives were considered.

The first one is The Emotion API (see Verma 2019 [109]), a service given from the Microsoft Cognitive Service Pack, which uses machine learning to detect facial expressions and classify them into Ekman's seven basic emotions (happiness, sadness, fear, anger, disgust, surprise and neutral). The API is built on deep learning models that are trained using large and labeled datasets of facial images (such as FER2013 and AffectNet), that contain various facial expressions with their corresponding emotional categories labeled. The training process involves putting these emotionally labeled images into a convolutional neural network (CNN), which automatically detects key features (including the position and shape of the mouth, eyes, and eyebrows) corresponding to each different emotion. In particular, the model identifies 68 critical points, called facial landmarks, such as eyes (that can change size), mouth (that can change alignment), eyebrows, and chin (that can change position). The CNN can learn these patterns over time through multiple layers of feature extraction. Each layer of the network represents a higher level of facial data, going from basic shapes (such as edges) to complex facial structures. Once the landmarks are detected, the extracted features are passed into a pre-trained deep neural network, which calculates the likelihood that a current facial expression fits into all seven of Ekman's basic emotions. The final output is the emotion with the highest core. The Emotion API has been validated on the same training datasets, FER2013 and AffectNet. During the validation phase, API's predictions were compared with the ground truth labels provided in the dataset, and metrics like accuracy, precision, recall, and F1 score were used to assess the model's performance, which is very promising. However, it does not operate in real-time and this is a big con since this project aspires to provide an online device that works in real-time to continuously monitor the emotional state.

Another option is Affidex (see McDuff 2016 [72]), developed by MIT's Affective Computing Group. This tool is used for recognizing and analyzing emotional expressions through advanced computer vision and machine learning techniques. It works in this way: first of all, as for The Emotion API, it used CNN to detect facial landmarks and extracts both global (overall face shape) and local features (specific areas such as mouth or eyes) from the 68 detected critical points. Also in this case, the system's algorithm then maps these features to emotional states using probability scores, calculated for each emotion. This process allows not only the categorization of static facial features into an emotional state but also registers their dynamic changes over time. Indeed, the strength of Affidex is its ability to work in real-time, this is possible thanks to its low-latency CNN architecture. Additionally, it is open-source and offers an SDK (Software Development Kit) for multiple platforms, including Windows, Java, and JavaScript, which makes it very versatile for many applications. Affidex is trained using a combination of posed and spontaneous emotional expressions taken from large datasets, including FER2013, CK+, and MIT's proprietary datasets. Also here, as for API, all these datasets contain face images with a labeled emotion. This allows the CNN to learn patterns of facial movement associated with each emotion. Subsequently, it was validated using standard datasets, such as CK+ and AffectNet, calculating metrics such as accuracy, precision, recall, and F1 score to evaluate the performance. However, its accuracy and reliability, particularly in subtle emotional recognition low-light conditions, were found to be inconsistent according to validation studies (see Stöckli 2018 [111] and Magdin 2019 [112]).

Given the limitations of these options, the choice fell to Face-api.js (see face-api.js 2014 [113]), a JavaScript-based library that operates entirely in the web browser, which makes it particularly well-suited for real-time applications. Like the previous algorithms, also Face-api.js uses CNN to detect 68 facial landmarks, covering critical facial points. In particular, it uses a special CNN architecture, called MobileNet, which is lightweight and efficient. It's designed on purpose to maximize efficiency and reduce computational complexity while maintaining a high level of accuracy. This is possible thanks to a technique called depthwise separable convolutions, based on the separation of spatial filtering (depthwise convolution) and depth filtering (pointwise convolution), while standard convolution does both simultaneously, enhancing computational cost. Depthwise and pointwise convolution split the traditional convolution process to make it more efficient and it works in this way: for instance, an image is composed of channels, such as the red, green, and blue (RGB) channels in a color image, where each channel holds pixel intensity values for that specific color. The depthwise convolution applies a filter to each channel to detect, for example, edges in the red, green, and blue channels separately. Then the pointwise convolution filter combines the information from all channels, allowing the integration of information across all colors (see Howard et al. 2017 [114]). The separation of depthwise and pointwise convolution reduces drastically the number of operations and parameters involved, allowing for a faster run with minimal loss of performance. This efficiency makes MobileNet particularly suitable for real-time and resource-constrained environments, like web browsers or mobile devices, where computational power and memory capacity are limited. Despite its lightweight nature, MobileNet still provides high power in capturing complex facial features and expressions by identifying 68 key facial landmarks (such as eyes, nose, mouth, and eyebrows) that are crucial for emotion recognition. The latter is possible using models pre-trained on datasets like FER2013 and AffectNet. Although its accuracy might not fully reach more complex offline tools (such as Affidex) anyway it offers a perfect balance between speed and performance. Finally, Face-api.js was validated on the same datasets of the training phase (FER2013 and AffectNet). Although this system may show some limitations in identifying subtle or nuanced emotions (especially in low-light or ambiguous conditions), its overall accuracy is strong for basic emotion recognition in real-time and this is sufficient for many applications, also thanks to its open-source nature.

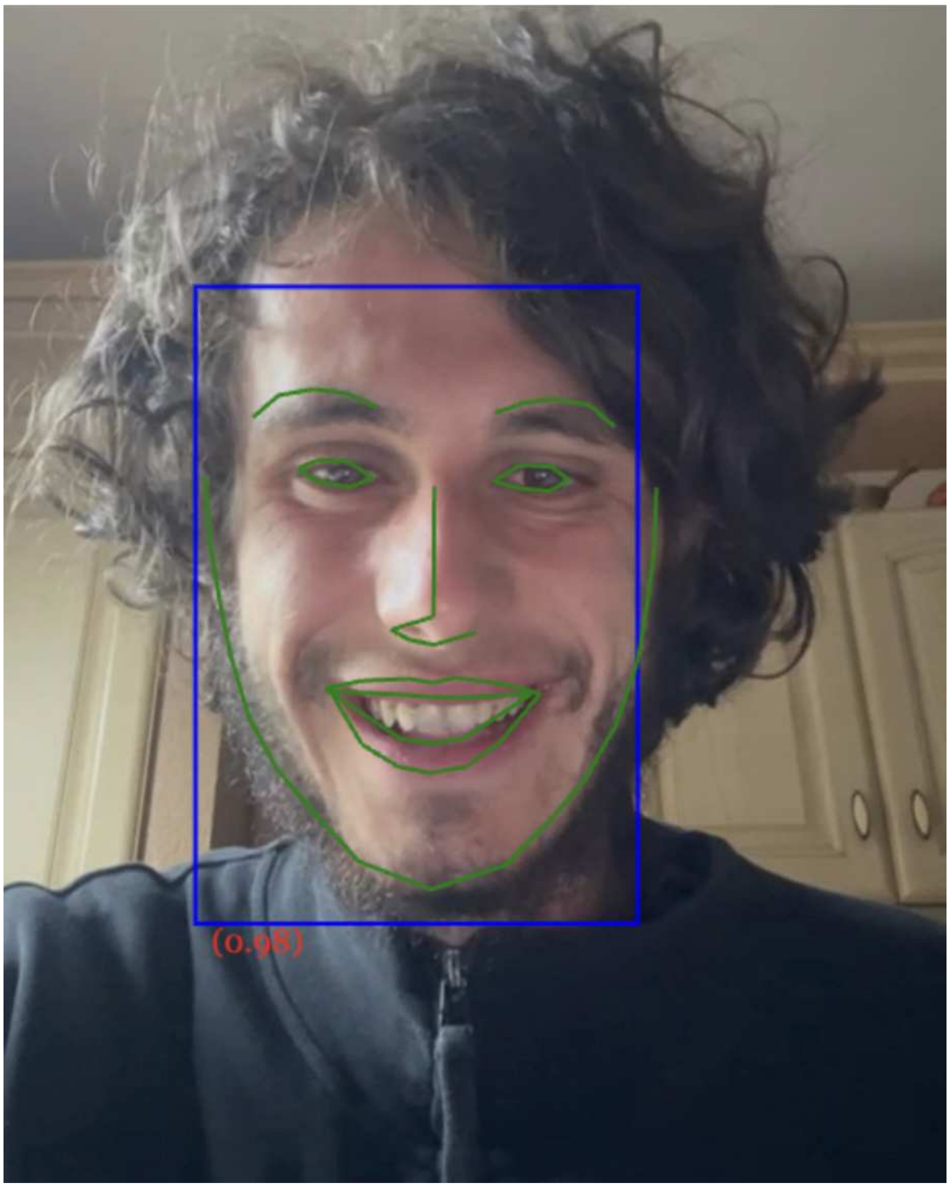

Figure 2.4: An example of a face detection and landmark extraction output.

2.4.1 Adaptation of the Face-api.js Library

To use the face-api-js library in this experiment, it was necessary to implement an HTML code that manages a web page where the video containing participant's facial response to each stimulus is loaded and the facial emotion recognition (FER) algorithm is applied. The final purpose of the HTML code is to generate a CSV file that includes a timestamp column and seven additional columns (one for each emotion), where the probability of detecting an emotion is calculated for each instant. To create the timestamp column a sampling rate of 100 milliseconds was set to consider the "high road" of LeDoux's theory which typically has a response lasting 250-300 milliseconds as seen in the first chapter (see LeDoux 1996 [30] and Pessoa 2010 [37]). This setup enables a systematic capture and analysis of emotional responses over time, bringing to the evaluation of the task and FER algorithm.

2.4.2 Video Collection and Pre-Processing

The video data were collected from 14 students, 6 of them identifying as female and 8 as male with ages going from 22 to 28. In this setup, 15 participants were shown 15 images. The facial expressions' videos of participants were collected using a setup involving two computers connected via a Zoom call. The first computer presented the online task to be completed while simultaneously capturing the participant's face during the whole experiment. The second computer recorded the Zoom call, including the participant's video. Once the full video (approximately 4 minutes long) was extracted, a pre-processing phase followed. Each video was segmented into short clips containing only the 4-second intervals during which the stimulus was presented. These short clips were then processed by a Facial Emotion Recognition (FER) algorithm based on the face-api.js library and managed via HTML code. For each mini clip (each one of them is corresponding to a specific image), a CSV file was generated, containing timestamps and emotion probabilities.

2.4.3 Data Processing

At this stage, it has been implemented a specific further MATLAB code to process the individual mini videos, grouping them for each image. Finally, for each image, the emotion probability signals captured by the FER algorithm for all participants were plotted on the same graph. Each signal is the plot of the probability of a certain emotion over time for each participant. Additionally, the mean and standard deviation of all the signals for each specific emotion were also plotted. This provided a clearer representation of the overall emotional response patterns, alongside with the variability of these responses across participants for each emotion.

2.4.4 Area Under Curve (AUC) and Emotion Classification

The Area Under Curve (AUC) was chosen to evaluate the predominance of emotions leading to the emotion classification using the FER algorithm. Specifically, the emotion with the highest AUC was considered as the classified emotion for a given image. However, it was observed that, with this method, the neutral emotion is frequently predominant. Therefore, to overcome this potential bias, a subsequent analysis was conducted excluding all neutral signals. Finally, a third analysis was performed by grouping the emotions into just two categories: positive (happiness) and negative (sadness, fear, anger, disgust), in order to assess how the results change with a simplified interpretation of the data. For each one of the three AUC-based classification analyses, a confusion matrix was generated using quiz responses as a ground truth, and accuracy and precision values were derived.

CHAPTER 3

Results

3.1 Online Task Results

The results of the two versions of the online task will be presented separated by gender to assess a potential presence of gender-related differences in emotional perception. This gender-specific analysis is essential to understand if participants identifying as male or female will react differently to visual stimuli, while for participant identifying as "other" data was splitted between "female" and "male" basing on the specific stimuli, since stimuli were selected to be different between genders.

3.1.1 Reduced Version

First of all, a histogram was plotted to show the frequencies of emotions associated with each single image. This allows a preliminary visual inspection of the data quality. Figures 3.1 and 3.2, down below, present the results for female and male participants, respectively.

