Face perception foundations for pattern recognition algorithms

FACE PERCEPTION FOUNDATIONS FOR PATTERN RECOGNITION ALGORITHMS

F. MARCOLIN, E. VEZZETTI
Department of management and Production Engineering
Politecnico di Torino, corso Duca degli Abruzzi 24 - 10129 Torino, Italy

M.G. MONACI
Department of Human and Social Sciences
Università della Valle d'Aosta, Strada Cappuccini 2A - 11100 Aosta, Italy

ABSTRACT

Pattern recognition system developers have looked in multiple directions over the years and designed a broad spectrum of methodologies for face identification and verification, both in 2D and 3D. These techniques rely on sound methods and experimentations, and currently give high to excellent recognition rates in terms of performance. Nonetheless, it seems that the most performing face recognition system, especially when familiar faces are involved, is still the human being, able to detect known faces in the wild, in presence of occlusions or extreme light contrast, caricatures, sketches, partial views, blurred images. This is one of the manifold reasons why the human visual system at eye and brain level and face perception techniques are currently being studied by neuroscientists and psychologists, with the aim to uncover the processes underneath the human vision.

The purpose of this work is to review the current literature about perception foundations and related biologically-inspired methodologies for face recognition.

KEYWORDS

Face recognition; face perception; face expression recognition; human vision; visual system; biologically-inspired computing.

1 INTRODUCTION

From a computational viewpoint, face recognition (FR) process is defined as: given still or video images of a scene, identify one or more persons in the scene using a stored database of faces [1]. Human face recognition (FR) is somehow the same, and the stored database of faces is in our brain, with different degrees of knowledge depending on whether a face belongs to a friend, a famous person, or a person we know by sight.

Facial expression recognition (FER) regards the identification of the emotion felt by a person thanks to the codification of the muscle movements of his/her face. Algorithms try to automatically perform this task with the aim to emulate an innate process of our brains. From an anthropologic perspective, the

study of the expression of emotions has its roots in Darwin's studies [2] and, before him, in Charles Bell's [3], about the anatomy of facial expressions.

The process underlying human face perception and recognition has being studied from the Thirties from different perspectives to understand what happens at eye and brain level when we identify a face as belonging to a specific subject or a person's emotion thanks to the movements of the face. Even if the current modern technologies have allowed to understand more in depth how this process works, several fundamental issues on human face perception (FP) are still unresolved [4]. The main research veins of face perception are briefly presented here below.

The first consideration regards how faces are processed. Two trains of thoughts have been developed. The first and mainly acknowledged is the one assessing the face perception is a **holistic** (or "whole") process. Faces are stored in a relatively holistic form in memory, i.e. faces are initially encoded holistically during perception and perceived holistically [5], meaning that a face is perceived as a non decomposable whole, at a single glance, rather than as a collection of individual features [6]. In other words, face processing is said to be distinct from non-face object in that it is more holistic; that is, faces are represented as non-decomposed wholes rather than as a combination of independently-represented component parts (eyes, nose, mouth), and the relations between them [7]. This theory has been endorsed and partially supported by different researchers [8] [4] [9] [8] [10] [11]. For instance, Van Belle et al. saw that prosopagnosics, i.e. patients affected by a damage in the right brain hemisphere, process feature-by-feature, while normal people process holistically, thus proving that "standard" face processing is holistic [6]. Other studies even revealed that holistic/configural processing is limited to faces [12].

Another evidence is given by the face inversion effect [13]. Recognizing an upside-down face is not trivial for us, because we use a super quick look to recognize a face [14], which works only when the face is in an upward position. This pattern of results takes many years to develop (six-year-old children are not affected by inversion, ten-year-olds exhibit near adult-like performance) [10]. Recognizing an upside-down face is not impossible though; it requires training, which starts with learning how to process a face feature by feature. The so-called Thatcher illusion is related to the face inversion effect. In 1980 Thompson showed that, with the inversion of the facial features (of Margaret Thatcher's face, for instance), we do not just process individual traits, but also the positions and relationships between them [15]. Tanaka and Gordon listed the single effects of face inversion: affects configural coding of the vertical position of eyes/eyebrows; affects vertical spacing between nose and mouth features; affects featural coding of size and shape of the mouth; though, does not perturb the featural perception of eye shape and size; leaves intact configural perception involving the horizontal spacing between the eyes. Thus, featural and configural information in the eye region is selectively spared during inversion, whereas information outside this area is severely impared [16]. The considerations on inversion support the fact the face is processed as a whole.

The alternative **local** face processing theory has been cited by Ellis [14] and consists in an analytic view, implemented feature by feature. It is the way prosopagnosia-affected people [6], passport officers, and children use to identify faces. In particular, the training path undertaken by passport officers relies on a feature-by-feature approach [17]. Similarly, the visual system progresses from piecemeal (local) to holistic in the first several years of life [10], meaning that for children under ten years of age, the process is feature-based, then it changes into holistic [1]. Another relevant contribution in the "debate" which leans away from the holistic approach is given by Schyns's group. They stress the importance of specific facial features for FER (featural approach), in particular with the assertion that facial information tends to be integrated from the eyes downward in the face [18]. "Information processing gradually shifts from encoding common face features across all spatial scales (such as the eyes) to representing only the finer scales of the diagnostic features that are richer in useful information for behavior (e.g. the wide opened eyes in 'fear' expression)" [19]. Their studies also report that facial features are first processed in isolation in extrastriate regions of the brain, then in combinations in the occipito-temporal regions, thus suggesting that both featural and holistic processings happen, at different times [20]. The role of specific facial features will be reprised in section 2.1.4 regarding the hierarchy of them.

Overall, we can say that human FR generally is a holistic process involving an interdependency between featural and configural information [10]. Face is of utmost importance, as our visual system starts with a rudimentary preference for face-like patterns [4] [10], even if our visual attention is not captured by fake faces (pareidolia faces) [21]. The reason of the weight given to faces by our brain is that body structure and gait information are much less useful for identification than facial information [10].

Another issue in the field which supports the understanding of human facial perception and recognition is the study of performances of human FR on unfamiliar faces, with the aim of answering to the question "Is human face recognition for **unfamiliar faces** poor?" [8]. Recent results have showed that we are not excellent in recognizing faces which do not belong to people we know, even if the recognition is made by trained staff [22]. Also, it has been shown that this ability cannot be improved with experience [10], but is innate in some people. In other words, some people are particularly gifted in recognizing unfamiliar faces, while others are not, even if they undergo specific trainings [17]. Generally speaking, unfamiliar faces are more difficult to be recognized; recognition performances are not 100%, which instead is the recognition rate for familiar faces. Indeed, the ability to tolerate facial image degradation increases with familiarity [4] [10]. Structural codes, which capture those aspects of the structure of a face essential to distinguish it from other faces, for familiar faces differ from those formed to unfamiliar faces; structural codes for familiar faces emphasize the more informative and less changeable (hairstyle) regions of the face [23]. Some observations have also arisen regarding the different types of familiarity (friends, famous individuals, close relatives such as parents), with different levels of responses in the brain [24]. As a consequence, to understand the key difference between familiar and unfamiliar faces, and how differently they are managed by the brain is core for the research [22].

These core considerations help framing and understanding the vast topic of human FP and FR, and are used here to define and tune the scenario. The aim of this work is to offer a thorough survey about human FP and on how visual systems dynamics have been applied to automatic face and facial expression recognition methodologies in the computer vision and pattern recognition field. Sinha et al. in 2006 collected 19 fundamental results on FP with the hope that computer vision researchers could take advantage of them for implementing their algorithms [10]. The present work, with a similar purpose, offers the step forward, i.e., an up-to-date review about significant FP principles and on how biological face recognitions dynamics have been incorporated into automatic methodologies in these years, showing that these fields are interlinked and there is still room for inspiration and further investigation to connect them more deeply. This is somehow urgent, considering that various contributions emerged recently, thanks to the advances in technology, describing the brain functioning in specific tasks.

This study is written for readers working on face analysis, whether they be scientists or nonscientists, who do not need to have a deep knowledge of the various theories, debates and broad range of findings in the science of face and facial expressions, with sufficient pointers to those discussions if they are of interest. (Barrett et al. aimed at such an effect when they discussed the challenges to inferring emotions from human facial movements [25].) Also, it is important to state that it is out of the scope of this study to review all algorithms which are biologically plausible or replicating general brain dynamics; only those inspired by human face processing are taken into consideration.

The main foundations of face perception and processes of the human visual system are presented and introduced in the following section. Then, biologically-inspired face recognition algorithms are presented. Lastly, a discussion and future research hints conclude the work.

2 MAIN FOUNDATIONS OF FACE PERCEPTION

Face perception is an individual's understanding and interpretation of the face, particularly the human face, especially in relation to the associated information processing in the brain [26]. The face perception research field includes studies based on psychology/behaviour and neuro-imaging. The main findings concerning recognition are summed-up here. A detailed review of relevant studies on neuroscience and psychophysics is beyond the scope of this paper. We mainly summarize findings which are potentially usable for the design of face recognition systems, similarly to Chellappa et al.' study in 1995 [1]. Section 2.1 exposes the objective and impartial face perception outcomes, which do not depend on the recognizing person; section 2.2 briefly overviews the behavioural and subjective/individual aspects of human FR.

2.1 EXTERNAL DYNAMICS

Research on facial perception processes has demonstrated that there are patterns which belong, without distinction, to all human beings. These common processes mainly concern the weight given to the facial features (eyes, nose,…), the ability to recognize caricatures and "normal" faces, and the manipulation to the facial images in terms of definition and colour. There are also two very popular research questions that still have no final answer. This section presents, firstly, these two issues, then extrapolates the main findings regarding typical recognition mechanisms, regardless of the individual and his/her personality and characteristics.

2.1.1 Is face recognition a special recognition process for the brain?

The first issue arising from psychologists is to determine whether or not **face recognition is a special recognition process for the brain** [14] [22] [27]. Is human face a special object for us? If yes, how is it processed differently from other objects? Are there dedicated neurons for that? Is there a specific look sequence of our eyes to scan a face for recognition purpose? These questions are motivated by the fact that faces are more easily remembered by humans than other objects when presented in an upright position [1].

Ellis was the first addressing this issue, in 1975; he supported both the 'yes' and 'no' theories relying on meaningful experimentations. Ellis first studied the recognition abilities of prosopagnosia-affected patients. Prosopagnosia is a cognitive disorder caused by lesions in various parts of the inferior occipital areas (occipital face area, OFA), fusiform gyrus (fusiform face area, FFA), and the anterior temporal cortex, which are areas activated specifically in response to face stimuli [28]. The problems affecting these patients in recognizing other faces (even their own) and the fact that they are normally perfectly able to identify other objects seem to demonstrate that there is an area of the brain dedicated to face processing. Other studies support this thesis relying on experimentations on prosopagnosia [29] or on patients with brain damage at the right hemisphere [1].

Ellis also observed that babies are able to recognize their mother at two weeks age. Chellappa et al. argued that infants come into the world pre-wired to be attracted by faces [1]. For filial and sexual imprinting reasons [30], physiognomic details are innate in humans, who result in an innate proclivity for analyzing faces; this phenomenon is identified as "ontogeny of interest in faces" [14] or "innate knowledge" of faces [31]. Also, what makes face processing special is that it is gated by an obligatory detection process [7].

Other arguments concern the already mentioned holistic dynamics in processing faces. For faces we are global/holistic; for other objects we are analytical and detailed [14], namely face processing is said to be distinct from non-face object in that it is more holistic [7]. A shrewd view is mentioned by Tanaka

and Gordon; FR “is not distinctive because it depends exclusively on holistic representations. It is only distinguished because its representation is disproportionately more holistic than the representation of most other types of objects” [16].

The last evidence of the 'yes' thesis is given by neuroscience. Different research studies report that there is a large number of neurons devoted to the class of face stimuli [32] [4] [10] [33]. In particular, "the principles underlying the processing of faces and other objects may be similar, but more neurons may be allocated to represent different aspects of faces because of the need to recognize faces of many different individuals, that is to identify many individuals within the category of faces" [32]. Chellappa et al. even assert the existence of a "grandmother" neuron for face recognition [1]. Furthermore, the existence of face recognition cells (encoding precise face templates) and face detector cells (detection, segmentation, alignment) has been affirmed [7]. Finally, specific studies answering to the research question Are faces special? conclude by endorsing the ‘yes’ thesis by focusing on the within-class discrimination hypothesis and the derived expertise hypothesis [12]. This hypothesis proposes that face-specific mechanisms are engaged by individual-level recognition of objects, just as they are by individual-level recognition of faces. Evidence argued that neither within-class discrimination hypothesis nor individual–level categorization with extensive expertise leads to face-like processing for objects [12].

The 'no' theory is not very popular. In 1975, Ellis stated that the 'yes' theory was actually impossible to be proven, and this was one of the reasons why the 'no' theory could have been taken up. He also saw that prosopagnosia is associated to other deficits in some cases [14], thus this defect cannot be fully adopted to study human face perception/recognition processes. Table 1 provides a sum-up about the two theories.

In the years, this landmark study of Ellis has been contested because experiments at that time have been done on photographs. Thus, Ellis's considerations could be considered valuable for picture recognition field and not specifically for face recognition [22].

At the moment, the majority of sources, including general informative sources, agrees with the 'yes' theory. Studies based on neuropsychology, behaviour, electrophysiology, and neuro-imaging have supported the notion of a specialized mechanism for perceiving faces [26]. The characterization of the FFA as a dedicated face processing module appears very strong [10] [7], and specialized face processing mechanisms in the human visual system are very real possibility [10]. Nonetheless, Haxby and Gobbini affirmed that the two perspectives are not necessarily incompatible [24].

Table 1. Arguments and research outcomes supporting the 'yes' and 'no' theories about whether FR is a special recognition for the brain or not.