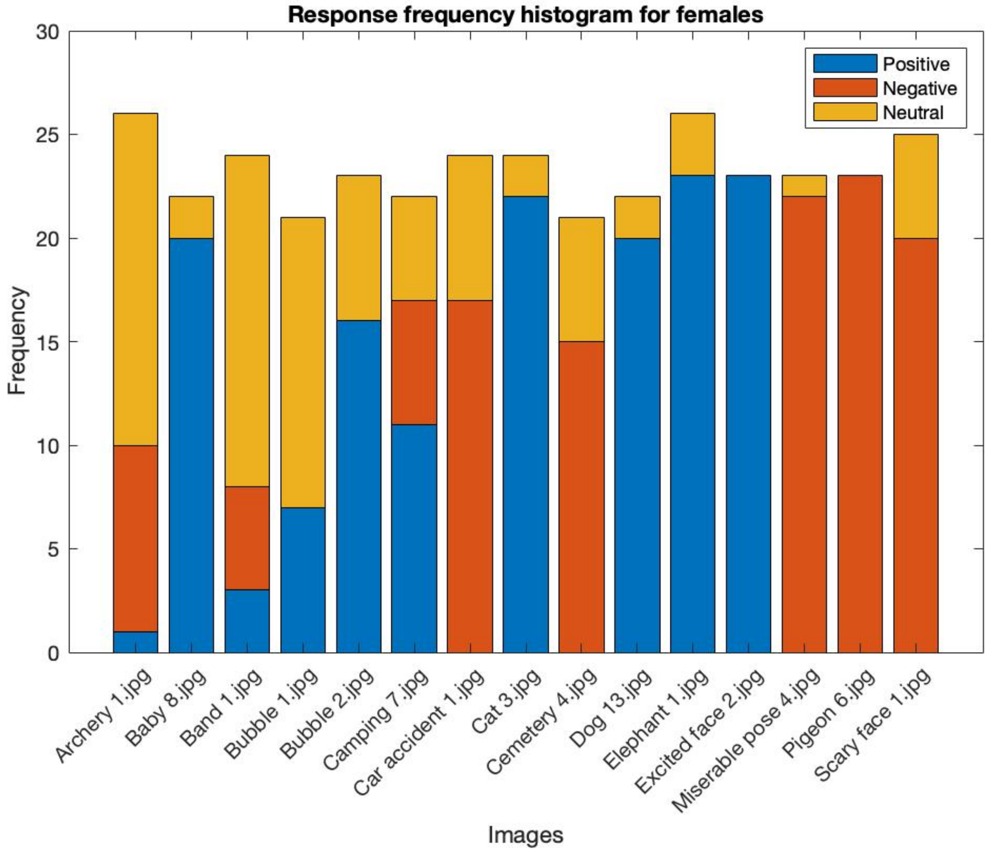

Figure 3.1: Histogram of emotions for female participants in the reduced version.

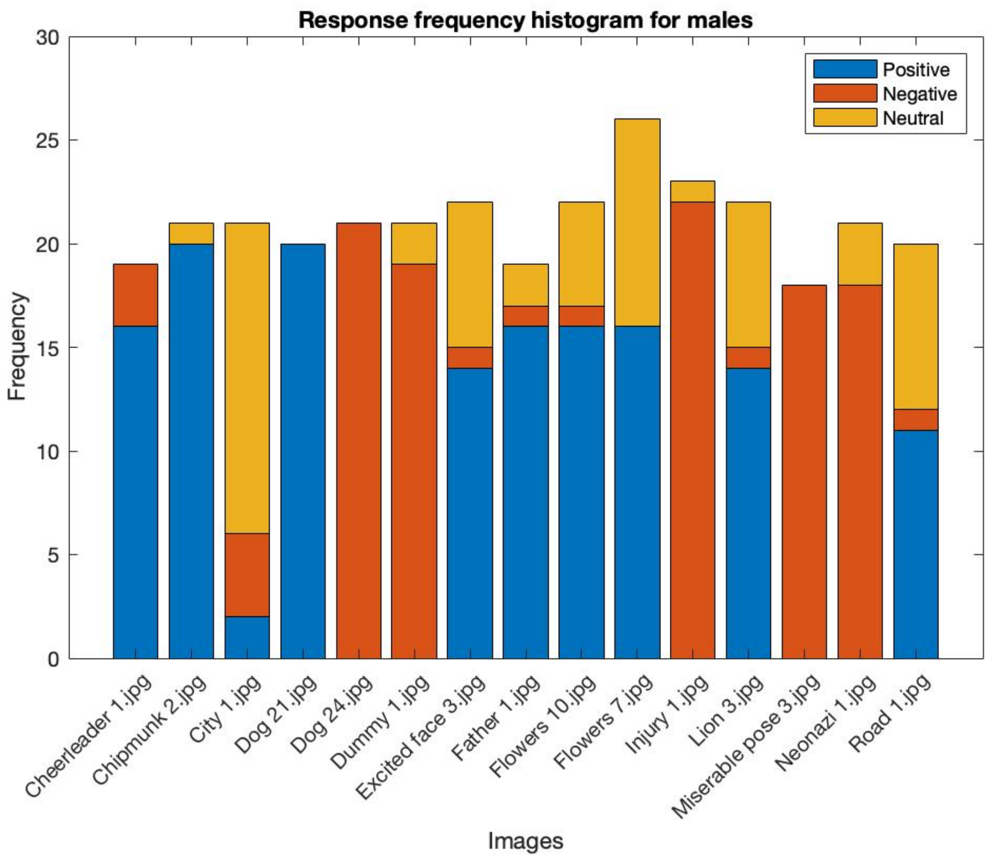

Figure 3.2: Histogram of emotions for male participants in the reduced version.

The figure down below 3.3. presents a confusion matrix and a representation of accuracy and precision generated from the data collected from the reduced version of the online task for female participants. In particular, (A) is the confusion matrix representing the distribution of perceived emotions compared to the expected ones. Each row represents the expected emotion (considered the ground truth in this first phase), while each column represents the perceived emotion (given from the quiz response). The diagonal shows the number of times the perceived emotion matched the expected one, while outside the diagonal mismatches are reported. The table (B) shows the row-normalization of the confusion matrix, highlighting accuracy for each class (blue cells). On the other hand, the table in (C) shows the column-normalization representing the precision of this classification (blue cells). Finally, the total accuracy was calculated and it's equal to 75,36%.

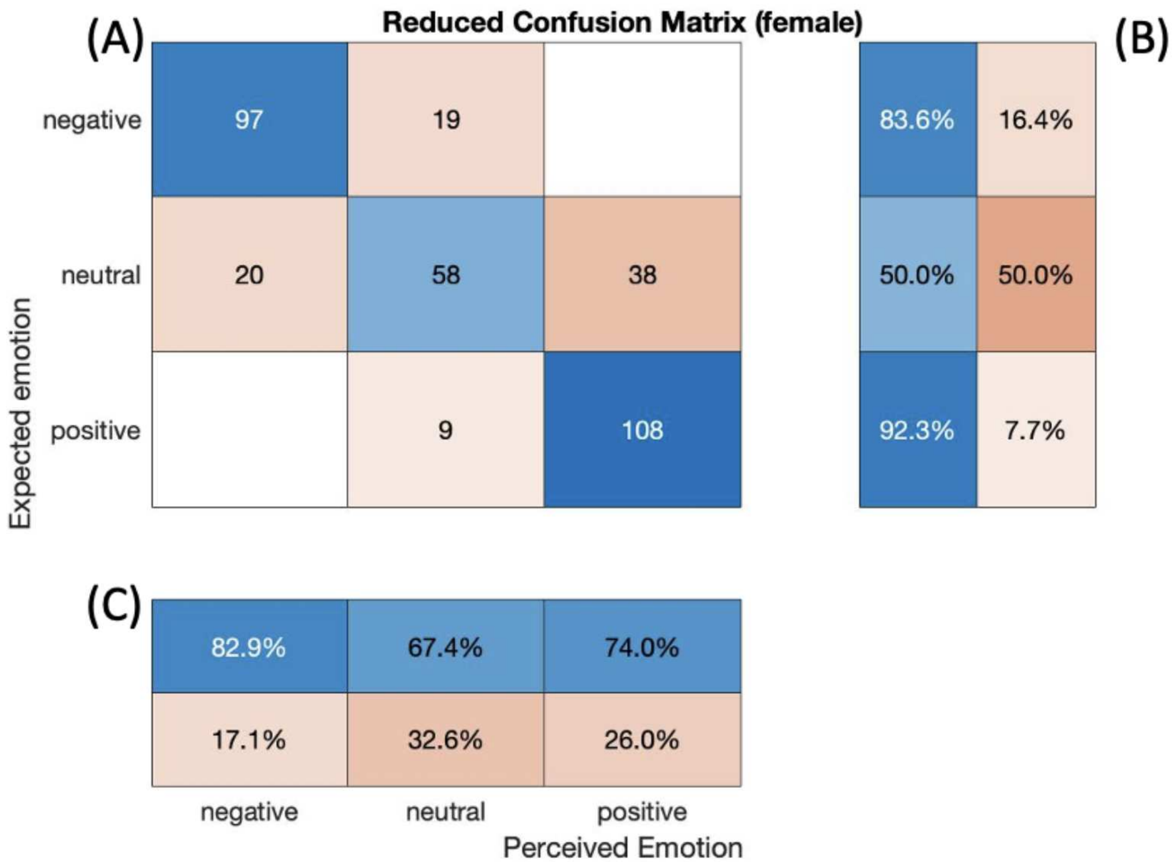

Figure 3.3: (A) Confusion matrix of the results from the reduced version of the online task for female participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

Down below (figure 3.4), the results of the same analysis for male participants are reported in figure XX. Moreover, total accuracy has been also calculated and it's equal to 76,68%.

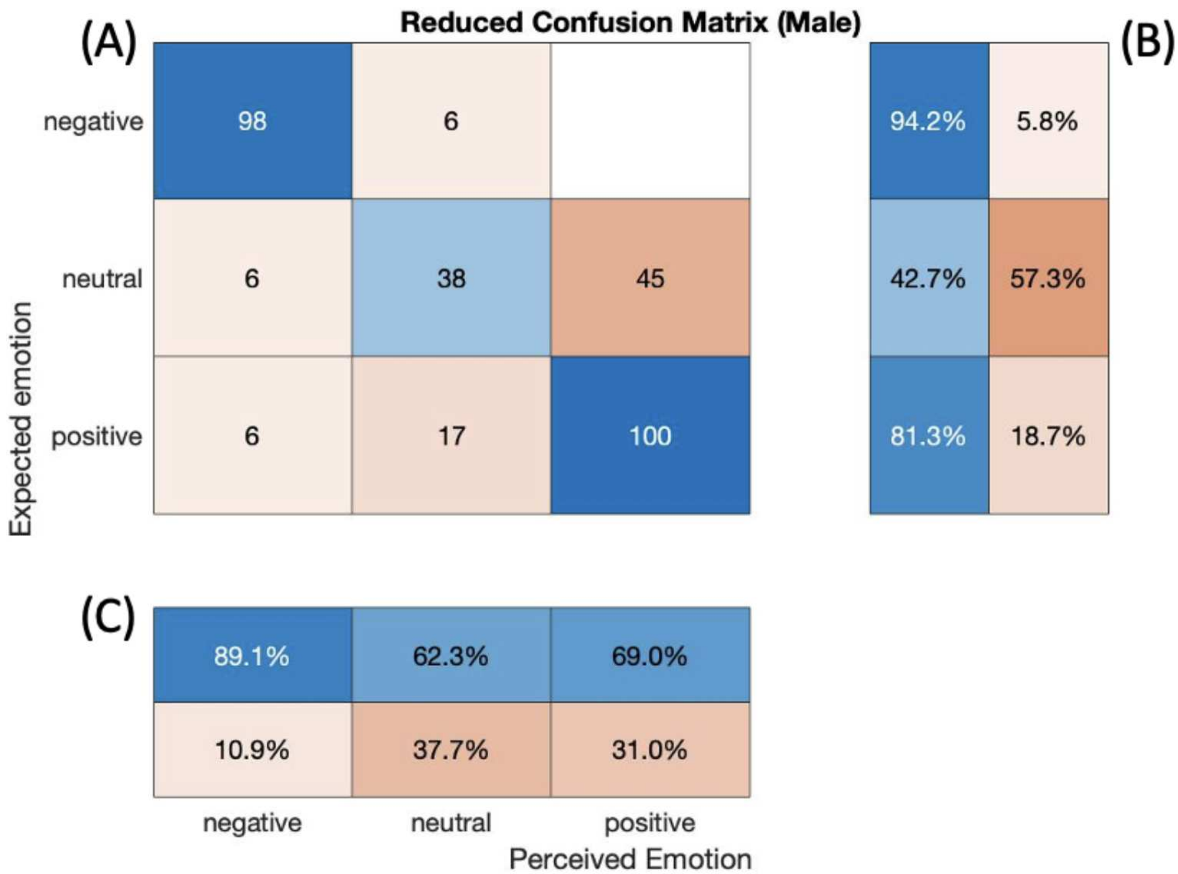

Figure 3.4: (A) Confusion matrix of the results from the reduced version of the online task for male participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

3.1.2 Complete Version

The results of the complete version of the task are presented in the same way as the reduced version, with the only difference in the emotional categories. Indeed, the categories "positive," "negative," and "neutral" have been expanded into Ekman's seven basic emotions, making the data more complex. Figures 3.5 and 3.6 display histograms of response frequencies for each image, respectively for females and males.