Is human face recognition a special recognition process for the brain? [14] [22] Critic: Experiments at Ellis time have been done on photographs. Thus, it may not be FR, but picture recognition [22]	
YES	**NO**
Prosopagnosia considerations	
Impairments in FR [14] [1] [7] Prosopagnosia is not necessarily associated to other defects [29]	In some cases the prosopagnosia is associated to other deficits [14]
Behavioural considerations	
Babies recognize their mother at two weeks age [14] It is argued that infants come into the world pre-wired to be attracted by faces [1] For filial and sexual imprinting reasons, physiognomic details are innate in humans [14]: • innate proclivity for analyzing faces • ontogeny of interest in faces What makes face processing special is that it is gated by an obligatory detection process [7]	
Face processing considerations	
For faces we are global/holistic; for other objects we are analytical & detailed [14] Face processing is said to be distinct from non-face object in that it is more holistic [7]	
Neuroscientific considerations	
There are specialized neural resources devoted to face perception [4] [10] [33] FFA is activated specifically by faces [7] The characterization of the FFA as a dedicated face processing module appears very strong [10] There is a large number of neurons devoted to the class of face stimuli [32] In particular, "the principles underlying the processing of faces and other objects may be similar, but more neurons may become allocated to represent different aspects of faces because of the need to recognize faces of many different individuals, that is to identify many individuals within the category of faces." [32] There might exist a "grandmother" neuron for face recognition [1] Face recognition cells (encoding precise face templates) and face detector cells (detection, segmentation, alignment) exist [7] Within-class discrimination hypothesis and the derived expertise hypothesis [12]	
Other / generic considerations	
Faces are more easily remembered by humans than other objects when presented in an upright position [1] Considerations about face inversion effect show that human FR is a face-specific process [13] Studies based on neuropsychology, behaviour, electrophysiology, and neuro-imaging have supported the notion of a specialized mechanism for perceiving faces [26].	YES is impossible to be proved [14]

2.1.2 Are face recognition and face expression recognition two separate processes?

A key issue about face perception is to understand whether **face recognition and face expression recognition are two separate processes or not**. Pattern recognition and computer vision researchers work for developing algorithms which are able to process the query face and to understand both to which person of a gallery dataset belong and which expression that face is actually displaying. Generally, face and face expression recognition (FR and FER, respectively) algorithms are separate and rely on different methodologies. The human visual system may work similarly [24]. However, in the psychology field, considerations have arisen about the fact that some characteristic expressions belong to specific subjects, implying that the presence of a typical facial emotion could support the recognition of an individual. Similarly to the previous topic, the 'yes' and 'no' theories have been supported by different research groups.

The first observation supporting the first theory regards prosopagnosia. Some prosopagnostic patients, who have difficulties in identifying familiar faces, recognize emotional expressions. Patients suffering from 'organic brain syndrome' work the contrary [1]. The psychological perspective assesses that a familiar face is represented via an interlinked set of expression-independent structural codes for distinct head angles [23], except for characteristic expressions. From a physical viewpoint, studies affirm that FER and FR are separate processes, because they are managed by different brain regions/systems [29] [4] [10], even if a clear-cut separation between regions has not been quantified yet. In particular, the superior temporal sulcus manages changeable aspects of face (FER) and fusiform gyrus (inferior temporal) deals with identity (FR) [7] [33]. The fMRI works of Kanwisher's group support this thesis. They showed that the OFA and the FFA parts are more involved in recognition of individual identity, whereas fSTS (a face-selective region of the superior temporal sulcus, STS) deals with the recognition of social information in faces, i.e. eye gaze direction and expressions [28]. Also, some neurons have responses which reflect the facial expression but not the facial identity of the stimulus [32].

The 'no' theory has been firstly supported by studies which affirm that facial expressions may affect human recognition rates (HRR) [14], as many people have characteristic expressions which help the recognition of their identity. This thesis, supported by other studies [34], is endorsed by some brain and neuron considerations. Haxby et al. assert that a region of the brain (fusiform gyrus) may support both FER and FR because some expressions are associated to specific persons [29]. Sinha et al. also proposed the "identity expression" concept, affirming that a small subset of neurons respond both to identity and emotion [10]. Table 2 outlines the two theories and relative supporting studies.

By taking into consideration and examining two models supporting respectively the 'yes' and 'no' theories, Calder proposed an hybrid view asserting that at least certain aspects of facial identity and facial expression recognition involve the same visual route. The suggested framework "incorporates a ventral temporal route (including the fusiform gyrus) involved in the analysis of visual form associated with both facial identity and facial expression". Then, dynamic facial information is coded by a separate route including the STS [35].

Evidence of correlation between emotions and recognition comes also from psychological and neuroscientific considerations. Increasing psychosocial evidence supports that recognition memory for faces with both negative and positive emotional expressions is more accurate than for neutral faces. In one of the first studies, Galper e Hochberg [36] presented to their participants twice the same face but with different facial expressions and, in the memory task, recognition was better when the alternative emotional expression was presented immediately after compared to the identical face presented 5 days after. The emotional expressions seem to have a facilitatory effect on FR enhancing face processing, albeit is still under debate if more for positive or negative compared to neutral expressions [37]. Negative emotions may attract more attention because, from an evolutionary and adaptive point of view, are more threatening and dangerous; while, according to other studies [38], people process happy faces more than negative ones because they are more pleasurable. FR after incidental learning (asking to participants a gender discrimination task for each faces with different emotional expression) resulted

significantly better for negative faces, particularly fearful faces [39]. Examining, through neuroimaging, the neural correlates of labelling emotional expressive or neutral faces, Jehna et al. [40] saw that participants had more difficulties in identifying emotionally expressive faces (anger, fear and disgust) than neutral faces. With neutral faces and other stimuli (eg. buildings) similar brain zones were activated, that is the fusiform occipital temporal gyrus, the lateral occipital cortex and the lower frontal gyrus. On the contrary, images of emotionally expressive faces, opposed to neutral faces, activated different brain areas.

Table 2. Arguments and research outcomes supporting the 'yes' and 'no' theories about whether FER and FR are two separate processes or not.

Are FR and FER separate processes?	
YES	**NO**
Some prosopagnostic patients, who have difficulties in identifying familiar faces, recognize emotional expressions. Patients suffering from 'organic brain syndrome' work the contrary [1] A familiar face is represented via an interlinked set of expression-independent structural codes for distinct head angles [23] except for characteristic expressions FER and FR are separate processes because they are managed by different brain regions [29] even if the separation between regions has not been quantified yet FER and FR might be processed by separate systems [4] [10] Some neurons have responses which reflect the facial expression but not the facial identity of the stimulus [32] The superior temporal sulcus manages changeable aspects of face (FER) and fusiform gyrus deals with identity [7] [33] [28]	Different FEs may affect HRR [14] A region of the brain (fusiform gyrus) may support both FER and FR because some expressions are associated to specific persons [29] FER and FR are not separate processes [34] A small subset of neurons respond both to identity and emotion; concept of "identity expression" [10]

2.1.3 Best facial half

It is well known that faces are not perfectly symmetric. The two halves differ for details in normal faces. Thus, a few studies have been undertaken about which half is the one we memorize more in other people's faces. Ellis dealt with this aspect and asserted that we better memorize faces on the left side (effective right side) of people [14]. In other words, experimentations showed that if we reconstruct a face by mirroring the left half and the right half, the left side mirroring reconstruction is recognized better. The reason is that the left vision field arrives initially at the right hemisphere, which has an advantage in the reception and storage of facial information acquired by the visual system [1].

2.1.4 Facial features hierarchy

Facial features have different levels of importance for recognition. Most of the studies agree on the fact that the upper part of the face is more important recognition-wise [1]. The hierarchy proposed by Ellis in 1975 [14], given by 1) **eyes** at first position, then 2) **nose**, 3) **mouth**, 4) **lips/chin**, and 5) **ears**, has been endorsed by Zhan et al. in 2009 [8], and completed by **age**, **facial shape**, and **hair length** components by Davies and Young [22]. **Eyebrows** have been mentioned by Chellappa et al. as key

elements for recognition, for they are high-contrast and large facial features, less susceptible to shadow and illumination changes [4]. Also, their role is core as they convey emotions [10]. According to Chellappa et al., nose plays an insignificant role due to the fact that most of the studies have been done using frontal images [1]. They also asserted that both external and internal features are important for human FR, but internal features are dominant in the recognition of familiar faces [1].

Specific methodologies (Bubbles technique or reverse correlation method) have been developed for understanding which facial information and facial parts are used by the brain in order to make a judgment and a categorization about the expression displayed by a face [41] [42]. Concerning emotion recognition, the studies of Ekman [43] [44] have revealed that every movement of facial muscles can be codified and is related to specific emotions. Hence, facial features and in particular mouth, eyes, and eyebrows have a significant role in the recognition of emotions via facial expressions, as they convey the majority of muscular movements. Cross-race recognition studies have partially endorsed this hypothesis. Apparently, recognizing emotions on other-race faces is more difficult, not only because of the viewers lessen experience but also because the scanning strategies in fixating own- and other-race faces are different [45], thus resulting in different emotion appraisal because different face regions provide visual information related to different categories of emotion (for example, the eye region is important for recognizing sad expressions, whereas the mouth for recognizing both happy and sad expressions [46]). Eye-tracking studies demonstrated, for instance, that Chinese fixated more on the eye regions of Caucasian faces than on eye regions of Chinese faces, while they fixated more on nose regions of Chinese faces than on that of Caucasian faces [45], and confirmed the differential emotion attribution to faces of different races.

2.1.5 Caricatures and normotype

Caricatures are sketches of people/faces in which the characteristic features of the subject are exaggerated. In particular, caricatures are created to exaggerate deviations in shape or both shape and pigmentation cues [10]. Experiments have shown that subjects in caricatures are often recognized more than in veridical representations, even when faces are unfamiliar, as "accuracy in recognition occurs where the facial detail differs from the exemplar" [8] [47] [10] [1] [7].

This is connected to the concepts of normotype and beauty. The facial normotype is a prototypical face where the facial traits are 'normal' and regular. It has been shown that faces far from the normotype are easier to be recognized, "perhaps because they reside in a less populated area of face space" [48], while typical faces are more easily confused with other faces [49]. In particular, beautiful faces are the easiest to be recognized, followed by the ugly and eventually by the standard ones [14]. This also asserts that the normotype is not necessarily associated to beauty [1], contrarily to the thesis of Zheng et al. affirming that attractive faces are more similar to the norm/average face [48]. As a completion of this topic, Bittner and Gold asserted that facial symmetry improves the impression of beauty but does not improve HRR [50].

2.1.6 Role of image quality, manipulation, and colours

Researchers investigated on whether human recognition rates may be affected by the quality of the facial image, its colours and manipulation. If images are out-of-focus or processed with a mean filter, for the purpose of smoothing, i.e. reducing the amount of intensity variation between one pixel and the next, we have **low spatial frequency** images. These images are blurred pictures in which all fine detail of features has been removed. Low spatial frequency bands play a dominant role in FR, as they contribute to the global description, while high frequency components contribute to the finer details required in the identification task [1]. Experiments show that we are able to distinguish very well celebrities' (and familiar) faces even in low spatial frequencies [23]. Also, even if faces are distorted and compressed, human FR does not decrease [10].

Other evidences have emerged about **colours** and **light**. Light is said to be more important in human FR than colours and pose [22] [10] [14]. Other sources assert colours, pigmentation, illumination are all significant [10]. In particular, pigmentation and shape cues are equally important, but with colours human FR is higher, as colour cues are specific to every individual [10]. Also, edges capture the most important aspects of images, i.e. the discontinuities, but photometric cues need to be added to have high human RRs [10].

The **contrast polarity inversion** has also been examined. Faces are particularly difficult to be recognized when viewed in reverse contrast, as in photographic negatives [51] [4] [10]. The curious aspect of this pattern is that no information is lost in negatives, i.e. photographies and its negatives contain the same amount of information.

2.1.7 At the brain level

The interest in face perception does not only regard results obtained via experimentations with subjects asked to recognize faces in different conditions, but also biological studies concerning the anatomy and the functioning of visual system at eye and brain level.

Much is known about the neural systems that subserve face recognition in adult humans and primates using functional magnetic resonance imaging (fMRI) [52] [53] [54] and electrophysiology [55] [56] [57]. The primary locus for human face processing is the lateral fusiform gyrus [29] and, in general, the FFA [10] [7], together with OFA, fSTS, AFP (anterior face patch), and regions in the frontal lobe and orbitofrontal cortex [56]. FFA deals with face classification [7], with the holistic processing of faces (low spatial frequency), and mainly serves facial identity recognition [28]. Prospagnosics have this part injured, thus they can only process faces feature-by-feature. The occipital face area (OFA) deals with face measurement [7] and with local face processing, i.e. with the identification of facial parts (high spatial frequency) [28]. Together with FFA, it is meant to support identity recognition via face processing. fSTS is a face-selective part of the superior temporal sulcus, which is being activated by eye gaze (although not from the physical properties of the eyes) and facial expressions [24] [28]. Specific considerations regarding the activation of these areas for eye gaze processing have been reported by Perlphrey and Wyk [58]. Figure 1 shows the brain regions involved in face processing according to the works of Kanwisher's [28] and Tsao's groups [56].

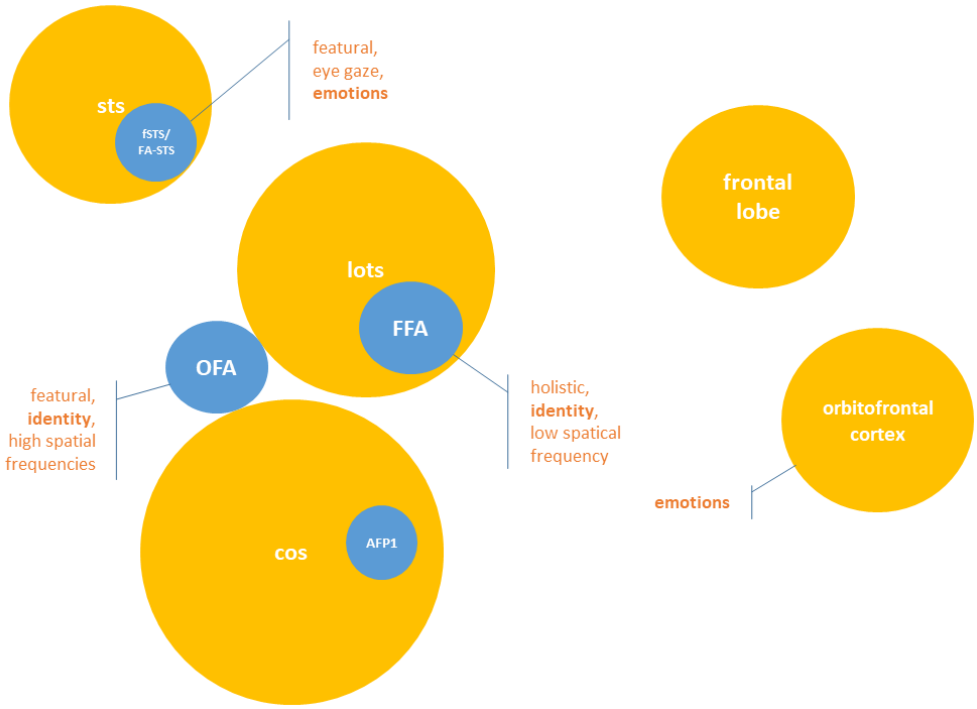


Figure 1. Face parts (or "face patches") in the human brain with respective tasks according to [28] and [56]. The scheme is just recapitulatory for brain parts and functions, not anatomical (for respective anatomical figures see [28] and [56], and related works). Abbreviations: superior temporal sulcus (sts), lateral occipito temporal sulcus (lots), collateral sulcus (cos), superior temporal sulcus face area (fSTS or FA-STS), fusiform face area (FFA), occipital face area (OFA), anterior face patch 1 (AFP1).

This model could be completed by the "general model of the distributed neural systems for face perception" proposed by Gobbini and Haxby [59], which identifies the face parts described in Figure 1 as the "core system" and complete the model with an "extended system" including the brain parts dealing with personal information and these involved in the emotions (like the amygdala).

To complete this framework, ERP components should be mentioned. Event-related brain potentials (ERPs) are measured brain responses to specific sensory, cognitive, or motor events. ERP studies have uncovered several components that are linked to different stages in FP, FR, and the processing of FER. The most studied ERP component is N170, which belongs to the family of visually evoked N1 components that are elicited over visual brain areas in response to most types of visual stimuli. "The N170 component has acquired its status as a face-sensitive component because ERP amplitudes elicited at occipitotemporal electrodes between 140 ms and 200 ms after stimulus onset are virtually always larger in response to faces than in response to non-face objects" [60]. Schweinberger assessed that N170 predominantly reflects the categorization of a stimulus as a face and deals with structural encoding of faces [61]. Other components have been labelled as face-sensitive: P100, reflecting top-down attentional processes related to FP; N250 and N250r, dealing with familiarity and repetitions [62]; N400 and N400f, regarding semantic information about people [61]. The study of these neurophysiological responses and of their connections with different aspects of face processing is encouraged by the authors working in the field. Schematically,

N170 is the face-sensitive component, reflects the categorization of a stimulus as a face and deals with structural encoding of faces [61];

P100 reflects top-down attentional processes related to FP [62];

N250 and N250r deal with familiarity and repetitions [62];

N400 and N400f regard semantic information about people [61].

Generally speaking, the literature offers good but different explanations about how the human neural system works for vision and perception [63] [29] [64] [65]. Figure 2 and Table 3 sum up the role of the main regions of the human visual system. Figure 3 shows a concise anatomical representation of the brain [66] and specifically of Brodmann's areas [67] involved in vision and face processing.

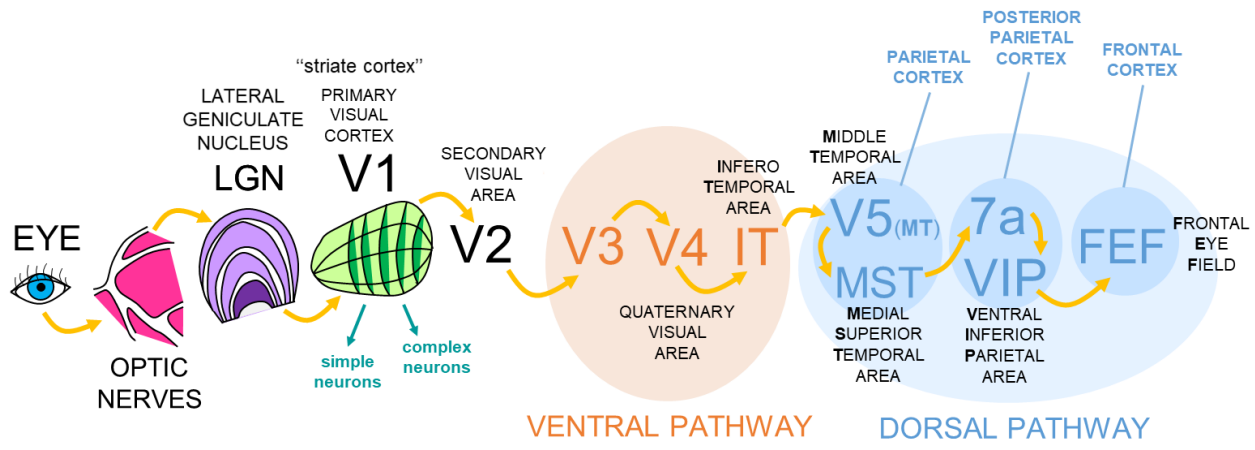
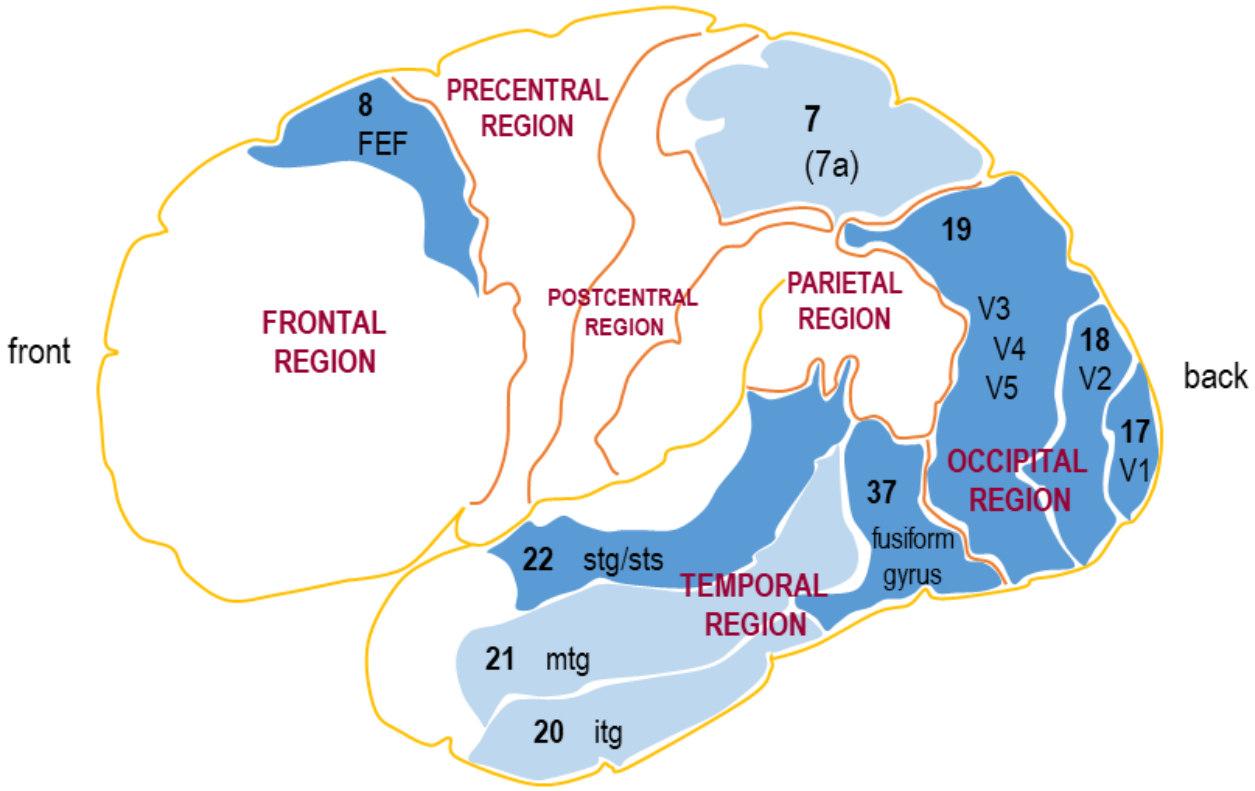


Figure 2. Scheme of the information flow in the Human Visual System according to [63] **and** [64]**.**

Table 3. Human Visual System basic structure and functioning according to [63] and [64]. The ventral pathway includes regions V3, V4, and IT. The dorsal pathway includes the parietal cortex, the posterior parietal cortex, and the frontal cortex.

Region		**Role**
Eye		Acquires visual signals
Optic Nerves		They run from an eye and lead to LGN [63]
Lateral Geniculate Nucleus (LGN)		Relays information from the optic nerve to V1
V1 / primary visual cortex / “striate cortex”		Receives information from LGN [63] Functions: Angular selectivity Detects oriented line Codes object contour orientation [63] Its neurons are simple and complex • Simple cells have receptive fields containing oriented subregions acting like edge filters • Complex cells respond primarly to oriented edges and gratings, with degree of spatial invariance [64]
V2 / secondary visual area		Receives information from V1 Topographic organization map of the visual field [63]
Ventral pathway		Receives information from V2 Visual memory and recognition with spatial localization [63]
V3		Preserves topographic organization Perception and localization of patterns [63]
V4 / quaternary visual area		Object discrimination Consciousness (sensible to anaesthesia!) [63]
Infero-Temporal area (IT) / Infero-Temporal cortex	Posterior Infero-Temporal area (PIT)	Operates in object recognition using visual memory [63]
	Anterior Infero-Temporal area (AIT)	
Dorsal pathway		Receives information from V2 Temporal treatments [63]
Parietal cortex	V5 / Middle Temporal area (MT)	Detect movements by angular direction [63]
	Medial Superior Temporal area (MST)	
Posterior parietal cortex	7a	Visual attention [63]
	Ventral Inferior Parietal (VIP)	
Frontal cortex	Frontal Eye Field (FEF)	Controls the mechanisms of the eye movements [63]



OCCIPITAL REGION
area 17 *striate area* or *primary visual cortex* V1
area 18 *occipital area* or *secondary visual cortex* V2
area 19 *preoccipital area* or *associative visual cortex*
includes V3, V4, and V5

TEMPORAL REGION
area 20 *inferior temporal area/gyrus* (itg)
area 21 *middle temporal area/gyrus* (mtg)
area 22 *superior temporal area/gyrus* (stg/sts)
includes *superior temporal sulcus face area* (fSTS or FA-STS)
area 37 *occipitotemporal area*
includes *fusiform gyrus/fusiform face area* (FFA)

PARIETAL REGION
area 7 *superior parietal area*
includes area 7a

FRONTAL REGION
area 8 *intermediate frontal area*
includes *Frontal Eye Fields* (FEF)

Figure 3. Compact representation of the regions of the brain and respective Brodmann areas [67] involved in the visual signal and face processing. The regions are identified by numbers, according to Brodmann's model. The OCCIPITAL REGION includes the *occipital face area* (OFA) mentioned in Figure 1. The area called *anterior face patch 1* (AFP1) is located between the occipital and temporal regions. The *orbitofrontal cortex* is located in the FRONTAL REGION.

A stimulating and inspiring face recognition branch regards the building of face mental representation. Studies on **representation of visual stimuli** in the temporal cortical visual areas showed that "one of the major problems which must be solved by a visual system is the building of a representation of visual information which allows recognition to occur relatively independently of size, contrast, spatial frequency, position on the retina, angle of view" [32]. Thus, a face can be recognized as familiar when there is a match between its encoded representation and a stored structural code [23]. In particular, there are three representational stages beyond the retinal image: the primal sketch, which makes explicit the intensity changes present in the image; the viewer-centred representation, that describes the surface layout of a viewed scene relative to the viewer; the object-centred representation, allowing the recognition of objects from any viewpoint [23]. Apparently, the performance on face memory tasks only weakly predicts face matching ability, meaning that these two modes of face identification rely on rather different cognitive processes [17].

Timewise, the steps of human FR process have been studied and identified by Zheng et al. in 2012 [48]. If at zero milliseconds a face is seen, at 100th ms the first preprocessing happens; at 150th ms the visual stymulus is identified as a face. Then, at 150th-200th ms the facial features are processed (structural encoding), and at 200th ms a facial representation is created, to be compared with other representations stored in the brain. At 300th ms the process is terminated and the face is recognized or not. The neural computation is to be considered feedforward-working [4] [10].

Besson et al. also worked on timing of human FR, but considering the levels of facial image processing. According to their study, there are three **face processing levels** relying on minimal reaction times (minRT) [13]. The first is the SUPerordinate (*human*), called human face categorization (HFC), which corresponds to the process in which the visual system concludes "this face is human"; the minRT is 260 ms. The second is the SUBordinate (*familiarity*), called familiar face recognition (FFR); the visual system takes 360ms (minRT) to state that "this face is familiar". The last is SUBordinate (recognition/*finding someone in the crowd*), called individual face recognition (IFR). The process ends up with asserting "this face is Brad Pitt"; this part is 120/140ms faster than FFR and 20/30ms longer than HFC, i.e. its minRT is approximately 0.25 seconds. The important findings of this work are: i) even if they are both considered as FR, FFR and IFR are actually different processes; ii) the *familiarity* process is the least accurate and more difficult, while *human* one is the most accurate and easy; iii) HFC is the least effected by facial inversion effect. In a nutshell,

accuracy(FFR) < accuracy(IFR) < accuracy(HFC);
difficulty(FFR) > difficulty(IFR) > difficulty (HFC).

Meaning:
accuracy(“this face is familiar”) < accuracy(“this face is Brad Pitt”) < accuracy(“this face is human”);
difficulty(“this face is familiar”) > difficulty(“this face is Brad Pitt”) > difficulty(“this face is human”).

In presence of face inversion:
accuracy(FFR) < accuracy(IFR) < accuracy(HFC).

Besson, et al., 2017 [13]

2.2 PSYCHOLOGICAL DYNAMICS

Besides identity, faces convey other information such as gender and ethnicity which influence our categorization process and our daily interactions. Even if we are relatively accurate in recognition of familiar faces, a number of factors influence our ability particularly for unfamiliar faces. First, there are large individual differences; some people are definitely more able in face recognition and memory, although people are often not aware of it, as adults with typical face recognition ability have modest insight into their ability [68]. Besides individual differences, evidence suggests people are significantly more accurate at recognizing faces of their own group compared to outgroup faces. Even the so-called super-recognizers, people with extraordinary ability in FR, are subject to the same biases as typical perceivers [69]. The most investigated own-group or in-group biases are relative to gender, age, and race or ethnic groups. However, biases have also been found for groups even without physiognomic facial markers (i.e. sexual orientation, students of same university, political affiliation, experimentally induced ingroup/outgroup memberships) [70] [71] [72].