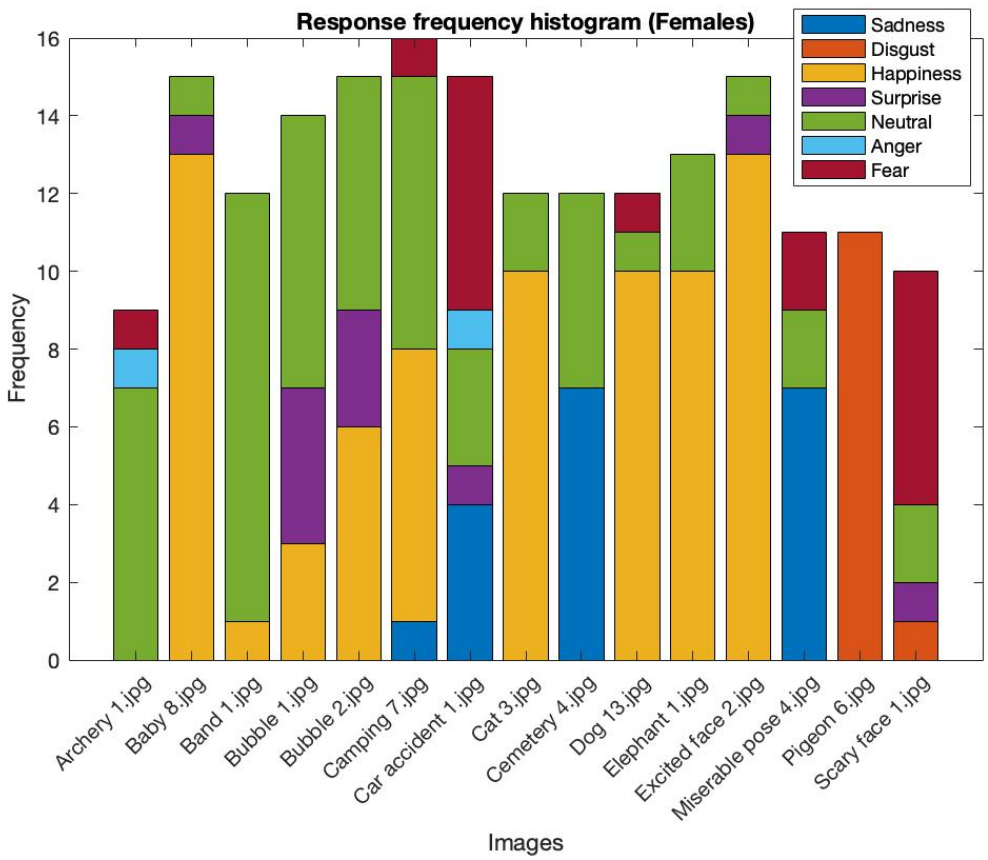

Figure 3.5: Histogram of emotions for female participants in the complete version.

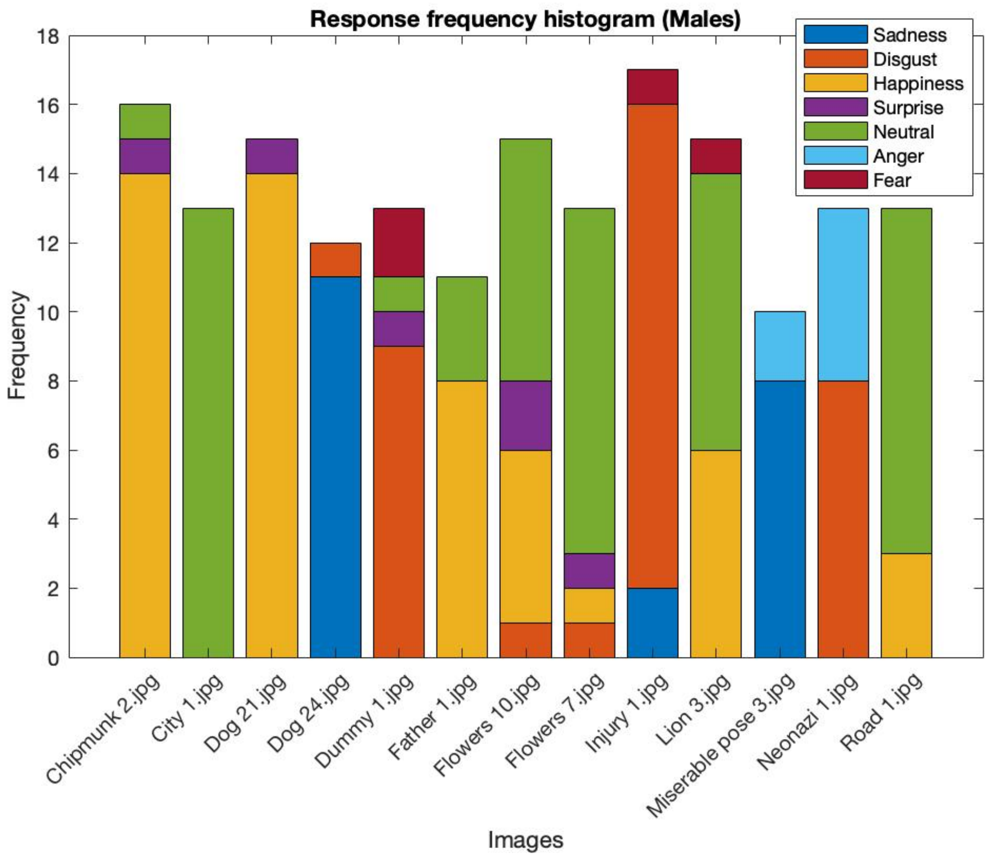

Figure 3.6: Histogram of emotions for male participants in the complete version.

Meanwhile, Figures 3.7 and 3.8 show tables with visual representations of the confusion matrices, accuracy, and precision, as seen previously. One more time, the results are separated by gender. Moreover, the total accuracy was calculated for each gender, resulting as 68,23% for female and 70.87% for male.

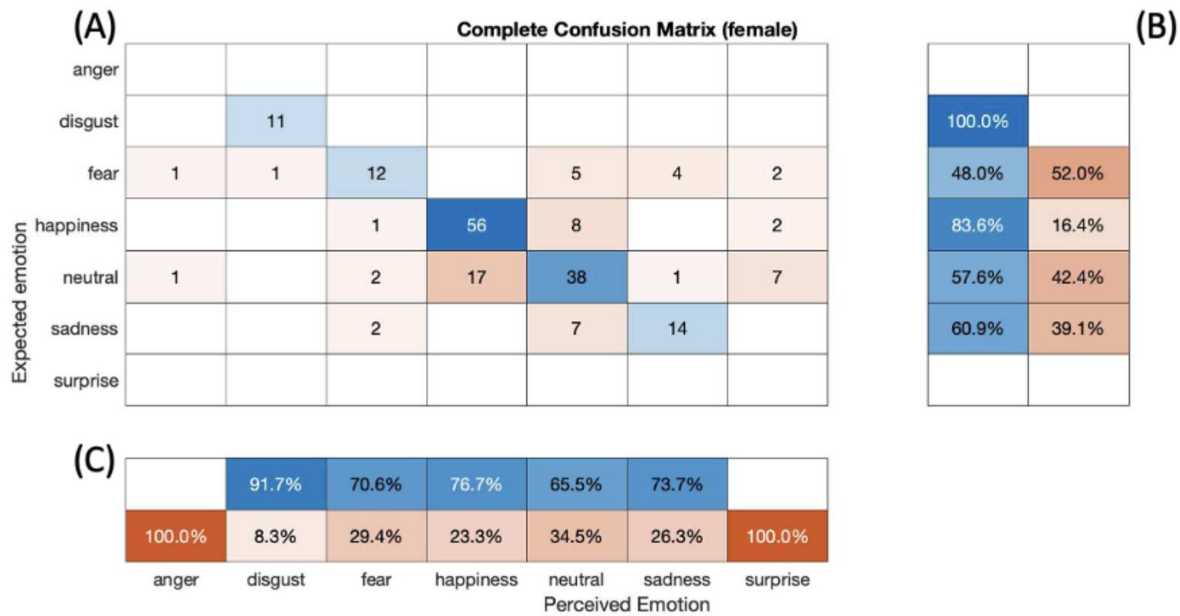

Figure 3.7: (A) Confusion matrix of the results from the complete version of the online task for female participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

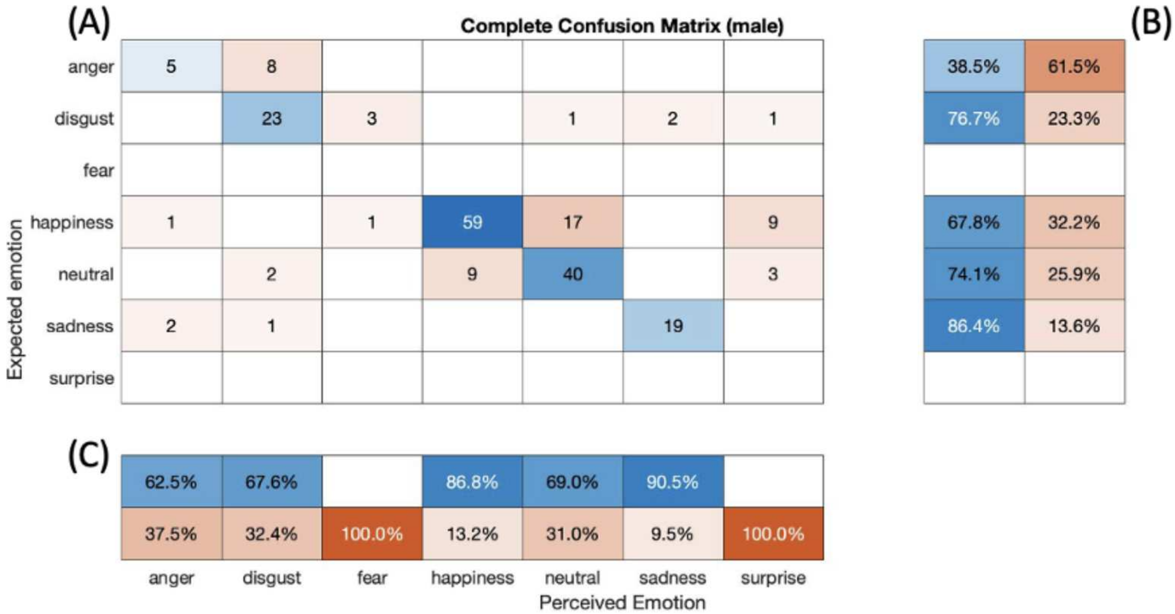

Figure 3.8: (A) Confusion matrix of the results from the complete version of the online task for male participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

3.2 Offline FER Algorithm Results

The results of the offline FER Algorithm are divided into two parts. The first part involves a simulated test performed on a video of an expert actor mimicking the seven basic emotions without any external stimuli. This phase serves to validate the accuracy of the FER algorithm by ensuring that it can correctly recognize and classify facial expressions corresponding to each emotion. The second part tests the algorithm in a real scenario, by analyzing video recordings of participants' facial responses triggered by visual stimuli. This phase evaluates the algorithm's performance in detecting emotional reactions in real-time, providing a more practical assessment of its effectiveness in capturing emotional expressions under experimental conditions.

3.2.1 Performance on Actor

The performance on the expressions simulated by an actor is promising. In fact, the FER algorithm was able to identify the face, extract the landmarks and classify the emotions with a probability always higher than 0.95. In figure 3.10 an example is given but all the other emotions have similar trends.

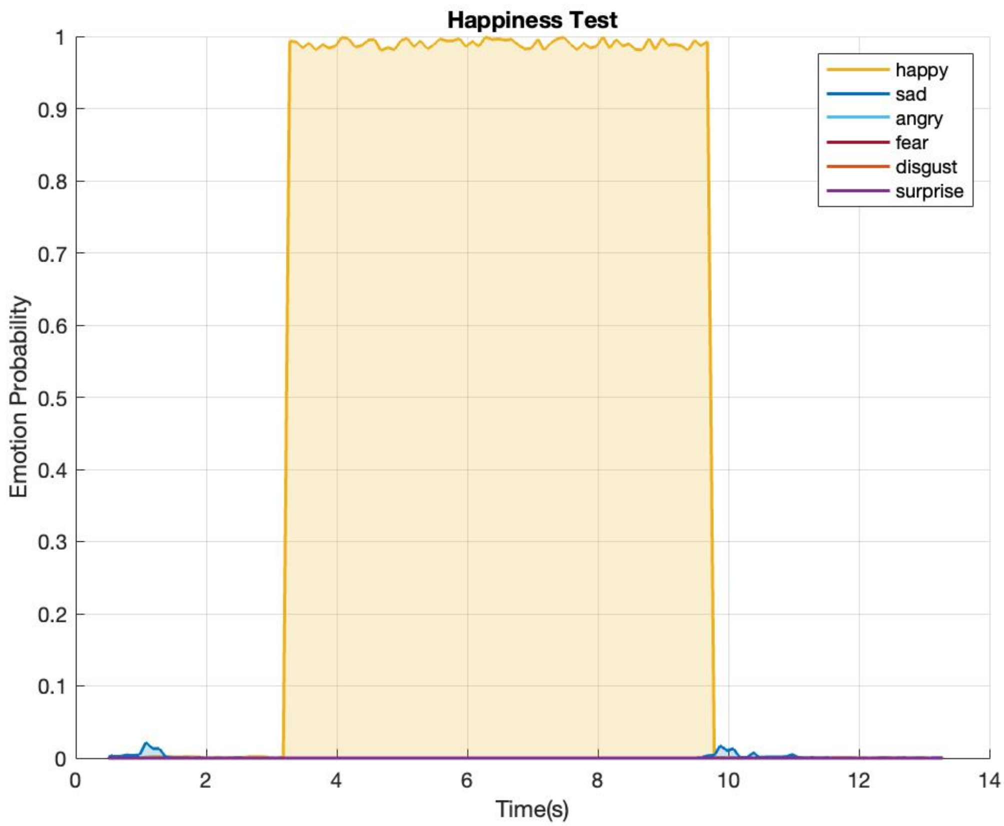

Figure 3.9: Happiness probability signal from a simulation on actor.

3.2.2 Predominant Emotion Detection Across Real Participants

The results of the FER algorithm for predominant emotion's detection are presented similarly to those of the online task. A visual representation of the confusion matrix is shown, alongside with the accuracy and precision value for each class. However, in this case, the rows display the emotions perceived (based on quiz responses), while the columns show the emotions detected by the FER algorithm. The detected emotion is the one with the highest AUC from the probability signals for each image.

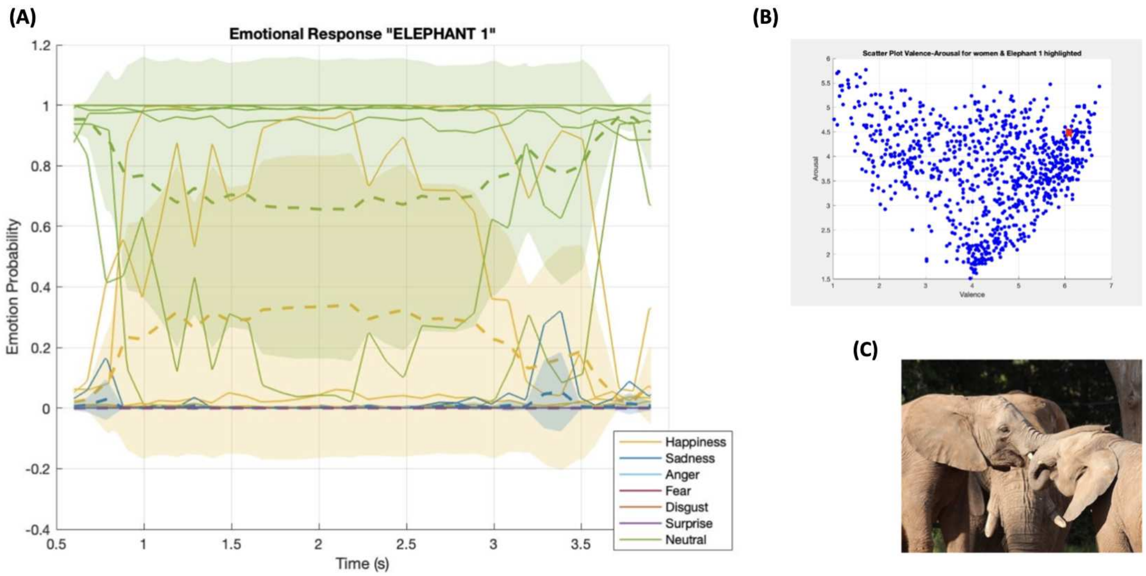

Figure 3.10: (A) An example of the graph showing all the probability signals for each emotion for every participant. The dashed line represents the average of the signals for each emotion, while the shaded areas represent the standard deviation. (B) Plot of the entire OASIS dataset with the image used as the stimulus marked in red. (C) The image used as the stimulus in this case.

While figures 3.11 and 3.12 display results for females and males, respectively.

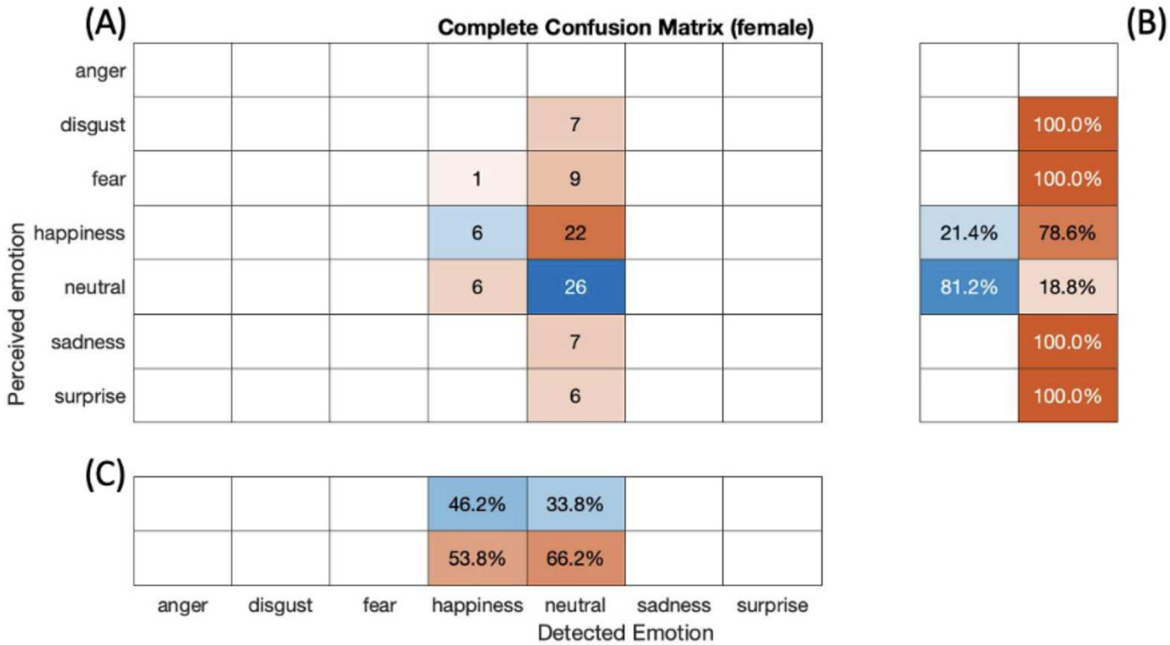

Figure 3.11: (A) Confusion matrix of the results from the complete version of the online task for female participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

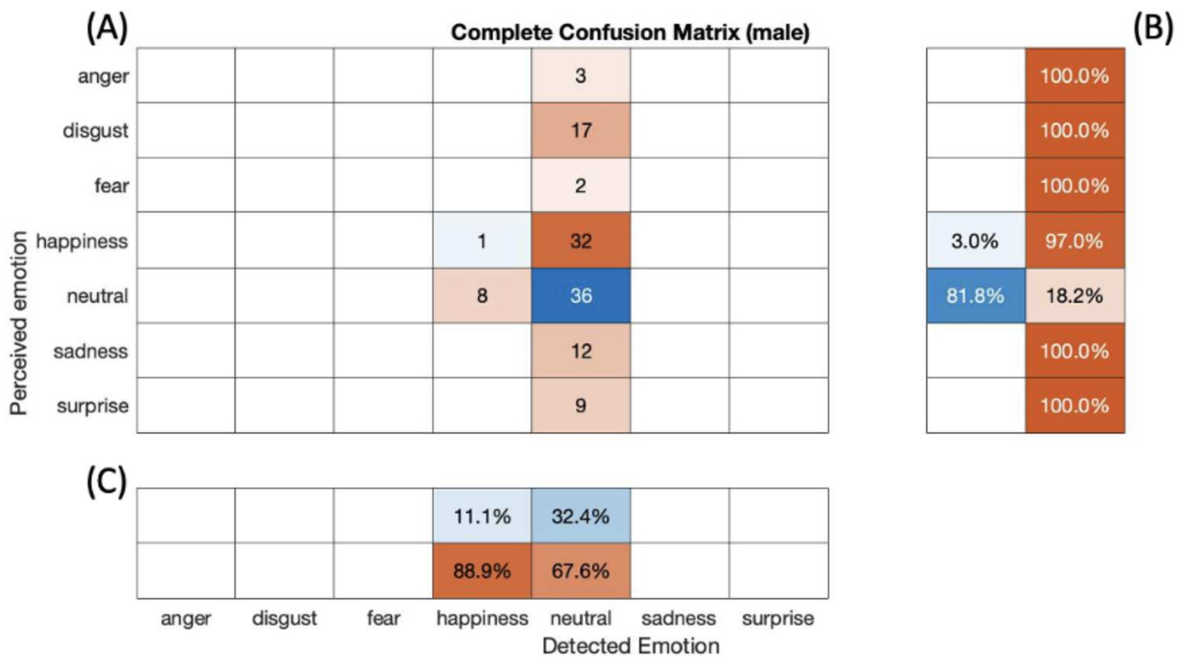

Figure 3.12: (A) Confusion matrix of the results from the complete version of the online task for male participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

Additionally, the total accuracy is 30,83% for male and 35,56% for female. Since the overall accuracy rates were quite low, the "neutral" emotion was excluded to assess if the emotional classification improved, while still using the AUC-based approach to establish the predominant emotion. Figures 3.13 and 3.14 show the results of this filtered version for female and male participants, respectively.