Recent studies explored these biases often jointly, employing techniques as evocated potential or neuroimaging to highlight the underlying neural mechanisms, or eye tracking to investigate the face processing strategies, with the aim to develop a comprehensive model and to disambiguate between the two most proposed accounts of the in-group biases. The first account refers to familiarity and greater perceptual expertise. Frequent contacts with own-groups people shape the "face space" [73] and allow a greater accuracy of the face representational system for distinguish individual faces of the in-group [70]. A second competing account of these biases refer to sociocognitive motivational mechanism. Faces perceived as belonging to an out-group are processed at a categorial level, whereas in-group faces are individualized [74], with a mechanism similar to the use of stereotypes and at the origin of prejudice; therefore, it is a different level of motivation which results in deeper processing and subsequent greater recognition and memory.

After a brief definition of the most studied biases, we summarize the recent empirical evidence focusing on the underlying mechanisms.

2.2.1 The gender bias

Women are better at recognizing female than male faces, while evidence on men is more contradictory [75] [76]. In addition to this asymmetrical own-gender bias (OGB; see [77], for a metananalysis), female faces are recognized more accurately than male faces both by female and male participants [78]. Also, women are generally more accurate at recognizing both female and male faces compared to men [79] [80]. The superior ability of women in face recognition is still under debate [81] [79] [82] but there is evidence that neonatal girls already spend more time looking at a face than neonatal boys [83].

The pionieristic study of Cross et al. proposed an explanation to the gender bias arguing that women are more skilled at recognizing female faces because continually exposed to that kind of stimulus [84]. Herlitz and Lovén suggest that men and women are both exposed to female faces to a greater extent during infancy than to male faces (mother, caregivers). This potentially leads to deeper processing and to greater recognition accuracy [77]. The asymmetry in the OGB may therefore derive from an early perceptual exposure to female faces [83] which in females persists throughout development because of larger number of reciprocal interactions with other females. The OGB in women remains also for less familiar faces, such as other-race and other age-faces, supporting that women have acquired a generalized perceptual expertise for female faces [78]. Boys and men, on the other side, may become more familiar with other males but this at most could weaken their original ability for processing female faces, thus resulting in the lack of gender bias often found in males.

Palmer et al. support instead a motivational basis for the gender bias; they used a divided-attention paradigm during the encoding of faces and found a reduced recognition particularly for males, suggesting that women play more attention at encoding own-gender faces [76]. Similar support in term

of motivational underlying mechanism comes from studies with participants with diverse sexual orientation [75] [72], and the findings that a pro-female gender bias in face recognition was found in all participants with the exception of gay men, who showed a pro-male gender bias. Lovén, Herlitz and Rehnman also manipulated in three experiments the level of attention reducing the processing of faces during encoding, but their results did not supported the hypothesis that own-gender women bias would be reduced; therefore, their findings point to the role of women superior perceptual experience for female faces, which subsequently facilitate recognition memory [85].

Exploring with fMRI the neural substrates of gender differences [86] findings showed, besides supporting the superiority in recognizing female faces both in men and women, greater activation to female vs male faces and a reduced activation of women brains, according to the author suggesting a greater efficiency in the women relevant neural system. Using event-related brain potentials (ERP), measured during tasks of face orientation and gender identification of presented faces, Sun, Gao & Han found an activation of specific brain area (larger for N170 in orientation identification, P3 in gender identification), with stronger effects difference in women than in men, suggesting that attention to social information in faces assume more relevance in women than in men [87]. Using eye-tracking [88], different scanning when processing own- and other-gender faces emerged, with more and longer fixations to the eyes when viewing own-gender faces.

Comparing the behavioral and neural processes underlying the OGB, both in the learning (exposure to unfamiliar faces to memorize and categorize by gender) and in the test phase (recognizing the presented face as new or learned), the ERP did not reveal strong evidence for different processing of own- and other-gender faces, although the gender biases have been confirmed and, in this study, both for male and female participants [74]. However, a difference in the learning phase in some neural correlates related to memory effects (occipito-temporal P2 and central N200) was potentially linked to the own-gender bias. In addition, the memory bias did not correlate either with quality or frequency of contact with female or male individuals, and the authors conclude that the own-gender bias largely rely on perceived ingroup vs outgroup membership, therefore to social categorization and not to perceptual expertise.

2.2.2 The age bias

Evidence supports the existence of a similar own-age bias, with superior recognition for faces of one's own age group. A meta-analysis [89] confirms that the discrimination ability is better for same-age compared to other-age faces at all ages, even in children. A possible explanation refers to the role of recent experience with people of similar age and is supported by the presence of the bias even in groups, such as older people, with prior experience with other age groups. However, a social categorization task of faces in old/young age group was found to be correlated with the OAB, pointing to a motivational account [90].

To test whether generally FR increases with age, Cross et al. compared 7, 12, 17 yo children, but findings were not positive [84]. Faces constitute special stimuli already in neonates [91], who are able to recognize the primary components that distinguish a face (nose, eyes and mouth), but also have considerable expertise in identifying the spatial relations between the characteristics, as the distance between the eyes [92]. Despite numerous works in support, there is however currently no empirical evidence demonstrating in an irrefutable manner that the recognition of faces is governed by innate factors. This conclusion is difficult to be proved because from the earliest days of life the child will come in contact with many facial stimuli rather than with other types of visual stimuli. Wolff showed that they smiled in response to a face starting from the fourth week of life [93]; besides, since the second week they manage to discriminate between the mother's face and another face [94] and prefer the face of the mother than any other feminine face, familiar or not [95].

2.2.3 Cross race

The most investigated bias in face recognition is the cross-ethnic effect (in the literature, usually termed the Cross-Race Effect, CRE, or Other-Race Effect, ORE) which refers to the superior identification of same-race compared to other-race faces.In some people, the ORE effect reached almost a clinical level, arriving to be considered face-blind for other-race faces [96], probably because of a lack on interethnic contact and a general low face recognition ability.

Albeit ORE is sometimes found to be stronger in White people [84] [97] [73], it exists broadly across several ethno-cultural groups [98]. The level of interethnic contact explained part of the findings, but even the attitudes toward ethnic groups probably play a role. A perceptual-social linkage hypothesis [99] suggests that the greater exposure to own-race faces together with the mostly positive experiences with these faces may lead to implicit racial bias. According to Galper [100], positive behaviors toward people of different ethnic groups may contribute to an improvement of the memory performances, to achieve comparable results to people belonging to that same ethnicity, as observed in a group of white students attending courses with black students, who were very skillful in memorizing the faces of their black colleagues.

Time ago, Malpass [101] demonstrated the possibility to improve the recognition performance through training, teaching the subjects to focus on the salient but at the same time discriminating characteristics of a face. For a white subject the most salient feature of a black face is the color of the skin but it is a little characteristic element within a group of black individuals [1]. Attention, to facilitate memory, should be focused on other elements like the shape of the nose and mouth [102]. Eye tracking confirmed that internal facial features (particularly eyes, then nose and mouth) are fundamental for face recognition as these features receive more fixations than others even in children, especially for upright faces [103]. A further question may refer to the relative importance of specific diagnostic facial region in other- and same- race faces. For instance, Caucasian children and adults use eye-centric scanning for both, while Asian children and adults use nose-centric scanning for own-race faces and eye-centric scanning for other-race faces [104]. A study [105] which disguised the upper and lower regions (with sunglasses and a bandana) confirm the superiority of the eye region; when disguised, recognition memory was impaired with similar effects on recognition of same- and cross-race faces. However, whereas the upper facial region is more diagnostic for white faces, the lower facial region is more diagnostic for black faces. The context also plays a relevant role, and experiments comparing same- and other-race presentation of faces individually or in groups found that the group presentation impaired the recognition memory of cross-race but not of own-group faces, pointing to the influence of viewing conditions in real world situations [106], as in the case of eyewitness misidentification of other-race offenders leading to wrongful imprisonment [96].

According to socio-cognitive motivational account, processes of categorization vs individuation at encoding account for differential recognition of same- and cross-race faces [107], playing a role even at post-encoding, at recognition [108]. That is, faces are quickly coded by their group status and ingroup faces are processed deeply whereas outgroup faces are processed shallowly [88]. Support derives from studies which use divided or reduced attention [76], labeling of own-race faces as belonging to an ingroup/outgroup as same/other university [109], or instruction to pay attention to outgroup faces during learning [110].

If motivation is the underlying mechanism, the amount of effort people applies in recognizing own- and other-race faces should be different. With this hypothesis, Crookes & Rhodes [111] compared recognition of Australian and Chinese faces in Caucasian participants in two conditions: with standard timing or with self-paced timing. If people are less motivated in recognizing other-race faces, they should apply less effort and spend less time studying other-race than own-race faces. Evidence did not support the hypothesis, and actually participant reported more effort in telling apart other-race than own-race faces. These findings are therefore contrary to the social-cognitive explanations, and pointing to a perceptual-experience account. It has been argued that the evidence supporting the social–motivational

account derives from a particular cultural setting: a high socio-economic status group (typically US whites) looking at the faces of a lower status group (US blacks) with whom observers typically have at least moderate perceptual experience. Five studies [112] tested the relative role of motivation-to-individuate instructions in participants with wide range of perceptual experience but with equal socio-economic status (Asian and Caucasian participants), and found no support to the motivational account; that is no reduction in the ORE with motivation instructions. The level of ORE was predicted mainly by the level of contact with the other races.

Contact may however not completely explain the ORE, given that it has been also reduced by labelling own- and other-race face as members of the ingroup, e.g., same vs other university [113]. This points to a different level of motivation in encoding faces after the initial simple categorization of face as ingroup or outgroup members compared to a facial identity task; this latter may increase motivation to individuate the faces [74].

Further evidence supports the role of experience. ORE has been reversed if an individual has earlier exposure to people from another ethnic group, e.g. Korean adults adopted by French families recognized Caucasian faces more accurately than Asian faces [114]. Infants at 3 months initially recognize both own- and other-race face but 9-months olds recognized only the familiar own–race faces. Two alternative explanations refer to a perceptual-learning mechanism or to a perceptual narrowing one. The first suggests that initially newborns are unable to process face identity of all races, but with the extensive exposure to own-race faces they develop better ability in recognizing own-race faces, while the narrowing maintains that the ability to recognize other-race faces declines because of the lack of experience [115].

The exposure to faces may influence, starting from early infancy, scanning and recognition of faces. Infants as early as at 6 months of age scan own- and other faces differently and already show a superior recognition of own-race faces compared to other-race faces [116] [117]; using eye tracking, after familiarization with dynamic faces, children recognized later own-race static faces but not other-race faces, with better recognition with more scanning. In addition, children fixated longer on the nose of own race faces. The familiarity that infants have with the faces they encounter daily determine that 3-months old look significantly longer at own- over other-race faces, 6 month-olds look at both equally, and 9-month-olds look significantly longer at other- over own race faces [116]. This is probably due to the fact that they increasingly process more rapidly information from own-race faces and spend more time exploring less familiar faces.

In two ethnic groups with high degree of reciprocal contact (Andalusian Romani and Andalusian Caucasians), the recognition performance of East Asian, Caucasian, and Romani faces support that the ORE is due to lack of contact. In addition, eye movements were monitored and, although with similar face recognition performance, Caucasian and Romani observers employed different visual exploration strategies; Romani focused their attention on the eyes, while Caucasians fixated more on the nose [118]. Finally, ORE was also reversed, arriving to an other-race advantage, when Caucasian participants recognized black male faces better than white male faces when the facial expressions indicated anger or fear [119], thus probably interpreting the faces as potentially threatening.

Examining with eye-tracking visual attention and memory for faces of different race and gender in women and men, thus studying at the same time both ORE and OGB, a study [78] highlighted that own-race faces were generally viewed longer than other-race faces and subsequently recognition memory was higher, as it was for female faces than for male faces, and that male faces of other race received more time viewing, probably because they may be perceived as a potential threat. However, longer viewing time influenced later memory for faces but did not reduce the magnitude of ORE. The proposed explanations point mainly to a dual process, that is social-motivation – lower in recognizing other-race faces - and perceptual experience - greater for own-race faces because of greater exposure and familiarity. Similarly, the Categorization-Individuation model (CIM) [70] proposes that social categorization, perceiver motivation, and perceiver experience are the three causes at work to different extents in the own-group biases. Categorization involves a focus on shared characteristics among a class

of exemplar (i.e., category-diagnostic information, for instance the skin color for race), while individuation involves an attentional focus on characteristics unique to a particular exemplar (i.e., identity-diagnostic information). The first occurs quickly, effortlessly, and spontaneously for novel faces and tend to be more salient for outgroups than for outgroups, while individuating characteristics may be more relevant for ingroup faces. An experimentally-induced focus on identity cues has shown to be effective in reducing the OAB, but it not reduced the ORE [120], supporting that the categorization-individuation processes do not play the same role in the two biases. A dual-route approach to ORE is supported also by the findings of Wan et al. [112] who propose that lack of motivation and lack of experience contribute differently across varying world locations and cultural settings.

A final consideration about cross-race biases regards emotion recognition. An eye-tracking study of Hu et al. [45] showed that viewers did not attribute more own-race than other-race neutral faces as neutral; more Chinese neutral faces were perceived as neutral by all participants. Also, the attribution of positive or negative expressions to the presented faces was influenced by stereotypes, with greater attribution of negative expression to Chinese faces and of positive expressions to Caucasian faces; in addition, participants' pupil size was larger when processing other-race faces, suggesting greater cognitive load in processing unfamiliar other-race faces.

Interestingly, also automatic FR systems are affected by cross-race bias. Existing training and testing databases consist of almost Caucasian subjects, implying that "there are still no independent testing databases to evaluate racial bias and even no training databases and methods to reduce it" [121]. This means that racial bias in databases reflects in algorithms; thus, the performance of FR methods depends on the race. Wang et al. addressed this issue by introducing a new dataset with Caucasian, Asian, Indian, and African subjects [121]. This suggests that putting in contact perception foundations and automatic FR algorithms could contribute to answer to the research questions and make the two fields nurture each other.

3 BIOLOGICALLY-INSPIRED FACE RECOGNITION ALGORITHMS

Several findings of neuroscientists studying human FP have important consequences for engineers who design algorithms and systems for machine recognition of human faces [1]. Increased knowledge about the way people recognize each other may help to guide efforts to develop practical automatic face recognition systems [10]. This trend has emerged only recently. It has been reported that, in 1995, barring a few exceptions, research on machine recognition of faces had developed independent of studies in psychophysics and neurophysiology. The reason is that at that time it was considered futile (impossible with that technology) to attempt to develop a system which could mimic all human traits [1]. Now, especially considering the studying the human brain is a hot research topic, the time has come for the technology to consider the functioning of the brain as a stimulus and inspiration for the algorithm conceptualization. This section sums up the most relevant contributions published after 2000, making a few exceptions of previous works which have been evaluated relevant and significant for the general understanding of the topic. The search keywords on Scopus were "face perception" and "face recognition"; the contributions have been then selected in the research areas of Computer Science and Engineering.

Two approaches are used in algorithms: **bottom-up** and **top-down**. Bottom-up is try to reverse-engineer the visual system by reconstructing the brain piece by piece so that artificial intelligence will emerge. Top-down is feed-forward engineering, i.e. computer vision and machine learning, which happens by building formal theories of visual perception and applying the biologically-plausible implementations of these theories to the neural circuits. Most of the biologically-inspired methods for FR are top-down models [65]. They respectively correspond to hierarchical methods, as they are based on hierarchical representations, and techniques relying on efficient coding hypothesis.

Neural networks and deep learning are the primary examples of how some functions of the human being, the **neural system** in this case, could provide a hint to researchers in Computer Science for developing computational programs. The Computer Science term 'neural network' "has its origins in attempts to find mathematical representations of information processing in biological systems", although in some cases their biological plausibility has been the subject of discussion [122].

Descriptions of neural system's functioning and of how could be simulated by computers have been provided in distinguished works exploring the topic from different perspectives. In 1943, McCullogh and Pitts proposed a neural system model adopting psychological functions (the activity of a single neuron is called "psychon") as starting point and focusing on the "all-or-none" law of nervous activity [123]. In 1962, Rosenblatt designed and tested perceptron brain models; a perceptron consists of a set of signal generating units (called "neuro-mimes") connected together to form a network [124]. Differently from McCullogh and Pitts's study, who focused on psychological functions, physiological principles form the basis of the perceptron models. Necessarily, the probabilistic properties of a class of physical sistems was here the departure point. An early primary contribution about adaptive or "learning" systems for pattern recognition was given by Widrow and Hoff in 1960 [125]. In this work, an adjustable neuron model is adopted for the construction of an adaptive pattern classification machine called Adaline (for <u>ada</u>ptive <u>line</u>ar).

Among the many types of artificial neural networks developed in the last decades, Convolutional Neural Networks (CNNs) are worth being taken into consideration in this study for their attempt to specifically resemble the **virtual cortex**, which makes them be suitable for applications in pattern recognition and, thus, face recognition. The origins of the model rely on Hubel and Wiesel's discoveries in the Sixties on two major cell types in the primary visual cortex (V1) of cats: simple cells and complex cells, the second receiving input from the first [126]. These findings have been transformed into a functioning model called Neocognitron by Fukushima in the Eighties [127], that relies on S-cells and C-cells named after simple and complex ones. "After a layer of simple and complex cells representing the basic computations of V1, the Neocognitron simply repeats the process again. That is, the output of the first layer of complex

cells serves as the input to the second simple cell layer, and so on. With several repeats, this creates a hierarchical model that mimics not just the operations of V1 but the ventral visual pathway as a whole. The network is "self-organized," meaning weights change with repeated exposure to unlabeled image" [128]. This model is considered as the precursor of modern CNNs, that use convolution mathematical operation in place of general matrix multiplication in at least one of their layers. Images fed into these networks are usually first normalized and separated into the three color channels RGB, that capture certain computations done by the retina. Each "stacked bundle of convolution–nonlinearity–pooling" can then be thought of as an approximation to a single visual area (usually V1, V2, V4, IT). Thus, standard feedforward CNN architectures are believed to represent the very initial stages of visual processing [128]. Lindsay's study [128] is a recent contribution focusing on how traditional CNN models and alternative architectures imitate the visual system.

For their success in image classification, CNNs are largely used for automatic face categorization tasks like FER [129] [130] [131], FR [132] and medical applications [133], with AlexNet, ResNet and VGG being the mostly adopted [130] [134]. Detailed literature reviews exist [134] [135] [136] [137] [138] [139] [140] [141] [142], for interested readers. CNNs can be considered a mainstream for computer-based face recognition and a key computational approach imitating the visual cortex.

Besides CNNs, the **human visual system** has been inspirational for other approaches to face categorization. Carnec and Barba presented a pattern recognition methodology based on perceptual data, simulating the functioning of the human visual system [63]. The implemented method reproduces a simpler version of the visual system conserving these specifications: i) the adoption of the eye as the starting point, represented by RGB images; ii) detection of oriented contrasts, simulating the role of V1 area; iii) topographic organization of information, simulating the role of V2 (see Table 3 for more details on these parts). Success rates range between 85 and 100%. Fu et al. adopted spiking neural network to simulate the visual cortex mechanism for facial expression recognition. Recent evidences from neurobiology have led researchers to build cortex-like scheme based model with single spiking neurons acting as computation units (most cortical physiologists believe that most neurons are in cortex spike), the Spiking Neural Networks (SNNs). SNN utilizes information representation as trains of spikes, embedded with spatiotemporal characteristics. Following the standard model of visual cortex, from the sensory/input layer to the final classification layer, the overall system consists of three main blocks: (1) the sensory/ receptive layer, which consists of simple cell behavior simulator and complex cell behavior simulator, data preprocessing, feature extraction; (2) the learning layer, which consists of only excitatory neurons; (3) the classification layer, which accumulates all the outputs from the learning layer [144]. Zhang et al. applied a hierarchical cortical model to facial expression recognition. The backbone idea relies on Poggio's cortex mechanism like system. "The core of the model is the hypothesis that the main function of the ventral stream is a mechanism which has evolved to achieve the trade-off between selectivity and invariance in IT area for fast and accurate object of interest recognition tasks, which is done through a underlying hierarchical structure (from retina to IT). The model has been successfully tested on the JEFFE database, proving that the cortical like mechanism for facial expression recognition should be exploited in great consideration [64]. Tistarelli et al. presented a pseudo-hierarchical Hidden Markov Model (HMM) to build a double layer architecture to extract shape and motion information from face sequences, inspired by the fact that we capture both facial traits (static) and behavioural (dynamic) features. The main outcome is the capability of modelling both physiological and behavioural features, so that identification and verification tasks could be gained effectively [52].

In 1991 Cottrell and Metcalfe introduced *holons* as features for FR [145]. Holons are given to one and two layer back propagation networks that are trained to classify the input features for identity, emotions, and gender. The developed system has been called Emotion PATern recognition using Holons (EMPATH); it is a biologically plausible neural network model for categorizing facial expressions [144]. In 2002, Dailey et al. showed that EMPATH, trained to classify facial expressions into six basic emotions, predicts data used to support both the discrete and continuous theories about facial expression perception. The system is a feed-forward network consisting of three layers common to most object recognition models. The first layer of the model is a set of neurons whose response properties are similar to those of complex cells in the visual cortex. Gabor filters are used to model complex cells. "As

a feature detector, the Gabor filter has the advantage of moderate translation invariance over pixel-based representations. This means that features can move in the receptive field without dramatically affecting the detector's response, making the first layer of our model robust to small image changes" [146]. Other neurocomputational models of face processing have been investigated in 2011 in a dedicated study by Cottrell and Hsiao [147]. They resumed the features of three biologically plausible models developed in the last years of the Nineties: (i) O'Really and Munataka's model included V1, V2, V4 and the lateral gineculate nucleus and was considered the most biologically plausible; (ii) "the standard model" (TSM) included V1, V2, V4, inferior temporal cortex layer and prefrontal cortex, was developed by Riesenhuber and Poggio and involved Gabor filters; (iii) Cottrell and Dailey's "the model" (TM) was considered the least biologically plausible, including V1 and FFA, with the use of Gabor filters. TSM was considered by the authors "neurally realistic, while requiring a search for parameters to fit the data, and TM being the less realistic, but using learning on a realistic task to set the parameters" [147]. The development of models which could be biologically plausible in both their learning methods and their neural processing mechanisms is seen by the authors as the next crucial research step.

The adoption of Gabor representation for approximating the shape of the receptive fields in the primary visual cortex was firstly introduced by Marčelja in 1980 [148] relying on a description of simple cortical cells given by Hubel and Wiesel [126]. Petkov et al. applied Gabor functions to FR. They developed a biologically motivated FR approach with the adoption of 2D Gabor functions resembling the receptive fields of simple cells in the primary visual cortex V1. Data obtained by projecting a 2D signal (image) onto a set of Gabor functions is interpreted as the activities of individual cells in the V1. A 94% RR was achieved on a database of 205 facial images of 30 subjects [149]. More recently, Gabor representations were adopted by Sato et al., that proposed a computational model of the cognitive processing in the brain involving both configurational and featural processes [150]. According to the authors, the first cerebral step is the configurational process, which is the perception of internal relationships among features, that are codified by low spatial frequencies. The second step is the featural process; it relies on the perception of contour details of nose, eyelid, mouth, and eyes. These contours are codified at high frequencies. Low and high frequencies play a key part in human FR, as they define the framework of human visual system and the concepts of holistic and local processing, respectively. The possibility to emulate this framework in the algorithms encouraged the authors to employ Gabor pyramids. The purpose of the study was to explain why human observers have difficulty in distinguishing two extremely similar faces of different genders, for applications in Human-Computer Interaction context. The literature on the adoption of Gabor features for pattern recognition applications is vast and varied, and has been reviewed in some *ad hoc* studies [151] [152] [153].

Deng et al. argued that Gabor wavelet representations were not enough to gain acceptable RRs [154]. They developed an uncontrolled FR system aiming at emulating early processing, face coding, and cue-fusion strategies of human FR. They adopted a set of biologically inspired **features** derived from the response properties of neurons in the early stages of visual pathways (spectral and spatial frequencies, orientation, color opponent); developed a **dimensionality reduction module** relying on an incremental robust discriminant module with the objective of emulating the perceptual learning that happens in later stages of visual pathway to extract higher level (identity-specific) codes; processed **face coding** on different facial cues, yielding separated visual pathways, in order to predigest the highly complex cue-fusion strategy of human FR. The authors obtained 93% RR and encouraged researchers to search more informative features and perceptually optimal face codes in order to build a valid perception-inspired computational model.

The human visual system has also been inspirational for the **visual attention mechanism**, with successful results in FER tasks. Visual attention mechanism relies on the functioning of the human visual system, meaning that, when processing a complex visual scene, humans have the ability to rapidly direct the gaze towards objects of interest in the scene. The simulated attention mechanism is often embedded into a CNN architecture aimed at classifying facial expressions with automatic determination of regions of interest [155], occlusion identification [156], and the adoption of 3D data (depth maps) [157].

Puri and Lall relied on the role of abstractions to develop a methodology for head pose estimation in real time [158]. Both the brain and the algorithms work faster with **abstractions** of images, i.e. with features, which must be extracted from images before the matching process. Humans do this process automatically, while computers need to be provided with features [158] [159]. Generally speaking, **features** are largely adopted in the Computer Vision field for recognition purposes [160] [161], but not often consciously connected to FP. Similarly, the ability to abstract from **invariants** has been investigated. While modelling differences in perception between autistic and non-autistic children, Wu et al. made some considerations about the role of invariants in the human perception. "Real-world invariants may often be hard to detect because they are hidden behind several stages of processing. The problem here is quite similar to that of dimensionality reduction in machine learning: knowledge of symmetries and other invariants allows one to reduce the search space and is a basic element of 'imagination' and planning. Thus, such knowledge is essential in real-world intelligent agents" [159].

Zhan et al. took from the FP context the concepts of **familiarity** (role of memory) and the hierarchy of importance of facial components to propose a FR method to deal with the "one sample problem". Particularly, quantized local features which learnt from generic face dataset are used to mimic the prototype effect of human FR [8]. Every local feature is quantified according to a codebook to mimic the prototype effect. The codebook is built using clustering approach based on a generic facial dataset which consists of images from subjects other than those under consideration. In this method, the authors also adopted the **caricature concept** to weight the distinctiveness in the comparison with the average face. The Mahalanobis distance is used to measure the distinctiveness of every facial part (mouth is excluded due to its shape variations). Similarly, Dagnes et al. used the average face to evaluate the distance of the query face to the typical one, thus reprising the caricature idea. For this purpose, a novel descriptor based on the shape index and the average face, called 'personal shape index', has been adopted as discriminating feature for a novel FR methodology [162] [163].

Another FP aspect which has been adopted in the algorithms is **Weber's law** [164] [165]. As the ability of human perception to dissolve intensity change ΔI depends proportionally to the absolute intensity I, "the human perception of local micro-pattern is defined by the ratio between two terms: one is relative intensity differences of a central pixel against its neighbors and the other is intensity of local central pixel" [165]. Xie et al. proposed a face representation based on this law and a FR methodology relying on Local Binary Patterns (LBPs) on images. Dunker and Keller dealt with illumination normalization for FR relying on the same law [164].

A perception-driven approach was constructed by Yu et al. to synthesize facial expressions in the absence of actor performance data for that expression [166]. The authors blended 3D Morphable Model computer graphics technique with reverse correlation, a method imported form human psychophysics. By studying visual properties during stimulus classification tasks using behavioural and brain responses, reverse correlation uses the response of human observers to filter the parametric values that correspond to a specific expression. This perception principle could be considered belonging to the already mentioned internal **mental representation**.

Some hints for specific computational models for the "psychology of face recognition" come from neuroscientific studies. Active Appearance Models (AAMs) are being considered by Zhan et al. for modelling the **mental representation** of a face identity [167]. AAM is a mathematical model used in the context of Pattern Recognition which contains full 3D surface and 2D texture information of the face. The general aim of the study is "to model the face identity contents in the generative 3D space of faces and to use these models to generate identification information in resemblance tasks that test the generalizability of identity information." This research may open new avenues for the interplay between visual information, categorization tasks, and their implementation.

Table 4 sums the content of this section by reporting respective aims, methods, and results in terms of recognition rates.

Table 4. Schematization of the works adopting FP concepts in FR and FER algorithms.