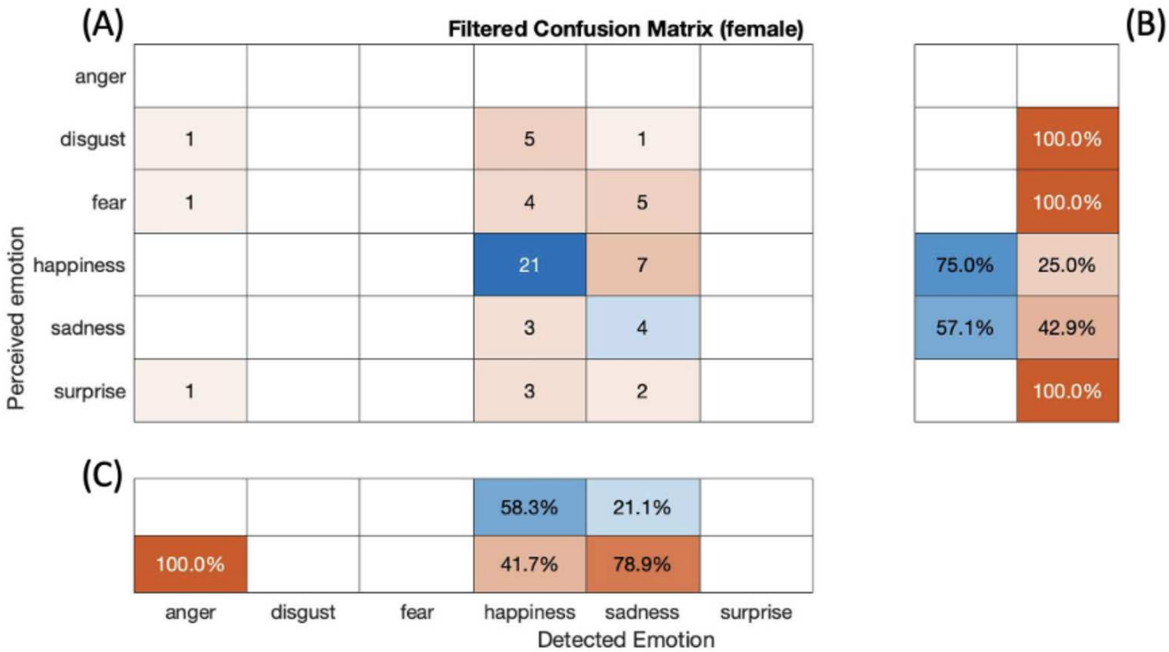

Figure 3.13: (A) Confusion matrix of the results from the filtered version of the online task for female participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

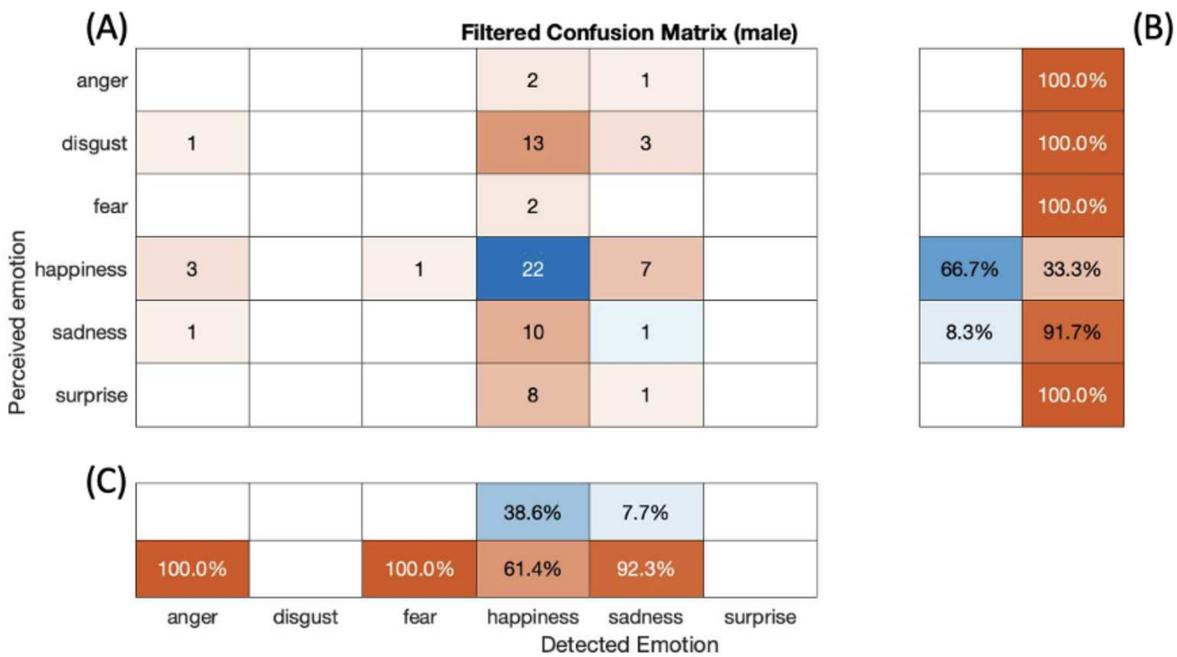

Figure 3.14: (A) Confusion matrix of the results from the filtered version of the online task for male participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

Unfortunately, overall accuracy only improved in the case of female participants, while it actually became worse for males. In particular, it increases to 43.10% for females but drops to 30.26% for males. This led to an additional analysis, where a simplified model based only on two emotion categories ("positive" and "negative") was adopted. This approach resulted in improving class accuracy, as shown in figures 3.15 and 3.16, with overall accuracy rising to 63.46% for females and 43.28% for males.

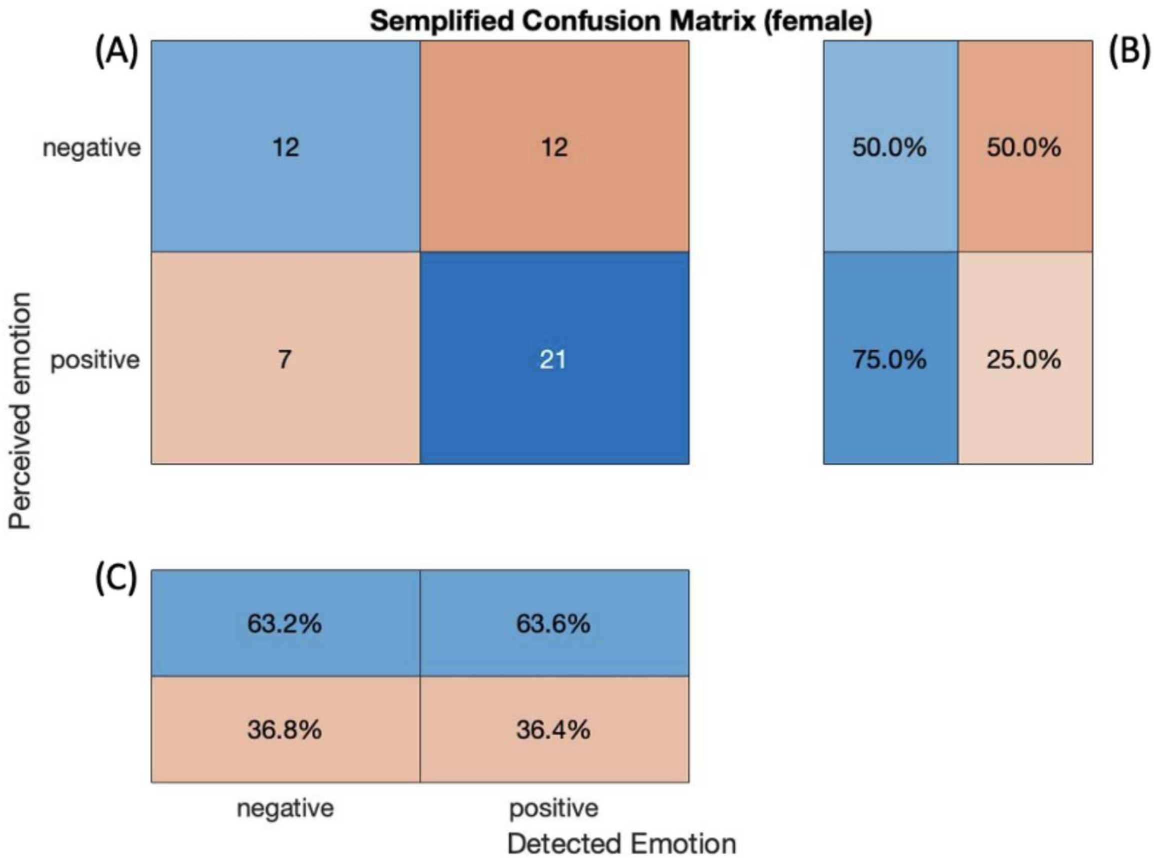

Figure 3.15: (A) Confusion matrix of the results from the simplified version of the online task for female participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

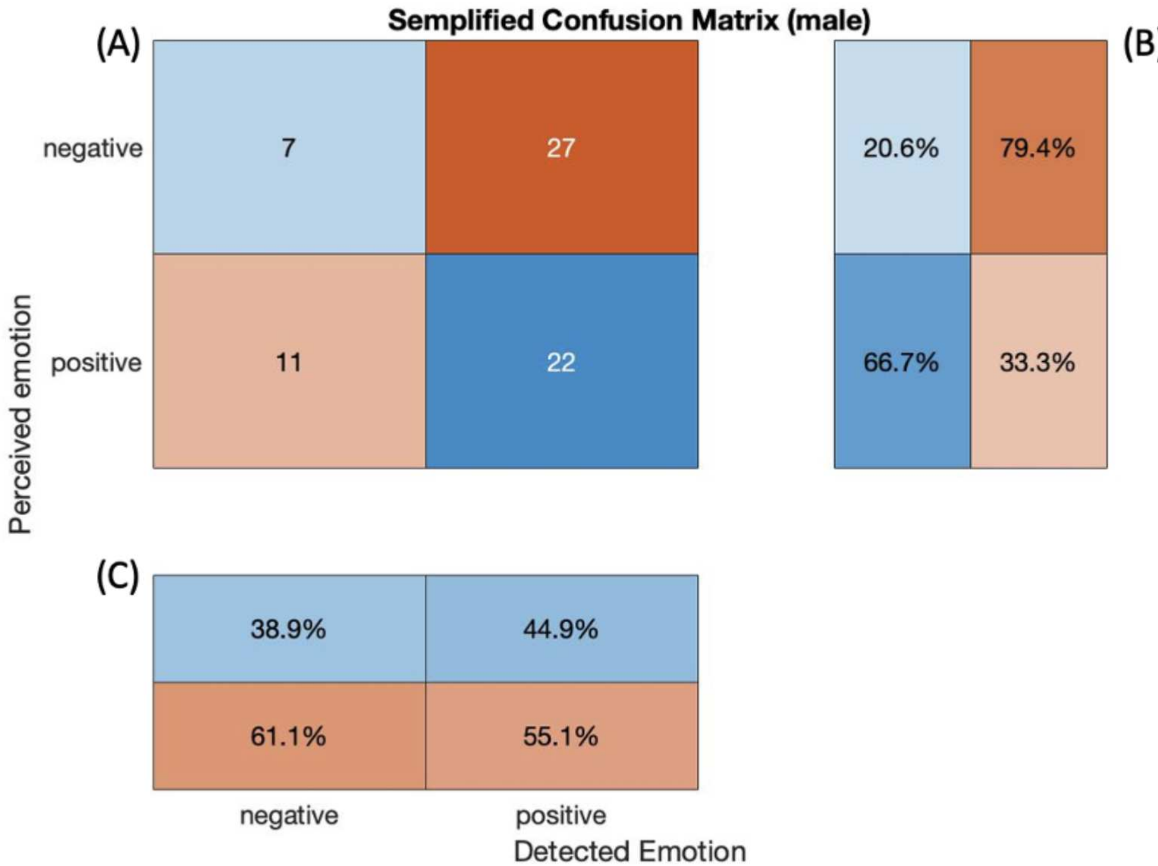

Figure 3.16: (A) Confusion matrix of the results from the simplified version of the online task for male participants. (B) Accuracy table (blue cells). (C) Precision table (blue cells) and the Error Rate (red cells).

Although these results may not seem very promising, the following chapter will show that by shifting the analytical perspective, some interesting insights can still be uncovered.

CHAPTER 4

Discussion

In this thesis, the findings demonstrated that the online task successfully revalidated a dataset of images (OASIS), confirming that participants could reliably associate the stimuli with recognizable emotions. High accuracy was achieved for emotions such as happiness, disgust, and sadness, validating the dataset for further experimental use. However, when these same validated images were used to test the FER algorithm, the performance dropped significantly. The algorithm struggled to reach the accuracy levels achieved in the online task, particularly for negative emotions like fear and anger. While happiness and neutral emotions were detected with better accuracy, the results highlighted a significant gap between human emotional perception and the algorithm's ability to recognize spontaneous emotional reactions. This discrepancy emphasizes the limitations of current FER systems and suggests improving future works.

4.1 Online Task Validation

4.1.1 Reduced Version

The reduced version of the online task, which classified emotions into "positive," "negative," and "neutral" categories, showed promising results. Both male and female stimuli reveal no key differences in emotional perception accuracy. Indeed, both of them demonstrated high accuracy in detecting "negative" emotions (F = 83.6%, M = 94.2%) and "positive" emotions (F = 92.3%, M = 81,3%), which is optimal. However, the "neutral" emotion was not at these levels (F = 50.0%, M = 57.3%). This could be due to several psychological and cognitive factors. First, the ambiguity in emotional perception, according to which, participants may find it more difficult to interpret neutral stimuli because it seems ambiguous. Research shows that individuals often associate ambiguous stimuli with stronger emotional reactions (positive or negative) rather than perceiving them as neutral (see Reynolds et al. 2019 [115] and Everaert et al. 2012 [116]). Indeed, participants may have oversimplified complex or ambiguous emotions by choosing "positive" or "negative" labels instead of "neutral" since these "extreme" emotions seem to be more distinct and easier to recognize (see Crane & Gross 2013 [118]). Second, it's what is called the "negativity bias", which is the tendency in psychological experiments for participants to prioritize negative emotions over neutral or positive ones (see Xu et al. 2021 [117]). However, while considering these issues, the overall accuracy for both male and female groups is still acceptable (F = 75,36%, M = 76,68%). Additionally, precision values are relatively high, which is index of a limited dispersion of the data, suggesting that the dataset extracted from OASIS is sufficiently valid to be used for this experimental protocol.

4.1.2 Complete Version

In the complete version of the task, emotions were categorized using Ekman's seven basic emotions rather than just the three categories (positive, negative, neutral) used in the reduced version. The data reveals some interesting gender-based differences in emotional accuracy. First, females had notably higher accuracy in detecting disgust (100%) compared to males (76.7%). This result is not surprising since previous researchers investigated female's height sensitivity to emotions related to danger or aversion, such as disgust. This might have an evolutionary root; indeed, disgust can operate as a protective mechanism (see Wager et al. 2003 [119]). Vice versa, males showed higher accuracy in detecting sadness (86.4%) compared to females (60.9%) and neutral emotions (74.1%) compared to females (57.6%). This may be due to the "negativity bias" (see Xu et al. 2021 [117]) and social or cognitive factors controlling the gender emotional expression (see Brody & Hall 2008 [120]), respectively. Meanwhile, happiness accuracy was comparable between both groups (F = 83.6%, M = 67.8%), with females performing slightly better, supporting what is already known from the literature, i.e., women tend to identify positive emotions more easily (Hoffmann et al. 2010 [121]). Finally, as for the reduced task, the overall accuracy is sufficient to guarantee the validation of the dataset of images extracted from OASIS (F = 68.23%, M = 70.87%). This validation is also supported by the precision values for each class.

4.2 Offline FER Algorithm Validation

4.2.1 Performance on Actor-generated Videos vs Real Participants

The FER algorithm initially demonstrated strong performance when tested on actor-generated videos, where emotions were mimicked without the presence of any external stimuli. This controlled environment allowed the algorithm to identify emotional expressions, validating its capability to process idealized facial expressions accurately. The results of this first test agree with other similar studies which have shown that the algorithm performs best in controlled situations where facial expressions are clear and sometimes exaggerated if compared to a real situation (see Martinez et al. 2017 [122]).

However, when transitioning to real-world scenarios, represented by the experimental protocol illustrated previously, in which real participants had to respond to emotional stimuli, the algorithm's performance dropped significantly. Indeed, for female participants, the overall accuracy was 35.56%, while for males it was 30.83%. In both groups, the algorithm failed to detect fear, disgust, anger, sadness, or surprise. Instead, happiness and neutral emotions were the only reliably detected categories. Specifically, happiness was identified with an accuracy of 21.4% for females and 3% for males, while neutral emotions achieved 81.2% and 81.8% accuracy for females and males, respectively. The precision scores also reflected this trend, with happiness for females showing a 46.2% precision and neutral at 33.8%. In males, happiness had a very low precision (11.1%), while neutral emotions were identified with a precision of 32.4%. These first results are not surprising, because they reflect studies that highlight the challenges of

FER in spontaneous emotional responses. Indeed, real-world emotional responses tend to be more nuanced and less exaggerated, making them harder to decode (see McDuff et al. 2015 [123], Kulke et al. 2020 [124]). However, despite the paucity of these initial results, a gender-related difference in the detection of the "happiness" emotion is starting to glimpse. To highlight these fine results, it was necessary to apply a first filter.

Therefore, it has been chosen to exclude the "neutral" emotion to study the predominance of other emotions and see if the results improve. As hoped, accuracy improved for both genders. The first effect of the filter shows a gender-related differences in emotion detection, particularly for "happiness" and "sadness". For females, overall accuracy rose to 43.10%, with "happiness" detected at 75.0%, and "sadness" at 57.1%. While precision was more moderate for "happiness" achieving 58.3% against the 21,1% of "sadness". For males, the results were similar but less encouraging. Indeed, the overall accuracy was lower at 30.26%, with "happiness" detected at 66.7% and "sadness" at only 8.3%. While precision for males was particularly poor, with "happiness" and "sadness" equal to 38.6% and 7.7%, respectively. However, for both genders, the detection of "fear", "disgust", and "anger" remained poor. This suggests that, while the algorithm performs well for positive emotions, it struggles with negative (or more complex) emotional states. And this theory aligns with current researches, which show that FER algorithms often have difficulties recognizing negative emotions, especially in real-world scenarios (see Kosti et al. 2019 [125]).

At this point a further simplification has been introduced in the interpretation of the data by grouping emotions into two categories (positive = happiness and negative = sadness, fear, disgust, and anger). In this case, precision values improved especially for females (positive = 63,6%, negative = 63,2%) in confront to male (positive = 44,9%, negative = 38,9%). While a significant improvement in accuracy was observed. In particular, female participants had an overall accuracy of 63.46% (against the 43,28% of the male). Additionally, positive emotions were detected more accurately for both genders (F = 75%, M = 66.7%). However, negative emotions remain more challenging to identify, especially for males (F = 50%, M = 20,6%). This aligns with studies showing that FER systems often perform better with positive emotions compared to negative ones (see Cowen et al. 2019 [126]).

4.2.2 Potential Factors Affecting FER Accuracy

In addition to gender differences and the difficulty of coding complex emotions via FER, there are other factors that may have influenced the accuracy of the results. It's known from literature that FER accuracy is linked to the quality of the images used as a visual stimulus. Indeed, Image's clarity and resolution influence the algorithm's performance (see Wang et al. 2004 [127]). This leads to some interesting considerations.

First, some participants informally mentioned that they expected stronger stimuli. This is understandable, considering that modern individuals are exposed daily to high-intensity emotional stimuli from social media, movies, and news, which raises the emotional sensitivity threshold (see Gross 2002 [128]). Indeed, research shows that the constant exposure to high emotional stimuli in media have a desensitization effect on individuals, making it harder for them to express emotionally their feelings in experimental settings, especially subtle emotional cues (see Potter & Bolls 2012 [129]). For instance, social media and news platforms use emotionally charged content to increase engagement, since the latter allows monetization, which is often their dominant decisive variable. Also, movies and television use the same strategy to captivate audiences. The problem is that, over time, this may reduce the ability to perceive (and react) to a lower-intensity emotion in more natural situations (see Furl et al. 2012 [130]). This phenomenon is also known as emotional habituation, and it could be the major factor of an underperformance of FER systems (see Wirth & Schramm 2005 [131]).

Besides, other participants informally commented that the visual stimuli were too USA-centered, which may have reduced their emotional resonance for an Italian cohort. This aligns with research showing that the emotional response is influenced by cultural context. Indeed, people from different cultures interpret emotion differently (see Ekman et al. 1969 [132]). This cultural discordance between stimuli and participants can lower the interpretation and affect FER's performance (see Mesquita 2001 [133]).

TThe other factor that might have affected the FER's accuracy is the "freezing" phenomenon, that consists in exhibiting neutral expressions in response to strong negative stimuli. This can be interpreted as an emotional regulator or survival mechanism, where individuals suppress their visible negative reactions under extreme stress (see Roelofs 2017 [134]). In this case, participants exhibited a neutral expression in response to stimuli expected to evoke a strong negative emotional response. This outcome is in contrast with the big variability shown from weaker stimuli's reactions. This suggests that the absence of emotional expression may not lead to an emotional disengagement, but rather a "freezing" response. Another possible interpretation of the "freezing" phenomenon is that Facial Emotion Recognition (FER) may not be sufficient to fully capture the emotional states of individuals. In this case, other physiological measures (explored in the first sections of the thesis), such as heart rate, skin conductance (GSR), or even pupil dilation might be necessary to provide a more comprehensive understanding of emotional responses. These physiological markers often capture unconscious emotional arousal, even when facial expressions remain neutral, indicating the presence of stress or emotional regulation mechanisms (see McDuff et al. 2016 [72]). An example of this phenomenon is shown in Figure 4.1.