Work	Inspiring perception concept	Aim	Method	Results
Carnec and Barba [63]	**visual cortex mechanism**: i) the adoption of the eye as the starting point, represented by RGB images; ii) detection of oriented contrasts, simulating the role of V1 area; iii) topographic organization of information, simulating the role of V2	pattern recognition	*ad hoc* (pseudocode written in the paper)	85-100% RR
Fu et al. [144]	**visual cortex mechanism**: (1) the sensory/ receptive layer, which consists of simple cell behavior simulator and complex cell behavior simulator, data preprocessing, feature extraction; (2) the learning layer, which consists of only excitatory neurons; (3) the classification layer, which accumulates all the outputs from the learning layer	FER	spiking neural network	97.35% RR
Zhang et al. [64]	**visual cortex mechanism**	FER	hierarchical cortical model	95.1%, 95.3%, 92.4%, 89.5% RR depending on the database
Tistarelli et al. [52]	**visual cortex mechanism**; capture of both facial traits (static) and behavioural (dynamic) features.	extract shape and motion information from face sequences for FR	pseudo-hierarchical Hidden Markov Model (HMM)	100%
Dailey et al. [146]	discrete and continuous theories about facial expression perception	FER	Emotion PATern recognition using Holons (EMPATH), a biologically plausible neural network model. Gabor filters	90% RR
Petkov et al. [149]	functioning of the receptive fields of simple cells in the primary visual cortex V1	FR	Gabor representation	94% RR
Sato et al., [150]	cognitive processing in the brain: configurational and featural processes	explain why human observers have difficulty in distinguishing two extremely similar faces of different genders, for applications in Human-Computer Interaction context	Gabor pyramids	/
Sun et al. [155]	**visual attention mechanism** of human visual system	FER	novel Convolutional Neural Networks	depending on the database

Li et al. [156]	**visual attention mechanism** of human visual system	FER	novel Convolutional Neural Networks	depending on the database and on CNN version
Li et al. [157]	**visual attention mechanism** of human visual system	FER	novel Convolutional Neural Networks	94.63%, 98.52%, 94.33% RR depending on the database
Deng et al. [154]	early processing, face coding, and cue-fusion strategies of human FR	FR	specific method	93% RR
Puri and Lall [158]	**abstractions**	head pose estimation in real time	regional segmentation and stroke rendering	80-96% depending on the angle
Wu et al. [159]	**invariants**	FR	specific method	/
Zhan et al. [8]	**familiarity** (role of memory); hierarchy of importance of **facial components**; **caricature concept**	FR	landmark and feature based method	92%, 77%, 75%, 71% RR depending on the dataset
Dagnes et al. [162]	**caricature concept**	FR	specific method	95.39% RR
Xie et al. [164] [165]	**Weber's law**	Face representation and FR	LBPs	98% RR
Yu et al. [166]	**mental representation** of the face	face expression synthesis	3D Morphable Model + reverse correlation	95.4% by observers

4 DISCUSSION AND FUTURE WORK

In 1995, Chellappa et al. suggested engineers to use FP concepts when designing FR methods: the role of holistic and features; the significance of facial components (eyes ...); the caricature concept; the role of spatial frequency analysis. In particular, "designers should include both global and local features for representing and recognizing faces" [1]. Although some of these suggestions have been taken into consideration in the last decades, as Section 3 shows, perception foundations are still seen as something only for psychologists. The caricature concept was particularly mentioned as a suggestion for future work in pattern recognition [10] [8]. Other hints concern adaptation, invariant scale and position features and in general more informative biologically-inspired features [154], probabilistic parameters [144], the equal importance of pigmentation and shape cues [10], perceptually optimal face codes [154].

Concerning the model types, according to Fu et al., bottom-up models have proven to be qualitatively constrained by the anatomy and physiology of the visual cortex and may not be actually suitable for practical computer vision systems, and a more high-level computation framework is required. Hierarchical models can be viewed as conceptual tools rather than computational means; how to combine the logic structure of the hierarchy with the computation unit in vivo should be considered with great attention [144].

The main concerns of researchers in the field regard the little we know of the functioning of the visual system and general in the field of perception. Specific researches for human facial information and facial expression recognition in the field of biologically-motivated models are missing. Surprisingly, little works attempted to explore the cognition mechanism using computing units emulating the visual system [144], and a framework for a biologically-inspired model mimicking brain functioning is still missing [64]. Also, the nature of FR e FER impairments in autistics remains unclear [6]; similarly to prosopagnosia, knowing how these impairments work could significantly foster the research on perception under normal conditions. Furthermore, studying the errors of recognition made by normal people both in everyday life and under laboratory conditions could help the general understanding of the FP process [23]. Anatomically speaking, "lesion data can answer the crucial question of whether a given brain region is necessary for a given computation" [168], supporting, this way, the design of related computational models. Another psychological concern regards how cognitive and developmental aspects are correlated to human FR performances. Psychiatric disorders such as eating disorders, phobias, obsessive-compulsive patterns, and depression may imply different facial perception attitudes and performances. For instance, it is believed that people with little self-concept are very good observers and thus develop more structured and performing human FR. The influence of personality traits in face processing [169] and the role of cultural and social factors [170] could also be more deeply investigated in this sense.

A further stimulus is given by the research question Are computers better than humans?. According to O'Toole et al.'s experimentations, for non-familiar faces, nowadays algorithms are more accurate than humans in the good and moderate facial image condition; comparable to humans in the poor accuracy condition. For familiar faces, humans beat computers [171]. Sinha et al. have confirmed this thesis by asserting that the appropriate benchmarking for evaluating machine-based face recognition systems is still human performance with familiar faces [10]. An endorsement is given by O'Toole et al., as they affirmed that the next generation of FR algorithms will have to operate with levels of robustness comparable to those humans show in recognizing familiar faces [171]. Thus, humans undoubtedly win, but fusing the machines and humans could give the best RRs [4], in particular when similarly looking people are involved [172] [173]. Dunker and Keller, indeed, conclude their comparative study (between conventional and perception-inspired methods) by asserting that the algorithms based on human perception outperformed nearly all other algorithms [164]. A similar assertion was made by Burtons and Jenkins [174], who suggested to computer scientists to base the FR match on abstracted representation of the face or to store multiple exemplars, in order to emulate human familiarity in the automatic systems. Investigating how we come to be expert is some faces is a key question to be answered for them.

The methodologies for FR and FER involving knowledge on perception or on the visual system are only a few compared to the vast literature within the pattern recognition field. In particular, it seems that biological hints are used unconsciously, such as for Gabor functions. Neural networks are a well-known example about how a human mechanism could be computationally copied. Indeed, some work has been done to emulate parts of the brain with deep learning, namely with deep neural networks (DNN). An example is given by the Blue Brain Project [175]. It is a bottom-up approach about the simulation of the whole neocortical column. The goal is attractive and elusive, i.e. building a brain-like device which contains simulated brain regions in order to emulate the brain-like process [64]. Nonetheless, it does not deal with the visual cortex alone, thus it cannot specifically reproduce human FR. Besides the common concern on artificial neural networks regarding the vagueness of their resemblance of the nervous system and the excessive conceptual simplifications which make them unsuitable to describe it ("the brain is given to us, the product of a long evolution — we do not want to know how it may work but how it actually does work") [176], computational neuroscientists have tried to model biological vision with DNN [177]. Though, no model exists for outlining the brain functioning for face processing specifically. Regarding these types of models, some researchers have suggested that computer science methodologies such as DNN could be a promising framework to investigate how visual stimuli emerge in the human brain [178]. This way, the union between machine learning and neuroscience can have twofold purpose: use evidence from neuroscience to help computer scientists to devise better the face and the facial emotions; use approaches in computer science to help neuroscientists to understand relevant neural processes more in-depth. The main aim of this paper was to address the first issue.

In conclusions, despite it could be considered as necessary that researchers in Computer Vision rely on FP for designing their algorithms, it could beneficial for them to be informed about the relevant findings, especially when human interaction is involved [1] [179] [41]. We have a gift from nature that should be deepened and exploited [7]. In the next future, we see as inevitable step of the research pathways in these fields (computer vision and pattern recognition, neuroscience, psychology) to start collaborations/projects between these and assemble interdisciplinary research groups with these knowledge, with the purpose to uncover together the meticulous functioning of the human visual system for FP and mutually juice up the respective expertise to create a common framework of knowledge in the area. Human-computer interaction applications are booming now, and perception could be the key factor which could make the big change in this sense.

5 CONCLUSIONS

This study presents the main foundations of facial perception and their application to pattern recognition algorithms such as face recognition and facial expression recognition. Despite the literature in the field of face recognition is vast and varied, only a few contributions involve findings coming from perception or inspired by the functioning of the human visual system. Gabor functions and features seem to be the most renowned example, but their adoption in face recognition is hardly due to a specific knowledge about complex cells in the primary visual cortex or the use of features at the brain level, respectively. The possibility to study the perception under different perspectives is fostered by the recent advances in neuroscience and the modern technology tools, which could uncover human face recognition processes and make this knowledge available to other disciplines.

ACKNOWLEDGEMENTS

We thank psychologist Marianna Sancineto for the consultancy regarding cognitive and developmental aspects; Giansalvo Cirrincione for the hints provided regarding neural networks; Piergianni Olivetti for the material concerning brain anatomy and Brodmann areas.

REFERENCES

[1] R. Chellappa, C. L. Wilson e S. Sirohey, «Human and machine recognition of faces: A survey,» *Proceedings of the IEEE,* vol. 83, n. 5, pp. 705-741, 1995.

[2] C. Darwin, The expression of the emotions in man and animals, London: John Murray, 1872.

[3] C. Bell, The Anatomy and Philosophy of Expression: As Connected with the Fine Arts, London, 1806.

[4] R. Chellappa, P. Sinha e P. J. Phillips, «Face recognition by computers and humans,» *Computer,* vol. 43, n. 2, 2010.

[5] M. J. Farah, K. D. Wilson, M. Drain e J. N. Tanaka, «What is" special" about face perception?,» *Psychological review,* vol. 105, n. 3, pp. 482-498, 1998.

[6] G. Van Belle, P. De Graef, K. Verfaillie, T. Busigny e B. Rossion, «Whole not hole: Expert face recognition requires holistic perception,» *Neuropsychologia,* vol. 48, n. 9, pp. 2620-2629, 2010.

[7] D. Y. Tsao e M. S. Livingstone, «Mechanisms of face perception,» *Annu. Rev. Neurosci.,* vol. 31, pp. 411-437, 2008.

[8] C. Zhan, W. Li e P. Ogunbona, «Face recognition from single sample based on human face perception,» *24th International Conference In Image and Vision Computing New Zealand (IVCNZ'09),* pp. 56-61, November 2009.

[9] B. Meinhardt-Injac, I. Boutet, M. Persike, G. Meinhardt e M. Imhof, «From development to aging: holistic face perception in children, younger and older adults,» *Cognition,* vol. 158, pp. 134-146, 2017.

[10] P. Sinha, B. Balas, Y. Ostrovsky e R. Russell, «Face recognition by humans: Nineteen results all computer vision researchers should know about,» *Proceedings of the IEEE,* vol. 94, n. 11, pp. 1948-1962, 2006.

[11] B. Meinhardt-Injac, I. Boutet, M. Persike, G. Meinhardt e M. Imhof, «From development to aging: holistic face perception in children, younger and older adults,» *Cognition,* vol. 158, pp. 134-146, 2017.

[12] E. McKone e R. Robbins, «Are Faces Special?,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 149-176.

[13] G. Besson, G. Barragan-Jason, S. J. Thorpe, M. Fabre-Thorpe, S. Puma, M. Ceccaldi e E. J. Barbeau, «From face processing to face recognition: Comparing three different processing levels,» *Cognition,* vol. 158, pp. 33-43, 2017.

[14] H. D. Ellis, «Recognizing faces,» *British Journal of Psychology,* vol. 66, n. 4, pp. 409-426, 1975.

[15] P. Thompson, «Margaret Thatcher: a new illusion,» *Perception,* 1980.

[16] J. W. Tanaka e I. Gordon, «Features, Configuration, and Holistic Face Processing,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 177-194.

[17] D. White, R. I. Kemp, R. Jenkins, M. Matheson e A. M. Burton, «Passport officers' errors in face matching,» *PloS one,* vol. 9, n. 8, p. e103510, 2014.

[18] P. G. Schyns, L. S. Petro e M. L. Smith, «Dynamics of visual information integration in the brain for categorizing facial expressions,» *Current biology,* vol. 17, n. 18, pp. 1580-1585, 2007.

[19] N. J. van Rijsbergen e P. G. Schyns, «Dynamics of trimming the content of face representations for categorization in the brain,» *PLoS Computational Biology,* vol. 5, n. 11, p. e1000561, 2009.

[20] M. L. Smith, P. Fries, F. Gosselin, R. Goebel e P. G. Schyns, «Inverse mapping the neuronal substrates of face categorizations,» *Cerebral Cortex,* vol. 19, n. 10, pp. 2428-2438, 2009.

[21] A. Ariga e K. Arihara, «Visual attention is captured by task-irrelevant faces, but not by pareidolia faces,» *9th International Conference on Knowledge and Smart Technology (KST),* pp. 266-269, February 2017.

[22] G. M. Davies e A. W. Young, «Research on face recognition: The Aberdeen influence,» *British Journal of Psychology,* vol. 108, n. 4, pp. 812-830, 2017.

[23] V. Bruce e A. Young, «Understanding face recognition,» *British journal of psychology,* vol. 77, n. 3, pp. 305-327, 1986.

[24] J. V. Haxby e M. I. Gobbini, «Distributed Neural Systems for Face Perception,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 93-110.

[25] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez e S. D. Pollak, «Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements,» *Psychological science in the public interest,* vol. 20, n. 1, pp. 1-68., 2019.

[26] A. J. Calder, G. Rhodes, M. H. Johnson e J. V. Haxby, The Oxford Handbook of Face Perception, New York: Oxford University Press, 2011.

[27] K. Dobs, A. Kell, I. Palmer, M. Cohen e N. Kanwisher, «Why Are Face and Object Processing Segregated in the Human Brain? Testing Computational Hypotheses with Deep Convolutional Neural Networks,» in *Cognitive Computational Neuroscience Conference*, Berlin, Germany, 2019.

[28] J. Liu, A. Harris e N. Kanwisher, «Perception of face parts and face configurations: an fMRI study,» *Journal of cognitive neuroscience,* vol. 22, n. 1, pp. 203-211, 2010.

[29] J. V. Haxby, E. A. Hoffman e M. I. Gobbini, «The distributed human neural system for face perception,» *Trends in cognitive sciences,* vol. 4, n. 6, pp. 223-233, 2000.

[30] K. Immelmann, «Sexual and other long-term aspects of imprinting in birds and other species,» *Advances in the Study of Behavior,* vol. 4, pp. 147-174, 1972.

[31] M. Johnson, «Face Perception: a Developmental Perspective,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 3-14.

[32] E. T. Rolls, «Learning mechanisms in the temporal lobe visual cortex,» *Behavioural brain research,* vol. 66, n. 1-2, pp. 177-185, 1995.

[33] E. Rolls, «Face Neurons,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 51-75.

[34] A. S. Redfern e C. P. Benton, «Expression dependence in the perception of facial identity,» *i-Perception,* vol. 8, n. 3, 2017.

[35] A. J. Calder, «Does Facial Identity and Facial Expression Recognition Involve Separate Visual Routes?,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 427-448.

[36] R. E. Galper e J. Hochberg, «Recognition memory for photographs of faces,» *The American journal of psychology,* 1971.

[37] S. Bate, B. Parris, C. Haslam e J. Kay, «Socio-emotional functioning and face recognition ability in the normal population,» *Personality and Individual Differences,* vol. 48, n. 2, pp. 239-242, 2010.

[38] J. D. LaBarbera, C. E. Izard, P. Vietze e S. A. Parisi, «Four-and six-month-old infants' visual responses to joy, anger, and neutral expressions,» *Child development,* pp. 535-538, 1976.

[39] B. Wang, «Facial expression influences recognition memory for faces: Robust enhancement effect of fearful expression,» *Memory,* vol. 21, n. 3, pp. 301-314, 2013.

[40] M. Jehna, C. Neuper, A. Ischebeck, M. Loitfelder, S. Ropele, C. Langkammer e C. Enzinger, «The functional correlates of face perception and recognition of emotional facial expressions as evidenced by fMRI,» *Brain Research,* vol. 1393, pp. 73-83, 2011.

[41] R. E. Jack e P. G. Schyns, «Toward a social psychophysics of face communication,» *Annual review of psychology,* vol. 68, pp. 269-297, 2017.

[42] P. G. Schyns, F. Gosselin e M. L. Smith, «Information processing algorithms in the brain,» *Trends in Cognitive Sciences,* vol. 13, n. 1, pp. 20-26, 2009.

[43] R. Ekman, What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), USA: Oxford University Press, 1997.

[44] J. F. Cohn, Z. Ambadar e P. Ekman, «Observer-based measurement of facial expression with the Facial Action Coding System,» in *The handbook of emotion elicitation and assessment*, 2007, pp. 203-221.

[45] C. S. Hu, Q. Wang, T. Han, E. Weare e G. Fu, «Differential emotion attribution to neutral faces of own and other races,» *Cognition and Emotion,* vol. 31, n. 2, pp. 360-368, 2017.

[46] M. Nusseck, D. W. Cunningham, C. Wallraven e H. H. Bülthoff, «The contribution of different facial regions to the recognition of conversational expressions,» *Journal of vision,* vol. 8, n. 8, p. 1, 2008.

[47] L. Tian e S. Xiao, «Facial feature exaggeration according to social psychology of face perception,» *Computer Graphics Forum,* vol. 35, n. 7, pp. 391-399, October 2016.

[48] X. Zheng, C. J. Mondloch e S. J. Segalowitz, «The timing of individual face recognition in the brain,» *Neuropsychologia,* vol. 50, n. 7, pp. 1451-1461, 2012.

[49] A. O'Toole, «Cognitive and Computational Approaches to Face Recognition,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 15-30.

[50] J. L. Bittner e J. M. Gold, «The Impact of Symmetry on the Efficiency of Human Face Perception,» *Perception,* vol. 46, n. 7, pp. 830-859, 2017.

[51] A. M. Martinez, «Computational models of face perception,» *Current directions in psychological science,* vol. 26, n. 3, pp. 263-269, 2017.

[52] M. Tistarelli, M. Bicego e E. Grosso, «Recognizing People's Faces: from Human to Machine Vision,» *9th International Conference on Control, Automation, Robotics and Vision (ICARCV'06),* pp. 1-7, December 2006.

[53] N. Kanwisher e G. Yovel, «The fusiform face area: a cortical region specialized for the perception of faces,» *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 361, n. 1476, pp. 2109-2128, 2006.

[54] Grill-Spector, K. N. K. e N. Kanwisher, «The fusiform face area subserves face perception, not generic within-category identification,» *Nature neuroscience,* vol. 7, n. 5, p. 555, 2004.

[55] D. Y. Tsao, W. A. Freiwald, R. B. Tootell e M. S. Livingstone, «A cortical region consisting entirely of face-selective cells,» *Science,* vol. 311, n. 5761, pp. 670-674, 2006.

[56] W. Freiwald e D. Tsao, «Taking Apart the Neural Machinery of Face Processing,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 707-718.

[57] P. Grimaldi, K. S. Saleem e D. Tsao, «Anatomical connections of the functionally defined “face patches” in the macaque monkey,» *Neuron,* vol. 90, n. 6, pp. 1325-1342, 2016.

[58] K. A. Pelphrey e B. C. V. Wyk, «Functional and Neural Mechanisms for Eye Gaze Processing,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 591-603.

[59] M. I. Gobbini e J. V. Haxby, «Neural systems for recognition of familiar faces,» *Neuropsychologia,* vol. 45, n. 1, pp. 32-41, 2007.

[60] M. Eimer, «The Face-Sensitive N170 Component of the Event-Related Brain Potential,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 329-344.

[61] S. R. Schweinberger, «Neurophysiological Correlates of Face Recognition,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 345-366.

[62] M. L. Smith, F. Gosselin e P. G. Schyns, «From a face to its category via a few information processing states in the brain,» *NeuroImage,* vol. 37, pp. 074-984, 2007.

[63] M. Carnec e D. Barba, «Faces recognition based both on perception and behavior of the human visual system,» *14th International Conference on Digital Signal Processing,* vol. 1, pp. 175-179, 2002.

[64] T. Zhang, G. Yang e X. Kuai, «On cortex mechanism hierarchy model for facial expression recognition: multi-database evaluation results,» *International Symposium on Neural Networks,* pp. 141-148, July 2012.

[65] T. Poggio e T. Serre, «Models of visual cortex,» *Scholarpedia,* vol. 8, n. 4, 2013.

[66] E. Pernkopf, Atlante di anatomia sistematica e topografica dell'uomo, Volume I, Testa e collo, Padova: Piccin editore, 1963.

[67] K. Brodmann, Brodmann's: Localisation in the cerebral cortex, Springer Science & Business Media, 2007.

[68] R. Palermo, B. Rossion, G. Rhodes, R. Laguesse, T. Tez, B. Hall e S. Al-Janabi, «Do people have insight into their face recognition abilities?,» *The Quarterly Journal of Experimental Psychology,* vol. 70, n. 2, pp. 218-233, 2017.

[69] S. Bate, R. Bennetts, N. Hasshim, E. Portch, E. Murray, E. Burns e G. Dudfield, «The limits of super recognition: An other-ethnicity effect in individuals with extraordinary face recognition skills,» *Journal of Experimental Psychology: Human Perception and Performance,* vol. 45, n. 3, pp. 363-377, 2019.

[70] K. Hugenberg, J. P. Wilson, P. E. See e S. G. Young, «Towards a synthetic model of own group biases in face memory,» *Visual Cognition,* vol. 21, n. 9/10, p. 1392–1417, 2013.

[71] D. G. Ray, «Wondering is enough: Uncertainty about category information undermines face recognition,» *Journal of Experimental Social Psychology,* vol. 82, p. 16–25, 2019.

[72] P. J. Hills, J. M. Pake, J. R. Dempsey e M. B. Lewis, «Exploring the contribution of motivation and experience in the postpubescent own-gender bias in face recognition,» *Journal of Experimental Psychology: Human Perception and Performance,* vol. 44, n. 9, 2018.

[73] T. Valentine e M. Endo, «Towards an exemplar model of face processing: The effects of race and distinctiveness,» *The Quarterly Journal of Experimental Psychology,* vol. 44, n. 4, pp. 671-703, 1992.

[74] N. Wolff, K. Kemter, S. R. Schweinberger e H. Wiese, «What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias,» *Social Cognitive and Affective Neuroscience,* vol. 9, n. 5, p. 580–590, 2014.

[75] M. C. Steffens, S. Landmann e S. Mecklenbräuker, «Participant Sexual Orientation Matters: New Evidence on the Gender Bias in Face Recognition,» *Experimental Psychology,* vol. 60, n. 5, pp. 362-367, 2013.

[76] M. A. Palmer, N. Brewer e R. Horry, «Understanding gender bias in face recognition: Effects of divided attention at encoding,» *Acta Psychologica,* vol. 142, n. 3, p. 362–369, 2013.

[77] A. Herlitz e J. Lovén, «Sex differences and the own-gender bias in face recognition: A meta-analytic review,» *Visual Cognition,* vol. 21, n. 9/10, p. 1306–1336, 2013.

[78] J. Lovén, J. Rehnman, S. Wiens, T. Lindholm, N. Peira e A. Herlitz, «Who are you looking at? The influence of face gender on visual attention and memory for own- and other-race faces,» *Memory,* vol. 20, n. 4, p. 321–331, 2012.

[79] J. Rehnman e A. Herlitz, «Higher face recognition ability in girls: Magnified by own-sex and own-ethnicity bias,» *Memory,* vol. 14, n. 3, p. 289–296, 2006.

[80] A. Cellerino, D. Borghetti e F. Sartucci, «Sex differences in face gender recognition in humans,» *Brain research bulletin,* vol. 63, n. 6, pp. 443-449, 2004.

[81] C. Lewin e A. Herlitz, «Sex differences in face recognition--women's faces make the difference,» *Brain and Cognition,* vol. 50, n. 1, p. 121–128, 2002.

[82] S. J. McKelvie, «Sex differences in memory for faces,» *Journal of Psychology,* vol. 107, pp. 109-125, 1981.

[83] J. Connellan, S. Baron-Cohen, S. Wheelwright, A. Batki e J. Ahluwalia, «Sex differences in human neonatal social perception,» *Infant behavior and Development,* vol. 23, n. 1, pp. 113-118, 2000.

[84] J. F. Cross, J. Cross e J. Daly, «Sex, race, age, and beauty as factors in recognition of faces,» *Perception & Psychophysics,* vol. 10, n. 6, p. 393–396, 1971.

[85] J. Lovén, A. Herlitz e J. Rehnman, «Women's own-gender bias in face recognition memory: The role of attention at encoding,» *Experimental Psychology,* vol. 58, n. 4, p. 333–340, 2011.

[86] T. Ino, R. Nakai, T. Azuma, T. Kimura e H. Fukuyama, «Gender differences in brain activation during encoding and recognition of male and female faces,» *Brain Imaging and Behavior,* vol. 4, n. 1, p. 55–67, 2010.

[87] Y. Sun, X. Gao e S. Han, «Sex differences in face gender recognition: An event-related potential study,» *Brain Research,* vol. 1327, p. 69–76.

[88] T. W. Man e P. J. Hills, «Eye-tracking the own-gender bias in face recognition: Other-gender faces are viewed differently to own-gender faces,» *Visual Cognition,* vol. 24, n. 9-10, pp. 447-458, 2016.

[89] M. G. Rhodes e J. S. Anastasi, «The own-age bias in face recognition: A meta-analytic and theoretical review,» *Psychological Bulletin,* vol. 138, n. 1, p. 146–174, 2012.

[90] B. M. Craig e E. M. Thorne, «Social categorization and individuation in the own-age bias,» *British Journal of Psychology,* 2019.

[91] A. Tzavaras, H. Hecaen e H. Le Bras, «The problem of specificity of inadequate recognition of the human face after hemispheric unilateral lesions,» *Neuropsychologia,* vol. 8, n. 4, p. 403–416, 1970.

[92] I. Leo e F. Simion, «Newborns' Mooney-face perception,» *Infancy,* vol. 14, n. 6, p. 641–653, 2009.

[93] P. Wolff, «Observations on the early development of smiling,» in *Determinants of Infant Behaviour II*, London, Methuen, 1963.

[94] G. Carpenter, «Mother-Stranger Discrimination in the Early Weeks of Life,» ERIC, London, UK, 1973.

[95] I. W. R. Bushnell, «Mother's face recognition in newborn infants: learning and memory,» *Infant & Child Development,* vol. 10, n. 1/2, p. 67–74, 2001.

[96] L. Wan, K. Crookes, A. Dawel, M. Pidcock, A. Hall e E. McKone, «Face-blind for other-race faces: Individual differences in other-race recognition impairments,» *Journal of Experimental Psychology,* vol. 146, n. 1, p. 102–122, 2017.

[97] R. S. Malpass e J. Kravitz, «Recognition for faces of own and other race,» *Journal of Personality and Social Psychology,* vol. 13, n. 4, p. 330–334, 1969.

[98] T. F. Gross, «Own-Ethnicity Bias in the Recognition of Black, East Asian, Hispanic, and White Faces,» *Basic & Applied Social Psychology,* vol. 31, n. 2, p. 128–135, 2009.

[99] K. Lee, P. C. Quinn e O. Pascalis, «Face race processing and racial bias in early development: A perceptual-social linkage,» *Current Directions in Psychological Science,* vol. 26, n. 3, p. 256–262, 2017.

[100] R. E. Galper, «"Functional race membership" and recognition of faces,» *Perceptual and Motor Skills,* vol. 37, n. 2, p. 455–462, 1973.

[101] R. S. Malpass, «Racial bias in eyewitness identification,» *Personality and Social Psychology Bulletin,* vol. 1, n. 1, p. 42–44, 1974.

[102] H. Ellis, J. Shepherd e A. Bruce, «The effects of age and sex upon adolescents' recognition of faces,» *The Journal of Genetic Psychology: Research and Theory on Human Development,* vol. 123, n. 1, p. 173–174, 1973.

[103] P. J. Hills e S. F. Willis, «Children view own-age faces qualitatively differently to other-age faces,» *Journal of Cognitive Psychology,* vol. 28, n. 5, pp. 601-610, 2016.

[104] C. Hu, Q. Wang, G. Fu, P. C. Quinn e K. Lee, «Both children and adults scan faces of own and other races differently,» *Vision Research,* vol. 102, p. 1–10, 2014.

[105] T. B. Nguyen e K. Pezdek, «Memory for disguised same- and cross-race faces: The eyes have it,» *Visual Cognition,* vol. 25, n. 7/8, p. 762–769, 2017.

[106] K. Pezdek, M. O'Brien e C. Wasson, «Cross-race (but not same-race) face identification is impaired by presenting faces in a group rather than individually,» *Law and Human Behavior,* vol. 36, n. 6, p. 488–495, 2012.

[107] M. Stelter e J. Degner, «Investigating the other-race effect in working memory,» *British Journal of Psychology,* vol. 109, n. 4, p. 777–798, 2018.

[108] M. R. Ho e K. Pezdek, «Postencoding cognitive processes in the cross-race effect: Categorization and individuation during face recognition,» *Psychonomic Bulletin & Review,* vol. 23, n. 3, p. 771–780, 2016.

[109] K. Hugenberg, J. Miller e H. M. Claypool, «Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem,» *Journal of Experimental Social Psychology,* vol. 43, n. 2, p. 334–340, 2007.