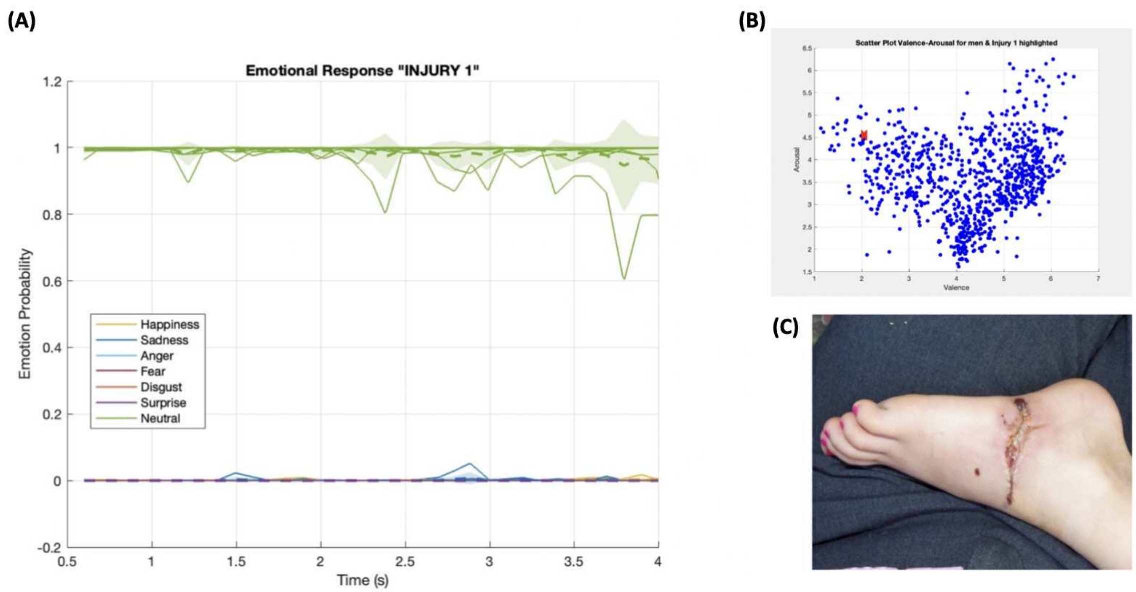

Figure 4.1: (A) An example of the graph showing the "freezing effect". In particular it shows all the probability signals for each emotion for every participant. The dashed line represents the average of the signals for each emotion, while the shaded areas represent the standard deviation. (B) Plot of the entire OASIS dataset with the image used as the stimulus marked in red. (C) The image used as the stimulus in this case.

4.3 Limitations

4.3.1 Image Dataset

The first limitation is about the "power" of the images of the OASIS dataset. As mentioned earlier, on one hand, some images were not intense enough to evoke strong emotional responses. Although, on the other hand, apparently strong images may have evoked a "freezing" effect precisely because of their strength. Additionally, a cultural bias was observed, as several participants informally noted that the images seemed too centered on a USA's cultural context, which is different from the cohort nationality.

4.3.2 Experimental Conditions

The experiment took place in a controlled setting to ensure uniformity in data collection. However, this controlled setting may have been too rigid. Indeed, the researcher was constantly present during the experiment to ensure the collection of a high-quality video and solve any possible problem with the online task. This could have unintentionally influenced the participants' emotional responses by making them feel observed and judged, creating a less "natural" environment. Additionally, probably the participants' emotional baseline was not ideal for accurate data collection. Indeed, the videos were recorded right before lunch break, and some of the participants appeared visibly tired.

4.3.3 Demographics

The online task and FER dataset is completely composed of university students, all of them, with only one exception, are Italian. Additionally, participants are all similar ages (ranging from 22 to 28) and have the same level of education. This is a clear limitation related to the WEIRD (Western, Educated, Industrialized, Rich, and Democratic) bias (see Henrich et al. 2010 [135]). This type of demographic homogeneity makes it impossible to widen the results across different populations. And this makes this study absolutely unrepresentative of the real population.

4.3.4 Artificial Intelligence Regulation and Ethical Considerations

The European Union's AI Act (see European Commission 2024 [136]), which introduces a comprehensive legal framework for the regulation of AI systems, includes specific directives for FER systems. Emotion recognition systems are defined as AI technologies that identify emotions basing on biometric data, such as facial expressions. Generally, these types of systems fall under the "high-risk" category due to their reliance on biometric features. Therefore, these systems are subject to rigid regulation including transparency, data privacy, and ethical considerations, especially when integrated into public tools. In particular, the AI Act definitively prohibits the use of AI systems that build facial recognition databases through untargeted scraping of images from the internet or CCTV footage, that's because this practice can lead to significant privacy violations and contribute to a dangerous mass surveillance. Furthermore, it is important to underline that the regulation distinguishes between biometric categorization and emotion recognition. As the biometric categorization is about assigning individuals to categories based on traits like age, gender, or behavioral traits while in the emotion recognition the focus is on identifying just the emotional states. The FER system employed in this thesis does not directly violate the AI Act, although it operates in a framework with stringent regulations that need careful consideration. First, the video data collection process includes biometric data, thus it falls under the "high-risk" category. Therefore, it's fundamental to ensure transparency, consent, and privacy protection during both data collection pre-processing and processing phases. That's why all participants were aware of the nature of the experiment and consented to their participation. Additionally, the emotional analysis was just landmark-based, and all videos were cancelled right after the acquisition of the required data to track the emotional probability signal. Moreover, it is important to emphasize that this work lies in a restricted academic research area, aimed at understanding the mechanisms behind an emotional response and developing a protocol to measure this response accurately and precisely. All future applications of these FER systems, particularly in public or commercial contexts, must carefully consider ethical aspects, such emotion manipulation, invasion of privacy, and misuse of sensitive biometric data. It's crucial for all researchers and engineers to remember that the purpose of their work (non-only in emotion recognition) must always align with the user's safeguard of personal rights, respecting their dignity and autonomy.

4.4 Future Work

4.4.1 Online Task and Offline FER Algorithm Integration

One of the key aims for future work is integrating the online task with the offline FER algorithm to create a unified fluid platform. The integration of these two parts would allow for a real-time analysis of both self-reported data and automatic emotion classification, providing a valid tool for a richer and deeper understanding of emotional responses. This would enable immediate validation of the FER algorithm by comparing self-reported emotions and facially detected emotions in real time, and this would improve accuracy and reduce the timing of post-experiment processing. A high-quality FER tool would have multiple applications across various fields. For instance, it could be used in healthcare for monitoring patients' emotional well-being, such as detecting signs of depression or anxiety in clinical settings (see Girard et al. 2013 [137]). While, in education, it could help educators to track students' engagement and attention state by monitoring the emotional responses during the learning activities (see D'Mello & Graesser 2012 [138]). Moreover, in human-computer interaction, FER systems can enable online assistant's bots to enhance their empathetic response to users (see McDuff et al. 2016 [72]). Additionally, FER could be useful in security and law area, by detecting suspicious or stress-related behaviors (see Dhall et al. 2016 [139]). Naturally, these applications often represent a potentially major violation of privacy, so all ethical standards must be respected. In this last consideration, the FER system itself can be useful in identifying and mitigating the misuse or manipulation of emotion detection technologies. Indeed, FER can recognize a potential manipulation through misperception in fields like marketing, advertising, or politics, as illustrated in the previous sections. By continuously monitoring and analyzing emotional responses, FER systems can flag situations where emotional manipulation may be taking place, particularly if it detects patterns of extreme or exaggerated emotional responses (see Calvo & D'Mello 2010 [84]).

4.4.2 Datasets Differentiation

Future work should also focus on exploring and testing other datasets beyond OASIS, such as IAPS and NAPS and not only, since using a single dataset can emphasize the limitations already discussed. First of all, incorporating multiple datasets would allow the algorithm to handle cultural differences and avoid biases linked to the overrepresentation of the Western culture. Moreover, it would enhance the algorithm's ability to recognize more nuanced emotions that are not well-represented in a single dataset. Integrating different datasets, it would be possible to design specific visual stimuli aim to elicit stronger emotional reactions, helping to go deeper into the study of "freezing" mechanism.

CHAPTER 5

Conclusions

The various phases of this thesis allowed the development and validation of an experimental protocol for Facial Emotion Recognition (FER), including the creation of an online task designed to present visual stimuli and capture a subjective emotional response (through a simple quiz) combined with an objective emotional response (through video recordings).

Through the online task, the dataset extracted from OASIS was tested and validated. In particular, the dataset had high accuracy in detecting "positive" and "negative" emotions in both genders but faced challenges with "neutral" ones. This may reflect the difficulty of participants to interpret and classify an ambiguous stimulus in this kind of experiment. The online task also allowed to validate the image dataset also for a more complex classification (Ekman's seven basic emotion states), where a first gender-based differences started to come out. For instance, females exhibited higher accuracy in detecting disgust and happiness, whereas males were better at identifying sadness. This gender difference aligns with other research, suggesting that women may be more sensitive to certain emotions due to evolutionary and social factors. Finally, despite some emotions with nuanced accuracy results, the OASIS dataset was validated for this kind of experiment, affirming its robustness considering the overall accuracy.

The offline FER algorithm performed successfully with simulated and exaggerated expression (actor-generated videos). While, on the other hand, it faced significant challenges when applied to ”real” experimental data. In particular, the algorithm struggled to accurately detect complex emotions and this leaded to the need to simplify the analysis in two steps. First, filtering the ”neutral” emotion to better highlight the results for other emotions and its gender differences. In this way, it was possible to observe an increase for both genders' accuracy, especially for happiness and sadness. However, male participants, in agreement with other similar researches, continued to exhibit lower accuracy in detecting negative emotions. Finally, a further simplification was introduced by classifying emotions into only two categories (”positive” and ”negative”). This provided a clearer picture improving the overall accuracy and highlighting gender differences in emotional perception, with females more likely to express emotions while males are more rigid, especially when it comes to negative emotions. All these results are in agreement with similar studies, discussed in the previous section. Moreover, several external factors contributed to the lower accuracy rates of the FER algorithm observed in this experiment, such as the weakness of visual stimuli (probably related to the phenomenon of emotional habituation) or a potential cultural bias related to a ”US-centric” images used as a stimulus for an Italian cohort. Additionally, a freezing phenomenon was observed in response to the strongest stimuli.

In conclusion, this thesis successfully demonstrated that Facial Emotion Recognition (FER) can serve as a reliable biomarker for quantifying emotional responses, by developing and validating an experimental pipeline for emotion measurement, combining self-reported emotions with detected facial emotions. In particular, happiness was consistently the most accurate emotion detected by the FER algorithm; this suggests that positive emotions might be more easily identified by FER systems rather than negative or more complex emotional states. However, this is only a first step of a much larger and more ambitious project. Indeed, an integration of the online task with the offline FER algorithm solidly validated, using different testing datasets, would significantly expand the potential for several applications across various fields. For instance, this could enable real-time monitoring of emotional states, helping clinicians to detect signs of emotional disorders (such as depression or anxiety) or providing teachers a tool to evaluate student's engagement during a learning activity. Furthermore, FER systems can guarantee transparency and privacy, identifying a potential misuse of this technology that would lead to a dangerous misperception of the objective reality through an emotional manipulation (for instance for marketing, politics, or media reasons). The balance between technological advancement and ethical responsibility is the key to unlock the full potential of FER systems across various fields.

Acknowledgments and Dedications

Of course, I thank my family—my mom, my dad, and my brother—for always being my stronghold, my safe haven in times of challenge.

I would like to once again thank Professors Mattia Veronese and Andrea Loreggia, who believed in this research project from the very beginning, helping me to shape this first step, scaling down my overly ambitious initial idea without ever diminishing my work or my vision.

A special thanks to all my dear friends, who make me feel valued every day, offering me unconditional support, for which I am deeply grateful.

I would also like to thank the NONA squadra of Croce Verde APM for teaching me how to handle both high-intensity emergency situations and quieter, yet equally important, moments.

Thank you to casaserpi, the most beautiful home in the world, made special by the people who fill every moment with joy and craziness.

And another big thank you to the Murga di Padova di Padova for teaching me to express myself in alternative ways, but above all, for empowering me to act instead of remaining silent.

To all the precious people I have met along this journey, even if our paths have diverged, thank you very much.

I am deeply touched by the overwhelming love and affection I have received over the years, and I hope I can repay even a small part of it.

Finally, I dedicate this thesis to the real victims of the ”manipulation and misperception” discussed within its pages. I send a thought to all the victims of current conflicts, rendered invisible by our media, but sustained through our money, our forces, our resources, and, most of all, through our silence. I hope that one day a sophisticated engineering tool will allow us to recognize a genocide, a war crime, gender violence, or an abuse of power. Or maybe, we just need a little more humanity.

May they all rest in peace.

Bibliography

[1] Hornung, E. (1999). *The Ancient Egyptian Books of the Afterlife*. Cornell University Press.

[2] Nussbaum, M. (2001). *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge University Press.

[3] Ames, R.T., & Hall, D.L. (2003). *Dao De Jing: A Philosophical Translation*. Ballantine Books.