[110] S. G. Young, M. J. Bernstein e K. Hugenberg, «When do own-group biases in face recognition occur? Encoding versus post-encoding,» *Social Cognition,* vol. 28, n. 2, p. 240–250, 2010.

[111] K. Crookes e G. Rhodes, «Poor recognition of other-race faces cannot always be explained by a lack of effort,» *Visual Cognition,* vol. 25, n. 4-6, p. 430–441, 2017.

[112] L. Wan, K. Crookes, A. Dawel, M. Pidcock, A. Hall e E. McKone, «Face-blind for other-race faces: Individual differences in other-race recognition impairments,» *Journal of Experimental Psychology: General,* vol. 146, n. 1, p. 102, 2017.

[113] E. Hehman, E. W. Mania e S. L. Gaertner, «Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect,» *Journal of Experimental Social Psychology,* vol. 46, n. 2, p. 445–448, 2010.

[114] S. Sangrigoli, C. Pallier, A.-M. Argenti, V. A. G. Ventureyra e S. de Schonen, «Reversibility of the Other-Race Effect in Face Recognition During Childhood,» *Psychological Science,* vol. 16, n. 6, p. 440–444, 2005.

[115] S. H.-L. Chien, J.-F. Wang e T.-R. Huang, «Developing the own-race advantage in 4-, 6-, and 9-month-old Taiwanese infants: A perceptual learning perspective,» *Frontiers in Psychology,* vol. 7, 2016.

[116] C. Liu, W. Chen e J. Ward, «Effects of exposure to facial expression variation in face learning and recognition,» *Psychological Research,* vol. 79, n. 6, p. 1042–1053, 2015.

[117] N. A. Sugden e A. R. Marquis, «Meta-analytic review of the development of face discrimination in infancy: Face race, face gender, infant age, and methodology moderate face discrimination,» *Psychological Bulletin,* vol. 143, n. 11, p. 1201–1244, 2017.

[118] A. Megias, I. Rzeszewska, L. Aguado e A. Catena, «Influence of cross-ethnic social experience on face recognition accuracy and the visual perceptual strategies involved,» *International Journal of Intercultural Relations,* vol. 65, pp. 42-50, 2018.

[119] E. G. Krumhuber e A. S. Manstead, «When memory is better for out-group faces: On negative emotions and gender roles,» *Journal of Nonverbal Behavior,* vol. 35, n. 1, pp. 51-61, 2011.

[120] V. Proietti, S. Laurence, C. M. Matthews, X. Zhou e C. J. Mondloch, «Attending to identity cues reduces the own-age but not the own-race recognition advantage,» *Vision Research,* vol. 157, p. 184–191, 2019.

[121] M. Wang, W. Deng, J. Hu, J. Peng, X. Tao e Y. Huang, «Racial faces in-the-wild: Reducing racial bias by deep unsupervised domain adaptation,» *arXiv:1812.00194,* December 2018.

[122] C. M. Bishop, Pattern Recognition and Machine Learning, New York: Springer Science+Business Media, LLC, 2006.

[123] W. S. McCulloch e W. Pitts, «A logical calculus of the ideas immanent in nervous activity,» *The bulletin of mathematical biophysics,* vol. 5, n. 4, pp. 115-133, 1943.

[124] F. Rosenblatt, Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms, Washington D.C.: Spartan Books, 1962.

[125] B. Widrow e M. E. Hoff, «Adaptive switching circuits,» Stanford Univ Ca Stanford Electronics Labs, Stanford, California, 1960.

[126] D. H. Hubel e T. N. Wiesel, «Receptive fields, binocular interaction and functional architecture in the cat's visual cortex,» *The Journal of physiology,* vol. 160, n. 1, pp. 106-154, 1962.

[127] K. Fukushima, «Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,» *Biological Cybernetics,* vol. 36, pp. 193-202, 1980.

[128] G. W. Lindsay, «Convolutional neural networks as a model of the visual system: past, present, and future,» *Journal of cognitive neuroscience,* pp. 1-15, 2020.

[129] F. Nonis, P. Barbiero, G. Cirrincione, E. Olivetti, F. Marcolin e E. Vezzetti, «Understanding Abstraction in Deep CNN: An Application on Facial Emotion Recognition,» in *Progresses in Artificial Intelligence and Neural Systems*, Singapore, Springer, 2020, pp. 281-290.

[130] Z. Fei, E. L. Yang, B. D. D. U., I. W. S., X. Li e H. Zhou, «Deep convolution network based emotion analysis towards mental health care,» *Neurocomputing,* vol. 388, pp. 212-227, 2020.

[131] Y. Guo, Y. Xia, J. Y. H. Wang e R. C. Chen, «Real-time facial affective computing on mobile devices,» *Sensors,* vol. 20, n. 3, p. 870, 2020.

[132] E. C. Olivetti, J. Ferretti, G. Cirrincione, F. Nonis, S. Tornincasa e F. Marcolin, «Deep CNN for 3D face recognition,» in *International Conference on Design, Simulation, Manufacturing: The Innovation Exchange*, Springer, 2019, pp. 665-674.

[133] X. Liu, Y. Xia, H. Yu, J. Dong, M. Jian e T. D. Pham, «Region Based Parallel Hierarchy Convolutional Neural Network for Automatic Facial Nerve Paralysis Evaluation,» *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 28, n. 10, pp. 2325-2332, 2020.

[134] S. Li e W. Deng, « Deep facial expression recognition: A survey,» *arXiv:1804.08348,* 2018.

[135] M. M. Kasar, D. Bhattacharyya e T. H. Kim, «Face recognition using neural network: a review,» *International Journal of Security and Its Applications,* vol. 10, n. 3, pp. 81-100, 2016.

[136] I. Adjabi, A. Ouahabi, A. Benzaoui e A. Taleb-Ahmed, «Past, present, and future of face recognition: A review,» *Electronics,* vol. 9, n. 8, p. 1188, 2020.

[137] P. Payal e M. M. Goyani, «A comprehensive study on face recognition: methods and challenges,» *The Imaging Science Journal,* vol. 68, n. 2, pp. 114-127, 2020.

[138] Y. Kortli, M. Jridi, A. Al Falou e M. Atri, «Face recognition systems: A Survey,» *Sensors,,* vol. 20, n. 2, p. 342, 2020.

[139] D. Canedo e A. J. Neves, «Facial expression recognition using computer vision: A systematic review,» *Applied Sciences,* vol. 9, n. 21, p. 4678, 2019.

[140] F. Nonis, N. Dagnes, F. Marcolin e E. Vezzetti, «3D Approaches and Challenges in Facial Expression Recognition Algorithms—A Literature Review,» *Applied Sciences,* vol. 9, n. 3904, pp. 1-33, 2019.

[141] M. R. Reshma e B. Kannan, «Approaches on partial face recognition: a literature review,» *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), IEEE,* pp. 538-544, April 2019.

[142] D. S. Trigueros, L. Meng e M. Hartnett, «Face recognition: From traditional to deep learning methods,» *arXiv preprint arXiv:1811.00116,* 2018.

[143] M. Wang e W. Deng, «Deep face recognition: A survey,» *arXiv:1804.06655,* 2018.

[144] S. Y. Fu, G. S. Yang e X. K. Kuai, «A spiking neural network based cortex-like mechanism and application to facial expression recognition,» *Computational intelligence and neuroscience,* vol. 19, 2012.

[145] G. W. Cottrell e J. Metcalfe, «EMPATH: Face, emotion, and gender recognition using holons,» *Advances in neural information processing systems,* pp. 564-571, 1991.

[146] M. N. Dailey, G. W. Cottrell, C. Padgett e R. Adolphs, «EMPATH: A neural network that categorizes facial expressions,» *Journal of cognitive neuroscience,* vol. 14, n. 8, pp. 1158-1173, 2002.

[147] G. W. Cottrell e J. H. Hsiao, «Neurocomputational Models of Face Processing,» in *The Oxford Handbook of Face Perception*, New York, The Oxford Handbook of Face Perception, 2011, pp. 401-423.

[148] S. Marĉelja, «Mathematical description of the responses of simple cortical cells,» *JOSA,* vol. 70, n. 11, pp. 1297-1300, 1980.

[149] N. Petkov, P. Kruizinga e T. Lourens, «Biologically motivated approach to face recognition,» *International Workshop on Artificial Neural Networks,* pp. 68-77, June 1993.

[150] Y. D. Sato, T. Nagatomi, K. Horio e H. Miyamoto, «Cognitive Computation,» *The cognitive mechanisms of multi-scale perception for the recognition of extremely similar faces,* vol. 7, n. 5, pp. 501-508, 2015.

[151] L. Shen e L. Bai, «A review on Gabor wavelets for face recognition,» *Pattern analysis and applications,* vol. 9, n. 2-3, pp. 273-292, 2006.

[152] V. J. Mistry e M. M. Goyani, «A literature survey on facial expression recognition using global features,» *Int. J. Eng. Adv. Technol,* vol. 2, n. 4, pp. 653-657, 2013.

[153] Á. Serrano, I. M. de Diego, C. Conde e E. Cabello, «Recent advances in face biometrics with Gabor wavelets: A review,» *Pattern Recognition Letters,* vol. 31, n. 5, pp. 372-381, 2010.

[154] W. Deng, J. Hu, J. Guo, W. Cai e D. Feng, «Emulating biological strategies for uncontrolled face recognition,» *Pattern Recognition,* vol. 43, n. 6, pp. 2210-2223, 2010.

[155] W. Sun, H. Zhao e Z. Jin, «A visual attention based ROI detection method for facial expression recognition,» *Neurocomputing,* vol. 296, n. 12-22, 2018.

[156] Y. Li, J. Zeng, S. Shan e X. Chen, «Occlusion aware facial expression recognition using CNN with attention mechanism,» *IEEE Transactions on Image Processing,* vol. 28, n. 5, pp. 2439-2450, 2018.

[157] J. Li, K. Jin, D. Zhou, N. Kubota e Z. Ju, «Attention mechanism-based CNN for facial expression recognition,» *Neurocomputing,* vol. 411, pp. 340-350, 2020.

[158] A. V. Puri e B. Lall, «Exploiting perception for face analysis: image abstraction for head pose estimation,» *European Conference on Computer Vision,* pp. 319-329, October 2012.

[159] H. Wu, W. C. Chen e N. M. Mayer, «A face recognition system that simulates perception impairments of autistic children,» *Neurocomputing,* vol. 168, pp. 770-776, 2015.

[160] E. Vezzetti, F. Marcolin e G. Fracastoro, «3D face recognition: An automatic strategy based on geometrical descriptors and landmarks,» *Robotics and Autonomous Systems,* vol. 62, n. 12, pp. 1768-1776, 2014.

[161] E. Vezzetti, F. Marcolin, S. Tornincasa, L. Ulrich and N. Dagnes, "3D geometry-based automatic landmark localization in presence of facial occlusions.," *Multimedia Tools and Applications,* 2017.

[162] N. Dagnes, F. Marcolin, F. Nonis, S. Tornincasa e E. Vezzetti, «3D geometry-based face recognition in presence of eye and mouth occlusions,» *International Journal on Interactive Design and Manufacturing (IJIDeM),* pp. 1-19, 2019.

[163] G. Ciravegna, G. Cirrincione, F. Marcolin, P. Barbiero, N. Dagnes e E. Piccolo, «Assessing discriminating capability of geometrical descriptors for 3D face recognition by using the GH-EXIN neural network,» in *WIRN2018*, Vietri sul mare (Salerno, Italy), 2018.

[164] P. Dunker e M. Keller, «Illumination Normalization for Face Recognition-A Comparative Study of Conventional vs. Perception-inspired Algorithms,» *BIOSIGNALS,* pp. 237-243, 2008.

[165] Z. Xie, G. Liu e Z. Fang, «Face recognition based on combination of human perception and local binary pattern,» *International Conference on Intelligent Science and Intelligent Data Engineering,* pp. 365-373, October 2011.

[166] H. Yu, O. G. Garrod e P. G. Schyns, «Perception-driven facial expression synthesis,» *Computers & Graphics,* vol. 36, n. 3, pp. 152-162, 2012.

[167] J. Zhan, O. G. Garrod, N. van Rijsbergen e P. G. Schyns, «Modelling face memory reveals task-generalizable representations,» *Nature human behaviour,* vol. 1, 2019.

[168] N. Kanwisher e J. J. S. Barton, «The Functional Architecture of the Face System: Integrating Evidence from fMRI and Patient Studies,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 111-129.

[169] E. Fox e K. Zougkou, «Influence of Personality Traits on Processing of Facial Expressions,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 515-534.

[170] N. Ambady e M. Weisbuch, «On Perceiving Facial Expressions: the Role of Culture and Context,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 479-488.

[171] A. J. O'Toole, X. An, J. Dunlop, V. Natu e P. J. Phillips, «Comparing face recognition algorithms to humans on challenging tasks,» *ACM Transactions on Applied Perception (TAP),* vol. 9, n. 4, p. 16, 2012.

[172] Y. Zhong e W. Deng, «Deep Difference Analysis in Similar-looking Face recognition,» in *24th International Conference on Pattern Recognition (ICPR)*, Beijing, 2018.

[173] W. Deng, J. Hu, N. Zhang, B. Chen e J. Guo, «Fine-grained face verification: Fglfw database, baselines, and human-dcmn partnership,» *Pattern Recognition,* vol. 66, pp. 63-73, 2017.

[174] A. M. Burtons e R. Jenkins, «Unfamiliar Face Perception,» in *The Oxford Handbook of Face Perception*, New York, Oxford University Press, 2011, pp. 287-306.

[175] H. Markram, «The blue brain project,» *Nature Reviews Neuroscience,* vol. 7, n. 2, pp. 153-160, 2006.

[176] F. Crick, «The recent excitement about neural networks,» *Nature,* vol. 337, n. 6203, pp. 129-132, 1989.

[177] N. Kriegeskorte, «Deep neural networks: a new framework for modeling biological vision and brain information processing,» *Annual review of vision science,* vol. 1, pp. 417-446, 2015.

[178] R. M. Cichy, A. Khosla, D. Pantazis e A. Oliva, «Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks,» *NeuroImage,* vol. 153, pp. 346-358, 2017.

[179] R. E. Jack e P. G. Schyns, «The human face as a dynamic tool for social communication,» *Current Biology,* vol. 25, n. 14, pp. R621-R634, 2015.

[180] T. F. Gross, «Face recognition and own-ethnicity bias in Black, East/Southeast Asian, Hispanic, and White children,» *Asian American Journal of Psychology,* vol. 5, n. 3, p. 181–189, 2014.