[4] Knuuttila, S. (2004). *Emotions in Ancient and Medieval Philosophy*. Oxford University Press.

[5] Gutas, D. (2001). *Avicenna and the Aristotelian Tradition*. Brill.

[6] Descartes, R. (1649). *The Passions of the Soul*. Hackett Publishing Company.

[7] Hume, D. (1739). *A Treatise of Human Nature*. Clarendon Press.

[8] Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. John Murray.

[9] James, W. (1884). What is an Emotion? *Mind*, 9(34), 188-205.

[10] Lange, C. (1885). *Om Sindsbevægelser: Et Psyko-fysiologisk Studie*. Jacob Lunds Forlag.

[11] Freud, S. (1923). *The Ego and the Id*. W.W. Norton & Company.

[12] Cannon, W. B. (1927). The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *American Journal of Psychology.*

[13] Fulton, J. F. (1935). The Prefrontal Cortex: Its Relations to Emotion. *Psychosomatic Medicine.*

[14] Papez, J. W. (1937). A Proposed Mechanism of Emotion. *Archives of Neurology & Psychiatry.*

[15] LeDoux, J. E. (2000). Emotion Circuits in the Brain.

[16] Descartes, R. (1641). *Meditations on First Philosophy.*

[17] Bunge, M. (2010). *Matter and Mind: A Philosophical Inquiry.*

[18] Pylyshyn, Z. (1984). *Computation and Cognition: Toward a Foundation for Cognitive Science.*

[19] Kihlstrom, J. F. (1987). *The Cognitive Unconscious.*

[20] Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.*

[21] Davis, M. (1992). The role of the amygdala in conditioned fear.

[22] Neisser, U. (1967). *Cognitive Psychology.* Appleton-Century-Crofts.

[23] Gardner, H. (1987). *The Mind's New Science: A History of the Cognitive Revolution.* Basic Books.

[24] MacLean, P. D. (1949). Psychosomatic Disease and the "Visceral Brain".

[25] MacLean, P. D. (1952). Some Psychiatric Implications of Physiological Studies on the Limbic System.

[26] LeDoux, J. E. (1987). Emotion: Clues from the Brain.

[27] LeDoux, J. E. (1991). Neural Systems Underlying Emotions.

[28] Bargh, J. A., & Chartrand, T. L. (1999). The Unbearable Automaticity of Being.

[29] Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions.*

[30] LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life.*

[31] Pessoa, L. (2008). On the relationship between emotion and cognition.

[32] Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A. S., McNamara, J. O., & Williams, S. M. (2012). *Neuroscience.*

[33] Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2013). *Principles of Neural Science.*

[34] Bayram, A., Karahan, E., Bilgiç, B., Ademoglu, A., & Demiralp, T. (2014). Achromatic temporal-frequency responses of human lateral geniculate nucleus and primary visual cortex.

[35] Bechara, A., Damasio, H., & Damasio, A. R. (1995). Amygdala damage and decision-making deficits in humans.

[36] Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning.

[37] Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance.

[38] Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain.

[39] Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention.

[40] Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception.

[41] De Houwer, J., & Hermans, D. (2010). *Cognition and Emotion: Reviews of Current Research and Theories.*

[42] Gross, J. J. (2014). Emotion regulation: Conceptual and practical issues.

[43] Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence.

[44] Mesquita, B., & Walker, R. (2003). Cultural differences in emotions: A context for interpreting emotional experiences.

[45] Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States.

[46] Norris, C. J., Larsen, J. T., & Cacioppo, J. T. (2011). Neuroimaging of emotionally evocative pictures: The negativity bias in the amygdala.

[47] Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making.

[48] Murray, C., Manrai, A. K., & Manrai, L. A. (2018). The influence of emotions on consumer behavior.

[49] Brader, T. (2006). *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work*.

[50] Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization.

[51] Lodge, M., & Taber, C. S. (2013). *The Rationalizing Voter*.

[52] Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization.

[53] Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment.

[54] Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood.

[55] Likert, R. (1932). A technique for the measurement of attitudes.

[56] Diener, E., Larsen, R. J., Levine, S., & Emmons, R. A. (1985). Intensity and frequency: Dimensions underlying positive and negative affect.

[57] Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*.

[58] Bradburn, N. M. (1969). *The Structure of Psychological Well-Being*.

[59] Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion.

[60] Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications.

[61] Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential.

[62] James, W. (1884). What is an emotion?

[63] Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory.

[64] Jung, C. G. (1906). Studies in Word Association.

[65] Lykken, D. T. (1959). The GSR in the detection of guilt.

[66] Hess, E. H. (1965). Attitude and pupil size.

[67] Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS).*

[68] Darwin, C. (1872). *The Expression of the Emotions in Man and Animals.*

[69] Cohn, J. F., & Ekman, P. (2005). Measuring facial action by manual coding, facial EMG, and automatic facial image analysis.

[70] Mehrabian, A. (1968). Communication without words.

[71] Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research.

[72] McDuff, D., Gontarek, S., & Picard, R. W. (2016). Remote measurement of cognitive stress via heart rate variability.

[73] Healey, J., & Picard, R. W. (2005). Detecting Stress During Real-World Driving Tasks Using Physiological Sensors.

[74] Picard, R. W., & Healey, J. (1997). Affective Wearables.

[75] Setz, C., et al. (2010). Discriminating Stress from Cognitive Load using a Wearable EDA Device.

[76] Bradley, M. M., et al. (2008). Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing.

[77] Karg, M., Kühnlenz, K., & Buss, M. (2013). Recognition of Affect Based on Gait Patterns.

[78] Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions.

[79] Salah, A. A., Gevers, T., Sebe, N., & Vinciarelli, A. (2011). Challenges of Human Behavior Understanding.

[80] Cowie, R., et al. (2001). Emotion Recognition in Human-Computer Interaction.

[81] Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis.

[82] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*.

[83] Ekman, P. (1993). Facial expression and emotion.

[84] Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications.

[85] Matsumoto, D., & Hwang, H. C. (2011). Evidence for universal facial expressions of emotion.

[86] Mandryk, R. L., & Inkpen, K. M. (2004). Physiological indicators for the evaluation of co-located collaborative play.

[87] Lewinski, P., den Uyl, T., & Butler, C. (2014). Automated Facial Coding: Validation of Basic Emotions and FACS AUs in FaceReader.

[88] Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. (2017). Facial expression analysis with AFFDEX and FACET: A validation study.

[89] Mollahosseini, A., Chan, D., & Mahoor, M. H. (2017). Going deeper in facial expression recognition using deep neural networks.

[90] Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*.

[91] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features.

[92] Shih-Chung Hsu, Hsin-Hui Huang, Chung-Lin Huang (2017). Facial Expression Recognition for Human-Robot Interaction.

[93] Fuzail Khan (2018). Facial Expression Recognition using Facial Landmark Detection and Feature Extraction via Neural Networks.

[94] Sariyanidi, E., Gunes, H., & Cavallaro, A. (2015). Automatic analysis of facial affect: A survey of registration, representation, and recognition.

[95] Bartlett, M. S., Littlewort, G., Frank, M. G., Lainscsek, C., Fasel, I., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions.

[96] Arriaga, O., Valdenegro-Toro, M., & Plöger, P. (2017). Real-time convolutional neural networks for emotion and gender classification.

[97] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2016). AffectNet: A database for facial expression, valence, and arousal computing in the wild.

[98] Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International Affective Picture System (IAPS): Technical manual and affective ratings.

[99] Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the International Affective Picture System.

[100] Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual.

[101] Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database.

[102] Balsamo, M., Carlucci, L., Padulo, C., Perfetti, B., & Fairfield, B. (2020). A Bottom-Up Validation of the IAPS, GAPED, and NAPS Affective Picture Databases: Differential Effects on Behavioral Performance.

[103] Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS).

[104] Redies, C., Grebenkina, M., Mohseni, M., & Kaduhm, A. (2020). Global Image Properties Predict Ratings of Affective Pictures.

[105] Jović, A., & Burnik, K. (2022). Investigation of Relationships between Discrete and Dimensional Emotion Models in Affective Picture Databases Using Unsupervised Machine Learning.

[106] Russell, J. A. (1980). A Circumplex Model of Affect.

[107] Sieun An, Li-Jun Ji, Michael Marks, & Zhiyong Zhang. (2017). Two sides of emotion: Exploring positivity and negativity in six basic emotions across cultures.

[108] La Monica, L., Keller, F., Cenerini, C., Vollero, L., Pennazza, G., & Santonico, M. (2023). Development of a Universal Validation Protocol and an Open-Source Database for Multi-Contextual Facial Expression Recognition.

[109] Verma, A., Malla, D., Choudhary, A. K., & Arora, V. (2019). A Detailed Study of Azure Platform & Its Cognitive Services.

[110] McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. E. (2016). AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit.

[111] Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial Expression Analysis with AFFDEX and FACET: A Validation Study.

[112] Magdin, M., Benko, L., & Koprda, Š. (2019). A Case Study of Facial Emotion Classification Using AFFDEX.

[113] Justadudewhohacks. JavaScript API for Face Detection and Face Recognition in the Browser and Node.js with Tensorflow.js. Available online: https://github.com/justadudewhohacks/face-api.js (accessed October 10, 2024).

[114] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications.

[115] Reynolds, R. M., Novotny, E., Lee, J., Roth, D., & Bente, G. (2019). Ambiguous Bodies: The Role of Displayed Arousal in Emotion [Mis]Perception.

[116] Everaert, J., Koster, E. H. W., & Derakshan, N. (2012). The combined cognitive bias hypothesis in depression.

[117] Xu, Q., Ye, C., Gu, S., Hu, Z., Lei, Y., Li, X., Huang, L., & Liu, Q. (2021). Negative and Positive Bias for Emotional Faces: Evidence from the Attention and Working Memory Paradigms.

[118] Crane, E. A., & Gross, M. M. (2013). Effort-shape characteristics of emotion-related body movement.

[119] Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: A meta-analysis of findings from neuroimaging.

[120] Brody, L. R., & Hall, J. A. (2008). Gender and emotion in context: Expressive accuracy, emotion regulation, and perception.

[121] Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men.

[122] Martinez, B., Valstar, M., Jiang, B., & Pantic, M. (2017). Automatic analysis of facial actions: A survey.

[123] McDuff, D., El Kaliouby, R., Cohn, J. F., & Picard, R. W. (2015). Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads.

[124] Kulke, L., Feyerabend, D., & Schacht, A. (2020). A comparison of the affectiva iMotions facial expression analysis software with EMG for identifying facial expressions of emotion.

[125] Kosti, R., Alvarez, J. M., Recasens, A., & Lapedriza, A. (2019). Emotion recognition in context.

[126] Cowen, A. S., Sauter, D. A., Tracy, J. L., & Keltner, D. (2019). Mapping the passions: Toward a high-dimensional taxonomy of emotional experience and expression.

[127] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity.

[128] Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences.

[129] Potter, R. F., & Bolls, P. D. (2012). *Psychophysiological Measurement and Meaning: Cognitive and Emotional Processing of Media*.

[130] Furl, N., Gallagher, S., & Averbeck, B. B. (2012). A Selective Emotional Decision-Making Bias Elicited by Facial Expressions.

[131] Wirth, W., & Schramm, H. (2005). Media and Emotions.

[132] Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-Cultural Elements in Facial Displays of Emotion.

[133] Mesquita, B. (2001). Culture and Emotion: Different Approaches to the Question.

[134] Roelofs, K. (2017). Freeze for Action: Neurobiological Mechanisms in Animal and Human Freezing.

[135] Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?

[136] European Commission. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council on Artificial Intelligence (Artificial Intelligence Act).

[137] Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S., & Rosenwald, D. P. (2013). Social risk and depression: Evidence from manual and automatic facial expression analysis.

[138] D'Mello, S., & Graesser, A. (2012). Emotions during learning with AutoTutor.

[139] Dhall, A., Goecke, R., Gedeon, T., & Sebe, N. (2016). Emotion recognition in the wild.